Sports-études année 1

COURS DE MATHÉMATIQUES

Deuxième partie : algèbre linéaire

Aimé LACHAL

Centre de Mathématiques

Institut National des Sciences Appliquées de Lyon

Avertissement

Ceci est la deuxième partie du cours de mathématiques s'adressant aux élèves de l'année 1 de la section sportive de haut niveau. Elle concerne la théorie des espaces vectoriels et se déroule au second semestre. On donne les bases de la structure d'espace vectoriel (prolongeant ou complétant les structures algébriques abordées en année 0 : groupes, anneaux, corps), puis on introduit les applications linéaires entre espaces vectoriels qui transportent cette structure et qui permettent de définir un calcul matriciel. L'objectif final est la diagonalisation des matrices et la réduction des endomorphismes conduisant à des applications importantes :

— résolution de systèmes différentiels linéaires ;
— étude de suites définies par des relations de récurrence linéaires ;
— réduction des coniques.

Elle sera complétée en année 2 par un chapitre sur les espaces vectoriels normés. Le calcul matriciel sera utilisé en calcul différentiel à plusieurs variables ainsi que dans l'étude des quadriques de l'espace.

Chaque chapitre est suivi d'exercices traités en séances de travaux dirigés. Ce document écrit est plus complet que l'exposé du cours oral. En particulier, les parties dactylographiées en petits caractères sont soit des parties de démonstration délicates ou hors-programme, qui sont insérées dans ce cours par souci de complétude, soit des remarques ou compléments anecdotiques, pour satisfaire les esprits curieux. Elles pourront être omises en première lecture.

Table des matières

Chapitre 1

Espaces vectoriels

Les structures algébriques concernent des ensembles pour lesquels on dispose de certaines opérations. Elles apparaissent naturellement et progressivement dans la construction des nombres usuels : entiers, rationnels, réels, complexes... avec les opérations d'addition et de multiplication. Ces structures ont pour vocation d'unifier des propriétés typiques (telles que la commutativité, l'associativité, la distributivité...) et d'ériger des règles de calcul qui s'avèrent communes à de nombreux exemples de natures *a priori* différentes. La structure d'espace vectoriel est née avec l'étude géométrique des vecteurs du plan ou de l'espace dans lequel on vit. Malgré l'aspect géométrique de cette origine, les propriétés des espaces vectoriels sont de nature algébrique et s'appliquent dans divers contextes éloignés, en apparence, de la géométrie (espaces de fonctions, de polynômes, de séries de Fourier...)

Par exemple, pour les trois ensembles
— $\overrightarrow{\mathcal{P}}$ des vecteurs du plan,
— $\mathbb{R}_1[X]$ des binômes de degré au plus 1, $\alpha X + \beta$ avec $\alpha, \beta \in \mathbb{R}$,
— $\mathcal{F}_{\text{trigo}}$ des fonctions trigonométriques de la forme $x \longmapsto \alpha \cos x + \beta \sin x$ avec $\alpha, \beta \in \mathbb{R}$,
on peut définir des opérations tout à fait similaires pour lesquels des règles de calcul sont rigoureusement les mêmes. Ces opérations sont les suivantes :

$$\begin{array}{ccc} \overrightarrow{\mathcal{P}} \times \overrightarrow{\mathcal{P}} & \longrightarrow & \overrightarrow{\mathcal{P}} \\ (\vec{u}, \vec{v}) & \longmapsto & \vec{u} + \vec{v} \\ (\alpha, \vec{u}) & \longmapsto & \alpha \,.\, \vec{u} \end{array} \qquad \begin{array}{ccc} \mathbb{R}_1[X] \times \mathbb{R}_1[X] & \longrightarrow & \mathbb{R}_1[X] \\ (P, Q) & \longmapsto & P + Q \\ (\alpha, P) & \longmapsto & \alpha \,.\, P \end{array} \qquad \begin{array}{ccc} \mathcal{F}_{\text{trigo}} \times \mathcal{F}_{\text{trigo}} & \longrightarrow & \mathcal{F}_{\text{trigo}} \\ (f, g) & \longmapsto & f + g \\ (\alpha, f) & \longmapsto & \alpha \,.\, f \end{array}$$

— Pour le plan $\overrightarrow{\mathcal{P}}$, si l'on dispose d'une base $\{\vec{i}, \vec{j}\}$, on peut écrire n'importe quel vecteur $\vec{u}$ de $\overrightarrow{\mathcal{P}}$ sous la forme $\vec{u} = \alpha\vec{i} + \beta\vec{j}$ avec $\alpha, \beta \in \mathbb{R}$;
— l'ensemble $\mathbb{R}_1[X]$ est constitué de binômes de la forme $\alpha X + \beta 1$ avec $\alpha, \beta \in \mathbb{R}$, 1 étant le polynôme constant égal à 1 ;
— l'ensemble $\mathcal{F}_{\text{trigo}}$ est constitué de fonctions de la forme $\alpha \cos + \beta \sin$ avec $\alpha, \beta \in \mathbb{R}$.

Malgré la différence de nature de tous ces objets, dans les trois cas, les paires $\{\vec{i}, \vec{j}\}$, $\{X, 1\}$ et $\{\cos, \sin\}$ jouent le même rôle vis-à-vis des opérations + et ., à savoir que l'on peut écrire dans chaque situation un objet de l'ensemble uniquement à l'aide de la paire correspondante et de facteurs d'échelle α, β ; on parle de *combinaisons linéaires* (*e.g.* $\alpha\vec{i} + \beta\vec{j}$) et de *bases* (*e.g.* $\{\vec{i}, \vec{j}\}$). Les ensembles $\overrightarrow{\mathcal{P}}$, $\mathbb{R}_1[X]$, $\mathcal{F}_{\text{trigo}}$ sont ainsi rapprochés les uns les autres selon un même jeu de calcul « vectoriel » et seront qualifiés d'*espaces vectoriels isomorphes*.

1 Rappel : structures de groupe, anneau et corps

Définition 1.1 *Soit E un ensemble. Une loi interne sur E notée $*$ est une application de $E \times E$ à valeurs dans E :*

$$\begin{array}{rcl} * : E \times E & \longrightarrow & E \\ (x, y) & \longmapsto & x * y \end{array}$$

1.1 Groupes

Définition 1.2 *Soit E un ensemble muni d'une loi interne $*$. Le couple $(E, *)$ est un groupe lorsque*

1. *la loi $*$ est* associative *:*

$$\forall x, y, z \in E,\ (x * y) * z = x * (y * z)\,;$$

2. *la loi $*$ admet un élément neutre noté e :*

$$\exists e \in E, \forall x \in E,\ x * e = e * x = x\,;$$

3. *tout élément x de E admet un* symétrique *noté x' pour la loi $*$:*

$$\forall x \in E, \exists x' \in E,\ x * x' = x' * x = e.$$

Si de plus la loi $*$ est commutative *:*

$$\forall x, y \in E,\ x * y = y * x,$$

on dit que le groupe $(E, *)$ est commutatif *ou* abélien.

Remarques.

1. Dans un groupe, on peut « simplifier » des égalités :

$$x * y = x * z \Longrightarrow y = z \quad \text{et} \quad y * x = z * x \Longrightarrow y = z.$$

En effet,

$$x * y = x * z \Longrightarrow x' * (x * y) = x' * (x * z) \Longrightarrow (x' * x) * y = (x' * x) * z \Longrightarrow e * y = e * z \Longrightarrow y = z.$$

2. Un groupe admet un unique élément neutre et tout élément admet un seul symétrique. En effet,
   — si e et e' sont deux éléments neutres, on a $e * e' = e = e'$;
   — si x' et x'' sont deux symétriques de x, on a $x * x' = x * x'' = e$ et donc en simplifiant par x on trouve $x' = x''$.

Lorsque la loi $*$ est commutative, l'usage veut qu'on la note additivement : $+$ (*addition*), et dans ce cas l'élément neutre est noté 0 (*zéro*), le symétrique de x est noté $-x$ (*opposé de* x). Dans d'autres cas, il arrive qu'on note la loi $*$ multiplicativement : $\times$ (*multiplication*), l'élément neutre est alors noté 1 (*un*), le symétrique de x est noté x^{-1} (*inverse* de x). On note dans le même esprit $x + (-y) = x - y$ et $x \times y^{-1} = x/y$. La notation $\dfrac{x}{y}$ n'a de sens que si la loi $\times$ est commutative.

EXEMPLES.

1. Les couples $(\mathbb{Z},+)$, $(\mathbb{Q},+)$, $(\mathbb{Q}^*,x)$ sont des groupes commutatifs.
2. Soit $\mathcal{B}(E)$ l'ensemble des bijections d'un ensemble E dans lui-même (ce sont les *permutations* de E). Le couple $(E,\circ)$ est un groupe en général non commutatif. En effet,
   — la loi $\circ$ est associative : soient $f,g,h\in\mathcal{B}(E)$. On a
   $$\forall x\in E,\ ((f\circ g)\circ h)(x)=(f\circ g)(h(x))=f(g(h(x)))=f((g\circ h)(x))=(f\circ(g\circ h))(x)$$
   donc $(f\circ g)\circ h=f\circ(g\circ h)$;
   — la loi $\circ$ admet pour élément neutre l'application identité id_E : soit $f\in\mathcal{B}(E)$. On a
   $$\forall x\in E,\ (f\circ \mathrm{id}_E)(x)=f(\mathrm{id}_E(x))=f(x)=\mathrm{id}_E(f(x))=(\mathrm{id}_E\circ f)(x)$$
   donc $f\circ\mathrm{id}_E=\mathrm{id}_E\circ f$;
   — l'élément f de $\mathcal{B}(E)$ admet pour symétrique sa bijection réciproque f^{-1} pour la loi $\circ$:
   $$f\circ f^{-1}=f^{-1}\circ f=\mathrm{id}_E.$$

1.2 Anneaux

Définition 1.3 *Soit E un ensemble muni de deux lois internes notées $+$ et $\times$. Le triplet $(E,+,\times)$ est un anneau lorsque*

1. *$(E,+)$ est un groupe commutatif ;*
2. *la loi $\times$ est associative ;*
3. *la loi $\times$ est* distributive *sur la loi $+$:*
   $$\forall x,y,z\in E,\ (x+y)\times z=(x\times z)+(y\times z)\quad et\quad z\times(x+y)=(z\times x)+(z\times y).$$

— *Si de plus la loi $\times$ est commutative (resp. admet un élément neutre, appelé* élément unité*), on dit que l'anneau $(E,+,\times)$ est* commutatif *(resp.* unitaire*).*
— *Si l'on a*
$$\forall x,y\in E,\ x\times y=0\Longrightarrow x=0\ ou\ y=0,$$
c'est-à-dire qu'il n'y a pas de diviseurs de zéro*, on dit que l'anneau est $(E,+,\times)$* intègre.

REMARQUES.

1. Dans un anneau, l'élément neutre 0 pour la première loi $+$ est *absorbant* pour la deuxième $\times$:
   $$\forall x\in E,\ x\times 0=0\times x=0.$$
   En effet, par distributivité on a par exemple $x\times 0=x\times(0+0)=(x\times 0)+(x\times 0)$, et, en simplifiant par $x\times 0$, on trouve $x\times 0=0$.
2. Dans un anneau, on adopte la règle de priorité suivante : la multiplication est prioritaire sur l'addition, c'est-à-dire que l'on note l'expression $(a\times b)+(c\times d)$ sans parenthèses : $a\times b+c\times d$.

EXEMPLE.

Les triplets $(\mathbb{Z},+,\times)$ et $(\mathbb{R}[X],+,\times)$ sont des anneaux commutatifs.

1.3 Corps

Définition 1.4 *Soit E un ensemble muni de deux lois internes notées $+$ et $\times$. Le triplet $(E, +, \times)$ est un corps lorsque*

1. *$(E, +)$ est un groupe commutatif ;*
2. *$(E \setminus \{0\}, \times)$ est un groupe ;*
3. *la loi $\times$ est distributive sur la loi $+$.*

Si de plus la loi $\times$ est commutative, on dit que le corps $(E, +, \times)$ est commutatif.

Exemple.

Les triplets $(\mathbb{Q}, +, \times)$, $(\mathbb{R}, +, \times)$, $(\mathbb{C}, +, \times)$ et $(\mathbb{R}(X), +, \times)$ sont des corps commutatifs.

2 Espaces vectoriels

Définition 2.1 *Soit E un ensemble et $\mathbb{K}$ un corps commutatif. Une loi externe sur E relativement à $\mathbb{K}$ notée $\odot$ est une application de $\mathbb{K} \times E$ à valeurs dans E :*

$$\begin{array}{rcl} \odot : \mathbb{K} \times E & \longrightarrow & E \\ (\alpha, x) & \longmapsto & \alpha \odot y \end{array}$$

Dans tout ce cours, $(\mathbb{K}, +, \times)$ désignera un corps commutatif. Son élément neutre et son élément unité seront respectivement notés 0 et 1. L'opposé d'un élément $\alpha \in \mathbb{K}$ sera noté $-\alpha$ et l'inverse d'un élément $\alpha \in \mathbb{K} \setminus \{0\}$ sera noté $\dfrac{1}{\alpha}$ (ou $1/\alpha$ ou α^{-1}). L'entier n sera assimilé à l'élément $\underbrace{1 + \cdots + 1}_{n \text{ fois}}$ de $\mathbb{K}$.

2.1 Axiomes

Définition 2.2 *Soit E un ensemble muni d'une loi interne $\oplus$* (addition vectorielle) *et d'une loi externe $\odot$* (multiplication scalaire) *relativement au corps $\mathbb{K}$. Le triplet $(E, \oplus, \odot)$ est un* espace vectoriel sur $\mathbb{K}$, *ou encore un* $\mathbb{K}$-espace vectoriel, *lorsque*

1. *$(E, \oplus)$ est un groupe commutatif ;*
2. *$\forall \alpha, \beta \in \mathbb{K}$, $\forall \vec{u}, \vec{v} \in E$,*

$$\begin{array}{rcl} \alpha \odot (\vec{u} \oplus \vec{v}) & = & (\alpha \odot \vec{u}) \oplus (\alpha \odot \vec{v}), \\ (\alpha + \beta) \odot \vec{u} & = & (\alpha \odot \vec{u}) \oplus (\beta \odot \vec{u}), \\ (\alpha \times \beta) \odot \vec{u} & = & \alpha \odot (\beta \odot \vec{u}), \\ 1 \odot \vec{u} & = & \vec{u}. \end{array}$$

Les éléments de E sont appelés vecteurs *et ceux de $\mathbb{K}$* scalaires. *L'élément neutre de l'addition vectorielle est noté $\vec{0}$ et le symétrique du vecteur $\vec{u}$ pour cette addition est noté $\ominus\vec{u}$.*

Remarques.

1. On dénote dans ce paragraphe les deux lois d'espace vectoriel provisoirement par les symboles $\oplus$ et $\odot$ de façon à bien les différencier des deux lois de corps, mais on adoptera dès le paragraphe suivant les notations classiques $+$ et . au lieu de $\oplus$ et $\odot$, le symbole $\times$ sera omis et même le symbole . sera souvent omis. On notera également $\vec{u} \oplus (\ominus \vec{v}) = \vec{u} \ominus \vec{v}$ ou encore avec les futures notations $\vec{u} + (-\vec{v}) = \vec{u} - \vec{v}$.

2. Dans un espace vectoriel on adopte la règle de priorité suivante : la multiplication scalaire est prioritaire sur l'addition vectorielle. Par exemple, avec les futures notations, on notera l'expression $(\alpha\vec{u}) + (\beta\vec{v})$ sans parenthèses : $\alpha\vec{u} + \beta\vec{v}$.

3. Dans la suite, on ne considérera que des espaces vectoriels sur $\mathbb{R}$ ou $\mathbb{C}$.

Les dessins ci-dessous illustrent les axiomes de la définition 2.2 dans le cadre des vecteurs géométriques du plan.

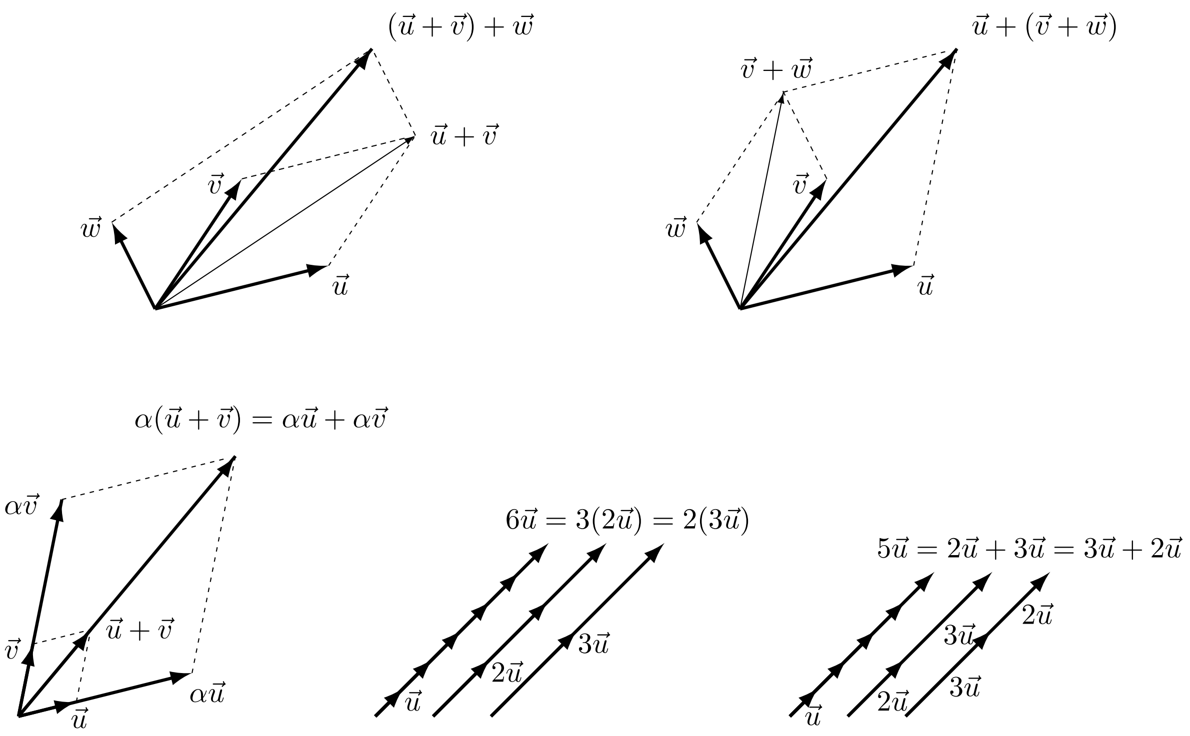

Figure 1.1 – Illustration des axiomes d'un espace vectoriel

Définition 2.3 *On peut construire à l'aide de n vecteurs $\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_n$ et n scalaires $\alpha_1, \alpha_2, \ldots, \alpha_n$ le vecteur $(\alpha_1 \odot \vec{u}_1) \oplus (\alpha_2 \odot \vec{u}_2) \oplus \cdots \oplus (\alpha_n \odot \vec{u}_n)$. C'est la* combinaison linéaire *des vecteurs $\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_n$ à coefficients $\alpha_1, \alpha_2, \ldots, \alpha_n$*

2.2 Propriétés immédiates

Proposition 2.4 *Soit E un $\mathbb{K}$-espace vectoriel. On a $\forall n \in \mathbb{N}^*$, $\forall \alpha \in \mathbb{K}$, $\forall \vec{u} \in E$,*

$$0 \odot \vec{u} = \vec{0}, \quad \alpha \odot \vec{0} = \vec{0}, \quad \alpha \odot \vec{u} = \vec{0} \Longrightarrow \alpha = 0 \ \text{ou}\ \vec{u} = \vec{0},$$

$$(-\alpha) \odot \vec{u} = \ominus(\alpha \odot \vec{u}), \quad n \odot \vec{u} = \underbrace{\vec{u} \oplus \cdots \oplus \vec{u}}_{n \text{ fois}}.$$

Démonstration.

En utilisant les axiomes de la structure d'espace vectoriel on obtient

— $0 \odot \vec{u} = (0+0) \odot \vec{u} = 0 \odot \vec{u} \oplus 0 \odot \vec{u}$ et donc en simplifiant : $0 \odot \vec{u} = \vec{0}$;

— $\alpha \odot \vec{0} = \alpha \odot (\vec{0} \oplus \vec{0}) = \alpha \odot \vec{0} \oplus \alpha \odot \vec{0}$ et donc en simplifiant : $\alpha \odot \vec{0} = \vec{0}$;

— en supposant que $\alpha \odot \vec{u} = \vec{0}$ et $\alpha \neq 0$ et en multipliant par α^{-1}, on trouve d'une part que $\alpha^{-1} \odot (\alpha \odot \vec{u}) = \alpha^{-1} \odot \vec{0} = \vec{0}$ et d'autre part $\alpha^{-1} \odot (\alpha \odot \vec{u}) = (\alpha^{-1} \times \alpha) \odot \vec{u} = 1 \odot \vec{u} = \vec{u}$. D'où $\vec{u} = \vec{0}$;

— $((-\alpha) \odot \vec{u}) \oplus (\alpha \odot \vec{u}) = (-\alpha + \alpha) \odot \vec{u} = 0 \odot \vec{u} = \vec{0}$, ce qui montre que $(-\alpha) \odot \vec{u}$ est le symétrique de $\alpha \odot \vec{u}$ pour $\oplus$;

— $n \odot \vec{u} = \underbrace{(1 + \cdots + 1)}_{n \text{ fois}} \odot \vec{u} = \underbrace{(1 \odot \vec{u}) \oplus \cdots \oplus (1 \odot \vec{u})}_{n \text{ fois}} = \underbrace{\vec{u} \oplus \cdots \oplus \vec{u}}_{n \text{ fois}}$. □

2.3 Exemple fondamental

Soit $\mathbb{K}^n$ le produit cartésien de n copies de $\mathbb{K}$: $\mathbb{K}^n = \underbrace{\mathbb{K} \times \cdots \times \mathbb{K}}_{n \text{ fois}}$. C'est l'ensemble des n-uplets constitués d'éléments de $\mathbb{K}$: $\mathbb{K}^n = \{(x_1, \ldots, x_n),\ x_1, \ldots, x_n \in \mathbb{K}\}$. On définit sur $\mathbb{K}^n$

— l'*égalité* : $(x_1, \ldots, x_n) = (y_1, \ldots, y_n) \Longleftrightarrow x_1 = y_1, \ldots, x_n = y_n$;

— l'*addition* : $(x_1, \ldots, x_n) \oplus (y_1, \ldots, y_n) = (x_1 + y_1, \ldots, x_1 + y_n)$;

— la *multiplication scalaire* : $\alpha \odot (x_1, \ldots, x_n) = (\alpha \times x_1, \ldots, \alpha \times x_n)$.

On peut montrer (la preuve est très simple mais longue) que le triplet $(\mathbb{K}^n, \oplus, \odot)$ est un $\mathbb{K}$-espace vectoriel. Le vecteur nul est ici $(0, \ldots, 0)$ et l'opposé du vecteur $(x_1, \ldots, x_n)$ est $\ominus(x_1, \ldots, x_n) = (-x_1, \ldots, -x_n)$.

Autres exemples.

1. L'ensemble des polynômes $\mathbb{K}[X]$ muni des lois

$$\oplus : \begin{array}{ccc} \mathbb{K}[X] \times \mathbb{K}[X] & \longrightarrow & \mathbb{K}[X] \\ (P, Q) & \longmapsto & P + Q \end{array} \quad \text{et} \quad \odot : \begin{array}{ccc} \mathbb{K} \times \mathbb{K}[X] & \longrightarrow & \mathbb{K}[X] \\ (\alpha, Q) & \longmapsto & \alpha Q \end{array}$$

   est un $\mathbb{K}$-espace vectoriel. L'ensemble $\mathbb{K}_n[X]$ des polynômes de degré au plus n muni des mêmes lois (en remplaçant $\mathbb{K}[X]$ par $\mathbb{K}_n[X]$) est aussi un $\mathbb{K}$-espace vectoriel.

2. L'ensemble $E = \mathcal{A}(F, \mathbb{K})$ des applications d'un ensemble F à valeurs dans $\mathbb{K}$ muni des lois

$$\oplus : \begin{array}{ccc} E \times E & \longrightarrow & E \\ (f, g) & \longmapsto & \left[f \oplus g : \begin{array}{ccc} F & \longrightarrow & \mathbb{K} \\ x & \longmapsto & f(x) + g(x) \end{array} \right] \end{array}$$

et

$$\begin{array}{rccl} \odot : & \mathbb{K} \times E & \longrightarrow & E \\ & (\alpha, f) & \longmapsto & \left[\alpha \odot f : \begin{array}{ccc} F & \longrightarrow & \mathbb{K} \\ x & \longmapsto & \alpha \times f(x) \end{array} \right] \end{array}$$

est un $\mathbb{K}$-espace vectoriel.

3 Sous-espaces vectoriels

Dans toute la suite de ce chapitre, E désigne un $\mathbb{K}$-espace vectoriel. La structure d'espace vectoriel est une structure riche puisqu'elle possède de nombreuses propriétés. En revanche, il est parfois lourd de vérifier tous les axiomes de cette structure et il existe des critères permettant de décider si un ensemble est un espace vectoriel ou non à l'aide de peu de vérification. Ces critères reposent sur la notion de sous-espace vectoriel et s'appliquent notamment à des ensembles qui sont des sous-ensembles d'un espace vectoriel ou construits à partir de plusieurs espaces vectoriels (intersection, somme). En pratique, disposant de quelques espaces vectoriels « classiques », on aura très souvent affaire à des sous-ensembles de ceux-ci.

3.1 Sous-espaces vectoriels

Si F est une partie non vide de E, on peut considérer les restrictions des lois de E à F : $+|_{F\times F}$ et $.|_{\mathbb{K}\times F}$. Elles sont définies par

$$\begin{array}{rccl} +|_{F\times F} : & F \times F & \longrightarrow & E \\ & (\vec{u}, \vec{v}) & \longmapsto & \vec{u} + \vec{v} \end{array} \quad \text{et} \quad \begin{array}{rccl} .|_{\mathbb{K}\times F} : & \mathbb{K} \times F & \longrightarrow & E \\ & (\alpha, \vec{u}) & \longmapsto & \alpha \, . \, \vec{u} \end{array}$$

Pour que F muni de ces deux lois soit un espace vectoriel, il est nécessaire que leur ensemble d'arrivée soit contenu dans F de façon à en faire des lois respectivement interne et externe.

Définition 3.1 *Une partie non vide F de E est un sous-espace vectoriel de E lorsque $(F, +|_{F\times F}, .|_{\mathbb{K}\times F})$ est un espace vectoriel.*

Le résultat ci-dessous permet de voir très simplement que F est un sous-espace vectoriel de E sans vérifier tous les axiomes de la définition d'un espace vectoriel.

Théorème 3.2 *Soit F une partie non vide de E. Alors F est un sous-espace vectoriel de E si et seulement si*

$$\forall \alpha \in \mathbb{K}, \forall \vec{u}, \vec{v} \in F,\ \vec{u}+\vec{v} \in F \quad \text{et} \quad \alpha \,.\, \vec{u} \in F$$

ou encore

$$\forall \alpha, \beta \in \mathbb{K}, \forall \vec{u}, \vec{v} \in F,\ \alpha \,.\, \vec{u} + \beta \,.\, \vec{v} \in F.$$

On dit que F est stable par combinaison linéaire.

Démonstration.

La preuve est immédiate. Les assertions du théorème garantissent que les opérations restreintes $+|_{F\times F}$ et $.|_{F\times F}$ sont des lois respectivement interne et externe. De plus, en faisant $\alpha = 0$ et $\alpha = -1$ dans l'assertion $\forall \alpha \in \mathbb{K}, \forall \vec{u} \in F,\ \alpha \,.\, \vec{u} \in F$, on voit que $\vec{0} \in F$ et $-\vec{u} \in F$ prouvant ainsi que l'élément neutre de $+$ et le symétrique d'un vecteur de F sont dans F. Les autres propriétés relatives aux lois $+$ et . (commutativité, associativité, etc.) se transmettent naturellement aux lois $+|_{F\times F}$ et $.|_{F\times F}$. □

Exemples.

1. Considérons dans le $\mathbb{K}$-espace vectoriel $E = \mathbb{K}^n$ les sous-ensembles F et G définis par $F = \{(\lambda a_1, \ldots, \lambda a_n),\ \lambda \in \mathbb{K}\}$ et $G = \{(x_1, \ldots, x_n) \in E : a_1x_1 + \cdots + a_nx_n = 0\}$ où les scalaires $a_1, \ldots, a_n$ sont fixés.
   (a). Soit $\alpha, \beta \in \mathbb{K}$ et $\vec{u}, \vec{v} \in F$. Les vecteurs $\vec{u}$ et $\vec{v}$ sont de la forme $\vec{u} = (\lambda a_1, \ldots, \lambda a_n)$ et $\vec{v} = (\mu a_1, \ldots, \mu a_n)$ avec $\lambda, \mu \in \mathbb{K}$. On a alors
   $$\alpha\vec{u} + \beta\vec{v} = \alpha(\lambda a_1, \ldots, \lambda a_n) + \beta(\mu a_1, \ldots, \mu a_n) = ((\alpha\lambda + \beta\mu)a_1, \ldots, (\alpha\lambda + \beta\mu)a_n) \in F$$
   prouvant ainsi que F est un sous-espace vectoriel de E.
   (b). Soit $\alpha, \beta \in \mathbb{K}$ et $(x_1, \ldots, x_n), (y_1, \ldots, y_n) \in G$. On a les relations $a_1x_1 + \cdots + a_nx_n = 0$ et $a_1y_1 + \cdots + a_ny_n = 0$. La combinaison linéaire $(z_1, \ldots, z_n) = \alpha(x_1, \ldots, x_n) + \beta(y_1, \ldots, y_n) = (\alpha x_1 + \beta y_1, \ldots, \alpha x_n + \beta y_n)$ vérifie
   $$a_1z_1 + \cdots + a_nz_n = \alpha(a_1x_1 + \cdots + a_nx_n) + \beta(a_1y_1 + \cdots + a_ny_n) = 0.$$
   On a donc $(z_1, \ldots, z_n) \in G$ et G est un sous-espace vectoriel de E.
2. L'ensemble $\mathbb{K}_n[X]$ est un sous-espace vectoriel du $\mathbb{K}$-espace vectoriel $\mathbb{K}[X]$. En effet, si P et Q sont des polynômes de degré inférieur ou égal à n, alors toute combinaison linéaire de P et Q le sera également, ce qui signifie que $\mathbb{K}_n[X]$ est stable par combinaison linéaire.
3. Dans le $\mathbb{K}$-espace vectoriel $E = \mathcal{A}(\mathbb{K}, \mathbb{K})$ des applications de $\mathbb{K}$ dans $\mathbb{K}$, on vérifie très facilement que les sous-ensembles des applications paires $\mathcal{A}_P(\mathbb{K}, \mathbb{K})$ et des applications impaires $\mathcal{A}_I(\mathbb{K}, \mathbb{K})$ sont stables par combinaison linéaire et sont par conséquent des sous-espaces vectoriels de E.

Contre-exemples.Dans le $\mathbb{R}$-espace vectoriel $E = \mathbb{R}^2$, les sous-ensembles $F = \{(\lambda, \lambda + 1),\ \lambda \in \mathbb{K}\}$, $G = \{(\lambda, \lambda^2),\ \lambda \in \mathbb{K}\}$, $H = \{(x, y) \in E : x + y = 1\}$ et $K = \{(x, y) \in E : x^3 - y = 0\}$ ne sont pas des sous-espaces vectoriels comme on peut le voir facilement en constatant que F et G ne contiennent pas le vecteur nul et G et H ne sont pas stables par multiplication scalaire.

À partir de plusieurs sous-espaces vectoriels, on peut construire facilement d'autres espaces vectoriels par intersection, somme, sous-espace vectoriel engendré... Ces diverses notions sont présentées ci-après.

3.2 Intersection de deux sous-espaces vectoriels

Proposition 3.3 *Soit F et G deux sous-espaces vectoriels de E. Alors $F \cap G$ est un sous-espace vectoriel de E.*

DÉMONSTRATION.

Soit $\alpha, \beta \in \mathbb{K}$ et $\vec{u}, \vec{v} \in F \cap G$. Comme $\vec{u}, \vec{v} \in F$, on $\alpha\vec{u} + \beta\vec{v} \in F$. De même $\vec{u}, \vec{v} \in G$ et donc $\alpha\vec{u} + \beta\vec{v} \in G$. Ainsi $\alpha\vec{u} + \beta\vec{v} \in F \cap G$, ce qui montre que $F \cap G$ est un sous-espace vectoriel de E. □

REMARQUE.

Plus généralement, si $F_1, \ldots, F_n$ sont des sous-espaces vectoriels de E, alors $F_1 \cap \ldots \cap F_n$ est un sous-espace vectoriel de E.

EXEMPLE.

Considérons dans le $\mathbb{K}$-espace vectoriel $E = \mathbb{K}^n$ le sous-ensemble

$$F = \{(x_1, \ldots, x_n) \in E : a_1x_1 + \cdots + a_nx_n = 0 \text{ et } b_1x_1 + \cdots + b_nx_n = 0\}$$

où les scalaires $a_1, \ldots, a_n$ sont non tous nuls ainsi que les scalaires $b_1, \ldots, b_n$. En écrivant F sous la forme $G \cap H$ avec

$$G = \{(x_1, \ldots, x_n) \in E : a_1x_1 + \cdots + a_nx_n = 0 \quad \text{et} \quad H = \{(x_1, \ldots, x_n) \in E : a_1x_1 + \cdots + a_nx_n = 0\},$$

G et H étant des sous-espaces vectoriels de E, on voit que F est un sous-espace vectoriel de E.

La réunion de deux sous-espaces vectoriels n'est en général pas un espace vectoriel (sauf si l'un des deux est contenu dans l'autre). On va alors définir un sous-espace vectoriel « minimal » contenant l'ensemble $F \cup G$. Le terme minimal est à prendre au sens suivant : si H est ce sous-espace vectoriel minimal et si K est un sous-espace vectoriel contenant $F \cup G$, alors $H \subset K$. On précise cette notion dans le paragraphe suivant.

3.3 Sous-espace vectoriel engendré par une partie

Définition 3.4 *Soit A une partie de E. Le sous-espace vectoriel engendré par A, noté* $\mathrm{vect}(A)$*, est le plus petit sous-espace vectoriel de E contenant A. En d'autres termes,* $\mathrm{vect}(A)$ *est caractérisé par*

— $\mathrm{vect}(A)$ *est un sous-espace vectoriel de E ;*

— *si F est un sous-espace vectoriel de E contenant A, alors F contient aussi* $\mathrm{vect}(A)$ *:*

$$A \subset F \Longrightarrow \mathrm{vect}(A) \subset F$$

ou encore

$$A \subset F \subset \mathrm{vect}(A) \Longrightarrow F = \mathrm{vect}(A).$$

Propriétés immédiates.

1. Si F est un sous-espace vectoriel de E, on a $\mathrm{vect}(F) = F$.
2. Si A et B sont deux parties de E telles que $A \subset B$, on a $\mathrm{vect}(A) \subset \mathrm{vect}(B)$.
3. Le sous-espace vectoriel $\mathrm{vect}(A)$ est en fait l'intersection de tous les sous-espaces vectoriels de E contenant A.

Considérons le cas particulier où $A = \{\vec{u}_1, \ldots, \vec{u}_n\}$, les vecteurs $\vec{u}_1, \ldots, \vec{u}_n$ étant fixés.

— On a $\vec{u}_1, \ldots, \vec{u}_n \in \mathrm{vect}(A)$ donc $\forall \lambda_1, \ldots, \lambda_n \in \mathbb{K},\ \lambda_1\vec{u}_1 + \cdots + \lambda_n\vec{u}_n \in \mathrm{vect}(A)$. Introduisons alors l'ensemble $F = \{\lambda_1\vec{u}_1 + \cdots + \lambda_n\vec{u}_n, \lambda_1, \ldots, \lambda_n \in \mathbb{K}\}$. On vient de voir que $F \subset \mathrm{vect}(A)$. Bien sûr, F contient A (choisir des coefficients $\lambda_1, \ldots, \lambda_n$ tous nuls excepté un que l'on prendra égal à 1) et donc

$$A \subset F \subset \mathrm{vect}(A).$$

— Montrons que F est un sous-espace vectoriel de E. Soit $\alpha, \beta \in \mathbb{K}$ et $\vec{v} = \lambda_1\vec{u}_1 + \cdots + \lambda_n\vec{u}_n, \vec{w} = \mu_1\vec{u}_1 + \cdots + \mu_n\vec{u}_n \in F$. On a

$$\alpha\vec{v} + \beta\vec{w} = (\alpha\lambda_1 + \beta\mu_1)\vec{u}_1 + \cdots + (\alpha\lambda_n + \beta\mu_n)\vec{u}_n \in F$$

ce qui prouve l'assertion annoncée.

En conclusion, on a trouvé un sous-espace vectoriel F de E contenant A et contenu dans $\mathrm{vect}(A)$. Par définition de $\mathrm{vect}(A)$, F coïncide nécessairement avec $\mathrm{vect}(A)$. On a démontré le résultat ci-dessous.

Proposition 3.5 *Soit $\vec{u}_1, \ldots, \vec{u}_n \in E$. Le sous-espace vectoriel de E engendré par les vecteurs $\vec{u}_1, \ldots, \vec{u}_n$ est l'ensemble des combinaisons linéaires de ces vecteurs :*

$$\mathrm{vect}(\{\vec{u}_1, \ldots, \vec{u}_n\}) = \{\lambda_1\vec{u}_1 + \cdots + \lambda_n\vec{u}_n, \lambda_1, \ldots, \lambda_n \in \mathbb{K}\}.$$

Exemples.

1. Si $\vec{u}$ est un vecteur non nul de E, $\mathrm{vect}(\vec{u}) = \{\lambda\vec{u}, \lambda \in \mathbb{K}\}$; c'est la *droite vectorielle* engendrée par $\vec{u}$.
2. Si $\vec{u}$ et $\vec{v}$ sont deux vecteurs non colinéaires de E, $\mathrm{vect}(\vec{u}, \vec{v}) = \{\lambda\vec{u} + \mu\vec{v}, \lambda, \mu \in \mathbb{K}\}$; c'est le *plan vectoriel* engendré par $\vec{u}$ et $\vec{v}$.

DESSIN A FAIRE

3.4 Somme de deux sous-espaces vectoriels

Soit F et G deux sous-espaces vectoriels de E. On a déjà signalé que la réunion $F \cup G$ n'est en général pas un espace vectoriel. Essayons donc de caractériser $\mathrm{vect}(F \cup G)$.

— Si $\vec{u} \in F$ et $\vec{v} \in G$, alors $\vec{u}, \vec{v} \in F \cup G$ et donc $\vec{u} + \vec{v} \in \mathrm{vect}(F \cup G)$. Introduisons alors $H = \{\vec{u} + \vec{v}, (\vec{u}, \vec{v}) \in F \times G\}$. On vient de voir que $H \subset \mathrm{vect}(F \cup G)$. D'autre part, F et G sont évidemment contenus dans H (écrire $\vec{u} = \vec{u} + \vec{0}$ pour $\vec{u} \in F$ et $\vec{v} = \vec{0} + \vec{v}$ pour $\vec{v} \in G$) et donc
$$F \cup G \subset H \subset \mathrm{vect}(F \cup G).$$

— Montrons que H est un sous-espace vectoriel de E. Soit $\alpha_1, \alpha_2 \in \mathbb{K}$ et $\vec{w}_1, \vec{w}_2 \in H$. Les vecteurs $\vec{w}_1$ et $\vec{w}_2$ s'écrivent sous la forme $\vec{w}_1 = \vec{u}_1 + \vec{v}_1$ et $\vec{w}_2 = \vec{u}_2 + \vec{v}_2$ avec $(\vec{u}_1, \vec{v}_1) \in F \times G$ et $(\vec{u}_2, \vec{v}_2) \in F \times G$. Puisque $\alpha_1 \vec{u}_1 + \alpha_2 \vec{u}_2 \in F$ et $\alpha_1 \vec{v}_1 + \alpha_2 \vec{v}_2 \in G$, on a
$$\alpha_1 \vec{w}_1 + \alpha_2 \vec{w}_2 = (\alpha_1 \vec{u}_1 + \alpha_2 \vec{u}_2) + (\alpha_1 \vec{v}_1 + \alpha_2 \vec{v}_2) \in H$$
ce qui montre bien que H est un sous-espace vectoriel de E.

On a ainsi construit un sous-espace vectoriel H de E contenant $F \cup G$ et contenu dans $\mathrm{vect}(F \cup G)$, d'où $H = \mathrm{vect}(F \cup G)$.

Définition 3.6 *Soit F et G deux sous-espaces vectoriels de E. La somme de F et G est le sous-ensemble défini par $F + G = \{\vec{u} + \vec{v}, (\vec{u}, \vec{v}) \in F \times G\}$. C'est le sous-espace vectoriel engendré par $F \cup G$: $F + G = \mathrm{vect}(F \cup G)$.*

Définition 3.7 *Soit F et G deux sous-espaces vectoriels de E. La somme de $F+G$ est directe lorsque la décomposition $\vec{w} = \vec{u} + \vec{v}$ avec $(\vec{u}, \vec{v}) \in F \times G$ d'un vecteur quelconque $\vec{w}$ de E est unique. On note $F + G$ dans ce cas $F \oplus G$.*

On définit plus généralement la *somme de n sous-espaces vectoriels* $F_1, \ldots, F_n$ de E par
$$F_1 + \cdots + F_n = \mathrm{vect}(F_1 \cup \ldots \cup F_n) = \{\vec{u}_1 + \cdots + \vec{u}_n, (\vec{u}_1, \ldots, \vec{u}_n) \in F_1 \times \cdots \times F_n\}.$$
Lorsque la décomposition $\vec{w} = \vec{u}_1 + \cdots + \vec{u}_n$ avec $(\vec{u}_1, \ldots, \vec{u}_n) \in F_1 \times \cdots \times F_n$ d'un vecteur quelconque $\vec{w}$ de E est unique, on dit que la somme $F_1 + \cdots + F_n$ est *directe* et on la note dans ce cas $F_1 \oplus \cdots \oplus F_n$.

Proposition 3.8 *La somme $F + G$ est directe si et seulement si $F \cap G = \{\vec{0}\}$.*

Démonstration.

1. Supposons la somme $F + G$ directe. Soit $\vec{u} \in F \cap G$. On a les décompositions triviales $\vec{u} = \vec{u} + \vec{0} = \vec{0} + \vec{u}$ avec $(\vec{u}, \vec{0}) \in F \times G$ et $(\vec{0}, \vec{u}) \in F \times G$. Par unicité, on a nécessairement $\vec{u} = \vec{0}$ ce qui montre que $F \cap G = \{\vec{0}\}$.
2. Supposons réciproquement que $F \cap G = \{\vec{0}\}$. Soit $\vec{w} \in F + G$. Si $\vec{w}$ admet deux décompositions de la forme $\vec{w} = \vec{u}_1 + \vec{v}_1 = \vec{u}_2 + \vec{v}_2$ avec $(\vec{u}_1, \vec{v}_1) \in F \times G$ et $(\vec{u}_2, \vec{v}_2) \in F \times G$, alors $\vec{u}_1 - \vec{u}_2 = \vec{v}_2 - \vec{v}_1$. Comme $\vec{u}_1 - \vec{u}_2 \in F$ et $\vec{v}_2 - \vec{v}_1 \in G$, on a $\vec{u}_1 - \vec{u}_2, \vec{v}_2 - \vec{v}_1 \in F \cap G$. Or $F \cap G = \{\vec{0}\}$, donc $\vec{u}_1 - \vec{u}_2 = \vec{v}_2 - \vec{v}_1 = \vec{0}$, soit $(\vec{u}_1, \vec{v}_1) = (\vec{u}_2, \vec{v}_2)$, ce qui montre que la somme $F + G$ est directe. □

Exemple.

Dans le $\mathbb{K}$-espace vectoriel $E = \mathbb{K}^n$, étudions la somme des deux sous-espaces vectoriels $F = \{(\lambda a_1, \ldots, \lambda a_n),\ \lambda \in \mathbb{K}\}$ et $G = \{(x_1, \ldots, x_n) \in E : b_1 x_1 + \cdots + b_n x_n = 0\}$ où les scalaires $a_1, \ldots, a_n$ sont non tous nuls ainsi que les scalaires $b_1, \ldots, b_n$. Pour cela, on détermine l'intersection $F \cap G$. Soit $\vec{u} = (x_1, \ldots, x_n) \in F \cap G$. Le vecteur $\vec{u}$ est en particulier de la forme $\vec{u} = (\lambda a_1, \ldots, \lambda a_n)$ avec $\lambda \in \mathbb{K}$, donc, en reportant $x_1 = \lambda a_1, \ldots, x_n = \lambda a_n$ dans l'équation caractérisant G, on trouve

$$b_1 x_1 + \cdots + b_n x_n = \lambda(a_1 b_1 + \cdots + a_n b_n) = 0.$$

Ainsi,
- si $a_1 b_1 + \cdots + a_n b_n \neq 0$, on a $\lambda = 0$, donc $\vec{u} = \vec{0}$ et $F \cap G = \{\vec{0}\}$. Dans ce cas la somme $F + G$ est directe ;
- si $a_1 b_1 + \cdots + a_n b_n = 0$, on a $F \subset G$ et alors $F + G = G$.

3.5 Sous-espaces vectoriels supplémentaires

Définition 3.9 *Les sous-espaces vectoriels F et G sont supplémentaires dans E lorsque $F \oplus G = E$; en d'autres termes,*

$$\forall \vec{w} \in E,\ \exists!(\vec{u}, \vec{v}) \in F \times G,\ \vec{w} = \vec{u} + \vec{v}.$$

Plus généralement, on dit que n sous-espaces vectoriels $F_1, \ldots, F_n$ sont *supplémentaires dans E* lorsque $F_1 \oplus \cdots \oplus F_n = E$.

D'après le paragraphe précédent, on a immédiatement l'assertion suivante.

Proposition 3.10 *Les sous-espaces vectoriels F et G sont supplémentaires dans E si et seulement si $F + G = E$ et $F \cap G = \{\vec{0}\}$.*

Exemples.

1. Soit, dans le $\mathbb{R}$-espace vectoriel $E = \mathbb{R}^2$, les sous-espaces vectoriels F et G définis par $F = \{(2\lambda, 3\lambda),\ \lambda \in \mathbb{R}\}$ et $G = \{(x, y) \in E : 4x - y = 0\}$. Soit $\vec{w} = (x, y) \in E$. Cherchons s'il existe un couple $(\vec{u}, \vec{v}) \in F \times G$ tel que $\vec{w} = \vec{u} + \vec{v}$. En écrivant G sous la forme $\{(\mu, 4\mu),\ \mu \in \mathbb{R}\}$, on cherche s'il existe une décomposition de la forme $\vec{w} = \vec{u} + \vec{v}$ avec $\vec{u} = (2\lambda, 3\lambda)$ et $\vec{v} = (\mu, 4\mu)$. Cela conduit au système $\begin{cases} 2\lambda + \mu = x \\ 3\lambda + 4\mu = y \end{cases}$ dont la solution est $\lambda = \dfrac{4x - y}{5}$, $\mu = \dfrac{-3x + 2y}{5}$, ce qui démontre l'existence et l'unicité du couple $(\vec{u}, \vec{v}) \in F \times G$ tel que $\vec{w} = \vec{u} + \vec{v}$. D'où $F \oplus G = E$ et F et G sont supplémentaires dans E.
2. Dans le $\mathbb{R}$-espace vectoriel $E = \mathcal{A}(\mathbb{R}, \mathbb{R})$ des applications de $\mathbb{R}$ dans $\mathbb{R}$, on considère le sous-espace vectoriel des applications paires $F = \mathcal{A}_P(\mathbb{R}, \mathbb{R})$ et celui des applications

impaires $G = \mathcal{A}_I(\mathbb{R}, \mathbb{R})$. Soit $h \in E$. Cherchons s'il existe un couple $(f, g) \in F \times G$ tel que $h = f + g$. Si une telle décomposition existe, on a nécessairement

$$h(x) = f(x) + g(x) \quad \text{et} \quad h(-x) = f(x) - g(x),$$

ce qui entraîne

$$f(x) = \frac{h(x) + h(-x)}{2} \quad \text{et} \quad g(x) = \frac{h(x) - h(-x)}{2}.$$

On vérifie bien que les applications f et g ainsi construites sont *a posteriori* respectivement paire et impaire. On a également prouvé l'unicité de la décomposition $h = f + g$ et donc $E = F \oplus G$.

4 Systèmes de vecteurs, dimension

Avec un système de deux vecteurs non colinéaires $\mathcal{S} = \{\vec{u}_1, \vec{u}_2\}$, il est possible de reconstruire tout l'ensemble des vecteurs du plan vectoriel $E = \mathrm{vect}(\mathcal{S})$ par combinaisons linéaires de $\vec{u}_1$ et $\vec{u}_2$. Si on enlève un vecteur au système $\mathcal{S}$, on ne peut plus reconstruire ce plan vectoriel (on dit que $\mathcal{S}$ est un *système générateur minimal*) ; si on ajoute un vecteur $\vec{u}_3$ à $\mathcal{S}$, on peut de nouveau reconstruire ce plan vectoriel mais $\vec{u}_3$ n'est pas indispensable à cette reconstruction puisqu'il est en fait combinaison linéaire de $\vec{u}_1$ et $\vec{u}_2$ (on dit que $\mathcal{S}$ est un *système libre maximal*). En fait le système $\mathcal{S}$ est une *base* du plan vectoriel E. On va étendre ces trois notions de systèmes libres, systèmes générateurs et de bases à un espace vectoriel quelconque.

Dans tout ce paragraphe, E désigne un $\mathbb{K}$-espace vectoriel et tous les systèmes de vecteurs de E considérés seront supposés **finis**.

Remarque.

Il y a une ambiguïté de notation en adoptant l'écriture ensembliste $\{\vec{u}_1, \dots, \vec{u}_n\}$ car un système de vecteurs peut contenir des vecteurs identiques contrairement à la coutume générale de la théorie des ensembles qui stipule que l'on ne doit écrire un élément qu'une seule fois dans un ensemble. De plus, l'ordre dans lequel sont écrits les vecteurs d'un système sera important notamment dans le cas des bases. Certains auteurs adoptent une écriture de type liste $(\vec{u}_1, \dots, \vec{u}_n)$ qui autorise maintenant la répétition de vecteurs identiques et respecte l'ordre d'écriture mais qui pose des problèmes de cohérence pour construire des intersections ou des réunions de systèmes de vecteurs...

4.1 Systèmes libres, liés

Définition 4.1 *Le système de vecteurs $\mathcal{S} = \{\vec{u}_1, \dots, \vec{u}_p\}$ est libre lorsque*

$$\forall \alpha_1, \dots, \alpha_p \in \mathbb{K}, \ [(\alpha_1 \vec{u}_1 + \cdots + \alpha_p \vec{u}_p = \vec{0} \Longrightarrow \alpha_1 = \cdots = \alpha_p = 0.]$$

On dit aussi que les vecteurs $\vec{u}_1, \dots, \vec{u}_p$ sont linéairement indépendants.
Dans le cas contraire, le système $\mathcal{S}$ est lié. Cela signifie qu'il existe des scalaires $\alpha_1, \dots, \alpha_p \in \mathbb{K}$ non tous nuls tels que $\alpha_1 \vec{u}_1 + \cdots + \alpha_p \vec{u}_p = \vec{0}$.

Propriétés immédiates.

1. Si le système $\mathcal{S}$ contient le vecteur nul, alors il est lié. S'il contient deux vecteurs identiques, il est aussi lié.
2. Une autre façon de dire que le système $\mathcal{S}$ est lié est qu'il existe un vecteur de $\mathcal{S}$ qui est combinaison linéaire des autres vecteurs de $\mathcal{S}$.
3. Si le système $\mathcal{S}$ est libre, alors

$$\alpha_1\vec{u}_1 + \cdots + \alpha_p\vec{u}_p = \beta_1\vec{u}_1 + \cdots + \beta_p\vec{u}_p \Longrightarrow \alpha_1 = \beta_1, \ldots, \alpha_p = \beta_p.$$

4. — Pour $n = 1$, le système $\mathcal{S} = \{\vec{u}_1\}$ est libre si et seulement si le vecteur $\vec{u}_1$ est non nul.
   — Pour $n = 2$, le système $\mathcal{S} = \{\vec{u}_1, \vec{u}_2\}$ est libre si et seulement si les vecteurs $\vec{u}_1, \vec{u}_2$ ne sont pas colinéaires.
   — Pour $n = 3$, le système $\mathcal{S} = \{\vec{u}_1, \vec{u}_2, \vec{u}_3\}$ est libre si et seulement si les vecteurs $\vec{u}_1, \vec{u}_2, \vec{u}_3$ ne sont pas coplanaires.

Proposition 4.2 *Soit $\mathcal{L}$ un système libre et $\vec{u} \in E$.*

1. *Si $\vec{u} \in \mathcal{L}$, le système $\mathcal{L} \setminus \{\vec{u}\}$ est encore libre.*
2. *Le système $\mathcal{L} \cup \{\vec{u}\}$ est libre si et seulement si $\vec{u}$ n'est pas combinaison linéaire des vecteurs de $\mathcal{L}$, i.e. $\vec{u} \notin \mathrm{vect}(\mathcal{L})$.*

Démonstration.

Posons $\mathcal{L} = \{\vec{u}_1, \ldots, \vec{u}_p\}$.

1. Supposons par exemple que $\vec{u} = \vec{u}_p$. Soit $\alpha_1, \ldots, \alpha_{p-1} \in \mathbb{K}$ tels que $\alpha_1\vec{u}_1 + \cdots + \alpha_{p-1}\vec{u}_{p-1} = \vec{0}$. On a, en posant $\alpha_p = 0$, la relation $\alpha_1\vec{u}_1 + \cdots + \alpha_p\vec{u}_p = \vec{0}$ de laquelle on déduit, puisque le système $\mathcal{L}$ est libre, $\alpha_1 = \cdots = \alpha_p = 0$ et le système $\mathcal{L} \setminus \{\vec{u}\} = \{\vec{u}_1, \ldots, \vec{u}_{p-1}\}$ est libre.
2. On va montrer la contraposée de la seconde assertion : le système $\mathcal{L} \cup \{\vec{u}\}$ est lié si et seulement si $\vec{u}$ est combinaison linéaire des autres vecteurs de $\mathcal{L}$.
   (a). Supposons le système $\mathcal{L} \cup \{\vec{u}\}$ lié. Il existe des scalaires $\alpha_1, \ldots, \alpha_{p+1} \in \mathbb{K}$ non tous nuls tels que $\alpha_1\vec{u}_1 + \cdots + \alpha_p\vec{u}_p + \alpha_{p+1}\vec{u} = \vec{0}$. Le scalaire α_{p+1} est nécessairement non nul, sinon on aurait $\alpha_1\vec{u}_1 + \cdots + \alpha_p\vec{u}_p = \vec{0}$ et le système $\mathcal{L}$ serait lié. Donc $\vec{u} = -\dfrac{\alpha_1}{\alpha_{p+1}}\vec{u}_1 - \cdots - \dfrac{\alpha_p}{\alpha_{p+1}}\vec{u}_p$ ce qui signifie que $\vec{u}$ est combinaison linéaire des autres vecteurs de $\mathcal{L}$.
   (b). Supposons que $\vec{u}$ soit combinaison linéaire des vecteurs de $\mathcal{L}$. On sait que le système $\mathcal{L} \cup \{\vec{u}\}$ est alors lié. □

4.2 Systèmes générateurs

Définition 4.3 *Le système de vecteurs $\mathcal{S} = \{\vec{u}_1, \ldots, \vec{u}_p\}$ est générateur de E lorsque*

$$\forall \vec{u} \in E,\ \exists \alpha_1, \ldots, \alpha_p \in \mathbb{K},\ \vec{u} = \alpha_1\vec{u}_1 + \cdots + \alpha_p\vec{u}_p.$$

Cela s'écrit encore $\mathrm{vect}(\mathcal{S}) = E$. On dit aussi que le système $\mathcal{S}$ engendre E.

Propriété immédiate.

Le système $\mathcal{S}$ est générateur du sous-espace vectoriel qu'il engendre, *i.e.* $\mathrm{vect}(\mathcal{S})$, puisque ce dernier est l'ensemble des combinaisons linéaires des vecteurs de $\mathcal{S}$.

Proposition 4.4 *Soit $\mathcal{G}$ un système générateur de E et $\vec{u} \in E$.*

1. *Le système $\mathcal{G} \cup \{\vec{u}\}$ est encore générateur de E.*
2. *Si $\vec{u} \in \mathcal{G}$, le système $\mathcal{G} \setminus \{\vec{u}\}$ est générateur de E si et seulement si $\vec{u}$ est combinaison linéaire des autres vecteurs de $\mathcal{G}$, i.e. $\vec{u} \in \mathrm{vect}(\mathcal{G} \setminus \{\vec{u}\})$.*

Démonstration.

Posons $\mathcal{G} = \{\vec{u}_1, \dots, \vec{u}_p\}$.

1. Soit $\vec{v} \in E$. Le système $\mathcal{G}$ étant générateur de E, il existe des scalaires $\alpha_1, \dots, \alpha_p \in \mathbb{K}$ tels que $\vec{v} = \alpha_1\vec{u}_1 + \cdots + \alpha_p\vec{u}_p$ et donc, en posant $\alpha_{p+1} = 0$,
$$\vec{v} = \alpha_1\vec{u}_1 + \cdots + \alpha_p\vec{u}_p + \alpha_{p+1}\vec{u}.$$
Cela prouve que le système $\mathcal{G} \cup \{\vec{u}\}$ est encore générateur de E.
2. Admettons par exemple que $\vec{u} = \vec{u}_p$.
   (a). Supposons le système $\mathcal{G} \setminus \{\vec{u}\} = \{\vec{u}_1, \dots, \vec{u}_{p-1}\}$ générateur de E. Tout vecteur de E est combinaison linéaire des vecteurs de $\mathcal{G} \setminus \{\vec{u}\}$. En particulier, le vecteur $\vec{u}$ est combinaison linéaire des vecteurs de $\mathcal{G} \setminus \{\vec{u}\}$.
   (b). Supposons que $\vec{u}$ soit combinaison linéaire des vecteurs de $\mathcal{G} \setminus \{\vec{u}\}$. Il existe donc des scalaires $\alpha_1, \dots, \alpha_{p-1} \in \mathbb{K}$ tels que $\vec{u} = \alpha_1\vec{u}_1 + \cdots + \alpha_{p-1}\vec{u}_{p-1}$. Soit maintenant $\vec{v} \in E$. $\mathcal{G}$ étant un système générateur de E, il existe des scalaires $\beta_1, \dots, \beta_p \in \mathbb{K}$ tels que $\vec{v} = \beta_1\vec{u}_1 + \cdots + \beta_p\vec{u}_p$. On a alors
$$\begin{aligned}\vec{v} &= \beta_1\vec{u}_1 + \cdots + \beta_{p-1}\vec{u}_{p-1} + \beta_p(\alpha_1\vec{u}_1 + \cdots + \alpha_{p-1}\vec{u}_{p-1}) \\ &= (\beta_1 + \alpha_1\beta_p)\vec{u}_1 + \cdots + (\beta_{p-1} + \alpha_{p-1}\beta_p)\vec{u}_{p-1}.\end{aligned}$$
   Ainsi, $\vec{v}$ est combinaison linéaire des vecteurs de $\mathcal{G} \setminus \{\vec{u}\}$ et $\mathcal{G} \setminus \{\vec{u}\}$ est un système générateur de E. □

Proposition 4.5 *Soit F et G des sous-espaces vectoriels de E. Si $\mathcal{G}_F$ et $\mathcal{G}_G$ sont des systèmes générateurs de F et G respectivement, alors $\mathcal{G}_F \cup \mathcal{G}_G$ est un système générateur de $F + G$.*

Démonstration.

Posons $\mathcal{G}_F = \{\vec{u}_1, \dots, \vec{u}_p\}$ et $\mathcal{G}_G = \{\vec{v}_1, \dots, \vec{v}_q\}$. On a $\mathcal{G}_F \cup \mathcal{G}_G = \{\vec{u}_1, \dots, \vec{u}_p, \vec{v}_1, \dots, \vec{v}_q\}$. Tout vecteur de $F + G$ est de la forme $\vec{w} = \vec{u} + \vec{v}$ avec $(\vec{u}, \vec{v}) \in F \times G$. De plus, il existe des scalaires $\alpha_1, \dots, \alpha_p, \beta_1, \dots, \beta_q$ tels que $\vec{u} = \alpha_1\vec{u}_1 + \cdots + \alpha_p\vec{u}_p$ et $\vec{v} = \beta_1\vec{v}_1 + \cdots + \beta_p\vec{v}_q$ et donc
$$\vec{w} = \alpha_1\vec{u}_1 + \cdots + \alpha_p\vec{u}_p + \beta_1\vec{v}_1 + \cdots + \beta_p\vec{v}_q$$
ce qui montre bien que le système $\mathcal{G}_F \cup \mathcal{G}_G$ est générateur de $F + G$. □

4.3 Bases, dimension

Le nombre de vecteurs dans un système est fondamental dans l'étude des espaces vectoriels. La théorie de la dimension a pour objectif de définir un nombre optimal (à la fois minimal et maximal) de vecteurs nécessaires pour reconstruire l'espace E tout entier par combinaison linéaire et ce sans redondance. Cette théorie repose sur des inégalités entre le nombre de vecteurs d'un système libre et celui d'un système générateur de E.

Lemme fondamental 4.6 *Soit $k \in \mathbb{N}^*$. Alors $k+1$ vecteurs combinaisons linéaires de k vecteurs forment un système lié.*

Démonstration.

Examinons d'abord quelques exemples.

— *Cas $k = 1$:* considérons le système $\mathcal{S} = \{\vec{u}, \vec{v}\}$ avec $\vec{u} = \alpha\vec{i}$ et $\vec{v} = \beta\vec{i}$. Si $\beta = 0$, on a $\vec{v} = \vec{0}$ et le système $\mathcal{S}$ est lié. Si $\beta \neq 0$, on a $\vec{i} = \dfrac{1}{\beta}\,\vec{v}$ puis $\vec{u} = \dfrac{\alpha}{\beta}\,\vec{v}$ et le système $\mathcal{S}$ est encore lié.

— *Cas $k = 2$:* considérons le système $\mathcal{S} = \{\vec{u}, \vec{v}, \vec{w}\}$ avec $\vec{u} = \alpha\vec{i} + \beta\vec{j}$, $\vec{v} = \alpha'\vec{i} + \beta'\vec{j}$ et $\vec{w} = \alpha''\vec{i} + \beta''\vec{j}$. Si $\alpha'' = \beta'' = 0$, on a $\vec{w} = \vec{0}$ et le système $\mathcal{S}$ est lié. Sinon, supposons par exemple $\beta'' \neq 0$. On a $\vec{j} = \dfrac{1}{\beta''}\,\vec{w} - \dfrac{\alpha''}{\beta''}\,\vec{i}$ que l'on reporte dans $\vec{u}$ et $\vec{v}$:

$$\vec{u} = (\alpha - \frac{\beta}{\beta''}\,\alpha'')\vec{i} + \frac{\beta}{\beta''}\,\vec{w} \quad \text{et} \quad \vec{v} = (\alpha' - \frac{\beta'}{\beta''}\,\alpha'')\vec{i} + \frac{\beta'}{\beta''}\,\vec{w}.$$

On voit alors que les vecteurs $\vec{u} - \dfrac{\beta}{\beta''}\,\vec{w}$ et $\vec{v} - \dfrac{\beta'}{\beta''}\,\vec{w}$ sont des combinaisons linéaires du vecteur $\vec{i}$ et, d'après le cas précédent, ils forment un système lié :

$$\exists a, b \in \mathbb{K} \text{ non tous nuls, } a(\vec{u} - \frac{\beta}{\beta''}\,\vec{w}) + b(\vec{v} - \frac{\beta'}{\beta''}\,\vec{w}) = \vec{0}.$$

Cette relation est de la forme $a\vec{u} + b\vec{v} + c\vec{w} = \vec{0}$ où les scalaires a, b, c sont non tous nuls et le système $\mathcal{S}$ est donc lié.

— *Cas général :* La démonstration se fait par récurrence sur k. Admettons que le résultat soit vrai pour $k-1$. Soit alors $\mathcal{S} = \{\vec{u}_1, \dots, \vec{u}_{k+1}\}$ un système de $k+1$ vecteurs combinaisons linéaires de k vecteurs $\vec{e}_1, \dots, \vec{e}_k$:

$$\begin{array}{rcl} \vec{u}_1 & = & \alpha_{11}\,\vec{e}_1 + \cdots + \alpha_{1k}\,\vec{e}_k \\ & \vdots & \\ \vec{u}_{k+1} & = & \alpha_{k+1\,1}\,\vec{e}_1 + \cdots + \alpha_{k+1\,k}\,\vec{e}_k \end{array}$$

Si tous les coefficients de la dernière ligne $\alpha_{k+1\,1}, \alpha_{k+1\,2}, \dots, \alpha_{k+1\,k}$ sont nuls, on a $\vec{u}_{k+1} = \vec{0}$ et le système $\mathcal{S}$ est lié. Sinon supposons par exemple que $\alpha_{k+1\,k} \neq 0$. On tire alors

$$\vec{e}_k = \frac{1}{\alpha_{k+1k}}\,\vec{u}_{k+1} - \frac{\alpha_{k+11}}{\alpha_{k+1k}}\,\vec{e}_1 - \cdots - \frac{\alpha_{k+1k-1}}{\alpha_{k+1k}}\,\vec{e}_{k-1}$$

et, en reportant dans les autres égalités, on voit que les k vecteurs $\vec{u}_1 - \dfrac{\alpha_{1k}}{\alpha_{k+1k}}\,\vec{u}_{k+1}, \dots, \vec{u}_k - \dfrac{\alpha_{kk}}{\alpha_{k+1k}}\,\vec{u}_{k+1}$ sont des combinaisons linéaires des $k-1$ vecteurs $\vec{e}_1, \dots, \vec{e}_{k-1}$. Ils forment un système lié d'après l'hypothèse de récurrence. Par conséquent le système $\mathcal{S}$ est lié. □

Corollaire 4.7 *Soit $\mathcal{L}$ un système libre et $\mathcal{G}$ un système générateur de E. On a nécessairement*

$$\mathrm{card}(\mathcal{L}) \leqslant \mathrm{card}(\mathcal{G}).$$

Démonstration.

Posons $p = \text{card}(\mathcal{L})$ et $q = \text{card}(\mathcal{G})$. Le système $\mathcal{G}$ étant générateur de E, les p vecteurs de $\mathcal{L}$ sont des combinaisons linéaires des q vecteurs de $\mathcal{G}$. Si on avait $p > q$ (*i.e.* $p \geqslant q+1$), le système $\mathcal{L}$ contiendrait un sous-système de $q+1$ vecteurs combinaisons linéaires de q vecteurs. Ce dernier est lié d'après le lemme fondamental et $\mathcal{L}$ serait également lié, ce qui est contraire à l'hypothèse. Ainsi $p \leqslant q$. □

Définition 4.8 *Une base de E est un système libre et générateur de E.*

Proposition 4.9 *Le système de vecteurs $\mathcal{S} = \{\vec{u}_1, \dots, \vec{u}_n\}$ est une base de E si et seulement si*

$$\forall \vec{u} \in E,\ \exists!(\alpha_1, \dots, \alpha_n) \in \mathbb{K}^n,\ \vec{u} = \alpha_1\vec{u}_1 + \dots + \alpha_n\vec{u}_n.$$

Les coefficients $\alpha_1, \dots, \alpha_n$ sont les coordonnées du vecteur $\vec{u}$ relativement à la base $\mathcal{S}$.

Remarque.

L'ordre dans lequel sont écrits les vecteurs $\vec{u}_1, \dots, \vec{u}_n$ de la base $\mathcal{S}$ est important et les coordonnées d'un vecteur de E doivent respecter le même ordre.

Exemples.

1. Dans le $\mathbb{K}$-espace vectoriel $\mathbb{K}^n$, on peut écrire

$$\begin{aligned}(x_1, x_2, \dots, x_n) &= x_1(1,0,0,\dots,0) + x_2(0,1,0,\dots,0) + \dots + x_n(0,0,0,\dots,1) \\ &= x_1\vec{e}_1 + x_2\vec{e}_2 + \dots + x_n\vec{e}_n\end{aligned}$$

où

$$\vec{e}_i = (0,\dots,0,\underset{\substack{\uparrow \\ i^{\text{e}}\text{position}}}{1},0,\dots,0) \text{ pour } 1 \leqslant i \leqslant n.$$

Cela montre que le système $\mathcal{S} = \{\vec{e}_1, \dots, \vec{e}_n\}$ est générateur de $\mathbb{K}^n$. Par ailleurs, si $\alpha_1\vec{e}_1 + \dots + \alpha_n\vec{e}_n = \vec{0}$, alors, d'après le calcul précédent, on a $(\alpha_1, \dots, \alpha_n) = \vec{0}$ et donc $\alpha_1 = \dots = \alpha_n = 0$. Ainsi, le système $\mathcal{S}$ est libre. En conclusion, le système $\{\vec{e}_1, \dots, \vec{e}_n\}$ est une base de $\mathbb{K}^n$, on l'appelle la *base canonique* de $\mathbb{K}^n$. Dans l'écriture $\vec{u} = (x_1, x_2, \dots, x_n)$, les scalaires $x_1, \dots, x_n$ sont les coordonnées du vecteurs $\vec{u}$ relativement à la base canonique.

2. Tout polynôme de $\mathbb{K}_n[X]$ s'écrit par définition sous la forme $\displaystyle\sum_{i=0}^{n} a_i X^i$. Ceci montre que le système $\mathcal{S} = \{1, X, X^2, \dots, X^n\}$ est générateur du $\mathbb{K}$-espace vectoriel $\mathbb{K}_n[X]$. D'autre part, si $\displaystyle\sum_{i=0}^{n} a_i X^i = 0$, par définition de l'égalité de deux polynômes —

$$\sum_{i=0}^{n} a_i X^i = \sum_{i=0}^{n} b_i X^i \Longrightarrow a_0 = b_0, a_1 = b_1, \dots, a_n = b_n,$$

— on trouve immédiatement $a_0 = a_1 = \cdots = a_n = 0$ et le système $\mathcal{S}$ est libre. Ainsi, le système $\{1, X, X^2, \ldots, X^n\}$ est une base de $\mathbb{K}_n[X]$ appelée *base canonique* de $\mathbb{K}_n[X]$. Dans l'écriture $P = \sum_{i=0}^{n} a_i X^i$, les scalaires $a_0, \ldots, a_n$ sont les coordonnées du vecteur P relativement à la base canonique.

Proposition 4.10

1. *Les bases de E sont les systèmes libres maximaux de E.*
2. *Les bases de E sont les systèmes générateurs minimaux de E.*

Démonstration.

On utilise les propositions 4.2 et 4.4.

1.(a). Soit $\mathcal{B}$ une base de E ; c'est un système libre. Supposons qu'il ne soit pas maximal, c'est-à-dire qu'il existerait un vecteur $\vec{u} \in E \setminus \mathcal{B}$ tel que le système « augmenté » $\mathcal{B} \cup \{\vec{u}\}$ soit encore libre. Le vecteur $\vec{u}$ ne serait alors pas combinaison linéaire des vecteurs de $\mathcal{B}$, ce qui contredirait le fait que $\mathcal{B}$ est un système générateur de E. Le système libre $\mathcal{B}$ est donc maximal.

(b). Réciproquement, soit $\mathcal{B}$ un système libre maximal. Supposons qu'il ne soit pas générateur de E. Il existerait donc un vecteur $\vec{u} \in E$ qui ne serait pas combinaison linéaire des vecteurs de $\mathcal{B}$. Alors $\mathcal{B} \cup \{\vec{u}\}$ serait encore un système libre, ce qui contredirait le fait que $\mathcal{B}$ est un système libre maximal. Le système libre $\mathcal{B}$ est donc générateur de E et c'est une base de E.

2.(a). Soit $\mathcal{B}$ une base de E ; c'est un système générateur de E. Supposons qu'il ne soit pas minimal, c'est-à-dire qu'il existerait un vecteur $\vec{u} \in \mathcal{B}$ tel que le système « diminué » $\mathcal{B} \setminus \{\vec{u}\}$ est encore générateur de E. Le vecteur $\vec{u}$ serait alors combinaison linéaire des autres vecteurs de $\mathcal{B}$, ce qui contredirait le fait que $\mathcal{B}$ est un système libre. Le système générateur $\mathcal{B}$ est donc minimal.

(b). Réciproquement, soit $\mathcal{B}$ un système générateur minimal. Supposons qu'il ne soit pas libre. Il existerait donc un vecteur $\vec{u} \in \mathcal{B}$ qui serait combinaison linéaire des autres vecteurs de $\mathcal{B}$. Alors $\mathcal{B} \setminus \{\vec{u}\}$ serait encore un système générateur de E, ce qui contredirait le fait que $\mathcal{B}$ est un système générateur minimal. Le système générateur $\mathcal{B}$ est donc libre et c'est une base de E. □

Définition 4.11 *Le $\mathbb{K}$-espace vectoriel E est de dimension finie lorsqu'il admet un système générateur fini. Dans le contraire, E est de dimension infinie.*

Théorème 4.12 (Existence de bases en dimension finie) *Soit E un $\mathbb{K}$-espace vectoriel de dimension finie non réduit à son vecteur nul. Alors E admet au moins une base. De plus toutes les bases de E contiennent le même nombre de vecteurs, ce nombre commun est appelé dimension de E et noté* $\dim(E)$.

Si $E = \{\vec{0}\}$, on pose $\dim(E) = 0$. Lorsque $\dim(E) = 1$, on dit que E est une *droite vectorielle* et lorsque $\dim(E) = 2$, E est un *plan vectoriel*.

Démonstration.

1. Soit $\mathcal{S}$ un système générateur fini de E. Si $\mathcal{S}$ est minimal, c'est une base de E. Sinon, il n'est pas minimal et il existe alors un vecteur $\vec{u}_1 \in \mathcal{S}$ tel que le système $\mathcal{S}_1 = \mathcal{S} \setminus \{\vec{u}_1\}$ soit encore générateur de E. Si $\mathcal{S}_1$ est minimal, c'est une base de E. Sinon, il n'est pas minimal et il existe alors un vecteur $\vec{u}_2 \in \mathcal{S}_1$ tel que le système $\mathcal{S}_2 = \mathcal{S}_1 \setminus \{\vec{u}_2\}$ soit encore générateur de E. On construit de proche en proche une suite décroissante (au sens de l'inclusion) de systèmes $(\mathcal{S}_n)_{n\in\mathbb{N}}$ générateurs de E :
$$\mathcal{S} \supset \mathcal{S}_1 \supset \mathcal{S}_2 \supset \cdots \supset \emptyset.$$
La suite des cardinaux correspondants $(\mathrm{card}(\mathcal{S}_n))_{n\in\mathbb{N}}$ est donc une suite d'entiers décroissante minorée par 0, elle est nécessairement stationnaire :
$$\exists n_0 \in \mathbb{N},\, \forall n \geqslant n_0,\ \mathrm{card}(\mathcal{S}_n) = \mathrm{card}(\mathcal{S}_{n_0}).$$
Le système générateur $\mathcal{S}_{n_0}$ est alors minimal, c'est une base de E.
2. Soit $\mathcal{B}$ et $\mathcal{B}'$ deux bases de E. Notons $n = \mathrm{card}(\mathcal{B})$ et $n' = \mathrm{card}(\mathcal{B}')$.
   — D'une part, le système $\mathcal{B}$ étant libre et le système $\mathcal{B}'$ étant générateur, on a $n \leqslant n'$ d'après le corollaire 4.7.
   — D'autre part, le système $\mathcal{B}$ étant générateur et le système $\mathcal{B}'$ étant libre, on a de la même manière $n' \leqslant n$.

   Ainsi $n = n'$. □

Exemple.

Le $\mathbb{K}$-espace vectoriel $\mathbb{K}^n$ est de dimension n et le $\mathbb{K}$-espace vectoriel $\mathbb{K}_n[X]$ est de dimension $n+1$. Le $\mathbb{K}$-espace vectoriel $\mathbb{K}[X]$ contenant tous les sous-espaces vectoriels $\mathbb{K}_n[X]$ contient des systèmes libres de cardinal quelconque, il est donc de dimension infinie.

Théorème 4.13 *Soit E un $\mathbb{K}$-espace vectoriel de dimension n.*

1. *Soit $\mathcal{L}$ un système libre. On a $\mathrm{card}(\mathcal{L}) \leqslant n$ et*
$$\mathcal{L} \text{ est une base de } E \text{ si et seulement si } \mathrm{card}(\mathcal{L}) = n.$$
2. *Soit $\mathcal{G}$ un système générateur de E. On a $\mathrm{card}(\mathcal{G}) \geqslant n$ et*
$$\mathcal{G} \text{ est une base de } E \text{ si et seulement si } \mathrm{card}(\mathcal{G}) = n.$$

Démonstration.

Soit $\mathcal{B}$ une base de E. On a $\mathrm{card}(\mathcal{B}) = n$.

1.(a). Si $\mathcal{L}$ est un système libre, d'après le corollaire 4.7, $\mathcal{B}$ étant un système générateur de E, on a $\mathrm{card}(\mathcal{L}) \leqslant \mathrm{card}(\mathcal{B}) = n$.

(b). Si $\mathcal{L}$ est un système libre de n vecteurs, il est maximal et c'est une base de E.

2.(a). Si $\mathcal{G}$ est un système générateur de E, $\mathcal{B}$ étant un système libre, on a de manière analogue $\mathrm{card}(\mathcal{G}) \geqslant \mathrm{card}(\mathcal{B}) = n$.

(b). Si $\mathcal{G}$ est un système générateur de E de n vecteurs, il est minimal et c'est une base de E. □

Commentaire important.

Il suffit de l'**une des deux** propriétés « libre » ou « générateur » pour en déduire l'autre dès que le nombre de vecteurs du système étudié coïncide avec la dimension de E.

Corollaire 4.14 *Soit E un $\mathbb{K}$-espace vectoriel de dimension n.*

1. *Si $\mathrm{card}(\mathcal{S}) < n$, le système $\mathcal{S}$ n'est pas générateur de E.*
2. *Si $\mathrm{card}(\mathcal{S}) > n$, le système $\mathcal{S}$ est lié.*

Exemple.

Considérons dans le $\mathbb{K}$-espace vectoriel $\mathbb{K}_n[X]$ des polynômes $P_0, P_1 \ldots, P_n$ tels que $\forall i \in \{0, 1, \ldots, n\}$, $\deg(P_i) = i$. Le système $\mathcal{S} = \{P_0, P_1 \ldots, P_n\}$ est libre. En effet, si $\alpha_0, \alpha_1, \ldots, \alpha_n$ sont des scalaires tels que $\displaystyle\sum_{i=0}^{n} \alpha_i P_i = 0$, alors on trouve de proche en proche $\alpha_n = \alpha_{n-1} = \cdots = \alpha_1 = \alpha_0 = 0$. Par exemple, si on avait $\alpha_n \neq 0$, le polynôme $\displaystyle\sum_{i=0}^{n} \alpha_i P_i$ serait de degré n et ne pourrait être nul ; donc $\alpha_n = 0$ et le polynôme devient $\displaystyle\sum_{i=0}^{n-1} \alpha_i P_i$. On poursuit le raisonnement ainsi de suite. D'autre part, $\mathrm{card}(\mathcal{S}) = n + 1 = \dim(\mathbb{K}_n[X])$; le système $\mathcal{S}$ est donc une base de $\mathbb{K}_n[X]$.

Théorème 4.15 (de la base incomplète) *Soit E un $\mathbb{K}$-espace vectoriel de dimension n et $\mathcal{L}$ un système libre contenant p vecteurs. On peut compléter $\mathcal{L}$ en une base $\mathcal{B}$ de E : si $\mathcal{L} = \{\vec{u}_1, \ldots, \vec{u}_p\}$, il existe $n - p$ vecteurs $\vec{u}_{p+1}, \ldots, \vec{u}_n$ tels que le système $\mathcal{B} = \{\vec{u}_1, \ldots, \vec{u}_n\}$ soit une base de E.*

Démonstration.

On sait que E admet une base finie $\mathcal{B}'$. Il est possible d'en extraire $n-p$ vecteurs $\vec{u}_{p+1}, \ldots, \vec{u}_n$ tel que le système $\mathcal{L} \cup \{\vec{u}_{p+1}, \ldots, \vec{u}_n\}$ (contenant n vecteurs) soit une base de E. En effet,

- si $p = n$, le système $\mathcal{L}$ est une base de E, cette base est complète ;
- si $p < n$, le système $\mathcal{L}$ n'est pas générateur de E. Il existe donc un vecteur $\vec{u}_{p+1} \in \mathcal{B}'$ non combinaison linéaire des vecteurs de $\mathcal{L}$, sinon les n vecteurs de $\mathcal{B}'$ seraient des combinaisons linéaires des p vecteurs de $\mathcal{L}$, ce qui entraînerait que $\mathcal{B}'$ est un système lié en vertu du lemme 4.6. Le système $\mathcal{L}_1 = \mathcal{L} \cup \{\vec{u}_{p+1}\} = \{\vec{u}_1, \ldots, \vec{u}_{p+1}\}$ est libre ;
  - ∗ si $p + 1 = n$, le système $\mathcal{L}_1$ est une base de E et on a complété le système $\mathcal{L}$ en une base de E ;
  - ∗ si $p + 1 < n$, le système $\mathcal{L}_1$ n'est pas générateur de E. Comme précédemment, il existe un vecteur $\vec{u}_{p+2} \in \mathcal{B}'$ non combinaison linéaire des vecteurs de $\mathcal{L}_1$, et le système $\mathcal{L}_2 = \mathcal{L}_1 \cup \{\vec{u}_{p+2}\} = \{\vec{u}_1, \ldots, \vec{u}_{p+2}\}$ est libre ;
    - . si $p + 2 = n$, le système $\mathcal{L}_2$ est une base de E et on a complété le système $\mathcal{L}$ en une base de E ;
    - . sinon on réitère le procédé.

On construit ainsi de proche en proche une suite de systèmes libres $\mathcal{L} \subset \mathcal{L}_1 \subset \mathcal{L}_2 \subset \cdots \subset \mathcal{L}_{n-p}$ À chaque étape, on choisit un vecteur $\vec{u}_i \in \mathcal{B}'$ tel que le système $\mathcal{L}_i = \mathcal{L}_{i-1} \cup \{\vec{u}_i\}$ soit libre. Finalement, le système $\mathcal{B} = \mathcal{L}_{n-p} = \{\vec{u}_1, \ldots, \vec{u}_n\}$ est libre et comporte n vecteurs, c'est une base complétant le système libre $\mathcal{L}$. □

Corollaire 4.16 *(Existence de supplémentaires) Soit E un $\mathbb{K}$-espace vectoriel de dimension finie. Tout sous-espace vectoriel de E admet un supplémentaire dans E.*

DÉMONSTRATION.

Soit F un sous-espace vectoriel de E.

1. Vérifions d'abord que l'espace vectoriel F est de dimension finie. Un système libre de vecteurs de F est aussi un système libre de vecteurs de E, il a donc un cardinal majoré par $\dim(E)$. Le sous-espace vectoriel F admet alors au moins un système libre maximal de cardinal inférieur ou égal à $\dim(E)$, c'est une base de F.
2. Soit $\mathcal{L}$ une base de E. C'est un système libre de vecteurs de E que l'on peut compléter en une base $\mathcal{B} = \mathcal{L} \cup \mathcal{L}'$ de E. Le système $\mathcal{L}'$ étant contenu dans une base est nécessairement libre. Soit $G = \mathrm{vect}(\mathcal{L}')$. Montrons que F et G sont des sous-espaces vectoriels supplémentaires dans E.

   (a). Posons $\mathcal{L} = \{\vec{u}_1, \dots, \vec{u}_p\}$ et $\mathcal{L}' = \{\vec{u}_{p+1}, \dots, \vec{u}_n\}$. Soit $\vec{w} \in E$. Décomposons $\vec{w}$ relativement à la base $\mathcal{B}$:

   $$\vec{w} = (\alpha_1\vec{u}_1 + \cdots + \alpha_p\vec{u}_p) + (\alpha_{p+1}\vec{u}_{p+1} + \cdots + \alpha_n\vec{u}_n) = \vec{u} + \vec{v}$$

   avec

   $$\vec{u} = \alpha_1\vec{u}_1 + \cdots + \alpha_p\vec{u}_p \in F \quad \text{et} \quad \vec{w} = \alpha_{p+1}\vec{u}_{p+1} + \cdots + \alpha_n\vec{u}_n \in G.$$

   Ceci montre que $F + G = E$.

   (b). Soit maintenant $\vec{w} \in F \cap G$. Le vecteur $\vec{w}$ admet deux décompositions

   $$\vec{w} = \alpha_1\vec{u}_1 + \cdots + \alpha_p\vec{u}_p = \alpha_{p+1}\vec{u}_{p+1} + \cdots + \alpha_n\vec{u}_n,$$

   ce qui entraîne

   $$\alpha_1\vec{u}_1 + \cdots + \alpha_p\vec{u}_p - \alpha_{p+1}\vec{u}_{p+1} - \cdots - \alpha_n\vec{u}_n = \vec{0},$$

   soit encore, puisque $\mathcal{B}$ est une base, $\alpha_1 = \cdots = \alpha_p = \alpha_{p+1} = \cdots = \alpha_n = 0$. Donc $\vec{w} = \vec{0}$ puis $F \cap G = \{\vec{0}\}$.

On a finalement $F \oplus G = E$. □

Proposition 4.17 *Soit F et G des sous-espaces vectoriels de E tels que la somme $F + G$ soit directe, et $\mathcal{B}_F$ et $\mathcal{B}_G$ des bases de F et G respectivement. Alors $\mathcal{B}_F \cup \mathcal{B}_G$ est une base de $F \oplus G$.*

DÉMONSTRATION.

Rappelons tout d'abord que, dans le cas général (la somme $F+G$ n'étant pas nécessairement directe), $\mathcal{B}_F \cup \mathcal{B}_G$ est un système générateur de $F + G$ (voir la proposition 4.5). D'autre part, si la somme $F + G$ étant directe, on a $F \cap G = \{\vec{0}\}$, donc $\mathcal{B}_F \cup \mathcal{B}_G$ est un système libre. C'est bien une base de $F \oplus G$. □

4.4 Dimension d'un sous-espace vectoriel

Rappelons, comme cela a été vu au cours de la démonstration du corollaire 4.16, que tout sous-espace vectoriel d'un espace vectoriel de dimension finie est un espace vectoriel de dimension finie.

Théorème 4.18 *Soit E un $\mathbb{K}$-espace vectoriel de dimension finie et F et G deux sous-espaces vectoriels de E.*

1. *(a) Si $F \subset G$, alors $\dim(F) \leqslant \dim(G)$.*
   (b) Si $F \subset G$ et $\dim(F) = \dim(G)$, alors $F = G$.
2. *On a (formule de Grassmann[1])*

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

En particulier, si la somme $F+G$ est directe,

$$\dim(F \oplus G) = \dim(F) + \dim(G).$$

3. *Les sous-espaces vectoriels F et G sont supplémentaires dans E si et seulement si*

$$F \cap G = \{\vec{0}\} \quad \text{et} \quad \dim(F) + \dim(G) = \dim(E),$$

ou encore, si et seulement si

$$F + G = E \quad \text{et} \quad \dim(F) + \dim(G) = \dim(E).$$

Démonstration.

1. Supposons $F \subset G$. Soit $\mathcal{B}$ une base de F. C'est en particulier un système libre de vecteurs de G et donc $\dim(F) = \operatorname{card}(\mathcal{B}) \leqslant \dim(G)$. Si on a de plus $\dim(F) = \dim(G)$, alors $\operatorname{card}(\mathcal{B}) = \dim(G)$ et le système libre $\mathcal{B}$ est une base de G. D'où $F = G$.
2. Posons $p = \dim(F)$, $q = \dim(G)$ et $r = \dim(F \cap G)$. Soit $\{\vec{u}_1, \dots, \vec{u}_r\}$ une base de $F \cap G$. On utilise le théorème 4.15 :
   — c'est un système libre de F, on le complète en une base $\{\vec{u}_1, \dots, \vec{u}_r, \vec{u}\,'_{r+1}, \dots, \vec{u}\,'_p\}$ de F ;
   — c'est un système libre de G, on le complète en une base $\{\vec{u}_1, \dots, \vec{u}_r, \vec{u}\,''_{p+1}, \dots, \vec{u}\,''_{p+q-r}\}$ de G.

   Considérons alors le système $\mathcal{B} = \{\vec{u}_1, \dots, \vec{u}_r, \vec{u}\,'_{r+1}, \dots, \vec{u}\,'_p, \vec{u}\,''_{p+1}, \dots, \vec{u}\,''_{p+q-r}\}$. Montrons que c'est une base de $F+G$.

   (a). C'est un système générateur de $F+G$. En effet, soit $\vec{w} \in F+G$. Il existe $(\vec{u}, \vec{v}) \in F \times G$ tel que $\vec{w} = \vec{u} + \vec{v}$. Le vecteur $\vec{u}$ est de la forme $\displaystyle\sum_{i=1}^{r} \alpha_i \vec{u}_i + \sum_{i=r+1}^{p} \alpha'_i \vec{u}\,'_i$ et le vecteur $\vec{v}$ est de la forme $\displaystyle\sum_{i=1}^{r} \beta_i \vec{u}_i + \sum_{i=p+1}^{p+q-r} \beta''_i \vec{u}\,''_i$. Donc

$$\vec{w} = \sum_{i=1}^{r} (\alpha_i + \beta_i) \vec{u}_i + \sum_{i=r+1}^{p} \alpha'_i \vec{u}\,'_i + \sum_{i=p+1}^{p+q-r} \beta''_i \vec{u}\,''_i$$

   et $\vec{w}$ est bien combinaison linéaire des vecteurs de $\mathcal{B}$.

1. Grassmann, Hermann : mathématicien polonais (Szczecin 1809–Szczecin 1877)

(b). C'est un système libre. En effet, soient $\alpha_1,\dots,\alpha_r,\alpha'_{r+1},\dots,\alpha'_p,\alpha''_{p+1},\dots,\alpha''_{p+q-r}$ des scalaires tels que

$$\sum_{i=1}^{r}\alpha_i\vec{u}_i+\sum_{i=r+1}^{p}\alpha'_i\vec{u}'_i+\sum_{i=p+1}^{p+q-r}\alpha''_i\vec{u}''_i=\vec{0}.$$

On a

$$\underbrace{\sum_{i=p+1}^{p+q-r}\alpha''_i\vec{u}''_i}_{\in G}=\underbrace{-\sum_{i=1}^{r}\alpha_i\vec{u}_i-\sum_{i=r+1}^{p}\alpha'_i\vec{u}'_i}_{\in F}\in F\cap G,$$

il existe donc des scalaires $\beta_1,\dots,\beta_r$ tels que

$$\sum_{i=p+1}^{p+q-r}\alpha''_i\vec{u}''_i=\sum_{i=1}^{r}\beta_i\vec{u}_i.$$

Le système $\{\vec{u}_1,\dots,\vec{u}_r,\vec{u}''_{p+1},\dots,\vec{u}''_{p+q-r}\}$ étant libre (c'est une base de G), les coefficients $\alpha''_{p+1},\dots,\alpha''_{p+q-r},\beta_1,\dots,\beta_r$ sont tous nuls. On en tire dès lors

$$\sum_{i=1}^{r}\alpha_i\vec{u}_i+\sum_{i=r+1}^{p}\alpha'_i\vec{u}'_i=\vec{0}$$

et comme le système $\{\vec{u}_1,\dots,\vec{u}_r,\vec{u}'_{r+1},\dots,\vec{u}'_p\}$ est libre (c'est une base de F), les coefficients $\alpha_1,\dots,\alpha_r,\alpha'_{r+1},\dots,\alpha'_p$ sont tous nuls.

On a ainsi prouvé que le système $\mathcal{B}$ est une base de $F+G$. En conclusion,

$$\dim(F+G)=\mathrm{card}(\mathcal{B})=p+q-r=\dim(F)+\dim(G)-\dim(F\cap G).$$

3. Le troisième point est une conséquence du précédent. Supposons que $\dim(F)+\dim(G)=\dim(E)$.

(a). Si $F\cap G=\{\vec{0}\}$, on a $\dim(F+G)=\dim(F)+\dim(G)=\dim(E)$. D'après le premier point, le seul sous-espace vectoriel de E de dimension $\dim(E)$ est E lui-même, donc $F+G=E$.

(b). Si $F+G=E$, on a $\dim(F\cap G)=\dim(E)-\dim(F)-\dim(G)=0$ et donc $F\cap G=\{\vec{0}\}$.

Dans chacun des deux cas ci-dessus, les sous-espaces vectoriels F et G sont supplémentaires dans E. □

Remarque.

La formule de Grassmann fournit des encadrements pour les dimensions des sous-espaces vectoriels $F \cap G$ et $F + G$:

— puisque $F \cap G$ est contenu à la fois dans F et dans G, on a

$$0 \leqslant \dim(F \cap G) \leqslant \min(\dim(F), \dim(G)) \ ;$$

— de même, $F + G$ étant contenu dans E et contenant F et G, on a

$$\max(\dim(F), \dim(G)) \leqslant \dim(F + G) \leqslant \dim(E).$$

Exemples.

1. *Droites et hyperplans :* considérons dans le $\mathbb{K}$-espace vectoriel $E = \mathbb{K}^n$ les deux sous-espaces vectoriels

$$F = \{(\lambda a_1, \dots, \lambda a_n),\ \lambda \in \mathbb{K}\} \quad \text{et} \quad G = \{(x_1, \dots, x_n) \in E : b_1 x_1 + \dots + b_n x_n = 0\}$$

où les scalaires $a_1, \dots, a_n$ sont non tous nuls ainsi que les scalaires $b_1, \dots, b_n$. On a vu dans le paragraphe 3.4 que $F \cap G = \{\vec{0}\}$ si et seulement si $a_1 b_1 + \dots + a_n b_n \neq 0$. D'autre part, $F = \mathrm{vect}(\{(a_1, \dots, a_n)\})$ donc $\dim(F) = 1$. En ce qui concerne G, en supposant par exemple $\beta_n \neq 0$,

$$(x_1, \dots, x_n) \in G \iff x_n = -\frac{b_1}{b_n}\, x_1 - \dots - \frac{b_{n-1}}{b_n}\, x_{n-1}$$

et alors

$$(x_1, x_2, \dots, x_n) = x_1(1, 0, \dots, 0, -\frac{b_1}{b_n}) + x_2(0, 1, \dots, 0, -\frac{b_2}{b_n}) + \dots + x_{n-1}(0, 0, \dots, 1, -\frac{b_{n-1}}{b_n}).$$

On a ainsi obtenu un système générateur de G :

$$\{(b_n, 0, \dots, 0, -b_1), (0, b_n, \dots, 0, -b_2), \dots, (0, 0, \dots, b_n, -b_{n-1})\}.$$

Il est immédiat de voir qu'il est libre, c'est donc une base de G et $\dim(G) = n-1$. Les sous-espaces vectoriels de dimension $n - 1$ d'un espace vectoriel de dimension n sont appelés *hyperplans vectoriels*. Ils sont caractérisés par une *équation cartésienne* ; ici l'hyperplan G est caractérisé par l'équation cartésienne $b_1 x_1 + \dots + b_n x_n = 0$. On a $\dim(F) + \dim(G) = n = \dim(E)$ est par conséquent, les sous-espaces vectoriels F et G sont supplémentaires dans E si et seulement si $a_1 b_1 + \dots + a_n b_n \neq 0$. Il convient de noter qu'il serait assez compliqué de vérifier directement l'égalité $F + G = E$. Ici, on a utilisé la théorie de la dimension pour éluder cette difficulté.

2. *Étude détaillée d'un exemple dans $\mathbb{R}^4$:* dans le $\mathbb{R}$-espace vectoriel $E = \mathbb{R}^4$, on considère les deux sous-espaces vectoriels

$$F = \{(\lambda + 2\mu, 2\lambda + \mu, 3\mu, \lambda + \mu),\ \lambda, \mu \in \mathbb{R}\} \quad \text{et} \quad G = \{(x, y, z, t) \in E : y - 2z + t = 0\}.$$

On voudrait déterminer les sous-espaces vectoriels $F \cap G$ et $F + G$. Pour cela, on commence par chercher des bases de F et G. Soit $\vec{u} = (x, y, z, t)$ un vecteur générique de E.

— Pour F, on a

$$\vec{u} \in F \iff \exists \lambda, \mu \in \mathbb{R},\ \vec{u} = (\lambda + 2\mu, 2\lambda + \mu, 3\mu, \lambda + \mu) \iff \exists \lambda, \mu \in \mathbb{R}, \left\{ \begin{array}{rcr} x &=& \lambda + 2\mu \\ y &=& 2\lambda + \mu \\ z &=& 3\mu \\ t &=& \lambda + \mu \end{array} \right..$$

Le système d'équations précédent est une *représentation paramétrique* de F, de laquelle on déduit

$$\begin{aligned} F &= \{(\lambda+2\mu, 2\lambda+\mu, 3\mu, \lambda+\mu),\ \lambda,\mu\in\mathbb{R}\} \\ &= \{\lambda(1,2,0,1)+\mu(2,1,3,1),\ \lambda,\mu\in\mathbb{R}\} \\ &= \mathrm{vect}(\{(1,2,0,1),(2,1,3,1)\}). \end{aligned}$$

On obtient ainsi un système générateur de F : $\mathcal{B}_F = \{(1,2,0,1),(2,1,3,1)\}$ qui est clairement une base de F puisque les deux vecteurs $(1,2,0,1)$ et $(2,1,3,1)$ ne sont pas colinéaires. On a donc $\dim(F)=2$ et F est un plan vectoriel.

— De manière analogue, partant de l'équation $y-2z+t=0$ appelée *représentation cartésienne* de G, on trouve

$$\vec{u}\in G \Longleftrightarrow y-2z+t=0 \Longleftrightarrow t=-y+2z \Longleftrightarrow \vec{u}=(x,y,z,-y+2z),$$

d'où l'on tire

$$\begin{aligned} G &= \{(\lambda,\mu,\nu,-\mu+2\nu),\ \lambda,\mu,\nu\in\mathbb{R}\} \\ &= \{\lambda(1,0,0,0)+\mu(0,1,0,-1)+\nu(0,0,1,2),\ \lambda,\mu,\nu\in\mathbb{R}\} \\ &= \mathrm{vect}(\{(1,0,0,0),(0,1,0,-1),(0,0,1,2)\}). \end{aligned}$$

On obtient ainsi un système générateur de G : $\mathcal{B}_G = \{(1,0,0,0),(0,1,0,-1),(0,0,1,2)\}$ qui s'avère être une base de G. Vérifions que $\mathcal{B}_G$ est effectivement un système libre. Soit $\alpha,\beta,\gamma\in\mathbb{R}$ des scalaires tels que $\alpha(1,0,0,0)+\beta(0,1,0,-1)+\gamma(0,0,1,2)=\vec{0}$. On a alors $(\alpha,\beta,\gamma,-\beta+2\gamma)=(0,0,0,0)$ et nécessairement $\alpha=\beta=\gamma=0$. Les vecteurs $(1,0,0,0),(0,1,0,-1),(0,0,1,2)$ ne sont pas coplanaires, $\dim(G)=3$ et G est un hyperplan vectoriel.

— Pour obtenir $F\cap G$, on utilise la représentation paramétrique de F ainsi que la représentation cartésienne de G :

$$\vec{u}\in F\cap G \Longleftrightarrow \exists\lambda,\mu\in\mathbb{R},\ \left\{\begin{array}{rcl} x &=& \lambda+2\mu \\ y &=& 2\lambda+\mu \\ z &=& 3\mu \\ t &=& \lambda+\mu \end{array}\right. \quad \text{et} \quad y-2z+t=0.$$

Injectant les équations paramétriques (contenant les paramètres λ,μ) dans l'équation cartésienne $y-2z+t=0$, on obtient la relation $3\lambda-4\mu=0$, ou encore $\mu=\dfrac{3}{4}\lambda$, puis en reportant cette dernière dans les équations paramétriques, on trouve, après multiplication par 4 ($\lambda=4\nu,\mu=3\nu$),

$$\vec{u}\in F\cap G \Longleftrightarrow \exists\nu\in\mathbb{R},\ \left\{\begin{array}{rcl} x &=& 10\nu \\ y &=& 11\nu \\ z &=& 9\nu \\ t &=& 7\nu \end{array}\right. .$$

On voit donc que

$$F\cap G = \{(10\nu,11\nu,9\nu,7\nu),\ \nu\in\mathbb{R}\} = \mathrm{vect}(\{(10,11,9,7)\})$$

et alors $F\cap G$ est la droite vectorielle engendrée par le vecteur $(10,11,9,7)$.

— Pour obtenir $F+G$, la formule de Grassmann donne une première information :

$$\dim(F+G) = \dim(F) + \dim(G) - \dim(F \cap G) = 3 + 2 - 1 = 4.$$

En fait, cette information est fort précieuse puisque, du fait que le seul sous-espace vectoriel de E de dimension 4 est E lui-même, elle donne immédiatement $F+G=E$. On aurait pu aussi utiliser les bases $\mathcal{B}_F$ et $\mathcal{B}_G$ pour obtenir un système générateur de $F+G$ (voir la proposition 4.5) :

$$\mathcal{B}_F \cup \mathcal{B}_G = \{(1,2,0,1),(2,1,3,1),(1,0,0,0),(0,1,0,-1),(0,0,1,2)\}.$$

Comme $\mathrm{card}(\mathcal{B}_F \cup \mathcal{B}_G) = 5 > 4 = \dim(E)$, il est clair que ce système est lié. Notons que, en utilisant l'information que $(10,11,9,7) \in F \cap G$, on a des combinaisons linéaires de la forme

$$(10,11,9,7) = \lambda(1,2,0,1) + \mu(2,1,3,1) = \alpha(1,0,0,0) + \beta(0,1,0,-1) + \gamma(0,0,1,2)\,;$$

les coefficients sont donnés par $\lambda = 4$, $\mu = 3$, $\alpha = 10$, $\beta = 11$, $\gamma = 9$. On voit donc que $(1,2,0,1)$ par exemple est combinaison linéaire des autres vecteurs de $\mathcal{B}_F \cup \mathcal{B}_G$ et alors $(\mathcal{B}_F \cup \mathcal{B}_G) \setminus \{(1,2,0,1)\} = \{(2,1,3,1),(1,0,0,0),(0,1,0,-1),(0,0,1,2)\}$ est encore un système générateur de $F+G$. On peut vérifier qu'il est libre et que c'est par conséquent une base de $F+G$.

4.5 Représentations paramétrique et cartésienne d'un sous-espace vectoriel

Soit E un $\mathbb{K}$-espace vectoriel de dimension n, $\mathcal{B} = \{\vec{e}_1, \dots, \vec{e}_n\}$ une base de E et $\mathcal{S} = \{\vec{u}_1, \dots, \vec{u}_p\}$ un système libre de E. Introduisons les coordonnées des vecteurs $\vec{u}_1, \dots, \vec{u}_p$ relativement à la base $\mathcal{B}$: $\vec{u}_j = \sum_{i=1}^{n} \alpha_{ij} \vec{e}_i$, $1 \leqslant j \leqslant p$. Soit enfin F le sous-espace vectoriel de E engendré par $\mathcal{S}$: $F = \mathrm{vect}(\mathcal{S})$. Pour un vecteur générique $\vec{u} = \sum_{i=1}^{n} x_i \vec{e}_i$ de E, on a

$$\vec{u} \in F \iff \exists \lambda_1, \dots, \lambda_p \in \mathbb{K},\ \vec{u} = \sum_{j=1}^{p} \lambda_j \vec{u}_j \iff \exists \lambda_1, \dots, \lambda_p \in \mathbb{K},\ \forall i \in \{1, \dots, n\},\ x_i = \sum_{j=1}^{p} \alpha_{ij} \lambda_j.$$

Le système d'équations

$$\left\{ \begin{array}{rcl} x_1 & = & \alpha_{11}\,\lambda_1 + \alpha_{12}\,\lambda_2 + \cdots + \alpha_{1p}\,\lambda_p \\ x_2 & = & \alpha_{21}\,\lambda_1 + \alpha_{22}\,\lambda_2 + \cdots + \alpha_{2p}\,\lambda_p \\ & \vdots & \\ x_n & = & \alpha_{n1}\,\lambda_1 + \alpha_{n2}\,\lambda_2 + \cdots + \alpha_{np}\,\lambda_p \end{array} \right.$$

est une *représentation paramétrique* du sous-espace vectoriel F (qui est de dimension p). Elle contient p paramètres. En éliminant les p paramètres $\lambda_1, \dots, \lambda_p$, il est possible d'obtenir un système de $n-p$ équations reliant les coordonnées $x_1, \dots, x_n$:

$$\left\{ \begin{array}{rcl} \beta_{11}\,x_1 + \quad \beta_{12}\,x_2 + \cdots + \quad \beta_{1n}\,x_n & = & 0 \\ \beta_{21}\,x_1 + \quad \beta_{22}\,x_2 + \cdots + \quad \beta_{2n}\,x_n & = & 0 \\ & \vdots & \\ \beta_{n-p\,1}\,x_1 + \beta_{n-p\,2}\,x_2 + \cdots + \beta_{n-p\,n}\,x_n & = & 0 \end{array} \right.$$

Ce système est une *représentation cartésienne* du sous-espace vectoriel F.

Exemple.

Droites et hyperplans vectoriels : le discours précédent nous enseigne que
- pour une droite vectorielle de E, correspondant au cas $p = 1$, les représentations cartésiennes contiennent $n-1$ équations ;
- pour un hyperplan vectoriel de E, correspondant au cas $p = n-1$, les représentations cartésiennes contiennent 1 équation ; il est donc caractérisé par une équation cartésienne.

Considérons par exemple dans le $\mathbb{R}$-espace vectoriel $E = \mathbb{R}^3$ la droite vectorielle D engendrée par le vecteur $\vec{u} = (1, 3, -2)$: $D = \mathrm{vect}(\{\vec{u}\})$, et le plan vectoriel P engendré par les vecteurs $\vec{u}_1 = (1, 1, -1)$ et $\vec{u}_2 = (2, 5, -1)$: $P = \mathrm{vect}(\{\vec{u}_1, \vec{u}_2\})$.
- On a $D = \{\lambda(1, 3, -2),\ \lambda \in \mathbb{R}\} = \{(\lambda, 3\lambda, -2\lambda),\ \lambda \in \mathbb{R}\}$. Un vecteur générique $\vec{v} = (x, y, z)$ de E appartient à D si et seulement si il existe un scalaire $\lambda \in \mathbb{R}$ tel que $(x, y, z) = (\lambda, 3\lambda, -2\lambda)$. Ceci conduit à la représentation paramétrique de la droite D suivante :
$$\left\{\begin{array}{rcr} x & = & \lambda \\ y & = & 3\lambda \\ z & = & -2\lambda \end{array}\right.,\ \lambda \in \mathbb{R}.$$
On en obtient une représentation cartésienne en éliminant le paramètre λ : de la première équation, on tire $\lambda = x$ que l'on injecte dans les deux autres équations pour arriver à
$$\left\{\begin{array}{rcl} 3x - y & = & 0 \\ 2x + z & = & 0 \end{array}\right..$$
- De même, en écrivant
$$P = \{\lambda(1, 1, -1) + \mu(2, 5, -1),\ \lambda, \mu \in \mathbb{R}\} = \{(\lambda + 2\mu, \lambda + 5\mu, -\lambda - \mu),\ \lambda, \mu \in \mathbb{R}\},$$
on obtient une représentation paramétrique de P :
$$\left\{\begin{array}{rcl} x & = & \lambda + 2\mu \\ y & = & \lambda + 5\mu \\ z & = & -\lambda - \mu \end{array}\right.,\ \lambda, \mu \in \mathbb{R},$$
de laquelle on déduit une représentation cartésienne après élimination des paramètres λ et μ. En effet, des première et troisième équations, on tire
$$\left\{\begin{array}{rcl} \mu & = & x + z \\ \lambda & = & -x - 2z \end{array}\right.$$
que l'on injecte dans la deuxième équation pour obtenir la représentation cartésienne de P suivante :
$$4x - y + 3z = 0.$$

4.6 Rang d'un système de vecteurs

Définition 4.19 *Soit $\mathcal{S}$ un système de vecteurs de E. Le rang du système $\mathcal{S}$ est la dimension du sous-espace vectoriel de E engendré par $\mathcal{S}$:*
$$\mathrm{rang}(\mathcal{S}) = \dim(\mathrm{vect}(\mathcal{S})).$$

Propriétés immédiates.

— Si le système $\mathcal{S}$ contient le vecteur nul, $\operatorname{rang}(\mathcal{S}) = \operatorname{rang}(\mathcal{S} \setminus \{\vec{0}\})$.

— Le rang de $\mathcal{S}$ ne dépend pas de l'ordre dans lequel sont écrits les vecteurs de $\mathcal{S}$.

— Le système $\mathcal{S}$ étant générateur de $\operatorname{vect}(\mathcal{S})$, on a $\operatorname{card}(\mathcal{S}) \geqslant \dim(\operatorname{vect}(\mathcal{S}))$. D'autre part, $\operatorname{vect}(\mathcal{S})$ étant un sous-espace vectoriel de E, on a $\dim(\operatorname{vect}(\mathcal{S})) \leqslant \dim(E)$. En conséquence, on a l'inégalité
$$\operatorname{rang}(\mathcal{S}) \leqslant \min(\operatorname{card}(\mathcal{S}), \dim(E)).$$

Proposition 4.20 *Soit $\mathcal{S}$ un système de vecteurs de E. On a*

— $\operatorname{rang}(\mathcal{S}) = \operatorname{card}(\mathcal{S})$ *si et seulement si le système $\mathcal{S}$ est libre ;*

— $\operatorname{rang}(\mathcal{S}) = \dim(E)$ *si et seulement si le système $\mathcal{S}$ est générateur de E.*

Démonstration.

Pour la première assertion, il suffit d'observer que $\operatorname{rang}(\mathcal{S}) = \operatorname{card}(\mathcal{S})$ si et seulement si le système $\mathcal{S}$ est une base de $\operatorname{vect}(\mathcal{S})$. La deuxième est évidente. □

Proposition 4.21 *Soit $\alpha_1, \ldots, \alpha_p \in \mathbb{K}$ avec $\alpha_i \neq 0$, $\vec{u}_1, \ldots, \vec{u}_p \in E$ et $\vec{u}_i' = \sum_{j=1}^{p} \alpha_j \vec{u}_j$ pour un i fixé. Le rang du système de vecteurs $\{\vec{u}_1, \ldots, \vec{u}_p\}$ est inchangé lorsqu'on remplace $\vec{u}_i$ par $\vec{u}_i'$:*
$$\operatorname{rang}(\{\vec{u}_1, \ldots, \vec{u}_i, \ldots, \vec{u}_p\}) = \operatorname{rang}(\{\vec{u}_1, \ldots, \vec{u}_i', \ldots, \vec{u}_p\}).$$

Démonstration.

Posons $\mathcal{S} = \{\vec{u}_1, \ldots, \vec{u}_i, \ldots, \vec{u}_p\}$ et $\mathcal{S}' = \{\vec{u}_1, \ldots, \vec{u}_i', \ldots, \vec{u}_p\}$. On a $\vec{u}_i' \in \operatorname{vect}(\mathcal{S}')$ et donc $\operatorname{vect}(\mathcal{S}') \subset \operatorname{vect}(\mathcal{S})$. Réciproquement, puisque $\alpha_i \neq 0$,
$$\vec{u}_i = \sum_{\substack{j=1 \\ j \neq i}}^{p} -\frac{\alpha_j}{\alpha_i}\, \vec{u}_j + \frac{1}{\alpha_i}\, \vec{u}_i' \in \operatorname{vect}(\mathcal{S}')$$
et alors $\operatorname{vect}(\mathcal{S}) \subset \operatorname{vect}(\mathcal{S}')$. On en tire $\operatorname{vect}(\mathcal{S}) = \operatorname{vect}(\mathcal{S}')$ puis $\operatorname{rang}(\mathcal{S}) = \operatorname{rang}(\mathcal{S}')$. □

Proposition 4.22 *Soit $\{\vec{e}_1, \ldots, \vec{e}_n\}$ une base de E, $p \leqslant n$ et $(\alpha_{ij})_{\substack{1 \leqslant i \leqslant n \\ 1 \leqslant j \leqslant p}}$ un tableau de scalaires tels que*
$$\forall j \in \{1, \ldots, p\},\ \alpha_{jj} \neq 0 \quad \textit{et} \quad \forall i \in \{1, \ldots, n\}, \forall j \in \{1, \ldots, p\},\ i < j \Longrightarrow \alpha_{ij} = 0.$$
On pose $\vec{u}_j = \sum_{i=j}^{n} \alpha_{ij} \vec{e}_i$ pour $1 \leqslant j \leqslant p$. Alors, le système de vecteurs $\{\vec{u}_1, \ldots, \vec{u}_p\}$ est libre, donc de rang p.

Il est commode de disposer les coefficients α_{ij} sous forme d'un tableau dont les colonnes représentent les vecteurs $\{\vec{u}_1, \ldots, \vec{u}_p\}$ relativement à la base $\{\vec{e}_1, \ldots, \vec{e}_n\}$. La condition $\alpha_{ij} = 0$ pour $i < j$ donne un tableau de la forme

$$\begin{array}{c c}
\begin{array}{ccccc} \vec{u}_1 & \vec{u}_2 & \vec{u}_3 & \ldots & \vec{u}_p \end{array} & \\
\begin{pmatrix}
\alpha_{11} & 0 & 0 & & 0 \\
\alpha_{21} & \alpha_{22} & 0 & & 0 \\
\alpha_{31} & \alpha_{32} & \alpha_{33} & & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
 & & & & \alpha_{pp} \\
\vdots & \vdots & \vdots & & \vdots \\
\alpha_{n1} & \alpha_{n2} & \alpha_{n3} & & \alpha_{np}
\end{pmatrix} & \begin{array}{c} \vec{e}_1 \\ \vec{e}_2 \\ \vec{e}_3 \\ \vdots \\ \vec{e}_p \\ \vdots \\ \vec{e}_n \end{array}
\end{array}$$

et la condition $\alpha_{jj} \neq 0$ pour tout j signifie que les termes diagonaux du tableau ci-dessus sont tous non nuls.

Démonstration.

Soit $\lambda_1, \ldots, \lambda_p \in \mathbb{K}$ tels que $\sum_{j=1}^{p} \lambda_j \vec{u}_j = \vec{0}$. Comme

$$\sum_{j=1}^{p} \lambda_j \vec{u}_j = \sum_{j=1}^{p} \lambda_j \left(\sum_{i=j}^{n} \alpha_{ij} \vec{e}_i \right) = \sum_{i=1}^{n} \left(\sum_{j=1}^{\min(i,p)} \alpha_{ij} \lambda_j \right) \vec{e}_i,$$

on est conduit au système triangulaire d'équations $\sum_{j=1}^{i} \alpha_{ij} \lambda_j = 0$, $1 \leqslant i \leqslant p$, soit plus précisément

$$\left\{ \begin{array}{ll}
\alpha_{11}\lambda_1 & = 0 \\
\alpha_{21}\lambda_1 + \alpha_{22}\lambda_2 & = 0 \\
 & \vdots \\
\alpha_{p1}\lambda_1 + \alpha_{p2}\lambda_2 + \cdots + \alpha_{pp}\lambda_p & = 0
\end{array} \right.$$

duquel on tire successivement, puisque les coefficients $\alpha_{11}, \ldots, \alpha_{pp}$ sont non nuls, $\lambda_1 = \cdots = \lambda_p = 0$. Le système $\mathcal{S}$ est donc libre. □

Détermination pratique du rang : méthode des zéros échelonnés

Soit $\{\vec{e}_1, \ldots, \vec{e}_n\}$ une base de E, $p \in \mathbb{N}^*$ et $\mathcal{S} = \{\vec{u}_1, \ldots, \vec{u}_p\}$ un système de vecteurs de E donnés par $\vec{u}_j = \sum_{i=1}^{n} \alpha_{ij} \vec{e}_i$ pour $1 \leqslant j \leqslant p$. Le principe de la méthode des zéros échelonnés consiste à remplacer les vecteurs $\vec{u}_j$ par des combinaisons linéaires de façon à obtenir un système de vecteurs $\mathcal{S}' = \{\vec{u}'_1, \dots, \vec{u}'_p\}$ associé à un tableau de coordonnées de la forme

$$\begin{array}{c} \begin{array}{cccccccc} \vec{u}'_1 & \vec{u}'_2 & \vec{u}'_3 & \dots & \vec{u}'_r & \vec{u}'_{r+1} & \dots & \vec{u}'_p \end{array} \\ \begin{pmatrix} \alpha'_{11} & 0 & 0 & & 0 & 0 & & 0 \\ \alpha'_{21} & \alpha'_{22} & 0 & & 0 & 0 & & 0 \\ \alpha'_{31} & \alpha'_{32} & \alpha'_{33} & & 0 & 0 & & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & & \vdots \\ & & & & \alpha'_{rr} & 0 & & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots & & \vdots \\ \alpha'_{n1} & \alpha'_{n2} & \alpha'_{n3} & & \alpha'_{nr} & 0 & & 0 \end{pmatrix} \begin{array}{c} \vec{e}_1 \\ \vec{e}_2 \\ \vec{e}_3 \\ \vdots \\ \vec{e}_r \\ \vdots \\ \vec{e}_n \end{array} \end{array}$$

où les scalaires $\alpha'_{11}, \dots, \alpha'_{rr}$ sont non nuls. On obtient ainsi, en vertu des propositions 4.21 et 4.22, une base $\{\vec{u}'_1, \dots, \vec{u}'_r\}$ de $\mathrm{vect}(\mathcal{S})$ et on a $\mathrm{rang}(\mathcal{S}) = r$. De plus, lorsque $p > r$, les $p - r$ vecteurs nuls $\vec{u}'_{r+1}, \dots, \vec{u}'_p$ fournissent $p - r$ combinaisons linéaires non triviales liant les vecteurs $\vec{u}_1, \dots, \vec{u}_p$.

Indiquons les premières étapes de la méthode.

1. Supposons $\alpha_{11} \neq 0$ et posons $\vec{u}'_1 = \vec{u}_1$. On cherche un vecteur $\vec{u}'_2$ sous la forme $\vec{u}'_2 = \vec{u}_2 - \lambda \vec{u}_1$ de telle sorte que la coordonnée sur $\vec{e}_1$ soit nulle. Cela donne $\lambda = \alpha_{12}/\alpha_{11}$, soit
$$\vec{u}'_2 = \vec{u}_2 - \frac{\alpha_{12}}{\alpha_{11}}\,\vec{u}_1,$$
donc $\vec{u}'_2$ est de la forme $\vec{u}'_2 = \sum_{i=2}^{n} \alpha'_{i2}\vec{e}_i$ et l'on a $\mathrm{vect}(\{\vec{u}_1, \vec{u}_2\}) = \mathrm{vect}(\{\vec{u}'_1, \vec{u}'_2\})$.
Si $\alpha_{11} = 0$, on permute les vecteurs $\vec{u}_1, \dots, \vec{u}_p$ (colonnes) ou $\vec{e}_1, \dots, \vec{e}_n$ (lignes) de manière à placer un terme non nul à l'intersection des première ligne et première colonne.
2. Supposons ensuite $\alpha'_{22} \neq 0$. On cherche $\vec{u}'_3 = \vec{u}_3 - \mu \vec{u}'_1 - \nu \vec{u}'_2$ avec des scalaires μ et ν de manière à ce que cette combinaison linéaire fournisse un vecteur dont les coordonnées sur $\vec{e}_1$ et $\vec{e}_2$ soient nulles. On trouve
$$\vec{u}'_3 = \vec{u}_3 - \frac{\alpha_{13}}{\alpha'_{11}}\,\vec{u}'_1 - \frac{\alpha'_{11}\alpha_{23} - \alpha'_{21}\alpha_{13}}{\alpha'_{11}\alpha'_{22}}\,\vec{u}'_2,$$
donc $\vec{u}'_2$ est de la forme $\vec{u}'_2 = \sum_{i=3}^{n} \alpha'_{i3}\vec{e}_i$ et l'on a $\mathrm{vect}(\{\vec{u}_1, \vec{u}_2, \vec{u}_3\}) = \mathrm{vect}(\{\vec{u}'_1, \vec{u}'_2, \vec{u}'_3\})$.
Si $\alpha'_{22} = 0$, on permute les vecteurs $\vec{u}'_2, \dots, \vec{u}'_p$ (colonnes) ou $\vec{e}_2, \dots, \vec{e}_n$ (lignes) de manière à placer un terme non nul à l'intersection des deuxième ligne et deuxième colonne. Et ainsi de suite.

Exemple.

Dans le $\mathbb{K}$-espace vectoriel $E = \mathbb{K}^4$ rapporté à sa base canonique $\{\vec{e}_1, \vec{e}_2, \vec{e}_3, \vec{e}_4\}$, on considère le système de vecteurs $\mathcal{S} = \{\vec{u}_1, \vec{u}_2, \vec{u}_3, \vec{u}_4, \vec{u}_5\}$ où $\vec{u}_1 = (1, 1, 0, 5)$, $\vec{u}_2 = (2, -1, 2, -4)$, $\vec{u}_3 = (4, 1, 2, 6)$, $\vec{u}_4 = (1, 0, 1, 0)$, $\vec{u}_5 = (4, 0, 3, 1)$. On remarque que $\mathrm{card}(\mathcal{S}) = 5 > 4 = \dim(E)$, le système $\mathcal{S}$ est donc lié. Pour déterminer le rang de $\mathcal{S}$, on dispose ses vecteurs sous forme d'un tableau :

— Faisons apparaître des zéros sur la première ligne à partir de la deuxième colonne :

$$\begin{array}{cc} \begin{array}{ccccc} \vec{u}_1' & \vec{u}_2' & \vec{u}_3' & \vec{u}_4' & \vec{u}_5' \end{array} & \\ \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 1 & -3 & -3 & -1 & -4 \\ 0 & 2 & 2 & 1 & 3 \\ 5 & -14 & -14 & -5 & -19 \end{pmatrix} & \begin{array}{c} \vec{e}_1 \\ \vec{e}_2 \\ \vec{e}_3 \\ \vec{e}_4 \end{array} \end{array}$$

avec $\vec{u}_1' = \vec{u}_1$, $\vec{u}_2' = \vec{u}_2 - 2\vec{u}_1$, $\vec{u}_3' = \vec{u}_3 - 4\vec{u}_1$, $\vec{u}_4' = \vec{u}_4 - \vec{u}_1$ et $\vec{u}_5' = \vec{u}_5 - 4\vec{u}_1$.

— Faisons apparaître des zéros sur la deuxième ligne à partir de la troisième colonne :

$$\begin{array}{cc} \begin{array}{ccccc} \vec{u}_1'' & \vec{u}_2'' & \vec{u}_3'' & \vec{u}_4'' & \vec{u}_5'' \end{array} & \\ \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 1 & -3 & 0 & 0 & 0 \\ 0 & 2 & 0 & 1 & 1 \\ 5 & -14 & 0 & -1 & -1 \end{pmatrix} & \begin{array}{c} \vec{e}_1 \\ \vec{e}_2 \\ \vec{e}_3 \\ \vec{e}_4 \end{array} \end{array}$$

avec $\vec{u}_1'' = \vec{u}_1'$, $\vec{u}_2'' = \vec{u}_2'$, $\vec{u}_3'' = \vec{u}_3' - \vec{u}_2'$, $\vec{u}_4'' = 3\vec{u}_4' - \vec{u}_2'$ et $\vec{u}_5'' = 3\vec{u}_5' - 4\vec{u}_2'$.

— Le terme apparaissant à l'intersection des troisième ligne et troisième colonne étant nul, on permute par exemple les vecteurs $\vec{u}_3''$ et $\vec{u}_5''$:

$$\begin{array}{cc} \begin{array}{ccccc} \vec{u}_1'' & \vec{u}_2'' & \vec{u}_5'' & \vec{u}_4'' & \vec{u}_3'' \end{array} & \\ \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 1 & -3 & 0 & 0 & 0 \\ 0 & 2 & 1 & 1 & 0 \\ 5 & -14 & -1 & -1 & 0 \end{pmatrix} & \begin{array}{c} \vec{e}_1 \\ \vec{e}_2 \\ \vec{e}_3 \\ \vec{e}_4 \end{array} \end{array}$$

et on fait maintenant apparaître des zéros sur la troisième ligne à partir de la quatrième colonne :

$$\begin{array}{cc} \begin{array}{ccccc} \vec{u}_1''' & \vec{u}_2''' & \vec{u}_3''' & \vec{u}_4''' & \vec{u}_5''' \end{array} & \\ \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 1 & -3 & 0 & 0 & 0 \\ 0 & 2 & 1 & 0 & 0 \\ 5 & -14 & -1 & 0 & 0 \end{pmatrix} & \begin{array}{c} \vec{e}_1 \\ \vec{e}_2 \\ \vec{e}_3 \\ \vec{e}_4 \end{array} \end{array}$$

avec $\vec{u}_1''' = \vec{u}_1''$, $\vec{u}_2''' = \vec{u}_2''$, $\vec{u}_3''' = \vec{u}_5''$, $\vec{u}_4''' = \vec{u}_4'' - \vec{u}_5''$ et $\vec{u}_5''' = \vec{u}_3''$.

Ainsi, $\mathrm{vect}(\mathcal{S}) = \mathrm{vect}(\{\vec{u}_1''', \vec{u}_2''', \vec{u}_3'''\}) = \mathrm{vect}(\{\vec{u}_1, \vec{u}_2 - 2\vec{u}_1, 3\vec{u}_5 - 4\vec{u}_2 - 4\vec{u}_1\})$, le système $\{(1,1,0,5), (0,-3,2,-14), (0,0,1,-1)\}$ est une base de $\mathrm{vect}(\mathcal{S})$ et $\mathrm{rang}(\mathcal{S}) = 3$. De plus les égalités $\vec{u}_4''' = \vec{u}_5''' = \vec{0}$ conduisent progressivement à $\vec{u}_4'' - \vec{u}_5'' = \vec{u}_3'' = \vec{0}$, puis $\vec{u}_2' + \vec{u}_4' - \vec{u}_5' = \vec{u}_3' - \vec{u}_2' = \vec{0}$ et enfin $\vec{u}_3 = 2\vec{u}_1 + \vec{u}_2$ et $\vec{u}_5 = \vec{u}_1 + \vec{u}_2 + \vec{u}_4$.

5 Exercices sur les espaces vectoriels

Exercice 1 Dans le $\mathbb{R}$-espace vectoriel $E = \mathbb{R}^3$, les systèmes de vecteurs suivants sont-ils libres ? générateurs de E ? des bases de E ?

$a)$ $\{(1, -1, 3,), (0, 1, -1), (0, 3, -2)\}$; $b)$ $\{(1, 3, -5), (1, -2, 3), (1, 8, -13)\}$;
$c)$ $\{(1, a, b), (1, 1, 0), (0, a - 1, c)\}$.

Exercice 2 Dans le $\mathbb{R}$-espace vectoriel $E = \mathbb{R}^3$ rapporté à sa base canonique $\mathcal{B} = \{\vec{i}, \vec{j}, \vec{k}\}$, on considère le système de vecteurs $\mathcal{B}' = \{\vec{i} + \vec{j} + \vec{k}, \vec{j} + 2\vec{k}, 2\vec{j} + \vec{k}\}$.

1. Montrer que $\mathcal{B}'$ est une base de E.
2. Déterminer l'ensemble F des vecteurs de E qui ont mêmes coordonnées dans les bases $\mathcal{B}$ et $\mathcal{B}'$.
3. Montrer que F est un sous-espace vectoriel de E.

Exercice 3 Dans le $\mathbb{R}$-espace vectoriel $E = \mathbb{R}^3$, on considère les sous-ensembles suivants :

$$F_1 = \{(-2\lambda, 3\lambda, 4\lambda),\ \lambda \in \mathbb{R}\}, \quad F_2 = \{(x, y, z) \in E : x + 6y - z = 0\},$$
$$F_3 = \{(x, y, z) \in E : (x + 2y = 0) \text{ et } (3x + 2y + z = 0)\}.$$

1. Montrer que F_1, F_2 et F_3 sont des sous-espaces vectoriels de E dont on déterminera pour chacun une base.
2. Déterminer les sous-espaces vectoriels $F_i \cap F_j$ et $F_i + F_j$ pour $i \neq j$, $i, j \in \{1, 2, 3\}$.

Exercice 4 Dans le $\mathbb{R}$-espace vectoriel $E = \mathbb{R}^4$, on considère les sous-ensembles suivants :

$$\begin{aligned} F_1 &= \{(\lambda + \mu, -\lambda, \lambda + 2\mu, \mu + \nu),\ \lambda, \mu, \nu \in \mathbb{R}\}, \\ F_2 &= \{(x, y, z, t) \in E : (3x + 4y + z + 2t = 0) \\ &\qquad \text{et } (6x + 8y + 2z + 5t = 0) \text{ et } (9x + 12y + 3z + 10t = 0)\}. \end{aligned}$$

1. Montrer que F_1 et F_2 sont des sous-espaces vectoriels de E dont on déterminera pour chacun une base ainsi que sa dimension.
2. Déterminer les sous-espaces vectoriels $F_1 \cap F_2$ et $F_1 + F_2$. Les sous-espaces vectoriels F_1 et F_2 sont-ils supplémentaires dans E ?
3. Déterminer un sous-espace vectoriel G_1 de F_1 tel que G_1 et F_2 soient des sous-espaces vectoriels supplémentaires dans E. Déterminer de même un sous-espace vectoriel G_2 de F_2 tel que F_1 et G_2 soient des sous-espaces vectoriels supplémentaires dans E.

Exercice 5 Dans le $\mathbb{R}$-espace vectoriel $E = \mathbb{R}^3$ rapporté à sa base canonique $\mathcal{B} = \{\vec{i}, \vec{j}, \vec{k}\}$, on considère les vecteurs

$$\vec{u}_1 = -\vec{i} + \vec{j} - \vec{k}, \quad \vec{u}_2 = \vec{i} + 2\vec{j} + 4\vec{k}, \quad \vec{u}_3 = 3\vec{i} - \vec{j} + a\vec{k}, \quad \vec{u}_4 = 2\vec{i} + 3\vec{j} + b\vec{k}.$$

Déterminer les réels a et b de telle sorte que les systèmes de vecteurs $\{\vec{u}_1, \vec{u}_2\}$ et $\{\vec{u}_3, \vec{u}_4\}$ engendrent le même sous-espace vectoriel de E.

Exercice 6 Déterminer le rang des systèmes de vecteurs suivants :

a) $\{(2,-1,3),(-1,0,1),(7,5,4),(-2,6,8)\}$;

b) $\{(1,1,\alpha),(1,\alpha,1),(\alpha,1,1)\}$;

c) $\{(1,1,0,0),(0,1,1,0),(0,0,1,1),(1,0,0,1)\}$;

d) $\{(1,0,2,3),(7,4,-2,1),(5,2,4,7),(3,2,0,1),(0,0,1,1)\}$.

Exercice 7 Soit a et b deux réels distincts. On se place dans le $\mathbb{R}$-espace vectoriel $E=\mathbb{R}_n[X]$ et l'on pose pour tout $\lambda\in\mathbb{R}$, $F_\lambda=\{P\in E : P(\lambda)=0\}$.

1. Montrer que F_a est un sous-espace vectoriel de E dont on déterminera une base ainsi que la dimension.
2. Déterminer une base ainsi que la dimension du sous-espace vectoriel $F_a\cap F_b$.
3. En déduire la dimension du sous-espace vectoriel F_a+F_b. A-t-on $E=F_a+F_b$? La somme est-elle directe ?
4. Donner tous les sous-espaces vectoriels supplémentaires de F_a dans E.

Exercice 8 On se place dans le $\mathbb{R}$-espace vectoriel E des fonctions définies sur $\mathbb{R}$ à valeurs réelles et on considère le sous-ensemble F des fonctions $f_{a,\varphi}\in E$ définies par $\forall x\in\mathbb{R}, f_{a,\varphi}(x)=a\cos(x+\varphi)$.

1. Montrer que F est un sous-espace vectoriel de E ; on pourra écrire $f_{a,\varphi}$ sous une autre forme.
2. Déterminer une base de F et calculer sa dimension.

Exercice 9 Soit $a,b\in\mathbb{C}$ et E le $\mathbb{C}$-espace vectoriel des fonctions définies sur $\mathbb{R}$ à valeurs complexes de classe C^∞.

1. Déterminer l'ensemble F des solutions dans E de l'équation différentielle $y''+ay'+by=0$.
2. Montrer que F est un sous-espace vectoriel de E.
3. Déterminer une base de F, puis en déduire sa dimension.

Exercice 10 Soit $E=\mathbb{C}^2$.

1. Définir sur E la structure d'espace vectoriel « canonique » sur le corps $\mathbb{C}$.
2. Déterminer une base du $\mathbb{C}$-espace vectoriel E.
3. Le sous-ensemble $\{(z_1,z_2)\in E : z_1+z_2\in\mathbb{R}\}$ est-il un sous-espace vectoriel de E ? Dans l'affirmative, en déterminer une base.
4. Le système de vecteurs $\{(1,0),(1,i),(1,2)\}$ est-il libre ? générateur de E ?
5. Reprendre les questions précédentes en prenant $\mathbb{R}$ au lieu de $\mathbb{C}$ pour corps de scalaires.

Exercice 11 Soit E le $\mathbb{R}$-espace vectoriel des fonctions définies sur $\mathbb{R}$ à valeurs dans $\mathbb{R}$, $n\geqslant 1$ et $\alpha_1,\dots,\alpha_n$ des réels ordonnés : $\alpha_1<\dots<\alpha_n$.

1. On pose pour tout $x\in\mathbb{R}$ et tout $\alpha\in\mathbb{R}$, $e_\alpha(x)=e^{\alpha x}$. Montrer que le système de vecteurs $\{e_{\alpha_1},\dots,e_{\alpha_n}\}$ de E est libre. On pourra utiliser le comportement à l'infini des fonctions e_α.

2. On pose pour tout $x \in \mathbb{R}$ et tout $\alpha \in \mathbb{R}$, $s_\alpha(x) = \sin(\alpha x)$. Montrer que le système de vecteurs $\{s_{\alpha_1}, \ldots, s_{\alpha_n}\}$ de E est libre. On pourra raisonner par récurrence sur n et utiliser la relation $s''_\alpha = -\alpha^2 s_\alpha$.
3. Que dire de la dimension du $\mathbb{R}$-espace vectoriel E ?

Chapitre 2

Applications linéaires

Soit E et F deux $\mathbb{K}$-espaces vectoriels. On considère les applications entre E et F qui sont « compatibles » avec la structure d'espace vectoriel. Ce sont les *applications linéaires* ; elles « transportent » les combinaisons linéaires. Les plus simples sont les applications de $\mathbb{K}$ dans $\mathbb{K}$ de la forme $f: \begin{array}{ccc} \mathbb{K} & \longrightarrow & \mathbb{K} \\ x & \longmapsto & ax \end{array}$. Elles vérifient les relations $f(x+y) = f(x)+f(y)$ et $f(\alpha x) = \alpha f(x)$.
Un tel exemple se rencontre dans diverses situations. Par exemple, en électricité, les lois d'Ohm [1] $U = RI$, $Q = CU$ expriment des dépendances *linéaires* entre intensité ou charge électrique et tension. Un autre exemple issu du même domaine concerne les quadripôles dits « linéaires » pour lesquels on dispose d'un couple courant-tension $(I_\mathrm{e}, U_\mathrm{e})$ en entrée et d'un couple courant-tension $(I_\mathrm{s}, U_\mathrm{s})$ en sortie. Les lois de Kirchhoff [2] expriment des relations *linéaires* entre ces deux couples :

$$\left\{\begin{array}{rcl} I_\mathrm{s} & = & aI_\mathrm{e} + bU_\mathrm{e} \\ U_\mathrm{s} & = & cI_\mathrm{e} + dU_\mathrm{e} \end{array}\right. .$$

C'est un exemple d'application linéaire de $\mathbb{R}^2$ dans $\mathbb{R}^2$. De nombreux phénomènes de la nature sont modélisés par des relations entre diverses variables, notamment de la forme $y = f(x_1, \dots, x_n)$, f étant une fonction de n variables à valeurs réelles. De telles fonctions f sont en général très complexes et des approximations sont indispensables pour résoudre les problèmes concernés, lesquels sont qualifiés de manière floue de « non-linéaires ». Le calcul différentiel propose dans cet état d'esprit des approximations des premier, deuxième, etc., ordres. Une approximation du premier ordre au voisinage d'un point $x = (x_1, \dots, x_n)$ est fournie par la différentielle de f, c'est typiquement une application linéaire d'une variable vectorielle $\vec{h} = (h_1, \dots, h_n)$ définie par

$$D_x f(\vec{h}) = h_1 \frac{\partial f}{\partial x_1} + \cdots + h_n \frac{\partial f}{\partial x_n}.$$

Notons que ce n'est en fait qu'une fonction polynomiale des n variables $h_1, \dots, h_n$ de degré un relativement à chacune d'elles sans terme constant, et qu'elle vérifie les relations $f(\vec{h} + \vec{k}) = f(\vec{h}) + f(\vec{k})$ et $f(\alpha\vec{h}) = \alpha f(\vec{h})$. Ces relations vont servir de définition pour une application linéaire dans un contexte vectoriel quelconque.

Dans ce chapitre, E et F désignent deux $\mathbb{K}$-espaces vectoriels. Leurs vecteurs nuls seront notés respectivement $\vec{0}_E$ et $\vec{0}_F$.

1. Ohm, Georg Simon : physicien allemand (Erlangen 1789–Münich 1854)
2. Kirchhoff, Gustav Robert : physicien allemand (Königsberg 1824–Berlin 1887)

1 Applications linéaires

1.1 Définitions, propriétés

Définition 1.1 *Une application $f : E \longrightarrow F$ est linéaire lorsque*

$$\forall \vec{u}, \vec{v} \in E,\ f(\vec{u}+\vec{v}) = f(\vec{u}) + f(\vec{v})$$

et

$$\forall \alpha \in \mathbb{K}, \forall \vec{u} \in E,\ f(\alpha\vec{u}) = \alpha f(\vec{u}),$$

ou encore, de manière équivalente,

$$\forall \alpha, \beta \in \mathbb{K}, \forall \vec{u}, \vec{v} \in E,\ f(\alpha\vec{u}+\beta\vec{v}) = \alpha f(\vec{u}) + \beta f(\vec{v}).$$

On note $\mathcal{L}(E,F)$ l'ensemble des applications linéaires de E dans F ; lorsque $E = F$, on le note plus simplement $\mathcal{L}(E)$.

Propriétés immédiates.

Si $f \in \mathcal{L}(E,F)$, on a
— $f(\vec{0}_E) = \vec{0}_F$;
— $\forall \vec{u} \in E,\ f(-\vec{u}) = -f(\vec{u})$;
— $\forall \alpha_1, \dots, \alpha_n \in \mathbb{K}, \forall \vec{u}_1, \dots, \vec{u}_n \in E,\ f(\alpha_1\vec{u}_1 + \cdots + \alpha_n\vec{u}_n) = \alpha_1 f(\vec{u}_1) + \cdots + \alpha_n f(\vec{u}_n)$.

Définition 1.2 *Dans certains cas, une application linéaire porte un nom particulier :*
- *une* forme linéaire *sur E est une application linéaire de E sur son corps de base $\mathbb{K}$;*
- *un* endomorphisme *de E est une application linéaire de E dans E ;*
- *un* isomorphisme *de E sur F est une bijection linéaire de E dans F. Dans ce cas, on dit que les $\mathbb{K}$-espaces vectoriels E et F sont isomorphes ;*
- *un* automorphisme *de E est un isomorphisme de E sur lui-même.*

Exemple.

Homothéties de E : soit $\lambda \in \mathbb{K}$. L'*homothétie de rapport λ* est l'application $h_\lambda : E \longrightarrow E$ définie par $\forall \vec{u},\ h_\lambda(\vec{u}) = \lambda\vec{u}$. Il est clair que si le scalaire λ est non nul, l'application h_λ est un automorphisme de E. En particulier pour $\lambda = 1$, h_1 est l'application *identité* de E notée id_E : $\forall \vec{u},\ \mathrm{id}_E(\vec{u}) = \vec{u}$.

Proposition 1.3
1. *Soit E, F, G trois $\mathbb{K}$-espaces vectoriels et $f \in \mathcal{L}(F,G)$, $g \in \mathcal{L}(E,F)$. Alors $f \circ g \in \mathcal{L}(E,G)$.*
2. *Si f est un isomorphisme de E sur F, alors $f^{-1} \in \mathcal{L}(F,E)$.*

Démonstration.

1. Soit $\alpha, \beta \in \mathbb{K}$ et $\vec{u}, \vec{v} \in E$. On a

$$\begin{aligned}(f \circ g)(\alpha\vec{u}+\beta\vec{v}) &= f(g(\alpha\vec{u}+\beta\vec{v})) = f(\alpha g(\vec{u}) + \beta g(\vec{v})) \\ &= \alpha f(g(\vec{u})) + \beta f(g(\vec{v})) = \alpha(f \circ g)(\vec{u}) + \beta(f \circ g)(\vec{v})\end{aligned}$$

et donc $f \circ g \in \mathcal{L}(E, G)$.

2. Soit f un isomorphisme de E sur F. Soit $\alpha, \beta \in \mathbb{K}$ et $\vec{u}, \vec{v} \in E$. On doit montrer $f^{-1}(\alpha\vec{u}+\beta\vec{v}) = \alpha f^{-1}(\vec{u}) + \beta f^{-1}(\vec{v})$. Cela est équivalent, puisque f est bijective, à montrer que ces deux vecteurs ont même image par f. On a

$$\begin{aligned} f(\alpha f^{-1}(\vec{u}) + \beta f^{-1}(\vec{v})) &= \alpha f(f^{-1}(\vec{u})) + \beta f(f^{-1}(\vec{v})) = \alpha(f \circ f^{-1})(\vec{u}) + \beta(f \circ f^{-1})(\vec{v}) \\ &= \alpha\vec{u} + \beta\vec{v} = (f \circ f^{-1})(\alpha\vec{u} + \beta\vec{v}) = f(f^{-1}(\alpha\vec{u} + \beta\vec{v})), \end{aligned}$$

d'où le résultat. □

Exemples.

1. Considérons les $\mathbb{R}$-espaces vectoriels $E = \mathbb{R}^3$ et $F = \mathbb{R}^2$ et l'application

$$f : \begin{array}{rcl} E & \longrightarrow & F \\ (x, y, z) & \longmapsto & (2x + 3y + z, x - 4y - z) \end{array}$$

Cette application est donc définie par la relation $f((x, y, z)) = (2x + 3y + z, x - 4y - z)$.
Afin d'alléger les notations, on note $f((x, y, z))$ plutôt $f(x, y, z)$.
Soit $\alpha, \beta \in \mathbb{K}$ et $\vec{u}, \vec{v} \in E$. Les vecteurs $\vec{u}, \vec{v}$ sont de la forme $\vec{u} = (x, y, z)$ et $\vec{v} = (x', y', z')$.
On a

$$\begin{aligned} &f(\alpha\vec{u} + \beta\vec{v}) \\ &\quad = f(\alpha x + \beta x', \alpha y + \beta y', \alpha z + \beta z') \\ &\quad = (2(\alpha x + \beta x') + 3(\alpha y + \beta y') + (\alpha z + \beta z'), (\alpha x + \beta x') - 4(\alpha y + \beta y') - (\alpha z + \beta z')) \\ &\quad = (\alpha(2x + 3y + z) + \beta(2x' + 3y' + z'), \alpha(x - 4y - z) + \beta(x' - 4y' - z')) \\ &\quad = \alpha(2x + 3y + z, x - 4y - z) + \beta(2x' + 3y' + z', x' - 4y' - z') \\ &\quad = \alpha f(\vec{u}) + \beta f(\vec{v}). \end{aligned}$$

Ainsi $f \in \mathcal{L}(E, F)$.

2. Dans le $\mathbb{K}$-espace vectoriel $E = \mathbb{K}[X]$, les applications « dérivation » et « intégration » respectivement définies par $P(X) \longmapsto P'(X)$ et $P(X) \longmapsto \displaystyle\int_0^X P(x)\,dx$ sont des endomorphismes de E. L'application « translation » définie par $P(X) \longmapsto P(X + a)$ est un automorphisme de E de réciproque $P(X) \longmapsto P(X - a)$.

1.2 Matrice et représentation analytique d'une application linéaire en dimension finie

Supposons que $\dim(E) = n$ et $\dim(F) = p$ et introduisons des bases de E et F : $\mathcal{B}_E = \{\vec{e}_1, \ldots, \vec{e}_n\}$ et $\mathcal{B}_F = \{\vec{\varepsilon}_1, \ldots, \vec{\varepsilon}_p\}$. Soit $\vec{u} \in E$ et $f \in \mathcal{L}(E, F)$. En décomposant $\vec{u}$ relativement à la base $\mathcal{B}_E$, $\vec{u} = \displaystyle\sum_{j=1}^{n} x_j \vec{e}_j$, et en utilisant les propriétés de l'application linéaire f, on voit que

$$f(\vec{u}) = f\left(\sum_{j=1}^{n} x_j \vec{e}_j\right) = \sum_{j=1}^{n} x_j f(\vec{e}_j).$$

La relation ci-dessus signifie que l'image d'un vecteur quelconque de E est déterminée dès que l'on connaît les images des vecteurs de la base $\mathcal{B}_E$: $f(\vec{e}_1), \ldots, f(\vec{e}_n)$. On a donc prouvé le résultat fondamental suivant.

Théorème 1.4 *Une application linéaire $f \in \mathcal{L}(E,F)$ est entièrement caractérisée par la donnée de l'image d'une base de E par f.*

Introduisons maintenant les coordonnées des vecteurs de $f(\mathcal{B}_E)$ dans la base $\mathcal{B}_F$:

$$f(\vec{e}_j) = \sum_{i=1}^{p} a_{ij}\vec{\varepsilon}_i, \quad 1 \leqslant j \leqslant n,$$

ainsi que celles du vecteur-image de $\vec{u}$:

$$f(\vec{u}) = \sum_{i=1}^{p} y_i\vec{\varepsilon}_i.$$

On a

$$f(\vec{u}) = \sum_{j=1}^{n} x_j f(\vec{e}_j) = \sum_{j=1}^{n} x_j \left(\sum_{i=1}^{p} a_{ij}\vec{\varepsilon}_i\right) = \sum_{i=1}^{p} \left(\sum_{j=1}^{n} a_{ij}x_j\right)\vec{\varepsilon}_i$$

ce qui fournit

$$\boxed{y_i = \sum_{j=1}^{n} a_{ij}x_j, \quad 1 \leqslant i \leqslant p.}$$

Ceci s'écrit de manière plus détaillée

$$\left\{\begin{array}{rcl} y_1 & = & a_{11}\,x_1 + \cdots + a_{1n}\,x_n \\ \vdots & & \\ y_p & = & a_{p1}\,x_1 + \cdots + a_{pn}\,x_n \end{array}\right. .$$

Le système précédent est la *représentation analytique* de l'application linéaire f. Elle est caractérisée par le tableau rectangulaire de scalaires $A = (a_{ij})_{\substack{1\leqslant i\leqslant p\\1\leqslant j\leqslant n}}$ qui est la *matrice* de f relativement aux bases $\mathcal{B}_E$ et $\mathcal{B}_F$. On la note $\mathrm{Mat}(f;\mathcal{B}_E,\mathcal{B}_F)$ et plus simplement $\mathrm{Mat}(f;\mathcal{B}_E)$ lorsque $E = F$ et $\mathcal{B}_E = \mathcal{B}_F$, et on la dispose de la manière suivante :

$$A = \mathrm{Mat}(f;\mathcal{B}_E,\mathcal{B}_F) = \begin{array}{c} \begin{array}{cccc} f(\vec{e}_1) & f(\vec{e}_2) & \ldots & f(\vec{e}_n) \end{array} \\ \begin{pmatrix} a_{11} & a_{12} & & a_{1n} \\ a_{21} & a_{22} & & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{p1} & a_{p2} & & a_{pn} \end{pmatrix} \end{array} \begin{array}{c} \\ \vec{\varepsilon}_1 \\ \vec{\varepsilon}_2 \\ \vdots \\ \vec{\varepsilon}_p \end{array} .$$

C'est un tableau à p lignes et n colonnes qui contient toute l'information de l'application linéaire f. On dit que la matrice est de *type* $p \times n$. La coutume veut que le premier (resp. deuxième) indice du scalaire a_{ij} désigne le numéro de la ligne (resp. colonne) sur laquelle il se situe. Les coordonnées des vecteurs-images $f(\vec{e}_j)$, $1 \leqslant j \leqslant n$, sont disposées en colonnes. La lecture de ce tableau est donc **verticale**. Pour un vecteur, on définit également une *matrice-colonne* contenant ses coordonnées relativement à une base. Par exemple, dans notre étude, on écrit

$$X = \mathrm{Mat}(\vec{u};\mathcal{B}_E) = \begin{array}{c} \vec{u} \\ \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \end{array} \begin{array}{c} \\ \vec{e}_1 \\ \vec{e}_2 \\ \vdots \\ \vec{e}_n \end{array} \quad \text{et} \quad Y = \mathrm{Mat}(f(\vec{u});\mathcal{B}_F) = \begin{array}{c} f(\vec{u}) \\ \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_p \end{pmatrix} \end{array} \begin{array}{c} \\ \vec{\varepsilon}_1 \\ \vec{\varepsilon}_2 \\ \vdots \\ \vec{\varepsilon}_p \end{array} .$$

Au vu de l'écriture analytique de l'application linéaire f, on définit enfin le *produit matriciel* de la matrice-rectangle A (de type $p \times n$) par la matrice-colonne X (de type $n \times 1$) comme étant la matrice-colonne (de type $p \times 1$)

$$AX = \begin{pmatrix} \displaystyle\sum_{j=1}^{n} a_{1j}x_j \\ \displaystyle\sum_{j=1}^{n} a_{2j}x_j \\ \vdots \\ \displaystyle\sum_{j=1}^{n} a_{pj}x_j \end{pmatrix} .$$

Ainsi, la représentation analytique de l'application linéaire f s'écrit de manière condensée

$$\boxed{Y = AX}$$

qui est la simple généralisation de la relation $y = ax$ pour les applications linéaires de $\mathbb{K}$ dans $\mathbb{K}$ (dimension 1). La représentation analytique de f correspond à une lecture **horizontale** de la matrice : la i^{e} équation du système $Y = AX$ s'écrit $y_i = a_{i1}x_1 + \cdots + a_{in}x_n$, elle fait intervenir les coefficients $a_{i1}, \ldots, a_{in}$ qui constituent la i^{e} ligne de A.

Exemples.

1. Considérons les $\mathbb{R}$-espaces vectoriels $E = \mathbb{R}^3$ et $F = \mathbb{R}^2$ respectivement rapportés à leurs bases canoniques $\mathcal{B}_E = \{\vec{i}, \vec{j}, \vec{k}\}$ et $\mathcal{B}_F = \{\vec{\varepsilon}_1, \vec{\varepsilon}_2\}$. Soit $f \in \mathcal{L}(E, F)$ l'application linéaire telle que $f(\vec{i}) = 2\vec{\varepsilon}_1 + \vec{\varepsilon}_2$, $f(\vec{j}) = 3\vec{\varepsilon}_1 - 4\vec{\varepsilon}_2$, $f(\vec{k}) = \vec{\varepsilon}_1 - \vec{\varepsilon}_2$. La matrice de f relativement aux bases $\mathcal{B}_E$ et $\mathcal{B}_F$ est donnée par

$$A = \mathrm{Mat}(f;\mathcal{B}_E, \mathcal{B}_F) = \begin{array}{c} \begin{array}{ccc} f(\vec{i}) & f(\vec{j}) & f(\vec{k}) \end{array} \\ \begin{pmatrix} 2 & 3 & 1 \\ 1 & -4 & -1 \end{pmatrix} \end{array} \begin{array}{c} \\ \vec{\varepsilon}_1 \\ \vec{\varepsilon}_2 \end{array} .$$

On obtient ensuite l'image $(x_1, x_2) = x_1\vec{\varepsilon}_1 + x_2\vec{\varepsilon}_2$, de matrice-colonne $Y = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$, d'un vecteur $(x, y, z) = x\vec{i} + y\vec{j} + z\vec{k}$ en multipliant la matrice-rectangle A par la matrice-colonne

$X = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$:

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = Y = AX = \begin{pmatrix} 2x+3y+z \\ x-4y-z \end{pmatrix},$$

ce qui fournit la représentation analytique de f :

$$\begin{cases} x_1 = 2x+3y+z \\ x_2 = x-4y-z \end{cases},$$

c'est-à-dire $f(x,y,z) = (2x+3y+z, x-4y-z)$.

2. Pour l'application identité de E, les vecteurs $\mathrm{id}_E(\vec{e}_i) = \vec{e}_i$, $1 \leqslant i \leqslant n$, ont pour coordonnées $(0,\dots,0,\underset{\substack{\uparrow \\ i^{\text{e}}\text{position}}}{1},0,\dots,0)$. La matrice de id_E relativement aux bases $\mathcal{B}_E$ et $\mathcal{B}_E$ est donc donnée par

$$\mathrm{Mat}(\mathrm{id}_E;\mathcal{B}_E) = \begin{pmatrix} 1 & 0 & & 0 \\ 0 & 1 & & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & & 1 \end{pmatrix}$$

que l'on appelle *matrice-unité d'ordre* n et que l'on note I_n.

3. Rappelons qu'une forme linéaire φ sur E est une application de E dans $F = \mathbb{K}$. On choisira pour base de F sa base canonique $\mathcal{B}_F = \{1\}$. Elle est entièrement déterminée par la donnée des scalaires $\varphi(\vec{e}_1) = a_1, \varphi(\vec{e}_2) = a_2, \dots, \varphi(\vec{e}_n) = a_n$. La matrice de φ relativement aux bases $\mathcal{B}_E$ et $\mathcal{B}_F$ est donnée par

$$A = \mathrm{Mat}(\varphi;\mathcal{B}_E,\mathcal{B}_F) = \begin{array}{cc} \begin{array}{cccc} \varphi(\vec{e}_1) & \varphi(\vec{e}_2) & \dots & \varphi(\vec{e}_n) \end{array} & \\ \begin{pmatrix} a_1 & a_2 & \dots & a_n \end{pmatrix} & 1 \end{array}.$$

On obtient ensuite l'image y, de matrice-colonne $Y = \begin{pmatrix} y \end{pmatrix}$, d'un vecteur $x_1\vec{e}_1 + x_2\vec{e}_2 + \dots + x_n\vec{e}_n$ en multipliant la matrice-ligne A par la matrice-colonne $X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$:

$$\begin{pmatrix} y \end{pmatrix} = Y = AX = \begin{pmatrix} a_1x_1 + a_2x_2 + \dots + a_nx_n \end{pmatrix},$$

ce qui fournit l'écriture analytique de φ :

$$\varphi\left(\sum_{i=1}^n x_i\vec{e}_i\right) = \sum_{i=1}^n a_ix_i.$$

Pour $E = F = \mathbb{K}$, on retrouve les applications linéaires les plus simples $\begin{array}{ccc} \mathbb{K} & \longrightarrow & \mathbb{K} \\ x & \longmapsto & ax \end{array}$.

1.3 Composition d'applications linéaires

Soit E, F, G trois $\mathbb{K}$-espaces vectoriels de dimensions respectives n, p, q, et $f \in \mathcal{L}(F,G)$, $g \in \mathcal{L}(E,F)$. Il a déjà été vu que $f \circ g \in \mathcal{L}(E,G)$. On va essayer d'exprimer la matrice de $f \circ g$ à l'aide des matrices de f et g. Introduisons donc des bases $\mathcal{B}_E = \{\vec{e}_1, \ldots, \vec{e}_n\}$, $\mathcal{B}_F = \{\vec{\varepsilon}_1, \ldots, \vec{\varepsilon}_p\}$, $\mathcal{B}_G = \{\vec{\eta}_1, \ldots, \vec{\eta}_q\}$ de E, F, G respectivement, puis les matrices de f, g et $f \circ g$ relatives à celles-ci :

$$A = \mathrm{Mat}(f; \mathcal{B}_F, \mathcal{B}_G) = (a_{ij})_{\substack{1 \leqslant i \leqslant q \\ 1 \leqslant j \leqslant p}}, \quad B = \mathrm{Mat}(g; \mathcal{B}_E, \mathcal{B}_F) = (b_{ij})_{\substack{1 \leqslant i \leqslant p \\ 1 \leqslant j \leqslant n}}$$

et

$$C = \mathrm{Mat}(f \circ g; \mathcal{B}_E, \mathcal{B}_G) = (c_{ij})_{\substack{1 \leqslant i \leqslant q \\ 1 \leqslant j \leqslant n}}.$$

La situation est schématisée par le diagramme suivant dans lequel on a mis en indice les dimensions des espaces vectoriels :

$$\begin{array}{c} (E_n, \mathcal{B}_E) \xrightarrow{g} (F_p, \mathcal{B}_F) \xrightarrow{f} (G_q, \mathcal{B}_G) \\ \xrightarrow[f \circ g]{\hspace{6em}} \end{array}.$$

On a donc, par définition des matrices A, B et C,

$$f(\vec{\varepsilon}_k) = \sum_{i=1}^{q} a_{ik} \vec{\eta}_i, \quad 1 \leqslant k \leqslant p, \qquad g(\vec{e}_j) = \sum_{k=1}^{p} b_{kj} \vec{\varepsilon}_k, \quad 1 \leqslant j \leqslant n,$$

et

$$(f \circ g)(\vec{e}_j) = \sum_{i=1}^{q} c_{ij} \vec{\eta}_i, \quad 1 \leqslant j \leqslant n.$$

On a alors

$$(f \circ g)(\vec{e}_j) = f\left(\sum_{k=1}^{p} b_{kj} \vec{\varepsilon}_k\right) = \sum_{k=1}^{p} b_{kj} f(\vec{\varepsilon}_k) = \sum_{k=1}^{p} b_{kj} \left(\sum_{i=1}^{q} a_{ik} \vec{\eta}_i\right) = \sum_{i=1}^{q} \left(\sum_{k=1}^{p} a_{ik} b_{kj}\right) \vec{\eta}_i$$

d'où l'on tire

$$\boxed{c_{ij} = \sum_{k=1}^{p} a_{ik} b_{kj}, \quad 1 \leqslant i \leqslant q, 1 \leqslant j \leqslant n.}$$

La relation ci-dessus définit une matrice C appelée *produit* des matrices A et B qui est notée $A \times B$ ou encore AB. On a donc obtenu la matrice de la composée de deux applications linéaires.

Proposition 1.5 *Soit E, F, G trois $\mathbb{K}$-espaces vectoriels de dimension finie et $f \in \mathcal{L}(F, G)$, $g \in \mathcal{L}(E, F)$. On a*

$$\mathrm{Mat}(f \circ g; \mathcal{B}_E, \mathcal{B}_G) = \mathrm{Mat}(f; \mathcal{B}_F, \mathcal{B}_G) \times \mathrm{Mat}(g; \mathcal{B}_E, \mathcal{B}_F).$$

DESSIN A FAIRE

Remarques.

1. Notons que le produit de matrices n'est pas commutatif puisque la loi de composition $\circ$ n'est pas commutative.
2. Pour pouvoir calculer le produit de deux matrices, il faut que le nombre de colonnes de la matrice de gauche coïncide avec le nombre de lignes de celle de droite, ce qui provient du fait que pour pouvoir composer f et g, il faut que l'ensemble d'arrivée de g coïncide avec l'ensemble de départ de f.
3. La proposition précédente appliquée aux endomorphismes $f \in \mathcal{L}(E)$ et $\mathrm{id}_E \in \mathcal{L}(E)$ donne (ici $E = F$ et $\mathcal{B}_E = \mathcal{B}_F$)

$$\mathrm{Mat}(f;\mathcal{B}_E) \times \mathrm{Mat}(\mathrm{id}_E;\mathcal{B}_E) = \mathrm{Mat}(f \circ \mathrm{id}_E;\mathcal{B}_E) = \mathrm{Mat}(f;\mathcal{B}_E)$$

et

$$\mathrm{Mat}(\mathrm{id}_E;\mathcal{B}_E) \times \mathrm{Mat}(f_E;\mathcal{B}_E) = \mathrm{Mat}(\mathrm{id}_E \circ f;\mathcal{B}_E) = \mathrm{Mat}(f;\mathcal{B}_E),$$

soit encore

$$A \times I_n = I_n \times A = A.$$

Cette égalité signifie que I_n est l'*élément neutre* de la multiplication des matrices carrées de type $n \times n$.

Exemple.

Considérons dans les $\mathbb{R}$-espaces vectoriels $E = \mathbb{R}^2$ et $F = \mathbb{R}^4$ munis de leurs bases canoniques $\mathcal{B}_E = \{\vec{i},\vec{j}\}$ et $\mathcal{B}_F = \{\vec{e}_1,\vec{e}_2,\vec{e}_3,\vec{e}_4\}$ les applications linéaires $f \in \mathcal{L}(E,F)$ et $g \in \mathcal{L}(F,E)$ définies par

$$f : \begin{array}{rcl} E & \longrightarrow & F \\ (x,y) & \longmapsto & (x+y, 2x-y, x+3y, -4x+y) \end{array}$$

et

$$g : \begin{array}{rcl} F & \longrightarrow & E \\ (x,y,z,t) & \longmapsto & (x+4y+z+t, 2x+3y-z+t) \end{array}.$$

Les matrices de f et g relativement aux bases $\mathcal{B}_E$ et $\mathcal{B}_F$ s'écrivent

$$A = \begin{array}{cc} \begin{array}{cc} f(\vec{i}) & f(\vec{j}) \end{array} & \\ \begin{pmatrix} 1 & 1 \\ 2 & -1 \\ 1 & 3 \\ -4 & 1 \end{pmatrix} & \begin{array}{c} \vec{e}_1 \\ \vec{e}_2 \\ \vec{e}_3 \\ \vec{e}_4 \end{array} \end{array} \quad \text{et} \quad B = \begin{array}{cc} \begin{array}{cccc} f(\vec{e}_1) & f(\vec{e}_2) & f(\vec{e}_3) & f(\vec{e}_4) \end{array} & \\ \begin{pmatrix} 1 & 4 & 1 & 1 \\ 2 & 3 & -1 & 1 \end{pmatrix} & \begin{array}{c} \vec{i} \\ \vec{j} \end{array} \end{array},$$

celle de $f \circ g \in \mathcal{L}(F)$ est donnée par

$$C = A \times B = \begin{pmatrix} 3 & 7 & 0 & 2 \\ 0 & 5 & 3 & 1 \\ 7 & 10 & -2 & 4 \\ -2 & -13 & -5 & -3 \end{pmatrix}$$

et celle de $g \circ f \in \mathcal{L}(E)$ par

$$D = B \times A = \begin{pmatrix} 6 & 1 \\ 3 & -3 \end{pmatrix}.$$

Des matrices C et D, on récupère les expressions explicites de $f \circ g$ et $g \circ f$:

— pour $f \circ g$, on effectue le produit matriciel $Y = CX$ qui fournit

$$(f \circ g)(x, y, z, t) = (3x + 7y + 2t, 5y + 3z + t, 7x + 10y - 2z + 4t, -2x - 13y - 5z - 3t)\ ;$$

— pour $g \circ f$, le produit matriciel $Y = DX$ conduit à

$$(g \circ f)(x, y) = (6x + y, 3x - 3y).$$

EXEMPLES OPTIQUE, QUADRIPOLES A FAIRE

Supposons maintenant que f soit un isomorphisme de E sur F et appliquons l'étude précédente aux applications linéaires $f \in \mathcal{L}(E, F)$ et $f^{-1} \in \mathcal{L}(F, E)$. On verra ultérieurement que les espaces vectoriels E et F sont nécessairement de même dimension n. On a donc

$$\mathrm{Mat}(f^{-1}; \mathcal{B}_F, \mathcal{B}_E) \times \mathrm{Mat}(f; \mathcal{B}_E, \mathcal{B}_F) = \mathrm{Mat}(f^{-1} \circ f; \mathcal{B}_E, \mathcal{B}_E) = \mathrm{Mat}(\mathrm{id}_E; \mathcal{B}_E) = I_n$$

et

$$\mathrm{Mat}(f; \mathcal{B}_E, \mathcal{B}_F) \times \mathrm{Mat}(f^{-1}; \mathcal{B}_F, \mathcal{B}_E) = \mathrm{Mat}(f \circ f^{-1}; \mathcal{B}_F, \mathcal{B}_F) = \mathrm{Mat}(\mathrm{id}_F; \mathcal{B}_F) = I_n.$$

Les deux égalités ci-dessus traduisent le fait que les matrices $\mathrm{Mat}(f; \mathcal{B}_E, \mathcal{B}_F)$ et $\mathrm{Mat}(f^{-1}; \mathcal{B}_F, \mathcal{B}_E)$ sont *inverses* l'une de l'autre pour la multiplication des matrices carrées et on note dans ce cas $\mathrm{Mat}(f^{-1}; \mathcal{B}_F, \mathcal{B}_E) = \mathrm{Mat}(f; \mathcal{B}_F, \mathcal{B}_E)^{-1}$.

Proposition 1.6 *Soit E, F deux $\mathbb{K}$-espaces vectoriels de dimension finies et f un isomorphisme de E sur F. On a*

$$\mathrm{Mat}(f^{-1}; \mathcal{B}_F, \mathcal{B}_E) = \mathrm{Mat}(f; \mathcal{B}_F, \mathcal{B}_E)^{-1}.$$

Exemple.

Considérons le $\mathbb{R}$-espace vectoriel $E = \mathbb{R}^2$ rapporté à sa base canonique $\mathcal{B}_E = \{\vec{i}, \vec{j}\}$. Soit $f \in \mathcal{L}(E)$ l'endomorphisme défini par

$$f : \begin{array}{ccc} E & \longrightarrow & E \\ (x, y) & \longmapsto & (2x + y, 3x + y) \end{array}.$$

La matrice de f relativement à la base $\mathcal{B}_E$ est donnée par

$$A = \mathrm{Mat}(f; \mathcal{B}_E) = \begin{array}{cc} \begin{array}{cc} f(\vec{i}) & f(\vec{j}) \end{array} & \\ \begin{pmatrix} 2 & 1 \\ 3 & 1 \end{pmatrix} & \begin{array}{c} \vec{i} \\ \vec{j} \end{array} \end{array}.$$

Montrons que f est un automorphisme de E et déterminons son automorphisme réciproque.

— Une première méthode consiste à résoudre l'équation d'inconnue (x, y), $f(x, y) = (x', y')$, où (x', y') est un vecteur de E donné. On a

$$\left\{ \begin{array}{rcl} 2x + y & = & x' \\ 3x + y & = & y' \end{array} \right.,$$

ce système admet une unique solution qui est donnée par $(x, y) = (-x' + y', 3x' - 2y')$. Ainsi, l'application linéaire f est bijective, son application réciproque est

$$f^{-1} : \begin{array}{ccc} E & \longrightarrow & E \\ (x, y) & \longmapsto & (-x + y, 3x - 2y) \end{array}$$

et sa matrice relativement à la base $\mathcal{B}_E$ est

$$A^{-1} = \mathrm{Mat}(f^{-1};\mathcal{B}_E) = \begin{array}{cc} \begin{array}{cc} f^{-1}(\vec{\imath}) & f^{-1}(\vec{\jmath}) \end{array} & \\ \begin{pmatrix} -1 & 1 \\ 3 & -2 \end{pmatrix} & \begin{array}{c} \vec{\imath} \\ \vec{\jmath} \end{array} \end{array}.$$

— Une deuxième méthode consiste à exprimer $f^{-1}(\vec{\imath})$ et $f^{-1}(\vec{\jmath})$ en fonction de $\vec{\imath}$ et $\vec{\jmath}$. On part du système d'équations vectorielles

$$\left\{\begin{array}{rcl} f(\vec{\imath}) & = & 2\vec{\imath}+3\vec{\jmath} \\ f(\vec{\jmath}) & = & \vec{\imath}+\vec{\jmath} \end{array}\right.$$

dont la solution est

$$\left\{\begin{array}{rcrcl} \vec{\imath} & = & -f(\vec{\imath})+3f(\vec{\jmath}) & = & f(-\vec{\imath}+3\vec{\jmath}) \\ \vec{\jmath} & = & f(\vec{\imath})+2f(\vec{\jmath}) & = & f(\vec{\imath}+2\vec{\jmath}) \end{array}\right.,$$

ce qui fournit finalement, en admettant que f est bijective,

$$\left\{\begin{array}{rcl} f^{-1}(\vec{\imath}) & = & -\vec{\imath}+3\vec{\jmath} \\ f^{-1}(\vec{\jmath}) & = & \vec{\imath}+2\vec{\jmath} \end{array}\right.$$

et l'on retrouve la matrice de f^{-1} précédemment obtenue.

1.4 Changement de bases

Étant donnée une application linéaire, il existe parfois des bases dans lesquelles elle admet une matrice particulièrement simple alors que ce n'était peut-être pas le cas dans une base initialement définie. C'est le cas notamment des projections et des symétries vectorielles, et plus généralement des endomorphismes diagonalisables qui seront étudiés dans ce cours ultérieurement. Pour ces applications linéaires-là, il existe en effet des bases dans lesquelles leur matrice est diagonale, donc facilement exploitable au niveau des calculs. L'objectif de cette partie est de donner un procédé de transformation de matrice d'applications linéaires lors de changements de bases dans les espaces vectoriels de départ et d'arrivée de ces applications.

1.4.1 Changement de bases pour les vecteurs

Soit E un $\mathbb{K}$-espace vectoriel de dimension n. On va examiner l'effet d'un changement de bases sur les coordonnées d'un vecteur de E. On se donne deux bases de E :

$$\mathcal{B}_E = \{\vec{e}_1,\ldots,\vec{e}_n\} \quad \text{et} \quad \mathcal{B}'_E = \{\vec{e}'_1,\ldots,\vec{e}'_n\}.$$

Dans cette situation, la base $\mathcal{B}_E$ est considérée comme une base de référence (ancienne base) et les vecteurs de la base $\mathcal{B}'_E$ (nouvelle base) sont exprimés à l'aide de ceux de $\mathcal{B}_E$:

$$\vec{e}'_j = \sum_{i=1}^{n} p_{ij}\vec{e}_i, \quad 1 \leqslant j \leqslant n.$$

On définit alors la *matrice de passage $P_{\mathcal{B}_E \to \mathcal{B}'_E}$ de la base $\mathcal{B}_E$ vers la base $\mathcal{B}'_E$* selon

$$P = P_{\mathcal{B}_E \to \mathcal{B}'_E} = \begin{array}{c} \begin{array}{cccc} \vec{e}\,'_1 & \vec{e}\,'_2 & \dots & \vec{e}\,'_n \end{array} \\ \begin{pmatrix} p_{11} & p_{12} & & p_{1n} \\ p_{21} & p_{22} & & p_{2n} \\ \vdots & \vdots & \dots & \vdots \\ p_{n1} & p_{n2} & & p_{nn} \end{pmatrix} \end{array} \begin{array}{c} \vec{e}_1 \\ \vec{e}_2 \\ \vdots \\ \vec{e}_n \end{array}.$$

En étendant la définition du produit matriciel à des matrices-lignes de vecteurs, on pourrait écrire la relation entre les deux bases $\mathcal{B}_E$ et $\mathcal{B}'_E$ comme suit :

$$\begin{pmatrix} \vec{e}\,'_1 & \vec{e}\,'_2 & \dots & \vec{e}\,'_n \end{pmatrix} = \begin{pmatrix} \vec{e}_1 & \vec{e}_2 & \dots & \vec{e}_n \end{pmatrix} P_{\mathcal{B}_E \to \mathcal{B}'_E}.$$

Examinons comment se transforment les coordonnées d'un vecteur lors de ce changement de bases. Soit $\vec{u} \in E$, $x_1, \dots, x_n$ ses coordonnées relativement à la base $\mathcal{B}_E$ et $x'_1, \dots, x'_n$ ses coordonnées relativement à la base $\mathcal{B}'_E$. On a d'une part $\vec{u} = \displaystyle\sum_{i=1}^{n} x_i \vec{e}_i$ et d'autre part

$$\vec{u} = \sum_{j=1}^{n} x'_j \vec{e}\,'_j = \sum_{j=1}^{n} x'_j \left(\sum_{i=1}^{n} p_{ij} \vec{e}_i \right) = \sum_{i=1}^{n} \left(\sum_{j=1}^{n} p_{ij} x'_j \right) \vec{e}_i,$$

ce qui fournit des relations entre les nouvelles coordonnées $x'_1, \dots, x'_n$ et les anciennes $x_1, \dots, x_n$:

$$x_i = \sum_{j=1}^{n} p_{ij} x'_j, \quad 1 \leqslant i \leqslant n.$$

En introduisant les matrices-colonnes des coordonnées du vecteur $\vec{u}$ relativement aux bases $\mathcal{B}_E$ et $\mathcal{B}'_E$,

$$X = \operatorname{Mat}(\vec{u}; \mathcal{B}_E) = \begin{array}{c} \vec{u} \\ \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \end{array} \begin{array}{c} \vec{e}_1 \\ \vec{e}_2 \\ \vdots \\ \vec{e}_n \end{array} \quad \text{et} \quad X' = \operatorname{Mat}(\vec{u}; \mathcal{B}'_E) = \begin{array}{c} \vec{u} \\ \begin{pmatrix} x'_1 \\ x'_2 \\ \vdots \\ x'_n \end{pmatrix} \end{array} \begin{array}{c} \vec{e}\,'_1 \\ \vec{e}\,'_2 \\ \vdots \\ \vec{e}\,'_n \end{array},$$

cela s'écrit

$$\boxed{X = P_{\mathcal{B}_E \to \mathcal{B}'_E} X'.}$$

Remarque.

Cette formule exprime les **anciennes** coordonnées en fonction des **nouvelles**.

La question du changement de bases inverse, c'est-à-dire de $\mathcal{B}'_E$ vers $\mathcal{B}_E$, se pose également. Notons tout d'abord que $P_{\mathcal{B}_E \to \mathcal{B}'_E} = \operatorname{Mat}(\mathrm{id}_E; \mathcal{B}'_E, \mathcal{B}_E)$ et que réciproquement, puisque bien sûr $\mathrm{id}_E^{-1} = \mathrm{id}_E$,

$$\operatorname{Mat}(\mathrm{id}_E; \mathcal{B}_E, \mathcal{B}'_E) = \operatorname{Mat}(\mathrm{id}_E^{-1}; \mathcal{B}_E, \mathcal{B}'_E) = \operatorname{Mat}(\mathrm{id}_E; \mathcal{B}'_E, \mathcal{B}_E)^{-1}$$

soit encore

$$\boxed{P_{\mathcal{B}'_E \to \mathcal{B}_E} = (P_{\mathcal{B}_E \to \mathcal{B}'_E})^{-1}.}$$

Ainsi, les matrices de passage de la base $\mathcal{B}_E$ vers la base $\mathcal{B}'_E$ et de la base $\mathcal{B}'_E$ vers la base $\mathcal{B}_E$ sont inverses l'une de l'autre.

1.4.2 Changement de bases pour les applications linéaires

Soit maintenant E et F deux $\mathbb{K}$-espaces vectoriels de dimensions respectives n et p et $f \in \mathcal{L}(E,F)$. On va examiner cette fois l'effet d'un changement de bases sur la matrice de f. On se donne deux bases de E :

$$\mathcal{B}_E = \{\vec{e}_1, \dots, \vec{e}_n\} \quad \text{et} \quad \mathcal{B}'_E = \{\vec{e}'_1, \dots, \vec{e}'_n\},$$

ainsi que deux bases de F :

$$\mathcal{B}_F = \{\vec{\varepsilon}_1, \dots, \vec{\varepsilon}_p\} \quad \text{et} \quad \mathcal{B}'_F = \{\vec{\varepsilon}'_1, \dots, \vec{\varepsilon}'_p\}.$$

On introduit les matrices de passage de $\mathcal{B}_E$ vers $\mathcal{B}'_E$ et de $\mathcal{B}_F$ vers $\mathcal{B}'_F$, puis les matrices de f relativement aux bases $\mathcal{B}_E$ et $\mathcal{B}_F$ et relativement aux bases $\mathcal{B}'_E$ et $\mathcal{B}'_F$:

$$P = P_{\mathcal{B}_E \to \mathcal{B}'_E}, \quad Q = P_{\mathcal{B}_F \to \mathcal{B}'_F}, \quad A = \mathrm{Mat}(f; \mathcal{B}_E, \mathcal{B}_F), \quad A' = \mathrm{Mat}(f; \mathcal{B}'_E, \mathcal{B}'_F).$$

En interprétant les matrices de passage en termes de matrices d'applications identités, à savoir

$$\begin{array}{ll} P = \mathrm{Mat}(\mathrm{id}_E; \mathcal{B}'_E, \mathcal{B}_E), & P^{-1} = \mathrm{Mat}(\mathrm{id}_E; \mathcal{B}_E, \mathcal{B}'_E), \\ Q = \mathrm{Mat}(\mathrm{id}_F; \mathcal{B}'_F, \mathcal{B}_F), & Q^{-1} = \mathrm{Mat}(\mathrm{id}_F; \mathcal{B}_F, \mathcal{B}'_F), \end{array}$$

on observe, en écrivant l'égalité triviale $f = \mathrm{id}_F \circ f \circ \mathrm{id}_E$, mais lue dans les bases ad hoc, que

$$\begin{aligned} A' &= \mathrm{Mat}(f; \mathcal{B}'_E, \mathcal{B}'_F) = \mathrm{Mat}(\mathrm{id}_F \circ f \circ \mathrm{id}_E; \mathcal{B}'_E, \mathcal{B}'_F) \\ &= \mathrm{Mat}(\mathrm{id}_F; \mathcal{B}_F, \mathcal{B}'_F) \times \mathrm{Mat}(f; \mathcal{B}_E, \mathcal{B}_F) \times \mathrm{Mat}(\mathrm{id}_E; \mathcal{B}'_E, \mathcal{B}_E) = Q^{-1}AP. \end{aligned}$$

De la même manière, on a réciproquement

$$\begin{aligned} A &= \mathrm{Mat}(f; \mathcal{B}_E, \mathcal{B}_F) = \mathrm{Mat}(\mathrm{id}_F \circ f \circ \mathrm{id}_E; \mathcal{B}_E, \mathcal{B}_F) \\ &= \mathrm{Mat}(\mathrm{id}_F; \mathcal{B}'_F, \mathcal{B}_F) \times \mathrm{Mat}(f; \mathcal{B}'_E, \mathcal{B}'_F) \times \mathrm{Mat}(\mathrm{id}_E; \mathcal{B}_E, \mathcal{B}'_E) = QA'P^{-1}. \end{aligned}$$

On a donc obtenu

$$\boxed{A' = Q^{-1}AP \quad \text{et} \quad A = QA'P^{-1}.}$$

Les schémas ci-dessous illustrent ces relations.

$$\begin{array}{ccc} (E, \mathcal{B}_E) & \xrightarrow[A]{f} & (F, \mathcal{B}_F) \\ {\scriptstyle \mathrm{id}_E}\big\uparrow{\scriptstyle P} & & {\scriptstyle Q^{-1}}\big\downarrow{\scriptstyle \mathrm{id}_F} \\ (E, \mathcal{B}'_E) & \xrightarrow[f]{A'} & (F, \mathcal{B}'_F) \end{array} \quad \text{et} \quad \begin{array}{ccc} (E, \mathcal{B}_E) & \xrightarrow[A]{f} & (F, \mathcal{B}_F) \\ {\scriptstyle \mathrm{id}_E}\big\downarrow{\scriptstyle P^{-1}} & & {\scriptstyle Q}\big\uparrow{\scriptstyle \mathrm{id}_F} \\ (E, \mathcal{B}'_E) & \xrightarrow[f]{A'} & (F, \mathcal{B}'_F) \end{array}.$$

2 Image, noyau, rang

2.1 Rappels : injections, surjections, bijections

Soit $\mathcal{E}$ et $\mathcal{F}$ deux ensembles et $f : \mathcal{E} \longrightarrow \mathcal{F}$ une application.

— l'application f est *injective* (ou une *injection*) lorsque tout élément de $\mathcal{F}$ admet **au plus** un antécédent dans $\mathcal{E}$ par f. Ceci signifie encore que deux éléments distincts de $\mathcal{E}$ ont des images distinctes dans $\mathcal{F}$ par f et s'écrit mathématiquement

$$\forall x, x' \in \mathcal{E},\ [x \neq y \Longrightarrow f(x) \neq f(x')],$$

ce qui est aussi équivalent, par contraposition, à

$$\forall x, x' \in \mathcal{E},\ [f(x) = f(x') \Longrightarrow x = x'].$$

— l'application f est *surjective* (ou une *surjection*) lorsque tout élément de $\mathcal{F}$ admet **au moins** un antécédent dans $\mathcal{E}$ par f. Ceci s'écrit mathématiquement

$$\forall y \in \mathcal{F}, \exists x \in \mathcal{E},\ y = f(x).$$

— l'application f est *bijective* (ou une *bijection*) lorsque tout élément de $\mathcal{F}$ admet **exactement** un antécédent dans $\mathcal{E}$ par f. Ceci s'écrit mathématiquement

$$\forall y \in \mathcal{F}, \exists! x \in \mathcal{E},\ y = f(x).$$

On note $f^{-1} : B \longrightarrow A$ la bijection réciproque. Elle est caractérisée par l'équivalence

$$\forall x \in \mathcal{E}, \forall y \in \mathcal{F},\ [y = f(x) \Longleftrightarrow x = f^{-1}(y)].$$

Soit A une partie de $\mathcal{E}$ et B une partie de $\mathcal{F}$. On définit l'*image (directe)* de A et l'*image réciproque* de B par

$$f(A) = \{f(x),\ x \in A\} \quad \text{et} \quad f^{-1}(B) = \{x \in \mathcal{E} : f(x) \in B\}.$$

Lorsque f est bijective, on a $f^{-1}(B) = \{f^{-1}(y),\ y \in B\}$ et en particulier $f^{-1}(\{y\}) = \{f^{-1}(y)\}$. Remarquons que dans cette égalité, la notation f^{-1} a un double sens : $f^{-1}(\{y\})$ représente un ensemble, celui des antécédents de y par f et ne nécessite par la bijectivité de f, alors que $f^{-1}(y)$ désigne l'(unique) antécédent de y par f et nécessite la bijectivité de f.

2.2 Image et noyau d'une application linéaire

Dans le cas des applications linéaires entre deux espaces vectoriels E et F, les notions précédentes peuvent s'étudier d'une manière spécifique à la structure linéaire. Les ensembles définis ci-dessous sont d'une importance capitale dans cette étude.

Définition 2.1 *Soit $f \in \mathcal{L}(E, F)$.*

— *L'image de f est le sous-ensemble $f(E)$ de F. Il est noté* $\operatorname{Im} f$ *:*

$$\operatorname{Im} f = \{f(\vec{u}),\ \vec{u} \in E\}.$$

— *Le noyau de f est le sous-ensemble $f^{-1}(\{\vec{0}_F\})$ de E. Il est noté* $\operatorname{Ker} f$ *:*

$$\operatorname{Ker} f = \{\vec{u} \in E : f(\vec{u}) = \vec{0}_F\}.$$

Proposition 2.2 *Soit $f \in \mathcal{L}(E, F)$.*

1. *L'ensemble* $\operatorname{Ker} f$ *est un sous-espace vectoriel de E.*
2. *L'ensemble* $\operatorname{Im} f$ *est un sous-espace vectoriel de F. Lorsque E est de dimension finie, si $\mathcal{B}_E$ est une base de E, alors le sous-espace vectoriel* $\operatorname{Im} f$ *est engendré par le système de vecteurs $f(\mathcal{B}_E)$:*
$$\operatorname{Im} f = \operatorname{vect}(f(\mathcal{B}_E))$$
et en particulier $\dim(\operatorname{Im} f) \leqslant \dim(E)$.

Démonstration.

1. Soit $\alpha, \beta \in \mathbb{K}$ et $\vec{u}, \vec{v} \in \operatorname{Ker} f$. Puisque f est linéaire et $f(\vec{u}) = f(\vec{v}) = \vec{0}_F$, on a
$$f(\alpha\vec{u} + \beta\vec{v}) = \alpha f(\vec{u}) + \beta f(\vec{v}) = \vec{0}_F$$
et donc $\alpha\vec{u} + \beta\vec{v} \in \operatorname{Ker} f$. Ceci montre que $\operatorname{Ker} f$ est un sous-espace vectoriel de E.
2. Soit $\alpha, \beta \in \mathbb{K}$ et $\vec{u}, \vec{v} \in \operatorname{Im} f$. Il existe alors des vecteurs $\vec{u}_1, \vec{v}_1 \in E$ tels que $f(\vec{u}_1) = \vec{u}$ et $f(\vec{v}_1) = \vec{v}$. Puisque f est linéaire, on a
$$\alpha\vec{u} + \beta\vec{v} = \alpha f(\vec{u}_1) + \beta f(\vec{v}_1) = f(\alpha\vec{u}_1 + \beta\vec{v}_1)$$
et donc $\alpha\vec{u} + \beta\vec{v} \in \operatorname{Im} f$. Ainsi, $\operatorname{Im} f$ est un sous-espace vectoriel de F.
Si de plus E est de dimension finie et $\mathcal{B}_E = \{\vec{e}_1, \ldots, \vec{e}_n\}$ est une base de E, tout vecteur $\vec{u}$ de E s'écrit sous la forme $\vec{u} = \displaystyle\sum_{i=1}^{n} x_i\vec{e}_i$, donc $f(\vec{u}) = \displaystyle\sum_{i=1}^{n} x_i f(\vec{e}_i)$, et alors
$$\begin{aligned}\operatorname{Im} f = \{f(\vec{u}),\ \vec{u} \in E\} &= \left\{\sum_{i=1}^{n} x_i f(\vec{e}_i),\ \ x_1, \ldots, x_n \in \mathbb{K}\right\}\\ &= \operatorname{vect}(\{f(\vec{e}_1), \ldots, f(\vec{e}_n)\}) = \operatorname{vect}(f(\mathcal{B}_E)).\end{aligned}$$
De là découle
$$\dim(\operatorname{Im} f) \leqslant \operatorname{card}(f(\mathcal{B}_E)) \leqslant \operatorname{card}(\mathcal{B}_E) = \dim(E).$$
□

Proposition 2.3 *Soit $f \in \mathcal{L}(E, F)$.*

1. *f est injective si et seulement si* $\operatorname{Ker} f = \{\vec{0}_E\}$.
2. *f est surjective si et seulement si* $\operatorname{Im} f = F$.

Démonstration.

1. Supposons f injective. On a, puisque $f(\vec{0}_E) = \vec{0}_F$,
$$\vec{u} \in \operatorname{Ker} f \iff f(\vec{u}) = f(\vec{0}_E) \iff \vec{u} = \vec{0}_E,$$
et donc $\operatorname{Ker} f = \{\vec{0}_E\}$. Réciproquement, supposons que $\operatorname{Ker} f = \{\vec{0}_E\}$. Soit $\vec{u}, \vec{v}$ des vecteurs de E tels que $f(\vec{u}) = f(\vec{v})$. On obtient par linéarité $f(\vec{u}) - f(\vec{v}) = f(\vec{u} - \vec{v}) = \vec{0}_F$. Donc $\vec{u} - \vec{v} \in \operatorname{Ker} f$, ce qui donne, compte tenu de l'hypothèse, $\vec{u} - \vec{v} = \vec{0}_E$, soit encore $\vec{u} = \vec{v}$. Cela prouve l'injectivité de f.
2. La seconde assertion n'est rien d'autre que la définition de la surjectivité et est indépendante du contexte vectoriel. □

Exemple.

Considérons dans les $\mathbb{R}$-espaces vectoriels $E = \mathbb{R}^3$ et $F = \mathbb{R}^4$ l'application linéaire $f \in \mathcal{L}(E, F)$ définie par

$$f(x, y, z) = (x + 2y + 4z, -x + 5y + 6z, 3x - y + 2z, -3x + 8y + 8z).$$

Sa matrice relativement aux bases canoniques $\mathcal{B}_E = \{\vec{\imath}, \vec{\jmath}, \vec{k}\}$ de E et $\mathcal{B}_F = \{\vec{e}_1, \vec{e}_2, \vec{e}_3, \vec{e}_4\}$ de F a pour expression

$$\begin{array}{cc} & \begin{array}{ccc} f(\vec{\imath}) & f(\vec{\jmath}) & f(\vec{k}) \end{array} \\ A = & \begin{pmatrix} 1 & 2 & 4 \\ -1 & 5 & 6 \\ 3 & -1 & 2 \\ -3 & 8 & 8 \end{pmatrix} \begin{array}{c} \vec{e}_1 \\ \vec{e}_2 \\ \vec{e}_3 \\ \vec{e}_4 \end{array} . \end{array}$$

— *Détermination du noyau :* soit $\vec{u} = (x, y, z)$ un vecteur générique de E. On a

$$\vec{u} \in \operatorname{Ker} f \iff f(\vec{u}) = \vec{0} \iff \left\{ \begin{array}{rcl} x + 2y + 4z &=& 0 \\ -x + 5y + 6z &=& 0 \\ 3x - y + 2z &=& 0 \\ -3x + 8y + 8z &=& 0 \end{array} \right. .$$

Des deux premières équations, on tire $7y + 10z = 0$, ou encore $z = -\dfrac{7}{10}\,y$, que l'on injecte dans les deux autres équations. Cela conduit à la seule équation $x = \dfrac{4}{5}\,y$, laquelle fournit, avec la relation $z = -\dfrac{7}{10}\,y$, la représentation paramétrique

$$\left\{ \begin{array}{rcl} x &=& 8\lambda \\ y &=& 10\lambda \\ z &=& -7\lambda \end{array} \right. ,\ \lambda \in \mathbb{R}.$$

Ainsi, $\operatorname{Ker} f$ est la droite vectorielle engendrée par le vecteur $(8, 10, -7)$:

$$\operatorname{Ker} f = \operatorname{vect}(\{(8, 10, -7)\}).$$

Ce noyau n'étant pas réduit au vecteur nul, l'application linéaire f n'est pas injective.

— *Détermination de l'image :* l'image de f est le sous-espace vectoriel de F engendré par les images des vecteurs de la base canonique de E par exemple :

$$\operatorname{Im} f = \operatorname{vect}(\{f(\vec{\imath}), f(\vec{\jmath}), f(\vec{k})\}),$$

les vecteurs $f(\vec{\imath}), f(\vec{\jmath}), f(\vec{k})$, que l'on notera respectivement $\vec{u}, \vec{v}, \vec{w}$, étant donnés par les colonnes de A :

$$\vec{u} = (1, -1, 3, -3), \vec{v} = (2, 5, -1, 8), \vec{w} = (4, 6, 2, 8).$$

Afin de déterminer au mieux $\operatorname{Im} f$, on commence par évaluer le rang du système de vecteurs $\{\vec{u}, \vec{v}, \vec{w}\}$ à l'aide de la méthode des zéros échelonnés : le système $\{\vec{u}, \vec{v}, \vec{w}\}$ a même rang que le système $\{\vec{u}, \vec{v}\,', \vec{w}\,'\}$ où les vecteurs $\vec{v}\,'$ et $\vec{w}\,'$ sont donnés par $\vec{v}\,' = \vec{v} - 2\vec{u}$, $\vec{w}\,' = \vec{w} - 4\vec{u}$, soit encore, sous forme matricielle,

$$\begin{array}{c} \begin{array}{ccc} \vec{u} & \vec{v}\,' & \vec{w}\,' \end{array} \\ \begin{pmatrix} 1 & 0 & 0 \\ -1 & 7 & 10 \\ 3 & -7 & -10 \\ -3 & 14 & 20 \end{pmatrix} \end{array} \begin{array}{c} \\ \vec{e}_1 \\ \vec{e}_2 \\ \vec{e}_3 \\ \vec{e}_4 \end{array},$$

lequel a même rang que le système libre $\{\vec{u}, \vec{v}\,'\}$, puisque $7\vec{w}\,' - 10\vec{v}\,' = \vec{0}$. En conclusion, $\operatorname{Im} f$ est le plan vectoriel engendré par les vecteurs $\vec{u}, \vec{v}\,'$, ou encore $\vec{u}, \vec{v}$:

$$\operatorname{Im} f = \operatorname{vect}(\{(1, -1, 3, -3), (2, 5, -1, 8)\}).$$

Puisque $\operatorname{Im} f \neq F$, l'application linéaire f n'est pas surjective. Notons que la relation $7\vec{w}\,' - 10\vec{v}\,' = 0$ entraîne la relation $8\vec{u} + 10\vec{v} - 7\vec{w} = \vec{0}$ laquelle s'écrit encore $f(8\vec{\imath} + 10\vec{\jmath} - 7\vec{k}) = \vec{0}$. Cette dernière met en évidence un vecteur particulier non nul du noyau de f, à savoir $8\vec{\imath} + 10\vec{\jmath} - 7\vec{k}$ que l'on avait déjà obtenu directement.

2.3 Image d'un système de vecteurs

Proposition 2.4 *Soit $f \in \mathcal{L}(E, F)$ et $\mathcal{S}$, $\mathcal{L}$ et $\mathcal{G}$ des systèmes de vecteurs de E.*

1. *Si $\mathcal{S}$ est un système lié, alors $f(\mathcal{S})$ est un système de vecteurs de F lié. Par contraposition, si le système $f(\mathcal{S})$ est libre, alors le système $\mathcal{S}$ est libre.*
2. *Si $\mathcal{L}$ est un système libre et si f est injective, alors $f(\mathcal{L})$ est un système de vecteurs de F libre.*
3. *Si $\mathcal{G}$ est un système générateur de E et si f est surjective, alors $f(\mathcal{G})$ est un système générateur de F.*

Démonstration.

1. Soit $\mathcal{S} = \{\vec{u}_1, \dots, \vec{u}_n\}$ un système de E lié. Il existe des scalaires non tous nuls $\alpha_1, \dots, \alpha_n$ tels que $\alpha_1\vec{u}_1 + \dots + \alpha_n\vec{u}_n = \vec{0}_E$ et par linéarité on a $\alpha_1 f(\vec{u}_1) + \dots + \alpha_n f(\vec{u}_n) = \vec{0}_F$. Le système $f(\mathcal{S}) = \{f(\vec{u}_1), \dots, f(\vec{u}_n)\}$ de F est donc lié.
2. Supposons f injective. Soit $\mathcal{L} = \{\vec{u}_1, \dots, \vec{u}_n\}$ un système de E libre. Soit $\alpha_1, \dots, \alpha_n$ des scalaires tels que $\alpha_1 f(\vec{u}_1) + \dots + \alpha_n f(\vec{u}_n) = \vec{0}_F$. On a alors $f(\alpha_1\vec{u}_1 + \dots + \alpha_n\vec{u}_n) = \vec{0}_F$, c'est-à-dire $\alpha_1\vec{u}_1 + \dots + \alpha_n\vec{u}_n \in \operatorname{Ker} f$. Comme f est injective, $\operatorname{Ker} f = \{\vec{0}_E\}$ et donc $\alpha_1\vec{u}_1 + \dots + \alpha_n\vec{u}_n = \vec{0}_E$, soit encore, puisque le système $\mathcal{L}$ est libre, $\alpha_1 = \dots = \alpha_n = 0$, ce qui montre que le système $f(\mathcal{L}) = \{f(\vec{u}_1), \dots, f(\vec{u}_n)\}$ de F est libre.
3. Supposons f surjective. Soit $\mathcal{G} = \{\vec{u}_1, \dots, \vec{u}_n\}$ un système générateur de E. Soit $\vec{v} \in F$. Par surjectivité, il existe un vecteur $\vec{u} \in E$ tel que $\vec{v} = f(\vec{u})$. Le système $\mathcal{G}$ étant générateur de E, il existe des scalaires $\alpha_1, \dots, \alpha_n$ tels que $\vec{u} = \alpha_1\vec{u}_1 + \dots + \alpha_n\vec{u}_n$. On a alors $\vec{v} = \alpha_1 f(\vec{u}_1) + \dots + \alpha_n f(\vec{u}_n)$ ce qui montre que le système $f(\mathcal{G}) = \{f(\vec{u}_1), \dots, f(\vec{u}_n)\}$ est générateur de F. $\square$

Proposition 2.5 *Supposons E et F de dimension finie. Soit $f \in \mathcal{L}(E,F)$.*

1. *Si $\mathcal{B}_E$ est une base de E et si f est injective, alors $f(\mathcal{B}_E)$ est une base de $\operatorname{Im} f$. Pour que f soit injective, il est nécessaire que $\dim(E) \leqslant \dim(F)$.*
2. *Pour que f soit surjective, il est nécessaire que $\dim(E) \geqslant \dim(F)$.*
3. *Soit $\mathcal{B}_E$ une base de E.*
   - *(a). Si le système $f(\mathcal{B}_E)$ est libre, alors f est injective.*
   - *(b). Si le système $f(\mathcal{B}_E)$ est générateur de F, alors f est surjective.*
   - *(c). Si le système $f(\mathcal{B}_E)$ est une base de F, alors f est bijective.*

Démonstration.

1. Soit $\mathcal{B}_E$ une base de E. Supposons f injective. D'après la proposition précédente, $f(\mathcal{B}_E)$ est un système libre de F. Par ailleurs, il est générateur de $\operatorname{Im} f$, c'en est donc une base. D'autre part, les systèmes $\mathcal{B}_E$ et $f(\mathcal{B}_E)$ ont même nombre de vecteurs et donc
$$\dim(E) = \operatorname{card}(\mathcal{B}_E) = \operatorname{card}(f(\mathcal{B}_E)) = \dim(\operatorname{Im} f) \leqslant \dim F.$$
2. Soit $\mathcal{B}_F = \{\vec{\varepsilon}_1, \dots, \vec{\varepsilon}_p\}$ une base de F. Supposons f surjective. Il existe des vecteurs $\vec{u}_1, \dots, \vec{u}_p$ de E tels que $f(\vec{u}_1) = \vec{\varepsilon}_1, \dots, f(\vec{u}_p) = \vec{\varepsilon}_p$. Le système de vecteurs $\mathcal{B}_F = \{f(\vec{u}_1), \dots, f(\vec{u}_p)\}$ de F étant libre, le système de vecteurs $\mathcal{L} = \{\vec{u}_1, \dots, \vec{u}_p\}$ de E est également libre d'après le théorème précédent et les deux systèmes ont même nombre de vecteurs. Ainsi
$$\dim(E) \geqslant \operatorname{card}(\mathcal{L}) = \operatorname{card}(\mathcal{B}_F) = \dim F.$$
3. Soit $\mathcal{B}_E = \{\vec{e}_1, \dots, \vec{e}_n\}$ une base de E.
   - (a). Supposons le système $f(\mathcal{B}) = \{f(\vec{e}_1), \dots, f(\vec{e}_n)\}$ libre. Soit $\vec{u}$ un vecteur de E. Il existe des scalaires $x_1, \dots, x_n$ tels que $\vec{u}_1 = x_1\vec{e}_1 + \cdots + x_n\vec{e}_n$. On a
   $$\vec{u} \in \operatorname{Ker} f \iff f(\vec{u}) = \vec{0}_F \iff x_1 f(\vec{e}_1) + \cdots + x_n f(\vec{e}_n) = \vec{0}_F$$
   $$\iff x_1 = \cdots = x_n = 0 \iff \vec{u} = \vec{0}_E.$$
   On vient de montrer que $\operatorname{Ker} f = \{\vec{0}_E\}$ et f est donc injective.
   - (b). Supposons le système $f(\mathcal{B}) = \{f(\vec{e}_1), \dots, f(\vec{e}_n)\}$ générateur de F. On a donc $F = \operatorname{vect}(f(\mathcal{B}))$. Par ailleurs, on sait que $\operatorname{Im} f = \operatorname{vect}(f(\mathcal{B}))$, donc $\operatorname{Im} f = F$ et f est surjective.
   - (c). Ce dernier point découle immédiatement des deux précédents. □

Théorème 2.6 *1. Si E et F sont des $\mathbb{K}$-espaces vectoriels de dimension finie et si f est un isomorphisme de E sur F, on a $\dim(E) = \dim(F)$.*

2. Tout $\mathbb{K}$-espace vectoriel de dimension n est isomorphe à $\mathbb{K}^n$.

3. Tous les $\mathbb{K}$-espaces vectoriels de même dimension sont isomorphes.

Démonstration.

1. L'isomorphisme f est une application linéaire à la fois injective et surjective. D'après la proposition précédente, on en déduit que $\dim(E) \leqslant \dim(F)$ et $\dim(F) \leqslant \dim(E)$, d'où l'égalité des dimensions de E et F.

2. Soit E un $\mathbb{K}$-espace vectoriel de dimension n, $\mathcal{B}_E = \{\vec{e}_1, \ldots, \vec{e}_n\}$ une base de E et $\mathcal{B} = \{\vec{\varepsilon}_1, \ldots, \vec{\varepsilon}_n\}$ la base canonique du $\mathbb{K}$-espace vectoriel $\mathbb{K}^n$. Il est clair que l'application linéaire $\varphi_E : E \longrightarrow \mathbb{K}^n$ définie par $\varphi_E(\vec{e}_i) = \vec{\varepsilon}_i$ pour tout $i \in \{1, \ldots, n\}$ est un isomorphisme. Donc E et $\mathbb{K}^n$ sont isomorphes.
3. Soit E et F deux $\mathbb{K}$-espaces vectoriels de dimension n. Avec les notations précédentes, $\varphi_F^{-1} \circ \varphi_E$ définit un isomorphisme de E sur F et E et F sont donc isomorphes. $\square$

Exemples.

1. Les droites vectorielles de E sont isomorphes au corps de base $\mathbb{K}$ considéré comme $\mathbb{K}$-espace vectoriel.
2. Les hyperplans vectoriels de E sont isomorphes au $\mathbb{K}$-espace vectoriel $\mathbb{K}^{n-1}$.
3. Dans le $\mathbb{K}$-espace vectoriel $\mathbb{K}^3$ rapporté à sa base canonique $\{\vec{i}, \vec{j}, \vec{k}\}$, le sous-ensemble $P = \{(x, y, 0),\ x, y \in \mathbb{K}\}$ est le plan vectoriel engendré par $\{\vec{i}, \vec{j}\}$. Il est isomorphe au $\mathbb{K}$-espace vectoriel $\mathbb{K}^2$ via l'isomorphisme $(x, y, 0) \in P \longmapsto (x, y) \in \mathbb{K}^2$, mais ce n'est pas $\mathbb{K}^2$ lui-même puisque les ensembles $\mathbb{K}^2$ et $\mathbb{K}^3$ sont de natures différentes ($\mathbb{K}^2 \not\subset \mathbb{K}^3$).

2.4 Rang d'une application linéaire

On suppose ici que les espaces vectoriels E et F sont de dimension finie. Soit $\mathcal{B}$ une base de E et $f \in \mathcal{L}(E, F)$.

Définition 2.7 *Le rang de l'application linéaire f est la dimension de son image :*

$$\operatorname{rang}(f) = \dim(\operatorname{Im} f) = \operatorname{rang}(f(\mathcal{B})).$$

Remarques.

1. Cette définition ne dépend pas du choix de la base $\mathcal{B}$. En effet, si $\mathcal{B}'$ est une autre base de E, on a $\operatorname{Im} f = \operatorname{vect}(f(\mathcal{B})) = \operatorname{vect}(f(\mathcal{B}'))$ et donc $\operatorname{rang}(f(\mathcal{B})) = \operatorname{rang}(f(\mathcal{B}'))$.
2. On a immédiatement $\dim(\operatorname{Im} f) \leqslant \dim(F)$. Par ailleurs, on a $\operatorname{rang}(f(\mathcal{B})) = \operatorname{card}(f(\mathcal{B})) \leqslant \operatorname{card}(\mathcal{B}) = \dim(E)$. On obtient ainsi la majoration $\operatorname{rang}(f) \leqslant \min(\dim(E), \dim(F))$.

Théorème 2.8 (du rang) *Les dimensions des espaces vectoriels* $\operatorname{Ker} f$ *et* $\operatorname{Im} f$ *sont reliées par*

$$\dim(\operatorname{Ker} f) + \dim(\operatorname{Im} f) = \dim(E).$$

Démonstration.

On introduit un sous-espace vectoriel G supplémentaire de $\operatorname{Ker} f$ dans E :

$$E = \operatorname{Ker} f \oplus G,$$

puis la restriction g de f à G : $g = f|_G$. Montrons que g réalise un isomorphisme de G sur $\operatorname{Im} f$.

— L'application linéaire g est injective car

$$\operatorname{Ker} g = \operatorname{Ker} f \cap G = \{\vec{0}_E\}.$$

— L'application g a pour image $\operatorname{Im} f$. En effet, soit $\vec{v} \in \operatorname{Im} f$. Alors $\exists \vec{u} \in E,\ \vec{v} = f(\vec{u})$. D'autre part,
$$\exists (\vec{u}_1, \vec{u}_2) \in \operatorname{Ker} f \times G,\ \vec{u} = \vec{u}_1 + \vec{u}_2$$
d'où
$$\vec{v} = f(\vec{u}) = f(\vec{u}_1 + \vec{u}_2) = f(\vec{u}_1) + f(\vec{u}_2) = g(\vec{u}_2).$$
Ainsi $\vec{v} \in \operatorname{Im} g$ et $\operatorname{Im} f \subset \operatorname{Im} g$. L'inclusion inverse est immédiate par définition de g donc $\operatorname{Im} g = \operatorname{Im} f$.

En conclusion, $\dim(\operatorname{Im} f) = \dim G = \dim E - \dim(\operatorname{Ker} f)$ ce qui prouve le résultat annoncé. □

Exemple.

Reprenons l'exemple proposé à la fin du paragraphe 2.2 : pour l'application linéaire $f \in \mathcal{L}(\mathbb{R}^3, \mathbb{R}^4)$ définie par
$$f(x, y, z) = (x + 2y + 4z, -x + 5y + 6z, 3x - y + 2z, -3x + 8y + 8z),$$
on a trouvé que $\operatorname{Ker} f$ était une droite de $\mathbb{R}^3$ et $\operatorname{Im} f$ un plan de $\mathbb{R}^4$. Le théorème du rang est bien vérifié : $\dim(\operatorname{Ker} f) + \dim(\operatorname{Im} f) = 1 + 2 = 3 = \dim(\mathbb{R}^3)$. Notons d'ailleurs qu'avec cette information de dimension, la seule recherche de $\operatorname{Im} f$ qui conduisait à $\operatorname{Im} f = \operatorname{vect}(\{(1, -1, 3, -3), (2, 5, -1, 8)\})$, ainsi qu'à la relation $f(8\vec{i} + 10\vec{j} - 7\vec{k}) = \vec{0}$ assurant que $8\vec{i} + 10\vec{j} - 7\vec{k}$ est un vecteur particulier non nul du noyau de f, permet de conclure immédiatement (sans aucune autre recherche) que $\operatorname{Ker} f$ est exactement la droite vectorielle engendrée par le vecteur $8\vec{i} + 10\vec{j} - 7\vec{k}$.

Rappelons la majoration $\operatorname{rang}(f) \leqslant \min(\dim(E), \dim(F))$. Le théorème du rang donne des conditions pour que les égalités $\operatorname{rang}(f) = \dim(E)$ ou $\operatorname{rang}(f) = \dim(F)$ soient vérifiées.

Corollaire 2.9 *On a les équivalences suivantes :*
$$\operatorname{rang}(f) = \dim(E) \Longleftrightarrow f \text{ injective} \quad \text{et} \quad \operatorname{rang}(f) = \dim(F) \Longleftrightarrow f \text{ surjective.}$$

Démonstration.

1. D'après le théorème du rang, puisque $\operatorname{rang}(f) = \dim(\operatorname{Im} f)$, on a
$$\operatorname{rang}(f) = \dim(E) \Longleftrightarrow \dim(\operatorname{Ker} f) = 0 \Longleftrightarrow \operatorname{Ker} f = \{\vec{0}\} \Longleftrightarrow f \text{ injective.}$$
2. En ce qui concerne la deuxième équivalence, on a simplement
$$\operatorname{rang}(f) = \dim(F) \Longleftrightarrow \dim(\operatorname{Im} f) = \dim(F) \Longleftrightarrow \operatorname{Im} f = F \Longleftrightarrow f \text{ surjective.}$$ □

Corollaire 2.10 *Supposons que* $\dim(E) = \dim(F)$.
— *Si f est injective, alors elle est bijective.*
— *Si f est surjective, alors elle est bijective.*
En d'autres termes, on a les équivalences suivantes :
$$f \text{ injective} \Longleftrightarrow f \text{ surjective} \Longleftrightarrow f \text{ bijective.}$$

Démonstration.

1. Si f est injective, $\operatorname{Ker} f = \{\vec{0}_E\}$. Le théorème du rang donne $\dim(\operatorname{Im} f) = \dim E$, soit encore $\operatorname{Im} f = F$ et f est surjective.
2. Si f est surjective, $\operatorname{Im} f = F$. Le théorème du rang donne $\dim(\operatorname{Ker} f) = \dim E - \dim F = 0$, soit encore $\operatorname{Ker} f = \{\vec{0}_E\}$ et f est injective.
3. On a donc démontré l'équivalence
$$f \text{ injective} \Longleftrightarrow f \text{ surjective}. \qquad \square$$

Ce résultat est très important puisque, pour montrer qu'une application linéaire entre deux espaces vectoriels de même dimension finie est bijective, il **suffit** de vérifier l'une des deux qualités injectivité ou surjectivité. Par contre, il peut être mis en défaut en dimension infinie.

Contre-exemple.

Considérons dans le $\mathbb{R}$-espace vectoriel de dimension infinie $E = \mathbb{R}[X]$ les endomorphismes de E « dérivation » et « intégration » respectivement définis par $f(P(X)) = P'(X)$ et $g(P(X)) = \displaystyle\int_0^X P(x)\,dx$.

— On a clairement $\operatorname{Im} f = E$ puisque tout polynôme admet au moins pour antécédent une de ses primitives, et $\operatorname{Ker} f$ est la droite vectorielle des polynômes constants. Donc f est surjective mais pas injective.

— On a $\operatorname{Ker} g = \{0\}$ puisque, par dérivation, $\displaystyle\int_0^X P(x)\,dx = 0 \Longrightarrow P = 0$. Par ailleurs, puisque $\displaystyle\int_0^X P(x)\,dx$ est la primitive de P s'annulant en 0, $\operatorname{Im} g$ est le sous-espace vectoriel des polynômes de terme constant nul. Donc g est injective mais pas surjective.

— Notons enfin que $f \circ g = \mathrm{id}_E$ puisque $\left(\displaystyle\int_0^X P(x)\,dx\right)' = P(X)$. En revanche, on a $(g \circ f)(P(X)) = \displaystyle\int_0^X P'(x)\,dx = P(X) - P(0)$ et donc $g \circ f \neq \mathrm{id}_E$.

Exemples.

1. Considérons dans les $\mathbb{R}$-espaces vectoriels $E = \mathbb{R}_n[X]$ et $F = \mathbb{R}_{n-1}[X] \times \mathbb{R}$ l'application linéaire $f \in \mathcal{L}(E, F)$ définie par $f(P(X)) = (P'(X), P(0))$. Montrons que f est isomorphisme de E sur F. Comme $\dim(E) = \dim(F)$ $(= n+1)$, il suffit de vérifier que f est par exemple injective. À cet effet, on a
$$P(X) \in \operatorname{Ker} f \Longleftrightarrow [P'(X) = 0 \text{ et } P(0) = 0] \Longleftrightarrow P(X) = 0,$$
donc $\operatorname{Ker} f = \{\vec{0}\}$, ce qui assure l'injectivité de f puis sa bijectivité.
2. Reprenons la même application linéaire en choisissant pour ensembles de départ et d'arrivée les $\mathbb{R}$-espaces vectoriels $E = \mathbb{R}[X]$ et $F = \mathbb{R}[X] \times \mathbb{R}$. Les espaces vectoriels E et F étant à présent de dimension infinie, pour voir que f est un isomorphisme de E sur F, il ne suffit plus de vérifier seulement l'injectivité. Soit alors $(Q(X), a) \in F$ et cherchons ses antécédents $P(X)$ de $(Q(X), a)$ par f :
$$f(P(X)) = (Q(X), a) \Longleftrightarrow [P'(X) = Q(X) \text{ et } P(0) = a] \Longleftrightarrow P(X) = \int_0^X Q(x)\,dx + a$$
ce qui montre que $(Q(X), a)$ a un unique antécédent par f et que f est bien bijective. On peut également décrire son isomorphisme réciproque $f^{-1} \in \mathcal{L}(F, E)$: il est donné par $f^{-1}(Q(X), a) = \displaystyle\int_0^X Q(x)\,dx + a$.

3 Deux exemples géométriques : projections et symétries

Dans ce paragraphe, E est un $\mathbb{K}$-espace vectoriel de dimension finie et F et G sont des sous-espaces vectoriels supplémentaires de E : $E = F \oplus G$. Soit $\vec{w} \in E$. Alors il existe un unique couple de vecteurs $(\vec{u}, \vec{v}) \in F \times G$ tels que $\vec{w} = \vec{u} + \vec{v}$. Les vecteurs $\vec{u}$ et $\vec{v}$ sont respectivement les *projections* du vecteur $\vec{w}$ sur les sous-espaces vectoriels F et G parallèlement à G et F (voir Fig. 2.1).

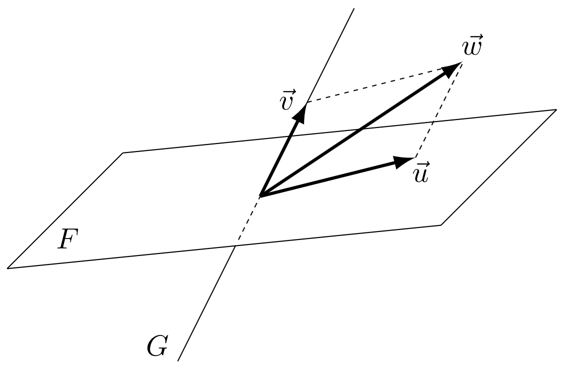

FIGURE 2.1 – Projections sur les sous-espaces vectoriels F et G

3.1 Projections vectorielles

Définition 3.1 *L'application* $\begin{array}{rccc} p: & E & \longrightarrow & E \\ & \vec{w} & \longmapsto & \vec{u} \end{array}$ *est la projection vectorielle de E sur F parallèlement G. De même, l'application* $\begin{array}{rccc} q: & E & \longrightarrow & E \\ & \vec{w} & \longmapsto & \vec{v} \end{array}$ *est la projection vectorielle de E sur G parallèlement F.*

Montrons que p et q sont des applications linéaires. Soit $\alpha, \beta \in \mathbb{K}$, $\vec{w}, \vec{w}\,' \in E$. Il existe des vecteurs $(\vec{u}, \vec{v}) \in F \times G$ et $(\vec{u}\,', \vec{v}\,') \in F \times G$ tels que $\vec{w} = \vec{u} + \vec{v}$ et $\vec{w}\,' = \vec{u}\,' + \vec{v}\,'$. Alors, $\alpha\vec{w} + \beta\vec{w}\,' = (\alpha\vec{u} + \beta\vec{u}\,') + (\alpha\vec{v} + \beta\vec{v}\,')$ avec $\alpha\vec{u} + \beta\vec{u}\,' \in F$ et $\alpha\vec{v} + \beta\vec{v}\,' \in G$. On a donc, par unicité,

$$p(\alpha\vec{w} + \beta\vec{w}\,') = \alpha\vec{u} + \beta\vec{u}\,' = \alpha p(\vec{w}) + \beta p(\vec{w}\,') \quad \text{et} \quad q(\alpha\vec{w} + \beta\vec{w}\,') = \alpha\vec{v} + \beta\vec{v}\,' = \alpha p(\vec{w}) + \beta p(\vec{w}\,'),$$

d'où l'on tire $p, q \in \mathcal{L}(E)$. Déterminons à présent les matrices des endomorphismes p et q relativement à une base $\mathcal{B}$ de E construite en réunissant une base $\mathcal{B}_F$ de F et une base $\mathcal{B}_G$ de G : $\mathcal{B} = \mathcal{B}_F \cup \mathcal{B}_G$. D'après les définitions de p et q, on a

$$\forall \vec{u} \in \mathcal{B}_F,\ p(\vec{u}) = \vec{u}, q(\vec{u}) = 0 \quad \text{et} \quad \forall \vec{v} \in \mathcal{B}_G,\ p(\vec{v}) = 0, q(\vec{v}) = \vec{v},$$

ce qui s'écrit en d'autres termes $p|_F = \mathrm{id}_F, p|_G = 0$, $q|_F = 0, q|_G = \mathrm{id}_G$. On obtient donc le résultat suivant.

Proposition 3.2 *Les projections vectorielles p et q sont des endomorphismes de E de matrices relativement à une base $\mathcal{B}$ de E de la forme $\mathcal{B}_F \cup \mathcal{B}_G$, où $\mathcal{B}_F$ est une base de F et $\mathcal{B}_G$ une base de G,*

$$\mathrm{Mat}(p;\mathcal{B}) = \left(\begin{array}{c|c} I_F & O \\ \hline O & O_G \end{array}\right) \quad \textit{et} \quad \mathrm{Mat}(q;\mathcal{B}) = \left(\begin{array}{c|c} O_F & O \\ \hline O & I_G \end{array}\right)$$

où par exemple I_F est la matrice unité d'ordre $\dim(F)$ *et O_G la matrice carrée nulle d'ordre* $\dim(G)$.

Détermination pratique d'une projection vectorielle.

Soit $\mathcal{B}_E$ une base de E. On souhaiterait déterminer analytiquement la projection vectorielle p relativement à la base $\mathcal{B}_E$.

— À l'aide d'un changement de bases, il est possible de calculer la matrice de p relativement à la base $\mathcal{B}_E$ à partir de la matrice relativement à $\mathcal{B}_F \cup \mathcal{B}_G$ qui est très simple :

$$\mathrm{Mat}(p;\mathcal{B}_E) = P_{(\mathcal{B}_E \to \mathcal{B}_F \cup \mathcal{B}_G)} \left(\begin{array}{c|c} I_F & O \\ \hline O & O_G \end{array}\right) P^{-1}_{(\mathcal{B}_E \to \mathcal{B}_F \cup \mathcal{B}_G)}.$$

La principale difficulté réside dans le calcul de l'inverse de la matrice de passage.

— À l'aide d'un système d'équations, il est possible de déterminer la représentation analytique de la projection vectorielle p relativement à la base $\mathcal{B}_E$: le vecteur image $\vec{u} = p(\vec{w})$ est caractérisé par les deux conditions

$$\vec{u} \in F \quad \text{et} \quad \vec{w} - \vec{u} \in G.$$

Si F et G sont donnés par des représentations paramétriques ou cartésiennes relativement à une base $\mathcal{B}_E$ de E, ces deux conditions conduisent à un système d'équations linéaires fournissant les coordonnées du vecteur $\vec{u}$ relativement à la base $\mathcal{B}_E$ à partir de celles du vecteur $\vec{w}$: c'est la représentation analytique de la projection p. La principale difficulté réside cette fois dans la résolution du système d'équations.

Exemple.

Soit dans le $\mathbb{R}$-espace vectoriel $E = \mathbb{R}^3$

— le plan vectoriel $F = \{(x,y,z) \in E : ax + by + cz = 0\}$;

— la droite vectorielle $G = \{(\lambda\alpha, \lambda\beta, \lambda\gamma),\ \lambda \in \mathbb{R}\}$.

On note $\mathcal{B}_E = \{\vec{i}, \vec{j}, \vec{k}\}$ la base canonique de E. Une base de F est donnée par $\mathcal{B}_F = \{(b,-a,0), (c,0,-a)\}$ et une base de G par $\mathcal{B}_G = \{(\alpha,\beta,\gamma)\}$. On a déjà vu que les sous-espaces vectoriels F et G sont supplémentaires dans E si et seulement si $a\alpha + b\beta + c\gamma \neq 0$, relation que l'on supposera dans la suite. On supposera également, pour simplifier la discussion, que $a \neq 0$. On peut donc définir la projection vectorielle p sur P parallèlement à G.

Première méthode : on détermine la matrice de la projection vectorielle p relativement à la base $\mathcal{B}_E$. Une base adaptée à la décomposition $E = F \oplus G$ est $\mathcal{B} = \mathcal{B}_F \cup \mathcal{B}_G = \{\vec{i}\,', \vec{j}\,', \vec{k}\,'\}$ avec $\vec{i}\,' = (b,-a,0), \vec{j}\,' = (c,0,-a), \vec{k}\,' = (\alpha,\beta,\gamma)$. Relativement à cette base, la matrice de p a la forme très simple

$$\mathrm{Mat}(p;\mathcal{B}) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

La matrice de passage de la base $\mathcal{B}$ vers la base $\mathcal{B}_E$ est

$$P = P_{\mathcal{B}\to\mathcal{B}_E} = \begin{pmatrix} b & c & \alpha \\ -a & 0 & \beta \\ 0 & -a & \gamma \end{pmatrix}.$$

On doit déterminer l'inverse de cette matrice. Rappelons que cet inverse n'est autre que la matrice de passage $P_{\mathcal{B}_E\to\mathcal{B}}$:

$$P^{-1}_{\mathcal{B}\to\mathcal{B}_E} = P_{\mathcal{B}_E\to\mathcal{B}}$$

et pour obtenir $P_{\mathcal{B}_E\to\mathcal{B}}$, on doit résoudre le système d'équations vectorielles

$$\left\{\begin{array}{l} b\vec{i} - a\vec{j} \phantom{+\gamma\vec{k}} = \vec{i}\,' \\ c\vec{i} \phantom{+\beta\vec{j}} - a\vec{k} = \vec{j}\,' \\ \alpha\vec{i} + \beta\vec{j} + \gamma\vec{k} = \vec{k}\,' \end{array}\right. .$$

Des deux premières équations, on tire les vecteurs $\vec{j}$ et $\vec{k}$ en fonction de $\vec{i}$:

$$\vec{j} = \frac{b}{a}\,\vec{i} - \frac{1}{a}\,\vec{i}\,', \quad \vec{k} = \frac{c}{a}\,\vec{i} - \frac{1}{a}\,\vec{j}\,'$$

et en reportant ces derniers dans la troisième équation, on obtient le vecteur $\vec{i}$ duquel on déduit ensuite $\vec{j}$ et $\vec{k}$:

$$\left\{\begin{array}{lcl} \vec{i} & = & \dfrac{1}{a\alpha+b\beta+c\gamma}\,(\beta\vec{i}\,' + \gamma\vec{j}\,' + a\vec{k}\,') \\ \vec{j} & = & \dfrac{1}{a(a\alpha+b\beta+c\gamma)}\,(-(a\alpha+c\gamma)\vec{i}\,' + b\gamma\vec{j}\,' + ab\vec{k}\,') \\ \vec{k} & = & \dfrac{1}{a(a\alpha+b\beta+c\gamma)}\,(c\beta\vec{i}\,' - (a\alpha+b\beta)\vec{j}\,' + ac\vec{k}\,') \end{array}\right.$$

d'où

$$P^{-1} = \begin{pmatrix} \dfrac{\beta}{a\alpha+b\beta+c\gamma} & -\dfrac{a\alpha+c\gamma}{a(a\alpha+b\beta+c\gamma)} & \dfrac{c\beta}{a(a\alpha+b\beta+c\gamma)} \\ \dfrac{\gamma}{a\alpha+b\beta+c\gamma} & \dfrac{b\gamma}{a(a\alpha+b\beta+c\gamma)} & -\dfrac{a\alpha+b\beta}{a(a\alpha+b\beta+c\gamma)} \\ \dfrac{a}{a\alpha+b\beta+c\gamma} & \dfrac{ab}{a(a\alpha+b\beta+c\gamma)} & \dfrac{ac}{a(a\alpha+b\beta+c\gamma)} \end{pmatrix}.$$

On peut maintenant terminer le calcul de la matrice de p relativement à la base $\mathcal{B}$:

$$\begin{aligned}
\mathrm{Mat}(p;\mathcal{B}_E,\mathcal{B}_E) &= P\begin{pmatrix}1&0&0\\0&1&0\\0&0&0\end{pmatrix}P^{-1}\\
&= \begin{pmatrix}b&c&0\\-a&0&0\\0&-a&0\end{pmatrix}\begin{pmatrix}\dfrac{\beta}{a\alpha+b\beta+c\gamma} & -\dfrac{a\alpha+c\gamma}{a(a\alpha+b\beta+c\gamma)} & \dfrac{c\beta}{a(a\alpha+b\beta+c\gamma)}\\ \dfrac{\gamma}{a\alpha+b\beta+c\gamma} & \dfrac{b\gamma}{a(a\alpha+b\beta+c\gamma)} & -\dfrac{a\alpha+b\beta}{a(a\alpha+b\beta+c\gamma)}\\ \dfrac{a}{a\alpha+b\beta+c\gamma} & \dfrac{ab}{a(a\alpha+b\beta+c\gamma)} & \dfrac{ac}{a(a\alpha+b\beta+c\gamma)}\end{pmatrix}\\
&= \begin{pmatrix}\dfrac{b\beta+c\gamma}{a\alpha+b\beta+c\gamma} & -\dfrac{b\alpha}{a\alpha+b\beta+c\gamma} & -\dfrac{c\alpha}{a\alpha+b\beta+c\gamma}\\ -\dfrac{a\beta}{a\alpha+b\beta+c\gamma} & \dfrac{a\alpha+c\gamma}{a\alpha+b\beta+c\gamma} & -\dfrac{c\beta}{a\alpha+b\beta+c\gamma}\\ -\dfrac{a\gamma}{a\alpha+b\beta+c\gamma} & -\dfrac{b\gamma}{a\alpha+b\beta+c\gamma} & \dfrac{a\alpha+b\beta}{a\alpha+b\beta+c\gamma}\end{pmatrix}.
\end{aligned}$$

Deuxième méthode : on détermine la représentation analytique de la projection vectorielle p relativement à la base $\mathcal{B}_E$. Soit $\vec{w}=(x,y,z)$ un vecteur de E et $\vec{u}=p(\vec{w})=(x',y',z')$ son image par p. On sait que le vecteur $\vec{u}$ est caractérisé par les conditions ci-dessous :

$$\begin{aligned}
\vec{u}=p(\vec{w}) &\iff (x',y',z')\in F \quad\text{et}\quad (x'-x,y'-y,z'-z)\in G\\
&\iff ax'+by'+cz'=0 \quad\text{et}\quad [\exists\lambda\in\mathbb{R},\ x'-x=\lambda\alpha,\ y'-y=\lambda\beta,\ z'-z=\lambda\gamma]\\
&\iff \exists\lambda\in\mathbb{R},\ x'=x+\lambda\alpha,\ y'=y+\lambda\beta,\ z'=z+\lambda\gamma\\
&\qquad\quad \text{et}\quad (ax+by+cz)+\lambda(a\alpha+b\beta+c\gamma)=0
\end{aligned}$$

ce qui donne pour λ la valeur $\lambda=-\dfrac{ax+by+cz}{a\alpha+b\beta+c\gamma}$, puis

$$\left\{\begin{aligned}
x' &= \frac{(b\beta+c\gamma)x-b\alpha y-c\alpha z}{a\alpha+b\beta+c\gamma}\\
y' &= \frac{-a\beta x+(a\alpha+c\gamma)y-c\beta z}{a\alpha+b\beta+c\gamma}\\
z' &= \frac{-a\gamma x-b\gamma y+(a\alpha+b\beta)z}{a\alpha+b\beta+c\gamma}
\end{aligned}\right. .$$

Énonçons maintenant quelques propriétés des projections vectorielles p et q.

Théorème 3.3 *Les projections vectorielles p et q vérifient*

$$p\circ p=p,\quad q\circ q=q,\quad p\circ q=q\circ p=0,\quad p+q=\mathrm{id}_E$$

(du fait des propriétés $p\circ p=p$, $q\circ q=q$, on dit que les applications p et q sont idempotentes), et

$$\mathrm{Im}\,p=\mathrm{Inv}\,p=F,\quad \mathrm{Ker}\,p=G,\quad \mathrm{Im}\,q=\mathrm{Inv}\,q=G,\quad \mathrm{Ker}\,q=F.$$

où $\mathrm{Inv}\,p$ *est l'ensemble des vecteurs invariants par p :*

$$\mathrm{Inv}\,p=\{\vec{w}\in E\ :p(\vec{w})=\vec{w}\}.$$

Remarque.

L'égalité $p(\vec{w}) = \vec{w}$ est équivalente à $(p - \mathrm{id}_E)(\vec{w}) = \vec{0}$ et l'ensemble des vecteurs invariants par p s'écrit encore $\mathrm{Inv}\, p = \mathrm{Ker}(p - \mathrm{id}_E)$.

Démonstration.

1. On démontre d'abord la première série de propriétés.
   — Soit $\vec{w} \in E$. En écrivant $p(\vec{w}) = p(\vec{w}) + \vec{0}$ avec $(p(\vec{w}), \vec{0}) \in F \times G$, on voit que $p(p(\vec{w})) = p(\vec{w})$ et $q(p(\vec{w})) = \vec{0}$, c'est-à-dire $p \circ p = p$ et $q \circ p = 0$. Ces propriétés subsistent en échangeant p et q.
   — D'autre part, on a, par définition de p et q, pour tout $\vec{w} \in E$: $\vec{w} = p(\vec{w}) + q(\vec{w}) = (p + q)(\vec{w})$, soit encore $p + q = \mathrm{id}_E$.

2. En ce qui concerne la deuxième série de propositions, on détermine par exemple les noyau et image de p.
   — Par définition de p, on a $\forall \vec{w} \in E$, $p(\vec{w}) \in F$ et donc $\mathrm{Im}\, p \subset F$. Si $\vec{w} \in \mathrm{Inv}\, p$, alors $\vec{w} = p(\vec{w}) \in F$ et donc $\mathrm{Inv}\, p \subset F$. Réciproquement, on a $\forall \vec{u} \in F$, $p(\vec{u}) = \vec{u}$ et l'on obtient à la fois $\vec{u} \in \mathrm{Im}\, p$ et $\vec{u} \in \mathrm{Inv}\, p$. D'où $F \subset \mathrm{Im}\, p$ et $F \subset \mathrm{Inv}\, p$. On a ainsi montré que $\mathrm{Im}\, p = \mathrm{Inv}\, p = F$.
   — Toujours par définition de p, on a

$$\vec{w} \in \mathrm{Ker}\, p \iff p(\vec{w}) = \vec{0} \iff \vec{u} = \vec{0} \iff \vec{w} = \vec{v}$$

   et donc $\vec{w} \in G$ et réciproquement. Il vient $\mathrm{Ker}\, p = G$. □

La propriété d'idempotence est en fait caractéristique des projections vectorielles.

Théorème 3.4 *Soit p un endomorphisme de E idempotent : $p \circ p = p$. Alors, on a $E = \mathrm{Im}\, p \oplus \mathrm{Ker}\, p$, $\mathrm{Im}\, p = \mathrm{Inv}\, p$ et p est la projection vectorielle de E sur $\mathrm{Im}\, p$ parallèlement à $\mathrm{Ker}\, p$.*

Démonstration.

1. Montrons d'abord que $\mathrm{Im}\, p = \mathrm{Inv}\, p$. Si $\vec{u} \in \mathrm{Im}\, p$, il existe $\vec{w} \in E$ tel que $p(\vec{w}) = \vec{u}$. On a alors $\vec{u} = p(\vec{w}) = (p \circ p)(\vec{w}) = p(\vec{u})$ et donc $\vec{u} \in \mathrm{Inv}\, p$. La réciproque est immédiate : si $\vec{u} \in \mathrm{Inv}\, p$, on a $\vec{u} = p(\vec{u}) \in \mathrm{Im}\, p$.

2. Montrons que $E = \mathrm{Inv}\, p \oplus \mathrm{Ker}\, p$.
   — Vérifions que $\mathrm{Im}\, p \cap \mathrm{Ker}\, p = \{\vec{0}\}$. Soit $\vec{u} \in \mathrm{Im}\, p \cap \mathrm{Ker}\, p$. On a vu précédemment que $\mathrm{Im}\, p = \mathrm{Inv}\, p$, donc $\vec{u} = p(\vec{u}) = \vec{0}$.
   — Soit $\vec{u} \in E$. On a la décomposition $\vec{u} = p(\vec{u}) + [\vec{u} - p(\vec{u})]$. On a $p(\vec{u}) \in \mathrm{Im}\, p$, puis $p(\vec{u} - p(\vec{u})) = p(\vec{u}) - (p \circ p)(\vec{u}) = p(\vec{u}) - p(\vec{u}) = \vec{0}$ et donc $\vec{u} - p(\vec{u}) \in \mathrm{Ker}\, p$. Ainsi, $E = \mathrm{Inv}\, p + \mathrm{Ker}\, p$.

   Notons que d'après le théorème du rang, pour montrer que les sous-espaces vectoriels $\mathrm{Im}\, p$ et $\mathrm{Ker}\, p$ sont supplémentaires dans E, il aurait suffit de vérifier seulement que $\mathrm{Im}\, p \cap \mathrm{Ker}\, p = \{\vec{0}\}$.

3. En utilisant à présent la somme directe $E = \mathrm{Inv}\, p \oplus \mathrm{Ker}\, p$, pour tout $\vec{w} \in E$, il existe un unique couple $(\vec{u}, \vec{v}) \in \mathrm{Inv}\, p \times \mathrm{Ker}\, p$ tel que $\vec{w} = \vec{u} + \vec{v}$. On a $p(\vec{w}) = p(\vec{u}) = \vec{u}$, ce qui montre bien que p est la projection vectorielle de E sur $\mathrm{Inv}\, p$ parallèlement à $\mathrm{Ker}\, p$. □

Exemple.

Considérons dans le $\mathbb{R}$-espace vectoriel $E = \mathbb{R}^3$ l'endomorphisme p de représentation analytique relativement à la base canonique $\mathcal{B}$ de E

$$\begin{cases} x' = 3x - 2y + 8z \\ y' = -x + 2y - 4z \\ z' = -x + y - 3z \end{cases}.$$

La matrice de p relativement à la base $\mathcal{B}$ est

$$\mathrm{Mat}(p; \mathcal{B}) = \begin{pmatrix} 3 & -2 & 8 \\ -1 & 2 & -4 \\ -1 & 1 & -3 \end{pmatrix}.$$

— Il est facile de voir que $\mathrm{Mat}(p; \mathcal{B})^2 = \mathrm{Mat}(p; \mathcal{B})$. Or $\mathrm{Mat}(p; \mathcal{B})^2$ est la matrice de $p \circ p$ relativement à la base $\mathcal{B}$, donc $p \circ p = p$ et p est une projection vectorielle de E. Déterminons ses sous-espaces vectoriels caractéristiques.
— Soit $\vec{w} = (x, y, z) \in E$. On a très facilement

$$\vec{w} \in \mathrm{Inv}\, p \Longleftrightarrow x - y + 4z = 0$$

et donc $F = \mathrm{Inv}\, p = \{(x, y, z) \in E : x - y + 4z = 0\}$; c'est un plan vectoriel.
— Remarquons que $\mathrm{Im}\, p = \mathrm{vect}(\{(3, -1, -1), (-2, 2, 1), (8, -4, -3)\}) = \mathrm{vect}(\{(3, -1, -1), (-2, 2, 1)\})$ puisque $(8, -4, -3) = 2(3, -1, -1) - (-2, 2, 1)$, et $\{(3, -1, -1), (-2, 2, 1)\}$ est une base de $\mathrm{Im}\, p$; c'est un plan vectoriel.
— Les coordonnées des vecteurs $(3, -1, -1)$ et $(-2, 2, 1)$ vérifient l'équation $x - y + 4z = 0$, cela montre que $\mathrm{Im}\, p \subset \mathrm{Inv}\, p$ et comme d'autre part $\mathrm{Im}\, p$ et $\mathrm{Inv}\, p$ ont même dimension, on a en fait $\mathrm{Im}\, p = \mathrm{Inv}\, p$, conformément au théorème 3.4.
— De même,

$$\vec{w} \in \mathrm{Ker}\, p \Longleftrightarrow \begin{cases} 3x - 2y + 8z = 0 \\ -x + 2y - 4z = 0 \\ -x + y - 3z = 0 \end{cases} \Longleftrightarrow \begin{cases} x = 2y - 4z \\ y = z \end{cases} \Longleftrightarrow \vec{w} = (-2y, y, y)$$

et donc $G = \mathrm{Ker}\, p = \{(-2\lambda, \lambda, \lambda),\ \lambda \in \mathbb{R}\}$; c'est la droite vectorielle de base $\{(-2, 1, 1)\}$.

En conclusion, p est la projection vectorielle de E sur F parallèlement à G.

3.2 Symétries vectorielles

Définition 3.5 *L'application* $\begin{array}{rcl} s : E & \longrightarrow & E \\ \vec{w} & \longmapsto & \vec{u} - \vec{v} \end{array}$ *est la symétrie vectorielle de E par rapport à F parallèlement G. De même, l'application* $\begin{array}{rcl} s' : E & \longrightarrow & E \\ \vec{w} & \longmapsto & \vec{v} - \vec{u} \end{array}$ *est la symétrie vectorielle de E par rapport G parallèlement F.*

On observe immédiatement à partir des définitions des projections et symétries que les applications p, q, s, s' sont reliées par

$$s' = -s, \quad p = \frac{1}{2}(s + \mathrm{id}_E), \quad q = \frac{1}{2}(s' + \mathrm{id}_E),$$

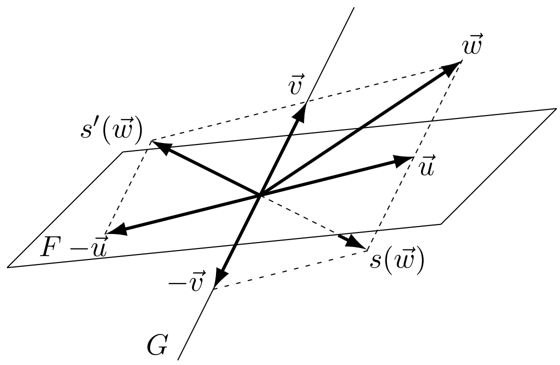

FIGURE 2.2 – Symétries par rapport aux sous-espaces vectoriels F et G

et inversement

$$s = 2p - \mathrm{id}_E, \quad s' = 2q - \mathrm{id}_E.$$

Il est donc clair que s est un endomorphisme de E. La définition de s montre également que $s \circ s = \mathrm{id}_E$ (s est une application *involutive* ou une *involution*) et alors s est un automorphisme de E d'automorphisme réciproque lui-même : $s^{-1} = s$.

REMARQUE.

La notation $\frac{1}{2}$ fait référence à l'inverse de l'élément $2 = 1 + 1$ du corps $\mathbb{K}$ pour la loi $\times$. Signalons qu'il existe des corps dans lesquels on a $1 + 1 = 0$ et donc 2 n'est pas inversible. On suppose donc ici que 2 est inversible, ce qui est notamment le cas pour $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$.

Proposition 3.6 *Les symétries vectorielles s et s' sont des automorphismes de E de matrices relativement à une base $\mathcal{B}$ de E de la forme $\mathcal{B}_F \cup \mathcal{B}_G$, où $\mathcal{B}_F$ est une base de F et $\mathcal{B}_G$ est une base de G,*

$$\mathrm{Mat}(s;\mathcal{B}) = \left(\begin{array}{c|c} I_F & O \\ \hline O & -I_G \end{array}\right) \quad \textit{et} \quad \mathrm{Mat}(s';\mathcal{B}) = \left(\begin{array}{c|c} -I_F & O \\ \hline O & I_G \end{array}\right).$$

DÉTERMINATION PRATIQUE D'UNE SYMÉTRIE VECTORIELLE.

Soit $\mathcal{B}_E$ une base de E. On souhaite déterminer analytiquement la symétrie vectorielle s relativement à la base $\mathcal{B}_E$.

— À l'aide d'un changement de bases, il est possible de calculer la matrice de s relativement à la base $\mathcal{B}_E$ à partir de la matrice relativement à $\mathcal{B}_F \cup \mathcal{B}_G$ qui est très simple :

$$\mathrm{Mat}(s;\mathcal{B}_E) = P_{(\mathcal{B}_E \to \mathcal{B}_F \cup \mathcal{B}_G)} \left(\begin{array}{c|c} I_F & O \\ \hline O & -I_G \end{array}\right) P^{-1}_{(\mathcal{B}_E \to \mathcal{B}_F \cup \mathcal{B}_G)}.$$

— À l'aide d'un système d'équations, il est possible de déterminer la représentation analytique de la symétrie vectorielle s relativement à la base $\mathcal{B}_E$: le vecteur image $\vec{w}\,' = s(\vec{w})$ est caractérisé par les deux conditions

$$\vec{w}\,' + \vec{w} \in F \quad \text{et} \quad \vec{w}\,' - \vec{w} \in G.$$

Si F et G sont donnés par des représentations paramétriques ou cartésiennes dans une base $\mathcal{B}_E$ de E, ces deux conditions conduisent à un système d'équations linéaires fournissant les coordonnées du vecteur $\vec{w}'$ relativement à la base $\mathcal{B}_E$ à partir de celles du vecteur $\vec{w}$: c'est la représentation analytique de la symétrie s.

Comme dans le cas des projections, on peut caractériser les sous-espaces vectoriels F et G à l'aide de vecteurs particuliers au regard de l'application s. En effet,

$$s(\vec{w}) = \vec{w} \Longleftrightarrow \vec{u} - \vec{v} = \vec{u} + \vec{v} \Longleftrightarrow \vec{v} = \vec{0} \Longleftrightarrow \vec{w} \in F$$

et donc $F = \operatorname{Inv} s = \operatorname{Ker}(s - \mathrm{id}_E)$. De même,

$$s(\vec{w}) = -\vec{w} \Longleftrightarrow \vec{u} - \vec{v} = -\vec{u} - \vec{v} \Longleftrightarrow \vec{u} = \vec{0} \Longleftrightarrow \vec{w} \in G$$

et donc $G = \operatorname{Ker}(s + \mathrm{id}_E)$.

Proposition 3.7 *La symétrie vectorielle s est un automorphisme involutif de E et*

$$F = \operatorname{Inv} s = \operatorname{Ker}(s - \mathrm{id}_E) \quad \textit{et} \quad G = \operatorname{Ker}(s + \mathrm{id}_E).$$

On a une réciproque à ces propriétés.

Théorème 3.8 *Soit s un endomorphisme de E involutif : $s \circ s = \mathrm{id}_E$. Alors, on a $E = \operatorname{Ker}(s - \mathrm{id}_E) \oplus \operatorname{Ker}(s + \mathrm{id}_E)$ et s est la symétrie vectorielle de E par rapport à $\operatorname{Ker}(s - \mathrm{id}_E)$ parallèlement à $\operatorname{Ker}(s + \mathrm{id}_E)$.*

DÉMONSTRATION.

En introduisant $p = \frac{1}{2}(s + \mathrm{id}_E)$, on voit que p est un endomorphisme idempotent :

$$p \circ p = \frac{1}{4}\,(s \circ s + 2s + \mathrm{id}_E) = \frac{1}{4}\,(2s + 2\mathrm{id}_E) = p.$$

On sait alors que p est la projection vectorielle sur $\operatorname{Inv} p$ parallèlement à $\operatorname{Ker} p$. Or $\operatorname{Inv} p = \operatorname{Ker}(p - \mathrm{id}_E) = \operatorname{Ker}(s - \mathrm{id}_E)$ et $\operatorname{Ker} p = \operatorname{Ker}(s + \mathrm{id}_E)$ et le résultat annoncé est ainsi prouvé. □

4 Exercices sur les applications linéaires

Exercice 1 Dans les $\mathbb{R}$-espaces vectoriels $E = \mathbb{R}^2$ et $F = \mathbb{R}^3$ rapportés à leurs bases canoniques respectives $\mathcal{B} = \{\vec{\imath}, \vec{\jmath}\}$ et $\mathcal{B}' = \{\vec{e}_1, \vec{e}_2, \vec{e}_3\}$, on considère les applications $f : E \to F$ et $g : F \to E$ définies par

$$\begin{aligned} f(x,y) &= (2x+y, x-y, x+2y), \\ g(x,y,z) &= (x+y+z, x-y-z). \end{aligned}$$

1. Montrer que f et g sont des applications linéaires.
2. Écrire les matrices de f et g relativement aux bases $\mathcal{B}$ et $\mathcal{B}'$.
3. Déterminer les noyaux et images de f et g.
4. Calculer $f \circ g$ et $g \circ f$, puis déterminer leurs matrices relativement aux bases $\mathcal{B}$ et $\mathcal{B}'$, ainsi que leurs noyaux et images que l'on comparera (au sens de l'inclusion) à ceux de f et g.

Exercice 2 Dans les $\mathbb{R}$-espaces vectoriels $E = \mathbb{R}^3$ et $F = \mathbb{R}^4$ rapportés à leurs bases canoniques respectives $\mathcal{B} = \{\vec{\imath}, \vec{\jmath}, \vec{k}\}$ et $\mathcal{B}' = \{\vec{e}_1, \vec{e}_2, \vec{e}_3, \vec{e}_4\}$, on considère l'application linéaire $f : E \to F$ telle que

$$f(\vec{\imath}) = \vec{e}_1 + \vec{e}_2 + \vec{e}_3 + \vec{e}_4, \quad f(\vec{\jmath}) = \vec{e}_1 - \vec{e}_2 + \vec{e}_3 - \vec{e}_4, \quad f(\vec{k}) = \vec{e}_1 + \vec{e}_3.$$

1. Donner l'expression analytique de f.
2. Déterminer les noyau et image de f en précisant pour chacun une base et la dimension.

Exercice 3 On se place dans le $\mathbb{R}$-espace vectoriel $E = \mathbb{R}_n[X]$ des polynômes à coefficients réels de degré au plus n.

1. Montrer que le système de vecteurs $\mathcal{B}_a = \{1, X-a, (X-a)^2, \dots, (X-a)^n\}$ est une base de E, a étant un réel fixé.
2. Calculer les coordonnées de X^k relativement à la base $\mathcal{B}_a$ pour $0 \leqslant k \leqslant n$ (on pourra utiliser la formule du binôme de Newton, ou la celle de Taylor).
3. Écrire les deux matrices de passage entre les bases $\mathcal{B}_0$ et $\mathcal{B}_a$. Quelle relation y a-t-il entre ces deux matrices ? On pourra introduire la base $\mathcal{B}_{-a}$.
4. Montrer que l'application f définie sur E par $f(P(X)) = P(X+a)$ est un automorphisme de E dont on déterminera l'automorphisme réciproque.
5. Écrire les matrices $M(f; \mathcal{B}_a, \mathcal{B}_0)$, $M(f; \mathcal{B}_0)$ et $M(f; \mathcal{B}_a)$.

Exercice 4 Dans le $\mathbb{R}$-espace vectoriel E des fonctions définies sur $\mathbb{R}$ à valeurs réelles, on considère les fonctions f_1 et f_2 définies par

$$f_1(x) = e^{ax}\cos(bx) \quad \text{et} \quad f_2(x) = e^{ax}\sin(bx)$$

où a et b sont des réels fixés, b non nul. Soit F le sous-espace vectoriel de E engendré par les vecteurs f_1 et f_2, puis $\varphi : F \to E$ l'application définie par $\varphi(f) = f'$.

1. Montrer que le système de vecteurs $\mathcal{B} = \{f_1, f_2\}$ est une base de F.
2. Montrer que φ est un automorphisme de F dont on calculera la matrice relativement à la base $\mathcal{B}$ ainsi que l'automorphisme réciproque φ^{-1}.
3. Donner une primitive de la fonction $x \mapsto e^{ax}\cos(bx+\theta)$ sans calcul d'intégrale.

Exercice 5 On se place dans le $\mathbb{R}$-espace vectoriel $E = \mathbb{R}^3$.

1. Soit p l'endomorphisme de E défini par
$$p(x, y, z) = (x, x + \tfrac{1}{2}y - \tfrac{1}{2}z, x - \tfrac{1}{2}y + \tfrac{1}{2}z).$$
Montrer que p est une projection vectorielle dont on précisera les sous-espaces vectoriels caractéristiques.
2. Soit s l'endomorphisme de E défini par
$$s(x, y, z) = (\tfrac{1}{3}(5x + 2y - 4z), y, \tfrac{1}{3}(4x + 4y - 5z)).$$
Montrer que s est une symétrie vectorielle dont on précisera les sous-espaces vectoriels caractéristiques.

Exercice 6 On se place dans le $\mathbb{R}$-espace vectoriel $E = \mathbb{R}^3$.

1. Soit P le plan vectoriel de E d'équation $x - 2y + 3z = 0$ et D la droite vectorielle de E engendrée par le vecteur $\vec{i}+\vec{j}+\vec{k}$. Déterminer la représentation analytique de la projection vectorielle p de E sur P parallèlement à D. En déduire celle de la symétrie vectorielle s par rapport à P parallèlement à D.
2. Soit P le plan vectoriel de E d'équation $x + 2y + 3z = 0$ et D la droite vectorielle de E engendrée par le vecteur $\vec{i} - \vec{j} - \vec{k}$. Déterminer la représentation analytique de la symétrie vectorielle s par rapport à P parallèlement à D. En déduire celle de la projection vectorielle s de E sur P parallèlement à D.

Exercice 7 Dans le $\mathbb{R}$-espace vectoriel $E = \mathbb{R}_3[X]$ des polynômes à coefficients réels de degré au plus 3, on considère l'application f définie sur E par
$$f(P(X)) = \tfrac{1}{2}[P(X) + P(1 - X)].$$

1. Montrer que l'application f est un endomorphisme de E dont on déterminera la matrice relativement à la base canonique de E.
2. Déterminer les noyau et image de f.
3. Calculer $f \circ f$. Préciser alors géométriquement l'endomorphisme f.

Exercice 8 Soit E, F, G des $\mathbb{K}$-espaces vectoriels et $f \in \mathcal{L}(E, F)$, $g \in \mathcal{L}(F, G)$.

1. Montrer les inclusions $\operatorname{Ker} f \subset \operatorname{Ker}(g \circ f)$ et $\operatorname{Im}(g \circ f) \subset \operatorname{Im} g$.
2. En déduire l'inégalité $\operatorname{rang}(g \circ f) \leqslant \min(\operatorname{rang}(f), \operatorname{rang}(g))$.
3. Montrer les équivalences
   (a). $\operatorname{Ker} f = \operatorname{Ker}(g \circ f) \iff \operatorname{Im} f \cap \operatorname{Ker} g = \{\vec{0}\}$. Cette condition est vérifiée en particulier lorsque g est injective.
   (b). $\operatorname{Im}(g \circ f) = \operatorname{Im} g \iff \operatorname{Im} f + \operatorname{Ker} g = E$. Cette condition est vérifiée en particulier lorsque f est surjective.
4. En déduire que
   — si g est injective, alors $\operatorname{rang}(g \circ f) = \operatorname{rang}(f)$;
   — si f est surjective, alors $\operatorname{rang}(g \circ f) = \operatorname{rang}(g)$.

Chapitre 3

Matrices

Dans le chapitre précédent, on a défini la matrice d'une application linéaire $f \in \mathcal{L}(E,F)$ entre deux $\mathbb{K}$-espaces vectoriels E et F comme étant un tableau rectangulaire dont les colonnes sont les coordonnées des images par f des vecteurs d'une base de E dans une base de F. Rappelons l'importance de ce tableau qui au travers d'un certain nombre de salaires caractérise complètement l'application linéaire f. De manière plus générale, on peut définir une matrice indépendamment de toute application linéaire comme étant un tableau rectangulaire de scalaires. Gardant toutefois en mémoire le contexte fonctionnel sous-jacent, on va définir naturellement des opérations sur les matrices conduisant à un calcul matriciel très puissant qui trouvera des applications dans une grande variété de domaines (voir le chapitre 5 traitant de la diagonalisation pour quelques exemples d'applications).

Dans tout ce chapitre, E et F désignent deux $\mathbb{K}$-espaces vectoriels.

1 Définition et opérations

1.1 Définition

Définition 1.1 *Une matrice de type $p \times n$ est un tableau rectangulaire de scalaires à p lignes et n colonnes. Une telle matrice a la forme*

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{p1} & a_{p2} & \dots & a_{pn} \end{pmatrix}.$$

On note encore ce tableau sous la forme abrégée $A = (a_{ij})_{\substack{1\leqslant i\leqslant p \\ 1\leqslant j\leqslant n}}$. Le premier (resp. deuxième) indice du scalaire a_{ij} désigne habituellement le numéro de la ligne (resp. colonne) sur laquelle il se situe. On note $\mathcal{M}_{p\times n}(\mathbb{K})$ l'ensemble des matrices de type $p \times n$ à coefficients dans $\mathbb{K}$ et l'on note plus simplement $\mathcal{M}_n(\mathbb{K})$ l'ensemble des matrices carrées de type $n \times n$ (dites d'ordre n).

Exemple.

On a défini dans le chapitre précédent des matrices dans divers contextes.

— *Matrice des coordonnées d'un vecteur dans une base :* pour un vecteur $\vec{u}$ de coordonnées $(x_1, \dots, x_n)$ relativement à une base $\mathcal{B}_E$ de E, sa matrice de coordonnées est la matrice-colonne de type $n \times 1$
$$X = \mathrm{Mat}(\vec{u}; \mathcal{B}_E) = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

— *Matrice d'une application linéaire :* pour une application linéaire $f \in \mathcal{L}(E, F)$ donnée par l'image d'une base $\mathcal{B}_E = \{\vec{e}_1, \dots, \vec{e}_n\}$ de E relativement à une base $\mathcal{B}_F = \{\vec{\varepsilon}_1, \dots, \vec{\varepsilon}_p\}$ de F selon $f(\vec{e}_j) = \displaystyle\sum_{i=1}^{p} a_{ij}\vec{\varepsilon}_i$, $1 \leqslant j \leqslant n$, la matrice de f relativement aux bases $\mathcal{B}_E$ et $\mathcal{B}_F$ est la matrice de type $p \times n$
$$A = \mathrm{Mat}(f; \mathcal{B}_E, \mathcal{B}_F) = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{p1} & \dots & a_{pn} \end{pmatrix}.$$

— *Matrice de passage :* si l'on dispose de deux bases $\mathcal{B}_E = \{\vec{e}_1, \dots, \vec{e}_n\}$ et $\mathcal{B}'_E = \{\vec{e}\,'_1, \dots, \vec{e}\,'_n\}$ de E, les vecteurs de la deuxième base étant exprimés en fonction de la première : $\vec{e}\,'_j = \displaystyle\sum_{i=1}^{n} p_{ij}\vec{e}_i$, $1 \leqslant j \leqslant n$, la matrice de passage de la base $\mathcal{B}_E$ vers la base $\mathcal{B}'_E$ est la matrice carrée d'ordre n
$$P = P_{\mathcal{B}_E \to \mathcal{B}'_E} = \begin{pmatrix} p_{11} & \dots & p_{1n} \\ \vdots & & \vdots \\ p_{n1} & \dots & p_{nn} \end{pmatrix}.$$

1.2 Opérations

Définissons tout d'abord l'*égalité* de deux matrices : pour $A = (a_{ij})_{\substack{1 \leqslant i \leqslant p \\ 1 \leqslant j \leqslant n}} \in \mathcal{M}_{p \times n}(\mathbb{K})$ et $B = (b_{ij})_{\substack{1 \leqslant i \leqslant p \\ 1 \leqslant j \leqslant n}} \in \mathcal{M}_{p \times n}(\mathbb{K})$, on pose
$$A = B \iff \forall (i,j) \in \{1, \dots, p\} \times \{1, \dots, n\},\ a_{ij} = b_{ij}.$$

Définissons à présent les opérations fondamentales du calcul matriciel.

— *Addition :* pour $A = (a_{ij})_{\substack{1 \leqslant i \leqslant p \\ 1 \leqslant j \leqslant n}} \in \mathcal{M}_{p \times n}(\mathbb{K})$ et $B = (b_{ij})_{\substack{1 \leqslant i \leqslant p \\ 1 \leqslant j \leqslant n}} \in \mathcal{M}_{p \times n}(\mathbb{K})$, on pose
$$\boxed{A + B = (a_{ij} + b_{ij})_{\substack{1 \leqslant i \leqslant p \\ 1 \leqslant j \leqslant n}} \in \mathcal{M}_{p \times n}(\mathbb{K}).}$$

— *Multiplication par un scalaire :* pour $\alpha \in \mathbb{K}$ et $A = (a_{ij})_{\substack{1 \leqslant i \leqslant p \\ 1 \leqslant j \leqslant n}} \in \mathcal{M}_{p \times n}(\mathbb{K})$, on pose
$$\boxed{\alpha \,.\, A = (\alpha a_{ij})_{\substack{1 \leqslant i \leqslant p \\ 1 \leqslant j \leqslant n}} \in \mathcal{M}_{p \times n}(\mathbb{K}).}$$

On note encore ce produit plus simplement αA.

— *Multiplication :* pour $A=(a_{ij})_{\substack{1\leqslant i\leqslant q\\1\leqslant j\leqslant p}}\in\mathcal{M}_{p\times n}(\mathbb{K})$ et $B=(b_{ij})_{\substack{1\leqslant i\leqslant p\\1\leqslant j\leqslant n}}\in\mathcal{M}_{p\times n}(\mathbb{K})$, on pose

$$\boxed{A\times B=(c_{ij})_{\substack{1\leqslant i\leqslant q\\1\leqslant j\leqslant n}}\in\mathcal{M}_{p\times n}(\mathbb{K})\text{ avec }c_{ij}=\sum_{k=1}^{p}a_{ik}b_{kj}.}$$

On note encore ce produit plus simplement AB.

— *Transposition :* pour $A=(a_{ij})_{\substack{1\leqslant i\leqslant p\\1\leqslant j\leqslant n}}\in\mathcal{M}_{p\times n}(\mathbb{K})$, on pose

$$\boxed{{}^tA=(a_{ji})_{\substack{1\leqslant i\leqslant n\\1\leqslant j\leqslant p}}\in\mathcal{M}_{p\times n}(\mathbb{K}).}$$

1.3 Quelques exemples de matrices

— *Matrices symétriques, antisymétriques :*

Définition 1.2 *Soit $A\in\mathcal{M}_n(\mathbb{K})$. La matrice A est symétrique lorsque ${}^tA=A$ et antisymétrique lorsque ${}^tA=-A$. On note $\mathcal{S}_n(\mathbb{K})$ (resp. $\mathcal{A}_n(\mathbb{K})$) l'ensemble des matrices carrées symétriques d'ordre n (resp. antisymétriques).*

Une matrice symétrique A est de la forme

$$A=\begin{pmatrix}a_{11}&a_{12}&a_{13}&\dots&a_{1n}\\a_{12}&a_{22}&a_{23}&\dots&a_{2n}\\a_{13}&a_{23}&a_{33}&\dots&a_{3n}\\\vdots&\vdots&\vdots&\ddots&\vdots\\a_{1n}&a_{2n}&a_{3n}&\dots&a_{nn}\end{pmatrix}$$

et lorsque $\mathbb{K}=\mathbb{R}$ ou $\mathbb{C}$, une matrice antisymétrique A est de la forme

$$A=\begin{pmatrix}0&a_{12}&a_{13}&\dots&a_{1n}\\-a_{12}&0&a_{23}&\dots&a_{2n}\\-a_{13}&-a_{23}&0&&a_{3n}\\\vdots&\vdots&&\ddots&\vdots\\-a_{1n}&-a_{2n}&-a_{3n}&\dots&0\end{pmatrix}.$$

Les termes diagonaux sont en effet nuls puisque dans $\mathbb{R}$ ou $\mathbb{C}$ on a $a_{ii}=-a_{ii}\Longrightarrow 2a_{ii}=0\Longrightarrow a_{ii}=0$. Il existe des corps dans lesquels on peut avoir $2x=0$ et $x\neq0$...

— *Matrices triangulaires, diagonales :* une matrice *triangulaire supérieure* (resp. *inférieure*) est une matrice carrée de la forme

$$\begin{pmatrix}a_{11}&a_{12}&a_{13}&\dots&a_{1n}\\0&a_{22}&a_{23}&\dots&a_{2n}\\0&0&a_{33}&\dots&a_{3n}\\\vdots&\vdots&&\ddots&\vdots\\0&0&0&\dots&a_{nn}\end{pmatrix}\quad\text{(resp. }\begin{pmatrix}a_{11}&0&0&\dots&0\\a_{21}&a_{22}&0&\dots&0\\a_{31}&a_{32}&a_{33}&&0\\\vdots&\vdots&\vdots&\ddots&\vdots\\a_{n1}&a_{n2}&a_{n3}&\dots&a_{nn}\end{pmatrix}\text{)}.$$

Une matrice *diagonale* est à la fois triangulaire supérieure et inférieure ; elle est donc de la forme

$$\operatorname{diag}(a_{11},a_{22},\dots,a_{nn})=\begin{pmatrix}a_{11}&0&\dots&0\\0&a_{22}&&0\\\vdots&&\ddots&\vdots\\0&0&\dots&a_{nn}\end{pmatrix}.$$

La matrice diagonale

$$I_n = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix}$$

est la *matrice unité* de $\mathcal{M}_n(\mathbb{K})$. Elle joue le rôle d'élément neutre pour la multiplication. Dans le cas particulier où tous les termes diagonaux coïncident, on a une *matrice scalaire* :

$$aI_n = \begin{pmatrix} a & 0 & \dots & 0 \\ 0 & a & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \dots & a \end{pmatrix} ;$$

c'est la matrice d'une homothétie vectorielle relativement à une base quelconque de E.

— *Matrices d'extraction :* il est possible d'extraire un terme d'une matrice par un procédé multiplicatif simple. En effet, pour une matrice $A = (a_{ij})_{\substack{1\leqslant i\leqslant p \\ 1\leqslant j\leqslant n}}$, on récupère le terme a_{ij} grâce à la multiplication de A par des matrices d'*extraction* :

$$(a_{ij}) = \begin{array}{c} \begin{pmatrix} 0 & \dots & 0 & 1 & 0 & \dots & 0 \end{pmatrix} \\ \begin{array}{ccccccc} & & & \uparrow & & & \\ & & & i & & & \end{array} \end{array} A \begin{array}{c} \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \end{array} \begin{array}{c} \\ \\ \\ \leftarrow j \\ \\ \\ \\ \end{array} .$$

La matrice-ligne ci-dessus est de type $1 \times p$ et la matrice-colonne est de type $n \times 1$.

— *Matrices de permutation :* une matrice de *transposition* s'obtient en permutant deux lignes ou deux colonnes (i^{e} et j^{e} lignes ou i^{e} et j^{e} colonnes) de la matrice unité I_n. Elle est de la forme

$$P_{ij}^{(n)} = \begin{array}{c} \begin{array}{ccccccccc} & & & i & & & j & & \\ & & & \downarrow & & & \downarrow & & \end{array} \\ \begin{pmatrix} 1 & & & & & & & & \\ & \ddots & & & & & & & \\ & & 1 & & & & & & \\ & & & 0 & \dots & \dots & 1 & & \\ & & & \vdots & 1 & & \vdots & & \\ & & & \vdots & & \ddots & \vdots & & \\ & & & \vdots & & 1 & \vdots & & \\ & & & 1 & \dots & \dots & 0 & & \\ & & & & & & & 1 & \\ & & & & & & & & \ddots & \\ & & & & & & & & & 1 \end{pmatrix} \end{array} \begin{array}{c} \\ \\ \\ \\ \\ \leftarrow i \\ \\ \\ \\ \\ \leftarrow j \\ \\ \\ \\ \end{array}$$

où tous les termes omis sont nuls.

- Pour permuter les i^{e} et j^{e} lignes d'une matrice $A = (a_{ij})_{\substack{1\leqslant i\leqslant p\\1\leqslant j\leqslant n}} \in \mathcal{M}_{p\times n}(\mathbb{K})$, on multiplie A par $P_{ij}^{(p)}$ à gauche : $P_{ij}^{(p)}A$.
- Pour permuter les i^{e} et j^{e} colonnes d'une matrice $A = (a_{ij})_{\substack{1\leqslant i\leqslant p\\1\leqslant j\leqslant n}} \in \mathcal{M}_{p\times n}(\mathbb{K})$, on multiplie A par $P_{ij}^{(n)}$ à droite : $AP_{ij}^{(n)}$.

Il est facile de prouver par exemple la deuxième assertion en interprétant A comme étant la matrice d'une application linéaire $f \in \mathcal{L}(E,F)$ et P_{ij} comme celle d'un système de vecteurs de E. Précisons cela : si $\mathcal{B}_E = \{\vec{e}_1, \dots, \vec{e}_n\}$,

$$\begin{aligned} A &= \mathrm{Mat}(f;\mathcal{B}_E,\mathcal{B}_F) = \mathrm{Mat}(\{f(\vec{e}_1),\dots,f(\vec{e}_n)\};\mathcal{B}_F), \\ P_{ij} &= \mathrm{Mat}(\{\vec{e}_1,\dots,\vec{e}_{i-1},\vec{e}_j,\vec{e}_{i+1},\dots,\vec{e}_{j-1},\vec{e}_i,\vec{e}_{j+1},\dots,\vec{e}_n\};\mathcal{B}_E). \end{aligned}$$

En fait P_{ij} est la matrice de passage de la base $\mathcal{B}_E$ vers la base

$$\{\vec{e}_1,\dots,\vec{e}_{i-1},\vec{e}_j,\vec{e}_{i+1},\dots,\vec{e}_{j-1},\vec{e}_i,\vec{e}_{j+1},\dots,\vec{e}_n\}.$$

Dans ce contexte, on obtient

$$AP_{ij} = \mathrm{Mat}(\{f(\vec{e}_1),\dots,f(\vec{e}_{i-1}),f(\vec{e}_j),f(\vec{e}_{i+1}),\dots,f(\vec{e}_{j-1}),f(\vec{e}_i),f(\vec{e}_{j+1}),\dots,f(\vec{e}_n)\};\mathcal{B}_F),$$

cette dernière matrice étant la matrice déduite de A en permutant les i^{e} et j^{e} colonnes.
Plus généralement, une matrice de *permutation* s'obtient en permutant plusieurs lignes ou plusieurs colonnes de la matrice unité I_n. Elle peut s'écrire comme un produit de matrices de tranposition.

2 Structures d'espace vectoriel et d'anneau

L'ensemble des matrices $\mathcal{M}_{p\times n}(\mathbb{K})$ jouit de propriétés de nature vectorielle (c'est un espace vectoriel) et de nature arithmétique (c'est un anneau lorsque $n = p$).

2.1 Structure d'espace vectoriel

Théorème 2.1 *Le triplet $(\mathcal{M}_{p\times n}(\mathbb{K}), +, .)$ est un $\mathbb{K}$-espace vectoriel de dimension pn. La base dite canonique de cet espace vectoriel est donnée par $\{E_{ij}, 1 \leqslant i \leqslant p, 1 \leqslant j \leqslant n\}$ où la matrice E_{ij} ne comporte que des termes nuls excepté celui qui se situe à l'intersection des i^{e} ligne et j^{e} colonne et qui vaut 1 (les matrices E_{ij} sont dites* matrices élémentaires*).*

Démonstration.

1. Notons O_{pn} la matrice de $\mathcal{M}_{p\times n}(\mathbb{K})$ dont tous les termes sont nuls et pour $A = (a_{ij})_{\substack{1\leqslant i\leqslant p\\1\leqslant j\leqslant n}} \in \mathcal{M}_{p\times n}(\mathbb{K})$, $-A = (-a_{ij})_{\substack{1\leqslant i\leqslant p\\1\leqslant j\leqslant n}}$. Les propriétés suivantes sont immédiates :

$$\forall A,B,C \in \mathcal{M}_{p\times n}(\mathbb{K}),\quad \begin{array}{ll} A+B = B+A, & (A+B)+C = A+(B+C), \\ A+O_{pn} = O_{pn}+A = A, & A+(-A) = (-A)+A = O_{pn}. \end{array}$$

De même,

$$\forall \alpha,\beta \in \mathbb{K},\ \forall A,B \in \mathcal{M}_{p\times n}(\mathbb{K}),\quad \begin{array}{ll} 1A = A, & \alpha(\beta A) = (\alpha\beta)A, \\ \alpha(A+B) = \alpha A+\alpha B, & (\alpha+\beta)A = \alpha A+\beta A. \end{array}$$

Tout ceci montre que $(\mathcal{M}_{p\times n}(\mathbb{K}), +, .)$ est un $\mathbb{K}$-espace vectoriel ; O_{pn} en est le vecteur nul et $-A$ est le vecteur opposé de A. Lorsque $n = p$, on notera O_n la matrice carrée d'ordre n nulle.

2. En écrivant une matrice $A = (a_{ij})_{\substack{1\leqslant i\leqslant p\\1\leqslant j\leqslant n}} \in \mathcal{M}_{p\times n}(\mathbb{K})$ de manière naturelle

$$A = \sum_{\substack{1\leqslant i\leqslant p\\1\leqslant j\leqslant n}} a_{ij}E_{ij},$$

on observe que le système $\{E_{ij}, 1 \leqslant i \leqslant p, 1 \leqslant j \leqslant n\}$ est générateur de $\mathcal{M}_{p\times n}(\mathbb{K})$. C'est aussi un système libre puisque, si $(a_{ij})_{\substack{1\leqslant i\leqslant p\\1\leqslant j\leqslant n}}$ est une famille de scalaires telle que $\sum_{\substack{1\leqslant i\leqslant p\\1\leqslant j\leqslant n}} a_{ij}E_{ij} = O_{pn}$, on a $(a_{ij})_{\substack{1\leqslant i\leqslant p\\1\leqslant j\leqslant n}} = O_{pn}$ et donc tous les a_{ij} sont nuls. Le système $\{E_{ij}, 1 \leqslant i \leqslant p, 1 \leqslant j \leqslant n\}$ est une base de $\mathcal{M}_{p\times n}(\mathbb{K})$ qui contient pn matrices et alors $\dim(\mathcal{M}_{p\times n}(\mathbb{K})) = pn$. □

Proposition 2.2 *La transposition* $\begin{array}{ccc}\mathcal{M}_{p\times n}(\mathbb{K}) & \longrightarrow & \mathcal{M}_{n\times p}(\mathbb{K})\\ A & \longmapsto & {}^tA\end{array}$ *est une application linéaire vérifiant* $\forall A \in \mathcal{M}_{p\times n}(\mathbb{K}),\ {}^t({}^tA) = A$. *En d'autres termes,*

$$\forall A, B \in \mathcal{M}_{p\times n}(\mathbb{K}),\ {}^t(A + B) = {}^tA + {}^tB$$

et

$$\forall \alpha \in \mathbb{K}, \forall A \in \mathcal{M}_{p\times n}(\mathbb{K}),\ {}^t(\alpha A) = \alpha\, {}^tA.$$

Lorsque $p = n$, *c'est la symétrie vectorielle par rapport à* $\mathcal{S}_n(\mathbb{K})$ *parallèlement à* $\mathcal{A}_n(\mathbb{K})$.

Démonstration.

Il est très facile de voir que la transposition est linéaire. Supposons maintenant que $p = n$. On a $\forall A \in \mathcal{M}_n(\mathbb{K}),\ {}^t({}^tA) = A$; la transposition est donc un endomorphisme involutif, c'est une symétrie vectorielle. L'ensemble des matrices invariantes est précisément celui des matrices symétriques et l'ensemble des matrices anti-invariantes est celui des matrices antisymétriques. □

Proposition 2.3 *Lorsque* $\mathbb{K} = \mathbb{R}$ *ou* $\mathbb{C}$, *les ensembles des matrices carrées symétriques et antisymétriques de type* $n \times n$ *sont des sous-espaces vectoriels du* $\mathbb{K}$*-espace vectoriel* $\mathcal{M}_n(\mathbb{K})$ *supplémentaires de dimensions respectives* $n(n+1)/2$ *et* $n(n-1)/2$: $\mathcal{M}_n(\mathbb{K}) = \mathcal{S}_n(\mathbb{K}) \oplus \mathcal{A}_n(\mathbb{K})$.

Démonstration.

1. Une matrice symétrique peut s'écrire sous la forme

$$A = \sum_{1\leqslant i\leqslant n} a_{ii}E_{ii} + \sum_{1\leqslant i<j\leqslant n} a_{ij}(E_{ij} + E_{ji}).$$

L'ensemble $\mathcal{S}_n(\mathbb{K})$ est donc constitué de combinaisons linéaires des matrices du système $\mathcal{S} = \{E_{ii}, 1 \leqslant i \leqslant n\} \cup \{E_{ij} + E_{ji}, 1 \leqslant i < j \leqslant n\}$. Cet ensemble est alors un sous-espace vectoriel de $\mathcal{M}_n(\mathbb{K})$ et le système $\mathcal{S}$ est générateur de $\mathcal{S}_n(\mathbb{K})$. Il est immédiat que c'est en fait une base. On en déduit la dimension de $\mathcal{S}_n(\mathbb{K})$:

$$\dim(\mathcal{S}_n(\mathbb{K})) = \mathrm{card}(\mathcal{S}) = n + (n-1) + (n-2) + \cdots + 3 + 2 + 1 = \frac{n(n+1)}{2}.$$

2. De même, une matrice antisymétrique peut s'écrire sous la forme (lorsque $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$)

$$A = \sum_{1 \leqslant i < j \leqslant n} a_{ij}(E_{ij} - E_{ji})$$

On voit que l'ensemble $\mathcal{A}_n(\mathbb{K})$ est constitué de combinaisons linéaires des matrices du système $\mathcal{S}' = \{E_{ij} - E_{ji}, 1 \leqslant i < j \leqslant n\}$. Cet ensemble est donc un sous-espace vectoriel de $\mathcal{M}_n(\mathbb{K})$, le système $\mathcal{S}'$ est générateur de $\mathcal{A}_n(\mathbb{K})$ et c'est même une base. D'où la dimension de $\mathcal{A}_n(\mathbb{K})$:

$$\dim(\mathcal{A}_n(\mathbb{K})) = \mathrm{card}(\mathcal{S}') = (n-1) + (n-2) + \cdots + 3 + 2 + 1 = \frac{n(n-1)}{2}.$$

3. Soit $A \in \mathcal{S}_n(\mathbb{K}) \cap \mathcal{A}_n(\mathbb{K})$. On a ${}^tA = A = -A$, donc $2A = O_n$ puis, lorsque $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$, $A = O_n$. Il vient $\mathcal{S}_n(\mathbb{K}) \cap \mathcal{A}_n(\mathbb{K}) = \{O_n\}$. On a d'autre part

$$\dim(\mathcal{S}_n(\mathbb{K})) + \dim(\mathcal{A}_n(\mathbb{K})) = n^2 = \dim(\mathcal{M}_n(\mathbb{K})),$$

On en conclut que les sous-espace vectoriels $\mathcal{S}_n(\mathbb{K})$ et $\mathcal{A}_n(\mathbb{K})$ sont supplémentaires dans $\mathcal{M}_n(\mathbb{K})$. □

Remarque.

La méthode utilisée ci-dessus pour prouver que $\mathcal{M}_n(\mathbb{K}) = \mathcal{S}_n(\mathbb{K}) \oplus \mathcal{A}_n(\mathbb{K})$ repose sur un argument de dimension, évitant ainsi de rechercher pour toute matrice de $\mathcal{M}_n(\mathbb{K})$ une décomposition en la somme d'une matrice symétrique et d'une matrice antisymétrique. Ce procédé n'est pas constructif et il peut être parfois utile de savoir exhiber une telle décomposition. Dans le cas présent, on sait effectivement écrire une matrice A sous la forme $A = B + C$ avec $(B, C) \in \mathcal{S}_n(\mathbb{K}) \times \mathcal{A}_n(\mathbb{K})$. On a nécessairement ${}^tA = {}^tB + {}^tC = B - C$ et on tire immédiatement $B = \frac{1}{2}(A + {}^tA)$ et $C = \frac{1}{2}(A - {}^tA)$. On vérifie aisément *a posteriori* que $B \in \mathcal{S}_n(\mathbb{K})$ et $C \in \mathcal{A}_n(\mathbb{K})$.

2.2 Structure d'anneau

Théorème 2.4 *Pour $n \geqslant 2$, le triplet $(\mathcal{M}_n(\mathbb{K}), +, \times)$ est un anneau non commutatif, unitaire et non intègre.*

Démonstration.

1. Signalons tout d'abord des propriétés d'associativité et de distributivité générales :

$$\forall A \in \mathcal{M}_{r\times q}(\mathbb{K}),\ \forall B \in \mathcal{M}_{q\times p}(\mathbb{K}),\ \forall C \in \mathcal{M}_{p\times n}(\mathbb{K}),\ (AB)C = A(BC),$$

$$\forall A, B \in \mathcal{M}_{q\times p}(\mathbb{K}),\ \forall C, D \in \mathcal{M}_{p\times n}(\mathbb{K}),\ (A+B)C = AC+BC \quad \text{et} \quad A(C+D) = AC+AD.$$

Vérifions par exemple la propriété d'associativité (celle de distributivité se vérifie de manière analogue).

— Les termes généraux des matrices $(AB)C$ et $A(BC)$ sont égaux :

$$\sum_{l=1}^{p} \left(\sum_{k=1}^{q} a_{ik} b_{kl} \right) c_{lj} = \sum_{k=1}^{q} a_{ik} \left(\sum_{l=1}^{p} b_{kl} c_{lj} \right).$$

— Une autre manière de prouver l'associativité consiste à introduire des applications linéaires $f \in \mathcal{L}(G, H)$, $g \in \mathcal{L}(F, G)$ et $h \in \mathcal{L}(E, F)$ associées aux matrices A, B et C :

$$A = \mathrm{Mat}(f; \mathcal{B}_G, \mathcal{B}_H),\ B = \mathrm{Mat}(g; \mathcal{B}_F, \mathcal{B}_G),\ C = \mathrm{Mat}(h; \mathcal{B}_E, \mathcal{B}_F).$$

On a

$$(AB)C = \mathrm{Mat}((f \circ g) \circ h; \mathcal{B}_E, \mathcal{B}_H),\ A(BC) = \mathrm{Mat}(f \circ (g \circ h); \mathcal{B}_E, \mathcal{B}_H),$$

et comme $(f \circ g) \circ h = f \circ (g \circ h)$, on voit que $(AB)C = A(BC)$.

2. Dans le cas des matrices carrées, les propriétés décrites ci-dessus sont de « véritables » associativité et distributivité (pour des lois de composition internes). On a donc en plus d'une structure d'espace vectoriel une structure d'anneau. De plus, il admet un élément unité qui est la matrice unité I_n :

$$\forall A \in \mathcal{M}_n(\mathbb{K}),\ AI_n = I_nA = A.$$

Cet anneau n'est ni commutatif ni intègre dès que $n \geqslant 2$. Par exemple, pour $n = 2$, on a

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \quad \text{et} \quad \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix},$$

les matrices $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ et $\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$ ne commutent pas. On a également

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = O_2$$

ce qui montre que les matrices $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ et $\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$ sont des diviseurs de O_2. Dans le cas général, on peut construire des matrices carrées d'ordre n qui ne commutent pas, ainsi que des diviseurs de O_n en complétant les deux précédentes par des 0. □

Remarque.

En combinant toutes les propriétés décrites dans ces deux paragraphes, et en rajoutant une dernière propriété reliant les lois . et $\times$:

$$\forall \alpha, \beta \in \mathbb{K}, \forall A, B \in \mathcal{M}_n(\mathbb{K}),\ (\alpha \,.\, A) \times (\beta \,.\, B) = (\alpha\beta)(A \times B)$$

on obtient pour le quadruplet $(\mathcal{M}_n(\mathbb{K}), +, ., \times)$ une structure d'*algèbre unitaire*.

On vérifie sans peine le résultat suivant.

Proposition 2.5 *Soit $m \in \mathbb{N}$ et A, B deux matrices diagonales d'ordre n données par $A = \mathrm{diag}(a_1, \ldots, a_n)$ et $B = \mathrm{diag}(b_1, \ldots, b_n)$. Le produit AB et la puissance m^e de A sont les matrices diagonales*

$$AB = \mathrm{diag}(a_1b_1, \ldots, a_nb_n) \quad \textit{et} \quad A^m = \mathrm{diag}(a_1^m, \ldots, a_n^m).$$

Théorème 2.6 (Formule du binôme de Newton) *Si les matrices carrées A et B commutent (i.e. $AB = BA$), on a pour tout entier naturel m*

$$(A+B)^m = \sum_{i=0}^{m} C_m^i A^i B^{m-i} \quad \textit{avec} \quad C_m^i = \frac{m!}{i!(m-i)!} = \frac{m(m-1)\dots(m-i+1)}{i!}.$$

La démonstration de cette formule se fait par récurrence sur m comme dans le cas classique. Insistons sur le fait que la multiplication matricielle n'étant pas commutative, la formule du binôme peut être mise en défaut. Examinons le cas particulier $m = 2$:

$$(A+B)^2 = (A+B)(A+B) = A^2 + AB + BA + B^2,$$

et pour avoir $(A+B)^2 = A^2 + 2AB + B^2$, il faut et il suffit que $AB = BA$.

Exemple.

Soit $A = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 4 \\ 0 & 0 & 1 \end{pmatrix}$. En écrivant $A = I_3 + B$ avec $B = \begin{pmatrix} 0 & 2 & 3 \\ 0 & 0 & 4 \\ 0 & 0 & 0 \end{pmatrix}$ et en observant que $B^2 = \begin{pmatrix} 0 & 0 & 8 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ et $B^m = O_3$ pour tout $m \geqslant 3$, la formule du binôme donne, puisque la matrice unité I_3 commute avec n'importe quelle matrice,

$$A^m = I_3 + mB + \frac{m(m-1)}{2} B^2 = \begin{pmatrix} 1 & 2m & m(4m-1) \\ 0 & 1 & 4m \\ 0 & 0 & 1 \end{pmatrix}.$$

Proposition 2.7 *On a*

$$\forall A \in \mathcal{M}_{n\times p}(\mathbb{K}),\ \forall B \in \mathcal{M}_{p\times q}(\mathbb{K}),\ {}^t(AB) = {}^tB\,{}^tA \quad \textit{et} \quad \forall A \in \mathcal{M}_n(\mathbb{K}),\ \forall m \in \mathbb{N}^*,\ {}^t(A^m) = ({}^tA)^m.$$

Démonstration.

Posons $A = (a_{ij})_{\substack{1\leqslant i\leqslant n\\ 1\leqslant j\leqslant p}}$ et $B = (b_{ij})_{\substack{1\leqslant i\leqslant p\\ 1\leqslant j\leqslant q}}$. On a

$$AB = \left(\sum_{k=1}^{p} a_{ik} b_{kj}\right)_{\substack{1\leqslant i\leqslant n\\ 1\leqslant j\leqslant q}} \quad \text{puis} \quad {}^t(AB) = \left(\sum_{k=1}^{p} b_{ki} a_{jk}\right)_{\substack{1\leqslant i\leqslant q\\ 1\leqslant j\leqslant n}} = {}^tB\,{}^tA. \qquad \square$$

2.3 Matrices inversibles

2.3.1 Matrices inversibles, calcul de l'inverse

Définition 2.8 *Une matrice carrée $A \in \mathcal{M}_n(\mathbb{K})$ est inversible lorsqu'il existe une matrice $A' \in \mathcal{M}_n(\mathbb{K})$ telle que $AA' = A'A = I_n$. Dans ce cas, la matrice A admet un unique inverse noté A^{-1}.*

Le calcul de l'inverse d'une matrice est en général très délicat. On dispose de plusieurs méthodes. En introduisant un endomorphisme $f \in \mathcal{L}(E)$ dont la matrice relativement à une base $\mathcal{B}$ est A, on voit que A est inversible si et seulement si f est un automorphisme de E. En effet si A est inversible et si $g \in \mathcal{L}(E)$ est l'endomorphisme dont la matrice relativement à $\mathcal{B}$ est A^{-1}, les égalités $A\times A^{-1} = A^{-1}\times A = I_n$ sont équivalentes à $\mathrm{Mat}(f\circ g;\mathcal{B}) = \mathrm{Mat}(g\circ f;\mathcal{B}) = \mathrm{Mat}(\mathrm{id}_E;\mathcal{B})$, soit encore $f\circ g = g\circ f = \mathrm{id}_E$. Ces dernières relations traduisent le fait que f est bijective de réciproque g.

Voici deux méthodes pour déterminer f^{-1} :
— la recherche de la représentation analytique de f^{-1} conduit à un système de n équations à n inconnues scalaires ;
— la recherche de l'image d'une base conduit à un système de n équations à n inconnues vectorielles.
En fait les deux systèmes précédents ont des matrices de coefficients transposées l'une de l'autre.

Exemple.

Inverse d'une matrice diagonale : considérons la matrice diagonale $A = \mathrm{diag}(a_1,\dots,a_n)$. Introduisons l'endomorphisme f dont A est la matrice relativement à une base fixée $\mathcal{B} = \{\vec{e}_1,\dots,\vec{e}_n\}$. Celui-ci est caractérisé par $\forall i \in \{1,\dots,n\},\ f(\vec{e}_i) = a_i\vec{e}_i$. Il est immédiat que le rang de f est égal au nombre de termes a_i non nuls. Ainsi, f est bijective si et seulement si $\forall i \in \{1,\dots,n\},\ a_i \neq 0$, auquel cas l'endomorphisme réciproque f^{-1} est caractérisé par $\forall i \in \{1,\dots,n\},\ f^{-1}(\vec{e}_i) = \dfrac{1}{a_i}\vec{e}_i$. Ceci conduit à la matrice inverse de A :

$$\boxed{\mathrm{diag}(a_1,\dots,a_n)^{-1} = \mathrm{diag}\left(\frac{1}{a_1},\dots,\frac{1}{a_n}\right).}$$

Dans le cas des matrices carrées d'ordre 2, on dispose d'une écriture explicite simple pour la matrice inverse.

Proposition 2.9 *La matrice* $\begin{pmatrix} a & c \\ b & d \end{pmatrix}$ *est inversible si et seulement si* $ad-bc \neq 0$ *et dans ce cas,*

$$\begin{pmatrix} a & c \\ b & d \end{pmatrix}^{-1} = \frac{1}{ad-bc}\begin{pmatrix} d & -c \\ -b & a \end{pmatrix}.$$

Proposition 2.10 *Soit A et B deux matrices inversibles. Les matrices AB et ${}^tA$ sont inversibles et*

$$(AB)^{-1} = B^{-1}A^{-1} \quad \textit{et} \quad ({}^tA)^{-1} = {}^t(A^{-1}).$$

Démonstration.

Il suffit de voir que

$$\begin{aligned} (AB)\times(B^{-1}A^{-1}) &= A\times(BB^{-1})\times A^{-1} &= AA^{-1} &= I_n, \\ (B^{-1}A^{-1})\times(AB) &= B^{-1}\times(A^{-1}A)\times B &= B^{-1}B &= I_n, \end{aligned}$$

et que

$$ {}^tA \times {}^t(A^{-1}) = {}^t(A^{-1}A) = {}^tI_n = I_n, \qquad {}^t(A^{-1}) \times {}^tA = {}^t(AA^{-1}) = {}^tI_n = I_n. \qquad \square $$

Remarques.

1. Pour un entier positif $m \in \mathbb{N}^*$ et une matrice inversible A, on peut définir A^{-m} comme étant l'inverse de la matrice A^m ou encore la puissance m^{e} de la matrice A^{-1} : $A^{-m} = (A^m)^{-1} = (A^{-1})^m$.
2. L'ensemble des matrices carrées d'ordre n inversibles possède une structure de groupe pour la multiplication appelée *groupe linéaire de* $\mathbb{K}^n$.

2.3.2 Changement de bases, matrices équivalentes, semblables

Dans le chapitre 2, on a déjà étudié l'effet d'un changement de base sur un vecteur et sur une application linéaire. On reprend cette notion ici dans le contexte du calcul matriciel.

— *Changement de bases pour les vecteurs :* soient $\mathcal{B}_E$ et $\mathcal{B}'_E$ des bases d'un $\mathbb{K}$-espace vectoriel E de dimension n, P la matrice de passage de la base $\mathcal{B}_E$ vers la base $\mathcal{B}'_E$: $P = P_{\mathcal{B}_E \to \mathcal{B}'_E}$. Rappelons que la matrice de passage de $\mathcal{B}'_E$ vers $\mathcal{B}_E$ est $P^{-1} = P_{\mathcal{B}'_E \to \mathcal{B}_E}$. Introduisons pour un vecteur $\vec{u} \in E$ ses matrices-colonnes de coordonnées X et X' respectivement dans les bases $\mathcal{B}_E$ et $\mathcal{B}'_E$. On sait que

$$\boxed{X = PX'.}$$

On obtient par multiplication à gauche par P^{-1} :

$$P^{-1}X = P^{-1}(PX') = (P^{-1}P)X' = I_nX' = X'$$

et l'on retrouve la formule inverse

$$\boxed{X' = P^{-1}X.}$$

— *Changement de bases pour les applications linéaires :* soient maintenant E et F des $\mathbb{K}$-espaces vectoriels de dimensions respectives n et p, $\mathcal{B}_E$ et $\mathcal{B}'_E$ des bases de E, $\mathcal{B}_F$ et $\mathcal{B}'_F$ des bases de F et $f \in \mathcal{L}(E, F)$. Désignons par A et A' les matrices de f relativement aux bases $\mathcal{B}_E$, $\mathcal{B}_F$ et $\mathcal{B}'_E$, $\mathcal{B}'_F$,

$$A = \mathrm{Mat}(f; \mathcal{B}_E, \mathcal{B}_F) \quad \text{et} \quad A' = \mathrm{Mat}(f; \mathcal{B}'_E, \mathcal{B}'_F),$$

puis

$$P = P_{\mathcal{B}_E \to \mathcal{B}'_E} \in \mathcal{M}_n(\mathbb{K}) \quad \text{et} \quad Q = P_{\mathcal{B}_F \to \mathcal{B}'_F} \in \mathcal{M}_p(\mathbb{K}).$$

Introduisons pour un vecteur $\vec{u} \in E$ ses matrices-colonnes de coordonnées X et X' respectivement dans les bases $\mathcal{B}_E$ et $\mathcal{B}'_E$, ainsi que les matrices-colonnes Y et Y' des coordonnées de $f(\vec{u})$ respectivement dans les bases $\mathcal{B}_F$ et $\mathcal{B}'_F$. D'après ce qui précède, on a $X = PX'$ et $Y = QY'$. On sait par ailleurs que $Y = AX$ et $Y' = A'X'$ (ce sont les écritures analytiques de l'application linéaire f dans les différentes bases considérées). On a alors

$$AX = Y = QY' = Q(A'X') = (QA')X' = (QA') \times (P^{-1}X) = (QA'P^{-1})X$$

c'est-à-dire

$$\forall X \in \mathcal{M}_{n\times 1}(\mathbb{K}),\ AX = (QA'P^{-1})X,$$

d'où l'on tire

$$\boxed{A = QA'P^{-1}.}$$

Inversement, en multipliant cette relation à gauche par Q^{-1} et à droite par P, on trouve

$$Q^{-1}AP = Q^{-1} \times (QA'P^{-1}) \times P = (Q^{-1}Q) \times A' \times (P^{-1}P) = I_p A' I_n = A'$$

et donc

$$\boxed{A' = Q^{-1}AP.}$$

Définition 2.11 1. *Les matrices A et A' de type $p \times n$ sont équivalentes lorsqu'il existe des matrices inversibles $P \in \mathcal{M}_n(\mathbb{K})$ et $Q \in \mathcal{M}_p(\mathbb{K})$ telles que $A' = Q^{-1}AP$.*

2. *Les matrices A et A' carrées d'ordre n sont semblables lorsqu'il existe une matrice inversible $P \in \mathcal{M}_n(\mathbb{K})$ telle que $A' = P^{-1}AP$.*

Des matrices semblables sont bien sûr équivalentes. L'objectif de la diagonalisation d'une matrice donnée sera de trouver une matrice diagonale semblable à cette matrice. Le chapitre 5 sera consacré à ce problème.

Proposition 2.12 1. *Soit A et A' deux matrices équivalentes. Alors les matrices ${}^tA$ et ${}^tA'$ sont équivalentes. Si de plus A est inversible, alors A' est aussi inversible et A^{-1} et A'^{-1} sont équivalentes.*

2. *Soit A et A' deux matrices semblables. Les résultats ci-dessus sont valides et de plus, $\forall m \in \mathbb{N}$, A^m et A'^m sont semblables.*

Démonstration.

1. Si $A' = Q^{-1}AP$ où P et Q sont deux matrices inversibles, alors ${}^tP$ et ${}^tQ$ sont également inversibles et ${}^tA' = {}^tP\,{}^tA({}^tQ)^{-1}$. Les matrices ${}^tA$ et ${}^tA'$ sont donc équivalentes.
   Si A est inversible, il est clair que $A' = Q^{-1}AP$ l'est aussi et l'on a $A'^{-1} = P^{-1}A^{-1}Q$. Les matrices A^{-1} et A'^{-1} sont donc équivalentes.
2. La démonstration précédente s'étend très facilement au cas des matrices semblables. Examinons le cas de la puissance m. Si $A' = P^{-1}AP$, alors

$$\begin{aligned} A'^m &= \underbrace{(P^{-1}AP) \times (P^{-1}AP) \times \cdots \times (P^{-1}AP)}_{m \text{ fois}} \\ &= P^{-1} \times A \times (PP^{-1}) \times A \times \cdots \times A \times (PP^{-1}) \times A \times P \\ &= P^{-1}A^m P. \end{aligned}$$

   Les matrices A^m et A'^m sont donc semblables. □

2.3.3 Rang d'une matrice

Définition 2.13 *Soit A une matrice. Introduisons une application linéaire f dont A est la matrice relativement à des bases fixées. Le rang de la matrice A est le rang de l'application linéaire f :*

$$\operatorname{rang}(A) = \operatorname{rang}(f) = \dim(\operatorname{Im} f).$$

Théorème 2.14 *1. Deux matrices équivalentes ont même rang.*

2. Soit $A \in \mathcal{M}_{p\times n}(\mathbb{K})$ une matrice de rang r. Alors A est équivalente à la matrice

$$J_{pnr} = \left(\begin{array}{c|c} I_r & O_{r\,n-r} \\ \hline O_{p-r\,r} & O_{p-r\,n-r} \end{array}\right) = \begin{array}{l} \left(\begin{array}{ccccccc} 1 & & 0 & 0 & \dots & 0 \\ & \ddots & & \vdots & & \vdots \\ 0 & & 1 & 0 & \dots & 0 \\ 0 & \dots & 0 & 0 & \dots & 0 \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \dots & 0 & 0 & \dots & 0 \end{array}\right) \begin{array}{l} \left.\vphantom{\begin{array}{c}1\\ \vdots\\ 0\end{array}}\right\} r \\ \left.\vphantom{\begin{array}{c}0\\ \vdots\\ 0\end{array}}\right\} p-r \end{array} \\ \quad\ \underbrace{\hphantom{1\ \dots\ 0}}_{r}\ \ \underbrace{\hphantom{0\ \dots\ 0}}_{n-r} \end{array}.$$

En conséquence, deux matrices de même type et de même rang sont équivalentes.

3. Une matrice carrée d'ordre n est inversible si et seulement si son rang est égal à n.

4. Une matrice et sa transposée ont même rang :

$$\operatorname{rang}({}^tA) = \operatorname{rang}(A).$$

DÉMONSTRATION.

1. Remarquons tout simplement que deux matrices équivalentes représentent la même application linéaire relativement à des bases différentes et que le rang d'une application linéaire ne dépend naturellement pas des bases choisies. Deux matrices équivalentes ont donc même rang.

2.(a). Soit $f \in \mathcal{L}(E,F)$ dont la matrice relativement à des bases $\mathcal{B}_E = \{\vec{e}_1,\dots,\vec{e}_n\}$ de E et $\mathcal{B}_F = \{\vec{\varepsilon}_1,\dots,\vec{\varepsilon}_p\}$ de F est A :

$$\operatorname{Mat}(f;\mathcal{B}_E,\mathcal{B}_F) = A.$$

L'application linéaire f est de rang r, donc, d'après le théorème du rang, $\dim(\operatorname{Ker} f) = n-r$ et $\dim(\operatorname{Im} f) = r$. En reprenant la démonstration du-dit théorème, on introduit un sous-espace vectoriel supplémentaire G de $\operatorname{Ker} f$ dans E : $E = G \oplus \operatorname{Ker} f$, puis on construit une base $\mathcal{B}'_E$ de E en réunissant une base $\{\vec{e}\,'_1,\dots,\vec{e}\,'_r\}$ de G et une base $\{\vec{e}\,'_{r+1},\dots,\vec{e}\,'_n\}$ de $\operatorname{Ker} f$:

$$\mathcal{B}'_E = \{\vec{e}\,'_1,\dots,\vec{e}\,'_r,\vec{e}\,'_{r+1},\dots,\vec{e}\,'_n\}.$$

On a vu que la restriction de f au sous-espace vectoriel G est injective, donc, en posant $\vec{\varepsilon}\,'_i = f(\vec{e}\,'_i)$ pour $1 \leqslant i \leqslant r$, le système de vecteurs $\{f(\vec{\varepsilon}\,'_1),\dots,f(\vec{\varepsilon}\,'_r)\}$ de F est libre. On le complète en une base $\mathcal{B}'_F$ de F :

$$\mathcal{B}'_F = \{\vec{\varepsilon}\,'_1,\dots,\vec{\varepsilon}\,'_r,\vec{\varepsilon}\,'_{r+1},\dots,\vec{\varepsilon}\,'_n\}.$$

Ayant $f(\vec{e}\,'_i) = \begin{cases} \vec{\varepsilon}\,'_i & \text{si } 1 \leqslant i \leqslant r \\ 0 & \text{si } r+1 \leqslant i \leqslant n \end{cases}$, il est clair que la matrice A' de f relativement aux bases $\mathcal{B}'_E$ et $\mathcal{B}'_F$ est donnée par la matrice dénotée par J_{pnr} suivante :

$$A' = J_{pnr} = \begin{array}{c} \quad f(\vec{e}'_1)\dots f(\vec{e}'_r) \quad f(\vec{e}'_{r+1})\dots f(\vec{e}'_n) \\ \left(\begin{array}{c|c} I_r & O_{r\,n-r} \\ \hline O_{p-r\,r} & O_{p-r\,n-r} \end{array}\right) \end{array} \begin{array}{c} \vec{\varepsilon}'_1 \\ \vdots \\ \vec{\varepsilon}'_r \\ \vec{\varepsilon}'_{r+1} \\ \vdots \\ \vec{\varepsilon}'_n \end{array}.$$

Relions enfin les matrices de f relativement aux différentes bases introduites à l'aide des matrices de passage $P = P_{\mathcal{B}_E \to \mathcal{B}'_E}$ et $Q = P_{\mathcal{B}_F \to \mathcal{B}'_F}$: $A' = Q^{-1}AP$. Les matrices A et A' sont équivalentes.

(b). Si A et A' sont deux matrices de type $p \times n$ et de même rang r, alors il existe des matrices inversibles $P, P' \in \mathcal{M}_p(\mathbb{K})$ et $Q, Q' \in \mathcal{M}_n(\mathbb{K})$ telles que $A = Q^{-1}J_{pnr}P$ et $A' = Q'^{-1}J_{pnr}P'$, et donc $A' = (Q'^{-1}Q)A(P^{-1}P')$. Les matrices A et A' sont équivalentes.

3. Soit $A \in \mathcal{M}_n(\mathbb{K})$. D'après le point précédent, la matrice A est inversible si et seulement si la matrice J_{nnr} l'est, donc si et seulement si $J_{nnr} = I_n$ c'est-à-dire $r = n$.
4. Une matrice $A \in \mathcal{M}_{p\times n}(\mathbb{K})$ de rang r peut s'écrire sous la forme $A = Q^{-1}J_{pnr}P$. En transposant, on trouve ${}^tA = {}^tP J_{npr} {}^tQ^{-1}$. Il est clair que les matrices élémentaires J_{pnr} et J_{npr} ont même rang et il en est donc de même pour les matrices A et ${}^tA$. □

Exemples.

1. Considérons la matrice
$$A = \begin{pmatrix} 1 & 2 & 4 \\ 1 & 1 & 3 \\ -1 & 0 & -2 \\ 2 & 1 & 5 \end{pmatrix}$$
et cherchons une matrice A' équivalente à A de la forme décrite dans le théorème ci-dessus. La méthode des zéros échelonnés sur les colonnes de A permet d'abord de déterminer le rang de A. Introduisons les vecteurs colonnes $\vec{u}_1, \vec{u}_2, \vec{u}_3$ de A relativement à une base $\{\vec{e}_1, \vec{e}_2, \vec{e}_3, \vec{e}_4\}$:
$$A = \begin{array}{c} \begin{array}{ccc} \vec{u}_1 & \vec{u}_2 & \vec{u}_3 \end{array} \\ \begin{pmatrix} 1 & 2 & 4 \\ 1 & 1 & 3 \\ -1 & 0 & -2 \\ 2 & 1 & 5 \end{pmatrix} \end{array} \begin{array}{c} \vec{e}_1 \\ \vec{e}_2 \\ \vec{e}_3 \\ \vec{e}_4 \end{array}.$$
On adopte une lecture **verticale** de la matrice A.
— Faisons apparaître des zéros sur la première ligne à partir de la deuxième colonne :
$$A_1 = \begin{array}{c} \begin{array}{ccc} \vec{u}'_1 & \vec{u}'_2 & \vec{u}'_3 \end{array} \\ \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ -1 & -2 & -2 \\ 2 & 3 & 3 \end{pmatrix} \end{array} \begin{array}{c} \vec{e}_1 \\ \vec{e}_2 \\ \vec{e}_3 \\ \vec{e}_4 \end{array}$$

avec $\vec{u}'_1 = \vec{u}_1$, $\vec{u}'_2 = -\vec{u}_2 + 2\vec{u}_1$, $\vec{u}'_3 = -\vec{u}_3 + 4\vec{u}_1$. Cette manipulation revient à multiplier à droite A par une matrice inversible P_1 :

$$A_1 = AP_1 \quad \text{avec} \quad P_1 = \begin{array}{c} \begin{array}{ccc} \vec{u}'_1 & \vec{u}'_2 & \vec{u}'_3 \end{array} \\ \begin{pmatrix} 1 & 2 & 4 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \end{array} \begin{array}{c} \\ \vec{u}_1 \\ \vec{u}_2 \\ \vec{u}_3 \end{array} .$$

— Faisons apparaître des zéros sur la deuxième ligne à partir de la troisième colonne :

$$A_2 = \begin{array}{c} \begin{array}{ccc} \vec{u}''_1 & \vec{u}''_2 & \vec{u}''_3 \end{array} \\ \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ -1 & -2 & 0 \\ 2 & 3 & 0 \end{pmatrix} \end{array} \begin{array}{c} \\ \vec{e}_1 \\ \vec{e}_2 \\ \vec{e}_3 \\ \vec{e}_4 \end{array}$$

avec $\vec{u}''_1 = \vec{u}'_1$, $\vec{u}''_2 = \vec{u}'_2$, $\vec{u}''_3 = \vec{u}'_3 - \vec{u}'_2$. Cette manipulation revient à multiplier à droite A_1 par une matrice inversible P_2 :

$$A_2 = A_1P_2 \quad \text{avec} \quad P_2 = \begin{array}{c} \begin{array}{ccc} \vec{u}''_1 & \vec{u}''_2 & \vec{u}''_3 \end{array} \\ \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix} \end{array} \begin{array}{c} \\ \vec{u}'_1 \\ \vec{u}'_2 \\ \vec{u}'_3 \end{array} .$$

À ce stade, on peut conclure que $\mathrm{rang}(A) = 2$. On va maintenant utiliser une méthode de type zéros échelonnés sur les lignes de A_2. Pour cela, on adopte une lecture **horizontale** de la matrice A_2. Introduisons donc les vecteurs lignes $\vec{v}_1, \vec{v}_2, \vec{v}_3, \vec{v}_4$ de A_2 relativement à une base $\{\vec{\varepsilon}_1, \vec{\varepsilon}_2, \vec{\varepsilon}_3\}$:

$$A_2 = \begin{array}{c} \vec{v}_1 \\ \vec{v}_2 \\ \vec{v}_3 \\ \vec{v}_4 \\ \\ \end{array} \begin{array}{c} \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ -1 & -2 & 0 \\ 2 & 3 & 0 \end{pmatrix} \\ \begin{array}{ccc} \vec{\varepsilon}_1 & \vec{\varepsilon}_2 & \vec{\varepsilon}_3 \end{array} \end{array} .$$

— Faisons apparaître des zéros sur la première colonne à partir de la deuxième ligne :

$$A'_1 = \begin{array}{c} \vec{v}'_1 \\ \vec{v}'_2 \\ \vec{v}'_3 \\ \vec{v}'_4 \\ \\ \end{array} \begin{array}{c} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \end{pmatrix} \\ \begin{array}{ccc} \vec{\varepsilon}_1 & \vec{\varepsilon}_2 & \vec{\varepsilon}_3 \end{array} \end{array}$$

avec $\vec{v}'_1 = \vec{v}_1$, $\vec{v}'_2 = \vec{v}_2 - \vec{v}_1$, $\vec{v}'_3 = -\frac{1}{2}\vec{v}_3 - \frac{1}{2}\vec{v}_1$, $\vec{v}'_4 = \vec{v}_4 - 2\vec{v}_2$. Cette manipulation revient à multiplier à gauche A' par une matrice inversible Q_1 :

$$A'_1 = Q_1 A_2 \quad \text{avec} \quad Q_1 = \begin{array}{c} \vec{v}'_1 \\ \vec{v}'_2 \\ \vec{v}'_3 \\ \vec{v}'_4 \\ \end{array} \begin{array}{c} \begin{pmatrix} 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ -1/2 & 0 & -1/2 & 0 \\ 0 & -2 & 0 & 1 \end{pmatrix} \\ \begin{array}{cccc} \vec{v}_1 & \vec{v}_2 & \vec{v}_3 & \vec{v}_4 \end{array} \end{array}.$$

— Faisons apparaître des zéros sur la deuxième colonne à partir de la troisième ligne :

$$A' = \begin{array}{c} \vec{v}''_1 \\ \vec{v}''_2 \\ \vec{v}''_3 \\ \vec{v}''_4 \\ \end{array} \begin{array}{c} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \\ \begin{array}{ccc} \vec{\varepsilon}_1 & \vec{\varepsilon}_2 & \vec{\varepsilon}_3 \end{array} \end{array}$$

avec $\vec{v}''_1 = \vec{v}'_1$, $\vec{v}''_2 = \vec{v}'_2$, $\vec{v}''_3 = \vec{v}'_3 - \vec{v}'_2$, $\vec{v}''_4 = \vec{v}'_4 - \vec{v}'_2$. Cette manipulation revient à multiplier à gauche A'_1 par une matrice inversible Q_2 :

$$A' = Q_2 A'_1 \quad \text{avec} \quad Q_2 = \begin{array}{c} \vec{v}''_1 \\ \vec{v}''_2 \\ \vec{v}''_3 \\ \vec{v}''_4 \\ \end{array} \begin{array}{c} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & -1 & 0 & 1 \end{pmatrix} \\ \begin{array}{cccc} \vec{v}'_1 & \vec{v}'_2 & \vec{v}'_3 & \vec{v}'_4 \end{array} \end{array}.$$

En conclusion, on a obtenu

$$A' = Q_2 A'_1 = Q_2 Q_1 A_2 = Q_2 Q_1 A_1 P_2 = Q_2 Q_1 A P_1 P_2,$$

soit encore, en posant $P = P_1 P_2 = \begin{pmatrix} 1 & 2 & 2 \\ 0 & -1 & 1 \\ 0 & 0 & -1 \end{pmatrix}$ et $Q^{-1} = Q_2 Q_1 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 1/2 & -1 & -1/2 & 0 \\ 1 & -3 & 0 & 1 \end{pmatrix}$,

$$A' = Q^{-1} A P = \left(\begin{array}{cc|c} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \hline 0 & 0 & 0 \\ 0 & 0 & 0 \end{array} \right)$$

2. Vérifions sur l'exemple précédent l'égalité $\operatorname{rang}({}^tA) = \operatorname{rang}(A)$. On a vu que $\operatorname{rang}(A) = 2$. Posons

$$B = {}^tA = \begin{pmatrix} 1 & 1 & -1 & 2 \\ 2 & 1 & 0 & 1 \\ 4 & 3 & -2 & 5 \end{pmatrix}.$$

La méthode des zéros échelonnés fournit successivement les matrices équivalentes à B suivantes :

$$B_1 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 2 & -1 & 2 & -3 \\ 4 & -1 & 2 & -3 \end{pmatrix} \quad \text{puis} \quad B_2 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 2 & -1 & 0 & 0 \\ 4 & -1 & 0 & 0 \end{pmatrix}.$$

Il est clair maintenant que $\operatorname{rang}(B_2) = 2$ et donc que $\operatorname{rang}(B) = 2 = \operatorname{rang}(A)$.

Remarque.

L'assertion $\mathrm{rang}({}^tA) = \mathrm{rang}(A)$ signifie que rang de A coïncide non seulement avec le rang des vecteurs-colonnes de A mais également avec celui des vecteurs-lignes de A.

Corollaire 2.15 *Soit A une matrice de rang r. On peut alors extraire de A une sous-matrice carrée d'ordre r inversible et toute sous-matrice carrée de A d'ordre supérieur à r n'est pas inversible. En d'autres termes, r est l'ordre maximum d'une matrice carrée inversible extraite de A.*

Démonstration.

Soit $A \in \mathcal{M}_{p\times n}(\mathbb{K})$ une matrice de type $p \times n$. Introduisons une base $\mathcal{B}_p$ d'un $\mathbb{K}$-espace vectoriel de dimension p et un système de vecteurs $\mathcal{S}_n$ de n vecteurs exprimés dans la base $\mathcal{B}_p$ selon la matrice A :

$$A = \mathrm{Mat}(\mathcal{S}_n; \mathcal{B}_p).$$

1. Supposons A de rang supérieur ou égal à r :

$$\mathrm{rang}(A) \geqslant r.$$

   On sait que $r \leqslant \min(p, n)$. On a $\mathrm{rang}(\mathcal{S}_n) \geqslant r$ et il existe un sous-système libre $\mathcal{S}'_r$ de $\mathcal{S}_n$ contenant r vecteurs. Soit alors

$$A' = \mathrm{Mat}(\mathcal{S}'_r; \mathcal{B}_p) \in \mathcal{M}_{p\times r}(\mathbb{K}).$$

   La matrice A' est de rang r : $\mathrm{rang}(A') = r$. Il en est de même de sa transposée : $\mathrm{rang}({}^tA') = r$. Introduisons à présent une base $\tilde{\mathcal{B}}_r$ d'un $\mathbb{K}$-espace vectoriel de dimension r et un système de vecteurs $\tilde{\mathcal{S}}'_p$ de p vecteurs exprimés dans la base $\tilde{\mathcal{B}}_r$ selon la matrice ${}^tA'$:

$${}^tA' = \mathrm{Mat}(\tilde{\mathcal{S}}'_p; \tilde{\mathcal{B}}_r) \in \mathcal{M}_{r\times p}(\mathbb{K}).$$

   Comme $\mathrm{rang}({}^tA') = r$, il existe un sous-système libre $\tilde{\mathcal{S}}_r$ de $\tilde{\mathcal{S}}'_p$ contenant r vecteurs. Soit alors

$$\tilde{A} = \mathrm{Mat}(\tilde{\mathcal{S}}_r; \tilde{\mathcal{B}}_r) \in \mathcal{M}_r(\mathbb{K}).$$

   La matrice carrée d'ordre r $\tilde{A}$ est de rang r, elle est donc inversible et il en est de même pour la matrice ${}^t\tilde{A}$ qui est extraite de A (voir Fig. 3.1).

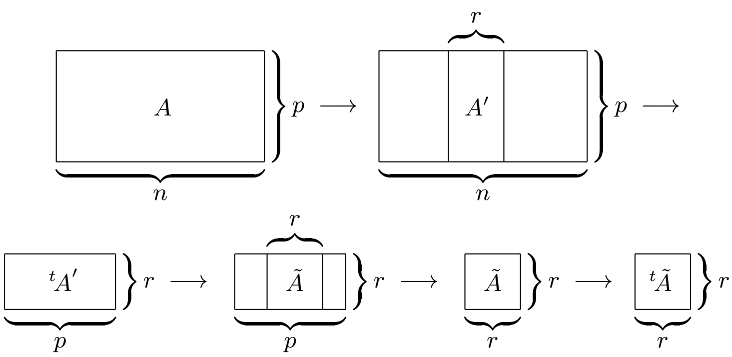

Figure 3.1 – Extraction d'une sous-matrice de A inversible (carrée d'ordre r)

2. Supposons A de rang inférieur ou égal à r :

$$\mathrm{rang}(A) \leqslant r.$$

   Tout sous-système $\mathcal{S}'$ de $\mathcal{S}_n$ contenant plus de r vecteurs, disons r' vecteurs, est nécessairement lié. Il en est bien sûr de même pour tout système dont les vecteurs sont déduits de ceux de $\mathcal{S}'$ en ne conservant que les coordonnées sur r' vecteurs de la base $\mathcal{B}_p$ (ces vecteurs appartiennent donc à un $\mathbb{K}$-espace vectoriel de dimension r'). Ainsi, toute matrice carrée d'ordre $r' > r$ extraite de A n'est pas inversible.
3. En conclusion, si A est de rang r, A possède une sous-matrice carrée d'ordre r inversible et toute sous-matrice carrée de A d'ordre supérieur à r n'est pas inversible. □

3 Exercices sur les matrices

Exercice 1 Soit A la matrice de $\mathcal{M}_3(\mathbb{R})$ donnée par $A = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$. On pose $B = A - I_3$.

1. Calculer B^n pour tout $n \in \mathbb{N}$.
2. En déduire A^n pour tout $n \in \mathbb{N}$.

Exercice 2 Calculer A^{100} pour la matrice donnée par $A = \begin{pmatrix} 5 & -4 \\ 4 & -3 \end{pmatrix}$.

Exercice 3 Soit A la matrice de $\mathcal{M}_3(\mathbb{R})$ donnée par $A = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$.

1. Calculer A^2 et l'exprimer en fonction de A et I_3. En déduire que A est inversible et déterminer son inverse A^{-1}.
2. Montrer que, pour tout $n \in \mathbb{N}$, il existe un couple de réels (α_n, β_n) tels que $A^n = \alpha_n A + \beta_n I_3$. On trouvera des relations de récurrence entre α_n, β_n, α_{n+1} et β_{n+1}.
3. Écrire une relation de récurrence entre α_n, α_{n+1} et α_{n+2}. Déterminer alors explicitement α_n et β_n.
4. En déduire A^n pour tout $n \in \mathbb{N}$.

Exercice 4 On désigne par F le sous-ensemble de $E = \mathcal{M}_n(\mathbb{R})$ des matrices de la forme

$$M_{ab} = \begin{pmatrix} a & b & \dots & b & b \\ b & a & \ddots & & b \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ b & & \ddots & a & b \\ b & b & \dots & b & a \end{pmatrix}, \quad \text{où } a, b \in \mathbb{R}.$$

1. Montrer que $(F, +, .)$ est-un sous-espace vectoriel du $\mathbb{R}$-espace vectoriel $(E, +, .)$ dont on donnera une base.
2. Montrer que $(F, +, \times)$ est-un sous-anneau de l'anneau $(E, +, \times)$.
3. Déterminer les éléments inversibles de F.
4. Déterminer les noyaux, images et rangs des endomorphismes de $\mathbb{R}^n$ associés aux matrices non-inversibles de F.

Exercice 5 (Exponentielle d'une matrice) On dit qu'une matrice carrée A est *nilpotente* s'il existe un entier $k \geqslant 1$ tel que $A^k = 0$. Le plus petit tel entier s'appelle *indice de nilpotence* de A. Pour une matrice nilpotente A d'indice de nilpotence p, on définit son *exponentielle* par

$$\boxed{\exp(A) = \sum_{i=0}^{p-1} \frac{1}{i!} A^i = \sum_{i=0}^{+\infty} \frac{1}{i!} A^i}$$

où la deuxième somme ne contient qu'un nombre fini de termes non nuls.

Soit A et B deux matrices nilpotentes qui commutent.

1. Montrer que les matrices AB et $A+B$ sont nilpotentes.
2. Montrer que $\exp(A+B) = \exp(A)\exp(B)$.
3. Montrer que $\exp(A)$ est une matrice inversible et déterminer son inverse $[\exp(A)]^{-1}$.
4. Montrer que si les matrices A et B sont semblables, alors il en est de même pour les matrices $\exp(A)$ et $\exp(B)$.
5. Soit $A \in \mathcal{M}_n(\mathbb{R})$ la matrice donnée par

$$A = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ \vdots & & & 0 & 1 \\ 0 & \dots & \dots & 0 & 0 \end{pmatrix}.$$

   Calculer A^i pour tout $i \in \mathbb{N}$. En déduire $\exp(A)$.

Exercice 6 (Trace d'une matrice) Soit $\mathbb{K}$ un corps commutatif. Pour toute matrice $A = (a_{ij})_{1\leqslant i,j\leqslant n}$ de $\mathcal{M}_n(\mathbb{K})$, on appelle *trace* de A le scalaire

$$\boxed{\operatorname{tr}(A) = \sum_{i=1}^{n} a_{ii}.}$$

1. Montrer que l'application $\operatorname{tr} : \mathcal{M}_n(\mathbb{K}) \to \mathbb{K}$ ainsi définie est une forme linéaire sur le $\mathbb{K}$-espace vectoriel $\mathcal{M}_n(\mathbb{K})$ dont on précisera le noyau.
2. Montrer que si A et B sont deux matrices de $\mathcal{M}_n(\mathbb{K})$, on a l'égalité $\operatorname{tr}(AB) = \operatorname{tr}(BA)$. Que peut-on dire des traces de deux matrices semblables ?
3. Existe-t-il des matrices A et B telles que $AB - BA = I_n$?
4. Soit $A = (a_{ij})_{\substack{1\leqslant i\leqslant m\\1\leqslant j\leqslant n}}$ une matrice de $\mathcal{M}_{m,n}(\mathbb{K})$. Calculer $\operatorname{tr}({}^tAA)$ et $\operatorname{tr}(A\,{}^tA)$.

Chapitre 4

Déterminants

La notion de déterminant est bien connue en dimension 2 et utilisée

— pour caractériser la colinéarité de deux vecteurs $\left\{ \begin{array}{l} \vec{u} = a\vec{i} + b\vec{j} \\ \vec{v} = c\vec{i} + d\vec{j} \end{array} \right.$:

$$\vec{u} \text{ et } \vec{v} \text{ sont colinéaires } \iff ad - bc \neq 0.$$

La quantité d'intérêt est $ad - bc$, appelée *déterminant du système de vecteurs* $\{\vec{u}, \vec{v}\}$ *relativement à la base de référence* $\{\vec{i}, \vec{j}\}$;

— pour écrire l'inverse d'une matrice carrée inversible $A = \begin{pmatrix} a & c \\ b & d \end{pmatrix}$:

$$A \text{ est inversible } \iff ad - bc \neq 0.$$

La quantité d'intérêt est de nouveau $ad - bc$, appelée *déterminant de la matrice* A. Dans le cas ou celle-ci est non nulle, l'inverse de la matrice A s'écrit

$$A^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -c \\ -b & a \end{pmatrix}.$$

— pour résoudre des systèmes (S) de deux équations linéaires à deux inconnues $\left\{ \begin{array}{l} ax + by = \alpha \\ cx + dy = \beta \end{array} \right.$:

$$\text{(S) admet une unique solution } \iff ad - bc \neq 0.$$

Une fois de plus, la quantité d'intérêt est $ad - bc$, appelée cette fois *déterminant principal du système* (S). De plus, dans le cas ou celle-ci est non nulle, la solution est donnée par

$$x = \frac{\alpha d - \beta b}{ad - bc}, \quad y = \frac{a\beta - \alpha c}{ad - bc}.$$

Elle fait apparaître d'autres quantités analogues : $\alpha d - \beta b$ et $a\beta - \alpha c$ qui sont appelées *déterminants relatifs du système* (S).

En dimension 2, le déterminant, considéré comme fonction de deux variables vectorielles, possède des propriétés de linéarité par rapport à chacune de ses variables ainsi que la propriété de s'annuler lorsque ses variables coïncident. Ces propriétés conduisent à introduire la notion de *forme bilinéaire alternée*.

Le définition du déterminant s'étend au cas de la dimension n et peut se faire au travers de la notion de *forme n-linéaire alternée*. On définira successivement

— le déterminant d'un système de n vecteurs relativement à une base de référence fixée ;
— le déterminant d'une matrice carrée d'ordre n ;
— le déterminant d'un endomorphisme ;
— le déterminant d'un système de n équations linéaires à n inconnues.

Puis on utilisera la théorie des déterminants :

— pour tester l'indépendance linéaire de n vecteurs ;
— pour tester l'inversibilité d'une matrice carrée d'ordre n ;
— pour résoudre un système de n équations linéaires à n inconnues.

Ce sera un outil d'un intérêt plus théorique que pratique, puisque nécessitant un nombre colossal de calculs dès que $n \geqslant 5$, et des méthodes numériques beaucoup plus économiques lui seront préférées. Cela étant, le déterminant présente l'avantage d'offrir une approche synthétique et formelle aux problèmes précédemment décrits.

On commencera par définir et traiter en détail les formes bilinéaires alternées en dimension 2 et les formes trilinéaires alternées en dimension 3 avant d'aborder le cas d'une dimension finie quelconque.

1 Préliminaires : cas des dimensions 2 et 3

1.1 Formes bilinéaires alternées en dimension 2

Soit E un $\mathbb{K}$-espace vectoriel de dimension 2.

Définition 1.1 1. *Une application* $\varphi : \begin{array}{ccc} E \times E & \longrightarrow & \mathbb{K} \\ (\vec{u}, \vec{v}) & \longmapsto & \varphi(\vec{u}, \vec{v}) \end{array}$ *est une forme bilinéaire sur E lorsque les propriétés suivantes sont satisfaites :*

— $\forall \vec{v} \in E$, *l'application* $\begin{array}{ccc} E & \longrightarrow & \mathbb{K} \\ \vec{u} & \longmapsto & \varphi(\vec{u}, \vec{v}) \end{array}$ *est une forme linéaire sur E ;*
— $\forall \vec{u} \in E$, *l'application* $\begin{array}{ccc} E & \longrightarrow & \mathbb{K} \\ \vec{v} & \longmapsto & \varphi(\vec{u}, \vec{v}) \end{array}$ *est une forme linéaire sur E.*

2. *Soit φ une forme bilinéaire sur E.*

— *La forme φ est* alternée *lorsqu'elle s'annule sur les couples de vecteurs identiques :*

$$\forall \vec{u} \in E,\ \varphi(\vec{u}, \vec{u}) = 0.$$

— *La forme φ est* antisymétrique *lorsque*

$$\forall \vec{u}, \vec{v} \in E,\ \varphi(\vec{v}, \vec{u}) = -\varphi(\vec{u}, \vec{v}).$$

Proposition 1.2 1. *Une forme bilinéaire alternée est antisymétrique.*

2. *Si $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$, une forme bilinéaire antisymétrique est alternée.*

DÉMONSTRATION.

Soit φ une forme bilinéaire sur E.

1. Supposons φ alternée. Soit $\vec{u}, \vec{v} \in E$. On a d'une part $\varphi(\vec{u}+\vec{v}, \vec{u}+\vec{v}) = 0$. D'autre part, par bilinéarité,
$$\varphi(\vec{u}+\vec{v}, \vec{u}+\vec{v}) = \varphi(\vec{u}, \vec{u}) + \varphi(\vec{u}, \vec{v}) + \varphi(\vec{v}, \vec{u}) + \varphi(\vec{v}, \vec{v}) = \varphi(\vec{u}, \vec{v}) + \varphi(\vec{v}, \vec{u}),$$
puisque $\varphi(\vec{u}, \vec{u}) = \varphi(\vec{v}, \vec{v}) = 0$. Ainsi $\varphi(\vec{u}, \vec{v}) + \varphi(\vec{v}, \vec{u}) = 0$, soit encore $\varphi(\vec{v}, \vec{u}) = -\varphi(\vec{u}, \vec{v})$ et φ est antisymétrique.
2. Supposons φ antisymétrique. Soit $\vec{u} \in E$. On a $\varphi(\vec{u}, \vec{u}) = -\varphi(\vec{u}, \vec{u})$, donc $2\varphi(\vec{u}, \vec{u}) = 0$. Cette dernière égalité entraîne, lorsque $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$, après multiplication par $\frac{1}{2}$, $\varphi(\vec{u}, \vec{u}) = 0$ qui montre que φ est alternée. □

Cherchons l'écriture d'une forme bilinéaire φ dans une base $\mathcal{B} = \{\vec{i}, \vec{j}\}$ de E. Soit $\vec{u} = x\vec{i} + y\vec{j}$ et $\vec{v} = x'\vec{i} + y'\vec{j}$. En utilisant les propriétés de bilinéarité de φ, on trouve
$$\begin{aligned}\varphi(\vec{u}, \vec{v}) = \varphi(x\vec{i} + y\vec{j}, \vec{v}) &= x\varphi(\vec{i}, x'\vec{i} + y'\vec{j}) + y\varphi(\vec{j}, x'\vec{i} + y'\vec{j})\\ &= xx'\varphi(\vec{i}, \vec{i}) + xy'\varphi(\vec{i}, \vec{j}) + x'y\varphi(\vec{j}, \vec{i}) + yy'\varphi(\vec{j}, \vec{j}).\end{aligned}$$
Dans le cas où la forme bilinéaire φ est alternée, la formule précédente se simplifie en
$$\boxed{\varphi(\vec{u}, \vec{v}) = (xy' - x'y)\varphi(\vec{i}, \vec{j}).}$$
La forme bilinéaire alternée φ est donc entièrement déterminée par la donnée du scalaire $\varphi(\vec{i}, \vec{j})$. On note
$$xy' - x'y = \begin{vmatrix} x & x' \\ y & y' \end{vmatrix},$$
c'est le *déterminant* du système de vecteurs $\{\vec{u}, \vec{v}\}$ relativement à la base $\{\vec{i}, \vec{j}\}$.

REMARQUE.

Dans le plan vectoriel euclidien orienté $E = \mathbb{R}^2$ rapporté à sa base canonique orthonormée directe $(\vec{i}, \vec{j})$, le scalaire $xy' - x'y$ est l'aire algébrique du parallélogramme construit sur les vecteurs $\vec{u}$ et $\vec{v}$, celle du parallélogramme construit sur les vecteurs $\vec{i}$ et $\vec{j}$ étant prise en référence égale à 1 (voir Fig. 4.1). En effet, introduisons les angles $\theta = \widehat{(\vec{i}, \vec{u})}$ et $\varphi = \widehat{(\vec{u}, \vec{v})}$, puis les normes $\|\vec{u}\| = r$ et $\|\vec{v}\| = r'$. On a $\theta + \varphi = \widehat{(\vec{i}, \vec{v})}$ et
$$x = r\cos\theta,\ y = \rho\sin\theta,\ x' = r'\cos(\theta + \varphi),\ y' = r'\sin(\theta + \varphi).$$
D'où
$$xy' - x'y = rr'[\cos\theta\, \sin(\theta + \varphi) - \sin\theta\, \cos(\theta + \varphi)] = rr'\sin\varphi$$
qui s'écrit de manière plus géométrique *base*×*hauteur*. Notons que l'on a aussi, en terme de produit vectoriel dans l'espace $E = R^3$ en introduisant $\vec{k} = \vec{i} \wedge \vec{j}$,
$$\vec{u} \wedge \vec{v} = (xy' - x'y)\vec{k}.$$

1.2 Formes trilinéaires alternées en dimension 3

Soit E un $\mathbb{K}$-espace vectoriel de dimension 3.

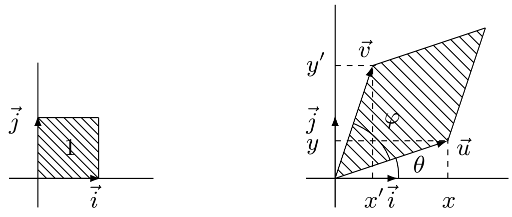

FIGURE 4.1 – Aire d'un parallélogramme

Définition 1.3 1. *Une application* $\varphi : \begin{array}{ccc} E\times E\times E & \longrightarrow & \mathbb{K} \\ (\vec{u},\vec{v},\vec{w}) & \longmapsto & \varphi(\vec{u},\vec{v},\vec{w}) \end{array}$ *est une forme trilinéaire sur* E *lorsque les propriétés suivantes sont satisfaites :*

— $\forall \vec{v},\vec{w}\in E$, *l'application* $\begin{array}{ccc} E & \longrightarrow & \mathbb{K} \\ \vec{u} & \longmapsto & \varphi(\vec{u},\vec{v},\vec{w}) \end{array}$ *est une forme linéaire sur* E *;*

— $\forall \vec{u},\vec{w}\in E$, *l'application* $\begin{array}{ccc} E & \longrightarrow & \mathbb{K} \\ \vec{v} & \longmapsto & \varphi(\vec{u},\vec{v},\vec{w}) \end{array}$ *est une forme linéaire sur* E *;*

— $\forall \vec{u},\vec{v}\in E$, *l'application* $\begin{array}{ccc} E & \longrightarrow & \mathbb{K} \\ \vec{w} & \longmapsto & \varphi(\vec{u},\vec{v},\vec{w}) \end{array}$ *est une forme linéaire sur* E.

2. *Soit* φ *une forme trilinéaire sur* E.

— *La forme* φ *est* alternée *lorsqu'elle s'annule dès que deux des trois vecteurs sont identiques :*

$$\forall \vec{u},\vec{v}\in E,\ \varphi(\vec{u},\vec{u},\vec{v})=\varphi(\vec{u},\vec{v},\vec{u})=\varphi(\vec{u},\vec{v},\vec{v})=0.$$

— *La forme* φ *est* antisymétrique *lorsque*

$$\forall \vec{u},\vec{v},\vec{w}\in E,\ \varphi(\vec{v},\vec{u},\vec{w})=\varphi(\vec{w},\vec{v},\vec{u})=\varphi(\vec{u},\vec{w},\vec{v})=-\varphi(\vec{u},\vec{v},\vec{w}).$$

Pour une forme trilinéaire φ, l'antisymétrie signifie que lorsqu'on échange deux vecteurs parmi les trois, la valeur par φ est changée en sa valeur opposée. Lorsqu'on permute les trois vecteurs, cela revient à effectuer deux échanges de deux vecteurs et la valeur par φ est inchangée :

$$\forall \vec{u},\vec{v},\vec{w}\in E,\ \varphi(\vec{v},\vec{w},\vec{u})=\varphi(\vec{w},\vec{u},\vec{v})=\varphi(\vec{u},\vec{v},\vec{w}).$$

Comme dans le cas de la dimension 2, il y a équivalence, au moins lorsque $\mathbb{K}=\mathbb{R}$ ou $\mathbb{C}$, entre « φ antisymétrique » et « φ alternée ».

Cherchons l'écriture d'une forme trilinéaire alternée φ dans une base $\mathcal{B}=\{\vec{i},\vec{j},\vec{k}\}$ de E. Soit $\vec{u}=x\vec{i}+y\vec{j}+z\vec{k}$, $\vec{v}=x'\vec{i}+y'\vec{j}+z'\vec{k}$ et $\vec{w}=x''\vec{i}+y''\vec{j}+z''\vec{k}$. En utilisant les propriétés de trilinéarité et le caractère alterné de φ, on trouve

$$\begin{aligned}
\varphi(\vec{u},\vec{v},\vec{w}) &= \varphi(x\vec{i}+y\vec{j}+z\vec{k},\vec{v},\vec{w}) \\
&= x\varphi(\vec{i},x'\vec{i}+y'\vec{j}+z'\vec{k},\vec{w})+y\varphi(\vec{j},x'\vec{i}+y'\vec{j}+z'\vec{k},\vec{w})+z\varphi(\vec{k},x'\vec{i}+y'\vec{j}+z'\vec{k},\vec{w}) \\
&= x[y'\varphi(\vec{i},\vec{j},\vec{w})+z'\varphi(\vec{i},\vec{k},\vec{w})]+y[x'\varphi(\vec{j},\vec{i},\vec{w})+z'\varphi(\vec{j},\vec{k},\vec{w})]+z[x'\varphi(\vec{k},\vec{i},\vec{w})+y'\varphi(\vec{k},\vec{j},\vec{w})].
\end{aligned}$$

Or, toujours d'après les propositions de φ,

$$\begin{array}{rclcl}
\varphi(\vec{i},\vec{j},\vec{w}) & = & z''\varphi(\vec{i},\vec{j},\vec{k}) & & \\
\varphi(\vec{i},\vec{k},\vec{w}) & = & y''\varphi(\vec{i},\vec{k},\vec{j}) & = & -y''\varphi(\vec{i},\vec{j},\vec{k}) \\
\varphi(\vec{j},\vec{i},\vec{w}) & = & z''\varphi(\vec{j},\vec{i},\vec{k}) & = & -z''\varphi(\vec{i},\vec{j},\vec{k}) \\
\varphi(\vec{j},\vec{k},\vec{w}) & = & x''\varphi(\vec{j},\vec{k},\vec{i}) & = & x''\varphi(\vec{i},\vec{j},\vec{k}) \\
\varphi(\vec{k},\vec{i},\vec{w}) & = & y''\varphi(\vec{k},\vec{i},\vec{j}) & = & y''\varphi(\vec{i},\vec{j},\vec{k}) \\
\varphi(\vec{k},\vec{j},\vec{w}) & = & x''\varphi(\vec{k},\vec{j},\vec{i}) & = & -x''\varphi(\vec{i},\vec{j},\vec{k})
\end{array}$$

d'où l'on déduit

$$\boxed{\varphi(\vec{u},\vec{v},\vec{w}) = [x(y'z''-y''z') - y(x'z''-x''z') + z(x'y''-x''y')]\varphi(\vec{i},\vec{j},\vec{k}).}$$

La forme trilinéaire alternée φ est donc entièrement déterminée par la donnée du scalaire $\varphi(\vec{i},\vec{j},\vec{k})$.

On note

$$\begin{array}{rcl}
\begin{vmatrix} x & x' & x'' \\ y & y' & y'' \\ z & z' & z'' \end{vmatrix} & = & x(y'z''-y''z') - y(x'z''-x''z') + z(x'y''-x''y') \\
 & = & x\begin{vmatrix} y' & y'' \\ z' & z'' \end{vmatrix} - y\begin{vmatrix} x' & x'' \\ z' & z'' \end{vmatrix} + z\begin{vmatrix} x' & x'' \\ y' & y'' \end{vmatrix},
\end{array}$$

c'est le *déterminant* du système de vecteurs $\{\vec{u},\vec{v},\vec{w}\}$ relativement à la base $\{\vec{i},\vec{j},\vec{k}\}$.

REMARQUES.Plaçons-nous dans le $\mathbb{R}$-espace vectoriel euclidien euclidien orienté $E = \mathbb{R}^3$ rapporté à sa base canonique orthonormée directe $\mathcal{B} = \{\vec{i},\vec{j},\vec{k}\}$. Soit trois vecteurs $\vec{u} = x\vec{i} + y\vec{j} + z\vec{k}$, $\vec{v} = x'\vec{i} + y'\vec{j} + z'\vec{k}$ et $\vec{w} = x''\vec{i} + y''\vec{j} + z''\vec{k}$ de E.

1. Rappelons que le produit vectoriel de $\vec{v}$ et $\vec{w}$ est le vecteur donné par

$$\vec{v}\wedge\vec{w} = (y'z''-y''z')\vec{i} + (x''z'-x'z'')\vec{j} + (x'y''-x''y')\vec{k}$$

et que le produit mixte des trois vecteurs $\vec{u},\vec{v},\vec{w}$ est le réel donné par

$$\vec{u}.(\vec{v}\wedge\vec{w}) = x(y'z''-y''z') + y(x''z'-x'z'') + z(x'y''-x''y').$$

On observe alors l'égalité entre déterminant et produit mixte :

$$\det_{\mathcal{B}}(\vec{u},\vec{v},\vec{w}) = \vec{u}.(\vec{v}\wedge\vec{w}) = (\vec{u}\wedge\vec{v}).\vec{w}.$$

2. Le scalaire $\begin{vmatrix} x & x' & x'' \\ y & y' & y'' \\ z & z' & z'' \end{vmatrix}$ est le volume algébrique du parallélépipède construit sur les vecteurs $\vec{u}$, $\vec{v}$ et $\vec{w}$, celui du parallélépipède construit sur les vecteurs $\vec{i}$, $\vec{j}$ et $\vec{k}$ étant pris en référence égal à 1 (voir Fig. 4.2). Ceci peut se voir facilement en introduisant l'angle $\theta = (\widehat{\vec{u}\wedge\vec{v},\vec{w}})$:

$$(\vec{u}\wedge\vec{v}).\vec{w} = \|\vec{u}\wedge\vec{v}\| \times (\|\vec{w}\| \times \cos\theta) = \mathit{base}\times\mathit{hauteur}.$$

2 Formes n-linéaires alternées en dimension n

On va étendre les notions introduites précédemment au cas d'une dimension quelconque. Soit E un $\mathbb{K}$-espace vectoriel de dimension n.

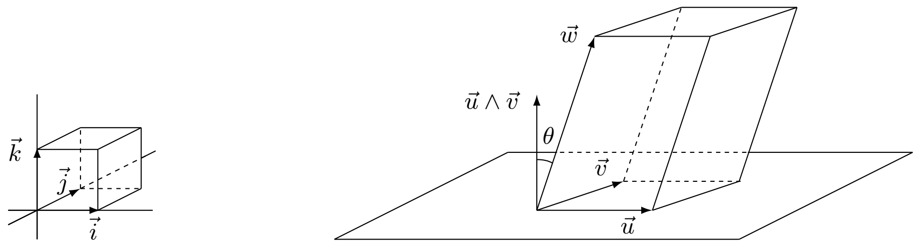

FIGURE 4.2 – Volume d'un parallélépipède

Définition 2.1 1. *Une application* $\varphi : \begin{array}{ccc} E \times \cdots \times E & \longrightarrow & \mathbb{K} \\ (\vec{u}_1, \ldots, \vec{u}_n) & \longmapsto & \varphi(\vec{u}_1, \ldots, \vec{u}_n) \end{array}$ *est une forme* n-linéaire *sur E lorsque pour tout indice $i \in \{1, \ldots, n\}$ et tous vecteurs $\vec{u}_1, \ldots, \vec{u}_n \in E$, l'application* $\begin{array}{ccc} E & \longrightarrow & \mathbb{K} \\ \vec{u} & \longmapsto & \varphi(\vec{u}_1, \ldots, \vec{u}_{i-1}, \vec{u}, \vec{u}_{i+1}, \ldots, \vec{u}_n) \end{array}$ *est une forme linéaire sur E.*

2. *Soit φ une forme n-linéaire sur E.*
   — *La forme φ est* alternée *lorsqu'elle s'annule dès que deux des n vecteurs sont identiques :*
   $$\forall i, j \in \{1, \ldots, n\} \text{ tels que } i < j,\ \forall \vec{u}_1, \ldots, \vec{u}_n \in E,\ [\vec{u}_i = \vec{u}_j \Longrightarrow \varphi(\vec{u}_1, \ldots, \vec{u}_n) = 0].$$
   — *La forme φ est* antisymétrique *lorsque*
   $$\forall i, j \in \{1, \ldots, n\} \text{ tels que } i < j,\ \forall \vec{u}_1, \ldots, \vec{u}_n \in E,$$
   $$\varphi(\vec{u}_1, \ldots, \vec{u}_j, \ldots, \vec{u}_i, \ldots, \vec{u}_n) = -\varphi(\vec{u}_1, \ldots, \vec{u}_i, \ldots, \vec{u}_j, \ldots, \vec{u}_n).$$

On prouve de manière analogue au cas de la dimension 2 que les notions de forme n-linéaire alternée et de forme n-linéaire antisymétrique coïncident dans de nombreux cas.

Proposition 2.2 1. *Une forme n-linéaire alternée est antisymétrique.*
2. *Si $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$, une forme n-linéaire antisymétrique est alternée.*

On va étudier l'effet d'un changement d'ordre dans le n-uplet de vecteurs $(\vec{u}_1, \ldots, \vec{u}_n)$ produit sur la valeur $\varphi(\vec{u}_1, \ldots, \vec{u}_n)$. À ce propos, rappelons qu'une *permutation* des indices $1, \ldots, n$ est une bijection de l'ensemble $\{1, \ldots, n\}$ sur lui-même. On note $\mathfrak{S}_n$ l'ensemble des permutations de $1, \ldots, n$. Une *transposition* est une permutation qui échange deux indices, laissant les autres invariants.

Proposition 2.3 *Soit φ une forme n-linéaire alternée sur E. On a*
$$\forall \sigma \in \mathfrak{S}_n,\ \varphi(\vec{u}_{\sigma(1)}, \ldots, \vec{u}_{\sigma(n)}) = (-1)^{N(\sigma)} \varphi(\vec{u}_1, \ldots, \vec{u}_n),$$
où $N(\sigma)$ est le nombre de transpositions nécessaires pour transformer l'ensemble d'indices $\{\sigma(1), \ldots, \sigma(n)\}$ en l'ensemble ordonné $\{1, \ldots, n\}$.

Cherchons maintenant l'écriture d'une forme n-linéaire alternée φ dans une base $\mathcal{B} = \{\vec{e}_1, \ldots, \vec{e}_n\}$

de E. Soit $\vec{u}_j = \sum_{i=1}^{n} x_{ij}\vec{e}_i$, $1 \leqslant j \leqslant n$, n vecteurs de E. Par n-linéarité, on trouve

$$\varphi(\vec{u}_1,\ldots,\vec{u}_n) = \varphi\left(\sum_{i_1=1}^{n} x_{i_1 1}\vec{e}_{i_1},\ldots,\sum_{i_n=1}^{n} x_{i_n n}\vec{e}_{i_n}\right) = \sum_{1\leqslant i_1,\ldots,i_n\leqslant n} x_{i_1 1}\ldots x_{i_n n}\,\varphi(\vec{e}_{i_1},\ldots,\vec{e}_{i_n}).$$

Le caractère alterné de φ assure que $\varphi(\vec{e}_{i_1},\ldots,\vec{e}_{i_n}) = 0$ dès que deux indices i_j coïncident et alors que la somme ci-dessus se réduit à la somme portant sur les indices $i_1,\ldots,i_n$ tous distincts, c'est-à-dire les permutations de $1,\ldots,n$:

$$\varphi(\vec{u}_1,\ldots,\vec{u}_n) = \sum_{\sigma\in\mathfrak{S}_n} x_{\sigma(1)1}\ldots x_{\sigma(n)n}\,\varphi(\vec{e}_{\sigma(1)},\ldots,\vec{e}_{\sigma(n)}).$$

On obtient finalement l'écriture suivante.

Théorème 2.4 *Soit φ une forme n-linéaire alternée sur E et $\mathcal{B} = \{\vec{e}_1,\ldots,\vec{e}_n\}$ une base de E.*

$$\varphi(\vec{u}_1,\ldots,\vec{u}_n) = \left[\sum_{\sigma\in\mathfrak{S}_n} (-1)^{N(\sigma)} x_{\sigma(1)1}\ldots x_{\sigma(n)n}\right]\varphi(\vec{e}_1,\ldots,\vec{e}_n).$$

Ainsi, la forme φ est entièrement déterminée par la donnée du scalaire $\varphi(\vec{e}_1,\ldots,\vec{e}_n)$.

Il en découle immédiatement le résultat important suivant.

Corollaire 2.5 *Deux forme n-linéaires alternées sont proportionnelles.*

Dans le reste de ce chapitre, on attache un intérêt particulier au cas de la forme n-linéaire alternée, prise en référence, telle que $\varphi(\vec{e}_1,\ldots,\vec{e}_n) = 1$. Elle seule suffit à représenter toutes les autres formes n-linéaires alternées (qui lui sont proportionnelles) définies sur le même espace vectoriel.

3 Déterminants

Soit E un $\mathbb{K}$-espace vectoriel de dimension n et $\mathcal{B} = \{\vec{e}_1,\ldots,\vec{e}_n\}$ une base de E.

3.1 Déterminant d'un système de n vecteurs

Définition 3.1 *Le déterminant relativement à la base $\mathcal{B}$ est l'unique forme n-linéaire alternée φ sur E telle que $\varphi(\vec{e}_1,\ldots,\vec{e}_n) = 1$. On note cette forme* $\det_{\mathcal{B}}$ *et pour un système de vecteurs $\mathcal{S} = \{\vec{u}_1,\ldots,\vec{u}_n\}$, on note* $\det_{\mathcal{B}}(\mathcal{S}) = \det_{\mathcal{B}}(\vec{u}_1,\ldots,\vec{u}_n)$.

Remarque.

On prendra garde à bien respecter l'ordre dans lesquels apparaissent les vecteurs dans le système $\mathcal{S}$, puisque le signe du déterminant en dépend.

Soit $\mathcal{S} = \{\vec{u}_1, \dots, \vec{u}_n\}$ un système de n vecteurs de E exprimés dans la base $\mathcal{B}$ selon $\vec{u}_j = \sum_{i=1}^{n} x_{ij}\vec{e}_i$, $1 \leqslant j \leqslant n$. D'après le théorème 2.4, on a explicitement, en introduisant la notation matricielle ci-dessous,

$$\det_{\mathcal{B}}(\mathcal{S}) = \begin{vmatrix} x_{11} & \dots & x_{1n} \\ \vdots & & \vdots \\ x_{n1} & \dots & x_{nn} \end{vmatrix} = \sum_{\sigma \in \mathfrak{S}_n} (-1)^{N(\sigma)} x_{\sigma(1)1} \dots x_{\sigma(n)n}.$$

Signalons immédiatement que cette formule, tout aussi explicite soit-elle, est très rarement employée en pratique, car elle comporte un très grand nombre d'opérations (rappelons que $\operatorname{card}(\mathfrak{S}_n) = n!$). Même avec les ordinateurs les plus puissants, une telle formule occasionnerait un temps de calcul et un coût fort élevés voire complètement déraisonnable, et d'autres méthodes de calcul sont utilisées. On en présentera quelques-unes ultérieurement.

Le principal résultat qu'on utilisera par la suite est le suivant.

Théorème 3.2 *Soit $\mathcal{S}$ un système de n vecteurs de E. Le système $\mathcal{S}$ est une base de E si et seulement si $\det_{\mathcal{B}}(\mathcal{S}) \neq 0$. Dans ce cas, on a*

$$\det_{\mathcal{S}}(\mathcal{B}) = \frac{1}{\det_{\mathcal{B}}(\mathcal{S})}.$$

Démonstration.

Posons $\mathcal{S} = \{\vec{u}_1, \dots, \vec{u}_n\}$.

1. Supposons le système $\mathcal{S}$ lié. Un vecteur de $\mathcal{S}$ est donc combinaison linéaire des autres vecteurs de $\mathcal{S}$; disons par exemple que $\vec{u}_n = \sum_{i=1}^{n-1} \alpha_i \vec{u}_i$. On a alors

$$\det_{\mathcal{B}}(\mathcal{S}) = \sum_{i=1}^{n-1} \alpha_i \det_{\mathcal{B}}(\vec{u}_1, \dots, \vec{u}_{n-1}, \vec{u}_i).$$

   Il est clair que pour tout $i \in \{1, \dots, n-1\}$, $\det_{\mathcal{B}}(\vec{u}_1, \dots, \vec{u}_{n-1}, \vec{u}_i) = 0$ puisque le système $\{\vec{u}_1, \dots, \vec{u}_{n-1}, \vec{u}_i\}$ comporte deux vecteurs identiques, et donc $\det_{\mathcal{B}}(\mathcal{S}) = 0$.

2. Supposons le système $\mathcal{S}$ libre. Puisqu'il contient n vecteurs et que $\dim(E) = n$, c'est une base de E. On peut donc considérer la forme n-linéaire alternée $\det_{\mathcal{S}}$ et la comparer à $\det_{\mathcal{B}}$. En vertu du corollaire 2.5, toutes les formes n-linéaires alternées sont proportionnelles et alors il existe un scalaire $\lambda \in \mathbb{K}$ tel que $\det_{\mathcal{B}} = \lambda \det_{\mathcal{S}}$. Le coefficient de proportionnalité λ est non nul car les formes $\det_{\mathcal{B}}$ et $\det_{\mathcal{S}}$ sont non identiquement nulles. En effet, elles valent 1 respectivement en $\mathcal{B}$ et $\mathcal{S}$, ce qui donne

$$\det_{\mathcal{B}}(\mathcal{S}) = \lambda \det_{\mathcal{S}}(\mathcal{S}) = \lambda \quad \text{et} \quad 1 = \det_{\mathcal{B}}(\mathcal{B}) = \lambda \det_{\mathcal{S}}(\mathcal{B}),$$

   d'où l'on tire

$$\det_{\mathcal{B}}(\mathcal{S}) = \lambda \neq 0 \quad \text{et} \quad \det_{\mathcal{S}}(\mathcal{B}) = \frac{1}{\lambda} = \frac{1}{\det_{\mathcal{B}}(\mathcal{S})}.$$ □

Application : équation cartésienne d'un plan vectoriel en dimension 3.

Soit E un $\mathbb{K}$-espace vectoriel de dimension 3, $\mathcal{B} = \{\vec{i}, \vec{j}, \vec{k}\}$ une base de E et $\vec{u} = a\vec{i} + b\vec{j} + c\vec{k}$ et $\vec{u}' = a'\vec{i} + b'\vec{j} + c'\vec{k}$ deux vecteurs de E non colinéaires. On cherche une caractérisation du plan vectoriel $\mathcal{P}$ engendré par les vecteurs $\vec{u}$ et $\vec{u}'$: $\mathcal{P} = \mathrm{vect}(\{\vec{u}, \vec{u}'\})$. Soit $\vec{v} = x\vec{i} + y\vec{j} + z\vec{k}$ un vecteur générique de E. On a les équivalences suivantes :

$$\vec{v} \in \mathcal{P} \iff \text{ le système } (\vec{u}, \vec{u}', \vec{v}) \text{ est lié } \iff \det_{\mathcal{B}}(\vec{u}, \vec{u}', \vec{v}) = 0 \iff \begin{vmatrix} a & a' & x \\ b & b' & y \\ c & c' & z \end{vmatrix} = 0$$

ce qui conduit à la représentation cartésienne de $\mathcal{P}$:

$$Ax + By + Cz = 0 \quad \text{avec} \quad A = bc' - b'c,\ B = a'c - ac',\ C = ab' - a'b.$$

Proposition 3.3 *Soit $\mathcal{B}$, $\mathcal{B}'$ deux bases de E et $\mathcal{S}$ un système de n vecteurs de E. On a*

$$\det_{\mathcal{B}'}(\mathcal{S}) = \det_{\mathcal{B}'}(\mathcal{B}) \times \det_{\mathcal{B}}(\mathcal{S}).$$

Démonstration.

Les formes n-linéaires alternées $\det_{\mathcal{B}}$ et $\det_{\mathcal{B}'}$ sont proportionnelles : il existe $\mu \in \mathbb{K}$ tel que $\det_{\mathcal{B}'} = \mu \det_{\mathcal{B}}$ et le coefficient de proportionnalité μ est précisément $\det_{\mathcal{B}'}(\mathcal{B})$. □

3.2 Déterminant d'une matrice carrée

En s'inspirant de la formule du théorème 2.4, on peut définir le déterminant d'une matrice carrée indépendamment de tout contexte vectoriel.

Définition 3.4 *Soit $A = (a_{ij})_{1 \leqslant i,j \leqslant n} \in \mathcal{M}_n(\mathbb{K})$. Le déterminant de la matrice carrée A est le scalaire*

$$\det(A) = \begin{vmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{vmatrix} = \sum_{\sigma \in \mathfrak{S}_n} (-1)^{N(\sigma)} a_{\sigma(1)1} \dots a_{\sigma(n)n}.$$

On peut interpréter le déterminant de A comme étant le déterminant du système de n vecteurs $\mathcal{S}$ dont la matrice des coordonnées relativement à une base $\mathcal{B}$ d'un $\mathbb{K}$-espace vectoriel de dimension n est précisément A : si $A = \mathrm{Mat}(\mathcal{S}; \mathcal{B})$, alors

$$\det(A) = \det_{\mathcal{B}}(\mathcal{S}).$$

Exemple.

Le déterminant d'une matrice triangulaire (ou diagonale) est égale au produit de ses termes diagonaux :

$$\boxed{\begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ 0 & a_{22} & \dots & a_{2n} \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \dots & a_{nn} \end{vmatrix} = \begin{vmatrix} a_{11} & 0 & \dots & 0 \\ a_{21} & a_{22} & & 0 \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} = \prod_{i=1}^{n} a_{ii}.}$$

En particulier, $\det(I_n) = 1$. En effet, supposons par exemple que A soit une matrice triangulaire supérieure. Si σ est une permutation des indices $1, \ldots, n$ autre que l'identité, il existe un indice i tel que $\sigma(i) < i$ et donc $a_{\sigma(i)i} = 0$. Sinon, on aurait pour tout $i \in \{1, \ldots, n\}$, $\sigma(i) \geqslant i$, c'est-à-dire $\sigma(1) \geqslant 1, \ldots, \sigma(n) \geqslant n$, inégalités qui entraîneraient successivement $\sigma(n) = n, \sigma(n-1) = n-1, \ldots, \sigma(1) = 1$, et σ serait en fait l'identité. Ainsi, la somme définissant $\det(A)$ ne contient au plus qu'un terme non nul, celui correspondant à $\sigma = \mathrm{id}$, c'est le produit $a_{11} \ldots a_{nn}$.

Proposition 3.5 *Soit $\alpha \in \mathbb{K}$ et $A, B \in \mathcal{M}_n(\mathbb{K})$. On a*

$$\det(\alpha A) = \alpha^n \det(A), \quad \det(AB) = \det(A)\det(B), \quad \det({}^tA) = \det(A).$$

Démonstration.

1. On a très simplement

$$\det(\alpha A) = \sum_{\sigma\in\mathfrak{S}_n} (-1)^{N(\sigma)} (\alpha a_{\sigma(1)1}) \ldots (\alpha a_{\sigma(n)n})$$
$$= \alpha^n \sum_{\sigma\in\mathfrak{S}_n} (-1)^{N(\sigma)} a_{\sigma(1)1} \ldots a_{\sigma(n)n} = \alpha^n \det(A).$$

On peut aussi utiliser le caractère n-linéaire du déterminant : si $A = \mathrm{Mat}(\{\vec{u}_1, \ldots, \vec{u}_n\}; \mathcal{B})$, alors $\alpha A = \mathrm{Mat}(\{\alpha\vec{u}_1, \ldots, \alpha\vec{u}_n\}; \mathcal{B})$ et

$$\det(A) = \det_{\mathcal{B}}(\alpha\vec{u}_1, \ldots, \alpha\vec{u}_n) = \alpha^n \det_{\mathcal{B}}(\vec{u}_1, \ldots, \vec{u}_n) = \alpha^n \det(A).$$

2. Le calcul de $\det(AB)$ est plus délicat et nécessite une interprétation judicieuse du produit AB.
   — Supposons A inversible. On peut donc considérer A comme étant la matrice de passage d'une base $\mathcal{B} = \{\vec{e}_1, \ldots, \vec{e}_n\}$ de E vers une autre $\mathcal{B}' = \{\vec{e}'_1, \ldots, \vec{e}'_n\}$: $A = P_{\mathcal{B}\to\mathcal{B}'}$. On considère ensuite B comme étant la matrice d'un système de vecteurs $\mathcal{S} = \{\vec{u}_1, \ldots, \vec{u}_n\}$ relativement à la base $\mathcal{B}'$: $B = \mathrm{Mat}(\mathcal{S}; \mathcal{B}')$. Le produit AB est alors la matrice du système de vecteurs $\mathcal{S}$ relativement à la base $\mathcal{B}$: $AB = \mathrm{Mat}(\mathcal{S}; \mathcal{B})$. En effet, dans ce contexte, on a
   $$\vec{e}'_k = \sum_{i=1}^n a_{ik}\vec{e}_i, \quad \vec{u}_j = \sum_{k=1}^n b_{kj}\vec{e}'_k$$
   et donc
   $$\vec{u}_j = \sum_{i=1}^n \left(\sum_{k=1}^n a_{ik}b_{kj}\right)\vec{e}_i = \sum_{i=1}^n c_{ij}\vec{e}_i$$
   où c_{ij} est le terme générique de la matrice AB. Ainsi, en utilisant la proposition 3.3,
   $$\det(AB) = \det_{\mathcal{B}}(\mathcal{S}) = \det_{\mathcal{B}}(\mathcal{B}') \times \det_{\mathcal{B}'}(\mathcal{S}) = \det(A) \times \det(B).$$
   — Supposons A non inversible. On a $\det(A) = 0$. La matrice AB est nécessairement non inversible. Sinon, il existerait une matrice C telle que $(AB)C = I_n$ soit encore $A(BC) = I_n$ et A serait inversible. Donc $\det(AB) = 0$ et finalement $\det(AB) = \det(A)\det(B)$ également dans ce cas.
3. Par définition,
   $$\det({}^tA) = \sum_{\sigma\in\mathfrak{S}_n} (-1)^{N(\sigma)} a_{1\sigma(1)} \ldots a_{n\sigma(n)}.$$
   Remarquons que le produit $a_{1\sigma(1)} \ldots a_{n\sigma(n)}$ ordonné relativement à l'indice de ligne, peut s'écrire, en réordonnant cette fois relativement à l'indice de colonne,
   $$a_{1\sigma(1)} \ldots a_{n\sigma(n)} = a_{\sigma^{-1}(1)1} \ldots a_{\sigma^{-1}(n)n}$$
   et donc
   $$\det({}^tA) = \sum_{\sigma\in\mathfrak{S}_n} (-1)^{N(\sigma)} a_{\sigma^{-1}(1)1} \ldots a_{\sigma^{-1}(n)n} = \sum_{\sigma'\in\mathfrak{S}_n} (-1)^{N(\sigma')} a_{\sigma'(1)1} \ldots a_{\sigma'(n)n}$$
   où l'on a effectué dans la dernière somme le changement d'indice $\sigma' = \sigma^{-1}$ et utilisé $N(\sigma') = N(\sigma)$. On récupère ainsi le déterminant de A. □

Remarque.

Voici une démonstration directe du deuxième point utilisant la définition intrinsèque du déterminant d'une matrice :

$$\begin{aligned}
\det(AB) &= \sum_{\sigma\in\mathfrak{S}_n}(-1)^{N(\sigma)}c_{\sigma(1)1}\dots c_{\sigma(n)n}\\
&= \sum_{\sigma\in\mathfrak{S}_n}(-1)^{N(\sigma)}\sum_{k_1=1}^{n}a_{\sigma(1)k_1}b_{k_1 1}\dots\sum_{k_n=1}^{n}a_{\sigma(n)k_n}b_{k_n n}\\
&= \sum_{\sigma\in\mathfrak{S}_n}(-1)^{N(\sigma)}\sum_{1\leqslant k_1,\dots,k_n\leqslant n}a_{\sigma(1)k_1}b_{k_1 1}\dots a_{\sigma(n)k_n}b_{k_n n}\\
&= \sum_{1\leqslant k_1,\dots,k_n\leqslant n}\left[\sum_{\sigma\in\mathfrak{S}_n}(-1)^{N(\sigma)}a_{\sigma(1)k_1}\dots a_{\sigma(n)k_n}\right]b_{k_1 1}\dots b_{k_n n}.
\end{aligned}$$

Observons que si l'on interprète A comme étant la matrice d'un système de vecteurs $\{\vec{u}_1,\dots,\vec{u}_n\}$ relativement à une base $\mathcal{B}$, la somme $\sum_{\sigma\in\mathfrak{S}_n}(-1)^{N(\sigma)}a_{\sigma(1)k_1}\dots a_{\sigma(n)k_n}$ représente le déterminant $\det_{\mathcal{B}}(\vec{u}_{k_1},\dots,\vec{u}_{k_n})$. Ce dernier comporte deux colonnes identiques, donc s'annule, dès que deux indices k_j coïncident et pour qu'il soit non nul, il est nécessaire que les indices $k_1,\dots,k_n$ soient une permutation de $1,\dots,n$. Dans ce cas, on peut écrire $k_1=\tau(1),\dots,k_n=\tau(n)$ avec $\tau\in\mathfrak{S}_n$, et alors, succinctement,

$$\begin{aligned}
\det(AB) &= \sum_{\tau\in\mathfrak{S}_n}\left[\sum_{\sigma\in\mathfrak{S}_n}(-1)^{N(\sigma)}a_{\sigma(1)\tau(1)}\dots a_{\sigma(n)\tau(n)}\right]b_{\tau(1)1}\dots b_{\tau(n)n}\\
&= \sum_{\tau\in\mathfrak{S}_n}\left[\sum_{\sigma\in\mathfrak{S}_n}(-1)^{N(\sigma)}a_{(\sigma\circ\tau^{-1})(1)1}\dots a_{(\sigma\circ\tau^{-1})(n)n}\right]b_{\tau(1)1}\dots b_{\tau(n)n}\\
&= \sum_{\tau\in\mathfrak{S}_n}\left[\sum_{\sigma'\in\mathfrak{S}_n}(-1)^{N(\sigma'\circ\tau)}a_{\sigma'(1)1}\dots a_{\sigma'(n)n}\right]b_{\tau(1)1}\dots b_{\tau(n)n}\\
&= \left[\sum_{\sigma\in\mathfrak{S}_n}(-1)^{N(\sigma)}a_{\sigma(1)1}\dots a_{\sigma(n)n}\right]\left[\sum_{\tau\in\mathfrak{S}_n}(-1)^{N(\tau)}b_{\tau(1)1}\dots b_{\tau(n)n}\right]\\
&= \det(A)\det(B).
\end{aligned}$$

Corollaire 3.6 *1. Soit $A\in\mathcal{M}_n(\mathbb{K})$. La matrice A est inversible si et seulement si $\det(A)\neq 0$ et dans ce cas,*

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

2. Deux matrices semblables ont même déterminant.

Démonstration.

1. — Si A est inversible, il existe une matrice B telle que $AB=I_n$. On a alors $\det(A)\det(B)=\det(AB)=\det(I_n)=1$ ce qui montre que $\det(A)\neq 0$. En fait $B=A^{-1}$ et on voit par ce raisonnement que $\det(A^{-1})=1/\det(A)$.
   — Si A n'est pas inversible, $\operatorname{rang}(A)=r<n$ et A est équivalente à la matrice diagonale $A'=\left(\begin{array}{c|c} I_r & O_{r\,n-r}\\ \hline O_{n-r\,r} & O_{n-r}\end{array}\right)$: il existe des matrices carrées d'ordre n inversibles P et Q telles que $A=QA'P^{-1}$. On a alors $\det(A)=\det(Q)\det(A')\det(P)^{-1}=0$ puisque $\det(A')=0$.
2. Si A et B sont semblables, il existe une matrice carrée d'ordre n inversible P telle que $A=PBP^{-1}$. On a alors $\det(A)=\det(P)\det(B)\det(P)^{-1}=\det(B)$. □

Remarque.

Une autre démonstration de la caractérisation d'inversibilité en terme de non-nullité du déterminant peut se faire en interprétant A comme étant la matrice des coordonnées d'un système de n vecteurs $\mathcal{S}$ relativement à une base $\mathcal{B}$ d'un $\mathbb{K}$-espace vectoriel de dimension n. Dans ce contexte, A est inversible si et seulement si $\mathcal{S}$ est une base et l'on a vu (théorème 3.2) que cette dernière assertion est équivalente à $\det_{\mathcal{B}}(\mathcal{S}) \neq 0$, soit encore $\det(A) \neq 0$.

3.3 Déterminant d'un endomorphisme

Définition 3.7 *Le déterminant d'un endomorphisme f de E est le déterminant de sa matrice relativement à une base (quelconque) $\mathcal{B}$ de E :*

$$\det(f) = \det_{\mathcal{B}}(\mathrm{Mat}(f;\mathcal{B})) = \det_{\mathcal{B}}(f(\mathcal{B})).$$

Cette définition est cohérente car l'égalité ci-dessus s'avère être indépendante de la base $\mathcal{B}$. En effet, si $\mathcal{B}'$ est une autre base de E, on sait que les matrices de f relativement aux bases $\mathcal{B}$ et $\mathcal{B}'$ sont semblables : il existe une matrice de passage P telle que $\mathrm{Mat}(f;\mathcal{B}) = P\mathrm{Mat}(f;\mathcal{B}')P^{-1}$, elles ont donc même déterminant d'après le corollaire 3.6. Les propriétés des proposition 3.5 et corollaire 3.6 s'étendent sans difficulté au cas des endomorphismes.

Proposition 3.8 *Soit f et g des endomorphismes de E. On a*

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

L'endomorphisme f est un automorphisme de E si et seulement si $\det(f) \neq 0$ *et dans ce cas*

$$\det(f^{-1}) = \frac{1}{\det(f)}.$$

3.4 Calcul d'un déterminant

3.4.1 Développement par rapport à une ligne ou une colonne

Théorème 3.9 (Développement par rapport à la j^{e} colonne) *On a pour n'importe quel $j \in \{1,\ldots,n\}$*

$$\det(A) = \sum_{i=1}^{n} a_{ij}A_{ij} \quad \textit{avec} \quad A_{ij} = (-1)^{i+j}\det\left((a_{i'j'})_{\substack{1\leqslant i',j'\leqslant n\\ i'\neq i, j'\neq j}}\right).$$

Le scalaire A_{ij} est le *cofacteur* du terme a_{ij} dans ce développement et le déterminant $M_{ij} = \det\left((a_{i'j'})_{\substack{1\leqslant i',j'\leqslant n\\ i'\neq i,j'\neq j}}\right)$ est le *mineur* relatif à a_{ij}. Ce dernier est obtenu en retirant à $\det(A)$ les i^{e} ligne et j^{e} colonne. Cofacteurs et mineurs sont reliés par $A_{ij} = (-1)^{i+j}M_{ij}$.

Démonstration.

En utilisant la linéarité du déterminant par rapport à la j^{e} variable (j fixé) ainsi que le caractère alterné, on obtient

$$\begin{aligned}\det(A) &= \sum_{i=1}^{n} a_{ij}\det\nolimits_{\mathcal{B}}(\vec{u}_1,\dots,\vec{u}_{j-1},\vec{e}_i,\vec{u}_{j+1},\dots,\vec{u}_n)\\ &= \sum_{i=1}^{n}(-1)^{n-j}a_{ij}\det\nolimits_{\mathcal{B}}(\vec{u}_1,\dots,\vec{u}_{j-1},\vec{u}_{j+1},\dots,\vec{u}_n,\vec{e}_i).\end{aligned}$$

Le déterminant dans la somme précédente peut s'écrire en modifiant l'ordre des vecteurs de la base $\mathcal{B}$ comme suit :

$$\det\nolimits_{\mathcal{B}}(\vec{u}_1,\dots,\vec{u}_{j-1},\vec{u}_{j+1},\dots,\vec{u}_n,\vec{e}_i) = (-1)^{n-i}\det\nolimits_{\mathcal{B}_i}(\vec{u}_1,\dots,\vec{u}_{j-1},\vec{u}_{j+1},\dots,\vec{u}_n,\vec{e}_i)$$

où $\mathcal{B}_i = \{\vec{e}_1,\dots,\vec{e}_{i-1},\vec{e}_{i+1},\dots,\vec{e}_n,\vec{e}_i\}$, et donc

$$\det(A) = \sum_{i=1}^{n}(-1)^{i+j}a_{ij}M_{ij} \quad \text{avec} \quad M_{ij} = \det\nolimits_{\mathcal{B}_i}(\vec{u}_1,\dots,\vec{u}_{j-1},\vec{u}_{j+1},\dots,\vec{u}_n,\vec{e}_i).$$

On a plus précisément,

$$\begin{array}{ccc}
 & \begin{array}{ccccccc} \vec{u}_1 & \dots & \vec{u}_{j-1} & \vec{u}_{j+1} & \dots & \vec{u}_n & \vec{e}_i \end{array} & \\
M_{ij} = & \begin{vmatrix} a_{11} & \dots & a_{1\,j-1} & a_{1\,j+1} & \dots & a_{1n} & 0\\ \vdots & & \vdots & \vdots & & \vdots & \vdots\\ a_{i-1\,1} & \dots & a_{i-1\,j-1} & a_{i-1\,j+1} & \dots & a_{i-1\,n} & 0\\ a_{i+1\,1} & \dots & a_{i+1\,j-1} & a_{i+1\,j+1} & \dots & a_{i+1\,n} & 0\\ \vdots & & \vdots & \vdots & & \vdots & \vdots\\ a_{n1} & \dots & a_{n\,j-1} & a_{n\,j+1} & \dots & a_{nn} & 0\\ a_{i1} & \dots & a_{i\,j-1} & a_{i\,j+1} & \dots & a_{in} & 1 \end{vmatrix} & \begin{array}{c} \vec{e}_1\\ \vdots\\ \vec{e}_{i-1}\\ \vec{e}_{i+1}\\ \vdots\\ \vec{e}_n\\ \vec{e}_i \end{array} .
\end{array}$$

Le déterminant étant une forme alternée, A_{ij} est invariant lorsqu'on rajoute aux vecteurs $\vec{u}_1,\dots,\vec{u}_{j-1},\vec{u}_{j+1},\dots,\vec{u}_n$ une combinaison linéaire de $\vec{e}_i$; en introduisant les vecteurs

$$\vec{u}'_1 = \vec{u}_1 - a_{i1}\vec{e}_i,\dots,\vec{u}'_{j-1} = \vec{u}_{j-1} - a_{i\,j-1}\vec{e}_i, \vec{u}'_{j+1} = \vec{u}_{j+1} - a_{i\,j+1}\vec{e}_i,\dots,\vec{u}'_n = \vec{u}_n - a_{in}\vec{e}_i,$$

on a

$$M_{ij} = \det\nolimits_{\mathcal{B}_i}(\vec{u}'_1,\dots,\vec{u}'_{j-1},\vec{u}'_{j+1},\dots,\vec{u}'_n,\vec{e}_i).$$

Ceci s'écrit plus clairement selon

$$\begin{array}{ccc}
 & \begin{array}{ccccccc} \vec{u}'_1 & \dots & \vec{u}'_{j-1} & \vec{u}'_{j+1} & \dots & \vec{u}'_n & \vec{e}_i \end{array} & \\
M_{ij} = & \begin{vmatrix} a_{11} & \dots & a_{1\,j-1} & a_{1\,j+1} & \dots & a_{1n} & 0\\ \vdots & & \vdots & \vdots & & \vdots & \vdots\\ a_{i-1\,1} & \dots & a_{i-1\,j-1} & a_{i-1\,j+1} & \dots & a_{i-1\,n} & 0\\ a_{i+1\,1} & \dots & a_{i+1\,j-1} & a_{i+1\,j+1} & \dots & a_{i+1\,n} & 0\\ \vdots & & \vdots & \vdots & & \vdots & \vdots\\ a_{n1} & \dots & a_{n\,j-1} & a_{n\,j+1} & \dots & a_{nn} & 0\\ 0 & \dots & 0 & 0 & \dots & 0 & 1 \end{vmatrix} & \begin{array}{c} \vec{e}_1\\ \vdots\\ \vec{e}_{i-1}\\ \vec{e}_{i+1}\\ \vdots\\ \vec{e}_n\\ \vec{e}_i \end{array}
\end{array}$$

soit encore, en changeant de notations,

$$M_{ij} = \begin{vmatrix} b_{11} & \dots & b_{1\,n-1} & 0\\ \vdots & & \vdots & \vdots\\ b_{n-1\,1} & \dots & b_{n-1\,n-1} & 0\\ 0 & \dots & 0 & 1 \end{vmatrix}.$$

En invoquant la formule sommatoire du déterminant et en observant que la dernière colonne de M_{ij} ne comporte qu'un seul terme non nul (qui vaut 1), il vient

$$M_{ij} = \sum_{\sigma \in \mathfrak{S}_n : \sigma(n)=n} (-1)^{N(\sigma)} b_{\sigma(1)1} \dots b_{\sigma(n-1)n-1}.$$

La somme ci-dessus porte sur les permutations de $\{1, \dots, n\}$ laissant n invariant : $\sigma(n) = n$ puisque $b_{\sigma(n)n} = 0$ si $\sigma(n) \neq n$ et $b_{\sigma(n)n} = 1$ si $\sigma(n) = n$ (les $b_{\sigma(n)n}$ sont les termes de la dernière colonne du dernier déterminant). Cela revient à ne prendre en compte que les permutations de $\{1, \dots, n-1\}$:

$$M_{ij} = \sum_{\sigma \in \mathfrak{S}_{n-1}} (-1)^{N(\sigma)} b_{\sigma(1)1} \dots b_{\sigma(n-1)\,n-1},$$

et l'on reconnaît en cette dernière somme le déterminant d'ordre $n-1$

$$M_{ij} = \begin{vmatrix} a_{11} & \dots & a_{1\,j-1} & a_{1\,j+1} & \dots & a_{1n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{i-1\,1} & \dots & a_{i-1\,j-1} & a_{i-1\,j+1} & \dots & a_{i-1\,n} \\ a_{i+1\,1} & \dots & a_{i+1\,j-1} & a_{i+1\,j+1} & \dots & a_{i+1\,n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{n1} & \dots & a_{n\,j-1} & a_{n\,j+1} & \dots & a_{nn} \end{vmatrix} = \det\Big((a_{i'j'})_{\substack{1\leqslant i',j'\leqslant n \\ i'\neq i, j'\neq j}} \Big).$$

On obtient bien le développement requis. □

Par transposition, puisque $\det({}^tA) = \det(A)$, on peut également développer $\det(A)$ par rapport à une ligne quelconque.

Théorème 3.10 (Développement par rapport à la i^{e} ligne) *On a pour n'importe quel $i \in \{1, \dots, n\}$*

$$\det(A) = \sum_{j=1}^{n} a_{ij} A_{ij} \quad \textit{avec} \quad A_{ij} = (-1)^{i+j} \det\Big((a_{i'j'})_{\substack{1\leqslant i',j'\leqslant n \\ i'\neq i, j'\neq j}} \Big).$$

EXEMPLE.

Déterminant d'ordre 3 : considérons la matrice $A = \begin{pmatrix} x & x' & x'' \\ y & y' & y'' \\ z & z' & z'' \end{pmatrix}$ et développons son déterminant par rapport à sa première colonne :

$$\det(A) = \alpha x + \beta y + \gamma z \quad \text{avec} \quad \alpha = \begin{vmatrix} y' & y'' \\ z' & z'' \end{vmatrix},\ \beta = -\begin{vmatrix} x' & x'' \\ z' & z'' \end{vmatrix},\ \gamma = \begin{vmatrix} x' & x'' \\ y' & y'' \end{vmatrix}.$$

On récupère la formule déjà obtenue en préliminaire de ce chapitre. À titre de vérification, développons maintenant ce déterminant par rapport à sa première ligne :

$$\det(A) = \alpha x + \alpha' x' + \alpha'' x'' \quad \text{avec} \quad \alpha = \begin{vmatrix} y' & y'' \\ z' & z'' \end{vmatrix},\ \alpha' = -\begin{vmatrix} y & y'' \\ z & z'' \end{vmatrix},\ \alpha'' = \begin{vmatrix} y & y' \\ z & z' \end{vmatrix}.$$

Un simple calcul montre l'égalité des deux résultats.

3.4.2 Méthode des zéros échelonnés

Une méthode fortement recommandée consiste à remplacer certaines lignes ou colonnes par des combinaisons linéaires d'autres lignes ou colonnes de manière à faire apparaître des zéros à l'aide de la méthode des zéros échelonnés. Cette méthode repose sur les deux résultats suivants qui découlent aisément de la propriété de n-linéarité et du caractère alterné du déterminant ; on pourra rapprocher ces deux résultats de ceux énoncés dans les propositions 4.21 et 4.22 du chapitre 1. On obtient à la fin de cette procédure le déterminant d'une matrice triangulaire qui est très simple à évaluer.

Proposition 3.11 *Soit $\{\vec{u}_1,\dots,\vec{u}_n\}$ un système de vecteurs de E, $\alpha_1,\dots,\alpha_{i-1},\alpha_{i+1},\dots,\alpha_n \in \mathbb{K}$ et $\vec{u}_i' = \vec{u}_i + \sum_{\substack{1\leqslant j\leqslant n\\ j\neq i}} \alpha_j\vec{u}_j$ pour un i fixé. Le déterminant du système de vecteurs $\{\vec{u}_1,\dots,\vec{u}_n\}$ est inchangé lorsqu'on remplace $\vec{u}_i$ par $\vec{u}_i'$:*

$$\det_{\mathcal{B}}(\vec{u}_1,\dots,\vec{u}_i,\dots,\vec{u}_n) = \det_{\mathcal{B}}(\vec{u}_1,\dots,\vec{u}_i',\dots,\vec{u}_n).$$

Démonstration.

En utilisant la linéarité par rapport à la i^{e} variable de $\det_{\mathcal{B}}$, on obtient

$$\det_{\mathcal{B}}(\vec{u}_1,\dots,\vec{u}_i',\dots,\vec{u}_n) = \det_{\mathcal{B}}(\vec{u}_1,\dots,\vec{u}_i,\dots,\vec{u}_n) + \sum_{\substack{1\leqslant j\leqslant n\\ j\neq i}} \alpha_j \det_{\mathcal{B}}(\vec{u}_1,\dots,\underset{\substack{\uparrow\\ i^{\text{e}}\text{ place}}}{\vec{u}_j},\dots,\vec{u}_n).$$

Le système de vecteurs $\{\vec{u}_1,\dots,\vec{u}_{i-1},\vec{u}_j,\vec{u}_{i+1},\dots,\vec{u}_n\}$ comportant au moins deux fois le vecteur $\vec{u}_j$ lorsque $j\neq i$, le déterminant figurant dans la somme précédente s'annule et le résultat annoncé s'ensuit aussitôt. □

En appliquant ce principe plusieurs fois, on obtient le suivant qui repose sur une méthode de zéros échelonnés avec des pivots pris égaux à 1.

Proposition 3.12 *Soit $(\alpha_{ij})_{1\leqslant i<j\leqslant n}$ un tableau triangulaire supérieur de scalaires et $\vec{u}_j' = \vec{u}_j + \sum_{i=1}^{j-1} \alpha_{ij}\vec{u}_i$ pour $1\leqslant j\leqslant n$ ($\vec{u}_1' = \vec{u}_1$). Alors les déterminants des systèmes de vecteurs $\{\vec{u}_1,\dots,\vec{u}_n\}$ et $\{\vec{u}_1',\dots,\vec{u}_n'\}$ sont identiques :*

$$\det_{\mathcal{B}}(\vec{u}_1,\dots,\vec{u}_n) = \det_{\mathcal{B}}(\vec{u}_1',\dots,\vec{u}_n').$$

Ce résultat subsiste avec un tableau triangulaire inférieur de scalaires $(\alpha_{ij})_{1\leqslant j<i\leqslant n}$ et les vecteurs $\vec{u}_j' = \vec{u}_j + \sum_{i=j+1}^{n} \alpha_{ij}\vec{u}_i$ pour $1\leqslant j\leqslant n$ ($\vec{u}_n' = \vec{u}_n$).

Remarques.

1. On ne peut pas remplacer simultanément les vecteurs $\vec{u}_1,\dots,\vec{u}_n$ par n'importe quelles combinaisons linéaires de ces vecteurs. Par exemple, si on choisit $\vec{u}_1' = \vec{u}_1+\vec{u}_2$ et $\vec{u}_2' = \vec{u}_2+\vec{u}_1$ on aurait bien sûr $\det_{\mathcal{B}}(\vec{u}_1',\dots,\vec{u}_n') = 0$ alors que $\det_{\mathcal{B}}(\vec{u}_1,\dots,\vec{u}_n)$ pourrait ne pas être nul.
2. Si on prend des pivots différents de 1, la valeur du déterminant s'en trouve modifiée. Par exemple, si on choisit $\vec{u}_1' = 2\vec{u}_1+\vec{u}_2+\cdots+\vec{u}_n, \vec{u}_2' = \vec{u}_2,\dots,\vec{u}_n' = \vec{u}_n$, alors

$$\det_{\mathcal{B}}(\vec{u}_1',\dots,\vec{u}_n') = 2\det_{\mathcal{B}}(\vec{u}_1,\dots,\vec{u}_n).$$

Exemple.

Soit la matrice $A = \begin{pmatrix} 1 & 1 & -1 \\ 0 & -1 & 1 \\ -1 & 2 & 0 \end{pmatrix}$. En écrivant dans un contexte vectoriel évident $A = \mathrm{Mat}(\{\vec{u}_1, \vec{u}_2, \vec{u}_3\}; \mathcal{B})$ et en choisissant $\vec{u}_2' = \vec{u}_2 - \vec{u}_1$ et $\vec{u}_3' = \vec{u}_3 + \vec{u}_1$, on obtient d'abord

$$A = \det_{\mathcal{B}}(\vec{u}_1, \vec{u}_2', \vec{u}_3') = \begin{vmatrix} 1 & 0 & 0 \\ 0 & -1 & 1 \\ -1 & 3 & -1 \end{vmatrix}.$$

Puis en prenant $\vec{u}_3'' = \vec{u}_3' + \vec{u}_2'$, il vient

$$A = \det_{\mathcal{B}}(\vec{u}_1, \vec{u}_2', \vec{u}_3'') = \begin{vmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ -1 & 2 & 2 \end{vmatrix} = -2.$$

4 Applications

4.1 Inverse d'une matrice carrée

On dispose pour une matrice inversible d'une représentation explicite de son inverse en termes de déterminants. Cela étant, cette représentation n'est exploitable que dans les cas les plus simples, notamment lorsque la matrice étudiée comporte beaucoup de termes nuls.

Proposition 4.1 *Soit A une matrice inversible (donc $\det(A) \neq 0$). Son inverse est donné par*

$$A^{-1} = \frac{1}{\det(A)}\, {}^t(\mathrm{co}(A))) \quad \textit{avec} \quad \mathrm{co}(A) = (A_{ij})_{1 \leqslant i,j \leqslant n}.$$

La matrice $\mathrm{co}(A)$ est la matrice des cofacteurs de A, c'est la comatrice de A.

Démonstration.

On va vérifier que $\frac{1}{\det(A)}\, {}^t(\mathrm{co}(A)))$ est l'inverse de A en effectuant le produit $A \times {}^t(\mathrm{co}(A))$ pour montrer que $A \times \frac{1}{\det(A)}\, {}^t(\mathrm{co}(A)) = I_n$. Cela revient à vérifier que

$$\sum_{k=1}^{n} a_{ik} A_{jk} = \begin{cases} \det(A) & \text{si } i = j \\ 0 & \text{si } i \neq j \end{cases}.$$

Le cas $i = j$ correspond au développement de $\det(A)$ par rapport à la i^{e} ligne comme cela a été vu dans le paragraphe précédent :

$$\sum_{k=1}^{n} a_{ik} A_{ik} = \det(A).$$

Lorsque $i \neq j$, la somme $\sum_{k=1}^{n} a_{ik} A_{jk}$ représente le développement par rapport à sa i^{e} ligne du déterminant de la matrice déduite de A en remplaçant sa j^{e} ligne par sa i^{e}. Cette dernière matrice a donc deux lignes identiques (les i^{e} et j^{e}), elle n'est pas inversible et son déterminant est nul :

$$\sum_{k=1}^{n} a_{ik}A_{jk} = \sum_{k=1}^{n}(-1)^{j+k}a_{ik}\begin{vmatrix} a_{11} & \dots & a_{1\,k-1} & a_{1\,k+1} & \dots & a_{1n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{j-1\,1} & \dots & a_{j-1\,k-1} & a_{j-1\,k+1} & \dots & a_{j-1\,n} \\ a_{j+1\,1} & \dots & a_{j+1\,k-1} & a_{j+1\,k+1} & \dots & a_{j+1\,n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{n1} & \dots & a_{n\,k-1} & a_{n\,k+1} & \dots & a_{nn} \end{vmatrix}$$

$$= \begin{vmatrix} a_{11} & \dots & a_{1\,k-1} & a_{1k} & a_{1\,k+1} & \dots & a_{1n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{i1} & \dots & a_{i\,k-1} & a_{ik} & a_{i\,k+1} & \dots & a_{in} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{j-1\,1} & \dots & a_{j-1\,k-1} & a_{j-1\,k} & a_{j-1\,k+1} & \dots & a_{j-1\,n} \\ a_{i1} & \dots & a_{i\,k-1} & a_{ik} & a_{i\,k+1} & \dots & a_{in} \\ a_{j+1\,1} & \dots & a_{j+1\,k-1} & a_{j+1\,k} & a_{j+1\,k+1} & \dots & a_{j+1\,n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & \dots & a_{n\,k-1} & a_{nk} & a_{n\,k+1} & \dots & a_{nn} \end{vmatrix} = 0. \qquad \square$$

Exemple.

Soit $A = \begin{pmatrix} 1 & 0 & -1 \\ 0 & -1 & 1 \\ -1 & 2 & 0 \end{pmatrix}$. La liste des cofacteurs de A est la suivante :

$$A_{11} = \begin{vmatrix} -1 & 1 \\ 2 & 0 \end{vmatrix} = -2, \qquad A_{12} = -\begin{vmatrix} 0 & 1 \\ -1 & 0 \end{vmatrix} = -1, \quad A_{13} = \begin{vmatrix} 0 & -1 \\ -1 & 2 \end{vmatrix} = -1,$$

$$A_{21} = -\begin{vmatrix} 0 & -1 \\ 2 & 0 \end{vmatrix} = -2, \quad A_{22} = \begin{vmatrix} 1 & -1 \\ -1 & 0 \end{vmatrix} = -1, \quad A_{23} = -\begin{vmatrix} 1 & 0 \\ -1 & 2 \end{vmatrix} = -2,$$

$$A_{31} = \begin{vmatrix} 0 & -1 \\ -1 & 1 \end{vmatrix} = -1, \quad A_{32} = -\begin{vmatrix} 1 & -1 \\ 0 & 1 \end{vmatrix} = -1, \quad A_{33} = \begin{vmatrix} 1 & 0 \\ 0 & -1 \end{vmatrix} = -1.$$

On obtient alors le déterminant de A par développement par rapport à la première ligne par exemple :

$$\det(A) = a_{11}A_{11} + a_{12}A_{12} + a_{13}A_{13} = -1.$$

Ce déterminant est non nul et la matrice A est alors inversible. On peut vérifier à titre d'exercice que les développements par rapport aux autres lignes donnent le même résultat :

$$a_{21}A_{21} + a_{22}A_{22} + a_{23}A_{23} = a_{31}A_{31} + a_{32}A_{32} + a_{33}A_{33} = -1,$$

ainsi que ceux par rapport aux colonnes :

$$a_{11}A_{11} + a_{21}A_{21} + a_{31}A_{31} = a_{12}A_{12} + a_{22}A_{22} + a_{32}A_{32} = a_{13}A_{13} + a_{23}A_{23} + a_{33}A_{33} = -1.$$

Écrivons ensuite la comatrice de A :

$$\operatorname{co}(A) = \begin{pmatrix} -2 & -1 & -1 \\ -2 & -1 & -2 \\ -1 & -1 & -1 \end{pmatrix}$$

et enfin la matrice inverse de A :

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

4.2 Rang d'une matrice et déterminants

D'après le corollaire 2.15, on a la caractérisation suivante du rang en termes de mineurs extraits.

Proposition 4.2 *Le rang d'une matrice est égal à l'ordre maximum d'un mineur non nul extrait de cette matrice :*

$$\mathrm{rang}(A) = \max\{\rho : \det(A') \neq 0, A' \in \mathcal{M}_\rho(\mathbb{K}) \text{ extraite de } A\}.$$

Exemple.

Considérons la matrice de type 4×5

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

Les mineurs d'ordre 4 de A sont les déterminants

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

On vérifie qu'ils sont tous nuls et la matrice A a un rang strictement inférieur à 4. Il est en fait inutile de calculer la totalité des mineurs d'ordre 4, il suffit de calculer les mineurs d'ordre 4 « glissants » car, si ceux-ci sont nuls, tous les autres le sont également. Dans cet exemple, ce sont les déterminants suivants :

$$\begin{vmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & -1 \\ 1 & -1 & 0 & 0 \\ 1 & -1 & 0 & 1 \end{vmatrix} \quad \text{et} \quad \begin{vmatrix} 0 & 1 & 1 & 1 \\ 1 & 1 & -1 & 2 \\ -1 & 0 & 0 & 3 \\ -1 & 0 & 1 & 1 \end{vmatrix}.$$

Par ailleurs, le mineur de A d'ordre 3 $\begin{vmatrix} 1 & 1 & 1 \\ 1 & -1 & 2 \\ 0 & 0 & 3 \end{vmatrix}$ est clairement non nul et le rang de A est supérieur ou égal à 3. En conclusion, $\mathrm{rang}(A) = 3$.

Application : Représentation cartésienne d'un sous-espace vectoriel.

Soit E un $\mathbb{K}$-espace vectoriel de dimension n, $\{\vec{e}_1, \dots, \vec{e}_n\}$ une base de E et F le sous-espace vectoriel de E engendré par le système libre de vecteurs $\{\vec{u}_1, \dots, \vec{u}_r\}$ donnés par

$$\begin{cases} \vec{u}_1 = a_{11}\,\vec{e}_1 + \cdots + a_{n1}\,\vec{e}_n \\ \quad\vdots \\ \vec{u}_r = a_{1r}\,\vec{e}_1 + \cdots + a_{nr}\,\vec{e}_n \end{cases}.$$

Pour obtenir une représentation cartésienne de $F = \mathrm{vect}(\{\vec{u}_1, \dots, \vec{u}_r\})$, on utilise la caractérisation suivante :

$$\vec{u} \in F \Longleftrightarrow \mathrm{rang}\{\vec{u}_1, \dots, \vec{u}_r, \vec{u}\} = r \Longleftrightarrow \mathrm{rang}\begin{pmatrix} a_{11} & \dots & a_{1r} & x_1 \\ \vdots & & \vdots & \vdots \\ a_{n1} & \dots & a_{nr} & x_n \end{pmatrix} = r.$$

Examinons d'abord deux cas particuliers très importants.

— *Cas d'un hyperplan vectoriel $\mathcal{H}$:* on a $\dim(\mathcal{H}) = r = n-1$. La matrice $\begin{pmatrix} a_{11} & \dots & a_{1\,n-1} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{n\,n-1} \end{pmatrix}$ étant celle d'une base de $\mathcal{H}$ est de rang $n-1$. On a alors

$$\mathrm{rang}\begin{pmatrix} a_{11} & \dots & a_{1\,n-1} & x_1 \\ \vdots & & \vdots & \vdots \\ a_{n1} & \dots & a_{n\,n-1} & x_n \end{pmatrix} = r \Longleftrightarrow \begin{vmatrix} a_{11} & \dots & a_{1\,n-1} & x_1 \\ \vdots & & \vdots & \vdots \\ a_{n1} & \dots & a_{n\,n-1} & x_n \end{vmatrix} = 0.$$

En développant le déterminant ci-dessus par rapport à la dernière colonne, on voit que $\mathcal{H}$ peut être représenté par une seule équation cartésienne qui a la forme suivante :

$$\boxed{A_1x_1 + A_2x_2 + \dots + A_nx_n = 0.}$$

— *Cas d'une droite vectorielle $\mathcal{D}$:* soit $\mathcal{D}$ la droite engendrée par le vecteur $a_1\vec{e}_1 + \dots + a_n\vec{e}_n$. On a $\dim(\mathcal{D}) = r = 1$ et $\mathcal{D}$ est caractérisée par les équivalences suivantes (en ne considérant que les mineurs « glissants ») :

$$\vec{u} \in \mathcal{D} \Longleftrightarrow \mathrm{rang}\begin{pmatrix} a_1 & x_1 \\ \vdots & \vdots \\ a_n & x_n \end{pmatrix} = 1 \Longleftrightarrow \forall k \in \{1, \dots, n-1\}, \begin{vmatrix} a_k & x_k \\ a_{k+1} & x_{k+1} \end{vmatrix} = 0.$$

Elle admet donc une représentation cartésienne à $n-1$ équations :

$$\boxed{\left\{\begin{array}{rcl} a_1\,x_2 - \quad a_2\,x_1 & = & 0 \\ a_2\,x_3 - \quad a_3\,x_2 & = & 0 \\ & \vdots & \\ a_{n-1}\,x_n - a_n\,x_{n-1} & = & 0 \end{array}\right.}\,.$$

Reprenons le cas général. La matrice $\begin{pmatrix} a_{11} & \dots & a_{1r} & x_1 \\ \vdots & & \vdots & \vdots \\ a_{n1} & \dots & a_{nr} & x_n \end{pmatrix}$ est de type $n \times (r+1)$ avec $r \leqslant n$, et la matrice extraite $\begin{pmatrix} a_{11} & \dots & a_{1r} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nr} \end{pmatrix}$ étant celle d'un système libre est de rang r. Ainsi, d'après la proposition 4.2, la matrice $\begin{pmatrix} a_{11} & \dots & a_{1r} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nr} \end{pmatrix}$ est de rang r si et seulement si tous les mineurs d'ordre $r+1$ (donc contenant toutes les colonnes) sont nuls. En termes mathématiques, cela s'écrit

$$\mathrm{rang}\begin{pmatrix} a_{11} & \dots & a_{1r} & x_1 \\ \vdots & & \vdots & \vdots \\ a_{n1} & \dots & a_{nr} & x_n \end{pmatrix} = r \Longleftrightarrow \forall \sigma \in \mathfrak{S}_{r+1,n}, \begin{vmatrix} a_{\sigma(1)1} & \dots & a_{\sigma(1)r} & x_{\sigma(1)} \\ \vdots & & \vdots & \vdots \\ a_{\sigma(r+1)1} & \dots & a_{\sigma(r+1)r} & x_{\sigma(r+1)} \end{vmatrix} = 0$$

où $\mathfrak{S}_{r+1,n}$ est l'ensemble des applications de $\{1,\ldots,r+1\}$ dans $\{1,\ldots,n\}$ strictement croissantes. En fait, comme cela a été signalé plus haut, il est inutile d'envisager tous les mineurs d'ordre $r+1$, il suffit seulement de garder en mémoire les $n-r$ mineurs d'ordre $r+1$ glissants, ce qui conduit à la représentation cartésienne de F à $n-r$ équations suivante :

$$\vec{u} \in F \iff \forall k \in \{1,\ldots,n-r\}, \begin{vmatrix} a_{k\,1} & \ldots & a_{k\,r} & x_k \\ \vdots & & \vdots & \vdots \\ a_{k+r\,1} & \ldots & a_{k+r\,r} & x_{k+r} \end{vmatrix} = 0.$$

4.3 Systèmes d'équations linéaires

Le déterminant est un outil théorique important pour l'étude qualitative des systèmes linéaires d'équations. Il permet également dans certains cas de résoudre ce type de système. Les théories de Cramer et de Rouché-Fontené présentées ci-après vont dans ce sens. Cela étant, de nombreuses situations conduisent à des systèmes linéaires comportant de très grands nombres d'équations et d'inconnues, et des méthodes numériques lui seront avantageusement préférées. Une telle méthode (méthode de Gauss) est présentée à la fin de ce paragraphe.

Considérons le système linéaire avec second membre de p équations à n inconnues suivant :

$$\text{(S)} \qquad \begin{cases} a_{11}\,x_1 + a_{12}\,x_2 + \cdots + a_{1n}\,x_n &= b_1 \\ a_{21}\,x_1 + a_{22}\,x_2 + \cdots + a_{2n}\,x_n &= b_2 \\ & \vdots \\ a_{p1}\,x_1 + a_{p2}\,x_2 + \cdots + a_{pn}\,x_n &= b_p \end{cases}$$

dans lequel les coefficients a_{ij}, $1 \leqslant i \leqslant p, 1 \leqslant j \leqslant n$, et b_i, $1 \leqslant i \leqslant p$, sont des scalaires de $\mathbb{K}$ donnés et les $x_1,\ldots,x_n$ sont des scalaires inconnus. L'inconnue du système (S) est en fait le n-uplet $(x_1,\ldots,x_n)$. Deux questions se posent :
— le système (S) admet-il des solutions, et dans l'affirmative, combien en admet-il ? On verra qu'il y a trois possibilités : (S) admet zéro, une seule ou une infinité de solutions ;
— quelle est la solution de (S) lorsqu'elle est unique, et quelle est la forme générale des solutions de (S) lorsqu'il y en a une infinité ?

Lorsque les $b_1,\ldots,b_p$ sont tous nuls, on dit que (S) est un *système homogène*. Un tel système admet au moins une solution, la solution nulle : $(x_1,\ldots,x_n) = (0,\ldots,0)$.

4.3.1 Interprétations

— *Interprétation vectorielle :* introduisons les n vecteurs $\vec{u}_1,\ldots,\vec{u}_n$ exprimés dans une base $\mathcal{B} = \{\vec{e}_1,\ldots,\vec{e}_p\}$ d'un $\mathbb{K}$-espace vectoriel E de dimension p selon

$$\begin{cases} \vec{u}_1 &= a_{11}\,\vec{e}_1 + \cdots + a_{p1}\,\vec{e}_p \\ & \vdots \\ \vec{u}_n &= a_{1n}\,\vec{e}_1 + \cdots + a_{pn}\,\vec{e}_p \end{cases}$$

ainsi que le vecteur

$$\vec{v} = b_1\,\vec{e}_1 + \cdots + b_p\,\vec{e}_p.$$

Le système (S) s'écrit

$$\sum_{j=1}^{n} x_j \vec{u}_j = \vec{v}.$$

Cette égalité signifie que le vecteur $\vec{v}$ est la combinaison linéaire des vecteurs $\vec{u}_1, \ldots, \vec{u}_n$ dont les coefficients sont les inconnues. Ainsi, (S) admet des solutions si et seulement si $\vec{v} \in \mathrm{vect}(\{\vec{u}_1, \ldots, \vec{u}_n\})$. On peut faire alors la discussion suivante relativement au système de vecteurs $\mathcal{S} = \{\vec{u}_1, \ldots, \vec{u}_n\}$:
(i) si $\mathcal{S}$ est libre ($n \leqslant p$), (S) admet au plus une solution $(x_1, \ldots, x_n)$;
(ii) si $\mathcal{S}$ est générateur de E ($n \geqslant p$), (S) admet au moins une solution $(x_1, \ldots, x_n)$. Il admet soit une solution unique, soit une infinité de solutions ;
(iii) si $\mathcal{S}$ est une base de E ($n = p$), (S) admet une solution unique $(x_1, \ldots, x_n)$.

— *Interprétation fonctionnelle :* soient E et F deux $\mathbb{K}$ espaces vectoriels rapportés à des bases respectives $\{\vec{e}_1, \ldots, \vec{e}_n\}$ et $\{\vec{\varepsilon}_1, \ldots, \vec{\varepsilon}_p\}$, $f \in \mathcal{L}(E, F)$ l'application linéaire donnée par

$$\left\{\begin{array}{rcl} f(\vec{e}_1) & = & a_{11}\,\vec{\varepsilon}_1 + \cdots + a_{p1}\,\vec{\varepsilon}_p \\ & \vdots & \\ f(\vec{e}_n) & = & a_{1n}\,\vec{\varepsilon}_1 + \cdots + a_{pn}\,\vec{\varepsilon}_p \end{array}\right.$$

et les vecteurs $\vec{u} \in E$ et $\vec{v} \in F$ donnés par

$$\vec{u} = x_1\vec{e}_1 + \cdots + x_n\vec{e}_n, \quad \vec{v} = b_1\vec{\varepsilon}_1 + \cdots + b_p\vec{\varepsilon}_p.$$

Le système (S) s'écrit, $\vec{u}$ étant le vecteur inconnu,

$$f(\vec{u}) = \vec{v}.$$

Cette égalité signifie que l'ensemble des solutions est l'ensemble des antécédents de $\vec{v}$ par f : $f^{-1}(\{\vec{v}\})$. Ainsi, (S) admet des solutions si et seulement si $\vec{v} \in \mathrm{Im}\, f$. On peut faire alors la discussion suivante relativement à la qualité de f :
(i) si f est injective ($n \leqslant p$), $\vec{v}$ admet au plus un antécédent par f et (S) admet au plus une solution $(x_1, \ldots, x_n)$;
(ii) si f est surjective ($n \geqslant p$), $\vec{v}$ admet au moins un antécédent par f et (S) admet au moins une solution $(x_1, \ldots, x_n)$. Il admet soit une solution unique, soit une infinité de solutions ;
(iii) si f est bijective ($n = p$), $\vec{v}$ admet exactement un antécédent par f et (S) admet une solution unique $(x_1, \ldots, x_n)$.

— *Interprétation matricielle :* soit $A \in \mathcal{M}_{p\times n}(\mathbb{K})$, $B \in \mathcal{M}_{p1}(\mathbb{K})$ et $X \in \mathcal{M}_{n1}(\mathbb{K})$ les matrices données par

$$A = \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \vdots \\ a_{p1} & \ldots & a_{pn} \end{pmatrix}, \quad B = \begin{pmatrix} b_1 \\ \vdots \\ b_p \end{pmatrix}, \quad X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

Le système (S) s'écrit, X étant la matrice inconnue,

$$AX = B.$$

La discussion se fait à présent relativement à la matrice A :
(i) si la matrice A est inversible ($n = p$), on a $X = A^{-1}B$ et (S) admet une solution unique $(x_1, \ldots, x_n)$;
(ii) si la matrice A n'est pas inversible, il est plus difficile de conclure.

Lorsque le système (S) est homogène, il admet une unique solution (la solution nulle en l'occurrence) si et seulement si la matrice A est inversible.

Définition 4.3 *Le rang du système linéaire* (S) *est le rang commun du système de vecteurs* $\mathcal{S}$*, de l'application linéaire* f *et de la matrice* A *:*

$$\mathrm{rang(S)} = \mathrm{rang}(\mathcal{S}) = \mathrm{rang}(f) = \mathrm{rang}(A).$$

4.3.2 Système de Cramer

Définition 4.4 *Le système linéaire* (S) *est un système de Cramer*[1] *lorsque* $\mathrm{rang(S)} = n = p$.

Le système linéaire (S) est un système de er si et seulement si le système de vecteurs $\mathcal{S}$ est une base de E, on encore l'application linéaire f est bijective, ou encore la matrice A est inversible : $\det(A) \neq 0$. Dans ce cas, il est possible d'exprimer la solution de manière explicite à l'aide de déterminants.

Supposons donc que (S) soit un système de Cramer. Calculons le déterminant du système de vecteurs déduit de $\mathcal{S}$ en remplaçant le i^{e} vecteur $\vec{u}_i$ par $\vec{v}$:

$$\begin{aligned} \det_{\mathcal{B}}(\vec{u}_1, \dots, \vec{u}_{i-1}, \vec{v}, \vec{u}_{i+1}, \dots, \vec{u}_n) &= \det_{\mathcal{B}}(\vec{u}_1, \dots, \vec{u}_{i-1}, \sum_{j=1}^{n} x_j \vec{u}_j, \vec{u}_{i+1}, \dots, \vec{u}_n) \\ &= \sum_{j=1}^{n} x_j \det_{\mathcal{B}}(\vec{u}_1, \dots, \vec{u}_{i-1}, \vec{u}_j, \vec{u}_{i+1}, \dots, \vec{u}_n). \end{aligned}$$

Lorsque $j \neq i$, le déterminant $\det_{\mathcal{B}}(\vec{u}_1, \dots, \vec{u}_{i-1}, \vec{u}_j, \vec{u}_{i+1}, \dots, \vec{u}_n)$ contient deux fois le vecteur $\vec{u}_j$ et il s'annule. On a alors

$$\det_{\mathcal{B}}(\vec{u}_1, \dots, \vec{u}_{i-1}, \vec{v}, \vec{u}_{i+1}, \dots, \vec{u}_n) = x_i \det_{\mathcal{B}}(\vec{u}_1, \dots, \vec{u}_{i-1}, \vec{u}_i, \vec{u}_{i+1}, \dots, \vec{u}_n).$$

Le système d'équations (S) étant de Cramer, le système de vecteurs $\mathcal{S}$ est une base de E et son déterminant est non nul : $\det_{\mathcal{B}}(\vec{u}_1, \dots, \vec{u}_n) \neq 0$. On obtient ainsi la solution unique de (S) :

$$\boxed{x_i = \frac{\det_{\mathcal{B}}(\vec{u}_1, \dots, \vec{u}_{i-1}, \vec{v}, \vec{u}_{i+1}, \dots, \vec{u}_n)}{\det_{\mathcal{B}}(\vec{u}_1, \dots, \vec{u}_n)},\ 1 \leqslant i \leqslant n.}$$

Exemples.

1. *Système de deux équations à deux inconnues :* soit

$$\text{(S)} \qquad \begin{cases} a\,x + b\,y = c & \text{(E)} \\ a'\,x + b'\,y = c' & \text{(E}'\text{)} \end{cases}.$$

1. Cramer, Gabriel : mathématicien suisse (Genève 1704–Bagnols-sur-Cèze 1752)

Le système (S) est de Cramer si et seulement si $\begin{vmatrix} a & b \\ a' & b' \end{vmatrix} \neq 0$. Dans ce cas, la solution s'écrit, en posant $\Delta = \begin{vmatrix} a & b \\ a' & b' \end{vmatrix}$, $\Delta_x = \begin{vmatrix} c & b \\ c' & b' \end{vmatrix}$, $\Delta_y = \begin{vmatrix} a & c \\ a' & c' \end{vmatrix}$,

$$\boxed{x = \frac{\Delta_x}{\Delta} = \frac{b'c - bc'}{ab' - a'b}, \qquad y = \frac{\Delta_y}{\Delta} = \frac{ac' - a'c}{ab' - a'b}.}$$

On retrouve ainsi la solution obtenue par la méthode élémentaire des « multiplicateurs » :

- en multipliant les équations (E) par a' et (E$'$) par a, puis en les soustrayant, on trouve $(ab' - a'b)y = ac' - a'c$;
- en multipliant les équations (E) par b' et (E$'$) par b, puis en les soustrayant, on trouve $(ab' - a'b)x = b'c - bc'$.

2. *Système de trois équations à trois inconnues :* soit

$$\text{(S)} \qquad \left\{ \begin{array}{rl} a\,x + \;\; b\,y + \;\; c\,z = & d \\ a'\,x + b'\,y + c'\,z = & d' \\ a''\,x + b''\,y + c''\,z = & d'' \end{array} \right. .$$

Le système (S) est de Cramer si et seulement si $\begin{vmatrix} a & b & c \\ a' & b' & c' \\ a'' & b'' & c'' \end{vmatrix} \neq 0$. Dans ce cas, en posant

$$\Delta = \begin{vmatrix} a & b & c \\ a' & b' & c' \\ a'' & b'' & c'' \end{vmatrix}, \quad \Delta_x = \begin{vmatrix} d & b & c \\ d' & b' & c' \\ d'' & b'' & c'' \end{vmatrix}, \quad \Delta_y = \begin{vmatrix} a & d & c \\ a' & d' & c' \\ a'' & d'' & c'' \end{vmatrix}, \quad \Delta_z = \begin{vmatrix} a & b & d \\ a' & b' & d' \\ a'' & b'' & d'' \end{vmatrix},$$

la solution s'écrit

$$\boxed{x = \frac{\Delta_x}{\Delta}, \quad y = \frac{\Delta_y}{\Delta}, \quad z = \frac{\Delta_z}{\Delta}.}$$

4.3.3 Théorie de Rouché-Fontené

Dans ce paragraphe, on se place dans le cas où la matrice A n'est pas inversible. Notons r son rang. D'après le corollaire 2.15 on peut extraire de A une sous-matrice carrée A' d'ordre r inversible. Supposons pour simplifier les notations que $A' = (a_{ij})_{1 \leqslant i,j \leqslant r}$.

— *Cas $r = p < n$:* le système (S) a plus d'inconnues que d'équations. Écrivons-le sous la forme

$$\text{(S}'\text{)} \qquad \left\{ \begin{array}{rcl} a_{11}\,x_1 + \cdots + a_{1p}\,x_p & = & b_1 - a_{1\,p+1}\,x_{p+1} - \cdots - a_{1n}\,x_n \\ & \vdots & \\ a_{p1}\,x_1 + \cdots + a_{pp}\,x_p & = & b_p - a_{p\,p+1}\,x_{p+1} - \cdots - a_{pn}\,x_n \end{array} \right. .$$

Le système (S$'$) (dit *principal*) est un système de Cramer relativement aux inconnues (dites *principales*) $x_1, \ldots, x_p$, il admet donc une unique solution $(x_1, \ldots, x_p)$ s'exprimant en fonction des autres inconnues (*non principales*) $x_{p+1}, \ldots, x_n$ que l'on peut considérer comme paramètres. Ainsi, le système (S) admet une infinité de solutions dépendant de $n - p$ paramètres.

— *Cas $r = n < p$:* le système (S) a plus d'équations que d'inconnues. Réécrivons les n premières équations (dites *principales*) du système (S) :

$$(\mathrm{S}') \qquad \begin{cases} a_{11}\,x_1 + \cdots + a_{1n}\,x_n \;=\; b_1 \\ \qquad\qquad\qquad\quad \vdots \\ a_{n1}\,x_1 + \cdots + a_{nn}\,x_n \;=\; b_n \end{cases}.$$

Le système (S′) (dit *principal*) est un système de Cramer, il admet donc une unique solution $(x_1, \dots, x_n)$. Il reste à voir si cette solution vérifie également les $p - n$ dernières équations (*non principales*) du système (S) :

$$(\mathrm{E_i}) \qquad a_{i1}\,x_1 + \cdots + a_{in}\,x_n \;=\; b_i, \quad n+1 \leqslant i \leqslant p.$$

On a l'alternative suivante :

- ou bien toutes les équations non principales sont vérifiées et le système (S) admet une solution unique. On dit que le système (S) est *compatible* ;
- ou bien l'une des équations non principales n'est pas vérifiée et le système (S) n'a pas de solution. On dit que le système (S) est *incompatible*.

La compatibilité peut bien sûr se vérifier en remplaçant directement la solution $(x_1, \dots, x_n)$ dans les équations non principales, à condition de l'avoir calculée. Néanmoins, il est possible d'opérer cette vérification sans détermination de la solution $(x_1, \dots, x_n)$. Une caractérisation d'un système compatible est la suivante. Le système (S) est compatible si et seulement si les $p - n$ déterminants suivants (appelés *bordants*) sont nuls :

$$\forall i \in \{n+1, \dots, p\}, \quad \begin{vmatrix} a_{11} & \dots & a_{1n} & b_1 \\ \vdots & & \vdots & \vdots \\ a_{n1} & \dots & a_{nn} & b_n \\ a_{i1} & \dots & a_{in} & b_i \end{vmatrix} = 0.$$

En effet, en adjoignant au système (S′) l'équation (Ei), on obtient le système

$$(\mathrm{S'_i}) \qquad \begin{cases} a_{11}\,x_1 + \cdots + a_{1n}\,x_n \;=\; b_1 \\ \qquad\qquad\qquad\quad \vdots \\ a_{n1}\,x_1 + \cdots + a_{nn}\,x_n \;=\; b_n \\ a_{i1}\,x_1 + \cdots + a_{in}\,x_n \;=\; b_i \end{cases}$$

que l'on peut interpréter, en introduisant les vecteurs (d'un $\mathbb{K}$-espace vectoriel de dimension $n+1$ de base $\mathcal{B}' = \{\vec{e}^{\,\prime}_1, \dots, \vec{e}^{\,\prime}_{n+1}\}$)

$$\begin{cases} \vec{u}^{\,\prime}_1 \;=\; a_{11}\,\vec{e}^{\,\prime}_1 + \cdots + a_{n1}\,\vec{e}^{\,\prime}_n + a_{i1}\,\vec{e}^{\,\prime}_{n+1} \\ \quad\vdots \\ \vec{u}^{\,\prime}_n \;=\; a_{1n}\,\vec{e}^{\,\prime}_1 + \cdots + a_{nn}\,\vec{e}^{\,\prime}_n + a_{in}\,\vec{e}^{\,\prime}_{n+1} \end{cases}$$

ainsi que le vecteur

$$\vec{v}^{\,\prime}_i \;=\; b_1\,\vec{e}^{\,\prime}_1 + \cdots + b_n\,\vec{e}^{\,\prime}_n + b_i\,\vec{e}^{\,\prime}_{n+1},$$

selon l'égalité vectorielle

$$\sum_{j=1}^{n} x_j \vec{u}^{\,\prime}_{ij} = \vec{v}^{\,\prime}_i.$$

Le système d'équations (S′i) est donc compatible si et seulement si le vecteur $\vec{v}^{\,\prime}_i$ est combinaison linéaire des vecteurs $\vec{u}^{\,\prime}_{i1}, \dots, \vec{u}^{\,\prime}_{in}$. Par hypothèse, la matrice A du système (S) est de rang n ainsi que la matrice principale A' ; cela assure que le système de vecteurs $\{\vec{u}^{\,\prime}_1, \dots, \vec{u}^{\,\prime}_n\}$ est de rang n, et dans ces conditions :

$$\vec{v}^{\,\prime}_i \in \mathrm{vect}(\{\vec{u}^{\,\prime}_{i1}, \dots, \vec{u}^{\,\prime}_{in}\}) \Longleftrightarrow \mathrm{rang}(\{\vec{u}^{\,\prime}_{i1}, \dots, \vec{u}^{\,\prime}_{in}, \vec{v}^{\,\prime}_i\}) = n \Longleftrightarrow \det_{\mathcal{B}'}(\vec{u}^{\,\prime}_{i1}, \dots, \vec{u}^{\,\prime}_{in}, \vec{v}^{\,\prime}_i) = 0,$$

et l'on débouche sur la caractérisation énoncée en termes de bordants.

— *Cas* $r < \min(p,n)$ *:* réécrivons les r premières équations (dites *principales*) du système (S) :

$$(\mathrm{S}')\qquad \begin{cases} a_{11}\,x_1 + \cdots + a_{1r}\,x_r \;=\; b_1 - a_{1\,r+1}\,x_{r+1} - \cdots - a_{1n}\,x_n \\ \qquad\qquad\vdots \\ a_{r1}\,x_1 + \cdots + a_{rr}\,x_r \;=\; b_r - a_{r\,r+1}\,x_{r+1} - \cdots - a_{rn}\,x_n \end{cases}.$$

Le système (S′) (dit *principal*) est un système de Cramer relativement aux r inconnues (dites *principales*) $x_1,\dots,x_r$, il admet donc une unique solution $(x_1,\dots,x_r)$ s'exprimant en fonction des autres inconnues (*non principales*) $x_{r+1},\dots,x_n$ que l'on peut considérer comme paramètres. Il reste à voir si cette solution vérifie également les $p-r$ dernières équations (*non principales*) du système (S) :

$$(\mathrm{E_i})\qquad a_{i1}\,x_1 + \cdots + a_{ir}\,x_r \;=\; b_i - a_{i\,r+1}\,x_{r+1} - \cdots - a_{in}\,x_n.$$

Comme dans le cas précédent, on a l'alternative suivante :

- ou bien toutes les équations non principales sont vérifiées et le système (S) admet une infinité de solutions dépendant de $n-r$ paramètres. Le système (S) est *compatible* ;
- ou bien l'une des équations non principales n'est pas vérifiée et le système (S) n'a pas de solution. Le système (S) est *incompatible*.

De la même manière, la compatibilité du système (S) peut se tester en calculant des bordants. Le système (S) est compatible si et seulement si les $p-r$ bordants ci-dessous sont nuls :

$$\forall i \in \{r+1,\dots,p\},\quad \begin{vmatrix} a_{11} & \dots & a_{1r} & b_1 \\ \vdots & & \vdots & \vdots \\ a_{r1} & \dots & a_{rr} & b_r \\ a_{i1} & \dots & a_{ir} & b_i \end{vmatrix} = 0.$$

En effet, introduisons les vecteurs, (d'un $\mathbb{K}$-espace vectoriel de dimension $r+1$ de base $\mathcal{B}' = \{\vec{e}_1{}',\dots,\vec{e}_{r+1}{}'\}$) pour $r+1 \leqslant i \leqslant p$,

$$\begin{cases} \vec{u}_{i1}' \;=\; a_{11}\,\vec{e}_1{}' + \cdots + a_{r1}\,\vec{e}_r{}' + a_{i1}\,\vec{e}_{r+1}{}' \\ \quad\vdots \\ \vec{u}_{ir}' \;=\; a_{1r}\,\vec{e}_1{}' + \cdots + a_{rr}\,\vec{e}_r{}' + a_{ir}\,\vec{e}_{r+1}{}' \end{cases}$$

et

$$\vec{v}_i' \;=\; b_1\,\vec{e}_1{}' + \cdots + b_r\,\vec{e}_r{}' + b_i\,\vec{e}_{r+1}{}'.$$

Le système de vecteurs $\{\vec{u}_{i1}',\dots,\vec{u}_{ir}'\}$ est de rang r et alors la compatibilité du système (S′)–(Eᵢ) se traduit par

$$\begin{aligned} &\vec{v}_i' - a_{i\,r+1}\vec{u}_{i\,r+1}' - \cdots - a_{in}\vec{u}_{in}' \in \mathrm{vect}(\{\vec{u}_{i1}',\dots,\vec{u}_{ir}'\}) \\ &\qquad\iff \mathrm{rang}(\{\vec{u}_{i1}',\dots,\vec{u}_{ir}',\vec{v}_i' - a_{i\,r+1}\vec{u}_{i\,r+1}' - \cdots - a_{in}\vec{u}_{in}'\}) = r \\ &\qquad\iff \det\nolimits_{\mathcal{B}'}(\vec{u}_{i1}',\dots,\vec{u}_{ir}',\vec{v}_i' - a_{i\,r+1}\vec{u}_{i\,r+1}' - \cdots - a_{in}\vec{u}_{in}') = 0. \end{aligned}$$

Par linéarité par rapport à la dernière variable du déterminant, on a

$$\begin{aligned} &\det\nolimits_{\mathcal{B}'}(\vec{u}_{i1}',\dots,\vec{u}_{ir}',\vec{v}_i' - a_{i\,r+1}\vec{u}_{i\,r+1}' - \cdots - a_{in}\vec{u}_{in}') \\ &\qquad = \det\nolimits_{\mathcal{B}'}(\vec{u}_{i1}',\dots,\vec{u}_{ir}',\vec{v}_i') \\ &\qquad\quad -a_{i\,r+1}\det\nolimits_{\mathcal{B}'}(\vec{u}_{i1}',\dots,\vec{u}_{ir}',\vec{u}_{i\,r+1}') - \cdots - a_{in}\det\nolimits_{\mathcal{B}'}(\vec{u}_{i1}',\dots,\vec{u}_{ir}',\vec{u}_{in}') \\ &\qquad = \det\nolimits_{\mathcal{B}'}(\vec{u}_{i1}',\dots,\vec{u}_{ir}',\vec{v}_i'). \end{aligned}$$

La dernière égalité découle du fait que la matrice A est de rang r et donc que tous les mineurs extraits de A d'ordre $r+1$, ici représentés par $\det_{\mathcal{B}'}(\vec{u}_{i1}',\dots,\vec{u}_{ir}',\vec{u}_{ij}')$, $r+1 \leqslant j \leqslant n$, sont nuls. En conclusion, la compatibilité du système (S) est équivalente à $\det_{\mathcal{B}'}(\vec{u}_{i1}',\dots,\vec{u}_{ir}',\vec{v}_i') = 0$ pour $r \leqslant i \leqslant p$, comme annoncé.

Résumons l'analyse faite.

Théorème 4.5 (Théorie de Rouché [2]-Fontené [3]) *Soit* (S) *un système de p équations à n inconnues de rang r. Alors,*

- *si $r = p = n$,* (S) *est un système de Cramer et il admet une solution unique ;*
- *si $r = p < n$, le système* (S) *admet une infinité de solutions dépendant de $n - p$ paramètres ;*
- *si $r = n < p$, soit le système* (S) *admet une unique solution, soit il n'en admet pas ;*
- *si $r < \min(p, n)$, soit le système* (S) *admet une infinité de solutions dépendant de $n-r$ paramètres, soit il n'en admet pas.*

Dans les deux derniers cas, on choisit un système principal de r équations à r inconnues et l'on évalue les $p - r$ bordants du mineur principal choisi. Le système (S) *est alors compatible si et seulement si ces $p - r$ bordants sont nuls.*

Exemples.

1. Considérons le système de trois équations à trois inconnues suivant où a est un paramètre réel :

$$(\mathrm{S}_1) \qquad \left\{\begin{array}{rcl} 3\,x + 2\,y - 5\,z &=& 1 \\ x - \,y + 5\,z &=& 0\,. \\ x + \,y - 3\,z &=& a \end{array}\right.$$

La matrice associée à (S_1) est $A_1 = \begin{pmatrix} 3 & 2 & -5 \\ 1 & -1 & 5 \\ 1 & 1 & -3 \end{pmatrix}$; elle est de rang 2. La matrice carrée d'ordre 2 extraite de A_1 : $A_1' = \begin{pmatrix} 3 & 2 \\ 1 & -1 \end{pmatrix}$, est inversible puisque $\det(A_1') = \begin{vmatrix} 3 & 2 \\ 1 & -1 \end{vmatrix} = -5 \neq 0$. On peut donc choisir x et y pour inconnues principales et les deux premières équations comme équations principales :

$$\left\{\begin{array}{rcr} 3\,x + 2\,y &=& 5\,z + 1 \\ x - \,y &=& -5\,z \end{array}\right. .$$

On a affaire à un système de Cramer d'inconnues x, y dont la solution est

$$x = \frac{\begin{vmatrix} 5z+1 & 2 \\ -5z & -1 \end{vmatrix}}{\begin{vmatrix} 3 & 2 \\ 1 & -1 \end{vmatrix}} = -z + \frac{1}{5}, \qquad y = \frac{\begin{vmatrix} 3 & 5z+1 \\ 1 & -5z \end{vmatrix}}{\begin{vmatrix} 3 & 2 \\ 1 & -1 \end{vmatrix}} = 4z + \frac{1}{5}.$$

Enfin, pour voir si le système est compatible, on calcule le bordant de A_1' (développement par rapport à la troisième colonne) :

$$\begin{vmatrix} \boxed{\begin{matrix} 3 & 2 \\ 1 & -1 \end{matrix}} & \begin{matrix} 1 \\ 0 \end{matrix} \\ \begin{matrix} 1 & 1 \end{matrix} & a \end{vmatrix} = \begin{vmatrix} 1 & -1 \\ 1 & 1 \end{vmatrix} + a \begin{vmatrix} 3 & 2 \\ 1 & -1 \end{vmatrix} = 2 - 5a.$$

Le système (S_1) est donc compatible si et seulement si le bordant ci-dessus est nul : $a = 2/5$. En conséquence, son ensemble de solutions est $\{(-z + \frac{1}{5}, 4z + \frac{1}{5}, z),\ z \in \mathbb{R}\}$ lorsque $a = 2/5$ et vide lorsque $a \neq -2/5$.

2. Rouché, Eugène : mathématicien français (Sommières 1832–Lunel 1910)
3. Fontené, Georges : mathématicien français (Rousies 1848– Paris 1923)

2. Soit le système de deux équations à trois inconnues :

$$(\mathrm{S}_2) \qquad \begin{cases} 2\,x - 3\,y + z = 2 \\ -x + y + 3\,z = 1 \end{cases}.$$

Ce système est associé à la matrice $A_2 = \begin{pmatrix} 2 & -3 & 1 \\ -1 & 1 & 3 \end{pmatrix}$ qui est de rang 2. On peut extraire de A_2 une matrice carrée d'ordre 2 inversible ; par exemple $A_2' = \begin{pmatrix} 2 & -3 \\ -1 & 1 \end{pmatrix}$. On peut donc choisir x et y pour inconnues principales et les deux équations du système sont principales :

$$\begin{cases} 2\,x - 3\,y = -z + 2 \\ -x + y = -3\,z + 1 \end{cases}.$$

Ceci est un système de Cramer d'inconnues x, y dont la solution est

$$x = \frac{\begin{vmatrix} -z+2 & -3 \\ -3z+1 & 1 \end{vmatrix}}{\begin{vmatrix} 2 & -3 \\ -1 & 1 \end{vmatrix}} = 10z - 5, \qquad y = \frac{\begin{vmatrix} 2 & -z+2 \\ -1 & -3z+1 \end{vmatrix}}{\begin{vmatrix} 2 & -3 \\ -1 & 1 \end{vmatrix}} = 7z - 4.$$

Il n'y a pas de problème de compatibilité et le système (S_2) admet une infinité de solutions données par $(10z - 5, 7z - 4, z)$, $z \in \mathbb{R}$.

3. Soit le système de trois équations à deux inconnues de paramètre réel b :

$$(\mathrm{S}_3) \qquad \begin{cases} x + 2\,y = 1 \\ x - y = 2 \\ x - 2\,y = b \end{cases}.$$

Ce système est associé à la matrice $A_3 = \begin{pmatrix} 1 & 2 \\ 1 & -1 \\ 1 & -2 \end{pmatrix}$ qui est de rang 2. On peut extraire de A_3 une matrice carrée d'ordre 2 inversible ; par exemple $A_3' = \begin{pmatrix} 1 & 2 \\ 1 & -1 \end{pmatrix}$. On peut donc choisir x et y pour inconnues principales et les deux premières équations comme équations principales :

$$\begin{cases} x + 2\,y = 1 \\ x - y = 2 \end{cases}.$$

On a affaire à un système de Cramer d'inconnues x, y dont la solution est

$$x = \frac{\begin{vmatrix} 1 & 2 \\ 2 & -1 \end{vmatrix}}{\begin{vmatrix} 1 & 2 \\ 1 & -1 \end{vmatrix}} = \frac{5}{3}, \qquad y = \frac{\begin{vmatrix} 1 & 1 \\ 1 & 2 \end{vmatrix}}{\begin{vmatrix} 1 & 2 \\ 1 & -1 \end{vmatrix}} = -\frac{1}{3}.$$

Enfin, pour voir si le système est compatible, on calcule le bordant de A_3' (développement par rapport à la troisième colonne) :

$$\begin{vmatrix} 1 & 2 & 1 \\ 1 & -1 & 2 \\ 1 & -2 & b \end{vmatrix} = \begin{vmatrix} 1 & -1 \\ 1 & -2 \end{vmatrix} - 2 \begin{vmatrix} 1 & 2 \\ 1 & -2 \end{vmatrix} + b \begin{vmatrix} 1 & 2 \\ 1 & -1 \end{vmatrix} = 7 - 3b.$$

Le système (S_3) est donc compatible si et seulement si le bordant est nul : $b = 7/3$. En conséquence, il admet une unique solution $(x, y) = (\frac{5}{3}, -\frac{1}{3})$ lorsque $b = 7/3$ et n'admet pas de solution lorsque $b \neq 7/3$.

4.3.4 Méthode de Gauss

La résolution d'un système de Cramer par la méthode des déterminants nécessite une quantité de calculs absolument faramineuse dès que la taille du système est élevée. En effet, le développement d'un déterminant d'ordre n selon sa formule sommatoire demande $n! - 1$ additions et $(n-1)n!$ produits, ce qui fait $nn! - 1$ opérations pour un seul déterminant. La méthode de Cramer pour un système de n équations à n inconnues fait intervenir $n+1$ déterminants et nécessite $(n+1)(nn!-1) \simeq n(n+1)!$ opérations. Pour $n = 2, 3, 4, 5$, cela fait respectivement $9, 68, 475, 3594$ opérations.

Évaluons ce nombre pour $n = 100$ par exemple. L'approximation de Stirling $n! \sim \sqrt{2\pi}\, n^{n+1/2} e^{-n}$ permet d'effectuer cette estimation : avec $e^{-100} = 10^{-100\log_{10} e} \simeq 10^{-43,43}$, on trouve

$$100 \times 101! \simeq 100 \times 101 \times 100^{100,5} \times 10^{-43,43}\sqrt{2\pi} \simeq 9,4 \times 10^{161}.$$

Un ordinateur fonctionnant à 100 mégaflops (effectuant 10^8 opérations par seconde) pourrait réaliser $10^8 \times 365 \times 86400 \simeq 3,1 \times 10^{15}$ opérations par an. Il faudrait environ 3×10^{146} années pour faire résoudre ce système, tout en sachant que l'Univers a quinze milliard d'années d'âge, il faudrait 2×10^{136} fois l'âge de l'Univers pour obtenir la solution...

D'où la nécessité de trouver des méthodes de résolution plus économiques en calculs. On en présente une ci-dessous : la *méthode du pivot (ou d'élimination) de Gauss*[4].

— *Principe de la méthode :* la méthode de Gauss ramène la résolution à celle d'un système triangulaire. Son principe est le même que celui de la méthode des zéros échelonnés et consiste à faire apparaître progressivement des coefficients nuls dans chaque équation à l'aide de combinaisons linéaires judicieuses des équations du système (S). Plus précisément, en supposant par exemple que $r < \min(p, n)$, on se ramène à un système (S′) équivalent à (S) de la forme

$$(\mathrm{S}') \quad \left\{\begin{array}{rcll}
a'_{11}\,x'_1 + a'_{12}\,x'_2 + a'_{13}\,x'_3 + \cdots + a'_{1r}\,x'_r + \cdots + a'_{1n}\,x'_n &=& b'_1 & (\mathrm{E}'_1)\\
a'_{22}\,x'_2 + a'_{23}\,x'_3 + \cdots + a'_{2r}\,x'_r + \cdots + a'_{2n}\,x'_n &=& b'_2 & (\mathrm{E}'_2)\\
a'_{33}\,x'_3 + \cdots + a'_{3r}\,x'_r + \cdots + a'_{3n}\,x'_n &=& b'_3 & (\mathrm{E}'_3)\\
\ddots \qquad\qquad & & \vdots & \vdots\\
a'_{rr}\,x'_r + \cdots + a'_{rn}\,x'_n &=& b'_r & (\mathrm{E}'_\mathrm{r})\\
0 &=& b'_{r+1} & (\mathrm{E}'_{\mathrm{r}+1})\\
& & \vdots & \vdots\\
0 &=& b'_p & (\mathrm{E}'_\mathrm{p})
\end{array}\right.$$

où les coefficients $a'_{11}, \dots, a'_{rr}$ sont tous non nuls et $x'_1, \dots, x'_n$ est une permutation éventuelle des inconnues $x_1, \dots, x_n$. Le système étant de rang r, les $p - r$ équations restantes ne contiennent plus d'inconnue et constituent de fait des égalités soit vraies soit fausses ; ce sont les conditions de compatibilité. Supposons les égalités $(\mathrm{E}'_{\mathrm{r}+1})$–$(\mathrm{E}'_\mathrm{p})$ satisfaites. Le système est donc compatible. Sa résolution se fait de manière progressive en exprimant x'_r en fonction de $x'_{r+1}, \dots, x_n$ grâce à $(\mathrm{E_r})$:

$$x'_r = \frac{b'_r - a'_{r\,r+1}x'_{r+1} - \cdots - a'_{rn}x'_n}{a'_{rr}},$$

4. Gauss, Carl Friedrich : astronome, mathématicien et physicien allemand (Brunswick 1777–Göttingen 1855)

que l'on reporte dans $(\mathrm{E_{r-1}})$, équation de laquelle on tire x'_{r-1} :

$$x'_{r-1} = \frac{b'_{r-1} - a'_{r-1\,r}x'_r - \cdots - a'_{r-1\,n}x'_n}{a'_{r-1\,r-1}},$$

et on remonte ainsi de suite jusqu'à $(\mathrm{E_1})$ pour en tirer x'_1 :

$$x'_1 = \frac{b'_1 - a'_{12}x'_2 - \cdots - a'_{1n}x'_n}{a'_{11}}.$$

On obtient ainsi $x'_1, \ldots, x'_r$ en fonction de $x'_{r+1}, \ldots, x'_n$.

— *Détails de la méthode :* indiquons les étapes de l'algorithme partant du système initial

$$(\mathrm{S}) \qquad \left\{\begin{array}{rcll} a_{11}\,x_1 + a_{12}\,x_2 + \cdots + a_{1n}\,x_n &=& b_1 & (\mathrm{E_1}) \\ a_{21}\,x_1 + a_{22}\,x_2 + \cdots + a_{2n}\,x_n &=& b_2 & (\mathrm{E_2}) \\ &\vdots& & \vdots \\ a_{p1}\,x_1 + a_{p2}\,x_2 + \cdots + a_{pn}\,x_n &=& b_p & (\mathrm{E_p}) \end{array}\right.$$

- Posons $a^1_{ij} = a_{ij}$, $b^1_i = b_i$ et $(\mathrm{E^1_i}) = (\mathrm{E_i})$ pour $1 \leqslant i \leqslant p$, $1 \leqslant j \leqslant n$.
- Supposons $a^1_{11} \neq 0$. On conserve l'équation $(\mathrm{E^1_1})$. On remplace les équations $(\mathrm{E^1_i})$, $2 \leqslant i \leqslant p$, par $(\mathrm{E^2_i}) = (\mathrm{E^1_i}) - \dfrac{a^1_{i1}}{a^1_{11}}\,(\mathrm{E^1_1})$ ce qui fournit les coefficients

$$\begin{array}{rcl} a^2_{ij} &=& a^1_{ij} - \dfrac{a^1_{i1}}{a^1_{11}}\,a^1_{1j},\ 2 \leqslant i \leqslant p,\ 2 \leqslant j \leqslant n, \\ b^2_i &=& b^1_i - \dfrac{a^1_{i1}}{a^1_{11}}\,b^1_1,\ 2 \leqslant i \leqslant p, \end{array}$$

et le système correspondant

$$(\mathrm{S^2}) \qquad \left\{\begin{array}{rcll} a^1_{11}\,x_1 + a^1_{12}\,x_2 + \cdots + a^1_{1n}\,x_n &=& b^1_1 & (\mathrm{E^1_1}) \\ a^2_{22}\,x_2 + \cdots + a^2_{2n}\,x_n &=& b^2_2 & (\mathrm{E^2_2}) \\ \vdots \qquad\qquad &\vdots& & \vdots \\ a^2_{p2}\,x_2 + \cdots + a^2_{pn}\,x_n &=& b^2_p & (\mathrm{E^2_p}) \end{array}\right.$$

Si $a^1_{11} = 0$, on permute les équations $(\mathrm{E^1_1}), \ldots, (\mathrm{E^1_p})$ ou les inconnues $x_1, \ldots, x_n$ de façon que le premier coefficient de la première équation soit non nul (et on peut alors le choisir comme pivot).
- Supposons $a^2_{22} \neq 0$. On conserve les équations $(\mathrm{E^1_1})$ et $(\mathrm{E^2_2})$. On remplace $(\mathrm{E^2_i})$, $3 \leqslant i \leqslant p$, par $(\mathrm{E^3_i}) = (\mathrm{E^2_i}) - \dfrac{a^2_{i2}}{a^2_{22}}\,(\mathrm{E^2_2})$ ce qui fournit les coefficients

$$\begin{array}{rcl} a^3_{ij} &=& a^2_{ij} - \dfrac{a^2_{i2}}{a^2_{22}}\,a^2_{2j},\ 3 \leqslant i \leqslant p,\ 3 \leqslant j \leqslant n, \\ b^3_i &=& b^2_i - \dfrac{a^2_{i2}}{a^2_{22}}\,b^2_2,\ 3 \leqslant i \leqslant p, \end{array}$$

puis le système

$$(\mathrm{S^3}) \qquad \left\{\begin{array}{rcll} a^1_{11}\,x_1 + a^1_{12}\,x_2 + a^1_{13}\,x_3 + \cdots + a^1_{1n}\,x_n &=& b^1_1 & (\mathrm{E^1_1}) \\ a^2_{22}\,x_2 + a^2_{23}\,x_3 + \cdots + a^2_{2n}\,x_n &=& b^2_2 & (\mathrm{E^2_2}) \\ a^3_{33}\,x_3 + \cdots + a^3_{3n}\,x_n &=& b^3_3 & (\mathrm{E^3_3}) \\ \vdots \qquad\qquad &\vdots& & \vdots \\ a^3_{p3}\,x_3 + \cdots + a^3_{pn}\,x_n &=& b^3_p & (\mathrm{E^3_p}) \end{array}\right.$$

Si $a_{22}^2 = 0$, on permute les équations restantes ou les inconnues restantes de façon à remplacer a_{22}^2 par un coefficient non nul (que l'on peut alors choisir comme pivot).

- Supposons que l'on ait ainsi construit au bout de k étapes le système

$$(\mathrm{S^k}) \quad \left\{ \begin{array}{rcll} a_{11}^1\, x_1 + a_{12}^1\, x_2 + a_{13}^1\, x_3 + \cdots + a_{1k}^1\, x_k + \cdots + a_{1n}^1\, x_n & = & b_1^1 & (\mathrm{E_1^1}) \\ a_{22}^2\, x_2 + a_{23}^2\, x_3 + \cdots + a_{2k}^2\, x_k + \cdots + a_{2n}^2\, x_n & = & b_2^2 & (\mathrm{E_2^2}) \\ a_{33}^3\, x_3 + \cdots + a_{3k}^3\, x_k + \cdots + a_{3n}^3\, x_n & = & b_3^3 & (\mathrm{E_3^3}) \\ \ddots \qquad & & \vdots & \vdots \\ a_{kk}^k\, x_k + \cdots + a_{kn}^k\, x_n & = & b_k^k & (\mathrm{E_k^k}) \\ a_{k+1\,k}^k\, x_k + \cdots + a_{k+1\,n}^k\, x_n & = & b_{k+1}^k & (\mathrm{E_{k+1}^k}) \\ \vdots \qquad & & \vdots & \vdots \\ a_{pk}^k\, x_k + \cdots + a_{pn}^k\, x_n & = & b_p^k & (\mathrm{E_p^k}) \end{array} \right.$$

L'étape suivante consiste à éliminer le coefficient de x_k dans les équations $(\mathrm{E_i})$, $k+1 \leqslant i \leqslant p$. Supposons $a_{kk}^k \neq 0$. On remplace $(\mathrm{E_i^k})$, $k+1 \leqslant i \leqslant p$, par $(\mathrm{E_i^{k+1}}) = (\mathrm{E_i^k}) - \dfrac{a_{ik}^k}{a_{kk}^k}\,(\mathrm{E_k^k})$ dont les coefficients valent

$$\begin{aligned} a_{ij}^{k+1} &= a_{ij}^k - \frac{a_{ik}^k}{a_{kk}^k}\, a_{kj}^k,\ k+1 \leqslant i \leqslant p,\ k+1 \leqslant j \leqslant n, \\ b_i^{k+1} &= b_i^k - \frac{a_{ik}^k}{a_{kk}^k}\, b_k^k,\ k+1 \leqslant i \leqslant n. \end{aligned}$$

On reproduit ainsi les étapes précédentes jusqu'à la r^{e} équation pour obtenir un système $(\mathrm{S^r})$ triangulaire équivalent à (S) :

$$(\mathrm{S^r}) \quad \left\{ \begin{array}{rcll} a_{11}^1\, x_1 + a_{12}^1\, x_2 + a_{13}^1\, x_3 + \cdots + a_{1r}^1\, x_r + \cdots + a_{1n}^1\, x_n & = & b_1^1 & (\mathrm{E_1^1}) \\ a_{22}^2\, x_2 + a_{23}^2\, x_3 + \cdots + a_{2r}^2\, x_r + \cdots + a_{2n}^2\, x_n & = & b_2^2 & (\mathrm{E_2^2}) \\ a_{33}^3\, x_3 + \cdots + a_{3r}^3\, x_r + \cdots + a_{3n}^3\, x_n & = & b_3^3 & (\mathrm{E_3^3}) \\ \ddots \qquad & & \vdots & \vdots \\ a_{rr}^r\, x_r + \cdots + a_{rn}^r\, x_n & = & b_r^r & (\mathrm{E_r^r}) \\ 0 & = & b_{r+1}^r & (\mathrm{E_{r+1}^r}) \\ & & \vdots & \vdots \\ 0 & = & b_p^r & (\mathrm{E_p^r}) \end{array} \right.$$

où les coefficients $a_{11}^r, \ldots, a_{rr}^r$ sont tous non nuls.

— *Coût de la méthode :* le nombre d'opérations effectuées au cours de la résolution d'un système de Cramer de n équations à n inconnues par cette méthode est de l'ordre de $2n^3/3$ lorsque n est grand. Par exemple pour $n = 100$, un ordinateur fonctionnant à 100 mégaflops prendra 6,7 ms pour réaliser ce calcul.
En effet, évaluons le nombre exact d'opérations effectuées. Lors du passage de la k^{e} à la $(k+1)^{\mathrm{e}}$ étape, le calcul des a_{ij}^{k+1} en fonction des a_{ij}^k nécessite une division (de a_{ik}^k par a_{kk}^k) commune à toute la i^{e} équation, une soustraction et une multiplication pour chaque i, j. De même pour le calcul des b_i^{k+1} en fonction des b_i^k. En faisant varier j de $k+1$ à n, cela donne pour la i^{e} équation $2(n-k)+1$ opérations, puis en faisant varier i de $k+1$ à n, la k^{e} itération demande $(n-k)(2(n-k)+1)$ opérations. Enfin, en faisant varier k de 1 à $n-1$, le nombre d'opérations utilisées par l'algorithme de transformation du système est

$$\sum_{k=1}^{n-1} [2(n-k)^2 + (n-k)] = 2\sum_{k=1}^{n-1} k^2 + \sum_{k=1}^{n-1} k = \frac{(n-1)n(2n-1)}{3} + \frac{(n-1)n}{2} = \frac{2n^3}{3} - \frac{n^2}{2} - \frac{n}{6}.$$

Par ailleurs, la résolution du système triangulaire obtenu utilise pour le calcul de x_k en fonction de $x_{k-1}, \dots, x_1$ une division, $n-k$ multiplications et $n-k$ soustractions, soit au total

$$\sum_{k=1}^{n}[2(n-k)+1] = (n-1)n + (n-1) = n^2 - 1,$$

d'où le coût final de $\dfrac{2n^3}{3} + \dfrac{n^2}{2} - \dfrac{n}{6} - 1 \sim \dfrac{2n^3}{3}$ opérations.

Exemple.

Reprenons l'exemple du paragraphe précédent :

$$(\mathrm{S}^1) \qquad \left\{\begin{array}{rcl} 3\,x+2\,y-5\,z &=& 1 \\ x-\ \ y+5\,z &=& 0 \\ x+\ \ y-3\,z &=& a \end{array}\right. \qquad \begin{array}{r} (\mathrm{E}_1^1) \\ (\mathrm{E}_2^1) \\ (\mathrm{E}_3^1) \end{array}$$

Les combinaisons linéaires des équations (E_1^1)–(E_2^1) suivantes :

$$(\mathrm{E}_2^2) = (\mathrm{E}_2^1) - \frac{1}{3}\,(\mathrm{E}_1^1) \quad \text{et} \quad (\mathrm{E}_3^2) = (\mathrm{E}_3^1) - \frac{1}{3}\,(\mathrm{E}_1^1)$$

conduisent au système (S^2) équivalent à (S^1)

$$(\mathrm{S}^2) \qquad \left\{\begin{array}{rcl} 3\,x+2\,y-\ \ 5\,z &=& 1 \\ 5\,y-20\,z &=& 1 \\ y-\ \ 4\,z &=& 3a-1 \end{array}\right. \qquad \begin{array}{r} (\mathrm{E}_1^1) \\ (\mathrm{E}_2^2) \\ (\mathrm{E}_3^2) \end{array}$$

puis la combinaison linéaire des équations (E_2^2) et (E_3^2) suivante :

$$(\mathrm{E}_3^3) = (\mathrm{E}_3^2) - \frac{1}{5}\,(\mathrm{E}_2^2)$$

conduit au système triangulaire (S^3) équivalent à (S^1)

$$(\mathrm{S}^3) \qquad \left\{\begin{array}{rcl} 3\,x+2\,y-\ \ 5\,z &=& 1 \\ 5\,y-20\,z &=& 1 \\ 0 &=& 5a-2 \end{array}\right. \qquad \begin{array}{r} (\mathrm{E}_1^1) \\ (\mathrm{E}_2^2) \\ (\mathrm{E}_3^3) \end{array}$$

La condition de compatibilité est $5a-2=0$. Supposons donc que $a = 2/5$. On tire de (E_2^2) $y = 4z + \frac{1}{5}$ puis, en reportant dans (E_1^1), $x = -z + \frac{1}{5}$.

5 Exercices sur les déterminants

Exercice 1 Soit a, b, c, d des nombres réels. Calculer les déterminants suivants (on exprimera leur valeur sous une forme factorisée) :

$$\begin{vmatrix} 1 & 1 & 1 \\ b+c & c+a & a+b \\ bc & ca & ab \end{vmatrix} ; \quad \begin{vmatrix} a-b-c & 2a & 2a \\ 2b & b-c-a & 2b \\ 2c & 2c & c-a-b \end{vmatrix} ; \quad \begin{vmatrix} a & b & c & d \\ c & d & a & b \\ b & a & d & c \\ d & c & b & a \end{vmatrix}.$$

Soit $A = \begin{pmatrix} a & -b & -c & -d \\ b & a & d & -c \\ c & -d & a & b \\ d & c & -b & a \end{pmatrix}$. Calculer ${}^t\!AA$ puis en déduire $\det(A)$.

Exercice 2 Soit a_1, a_2, …, a_n, a et b des nombres réels. Calculer les déterminants

$$\begin{vmatrix} a_1 & a_1 & a_1 & \dots & a_1 \\ a_1 & a_2 & a_2 & \dots & a_2 \\ a_1 & a_2 & a_3 & \dots & a_3 \\ \vdots & \vdots & \vdots & & \vdots \\ a_1 & a_2 & a_3 & \dots & a_n \end{vmatrix} \quad \text{et} \quad \begin{vmatrix} a & b & \dots & b & b \\ b & a & \ddots & & b \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ b & & \ddots & a & b \\ b & b & \dots & b & a \end{vmatrix}.$$

Exercice 3 (Déterminant de Vandermonde[5]) Soit a_1, a_2, …, a_n des nombres réels.

1. Calculer le déterminant

$$D_n(a_1, a_2, \dots, a_n) = \begin{vmatrix} 1 & a_1 & a_1^2 & \dots & a_1^{n-1} \\ 1 & a_2 & a_2^2 & \dots & a_2^{n-1} \\ 1 & a_3 & a_3^2 & \dots & a_3^{n-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & a_n & a_n^2 & \dots & a_n^{n-1} \end{vmatrix}.$$

   Indication : si $\vec{u}_i$ représente la i^{e} colonne, alors

$$\det(\vec{u}_1, \vec{u}_2, \dots, \vec{u}_n) = \det(\vec{u}_1, \vec{u}_2 - a_n\vec{u}_1, \dots, \vec{u}_i - a_n\vec{u}_{i-1}, \dots, \vec{u}_n - a_n\vec{u}_{n-1}),$$

   ce qui fournit une relation de récurrence pour la suite $(D_n(a_1, a_2, \dots, a_n))_{n\geqslant 1}$.
   Résultat : $D_n(a_1, a_2, \dots, a_n) = \displaystyle\prod_{1\leqslant i<j\leqslant n} (a_j - a_i)$.

2. Application 1.
   (a). En développant $D_n(a_1, a_2, \dots, a_n)$ considéré comme un polynôme de la variable a_i par rapport à la j^{e} ligne, calculer le $(i,j)^{\text{e}}$ cofacteur de ce déterminant. On utilisera les relations entre coefficients et racines d'un polynôme.

5. Vandermonde, Alexandre-Théophile : mathématicien français (Paris 1735–Paris 1796)

(b). On suppose les réels a_1, a_2, ..., a_n tous distincts. En déduire la solution du système de Cramer

$$\left\{\begin{array}{rcl} x_1 + \quad x_2 + \cdots + \quad x_n & = & \alpha \\ a_1 x_1 + \quad a_2 x_2 + \cdots + \quad a_n x_n & = & 0 \\ a_1^2 x_1 + \quad a_2^2 x_2 + \cdots + \quad a_n^2 x_n & = & 0 \\ & \vdots & \\ a_1^{n-2} x_1 + a_2^{n-2} x_2 + \cdots + a_n^{n-2} x_n & = & 0 \\ a_1^{n-1} x_1 + a_2^{n-1} x_2 + \cdots + a_n^{n-1} x_n & = & \beta \end{array}\right.$$

3. Application 2.
   On désigne par E le $\mathbb{R}$-espace vectoriel des fonctions de $\mathbb{R}$ dans $\mathbb{R}$. On dit qu'un système infini $\mathcal{S}$ de vecteurs de E est libre si tout système fini extrait de $\mathcal{S}$ est libre.
   Pour tout $a \in \mathbb{R}$, on définit le vecteur e_a de E par $\forall x \in \mathbb{R}$, $e_a(x) = e^{ax}$. Montrer que le système infini de vecteurs $\{e_a, a \in \mathbb{R}\}$ est libre.

Exercice 4 Soit $n \in \mathbb{N} \setminus \{0,1\}$ et $\{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n\}$ une base du $\mathbb{R}$-espace vectoriel $\mathbb{R}^n$.

1. Étudier, en fonction du paramètre réel a et de la parité de l'entier n, le rang du système de vecteurs $\{\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_n\}$ défini par

$$\vec{u}_1 = \vec{e}_1 + a\vec{e}_n \quad \text{et} \quad \forall i \in \{2,3,\ldots,n\},\ \vec{u}_i = a\vec{e}_{i-1} + \vec{e}_i.$$

2. Déterminer, suivant les valeurs des réels b_1, b_2, ..., b_n et la parité de n, le nombre de solutions du système linéaire

$$\left\{\begin{array}{lcl} x_i + x_{i+1} & = & b_i, \quad 1 \leqslant i \leqslant n-1, \\ x_1 + x_n & = & b_n. \end{array}\right.$$

3. Application.
   Dans le plan réel rapporté à un repère $(O;\vec{i},\vec{j})$, déterminer

(a). les triangles ayant pour milieux des côtés les trois points $A(1,0)$, $B(0,0)$ et $C(0,1)$;

(b). les quadrilatères ayant pour milieux des côtés les quatre points A, B, C et $D(1,1)$. Plus généralement, de quel type doit être le quadrilatère $ABCD$ pour qu'il existe des quadrilatères ayant $ABCD$ pour milieux des côtés ?

Chapitre 5

Diagonalisation

1 Position du problème

L'objectif de ce chapitre et de proposer une méthode de calcul des puissances d'une matrice carrée. Notons tout d'abord un cas très simple pour lequel ce calcul est facile : il s'agit du cas des matrices diagonales. Si $D = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$, alors pour tout entier positif m, on a $D^m = \mathrm{diag}(\lambda_1^m, \ldots, \lambda_n^m)$. Considérons maintenant une matrice carrée $A \in \mathcal{M}_n(\mathbb{K})$ et introduisons le contexte vectoriel associé : soit E un $\mathbb{K}$-espace vectoriel de dimension n, $\mathcal{B}$ une base de référence de E et $f \in \mathcal{L}(E)$ un endomorphisme dont la matrice relativement à $\mathcal{B}$ est $\mathrm{Mat}(f; \mathcal{B}) = A$. Supposons qu'il existe une base $\mathcal{B}' = \{\vec{e}_1{}', \ldots, \vec{e}_n{}'\}$ de E relativement à laquelle f admette pour matrice

$$\mathrm{Mat}(f; \mathcal{B}') = D = \begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & & \vdots \\ \vdots & & \ddots & \vdots \\ 0 & \ldots & \ldots & \lambda_n \end{pmatrix}.$$

Le lien entre les deux matrices A et D est bien connu : $A = PDP^{-1}$ où P est la matrice de passage de la base $\mathcal{B}$ vers la base $\mathcal{B}'$ (les matrices A et D sont semblables). Si une telle base $\mathcal{B}'$ existe, on a nécessairement

$$\forall i \in \{1, \ldots, n\},\ f(\vec{e}_i{}') = \lambda_i \vec{e}_i{}'.$$

Ceci s'écrit $(f - \lambda_i \mathrm{id}_E)(\vec{e}_i{}') = \vec{0}$ ou encore $\vec{e}_i{}' \in \mathrm{Ker}(f - \lambda_i \mathrm{id}_E)$; les noyaux $\mathrm{Ker}(f - \lambda_i \mathrm{id}_E)$ ne sont donc pas réduits au vecteur nul et l'endomorphisme $f - \lambda_i \mathrm{id}_E$ n'est pas injectif. On dit que les λ_i sont des *valeurs propres* de f et les $\vec{e}_i{}'$ des *vecteurs propres associés aux* λ_i. En conséquence, pour qu'il existe une base de E relativement à laquelle la matrice de f soit diagonale, il faut pouvoir construire une base constituée de vecteurs propres de f. Une telle base ainsi construite, il devient facile de calculer la puissance m^{e} de l'endomorphisme f définie par $f^m = \underbrace{f \circ \cdots \circ f}_{m \text{ fois}}$:

$$\mathrm{Mat}(f^m; \mathcal{B}') = D^m = \begin{pmatrix} \lambda_1^m & 0 & \ldots & 0 \\ 0 & \lambda_2^m & & \vdots \\ \vdots & & \ddots & \vdots \\ 0 & \ldots & \ldots & \lambda_n^m \end{pmatrix},$$

puis celle de A :

$$A^m = \mathrm{Mat}(f^m; \mathcal{B}) = PD^mP^{-1}.$$

Ce chapitre fournit des conditions nécessaires et suffisantes pour que cette opération puisse être réalisable.

2 Valeurs propres, vecteurs propres

Soit E un $\mathbb{K}$-espace vectoriel et $f \in \mathcal{L}(E)$ un endomorphisme de E.

Définition 2.1 *1. — Un scalaire $\lambda \in \mathbb{K}$ est une* valeur propre *de f lorsque l'endomorphisme $f - \lambda\mathrm{id}_E$ n'est pas injectif, soit encore*

$$\exists \vec{u} \in E \setminus \{\vec{0}\},\ f(\vec{u}) = \lambda\vec{u}.$$

— *Si E est de dimension finie n,*

$$\lambda \text{ est une valeur propre de } f \iff \det(f - \lambda\mathrm{id}_E) = 0.$$

La fonction $\Delta_f : \begin{array}{ccc} \mathbb{K} & \longrightarrow & \mathbb{K} \\ \lambda & \longmapsto & \det(f - \lambda\mathrm{id}_E) \end{array}$ *est une fonction polynomiale sur $\mathbb{K}$ de degré n dont les racines sont les valeurs propres de f. Elle est appelée* polynôme caractéristique *de f.*

— *L'ensemble des valeurs propres est appelé* spectre *de f.*

2. Soit λ une valeur propre de f.

— *Un vecteur non nul $\vec{u} \in E\setminus\{\vec{0}\}$ est un* vecteur propre associé *à λ lorsque $f(\vec{u}) = \lambda\vec{u}$.*

— *L'ensemble des vecteurs propres associés à λ est appelé* sous-espace propre associé *à λ, c'est le sous-espace vectoriel $E_\lambda = \mathrm{Ker}(f - \lambda\mathrm{id}_E)$.*

En dimension finie, si A est la matrice de f relativement à une base de E, on parle indifféremment des valeurs propres et des polynômes caractéristiques de f ou de A. La définition matricielle est cohérente puisque si B est la matrice de f relativement à une autre base de E, les matrices A et B sont semblables : il existe une matrice de passage P telle que $A = PBP^{-1}$, et dans ce cas,

$$\begin{aligned}\Delta_A(\lambda) &= \det(A - \lambda I_n) = \det(PBP^{-1} - \lambda PP^{-1}) = \det(P(B - \lambda I_n)P^{-1}) \\ &= \det(P) \times \det(B - \lambda I_n) \times \det(P)^{-1} = \det(B - \lambda I_n) = \Delta_B(\lambda).\end{aligned}$$

Ainsi, A et B ont même polynôme caractéristique et les définitions matricielles ne dépendent pas du choix de la base de référence.

Énonçons quelques propriétés sur les valeurs propres et sous-espaces propres.

Théorème 2.2 *Soit $\lambda_1, \ldots, \lambda_k$ des valeurs propres distinctes de f dans $\mathbb{K}$. La somme des sous-espaces propres associés est directe :*

$$E_{\lambda_1} + \cdots + E_{\lambda_k} = E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_k}.$$

DÉMONSTRATION.

Soit $\vec{u} \in E_{\lambda_1} + \cdots + E_{\lambda_k}$. Le vecteur $\vec{u}$ admet une décomposition $\vec{u} = \vec{u}_1 + \cdots + \vec{u}_k$ avec $(\vec{u}_1, \ldots, \vec{u}_k) \in E_{\lambda_1} \times \cdots \times E_{\lambda_k}$. Montrons que cette décomposition est unique. Appliquons f plusieurs fois de suite à cette décomposition. Puisque $f(\vec{u}_1) = \lambda_1 \vec{u}_1, \ldots, f(\vec{u}_k) = \lambda_k \vec{u}_k$, cela conduit au système d'équations vectorielles, en rappelant la notation $f^i = \underbrace{f \circ \cdots \circ f}_{i \text{ fois}}$,

$$\left\{\begin{array}{rcl} \vec{u} &=& \vec{u}_1 + \cdots + \vec{u}_k \\ f(\vec{u}) &=& \lambda_1 \vec{u}_1 + \cdots + \lambda_k \vec{u}_k \\ f^2(\vec{u}) &=& \lambda_1^2 \vec{u}_1 + \cdots + \lambda_k^2 \vec{u}_k \\ &\vdots& \\ f^{k-1}(\vec{u}) &=& \lambda_1^{k-1} \vec{u}_1 + \cdots + \lambda_k^{k-1} \vec{u}_k \end{array}\right. .$$

La matrice de ce système est $\begin{pmatrix} 1 & \ldots & 1 \\ \lambda_1 & \ldots & \lambda_k \\ \lambda_1^2 & \ldots & \lambda_k^2 \\ \vdots & & \vdots \\ \lambda_1^{k-1} & \ldots & \lambda_k^{k-1} \end{pmatrix}$, c'est une matrice de Vandermonde dont le déterminant est donné par

$$\begin{vmatrix} 1 & \ldots & 1 \\ \lambda_1 & \ldots & \lambda_k \\ \lambda_1^2 & \ldots & \lambda_k^2 \\ \vdots & & \vdots \\ \lambda_1^{k-1} & \ldots & \lambda_k^{k-1} \end{vmatrix} = \prod_{1 \leqslant i < j \leqslant k} (\lambda_j - \lambda_i).$$

Il est non nul puisque les valeurs propres $\lambda_1, \ldots, \lambda_k$ sont supposées distinctes. Le système précédent est donc un système de Cramer et admet une solution unique pour le k-uplet $(\vec{u}_1, \ldots, \vec{u}_k)$. Les vecteurs $\vec{u}_1, \ldots, \vec{u}_k$ sont d'ailleurs des combinaisons linéaires des vecteurs $\vec{u}, f(\vec{u}), \ldots, f^{k-1}(\vec{u})$. Ainsi, la somme $E_{\lambda_1} + \cdots + E_{\lambda_k}$ est directe. □

Les valeurs propres sont les racines du polynôme caractéristique de f. Ces racines peuvent être multiples. On appelle *multiplicité* d'une valeur propre sa multiplicité dans le polynôme caractéristique.

Théorème 2.3 *Soit λ_i une valeur propre de f de multiplicité n_i. On a $1 \leqslant \dim(E_{\lambda_i}) \leqslant n_i$.*

DÉMONSTRATION.

Notons m_i la dimension de E_{λ_i}. Soit F un sous-espace vectoriel supplémentaire de E_{λ_i} dans E : $E = E_{\lambda_i} \oplus F$. Introduisons des bases $\mathcal{B}_{\lambda_i}$ de E_{λ_i} et $\mathcal{B}_F$ de F ; la réunion $\mathcal{B} = \mathcal{B}_{\lambda_i} \cup \mathcal{B}_F$ est une base de E relativement à laquelle la matrice de l'endomorphisme f a la forme

$$\begin{array}{cc} \begin{array}{cc} f(\mathcal{B}_{\lambda_i}) & f(\mathcal{B}_F) \end{array} & \\ \left(\begin{array}{c|c} \lambda I_{m_i} & A \\ \hline O_{n-m_i,m_i} & B \end{array}\right) & \begin{array}{c} \mathcal{B}_{\lambda_i} \\ \mathcal{B}_F \end{array} \end{array}$$

avec $A \in \mathcal{M}_{m_i \times (n-m_i)}(\mathbb{K})$ et $B \in \mathcal{M}_{m_i}(\mathbb{K})$. On déduit, en développant le déterminant de cette matrice successivement par rapport à ses premières colonnes, que le polynôme caractéristique de

f peut se factoriser selon

$$\Delta_f(\lambda) = \left|\begin{array}{cccc|c} \lambda_i-\lambda & 0 & \dots & 0 & \\ 0 & \lambda_i-\lambda & & 0 & A \\ \vdots & & \ddots & \vdots & \\ 0 & 0 & \dots & \lambda_i-\lambda & \\ \hline 0 & 0 & \dots & 0 & \\ \vdots & \vdots & & \vdots & B-\lambda I_{n-m_i} \\ 0 & 0 & \dots & 0 & \end{array}\right| = (\lambda_i-\lambda)^{m_i}\Delta_B(\lambda)$$

où Δ_B est le polynôme caractéristique (de degré $n-m_i$) de la matrice B. Cette factorisation prouve que la multiplicité de la racine λ_i est au moins m_i. Par hypothèse, celle-là est égale à n_i et donc $m_i \leqslant n_i$. □

3 Diagonalisation

On suppose dorénavant que E est un $\mathbb{K}$-espace vectoriel de dimension finie n.

Définition 3.1 1. *L'endomorphisme f de E est* diagonalisable *lorsqu'il existe une base de E relativement à laquelle la matrice de f est diagonale, ou encore de manière équivalente, lorsque que E admet une base formée de vecteurs propres de f.*
2. *La matrice carrée A est* diagonalisable *lorsqu'elle est semblable à une matrice diagonale.*

Théorème 3.2 *L'endomorphisme f de E est diagonalisable si et seulement si ses sous-espaces propres sont supplémentaires dans E, i.e., si $\lambda_1, \dots, \lambda_k$ sont les valeurs propres distinctes de f dans $\mathbb{K}$,*

$$E = E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_k}.$$

Démonstration.

1. Supposons f diagonalisable. Soit $\mathcal{B}'$ une base de E constituée de vecteurs propres de f. En rassemblant dans $\mathcal{B}'$ les vecteurs propres de f associés à la valeur propre λ_1, ceux associés à la valeur propre λ_2, etc., on obtient une partition de $\mathcal{B}'$ en $\mathcal{L}_{\lambda_1} \cup \cdots \cup \mathcal{L}_{\lambda_k}$, les $\mathcal{L}_{\lambda_i}$ étant formés de vecteurs propres associés à λ_i. Pour chaque $i \in \{1, \dots, k\}$, $\mathcal{L}_{\lambda_i}$ est un système libre de vecteurs de E_{λ_i} (puisque contenu dans une base), il engendre un sous-espace vectoriel F_{λ_i} de E_{λ_i}. La relation $\mathcal{B}' = \mathcal{L}_{\lambda_1} \cup \cdots \cup \mathcal{L}_{\lambda_k}$ entraîne

$$E = F_{\lambda_1} \oplus \cdots \oplus F_{\lambda_k} \subset E_{\lambda_1} + \cdots + E_{\lambda_k}.$$

D'autre part, on sait que la somme $E_{\lambda_1} + \cdots + E_{\lambda_k}$ est directe (théorème 2.2), d'où

$$E = E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_k}.$$

On en déduit d'ailleurs que $F_{\lambda_i} = E_{\lambda_i}$ et que $\mathcal{L}_{\lambda_i}$ est en fait une base de E_{λ_i} pour chaque i.

2. Supposons réciproquement que $E = E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_k}$. Notons m_i la dimension de E_{λ_i} pour tout $i \in \{1, \ldots, k\}$. Soit $\mathcal{B}_{\lambda_1}, \ldots, \mathcal{B}_{\lambda_k}$ des bases respectives de $E_{\lambda_1}, \ldots, E_{\lambda_k}$, puis $\mathcal{B}' = \mathcal{B}_{\lambda_1} \cup \cdots \cup \mathcal{B}_{\lambda_k}$. Le système de vecteurs $\mathcal{B}'$ est une base de E et la matrice de f relativement à $\mathcal{B}'$ est

$$\begin{array}{cc} \begin{array}{cccc} f(\mathcal{B}_{\lambda_1}) & f(\mathcal{B}_{\lambda_2}) & \ldots & f(\mathcal{B}_{\lambda_k}) \end{array} & \\ \begin{pmatrix} \boxed{\lambda_1 I_{m_1}} & O & & O \\ O & \boxed{\lambda_2 I_{m_2}} & & O \\ \vdots & \vdots & \ddots & \vdots \\ O & O & & \boxed{\lambda_k I_{m_k}} \end{pmatrix} & \begin{array}{c} \mathcal{B}_{\lambda_1} \\ \mathcal{B}_{\lambda_2} \\ \vdots \\ \mathcal{B}_{\lambda_k} \end{array} \end{array}$$

où chaque bloc diagonal est une matrice carrée d'ordre m_i diagonale :

$$\lambda_i I_{m_i} = \begin{pmatrix} \lambda_i & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & \lambda_i \end{pmatrix} \in \mathcal{M}_{m_i}(\mathbb{K}).$$

La matrice $\mathrm{Mat}(f; \mathcal{B}')$ est diagonale et l'endomorphisme f est alors diagonalisable. □

De ce théorème, on déduit une condition nécessaire et suffisante importante en pratique pour qu'un endomorphisme soit diagonalisable.

Corollaire 3.3 *L'endomorphisme f de E est diagonalisable si et seulement si son polynôme caractéristique admet toutes ses racines dans $\mathbb{K}$ et, si $\lambda_1, \ldots, \lambda_k$ sont les valeurs propres distinctes de f de multiplicités respectives $n_1, \ldots, n_k$, $\forall i \in \{1, \ldots, k\}$, $\dim(E_{\lambda_i}) = n_i$.*

Remarque.

La condition $\dim(E_{\lambda_i}) = n_i$ peut encore s'écrire $\mathrm{rang}(f - \lambda_i \mathrm{id}_E) = n - n_i$. Cette dernière peut se vérifier à l'aide de la méthode des zéros échelonnés et évite la recherche de la détermination explicite du sous-espace propre E_{λ_i}.

Démonstration.

Notons tout d'abord que $n_1 + \cdots + n_k = n$ puisque le polynôme caractéristique de f se factorise selon $\Delta_f(\lambda) = \prod_{i=1}^{k} (\lambda_i - \lambda)^{n_i}$. On a

$$\begin{aligned} E = E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_k} &\iff \dim(E_{\lambda_1}) + \cdots + \dim(E_{\lambda_k}) = n = n_1 + \cdots + n_k \\ &\iff [n_1 - \dim(E_{\lambda_1})] + \cdots + [n_k - \dim(E_{\lambda_k})] = 0. \end{aligned}$$

D'après le théorème 2.3, les nombres $n_i - \dim(E_{\lambda_i})$, $1 \leqslant i \leqslant k$, sont positifs ou nuls. Leur somme est donc nulle si et seulement si ils sont tous nuls, soit :

$$E = E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_k} \iff \forall i \in \{1, \ldots, k\},\ \dim(E_{\lambda_i}) = n_i,$$

ce qui permet de conclure grâce au théorème 3.2. □

Exemple.

Soit $A = \begin{pmatrix} 2 & 0 & 0 \\ -3 & -1 & 3 \\ 3 & 3 & -1 \end{pmatrix}$ et f l'endomorphisme de $E = \mathbb{K}^3$ dont A est la matrice relativement à la base canonique $\mathcal{B}$ de E : $A = \mathrm{Mat}(f;\mathcal{B})$.

— *Recherche des valeurs propres :* le polynôme caractéristique de A est donné par

$$\Delta_A(\lambda) = \begin{vmatrix} 2-\lambda & 0 & 0 \\ -3 & -1-\lambda & 3 \\ 3 & 3 & -1-\lambda \end{vmatrix} = (2-\lambda) \times \begin{vmatrix} -1-\lambda & 3 \\ 3 & -1-\lambda \end{vmatrix} = (2-\lambda)^2(-4-\lambda).$$

La matrice A admet donc la valeur propre simple -4 et la valeur propre double 2.

— *Recherche des sous-espaces propres :*

- pour la valeur propre -4, on a

$$A + 4I_3 = \mathrm{Mat}(f + 4\mathrm{id}_E;\mathcal{B}) = \begin{pmatrix} 6 & 0 & 0 \\ -3 & 3 & 3 \\ 3 & 3 & 3 \end{pmatrix}.$$

Le sous-espace propre associé est le noyau $\mathrm{Ker}(f + 4\mathrm{id}_E)$ caractérisé par le système d'équations

$$\begin{cases} 6\,x = 0 \\ -3\,x + 3\,y + 3\,z = 0 \\ 3\,x + 3\,y + 3\,z = 0 \end{cases}$$

ou encore $x = 0$ et $y + z = 0$. Ainsi

$$E_{-4} = \mathrm{vect}\{(0,1,-1)\}\,;$$

- pour la valeur propre 2, on a

$$A - 2I_3 = \mathrm{Mat}(f - 2\mathrm{id}_E;\mathcal{B}) = \begin{pmatrix} 0 & 0 & 0 \\ -3 & -3 & 3 \\ 3 & 3 & -3 \end{pmatrix}.$$

Cette matrice est clairement de rang 1 qui coïncide avec $3 - 2$, ce qui montre que la matrice A est diagonalisable en vertu de la remarque suivant le corollaire 3.3. Le sous-espace propre associé est $\mathrm{Ker}(f - 2\mathrm{id}_E)$ caractérisé par le système d'équations $\begin{cases} -3\,x - 3\,y + 3\,z = 0 \\ 3\,x + 3\,y - 3\,z = 0 \end{cases}$ ou encore $x + y - z = 0$. Ainsi, E_2 est le plan vectoriel

$$E_2 = \mathrm{vect}\{(1,0,1),(0,1,1)\},$$

et l'on vérifie de nouveau que la matrice A est diagonalisable puisque $\dim(E_2)$ coïncide avec la multiplicité de la valeur propre 2.

— *Diagonalisation :* on obtient une base de E formée de vecteurs propres de A en réunissant une base de E_{-4} et une base de E_2 : $\mathcal{B}' = \{(0,1,-1),(1,0,1),(0,1,1)\}$. On pose alors

$$P = P_{\mathcal{B}\to\mathcal{B}'} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ -1 & 1 & 1 \end{pmatrix} \quad \text{et} \quad D = \mathrm{Mat}(f;\mathcal{B}') = \begin{pmatrix} -4 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

et l'on a $A = PDP^{-1}$. On veillera à bien respecter le même ordre dans lequel sont écrits les valeurs propres dans D et les vecteurs propres dans P.

Cas particuliers.

1. Si les valeurs propres de f sont simples, leur multiplicité vaut 1 et coïncide avec la dimension de leur sous-espace propre associé (qui est une droite vectorielle dans ce cas). L'endomorphisme f est diagonalisable. Si $\lambda_1, \ldots, \lambda_n$ sont ses valeurs propres distinctes, la matrice de f dans une base de vecteurs propres est la matrice diagonale

$$D = \begin{pmatrix} \lambda_1 & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & \lambda_n \end{pmatrix}$$

2. Si $\dim(E) = 2$, l'endomorphisme f a une matrice relativement à une base de E de la forme $A = \begin{pmatrix} a & c \\ b & d \end{pmatrix}$. Son polynôme caractéristique est donné par $\Delta_A(\lambda) = \lambda^2 - (a+d)\lambda$ $+(ad - bc)$, soit encore, indépendamment des coefficients de la matrice,

$$\boxed{\Delta_A(\lambda) = \lambda^2 - \mathrm{tr}(A)\lambda + \det(A).}$$

— Si A admet deux valeurs propres distinctes λ, μ dans $\mathbb{K}$, elle est diagonalisable et semblable à la matrice diagonale $\begin{pmatrix} \lambda & 0 \\ 0 & \mu \end{pmatrix}$.

— Si A est une matrice réelle ($a, b, c, d \in \mathbb{R}$) admettant deux valeurs propres complexes non réelles, celles-ci sont nécessairement conjuguées. Notons-les $\lambda + i\mu$ et $\lambda - i\mu$ avec $\lambda \in \mathbb{R}, \mu \in \mathbb{R}^*$ et posons $D = \begin{pmatrix} \lambda + i\mu & 0 \\ 0 & \lambda - i\mu \end{pmatrix}$. On peut également trouver une base de vecteurs propres à coordonnées complexes conjuguées : $(\alpha + i\beta, \gamma + i\delta)$ et $(\alpha - i\beta, \gamma - i\delta)$ où $\alpha, \beta, \gamma, \delta \in \mathbb{R}$ vérifient $\alpha\delta - \beta\gamma \neq 0$ conduisant à une matrice de passage de la forme $P = \begin{pmatrix} \alpha + i\beta & \alpha - i\beta \\ \gamma + i\delta & \gamma - i\delta \end{pmatrix}$. On obtient dès lors la diagonalisation de A sur $\mathbb{C}$ suivante : $A = PDP^{-1}$. En décomposant P et P^{-1} selon

$$P = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \begin{pmatrix} 1 & 1 \\ i & -i \end{pmatrix} \quad \text{et} \quad P^{-1} = \frac{1}{2(\alpha\delta - \beta\gamma)} \begin{pmatrix} 1 & -i \\ 1 & i \end{pmatrix} \begin{pmatrix} \delta & -\beta \\ -\gamma & \alpha \end{pmatrix},$$

on trouve la factorisation

$$\begin{aligned} A &= \left[\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \begin{pmatrix} 1 & 1 \\ i & -i \end{pmatrix}\right] \times \begin{pmatrix} \lambda + i\mu & 0 \\ 0 & \lambda - i\mu \end{pmatrix} \times \left[\frac{1}{2(\alpha\delta - \beta\gamma)} \begin{pmatrix} 1 & -i \\ 1 & i \end{pmatrix} \begin{pmatrix} \delta & -\beta \\ -\gamma & \alpha \end{pmatrix}\right] \\ &= \frac{1}{2(\alpha\delta - \beta\gamma)} \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \times \left[\begin{pmatrix} 1 & 1 \\ i & -i \end{pmatrix} \begin{pmatrix} \lambda + i\mu & 0 \\ 0 & \lambda - i\mu \end{pmatrix} \begin{pmatrix} 1 & -i \\ 1 & i \end{pmatrix}\right] \times \begin{pmatrix} \delta & -\beta \\ -\gamma & \alpha \end{pmatrix} \\ &= \frac{1}{\alpha\delta - \beta\gamma} \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \begin{pmatrix} \lambda & \mu \\ -\mu & \lambda \end{pmatrix} \begin{pmatrix} \delta & -\beta \\ -\gamma & \alpha \end{pmatrix}. \end{aligned}$$

On débouche ainsi sur la décomposition réelle de A suivante :

$$A = \tilde{P}\tilde{D}\tilde{P}^{-1} \quad \text{avec} \quad \tilde{P} = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \quad \text{et} \quad \tilde{D} = \begin{pmatrix} \lambda & \mu \\ -\mu & \lambda \end{pmatrix}.$$

Les matrices $\tilde{P}$ et $\tilde{D}$ sont réelles, $\tilde{P}$ représente la matrice de passage de la base initiale vers la base de vecteurs à coordonnées réelles (α, γ) et (β, δ), $\tilde{D}$ représente dans une base orthonormée la matrice d'une *similitude plane directe* *i.e.* la composée d'une homothétie vectorielle par une rotation vectorielle.

— Si A admet une valeur propre double λ dans $\mathbb{K}$, elle est diagonalisable uniquement lorsque $A = \lambda I_2$, c'est-à-dire lorsque f est une homothétie vectorielle. En effet, si A est diagonalisable, elle est semblable à la matrice diagonale $\begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix} = \lambda I_2$ et il existe une matrice de passage P telle que $A = P(\lambda I_2)P^{-1} = \lambda I_2$. Si f n'est pas une homothétie vectorielle, A est semblable à la matrice triangulaire $T = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}$. En effet, on peut choisir une base $\{\vec{u}, \vec{v}\}$ relativement à laquelle f admette T pour matrice ; elle est telle que $f(\vec{u}) = \lambda\vec{u}$ et $f(\vec{v}) = \vec{u} + \lambda\vec{v}$. Ainsi, $\vec{u} \in E_\lambda$ et $\vec{v} \in (f - \lambda \mathrm{id}_E)^{-1}(\vec{u})$. On dit que $\vec{v}$ est un vecteur *spectral*.

3. De manière plus générale, lorsque $\dim(E) = n$, on peut écrire

$$\Delta_A(\lambda) = (-\lambda)^n + \mathrm{tr}(A)(-\lambda)^{n-1} + \cdots + \det(A).$$

Notons au passage que, d'après les relations entre coefficients et racines d'un polynôme, si A admet toutes ses valeurs propres $\lambda_1, \dots, \lambda_n$ (éventuellement confondues) dans $\mathbb{K}$,

$$\boxed{\mathrm{tr}(A) = \sum_{i=1}^{n} \lambda_i \quad \text{et} \quad \det(A) = \prod_{i=1}^{n} \lambda_i.}$$

Signalons également que A est non inversible si et seulement si elle admet la valeur propre nulle, et que le rang de A est égal au nombre de valeurs propres non nulles (répétées avec leur ordre de multiplicité) à condition qu'elles soient toutes dans $\mathbb{K}$.

4. Pour les matrices symétriques réelles, on dispose d'un résultat particulièrement intéressant. On se place dans un $\mathbb{R}$-espace vectoriel E de dimension finie muni d'un produit scalaire (on parle d'*espace vectoriel euclidien*[1]).

Définition 3.4 *Soit $P \in \mathcal{M}_n(\mathbb{K})$. La matrice P est orthogonale lorsque $P \times {}^tP = {}^tP \times P = I_n$, ce qui est équivalent à ${}^tP = P^{-1}$.*

Si $\mathcal{B} = \{\vec{e}_1, \dots, \vec{e}_n\}$ est une base orthonormée de E et $P = \mathrm{Mat}(\{\vec{e}\,'_1, \dots, \vec{e}\,'_n\}; \mathcal{B})$ où $\vec{e}\,'_j = \sum_{i=1}^{n} a_{ij}\vec{e}_i$, $1 \leqslant j \leqslant n$, on a

$${}^tP \times P = \left(\sum_{k=1}^{k} a_{ki}a_{kj}\right)_{1 \leqslant i,j \leqslant n} = (\vec{e}\,'_i.\vec{e}\,'_j)_{1 \leqslant i,j \leqslant n},$$

et alors la matrice P est orthogonale si et seulement si $(\vec{e}\,'_i.\vec{e}\,'_j)_{1 \leqslant i,j \leqslant n} = I_n$ ou encore

$$\forall i, j \in \{1, \dots, n\},\ \vec{e}\,'_i.\vec{e}\,'_j = 0 \text{ si } i \neq j \quad \text{et} \quad \|\vec{e}\,'_i\| = 1.$$

Les relations ci-dessus expriment que le système de vecteurs $\{\vec{e}\,'_1, \dots, \vec{e}\,'_n\}$ est une base orthonormée de E (cf. chapitre 6). On a ainsi prouvé le résultat suivant.

Proposition 3.5 *Soit $\mathcal{B}$ une base orthonormée de E, $\mathcal{B}'$ une base de E et $P = P_{\mathcal{B} \to \mathcal{B}'}$ la matrice de passage de la base $\mathcal{B}$ vers la base $\mathcal{B}'$. La matrice P est orthogonale si et seulement si la base $\mathcal{B}'$ est orthonormée.*

1. Euclide : mathématicien grec du III^{e}siècle av. J.-C.

La relation $P^{-1} = {}^tP$ s'écrit également $P_{\mathcal{B}'\to\mathcal{B}} = {}^tP_{\mathcal{B}\to\mathcal{B}'}$, cette dernière égalité exprime que le tableau $(a_{ij})_{1\leqslant i,j\leqslant n}$, non seulement lu **verticalement** fournit les coordonnées des vecteurs de $\mathcal{B}'$ en fonction de ceux de $\mathcal{B}$ (comme d'habitude), mais lu aussi **horizontalement** fournit les coordonnées des vecteurs de $\mathcal{B}$ en fonction de ceux de $\mathcal{B}'$. Énonçons enfin le résultat fort important concernant les matrices symétriques réelles.

Théorème 3.6 *Soit A une matrice symétrique réelle. Alors, toutes les valeurs propres de A sont réelles et les sous-espaces propres associés sont deux à deux orthogonaux. En d'autres termes, E admet une base orthonormée constituée de vecteurs propres de A, ou encore la matrice A est diagonalisable via une matrice de passage orthogonale.*

Exemple.

Considérons la matrice symétrique réelle $A = \begin{pmatrix} 3 & 1 & 1 \\ 1 & 3 & 1 \\ 1 & 1 & 3 \end{pmatrix}$.

— *Recherche des valeurs propres :* évaluons le polynôme caractéristique de A. Dans les calculs qui suivent, le deuxième déterminant se déduit du premier en remplaçant la première colonne par la somme des trois colonnes, et le troisième se déduit du second en retranchant la première ligne aux deuxième et troisième lignes :

$$\begin{aligned}\Delta_A(\lambda) &= \begin{vmatrix} 3-\lambda & 1 & 1 \\ 1 & 3-\lambda & 1 \\ 1 & 1 & 3-\lambda \end{vmatrix} = \begin{vmatrix} 5-\lambda & 1 & 1 \\ 5-\lambda & 3-\lambda & 1 \\ 5-\lambda & 1 & 3-\lambda \end{vmatrix} \\ &= \begin{vmatrix} 5-\lambda & 1 & 1 \\ 0 & 2-\lambda & 0 \\ 0 & 0 & 2-\lambda \end{vmatrix} = (5-\lambda)(2-\lambda)^2.\end{aligned}$$

On voit donc que la matrice A admet la valeur propre simple 5 et la valeur propre double 2 ; toutes ses valeurs propres sont réelles.

— *Recherche des sous-espaces propres :*

- pour la valeur propre 5, on a

$$A - 5I_3 = \begin{pmatrix} -2 & 1 & 1 \\ 1 & -2 & 1 \\ 1 & 1 & -2 \end{pmatrix}.$$

Le sous-espace propre associé est caractérisé par le système d'équations

$$\left\{\begin{aligned} -2\,x + y + z &= 0 \\ x - 2\,y + z &= 0 \\ x + y - 2\,z &= 0 \end{aligned}\right.$$

ou encore $x = y = z$, c'est la droite vectorielle

$$E_5 = \operatorname{vect}\{(1,1,1)\}\,;$$

- pour la valeur propre 2, on a

$$A - 2I_3 = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}.$$

Le sous-espace propre associé est caractérisé par l'équation $x+y+z=0$, c'est le plan vectoriel
$$E_2 = \mathrm{vect}\{(1,-1,0),(0,1,-1)\}.$$
La dimension de E_2 coïncide avec la multiplicité de la valeur propre 2 et ceci montre bien que A est diagonalisable.

— *Diagonalisation :* on constate effectivement que les sous-espaces propres E_5 et E_2 sont orthogonaux. On peut donc choisir une base orthonormée formée de vecteurs propres, que l'on va construire à partir de la base de vecteurs propres $\{(1,1,1),(1,-1,0),(0,1,-1)\}$. Cherchons un vecteur (x,y,z) du plan vectoriel E_2 orthogonal au vecteur propre $(1,-1,0)$: il vérifie le système d'équations $\begin{cases} x-y & = 0 \\ x+y+z & = 0 \end{cases}$ qui se simplifie selon $\begin{cases} y & = x \\ z & = -2x \end{cases}$. Par exemple, $(1,1,-2)$ est un vecteur orthogonal à $(1,-1,0)$ appartenant à E_2. Les trois vecteurs $(1,1,1),(1,-1,0),(1,1,-2)$ sont donc des vecteurs propres deux à deux orthogonaux. Pour en déduire une base orthonormée, il suffit de les normer. La norme $(1,1,1)$ est $\sqrt{3}$, celle de $(1,-1,0)$ est $\sqrt{2}$ et celle de $(1,1,-2)$ est $\sqrt{6}$; il suffit alors de les multiplier respectivement par $1/\sqrt{3}$, $1/\sqrt{2}$ et $1/\sqrt{6}$. Cela donne la matrice de passage orthogonale
$$P = \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & 0 & -\frac{2}{\sqrt{6}} \end{pmatrix}$$
et la décomposition recherchée de A est donc
$$A = PD\,{}^tP \quad \text{avec} \quad D = \begin{pmatrix} 5 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix} \quad \text{et } P^{-1} = {}^tP.$$

Compléments

On suppose jusqu'à la fin de ce paragraphe que $\mathbb{K} = \mathbb{C}$.

1. *Trigonalisation, réduction de Jordan*
   Lorsque l'endomorphisme f n'est pas diagonalisable, il est toutefois possible de trouver une base relativement à laquelle la matrice de f est triangulaire. On dit que l'endomorphisme f est *trigonalisable*. Notons $\lambda_1,\dots,\lambda_k$ les valeurs propres complexes distinctes de f et $n_1,\dots,n_k$ leurs multiplicités respectives. La théorie de Jordan[2] stipule qu'il existe une base de E relativement à laquelle la matrice de f est
$$T = \begin{pmatrix} \boxed{J_{\lambda_1,n_1}} & O & \dots & O \\ O & \boxed{J_{\lambda_2,n_2}} & & \vdots \\ \vdots & & \ddots & \vdots \\ O & \dots & \dots & \boxed{J_{\lambda_k,n_k}} \end{pmatrix}$$
   où les blocs diagonaux sont donnés par des matrices de Jordan :
$$J_{\lambda_i,n_i} = \begin{pmatrix} \lambda_i & \varepsilon_{i,1} & 0 & \dots & 0 \\ 0 & \lambda_i & \varepsilon_{i,2} & & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ \vdots & & & \lambda_i & \varepsilon_{i,n_i-1} \\ 0 & \dots & \dots & 0 & \lambda_i \end{pmatrix} \in \mathcal{M}_{n_i}(\mathbb{K}),$$

2. Jordan, Camille : mathématicien français (Lyon 1838–Paris 1922)

les $\varepsilon_{i,j}$, $1 \leqslant j \leqslant n_i - 1$, valant 0 ou 1. De manière plus précise, en posant $m_i = \dim(E_{\lambda_i})$,

$$J_{\lambda_i,n_i} = \begin{pmatrix} \boxed{J_{\lambda_i}^{(1)}} & O & \dots & O \\ O & \boxed{J_{\lambda_i}^{(2)}} & & \vdots \\ \vdots & & \ddots & \vdots \\ O & \dots & \dots & \boxed{J_{\lambda_i}^{(m_i)}} \end{pmatrix}$$

où les $J_{\lambda_i}^{(j)}$ sont des matrices de Jordan élémentaires de la forme

$$\begin{pmatrix} \lambda_i & 1 & 0 & \dots & 0 \\ 0 & \lambda_i & 1 & & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ \vdots & & & \lambda_i & 1 \\ 0 & \dots & \dots & 0 & \lambda_i \end{pmatrix}.$$

On peut écrire la matrice triangulaire T comme étant la somme d'une matrice diagonale D et d'une matrice triangulaire N n'ayant que des zéros sur la diagonale (*décomposition de Dunford*[3]) : $T = D + N$ avec

$$D = \begin{pmatrix} \boxed{\lambda_1 I_{n_1}} & O & \dots & O \\ O & \boxed{\lambda_2 I_{n_2}} & & \vdots \\ \vdots & & \ddots & \vdots \\ O & \dots & \dots & \boxed{\lambda_k I_{n_k}} \end{pmatrix} \quad \text{et} \quad N = \begin{pmatrix} \boxed{N_{\lambda_1,n_1}} & O & \dots & O \\ O & \boxed{N_{\lambda_2,n_2}} & & \vdots \\ \vdots & & \ddots & \vdots \\ O & \dots & \dots & \boxed{N_{\lambda_k,n_k}} \end{pmatrix}$$

où

$$N_{\lambda_i,n_i} = J_{\lambda_i,n_i} - \lambda_i I_{n_i} = \begin{pmatrix} 0 & \varepsilon_{i,1} & \dots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \varepsilon_{i,n_i-1} \\ 0 & \dots & \dots & 0 \end{pmatrix} \in \mathcal{M}_{n_i}(\mathbb{K}).$$

Globalement, N est de la forme

$$N = \begin{pmatrix} 0 & \varepsilon_1 & \dots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \varepsilon_{n-1} \\ 0 & \dots & \dots & 0 \end{pmatrix} \in \mathcal{M}_n(\mathbb{K})$$

où les $\varepsilon_1, \dots, \varepsilon_n$ valent 0 ou 1. Les matrices N_{λ_i,n_i} sont nilpotentes : $N_{\lambda_i,n_i}^{n_i} = O_{n_i}$, et il en est de même pour la matrice N : en posant $m = \max(n_1, \dots, n_k)$, on a $N^m = O_n$. De plus, les matrices D et N commutent, cette information est cruciale pour pouvoir calculer les puissances de T à l'aide de la formule du binôme.

Exemples.

(a). Considérons la matrice $A = \begin{pmatrix} 2 & 1 & -2 \\ 15 & 6 & -11 \\ 14 & 6 & -11 \end{pmatrix}$. Introduisons le $\mathbb{R}$-espace vectoriel $E = \mathbb{R}^3$ rapporté à sa base canonique $\mathcal{B}$ et $f \in \mathcal{L}(E)$ l'endomorphisme dont la matrice relativement à $\mathcal{B}$ est A : $A = \mathrm{Mat}(f; \mathcal{B})$. Calculons le polynôme caractéristique de A. On a indiqué à chaque étape les opérations

3. Dunford, Nelson : mathématicien américain (Saint-Louis 1906–Sarasota 1986)

effectuées sur les lignes ou colonnes :

$$\Delta_A(\lambda)=\begin{vmatrix}2-\lambda & 1 & -2\\ 15 & 6-\lambda & -11\\ 14 & 6 & -11-\lambda\end{vmatrix} \underset{\downarrow}{\overset{(L_2\to L_2-L_3)}{=}} \begin{vmatrix}2-\lambda & 1 & -2\\ 1 & -\lambda & \lambda\\ 14 & 6 & -11-\lambda\end{vmatrix}$$

$$\overset{\substack{(C_2\to C_2+\lambda C_1\\ C_3\to C_3+C_2)\\ \downarrow}}{=} \begin{vmatrix}2-\lambda & 1+2\lambda-\lambda^2 & -1\\ 1 & 0 & 0\\ 14 & 6+14\lambda & -5-\lambda\end{vmatrix} \overset{\substack{\text{(développement par}\\ \text{rapport à la 2}^{\text{e}}\text{ ligne)}\\ \downarrow}}{=} -\begin{vmatrix}1+2\lambda-\lambda^2 & -1\\ 6+14\lambda & -5-\lambda\end{vmatrix}$$

$$\overset{\substack{(C_1\to C_1-2C_2)\\ \downarrow}}{=} \begin{vmatrix}3+2\lambda-\lambda^2 & 1\\ 16+16\lambda & 5+\lambda\end{vmatrix} \overset{\substack{\text{(factorisation}\\ \text{par }1+\lambda)\\ \downarrow}}{=} (1+\lambda)\begin{vmatrix}3-\lambda & 1\\ 16 & 5+\lambda\end{vmatrix}$$

$$\overset{\substack{(C_1\to C_1-4C_2)\\ \downarrow}}{=} (1+\lambda)\begin{vmatrix}-1-\lambda & 1\\ -4-4\lambda & 5+\lambda\end{vmatrix} \overset{\substack{\text{(factorisation}\\ \text{par }1+\lambda)\\ \downarrow}}{=} -(1+\lambda)^2\begin{vmatrix}1 & 1\\ 4 & 5+\lambda\end{vmatrix}$$

$$= -(1+\lambda)^3.$$

La matrice A admet la valeur propre triple -1. Il est clair qu'elle n'est pas diagonalisable sinon elle serait semblable à $-I_3$ et serait donc égale à $-I_3$ ce qui n'est pas le cas. Déterminons le sous-espace propre associé à -1. On a $A+I_3=\begin{pmatrix}3 & 1 & -2\\ 15 & 7 & -11\\ 14 & 6 & -10\end{pmatrix}$ et E_{-1} est caractérisé par le système ci-dessous :

$$\left\{\begin{aligned}3x+\ y-\ 2z&=0\\ 15x+7y-11z&=0\\ 14x+6y-10z&=0\end{aligned}\right. \iff \left\{\begin{aligned}3x+\ y&=2z\\ 15x+7y&=11z\end{aligned}\right. \iff x=y=z/2.$$

Le sous-espace propre E_{-1} est donc la droite vectorielle engendrée par $(1,1,2)$. La théorie de Jordan nous enseigne que la matrice A est semblable soit à la matrice $\begin{pmatrix}-1 & 1 & 0\\ 0 & -1 & 0\\ 0 & 0 & -1\end{pmatrix}$ soit à la matrice $\begin{pmatrix}-1 & 1 & 0\\ 0 & -1 & 1\\ 0 & 0 & -1\end{pmatrix}$ (notons que $\begin{pmatrix}-1 & 1 & 0\\ 0 & -1 & 0\\ 0 & 0 & -1\end{pmatrix}$ est semblable à $\begin{pmatrix}-1 & 0 & 0\\ 0 & -1 & 1\\ 0 & 0 & -1\end{pmatrix}$ comme cela se voit aisément par échanges de lignes et de colonnes). La première matrice conduit à un sous-espace propre de dimension 2, elle est par conséquent à exclure. Recherchons donc une base $\mathcal{B}'=\{\vec{u},\vec{v},\vec{w}\}$ de E relativement à laquelle f admette pour matrice $T=\begin{pmatrix}-1 & 1 & 0\\ 0 & -1 & 1\\ 0 & 0 & -1\end{pmatrix}$. Les vecteurs $\vec{u},\vec{v},\vec{w}$ vérifient nécessairement les équations

$$\left\{\begin{aligned}f(\vec{u})&=-\vec{u}\\ f(\vec{v})&=-\vec{v}+\vec{u}\\ f(\vec{w})&=-\vec{w}+\vec{v}\end{aligned}\right..$$

La première équation signifie que $\vec{u}\in E_{-1}$. Prenons par exemple $\vec{u}=(1,1,2)$. Cherchons $\vec{v}=(x,y,z)$ tel que $f(\vec{v})=-\vec{v}+\vec{u}$. Cela conduit au système

$$\left\{\begin{aligned}3x+\ y-\ 2z&=1\\ 15x+7y-11z&=1\\ 14x+6y-10z&=2\end{aligned}\right. \iff \left\{\begin{aligned}x&=\frac{z}{2}+1\\ y&=\frac{z}{2}-2\end{aligned}\right..$$

Par exemple, $\vec{v}=(1,-2,0)$ convient. Cherchons enfin $\vec{w}=(x,y,z)$ tel que $f(\vec{w})=-\vec{w}+\vec{v}$. On a cette fois le système

$$\left\{\begin{aligned}3x+\ y-\ 2z&=1\\ 15x+7y-11z&=-2\\ 14x+6y-10z&=0\end{aligned}\right. \iff \left\{\begin{aligned}x&=\frac{z+3}{2}\\ y&=\frac{z-7}{2}\end{aligned}\right.$$

qui fournit par exemple $\vec{w}=(2,-3,1)$. Il reste à vérifier que le système de vecteurs $\{\vec{u},\vec{v},\vec{w}\}$ est une base de E. Il suffit de voir qu'il est libre. On peut le vérifier soit directement à partir des expressions numériques des vecteurs $\vec{u},\vec{v},\vec{w}$, soit à partir des équations satisfaites par $\vec{u},\vec{v},\vec{w}$. Adoptons la deuxième approche qui a le mérite de pouvoir se généraliser le cas échéant. Soit $\alpha,\beta,\gamma\in\mathbb{R}$ tels que

$$\alpha\vec{u}+\beta\vec{v}+\gamma\vec{w}=\vec{0}.$$

Appliquons $(f+\mathrm{id}_E)$ à cette égalité. Puisque $(f+\mathrm{id}_E)(\vec{u})=\vec{0}$, $(f+\mathrm{id}_E)(\vec{v})=\vec{u}$, $(f+\mathrm{id}_E)(\vec{w})=\vec{v}$, on trouve

$$\beta\vec{v}+\gamma\vec{w}=\vec{0}.$$

Appliquons une deuxième fois $(f+\mathrm{id}_E)$: cela donne $\gamma\vec{u}=\vec{0}$, d'où l'on tire $\gamma=0$, puis successivement $\beta=0$ et $\alpha=0$. Le système $\mathcal{B}'=\{\vec{u},\vec{v},\vec{w}\}$ est donc une base de E et la matrice de passage de $\mathcal{B}$ vers $\mathcal{B}'$ est donnée par

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

On a ainsi obtenu la factorisation

$$A=PTP^{-1}\quad\text{avec}\quad P=\begin{pmatrix}1&1&2\\1&-2&-3\\2&0&1\end{pmatrix},\ T=\begin{pmatrix}-1&1&0\\0&-1&1\\0&0&-1\end{pmatrix},\ P^{-1}=\begin{pmatrix}2&1&-1\\7&3&-5\\-4&-2&3\end{pmatrix}.$$

(b). Considérons la matrice $A=\begin{pmatrix}-5&-2&3\\8&3&-6\\0&0&-1\end{pmatrix}$. Son polynôme caractéristique s'évalue comme suit :

$$\Delta_A(\lambda)=\begin{vmatrix}-5-\lambda&-2&3\\8&3-\lambda&-6\\0&0&-1-\lambda\end{vmatrix}\overset{(C_1\to C_1-2C_2)}{\underset{\downarrow}{=}}\begin{vmatrix}-1-\lambda&-2&3\\2+2\lambda&3-\lambda&-6\\0&0&-1-\lambda\end{vmatrix}$$

$$\overset{\text{(factorisation par }\lambda+1}{\underset{\downarrow}{=}}\quad(\lambda+1)\begin{vmatrix}-1&-2&3\\2&3-\lambda&-6\\0&0&-1-\lambda\end{vmatrix}\overset{\substack{\text{(développement par}\\\text{rapport à la }3^{\text{e}}\text{ ligne)}}}{\underset{\downarrow}{=}}-(\lambda+1)^2\begin{vmatrix}-1&-2\\2&3-\lambda\end{vmatrix}$$

$$=\quad-(\lambda+1)^3.$$

On voit que A admet la valeur propre triple -1 et comme dans le cas précédent, A n'est pas diagonalisable. Le sous-espace propre E_{-1} est caractérisé par

$$\begin{cases}-4x-2y+3z=0\\8x+4y-6z=0\end{cases}\iff 4x+2y-3z=2z.$$

Le sous-espace propre E_{-1} est donc le plan vectoriel engendré par les vecteurs $(3,0,4)$ et $(1,-2,0)$. D'après la discussion faite dans le cas précédent, il ressort que cette fois A est semblable à la matrice $T=\begin{pmatrix}-1&0&0\\0&-1&1\\0&0&-1\end{pmatrix}$. Cherchons donc une base $\{\vec{u},\vec{v},\vec{w}\}$ de E relativement à laquelle f admette T pour matrice. Les vecteurs $\vec{u},\vec{v},\vec{w}$ vérifient le système

$$\begin{cases}f(\vec{u})=-\vec{u}\\f(\vec{v})=-\vec{v}\\f(\vec{w})=-\vec{w}+\vec{v}\end{cases}.$$

Il apparaît clairement que $\vec{u},\vec{v}\in E_{-1}$. Choisissons par exemple $\vec{u}=(3,0,4)$ et $\vec{v}=(1,-2,0)$ (en fait ce choix n'est pas complètement arbitraire), puis cherchons $\vec{w}=(x,y,z)$ tel que $f(\vec{w})=-\vec{w}+\vec{v}$. On tombe sur le système

$$\begin{cases}-4x-2y+3z=1\\8x+4y-6z=-2\\0=0\end{cases}\iff 4x+2y-3z=-1.$$

Signalons que la dernière équation du système ci-dessus $(0=0)$ limite l'arbitraire du choix du vecteur $\vec{v}$; sa troisième coordonnée devait être nécessairement nulle. Maintenant, on choisit $\vec{w}=(0,1,1)$. On peut démontrer par une méthode similaire à celle précédemment utilisée que le système de vecteurs $\{\vec{u},\vec{v},\vec{w}\}$ est une base de E. Une autre technique consiste à remarquer que, puisque $\vec{u}\in E_{-1}$,

$$f(\vec{w})=-\vec{w}+\vec{v}\iff(f+\mathrm{id}_E)(\vec{w})=\vec{v}\Longrightarrow(f+\mathrm{id}_E)^2(\vec{w})=\vec{0}\Longrightarrow\vec{w}\in\mathrm{Ker}(f+\mathrm{id}_E)^2.$$

Ici, $(A+I_3)^2=\begin{pmatrix}-4&-2&3\\8&4&-6\\0&0&0\end{pmatrix}^2=O_3$ et donc $\mathrm{Ker}(f+\mathrm{id}_E)^2=E$. Ceci amène à choisir $\vec{w}\in\mathrm{Ker}(f+\mathrm{id}_E)^2\setminus\mathrm{Ker}(f+\mathrm{id}_E)$, par exemple $\vec{w}=(0,1,1)$, puis poser $\vec{v}=(f+\mathrm{id}_E)(\vec{w})=(1,-2,0)$, et l'on trouve une base convenable. D'où la décomposition

$$A=PTP^{-1}\quad\text{avec}\quad P=\begin{pmatrix}3&1&0\\0&-2&1\\4&0&1\end{pmatrix},\ T=\begin{pmatrix}-1&0&0\\0&-1&1\\0&0&-1\end{pmatrix},\ P^{-1}=\frac{1}{2}\begin{pmatrix}2&1&-1\\-4&-3&3\\-8&-4&6\end{pmatrix}.$$

2. *Théorème de Cayley-Hamilton*

Si P est un polynôme défini par $P(\lambda) = \sum_{i=0}^{n} a_i \lambda^i$, on peut étendre la définition de ce polynôme aux matrices carrées M en posant $P(M) = \sum_{i=0}^{n} a_i M^i$.

Théorème 3.7 (Théorème de Cayley[4]-Hamilton[5]) *Pour toute matrice carrée A, on a $\Delta_A(A) = O$.*

Démonstration.

Utilisons la forme factorisée du polynôme Δ_A : $\Delta_A(\lambda) = \prod_{i=1}^{k} (\lambda_i - \lambda)^{n_i}$. On a $\Delta_A(A) = \prod_{i=1}^{k} (\lambda_i I_n - A)^{n_i}$. Invoquant la décomposition de Jordan précédente, on trouve

$$A - \lambda_1 I_n = \begin{pmatrix} \boxed{N_{\lambda_1,n_1}} & O & \dots & O \\ O & \boxed{N_{\lambda_2,n_2} + (\lambda_2 - \lambda_1) I_{n_2}} & & \vdots \\ \vdots & & \ddots & \vdots \\ O & \dots & \dots & \boxed{N_{\lambda_k,n_k} + (\lambda_k - \lambda_1) I_{n_k}} \end{pmatrix},$$

et donc, puisque $N_{\lambda_1,n_1}^{n_1} = 0$,

$$(A - \lambda_1 I_n)^{n_1} = \begin{pmatrix} \boxed{O_{n_1}} & O & \dots & O \\ O & \boxed{(N_{\lambda_2,n_2} + (\lambda_2 - \lambda_1) I_{n_2})^{n_1}} & & \vdots \\ \vdots & & \ddots & \vdots \\ O & \dots & \dots & \boxed{(N_{\lambda_k,n_k} + (\lambda_k - \lambda_1) I_{n_k})^{n_1}} \end{pmatrix}.$$

De même,

$$(A - \lambda_2 I_n)^{n_2} = \begin{pmatrix} \boxed{(N_{\lambda_1,n_1} + (\lambda_1 - \lambda_2) I_{n_1})^{n_2}} & O & \dots & O \\ O & \boxed{O_{n_2}} & & \vdots \\ \vdots & & \ddots & \vdots \\ O & \dots & \dots & \boxed{(N_{\lambda_k,n_k} + (\lambda_k - \lambda_2) I_{n_k})^{n_2}} \end{pmatrix}$$

et ainsi de suite jusqu'à

$$(A - \lambda_k I_n)^{n_k} = \begin{pmatrix} \boxed{(N_{\lambda_1,n_1} + (\lambda_1 - \lambda_k) I_{n_1})^{n_k}} & O & \dots & O \\ O & \boxed{(N_{\lambda_2,n_2} + (\lambda_2 - \lambda_k) I_{n_2})^{n_k}} & & \vdots \\ \vdots & & \ddots & \vdots \\ O & \dots & \dots & \boxed{O_{n_k}} \end{pmatrix}.$$

Finalement, effectuer le produit $(A - \lambda_1 I_n)^{n_1} \times (A - \lambda_2 I_n)^{n_2} \times \dots \times (A - \lambda_k I_n)^{n_k}$ revient à effectuer k produits de blocs diagonaux, chacun de ces produits contenant un bloc nul : tous ces produits sont nuls. D'où

$$\Delta_A(A) = (-1)^n (A - \lambda_1 I_n)^{n_1} \times (A - \lambda_2 I_n)^{n_2} \times \dots \times (A - \lambda_k I_n)^{n_k} = O_n. \qquad \square$$

Application : expression polynomiale de l'inverse d'une matrice

En utilisant la forme développée de Δ_A, on trouve la relation

$$\Delta_A(A) = a_n A^n + a_{n-1} A^{n-1} + \dots + a_0 I_n = O_n$$

4. Cayley, Arthur : mathématicien anglais (Richmond 1821–Cambridge 1895)
5. Hamilton, sir William Rowan : astronome et mathématicien irlandais (Dublin 1805–Dublin 1865)

avec
$$a_n = (-1)^n,\ a_{n-1} = (-1)^{n-1}\mathrm{tr}(A), \cdots, a_0 = \det(A).$$
Supposons la matrice A inversible ; on a donc $\det(A) \neq 0$. La relation précédente fournit une expression polynomiale de A^{-1}. En effet, en écrivant
$$A \times [a_n A^{n-1} + a_{n-1} A^{n-2} + \cdots + a_1 I_n] = -\det(A) I_n,$$
on tire
$$A^{-1} = -\frac{1}{\det(A)}\,[a_n A^{n-1} + a_{n-1} A^{n-2} + \cdots + a_1 I_n].$$

Remarque.

En identifiant les formes factorisée et développée de Δ_A, on voit que le coefficient a_1 est l'opposé de la somme des produits des valeurs propres de A prises $(n-1)$ à $(n-1)$, soit encore le produit de toutes les valeurs propres multiplié par la somme des inverses des valeurs propres. Or cette dernière n'est autre que la trace de T^{-1}. Comme A est semblable à T, A^{-1} est semblable à T^{-1} et $\mathrm{tr}(T^{-1}) = \mathrm{tr}(A^{-1})$. En conséquence, $a_1 = -\det(A)\,\mathrm{tr}(A^{-1})$.

Un exemple de matrice nilpotente

Une matrice carrée A d'ordre n est *nilpotente* lorsqu'il existe un entier $m \in \mathbb{N}^*$ tel que $A^m = O_n$. Un endomorphisme f associé à une matrice nilpotente est dit *nilpotent* ; il vérifie de même $f^m = 0$ pour un certain entier m. Soit $A \in \mathcal{M}_n(\mathbb{C})$ une matrice nilpotente et f un endomorphisme (nilpotent) de E dont A est la matrice relativement à une base de E.

— Soit λ une valeur propre de f. Il existe un vecteur $\vec{u} \in E \setminus \{\vec{0}\}$ tel que $f(\vec{u}) = \lambda\vec{u}$. En appliquant $(m-1)$ fois f à cette égalité, on trouve la relation $\vec{0} = f^m(\vec{u}) = \lambda^m \vec{u}$ de laquelle on tire $\lambda = 0$. On en déduit que 0 est la seule valeur propre de A et qu'elle est donc de multiplicité n. La matrice A est diagonalisable uniquement dans le cas où elle est nulle. Le polynôme caractéristique de A s'écrit simplement $\Delta_A(\lambda) = (-\lambda)^n$ et le théorème de Cayley-Hamilton assure alors que $A^n = O_n$. Le plus petit entier $m \in \mathbb{N}^*$ tel que $A^m = O_n$ est l'*indice de nilpotence de* A, on voit qu'il est nécessairement inférieur ou égal à l'ordre n de la matrice A.

— Supposons la matrice A d'indice de nilpotence exactement n ; on a donc $A^n = O_n$ et $A^{n-1} \neq O_n$, ou encore $f^n = 0$ et $f^{n-1} \neq 0$. On va construire une base de E relativement à laquelle la matrice de f est très simple. Dans ce but, choisissons un vecteur $\vec{u}_n \in E \setminus \mathrm{Ker}(f^{n-1})$, puis posons successivement
$$\begin{array}{lcl}
\vec{u}_{n-1} & = & f(\vec{u}_n),\\
\vec{u}_{n-2} & = & f(\vec{u}_{n-1}) = f^2(\vec{u}_n),\\
\vec{u}_{n-3} & = & f(\vec{u}_{n-2}) = f^2(\vec{u}_{n-1}) = f^3(\vec{u}_n),\\
\ \vdots & & \\
\ \vec{u}_1 & = & f(\vec{u}_2) = f^2(\vec{u}_3) = f^3(\vec{u}_4) = \cdots = f^{n-1}(\vec{u}_n).
\end{array}$$
Notons que $f(\vec{u}_1) = f^n(\vec{u}_n) = \vec{0}$ et que, ayant choisi $\vec{u}_n \in E \setminus \mathrm{Ker}(f^{n-1})$, on a $\vec{u}_1 = f^{n-1}(\vec{u}_n) \neq \vec{0}$. Cette remarque permet de voir que le système de vecteurs $\mathcal{B}' = \{\vec{u}_1, \ldots, \vec{u}_n\}$ est une base de E. En effet, il contient n vecteurs et il est libre : si $\alpha_1, \alpha_2, \ldots, \alpha_n$ sont des scalaires tels que $\alpha_1\vec{u}_1 + \alpha_2\vec{u}_2 + \cdots + \alpha_n\vec{u}_n = \vec{0}$, alors en appliquant f plusieurs fois de suite, on obtient
$$\begin{array}{rcl}
\alpha_1\vec{u}_1 + \alpha_2\vec{u}_2 + \cdots + \alpha_n\vec{u}_n = \vec{0} & \Longrightarrow & \alpha_2\vec{u}_1 + \alpha_3\vec{u}_2 + \cdots + \alpha_n\vec{u}_{n-1} = \vec{0}\\
& \Longrightarrow & \alpha_3\vec{u}_1 + \alpha_4\vec{u}_2 + \cdots + \alpha_n\vec{u}_{n-2} = \vec{0}\\
& \cdots & \\
& \Longrightarrow & \alpha_{n-1}\vec{u}_1 + \alpha_n\vec{u}_2 = \vec{0}\\
& \Longrightarrow & \alpha_n\vec{u}_1 = \vec{0}.
\end{array}$$
On a signalé que $\vec{u}_1 \neq \vec{0}$ et par conséquent $\alpha_n = 0$. En remontant ensuite les égalités précédentes, on déduit de proche en proche que tous les α_i sont nuls. Par ailleurs, par construction des vecteurs $\vec{u}_1, \ldots, \vec{u}_n$, la matrice de f dans la base $\mathcal{B}'$ est donnée par
$$\mathrm{Mat}(f;\mathcal{B}') = \begin{pmatrix}
0 & 1 & 0 & \ldots & 0\\
0 & 0 & 1 & & \vdots\\
\vdots & & \ddots & \ddots & 0\\
\vdots & & & 0 & 1\\
0 & \ldots & \ldots & 0 & 0
\end{pmatrix}.$$

C'est une matrice de Jordan élémentaire à laquelle est semblable la matrice A (qui était nilpotente d'indice de nilpotence n).

4 Applications

4.1 Puissances d'une matrice, suites récurrentes linéaires

Si A et B sont des matrices semblables, A^m et B^m le sont aussi via la même matrice de passage : si $A = PBP^{-1}$,

$$\boxed{A^m = PB^mP^{-1}.}$$

— Si A est diagonalisable, on a une décomposition de la forme $A = PDP^{-1}$ avec $D = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$. On a déjà vu que

$$A^m = PD^mP^{-1} \quad \text{avec} \quad D^m = \begin{pmatrix} \lambda_1^m & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & \lambda_n^m \end{pmatrix}.$$

— Si A est trigonalisable, on a une décomposition de la forme $A = PTP^{-1}$ avec $T = D + N$, D étant diagonale, N triangulaire nilpotente, vérifiant $DN = ND$. On peut appliquer la formule du binôme :

$$A^m = PT^mP^{-1} \quad \text{avec} \quad T^m = \sum_{k=0}^{\min(m,n-1)} C_m^k D^{m-k} N^k.$$

Application : systèmes linéaires de suites récurrentes.

Considérons le système de suites récurrentes donné par

$$\begin{cases} u_{m+1}^{(1)} = a_{11}\, u_m^{(1)} + \cdots + a_{1n}\, u_m^{(n)} \\ \quad\vdots \\ u_{m+1}^{(n)} = a_{n1}\, u_m^{(1)} + \cdots + a_{nn}\, u_m^{(n)} \end{cases}.$$

Le système ci-dessus exprime une dépendance linéaire entre les termes de rangs m et $m+1$ des n suites $(u_m^{(1)})_{m\in\mathbb{N}}, \ldots, (u_m^{(n)})_{m\in\mathbb{N}}$. Les coefficients a_{ij}, $1 \leqslant i, j \leqslant n$, sont des scalaires donnés ainsi que les conditions initiales $u_0^{(1)}, \ldots, u_0^{(n)}$. Introduisons les matrices

$$A = \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \ldots & a_{nn} \end{pmatrix} \quad \text{et} \quad U_m = \begin{pmatrix} u_m^{(1)} \\ \vdots \\ u_m^{(n)} \end{pmatrix}.$$

Le système s'écrit sous forme matricielle

$$U_{m+1} = AU_m.$$

Il s'agit d'une suite géométrique matricielle (ou vectorielle) de raison la matrice A et de premier terme U_0. En itérant cette dernière relation, il vient aisément

$$\boxed{U_m = A^mU_0.}$$

D'où l'intérêt de savoir calculer A^m.

Exemples.

1. *Récurrence linéaire à trois indices :* soit $(u_m)_{m\in\mathbb{N}}$ la suite définie sur $\mathbb{C}$ par la relation de récurrence $u_{m+2} = au_{m+1} + bu_m$ avec données initiales u_0 et u_1. On ramène cette récurrence à un système de suites récurrentes à deux indices en introduisant une suite auxiliaire $v_m = u_{m+1}$, ce qui conduit à

$$\left\{\begin{array}{rcr} u_{m+1} & = & v_m \\ v_{m+1} & = & b\,u_m + a\,v_m \end{array}\right. .$$

Ce système est associé à la matrice $A = \begin{pmatrix} 0 & 1 \\ b & a \end{pmatrix}$. Son polynôme caractéristique est $\Delta_A(\lambda) = \lambda^2 - a\lambda - b$. Les valeurs propres sont donc les racines de l'*équation caractéristique*

$$\boxed{\lambda^2 - a\lambda - b = 0.} \tag{EC}$$

— Si (EC) admet deux racines distinctes λ et μ, la matrice A est diagonalisable sur $\mathbb{C}$:

$$A = PDP^{-1} \quad \text{avec} \quad P = \begin{pmatrix} 1 & 1 \\ \lambda & \mu \end{pmatrix},\ D = \begin{pmatrix} \lambda & 0 \\ 0 & \mu \end{pmatrix},\ P^{-1} = \frac{1}{\mu-\lambda}\begin{pmatrix} \mu & -1 \\ -\lambda & 1 \end{pmatrix}.$$

On trouve

$$\begin{aligned} A^m &= PD^mP^{-1} = \frac{1}{\mu-\lambda}\begin{pmatrix} 1 & 1 \\ \lambda & \mu \end{pmatrix} \times \begin{pmatrix} \lambda^m & 0 \\ 0 & \mu^m \end{pmatrix} \times \begin{pmatrix} \mu & -1 \\ -\lambda & 1 \end{pmatrix} \\ &= \frac{1}{\mu-\lambda}\begin{pmatrix} \lambda\mu(\lambda^{m-1} - \mu^{m-1}) & \mu^m - \lambda^m \\ \lambda\mu(\lambda^m - \mu^m) & \mu^{m+1} - \lambda^{m+1} \end{pmatrix} \end{aligned}$$

puis

$$U_m = A^mU_0 = \frac{1}{\mu-\lambda}\begin{pmatrix} \lambda\mu(\lambda^{m-1} - \mu^{m-1}) & \mu^m - \lambda^m \\ \lambda\mu(\lambda^m - \mu^m) & \mu^{m+1} - \lambda^{m+1} \end{pmatrix} \times \begin{pmatrix} u_0 \\ u_1 \end{pmatrix}.$$

En fait, on a seulement besoin de u_m qui est le premier terme de la matrice U_m ; on le sort de celle-là en effectuant le produit $\begin{pmatrix} 1 & 0 \end{pmatrix} \times U_m$. Ainsi,

$$\begin{aligned} (u_m) &= \frac{1}{\mu-\lambda}\begin{pmatrix} 1 & 0 \end{pmatrix} \times \begin{pmatrix} \lambda\mu(\lambda^{m-1} - \mu^{m-1}) & \mu^m - \lambda^m \\ \lambda\mu(\lambda^m - \mu^m) & \mu^{m+1} - \lambda^{m+1} \end{pmatrix} \times \begin{pmatrix} u_0 \\ u_1 \end{pmatrix} \\ &= \left(-\frac{\mu\lambda^m - \lambda\mu^m}{\lambda-\mu}\,u_0 + \frac{\lambda^m - \mu^m}{\lambda-\mu}\,u_1 \right). \end{aligned}$$

Le résultat est de la forme

$$\boxed{u_m = \alpha\lambda^m + \beta\mu^m.}$$

— Si (EC) admet une racine double λ, la matrice A est n'est pas diagonalisable, en revanche elle est trigonalisable sur $\mathbb{C}$:

$$A = PTP^{-1} \quad \text{avec} \quad P = \begin{pmatrix} 1 & 0 \\ \lambda & 1 \end{pmatrix},\ T = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix},\ P^{-1} = \begin{pmatrix} 1 & 0 \\ -\lambda & 1 \end{pmatrix}.$$

On a

$$A^m = PT^mP^{-1} = \begin{pmatrix} 1 & 0 \\ \lambda & 1 \end{pmatrix} \times \begin{pmatrix} \lambda^m & m\lambda^{m-1} \\ 0 & \lambda^m \end{pmatrix} \times \begin{pmatrix} 1 & 0 \\ -\lambda & 1 \end{pmatrix} = \begin{pmatrix} (1-m)\lambda^m & m\lambda^{m-1} \\ -m\lambda^{m+1} & (m+1)\lambda^m \end{pmatrix}$$

puis

$$(u_m) = \begin{pmatrix} 1 & 0 \end{pmatrix} \times \begin{pmatrix} (1-m)\lambda^m & m\lambda^{m-1} \\ -m\lambda^{m+1} & (m+1)\lambda^m \end{pmatrix} \times \begin{pmatrix} u_0 \\ u_1 \end{pmatrix} = (1-m)\lambda^m u_0 + m\lambda^{m-1}u_1.$$

Le résultat est de la forme

$$\boxed{u_m = (\alpha m + \beta)\lambda^m.}$$

2. *Émission d'un signal binaire dans un canal bruité :* une source émet des signaux binaires à travers un canal bruité. La proportion de 0 émis est p ($p \in]0,1[$) et celle de 1 émis est $1-p$. Cette répartition fréquentielle induit un modèle probabiliste de Bernoulli[6] pour l'émission d'un signal générique : la probabilité d'émettre un 0 est p et celle d'émettre un 1 est $1-p$. Le canal est constitué d'une succession de m émetteurs-récepteurs identiques reliés en série qui, agissant comme des sortes de boîtes noires, déforment les signaux avec une certaine probabilité $\varepsilon \in]0,1[$: un 0 (resp. 1) reçu est réémis sans transformation avec probabilité $q = 1-\varepsilon$ et réémis avec modification en 1 (resp. 0) avec probabilité ε (voir Fig. 5.1). On cherche à évaluer les proportions 0 et 1 à la sortie du canal, ainsi que la proportion de signaux correctement restitués.

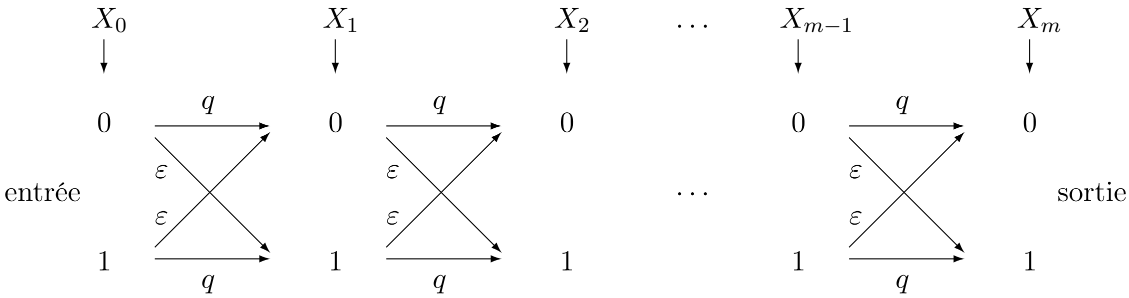

FIGURE 5.1 – Émission d'un signal binaire dans un canal bruité

Notons X_0 le signal émis par la source à l'entrée du canal et pour $1 \leqslant i \leqslant m$, X_i le signal reçu à l'issue du i^{e} émetteur-récepteur. Ce sont des variables aléatoires prenant les deux valeurs 0 et 1. Le problème posé consiste à déterminer les probabilités $\mathbb{P}(X_m = 0)$, $\mathbb{P}(X_m = 1)$ et $\mathbb{P}(X_m = X_0)$. Le transit du signal dans cette i^{e} boîte est schématisé par la probabilité conditionnelle

$$\mathbb{P}(X_i = b | X_{i-1} = a) = \begin{cases} q & \text{si } a = b \\ \varepsilon & \text{si } a \neq b \end{cases}.$$

Rappelons la définition d'une *probabilité conditionnelle* :

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} \quad \text{lorque } \mathbb{P}(B) \neq 0.$$

En écrivant la partition

$$A = A \cap (B \cup \overline{B}) = (A \cap B) \cup (A \cap \overline{B})$$

où $\overline{B}$ dénote le complémentaire de B, on déduit la *formule des probabilités totales*

$$\mathbb{P}(A) = \mathbb{P}(A \cap B) + \mathbb{P}(A \cap \overline{B}) = \mathbb{P}(A|B)\mathbb{P}(B) + \mathbb{P}(A|\overline{B})\mathbb{P}(\overline{B}).$$

6. Bernoulli, Jakob : mathématicien suisse (Bâle 1654–Bâle 1705)

Appliquons cette relation aux événements $A = (X_i = a)$ pour $a \in \{0, 1\}$ et $B = (X_{i-1} = 0)$:

$$\mathbb{P}(X_i = a) = \mathbb{P}(X_i = a | X_{i-1} = 0)\mathbb{P}(X_{i-1} = 0) + \mathbb{P}(X_i = a | X_{i-1} = 1)\mathbb{P}(X_{i-1} = 1).$$

D'après les données, cela donne, en posant $u_i = \mathbb{P}(X_i = 0)$ et $v_i = \mathbb{P}(X_i = 1)$,

$$\left\{ \begin{array}{rcl} u_i &=& q u_{i-1} + \varepsilon v_{i-1} \\ v_i &=& \varepsilon u_{i-1} + q v_{i-1} \end{array} \right. \quad \text{avec} \quad \left\{ \begin{array}{rcl} u_0 &=& p \\ v_0 &=& 1 - p \end{array} \right. ,$$

ou encore sous forme matricielle

$$U_i = A U_{i-1}$$

avec

$$U_i = \begin{pmatrix} u_i \\ v_i \end{pmatrix}, \ A = \begin{pmatrix} q & \varepsilon \\ \varepsilon & q \end{pmatrix}, \ U_0 = \begin{pmatrix} p \\ 1-p \end{pmatrix}.$$

On a $U_m = A^m U_0$. Le polynôme caractéristique de A est $\Delta_A(\lambda) = \lambda^2 - 2q\lambda + (q^2 - \varepsilon^2)$, ce qui fournit les valeurs propres : $q + \varepsilon = 1$ et $q - \varepsilon = 2q - 1$. La diagonalisation de A donne $A = PDP^{-1}$ avec

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

On a

$$\begin{aligned} A^m &= PD^m P^{-1} = \frac{1}{2} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \times \begin{pmatrix} 1 & 0 \\ 0 & (2q-1)^m \end{pmatrix} \times \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \\ &= \frac{1}{2} \begin{pmatrix} 1 + (2q-1)^m & 1 - (2q-1)^m \\ 1 - (2q-1)^m & 1 + (2q-1)^m \end{pmatrix} \end{aligned}$$

et alors

$$U_m = A^m U_0 = \frac{1}{2} \begin{pmatrix} (1-p)[1 + (2q-1)^m] + p[1 - (2q-1)^m] \\ (1-p)[1 - (2q-1)^m] + p[1 + (2q-1)^m] \end{pmatrix} = \begin{pmatrix} \frac{1}{2} - (p - \frac{1}{2})(2q-1)^m \\ \frac{1}{2} + (p - \frac{1}{2})(2q-1)^m \end{pmatrix}.$$

On obtient ainsi la loi de probabilité du signal reçu à la sortie du canal :

$$\boxed{\left\{ \begin{array}{rcl} \mathbb{P}(X_m = 0) &=& \dfrac{1}{2} - \left(p - \dfrac{1}{2}\right)(2q-1)^m \\[2ex] \mathbb{P}(X_m = 1) &=& \dfrac{1}{2} + \left(p - \dfrac{1}{2}\right)(2q-1)^m \end{array} \right. .}$$

Statistiquement, la proportion de 0 en sortie est de $\frac{1}{2} - (p - \frac{1}{2})(2q-1)^m$ et celle de 1 est de $\frac{1}{2} + (p - \frac{1}{2})(2q-1)^m$. Notons que lorsque $m \longrightarrow +\infty$, ces proportions tendent vers $\frac{1}{2}$ puisque $|2q - 1| < 1$ et donc, quelques soient les proportions initiales de 0 et 1, les proportions finales sont équilibrées pour m grands et le signal a été complètement brouillé.

Évaluons enfin la probabilité que le signal initial soit correctement restitué en sortie :

$$\begin{aligned} \mathbb{P}(X_m = X_0) &= \mathbb{P}((X_m = 0) \cap (X_0 = 0)) + \mathbb{P}((X_m = 1) \cap (X_0 = 1)) \\ &= \mathbb{P}(X_m = 0 | X_0 = 0)\mathbb{P}(X_0 = 0) + \mathbb{P}(X_m = 1 | X_0 = 1)\mathbb{P}(X_0 = 1). \end{aligned}$$

On est amené à calculer les probabilités $\mathbb{P}(X_m = 0|X_0 = 0)$ et $\mathbb{P}(X_m = 1|X_0 = 1)$. Posons à cet effet $u_i = \mathbb{P}(X_i = 0|X_0 = a)$ et $v_i = \mathbb{P}(X_i = 1|X_0 = a)$ pour $a \in \{0,1\}$. On va utiliser une formule des probabilités totales conditionnelles. Partant de la partition $A \cap B = (A \cap B \cap C) \cup (A \cap B \cap \overline{C})$, on tire

$$\begin{aligned}
\mathbb{P}(A \cap B) &= \mathbb{P}(A \cap B \cap C) + \mathbb{P}(A \cap B \cap \overline{C}) \\
&= \mathbb{P}(A|B \cap C)\mathbb{P}(B \cap C) + \mathbb{P}(A|B \cap \overline{C})\mathbb{P}(B \cap \overline{C}) \\
&= \mathbb{P}(A|B \cap C)\mathbb{P}(C|B)\mathbb{P}(B) + \mathbb{P}(A|B \cap \overline{C})\mathbb{P}(\overline{C}|B)\mathbb{P}(B),
\end{aligned}$$

puis

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} = \mathbb{P}(A|B \cap C)\mathbb{P}(C|B) + \mathbb{P}(A|B \cap \overline{C})\mathbb{P}(\overline{C}|B).$$

L'application de cette dernière relation aux événements $A = (X_i = b)$, $B = (X_0 = a)$ et $C = (X_{i-1} = 0)$ $(a, b \in \{0,1\})$ conduit à

$$\begin{aligned}
\mathbb{P}(X_i = b|X_0 = a) &= \mathbb{P}(X_i = b|X_{i-1} = 0 \text{ et } X_0 = a)\mathbb{P}(X_{i-1} = 0|X_0 = a) \\
&\quad +\mathbb{P}(X_i = b|X_{i-1} = 1 \text{ et } X_0 = a)\mathbb{P}(X_{i-1} = 1|X_0 = a).
\end{aligned}$$

En supposant que chaque signal en sortie du i^{e} émetteur-récepteur ne subit que l'influence de ce dernier et non des précédents, on peut écrire

$$\mathbb{P}(X_i = b|X_{i-1} = c \text{ et } X_0 = a) = \mathbb{P}(X_i = b|X_{i-1} = c) \text{ pour } c \in \{0,1\}.$$

D'autre part,

$$\mathbb{P}(X_i = b|X_{i-1} = a) = \mathbb{P}(X_1 = b|X_0 = a) = \begin{cases} q & \text{si } a = b \\ \varepsilon & \text{si } a \neq b \end{cases}.$$

Pour $b = 0$ et $b = 1$, cela donne

$$\begin{cases} u_i = qu_{i-1} + \varepsilon v_{i-1} \\ v_i = \varepsilon u_{i-1} + qv_{i-1} \end{cases}.$$

On est ainsi confronté au même système de suites récurrentes mais avec des conditions initiales différentes. Ici,

$$u_0 = \begin{cases} 1 & \text{si } a = 0 \\ 0 & \text{si } a = 1 \end{cases} \quad \text{et} \quad v_0 = \begin{cases} 0 & \text{si } a = 0 \\ 1 & \text{si } a = 1 \end{cases}.$$

La solution est donnée cette fois par

$$\begin{aligned}
\begin{pmatrix} u_m \\ v_m \end{pmatrix} = A^m \begin{pmatrix} 1 \\ 0 \end{pmatrix} &= \frac{1}{2}\begin{pmatrix} 1 + (2q-1)^m \\ 1 - (2q-1)^m \end{pmatrix} \text{ si } a = 0, \\
\begin{pmatrix} u_m \\ v_m \end{pmatrix} = A^m \begin{pmatrix} 0 \\ 1 \end{pmatrix} &= \frac{1}{2}\begin{pmatrix} 1 - (2q-1)^m \\ 1 + (2q-1)^m \end{pmatrix} \text{ si } a = 1.
\end{aligned}$$

La probabilité $\mathbb{P}(X_m = 0|X_0 = 0)$ correspond à u_m avec $a = 0$ alors que la probabilité $\mathbb{P}(X_m = 1|X_0 = 1)$ correspond à v_m avec $a = 1$. En conséquence,

$$\mathbb{P}(X_m = 0|X_0 = 0) = \mathbb{P}(X_m = 1|X_0 = 1) = \frac{1}{2}\,[1 + (2q-1)^m]$$

d'où il vient immédiatement

$$\boxed{\mathbb{P}(X_m = X_0) = \frac{1}{2}\,[1 + (2q-1)^m].}$$

4.2 Exponentielle d'une matrice, systèmes différentiels linéaires

Rappelons la définition de l'exponentielle d'une matrice carrée A (on admettra la convergence de la série) :

$$\exp(A) = \sum_{m=0}^{\infty} \frac{1}{m!} A^m.$$

Si A et B sont des matrices semblables, $\exp(A)$ et $\exp(B)$ le sont aussi via la même matrice de passage : si $A = PBP^{-1}$,

$$\exp(A) = \sum_{m=0}^{\infty} \frac{1}{m!} A^m = P\left(\sum_{m=0}^{\infty} \frac{1}{m!} B^m\right)P^{-1},$$

soit encore

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

— Si A est diagonalisable, on a une décomposition de la forme $A = PDP^{-1}$ avec $D = \mathrm{diag}(\lambda_1, \dots, \lambda_n)$, et alors

$$\exp(A) = P\exp(D)P^{-1} \text{ où } \exp(D) = \begin{pmatrix} e^{\lambda_1} & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & e^{\lambda_n} \end{pmatrix}.$$

— Si A est une matrice carrée d'ordre 2 réelle admettant deux valeurs propres complexes non réelles (elles sont alors nécessairement conjuguées, notons-les $\lambda + i\mu$ et $\lambda - i\mu$ avec $\lambda \in \mathbb{R}$, $\mu \in \mathbb{R}^*$), alors elle est diagonalisable sur $\mathbb{C}$ via une matrice de passage P contenant deux vecteurs conjugués et l'on a $A = PDP^{-1}$ avec $D = \mathrm{diag}(\lambda+i\mu, \lambda-i\mu)$. En introduisant la matrice de passage réelle $\tilde{P}$ telle que $P = \tilde{P}\begin{pmatrix} 1 & 1 \\ i & -i \end{pmatrix}$ (ou encore de manière équivalente $\tilde{P} = \frac{1}{2}P\begin{pmatrix} 1 & -i \\ 1 & i \end{pmatrix}$), on débouche sur

$$\begin{aligned} \exp(A) = P\exp(D)P^{-1} &= \left[\tilde{P}\begin{pmatrix} 1 & 1 \\ i & -i \end{pmatrix}\right] \exp\begin{pmatrix} \lambda + i\mu & 0 \\ 0 & \lambda - i\mu \end{pmatrix} \left[\frac{1}{2}\begin{pmatrix} 1 & -i \\ 1 & i \end{pmatrix} \tilde{P}^{-1}\right] \\ &= \tilde{P}\left[\frac{1}{2}\begin{pmatrix} 1 & 1 \\ i & -i \end{pmatrix}\begin{pmatrix} e^{\lambda+i\mu} & 0 \\ 0 & e^{\lambda-i\mu} \end{pmatrix}\begin{pmatrix} 1 & -i \\ 1 & i \end{pmatrix}\right]\tilde{P}^{-1} = e^{\lambda}\tilde{P}\begin{pmatrix} \cos(\mu) & \sin(\mu) \\ -\sin(\mu) & \cos(\mu) \end{pmatrix}\tilde{P}^{-1}. \end{aligned}$$

Par ailleurs, en écrivant

$$A = PDP^{-1} = \tilde{P}\begin{pmatrix} \lambda & \mu \\ -\mu & \lambda \end{pmatrix}\tilde{P}^{-1}$$

on voit par identification que

$$\exp\begin{pmatrix} \lambda & \mu \\ -\mu & \lambda \end{pmatrix} = e^{\lambda}\begin{pmatrix} \cos(\mu) & \sin(\mu) \\ -\sin(\mu) & \cos(\mu) \end{pmatrix}.$$

— Si A est trigonalisable, on a une décomposition de la forme $A = PTP^{-1}$ avec $T = D+N$, D étant diagonale, N triangulaire nilpotente, vérifiant $DN = ND$. On a aussi dans ce cas $\exp(A) = P\exp(T)P^{-1}$. On peut appliquer la formule du produit de deux exponentielles :

$$\exp(T) = \exp(D+N) = \exp(D)\exp(N) \quad \text{avec} \quad \exp(N) = \sum_{m=0}^{n-1} \frac{1}{m!} N^m.$$

La matrice triangulaire T est constituée de blocs de Jordan du type

$$J_\lambda = \begin{pmatrix} \lambda & 1 & 0 & \dots & 0 \\ 0 & \lambda & 1 & & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ \vdots & & & \lambda & 1 \\ 0 & \dots & \dots & 0 & \lambda \end{pmatrix} = \lambda I_i + J_0 \in \mathcal{M}_i(\mathbb{K}).$$

Calculons l'exponentielle d'un tel bloc :

$$\exp(J_\lambda) = \exp(\lambda I_i + J_0) = \exp(\lambda I_i)\exp(J_0).$$

On a d'une part

$$\exp(\lambda I_i) = \begin{pmatrix} e^\lambda & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & e^\lambda \end{pmatrix} = e^\lambda I_i$$

et d'autre part, en observant que

$$J_0 = \begin{pmatrix} 0 & 1 & 0 & \dots & \dots & 0 \\ 0 & 0 & 1 & & & \vdots \\ \vdots & & \ddots & \ddots & & \vdots \\ \vdots & & & \ddots & 1 & 0 \\ \vdots & & & & 0 & 1 \\ 0 & \dots & \dots & \dots & 0 & 0 \end{pmatrix}, \quad J_0^2 = \begin{pmatrix} 0 & 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 0 & \ddots & & \vdots \\ \vdots & & & \ddots & \ddots & 0 \\ \vdots & & & & 0 & 1 \\ \vdots & & & & 0 & 0 \\ 0 & \dots & \dots & \dots & 0 & 0 \end{pmatrix}, \dots,$$

$$\dots, J_0^{i-2} = \begin{pmatrix} 0 & \dots & 0 & 1 & 0 \\ \vdots & & & 0 & 1 \\ \vdots & & & & 0 \\ \vdots & & & & \vdots \\ 0 & \dots & \dots & \dots & 0 \end{pmatrix}, J_0^{i-1} = \begin{pmatrix} 0 & \dots & 0 & 1 \\ \vdots & & & 0 \\ \vdots & & & \vdots \\ 0 & \dots & \dots & 0 \end{pmatrix} \quad \text{et} \quad J_0^i = O_i,$$

il vient

$$\exp(J_0) = \sum_{m=0}^{i-1} \frac{1}{m!} J_0^m = \begin{pmatrix} 1 & \frac{1}{1!} & \frac{1}{2!} & \frac{1}{3!} & \dots & \frac{1}{(i-2)!} & \frac{1}{(i-1)!} \\ 0 & 1 & \frac{1}{1!} & \frac{1}{2!} & & & \frac{1}{(i-2)!} \\ \vdots & & \ddots & \ddots & \ddots & & \vdots \\ \vdots & & & 1 & \frac{1}{1!} & \frac{1}{2!} & \frac{1}{3!} \\ \vdots & & & & 1 & \frac{1}{1!} & \frac{1}{2!} \\ \vdots & & & & & 1 & \frac{1}{1!} \\ 0 & \dots & \dots & \dots & \dots & 0 & 1 \end{pmatrix}$$

et finalement

$$\exp(J_\lambda) = e^\lambda \exp(J_0).$$

Application : systèmes différentiels linéaires à coefficients constants dans $\mathbb{C}$.

Considérons le système différentiel suivant :

$$\left\{ \begin{array}{rcl} x_1'(t) & = & a_{11}\, x_1(t) + \dots + a_{1n}\, x_n(t) \\ & \vdots & \\ x_n'(t) & = & a_{n1}\, x_1(t) + \dots + a_{nn}\, x_n(t) \end{array} \right. . \qquad \text{(S)}$$

Le système (S) exprime une dépendance linéaire entre les n fonctions $x_1, \ldots, x_n$ et leurs dérivées. Les coefficients a_{ij}, $1 \leqslant i, j \leqslant n$, sont des nombres complexes donnés ainsi que les conditions initiales $x_1(0), \ldots, x_n(0)$. Introduisons les matrices

$$A = \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \ldots & a_{nn} \end{pmatrix} \quad \text{et} \quad X(t) = \begin{pmatrix} x_1(t) \\ \vdots \\ x_n(t) \end{pmatrix}.$$

Le système (S) s'écrit sous forme matricielle

$$X'(t) = AX(t). \tag{E}$$

Il s'agit d'une équation différentielle matricielle (ou vectorielle) linéaire d'ordre 1 à coefficient constant (le coefficient est la matrice $A \in \mathcal{M}_n(\mathbb{C})$).

— *Première approche :* dans le cas de la dimension 1, on sait que la solution s'exprime à l'aide d'une fonction exponentielle. Par analogie, on va utiliser l'exponentielle d'une matrice. Faisons un calcul préliminaire : dérivons $\exp(tA) = \sum_{m=0}^{\infty} \frac{t^m}{m!} A^m$. Cela donne, sans se soucier de problèmes de convergence,

$$\frac{d}{dt}[\exp(tA)] = \sum_{m=1}^{\infty} \frac{t^{m-1}}{(m-1)!} A^m = \sum_{m=0}^{\infty} \frac{t^m}{m!} A^{m+1} = A\exp(tA) = \exp(tA)A.$$

En remarquant alors que $\frac{d}{dt}[\exp(-tA)X(t)] = \exp(-tA)[X'(t) - AX(t)]$, l'équation (E) s'écrit $\frac{d}{dt}[\exp(-tA)X(t)] = O_n$ d'où l'on déduit aisément

$$\boxed{X(t) = \exp(tA)X(0).}$$

On a ainsi exprimé la solution de (E) en fonction de $X(0)$. Si l'on précise des conditions initiales pour $x_1(0), \ldots, x_n(0)$, la solution correspondante s'en déduit aussitôt. Cette formule justifie notre intérêt pour la fonction $t \longmapsto \exp(tA)$ et il est important de savoir calculer $\exp(tA)$.

— *Deuxième approche :* il est souvent possible de transformer le système (S) en un système plus simple dans lequel les équations ne contiennent qu'une fonction inconnue, et ne pas avoir à utiliser l'exponentielle d'une matrice.

- *Cas diagonalisable :* on suppose A diagonalisable de valeurs propres $\lambda_1, \ldots, \lambda_n$, donc $A = PDP^{-1}$ avec $D = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$. Posons $X(t) = PY(t)$ avec $Y(t) = \begin{pmatrix} y_1(t) \\ \vdots \\ y_n(t) \end{pmatrix}$.

  On a $X'(t) = PY'(t)$ que l'on reporte dans l'équation (E) :

  $$Y'(t) = P^{-1}X'(t) = P^{-1}AX(t) = (P^{-1}AP)Y(t) = DY(t).$$

  En écrivant $X'(t) = AX(t) \Longleftrightarrow Y'(t) = DY(t)$, on tombe sur un système différentiel simple, dans lequel les équations ne comportent qu'une seule fonction inconnue,

  $$\left\{ \begin{array}{rcl} y_1'(t) & = & \lambda_1\, y_1(t) \\ & \vdots & \\ y_n'(t) & = & \lambda_n\, y_1(t) \end{array} \right.$$

dont la solution est donnée par

$$\left\{\begin{array}{rcl} y_1(t) & = & \alpha_1\, e^{\lambda_1 t} \\ & \vdots & \\ y_n'(t) & = & \alpha_n\, e^{\lambda_n t} \end{array}\right., \quad \alpha_1,\ldots,\alpha_n \in \mathbb{C},$$

soit encore, en posant $E_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, E_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \ldots, E_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}$,

$$Y(t) = \alpha_1\, e^{\lambda_1 t}\, E_1 + \alpha_2\, e^{\lambda_2 t}\, E_2 + \cdots + \alpha_n\, e^{\lambda_n t}\, E_n.$$

On récupère la solution générale du système différentiel (S) en reportant la solution précédente dans $X(t) = PY(t)$ en observant que $PE_i = V_i$ n'est autre que la matrice-colonne du i^{e} vecteur propre de A choisi dans la matrice de passage P :

$$\boxed{X(t) = \sum_{i=1}^{n} \alpha_i\, e^{\lambda_i t} V_i \quad \text{avec} \quad AV_i = \lambda_i V_i.}$$

Si l'on recherche un solution vérifiant des conditions initiales fixées au préalable (en $t = 0$ par exemple), on aura à déterminer les constantes $\alpha_1, \ldots, \alpha_n$ telles que

$$\sum_{i=1}^{n} \alpha_i V_i = X(0).$$

Il est facile de voir que $\sum_{i=1}^{n} \alpha_i V_i = P \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}$ et donc que ce problème se ramène à l'inversion de la matrice de passage P.

- *Cas non diagonalisable :* lorsque la matrice A n'est pas diagonalisable, elle est trigonalisable, soit $A = PTP^{-1}$. La même méthode conduit à des systèmes différentiels intermédiaires, pour chaque bloc de Jordan

$$J_\lambda = \begin{pmatrix} \lambda & 1 & 0 & \ldots & 0 \\ 0 & \lambda & 1 & & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ \vdots & & & \lambda & 1 \\ 0 & \ldots & \ldots & 0 & \lambda \end{pmatrix} \in \mathcal{M}_i(\mathbb{C}),$$

de la forme

$$\left\{\begin{array}{rcl} y_1'(t) & = & \lambda\, y_1(t) + y_2(t) \\ y_2'(t) & = & \lambda\, y_2(t) + y_3(t) \\ & \vdots & \\ y_{i-1}'(t) & = & \lambda\, y_{i-1}(t) + y_i(t) \\ y_i'(t) & = & \lambda\, y_i(t) \end{array}\right.$$

que l'on résout progressivement en partant de la dernière équation jusqu'à la première.

Complément : systèmes différentiels avec second membre.

Considérons le système différentiel avec second membre

$$X'(t) = AX(t) + B(t),$$

où

— $t \in \mathbb{R} \longmapsto X(t) \in \mathcal{M}_{n1}(\mathbb{C})$ est la fonction inconnue,
— $A \in \mathcal{M}_n(\mathbb{C})$ est une matrice constante donnée,
— $t \in \mathbb{R} \longmapsto B(t) \in \mathcal{M}_{n1}(\mathbb{C})$ est une fonction donnée.

En introduisant le facteur intégrant $\exp(-tA)$, on écrit ce système sous la forme

$$\frac{d}{dt}[\exp(-tA)X(t)] = \exp(-tA)[X'(t) - AX(t)] = \exp(-tA)B(t)$$

de laquelle on tire

$$\exp(-tA)X(t) = X(0) + \int_0^t \exp(-sA)B(s)\,ds.$$

La solution est donc

$$\boxed{X(t) = \exp(tA)X(0) + \int_0^t \exp((t-s)A)B(s)\,ds.}$$

Exemples.

1. *Portraits de phases en dimension 2 sur* $\mathbb{R}$.

(a). *Nœud.* Considérons le système

$$\left\{\begin{array}{rcl} x'(t) &=& 3\,x(t) + 3\,y(t) \\ y'(t) &=& -3\,x(t) + 13\,y(t) \end{array}\right. . \tag{S}$$

La matrice associée est $A = \begin{pmatrix} 3 & 3 \\ -3 & 13 \end{pmatrix}$. Son polynôme caractéristique est donné par $\Delta_A(\lambda) = \lambda^2 - 16\lambda + 48$. On trouve les valeurs propres 4 et 12, puis les sous-espaces propres

$$E_4 = \text{vect}\{(3,1)\} \quad \text{et} \quad E_{12} = \text{vect}\{(1,3)\}.$$

Introduisons donc la matrice de passage $P = \begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix}$ ainsi que la matrice diagonale $\begin{pmatrix} 4 & 0 \\ 0 & 12 \end{pmatrix}$ à laquelle est semblable A, et posons $\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = P \begin{pmatrix} u(t) \\ v(t) \end{pmatrix}$. Les nouvelles fonctions inconnues u, v satisfont au système différentiel

$$\left\{\begin{array}{rcl} u'(t) &=& 4\,u(t) \\ v'(t) &=& 12\,v(t) \end{array}\right.$$

dont la solution est

$$\left\{\begin{array}{rcl} u(t) &=& \alpha\, e^{4t} \\ v(t) &=& \beta\, e^{12t} \end{array}\right.$$

où α, β sont des constantes réelles. Finalement, la solution générale du système (S) est

$$\begin{cases} x(t) = 3\alpha\, e^{4t} + \beta\, e^{12t} \\ y(t) = \alpha\, e^{4t} + 3\beta\, e^{12t} \end{cases}, \quad \alpha, \beta \in \mathbb{R}.$$

Dans la base intermédiaire, les courbes admettent des représentations cartésiennes de la forme

$$v = \beta \left(\frac{u}{\alpha}\right)^3,$$

elles ont une allure parabolique et passent toutes par l'origine O qui est appelée *nœud* (voir Fig. 5.2).

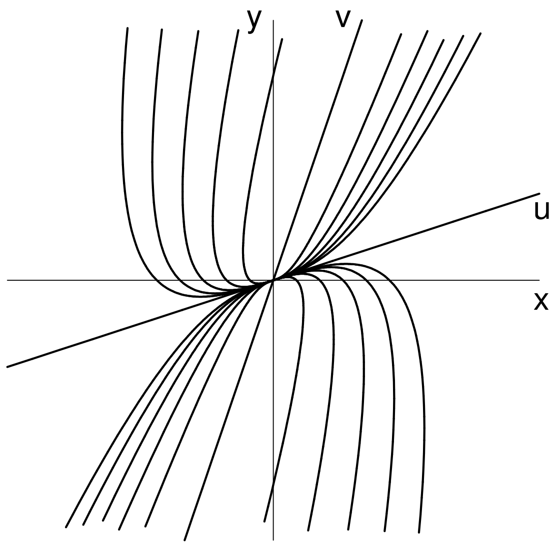

FIGURE 5.2 – Nœud

En adoptant l'approche exponentielle, on aurait obtenu, partant de la décomposition

$$A = \begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix} \times \begin{pmatrix} 4 & 0 \\ 0 & 12 \end{pmatrix} \times \frac{1}{8} \begin{pmatrix} 3 & -1 \\ -1 & 3 \end{pmatrix},$$

$$\exp(tA) = \frac{1}{8} \begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} e^{4t} & 0 \\ 0 & e^{12t} \end{pmatrix} \begin{pmatrix} 3 & -1 \\ -1 & 3 \end{pmatrix} = \frac{1}{8} \begin{pmatrix} 9\, e^{4t} - e^{12t} & -3\, e^{4t} + 3\, e^{12t} \\ 3\, e^{4t} - 3\, e^{12t} & -e^{4t} + 9\, e^{12t} \end{pmatrix}.$$

La solution du système (S) est directement donnée par

$$\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = \frac{1}{8} \begin{pmatrix} 9\, e^{4t} - e^{12t} & -3\, e^{4t} + 3\, e^{12t} \\ 3\, e^{4t} - 3\, e^{12t} & -e^{4t} + 9\, e^{12t} \end{pmatrix} \begin{pmatrix} x(0) \\ y(0) \end{pmatrix},$$

ou encore

$$\begin{cases} x(t) = \frac{1}{8} [(9x(0) - 3y(0))\, e^{4t} + (-x(0) + 3y(0))\, e^{12t}] \\ y(t) = \frac{1}{8} [(3x(0) - y(0))\, e^{4t} + (-3x(0) + 9y(0))\, e^{12t}] \end{cases}$$

qui coïncide avec l'expression précédemment obtenue en prenant $\alpha = \frac{1}{8}(3x(0) - y(0))$ et $\beta = \frac{1}{8}(-x(0) + 3y(0))$.

(b). *Point selle ou col.* Soit le système différentiel

$$\begin{cases} x'(t) = 3\, x(t) - 4\, y(t) \\ y'(t) = 2\, x(t) - 6\, y(t) \end{cases}. \tag{S}$$

Sa matrice associée est $A = \begin{pmatrix} 3 & -4 \\ 2 & -6 \end{pmatrix}$, elle admet pour polynôme caractéristique $\Delta_A(\lambda) = \lambda^2 + 3\lambda - 10$. On trouve les valeurs propres 2 et -5, puis les sous-espaces propres

$$E_2 = \mathrm{vect}\{(4,1)\} \quad \text{et} \quad E_6 = \mathrm{vect}\{(1,2)\},$$

ainsi que la matrice de passage $P = \begin{pmatrix} 4 & 1 \\ 1 & 2 \end{pmatrix}$ et la matrice diagonale $\begin{pmatrix} 2 & 0 \\ 0 & -5 \end{pmatrix}$ correspondantes. En posant $\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = P \begin{pmatrix} u(t) \\ v(t) \end{pmatrix}$, on voit que les nouvelles fonctions inconnues u, v satisfont au système différentiel

$$\begin{cases} u'(t) = 2\,u(t) \\ v'(t) = -5\,v(t) \end{cases}.$$

La solution de ce dernier s'écrit

$$\begin{cases} u(t) = \alpha\, e^{2t} \\ v(t) = \beta\, e^{-5t} \end{cases}$$

où α, β sont des constantes réelles. Enfin, la solution générale du système (S) est

$$\begin{cases} x(t) = 4\alpha\, e^{2t} + \beta\, e^{-5t} \\ y(t) = \alpha\, e^{2t} + 2\beta\, e^{-5t} \end{cases}, \quad \alpha, \beta \in \mathbb{R}.$$

Dans la base intermédiaire, les courbes admettent des représentations cartésiennes de la forme

$$v = \beta \left(\frac{u}{\alpha}\right)^{-5/2},$$

elles ont une allure hyperbolique de centre l'origine O. Le point O porte le nom de *col* ou de *point selle* (voir Fig. 5.3).

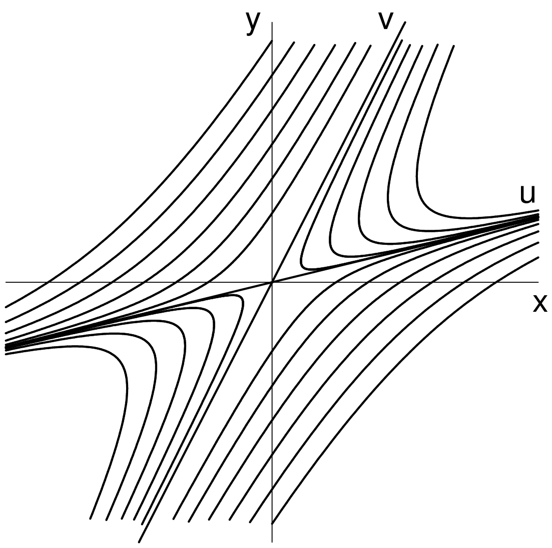

FIGURE 5.3 – Point selle

(c). *Autre nœud.* Considérons le système

$$\begin{cases} x'(t) = -x(t) + y(t) \\ y'(t) = -9\,x(t) + 5\,y(t) \end{cases}. \tag{S}$$

La matrice associée à (S) est $A = \begin{pmatrix} -1 & 1 \\ -9 & 5 \end{pmatrix}$, dont le polynôme caractéristique est $\Delta_A(\lambda) = \lambda^2 - 4\lambda + 4$. On tire la valeur propre double 2, puis le sous-espace propre $E_2 = \mathrm{vect}\{(1,3)\}$. La matrice A n'est donc pas diagonalisable. La théorie de Jordan nous enseigne que la matrice A est toutefois trigonalisable et qu'elle est semblable à la matrice triangulaire $T = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}$. En introduisant l'endomorphisme f du $\mathbb{R}$-espace vectoriel $\mathbb{R}^2$ dont la matrice relativement à la base canonique est A, on est amené à rechercher une base $\{\vec{u}, \vec{v}\}$ relativement à laquelle la matrice de f est T. Ceci donne les équations $f(\vec{u}) = 2\vec{u} + \vec{v}$ et $f(\vec{v}) = 2\vec{v}$. Il suffit de choisir pour $\vec{v}$ un vecteur propre de f associé à la valeur propre 2, *e.g.* $\vec{v} = (1,3)$, puis de trouver un vecteur $\vec{u}$ tel que $(f - \mathrm{id})(\vec{u}) = \vec{v}$. Par exemple, $\vec{u} = (0,1)$ convient. La matrice de passage de la base canonique vers la base $\{\vec{u}, \vec{v}\}$ est $P = \begin{pmatrix} 1 & 0 \\ 3 & 1 \end{pmatrix}$. Effectuons le changement de fonctions $\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = P \begin{pmatrix} u(t) \\ v(t) \end{pmatrix}$. Les nouvelles fonctions inconnues u, v satisfont au système différentiel

$$\begin{cases} u'(t) = 2\,u(t) + v(t) \\ v'(t) = 2\,v(t) \end{cases}.$$

De la deuxième équation on tire $v(t) = \beta\, e^{2t}$ que l'on reporte dans la première :

$$u'(t) = 2\,u(t) + \beta\, e^{2t}.$$

La solution de l'équation homogène associée $u_H'(t) = 2u_H(t)$ est $u_H(t) = \alpha\, e^{2t}$. Le second membre $\beta\, e^{2t}$ étant solution de l'équation homogène, on est dans le cas de résonance et l'on doit rechercher une solution particulière de la forme $u_P(t) = \gamma t\, e^{2t}$. On trouve facilement $\gamma = \beta$, d'où

$$\begin{cases} u(t) = (\alpha + \beta t)\, e^{2t} \\ v(t) = \beta\, e^{2t} \end{cases}$$

où α, β sont des constantes réelles. La solution générale du système (S) est alors

$$\begin{cases} x(t) = (\alpha + \beta t)\, e^{2t} \\ y(t) = (3\alpha + \beta + 3\beta t)\, e^{2t} \end{cases}, \quad \alpha, \beta \in \mathbb{R}.$$

On a en général

$$\lim_{t\to\infty} \frac{y(t)}{x(t)} = 3 \quad \text{et} \quad \lim_{t\to\infty} [y(t) - 3x(t)] = \infty.$$

Les courbes ont une allure asymptotiquement parabolique et l'origine O porte de nouveau le nom de *nœud* (voir Fig. 5.4).

À l'aide de la factorisation

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

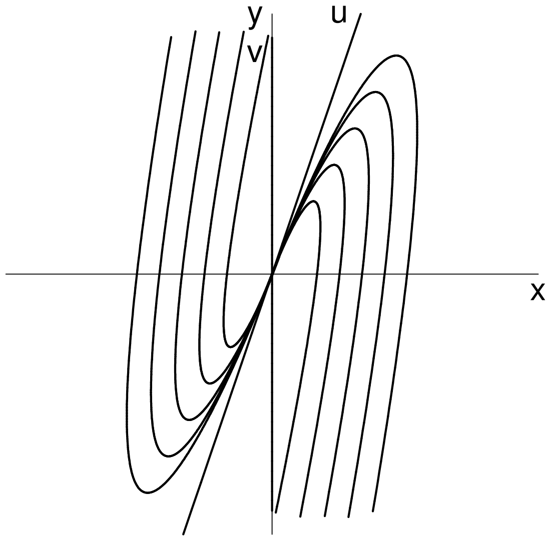

FIGURE 5.4 – Nœud

et de la décomposition

$$T = 2I_2 + \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix},$$

l'approche exponentielle conduit successivement à

$$\exp(tT) = \exp(2tI_2) \times \exp\left[t\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\right] = e^{2t}\left[I_2 + t\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\right] = \begin{pmatrix} e^{2t} & t\,e^{2t} \\ 0 & e^{2t} \end{pmatrix},$$

puis à

$$\exp(tA) = \begin{pmatrix} 1 & 0 \\ 3 & 1 \end{pmatrix}\begin{pmatrix} e^{2t} & t\,e^{2t} \\ 0 & e^{2t} \end{pmatrix}\begin{pmatrix} 1 & 0 \\ -3 & 1 \end{pmatrix} = \begin{pmatrix} (-3t+1)\,e^{2t} & t\,e^{2t} \\ -9t\,e^{2t} & (3t+1)\,e^{2t} \end{pmatrix}.$$

On récupère la solution en fonction de la condition initiale :

$$\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = \begin{pmatrix} (-3t+1)\,e^{2t} & t\,e^{2t} \\ -9t\,e^{2t} & (3t+1)\,e^{2t} \end{pmatrix}\begin{pmatrix} x(0) \\ y(0) \end{pmatrix},$$

ou encore

$$\begin{cases} x(t) = [(-3\,x(0) + y(0))t + x(0)]\,e^{2t} \\ y(t) = [(-9\,x(0) + 3\,y(0))t + y(0)]\,e^{2t} \end{cases}$$

qui coïncide avec l'expression de la première partie avec $\alpha = x(0)$ et $\beta = -3\,x(0) + y(0)$).

(d). *Foyer.* Soit le système différentiel

$$\begin{cases} x'(t) = x(t) - 5\,y(t) \\ y'(t) = 4\,x(t) + 5\,y(t) \end{cases}. \tag{S}$$

Le système (S) est associé à la matrice $A = \begin{pmatrix} 1 & -5 \\ 4 & 5 \end{pmatrix}$, dont le polynôme caractéristique est $\Delta_A(\lambda) = \lambda^2 - 6\lambda + 25$. On obtient les valeurs propres complexes conjuguées $3+4i$ et $3-4i$ et donc la matrice A est diagonalisable sur $\mathbb{C}$. Ses sous-espaces propres sont

$$E_{3+4i} = \text{vect}\{(1-2i, -2)\} \quad \text{et} \quad E_{3-4i} = \text{vect}\{(1+2i, -2)\}.$$

En introduisant la matrice de passage $P = \begin{pmatrix} 1-2i & 1+2i \\ -2 & -2 \end{pmatrix}$ ainsi que la matrice diagonale $\begin{pmatrix} 3+4i & 0 \\ 0 & 3-4i \end{pmatrix}$, et en posant $\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = P\begin{pmatrix} u(t) \\ v(t) \end{pmatrix}$, on trouve le système différentiel complexe intermédiaire

$$\begin{cases} u'(t) = (3+4i)\,u(t) \\ v'(t) = (3-4i)\,v(t) \end{cases}.$$

La solution de ce dernier s'écrit

$$\begin{cases} u(t) = a\,e^{(3+4i)t} \\ v(t) = b\,e^{(3-4i)t} \end{cases},\ a,b\in\mathbb{C}.$$

La solution générale du système (S) sur $\mathbb{C}$ est alors

$$\begin{cases} x(t) = (1-2i)a\,e^{(3+4i)t} + (1+2i)b\,e^{(3-4i)t} \\ y(t) = -2a\,e^{(3+4i)t} - 2b\,e^{(3-4i)t} \end{cases},\ a,b\in\mathbb{C}.$$

Le système (S) étant initialement écrit sur $\mathbb{R}$, il est naturel de ne conserver que les solutions réelles. En écrivant

$$e^{(3+4i)t} = e^{3t}[\cos(4t) + i\sin(4t)] \quad \text{et} \quad e^{(3-4i)t} = e^{3t}[\cos(4t) - i\sin(4t)]$$

on trouve

$$\begin{cases} x(t) = [(1-2i)a + (1+2i)b]\,e^{3t}\cos(4t) + i[(1-2i)a - (1+2i)b]\,e^{3t}\sin(4t) \\ \qquad = [(a+b) - 2i(a-b)]\,e^{3t}\cos(4t) + [2(a+b) + i(a-b)]\,e^{3t}\sin(4t) \\ y(t) = -2(a+b)\,e^{3t}\cos(4t) - 2i(a-b)\,e^{3t}\sin(4t) \end{cases},\ a,b\in\mathbb{C}.$$

soit encore, en posant $a+b=\alpha$ et $i(a-b)=\beta$,

$$\begin{cases} x(t) = e^{3t}\,[(\alpha-2\beta)\,\cos(4t) + (2\alpha+\beta)\,\sin(4t)] \\ y(t) = -2\,e^{3t}\,[\alpha\,\cos(4t) + \beta\,\sin(4t)] \end{cases},\ \alpha,\beta\in\mathbb{C}.$$

Enfin, pour que la solution générale soit réelle, il est nécessaire que $x(0), y(0)\in\mathbb{R}$, ce qui impose les contraintes $\alpha+\beta, \alpha-\beta\in\mathbb{R}$, *i.e.* $\alpha,\beta\in\mathbb{R}$. On récupère donc la solution générale sur $\mathbb{R}$ en limitant le domaine de variation des constantes α et β à $\mathbb{R}$:

$$\begin{cases} x(t) = e^{3t}\,[(\alpha-2\beta)\,\cos(4t) + (2\alpha+\beta)\,\sin(4t)] \\ y(t) = -2\,e^{3t}\,[\alpha\,\cos(4t) + \beta\,\sin(4t)] \end{cases},\ \alpha,\beta\in\mathbb{R}.$$

Une autre approche consiste à introduire une base intermédiaire constituée de vecteurs réels. Pour cela, on isole les parties réelle et imaginaire du vecteur propre $(1-2i,-2)$ associé à la valeur propre $3+4i$:

$$(1-2i,-2) = (1,-2) + i(-2,0) = \vec{u} + i\vec{v} \quad \text{avec} \quad \vec{u} = (1,-2),\ \vec{v} = (-2,0).$$

On a, en utilisant l'endomorphisme $f\in\mathcal{L}(\mathbb{R}^2)$ associé à la matrice A,

$$f(\vec{u}+i\vec{v}) = (3+4i)\,(\vec{u}+i\vec{v}) = (3\vec{u}-4\vec{v}) + i(4\vec{u}+3\vec{v})$$

et, en séparant parties réelles et imaginaires,

$$f(\vec{u}) = 3\vec{u}-4\vec{v},\ f(\vec{v}) = 4\vec{u}+3\vec{v}.$$

La matrice de f dans la base $\{\vec{u},\vec{v}\}$ est alors $\begin{pmatrix} 3 & 4 \\ -4 & 3 \end{pmatrix}$. Cela conduit à la factorisation

$$A = \tilde{P}B\tilde{P}^{-1} \quad \text{avec} \quad \tilde{P} = \begin{pmatrix} 1 & -2 \\ -2 & 0 \end{pmatrix},\ B = \begin{pmatrix} 3 & 4 \\ -4 & 3 \end{pmatrix},\ \tilde{P}^{-1} = -\frac{1}{4}\begin{pmatrix} 0 & 2 \\ 2 & 1 \end{pmatrix}.$$

Comme
$$\exp(tB) = e^{3t}\begin{pmatrix} \cos(4t) & \sin(4t) \\ -\sin(4t) & \cos(4t)\end{pmatrix},$$
on trouve
$$\exp(tA) = \frac{1}{4}\,e^{3t}\begin{pmatrix} 4\,\cos(4t) - 2\,\sin(4t) & -5\,\sin(4t) \\ 4\,\sin(4t) & 4\cos(4t) + 2\,\sin(4t)\end{pmatrix}.$$
La solution du système (S) est directement donnée (sous forme réelle) par
$$\begin{pmatrix} x(t) \\ y(t)\end{pmatrix} = \exp(tA)\begin{pmatrix} x(0) \\ y(0)\end{pmatrix},$$
soit, explicitement,
$$\left\{\begin{array}{rcl} x(t) &=& \frac{1}{4}\,e^{3t}[4x(0)\,\cos(4t) - (2x(0) + 5y(0))\,\sin(4t)] \\ y(t) &=& \frac{1}{2}\,e^{3t}[2y(0)\,\cos(4t) + \ \ (2x(0) + y(0))\,\sin(4t)] \end{array}\right..$$
Afin d'analyser le comportement asymptotique des courbes intégrales, il est utile de se placer dans la base $\{\vec{u}, \vec{v}\}$. Dans cette dernière la fonction $Y(t) = \begin{pmatrix} u(t) \\ v(t)\end{pmatrix} = Q^{-1}X(t)$ est donnée par $Y(t) = \exp(tB)Y(0)$, soit
$$\left\{\begin{array}{rcl} u(t) &=& e^{3t}[u(0)\,\cos(4t) + v(0)\,\sin(4t)] \\ v(t) &=& e^{3t}[v(0)\,\cos(4t) - u(0)\,\sin(4t)] \end{array}\right..$$
En posant $u(0) = A\cos\varphi$ et $v(0) = A\sin\varphi$, il vient plus simplement
$$\left\{\begin{array}{rcl} u(t) &=& A\,e^{3t}\,\cos(4t - \varphi) \\ v(t) &=& -A\,e^{3t}\,\sin(4t - \varphi) \end{array}\right..$$
Sous cette forme, on voit que la courbe a un mouvement de rotation dans le sens des aiguilles d'une montre et que
$$u(t)^2 + v(t)^2 = [u(0)^2 + v(0)^2]\,e^{6t}.$$
Cette dernière relation montre que
$$\lim_{t\to-\infty}[u(t)^2 + v(t)^2] = 0 \quad \text{et} \quad \lim_{t\to+\infty}[u(t)^2 + v(t)^2] = +\infty.$$
La courbe a une allure de spirale s'enroulant autour de l'origine O lorsque $t \to -\infty$. Le point O est appelé *foyer* (voir Fig. 5.5).

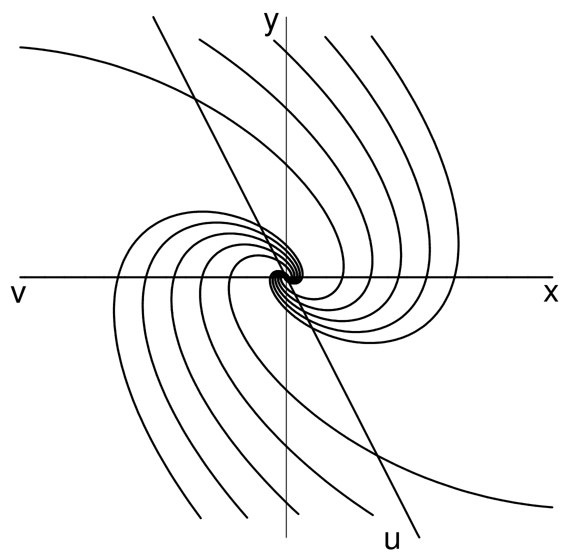

FIGURE 5.5 – Foyer

(e). *Autre foyer.* Le système différentiel

$$\begin{cases} x'(t) = -3\,x(t) + 5\,y(t) \\ y'(t) = -5\,x(t) + 3\,y(t) \end{cases} \tag{S}$$

est associé à la matrice $A = \begin{pmatrix} -3 & 5 \\ -5 & 3 \end{pmatrix}$. Son polynôme caractéristique est donné par $\Delta_A(\lambda) = \lambda^2 + 16$, et ses valeurs propres sont les nombres complexes conjugués $4i$ et $-4i$. La matrice A est diagonalisable sur $\mathbb{C}$. Ses sous-espaces propres sont

$$E_{4i} = \text{vect}\{(5, 3+4i)\} \quad \text{et} \quad E_{-4i} = \text{vect}\{(5, 3-4i)\}.$$

Par une méthode similaire à celle utilisée dans l'exemple précédent, on introduit la base intermédiaire $\{\vec{u}, \vec{v}\}$ construite à partir des parties réelles et imaginaires des vecteurs propres conjugués $(5, 3+4i)$ et $(5, 3-4i)$: $\vec{u} = (5, 3), \vec{v} = (0, 4)$. L'endomorphisme $f \in \mathcal{L}(\mathbb{R}^2)$ associé à la matrice A vérifie

$$f(\vec{u} + i\vec{v}) = 4i\,(\vec{u} + i\vec{v}) = -4\vec{v} + 4i\vec{u}$$

soit encore

$$f(\vec{u}) = -4\vec{v},\ f(\vec{v}) = 4\vec{u}.$$

La matrice de f relativement la base $\{\vec{u}, \vec{v}\}$ est alors $\begin{pmatrix} 0 & 4 \\ -4 & 0 \end{pmatrix}$. Dans cette même base, le système différentiel s'écrit, en posant $\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = P \begin{pmatrix} u(t) \\ v(t) \end{pmatrix}$ avec $P = \begin{pmatrix} 5 & 0 \\ 3 & 4 \end{pmatrix}$,

$$\begin{cases} u'(t) = 4\,v(t) \\ v'(t) = -4\,u(t) \end{cases}.$$

Une dérivation fournit des équations différentielles linéaires du second ordre ne comportant qu'une seule fonction inconnue :

$$\begin{cases} u''(t) = -16\,u(t) \\ v''(t) = -16\,v(t) \end{cases}$$

dont les solutions générales sont

$$\begin{cases} u(t) = \alpha\,\cos(4t) + \gamma\,\sin(4t) \\ v(t) = \beta\,\cos(4t) + \delta\,\sin(4t) \end{cases},\ \alpha, \beta, \gamma, \delta \in \mathbb{R}.$$

En reportant ces expressions dans le système du premier ordre, on trouve des relations entre les constantes $\alpha, \beta, \gamma, \delta$, à savoir $\gamma = -\beta$ et $\delta = \alpha$. D'où

$$\begin{cases} u(t) = \alpha\,\cos(4t) - \beta\,\sin(4t) \\ v(t) = \beta\,\cos(4t) + \alpha\,\sin(4t) \end{cases},\ \alpha, \beta \in \mathbb{R},$$

puis

$$\begin{cases} x(t) = 5\alpha\,\cos(4t) - 5\beta\,\sin(4t) \\ y(t) = (3\alpha + 4\beta)\,\cos(4t) + (4\alpha - 3\beta)\,\sin(4t) \end{cases},\ \alpha, \beta \in \mathbb{R}.$$

Il est clair que les fonctions u, v vérifient

$$u^2(t) + v^2(t) = \alpha^2 + \beta^2,$$

ce qui entraîne

$$5x^2(t) - 6x(t)y(t) + 5y^2(t) = c$$

pour une certaine constante c. Cela montre que les trajectoires sont portées par des ellipses concentriques de centre commun le point O appelé encore dans ce cas *foyer* (voir Fig. 5.6).

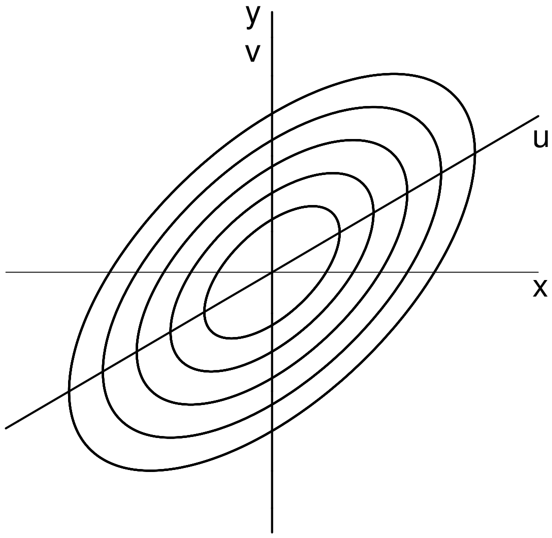

Figure 5.6 – Foyer

2. *En dimension 3 sur* $\mathbb{R}$.

(a). *Trois valeurs propres réelles distinctes.* Considérons le système

$$\left\{\begin{array}{rcl} x'(t) &=& x(t) + y(t) + 2\,z(t) \\ y'(t) &=& 2\,x(t) \\ z'(t) &=& x(t) - y(t) \end{array}\right. . \tag{S}$$

La matrice associée est $A = \begin{pmatrix} 1 & 1 & 2 \\ 2 & 0 & 0 \\ 1 & -1 & 0 \end{pmatrix}$. Son polynôme caractéristique est donné par $\Delta_A(\lambda) = \lambda^3 - \lambda^2 - 4\lambda + 4$. On trouve les valeurs propres 2, 1 et -2, puis les sous-espaces propres

$$E_2 = \mathrm{vect}\{(1,1,0)\}, \quad E_1 = \mathrm{vect}\{(1,2,-1)\}, \quad E_{-2} = \mathrm{vect}\{(1,-1,-1)\}.$$

Introduisons donc la matrice de passage $P = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & -1 \\ 0 & -1 & -1 \end{pmatrix}$ ainsi que la matrice diagonale $\begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}$ à laquelle est semblable A, et posons $\begin{pmatrix} x(t) \\ y(t) \\ z(t) \end{pmatrix} = P \begin{pmatrix} u(t) \\ v(t) \\ w(t) \end{pmatrix}$. Les nouvelles fonctions inconnues u, v, w satisfont au système différentiel

$$\left\{\begin{array}{rcr} u'(t) &=& 2\,u(t) \\ v'(t) &=& v(t) \\ w'(t) &=& -2\,w(t) \end{array}\right.$$

dont la solution est

$$\begin{cases} u(t) = \alpha\, e^{2t} \\ v(t) = \beta\, e^{t} \\ w(t) = \gamma\, e^{-2t} \end{cases}$$

où α, β, γ sont des constantes réelles. Finalement, la solution générale du système (S) est

$$\begin{cases} x(t) = \alpha\, e^{2t} + \beta\, e^{t} + \gamma\, e^{-2t} \\ y(t) = \alpha\, e^{2t} + 2\beta\, e^{t} - \gamma\, e^{-2t} \\ z(t) = \qquad\quad - \beta\, e^{t} - \gamma\, e^{-2t} \end{cases}, \ \alpha, \beta, \gamma \in \mathbb{R}.$$

En adoptant l'approche exponentielle, partant de la décomposition

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & -1 \\ 0 & -1 & -1 \end{pmatrix} \times \begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix} \times \frac{1}{3} \begin{pmatrix} 3 & 0 & 3 \\ -1 & 1 & -2 \\ 1 & -1 & -1 \end{pmatrix},$$

on aurait obtenu l'exponentielle

$$\begin{aligned} \exp(tA) &= \frac{1}{3} \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & -1 \\ 0 & -1 & -1 \end{pmatrix} \begin{pmatrix} e^{2t} & 0 & 0 \\ 0 & e^{t} & 0 \\ 0 & 0 & e^{-2t} \end{pmatrix} \begin{pmatrix} 3 & 0 & 3 \\ -1 & 1 & -2 \\ 1 & -1 & -1 \end{pmatrix} \\ &= \frac{1}{3} \begin{pmatrix} 3\, e^{2t} - e^{t} + e^{-2t} & e^{t} - e^{-2t} & 3\, e^{2t} - 2\, e^{t} - e^{-2t} \\ 3\, e^{2t} - 2\, e^{t} - e^{-2t} & 2\, e^{t} + e^{-2t} & 3\, e^{2t} - 4\, e^{t} + e^{-2t} \\ e^{t} - e^{-2t} & -e^{t} + e^{-2t} & 2\, e^{t} + e^{-2t} \end{pmatrix}. \end{aligned}$$

La solution du système (S) est directement donnée par

$$\begin{pmatrix} x(t) \\ y(t) \\ z(t) \end{pmatrix} = \frac{1}{3} \begin{pmatrix} 3\, e^{2t} - e^{t} + e^{-2t} & e^{t} - e^{-2t} & 3\, e^{2t} - 2\, e^{t} - e^{-2t} \\ 3\, e^{2t} - 2\, e^{t} - e^{-2t} & 2\, e^{t} + e^{-2t} & 3\, e^{2t} - 4\, e^{t} + e^{-2t} \\ e^{t} - e^{-2t} & -e^{t} + e^{-2t} & 2\, e^{t} + e^{-2t} \end{pmatrix} \begin{pmatrix} x(0) \\ y(0) \\ z(0) \end{pmatrix},$$

soit encore

$$\begin{cases} x(t) = (x(0) + z(0))\, e^{2t} + \dfrac{1}{3}(-x(0) + y(0) - 2\, z(0))\, e^{t} + \dfrac{1}{3}(x(0) - y(0) - z(0))\, e^{-2t} \\ y(t) = (x(0) + z(0))\, e^{2t} + \dfrac{1}{3}(-2\, x(0) + 2\, y(0) - 4\, z(0))\, e^{t} + \dfrac{1}{3}(-x(0) + y(0) + z(0))\, e^{-2t} \\ z(t) = \dfrac{1}{3}(x(0) - y(0) + 2\, z(0))\, e^{t} + \dfrac{1}{3}(-x(0) + y(0) + z(0))\, e^{-2t} \end{cases}$$

qui coïncide avec l'expression précédemment obtenue en choisissant pour coefficients $\alpha = x(0) + z(0)$, $\beta = \frac{1}{3}(-x(0) + y(0) - 2z(0))$ et $\gamma = \frac{1}{3}(x(0) - y(0) - z(0))$.

(b). *Deux valeurs propres réelles distinctes (cas diagonalisable).* Considérons le système

$$\begin{cases} x'(t) = x(t) + y(t) + 2\, z(t) \\ y'(t) = 2\, y(t) \\ z'(t) = x(t) - y(t) \end{cases}. \tag{S}$$

La matrice associée est $A = \begin{pmatrix} 1 & 1 & 2 \\ 0 & 2 & 0 \\ 1 & -1 & 0 \end{pmatrix}$. Son polynôme caractéristique est donné par $\Delta_A(\lambda) = \lambda^3 - 3\lambda^2 + 4$. On trouve la valeur propre double 2 et la valeur propre simple -1, puis les sous-espaces propres

$$E_2 = \text{vect}\{(2, 0, 1), (1, 1, 0)\} \quad \text{et} \quad E_{-1} = \text{vect}\{(1, 0, -1)\}.$$

Introduisons donc la matrice de passage $P = \begin{pmatrix} 2 & 1 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & -1 \end{pmatrix}$ ainsi que la matrice diagonale $\begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -1 \end{pmatrix}$ à laquelle est semblable A, et posons $\begin{pmatrix} x(t) \\ y(t) \\ z(t) \end{pmatrix} = P \begin{pmatrix} u(t) \\ v(t) \\ w(t) \end{pmatrix}$.
Les nouvelles fonctions inconnues u, v, w satisfont au système différentiel

$$\begin{cases} u'(t) &= \ 2\,u(t) \\ v'(t) &= \ 2\,v(t) \\ w'(t) &= -w(t) \end{cases}$$

dont la solution est

$$\begin{cases} u(t) &= \alpha\, e^{2t} \\ v(t) &= \beta\, e^{2t} \\ w(t) &= \gamma\, e^{-t} \end{cases}$$

où α, β, γ sont des constantes réelles. Finalement, la solution générale du système (S) est

$$\begin{cases} x(t) &= (2\alpha+\beta)\, e^{2t} + \gamma\, e^{-t} \\ y(t) &= \qquad \beta\, e^{2t} \\ z(t) &= \qquad \alpha\, e^{2t} - \gamma\, e^{-t} \end{cases}, \ \alpha, \beta, \gamma \in \mathbb{R}.$$

(c). *Deux valeurs propres réelles distinctes (cas non diagonalisable).* Considérons le système

$$\begin{cases} x'(t) &= x(t) \\ y'(t) &= x(t) + y(t)\,. \\ z'(t) &= \qquad y(t) \end{cases} \tag{S}$$

La matrice associée est $A = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 1 & 0 \end{pmatrix}$. Son polynôme caractéristique est donné par $\Delta_A(\lambda) = \lambda^3 - 2\lambda^2 + \lambda$. On trouve la valeur propre double 1 et la valeur propre simple 0, puis les sous-espaces propres

$$E_1 = \mathrm{vect}\{(0,1,1)\} \quad \text{et} \quad E_0 = \mathrm{vect}\{(0,0,1)\}.$$

À l'aide de la réduction de Jordan, en introduisant le vecteur intermédiaire $(1,1,0)$ entre les vecteurs propres $(0,1,1)$ et $(0,0,1)$ puis la matrice de passage correspondante $P = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}$, on trouve que $A = PTP^{-1}$ où T est la matrice triangulaire $\begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$. Posons alors $\begin{pmatrix} x(t) \\ y(t) \\ z(t) \end{pmatrix} = P \begin{pmatrix} u(t) \\ v(t) \\ w(t) \end{pmatrix}$. Les nouvelles fonctions inconnues u, v, w satisfont au système différentiel

$$\begin{cases} u'(t) &= u(t) + v(t) \\ v'(t) &= \qquad v(t) \\ w'(t) &= \ 0 \end{cases}$$

dont la solution est

$$\begin{cases} u(t) = \alpha\, e^t + \beta t\, e^t \\ v(t) = \beta\, e^t \\ w(t) = \gamma \end{cases}$$

où α, β, γ sont des constantes réelles. Finalement, la solution générale du système (S) est

$$\begin{cases} x(t) = \beta\, e^t \\ y(t) = ((\alpha+\beta)+\beta t)\, e^t \\ z(t) = (\alpha+\beta t)\, e^t + \gamma \end{cases}, \ \alpha, \beta, \gamma \in \mathbb{R}.$$

(d). *Une seule valeur propre (cas non diagonalisable).* Considérons le système

$$\begin{cases} x'(t) = x(t) + y(t) + 2\, z(t) \\ y'(t) = 2\, y(t) + z(t) \\ z'(t) = -x(t) + y(t) + 3\, z(t) \end{cases}. \tag{S}$$

La matrice associée est $A = \begin{pmatrix} 1 & 1 & 2 \\ 0 & 2 & 1 \\ -1 & 1 & 3 \end{pmatrix}$. Son polynôme caractéristique est donné par $\Delta_A(\lambda) = \lambda^3 - 6\lambda^2 + 12\lambda - 8$. On trouve la valeur propre triple 2, puis le sous-espace propre

$$E_2 = \text{vect}\{(1,1,0)\}.$$

À l'aide de la réduction de Jordan, en rajoutant les vecteurs intermédiaires $(1,0,1)$ et $(-1,0,0)$ au vecteur propre $(1,1,0)$ puis la matrice de passage correspondante $P = \begin{pmatrix} 1 & 1 & -1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$, on trouve que $A = PTP^{-1}$ où T est la matrice triangulaire $\begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}$. Posons alors $\begin{pmatrix} x(t) \\ y(t) \\ z(t) \end{pmatrix} = P \begin{pmatrix} u(t) \\ v(t) \\ w(t) \end{pmatrix}$. Les nouvelles fonctions inconnues u, v, w satisfont au système différentiel

$$\begin{cases} u'(t) = 2\, u(t) + v(t) \\ v'(t) = 2\, v(t) + w(t) \\ w'(t) = 2\, w(t) \end{cases}$$

dont la solution est

$$\begin{cases} u(t) = \alpha\, e^{2t} + \beta t\, e^{2t} + \frac{1}{2}\gamma t^2\, e^{2t} \\ v(t) = \beta\, e^{2t} + \gamma t\, e^{2t} \\ w(t) = \gamma\, e^{2t} \end{cases}$$

où α, β, γ sont des constantes réelles. Finalement, la solution générale du système (S) est

$$\begin{cases} x(t) = ((\alpha+\beta-\gamma)+(\beta+\gamma)t + \frac{1}{2}\gamma t^2)\, e^{2t} \\ y(t) = (\alpha+\beta t + \frac{1}{2}\gamma t^2)\, e^{2t} \\ z(t) = (\beta+\gamma t)\, e^{2t} \end{cases}, \ \alpha, \beta, \gamma \in \mathbb{R}.$$

En adoptant l'approche exponentielle, partant de la décomposition

$$\begin{aligned} A &= \begin{pmatrix} 1 & 1 & -1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \times \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix} \times \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -1 & 1 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 1 & -1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \times \left[2 \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \right] \times \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -1 & 1 & 1 \end{pmatrix} \end{aligned}$$

on aurait obtenu l'exponentielle

$$\begin{aligned} \exp(tA) &= \begin{pmatrix} 1 & 1 & -1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \times e^{2t} \begin{pmatrix} 1 & t & \frac{1}{2}t^2 \\ 0 & 1 & t \\ 0 & 0 & 1 \end{pmatrix} \times \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -1 & 1 & 1 \end{pmatrix} \\ &= e^{2t} \begin{pmatrix} -\frac{1}{2}t^2 - t + 1 & \frac{1}{2}t^2 + t & \frac{1}{2}t^2 + 2\,t \\ -\frac{1}{2}t^2 & \frac{1}{2}t^2 + 1 & \frac{1}{2}t^2 + t \\ -t & t & t+1 \end{pmatrix}. \end{aligned}$$

La solution du système (S) est directement donnée par

$$\begin{pmatrix} x(t) \\ y(t) \\ z(t) \end{pmatrix} = e^{2t} \begin{pmatrix} -\frac{1}{2}t^2 - t + 1 & \frac{1}{2}t^2 + t & \frac{1}{2}t^2 + 2\,t \\ -\frac{1}{2}t^2 & \frac{1}{2}t^2 + 1 & \frac{1}{2}t^2 + t \\ -t & t & t+1 \end{pmatrix} \begin{pmatrix} x(0) \\ y(0) \\ z(0) \end{pmatrix},$$

soit encore

$$\left\{ \begin{aligned} x(t) &= \left[\frac{1}{2}\,(-x(0) + y(0) + z(0))\,t^2 + (-x(0) + y(0) + 2\,z(0))\,t + x(0) \right] e^{2t} \\ y(t) &= \left[\frac{1}{2}\,(-x(0) + y(0) + z(0))\,t^2 + z(0))\,t + y(0) \right] e^{2t} \\ z(t) &= \big[(-x(0) + y(0) + z(0))\,t + z(0) \big] e^{2t} \end{aligned} \right.$$

qui coïncide avec l'expression précédemment obtenue en choisissant les coefficients $\alpha = y(0)$, $\beta = z(0)$ et $\gamma = -x(0) + y(0) + z(0)$.

(e). *Une valeur propre réelle et deux valeurs propres complexes non réelles conjuguées.* Considérons le système

$$\left\{ \begin{aligned} x'(t) &= x(t) + 2\,z(t) \\ y'(t) &= x(t) + y(t) \\ z'(t) &= y(t) \end{aligned} \right. . \qquad \text{(S)}$$

La matrice associée est $A = \begin{pmatrix} 1 & 0 & 2 \\ 1 & 1 & 0 \\ 0 & 1 & 0 \end{pmatrix}$. Son polynôme caractéristique est donné par $\Delta_A(\lambda) = \lambda^3 - 2\lambda^2 + \lambda - 2$. On trouve la valeur propre réelle 2 et les valeurs propres complexes conjuguées i et $-i$, puis les sous-espaces propres

$$E_2 = \mathrm{vect}\{(2,2,1)\}, \quad E_i = \mathrm{vect}\{(1+i, -i, -1)\}, \quad E_{-i} = \mathrm{vect}\{(1-i, i, -1)\}.$$

Introduisons donc la matrice de passage $P = \begin{pmatrix} 2 & 1+i & 1-i \\ 2 & -i & i \\ 1 & -1 & -1 \end{pmatrix}$ ainsi que la matrice diagonale $\begin{pmatrix} 2 & 0 & 0 \\ 0 & i & 0 \\ 0 & 0 & -i \end{pmatrix}$ à laquelle est semblable A, et posons $\begin{pmatrix} x(t) \\ y(t) \\ z(t) \end{pmatrix} = P \begin{pmatrix} u(t) \\ v(t) \\ w(t) \end{pmatrix}$. Les nouvelles fonctions inconnues à valeurs complexes u, v, w satisfont au système différentiel

$$\left\{ \begin{aligned} u'(t) &= 2\,u(t) \\ v'(t) &= i\,v(t) \\ w'(t) &= -i\,w(t) \end{aligned} \right.$$

dont la solution est

$$\begin{cases} u(t) &= \alpha\, e^{2t} \\ v(t) &= \beta\, e^{it} \\ w(t) &= \gamma\, e^{-it} \end{cases}$$

où α, β, γ sont des constantes complexes. Finalement, la solution générale du système (S) sur $\mathbb{C}$ est

$$\begin{cases} x(t) = 2\alpha\, e^{2t} + (1+i)\beta\, e^{it} + (1-i)\gamma\, e^{-it} \\ y(t) = 2\alpha\, e^{2t} - \quad i\beta\, e^{it} + \quad i\gamma\, e^{-it} \\ z(t) = \quad \alpha\, e^{2t} - \quad \beta\, e^{it} - \quad \gamma\, e^{-it} \end{cases}, \ \alpha, \beta, \gamma \in \mathbb{C}.$$

puis la solution générale sur $\mathbb{R}$, en remarquant que l'ensemble des combinaisons linéaires des vecteurs $t \mapsto ((1+i)\, e^{it}, -i\, e^{it}, -e^{it})$ et $t \mapsto ((1-i)\, e^{-it}, i\, e^{-it}, -e^{-it})$ coïncide avec celui des combinaisons linéaires des vecteurs $t \mapsto (\cos(t) - \sin(t), \sin(t), -\cos(t))$ et $t \mapsto (\cos(t) + \sin(t), -\cos(t), -\sin(t))$:

$$\begin{cases} x(t) = 2\alpha\, e^{2t} + (\beta+\gamma)\cos(t) + (-\beta+\gamma)\sin(t) \\ y(t) = 2\alpha\, e^{2t} - \quad \gamma\cos(t) + \quad \beta\sin(t) \\ z(t) = \quad \alpha\, e^{2t} - \quad \beta\cos(t) - \quad \gamma\sin(t) \end{cases}, \ \alpha, \beta, \gamma \in \mathbb{R}.$$

En adoptant l'approche exponentielle, partant de la décomposition

$$A = \begin{pmatrix} 2 & 1+i & 1-i \\ 2 & -i & i \\ 1 & -1 & -1 \end{pmatrix} \times \begin{pmatrix} 2 & 0 & 0 \\ 0 & i & 0 \\ 0 & 0 & -i \end{pmatrix} \times \frac{1}{5} \begin{pmatrix} 1 & 1 & 1 \\ \frac{1}{2} - i & \frac{1}{2} + \frac{3}{2}i & -2-i \\ \frac{1}{2} + i & \frac{1}{2} - \frac{3}{2}i & -2+i \end{pmatrix}$$

ou encore, en introduisant les parties réelle et imaginaire de la matrice de passage, la décomposition sous forme réelle

$$\begin{aligned} A &= \frac{1}{5} \left[\begin{pmatrix} 2 & 1 & 1 \\ 2 & 0 & -1 \\ 1 & -1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & i & -i \end{pmatrix} \right] \times \begin{pmatrix} 2 & 0 & 0 \\ 0 & i & 0 \\ 0 & 0 & -i \end{pmatrix} \times \left[\begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2} & -\frac{i}{2} \\ 0 & \frac{1}{2} & \frac{i}{2} \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & -4 \\ 2 & -3 & 2 \end{pmatrix} \right] \\ &= \frac{1}{5} \begin{pmatrix} 2 & 1 & 1 \\ 2 & 0 & -1 \\ 1 & -1 & 0 \end{pmatrix} \times \left[\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & i & -i \end{pmatrix} \begin{pmatrix} 2 & 0 & 0 \\ 0 & i & 0 \\ 0 & 0 & -i \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2} & -\frac{i}{2} \\ 0 & \frac{1}{2} & \frac{i}{2} \end{pmatrix} \right] \times \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & -4 \\ 2 & -3 & 2 \end{pmatrix} \\ &= \frac{1}{5} \begin{pmatrix} 2 & 1 & 1 \\ 2 & 0 & -1 \\ 1 & -1 & 0 \end{pmatrix} \begin{pmatrix} 2 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & -4 \\ 2 & -3 & 2 \end{pmatrix}, \end{aligned}$$

on aurait obtenu l'exponentielle

$$\begin{aligned} \exp(tA) &= \frac{1}{5} \begin{pmatrix} 2 & 1 & 1 \\ 2 & 0 & -1 \\ 1 & -1 & 0 \end{pmatrix} \begin{pmatrix} e^{2t} & 0 & 0 \\ 0 & \cos(t) & \sin(t) \\ 0 & -\sin(t) & \cos(t) \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & -4 \\ 2 & -3 & 2 \end{pmatrix} \\ &= \frac{1}{5} \begin{pmatrix} 2\, e^{2t} + 3\cos(t) + \sin(t) & 2\, e^{2t} - 2\cos(t) - 4\sin(t) & 2\, e^{2t} - 2\cos(t) + 6\sin(t) \\ 2\, e^{2t} - 2\cos(t) + \sin(t) & 2\, e^{2t} + 3\cos(t) + \sin(t) & 2\, e^{2t} - 2\cos(t) - 4\sin(t) \\ e^{2t} - \cos(t) - 2\sin(t) & e^{2t} - \cos(t) + 3\sin(t) & e^{2t} + 4\cos(t) - 2\sin(t) \end{pmatrix}. \end{aligned}$$

La solution du système (S) est directement donnée par $\begin{pmatrix} x(t) \\ y(t) \\ z(t) \end{pmatrix} = \exp(tA) \begin{pmatrix} x(0) \\ y(0) \\ z(0) \end{pmatrix}$, soit encore

$$\begin{cases} x(t) = \\ \frac{2}{5}(x(0)+y(0)+z(0))\, e^{2t} + \frac{1}{5}(3\,x(0) - 2\,y(0) - 2\,z(0))\cos(t) + \frac{1}{5}(x(0) - 4\,y(0) + 6\,z(0))\sin(t) \\ y(t) = \\ \frac{2}{5}(x(0)+y(0)+z(0))\, e^{2t} - \frac{1}{5}(2\,x(0) - 3\,y(0) + 2\,z(0))\cos(t) + \frac{1}{5}(x(0) + y(0) - 4\,z(0))\sin(t) \\ z(t) = \\ \frac{1}{5}(x(0)+y(0)+z(0))\, e^{2t} - \frac{1}{5}(x(0) + y(0) - 4\,z(0))\cos(t) - \frac{1}{5}(2\,x(0) - 3\,y(0) + 2\,z(0))\sin(t) \end{cases}$$

qui coïncide avec l'expression précédemment obtenue en choisissant $\alpha = \dfrac{1}{5}\,(x(0) + y(0) + z(0))$, $\beta = \dfrac{1}{5}\,(x(0) + y(0) - 4\,z(0))$ et $\gamma = \dfrac{1}{5}\,(2\,x(0) - 3\,y(0) + 2\,z(0))$.

4.3 Réduction des coniques

Considérons, dans un repère orthonormé direct $(O, \vec{\imath}, \vec{\jmath})$ du plan, la conique $(\mathcal{C})$ d'équation

$$ax^2 + 2bxy + cy^2 + 2dx + 2ey + f = 0, \tag{E}$$

où a, b, c, d, e, f sont des réels, $a, b, c,$ étant non tous nuls. Posons $q(x, y) = ax^2 + 2bxy + cy^2$ et réécrivons $q(x, y)$ comme suit :

$$q(x, y) = x(ax + by) + y(bx + cy) = \begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} ax + by \\ bx + cy \end{pmatrix} = \begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} a & b \\ b & c \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

On a commis un léger abus d'écriture dans les égalité précédentes puisque les deux premiers membres sont des réels alors que les deux derniers sont des matrices de type 1×1 (comportant un unique terme). On identifiera donc les deux objets : $(x) = x$. Introduisons $X = \begin{pmatrix} x \\ y \end{pmatrix}$ et $A = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$. Ceci prend la forme matricielle suivante :

$$q(x, y) = {}^tXAX.$$

La matrice A étant réelle et symétrique, elle est diagonalisable via une matrice de passage orthogonale :

$$A = PD\,{}^tP \quad \text{avec} \quad D = \begin{pmatrix} \lambda & 0 \\ 0 & \mu \end{pmatrix} \text{ et } {}^tP = P^{-1}.$$

On peut même choisir la matrice P telle que $\det(P) = 1$. C'est la matrice de passage de la base orthonormée directe $\mathcal{B} = (\vec{\imath}, \vec{\jmath})$ vers une autre base orthonormée directe $\mathcal{B}' = (\vec{\imath}', \vec{\jmath}')$ formée de vecteurs propres pour A. On peut également la percevoir comme la matrice d'une rotation vectorielle du plan vectoriel : $P = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$. Effectuons le changement de coordonnées associé : $X = PY$ où $Y = \begin{pmatrix} x' \\ y' \end{pmatrix}$. On a avec ces nouvelles coordonnées

$$q(x, y) = {}^tXAX = {}^t(PY)(PD\,{}^tP)(PY) = {}^tY({}^tPP)D({}^tPP)Y = {}^tYDY,$$

soit

$$q(x, y) = \lambda x'^2 + \mu y'^2.$$

La conique $(\mathcal{C})$ admet pour équation relativement au repère $(O, \vec{\imath}\,', \vec{\jmath}\,')$

$$\lambda x'^2 + \mu y'^2 + 2d'x' + 2e'y' + f = 0.$$

Les coefficients d', e' et f peuvent s'obtenir à partir de la relation

$$\begin{pmatrix} d' & e' \end{pmatrix} = \begin{pmatrix} d & e \end{pmatrix} P$$

puisque

$$dx + ey = \begin{pmatrix} d & e \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} d & e \end{pmatrix} P \begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} d' & e' \end{pmatrix} \begin{pmatrix} x' \\ y' \end{pmatrix} = d'x' + e'y'.$$

— *Cas où $\lambda,\mu\neq 0$:* en effectuant enfin la translation $x''=x'+\dfrac{d'}{\lambda}, y''=y'+\dfrac{e'}{\mu}$, l'équation prend, dans le repère orthonormé direct $(O',\vec{i}\,',\vec{j}\,')$ de nouvelle origine le point O' de coordonnées $\Big(-\dfrac{d'}{\lambda},-\dfrac{e'}{\mu}\Big)$ relativement au repère $(O,\vec{i}\,',\vec{j}\,')$, la forme simple (où $\alpha=\dfrac{d'^2}{\lambda}+\dfrac{e'^2}{\mu}-f$)

$$\lambda x''^2+\mu y''^2=\alpha.$$

- *Lorsque λ et μ sont de mêmes signes :*
  - $*$ si α a le même signe que λ et μ, alors, en posant $a=\sqrt{\alpha/\lambda}$ et $b=\sqrt{\alpha/\mu}$, l'équation (E) prend la forme réduite

$$\boxed{\frac{x''^2}{a^2}+\frac{y''^2}{b^2}=1.}$$

La conique $\mathcal{C}$ est l'ellipse de centre O', d'axes dirigés et orientés par les vecteurs $\vec{i}\,'$ et $\vec{j}\,'$ (ce sont les *directions propres*) et de demi-axes a et b (voir Fig. 5.7) ;

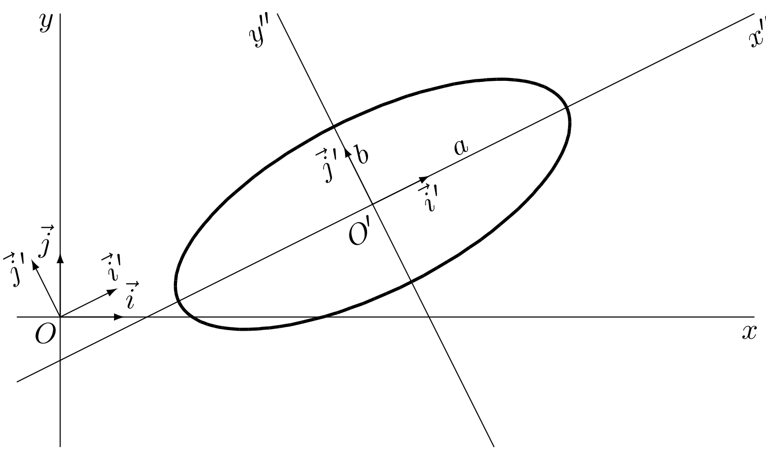

FIGURE 5.7 – Ellipse

  - $*$ si $\alpha=0$, on a

$$\frac{x''^2}{a^2}+\frac{y''^2}{b^2}=0\Longrightarrow x''=y''=0$$

et la courbe est réduite au point O' ;
  - $*$ si α a le signe opposé de λ et μ, la courbe est vide.
- *Lorsque λ et μ sont de signes contraires :*
  - $*$ si α a le même signe que λ, alors, en posant $a=\sqrt{\alpha/\lambda}$ et $b=\sqrt{-\alpha/\mu}$, l'équation (E) prend la forme réduite

$$\boxed{\frac{x''^2}{a^2}-\frac{y''^2}{b^2}=1.}$$

La conique $\mathcal{C}$ est une hyperbole de centre O' et d'axes dirigés et orientés par les vecteurs $\vec{i}\,'$ et $\vec{j}\,'$, $O'x''$ étant l'axe transverse de $\mathcal{C}$ (voir Fig. 5.8) ;

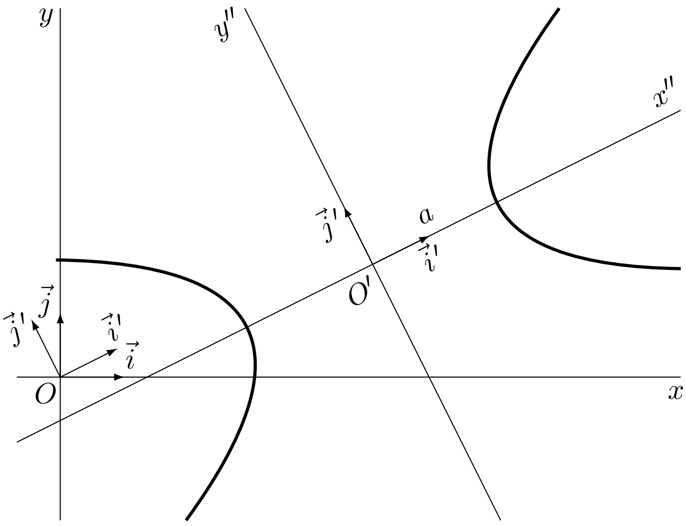

FIGURE 5.8 – Hyperbole

∗ si α a le même signe que μ, alors, en posant $a=\sqrt{-\alpha/\lambda}$ et $b=\sqrt{\alpha/\mu}$, l'équation (E) prend la forme réduite

$$\boxed{\frac{x''^2}{a^2}-\frac{y''^2}{b^2}=-1.}$$

La conique $\mathcal{C}$ est une hyperbole de mêmes caractéristiques que la précédente excepté l'axe transverse qui est cette fois l'axe $O'y''$;

∗ si $\alpha=0$, on a

$$\frac{x''^2}{a^2}-\frac{y''^2}{b^2}=0.$$

La courbe est la réunion des deux droites d'équations $y''=\pm\dfrac{a}{b}\,x''$.

— *Cas où l'un des deux nombres λ et μ est nul :* supposons par exemple $\lambda=0$ et $\mu\neq 0$.

- Si $d'\neq 0$, en effectuant la translation $x''=x'+\dfrac{f}{2d'}-\dfrac{e'^2}{2\mu d'}, y''=y'+\dfrac{e'}{\mu}$, l'équation (E) prend, dans le repère orthonormé direct $(O',\vec{i}\,',\vec{j}\,')$ de nouvelle origine le point O' de coordonnées $\left(\dfrac{e'^2-\mu f}{2\mu d'},-\dfrac{e'}{\mu}\right)$ relativement au repère $(O,\vec{i}\,',\vec{j}\,')$, la forme réduite (où $p=-2d'$)

$$\boxed{y''^2=2px''.}$$

La conique $\mathcal{C}$ est la parabole d'axe $O'x''$, de sommet O' et de paramètre p (voir Fig. 5.9).

- Si $d'=0$, on a $\mu y'^2+2e'y'+f=0$, ce qui donne zéro, une ou deux valeurs pour y' (sans condition pour x') et la courbe est donc soit vide, soit une droite de direction $\vec{i}\,'$, soit la réunion de deux droites de direction $\vec{i}\,'$.

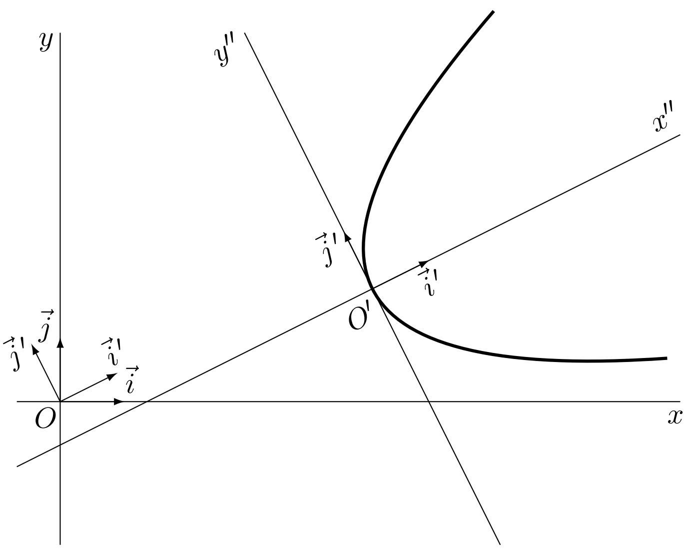

Figure 5.9 – Parabole

5 Exercices sur la diagonalisation

Exercice 1 Les matrices ci-dessous sont-elles diagonalisables sur $\mathbb{R}$? Dans l'affirmative, donner une base de vecteurs propres.

$$\text{a)} \begin{pmatrix} 9 & 0 & 0 \\ -5 & 4 & 0 \\ -8 & 0 & 1 \end{pmatrix} \quad \text{b)} \begin{pmatrix} 2 & 0 & 0 \\ -3 & -1 & 1 \\ 3 & 4 & -1 \end{pmatrix} \quad \text{c)} \begin{pmatrix} -3 & 1 & 0 \\ 0 & 1 & 0 \\ 4 & -8 & 2 \end{pmatrix}$$

$$\text{d)} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 2 \end{pmatrix} \quad \text{e)} \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & -3 \\ 0 & 1 & 3 \end{pmatrix}$$

Exercice 2 Soit a et b des nombres réels non nuls. Diagonaliser la matrice $A \in \mathcal{M}_n(\mathbb{R})$ donnée par

$$A = \begin{pmatrix} a & b & \dots & b & b \\ b & a & \ddots & & b \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ b & & \ddots & a & b \\ b & b & \dots & b & a \end{pmatrix}.$$

On rappelle que $\det(A) = (a-b)^{n-1}(a+(n-1)b)$ (voir la fiche sur les déterminants).

Exercice 3 (Une matrice de permutation) Diagonaliser la matrice $A \in \mathcal{M}_n(\mathbb{C})$ donnée par

$$A = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & & & \ddots & 1 \\ 1 & 0 & \dots & \dots & 0 \end{pmatrix}.$$

On pourra introduire la notation $\omega = e^{i\,\frac{2\pi}{n}}$.

Exercice 4 Diagonaliser la matrice

$$A = \begin{pmatrix} \frac{14}{5} & -\frac{2}{5} \\ -\frac{2}{5} & \frac{11}{5} \end{pmatrix}.$$

Applications.

1. Déterminer les suites $(u_n)_{n\in\mathbb{N}}$ et $(v_n)_{n\in\mathbb{N}}$ telles que

$$\begin{cases} u_{n+1} = \frac{14}{5}\,u_n - \frac{2}{5}\,v_n \\ v_{n+1} = -\frac{2}{5}\,u_n + \frac{11}{5}\,v_n \end{cases}$$

en fonction de leurs premiers termes u_0 et v_0.

2. Déterminer les fonctions x et y telles que

$$\left\{\begin{array}{rcrcr} x'(t) & = & \dfrac{14}{5}\,x(t) & - & \dfrac{2}{5}\,y(t) \\ y'(t) & = & -\dfrac{2}{5}\,x(t) & + & \dfrac{11}{5}\,y(t) \end{array}\right.$$

en fonction de $x(0)$ et $y(0)$. Tracer les courbes intégrales dans un repère orthonormé.

3. Déterminer la nature de la conique d'equation $14x^2 - 4xy + 11y^2 = 5$, puis la représenter dans un repère orthonormé.

Reprendre l'exercice avec la matrice

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

Exercice 5 On se place dans le $\mathbb{R}$-espace vectoriel $E = \mathbb{R}^3$ rapporté à une base $\mathcal{B} = \{\vec{e}_1, \vec{e}_2, \vec{e}_3\}$. Soit f l'endomorphisme de E dont la matrice dans la base $\mathcal{B}$ est donnée par $A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -2 & 3 & 0 \end{pmatrix}$.

1. L'endomorphisme f est-il diagonalisable ?
2. Déterminer une base dans laquelle f admet pour matrice $\begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}$.
3. Applications.

(a). Déterminer les suites $(u_n)_{n\in\mathbb{N}}$, $(v_n)_{n\in\mathbb{N}}$ et $(w_n)_{n\in\mathbb{N}}$ telles que

$$\left\{\begin{array}{rcl} u_{n+1} & = & v_n \\ v_{n+1} & = & w_n \\ w_{n+1} & = & -2u_n + 3v_n \end{array}\right.$$

en fonction de leurs premiers termes u_0, v_0 et w_0.

(b). Déterminer la solution générale du système différentiel

$$\left\{\begin{array}{rcl} x'(t) & = & y(t) \\ y'(t) & = & z(t) \\ z'(t) & = & -2x(t) + 3y(t) \end{array}\right.$$

Chapitre 6

Algèbre bilinéaire réelle

Alors que l'algèbre linéaire concerne des vecteurs d'un espace vectoriel « pris un à un », l'*algèbre bilinéaire* porte sur les couples de vecteurs (vecteurs « pris deux à deux »). Les *applications bilinéaires* agissent sur des couples de vecteurs et possèdent des propriétés de linéarité classiques, mais ce relativement à chacune de leurs deux variables vectorielles. Un cas déjà rencontré dans ce cours est celui du déterminant d'ordre 2, fournissant un exemple de forme bilinéaire alternée sur un plan vectoriel. Ce chapitre est consacré à l'étude des formes bilinéaires symétriques et des formes quadratiques sur un $\mathbb{R}$-espace vectoriel de dimension finie quelconque. La notion de *produit scalaire*, forme bilinéaire symétrique particulière, permet de définir une *géométrie euclidienne* et d'y formaliser les notions familières de *distance* et d'*angle*.

Dans tout ce chapitre, E désigne un $\mathbb{R}$-espace vectoriel.

1 Formes bilinéaires, formes quadratiques

1.1 Définitions, propriétés

Définition 1.1 *1. Une forme bilinéaire sur E est une application* $\varphi : \begin{array}{ccc} E \times E & \longrightarrow & \mathbb{R} \\ (\vec{u}, \vec{v}) & \longmapsto & \varphi(\vec{u}, \vec{v}) \end{array}$ *telle que pour tous scalaires $\alpha, \beta \in \mathbb{R}$ et tous vecteurs $\vec{u}, \vec{u}_1, \vec{u}_2, \vec{v}, \vec{v}_1, \vec{v}_2 \in E$,*

$$\begin{array}{rcl} \varphi(\alpha\vec{u}_1 + \beta\vec{u}_2, \vec{v}) & = & \alpha\varphi(\vec{u}_1, \vec{v}) + \beta\varphi(\vec{u}_2, \vec{v}), \\ \varphi(\vec{u}, \alpha\vec{v}_1 + \beta\vec{v}_2) & = & \alpha\varphi(\vec{u}, \vec{v}_1) + \beta\varphi(\vec{u}, \vec{v}_2). \end{array}$$

2. *Soit φ une forme bilinéaire sur E. La forme quadratique associée à φ est l'application* $q_\varphi : \begin{array}{ccc} E & \longrightarrow & \mathbb{R} \\ \vec{u} & \longmapsto & \varphi(\vec{u}, \vec{u}) \end{array}$. *Plus généralement, une forme quadratique sur E (sans précision supplémentaire) est la forme quadratique associée à une forme bilinéaire sur E.*
3. *Une forme bilinéaire φ sur E est symétrique lorsque*

$$\forall \vec{u}, \vec{v} \in E,\ \varphi(\vec{u}, \vec{v}) = \varphi(\vec{v}, \vec{u}).$$

Exemple.

Considérons le $\mathbb{R}$ espace vectoriel $E = \mathbb{K}^2$ rapporté à sa base canonique $\mathcal{B} = \{\vec{i}, \vec{j}\}$.

— Recherchons les formes bilinéaires sur E. Si φ est une forme bilinéaire sur E, on a nécessairement, pour $\vec{u} = (x,y) = x\vec{i} + y\vec{j}$ et $\vec{v} = (x',y') = x'\vec{i} + y'\vec{j}$,

$$\begin{aligned} \varphi(\vec{u},\vec{v}) &= \varphi(x\vec{i} + y\vec{j}, \vec{v}) \\ &= x\varphi(\vec{i}, x'\vec{i} + y'\vec{j}) + y\varphi(\vec{j}, x'\vec{i} + y'\vec{j}) \\ &= xx'\varphi(\vec{i},\vec{i}) + xy'\varphi(\vec{i},\vec{j}) + x'y\varphi(\vec{j},\vec{i}) + yy'\varphi(\vec{j},\vec{j}) \\ &= axx' + bxy' + cx'y + dyy' \end{aligned}$$

avec $a = \varphi(\vec{i},\vec{i}), b = \varphi(\vec{i},\vec{j}), c = \varphi(\vec{j},\vec{i}), d = \varphi(\vec{j},\vec{j})$. Inversement, on vérifie aisément qu'une application donnée par $\varphi((x,y),(x',y')) = axx' + bxy' + cx'y + dyy'$ pour tous $(x,y),(x',y') \in E$ est effectivement bilinéaire.

— On voit ainsi qu'une forme bilinéaire sur E est entièrement caractérisée par la donnée des scalaires $\varphi(\vec{i},\vec{i}), \varphi(\vec{i},\vec{j}), \varphi(\vec{j},\vec{i}), \varphi(\vec{j},\vec{j})$. Notons par ailleurs que la quantité $axx' + bxy' + cx'y + dyy'$ peut s'écrire sous la forme matricielle suivante :

$$\begin{aligned} axx' + bxy' + cx'y + dyy' &= x(ax' + by') + y(cx' + dy') \\ &= \begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} ax' + by' \\ cx' + dy' \end{pmatrix} = \begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x' \\ y' \end{pmatrix}. \end{aligned}$$

Dans ces égalités, on a commis un léger abus d'écriture en identifiant une matrice contenant un seul élément avec cet élément : $(x) = x$, et l'on a également utilisé la relation élémentaire $xx' + yy' = \begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} x' \\ y' \end{pmatrix}$. En posant $X = \operatorname{Mat}(\vec{u};\mathcal{B}) = \begin{pmatrix} x \\ y \end{pmatrix}$, $Y = \operatorname{Mat}(\vec{v};\mathcal{B}) = \begin{pmatrix} x' \\ y' \end{pmatrix}$ et $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, ceci s'écrit simplement

$$\varphi(\vec{u},\vec{v}) = {}^tXAY.$$

La matrice carrée A est la *matrice de la forme bilinéaire φ relativement à la base $\mathcal{B}$* :

$$A = \operatorname{Mat}(\varphi;\mathcal{B}) = \begin{pmatrix} \varphi(\vec{i},\vec{i}) & \varphi(\vec{i},\vec{j}) \\ \varphi(\vec{j},\vec{i}) & \varphi(\vec{j},\vec{j}) \end{pmatrix}.$$

— La forme bilinéaire φ est symétrique si et seulement si

$$\forall (x,y),(x',y') \in E,\ axx' + bxy' + cx'y + dyy' = ax'x + bx'y + cxy' + dy'y,$$

soit encore si et seulement si $b = c$. En termes matriciels, φ est symétrique si et seulement si A est une matrice symétrique.

— La forme quadratique associée à φ est donnée par

$$q_\varphi(\vec{u}) = \varphi(\vec{u},\vec{u}) = ax^2 + (b+c)xy + dy^2.$$

Par souci de simplification d'écriture, on note $q_\varphi(\vec{u}) = q_\varphi((x,y))$ plutôt $q_\varphi(x,y)$. Elle est de la forme $q_\varphi(x,y) = \alpha x^2 + 2\beta xy + \gamma y^2$. Réciproquement, soit q une application donnée par $q(x,y) = \alpha x^2 + 2\beta xy + \gamma y^2$ pour tout $(x,y) \in E$, avec $\alpha, \beta, \gamma \in \mathbb{R}$. C'est la forme générale d'une forme quadratique sur E sans avoir défini de forme bilinéaire au préalable. Il est alors naturel de se poser la question suivante : l'application q peut-elle être considérée comme la forme quadratique associée à une forme bilinéaire ? Si φ est une telle forme bilinéaire, la question s'énonce comme suit : a-t-on $q_\varphi = q$? Soit $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ la matrice de φ relativement à la base canonique de E. On a alors les équivalences

$$\begin{aligned} q_\varphi = q &\iff \forall (x,y) \in E,\ ax^2 + (b+c)xy + dy^2 = \alpha x^2 + 2\beta xy + \gamma y^2 \\ &\iff a = \alpha,\ b + c = 2\beta,\ c = \gamma. \end{aligned}$$

On voit qu'il existe une infinité de formes bilinéaires répondant à la question posée. Si l'on restreint à présent la recherche aux formes bilinéaires symétriques, on a avec les notations précédentes $b = c$ et dans ce cas

$$q_\varphi = q \iff \forall (x,y) \in E,\ a = \alpha,\ b = c = \beta,\ d = \gamma,$$

Cette fois, il existe une unique forme bilinéaire symétrique $\tilde{\varphi}_q$ dont la forme quadratique associée est q ; elle est appelée *forme polaire de* q, elle est donnée par

$$\forall (x,y), (x',y') \in E,\ \tilde{\varphi}_q((x,y),(x',y')) = \alpha xx' + \beta(xy' + x'y) + \gamma yy'$$

et sa matrice est $A = \begin{pmatrix} \alpha & \beta \\ \beta & \gamma \end{pmatrix}$. C'est indifféremment la matrice de $\tilde{\varphi}_q$ ou q : $A = \mathrm{Mat}(\tilde{\varphi}_q; \mathcal{B}) = \mathrm{Mat}(q; \mathcal{B})$. On a les deux relations fondamentales

$$\tilde{\varphi}_q((x,y),(x',y')) = {}^tXAY \quad \text{et} \quad q(x,y) = {}^tXAX.$$

1.2 Matrice d'une forme bilinéaire et d'une forme quadratique

L'exemple de la section précédente correspondant au cas de la dimension 2 peut s'étendre facilement au cas d'un espace vectoriel de dimension finie quelconque. Supposons l'espace vectoriel E de dimension n. Soit $\mathcal{B} = \{\vec{e}_1, \ldots, \vec{e}_n\}$ une base de E et φ une forme bilinéaire sur E. Introduisons deux vecteurs génériques $\vec{u}, \vec{v} \in E$ exprimés relativement à la base $\mathcal{B}$ selon $\vec{u} = \sum_{i=1}^{n} x_i \vec{e}_i$ et $\vec{v} = \sum_{i=1}^{n} y_i \vec{e}_i$. On a

$$\begin{aligned} \varphi(\vec{u}, \vec{v}) = \varphi\left(\sum_{i=1}^{n} x_i \vec{e}_i, \vec{v}\right) &= \sum_{i=1}^{n} x_i \varphi\left(\vec{e}_i, \sum_{j=1}^{n} y_j \vec{e}_j\right) \\ &= \sum_{i=1}^{n} x_i \left(\sum_{j=1}^{n} y_j \varphi(\vec{e}_i, \vec{e}_j)\right) = \sum_{1 \leqslant i,j \leqslant n} x_i y_j \varphi(\vec{e}_i, \vec{e}_j). \end{aligned}$$

On a donc le résultat ci-dessous.

Théorème 1.2 *La forme bilinéaire φ est entièrement caractérisée par les scalaires $\varphi(\vec{e}_i, \vec{e}_j)$, $1 \leqslant i, j \leqslant n$.*

Posons alors $a_{ij} = \varphi(\vec{e}_i, \vec{e}_j)$, puis $A = (a_{ij})_{1\leqslant i,j\leqslant n}$. La matrice A est la *matrice de la forme bilinéaire φ relativement à la base $\mathcal{B}$* ; on note $A = \mathrm{Mat}(\varphi; \mathcal{B})$. L'*écriture analytique de φ relativement à la base $\mathcal{B}$* est

$$\boxed{\varphi(\vec{u}, \vec{v}) = \sum_{1\leqslant i,j\leqslant n} a_{ij} x_i y_j.}$$

En introduisant les matrices-colonnes $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ et $Y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$ des vecteurs $\vec{u}, \vec{v}$, on tire une expression matricielle pour $\varphi(\vec{u}, \vec{v})$:

$$\varphi(\vec{u}, \vec{v}) = \sum_{1\leqslant i,j\leqslant n} a_{ij} x_i y_j = \sum_{i=1}^{n} x_i (a_{i1} y_1 + \cdots + a_{in} y_n) = \begin{pmatrix} x_1 & \ldots & x_n \end{pmatrix} \begin{pmatrix} a_{11} y_1 + \cdots + a_{1n} y_n \\ \vdots \\ a_{n1} y_1 + \cdots + a_{nn} y_n \end{pmatrix}.$$

Remarquant enfin que

$$\begin{pmatrix} a_{11} y_1 + \cdots + a_{1n} y_n \\ \vdots \\ a_{n1} y_1 + \cdots + a_{nn} y_n \end{pmatrix} = AY,$$

on trouve

$$\boxed{\varphi(\vec{u}, \vec{v}) = {}^t\!XAY.}$$

Proposition 1.3 *Une forme bilinéaire est symétrique si et seulement si sa matrice relativement à une base de E est symétrique.*

Démonstration.

Avec les notations précédentes, on a $\varphi(\vec{u}, \vec{v}) = {}^t\!XAY$ et $\varphi(\vec{v}, \vec{u}) = {}^t Y A X$. Observons que la matrice ${}^t Y A X$ se réduit à un scalaire et qu'elle est par conséquent symétrique : ${}^t Y A X = {}^t\!X {}^t\!A Y$. On a alors les équivalences

$$\begin{aligned} \varphi \text{ symétrique } &\iff \forall X, Y \in \mathcal{M}_{n\times 1}(\mathbb{R}),\ {}^t\!XAY = {}^t Y A X \\ &\iff \forall X, Y \in \mathcal{M}_{n\times 1}(\mathbb{R}),\ {}^t\!XAY = {}^t\!X {}^t\!A Y \end{aligned}$$

soit encore

$$\varphi \text{ symétrique } \iff {}^t\!A = A.$$

□

Soit φ une forme bilinéaire de matrice A relativement à la base $\mathcal{B}$. La forme quadratique associée q_φ est alors caractérisée par les relations analytique et matricielle suivantes :

$$\boxed{q_\varphi(\vec{u}) = \sum_{1\leqslant i,j\leqslant n} a_{ij} x_i x_j = {}^t\!XAX.}$$

En isolant les carrés x_i^2 et en rassemblant les produits $x_i x_j$ et $x_j x_i$ dans la somme $\displaystyle\sum_{1\leqslant i,j\leqslant n} a_{ij} x_i x_j$, on peut réécrire

$$q_\varphi(\vec{u}) = \sum_{1\leqslant i,j\leqslant n} a_{ij} x_i x_j = \sum_{i=1}^{n} a_{ii} x_i^2 + \sum_{1\leqslant i<j\leqslant n} (a_{ij} + a_{ji}) x_i x_j.$$

L'expression d'une forme quadratique générale relativement à la base $\mathcal{B}$ est donc

$$q(\vec{u}) = \sum_{i=1}^{n} \alpha_{ii} x_i^2 + 2 \sum_{1 \leqslant i < j \leqslant n} \alpha_{ij} x_i x_j,$$

où les α_{ij}, $1 \leqslant i \leqslant j \leqslant n$ sont des scalaires. On voit qu'une forme quadratique est en fait une fonction polynôme des coordonnées $x_1, \ldots, x_n$ ne comportant que des monômes de degré 2.

Proposition 1.4 *Une forme quadratique q est associée à une unique forme bilinéaire symétrique $\tilde{\varphi}_q$: $q = q_{\tilde{\varphi}_q}$. Cette dernière est appelée forme polaire de la forme quadratique q ; elle est donnée par*

$$\forall \vec{u}, \vec{v} \in E,\ \tilde{\varphi}_q(\vec{u}, \vec{v}) = \frac{1}{2}\left[q(\vec{u} + \vec{v}) - q(\vec{u}) - q(\vec{v})\right]$$

ou encore

$$\forall \vec{u}, \vec{v} \in E,\ \tilde{\varphi}_q(\vec{u}, \vec{v}) = \frac{1}{4}\left[q(\vec{u} + \vec{v}) - q(\vec{u} - \vec{v})\right].$$

Démonstration.

— Supposons qu'il existe une forme bilinéaire symétrique φ dont q soit la forme quadratique associée. Sous cette condition, on a

$$q(\vec{u} + \vec{v}) = \varphi(\vec{u} + \vec{v}, \vec{u} + \vec{v}) = \varphi(\vec{u}, \vec{u}) + \varphi(\vec{u}, \vec{v}) + \varphi(\vec{v}, \vec{u}) + \varphi(\vec{v}, \vec{v}),$$

c'est-à-dire, puisque φ est symétrique,

$$q(\vec{u} + \vec{v}) = q(\vec{u}) + q(\vec{v}) + 2\varphi(\vec{u}, \vec{v}).$$

On obtient de même

$$q(\vec{u} - \vec{v}) = q(\vec{u}) + q(\vec{v}) - 2\varphi(\vec{u}, \vec{v}).$$

On tire de ces deux dernières égalités

$$\varphi(\vec{u}, \vec{v}) = \frac{1}{2}\left[q(\vec{u} + \vec{v}) - q(\vec{u}) - q(\vec{v})\right] = \frac{1}{4}\left[q(\vec{u} + \vec{v}) - q(\vec{u} - \vec{v})\right],$$

ce qui prouve l'unicité d'une éventuelle forme bilinéaire symétrique associée à q. Notons $\tilde{\varphi}_q$ l'application définie sur $E \times E$ par

$$\forall \vec{u}, \vec{v} \in E,\ \tilde{\varphi}_q(\vec{u}, \vec{v}) = \frac{1}{2}\left[q(\vec{u} + \vec{v}) - q(\vec{u}) - q(\vec{v})\right] = \frac{1}{4}\left[q(\vec{u} + \vec{v}) - q(\vec{u} - \vec{v})\right],$$

— Il reste à vérifier que l'application $\tilde{\varphi}_q$ définie par les égalités précédentes est bien *a posteriori* une forme bilinéaire symétrique sur E. Par définition, l'application q est la forme quadratique associée à une forme bilinéaire (pas nécessairement symétrique) ψ. Un calcul identique à celui effectué dans l'alinéa précédent conduit à

$$\psi(\vec{u}, \vec{v}) + \psi(\vec{v}, \vec{u}) = q(\vec{u} + \vec{v}) - q(\vec{u}) - q(\vec{v}).$$

On a alors

$$\tilde{\varphi}_q(\vec{u}, \vec{v}) = \frac{1}{2}\left[q(\vec{u} + \vec{v}) - q(\vec{u}) - q(\vec{v})\right] = \frac{1}{2}\left[\psi(\vec{u}, \vec{v}) + \psi(\vec{v}, \vec{u})\right].$$

Il est clair sous cette forme que $\tilde{\varphi}_q$ une forme bilinéaire symétrique sur E. C'est la forme polaire de q. □

Règle de dédoublement des produits.

Considérons une forme quadratique sur E définie relativement à une base $\mathcal{B}$ selon

$$q(\vec{u}) = \sum_{i=1}^{n} \alpha_{ii} x_i^2 + 2 \sum_{1 \leqslant i < j \leqslant n} \alpha_{ij} x_i x_j.$$

Sa forme polaire $\tilde{\varphi}_q$ se calcul comme suit :

$$\begin{aligned}
\tilde{\varphi}_q(\vec{u}, \vec{v}) &= \frac{1}{4} \left[q(\vec{u}+\vec{v}) - q(\vec{u}-\vec{v}) \right] \\
&= \frac{1}{4} \Bigg[\sum_{i=1}^{n} \alpha_{ii}(x_i+y_i)^2 + 2 \sum_{1 \leqslant i < j \leqslant n} \alpha_{ij}(x_i+y_i)(x_j+y_j) \\
&\quad - \sum_{i=1}^{n} \alpha_{ii}(x_i-y_i)^2 - 2 \sum_{1 \leqslant i < j \leqslant n} \alpha_{ij}(x_i-y_i)(x_j-y_j) \Bigg] \\
&= \frac{1}{4} \Bigg[\sum_{i=1}^{n} \alpha_{ii}(x_i^2 + 2x_iy_i + y_i^2) + 2 \sum_{1 \leqslant i < j \leqslant n} \alpha_{ij}(x_ix_j + x_iy_j + x_jy_i + x_jy_j) \\
&\quad - \sum_{i=1}^{n} \alpha_{ii}(x_i^2 - 2x_iy_i + y_i^2) - 2 \sum_{1 \leqslant i < j \leqslant n} \alpha_{ij}(x_ix_j - x_iy_j - x_jy_i + x_jy_j) \Bigg] \\
&= \sum_{i=1}^{n} \alpha_{ii} x_i y_i + \sum_{1 \leqslant i < j \leqslant n} \alpha_{ij}(x_iy_j + x_jy_i).
\end{aligned}$$

Le résultat ainsi obtenu conduit à la *règle de dédoublement des produits* : les carrés x_i^2 contenus dans q se transforment en x_iy_i dans $\tilde{\varphi}_q$ et les produits x_ix_j, $i \neq j$, en $\frac{1}{2}(x_iy_j + x_jy_i)$. Le schéma de transformation est donc

$$x_i^2 \longrightarrow x_iy_i \quad \text{et} \quad x_ix_j \longrightarrow \frac{1}{2}(x_iy_j + x_jy_i).$$

À partir de la matrice (non nécessairement symétrique) A de φ, on obtient la matrice (symétrique) de la forme polaire de q_φ selon la formule $B = \frac{1}{2}(A + {}^tA)$.

Exemple.

Soit $E = \mathbb{R}^2$ et φ la forme bilinéaire dont la matrice relativement à la base canonique est $A = \begin{pmatrix} 5 & 1 \\ 3 & 6 \end{pmatrix}$. Les écritures explicites de φ et de sa forme quadratique associée q_φ sont

$$\varphi((x,y),(x',y')) = 5xx' + xy' + 3x'y + 6yy' \quad \text{et} \quad q_\varphi(x,y) = 5x^2 + 4xy + 6y^2.$$

La règle de dédoublement des produits s'applique ici en transformant les carrés figurant dans q_φ selon les schémas $x^2 \longrightarrow xx'$ et $y^2 \longrightarrow yy'$, et le produit xy selon $xy \longrightarrow \frac{1}{2}(xy' + x'y)$. Cela fournit la forme polaire de la forme quadratique q_φ :

$$\varphi_{q_\varphi}((x,y),(x',y')) = 5xx' + 2(xy' + x'y) + 6yy'$$

dont la matrice relativement à la base canonique est $B = \begin{pmatrix} 5 & 2 \\ 2 & 6 \end{pmatrix} = \dfrac{1}{2}(A + {}^tA)$.

1.3 Changement de bases

Étudions l'effet d'un changement de base occasionné sur une forme bilinéaire en dimension finie. Soit $\mathcal{B}$ et $\mathcal{B}'$ deux bases de E et φ une forme bilinéaire sur E. Introduisons les matrices de φ relativement à chacune des bases $\mathcal{B}$ et $\mathcal{B}'$: $A = \mathrm{Mat}(\varphi; \mathcal{B})$ et $A' = \mathrm{Mat}(\varphi; \mathcal{B}')$, ainsi que la matrice de passage de $\mathcal{B}$ vers $\mathcal{B}'$: $P = P_{\mathcal{B} \to \mathcal{B}'}$. Pour deux vecteurs génériques $\vec{u}, \vec{v}$ de E de matrices-colonnes de coordonnées relativement aux bases $\mathcal{B}$ et $\mathcal{B}'$ les matrices $X = \mathrm{Mat}(\vec{u}; \mathcal{B})$, $X' = \mathrm{Mat}(\vec{u}; \mathcal{B}')$, $Y = \mathrm{Mat}(\vec{v}; \mathcal{B})$ et $Y' = \mathrm{Mat}(\vec{v}; \mathcal{B}')$, on a

$$\varphi(\vec{u}, \vec{v}) = {}^t\!XAY = {}^t\!X'A'Y'.$$

D'autre part, on sait que les matrices de coordonnées X et X' sont reliées par $X = PX'$. De même $Y = PY'$. Donc

$$\forall X, Y \in \mathcal{M}_{n\times 1}(\mathbb{R}),\ {}^t\!XAY = {}^t(PX')A(PY') = {}^t\!X'({}^t\!PAP)Y'$$

d'où l'on tire la relation de changement de base pour les formes bilinéaires :

$$\boxed{A' = {}^t\!PAP.}$$

Remarques.

1. On veillera à ne pas confondre cette formule de changement de base qui concerne les formes bilinéaires avec celle qui concerne les endomorphismes : $A' = P^{-1}AP$. Dans cette dernière, les matrices A et A' sont les matrices d'un endomorphisme respectivement relatives aux bases $\mathcal{B}$ et $\mathcal{B}'$.

2. Des matrices carrées A et A' liées par une relation de la forme $A' = {}^t\!PAP$ où P est une matrice inversible sont des *matrices congruentes*. Elles sont en particulier équivalentes, et ont entre autre le même rang. Si A est symétrique, il en de même pour A'.

Définition 1.5 *1. Soit φ une forme bilinéaire symétrique sur E, q la forme quadratique associée à φ. Les rangs de φ et q sont égaux au rang de la matrice de φ (et q) relativement à une base quelconque de E ; ils sont notés* $\mathrm{rang}(\varphi)$ *et* $\mathrm{rang}(q)$.

2. Les formes φ et q sont non dégénérées *lorsque* $\mathrm{rang}(\varphi) = \mathrm{rang}(q) = \dim(E)$ *; dans le cas contraire, elles sont* dégénérées.

2 Produit scalaire, espace vectoriel euclidien

2.1 Produit scalaire

Définition 2.1 1. *Une forme bilinéaire symétrique φ sur E est un produit scalaire sur E lorsqu'elle est* définie positive, *c'est-à-dire*

$$\forall \vec{u} \in E \setminus \{\vec{0}\},\ \varphi(\vec{u}, \vec{u}) > 0.$$

2. *Une forme quadratique est* définie positive *lorsque sa forme polaire est un produit scalaire.*
3. *Soit φ un produit scalaire sur E. La* norme associée *à φ est l'application $N_\varphi : \begin{array}{ccc} E & \longrightarrow & \mathbb{R}^+ \\ \vec{u} & \longmapsto & \sqrt{\varphi(\vec{u}, \vec{u})} \end{array}$. En d'autres termes, $N_\varphi = \sqrt{q_\varphi}$.*
4. *Un $\mathbb{R}$-espace vectoriel de dimension finie muni d'un produit scalaire est appelé* espace vectoriel euclidien.

Remarque.

La condition « définie positive » signifie que la forme quadratique associée à φ est positive ou nulle :

$$\forall \vec{u} \in E,\ q_\varphi(\vec{u}) \geqslant 0$$

et qu'elle ne s'annule que pour le vecteur nul :

$$\forall \vec{u} \in E,\ [q_\varphi(\vec{u}) = 0 \Longrightarrow \vec{u} = \vec{0}].$$

Exemple.

On sait que sur le $\mathbb{R}$-espace vectoriel $E = \mathbb{R}^2$, une forme bilinéaire symétrique φ a pour expression générale $\varphi((x, y), (x', y')) = axx' + b(xy' + x'y) + cyy'$. La forme quadratique associée à φ est donc déterminée par $q_\varphi(x, y) = ax^2 + 2bxy + cy^2$. Alors, φ est un produit scalaire si et seulement si on a $\forall (x, y) \in E \setminus \{(0, 0)\},\ ax^2 + 2bxy + cy^2 > 0$. Cette condition est équivalente aux suivantes :

$$\forall x \in \mathbb{R}^*, ax^2 > 0, \quad \forall y \in \mathbb{R}^*, cy^2 > 0,$$
$$\forall x \in \mathbb{R},\ \forall y \in \mathbb{R}^*, a\left(\frac{x}{y}\right)^2 + 2b\left(\frac{x}{y}\right) + c > 0$$

ou encore

$$a > 0,\ c > 0 \quad \text{et} \quad b^2 - ac < 0.$$

Du fait de la condition $b^2 - ac < 0$, laquelle entraîne $ac > 0$, une des inégalités $a > 0$ et $c > 0$ est redondante et une condition nécessaire et suffisante pour que φ soit un produit scalaire sur E est

$$\boxed{a > 0 \quad \text{et} \quad b^2 - ac < 0.}$$

Exemple fondamental : structure euclidienne canonique sur $\mathbb{R}^n$

Dans le $\mathbb{R}$-espace vectoriel $E = \mathbb{R}^n$, on considère l'application $\varphi : E \times E \longrightarrow \mathbb{R}$ définie pour $\vec{u} = (x_1, \ldots, x_n)$ et $\vec{v} = (y_1, \ldots, y_n)$ par

$$\varphi(\vec{u}, \vec{v}) = x_1 y_1 + \cdots + x_n y_n.$$

Il est immédiat que φ est une forme bilinéaire symétrique. De plus, elle est définie positive puisque

$$\varphi(\vec{u}, \vec{u}) = x_1^2 + \cdots + x_n^2 \geqslant 0$$

et

$$\varphi(\vec{u}, \vec{u}) = 0 \Longrightarrow x_1^2 + \cdots + x_n^2 = 0 \Longrightarrow x_1 = \cdots = x_n = 0 \Longrightarrow \vec{u} = \vec{0}.$$

C'est le *produit scalaire canonique sur* $\mathbb{R}^n$. On note usuellement

$$\vec{u}.\vec{v} = x_1 y_1 + \cdots + x_n y_n \quad \text{et} \quad \|\vec{u}\| = \sqrt{x_1^2 + \cdots + x_n^2}.$$

Autre exemple.

Dans le $\mathbb{R}$-espace vectoriel $E = (C[a, b], \mathbb{R})$ des fonctions réelles continues sur $[a, b]$, l'application $\varphi : E \times E \longrightarrow \mathbb{R}$ définie pour $f, g \in E$ par

$$\varphi(f, g) = \int_a^b f(x)\, g(x)\, dx$$

est un produit scalaire sur E. En effet, il est facile de vérifier que c'est une forme bilinéaire symétrique. Par ailleurs, elle est également définie positive puisque

$$\varphi(f, f) = \int_a^b f(x)^2\, dx \geqslant 0$$

et

$$\varphi(f, f) = 0 \Longrightarrow \int_a^b f(x)^2\, dx = 0 \Longrightarrow f = 0.$$

La dernière égalité ci-dessus nécessite une justification : si f n'était pas identiquement nulle sur $[a, b]$, il existerait un point $x_0 \in [a, b]$ tel que $f(x_0) \neq 0$ et par continuité on aurait $|f(x)| \geqslant |f(x_0)|/2$ sur un voisinage $[x_0 - \varepsilon, x_0 + \varepsilon]$ de x_0. Par suite, on aurait

$$\int_a^b f(x)^2\, dx \geqslant \int_{x_0-\varepsilon}^{x_0+\varepsilon} f(x_0)^2\, dx = 2\varepsilon f(x_0)^2 > 0,$$

ce qui contredirait l'hypothèse $\displaystyle\int_a^b f(x)^2\, dx = 0$. La norme associée est donnée par

$$\|f\|_2 = \sqrt{\int_a^b f(x)^2\, dx}.$$

DESSIN ?

Théorème 2.2 (Inégalité de Cauchy [1]-Schwarz [2]) *Soit φ un produit scalaire sur E. On a l'inégalité*

$$\forall \vec{u}, \vec{v} \in E,\ |\varphi(\vec{u}, \vec{v})| \leqslant N_\varphi(\vec{u}) \times N_\varphi(\vec{v}).$$

De plus, l'égalité $|\varphi(\vec{u}, \vec{v})| = N_\varphi(\vec{u}) \times N_\varphi(\vec{v})$ est satisfaite si et seulement si les vecteurs $\vec{u}$ et $\vec{v}$ sont colinéaires.

Démonstration.

Soit P la fonction définie sur $\mathbb{R}$ par $P(\lambda) = N_\varphi(\lambda\vec{u} + \vec{v})^2$ pour tout $\lambda \in \mathbb{R}$.

1. On a

$$P(\lambda)=\varphi(\lambda\vec{u}+\vec{v}, \lambda\vec{u}+\vec{v})=\lambda^2\varphi(\vec{u}, \vec{u})+2\lambda\varphi(\vec{u}, \vec{v})+\varphi(\vec{v}, \vec{v})=\lambda^2 N_\varphi(\vec{u})^2+2\lambda\varphi(\vec{u}, \vec{v})+N_\varphi(\vec{v}, \vec{v})^2.$$

Sous cette dernière forme, il apparaît que P est un polynôme de degré 2 positif. Cela signifie que son discriminant est négatif ou nul :

$$\varphi(\vec{u}, \vec{v})^2 - N_\varphi(\vec{u})^2 N_\varphi(\vec{v})^2 \leqslant 0,$$

condition équivalente à l'inégalité de Cauchy-Schwarz.

2. Considérons maintenant deux vecteurs $\vec{u}, \vec{v}$ tels que $|\varphi(\vec{u}, \vec{v})| = N_\varphi(\vec{u}) \times N_\varphi(\vec{v})$.
   — Si $\vec{u} = \vec{0}$, l'égalité ci-dessus est vérifiée et les vecteurs $\vec{u}$ et $\vec{v}$ sont colinéaires.
   — Si $\vec{u} \neq \vec{0}$, le polynôme P a un discriminant nul, il possède une racine double : $\exists\lambda \in \mathbb{R}$, $P(\lambda) = 0$. Cela s'écrit $\exists\lambda \in \mathbb{R},\ N_\varphi(\lambda\vec{u} + \vec{v}) = 0$, ou encore $\exists\lambda \in \mathbb{R},\ \lambda\vec{u} + \vec{v} = \vec{0}$, *i.e.* les vecteurs $\vec{u}$ et $\vec{v}$ sont colinéaires.

Théorème 2.3 (Inégalité triangulaire (ou de Minkowski [3])) *Soit φ un produit scalaire sur E. On a l'inégalité*

$$N_\varphi(\vec{u} + \vec{v}) \leqslant N_\varphi(\vec{u}) + N_\varphi(\vec{v})$$

De plus, l'égalité $N_\varphi(\vec{u} + \vec{v}) = N_\varphi(\vec{u}) + N_\varphi(\vec{v})$ est satisfaite si et seulement si les vecteurs $\vec{u}$ et $\vec{v}$ sont colinéaires et de même sens.

Démonstration.

1. Développons le carré ci-dessous :

$$[N_\varphi(\vec{u}) + N_\varphi(\vec{v})]^2 = N_\varphi(\vec{u})^2 + 2N_\varphi(\vec{u}) \times N_\varphi(\vec{v}) + N_\varphi(\vec{v})^2.$$

En utilisant l'inégalité de Cauchy-Schwarz, on trouve

$$[N_\varphi(\vec{u}) + N_\varphi(\vec{v})]^2 \geqslant N_\varphi(\vec{u})^2 + 2\varphi(\vec{u}, \vec{v}) + N_\varphi(\vec{v})^2 = N_\varphi(\vec{u} + \vec{v})^2$$

d'où l'inégalité triangulaire en prenant les racines carrées de chaque membre de cette inégalité.

1. Cauchy, Augustin Louis : mathématicien français (Paris 1789–Sceaux 1857)
2. Schwarz, Hermann Amandus : mathématicien allemand (Hermsdorf 1843–Berlin 1921)
3. Minkowski, Hermann : mathématicien allemand (Kovno 1864–Göttingen 1909)

2. Si les vecteurs $\vec{u}$ et $\vec{v}$ vérifient l'égalité $N_\varphi(\vec{u}) + N_\varphi(\vec{v}) = N_\varphi(\vec{u}+\vec{v})$, alors
$$[N_\varphi(\vec{u}) + N_\varphi(\vec{v})]^2 = N_\varphi(\vec{u}+\vec{v})^2,$$
relation équivalente à
$$\varphi(\vec{u},\vec{v}) = N_\varphi(\vec{u}) \times N_\varphi(\vec{v}).$$
C'est un cas particulier du cas d'égalité dans l'inégalité de Cauchy-Schwarz (notons l'absence de valeur absolue dans le premier membre). Les vecteurs $\vec{u}$ et $\vec{v}$ sont nécessairement colinéaires. Par exemple, $\exists\lambda \in \mathbb{R},\ \vec{v} = \lambda\vec{u}$. En reportant cela dans l'égalité ci-dessus, il vient la condition $|\lambda| = \lambda$ qui impose au coefficient de proportionnalité d'être positif : $\lambda \geqslant 0$. Les vecteurs $\vec{u}$ et $\vec{v}$ sont donc colinéaires de même sens. □

Remarque.

Pour les vecteurs $\vec{u} = (|x_1|, \ldots, |x_n|)$ et $\vec{v} = (|y_1|, \ldots, |y_n|)$ de $\mathbb{R}^n$, l'inégalité de Cauchy-Schwarz prend la forme remarquable
$$\boxed{\left|\sum_{i=1}^n x_i y_i\right| \leqslant \sum_{i=1}^n |x_i y_i| \leqslant \sqrt{\sum_{i=1}^n x_i^2}\sqrt{\sum_{i=1}^n y_i^2}}$$
et pour les fonctions $\vec{u} = |f|$ et $\vec{v} = |g|$ de $C([a,b],\mathbb{R})$, elle revêt la forme non moins remarquable
$$\boxed{\left|\int_a^b f(x)\,g(x)\,dx\right| \leqslant \int_a^b |f(x)\,g(x)|\,dx \leqslant \sqrt{\int_a^b f(x)^2\,dx}\sqrt{\int_a^b g(x)^2\,dx}.}$$
Les inégalités intermédiaires écrites en premiers membres proviennent de l'inégalité triangulaire sur $\mathbb{R}$ appliquée à n vecteurs.

On voit ainsi que la norme N_φ associée à un produit scalaire φ vérifie les propriétés d'une *norme* générale dont voici la définition.

Définition 2.4 *Une norme sur E est une application $N : E \longrightarrow \mathbb{R}^+$ vérifiant les propriétés suivantes :*
1. $\forall\vec{u} \in E \setminus \{\vec{0}\},\ N(\vec{u}) > 0$;
2. $\forall\lambda \in \mathbb{R}, \forall\vec{u} \in E,\ N(\lambda\vec{u}) = |\lambda| N(\vec{u})$;
3. $\forall\vec{u},\vec{v} \in E,\ N_\varphi(\vec{u}+\vec{v}) \leqslant N_\varphi(\vec{u}) + N_\varphi(\vec{v})$.

La norme N_φ étant issue d'un produit scalaire, à savoir φ, porte le nom de *norme euclidienne*. Une question naturelle qui se pose alors est la suivante : disposant d'une norme sur E, est-elle euclidienne, c'est-à-dire est-elle associée à un produit scalaire sur E ? La réponse n'est pas toujours positive et on démontre qu'une norme est euclidienne si et seulement si elle vérifie l'*identité du parallélogramme* (dont la démonstration est immédiate) énoncée ci-dessous.

Théorème 2.5 (Identité du parallélogramme) *La norme N_φ vérifie l'égalité*
$$N_\varphi(\vec{u}+\vec{v})^2 + N_\varphi(\vec{u}-\vec{v})^2 = 2[N_\varphi(\vec{u})^2 + N_\varphi(\vec{v})^2].$$

Cette égalité s'interprète comme suit : dans le parallélogramme formé par les vecteurs $\vec{u}$ et $\vec{v}$, la somme des carrés des longueurs des diagonales est égale à la somme des carrés des longueurs des côtés (voir Fig. 6.1).

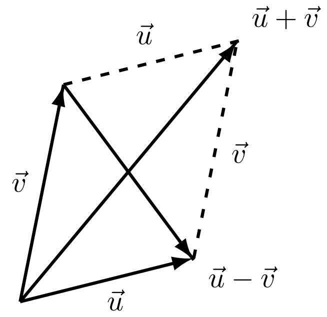

FIGURE 6.1 – Identité du parallélogramme

2.2 Orthogonalité

Dans cette partie, on se place dans un espace vectoriel euclidien (E, φ) et l'on définit diverses notions d'orthogonalité. Ces notions pourraient être étendues au cas d'une forme bilinéaire symétrique qui n'est pas définie positive mais cette situation ne sera pas examinée ici.

(a) *Orthogonalité de deux vecteurs, de deux sous-espaces vectoriels*

Définition 2.6*(a). Deux vecteurs $\vec{u}, \vec{v} \in E$ sont* orthogonaux *lorsque $\varphi(\vec{u}, \vec{v}) = 0$. On note cette propriété $\vec{u} \perp \vec{v}$.*

(b). Deux sous-espaces vectoriels F et G de E sont orthogonaux *lorsque $\forall(\vec{u}, \vec{v}) \in F \times G$, $\vec{u} \perp \vec{v}$.*

Théorème 2.7 (Pythagore[4]) *Soit $\vec{u}, \vec{v}$ deux vecteurs de E. On a l'équivalence*

$$\vec{u} \perp \vec{v} \iff N_\varphi(\vec{u} + \vec{v})^2 = N_\varphi(\vec{u})^2 + N_\varphi(\vec{v})^2.$$

DÉMONSTRATION.

Le développement du carré $N_\varphi(\vec{u} + \vec{v})^2$ donne

$$N_\varphi(\vec{u} + \vec{v})^2 = N_\varphi(\vec{u})^2 + 2\varphi(\vec{u}, \vec{v}) + N_\varphi(\vec{v})^2$$

et l'on voit que

$$N_\varphi(\vec{u} + \vec{v})^2 = N_\varphi(\vec{u})^2 + N_\varphi(\vec{v})^2 \iff \varphi(\vec{u}, \vec{v}) = 0 \iff \vec{u} \perp \vec{v}. \qquad \square$$

4. Pythagore : philosophe et mathématicien grec (Samos vers 570 av. J.-C–Métaponte, 480 av. J.-C)

(b) *Bases orthonormées, procédé d'orthogonalisation de Gram*[5]*-Schmidt*[6]

Définition 2.8 *Soit $\mathcal{S} = \{\vec{u}_1, \dots, \vec{u}_k\}$ un système de vecteurs de E.*
— *Le système $\mathcal{S}$ est* orthogonal *lorsque les vecteurs de $\mathcal{S}$ sont deux à deux orthogonaux :*

$$\forall i, j \in \{1, \dots, k\},\ i \neq j \Longrightarrow \vec{u}_i \perp \vec{u}_j.$$

— *Le système $\mathcal{S}$ est* orthonormé *lorsqu'il est orthogonal et tous les vecteurs de $\mathcal{S}$ sont* normés *(ou* unitaires*), c'est-à-dire de norme 1. Ces conditions se résument selon :*

$$\forall i, j \in \{1, \dots, k\},\ \varphi(\vec{u}_i, \vec{u}_j) = \begin{cases} 0 & \text{si } i \neq j \\ 1 & \text{si } i = j \end{cases}.$$

Proposition 2.9 *Un système orthogonal ne contenant pas le vecteur nul est libre.*

Démonstration.

Soit $\mathcal{S} = \{\vec{u}_1, \dots, \vec{u}_k\}$ un système de vecteurs de E orthogonal et $\alpha_1, \dots, \alpha_k$ des scalaires tels que $\alpha_1\vec{u}_1 + \dots + \alpha_k\vec{u}_k = \vec{0}$. On a alors

$$\forall i \in \{1, \dots, k\},\ \varphi(\alpha_1\vec{u}_1 + \dots + \alpha_k\vec{u}_k, \vec{u}_i) = 0.$$

Or

$$\varphi(\alpha_1\vec{u}_1 + \dots + \alpha_k\vec{u}_k, \vec{u}_i) = \alpha_1\varphi(\vec{u}_1, \vec{u}_i) + \dots + \alpha_k\varphi(\vec{u}_k, \vec{u}_i)\alpha_i\varphi(\vec{u}_i, \vec{u}_i) = \alpha_i N_\varphi(\vec{u}_i).$$

On a *a fortiori*

$$\forall i \in \{1, \dots, k\},\ \alpha_i N_\varphi(\vec{u}_i) = 0.$$

Si l'on suppose de plus que $\mathcal{S}$ ne contient pas le vecteur nul, on a $N_\varphi(\vec{u}_i) \neq 0$ et par suite $\forall i \in \{1, \dots, k\},\ \alpha_i = 0$, prouvant ainsi que le système $\mathcal{S}$ est libre. □

À ce niveau, une question naturelle qui se pose est la suivante : existe-t-il des bases de E orthonormées ? Examinons tout d'abord un exemple.

Exemple.

Dans le $\mathbb{R}$-espace vectoriel $\mathbb{R}^n$ muni de sa structure euclidienne canonique, la base canonique est une base orthonormée : si $\vec{e}_i = (0, \dots, 1, \dots, 0)$ pour $1 \leqslant i \leqslant n$, alors $\vec{e}_i.\vec{e}_j = 0$ si $i \neq j$ et $\|\vec{e}_i\| = 1$.

La réponse à la question posée ci-dessus est positive dans le cas général et le procédé d'orthogonalisation de Gram-Schmidt permet de construire une base orthonormée à partir d'une base quelconque de E.

Soit $\{\vec{u}_1, \dots, \vec{u}_n\}$ une base de E. On va construire une base orthogonale $\{\vec{e}_1, \dots, \vec{e}_n\}$ telle que pour tout $i \in \{1, \dots, n\}$, $\vec{e}_i$ soit combinaison linéaire de $\vec{u}_1, \dots, \vec{u}_i$, puis en déduire une base orthonormée $\{\vec{\varepsilon}_1, \dots, \vec{\varepsilon}_n\}$ en posant $\vec{\varepsilon}_i = \dfrac{1}{N_\varphi(\vec{e}_i)}$.

— On commence par poser simplement $\vec{e}_1 = \vec{u}_1$.

5. Gram, Jørgen Pedersen : mathématicien danois (Nustrup 1850–Copenhague 1916)
6. Schmidt, Erhard : mathématicien allemand (Dorpat 1876–Berlin 1959)

— On cherche ensuite le vecteur $\vec{e}_2$ sous la forme $\vec{e}_2 = \vec{u}_2 - \alpha_{12}\vec{e}_1$ où le scalaire α_{12} est tel que $\vec{e}_2 \perp \vec{e}_1$. On a

$$\varphi(\vec{e}_1, \vec{e}_2) = \varphi(\vec{e}_1, \vec{u}_2) - \alpha_{12}\varphi(\vec{e}_1, \vec{e}_1)$$

et la condition $\vec{e}_2 \perp \vec{e}_1$ est satisfaite pour le scalaire

$$\alpha_{12} = \frac{\varphi(\vec{e}_1, \vec{u}_2)}{\varphi(\vec{e}_1, \vec{e}_1)}.$$

D'où le vecteur

$$\vec{e}_2 = \vec{u}_2 - \frac{\varphi(\vec{e}_1, \vec{u}_2)}{\varphi(\vec{e}_1, \vec{e}_1)}\,\vec{e}_1.$$

DESSIN A FAIRE

— De manière analogue, on cherche le vecteur $\vec{e}_3$ sous la forme $\vec{e}_3 = \vec{u}_3 - \alpha_{13}\vec{e}_1 - \alpha_{23}\vec{e}_2$ où les scalaires α_{13} et α_{23} sont tels que $\vec{e}_3 \perp \vec{e}_1$ et $\vec{e}_3 \perp \vec{e}_2$. On a

$$\begin{aligned}
\varphi(\vec{e}_1, \vec{e}_3) &= \varphi(\vec{e}_1, \vec{u}_3) - \alpha_{13}\varphi(\vec{e}_1, \vec{e}_1) - \alpha_{23}\varphi(\vec{e}_1, \vec{e}_2)\\
\varphi(\vec{e}_2, \vec{e}_3) &= \varphi(\vec{e}_2, \vec{u}_3) - \alpha_{13}\varphi(\vec{e}_2, \vec{e}_1) - \alpha_{23}\varphi(\vec{e}_2, \vec{e}_2)
\end{aligned}$$

et les conditions $\vec{e}_3 \perp \vec{e}_1$ et $\vec{e}_3 \perp \vec{e}_2$ sont satisfaites avec les choix

$$\alpha_{13} = \frac{\varphi(\vec{e}_1, \vec{u}_3)}{\varphi(\vec{e}_1, \vec{e}_1)} \quad \text{et} \quad \alpha_{23} = \frac{\varphi(\vec{e}_2, \vec{u}_3)}{\varphi(\vec{e}_2, \vec{e}_2)}.$$

On a ainsi trouvé le vecteur

$$\vec{e}_3 = \vec{u}_3 - \frac{\varphi(\vec{e}_1, \vec{u}_3)}{\varphi(\vec{e}_1, \vec{e}_1)}\,\vec{e}_1 - \frac{\varphi(\vec{e}_2, \vec{u}_3)}{\varphi(\vec{e}2, \vec{e}_2)}\,\vec{e}_2.$$

DESSIN A FAIRE

— On construit ainsi de manière récurrente le vecteur $\vec{e}_j$ pour $j \leqslant n$ à l'aide du vecteur donné $\vec{u}_j$ et des vecteurs précédemment construits $\vec{e}_1, \dots, \vec{e}_{j-1}$ selon

$$\vec{e}_j = \vec{u}_j - \sum_{i=1}^{j-1} \frac{\varphi(\vec{e}_i, \vec{u}_j)}{\varphi(\vec{e}_i, \vec{e}_i)}\,\vec{e}_i.$$

— On dispose finalement d'un système orthogonal $\{\vec{e}_1, \dots, \vec{e}_n\}$. Il est libre et contient n vecteurs, c'est donc une base orthogonale de E. L'ultime étape consiste à normer chacun des vecteurs $\vec{e}_j$ en les remplaçant par $\vec{\varepsilon}_j = \dfrac{1}{N_\varphi(\vec{e}_j)}\,\vec{e}_j$ et le système $\{\vec{\varepsilon}_1, \dots, \vec{\varepsilon}_n\}$ est une base orthonormée de E.

Remarque.

Signalons une astuce pour le calcul de la norme $N_\varphi(\vec{e}_j)$ apparaissant dans l'ultime étape de l'algorithme précédent : on a

$$\begin{aligned}
N_\varphi(\vec{e}_j)^2 = \varphi(\vec{e}_j, \vec{e}_j) &= \varphi(\vec{e}_j, \vec{u}_j - \sum_{i=1}^{j-1} \frac{\varphi(\vec{e}_i, \vec{u}_j)}{\varphi(\vec{e}_i, \vec{e}_i)}\,\vec{e}_i)\\
&= \varphi(\vec{e}_j, \vec{u}_j) - \sum_{i=1}^{j-1} \frac{\varphi(\vec{e}_i, \vec{u}_j)}{\varphi(\vec{e}_i, \vec{e}_i)}\,\varphi(\vec{e}_j, \vec{e}_i) = \varphi(\vec{e}_j, \vec{u}_j)
\end{aligned}$$

puisque le vecteur $\vec{e}_j$ est orthogonal à tous les $\vec{e}_1, \dots, \vec{e}_{j-1}$. Compte tenu de l'expression relativement complexe du vecteurs $\vec{e}_j$, la quantité $\varphi(\vec{e}_j, \vec{u}_j)$ est en pratique plus simple à évaluer que $\varphi(\vec{e}_j, \vec{e}_j)$.

Exemple.

Soit, dans le $\mathbb{R}$-espace vectoriel $E = \mathbb{R}^3$ muni de sa structure euclidienne canonique, la base $\{\vec{u}_1, \vec{u}_2, \vec{u}_3\}$ donnée par

$$\vec{u}_1 = (1, 1, 1),\ \vec{u}_2 = (1, 1, 0),\ \vec{u}_3 = (1, 0, 0).$$

Construisons selon le procédé de Gram-Schmidt une base orthonormée $\{\vec{\varepsilon}_1, \vec{\varepsilon}_2, \vec{\varepsilon}_3\}$ de E à partir de la base donnée. On commence par construire une base orthogonale $\{\vec{e}_1, \vec{e}_2, \vec{e}_3\}$ de E comme suit :

$$\vec{e}_1 = \vec{u}_1, \quad \vec{e}_2 = \vec{u}_2 - \frac{\vec{e}_1.\vec{u}_2}{\vec{e}_1.\vec{e}_1}\,\vec{e}_1, \quad \vec{e}_3 = \vec{u}_3 - \frac{\vec{e}_1.\vec{u}_3}{\vec{e}_1.\vec{e}_1}\,\vec{e}_1 - \frac{\vec{e}_2.\vec{u}_3}{\vec{e}_2.\vec{e}_2}\,\vec{e}_2.$$

De simples calculs conduisent à

$$\vec{e}_1 = (1, 1, 1), \quad \vec{e}_2 = \left(\frac{1}{3}, \frac{1}{3}, -\frac{2}{3}\right), \quad \vec{e}_3 = \left(\frac{1}{2}, -\frac{1}{2}, 0\right).$$

Enfin, on normalise cette base en posant

$$\vec{\varepsilon}_1 = \frac{1}{\|\vec{e}_1\|}\,\vec{e}_1, \quad \vec{\varepsilon}_2 = \frac{1}{\|\vec{e}_2\|}\,\vec{e}_2, \quad \vec{\varepsilon}_3 = \frac{1}{\|\vec{e}_3\|}\,\vec{e}_3$$

et l'on obtient explicitement

$$\vec{\varepsilon}_1 = \left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right), \quad \vec{\varepsilon}_1 = \left(\frac{1}{\sqrt{6}}, \frac{1}{\sqrt{6}}, -\frac{2}{\sqrt{6}}\right), \quad \vec{\varepsilon}_1 = \left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}, 0\right).$$

Théorème 2.10 *Soit $\mathcal{B} = \{\vec{e}_1, \dots, \vec{e}_n\}$ une base orthonormée de E.*

(a). On a

$$\forall \vec{u} \in E,\ \vec{u} = \sum_{i=1}^{n} \varphi(\vec{u}, \vec{e}_i)\vec{e}_i.$$

En d'autres termes, les coordonnées du vecteur $\vec{u}$ relativement à $\mathcal{B}$ sont les scalaires $\varphi(\vec{u}, \vec{e}_i)$, $1 \leqslant i \leqslant n$.

(b). On a

$$\forall \vec{u}, \vec{v} \in E,\ \varphi(\vec{u}, \vec{v}) = \sum_{i=1}^{n} \varphi(\vec{u}, \vec{e}_i)\varphi(\vec{v}, \vec{e}_i) \quad \text{et} \quad N_\varphi(\vec{u}) = \sqrt{\sum_{i=1}^{n} \varphi(\vec{u}, \vec{e}_i)^2}.$$

En posant $x_i = \varphi(\vec{u}, \vec{e}_i)$, $y_i = \varphi(\vec{v}, \vec{e}_i)$ et en introduisant les matrices-colonnes $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ et $Y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$, cela s'écrit plus simplement

$$\forall \vec{u}, \vec{v} \in E,\ \varphi(\vec{u}, \vec{v}) = \sum_{i=1}^{n} x_i y_i = {}^tXY = {}^tYX \quad \text{et} \quad N_\varphi(\vec{u}) = \sqrt{\sum_{i=1}^{n} x_i^2}.$$

(c). La matrice du produit scalaire φ relativement à la base orthonormée $\mathcal{B}$ est la matrice unité :

$$\mathrm{Mat}(\varphi; \mathcal{B}) = I_n.$$

En particulier, le rang de φ est n et la forme bilinéaire φ est non dégénérée.

(c) *Orthogonal d'un sous-espace vectoriel*

Définition 2.11 *Soit F un sous-espace vectoriel de E. L'orthogonal de F est le sous-ensemble de E noté $F^\perp$ défini par*

$$F^\perp = \{\vec{u} \in E \ :\ \forall \vec{v} \in F, \vec{u} \perp \vec{v}\}.$$

Théorème 2.12 *Soit F un sous-espace vectoriel de E.*

(a). L'orthogonal $F^\perp$ de F est un sous-espace vectoriel de E.

(b). Les sous-espaces vectoriels F et $F^\perp$ sont supplémentaires dans E, et donc

$$\dim(F^\perp) = \dim(E) - \dim(F).$$

Le sous-espace vectoriel $F^\perp$ est le supplémentaire orthogonal de F.

(c). L'orthogonal de $F^\perp$ est F : $(F^\perp)^\perp = F$.

DÉMONSTRATION.

— Soit $\alpha, \beta \in \mathbb{R}$ et $\vec{u}, \vec{v} \in F^\perp$. On a pour tout $\vec{w} \in F$

$$\varphi(\alpha\vec{u} + \beta\vec{v}, \vec{w}) = \alpha\varphi(\vec{u}, \vec{w}) + \beta\varphi(\vec{v}, \vec{w}) = 0$$

puisque $\vec{u} \perp \vec{w}$ et $\vec{v} \perp \vec{w}$. Donc $\alpha\vec{u} + \beta\vec{v} \perp \vec{w}$, *i.e.* $\alpha\vec{u} + \beta\vec{v} \in F^{\perp}$, ce qui prouve que $F^{\perp}$ est un sous-espace vectoriel de E.

— Soit $\mathcal{B}_F$ une base orthonormée de F. En utilisant le procédé d'orthogonalisation de Gram-Schmidt, on peut compléter le système orthonormé $\mathcal{B}_F$ en une base orthonormée $\mathcal{B}$ de E : $\mathcal{B} = \mathcal{B}_F \cup \mathcal{B}'$. Dans cette construction, les vecteurs de $\mathcal{B}'$ sont orthogonaux à ceux de $\mathcal{B}_F$, donc $\mathcal{B}' \subset F^{\perp}$ puis

$$E = \mathrm{vect}(\mathcal{B}) = \mathrm{vect}(\mathcal{B}_F \cup \mathcal{B}') = \mathrm{vect}(\mathcal{B}_F) + \mathrm{vect}(\mathcal{B}') \subset F + F^{\perp}.$$

Ayant évidemment l'inclusion inverse $F + F^{\perp} \subset E$, on a exactement l'égalité

$$E = F + F^{\perp}.$$

Soit d'autre part $\vec{u} \in F \cap F^{\perp}$. On a $\vec{u} \in F^{\perp}$, donc $\forall \vec{v} \in F,\ \varphi(\vec{u}, \vec{v}) = 0$. En particulier, puisque $\vec{u} \in F$, $\varphi(\vec{u}, \vec{u}) = 0$. L'application φ étant un produit scalaire, seul le vecteur vérifie cette relation ; $\vec{u} = \vec{0}$ et par conséquent

$$F \cap F^{\perp} = \{\vec{0}\}.$$

On a ainsi prouvé que

$$E = F \oplus F^{\perp},$$

En d'autres termes, les sous-espace vectoriel F et $F^{\perp}$ sont supplémentaires dans E.

— Soit $\vec{u} \in F$. On a $\forall \vec{v} \in F,\ \varphi(\vec{u}, \vec{v}) = 0$, et alors $\vec{u} \in (F^{\perp})^{\perp}$. Ceci démontre l'inclusion $F \subset (F^{\perp})^{\perp}$. Par ailleurs, en utilisant le point précédent,

$$\dim((F^{\perp})^{\perp}) = \dim(E) - \dim(F^{\perp}) = \dim(E) - [\dim(E) - \dim(F)] = \dim(F).$$

Cela permet de conclure à l'égalité $F = (F^{\perp})^{\perp}$. □

Remarque.

En dimension infinie, le *biorthogonal de* F, noté $F^{\perp\perp}$, est parfois plus grand que F.

Soit par exemple $E = \ell^2$ l'ensemble des suites réelles $(u_n)_{n\in\mathbb{N}}$ telles que la série $(\sum u_n^2)$ soit convergente. Si $u = (u_n)_{n\in\mathbb{N}}$ et $v = (v_n)_{n\in\mathbb{N}}$ sont deux suites de E, alors l'inégalité élémentaire $(x+y)^2 \leqslant 2(x^2+y^2)$ montre que la série $(\sum (u_n+v_n)^2)$ est convergente et donc que la suite $u+v$ est un élément de E. De plus si $\lambda \in \mathbb{R}$, alors la suite λu est aussi un élément de E. Ainsi E est un $\mathbb{R}$-espace vectoriel.

Par ailleurs, l'inégalité de Cauchy-Schwarz assure que la série $(\sum u_n v_n)$ est convergente. En effet, pour tout $n \in \mathbb{N}$:

$$\sum_{k=0}^{n} |u_k v_k| \leqslant \sqrt{\sum_{k=0}^{n} u_k^2}\sqrt{\sum_{k=0}^{n} v_k^2} \leqslant \sqrt{\sum_{k=0}^{\infty} u_k^2}\sqrt{\sum_{k=0}^{\infty} v_k^2}$$

donc la suite des sommes partielles de la série à termes positifs $(\sum |u_k v_k|)$ est bornée, cette série est donc convergente. La série $(\sum u_n v_n)$ est absolument convergente, donc convergente.

On pose alors

$$\forall u, v \in E,\ \varphi(u, v) = \sum_{k=0}^{\infty} u_k v_k.$$

On vérifie facilement que φ définit un produit scalaire sur E.

Considérons à présent dans E le sous-ensemble F des suites à support fini, c'est-à-dire qui n'ont qu'un nombre fini de termes non nuls. Il est clair que F est un sous-espace vectoriel de E qui contient les suites élémentaires $e_0, e_1, e_2, \ldots$ constituées d'un seul terme non nul et qui vaut 1 : $e_k = (0, \ldots, 0, 1, 0, \ldots)$, le 1 se trouvant à la k^{e} position ($k \in \mathbb{N}$).

Déterminons l'orthogonal du sous-espace vectoriel F relativement à φ. Un vecteur $u = (u_n)_{n\in\mathbb{N}}$ de E orthogonal à F est en particulier orthogonal à toutes les suites e_k, $k \in \mathbb{N}$: $\varphi(u, e_k) = u_k = 0$ et donc $u = 0$, *i.e.* u est la suite nulle. Ainsi $F^{\perp} = \{0\}$, puis clairement $(F^{\perp})^{\perp} = E$. On a un exemple de sous-espace vectoriel F pour lequel $F \subsetneqq (F^{\perp})^{\perp}$.

Proposition 2.13 *Soit F et G deux sous-espaces vectoriels de E.*
(a). Si $F \subset G$, alors $G^\perp \subset F^\perp$.
(b). Les orthogonaux de $F+G$ et $F \cap G$ sont respectivement donnés par

$$(F+G)^\perp = F^\perp \cap G^\perp \quad \text{et} \quad (F \cap G)^\perp = F^\perp + G^\perp.$$

Démonstration.

(a). Soit $\vec{u} \in E$. On a les implications

$$\vec{u} \in F^\perp \Longrightarrow \forall \vec{v} \in G,\ \vec{u} \perp \vec{v} \Longrightarrow \forall \vec{v} \in F,\ \vec{u} \perp \vec{v} \Longrightarrow \vec{u} \in G^\perp.$$

D'où l'inclusion $G^\perp \subset F^\perp$.

(b). Soit $\vec{u} \in E$. Les équivalences suivantes sont très claires :

$$\vec{u} \in (F+G)^\perp \Longleftrightarrow \vec{u} \perp (F+G) \Longleftrightarrow \vec{u} \perp F \text{ et } \vec{u} \perp G \Longleftrightarrow \vec{u} \in F^\perp \cap G^\perp.$$

Elles montrent l'égalité

$$(F+G)^\perp = F^\perp \cap G^\perp.$$

Puis, en considérant les orthogonaux des sous-espaces vectoriels ci-dessus, après application du théorème 2.12 et changement de notations,

$$(F'^\perp \cap G'^\perp)^\perp = [(F'+G')^\perp]^\perp = F' + G'.$$

Prenons pour F' et G' les sous-espaces vectoriels $F^\perp$ et $G^\perp$. Une nouvelle application du théorème 2.12 fournit $F'^\perp = F$ et $G'^\perp = G$ et la relation précédente s'écrit

$$(F \cap G)^\perp = F^\perp + G^\perp. \qquad \square$$

Exemple.

Soit $\vec{u} \in E$ et $\mathcal{D}$ la droite vectorielle engendrée par le vecteur $\vec{u}$. Déterminons l'orthogonal de $\mathcal{D}$. Fixons une base orthonormée $\mathcal{B} = \{\vec{e}_1, \dots, \vec{e}_n\}$ de E et introduisons les coordonnées de $\vec{u}$ relativement à $\mathcal{B}$: $\vec{u} = a_1\vec{e}_1 + \cdots + a_n\vec{e}_n$, puis $\vec{v}$ un vecteur générique de E : $\vec{v} = x_1\vec{e}_1 + \cdots + x_n\vec{e}_n$. On a

$$\vec{v} \in \mathcal{D}^\perp \Longleftrightarrow \forall \vec{w} \in \mathcal{D},\ \vec{w} \perp \vec{v} \Longleftrightarrow \vec{u} \perp \vec{v} \Longleftrightarrow \varphi(\vec{u}, \vec{v}) = 0 \Longleftrightarrow a_1x_1 + \cdots + a_nx_n = 0.$$

L'orthogonal de $\mathcal{D}$ est donc le sous-espace vectoriel

$$\boxed{\mathcal{D}^\perp = \mathcal{H} = \{x_1\vec{e}_1 + \cdots + x_n\vec{e}_n \ :\ a_1x_1 + \cdots + a_nx_n = 0\}.}$$

On a $\dim(\mathcal{H}) = \dim(E) - \dim(\mathcal{D}) = n-1$, $\mathcal{H}$ est donc un hyperplan vectoriel. Réciproquement, l'orthogonal de l'hyperplan vectoriel $\mathcal{H}$ est la droite vectorielle $\mathcal{D}$:

$$\boxed{\mathcal{H}^\perp = \mathcal{D} = \mathrm{vect}(a_1\vec{e}_1 + \cdots + a_n\vec{e}_n).}$$

(d) *Distance entre sous-espaces vectoriels*

A FAIRE

2.3 Réduction des formes quadratiques

Les *valeurs propres* d'une forme quadratique sont les valeurs propres de sa matrice relativement à une base orthonormée (quelconque) de E.

Corollaire 2.14 *Soit q une forme quadratique sur E. Notons $\lambda_1, \ldots, \lambda_n$ ses valeurs propres. Alors il existe une base orthonormée $\mathcal{B}' = \{\vec{e}'_1, \ldots, \vec{e}'_n\}$ de E relativement à laquelle on a*

$$\textit{pour tout } \vec{u} = \sum_{i=1}^{n} x_i \vec{e}_i \in E,\ q(\vec{u}) = \sum_{i=1}^{n} \lambda_i x_i^2.$$

Démonstration.

Soit $\mathcal{B} = \{\vec{e}_1, \ldots, \vec{e}_n\}$ une base orthonormée de E et $A = \mathrm{Mat}(q; \mathcal{B}) = (a_{ij})_{1 \leqslant i,j \leqslant n}$. La matrice A est symétrique réelle, donc diagonalisable via une matrice de passage orthogonale : $A = PD{}^tP$ où D est la matrice diagonale des valeurs propres de q, $D = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$, et P est une matrice orthogonale : ${}^tP = P^{-1}$ (cf. théorème 3.6). La matrice P peut donc être considérée comme la matrice de passage de la base orthonormée $\mathcal{B}$ vers une autre base orthonormée $\mathcal{B}'$: $P = P_{\mathcal{B} \to \mathcal{B}'}$ (cf. proposition 3.5). Introduisons alors les matrices-coordonnées d'un vecteur $\vec{u}$ générique de E : $X = \mathrm{Mat}(\vec{u}; \mathcal{B}) = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ et $X' = \mathrm{Mat}(\vec{u}; \mathcal{B}') = \begin{pmatrix} x'_1 \\ \vdots \\ x'_n \end{pmatrix}$. Rappelons pour mémoire que, relativement à la base $\mathcal{B}$,

$$q(\vec{u}) = \sum_{1 \leqslant i,j \leqslant n} x_i x_j = {}^tXAX.$$

Effectuons le changement de bases $X = PX'$. La formule de transformation correspondante au niveau de la forme quadratique donne immédiatement

$$q(\vec{u}) = {}^tXAX = {}^tX'DX' = \sum_{i=1}^{n} \lambda_i x_i'^2. \qquad \square$$

Corollaire 2.15 *Soit q une forme quadratique sur E. Alors q est définie positive si et seulement si toutes ses valeurs propres sont strictement positives.*

Démonstration.

Notons $\lambda_1, \ldots, \lambda_n$ les valeurs propres de q. D'après le corollaire 2.14, il existe une base orthonormée $\mathcal{B}'$ relativement à laquelle on a

$$q(\vec{u}) = \sum_{i=1}^{n} \lambda_i x_i'^2.$$

— Supposons les valeurs propres de q strictement positives. Dans ce cas, on a clairement

$$q(\vec{u}) = 0 \Longrightarrow \sum_{i=1}^{n} \lambda_i x_i'^2 = 0 \Longrightarrow x'_1 = \ldots = x'_n = 0 \Longrightarrow \vec{u} = \vec{0}$$

et la forme quadratique q est définie positive.

— Si q admet deux valeurs propres de signes opposés, disons par exemple $\lambda_1 > 0$ et $\lambda_2 < 0$, alors le vecteur non nul $\vec{u}_0$ de coordonnées $(\sqrt{-\lambda_2}, \sqrt{\lambda_1}, 0, \ldots, 0)$ relativement à la base $\mathcal{B}'$ vérifie $q(\vec{u}_0) = 0$. Ceci montre que la forme quadratique q n'est pas définie positive.

□

Remarque.

Les vecteurs $\vec{u} \in E$ vérifiant $q(\vec{u}) = 0$ sont dits *isotropes* pour q. Une forme quadratique est définie positive ou définie négative si et seulement si elle n'admet pas de vecteur isotrope non nul. Par ailleurs, on vérifie aisément que si $\vec{u}$ est un vecteur isotrope pour q, alors il en est de même pour tous les vecteurs colinéaires à $\vec{u}$. En d'autres termes, si l'ensemble des vecteurs isotropes de q contient un vecteur non nul $\vec{u}$, il contient la droite vectorielle engendrée par $\vec{u}$ toute entière. On parle alors de *cône isotrope de q*.

Définition 2.16 *La signature d'une forme quadratique q est le couple (r, s) où r et s désignent respectivement le nombre de valeurs propres strictement positives et strictement négatives.*

Le corollaire 2.15 s'énonce en termes de signature comme suit :

q est définie positive si et seulement si sa signature est $(n, 0)$.

Application : réduction des quadriques A FAIRE

Complément : décomposition de Gauss, factorisation de Cholesky.

On se propose d'étudier sur deux exemples particuliers d'autres problèmes relatifs aux formes quadratiques.

1. *Décomposition de Gauss :* considérons la forme quadratique q définie sur le $\mathbb{R}$-espace vectoriel $E = \mathbb{R}^3$ muni de sa structure euclidienne canonique par
$$q(x, y, z) = 4x^2 - 4y^2 - 12z^2 + 6xy - 8xz + 14yz.$$
Sa matrice relativement à la base canonique est donnée par
$$A = \begin{pmatrix} 4 & 3 & -4 \\ 3 & -4 & 7 \\ -4 & 7 & 12 \end{pmatrix},$$
dont le polynôme caractéristique est $\Delta_A(\lambda) = -\lambda(\lambda^2 + 12\lambda - 90)$. On voit donc que A admet trois valeurs propres distinctes : 0, $\lambda = -6 + 3\sqrt{14}$ et $\mu = -6 - 3\sqrt{14}$. Les valeurs propres λ et μ sont de signes opposés. Le rang de q est 2 et sa signature est $(1, 1)$ (une valeur propre positive et une négative). On sait que la forme réduite de q est, relativement à une base orthonormée appropriée,
$$q(x, y, z) = \lambda x'^2 + \mu y'^2.$$

En écrivant

$$\begin{aligned}
4x^2+6xy-8xz &= (2x)^2+2(2x)\left(\frac{3}{2}\,y-2z\right)\\
&= \left(2x+\frac{3}{2}\,y-2z\right)^2-\left(\frac{3}{2}\,y-2z\right)^2\\
&= \left(2x+\frac{3}{2}\,y-2z\right)^2-\frac{9}{4}\,y^2+6yz-4z^2
\end{aligned}$$

on peut réécrire q sous la forme

$$\begin{aligned}
q(x,y,z) &= \left(2x+\frac{3}{2}\,y-2z\right)^2+\left(-\frac{9}{4}\,y^2+6yz-4z^2\right)-4y^2-12z^2+14yz\\
&= \left(2x+\frac{3}{2}\,y-2z\right)^2-\frac{25}{4}\,y^2+20yz-16z^2\\
&= \left(2x+\frac{3}{2}\,y-2z\right)^2-\left(\frac{5}{2}\,y-4z\right)^2.
\end{aligned}$$

Finalement,

$$q(x,y,z)=L_1(x,y,z)^2-L_2(y,z)^2$$

avec

$$L_1(x,y,z)=2x+\frac{3}{2}\,y-2z \quad \text{et} \quad L_2(y,z)=\frac{5}{2}\,y-4z.$$

Cette décomposition fait apparaître, tout comme $q(x,y,z)=\lambda x'^2+\mu y'^2$, la différence des carrés de deux formes linéaires non proportionnelles. Ceci est un fait général, il constitue la *loi d'inertie de Sylvester*[7] : toutes les décompositions en combinaisons linéaires de carrés de formes linéaires linéairement indépendantes contiennent les mêmes nombres de coefficients positifs, négatifs et nuls.

2. *Factorisation de Cholesky*[8] *:* considérons la forme quadratique q définie sur le $\mathbb{R}$-espace vectoriel $E=\mathbb{R}^3$ muni de sa structure euclidienne canonique par

$$q(x,y,z)=4x^2+\frac{17}{2}\,y^2+24z^2+6xy-8xz-26yz.$$

Sa matrice relativement à la base canonique est donnée par

$$A=\begin{pmatrix} 4 & 3 & -4\\ 3 & \frac{17}{2} & -13\\ -4 & 13 & 24\end{pmatrix},$$

dont le polynôme caractéristique est $\Delta_A(\lambda)=-\lambda^3+\frac{73}{4}\lambda^2-140\lambda+100$. Une étude des variations de la fonction Δ permet de voir que A admet trois valeurs propres distinctes strictement positives λ,μ,ν. Le rang de q est 3 et sa signature est $(3,0)$. La forme quadratique q est définie positive. On sait que la forme réduite de q est, relativement à une base orthonormée appropriée,

$$q(x,y,z)=\lambda x'^2+\mu y'^2+\nu z'^2.$$

7. Sylvester, James Joseph : mathématicien anglais (Londres 1814-Londres 1897)
8. Cholesky, André-Louis : ingénieur français (Montguyon 1875–Bagneux 1918

L'évaluation du polynôme caractéristique de A étant assez pénible, et n'ayant pas obtenu explicitement les valeurs propres de A, la méthode de Gauss semble plus judicieuse pour étudier q. Cette méthode conduit ici à

$$\begin{aligned}
q(x,y,z) &= \left(2x+\frac{3}{2}\,y-2z\right)^2+\left(-\frac{9}{4}\,y^2+6yz-4z^2\right)+\frac{17}{2}\,y^2+24z^2-26yz\\
&= \left(2x+\frac{3}{2}\,y-2z\right)^2+\frac{25}{4}\,y^2-20yz+20z^2\\
&= \left(2x+\frac{3}{2}\,y-2z\right)^2+\left(\frac{5}{2}\,y-4z\right)^2+(2z)^2.
\end{aligned}$$

Finalement,

$$q(x,y,z)=L_1(x,y,z)^2+L_2(y,z)^2+L_3(z)^2.$$

avec

$$L_1(x,y,z)=2x+\frac{3}{2}\,y-2z,\quad L_2(y,z)=\frac{5}{2}\,y-4z,\quad L_3(z)=2z.$$

Sous cette forme, il est aisé de voir que q est définie positive. En effet,

$$\begin{aligned}
q(x,y,z)=0 &\implies L_1(x,y,z)^2+L_2(y,z)^2+L_3(z)^2=0\\
&\implies L_1(x,y,z)=0,\ L_2(y,z)=0,\ L_3(z)=0\\
&\implies \left\{\begin{aligned} 2x+\tfrac{3}{2}\,y-2z&=0\\ \tfrac{5}{2}\,y-4z&=0\\ 2z&=0\end{aligned}\right.\\
&\implies x=y=z=0.
\end{aligned}$$

L'application $\sqrt{q}$ définit donc une norme et la forme polaire de q est un produit scalaire. La décomposition ainsi obtenue conduit à une factorisation intéressante de la matrice A : partant de

$$L_1(x,y,z)^2+L_2(y,z)^2+L_3(z)^2=\begin{pmatrix}L_1(x,y,z) & L_2(y,z) & L_3(z)\end{pmatrix}\begin{pmatrix}L_1(x,y,z)\\ L_2(y,z)\\ L_3(z)\end{pmatrix},$$

puis

$$\begin{aligned}
\begin{pmatrix}L_1(x,y,z) & L_2(y,z) & L_3(z)\end{pmatrix} &= \begin{pmatrix}x & y & z\end{pmatrix}\begin{pmatrix}2 & 0 & 0\\ \frac{3}{2} & \frac{5}{2} & 0\\ -2 & -4 & 2\end{pmatrix}\\
\begin{pmatrix}L_1(x,y,z)\\ L_2(y,z)\\ L_3(z)\end{pmatrix} &= \begin{pmatrix}2 & \frac{3}{2} & -2\\ 0 & \frac{5}{2} & -4\\ 0 & 0 & 2\end{pmatrix}\begin{pmatrix}x\\ y\\ z\end{pmatrix},
\end{aligned}$$

on tire

$$q(x,y,z)=\begin{pmatrix}x & y & z\end{pmatrix}\begin{pmatrix}2 & 0 & 0\\ \frac{3}{2} & \frac{5}{2} & 0\\ -2 & -4 & 2\end{pmatrix}\begin{pmatrix}2 & \frac{3}{2} & -2\\ 0 & \frac{5}{2} & -4\\ 0 & 0 & 2\end{pmatrix}\begin{pmatrix}x\\ y\\ z\end{pmatrix}.$$

D'autre part, A étant la matrice de q dans la base canonique, on a aussi

$$q(x,y,z)=\begin{pmatrix}x & y & z\end{pmatrix}A\begin{pmatrix}x\\ y\\ z\end{pmatrix}.$$

La confrontation des deux représentations de q ci-dessus fournit la factorisation

$$A = {}^tTT \quad \text{avec} \quad T = \begin{pmatrix} 2 & \frac{3}{2} & -2 \\ 0 & \frac{5}{2} & -4 \\ 0 & 0 & 2 \end{pmatrix}.$$

Ce type de factorisation est possible pour toute matrice carrée symétrique réelle définie positive.

Voici une application de la manipulation effectuée ci-dessus à la résolution d'un système linéaire d'équations. Considérons le système d'inconnues x, y, z et de paramètres a, b, c

$$(\mathrm{S}) \qquad \left\{ \begin{array}{rl} 4x + 3y - 4z = & a \\ 3x + \frac{17}{2}\,y - 13z = & b \\ -4x - 13y + 24z = & c \end{array} \right. .$$

Il s'écrit matriciellement $AX = B$ avec $X = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$ et $B = \begin{pmatrix} a \\ b \\ c \end{pmatrix}$. Remarquons tout d'abord que

$$\det(A) = \det({}^tTT) = \det({}^tT)\det(T) = \det(T)^2 \neq 0$$

puisque $\det(T) = 2 \times \frac{5}{2} \times 2 = 10$. Le système (S) est donc un système de Cramer et possède une solution unique (x, y, z). La factorisation de Cholesky ramène la résolution de (S) à la celle de deux systèmes triangulaires. En effet, en posant $X' = TX = \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix}$, on a $AX = B \Longleftrightarrow {}^tTX' = B$, ce qui donne le système

$$(\mathrm{S}') \qquad \left\{ \begin{array}{rl} 2x' \qquad\qquad = & a \\ \frac{3}{2}\,x' + \frac{5}{2}\,y' \qquad = & b \\ -2x' - 4y' + 2z' = & c \end{array} \right. .$$

Le système (S$'$) étant triangulaire se résout simplement :

$$\left\{ \begin{array}{rcl} x' & = & \frac{1}{2}\,a \\ y' & = & -\frac{3}{10}\,a + \frac{2}{5}\,b \\ z' & = & -\frac{1}{10}\,a + \frac{4}{5}\,b + \frac{1}{2}\,c \end{array} \right. .$$

Il reste à résoudre le deuxième système $TX = X'$ où X' est la solution de (S$'$) qui vient d'être trouvée :

$$(\mathrm{S}'') \qquad \left\{ \begin{array}{rl} 2x + \frac{3}{2}\,y - 2z = & x' \\ \frac{5}{2}\,y - 4z = & y' \\ 2z = & z' \end{array} \right. .$$

Remarquons que la résolution du système (S$'$) était équivalente à la recherche de la matrice inverse de ${}^tT$ et que la résolution du système (S$''$) est équivalente à la recherche de la matrice inverse de T. On a en fait obtenu

$$({}^tT)^{-1} = \begin{pmatrix} \frac{1}{2} & 0 & 0 \\ -\frac{3}{10} & \frac{2}{5} & 0 \\ -\frac{1}{10} & \frac{4}{5} & \frac{1}{2} \end{pmatrix}$$

d'où l'on tire

$$T^{-1} = \begin{pmatrix} \frac{1}{2} & -\frac{3}{10} & -\frac{1}{10} \\ 0 & \frac{2}{5} & \frac{4}{5} \\ 0 & 0 & \frac{1}{2} \end{pmatrix}$$

et enfin

$$X = T^{-1}X' = \begin{pmatrix} \frac{1}{2} & -\frac{3}{10} & -\frac{1}{10} \\ 0 & \frac{2}{5} & \frac{4}{5} \\ 0 & 0 & \frac{1}{2} \end{pmatrix} \begin{pmatrix} \frac{1}{2}\,a \\ -\frac{3}{10}\,a + \frac{2}{5}\,b \\ -\frac{1}{10}\,a + \frac{4}{5}\,b + \frac{1}{2}\,c \end{pmatrix}$$

c'est-à-dire

$$\begin{cases} x = \frac{7}{20}\,a - \frac{1}{5}\,b - \frac{1}{20}\,c \\ y = -\frac{1}{5}\,a + \frac{4}{5}\,b + \frac{2}{5}\,c\,. \\ z = -\frac{1}{20}\,a + \frac{2}{5}\,b + \frac{1}{4}\,c \end{cases}$$

3 Isométries vectorielles

Dans ce paragraphe, on se place dans un espace vectoriel euclidien (E, φ) de dimension n. On note N la norme associée au produit scalaire φ. On s'intéresse aux endomorphismes de E qui conservent la structure euclidienne.

3.1 Définition, propriétés

Définition 3.1 *Soit $f \in \mathcal{L}(E)$ un endomorphisme de E. L'application f est une isométrie vectorielle de E lorsqu'elle conserve la norme :*

$$\forall \vec{u} \in E,\ N(f(\vec{u})) = N(\vec{u}).$$

Voici quelques propriétés immédiates.

Proposition 3.2
1. *Les isométries vectorielles de E sont des automorphismes de E*
2. *Soit f et g des isométries vectorielles de E. Alors les endomorphismes $f \circ g$ et f^{-1} sont également des isométries vectorielles de E.*

DÉMONSTRATION.

1. Soit $\vec{u} \in E$ un vecteur tel que $f(\vec{u}) = 0$. On en déduit alors que $N(\vec{u}) = N(f(\vec{u})) = 0$, puis que $\vec{u} = \vec{0}$. Cela montre que $\operatorname{Ker} f = \{\vec{0}\}$, donc que f est injective et en fait bijective puisque f est un endomorphisme de E.
2. La deuxième assertion résulte aisément des égalités suivantes :

$$N((f \circ g)(\vec{u})) = N(g(\vec{u})) = N(\vec{u}) \quad \text{et} \quad N(f^{-1}(\vec{u})) = N(f(f^{-1}(\vec{u}))) = N(\vec{u}). \qquad \square$$

Proposition 3.3 *Soit $f \in \mathcal{L}(E)$. L'application f est une isométrie vectorielle de E si et seulement si elle conserve le produit scalaire :*

$$\forall \vec{u}, \vec{v} \in E,\ \varphi(f(\vec{u}), f(\vec{v})) = \varphi(\vec{u}, \vec{v}).$$

DÉMONSTRATION.

1. Supposons que φ conserve le produit scalaire. Alors, pour tout $\vec{u} \in E$,

$$N(f(\vec{u}))^2 = \varphi(f(\vec{u}), f(\vec{u})) = \varphi(\vec{u}, \vec{u}) = N(\vec{u})^2$$

et φ conserve la norme.

2. Supposons que φ conserve la norme. Partant de l'expression de φ au moyen de N suivante

$$\varphi(\vec{u}, \vec{v}) = \frac{1}{4}[N(\vec{u} + \vec{v})^2 - N(\vec{u} - \vec{v})^2],$$

on trouve successivement

$$\begin{aligned}
\varphi(f(\vec{u}), f(\vec{v})) &= \frac{1}{4}\,[N(f(\vec{u}) + f(\vec{v}))^2 - N(f(\vec{u}) - f(\vec{v}))^2] \\
&= \frac{1}{4}\,[N(f(\vec{u} + \vec{v}))^2 - N(f(\vec{u} - \vec{v}))^2] \\
&= \frac{1}{4}\,[N(\vec{u} + \vec{v})^2 - N(\vec{u} - \vec{v})^2] \\
&= \varphi(\vec{u}, \vec{v})
\end{aligned}$$

et φ conserve le produit scalaire. $\square$

Le théorème suivant énonce quelques propriétés caractéristiques des isométries vectorielles à l'aide de bases orthonormées.

Théorème 3.4 *Soit $f \in \mathcal{L}(E)$.*

1. *Si f est une isométrie vectorielle de E, alors pour toute base orthonormée $\mathcal{B}$ de E, l'image $f(\mathcal{B})$ est encore une base orthonormée de E.*
2. *S'il existe une base orthonormée $\mathcal{B}$ de E dont l'image $f(\mathcal{B})$ est une base orthonormée de E, alors l'endomorphisme f est une isométrie vectorielle de E.*
3. *Soit $\mathcal{B}$ une base orthonormée de E et $A = \mathrm{Mat}(f; \mathcal{B})$ la matrice de l'endormorphisme f relativement à la base $\mathcal{B}$. L'application f est une isométrie vectorielle de E si et seulement si la matrice A est orthogonale.*

Démonstration.

1. Supposons que f soit une isométrie vectorielle de E. Soit $\mathcal{B} = \{\vec{e}_1, \ldots, \vec{e}_n\}$ une base orthonormée de E. Vérifions que le système de vecteurs $f(\mathcal{B}) = \{f(\vec{e}_1), \ldots, f(\vec{e}_n)\}$ est une base orthonormée de E. On a
$$\varphi(f(\vec{e}_i), f(\vec{e}_j)) = \varphi(\vec{e}_i, \vec{e}_j) = \begin{cases} 0 & \text{si } i \neq j \\ 1 & \text{si } i = j \end{cases}$$
et alors $f(\mathcal{B})$ est un système orthonormé donc libre. Comme il contient n vecteurs, c'est une base orthonormée de E.
2. Soit $\mathcal{B} = \{\vec{e}_1, \ldots, \vec{e}_n\}$ une base orthonormée de E telle que le système de vecteurs $f(\mathcal{B}) = \{f(\vec{e}_1), \ldots, f(\vec{e}_n)\}$ soit une base orthonormée de E. Pour un vecteur générique $\vec{u} = \sum_{i=1}^{n} x_i \vec{e}_i$ de E, on a d'une part
$$N(\vec{u}) = \sqrt{\sum_{i=1}^{n} x_i^2},$$
et d'autre part
$$f(\vec{u}) = \sum_{i=1}^{n} x_i f(\vec{e}_i) \quad \text{puis} \quad N(f(\vec{u})) = \sqrt{\sum_{i=1}^{n} x_i^2}.$$
Ainsi $\forall \vec{u} \in E,\ N(f(\vec{u})) = N(\vec{u})$, donc f est une isométrie vectorielle.
3. Soit $A = \mathrm{Mat}(f; \mathcal{B}) = (a_{ij})_{1 \leqslant i,j \leqslant n}$ la matrice d'un endomorphisme f relativement à une base orthonormée $\mathcal{B}$ de E. Rappelons que les a_{ij}, $1 \leqslant i, j \leqslant n$, sont définis par les relations $f(\vec{e}_j) = \sum_{i=1}^{n} a_{ij} \vec{e}_i$, $1 \leqslant j \leqslant n$. Calculons le produit ${}^tAA$:
$${}^tAA = \left(\sum_{k=1}^{n} a_{ki} a_{kj} \right)_{1 \leqslant i,j \leqslant n}.$$
Observant que la somme $\sum_{k=1}^{n} a_{ki} a_{kj}$ n'est autre que le produit scalaire de $f(\vec{e}_i)$ par $f(\vec{e}_j)$:
$$\sum_{k=1}^{n} a_{ki} a_{kj} = \varphi(f(\vec{e}_i), f(\vec{e}_j)),$$
on obtient
$${}^tAA = (\varphi(f(\vec{e}_i), f(\vec{e}_j)))_{1 \leqslant i,j \leqslant n}.$$
Par suite, il vient les équivalences suivantes :
$$\begin{aligned} f \text{ est une isométrie vectorielle} &\iff \text{le système } \{f(\vec{e}_1), \ldots, f(\vec{e}_n)\} \text{ est une base de } E \\ &\iff \varphi(f(\vec{e}_i), f(\vec{e}_j)) = \varphi(\vec{e}_i, \vec{e}_j) = \begin{cases} 0 & \text{si } i \neq j \\ 1 & \text{si } i = j \end{cases} \\ &\iff {}^tAA = I_n \\ &\iff A \text{ est une matrice orthogonale.} \end{aligned}$$

□

Si $A \in \mathcal{M}_n(\mathbb{K})$ est une matrice orthogonale, on a

$$\det(A)^2 = \det({}^tA)\det(A) = \det({}^tAA) = \det(I_n) = 1$$

d'où

$$\det(A) = \pm 1.$$

Cette remarque conduit à faire une distinction entre les isométries vectorielles possédant un déterminant égal à 1 et celles possédant un déterminant égal à -1.

Définition 3.5 *Soit f une isométrie vectorielle de E.*
— *Si* $\det(f) = 1$*, l'application f est une* isométrie directe *(ou* positive*) ;*
— *Si* $\det(f) = -1$*, l'application f est une* isométrie indirecte *(ou* rétrograde *ou* négative*).*

Définition 3.6 *Soit $\mathcal{B}$ et $\mathcal{B}'$ deux bases orthonormées de E et f l'isométrie vectorielle envoyant $\mathcal{B}$ sur $\mathcal{B}'$.*

1. *Les bases $\mathcal{B}$ et $\mathcal{B}'$ ont* même orientation *(resp.* des orientations contraires*) lorsque f est une isométrie directe (resp. indirecte).*
2. *On oriente l'espace vectoriel euclidien (E, φ) en décrétant qu'une base orthonormée $\mathcal{B}$ de E choisie au préalable est* positivement orientée *; elle servira d'orientation de référence. Une autre base orthonormée de E est* directe *(resp.* indirecte*) si elle a même orientation que $\mathcal{B}$ (resp. une orientation contraire à celle de $\mathcal{B}$).*

Exemple.

Dans un espace vectoriel euclidien E de dimension 3, le produit vectoriel permet de construire facilement une base orthonormée directe à partir de deux vecteurs unitaires orthogonaux $\vec{u}, \vec{v} \in E$: le système de vecteurs $(\vec{u}, \vec{v}, \vec{u} \wedge \vec{v})$ est une base orthonormée directe de E, alors que $(\vec{u}, \vec{v}, -\vec{u} \wedge \vec{v})$ est une base orthonormée indirecte de E

3.2 Classification des isométries vectorielles en dimension $\leqslant 3$

Avant de décrire complètement les isométries vectorielles d'un $\mathbb{R}$-espace vectoriel de dimension $\leqslant 3$, faisons quelques remarques préliminaires.

Proposition 3.7 *Une isométrie vectorielle d'un espace vectoriel de dimension impaire admet au moins 1 ou -1 pour valeur propre.*

Démonstration.

Le polynôme Δ_A étant de degré impair, il admet nécessairement au moins une racine réelle (c'est une simple conséquence du théorème des valeurs intermédiaires : la fonction Δ_A est continue ayant des limites infinies de signes opposés à l'infini). De plus, si λ est une valeur propre de f et $\vec{u} \in E$ un vecteur propre associé à λ, on a $N(\vec{u}) = N(f(\vec{u})) = N(\lambda\vec{u}) = |\lambda| N(\vec{u})$ d'où l'on déduit $|\lambda| = 1$. Donc les seules valeurs propres réelles possibles sont 1 et -1. On peut en fait démontrer en se plaçant sur $\mathbb{C}$ que toutes les valeurs propres (complexes) de f sont de module 1. □

Proposition 3.8 *1. Si F et G sont des sous-espaces vectoriels orthogonaux de E, alors il en est de même des sous-espaces vectoriels $f(F)$ et $f(G)$.*

2. En particulier, si F est un sous-espace vectoriel de E stable par f, alors il en est de même du sous-espace vectoriel $f(F)$. De plus, la restriction de f a F induit une isométrie vectorielle de F : $f|_F : F \longrightarrow F$.

Démonstration.

1. Soit $(\vec{u},\vec{v}) \in F\times G$. Puisque f est une isométrie et $F \perp G$, on a $\varphi(f(\vec{u}), f(\vec{v})) = \varphi(\vec{u},\vec{v}) = 0$, ce qui implique
$$\forall(\vec{u},\vec{v}) \in F \times G,\ f(\vec{u}) \perp f(\vec{v}),$$
soit encore
$$f(F) \perp f(G).$$
2. Soit F un sous-espace vectoriel stable par f. Cela signifie que $f(F) \subset F$. En fait, on a ici plus particulièrement $f(F) = F$; ceci est dû à l'injectivité de l'isométrie f qui assure l'égalité des dimensions de $f(F)$ et F. En appliquant alors le résultat précédent à $G = F^{\perp}$, on voit que les sous-espaces vectoriels $f(F^{\perp})$ et $f(F) = F$ sont orthogonaux et donc $f(F^{\perp}) \subset F^{\perp}$. Par le même argument que celui évoqué ci-dessus, on a plus précisément
$$f(F^{\perp}) = F^{\perp}.$$
On peut donc considérer l'application restreinte $f|_F : \begin{array}{ccc} F & \longrightarrow & F \\ \vec{u} & \longmapsto & f(\vec{u}) \end{array}$ qui est évidemment une isométrie vectorielle de F. □

Un exemple de sous-espace vectoriel stable par f est un sous-espace propre de f. Notamment, lorsque f admet pour valeur propre 1 (resp. -1), le sous-espace propre associé à 1 : $\mathrm{Inv}\, f$ (resp. -1 : $\mathrm{Ker}(f + \mathrm{id}_E)$) est un sous-espace vectoriel stable par f, et son othogonal aussi. L'avantage de ce résultat est qu'il permet d'étendre des propriétés d'isométries d'un espace vectoriel de dimension inférieure à celle de E aux isométries de E.

Remarque.

Il ne faut pas confondre un sous-ensemble de E stable par f et le sous-ensemble des vecteurs de E invariants par f. En effet, dire que F est stable par f (on dit aussi que F est *globalement invariant par f*) signifie que
$$\forall \vec{u} \in F,\ f(\vec{u}) \in F,$$
mais on n'a pas nécessairement $f(\vec{u}) = \vec{u}$, c'est-à-dire le vecteur $\vec{u}$ n'est pas nécessairement par f.

3.2.1 Cas de la dimension 1

Supposons que E soit une droite vectorielle engendrée par un vecteur unitaire $\vec{i}$, *i.e.* $E = \mathrm{vect}\{\vec{i}\}$ avec $N(\vec{i}) = 1$. Un endomorphisme $f \in \mathcal{L}(E)$ est entièrement déterminé par $f(\vec{i}) = a\vec{i}$

pour un certain réel a. Le scalaire a est en fait une valeur propre de f, et d'après la proposition 3.7, si f est une isométrie vectorielle, a ne peut égaler que 1 ou -1 :

$$f = \pm \mathrm{id}_E.$$

Ainsi, les seules isométries vectorielles de E sont l'identité de E et son opposée.

3.2.2 Cas de la dimension 2

Supposons que E soit un plan vectoriel engendré par deux vecteurs orthogonaux et unitaires $\vec{i}$, $\vec{j}$, *i.e.* $E = \mathrm{vect}\{\vec{i}, \vec{j}\}$ où $\{\vec{i}, \vec{j}\}$ est une base orthonormée que l'on considérera directe. Un endomorphisme $f \in \mathcal{L}(E)$ est déterminée par sa matrice $A = \mathrm{Mat}(f; \mathcal{B}) = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Cherchons à quelles conditions f est une isométrie vectorielle, ou encore, en termes matriciels, A est une matrice orthogonale. Effectuons tout d'abord un petit calcul :

$${}^tAA = \begin{pmatrix} a^2+b^2 & ac+bd \\ ac+bd & c^2+d^2 \end{pmatrix}.$$

On a alors les équivalences

$$\begin{aligned} f \text{ est une isométrie vectorielle} &\iff a^2+b^2 = c^2+d^2 = 1 \quad \text{et} \quad ac+bd = 0 \\ &\iff \exists \theta, \varphi \in [0, 2\pi[, \ \left\{ \begin{array}{l} (a,b) = (\cos\theta, \sin\theta) \\ (c,d) = (\cos\varphi, \sin\varphi) \\ \cos\theta\cos\varphi + \sin\theta\sin\varphi = 0 \end{array} \right. . \end{aligned}$$

Or $\cos\theta\cos\varphi + \sin\theta\sin\varphi = \cos(\theta - \varphi)$, donc la dernière équation ci-dessus impose à θ et φ la condition $\varphi = \theta \pm \dfrac{\pi}{2}\,[\mathrm{mod.}\, 2\pi]$ et alors $(c,d) = \pm(-\sin\theta, \cos\theta)$. On tombe sur deux familles de matrices orthogonales (et donc d'isométries vectorielles).

(a) *Première famille :* elle correspond aux matrices de la forme

$$A = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$$

L'isométrie vectorielle f est la *rotation vectorielle d'angle* θ (voir Fig. 6.2).

REMARQUES.

1. Notons que l'on a $\det(A) = 1$, donc les rotations sont des isométries vectorielles directes.
2. Le polynôme caractéristique de A est donné par $\Delta_A(\lambda) = \lambda^2 - 2\lambda\cos\theta + 1$ et A admet les valeurs propres complexes conjuguées $e^{i\theta}$ et $e^{-i\theta}$. La matrice A est diagonalisable sur $\mathbb{C}$ mais ne l'est pas sur $\mathbb{R}$ excepté les cas triviaux où $\theta = 0\,[\mathrm{mod.}\, 2\pi]$ et $\theta = \pi\,[\mathrm{mod.}\, 2\pi]$ correspondant à $\pm\mathrm{id}_E$.
3. En posant $A_\theta = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$ pour tenir compte de la dépendance du paramètre θ, on a les relations $A_\theta A_\varphi = A_{\theta+\varphi}$ et $A_\theta^{-1} = A_{-\theta}$ qui montrent que l'ensemble des matrices A_θ, $\theta \in \mathbb{R}$, est un sous-groupe commutatif du groupe multiplicatif des matrices carrées d'ordre 2 inversibles.

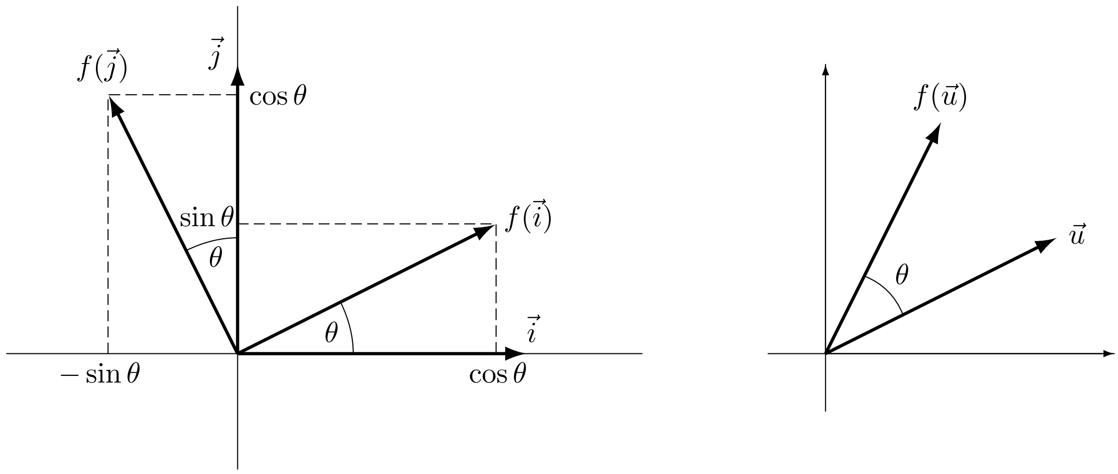

FIGURE 6.2 – Rotation vectorielle en dimension 2

(b) *Deuxième famille :* elle correspond aux matrices

$$A = \begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix}.$$

On a $A^2 = I_2$, ce qui montre que l'isométrie f est une symétrie vectorielle de E. Ses sous-espaces vectoriels caractéristiques sont les suivants :

— l'ensemble des vecteurs invariants de f, $\operatorname{Inv} f = \operatorname{Ker}(f - \mathrm{id}_E)$, qui est caractérisé par le système

$$\begin{cases} (\cos\theta - 1)\,x + \sin\theta\, y = 0 \\ \sin\theta\, x + (\cos\theta + 1)\, y = 0 \end{cases}.$$

Se souvenant des formules trigonométriques

$$1 + \cos\theta = 2\cos^2\frac{\theta}{2},\ 1 - \cos\theta = 2\sin^2\frac{\theta}{2},\ \sin\theta = 2\sin\frac{\theta}{2}\cos\frac{\theta}{2},$$

le système précédent est équivalent à l'équation

$$x\sin\frac{\theta}{2} - y\cos\frac{\theta}{2} = 0.$$

Il s'agit de la droite vectorielle d'angle polaire $\theta/2$:

$$\operatorname{Inv} f = \mathcal{D} = \operatorname{vect}\Big\{\cos\frac{\theta}{2}\,\vec{i} + \sin\frac{\theta}{2}\,\vec{j}\Big\}\,;$$

— l'ensemble des vecteurs anti-invariants de f, $\operatorname{Ker}(f + \mathrm{id}_E)$, qui est caractérisé, de manière analogue, par l'équation

$$x\cos\frac{\theta}{2} + y\sin\frac{\theta}{2} = 0.$$

C'est la droite vectorielle d'angle polaire $(\theta + \pi)/2$:

$$\Delta = \operatorname{vect}\Big\{-\sin\frac{\theta}{2}\,\vec{i} + \cos\frac{\theta}{2}\,\vec{j}\Big\}.$$

Elle est orthogonale à $\mathcal{D}$, ce qui était prévisible en vertu de la proposition 3.8. Un autre argument en faveur de l'orthogonalité des droites $\mathcal{D}$ et Δ est que la matrice A est symétrique réelle, donc diagonalisable sur $\mathbb{R}$ (de valeurs propres propres 1 et -1) et ses sous-espaces propres sont orthogonaux : $\mathcal{D} \perp \Delta$.

En conclusion, l'isométrie f est la *symétrie vectorielle orthogonale par rapport à la droite d'angle polaire* $\theta/2$ (voir Fig. 6.3).

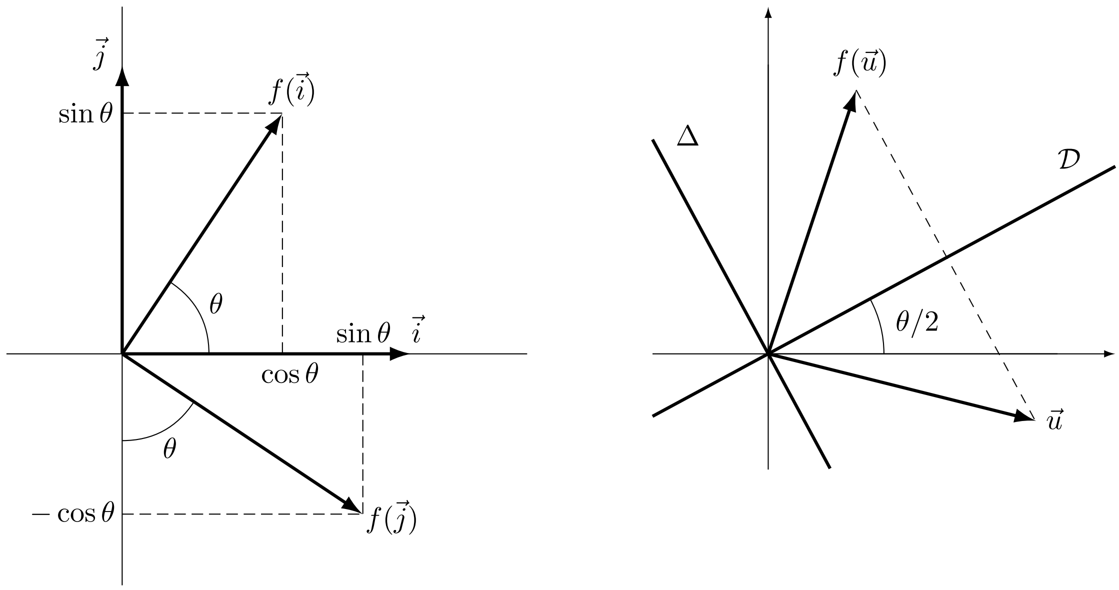

FIGURE 6.3 – Symétrie vectorielle en dimension 2

3.2.3 Cas de la dimension 3

Soit E un $\mathbb{R}$-espace vectoriel de dimension 3 et $\mathcal{B} = \{\vec{i}, \vec{j}, \vec{k}\}$ une base orthonormée que l'on considérera directe. La classification des isométries de E proposée ci-dessous est faite selon les différentes valeurs que peut prendre la dimension du sous-espace vectoriel des vecteurs invariants par f.

(a) *Premier cas :* $\dim(\operatorname{Inv} f) = 3$.

On a $\operatorname{Inv} f = E$ et alors $f = \mathrm{id}_E$.

(b) *Deuxième cas :* $\dim(\operatorname{Inv} f) = 2$.

Le sous-espace vectoriel $\mathcal{P} = \operatorname{Inv} f$ est un plan vectoriel stable par f et d'après la proposition 3.8 la droite vectorielle $\mathcal{D} = \mathcal{P}^\perp$ est également stable par f. On obtient alors par restriction de f à $\mathcal{D}$ une isométrie de la droite $\mathcal{D}$. En se référant au cas précédent, on n'a que les deux possibilités $f|_\mathcal{D} = \pm\mathrm{id}_\mathcal{D}$. Le cas $f|_\mathcal{D} = \mathrm{id}_\mathcal{D}$ est à écarter car, étant donné que $f|_\mathcal{P} = \mathrm{id}_\mathcal{P}$ et que $\mathcal{P} \oplus \mathcal{D} = E$, on aurait $f = \mathrm{id}_E$ ce qui est contraire à l'hypothèse $\dim(\operatorname{Inv} f) = 2$. On a donc $f|_\mathcal{D} = -\mathrm{id}_\mathcal{D}$, puis $\mathcal{D} = \operatorname{Ker}(f + \mathrm{id}_E)$; en d'autre termes, les sous-espaces vectoriels supplémentaires $\mathcal{P}$ et $\mathcal{D}$ sont respectivement les ensembles des vecteurs invariants et anti-invariants de E. On reconnaît donc en f une symétrie vectorielle de E par rapport à $\mathcal{P}$ et de direction $\mathcal{D}$: c'est la *symétrie vectorielle orthogonale par rapport à* $\mathcal{P}$ (voir Fig. 6.4). Relativement à une base adaptée à la décomposition en somme directe $E = \mathcal{P} \oplus \mathcal{D}$, la matrice de f s'écrit

$$\mathrm{Mat}(p; \mathcal{B}) = \left(\begin{array}{cc|c} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \hline 0 & 0 & -1 \end{array}\right).$$

Notons que $\det(f) = -1$ et donc que f est une isométrie indirecte.

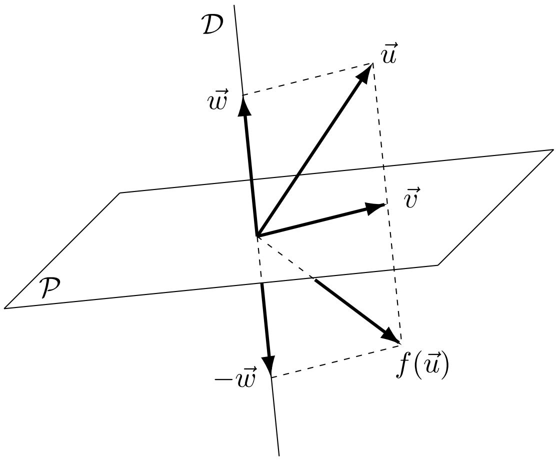

FIGURE 6.4 – Symétrie vectorielle orthogonale en dimension 3

(c) *Troisième cas :* $\dim(\mathrm{Inv}\, f) = 1$.
Le sous-espace vectoriel $\mathcal{D} = \mathrm{Inv}\, f$ est une droite vectorielle stable par f et d'après la proposition 3.8 le plan vectoriel $\mathcal{P} = \mathcal{D}^{\perp}$ est également stable par f. La restriction de f à $\mathcal{P}$ est alors une isométrie du plan $\mathcal{P}$ sans vecteur invariant non nul. En se référant au cas précédent, $f|_{\mathcal{P}}$ est une rotation vectorielle de P. Notons θ son angle. L'isométrie f est la *rotation vectorielle de E d'axe $\mathcal{D}$ et d'angle θ*. Relativement à une base adaptée à la décomposition en somme directe $E = \mathcal{P} \oplus \mathcal{D}$, la matrice de f s'écrit

$$\mathrm{Mat}(p; \mathcal{B}) = \left(\begin{array}{cc|c} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ \hline 0 & 0 & 1 \end{array}\right).$$

La construction géométrique de l'image d'un vecteur $\vec{u} \in E$ s'effectue comme suit. Soit $\vec{n}$ un vecteur unitaire de $\mathcal{D}$. On décompose $\vec{u}$ en la somme $\vec{u} = \vec{v} + \vec{w}$ avec $(\vec{v}, \vec{w}) \in \mathcal{P} \times \mathcal{D}$. Les vecteurs $\vec{v}$ et $\vec{w}$ sont les projections orthogonales respectives de $\vec{u}$ sur $\mathcal{P}$ et $\mathcal{D}$. Elles sont données par

$$\vec{w} = \vec{u}.\vec{n} \quad \text{et} \quad \vec{v} = \vec{u} - (\vec{u}.\vec{n})\vec{n}.$$

On a ensuite $f(\vec{u}) = f(\vec{v}) + f(\vec{w})$. D'une part $f(\vec{w}) = \vec{w}$ puisque $\vec{w} \in \mathrm{Inv}\, f$. D'autre part, $f(\vec{v})$ est l'image du vecteur $\vec{v} \in \mathcal{P}$ par la rotation de $\mathcal{P}$ d'angle θ, et alors $f(\vec{u}) = f(\vec{v}) + \vec{w}$ (voir Fig. 6.5). Plus précisément, notant que le système de vecteurs $(\frac{1}{\|\vec{v}\|}\, \vec{v}, \frac{1}{\|\vec{v}\|}\, \vec{n} \wedge \vec{v}, \vec{n})$ est une base orthormée directe de E, on dispose d'une base orthormée directe $(\frac{1}{\|\vec{v}\|}\, \vec{v}, \frac{1}{\|\vec{v}\|}\, \vec{n} \wedge \vec{v})$ de $\mathcal{P}$ relativement à laquelle on peut écrire (voir Fig. 6.6)

$$f(\vec{v}) = (\cos\theta)\, \vec{v} + (\sin\theta)\, \vec{n} \wedge \vec{v}.$$

En conséquence, on trouve

$$f(\vec{u}) = (\cos\theta)\, [\vec{u} - (\vec{u}.\vec{n})\vec{n}] + (\sin\theta)\, \vec{n} \wedge \vec{v} + \vec{u}.\vec{n},$$

ou encore, puisque $\vec{n} \wedge \vec{v} = \vec{n} \wedge \vec{u}$,

$$\boxed{f(\vec{u}) = (\cos\theta)\,\vec{u} + (\sin\theta)\,\vec{n} \wedge \vec{u} + (1 - \cos\theta)\,(\vec{u}.\vec{n})\,\vec{n}.}$$

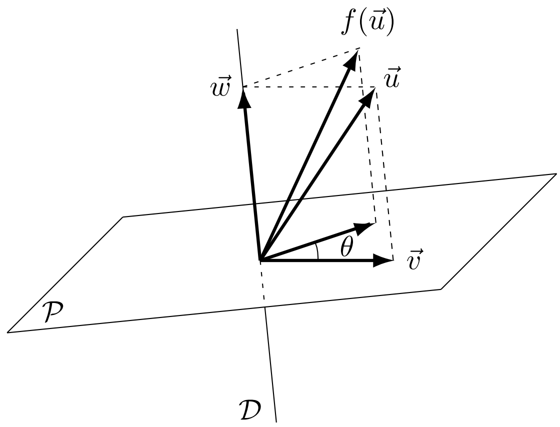

Figure 6.5 – Rotation vectorielle en dimension 3

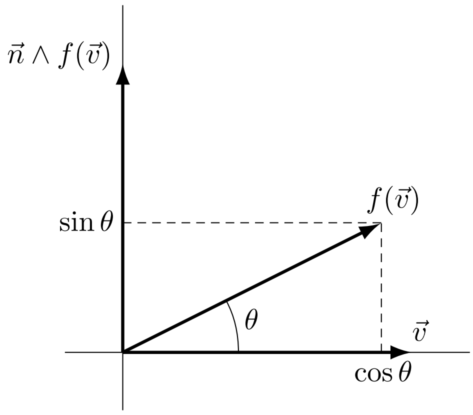

Figure 6.6 – Rotation vectorielle en dimension 3 (détail)

Remarques.

1. Notons que $\det(f) = 1$ et donc que f est une isométrie directe.
2. Une rotation vectorielle est caractérisée par son axe qui est l'ensemble des vecteurs invariants et pas son angle. La détermination de l'angle peut se faire comme suit. Observons tout d'abord que $\operatorname{tr}(f) = 2\cos\theta + 1$, ce qui permet de trouver immédiatement

$$\cos\theta = \frac{1}{2}\,[\operatorname{tr}(A) - 1].$$

Le cosinus à lui seul ne suffit pas à déterminer l'angle θ puisqu'il subsiste une ambiguïté au niveau du signe. Le sinus permet de pallier à ce problème. En effet, en choisissant un vecteur $\vec{u} \in \mathcal{P}$ unitaire, puis en posant $\vec{v} = \vec{n} \wedge \vec{u}$, le système de vecteurs $(\vec{u}, \vec{v}, \vec{n})$

constitue une base orthonormée directe de E. De plus, $f(\vec{u}) \in \mathcal{P}$ et $f(\vec{u}) = (\cos\theta)\,\vec{u} + (\sin\theta)\,\vec{v}$, donc

$$\vec{u} \wedge f(\vec{u}) = (\sin\theta)\,\vec{n},$$

identité de laquelle on tire $\sin\theta$ comme coefficient de colinéarité entre les vecteurs $\vec{u} \wedge f(\vec{u})$ et $\vec{n}$. En fait,

$$\sin\theta = \varphi(\vec{u} \wedge f(\vec{u}), \vec{n}) = \det_{\mathcal{B}}(\vec{u}, f(\vec{u}), \vec{n}).$$

Enfin, la connaissance de $\cos\theta$ et de $\sin\theta$ permet de trouver l'angle θ.

(d) *Quatrième cas :* $\dim(\operatorname{Inv} f) = 0$.
On a $\operatorname{Inv} f = \{\vec{0}\}$. Cela signifie que 1 n'est pas une valeur propre de f et dans ce cas, d'après la proposition 3.7, -1 est une valeur propre de f. Sa multiplicité est nécessairement impaire (1 ou 3), sinon elle vaudrait 2, le polynôme caractéristique Δ_f de f admettrait une racine complexe non réelle λ. Le polynôme Δ_f étant à coefficients réels, il admettrait également $\overline{\lambda}$ comme autre racine, ce qui ferait un total d'au moins quatre racines (en comptant les multiplicités) et conduirait à une absurdité puisque $\deg(D_f) = 3$. Deux situations sont donc envisageables :
— ou bien, la valeur propre -1 est triple et dans ce cas $f = -\mathrm{id}_E$;
— ou bien, la valeur propre -1 est simple et alors le sous-espace propre associé $\mathcal{D} = \operatorname{Ker}(f + \mathrm{id}_E)$ est une droite vectorielle. Le sous-espace vectoriel $\mathcal{D}$ étant stable par f, la proposition 3.8 assure le plan vectoriel $\mathcal{P} = \mathcal{D}^\perp$ est également stable par f. La restriction de f à $\mathcal{P}$ est alors une isométrie du plan $\mathcal{P}$ sans vecteur invariant non nul, $f|_{\mathcal{P}}$ est une rotation vectorielle de P. Notons θ son angle. Relativement à une base adaptée à la décomposition en somme directe $E = \mathcal{P} \oplus \mathcal{D}$, la matrice de f s'écrit

$$\operatorname{Mat}(p; \mathcal{B}) = \left(\begin{array}{cc|c} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ \hline 0 & 0 & -1 \end{array}\right).$$

La factorisation élémentaire

$$\begin{aligned} \left(\begin{array}{cc|c} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ \hline 0 & 0 & -1 \end{array}\right) &= \left(\begin{array}{cc|c} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ \hline 0 & 0 & 1 \end{array}\right) \left(\begin{array}{cc|c} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \hline 0 & 0 & -1 \end{array}\right) \\ &= \left(\begin{array}{cc|c} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \hline 0 & 0 & -1 \end{array}\right) \left(\begin{array}{cc|c} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ \hline 0 & 0 & 1 \end{array}\right) \end{aligned}$$

conduit à interpréter l'isométrie f comme étant la composée de la rotation vectorielle $r_{D,\theta}$ de E d'axe $\mathcal{D}$ et d'angle θ et la symétrie vectorielle orthogonale s_P de E par rapport à $\mathcal{P}$, la composition étant commutative :

$$f = r_{D,\theta} \circ s_P = s_P \circ r_{D,\theta}.$$

La construction géométrique de l'image d'un vecteur $\vec{u} \in E$ se fait comme suit. En décomposant $\vec{u}$ en la somme $\vec{u} = \vec{v} + \vec{w}$ avec $(\vec{v}, \vec{w}) \in \mathcal{P} \times \mathcal{D}$, on obtient $f(\vec{u}) = f(\vec{v}) + f(\vec{w})$. D'une part $f(\vec{w}) = -\vec{w}$ puisque $\vec{w} \in \operatorname{Ker}(f + \mathrm{id}_E)$. D'autre part, $f(\vec{v})$ est l'image du vecteur $\vec{v} \in \mathcal{P}$ par la rotation de $\mathcal{P}$ d'angle θ. D'où $f(\vec{u}) = f(\vec{v}) - \vec{w}$ (voir Fig. 6.7).

Remarques.

1. Notons que $\det(f) = -1$ et donc que f est une isométrie indirecte.

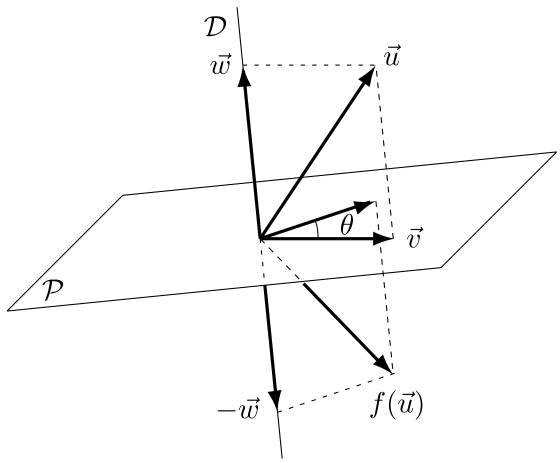

FIGURE 6.7 – Isométrie vectorielle sans invariant non nul en dimension 3

2. Les éléments caractéristiques de f son l'axe $\mathcal{D}$ et l'angle θ de la rotation et le plan $\mathcal{P}$ de la symétrie qui interviennent dans la décomposition ci-dessus. L'axe $\mathcal{D}$ est l'ensemble des vecteurs anti-invariants, le plan $\mathcal{P}$ est l'orthogonal. La détermination de l'angle peut se faire comme dans le cas précédent. Remarquons tout d'abord que $\operatorname{tr}(f) = 2\cos\theta - 1$, ce qui permet de trouver immédiatement

$$\cos\theta = \frac{1}{2}\,[\operatorname{tr}(A) + 1].$$

Le sinus de θ peut se calculer en choisissant un vecteur $\vec{u} \in \mathcal{P}$, puis en posant $\vec{v} = \vec{n} \wedge \vec{u}$. Le système de vecteurs $(\vec{u}, \vec{v}, \vec{n})$ constitue une base orthonormée directe de E. De plus, $f(\vec{u}) \in \mathcal{P}$ et $f(\vec{u}) = (\cos\theta)\,\vec{u} + (\sin\theta)\,\vec{v}$, donc

$$\vec{u} \wedge f(\vec{u}) = (\sin\theta)\,\vec{n}$$

d'où l'on tire $\sin\theta$, puis θ.

4 Exercices sur l'algèbre bilinéaire

Exercice 1 Caractériser géométriquement les isométries du $\mathbb{R}$-espace vectoriel euclidien $\mathbb{R}^3$ dont les matrices relativement à une base orthonormée directe sont données par

$$\text{a)}\ \frac{1}{3}\begin{pmatrix} 2 & 1 & 2 \\ -2 & 2 & 1 \\ -1 & -2 & 2 \end{pmatrix} \quad \text{b)}\ \frac{1}{3}\begin{pmatrix} 1 & -2 & -2 \\ -2 & 1 & -2 \\ -2 & -2 & 1 \end{pmatrix} \quad \text{c)}\ \begin{pmatrix} 0 & -1 & 0 \\ 0 & 0 & -1 \\ -1 & 0 & 0 \end{pmatrix}$$

Exercice 2 Soit les matrices définies par

$$A = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} \ \text{ et } \ B = \begin{pmatrix} -1 & 2 & 2 \\ 2 & -1 & 2 \\ 2 & 2 & -1 \end{pmatrix}.$$

1. En interprétant A comme la matrice d'une rotation, trouver une matrice de rotation $M \in \mathcal{M}_2(\mathbb{R})$ telle que $M^2 = A$.
2. En déduire une matrice $N \in \mathcal{M}_3(\mathbb{R})$ telle que $N^2 = B$. On pourra diagonaliser la matrice B.

Exercice 3

1. Déterminer la nature de la conique d'équation $P(x, y) = 0$, puis la représenter dans un repère orthonormé dans chacun des cas suivants :
   (a). $P(x, y) = 13x^2 - 32xy + 37y^2 - 2x + 14y - 5$;
   (b). $P(x, y) = xy + 3x - 5y - 4$;
   (c). $P(x, y) = 4x^2 - 4xy + y^2 - 2x - 14y + 7$.
2. Déterminer pour chacun des cas ci-dessus le gradient $\overrightarrow{\text{grad}}\, P(x, y)$, puis résoudre l'équation $\overrightarrow{\text{grad}}\, P(x, y) = \vec{0}$.

Exercice 4 On se place dans le $\mathbb{R}$-espace vectoriel E des fonctions continues de $[0, 1]$ dans $\mathbb{R}$ et l'on considère le sous-ensemble

$$F = \{f \in E : \int_0^1 f(x)^2 \,\frac{dx}{x} \text{ est convergente}\}.$$

1. Montrer que F est un sous-espace vectoriel de E. On pourra utiliser l'inégalité élémentaire $2|ab| \leqslant a^2 + b^2$.
2. On pose pour $(f, g) \in F \times F$, $\varphi(f, g) = \displaystyle\int_0^1 f(x)g(x)\,\frac{dx}{x}$. Montrer que l'application φ définit un produit scalaire sur F.
3. Soit $(P_n)_{n\geqslant 1}$ la suite de polynômes définie par $P_n(x) = x^n$.
   (a). Trouver les trois premiers polynômes Q_1, Q_2, Q_3 de la suite de polynômes orthogonaux $(Q_n)_{n\geqslant 1}$ construite à partir de la suite de polynômes $(P_n)_{n\geqslant 1}$ par le procédé de Gram-Schmidt :
   $$Q_1 = P_1, \quad Q_2 = P_2 - \lambda Q_1, \quad Q_3 = P_3 - \mu Q_1 - \nu Q_2.$$

(b). En déduire les trois premiers polynômes $\tilde{Q}_1$, $\tilde{Q}_2$, $\tilde{Q}_3$ de la suite orthonormée $(\tilde{Q}_n)_{n\geqslant 1}$ générée par la suite de polynômes $(P_n)_{n\geqslant 1}$.

Exercice 5 On se place dans le $\mathbb{R}$-espace vectoriel euclidien $E = \mathbb{R}^2$ muni de son produit scalaire canonique et on considère, pour tout $a \in \mathbb{R}$, la forme quadratique

$$q_a(x, y) = (\operatorname{ch} a)\, x^2 + 2\,(\operatorname{sh} a)\, xy + (\operatorname{ch} a)\, y^2.$$

Soit φ_a la forme bilinéaire symétrique associée à la forme quadratique q_a et A_a sa matrice dans la base canonique de E.

1. Écrire la matrice A_a ainsi que la forme bilinéaire φ_a.
2. Montrer que φ_a définit un produit scalaire sur E.

3.(a). Montrer que l'ensemble des matrices A_a, $a \in \mathbb{R}$, est un sous-groupe commutatif du groupe multiplicatif des matrices carrées d'ordre 2 inversibles. Déterminer A_a^n pour tout $n \in \mathbb{Z}$, puis $\exp(tA)$ pour tout $t \in \mathbb{R}$.

(b). Résoudre le système différentiel

$$\left\{\begin{array}{rcl} x'(t) & = & (\operatorname{ch} a)\, x(t) + (\operatorname{sh} a)\, y(t) \\ y'(t) & = & (\operatorname{sh} a)\, x(t) + (\operatorname{ch} a)\, y(t) \end{array}\right.$$

avec les conditions initiales $x(0) = 2$ et $y(0) = 0$.

4.(a). Donner la forme réduite de la forme quadratique q_a dans une base orthonormée.

(b). Déterminer la nature de la conique $(\mathcal{C}_a)$ d'équation $q_a(x, y) = 1$, puis représenter les courbes $(\mathcal{C}_a)$ pour $a \in \{-\ln 8, \ln 2, 0, \ln 2, \ln 8\}$ dans un même repère orthonormé du plan.