Algèbre linéaire

4e édition

Joseph GRIFONE

CÉPADUÈS-ÉDITIONS
111, rue Nicolas Vauquelin
31100 Toulouse – France
Tél. : 05 61 40 57 36 – Fax : 05 61 41 79 89
www.cepadues.com
Courriel : cepadues@cepadues.com
Coordonnées GPS en WGS 84
N 43° 34'43,2''
E 001° 24'21,5''

Chez le même éditeur

ISBN : 978.2.85428.962.6
1re éd. 978.2.85428.239.9

Dépôt légal : mai 2011

N° éditeur : 962

Sommaire

AVANT-PROPOS

L'Algèbre Linéaire a une place spéciale parmi les disciplines enseignées en premier cycle d'Université. D'une part parce que, étant utilisée pratiquement dans toutes les branches scientifiques, sa connaissance fait partie du bagage indispensable au futur chercheur, ingénieur ou agrégatif. D'autre part parce que l'algèbre et la géométrie se mêlent constamment, l'imagination est sans cesse sollicitée et, de ce fait, elle est très utile à la formation de l'esprit mathématique.

Malheureusement les programmes actuels de l'enseignement dans le secondaire ne comportent presque plus d'Algèbre Linéaire. N'ayant ni le recul, ni le langage de base, les étudiants abordent souvent cette discipline d'une manière abstraite et formelle : il s'ensuit un décalage entre le niveau requis et les résultats atteints, décalage que tout enseignant a pu constater ces dernières années.

Dans cet ouvrage, fruit d'une expérience de plusieurs années d'enseignement de l'Algèbre Linéaire en premier cycle d'Université, j'ai essayé de combler cette lacune. Je me suis efforcé de présenter les différentes notions en mettant en évidence leur raison d'être et leur intérêt, en soulignant leur signification géométrique et en illustrant chaque notion et chaque énoncé par des exemples et des exercices résolus.

J'ai évité, cependant, les originalités. En particulier, le plan est des plus classiques, avec une seule exception : les espaces euclidiens et les espaces hermitiens sont traités avant les formes quadratiques et les formes hermitiennes, contrairement à ce qui est fait dans la plupart des manuels. L'expérience montre que, malgré les redites, ces parties du cours sont beaucoup mieux comprises.

J'ai aussi essayé d'aller à l'essentiel. Pratiquement toute l'algèbre linéaire est contenue dans le cadre où le corps des scalaires est $\mathbb{R}$ ou $\mathbb{C}$. Non pas que l'algèbre linéaire sur les corps finis, par exemple, ne soit pas intéressante, mais elle n'est pas essentielle à la compréhension profonde de la théorie. C'est pourquoi j'ai relégué en appendice tout le bagage sur l'algèbre générale (notions de groupes, anneaux, corps, polynômes formels, quotients, etc) qui souvent assomme l'étudiant en début d'année. Ici aussi l'expérience prouve que, lors une seconde lecture, celui qui a déjà assimilé la substance, le mécanisme et la vision géométrique de l'algèbre linéaire sur $\mathbb{R}$ ou $\mathbb{C}$ passe sans difficultés à l'algèbre linéaire sur les corps plus généraux. Cependant toutes les définitions et les énoncés sont donnés pour un corps quelconque de manière à ce que l'exposé ne souffre pas de cette restriction.

Les appendices jouent d'ailleurs un rôle important dans le plan de cet ouvrage. J'ai voulu éviter un livre trop dense, une sorte d'encyclopédie de résultats, où l'étudiant a du mal à dégager ce qui est essentiel et ce qui est secondaire. Aussi les différents chapitres sont consacrés à la compréhension de ce qui est essentiel à la théorie, alors que les résultats annexes et les compléments sont traités en Appendice. Par ailleurs les appendices peuvent servir efficacement à ouvrir l'esprit sur le vaste champ auquel l'algèbre linéaire est appliquée. Pour ne citer qu'un exemple, j'ai renvoyé en appendice l'exponentielle des matrices, non seulement parce que son introduction dans le cours de base n'est pas vraiment essentielle, mais aussi parce que l'on peut ainsi mieux mettre en évidence l'importance et l'intérêt que cette notion présente, par exemple pour les systèmes dynamiques, voire les algèbres de Lie et les groupes continus de transformations. Dans ce même esprit j'ai traité dans les appendices certains thèmes que l'on aborde en général en dernière année de Licence ou en Master, comme les inverses généralisées, la méthode des moindres carrés, les méthodes directes de calcul numérique, les espaces affines, les systèmes dynamiques. Il ne s'agit pas évidemment d'exposés complets, mais d'introductions, qui, je l'espère, éveilleront la curiosité et l'intérêt du lecteur.

A la fin de chaque chapitre sont proposés un certain nombre d'exercices et de problèmes. Les plus difficiles sont précédés d'un ou plusieurs *. Leur finalité est de faciliter une meilleure compréhension du cours et d'apprendre, graduellement, à se servir des notions acquises. J'ai choisi de ne pas donner la solution complète, afin d'encourager le lecteur à chercher lui-même la réponse : tout enseignant sait qu'il est bien plus profitable pour un étudiant de réfléchir sur un exercice, voire d'échouer, que de lire plusieurs solutions. Cependant, l'expérience montre aussi que, souvent, le fait de ne disposer d'aucune indication peut amener à des impasses et décourager la réflexion. Ce qui m'a semblé le plus utile, est de guider la réflexion en donnant à part des indications sur la façon d'aborder les questions et de progresser dans le raisonnement, en fournissant aussi les résultats des calculs. Pour un bon recueil d'exercices et de problèmes, le lecteur est invité à consulter les livres de Jean-Marie Morvan et al. (Ed. Cépaduès), rédigés dans le même esprit que ce manuel et avec les mêmes notations. [1]

À l'occasion de cette nouvelle édition je voudrais tout d'abord renouveler ma gratitude envers tous ceux qui ont eu la gentillesse de m'aider par leurs indications et leurs commentaires, ou qui ont pris la peine de me signaler les coquilles et les corrections. Puisque je ne me fais pas d'illusions et que des imprécisions se sont probablement glissées aussi dans cette version, je suis d'avance reconnaissant envers tous ceux qui auront l'amabilité de me les signaler ou qui voudront bien me faire part de leurs remarques et de leurs suggestions pour améliorer cet ouvrage [2].

Je voudrais remercier tout spécialement ceux de mes collègues qui m'ont aidé par leurs conseils tout au long des diverses éditions de ce livre, en consacrant souvent une partie non négligeable de leur temps. MM. Jean-Baptiste Hiriart-Urruty, Lubomir Gavrilov, Pierre Damphousse, ont droit tout particulièrement à ma reconnaissance.

Pour terminer, je voudrais répéter ce que je disais lors de la première parution de cet ouvrage : que je n'aurais jamais songé à publier ce livre – car il existe de nombreux et excellents manuels [3] – si un certain nombre de mes étudiants ne me l'avait d'abord suggéré, puis demandé. C'est à eux que je le dédie.

Note sur la quatrième édition.

Dans cette nouvelle édition, j'ai ajouté quelques exercices et problèmes, ainsi que des nouveaux appendices afin de mieux faire comprendre les relations particulièrement étroites entre l'Algèbre Linéaire et la Géométrie. En particulier, le lecteur trouvera en appendice une étude plus fine du groupe orthogonal, la description du groupe des isométries en dimension trois et une introduction aux groupes cristallographiques par la détermination des groupes discrets de symétries d'un sous-ensemble borné de $\mathbb{R}^3$.

Toulouse, le 9 janvier 2011.

[1] Collection *Bien débuter en Mathématiques*, en particulier : *Espaces vectoriels et Matrices ; Réduction des endomorphismes ; Dualité, Formes quadratiques, Formes hermitiennes ; Problèmes de Mathématiques (3), Algèbre linéaire ; Topologie des espaces vectoriels normés.*

[2] email : jgrifone@gmail.com

[3] cf. par exemple, les deux tomes de J.M. Monier : *Algèbre 1 et 2*, Ed. Dunod.

Chapitre 1

Espaces Vectoriels

1.1 Introduction

Beaucoup de problèmes de mathématiques ou de physique vérifient la propriété suivante : si u et v sont deux solutions, alors $u + v$ est aussi une solution, ainsi que $k\, u$, k étant un nombre réel ou complexe. De tels problèmes sont dits *linéaires* [1] et ils sont habituellement plus faciles à résoudre que les problèmes plus généraux (appelés "non linéaires" précisément).

Les physiciens, par exemple, rencontrent souvent des problèmes linéaires : le principe de superposition exprime justement le fait que les équations de la chaleur, des cordes vibrantes, de l'électricité, etc. sont linéaires. Beaucoup d'autres problèmes sont linéaires *en première approximation* : l'équation des oscillations du pendule, par exemple, n'est pas linéaire, mais si l'on s'intéresse aux petites oscillations, elle peut être approchée par une équation linéaire.

En fait, un grand nombre de problèmes provenant de toutes les branches des mathématiques, ainsi que des applications à la physique, à la chimie, à l'économie, etc., sont linéaires, du moins en première approximation. On comprend dès lors l'intérêt qu'il peut y avoir à dégager un cadre mathématique commun à ce type de problèmes, de manière à pouvoir déterminer des méthodes et des algorithmes adaptés. Ce cadre mathématique commun est la notion d'*espace vectoriel.*

Avant de commencer l'étude abstraite, considérons un exemple géométrique qui va nous permettre de visualiser, d'une certaine manière, les constructions que nous allons élaborer.

On sait que les physiciens représentent certaines grandeurs par des segments

[1] Nous verrons dans la suite une définition plus précise de cette notion.

orientés que l'on appelle *vecteurs*. Une force, par exemple, n'est pas déterminée uniquement par son intensité, mais aussi :

- par son point d'application ;
- par la direction et le sens suivant lesquels elle s'exerce.

On représente cela par une flèche ayant comme origine le point d'application, de longueur égale à l'intensité de la force et dont le sens et la direction sont ceux de la force.

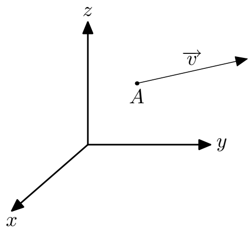


Figure 1

Sur les vecteurs de ***même origine*** on peut définir deux opérations :

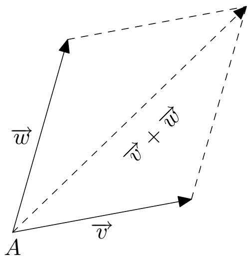


Figure 2

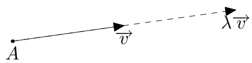


Figure 3

- ***l'addition***, qui exprime la loi de composition des forces ou des vitesses, définie par la «*règle du parallélogramme*» (cf. fig. 2)) ;

- ***le produit*** d'un vecteur par un nombre réel λ, qui donne un vecteur homothétique de rapport λ, c'est-à-dire un vecteur ayant la même direction, de même sens si $\lambda > 0$ et de sens contraire si $\lambda < 0$, et dont la longueur est multipliée par $|\lambda|$ (cf. fig. 3)).

REMARQUE. – ***On n'additionne pas de vecteurs d'origines différentes.***
La théorie des espaces vectoriels reflète justement cette situation ; aussi si l'on veut avoir une visualisation géométrique du problème, il faudra considérer toujours uniquement des vecteurs ***ayant tous la même origine.***

Pour pouvoir faire des calculs, on part de l'observation suivante. Considérons le plan ou l'espace ordinaire[2], que nous noterons $\mathcal{E}$ et *fixons un point* $O \in \mathcal{E}$. A tout point $P \in \mathcal{E}$ est associé un et un seul vecteur d'origine O (celui qui a P comme extrémité). Si l'on choisit un système de coordonnées d'origine O, on peut donc repérer les vecteurs du plan issus de O par les coordonnées de P, c'est-à-dire par les couples $(x_1, x_2) \in \mathbb{R}^2$. D'une manière analogue, les vecteurs issus d'un point de l'espace peuvent être repérés par les triplets $(x_1, x_2, x_3) \in \mathbb{R}^3$.

[2]Nous utiliserons les mots plan, espace, point, système de coordonnées, etc. au sens de la géométrie élémentaire.

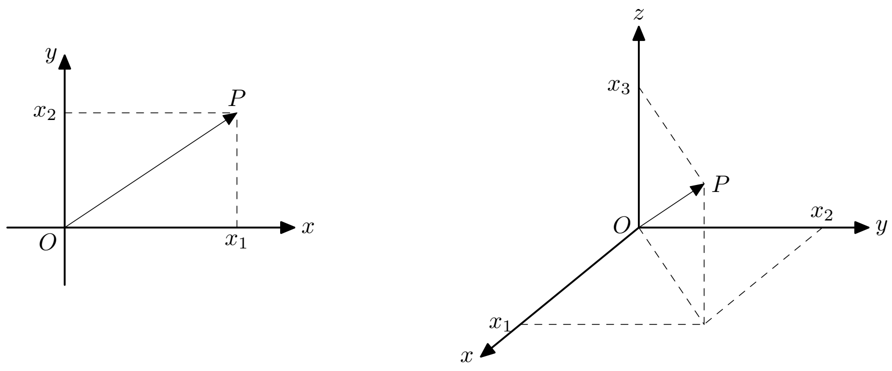


Figure 4

On vérifie facilement que si $\vec{v} = (x_1, x_2)$, $\vec{w} = (y_1, y_2)$ et $\lambda \in \mathbb{R}$, alors :

$$\begin{aligned} \vec{v} + \vec{w} &= (x_1 + y_1, x_2 + y_2) \\ \lambda\,\vec{v} &= (\lambda\,x_1, \lambda\,x_2) \end{aligned}$$

A l'aide de ces formules, on peut vérifier les propriétés suivantes :

A) L'addition est commutative et associative, c'est-à-dire :

$$\begin{aligned} \vec{v} + \vec{w} &= \vec{w} + \vec{v} && (\textit{commutativité}) \\ \vec{u} + (\vec{v} + \vec{w}) &= (\vec{u} + \vec{v}) + \vec{w} && (\textit{associativité}) \end{aligned}$$

De plus, il existe un vecteur noté $\overrightarrow{0}$ tel que $\vec{0} + \vec{v} = \vec{v}$ pour tout vecteur $\vec{v}$ (il s'agit évidemment du vecteur de longueur nulle $\vec{0} = (0, 0)$). Enfin, pour tout vecteur $\vec{v}$, il existe un vecteur $\vec{w}$ tel que $\vec{v} + \vec{w} = \vec{0}$ (si $\vec{v} = (x_1, x_2)$ alors $\vec{w} = (-x_1, -x_2)$).

B) Quant au produit par un réel, il vérifie les propriétés suivantes :

$$\begin{cases} \lambda\,(\mu\,\vec{v}) &= (\lambda\,\mu)\,\vec{v} \\ (\lambda + \mu)\,\vec{v} &= \lambda\,\vec{v} + \mu\,\vec{v} \\ \lambda\,(\vec{v} + \vec{w}) &= \lambda\,\vec{v} + \lambda\,\vec{w} \\ 1\,.\,\vec{v} &= \vec{v} \end{cases}$$

Il y a, bien entendu, beaucoup d'autres propriétés qui sont vérifiées mais, comme nous le verrons, celles que nous venons de signaler constituent justement «*le cadre mathématique commun à tous les problèmes linéaires*». En d'autres termes, toutes les propriétés essentielles des problèmes linéaires peuvent être dégagées à partir de ces propriétés.

Remarque. – L'exemple que l'on vient d'étudier est très utile, car il permet d'avoir présent à l'esprit un modèle géométrique qui peut servir de support à l'intuition. Cependant, il est important de comprendre que cette interprétation, même si elle est suggestive, n'est pas essentielle à la théorie. D'abord parce que nous ne considérerons pas seulement des espaces de «dimension» 2 ou 3 comme $\mathbb{R}^2$ ou $\mathbb{R}^3$ mais aussi des espaces de dimension supérieure, comme $\mathbb{R}^n$, ou même de dimension infinie. D'une manière approximative on peut dire que la *dimension*, dont nous donnerons la définition précise par la suite, est liée au nombre des paramètres qui interviennent dans le problème, nombre qui peut être très grand : dans ce cas, l'analogie avec les vecteurs de l'espace ordinaire risque de ne pas être d'un grand secours. D'autre part, on considérera aussi la multiplication des vecteurs par des nombres complexes ou, en général, appartenant à un corps commutatif quelconque et dans ce cas l'interprétation géométrique de la loi de produit n'est pas évidente. Ceci dit, le support géométrique est particulièrement important en algèbre linéaire et, en règle générale, il ne faut pas se priver d'y faire appel.

1.2 Espaces vectoriels

Définition 1.1 – *Soit K un corps commutatif* [3]*. On appelle* **espace vectoriel** *sur K un ensemble E sur lequel on a défini deux lois de composition :*

A) *Une loi interne [c'est-à-dire une application $E \times E \to E$] dite* addition*, notée $+$, et vérifiant :*

A.1) $(x+y)+z = x+(y+z)\ ,\quad \forall\, x,y,z \in E\,;$

A.2) $x+y = y+x\ ,\quad \forall\, x,y \in E\,;$

A.3) *Il existe un élément de E noté 0_E, ou plus simplement 0, dit* neutre*, tel que* $\forall x \in E :\ x + 0_E = x\,;$

A.4) *Pour tout $x \in E$, il existe un élément de E noté $(-x)$, dit* opposé de x*, tel que :* $x+(-x) = 0_E$ [4]*.*

B) *Une loi externe de domaine K [c'est-à-dire une application $K \times E \to E$; on note λx (ou $\lambda \cdot x$) l'image dans E du couple $(\lambda, x) \in K \times E$], qui vérifie :*

B.1) $\lambda\,(\mu\, x) = (\lambda\,\mu) x\ ,\quad \forall\, \lambda,\mu \in K\,,\quad \forall\, x \in E\,;$

B.2) $(\lambda+\mu)x = \lambda\, x + \mu\, x\ ,\quad \forall\, \lambda,\mu \in K\,,\quad \forall\, x \in E\,;$

B.3) $\lambda\,(x+y) = \lambda\, x + \lambda\, y\ ,\quad \forall\, \lambda \in K\,,\quad \forall\, x,y \in E\,;$

B.4) $1\,.\,x = x\ ,\quad \forall\, x \in E$
(1 étant l'élément neutre de la multiplication dans K).

Les éléments de K sont dits **scalaires** *et ceux de E* **vecteurs***.*

[3] Pour ne pas alourdir l'exposé, la définition de corps est donnée en Appendice 1. Le lecteur qui n'est pas familiarisé avec cette notion pourra supposer que $K = \mathbb{R}$ ou $\mathbb{C}$.

[4] On exprime ces quatre propriétés en disant que $(E, +)$ est un groupe abélien (ou commutatif) (cf. Appendice 1).

Exemple 1. –
$E = \mathbb{R}^n$ muni des lois suivantes est un espace vectoriel sur $\mathbb{R}$:

$$\begin{array}{rcl} (x_1,\ldots,x_n) + (y_1,\ldots,y_n) & := & (x_1+y_1,\ldots,x_n+y_n) \\ \lambda(x_1,\ldots,x_n) & := & (\lambda\, x_1,\ldots,\lambda\, x_n) \end{array} \tag{1}$$

Ici $0_{\mathbb{R}^n} = (0,\ldots,0)$; l'opposé $(-x)$ de $x = (x_1,\ldots,x_n)$ est $(-x_1,\ldots,-x_n)$. De même $\mathbb{C}^n$ est muni d'une structure d'espace vectoriel sur $\mathbb{C}$ et, plus généralement, K^n est muni d'une structure d'espace vectoriel sur K avec les lois définies par les formules (1).

Exemple 2. –
L'ensemble $\mathbb{R}_n[x]$ des *fonctions polynômes* à cœfficients dans $\mathbb{R}$ de degré $\leq n$, c'est-à-dire :

$$\mathbb{R}_n[x] = \{P : \mathbb{R} \longrightarrow \mathbb{R} \mid P(x) = a_0 + a_1\, x + \cdots + a_n\, x^n \ ; \ a_i \in \mathbb{R}\}$$

est un espace vectoriel sur $\mathbb{R}$ pour les lois :

$$\begin{array}{rcl} (a_0 + \cdots + a_n x^n) + (b_0 + \cdots + b_n x^n) & := & (a_0+b_0) + \cdots + (a_n+b_n)x^n \\ \lambda(a_0 + a_1 x + \cdots + a_n x^n) & := & (\lambda\, a_0 + \lambda\, a_1 x + \cdots + \lambda\, a_n x^n) \end{array} \tag{2}$$

Plus généralement, l'ensemble $\mathbb{R}[x]$ des fonctions polynômes *de tous les degrés possibles* à cœfficients dans $\mathbb{R}$ est un espace vectoriel sur $\mathbb{R}$ avec les lois :

$$\begin{array}{rcl} \sum_k (a_k x^k) + \sum_k (b_k x^k) & := & \sum_k (a_k + b_k) x^k \\ \lambda\left(\sum_k a_k x^k\right) & := & \sum_k (\lambda\, a_k) x^k \end{array} \tag{3}$$

(les sommes ne comportent qu'un nombre fini de termes non nuls).

Exemple 3.–
Soit $\mathcal{M}_2(K) = \{$ matrices carrées d'ordre 2 à cœfficients dans $K\}$. On définit sur $\mathcal{M}_2(K)$ une structure d'espace vectoriel sur K, en posant :

$$\begin{array}{rcl} \begin{pmatrix} a & b \\ c & d \end{pmatrix} + \begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix} & = & \begin{pmatrix} a+a' & b+b' \\ c+c' & d+d' \end{pmatrix} \\ \lambda \begin{pmatrix} a & b \\ c & d \end{pmatrix} & = & \begin{pmatrix} \lambda\, a & \lambda\, b \\ \lambda\, c & \lambda\, d \end{pmatrix} \end{array} \tag{4}$$

L'élément neutre est évidemment la matrice nulle $\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$ et l'opposée de $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ est $\begin{pmatrix} -a & -b \\ -c & -d \end{pmatrix}$

Exemple 4.–
Soit E un espace vectoriel sur K, $\mathcal{A}$ un ensemble quelconque non vide, et :

$$\mathcal{E} = \{\text{applications } f : \mathcal{A} \to E\}$$

On peut définir sur $\mathcal{E}$ une structure d'espace vectoriel sur K par les lois suivantes :

$$\text{si} \quad f, g \in \mathcal{E} : \qquad f+g : \begin{array}{ccc} \mathcal{A} & \longrightarrow & E \\ a & \longmapsto & f(a)+g(a) \end{array}$$

$$\text{si} \quad \lambda \in K\,,\ f \in \mathcal{E} : \qquad \lambda\, f : \begin{array}{ccc} \mathcal{A} & \longrightarrow & E \\ a & \longmapsto & \lambda\, f(a) \end{array}$$

c'est-à-dire :

$$\begin{aligned} (f+g)\,(a) \quad &:= f\,(a) + g\,(a) \\ (\lambda\, f)\,(a) \quad &:= \lambda\, f\,(a) \end{aligned} \tag{5}$$

Ici l'élément neutre est l'application constante nulle et l'opposée de f est définie par $(-f)\,(a)\ := -\Big(f\,(a)\Big)$.

Exemple 5.–
Soient E_1 et E_2 deux espaces vectoriels sur le même corps K. On définit une structure d'espace vectoriel sur $E_1 \times E_2$ par :

$$\begin{aligned} (x_1, x_2) + (y_1, y_2) \quad &:= (x_1 + y_1, x_2 + y_2) \\ \lambda\,(x_1, x_2) \quad &:= (\lambda\, x_1, \lambda\, x_2) \qquad (\lambda \in K) \end{aligned} \tag{6}$$

D'une manière analogue, on définit une structure d'espace vectoriel sur $E_1 \times \cdots \times E_n$, si $E_1, \ldots, E_n$ sont des espaces vectoriels sur le même corps K.

Proposition 1.2 – *Pour tout $\lambda \in K$ et pour tout $x \in E$, on a :*

1. $\lambda\, 0_{_E} = 0_{_E} \qquad$ et $\qquad 0\, x = 0_{_E}$.
2. $\lambda\, x = 0_{_E} \quad \Longrightarrow \quad \lambda = 0$ ou $x = 0_{_E}$.
3. $(-\lambda)\, x = \lambda(-x) = -(\lambda\, x)$.

Démonstration : -

1. $\lambda\, 0_{_E} = \lambda\,(0_{_E} + 0_{_E}) = \lambda\, 0_{_E} + \lambda\, 0_{_E}$, d'où : $0_{_E} = \lambda\, 0_{_E}$
   $0\, x = (0+0)\, x = 0\, x + 0\, x$ d'où : $0\, x = 0_{_E}$.
2. Supposons $\lambda\, x = 0_{_E}$ et $\lambda \neq 0$. En multipliant par λ^{-1}, on obtient $\lambda^{-1}(\lambda\, x) = \lambda^{-1}\, 0_{_E}$ c'est-à-dire, d'après 1. : $\lambda^{-1}(\lambda\, x) = 0_{_E}$ d'où, compte tenu de l'axiome B_1, $(\lambda^{-1}\lambda)\, x = 0_{_E}$, c'est-à-dire, $1\, x = 0_{_E}$ et donc $x = 0_{_E}$.
3. Facile, laissée en exercice. □

NOTA. – Dans la suite, $(-\lambda)\, x$ sera noté $-\lambda\, x$ et $x + (-y)$ sera noté $x - y$.

Exercices 1. 2. 3.

1.3 Sous-espaces vectoriels

Définition 1.3 – *Soit E un espace vectoriel et F une partie* non vide *de E. On dit que F est un* ***sous-espace vectoriel*** *de E, si la restriction des lois de E à F fait de F un espace vectoriel.*

En principe, pour montrer que F est un sous-espace vectoriel, il faudrait vérifier les huit axiomes de la Définition 1.1. En fait, il suffit de vérifier la «stabilité» des lois de composition comme l'affirme la proposition suivante :

Proposition 1.4 – *Soit E un espace vectoriel et $F \subset E$. Alors F est un sous-espace vectoriel de E si et seulement si :*

1. $F \neq \emptyset$
2. (a) $x, y \in F \implies x + y \in F$
   (b) $x \in F,\ \lambda \in K \implies \lambda\, x \in F$

Démonstration : Les conditions sont évidemment nécessaires. Supposons qu'elles sont satisfaites et montrons que F est un sous-espace vectoriel.
Les axiomes $\mathtt{A}_1, \mathtt{A}_2, \mathtt{B}_1, \mathtt{B}_2, \mathtt{B}_3, \mathtt{B}_4$ sont vérifiés pour tous les éléments de E et donc ils le sont en particulier par les éléments de F. Il reste à montrer que F admet un élément neutre et que tout élément de F admet un opposé (dans F).
Soit $x \in F$; on a $0\, x \in F$; mais $0\, x = 0_E$ donc $0_E \in F$. Ainsi *l'élément neutre* 0_E *de* E *appartient à* F.
D'autre part, $x + 0_E = x$, $\forall\, x \in F$, car cela est vrai pour tous les $x \in E$; aussi 0_E est l' élément neutre de F. L'axiome $\mathtt{A}_3$ est donc vérifié.
Enfin, si $x \in F$, on a $(-1)x \in F$, c'est-à-dire $-x \in F$. Ainsi, $\mathtt{A}_4$ est vérifié. $\square$

REMARQUE. – Comme on vient de le voir, l'élément neutre 0_E de E coïncide avec l'élément neutre 0_F de chaque sous-espace F. C'est pour cette raison que, dans la suite, on le notera simplement 0 (au lieu de 0_E ou 0_F), s'il n'y a pas de risque de confusion.

La proposition suivante se démontre facilement et elle est laissée en exercice :

Proposition 1.5 – *F est un sous-espace vectoriel de E si et seulement si :*

1. $F \neq \emptyset$
2. $x, y \in F$; $\lambda, \mu \in K \implies \lambda\, x + \mu\, y \in F$

REMARQUE. – Comme on l'a vu au cours de la démonstration, si F est un sous-espace vectoriel, alors F contient nécessairement le vecteur nul.

Exemples fondamentaux de sous-espaces vectoriels

1. ***Droite vectorielle.***

   Soit $v \in E,\ v \neq 0$; alors :

   $$F = \{y \in E \mid \exists\, \lambda \in K : y = \lambda\, v\}$$

   est un sous-espace vectoriel de E dit *droite vectorielle engendrée par v (cf. Fig. 5).*
   En effet $F \neq \emptyset$, car $v \in F$. De plus, F est stable pour les lois de E, car si $x, y \in F$ (c'est-à-dire : $x = \lambda\, v,\ y = \mu\, v$), on a :

   $$x + y = \lambda\, v + \mu\, v = (\lambda + \mu)\, v \in F$$

   De même, si $x \in F$ (c'est-à-dire $x = \lambda\, v$), on a : $\mu\, x = \mu\,(\lambda\, v) = (\mu\, \lambda)\, v \in F$.

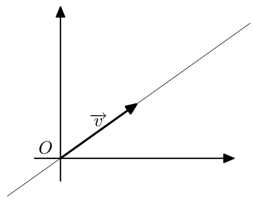


Figure 5

2. ***Plan vectoriel.***

   Soient $x_1, x_2 \in E$ et

   $$F = \{y \in E \mid \exists \lambda_1, \lambda_2 \in K : y = \lambda_1\, x_1 + \lambda_2\, x_2\}\,.$$

   F est un sous-espace vectoriel de E, dit *sous-espace engendré* par x_1 et x_2. Si x_1 et x_2 ne sont pas nuls et x_2 n'appartient pas à la droite vectorielle engendrée par x_1, F est dit *plan vectoriel* engendré par x_1 et x_2.

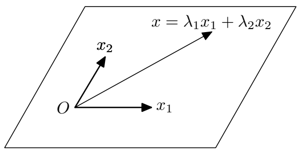


Figure 6

3. ***Sous-espace engendré.***

   Plus généralement, si $x_1, \dots, x_p \in E$ alors :

   $$F = \{y \in E \mid \exists\, \lambda_1, \dots, \lambda_p \in K : y = \lambda_1\, x_1 + \cdots + \lambda_p\, x_p\}$$

   est un sous-espace vectoriel de E noté $\mathrm{Vect}\{x_1, \dots, x_p\}$, dit ***sous-espace engendré*** par $x_1, \dots, x_p$, ou aussi ***espace des combinaisons linéaires*** de $x_1, \dots, x_p$. Nous verrons par la suite que, au fond, tous les sous-espaces vectoriels sont de ce type, c'est-à-dire obtenus par des "combinaisons linéaires" d'une famille d'éléments de E.

Remarque. – (Important) – Reprenons l'interprétation géométrique du § 1, E étant l'espace vectoriel des vecteurs d'origine O. Une droite vectorielle est une droite passant par O. De même, un plan vectoriel est un plan passant par O. Plus généralement, un sous-espace vectoriel de $\mathbb{R}^n$ peut être visualisé comme un "plan de dimension p" passant par O. On pourrait donner un sens précis à la notion de "plan de dimension p", mais cela n'est pas nécessaire. Retenons pour le moment le fait ***qu'il doit passer par*** O, car tout sous-espace vectoriel doit contenir le vecteur nul. Ainsi, par exemple, une droite ne passant pas par O n'est pas un sous-espace vectoriel : les points de la droite sont les extrémités des vecteurs issus de O et le vecteur nul n'est pas parmi eux.[5]

[5]Il s'agit en fait d'un «sous-espace affine», cf. Appendice 3.

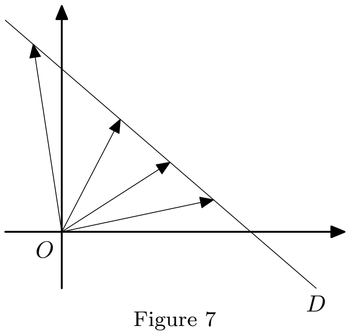


Figure 7

Exemple 1. –
Soit
$$F = \left\{(x, y, z) \in \mathbb{R}^3 \mid 2\,x + y + 3\,z = 0\right\}.$$
F est un sous-espace vectoriel de $\mathbb{R}^3$. En effet, soient $v_1 = (x_1, y_1, z_1)$ et $v_2 = (x_2, y_2, z_2) \in F$; on a :
$$2x_1 + y_1 + 3z_1 = 0 \quad \text{et} \quad 2\,x_2 + y_2 + 3\,z_2 = 0$$
d'où, en additionnant : $2\,(x_1 + x_2) + (y_1 + y_2) + 3\,(z_1 + z_2) = 0$, c'est-à-dire :
$$v_1 + v_2 = (x_1 + x_2,\ y_1 + y_2,\ z_1 + z_2) \in F$$
De même, on voit que $\lambda\, v \in F$ si $\lambda \in \mathbb{R}$ et $v \in F$.

Exemple 2. –
$\mathcal{E} = \left\{(x, y, z) \in \mathbb{R}^3 \mid 2\,x - y - z = 1\right\}$ n'est pas un sous-espace vectoriel de $\mathbb{R}^3$ car $0 = (0, 0, 0) \notin \mathcal{E}$.

Proposition 1.6 – *Soient F et G deux sous-espaces vectoriels de E.*

1. *$F \cap G$ est un sous-espace vectoriel de E.*
2. *$F \cup G$ n'est pas en général un sous-espace vectoriel de E.*
3. *Le complémentaire $E \setminus F$ d'un sous-espace vectoriel F n'est pas un sous-espace vectoriel de E.*

Démonstration :

1. On a d'abord $F \cap G \neq \emptyset$, car $0_E \in F \cap G$.
   Soient $x, y \in F \cap G$; on a $x, y \in F$ donc $x + y \in F$. De même, si $x, y \in G$, $x + y \in G$ et par conséquent $x + y \in F \cap G$.
   Si $\lambda \in K$ et $x \in F \cap G$, on a : $x \in F$, donc $\lambda\, x \in F$, et $x \in G$, donc $\lambda\, x \in G$; d'où : $\lambda\, x \in F \cap G$.

2. Cela tient au fait qu'en général $F \cup G$ n'est pas stable par la somme. Par exemple, soient $E = \mathbb{R}^2$, F la droite vectorielle engendrée par $(1, 0)$ et G la droite vectorielle engendrée par $(0, 1)$. On a :
   $(1, 0) \in F \quad \text{donc} \quad (1, 0) \in F \cup G$

   $(0, 1) \in G \quad \text{donc} \quad (0, 1) \in F \cup G$

   mais : $w = (1, 0) + (0, 1) = (1, 1) \notin F \cup G$

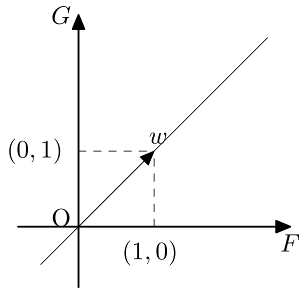


Figure 8

3. $E \setminus F$ ne contient pas 0, donc il n'est pas un sous-espace vectoriel. $\square$

REMARQUE. – C'est à cause de cela qu'en algèbre linéaire on considère rarement les réunions d'espaces vectoriels et encore moins les complémentaires, car on perd la structure d'espace vectoriel et donc toute la richesse de la théorie. Notamment, les raisonnements *par la contraposée*, c'est-à-dire du type «*supposons que* $x \notin E$» (E étant un espace vectoriel), sont en général à éviter.

Exercices 4. 5. 6.

1.4 Bases (en dimension finie)

Définition 1.7 – *Une famille de vecteurs* $\{v_1, \dots, v_p\}$ *d'un espace vectoriel* E *est dite* ***génératrice****, si* $E = \mathrm{Vec}\{v_1, \dots, v_p\}$*, ce qui veut dire (cf. Exemple 3. page 8) que*

$$\forall\, x \in E \;\; \exists\; \lambda_1, \dots, \lambda_p \in K \quad \text{tels que} \;:\quad x = \lambda_1\, v_1 + \cdots + \lambda_p\, v_p$$

On dit aussi, parfois, que tout $x \in E$ ***se décompose*** *sur les vecteurs* v_i *, ou encore que tout* $x \in E$ *est* ***combinaison linéaire*** *des vecteurs* v_i*.*

REMARQUE. – Une telle famille (finie) n'existe pas toujours. Considérons par exemple $\mathbb{R}[x]$ avec la structure d'espace vectoriel définie par les lois (3) page 5 et soit $\{P_1, \dots, P_n\}$ une famille finie de polynômes. Elle ne peut être génératrice car, en effectuant des combinaisons linéaires, on n'obtiendra que des polynômes de degré $\leq$ Sup{ degrés des (P_i)}.

Exemple 1. – Dans $\mathbb{R}^2$; la famille $\{e_1 = (1,0),\; e_2 = (0,1)\}$ est une famille génératrice car tout $x = (x_1, x_2) \in \mathbb{R}^2$, s'écrit :

$$(x_1, x_2) = x_1\,(1,0) + x_2\,(0,1) = x_1\, e_1 + x_2\, e_2.$$

Plus généralement dans K^n les vecteurs

$$e_1 = (1, 0, \dots, 0),\; \dots, e_k = (0, \dots 0, \underbrace{1}_{k^e \text{ rang}}, 0, \dots 0), \dots\dots, e_n = (0, \dots, 0, 1)$$

forment une famille génératrice, car pour tout $x = (x_1, \dots, x_n) \in K^n$, on peut s'écrire :

$$(x_1, \dots, x_n) = x_1\, e_1 + \cdots + x_n\, e_n$$

Exemple 2. – Dans $\mathbb{R}^2$, considérons une famille contenant les vecteurs e_1 et e_2 de l'exemple ci-dessus, par exemple la famille $\{e_1, e_2, v\}$, avec $v = (1, 2)$. Elle est évidemment génératrice, car pour tout $(x_1, x_2) \in \mathbb{R}^2$ on peut écrire :

$$(x_1, x_2) = x_1 e_1 + x_2 e_2 + 0\, v$$

Plus généralement : *toute famille contenant une famille génératrice est génératrice.*

Exemple 3. – Dans $\mathbb{R}^2$, soient $v_1 = (1, 1)$ et $v_2 = (1, -1)$. Montrons que $\{v_1, v_2\}$ est génératrice.
Soit $x = (a, b) \in \mathbb{R}^2$ avec a, b arbitraires : il s'agit de montrer qu'il existe $x_1, x_2 \in \mathbb{R}$ tels que $x = x_1 v_1 + x_2 v_2$, c'est-à-dire :

$$(a, b) = (x_1, x_1) + (x_2, -x_2) \equiv (x_1 + x_2,\ x_1 - x_2)$$

Ceci signifie que *quels que soient* a et $b \in \mathbb{R}$, il existe $x_1, x_2 \in \mathbb{R}$ vérifiant le système :

$$\begin{cases} x_1 + x_2 = a \\ x_1 - x_2 = b \end{cases}$$

En résolvant, on trouve en effet :

$$x_1 = \frac{a+b}{2}\ , \quad x_2 = \frac{a-b}{2}$$

solution définie *pour a, b arbitraires.* Donc $\{v_1, v_2\}$ est génératrice.

Définition 1.8 – *Un espace vectoriel est dit de **dimension finie**, s'il existe une famille génératrice finie ; dans le cas contraire, on dit qu'il est de dimension infinie.*

Comme on vient de le voir K^n est de dimension finie, alors que $\mathbb{R}[x]$ est de dimension infinie. De même, $\mathcal{E} = \{\text{applications}\, \mathbb{R} \to \mathbb{R}\}$ (avec la structure définie par les lois (5) page 5), est de dimension infinie car $\mathcal{E} \supset \mathbb{R}[x]$.

Définition 1.9 – *Soit $\{v_1, \ldots, v_p\}$, une famille finie d'éléments de E. On dit qu'elle est **libre**, si :*

$$\lambda_1\, v_1 + \cdots + \lambda_p\, v_p = 0 \implies \lambda_1 = 0, \ldots, \lambda_p = 0.$$

*On dit aussi que les vecteurs $v_1, \ldots, v_p$ sont **linéairement indépendants**. Une famille qui n'est pas libre est dite **liée** (on dit aussi que ses vecteurs sont liés ou linéairement dépendants).*

Exemple 1. – Dans $\mathbb{R}^3$, les vecteurs $v_1 = (1, 2, 1)$, $v_2 = (-1, 3, 1)$ et $v_3 = (-1, 13, 5)$ sont liés car on a $2\, v_1 + 3\, v_2 - v_3 = 0$.

Exemple 2. – Dans $\mathbb{R}^3$, les vecteurs $v_1 = (1, 1, -1)$, $v_2 = (0, 2, 1)$, $v_3 = (0, 0, 5)$ sont libres. En effet, supposons qu'il existe des réels $\lambda_1, \lambda_2, \lambda_3$ tels que $\lambda_1 v_1 + \lambda_2 v_2 + \lambda_3 v_3 = 0$, c'est-à-dire :

$$\lambda_1\ (1, 1, -1) + \lambda_2\ (0, 2, 1) + \lambda_3\ (0, 0, 5) = 0$$

On aura :

$$(\lambda_1,\ \lambda_1 + 2\,\lambda_2,\ -\lambda_1 + \lambda_2 + 5\,\lambda_3) = 0$$

c'est-à-dire :

$$\begin{cases} \lambda_1 & & & = 0 \\ \lambda_1 & + 2\,\lambda_2 & & = 0 \\ -\lambda_1 & + \ \lambda_2 & + 5\,\lambda_3 & = 0 \end{cases}$$

ce qui donne immédiatement : $\lambda_1 = \lambda_2 = \lambda_3 = 0$.

Proposition 1.10 – *Une famille $\{v_1, \ldots, v_p\}$ est liée si et seulement si l'un au moins des vecteurs v_i s'écrit comme combinaison linéaire des autres vecteurs de la famille (c'est-à-dire si l'un au moins des v_i appartient à l'espace vectoriel engendré par les autres).*

Ou, d'une manière équivalente :

une famille $\{v_1, \ldots, v_p\}$ est libre si et seulement si aucun des vecteurs v_i n'appartient à l'espace engendré par les autres.

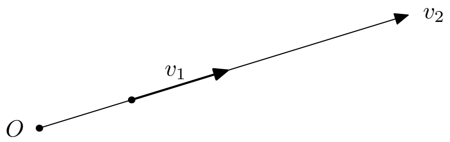


Figure 9-a

Les vecteurs v_1 et v_2 forment une famille liée : v_2 appartient à l'espace engendré par v_1.

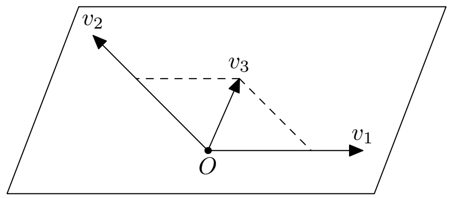


Figure 9-b

Les vecteurs v_1, v_2 et v_3 forment une famille liée : v_3 appartient à l'espace engendré par v_1 et v_2.

Démonstration : – Si $\{v_1, \ldots, v_p\}$ est liée, il existe $\lambda_1, \ldots, \lambda_p$ non tous nuls tels que $\lambda_1 v_1 + \cdots + \lambda_p v_p = 0$. Si, par exemple $\lambda_1 \neq 0$ (ce que l'on peut toujours supposer quitte à changer la numérotation), on pourra écrire :

$$v_1 = -\frac{\lambda_2}{\lambda_1} v_2 + \cdots + \frac{-\lambda_p}{\lambda_1} v_p$$

Réciproquement, supposons par exemple, que v_1 est combinaison linéaire des vecteurs $v_2, \ldots, v_p$; alors il existe $\mu_2, \ldots, \mu_p \in K$ tels que $v_1 = \mu_2 v_2 + \cdots + \mu_p v_p$, c'est-à-dire :

$$v_1 - \mu_2 v_2 - \cdots - \mu_p v_p = 0.$$

Il existe donc une combinaison linéaire des vecteurs $\{v_1, \ldots, v_p\}$ qui est nulle, sans que les cœfficients soient tous nuls. Donc la famille est liée. $\square$

L'intérêt de la notion de famille libre réside dans la propriété suivante :

Proposition 1.11 – *Soit $\{v_1, \ldots, v_p\}$ une famille libre et x un vecteur quelconque de l'espace engendré par les vecteurs v_i (c'est-à-dire x est combinaison linéaire des v_i). Alors la décomposition de x sur les v_i est unique.*

En effet, soient :

$$\begin{aligned} x &= \lambda_1 v_1 + \cdots + \lambda_p v_p \\ x &= \mu_1 v_1 + \cdots + \mu_p v_p \end{aligned}$$

deux décompositions de x. En faisant la différence on trouve :

$$0 = (\lambda_1 - \mu_1) v_1 + \cdots + (\lambda_p - \mu_p) v_p.$$

Puisque la famille est libre, on a $\lambda_1 - \mu_1 = 0, \ldots, \lambda_p - \mu_p = 0$ c'est-à-dire : $\lambda_1 = \mu_1, \ldots, \lambda_p = \mu_p$. $\square$

Définition 1.12 – *On appelle **base** une famille à la fois libre et génératrice.*

On a immédiatement :

Proposition 1.13 – *Une famille $\{v_1, \ldots, v_n\}$ est une base de E si et seulement si tout $x \in E$ se décompose d'une façon unique sur les v_i. C'est-à-dire :*

$\forall\, x \in E$ *il existe un **unique** n-uplet $(\lambda_1, \ldots, \lambda_n) \in K^n$ tel que :*

$$x = \lambda_1\, v_1 + \cdots + \lambda_n\, v_n.$$

(l'existence de la décomposition pour *tout* $x \in E$ équivaut au fait que la famille est génératrice ; l'unicité au fait que la famille est libre).

On peut aussi énoncer la proposition précédente de la manière suivante :

Proposition 1.14 – *Soit $B = \{v_1 \ldots v_n\}$ une base de E. Il existe alors une bijection :*

$$\varphi_B : \quad \underset{x = x_1\, e_1 + \cdots + x_n\, e_n}{E} \quad \underset{\longmapsto}{\longrightarrow} \quad \underset{(x_1, \ldots, x_n)}{K^n}$$

*Les scalaires x_i sont dits **composantes de** x dans la base $\{v_1, \ldots, v_n\}$.*

Exemple 1. – **Base canonique de $\mathbf{K^n}$** .

Soient les vecteurs :

$$\begin{aligned} e_1 &= (1, 0, \ldots, 0), \quad e_2 = (0, 1, 0 \ldots 0), \ldots\ldots\ldots, \\ e_k &= (0 \ldots 0, \underbrace{1}_{k^e rang}, 0 \ldots 0), \ldots\ldots\ldots, e_n = (0, \ldots, 0, 1) \end{aligned}$$

On sait déjà qu'ils forment une famille génératrice. Montrons qu'elle est libre. On a :

$$\begin{aligned} &\lambda_1\, e_1 + \cdots + \lambda_n\, e_n = 0 \iff \lambda_1\, (1, 0 \ldots 0) + \cdots + \lambda_n\, (0 \ldots 0, 1) = 0 \\ &\iff (\lambda_1, 0, \ldots, 0) + \cdots + (0, \ldots, 0, \lambda_n) = (0, \ldots, 0) \iff (\lambda_1, \ldots, \lambda_n) = (0, \ldots, 0) \\ &\iff \lambda_1 = 0, \ldots, \lambda_n = 0 \end{aligned}$$

Par conséquent $\{e_1, \ldots, e_n\}$ est une base de K^n, dite *base canonique.*

Exemple 2. – **Base canonique de $\mathbb{R}_\mathbf{n}\,[\mathbf{x}]$.**

La famille $\mathcal{B} = \left\{1, x, x^2, \ldots, x^n\right\}$ [6] est une base de $\mathbb{R}_n\,[x]$.
En effet, tout $P \in \mathbb{R}_n\,[x]$ s'écrit $P = a_0\, 1 + a_1\, x + \cdots + a_n\, x^n$ avec, $a_i \in \mathbb{R}$; $\mathcal{B}$ est donc génératrice. De plus :

$$\lambda_0\, 1 + \lambda_1\, x + \cdots + \lambda_n\, x^n = 0 \quad \Longrightarrow \quad \lambda_0 = 0,\ \lambda_1 = 0, \ldots, \lambda_n = 0.$$ [7]

[6] Plus précisément $\mathcal{B} = \{P_0, P_1, \ldots, P_n\}$ où $P_k : x \mapsto x^k$. Nous utiliserons souvent cet abus de notation.

[7] D'après la définition de polynômes : deux polynômes sont égaux si et seulement si les cœfficients des termes de même degré sont égaux (le 0 au second membre est le polyôme nul).

Exemple 3. – Dans $\mathcal{M}_2(K)$ (cf. exemple 2, page 5) soient les matrices :

$$E_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad E_2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad E_3 = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad E_4 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

Toute matrice $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathcal{M}_2(K)$, peut s'écrire :

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} = a\,E_1 + b\,E_2 + c\,E_3 + d\,E_4$$

ce qui montre que $\mathcal{B} = \{E_1, E_2, E_3, E_4\}$ est une famille génératrice.
D'autre part :

$$\begin{aligned} \lambda_1\,E_1 + \lambda_2\,E_2 + \lambda_3\,E_3 + \lambda_4\,E_4 = 0 \quad & \Longleftrightarrow \begin{pmatrix} \lambda_1 & \lambda_2 \\ \lambda_3 & \lambda_4 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \\ & \Longleftrightarrow \lambda_1 = \lambda_2 = \lambda_3 = \lambda_4 = 0 \end{aligned}$$

Donc $\mathcal{B}$ est une base de $\mathcal{M}_2(K)$.

Exemple 4. – Soit $F = \{(x, y, z) \in \mathbb{R}^3 \mid 2\,x + y + 3\,z = 0\}$. Chercher une base de F.
On a vu F est un sous-espace vectoriel de $\mathbb{R}^3$ (cf. exemple 1., page 9). On a : $v = (x, y, z) \in F \iff y = -2\,x - 3\,z$; donc :

$$v \in F \iff v = (x,\, -2\,x - 3\,z,\, z) \iff v = x\,(1, -2, 0) + z\,(0, -3, 1).$$

Par conséquent les vecteurs $v_1 = (1, 2, 0)$, $v_2 = (0, -3, 1)$ forment une famille génératrice de F. D'autre part :

$$\begin{aligned} & \lambda_1\,v_1 + \lambda_2\,v_2 = 0 \iff \lambda_1\,(1, -2, 0) + \lambda_2\,(0, -3, 1) = (0, 0, 0) \\ & \iff (\lambda_1,\, -2\,\lambda_1 - 3\,\lambda_2,\, \lambda_2) = (0, 0, 0) \iff \begin{cases} \lambda_1 & = 0 \\ -2\,\lambda_1 - 3\,\lambda_2 & = 0 \\ \lambda_2 & = 0 \end{cases} \end{aligned}$$

ce qui est équivalent à $\lambda_1 = \lambda_2 = 0$. Donc $\{v_1, v_2\}$ est libre et par conséquent elle est une base de F.

Proposition 1.15 –

1. *$\{x\}$ est une famille libre $\iff x \neq 0$.*
2. *Toute famille contenant une famille génératrice est génératrice.*
3. *Toute sous-famille d'une famille libre est libre.*
4. *Toute famille contenant une famille liée est liée.*
5. *Toute famille $\{v_1, \ldots, v_p\}$ dont l'un des vecteurs v_i est nul, est liée.*

Démonstration :

1. D'après la proposition 1.2 (2), $\lambda\,x = 0 \Longrightarrow \lambda = 0$ ou $x = 0$. Donc, si $x \neq 0$, $\lambda\,x = 0$ implique $\lambda = 0$, ce qui signifie que $\{x\}$ est une famille libre.
Réciproquement, supposons $\{x\}$ libre. Alors, d'après la définition de famille libre, si $\lambda\,x = 0$ on a nécessairement $\lambda = 0$, ce qui signifie, toujours d'après la proposition 1.2 (2), que $x \neq 0$.

2. Soit $\{v_1 \dots v_p\}$ une famille génératrice et $x = \lambda_1 v_1 + \cdots + \lambda_p v_p$ un élément arbitraire de E. On peut aussi écrire :

$$x = \lambda_1 v_1 + \cdots + \lambda_p v_p + 0\, w_1 + \cdots + 0\, w_q, \quad \text{avec} \quad w_1, \dots, w_q \in E.$$

   Donc tout $x \in E$ est combinaison linéaire de $v_1, \dots, v_p, w_1, \dots, w_q$.

3. Soit $\mathcal{F} = \{v_1 \dots v_p\}$ une famille libre et $\mathcal{F}'$ une sous-famille de $\mathcal{F}$. Quitte à changer la numérotation, on peut supposer que $\mathcal{F}' = \{v_1, \dots, v_k\}$ (avec $k \leq p$). Si $\mathcal{F}'$ était liée, l'un des vecteurs $v_1, \dots, v_k$ serait combinaison linéaire des autres. Il existerait donc un élément de $\mathcal{F}$ qui s'écrirait comme combinaison linéaire de certains éléments de $\mathcal{F}$. Or, cela est impossible car $\mathcal{F}$ est libre (cf. proposition 1.10).

4. Soit $\mathcal{F} = \{v_1, \dots, v_p\}$ une famille liée et $\mathcal{G} = \{v_1, \dots, v_p, w_1, \dots, w_q\}$. D'après la proposition 1.10, l'un des v_i est combinaison linéaire des autres. Or, les vecteurs v_i appartiennent à $\mathcal{G}$; donc l'un des éléments de $\mathcal{G}$ est combinaison linéaire des autres, et par conséquent $\mathcal{G}$ est liée.

5. Évident d'après 4., car il s'agit d'une famille contenant $\{0\}$, et $\{0\}$ est liée, d'après 1. $\square$

Exercices 7. 8. 9.

1.5 Existence de bases (en dimension finie)

Dans ce paragraphe nous allons montrer que dans tout espace vectoriel $E \neq \{0\}$, de dimension finie, il existe des bases.
Soit $E \neq \{0\}$ un espace vectoriel de dimension finie et $G = \{v_1, \dots, v_p\}$ une famille génératrice : pour tout $x \in E$, il existe $\alpha_1, \dots, \alpha_p \in K$ tels que :

$$x = \alpha_1 v_1 + \cdots + \alpha_p v_p.$$

Notons que G *contient certainement des familles libres* : il suffit de prendre, par exemple, $L = \{v_i\}$ avec $v_i \in G$, $v_i \neq 0$ [8] ; d'après la proposition 1.15, 1., L est libre.

Théorème d'existence . 1.16 –
Soit $E \neq \{0\}$ un espace vectoriel de dimension finie et G une famille génératrice. Considérons une famille libre $L \subset G$. Il existe alors une base $\mathcal{B}$ telle que $L \subset \mathcal{B} \subset G$.

Démonstration : Soit $G = \{v_1, \dots, v_p\}$ une famille génératrice et L_1 une famille libre contenue dans G. Quitte à changer la numérotation, on peut supposer que $L_1 = \{\, v_1, \dots v_r \,\}$. Si L_1 est génératrice, elle est une base et le théorème est démontré. Supposons donc que L_1 n'est pas génératrice.

a) Montrons tout d'abord qu'il existe $v_{i_1} \in \{v_{r+1}, \dots, v_p\}$ tel que la famille $L_2 = \{v_1, \dots, v_r,\, v_{i_1}\}$ est libre.

[8] G contient au moins un élément non nul, autrement on aurait $x = 0$ pour tout $x \in E$, ce qui est exclu car $E \neq \{0\}$.

En effet, si ce n'était pas le cas, chaque vecteur de $\{v_{r+1},\dots,v_p\}$ serait combinaison linéaire des vecteurs $\{v_1,\dots,v_r\}$ de L_1. Mais ceci est impossible, car si x est un vecteur quelconque de E, $x=\alpha_1\,v_1+\cdots+\alpha_r\,v_r+\alpha_{r+1}\,v_{r+1}+\cdots+\alpha_p\,v_p$, on pourrait remplacer dans cette expression les vecteurs $v_{r+1},\dots,v_p$ par leur expression en fonction des vecteurs $v_1,\dots,v_r$ et x s'écrirait comme combinaison linéaire des vecteurs $v_1,\dots,v_r$. Puisque x est arbitraire, cela signifierait que L_1 est génératrice, ce que nous avons exclu.
Nous avons donc prouvé que l'on peut *agrandir* la famille libre $L_1=\{v_1,...,v_r\}$ en lui ajoutant un vecteur $v_{i_1}\in\{v_{r+1},...,v_p\}$ de manière que la famille $L_2=\{L_1,v_{i_1}\}$ soit libre.

b) Si L_2 est génératrice, elle est une base et le théorème est démontré. Dans le cas contraire, en répétant le raisonnement, on voit qu'il existe $v_{i_2}\in G$, $v_{i_2}\notin L_2$, tel que la famille $L_3=\{L_2,v_{i_2}\}$ est libre. On construit ainsi une suite :

$$L_1 \underset{\neq}{\subset} L_2 \underset{\neq}{\subset} L_3 \underset{\neq}{\subset} \ldots\ldots\ldots \subset G$$

de familles libres et le processus peut être continué tant que L_k n'est pas génératrice. Mais G est une famille finie et par conséquent, le processus doit s'arrêter, éventuellement pour $L_k=G$. Il existe donc une famille L_k libre et génératrice, c'est-à-dire une base et $L_1\subset L_k\subset G$. $\square$

Ce théorème peut s'exprimer aussi de la manière suivante :

Théorème 1.17 –
Soit $E\neq\{0\}$ un espace vectoriel de dimension finie ; alors :

1. *De toute famille génératrice on peut extraire une base.*

2. (***Théorème de la base incomplète***). *Toute famille libre peut être complétée de manière à former une base.*

Démonstration :

1. C'est en fait ce nous avons établi dans la démonstration ci-dessus.

2. Soit $L=\{v_1,...,v_p\}$ une famille libre et $G=\{w_1,...,w_q\}$ une famille génératrice quelconque. La famille $G'=G\cup L$ est génératrice, car elle est une sur-famille d'une famille génératrice, et elle contient la famille L. D'après le théorème d'existence, il existe une base $\mathcal{B}$ telle que $L\subset\mathcal{B}\subset G'$ $\square$

Remarque. – Nous avons démontré en fait non seulement que toute famille libre peut être complétée en une base, mais qu'elle peut complétée en une base en choisissant les vecteurs *dans une famille génératrice arbitraire choisie à l'avance.*

1.6 Les théorèmes fondamentaux sur la dimension

Les résultats de ce paragraphe sont particulièrement importants.

Théorème 1.18 – *Dans un espace vectoriel E sur K de dimension finie, toutes les bases ont le même nombre d'éléments. Ce nombre est appelé* ***dimension*** *de E sur K et est noté* $\dim_K E$.

Démonstration : Pour la démonstration on a besoin du lemme suivant :

Lemme

Dans un espace vectoriel engendré par n éléments, toute famille contenant plus de n éléments est liée.

Démonstration du Lemme.– Soit $\mathcal{F} = \{v_1, \ldots, v_n\}$ une famille génératrice et $\mathcal{F}' = \{w_1, \ldots, w_m\}$ une famille de m vecteurs, avec $m > n$. Il s'agit de montrer que $\mathcal{F}'$ est liée.

a) Si l'un des w_i est nul $\mathcal{F}'$ est liée (cf. proposition 1.15, 5). Supposons donc que tous les w_i sont non nuls. Puisque $\mathcal{F}$ est génératrice, w_1 s'écrit :

$$w_1 = \alpha_1\, v_1 + \cdots + \alpha_n\, v_n$$

Comme $w_1 \neq 0$, il existe, parmi les $\alpha_1, \ldots, \alpha_n$, un α_i non nul. Supposons, quitte à changer la numérotation, que $\alpha_1 \neq 0$. On pourra alors écrire :

$$v_1 = \frac{1}{\alpha_1}\, w_1 - \left(\frac{\alpha_2}{\alpha_1}\, v_2 + \cdots + \frac{\alpha_n}{\alpha_1}\, v_n\right)$$

Soit x un élément quelconque de E : $x = \lambda_1\, v_1 + \cdots + \lambda_n\, v_n$. En remplaçant v_1 par l'expression trouvée, on voit que x est combinaison linéaire de $w_1, v_2, \ldots, v_n$. Puisque x est arbitraire cela signifie que la famille $\{w_1, v_2, \ldots, v_n\}$ est génératrice.
Nous avons donc remplacé dans la famille génératrice $\mathcal{F}$ l'un de ses vecteurs par l'un des vecteurs de $\mathcal{F}'$ et la famille ainsi obtenue est encore génératrice.

b) Considérons le vecteur w_2. Puisque $\{w_1, v_2, \ldots, v_n\}$ est génératrice, on peut écrire :

$$w_2 = \beta_1\, w_1 + \beta_2\, v_2 + \cdots + \beta_n\, v_n$$

Si $\beta_2 = \beta_3 = \cdots = \beta_n = 0$, on a $w_2 = \beta_1\, w_1$: la famille $\mathcal{F}' = \{w_1, w_2, \ldots, w_n\}$ est donc liée (car elle contient une famille liée) et le théorème est démontré.
Dans le cas contraire il existe un β_i parmi $\beta_2, \ldots, \beta_n$ différent de 0. Supposons, pour fixer les idées que $\beta_2 \neq 0$; on pourra alors écrire :

$$v_2 = \frac{1}{\beta_2}\, w_2 - \frac{1}{\beta_2}\, (\beta_1\, w_1 + \beta_3\, v_3 + \cdots + \beta_n\, v_n)$$

En raisonnant comme ci-dessus, on voit que la famille $\{\, w_1, w_2,\ v_3, \ldots, v_n\,\}$ est génératrice.

c) Ainsi, de proche en proche, on arrive à remplacer dans la famille $\mathcal{F} = \{v_1, \ldots, v_n\}$ les vecteurs $v_1, v_2, \ldots, v_n$ par les vecteurs $w_1, \ldots, w_n$ en obtenant à chaque étape une famille génératrice [9]. Au terme du processus, la famille $\{w_1, \ldots, w_n\}$ est génératrice. En particulier, w_{n+1} est combinaison linéaire des vecteurs $w_1, \ldots, w_n$, c'est-à-dire la famille $\{w_1, ..., w_{n+1}\}$ est liée. $\mathcal{F}'$ sera donc liée, car elle contient une famille liée. Le lemme est ainsi démontré. ◇

[9] Cette propriété est appelée parfois « *Théorème de l'échange* ».

Le théorème 1.18 s'en déduit immédiatement. En effet, soient $\mathcal{B}$ et $\mathcal{B}'$ deux bases. Comme $\mathcal{B}$ est génératrice, si $\mathcal{B}'$ avait plus d'éléments que $\mathcal{B}$ elle ne serait pas libre en vertu du Lemme. Or cela est exclu, car $\mathcal{B}'$ est une base. On a donc $\operatorname{Card}\mathcal{B}' \leq \operatorname{Card}\mathcal{B}$. De la même manière on voit que $\operatorname{Card}\mathcal{B} \leq \operatorname{Card}\mathcal{B}'$. □

Corollaire 1.19 –

1. *Dans un espace vectoriel de dimension n, toute famille ayant* ***plus*** *de n éléments est liée.*
2. *Dans un espace vectoriel de dimension n, les familles ayant* ***moins*** *de n éléments ne peuvent être génératrices.*

La première partie du corollaire est exactement ce que nous avons démontré dans le lemme ci-dessus.
La deuxième partie vient du corollaire 1.17, 1 : si $\mathcal{F}$ était une famille génératrice ayant moins de n éléments, on pourrait en extraire une base ayant moins de n éléments, ce qui est en contradiction avec le théorème 1.18. □

Remarques. –

1. Si $E = \{0\}$ on pose : $\dim_K E = 0$. On a évidemment :
$$E = \{0\} \iff \dim_K E = 0$$
2. $$\boxed{\dim_K K^n = n}$$
En effet, comme nous l'avons vu, la famille $\{e_1, ..., e_n\}$, avec $e_1 = (1, 0, \ldots, 0), \ldots,$ $e_n = (0, \ldots, 0, 1)$ est une base.
3. La dimension d'un espace vectoriel E dépend non seulement de E mais aussi du corps de base K (d'où la notation $\dim_K E$)[10].
En effet, considérons, par exemple, $\mathbb{C}$ muni de la structure d'espace vectoriel sur $\mathbb{R}$ définie par les lois :
$$\begin{array}{l} (a + i\,b) + (c + i\,d) = (a + c) + i\,(b + d) \\ \lambda\,(a + i\,b) = \lambda\,a + i\,\lambda\,b \qquad\qquad (\lambda \in \mathbb{R}) \end{array}$$
Tout élément $z \in \mathbb{C}$ s'écrit d'une manière unique : $z = a\,1 + b\,i$ avec $a, b \in \mathbb{R}$, ce qui signifie que $\{1, i\}$ est une base et donc $\dim_{\mathbb{R}} \mathbb{C} = 2$.
En revanche, si on considère $\mathbb{C}$ comme espace vectoriel sur $\mathbb{C}$, on a : $\dim_{\mathbb{C}} \mathbb{C} = 1$ (cf. remarque ci-dessus 2.).

Proposition 1.20 –
Soient $E_1, \ldots, E_p$ des espaces vectoriels de dimension finie sur le même corps K. Alors :
$$\dim_K(E_1 \times \cdots \times E_p) \;=\; \dim_K E_1 + \cdots + \dim_K E_p$$

En effet, soient $\{a_1, \ldots, a_{n_1}\}$, $\{b_1, \ldots, b_{n_2}\}, \ldots, \{l_1, \ldots, l_{n_p}\}$ des bases de E_1, $E_2, \ldots, E_p$. On vérifie facilement que la famille
$$\big\{ (a_i, 0, \ldots, 0\,)_{i=1\ldots n_1},\;\; (0, b_i, 0 \ldots 0\,)_{i=1\ldots n_2}, \ldots, \ldots (0, \ldots, 0, l_i\,)_{i=1\ldots n_p} \big\}$$
est une base de $E_1 \times \cdots \times E_p$. □

[10]Cependant, lorsque le contexte sera clair, on écrira aussi simplement $\dim E$ au lieu de $\dim_K E$.

En particulier, on a :

$$\dim_{\mathbb{C}} \mathbb{C}^n = n \ , \qquad \dim_{\mathbb{R}} \mathbb{C}^n = 2n.$$

En général, pour montrer qu'une famille est une base, il faut montrer qu'elle est libre et qu'elle est génératrice. Cependant, si la famille a exactement autant d'éléments que la dimension de l'espace, on a le théorème suivant qui est d'un usage fréquent :

Théorème 1.21 – *Soit E un espace vectoriel de dimension n. Alors :*
- *Toute famille génératrice ayant n éléments est une base.*
- *Toute famille libre ayant n éléments est une base.*

En effet, soit G une famille génératrice ayant n éléments. D'après le théorème 1.17 on peut en extraire une base. Mais cette base extraite doit avoir n éléments car la dimension de E est n : il s'agit donc de G elle-même.
De même, soit L est une famille libre ayant n éléments. D'après le théorème de la base incomplète on peut éventuellement lui ajouter certains vecteurs de manière à former une base. Mais si on lui ajoutait effectivement certains vecteurs, on aurait une base formée de plus de n éléments, ce qui est exclu. Donc L est elle-même une base.
□

Proposition 1.22 – *Soit E un espace vectoriel de dimension finie et F un sous-espace vectoriel de E. Alors F est de dimension finie, et de plus :*

1. $\dim_K F \leq \dim_K E$
2. $\dim_K F = \dim_K E \iff F = E$.

Démonstration :

1. Supposons $F \neq \{0\}$ et soit $x_1 \in F$, $x_1 \neq 0$. La famille $L_1 = \{x_1\}$ est libre. En raisonnant exactement comme dans la démonstration du Théorème d'existence 1.16, on construit une suite de familles libres de F :

   $$L_1 \underset{\neq}{\subset} L_2 \underset{\neq}{\subset} L_3 \underset{\neq}{\subset} \ldots\ldots\ldots \subset F$$

   et le processus ne s'arrête que si L_k est génératrice de F. Montrons d'abord que F est de dimension finie.
   Pour cela supposons que F n'admet pas de famille génératrice finie : il existe alors un indice $k > n$ (où n est la dimension de E) tel que L_k n'est pas génératrice. On aurait ainsi construit une famille libre de F formée de $k > n$ éléments. Or ceci est impossible, car les familles libres de F sont des familles libres de E et donc elles ne peuvent avoir plus de n éléments.

   F est donc de dimension finie. Plus précisément, on vient de voir que F admet nécessairement une famille génératrice formée d'au plus n éléments. En extrayant une base de cette famille, cette base aura au plus n éléments, donc : $\dim_K F \leq \dim_K E$.

2. Si $\dim_K F = \dim_K E$, il existe une base $\mathcal{B}$ de F ayant n éléments ($n = \dim_K E$). Mais $F \subset E$: $\mathcal{B}$ est donc une famille de E qui est libre et qui a n éléments. En vertu du théorème 1.21 , elle est une base de E et donc elle engendre E, c'est-à-dire $\mathrm{Vect}\{\mathcal{B}\} = E$. Mais $\mathrm{Vect}\{\mathcal{B}\} = F$, donc $F = E$. □

Remarque. – La partie 2. de la proposition ci-dessus est utilisée souvent pour montrer que deux espaces vectoriels E et F sont égaux. En effet, pour montrer que $E = F$, il faudrait, en principe, vérifier que $F \subset E$ et $E \subset F$. D'après la proposition, il suffit de vérifier que $F \subset E$ et que $\dim_K F = \dim_K E$ ce qui, en général, est plus facile.

Exercices 10. 11. 12. 13. 19. 20.

1.7 Bases en dimension infinie

Définition 1.23 – *Soit E un espace vectoriel, $\mathcal{F} = \{x_\alpha\}_{\alpha\in A}$ une famille d'éléments de E, non nécessairement finie, ni dénombrable. On appelle* ***combinaison linéaire finie*** *(ou simplement combinaison linéaire) de la famille, toute expression du type :*

$$\sum_{i\in I} \lambda_i\, x_i\,, \quad \text{où } I \text{ est une sous-famille } \textbf{\textit{finie}} \text{ de } A.$$

On appelle ***sous-espace engendré*** *par $\mathcal{F}$ le sous-espace vectoriel de E noté* $\mathrm{Vect}\{\mathcal{F}\}$*, formé par toutes les combinaisons linéaires finies des éléments de $\mathcal{F}$.*

On vérifie immédiatement que $\mathrm{Vect}\{\mathcal{F}\}$ est un sous-espace vectoriel de E : si $u = \lambda_1\, x_1 + \cdots + \lambda_p\, x_p \in \mathrm{Vect}\{\mathcal{F}\}$ et $v = \mu_1\, y_1 + \cdots + \mu_q\, y_q \in \mathrm{Vect}\{\mathcal{F}\}$ (avec $x_i, y_i \in \mathcal{F}$) on a : $\lambda\, u + \mu\, v =$ combinaison linéaire finie d'éléments de $\mathcal{F}$.
Plus précisément, on a :

Proposition 1.24 – $\mathrm{Vect}\{\mathcal{F}\}$ *est le plus petit sous-espace vectoriel de E contenant $\mathcal{F}$. En particulier, si F est un sous-espace vectoriel de E,* $\mathrm{Vect}\{F\} = F$.

En effet, si G est un sous-espace vectoriel de E contenant $\mathcal{F}$, il doit contenir, à cause de la stabilité, toutes les combinaisons linéaires finies des éléments de $\mathcal{F}$, c'est-à-dire $G \supset \mathrm{Vect}\{\mathcal{F}\}$. $\square$

Définition 1.25 –

1. *Une famille $\mathcal{F} = \{x_\alpha\}_{\alpha\in A}$ d'éléments de E est dite* ***génératrice*** *si* $\mathrm{Vect}\{\mathcal{F}\} = E$, *c'est-à-dire, si $\forall\, x \in E$ il existe une sous-famille* finie $\{x_1, \ldots, x_p\} \subset \mathcal{F}$ *telle que :*
$$x = \lambda_1\, x_1 + \cdots + \lambda_p\, x_p.$$
2. *La famille $\mathcal{F}$ est dite* ***libre****, si toute sous-famille* finie *est libre, c'est-à-dire si :* $\forall\, I \subset A$, I finie *:*
$$\sum_{i\in I} \lambda_i\, x_i = 0 \quad \Longrightarrow \quad \lambda_i = 0\,, \qquad \forall\, i \in I.$$
3. *Une famille $\mathcal{F}$ est dite* ***base*** *si elle est libre et génératrice.*

On vérifie immédiatement :

Proposition 1.26 – *$\mathcal{B}$ est une base de E si et seulement si tout élément de E s'écrit d'une manière unique comme combinaison linéaire finie d'éléments de $\mathcal{B}$.*

Exemple 1.– **Base canonique de $\mathbb{R}[\mathbf{x}]$.**
$\{1, x, x^2, \ldots, x^n \ldots\}_{n\in\mathbb{N}}$ *est une base de* $\mathbb{R}\,[x]$.

En effet, la famille est génératrice car tout polynôme s'écrit comme combinaison linéaire finie de $\{1, x, x^2, \ldots, x^n \ldots\}$. D'autre part, elle est libre car si l'on considère une combinaison linéaire finie nulle : $\lambda_0 1 + \lambda_1\, x + \cdots + \lambda_p\, x^p = 0$, on a : $\lambda_0 = 0, \ldots, \lambda_p = 0$.

Exemple 2. –
Soit $\mathbb{R}^{\mathbb{N}} = \left\{ u : \underset{n \longmapsto u_n}{\mathbb{N} \longrightarrow \mathbb{R}} \right\}$ l'espace vectoriel de toutes les suites réelles. Considérons la famille $\mathcal{F} = \{e_1, e_2, \ldots, e_n, \ldots\}_{n \in \mathbb{N}}$ où :

$$e_1 = (1, 0, \ldots, 0, \ldots) \ , \ e_2 = (0, 1, 0, \ldots), \ldots, e_n = (0, \ldots, 0, \underbrace{1}_{n^e \text{ rang}}, 0, \ldots)$$

c'est-à-dire e_n est la suite nulle partout sauf au rang n.
On vérifie immédiatement qu'il s'agit d'une famille libre, *mais elle n'est pas une base*, car elle n'est pas génératrice. En effet, par combinaisons linéaires *finies*, on engendre des suites $(u_0, u_1, \ldots, u_n \ldots)$ dans lesquelles seulement un nombre fini des u_i sont non nuls (c'est-à-dire les suites nulles à partir d'un certain rang) [11].

Pour les espaces de dimension infinie, on peut démontrer aussi un théorème d'existence de bases, mais la démonstration dépasse le cadre de ce cours[12] :

Théorème 1.27 – *Tout espace vectoriel non réduit à $\{0\}$ admet une base. Plus précisément :*

1. *De toute famille génératrice on peut extraire une base.*
2. *Toute famille libre peut être complétée en une base.*

Exercice 14.

1.8 Somme, somme directe, sous-espaces supplémentaires

Définition 1.28 – *Soient E_1, E_2 deux sous-espaces vectoriels d'un espace vectoriel E. On appelle* ***somme*** *de E_1 et E_2 le sous-espace de E défini par :*

$$E_1 + E_2 = \{x \in E \mid \exists\, x_1 \in E, \exists\, x_2 \in E_2 : x = x_1 + x_2\}$$

Il s'agit effectivement d'un sous-espace vectoriel de E comme on le vérifie facilement. Notons $\mathcal{E} = E_1 + E_2$. D'après la définition, tout élément de $\mathcal{E}$ est somme d'un élément de E_1 et d'un élément de E_2. Mais cette décomposition en général n'est pas unique. En effet, supposons $E_1 \cap E_2 \neq \{0\}$ et soit $x_0 \neq 0$, $x_0 \in E_1 \cap E_2$. Si $x = x_1 + x_2$ (avec $x_i \in E_i$) on a aussi la décomposition :

$$x = (x_1 + x_0) + (x_2 - x_0)$$

avec $x_1 + x_0 \in E_1$, $x_2 - x_0 \in E_2$. On voit donc qu'une condition nécessaire pour que la décomposition soit unique, est que $E_1 \cap E_2 = \{0\}$.
Cette condition est aussi suffisante. Supposons, en effet, que $E_1 \cap E_2 = \{0\}$, et soient $x = x_1 + x_2$, $x = x_1' + x_2'$ deux décompositions de x sur E_1 et E_2. Par soustraction, on a :

$$x_1 - x_1' = x_2' - x_2$$

Soit y cet élément. On a $y = x_1 - x_1' \in E_1$ et $y = x_2' - x_2 \in E_2$; donc $y \in E_1 \cap E_2 = \{0\}$ c'est-à-dire $y = 0$, d'où : $x_1 = x_1'$ et $x_2 = x_2'$.
Nous avons ainsi démontré :

[11] Il s'agit des *polynômes formels* (cf. Appendice A.2).
[12] Elle s'appuie sur l'axiome du choix.

Proposition 1.29 – *Soient E_1 et E_2 deux sous-espaces vectoriels de E, et soit $\mathcal{E}=E_1+E_2$. La décomposition de tout élément de $\mathcal{E}$ en somme d'un élément de E_1 et d'un élément de E_2 est unique, si et seulement si $E_1 \cap E_2 = \{0\}$. On écrit alors :*

$$\mathcal{E} = E_1 \oplus E_2$$

*et on dit que $\mathcal{E}$ est **somme directe** de E_1 et E_2.*

En d'autres termes :

$$\mathcal{E} = E_1 \;\oplus\; E_2 \iff \left\{\begin{array}{c} \mathcal{E} = E_1 + E_2 \\ \text{et} \\ E_1 \cap E_2 = \{0\} \end{array}\right. \iff \left\{\begin{array}{c} \mathcal{E} = E_1 + E_2 \quad \text{et} \\ \text{tout élément de } \mathcal{E} \text{ s'écrit} \\ \text{d'une manière } \textit{unique} \\ x = x_1 + x_2 \\ \text{avec } x_1 \in E_1 \;,\; x_2 \in E_2 \end{array}\right.$$

Proposition 1.30 – *Soit E un espace vectoriel et E_1, E_2 deux sous-espaces vectoriels de E. Alors $E = E_1 \oplus E_2$ si et seulement si pour toute base $\mathcal{B}_1$ et E_1 et toute base $\mathcal{B}_2$ de E_2, $\{\mathcal{B}_1, \mathcal{B}_2\}$ est une base de E.*

En effet, soient $\mathcal{B}_1 = \{v_\alpha\}_{\alpha\in A}$ et $\mathcal{B}_2 = \{w_\beta\}_{\beta\in B}$ des bases de E_1 et E_2 respectivement et supposons que $\{v_\alpha, w_\beta\}_{(\alpha,\beta)\in A\times B}$ est une base de E. Alors tout $x \in E$ s'écrit d'une manière *unique* :

$$x = \underbrace{\lambda_1\, v_{\alpha_1} + \cdots + \lambda_p\, v_{\alpha_p}}_{\in E_1} \; + \; \underbrace{\mu_1\, w_{\beta_1} + \cdots + \mu_q\, w_{\beta_q}}_{\in E_2}$$

c'est-à-dire tout $x \in E$ s'écrit *d'une manière unique* $x = x_1 + x_2$ avec $x_1 \in E_1$ et $x_2 \in E_2$; donc $E = E_1 \oplus E_2$. Réciproquement, si $E = E_1 \oplus E_2$, tout $x \in E$ se décompose d'une manière unique sur E_1 et E_2 et, par conséquent, sur la famille $\mathcal{B} = \{\mathcal{B}_1, \mathcal{B}_2\}$. On en déduit que $\mathcal{B}$ est une base de E. $\square$

Définition 1.31 – *Soit E un espace vectoriel et E_1, E_2 deux sous-espaces vectoriels de E. On dit que E_1 et E_2 sont **supplémentaires** (ou que E_2 est un supplémentaire de E_1), si $E = E_1 \oplus E_2$.*

Corollaire 1.32 *Soit E un espace vectoriel. Pour tout sous-espace vectoriel E_1, il existe toujours un supplémentaire. Le supplémentaire de E_1 n'est pas unique, mais si E est de dimension finie, tous les supplémentaires de E_1 ont même dimension.*

Pour plus de clarté, nous donnons la démonstration en dimension finie. En dimension infinie, seules les notations sont un peu plus lourdes.
Soit $\{v_1, \ldots, v_p\}$ une base de E_1 et soit $n = \dim E$. D'après le théorème de la base incomplète, il existe $w_{p+1}, \ldots, w_n$ tels que $\{v_1, \ldots, v_p,\ w_{p+1}, \ldots, w_n\}$ est une base de E. En posant $E_2 = \mathrm{Vect}\{w_{p+1}, \ldots, w_n\}$, on obtient, d'après le théorème précédent, un supplémentaire de E_1 dans E.
Puisque le choix de $w_{p+1}, \ldots, w_n$ n'est pas unique, le supplémentaire de E_1 n'est pas unique ; cependant, tous les supplémentaires de E_1 sont de dimension $n-p$, p étant la dimension de E_1. $\square$
Le résultat suivant est d'un usage fréquent :

Théorème 1.33 – *Soit E un espace vectoriel de* dimension finie. *Alors*

$$E = E_1 \oplus E_2 \iff \begin{cases} \textit{1.} \quad E_1 \cap E_2 = \{0\} \\ \textit{2.} \quad \dim E = \dim E_1 + \dim E_2. \end{cases}$$

Démonstration : Les conditions sont nécessaires (d'après la proposition 1.29 et le corollaire 1.32). Montrons qu'elles sont suffisantes. Soient $\{v_1, \ldots, v_p\}$ une base de E_1 et $\{w_{p+1}, \ldots, w_n\}$ une base de E_2, n étant la dimension de E. Pour montrer que $E = E_1 \oplus E_2$ il suffit de montrer, d'après 1.30, que la famille $\{v_1, ..., v_p, w_1, ..., w_q\}$ est une base. Soit $\lambda_1 v_1 + \cdots + \lambda_p v_p + \mu_{p+1} w_{p+1} + \cdots + \mu_n w_n = 0$. On a :

$$\underbrace{\lambda_1 v_1 + \cdots + \lambda_p v_p}_{\in E_1} = -\underbrace{(\mu_{p+1} w_{p+1} + \cdots + \mu_n w_n)}_{\in E_2}$$

Soit y ce vecteur ; $y \in E_1 \cap E_2$ donc $y = 0$. D'où : $\lambda_1 v_1 + \cdots + \lambda_p v_p = 0$ et, comme la famille $\{v_1, \ldots, v_p\}$ est libre : $\lambda_1 = 0, \ldots, \lambda_p = 0$. De même, $\mu_{p+1} w_{p+1} + \cdots + \mu_n w_n = 0$ et donc $\mu_{p+1} = 0, \ldots, \mu_n = 0$. Donc, la famille $\{v_1, \ldots, v_p,\ w_{p+1}, \ldots, w_n\}$ est libre. Puisqu'elle a n éléments, elle est une base de E, d'après le théorème 1.21. On en déduit que $E = E_1 \oplus E_2$. □

Exemple 1. – Dans $\mathbb{R}^2$ soient v et w sont deux vecteurs indépendants, $E_1 = \text{Vect}\{v\}$ et $E_2 = \text{Vect}\{w\}$ les droites vectorielles engendrées par v et par w On a $\mathbb{R}^2 = E_1 \oplus E_2$ car $\{v, w\}$ est une base. Comme on le voit d'ailleurs sur la figure tout $x \in \mathbb{R}^2$ se décompose d'une manière unique sur E_1 et E_2.

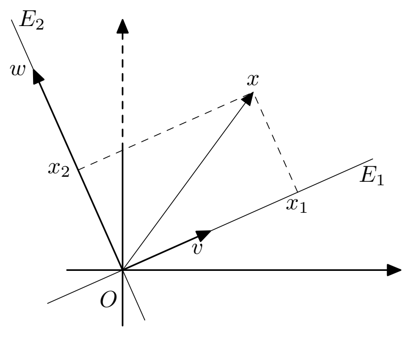


Figure 10

Exemple 2 – Dans $\mathbb{R}^3$, soit π un plan vectoriel et v un vecteur non contenu dans ce plan. On a :

$$\mathbb{R}^3 = \pi \oplus \text{Vect}\{v\}$$

car si $\{e_1, e_2\}$ est une base de π, alors $\{e_1, e_2, v\}$, est une base de $\mathbb{R}^3$.

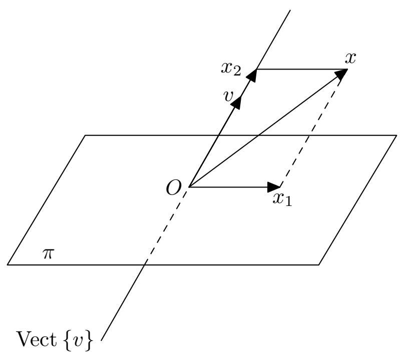


Figure 11

Exemple 3. – Soit $E = \mathcal{M}_2(\mathbb{R})$ et

$$E_1 = \left\{ A = \begin{pmatrix} a & 0 \\ 0 & 0 \end{pmatrix} \ , \ a \in \mathbb{R} \right\} \quad , \quad E_2 = \left\{ B = \begin{pmatrix} 0 & b \\ c & d \end{pmatrix} \ , \ b, c, d \in \mathbb{R} \right\}$$

E_1 et E_2 sont deux sous-espaces de E et $E_1 \cap E_2 = \{0\}$. De plus, toute matrice de E est la somme d'une matrice de E_1 et d'une matrice de E_2. Donc $E = E_1 \oplus E_2$.

Proposition 1.34 – *Soit E un espace vectoriel de dimension finie et E_1, E_2 deux sous-espaces vectoriels de E. On a :*

$$\boxed{\dim(E_1 + E_2) = \dim E_1 + \dim E_2 - \dim(E_1 \cap E_2)}$$

En particulier :

$$\boxed{\dim(E_1 \oplus E_2) = \dim E_1 + \dim E_2}$$

Démonstration : Supposons que $\dim E_1 = p$, $\dim E_2 = q$ et $\dim E_1 \cap E_2 = r$. Notons que, puisque $E_1 \cap E_2$ est un sous-espace vectoriel de E_1 et de E_2, on a $r \leq p$ et $r \leq q$. Considérons une base $\{a_1, \ldots, a_r\}$ de $E_1 \cap E_2$. Puisque la famille $\{a_1, \ldots, a_r\}$ est libre, on peut la compléter en une base de E_1 et aussi en une base de E_2. On peut donc construire :

- une base de E_1 du type $\{a_1, \ldots, a_r,\ e_{r+1}, \ldots, e_p\}$,
- une base de E_2 du type $\{a_1, \ldots, a_r,\ \varepsilon_{r+1}, \ldots, \varepsilon_q\}$.

Or tout vecteur de $E_1 + E_2$ s'écrit comme somme d'un vecteur de E_1, et d'un vecteur de E_2 et donc il est de la forme :

$$\begin{aligned} x = \ & \lambda_1\, a_1 + \cdots + \lambda_r\, a_r + \lambda_{r+1}\, e_{r+1} + \cdots + \lambda_p\, e_p \\ & + \mu_1\, a_1 + \cdots + \mu_r\, a_r + \mu_{r+1}\, \varepsilon_{r+1} + \cdots + \mu_q\, \varepsilon_q \end{aligned}$$

c'est-à-dire, en posant $\nu_i = \lambda_i + \mu_i$, pour $i = 1, \ldots, r$:

$$x = \nu_1\, a_1 + \cdots + \nu_r\, a_r + \lambda_{r+1}\, e_{r+1} + \cdots + \lambda_p\, e_p + \mu_{r+1}\, \varepsilon_{r+1} + \cdots + \mu_q\, \varepsilon_q$$

Par conséquent, la famille $\{a_1, \ldots, a_r,\ e_{r+1}, \ldots, e_p,\ \varepsilon_{r+1}, \ldots, \varepsilon_q\}$ engendre $E_1 + E_2$. Montrons qu'elle est libre.
Soit une combinaison linéaire nulle :

$$\underbrace{\nu_1 a_1 + \cdots + \nu_r a_r}_{x \in E_1 \cap E_2} + \underbrace{\lambda_{r+1} e_{r+1} + \cdots + \lambda_p e_p}_{y \in E_1} + \underbrace{\mu_{r+1} \varepsilon_{r+1} + \cdots + \mu_q \varepsilon_q}_{z \in E_2} = 0 \qquad (*)$$

On a $x + y + z = 0$, donc $z = -(x + y)$. Or $z \in E_2$ et $x + y \in E_1$, donc $z \in E_1 \cap E_2$.

Par conséquent, z peut s'écrire comme combinaison linéaire des a_i :

$$\mu_{r+1}\, \varepsilon_{r+1} + \cdots + \mu_q\, \varepsilon_q = \rho_1\, a_1 + \cdots + \rho_r\, a_r$$

Mais $\{a_1, \ldots, a_r,\ \varepsilon_{r+1}, \ldots, \varepsilon_q\}$ est une base de E_2, donc tous les cœfficients de cette combinaison linéaire doivent être nuls. En particulier, $\mu_{r+1} = 0, \ldots, \mu_q = 0$. On voit de

même que : $\lambda_{r+1} = 0, \ldots, \lambda_p = 0$. De (*) on déduit alors que : $\nu_1\, a_1 + \cdots + \nu_r\, a_r = 0$. Or la famille $\{a_1, \ldots, a_r\}$ est libre, donc $\nu_1 = 0, \ldots, \nu_r = 0$.
Ainsi la famille $\{a_1, \ldots, a_r,\ e_{r+1}, \ldots, e_p,\ \varepsilon_{r+1}, \ldots, \varepsilon_q\}$ est libre et donc elle est une base de $E_1 + E_2$. On en déduit :

$$\begin{aligned} \dim (E_1 + E_2) &= r + (p - r) + (q - r) = p + q - r \\ &= \dim E_1 + \dim E_2 - \dim (E_1 \cap E_2) \end{aligned} \qquad \square$$

Exercices 15. 16. 17. 18.

1.9 Somme et somme directe de plusieurs sous-espaces

Nota – La théorie et les résultats de ce paragraphe ne seront utilisés qu'au chapitre 6. On peut donc en différer l'étude jusqu'au commencement du chapitre 6.

Définition 1.35 – *Soient $E_1, \ldots, E_p$ des sous-espaces vectoriels d'un même espace vectoriel E. On note :*

$$E_1 + \cdots + E_p = \{x \in E \mid \exists\, x_1 \in E_1, \ldots, \exists\, x_p \in E_p\ :\ x = x_1 + \cdots + x_p\}\,.$$

*$E_1 + \cdots + E_p$ est un sous-espace vectoriel de E dit **somme des sous-espaces** E_i.*

Proposition 1.36 – *Si $\mathcal{G}_1, \ldots, \mathcal{G}_p$ sont des familles génératrices respectivement de $E_1, \ldots, E_p$, alors $\{\mathcal{G}_1, \ldots, \mathcal{G}_p\}$ est une famille génératrice de $E_1 + \cdots + E_p$.*

Démonstration : Soit $x = x_1 + \cdots + x_p$ un élément quelconque de $E_1 + \cdots + E_p$, avec $x_i \in E_i$. On voit immédiatement que, puisque pour chaque i x_i est combinaison linéaire des éléments de la famille $\mathcal{G}_i$, x est combinaison linéaire des éléments de la famille $\{\mathcal{G}_1, \ldots, \mathcal{G}_p\}$. $\square$

D'après la définition, tout $x \in E_1 + \cdots + E_p$ se décompose sur les E_i ; mais, en général, la décomposition n'est pas unique. Si, par exemple, $E_1 \cap E_2 \neq \{0\}$ et $x = x_1 + \cdots + x_p$ avec $x_i \in E_i$. En prenant $x_0 \neq 0$, $x_0 \in E_1 \cap E_2$ on peut écrire aussi :

$$x = (x_1 - x_0) + (x_2 + x_0) + x_3 + \cdots + x_p$$

et l'on aurait deux décompositions différentes de x.

Définition 1.37 – *Soient $E_1, \ldots, E_p$ des sous-espaces vectoriels de E. On dit **qu'ils sont en somme directe** si tout vecteur de $\mathcal{E} = E_1 + \cdots + E_p$ se décompose d'une manière **unique** en somme d'un vecteur de E_1, un vecteur de $E_2, \ldots,$ un vecteur de E_p. On écrit alors :*

$$\mathcal{E} = E_1 \oplus \ldots \oplus E_p$$

*et on dit que $\mathcal{E}$ est la **somme directe** des E_i.*

La meilleure manière de comprendre ce qu'est une somme directe est donnée par la caractérisation suivante qui généralise la proposition 1.30.

Théorème 1.38 – *Soient E_1,...,E_p des sous-espaces vectoriels de E. Alors ces sous-espaces E_1,...,E_p sont en somme directe si et seulement si pour toutes bases $\mathcal{B}_1, \ldots, \mathcal{B}_p$ de $E_1, \ldots, E_p$ respectivement, la famille $\{\mathcal{B}_1, \ldots, \mathcal{B}_p\}$ est libre.*

En d'autres termes, en vertu de la proposition 1.36 :

> *$E = E_1 \oplus \ldots \oplus E_p$ si et seulement si pour toutes bases $\mathcal{B}_1, \ldots, \mathcal{B}_p$ respectivement de $E_1, \ldots, E_p$, la famille $\{\mathcal{B}_1, \ldots, \mathcal{B}_p\}$ est une base de E.*

Démonstration : (Pour plus de clarté, nous donnons ici la démonstration en dimension finie. La démonstration en dimension infinie est exactement du même type et seules les notations sont un peu plus lourdes).

Soient $\mathcal{B}_1 = \{v_1, \ldots, v_r\}, \ldots, \mathcal{B}_p = \{w_1, \ldots, w_s\}$ des bases de $E_1, \ldots, E_p$ et supposons que la famille $\mathcal{B} = \{\mathcal{B}_1, \ldots, \mathcal{B}_p\}$ est libre. Puisque elle engendre $\mathcal{E} = E_1 + \cdots + E_p$ (cf. 1.36), elle est une base de $\mathcal{E}$. Donc, tout $x \in \mathcal{E}$ s'écrit d'une manière unique :

$$x = \underbrace{\lambda_1 v_1 + \cdots + \lambda_r v_r}_{\in E_1} + \cdots + \underbrace{\mu_1 w_1 + \cdots + \mu_s w_s}_{\in E_p}$$

Les λ_i et les μ_i étant uniques, x se décompose d'une manière unique sur les E_i, ce qui veut dire que les E_i sont en somme directe.

Réciproquement, supposons les E_i en somme directe et montrons que $\mathcal{B} = \{\mathcal{B}_1, \ldots, \mathcal{B}_p\}$ est libre. Soit la combinaison linéaire nulle :

$$\underbrace{\lambda_1 v_1 + \cdots + \lambda_r v_r}_{\in E_1} + \cdots + \underbrace{\mu_1 w_1 + \cdots \mu_s w_s}_{\in E_p} = 0 \qquad a)$$

et remarquons que :

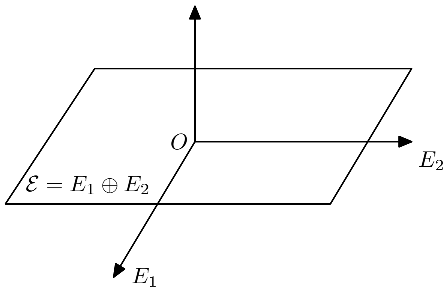
$$0_E = 0_{E_1} + \cdots + 0_{E_p} \qquad b)$$

(où on a noté 0_{E_i} l'élément neutre dans E_i, qui n'est autre que 0_E, pour souligner le fait qu'on le considère comme élément de E_i).
a) et b) fournissent deux décompositions de 0_E sur $E_1, \ldots, E_p$. Mais comme les E_i sont en somme directe, la décomposition est unique et donc :

$$\lambda_1 v_1 + \cdots + \lambda_r v_r = 0_{E_1}\ , \ldots\ldots,\ \mu_1 w_1 + \cdots + \mu_s w_s = 0_{E_s}$$

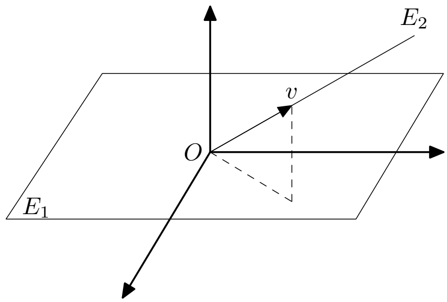
Or, les familles $\mathcal{B}_1, \ldots \mathcal{B}_p$ sont libres, et, par conséquent, $\lambda_1 = 0, \ldots, \lambda_r = 0, \ldots,$ $\mu_1 = 0, \ldots, \mu_s = 0.$ $\square$

Notons que $\mathcal{E} = E_1 \oplus \cdots \oplus E_p$ est un sous-espace vectoriel de E et peut, bien entendu, être différent de E.



$E_1 =$ axe Ox, $E_2 =$ axe Oy.
$\mathcal{E} = E_1 \oplus E_2 =$ plan xOy. On a : $\mathcal{E} \underset{\neq}{\subset} \mathbb{R}^3$.

Figure 12



$E_1 =$ plan Oxy, $E_2 = \mathrm{Vect}\{v\}$, $(v \notin E_1)$.
$\mathcal{E} = E_1 \oplus E_2$. On a : $\mathcal{E} = \mathbb{R}^3$.

Figure 13

Il faut donc bien distinguer les notions «les E_i sont en somme directe» (cf. Fig. 12) et «E est la somme directe des E_i» (cf. Fig. 13) ce qui signifie :

- les E_i sont en somme directe
- et, de plus, la somme des E_i "remplit" E tout entier.

Comme nous le verrons au chapitre 6, cette distinction est importante [13].

Corollaire 1.39 – *Si E est de dimension finie :*

$$\boxed{\dim(E_1 \oplus \cdots \oplus E_p) = \dim\, E_1 + \cdots + \dim\, E_p}$$

Particulièrement important est le corollaire suivant :

Corollaire 1.40 – *Soit E de dimension finie. Alors $E = E_1 \oplus \cdots \oplus E_p$, si et seulement si :*

1. $E = E_1 + \cdots + E_p$.
2. $\dim\, E = \dim\, E_1 + \cdots + \dim\, E_p$.

Démonstration : Les conditions sont évidemment nécessaires. Supposons qu'elles sont satisfaites et soient $\mathcal{B}_1, \ldots, \mathcal{B}_p$, des bases respectivement de $E_1, \ldots, E_p$. La famille $\mathcal{B} = \{\mathcal{B}_1, \ldots, \mathcal{B}_p\}$ est génératrice de E (d'après la condition 1. et la proposition 1.36). D'autre part, la condition 2. implique que le nombre d'éléments de $\mathcal{B}$ est égal à la dimension de E. Il s'ensuit, d'après le théorème 1.21, que $\mathcal{B}$ est une base et donc que $E = E_1 \oplus \cdots \oplus E_p$. $\square$

La proposition 1.29, page 22, se généralise de la manière suivante :

[13] Comme nous le verrons, l'obstacle au fait que l'on puisse *diagonaliser* une matrice consiste en ce que certains sous-espaces (les espaces propres), qui sont *toujours* en somme directe, peuvent avoir une somme *strictement* incluse dans E.

Théorème 1.41 – *Les sous-espaces $E_1, \dots, E_p$ sont en somme directe, si et seulement si*

$$E_1 \cap E_2 = \{0\} \ , \ (E_1 + E_2) \cap E_3 = \{0\} \ ,$$
$$(E_1 + E_2 + E_3) \cap E_4 = \{0\} \ , \dots\dots , (E_1 + E_2 + \cdots + E_{p-1}) \cap E_p = \{0\}$$

Démonstration : Supposons qu'il existe $x_0 \neq 0$, $x_0 \in (E_1 + \cdots + E_k) \cap E_{k+1}$ pour un certain $k \in \{1, \dots, p-1\}$ et soit $x = x_1 + \cdots + x_p$ (avec $x_i \in E_i$). On peut écrire :

$$\underbrace{x = x_1 + \cdots + x_k + x_0}_{\in E_1 + \cdots + E_k} + \underbrace{(x_{k+1} - x_0)}_{\in E_{k+1}} + x_{k+2} + \cdots + x_p$$

On aurait donc deux décompositions différentes de x sur les E_i, et par conséquent les E_i ne seraient pas en somme directe.
Réciproquement, supposons les conditions satisfaites et soient

$$x = x_1 + \cdots + x_p$$
$$x = x'_1 + \cdots + x'_p$$

deux décompositions de x sur les E_i. Par soustraction, on obtient :

$$(x_1 - x'_1) + \cdots + (x_{p-1} - x'_{p-1}) = x'_p - x_p$$

Notons y cet élément. On a $y \in (E_1 + \cdots + E_{p-1}) \cap E_p$ et donc, d'après la dernière condition, $y = 0$. Par conséquent :

$$x_p = x_{p'} \quad \text{et} \quad (x_1 - x'_1) + \cdots + (x_{p-2} - x'_{p-2}) = x'_{p-1} - x_{p-1}$$

Le même raisonnement montre que $x_{p-1} = x'_{p-1}$ et ainsi de suite jusqu'à montrer que la décomposition de x est unique, ce qui signifie que $E_1, \dots, E_p$ sont en somme directe. $\square$

Remarque. – Comme le montre la figure 14. la condition :

$$E_1 \cap E_2 \cap \cdots \cap E_p = \{0\}$$

ne suffit pas pour assurer que les sous-espaces $E_1, \dots, E_p$ sont en somme directe. Il en est de même pour les conditions

$$E_i \cap E_j = \{0\}$$

(pour $i \neq j$).

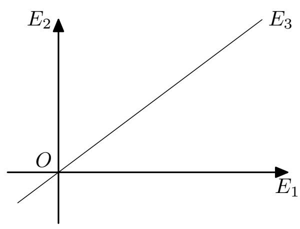


Figure 14

EXERCICES

1 On munit $\mathbb{R}^2$ des lois :

$$(x_1,\, x_2) + (y_1,\, y_2) \; : \; = \; (x_1 + y_1,\, x_2 + y_2)$$

$$\lambda\,(x_1,\, x_2) \; : \; = \; (\lambda\, x_1,\, 0) \qquad (\lambda \in \mathbb{R})$$

E est-il un espace vectoriel ?

2 On note $\mathbb{R}^+ \backslash \{0\}$, l'ensemble des nombres réels strictement positifs. Montrer que les lois :

$$x \oplus y \, := x\, y \qquad \lambda \,.\, x \, := x^{\lambda} \qquad (x,\, y \in \mathbb{R}_+ \backslash \{0\} \;,\; \lambda \in \mathbb{R})$$

confèrent à $\mathbb{R}^+ \backslash \{0\}$ une structure d'espace vectoriel sur $\mathbb{R}$.

*

3 Montrer que dans la définition d'espace vectoriel (définition 1.1), l'axiome `A.2` (commutativité de la somme) peut être déduit des autres axiomes.

4 Dans $\mathbb{R}^3$ muni des lois habituelles, lequel de ces sous-ensembles est-il un sous-espace vectoriel ?

$$F = \left\{ (x,y,z) \in \mathbb{R}^3 \;\middle|\; \begin{cases} x - y + z = 0 \\ 2\,x - y \quad = 0 \end{cases} \right\}$$

$$G = \{(x,y,z) \in \mathbb{R}^3 \mid (x-y)^2 = 2x + y\}$$

5 Dans $\mathcal{M}_2\,(\mathbb{R})$ muni les lois de l'exemple 3, page 5, lequel de ces sous-ensembles est-il un sous-espace vectoriel ?

$$E = \left\{ A \in \mathcal{M}_2\,(\mathbb{R}) \mid A = \begin{pmatrix} a+b & c \\ 2\,c & -b \end{pmatrix} \right\}$$

$$F = \left\{ A \in \mathcal{M}_2\,(\mathbb{R}) \mid A = \begin{pmatrix} 1 & a \\ 0 & b \end{pmatrix} \right\}$$

6 Soit $\mathbb{R}^{[a,b]} = \{f : [a,b] \longrightarrow \mathbb{R}\}$ muni des lois de l'exemple 4, page 5. Lequel de ces sous-ensembles est-il un sous-espace vectoriel ?

1. $\mathcal{C}^\circ\,([a,b],\mathbb{R}) = \{$fonctions continues $f : [a,b] \longrightarrow \mathbb{R}\}$;
2. ensemble des applications surjectives (resp. : injectives) $f : [a,b] \longrightarrow \mathbb{R}$;
3. ensemble des applications $f : [a,b] \longrightarrow \mathbb{R}$ telles que $2f\,(a) = f\,(b)$;
4. ensemble des applications $f : [a,b] \longrightarrow \mathbb{R}$ telles que $f\,(a) = f\,(b) + 1$.

7 Montrer que la famille $\{v_1, v_2\}$ où $v_1 = (1,2),\; v_2 = (-1,1)$ engendre $\mathbb{R}^2$. Est-elle une base ?

8

1. La famille $\{v_1, v_2, v_3\} \subset \mathbb{R}^3$ où $v_1 = (1,1,-1), v_2 = (2,1,3), v_3 = (0,-1,5)$ est-elle libre ? Quelle relation linéaire lie ces vecteurs ? Est-elle génératrice ?
2. Mêmes questions pour la famille $\{v_1, v_2, v_3\}$ où $v_1 = (1,1,0),\; v_2 = (1,2,1),$ $v_3 = (2,3,2)$.

9 Déterminer une base du sous-espace E de $\mathcal{M}_2\,(\mathbb{R})$ de l'exercice 5.

* **10** Soit $\mathcal{M}_n(K)$, l'ensemble des matrices carrées d'ordre n à cœfficients dans K, c'est-à-dire des tableaux d'éléments de K disposés sur n lignes et n colonnes. On définit sur $\mathcal{M}_n(K)$, les lois (analogues à celles de l'exemple 3, page 5) :

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & \dots & \dots & a_{2n} \\ \vdots & & & \vdots \\ a_{n1} & \dots & \dots & a_{nn} \end{pmatrix} + \begin{pmatrix} b_{11} & b_{12} & \dots & b_{1n} \\ b_{21} & \dots & \dots & b_{2n} \\ \vdots & & & \vdots \\ b_{n1} & \dots & \dots & b_{nn} \end{pmatrix} := \begin{pmatrix} a_{11}+b_{11} & a_{12}+b_{12} & \dots & a_{1n}+b_{1n} \\ a_{21}+b_{21} & \dots & \dots & a_{2n}+b_{2n} \\ \vdots & & & \vdots \\ a_{n1}+b_{n1} & \dots & \dots & a_{nn}+b_{nn} \end{pmatrix}$$

$$\lambda \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & \dots & \dots & a_{2n} \\ \vdots & & & \vdots \\ a_{n1} & \dots & \dots & a_{nn} \end{pmatrix} := \begin{pmatrix} \lambda a_{11} & \lambda a_{12} & \dots & \lambda a_{1n} \\ \lambda a_{21} & \dots & \dots & \lambda a_{2n} \\ \vdots & & & \vdots \\ \lambda a_{n1} & \dots & \dots & \lambda a_{nn} \end{pmatrix}$$

1. Déterminer une base et la dimension de cet espace.
2. Montrer que les matrices triangulaires supérieures, c'est-à-dire du type

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ 0 & a_{22} & \ddots & \vdots \\ \vdots & 0 & \ddots & \vdots \\ 0 & \dots & \dots & a_{nn} \end{pmatrix}$$

forment un sous-espace de $\mathcal{M}_n(K)$. En donner une base et la dimension.

* **11** On appelle ***trace*** d'une matrice $A \in \mathcal{M}_n(K)$, la somme des termes de la *diagonale principale* : $\operatorname{Tr} A = a_{11} + a_{22} + \ldots + a_{nn}$.

1. Montrer que l'ensemble $\mathcal{E} = \{A \in \mathcal{M}_n(K) \mid \operatorname{Tr}(A) = 0\}$ est un sous-espace vectoriel de $\mathcal{M}_n(K)$. En donner une base pour $n = 2$.
2. Soit $\mathcal{E}$ l'ensemble des matrices de $\mathcal{M}_n(K)$ à trace nulle et telles que la somme des éléments de chaque ligne est nulle. Montrer que $\mathcal{E}$ est un sous-espace vectoriel de $\mathcal{M}_n(K)$. En donner une base pour $n = 3$.

12 Soit $\mathcal{E} = \left\{P \in \mathbb{R}_2[x] : P = \lambda + (2\lambda - 3\mu)\,x + \mu\, x^2, \ \ (\lambda, \mu \in \mathbb{R})\right\}$. Montrer que $\mathcal{E}$ est un sous-espace vectoriel de $\mathbb{R}_2[x]$ et en donner une base.

* **13** Soit $P(x)$ un polynôme à cœfficients réels, de degré n. Montrer que P et ses n dérivées forment une base de $\mathbb{R}_n[x]$.

14 Soit $\mathcal{C}^\infty([a,b], \mathbb{R})$ l'ensemble des application $\mathcal{C}^\infty$ de $[a,b]$ dans $\mathbb{R}$. Montrer que les familles suivantes sont libres :
a) $\{x,\ e^x\}$ b) $\left\{e^x,\ e^{2x}\right\}$ c) $\{x,\ \sin x\}$ d) $\{\sin x,\ \cos x\}$
e) $\left\{1,\ e^x,\ e^{2x},\ \ldots,\ e^{nx}\right\}$ f) $\{1,\ \sin x,\ \sin 2x,\ \ldots,\ \sin nx\}$.

15 Soit E un espace vectoriel de dimension finie et F, G deux sous-espaces vectoriels de E. On note $\operatorname{Vect}(F \cup G)$ l'espace vectoriel engendré par $F \cup G$. Montrer que $\operatorname{Vect}(F \cup G) = F + G$.
Peut-on généraliser ce résultat au cas de n sous-espaces, c'est-à-dire : $\operatorname{Vect}(F_1 \cup F_2 \cup \ldots \cup F_n) = F_1 + \ldots + F_n$? Peut-on généraliser le résultat en dimension infinie ?

16 1. Soient F_1, F_2, F_3 trois sous-espaces vectoriels d'un espace vectoriel E. Montrer que :

$$F_1 \cap (F_2 + F_3) \supset F_1 \cap F_2 + F_1 \cap F_3$$

2. A-t-on l'inclusion contraire ?

* **17** Soient :

$$F = \left\{ A \in \mathcal{M}_2(\mathbb{R}) \mid A = \begin{pmatrix} a & 2a+b \\ -b & -a \end{pmatrix} \right\}$$

$$G = \left\{ A \in \mathcal{M}_2(\mathbb{R}) \mid A = \begin{pmatrix} a & 3a+b \\ -b & -2a+b \end{pmatrix} \right\}$$

Montrer que $\mathcal{M}_2(\mathbb{R}) = F \oplus G$.

* **18** Soit $A \in \mathcal{M}_n(K)$ et ${}^tA$ la matrice de $\mathcal{M}_n(K)$ dont les lignes sont les colonnes de A.
Par exemple, si $A = \begin{pmatrix} 1 & 2 & -1 \\ 0 & 1 & 3 \\ 8 & 1 & -1 \end{pmatrix}$ alors ${}^tA = \begin{pmatrix} 1 & 0 & 8 \\ 2 & 1 & 1 \\ -1 & 3 & -1 \end{pmatrix}$.
${}^tA$ est dite ***transposée*** de A. On dit que A est ***symétrique*** (resp. :***antisymétrique***) si ${}^tA = A$ (resp. : ${}^tA = -A$).

1. Montrer que les ensembles $\mathcal{S}_n(K)$ et $\mathcal{A}_n(K)$ des matrices respectivement symétriques et antisymétriques sont des sous-espaces vectoriels de $\mathcal{M}_n(K)$.
2. Déterminer une base et la dimension de $\mathcal{S}_n(K)$ et de $\mathcal{A}_n(K)$ pour $n = 3$. Généraliser à n quelconque.
3. Montrer que $\mathcal{M}_n(K) = \mathcal{S}_n(K) \oplus \mathcal{A}_n(K)$

** **19** **Espace vectoriel des suites récurrentes**

I. Soit E l'espace vectoriel des suites de nombres réels et $\mathcal{E} \subset E$ l'ensemble des suites vérifiant la relation de récurrence :

$$u_{n+2} = u_{n+1} + 2\,u_n \qquad (n \geq 0).$$

1. Montrer que $\mathcal{E}$ est un sous-espace vectoriel de E.
2. Montrer que les suites de terme général $a_n = (-1)^n$ et $b_n = 2^n$ forment une famille libre de $\mathcal{E}$.
3. Tenant compte du fait que les suites de $\mathcal{E}$ sont univoquement déterminées si on connaît u_0 et u_1, montrer que (a_n) et (b_n) forment une base de $\mathcal{E}$.
4. Déterminer les suites de $\mathcal{E}$ telles que $u_0 = 1$ et $u_1 = 2$.

II. *Généralisation.* Soit $\mathcal{E}$ l'ensemble des suites de nombres réels vérifiant :

$$u_{n+2} = \alpha\, u_{n+1} + \beta\, u_n \qquad (\alpha, \beta \in \mathbb{R}).$$

On suppose que α et β ne sont pas tous deux nuls.

1. Montrer que $\mathcal{E}$ est un espace vectoriel de dimension 2 sur $\mathbb{R}$.
2. Chercher une suite de $\mathcal{E}$ du type $a_n = r^n$. Montrer que r doit vérifier une équation de 2$^{\text{e}}$ degré. Lorsque cette équation admet deux racines réelles et distinctes r_1 et r_2, en déduire l'expression de u_n en fonction de u_0 et de u_1.
3. Montrer que lorsque l'équation de 2$^{\text{e}}$ degré admet une racine double r, les suites $a_n = r^n$ et $b_n = n\,r^n$ forment une base de $\mathcal{E}$.
4. Montrer que si λ_1 et λ_2 sont deux racines complexes conjuguées de l'équation de 2$^{\text{e}}$ degré et $\lambda_1 = \rho\, e^{i\theta}$, alors les suites $a_n = \rho^n \cos n\theta$ et $b_n = \rho^n \sin n\theta$ forment une base de $\mathcal{E}$.

III. *Application.* Déterminer le terme général des suites définies par la relation de récurrence :

$$u_{n+2} = 5\,u_{n+1} - 6\,u_n$$
$$u_{n+2} = 4\,u_{n+1} - 4\,u_n$$
$$u_{n+2} + u_{n+1} + u_n = 0.$$

** **20** **Équations différentielles linéaires à cœfficients constants**

I. Soit l'équation différentielle :

$$y'' + a\,y' + b\,y = 0$$

où $a, b \in \mathbb{R}$ et $y = y(x)$ la fonction inconnue (réelle de variable réelle.)

1. Montrer que les solutions forment un espace vectoriel $\mathcal{S}$ sur $\mathbb{R}$.
2. A l'aide du théorème d'existence et unicité, montrer que $\dim_{\mathbb{R}} \mathcal{S} = 2$.

II. Soit l'équation (dite équation caractéristique)

$$r^2 + ar + b = 0$$

1. Montrer que si elle admet deux racines réelles r_1, r_2, alors la solution générale est :
$$y = A\,e^{r_1 x} + B\,e^{r_2 x}$$
2. Montrer que si l'équation caractéristique admet une racine double r, la solution générale est :
$$y = (A + B\,x)\,e^{r\,x}$$
3. Montrer que si l'équation caractéristique admet deux racines complexes conjuguées $\alpha \pm i\,\beta$, la solution générale est :
$$y = e^{\alpha x}\,(A\cos\beta\,x + B\sin\beta\,x)$$

INDICATIONS

1 Seul l'axiome `B.4` n'est pas satisfait (ce qui montre qu'il ne peut être déduit des autres axiomes).

2 Vérifier les axiomes de la définition 1.1.

3 On a : $(1+1)\cdot(x+y\,) \underset{\mathtt{B_2}}{=} 1\cdot(x+y\,) + 1\cdot(x+y\,) \underset{\mathtt{B.4}}{=} (x+y\,) + (x+y\,)$.
D'autre part :

$$(1+1)\cdot(x+y\,) \underset{\mathtt{B.3}}{=} (1+1)\cdot x + (1+1)\cdot y \underset{\mathtt{B.2}}{=} (1\cdot x + 1\cdot x) + (1\cdot y + 1\cdot y\,) \underset{\mathtt{B.4}}{=} (x+x) + (y+y\,)$$

d'où : $(x+y\,) + (x+y\,) = (x+x) + (y+y\,)$.
En ajoutant à gauche $-x$ et à droite $-y$ (axiomes `A.4, A.3`) et en tenant compte de l'associativité (axiome `A.1`), on obtient $x + y = y + x$.

4 F oui ; G non.

5 Pour E, vérifier la stabilité (proposition 1.4) ou, mieux, écrire $A = a\begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix} + b\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix} + c\begin{pmatrix}0 & 1\\ 2 & 0\end{pmatrix}$: E est l'espace engendré par les trois matrices mises en évidence et donc il est un sous-espace vectoriel de $\mathcal{M}_2(\mathbb{R})$.
F ne contient pas la matrice nulle.

6

1. Oui.
2. Ne contient pas l'application nulle.
3. Oui.
4. Ne contient pas l'application nulle.

7 Le système $\begin{cases} x_1 - x_2 & = a \\ 2\,x_1 + x_2 & = b \end{cases}$ admet une solution *unique* pour tous a et b.

8

1. Le système $x_1\,v_1 + x_2\,v_2 + x_3\,v_3 = 0$ a comme solution $x_1 = 2\lambda$, $x_2 = -\lambda$, $x_3 = \lambda$ $(\lambda \in \mathbb{R})$. La famille n'est pas libre. Relation de dépendance : $2\,v_1 - v_2 + v_3 = 0$. La famille n'est pas génératrice, car le vecteur $v = (a, b, c)$ ne peut s'écrire sous la forme $v = x_1\,v_1 + x_2\,v_2 + x_3\,v_3$ que si : $4\,a - 5\,b - c = 0$.
2. Il s'agit d'une base.

9 $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 2 & 0 \end{pmatrix}$ est génératrice et libre.

10

1. $$E_{11} = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 0 & & \vdots \\ \vdots & & 0 & \vdots \\ 0 & \dots & \dots & 0 \end{pmatrix}, \quad E_{12} = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & & & \vdots \\ \vdots & & & 0 & \vdots \\ 0 & \dots & \dots & \dots & 0 \end{pmatrix} \quad \dots$$
   $$E_{nn} = \begin{pmatrix} 0 & \dots & 0 \\ \vdots & 0 & \vdots \\ 0 & \dots & 1 \end{pmatrix}$$
   (plus précisément, E_{ij} est la matrice dont tous les termes sont nuls sauf celui de la i-ième ligne et j-ième colonne, qui vaut 1). $\dim \mathcal{M}_n\,(K) = n^2$.
   Cette base est dite ***base canonique*** de $\mathcal{M}_n(\mathbb{R})$.
2. Engendré par $\{E_{ij}\}_{i \le j}$. La dimension est $1 + 2 + \dots + n = \dfrac{n\,(n+1)}{2}$.

11 Vérifier la stabilité des lois.

Pour $n = 2$, $A \in \mathcal{E}$ si et seulement si $A = \begin{pmatrix} a & b \\ c & -a \end{pmatrix}$ c'est-à-dire si et seulement si :

$$A = a \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} + b \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} + c \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}; \quad \dim \mathcal{E} = 3$$

12 $P \in \mathcal{E} \iff P = \lambda\,(1 + 2\,x) + \mu\,(-3\,x + x^2)$.
$P_1 = 1 + 2\,x$ et $P_2 = -3\,x + x^2$ forment une famille génératrice. Vérifier qu'il s'agit d'une famille libre.

13 Ecrire $\lambda_0\,P + \lambda_1\,P' + \lambda_2\,P'' + \dots + \lambda_n\,P^{(n)} = 0$. En égalant à 0 les cœfficients, il vient d'abord $\lambda_0 = 0$ car P est le seul polynôme de degré n. Ensuite $\lambda_1 = 0$ (cœfficient de x^{n-1}), etc. Appliquer ensuite le théorème 1.21.

14 a) $\lambda\, x + \mu\, e^x = 0 \quad \Longrightarrow \quad$ (en dérivant) $\lambda + \mu\, e^x = 0$. En faisant $x = 0$ dans ces deux équations, on obtient $\begin{cases} \mu & = 0 \\ \lambda + \mu & = 0 \end{cases}$, d'où : $\lambda = \mu = 0$.

On peut aussi donnerà x deux valeurs différentes. Par exemple $x = 0$ et $x = 1$. On obtient le système $\begin{cases} \mu & = 0 \\ \lambda + \mu\, e & = 0 \end{cases}$ qui donne : $\lambda = 0,\ \mu = 0$.

Raisonnement analogue pour les autres.

15 Utiliser la proposition 1.36.

16

1. Si $x \in (F_1 \cap F_2) + (F_1 \cap F_3)$ alors $x = y_2 + y_3$, avec $y_2 \in F_1 \cap F_2$ et $y_3 \in F_1 \cap F_3$. Donc $x \in F_1 \cap (F_2 + F_3)$.
2. Non. Considérer, par exemple, les sous-espaces de la figure 14, page 29.

17 Base de F : $\{A_1, A_2\}$ avec $A_1 = \begin{pmatrix} 1 & 2 \\ 0 & -1 \end{pmatrix}$ $A_2 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$.

Base de G : $\{B_1, B_2\}$ avec $B_1 = \begin{pmatrix} 1 & 3 \\ 0 & -2 \end{pmatrix}$ $B_2 = \begin{pmatrix} 0 & 1 \\ -1 & 1 \end{pmatrix}$.

Montrer que $\{A_1, A_2, B_1, B_2\}$ est une famille libre.

Appliquer les théorèmes 1.21 et 1.38.

18

1. Montrer d'abord que $\forall\, \lambda \in K,\ \forall\, A, B \in \mathcal{M}_n(K)$, on a ${}^t(\lambda\, A) = \lambda\, {}^tA$ et ${}^t(A+B) = {}^tA + {}^tB$. Se servir de ces propriétés pour montrer la stabilité des lois dans $\mathcal{S}_n(K)$ et dans $\mathcal{A}_n(K)$.
2. $A \in \mathcal{A}_3(K)$ si et seulement si :
$$\begin{aligned} A &= \begin{pmatrix} 0 & a & b \\ -a & 0 & c \\ -b & -c & 0 \end{pmatrix} \\ &= a \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} + b \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix} + c \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix} \end{aligned}$$
En déduire que $\dim \mathcal{A}_3(K) = 3$. De même : $\dim \mathcal{S}_3(K) = 6$.
Dans le cas général, les matrices antisymétriques sont caractérisées par les cœfficients strictement au-dessus de la diagonale principale (la diagonale principale est formée par des 0), et les matrices symétriques par les cœfficients au-dessus de la diagonale principale, la diagonale comprise. Par conséquent :
$$\dim \mathcal{S}_n(K) = \frac{n\,(n+1)}{2}, \quad \dim \mathcal{A}_n(K) = \frac{n\,(n-1)}{2}$$
3. Remarquer que $\forall\, A \in \mathcal{M}_n(K)$: $\quad A = \left(\dfrac{A + {}^tA}{2}\right) + \left(\dfrac{A - {}^tA}{2}\right)$.
Utiliser la proposition 1.29.

19 I.

1. Vérifier la stabilité des lois.
2. Ecrire $\lambda\,(-1)^n + \mu\, 2^n = 0$; faire $n = 0$, $n = 1$.
3. Pour toute suite $(u_n) \in \mathcal{E}$, montrer qu'il existe λ et μ tels que $u_n = \lambda\,(-1)^n + \mu\, 2^n$ faire $n = 0$, puis $n = 1$ et montrer que l'on peut résoudre le système en λ et μ.
4. $u_n = \dfrac{4}{3}\,(-1)^n - \dfrac{1}{3}\, 2^n$.

II. 1. Soient a_0, a_1, b_0, b_1 des réels tels que $a_0\, b_1 \neq a_1\, b_0$. On considère les suites (a_n), (b_n) de $\mathcal{E}$ dont les premiers termes sont respectivement a_0, a_1 et b_0, b_1. Soit $(u_n) \in \mathcal{E}$. Montrer qu'il existe un et un seul couple λ, μ de réels tels que :
$$\begin{cases} u_0 & = \lambda a_0 + \mu b_0 \\ u_1 & = \lambda a_1 + \mu b_1 \end{cases}$$
En déduire que (a_n) et (b_n) forment une base de $\mathcal{E}$.

2. En excluant la solution triviale $r = 0$, on a : $r^2 - \alpha r - \beta = 0$. Si r_1 et r_2 sont les deux racines réelles distinctes, $a_n = r_1^n$ et $b_n = r_2^n$ forment une base de $\mathcal{E}$ et pour toute suite de $\mathcal{E}$, il existe $\lambda, \mu \in \mathbb{R}$, tels que :

$$u_n = \lambda\, r_1^n + \mu\, r_2^n.$$

En faisant $n = 0$, puis $n = 1$, on détermine λ et μ en fonction de u_0 et u_1.

3. Même raisonnement.

4. On considère d'abord l'espace E' des suites complexes et $\mathcal{E}'$ le sous-espace des suites de E' vérifiant la relation de récurrence. On obtient deux suites $\varphi_n = \rho^n\, e^{in\theta}$ $\quad \psi_n = \rho^n\, e^{-in\theta}$ qui forment une base de $\mathcal{E}'$. La partie réelle et la partie imaginaire de φ_n sont :

$$a_n = \mathcal{R}e\ \varphi_n = \frac{\varphi_n + \psi_n}{2}, \quad b_n = \mathcal{I}m\ \varphi_n = \frac{\varphi_n - \psi_n}{2\,i}.$$

Etant combinaisons linéaires de φ_n et ψ_n, elles appartiennent à $\mathcal{E}'$ et, puisqu'elles sont réelles, elles appartiennent à $\mathcal{E}$. Vérifier qu'elles forment un système libre.

III.1. $u_n = \lambda\, 2^n + \mu\, 3^n$.

2. $u_n = (\lambda + n\,\mu)\, 2^n$.

3. $u_n = \lambda\ \cos\ \dfrac{2\pi}{3}\, n + \mu\ \sin\ \dfrac{2\pi\, n}{3}$.

20 I. 1. Montrer que $\mathcal{S}$ est un sous-espace vectoriel de l'espace des applications de $\mathbb{R}$ dans $\mathbb{R}$ muni des lois de l'exemple 4, page 5.

2. Le théorème d'existence et unicité affirme que $\forall\, y_0, z_0 \in \mathbb{R}$ il existe une et une seule solution $y = y(x)$ telle que $y(0) = y_0$ et $y'(0) = z_0$, dite «*solution de conditions initiales*» (y_0, z_0).
Soient $y_0, z_0,\ v_0, w_0 \in \mathbb{R}$ tels que $y_0\, w_0 \neq z_0\, v_0$ et y_1, y_2 les solutions de conditions initiales respectivement (y_0, z_0), (v_0, w_0). Soit y une solution quelconque, de condition initiale (α, β). Montrer qu'il existe un et un seul couple $A, B \in \mathbb{R}$, tel que :

$$\begin{cases} A\, y_0 + B\, v_0 = \alpha \\ A\, z_0 + B\, w_0 = \beta \end{cases}$$

En déduire que y_1 et y_2 est une base de $\mathcal{S}$.

II. 1. Vérifier que $\{e^{r_1 x}\,,\, e^{r_2 x}\}$ est une famille libre de solutions.

2. 3. Vérifications analogues.

Chapitre 2

La méthode du pivot (ou méthode d'élimination de Gauss)

L'étude de l'indépendance ou de la dépendance linéaire d'un système de vecteurs intervient d'une manière ou d'une autre dans tous les problèmes qui se posent en algèbre linéaire : détermination de bases, calcul de la dimension, mise en évidence de sommes directes, etc. Comme nous l'avons vu sur des exemples, le problème se ramène souvent à l'étude d'un système d'équations linéaires : savoir s'il admet des solutions, s'il y a une solution unique, voire déterminer les solutions. Aussi, l'étude des systèmes linéaires est-elle sous-jacente à tous les aspects de la théorie.

Dans ce chapitre nous expliquons la méthode du pivot qui fournit un algorithme simple et pratique pour résoudre ce type de problèmes. Une présentation rigoureuse de l'étude des systèmes linéaires sera donnée au chapitre 5. La finalité étant essentiellement d'ordre technique – disposer d'un outil de travail – nous nous limiterons le plus souvent, à faire comprendre la théorie sur des exemples, ce qui dans ce cas est largement suffisant, car les aspects théoriques sont d'une grande simplicité conceptuelle. Les raffinements de la méthode et les questions liées au choix du pivot seront traités dans l'Appendice A.4.

2.1 Etude d'un système d'équations linéaires par la méthode du pivot

On appelle système d'équations linéaires un système du type :

$$\begin{cases} a_{11}\,x_1 + a_{12}\,x_2 + \cdots + a_{1n}\,x_n = b_1 \\ a_{21}\,x_1 + a_{22}\,x_2 + \cdots + a_{2n}\,x_n = b_2 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{p1}\,x_1 + a_{p2}\,x_2 + \cdots + a_{pn}\,x_n = b_p \end{cases}$$

Les a_{ij} et les b_i sont des éléments de K, donnés. Les x_i sont dites '*inconnues* et résoudre le système signifie déterminer les $x_i \in K$, s'il y en a, qui vérifient toutes les équations.

La méthode du pivot est fondée sur la remarque suivante :

Propriété 2.1 – *L'ensemble des solutions d'un système linéaire ne change pas si l'on effectue sur les équations les « opérations élémentaires» suivantes :*

1. *changer l'ordre des équations ;*

2. multiplier une équation (premier et second membre) par un scalaire **non nul** *;*

3. ajouter à une équation une combinaison linéaire des **autres** *équations.*

Quelques exemples suffiront pour comprendre comment exploiter cette propriété, avant d'expliquer d'une manière plus précise la technique.

Exemple 1. –

$$\begin{array}{l} L_1 \\ L_2 \\ L_3 \end{array} \left\{ \begin{array}{l} 2x + y - 2z + 3w = 1 \\ \boxed{3}\,x + 2y - z + 2w = 4 \\ \boxed{3}\,x + 3y + 3z - 3w = 5 \end{array} \right.$$

On effectue des opérations élémentaires de manière à faire disparaître les deux coefficients encadrés. Par exemple :

$$\begin{array}{l} L_1 \\ L_2'=2L_2-3L_1 \\ L_3'=2L_3-3L_1 \end{array} \left\{ \begin{array}{rrrrl} 2x + & y & - 2z + & 3w & = 1 \\ & y & + 4z - & 5w & = 5 \\ & \boxed{3}\,y & +12z - & 15w & = 7 \end{array} \right.$$

L'ensemble des solutions n'a pas changé. On répète maintenant l'opération sur les deux dernières équations :

$$\begin{array}{l} L_1 \\ L_2' \\ L_3'-3L_2' \end{array} \left\{ \begin{array}{rl} 2x + y - 2z + 3w = & 1 \\ \boxed{}\,y + 4z - 5w = & 5 \\ \boxed{}\,0 = & -8 \end{array} \right.$$

On voit immédiatement que le système n'admet pas de solution.

Exemple 2. –

$$\begin{array}{l} L_1 \\ L_2 \\ L_3 \end{array} \left\{ \begin{array}{l} x + 2y - z = 1 \\ 2x + 3y + z = 2 \\ x + 4y - 6z = 2 \end{array} \right.$$

$$\begin{array}{l} L_1 \\ L_2'=L_2-2L_1 \\ L_3'=L_3-L_1 \end{array} \left\{ \begin{array}{r} x + 2y - z = 1 \\ - y + 3z = 0 \\ 2y - 5z = 1 \end{array} \right. \longrightarrow \begin{array}{l} L_1 \\ L_2' \\ L_3'+2L_2' \end{array} \left\{ \begin{array}{r} x + 2y - z = 1 \\ \boxed{} - y + 3z = 0 \\ \boxed{}\,z = 1 \end{array} \right.$$

On a donc $z = 1$. En reportant dans l'équation L_2' on obtient la valeur de y : $y = 3$. En remontant maintenant dans l'équation L_1 on trouve x : $x = -2y+z+1 = -6+1+1 = 4$. Le système admet donc la solution unique $x = -4,\ y = 3,\ z = 1$.

Comme on le voit, la méthode consiste à mettre le système «*sous forme échelonnée*» de manière à pouvoir, en partant de la solution de la dernière équation et en remontant, résoudre toutes les équations. Précisons cela.

Considérons une matrice, c'est-à-dire un tableau d'éléments de K rangés en lignes et colonnes. Pour un système d'équations linéaires on appelle matrice du système la matrice des cœfficients des premiers membres des équations. Par exemple, pour l'exemple 1., la matrice du système est :

$$\begin{array}{c}L_1\\L_2\\L_3\end{array}\begin{bmatrix}2&1&-2&3\\3&2&-1&2\\3&3&3&-3\end{bmatrix}$$

On notera L_1, $L_2, \ldots$ L_k les différentes lignes de la matrice. La ligne L_i sera dite $i^{\text{ème}}$ ligne ou *ligne d'indice i.*

On dit qu'une matrice est ***échelonnée*** si les lignes commencent par un nombre de zéros strictement croissant à mesure que l'indice augmente (c'est-à-dire, par exemple, la ligne L_3 commence par un nombre de zéros strictement plus grand que la ligne L_2 et celle-ci par un nombre de zéros strictement plus grand que la ligne L_1). Par exemple :

$$\begin{bmatrix}2&3&1&0&1&2&6\\0&0&7&1&8&2&0\\0&0&0&0&0&6&2\\0&0&0&0&0&0&0\end{bmatrix}$$

est une matrice échelonnée, alors que la matrice

$$\begin{bmatrix}1&3&1&0&1&5&9\\0&0&7&1&8&2&0\\0&0&5&1&0&6&2\\0&0&0&3&2&1&5\end{bmatrix}$$

n'est pas échelonnée.

Il est facile de se convaincre (la démonstration rigoureuse est lourde, mais elle n'est pas difficile) que par des opérations élémentaires, on peut toujours mettre le système sous forme échelonnée, c'est-à-dire on peut remplacer le système par un autre dont la matrice est échelonnée (l'ensemble des solutions n'aura pas changé).

Pour cela, il faut d'abord que la première équation commence par un cœfficient non nul, ce qu'on peut toujours faire en changeant éventuellement l'ordre des équations ($1^{ère}$ opération élémentaire). Ce cœfficient non nul est dit ***pivot***. (Si l'on reprend, par exemple, l'exemple 1., le pivot est 2).

Comme on le voit sur l'exemple 1., on peut échelonner les deux premières lignes en multipliant la ligne L_2 par le pivot (ce qu'on a le droit de faire car le pivot est non nul – cf. opération élémentaire b)) et en lui ajoutant l'équation L_1 multipliée par un cœfficient adéquat. On procède ainsi jusqu'à échelonner tout le système.

Exemple 3. –

$$\begin{array}{l} L_1 \\ L_2 \\ L_3 \\ L_4 \end{array} \left\{ \begin{array}{rcl} x + 2y - 3z &=& 4 \\ x + 3y + z &=& 11 \\ 2x + 5y - 4z &=& 13 \\ 4x + 11y &=& 37 \end{array} \right. \longrightarrow \begin{array}{l} L_1 \\ L_2' = L_2 - L_1 \\ L_3' = L_3 - 2L_1 \\ L_4' = L_4 - 4L_1 \end{array} \left\{ \begin{array}{rcl} x + 2y - 3z &=& 4 \\ \boxed{}\ y + 4z &=& 7 \\ \boxed{}\ y + 2z &=& 5 \\ \boxed{}\ 3y + 12z &=& 21 \end{array} \right.$$

Il s'agit maintenant d'échelonner les trois dernières équations. Le pivot est à présent le cœfficient de y dans l'équation L_2' (c'est-à-dire 1).

$$\begin{array}{l} L_1 \\ L_2' \\ L_3' - L_2' \\ L_4' - 3L_2' \end{array} \left\{ \begin{array}{rcl} x + 2y - 3z &=& 4 \\ \boxed{}\ y + 4z &=& 7 \\ \boxed{}\ 2z &=& 2 \\ \boxed{0} &=& 0 \end{array} \right.$$

Le système est maintenant sous forme échelonnée. On a immédiatement $z = 1$; en reportant dans L_2' : $y = 7 - 4 = 3$ et dans L_1 : $x = 4 - 6 + 3 = 1$. Le système admet donc la solution unique $x = 1$, $y = 3$, $z = 1$.

Exemple 4. –

$$\begin{array}{l} L_1 \\ L_2 \\ L_3 \end{array} \left\{ \begin{array}{rcl} x - 3y + 4z - 2w &=& 5 \\ x - y + 9z - w &=& 7 \\ x - 2y + 7z - 2w &=& 9 \end{array} \right.$$

$$\begin{array}{l} L_1 \\ L_2' = L_2 - L_1 \\ L_3' = L_3 - L_1 \end{array} \left\{ \begin{array}{rcl} x - 3y + 4z - 2w &=& 5 \\ 2y + 5z + w &=& 2 \\ y + 3z &=& 4 \end{array} \right. \longrightarrow \begin{array}{l} L_1 \\ L_2' \\ 2L_3' - L_2' \end{array} \left\{ \begin{array}{|r|l} \hline x - 3y + 4z & -2w = 5 \\ 2y + 5z & + w = 2 \\ z & - w = 6 \\ \hline \end{array} \right.$$

Le système est sous forme échelonnée. En donnant à w une valeur arbitraire λ, on met en évidence le système dont la matrice est la matrice encadrée, qui admet une solution unique pour chaque choix de λ :

$$\left\{ \begin{array}{|r|l} \hline x - 3y + 4z & = 5 + 2\lambda \\ 2y + 5z & = 2 - \lambda \\ z & = 6 + \lambda \\ \hline \end{array} \right.$$

En résolvant, on trouve :

$$x = -26 - 8\lambda \quad , \quad y = -9 - 3\lambda \quad , \quad z = 4 + \lambda \quad , \quad w = \lambda$$

$$x = -61 - 11\lambda \quad , \quad y = -14 - 3\lambda \quad , \quad z = 6 + \lambda \quad , \quad w = \lambda$$

Le système admet donc une infinité à un paramètre de solutions.

Méthode d'élimination pour la résolution d'un système d'équations linéaires

Les exemples que nous venons de traiter illustrent suffisamment la méthode pour résoudre un système d'équations linéaires.

1. Il faut tout d'abord *préparer* le système en échangeant éventuellement l'ordre des équations et des variables de manière à ce que le pivot soit non nul [1] ;

2. On échelonne ensuite le système.
   Deux cas peuvent se donner :

   (a) : *il se présente une équation du type :*

   $$0x_1 + 0x_2 + \cdots + 0x_n = b \quad \text{avec} \quad b \neq 0 \tag{$*$}$$

   Dans ce cas, il n'y a pas de solution : on dit que le système est ***incompatible.***

   (b) : *il n'y a pas d'équation du type (*).*

   S'il se présente une équation du type

   $$0\,x_1 + 0\,x_2 + \cdots + 0\,x_n = 0$$

   elle peut être écartée. Une fois éliminées toutes ces équations, on aboutit à un système du type :

   $$\left\{\begin{array}{|lllll|l} \hline \boxed{a_{11}}\,x_1 + a_{12}\,x_2 + \cdots\cdots\cdots & & & & & \cdots\cdots + a_{1n}\,x_n = b_1 \\ & \cdots\cdots\cdots\cdots & & & & \cdots\cdots\cdots\cdots \\ & & \boxed{a_{1r}}\,x_r + a_{1\,r+1}\,x_{r+1} + & & & \cdots\cdots + a_{in}\,x_n = b_i \\ & & & \cdots\cdots & & \cdots\cdots\cdots\cdots \\ & & & & \boxed{a_{ps}}\,x_s + & \cdots\cdots + a_{pn}\,x_n = b_p \\ \hline \end{array}\right.$$

   où les cœfficients encadrés (ceux qui sont au début de chaque équation) sont tous non nuls.
   On met ainsi en évidence une matrice échelonnée dont la dernière ligne n'a qu'un élément non nul (cf. matrice encadrée).
   Les inconnues qui apparaissent au début de chaque équation (celles correspondantes aux cœfficients encadrés) sont dites ***inconnues principales***, les autres (s'il y en a) ***variables libres*** (dans les exemples 2 et 3 il n'y a pas de variables libres ; dans l'exemple 4., x, y, z sont les inconnues principales, w la variable libre).

Deux cas sont possibles :

i) Il n'y a pas de variable libre.

[1] Nous verrons en Appendice A.4 les précautions qu'il faut prendre dans le choix du pivot, de manière à contrôler les erreurs dues aux approximations numériques dans les calculs.

Alors le système admet ***une et une seule solution*** que l'on obtient en résolvant d'abord la dernière équation et en remontant jusqu'à la première.

ii) Il y a des variables libres.

On donne alors aux variables libres des valeurs arbitraires $\lambda_1, \ldots, \lambda_m$, on porte les λ_i au second membre et l'on est ramené au cas précédent : il existe alors une et une seule solution pour chaque choix de $\lambda_1, \ldots, \lambda_m$ c'est-à-dire, ***une infinité de solutions dépendante de m paramètres*** (m = nombre des variables libres).

Exemple :

$$\begin{array}{l} {\scriptstyle L_1} \\ {\scriptstyle L_2} \\ {\scriptstyle L_3} \end{array} \left\{ \begin{array}{l} x + 2y - 2z + 3w = 2 \\ 2x + 4y - 3z + 4w = 5 \\ 5x + 10y - 8z + 11w = 12 \end{array} \right. \qquad \begin{array}{l} {\scriptstyle L_1} \\ {\scriptstyle L_2' = L_2 - 2L_1} \\ {\scriptstyle L_3' = L_3 - 5L_1} \end{array} \left\{ \begin{array}{r} x + 2y - 2z + 3w = 2 \\ z - 2w = 1 \\ 2z - 4w = 2 \end{array} \right.$$

$$\begin{array}{l} {\scriptstyle L_1} \\ {\scriptstyle L_2 - 2L_1} \\ {\scriptstyle L_3' - 2L_2'} \end{array} \left\{ \begin{array}{r} x + 2y - 2z + 3w = 2 \\ z - 2w = 1 \\ 0 = 0 \end{array} \right.$$

Aussi, le système est-il équivalent au système suivant :

$$\left\{ \begin{array}{|rr|l} \boxed{x} + 2y - 2z & & +3w = 2 \\ & \boxed{z} & -2w = 1 \end{array} \right.$$

x et z sont les inconnues principales, y et w les variables libres. En posant $y = \lambda$ et $w = \mu$, on a :

$$\left\{ \begin{array}{|r|l} x - 2z & = 2 - 2\lambda - 3\mu \\ z & = 1 \qquad + 2\mu \end{array} \right.$$

d'où les solutions $x = 4 - 2\lambda + \mu, \quad y = \lambda, \quad z = 1 + 2\mu, \qquad w = \mu.$

Exercices 1. 2. 3.

2.2 Cas des systèmes linéaires homogènes

On appelle système linéaire homogène un système d'équations linéaires dont le second membre est nul :

$$\left\{ \begin{array}{l} a_{11}\, x_1 + a_{12}\, x_2 + \cdots + a_{1n}\, x_n = 0 \\ a_{21}\, x_1 + a_{22}\, x_2 + \cdots + a_{2n}\, x_n = 0 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{p1}\, x_1 + a_{p2}\, x_2 + \cdots + a_{pn}\, x_n = 0 \end{array} \right.$$

Un tel système a toujours, bien entendu, au moins une solution, la solution nulle : $x_1 = x_2 = \cdots = x_n = 0$, dite solution triviale. Il est intéressant de savoir si le système admet des solutions non nulles. A cet effet, on a :

Théorème 2.2 – *L'ensemble des solutions d'un système linéaire homogène à n inconnues est un sous-espace vectoriel de K^n. Si le système sous forme échelonnée comporte k équations, l'espace des solutions est de dimension $n-k$.*
En particulier un système homogène avec plus d'inconnues que d'équations $(n > p)$ admet des solutions non nulles.

REMARQUE –La dimension de l'espace des solutions est égale au nombre de variables libres qui apparaissent dans la forme échelonnée.

Démonstration : La vérification de la stabilité des lois pour l'ensemble des solutions est très facile et elle est laissée en exercice.
Supposons avoir mis le système sous forme échelonnée et avoir trouvé k équations. Il y aura donc k inconnues principales et $n-k$ variables libres. Quitte à changer la numérotation des x_i, on peut supposer que les variables libres sont $x_1, \ldots, x_{n-k}$ et les variables principales $x_{n-k+1}, \ldots, x_n$. Une solution sera donc du type :

$$S = (\lambda_1, \ldots, \lambda_{n-k},\ x_{n-k+1}, \ldots, x_n)$$

où les $x_{n-k+1}, \ldots, x_n$ s'expriment linéairement à l'aide des λ_i :

$$\begin{array}{rl} x_{n-k+1} & = a_1\,\lambda_1 \,+\, \cdots \,+\, a_{n-k}\,\lambda_{n-k} \\ x_{n-k+2} & = b_1\,\lambda_1 \,+\, \cdots \,+\, b_{n-k}\,\lambda_{n-k} \\ \cdots\cdots & \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ x_n & = l_1\,\lambda_1 \,+\, \cdots \,+\, l_{n-k}\,\lambda_{n-k} \end{array}$$

donc la solution générale sera le n-uplet

$$S = \left(\lambda_1, \ldots, \lambda_{n-k},\ \sum_{i=1}^{n-k} a_i\,\lambda_i,\ \sum_{i=1}^{n-k} b_i\,\lambda_i, \ldots,\ \sum_{i=1}^{n-k} l_i\,\lambda_i\right)$$

que l'on peut écrire sous la forme :

$$\begin{array}{rl} S = & \lambda_1\,(1, 0, \ldots, 0,\ a_1,\ b_1, \ldots, l_1) + \lambda_2\,(0, 1, 0, \ldots, 0,\ a_2,\ b_2, \ldots, l_2) \\ & \qquad \cdots + \lambda_{n-k}\,(0, 0, \ldots, 1, a_{n-k}, b_{n-k}, \ldots, l_{n-k}) \end{array}$$

On met aussi en évidence $n-k$ solutions :

$$\begin{array}{c} (\,1, 0, \ldots, 0, \quad a_1, \quad b_1, \quad \ldots\ldots, l_1\,) \\ (\,0, 1, \ldots, 0, \quad a_2, \quad b_2, \quad \ldots\ldots, l_2\,) \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ (\,0, 0, \ldots, 1,\ a_{n-k},\ b_{n-k}, \ldots, l_{n-k}\,) \end{array}$$

(obtenues en donnant au $(n-k)$-uplet $(\lambda_1, \ldots, \lambda_{n-k})$ successivement les valeurs $(1, 0, \ldots, 0)$, $(0, 1, \ldots, 0) \ldots (0, \ldots, 0, 1)$). Elles forment un système de générateurs, car la solution générale est combinaison linéaire de ces solutions. Il est facile de voir qu'il s'agit d'une famille libre et donc d'une base de l'espace des solutions[2]. $\square$

[2]Comme nous le verrons au paragraphe 4, cela tient du fait que la "*matrice engendrée*" est échelonnée.

Remarque – Comme on l'a vu au cours de la démonstration, on obtient une base de solutions en donnant aux variables libres successivement les valeurs $(1,0,\ldots,0)$, $(0,1,0,\ldots,0)\ldots(0,\ldots,0,1)$.

Exemple :

$$\begin{cases} x+2\,y-3\,z+t-w=0 \\ x+3\,y-4\,z \qquad +w=0 \\ 2\,x+5\,y-7\,z+t \qquad =0 \end{cases}$$

On ramène facilement le système à la forme échelonnée :

$$\begin{cases} \boxed{\begin{matrix} x+2\,y \\ y \end{matrix}} \begin{matrix} -3\,z+t-\ \ w=0 \\ -\ \ z-t+2w=0 \end{matrix} \end{cases}$$

Les variables libres sont z, t, w ; l'ensemble des solutions est donc un sous-espace vectoriel de dimension 3 de $\mathbb{R}^5$. La solution générale est :

$$x=\lambda-3\,\mu+5\,\nu\,,\quad y=\lambda+\mu-2\,\nu\,,\quad z=\lambda\,,\quad t=\mu\,,\quad w=\nu$$

c'est-à-dire :

$$\begin{aligned} S &= (\lambda-3\,\mu+5\,\nu,\ \lambda+\mu-2\,\nu,\ \lambda,\ \mu,\ \nu) \\ &= \lambda\,(1,1,1,0,0)+\mu\,(-3,1,0,1,0)+\nu\,(5,-2,0,0,1) \end{aligned}$$

Donc une base de l'espace des solutions est :

$$(1,1,1,0,0),\ (-3,1,0,1,0),\ (5,-2,0,0,1)$$

Exercices 4. 5. 6.

2.3 Application aux familles libres et aux familles génératrices

I $\boxed{\textit{La famille } \{v_1,\ldots,v_p\} \textit{ est-elle libre ?}}$

Exemple : *Vérifier que les vecteurs de* $\mathbb{R}^3$

$$v_1=(1,-2,-3)\,,\quad v_2=(2,3,-1)\,,\quad v_3=(3,2,1)$$

forment une famille libre.

Il s'agit de savoir si $x_1\,v_1+x_2\,v_2+x_3\,v_3=0 \implies x_1=0,\ x_2=0,\ x_3=0$ c'est-à-dire si le système :

$$\begin{cases} x_1+2\,x_2+3\,x_3= & 0 \\ -2\,x_1+3\,x_2+2\,x_3= & 0 \\ -3\,x_1-\ \ x_2+\ \ x_3= & 0 \end{cases}$$

admet comme seule solution, la solution nulle. On a :

$$\begin{array}{c} L_1 \\ L_2'{=}L_2{+}2L_1 \\ L_3'{=}L_3{+}3L_1 \end{array} \begin{cases} x_1+2\,x_2+\ \ 3\,x_3=0 \\ \qquad 7\,x_2+\ \ 8\,x_3=0 \\ \qquad 5\,x_2+10\,x_3=0 \end{cases} \qquad \begin{array}{c} L_1 \\ L_2' \\ 7L_3-5L_2' \end{array} \begin{cases} x_1+2\,x_2+\ \ 3\,x_3=0 \\ \qquad 7\,x_2+\ \ 8\,x_3=0 \\ \qquad\qquad 30\,x_3=0 \end{cases}$$

d'où : $x_1=0,\ x_2=0,\ x_3=0$. La famille est donc libre.

II **Détermination les relations linéaires liant une famille de vecteurs**

Exemple : *Déterminer les relations linéaires liant les vecteurs de* $\mathbb{R}^4$ *suivants :*

$$v_1 = (1,1,0,2) \quad v_2 = (-1,0,2,1) \quad v_3 = (0,1,2,3) \quad v_4 = (1,3,4,8)$$

Il s'agit de déterminer les x_i $(i = 1, \ldots, 4)$ tels que

$$x_1\, v_1 + x_2\, v_2 + x_3\, v_3 + x_4\, v_4 = 0$$

c'est-à-dire les solutions du système :

$$\begin{array}{l} L_1 \\ L_2 \\ L_3 \\ L_4 \end{array}\left\{\begin{array}{rcl} x_1 - x_2 \qquad + x_4 &=& 0 \\ x_1 \qquad + x_3 + 3\,x_4 &=& 0 \\ 2\,x_2 + 2\,x_3 + 4\,x_4 &=& 0 \\ 2\,x_1 + x_2 + 3\,x_3 + 8\,x_4 &=& 0 \end{array}\right. \qquad \begin{array}{l} L_1 \\ L_2' = L_2 - L_1 \\ L_3 \\ L_4' = L_4 - 2L_1 \end{array}\left\{\begin{array}{rcl} x_1 - x_2 \qquad + x_4 &=& 0 \\ x_2 + x_3 + 2\,x_4 &=& 0 \\ 2\,x_2 + 2\,x_3 + 4\,x_4 &=& 0 \\ 3\,x_2 + 3\,x_3 + 6\,x_4 &=& 0 \end{array}\right.$$

$$\begin{array}{l} L_1 \\ L_2' \\ L_3 - 2L_2' \\ L_4' - 3L_2' \end{array}\left\{\begin{array}{rcl} \boxed{\begin{array}{r} x_1 - x_2 \\ x_2 \end{array}} \begin{array}{l} + x_4 \\ + x_3 + 2\,x_4 \end{array} & \begin{array}{l} = 0 \\ = 0 \end{array} \\ 0 &= 0 \\ 0 &= 0 \end{array}\right.$$

En posant $x_3 = \lambda \quad x_4 = \mu$, on trouve :

$$\begin{aligned} x_2 &= -\lambda - 2\,\mu \\ x_1 &= -\lambda - 3\,\mu \end{aligned}$$

d'où les relations :

$$(\lambda + 3\,\mu)\, v_1 + (\lambda + 2\,\mu)\, v_2 - \lambda\, v_3 - \mu\, v_4 = 0 \qquad (\lambda,\, \mu \in \mathbb{R})$$

En donnant au couple (λ, μ) les valeurs $(1, 0)$ et $(0, 1)$ on obtient les solutions indépendantes :

$$v_1 + v_2 - v_3 = 0 \qquad \text{et} \qquad 3\,v_1 + 2\,v_2 - v_4 = 0.$$

III **Vérifier si un vecteur appartient à l'espace engendré par $\{v_1, ..., v_p\}$ et déterminer, le cas échéant son expression en fonction de $v_1, ..., v_p$**

Exemple : *Soit* $v = (3, 9, -4, -2)$. *Montrer que* $v \in \mathrm{Vect}\{v_1,\, v_2,\, v_3\}$, *où on a* $v_1 = (1, -2, 0, 3)$, $v_2 = (2, 3, 0, -1)$, $v_3 = (2, -1, 2, 1)$ *et exprimer* v *en fonction de* v_1, v_2, v_3.

Il faut que l'on puisse trouver $x_1,\, x_2,\, x_3,\, x_4 \in \mathbb{R}$ tels que :

$$v = x_1\, v_1 + x_2\, v_2 + x_3\, v_3;$$

c'est-à-dire il faut que le système

$$\left\{\begin{array}{rcr} x_1 + 2\,x_2 + 2\,x_3 &=& 3 \\ -2\,x_1 + 3\,x_2 - x_3 &=& 9 \\ 2\,x_3 &=& -4 \\ 3\,x_1 - x_2 + x_3 &=& -2 \end{array}\right.$$

admette au moins une solution. On trouve facilement : $x_1 = 1,\ x_2 = 3,\ x_3 = -2$. Donc $v \in \mathrm{Vect}\{v_1,\, v_2,\, v_3\}$ et $v = v_1 + 3\,v_2 - 2\,v_3$.

IV La famille $\{v_1, \ldots, v_p\}$ est elle génératrice ?

Exemple : *Montrer que la famille* $v_1 = (1,1,1)$, $v_2 = (0,1,1)$, $v_3 = (0,1,-1)$ *engendre* $\mathbb{R}^3$*.*

Il faut montrer que *tout* vecteur $v \in \mathbb{R}^3$ appartient à $\mathrm{Vect}\{v_1, v_2, v_3\}$ c'est-à-dire que $\forall (a, b, c) \in \mathbb{R}^3$, le système $(a, b, c) = x_1 v_1 + x_2 v_2 + x_3 v_3$ admet au moins une solution. On a :

$$\begin{array}{l} {\scriptstyle L_1} \\ {\scriptstyle L_2} \\ {\scriptstyle L_3} \end{array} \left\{ \begin{array}{l} x_1 \qquad\qquad = a \\ x_1 + x_2 + x_3 = b \\ x_1 + x_2 - x_3 = c \end{array} \right. \longrightarrow \begin{array}{l} {\scriptstyle L_1} \\ {\scriptstyle L_2 - L_1} \\ {\scriptstyle L_3 - L_1} \end{array} \left\{ \begin{array}{l} x_1 \qquad\qquad = a \\ \qquad x_2 + x_3 = b - a \\ \qquad x_2 - x_3 = c - a \end{array} \right.$$

$$\longrightarrow \left\{ \begin{array}{l} x_1 = a \\ x_2 = \dfrac{b + c - 2a}{2} \\ x_3 = \dfrac{b - c}{2} \end{array} \right.$$

On a donc une solution $x_1\, x_2$, x_3 *pour tous* a, b, c ; la famille $\{v_1, v_2, v_3\}$ est donc génératrice.

V Déterminer une base de $F \cap G$.

Exemple : *Soient F et G les sous-espaces vectoriels de $\mathbb{R}^4$ engendrés respectivement par $\{v_1, v_2\}$ et $\{w_1, w_2\}$, où :*

$$v_1 = (1,-1,0,2), \;\; v_2 = (2,1,3,1), \;\; w_1 = (1,1,1,1), \;\; w_2 = (3,-4,4,2).$$

Déterminer une base de $F \cap G$.

Un vecteur quelconque $v \in F$ est du type $v = \lambda\, v_1 + \mu\, v_2$ et un vecteur quelconque $w \in G$ est du type $w = \alpha\, w_1 + \beta\, w_2$. Il faut donc trouver λ, μ, α, $\beta \in \mathbb{R}$, tels que :

$$\lambda\, v_1 + \mu\, v_2 = \alpha\, w_1 + \beta\, w_2$$

c'est-à-dire :

$$(\lambda + 2\,\mu, \; -\lambda + \mu, \; 3\,\mu, \; 2\,\lambda + \mu) = (\alpha + 3\,\beta, \; \alpha - 4\,\beta, \; \alpha + 4\,\beta, \; \alpha + 2\,\beta) \qquad (*)$$

ce qui donne le système :

$$\left\{ \begin{array}{rl} \lambda + 2\,\mu & = \alpha + 3\,\beta \\ -\lambda + \;\; \mu & = \alpha - 4\,\beta \\ 3\,\mu & = \alpha + 4\,\beta \\ 2\,\lambda + \mu & = \alpha + 2\,\beta \end{array} \right.$$

dont les inconnues sont λ, μ, α, β. En portant α et β à premier membre :

$$\begin{array}{l} L_1 \\ L_2 \\ L_3 \\ L_4 \end{array}\left\{\begin{array}{r} \lambda+2\,\mu-\alpha-3\,\beta=0 \\ -\lambda+\quad\mu-\alpha+4\,\beta=0 \\ 3\,\mu-\alpha-4\,\beta=0 \\ 2\,\lambda+\quad\mu-\alpha-2\,\beta=0 \end{array}\right. \longrightarrow \begin{array}{l} L_1 \\ L_1+L_2 \\ L_3 \\ L_4-2\,L_1 \end{array}\left\{\begin{array}{r} \lambda+2\,\mu-\quad\alpha-3\,\beta=0 \\ 3\,\mu-2\,\alpha+\quad\beta=0 \\ 3\,\mu-\quad\alpha-4\,\beta=0 \\ -3\,\mu+\quad\alpha+4\,\beta=0 \end{array}\right.$$

$$\longrightarrow \left\{\begin{array}{r} \boxed{\lambda}+2\,\mu-\quad\alpha\quad-3\,\beta=0 \\ \boxed{3\,\mu}\;-2\,\alpha\;+\quad\beta=0 \\ \boxed{\alpha}\quad-5\,\beta=0 \end{array}\right.$$

λ, μ, α sont les inconnues principales, β la variable libre. On trouve facilement, en fonction de β : $\alpha=5\,\beta,\ \mu=3\,\beta,\ \lambda=2\,\beta$.
$F\cap G$ est donc défini par l'un des deux membres de (*) où λ, μ, α sont exprimés en fonction de β. On trouve : $(8\,\beta,\ \beta,\ 9\,\beta,\ 7\,\beta)$ ce qui veut dire que $F\cap G$ est engendré par le vecteur $(8,1,9,7)$.

VI *Équations d'un sous-espace*

Le problème est le suivant. On se donne une famille de vecteurs $\mathcal{F}=\{v_1,\ldots,v_p\}$ et soit F l'espace engendré. Il s'agit de déterminer un système d'équations linéaires dont les solutions sont exactement les vecteurs de F.

Exemple : *Déterminer les équations du sous-espace de $\mathbb{R}^5$ engendré par les vecteurs $v_1=(1,3,-2,2,3)$, $v_2=(1,4,-3,4,2)$, $v_3=(2,3,-1,-2,9)$.*

Une équation quelconque du système est du type :

$$a\,x_1+b\,x_2+c\,x_3+d\,x_4+e\,x_5=0$$

En imposant que les composantes de v_1, v_2, v_3 vérifient cette équation, on trouve le système :

$$\begin{array}{ll} (\text{pour } v_1) & L_1 \\ (\text{pour } v_2) & L_2 \\ (\text{pour } v_3) & L_3 \end{array}\left\{\begin{array}{r} a+3\,b-2\,c+2\,d+3\,e=0 \\ a+4\,b-3\,c+4\,d+2\,e=0 \\ 2\,a+3\,b-\quad c-2\,d+9\,e=0 \end{array}\right.$$

$$\longrightarrow \begin{array}{l} L_1 \\ L_2-L_1 \\ L_3-2L_1 \end{array}\left\{\begin{array}{r} a+3\,b-2\,c+2\,d+3\,e=0 \\ b-\quad c+2\,d-\quad e=0 \\ -3\,b+3\,c-6\,d+3\,e=0 \end{array}\right. \longrightarrow \left\{\begin{array}{r} \boxed{a}\;+3\,b\;-2\,c+2\,d+3\,e=0 \\ \boxed{b}\;-\quad c+2\,d-\quad e=0 \\ 0\qquad=0 \end{array}\right.$$

En donnant aux variables libres, c, d, e les valeurs $\{1,0,0\}$, $\{0,1,0\}$ et $\{0,0,1\}$ on trouve le système :

$$\left\{\begin{array}{r} -x_1+\quad x_2+x_3=0 \\ 4\,x_1-2\,x_2+x_4=0 \\ -6\,x_1+\quad x_2+x_5=0 \end{array}\right.$$

Exercices : 7. 8. 9. 10. 11.

2.4 Utilisation pratique de la méthode du pivot

Comme on vient de le voir, les différents problèmes sur la dépendance linéaire reposent sur l'étude d'un système d'équations linéaires. Cependant dans la pratique, la méthode du pivot fournit une technique qui permet de s'affranchir de l'étude du système et de travailler directement sur les familles de vecteurs[3].

Théorème 2.3 – ***(Opérations élémentaires sur une famille de vecteurs).***
Soit $\{v_1, \ldots, v_p\}$ une famille de vecteurs. L'espace qu'ils engendrent ne change pas si l'on effectue sur les vecteurs de la famille l'une des «opérations élémentaires» suivantes :

a) changer l'ordre des vecteurs ;
b) multiplier un vecteur par un scalaire non nul ;
c) ajouter à un vecteur une combinaison linéaire des autres vecteurs.

Démonstration :

a) Evidente.

b) D'après a) il suffit de montrer que si

$$E = \mathrm{Vect}\{v_1, ..., v_p\} \quad \text{et} \quad F = \mathrm{Vect}\{k\, v_1, ..., v_p\} \quad (\text{avec } k \neq 0),$$

alors on a : $E = F$.
Soit $y \in F$: $y = \mu_1\,(k\, v_1) + \mu_2\, v_2 + \cdots + \mu_p\, v_p$; on a immédiatement $y = (\mu_1\, k)\, v_1 + \mu_2\, v_2 + \cdots + \mu_p\, v_p$ c'est-à-dire $y \in E$.
Réciproquement, si $y \in E$, c'est-à-dire y est de la forme $y = \lambda_1\, v_1 + \cdots + \lambda_p\, v_p$, on a :

$$y = \frac{\lambda_1}{k}\,(k\, v_1) + \lambda_2\, v_2 + \cdots + \lambda_p\, v_p$$

c'est-à-dire $y \in F$.

c) D'après a) et b) il suffit de montrer que si l'on ajoute à v_1 le vecteur v_2, l'espace engendré ne change pas. C'est-à-dire que $\mathrm{Vect}\{v_1,\, v_2, \ldots, v_p\} = \mathrm{Vect}\{v_1 + v_2,\ v_2, \ldots, v_p\}$.
Soit $y \in \mathrm{Vect}\{v_1 + v_2,\ v_2, \ldots, v_p\}$:

$$y = \mu_1\,(v_1 + v_2) + \mu_2\, v_2 + \cdots + \mu_p\, v_p$$

On a :

$$y = \mu_1\, v_1 + (\mu_2 + \mu_1)\, v_2 + \cdots + \mu_p\, v_p$$

c'est-à-dire : $y \in \mathrm{Vect}\{v_1,\, v_2, \ldots, v_p\}$.

Réciproquement, soit $y \in \mathrm{Vect}\{v_1,\, v_2, \ldots, v_p\}$: $y = \lambda_1\, v_1 + \cdots + \lambda_p\, v_p$. On peut écrire :

$$y = \lambda_1\,(v_1 + v_2) + (\lambda_2 - \lambda_1)\, v_2 + \lambda_3\, v_3 + \cdots + \lambda_p\, v_p,$$

et donc $y \in \mathrm{Vect}\{v_1 + v_2,\ v_2, \ldots, v_p\}$. $\square$

[3] Bien que l'algorithme que nous allons expliquer soit extrêmement simple et pratique, l'on ne doit pas perdre de vue les exemples du paragraphe 3. : on peut oublier une technique, il ne faut pas oublier ce qui est la démarche naturelle du raisonnement.

Soit E un espace vectoriel de dimension finie, $\{e_1, \ldots, e_n\}$ une base de E (que nous noterons dans la suite simplement $\{e_i\}$). Comme nous l'avons vu au chapitre 1, (cf. proposition 1.13), tout vecteur $v \in E$ peut être caractérisé par un n-uplet d'éléments de K (ses composants dans la base $\{e_i\}$).

Définition 2.4 – *Soit $\{v_1, \ldots, v_p\}$ une famille de vecteurs de E. On appelle* ***matrice engendrée*** *par les vecteurs v_i dans la base $\{e_i\}$ (ou plus simplement : matrice des vecteurs v_i) la matrice dont les lignes sont les composantes des vecteurs $v_1, \ldots, v_p$ dans la base $\{e_i\}$*[4].

Il est clair qu'aux opérations élémentaires du théorème 2.3 correspondent les opérations élémentaires sur les lignes de la matrice engendrée, que nous avons vues au §1 (changer les lignes entre elles, multiplier une ligne par un scalaire non nul, ajouter à une ligne une combinaison linéaire des autres lignes).
Comme nous l'avons vu au paragraphe 1., on peut toujours effectuer des opérations élémentaires sur une matrice de manière à la mettre sous forme échelonnée. Du théorème 2.3 on a immédiatement :

Corollaire 2.5 – *Les vecteurs lignes d'une matrice et les vecteurs lignes de sa réduite échelonnée engendrent le même espace vectoriel.*

Le résultat suivant donne la clé de la méthode :

Théorème 2.6 – *Soit $\{v_1, \ldots, v_p\}$ une famille de vecteurs, A la matrice engendrée (dans une base quelconque) et A' une réduite échelonnée de A. Alors les lignes non nulles de A' donnent une base de* $\mathrm{Vect}\{v_1, ..., v_p\}$.

Démonstration : Remarquons d'abord que les lignes non nulles de A' fournissent une famille génératrice de $\mathrm{Vect}\{v_1, ..., v_p\}$, d'après le corollaire 2.5. Il suffira donc de montrer qu'elles forment un système libre. La démonstration (cf. exercice 15) est plus longue à écrire qu'à comprendre. Vérifions-la dans un cas particulier. Soit :

$$A' = \begin{array}{c} \begin{bmatrix} 4 & a & b & c & d \\ 0 & 0 & 2 & c' & d' \\ 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \end{array} \begin{array}{c} v_1 \\ v_2 \\ v_3 \\ v_4 \end{array}$$

Vérifions que $\{v_1,\, v_2,\, v_3\}$ est une famille libre, c'est-à-dire que :

$$x_1\, v_1 + x_2\, v_2 + x_3\, v_3 = 0 \quad \Longrightarrow \quad x_1 = 0,\ x_2 = 0,\ x_3 = 0.$$

[4] Au chapitre suivant, nous construirons des matrices en mettant les composantes de certains vecteurs en *colonnes*. On aurait pu adopter ici la même convention avec les changements opportuns dans les énoncés des résultats. Cependant nous avons préféré adopter cette définition d'une part pour suivre de près la théorie développée au paragraphe 1. (où la notation en ligne s'impose) et d'autre part pour souligner la signification différente des matrices que nous introduirons (et éviter, par exemple, que l'on soit tenté d'effectuer des opérations élémentaires sur les matrices qui seront introduites au chapitre suivant).

Cela donne le système :

$$\begin{cases} 4\,x_1 & = 0 \\ a\,x_1 & = 0 \\ b\,x_1 + 2\,x_2 & = 0 \\ c\,x_1 + c'\,x_2 & = 0 \\ d\,x_1 + d'\,x_2 - x_3 & = 0 \end{cases}$$

La première équation donne $x_1 = 0$; en reportant dans la troisième on a $x_2 = 0$ et, en reportant dans la cinquième, $x_3 = 0$.
Il est facile de se convaincre que, dans le cas général, on obtient un système échelonné qui donne, de proche en proche : $x_1 = 0$, puis $x_2 = 0, \ldots$, etc. jusqu'à $x_p = 0$.

APPLICATIONS

I

> *Extraire une base d'une famille génératrice et déterminer les relations liant les vecteurs.*

Exemple :
Déterminer une base du sous-espace de $\mathbb{R}^4$ *engendré par les vecteurs*

$$v_1 = (1,1,0,-1)\,,\ v_2 = (-1,1,1,0)\,,\ v_3 = (0,2,1,-1)$$

et les éventuelles relations linéaires.

$$\begin{array}{c} \begin{bmatrix} 1 & 1 & 0 & -1 \\ -1 & 1 & 1 & 0 \\ 0 & 2 & 1 & -1 \end{bmatrix} \end{array} \begin{array}{l} v_1 \\ v_2 \\ v_3 \end{array} \longrightarrow \begin{bmatrix} 1 & 1 & 0 & -1 \\ 0 & 2 & 1 & -1 \\ 0 & 2 & 1 & -1 \end{bmatrix} \begin{array}{l} v_1 \\ v_2' = v_1 + v_2 \\ v_3 \end{array}$$

$$\longrightarrow \begin{bmatrix} 1 & 1 & 0 & -1 \\ 0 & 2 & 1 & -1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{array}{l} v_1 \\ v_2' \\ v_3' = v_2' - v_3 \end{array}$$

Ainsi $v_1 = (1,1,0,-1)$ et $v_2' = (0,2,1,-1)$ forment une base de $\mathrm{Vect}\{v_1,\, v_2,\, v_3\}$.
D'autre part $v_3' = 0$ c'est-à-dire $v_2' - v_3 = 0$ ou encore $v_1 + v_2 - v_3 = 0$, ce qui donne la relation linéaire liant $\{v_1,\, v_2,\, v_3\}$.

II

> *Compléter une famille libre en une base.*
> *Détermination d'un supplémentaire.*

Exemple :
Montrer que les vecteurs

$$v_1 = (1,2,-1,0,1), \quad v_2 = (2,1,1,1,1), \quad v_3 = (3,2,0,1,2)$$

forment une famille libre de $\mathbb{R}^5$*. Déterminer deux vecteurs* w_1, w_2 *de* $\mathbb{R}^5$ *de manière à ce que* $\{v_1, v_2, v_3, w_1, w_2\}$ *soit une base de* $\mathbb{R}^5$*.*

$$\begin{array}{c}\begin{bmatrix}1 & 2 & -1 & 0 & 1\\ 2 & 1 & 1 & 1 & 1\\ 3 & 2 & 0 & 1 & 2\end{bmatrix}\end{array}\begin{array}{l}v_1\\ v_2\\ v_3\end{array} \longrightarrow \begin{bmatrix}1 & 2 & -1 & 0 & 1\\ 0 & -3 & 3 & 1 & -1\\ 0 & -4 & 3 & 1 & -1\end{bmatrix}\begin{array}{l}v_1\\ v_2'=v_2-2v_1\\ v_3'=v_3-3v_1\end{array}$$

$$\longrightarrow \begin{bmatrix}1 & 2 & -1 & 0 & 1\\ 0 & -3 & 3 & 1 & -1\\ 0 & 0 & -3 & -1 & 1\end{bmatrix}\begin{array}{l}v_1\\ v_2'\\ 3v_3'-4v_2'=v_3''\end{array}$$

Puisque l'on n'obtient pas de lignes nulles dans la matrice échelonnée, la dimension de l'espace F engendré par v_1, v_2, v_3 est 3. $\{v_1, v_2, v_3\}$ est donc une famille libre, car elle est une base de F (notons que $\{v_1, v_2', v_3''\}$ est aussi une base de F). Considérons la matrice :

$$\begin{bmatrix}1 & 2 & -1 & 0 & 1\\ 0 & -3 & 3 & 1 & -1\\ 0 & 0 & -3 & -1 & 1\\ 0 & 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 0 & 1\end{bmatrix}\begin{array}{l}v_1\\ v_2'\\ v_3''\\ w_1\\ w_2\end{array}$$

Cette matrice est échelonnée, donc la famille $\{v_1, v_2', v_3'', w_1, w_2\}$ est une base de $\mathbb{R}^5$. D'après la proposition 1.30, page 22, $\{w_1, w_2\}$ est une base d'un supplémentaire G de F :

$$\mathbb{R}^5 = F \oplus G$$

Donc $\mathcal{B} = \{\underbrace{v_1,\, v_2,\, v_3}_{\text{base de } F},\ \underbrace{w_1,\, w_2}_{\text{base de } G}\}$ est une base de $\mathbb{R}^5$ (toujours d'après 1.30).

III $\boxed{\textit{Déterminer une base de } F+G.}$

Notons que si $\mathcal{G}_1$ et $\mathcal{G}_2$ sont des familles génératrices de F et G respectivement, la famille $\mathcal{G}_1 \cup \mathcal{G}_2$ engendre $F+G$ (cf. proposition 1.36, page 25). Il suffira donc d'appliquer la technique de l'exemple 1. pour extraire de $\mathcal{G}_1 \cup \mathcal{G}_2$, une base de $F+G$.

Exemple :

Soit $F = \mathrm{Vect}\{v_1, v_2, v_3\}$, *où :* $v_1 = (1, -1, 0, 2)$, $v_2 = (2, 1, 3, 1), v_3 = (4, 5, 9, -1)$ *et* $G = \mathrm{Vect}\{w_1, w_2\}$ *où :* $w_1 = (1, 1, 1, 1)$ *et* $w_2 = (3, -4, 4, 2)$.
Déterminer une base $F + G$.

On a : $F + G = \mathrm{Vect}\{v_1,\, v_2,\, v_3,\, w_1,\, w_2\}$. On commence par extraire une base de F et une base de G.

Base de F :

$$\left[\begin{array}{cccc} 1 & -1 & 0 & 2 \\ 2 & 1 & 3 & 1 \\ 4 & 5 & 9 & -1 \end{array}\right]\begin{array}{l} v_1 \\ v_2 \\ v_3 \end{array} \longrightarrow \left[\begin{array}{cccc} 1 & -1 & 0 & 2 \\ 0 & 3 & 3 & -3 \\ 0 & 9 & 9 & -9 \end{array}\right]\begin{array}{l} v_1 \\ v_2-2\,v_1=v_2' \\ v_3-4\,v_1=v_3' \end{array}$$

$$\longrightarrow \left[\begin{array}{cccc} 1 & -1 & 0 & 2 \\ 0 & 1 & 1 & -1 \\ 0 & 0 & 0 & 0 \end{array}\right]\begin{array}{l} v_1 \\ \frac{v_2'}{3}=v_2'' \\ v_3'-3\,v_2' \end{array}$$

Ainsi $\mathcal{B}_1 = \{v_1,\, v_2''\}$ est une base de F.

Base de G :

w_1 et w_2 sont évidemment libres et donc $\mathcal{B}_2 = \{w_1,\, w_2\}$ est une base de G. Ainsi $F+G = \text{Vect}\{v_1,\, v_2'',\, w_1,\, w_2\}$.

Base de F + G :

$$\left[\begin{array}{cccc} 1 & -1 & 0 & 2 \\ 0 & 1 & 1 & -1 \\ 1 & 1 & 1 & 1 \\ 3 & -4 & 4 & 2 \end{array}\right]\begin{array}{l} v_1 \\ v_2'' \\ w_1 \\ w_2 \end{array} \longrightarrow \left[\begin{array}{cccc} 1 & -1 & 0 & 2 \\ 0 & 1 & 1 & -1 \\ 0 & 2 & 1 & -1 \\ 0 & -1 & 4 & -4 \end{array}\right]\begin{array}{l} v_1 \\ v_2'' \\ u_1=w_1-v_1 \\ u_2=w_2-3v_1 \end{array}$$

$$\longrightarrow \left[\begin{array}{cccc} 1 & -1 & 0 & 2 \\ 0 & 1 & 1 & -1 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 5 & -5 \end{array}\right]\begin{array}{l} v_1 \\ v_2'' \\ u_1-2u_2'' \\ u_2+v_2'' \end{array} \longrightarrow \left[\begin{array}{cccc} 1 & -1 & 0 & 2 \\ 0 & 1 & 1 & -1 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & 0 \end{array}\right]\begin{array}{l} v_1 \\ v_2'' \\ u_1-2v_2'' \\ u_2+v_2''+5\,(u_1-2v_2'') \end{array}$$

Donc une base de $F+G$ est donnée par les vecteurs $(1,-1,0,2)$, $(0,1,1,-1)$ et $(0,0,-1,1)$.

IV Déterminer une base de $F \cap G$

Exemple :

Déterminer une base de $F \cap G$ où F et G sont les espaces de l'exemple III. ci-dessus.

On vient de voir que, dim $F+G=3$, donc :

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

Il suffira donc de déterminer un vecteur non nul de $F \cap G$. Comme on l'a vu ci-dessus, dans la dernière matrice échelonnée le vecteur $u_2 + v_2'' + 5\,(u_1 - 2\,v_2'')$ est nul. On a ainsi, en exprimant u_1 et u_2 en fonction des vecteurs de F et de G :

$$w_2 - 3\,v_1 + v_2'' + 5\,(w_1 - v_1 - 2\,v_2'') = 0$$

Et donc, en portant à premier membre les vecteurs de F et à second membre les vecteurs de G :

$$\underbrace{2\,v_1 - 9\,v_2''}_{\in\, F} = \underbrace{5\,w_1 + w_2}_{\in\, G}$$

Notons z le vecteur donné par l'un des deux membres de cette équation. On a $z \in F$ et $z \in G$ donc $z \in F \cap G$. Ainsi, le vecteur $z = 5\,w_1 + w_2 = (8,1,9,7)$ est une base de $F \cap G$.

Exercices 12. 13. 14. 15.

EXERCICES

1 Résoudre les systèmes suivants :

a) $\left\{\begin{array}{rl} x+2\,y-3\,z &= -1 \\ 3\,x-\;\;y+2\,z &= \;\;7 \\ 8\,x+2\,y-2\,z &= \;\;9 \end{array}\right.$ b) $\left\{\begin{array}{rl} 2\,x+\;\;y-2\,z &= \;\;10 \\ x+\;\;y+4\,z &= -9 \\ 7\,x+5\,y+\;\;z &= \;\;14 \end{array}\right.$

c) $\left\{\begin{array}{rl} x-3\,y+7\,z &= -4 \\ x+2\,y-3\,z &= \;\;6 \\ 7\,x+4\,y-\;\;z &= \;\;22 \end{array}\right.$

2 Déterminer les valeurs de a pour lesquelles le système

$$\left\{\begin{array}{rl} x+\;\;y-\;\;z &= 1 \\ x+2\,y+a\,z &= 2 \\ 2\,x+a\,y+2\,z &= 3 \end{array}\right.$$

1. n'a pas de solution ;
2. a une infinité de solutions ;
3. a une solution unique.

3 Pour quelles valeurs de a, b, c le système suivant admet-il au moins une solution ?

$$\left\{\begin{array}{l} x+2\,y-\;\;3\,z = a \\ 3\,x+8\,y-14\,z = b \\ 2\,x \qquad\;\; +\;\;4\,z = c \end{array}\right.$$

4 Le système suivant

$$\left\{\begin{array}{rl} 2\,x-3\,y+3\,z+t &= 0 \\ x-4\,y \qquad\;\; +t &= 0 \\ x-\;\;y+2\,z+t &= 0 \end{array}\right.$$

admet-il une solution non triviale ?

5 Déterminer la solution générale et une base de l'espace des solutions du système

$$\left\{\begin{array}{rl} x-\;\;y+\;\;z-\;\;t+\;\;w &= 0 \\ x+\;\;y+2\,z-\;\;t \qquad\;\; &= 0 \\ 2\,x-2\,y+3\,z-\;\;t+2\,w &= 0 \\ 4\,x-2\,y+6\,z-3\,t+3\,w &= 0 \end{array}\right.$$

6 Montrer qu'un système homogène dans lequel les cœfficients d'une inconnue sont tous nuls, admet une solution non triviale.

* **7** Montrer, à l'aide de la théorie sur les systèmes homogènes le résultat du corollaire 1.19, 1. : *dans un espace de dimension n, toute famille ayant plus de n vecteurs est liée.*

8

1. Pour quelles valeurs de α les vecteurs

   $$v_1 = (1,-1,0,2),\;\; v_2 = (1,0,1,2),\;\; v_3 = (1,3,5,7),\;\; v_4 = (0,2,3,\alpha)$$

   forment-ils une base de $\mathbb{R}^4$?

2. Dans le cas où la famille est liée, déterminer les relations linéaires qui lient ces vecteurs. Quelle est la dimension de l'espace engendré ?
3. Soit $v = (-2, k, 1, 3)$. Pour quelles valeurs de k, $v \in \mathrm{Vect}\{v_1, v_2, v_3, v_4\}$? Déterminer, dans ce cas, les composantes de v dans une base de $\mathrm{Vect}\{v_1, v_2, v_3, v_4\}$.

9 Montrer que les polynômes $P_1 = 1$, $P_2 = t-1$, $P_3 = (t-1)^2$ forment une base de $\mathbb{R}_2[t]$. Déterminer les coordonnées du vecteur $P = 2\,t^2 - 5\,t + 6$ dans cette base.

10 Déterminer les éventuelles relations linéaires liant les polynômes suivants de $\mathbb{R}_3\,[t]$:

$$P_1 = t^3 + 4\,t^2 - 2\,t + 3,\ \ P_2 = 2\,t^3 + 10\,t^2 - 3\,t + 7,\ \ P_3 = 2\,t^3 + 4\,t^2 - 6\,t + 4.$$

11 Montrer que les matrices $A = \begin{pmatrix} 1 & -2 \\ -2 & 1 \end{pmatrix}$, $B = \begin{pmatrix} 1 & 3 \\ 3 & 6 \end{pmatrix}$, $C = \begin{pmatrix} -1 & 1 \\ 1 & -3 \end{pmatrix}$ forment une base de l'espace vectoriel $\mathcal{S}_2\,(\mathbb{R})$ des matrices symétriques d'ordre 2. Décomposer sur cette base la matrice $M = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$.

12 Soit F le sous-espace vectoriel de $\mathbb{R}^5$ engendré par les vecteurs

$$v_1 = (1, 3, -2, 2, 3),\ \ v_2 = (2, 7, -5, 6, 5),\ \ v_3 = (1, 2, -1, 0, 4)$$

et G le sous-espace vectoriel de $\mathbb{R}^5$ engendré par

$$w_1 = (1, 3, 0, 2, 1),\ \ w_2 = (2, 7, -3, 6, 3),\ \ w_3 = (1, 1, 6, -2, -1).$$

1. Déterminer une base de F, une base de G, une base de $F + G$ et une base de $F \cap G$.
2. Déterminer les équations de $F + G$.

13 Traiter les exercices 8, 9, 10, 11, 12, par les méthodes illustrées au paragraphe 4.

* **14** Montrer que les vecteurs suivants de $\mathbb{C}^3$

$$\begin{array}{rcl} v_1 & = & (1+i\,,\quad 1+2\,i\,,\quad i\,) \\ v_2 & = & (\;2\,,\quad 4-\,i\,,\quad -1\;) \\ v_3 & = & (\;0\,,\quad -1+2\,i\,,\quad 2+i\,) \end{array}$$

sont liés si $\mathbb{C}^3$ est considéré comme espace vectoriel sur $\mathbb{C}$ et sont libres si $\mathbb{C}^3$ est considéré comme espace vectoriel sur $\mathbb{R}$.

* **15** Démontrer le théorème 2.6 : les lignes non nulles d'une matrice échelonnée forment un système libre.

INDICATIONS

1 a) Incompatible.
b) Solution unique $(1, 2, -3)$.
c) Une infinité de solutions à 1 paramètre : $(2-\lambda,\ 2+2\,\lambda,\ \lambda)$.

2 En échelonnant, on trouve :

$$\left\{ \begin{array}{rl} x + y - \qquad\qquad z & = 1 \\ y + \quad (a+1)\,z & = 1 \\ (2+a)\,(3-a)\,z & = 3-a \end{array} \right.$$

Solution unique si $a \neq -2$ et $a \neq 3$; infinité de solutions si $a = 3$; système incompatible pour $a = -2$.

3 $c + 2\,b - 8\,a = 0.$

4 Le système a plus d'inconnues que d'équations.

5 $(x, y, z, t) = (5\,\lambda - \mu,\ \lambda + \mu,\ -2\,\lambda,\ 2\,\lambda,\ 2\,\mu) = \lambda\,(5, 1, -2, 2, 0) + \mu\,(-1, 1, 0, 0, 2).$

6 Si, par exemple, l'inconnue x_j n'apparaît pas dans le système, le système admet la solution $(0, \ldots, 0, \underbrace{1}_{\text{j}^{\text{ème}}\ \text{rang}}, 0, \ldots, 0)$.

7 Soit $\{v_1, \ldots, v_m\}$ avec $m > n$. En étudiant $x_1\,v_1 + \cdots + x_m\,v_m = 0$ on est amené à un système homogène en n équations et $m > n$ inconnues, ayant donc une solution non triviale.

8

1. On est amené à étudier le système :
$$\begin{cases} x_1 + x_2 + x_3 &= 0 \\ -x_1 + 3\,x_3 + 2\,x_4 &= 0 \\ x_2 + 5\,x_3 + 3\,x_4 &= 0 \\ 2\,x_1 + 2\,x_2 + 7\,x_3 + \alpha\,x_4 &= 0 \end{cases}$$
On trouve $\alpha \neq 5$.
2. Pour $\alpha = 5$ on trouve $v_1 - 2\,v_2 + v_3 - v_4 = 0$. Donc $\dim \operatorname{Vect}\{v_1, \ldots, v_4\} = 3$ et $\{v_1,\, v_2,\, v_3\}$ est une base.
3. Si $\alpha \neq 5$ $\{v_1, \ldots, v_4\}$ est une base de $\mathbb{R}^4$ et $v \in \operatorname{Vect}\{v_1, \ldots, v_4\}$ pour tout $k \in \mathbb{R}$. Pour les composantes, on résout le système trouvé en 1. avec second membre $(-2,\, k,\, 1,\, 3)$. On trouve :
$$x_1 = -9 + 4\,k + 5\,r, \quad x_2 = 16 + 5\,k + 10\,r, \quad x_3 = -3 - k - 5\,r,\ x_4 = 5\,r$$
avec $r = \dfrac{2 + k}{\alpha - 5}$.
Pour $\alpha = 5$: $v \in \operatorname{Vect}\{v_1,\, v_2,\, v_3,\, v_4\} \equiv \operatorname{Vect}\{v_1,\, v_2,\, v_3\}$ si et seulement si $k = -2$. On trouve :
$$v = -v_1 + 6\,v_2 - v_3.$$

9 Comme $\dim \mathbb{R}_2\,[\,t\,] = 3$, il suffit de vérifier que la famille est libre ou qu'elle est génératrice. Puisque on demande les composants de P il est préférable de vérifier qu'elle est génératrice (pour ne pas avoir à refaire les calculs), c'est-à-dire que pour tous $a,\, b,\, c \in \mathbb{R}$ il existe $x_1,\, x_2,\, x_3 \in \mathbb{R}$ tels que
$$a\,t^2 + b\,t + c = x_1 + x_2\,(t - 1) + x_3\,(t - 1)^2.$$
En identifiant les coefficicents de $1,\, t,\, t^2$ on est amené au système :
$$\begin{cases} x_1 - x_2 + x_3 = c \\ x_2 - 2\,x_3 = b \\ x_3 = a \end{cases}$$
qui a une solution pour tous $a,\, b,\, c$. On trouve $P = 3 - (t - 1) + 2\,(t - 1)^2$.

10 $6\,P_1 - 2\,P_2 - P_3 = 0.$

11 Puisque $\dim \mathcal{S}_3\,(\mathbb{R}) = 3$ (cf. exercice 18, chapitre 1), il suffit de vérifier que la famille est génératrice. En écrivant $x\,A + y\,B + z\,C = M$ on est amené au système :
$$\begin{cases} x + y - z = a \\ -2\,x + 3\,y + z = b \\ x + 6\,y - 3\,z = c \end{cases}$$
On trouve :
$$M = \frac{1}{5}\,(15\,a + 3\,b - 4\,c)\,A + \frac{1}{5}\,(5\,a + 2\,b - c)\,B + (3\,a + b - c)\,C$$

12 On a : $3\,v_1 - v_2 - v_3 = 0$, $5\,w_1 - 2\,w_2 - w_3 = 0$, $\dim F = 2$, $\dim G = 2$. Une famille génératrice de $F + G$ est, par exemple, $\{v_1,\, v_2,\, w_1,\, w_2\}$. On trouve la relation $v_1 - v_2 - w_1 + w_2 = 0$. Donc $\dim(F + G) = 3$ et par conséquent $\dim F \cap G = 1$. Une base de $F + G$ est, par exemple, $\{v_1,\, v_2,\, w_1\}$. Base de $F \cap G$: $(1,\, 4,\, -3,\, 4,\, 2)$.

Equations du sous-espace $F+G$: $a\,x_1 + b\,x_2 + c\,x_3 + d\,x_4 + c\,x_5 = 0$ avec a, b, c vérifiant le système :

$$\begin{cases} a + 3\,b - 2\,c + 2\,d + 3\,e = 0 \\ 2\,a + 7\,b - 5\,c + 6\,d + 5\,e = 0 \\ a + 3\,b \phantom{{}- 5\,c} + 2\,d + e = 0 \end{cases}$$

On trouve $a = -7\,e + 4\,d$, $b = 2\,e - 2\,d$ (avec e et d variables libres) ce qui donne les équations de $F + G$:

$$\begin{cases} 7\,x_1 + 2\,x_2 + x_3 + x_5 = 0 \\ 4\,x_1 - 2\,x_2 + x_4 = 0 \end{cases}$$

13 Calculs standard.

14 En échelonnant la matrice

$$\begin{array}{cl} \begin{bmatrix} 1+i & 1+2\,i & i \\ 2 & 4-i & -1 \\ 0 & -1+2\,i & 2+i \end{bmatrix} & \begin{matrix} v_1 \\ v_2 \\ v_3 \end{matrix} \end{array}$$

on trouve $v_3 = (1 - i\,)\,v_1 - v_2$.

Si $\mathbb{C}^3$ est considéré espace vectoriel sur $\mathbb{R}$, dans la base

$$(1,0,0),\ \ (i,0,0),\ \ (0,1,0),\ \ (0,i,0),\ \ (0,0,1),\ (0,0,i)$$

la matrice engendrée par v_1, v_2, v_3, est :

$$\begin{array}{cl} \begin{bmatrix} 1 & 1 & 1 & 2 & 0 & 1 \\ 2 & 0 & 4 & -1 & -1 & 0 \\ 0 & 0 & -1 & 2 & 2 & 1 \end{bmatrix} & \begin{matrix} v_1 \\ v_2 \\ v_3 \end{matrix} \end{array}$$

$$\longrightarrow \begin{array}{cl} \begin{bmatrix} 1 & 1 & 1 & 2 & 0 & 1 \\ 0 & -2 & 2 & -5 & -1 & -2 \\ 0 & 0 & -1 & 2 & 2 & 1 \end{bmatrix} & \begin{matrix} v_1 \\ v_2 - 2v_1 \\ v_3 \end{matrix} \end{array}$$

15 Par récurrence sur le nombre n de lignes de la matrice échelonnée. Pour $n = 1$ trivial. Soient $v_1, \ldots, v_n$ les vecteurs lignes d'une matrice échelonnée. La matrice engendrée par $v_2, \ldots, v_n$ est échelonnée ; donc $\dim \operatorname{Vect}\{v_2, \ldots, v_n\} = n - 1$, d'après l'hypothèse de récurrence. Montrer que $v_1 \notin \operatorname{Vect}\{v_2, \ldots, v_n\}$ et que $\dim \operatorname{Vect}\{v_1,\, v_2, \ldots, v_n\} = n$.

Chapitre 3

Applications linéaires et matrices

La structure d'espace vectoriel ne devient vraiment intéressante que si l'on introduit la notion d'application linéaire. Il s'agit des applications entre espaces vectoriels qui, dans un sens que nous allons préciser, «conservent la structure d'espace vectoriel».
Dans ce chapitre, qui est un peu l'axe de tout le reste du cours, nous allons donner essentiellement les définitions et les résultats élémentaires de base.

3.1 Applications linéaires

Définition 3.1 – *Soient E et E' deux espaces vectoriels sur le* même *corps K et f une application de E dans E'. On dit que f est* ***linéaire****, si :*

a) $f(v+w) = f(v) + f(w)\,, \quad \forall\, v,\, w \in E\,;$
b) $f(\lambda v) = \lambda f(v)\,, \qquad \forall\, v \in E,\ \forall\, \lambda \in K\,.$

L'ensemble des applications linéaires de E dans E' est noté $\mathcal{L}_K(E,\, E')$ ou, plus simplement, $\mathcal{L}(E,\, E')$.

Remarque - Si f est linéaire, on a : $f(0) = 0$.
Il suffit, en effet, de faire $\lambda = 0$ dans $f(\lambda x) = \lambda f(x)$.

Certains types d'applications linéaires sont particulièrement importants ; nous y reviendrons largement dans la suite. En voici les définitions :

Définition 3.2 – *On appelle* ***endomorphisme*** *de E, une application linéairede E dans E (même espace de départ et d'arrivée). L'ensemble des endomorphismes de E est noté $\mathrm{End}_K(E)$ ou, plus simplement $\mathrm{End}\,(E)$*[1]*.*
On appelle ***isomorphisme*** *de E sur E' une application linéaire* bijective *de E dans E'.*

Exemple 1. –

$$f: \begin{array}{ccc} E & \longrightarrow & E' \\ v & \longmapsto & 0 \end{array}$$

est une application linéaire dite *application nulle.*

[1]Une autre notation courante est $\mathcal{L}_K(E)$ ou encore $\mathcal{L}(E)$.

Exemple 2. –

$$\begin{array}{rccc} \mathrm{id}_E : & E & \longrightarrow & E \\ & v & \longmapsto & v \end{array}$$

est un endomorphisme de E dit *identité sur* E.

Exemple 3. –

$$\begin{array}{rccc} f : & \mathbb{R}^3 & \longrightarrow & \mathbb{R}^2 \\ & (x_1, x_2, x_3) & \longmapsto & (2x_1+x_2,\, x_2-x_3) \end{array}$$

est une application linéaire.

En effet, si $v = (x_1,\, x_2,\, x_3)$ et $w = (y_1,\, y_2,\, y_3)$, on a :

$$\begin{aligned} f\,(v+w) &= f\Big((x_1+y_1,\ x_2+y_2,\ x_3+y_3)\Big) \\ &= \Big(2\,(x_1+y_1)+(x_2+y_2),\ (x_2+y_2)-(x_3+y_3)\Big) \\ &= \Big((2x_1+x_2)+(2y_1+y_2),\ (x_2-x_3)+(y_2-y_3)\Big) = f\,(v)+f\,(w) \\ f\,(\lambda\, v) &= f\Big((\lambda\, x_1,\ \lambda\, x_2,\ \lambda\, x_3)\Big) = f\Big((2\,\lambda\, x_1+\lambda\, x_2,\ \lambda\, x_2-\lambda\, x_3)\Big) \\ &= f\Big(\lambda\,(2x_1+x_3,\ x_2-x_3)\Big) = \lambda\, f\,\Big((2x_1+x_3,\ x_2-x_3)\Big) = \lambda\, f\,(v) \end{aligned}$$

Comme on peut s'en rendre compte par cet exemple, la linéarité de f tient au fait que les composantes x_i dans l'espace d'arrivée (ici $\mathbb{R}^2$) *apparaissent toutes à la puissance 1* : plus précisément chaque composante dans l'espace d'arrivée est un ***polynôme homogène de degré 1*** en les x_i. Nous verrons cela d'une manière plus précise dans la suite.
Ainsi, par exemple, l'application

$$\begin{array}{rccc} f : & \mathbb{R}^3 & \longrightarrow & \mathbb{R}^2 \\ & (x_1, x_2, x_3) & \longmapsto & (x_1^2-x_2,\, x_2+x_3) \end{array}$$

n'est pas linéaire (ni a), ni b) de la définition 3.1 ne sont satisfaites à cause du terme au carré). De même :

$$\begin{array}{rccc} f : & \mathbb{R}^3 & \longrightarrow & \mathbb{R}^2 \\ & (x_1, x_2, x_3) & \longmapsto & (2\,x_1x_2\ ,\ x_2+3\,x_1) \end{array}$$

n'est pas linéaire $(2\,x_1x_2$ est un polynôme homogène de degré 2 en x_1 et en $x_2)$, pas plus que :

$$\begin{array}{rccc} f : & \mathbb{R}^3 & \longrightarrow & \mathbb{R}^2 \\ & (x_1, x_2, x_3) & \longmapsto & (2\,x_1+3\,x_2+x_3+5,\ x_1-2\,x_2+x_3) \end{array}$$

$(2\,x_1+3\,x_2+x_3+5$ n'est pas homogène, étant somme d'un polynôme de degré 1 et d'un polynôme de degré 0).

Exemple 4. –

$$\begin{array}{rccc} f : & \mathbb{R}^3 & \longrightarrow & \mathbb{R} \\ & (x_1, x_2, x_3) & \longmapsto & -x_1+2\,x_2+5\,x_3 \end{array}$$

est une application linéaire.

Exemple 5. – Soient $\mathcal{C}^0([0,1],\mathbb{R})$ et $\mathcal{C}^1([0,1],\mathbb{R})$ les espaces vectoriels des applications $f:[0,1]\longrightarrow\mathbb{R}$ respectivement continues et continues à dérivée continue. L'application :

$$\begin{array}{rccc} D: & \mathcal{C}^1([0,1],\mathbb{R}) & \longrightarrow & \mathcal{C}^0([0,1],\mathbb{R}) \\ & f & \longmapsto & f' \end{array}$$

est une application linéaire, puisque :

$$\begin{gathered} D(f+g)=(f+g)'=f'+g'=Df+Dg \\ D(\lambda f)=(\lambda f)'=\lambda f' \end{gathered}$$

si $\lambda\in\mathbb{R}$ et $f,\,g\in\mathcal{D}([0,1],\mathbb{R})$

Exemple 6. – Soit $E=E_1\oplus E_2$; alors tout vecteur $x\in E$ s'écrit d'une manière unique $x=x_1+x_2$ où $x_1\in E_1$ et $x_2\in E_2$.
L'application :

$$\begin{array}{rccc} \mathrm{pr}_1: & E & \longrightarrow & E \\ & x_1+x_2 & \longmapsto & x_1 \end{array}$$

est une application linéaire dite ***projecteur sur*** E_1 ***parallèlement à*** E_2.

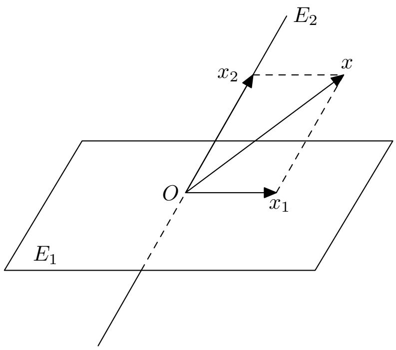


Figure 1

Exemple 7. – Soit $v_0\neq 0$ un vecteur de E ; l'application ***translation*** définie par

$$\begin{array}{rccc} t: & E & \longrightarrow & E \\ & v & \longmapsto & v+v_0 \end{array}$$

n'est pas linéaire (noter, par exemple, que : $t(0)=v_0\neq 0$)

Exercices 1. 2. 3. 4.

3.2 Image et noyau. Image d'une famille de vecteurs

Proposition 3.3 – *Soit $f:E\longrightarrow E'$ une application linéaire et F un sous-espace vectoriel de E. Alors $f(F)$ est un sous-espace vectoriel de E'.*
*En particulier $f(E)$ est un sous-espace de E' appelé **image** de f et noté $\mathrm{Im}\,f$. Sa dimension est appelée **rang** de f et est notée $\mathrm{rg}\,f$:*

$$\mathrm{rg}\,f:=\dim(\mathrm{Im}\,f)$$

En effet, soient $y_1, y_2 \in f(F)$; il existe alors $x_1, x_2 \in F$ tels que $y_1 = f(x_1)$, $y_2 = f(x_2)$. On a :

$$y_1 + y_2 = f(x_1) + f(x_2) = f(x_1 + x_2) \ ; \qquad \text{donc :} \quad y_1 + y_2 \in f(F)$$

De même si $y \in f(F)$ $\Big(y = f(x) \text{ avec } x \in F\Big)$, on a :

$$\lambda y = \lambda f(x) = f(\lambda x) \ ; \qquad \text{donc :} \quad \lambda y \in f(F) \qquad \square$$

Proposition 3.4 – *Soit $f \in \mathcal{L}(E, E')$ et*

$$\operatorname{Ker} f := \{x \in E \mid f(x) = 0\}$$

$\operatorname{Ker} f$ *est un sous-espace vectoriel de E appelé* ***noyau*** *de f.* [2]

En effet si $x, y \in \operatorname{Ker} f$, on a :

$$\begin{array}{lll} f(x+y) = f(x) + f(y) = 0 + 0 = 0 \ ; & \text{donc :} & x + y \in \operatorname{Ker} f \\ f(\lambda x) = \lambda f(x) = \lambda 0 = 0 \ ; & \text{donc :} & \lambda x \in \operatorname{Ker} f \qquad \square \end{array}$$

Proposition 3.5 – *Soit $f \in \mathcal{L}(E, E')$. Alors f est* ***injective*** *si et seulement si* $\operatorname{Ker} f = \{0\}$.

En effet, soit $\operatorname{Ker} f = \{0\}$ et $x, y \in E$ tels que $f(x) = f(y)$. On a : $f(x) - f(y) = 0$ d'où $f(x - y) = 0$. Ainsi $x - y \in \operatorname{Ker} f = \{0\}$ et donc $x = y$ c'est-à-dire f est injective.
Réciproquement, supposons f injective et soit $x \in \operatorname{Ker} f$, c'est-à-dire $f(x) = 0$. Puisque $f(0) = 0$ pour toute application linéaire, on a $f(x) = f(0)$. Comme f est injective, $x = 0$ et donc $\operatorname{Ker} f = \{0\}$. $\square$

Exemple 1 – Soit $E = E_1 \oplus E_2$ et pr_1 le projecteur sur E_1 parallèlement à E_2 (cf. exemple 6., page 59). On a : $\operatorname{Im} \mathrm{pr}_1 = E_1$ et $\operatorname{Ker} \mathrm{pr}_1 = E_2$, comme on le vérifie immédiatement.

Exemple 2 – Soit :

$$D : \underset{P}{\mathbb{R}[x]} \underset{\longmapsto}{\longrightarrow} \underset{P'}{\mathbb{R}[x]}$$

Le noyau de D est formé par les polynômes constants. D'autre part, $\operatorname{Im} D = \mathbb{R}[x]$, car si $P \in \mathbb{R}[x]$, $Q(x) := \displaystyle\int_0^x P(t)\, dt$ est un polynôme et on a $Q' = P$ c'est-à-dire $DQ = P$.

Exemple 3 – Soit :

$$f : \begin{array}{ccc} \mathbb{R}^3 & \longrightarrow & \mathbb{R}^3 \\ {\scriptstyle (x,y,z)} & \longmapsto & {\scriptstyle (x',y',z')} \end{array} \qquad \text{où :} \quad \left\{ \begin{array}{rcl} x' & = & x + y - z \\ y' & = & 2x + y - 3z \\ z' & = & 3x + 2y - 4z \end{array} \right.$$

$\operatorname{Ker} f$ est l'ensemble des triplets $(x, y, z) \in \mathbb{R}^3$ qui vérifient le système :

$$\left\{ \begin{array}{rcrcrcl} x & + & y & - & z & = & 0 \\ 2x & + & y & - & 3z & = & 0 \\ 3x & + & 2y & - & 4z & = & 0 \end{array} \right.$$

[2] Ker pour kernel = noyau, en anglais.

On trouve facilement $x = 2\lambda,\ y = -\lambda,\ z = \lambda$; c'est-à-dire $\operatorname{Ker} f$ est la droite vectorielle engendrée par le vecteur $(2, -1, 1)$.
Pour ce qui est de $\operatorname{Im} f$, on a :
$(x',\, y',\, z') \in \operatorname{Im} f$, si et seulement si, il existe $(x,\, y,\, z) \in \mathbb{R}^3$ vérifiant le système :

$$\left\{\begin{array}{rcrcrcl} x & + & y & - & z & = & x' \\ 2\,x & + & y & - & 3\,z & = & y' \\ 3\,x & + & 2\,y & - & 4\,z & = & z' \end{array}\right.$$

Il s'agit donc de savoir pour quelles valeurs de x', y', z' ce système est compatible. En échelonnant, on trouve :

$$\left\{\begin{array}{ll} x+y-z & = x' \\ \quad -y-z & = y' - 2\,x' \\ \quad -y-z & = z' - 3\,x' \end{array}\right. \quad \longrightarrow \quad \left\{\begin{array}{ll} x-y-z & = x' \\ \quad y+z & = 2x' - y' \\ \qquad 0 & = 2x' - y' + z' - 3x' \end{array}\right.$$

la condition de compatibilité est $2\,x' - y' + z' - 3\,x' = 0$ c'est-à-dire $x' + y' - z' = 0$.
L'image de f est donc le plan de $\mathbb{R}^3$ d'équation $x' + y' - z' = 0$.

Proposition 3.6 – *Soit $f \in \mathcal{L}(E, E')$ et $\{v_i\}_{i\in I}$ une famille de vecteurs de E.*

1. *Si f est* injective *et la famille de E $\{v_i\}_{i\in I}$ est libre, alors la famille $\{f(v_i)\}_{i\in I}$ de E' est libre.*
2. *Si f est* surjective *et la famille $\{v_i\}_{i\in I}$ est génératrice de E alors la famille $\{f(v_i)\}_{i\in I}$ est génératrice de E'.*

En particulier si f est bijective *l'image d'une base de E est une base de E'.*

Démonstration :

1. Supposons la famille $\{v_i\}_{i\in I}$ libre et soit f injective. Pour toute famille extraite $\{v_{\alpha_1}, ..., v_{\alpha_q}\}$, la relation
$$\lambda_1\, f(v_{\alpha_1}) + \cdots + \lambda_q\, f(v_{\alpha_q}) = 0$$
implique $f(\lambda_1\, v_{\alpha_1} + \cdots + \lambda_q\, v_{\alpha_q}) = 0$, c'est-à-dire $\lambda_1\, v_{\alpha_1} + \cdots + \lambda_q\, v_{\alpha_q} \in \operatorname{Ker} f$. Or $\operatorname{Ker} f = \{0\}$, donc
$$\lambda_1\, v_{\alpha_1} + \cdots + \lambda_q\, v_{\alpha_q} = 0$$
et puisque la famille $\{v_i\}_{i\in I}$ est libre, on a $\lambda_1 = 0, \ldots, \lambda_q = 0$. Donc la famille $\{f(v_i)\}_{i\in I}$ est libre.

2. Soit $y \in E'$ quelconque ; puisque f est surjective, il existe $x \in E$ tel que $y = f(x)$. D'autre part la famille $\{v_i\}_{i\in I}$ est génératrice, donc x est de la forme
$$x = \lambda_1\, v_{\alpha_1} + \cdots + \lambda_p\, v_{\alpha_p}$$
d'où : $f(x) = \lambda_1\, f(v_{\alpha_1}) + \cdots + \lambda_p\, f(v_{\alpha_p})$. y est donc combinaison linéaire d'éléments de la famille $\{f(v_i)\}_{i\in I}$ et, puisqu'il est choisi arbitrairement dans E', la famille $\{f(v_i)\}_{i\in I}$ est génératrice. $\square$

Théorème 3.7 – *Deux espaces vectoriels de dimension finie sont isomorphes, si et seulement si, ils ont même dimension.*

En effet, s'il existe un isomorphisme $f : E \longrightarrow E'$, l'image par f d'une base de E est une base de E', donc E et E' ont même dimension. Réciproquement, supposons que $\dim E = \dim E'$ et soient $\{e_1, \ldots, e_n\}$, $\{e'_1, \ldots, e'_n\}$ deux bases respectivement de E et E'. Considérons l'application $f : E \longrightarrow E'$ construite de la manière suivante :

- pour $k = 1, \ldots, n$ on pose : $f(e_k) = e'_k$;
- si $x = \sum_{k=1}^{n} x_k\, e_k$ on pose : $f(x) = \sum_{k=1}^{n} x_k\, f(e_k) = \sum_{k=1}^{n} x_k\, e'_k$

(en d'autres termes, on définit f sur la base de E et on la prolonge par linéarité sur E tout entier). On vérifie facilement que f est linéaire et bijective (la vérification est laissée en exercice). □

Remarque. – Comme on le voit de la démonstration, l'isomorphisme de E sur E' dépend du choix des bases dans E et dans E' et en général il n'y a pas d'isomorphisme canonique.

Dans le cas où les espaces E et E' sont de dimension finie, les dimensions du noyau et de l'image de f sont liées par la relation donnée dans le théorème suivant, l'un des plus importants en Algèbre Linéaire :

Théorème du rang . 3.8 – *Soient E et E' deux espaces vectoriels de dimension finie et $f : E \longrightarrow E'$ une application linéaire. On a alors :*

$$\boxed{\dim E = \operatorname{rg} f + \dim(\operatorname{Ker} f)}$$

Démonstration :
Supposons $\dim E = n$, $\dim \operatorname{Ker} f = r$ et montrons que $\dim(\operatorname{Im} f) = n - r$.
Soit $\{w_1, \ldots, w_r\}$ une base de $\operatorname{Ker} f$, et $\{v_1, \ldots, v_{n-r}\}$ une famille de vecteurs telle que $\{w_1, \ldots, w_r,\ v_1, \ldots, v_{n-r}\}$ soit une base de E. Soit $\mathcal{B} = \{f(v_1), \ldots, f(v_{n-r})\}$.
Montrons que $\mathcal{B}$ est une base de $\operatorname{Im} f$.

- $\mathcal{B}$ engendre $\operatorname{Im} f$.

Soit, en effet $y = f(x) \in \operatorname{Im} f$. Comme $x \in E$, x est de la forme $x = a_1 w_1 + \cdots + a_r w_r + b_1 v_1 + \cdots + b_{n-r} v_{n-r}$. On a donc :

$$\begin{aligned} y &= a_1 f(w_1) + \cdots + a_r f(w_r) + b_1 f(v_1) + \cdots + b_{n-r} f(v_{n-r}) \\ &= b_1 f(v_1) + \cdots + b_{n-r} f(v_{n-r}) \end{aligned}$$

ce qui montre que $\mathcal{B}$ engendre $\operatorname{Im} f$.

- $\mathcal{B}$ est libre.

Supposons que $\lambda_1 f(v_1) + \cdots + \lambda_{n-r} f(v_{n-r}) = 0$; on aura $f(\lambda_1 v_1 + \cdots + \lambda_{n-r} v_{n-r}) = 0$ donc :

$$\lambda_1 v_1 + \cdots + \lambda_{n-r} v_{n-r} \in \operatorname{Ker} f$$

Par conséquent, il existe $a_1, \ldots, a_r \in K$ tels que :

$$\lambda_1 v_1 + \cdots + \lambda_{n-r} v_{n-r} = a_1 w_1 + \cdots + a_r w_r$$

c'est-à-dire :

$$\lambda_1 v_1 + \cdots + \lambda_{n-r} v_{n-r} - a_1 w_1 - \cdots - a_r w_r = 0$$

Puisque la famille $\{v_1, \ldots, v_{n-r},\ w_1, \ldots, w_r\}$ est libre, les coefficients de cette combinaison linéaire sont tous nuls ; en particulier : $\lambda_1 = 0, \ldots, \lambda_{n-r} = 0$, c'est-à-dire $\mathcal{B}$ est libre. □

Ce théorème a un corollaire important. Pour montrer qu'une application linéaire est bijective, il faut montrer qu'elle est injective et surjective ; cependant, dans le cas de dimension finie, si la dimension de l'espace de départ et celle de l'espace d'arrivée sont les mêmes, il suffit de démontrer l'une des deux propriétés – soit l'injectivité, soit la surjectivité :

Corollaire 3.9 – *Soit $f \in \mathcal{L}(E, E')$, E, E' étant deux espaces vectoriels* ***de même dimension finie*** *(en particulier, par exemple, si $f \in \operatorname{End} E$, avec E de dimension finie). Alors les propriétés suivantes sont équivalentes :*

1. *f est injective.*
2. *f est surjective.*
3. *f est bijective.*

Démonstration : Il suffit, bien entendu de montrer que 1. est équivalent à 2..
Comme on l'a vu (cf. proposition 3.5), f est injective si et seulement si $\operatorname{Ker} f = \{0\}$. Puisque $\dim E = \operatorname{rg} f + \dim(\operatorname{Ker} f)$, f est injective si et seulement si $\dim E = \operatorname{rg} f$, c'est-à-dire $\dim E = \dim(\operatorname{Im} f)$. Or, par hypothèse, $\dim E = \dim E'$, donc f est injective si et seulement si $\dim(\operatorname{Im} f) = \dim E'$. Puisque $\operatorname{Im} f \subset E'$ cela équivaut à $\operatorname{Im} f = E'$ (cf. proposition 1.22, page 19), c'est-à-dire f surjective. □

Remarque. – Ce résultat est faux en dimension infinie. En voici un contre-exemple : l'application

$$D: \begin{array}{ccc} \mathbb{R}[x] & \longrightarrow & \mathbb{R}[x] \\ P & \longmapsto & P' \end{array}$$

est surjective et non injective.

Exercices 5. 6. 7. 8. 9. 10. 11. 12. 13. 14. 15. 16.

3.3 Matrices et applications linéaires

Si E' est un espace vectoriel sur K et A un ensemble quelconque $A \neq \emptyset$, l'ensemble des applications $f : A \longrightarrow E'$ est muni d'une structure d'espace vectoriel sur K par les lois :

$$f+g: \begin{array}{ccc} A & \longrightarrow & E' \\ x & \longmapsto & f(x)+g(x) \end{array} \qquad \lambda f: \begin{array}{ccc} A & \longrightarrow & E' \\ x & \longmapsto & \lambda f(x) \end{array}$$

(cf. exemple 4. page 5). En particulier $\mathcal{L}_K(E, E')$ est un sous-espace de l'espace vectoriel des applications $f : E \longrightarrow E'$:

Proposition 3.10 – *$\mathcal{L}_K(E, E')$ muni des lois :*

$$f+g: \begin{array}{ccc} E & \longrightarrow & E' \\ x & \longmapsto & f(x)+g(x) \end{array} \qquad \lambda f: \begin{array}{ccc} E & \longrightarrow & E' \\ x & \longmapsto & \lambda f(x) \end{array}$$

est un espace vectoriel sur K. De plus, si $f \in \mathcal{L}_K(E, E')$ et $g \in \mathcal{L}_K(E', E'')$ alors $g \circ f \in \mathcal{L}_K(E, E'')$ et on a :

$$\begin{array}{lll} g\circ(f+h) & = g\circ f+g\circ h & \\ (g+k)\circ f & = g\circ f+k\circ f & \\ g\circ(\lambda f) & = \lambda\, g\circ f & \forall\, f,h\in\mathcal{L}_K(E,E'),\ \forall\, g,k\in\mathcal{L}_K(E',E''),\ \forall\,\lambda\in K \end{array}$$

(avec la terminologie de l'Appendice A.1, $\mathrm{End}_K(E)$ *est donc un anneau pour les lois + et ∘).*
Enfin, si f est bijective, f^{-1} est linéaire.

La vérification est laissée en exercice (cf. exercice 2).

a) Matrices

Dans le cas où $\mathcal{L}_K\,(E,E')$ est de dimension finie, disons de dimension r, moyennant le choix d'une base on peut associer à toute application linéaire f un r-uplet d'éléments de K, les composantes de f. Pour des raisons que nous verrons par la suite, ces composantes sont rangées non pas sur une ligne mais sur un tableau ayant un certain nombre de lignes et de colonnes, que l'on appelle matrice associée à l'application linéaire f.

Définition 3.11 – *On appelle* ***matrice*** *de type $(p,\ n)$ à cœfficients dans K un tableau A de $p\,n$ éléments de K rangés sur p lignes et n colonnes :*

$$A=\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n}\\ a_{21} & a_{22} & \dots & a_{2n}\\ \dots & \dots & \dots & \dots\\ a_{p1} & a_{p2} & \dots & a_{pn}\end{pmatrix} \quad \text{ou, en abrégé : } A=(a_{ik})\ \text{, ou aussi : } A=||a_{ik}||.$$

L'ensemble des matrices à p lignes et n colonnes est noté $\mathcal{M}_{p,n}\,(K)$. Si $n=p$, $\mathcal{M}_{n,n}\,(K)$ est noté : $\mathcal{M}_n\,(K)$.

Ainsi, par exemple :

$$\begin{pmatrix} 1 & 3 & -1\\ 0 & 1 & 2\end{pmatrix}\in\mathcal{M}_{2,3}\,(\mathbb{R}) \qquad \begin{pmatrix} 1 & 2-i & 3+i\\ 0 & 1+i & i\\ -i & 2 & 1\end{pmatrix}\in\mathcal{M}_3\,(\mathbb{C})$$

Remarquons que dans la notation que nous venons d'adopter, a_{ik} désigne l'élément de la i-ème ligne et de la k-ème colonne [3] :

$$a_{\square\square}$$

indice de ligne   indice de colonne

Une autre notation que nous emploierons aussi dans la suite, est la «notation par colonnes» :

$$A=\|c_1,\dots,c_n\|\ , \qquad \text{où}\quad c_k=\begin{pmatrix} a_{1k}\\ a_{2k}\\ \vdots\\ a_{pk}\end{pmatrix} \quad \text{est la k}^{\text{ème}}\text{ colonne}$$

Sur l'ensemble $\mathcal{M}_{p,\,n}\,(K)$ on définit les lois :

[3] Cette notation est appelée parfois «notation LICO» pour : LIgne-COlonne

• *addition :* si $A = (a_{ik})$, $B = (b_{ik})$, on note $C = A + B$ la matrice (c_{ik}) telle que :

$$c_{ik} = a_{ik} + b_{ik}\,, \quad \forall\, i,\, k$$

• *produit par un scalaire :* si $A = (a_{ik})$ et $\lambda \in K$ on note $\lambda\, A$ la matrice $(\lambda\, a_{ik})$ c'est-à-dire la matrice obtenue en multipliant tous les éléments par λ.

Par exemple :

$$\begin{pmatrix} 2 & -1 & 0 & 3 \\ 1 & 2 & 1 & -1 \end{pmatrix} + \begin{pmatrix} 1 & 2 & 0 & 4 \\ 3 & 1 & -1 & 2 \end{pmatrix} = \begin{pmatrix} 3 & 1 & 0 & 7 \\ 4 & 3 & 0 & 1 \end{pmatrix}$$

$$5 \begin{pmatrix} 2 & -1 & 0 & 3 \\ 1 & 2 & 1 & -1 \end{pmatrix} = \begin{pmatrix} 10 & -5 & 0 & 15 \\ 5 & 10 & 5 & -5 \end{pmatrix}$$

Il est facile de voir que, muni de ces lois, $\mathcal{M}_{p,\,n}\,(K)$ est un espace vectoriel sur K. L'élément neutre est la matrice dont tous les éléments sont nuls, dite *matrice nulle*, notée 0. L'opposée de la matrice (a_{ik}) est la matrice $(-\,a_{ik})$.

Proposition 3.12 –

$$\boxed{\dim_{K} \mathcal{M}_{p,\,n}\,(K) = p\,n}$$

En effet, on vérifie facilement que les $p\,n$ matrices, dites ***matrices élémentaires***

$$E_{11} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & \cdots & \cdots & 0 \\ \vdots & & & \vdots \\ 0 & \cdots & \cdots & 0 \end{pmatrix}, \ldots, E_{ik} = \begin{array}{c} \begin{pmatrix} 0 & \cdots & 0 & \cdots & 0 \\ \vdots & & \vdots & & \vdots \\ 0 & \cdots & 1 & \cdots & 0 \\ \vdots & & \vdots & & \vdots \\ 0 & \cdots & 0 & \cdots & 0 \end{pmatrix} \leftarrow \text{i}^{\text{ème}} \text{ ligne}, \ldots, \\ \uparrow \\ k^{\text{ème}} \text{ colonne} \end{array}$$

$$E_{p\,n} = \begin{pmatrix} 0 & \cdots\cdots & 0 \\ 0 & \cdots\cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots\cdots & 1 \end{pmatrix}$$

forment une base de $\mathcal{M}_{p,n}\,(K)$ dite ***base canonique*** (cf. exercice 10, chapitre 1).

b) Matrices associées aux applications linéaires

Soient E et E' deux espaces vectoriels sur K, de dimension n et p respectivement, et $f : E \longrightarrow E'$ une application linéaire. Choisissons une base $\{e_1, \ldots, e_n\}$ de E et une base $\{\varepsilon_1, \ldots, \varepsilon_p\}$ de E'. Les images par f des vecteurs $e_1, \ldots, e_n$ se décomposent sur la base $\{\varepsilon_1, \ldots, \varepsilon_p\}$:

$$\begin{array}{ll} f\,(e_1) & = a_{11}\,\varepsilon_1 + a_{21}\,\varepsilon_2 + \cdots + a_{p1}\,\varepsilon_p \\ f\,(e_2) & = a_{12}\,\varepsilon_1 + a_{22}\,\varepsilon_2 + \cdots + a_{p2}\,\varepsilon_p \\ \ \ \cdots & \ \ \cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ f\,(e_n) & = a_{1n}\,\varepsilon_1 + a_{2n}\,\varepsilon_2 + \cdots + a_{p\,n}\,\varepsilon_p \end{array}$$

Définition 3.13 – *On appelle* ***matrice de f*** *dans les bases* $\{e_1, \ldots, e_n\}$, $\{\varepsilon_1, \ldots, \varepsilon_p\}$ *la matrice notée* $M(f)_{e_i, \varepsilon_j}$ *appartenant à* $\mathcal{M}_{p,n}(K)$ *dont les colonnes sont les composantes des vecteurs* $f(e_1), \ldots, f(e_n)$ *dans la base* $\{\varepsilon_1, \ldots, \varepsilon_p\}$ *:*

$$M(f)_{e_i, \varepsilon_j} = \begin{array}{c} \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & & & \vdots \\ a_{p1} & a_{p2} & \ldots & a_{pn} \end{pmatrix} \\ \begin{array}{cccc} \uparrow & \uparrow & & \uparrow \\ f(e_1) & f(e_2) & & f(e_n) \end{array} \end{array}$$

On utilisera aussi la notation : $\| f(e_1), \ldots, f(e_n) \|_{\varepsilon_j}$.

S'il n'y a pas d'ambiguïté possible, on écrira aussi $M(f)$ à la place de $M(f)_{e_i, \varepsilon_j}$, mais il est clair que la matrice associée à f dépend du choix des bases de E et E'.

Dans le cas où f est un endomorphisme, *on peut choisir la même base dans E considéré comme espace de départ et d'arrivée.* Dans ce cas, on notera $M(f)_{e_i}$ au lieu de $M(f)_{e_i, e_j}$.

Proposition 3.14 – *Soient* E *et* E' *deux espaces vectoriels sur* K *de dimension* n *et* p *respectivement,* $\{e_i\}$ *et* $\{\varepsilon_j\}$ *des bases de* E *et* E'. *Alors l'application :*

$$\begin{array}{rrcl} M: & \mathcal{L}_K(E, E') & \longrightarrow & \mathcal{M}_{p,n}(K) \\ & f & \longmapsto & M(f)_{e_i, \varepsilon_j} \end{array}$$

est un isomorphisme d'espaces vectoriels c'est-à-dire :

$$\left\{ \begin{array}{rcl} M(f+g) & = & M(f) + M(g) \\ M(\lambda f) & = & \lambda M(f) \end{array} \right. \qquad \textit{et } M \textit{ est bijective.}$$

En particulier : $\dim_K \mathcal{L}(E, E') = n\,p$

Démonstration : On a en effet :

$$\begin{aligned} M(f+g)_{e_i, \varepsilon_j} &= \|(f+g)(e_1), \ldots, (f+g)(e_n)\|_{\varepsilon_j} \\ &= \|f(e_1) + g(e_1), \ldots, f(e_n) + g(e_n)\|_{\varepsilon_j} \\ &= \|f(e_1), \ldots, f(e_n)\|_{\varepsilon_j} + \|g(e_1), \ldots, g(e_n)\|_{\varepsilon_j} \end{aligned}$$

d'après la définition de l'addition des matrices, c'est-à-dire :

$$M(f+g)_{e_i, \varepsilon_j} = M(f)_{e_i, \varepsilon_j} + M(g)_{e_i, \varepsilon_j}.$$

De même, si $\lambda \in K$:

$$\begin{aligned} M(\lambda f)_{e_i, \varepsilon_j} &= \|(\lambda f)(e_1), \ldots, (\lambda f)\, e_n\|_{\varepsilon_j} = \|\lambda f(e_1), \ldots, \lambda f(e_n)\|_{e_j} \\ &= \lambda \|f(e_1), \ldots, f(e_n)\|_{\varepsilon_j} \end{aligned}.$$

Donc M est linéaire.

D'autre part M est surjective. Soient, en effet :

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2n} \\ \vdots & & \vdots \\ a_{p1} & \cdots & a_{p\,n} \end{pmatrix} \in \mathcal{M}_{p\,n}(K)$$

et $f \in \mathcal{L}(E,\, F)$ définie de la manière suivante. On pose d'abord :

$$\begin{array}{ll} f(e_1) & = a_{11}\,\varepsilon_1 + a_{21}\,\varepsilon_2 + \cdots + a_{p1}\,\varepsilon_p \\ \cdots & \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ f(e_n) & = a_{1n}\,\varepsilon_1 + a_{2n}\,\varepsilon_2 + \cdots + a_{p\,n}\,\varepsilon_p \end{array}$$

On prolonge, ensuite, f par linéarité sur E, c'est-à-dire, si :

$$x = \lambda_1\, e_1 + \cdots + \lambda_n\, e_n \in E,$$

on pose :

$$f(x) = \lambda_1\, f(e_1) + \cdots + \lambda_n\, f(e_n).$$

Il est facile de vérifier que f est linéaire et que $A = M(f)_{e_i,\, \varepsilon_j}$.
Enfin M est injective. Soit en effet $f \in \operatorname{Ker} M$:

$$M(f) = \begin{array}{c} \begin{pmatrix} 0 & \cdots & 0 \\ 0 & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & 0 \end{pmatrix} \\ \begin{array}{cc} \uparrow & \uparrow \\ f(e_1) & f(e_n) \end{array} \end{array}$$

Ceci signifie que $f(e_1) = 0, \ldots, f(e_n) = 0$. Donc, si $x = \lambda_1\, e_1 + \cdots + \lambda_n\, e_n \in E$, on aura $f(x) = \lambda_1\, f(e_1) + \cdots + \lambda_n\, f(e_n) = 0$, c'est-à-dire $f = 0$. D'après la proposition 3.5, f est injective. □

Exemple 1 –

Soit E de dimension n et :
$$\begin{array}{rccc} \mathrm{id}_E : & E & \longrightarrow & E \\ & x & \longmapsto & x \end{array}$$

Considérons une base $\{e_i\}$. On a : $\mathrm{id}_E(e_i) = e_i$. Donc :

$$M(\mathrm{id}_E)_{e_i} = \begin{pmatrix} 1 & 0 & \ldots & 0 \\ 0 & 1 & 0 & \vdots \\ \vdots & 0 & \ddots & 0 \\ 0 & \ldots & 0 & 1 \end{pmatrix} \qquad (1 : \text{élément unité de } K)$$

Cette matrice est notée I_n ou simplement I et est appelée *matrice unité* de $\mathcal{M}_n(K)$.

Exemple 2 –

Soit $E = \mathbb{R}^2$ et :

$$pr_1 : \begin{array}{ccc} \mathbb{R}^2 & \longrightarrow & \mathbb{R}^2 \\ (x,y) & \longmapsto & (x,0) \end{array}$$

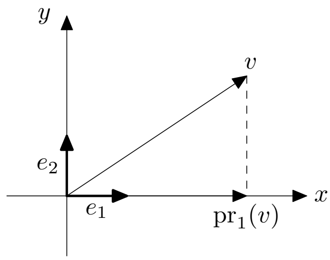


Figure 2

Considérons la base canonique de $\mathbb{R}^2$. On a :

$$\begin{array}{ll} \mathrm{pr}_1\,(e_1) & = e_1 \\ \mathrm{pr}_1\,(e_2) & = 0 \end{array}$$

Donc : $$M(\mathrm{pr}_1)_{e_i} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$$

Exemple 3 –

Soit $E = \mathbb{R}^2$ et f la projection sur la première bissectrice parallèlement à la seconde bissectrice. $\{e_1,\, e_2\}$ étant la base canonique de E, on a :

$$f\,(e_1) = f\,(e_2) = \varepsilon = \frac{1}{2}\,(e_1 + e_2)$$

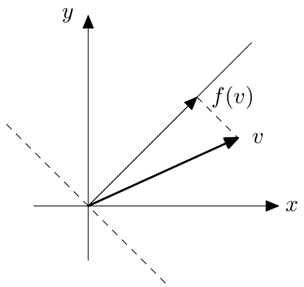


Figure 3

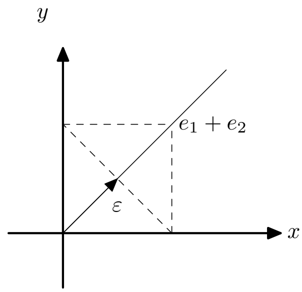


Figure 4

donc : $$M\,(f)_{e_i} = \tfrac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$$

Exemple 4 –

Soit $E = \mathbb{R}^2$ et f la symétrie par rapport à l'axe $\overrightarrow{Ox}$ parallèlement à l'axe $\overrightarrow{Oy}$. Soit $\{e_1,\, e_2\}$ la base canonique ; on a :

$$f\,(e_1) = e_1,\ f\,(e_2) = -e_2$$

Donc :

$$M\,(f)_{e_i} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

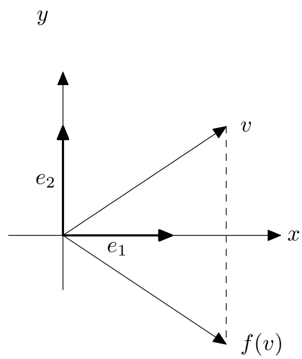


Figure 5

Exemple 5 –

Dans le plan $\mathbb{R}^2$ rapporté à la base canonique $\{e_1,\, e_2\}$, on considère la rotation de centre O et d'angle θ (nous verrons au chapitre 7 la définition précise de rotation ; pour le moment, on peut se contenter de la notion intuitive).
On voit facilement (cf. figure 6) que :

$$\begin{array}{lcl} f\,(e_1) & = & \cos\theta\, e_1 + \sin\theta\, e_2 \\ f\,(e_2) & = - & \sin\theta\, e_1 + \cos\theta\, e_2 \end{array}$$

et donc :

$$M\,(f)_{e_i} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

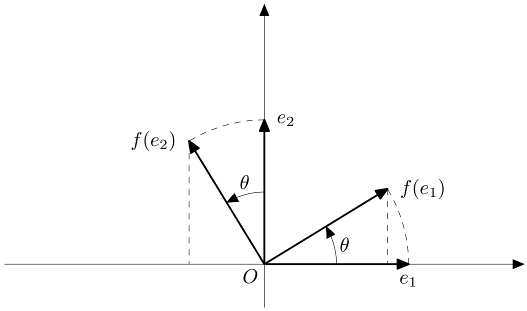


Figure 6

Exemple 6 –
Soit $\{\varepsilon_1,\, \varepsilon_2\}$ la base canonique de $\mathbb{R}^2$ et $\{e_1,\, e_2,\, e_3\}$ la base canonique de $\mathbb{R}^3$. On considère l'application linéaire :

$$f : \begin{array}{ccc} \mathbb{R}^3 & \longrightarrow & \mathbb{R}^2 \\ {\scriptstyle (x,y,z)} & \longmapsto & {\scriptstyle (x-y,\, z-y)} \end{array}$$

On a :

$$\begin{array}{lclcl} f\,(e_1) & = f\,(1,\,0,\,0) = & (1,\,0) & = & \varepsilon_1 \\ f\,(e_2) & = f\,(0,\,1,\,0) = & (-1,\,-1) & = - & \varepsilon_1 - \varepsilon_2 \\ f\,(e_3) & = f\,(0,\,0,\,1) = & (0,\,1) & = & \varepsilon_2 \end{array}$$

Donc :

$$M\,(f)_{e_i,\,\varepsilon_j} = \begin{pmatrix} 1 & -1 & 0 \\ 0 & -1 & 1 \end{pmatrix}$$

Exemple 7 –
On considère l'application :

$$\omega : \begin{array}{ccc} \mathbb{R}^n & \longrightarrow & \mathbb{R} \\ (x_1,\ldots,x_n) & \longmapsto & a_1\,x_1 + \cdots + a_n\,x_n \end{array}$$

En munissant $\mathbb{R}^n$ de la base canonique $\{e_1,\ldots,e_n\}$ et $\mathbb{R}$ de la base canonique $\{\varepsilon\}$ (c'est-à-dire $\varepsilon = 1$), on a :

$$\begin{array}{lll} \omega\,(e_1) = & f\,(1,0,\ldots,0) & = a_1 = a_1\,\varepsilon \\ \omega\,(e_2) = & f\,(0,1,0,\ldots,0) & = a_2 = a_2\,\varepsilon \\ \ldots & \ldots\ldots\ldots\ldots & \ldots\ldots\ldots \\ \omega\,(e_n) = & f\,(0,\ldots,0,1) & = a_n = a_n\,\varepsilon \end{array}$$

Ainsi :

$$M\,(\omega)_{e_i,\,\varepsilon} = (a_1,\ldots,a_n)$$

Exercices 17. 18. 19.

3.4 Produit de deux matrices

Au paragraphe précédent, nous avons défini sur les matrices les opérations d'addition et de produit par un scalaire. En vertu de la proposition 3.14, ces opérations correspondent aux opérations analogues sur les applications linéaires, c'est-à-dire on a :

$$\begin{aligned} M(f+g) &= M(f) + M(g) \\ M(\lambda f) &= \lambda M(f) \end{aligned}$$

Nous allons maintenant définir une nouvelle opération, le produit de matrices. Comme nous le verrons (cf. proposition 3.19), elle correspond à la composition des applications, en ce sens que :

$$M(f \circ g) = M(f) \cdot M(g)$$

Tout d'abord, remarquons que la composition des applications ne peut se faire pour tout couple d'applications, mais uniquement si l'espace d'arrivée de g est inclus dans l'espace de départ de f. Cette situation va se retrouver sur les matrices : le produit ne peut s'effectuer qu'entre matrices d'un certain type.

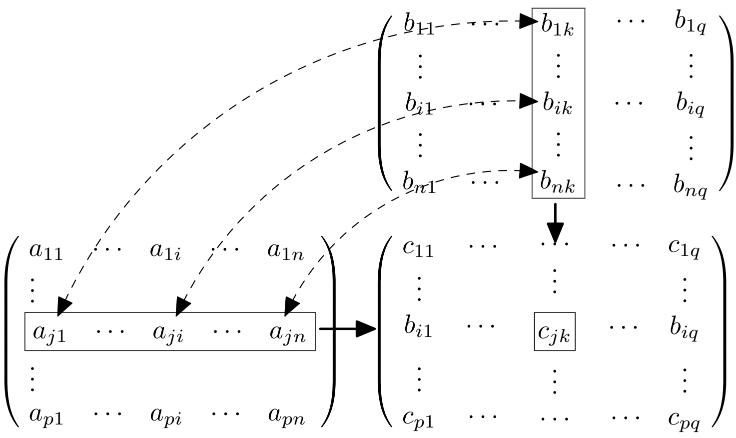
Définition 3.15 – *On appelle produit de matrices l'application :*

$$\begin{array}{ccccc} \mathcal{M}_{p,n}(K) & \times & \mathcal{M}_{n,q}(K) & \longrightarrow & \mathcal{M}_{p,q}(K) \\ {\scriptstyle (a_{ji})} & , & {\scriptstyle (b_{mk})} & \longmapsto & {\scriptstyle (c_{jk})} \end{array}$$

où :

$$c_{jk} = a_{j1}\, b_{1k} + a_{j2}\, b_{2k} + \cdots + a_{jn}\, b_{nk}.$$

En d'autres termes, l'élément c_{jk} de la j^{éme} ligne et k^{éme} colonne du produit $C = AB$ est la somme des produits des éléments de la j^{éme} ligne de A par les éléments de même rang de la k^{éme} colonne de B. Brièvement, on dit que le produit de deux matrices s'effectue «***lignes par colonnes***». Voici le schéma de cette définition.



Remarque. – Le produit AB ne peut s'effectuer que si le nombre des colonnes de A est égal au nombre des lignes de B.

Exemple 1 – Soit $A = \begin{pmatrix} 2 & -1 & 1 \\ 1 & 0 & 2 \end{pmatrix}$, $B = \begin{pmatrix} 2 & 0 & 1 & 0 \\ 1 & 1 & 2 & 2 \\ 1 & 2 & 1 & 0 \end{pmatrix}$

On a :

$$\begin{array}{ccc} & & \times \overbrace{\begin{pmatrix} 2 & 0 & 1 & 0 \\ 1 & 1 & 2 & 2 \\ 1 & 2 & 1 & 0 \end{pmatrix}}^{B} \\ \underbrace{\begin{pmatrix} 2 & -1 & 1 \\ 1 & 0 & 2 \end{pmatrix}}_{A} & = & \underbrace{\begin{pmatrix} 4 & 1 & 1 & -2 \\ 4 & 4 & 3 & 0 \end{pmatrix}}_{C=AB} \end{array}$$

Exemple 2 –

Soit $A = \begin{pmatrix} 2 & 3 \\ 4 & 1 \\ 0 & 2 \end{pmatrix}$ et $B = \begin{pmatrix} 1 \\ 3 \end{pmatrix}$; on a :

$$\begin{array}{ccc} & & \overbrace{\begin{pmatrix} 1 \\ 3 \end{pmatrix}}^{B} \\ \underbrace{\begin{pmatrix} 2 & 3 \\ 4 & 1 \\ 0 & 2 \end{pmatrix}}_{A} & = & \underbrace{\begin{pmatrix} 11 \\ 7 \\ 6 \end{pmatrix}}_{AB} \end{array}$$

Les remarques suivantes sont importantes :

Remarques –

1. *On peut avoir $AB = 0$ sans que A ou B soient nulles.*
2. *$AB = AC$ avec $A \neq 0$ n'implique pas nécessairement $B = C$ (c'est-à-dire, en général on ne peut pas "simplifier" par A, même si $A \neq 0$).*
3. *En général on a $AB \neq BA$ (c'est-à-dire : la multiplication entre matrices n'est pas commutative).*

Exemple :
Soient $A = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$, $B = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$ et $C = \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix}$.
En effectuant les produits, on trouve :

$$AB = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \qquad \text{(ce qui montre 1.)}$$

$$BA = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \qquad \text{donc } BA \neq AB \text{ (ce qui montre 3.)}$$

et $AB = AC$ (ce qui montre 2. puisque on a $B \neq C$).

Proposition 3.16 – *La multiplication est associative, c'est-à-dire :*

$$A\,(BC) = (AB)\,C\,, \qquad (\forall\, A \in \mathcal{M}_{p,n},\ \forall\, B \in \mathcal{M}_{n,q},\ \forall\, C \in \mathcal{M}_{q,m})$$

La multiplication est distributive à gauche et à droite par rapport à l'addition, c'est-à-dire :

$$\begin{aligned} A\,(B+C) &= AB+AC \\ (A+D)\,B &= AB+DB\ , \quad (\forall\, A,\ D \in \mathcal{M}_{p,n},\ \forall\, B,\ C \in \mathcal{M}_{n,q}) \end{aligned}$$

La vérification est laissée en exercice.
Remarquons enfin que la multiplication est une loi interne sur l'ensemble $\mathcal{M}_n\,(K)$ des matrices carrées d'ordre n, c'est-à-dire elle est une application :

$$\mathcal{M}_n\,(K) \times \mathcal{M}_n\,(K) \longrightarrow \mathcal{M}_n\,(K).$$

On vérifie immédiatement que la matrice I_n est l'élément neutre de la multiplication, c'est-à-dire :

$$\forall\, A \in \mathcal{M}_n\,(K) : \quad I_n\, A = A\, I_n = A.$$

Selon la terminologie de l'Appendice A.1, les lois de somme et de produit confèrent à $\mathcal{M}_n\,(K)$ une *structure d'anneau.*

Exercices 20. 21. 22. 23. 24.

3.5 Matrice d'un vecteur. Calcul de l'image d'un vecteur

Définition 3.17 – *Soit E un espace vectoriel de dimension n, $\{e_1, \ldots, e_n\}$ une base de E et $x = x_1\, e_1 + \cdots + x_n\, e_n$ un vecteur de E. On appelle matrice de x dans la base $\{e_i\}$ la matrice colonne des composantes de x dans la base $\{e_i\}$:*

$$M(x)_{e_i} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

$\big($*notée aussi* $M(x)\big)$.

Remarque. – Cette définition est cohérente avec la définition de matrice associée à une application linéaire. En effet, on peut identifier tout vecteur de E à une application linéaire de K dans E : à tout x de E on associe l'application linéaire

$$f: \begin{array}{ccc} K & \longrightarrow & E \\ \lambda & \longmapsto & \lambda\, x \end{array}$$

Si l'on écrit la matrice de f dans la base $\varepsilon = 1$ de K et $\{e_i\}$ de E, on a :

$$f\,(\varepsilon) = x = x_1\, e_1 + \cdots + x_n\, e_n$$

donc :

$$M\,(f)_{\varepsilon, e_i} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

Proposition 3.18 – *Soient E, F deux espaces vectoriels sur K, $\{e_1, .., e_n\}$ et $\{\varepsilon_1, ..., \varepsilon_p\}$ deux bases de E et F respectivement. Pour toute application $f \in \mathcal{L}_K(E, F)$ et pour tout $x \in E$, on a :*

$$M\big(f(x)\big)_{\varepsilon_j} = M(f)_{e_i,\,\varepsilon_j}\, M(x)_{e_i}$$

ou plus brièvement : $M\big(f(x)\big) = M(f)\, M(x)$.

Démonstration : Soit $M(f)_{e_i,\,\varepsilon_j} = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{p1} & \dots & a_{pn} \end{pmatrix}$, ce qui veut dire que : $f(e_j) = a_{1j}\,\varepsilon_1 + \cdots + a_{pj}\,\varepsilon_p = \displaystyle\sum_{k=1}^{p} a_{kj}\,\varepsilon_k$. On a :

$$\begin{aligned} f(x) &= f(x_1\, e_1 + \cdots + x_n\, e_n) = \textstyle\sum_{j=1}^{n} x_j\, f(e_j) = \sum_{j=1}^{n} x_j \sum_{k=1}^{p} a_{kj}\,\varepsilon_k \\ &= \textstyle\sum_{k=1}^{p} \underbrace{\left(\displaystyle\sum_{j=1}^{n} a_{kj}\, x_j\right)}_{y_k} \varepsilon_k = \sum_{k=1}^{p} y_k\,\varepsilon_k \end{aligned}$$

Donc : $M\big(f(x)\big)_{\varepsilon_j} = \begin{pmatrix} y_1 \\ \vdots \\ y_p \end{pmatrix}$, avec $y_k = \displaystyle\sum_{j=1}^{n} a_{kj}\, x_j$. D'autre part :

$$M(f)_{e_i,\varepsilon_j} \cdot M(x)_{e_i} = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{p1} & \dots & a_{pn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \sum_{j=1}^{n} a_{1j}\ x_j \\ \vdots \\ \sum_{j=1}^{n} a_{pj}\ x_j \end{pmatrix} = \begin{pmatrix} y_1 \\ \vdots \\ y_p \end{pmatrix}$$

donc : $M(f)_{e_i,\,\varepsilon_j} \cdot M(x)_{e_i} = M\big(f(x)\big)_{\varepsilon_j}$

Exemple :
Soit le plan $\mathbb{R}^2$ rapporté à sa base canonique. Déterminer l'image du vecteur $x = (3,\, 2)$ par la rotation de centre O et d'angle $\pi/6$.

On a :

$$\begin{aligned} M\Big(f(x)\Big) &= M(f) \cdot M(x) \\ &= \begin{pmatrix} \cos\pi/6 & -\sin\pi/6 \\ \sin\pi/6 & \cos\pi/6 \end{pmatrix} \begin{pmatrix} 3 \\ 2 \end{pmatrix} \end{aligned}$$

$$= \begin{pmatrix} \frac{\sqrt{3}}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{\sqrt{3}}{2} \end{pmatrix} \begin{pmatrix} 3 \\ 2 \end{pmatrix} = \begin{pmatrix} \frac{3\sqrt{3}-2}{2} \\ \frac{3+2\sqrt{3}}{2} \end{pmatrix}$$

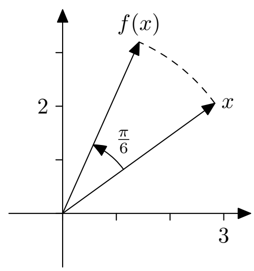


Figure 7

Exercice 25.

3.6 Produits de matrices. Matrice de l'inverse d'une application

Proposition 3.19 – *Soient E, F, G trois espaces vectoriels de dimension finie sur K, $\{e_1, \dots, e_n\}$, $\{\varepsilon_1, \dots, \varepsilon_p\}$, $\{\eta_1, \dots, \eta_q\}$ des bases de E, F et G respectivement. Si $g \in \mathcal{L}(E, F)$ et $f \in \mathcal{L}(F, G)$ (c'est-à-dire : $E \xrightarrow{g} F \xrightarrow{f} G$), on a :*

$$M(f \circ g)_{e_i,\, \eta_k} = M(f)_{\varepsilon_j,\, \eta_k}\, M(g)_{e_i,\, \varepsilon_j}$$

ou, plus brièvement :

$$M(f \circ g) = M(f)\, M(g).$$

Démonstration : Soit $x \in E$ arbitraire. En utilisant le résultat de la proposition 3.18, on a :

$$\begin{aligned} M(f \circ g)\, M(x) &= M\big((f \circ g)(x)\big) = M\Big(f\big(g(x)\big)\Big) \\ &= M(f)\, M\big(g(x)\big) = M(f)\, M(g)\, M(x) \end{aligned}$$

d'où, puisque x est arbitraire :

$$M(f \circ g) = M(f)\, M(g). \qquad \square$$

Exemple :

Déterminer dans la base canonique de $\mathbb{R}^2$ la matrice de l'application h qui est la composée de la rotation g autour de O d'angle θ, suivie de la projection f sur la première bissectrice parallèlement à la seconde bissectrice.

On a :

$$\begin{aligned} M(h) &= M(f \circ g) = M(f)\, M(g) \\ &= \tfrac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \\ &= \tfrac{1}{2}\begin{pmatrix} \cos\theta + \sin\theta & -\sin\theta + \cos\theta \\ \cos\theta + \sin\theta & -\sin\theta + \cos\theta \end{pmatrix} \end{aligned}$$

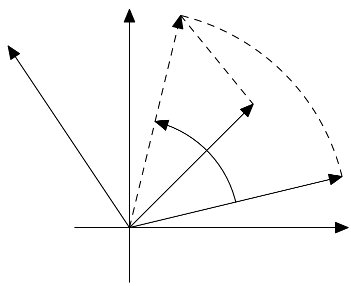


Figure 8

Définition 3.20 – *Une matrice carrée $A \in \mathcal{M}_n(K)$ est dite **inversible** s'il existe une matrice $A' \in \mathcal{M}_n(K)$ telle que :*

$$AA' = A'A = I.$$

*A' est dite **inverse** de A et est notée A^{-1}.*

Par exemple, la matrice $A = \begin{pmatrix} 1 & 2 \\ 1 & 3 \end{pmatrix}$ est inversible : son inverse est $A^{-1} = \begin{pmatrix} 3 & -2 \\ -1 & 1 \end{pmatrix}$ comme on le vérifie immédiatement en effectuant les produits AA^{-1} et $A^{-1}A$.

Il existe des matrices non inversibles, par exemple la matrice nulle. Mais la matrice nulle n'est pas la seule matrice non inversible. Considérons par exemple la matrice $A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$. S'il existait une matrice $A' = \begin{pmatrix} x & y \\ z & t \end{pmatrix}$ telle que $AA' = I$, on aurait :

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x & y \\ z & t \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

c'est-à-dire :
$$\begin{pmatrix} x & y \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

ce qui, évidemment, est impossible.

En fait, les matrices inversibles sont les matrices qui représentent les applications linéaires bijectives. On a en effet :

Proposition 3.21 – *Soient E et F deux espaces vectoriels de même dimension n sur K, $\{e_i\}$ une base de E, $\{\varepsilon_j\}$ une base de F. Une application linéaire $f : E \longrightarrow F$ est bijective (c'est-à-dire est un isomorphisme) si et seulement si $M(f)_{e_i, \varepsilon_j}$ est inversible. De plus :*

$$M(f)^{-1}_{e_i, \varepsilon_j} = M(f^{-1})_{\varepsilon_j, e_i}$$

ou, d'une manière plus concise : $M(f^{-1}) = M(f)^{-1}$.

Démonstration : On a $f^{-1} \circ f = \mathrm{id}_E$; d'où $M(f^{-1} \circ f)_{e_i, e_i} = M(\mathrm{id}_E)_{e_i, e_i}$
Donc, d'après la proposition 3.19 :

$$M(f^{-1})_{\varepsilon_j, e_i} M(f)_{e_i, \varepsilon_j} = I$$

De même, on voit que $M(f)_{e_i, \varepsilon_j} M(f^{-1})_{\varepsilon_j, e_i} = I$. $\square$

Calcul de l'inverse d'une matrice

Il existe différentes méthodes pour calculer l'inverse d'une matrice, sur lesquelles nous reviendrons. Pour le moment, on peut retenir la suivante qui est d'ailleurs d'un usage courant.
Soit $A \in \mathcal{M}_n(K)$, x et $x' \in K^n$ et X, X' les matrices colonnes qui représentent x et x' dans la base canonique de K^n. Considérons l'équation matricielle :

$$X' = A\,X \qquad (*)$$

Si A est inversible, en multipliant les deux membres à gauche par A^{-1} on obtient $A^{-1} X' = (A^{-1} A)\,X$, c'est-à-dire :

$$X = A^{-1} X'$$

Donc A^{-1} est la matrice du système obtenu en résolvant le système (*) en les composantes x_i de x.

Exemple :
Calculer l'inverse de la matrice

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

Ecrivons l'équation matricielle (*) avec $X = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ et $X' = \begin{pmatrix} x'_1 \\ x'_2 \end{pmatrix}$.

On a :
$\begin{pmatrix} x'_1 \\ x'_2 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ ce qui est équivalent au système $\begin{cases} x'_1 &= x_1 + 2\,x_2 \\ x'_2 &= x_1 + 3\,x_2 \end{cases}$

En résolvant en x_1 et x_2, on trouve :

$$\begin{cases} x_1 &= 3\,x'_1 - 2\,x'_2 \\ x_2 &= -\ x'_1 + \ x'_2 \end{cases}$$

c'est-à-dire $\quad X = \begin{pmatrix} 3 & -2 \\ -1 & 1 \end{pmatrix} X'$, $\quad$ donc $\quad A^{-1} = \begin{pmatrix} 3 & -2 \\ -1 & 1 \end{pmatrix}$.

Définition 3.22 – *L'ensemble des matrices inversibles de $\mathcal{M}_n\,(K)$ est noté* GL(n, K) *et est dit* ***groupe linéaire*** [4]*.*

Exercices 26. 27.

3.7 Changement de base

La matrice qui représente une application linéaire a été construite à l'aide d'un choix des bases dans l'espace de départ et dans l'espace d'arrivée. Dans ce paragraphe, nous allons voir comment sont reliées deux matrices qui représentent la *même* application linéaire en des bases *différentes*.

a) Matrice de passage

Soit E un espace vectoriel de dimension n, $\{e_1, \dots, e_n\}$ et $\{e'_1, \dots, e'_n\}$ deux bases de E. Les vecteurs e'_i s'écrivent comme combinaisons linéaires des vecteurs e_i :

$$\begin{array}{ll} e'_1 & = p_{11}\,e_1 + p_{21}\,e_2 + \cdots + p_{n1}\,e_n \\ e'_2 & = p_{12}\,e_1 + p_{22}\,e_2 + \cdots + p_{n2}\,e_n \\ \cdots & \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ e'_n & = p_{1n}\,e_1 + p_{2n}\,e_2 + \cdots + p_{nn}\,e_n \end{array}$$

Définition 3.23 – *On appelle* ***matrice de passage*** *de la base $\{e_i\}$ à la base $\{e'_i\}$ la matrice notée $P_{e_i \longrightarrow e'_i}$ dont les colonnes sont les composantes des vecteurs e'_i dans la base $\{e_i\}$:*

$$P_{e_i \longrightarrow e'_i} = \|e'_1, \dots, e'_n\|_{e_i} = \begin{pmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{21} & p_{22} & \cdots & p_{2n} \\ \vdots & \vdots & & \vdots \\ p_{n1} & p_{n2} & \cdots & p_{nn} \end{pmatrix}$$

[4] Le produit de matrices confère, en effet, à GL(n, K) une structure de groupe (selon la terminologie de l'Appendice A.1.), comme on le vérifie facilement.

On a, bien entendu :

$$\boxed{P_{e_i \longrightarrow e'_i} \;=\; M\,(\mathrm{id}_E)_{e'_i,\,e_i}}$$

$$\boxed{P_{e_i \longrightarrow e'_i} \;=\; M\,(\mathrm{id}_E)_{e'_i,\,e_i}}$$

On en déduit immédiatement (cf. propositions 3.19 et 3.21) :

Proposition 3.24 *1. Propriété transitivité :*

$$\boxed{P_{e_i \to e_{i'}}\, P_{e_{i'} \to e_{i"}} \;=\; P_{e_i \to e_{i"}}}$$

2. Les matrices de passage sont inversibles et on a :

$$\boxed{(P_{e_i \longrightarrow e'_i})^{-1} \;=\; P_{e'_i \longrightarrow e_i}}$$

b) Action du changement de base sur les composantes d'un vecteur

Soit $x \in E$, de composantes $(x_1, \ldots, x_n)$ dans la base $\{e_i\}$ et de composantes $(x'_1, \ldots, x'_n)$ dans la base $\{e'_i\}$. Il est facile de déterminer les relations entre les x_i et x'_i à l'aide de la matrice de passage $P_{e_i \longrightarrow e'_i}$.
Notons

$$X := \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = M\,(x)_{e_i}\;, \quad X' := \begin{pmatrix} x'_i \\ \vdots \\ x'_n \end{pmatrix} = M\,(x)_{e'_i} \qquad \text{et} \quad P = P_{e_i \longrightarrow e'_i}$$

On a (cf. proposition 3.18) :

$$P\,X' = M\,(\mathrm{id}_E)_{e'_i,\,e_i} \;\times\; M\,(x)_{e'_i} = M\left(\mathrm{id}_E\,(x)\right)_{e_i} = M\,(x)_{e_i} = X$$

c'est-à-dire $P\,X' = X$, d'où $X' = P^{-1}\,X$.

Nous avons donc démontré la relation importante :

Proposition 3.25 – *Soient $x \in E$, $\{e_i\}$ et $\{e'_i\}$ deux bases de E, $P = P_{e_i \longrightarrow e'_i}$ et $X = M\,(x)_{e_i}$, $X' = M\,(x)_{e'_i}$. Alors :*

$$\boxed{X' = P^{-1}\,X}$$

NOTA. – On dit que les composantes d'un vecteur x *se transforment d'une manière "contravariante"* pour exprimer le fait que si les bases sont transformées par la matrice P, les composantes de x sont transformées par la matrice P^{-1}.

Exemple :
Soit $\mathbb{R}^2$ muni de deux bases : la base canonique $\{e_1, e_2\}$ et la base $\{e'_1, e'_2\}$ définie par :

$$\begin{cases} e'_1 &= 2\,e_1 + \ e_2 \\ e'_2 &= 3\,e_1 + 2e_2 \end{cases} \qquad (*)$$

Soit $x = 2\,e_1 + 3\,e_2$. On a deux méthodes pour calculer les composantes de x dans la base $\{e'_1, e'_2\}$.

1ère méthode (méthode matricielle)

On a : $P = \begin{pmatrix} 2 & 3 \\ 1 & 2 \end{pmatrix}$, $P^{-1} = \begin{pmatrix} 2 & -3 \\ -1 & 2 \end{pmatrix}$ et $X = \begin{pmatrix} 2 \\ 3 \end{pmatrix}$

En appliquant la relation de la proposition 3.25, on trouve :

$$X' = P^{-1}\,X = \begin{pmatrix} 2 & -3 \\ -1 & 2 \end{pmatrix}\begin{pmatrix} 2 \\ 3 \end{pmatrix} = \begin{pmatrix} -5 \\ 4 \end{pmatrix}$$

donc $x = -5\,e'_1 + 4\,e'_2$.

2ème méthode (calcul direct)

On exprime e_1 et e_2 en fonction de e'_1 et e'_2 en résolvant le système (*).

On a : $$\begin{cases} e_1 &= \ \ 2\,e'_1 - \ e'_2 \\ e_2 &= -3\,e'_1 + 2e'_2 \end{cases}$$

En remplaçant dans l'expression de x, on trouve :

$$x = 2\,e_1 + 3\,e_2 = 2\,(2\,e'_1 - e'_2) + 3\,(-3\,e'_1 + 2\,e'_2) = -5\,e'_1 + 4\,e'_2.$$

c) Action du changement de base sur la représentation matricielle d'une application linéaire.

Proposition 3.26 – *Soient* $f \in \mathcal{L}(E, F)$, $\{e_1, \ldots, e_n\}$ *et* $\{e'_1, \ldots, e'_n\}$ *deux bases de* E *et* $\{\varepsilon_1, \ldots, \varepsilon_p\}$, $\{\varepsilon'_1, \ldots, \varepsilon'_p\}$ *deux bases de* F*. Notons :*

$$A = M\,(f)_{e_i,\,\varepsilon_j} \qquad A' = M\,(f)_{e'_i,\,\varepsilon'_j}\ , \quad P = P_{e_i \longrightarrow e'_i}\ , \quad Q = P_{\varepsilon_j \longrightarrow \varepsilon'_j}$$

On a alors :

$$\boxed{A' = Q^{-1}\,A\,P}$$

Démonstration : Soit $x \in E$ un vecteur arbitraire. D'après la proposition 3.25, on a :

$$M\Big(f\,(x)\Big)_{\varepsilon'_j} = Q^{-1}\,M\Big(f\,(x)\Big)_{\varepsilon_j} = Q^{-1}\,M\,(f)_{e_i,\,\varepsilon_j}\,M\,(x)_{e_i} = Q^{-1}\,A\,X$$

où on a posé $X = M\,(x)_{e_i}$. D'autre part, si $X' = M\,(x)_{e'_i}$:

$$M\Big(f\,(x)\Big)_{\varepsilon'_j} = M\,(f)_{e'_i,\,\varepsilon'_j}\,M\,(x)_{e'_i} = A'\,X' = A'\,P^{-1}\,X$$

Donc :

$$A'\,P^{-1}\,X = Q^{-1}\,A\,X$$

Comme x est arbitraire, cela implique que $A'\,P^{-1} = Q^{-1}\,A$, d'où : $A' = Q^{-1}\,A\,P$. □

Comme nous le verrons au chapitre 6, le cas des endomorphismes est particulièrement important, en particulier lorsqu'on prend la même base dans l'espace de départ et d'arrivée [5]. On a immédiatement :

Corollaire 3.27 – *Soit* $f = E \longrightarrow E$ *un endomorphisme de* E *et* $\{e_1, \ldots, e_n\}$, $\{e'_1, \ldots, e'_n\}$ *deux bases de* E. *Notons :*

$$A = M(f)_{e_i}, \qquad A' = M(f)_{e'_i} \qquad \text{et} \qquad P = P_{e_i \longrightarrow e'_i}.$$

On a alors :

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

Définition 3.28 – *Deux matrices* $A, A' \in \mathcal{M}_n(K)$ *sont dites* ***semblables*** *s'il existe une matrice inversible* $P \in \mathcal{M}_n(K)$ *telle que :*

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

Il est clair que ***deux matrices semblables représentent le même endomorphisme en des bases différentes.***

Exemple :

Soit f *l'endomorphisme de* $\mathbb{R}^3$ *qui dans la base canonique* $\{e_i\}$ *est représenté par la matrice :*

$$A = M(f)_{e_i} = \begin{pmatrix} 3 & -1 & 1 \\ 0 & 2 & 0 \\ 1 & -1 & 3 \end{pmatrix}$$

Déterminer la matrice A' *qui représente* f *dans la base* $\{e'_i\}$ *où :*

$$\left\{ \begin{array}{lcl} e'_1 & = & (1,\ 0,\ -1) \\ e'_2 & = & (0,\ 1,\ \ \ 1) \\ e'_3 & = & (1,\ 0,\ \ \ 1) \end{array} \right. .$$

On a $A' = P^{-1}\, A\, P$ avec $P = \|e'_1,\, e'_2,\, e'_3\|_{e_i} = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ -1 & 1 & 1 \end{pmatrix}$.

D'après la proposition 3.24 : $P^{-1} = P_{e'_i \longrightarrow e_i} = \|e_1,\, e_2,\, e_3\|_{e'_i}$. Il s'agit donc d'exprimer e_1, e_2, e_3 dans la base $\{e'_1,\, e'_2,\, e'_3\}$. Or :

$$\left\{ \begin{array}{lcl} e'_1 & = & e_1 - e_3 \\ e'_2 & = & e_2 + e_3 \\ e'_3 & = & e_1 + e_3 \end{array} \right.$$

En résolvant en e_1, e_2, e_3 :

$$\left\{ \begin{array}{lcl} e_1 & = & \frac{1}{2}(e'_1 + e'_3) \\ e_2 & = & \frac{1}{2}(e'_1 + 2e'_2 - e'_3) \\ e_3 & = & \frac{1}{2}(-e'_1 + e'_3) \end{array} \right.$$

[5] Dans ce cas, comme nous l'avons déjà signalé, on note $M(f)_{e_i}$ au lieu de $M(f)_{e_i,\, e_i}$.

donc

$$P^{-1} = \|e_1,\, e_2,\, e_3\|_{e'_i} = \frac{1}{2}\begin{pmatrix} 1 & 1 & -1 \\ 0 & 2 & 0 \\ 1 & -1 & 1 \end{pmatrix}.$$

En effectuant le produit $A' = P^{-1}\, A\, P$, on trouve :

$$A' = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 4 \end{pmatrix}$$

REMARQUE. – Puisque $A' = M\,(f)_{e'_i}$, ceci veut dire que :

$$f\,(e'_1) = 2e'_1 \quad , \quad f\,(e'_2) = 2e'_2 \quad , \quad f\,(e'_3) = 4e'_3$$

comme d'ailleurs on le vérifie directement. On a, en effet :

$$f\,(e'_1) = f\,(e_1 - e_3) = f\,(e_1) - f\,(e_3)$$

Or $f\,(e_1) = 3\,e_1 + e_3$ (cf. la première colonne de la matrice A) ; de même $f\,(e_3) = e_1 + 3\,e_3$. Donc : $f\,(e'_1) = 2\,e_1 - 2\,e_3 = 2\,e'_1$, etc.

Exercices 28. 29. 30.

3.8 Rang d'une application linéaire et rang d'une matrice

Comme nous l'avons vu (cf. définition 3.3), on appelle rang d'une application linéaire f la dimension de $\operatorname{Im} f$. Puisque $\mathcal{L}_K(E,\, F)$ est isomorphe à $\mathcal{M}_{p,n}(K)$, il faut s'attendre à ce que l'on puisse calculer le rang de f à l'aide de la matrice associée à f. C'est ce que nous allons voir dans ce paragraphe.

Définition 3.29 –

1. *Soit $\mathcal{F} = \{v_i\}_{i \in I}$ une famille de vecteurs. On appelle* ***rang de la famille*** *la dimension de l'espace engendré par les vecteurs $\{v_i\}_{i \in I}$.*
2. *Soit $A \in \mathcal{M}_{p,n}\,(K)$, $A = \|c_1, \ldots, c_n\|$ où l'on a noté c_i les vecteurs colonnes de A (notons que $c_i \in K^p$). On appelle* ***rang de la matrice*** *A le rang de la famille des vecteurs colonnes de A :*

$$\operatorname{rg} \|c_1, \ldots, c_n\| = \operatorname{rg}\,\{c_1, \ldots, c_n\} = \dim \operatorname{Vect}\{c_1, \ldots, c_n\}$$

Proposition 3.30 – *Soient E et F deux espaces vectoriels de dimension finie et $f \in \mathcal{L}\,(E,\, F)$. Soient $\{e_1, \ldots, e_n\}$, $\{\varepsilon_1, \ldots, \varepsilon_p\}$ deux bases quelconques de E et F respectivement, et $A = M\,(f)_{e_i,\,\varepsilon_j}$. On a alors :*

$$\operatorname{rg} f = \operatorname{rg} A$$

En particulier : deux matrices qui représentent la même application linéaire en des bases différentes ont même rang. En particulier, deux matrices semblables ont même rang.

En effet :

$$A = M\,(f)_{e_i,\,\varepsilon_j} = ||\, f(e_1), ..., f(e_n)\,||_{\varepsilon_j}$$

donc

$$\operatorname{rg} A = \dim\,(\,\operatorname{Vect}\{f(e_1), ..., f(e_n)\}\,) = \dim\,(\operatorname{Im} f) = \operatorname{rg} f \quad \Box$$

Soit $A \in \mathcal{M}_{p,n}(K)$ et ${}^tA$ la matrice dont les lignes sont les colonnes de A (cf. exercice 18, chapitre 1.). Par exemple, si :

$$A = \begin{pmatrix} 2 & -1 \\ 0 & 3 \\ 1 & 8 \end{pmatrix}, \quad {}^tA = \begin{pmatrix} 2 & 0 & 1 \\ -1 & 3 & 8 \end{pmatrix}.$$

${}^tA$ est dite ***transposée*** de A. Nous démontrerons par la suite (conséquence immédiate du théorème 4.12 page 114 et du théorème (E) page 127) la proposition suivante (cf. aussi exercice 37) :

Proposition 3.31 – *Pour toute matrice A, on a :*

$$\boxed{\operatorname{rg} A = \operatorname{rg} {}^tA}$$

Il s'ensuit que le rang d'une matrice est aussi égal au rang de la famille des vecteurs lignes.

Exemple :

Calculer le rang de la matrice :

$$A = \begin{pmatrix} 1 & -1 & 3 & 5 & 1 \\ 2 & 0 & -1 & 3 & 1 \\ 3 & -1 & 2 & 8 & 2 \end{pmatrix}$$

D'après la définition, il faudrait calculer le rang des vecteurs colonnes en échelonnant la matrice :

$$\begin{bmatrix} 1 & 2 & 3 \\ -1 & 0 & -1 \\ 3 & -1 & 2 \\ 5 & 3 & 8 \\ 1 & 1 & 2 \end{bmatrix}$$

Mais d'après la proposition 3.31, cela revient à calculer le rang des vecteurs lignes (c'est-à-dire à échelonner la matrice elle-même), ce qui est plus simple a priori. On voit, d'ailleurs, que la troisième ligne est la somme des deux premières lignes qui sont indépendantes entre elles. Donc $\operatorname{rg} A = 2$.

Exercice 31.

3.9 Espace dual

Comme nous l'avons vu en des différents exemples, il existe deux manières de caractériser un sous-espace vectoriel : en se donnant une famille génératrice ou en se donnant des équations linéaires (cf. par exemple, exemple 3. page 8 ; exemple VI page 47, etc). Cette seconde caractérisation peut être plus utile pour certains types de problèmes.

Par exemple considérons le sous-espace F de $\mathbb{R}^5$ engendré par les vecteurs

$$v_1 = (1,3,-2,2,3),\ v_2 = (1,4,-3,4,2),\ v_3 = (2,3,-1,-2,9)\ ;$$

on se demande si le vecteur $v = (1, 2, -2, 0, 1)$ appartient à F. Pour répondre cette question, il faut voir si l'on peut trouver des scalaires $\lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5$ tels que : $v = \lambda_1 v_1 + \lambda_2 v_2 + \lambda_3 v_3 + \lambda_4 v_4 + \lambda_5 v_5$, ce qui exige la résolution d'un système de 5 équations en 5 inconnues. En revanche, si le sous-espace est défini par des équations, cela est beaucoup plus simple. Comme nous l'avons vu (cf. exemple page 47) F peut être caractérisé par les équations

$$\begin{cases} -x_1 + x_2 + x_3 = 0 \\ 4x_1 - 2x_2 + x_4 = 0 \\ -6x_1 + x_2 + x_5 = 0 \end{cases}$$

On voit immédiatement que $v = (1, 2, -2, 0, 1) \notin F$ (les composantes de v ne vérifient pas la première équation.)

La notion d'espace dual permet de définir d'une manière plus précise ce que l'on entend par espace «caractérisé par des équations» et de travailler d'une manière systématique avec des équations plutôt qu'avec des familles de vecteurs.

Définition 3.32 – *On appelle **forme linéaire** sur E une application linéaire* [6]

$$\omega : E \longrightarrow K.$$

*L'ensemble des formes linéaires $\mathcal{L}_K(E, K)$ est noté E^* et est dit **espace dual** de E.*

Soit E est de dimension finie, $\{e_1, ..., e_n\}$ est une base de E et $x = x_1 e_1 + \cdots + x_n e_n$ un vecteur arbitraire de E. Si $\omega \in E^*$, on a :

$$\omega(x) = x_1 \omega(e_1) + \cdots + x_n \omega(e_n) = a_1 x_1 + \cdots + a_n x_n$$

où les $a_i = \omega(e_i)$ sont des scalaires. Ainsi les formes linéaires sur E sont les applications du type :

$$\omega : \quad \begin{array}{ccc} E & \longrightarrow & K \\ {\scriptstyle x = x_1 e_1 + \cdots + x_n e_n} & \longmapsto & {\scriptstyle a_1 x_1 + \cdots + a_n x_n \quad (a_i \in K)} \end{array} \qquad (*)$$

Remarque. – La matrice qui représente ω dans les bases $\{e_i\}$ et $\{1\}$ de K est une matrice ligne :

$$M(\omega)_{e_i, 1} = (a_1, \ldots, a_n)$$

Soit $\omega \in E^*$ une forme linéaire *non nulle*. On a $\operatorname{Im}\omega \subset K$ donc :

$$\dim(\operatorname{Im}\omega) \leq \dim_K K = 1$$

Puisque $\omega \neq 0$ on aura $\dim(\operatorname{Im}\omega) = 1$ et donc $\dim(\operatorname{Ker}\omega) = n - 1$, n étant la dimension de E. On a donc :

Proposition 3.33 – *Soit E de dimension finie n et $\omega \in E^*$, $\omega \neq 0$. On a :*

$$\dim(\operatorname{Ker}\omega) = n - 1$$

*Le noyau de ω est dit **hyperplan de E** déterminé par ω.*

[6] On note en général les formes linéaires par des lettres grecques : ω, θ, φ, etc.

Si ω a l'expression ci-dessus (*), l'hyperplan déterminé par ω est l'ensemble des vecteurs $x \in E$ dont les composantes vérifient l'équation linéaire :

$$a_1\, x_1 + \cdots + a_n\, x_n = 0.$$

L'espace des solutions d'un système d'équations linéaires peut donc être vu comme une intersection d'hyperplans. En d'autres termes, travailler sur l'espace dual revient à travailler sur des équations linéaires (plus précisément : sur les premiers membres des équations linéaires).

Puisque l'ensemble $\mathcal{L}_K(E, E')$ est un espace vectoriel sur K, et K est lui-même un espace vectoriel [7], $E^* = \mathcal{L}_K(E, K)$ est un espace vectoriel et, d'après la proposition 3.14 :

$$\dim_K \mathcal{L}\,(E,\, K) = \dim_K E \times \dim_K K = \dim_K E.$$

On a donc :

Proposition 3.34 – *Si E est de dimension finie, alors :*

$$\boxed{\dim E = \dim E^*}$$

En particulier, E et E^* sont isomorphes [8].

Notons que l'isomorphisme de E sur E^* n'est pas *canonique*, en ce sens qu'il dépend du choix d'une base dans E et d'une base dans E^* (cf. la démonstration du théorème 3.7 page 61).

a) Base duale

Comme nous allons le voir, si l'on choisit une base $\{e_1, \ldots, e_n\}$ de E on peut construire ***canoniquement*** une base $\{\theta_1, \ldots, \theta_n\}$ de E^* :

Théorème 3.35 – *Soit E de dimension finie n et $\{e_1, \ldots, e_n\}$ une base de E. Considérons les formes linéaires $\{\theta_1, \ldots, \theta_n\}$ définies par*

$$\theta_i\,(e_k) = \delta_{ik}$$

où $$\delta_{ik} = \begin{cases} 0 & si\ \ i \neq k \\ 1 & si\ \ i = k \end{cases}$$ *(δ_{ik} est dit symbole de Kronecker).*

Alors $\{\theta_1, \ldots, \theta_n\}$ est une base de E^* dite ***base duale*** *de $\{e_1, \ldots, e_n\}$.*

Démonstration : Remarquons qu'une forme linéaire ω est parfaitement déterminée si on connaît l'image des vecteurs d'une base, car d'après la linéarité de ω on a : $\omega\,(x) = \omega\,(x_1\, e_1 + \cdots + x_n\, e_n) = x_1\, \omega\,(e_1) + \cdots + x_n\, \omega\,(e_n)$. Donc, si on connaît $\omega\,(e_1), \ldots, \omega\,(e_n)$, ω est connue en tout x. En particulier, la définition ci-dessus détermine parfaitement les θ_i. Plus précisément, si $x = x_1\, e_1 + \cdots + x_n\, e_n$, on a :

$$\theta_i\,(x) = x_1\, \theta_i\,(e_1) + \cdots + x_n\, \theta_i\,(e_n) = x_i.$$

[7] Rappelons que K^n est un espace vectoriel sur K, de dimension n, cf. exemple 1, page 5 et remarque 2 page 18.

[8] cf. théorème 3.7.

Donc :

θ_i est l'application qui au vecteur x associe sa i-ième composante dans la base $\{e_1, \dots, e_n\}$.

Pour montrer que $\{\theta_1, \dots, \theta_n\}$ est une base, il suffit de montrer qu'elle est une famille libre, car on sait que $\dim_K E^* = n$.
Supposons qu'il existe $\lambda_1, \dots, \lambda_n \in K$ tels que $\lambda_1\,\theta_1 + \cdots + \lambda_n\,\theta_n = 0$. Ceci veut dire que pour tout vecteur $x \in E$, on a :

$$\lambda_1\,\theta_1\,(x) + \cdots + \lambda_n\,\theta_n\,(x) = 0.$$

En particulier, pour $x = e_k \quad (k = 1, \dots, n)$, on a :

$$\lambda_1\theta_1(e_k) + \cdots + \lambda_k\theta_k(e_k) + \cdots + \lambda_n\theta_n(e_k) = 0,$$

c'est-à-dire $\lambda_k = 0$, pour $k = 1, \dots, n$. Donc la famille $\{\theta_1, \dots, \theta_n\}$ est libre, et, par conséquent, elle est une base. $\square$

Exemple :
Soit $\{e_1,\, e_2,\, e_3\}$ la base de $\mathbb{R}^3$, où :

$$e_1 = (1,\,1,\,1) \quad , \quad e_2 = (1,\,0,\,-1) \quad , \quad e_3 = (0,\,1,\,1).$$

Déterminer la base de $(\mathbb{R}^3)^*$ duale de $\{e_1,\, e_2,\, e_3\}$.

On a :

$$\theta_1\,(e_1) = 1 \quad , \quad \theta_1\,(e_2) = 0 \quad , \quad \theta_1\,(e_3) = 0.$$

Si $\theta_1\,(x) = a\,x_1 + b\,x_2 + c\,x_3$, pour $x = (x_1,\, x_2,\, x_3)$, on doit avoir :

$$\begin{array}{ll} \theta_1\,(e_1) & = 1 \longrightarrow \\ \theta_1\,(e_2) & = 0 \longrightarrow \\ \theta_1\,(e_3) & = 0 \longrightarrow \end{array} \left\{ \begin{array}{rcl} a + b + c & = & 1 \\ a \qquad - c & = & 0 \\ b + c & = & 0 \end{array} \right.$$

ce qui donne $a = 1$, $b = -1$, $c = 1$, c'est-à-dire :

$$\theta_1\,(x) = x_1 - x_2 + x_3.$$

De même, pour θ_2, en imposant $\theta_2\,(e_1) = 0$, $\theta_2\,(e_2) = 1$, $\theta_2\,(e_3) = 0$, on trouve :

$$\left\{ \begin{array}{rcl} a + b + c & = & 0 \\ a \qquad - c & = & 1 \\ b + c & = & 0 \end{array} \right.$$

c'est-à-dire : $a = 0$, $b = 1$, $c = -1$, donc :

$$\theta_2\,(x) = x_2 - x_3.$$

Enfin pour θ_3, en imposant $\theta_3\,(e_1) = 0$, $\theta_3\,(e_2) = 0$, $\theta_3\,(e_3) = 1$, on obtient :

$$\left\{ \begin{array}{rcl} a + b + c & = & 0 \\ a \qquad - c & = & 0 \\ b + c & = & 1 \end{array} \right.$$

c'est-à-dire : $a = -1$, $b = 2$, $c = -1$; donc :

$$\theta_3\,(x) = -x_1 + 2x_2 - x_3$$

La base duale de $\{e_1,\, e_2,\, e_3\}$ est donc donnée par les formes linéaires θ_1, θ_2, θ_3 définies par :

$$\begin{array}{rl} \theta_1(x) = & x_1 - x_2 + x_3 \\ \theta_2(x) = & x_2 - x_3 \\ \theta_3(x) = & -x_1 + 2x_2 - x_3 \end{array}$$

b) Indépendance de formes linéaires

Considérons une famille de formes linéaires $\{\varphi_1, \dots, \varphi_p\}$. Il s'agit de savoir si cette famille est libre, c'est-à-dire si :

$$\lambda_1 \varphi_1 + \cdots + \lambda_p \varphi_p = 0 \quad \Longrightarrow \quad \lambda_1 = 0, \dots, \lambda_p = 0.$$

Le plus simple, en général, est de décomposer les φ_i sur une base de E^* et d'utiliser ensuite les méthodes vues au chapitre 2.

Exemple :

Soient $\varphi_1, \varphi_2, \varphi_3$ *les formes linéaires sur* $\mathbb{R}^4$ *définies par :*

$$\left\{ \begin{array}{rcl} \varphi_1(x) & = & x_1 - 3x_2 + 2x_3 - x_4 \\ \varphi_2(x) & = & x_1 + 2x_2 - x_4 \\ \varphi_3(x) & = & 2x_1 - x_2 + 2x_3 + x_4 \end{array} \right.$$

Montrer que $\{\varphi_1, \varphi_2, \varphi_3\}$ *est une famille de libre de* $(\mathbb{R}^4)^*$.

Soit $\{e_1, \dots, e_4\}$ la base canonique de $\mathbb{R}^4$ et $\{\theta_1, \dots, \theta_4\}$ sa base duale. On a $\theta_i(x) = x_i$, donc :

$$\varphi_1(x) = \theta_1(x) - 3\theta_2(x) + 2\theta_3(x) - \theta_4(x)$$

c'est-à-dire :

$$\varphi_1(x) = (\theta_1 - 3\theta_2 + 2\theta_3 - \theta_4)(x)$$

Puisque cela est vrai pour tout x, ceci signifie que :

$$\varphi_1 = \theta_1 - 3\theta_2 + 2\theta_3 - \theta_4$$

c'est-à-dire les composantes de φ_1 sur la base $\{\theta_1, \dots, \theta_4\}$ sont $(1, -3, 2, -1)$ (en fait les coefficients des x_i dans les $\varphi_i(x)$).
De même :

$$\begin{array}{rl} \varphi_2 = & \theta_1 + 2\theta_2 - \theta_4 \\ \varphi_3 = & 2\theta_1 - \theta_2 + 2\theta_3 + \theta_4 \end{array}$$

En utilisant la méthode du pivot, cela revient donc à échelonner la matrice du système d'équations dont les premiers membres sont les φ_i :

$$\begin{bmatrix} 1 & -3 & 2 & -1 \\ 1 & 2 & 0 & -1 \\ 2 & -1 & 2 & 1 \end{bmatrix} \begin{array}{l} \varphi_1 \\ \varphi_2 \\ \varphi_3 \end{array} \longrightarrow \begin{bmatrix} 1 & -3 & 2 & -1 \\ 0 & 5 & -2 & 0 \\ 0 & 5 & -2 & 3 \end{bmatrix} \begin{array}{l} \varphi_1 \\ \varphi_2' = \varphi_2 - \varphi_1 \\ \varphi_3' = \varphi_3 - 2\varphi_1 \end{array}$$

$$\longrightarrow \begin{bmatrix} 1 & -3 & 2 & -1 \\ 0 & 5 & -2 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix} \begin{array}{l} \varphi_1 \\ \varphi_2' \\ \varphi_3'' = \varphi_3' - \varphi_2' \end{array}$$

Donc $\{\varphi_1, \varphi_2, \varphi_3\}$ est une famille libre.

c) Bidual

Puisque E^* est un espace vectoriel sur K, on peut considérer l'ensemble des formes linéaires sur E^*, c'est-à-dire le dual de E^* dit «bidual» :

$$E^{**} = \mathcal{L}_K\,(E^*,\, K)$$

En dimension finie, on a $\dim_K E^{**} = \dim_K E^* = \dim_K E$, donc, tout comme E^*, E^{**} est isomorphe à E. On a cependant ici une propriété plus forte :

Proposition 3.36 – *En dimension finie, E^{**} est **canoniquement** isomorphe à E.*

> Ceci veut dire que l'on peut construire un isomorphisme de E sur E^{**} sans faire appel à des choix de bases. Ainsi, lorsque la dimension est finie, on peut, en quelque sorte, ***identifier** E **avec** E^{**}*[9].

Démonstration : Considérons l'application

$$\begin{aligned} \Phi :\quad E &\longrightarrow E^{**} \\ x &\longmapsto \Phi_x \end{aligned}$$

où Φ_x est l'application :

$$\begin{array}{rccc} \Phi_x : & E^* & \longrightarrow & K \\ & \omega & \longmapsto & \omega\,(x) \end{array}$$

Montrons d'abord que, pour chaque x fixé, Φ_x est bien une application linéaire $\Big($donc $\Phi_x \in \mathcal{L}_K\,(E^*,\, K) = E^{**}\Big)$. On a, pour $\omega_1,\, \omega_2 \in E^*$:

$$\Phi_x\,(\omega_1 + \omega_2) = (\omega_1 + \omega_2)\,(x) = \omega_1\,(x) + \omega_2\,(x) = \Phi_x\,(\omega_1) + \Phi_x\,(\omega_2)$$

et si $\omega \in E^*$ et $\lambda \in K$:

$$\Phi_x\,(\lambda\,\omega) = (\lambda\,\omega)\,(x) = \lambda\,\Big(\omega\,(x)\Big) = \lambda\,\Phi_x\,(\omega)$$

Donc $\Phi_x \in E^{**}$.
D'autre part Φ est linéaire. Soient en effet $x,\, y \in E$ et $\omega \in E^*$ arbitraire ; on a :

$$\Phi_{x+y}\,(\omega) = \omega\,(x+y) = \omega\,(x) + \omega\,(y) = \Phi_x\,(\omega) + \Phi_y\,(\omega) = (\Phi_x + \Phi_y)\,(\omega)$$

Puisque ω est arbitraire : $\Phi_{x+y} = \Phi_x + \Phi_y$. De même, si $\lambda \in K$, on a pour tout $\omega \in E^*$:

$$\Phi_{\lambda\,x}\,(\omega) = \omega\,(\lambda\,x) = \lambda\,\omega\,(x) = \lambda\,\Phi_x\,(\omega),$$

c'est-à-dire $\Phi_{\lambda\,x} = \lambda\,\Phi_x$. Φ est donc une application linéaire.
Il ne reste plus qu'à montrer que Φ est bijective et pour cela il suffira de montrer que Φ est injective, car on sait que $\dim E^{**} = \dim E$ (cf. corollaire 3.9).
Soit $x \in E$ tel que $\Phi_x = 0$; il s'agit de montrer que $x = 0$. Or $\Phi_x = 0$ signifie que $\forall\, \omega \in E^* \quad \Phi_x\,(\omega) = 0$. Soit $\{e_1, \ldots, e_n\}$ une base de E et $\{\theta_1, \ldots, \theta_n\}$ la base duale ; puisque $\theta_i \in E^*$ on aura $\Phi_x\,(\theta_i) = 0$ c'est-à-dire $\theta_i\,(x) = 0\,, \quad \forall\, i = 1, \ldots, n$. Mais $\theta_i\,(x)$ est la i$^{\text{ème}}$ composante de x sur la base $\{e_1, \ldots, e_n\}$; donc x a toutes ses composantes nulles et, par conséquent $x = 0$. $\square$

> Remarque. – En dimension infinie on peut montrer que Φ est injectif, mais *il n'est pas surjectif*. Donc, si la dimension est infinie, E n'est pas isomorphe à E^{**}, mais uniquement à son image par Φ.

Exercices 32. 33. 34. 35.

[9] En revanche, bien que E et E^* soient isomorphes, on ne peut pas vraiment les identifier car l'isomorphisme $\varphi : E \to E^*$ dépend du choix des bases ; l'élément de E^* qui par φ correspond à un élément de E n'est pas parfaitement défini, car il dépendra justement d'un choix préalable des bases.

3.10 Annulateur d'un sous-espace

Puisqu'en dimension finie E et E^* sont isomorphes, les sous espaces vectoriels de E peuvent être caractérisés par des sous-espaces vectoriels de E^* ; c'est ce que nous allons voir dans ce paragraphe. Comme nous l'avons dit, cette façon de caractériser un sous-espace de E par des équations peut être très utile (cf. page 81).
Étant donné, donc, un sous-espace vectoriel F de E engendré par les vecteurs $\{v_1,\ldots,v_p\}$, nous nous proposons de déterminer un système d'équations linéaires :

$$\begin{cases} a_1\, x_1 \,+\, \cdots \,+\, a_n\, x_n & = 0 \\ \ldots\ldots\ldots\ldots\ldots\ldots & \\ l_1\, x_1 \,+\, \cdots \,+\, l_n\, x_n & = 0 \end{cases}$$

dont les solutions sont exactement tous les vecteurs de F.

Proposition 3.37 – *Soit E un espace vectoriel et F un sous-espace vectoriel de E. L'ensemble F^0 des formes linéaires qui s'annulent sur tous les vecteurs de F,*

$$F^0 = \{\omega \in E^* \,|\, \omega\,(v) = 0\,, \quad \forall\, v \in F\}\ ,$$

est un sous-espace vectoriel de E^*, dit ***annulateur*** *de F.*

La démonstration est immédiate. Soient ω_1 et $\omega_2 \in F^0$; on a, pour tout $v \in F$:

$$(\omega_1 + \omega_2)\,(v) \;=\; \omega_1\,(v) + \omega_2\,(v) \;=\; 0$$

donc $\omega_1 + \omega_2 \in F^0$. De même on montre la stabilité de la loi externe. □

Les propositions suivantes permettent d'effectuer les calculs :

Proposition 3.38 – *Soit $\{v_1, \ldots, v_p\}$ une base de F. Alors :*

$$F^0 = \{\omega \in E^* \mid \omega\,(v_1) = 0, \ldots, \omega\,(v_p) = 0\}$$

En effet si $\omega \in F^0$ on a évidemment $\omega\,(v_1) = 0, \ldots, \omega\,(v_p) = 0$ car les v_i appartiennent à F.
Réciproquement, supposons que :

$$\omega\,(v_1) = 0, \ldots, \omega\,(v_p) = 0.$$

Alors pour tout $v \in F,\ v = \lambda_1\, v_1 + \cdots + \lambda_p\, v_p$, on a :

$$\omega\,(v) = \omega\,(\lambda_1\, v_1 + \cdots + \lambda_p\, v_p) = \lambda_1\, \omega\,(v_1) + \cdots + \lambda_p\, \omega\,(v_p) = 0$$

donc $\omega \in F^0$. □

Proposition 3.39 – *Soit E de dimension finie et F un sous-espace vectoriel de E. On a alors :*

$$\boxed{\dim\, E \,=\, \dim\, F \,+\, \dim\, F^0}$$

Démonstration : Supposons que dim $E = n$ et soit $\{v_1, \dots, v_p\}$ une base de F ; complétons-la en une base de E : $\mathcal{B} = \{v_1, \dots, v_p,\ v_{p+1}, \dots, v_n\}$. Soit $\{\theta_1 \dots, \theta_p,\ \theta_{p+1},$
$\dots \theta_n\}$ la base de $E^\star$ duale de $\mathcal{B}$ et montrons que $\{\theta_{p+1}, \dots, \theta_n\}$ est une base de F^0 (on aura alors : dim $F^0 = n - p =$ dim $E -$ dim F).
Tout d'abord $\theta_{p+1}, \dots, \theta_n \in F^0$ car, $\theta_k\,(v_1) = 0 \dots, \theta_k\,(v_p) = 0$, pour tout pour tout $k = p+1, \dots, n$ d'après la définition de base duale.
D'autre part $\{\theta_{p+1}, \dots, \theta_n\}$ est une famille libre car elle est extraite d'une base. Il ne reste plus qu'à montrer que $\{\theta_{p+1}, \dots, \theta_n\}$ engendre F^0.
Soit $\omega \in F^0$ quelconque ; montrons qu'il existe $\lambda_{p+1}, \dots, \lambda_n \in K$ tels que :

$$\omega = \lambda_{p+1}\,\theta_{p+1} + \cdots + \lambda_n\,\theta_n.$$

Soit $x \in E$: $x = x_1\,v_1 + \cdots + x_p\,v_p + x_{p+1}\,v_{p+1} + \cdots + x_n\,v_n$. Puisque $v_1, \dots, v_p \in F$ et $\omega \in F^0$, on a :

$$\omega\,(x) = x_{p+1}\,\omega\,(v_{p+1}) + \cdots + x_n\,\omega\,(v_n).$$

En posant $\omega\,(v_{p+1}) = \lambda_{p+1}\,, \dots,\ \omega\,(v_n) = \lambda_n$, on a :

$$\omega\,(x) = \lambda_{p+1}\,x_{p+1} + \cdots + \lambda_n\,x_n = \lambda_{p+1}\,\theta_{p+1}\,(x) + \cdots + \lambda_n\,\theta_n\,(x)$$

car $\{\theta_i\}$ est la base duale de $\{v_i, \dots, v_n\}$ et donc $\theta_i\,(x) = x_i$. Puisque cela est vrai pour tout $x \in E$, on a : $\omega = \lambda_{p+1}\,\theta_{p+1} + \cdots + \lambda_n\,\theta_n$. $\square$

Exemple :

Déterminer l'annulateur F^0 du sous-espace vectoriel F de $\mathbb{R}^5$ engendré par les vecteurs

$$v_1 = (1,3,-2,2,3)\ ,\ v_2 = (1,4,-3,4,2)\ ,\ v_3 = (2,3,-1,-2,9)$$

(cf. exemple VI. page 47).

On extrait d'abord une base de $\{v_1, v_2, v_3\}$:

$$\begin{bmatrix} 1 & 3 & -2 & 2 & 3 \\ 1 & 4 & -3 & 4 & 2 \\ 2 & 3 & -1 & -2 & 9 \end{bmatrix} \begin{array}{l} v_1 \\ v_2 \\ v_3 \end{array} \longrightarrow \begin{bmatrix} 1 & 3 & -2 & 2 & 3 \\ 0 & 1 & -1 & 2 & -1 \\ 0 & -3 & 3 & -6 & 3 \end{bmatrix} \begin{array}{l} v_1 \\ v_2 - v_1 = v_2' \\ v_3 - 2\,v_1 = v_3' \end{array}$$

$$\longrightarrow \begin{bmatrix} 1 & 3 & -2 & 2 & 3 \\ 0 & 1 & -1 & 2 & -1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{array}{l} v_1 \\ v_2' \\ v_3' - 3\,v_2' \end{array}$$

Donc $v_1 = (1,3,-2,2,3)$ et $v_2' = (0,1,-1,2,-1)$ forment une base de F.

D'après 3.38 : $F^0 = \{\omega \in E^* \,|\, \omega\,(v_1) = 0,\ \omega\,(v_2') = 0\}$

et d'après 3.39 : dim $F^0 = 5 - 2 = 3$.

Il s'agit donc de déterminer trois formes linéaires indépendantes qui s'annulent sur v_1 et v_2'.
Soit $\omega\,(x) = a_1\,x_1 + \cdots + a_5\,x_5$ une forme linéaire. En imposant $\omega\,(v_1) = 0$, on trouve : $a_1 + 3\,a_2 - 2\,a_3 + 2\,a_4 + 3\,a_5 = 0$; en imposant $\omega\,(v_2') = 0$, on a : $a_2 - a_3 + 2\,a_4 - a_5 = 0$.
Il faut donc résoudre le système :

$$\left\{ \begin{array}{rcl} a_1 + 3\,a_2 - 2\,a_3 + 2\,a_4 + 3\,a_5 & = & 0 \\ a_2 - \ \ a_3 + 2\,a_4 - \ \ a_5 & = & 0 \end{array} \right.$$

Le système est déjà échelonné avec variables libres $a_3,\ a_4,\ a_5$; la solution est :

$$a_1 = -\lambda + 4\,\mu - 6\,\nu \quad , \quad a_2 = \lambda - 2\,\mu + \nu \quad , \quad a_3 = \lambda \quad , \quad a_4 = \mu \quad , \quad a_5 = \nu$$

ce qui donne toutes les formes de F^0.
Si l'on veut trois formes indépendantes, il suffit de donner au triplet $(\lambda,\ \mu,\ \nu)$ successivement les valeurs $(1,0,0)$, $(0,1,0)$, $(0,0,1)$ conformément à la théorie développée au chapitre 2. On trouve alors :

$$\begin{cases} \varphi_1\,(x) = & -\ x_1 + \ \ x_2 + x_3 \\ \varphi_2\,(x) = & \ \ 4\,x_1 - 2\,x_2 + x_4 \\ \varphi_3\,(x) = & -6\,x_1 + \ \ x_2 + x_5 \end{cases}$$

(cf. page 43).

Exercices 36. 37.

EXERCICES

1

1. Soit E l'espace vectoriel des applications continues de $[0,1]$ dans $\mathbb{R}$.
   Montrer que l'application :
   $$\begin{array}{rccc} \psi: & E & \longrightarrow & E \\ & f & \longmapsto & F \end{array}$$
   où $F\,(x) = \displaystyle\int_0^x f\,(t)\,\mathrm{d}t$, est un endomorphisme de E.
2. Soit $f_0 \in E$ fixé. Montrer que l'application :
   $$\begin{array}{rccc} \varphi: & E & \longrightarrow & \mathbb{R} \\ & f & \longmapsto & \int_0^1 f_0\,(t)\,f\,(t)\,\mathrm{d}t \end{array}$$
   est une application linéaire sur E.
3. Montrer que l'application :
   $$\begin{array}{rccc} \delta: & E & \longrightarrow & \mathbb{R} \\ & f & \longmapsto & f\,(0) \end{array}$$
   est une application linéaire (elle est appelée *fonctionnelle de Dirac*).

2 Soit f une application linéaire bijective de E dans F (c'est-à-dire un isomorphisme).
Montrer que f^{-1} est une application linéaire.

3 Soit $h \in \mathbb{R}$ et :

$$\begin{array}{rccc} \varphi_h: & \mathbb{R}_n\,[x] & \longrightarrow & \mathbb{R}_n\,[x] \\ & P & \longmapsto & Q \end{array} \qquad \text{avec} \qquad Q\,(x) = P\,(x+h)$$

Montrer que φ_h est un isomorphisme.

* **4** **Transport de structure**

1. Soit E un espace vectoriel sur un corps K, $\mathcal{A}$ un ensemble non vide. On suppose qu'il existe une bijection $f : \mathcal{A} \longrightarrow E$. On définit alors sur $\mathcal{A}$ les lois suivantes :
   - loi d'addition :
   $$\begin{array}{ccc} \mathcal{A}\times\mathcal{A} & \longrightarrow & \mathcal{A} \\ (a\,,\,b) & \longmapsto & a\oplus b \end{array} \qquad \text{où} \qquad a\oplus b := f^{-1}\Big(f(a)+f(b)\Big)$$

- loi externe :

$$\begin{array}{ccc} K \times \mathcal{A} & \longrightarrow & \mathcal{A} \\ (\lambda,\, a) & \longmapsto & \lambda \cdot a \end{array} \qquad \text{où} \qquad \lambda \cdot a := f^{-1}\Big(\lambda\, f(a)\Big)$$

(en d'autres termes : on prend les images par f des éléments de $\mathcal{A}$, on les compose avec la structure d'espace vectoriel de E et on revient dans $\mathcal{A}$ par l'application f^{-1}).

Montrer que, muni de ces lois $\mathcal{A}$ est un espace vectoriel sur K isomorphe à E. (On dit que la structure de E a été transportée par isomorphisme sur $\mathcal{A}$, ou aussi que la structure de $\mathcal{A}$ est induite par f de celle de E).

2. Soit $\mathbb{R}$ muni de la structure habituelle d'espace vectoriel sur $\mathbb{R}$, $\mathcal{A} = \mathbb{R}_+ \setminus \{0\}$ l'ensemble des nombres réels strictement positifs et :

$$f : \begin{array}{ccc} \mathcal{A} & \longrightarrow & \mathbb{R} \\ a & \longmapsto & \operatorname{Log} a \end{array}$$

Déterminer la structure d'espace vectoriel sur $\mathcal{A}$ obtenue par transport de la structure d'espace vectoriel de $\mathbb{R}$.

3. Soit :

$$f : \begin{array}{ccc}]0,1[& \longrightarrow & \mathbb{R} \\ x & \longmapsto & \operatorname{Log} \dfrac{x}{1-x} \end{array}$$

Montrer que f est une bijection. Déterminer la structure d'espace vectoriel sur $]0,1[$ induite par f.

4. Peut-on mettre sur $\mathbb{R}$ une structure de $\mathbb{R}$-espace vectoriel de dimension 2 ?

5 Soit :

$$f : \begin{array}{ccc} \mathbb{R}^4 & \longrightarrow & \mathbb{R}^3 \\ (x,y,z,t) & \longmapsto & (x-y+z+t\,,\ x+2\,z-t\,,\ x+y+3\,z-3\,t\,) \end{array}$$

Déterminer une base de $\operatorname{Ker} f$ et une base de $\operatorname{Im} f$.

* **6** Soit :

$$M = \begin{pmatrix} 1 & 2 \\ 0 & 3 \end{pmatrix} \qquad \text{et} \qquad f\ : \begin{array}{ccc} \mathcal{M}_2\,(\mathbb{R}) & \longrightarrow & \mathcal{M}_2\,(\mathbb{R}) \\ A & \longmapsto & AM-MA \end{array}$$

Déterminer une base de $\operatorname{Ker} f$ et une base de $\operatorname{Im} f$.

* **7** Soit :

$$\varphi : \begin{array}{ccc} \mathbb{R}_n\,[x] & \longrightarrow & \mathbb{R}^{n+1} \\ P & \longmapsto & \varphi\,(P) = (\,P\,(0),\ P\,(1),\ P\,(2), \ldots, P\,(n)\,) \end{array}$$

Montrer que φ est un isomorphisme.

8 Soient $f : E \longrightarrow F$ et $g = G \longrightarrow E$ linéaires avec g surjective. Montrer que :

$$\operatorname{rg} f = \operatorname{rg}(f \circ g)$$

De même si $f : E \longrightarrow F$ et $h : F \longrightarrow G$ sont linéaires et h est injective, montrer que :

$$\operatorname{rg} f = \operatorname{rg}(h \circ f)$$

En particulier : *en composant à gauche ou à droite par une application linéaire bijective, le rang ne change pas.*

* **9** Soit E un espace vectoriel. On appelle ***projecteur*** un endomorphisme p de E tel que $p^2 = p$.

1. Montrer que si p est un projecteur alors : $E = \operatorname{Ker} p \oplus \operatorname{Im} p$

2. Montrer $\operatorname{Ker} p = \operatorname{Im}(\mathrm{id} - p)$ et que donc : $E = \operatorname{Im} p \oplus \operatorname{Im} q$, où $q = \mathrm{id} - p$.

** **10** **Système complet de projecteurs**

Soit $E = F_1 \oplus \cdots \oplus F_q$: tout $x \in E$ s'écrit d'une manière unique $x = x_1 + \cdots + x_q$. On pose : $P_i(x) := x_i$.

1. Montrer que P_i est un projecteur et que :

$$\begin{cases} P_i \circ P_j &= 0 \qquad (i, j \in \{1, \cdots, p\}, \quad i \neq j) \\ P_1 + \cdots + P_q &= \mathrm{id} \end{cases}$$

Une famille de projecteurs vérifiant les deux propriétés ci-dessus, est dite *système complet de projecteurs*.

2. Réciproquement, on suppose qu'il existe un système complet de projecteurs $\{P_1, \cdots, P_q\}$. Montrer que E est somme directe des images des P_i.

* **11** Soient E et F deux espaces vectoriels sur le même corps K, $f \in \mathcal{L}(E, F)$ et $g \in \mathcal{L}(F, E)$, telles que :

$$f \circ g \circ f = f \qquad \text{et} \qquad g \circ f \circ g = g$$

Montrer que $E = \operatorname{Ker} f \oplus \operatorname{Im} g$ (cf. Appendice 6).

* **12** Soit $f \in \operatorname{End}(E)$ tel que $f^3 = f$. Montrer que $E = \operatorname{Ker} f \oplus E_1 \oplus E_{-1}$ où $E_1 = \{x \in E \mid f(x) = x\}$ et $E_{-1} = \{x \in E \mid f(x) = -x\}$.

13 Existe-t-il des applications linéaires injectives de $\mathbb{R}^2$ dans $\mathbb{R}$? des applications linéaires surjectives de $\mathbb{R}$ dans $\mathbb{R}^2$? Généraliser.

* **14**

1. Soit f un endomorphisme de E. Montrer que pour tout $p \in \mathbb{N}^*$, on a :

$$\operatorname{Ker} f \subset \operatorname{Ker} f^2 \subset \operatorname{Ker} f^3 \subset \cdots \subset \operatorname{Ker} f^p$$

$$\operatorname{Im} f \supset \operatorname{Im} f^2 \supset \operatorname{Im} f^3 \supset \cdots \supset \operatorname{Im} f^p$$

2. On suppose E de dimension finie. Montrer que les propriétés suivantes sont équivalentes :
   (a) $\operatorname{Ker} f = \operatorname{Ker} f^2$
   (b) $\operatorname{Im} f = \operatorname{Im} f^2$
   (c) $E = \operatorname{Im} f \oplus \operatorname{Ker} f$

   (Pour une généralisation en dimension infinie, cf. exercice 16).

15 Soit $f \in \operatorname{End}(\mathbb{R}^3)$ tel que $f^2 = 0$. Montrer qu'il existe $v \in \mathbb{R}^3$ et $g \in \mathcal{L}(\mathbb{R}^3, \mathbb{R})$ tels que, pour tout $x \in \mathbb{R}^3$, $f(x) = g(x)\, v$.

** **16** **Caractère d'un endomorphisme**

Un endomorphisme f est dit ***nilpotent*** s'il existe un entier p tel que $f^p = 0$. Le plus petit entier p tel que $f^p = 0$ est dit *indice de nilpotence de f*.

Soit E un espace vectoriel (non nécessairement de dimension finie). On pose $J_n = \operatorname{Im} f^n$, $K_n = \operatorname{Ker} f^n$. On a (cf. exercice 14) :

$$\begin{array}{ccccccccc} J_1 & \supset & J_2 & \supset & \cdots & \supset & J_n & \supset & \cdots \\ K_1 & \subset & K_2 & \subset & \cdots & \subset & K_n & \subset & \cdots \end{array}$$

I. 1. Montrer que s'il existe m tel que $J_m = J_{m+1}$, alors : $J_m = J_{m+p}$ pour tout $p \in \mathbb{N}$.

2. Montrer que s'il existe m tel que $K_m = K_{m+1}$, alors : $K_m = K_{m+p}$ pour tout $p \in \mathbb{N}$.

II. On pose $J = \bigcap\limits_{n=1}^{\infty} J_n \qquad \mathcal{K} = \bigcup\limits_{n=1}^{\infty} K_n.$

1. On suppose qu'il existe m tel que $J_m = J_{m+1}$ et soit r le plus petit entier tel que $J_r = J_{r+1}$. Montrer que :
   (a) $f(J) \subset J$
   (b) $E = J + K_r$
2. On suppose qu'il existe m tel que $K_m = K_{m+1}$ et soit s le plus petit entier tel que $K_s = K_{s+1}$. Montrer que :
   (a) $f|_{\mathcal{K}}$ est nilpotent
   (b) $f|_J$ est injectif.
   (c) $J_s \cap \mathcal{K} = \{0\}$.

III. On dit que f est de ***caractère fini*** si r et s existent.
Montrer que si f est de caractère fini, alors :

1. $E = J \oplus \mathcal{K}$.
2. $f|_J$ est un automorphisme (c'est-à-dire un endomorphisme bijectif).
3. $f|_{\mathcal{K}}$ est nilpotente.
4. $r = s$ (r est dit ***caractère*** de f).

IV. Réciproquement, on suppose que $E = F \oplus G$ et qu'il existe $f \in \operatorname{End}(E)$, tel que :

- F, G sont stables par f (c'est-à-dire : $f(F) \subset F$, $f(G) \subset G$) ;
- $f|_F$ est un automorphisme ;
- $f|_G$ est nilpotent.

Montrer que f est alors de caractère fini et $F = J$, $G = \mathcal{K}$.

V. Montrer qu'en dimension finie tout endomorphisme est de caractère fini.
Donner un contre-exemple en dimension infinie.

17 Soit f l'endomorphisme de $\mathbb{R}^3$ qui dans la base canonique est représenté par la matrice

$$A = \begin{pmatrix} -2 & 3 & 1 \\ 5 & 1 & 0 \\ 4 & 11 & 3 \end{pmatrix}$$

Déterminer une base de $\operatorname{Ker} f$, une base de $\operatorname{Im} f$ et l'équation linéaire définissant $\operatorname{Im} f$.

* **18**

1. Existe-t-il des applications linéaires $f : \mathbb{R}^4 \longrightarrow \mathbb{R}^3$ telles que $\operatorname{Ker} f$ soit engendré par le vecteur $v = (1, 1, 0, -1)$ et $\operatorname{Im} f$ soit le plan d'équation $x + y - z = 0$?
2. Déterminer la forme générale des matrices qui représentent dans les bases canoniques de $\mathbb{R}^4$ et de $\mathbb{R}^3$ les applications linéaires $f : \mathbb{R}^4 \longrightarrow \mathbb{R}^3$ pour lesquelles $\operatorname{Ker} f = \operatorname{Vect}\{v_1,\, v_2\}$, où : $v_1 = (1, 1, 0, -1)$, $v_2 = (0, 1, -1, 0)$ et $\operatorname{Im} f$ est le plan π d'équation $x + y - z = 0$.
3. Ces matrices forment-elles un espace vectoriel (pour les lois de composition habituelles) ?

* **19** Déterminer la matrice qui représente dans la base canonique de $\mathbb{R}^3$ la projection sur le plan π d'équation $x + 2\,y + 3\,z = 0$ parallèlement à la droite D d'équation $\dfrac{x}{3} = \dfrac{y}{2} = z$.

20 Soit $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ -1 & 4 \end{pmatrix}$. Déterminer, s'il y en a, toutes les matrices B telles que $BA = I$.

* **21** Soit $A = \begin{pmatrix} 1 & -1 & -1 \\ -1 & 1 & -1 \\ -1 & -1 & 1 \end{pmatrix}$. Montrer que A^n est de la forme $\begin{pmatrix} a_n & b_n & b_n \\ b_n & a_n & b_n \\ b_n & b_n & a_n \end{pmatrix}$.
Exprimer a_{n+1} et b_{n+1} en fonction de a_n et b_n. En déduire A^n.

22 Montrer que si $A, B \in \mathcal{M}_n(K)$ et $AB = BA$ on a la ***formule du binôme*** :

$$(A+B)^n = \sum_{k=0}^{n} C_n^k \, A^k \, B^{n-k}$$

A-t-on le même résultat si A et B ne commutent pas ?

* **23** Soient $A \in \mathcal{M}_{p,n}(K)$, $B \in \mathcal{M}_{n,q}(K)$, $B = \|c_1, \dots, c_q\| \quad (c_i \in K^n)$.
Montrer que $AB = \| A\,c_1, \dots, A\,c_q \|$.

* **24** **Produit par blocs**

1. Montrer que :

$$\begin{array}{cc} & \times \begin{pmatrix} B_1 & | B_2 \end{pmatrix} \\ \left(\begin{array}{c} A_1 \\ \hline A_2 \end{array}\right) & = \left(\begin{array}{c|c} A_1\,B_1 & A_1\,B_2 \\ \hline A_2\,B_1 & A_2\,B_2 \end{array}\right) \end{array}, \qquad \text{où :} \quad \begin{array}{ll} A_1 \in \mathcal{M}_{k,n} & A_2 \in \mathcal{M}_{h,n} \\ B_1 \in \mathcal{M}_{n,r} & B_2 \in \mathcal{M}_{n,s} \end{array}$$

2. Montrer que :

$$\begin{array}{cc} & \times \left(\begin{array}{c} B_1 \\ \hline B_2 \end{array}\right) \begin{array}{l} \} \, k \\ \\ \end{array} \\ (\underbrace{A_1}_{k} \mid A_2) & = A_1\,B_1 + A_2\,B_2 \end{array} \qquad \text{où} \quad \begin{array}{ll} A_1 \in \mathcal{M}_{n,k} & A_2 \in \mathcal{M}_{n,h} \\ B_1 \in \mathcal{M}_{k,p} & B_2 \in \mathcal{M}_{h,q} \end{array}$$

3. Montrer que :

$$\begin{array}{cc} & \times \quad k\{ \left(\begin{array}{c|c} B_1 & B_2 \\ \hline B_3 & B_4 \end{array}\right) \\ \left(\begin{array}{c|c} \overbrace{A_1}^{k} & A_2 \\ \hline A_3 & A_4 \end{array}\right) & = \left(\begin{array}{c|c} A_1\,B_1 + A_2\,B_3 & A_1\,B_2 + A_2\,B_4 \\ \hline A_3\,B_1 + A_4\,B_3 & A_3\,B_2 + A_4\,B_4 \end{array}\right) \end{array}$$

En définitive : *on peut toujours faire le produit par blocs, pourvu qu'on puisse le faire...*, c'est-à-dire pourvu que les coupures soient faites de manière à pouvoir effectuer les produits des matrices.

25 Dans $\mathbb{R}^n$ on appelle *longueur* du vecteur $v = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$ le scalaire $\|v\| = \sqrt{a_1^2 + \cdots + a_n^2}$. Montrer que les rotations de centre 0 dans le plan $\mathbb{R}^2$ conservent la longueur des vecteurs (la rotation étant définie par la matrice de l'exemple 5, page 69).

26 Soit $A = \begin{pmatrix} 1 & 2 & -1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}$. Calculer A^{-1}.

27 Soit N une matrice carrée. On dit que N est *nilpotente* s'il existe $p \in \mathbb{N}$ tel que $N^p = 0$ (cf. exercice 16) . Montrer que si N est nilpotente, $I - N$ est inversible.
Application. Calculer l'inverse de la matrice :

$$A = \begin{pmatrix} 1 & a & b & c \\ 0 & 1 & a & b \\ 0 & 0 & 1 & a \\ 0 & 0 & 0 & 1 \end{pmatrix}, \qquad \text{où :} \qquad a, b, c \in \mathbb{R}$$

28 Soit f la projection de $\mathbb{R}^3$ sur le plan π d'équation $x + 2\,y + 3\,z = 0$ parallèlement à la droite (d) d'équation $\dfrac{x}{3} = \dfrac{y}{2} = \dfrac{z}{1}$. Ecrire la matrice A' de f dans la base $\{v_1, v_2, v_3\}$ où $\{v_1, v_2\}$ est une base de π et v_3 est une base de (d). En déduire la matrice de f dans la base canonique de $\mathbb{R}^3$ et retrouver ainsi le résultat de l'exercice 19.

29 Soit $\{e_1, e_2, e_3\}$ une base de $\mathbb{R}^3$ et $e'_1 = e_1 + e_2 - e_3$, $e'_2 = e_1 - e_2 + e_3$, $e'_3 = -e_1 + e_2 + e_3$. Montrer que $\{e'_1, e'_2, e'_3\}$ est une base de $\mathbb{R}^3$.
Soit f l'endomorphisme de $\mathbb{R}^3$ qui dans la base $\{e_i\}$ est représenté par la matrice $A = \dfrac{1}{2}\begin{pmatrix} a-b & a+c & c-b \\ b-a & c-a & b+c \\ a+b & a-c & b-c \end{pmatrix}$. Calculer la matrice de f dans la base $\{e'_i\}$.

* **30** Soit A une matrice carrée d'ordre n. On appelle ***trace*** de A la somme des termes de la diagonale principale :

$$\operatorname{Tr} A = a_{11} + a_{22} + \cdots + a_{nn}$$

1. Montrer que l'application $\operatorname{Tr} : \mathcal{M}_n(K) \longrightarrow K$ est une application linéaire.
2. Montrer que $\operatorname{Tr}(AB) = \operatorname{Tr}(BA)$. En déduire que *deux matrices semblables ont même trace* et que, par conséquent, en dimension finie on peut définir la trace d'un endomorphisme f par $\operatorname{Tr} f = \operatorname{Tr} M(f)_{e_i}$, $\{e_i\}$ étant une base quelconque.
3. Montrer que l'on ne peut pas trouver de matrices $A, B \in \mathcal{M}_n(\mathbb{C})$ telles que :
$$AB - BA = I$$
4. Soit $\mathcal{E}$ l'espace vectoriel des fonctions dérivables de $\mathbb{R}$ dans $\mathbb{R}$ et D, D_1 les endomorphismes de $\mathcal{E}$ définis par
$$(Df)(x) = f'(x) \qquad D_1\,f(x) = x\,f(x)$$
Calculer le «commutateur» $[D, D_1]$ défini par $[D, D_1] := D \circ D_1 - D_1 \circ D$.

31 Soit l'application :

$$\begin{array}{rccc} f: & \mathbb{R}_2[x] & \longrightarrow & \mathbb{R}_2[x] \\ & P & \longmapsto & (ax+1)\,P + (bx^2 + c)\,P' \end{array}$$

Quelles relations doivent vérifier a, b, c pour que f soit un endomorphisme de $\mathbb{R}_2[x]$? Déterminer, dans ce cas, le rang de f

32 Soit la base de $\mathbb{R}^n$:

$$\begin{aligned} e_1 &= (1,1,\ldots,1),\ \ e_2 = (0,1,1,\ldots,1),\ \ e_3 = (0,0,1,1,\ldots,1),\ldots, \\ e_n &= (0,0,\ldots,0,1) \end{aligned}$$

Déterminer la base de $(\mathbb{R}^n)^*$ duale de celle-ci.

33 On considère les trois formes linéaires sur $\mathbb{R}^3$:

$$\varphi_1(x) = 2x_1 - x_2 + 3x_3\ ,\quad \varphi_2(x) = 3x_1 - 5x_2 + x_3\ ,\quad \varphi_3(x) = 4x_1 - 7x_2 + x_3$$

Forment-elles une base du dual de $\mathbb{R}^3$? Déterminer les éventuelles relations linéaires.

34 Montrer que les formes linéaires :

$$\varphi_1(x) = x_1 + 2x_2 + x_3\ ,\quad \varphi_2(x) = 2x_1 + 3x_2 + 3x_3\ ,\quad \varphi_3(x) = 3x_1 + 7x_2 + x_3$$

forment une base de $(\mathbb{R}^3)^*$. Déterminer la base duale de celle-ci [en identifiant $\mathbb{R}^3$ avec $(\mathbb{R}^3)^{**}$].

*

35 Soient E, F deux espaces vectoriels sur K et $f : E \longrightarrow F$ une application linéaire. On définit une application (dite ***transposée*** de f) :

$$^t f : F^* \longrightarrow E^*$$

en posant, pour toute $\varphi \in F^*$: $\ ^t f(\varphi) = \varphi \circ f$.

1. Montrer que $^t f$ applique bien F^* dans E^*.
2. Montrer que $^t f$ est une application linéaire.
3. Vérifier les relations suivantes :
   (a) $^t(f+g) = {}^t f + {}^t g$
   (b) $^t(\lambda f) = \lambda({}^t f)$
   (c) $^t(f \circ g) = {}^t g \circ {}^t f$
   (d) $^t({}^t f) = f$
   (e) si f est bijective, $^t f$ est aussi bijective et :
   $$^t(f^{-1}) = ({}^t f)^{-1}$$
4. On suppose que E et F sont de dimension finie et soient $\{e_i\}$, $\{\varepsilon_j\}$ des bases de E et F respectivement, $\{\varphi_i\}$, $\{\psi_j\}$ leurs bases duales. On note $A = M(f)_{e_i,\, \varepsilon_j}$. Montrer que :
   $$M({}^t f)_{\psi_j,\, \varphi_i} = {}^t A$$
5. En déduire les relations sur les matrices :
   $$\begin{aligned} {}^t(AB) &= {}^t B\, {}^t A \\ {}^t(A^{-1}) &= ({}^t A)^{-1} \end{aligned}$$

*

36 Soient E un espace vectoriel de dimension finie et F, G des sous-espaces vectoriels de E. Montrer que :

1. $F \subset G \quad \Longrightarrow \quad G^0 \subset F^0$
2. $(F+G)^0 = F^0 \cap G^0$
3. $F^{00} = F$
4. $(F \cap G)^0 = F^0 + G^0$

*

37 Soit $f : E \longrightarrow F$ une application linéaire. Montrer que :

$$(\operatorname{Im} f)^0 = \operatorname{Ker}({}^t f)$$

En déduire que si E et F sont de dimension finie :

$$\operatorname{rg} f = \operatorname{rg}\, {}^t f$$

et que, par conséquent, pour toute $A \in \mathcal{M}_{p,n}(K)$: $\operatorname{rg} A = \operatorname{rg}^t A$

INDICATIONS

1 1. La primitive d'une fonction continue est continue. Vérifier la linéarité en utilisant la linéarité de l'intégrale.
2. et 3. Simples vérifications.

2 Soient $z = f^{-1}(a+b)$ et $w = f^{-1}(a) + f^{-1}(b)$. On a

$$a + b = f(z) \quad \text{et} \quad f\Big(f^{-1}(a) + f^{-1}(b)\Big) = f(w).$$

En déduire d'abord $f(z) = f(w)$, puis $z = w$. Démonstration analogue pour la loi externe.

3 Remarquer que $\varphi_h^{-1} = \varphi_{-h}$.

4

1. Simples vérifications. Par construction $f(a \oplus b) = f(a) + f(b)$ et $f(\lambda \cdot a) = \lambda\, f(a)$, donc f est un isomorphisme
2. $a \oplus b = e^{\operatorname{Log} a + \operatorname{Log} b} = e^{\operatorname{Log} ab} = ab$
   $\lambda \cdot a = e^{\lambda \operatorname{Log} a} = a^{\lambda}$ (cf. exercice 2. chapitre 1).
3. On trouve :
$$a \oplus b = \frac{ab}{(1-a)(1-b) + ab} \qquad \lambda \cdot a = \frac{a^{\lambda}}{a^{\lambda} + (1-a)^{\lambda}}$$
4. Aussi surprenant que cela puisse paraître, il existe des bijections de $\mathbb{R}$ sur $\mathbb{R}^2$ (Cantor).

5 Pour $\operatorname{Ker} f$ on trouve le système :

$$\begin{cases} x - y + z + t = 0 \\ y + z - 2t = 0 \end{cases}$$

Une base est, par exemple, $v_1 = (2, 1, -1, 0)$, $v_2 = (1, 2, 0, 1)$.
Pour $\operatorname{Im} f$ on étudie la compatibilité du système :

$$\begin{cases} x - y + z + t = x' \\ x + 2z - t = y' \\ x + y + 3z - 3t = z' \end{cases}$$

On trouve $x' - 2y' + z' = 0$ et une base est, par exemple :

$$\{w_1 = (1, 1, 1)\,,\ w_2 = (0, 1, 2)\}$$

(cf. exercice 18. pour une méthode plus simple pour la détermination de $\operatorname{Im} f$).

6 On a :

$$f\begin{pmatrix} x & y \\ z & t \end{pmatrix} = \begin{pmatrix} 2z & 2x + 2y - 2t \\ -2z & 2z \end{pmatrix}$$

Base de $\operatorname{Ker} f$:

$$\left\{ M_1 = \begin{pmatrix} -1 & 1 \\ 0 & 0 \end{pmatrix} \quad , \quad M_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right\}$$

$$\dim(\operatorname{Im} f) = 4 - 2 = 2$$

$$B \in \operatorname{Im} f \Longleftrightarrow B = 2z \begin{pmatrix} -1 & 0 \\ -1 & 1 \end{pmatrix} + 2(x + y - t) \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

7 Vérifier d'abord la linéarité.
On a : $\dim \mathbb{R}_n[x] = \dim \mathbb{R}^{n+1}$. Il suffit donc de vérifier que φ est injective. Utiliser le fait qu'un polynôme non nul de degré n admet au plus n racines.

8 Si g est surjective $g(G) = E$ donc $f \circ g(G) = f(E)$ c'est-à-dire $\operatorname{Im}(f \circ g) = \operatorname{Im} f$.
Si h est injective, soit $\{v_i\}_{i \in I}$ avec $v_i = f(w_i)$ base de $\operatorname{Im} f$; on prend leurs images par h.

9 1. – Utiliser la propositions 1.29. Soit $y \in \operatorname{Ker} p \cap \operatorname{Im} p$ (c'est-à-dire : $p(y) = 0$ et il existe $x \in E$ tel que $y = p(x)$). Utiliser $p^2 = p$ pour montrer que $y = 0$. Pour tout $x \in E$ poser : $x = p(x) + (x - p(x))$. Vérifier que $x - p(x) \in \operatorname{Ker} p$.

2. – $p \circ (\mathrm{id} - p) = 0$, donc $\operatorname{Im}(\mathrm{id} - p) \subset \operatorname{Ker} p$. D'autre part, si $x \in \operatorname{Ker} p$, on a : $x = x - p(x)$, donc $x \in \operatorname{Im}(\mathrm{id} - p)$.

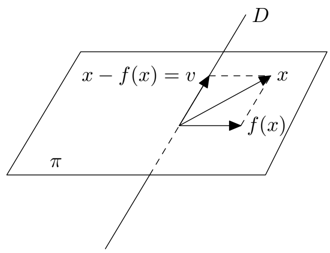


Figure 10

10 1. Simples vérifications.

2. Puisque $\mathrm{id} = P_1 + \cdots + P_q$ on a facilement $E = \operatorname{Im} P_1 + \cdots + \operatorname{Im} P_q$.
Soit $x \in E$ arbiraire, $x = x_1 + \cdots + x_q$ et $x = x'_1 + \cdots + x'_q$ deux décompositions de x. Par différence :

$$0 = (x_1 - x'_1) + \cdots + (x_q - x'_q)$$

En faisant agir P_i, on a $x_i - x'_i = 0$. Donc la décomposition est unique.

11 Employer la méthode dite ***"d'analyse et synthèse"*** . Il s'agit de montrer que tout $x \in E$ s'écrit d'une manière unique $x = y + z$ avec $y \in \operatorname{Ker} f$, $z \in \operatorname{Im} g$.
Analyse. *On suppose que la décomposition existe* pour tout $x \in E$. Soit $x = y + z$ avec $y \in \operatorname{Ker} f$ et $z \in \operatorname{Im} g$, c'est-à-dire : $f(y) = 0$ et $z = g(u)$. On essaie d'exprimer y et z en fonction de x.
De $x = y + z$, en appliquant f on obtient : $f(x) = f(z)$ c'est-à-dire $f(x) = f \circ g(u)$.
En appliquant g : $g \circ f(x) = g \circ f \circ g(u) = g(u)$ et donc : $g \circ f(x) = z$. On en déduit $y = x - g \circ f(x)$. Donc, si la décomposition existe, elle est nécessairement du type :

$$x = \underbrace{x - g\Big(f(x)\Big)}_{\in \operatorname{Ker} f} + \underbrace{g\Big(f(x)\Big)}_{\in \operatorname{Im} g}$$

ce qui montre *l'unicité*.
Synthèse. Il ne reste plus qu'à montrer que $g\Big(f(x)\Big) \in \operatorname{Im} g$ (ce qui est évident) et $x - g\Big(f(x)\Big) \in \operatorname{Ker} f$ (utiliser $f \circ g \circ f = f$). Ceci montre *l'existence* de la décomposition pour tout $x \in E$.
Ainsi l'analyse montre l'unicité de la décomposition et détermine les composantes de la décomposition ; la synthèse est une vérification qui montre l'existence.

12 Utiliser la méthode de l'exercice 11. On part de

$$x = x_0 + x_1 + x_{-1} \quad \text{avec} \quad x_0 \in \operatorname{Ker} f \ , \ x_1 \in E_1 \ , \ x_{-1} \in E_{-1}$$

Appliquer f, puis, à nouveau f. On trouve :

$$x_1 = \frac{f(x) + f^2(x)}{2} \quad , \quad x_2 = \frac{f^2(x) - f(x)}{2} \quad , \quad x_0 = x - f^2(x)$$

13 Non, en vertu du théorème du rang.
Si $m > n$ il n'existe ni d'applications linéaires injectives de $\mathbb{R}^m$ dans $\mathbb{R}^n$, ni d'applications linéaires surjectives de $\mathbb{R}^n$ dans $\mathbb{R}^m$.

14

1. $x \in \operatorname{Ker} f^k \implies f^k(x) = 0 \implies f^{k+1}(x) = 0 \implies x \in \operatorname{Ker} f^{k+1}$. Puisque f est un endomorphisme, on a : $f(E) \subset E$, d'où $f^k\Big(f(E)\Big) \subset f^k(E)$, c'est-à-dire $\operatorname{Im} f^{k+1} \subset \operatorname{Im} f^k$.
2. D'après 1. :

$$\begin{aligned}\text{(a))} \quad &\iff (\dim \operatorname{Ker} f = \dim \operatorname{Ker} f^2) \iff \Big(n - \operatorname{rg} f = n - \operatorname{rg} f^2\Big)\\ &\iff \operatorname{rg} f = \operatorname{rg} f^2\end{aligned}$$

ce qui, toujours d'après 1, équivaut à (b).

Pour (c) montrer que $\operatorname{Im} f \cap \operatorname{Ker} f = \{0\}$ et utiliser le théorème 1.33 ainsi que le théorème du rang.

15 Montrer d'abord que $f^2 = 0 \iff \operatorname{Im} f \subset \operatorname{Ker} f$.

Utiliser le théorème du rang pour montrer que $\operatorname{rg} f \leq 1$. Il existe donc $v \in \mathbb{R}^3$ tel que pour tout $x \in \mathbb{R}^3$ on peut trouver $\lambda_x \in \mathbb{R}$ tel que $f(x) = \lambda_x v$. Montrer que λ_x dépend linéairement de x.

16 I. 1. $f^m(E) = f^{m+1}(E) \Longrightarrow f^{m+1}(E) = f^{m+r}(E)$, etc.

2. Soit $x \in K_{m+2}$: $f^{m+2}(x) = 0 \implies f^{m+1}\Big(f(x)\Big) = 0 \implies f(x) \in K_{m+1} = K_m \implies f^{m+1}(x) = 0$, etc.

II. 1. (a) On a $J = J_r$ d'où $f(J) = f(J_r) = J_{r+1} = J_r = J$.

(b) Soit $x \in E$; $f^r(x) \in J_r = J_{2r} \implies f^r(x) = f^{2r}(x')$; soit $y = f^r(x')$ ($y \in J_r = J$) et $z = x - y$. On a $f^r(z) = f^r(x) - f^r(y) = 0$ donc $z \in K_r$.

2. (a) On a $\mathcal{K} = K_s$. Si $x \in \mathcal{K}$, $f^s(x) = 0$; donc $f|_{\mathcal{K}}$ est nilpotente.

(b) Soit $x \in J$ tel que $f(x) = 0$; puisque $J = \cap_{n=1}^{\infty} J_n : x \in J_n$, $\forall\, n \geq 1$. En particulier $x \in J_s : x = f^s(x')$. D'où :

$$0 = f(x) = f^{s+1}(x') \implies x' \in K_{s+1} \equiv K_s \Longrightarrow f^s(x') = 0 \implies x = 0.$$

(c) Soit $x \in J_s \cap \mathcal{K}$; on a $x = f^s(x')$ et $f^s(x) = 0$ d 'où :

$$f^{2s}(x') = 0 \implies x' \in K_{2s} \equiv K_s \implies f^s(x') = 0 \Longrightarrow x = 0$$

III.1. On a : $J = J_r = J_{r+s}$ et $\mathcal{K} = K_s = K_{r+s}$. Il faut montrer que $E = J_{r+s} \oplus K_{r+s}$. Utiliser les résultats de la partie II.

2. cf. II.
3. cf. II.
4. On a $E = J_r \oplus K_s$.

$$\begin{aligned} f^s(E) &= f^s(J_r) &&= J_r = J \\ f^{s+1}(E) &= f(J) &&= J \end{aligned}$$

donc $J_s = J_{s+1}$ et par conséquent $r \leq s$.
Soit $x \in K_s$; $f^r(x) \in J_r$ et $f^s\Big(f^r(x)\Big) = f^r\Big(f^s(x)\Big) = 0$ (car $x \in K_s$) ; donc $f^r(x) \in J_r \cap K_s = \{0\}$; ainsi $K_s \subset K_r$.
D'autre part $r \leq s$ donc $K_s \supset K_r$ et par conséquent $K_s = K_r$. Ceci implique que la suite des noyaux est stationnaire déjà à l'indice r et donc $s \leq r$.

IV. Soit r l'indice de nilpotence de $f|_G$. On a :

$$f^r(E) = f^r(F) + f^r(G) = f^r(F) = F$$

car $f|_F$ est un automorphisme ; de même : $f^{r+1}(E) = F$. Donc :

$$F = \operatorname{Im} f^r = \operatorname{Im} f^{r+1}$$

D'autre part : $G \subset \operatorname{Ker} f^r \subset \operatorname{Ker} f^{r+1}$. Soit $x \in \operatorname{Ker} f^{r+1}$, $x = x_F + x_G$ avec $x_F \in F$, $x_G \in G$. $f^{r+1}(x) = 0 \implies f^{r+1}(x_F) = 0$; comme $f|_F$ est un automorphisme, $f^{r+1}|_F$ est aussi un automorphisme, donc $x_F = 0$ et par conséquent $\operatorname{Ker} f^{r+1} = G$. Ce qui montre que $G = \operatorname{Ker} f^r = \operatorname{Ker} f^{r+1}$.

V. Si les dimensions des K_s croissaient strictement, elles finiraient par dépasser la dimension de E.
Pour le contre-exemple, considérer l'endomorphisme $f : \begin{array}{ccc} \mathbb{R}[x] & \longrightarrow & \mathbb{R}[x] \\ P & \longmapsto & P' \end{array}$. On a : $\mathcal{K} = \mathbb{R}[x]$ et $J = \mathbb{R}[x]$; donc $\mathcal{K}$ et J ne sont pas supplémentaires dans $\mathbb{R}[x]$.

17 Soit $v \in \operatorname{Ker} f$ et $V = M(v)_{e_i} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$. En imposant $AV = 0$ on trouve le système :

$$\begin{cases} -2x + \ \ 3y + z &= 0 \\ 5x + \quad y &= 0 \\ 4x + 11y + 3z &= 0 \end{cases}$$

ce qui donne $V = \begin{pmatrix} 1 \\ -5 \\ 17 \end{pmatrix}$ (à un facteur de proportionnalité près).
Donc $\dim \operatorname{Ker} f = 1$ et $\dim \operatorname{Im} f = 2$. Pour avoir une base de $\operatorname{Im} f$ il suffit de prendre deux vecteurs colonnes indépendants, par exemple $f(e_1)$ et $f(e_2)$ c'est-à-dire $\begin{pmatrix} -2 \\ 5 \\ 4 \end{pmatrix}$, $\begin{pmatrix} 3 \\ 1 \\ 11 \end{pmatrix}$ (on a identifié un vecteur de $\mathbb{R}^3$ avec sa matrice colonne dans la base canonique, ce que l'on fera souvent dans la suite).
L'équation de $\operatorname{Im} f$ est $3x + 2y - z = 0$.

18
1. Non, en vertu du théorème du rang.
2. Soit :

$$A = M(f)_{e_i,\varepsilon_j} = \begin{array}{c} \begin{pmatrix} a_1 & a_2 & a_3 & a_4 \\ b_1 & b_2 & b_3 & b_4 \\ c_1 & c_2 & c_3 & c_4 \end{pmatrix} \\ \begin{array}{cccc} \uparrow & \uparrow & \uparrow & \uparrow \\ f(e_1) & f(e_2) & f(e_3) & f(e_4) \end{array} \end{array}$$

   En imposant que $f(e_i) \in \pi$ on trouve : $c_i = a_i + b_i$ $(i = 1, \dots, 4)$. Ces conditions sont équivalentes à $\operatorname{Im} f \subset \pi$.
   Imposer, ensuite, $v_1, v_2 \in \operatorname{Ker} f$ (c'est-à-dire $\operatorname{Vect}\{v_1, v_2\} \subset \operatorname{Ker} f$). On trouve :

$$A = \begin{pmatrix} a_1 & a_2 & a_2 & a_1 + a_2 \\ b_1 & b_2 & b_2 & b_1 + b_2 \\ a_1 + b_1 & a_2 + b_2 & a_2 + b_2 & a_1 + a_2 + b_1 + b_2 \end{pmatrix}$$

   Il reste à imposer : $\operatorname{Im} f = \pi$ (ce qui, d'après le théorème du rang, implique $\dim \operatorname{Ker} f = 2$, donc $\operatorname{Vect}\{v_1, v_2\} = \operatorname{Ker} f$). Comme $\operatorname{Im} f \subset \pi$, ceci équivaut à $\operatorname{rg} f = \dim \pi = 2$, ce qui s'exprime par le fait que les colonnes ne sont pas toutes proportionnelles à l'une d'entre elles. On trouve les matrices ci-dessus avec la restriction : $a_1 b_2 - a_2 b_1 \neq 0$.
3. Non (il n'y a pas la matrice nulle).

19 $f(x) + v = x$, où, $v = (3\lambda,\ 2\lambda,\ \lambda)$. On détermine λ de manière à ce que $x - v \in \pi$.
On a :

$$(x - 3\lambda) + 2(y - 2\lambda) + 3(z - \lambda) = 0$$
$$\Longrightarrow \lambda = \frac{1}{10}(x + 2y + 3z)$$

Calculer $f(e_1)$, $f(e_2)$, $f(e_3)$. On trouve :

$$M(f)_{e_i} = \frac{1}{10}\begin{pmatrix} 7 & -6 & -9 \\ -2 & 6 & -6 \\ -1 & -2 & 7 \end{pmatrix}$$

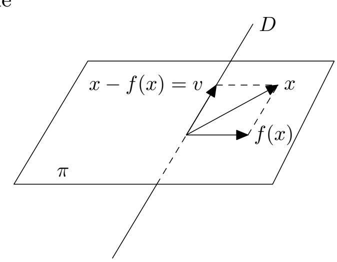


Figure 11

20 Nécessairement $I = I_2$ et $B \in \mathcal{M}_{2,3}(\mathbb{R})$. On trouve :

$$B = \begin{pmatrix} -2-8\,c & 1+3\,c & c \\ \dfrac{3}{2}-8\,c' & -\dfrac{1}{2}+3\,c' & c' \end{pmatrix}, \quad \text{avec } c, c' \in \mathbb{R}$$

21 On trouve :

$$\begin{cases} a_{n+1} = a_n - 2\,b_n \\ b_{n+1} = -a_n \end{cases} \quad \text{c'est-à-dire :} \quad \begin{cases} a_{n+1} = a_n + 2\,a_{n-1} \\ b_{n+1} = -a_n \end{cases}$$

En utilisant les résultats de l'exercice 19. du chapitre 1, on trouve :

$$a_n = \frac{1}{3}\left(2^{n+1} + (-1)^n\right) \qquad b_n = -\frac{1}{3}\left(2^n + (-1)^{n-1}\right)$$

22 Démonstration par récurrence. Si $AB \neq BA$, on a :

$$(A+B)^2 = A^2 + AB + BA + B^2 \neq A^2 + 2AB + B^2$$

23 Utiliser les notations suivantes. Soit $l = (a_1, \dots, a_n)$ une matrice ligne et $c = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$ une matrice colonne. On a $l\,c = a_1\,b_1 + \cdots + a_n\,b_n$. Si $l_1, \dots, l_p$ sont les lignes de A, employer la notation $A = \left\| \begin{matrix} l_1 \\ \vdots \\ l_p \end{matrix} \right\|$. On a alors : $A\,c = \begin{pmatrix} l_1\,c \\ \vdots \\ l_p\,c \end{pmatrix}$ et :

$$\begin{array}{cc} \times & \| \quad c_1, \dots, \quad c_q \; \| \\ \left\| \begin{matrix} l_1 \\ \vdots \\ l_p \end{matrix} \right\| & = \begin{pmatrix} l_1\,c_1 \dots l_1\,c_q \\ \dots\dots\dots\dots \\ l_p\,c_1 \dots l_p\,c_p \end{pmatrix} = \|\, A\,c_1, \dots, A\,c_q \,\| \end{array}$$

24 Utiliser les notations de l'exercice 23.

Pour 1. :

$$A_1 = \left\| \begin{matrix} l_1 \\ \vdots \\ l_k \end{matrix} \right\|, \; A_2 = \left\| \begin{matrix} m_1 \\ \vdots \\ m_h \end{matrix} \right\|, \; B_1 = \|\, v_1, \dots, v_r \,\|, \; B_2 = \|\, w_1, \dots, w_s \,\|$$

On a :

$$\begin{array}{cc} \times & \| \quad v_1, \dots, \quad v_r \qquad\qquad w_1, \dots, \quad w_s \quad \| \\ \left\| \begin{matrix} l_1 \\ \vdots \\ l_k \\ \hline m_1 \\ \vdots \\ m_h \end{matrix} \right\| & = \left(\begin{array}{c|c} l_1\,v_1 \dots l_1\,v_r & l_1\,w_1 \dots l_1\,w_s \\ \dots\dots\dots\dots & \dots\dots\dots\dots \\ l_k\,v_1 \dots l_k\,v_r & l_k\,w_1 \dots l_k\,w_s \\ \hline m_1\,v_1 \dots m_1\,v_r & m_1\,w_1 \dots m_1\,w_s \\ \dots\dots\dots\dots & \dots\dots\dots\dots \\ m_h\,v_1 \dots m_h\,v_r & m_h\,w_1 \dots m_h\,w_s \end{array} \right) \end{array}$$

Méthode analogue pour les autres.

25 Il faut vérifier que $\|f(v)\| = \|v\|$, $\forall v \in \mathbb{R}^2$. Dans la base canonique de $\mathbb{R}^2$, la matrice de f est $\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$. Si $v = \begin{pmatrix} a \\ b \end{pmatrix}$ on a $f(v) = \begin{pmatrix} a\cos\theta - b\sin\theta \\ a\sin\theta + b\cos\theta \end{pmatrix}$.

26 $X' = A\,X \iff \left\{ \begin{array}{rcl} x + 2\,y + z & = & x' \\ y + 2\,z & = & y' \\ z & = & z' \end{array} \right.$.

En résolvant par rapport à x, y, z :

$$\left\{ \begin{array}{rcr} x & = & x' - 2\,y' + 3\,z' \\ y & = & y' - 2\,z' \\ z & = & z' \end{array} \right. \qquad \text{donc} \qquad A^{-1} = \begin{pmatrix} 1 & -2 & 3 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{pmatrix}$$

27 On a :

$$(I - N)\,(I + N + N^2 + \cdots + N^{p-1}) = I$$

donc :

$$(I - N)^{-1} = I + N + N^2 + \cdots + N^{p-1}$$

On trouve :

$$A^{-1} = \begin{pmatrix} 1 & -a & a^2 - b & -a^3 + 2ab - c \\ 0 & 1 & -a & a^2 - b \\ 0 & 0 & 1 & -a \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

28 $A' = M(f)_{v_i} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$. En prenant, par exemple : $v_1 = \begin{pmatrix} 2 \\ -1 \\ 0 \end{pmatrix}$, $v_2 = \begin{pmatrix} 3 \\ 0 \\ -1 \end{pmatrix}$, $v_3 = \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix}$, on a : $P = P_{e_i \longrightarrow v_i} = \begin{pmatrix} 2 & 3 & 3 \\ -1 & 0 & 2 \\ 0 & 1 & 1 \end{pmatrix}$.

Calculer $A = P\,A'\,P^{-1}$.

29

$$M(f)_{e'_i} = \begin{pmatrix} 0 & 0 & c \\ a & 0 & 0 \\ 0 & b & 0 \end{pmatrix}$$

30

1. Vérifier que $\mathrm{Tr}(A + B) = \mathrm{Tr}\,A + \mathrm{Tr}\,B$ et $\mathrm{Tr}(\lambda\,A) = \lambda\ \mathrm{Tr}\,A$.
2. Soit $C = AB$, $C = (c_{ik})$, avec $c_{ik} = \sum_{j=1}^{n} a_{ij}\,b_{jk}$. On a : $\mathrm{Tr}\ C = \sum_{i,j=1}^{n} a_{ij}\,b_{ji}$.
   Soit $D = BA = (d_{ik})$, avec $d_{ik} = \sum_{j=1}^{n} b_{ij}\,a_{jk}$. On a : $\mathrm{Tr}\,D = \sum_{i,j=1}^{n} b_{ij}\,a_{ji}$.
   Changer dans cette expression les notations en remplaçant i par j et j par i et vérifier que $\mathrm{Tr}\,C = \mathrm{Tr}\,D$.
   $$\mathrm{Tr}(P^{-1}\,A\,P) = \mathrm{Tr}\Big(P^{-1}\,(A\,P)\Big) = \mathrm{Tr}\Big((A\,P)\,P^{-1}\Big) = \mathrm{Tr}(\,A\,P\,P^{-1}) = \mathrm{Tr}\,A$$
3. Si cela était possible, en passant aux traces...
4. On trouve : $[D, D_1] = \mathrm{id}$.
   NOTA. En dimension infinie il existe donc des endomorphismes A, B tels que $[A, B] = \mathrm{id}$ (ce qui est exclu en dimension finie). L'existence de cette relation permet de justifier le principe d'indétermination d'Heisemberg, fondamentale en Mécanique Quantique. Ce qui explique aussi pourquoi la formalisation de la Mécanique Quantique exige l'utilisation des espaces vectoriels de dimension infinie.

31 Calculer $f(1)$, $f(x)$, $f(x^2)$ et imposer qu'ils appartiennent à $\mathbb{R}_2\,[x]$. On trouve que c'est le cas si et seulement si $a + 2\,b = 0$. On a alors :

$$M(f)_{\{1,x,x^2\}} = \begin{pmatrix} 1 & c & 0 \\ -2b & 1 & 0 \\ 0 & -b & 1 + 2\,c \end{pmatrix}$$

Si $c = -\dfrac{1}{2}$ ou $b\,c = -\dfrac{1}{2}$ on a $\mathrm{rg}\,A = 2$; sinon $\mathrm{rg}\,A = 3$.

32 On trouve :

$$\theta_1(x) = x_1\ ,\quad \theta_2(x) = -x_1 + x_2\ ,\quad \theta_3(x) = -x_2 + x_3\ ,\ \ldots\ldots,$$
$$\theta_{k+1}(x) = -x_k + x_{k+1}, \ldots\ldots, \theta_n(x) = -x_{n-1} + x_n$$

33 $\varphi_1 - 10\,\varphi_2 + 7\,\varphi_3 = 0$

34 Notons $\{\omega_1,\, \omega_2,\, \omega_3\}$ la base de $\mathbb{R}^{3^{**}}$ duale de $\{\varphi_1,\, \varphi_2,\, \varphi_3\}$: $\omega_i\,(\varphi_j) = \delta_{ij}$.
Comme $\mathbb{R}_3^{**} \simeq \mathbb{R}^3$, avec les notations de page 86 il existe $v_1, v_2, v_3 \in \mathbb{R}^3$ tels que $\Phi_{v_i} = \omega_i$; donc $\Phi_{v_i}(\varphi_j) = \delta_{ij}$, c'est-à-dire $\varphi_i(v_j) = \delta_{ij}$. En d'autres termes $\{\varphi_1, \varphi_2, \varphi_3\}$ est la base duale de $\{v_1, v_2, v_3\}$.

En imposant $\varphi_1(v_1) = 1$, $\varphi_2(v_1) = 0$, $\varphi_3(v_1) = 0$, on trouve : $v_1 = \begin{pmatrix} -18 \\ 7 \\ 5 \end{pmatrix}$.

De la même manière, on trouve : $v_2 = \begin{pmatrix} 5 \\ -2 \\ -1 \end{pmatrix}$ et $v_3 = \begin{pmatrix} 3 \\ -1 \\ -1 \end{pmatrix}$.

35

1. Vérifier que $\varphi \circ f$ est linéaire.
2. Simples vérifications, sauf pour (d), qui elle en revanche se démontre très facilement à l'aide de la question 4. qui suit (*ce qui montre que parfois le point de vue des applications est plus commode (comme pour (c)) et parfois il est plus commode d'utiliser le point de vue des matrices.*
3. Pour (e) , montrer d'abord que ${}^t(\mathrm{id}_E) = \mathrm{id}_{E^*}$; appliquer ensuite (c) à $f \circ f^{-1} = \mathrm{id}_F$ et à $f^{-1} \circ f = \mathrm{id}_E$.
4. Soit :
$$A = \begin{pmatrix} a_{11} \ldots a_{1n} \\ \ldots\ldots\ldots \\ a_{n1} \ldots a_{nn} \end{pmatrix} \qquad \text{et} \qquad B = M\,({}^t f)_{\psi_j, \varphi_i} = \begin{pmatrix} b_{11} \ldots b_{1n} \\ \ldots\ldots\ldots \\ b_{n1} \ldots b_{nn} \end{pmatrix}$$
c'est-à-dire :
$$f\,(e_k) = \sum_{i=1}^{p} a_{ki}\,\varepsilon_i \qquad \text{et} \qquad ({}^t f)\,(\psi_j) = \sum_{\alpha=1}^{n} b_{j\alpha}\,\varphi_\alpha$$
En déduire : $(\psi_j \circ f)\,(e_k) = a_{kj}$ et $({}^t f)\,(\psi_j)\,(e_k) = b_{jk}$, d'où : $b_{jk} = a_{kj}$ et donc $B = {}^t A$.

36

1. Soit $\varphi \in G^0$; $\forall\, y \in G$ on a $\varphi\,(y) = 0$; en particulier : $\varphi\,(x) = 0\,, \forall\, x \in F$.
2. Soit $\varphi \in (F+G)^0$; $\forall\, x \in F$, $\forall\, y \in G$: $\varphi\,(x+y) = 0$; en particulier, en prenant $x = 0_F$, on voit que $\varphi \in G^0$; en prenant $y = 0_G$: $\varphi \in F^0$.
Réciproque facile.
3. Montrer d'abord que $F \subset F^{00}$; puis que $\dim F = \dim F^{00}$.
4. Remplacer dans 2. F et G par F^0 et G^0 et prendre l'annulateur des deux membres.

37 $\varphi \in (\mathrm{Im}\, f)^0 \iff \varphi\,(f\,(x)\,) = 0\,, \forall\, x \in E \iff (\,{}^t f \circ \varphi)\,(x) = 0\,, \forall\, x \in E$
$\iff {}^t f \circ \varphi = 0 \iff \varphi \in \mathrm{Ker}(\,{}^t f\,)$.
Appliquer le théorème du rang à ${}^t f$ et la proposition 3.39 page 87 au sous-espace $\mathrm{Im}\, f$.

Chapitre 4

Déterminants

Comme nous l'avons déjà fait remarquer, la notion de dépendance ou indépendance linéaire intervient, d'une manière ou d'une autre, dans les différents problèmes d'algèbre linéaire.
Jusqu'à présent, la seule méthode que nous avons vue pour savoir si un système de vecteurs est libre, est celle qui consiste à étudier un système d'équations linéaires par élimination, ou encore – ce qui au fond revient au même – à échelonner une certaine matrice.
Le principal intérêt des déterminants est de fournir des conditions *explicites* d'indépendance linéaire, en quelque sorte des formules qui permettent de savoir si une famille de vecteurs est libre ou non. Appliquée aux systèmes d'équations linéaires, la notion de déterminant permet d'obtenir les conditions de compatibilité sans que l'on soit obligé de chercher explicitement la solution et elle donne aussi des formules pour la solution. L'intérêt de cela apparaît clairement lorsque l'on étudie les systèmes d'équations avec paramètre (comme nous le ferons au chapitre 6. pour le calcul des vecteurs propres) : la théorie des déterminants permet de donner immédiatement les conditions sur les paramètres pour que le problème admette une solution, alors qu'une étude directe par la méthode d'élimination conduirait à des calculs compliqués et surtout ne permettrait que très artificiellement d'avoir des résultats généraux.
La présentation que nous donnons ici n'est pas la plus élégante, mais elle permet d'effectuer immédiatement les calculs et de se familiariser rapidement avec la notion de déterminant et ses applications. Il est important, cependant, de comprendre que le déterminant est une *n-forme multilinéaire alternée* car l'on voit ainsi d'une façon très naturelle le lien entre la notion de déterminant et les problèmes d'indépendance linéaire.
Nous donnerons cette caractérisation au paragraphe 2 et 4 (cf. théorèmes 4.2 et 4.11) ; au paragraphe 10 nous mettrons en évidence l'interprétation géométrique par le volume. La notion d'orientation est présentée au paragraphe 11.
Comme toujours, le corps K est supposé commutatif.

NOTA - Les démonstrations en petits caractères peuvent être sautées en première lecture.

4.1 Définition des déterminants par récurrence

Définition 4.1 – *Soit* $A = (a_{ij}) \in \mathcal{M}_n(K)$. *On définit, par récurrence, une application :* dét : $\mathcal{M}_n(K) \longrightarrow K$ *de la manière suivante :*

- Si $n = 1$, *c'est-à-dire si* $A = (a)$, *on pose* $\det A = a$ *;*
- si $n > 1$, *notons* A_{ij} *la matrice obtenue de* A *en supprimant la* $i^{\text{ème}}$ *ligne et la* $j^{\text{ème}}$ *ligne colonne (c'est-à-dire la ligne et la colonne qui passent par l'élément*

a_{ij}) ; on pose alors $\big($ puisque $A_{ij} \in \mathcal{M}_{n-1}(K)\big)$:

$$\boxed{\text{dét}\,A = a_{11}\,\text{dét}\,A_{11} + \cdots + (-1)^{k+1}a_{1k}\,\text{dét}\,A_{1k} + \cdots + (-1)^{n+1}\,a_{1n}\,\text{dét}\,A_{1n}}$$

Le scalaire dét A *est dit **déterminant de** A et le déterminant de la matrice* $\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}$ *est noté, habituellement :* $\begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix}$.

Exemple 1 :

$$\begin{vmatrix} 4 & -1 \\ 3 & 2 \end{vmatrix} = 4 \times 2 - (-1) \times 3 = 11$$

Plus généralement :

$$\begin{vmatrix} a & c \\ b & d \end{vmatrix} = ad - bc$$

Exemple 2 :

$$\begin{vmatrix} 1 & -2 & 3 \\ 2 & 1 & -1 \\ 1 & 5 & 1 \end{vmatrix} = 1\begin{vmatrix} 1 & -2 & 3 \\ 2 & 1 & -1 \\ 1 & 5 & 1 \end{vmatrix} - (-2)\begin{vmatrix} 1 & -2 & 3 \\ 2 & 1 & -1 \\ & 5 & 1 \end{vmatrix} + 3\begin{vmatrix} 1 & -2 & 3 \\ 2 & 1 & -1 \\ 1 & 5 & 1 \end{vmatrix}$$

$$= 1\begin{vmatrix} 1 & -1 \\ 5 & 1 \end{vmatrix} - (-2)\begin{vmatrix} 2 & -1 \\ 1 & 1 \end{vmatrix} + 3\begin{vmatrix} 2 & 1 \\ 1 & 5 \end{vmatrix}$$

$$= (1+5) + 2(2+1) + 3(10-1) = 39$$

Plus généralement :

$$\begin{aligned} \begin{vmatrix} a & a' & a'' \\ b & b' & b'' \\ c & c' & c'' \end{vmatrix} &= a\begin{vmatrix} b' & b'' \\ c' & c'' \end{vmatrix} - a'\begin{vmatrix} b & b'' \\ c & c'' \end{vmatrix} + a''\begin{vmatrix} b & b' \\ c & c' \end{vmatrix} \\ &= a\,(b'\,c'' - b''\,c') - a'\,(b\,c'' - b''\,c) + a''\,(b\,c' - b'\,c) \\ &= a\,b'\,c'' + a'\,b''\,c + a''\,b\,c' - a''\,b'\,c - a'\,b\,c'' - a\,b''\,c' \end{aligned}$$

Règle de Sarrus

Ce résultat permet d'énoncer la règle de Sarrus pour le calcul d'un déterminant d'ordre 3 :

le déterminant d'une matrice d'ordre 3 est la somme de six termes, trois affectés du signe + et trois du signe − :

- les produits affectés du signe + contiennent soit les trois termes de la diagonale principale, soit deux termes parallèles à cette diagonale ;
- pour les produits affectés du signe −, on procède de même en changeant de diagonale.

Plus précisément, voici le schéma de cette règle (dont le seul intérêt est d'effectuer plus vite les calculs) :

$$\underset{ab'c''}{\begin{vmatrix} \bullet & \cdot & \cdot \\ \cdot & \bullet & \cdot \\ \cdot & \cdot & \bullet \end{vmatrix}} + \underset{a'b''c}{\begin{vmatrix} \cdot & \bullet & \cdot \\ \cdot & \cdot & \bullet \\ \bullet & \cdot & \cdot \end{vmatrix}} + \underset{a''bc'}{\begin{vmatrix} \cdot & \cdot & \bullet \\ \bullet & \cdot & \cdot \\ \cdot & \bullet & \cdot \end{vmatrix}} - \underset{a''b'c}{\begin{vmatrix} \cdot & \cdot & \bullet \\ \cdot & \bullet & \cdot \\ \bullet & \cdot & \cdot \end{vmatrix}} - \underset{a'bc''}{\begin{vmatrix} \cdot & \bullet & \cdot \\ \bullet & \cdot & \cdot \\ \cdot & \cdot & \bullet \end{vmatrix}} - \underset{ab''c'}{\begin{vmatrix} \bullet & \cdot & \cdot \\ \cdot & \cdot & \bullet \\ \cdot & \bullet & \cdot \end{vmatrix}}$$

(Bien entendu, cette règle ne s'applique qu'aux matrices d'ordre 3).

Exemple 3 :

$$\begin{vmatrix} a_{11} & 0 & \cdots & 0 \\ a_{21} & a_{22} & 0 & \vdots \\ \vdots & * & \ddots & 0 \\ a_{n1} & \cdots & \cdots & a_{nn} \end{vmatrix} = a_{11} \begin{vmatrix} a_{22} & 0 & \cdots & 0 \\ a_{32} & a_{33} & 0 & \vdots \\ \vdots & * & \ddots & 0 \\ a_{n2} & \cdots & \cdots & a_{nn} \end{vmatrix}$$

$$= a_{11}\, a_{22} \begin{vmatrix} a_{33} & 0 & \cdots & 0 \\ \vdots & \ddots & 0 & \vdots \\ \vdots & * & \ddots & 0 \\ a_{n3} & \cdots & \cdots & a_{nn} \end{vmatrix} = \cdots = a_{11}\, a_{22}\, a_{33} \cdots a_{nn}$$

Le déterminant d'une matrice triangulaire (relativement à la diagonale principale) est donc égal au produit des termes de la diagonale.

Exercices 1. 2.

4.2 Les déterminants vus comme formes multilinéaires alternées

La propriété fondamentale des déterminants est exprimée par le théorème suivant :

Théorème 4.2 –

1. *Le déterminant est une application* ***linéaire par rapport à chaque colonne****, c'est-à-dire : si* $A = \| c_1, \dots, c_n \|$ *alors,* $\forall\, \lambda \in K$, $\forall\, k = 1, \dots, n$ *:*
   - *(a)* $\text{dét} \| c_1, \dots, \lambda\, c_k, \dots, c_n \| = \lambda\ \text{dét} \| c_1, \dots, c_k, \dots, c_n \|$
   - *(b)* $\text{dét} \| c_1, \dots, a_k + b_k, \dots, c_n \| =$
     $$\text{dét} \| c_1, \dots, a_k, \dots, c_n \| + \text{dét} \| c_1, \dots, b_k, \dots, c_n \|$$
2. *Si deux colonnes sont égales, le déterminant est nul* [1].

Nota. – Nous verrons une réciproque de cette propriété (cf. proposition 4.11) en ce sens que toute application *multilinéaire et alternée*, c'est-à-dire qui vérifie les propriétés 1. et 2. ci-dessus, est, dans une certaine base, le déterminant d'une matrice.

Démonstration : Par récurrence sur n. [2]

Pour $n = 1$ il n'y a rien à démontrer.
Pour $n = 2$ c'est une simple vérification. On a :

$$\begin{vmatrix} \lambda a & c \\ \lambda b & d \end{vmatrix} = \lambda\, a\, d - \lambda\, b\, c = \lambda\, (a\, d - b\, c) = \lambda \begin{vmatrix} a & c \\ b & d \end{vmatrix}$$

$$\begin{vmatrix} a + a' & c \\ b + b' & d \end{vmatrix} = (a + a')\, d - (b + b')\, c = (a\, d - b\, c) + (a'\, d - b'\, c)$$

$$= \begin{vmatrix} a & c \\ b & d \end{vmatrix} + \begin{vmatrix} a' & c \\ b' & d \end{vmatrix}$$

[1] Pour une raison que l'on verra par la suite (cf. corollaire 4.3 page 107), cette propriété s'exprime en disant que le déterminant est une application alternée.

[2] Cette démonstration est tirée de S. Lang : *Algèbre linéaire 1*, InterEditions, Paris, 1976.

ce qui montre la linéarité par rapport à la première colonne. De la même manière, on vérifie la linéarité par rapport à la seconde colonne.

Enfin $\begin{vmatrix} a & a \\ b & b \end{vmatrix} = a\,b - a\,b = 0$ et la propriété 2. est vérifiée.

Supposons le théorème vrai jusqu'à l'ordre $n-1$ et soit $A \in \mathcal{M}_n(K)$.

Montrons d'abord la linéarité de $\text{dét}\, A$ par rapport à la $k^{\text{ème}}$ colonne où $k \in \{1, \dots, n\}$.

D'après la définition, $\text{dét}\, A$ est une somme de termes du type :

$$(-1)^{1+j}\, a_{1j}\ \text{dét}\, A_{1j}$$

Il suffira de vérifier que chacun de ces termes dépend linéairement de la $k^{\text{ème}}$ colonne.

Si $j \neq k$, on a : $A_{1j} = \begin{pmatrix} a_{11} & \cdots & a_{1j} & \cdots & a_{1k} & \cdots & a_{1n} \\ \vdots & & \vdots & & \vdots & & \vdots \\ a_{n1} & \cdots & a_{nj} & \cdots & a_{nk} & \cdots & a_{nn} \end{pmatrix}$

Dans ce cas, le terme a_{1j} ne dépend pas de la $k^{\text{ème}}$ colonne, alors que la matrice A_{1j} contient la $k^{\text{ème}}$ colonne. D'après l'hypothèse de récurrence, $\text{dét}\, A_{1j}$ est une fonction linéaire de la $k^{\text{ème}}$ colonne. Ainsi, $a_{1j}\, \text{dét}\, A_{1j}$ dépend linéairement de la $k^{\text{ème}}$ colonne.

Si $j = k$, $a_{1j} = a_{1k}$ dépend linéairement de la $k^{\text{ème}}$ colonne, alors que A_{1k} ne contient pas la $k^{\text{ème}}$ colonne. Aussi, $a_{1k}\, \text{dét}\, A_{1k}$ dépend linéairement de la $k^{\text{ème}}$ colonne.

Ainsi dans tous les cas, $\text{dét}\, A$ dépend linéairement de la $k^{\text{ème}}$ colonne.

Il reste à montrer que si deux colonnes sont égales, le déterminant est nul.

Lemme (a) :

Si deux colonnes adjacentes sont égales, le déterminant est nul.

En effet, supposons par exemple que $c_k = c_{k+1}$ et soit j un indice différent de k et de $k+1$.

On a : $A_{1j} = \begin{pmatrix} a_{11} & \cdots & a_{1j} & \cdots & a_{1k} & a_{1k} & \cdots & a_{1n} \\ \vdots & & \vdots & & \vdots & \vdots & & \vdots \\ a_{n1} & \cdots & a_{nj} & \cdots & a_{nk} & a_{nk} & \cdots & a_{nn} \end{pmatrix}$

Cette matrice d'ordre $n-1$ a deux colonnes égales, donc, d'après l'hypothèse de récurrence, son déterminant est nul. Aussi, dans le développement de $\text{dét}\, A$ tous les termes du type $a_{1j}\ \text{dét}\, A_{1j}$ avec $j \neq k$ et $j \neq k+1$ sont nuls. Il reste :

$$\text{dét}\, A = (-1)^{1+k}\, a_{1k}\ \text{dét}\, A_{1k} \;+\; (-1)^{1+k+1}\, a_{1,k+1}\ \text{dét}\, A_{1,k+1}$$

Or $c_k = c_{k+1}$, donc $a_{1k} = a_{1,k+1}$ et $A_{1k} = A_{1,k+1}$; les deux termes se compensent et par conséquent $\text{dét}\, A = 0$. $\diamondsuit$

Lemme (b) :

Si dans un déterminant on échange entre elles deux colonnes adjacentes, le déterminant change de signe.

En effet, remplaçons dans la matrice A la $j^{\text{ème}}$ et la $(j+1)^{\text{ème}}$ colonne par leur somme $c_j + c_{j+1}$. D'après le lemme (a), on aura une matrice à déterminant nul :

$$\text{dét}\, \| c_1, \dots, c_j + c_{j+1}, c_j + c_{j+1}, \dots, c_n \| = 0$$

En utilisant la linéarité par rapport à chaque colonne que nous avons démontrée ci-dessus, on a :

$$\begin{aligned} 0 = \;& \text{dét}\, \| c_1, \dots, c_j, \;\; c_j, \dots, c_n \| \;+\; \text{dét}\, \| c_1, \dots, c_j, \; c_{j+1}, \dots, c_n \| \\ &+ \text{dét}\, \| c_1, \dots, c_{j+1}, \; c_j, \dots, c_n \| \;+\; \text{dét}\, \| c_1, \dots, c_{j+1}, \; c_{j+1}, \dots, c_n \| \end{aligned}$$

En utilisant le lemme (a), on voit que le premier et le quatrième termes sont nuls et donc que :
$$\text{dét}\,\|c_1,\ldots,c_j,\ c_{j+1},\ldots,c_n\| \;=\; -\,\text{dét}\,\|c_1,\ldots,c_{j+1},\ c_j,\ldots,c_n\|$$
ce qui montre le lemme . $\diamondsuit$

Revenons maintenant au théorème et supposons que deux colonnes quelconques soient égales, par exemple $c_i = c_k$. Par échanges successifs de colonnes adjacentes, on peut modifier la matrice jusqu'à ce que les deux colonnes égales soient adjacentes. Chaque fois que l'on fait un tel échange, le déterminant change de signe d'après le lemme (b) ce qui n'a pas d'incidence sur le fait qu'il soit nul ou non. Puisque, lorsque les colonnes égales sont adjacentes le déterminant est nul d'après le lemme (a), il le restera lorsque c_i et c_k sont deux colonnes quelconques égales. $\square$

Le théorème que nous venons de démontrer entraîne toute une série de propriétés importantes.

Corollaire 4.3 – *Si l'on échange entre elles deux colonnes, le déterminant change de signe.*

La démonstration est analogue à celle du lemme (b). Remplaçons dans la matrice A les colonnes c_i et c_k par leurs sommes. On a :

$$\begin{aligned} 0 = \quad & \text{dét}\,\|\,c_1,\ldots,c_i+c_k,\ldots,c_i+c_k,\ldots,c_n\,\| \;=\; \text{dét}\,\|\,c_1,\ldots,c_i,\ldots,c_i,\ldots,c_n\,\| \\ & + \text{dét}\,\|\,c_1,\ldots,c_i,\ldots,c_k,\ldots,c_n\,\| + \text{dét}\,\|\,c_1,\ldots,c_k,\ldots,c_i,\ldots,c_n\,\| \\ & + \text{dét}\,\|\,c_1,\ldots,c_k,\ldots,c_k,\ldots,c_n\,\| \end{aligned}$$

Le premier et le dernier termes étant nuls, d'après la propriété 2. du théorème 4.2, on obtient facilement le résultat.

Permutations des colonnes

Soit $A = \|c_1,\ldots,c_n\|$ et A' une matrice obtenue en changeant l'ordre des vecteurs colonnes de A (on dit que A' est obtenue de A par *permutation* des vecteurs colonnes).
Par exemple $A = \|c_1,c_2,c_3,c_4\|$ et $A' = \|c_4,c_3,c_1,c_2\|$.
Il est facile de se rendre compte qu'en échangeant entre elles deux colonnes on peut, par de telles opérations successives, passer de A' à A.
Il y a même une façon standard de le faire qui consiste à mettre d'abord la colonne c_1 à la première place, puis c_2 à la seconde, etc.
Par exemple :

$$A' = \|c_4,c_3,c_1,c_2\| \;\underset{\substack{\text{en échangeant} \\ c_1 \text{ et } c_4}}{\longrightarrow}\; \|c_1,c_3,c_4,c_2\| \;\underset{\substack{\text{en échangeant} \\ c_2 \text{ et } c_3}}{\longrightarrow}\; \|c_1,c_2,c_4,c_3\|$$
$$\underset{\substack{\text{en échangeant} \\ c_3 \text{ et } c_4}}{\longrightarrow}\; \|c_1,c_2,c_3,c_4\| = A$$

L'opération qui consiste à échanger entre elles deux colonnes en laissant fixes les autres est dite ***transposition des colonnes***. A chaque transposition le déterminant change de signe et donc, $\text{dét}\,A' = (-1)^\tau\,\text{dét}\,A$ où τ est le nombre de transpositions nécessaires pour passer de A à A'.[3] En particulier, si $\text{dét}\,A = 0$ alors $\text{dét}\,A' = 0$.
Dans l'exemple précédent : $\tau = 3$ donc $\text{dét}\,A' = -\,\text{dét}\,A$.

[3] Pour une présentation plus détaillée de ces questions, cf. paragraphe 3.

Exemple – Calculer :

$$\Delta = \begin{vmatrix} 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{vmatrix}$$

$$\Delta \;=\; \text{dét}\,\|e_2, e_4, e_5, e_1, e_3\| = -\,\text{dét}\,\|e_1, e_4, e_5, e_2, e_3\| = \text{dét}\,\|e_1, e_2, e_5, e_4, e_3\|$$

$$= \; -\,\text{dét}\,\|e_1, e_2, e_3, e_4, e_5\| = - \begin{vmatrix} 1 & 0 & \cdots & & 0 \\ 0 & \ddots & & & \vdots \\ \vdots & & \ddots & & 0 \\ & & & \ddots & \\ 0 & \cdots & & 0 & 1 \end{vmatrix} = -\,1$$

Ces remarques permettent de démontrer le théorème principal, celui qui motive l'introduction de la notion de déterminant.

Théorème 4.4 – *Soit* $A = \|c_1, \ldots, c_n\| \in \mathcal{M}_n(K)$. *Alors les vecteurs* $c_1, \ldots, c_n$ *forment une base de* K^n *si et seulement si* $\text{dét}\,A \neq 0$.

Ou, d'une manière équivalente, (en prenant la contraposée) :

Un déterminant est nul si et seulement si l'une des colonnes est combinaison linéaire des autres colonnes.

Démonstration : Supposons, quitte à changer l'ordre des colonnes (ce qui n'a pas d'incidence sur le fait que le déterminant soit nul ou non) que la première colonne soit une combinaison linéaire des autres :

$$A = \| \sum_{i=2}^{n} \lambda_i \, c_i, c_2, \ldots, c_n \, \|$$

En utilisant la linéarité, on a :

$$\text{dét}\,A = \sum_{i=2}^{n} \lambda_i \;\text{dét}\,\| \, c_i, c_2, \ldots, c_n \, \|$$

Or, chaque terme de la somme est un déterminant avec deux colonnes égales, donc nul, et par conséquent $\text{dét}\,A = 0$.

Réciproquement, soit $\{c_1, \ldots, c_n\}$ une base de K^n et montrons que si on avait $\text{dét}\,\| \, c_1, \ldots, c_n \, \| = 0$ l'on aurait une contradiction.

Puisque $\{c_1, \ldots, c_n\}$ est une base, tout vecteur v de K^n s'écrit comme combinaison linéaire de $c_1, \ldots, c_n$. Soient :

$$v_1 = \sum_{j=1}^{n} a_j \, c_j \quad , \quad v_2 = \sum_{k=1}^{n} b_k \, c_k \quad , \ldots , \quad v_n = \sum_{l=1}^{n} g_l \, c_l$$

n vecteurs quelconques de K^n. On a, en utilisant la linéarité par rapport à chaque colonne :

$$\text{dét}\,\|v_1, \ldots, v_n\| = \text{dét}\,\| \sum_{j=1}^{n} a_j \, c_j, \; \sum_{k=1}^{n} b_k \, c_k, \ldots, \sum_{l=1}^{n} g_l \, c_l \|$$

$$= \sum_{j=1}^{n} \sum_{k=1}^{n} \dots \sum_{l=1}^{n} a_j\, b_k \dots g_l \text{ dét} \|c_j, c_k, \dots, c_l\|$$

Or, chaque terme de cette somme est du type $\alpha \text{ dét} \|c_j, c_k, \dots, c_l\|$, où α est un scalaire et les indices $j, k, \dots, l$ sont l'un quelconque des nombres $1, 2, \dots, n$. Si deux indices sont égaux, $\|c_j, c_k, \dots, c_l\| = 0$; si tous les indices sont différents, $\|c_j, c_k, \dots, c_l\|$ est obtenue par permutation des colonnes de la matrice $\|c_1, \dots, c_n\|$. Puisque le déterminant de celles-ci a été supposé nul, le terme $\alpha \text{ dét} \|c_j, c_k, \dots, c_l\|$ est nul et, finalement, $\text{dét} \|v_1, \dots, v_n\| = 0$.
Mais ceci est impossible, car $v_1, \dots, v_n$ sont des vecteurs arbitraires (il suffit de prendre pour $\{v_1, \dots, v_n\}$ la base canonique $\{e_1, \dots, e_n\}$ de K^n : $\text{dét} \|e_1, \dots, e_n\| = \text{dét}\, I_n = 1$). $\square$

Exercices 3. 4. 5.

4.3 Permutations, transpositions, signature

Dans ce paragraphe, nous allons définir d'une manière plus précise certaines notions que nous avons introduites et utilisées au paragraphe précédent.
Comme nous l'avons vu, un rôle important dans la théorie des déterminants est joué par les permutations des colonnes. Voici la définition précise de la notion de permutation.

Définition 4.5 – *On appelle* ***permutation*** *de* $\{1, 2, \dots, n\}$, *toute application bijective :*

$$\sigma : \{1, 2, \dots, n\} \longrightarrow \{1, 2, \dots, n\}$$

L'ensemble des permutations de $\{1, 2, \dots, n\}$ *est noté* $\mathcal{S}_n$.

Puisque σ est bijective, l'ensemble des valeurs de σ, $\{\sigma(1), \sigma(2), \dots, \sigma(n)\}$, possède n éléments distincts et donc est composé des entiers $1, 2, \dots, n$ rangés dans un ordre éventuellement différent.
Par exemple, l'application $\sigma : \{1,\ 2,\ 3,\ 4\} \longrightarrow \{1,\ 2,\ 3,\ 4\}$ telle que $\sigma(1) = 4$, $\sigma(2) = 1$, $\sigma(3) = 3$, $\sigma(4) = 2$ est une permutation : $\sigma \in \mathcal{S}_4$.
Comme il est bien connu, $\mathcal{S}_n$ contient $n\,!$ éléments ($n\,! = 1 \cdot 2 \cdot 3 \cdot \ldots \cdot n$).
La notation habituelle pour une permutation consiste à écrire dans une première rangée les entiers $1, 2, \dots, n$ et au-dessous leurs images par σ :

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

Ainsi, la permutation de l'exemple ci-dessus s'écrit :

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

Si σ_1 et σ_2 sont deux permutations dans $\mathcal{S}_n$, on peut former la permutation produit $\sigma_1 \circ \sigma_2$.
Par exemple, si :

$$\sigma_1 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 1 & 3 & 2 \end{pmatrix} \qquad \text{et} \qquad \sigma_2 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 1 & 2 & 4 \end{pmatrix}$$

on a :

$$\sigma_1 \circ \sigma_2\,(1) = \sigma_1(3) = 3\ , \qquad \sigma_1 \circ \sigma_2\,(2) = \sigma_1(1) = 4$$
$$\sigma_1 \circ \sigma_2\,(3) = \sigma_1(2) = 1\ , \qquad \sigma_1 \circ \sigma_2\,(4) = \sigma_1(4) = 2$$

donc :

$$\sigma_1 \circ \sigma_2 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 4 & 1 & 2 \end{pmatrix}$$

On écrira aussi, d'une manière naturelle malgré l'abus de notation :

$$\sigma_1 \circ \sigma_2 = \begin{array}{c} \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 1 & 2 & 4 \\ 3 & 4 & 1 & 2 \end{pmatrix} \end{array} \begin{array}{l} \\ {\scriptstyle \sigma_2} \\ {\scriptstyle \sigma_1 \circ \sigma_2} \end{array}$$

De même, puisque σ est bijective, on peut considérer la permutation inverse σ^{-1}. Dans l'exemple précédent :

$$\sigma_1^{-1} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 3 & 1 \end{pmatrix}$$

Définition 4.6 – *On appelle* ***transposition*** *une permutation qui échange entre eux deux éléments différents et laisse fixes les autres. Les transpositions seront notées $\tau, t, \ldots$*

Par exemple :

$$\tau = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 4 & 3 & 2 & 5 \end{pmatrix}$$

est la transposition de $\mathcal{S}_5$ qui échange 2 et 4.

Remarquons que si τ est une transposition, $\tau^2 = \mathrm{id}$ et donc $\tau^{-1} = \tau$.

Comme nous l'avons vu au paragraphe précédent, on peut obtenir une permutation en effectuant plusieurs transpositions successives. Par exemple :

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 1 & 3 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 2 & 3 & 1 \\ 4 & 1 & 3 & 2 \end{pmatrix} \begin{array}{l} \\ \text{par } \tau_1 \text{ qui échange 1 et 4} \\ \text{par } \tau_2 \text{ qui échange 1 et 2} \end{array}$$

donc $\sigma = \tau_2 \circ \tau_1$.

Proposition 4.7 – *Toute permutation peut se décomposer en produit de transpositions.*

Démonstration : Soit $\sigma \in \mathcal{S}_n$. Raisonnons par récurrence sur n.
Pour $n = 1$ il n'y a rien à démontrer.
Supposons la propriété vraie jusqu'à l'ordre $n-1$ et considérons une permutation $\sigma \in \mathcal{S}_n$.
Soit $k = \sigma(n)$, c'est-à-dire :

$$\sigma = \begin{pmatrix} 1 & 2 & \cdots & n \\ * & * & \cdots & k \end{pmatrix}$$

Deux possibilités peuvent se présenter :

1. $k = n$. Ceci veut dire que l'élément n est fixe et que donc σ est une permutation de $\{1, 2, \ldots, n-1\}$. D'après l'hypothèse de récurrence σ est un produit de transpositions.
2. $k \neq n$. Considérons alors la transposition τ qui échange k et n et soit $\sigma_1 = \tau \circ \sigma$. On a :
$$\sigma_1\,(n) = \tau\sigma(n) = \tau(k) = n$$
Donc σ_1 laisse fixe n, et d'après l'analyse du cas précédent, elle se décompose en produit de transpositions : $\sigma_1 = \tau_1 \circ \cdots \circ \tau_p$.
On aura donc : $\tau \circ \sigma = \tau_1 \circ \cdots \circ \tau_p$, d'où, en composant avec τ à gauche :
$$\sigma = \tau \circ \tau_1 \circ \cdots \circ \tau_p$$
car $\tau^2 = \mathrm{id}$. Donc σ est un produit de transpositions. $\square$

Remarquons que la façon de décomposer σ en produit de transpositions n'est pas unique.

Par exemple, si :

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 3 & 5 & 1 & 2 \end{pmatrix}$$

on a :

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 2 & 3 & 1 & 5 \\ 4 & 5 & 3 & 1 & 2 \\ 4 & 3 & 5 & 1 & 2 \end{pmatrix} \quad \text{et aussi} \quad \sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 1 & 3 & 4 & 5 \\ 4 & 1 & 3 & 2 & 5 \\ 4 & 3 & 1 & 2 & 5 \\ 4 & 3 & 5 & 2 & 1 \\ 4 & 3 & 5 & 1 & 2 \end{pmatrix}$$

Dans le premier cas, σ est produit de 3 transpositions, dans le second cas elle est le produit de 5 transpositions.

Cependant, ce qu'il y a de remarquable, et important pour la suite de la théorie, est que la parité du nombre de transpositions (c'est-à-dire le fait que ce nombre soit pair ou impair) est indépendante de la décomposition. On a :

Théorème 4.8 – *Soit σ une permutation. S'il existe une décomposition de σ en un nombre pair (respect. : impair) de transpositions, alors toute autre décomposition de σ comportera un nombre pair (respect. : impair) de transpositions. Par conséquent, le nombre :*

$$\varepsilon(\sigma) := (-1)^p$$

(où p est le nombre de transpositions dans une décomposition de σ) est indépendant de la décomposition. Ce nombre est appelé ***signature*** *de σ.*

Démonstration : Il existe une démonstration directe de ce résultat (par des raisonnements de combinatoire). Cependant, on peut le démontrer facilement en utilisant les résultats du paragraphe 2. [4]
Le corollaire 4.3 dit en effet que si l'on échange entre elles deux colonnes le déterminant change de signe. Avec les définitions de ce paragraphe, cela peut s'énoncer de la manière suivante :

Soit τ une transposition de $\{1, 2, \dots, n\}$. Alors :

$$\text{dét}\, \|v_{\tau(1)}, \dots, v_{\tau(n)}\| = -\,\text{dét}\, \|v_1, \dots, v_n\|$$

Soit maintenant σ une permutation et considérons deux décompositions de σ en produit de transpositions :

$$\begin{aligned} \sigma &= \tau_1 \circ \cdots \circ \tau_p \\ \sigma &= t_1 \circ \cdots \circ t_q \end{aligned}$$

Il s'agit de montrer que : $(-1)^p = (-1)^q$.
Soit $\{e_1, \dots, e_n\}$ la base canonique de K^n et $\Delta = \text{dét}\, \|e_{\sigma(1)}, \dots, e_{\sigma(n)}\|$. On a :

$$\begin{aligned} \Delta \quad &= \text{dét}\, \|e_{\tau_1\circ\tau_2\circ\cdots\circ\tau_p(1)}, \dots, e_{\tau_1\circ\tau_2\circ\cdots\circ\tau_p(n)}\| = -\text{dét}\, \|e_{\tau_2\circ\cdots\circ\tau_p(1)}, \dots, e_{\tau_2\circ\cdots\circ\tau_p(n)}\| \\ &= +\text{dét}\, \|e_{\tau_3\circ\cdots\circ\tau_p(1)}, \dots, e_{\tau_3\circ\cdots\circ\tau_p(n)}\| = \cdots = (-1)^p \text{dét}\, \|e_1, \dots, e_n\| = (-1)^p \end{aligned}$$

De même, en utilisant la deuxième décomposition, $\sigma = t_1 \circ \cdots \circ t_q$, on trouve $\Delta = (-1)^q$, d'où : $(-1)^p = (-1)^q$. $\square$

[4] Au paragraphe 2, nous n'avons admis que la proposition 4.7 que nous venons de démontrer.

On a immédiatement les propriétés suivantes :

Corollaire 4.9 –

1. *Si τ est une transposition,* $\varepsilon(\tau) = -1$.

2. $\varepsilon(\sigma_1 \circ \sigma_2) = \varepsilon(\sigma_1)\, \varepsilon(\sigma_2), \ \forall\, \sigma_1, \sigma_2 \in \mathcal{S}_n$

3. $\varepsilon(\sigma^{-1}) = \varepsilon(\sigma), \ \forall\, \sigma \in \mathcal{S}_n$

4. $\text{dét}\, \|v_{\sigma(1)}, \dots, v_{\sigma(n)}\| = \varepsilon(\sigma)\ \text{dét}\, \|v_1, \dots, v_n\|$.

Vérifions par exemple 2. Soient $\sigma_1 = \tau_1 \circ \cdots \circ \tau_p$ et $\sigma_2 = t_1 \circ \cdots \circ t_q$. On a : $\sigma_1 \circ \sigma_2 = \tau_1 \circ \cdots \circ \tau_p \circ t_1 \circ \cdots \circ t_q$, donc :

$$\varepsilon(\sigma_1 \circ \sigma_2) = (-1)^{p+q} = (-1)^p\, (-1)^q = \varepsilon(\sigma_1)\, \varepsilon(\sigma_2)$$

Les autres propriétés se démontrent tout aussi facilement.

Exercices 6. 7.

4.4 Une formule explicite pour le déterminant

A l'aide des résultats du paragraphe 3, on peut donner une formule pour calculer les déterminants, dont l'intérêt est surtout théorique : elle va nous permettre de démontrer certaines propriétés importantes des déterminants.

Commençons par le cas $n = 2$.

Soit $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = \|c_1, c_2\|$, où c_1, c_2 sont les vecteurs colonne de A. Si $\{e_1, e_2\}$ est la base canonique de K^2 :

$$\begin{aligned} c_1 &= a_{11}\, e_1 + a_{21}\, e_2 \\ c_2 &= a_{12}\, e_1 + a_{22}\, e_2 \end{aligned}$$

En utilisant les propriétés fondamentales vues au paragraphe 2., on a :

$$\begin{aligned} \text{dét}\ A &= \text{dét}\, \|a_{11}\, e_1 + a_{21}\, e_2,\ a_{12}\, e_1 + a_{22}\, e_2\| \\ &= \underbrace{a_{11}\, a_{12} \text{dét}\, \|e_1, e_1\|}_{=0} + a_{11}\, a_{22} \text{dét}\, \|e_1, e_2\| + a_{21}\, a_{12} \text{dét}\, \|e_2, e_1\| \\ &+ \underbrace{a_{21}\, a_{22} \text{dét}\, \|e_2, e_2\|}_{=0} = (a_{11}\, a_{22} - a_{21}\, a_{12}) \text{dét}\, \|e_1, e_2\| \\ &= a_{11}\, a_{22} - a_{21}\, a_{12} \end{aligned}$$

(on retrouve la formule vue au paragraphe 1).

Considérons maintenant le cas $n = 3$ et soit $A = \|c_1, c_2, c_3\|$, avec :

$$\begin{aligned} c_1 &= a_{11}\, e_1 + a_{21}\, e_2 + a_{31}\, e_3 = \textstyle\sum_{i=1}^3 a_{i1}\, e_i \\ c_2 &= a_{12}\, e_1 + a_{22}\, e_2 + a_{32}\, e_3 = \textstyle\sum_{j=1}^3 a_{j2}\, e_j \\ c_3 &= a_{13}\, e_1 + a_{23}\, e_2 + a_{33}\, e_3 = \textstyle\sum_{k=1}^3 a_{k3}\, e_k \end{aligned}$$

On a :

$$\text{dét}\ A = \text{dét}\, \|\sum_{i=1}^{3} a_{i1}\, e_i,\ \sum_{j=1}^{3} a_{j2}\, e_j,\ \sum_{k=1}^{3} a_{k3}\, e_k\| = \sum_{i,j,k=1}^{3} a_{i1}\, a_{j2}\, a_{k3}\ \text{dét}\, \|e_i, e_j, e_k\|$$

Lorsque deux des indices i, j, k sont égaux, le terme correspondant dans la somme est nul (car le déterminant qui y intervient a deux colonnes égales). Il reste les termes correspondant aux indices $\{i, j, k\}$ avec i, j, k variant entre 1 et 3 et différents entre eux, c'est-à-dire correspondant à toutes les permutations de $\{1, 2, 3\}$. Par exemple $\{i, j, k\}$ égal à $\{3, 1, 2\}$ ou à $\{1, 3, 2\}$, etc.

En changeant de notation et en posant :

$$i = \sigma(1) \quad , \quad j = \sigma(2) \quad , \quad k = \sigma(3) \qquad \text{avec } \sigma \in \mathcal{S}_3$$

on a :

$$\begin{aligned}
\text{dét}\, A &= \sum_{\sigma\in\mathcal{S}_3} a_{\sigma(1)1}\, a_{\sigma(2)2}\, a_{\sigma(3)3}\ \text{dét}\, \|e_{\sigma(1)},\, e_{\sigma(2)},\, e_{\sigma(3)}\| \\
&= \sum_{\sigma\in\mathcal{S}_3} \varepsilon(\sigma)\, a_{\sigma(1)1}\, a_{\sigma(2)2}\, a_{\sigma(3)3}\ \underbrace{\text{dét}\, \|e_1,\, e_2,\, e_3\,\|}_{=1} \\
&= \sum_{\sigma\in\mathcal{S}_3} \varepsilon(\sigma)\, a_{\sigma(1)1}\, a_{\sigma(2)2}\, a_{\sigma(3)3}
\end{aligned}$$

Plus généralement, on démontre exactement de la même façon :

Proposition 4.10 – *Si $A = (a_{ij})$, on a :*

$$\boxed{\text{dét}\, A = \sum_{\sigma\in\mathcal{S}_n} \varepsilon(\sigma)\, a_{\sigma(1)1}\ \ a_{\sigma(2)2}\ \cdots\ a_{\sigma(n)n}}$$

Exemple – Explicitons le résultat pour $n = 3$. Cherchons d'abord toutes les permutations de $\{1, 2, 3\}$. On a :

$$\mathcal{S}_3 = \{\sigma_1,\, \sigma_2,\, \sigma_3,\, \tau_1,\, \tau_2,\, \tau_3\}$$

où :

$$\sigma_1 = \text{id} \ = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix} \quad , \quad \sigma_2 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} \quad , \quad \sigma_3 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}$$

$$\tau_1 \ = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} \quad , \quad \tau_2 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} \quad , \quad \tau_3 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}$$

On vérifie facilement que $\varepsilon(\sigma_i) = 1$ et $\varepsilon(\tau_i) = -1 \quad (i = 1, 2, 3)$, d'où :

$$\begin{aligned}
\text{dét}\, A = \quad & a_{11}\, a_{22}\, a_{33} + a_{21}\, a_{32}\, a_{13} + a_{31}\, a_{12}\, a_{23} \\
& - a_{11}\, a_{32}\, a_{23} - a_{31}\, a_{22}\, a_{13} - a_{21}\, a_{12}\, a_{33}
\end{aligned}$$

(on retrouve ainsi la formule de Sarrus).

La formule ci-dessus nous permet de montrer que toute application qui vérifie les propriétés 1. et 2. du théorème 4.2 est en fait le déterminant d'une matrice dans une base opportune. Commençons par le cas $n = 2$ et considérons une application $f : \mathbb{R}^2 \times \mathbb{R}^2 \longrightarrow \mathbb{R}$

- *bilinéaire* (c'est-à-dire linéaire en chaque facteur)
- et *alternée*, c'est-à-dire telle que $f(v,\, v) = 0\ \forall\, v \in \mathbb{R}^2$
  (ce qui équivaut à $f(v,\, w) = -\, f(w,\, v), \quad \forall\, v, w \in \mathbb{R}^2$) (cf. corollaire 4.3 page 107).

Soit $\{e_1, e_2\}$ la base canonique de $\mathbb{R}^2$ et $v = a\, e_1 + b\, e_2,\ w = c\, e_1 + d\, e_2,$ deux vecteurs quelconques de $\mathbb{R}^2$. On a :

$$\begin{aligned}
f(v,\, w) &= f(a\, e_1 + b\, e_2,\ c\, e_1 + d\, e_2) \\
&= a\, c\, f(e_1,\, e_1) + a\, d\, f(e_1,\, e_2) + b\, c\, f(e_2,\, e_1) + b\, d\, f(e_2,\, e_2) \\
&= (a\, d\ -\ b\, c)\ f(e_1,\, e_2).
\end{aligned}$$

c'est-à-dire

$$f(v, w) = \text{dét}\, \|\, v, w\, \|_{e_i}\, f(e_1, e_2)$$

Ceci montre que si $f(e_1, e_2) = 0$ alors f est identiquement nulle, ou, en prenant la contraposée, que si f n'est pas l'application nulle, $f(e_1, e_2) \neq 0$. Si donc $f \neq 0$, en remplaçant f par l'application $\tilde{f} = \dfrac{f}{f(e_1, e_2)}$, on a :

$$\text{dét}\, \|v, w\|_{e_i} = \tilde{f}(v, w)$$

On peut énoncer ce résultat aussi de la manière suivante :

Soit $f : \mathbb{R}^2 \times \mathbb{R}^2 \longrightarrow \mathbb{R}$ *une application bilinéaire alternée telle que* $f(e_1, e_2) = 1$, $\{e_1, e_2\}$ *étant la base canonique. Alors*

$$f(v, w) = \text{dét}\, A \ , \quad \textit{où :} \quad A = \| \, v, w \, \|$$

De la même manière, en utilisant la formule de la proposition 4.10, on prouve :

Proposition 4.11 – *Soit* E *un espace vectoriel de dimension* n *sur* K *et*

$$f : \underbrace{E \times \cdots \times E}_{n fois} \longrightarrow K$$

une application multilinéaire et alternée, c'est-à-dire vérifiant :

a) $f\,(v_1, \ldots, \lambda\, v_i + \mu\, w_i, \ldots, v_n\,) = \lambda\, f\,(v_1, \ldots, v_i, \ldots, v_n\,) \ + \ \mu\, f\,(v_1, \ldots, w_i, \ldots, v_n\,)$

b) $f\,(v_1, \ldots, v_i, \ldots, v_i, \ldots, v_n\,) = 0$

Soit $\{e_1, ..., e_n\}$ *une base de* E *et* $\{\, v_1, ..., v_n\}$ n *vecteurs de* E*. On a alors :*

$$f\,(\,v_1, \ldots, v_n\,) = \text{dét}\, \|\, v_1, \ldots, v_n \,\|_{(e_i)}\, f(e_1, ..., e_n\,)$$

Exercice 8.

4.5 Déterminant de la transposée d'une matrice

Théorème 4.12 – *Pour toute matrice* $A \in \mathcal{M}_n(K)$*, on a :*

$$\boxed{\text{dét}\ ({}^t A\,) = \text{dét}\, A}$$

Démonstration : Soit $A = (a_{ij})$. Comme nous l'avons vu :

$$\text{dét}\ A = \sum_{\sigma \in \mathcal{S}_n} \varepsilon\,(\sigma)\ a_{\sigma(1)1}\ \ a_{\sigma(2)2}\ \cdots\ a_{\sigma(n)n}$$

Puisque le corps K est commutatif, on peut effectuer une permutation quelconque sur les termes à second membre. Par exemple, on peut permuter les deux premiers facteurs :

$$a_{\sigma(1)1}\ \ a_{\sigma(2)2}\ \cdots\ a_{\sigma(n)n}\ =\ a_{\sigma(2)2}\ \ a_{\sigma(1)1}\ \cdots\ a_{\sigma(n)n}$$

Si τ est la transposition qui échange 1 et 2, on peut écrire cela de la manière suivante :

$$a_{\sigma(1)1}\ \ a_{\sigma(2)2}\ \cdots\ a_{\sigma(n)n}\ =\ a_{\sigma\circ\tau(1)\,\tau(1)}\ \ a_{\sigma\circ\tau(2)\,\tau(2)}\ \cdots\ a_{\sigma\circ\tau(n)\,\tau(n)}$$

Plus généralement, pour toute permutation $\rho \in \mathcal{S}_n$:

$$a_{\sigma(1)1}\ \ a_{\sigma(2)2}\ \cdots\ a_{\sigma(n)n}\ =\ a_{\sigma\circ\rho(1)\,\rho(1)}\ \ a_{\sigma\circ\rho(2)\,\rho(2)}\ \cdots\ a_{\sigma\circ\rho(n)\,\rho(n)}$$

En particulier, pour $\rho = \sigma^{-1}$

$$a_{\sigma(1)1}\ \ a_{\sigma(2)2}\ \cdots\ a_{\sigma(n)n}\ =\ a_{1\,\rho(1)}\ \cdots\ a_{n\,\rho(n)}$$

D'autre part $\varepsilon\,(\rho) = \varepsilon\,(\sigma^{-1}) = \varepsilon\,(\sigma)$ donc :

$$\varepsilon\,(\sigma)\ a_{\sigma(1)1}\ \cdots\ a_{\sigma(n)n}\ =\ \varepsilon\,(\rho)\ a_{1\,\rho(1)}\ \cdots\ a_{n\,\rho(n)}$$

Lorsque σ décrit $\mathcal{S}_n$, $\rho = \sigma^{-1}$ décrit aussi $\mathcal{S}_n$ et donc :

$$\text{dét}\, A = \sum_{\rho \in \mathcal{S}_n} \varepsilon(\rho)\, a_{1\,\rho(1)} \cdots a_{n\,\rho(n)}$$

Or ceci est exactement dét $({}^tA)$. En effet, ${}^tA = (b_{ij})$ avec $b_{ij} = a_{ji}$; donc :

$$\text{dét}\,({}^tA) = \sum_{\rho \in \mathcal{S}_n} \varepsilon(\rho)\, b_{\rho(1)\,1} \cdots b_{\rho(n)\,n} = \sum_{\rho \in \mathcal{S}_n} \varepsilon(\rho)\, a_{1\,\rho(1)} \cdots a_{n\,\rho(n)} \quad \square$$

On en déduit immédiatement :

Corollaire 4.13 ***Règle de «dualité»*** –
Toutes les propriétés des déterminants relatives aux colonnes peuvent être affirmées pour les lignes.

Ainsi par exemple, les propriétés exprimées par les théorèmes 4.2, 4.3, 4.4 admettent leur version duale :

Théorème 4.14 –

1. *Le déterminant est une fonction multilinéaire de chaque ligne.*
2. *Si une matrice a deux lignes égales, le déterminant est nul.*
3. *Si l'on échange entre elles deux lignes, le déterminant change de signe.*
4. *Le déterminant d'une matrice est non nul si et seulement si les vecteurs lignes sont indépendants,*

   ou (d'une manière équivalente) :

   un déterminant est nul si et seulement si l'une des lignes est une combinaison linéaire des autres lignes.

Exercice 9.

4.6 Calcul des déterminants

Nous avons déjà calculé le déterminant par le développement selon la I^{ère} ligne (cf. définition 4.1). Compte tenu du fait que si l'on échange entre elles des lignes le déterminant est le même au signe près, et compte tenu aussi de la règle de dualité, on peut imaginer que le rôle de la première ligne peut être tenu par l'une quelconque des lignes ou des colonnes, quitte à tenir compte du changement de signe.

Définition 4.15 – *On appelle* ***cofacteur*** *de l'élément a_{ij} le scalaire :*

$$\text{cof}\,(a_{ij}) \;:=\; (-1)^{i+j}\, \text{dét}\, A_{ij}$$

où A_{ij} est la matrice obtenue en supprimant la $i^{\text{ème}}$ ligne et la $j^{\text{ème}}$ colonne.

Par exemple, si :

$$A \quad = \begin{pmatrix} 1 & 0 & -3 \\ 2 & 4 & -2 \\ 5 & -1 & 3 \end{pmatrix} \quad , \quad \text{cof}\,(1) = + \begin{vmatrix} 4 & -2 \\ -1 & 3 \end{vmatrix} = 10$$

$$\text{cof}\,(0) \quad = - \begin{vmatrix} 2 & -2 \\ 5 & 3 \end{vmatrix} = -\,16 \qquad \text{etc.}$$

Remarques. –

1. Le signe $(-1)^{i+j}$ dans la définition de cofacteur est déterminé par le schéma suivant (schéma *en damier*) :

$$\begin{pmatrix} + & - & + & - & + & \cdots \\ - & + & - & + & - & \cdots \\ + & - & + & - & + & \cdots \\ - & + & - & + & - & \cdots \\ \multicolumn{6}{c}{\dotfill} \end{pmatrix}$$

2. Avec cette notation, la formule de la définition 4.1, s'écrit :

$$\text{dét}\, A = a_{11} \operatorname{cof}(a_{11}) \ + \ a_{12} \operatorname{cof}(a_{12}) \ + \ \cdots \ + \ a_{1n} \operatorname{cof}(a_{1n})$$

Théorème 4.16 – *On a les formules suivantes :*

Développement du déterminant selon la $j^{\text{ème}}$ ligne :

$$\boxed{\text{dét}\, A = a_{j1} \ \operatorname{cof}(a_{j1}) \ + \ a_{j2} \ \operatorname{cof}(a_{j2}) \ + \ \cdots \ + \ a_{jn} \ \operatorname{cof}(a_{jn})}$$

Développement du déterminant selon la $j^{\text{ème}}$ colonne :

$$\boxed{\text{dét}\, A = a_{1j} \ \operatorname{cof}(a_{1j}) \ + \ a_{2j} \ \operatorname{cof}(a_{2j}) \ + \ \cdots \ + \ a_{nj} \ \operatorname{cof}(a_{nj})}$$

Démonstration : Il suffit de démontrer la formule du développement suivant la $j^{\text{ème}}$ ligne (le développement selon la $j^{\text{ème}}$ colonne s'obtient par dualité, c'est-à-dire en remplaçant A par ${}^tA$).

Soit $A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}$ et considérons la matrice B obtenue de A en échangeant la première ligne avec la $j^{\text{ème}}$ ligne (on aura : dét $B = -$ dét A) :

$$B = \begin{pmatrix} b_{11} & \cdots & b_{1k} & \cdots & b_{1n} \\ \vdots & & \vdots & & \vdots \\ b_{j1} & \cdots & b_{jk} & \cdots & b_{jn} \\ \vdots & & \vdots & & \vdots \\ b_{n1} & \cdots & b_{nk} & \cdots & b_{nn} \end{pmatrix} = \begin{array}{@{}c@{}l} \begin{pmatrix} a_{j1} & \cdots & a_{jk} & \cdots & a_{jn} \\ \vdots & & \vdots & & \vdots \\ a_{11} & \cdots & a_{1k} & \cdots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ b_{n1} & \cdots & a_{nk} & \cdots & a_{nn} \end{pmatrix} & \begin{array}{l} \\ \\ \longleftarrow j^{\text{ème}}\text{ligne} \\ \\ \\ \end{array} \end{array}$$

c'est-à-dire, pour tout $k = 1, \ldots, n$:

$$b_{1k} = a_{jk}\,, \qquad b_{jk} = a_{1k} \qquad \text{et} \qquad b_{ik} = a_{ik} \quad (\text{ si } \ i \neq 1 \ \text{ et } \ i \neq j)$$

Développons B suivant la première ligne selon la définition 4.1 :

$$\begin{aligned} \text{dét}\, B \quad &= \ b_{11} \operatorname{cof}(b_{11}) + b_{12} \operatorname{cof}(b_{12}) \ + \ \cdots \ + \ b_{1n} \operatorname{cof}(b_{1n}) \\ &= \ a_{j1} \operatorname{cof}(b_{11}) + a_{j2} \operatorname{cof}(b_{12}) \ + \ \cdots \ + \ a_{jn} \operatorname{cof}(b_{1n}) \end{aligned}$$

Considérons l'un des termes de cette somme, par exemple $a_{jk}\,\mathrm{cof}\,(b_{1k})$. On a :

$$\begin{aligned}\mathrm{cof}\,(b_{1k}) \quad &= (-1)^{1+k} \begin{vmatrix} b_{11} & \cdots & b_{1k} & \cdots & b_{1n} \\ a_{21} & \cdots & a_{2k} & \cdots & a_{2n} \\ \vdots & & \vdots & & \vdots \\ a_{11} & \cdots & a_{1k} & \cdots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \cdots & a_{nk} & \cdots & a_{nn} \end{vmatrix} \\ &= (-1)^{1+k} \begin{vmatrix} a_{21} & \cdots & a_{2k} & \cdots & a_{2n} \\ \vdots & & \vdots & & \vdots \\ a_{11} & \cdots & a_{1k} & \cdots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \cdots & a_{nk} & \cdots & a_{nn} \end{vmatrix}\end{aligned}$$

Si maintenant on amène la $(j-1)^{ème}$ ligne en $I^{ère}$ ligne en effectuant $(j-2)$ transpositions avec les $(j-2)$ lignes qui la précèdent, on aura :

$$\mathrm{cof}\,(b_{1k}) = (-1)^{1+k+(j-2)} = \begin{vmatrix} a_{11} & \cdots & a_{1k} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2k} & \cdots & a_{2n} \\ \vdots & & \vdots & & \vdots \\ a_{11} & \cdots & a_{1k} & \cdots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \cdots & a_{nk} & \cdots & a_{nn} \end{vmatrix} = -\mathrm{cof}\,(a_{jk})$$

donc :

$$\text{dét}\, B = -\Big(a_{j1}\,\mathrm{cof}\,(a_{j1}) + a_{j2}\,\mathrm{cof}\,(a_{j2}) + \ \cdots\ + a_{jn}\,\mathrm{cof}\,(a_{jn})\Big)$$

Puisque dét $B = -$ dét A, on a la formule. $\square$

Exemple – Soit :

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

– D'eveloppons A selon la III^e ligne :

$$\text{dét}\, A = 1\begin{vmatrix} -2 & 3 \\ 1 & 0 \end{vmatrix} - (-1)\begin{vmatrix} 1 & 3 \\ 2 & 0 \end{vmatrix} + 2\begin{vmatrix} 1 & -2 \\ 2 & 1 \end{vmatrix} = -3 - 6 + 10 = 1$$

– Développons A selon la III^e colonne :

$$\text{dét}\, A = +3\begin{vmatrix} 2 & 1 \\ 1 & -1 \end{vmatrix} + 2\begin{vmatrix} 1 & -2 \\ 2 & 1 \end{vmatrix} = -9 + 10 = 1$$

Le calcul d'un déterminant ne présente pas de difficultés, mais il peut être très long. La somme qui apparaît dans la formule 4.10 comporte autant de termes qu'il y a de permutations dans $\mathcal{S}_n$ c'est-à-dire $n!$ (pour $n = 5$ cela donne 120 termes, pour n = 10 plus de 3 millions...). Cependant, le théorème fondamental 4.2 et sa version duale 4.14 permettent de simplifier les calculs. On utilise pour cela la propriété suivante :

Proposition 4.17 – *Le déterminant ne change pas si à une ligne (respect. à une colonne) on ajoute une combinaison linéaire des autres lignes (respect. des autres colonnes).*

La démonstration est très simple. Ajoutons par exemple à la $i^{ème}$ colonne une combinaison linéaire des autres colonnes, on a :

$$\text{dét}\,\|v_1,\ldots,v_i+\textstyle\sum_{j\neq i}\lambda_j\,v_j,\ldots,v_n\,\| \;=\; \text{dét}\,\|v_1,\ldots,v_i,\ldots,v_n\,\|$$
$$+\textstyle\sum_{j\neq i}\lambda_j\|v_1,\ldots,v_{i-1},v_j,v_{i+1},\ldots,v_n\,\| \;=\; \text{dét}\,\|v_1,\ldots,v_i,\ldots,v_n\,\|$$

car dans la somme à second membre chaque terme est nul puisqu'il comporte deux colonnes égales. $\square$

La méthode pratique pour le calcul des déterminants consiste à utiliser la propriété ci-dessus de manière à faire paraître le plus possible de zéros sur les lignes ou sur les colonnes.

Exemple – Calculer :

$$\text{dét}\,A=\begin{vmatrix}1&-2&1&3&4\\1&-1&0&2&4\\2&1&3&1&2\\-1&0&1&1&3\\0&1&-1&1&3\end{vmatrix}$$

On a :

$$\text{dét}\,A=\begin{array}{c}\begin{vmatrix}1&-2&1&3&4\\0&1&-1&-1&0\\0&5&1&-5&-6\\0&-2&2&4&7\\0&1&-1&1&3\end{vmatrix}\end{array}\begin{array}{l}\\ {\scriptstyle L_2-L_1}\\ {\scriptstyle L_3-2\,L_1}\\ {\scriptstyle L_4+L_1}\\ \\ \end{array}$$

En développant selon la première colonne :

$$\text{dét}\,A=\begin{vmatrix}1&-1&-1&0\\5&1&-5&-6\\-2&2&4&7\\1&-1&1&3\end{vmatrix}\;=\;\begin{array}{c}\begin{vmatrix}1&0&0&0\\5&6&0&-6\\-2&0&2&7\\1&0&2&3\end{vmatrix}\\ \begin{array}{cccc}&{\scriptstyle C_2+C_1}&&{\scriptstyle C_3+C_1}\end{array}\end{array}$$

en développant selon la I^{ère} ligne :

$$\text{dét}\,A\;=\;\begin{vmatrix}6&0&-6\\0&2&7\\0&2&3\end{vmatrix}\;=\;6\begin{vmatrix}2&7\\2&3\end{vmatrix}\;=\;6\,(6\;-\;14)\;=\;-\,48$$

Exercices 10. 11. 12. 13.

4.7 Déterminant du produit de matrices. Déterminant d'un endomorphisme

Théorème 4.18 – *Pour toutes matrices* $A, B \in \mathcal{M}_n(K)$, *on a :*

$$\boxed{\text{dét}\,(AB) = (\text{dét}\,A)\,(\text{dét}\,B)}$$

Pour toutes matrices $A, B \in \mathcal{M}_n(K)$

Démonstration : Soient :

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} = \|a_1, \ldots, a_n\|\ ,\ B = \begin{pmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & & \vdots \\ b_{n1} & \cdots & b_{nn} \end{pmatrix} = \|b_1, \ldots, b_n\|$$

et :
$$C = AB = \begin{pmatrix} c_{11} & \cdots & c_{1n} \\ \vdots & & \vdots \\ c_{n1} & \cdots & c_{nn} \end{pmatrix} = \|c_1, \ldots, c_n\|$$

On a, d'après la définition du produit de matrices :

$$c_{ik} = \sum_{j=1}^{n} a_{ij}\, b_{jk}$$

donc :

$$c_k = \begin{pmatrix} \displaystyle\sum_{j=1}^{n} a_{1j}\, b_{jk} \\ \vdots \\ \displaystyle\sum_{j=1}^{n} a_{nj}\, b_{jk} \end{pmatrix} = \sum_{j=1}^{n} b_{jk} \begin{pmatrix} a_{1j} \\ \vdots \\ a_{nj} \end{pmatrix} = b_{1k}\, a_1 \ + \ \cdots \ + \ b_{nk}\, a_n$$
$$= \sum_{j=1}^{n} b_{jk}\, a_j$$

d'où :

$$\text{dét}\,(AB) = \text{dét}\,\|c_1, \ldots, c_n\| = \text{dét}\,\|\sum_{j=1}^{n} b_{j1}\, a_j,\ \sum_{l=1}^{n} b_{l2}\, a_l, \ldots, \sum_{i=1}^{n} b_{in}\, a_i\|$$

$$= \sum_{j,l,\ldots,i=1}^{n} b_{j1}\, b_{l2}, \ldots, b_{in}\ \text{dét}\,\|a_j, a_l, \ldots, a_i\|$$

En raisonnant comme dans la démonstration de la proposition 4.10, on voit que les termes de cette somme avec $j, l, \ldots, i$ non tous distincts, sont nuls ; il reste les termes avec $j, l, \ldots, i$ variant de 1 à n et tous différents entre eux, c'est-à-dire $\{j, l, \ldots, i\}$ permutation de $\{1, 2, \ldots, n\}$. En écrivant, donc $j = \sigma(1),\ l = \sigma(2),\ \ldots,\ i = \sigma(n)$ avec $\sigma \in \mathcal{S}_n$, on aura :

$$\begin{aligned} \text{dét}\,(AB) &= \sum_{\sigma \in \mathcal{S}_n} b_{\sigma(1)1} \ldots b_{\sigma(n)n}\ \text{dét}\,\|a_{\sigma(1)}, \ldots, a_{\sigma(n)}\| \\ &= \sum_{\sigma \in \mathcal{S}_n} \varepsilon(\sigma)\, b_{\sigma(1)1} \ldots b_{\sigma(n)n}\ \text{dét}\,\|a_1, \ldots, a_n\| = (\text{dét}\,B)\,(\text{dét}\,A) \quad \square \end{aligned}$$

Corollaire 4.19 – *$A \in \mathcal{M}_n(K)$ est inversible si et seulement si* $\text{dét}\, A \neq 0$ *et l'on a alors :*

$$\boxed{\text{dét}\,(A^{-1}) = \frac{1}{\text{dét}\, A}}$$

En effet, supposons A inversible. Il existe alors $A^{-1} \in \mathcal{M}_n(K)$ telle que $A\,A^{-1} = I$; d'où $\text{dét}\,(A\,A^{-1}) = 1$ et par conséquent $(\text{dét}\, A)\,\text{dét}\,(A^{-1}) = 1$. Donc $\text{dét}\, A \neq 0$ et $\text{dét}\,(A^{-1}) = \dfrac{1}{\text{dét}\, A}\cdot$

Réciproquement, soit $A = \|c_1, \ldots, c_n\|$ et supposons que $\text{dét}\, A \neq 0$. Les vecteurs colonnes de A forment donc une base de K^n (cf. théorème 4.4) et A est la matrice de passage de la base canonique $\{e_i\}$ à la base $\{c_i\}$: $A = P_{e_i \longrightarrow c_i}$. Or, on sait que les matrices de passage sont inversibles, (cf. proposition 3.24 page 77) donc A est inversible. □

Corollaire 4.20 – *Si A et A' sont deux matrices semblables (c'est-à-dire, si elles représentent le même endomorphisme en des bases différentes), alors* $\text{dét}\, A = \text{dét}\, A'$.

En effet, puisque $A' = P^{-1}\,A\,P$ on a :

$$\begin{aligned} \text{dét}\, A' &= \text{dét}\,(P^{-1}\,A\,P) = \text{dét}\,(P^{-1})\ (\text{dét}\, A)\ (\text{dét}\, P) \\ &= \frac{\text{dét}\, A\ \cdot\ \text{dét}\, P}{\text{dét}\, P} = \text{dét}\, A \end{aligned}$$

□

Ce corollaire permet de poser la définition suivante :

Définition 4.21 – *Soit E un espace vectoriel de dimension finie et f un endomorphisme de E. On appelle* ***déterminant de*** *f le déterminant de la matrice qui représente f dans une base (quelconque) de E :*

$$\boxed{\text{dét}\, f = \text{dét}\, M\,(f)_{e_i}}$$

$\{e_i\}$ étant une base quelconque de E.

Exemple –

Calculer le déterminant de la matrice A qui représente dans la base canonique de $\mathbb{R}^3$ la symétrie par rapport au plan π d'équation $x + 2y + 3z = 0$, parallèlement à la droite d'équation $\dfrac{x}{2} = \dfrac{y}{3} = z$

On pourrait calculer la matrice comme dans l'exercice 19 du chapitre 3. On trouve :

$$A = \frac{1}{3}\begin{pmatrix} -1 & -4 & 8 \\ -6 & -3 & 12 \\ -2 & -2 & 7 \end{pmatrix}$$

d'où $\text{dét}\, A = -1$, comme on le vérifie en effectuant les calculs. Cependant, on peut procéder plus facilement de la manière suivante.

On considère la base $\{v_1, v_2, v_3\}$ ou v_1 est un vecteur directeur de D et $\{v_2, v_3\}$ sont deux vecteurs indépendants de π. Si g est la symétrie par rapport à π, on a :

$$g(v_1) = -v_1 \qquad g(v_2) = v_2 \qquad g(v_3) = v_3$$

d'où :

$$M(g)_{v_i} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad \text{et donc} \qquad \text{dét}\, g = -1$$

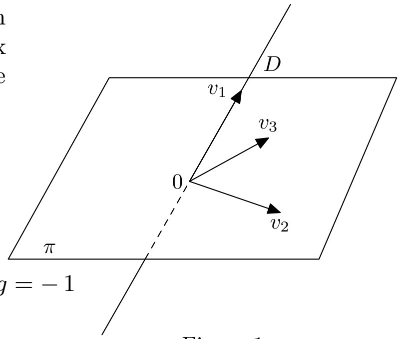


Figure 1

Exercices 14. 15. 16. 17. 18.

4.8 Calcul de l'inverse d'une matrice

Théorème 4.22 – *Soit $A \in \mathcal{M}_n(K)$ et $\mathrm{cof}(A)$ la* **comatrice** *de A, c'est-à-dire la matrice obtenue de A en remplaçant chaque élément par son cofacteur.*
On a alors :

$$\boxed{A\ {}^t\mathrm{cof}(A) \;=\; {}^t\mathrm{cof}(A)\ A \;=\; (\text{dét}\, A)\, I}$$

En particulier, **si A est inversible** *(c'est-à-dire $\text{dét}\, A \neq 0$), on a :*

$$\boxed{A^{-1} \;=\; \frac{1}{\text{dét}\, A}\ {}^t\mathrm{cof}(A)}$$

Démonstration : Considérons l'expression :

$$a_{j1}\ \mathrm{cof}(a_{k1}) + a_{j2}\ \mathrm{cof}(a_{k2}) + \cdots + a_{jn}\ \mathrm{cof}(a_{kn})$$

– si $k = j$ elle vaut $\text{dét}\, A$ (il s'agit du développement selon la $j^{\text{ème}}$ ligne) ;
– si $k \neq j$ elle vaut 0. En effet, soit B la matrice obtenue de A en remplaçant la $k^{\text{ème}}$ ligne par la $j^{\text{ème}}$ ligne. Son déterminant vaut 0, car B a deux lignes égales :

$$\begin{aligned} 0 \;=\; \text{dét}\, B \;&=\; a_{k1}\ \mathrm{cof}(a_{k1}) + \cdots + a_{kn}\ \mathrm{cof}(a_{kn}) \\ &=\; a_{j1}\ \mathrm{cof}(a_{k1}) + \cdots + a_{jn}\ \mathrm{cof}(a_{kn}) \end{aligned}$$

On a donc

$$a_{j1}\ \mathrm{cof}(a_{k1}) + a_{j2}\ \mathrm{cof}(a_{k2}) + \cdots + a_{jn}\ \mathrm{cof}(a_{kn}) \;=\; \delta_{jk}\ \text{dét}\, A \qquad (*)$$

où δ_{jk} est le symbole de Kronecker. On en déduit : $A\ {}^t\mathrm{cof}(A) = (\text{dét}\, A)\, I$. En effet, dans l'expression (*) le premier membre est le terme de la $j^{\text{ème}}$ ligne et $k^{\text{ème}}$ colonne de $A\ {}^t\mathrm{cof}(A)$ et le second membre le terme de la $j^{\text{ème}}$ ligne et $k^{\text{ème}}$ colonne de $(\text{dét}\, A)\, I$.
D'une manière tout à fait analogue, on démontre que ${}^t\mathrm{cof}(A)\ A = I$ $\square$.

Exemple 1 – Soit $A = \begin{pmatrix} 1 & 2 \\ -1 & 3 \end{pmatrix}$. On a dét $A = 5 \neq 0$, donc A est inversible. On a :

$$\operatorname{cof}(3) = 1\ ,\ \operatorname{cof}(-1) = -2\ ,\ \operatorname{cof}(2) = 1\ ,\ \operatorname{cof}(1) = 3.$$

Donc $\operatorname{cof}(A) = \begin{pmatrix} 3 & 1 \\ -2 & 1 \end{pmatrix}$ et $A^{-1} = \frac{1}{5}\begin{pmatrix} 3 & -2 \\ 1 & 1 \end{pmatrix}$.

Plus généralement, si $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, avec $a\,d - b\,c \neq 0$, on a :

$$A^{-1} = \frac{1}{a\,d - b\,c}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

Exemple 2 – $A = \begin{pmatrix} 1 & 2 & 0 \\ -1 & 3 & 0 \\ 0 & 1 & -1 \end{pmatrix}$, dét $A = -3 - 2 = -5 \neq 0$. Donc A est inversible.

Les cofacteurs des coefficients de la première ligne sont :

$$\operatorname{cof}(1) = \begin{vmatrix} 3 & 1 \\ 0 & -1 \end{vmatrix} = -3, \quad \operatorname{cof}(2) = -\begin{vmatrix} -1 & 0 \\ 0 & -1 \end{vmatrix} = -1, \quad \operatorname{cof}(0) = \begin{vmatrix} -1 & 3 \\ 0 & 1 \end{vmatrix} = -1.$$

Après calcul des autres cofacteurs, on trouve :

$$A^{-1} = -\frac{1}{5}\begin{pmatrix} -3 & 2 & 0 \\ -1 & -1 & 0 \\ -1 & -1 & 5 \end{pmatrix}$$

Exercices 19. 20.

4.9 Application des déterminants à la théorie du rang

La principale application des déterminants concerne la théorie du rang.

(A) $\boxed{\text{Caractérisation des bases}}$

Nous avons déjà vu le théorème fondamental 4.4 (page 108) qui affirme que les vecteurs $\{v_1, \ldots, v_n\}$ de K^n forment une base si et seulement si dét$\|v_1, \ldots, v_n\| \neq 0$. Ce résultat s'étend immédiatement à un espace vectoriel quelconque E de dimension finie. En effet, le choix d'une base $\{e_1, \ldots, e_n\}$ de E permet d'identifier E à K_n par l'isomorphisme.

$$\begin{array}{rccc} \varphi : & E & \longrightarrow & K^n \\ & v = x_1\,e_1 + \cdots + x_n\,e_n & \longmapsto & (x_1, \ldots, x_n) \end{array}$$

(cf. proposition 1.14). Puisque φ est un isomorphisme, les vecteurs $v_1, \ldots, v_n$ forment une base de E si et seulement si les vecteurs $\varphi(v_1), \ldots, \varphi(v_n)$, c'est-à-dire les composantes des vecteurs v_i, forment une base de K^n (et, plus généralement : $\operatorname{rg}\{v_1, \ldots, v_n\} = \operatorname{rg}\{\varphi(v_1), \ldots, \varphi(v_n)\}$). Aussi le théorème 4.4 peut s'énoncer :

Théorème 4.23 – **(A)**
Soit E un espace vectoriel de dimension n. Les vecteurs $\{v_1, \ldots, v_n\}$ de E forment une base si et seulement si :

$$\text{dét}\, \|v_1, \ldots, v_n\|_{e_i} \neq 0$$

où $\|v_1, \ldots, v_n\|_{e_i}$ désigne la matrice dont les colonnes sont les composantes des vecteurs $v_1, \ldots, v_n$ dans une base $\{e_i\}$ (quelconque) de E.

Nota. – Dans la suite, si le contexte est clair, on notera simplement $\|v_1, \ldots, v_n\|$ la matrice dont les colonnes sont les composantes des vecteurs $v_1, \ldots, v_n$ dans une base de E.

(B) **Comment reconnaître si une famille de vecteurs est libre**

On appelle ***mineur*** d'ordre r d'une matrice A le déterminant d'une matrice d'ordre r extraite de A obtenue en choisissant r lignes et r colonnes.

Par exemple, en choisissant la $2^{ème}$ et $3^{ème}$ ligne et la $2^{ème}$ et $4^{ème}$ colonne dans la matrice

$$\begin{pmatrix} 1 & 2 & 7 & 3 & 2 \\ 0 & \boxed{3} & 5 & \boxed{4} & 1 \\ 2 & \boxed{1} & 3 & \boxed{6} & 1 \\ 2 & 1 & -1 & 3 & 2 \end{pmatrix} \qquad \text{on obtient le mineur} \quad \delta = \begin{vmatrix} 3 & 4 \\ 1 & 6 \end{vmatrix} = 14$$

Théorème 4.24 – **(B)**
Soient $\{v_1, \ldots, v_r\}$ r vecteurs d'un espace vectoriel E de dimension n $(r \leq n)$ et $A = \|v_1, \ldots, v_r\|$ la matrice dont les colonnes sont les composantes des vecteurs $v_1, \ldots, v_r$ dans une base quelconque de E.
La famille $\{v_1, \ldots, v_r\}$ est libre *si et seulement si on peut extraire de A un mineur d'ordre r non nul.*

D'après la remarque ci-dessus on peut se limiter à démontrer ce théorème, ainsi que ceux qui suivent, pour les vecteurs de K^n, ce qui simplifiera les notations.

Démonstration : Montrons d'abord que la condition est suffisante.
Soit :

$$A = \begin{pmatrix} \boxed{\begin{matrix} a_{11} & \cdots & a_{1r} \\ \vdots & & \vdots \\ a_{r1} & \cdots & a_{rr} \end{matrix}} \\ \begin{matrix} \vdots & & \vdots \\ a_{n1} & \cdots & a_{nr} \end{matrix} \end{pmatrix}$$

Quitte à permuter les lignes – ce qui n'affecte pas le rang de la matrice – on peut supposer que le mineur encadré est non nul.

Considérons l'application :

$$p: \begin{array}{ccc} K^n & \longrightarrow & K^r \\ \begin{pmatrix} b_1 \\ \vdots \\ b_r \\ \vdots \\ b_n \end{pmatrix} & \longmapsto & \begin{pmatrix} b_1 \\ \vdots \\ b_r \end{pmatrix} \end{array}$$

Il s'agit d'une application linéaire. On a :

$$\| p\,(v_1), \ldots, p\,(v_r) \| = \begin{pmatrix} a_{11} & \cdots & a_{1r} \\ \vdots & & \vdots \\ a_{r1} & \cdots & a_{rr} \end{pmatrix}$$

Donc

$$\text{dét}\, \| p\,(v_1), \ldots, p\,(v_r) \| \neq 0$$

D'après le théorème (A), $\{p\,(v_1), \ldots, p\,(v_r)\}$ est une base de K^r, donc une famille libre. On en déduit que $\{v_1, \ldots, v_r\}$ est libre. En effet si $\lambda_1\, v_1 + \cdots + \lambda_r\, v_r = 0$, en prenant l'image par p, on a : $\lambda_1\; p\,(v_1) + \cdots + \lambda_r\; p\,(v_r) = 0$ et donc $\lambda_1 = 0, \ldots, \lambda_r = 0$. Réciproquement, supposons la famille $\{v_1, \ldots, v_r\}$ libre. On peut alors compléter cette famille par $n-r$ vecteurs de la base canonique de K^n de manière à former une base[5]. Quitte à changer la numérotation de la base canonique, on peut supposer que $\{v_1, \ldots, v_r,\ e_{r+1}, \ldots, e_n\}$ est une base. Soit :

$$B = \| v_1, \ldots, v_r,\ e_{r+1}, \ldots, e_n \|_{e_i} = \begin{pmatrix} a_{11} & \cdots & a_{1r} & 0 & \cdots & \cdots & 0 \\ \vdots & & \vdots & \vdots & & & \vdots \\ a_{r1} & \cdots & a_{rr} & 0 & \cdots & \cdots & 0 \\ a_{r+1,\,1} & \cdots & a_{r+1,\,r} & 1 & 0 & \cdots & 0 \\ \vdots & & \vdots & 0 & 1 & & \vdots \\ \vdots & & \vdots & \vdots & \vdots & \ddots & 0 \\ a_{n1} & \cdots & a_{nr} & 0 & 0 & \cdots & 1 \end{pmatrix}$$

D'après le théorème (A), $\text{dét}\, B \neq 0$. En développant $\text{dét}\, B$ selon la dernière colonne et en répétant successivement cette opération, on trouve :

$$\text{dét}\, B = \begin{vmatrix} a_{11} & \cdots & a_{1r} \\ \vdots & & \vdots \\ a_{r1} & \cdots & a_{rr} \end{vmatrix}$$

On peut donc extraire de A un mineur d'ordre r non nul. $\square$

[5]Le théorème de la base incomplète dit non seulement que toute famille libre peut être complétée en une base, mais qu'elle peut être complétée en choisissant les vecteurs dans une famille génératrice quelconque fixée au préalable (cf. remarque page 16.)

Exemple –

Les vecteurs :

$$v_1 = \begin{pmatrix} 1 \\ 2 \\ 3 \\ 3 \\ 5 \end{pmatrix} \qquad v_2 = \begin{pmatrix} 0 \\ 1 \\ 2 \\ 4 \\ 0 \end{pmatrix} \qquad v_3 = \begin{pmatrix} 1 \\ 5 \\ 9 \\ -2 \\ 0 \end{pmatrix}$$

forment une famille libre de $\mathbb{R}^5$ car dans la matrice :

$$A = \left(\begin{array}{ccc} 1 & 0 & 1 \\ 2 & 1 & 5 \\ \hline \multicolumn{1}{|c}{3} & 2 & \multicolumn{1}{c|}{9} \\ \multicolumn{1}{|c}{3} & 4 & \multicolumn{1}{c|}{-2} \\ \multicolumn{1}{|c}{5} & 0 & \multicolumn{1}{c|}{0} \\ \hline \end{array}\right)$$

le mineur encadré est non nul.

(C) **COMMENT RECONNAÎTRE SI UN VECTEUR APPARTIENT À L'ESPACE ENGENDRÉ PAR D'AUTRES VECTEURS**

Définition 4.25 – *Soit A une matrice et δ un mineur d'ordre r extrait de A. On appelle* ***bordant*** *de δ tout mineur d'ordre $r+1$ extrait de A, dont δ est un déterminant extrait.*

Par exemple, soit : $\left(\begin{array}{ccccc} \hline \multicolumn{1}{|c}{1} & \multicolumn{1}{c|}{2} & 2 & -1 & 0 \\ \multicolumn{1}{|c}{1} & \multicolumn{1}{c|}{3} & 4 & 2 & -1 \\ \hline 3 & -1 & 2 & -1 & 0 \\ 4 & 1 & 4 & 2 & -1 \end{array}\right)$ et $\delta = \begin{vmatrix} 1 & 2 \\ 1 & 3 \end{vmatrix}$. Les bordants de δ sont :

en bordant avec la III^{eme} ligne $\longrightarrow$ $\begin{vmatrix} 1 & 2 & 2 \\ 1 & 3 & 4 \\ 3 & -1 & 2 \end{vmatrix}$ $\begin{vmatrix} 1 & 2 & -1 \\ 1 & 3 & 2 \\ 3 & -1 & -1 \end{vmatrix}$ $\begin{vmatrix} 1 & 2 & 0 \\ 1 & 3 & -1 \\ 3 & -1 & 0 \end{vmatrix}$

en bordant avec la IV^{eme} ligne $\longrightarrow$ $\begin{vmatrix} 1 & 2 & 2 \\ 1 & 3 & 4 \\ 4 & 1 & 4 \end{vmatrix}$ $\begin{vmatrix} 1 & 2 & -1 \\ 1 & 3 & 2 \\ 4 & 1 & 2 \end{vmatrix}$ $\begin{vmatrix} 1 & 2 & 0 \\ 1 & 3 & -1 \\ 4 & 1 & -1 \end{vmatrix}$

$\uparrow$ en bordant avec la III^{eme} colonne — $\uparrow$ en bordant avec la IV^{eme} colonne — $\uparrow$ en bordant avec la V^{eme} colonne

On voit immédiatement que si $A \in \mathcal{M}_{p,\,n}$ et δ est un mineur d'ordre r il y a exactement $(p-r)\,(n-r)$ bordants de δ dans A.

Théorème 4.26 – **(C)**

Soient $\{v_1, \ldots, v_r\}$ r vecteurs linéairement indépendants, $A = \|v_1, \ldots, v_r\|$ la matrice dont les colonnes sont les composantes de ces vecteurs dans une base quelconque, et δ un mineur d'ordre r non nul extrait de A.

Pour qu'un vecteur w appartienne à l'espace $\mathrm{Vect}\{v_1, \ldots, v_r\}$ *il faut et il suffit que tous les bordants de δ dans la matrice $B = \|v_1, \ldots, v_r, w\|$ soient nuls.*

Démonstration : Comme dans les théorèmes précédents, on peut se limiter à considérer les vecteurs de K^n.
Soient :

$$A = \begin{pmatrix} \boxed{\begin{matrix} a_{11} & \cdots & \cdots & a_{1r} \\ \vdots & & & \vdots \\ a_{r1} & \cdots & \cdots & a_{rr} \end{matrix}} \\ \begin{matrix} a_{r+1,1} & \cdots & a_{r+1,r} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{matrix} \end{pmatrix} \quad \text{et} \quad B = \begin{pmatrix} \boxed{\begin{matrix} a_{11} & \cdots & \cdots & a_{1r} \\ \vdots & & & \vdots \\ a_{r1} & \cdots & \cdots & a_{rr} \end{matrix}} & \begin{matrix} b_1 \\ \vdots \\ b_r \end{matrix} \\ \begin{matrix} a_{r+1,1} & \cdots & a_{r+1,r} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{matrix} & \begin{matrix} b_{r+1} \\ \vdots \\ b_n \end{matrix} \end{pmatrix}$$

Quitte changer l'ordre des lignes et des colonnes, on peut supposer que le mineur δ non nul est le mineur encadré.
Supposons que $w \in \text{Vect}\{v_1, \ldots, v_r\}$. Si l'un des bordants était non nul, la famille $\{v_1, \ldots, v_r, w\}$ serait libre $\big($cf. théorème (B) $\big)$, ce qui est exclu, car w appartient l'espace engendre par $\{v_1, .., v_r\}$. Donc tous les bordants de δ dans B sont nuls.
Réciproquement, supposons que tous les bordants de δ sont nuls.
En regardant les vecteurs ***lignes*** de la matrice B, on voit que les r premiers sont indépendants, car $\delta \neq 0$ $\big($théorème (B)$\big)$ et chacun des autres est lié aux r premiers $\big($théorème (A)$\big)$. Ainsi la matrice B a pour rang r. Donc les vecteurs colonnes de B, $\{v_1, \ldots, v_r, w\}$, forment une famille de rang r et comme $\{v_1, \ldots, v_r\}$ est une famille libre, $w \in \text{Vect}\{v_1, \ldots, v_r\}$. $\square$

Exemple – *Pour quelles valeurs de* $\alpha, \beta \in \mathbb{R}$, *le vecteur* $w = \begin{pmatrix} \alpha \\ 2 \\ 1 \\ \beta \end{pmatrix}$ *appartient-il au sous-espace de* $\mathbb{R}^4$ *engendré par les vecteurs* $v_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}$ *et* $v_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix}$ *?* On a :

$$A = \begin{pmatrix} \boxed{\begin{matrix} 1 & 0 \\ 0 & 1 \end{matrix}} \\ \begin{matrix} 1 & 0 \\ 0 & 1 \end{matrix} \end{pmatrix} \qquad \delta = \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} \qquad \textit{et} \quad B = \begin{pmatrix} \boxed{\begin{matrix} 1 & 0 \\ 0 & 1 \end{matrix}} & \begin{matrix} \alpha \\ 2 \end{matrix} \\ \begin{matrix} 1 & 0 \\ 0 & 1 \end{matrix} & \begin{matrix} 1 \\ \beta \end{matrix} \end{pmatrix}$$

Les bordants de δ sont :

$$\Delta_1 = \begin{vmatrix} \boxed{\begin{matrix} 1 & 0 \\ 0 & 1 \end{matrix}} & \begin{matrix} \alpha \\ 2 \end{matrix} \\ \begin{matrix} 1 & 0 \end{matrix} & 1 \end{vmatrix} = 1 - \alpha \quad \text{et} \quad \Delta_2 = \begin{vmatrix} \boxed{\begin{matrix} 1 & 0 \\ 0 & 1 \end{matrix}} & \begin{matrix} \alpha \\ 2 \end{matrix} \\ \begin{matrix} 0 & 1 \end{matrix} & \beta \end{vmatrix} = \beta - 2$$

Donc $w \in \text{Vect}\{v_1, v_2\}$ si et seulement si $\alpha = 1$ et $\beta = 2$.

D) Détermination du rang

Théorème 4.27 – **(D)** *Soit* $A \in \mathcal{M}_{p,n}(K)$. *Le rang de* A *est* r *si et seulement si on peut extraire de* A *un mineur* δ *d'ordre* r *non nul tel que tous les bordants de* δ *dans* A *sont nuls.*

Démonstration : Les conditions sont suffisantes. Soit, en effet :

$$A = \|v_1, \ldots, v_n\| = \begin{pmatrix} \boxed{\begin{matrix} a_{11} & \cdots & a_{1r} \\ \vdots & & \vdots \\ a_{r1} & \cdots & a_{rr} \end{matrix}} & \begin{matrix} \cdots & a_{1n} \\ & \vdots \\ \cdots & a_{rn} \end{matrix} \\ \begin{matrix} \vdots & & \vdots \\ a_{p1} & \cdots & a_{pr} \end{matrix} & \begin{matrix} & \vdots \\ \cdots & a_{pn} \end{matrix} \end{pmatrix}$$

et soit $\delta \neq 0$ le mineur encadré. Comme $\delta \neq 0$, les vecteurs $\{v_1, \ldots, v_r\}$ sont indépendants $\big($théorème (B)$\big)$. Puisque tous les bordants de δ sont nuls, chaque vecteur v_s $(s \geq r+1)$ appartient à l'espace $\mathrm{Vect}\{v_1, \ldots, v_r\}$ $\big($théorème (C)$\big)$. Donc les vecteurs colonnes engendrent un espace de dimension r et par conséquent $\mathrm{rg}\, A = r$.

Réciproquement, soit $\mathrm{rg}\, A = r$ et supposons, quitte à changer l'ordre des colonnes, que $\{v_1, \ldots, v_r\}$ est une famille libre. On peut alors extraire de la matrice formée par les r premières colonnes de A un mineur δ d'ordre r non nul $\big($théorème (B)$\big)$. Quitte à changer la numérotation des coordonnées, on peut supposer que δ est le mineur formé par les r premières lignes et colonnes de A. Or $\mathrm{rg}\, A = r$ donc $v_s \in \mathrm{Vect}\{v_1, \ldots, v_r\}$ pour $s \geq r+1$; d'après le théorème (C), tous les bordants de δ dans A sont nuls. $\square$

Théorème 4.28 – **(E)**

Le rang d'une matrice A est l'ordre maximal des mineurs non nuls extraits de A, c'est-à-dire :

$$\mathrm{rg}\, A = r \iff \begin{cases} \text{- } \textit{il existe un mineur d'ordre } r \textit{ non nul} \\ \text{- } \textit{tous les mineurs d'ordre } s > r \textit{ sont nuls} \end{cases}$$

Démonstration : Supposons que le rang de A est r ; s'il existait un mineur d'ordre $s > r$ non nul, on pourrait extraire des colonnes de A une famille libre formée de $s > r$ vecteurs $\big($théorème (B)$\big)$. Or ceci est impossible car les vecteurs colonnes de A engendrent un espace de dimension r.

La réciproque est une conséquence du théorème précédent. $\square$

Exercices 21. 22.

4.10 Interprétation géométrique du déterminant : volume dans $\mathbb{R}^n$

NOTA. – La présentation que nous donnons ici est tirée du livre de S. Lang : *Algèbre linéaire 1*, InterEditions, Paris, 1976.

Comme nous l'avons dit dans l'introduction, le déterminant possède une remarquable interprétation en termes de volume dans $\mathbb{R}^n$.

Etudions d'abord le cas de dimension 2 et considérons le parallélogramme engendré par deux vecteurs v et w de $\mathbb{R}^2$.

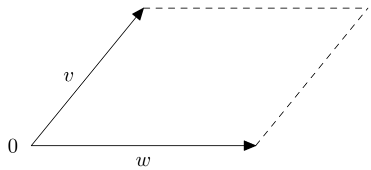


Figure 2

En regardant v et w comme deux vecteurs colonnes, on peut considérer le déterminant dét $\|v,\, w\|$. Etant donné que

$$\text{dét}\,\|v,\, w\| \;=\; -\,\text{dét}\,\|w,\, v\|$$

le déterminant lui-même ne peut être l'aire du parallélogramme, puisque cette aire doit être ≥ 0. On a cependant :

Théorème 4.29 – *L'aire $\mathcal{A}\,(v,\, w)$ du parallélogramme engendré par v et w est égale à la valeur absolue du déterminant* dét $\|v,\, w\|$. *C'est-à-dire :*

$$\boxed{\mathcal{A}\,(v,\, w) = \Big|\text{dét}\,\|\, v,\, w\,\|\Big|}$$

Démonstration : Posons :

$$\tilde{\mathcal{A}}\,(v,\, w) = \begin{cases} \mathcal{A}\,(v,\, w) & \text{si} \quad \text{dét}\,\|v,\, w\| \geq 0 \\ -\,\mathcal{A}(v,\, w) & \text{si} \quad \text{dét}\,\|v,\, w\| < 0 \end{cases}$$

$\tilde{\mathcal{A}}$ est dite *aire orientée*. Avec cette notation, il s'agit de démontrer que :

$$\tilde{\mathcal{A}}\,(v,\, w) = \text{dét}\,\|v,\, w\|.$$

Pour cela il suffira, d'après la proposition 4.11 page 114, de montrer que $\tilde{\mathcal{A}}$ est bilinéaire, alternée et que $\tilde{\mathcal{A}}\,(e_1,\, e_2) = 1$.

On a immédiatement :

$$\tilde{\mathcal{A}}\,(v,\, v) \;=\; \pm\,\mathcal{A}\,(v,\, v) \;=\; 0$$

et :

$$\tilde{\mathcal{A}}\,(e_1,\, e_2) \;=\; \frac{\text{dét}\,\|e_1,\, e_2\|}{\Big|\text{dét}\,\|e_1, e_2\|\Big|}\,\mathcal{A}\,(e_1,\, e_2) \;=\; \frac{1}{1} \times 1 \;=\; 1$$

Il ne reste plus à démontrer que $\tilde{\mathcal{A}}$ est bilinéaire.

Lemme. – *Pour tout $\lambda \in \mathbb{R}$, on a :*

$$\mathcal{A}\,(\lambda\, v,\; w) = |\,\lambda\,|\; \mathcal{A}\,(v,\, w)$$

Pour montrer ce lemme, il suffira de vérifier que :

$$\text{a)} \qquad \mathcal{A}\,(-v,\, w) \;=\; \mathcal{A}\,(v,\, w)$$

$$\text{b)} \qquad \mathcal{A}\,(c\,v,\, w) \;=\; c\,\mathcal{A}\,(v,\, w)\,, \quad \forall\, c \in \mathbb{R},\; c \geq 0$$

Si v et w sont liés, ou si $c = 0$, a) et b) sont évidentes car l'aire du parallélogramme engendré par deux vecteurs liés est nulle. Supposons donc que v et w sont indépendants et $c \neq 0$.

a) est immédiate : le parallélogramme engendré par $-v$ et w est le translaté du parallélogramme engendré par v et w (cf. fig. 3).

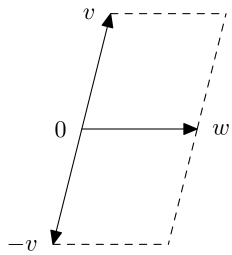


Figure 3

Montrons b). On a d'abord :

$\mathcal{A}(n\,v,\ w) \;=\; n\,\mathcal{A}(v,\,w)$ pour tout entier $n > 0$

car le parallélogramme défini par nv et w est la réunion de n parallélogrammes qui sont les translatés du parallélogramme défini par v et w (cf. fig. 4).

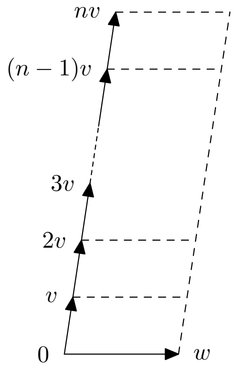


Figure 4

On en déduit, en remplaçant v par $\dfrac{1}{n}v$ dans la formule trouvée :

$$\mathcal{A}\left(\frac{1}{n}\,v,\,w\right) = \frac{1}{n}\,\mathcal{A}(v,\,w)$$

et donc :

$$\mathcal{A}\left(\frac{m}{n}\,v,\,w\right) \;=\; \frac{m}{n}\,\mathcal{A}(v,\,w)\,, \qquad \begin{array}{l}\forall\, m,\,n\in\mathbb{N}\\ m>0,\ n>0\end{array}$$

Soient maintenant r et r' deux nombres rationnels tels que :
$0 < r < c < r'$; on a (cf. fig. 5) :

$$\mathcal{A}(r\,v,\ w) \le \mathcal{A}(c\,v,\ w) \le \mathcal{A}(r'\,v,\ w)$$

donc :

$$r\,\mathcal{A}(v,\,w) \le \mathcal{A}(c\,v,\ w) \le r'\,\mathcal{A}(v,\,w)$$

en faisant tendre r et r' vers c, on obtient :

$$c\,\mathcal{A}(v,\,w) = \mathcal{A}(c,\,w)$$

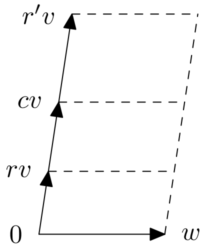


Figure 5

ce qui montre le lemme. $\diamondsuit$

En revenant à la proposition, on en déduit que :

$$\tilde{\mathcal{A}}(\lambda\,v,\ w) \;=\; \lambda\,\tilde{\mathcal{A}}(v,\,w)\,, \quad \forall\,\lambda\in\mathbb{R}.$$

En effet, si v et w sont liés, ou si λ est nul, cela est trivial. Autrement :

$$\begin{aligned}\tilde{\mathcal{A}}(\lambda\,v,\,w) \quad &= \frac{\text{dét}\,\|\lambda\,v,\,w\|}{\Big|\,\text{dét}\,\|\lambda\,v,\ w\|\,\Big|}\mathcal{A}(\lambda\,v,\ w)\\ &= \frac{\lambda}{|\,\lambda\,|}\ \frac{\text{dét}\,\|v,\,w\|}{\Big|\text{dét}\,\|v,\,w\|\Big|}\,|\,\lambda\,|\ \mathcal{A}(v,\,w) \;=\; \lambda\,\tilde{\mathcal{A}}(v,\,w)\end{aligned}$$

Il ne reste plus à démontrer que :

$$\mathcal{A}(v_1 + v_2,\, w) \;=\; \mathcal{A}(v_1,\, w) + \mathcal{A}(v_2,\, w)\,, \quad \forall\, v_1,\, v_2,\, w \in \mathbb{R}^2$$

(la linéarité dans le second argument se démontre de la même manière).
La démonstration se fait par les étapes suivantes :

(i) $\quad A(v + w,\, w) \;=\; \mathcal{A}(v,\, w)$

En effet :

$$\begin{aligned} \mathcal{A}(v + w,\, w) &= \mathrm{Aire}\, B \,+\, \mathrm{Aire}\, C \\ \mathcal{A}(v,\, w) &= \mathrm{Aire}\, A \,+\, \mathrm{Aire}\, B \end{aligned}$$

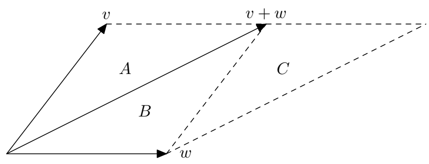


Figure 6

et d'autre part $\mathrm{Aire}\, A = \mathrm{Aire}\, C$, car le triangle C est le translaté du triangle A.

(ii) $\quad \tilde{\mathcal{A}}(v + w,\, w) \;=\; \tilde{\mathcal{A}}(v,\, w).$

Cela est évident si v et w sont liés. Si v et w sont libres :

$$\begin{aligned} \tilde{\mathcal{A}}(v + w,\, w) &= \frac{\Big|\mathrm{dét}\, \|v + w,\, w\|\Big|}{\mathrm{dét}\, \|v + w,\, w\|}\mathcal{A}(v + w,\, w) \\ &= \frac{\Big|\mathrm{dét}\, \|v,\, w\|\Big|}{\mathrm{dét}\, \|v,\, w\|}\mathcal{A}(v,\, w) \;=\; \tilde{\mathcal{A}}(v,\, w) \end{aligned}$$

(iii) $\quad \tilde{\mathcal{A}}(\lambda\, v + \mu\, w,\, w) = \lambda\, \tilde{\mathcal{A}}(v,\, w)\,, \quad \forall\, \lambda,\, \mu \in \mathbb{R}.$

Si $\mu = 0$, cela a déjà été démontré.
Supposons $\mu \neq 0$:

$$\begin{aligned} \tilde{\mathcal{A}}(\lambda\, v + \mu\, w,\, w) &= \tilde{\mathcal{A}}\left(\lambda\, v + \mu\, w,\, \frac{\mu}{\mu}\, w\right) = \frac{1}{\mu}\, \tilde{\mathcal{A}}(\lambda\, v + \mu\, w,\, \mu\, w) \\ &\underset{\text{(ii)}}{=} \tfrac{1}{\mu}\, \tilde{\mathcal{A}}(\lambda\, v,\, \mu\, w) = \tilde{\mathcal{A}}\left(\lambda\, v,\, \frac{\mu}{\mu}\, w\right) = \tilde{\mathcal{A}}(\lambda\, v,\, w) \end{aligned}$$

Soient maintenant $v_1 = a\, v + b\, w$ et $v_2 = c\, v + d\, w$, on a :

$$\begin{aligned} \tilde{\mathcal{A}}(v_1 + v_2,\, w) &= \tilde{\mathcal{A}}((a + c)\, v + (b + d)\, w,\, w) \underset{\text{(ii)}}{=} (a + c)\, \tilde{\mathcal{A}}(v,\, w) \\ &= a\, \tilde{\mathcal{A}}(v,\, w) + c\, \tilde{\mathcal{A}}(v,\, w) \;=\; \tilde{\mathcal{A}}(a\, v,\, w) + \tilde{\mathcal{A}}(c\, v,\, w) \\ &\underset{\text{(ii)}}{=} \tilde{\mathcal{A}}(v_1,\, w) + \tilde{\mathcal{A}}(v_2,\, w) \end{aligned}$$

□

L'intérêt de cette démonstration réside dans le fait qu'elle peut se généraliser facilement en dimension supérieure.
Notons que le parallélogramme engendré par v_1 et v_2 est l'ensemble :

$$\left\{ z \in \mathbb{R}^2 \,|\, z = \lambda_1\, v_1 + \lambda_2\, v_2 \quad \text{avec} \quad 0 \leq \lambda_i \leq 1\ ,\ i = 1,\, 2 \right\}$$

On a :

Théorème 4.30 – *Soient* $v_1, \dots, v_n$ n *vecteurs de* $\mathbb{R}^n$ *; on note* $\mathrm{Vol}\,(v_1, \dots, v_n)$ *le volume du parallélépipède engendré par* $v_1, \dots, v_n$*, c'est-à-dire de l'ensemble :*

$$\{\ z \in \mathbb{R}^n \mid z = \lambda_1\, v_1 + \cdots + \lambda_n\, v_n\ , \quad \text{avec :} \quad 0 \leq \lambda_i \leq 1\ ,\ i = 1, \dots, n\ \}$$

On a alors :

$$\mathrm{Vol}\ (v_1, \dots, v_n)\ =\ \Big|\text{dét}\, \|\, v_1, \dots, v_n\, \|\ \Big|$$

La démonstration se fait exactement de la même manière qu'en dimension 2. Celle-ci, en effet, repose essentiellement sur le fait que le déterminant est une forme multilinéaire alternée. Or les conditions qui expriment que le déterminant est une forme multilinéaire alternée f (cf. proposition 4.11 page 114) ne font intervenir que *deux* arguments de f en même temps. En gardant donc fixes tous les v_i sauf deux, on peut étendre sans trop de difficultés la démonstration donnée ci-dessus. □

Exercice 23.

4.11 Orientation

L'une des applications les plus importantes de la théorie des déterminants est qu'elle permet de définir la notion d'orientation.
Soit $\mathcal{B} = \{v_1, v_2, v_3\}$ une base ordonnée de $\mathbb{R}^3$; puisque $\mathcal{B}$ est une base, dét $\|v_1, v_2, v_3\| \neq 0$. On dit que $\mathcal{B}$ est orientée positivement si dét $\|v_1, v_2, v_3\|\ >\ 0$; dans le cas contraire, on dit que $\mathcal{B}$ est orientée négativement :

Figure 7

On sait l'importance que cette notion d'orientation a en physique (par exemple en électromagnétisme) et, plus généralement, dans tous les problèmes où il intervient des rotations. Elle sert à donner une signification précise à la notion de *sens de la rotation* ce qui revient à choisir un ordre dans les repères de référence.

La notion d'orientation se généralise à un espace vectoriel réel quelconque de la manière suivante.
Soit E un espace vectoriel **réel** de dimension n , $\{e_i\}$ et $\{e'_i\}$ deux bases *ordonnées* différentes[6] de E et $P = P_{e_i \longrightarrow e'_i}$ la matrice de passage. Puisque dét $P \neq 0$, on peut poser la définition suivante :

[6]«différentes» en ce sens que deux bases sont considérées différentes même si elles ne diffèrent que par l'ordre des éléments. Par la suite nous supposerons toujours cela, même si nous ne le disons pas explicitement.

Définition 4.31 – *On dit que les bases* $\{e_i\}$ *et* $\{e'_i\}$ *d'un espace vectoriel réel de dimension finie* ***ont même orientation*** *si* dét $P_{e_i \longrightarrow e'_i} > 0$ *(dans le cas contraire, on dit qu'elles ont une orientation opposée).*

REMARQUE – On peut poser cette définition car on a les propriétés suivantes :

- $\{e_i\}$ *a la même orientation que* $\{e_i\}$, puisque dét $P_{e_i \longrightarrow e_i} =$ dét $I = 1 > 0$;
- *si* $\{e_i\}$ *a la même orientation que* $\{e'_i\}$ *alors* $\{e'_i\}$ *a la même orientation que* $\{e_i\}$ car si dét $P_{e_i \longrightarrow e'_i} > 0$ alors dét $P_{e'_i \longrightarrow e_i} = (\text{dét } P_{e_i \longrightarrow e'_i})^{-1} > 0$;
- et, enfin, *si* $\{e_i\}$ *a la même orientation que* $\{e'_i\}$ *et* $\{e'_i\}$ *a la même orientation que* $\{e''_i\}$, *alors* $\{e_i\}$ *a la même orientation que* $\{e''_i\}$ (vérification laissée en exercice).

Le lecteur familiarisé avec la notion de quotient (cf. Appendice A.3.), aura remarqué qu'il s'agit d'une *relation d'équivalence.*

Proposition 4.32 – *L'ensemble* $\mathbb{B}$ *de toutes les bases de* E *se partage en deux sous-ensembles disjoints non vides :* $\mathbb{B} = \mathbb{B}_1 \cup \mathbb{B}_2$*. Toutes les bases de* $\mathbb{B}_1$ *(respect. : de* $\mathbb{B}_2$*) ont même orientation.* $\mathbb{B}_1$ *et* $\mathbb{B}_2$ *sont dites* ***classes d'orientation.***

La démonstration est très simple[7]. Choisissons une base $\mathcal{B}$ de E et notons :

$$
\begin{aligned}
C^+_{\mathcal{B}} : \quad &= \{ \text{ bases de } E, \text{ de même orientation que } \mathcal{B} \ \} \\
&= \{ \ \mathcal{B}' \in \mathbb{B} \mid \text{dét } P_{\mathcal{B} \longrightarrow \mathcal{B}'} > 0 \ \} \\
C^-_{\mathcal{B}} : \quad &= \{ \text{ bases de } E, \text{ d'orientation opposée à celle de } \mathcal{B} \ \} \\
&= \{ \ \mathcal{B}' \in \mathbb{B} \mid \text{dét } P_{\mathcal{B} \longrightarrow \mathcal{B}'} < 0 \ \}
\end{aligned}
$$

Il est clair que toute base de E est soit dans $\mathcal{C}^+_{\mathcal{B}}$ soit dans $\mathcal{C}^-_{\mathcal{B}}$ et qu'elle ne peut être dans les deux à la fois ; ainsi $\mathbb{B}$ est la réunion disjointe de $C^+_{\mathcal{B}}$ et de $C^-_{\mathcal{B}}$.
Si maintenant nous fixons une autre base $\mathcal{B}'$ de E ont voit immédiatement que :

$$
\begin{aligned}
C^+_{\mathcal{B}'} &= C^+_{\mathcal{B}} \qquad \text{si } \mathcal{B} \text{ et } \mathcal{B}' \text{ ont même orientation} \\
C^+_{\mathcal{B}'} &= C^-_{\mathcal{B}} \qquad \text{si } \mathcal{B} \text{ et } \mathcal{B}' \text{ ont une orientation opposée}
\end{aligned}
$$

ce qui montre que la partition construite à l'aide de $\mathcal{B}$ est indépendante du choix de $\mathcal{B}$. □

Définition 4.33 – *Soit* E *un espace vectoriel* réel *de dimension finie. On dit que l'****on a fixé une orientation*** *de* E *si l'on a choisi une classe d'orientation. On dira alors que les bases de cette classe sont* ***orientées positivement*** *(ou qu'elles sont* ***directes****) et celles de l'autre négativement.*

En d'autres termes, fixer une orientation revient simplement à affecter d'un signe $+$ une classe et d'un signe $-$ l'autre classe.

NOTA. – *Pour choisir une orientation, il suffit de choisir une base* $\mathcal{B}$ *de* E : seront orientées positivement (par définition) toutes les bases $\mathcal{B}'$ qui sont dans la même classe que $\mathcal{B}$, c'est-à-dire telles que dét $P_{\mathcal{B} \longrightarrow \mathcal{B}'} > 0$.

[7] Elle est immédiate si l'on fait appel à la notion de quotient (cf Appendice A.3).

Exemple – **Orientation canonique de $\mathbb{R}^n$.**

Puisque sur $\mathbb{R}^n$ il y a une base canonique $\{e_i\}$, $\mathbb{R}^n$ est muni d'une orientation canonique : la base $\{v_1, \ldots, v_n\}$ est orientée positivement si et seulement si $\text{dét}\, P_{e_i \longrightarrow v_i} > 0$, c'est-à-dire si $\text{dét}\, \|v_1, \ldots, v_n\| > 0$, comme nous l'avons dit ci-dessus pour $n = 3$.

De la même manière on peut orienter canoniquement $\mathbb{R}_n[X]$ par le choix de la base $\{1, X, ..., X^n\}$ ou encore $\mathcal{M}_{pn}(\mathbb{R})$, en choisissant la base formée par les matrices élémentaires $\{E_{ij}\}$ (cf. page 65.)

Orientation induite

Comme on vient de le voir, dans certains cas l'espace vectoriel peut être muni canoniquement d'une orientation. Cependant cela n'est pas toujours le cas. Par exemple si π est un plan vectoriel dans $\mathbb{R}^3$, il n'y a pas sur π de base canonique (c'est-à-dire privilégiée) et donc il n'y a pas d'orientation canonique.
On peut définir une orientation sur π en choisissant une base $\{v_1,\ v_2\}$ de π . Mais on peut le faire aussi d'une autre façon, qui sous certains aspects est plus pratique.
Pour cela, on choisit un vecteur $w \in \mathbb{R}^3 \setminus \{\pi\}$: on dira qu'une base $\mathcal{B} = \{v_1,\ v_2\}$ de π est orientée positivement si la base $\tilde{\mathcal{B}} = \{v_1,\ v_2,\ w\}$ est orientée positivement pour l'orientation canonique de $\mathbb{R}^3$.
Cela fixe effectivement une orientation de π. Montrons, en effet, que, conformément à la définition 4.31, deux bases $\mathcal{B} = \{v_1, v_2\}$, $\mathcal{B}' = \{v_1', v_2'\}$ sont dans la même classe si et seulement si $\text{dét}\, P_{\mathcal{B} \to \mathcal{B}'} > 0$.
Notons $P = P_{\mathcal{B} \to \mathcal{B}'}$ et soient $\tilde{\mathcal{B}} = \{v_1, v_2, w\}$, $\tilde{\mathcal{B}}' = \{v_1', v_2', w\}$; on a :

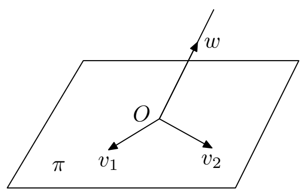


Figure 8

$$\text{dét}\, P_{\tilde{\mathcal{B}} \to \tilde{\mathcal{B}}'} = \text{dét}\, \|v_1', v_2',\ w\|_{\{v_1,, v_2,\ w\}} \ = \ \text{dét} \begin{pmatrix} \begin{array}{cc|c} \multicolumn{2}{c|}{P} & 0 \\ & & 0 \\ \hline 0 & 0 & 1 \end{array} \end{pmatrix} = \text{dét}\, P$$

Ainsi $\mathcal{B}$ et $\mathcal{B}'$ sont dans la même classe (au sens de l'orientation que l'on vient de définir, c'est-à-dire : $\text{dét}\, P_{\tilde{\mathcal{B}} \to \tilde{\mathcal{B}}'} > 0$) si et seulement si elles le sont au sens habituel, c'est-à-dire si $\text{dét}\, P_{\mathcal{B} \to \mathcal{B}'} > 0$.

Nous avons donc orienté π en choisissant un vecteur $w \notin \pi$. Cette définition se généralise immédiatement :

Définition 4.34 – *Soit E un espace vectoriel réel,* orienté, *de dimension n et F un sous-espace vectoriel de dimension $n-1$. Le choix d'un vecteur $w \notin F$, définit sur F une orientation dite* ***orientation induite*** *par w et par l'orientation de E : une base $\mathcal{B} = \{v_1, ..., v_{n-1}\}$ de F sera dite positivement orientée si la base $\tilde{\mathcal{B}} = \{v_1, ..., v_{n-1}, w\}$ de E est positivement orientée.*

Remarque – On pourrait penser qu'il y a une infinité d'orientations induites sur F, c'est-à-dire autant que de choix du vecteur $w \notin F$. Ceci est exclu, évidemment, car sur un espace vectoriel il n'y a que deux orientations possibles.

Il est facile de voir, en effet que l'orientation définie par w *ne dépend que du demi-espace auquel w appartient.*

Plus précisément : soient $w,\ w' \notin F$; puisque

$$E = F \oplus \text{Vect}\{w\}$$

w' s'écrit d'une manière unique

$$w' = x + \lambda\, w \qquad \text{avec } x \in F.$$

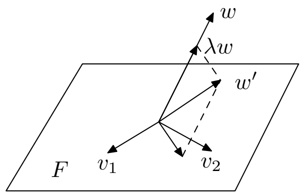


Figure 9

Si $\tilde{\mathcal{B}'}$ est une base de E et $\mathcal{B} = \{v_1, ..., v_{n-1}\}$ est une base de F, on a :

$$\begin{aligned} \text{dét}\,||v_1, ..., v_{n-1}, w'||_{\tilde{\mathcal{B}'}} &= \text{dét}\,||v_1, ..., v_{n-1}, x + \lambda w||_{\tilde{\mathcal{B}'}} = \text{dét}\,||v_1, ..., v_{n-1}, \lambda w||_{\tilde{\mathcal{B}'}} \\ &= \lambda\, \text{dét}\,||v_1, ..., v_{n-1}, w||_{\tilde{\mathcal{B}'}} \end{aligned}$$

ce qui montre que l'orientation définie par w' est la même que celle définie par w si $\lambda > 0$ et est l'opposée si $\lambda < 0$. $\square$

Exercice 24.

EXERCICES

1 Calculer les déterminants suivants :

$$\Delta_1 = \begin{vmatrix} 1 & a & -b \\ -a & 1 & c \\ b & -c & 1 \end{vmatrix} \qquad \Delta_2 = \begin{vmatrix} a & 0 & b & 0 \\ 0 & a & 0 & b \\ c & 0 & d & 0 \\ 0 & c & 0 & d \end{vmatrix} \qquad a,\ b,\ c,\ d \in \mathbb{R}.$$

2 Résoudre l'équation :

$$\begin{vmatrix} 1 & 1 & 1 & 1 \\ x & a & 0 & 0 \\ x & 0 & b & 0 \\ x & 0 & 0 & c \end{vmatrix} = 0 \qquad (a,\ b,\ c,\ d \in \mathbb{R} \text{ non nuls }).$$

3 Montrer que pour $n > 2$ le déterminant suivant est nul :

$$\Delta = \begin{vmatrix} a_1 - b_1 & a_1 - b_2 & \cdots & a_1 - b_n \\ a_2 - b_1 & a_2 - b_2 & \cdots & a_2 - b_n \\ \vdots & \vdots & & \vdots \\ a_n - b_1 & a_n - b_2 & \cdots & a_n - b_n \end{vmatrix}$$

4 Les nombres 136, 221 et 595 sont divisibles par 17. Montrer, sans le calculer, que le déterminant :

$$\begin{vmatrix} 1 & 3 & 6 \\ 2 & 2 & 1 \\ 5 & 9 & 5 \end{vmatrix}$$

est divisible par 17.

5 Pour quelles valeurs de $\lambda,\ \mu \in \mathbb{R}$ les vecteurs $\begin{pmatrix} 2 \\ \mu \\ 1 \end{pmatrix}, \begin{pmatrix} \lambda \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} \mu \\ \lambda \\ 0 \end{pmatrix}$ forment-ils une base de $\mathbb{R}^3$?

6 Calculer la signature des permutations suivantes de $\{1\ 2\ 3\ 4\ 5\}$:

$$\sigma_1 = (2\ 4\ 1\ 3\ 5) \quad , \quad \sigma_2 = (1\ 5\ 2\ 4\ 3) \quad , \quad \sigma_3 = (2\ 3\ 4\ 5\ 1)$$

Calculer $\sigma = \sigma_1 \circ \sigma_2 \circ \sigma_3$ et vérifier que $\varepsilon(\sigma) = \varepsilon(\sigma_1)\, \varepsilon(\sigma_2)\, \varepsilon(\sigma_3)$.

*

7 Soit $\sigma \in \mathcal{S}_n$. On dit que deux éléments $i, j \in \{1, 2, \ldots, n\}$ sont **inversés** dans σ si $i < k$ et $\sigma(i) > \sigma(k)$. Soit :

$$\sharp\, i := \text{nombre des } k \in \{1, 2, \ldots, n\} \text{ tels que : } \quad k < i \quad \text{et} \quad \sigma(k) > \sigma(i)$$

Montrer que : $\varepsilon(\sigma) = (-1)^{\sum_{i=1}^{n} (\sharp\, i)}$.

[Par exemple si $\sigma \in \mathcal{F}_6$: $\sigma = (2\ 4\ 1\ 3\ 6\ 5)$, on a : $\sharp 1 = 0$, $\sharp 2 = 0$, $\sharp 3 = 0$ $\sharp 4 = 2$, $\sharp 5 = 0$, $\sharp 6 = 1$, donc : $\varepsilon(\sigma) = 1$].

*

8 Soit A une matrice de terme général a_{ij}. Montrer que l'on ne change pas la valeur de son déterminant si l'on remplace les éléments a_{ij}, avec $i + j$ impair, par $-a_{ij}$.

9 Soit $A \in \mathcal{M}_n(K)$ antisymétrique. Montrer que si n est impair dét $A = 0$.

10 Calculer les déterminants suivants :

$$\Delta_1 = \begin{vmatrix} 3 & 16 & 24 & 33 \\ 1 & 5 & 7 & 9 \\ 5 & 27 & 36 & 55 \\ 7 & 38 & 51 & 78 \end{vmatrix} \qquad \Delta_2 = \begin{vmatrix} x+2 & 2x+3 & 3x+4 \\ 2x+3 & 3x+4 & 4x+5 \\ 3x+5 & 5x+8 & 10x+17 \end{vmatrix}$$

11 Calculer le déterminant :

$$\Delta_n = \begin{vmatrix} a_n & a_{n-1} & a_{n-2} & a_{n-3} & \ldots & \ldots & a_0 \\ -1 & x & 0 & 0 & \ldots & \ldots & 0 \\ 0 & -1 & x & 0 & \ldots & \ldots & 0 \\ 0 & 0 & -1 & x & \ldots & 0 & 0 \\ \vdots & & & & \ddots & & \\ 0 & 0 & 0 & 0 & \ldots & -1 & x \end{vmatrix}$$

*

12 Calculer le déterminant :

$$\Delta_n = \begin{vmatrix} 1+x^2 & x & 0 & \cdots & \cdots & \cdots & 0 \\ x & 1+x^2 & x & 0 & \cdots & \cdots & 0 \\ 0 & x & 1+x^2 & x & \cdots & \cdots & 0 \\ \vdots & & \ldots & & \ldots & \ldots & \ldots \\ 0 & 0 & \cdots & \cdots & x & 1+x^2 & x \\ 0 & 0 & \cdots & \cdots & 0 & x & 1+x^2 \end{vmatrix}$$

**

13 **Déterminant de Vandermonde**

Montrer que :

$$\begin{vmatrix} 1 & x_1 & (x_1)^2 & \ldots & (x_1)^{n-1} \\ 1 & x_2 & (x_2)^2 & \ldots & (x_2)^{n-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_n & (x_n)^2 & \ldots & (x_n)^{n-1} \end{vmatrix} = \prod_{1 \le i < j \le n} (x_j - x_i)$$

$*$ **14** Soit :

$$D = \begin{array}{c} \begin{array}{cc} \overbrace{}^{k} & \overbrace{}^{n-k} \end{array} \\ \left(\begin{array}{c|c} A & B \\ \hline 0 & C \end{array}\right) \end{array} \begin{array}{l} \}\, k \\ \}\, n-k \end{array} \qquad \begin{array}{ll} \text{avec} & A \in \mathcal{M}_k(K) \\ \text{et} & C \in \mathcal{M}_{n-k}(K) \end{array}$$

Montrer que $\text{dét}\, D = \text{dét}\, A \cdot \text{dét}\, C$.
Généraliser :

$$\text{dét} \begin{pmatrix} \boxed{A_1} & & & * \\ & \boxed{A_2} & & \\ & & \ddots & \\ 0 & & & \boxed{A_p} \end{pmatrix} = \begin{array}{l} \text{dét}\, A_1 \cdot \text{dét}\, A_2 \cdots \text{dét}\, A_p \\ (A_1, \cdots, A_p \text{ étant des matrices carrées}) \end{array}$$

$*$ **15** Montrer que, si $A, B \in \mathcal{M}_n(K)$:

$$\text{dét} \left(\begin{array}{c|c} A & B \\ \hline B & A \end{array}\right) = \text{dét}\,(A+B)\,\text{dét}\,(A-B)$$

$*$ **16** Montrer que, si $A, B \in \mathcal{M}_n(\mathbb{R})$ et $AB = BA$:

$$\text{dét} \left(\begin{array}{c|c} A & -B \\ \hline B & A \end{array}\right) = \text{dét}\,(A^2 + B^2)$$

17 Soient $v_1, \cdots, v_n$ n vecteurs d'un espace vectoriel E de dimension n et $\{e_i\}$, $\{e'_i\}$ deux bases de E. Montrer que :

$$\text{dét}\, \|v_1, \ldots v_n\|_{e_i} = (\text{dét}\, P_{e_i \to e'_i})\ \text{dét}\, \|v_1, \cdots, v_n\|_{e'_i}$$

18 Calculer le déterminant de la matrice qui représente dans la base canonique de $\mathbb{R}^3$ la projection f sur le plan $x - y + 5z = 0$ parallèlement à la droite $x = y = z$ suivie de la projection g sur le plan $x + y + z = 0$ parallèlement à la droite $x = y = -5z$.

19 Pour quelles valeurs de $a \in \mathbb{R}$ la matrice $A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 3 & a \end{pmatrix} \in \mathcal{M}_3\,(\mathbb{R})$ est-elle inversible ? Calculer, dans ce cas, son inverse.

$*$ **20** Soit $A \in \mathcal{M}_n\,(K)$. On note $A^+ = {}^t\text{cof}(A)$. Montrer que si A est inversible, $\text{dét}\, A^+ = (\text{dét}\, A)^{n-1}$. En déduire que si A est inversible, A^+ l'est aussi et que $A^{++} = (\text{dét}\, A)^{n-2}\, A$.
[Pour une généralisation au cas non inversible, cf. exercice 5 , chapitre 6.]

21 Soit la matrice $A = \begin{pmatrix} a & 2 & -1 & b \\ 3 & 0 & 1 & -4 \\ 5 & 4 & -1 & 2 \end{pmatrix}$. Montrer que $\text{rg}\, A \geq 2$.
Par quelles valeurs de a et b a-t-on $\text{rg}\, A = 2$?

$*$ **22** a) Soient $v_1 = (a, b, c)$ et $v_2 = (a', b', c')$ deux vecteurs indépendants de $\mathbb{R}^3$. Ecrire l'expression du plan vectoriel engendré par v_1 et v_2.
b) Même question pour $v_1 = (a, b, c, d)$ et $v_2 = (a', b', c', d')$, vecteurs indépendants de $\mathbb{R}^4$.

23 Montrer que l'aire du triangle de sommets $A = (x_1,\, y_1)$, $B = (x_2,\, y_2)$, $C = (x_3,\, y_3)$ est donnée par :

$$\left| \frac{1}{2} \begin{vmatrix} x_2 - x_1 & x_3 - x_1 \\ y_2 - y_1 & y_3 - y_1 \end{vmatrix} \right|$$

24

1. Soit $\mathcal{O}$ l'orientation de $\mathbb{R}^4$ définie par les vecteurs
$$\{v_1 = (1,1,0,1)\,,\ v_2 = (0,0,1,0)\,,\ v_3 = (0,-1,0,1)\,,\ v_4 = (0,0,0,1)\}$$
Coïncide-t-elle avec l'orientation canonique de $\mathbb{R}^4$?
2. Soit F le sous-espace de $\mathbb{R}^4$ engendré par les vecteurs $\{w_1 = (1,0,2,1),\, w_2 = (2,-1,1,7), w_3 = (1,0,1,1)\}$. Quelle est la dimension de F ? Déterminer un vecteur $w \notin F$ tel que l'orientation induite sur F par w et par $\mathcal{O}$ coïncide avec l'orientation définie par $\{w_1, w_2, w_3\}$.
3. Soit $\mathbb{R}^4$ muni de l'orientation canonique et F l'hyperplan défini par $x + y - z + 2\,t = 0$. On munit F de l'orientation induite par le vecteur $w = (1,1,1,1)$. La base $w_1 = (1,-1,0,0)$, $w_2 = (0,1,1,0)$, $w_3 = (0,0,2,1)$ de F est-elle orientée positivement ?

INDICATIONS

1 $\Delta_1 = 1 + a^2 + b^2 + c^2 \quad , \quad \Delta_2 = (ad - bc)^2$.

2 $\dfrac{1}{x} = \dfrac{1}{a} + \dfrac{1}{b} + \dfrac{1}{c}$.

3 Soient $\vec{a} = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$ et $\vec{v} = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$

$$\Delta = \text{dét}\, \| \, \vec{a} - b_1 \vec{v},\ \vec{a} - b_2 \vec{v}, \ldots, \vec{a} - b_n \vec{v} \, \|$$

Les vecteurs colonnes appartiennent à $\text{Vect}\{\,\vec{a},\ \vec{v}\,\}$ et donc, si $n > 2$, ils sont donc liés.

4 Ajouter à la III$^{\text{ème}}$ colonne la I^{ere} multipliée par 100 et la II$^{\text{ème}}$ multipliée par 10.

5 $\begin{vmatrix} 2 & \lambda & \mu \\ \mu & 0 & \lambda \\ 1 & 1 & 0 \end{vmatrix} = \lambda^2 + \mu^2 - 2\,\lambda \equiv (\lambda - 1)^2 + \mu^2 - 1$

Les vecteurs sont liés pour $\lambda = 1 + \cos\theta,\ \mu = \sin\theta \quad (\theta \in \mathbb{R})$.

6 $\varepsilon\,(\sigma_1) = -\,1\,, \quad \varepsilon\,(\sigma_2) = 1\,; \quad \varepsilon\,(\sigma_3) = 1\,; \quad \sigma = (5\,4\,3\,1\,2)\,, \quad \varepsilon\,(\sigma) = -\,1.$

7 Par récurrence sur n. Pour $n = 1$ il n'y a rien à démontrer.
Supposons la propriété vraie à l'ordre $n-1$ et soit $\sigma \in \mathcal{S}_n$. Soit $n = \sigma\,(k)$, c'est-à-dire :

$$\sigma = \begin{pmatrix} 1 & 2 \cdots & k & \cdots & n \\ * & * \cdots & n & \cdots & * \end{pmatrix}$$

Considérer la permutation σ_1 qui amène n à la dernière place en effectuant successivement des échanges avec les éléments qui le suivent. Montrer que $\varepsilon\,(\sigma_1) = (-\,1)^{\sharp\, n}$. Appliquer l'hypothèse de récurrence à $\sigma_2 = \sigma_1 \circ \sigma$ et utiliser le fait que $\varepsilon\,(\sigma_2) = \varepsilon\,(\sigma_1)\,\varepsilon\,(\sigma)$.

8 D'après la formule de la proposition 4.10, $\det A$ est une somme de termes du type $\varepsilon(j)\, a_{j_1 1}\, a_{j_2 2} \cdots a_{j_n n}$ où $j = (j_1, \ldots, j_n)$ est une permutation de $(1, 2, \ldots, n)$. Etant donné que $j_1 + j_2 + \cdots + j_n = \dfrac{n(n+1)}{2}$ on a : $(j_1+1)+(j_2+2)+\cdots+(j_n+n) = n(n+1)$ qui est un nombre pair. Il y a donc dans $a_{j_1 1}\, a_{j_2 2}, \ldots, a_{j_n n}$ un nombre ***pair*** de couples $(j_k,\, k)$ tel que $j_k + k$ est impair. Conclure.

9 $\det\left({}^tA\right) = \det(-A) = (-1)^n \det A \ldots$

10 $\Delta_1 = -2$.
Pour Δ_2 : soustraire la première ligne de la seconde et la seconde de la troisième. Mettre en facteur $(x+1)(x+2)$. On trouve : $\Delta_2 = -3\,(x+1)^2\,(x+2)$.

11 En développant selon la première colonne, on trouve $\Delta_n = a_n\, x^n + \Delta_{n-1}$; d'où : $\Delta_n = a_n\, x^n + a_{n-1}\, x^{n-1} + \cdots + a_1\, x + a_0$

12 $\Delta_n - (1 + x^2)\,\Delta_{n-1} + x^2\,\Delta_n = 0$. A l'aide des résultats de l'exercice 19 du chapitre 1, on trouve $\Delta = 1 + x^2 + x^4 + \cdots + x^{2n}$.

13 Considérer le déterminant comme un polynôme en les indéterminées $(x_i)_{i=1,\ldots,n}$. En donnant à x_i et x_j $(i \neq j)$ des valeurs égales, le déterminant s'annule (deux lignes égales). Donc $x_i - x_j$ divise le déterminant. Puisque les $x_i - x_j$ sont deux à deux premiers entre eux, le déterminant est divisible par leur produit, et, pour des raisons de degré, il vaut $k \prod\limits_{i<j} (x_j - x_i)$. En comparant les termes en $(x_i)^{n-1}$ on voit que $k = 1$.

14 Remarquer que :

$$D = \begin{pmatrix} I_k & 0 \\ 0 & C \end{pmatrix} \begin{pmatrix} A & B \\ 0 & I_{n-k} \end{pmatrix}$$

15 En ajoutant la (n+i)$^{\text{ème}}$ colonne à la i$^{\text{ème}}$ colonne (pour $i = 1, \ldots, n$) :

$$\det \left(\begin{array}{c|c} A & B \\ \hline B & A \end{array}\right) = \det \left(\begin{array}{c|c} A+B & B \\ \hline B+A & A \end{array}\right)$$

Soustraire la i$^{\text{ème}}$ ligne de la (n+i)$^{\text{ème}}$ ligne (pour $i = 1, \ldots, n$).

16 Comme dans l'exercice précédent :

$$\det \left(\begin{array}{c|c} A & -B \\ \hline B & A \end{array}\right) = \det \left(\begin{array}{c|c} A - i\,B & -B \\ \hline B + i\,A & A \end{array}\right) = \det \left(\begin{array}{c|c} A - i\,B & -B \\ \hline 0 & A + i\,B \end{array}\right)$$

17 Soients : $X_i = M\,(v_i)_{e_i}$, $X'_i = M\,(v_i)_{e'_i}$. On a :
$\det \|v_1, \ldots, v_n\|_{e'_i} = \det \|X'_1, \ldots, X'_n\| = \det \|P^{-1}\, X_1, \ldots, P^{-1}\, X_n\|$
$= \det\left(P^{-1}\, \|X_1, \ldots, X_n\|\right)$ (cf. chapitre 3, exercice 23).

Autre méthode : Si $v_1, ...v_n$ sont liés, les deux membres sont nuls. Si $\{v_1, ..., v_n\}$ est une base, alors, d'après la transitivité des matrices de passage (cf. proposition 3.24, page 77)

$$\text{dét}\,\|v_1, ..., v_n\| = \text{dét}\,P_{e_i \to v_i} = \text{dét}\,(P_{e_i \to e_i{}'} P_{e_i{}' \to v_i}) = \text{dét}\,P_{e_i \to e_i{}'}\ \text{dét}\,\|v_1, ..., v_n\|_{e_i{}'}.$$

18 On a $\text{dét}\, f = 0$, $\text{dét}\, g = 0$, car f et g sont non injectives.
Donc $\text{dét}\,(g \circ f) = (\text{dét}\, g)\,(\text{dét}\, f) = 0$.

19 $a \neq 7 \qquad A^{-1} = \dfrac{1}{a-7} \begin{pmatrix} 2a-12 & 3-a & 2 \\ 4-a & a-1 & -3 \\ 1 & -2 & 1 \end{pmatrix}.$

20 De la relation $A\,A^{+} = (\text{dét}\, A)\, I$ on obtient $(\text{dét}\, A)\,(\text{dét}\, A^{+}) = (\text{dét}\, A)^n$. Ecrire ensuite la relation $A\,A^{+} = (\text{dét}\, A)\, I$ en remplaçant A par A^{+}.

21 Dans $A = \begin{pmatrix} a & 2 & -1 & b \\ \boxed{\begin{matrix} 3 & 0 \\ 5 & 4 \end{matrix}} & & 1 & -4 \\ & & -1 & 2 \end{pmatrix}$ le mineur encadré est non nul, donc $\text{rg}\, A \geq 2$.

Calculer les bordants. On trouve $a = 1$ et $b = 3$.

22 a) $\begin{vmatrix} a & a' & x \\ b & b' & y \\ c & c' & z \end{vmatrix} = 0$. En effet cette équation exprime, d'après le théorème (A), que $\begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \text{Vect}\{v_1, v_2\}$.

b) Utiliser le théorème (B). Si par exemple $\begin{vmatrix} a & a' \\ b & b' \end{vmatrix} \neq 0$, le plan est défini par le système formé des équations :

$$\begin{vmatrix} a & a' & x \\ b & b' & y \\ c & c' & z \end{vmatrix} = 0 \qquad \begin{vmatrix} a & a' & x \\ b & b' & y \\ d & d' & u \end{vmatrix} = 0$$

23 L'aire étant invariante par translation, effectuer une translation de manière à ce que le sommet A soit l'origine.

24

1. On a $\text{dét}\,\|\, v_1, v_2, v_3, v_4 \,\| > 0$. Donc l'orientation définie par ces vecteurs coïncide avec l'orientation canonique.
2. Il suffit de choisir un vecteur $w \notin F$, tel que $\text{dét}\,\|\, w_1, w_2, w_3, w \,\| > 0$. On voit que, par exemple $w = (0, 0, 0, 1)$ convient.
3. On a $\text{dét}\,\|w_1, w_2, w_3, \| = 3$.

Chapitre 5

Systèmes d'équations linéaires

Au chapitre 2. nous avons appris à résoudre les systèmes d'équations linéaires par la méthode du pivot. Cette méthode, même si elle est très simple et très efficace [1], se révèle en général mal aisée lorsque dans le système interviennent des paramètres et surtout elle ne permet pas de savoir a priori si le système admet ou non des solutions : pour le savoir on est obligé d'entamer la résolution du système et voir si l'on obtient des équations du type $0 = a$ avec $a \neq 0$.
Les déterminants fournissent un outil efficace et indispensable pour la discussion des systèmes linéaires : ils permettent d'avoir les conditions de compatibilité sous forme de relations liant les coefficients et fournissent aussi des formules qui donnent explicitement la solution (formules de Cramer).

5.1 Définitions et interprétations

Considérons un système linéaire de p équations en n inconnues :

$$\begin{cases} a_{11}\, x_1 + a_{12}\, x_2 + \cdots\cdots + \ a_{1n}\, x_n = b_1 \\ \ \ \vdots \\ a_{p1}\, x_1 + a_{p2}\, x_2 + \cdots\cdots + \ a_{pn}\, x_n = b_p \end{cases} \tag{1}$$

où les a_{ij} et les b_{ij} appartiennent à un corps K (commutatif).
On appelle ***solution*** tout vecteur $x = (x_1, \ldots, x_n) \in K^n$ dont les composantes x_i satisfont toutes les équations. Le système est dit ***compatible*** s'il admet au moins une solution.

Expression matricielle
Soient :

$$A = (a_{ij}) = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{p1} & \cdots & a_{pn} \end{pmatrix} \in \mathcal{M}_{p,n}(K)$$

[1]Cf. en particulier les considérations que nous développons en Appendice 5.

$$B = \begin{pmatrix} b_1 \\ \vdots \\ b_p \end{pmatrix} \in \mathcal{M}_{p,1}(K) \quad \text{et} \quad X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathcal{M}_{n,1}(K)$$

Le système (1) peut s'écrire sous la forme matricielle :

$$\boxed{A\,X = B} \tag{1'}$$

comme on le voit facilement en effectuant le produit des matrices A et X.
On appelle ***rang du système*** le rang de la matrice A.

Expression vectorielle

Notons $\vec{c}_1, \ldots, \vec{c}_n$ les vecteurs colonnes de la matrice A :

$$\vec{c}_1 = \begin{pmatrix} a_{11} \\ \vdots \\ a_{p1} \end{pmatrix} \in K^p, \ldots\ldots, \vec{c}_n = \begin{pmatrix} a_{1n} \\ \vdots \\ a_{pn} \end{pmatrix} \in K^p$$

On a :

$$x_1\,\vec{c}_1 = \begin{pmatrix} a_{11}\,x_1 \\ \vdots \\ a_{p1}\,x_1 \end{pmatrix}, \ldots\ldots, x_n\,\vec{c}_n = \begin{pmatrix} a_{1n}\,x_n \\ \cdots \\ a_{pn}\,x_n \end{pmatrix}$$

Si donc $\vec{b} = \begin{pmatrix} b_1 \\ \vdots \\ b_p \end{pmatrix} \in K^p$, le système peut s'écrire :

$$\boxed{x_1\,\vec{c}_1 + \cdots + x_n\,\vec{c}_n = \vec{b}} \tag{1''}$$

Résoudre le système signifie déterminer les coefficients de la décomposition du vecteur $\vec{b} \in K^p$ sur les vecteurs $\{\vec{c}_1, \ldots, \vec{c}_n\}$ de K^p : donc, pour que le système soit compatible il faut et il suffit que $\overrightarrow{b}$ appartienne à l'espace engendré par les vecteurs $\vec{c}_1, \ldots, \vec{c}_n$.

5.2 Systèmes de Cramer

Définition 5.1 – *On appelle **système de Cramer** un système linéaire dont la matrice A est* carrée *et* inversible.

Il s'agit donc d'un système de n équations en n inconnues de rang n :

$$\left\{ \begin{array}{l} a_{11}\,x_1 + \cdots + a_{1n}\,x_n = b_1 \\ \quad\vdots \\ a_{n1}\,x_1 + \cdots + a_{nn}\,x_n = b_n \end{array} \right. \qquad \text{avec} \quad \text{dét}\, A \neq 0 \tag{2}$$

Sous forme matricielle, le système s'écrit :

$$A\,X \;=\; B \tag{2'}$$

Comme A est inversible, en multipliant par A^{-1} à gauche, on trouve :

$$X = A^{-1}\,B \tag{3}$$

Réciproquement, $X = A^{-1}\,B$ satisfait l'équation (2'). Aussi :

> ***un système de Cramer admet toujours une et une seule solution*** donnée par (3).

La solution peut être exprimée aussi par les formules de Cramer. Considérons l'interprétation vectorielle du système (1") ; la solution $(x_1, \ldots, x_n)$ est telle que :

$$x_1\,\vec{c}_1 + \cdots + x_n\,\vec{c}_n = \vec{b}$$

Or :

$$\begin{aligned} &\text{dét}\,\|\vec{c}_1, \ldots, \vec{c}_{i-1},\ \vec{b},\ \vec{c}_{i+1}, \ldots, \vec{c}_n\| = \\ &\quad = \text{dét}\,\|\vec{c}_1, \ldots, \vec{c}_{i-1},\ \textstyle\sum_{k=1}^{n} x_k\,\vec{c}_k,\ \vec{c}_{i+1}, \ldots, \vec{c}_n\| \\ &\quad = \textstyle\sum_{k=1}^{n} x_k\ \text{dét}\,\|\vec{c}_1, \ldots, \vec{c}_{i-1},\ \vec{c}_k,\ \vec{c}_{i+1}, \ldots, \vec{c}_n\| \end{aligned}$$

Pour $k \neq i$ les déterminants de cette somme sont nuls (deux colonnes égales). Il reste le terme avec $k = i$, c'est-à-dire x_i dét A. Ainsi :

$$\text{dét}\,\|\vec{c}_1, \ldots, \vec{c}_{i-1},\ \vec{b},\ \vec{c}_{i+1}, \ldots, \vec{c}_n\| \;=\; x_i\ \text{dét}\,A$$

d'où :

$$x_i \;=\; \frac{\text{dét}\,\|\vec{c}_1, \ldots, \vec{c}_{i-1},\ \vec{b},\ \vec{c}_{i+1}, \ldots, \vec{c}_n\|}{\text{dét}\,A}$$

On peut résumer les résultats dans le théorème suivant :

Théorème de Cramer . 5.2 – *Un système de Cramer :*

$$\begin{cases} a_{11}\,x_1 + \cdots + a_{1n}\,x_n = b_1 \\ \quad\vdots \\ a_{n1}\,x_1 + \cdots + a_{nn}\,x_n = b_n \end{cases}$$

$$(\text{avec} \quad A = (a_{ij}) = \|\vec{c}_1, \ldots, \vec{c}_n\|, \quad \text{dét}\,A \neq 0)$$

admet toujours une et une seule solution, quel que soit le vecteur $\vec{b} = (b_1, \ldots, b_n)$, *solution donnée par les formules de Cramer :*

$$\boxed{x_i \;=\; \frac{\text{dét}\,\|\ \vec{c}_1, \ldots, \vec{c}_{i-1},\ \vec{b},\ \vec{c}_{i+1}, \ldots, \vec{c}_n\ \|}{\text{dét}\,A}}$$

Exemple – Soit le système :

$$\begin{cases} 2\,x \;-\; 5\,y \;+\; 2\,z & = 7 \\ \;x \;+\; 2\,y \;-\; 4\,z & = 3 \\ 3\,x \;-\; 4\,y \;-\; 6\,z & = 5 \end{cases}$$

On a :

$$\text{dét}\,A \;=\; \begin{vmatrix} 2 & -5 & 2 \\ 1 & 2 & -4 \\ 3 & -4 & -6 \end{vmatrix} \;=\; -\,46$$

Le système est donc de Cramer.
Les formules de Cramer donnent :

$$x \;=\; -\,\frac{1}{46}\begin{vmatrix} 7 & -5 & 2 \\ 3 & 2 & -4 \\ 5 & -4 & -6 \end{vmatrix} = 5 \quad , \quad y \;=\; -\,\frac{1}{46}\begin{vmatrix} 2 & 7 & 2 \\ 1 & 3 & -4 \\ 3 & 5 & -6 \end{vmatrix} = 1$$

$$z \;=\; -\,\frac{1}{46}\begin{vmatrix} 2 & -5 & 7 \\ 1 & 2 & 3 \\ 3 & -4 & 5 \end{vmatrix} = 1$$

Exercices 1. 2.

5.3 Cas général. Le théorème de Rouché-Fontené

Considérons un système de p équations en n inconnues de rang r :

$$\left\{\begin{array}{l} \boxed{\begin{array}{lllll} a_{11}\,x_1 & + & \cdots & + & a_{1r}\,x_r \\ \;\vdots & & & & \\ a_{r1}\,x_1 & + & \cdots & + & a_{rr}\,x_r \end{array}} \begin{array}{lllll} + & \cdots & + & a_{1n}\,x_n & = \; b_1 \\ & & & & \\ + & \cdots & + & a_{rn}\,x_n & = \; b_r \end{array} \\ \;\;\vdots \\ \;a_{p1}\,x_1 \;+\; \cdots \;+\; a_{pr}\,x_r \;+\; \cdots \;+\; a_{pn}\,x_n \;=\; b_p \end{array}\right. \tag{4}$$

On peut supposer, quitte à changer l'ordre des équations et la numérotation des inconnues que le mineur encadré est non nul.
Sous forme vectorielle, avec les notations du § 1, le système s'écrit :

$$x_1\,\vec{c}_1 + \cdots + x_n\,\vec{c}_n = \vec{b}$$

Le système est compatible si et seulement si $\vec{b} \in \text{Vect}\,\{\vec{c}_1, \ldots, \vec{c}_n\}$. Or, la famille de vecteurs $\{\vec{c}_1, \ldots, \vec{c}_n\}$ a rang r et, puisque

$$\delta \;:=\; \begin{vmatrix} a_{11} & \cdots & a_{1r} \\ \vdots & & \vdots \\ a_{r1} & \cdots & a_{rr} \end{vmatrix} \;\neq\; 0\;,$$

$\{\vec{c}_1, \ldots, \vec{c}_r\}$ est une famille libre et donc une base de $\text{Vect}\,\{\vec{c}_1, \ldots, \vec{c}_n\}$ (cf. théorème (B)). D'après le théorème (C), $\vec{b} \in \text{Vect}\,\{\vec{c}_1, \ldots, \vec{c}_n\} \equiv \text{Vect}\,\{\vec{c}_1, \ldots, \vec{c}_r\}$ si et seulement si tous les bordants de δ sont nuls :

$$\Delta_s = \begin{vmatrix} \boxed{\begin{matrix} a_{11} & \cdots & a_{1r} \\ \vdots & & \vdots \\ a_{r1} & \cdots & a_{rr} \end{matrix}} & \begin{matrix} b_1 \\ \vdots \\ b_r \end{matrix} \\ \begin{matrix} a_{s1} & \cdots & a_{sr} \end{matrix} & b_s \end{vmatrix} = 0, \qquad s = r+1, ..., n$$

Les déterminants Δ_s sont dits ***déterminants caractéristiques***. Leur annulation est donc une condition nécessaire et suffisante pour que le système soit compatible.

Recherche effective de la solution

Considérons le système (4), dans lequel on suppose que la matrice encadrée a un déterminant non nul et supposons réalisée la condition de compatibilité. Soit la matrice :

$$B = \|\vec{c}_1, \ldots, \vec{c}_r, \ldots, \vec{c}_n, \vec{b}\| = \begin{pmatrix} \boxed{\begin{matrix} a_{11} & \cdots & a_{1r} \\ \vdots & & \vdots \\ a_{r1} & \cdots & a_{rr} \end{matrix}} & \begin{matrix} \cdots & a_{1n} & b_1 \\ & \vdots & \vdots \\ \cdots & a_{rn} & b_r \end{matrix} \\ \begin{matrix} \vdots & & \vdots \\ a_{p1} & \cdots & a_{pr} \end{matrix} & \begin{matrix} & \vdots & \vdots \\ \cdots & a_{pn} & b_p \end{matrix} \end{pmatrix}$$

Cette matrice a rang r, car $\vec{c}_{r+1}, \ldots, \vec{c}_n, \vec{b} \in \mathrm{Vect}\{\vec{c}_1, \ldots, \vec{c}_r\}$ et les vecteurs $\vec{c}_1, \ldots, \vec{c}_r$ forment un système libre. Donc les *lignes* de B forment une famille de rang r. Puisque les r premières lignes sont indépendantes (à cause de la présence du mineur encadré) les $p-r$ dernières lignes sont combinaisons linéaires des r premières. Ceci veut dire que dans le système (4) les dernières $p-r$ équations s'obtiennent comme combinaisons linéaires des autres et, par conséquent, *elles peuvent être éliminées*. Aussi le système (4) est équivalent au système :

$$\left\{ \begin{array}{l} \boxed{\begin{array}{lllll} a_{11}\,x_1 & + & \cdots & + & a_{1r}\,x_r \\ \ \vdots & & & & \\ a_{r1}\,x_1 & + & \cdots & + & a_{rr}\,x_r \end{array}} \begin{array}{lllll} + & \cdots & + & a_{1n}\,x_n & = b_1 \\ & & & & \\ + & \cdots & + & a_{rn}\,x_n & = b_r \end{array} \end{array} \right. \tag{4'}$$

Ces équations sont dites ***équations principales*** . Les inconnues qui y interviennent – c'est-à-dire $x_1, \ldots, x_r$ – sont dites ***inconnues principales*** ; les autres ***variables libres***.
Pour calculer la solution, on donne aux variables libres $x_{r+1}, \ldots, x_n$ des valeurs arbitraires :

$$x_{r+1} = \lambda_{r+1}, \ldots\ldots, x_n = \lambda_n \qquad (\lambda_i \in K)$$

Le système (4') s'écrit alors :

$$\left\{ \begin{array}{l} \boxed{\begin{array}{lllll} a_{1\,1}\,x_1 & + & \cdots & + & a_{1\,r}\,x_r \\ \ \vdots & & & & \\ a_{r\,1}\,x_1 & + & \cdots & + & a_{r\,r}\,x_r \end{array}} \begin{array}{l} = b_1 - a_{1\,r+1}\,\lambda_{r+1} - \cdots - a_{1\,n}\,\lambda_n \\ \\ = b_r - a_{r\,r+1}\,\lambda_{r+1} - \cdots - a_{r\,n}\,\lambda_r \end{array} \end{array} \right. \tag{4''}$$

On obtient ainsi un système de Cramer, lequel admet une et une seule solution pour chaque choix de $\lambda_{r+1}, \ldots, \lambda_n$. La solution dépendra donc des $n - r$ constantes que l'on a fixées. On dit que l'on a une ***indétermination d'ordre*** $n - r$.
On peut résumer les résultats dans le théorème suivant :

Théorème de Rouché-Fontené . 5.3 – *Soit :*

$$\left\{\begin{array}{l} \boxed{\begin{array}{lllll} a_{11}\,x_1 & + & \cdots & + & a_{1r}\,x_r \\ \vdots & & & & \\ a_{r1}\,x_1 & + & \cdots & + & a_{rr}\,x_r \end{array}} \begin{array}{lllll} + & \cdots & + & a_{1n}\,x_n & = b_1 \\ & & & & \\ + & \cdots & + & a_{rn}\,x_n & = b_r \end{array} \\ \quad\vdots \\ a_{p1}\,x_1 + \cdots + a_{pr}\,x_r + \cdots + a_{pn}\,x_n = b_p \end{array}\right.$$

un système de p équations en n inconnues, de rang r.
On extrait du système un mineur δ d'ordre r non nul ; quitte à changer la numérotation, on peut supposer que δ est le mineur encadré :

$$\delta = \begin{vmatrix} a_{11} & \cdots & a_{1r} \\ \vdots & & \vdots \\ a_{r1} & \cdots & a_{rr} \end{vmatrix} \neq 0$$

1. *Le système est compatible si et seulement si tous les déterminants caractéristiques associés à δ sont nuls :*

$$\Delta_s = \begin{vmatrix} \boxed{\begin{matrix} a_{11} & \cdots & a_{1r} \\ \vdots & & \vdots \\ a_{r1} & \cdots & a_{rr} \end{matrix}} & \begin{matrix} b_1 \\ \vdots \\ b_r \end{matrix} \\ \begin{matrix} a_{s1} & \cdots & a_{sr} \end{matrix} & b_s \end{vmatrix} = 0 \qquad (s = r+1, ..., n)$$

2. *Si cette condition est réalisée, le système est équivalent au système des équations principales :*

$$\left\{ \boxed{\begin{array}{lllll} a_{1\,1}\,x_1 & + & \cdots & + & a_{1\,r}\,x_r \\ \vdots & & & & \\ a_{r\,1}\,x_1 & + & \cdots & + & a_{r\,r}\,x_r \end{array}} \begin{array}{ll} = & b_1 - a_{1\,r+1}\,x_{r+1} - \cdots - a_{1\,n}\,x_n \\ & \\ = & b_r - a_{r\,r+1}\,x_{r+1} - \cdots - a_{r\,n}\,x_r \end{array} \right.$$

Il admet alors une infinité de solutions dépendantes de $n - r$ paramètres[2]*. Les solutions se calculent en résolvant le système de Cramer obtenu en donnant aux variables libres $x_{r+1}, \ldots, x_n$ des valeurs arbitraires.*

[2]Plus précisément, les solutions forment un espace affine de dimension $n - r$ (cf. Appendice A.8).

Exemple :

Discuter, d'après les valeurs de k, α, β, $\gamma \in \mathbb{R}$ le système :

$$\begin{cases} 2x + y - z = \alpha \\ y + 3z = \beta \\ 2x + ky + 2z = \gamma \end{cases}$$

On étudie d'abord le rang du système. La matrice du système est :

$$A = \begin{pmatrix} \boxed{\begin{matrix} 2 & 1 \\ 0 & 1 \end{matrix}} & \begin{matrix} -1 \\ 3 \end{matrix} \\ \begin{matrix} 2 & k \end{matrix} & 2 \end{pmatrix}$$

Deux cas se présentent :

1. $\boxed{k \neq 2}$. Le système est de Cramer. La solution s'obtient par les formules de Cramer :

$$x = \frac{\begin{vmatrix} \alpha & 1 & -1 \\ \beta & 1 & 3 \\ \gamma & k & 2 \end{vmatrix}}{6\,(2-k)} \quad , \quad y = \frac{\begin{vmatrix} 2 & \alpha & -1 \\ 0 & \beta & 3 \\ 2 & \gamma & 2 \end{vmatrix}}{6\,(2-k)} \quad , \quad z = \frac{\begin{vmatrix} 2 & 1 & \alpha \\ 0 & 1 & \beta \\ 2 & k & \gamma \end{vmatrix}}{6\,(2-k)}$$

   En effectuant les calculs, on trouve :

$$x = \frac{(2-3k)\,\alpha - (k+2)\,\beta + 4\gamma}{6\,(2-k)} , \; y = \frac{\alpha + \beta - \gamma}{2-k} , \; z = \frac{-\alpha + (1-k)\,\beta + \gamma}{3\,(2-k)}.$$

2. $\boxed{k = 2}$. Le système est de rang 2. On étudie alors la compatibilité à l'aide des déterminants caractéristiques. On extrait un mineur d'ordre 2 non nul, par exemple :

$$\delta = \begin{vmatrix} 2 & 1 \\ 0 & 1 \end{vmatrix}$$

   Le seul déterminant caractéristique est :

$$\Delta = \begin{vmatrix} \boxed{\begin{matrix} 2 & 1 \\ 0 & 1 \end{matrix}} & \begin{matrix} \alpha \\ \beta \end{matrix} \\ \begin{matrix} 2 & 2 \end{matrix} & \gamma \end{vmatrix} = 2\,(\gamma - \alpha - \beta)$$

Donc le système est compatible si et seulement si $\gamma = \alpha + \beta$.

Les solutions s'obtiennent alors en résolvant le système des équations principales :

$$\left\{ \boxed{\begin{matrix} 2x + & y \\ & y \end{matrix}} \begin{matrix} - z = \alpha \\ + 3z = \beta \end{matrix} \right.$$

La variable libre est z. En posant $z = \lambda$, on a :

$$x = \frac{\alpha - \beta + 4\lambda}{2} \quad , \quad y = \beta - 3\lambda \quad , \quad z = \lambda$$

En résumant :

1. si $k \neq 2$ le système admet une et une seule solution ;

2. si $k = 2$ deux cas peuvent se présenter :

a) $\gamma \neq \alpha + \beta$: le système n'admet pas de solution ;

b) $\gamma = \alpha + \beta$: le système admet une infinité de solutions dépendant d'un paramètre.

Exercices 3. 4. 5.

5.4 Cas des systèmes homogènes

On appelle ***système homogène*** un système du type :

$$A\,X \,=\, 0$$

c'est-à-dire un système linéaire dont le second membre est nul.
On voit immédiatement qu'un système homogène est toujours compatible : il admet au moins la solution nulle. Dans de nombreux problèmes (cf. par ex. chapitre 6) il est important de savoir s'il existe des solutions non nulles.
Remarquons tout d'abord que *l'ensemble des solutions forme un espace vectoriel.* On le voit immédiatement en considérant l'application : $f : K^n \longrightarrow K^p$ dont la matrice, dans les bases canoniques est A. Le système s'écrit :

$$f\,(x) \,=\, 0$$

et l'ensemble des solutions n'est pas autre chose que le noyau $\operatorname{Ker} f$ de f, qui est un espace vectoriel.
En utilisant le théorème du rang, on a :

$$n = \dim \operatorname{Ker} f + \operatorname{rg} f$$

Donc l'espace des solutions est de dimension $n - r$, où r est le rang du système. On a ainsi :

Proposition 5.4 – *Soit*

$$\left\{\begin{array}{cccccccccc} \boxed{\begin{array}{ccccc} a_{11}\,x_1 & + & \cdots & + & a_{1r}\,x_r \\ \vdots & & & & \\ a_{r1}\,x_1 & + & \cdots & + & a_{rr}\,x_r \end{array}} & \begin{array}{cccc} + & \cdots & + & a_{1n}\,x_n \\ & & & \\ + & \cdots & + & a_{rn}\,x_n \end{array} & \begin{array}{cc} = & 0 \\ & \\ = & 0 \end{array} \\ \begin{array}{ccccc} \vdots & & & & \\ a_{p1}\,x_1 & + & \cdots & + & a_{pr}\,x_r \end{array} & \begin{array}{cccc} & & & \\ + & \cdots & + & a_{pn}\,x_n \end{array} & \begin{array}{cc} & \\ = & 0 \end{array} \end{array}\right.$$

un système linéaire homogène de rang r. On suppose que le mineur encadré est non nul. Puisque les conditions de compatibilité du théorème de Rouché-Fontené sont trivialement satisfaites, le système est équivalent au système des équations principales :

$$\left\{\begin{array}{ccc} \boxed{\begin{array}{ccccc} a_{11}\,x_1 & + & \cdots & + & a_{1r}\,x_r \\ \vdots & & & & \\ a_{r1}\,x_1 & + & \cdots & + & a_{rr}\,x_r \end{array}} & \begin{array}{cccc} + & \cdots & + & a_{1n}\,x_n \\ & & & \\ + & \cdots & + & a_{rn}\,x_n \end{array} & \begin{array}{cc} = & 0 \\ & \\ = & 0 \end{array} \end{array}\right.$$

Les solutions forment un espace vectoriel de dimension $n - r$ (on dit brièvement que : r équations homogènes indépendantes (ou r hyperplans indépendants) dans un espace vectoriel de dimension n définissent un sous-espace vectoriel de dimension $n - r$ *).*

Remarques. –

1. Les solutions d'un système linéaire non homogène $A\,X = B$ avec $B \neq 0$ ne forment pas un espace vectoriel. En effet, si X_1 et X_2 sont deux solutions, on a $A\,X_1 = B$ et $A\,X_2 = B$, d'où $A\,(X_1 + X_2) = 2B \neq B$: donc la somme de deux solutions n'est pas une solution[3]
2. D'après la proposition ci-dessus, un système homogène admet une solution non nulle si et seulement si $n > r$. C'est le cas, par exemple, si $n > p$, c'est-à-dire s'il y a plus d'inconnues que d'équations (cf. aussi théorème 2.2 page 43).

Un cas particulièrement important est celui où la matrice A *est carrée*. Dans ce cas, la condition pour qu'il y ait des solutions non nulles $(r < n)$ est équivalente à ce qu'il ne soit pas de Cramer, c'est-à-dire dét $A = 0$. On a ainsi :

Proposition 5.5 – *Soit* $A\,X = 0$ *un système linéaire homogène où A est une matrice* carrée*,* $A \in \mathcal{M}_n\,(K)$*. Le système admet des solutions non nulles si et seulement si* dét $A = 0$.

Exercices 6. 7. 8. 9. 10.

EXERCICES

1 Résoudre les systèmes suivants :

$$\left\{\begin{array}{rcrcrcl} 3\,x & - & y & + & 2\,z & = & a \\ -x & + & 2\,y & - & 3\,z & = & b \\ x & + & 2\,y & + & z & = & c \end{array}\right. \qquad (a, b, c \in \mathbb{R})$$

$$\left\{\begin{array}{rcrcrcl} (1+i)\,x & + & (1-2\,i)\,y & + & (-1+3\,i)\,z & = & 2+i \\ x & - & 2\,y & + & z & = & 0 \\ i\,x & + & (2-i)\,y & - & 2\,z & = & 0 \end{array}\right.$$

2 Soient $x_1, \ldots, x_{n+1}$ $n+1$ nombres réels deux à deux distincts. Montrer qu'il existe un et un seul polynôme à coefficients réels de degré $\leq n$ qui prend en $x_1, \ldots, x_{n+1}$ des valeurs préfixées.

3 Etudier et éventuellement résoudre les systèmes suivants :

$$\left\{\begin{array}{rcrcrcl} 2\,x & - & 3\,y & + & z & = & 1 \\ x & + & 4\,y & - & 2\,z & = & 0 \\ 8\,x & - & y & - & z & = & 3 \end{array}\right. \qquad \left\{\begin{array}{rcrcrcl} x & - & y & - & z & = & 1 \\ 2\,x & + & 2\,y & - & 3\,z & = & 0 \\ x & + & 2\,y & - & z & = & 0 \\ 4\,x & + & 3\,y & - & 5\,z & = & 1 \end{array}\right.$$

$$\left\{\begin{array}{rcrcrcrcl} x & - & y & + & z & + & t & = & 2 \\ x & + & y & - & z & + & 2\,t & = & 1 \\ & & y & + & 2\,z & & & = & 0 \end{array}\right.$$

4 Discuter suivant les valeurs du paramètre réel α, et éventuellement résoudre, le système suivant :

$$\left\{\begin{array}{rcrcrcl} x & + & y & + & (1-2\,\alpha)\,z & = & 2\,(1+\alpha) \\ (1+\alpha)\,x & - & (1+\alpha)\,y & + & (2+\alpha)\,z & = & 0 \\ 2\,x & - & 2\,\alpha\,y & + & 3\,z & = & 2\,(1+\alpha) \end{array}\right.$$

[3] Elles forment un espace affine (cf. Appendice 3).)

5 Discuter selon les valeurs de a, b, $c \in \mathbb{R}$ (avec $a\,b\,c \neq 0$), le système :

$$\begin{cases} 4\,b\,c\,x - a\,c\,y + 3\,a\,b\,z = a \\ 5\,b\,c\,x + 2\,a\,c\,y + 7\,a\,b\,z = b \\ 3\,b\,c\,x + a\,c\,y + 4\,a\,b\,z = c \end{cases}$$

6 Résoudre le système suivant :

$$\begin{cases} (2-i)\,x + (1+i)\,y + (4+3\,i)\,z = 0 \\ (4+3\,i)\,x + (-1+3\,i)\,y + (-2+11\,i)\,z = 0 \end{cases}$$

7 Déterminer la dimension et une base du sous-espace vectoriel de $\mathbb{R}^5$ défini par les équations :

$$\begin{cases} x - y + 2\,z - t + u = 0 \\ 2\,x + y + z - 2\,t + 2\,u = 0 \\ x + z - t + u = 0 \end{cases}$$

8 Déterminer quatre points alignés M, N, P, Q connaissant les milieux A, B, C, D des segments MN, NP, PQ, QM. Quelle relation doivent satisfaire A, B, C, D pour que le problème admette une solution ?

9 Soit f l'endomorphisme de $\mathbb{R}^3$ qui dans la base canonique est représenté par la matrice :

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

Pour quelles valeurs de $\lambda \in \mathbb{R}$, existe-t-il des vecteurs non nuls $v \in \mathbb{R}^3$ tels que $f\,(v) = \lambda\,v$?

* **10** On suppose que le système :

$$\begin{cases} a\,x + b\,y + c\,z = 0 \\ a'\,x + b'\,y + c'\,z = 0 \end{cases}$$

est de rang 2.
Expliquer pourquoi une base de l'espace des solutions est donnée par $(x,\, y,\, z)$ avec :

$$x = \begin{vmatrix} b & c \\ b' & c' \end{vmatrix} \quad , \quad y = -\begin{vmatrix} a & c \\ a' & c' \end{vmatrix} \quad , \quad z = \begin{vmatrix} a & b \\ a' & b' \end{vmatrix}$$

Généraliser.

INDICATIONS

1 Pour le premier système :

$$x = \frac{8\,a + 5\,b - c}{8} \quad , \quad y = \frac{-\,2\,a + b + 7\,c}{8} \quad , \quad z = \frac{-\,4\,a - 7\,b + 5\,c}{8}$$

Pour le second système : $x = y = z = i$.

2 Soit $P = a_n\,x^n + a_{n-1}\,x^{n-1} + \cdots + a_1\,x + a_0$. Ecrire le système :

$$P\,(x_i) = \alpha_i \qquad i = 1, \ldots, n+1$$

On a un déterminant de Vandermonde (cf. chapitre 4, exercice 13).

3 $x = \dfrac{2\lambda + 4}{11}$, $y = \dfrac{5\lambda - 1}{11}$, $z = \lambda$ pour le premier système.

Solution unique pour le second système : $x = \dfrac{4}{3}$, $y = -\dfrac{1}{3}$, $z = \dfrac{2}{3}$.

Le troisième système est de rang 3 ; solutions : $x = \dfrac{3(\lambda - 1)}{2}$, $y = -\dfrac{(1+\lambda)}{3}$, $z = \dfrac{1+\lambda}{6}$, $t = \lambda$.

4 Si $\alpha(\alpha+1)(\alpha-1) \neq 0$ on a une solution unique :

$$x = \frac{1}{2(\alpha - 1)}, \quad y = \frac{2\alpha + 3}{2(1-\alpha)}, \quad z = \frac{1+\alpha}{1-\alpha}$$

Si $\alpha = 0$ le rang est 2 et le système est compatible :

$$x = 1 - \frac{3}{2}\lambda, \quad y = 1 + \frac{\lambda}{2}, \quad z = \lambda$$

Si $\alpha = -1$ le rang est 2 et le système est compatible :

$$x = \lambda, \quad y = -\lambda, \quad z = 0$$

Si $\alpha = 1$ le système est incompatible.

5 Le déterminant du système est :

$$\begin{vmatrix} 4bc & -ac & 3ab \\ 5bc & 2ac & 7ab \\ 3bc & ac & 4ab \end{vmatrix} = a^2 b^2 c^2 \begin{vmatrix} 4 & -1 & 3 \\ 5 & 2 & 7 \\ 3 & 1 & 4 \end{vmatrix} = 0$$

Le système est de rang 2. La condition de compatibilité est : $13c - a - 7b = 0$.

6 Système de rang 1.

7 Système de rang 2. En prenant la première et la troisième équation :

$$\left\{ \begin{array}{l} \boxed{\begin{array}{ll} x & -y \\ x & \end{array}} \begin{array}{llll} +2z & -t & +u & = 0 \\ +\ z & -t & +u & = 0 \end{array} \end{array} \right.$$

Solutions : $x = -z + t - u$, $y = z - u$ avec z, t, u arbitraires, c'est-à-dire :

$$X = \begin{pmatrix} -z+t-u \\ z-u \\ z \\ t \\ u \end{pmatrix} = z \begin{pmatrix} -1 \\ 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} + t \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} + u \begin{pmatrix} -1 \\ -1 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

Les trois colonnes mises en évidence forment une famille génératrice et libre, donc une base, de l'espace des solutions.

8 Soient x, y, z, t les abscisses de M, N, P, Q et a, b, c, d celles de A, B, C, D. On doit résoudre le système :

$$\begin{cases} x + y = 2a \\ y + z = 2b \\ z + t = 2c \\ t + x = 2d \end{cases}$$

Système de rang 3. La condition de compatibilité est $a + c = b + d$, c'est-à-dire : le milieu de AC coïncide avec le mielieu de BD.

9 Il faut et il suffit que le système $AX = \lambda X$, c'est-à-dire $(A - \lambda I)X = 0$ admette des solutions non nulles. Pour cela, il faut et il suffit que dét $(A - \lambda I) = 0$. On trouve $\lambda = 1$.

10 Si A est une matrice carrée et a_{ij} ses coefficients, on a :

$$a_{i1}\ \text{cof}\,(a_{k1}) + a_{i2}\ \text{cof}\,(a_{k2}) + \cdots + a_{in}\ \text{cof}\,(a_{kn}) \;=\; \delta_{ik}\ \text{dét}\ A$$

(cf. théorème 4.22). La somme à premier membre est donc nulle si $i \neq k$. On considère la matrice :

$$\begin{pmatrix} l & m & n \\ a & b & c \\ a' & b' & c' \end{pmatrix}$$

pour $i = 2$ et $k = 1$, la relation ci-dessus donne :

$$a\ \text{cof}\,(l)\ +\ b\ \text{cof}\,(m)\ +\ c\ \text{cof}\,(n)\ =\ 0$$

pour $i = 3$ et $k = 1$:

$$a'\ \text{cof}\,(l)\ +\ b'\ \text{cof}\,(m)\ +\ c'\ \text{cof}\,(n)\ =\ 0$$

Donc $\text{cof}\,(l),\ \text{cof}\,(m),\ \text{cof}\,(n)$ vérifient le système. L'un deux est non nul car le système est de rang 2. Puisque l'espace des solutions est de dimension 1, $\Big(\text{cof}\,(l),\ \text{cof}\,(m),\ \text{cof}\,(n)\Big)$ est une base de l'espace des solutions.

Généralisation :

Soit le système homogène à n équations et n inconnues :

$$\begin{cases} a_{11}\,x_1 & + & \cdots & + & a_{1n}\,x_n & = & 0 \\ \cdots & \cdots & \cdots & \cdots & & & \\ a_{n-1,\,1}\,x_1 & + & \cdots & + & a_{n-1,\,n}\,x_n & = & 0 \end{cases}$$

supposé de rang $n-1$. Une base de l'espace des solutions est donnée par $\Big(\text{cof}\,(\ell_1), \ldots, \text{cof}\,(\ell_n)\Big)$ dans la matrice :

$$\begin{pmatrix} \ell_1 & \ell_2 & \cdots & \ell_n \\ a_{11} & a_{12} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n-1,\,1} & a_{n-1,\,2} & \cdots & a_{n-1,\,n} \end{pmatrix}$$

Chapitre 6

Réduction des endomorphismes

6.1 Position du problème

Soit E un espace vectoriel sur K et f un endomorphisme de E. Si E est de dimension finie et $\{e_i\}$ est une base de E, on peut construire la matrice qui représente f dans cette base :

$$M(f)_{e_i} \;=\; \| f\,(e_1), \ldots, f\,(e_n) \|_{e_i}.$$

Comme on l'a vu, si l'on change de base, en général la matrice change : si $A = M\,(f)_{e_i}$, $A' = M\,(f)_{e'_i}$ et $P = P_{e_i \longrightarrow e'_i}$ est la matrice de passage de la base $\{e_i\}$ à la base $\{e'_i\}$, on a :

$$A' \;=\; P^{-1}\,A\,P$$

Dans ce chapitre, nous nous proposons de chercher des bases de E dans lesquelles la forme de la matrice est la plus simple possible, c'est-à-dire, par exemple, diagonale ou, éventuellement, triangulaire.

Plus précisément, on dira que f est ***diagonalisable*** s'il existe une base $\{e_i\}$, telle que :

$$M(f)_{e_i} \;=\; \begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & a_{nn} \end{pmatrix}$$

On dira que f est ***trigonalisable*** s'il existe une base $\{e_i\}$, telle que :

$$M(f)_{e_i} = \begin{pmatrix} a_{11} & & * & \\ 0 & a_{22} & & \\ \vdots & \ddots & \ddots & \\ 0 & \cdots & 0 & a_{nn} \end{pmatrix} \quad \text{ou} \quad M(f)_{e_i} = \begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ & a_{22} & \ddots & \vdots \\ & & \ddots & 0 \\ * & & & a_{nn} \end{pmatrix}$$

Ce problème est en tout point semblable à celui qui se pose, par exemple, en mécanique ou en physique lorsqu'on essaie de déterminer le repère dans lequel le problème a la forme la plus simple et la plus pratique. Par exemple, si l'on étudie la chute libre d'un corps il est naturel de prendre un repère dont l'un des axes est vertical, de manière à ce que le vecteur accélération n'ait qu'une composante non nulle. Il serait contraire au bon sens de prendre les axes obliques selon lesquels on serait obligé de décomposer le vecteur accélération. Pour les endomorphismes, c'est le même type de problème : il s'agit de déterminer les bases dans lesquelles la matrice de f (c'est-à-dire, au fond, l'ensemble des coordonnées de f) a *le plus de zéros possible.*

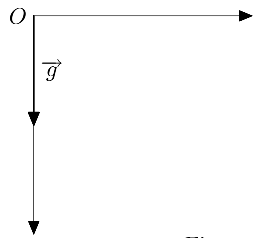


Figure 1

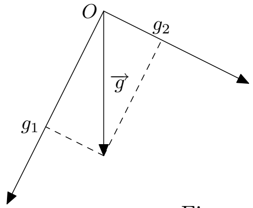


Figure 2

Le problème qui nous occupe est double :

1. Caractériser les endomorphismes diagonalisables (ou trigonalisables).
2. Déterminer effectivement, si elles existent, les bases dans lesquelles la matrice est diagonale ou triangulaire.

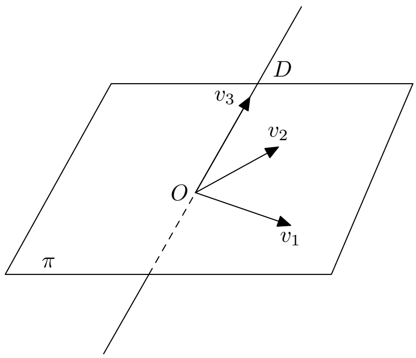
En termes de matrices, puisque deux matrices A et A' liées par la relation $A' = P^{-1}A\,P$ représentent le même endomorphisme en des bases différentes, le problème s'énonce ainsi :

1. Caractériser les matrices $A \in \mathcal{M}_n(K)$ pour lesquelles il existe $P \in \mathcal{M}_n(K)$ inversible, telles que $A' = P^{-1}A\,P$ soit diagonale (respectivement : triangulaire).
2. Déterminer effectivement P et A'.

Dans certains cas, le problème peut se résoudre assez facilement par des considérations géométriques. Soit, par exemple, f la projection de $\mathbb{R}^3$ sur le plan d'équation $x+2\,y+3\,z=0$ parallèlement à la droite D d'équation $\dfrac{x}{3}=\dfrac{y}{2}=z$. Dans la base canonique $\{e_i\}$, on a :

$$M(f)_{e_i} = \frac{1}{10}\begin{pmatrix} 7 & -6 & -9 \\ -2 & 6 & -6 \\ -1 & -2 & 7 \end{pmatrix}$$

(cf. chapitre 3, exercice 19).



En revanche dans la base $\{v_1, v_2, v_3\}$ où $\{v_1, v_2\} \in \pi$ et $v_3 \in D$, on a : $f(v_1)=v_1,\ f(v_2)=v_2,\ f(v_3)=0$, donc :

$$M(f)_{v_i} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

f est donc diagonalisable. De même, si l'on considère la symétrie g par rapport au plan π, parallèlement à la droite D, on a : $g(v_1) = v_1$, $g(v_2) = v_2$, $g(v_3) = -v_3$; donc :

$$M(g)_{v_i} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

Il est clair, cependant, que l'on ne peut pas espérer de résoudre toujours géométriquement ce type de questions. Dans les paragraphes qui suivent, nous allons donner la théorie et les méthodes de calcul qui fournissent la solution explicite du problème.

NOTA. – Dans tout ce chapitre, nous supposerons que E est un espace vectoriel de dimension finie sur K. Certaines définitions, cependant, restent valables en dimension infinie. Nous ne manquerons pas de le souligner et de donner des exemples.
Le lecteur qui n'est pas familiarisé avec les corps abstraits, pourra toujours continuer à supposer $K = \mathbb{R}$ ou $\mathbb{C}$. Cependant, la différence entre $\mathbb{R}$ et $\mathbb{C}$ est importante, comme nous le verrons.

6.2 Vecteurs propres

La clé de la diagonalisation est la notion de vecteur propre (dont la définition est valable en dimension infinie).

Définition 6.1 – *Soit $f \in \operatorname{End}_K(E)$. Un vecteur $v \in E$ est dit* **vecteur propre** *de f si :*

1. $v \neq 0$.
2. $\exists \lambda \in K : f(v) = \lambda\, v$.

Le scalaire λ est dit **valeur propre** *correspondante à v.*

REMARQUES. –

1. Alors que les vecteurs propres sont non nuls par définition[1], la valeur propre peut être nulle : les vecteurs (non nuls) de $\operatorname{Ker} f$ sont justement les vecteurs propres correspondant à $\lambda = 0$.
2. Si v est un vecteur propre correspondant à la valeur propre λ, alors pour tout $\mu \in K$, $\mu \neq 0$, μv est aussi vecteur propre correspondant à la *même* valeur propre λ. En effet :
$$f(\mu\, v) = \mu\, f(v) = \mu\,(\lambda\, v) = \lambda\,(\mu\, v)$$

Ainsi les vecteurs propres sont :
- soit les vecteurs du noyau ;
- soit les vecteurs qui ne changent pas de direction sous l'action de f.

En particulier, la droite vectorielle D engendrée par un vecteur propre de f est invariante par f, c'est-à-dire : $f(D) \subset D$.

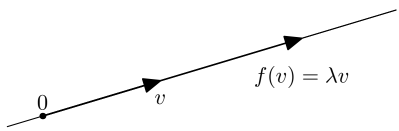


Figure 4

[1] La raison de ce choix est que les vecteurs propres sont utilisés pour construire des bases.

Exemple 1 –

Soit $E = \mathbb{R}^3$ et f la projection sur un plan π parallèlement à une droite D.
Pour tout vecteur $v \neq 0$ du plan, on a $f(v) = v$ et pour tout vecteur directeur w de la droite, on a $f(w) = 0$.
Ainsi, les vecteurs non nuls du plan sont des vecteurs propres correspondants à la valeur propre $+1$ et les vecteurs directeurs de la droite sont des vecteurs propres correspondants à la valeur propre $\lambda = 0$.
Ce sont là les seuls vecteurs propres – et donc aussi les seules valeurs propres – car tout autre vecteur change de direction sous l'action de f.

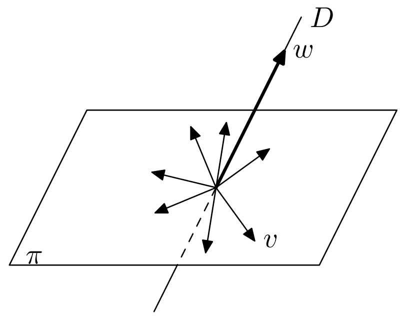


Figure 5

Exemple 2 –

Soit $f : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ la rotation d'angle θ et de centre 0.
Il est clair si $\theta \neq k\,\pi$ il n'y a pas de directions invariantes sous l'action de f, pas plus que de vecteurs du noyau (car f est bijective). Dans ce cas, il n'y a pas de vecteurs propres et donc pas de valeurs propres.

Exemple 3 –

Soit $k \in K$ et $h_k : \underset{x}{E} \underset{\longmapsto}{\longrightarrow} \underset{k\,x}{E}$ l'homothétie de rapport k.
Tout vecteur non nul de E est vecteur propre correspondant à la valeur propre k.

L'intérêt des vecteurs propres réside dans la propriété suivante :

Théorème (I) 6.2 – *$f \in \mathrm{End}_K(E)$ est diagonalisable si et seulement si il existe une base de E formée de vecteurs propres.*

La démonstration est immédiate[2]. Si $\{v_1, \ldots, v_n\}$ est une base formée de vecteurs propres correspondants aux valeurs propres $\lambda_1, \ldots, \lambda_n$, on a :

$$f\,(v_1) = \lambda_1\, v_1, \quad f\,(v_2) = \lambda_2\, v_2, \ldots\ldots\ldots, f\,(v_n) = \lambda_n\, v_n$$

ainsi :

$$M(f)_{v_i} \;=\; \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{pmatrix}$$

[2]Ce n'est en fait qu'une reformulation de la définition de vecteur propre. Si nous l'appelons «théorème», c'est pour en souligner l'importance : dans la pratique il faut toujours avoir présent que *diagonaliser c'est chercher une base de vecteurs propres.*

et donc f est diagonalisable.
Réciproquement, s'il existe une base $\{e_i\}$ telle que $M(f)_{e_i}$ est diagonale :

$$M(f)_{e_i} = \begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & a_{nn} \end{pmatrix}$$

En regardant les colonnes de la matrice, on voit que $f(e_1) = a_{11}e_1$, $f(e_2) = a_{22}e_2$, $\ldots$, $f(e_n) = a_{nn}\, e_n$, ce qui signifie que les vecteurs e_i sont des vecteurs propres. □
Ainsi, par exemple, les endomorphismes des exemples 1 et 3 sont diagonalisables, car dans ce cas on peut construire une base formée de vecteurs propres.
Remarquons que sur la diagonale principale de la matrice diagonale apparaissent justement les valeurs propres de l'endomorphisme.

Exercice 1. 2. 3.

6.3 Recherche des valeurs propres. Polynôme caractéristique

Soit λ une valeur propre de f ; il existe donc un vecteur $v \in E$, $v \neq 0$, tel que $f(v) = \lambda\, v$, c'est-à-dire $(f - \lambda \text{ id})\, v = 0$. Comme $v \neq 0$ cela signifie que l'endomorphisme $(f - \lambda \text{ id})$ n'est pas injectif, ce qui, en dimension finie, équivaut à :

$$\text{dét } (f - \lambda \text{ id}) = 0$$

Aussi, si $\{e_i\}$ est une base de E et :

$$M\,(f)_{e_i} = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}$$

la condition pour que λ soit une valeur propre s'écrit :

$$\begin{vmatrix} a_{11} - \lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} - \lambda & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} - \lambda \end{vmatrix} = 0.$$

En développant ce déterminant, on trouve une équation du type :

$$(-1)^n\, \lambda^n + a_{n-1}\, \lambda^{n-1} + \cdots + a_1\, \lambda + a_0 = 0$$

dont les racines (dans K) sont les valeurs propres de f.
Cette équation est dite *équation caractéristique* et le polynôme à premier membre, appelé polynôme caractéristique de f, est noté $P_f\,(\lambda)$. Nous avons ainsi démontré :

Proposition 6.3 – *Soit $f \in \text{End}_K(E)$, où E est un espace vectoriel de dimension finie n. Les valeurs propres de f sont les racines du polynôme :*

$$P_f\,(\lambda) := \text{ dét } (f - \lambda \text{ id}).$$

*$P_f\,(\lambda)$ est un polynôme de degré n en λ, appelé **polynôme caractéristique** de f.*

Exemple :
Soit $f : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ l'endomorphisme qui, dans la base canonique, est représenté par la matrice :

$$A = \begin{pmatrix} 1 & 2 \\ -1 & 4 \end{pmatrix}$$

On a :

$$P_f(\lambda) \ = \ \text{dét}\ (f - \lambda \ \text{id}) \ = \ \begin{vmatrix} 1-\lambda & 2 \\ -1 & 4-\lambda \end{vmatrix} \ = \ \lambda^2 \ - \ 5\,\lambda \ + \ 6 \ = \ (\lambda - 2)\,(\lambda - 3).$$

Donc, les valeurs propres de f sont $\lambda_1 = 2$ et $\lambda_2 = 3$.

Si $A = M\,(f)_{e_i}$, $P_f(\lambda) = \text{dét}\ (A - \lambda\, I)$ sera noté aussi $P_A\,(\lambda)$. Il est clair, cependant, que le polynôme caractéristique dépend de f et non pas de la base $\{e_i\}$ choisie, car le déterminant d'un endomorphisme ne dépend pas du choix de la base dans laquelle on le calcule (cf. définition 4.21 page 120).

L'ensemble des valeurs propres de f est dit ***spectre*** de f et est noté $\text{Sp}_K(f)$ (ou $\text{Sp}_K A$, si A est la matrice qui représente f dans une base).

Notons que $\lambda = 0$ est valeur propre de f si et seulement si $\text{dét}\, f = 0$, car $P_f(0) = \text{dét}\, f$.

REMARQUE –L'indice K dans la notation $\text{Sp}_K(A)$ est nécessaire pour la raison suivante. Considérons, par exemple, la matrice :

$$A \ = \ \begin{pmatrix} 2 & 1 \\ -5 & -2 \end{pmatrix}$$

$$\text{On a :} \qquad P_A(\lambda) \ = \ \begin{vmatrix} 2-\lambda & 1 \\ -5 & -2-\lambda \end{vmatrix} \ = \ \lambda^2 \ + \ 1$$

Si donc on considère A comme matrice de $\mathcal{M}_2\,(\mathbb{C})$, elle a deux valeurs propres : $+\,i$ et $-\,i$. Si, en revanche, A est considérée comme une matrice de $\mathcal{M}_2\,(\mathbb{R})$ elle n'a pas de valeurs propres. En d'autres termes :

$$\text{Sp}_{\,\mathbb{C}}\, A \ = \ \{-\,i, \ +\,i\,\} \quad , \quad \text{Sp}_{\,\mathbb{R}}\, A \ = \ \emptyset$$

Cependant, si le contexte est clair, on notera aussi $\text{Sp}\, f$ (ou $\text{Sp}\, A$).

Exercice 4. 5.

6.4 Digression sur les polynômes

Par la suite il sera souvent question de polynômes. Dans ce paragraphe, nous allons rappeler brièvement les définitions et les résultats dont nous nous servirons dans ce chapitre et dans les chapitres suivants. La théorie des polynômes est présentée avec plus de détails dans l'Appendice A.2.

Afin de nous placer dans le cadre le plus général englobant le cas où le corps K est fini, nous utiliserons la notation $K\,[X]$ (polynômes formels) au lieu de $K\,[x]$ (fonctions polynômes) (cf. Appendice A.2). Le lecteur qui n'est pas familiarisé avec ces notions peut, sans aucun inconvénient, continuer à supposer que $K = \mathbb{R}$ ou $\mathbb{C}$ et considérer les fonctions polynômes $K\,[x]$. [3]

[3]Dans le cas où le corps K contient une infinité d'éléments les deux notions (polynôme formel et fonction polynôme) coïncident, à un isomorphisme près : cf Appendice A.2.

1. Théorème de la division euclidienne

Soient A, $B \in K[X]$, $B \neq 0$. Il existe alors un et un seul couple de polynômes (Q, R) tels que :

$$A(X) = B(X)\,Q(X) + R(X) \quad \textit{avec} \quad d^\circ R < d^\circ B$$

où $d^\circ P$ désigne le degré de P.[4] *Q est dit* ***quotient*** *et R* ***reste*** *de la division euclidienne de A par B.*
Si le reste de la division euclidienne de A par B est nul on dit que B *divise* A (et on note : $B|A$).

2. Racines

Soit $P \in K[X]$; $a \in K$ est dite ***racine*** *de P si $P(a) = 0$* [$P(a)$ désignant la valeur en a de la fonction polynôme associée à P].

On a :

$$a \text{ est racine de } P \iff (X - a)\,|\,P$$

En effet, si on effectue la division euclidienne de P par $X - a$, on a :

$$P(X) = (X - a)\,Q(X) + R(X) \quad \text{avec} \quad d^\circ R < 1$$

donc R est un élément de K. En faisant $X = a$, on a immédiatement : $P(a) = 0 \iff R = 0$.

Donc a est racine de P si et seulement si on peut écrire :

$$P(X) = (X - a)\,Q(X) \quad \text{avec} \quad Q(X) \in K[X].$$

3. Racines multiples

Supposons que a est racine de P ; on peut donc écrire $P(X) = (X - a)Q_1(X)$. Si a est racine de Q_1, on peut mettre à facteur $X - a$ dans Q_1 et écrire $P(X) = (X - a)^2 Q_2(X)$. Si $Q_2(a) = 0$ on peut encore mettre $X - a$ à facteur et ainsi de suite, jusqu'à aboutir à une expression du type :

$$P(X) = (X - a)^k\, Q(X) \quad \text{avec} \quad Q(a) \neq 0.$$

On dit alors $a \in K$ est ***racine d'ordre*** **k** de P (en d'autres termes, $(X - a)^k$ divise P et $(X - a)^{k+1}$ ne divise pas P.)

Si $a_1, \ldots, a_p$ sont des racines deux à deux distinctes respectivement d'ordre $\alpha_1, \ldots, \alpha_p$, on peut écrire :

$$\boxed{\begin{array}{c} P(X) = (X - a_1)^{\alpha_1}\,(X - a_2)^{\alpha_2}, \ldots, (X - a_p)^{\alpha_p}\,Q(X) \\ \text{avec } Q(X) \in K[X]\text{ , } Q \text{ n'ayant pas de racine dans } K \end{array}} \qquad \text{(i)}$$

On voit de (i) que ***un polynôme de degré*** **n** ***admet au plus*** **n** ***racines*** (en comptant chaque racine autant de fois que son ordre de multiplicité).

4. Polynômes scindés. Factorisation des polynômes de $\mathbb{C}[X]$.

La terminologie de la définition suivante sera employée souvent par la suite :

Définition 6.4 – *Soit $P \in K[X]$ de degré n. On dit que P est* ***scindé dans*** **K** *si P admet n racines dans K (en comptant chaque racine avec sa multiplicité).*

[4]Par définition, on pose $d^0 0 = -\infty$.

Ainsi, par exemple :

$$P(X) \;=\; X^2 \;-\; 5X \;+\; 6 \;=\; (X-2)\,(X-3) \quad \text{est scindé dans } \mathbb{R}$$

De même :

$$P(X) \;=\; X^3 \;-\; 4X^2 \;+\; 5X \;-\; 2 \quad \text{est scindé dans } \mathbb{R},$$

car $P(X)$ s'écrit : $P(X) = (X-1)^2\,(X-2)$ et donc il a trois racines $\{1,\,1,\,2\}$.
En revanche $P(X) = X^2+1$ est scindé dans $\mathbb{C}$ mais pas dans $\mathbb{R}$.

En d'autres termes, un polynôme est scindé si et seulement si il peut s'écrire sous la forme

$$\boxed{\begin{array}{c} P(X) \;=\; a\,(X-a_1)^{\alpha_1}\,\cdots\,(X-a_p)^{\alpha_p} \\ \text{avec} \quad a \in K\,,\;\; a_i \neq a_j \;\; \text{pour} \;\; i \neq j \\ \text{et} \quad \alpha_1 \;+\; \cdots \;+\; \alpha_p \;=\; d^\circ\, P \end{array}} \qquad \text{(ii)}$$

Théorème de D'Alembert . 6.5 – *Tout polynôme de* $\mathbb{C}\,[X]$ *est scindé dans* $\mathbb{C}$.

Par conséquent, tout polynôme de $\mathbb{C}[X]$ se factorise sous la forme (ii).

5. Factorisation des polynômes de $\mathbb{R}[X]$.

Un polynôme P de $\mathbb{R}[X]$ peut être considéré comme un polynôme de $\mathbb{C}[X]$ et à ce titre il est scindé (dans $\mathbb{C}$) : il admettra donc n racines ($n = d^\circ P$) égales ou distinctes, réelles ou complexes.
0n voit facilement que :

$$\begin{array}{l} \text{si } \lambda \text{ est racine de } P \\ \text{d'ordre de multiplicité } \alpha \end{array} \quad \Longrightarrow \quad \begin{array}{l} \overline{\lambda} \text{ est racine de } P \\ \text{d'ordre de multiplicité } \alpha \end{array}$$

Si donc $a_1, \ldots, a_p$ sont les racines réelles de P de multiplicté $\alpha_1, \ldots, \alpha_p$ et $b_1, \ldots, b_q$, $\overline{b}_1, \ldots, \overline{b}_q$ les racines complexes de multiplicité $\beta_1, \ldots, \beta_q$, on a :

$$P = k\,(X-a_1)^{\alpha_1} \cdots (X-a_p)^{\alpha_p}(X-b_1)^{\beta_1}(X-\overline{b}_1)^{\beta_1} \cdots (X-b_q)^{\beta_q}(X-\overline{b}_q)^{\beta_q}$$

$$\text{or :} \quad (X-\lambda)(X-\overline{\lambda}) = X^2 + r\,X + s\,, \quad \text{avec} \quad r, s \in \mathbb{R},\; r^2 - 4\,s < 0.$$

On a donc :

Proposition 6.6 – *Soit* $P \in \mathbb{R}[X]$. *Alors* P *s'écrit d'une manière unique (à l'ordre des facteurs près) :*

$$\boxed{\begin{array}{c} P = k\,(X-a_1)^{\alpha_1} \cdots (X-a_p)^{\alpha_p}(X^2 + r_1 X + s_1)^{\beta_1} \cdots (X^2 + r_q\,X + s_q)^{\beta_q} \\ \text{avec :} \quad k,\, a_i,\, r_i,\, s_i \in \mathbb{R}\;, \quad r_i^2 - 4\,s_i < 0 \\ \text{et} \quad \alpha_1 + \cdots + \alpha_p + 2\,(\beta_1 + \cdots \beta_q) = n \quad (n = d^0 P) \end{array}} \qquad \text{(iii)}$$

6. Clôture algébrique. Sp et Sp'.

On peut montrer que pour tout corps K il existe un corps $K' \supset K$ tel que tout polynôme P de degré n à coefficients dans K' admette exactement n racines dans K' (c'est-à-dire P est scindé dans K'). En particulier, puisque $K \subset K'$, tout polynôme de degré n de $K[X]$ admet exactement n racines ***dans*** $\mathbf{K}'$. K' est dit ***clôture algébrique*** de K. (Ainsi $\mathbb{C}$ est la clôture algébrique de $\mathbb{R}$).
Tout polynôme $P \in K[X]$ peut donc s'écrire sous la forme (ii) à condition de considérer les racines $a_1, \cdots, a_p$ éventuellement dans K'. Ainsi, par exemple, $P = X^2+1$ s'écrit $P = (X-i)(X+i)$.

En appliquant ces considérations au cas du polynôme caractéristique d'un endomorphisme f, on voit immédiatement que si $\dim_K E = n$, alors f admet au plus n valeurs propres (si E est de dimension infinie le spectre de f peut être infini).
Si $P_f(X)$ est *scindé dans* K, il s'écrit :

$$P_f(X) = (-1)^n (X - \lambda_1)^{\alpha_1} \cdots (X - \lambda_p)^{\alpha_p}$$

où $\lambda_1, \cdots, \lambda_p \in K$ sont les valeurs propres deux à deux distinctes de f de multiplicité respectivement $\alpha_1, \cdots, \alpha_\rho$.

Dans la suite nous utiliserons la notation suivante : soit f un endomorphisme d'un espace vectoriel de dimension n sur K et $\lambda_1, ..., \lambda_n$ ses n valeurs propres, égales ou distinctes, éventuellement sur la clôture algébrique K' de K, [5] on pose

$$\mathrm{Sp}' f = \{ \lambda_1, ..., \lambda_n \}$$

Par exemple, si $P_f(X) = (-1)^n \, (X-2)^2 \, (X-4)^3$, on a :

$$\mathrm{Sp} f = \{2, 4\} \qquad \text{et} \qquad \mathrm{Sp}' f = \{ 2, 2, 4, 4, 4 \}$$

NOTA - La notation $\mathrm{Sp}' f$ n'est pas standard, mais nous ne nous gênerons par pour l'utiliser.

7. Polynômes premiers entre eux

Les polynômes $P_1, \ldots, P_r \in K[X]$ sont dits ***premiers entre eux*** [6] si les seuls polynômes qui divisent simultanément $P_1, \cdots, P_r$ sont les polynômes de degré 0 (ou «polynômes constants»).

Par exemple $P_1 = (X-2)\,(X-3)$, $P_2 = (X-1)\,(X-3)$ et $P_3 = X - 2$ sont premiers entre eux.

Il est clair que *si* $P_1, \cdots, P_r$ *sont scindés*, ils sont premiers entre eux si et seulement si il n'existe pas de racine qui soit commune à tous les P_i. En revanche, par exemple, les polynômes de $\mathbb{R}[X]$: $P_1 = (X-2)\,(X^2+1)$, et $P_2 = (X-3)\,(X^2+1)$ ne sont pas premiers entre eux (X^2+1 est diviseur commun) bien qu'ils n'aient pas de racine commune (dans $\mathbb{R}$!)

Théorème de Bezout . 6.7 –
Les polynômes $P_1, \cdots, P_r \in K[X]$ sont premiers entre eux si et seulement s'il existe des polynômes $U_1, \cdots, U_r \in K[X]$ tels que

$$P_1(X)\, U_1(X) + \cdots + P_r(X)\, U_r(X) = 1$$

On dit que les polynômes $P_1, \ldots, P_r$ sont ***deux à deux premiers entre eux*** si pour chaque $i \neq j$, $\{P_i, \, P_j\}$ sont premiers entre eux.

Par exemple si $P = a\,(X-a_1)^{\alpha_1} \cdots (X-a_r)^{\alpha_r}$ est un polynôme scindé, avec $a_i \neq a_j$ pour $i \neq j$, les polynômes $(X-a_1)^{\alpha_1}, \cdots, (X-a_r)^{\alpha_r}$ sont deux à deux premiers entre eux.

6.5 Recherche des vecteurs propres

Une fois calculées les valeurs propres on détermine les vecteurs propres en résolvant, dans le cas où la dimension est finie, un système linéaire : le système

$$(A - \lambda\, I)\, v = 0.$$

[5] Pour le lecteur qui n'est pas familiarisé avec les corps quelconques, considérer $K = \mathbb{R}$ ou $\mathbb{C}$ et $K' = \mathbb{C}$, et les n racines du polynôme caractéristique, $\lambda_1, ..., \lambda_n$, *égales ou distinctes*, ***réelles*** ou ***complexes***.

[6] On dit parfois *premiers entre eux dans leur ensemble*.

Exemple :

Reprenons l'exemple du § 3 : f est l'endomorphisme de $\mathbb{R}^3$ qui dans la base canonique est défini par :

$$A = \begin{pmatrix} 1 & 2 \\ -1 & 4 \end{pmatrix}$$

Les valeurs propres sont $\lambda_1 = 2$ et $\lambda_2 = 3$. Il existe donc deux vecteurs propres v_1, v_2 tels que :

$$f(v_1) = \lambda_1 v_1 \quad \text{et} \quad f(v_2) = \lambda_2 v_2$$

Calcul de v_1**.**

Il faut résoudre l'équation $f(v_1) = \lambda_1 v_1$, c'est-à-dire : $(f - \lambda\ \mathrm{id})\, v_1 = 0$. En écrivant cette équation dans la base canonique et en notant $v_1 = \begin{pmatrix} x \\ y \end{pmatrix}$ la matrice du vecteur v_1,

$$(A - 2\,I)\, v_1 = 0 \quad \text{équivaut à :} \quad \begin{pmatrix} -1 & 2 \\ -1 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

ce qui donne le système :

$$\begin{cases} -x + 2\,y & = 0 \\ -x + 2\,y & = 0 \end{cases}$$

dont les solutions sont engendrées par le vecteur $v_1 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$ (ou par tout autre vecteur non nul colinéaire à v_1).

Calcul de v_2**.**

Il faut résoudre $(f - 3\ \mathrm{id})\, v_2 = 0$. En posant $v_2 = \begin{pmatrix} x \\ y \end{pmatrix}$, on a, dans la base canonique :

$$(A - 3I)v_2 = 0\ , \quad \text{c'est-à-dire :} \begin{cases} -2\,x + 2\,y & = 0 \\ -2\,x + 2\,y & = 0 \end{cases}$$

et la solution est engendrée par $v_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$.

On a donc deux vecteurs propres $v_1 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$ et $v_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$. On voit immédiatement qu'ils forment une base de $\mathbb{R}^2$, car dét $\|v_1,\, v_2\| = \begin{vmatrix} 2 & 1 \\ 1 & 1 \end{vmatrix} \neq 0$. Par conséquent f est diagonalisable et

$$M(f)_{v_i} = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}$$

La matrice de passage $P_{e_i \to v_i}$ est la matrice :

$$P = \|v_1, v_2\|_{e_i} = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$$

On vérifie facilement que : $A' = P^{-1}AP$, où $A' = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}$.

Remarque. − Pour chercher les vecteurs propres il faut résoudre le système $(A - \lambda\, I)v = 0$. Puisque dét $(A - \lambda\, I) = 0$ (à cause du fait que λ est valeur propre) l'une au moins des lignes de la matrice $A - \lambda I$ est combinaison linéaire des autres lignes (cf. proposition 4.14). Par conséquent on obtiendra toujours un système linéaire homogène dans lequel une équation au moins est combinaison linéaire des autres, et donc elle peut être éliminée.

Notons enfin la propriété suivante :

Soit $A \in \mathcal{M}_n(K)$ *et* $\mathrm{Sp}'A = \{\lambda_1, ..., \lambda_n\}$ [7]*, alors :*

$$\boxed{\begin{array}{ll} \mathrm{Tr}\, A & = \lambda_1 + \cdots + \lambda_n \\ \text{dét}\, A & = \lambda_1 \cdot \cdots \cdot \lambda_n \end{array}}$$

Cela est clair pour ce qui est du déterminant, car

$$P_A(X) = \text{dét}\,(A - X\,I) = (-1)^n (X - \lambda_1) \cdots (X - \lambda_n).$$

En faisant $X = 0$ on trouve immédiatement $\text{dét}\, A = \lambda_1 \cdots \lambda_n$.
Pour ce qui est de la trace c'est un peu moins évident. Qu'il nous suffise ici de le voir dans le cas où A est diagonalisable : A est semblable à la matrice de $\mathcal{M}_n(K')$

$$A' = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$

et A et A' ont la même trace (cf. exercice 30, chapitre 3.)
Nous verrons que cette propriété est vraie dans le cas général (cf. corollaire 6.16, page 173). Le lecteur est invité cependant à s'en servir dès maintenant en prenant l'habitude de vérifier que la somme des valeurs propres qu'il a trouvée par le calcul est bien égale à la trace de la matrice.

Exercices 6. 7.

6.6 Caractérisation des endomorphismes diagonalisables

Dans l'exemple du § 5, nous avons vu que les vecteurs propres v_1 et v_2 correspondant aux deux valeurs propres λ_1, λ_2 *différentes* étaient linéairement indépendants. C'est là un fait général que nous allons expliquer dans ce paragraphe.

Définition 6.8 – *Soit* $\lambda \in K$. *On note :*

$$E_\lambda := \{v \in E \mid f(v) = \lambda\, v\}$$

E_λ *est un sous-espace vectoriel de* E *dit* **espace propre** *correspondant à* λ.

On vérifie immédiatement, en effet, que E_λ est stable pour les lois d'addition et de produit par un scalaire.
Remarquons que :

1. si λ *n'est pas valeur propre*, $E_\lambda = \{0\}$;
2. si λ *est valeur propre*, alors :
   $$E_\lambda = \{ \text{ vecteurs propres associés à } \lambda\} \cup \{0\} \quad \text{et} \quad \dim E_\lambda \geq 1.$$

La théorie sur la diagonalisation des endomorphismes se fonde sur deux propositions (A) et (B) (6.9 et 6.12) que nous allons expliquer maintenant.

[7] C'est-à-dire, par exemple si A est une matrice réelle, en prenant les racines réelles ou complexes égales ou distinctes du polynôme caractéristique.

Proposition (A) 6.9 – *Soient $\lambda_1, \ldots, \lambda_p$ des scalaires deux à deux distincts. Alors les espaces propres $E_{\lambda_1} \cdots, E_{\lambda_p}$ sont en somme directe.*

Ceci veut dire (cf. théorème 1.38, page 26) que si $\mathcal{B}_1, \ldots, \mathcal{B}_p$ sont des bases de $E_{\lambda_1}, \ldots, E_{\lambda_p}$, la famille $\{\mathcal{B}_1, \ldots, \mathcal{B}_p\}$ est libre dans E (non nécessairement génératrice).

Démonstration : Par récurrence sur p.

Pour $p = 1$ il n'y a rien à démontrer.

Supposons que $E_{\lambda_1}, \ldots, E_{\lambda_p}$ sont en somme directe et montrons que $E_{\lambda_1}, \ldots, E_{\lambda_p}$, $E_{\lambda_{p+1}}$ sont aussi en somme directe. Il s'agit de montrer que si $x \in (E_{\lambda_1} + \cdots + E_{\lambda_p}) \cap E_{\lambda_{p+1}}$, alors $x = 0$ (cf. proposition 1.41, page 28) : les autres propriétés sont automatiquement vérifiées d'après l'hypothèse de récurrence.
Soit donc $x = x_1 + \cdots + x_p \in (E_{\lambda_1} + \cdots + E_{\lambda_p}) \cap E_{\lambda_{p+1}}$ avec $x_k \in E_{\lambda_k}$. En prenant l'image par f on a :

$$f(x) = \lambda_1\, x_1 + \cdots + \lambda_p\, x_p$$

D'autre part, puisque $x \in E_{\lambda_{p+1}}$

$$f(x) = \lambda_{p+1}\, x \equiv \lambda_{p+1}\, x_1 + \cdots + \lambda_{p+1}\, x_p$$

En faisant la différence :

$$0 = (\lambda_1 - \lambda_{p+1})\, x_1 + \cdots + (\lambda_p - \lambda_{p+1})\, x_p$$

Or $E_{\lambda_1}, \ldots, E_{\lambda_p}$ sont en somme directe, donc $0_E = 0_{E_{\lambda_1}} + \cdots + 0_{E_{\lambda_p}}$ est la seule décomposition de 0, ce qui implique que :

$$(\lambda_k - \lambda_{p+1})\, x_k = 0_{E_k} \qquad \text{pour } k = 1, \ldots, p$$

Comme les λ_i sont deux à deux distinctes, on a $x_k = 0$, pour $k = 1, \cdots, p$, et donc $x = 0$. □

Ainsi donc les espaces propres sont toujours en somme directe, mais il peut se faire que leur somme «ne remplisse pas» E tout entier, c'est-à-dire que l'on ait

$$E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_p} \underset{\neq}{\subset} E$$

ce qui se produit si $\dim E_{\lambda_1} + \cdots + \dim E_{\lambda_p} < \dim E$ (cf. proposition 1.40, page 27). Tout le problème de la diagonalisation tient à cela. Plus précisément, comme nous allons le voir, f est diagonalisable si et seulement si $\dim E_{\lambda_1} + \cdots + \dim E_{\lambda_p} = \dim E$:

Théorème (II) 6.10 – *Soit $f \in \mathrm{End}\,(E)$ et $\lambda_1, \ldots, \lambda_p$ les valeurs propres de f. Alors les propriétés suivantes sont équivalentes :*

1. *f est diagonalisable.*
2. *E est somme directe des espaces propres : $E = E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_p}$.*
3. $\dim E_{\lambda_1} + \cdots + \dim E_{\lambda_p} = \dim E$.

Démonstration : Puisque $E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_p} \subset E$, il est clair que $(b) \iff (c)$.

D'autre part :

- si $E = E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_p}$ alors, en prenant des bases $\mathcal{B}_1, \ldots, \mathcal{B}_p$ de $E_{\lambda_1}, \ldots, E_{\lambda_p}$, la famille $\mathcal{B} := \{\mathcal{B}_1, \cdots, \mathcal{B}_p\}$ est une base de E ; puisque $\mathcal{B}$ est formée de vecteurs propres, f est diagonalisable. Donc $(b) \implies (a)$.
- Réciproquement, si f est diagonalisable, il existe une base $\mathcal{B}$ de E formée de vecteurs propres. Soit :

$$\mathcal{B} = \{\underbrace{v_1, \cdots, v_{n_1}}_{\in E_{\lambda_1}}, \cdots, \underbrace{w_1, \cdots, w_{n_p}}_{\in E_{\lambda_p}}\}.$$

Comme $\operatorname{card}(\mathcal{B}) = \dim E$, on a : $n_1 + \cdots + n_p = \dim E$, c'est-à-dire :

$$\dim E_{\lambda_1} + \cdots + \dim E_{\lambda_p} = \dim E \ .$$

Donc $(a) \implies (c)$ $\square$

Corollaire 6.11 – *Si f admet n valeurs propres deux à deux distinctes alors f est diagonalisable.*

En effet dans ce cas f admet n valeurs propres de multiplicité égale à 1. Il y aura donc n espaces propres de dimension 1.

Les dimensions des espaces propres interviennent donc d'une manière essentielle dans le problème de la diagonalisation. Un renseignement sur la dimension des espaces propres peut être lu directement sur le polynôme caractéristique. La proposition suivante permet d'améliorer le Théorème (II) :

Proposition (B) 6.12 – *Soit $f \in \operatorname{End}_K(E)$ et λ une valeur propre d'ordre α .*[8] *Alors*

$$\boxed{\dim\, E_\lambda \leq \alpha}$$

En effet, supposons que $\dim E_\lambda \geq \alpha + 1$, et soient $v_1, \ldots, v_{\alpha+1}$ des vecteurs propres indépendants correspondant à λ. Complétons cette famille en une base $\mathcal{B}$ de E : $\mathcal{B} = \{v_1, \ldots, v_{\alpha+1}, e_{\alpha+2}, \ldots, e_n\}$ On a :

$$M(f)_{\mathcal{B}} = \left(\begin{array}{ccc|c} \lambda & & 0 & \\ & \ddots & & A \\ 0 & & \lambda & \\ \hline & & 0 & B \end{array}\right)$$

d'où :

$$\begin{aligned} P_f(X) &= \text{dét} \left(\begin{array}{ccc|c} \lambda - X & & 0 & \\ & \ddots & & A \\ 0 & & \lambda - X & \\ \hline & & 0 & B - X\,I \end{array}\right) \\ &= (\lambda - X)^{\alpha+1}\, \text{dét}(B - XI) \end{aligned}$$

λ serait donc valeur propre d'ordre de multiplicité au moins égal à $\alpha + 1$, ce qui est exclu. $\square$

[8] c'est-à-dire λ est racine d'ordre α et de $P_f(\lambda)$.

Le théorème principal sur la diagonalisation est une conséquence des propositions (A) et (B).

Théorème (III) 6.13 – *Soit $f \in \mathrm{End}_K(E)$. Alors f est diagonalisable si et seulement si :*

1. *$P_f(X)$ est **scindé** dans K, ce qui veut dire que $P_f(X)$ s'écrit :*

$$\boxed{P_f(X) \;=\; (-1)^n\,(X-\lambda_1)^{\alpha_1}\cdots(X-\lambda_p)^{\alpha_p}}$$

   avec $\lambda_1,\ldots,\lambda_p \in K$ et $\alpha_1+\ldots+\alpha_p = n$

2. *Les dimensions des espaces propres sont **maximales**, c'est-à-dire : pour chaque valeur propre λ_i de multiplicité α_i, on a :*

$$\boxed{\dim E_{\lambda_i} = \alpha_i}$$

Démonstration : Si les conditions 1. et 2. sont satisfaites, on a :

$$\dim E_{\lambda_1} + \cdots + \dim E_{\lambda_p} = \alpha_1 + \cdots + \alpha_p = n$$

et donc, d'après le théorème (II), f est diagonalisable.
Réciproquement, les conditions sont nécessaires. Supposons en effet f diagonalisable, c'est-à-dire, d'après le théorème (II), $\dim E_{\lambda_1} + \cdots + \dim E_{\lambda_p} = n$. Si $P_f(X)$ n'était pas scindé il serait du type :

$$P_f(X) \;=\; Q(X)(X-\lambda_1)^{\alpha_1}\cdots(X-\lambda_k)^{\alpha_k} \qquad \text{avec} \qquad \alpha_1+\cdots+\alpha_k < n.$$

Par conséquent : $\dim E_{\lambda_1}+\cdots+\dim E_{\lambda_k} \;\le\; \alpha_1+\cdots+\alpha_k < n$, ce qui est exclu. $P_f(X)$ est donc scindé : $P_f(X) = (-1)^n(X-\lambda_1)^{\alpha_1}\cdots(X-\lambda_p)^{\alpha_p}$ avec $\alpha_1+\cdots+\alpha_p = n$.
D'autre part, s'il existait un λ_j tel que $\dim E_{\lambda_j} < \alpha_j$ on aurait :

$$\dim E_{\lambda_1} + \cdots + \dim E_{\lambda_p} < \alpha_1 + \cdots + \alpha_p = n$$

ce qui est exclu. Donc la condition 2. est vérifiée. □

Exemple 1 –

$$A = \begin{pmatrix} 2 & 0 & 4 \\ 3 & -4 & 12 \\ 1 & -2 & 5 \end{pmatrix}$$

On a : $P_A(\lambda) = -\lambda\,(\lambda-1)\,(\lambda-2)$.
A admet donc 3 valeurs propres distinctes et donc elle est diagonalisable.
Déterminons les sous-espaces propres.

$\mathbf{E_0}$: Soit $v_1 = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$; $v_1 \in E_0$ si et seulement si $A\,v_1 = 0$ ce qui donne le système :

$$\begin{cases} 2x \phantom{{}-4y} + 4z = 0 \\ 3x - 4y + 12z = 0 \\ x - 2y + 5z = 0 \end{cases}$$

on trouve : $v_1 = \begin{pmatrix} -4 \\ 3 \\ 2 \end{pmatrix}$ (ou tout autre vecteur colinéaire à celui-ci.)

$\mathbf{E_1}$: Il faut résoudre le système $(A - I)\, v_2 = 0$, c'est-à-dire :

$$\begin{cases} x \quad + \ 4\,z = 0 \\ 3\,x - 5\,y + 12\,z = 0 \\ x - 2\,y + \ 4\,z = 0 \end{cases} \quad \text{ce qui donne, par exemple, } v_2 = \begin{pmatrix} -4 \\ 0 \\ 1 \end{pmatrix}.$$

$\mathbf{E_2}$: On doit résoudre $(A - 2I)\, v_3 = 0$, c'est-à-dire :

$$\begin{cases} 4\,z = 0 \\ 3\,x - 6\,y + 12\,z = 0 \\ x - 2\,y + 3\,z = 0 \end{cases}$$

d'où $v_3 = \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}$, à un facteur de proportionnalité près.

Donc A est semblable à la matrice $A' = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$ et une matrice de passage est $A' = \begin{pmatrix} -4 & -4 & 0 \\ 3 & 0 & 1 \\ 2 & 1 & 0 \end{pmatrix}$.

Exemple 2 –

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

On trouve $P_A(\lambda) = -(\lambda - 1)(\lambda + 2)^2$. Donc A est diagonalisable si et seulement si l'espace propre correspondant à $\lambda = -2$ est de dimension 2.
E_{-2} est défini par le système $(A + 2I)\, v = 0$, c'est-à-dire :

$$\begin{cases} x + y + z = 0 \\ x + y + z = 0 \\ x + y + z = 0 \end{cases}$$

Il s'agit donc du plan vectoriel d'équation $x + y + z = 0$ qui est de dimension 2. f est donc diagonalisable.
Une base de E_{-2} est donnée, par exemple, par

$$v_1 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \qquad \text{et} \qquad v_2 = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}.$$

Pour E_1 on trouve la droite vectorielle d'équation :

$$\begin{cases} -2\,x + y + z = 0 \\ x - 2\,y + z = 0 \end{cases}$$

dont un générateur est $v_3 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$. Ainsi A est semblable à la matrice

$$A' = \begin{pmatrix} -2 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{et on a : } A' = P^{-1}AP \quad \text{avec} \quad P = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ -1 & -1 & 1 \end{pmatrix}.$$

Exemple 3 (cf. exercice 7) –

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

On a : $P_A(\lambda) = -\lambda\,(\lambda^2 + 3\,\lambda + 3)$ donc :

$$\mathrm{Sp}_{\,\mathbb{R}}\, A = \{0\} \qquad \mathrm{Sp}_{\,\mathbb{C}}\, A = \left\{0, \frac{-3+\sqrt{3}\,i}{2}\;,\; \frac{-3-\sqrt{3}\,i}{2}\right\}$$

Ainsi A n'est pas diagonalisable dans $\mathbb{R}$, car le polynôme caractéristique n'est pas scindé dans $\mathbb{R}$, mais elle est diagonalisable dans $\mathbb{C}$, car elle a trois valeurs propres complexes distinctes (cf. corollaire 6.11). C'est-à-dire A est semblable à une matrice diagonale complexe, mais pas à une matrice diagonale réelle.

Exemple 4 –

$$A = \begin{pmatrix} -4 & 0 & -2 \\ 0 & 1 & 0 \\ 5 & 1 & 3 \end{pmatrix}$$

On trouve $P_A(\lambda) = -(\lambda-1)^2(\lambda+2)$. Donc A est diagonalisable si et seulement si $\dim E_1 = 2$. La détermination de E_1 conduit au système

$$\begin{cases} -5\,x \qquad - 2\,z = 0 \\ \;\;5\,x + y - 2\,z = 0 \end{cases}$$

qui définit une droite vectorielle, car les deux équations sont indépendantes. On en déduit que $\dim E_1 = 1$ et, par conséquent, que A n'est diagonalisable ni dans $\mathbb{R}$ ni dans $\mathbb{C}$.

Exercices 8. 9. 10. 11. 12. 13. 14. 15. 16.

6.7 Trois applications

I. $\boxed{\text{Calcul de la puissance d'une matrice}}$

Soit $A \in \mathcal{M}_n(K)$. Supposons A diagonalisable ; il existe alors une matrice diagonale A' et une matrice inversible P telles que : $A' = P^{-1}AP$, c'est-à-dire $A = PA'\,P^{-1}$. Donc :

$$A^k \;=\; \underbrace{(PA'P^{-1})(PA'P^{-1})\cdots(PA'P^{-1})}_{k-\text{fois}} \;=\; P(A')^k\,P^{-1}$$

Or si $A' = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$, $(A')^k = \begin{pmatrix} \lambda_1^k & & 0 \\ & \ddots & \\ 0 & & \lambda_n^k \end{pmatrix}$ et donc A^k se calcule facilement par la formule

$$A^k = P \begin{pmatrix} \lambda_1^k & & 0 \\ & \ddots & \\ 0 & & \lambda_n^k \end{pmatrix} P^{-1}$$

Exemple –

Soit $A = \begin{pmatrix} 1 & -1 \\ 2 & 4 \end{pmatrix}$.

On a $P_A(\lambda) = \lambda^2 - 5\,\lambda + 6 = (\lambda - 2)(\lambda - 3)$. Ainsi $A' = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}$.

Pour les vecteurs propres on trouve :

$$\begin{array}{llll} E_2 : & -x - y = 0\,, & \text{donc :} & v_1 = \begin{pmatrix} 1 \\ -1 \end{pmatrix} \\ E_3 : & -2x - y = 0\,, & \text{donc :} & v_2 = \begin{pmatrix} 1 \\ -2 \end{pmatrix} \end{array}.$$

Ainsi : $P = \begin{pmatrix} 1 & 1 \\ -1 & -2 \end{pmatrix}$ et $P^{-1} = \begin{pmatrix} 2 & 1 \\ -1 & -1 \end{pmatrix}$.

En effectuant les calculs on obtient :

$$A^k = \begin{pmatrix} 2^{k+1} - 3^k & 2^{k+1} - 2\cdot 3^k \\ -2^k + 3^k & -2^k + 2\cdot 3^k \end{pmatrix}.$$

II. Résolution d'un système de suites récurrentes

Illustrons cela sur un exemple. Il s'agit de déterminer deux suites $(u_n)_{n\in\mathbb{N}}$, $(v_n)_{n\in\mathbb{N}}$ telles que :

$$(1) \quad \begin{cases} u_{n+1} &= u_n - v_n \\ v_{n+1} &= 2\,u_n + 4\,v_n \end{cases} \qquad \text{et telles que} \begin{cases} u_0 = 2 \\ v_0 = 1 \end{cases}$$

On pose $X_n = \begin{pmatrix} u_n \\ v_n \end{pmatrix}$. Le système (1) s'écrit :

$$X_{n+1} = A\,X_n \qquad \text{avec} \qquad A = \begin{pmatrix} 1 & -1 \\ 2 & 4 \end{pmatrix}$$

d'où, par récurrence :

$$X_n = A^n X_0 \qquad \text{avec} \qquad X_0 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}.$$

On est ainsi ramené au calcul de A^n. Dans notre cas, compte tenu du résultat ci-dessus :

$$\begin{aligned} \begin{pmatrix} u_n \\ v_n \end{pmatrix} &= \begin{pmatrix} 2^{n+1} - 3^n & 2^{n+1} - 2\cdot 3^n \\ -2^n + 3^n & -2^n + 2\cdot 3^n \end{pmatrix} \begin{pmatrix} 2 \\ 1 \end{pmatrix} \\ &= \begin{pmatrix} 2^{n+2} - 2\cdot 3^n + 2^{n+1} - 2\cdot 3^n \\ -2^{n+1} + 2\cdot 3^n - 2^n + 2\cdot 3^n \end{pmatrix} \end{aligned}$$

c'est-à-dire : $\begin{cases} u_n &= 3\cdot 2^{n+1} - 4\cdot 3^n \\ v_n &= -3\cdot 2^n + 4\cdot 3^n \end{cases}$.

III. Système différentiel linéaire à coefficients constants

Soit à résoudre le système différentiel

$$\left\{\begin{array}{l} \dfrac{\mathrm{d}x_1}{\mathrm{d}t} = a_{11}\,x_1 + \cdots + a_{1n}\,x_n \\ \quad\ \ldots\ldots\ldots\ldots\ldots\ldots \\ \dfrac{\mathrm{d}x_n}{\mathrm{d}t} = a_{n_1}\,x_1 + \cdots + a_{nn}\,x_n \end{array}\right.$$

avec $a_{ij} \in \mathbb{R}$ et $x_i \ : \ \mathbb{R} \longrightarrow \mathbb{R}$ dérivables.
Sous forme matricielle le système s'écrit :

$$(1) \qquad \frac{\mathrm{d}X}{\mathrm{d}t} = AX \quad , \quad \text{où} \qquad A = \Big(a_{ij}\Big), \ \ X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

Supposons A diagonalisable. Il existe alors A' matrice diagonale et P matrice inversible telles que : $A' = P^{-1}AP$. Si on considère A comme la matrice d'un endomorphisme f dans la base canonique, A' est la matrice de f dans la base de vecteurs propres $\{v_i\}$.
De même X est la matrice d'un vecteur $\vec{x}$ dans la base canonique et $X' = M(\vec{x})_{\,v_i}$ est liée à X par

$$X' = P^{-1}X$$

(cf. proposition 3.25, page 77.) En dérivant cette relation :

$$\frac{\mathrm{d}X'}{\mathrm{d}t} = P^{-1}\frac{\mathrm{d}X}{\mathrm{d}t}$$

(car A étant à coefficients constants, P sera aussi à coefficients constants). Donc :

$$\frac{\mathrm{d}X'}{\mathrm{d}t} \ = \ P^{-1}AX \ = \ \Big(P^{-1}AP\Big)X' \ = \ A'X'$$

Le système (1) est donc équivalent au système

$$\frac{\mathrm{d}X'}{\mathrm{d}t} \ = \ A'\,X'$$

Ce système s'intègre facilement, car A' est diagonale.
Ainsi, on peut résoudre le système $\dfrac{\mathrm{d}X}{\mathrm{d}t} = AX$ de la manière suivante :

a) On diagonalise A. Soit $A' = P^{-1}AP$ une matrice diagonale semblable à A ;

b) on intègre le système $\dfrac{\mathrm{d}X'}{\mathrm{d}t} = A'X'$;

c) on revient à X par $X = PX'$.

Exemple –

Soit le système

$$\left\{\begin{array}{l} \dfrac{\mathrm{d}x}{\mathrm{d}t} \quad = x \ \ - y \\ \dfrac{\mathrm{d}y}{\mathrm{d}t} \quad = 2\,x + 4\,y \end{array}\right.$$

On a $A' = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}$ et $P = \begin{pmatrix} 1 & 1 \\ -1 & -2 \end{pmatrix}$

Le système $\dfrac{\mathrm{d}X'}{\mathrm{d}t} = A'X'$ s'écrit :

$$\left\{\begin{array}{rcl} \dfrac{\mathrm{d}x'}{\mathrm{d}t} & = & 2\,x' \\ \dfrac{\mathrm{d}y'}{\mathrm{d}t} & = & 3\,y' \end{array}\right.$$

qui donne immédiatement $\left\{\begin{array}{rcl} x' & = & C_1\,e^{2t} \\ y' & = & C_2\,e^{3t} \end{array}\right.$ et donc, en revenant à X par $X = PX'$:

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ -1 & -2 \end{pmatrix} \begin{pmatrix} C_1\,e^{2t} \\ C_2\,e^{3t} \end{pmatrix} = \begin{pmatrix} C_1\,e^{2t} + C_2\,e^{3t} \\ -C_1\,e^{2t} - 2\,C_2\,e^{3t} \end{pmatrix}$$

c'est-à-dire :

$$\left\{\begin{array}{rcl} x & = & C_1\,e^{2t} \;+\; C_2\,e^{3t} \\ y & = & -C_1\,e^{2t} \;-\; 2\,C_2\,e^{3t} \end{array}\right.$$

Exercices 17. 18. 19. 20.

6.8 Trigonalisation

Une matrice $A \in \mathcal{M}_n(K)$ est dite *triangulaire supérieure* (resp. : *triangulaire inférieure*) si elle est de la forme

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ & a_{22} & & \vdots \\ & & \ddots & \\ 0 & & & a_{nn} \end{pmatrix} \qquad \left(\text{resp. :}\quad A = \begin{pmatrix} a_{11} & & & \\ a_{21} & a_{22} & & 0 \\ \vdots & & \ddots & \\ a_{n1} & \cdots & & a_{nn} \end{pmatrix}\right)$$

Remarque. – Toute matrice triangulaire supérieure est semblable à une matrice triangulaire inférieure.

En effet, soit A une matrice triangulaire supérieure et f l'endomorphisme de K^n qui dans la base canonique $\{e_1, \cdots, e_n\}$ est représenté par A, c'est-à-dire :

$$\left\{\begin{array}{rcl} f(e_1) & = & a_{11}\,e_1 \\ f(e_2) & = & a_{12}\,e_1 \;+\; a_{22}\,e_2 \\ \cdots & & \cdots\cdots\cdots\cdots\cdots\cdots \\ f(e_{n-1}) & = & a_{1,n-1}\,e_1 \;+\cdots+\; a_{n-1,n-1}\,e_{n-1} \\ f(e_n) & = & a_{1n}\,e_1 \;+\cdots+\; a_{nn}\,e_n \end{array}\right.$$

Considérons la base

$$\varepsilon_1 = e_n, \quad \varepsilon_2 = e_{n-1}, \quad \cdots\cdots, \quad \varepsilon_n = e_1.$$

On aura :

$$\left\{\begin{array}{rcl} f(\varepsilon_1) & = & a_{1n}\,\varepsilon_n \;+\; a_{2n}\,\varepsilon_{n-1} \;+\cdots+\; a_{nn}\,\varepsilon_1 \\ f(\varepsilon_2) & = & a_{1,n-1}\,\varepsilon_n \;+\cdots+\; a_{n-1,n-1}\,\varepsilon_2 \\ \cdots & & \cdots\cdots\cdots\cdots\cdots\cdots \\ f(\varepsilon_n) & = & a_{11}\,\varepsilon_n \end{array}\right.$$

donc

$$M(f)_{\varepsilon_i} = \begin{pmatrix} a_{nn} & \cdots & 0 \\ a_{n-1,n} & a_{n-1,n-1} & \\ \vdots & & \ddots \\ a_{1n} & \cdots\cdots & a_{11} \end{pmatrix} = A'$$

A et A' sont semblables car elles représentent le même endomorphisme en des bases différentes.

Le problème que nous allons étudier est de savoir quand une matrice $A \in \mathcal{M}_n(K)$ est semblable à une matrice triangulaire de $\mathcal{M}_n(K)$ (ou, en termes d'endomorphisme, quand un endomorphisme peut être représenté dans une certaine base par une matrice triangulaire). D'après la remarque, il suffira d'étudier le cas où A est semblable à une matrice triangulaire supérieure.

Théorème 6.14 – *Un endomorphisme est trigonalisable dans K si et seulement si son polynôme caractéristique est scindé dans K.*

NOTA –Ceci veut dire (cf. (ii) page 160) que le polynôme caractéristique a n racines dans K ($n = \dim E$) et donc s'écrit :

$$P_f(X) = (-1)^n\,(X-\lambda_1)^{\alpha_1}\cdots(X-\lambda_p)^{\alpha_p}$$

avec $\alpha_1 + \cdots \alpha_p = n$.

Corollaire 6.15 – *Toute matrice $A \in \mathcal{M}_n(\mathbb{C})$ est semblable à une matrice triangulaire de $\mathcal{M}_n(\mathbb{C})$.*

En particulier une matrice $A \in \mathcal{M}_n(\mathbb{R})$ peut être considérée comme matrice de $\mathcal{M}_n(\mathbb{C})$ et, à ce titre, elle est trigonalisable ; cependant les coefficients de la matrice triangulaire semblable à A ne sont pas en général des réels mais des complexes. On dit dans ce cas que A est trigonalisable dans $\mathbb{C}$ mais pas dans $\mathbb{R}$.

Démonstration : Supposons que l'endomorphisme f est trigonalisable et soit une base $\{e_1, \cdots, e_n\}$ telle que :

$$M(f)_{e_i} = \begin{pmatrix} a_{11} & & * \\ & \ddots & \\ 0 & & a_{nn} \end{pmatrix}$$

On a :

$$P_f(X) = \text{dét}\begin{pmatrix} a_{11}-X & & * \\ & \ddots & \\ 0 & & a_{nn}-X \end{pmatrix} = (a_{11}-X)(a_{22}-X)\cdots(a_{nn}-X).$$

donc $P_f(X)$ est scindé (remarquons que si A est triangulaire, les éléments de la diagonale a_{ii} sont les valeurs propres).
Réciproquement supposons $P_f(X)$ scindé et montrons par récurrence que f est trigonalisable.

Pour $n = 1$ il n'y a rien à montrer.
Supposons le résultat vrai à l'ordre $n-1$. Puisque $P_f(X)$ est scindé, il admet au moins une racine $\lambda \in K$ et donc il existe au moins un vecteur propre $\varepsilon_1 \in E_\lambda$.
Complétons $\{\varepsilon_1\}$ en une base : $\{\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n\}$. On a :

$$A := M(f)_{\varepsilon_i} = \begin{pmatrix} \lambda & b_2 & \cdots & b_n \\ 0 & & & \\ \vdots & & B & \\ 0 & & & \end{pmatrix}, \qquad \text{où : } B \in \mathcal{M}_{n-1}(K).$$

Soit $F = \operatorname{Vect}\{\varepsilon_2, \ldots, \varepsilon_n\}$ et $g : F \longrightarrow F$ l'unique endomorphisme de F tel que $M(g)_{\varepsilon_2,\ldots,\varepsilon_n} = B$. On a :

$$P_f(X) = \text{dét}(A - XI_n) = (\lambda - X)\,\text{dét}(B - XI_{n-1}) = (\lambda - X)\,P_g(X)$$

Puisque $P_f(X)$ est scindé, $P_g(X)$ l'est aussi et donc, d'après l'hypothèse de récurrence, B est trigonalisable, c'est-à-dire il existe une base $\{e_2, \ldots, e_n\}$ de F relativement à laquelle $M(g)_{e_2,\ldots,e_n}$ est triangulaire. Dans la base $\{\varepsilon_1,\, e_2, \ldots, e_n\}$ la matrice f est triangulaire. □

Remarques. –

1. Comme on l'a déjà dit, si A est trigonalisable, la matrice A' triangulaire semblable à A a sur la diagonale les valeurs propres de A.
2. Toute matrice $A \in \mathcal{M}_n(K)$ est trigonalisable sur la clôture algébrique K' de K.

Compte tenu de la remarque 1, on peut montrer dans le cas général la propriété concernant la relation liant la trace et le déterminant aux valeurs propres (cf. page 163) :

Corollaire 6.16 – *Soit $A \in \mathcal{M}_n(K)$ et $\mathrm{Sp}'A = \{\lambda_1, \cdots, \lambda_n\}$. On a alors :*

$$\boxed{\begin{aligned} \operatorname{Tr} A &= \lambda_1 + \cdots + \lambda_n \\ \text{dét}\, A &= \lambda_1 \cdot \cdots \cdot \lambda_n \end{aligned}}$$

En effet si $A' \in \mathcal{M}_n(K')$ est une matrice triangulaire semblable à A, on a $\operatorname{Tr} A' = \lambda_1 + \cdots + \lambda_n$ et $\text{dét}\, A' = \lambda_1 \cdot \cdots \cdot \lambda_n$. Puisque deux matrices semblables ont même trace et même déterminant, on a $\operatorname{Tr} A = \lambda_1 + \cdots + \lambda_n$ et $\text{dét}\, A = \lambda_1 \cdot \cdots \cdot \lambda_n$. □

Exemple –

$$\text{Soit} \quad A = \begin{pmatrix} -4 & 0 & -2 \\ 0 & 1 & 0 \\ 5 & 1 & 3 \end{pmatrix}$$

On a $P_A(X) = -(X-1)^2(X+2)$. Le polynôme caractéristique est scindé dans $\mathbb{R}$; donc A est trigonalisable dans $\mathbb{R}$. En regardant A comme la matrice d'un endomorphisme f de $\mathbb{R}^3$ dans la base canonique, on sait qu'il existe une base $\{v_i\}$ de $\mathbb{R}^3$ telle que

$$M(f)_{v_i} = \begin{array}{c} \begin{pmatrix} 1 & a & b \\ 0 & 1 & c \\ 0 & 0 & -2 \end{pmatrix} \\ \begin{array}{ccc} \uparrow & \uparrow & \uparrow \\ f(v_1) & f(v_2) & f(v_3) \end{array} \end{array}$$

ceci signifie que :

$$\begin{array}{c} \mathrm{I} \\ \mathrm{II} \\ \mathrm{III} \end{array} \qquad \left\{ \begin{array}{ll} f(v_1) & = v_1 \\ f(v_2) & = a\, v_1 + v_2 \\ f(v_3) & = b\, v_1 + c\, v_2 - 2\, v_3 \end{array} \right.$$

***Calcul de* $\mathbf{v_1}$** (v_1 est le vecteur propre correspondant à λ_1).

L'équation I , $(f - \mathrm{id})v_1 = 0$ (c'est-à-dire $(A - I)v_1 = 0$), donne :

$$\left\{ \begin{array}{rl} -5\,x - 2\,z & = 0 \\ 5\,x + y + 2\,z & = 0 \end{array} \right.$$

On peut prendre $v_1 = \begin{pmatrix} 2 \\ 0 \\ -5 \end{pmatrix}$

***Calcul de* $\mathbf{v_2}$.**

On résout l'équation II , $(f - \mathrm{id})\, v_2 = a\, v_1$, c'est-à-dire $(A - I)\, v_2 = a\, v_1$:

$$\left\{ \begin{array}{rl} -5\,x - z & = 2\,a \\ 5\,x + y + 2\,z & = -5\,a \end{array} \right.$$

d'où, par exemple, en prenant $a = 1$: $\quad v_2 = \begin{pmatrix} -2 \\ -3 \\ 4 \end{pmatrix}$.

***Calcul de* $\mathbf{v_3}$.**

On sait qu'il existe un vecteur propre v correspondant à la valeur propre -2, c'est-à-dire tel que $f(v) = -2\, v$. On peut prendre, donc : $v_3 = v$, $b = c = 0$.
En résolvant $(f + 2\ \mathrm{id})v_3 = 0$, c'est-à-dire $(A + 2\, I)\, v_3 = 0$, on trouve :

$$\left\{ \begin{array}{rl} -2\,x - 2\,z & = 0 \\ 3y & = 0 \end{array} \right.$$

qui donne, par exemple : $v_3 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$. Ainsi A est semblable à la matrice

$$A' = M(f)_{v_i} = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}$$

et la matrice de passage est la matrice

$$P = \|v_1, v_2, v_3\| = \begin{pmatrix} 2 & -2 & 1 \\ 0 & -3 & 0 \\ -5 & 4 & -1 \end{pmatrix}.$$

Exercice 21.

6.9 Polynômes annulateurs. Théorème de Cayley-Hamilton

Comme nous l'avons vu (théorème (III) 6.13) pour savoir si un endomorphisme est diagonalisable, on est amené à faire une étude des dimensions des espaces propres, étude qui peut être délicate, en particulier si des paramètres interviennent dans le problème. La théorie des polynômes annulateurs, que nous allons développer dans les paragraphes qui suivent, permet de s'affranchir de la recherche des espaces propres et de la discussion sur leur dimension en donnant d'une manière explicite les relations qui doivent relier les coefficients de la matrice pour qu'elle soit diagonalisable.

Soit E un espace vectoriel sur K et $Q \in K[X]$:

$$Q(X) = a_m X^m + a_{m-1} X^{m-1} + \cdots + a_1 X + a_0$$

Si $f \in \operatorname{End}_K(E)$, on note $Q(f)$ l'endomorphisme de E défini par :

$$Q(f) := a_m f^m + a_{m-1} f^{m-1} + \cdots + a_1 f + a_0 \operatorname{id}$$

où $\quad f^k = \underbrace{f \circ \cdots \circ f}_{k-\text{fois}}$

Remarque. – Bien que la loi de composition des endomorphismes ne soit pas commutative, on a cependant :

$$P(f) \circ Q(f) = Q(f) \circ P(f)$$

pour tout couple de polynômes.

Cela tient au fait que f est linéaire et que les différentes puissances de f commutent entre elles. Soient en effet $\quad P(f) = \sum_{i=1}^{n} a_i f^i \quad$ et $\quad Q(f) = \sum_{j=1}^{m} b_j f^j$; on a :

$$\begin{aligned} P(f) \circ Q(f) &= \left(\sum_{i=1}^n a_i f^i\right) \circ \left(\sum_{j=1}^m b_j f^j\right) = \sum_{i,j} a_i b_j f^i \circ f^j = \sum_{i,j} a_i b_j f^{i+j} \\ &= \sum_{i,j} b_j a_i f^j \circ f^i = \left(\sum_{j=1}^m b_j f^j\right) \circ \left(\sum_{i=1}^n a_i f^i\right) = Q(f) \circ P(f). \end{aligned}$$

Soit f un endomorphisme de E. Nous nous intéresserons par la suite aux polynômes $Q(X) \in K[X]$, *non nuls*, tels que $Q(f) = 0$ [9]. Par exemple, si f est un *projecteur* (c'est-à-dire si $f^2 = f$), on a $f^2 - f = 0$ donc le polynôme $Q(X) = X^2 - X$ vérifie $Q(f) = 0$.

Définition 6.17 – *Soit $f \in \operatorname{End}_K(E)$. Un polynôme $Q(X) \in K[X]$ est dit* ***annulateur*** *de f si $Q(f) = 0$.*

Dans les paragraphes qui suivent nous montrerons le résultat suivant

Un endomorphisme f est diagonalisable si et seulement si il existe un polynôme (non nul) annulateur de f scindé et n'ayant que des racines simples (c'est-à-dire d'ordre de multiplicité 1).

Ainsi, par exemple, un projecteur est nécessairement diagonalisable car le polynôme $Q(X) = X(X-1)$ l'annule, est scindé et n'a que des racines simples (cf. exercice 11 chapitre 3).
De même, considérons un endomorphisme f de E tel que $f^3 = f$. On peut montrer (cf. exercice 12 chapitre 3) que $E = E_0 \oplus E_1 \oplus E_{-1}$; donc f est diagonalisable car E est somme directe d'espaces propres. Le théorème que nous venons de citer donne directement le résultat, car le polynôme $Q(X) = X^3 - X = X(X-1)(X+1)$ est annulateur de f.

[9] Bien entendu, 0 désigne ici le l'endomorphisme nul.

Notons tout d'abord que la connaissance d'un polynôme annulateur donne immédiatement des renseignements sur le spectre de f. On a :

Proposition 6.18 – *Soit $Q(X)$ un polynôme annulateur de f. Alors les valeurs propres de f figurent* parmi *les racines de Q, c'est-à-dire :*

$$\mathrm{Sp}(f) \subset \mathrm{Rac}\, Q$$

En effet, si λ est une valeur propre de f, il existe un vecteur v non nul tel que $f(v) = \lambda\, v$. On a :

$$\begin{array}{ll} f^2(v) & = f(\lambda v) = \lambda f(v) = \lambda^2 v \\ f^3(v) & = \lambda^3 v \\ \cdots & \cdots \\ f^k(v) & = \lambda^k v \end{array}$$

Soit $Q(X) = a_m X^m + a_{m-1} X^{m-1} + \cdots + a_1 X + a_0$ un polynôme tel que $Q(f) = 0$, c'est-à-dire vérifiant :

$$a_m f^m + a_{m-1} f^{m-1} + \cdots + a_1 f + a_0 \;\mathrm{id} = 0$$

En appliquant cette relation au vecteur v on trouve :

$$(a_m\, f^m + a_{m-1}\, f^{m-1} + \cdots + a_1\, f + a_0\;\mathrm{id}\,)\, v = 0$$

c'est-à-dire

$$(a_m\, \lambda^m + a_{m-1}\, \lambda^{m-1} + \cdots + a_1\, \lambda + a_0)\, v = 0$$

Or $v \neq 0$, donc $a_m\, \lambda^m + \cdots + a_1\, \lambda + a_0 = 0$, c'est-à-dire $Q(\lambda) = 0$. $\square$

Ainsi, par exemple, un projecteur ne peut avoir que les valeurs propres 0 ou 1 (ou 0 *et* 1). Si f est un endomorphisme qui vérifie $f^3 = f$ ses valeurs propres ne peuvent être que 0, 1 ou -1.

Remarquons que toutes les racines de Q ne sont pas nécessairement valeurs propres. Par exemple id vérifie $\mathrm{id}^2 = \mathrm{id}$, donc $\mathrm{Q}(X) = X(X-1)$ est annulateur, et pourtant 0 n'est valeur propre.

La première question qui se pose est de savoir si, pour tout endomorphisme f, il existe un polynôme annulateur de f (autre que le polynôme nul).
La réponse est affirmative. En effet, si $\dim_K E = n$, on a $\dim_K \mathrm{End}_K(E) = n^2$. Donc toute famille de $n^2 + 1$ endormophismes est liée. En particulier les endomorphismes $\mathrm{id}, f, f^2, \cdots, f^{n^2}$ sont liés. Il existe donc $a_0, a_1, a_2, \cdots, a_{n^2} \in K$ non tous nuls, tels que :

$$a_0\;\mathrm{id} + a_1\, f + a_2\, f^2 + \cdots + a_{n^2}\, f^{n^2} = 0,$$

ce qui veut dire que le polynôme $Q(X) = a_0 + a_1 X + a_2 X^2 + \cdots + a_{n^2} X^{n^2}$ est annulateur de f.

Le théorème suivant dit que parmi les polynômes annulateurs de f il y a toujours le polynôme caractéristique.

Théorème de Cayley-Hamilton . 6.19 – *Soit $f \in \mathrm{End}_K(E)$ et $P_f(X)$ le polynôme caractéristique de f. On a alors :*

$$\boxed{P_f(f) = 0}$$

Remarque. – Comme nous le verrons dans la suite, le théorème de Cayley-Hamilton permet, entre autres, d'élaborer un algorithme qui donne *tous* les polynômes annulateurs de f. On peut ainsi savoir effectivement si parmi les polynômes annulateurs il y en a un qui n'a que des racines simples.

Démonstration : Supposons d'abord $K = \mathbb{C}$. Dans ce cas f est trigonalisable. Soit $\{e_i\}$ une base de E telle que :

$$M(f)_{e_i} = \begin{pmatrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$

On a :

$$P_f(X) = (\lambda_1 - X)(\lambda_2 - X)\cdots(\lambda_n - X).$$

Il s'agit de montrer que

$$(\lambda_1 \operatorname{id} - f) \circ (\lambda_2 \operatorname{id} - f) \circ \cdots \circ (\lambda_n \operatorname{id} - f) = 0.$$

Considérons l'application

$$g_i = (\lambda_1 \operatorname{id} - f) \circ \cdots \circ (\lambda_i \operatorname{id} - f)$$

Nous allons montrer par récurrence que g_i annule les vecteurs $e_1, e_2, \ldots, e_i$; pour $i = n$ on aura le théorème.
Pour i = 1, on a :

$$g_1(e_1) = (\lambda_1 \operatorname{id} - f\,)e_1 = \lambda_1 e_1 - f(e_1) = 0$$

(cf. la première colonne de $M(f)_{e_i}$). Supposons que $g_{i-1}(e_1) = 0, \ \ldots, \ g_{i-1}(e_{i-1}) = 0$ et montrons que :

$$g_i(e_1) = 0, \ \ldots, \ g_i(e_i) = 0.$$

On a : $g_i = g_{i-1} \circ (\lambda_i \operatorname{id} - f\,) = (\lambda_i \operatorname{id} - f) \circ g_{i-1}$, donc :

$$g_i(e_1) = 0, \ \ldots, \ g_i(e_{i-1}) = 0.$$

Il reste à montrer que $g_i(e_i) = 0$. On a : $g_i(e_i) = g_{i-1}(\lambda_i\, e_i - f(e_i))$. Or :

$$M(f)_{e_i} = \begin{pmatrix} \lambda_1 & & & a_1 & & & \\ 0 & \ddots & & & & & * \\ & & \ddots & a_{i-1} & & & \\ \vdots & & & \lambda_i & & & \\ & & & 0 & \ddots & & \\ \vdots & & & \vdots & & \ddots & \\ 0 & & & 0 & & & \lambda_n \end{pmatrix}$$
$$\begin{array}{c} \uparrow \\ f(e_i) \end{array}$$

donc
$$f(e_i) = a_1\, e_1 + \cdots + a_{i-1}\, e_{i-1} + \lambda_i\, e_i$$
d'où :
$$\begin{aligned} g_i(e_i) \quad &= g_{i-1}\big(\lambda_i\, e_i - (a_1\, e_1 + \cdots + a_{i-1}\, e_{i-1} + \lambda_i\, e_i)\big) \\ &= -g_{i-1}\,(\, a_1\, e_1 + a_2\, e_2 + \cdots + a_{i-1}\, e_{i-1}\,) = 0 \end{aligned}$$
car g_{i-1} s'annule sur $e_1, \ldots, e_{i-1}$. Ceci montre que si $A \in \mathcal{M}_n(\mathbb{C})$, $P_A(A) = 0$.

Si $A \in \mathcal{M}_n(\mathbb{R})$, on considère A comme une matrice complexe particulière et donc on a : $P_A(A) = 0$. Le théorème est donc démontré pour $K = \mathbb{R}$ et $K = \mathbb{C}$.

Si K est un corps quelconque, on considère la clôture algébrique K' de K (cf. page 160). Le théorème se démontre de la même manière en faisant jouer à K' le rôle de $\mathbb{C}$ et à K le rôle de $\mathbb{R}$. $\square$

Soit Q un un polynôme scindé :
$$Q(X) = (X - a_1)^{\alpha_1} \cdots (X - a_r)^{\alpha_r}.$$
Le polynôme
$$Q_1(X) = (X - a_1) \cdots (X - a_r)$$
est dit ***radical*** de Q. Il est facile de voir que *si f est un endomorphisme diagonalisable* non seulement $P_f(X)$, mais aussi son **radical** est un annulateur de f :

Proposition 6.20 – *Soit f un endomorphisme et*
$$P_f(X) = (-1)^n\, (X - \lambda_1)^{\alpha_1} \cdots (X - \lambda_p)^{\alpha_p}$$
le polynôme caractéristique. Alors, si f est diagonalisable, *le radical de $P_f(X)$, c'est-à-dire le polynôme $Q(X) = (X - \lambda_1) \cdots (X - \lambda_p)$, annule f.*

En effet, puisque f est diagonalisable, il existe une base $\mathcal{B} = \{v_1, \cdots, v_n\}$ formée de vecteurs propres. Si $\lambda_1, \cdots, \lambda_p$ sont les valeurs propres, pour tout $v \in \mathcal{B}$ il existe au moins une valeur propre λ_j telle que $(f - \lambda_j\,\mathrm{id})v = 0$; donc pour tout $v \in \mathcal{B}$, on a
$$(f - \lambda_1\,\mathrm{id}) \circ \cdots \circ (f - \lambda_p\,\mathrm{id})v = 0,$$
car les endomorphismes $(f - \lambda_k\,\mathrm{id})$ commutent (cf. remarque page 175).
Puisque l'endomorphisme $(f - \lambda_1\,\mathrm{id}) \circ \cdots \circ (f - \lambda_p\,\mathrm{id})$ s'annule sur une base, il s'annule sur tout vecteur. Donc le polynôme $Q(X) = (X - \lambda_1) \cdots (X - \lambda_p)$ est un annulateur de f. $\square$

REMARQUE. – Cette proposition montre en particulier que si f est diagonalisable, alors *il existe un polynôme annulateur qui est scindé et qui a toutes ses racines simples.* Dans le prochain paragraphe nous allons montrer la réciproque de cette propriété.

Exercices 22. 23. 24.

6.10 Le Lemme des noyaux

Lemme des Noyaux . 6.21 – *Soit $f \in \operatorname{End}_K(E)$ et*

$$Q(X) = Q_1(X) \cdot \cdots \cdot Q_p(X)$$

un polynôme factorisé en produit de polynômes deux à deux premiers entre eux.
Si $Q(f) = 0$ alors :

$$\boxed{E = \operatorname{Ker} Q_1(f) \oplus \cdots \oplus \operatorname{Ker} Q_p(f)}$$

Remarque. – Avant de donner la démonstration regardons d'abord comment ce lemme s'applique. Considérons, par exemple, un projecteur, c'est-à-dire un endomorphisme f tel que $f^2 = f$. Le polynôme $Q(X) = X(X-1)$ annule f. Comme X et $X-1$ sont premiers entre eux, on a :

$$E = \operatorname{Ker} f \oplus \operatorname{Ker}(f - \mathrm{id})$$

Donc E est somme directe d'espaces propres et, par conséquent, f est diagonalisable.
Plus généralement, si

$$Q(X) = (X - \lambda_1) \cdots (X - \lambda_p), \quad \text{avec} \quad \lambda_i \neq \lambda_j$$

annule f, on a :

$$\begin{aligned} E &= \operatorname{Ker}(f - \lambda_1 \mathrm{id}) \oplus \cdots \oplus \operatorname{Ker}(f - \lambda_p \mathrm{id}) \\ &= E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_p} \end{aligned}$$

et donc f est diagonalisable.

Plus précisément, on a :

Théorème (IV) Théorème des polynômes annulateurs. 6.22 – *Un endomorphisme f est diagonalisable si et seulement si il existe un polynôme annulateur de f, scindé et n'ayant que des racines simples.*

On vient de montrer en effet que la condition est suffisante. La réciproque est conséquence de la proposition 6.20 page 178. □

Démonstration du Lemme des Noyaux : par récurrence sur p.
Pour $p = 1$ le lemme est évident. En effet $Q(X) = Q_1(X)$ et, par hypothèse, $Q_1(f) = 0$; ce qui veut dire que $\forall x \in E \;:\; Q_1(f)(x) = 0$, c'est-à-dire $E \subset \operatorname{Ker} Q_1(f)$, donc $E = \operatorname{Ker} Q_1(f)$.
Cas où $p = 2$.
Soit $Q = Q_1 \cdot Q_2$. Puisque Q_1 et Q_2 sont premiers entre eux, d'après le théorème de Bézout il existe deux polynômes U_1 et U_2 tels que

$$U_1\, Q_1 + U_2\, Q_2 = 1$$

d'où :

$$U_1(f) \circ Q_1(f) + U_2(f) \circ Q_2(f) = \mathrm{id}$$

Ainsi :

$$\forall x \in E \ : \quad x = U_1(f) \circ Q_1(f)(x) + U_2(f) \circ Q_2(f)(x) \qquad (*)$$

c'est-à-dire :

$$E \subset \operatorname{Im} U_1(f) \circ Q_1(f) + \operatorname{Im} U_2(f) \circ Q_2(f)$$

et donc :

$$E = \operatorname{Im} U_1(f) \circ Q_1(f) + \operatorname{Im} U_2(f) \circ Q_2(f).$$

Or :

$$Q_2(f) \circ U_1(f) \circ Q_1(f) = 0 \quad (\text{car } Q_2(f) \circ Q_1(f) = 0)$$

donc :

$$\operatorname{Im} U_1(f) \circ Q_1(f) \subset \operatorname{Ker} Q_2(f).$$

De même :

$$\operatorname{Im} U_2(f) \circ Q_2(f) \subset \operatorname{Ker} Q_1(f)$$

et par conséquent :

$$E = \operatorname{Ker} Q_1(f) + \operatorname{Ker} Q_2(f).$$

D'autre part, si $x \in \operatorname{Ker} Q_1(f) \cap \operatorname{Ker} Q_2(f)$, d'après (*) on a $x = 0$ et donc :

$$E = \operatorname{Ker} Q_1(f) \ \oplus \ \operatorname{Ker} Q_2(f).$$

Supposons maintenant le théorème vrai jusqu'à l'ordre $p - 1$ et écrivons :

$$Q = \underbrace{Q_1 \cdot Q_2 \cdots Q_{p-1}}_{Q^*} \cdot Q_p$$

D'après le cas $p = 2$:

$$E = \operatorname{Ker} Q^*(f) \ \oplus \operatorname{Ker} Q_p$$

Il reste à montrer que :

$$\operatorname{Ker} Q^*(f) = \operatorname{Ker} Q_1(f) \ \oplus \cdots \ \oplus \operatorname{Ker} Q_{p-1}(f)$$

Soit $F = \operatorname{Ker} Q^*(f)$. D'après l'hypothèse de récurrence :

$$F = \widetilde{\operatorname{Ker}} Q_1(f) \ \oplus \ \cdots \ \oplus \widetilde{\operatorname{Ker}} Q_{p-1}(f)$$

où $\widetilde{\operatorname{Ker}}\, Q_i(f) \ = \ \{\, x \in F \,|\, Q_i(f)(x) = 0 \,\} \ \subset \ \operatorname{Ker} Q_i(f)$.

Or en fait $\widetilde{\operatorname{Ker}} Q_i(f) = \operatorname{Ker} Q_i(f)$. En effet si $x \in \operatorname{Ker} Q_i(f)$ on a $Q^*(f)(x) = 0$, donc $x \in F$ et donc $x \in \widetilde{\operatorname{Ker} Q_i}(f)$. Ainsi donc :

$$F = \operatorname{Ker} Q_1(f) \ \oplus \cdots \oplus \operatorname{Ker} Q_{p-1}(f). \qquad \square$$

Le Lemme des noyaux s'applique en particulier au polynôme caractéristique qui, d'après le théorème de Cayley-Hamilton, est annulateur de f :

Corollaire 6.23 – *Soit* $P_f(X) = (X-\lambda_1)^{\alpha_1} \cdots (X-\lambda_r)^{\alpha_r} Q(X)$ *(avec* $\lambda_i \neq \lambda_j$*) la décomposition de* $P_f(X)$ *en polynômes deux à deux premiers entre eux* [10] *(*Q *étant sans racines dans* K*). On a alors :*

$$E = \operatorname{Ker}(f - \lambda_1 \operatorname{id})^{\alpha_1} \oplus \cdots \oplus \operatorname{Ker}(f - \lambda_r \operatorname{id})^{\alpha_r} \oplus \operatorname{Ker} Q(f)$$

En particulier, si $P_f(X)$ *est scindé dans* K*, c'est-à-dire si*

$$P_f(X) = (-1)^n (X - \lambda_1)^{\alpha_1} \cdots (X - \lambda_p)^{\alpha_p},$$

alors : $$E = \operatorname{Ker}(f - \lambda_1 \operatorname{id})^{\alpha_1} \oplus \cdots \oplus \operatorname{Ker}(f - \lambda_p \operatorname{id})^{\alpha_p}$$

Cette décomposition en somme directe jouera un rôle important dans la suite.

Exercices 25. 26.

6.11 Recherche des polynômes annulateurs. Polynôme minimal

La théorie du polynôme minimal et le théorème de Cayley-Hamilton donnent une méthode de construction systématique de tous polynômes annulateurs de f.

Définition 6.24 – *On appelle* ***polynôme minimal*** *de* f *- noté* $m_f(X)$ *- le polynôme normalisé annulateur de* f *de degré le plus petit* [11].

On démontrera (cf. ci-dessous) que le polynôme minimal est unique.
Notons que, par définition, on a :

$$m_f(f) = 0$$

Il est clair que si un polynôme Q est un multiple de $m_f(X)$ (c'est-à-dire est divisible par $m_f(X)$) alors $Q(f) = 0$. En effet :

$$Q(X) = A(X)\, m_f(X) \quad \Longrightarrow \quad Q(f) = A(f)\, m_f(f) = 0$$

On a, en fait, la réciproque :

Proposition 6.25 – *Les polynômes annulateurs de* f *sont les polynômes du type*

$$Q(X) = A(X)\, m_f(X) \quad \text{avec } A(X) \in K[X]$$

Supposons en effet que $Q(f) = 0$. En effectuant la division euclidienne de Q par m_f, on a :

$$Q(X) = A(X)\, m_f(X) + R(X)$$

où $d^\circ R < d^\circ m_f$ (c'est-à-dire : ou $R = 0$ ou, si $R \neq 0$, $d^\circ R < d^\circ m_f$).
Puisque $Q(f) = 0$ et $m_f(f) = 0$, on a $R(f) = 0$. Donc R est un annulateur de f. Mais $m_f(X)$ est l'annulateur de degré le plus petit : ainsi $R \neq 0$ est impossible (car on aurait alors $d^\circ R < d^\circ m_f$). Donc $R = 0$ et m_f divise Q. $\square$

[10] cf. (i) page 159
[11] *normalisé* veut dire que le coefficient du terme de plus haut degré vaut 1.

Corollaire 6.26 – *$m_f(X)$ divise $P_f(X)$.*

Unicité du polynôme minimal

Soient m_1 et m_2 deux polynômes minimaux. Puisqu'ils sont annulateurs de f, m_1 divise m_2 et m_2 divise m_1. Donc $m_2 = km_1$ $(k \in K)$. Or m_1 et m_2 sont normalisés, donc : $m_1 = m_2$.

Recherche du polynôme minimal

La proposition 6.25 montre que pour connaître tous les polynômes annulateurs de f, il suffit de connaître le polynôme minimal. Nous allons expliquer maintenant comment on détermine le polynôme minimal.

Proposition 6.27 – *Les racines de $m_f(X)$ sont exactement les racines de $P_f(X)$, c'est-à-dire les valeurs propres, mais avec une multiplicité en général différente. En d'autres termes, si on considère $P_f(X)$ scindé (éventuellement sur la clôture algébrique K' de K), c'est-à-dire si*

$$P_f(X) = (-1)^n (X-\lambda_1)^{\alpha_1} \cdots (X-\lambda_p)^{\alpha_p} \qquad \text{avec : } \underset{\alpha_1+\cdots+\alpha_p=n}{\lambda_i \neq \lambda_j} \quad ,$$

alors :

$$m_f(X) = (X-\lambda_1)^{\beta_1} \cdots (X-\lambda_p)^{\beta_p} \quad \text{avec} \quad 1 \le \beta_i \le \alpha_i .$$

Démonstration : On sait que $P_f(X) = A(X)\, m_f(X)$; donc si λ est racine de m_f, alors elle est racine de P_f.
Réciproquement, soit λ racine de P_f, c'est-à-dire valeur propre de f, alors λ est racine de m_f, parce que m_f annule f (cf. proposition 6.18, page 176). $\square$

Exemples :

1. $A = \begin{pmatrix} 0 & 1 & 2 \\ 1 & 0 & 2 \\ 1 & 2 & 0 \end{pmatrix}$. On a : $P_A(X) = -(X+1)(X+2)(X-3)$,

   donc :

$$m_A(X) = (X+1)(X+2)(X-3).$$

2. $A = \begin{pmatrix} -1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 1 & -1 \end{pmatrix}$ $P_A(X) = -(X-1)(X+2)^2$.

   On a donc deux possibilités :

$$\begin{array}{lll} \text{soit} & m_f(X) & = (X-1)(X+2) \\ \text{soit} & m_f(X) & = (X-1)(X+2)^2 \end{array} .$$

   Calculons $(A-I)(A+2I)$; si l'on trouve la matrice nulle le polynôme minimal sera le premier, sinon ce sera le second.

$$(A-I)(A+2I) = \begin{pmatrix} -2 & 1 & 1 \\ 1 & -2 & 1 \\ 1 & 1 & -2 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

   Donc :

$$m_f(X) = (X-1)(X+2)$$

Théorème (V) 6.28 – *Un endomorphisme f est diagonalisable si et seulement si son polynôme minimal est scindé et il a toutes ses racines simples.*

En effet la condition est suffisante, d'après le théorème des polynômes annulateurs 6.22, car $m_f(f) = 0$.
Réciproquement, soit f diagonalisable ; alors $P_f(X)$ est scindé et, comme $m_f(X)$ divise $P_f(X)$, $m_f(X)$ est scindé. Soit $\mathrm{Sp}(f) = \{\lambda_1, ..., \lambda_p\}$ et $\mathcal{B} = \{v_1, ..., v_n\}$ une base de vecteurs propres. Comme on l'a vu (cf. page 178), le polynôme $Q(X) = (X - \lambda_1) \cdots (X - \lambda_p)$ est un annulateur de f. Il s'ensuit que $m_f(X)$ divise $Q(X)$. Puisque $Q(X)$ n'a que des racines simples, $m_f(X)$ n'a que des racines simples. $\square$

Exemples :

1. $A = \begin{pmatrix} -1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 1 & -1 \end{pmatrix}$. On a vu que $m_A(X) = (X+2)(X-1)$.
   Donc A est diagonalisable.

2. $A = \begin{pmatrix} 3 & 2 & -2 \\ -1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$ On a : $P_A(X) = -(X-1)^3$

   donc : $m_A(X) = \begin{cases} X - 1 \\ \text{ou :} \\ (X-1)^2 \\ \text{ou :} \\ (X-1)^3 \end{cases}$

   et A est diagonalisable $\iff m_A(X) = X - 1 \iff A - I = 0$.
   Comme A n'est pas la matrice unité I, A n'est pas diagonalisable.

3. $A = \begin{pmatrix} 3 & -1 & 1 \\ 2 & 0 & 1 \\ 1 & -1 & 2 \end{pmatrix}$. On a : $P_A(X) = -(X-1)(X-2)^2$

   donc : $m_A(X) = \begin{cases} (X-1)(X-2) \\ \text{ou :} \\ (X-1)(X-2)^2 \end{cases}$

   Ainsi :

   $$A \text{ est diagonalisable } \iff m_A(X) = (X-1)(X-2) \iff (A-I)(A-2\,I) = 0.$$

   En effectuant le produit $(A - I)(A - 2\,I)$ on trouve :

   $$(A-I)(A-2\,I) = \begin{pmatrix} 2 & -1 & 1 \\ 2 & -1 & 1 \\ 1 & -1 & 1 \end{pmatrix} \begin{pmatrix} 1 & -1 & 1 \\ 2 & -2 & 1 \\ 1 & -1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{pmatrix} \neq 0$$

   donc A n'est pas diagonalisable.

Exercices 27. 28. 29. 30. 31. 32. 33. 34. 35. 36.

6.12 Réduction en blocs triangulaires (ou réduction selon les espaces caractéristiques)

Comme nous l'avons vu, cf. théorème 6.14, si le polynôme caractéristique $P_f(X)$ est scindé il existe une base (v_i) telle que :

$$M(f)_{v_i} = \begin{pmatrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$

Une simplification supplémentaire est possible dans laquelle plusieurs des termes «*» sont nuls.

Définition 6.29 – *Soit* $P_f(X) = (-1)^n (X-\lambda_1)^{\alpha_1} \cdots (X-\lambda_p)^{\alpha_p}$ *(avec* $\lambda_i \neq \lambda_j$*). On appelle* ***espace caractéristique*** *associé à la valeur propre* λ_i *le sous-espace vectoriel :*

$$N_{\lambda_i} = \mathrm{Ker}(f - \lambda_i \ \mathrm{id})^{\alpha_i}$$

D'après le corollaire 6.23, *si* $P_f(X)$ *est scindé,* E *est somme directe d'espaces caractéristiques* (que f soit diagonalisable ou non) :

$$\boxed{E = N_{\lambda_1} \oplus \cdots \oplus N_{\lambda_p}}$$

Remarques. –

1. $N_\lambda \supset E_\lambda$

   En effet, si $x \in E_\lambda$, $(f - \lambda \ \mathrm{id})(x) = 0$ donc $(f - \lambda \ \mathrm{id})^\alpha(x) = 0$ (α étant l'ordre de multiplicité de λ), c'est-à-dire $x \in N_\lambda$.

2. Les sous-espaces caractéristiques sont *stables* par f, c'est-à-dire : $f(N_\lambda) \subset N_\lambda$.

   En effet, si $x \in N_\lambda$, on a : $(f - \lambda \ \mathrm{id})^\alpha(x) = 0$; d'où $f \circ (f - \lambda \ \mathrm{id})^\alpha(x) = 0$ et donc $(f - \lambda \ \mathrm{id})^\alpha \circ f(x) = 0$, c'est-à-dire $f(x) \in N_\lambda$.

Théorème 6.30 . (Réduction selon les sous-espaces caractéristiques)

Soit $f \in \mathrm{End}_K(E)$*. On suppose que le polynôme caractéristique est scindé :*

$$P_f(X) = (-1)^n (X-\lambda_1)^{\alpha_1} \cdots (X-\lambda_p)^{\alpha_p} \quad (\textit{avec } \lambda_i \neq \lambda_j)$$

Il existe alors une base $\mathcal{B}$ *de* E*,* $\mathcal{B} = \{\mathcal{B}_1, \mathcal{B}_2 \cdots \mathcal{B}_p\}$ *où* $\mathcal{B}_i$ *est une base de* N_{λ_i}*, telle* *que :*

$$M(f)_{\mathcal{B}} = \begin{pmatrix} \boxed{\begin{matrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_1 \end{matrix}} & & & \Large 0 \\ & \boxed{\begin{matrix} \lambda_2 & & * \\ & \ddots & \\ 0 & & \lambda_2 \end{matrix}} & & \\ & & \ddots & \\ \Large 0 & & & \boxed{\begin{matrix} \lambda_p & & * \\ & \ddots & \\ 0 & & \lambda_p \end{matrix}} \end{pmatrix}$$

$$\underbrace{\qquad\qquad}_{\alpha_1} \qquad \underbrace{\qquad\qquad}_{\alpha_2} \qquad\qquad \underbrace{\qquad\qquad}_{\alpha_p}$$

et

$$\begin{pmatrix} \lambda_i & & * \\ & \ddots & \\ 0 & & \lambda_i \end{pmatrix} = M(f|_{N_{\lambda_i}})_{\mathcal{B}_i}$$

La démonstration est très simple et repose sur le lemme suivant :

Lemme 6.31 – *Soit $E = E_1 \oplus \cdots \oplus E_p$, où les E_i sont des sous-espaces vectoriels stables par f. Si $\mathcal{B}_1, \cdots, \mathcal{B}_p$ sont des bases de $E_1, \cdots, E_p$, la matrice de f dans la base $\mathcal{B} = \{\mathcal{B}_1, \cdots, \mathcal{B}_p\}$ de E est :*

$$M(f)_{\mathcal{B}} = \begin{pmatrix} \boxed{M_1} & & & 0 \\ & \boxed{M_2} & & \\ & & \ddots & \\ 0 & & & \boxed{M_p} \end{pmatrix} \qquad \textit{où :} \qquad M_i = M\left(f|_{E_i}\right)_{\mathcal{B}_i}$$

En effet, soient $\mathcal{B}_1 = \{e_1, \cdots, e_{n_1}\}, \cdots, \mathcal{B}_p = \{\varepsilon_1, \cdots, \varepsilon_{n_p}\}$. Puisque $f(E_i) \subset E_i$, on a :

$$\begin{cases} f(e_1) = & a_{11}\, e_1 + \cdots + a_{1n_1}\, e_{n_1} \\ \cdots & \cdots\cdots\cdots\cdots\cdots\cdots \\ f(e_{n_1}) = & a_{n_1 1}\, e_1 + \cdots + a_{n_1 n_1}\, e_{n_1} \end{cases} , \quad \cdots \quad , \begin{cases} f(\varepsilon_1) = & b_{11}\, \varepsilon_1 + \cdots + b_{1n_p}\, \varepsilon_{n_p} \\ \cdots & \cdots\cdots\cdots\cdots\cdots\cdots \\ f(\varepsilon_{n_p}) = & b_{n_p 1}\, \varepsilon_1 + \cdots + b_{n_p n_p}\, \varepsilon_{n_p} \end{cases}$$

donc

$$M(f)_{\mathcal{B}} = \begin{pmatrix} \boxed{M_1} & & 0 \\ & \ddots & \\ 0 & & \boxed{M_p} \end{pmatrix} \qquad \text{où} \qquad M_1 = (\, a_{ij}\,)\ , \cdots ,\ M_p = (\, b_{lm}\,)$$

$\diamondsuit$

Démonstration du théorème :
Puisque les N_{λ_i} sont stables par f, d'après le lemme il existe une base $\mathcal{B}$ de E telle que

$$M(f)_{\mathcal{B}} = \begin{pmatrix} \boxed{M_1} & & 0 \\ & \ddots & \\ 0 & & \boxed{M_p} \end{pmatrix} \qquad \text{où :} \quad M_j = M(f_j),\ f_j = f|_{N_{\lambda_j}}.$$

Il s'agit de montrer que chaque matrice M_j est trigonalisable et ses valeurs propres sont $\{\underbrace{\lambda_j, \cdots, \lambda_j}_{\alpha_j \text{ fois}}\}$. On a :

$$N_{\lambda_j} = \operatorname{Ker}(f - \lambda_j \operatorname{id})^{\alpha_j}$$

donc $(f - \lambda_j \operatorname{id})^{\alpha_j}(x) = 0\,, \ \forall x \in N_{\lambda_j}$, c'est-à-dire : $(f_j - \lambda_j \operatorname{id})^{\alpha_j} = 0$. Ainsi le polynôme $(X - \lambda_j)^{\alpha_j}$ est un annulateur de f_j et par conséquent le polynôme minimal de f_i est du type :

$$m_{f_j}(X) = (X - \lambda_j)^{\gamma_j}\ , \qquad \text{avec} \quad 1 \leq \gamma_j \leq \alpha_j.$$

On en déduit :

$$P_{f_j}(X) = (-1)^{\delta_j}(X - \lambda_j)^{\delta_j}\ , \qquad \text{avec} \quad \delta_j \geq \gamma_j$$

Donc $P_{f_j}(X)$ est scindé et son spectre est $\{\underbrace{\lambda_j, \cdots, \lambda_j}_{\delta_j \text{fois}}\}$. Ceci prouve que M_j est trigonalisable.
Il reste à montrer que $\delta_i = \alpha_i$. On a :

$$\begin{aligned} P_f(X) &= \text{dét}\,(M_i - X\,I) \cdots \text{dét}\,(M_p - X\,I) \\ &= P_{f_1}(X) \cdots P_{f_p}(X) = (-1)^n (X - \lambda_1)^{\delta_i} \cdots (X - \lambda_p)^{\delta_p} \end{aligned}$$

et comme par ailleurs $P_f(X) = (-1)^n (X - \lambda_1)^{\alpha_1} \cdots (X - \lambda_p)^{\alpha_p}$, on a facilement $\delta_i = \alpha_i$. $\square$

Exemple – Soit

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

Considérons A comme la matrice d'un endomorphisme f de $\mathbb{R}^4$ dans la base canonique $\{e_i\}$. On a :

$$P_f(X) = X^2 (X-1)^2.$$

Il existe donc une base de $\mathbb{R}^4$ $\{v_1, \cdots, v_4\}$ telle que :

$$M(f)_{v_i} = \begin{pmatrix} \boxed{\begin{matrix} 1 & a \\ 0 & 1 \end{matrix}} & \begin{matrix} 0 & 0 \\ 0 & 0 \end{matrix} \\ \begin{matrix} 0 & 0 \\ 0 & 0 \end{matrix} & \boxed{\begin{matrix} 0 & b \\ 0 & 0 \end{matrix}} \end{pmatrix}$$

$$\begin{array}{cccc} \uparrow & \uparrow & \uparrow & \uparrow \\ f(v_1) & f(v_2) & f(v_3) & f(v_4) \end{array}$$

ce qui veut dire que :

$$\begin{cases} f(v_1) & = v_1 \\ f(v_2) & = a\, v_1 + v_2 \\ f(v_3) & = 0 \\ f(v_4) & = b\, v_3 \end{cases}$$

Calcul de $\mathbf{v}_1$. On résout le système $A\, v_1 = v_1$, avec $v_1 = \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix}$, c'est-à-dire :

$$\begin{cases} -y + 2z - 2t & = 0 \\ -y + z - t & = 0 \\ x - y & = 0 \\ x - y + z - t & = 0 \end{cases} \quad \to \quad \begin{cases} x & = y \\ y & = 0 \\ z & = t \end{cases}$$

On a donc un espace vectoriel de dimension 1 engendré par $v_1 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix}$.

Calcul de $\mathbf{v}_2$. On résout $f(v_2) = a\, v_1 + v_2$, c'est-à-dire : $(A - I)\, v_2 = a\, v_1$:

$$\begin{cases} -y + 2z - 2t & = 0 \\ -y + z - t & = 0 \\ x - y & = a \\ x - y + z - t & = a \end{cases} \quad \to \quad \begin{cases} x & = a \\ y & = 0 \\ z & = t \end{cases}$$

d'où, par exemple, en prenant $a = 1$: $v_2 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$.

(on ne peut pas prendre $a = 0$ sinon retrouve v_1)

Calcul de $\mathbf{v}_3$. $f(v_3) = 0$ donne le système :

$$\begin{cases} x - y + 2z - 2t & = 0 \\ z - t & = 0 \\ x - y + z & = 0 \\ x - y + z & = 0 \end{cases} \quad \text{On trouve : } v_3 = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}$$

Calcul de $\mathbf{v}_4$. On résout le système $f(v_4) = b\, v_3$:

$$\begin{cases} x - y + 2z - 2t & = b \\ z - t & = b \\ x - y + z & = 0 \\ x - y + z & = 0 \end{cases} \quad \text{qui donne : } \begin{cases} x & = y - b \\ t & = 0 \\ z & = b \end{cases},$$

En prenant $b = 1$, on a par exemple $v_4 = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix}$.

Ainsi :

$$M(f)_{v_i} = \begin{pmatrix} \boxed{\begin{matrix} 1 & 1 \\ 0 & 1 \end{matrix}} & \begin{matrix} 0 & 0 \\ 0 & 0 \end{matrix} \\ \begin{matrix} 0 & 0 \\ 0 & 0 \end{matrix} & \boxed{\begin{matrix} 0 & 1 \\ 0 & 0 \end{matrix}} \end{pmatrix}$$

et une matrice de passage est :

$$P = \|v_1, v_2, v_3, v_4\|_{e_i} = \begin{pmatrix} 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{pmatrix}$$

Notons que l'espace caractéristique N_1 est l'espace engendré par v_1 et v_2, et que N_0 est l'espace engendré par v_3 et v_4.

Application : calcul de la puissance d'une matrice.

Soit A trigonalisable, $A' = P^{-1}AP$ une réduite triangulaire de A par blocs triangulaires, P la matrice de passage.
On a : $A^k = P(A')^k P^{-1}$. Le calcul de $(A')^k$ est simple car :

$$\begin{pmatrix} \boxed{M_1} & & 0 \\ & \ddots & \\ 0 & & \boxed{M_p} \end{pmatrix}^k = \begin{pmatrix} \boxed{M_1}^k & & 0 \\ & \ddots & \\ 0 & & \boxed{M_p}^k \end{pmatrix}$$

On est ramené au calcul de : $\begin{pmatrix} \lambda & & * \\ & \ddots & \\ 0 & & \lambda \end{pmatrix}^k$. Posons :

$$B := \underbrace{\begin{pmatrix} \lambda & & * \\ & \ddots & \\ 0 & & \lambda \end{pmatrix}}_{n} = \lambda I + J \quad \text{où} \quad J := \underbrace{\begin{pmatrix} 0 & & * \\ & \ddots & \\ 0 & & 0 \end{pmatrix}}_{n}$$

On voit facilement que J est nilpotente, c'est-à-dire il existe $p \leq n$ tel que $J^p = 0$. En effet $P_J(X) = (-1)^n X^n$ et donc $m_J(X) = X^p$, avec $p \leq n$; comme $m_J(J) = 0$, on a $J^p = 0$. Puisque λI et J commutent, on calcule facilement B^k par la formule du binôme, à l'aide de J, $J^2, \cdots, J^{n-1}$.

Exercice 37.

6.13 Décomposition de Dunford

Un endomorphisme f est dit ***nilpotent*** (cf. exercice 3) s'il existe $k \in \mathbb{N}$ tel que $f^k = 0$. Le plus petit entier p tel que $f^p = 0$ est dit ***indice de nilpotence.***
Le théorème de décomposition selon les espaces caractéristiques permet de démontrer assez facilement le théorème suivant :

Théorème 6.32 Décomposition de Dunford. .
Soit f un endomorphisme dont le polynôme caractéristique est scindé. Alors f se décompose d'une manière unique *sous la forme :*

$$f = d + n$$

*où d est un endomorphisme **diagonalisable** et n un endomorphisme **nilpotent** qui **commutent.***

En particulier :

> *Toute matrice $A \in \mathcal{M}_n(\mathbb{C})$ se décompose d'une manière* unique *sous la forme $A = D + N$, où D et N sont deux matrices, respectivement* diagonalisable *et* nilpotente*, qui* commutent.

Démonstration :

a) Existence

Soit $\mathcal{B}$ une base dans laquelle f est réduit selon les sous-espaces caractéristiques :

$$A' = M(f)_{\mathcal{B}} = \begin{array}{c}\begin{pmatrix} \boxed{\begin{matrix}\lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_1\end{matrix}} & & & & 0 \\ & \boxed{\begin{matrix}\lambda_2 & & * \\ & \ddots & \\ 0 & & \lambda_2\end{matrix}} & & & \\ & & \ddots & & \\ & & & \ddots & \\ 0 & & & & \boxed{\begin{matrix}\lambda_p & & * \\ & \ddots & \\ 0 & & \lambda_p\end{matrix}} \end{pmatrix} \\ \underbrace{\qquad\quad}_{\alpha_1} \qquad \underbrace{\qquad\quad}_{\alpha_2} \qquad\qquad\qquad \underbrace{\qquad\quad}_{\alpha_p}\end{array}$$

$$= \begin{pmatrix} \boxed{\begin{matrix}\lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_1\end{matrix}} & & 0 \\ & \ddots & \\ 0 & & \boxed{\begin{matrix}\lambda_p & & 0 \\ & \ddots & \\ 0 & & \lambda_p\end{matrix}} \end{pmatrix} + \begin{pmatrix} \boxed{\begin{matrix}0 & & * \\ & \ddots & \\ 0 & & 0\end{matrix}} & & 0 \\ & \ddots & \\ 0 & & \boxed{\begin{matrix}0 & & * \\ & \ddots & \\ 0 & & 0\end{matrix}} \end{pmatrix}.$$

Notons D' la première matrice à second membre et N' la deuxième matrice :

$$A' = D' + N'$$

et soient d et n les endomorphismes qui dans la base $\mathcal{B}$ sont représentés respectivement par D' et N'. On aura :

$$f = d + n.$$

Remarquons que d est diagonalisable, car D' est diagonale, et que n nilpotent, car N' est nilpotente. De plus d et n commutent car D' et N' commutent. En effet, d'après l'exercice 24 du chapitre 3, on a la propriété suivante du produit par blocs :

$$\begin{pmatrix} \boxed{M_1} & & 0 \\ & \ddots & \\ 0 & & \boxed{M_p} \end{pmatrix} \cdot \begin{pmatrix} \boxed{N_1} & & 0 \\ & \ddots & \\ 0 & & \boxed{N_p} \end{pmatrix} = \begin{pmatrix} \boxed{M_1\, N_1} & & 0 \\ & \ddots & \\ 0 & & \boxed{M_p\, N_p} \end{pmatrix}$$

Par conséquent, il suffira de montrer que les matrices

$$\begin{pmatrix} \lambda_i & & 0 \\ & \ddots & \\ 0 & & \lambda_i \end{pmatrix} = \lambda_i\, I \qquad \text{et} \qquad \begin{pmatrix} 0 & & * \\ & \ddots & \\ 0 & & 0 \end{pmatrix}$$

commutent, ce qui est évident.
Aussi f est somme d'un endomorphisme diagonalisable et d'un endomorphisme nilpotent qui commutent.

b) Unicité

Pour montrer que la décomposition est unique, considérons deux décompositions

$$\begin{aligned} f &= d + n \\ &= d' + n' \end{aligned}$$

comme dans l'énonce, c'est-à-dire : d, d' diagonalisables, n, n' nilpotents et d (resp : d') commutant avec n (resp : avec n'). Par soustraction on a :

$$d - d' = n' - n.$$

Soit $h := d - d' = n' - n$. Nous allons montrer que $d - d'$ est diagonalisable que $n - n'$ est nilpotent : on aura alors $h = 0$ (cf. exercice 3), donc : $d = d'$ et $n = n'$.
La démonstration repose sur les lemmes suivants :

Lemme 1. - *La somme de deux endomorphismes diagonalisables* qui commutent*, est un endomorphisme diagonalisable.*

Pour la démonstration, cf. exercice 14.

Lemme 2. - *La somme de deux endomorphismes nilpotents* qui commutent*, est un endomorphisme nilpotent.*

En effet, soient n_1 et n_2 deux endomorphismes nilpotents d'indices de nilpotence respectivement r et s. Puisque n_1 et n_2 commutent, on peut utiliser la formule du binôme ; on a :

$$(n_1 + n_2)^{r+s} = \sum_{k=0}^{r+s} C_{r+s}^k\, n_1^k\, {n_2}^{r+s-k} = 0$$

car si $k \geq r$: $n_1^k = 0$

si $k < r$: alors $r + s - k > s$, donc ${n_2}^{r+s-k} = 0$

Lemme 3. - *d et d' commutent, de même que n et n'.*

Démonstration : - Notons tout d'abord que d commute avec n donc il commute avec $f = d + n$; de même d' commute avec f.
Considérons la décomposition de E en somme directe de sous-espaces caractéristiques :

$$E = N_{\lambda_1} \oplus \cdots \oplus N_{\lambda_p} \ ;$$

tout $x \in E$ s'écrit d'une manière unique

$$x = x_1 + \cdots x_p \quad \text{avec } \ x_i \in N_{\lambda_i}.$$

Soit q_i le projecteur sur N_{λ_i} : $q_i(x) = x_i$. Dans la base $\mathcal{B}$ adaptée à la décomposition on a :

$$M(q_i)_{\mathcal{B}} = \begin{pmatrix} 0 & & & & \Large 0 \\ & \ddots & & & \\ & & \boxed{\begin{matrix} 1 & & \\ & \ddots & \\ & & 1 \end{matrix}} & & \\ & & & \ddots & \\ \Large 0 & & & & 0 \end{pmatrix}$$

donc

$$d = \lambda_1 \, q_1 + \cdots \lambda_p \, q_p.$$

Considérons maintenant l'endomorphisme d' et montrons d'abord qu'il laisse stables les sous-espaces caractéristiques.
Soit $x \in N_{\lambda_i}$, c'est à-dire tel que $\mathrm{Ker}(f - \lambda_i \, \mathrm{id})^{\alpha_i}(x) = 0$. Puisque d' commute avec f, on a :

$$(f - \lambda_i \, \mathrm{id})^{\alpha_i} \circ d'(x) = d' \circ (f - \lambda_i \, \mathrm{id})^{\alpha_i}(x) = 0,$$

c'est-à-dire $d'(x) \in N_{\lambda_i}$. Les sous-espaces caractéristiques sont donc stables par d'.
Il s'ensuit que :

$$\begin{cases} q_i \circ d'\,(x_i) &= \ d'\,(x_i) \\ q_i \circ d'\,(x_j) &= \ 0 \qquad\quad \text{pour } i \neq j \end{cases}.$$

On peut montrer maintenant que d' commute avec tous les projecteurs q_i. En effet, pour tout $x \in E$, on a :

$$\begin{aligned} d' \circ q_i\,(x) &= d'\,(x_i) \qquad \text{et} \\ q_i \circ d'\,(x) &= q_i \circ d'\,(x_1 + \cdots + x_p) = q_i \circ d'\,(x_1) + \cdots q_i \circ d'\,(x_p) \\ &= q_i \circ d'\,(x_i) = d'\,(x_i) \end{aligned}$$

donc $d' \circ q_i = q_i \circ d'$ pour tout $i = 1, ..., p$. Par conséquent d' commute avec d, car $d = \lambda_1 \, q_1 + \cdots \lambda_p \, q_p$.
De la même manière on voit que n' commute avec d.
On a enfin :

$$n' \circ n = n' \circ (f - d) = (f - d) \circ n' = n \circ n' \qquad \diamondsuit$$

Le théorème est ainsi démontré. $\square$

Exercice 38.

6.14 La réduction de Jordan

La réduction par blocs triangulaires est, en général, suffisante pour les applications ; cependant une dernière réduction peut être effectuée à l'intérieur de chaque bloc jusquà parvenir à la forme qui, dans un certain sens, est la plus simple possible : la forme de Jordan.
Par exemple, soit :

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 2 \end{pmatrix} = M(f)_{e_i}$$

Cette matrice est sous forme triangulaire. On peut se demander si elle est diagonalisable. On trouve :

$$P_A(X) = -(X-1)^2(X-2) \quad \text{et} \quad m_A(X) = (X-1)(X-2)$$

A est donc diagonalisable : il existe une base $\{v_i\}$ telle que

$$M(f)_{v_i} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

En revanche, soit

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

On a :

$$P_B(X) = -(X-1)^2(X-2) \quad \text{et} \quad m_B(X) = (X-1)^2(X-2);$$

donc B ne peut pas être réduite à la forme diagonale. Ainsi pour certaines matrices triangulaires une ultérieure réduction est possible, pour d'autres non. La théorie de Jordan permet d'arriver à ce que l'on peut considérer la *dernière réduction*, aboutissant ainsi à une classification des matrices à la relation d'équivalence près "A est semblable à B" (ou si l'on veut : "A et B représentent le même endomorphisme").

Plus précisément le problème est le suivant. Comme on l'a déjà vu, si deux matrices sont semblables alors elles ont la même trace, le même déterminant et, plus généralement, les mêmes valeurs propres. En d'autres termes, si deux matrices ont, par exemple, une trace différente (ce qui est très facile à vérifier), alors elles ne représentent pas le même endomorphisme. Il est clair cependant que le spectre n'est pas le seul invariant de la classe des matrices qui représentent le même endomorphisme : le rang, par exemple, est un autre invariant. Mais le rang et le spectre eux-mêmes ne suffisent pas à caractériser les matrices semblables. Par exemple :

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

ont même rang et même spectre et pourtant elles ne sont pas semblables (vérifier qu'il n'existe pas de matrice inversible P telle que $AP = PB$).
La question se pose d'une manière naturelle de savoir s'il existe un ***système complet d'invariants***, c'est-à-dire si l'on peut associer à chaque matrice A un ensemble fini $\mathcal{S}_A$ de scalaires (tels que le spectre, le rang, etc) tel que si pour deux matrices $A, B \in \mathcal{M}_n(K)$ on a $\mathcal{S}_A = \mathcal{S}_B$, alors A et B sont semblables, c'est-à-dire elles représentent le même endomorphisme en des bases différentes. La réponse est donnée par le théorème de Jordan que nous allons expliquer dans ce paragraphe.

Définition 6.33 – *On appelle bloc de Jordan une matrice carrée du type :*

$$J(\lambda) = \begin{pmatrix} \lambda & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda \end{pmatrix}$$

$\Big[$ *pour les matrices de type* $(1,1)$ *:* $J(\lambda) := (\lambda)$ $\Big]$.

Propriété 6.34 – *Soit* $J(\lambda)$ *un bloc de Jordan d'ordre* n*, on a :*

$$\begin{array}{ll} P_J(X) & = (-1)^n (X-\lambda)^n \\ m_J(X) & = (X-\lambda)^n \\ \dim E_\lambda & = 1 \end{array}$$

La vérification est immédiate : il suffit d'effectuer les calculs.

Théorème de Jordan . 6.35 – *Soit* $f \in \operatorname{End}_K(E)$ *; on suppose que* $P_f(X)$ *est scindé.*

$\boxed{1}$ *Supposons d'abord que* f *n'a qu'une seule valeur propre et que :*

$$P_f(X) = (-1)^n (X-\lambda)^n, \qquad m_f(X) = (X-\lambda)^\beta, \qquad \dim E_\lambda = \gamma$$

Il existe alors une base B *de* E *telle que :*

$$M(f)_{\mathcal{B}} = \begin{pmatrix} \boxed{J_1(\lambda)} & & & 0 \\ & \boxed{J_2(\lambda)} & & \\ & & \ddots & \\ 0 & & & \boxed{J_\gamma(\lambda)} \end{pmatrix} \underset{notation}{=} \tilde{J}(\lambda)$$

où :
- *les* $J_k(\lambda)$ *sont des blocs de Jordan ;*
- *l'ordre du plus grand bloc est* β *;*
- *le nombre des blocs est* γ
  (ceci est évident, car dans chaque bloc il n'y a qu'un vecteur propre, d'après la proposition précédente).

$\boxed{2}$ *Si* f *admet les valeurs propres* $\lambda_1, \ldots, \lambda_p$ *de multiplicité* $\alpha_1, \ldots, \alpha_p$*, c'est-à-dire si*

$$P_f(X) = (-1)^n (X-\lambda_1)^{\alpha_1} \cdots (X-\lambda_p)^{\alpha_p} \qquad (\lambda_i \neq \lambda_j)$$

alors il existe une base $\mathcal{B}$ *de* E *telle que (avec les notations de* $\boxed{1}$*) :*

$$M(f)_{\mathcal{B}} = \begin{pmatrix} \boxed{\tilde{J}_1(\lambda)} & & & 0 \\ & \boxed{\tilde{J}_2(\lambda)} & & \\ & & \ddots & \\ 0 & & & \boxed{\tilde{J}_\gamma(\lambda)} \end{pmatrix}$$
$$\underbrace{\hphantom{\tilde{J}_1(\lambda)}}_{\alpha_1} \quad \underbrace{\hphantom{\tilde{J}_2(\lambda)}}_{\alpha_2} \quad \cdots \quad \underbrace{\hphantom{\tilde{J}_\gamma(\lambda)}}_{\alpha_p}$$

Par exemple : soit $\dim E = 5$, $P_f(X) = -(X-\lambda)^5$, $m_f(X) = (X-\lambda)^3$ et $\dim E_\lambda = 2$, il existe une base $\{v_i\}$ de E telle que :

$$M(f)_{v_i} = \begin{pmatrix} \boxed{\begin{matrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{matrix}} & \begin{matrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{matrix} \\ \begin{matrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{matrix} & \boxed{\begin{matrix} \lambda & 1 \\ 0 & \lambda \end{matrix}} \end{pmatrix}$$

Si $\dim E = 5$, $P_f(X) = -(X-\lambda)^5$, $m_f(X) = (X-\lambda)^3$ et $\dim E_\lambda = 3$, il existe une base $\{v_i\}$ de E telle que :

$$M(f)_{v_i} = \begin{pmatrix} \boxed{\begin{matrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{matrix}} & \begin{matrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{matrix} \\ \begin{matrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{matrix} & \begin{matrix} \boxed{\lambda} & 0 \\ 0 & \boxed{\lambda} \end{matrix} \end{pmatrix}$$

REMARQUE : Une matrice sous la forme de Jordan est diagonalisable si et seulement si elle est déjà sous forme diagonale.

En effet si elle n'est pas diagonale il y a un bloc de Jordan d'ordre $\beta > 1$, ce qui implique que dans le polynôme minimal il y a au moins un facteur du type $(X-\lambda)^\beta$, donc $m_f(X)$ n'a pas toutes ses racines simples.

Démonstration du théorème :

Soit $P_f(X) = (-1)^n (X-\lambda)^n$.
Puisque $P_f(X)$ est scindé, f est trigonalisable. Il existe donc une base $\mathcal{B}'$ telle que :

$$M(f)_{\mathcal{B}'} = \begin{pmatrix} \lambda & & & * \\ & \lambda & & \\ & & \ddots & \\ 0 & & & \lambda \end{pmatrix} \underset{notation}{=} A$$

Posons $A = \lambda I + N$ (ou, si l'on veut, $f = \lambda\,\mathrm{id} + u$) avec :

$$N = \begin{pmatrix} 0 & & * \\ & \ddots & \\ 0 & & 0 \end{pmatrix} = M(u)_{\mathcal{B}'}$$

Comme on l'a vu (cf. page 188), u est nilpotent. Puisque la matrice λI de $\lambda\,\mathrm{id}$ est la même en toute base, le *problème revient à étudier la réduction des endomorphismes nilpotents.*

Soit donc u un endomorphisme nilpotent et β son indice de nilpotence. On a $m_u(X) = X^\beta$. Donc u n'est diagonalisable que si $\beta = 1$, c'est-à-dire $u = 0$. Par la suite on supposera donc que $u \neq 0$.

Lemme 1. *Soit u un endomorphisme nilpotent d'indice de nilpotence β. Les propriétés suivantes sont équivalentes.*

1. *$\beta = n$, c'est-à-dire le polynôme minimal est égal (au signe près) au polynôme caractéristique :*
$$P_u(X) = (-1)^n X^n, \qquad m_u(X) = X^n.$$
2. *Il existe un vecteur $x \in E$, $x \neq 0$, tel que $\{x, u(x), u^2(x), \cdots, u^{n-1}(x)\}$ est une base de E (on dit que u est **cyclique**).*

3. Il existe une base $\mathcal{B}$ telle que :

$$M(u)_{\mathcal{B}} = \begin{pmatrix} 0 & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & 0 \end{pmatrix}$$

(c'est-à-dire u est représentable par un bloc de Jordan).

Démonstration :
L'équivalence de *2.* et *3.* est immédiate : $\mathcal{B}$ est justement la base $\{x, u(x), u^2(x), \cdots, u^{n-1}(x)\}$.
Si *3.* est vérifiée, $\beta = n$ (cf. propriété 6.34).
Supposons *1.* vérifiée. Puisque $u^{n-1} \neq 0$, il existe $x \in E$, $x \neq 0$, tel que $u^{n-1}(x) \neq 0$. Posons (cf. exercice 41) :

$$\begin{cases} v_n & = x \\ v_{n-1} & = u(x), \\ & \cdots \\ v_k & = u^{n-k}(x), \\ & \cdots \\ v_1 & = u^{n-1}(x) \end{cases}$$

On voit facilement que cette famille $\{v_1, \cdots, v_n\}$ est libre et donc elle est une base. En effet soit :

$$\sum_{k=1}^{n} \lambda_k v_k = 0$$

c'est-à-dire :

$$\lambda_1 u^{n-1}(x) + \lambda_2 u^{n-2}(x) + \cdots + \lambda_{n-1} u(x) + \lambda_n x = 0.$$

En prenant l'image par $u^{n-1}, u^{n-2}, \cdots, u$ on trouve successivement : $\lambda_n = 0$, $\lambda_{n-1} = 0, \cdots$, $\lambda_2 = 0$; d'où $\lambda_1 u^{n-1}(x) = 0$, et, puisque $u^{n-1}(x) \neq 0$, $\lambda_1 = 0$. Donc la famille $\{v_i\}$ est une base et

$$M(u)_{v_i} = \begin{pmatrix} 0 & 1 & & 0 \\ & \ddots & \ddots & \\ & & & 1 \\ 0 & & \ddots & 0 \end{pmatrix}$$

Le théorème est ainsi démontré pour $\beta = n$. $\diamondsuit$

D'après le lemme 1. et le lemme 6.31 page 185, le problème revient à démontrer que :

> *si u est un endomorphisme nilpotent, E est somme directe de sous-espaces stables par u, tels que la restriction de u à chacun de ces sous-espaces est un endomorphisme cyclique.*

Comme nous allons le voir, la construction de ces sous-espaces stables se fait comme dans le lemme 1. en choisissant opportunément certains vecteurs et en prenant leurs itérés par u.

Lemme 2. *En notant $K_p = \operatorname{Ker} u^p$, on a la suite d'inclusions strictes :*

$$\{0\} = K_0 \underset{\neq}{\subset} K_1 \underset{\neq}{\subset} K_2 \underset{\neq}{\subset} \cdots \underset{\neq}{\subset} K_{\beta-1} \underset{\neq}{\subset} K_\beta = E$$

En effet, $K_p \subset K_{p+1}$, car $u^p(x) = 0$ implique $u^{p+1}(x) = 0$.
D'autre part, s'il existait $p \in \{1, 2, \cdots, \beta - 1\}$ tel que $K_p = K_{p+1}$, on aurait :

$$K_p = K_{p+1} = K_{p+2} = \cdots = K_\beta = E$$

(cf. exercice 16, chapitre 3). Donc p serait l'indice de nilpotence, ce qui est exclu, car $p < \beta$.
$\diamondsuit$

D'après le lemme 2., pour tout $p \in \{1, \cdots, \beta\}$ il existe un sous-espace vectoriel $M_p \neq \{0\}$ tel que $K_p = K_{p-1} \oplus M_p$. On sait que le supplémentaire de K_{p-1} n'est pas unique, mais tous les supplémentaires ont même dimension. *Nous allons choisir les sous-espaces M_p de manière à ce que certaines conditions, nous permettant de mener à bien la construction, soient satisfaites.*

Lemme 3. *Il existe des sous-espaces vectoriels $M_1, M_2, \cdots, M_\beta$ non réduits à $\{0\}$, tels que :*

$$\begin{array}{lll} 1) & K_p = K_{p-1} \oplus M_p\,, & \text{pour} \quad p = 1, \cdots, \beta \\ 2) & u\,(M_p) \subset M_{p-1}\,, & \text{pour} \quad p = 2, \cdots, \beta \end{array}.$$

Démonstration : par récurrence "descendante" sur p.
– Si $p = \beta$, on choisit pour M_β un supplémentaire quelconque de $K_{\beta-1}$ dans K_β.
– Supposons avoir construit les sous-espaces $M_\beta, M_{\beta-1}, \cdots, M_p$ vérifiant 1) et 2) et montrons que l'on peut construire M_{p-1} vérifiant ces mêmes propriétés.
Remarquons tout d'abord que M_p vérifie :

$$\begin{array}{ll} a) & u\,(M_p) \subset K_{p-1} \\ b) & u\,(M_p) \cap K_{p-2} = \{0\} \end{array}$$

En effet, soit $x \in M_p$. Puisque $M_p \subset K_p$, on a $u^p(x) = 0$, d'où $u^{p-1}(u\,(x)\,) = 0$, c'est-à-dire $u\,(x) \in K_{p-1}$, ce qui prouve a).
Montrons b). Soit $y \in u\,(M_p) \cap K_{p-2}$: $y = u\,(x)$ avec $x \in M_p$ et $u^{p-2}(y) = 0$. On aura : $u^{p-1}(x) = 0$, c'est-à-dire $x \in K_{p-1}$ et par conséquent $x \in M_p \cap K_{p-1} = \{0\}$, d'où $x = 0$ et donc $y = 0$.

Ainsi $u\,(M_p)$ et K_{p-2} sont en somme directe (d'après b)) et $K_{p-2} \oplus u\,(M_p) \subset K_{p-1}$ (d'après a)).
Il s'ensuit qu'il existe un supplémentaire G_{p-1} de $K_{p-2} \oplus u\,(M_p)$ dans K_{p-1} :

$$K_{p-1} = K_{p-2} \oplus u\,(M_p) \oplus G_{p-1}$$

On pose alors :

$$M_{p-1} = u\,(M_p) \oplus G_{p-1}$$

M_{p-1} vérifie 1) et 2). $\diamondsuit$

Lemme 4.

$$E = M_1 \oplus M_2 \oplus \cdots \oplus M_\beta$$

En effet :

$$\begin{aligned} E = K_\beta = K_{\beta-1} \oplus M_\beta \quad &= K_{\beta-2} \oplus M_{\beta-1} \oplus M_\beta \\ &= K_{\beta-3} \oplus M_{\beta-2} \oplus M_{\beta-1} \oplus M_\beta \\ &\cdots\ \cdots\ \cdots\ \cdots\ \cdots\ \cdots \\ &= M_1 \oplus M_2 \oplus \cdots\ \cdots\ \cdots\ \cdots \oplus M_\beta \end{aligned}$$

$\diamondsuit$

Nous allons maintenant construire une base de E en choisissant, par un procédé itératif, une base sur chaque espace M_i : cela permet de mettre en évidence les sous-espaces stables sur lesquels u est cyclique.

Remarquons tout d'abord que $u(M_p) \subset M_{p-1}$ et que l'on a :

Lemme 5. *L'image par u d'une base de M_p est une famille libre de M_{p-1} (pour $p \geq 2$).*

En effet, soit $\{v_1, \cdots, v_r\}$ une base de M_p et soient $\lambda_1, \ldots \lambda_r$, tels que :

$$\lambda_1 \, u\,(v_1) + \cdots + \lambda_r \, u\,(v_r) = 0$$

on a $u(\lambda_1 \, v_1 + \cdots + \lambda_r \, v_r) = 0$, donc :

$$\lambda_1 \, v_1 + \cdots + \lambda_2 \, v_2 \in \operatorname{Ker} u = K_1 \subset K_{p-1}$$

Ainsi $\lambda_1 \, v_1 + \cdots + \lambda_r \, v_r \in M_p \cap K_{p-1} = \{0\}$ et, puisque $v_1, \cdots, v_r$ sont indépendants, on a :

$$\lambda_1 = \lambda_2 = \cdots\cdots = \lambda_r = 0 \qquad \diamondsuit$$

Venons maintenant à la construction de la base de E :

- Dans M_β (que nous allons noter G_β), on prend une base quelconque.
- Dans $M_{p-1} = u\,(M_p) \oplus G_{p-1}$ on prend l'image par u de la base construite sur M_p et on complète par une base quelconque de G_{p-1}.

On a ainsi les bases $\mathcal{B}_\beta, \mathcal{B}_{\beta-1}, \ldots, \mathcal{B}_1$ de $M_\beta, M_{\beta-1}, \ldots, M_1$:

$$\begin{aligned}
\mathcal{B}_\beta \; &= \Big\{ \underbrace{v_1, \ldots, v_{n_\beta}}_{G_\beta} \Big\} \\
B_{\beta-1} &= \Big\{ u(v_1), \ldots, u(v_{n_\beta}), \underbrace{w_1, \ldots, w_{n_{\beta-1}}}_{G_{\beta-1}} \Big\} \\
\mathcal{B}_{\beta-2} &= \Big\{ u^2(v_1), \ldots, u^2(v_{n_\beta}), u(w_1), \ldots, u(w_{n_{\beta-1}}), \underbrace{z_1, \ldots, z_{n_{\beta-2}}}_{G_{\beta-2}} \Big\} \\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
\mathcal{B}_1 \; &= \Big\{ u^{\beta-1}(v_1), \ldots, u^{\beta-1}(v_{n_\beta}), u^{\beta-2}(w_1), \ldots, u^{\beta-2}(w_{n_{\beta-1}}), \ldots, \underbrace{u(y_1), \ldots, u(y_{n_2})}_{u(G_2)}, \\
&\qquad \underbrace{x_1, \ldots, x_{n_1}}_{G_1} \Big\}
\end{aligned}$$

Introduisons la notation suivante : pour $k = \beta, \beta - 1, \ldots, 2, 1$ et $x \in G_k$, soit :

$$I_k(x) = \operatorname{Vect}\{x, u\,(x), u^2(x), \ldots, u^{k-1}(x)\,\}$$

Ainsi, par exemple :

$$\begin{aligned}
I_\beta(v_1) \quad &= \operatorname{Vect}\{v_1, u(v_1), u^2(v_1), \ldots, u^{\beta-1}(v_1)\,\} \\
I_{\beta-1}(w_1) \quad &= \operatorname{Vect}\{w_1, u(w_1), u^2(w_1), \ldots u^{\beta-2}(w_1)\,\} \qquad \text{etc.}
\end{aligned}$$

Lemme 6.

a) $\dim I_k(x) = k$

b) $I_k(x)$ *est stable par* u.

c) *La restriction de* u *à* $I_k(x)$ *est cyclique.*

La démonstration de ce lemme est une simple vérification. $\diamondsuit$

Il est clair que E est somme directe de $I_\beta(v_1), \ldots, I_\beta(v_{n_\beta}), I_{\beta-1}(w_1), \ldots, I_{\beta-1}(w_{n_{\beta-1}})$, $\ldots, I_1(x_1), \ldots, I_1(x_{n_1})$ d'où il suit immédiatement la partie 1. du théorème.

La partie 2. vient du fait que E est somme directe de sous-espaces caractéristiques et sur chaque espace caractéristique on applique la partie 1. du théorème. $\square$

Calcul des dimensions des blocs de Jordan

Dans l'énoncé du théorème nous avons donné des indications sur la dimension des blocs de Jordan, qui sont suffisantes, en général, pour déterminer la forme de Jordan lorsque la dimension de E n'est pas trop grande. En fait, en analysant la démonstration, on voit que l'on peut calculer explicitement à l'aide de f le nombre des blocs de Jordan de taille donnée qui apparaissent dans la réduction.

Proposition 6.36 – *Soit*

$n_p(\lambda)$ *: = nombre de blocs de Jordan d'ordre p pour la valeur propre λ ;*

$K_p(\lambda) := \mathrm{Ker}(f - \lambda\ \mathrm{id})^p$.

On a alors :

$$\boxed{n_p(\lambda) \;=\; 2\dim K_p(\lambda) \;-\; \dim K_{p-1}(\lambda) \;-\; \dim K_{p+1}(\lambda)}$$

En effet, comme dans la démonstration du théorème, on peut se limiter à démontrer cela pour un endomorphisme nilpotent. Dans ce cas, $n_p(0) = \dim G_p$. Or

$$K_p = K_{p-1} \oplus M_p\,, \qquad M_p = u(M_{p+1}) \oplus G_p$$

et $u|_{M_{p+1}}$ est injective. Ainsi :

$$\dim G_p = \dim M_p - \dim M_{p+1}$$

et

$$\dim M_p = \dim K_p - \dim K_{p-1},$$

d'où la formule. □

Corollaire 6.37 – *Deux matrices $A, B \in \mathcal{M}_n(K)$, dont les polynômes caractéristiques sont scindés, sont semblables si et seulement si elles ont la même réduite de Jordan (à l'ordre des blocs près).*

En effet si elles ont la même réduite de Jordan J, alors elles sont semblables à J et donc elles sont semblables.
Réciproquement, si A et B sont semblables, elles représentent le même endomorphisme, donc, en particulier, elles ont les mêmes valeurs propres. D'autre part, comme on vient de le voir, les dimensions des blocs de Jordan sont calculées à l'aide de l'endomorphisme par les formules ci-dessus et elles dépendant donc de l'endomorphisme et pas des matrices qui le représentent. Ainsi les deux matrices auront les mêmes blocs de Jordan.

On peut aussi exprimer le corollaire en disant que si A est une matrice carrée dont le polynôme caractéristique est scindé, alors :

l'ensemble des valeurs propres (avec leurs ordres) et les dimensions des blocs de Jordan associées à chaque valeur propre forment un ***système complet d'invariants****.*

Exercices 39. 40. 41. 42. 43.

EXERCICES

1 Soit E un espace vectoriel de dimension finie, F et G, deux sous-espaces vectoriels tels $E = F \oplus G$ et $\mathrm{p}_F : E \to E$ le projecteur de E sur F. Montrer que $\mathrm{Tr}\,\mathrm{p}_F = \dim F$.

2 Soit $E = \mathbb{R}^2$ et $f : \mathbb{R}^2 \to \mathbb{R}^2$ la symétrie orthogonale par rapport à la première bissectrice. Déterminer géométriquement les vecteurs propres, les valeurs propres, la trace et le déterminant de f.

3 Soit f un endomorphisme nilpotent (c'est-à-dire : il existe $p \in \mathbb{N}$ tel que $f^p = 0$). Montrer que f admet comme seule valeur propre 0. En déduire que si f est diagonalisable et nilpotent, alors $f = 0$.

* **4** Soit $P_f(X) = (-1)^n X^n + a_{n-1}\, X^{n-1} + \cdots + a_1\, X + a_0$. Montrer que $a_0 = \text{dét}\, f$ et $a_{n-1} = (-1)^{n-1}$ Tr f.
En déduire que pour $A \in \mathcal{M}_2(K)$:, $P_A(X) = X^2 - (\text{Tr}A)\, X + \text{dét}\, A$.

** **5** **Densité du groupe linéaire.**

Dans cet exercice on suppose connu le fait que $\mathcal{M}_n(K)$, *avec* $K = \mathbb{R}$ ou $\mathbb{C}$, *est un espace vectoriel normé. On prendra, par exemple pour norme de* $A = (a_{ij})$: $\|A\| = \sqrt{\sum_{i,j=1}^n a_{ij}}$ *(cf. Appendice A.6.) et on munira* $\mathcal{M}_n(K)$ *de la topologie associée à la norme.*

Montrer que le groupe linéaire des matrices inversibles, $\text{GL}(n, K)$, est dense dans $\mathcal{M}_n\,(K)$.
Application. Généraliser les propriétés de l'exercice 20. du chapitre 4 au cas non inversible, c'est-à-dire :

$$\text{dét}\, A^+ = (\text{dét}\, A)^{n-1} \quad , \quad A^{++} = (\text{dét}\, A)^{n-2}\, A$$

pour toute $A \in \mathcal{M}_n\,(K)$, où $A^+ = {}^t\text{cof}\,(A)$.

6 Soit $A = \begin{pmatrix} 7 & 4 & 0 & 0 \\ -12 & -7 & 0 & 0 \\ 20 & 11 & -6 & -12 \\ -12 & -6 & 6 & 11 \end{pmatrix}$. Montrer que A est diagonalisable ; déterminer une matrice réduite diagonale et une matrice de passage.

7 Soit $A \in \mathcal{M}_n(\mathbb{R})$. Montrer que si λ est une valeur propre complexe de A alors $\overline{\lambda}$ est aussi valeur propre de A, de même ordre de multiplicité, et que si v est un vecteur propre correspondant à λ alors $\overline{v}$ est un vecteur propre correspondant à $\overline{\lambda}$ (où $\overline{v}$ désigne le vecteur dont les composantes sont les conjuguées des composantes de v).
Soit $A = \begin{pmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \\ 1 & 0 & -1 \end{pmatrix}$. Déterminer une matrice dans $\mathcal{M}_n(\mathbb{C})$ diagonale et semblable à A, ainsi qu'une matrice de passage.

8 Soit f l'endomorphisme de $\mathbb{R}^3$ qui dans la base canonique est représenté par la matrice :

$$A = \begin{pmatrix} 1 & \text{Log}_a b & \text{Log}_a c \\ \text{Log}_b a & 1 & \text{Log}_b c \\ \text{Log}_c a & \text{Log}_c b & 1 \end{pmatrix}$$

où a, b, c sont trois nombres réels strictement positifs et différents de 1.
Montrer que f est diagonalisable et déterminer une base de vecteurs propres. En déduire la signification géométrique de f.

* **9** Soit

$$A_t = \begin{pmatrix} t & 1 & \cdots & \cdots & 1 \\ 1 & t & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 1 \\ 1 & \cdots & \cdots & 1 & t \end{pmatrix} \in \mathcal{M}_n(\mathbb{R}).$$

Sans calculer le polynôme caractéristique, montrer que $\lambda = t - 1$ est valeur propre. Déterminer E_{t-1} ; que peut-on dire de l'ordre de multiplicité de $t-1$ comme valeur propre ? En déduire $\mathrm{Sp}' A_t$. A_t est elle diagonalisable ? Pour quelles valeurs de t a-t-on $\det A_t = 0$?

* **10** Soit

$$A = \begin{pmatrix} 0 & a_1 & a_2 & a_3 & \cdots & a_n \\ a_1 & 0 & a_2 & a_3 & \cdots & a_n \\ a_1 & a_2 & 0 & a_3 & \cdots & a_n \\ \vdots & \vdots & & \ddots & & \vdots \\ \vdots & \vdots & & & \ddots & \vdots \\ a_1 & a_2 & a_3 & \cdots & a_n & 0 \end{pmatrix}$$

Déterminer, comme dans l'exercice 9, le spectre de A sans passer par le polynôme caractéristique. En déduire $\det A$.
Déduire de cet exemple une méthode simple pour construire une matrice diagonalisable non diagonale, dont on se donne les valeurs propres.

* **11** Soit E un espace vectoriel de dimension finie et $f \in \mathrm{End}_K(E)$ tel que $\mathrm{rg}\, f = 1$. Déterminer le spectre de f. Montrer que f est diagonalisable si et seulement si $\mathrm{Tr}\, f \neq 0$. Déduire de cet exemple une méthode simple pour construire une matrice non diagonalisable.

12 Soient f et g deux endomorphismes d'un espace vectoriel de dimension finie. Montrer que $f \circ g$ et $g \circ f$ ont mêmes valeurs propres.

* **13** Soient f et g deux endomorphismes d'un espace vectoriel de dimension n, tels que $f \circ g = g \circ f$. On suppose que g admet n valeurs propres distinctes. Montrer que f et g admettent une base commune de vecteurs propres. En déduire qu'il existe $a_0, a_1, a_2, \ldots, a_{n-1} \in K$ tels que :

$$f = a_0 \ \mathrm{id} + a_1\, g + a_2\, g^2 + \cdots + a_{n-1}\, g^{n-1}$$

** **14** **Diagonalisation simultanée.**
Soient f et g deux endomorphismes diagonalisables. On suppose que f et g commutent. On note $\lambda_1, \ldots, \lambda_p$ (resp : $\mu_1, \ldots, \mu_q$) les valeurs propres de f (resp : de g) et $F_1, \ldots, F_p$ les sous-espaces propres correspondants (resp : $G_1, \ldots, G_q$).

1. Montrer que chaque G_j est stable par f et chaque F_i est stable par g.
2. On pose : $H_{ij} := F_i \cap G_j \quad (i = 1, \ldots, p\,; \quad j = 1, \ldots, q)$. Montrer que, pour i donné, F_i est somme directe des $H_{ij},\ j = 1, \ldots, q$ (qui ne sont pas réduits à $\{0\}$). En déduire le résultat suivant :

   Étant donnés deux endomorphismes diagonalisables f et g, il existe une base de vecteurs propres commune à f et à g (c'est-à-dire f et g sont diagonalisables dans une même base) si et seulement si f et g commutent.

   En particulier :

   La somme de deux endomorphismes diagonalisables qui commutent est un endomorphisme diagonalisable.

** **15** **Décomposition spectrale**.

1. Soit f un endomorphisme diagonalisable de E, $\lambda_1, \ldots, \lambda_q$ les valeurs propres de f. Montrer qu'il existe un système complet de projecteurs $\{P_1, \ldots, P_q\}$ (cf. exercice 10, chapitre 3) tel que :

$$f = \lambda_1\, P_1 + \cdots + \lambda_q\, P_q \qquad (*)$$

2. Réciproquement, montrer que, s'il existe un système complet de projecteurs $\{P_1, \ldots, P_q\}$ et des scalaires $\lambda_1, \ldots, \lambda_q$ deux à deux distincts tels que $f = \lambda_1 P_1 + \ldots + \lambda_q P_q$, alors f est diagonalisable.

   L'expression (*) est dite ***décomposition spectrale*** de f.

3. Application. Déterminer la décomposition spectrale de

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

   En déduire A^n.

4. Montrer que dans la décomposition spectrale de f, le projecteurs P_i sont des polynômes de f dont les coefficients s'expriment en fonction des λ_j et peuvent être calculés explicitement.

16 **Exemple de vecteurs propres en dimension infinie.**

1. Soit E l'espace vectoriel des fonctions C^∞ sur $\mathbb{R}$ à valeurs dans $\mathbb{R}$.
   On considère l'endomorphisme Φ de E qui à f associe la fonction g définie par :

$$g(x) = \int_0^1 e^{x-t}\, f(t)\, \mathrm{d}t$$

   Déterminer les valeurs propres et les vecteurs propres de Φ.

2. Mêmes questions pour $\Phi(f) = g$ où :

$$g(x) = \int_{-\pi/2}^{\pi/2} \cos(x+t)\, f(t)\, \mathrm{d}t.$$

17 Soit $f : \mathbb{R}_2[x] \to \mathbb{R}_2[x]$ définie par :

$$f(Q) = (2\,x + 1)\, Q - (x^2 - 1)\, Q$$

Calculer $f^n(a_0 + a_1\, x + a_2\, x^2)$.

18 On note $\mathcal{E}$ l'espace vectoriel des matrices carrées réelles d'ordre 2 à trace nulle. Soit $B = \begin{pmatrix} 1 & 0 \\ 2 & 3 \end{pmatrix}$ et $f : \mathcal{E} \to \mathcal{E}$ définie par $f(M) = MB - BM$.

Calculer $f^n(A)$ où $A = \begin{pmatrix} a & b \\ c & -a \end{pmatrix}$.

19 Résoudre le système différentiel :

$$\begin{cases} \dfrac{\mathrm{d}x}{\mathrm{d}t} = 4\,x + 6\,y \\ \dfrac{\mathrm{d}y}{\mathrm{d}t} = -3\,x - 5\,y \\ \dfrac{\mathrm{d}z}{\mathrm{d}t} = -3\,x - 6\,y - 5\,z \end{cases}$$

* **20** **Résolution d'un système différentiel par la "méthode d'élimination".**

Soit $P \in \mathbb{R}_n[x]$, $P = a_0 + a_1 x + \cdots + a_n x^n$. On note $P\left[\frac{\mathrm{d}}{\mathrm{d}t}\right]$ l'opérateur différentiel $P\left[\frac{\mathrm{d}}{\mathrm{d}t}\right] : \mathcal{C}^\infty(\mathbb{R}) \to \mathcal{C}^\infty(\mathbb{R})$ défini par :

$$P\left[\frac{\mathrm{d}}{\mathrm{d}t}\right] = a_0\ \mathrm{id} + a_1 \frac{\mathrm{d}}{\mathrm{d}t} + \cdots + a_n \left(\frac{\mathrm{d}}{\mathrm{d}t}\right)^n.$$

Soit le système différentiel :

$$\left\{\begin{array}{ll} \dfrac{\mathrm{d}x_1}{\mathrm{d}t} & = a_{11}x_1 + \cdots + a_{1n}x_n \\ \cdots & \cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \dfrac{\mathrm{d}x_n}{\mathrm{d}t} & = a_{n1}x_1 + \cdots + a_{nn}x_n \end{array}\right. \qquad a_{ij} \in \mathbb{R}.$$

Montrer que chaque x_i vérifie l'équation différentielle

$$P_A\left[\frac{\mathrm{d}x_i}{\mathrm{d}t}\right] = 0$$

où P_A est le polynôme caractéristique de la matrice $A = (a_{ij})$.
Application. Résoudre le système différentiel :

$$\left\{\begin{array}{l} \dfrac{\mathrm{d}x_1}{\mathrm{d}t} = 3\,x_1 - 2\,x_2 \\ \dfrac{\mathrm{d}x_2}{\mathrm{d}t} = 2\,x_1 - x_2 \end{array}\right. .$$

21

1. Pour chacune des matrices ci-dessous, donner une matrice réduite triangulaire en précisant la matrice de passage :

$$A_1 = \begin{pmatrix} 3 & -1 & 1 \\ 2 & 0 & 1 \\ 1 & -1 & 2 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 3 & 2 & -2 \\ -1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}, \quad A_3 = \begin{pmatrix} 13 & -5 & -2 \\ -2 & 7 & -8 \\ -5 & 4 & 7 \end{pmatrix}.$$

2. Résoudre le système différentiel $\frac{\mathrm{d}X}{\mathrm{d}t} = A_1 X$.

22 Soit Q un polynôme et f un endomorphisme. Montrer que si f est diagonalisable alors $Q(f)$ est diagonalisable et, si $\mathrm{Sp} f = \{\lambda_1, \ldots, \lambda_p\}$ est le spectre de f, alors :

$$\mathrm{Sp}\, Q(f) = \{\, Q(\lambda_1), \ldots, Q(\lambda_p)\,\}$$

23 Donner une autre démonstration du résultat de l'exercice 20 en utilisant le théorème de Cayley-Hamilton.

* **24** Soient $A,\, B \in \mathcal{M}_n(\mathbb{C})$. On suppose que $\mathrm{Sp}_{\mathbb{C}}A \cap \mathrm{Sp}_{\mathbb{C}}B = \emptyset$.

1. Montrer que $\mathrm{dét}\, P_B(A) \neq 0$.
2. Soit $\varphi : \mathcal{M}_n(\mathbb{C}) \longrightarrow \mathcal{M}_n(\mathbb{C})$ définie par $\varphi(M) = A\,M - M\,B$. Montrer que si $M \in \mathrm{Ker}\,\varphi$, alors pour tout polynôme $Q(X)$, on a :
$$Q(A)\,M = M\,Q(B).$$
3. Déduire des questions précédentes que φ est injective et que
$$\forall\, C \in \mathcal{M}_n(\mathbb{C}) \ \ \exists!\, M \in \mathcal{M}_n(\mathbb{C}) \text{ telle que : } \ A\,M - M\,B = C.$$

* **25** **Calcul des projecteurs sur** $\mathrm{Ker} Q_1(f)$ **et** $\mathrm{Ker} Q_2(f)$.
Soit f un endomorphisme de E et $Q = Q_1\,Q_2$ un annulateur de f, où les polynômes Q_1 et Q_2 sont premiers entre eux. Soient U_1, U_2 deux polynômes tels que $U_1Q_1 + U_2Q_2 = 1$. Montrer que les projecteurs de $E = \mathrm{Ker}\,Q_1(f) \oplus \mathrm{Ker}\,Q_2(f)$ sur $\mathrm{Ker}\,Q_1(f)$ et $\mathrm{Ker}\,Q_2(f)$ sont respectivement

$$\mathrm{pr}_1 := U_2(f) \circ Q_2(f) \qquad \text{et} \qquad \mathrm{pr}_2 := U_1(f) \circ Q_1(f)\,.$$

En déduire un algorithme permettant d'exprimer les projecteurs sur les espaces $\mathrm{Ker}\,Q_1(f)$ et $\mathrm{Ker}\,Q_2(f)$ comme polynômes de f (*il s'agit donc d'une généralisation de l'exercice 15. 4*)

26 Soit la matrice

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

On admettra que i et 1 sont valeurs propres de A. En déduire que

$$\mathbb{R}^4 = \mathrm{Ker}(\,A - I\,)^2 \oplus \mathrm{Ker}(\,A^2 + I\,).$$

27 Calculer le polynôme minimal des matrices de l'exercice 21.

28 Soit $A = \begin{pmatrix} 1 & a & 1 \\ 0 & 1 & b \\ 0 & 0 & c \end{pmatrix}$, $(a, b, c \in \mathbb{R})$.

Pour quelles valeurs de a, b, c , A est-elle diagonalisable ?

29 Soit $A = \begin{pmatrix} 5 & 1 & 1 & -1 \\ 1 & 5 & 1 & -1 \\ 1 & 1 & 5 & -1 \\ -1 & -1 & -1 & 5 \end{pmatrix}$.

Montrer que A est diagonalisable et inversible. En déduire le polynôme minimal et, à l'aide du polynôme minimal, calculer A^{-1}.

** **30**

1. Pour quelles valeurs de a_1, a_2, $a_3 \in \mathbb{C}^3$, la matrice

$$A = \begin{pmatrix} 0 & 0 & a_1 \\ 0 & 0 & a_2 \\ a_1 & a_2 & a_3 \end{pmatrix}$$

est-elle diagonalisable ? (On suppose a_1 et a_2 non tous les deux nuls, autrement le problème est trivial).

2. Soit f l'endomorphisme de $\mathbb{C}^3$ qui dans la base canonique $(e_1, ..., e_n)$ est représenté par la matrice

$$M\,(f)_{e_i} = \begin{pmatrix} & & & a_1 \\ & 0 & & \vdots \\ & & & a_{n-1} \\ a_1 & \cdots & a_{n-1} & a_n \end{pmatrix}$$

On pose $\sigma = a_1^2 + \cdots + a_{n-1}^2$ et $v = a_1\, e_1 + \cdots + a_n\, e_n$.

(a) Calculer $f(e_i), f^2(e_i), f^3(e_i)$, pour $(i = 1, ..., n)$, en fonction de σ, v, a_n. En déduire un polynôme annulateur de f.

(b) Donner une condition nécessaire et suffisante pour que f soit diagonalisable.

* **31** Soit $\{e_1, \ldots, e_n\}$ la base canonique de $\mathbb{C}^n$ et f l'endomorphisme de $\mathbb{C}^n$ défini par :

$$\begin{cases} f(e_i) &= e_{i+1} \qquad \text{(pour } i = 1, \ldots, n-1) \\ f(e_n) &= e_1 \end{cases}$$

1. Calculer $f^k(e_i)$ pour $k = 2, \ldots, n$ et $i = 1, \ldots, n$.
En déduire que f est diagonalisable. Montrer que id, f, $f^2, \ldots, f^{n-1}$ sont linéairement indépendants. En déduire le polynôme minimal.
Déterminer les valeurs propres, le polynôme caractéristique et le déterminant de f.

2. Montrer que la matrice

$$A = \begin{pmatrix} a_0 & a_{n-1} & a_{n-2} & \cdots & a_2 & a_1 \\ a_1 & a_0 & a_{n-1} & \cdots & \cdots & a_2 \\ a_2 & a_1 & a_0 & \ddots & & \vdots \\ \vdots & \vdots & & \ddots & \ddots & \vdots \\ a_{n-1} & a_{n-2} & a_{n-3} & \cdots & a_1 & a_0 \end{pmatrix} \in \mathcal{M}_n(\mathbb{C})$$

est diagonalisable. Calculer dét A. [Utiliser l'exercice 22].

32 Soit $f \in \operatorname{End}(E)$ où E est un espace vectoriel de dimension n. Montrer que si $f^{50} = 0$, alors $f^n = 0$.

33 Soit $f : \mathbb{R}_n[x] \to \mathbb{R}_n[x]$ l'application qui à tout polynôme P associe le reste de la division euclidienne de P par $Q = X^2 - 32X + 7$. Montrer que f est un endomorphisme. Montrer que f est diagonalisable.

** **34**

1. Montrer que si f est diagonalisable, alors f^2 est diagonalisable.
   (*Dans la suite il s'agit d'étudier dans quel cas on a la réciproque. Pour simplifier, on distinguera les cas* dét $f \neq 0$ *et* dét $f = 0$)
2. (a) On suppose f^2 diagonalisable et dét $f \neq 0$. Soient $\lambda_1, ..., \lambda_p$ les valeurs propres de f^2. Montrer que le polynôme $(X^2 - \lambda_1) \cdots (X^2 - \lambda_p)$ est annulateur de f et en déduire que f est diagonalisable.
   (b) On suppose f^2 diagonalisable, dét $f = 0$ et $\operatorname{Ker} f = \operatorname{Ker} f^2$.
   Soient $0, \lambda_1, ..., \lambda_p$ les valeurs propres de f^2. Montrer que
   $$f^2 \circ (f^2 - \lambda_1 \operatorname{id}) \circ \cdots \circ (f^2 - \lambda_p \operatorname{id}) = 0.$$
   En déduire que
   $$\operatorname{Im}\Big[(f^2 - \lambda_1 \operatorname{id}) \circ \cdots \circ (f^2 - \lambda_p \operatorname{id})\Big] \subset \operatorname{Ker} f^2$$
   et que
   $$f \circ (f^2 - \lambda_1 \operatorname{id}) \circ \cdots (f^2 - \lambda_p \operatorname{id}) = 0.$$
   En déduire que f est diagonalisable.
3. Montrer que si f est diagonalisable alors $\operatorname{Ker} f = \operatorname{Ker} f^2$.
   En déduire le résultat suivant :

   Soit $f \in \operatorname{End}(\mathbb{C}^n)$. *Alors :*
   f *est diagonalisable* $\iff$ $\Big(f^2$ *est diagonalisable et* $\operatorname{Ker} f = \operatorname{Ker} f^2\Big)$
4. Application. Conditions nécessaires et suffisantes pour que la matrice
   $$A = \begin{pmatrix} 0 & \cdots & \cdots & a_1 \\ & & a_2 & \\ & \cdots & & \\ a_n & \cdots & \cdots & 0 \end{pmatrix} \in \mathcal{M}_n(\mathbb{C})$$
   soit diagonalisable.

35 Soit la matrice

$$A = \begin{pmatrix} 1 & -\alpha & -\alpha & 1 \\ 1-\beta & \alpha & \alpha - 1 & -\beta \\ \beta & -\alpha & 1-\alpha & 1+\beta \\ 0 & \alpha & \alpha & 0 \end{pmatrix}, \text{ où } \alpha, \beta \in \mathbb{R}.$$

1. Condition nécessaire et suffisante pour que A soit diagonalisable.

2. A l'aide du spectre de A, montrer que $A+I$ est inversible.
   Calculer $(A+I)^{-1}$ lorsque A est diagonalisable.

* **36** Déterminer toutes les matrices $A \in \mathcal{M}_2(\mathbb{R})$ telles que :

$$A^3 - 8\,A^2 + 21\,A - 18\,I = 0.$$

37 Soit A la matrice de l'exercice 35, avec $\alpha = 1$ et $\beta = 0$. Déterminer les sous-espaces caractéristiques. Calculer A^n.
Résoudre le système différentiel $\dfrac{\mathrm{d}X}{\mathrm{d}t} = AX$.

38 Donner la décomposition de Dunford de la matrice

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

39 Soit $A \in \mathcal{M}_6(K)$. On suppose que :

$$\begin{aligned} P_A(X) &= (X-1)^4\,(X-2)^2 \\ m_A(X) &= (X-1)^2\,(X-2) \end{aligned}$$

Que peut-on dire des dimensions des espaces propres ?
Quelles sont les formes de Jordan possibles ?

40 Soit la matrice :

$$A = \begin{pmatrix} 1 & \alpha & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 2 & 3 & 1 \\ -2 & -(4+\alpha) & -4 & -1 \end{pmatrix}, \qquad \text{où } \alpha \in \mathbb{R}.$$

En discutant selon les valeurs de α, donner une réduite de Jordan ainsi qu'une matrice de passage.

* **41** **Méthode pratique pour la construction d'une base de Jordan** (*cf. page 195*)

1. On considère la matrice

$$A_2 = \begin{pmatrix} 3 & 2 & -2 \\ -1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$$

   (cf. exercice 21). Quelle est sa réduite de Jordan ? (ordonner les blocs selon l'ordre décroissant en taille). En construisant la base de Jordan, peut-on choisir comme premier vecteur le vecteur $v = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$? Déterminer une base de Jordan.

2. On considère un endomorphisme f d'un espace vectoriel de dimension p et on suppose que sa réduite de Jordan est constituée par un seul bloc $J(\lambda)$. On pose $g = f - \lambda\,\mathrm{id}$. Montrer que l'on obtient une base $\mathcal{B} = \{v_1, ..., v_p\}$ dans laquelle la matrice de f est $J(\lambda)$, en prenant :

$$\begin{cases} v_p \in E \setminus \operatorname{Ker} g^{p-1} \\ v_{p-1} = g\,(v_p) \\ v_{p-2} = g\,(v_{p-1}) \\ \quad \vdots \\ v_1 \;= g\,(v_2) \end{cases}$$

3. *Exemple.* Déterminer une réduite de Jordan et une matrice de passage pour la matrice
$$A = \begin{pmatrix} 1 & 0 & 0 & 0 \\ -1 & 4 & 1 & -2 \\ 2 & 1 & 2 & -1 \\ 1 & 2 & 1 & 0 \end{pmatrix}$$

* **42** On suppose que :
$$\begin{aligned} P_A(X) &= (-1)^\alpha\,(X-\lambda)^\alpha \\ m_A(X) &= (X-\lambda)^\beta \\ \dim E_\lambda &= \gamma \end{aligned}$$
Montrer que :
$$\frac{\alpha}{\beta} \le \gamma \le 1 + \alpha - \beta.$$

* **43** Soit :
$$A = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & b & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & a & 0 \end{pmatrix} \qquad \text{avec} \quad \begin{aligned} ab &= 0 \\ a+b &= 1 \end{aligned} \quad \text{et}$$
Donner la forme de la matrice réduite de Jordan, en discutant selon les valeurs de a et b.

INDICATIONS

1 Écrire la matrice de p_F dans une base $\mathcal{B} = \{\mathcal{B}_1, \mathcal{B}_2\}$, où $\mathcal{B}_1$ est une base de F et $\mathcal{B}_2$ une base de G.

2 Vecteurs propres : $v_1 = \alpha \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, $\alpha \in \mathbb{R}$ et $v_2 = \beta \begin{pmatrix} 1 \\ -1 \end{pmatrix}$, $\beta \in \mathbb{R}$
Valeurs propres : $1, -1$. Tr $f = 0$; dét $f = -1$.

3 Soit v vecteur propre : $f(v) = \lambda v$. On a : $0 = f^p(v) = \lambda^p v$, d'où : $\lambda = 0$.
Si f est nilpotent et diagonalisable, il est représenté, dans une base de vecteurs propres, par une matrice diagonale ayant des zeros sur la diagonale ; donc $f = 0$.

4 $a_0 = P_f(0) = \text{dét}\,(f - 0\ \text{id}) = \text{dét}\, f$.
Développer dét $(f - X\ \text{id})$ selon les éléments de la première ligne : les cofacteurs de $a_{12}, a_{13}, \dots, a_{1n}$ sont des déterminants d'ordre $n-1$ qui contiennent chacun $n-2$ termes du type $a_{ii} - X$. Par suite, les termes de degré n et $n-1$ de $P_f(X)$ proviennent du produit des termes de la diagonale principale : $(a_{11} - X)(a_{22} - X)\cdots(a_{nn} - X)$.
Donc :
$$P_f(X) = (-1)^n X^n + (-1)^{n-1}\,(a_{11} + a_{22} + \cdots + a_{nn})X^{n-1} + \cdots$$

5 Il s'agit de montrer que
$$\forall\, A \in \mathcal{M}_n(K) \text{ et } \forall\, \varepsilon > 0,\ \exists\, B \in \mathrm{GL}(n,K) \text{ telle que } ||A - B\,|| < \varepsilon$$
Soit $\mathrm{Sp}(A) = \{\lambda_1, ..., \lambda_p\}$. Si tous les λ_i sont non nuls, dét $A \neq 0$. On peut donc prendre $B = A$.
Supposons, par exemple, que $\lambda_1 = 0$. Soit $\varepsilon > 0$ arbitraire et $r > 0$ tel que $r < \underset{\mathrm{i}=2,\dots,\mathrm{n}}{\mathrm{Min}} |\lambda_i|$ et $r < \dfrac{\varepsilon}{\sqrt{n}}$. Prenons $B = A + r\,I$; les valeurs propres de B sont $\{\,\lambda_1 + r, ..., \lambda_p + r\,\}$ donc non nulles, aussi $B \in \mathrm{GL}(n,K)$. D'autre part :
$$\|\,A - B\,\| = \|\,r\,I\,\| = r\sqrt{n} < \varepsilon.$$

Application. Les applications

$$A \in \mathcal{M}_n(K) \longmapsto \text{dét}\, A \quad \text{et} \quad A \in \mathcal{M}_n(K) \longmapsto A^+$$

sont continues, car $\text{dét}\, A$ est une fonction polynomiale des coefficients de A et A^+ se calcule en calculant des déterminants. Aussi les fonctions

$$f : \begin{array}{ccc} \mathcal{M}_n(K) & \longrightarrow & K \\ A & \longmapsto & \text{dét}\, A^+ \end{array} \quad \text{et} \quad g : \begin{array}{ccc} \mathcal{M}_n(K) & \longrightarrow & K \\ A & \longmapsto & (\text{dét}\, A)^{n-1} \end{array}$$

sont continues. Puisque elles sont égales sur une partie dense, elles sont égales sur $\mathcal{M}_n(K)$.
De la même manière on voit que $A^{++} = (\text{dét}\, A)^{n-2}\, A$.

6 $\text{Sp}A = \{-1, 1, 2, 3\}$; $\quad P = \begin{pmatrix} 1 & 2 & 0 & 0 \\ -2 & -3 & 0 & 0 \\ 2 & 1 & 3 & 4 \\ -1 & 0 & -2 & -3 \end{pmatrix}$

7 $P_A(X) \in \mathbb{C}[X]$ est à coefficients réels. Donc si λ est racine d'ordre α, $\overline{\lambda}$ est aussi racine d'ordre α.

$$A\,v = \lambda\, v \implies \overline{A\,v} = \overline{\lambda v} \implies \overline{A}\,\overline{v} = \overline{\lambda}\,\overline{v}$$

Puisque A est réelle : $A\overline{v} = \overline{\lambda}\,\overline{v}$.

$$\text{Sp}A = \left\{0,\ \frac{-3 + i\sqrt{3}}{2},\ \frac{-3 - i\sqrt{3}}{2}\right\}\ ,\quad P = \begin{pmatrix} 1 & 2 & 2 \\ 1 & -1 + i\sqrt{3} & -1 - i\sqrt{3} \\ 1 & -1 - i\sqrt{3} & -1 + i\sqrt{3} \end{pmatrix}$$

8 $P_A(X) = -X^2(X-3)$ $\left(\text{se servir de } \text{Log}_a b = \dfrac{\text{Log}_a b}{\text{Log} a}\right)$.

$E_0 = (\text{Log}\ a)\,x + (\text{Log}\ b)\,y + (\text{Log}\ c)\,z$ est de dimension 2. Donc A est diagonalisable. Dans une base de vecteurs propres :

$$M(f) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 3 \end{pmatrix}$$

f est la projection sur E_3 parallèlement à E_0, suivie d'une homothétie de rapport 3.

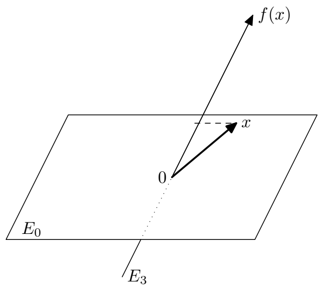


Figure 6

9 $P_{A_t}(\lambda) = \begin{vmatrix} t-\lambda & 1 & 1 & \cdots & 1 \\ 1 & t-\lambda & 1 & \cdots & 1 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \cdots & \ddots & \ddots & 1 \\ 1 & \cdots & \cdots & 1 & t-\lambda \end{vmatrix}$

Pour $t - \lambda = 1$ deux lignes (et même toutes) sont égales donc le déterminant est nul. Donc $\lambda = t - 1$ est valeur propre.
E_{t-1} est défini par $x_1 + \cdots + x_n = 0$ et donc est de dimension $n-1$. D'où : ordre $(t-1) \geq n-1$. Donc :

$$\text{Sp}' A_t = \{\underbrace{t-1, t-1, \ldots, t-1}_{n-1 \text{ fois}}, \lambda_n\}.$$

La dernière valeur propre se calcule par la trace : $\text{Tr}\ A_t = (n-1)(t-1) + \lambda_n$ d'où $\lambda_n = n + t - 1$.
On a $\lambda_n \neq t - 1$, donc ordre $(t-1) = n - 1 = \dim E_{t-1}$ et, par conséquent, A_t est diagonalisable.
$\text{dét}\, A_t = (t-1)^{n-1}(n+t-1)$; donc : $\text{dét}\, A_t = 0$ pour $t = 1$ et pour $t = 1 - n$.

10 Dans $P_A(\lambda)$ les deux premières lignes sont égales lorsque $\lambda = -a_1$. Ainsi $-a_1$ est valeur propre. De même, $-a_2, -a_3, \ldots, -a_n$ sont valeurs propres. Par la trace on a la dernière valeur propre : $a_1 + \cdots + a_n$. Donc

$$\mathrm{Sp}' A = \{-a_1, -a_2, \ldots, -a_n, a_1 + \cdots + a_n\} \text{ et } \det A = (-1)^n \left(\prod_{i=1}^{n} a_i\right)\left(\sum_{j=1}^{n} a_j\right).$$

Si les a_i sont tous deux à deux distincts et différents de $-(\sum_{i=1}^{n} a_i)$, A est diagonalisable. Pour construire une matrice diagonalisable, il suffit de prendre, par exemple, les a_i deux à deux distincts et ≥ 0. Par exemple :

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

11 $\dim E_0 = n-1$. Donc ordre$(0) \geq n-1$ et $\mathrm{Sp}' f = \{\underbrace{0, \ldots, 0,}_{n-1 \text{ fois}}\ \mathrm{Tr}\, f\}$.

f est diagonalisable si et seulement si ordre$(0) = n-1$ c'est-à-dire si et seulement si $\mathrm{Tr}\, f \neq 0$.
Pour construire une matrice non diagonalisable, on peut fixer arbitrairement la première ligne, puis se donner les autres lignes proportionnelles à la première en ayant soin de choisir le facteur de proportionnalité de la dernière ligne de manière que la trace soit nulle. Par exemple :

$$A = \begin{pmatrix} 1 & 2 & 4 & -1 \\ 2 & 4 & 8 & -2 \\ -1 & -2 & -4 & 1 \\ 1 & 2 & 4 & -1 \end{pmatrix}$$

12 Si $\lambda = 0$ est valeur propre de $f \circ g$, on a $\det(f \circ g) = 0$, donc $\det(g \circ f) = 0$; donc $\lambda = 0$ est valeur propre de $g \circ f$.
Si $\lambda \neq 0$ est valeur propre de $f \circ g$, il existe $v \neq 0$ tel que $(f \circ g)(v) = \lambda v$ d'où $g \circ f \circ g(v) = \lambda\, g(v)$. Or $g(v) \neq 0$, car si $g(v) = 0$, de l'égalité $(f \circ g)(v) = \lambda v$ on aurait $v = 0$; donc $g(v)$ vecteur propre de $g \circ f$ pour la valeur propre λ.

13 Montrer que si v est un vecteur propre de g pour la valeur propre λ $\left(v \in E_\lambda^{(g)}\right)$, on a : $f(v) \in E_\lambda(g)$. Utiliser $\dim E_\lambda^{(g)} = 1$.
Écrire l'équation $f = a_0 \,\mathrm{id} + a_1\, g + \cdots + a_{n-1}\, g^{n-1}$ dans une base de vecteurs propres commune à f et à g. On obtient un système dont les inconnues sont $a_0, a_1, \ldots, a_{n-1}$ et le déterminant est un déterminant de Vandermonde.

14

1. Trivial.
2. Soit $x_i \in F_i$. Comme $E = G_1 \oplus \cdots \oplus G_q$, $x_i = y_1 + \cdots + y_q$, avec $y_j \in G_j$. On a : $f(x_i) = \lambda_i\,(y_1 + \cdots + y_q) = f(y_1) + \cdots + f(y_q)$. Or $f(y_j) \in Gj$ (d'après 1.) et, en vertu de l'unicité de la décomposition : $f(y_j) = \lambda_i\, y_j$, donc $y_j \in F_i \cap G_j$. Ainsi, $F_i = F_i \cap G_1 + \cdots + F_i \cap G_q$.
   Montrer que la somme est directe comme dans la démonstration de la proposition 6.9, page 164.
   Puisque E est somme directe des F_i et pour chaque i F_i est somme directe des H_{ij} (non réduits à $\{0\}$), $j = 1, ..., q$, E sera somme directe des H_{ij}. En prenant une base dans chacun des H_{ij} qui ne sont pas réduits à $\{0\}$, on obtient une base de E formée de vecteurs qui sont propres aussi bien pour f que pour g. Aussi, si f et g commutent, ils sont simultanément diagonalisables.
   Réciproquement si f et g sont simultanément diagonalisables leurs matrices dans une base commune de vecteurs propres commutent, donc f et g commutent.

15

1. Puisque E est diagonalisable, E est somme directe des espaces propres :

$$E = E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_q}$$

Prendre pour P_i le projecteur sur E_{λ_i}.
Tout $x \in E$ s'écrit $x = x_1 + \cdots + x_q$ $(x_i \in E_{\lambda_i})$ d'où, en appliquant f :

$$f(x) = \lambda_1 x_1 + \cdots + \lambda_q x_q = \lambda_1 P_1 + \cdots \lambda_q P_q$$

2. On a $E = \operatorname{Im} P_1 \oplus \cdots \oplus \operatorname{Im} P_q$ (cf. exercice 10, chapitre 3)
Puisque $f = \lambda_1 P_1 + \cdots + \lambda_q P_q$ et $P_i \circ P_j$ pour $i \neq j$, on a pour tout $x \in \operatorname{Im} P_i$: $f(x) = \lambda_i P_i(x) = \lambda_i x$. Donc $\operatorname{Im} P_i \subset E_{\lambda_i}$. D'où :

$$E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_q} \supset \operatorname{Im} P_1 \oplus \cdots \oplus \operatorname{Im} P_q = E \ ;$$

donc E est somme directe d'espaces propres.

3. Les valeurs propres sont $\lambda_1 = +1$ (simple) et $\lambda_2 = -1$ (double). On a :

$$\begin{cases} I = P_1 + P_2 \\ A = P_1 - P_2 \end{cases}$$

d'où :

$$P_1 = \frac{I + A}{2} = \begin{pmatrix} 1 & 1 & -1 \\ 1 & 1 & -1 \\ 1 & 1 & -1 \end{pmatrix} \quad \text{et} \quad P_2 = \frac{I - A}{2} = \begin{pmatrix} 0 & -1 & 1 \\ -1 & 0 & 1 \\ -1 & -1 & 2 \end{pmatrix}$$

$A^n = P_1 + (-1)^n P_2$

4. Dans le cas général, on a :

$$\begin{cases} I & = & P_1 + \cdots + P_q \\ A & = & \lambda_1 P_1 + \cdots + \lambda_q P_q \\ \cdots & & \cdots \\ A^{q-1} & = & {\lambda_1}^{q-1} P_1 + \cdots + {\lambda_q}^{q-1} P_q \end{cases}$$

Le déterminant de ce système est un déterminant de Vandermonde (cf. exercice 13, chapitre 4). Ceci permet de calculer explicitement les projecteurs P_i.

Remarque - La relation $f = \lambda_1 P_1 + \cdots + \cdots \lambda_q P_q$ n'a rien de *mystérieux*. Si f est diagonalisable, on a, dans une base de vecteurs propres,

$$M(f)_{v_i} = \begin{pmatrix} \boxed{\begin{matrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_1 \end{matrix}} & & \\ & \ddots & \\ & & \boxed{\begin{matrix} \lambda_q & & \\ & \ddots & \\ & & \lambda_q \end{matrix}} \end{pmatrix} = \lambda_1 \begin{pmatrix} \boxed{\begin{matrix} 1 & & \\ & \ddots & \\ & & 1 \end{matrix}} & & \\ & \ddots & \\ & & 0 \end{pmatrix} +$$

$$\cdots + \lambda_q \begin{pmatrix} 0 & & \\ & \ddots & \\ & & \boxed{\begin{matrix} 1 & & \\ & \ddots & \\ & & 1 \end{matrix}} \end{pmatrix} = \lambda_1 P_1 + \cdots + \lambda_q P_q$$

16

1. ***Analyse***. En écrivant $\phi(f) = \lambda f$, on a :

$$\int_0^1 e^{x-t} f(t) \mathrm{d}t = \lambda f(x) \tag{1}$$

En dérivant, on trouve : $\lambda f'(x) = \lambda f(x)$. Si $\lambda \neq 0$ on a : $f(x) = Ce^x$.

Synthèse. Si $\lambda \neq 0$ et $f(x) = e^x$; en remplaçant dans (1), on trouve $\lambda = 1$. Si $\lambda = 0$ on a $\int_0^1 f(t) e^{-t} \mathrm{d}t = 0$.

Donc les valeurs propres sont 0 et 1 et les espaces propres :

$$E_0 = \left\{ f(t) \,\middle|\, \int_0^1 f(t) e^{-t} \mathrm{d}t = 0 \right\}, \quad E_1 = \{C e^x\}_{C \in \mathbb{R}}$$

2. L'analyse donne : $\lambda f'' + \lambda f = 0$.
   Pour la synthèse, utiliser l'indépendance des fonctions $\cos x$ et $\sin x$.
   On trouve les valeurs propres $0, \pi/2, -\pi/2$ avec espaces propres :

$$E_0 = \left\{ f(t) \;\middle|\; \int_{-\pi/2}^{\pi/2} f(t)\cos t \,\mathrm{d}t = \int_{-\pi/2}^{\pi/2} f(t)\sin t \,\mathrm{d}t = 0 \right\}$$

   $E_{\pi/2}$ est engendré par $\cos x$ et $E_{-\pi/2}$ est engendré par $\sin x$.

17 Écrire f dans la base $\mathcal{B} = \{1, x, x^2\}$. On trouve :

$$A := M(f)_{\mathcal{B}} = \begin{pmatrix} 1 & 1 & 0 \\ 2 & 1 & 2 \\ 0 & 1 & 1 \end{pmatrix}, \quad \mathrm{Sp}A = \{1, -1, 3\}$$

Une matrice de passage de $\mathcal{B}$ à une base de vecteurs propres est :

$$P = \frac{1}{4}\begin{pmatrix} 1 & 1 & 1 \\ 0 & -2 & 2 \\ -1 & 1 & 1 \end{pmatrix}$$

Calculer

$$\begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix} = A^n \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix}$$

et écrire le résultat sous la forme $\alpha + \beta\, x + \gamma\, x^2$.

18 $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathcal{E} \iff \quad a + d = 0 \iff \quad A = aE_1 + bE_2 + cE_3 \quad$ avec

$$E_1 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad E_2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad E_3 = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}.$$

$\mathcal{B} = \{E_1, E_2, E_3\}$ est une base de $\mathcal{E}$. Soit $C := M(f)_{\mathcal{B}} = \begin{pmatrix} 0 & 2 & 0 \\ 0 & 2 & 0 \\ -4 & 0 & -2 \end{pmatrix}$.

On trouve :

$$C^n = 2^n \begin{pmatrix} 0 & 1 & 0 \\ 0 & 1 & 0 \\ 2(-1)^n & -(1+(-1)^n) & (-1)^n \end{pmatrix}$$

d'où : $f^n(A) = 2^n \begin{pmatrix} b & b \\ 2\,a\,(-1)^n - b\,(1+(-1)^n) + c\,(-1)^n & -b \end{pmatrix}$

19 $x = C_2\, e^{-2t} + 2\, C_3\, e^t, \quad y = -C_2\, e^{-2t} - C_3\, e^t, \quad z = C_1\, e^{-5t} + C_2\, e^{-2t}.$

20 On a : $\quad {}^t\mathrm{Cof}(A - \lambda\, I) \cdot (A - \lambda\, I)\, X = P_A(\lambda)\, X \qquad (*)$

où $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ et le système s'écrit : $\left(A - I\,\frac{\mathrm{d}}{\mathrm{d}t}\right) X = 0.$

Prendre la i-ième composante de (*) et remplacer λ par $\frac{\mathrm{d}}{\mathrm{d}t}$.

Dans l'exemple x_1 et x_2 vérifient l'équation différentielle $y'' - 2\,y' + y = 0$, dont la solution générale est $y = e^t(c_1 + c_2\, t)$. Donc x_1 et x_2 sont de la forme :

$$\begin{cases} x_1 &= e^t\,(a + b\,t\,) \\ x_2 &= e^t\,(c + d\,t\,) \end{cases}$$

En remplaçant, par exemple, dans la première équation, on trouve :

$$\left\{\begin{array}{ll} 3\,b+2\,d & = b \\ 3\,a+2\,c & = a+b \end{array}\right. \quad \text{d'où :} \quad \left\{\begin{array}{ll} d & = -b \\ c & = -a+\dfrac{b}{2} \end{array}\right.$$

ce qui donne :

$$\left\{\begin{array}{ll} x_1 & = e^t\,(a+b\,t\,) \\ x_2 & = e^t\,(-a+\dfrac{b}{2}-b\,t\,) \end{array}\right.$$

Nota. Cette méthode permet de résoudre le système sans chercher les vecteurs propres (ou, le cas échéant, les espaces caractéristiques : cf exercice 37) .

21

1. $\mathrm{Sp}'A_1 = \{2, 2, 1\}$.
   Si f est l'endomorphisme associé à A_1 dans la base canonique, il existe une base $\{v_i\}$ telle que

$$M(f)_{v_i} = \begin{pmatrix} 2 & a & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

   E_2 est engendré par $v_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$; comme $\dim E_2 = 1$, on a $a \neq 0$.
   En prenant $a \neq 0$, on cherche v_2 tel que $f(v_2) = a\,v_1 + 2\,v_2$. En prenant par exemple $a = 1$, on trouve : $v_2 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$. E_1 est engendré par $v_3 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$. Donc :

$$A_1' = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad P = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 0 & 3 \\ 0 & 1 & 1 \end{pmatrix}$$

   $\mathrm{Sp}'A_2 = \{1, 1, 1\}$; $E_1 : x + y - z = 0$. On prend $v_1, v_2 \in E_1$, par exemple : $v_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$, $v_2 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$, et on complète en une base en choisissant, par exemple $v_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$. Ainsi : $A_2' = M(f)_{v_i} = \begin{pmatrix} 1 & 0 & a \\ 0 & 1 & b \\ 0 & 0 & 1 \end{pmatrix}$.
   On calcule a, b en imposant que $f(v_3) = a\,v_1 + b\,v_2 + v_3$. On obtient $a = -2, b = 1$.

   $\mathrm{Sp}'A_3 = \{9, 9, 9\}$. $\dim E_9 = 1$. E_9 est engendré par $v_1 = \begin{pmatrix} 2 \\ 2 \\ -1 \end{pmatrix}$. Il existe une base $\{v_i\}$ telle que :

$$M(f)_{v_i} = \begin{pmatrix} 9 & a & b \\ 0 & 9 & c \\ 0 & 0 & 9 \end{pmatrix}.$$

   En imposant $f(v_2) = a\,v_1 + 9\,v_2$ on voit que l'on peut prendre, par exemple, $a = 1$ et $v_2 = \begin{pmatrix} -1/3 \\ -2/3 \\ 0 \end{pmatrix}$. (*Attention* : on n'a pas le droit de multiplier v_2 par 3, car v_2 est solution d'un système non homogène).
   On complète $\{v_1, v_2\}$ en une base en prenant, par exemple : $v_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$.
   On trouve $b = 18$, $c = 2$.

2. On résout : $\dfrac{\mathrm{d}X'}{\mathrm{d}t} = A_1'\,X'$ et on revient à X par $X = PX'$.

$$\begin{cases} \frac{\mathrm{d}x'}{\mathrm{d}t} = & 2\,x' + y' \\ \frac{\mathrm{d}y'}{\mathrm{d}t} = & 2\,y' \\ \frac{\mathrm{d}z'}{\mathrm{d}t} = & z' \end{cases} \to \begin{cases} x' & = (C_2\,t + C_3)e^{2t} \\ y' & = C_2\,e^{2t} \\ z' & = C_1\,e^{t} \end{cases}$$

$$\to \begin{cases} x & = (C_2\,t + C_3)\,e^{2t} + C_3\,e^{t} \\ y & = (C_2\,t + C_3)\,e^{2t} + 3\,C_3\,e^{t} \\ z & = C_2\,e^{2t} + C_3\,e^{t} \end{cases}$$

22 Si v est vecteur propre de f pour la valeur λ, on a : $Q(f)(v) = Q(\lambda)\,v$.

23 En dérivant le système $\frac{\mathrm{d}X}{\mathrm{d}t} = AX$, on a par récurrence : $\left(\frac{\mathrm{d}}{\mathrm{d}t}\right)^p X = A^p X$. D'autre part, si $P_A(\lambda) = (-1)^n \lambda^n + a_{n-1}\lambda^{n-1} + \cdots + a_1\lambda + a_0$, on a : $(-1)^n\,A^n + a_{n-1}\,A^{n-1} + \cdots + a_1\,A + a_0\,I = 0$. Appliquer à X et tenir compte du fait que $A^p X = \left(\frac{\mathrm{d}}{\mathrm{d}t}\right)^p X$.

24
1. Si $\operatorname{Sp} A = \{\lambda_1, ..., \lambda_p\}$, on a : $\operatorname{Sp} P_B(A) = \{P_B(\lambda_1), ..., P_B(\lambda_p)\}$ (cf. exercice 22), donc toutes les valeurs propres de $P_B(A)$ sont non nulles.
2. On a $A^k\,M = M\,B^k$ pour tout $k \in \mathbb{N}$, etc.
3. Écrire l'égalité de 2. pour $Q(X) = P_B(X)$. Utiliser le théorème de Cayley-Hamilton et 1.

25 On a $\mathrm{pr}_1 \circ \mathrm{pr}_2 = 0$ puisque $Q_1(f) \circ Q_2(f) = 0$. D'autre part, d'après (*) page 180, on a : $x = \mathrm{pr}_1(x) + \mathrm{pr}_2(x)$, $\forall x \in E$, d'où $\mathrm{pr}_1(x) = \mathrm{pr}_1^2(x)$. Ceci montre que pr_1 est le projecteur sur $\operatorname{Ker} Q_1(f)$. Même démonstration pour pr_2.
Il est clair que l'algorithme de calcul des polynômes U_1 et U_2 du Théorème de Bézout permet de calculer facilement les projecteurs pr_1 et pr_2.

26 Puisque A est à coefficients réels, si i est valeur propre, alors $-i$ est aussi valeur propre. Donc $\operatorname{Sp}' A = \{i, -i, 1, \lambda\}$ et $\operatorname{Tr} A = i - i + 1 + \lambda$. D'où $\lambda = 1$. Le polynôme caractéristique de A est donc $P_A(X) = (X-1)^2(X^2+1)$. Puisque $X-1$ et X^2+1 sont premiers entre eux, on en déduit immédiatement que $\mathbb{R}^4 = \operatorname{Ker}(A-I)^2 \oplus \operatorname{Ker}(A^2+I)$.
Noter que le Lemme des Noyaux s'applique aussi si les polynômes Q_i ne sont pas scindés.

27 $m_{A_1}(X) = (X-2)^2(X-1)$, $\quad m_{A_2}(X) = (X-1)^2$, $\quad m_{A_3}(X) = (X-9)^3$.

28 $P_A(X) = -(X-1)^2(X-c)$.
Si $c \neq 1$: $m_A(X) = (X-1)\,(X-c)$ ou $m_A(X) = (X-1)^2\,(X-c)$. Donc A est diagonalisable si et seulement si $m_A(X) = (X-1)(X-c)$, c'est-à-dire si et seulement si $(A-I)(A-cI) = 0$. On trouve que A diagonalisable si et seulement si $a = 0$, et b quelconque.
Si $c = 1$: $m_A(X) = X - 1$ ou $(X-1)^2$ ou $(X-1)^3$. Donc A est diagonalisable si et seulement si $m_A(X) = X - 1$, c'est-à-dire si et seulement si $A = I$, ce qui est exclu.
En resumant : A est diagonalisable si et seulement si $c \neq 1$, $a = 0$.

29 Comme dans l'exercice 9, $\operatorname{Sp}' A = \{4, 4, 4, 8\}$, $\dim E_4 = 3$. Donc A est diagonalisable et par conséquent : $m_A(X) = (X-4)(X-8)$ d'où : $(A - 4\,I)(A - 8\,I) = 0$, c'est-à-dire : $A^2 - 12\,A + 32\,I = 0$.
Comme il n'y a pas de valeur propre nulle, A est inversible. En multipliant par A^{-1} cette équation, on trouve :

$$A^{-1} = \frac{1}{32}(12\,I - A) = \frac{1}{32}\begin{pmatrix} 7 & -1 & -1 & 1 \\ -1 & 7 & -1 & 1 \\ -1 & -1 & 7 & 1 \\ 1 & 1 & 1 & 7 \end{pmatrix}$$

30

1. $P_A(X) = -X\Big(X^2 - c\,X - (a^2+b^2)\Big)$. Si $a^2+b^2 \neq 0$ et $c^2 + 4\,(a^2+b^2) \neq 0$, A admet trois valeurs propres distinctes, donc elle est diagonalisable.
Soit $a^2+b^2 = 0$; on a : $P_A(X) = -X^2(X-c)$. Si $c=0$, A est diagonalisable si et seulement si $m_A(X) = X$, c'est-à-dire $A=0$, ce qui est exclu. Si $c \neq 0$, A est diagonalisable si et seulement si $A(A - c\,I) = 0$, ce qui est aussi exclu. Donc si $a^2+b^2=0$, A n'est pas diagonalisable.
Etude analogue pour $c^2 + 4\,(a^2+b^2) = 0$.
2. (a) On trouve $f^3 = a_n f^2 + \sigma f$. Polynôme annulateur :
$$Q(X) = X\,(X^2 - a_n\,X - \sigma)$$
(b) Si $\sigma \neq 0$ et $a_n^2 + 4\,\sigma \neq 0$, Q n'a que des racines simples et donc A est diagonalisable.
Si $\sigma = 0$, alors $Q(X) = X^2(X - a_n)$. Si $a_n = 0$, A est diagonalisable si et seulement si $A = 0$, ce qui est exclu, etc. (raisonner comme dans 1.).

31

1. $f^k(e_i) = e_{i+k\,[\mathrm{mod}\,n]}$. Donc $f^n(e_i) = e_i$, c'est-à-dire $f^n = \mathrm{id}$.
Donc $Q(X) = X^n - 1$ est annulateur de f. Puisqu'il est scindé dans $\mathbb{C}$ et il a toutes ses racines simples, f est diagonalisable.
Si $b_0\,\mathrm{id} + b_1\,f + b_2\,f^2 + \cdots + b_{n-1}\,f^{n-1} = 0$, en appliquant cette expression sur e_1, on trouve $b_0 = b_1 = b_2 = \cdots = b_{n-1} = 0$. Donc $d^0 m_f(X) = n$ et par conséquent : $m_f(X) = X^n - 1$ et $P_f(X) = (-1)^n(X^n - 1)$.
On en déduit :
$$\mathrm{Sp} f = \{\omega_0, \omega_1, \ldots, \omega_{n-1}\}, \quad \omega_k = e^{\frac{2i\pi}{k}}$$
2. Soit g l'endomorphisme associé à A dans la base canonique. On a : $g = P(f)$ avec $P = a_0 + a_1\,X + \cdots + a_{n-1}\,X^{n-1}$. Donc g est diagonalisable et :
$$\mathrm{Sp}\,g = \{P(\omega_0), P(\omega_1), \ldots, P(\omega_{n-1})\}\ , \quad \mathrm{dét}\,A = P(\omega_0)P(\omega_1)\cdots P(\omega_{n-1}).$$

32 X^{50} est annulateur. Donc $m_f(X) = X^\beta$ (avec $\beta \leq 50$) et, par conséquent $P_f(X) = (-1)^n X^n$. D'après Cayley-Hamilton : $f^n = 0$.

33 $f^2 = f$. Donc $X(X-1)$ est annulateur.

34

1. Tout vecteur propre de f est vecteur propre de f^2, donc, si f est diagonalisable, f^2 est diagonalisable (dans la même base de vecteurs propres).
2. (a) Puisque f^2 est diagonalisable, $m_{f^2}(X) = (X - \lambda_1)\cdots(X-\lambda_p)$. Utiliser $m_{f^2}(f^2) = 0$. Comme toutes les λ_i sont différentes de zéro, le polynôme
$$Q(X) = \Big(X - \sqrt{\lambda_1}\Big)\Big(X + \sqrt{\lambda_1}\Big)\cdots\cdots\Big(X - \sqrt{\lambda_p}\Big)\Big(X + \sqrt{\lambda_p}\Big)$$
a toutes ses racines simples.
(b) Puisque f^2 est diagonalisable, $m_{f^2}(X) = X(X-\lambda_1)\cdots(X-\lambda_p)$, d'ou $f^2 \circ (f^2 - \lambda_1\ \mathrm{id}) \circ \cdots \circ (f^2 - \lambda_p\ \mathrm{id}) = 0$, c'est-à-dire
$$\mathrm{Im}\left[(f^2 - \lambda_1\ \mathrm{id}) \circ \cdots \circ (f^2 - \lambda_p\ \mathrm{id})\right] \subset \mathrm{Ker}\,f^2$$
Utiliser $\mathrm{Ker}\,f = \mathrm{Ker}\,f^2$.
3. En ecrivant les matrices de f et f^2 dans une base de vecteurs propres, on voit facilement que $\mathrm{Ker}\,f = \mathrm{Ker}\,f^2$.
4. Application. A^2 est diagonale, donc A est diagonalisable si et seulement si $\mathrm{Ker}\,A = \mathrm{Ker}\,A^2$. Cette condition s'exprime sur les coefficients a_i par :
$$a_i = 0 \Longleftrightarrow a_{n-i+1} = 0 \qquad (i = 1, \ldots n)$$

35

1. $P_A(X) = X^2(X-1)^2$; donc A est diagonalisable si et seulement si $A\,(A-I) = 0$. Ce qui équivaut à $\alpha = 0$, $\beta = 0$.
2. $\mathrm{Sp}'(A+I) = \{1,1,2,2\}$, donc $\det(A+I) \neq 0$. Soit $B = A+I$; B est diagonalisable, donc $(B-I)\,(B-2I) = 0$; d'où : $B^{-1} = \dfrac{1}{2}(3\,I - B) = I - \dfrac{1}{2}A$.

36

$Q(X) = X^3 - 8X^2 + 21X - 18 = (X-2)(X-3)^2$ est annulateur ; donc $m_A(X)$ divise $Q(X)$ et $d^\circ m_A(X) \leq 2$. On a 4 cas possibles :

- $m_A(X) = X-2$, ce qui donne $A = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$.
- $m_A(X) = X-3$, qui donne $A = \begin{pmatrix} 3 & 0 \\ 0 & 3 \end{pmatrix}$
- $m_A(X) = (X-2)(X-3)$, d'où : $P_A(X) = (X-2)(X-3) = X^2 - 5\,X + 6$, ce qui équivaut à $\operatorname{Tr} A = 5$ et $\det A = 6$, c'est-à-dire :

$$A = \begin{pmatrix} a & b \\ c & 5-a \end{pmatrix} \quad \text{avec} \quad a\,(5-a) - b\,c = 6$$

- $m_A(X) = (X-3)^2$ ce qui équivaut à $\operatorname{Tr} A = 6$ et $\det A = 9$, etc.

37

A est semblable à $A' = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$ avec matrice de passage :

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

L'espace caractéristique N_0 est engendré par les deux premières colonnes de P et l'espace caractéristique N_1 par les deux dernières colonnes.

$$(A')^n = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad \text{pour } n \geq 2$$

Calculer $P(A')^n P^{-1}$.

38

On a $P_A(X) = -(X-1)\,(X-2)^2$. En effectuant une réduction en blocs triangulaires on voit que A est semblable à la matrice

$$A' = \begin{pmatrix} \boxed{\begin{matrix} 2 & 1 \\ 0 & 2 \end{matrix}} & \begin{matrix} 0 \\ 0 \end{matrix} \\ \begin{matrix} 0 & 0 \end{matrix} & \boxed{1} \end{pmatrix}$$

Donc

$$A' = D' + N' = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix} + \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

une matrice de passage étant $P = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}$. On a : $P^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 1 & -1 & 1 \\ -1 & 1 & 0 \end{pmatrix}$,

donc :

$$N = P\,N'\,P^{-1} = \begin{pmatrix} 1 & -1 & 1 \\ 1 & -1 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \quad \text{d'où :} \quad D = A - N = \begin{pmatrix} 2 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & -1 & 2 \end{pmatrix}$$

39 Les formes de Jordan possibles sont :

$$\begin{pmatrix} \boxed{\begin{matrix}1&1\\0&1\end{matrix}} & & & & \\ & \boxed{1} & & & \\ & & \boxed{1} & & \\ & & & \boxed{2} & \\ & & & & \boxed{2} \end{pmatrix} \quad \text{et} \quad \begin{pmatrix} \boxed{\begin{matrix}1&1\\0&1\end{matrix}} & & & \\ & \boxed{\begin{matrix}1&1\\0&1\end{matrix}} & & \\ & & \boxed{2} & \\ & & & \boxed{2} \end{pmatrix}$$

En particulier : $\dim E_1 = 3$ ou $\dim E_1 = 2$.

40 $P_A(X) = (X-1)^4, \quad m_A(X) = (X-1)^2$

$$E_{(1)} : \begin{cases} \alpha\, y = 0 \\ x + 2\,y + 2\,z + t = 0 \end{cases} \quad \text{et} \quad \dim E_{(1)} = \begin{cases} 3 & \text{si} \quad \alpha = 0 \\ 2 & \text{si} \quad \alpha \neq 0 \end{cases}$$

$$\text{Donc : si } \alpha = 0 \quad A \text{ est semblable à } A' = \begin{pmatrix} \boxed{\begin{matrix}1&1\\0&1\end{matrix}} & \begin{matrix}0&0\\0&0\end{matrix} \\ \begin{matrix}0&0\\0&0\end{matrix} & \begin{matrix}\boxed{1}&0\\0&\boxed{1}\end{matrix} \end{pmatrix}$$

$$\text{si } \alpha \neq 0 \quad A \text{ est semblable à } A' = \begin{pmatrix} \boxed{\begin{matrix}1&1\\0&1\end{matrix}} & \begin{matrix}0&0\\0&0\end{matrix} \\ \begin{matrix}0&0\\0&0\end{matrix} & \boxed{\begin{matrix}1&1\\0&1\end{matrix}} \end{pmatrix}$$

Matrice de passage :

$$\text{si } \alpha = 0 : \qquad P = \begin{pmatrix} 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & -1 & 0 \\ -2 & 0 & 0 & -1 \end{pmatrix}$$

$$\text{si } \alpha \neq 0 : \qquad P = \begin{pmatrix} 0 & 1 & 1 & -2/\alpha \\ 0 & 0 & 0 & 1/\alpha \\ 1 & 0 & 0 & 0 \\ -2 & 0 & -1 & 0 \end{pmatrix}$$

41

1. Réduite de Jordan : $\begin{pmatrix} \boxed{\begin{matrix}1&1\\0&1\end{matrix}} & \begin{matrix}0\\0\end{matrix} \\ \begin{matrix}0&0\end{matrix} & \boxed{1} \end{pmatrix}$.

   E_1 est le plan $x + y - z = 0$. Si $\{v_1, v_2, v_3\}$ est la base de Jordan on doit avoir $v_1,\, v_3 \in E_1$ et v_2 tel que $A\,v_2 = v_1 + v_2$. Si on prend $v_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$, on aboutit pour v_2 au système
   $$\begin{cases} 2\,x + 2\,y - 2\,z = 1 \\ -x \;\; - y \;\; + z = 0 \end{cases}$$
   qui est incompatible. Prendre $v_1 = \begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix}$, etc.

   Nota. C'est pour éviter ce type de problèmes que l'on utilise la construction ci-dessous.

2. Simple vérification

3. Réduite de Jordan :

$$A' = M(f)_{v_i} = \begin{pmatrix} \boxed{\begin{matrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{matrix}} & \begin{matrix} 0 \\ 0 \\ 0 \end{matrix} \\ \begin{matrix} 0 & 0 & 0 \end{matrix} & \boxed{1} \end{pmatrix}$$

Les vecteurs $\{v_1, v_2, v_3\}$ engendrent le sous-espace caractéristique N_2 ; le vecteur v_4 engendre le sous-espace propre E_1 : $v_4 = \begin{pmatrix} 1 \\ 1 \\ -4 \\ -1 \end{pmatrix}$.

Détermination de N_2. On a $N_2 = \operatorname{Ker}(A - 2\,I)^3$. En effectuant le calcul on voit que, dans les coordonnées $\{x, y, z, t\}$, N_2 est défini par l'équation $x = 0$.

Détermination de v_3. On prend pour v_3 un vecteur quelconque de $N_2 \setminus \operatorname{Ker}(A - 2I)^2$. Le calcul montre que $\operatorname{Ker}(A - 2\,I)^2$ est défini par le système :

$$\begin{cases} x & = 0 \\ -x \ + y - t = 0 \\ 3\,x \ + y - t = 0 \end{cases} \quad \text{c'est-à-dire :} \quad \begin{cases} x = 0 \\ y = t \end{cases} .$$

On peut prendre, par exemple : $v_3 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}$.

On prend ensuite $v_2 = (A - 2\,I)\,v_3 = \begin{pmatrix} 0 \\ 2 \\ 1 \\ 2 \end{pmatrix}$ et $v_1 = (A - 2\,I)\,v_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix}$.

Ainsi une matrice de passage à la forme de Jordan est :

$$P = \|\, v_1, v_2, v_3, v_4 \,\| = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 1 & 2 & 1 & 1 \\ 0 & 1 & 0 & -4 \\ 1 & 2 & 0 & -1 \end{pmatrix}.$$

42 La matrice est d'ordre α et décomposée en γ blocs d'ordre respectivement : $\beta_1 = \beta,\ \beta_2,\ \beta_3, \ldots,\ \beta_\gamma$. Or $\beta_i \geq 1$ (pour $i = 2, \ldots, \gamma$). Donc :

$$\alpha = \beta_1 + \beta_2 + \cdots + \beta_\gamma \geq \beta + \underbrace{1 + \cdots + 1}_{\gamma-1 \text{ fois}} = \beta + \gamma - 1$$

D'autre part : $\beta_i \leq \beta$; donc $\alpha \leq \beta\,\gamma$. Ainsi : $\beta + \gamma - 1 \leq \alpha \leq \beta\,\gamma$.

43 On a $A^3 = 0$; $P_A(X) = -X^7$, $m_A(X) = X^3$, $\dim E_0 = 3$ quels que soient a et b. Pour connaître la forme de Jordan, il faut calculer n_3, n_2 et n_1 à l'aide des formules de la proposition 6.36.
Si $a = 1$ et $b = 0$, on trouve $n_3 = 2$, donc $n_2 = 0$ et $n_1 = 1$;
si $a = 0$ et $b = 1$, on trouve $n_3 = 1$ et $n_2 = 2$, donc $n_1 = 0$.

Chapitre 7

Espaces euclidiens

La notion d'espace vectoriel ne constitue que le cadre général de l'algèbre linéaire : il s'agit de la structure minimale qui permet de traiter les problèmes linéaires. La richesse de la théorie s'accroît notablement si à la structure d'espace vectoriel on ajoute des structures supplémentaires qui permettent de rendre compte d'autres propriétés remarquables que l'on rencontre naturellement.
La plus importante de ces structures additionnelles est celle qui fait référence aux questions concernant la mesure des longueurs, des distances, des volumes, des angles, etc : c'est ce que l'on appelle les notions *métriques* (du grec $\mu\acute{\varepsilon}\tau\rho o\nu$ = mesure). De même que la notion d'espace vectoriel permet de donner un sens, de définir et étudier les phénomènes linéaires, la notion de *produit scalaire* permet de donner un sens, définir et étudier les propriétés métriques dans les espaces vectoriels.

7.1 Produit scalaire canonique dans $\mathbb{R}^2$ et $\mathbb{R}^3$

Dans ce paragraphe nous donnons les notions intuitives de "norme", "angle", "orientation", de manière à assimiler plus facilement les notions plus précises qui seront définies par la suite.
On sait qu'en physique ou en mécanique l'on introduit sur les vecteurs du plan $\mathbb{R}^2$ ou de l'espace $\mathbb{R}^3$ une opération dite produit scalaire définie de la manière suivante :
Si $x = (x_1, x_2, x_3)$ *et* $y = (y_1, y_2, y_3)$, *on appelle produit scalaire de* x *par* y *le scalaire :*

$$\langle x, y\rangle := x_1 y_1 + x_2 y_2 + x_3 y_3$$

Le produit scalaire jouit des propriétés suivantes :

i) Il est ***bilinéaire***, c'est-à-dire linéaire en chaque argument :

$$\begin{array}{rcl} \langle x+y, z\rangle &=& \langle x, z\rangle + \langle y, z\rangle \\ \langle \lambda x, z\rangle &=& \lambda\langle x, z\rangle \\ \langle x, y+z\rangle &=& \langle x, y\rangle + \langle x, z\rangle \\ \langle x, \lambda z\rangle &=& \lambda\langle x, z\rangle\ , \qquad\qquad \forall x, y, z \in \mathbb{R}^3,\ \forall \lambda \in \mathbb{R} \end{array} .$$

ii) Il est ***symétrique***, c'est-à-dire :

$$\langle x, y\rangle = \langle y, x\rangle\ , \qquad \forall x, y \in \mathbb{R}^3.$$

iii) Il est ***défini positif***, c'est-à-dire :

$$\begin{array}{ll} \langle x, x\rangle & \geq 0\ ,\ \forall\, x \in \mathbb{R}^3 \quad \text{et} \\ \langle x, x\rangle & = 0 \Longleftrightarrow x = 0 \end{array} .$$

Bien entendu, le produit scalaire vérifie d'autres propriétés, mais en fait toutes les notions "métriques" peuvent être introduites en tenant compte de celles-ci.
Nous allons rappeler ici brièvement quelques-unes de ces notions.

1. Norme

On appelle ***longueur*** ou ***norme*** du vecteur x le scalaire :

$$\boxed{\| x \| := \sqrt{\langle x,\, x \rangle} = \sqrt{x_1^2 + x_2^2 + x_3^2}}$$

La propriété iii), assure que l'on peut prendre la racine carrée et qu'un vecteur est nul si et seulement si sa longueur est nulle.
Cette notion correspond à celle intuitive déduite du théorème de Pythagore (cf. fig. 1).

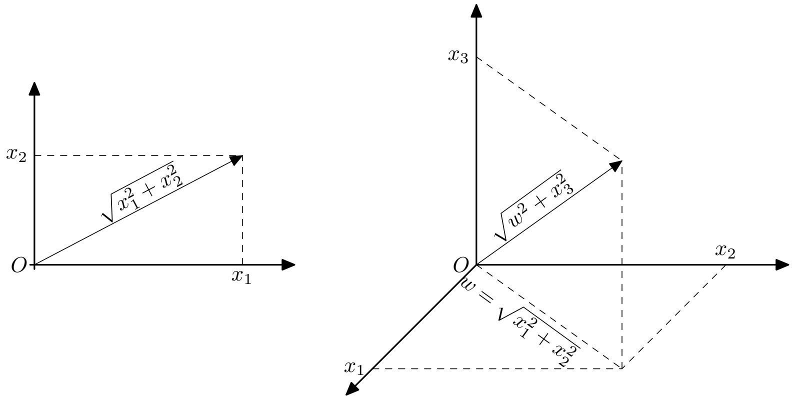


Figure 1

2. Orthogonalité

On dit que deux vecteurs v, w sont ***orthogonaux*** si

$$\boxed{\langle v,\, w \rangle \;=\; 0}$$

Cette définition coïncide avec la notion intuitive d'orthogonalité. Soient en effet v et w deux vecteurs ; on peut exprimer le fait qu'ils sont orthogonaux par la condition (cf. figure 2.) :

$$\|v + w\,\| = \|v - w\,\|$$

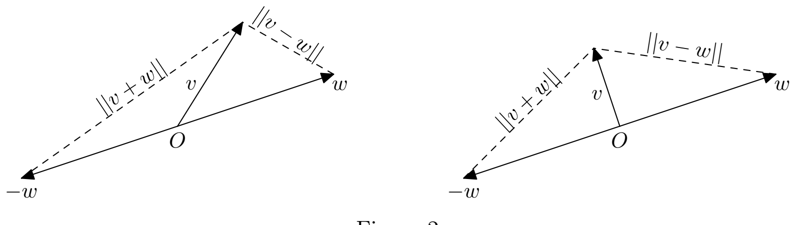


Figure 2

Or, d'après les propriétés i), ii) :

$$\begin{array}{rlll} & \|v + w\|^2 & = \langle v + w,\, v + w \rangle & = \|v\|^2 + \|w\|^2 + 2\,\langle v,\, w \rangle \\ \text{et} & \|v - w\|^2 & = \langle v - w,\, v - w \rangle & = \|v\|^2 + \|w\|^2 - 2\,\langle v,\, w \rangle \end{array}$$

Donc la condition $\|v + w\| = \|v - w\|$ est équivalente à $\langle v,\, w \rangle = 0$.

3. Angle (non orienté) dans $\mathbb{R}^2$ ou $\mathbb{R}^3$

Si v et w sont deux vecteurs non nuls, on définit d'abord la projection orthogonale de w sur v comme étant le scalaire λ tel que le vecteur $j = w - \lambda v$ soit orthogonal à v (cf fig. 3). On trouve :

$$\langle w - \lambda v,\, v\rangle = 0 \implies \lambda = \frac{\langle v,\, w\rangle}{\|v\|^2}.$$

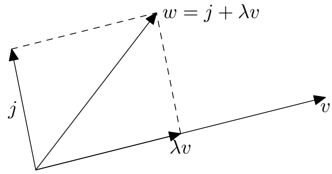


Figure 3

Si v et w sont deux vecteurs non nuls, l'angle non orienté entre v et w est l'angle compris entre 0 et π que forment les vecteurs de longueur 1

$$v_1 = \frac{v}{\|v\|} \quad \text{et} \quad w_1 = \frac{w}{\|w\|}$$

D'après la trigonométrie, le cosinus est défini par :

$$\cos\theta = \lambda$$

λ étant la projection orthogonale de w_1 sur v_1 :

$$\lambda \;=\; \frac{\langle v_1,\, w_1\rangle}{\|v_1\|^2} \;=\; \langle v_1,\, w_1\rangle \;=\; \frac{\langle v,\, w\rangle}{\|v\|\;\|w\|}$$

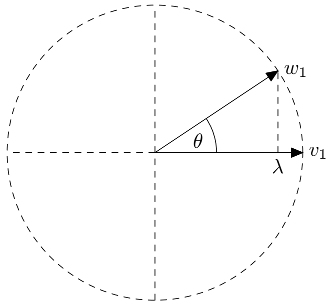


Figure 4

Donc

$$\boxed{\cos\theta = \frac{\langle v,\, w\rangle}{\|v\|\;\|w\|}} \qquad (1)$$

Ceci définit ***l'angle (non orienté)*** compris entre 0 et π. C'est cette formule qui inspirera, par la suite, la définition d'angle dans un cadre plus général (cf. page 233).

4. Angle orienté dans $\mathbb{R}^2$

Etant donné un couple de vecteurs non nuls de $\mathbb{R}^2$, $\{v, w\}$, on appelle ***angle orienté entre v et w*** celui qui est compté de v à w dans le sens trigonométrique. Cet angle, qui est compris entre 0 et 2π, est noté $(v,\, w)$.
On voit facilement que si $\theta = (v,\, w)$ et $\{e_i\}$ est la base canonique :

$$\boxed{\sin\theta = \frac{\text{dét}\,\|v,\, w\|_{e_i}}{\|v\|\;\|w\|}} \qquad (2)$$

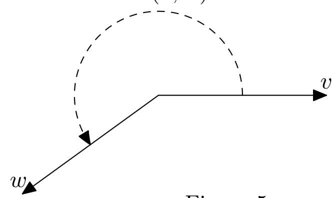


Figure 5

Soient en effet $v = \|v\| \begin{pmatrix} \cos\alpha \\ \sin\alpha \end{pmatrix}$ et $w = \|w\| \begin{pmatrix} \cos\beta \\ \sin\beta \end{pmatrix}$ (α et β étant les angles orientés que v et w forment avec

$$\begin{aligned} \text{dét}\,\|v,\, w\|_{e_i} &= \|v\|\,\|w\| \begin{vmatrix} \cos\alpha & \cos\beta \\ \sin\alpha & \sin\beta \end{vmatrix} \\ &= \|v\|\,\|w\|\,(\cos\alpha\,\cos\beta - \sin\alpha\,\sin\beta) \\ &= \|v\|\,\|w\|\,\sin(\beta-\alpha) = \|v\|\,\|w\|\,\sin\theta \end{aligned}$$

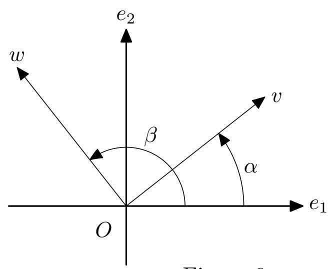


Figure 6

En particulier, si $\mathcal{A}(v,\, w)$ désigne l'aire du parallélogramme engendré par v et w, on a (cf. proposition 4.29) :

$$\boxed{\mathcal{A}(v,\, w) = \|v\|\,\|w\|\,|\sin\theta|} \tag{3}$$

(ou, ce qui revient au même : $\mathcal{A}(v,\, w) = \|v\|\,\|w\|\sin\theta$, θ désignant ici l'angle *non* orienté).

Remarque. – On a $\sin\theta > 0 \Longleftrightarrow \text{dét}\,\|v,\, w\|_{e_i} > 0$, c'est-à-dire $\sin\theta > 0$, si et seulement si la base $\{v,\, w\}$ a la même orientation que la base canonique (cf. définition 4.31.)

5. Angle orienté dans $\mathbb{R}^3$

La définition d'angle orienté ne s'étend pas immédiatement à deux vecteurs de $\mathbb{R}^3$.
Soient en effet v, w deux vecteurs indépendants de $\mathbb{R}^3$ et π le plan qu'ils engendrent. Pour pouvoir parler d'angle orienté il faut d'abord préciser ce que l'on doit entendre par «sens trigonométrique», car tout dépend – pour le dire en termes expressifs – du côté duquel on regarde le plan (cf fig. 7).

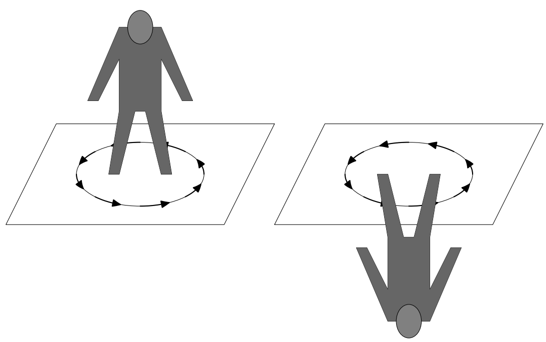


Figure 7

D'une manière plus précise, pour pouvoir parler d'angle orienté il faut fixer au préalable une orientation de la normale au plan $\pi = \text{Vect}\{v,\, w\}$, c'est-à-dire choisir un vecteur directeur $\vec{n}$ de la normale à π. Le choix de $\vec{n}$ détermine, conformément à la définition 4.34, une orientation de π : sont dites orientées positivement les bases $\{v_1,\, v_2\}$ de π telles que $\text{dét}\,\|v_1, v_2, \vec{n}\,\| > 0$. Si $\{v_1, v_2, \vec{n}\,\}$ est une telle base, on appelle *sens trigonométrique* ou *sens positif* celui qui est compté de v_1 à v_2 pour décrire un angle plus petit que π (c'est ce que l'on appelle en physique «la règle d'Ampère» ou aussi du «*tire-bouchon*» (cf. figure 8).

Ceci dit, une fois $\vec{n}$ choisi, on appelle ***angle orienté*** entre deux vecteurs non nuls v, w celui qui est compté de v à w dans le sens positif défini par $\vec{n}$. Cet angle est noté $(v,\, w)_n$.
Si $\theta = (v,\, w)_n$, on a :

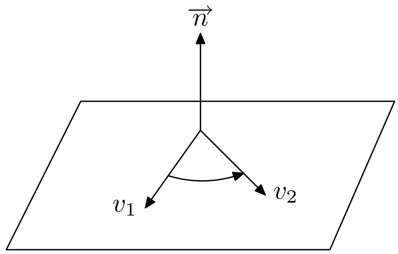


Figure 8

$$\boxed{\sin\theta = \frac{\text{dét}\,\|v, w, \vec{n}\,\|_{e_i}}{\|v\|\,\|w\|\,\|\vec{n}\,\|}} \tag{4}$$

En effet :

$$|\,\text{dét}\,\|v, w, \vec{n}\,\|_{e_i}\,| = \text{Vol}(v, w, \vec{n}\,) = \mathcal{A}(v, w)\cdot\|\vec{n}\,\| = |\,\sin\theta\,|\,\|v\,\|\,\|w\,\|\,\|\,\vec{n}\,\|$$

Or $\sin\theta > 0$ si et seulement si $\{v, w\}$ est orienté positivement (cf. remarque page 220), c'est-à-dire si $\text{dét}\,\|v, w, \vec{n}\,\|_{e_i} > 0$. On en déduit la formule.

REMARQUE – Bien entendu, le fait que l'on puisse définir sans ambiguïté l'angle orienté entre deux vecteurs de $\mathbb{R}^2$, tient à ce que $\mathbb{R}^2$ est muni d'une orientation canonique (l'orientation définie par la base canonique).
À la fin de ce chapitre nous donnerons une définition plus précise d'angle orienté (cf. définition 7.29)

Comme on le voit de ces quelques exemples, les différentes propriétés métriques de l'espace ordinaire peuvent être définies rien qu'en faisant appel aux propriétés i), ii), et iii). Aussi i), ii) et iii) peuvent être utilisées comme axiomes pour définir le produit scalaire dans les espaces vectoriels quelconques.

Exercices 1. 2. 3. 4.

7.2 Produit scalaire sur un espace vectoriel. Espaces euclidiens

Définition 7.1 – *Soit E un espace vectoriel* réel *; on appelle* ***produit scalaire*** *sur E une application $\langle\ ,\ \rangle : E\times E \to \mathbb{R}$ qui est :*
i) bilinéaire
ii) symétrique
iii) définie positive, c'est-à-dire :

$$\langle x, x\rangle \ \geq 0\,,\quad \forall\, x\in E \quad \text{et}$$
$$\langle x, x\rangle \ = 0 \quad \Longleftrightarrow x = 0$$

Un espace vectoriel réel de dimension finie*, muni d'un produit scalaire est dit* ***espace euclidien***[1]*.*

Exemple 1 – Soit $E = \mathbb{R}^n$ avec $\langle\ ,\ \rangle$ défini par :

$$\langle x, y\rangle \;:=\; x_1\,y_1 + \cdots + x_n\,y_n, \quad \text{où} : x = (x_1,\cdots,x_n),\ y = (y_1,\cdots,y_n)$$

[1] Si E est de dimension infinie, il est dit espace ***préhilbertien réel.***

Le produit scalaire ainsi défini est dit ***produit scalaire canonique***.
Cet exemple est important : nous verrons par la suite (cf. corollaire 7.7) que si E est un espace euclidien, moyennant le choix d'une base, on peut identifier E à $\mathbb{R}^n$ muni du produit scalaire canonique. En d'autres termes, *il s'agit, à un changement de base près, du seul exemple d'espace euclidien* [2].

Exemple 2 – Soit E un espace vectoriel réel de dimension n et $\{e_i\}$ une base de E. On définit alors un produit scalaire sur E, en posant :

$$\langle x,\, y\rangle_{e_i} = x_1\, y_1 + \cdots + x_n\, y_n$$

pour $x = \sum_{i=1}^n x_i\, e_i$, $y = \sum_{i=1}^n y_i\, e_i$.
$\langle x,\, y\rangle_{e_i}$ est dit ***produit scalaire associé à la base*** $\{e_i\}$.

Exemple 3 – Sur $\mathbb{R}^3$ on définit

$$\langle x, y\rangle \;:=\; x_1 y_1 + 2\, x_2 *, y_2 + 3\, x_3\, y_3 + x_1 y_2 + x_2 y_1$$

La symétrie est évidente. La bilinéarité vient du fait que tous les x_i et y_i apparaissent à la puissance 1 : plus précisément, le second membre est un polynôme homogène de degré 1 en les x_i et les y_i. Montrons que $\langle\, ,\, \rangle$ est défini positif.
On a :

$$\langle x, x\rangle \;:= x_1^2 + 2\, x_2^2 + 3\, x_3^2 + 2\, x_1 x_2$$

Ecrivons le premier et dernier terme à second membre comme le début d'un carré :

$$\langle x, x\rangle \;= \underbrace{(x_1 + x_2)^2 - x_2^2}_{x_1^2 + 2\, x_1\, x_2} + 2\, x_2^2 + 3\, x_3^2 = (x_1 + x_2)^2 + x_2^2 + 3\, x_3^2 \;\geq\; 0$$

et on a :

$$\langle x, x\rangle = 0 \iff \begin{cases} x_1 + x_2 & = 0 \\ \quad x_2 & = 0 \\ \qquad x_3 & = 0 \end{cases} \iff x_1 = x_2 = x_3 = 0$$

Donc $\langle\, ,\, \rangle$ est défini positif.

Exemple 4 – Sur $\mathcal{M}_2(\mathbb{R})$ on définit : $\langle A, B\rangle \;:= \mathrm{Tr}(\,{}^tAB)$.

Il s'agit d'un produit scalaire. En effet si $A = \begin{pmatrix} x_1 & x_2 \\ x_3 & x_4 \end{pmatrix}$ et $B = \begin{pmatrix} y_1 & y_2 \\ y_3 & y_4 \end{pmatrix}$, on a :

$$\langle A, B\rangle \;=\; x_1\, y_1 + x_2\, y_2 + x_3\, y_3 + x_4\, y_4$$

qui est bilinéaire, symétrique et définie positive.
Plus généralement sur $\mathcal{M}_n(\mathbb{R})$, $\langle A, B\rangle \;:=\; \mathrm{Tr}\,(\,{}^tA\, B)$ définit un produit scalaire, dit ***canonique***. Si $A = (\, x_{ij}\,)$ et $B = (\, y_{ij}\,)$, on a :

$$\langle A, B\rangle = \sum_{i,j=1}^{n} x_{ij}\, y_{ij}\,.$$

Il s'agit donc du produit scalaire associé à la base canonique E_{ij} de $\mathcal{M}_n(\mathbb{R})$ (cf. exercice 10, chapitre A.1).

[2] Ainsi, si l'on cherche des exemples ou contre-exemples il suffit de se limiter à ce cas.

Exemple 5 – Sur $\mathbb{R}[x]$ on définit : $\langle P, Q\rangle := \displaystyle\int_0^1 P(x)Q(x)\mathrm{d}x$.

La bilinéarité et la symétrie sont évidentes. D'autre part, $\langle P, P\rangle = \int_0^1 P(x)^2\mathrm{d}x \geq 0$. Supposons que $\langle P, P\rangle = 0$; puisque $P(x)$ est une fonction continue et $P(x)^2 \geq 0$, $\int_0^1 P(x)^2\mathrm{d}x = 0$ entraîne $P(x)|_{[0,1]} = 0$. Ainsi $P(x)$ est un polynôme ayant une infinité de racines et par conséquent, nécessairement, $P(x) \equiv 0$. Donc $\langle\, ,\,\rangle$ est un produit scalaire.

Notons que $\mathbb{R}[x]$, n'est pas de dimension finie et donc qu'il n'est pas euclidien, mais *pré-hilbertien réel*.

Exemple 6 – Si F est un sous-espace d'un espace vectoriel E muni d'un produit scalaire $\langle\, ,\,\rangle_E$, on définit sur F :

$$\langle x, y\rangle_F := \langle x, y\rangle_E\ , \quad (\forall x, y \in F)$$

On vérifie facilement que $\langle\, ,\,\rangle_F$ est un produit scalaire sur F, dit ***produit scalaire induit***. Sauf mention contraire, nous supposerons que les sous-espaces vectoriels d'un espace euclidien sont munis du produit scalaire induit. S'il n'y a pas de risque de confusion possible, nous noterons d'ailleurs $\langle\, ,\,\rangle$ aussi bien le produit scalaire $\langle\, ,\,\rangle_E$ que le produit scalaire induit $\langle\, ,\,\rangle_F$.

Exercices 5. 6. 7.

7.3 Méthode de Gauss pour la réduction en carrés

Dans ce paragraphe nous allons préciser la technique utilisée dans l'exemple 3. page 222 qui permet de savoir si une application bilinéaire symétrique est définie positive.

Définition 7.2 – *Soit E un espace vectoriel sur K (non nécessairement de dimension finie). On appelle* ***forme bilinéaire*** *une application $f : E \times E \longrightarrow K$ qui est linéaire dans les deux arguments, c'est-à-dire qui vérifie :*

- $f(x + y,\, z) = f(x,\, z) + f(y,\, z)$
- $f(x,\, y + z) = f(x,\, y) + f(x\,,z)$
- $f(\lambda\, x,\, y) = \lambda\, f(x,\, y) = f(x, \lambda\, y)\,, \qquad \forall\, x, y, z \in E,\ \forall\, \lambda \in K.$

Supposons E de dimension finie et soit $\{e_i\}$ une base de E. Si $x = \sum_{i=1}^n x_i\, e_i$ et $y = \sum_{j=1}^n y_j\, e_j$, on a :

$$f(x, y) = f\Big(\sum_{i=1}^n x_i\, e_i\,, \sum_{j=1}^n y_j\, e_j\Big) = \sum_{i,j=1}^n x_i\, y_j\, f(e_i, e_j)$$

Les $f(e_i, e_j)$ sont des éléments de K. Si on note $a_{ij} := f(e_i, e_j)$, l'expression de f dans la base $\{e_i\}$ est :

$$f(x, y) = \sum_{i,j=1}^n a_{ij}\, x_i\, y_j$$

c'est-à-dire f est du type :

$$f(x, y) = a_{11}\, x_1\, y_1 + a_{12}\, x_1\, y_2 + \cdots + a_{ij}\, x_i\, y_j + \cdots + a_{nn}\, x_n\, y_n$$

Autrement dit, on reconnaît une forme bilinéaire par le fait que, en l'écrivant dans une base, on obtient une somme de monômes dans lesquels x_i et y_j apparaissent à la puissance 1.

Ainsi, par exemple :

$$f(x,y) = x_1\, y_1 + 2\, x_2\, y_2 - x_3\, y_1 + x_1\, y_3$$

est bilinéaire, alors que :

$$f(x,y) = 3\, x_1\, y_1 + 3\, x_1^2\, y_2 + \cdots$$

ne l'est pas, pas plus que :

$$f(x,y) = 3\, x_2\, y_2 + 2\, x_2 + \cdots$$

(car dans le second monôme y_i n'apparait pas à la puissance 1).

La vérification de la symétrie est tout aussi facile : le fait que f ne change pas lorsqu'on échange le rôle de x et y équivaut, bien entendu, à $a_{ij} = a_{ji}$: c'est-à-dire le coefficient de $x_i y_j$ doit être égal à celui de $x_j y_i$. Ainsi, par exemple :

$$f(x,y) = 2\, x_1\, y_1 + x_2\, y_2 - 7\, x_1\, y_3 - 7\, x_3\, y_1 + 4\, x_2\, y_3 + 4\, x_3\, y_2$$

est symétrique.
Pour savoir si f est définie positive (dans le cas où $K = \mathbb{R}$), on procède comme dans l'exemple de page 222, selon la technique suivante due à Gauss.

Il s'agit de savoir si $f(x,x) \geq 0$ et si $f(x,x) = 0$ implique $x = 0$. En écrivant $f(x,x)$ on obtient une expression du type :

$$f(x,x) = a_{11}\, x_1^2 + \cdots + a_{nn}\, x_n^2 + \cdots + a_{ij}\, x_i\, x_j + \cdots$$

$f(x,x)$ se présente donc comme un polynôme homogène de degré 2 en les x_i. Les termes en x_i^2 sont dits «termes carrés», les termes en $x_i x_j$, avec $i \neq j$, «termes rectangles».

Remarque. – Si f est un produit scalaire, *tous* les coefficients a_{ii} des termes carrés sont *strictement positifs*.

En effet, s'il existe un $i \in \{1, \cdots, n\}$ tel que $a_{ii} \leq 0$; on a $a_{ii} = f(e_i, e_i) \leq 0$, donc f n'est pas définie positive.
Ainsi par exemple si $f(x,x) = 2\, x_1^2 + x_3^2 + 3\, x_1\, x_2 + 5\, x_2\, x_3$ on n'a pas pas un produit scalaire (le cœfficient de x_2^2 est nul).

Pour comprendre la méthode de Gauss, nous allons d'abord considérer quelques exemples :

Exemple 1 –
Soit $E = \mathbb{R}^3$, $\{e_i\}$ la base canonique, $x = x_1\, e_1 + x_2\, e_2 + x_3\, e_3$ et :

$$f(x,x) = x_1^2 + 2\, x_2^2 + 5\, x_3^2 + 2\, x_1\, x_2 - 4\, x_2\, x_3$$

Considérons un terme carré, par exemple x_1^2 (s'il n'y a pas de terme carré, f n'est pas un produit scalaire).

- on ordonne suivant la variable x_1 :

$$f(x,x) = \underbrace{x_1^2 + 2\,x_1\,x_2}_{\text{termes avec } x_1} + 2\,x_2^2 + 5\,x_3^2 - 4\,x_2\,x_3$$

– on écrit les termes contenant x_1 comme le début d'un carré :

$$f(x,x) = \underbrace{\underset{\substack{\uparrow\\ \text{terme}\\ \text{avec } x_1}}{(x_1 + x_2)^2} - \underset{\substack{\uparrow\\ \text{terme}\\ \text{correctif}}}{x_2{}^2}} + 2\,x_2^2 + 5\,x_3^2 - 4\,x_2\,x_3$$

– on obtient le carré d'une forme linéaire plus des termes qui ne contiennent pas x_1 :

$$f(x,x) = (x_1 + x_2)^2 + \underbrace{x_2^2 + 5\,x_3^2 - 4\,x_2\,x_3}_{\substack{\text{termes ne contenant}\\ \text{pas } x_1}}$$

– on recommence l'opération sur les termes qui ne contiennent pas x_1 :

$$\begin{aligned} f(x,x) &= (x_1 + x_2)^2 + \underbrace{x_2^2 - 4\,x_2\,x_3}_{\text{termes avec } x_2} + 5x_3^2 \\ &= (x_1 + x_2)^2 + \underbrace{(x_2 - 2\,x_3)^2 - \underset{\substack{\uparrow\\ \text{terme}\\ \text{correctif}}}{4\,x_3^2}} + 5\,x_3^2 \end{aligned}$$

– jusqu'à obtenir une somme de carrés de formes linéaires :

$$f(x,x) = (x_1 + x_2)^2 + (x_2 - 2\,x_3)^2 + x_3^2$$

Cette expression est dite ***réduction en carrés de Gauss***.

Notons que la réduction de Gauss n'est pas unique, car au lieu de partir de la variable x_1 on aurait pu partir de n'importe quelle autre variable.

Il est facile maintenant de savoir si f est définie positive. Dans notre exemple on a bien $f(x,x) \geq 0$. D'autre part $f(x,x) = 0$ équivaut au système :

$$\begin{cases} x_1 + x_2 \qquad\qquad = 0 \\ \qquad x_2 - 2\,x_3 \;\; = 0 \\ \qquad\qquad\quad x_3 \;\; = 0 \end{cases}$$

qui donne immédiatement $x_1 = x_2 = x_3 = 0$, c'est-à-dire $x = 0$. f est donc un produit scalaire.

Exemple 2 –

Soit E de dimension 3 et supposons que dans une base $\{e_i\}$:

$$f(x,x) = x_1^2 + 3\,x_2^2 + 5\,x_3^2 + 2\,x_1\,x_2 - 4\,x_1\,x_3 + 6\,x_2\,x_3$$

En appliquant la méthode de l'exemple 1, on trouve :

$$\begin{aligned} f(x,x) &= \underbrace{\Big(x_1 + (x_2 - 2\,x_3)\Big)^2}_{\substack{\uparrow \\ \text{termes avec } x_1}} - \underbrace{(x_2 - 2\,x_3)^2}_{\substack{\uparrow \\ \text{terme correctif}}} + 3\,x_2^2 + 5\,x_3^2 + 6\,x_2\,x_3 \\ &= (x_1 + x_2 - 2\,x_3)^2 + 2\,x_2^2 + x_3^2 + 10\,x_2\,x_3 \\ &= (x_1 + x_2 - 2\,x_3)^2 + (x_3 + 5\,x_2)^2 - 25\,x_2^2 + 2\,x_2^2 \\ &= (x_1 + x_2 - 2\,x_3)^2 + (x_3 + 5\,x_2)^2 - 23\,x_2^2 \end{aligned}$$

On voit immédiatement que f n'est pas un produit scalaire car cette expression peut s'annuler sans que x soit nul (ceci tient au fait que le dernier carré est affecté d'un signe "moins"). Il suffit, en effet, de prendre x_1, x_2, x_3 solution du système :

$$\begin{cases} x_1 + x_2 - 2x_3 &= 0 \\ x_3 - 5\,x_2 &= \sqrt{23}\,x_2 \end{cases}$$

système qui a des solutions non nulles, par exemple $x_2 = 1$, $x_3 = 5+\sqrt{23}$, $x_1 = 9+2\sqrt{23}$.

Il est facile de se convaincre de la propriété suivante :

Théorème 7.3 – *Soit E un espace vectoriel de* dimension n *sur $\mathbb{R}$, et f une forme bilinéaire symétrique ; alors f est définie positive si et seulement si la réduction de Gauss donne* ***n carrés affectés du signe "plus"***.

En effet si la réduction de Gauss donne n carrés affectés du signe $+$, en imposant $f(x,x) = 0$ on obtient un système homogène de n équations en n inconnues, *qui est échelonné* (car la méthode de Gauss permet, à chaque étape, d'éliminer une variable). Or un tel système n'admet que la solution nulle.
Réciproquement, s'il y a moins de n carrés, ou si l'un des carrés est affecté du signe "moins", il n'est pas difficile de se convaincre qu'il existe des solutions non nulles de $f(x,x) = 0$. En effet, soit

$$f(x,x) = a_1\,\varphi_1(x)^2 + a_2\,\varphi_2(x)^2 + \cdots + a_n\,\varphi_n(x)^2$$

la réduction de Gauss, où les $\varphi_i(x)$ sont des formes linéaires. Supposons, quitte à changer la numérotation, que $a_1 > 0$ et $a_2 \leq 0$ et considérons le système :

$$\begin{cases} \varphi_1(x) &= \sqrt{\dfrac{-a_2}{a_1}}\;\varphi_2(x) \\ \varphi_3(x) &= 0 \\ \quad\vdots & \\ \varphi_n(x) &= 0 \end{cases}$$

Ce système est homogène et comporte moins d'équations que d'inconnues. Il admet donc des solutions $x \neq 0$ (cf. theorème 2.2 page 43) ; de tels x vérifient $f(x,x) = 0$. $\square$

Exercice 8.

7.4 Le théorème fondamental des espaces euclidiens. Procédé d'orthonormalisation de Schmidt.

Soit $(E, \langle\, ,\, \rangle)$ un espace euclidien et soit $\{e_i\}$ une base de E. Si $x = \sum_{i=1}^{n} x_i\, e_i$ et $y = \sum_{j=1}^{n} y_j\, e_j$, on a :

$$\langle x\, ,\, y\rangle \;=\; \Big\langle \textstyle\sum_{i=1}^{n} x_i\, e_i \;,\; \sum_{j=1}^{n} y_j\, e_j \Big\rangle \;=\; \textstyle\sum_{i,j=1}^{n} x_i\, y_j\, \langle e_i\, ,\, e_j\rangle$$

Les $\langle e_i\, ,\, e_j\rangle$ sont des éléments de K. Si on pose $a_{ij} := \langle e_i\, ,\, e_j\rangle$, l'expression de $\langle\, ,\, \rangle$ dans la base $\{e_i\}$ est :

$$\boxed{\langle x\, ,\, y\rangle = \sum_{i,j=1}^{n} a_{ij}\, x_i\, y_j}$$

c'est-à-dire $\langle\, ,\, \rangle$ est du type :

$$\langle x\, ,\, y\rangle = a_{11}\, x_1\, y_1 + a_{12}\, x_1\, y_2 + \cdots + a_{ij}\, x_i\, y_j + \cdots + a_{nn}\, x_n\, y_n.$$

Dans ce chapitre nous allons montrer qu'en fait *il n'y a qu'un seul type d'espaces euclidiens.* Plus précisément, modulo un changement de base, tout produit scalaire peut s'écrire sous la forme [3] :

$$\langle\, x,\, y\,\rangle \;=\; x_1\, y_1 + \cdots + x_n\, y_n$$

Commençons par quelques définitions.

Définition 7.4 – *Soit $(E, \langle\, ,\, \rangle)$ un espace euclidien. Puisque $\langle x, x\rangle \;\geq 0$ pour tout $x \in E$, on peut considérer la racine carrée :*

$$\|\, x\,\| := \sqrt{\langle x,\, x\rangle}.$$

L'expression $\|x\,\|$ est dite ***norme*** [4] *de x.*
Deux vecteurs $x,\, y \in E$ sont dit ***orthogonaux*** *si $\langle x,\, y\rangle = 0$. On note cela aussi $x \perp y$.*

Définition 7.5 – *Soit E un espace euclidien. Une base $\{e_i\}$ est dite* ***orthogonale*** *si :*

$$\langle e_i,\, e_j\rangle = 0\,, \quad \forall\, i,\, j \text{ tels que } i \neq j.$$

Elle est dite ***orthonormée*** *si de plus chaque vecteur est de norme 1, c'est-à-dire si :*

$$\langle e_i,\, e_j\rangle = \delta_{ij}$$

où δ_{ij} est le "symbole de Kronecker" défini par : $\delta_{ij} = \begin{cases} 1 & si\ i = j \\ 0 & si\ i \neq j \end{cases}$

[3]En particulier, par exemple, si l'on demande de montrer une propriété sur les espaces euclidiens, on ne perd pas en généralité en la démontrant sur $\mathbb{R}^n$ muni du produit scalaire canonique, ce qui, bien entendu, est en général plus simple.

[4]La notion de norme et ses propriétés seront étudiées au §5.

Remarques. –

1. Il est clair que si $\{\varepsilon_i\}$ est une base orthogonale, les vecteurs $e_i := \dfrac{\varepsilon_i}{\|\varepsilon_i\|}$ définissent une base orthonormée.
2. Toute famille de vecteurs non nuls $\{\varepsilon_1, ..., \varepsilon_p\}$ deux à deux orthogonaux est libre.

   En effet, si $a_1\,\varepsilon_1 + \cdots + a_p\,\varepsilon_p = 0$; on a, pour $j = 1, ..., p$:
   $$0 = \langle\, e_j\ ,\ a_1\,\varepsilon_1 + \cdots + a_p\,\varepsilon_p\,\rangle = a_j$$
   ce qui montre que la famille est libre.

Soit $\{e_i\}$ une base orthonormée et $x = \sum_{i=1}^n x_i\,e_i,\ y = \sum_{j=1}^n y_j\,e_j$ deux vecteurs de E. On a :

$$\begin{aligned}\langle x,\, y\rangle &= \left\langle \sum_{i=1}^n x_i\,e_i\ ,\ \sum_{j=1}^n y_j\,e_j \right\rangle = \sum_{i,j=1}^n x_i\,y_j\ \langle e_i\,,\,e_j\rangle \\ &= \sum_{i,j=1}^n x_i\,y_j\,\delta_{ij} = x_1\,y_1 + \cdots + x_n\,y_n\end{aligned}.$$

Ainsi :

> *Dans une base orthonormée le produit scalaire prend la forme du produit scalaire canonique.*

On comprend dès lors l'intérêt de savoir s'il existe des bases orthonormées et, le cas échéant, comment on les construit.

Théorème Fondamental des Espaces Euclidiens . 7.6 .

Dans un espace euclidien il existe toujours des bases orthonormées.

Démonstration : D'après la remarque ci-dessus, il suffit de montrer qu'il existe des bases orthogonales. Montrons cela par récurrence sur la dimension n de E.
Pour $n = 1$ il n'y a rien à démontrer. Supposons le théorème vrai à l'ordre $n-1$ et soit $v \in E$, $v \neq 0$. Considérons l'ensemble de tous les vecteurs orthogonaux à v :

$$F = \{\, x \in E \mid \langle x\,,\,v\rangle = 0\,\}.$$

Il est facile de voir que F est un sous-espace vectoriel de E et que $\dim F = n-1$. En effet, $F = \operatorname{Ker}\omega$ où ω est la forme linéaire [5] définie par $\omega(x) = \langle x\,,v\rangle$. Comme $\omega(v) = \| v \| \neq 0$, $\omega \neq 0$ et donc $\dim F = \dim \operatorname{Ker}\omega = n-1$.

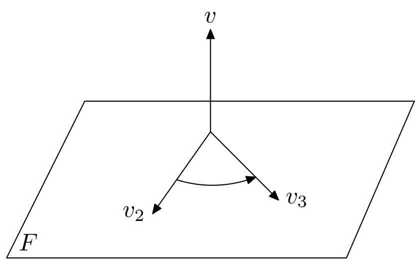


Figure 9

[5] On a bien $\omega \in E^*$, car le produit scalaire est linéaire en le premier argument.

Il s'ensuit, d'après l'hypothèse de récurrence, que sur F il existe des bases orthogonales. Par ailleurs, $v \notin F$, car si on avait $v \in F$, on aurait $\langle v \mid v \rangle = 0$ et donc $v = 0$. Aussi $E = \mathrm{Vect}\{v\} \oplus F$. Par conséquent, si $\{v_2, \cdots, v_n\}$ est une base orthogonale de F alors $\{v, v_2, \cdots, v_n\}$ est une base orthogonale de E (cf. figure). □

On en déduit :

Corollaire 7.7 *Le choix d'une base orthonormée $\mathcal{B} = \{e_i\}$ dans un espace euclidien, permet d'identifier E à $\mathbb{R}^n$ muni du produit scalaire canonique, par l'isomorphisme :*

$$\begin{array}{cccc} \varphi_{\mathcal{B}} : & E & \longrightarrow & \mathbb{R}^n \\ & x = \sum_{i=1}^n x_i e_i & \longmapsto & (x_1, \cdots, x_n) \end{array}$$

Construction de bases orthonormées. Procédé d'orthonormalisation de Schmidt

Nous allons expliquer ici une méthode due à Schmidt, qui, par un procédé standard, permet d'associer à toute base de E une base orthonormée [6].

Proposition 7.8 *Soit $\{v_1, \cdots, v_p\}$ une famille libre d'un espace euclidien E et $F =$ $\mathrm{Vect}\{v_1, \cdots, v_p\}$ le sous-espace engendré. Alors, par un procédé standard, on peut alors construire une base orthonormée de F à partir de $\{v_1, \cdots, v_p\}$.*
En particulier, à toute base de E on peut associer une base orthonormée de E.

Démonstration : La méthode consiste à construire d'abord, par récurrence, une base orthogonale $\{\varepsilon_1, \cdots, \varepsilon_p\}$ de F et ensuite à la normaliser $\Big(e_i = \dfrac{\varepsilon_i}{\|\varepsilon_i\|}\Big)$.
Pour cela on pose :

$$\begin{cases} \varepsilon_1 = v_1 \\ \varepsilon_2 = v_2 + \lambda\, \varepsilon_1\,, \qquad \text{avec } \lambda \text{ tel que } \varepsilon_2 \perp \varepsilon_1. \end{cases}$$

En imposant cette condition, on trouve :

$$0 \;=\; \langle v_2 + \lambda\, \varepsilon_1, \varepsilon_1 \rangle \;=\; \langle v_2, \varepsilon_1 \rangle + \lambda \|\varepsilon_1\|^2.$$

Comme $\varepsilon_1 \neq 0$, on obtient $\lambda = -\dfrac{\langle v_2, \varepsilon_1 \rangle}{\|\varepsilon_1\|^2}$. Notons que, puisque

$$\begin{cases} v_1 = \varepsilon_1 \\ v_2 = \varepsilon_2 - \lambda\, \varepsilon_1 \end{cases}$$

ε_1 et ε_2 engendrent le même espace que v_1 et v_2.
Une fois construit ε_2, on construit ε_3 en posant :

$$\begin{array}{l} \varepsilon_3 = v_3 + \mu\, \varepsilon_1 + \nu\, \varepsilon_2 \\ \text{avec}\, \mu \text{ et } \nu \text{ tels que : } \varepsilon_3 \perp \varepsilon_1 \text{ et } \varepsilon_3 \perp \varepsilon_2. \end{array}$$

Ceci donne

[6] Nous verrons au chapitre 9. (cf. page 307) une autre méthode fondée sur la réduction en carrés de Gauss.

$$0 = \langle v_3 + \mu\varepsilon_1 + \nu\varepsilon_2, \varepsilon_1\rangle = \langle v_3, \varepsilon_1\rangle + \mu \,\|\varepsilon_1\|^2 \text{ , car } \langle \varepsilon_1, \varepsilon_2\rangle = 0$$

d'où $\mu = -\dfrac{\langle v_3, \varepsilon_1\rangle}{\|\varepsilon_1\|^2}$. De même, en imposant que $\varepsilon_3 \perp \varepsilon_2$, on trouve $\nu = -\dfrac{\langle v_3,\, \varepsilon_2\rangle}{\|\varepsilon_2\|^2}$.
Comme

$$\begin{cases} v_1 = \varepsilon_1 \\ v_2 = \varepsilon_2 - \lambda\,\varepsilon_1 \\ v_3 = \varepsilon_3 - \mu\,\varepsilon_1 - \nu\,\varepsilon_2 \end{cases}$$

on a $\mathrm{Vect}\{\varepsilon_1, \varepsilon_2, \varepsilon_3\} = \mathrm{Vect}\{v_1, v_2, v_3\}$, c'est-à-dire $\{\varepsilon_1, \varepsilon_2, \varepsilon_3\}$ est une base orthogonale de l'espace engendré par v_1, v_2, v_3

On voit bien maintenant le procédé de récurrence.
Supposons avoir construit $\varepsilon_1, \cdots, \varepsilon_{k-1}$ pour $k \leq p$; on pose :

$$\begin{aligned} \varepsilon_k \quad &= v_k + \text{combinaison linéaire des vecteurs déjà trouvés} \\ &= v_3 + \lambda_1\,\varepsilon_1 + \cdots + \lambda_{k-1}\,\varepsilon_{k-1} \end{aligned}$$

Les conditions $\varepsilon_k \perp \varepsilon_i$ (pour $i = 1, \cdots, k-1$) sont équivalentes à :

$$\lambda_i = -\frac{\langle v_k,\, \varepsilon_i\rangle}{\|\varepsilon_i\|^2}$$

comme on le vérifie immédiatement.
Puisque $v_k = \varepsilon_k - \lambda_1\,\varepsilon_1 - \cdots - \lambda_{k-1}\,\varepsilon_{k-1}$, on voit facilement par récurrence que $\mathrm{Vect}\{\varepsilon_1, \ldots, \varepsilon_k\} = \mathrm{Vect}\{v_1, \ldots, v_k\}$, aussi $\{\varepsilon_1, \ldots, \varepsilon_k\}$ est une base orthogonale de $\mathrm{Vect}\{v_1, \ldots, v_k\}$.

Exemple –

Déterminer une base orthonormée du sous-espace vectoriel F de $\mathbb{R}^4$ engendré par les vecteurs :

$$v_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \; v_2 = \begin{pmatrix} 1 \\ 0 \\ -1 \\ 1 \end{pmatrix}, \; v_3 = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 1 \end{pmatrix}$$

($\mathbb{R}^4$ *étant muni du produit scalaire canonique*).
On vérifie d'abord que $\{v_1,\, v_2,\, v_3\}$ est une famille libre, ce qui est immédiat, car dans la matrice :

$$\|v_1,\, v_2,\, v_3\| = \begin{pmatrix} 1 \quad 1 \quad 0 \\ \boxed{\begin{matrix} 1 & 0 & 1 \\ 0 & -1 & 1 \\ 0 & 1 & 1 \end{matrix}} \end{pmatrix}$$

le mineur encadré est non nul. On pose :

$\varepsilon_1 = v_1$
$\varepsilon_2 = v_2 + \lambda\,\varepsilon_1$, avec $\lambda = -\dfrac{\langle v_2, \varepsilon_1\rangle}{\|\varepsilon_1\|^2} = -\dfrac{1}{2}$.

Donc $\varepsilon_2 = v_2 - \dfrac{1}{2}\varepsilon_1 = \begin{pmatrix} 1/2 \\ -1/2 \\ -1 \\ 1 \end{pmatrix} = \dfrac{1}{2}\begin{pmatrix} 1 \\ -1 \\ -2 \\ 2 \end{pmatrix}$.

Remarquons que l'on peut prendre comme second vecteur un vecteur ε_2' colinéaire à ε_2 : ε_1 et ε_2' resteront orthogonaux. On prend donc : $\varepsilon_2' = \begin{pmatrix} 1 \\ -1 \\ -2 \\ 2 \end{pmatrix}$ (cela simplifie les calculs).

On pose ensuite :

$$\varepsilon_3 = v_3 + \lambda_1\,\varepsilon_1 + \lambda_2\,\varepsilon_2'\ , \quad \text{avec} \quad \lambda_1 = -\frac{\langle v_3, \varepsilon_1\rangle}{\|\varepsilon_1\|^2} \quad \text{et} \quad \lambda_2 = -\frac{\langle v_3, \varepsilon_2'\rangle}{\|\varepsilon_2'\|^2}$$

On trouve : $\lambda_1 = -\dfrac{1}{2}$, $\lambda_2 = \dfrac{1}{10}$ d'où : $\varepsilon_3 = \dfrac{2}{5}\begin{pmatrix} -1 \\ 1 \\ 2 \\ 3 \end{pmatrix}$.

Ainsi ε_1, ε_2' et $\varepsilon_3' = \begin{pmatrix} -1 \\ 1 \\ 2 \\ 3 \end{pmatrix}$ forment une base orthogonale de F.

En normalisant, on trouve :

$$e_1 = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad e_2 = \frac{1}{\sqrt{10}}\begin{pmatrix} 1 \\ -1 \\ -2 \\ 2 \end{pmatrix}, \quad e_3 = \frac{1}{\sqrt{15}}\begin{pmatrix} -1 \\ 1 \\ 2 \\ 3 \end{pmatrix}$$

qui est une base orthonormée de F.

Exercice 9.

7.5 Norme d'un vecteur. Angle non orienté

1. Norme

Soit E un espace vectoriel réel muni d'un produit scalaire.

Proposition 7.9 – *L'application* $\|\ \| : E \longmapsto \mathbb{R}^+$ *définie par* $\|\,x\,\| := \sqrt{\langle x,\, x\rangle}$ *est une* ***norme****, c'est-à-dire elle vérifie les propriétés suivantes :*

1. $\|\lambda\,x\| = |\lambda|\|x\|\,, \quad \forall x \in E,\ \forall\lambda \in \mathbb{R}$;
2. $\|x\| = 0$ *si et seulement si* $x = 0$ *;*
3. Inégalité triangulaire *:*

$$\|x + y\,\| \leq \|x\| + \|y\,\|\,, \quad \forall x,\, y \in E$$

l'égalité ayant lieu si et seulement si il existe $\lambda \geq 0$ *tel que* $y = \lambda\,x$ *(ou* $x = \lambda\,y$*).*

Démonstration :
1. et 2. sont évidentes. Pour montrer 3. on a besoin du lemme suivant :

Inégalité de Cauchy-Schwarz . 7.10 – *Pour tous* $x,\, y \in E$*, on a :*

$$\langle x,\, y\,\rangle^2 \leq \|x\|^2\,\|y\,\|^2$$

l'égalité ayant lieu si et seulement si x *et* y *sont liés.*

En effet, si $\|x\| = \|y\| = 0$, on a $x = 0$ et $y = 0$ et la propriété est triviale. Supposons que $\|y\| \neq 0$ et soit $\lambda \in \mathbb{R}$ arbitraire ; on aura $\|x + \lambda y\|^2 \geq 0$, c'est-à-dire :

$$\|x\|^2 + \|\lambda y\|^2 + 2\langle x, \lambda y\rangle \geq 0\,, \quad \forall \lambda \in \mathbb{R},$$

donc :

$$\lambda^2 \|y\|^2 + 2\lambda \langle x, y\rangle + \|x\|^2 \geq 0\,, \quad \forall \lambda \in \mathbb{R}. \tag{$*$}$$

Pour que ce trinôme en λ soit ≥ 0 il faut et il suffit que le discriminant soit ≤ 0, c'est-à-dire :

$$\langle x, y\rangle^2 - \|x\|^2 \|y\|^2 \leq 0,$$

ce qui montre l'inégalité.
D'autre part, si l'on a $\langle x, y\rangle^2 = \|x\|^2 \|y\|^2$, le trinôme (*) a une racine (double) λ ; c'est-à-dire, il existe $\lambda \in \mathbb{R}$ tel que :

$$\lambda^2 \|y\|^2 + 2\lambda \langle x, y\rangle + \|x\|^2 = 0.$$

En remontant les calculs, on voit que cela signifie qu'il existe $\lambda \in \mathbb{R}$ tel que $\|x + \lambda y\| = 0$, c'est-à-dire : $x + \lambda y = 0$. $\diamondsuit$

Montrons maintenant l'inégalité triangulaire. On a :

$$\begin{aligned} \|x + y\|^2 &= \|x\|^2 + \|y\|^2 + 2\langle x, y\rangle \underset{(a)}{\leq} \|x\|^2 + \|y\|^2 + 2\,|\langle x, y\rangle| \\ &\underset{(b)}{\leq} \|x\|^2 + \|y\|^2 + 2\|x\|\,\|y\| = (\|x\| + \|y\|)^2 \end{aligned}$$

donc $$\|x + y\| \leq \|x\| + \|y\|.$$

Notons que si $y = \lambda x$ avec $\lambda \geq 0$ on a l'égalité. Réciproquement, supposons que l' on a l'égalité ; en remontant les calculs, on voit que :

- l'égalité en (b) implique que x et y sont colinéaires, d'après l'inégalité de Cauchy-Schwarz ; donc, si par exemple $x \neq 0$, $y = \lambda x$.
- on a l'égalité en (a) si $\langle x, y\rangle = |\langle x, y\rangle|$. Puisque $y = \lambda x$, on a $\lambda \|x\|^2 = |\lambda|\,\|x\|^2$ donc $\lambda \geq 0$. $\square$

Remarque. — L'inégalité triangulaire exprime le fait que dans un triangle la longueur d'un côté est inférieure ou égale à la somme des longueurs des deux autres côtés.

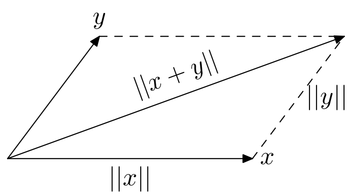


Figure 10

2. Angle

Soient x, y deux vecteurs non nuls ; on a, d'après l'inégalité de Cauchy-Schwarz :

$$\frac{|\langle x,\, y\rangle|}{\|x\|\,\|y\|} \leq 1.$$

Par conséquent, il existe un et un seul $\theta \in [0,\, \pi]$ tel que

$$\boxed{\cos\theta = \frac{\langle x,\, y\rangle}{\|x\|\,\|y\|}} \qquad (1)$$

θ est dit ***angle*** (non orienté) entre les vecteurs x et y (cf. (1) page 219).

Notons enfin la relation qui exprime le produit scalaire en fonction de la norme :

$$\boxed{\langle x,\, y\rangle = \frac{1}{2}\left(\|x+y\|^2 - \|x\|^2 - \|y\|^2\right)} \qquad (2)$$

dont la démonstration est immédiate. On a en effet :

$\|x+y\|^2 = \langle x+y,\, x+y\rangle = \|x\|^2 + \|y\|^2 + \langle x,\, y\rangle + \langle y,\, x\rangle = \|x\|^2 + \|y\|^2 + 2\,\langle x,\, y\rangle$

(en vertu de la symétrie du produit scalaire).

Exercices 10. 11.

7.6 Représentation matricielle du produit scalaire

Pour les calculs il est souvent pratique d'utiliser l'expression matricielle du produit scalaire.

Soit E un espace vectoriel de dimension finie n sur K et $b : E \times E \longrightarrow K$ une forme bilinéaire (cf. définition page 223. Si $\{e_i\}$ est une base de E et $x = \sum_{i=1}^n x_i\, e_i$, $y = \sum_{i=1}^n y_j\, e_j$, on a :

$$b(x,\, y) = \sum_{i=1}^{n} x_i\, y_j\, b(e_i,\, e_j).$$

b est donc déterminée par la connaissance des valeurs $b(e_i,\, e_j)$ sur une base.

Définition 7.11 *Soient b une forme bilinéaire sur E, et $\{e_i\}$ une base de E. On appelle* ***matrice de b*** *dans la base $\{e_i\}$ la matrice :*

$$M(b)_{e_i} = \begin{pmatrix} b(e_1,e_1) & b(e_1,e_2) & \cdots & b(e_1,e_n) \\ b(e_2,e_1) & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ b(e_n,e_1) & \cdots & \cdots & b(e_n,e_n) \end{pmatrix}$$

$$= \|b\,(\underset{\substack{\uparrow \\ \textit{ième} \\ \textit{ligne}}}{e_i}\;,\;\underset{\substack{\uparrow \\ \textit{jème} \\ \textit{colonne}}}{e_j})\|$$

Ainsi l'élément de la i^{eme} ligne et j^{eme} colonne est le coefficient de $x_i\, y_j$.

Exemple 1.– Si $b : \mathbb{R}^3 \times \mathbb{R}^3 \longrightarrow \mathbb{R}$ est la forme bilinéaire qui dans la base canonique $\{e_i\}$ s'écrit :

$$b(x,\, y\,) = 5\,x_1y_1 - 2\,x_2y_2 + 4\,x_3y_3 + 7\,x_1y_2 + 6\,x_1y_3 - 4\,x_3y_1 + 2\,x_2y_3 + 8\,x_3y_2$$

on a :

$$M(b)_{e_i} = \begin{pmatrix} 5 & 7 & 6 \\ 0 & -2 & 2 \\ -4 & 8 & 4 \end{pmatrix}$$

Exemple 2.– Si $\langle\ ,\ \rangle$ est le produit scalaire canonique sur $\mathbb{R}^n$, $\langle x,\, y\rangle = x_1y_1 + \cdots + x_ny_n$, on a :

$$M(\,\langle\,,\rangle\,)_{e_i} = \begin{pmatrix} 1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1 \end{pmatrix} = I$$

Proposition 7.12 *Soient b une forme bilinéaire sur E, $\{e_i\}$ une base, et*

$$A = M(b)_{e_i}, \quad X = M(x)_{e_i}, \quad Y = M(y)_{e_i} \qquad (x,\, y \in E).$$

On a :

$$\boxed{b(x,\, y) = {}^tXAY}$$

La démonstration est une simple vérification.

Exemple – Soit b est la forme bilinéaire ci-dessus et $X = \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix}$, $Y = \begin{pmatrix} 2 \\ 3 \\ 1 \end{pmatrix}$. On a :

$$\begin{array}{ccccccc} & & & \times & \begin{pmatrix} 5 & 7 & 6 \\ 0 & -2 & 2 \\ -4 & 8 & 4 \end{pmatrix} & \times & \left.\begin{pmatrix} 2 \\ 3 \\ 1 \end{pmatrix}\right\} \; Y \\ {}^tXAY & = & \underbrace{(1 \quad 2 \ -1)}_{{}^tX} & & \underbrace{(9 \ \ -5 \quad 6)}_{{}^tXA} & \rightarrow & \underset{{}^tXAY}{9} \end{array}$$

Donc $b(x,\, y) = 9$.

Remarque. – Une base $\{e_i\}$ *est orthonormée* si et seulement si :

$$\langle\, x, y\,\rangle_{e_i} = {}^tX\,Y$$

En effet, $\{e_i\}$ est orthonormée si et seulement si $\langle\, x, y\rangle_{e_i} = x_1\,y_1 + \cdots + x_n\,y_n$, ce qui équivaut au fait que la matrice dans la base $\{e_i\}$ est l'identité.

Il est clair qu'une forme bilinéaire b *est symétrique si et seulement si sa matrice est symétrique*. En particulier, la matrice d'un produit scalaire est symétrique (et, bien entendu, réelle, puisque dans ce cas, $K = \mathbb{R}$). La question qui se pose naturellement est de caractériser, parmi les matrices symétriques réelles celles qui sont associées à un produit scalaire [7]. Notons pour le moment que *la matrice qui représente un produit scalaire dans une base (quelconque) est nécessairement inversible.* Soit en effet A la matrice qui représente le produit scalaire dans une base $\{e_i\}$ et $X \in \mathcal{M}_{n,1}$, tel que $A\,X = 0$. En multipliant par ${}^tX$ à gauche, on aura :

$$ {}^tX\,A\,X = 0 $$

Cela signifie que si x est le vecteur de E qui dans la base $\{e_i\}$ est représenté par X, on a : $\langle x\,|\,x\rangle = 0$, d'où $x = 0$, c'est-à-dire $X = 0$. Donc $\operatorname{Ker} A = \{0\}$ et, par conséquent A est inversible.

Changement de base

Soient $\{e_i\}$ et $\{e'_i\}$ deux bases de E, $P = P_{e_i \to e'_i} = \|e'_1, \cdots, e'_n\|_{e_i}$ la matrice de passage, et $x, y \in E$. Si $X = M(x)_{e_i}$, $X' = M(x)_{e'_i}$, $Y = M(y)_{e_i}$ et $Y' = M(y)_{e'_i}$, on a :

$$ X' = P^{-1}X, \quad \text{d'où} \quad X = PX' \quad \text{et} \quad Y = PY' $$

(cf. proposition 3.25 page 77). Donc :

$$ b(x,y) = {}^tXAY = {}^t(PX')A(PY') = {}^tX'({}^tPAP)Y'. $$

D'autre part :

$$ b(x,\,y) = {}^tX'A'Y' \quad \text{où} \quad A' = M(b)_{e'_i} = \begin{pmatrix} b(e'_1,e'_1) & \cdots & b(e'_1,e'_n) \\ \vdots & & \vdots \\ b(e'_n,e'_1) & \cdots & b(e'_n,e'_n) \end{pmatrix} $$

On a donc :

$$ {}^tX'({}^tPAP)Y' = {}^tX'A'Y'\,, \quad \forall X'Y' \in \mathcal{M}_{n,1}(K) $$

et par conséquent (cf exercice 13) :

$$ \boxed{A' = {}^tPAP} $$

Remarque – Comparons cette formule avec la loi de changement de base des endomorphismes. Une fois fixée une base, à toute matrice sont associés un endomorphisme et une forme bilinéaire : dans le premier cas elle se transforme par $P^{-1}AP$ et dans le second par ${}^tPAP$.
En particulier, par exemple, il a un sens de parler de déterminant d'un endomorphisme. On le définit comme le déterminant de la matrice qui le représente dans une base quelconque, car

$$ \text{dét}\,A' = \text{dét}\,(P^{-1}AP) = (\text{dét}\,P)^{-1}\,\text{dét}\,A\,(\text{dét}\,P) = \text{dét}\,A. $$

En revanche, on ne peut pas définir le déterminant d'une forme bilinéaire comme le déterminant de la matrice qui la représente dans une base, car $\text{dét}\,A' = \text{dét}\,({}^tPAP) = (\text{dét}\,P)^2\,\text{dét}\,A$ et la définition dépendrait donc du choix de la base.

[7] Nous verrons par la suite (cf. Proposition 7.39) que la condition (nécessaire et suffisante) pour cela est *toutes les valeurs propres de la matrice (dont on montre qu'elles sont réelles) sont strictement positives.*

Exemple – *Écrire le produit scalaire canonique de* $\mathbb{R}^2$ *dans la base* $\{e'_1, e'_2\}$*, où :*

$$e'_1 = 2\,e_1 + 3\,e_2, \quad e'_2 = -8\,e_1 + 5\,e_2$$

Soit A' la matrice du produit scalaire dans la base $\{e'_1,\, e'_2\}$; on a :

$$A' \,=\, {}^tPIP \,=\, {}^tPP$$

Or

$$P = \begin{pmatrix} 2 & -8 \\ 3 & 5 \end{pmatrix}$$

Donc

$$A' \,=\, {}^tPP = \begin{pmatrix} 13 & -1 \\ -1 & 89 \end{pmatrix}$$

c'est-à-dire : $\langle x,\, y\rangle_{e'_i} = 13\,x'_1\,y'_1 - \,x'_1\,y'_2\, - x'_2\,y'_1 + 89\,x'_2\,y'_2.$

Exercices 12. 13.

7.7 Sous-espaces orthogonaux

Nous allons établir quelques propriétés relatives à la notion d'orthogonalité.

Définition 7.13 – *Soit* $A \subset E$ *(*A *un sous-ensemble quelconque) ; on pose :*

$$A^\perp := \{x \in E \mid \langle x\,,\, a\rangle = 0\,,\ \forall\, a \in A\}$$

NOTA – A cause de la symétrie, $A^\perp$ est aussi l'ensemble $\{x \in E \,|\, \langle a\,,\, x\rangle = 0\,,\ \forall\, a \in A\}$

On vérifie facilement que $A^\perp$ est un sous-espace vectoriel de E (même si A ne l'est pas), dit ***orthogonal*** de A.

On a :

$$\begin{array}{lrcl} 1. & \{0\}^\perp & = & E \\ 2. & E^\perp & = & \{0\} \end{array}$$

1. est évidente : $\forall x \in E$, on a $\langle x,\, 0\rangle = 0$, car l'application $\langle\,,\rangle$ est linéaire dans le second argument.
2. vient du fait que $\langle\,,\rangle$ est défini positif. Soit en effet $x \in E^\perp$; on a : $\langle x,\, y\rangle = 0\,,\ \forall y \in E$; en particulier $\langle x,\, x\rangle = 0$ et donc $x = 0$.

Proposition 7.14 – *Pour tout sous-espace vectoriel* F *d'un espace euclidien* E*, on a :*

1. $\dim E = \dim F + \dim F^\perp$;

2. $E = F \oplus F^\perp$;

3. $F^{\perp\perp} = F$

Démonstration : *1.* Soit $\{v_1, ..., v_p\}$ une base de F et complétons-la en une base $\mathcal{B} = \{v_1, ..., v_p, \varepsilon_{p+1}, ..., \varepsilon_n\}$ de E. Notons $(\,a_{ij}\,)$ la matrice du produit scalaire dans cette base :

$$\langle y, x\rangle = \sum_{i,j=1}^{n} a_{ij}\, y_i\, x_j$$

Puisque les composantes des vecteurs $v_1, v_2, ..., v_p$ dans la base $\mathcal{B}$ sont respectivement $(1,0,...,0)$, $(0,1,0,...,0)$, ... $(0,..., \underset{\text{rang p}}{1}, 0,...,0)$, on a :

$$\begin{array}{ccc} \langle v_1, x\rangle = 0 & \Longleftrightarrow & \sum_{j=1}^n a_{1j}x_j = 0 \\ \cdots & \cdots & \cdots \\ \langle v_p, x\rangle = 0 & \Longleftrightarrow & \sum_{j=1}^n a_{pj}x_j = 0 \end{array}$$

Donc $x \in F^\perp$ si et seulement si les composantes de x vérifient le système

$$\left\{ \begin{array}{l} a_{11}\, x_1 + \cdots + a_{1n}\, x_n = 0 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{p1}\, x_1 + \cdots + a_{pn}\, x_n = 0 \end{array} \right.$$

Or dét$\,(\, a_{ij}\,) \neq 0$ (cf. page 235), donc les lignes de la matrice $(\, a_{ij}\,)$ sont indépendantes. Il s'ensuit que ces équations sont indépendantes. Le rang du système est donc p et par conséquent l'ensemble des solutions est de dimension $n-p$, c'est--dire : $\dim F^\perp = n - p$.

2. D'après *1.* il suffira de montrer que $F \cap F^\perp = \{0\}$. Or cela est immédiat. Soit en effet $x \in F \cap F^\perp$. On a :

$$\langle \underset{\in F}{\underset{\uparrow}{x}}\ ,\ \underset{\in F^\perp}{\underset{\uparrow}{x}} \rangle = 0$$

donc $x = 0$, car le produit scalaire est défini positif.

3. On a, tout d'abord $F \subset F^{\perp\perp}$, car si $a \in F$, on a : $\langle a,\, x\rangle = 0\,,\ \forall x \in F^\perp$, c'est-à-dire $a \in F^{\perp\perp}$. D'autre part

$$\dim F^{\perp\perp} = \dim E - \dim F^\perp = \dim E - (\dim E - \dim F) = \dim F.$$

Puisque $F \subset F^{\perp\perp}$, on a $F = F^{\perp\perp}$ $\square$

La propriété 1. vient aussi de la proposition suivante. On sait qu'en dimension finie E et E^* sont isomorphes (car ils ont même dimension) ; mais en général il n'y a pas une manière canonique de construire l'isomorphisme. Cependant si E est euclidien, on peut construire un ***isomorphisme canonique*** entre E et E^*.

Proposition 7.15 – *Soit* $(E, \langle\ ,\ \rangle)$ *un espace euclidien. L'application*

$$\begin{array}{rccc} j: & E & \longrightarrow & E^* \\ & y & \longmapsto & j(y) \end{array} \qquad \text{où}\quad \begin{array}{rccc} j(y): & E & \longrightarrow & \mathbb{R} \\ & x & \longmapsto & \langle x, y\rangle \end{array}$$

est un ***isomorphisme*** *de* E *sur* E^*.
De plus, si F *est un sous-espace vectoriel de* E *et* F^0 *est l'annulateur de* F *(cf. 3.37 page 87), on a :*

$$j(F^\perp) = F^0$$

Démonstration : La linéarité de j vient de la linéarité de $\langle\ ,\ \rangle$ par rapport au premier argument. Puisque E et E^* sont de même dimension, pour montrer que j est un isomorphisme il suffira de montrer que j est injective.
Soit $y \in E$ tel que $j(y) = 0$, c'est-à-dire $\langle x, y\rangle = 0\,,\ \forall x \in E$. On aura, en particulier, $\langle y,\, y\,\rangle = 0$, donc $y = 0$. Ainsi j est injective et par conséquent bijective (car E est de dimension finie). D'autre part :

$$\begin{aligned} j^{-1}(F^0) &= \{y \in E \mid j(y) \in F^0\} = \{y \in E \mid j(y)(x) = 0\,,\quad \forall x \in F\} \\ &= \{y \in E \mid \langle x,\, y\,\rangle = 0\,,\ \forall x \in F\} = F^\perp \end{aligned}$$

Comme j est bijective, ceci est équivalent à $j(F^\perp) = F^0$. $\square$

Il s'ensuit, en particulier, que $\dim E = \dim F + \dim F^\perp$ (cf. proposition 3.39).

Exercices 14. 15. 16. 17. 18. 19. 20.

7.8 Endomorphisme adjoint

La formulation des énoncés et des démonstrations est parfois simplifiée par l'utilisation de la notion d'endomorphisme adjoint.

Proposition 7.16 – *Soit E un espace euclidien et $f \in \mathrm{End}\,(E)$. Il existe un et un seul endomorphisme f^* de E tel que*

$$\langle\, f(x),\, y\rangle \;=\langle\, x,\, f^*(y)\rangle\ , \quad \forall x, y \in E.$$

*f^* est dit **adjoint** de f. Si $\{e_i\}$ est une base* orthonormée *et $A = M(f)_{e_i}$, alors la matrice $A^* = M(f^*)_{e_i}$ est la transposée de A :*

$$A^* = {}^t\!A$$

Démonstration : Soit $\{e_i\}$ une base orthonormée de E et notons

$$A = M(f)_{e_i}, \quad A^* = M(f^*)_{e_i}, \qquad X = M(x)_{e_i}, \quad Y = M(y)_{e_i}$$

Puisque on est dans une base orthonormée, l'identité de l'énoncé s'écrit (cf. remarque page 234) :

$${}^t(A\,X)\,Y \;=\; {}^tX\,A^*\,Y\ , \quad \forall\, X,\, Y \in \mathcal{M}_{n,1}(\mathbb{R})$$

ce qui est équivalent à ${}^tA \;=\; A^*$ (cf. exercice 13). Ceci montre que A^* (donc f^*) est unique.
Réciproquement, si on définit $f^* \in \mathrm{End}\,(E)$ par $M(f^*) \;=\; {}^tA$ on voit, en remontant les calculs, que f^* vérifie l'identité de l'énoncé. $\square$

Nota. – Ceci donne une interprétation de la transposée d'une matrice :

> *Soit $A \in \mathcal{M}_n(\mathbb{R})$ et f l'endomorphisme de $\mathbb{R}^n$ qui* dans la base canonique *est représenté par A ; alors ${}^tA$ représente dans la base canonique l'endomorphisme f^* adjoint de f relativement au produit scalaire canonique sur $\mathbb{R}^n$, $\langle x\,,\,y\rangle = x_1\,y_1 + \cdots + x_n\,y_n$.*

On a les propriétés suivantes, qui se démontrent immédiatement dans une base orthonormée, compte tenu du fait que dans une base orthonormée $A^* \;=\; {}^tA$.

Proposition 7.17 – *Pour tous endomorphismes f et g et pour tout scalaire λ, on a :*

a) $f^{**} = f\ , \qquad (\mathrm{id})^* = \mathrm{id}\,.$

b) $(f+g)^* = f^* + g^*\ , \qquad (\lambda f)^* = \lambda f^*\ , \qquad (f\circ g)^* = g^*\circ f^*.$

c) $\mathrm{rg}\ f^* = \mathrm{rg}\ f\ , \qquad \mathrm{dét}\,f^* = \mathrm{dét}\,f.$

Exercices 21. 22. 23.

7.9 Groupe orthogonal

Le but de ce paragraphe est d'étudier les endomorphismes f d'un espace euclidien qui conservent la norme des vecteurs, c'est-à-dire tels que $\|f(x)\| \;=\; \|x\|$ pour tout $x \in E$. Il s'agit donc de la généralisation des *isométries vectorielles* de $\mathbb{R}^2$ et de $\mathbb{R}^3$ (c'est-à-dire les rotations et les réflexions) [8].

[8] *Attention* : les translations ne sont pas des isométries vectorielles, car elles ne sont pas des applications linéaires (elles sont des applications affines : cf. Appendice A.7)

Définition 7.18 – *Soit E un espace euclidien et $f \in \mathrm{End}\,(E)$. On dit que f est une* ***transformation orthogonale*** *si*

$$\langle f(x),\, f(y)\rangle \;=\; \langle x,\, y\rangle\;, \quad \forall x,\, y \in E.$$

On note $\mathrm{O}(E)$ l'ensemble des transformations orthogonales de E.

Proposition 7.19 – *Les propriétés suivantes sont équivalentes :*

1. $\langle f(x),\, f(y)\rangle \;=\; \langle x,\, y\rangle\;, \quad \forall x,\, y \in E.$
2. $\|f(x)\| = \|x\|\,, \qquad\qquad \forall x \in E.$
3. *Si $\{e_i\}$ est une* base orthonormée *et $A = M(f)_{e_i}$, alors ${}^tAA = I$ (ou, d'une manière équivalente, $A\,{}^tA = I$).*

Démonstration : 1. $\Longrightarrow$ 2. est évidente : il suffit de faire $x = y$.
Pour montrer que 2. $\Longrightarrow$ 1., on utilise l'identité (2), page 233 , qui exprime le produit scalaire en fonction de la norme. On a :

$$\begin{aligned}\langle f(x),\, f(y)\rangle \;&=\; \tfrac{1}{2}\big(\|f(x)+f(y)\|^2 - \|f(x)\|^2 - \|f(y)\|^2\big)\\ &=\; \tfrac{1}{2}\big(\|f(x+y)\|^2 - \|f(x)\|^2 - \|f(y)\|^2\big)\end{aligned}$$

donc, si f conserve la norme :

$$\langle f(x),\, f(y)\rangle = \frac{1}{2}\big(\|x+y\|^2 - \|x\|^2 - \|y\|^2\big) \;=\; \langle x,\, y\rangle\,.$$

Montrons que 1. $\Longleftrightarrow$ 3.
En écrivant 1. dans une base orthonormée, on a (cf. remarque page 234) :

$$\begin{aligned}\langle f(x),\, f(y)\rangle \;&=\; \langle x,\, y\rangle\;,\ \forall x, y \in E\\ \Longleftrightarrow\;& {}^t(AX)AY = {}^tXY,\ \ \forall X, Y \in \mathcal{M}_{n,1}(\mathbb{R})\\ \Longleftrightarrow\;& {}^tAA = I\end{aligned}$$

ce qui montre que 1. $\Longleftrightarrow$ 3.
D'autre part, puisque ${}^tAA = I$, on a $\mathrm{dét}\,({}^tAA) = 1$ c'est-à-dire :
$\mathrm{dét}\,({}^tA)(\mathrm{dét}\,A) = 1$. Or $\mathrm{dét}\,{}^tA = \mathrm{dét}\,A$; donc $(\mathrm{dét}\,A)^2 = 1$, et par conséquent

$$\mathrm{dét}\,A = \pm 1.$$

On en déduit que A est inversible. En multipliant par A^{-1} à droite, l'équation ${}^tAA = I$, on obtient $A^{-1} = {}^tA$; d'où, en multipliant à gauche par A : $A\,{}^tA = I$. $\square$

Remarque. — La propriété 3. montre que les transformations orthogonales f sont caractérisées par :

$$f^* \circ f = \mathrm{id}\,.$$

Propriétés 7.20 – *Soit f une transformation orthogonale ; alors :*

1. *Les valeurs propres de f sont $+1$ ou -1.*
2. $\mathrm{dét}\, f = \pm 1$. *En particulier f est bijective.*

Les transformations orthogonales de déterminant + 1 sont dites ***directes*** *; celle de déterminant – 1 sont dites* ***indirectes*** *ou "gauches" .*

En effet, si v est vecteur propre correspondant à λ, on a : $\|f(v)\| = \|\lambda\, v\| = |\,\lambda\,|\;\|v\|$; or f est orthogonale, donc $\|f(v)\| = \|v\|$ et, comme $v \neq 0$, on a $|\,\lambda\,| = 1$.
D'autre part, de $f^* \circ f = \mathrm{id}$, il vient immédiatement $(\mathrm{dét}\, f)^2 = 1$. $\square$

Remarque. − Si A est la matrice qui représente une transformation orthogonale, ses valeurs propres réelles sont donc $+1$ ou -1. Cependant, vue comme matrice de $\mathcal{M}_n(\mathbb{C})$, A peut avoir des valeurs propres complexes. Nous verrons par la suite (cf. remarque page 284) que les valeurs propres complexes de A sont de module 1, mais ceci ne résulte pas de la propriété ci-dessus, qui n'a de sens que dans le cas réel, puisque f est un endomorphisme d'un espace vectoriel réel.

Les transformations orthogonales peuvent être caractérisées aussi par la propriété suivante :

Proposition 7.21 – *f est orthogonale si et seulement si elle transforme toute base orthonormée en une base orthonormée. Pour cela il suffit qu'il existe* ***une*** *base orthonormée qui, par f, est transformée en une base orthonormée.*

En effet soit f orthogonale ; puisque f est bijective, elle transforme toute base en une base. Soit $\{e_i\}$ une base orthonormée, on a :

$$\langle f(e_i),\, f(e_j)\rangle \;=\; \langle e_i,\, e_j\rangle \;=\delta_{ij}$$

donc $\{f(e_i)\}$ est une base orthonormée.
Réciproquement, supposons qu'il existe une base orthonormée $\{e_i\}$ telle que $\{f(e_i)\}$ soit une base orthonormée, et soient $x = \sum_{i=1}^n x_i e_i$, $y = \sum_{i=1}^n y_i e_i$.
Puisque $\{e_i\}$ est orthonormée, on a :

$$\langle x,\, y\rangle = \textstyle\sum_{i=1}^n x_i\, y_i$$

D'autre part :

$$\begin{aligned}\langle f(x),\, f(y)\rangle \;&=\; \Big\langle \textstyle\sum_{i=1}^n x_i\, f(e_i),\, \sum_{j=1}^n y_j\, f(e_j)\Big\rangle \;=\; \sum_{i,j=1}^n x_i\, y_j\, \langle f(e_i),\, f(e_j)\rangle \\ &= \textstyle\sum_{i,j=1}^n x_i\, y_j\, \delta_{ij} \;=\; \sum_{i=1}^n x_i\, y_i = \langle x,\, y\rangle\end{aligned}$$

donc f est orthogonale. $\square$

Définition 7.22 – *L'ensemble*

$$\mathrm{O}(n,\mathbb{R}\,) := \{A \in \mathcal{M}_n(\mathbb{R}) \mid {}^tAA = I\} = \{A \in \mathcal{M}_n(\mathbb{R}) \mid A\,{}^tA = I\}$$

vérifie les propriétés suivantes :

(a) *si* $A,\, B \in \mathrm{O}(n,\mathbb{R}\,)$ *alors* $AB \in \mathrm{O}(n,\mathbb{R}\,)$ *;*

(b) $I \in \mathrm{O}(n,\mathbb{R}\,)$ *;*

(c) *si* $A \in \mathrm{O}(n,\mathbb{R}\,)$ *alors* $A^{-1} \in \mathrm{O}(n,\mathbb{R}\,)$.

En particulier, $\mathrm{O}(n,\mathbb{R}\,)$ *est un sous-groupe* [9] *du groupe linéaire* $\mathrm{GL}(n,\mathbb{R})$ *dit* ***groupe orthogonal.***

La vérification est laissée en exercice.

La signification des matrices orthogonales est claire : elles représentent ***dans une base orthonormée*** les transformations orthogonales d'un espace euclidien. On peut dire aussi qu'un matrice A est orthogonale si et seulement si elle représente une isométrie vectorielle de $\mathbb{R}^n$ dans la base canonique.

[9] cf. Appendice A.1

Il est clair que si $A \in O(n,\mathbb{R})$, dét $A = \pm 1$. On a donc deux types de matrices orthogonales : celles de déterminant $+1$, dites *matrices orthogonales directes*, et celles de déterminant -1, dites *matrices orthogonales gauches*. L'ensemble des matrices orthogonales directes est noté $SO(n,\mathbb{R})$. On vérifie facilement la propriété suivante :

Proposition 7.23 – *L'ensemble des matrices orthogonales directes*

$$SO(n,\mathbb{R}) = \{A \in O(n,\mathbb{R}) \,|\, \text{dét}\, A = 1\}$$

est un groupe, dit ***groupe spécial orthogonal.***

Exemple – La matrice

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

est orthogonale. On peut vérifier en effet que ${}^tAA = I$. Plus simplement, il suffit de vérifier que les vecteurs colonnes c_1, c_2, c_3 forment une famille orthonormée, c'est-à-dire :

$$\|c_i\|^2 = 1 \quad \text{et} \quad \langle c_i, c_j\rangle = 0 \quad \text{si } i \neq j.$$

En effet si l'on interprète A comme la matrice d'un endomorphisme f de $\mathbb{R}^3$ dans la base canonique, on a $c_i = f(e_i)$. Donc f transforme la base orthonormée $\{e_i\}$ en la base orthonormée $\{c_i\}$ et par conséquent, d'après la proposition 7.21, f est orthogonale. On voit immédiatement que dét $A = +1$; f est donc une transformation orthogonale directe.

Proposition 7.24 *La matrice de passage d'une base orthonormée à une base orthonormée est une matrice orthogonale.*

En effet, soient $\{e_i\}$ et $\{e_{i'}\}$ deux bases orthonormées et f l'endomorphisme défini par $f(e_i) = e'_i$ $(i = 1, \cdots, n)$. On a $M(f)_{e_i} = \| f(e_1), ..., f(e_n) \|_{e_i} = P_{e_i \to e'_i}$. D'après la proposition 7.21, f est orthogonale et donc $P_{e_i \to e'_i}$ est une matrice orthogonale.

Exercices 24. 25. 26.

7.10 Étude de $O(2,\mathbb{R})$ et $O(3,\mathbb{R})$

Dans cette section nous utilisons les notions intuitives et élémentaires de rotation, angle de rotation, angle orienté, etc. vues dans le paragraphe 7.1. Dans le paragraphe suivant nous donnerons les définitions précises de ces notions.

Étude du groupe $O(2,\mathbb{R})$

Soit

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

$A \in O(2,\mathbb{R})$ si et seulement si le système des vecteurs colonnes est orthonormé, c'est-à-dire :

$$\begin{cases} a^2 + c^2 = 1 \\ b^2 + d^2 = 1 \\ a\,b + c\,d = 0 \end{cases}$$

D'après la première équation, il existe $\theta \in \mathbb{R}$ tel que :

$$a = \cos\theta \quad , \quad c = \sin\theta.$$

De même, d'après la deuxième équation, il existe $\varphi \in \mathbb{R}$ tel que :

$$b = \cos\varphi \quad , \quad d = \sin\varphi.$$

Enfin,

$$a\,b + c\,d = 0 \Longleftrightarrow \cos\theta\cos\varphi + \sin\theta\sin\varphi = 0 \Longleftrightarrow \cos(\theta - \varphi) = 0$$

c'est-à-dire :

$$\varphi - \theta = (2k+1)\pi/2$$

Donc :

$$\begin{aligned} b &= \cos\Big(\theta + (2k+1)\pi/2\Big) = (-1)^{k+1}\sin\theta \\ d &= \sin\Big(\theta + (2k+1)\pi/2\Big) = (-1)^{k}\cos\theta \end{aligned}$$

Ainsi $A \in \mathrm{O}\,(2, \mathbb{R}\,)$ si et seulement si :

$$A = \begin{pmatrix} \cos\theta & (-1)^{k+1}\sin\theta \\ \sin\theta & (-1)^k\cos\theta \end{pmatrix}$$

Puisque $\text{dét}\,A = (-1)^k(\cos^2\theta + \sin^2\theta) = (-1)^k$, $A \in \mathrm{SO}\,(2, \mathbb{R}\,)$ si et seulement si k est pair. On a donc :

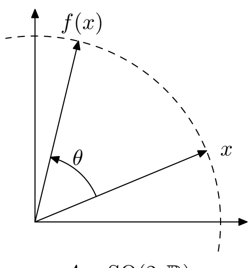
Proposition 7.25 – *Soit* $A \in \mathrm{O}(2, \mathbb{R})$; *alors :*

– *Soit* $A \in \mathrm{SO}\,(2, \mathbb{R}\,)$ *et dans ce cas* $A = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$

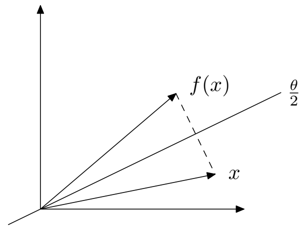
(rotation d'angle θ et de centre O, cf. Exemple 5, page 69).

– *Soit* $A \notin \mathrm{SO}\,(2, \mathbb{R}\,)$ *et alors* $A = \begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix}$

Dans ce cas A représente la symétrie orthogonale par rapport à la droite d'angle polaire $\theta/2$ (cf. exercice 3)



$A \in \mathrm{SO}(2, \mathbb{R})$



$A \notin \mathrm{SO}(2, \mathbb{R})$

Figure 11

Etude du groupe $\mathrm{O}(\mathbf{3}, \mathbb{R})$

Proposition 7.26 – *Soit $A \in \mathrm{O}(3, \mathbb{R})$ et f l'endomorphisme de $\mathbb{R}^3$ tel que $A = M(f)_{e_i}$, $\{e_i\}$ étant la base canonique. Il existe alors une base $\{e'_1, e'_2, e'_3\}$ de $\mathbb{R}^3$, orthonormée, telle que*

$$A' \equiv M(f)_{e'_i} = \begin{pmatrix} \boxed{\begin{matrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{matrix}} & \begin{matrix} 0 \\ 0 \end{matrix} \\ \begin{matrix} 0 & \quad 0 \end{matrix} & \varepsilon \end{pmatrix}$$

où $\varepsilon = +1$ *si* $\text{dét}\, A = 1$, *c'est-à-dire si* $A \in \mathrm{SO}(3, \mathbb{R})$ *;*
et $\varepsilon = -1$ *si* $\text{dét}\, A = -1$, *c'est-à-dire si* $A \notin \mathrm{SO}(3, \mathbb{R})$.

Démonstration : Si $A = \pm I$ le résultat est trivial (prendre pour $\{e'_i\}$ la base canonique et $\theta = 0$ si $A = I$, $\theta = \pi$ si $A = -I$).

Supposons donc $A \neq \pm I$. La démonstration se fait par les étapes suivantes :

Lemme 1. *Si* $\text{dét}\, A = 1$, $\lambda = 1$ *est valeur propre d'ordre 1 ou 3.*
Si $\text{dét}\, A = -1$, $\lambda = -1$ *est valeur propre d'ordre 1 ou 3.*

En effet, supposons $\text{dét}\, A = 1$ (le cas $\text{dét}\, A = -1$ se traite d'une manière analogue.) Si les trois valeurs propres $\lambda_1, \lambda_2, \lambda_3$ de A sont réelles, puisque $\lambda_1\lambda_2\lambda_3 = \text{dét}\, A = 1$ et $\lambda_i = \pm 1$, on a nécessairement : soit $\lambda_1 = 1$ et $\lambda_2 = \lambda_3 = -1$, soit $\lambda_1 = \lambda_2 = \lambda_3 = 1$. Si l'une des valeurs propres est complexe, alors, en la notant μ, $\bar\mu$ est aussi valeur propre, puisque A est réelle (cf. exercice 7 chapitre 6). Ainsi $\text{dét}\, A = \lambda\,\bar\mu$ où λ est l'autre valeur propre, nécessairement réelle, donc égale à $\lambda = \pm 1$. Or $\text{dét}\, A = 1$ donc $\lambda > 0$ et par conséquent $\lambda = 1$. $\diamondsuit$

Lemme 2. *Si* $\text{dét}\, A = 1$ *alors* $\dim E_1 = 1$.
Si $\text{dét}\, A = -1$ *alors* $\dim E_{-1} = 1$

En effet, supposons $\text{dét}\, A = 1$ ($\lambda = 1$ est alors valeur propre simple ou triple). Si $\dim E_1 = 3$, on a $E_1 = \mathbb{R}^3$ donc $A = I$, ce qui est exclu. Supposons par l'absurde que $\dim E_1 = 2$ ($\lambda = 1$ est alors valeur propre triple). Soit $\{v_1, v_2\}$ une base de E_1 et $w \in \mathrm{Vect}\{v_1, v_2\}^\perp, w \neq 0$. On a :

$$\langle f(w), v_i\rangle \underset{\substack{\uparrow \\ \text{car} \\ f(v_i)=v_i}}{=} \langle f(w), f(v_i)\rangle \underset{\substack{\uparrow \\ \text{car } f \\ \text{orthogonale}}}{=} \langle w, v_i\rangle = 0 \quad \text{pour } i = 1, 2\ ,$$

donc $f(w) \in \mathrm{Vect}\{v_1, v_2\}^\perp$, c'est-à-dire : $f(w) \in \mathrm{Vect}\{w\}$. Il existe donc $\lambda_0 \in \mathbb{R}$ tel que $f(w) = \lambda_0\, w$. Mais 1 est valeur propre triple, donc $\lambda_0 = 1$ et $f(w) = w$, c'est-à-dire $w \in E_1$, ce qui est exclu car $w \in E_1^\perp$. On a donc nécessairement $\dim E_1 = 1$.
Le cas $\text{dét}\, A = -1$ se traite d'une manière analogue. $\diamondsuit$

Lemme 3. *Si* $\text{dét}\, A = +1$ *(respect. :* $\text{dét}\, A = -1$*), alors le plan* $\pi = E_1^\perp$ *(respectiv. :* $\pi = {E_{-1}}^\perp$*) est invariant par* f *et la restriction de* f *à* π *est une rotation.*

Soit en effet $x \in \pi$, c'est-à-dire $\langle x, w\rangle = 0$. Comme f est orthogonale : $\langle f(x), f(w)\rangle = 0$. Or $f(w) = \pm w$, donc $\langle f(x), w\rangle = 0$, c'est-à-dire $f(x) \in \pi$. Ainsi π est stable par f.

Soit $\tilde f = f|_\pi$. $\forall x, y \in \pi$, on a :

$$\left\langle \tilde f(x), \tilde f(y)\right\rangle_\pi = \langle f(x), f(y)\rangle_E = \langle x, y\rangle_E = \langle x, y\rangle_\pi$$

donc $\tilde f$ est orthogonale.

Montrons que $\text{dét}\,\tilde f = 1$. Si $\{v_1, v_2\}$ est une base de π ; on a :

$$M(f)_{\{v_1,v_2,w\}} = \begin{pmatrix} a & b & 0 \\ c & d & 0 \\ 0 & 0 & \varepsilon \end{pmatrix}, \quad \text{où :} \quad \begin{pmatrix} a & b \\ c & d \end{pmatrix} = M(\tilde f)_{\{v_1,v_2\}}$$

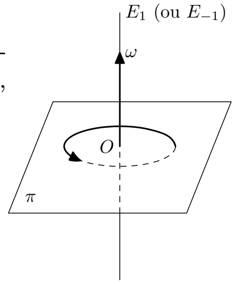


Figure 12

et $\varepsilon = \pm 1$ selon que $\text{dét}\,A = \pm 1$ (c'est-à-dire $\varepsilon = \text{dét}\,A$). Or le déterminant de cette matrice vaut $\text{dét}\,A$, car le déterminant est invariant par changement de base. Donc :

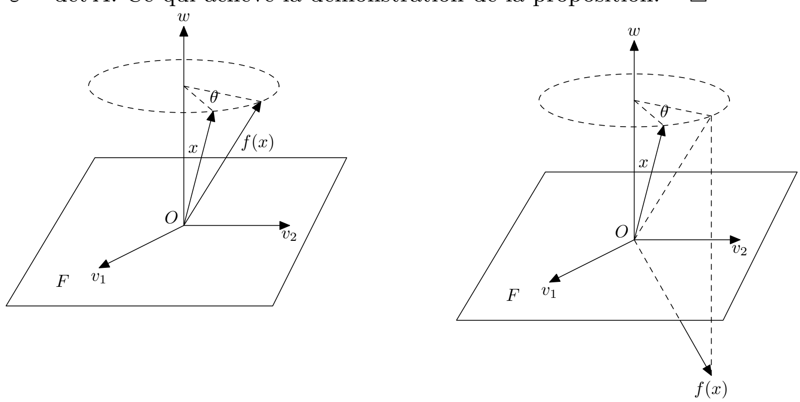
$$\text{dét}\,A = \begin{vmatrix} a & b \\ c & d \end{vmatrix} \cdot \varepsilon = \begin{vmatrix} a & b \\ c & d \end{vmatrix} \cdot \text{dét}\,A.$$

Ainsi $\begin{vmatrix} a & b \\ c & d \end{vmatrix} = 1$ et donc $\tilde f$ est une rotation. $\diamondsuit$

Il existe donc une base orthonormée $\{e'_1, e'_2, e'_3\}$ avec $e'_1, e'_2 \in \pi$ et $e'_3 \in E_1$ (respect : $e'_3 \in E_{-1}$) telle que :

$$M(f)_{e'_i} = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & \varepsilon \end{pmatrix}$$

avec $\varepsilon = \text{dét}\,A$. Ce qui achève la démonstration de la proposition. $\square$



$\text{dét}\,A = 1$ $\qquad$ $\text{dét}\,A = -1$

Figure 13

Ainsi, si $\text{dét}\,A = 1$, f représente une *rotation autour de l'axe* E_1. Puisque la trace est invariante par changement de base, l'angle (non orienté) de rotation est donné par la relation :

$$\text{Tr}\,A = 2\cos\theta + 1.$$

Si $\text{dét}\,A = -1$, on peut écrire $f = h \circ g$ avec :

$$M(g)_{e_i} = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{et} \quad M(h)_{e'_i} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

f est donc la *rotation autour de l'axe* E_{-1} *suivie de la symétrie orthogonale par rapport au plan* ${E_{-1}}^{\perp}$. L'angle de rotation est donné dans ce cas par

$$\operatorname{Tr} A = 2\cos\theta - 1.$$

En résumant :

Proposition 7.27 – *Soit* $A \in \mathrm{O}(3, \mathbb{R})$, $A \neq \pm I$.

Si $\operatorname{dét} A = 1$, *A représente dans la base canonique de $\mathbb{R}^3$ une* ***rotation autour de l'axe*** *E_1 ; l'angle (non orienté) de rotation est donné par :*

$$\operatorname{Tr} A = 2\cos\theta + 1.$$

Si $\operatorname{dét} A = -1$, *$A$ représente dans la base canonique de $\mathbb{R}^3$* ***une rotation autour de l'axe E_{-1} suivie de la symétrie orthogonale par rapport au plan*** *${E_{-1}}^{\perp}$; l'angle (non orienté) de rotation est donné par*

$$\operatorname{Tr} A = 2\cos\theta - 1.$$

En particulier, si $\operatorname{Tr} A = 1$ *on a* $\theta = 0$ *: A représente dans ce cas la symétrie orthogonale par rapport au plan ${E_{-1}}^{\perp}$, dite aussi* ***réflexion*** *par rapport à ${E_{-1}}^{\perp}$.*

(Pour une généralisation de ce résultat en dimension quelconque, cf. exercice 28)

Calcul de l'angle orienté de rotation.

Pour calculer l'angle orienté de rotation, il faut fixer une orientation du plan de rotation, ce que l'on peut faire en fixant l'orientation de la normale (cf. page 220). En choisissant donc un vecteur $\vec{n}$ de E_1 (respect. de E_{-1}) et un vecteur u du plan de rotation, l'angle orienté de rotation est $(u,\, f(u))$ et donc, d'après la formule (4) page 221 :

$$\sin\theta = \frac{\operatorname{dét} \|u,\, f(u),\, \vec{n}\,\|}{\|u\|^2\|\vec{n}\,\|}$$

car $\|f(u)\| = \|u\|$, puisque f est orthogonale.

Exemple –

Considérons la matrice

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

Nous avons vu (cf. page 241) qu'elle représente une rotation dans la base canonique de $\mathbb{R}^3$. L'axe de rotation est déterminé par un vecteur propre correspondant à la valeur propre $\lambda = 1$. On trouve facilement que E_1 est engendré par le vecteur $w = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$.

L'angle (non orienté) de rotation est déterminé par la relation : $\operatorname{Tr} A = 2\cos\theta + 1$, ce qui donne $\cos\theta = \dfrac{1}{2}$ et donc $\theta = \pm\pi/3$.

Pour déterminer l'angle orienté, on choisit un vecteur directeur de la normale, par exemple $w = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ et un vecteur $u \in E_1^{\perp}$ ($E_1^{\perp}$ est le plan $x + y + z = 0$), par exemple $u = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$. En calculant $f(u)$ on trouve : $f(u) = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$.

On a :

$$\sin\theta = \frac{1}{2\sqrt{3}}\begin{vmatrix} 1 & 1 & 1 \\ -1 & 0 & 1 \\ 0 & 1 & 1 \end{vmatrix} = \frac{\sqrt{3}}{2}.$$

Donc $\theta = \pi/3$.
Ainsi A représente, dans la base canonique de $\mathbb{R}^3$, la rotation d'angle $\pi/3$ autour de l'axe D dirigé et orienté par le vecteur $w = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$.

Exercices 27. 28. 29.

7.11 Rotations et angle dans un espace euclidien de dimension 2 ou 3

L'étude du paragraphe précédent utilisait les notions intuitives de rotation et angle. Elle nous permet de justifier les définitions précises de ces notions dans un espace euclidien de dimension 2 ou 3.

Définition 7.28 - *Soit E un espace euclidien* de dimension 2 ou 3. *On appelle* ***rotations*** *les endomorphismes $f \in \mathrm{SO}(E)$.*

Pour ce qui est de la notion d'angle, la définition d'angle non orienté est très simple. Soient $u, v \in E \setminus \{0\}$. D'après l'inégalité de Cauchy-Schwartz on a : $\frac{|\langle u,v \rangle|}{\|u\| \, \|v\|} \leq 1$. Il existe donc un unique $\theta \in [0, \pi]$, tel que

$$\boxed{\cos \theta = \frac{\langle u, v \rangle}{\|u\| \, \|v\|}} \qquad (*)$$

θ est dit *angle non orienté* entre les vecteurs u et v.
La notion d'angle orienté est plus délicate et elle fait intervenir, bien entendu, l'orientation du plan défini par les deux vecteurs (cf. page 220)[10].

Angle orienté en dimension 2

Soit $\left(\vec{E}_2, \langle\, , \,\rangle\right)$ un espace euclidien *orienté* de dimension 2 [11] et f une rotation. Comme on l'a vu (cf. Proposition 7.25), si $\mathcal{B}$ est une b.o.n directe,

$$M(f)_{\mathcal{B}} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

La valeur de $\theta \in]-\pi, \pi]$ [12] dans cette matrice est dite *angle de rotation*. Par la suite, une rotation d'angle θ sera notée $\mathcal{R}_\theta$.

[10] Si on a pu définir un angle orienté dans $\mathbb{R}^2$ ou $\mathbb{R}^3$ c'est parce que ces espaces sont naturellement orientés par leurs bases canoniques.
[11] Nous notons cet espace $\vec{E}_2$, plutt que E_2, justement pour souligner qu'il doit être orienté.
[12] ou, si l'on préfère : $\theta \in [0, 2\pi[$.

Proposition 7.29 *Soient* $u, v \in \vec{E}_2 \setminus \{0\}$ *et* $U = \dfrac{u}{\|u\|}$, $V = \dfrac{v}{\|v\|}$. *Il existe alors un unique* $\theta \in]-\pi, \pi]$ *tel que* $\mathcal{R}_\theta(U) = V$. θ *est appelé* ***angle orienté*** *entre u et v et est noté* (uv).

Démonstration : Soient $U = (U_1, U_2)$, $V = (V_1, V_2)$ les composantes de U, V dans une b.o.n. directe. Puisque

$$\begin{cases} U_1^2 + U_2^2 = 1 \\ V_1^2 + V_2^2 = 1 \end{cases}$$

il existe $\alpha, \beta \in \mathbb{R}$ tels que $U_1 = \cos\alpha$, $U_2 = \sin\alpha$ et $V_1 = \cos\beta$, $V_2 = \sin\beta$. On a :

$$\mathcal{R}_\theta(U) = V \iff \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} \cos\alpha \\ \sin\alpha \end{pmatrix} = \begin{pmatrix} \cos\beta \\ \sin\beta \end{pmatrix} \iff$$

$$\begin{cases} \cos\theta\cos\alpha - \sin\theta\sin\alpha = \cos\beta \\ \sin\theta\cos\alpha + \sin\alpha\cos\theta = \sin\beta \end{cases} \iff \begin{cases} \cos(\theta+\alpha) = \cos\beta \\ \sin(\theta+\alpha) = \sin\beta \end{cases} \Leftrightarrow \theta = \beta - \alpha \mod [2\pi]$$

□

Exemple . Soit $\mathbb{R}^2$ muni de l'orientation canonique. Chercher l'angle entre les vecteurs $u = (1, 1)$ et $v = (1 + \sqrt{3}, 1 - \sqrt{3})$.
On a : $U = \dfrac{1}{\sqrt{2}}(1, 1)$ et $V = \dfrac{1}{\sqrt{8}}(1 + \sqrt{3},\, 1 - \sqrt{3})$. Donc $R_\theta(U) = V$ si et seulement si :

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \frac{1}{2\sqrt{2}} \begin{pmatrix} 1+\sqrt{3} \\ 1-\sqrt{3} \end{pmatrix}$$

c'est-à-dire :

$$\begin{cases} \cos\theta - \sin\theta = \dfrac{1+\sqrt{3}}{\sqrt{2}} \\ \sin\theta + \cos\theta = \dfrac{1-\sqrt{3}}{\sqrt{2}} \end{cases}$$

ce qui donne : $\cos\theta = \dfrac{1}{2}$ et $\sin\theta = -\dfrac{\sqrt{3}}{2}$ c'est-à-dire : $\theta = -\dfrac{\pi}{3}$.

Pour calculer $\cos\theta$ et $\sin\theta$ on a, en fait, les formules suivantes :

$$\boxed{\cos\,\theta = \frac{\langle\, u, v\,\rangle}{\|u\|\,\|v\|}, \qquad \sin\theta = \frac{\text{dét}\,\|u, v\|_{(e_i)}}{\|u\|\,\|v\|}} \qquad (\star)$$

où (e_i) est une b.o.n.*directe*. En effet, il suffit de résoudre par rapport à $\cos\theta$ et $\sin\theta$ le système

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} U_1 \\ U_2 \end{pmatrix} = \begin{pmatrix} V_1 \\ V_2 \end{pmatrix}$$

Un petit calcul donne :

$$\cos\theta = U_1V_1 + U_2V_2 = \langle\, U,\, V\,\rangle \quad \text{et} \quad \sin\theta = U_1V_2 - U_2V_1 = \text{dét}\,\|U, V\|_{(e_i)}$$

d'où les formules $(\star)$.

Propriétés 7.30

1. $\mathcal{R}_\theta \circ \mathcal{R}_{\theta'} = \mathcal{R}_{\theta+\theta'}$.
2. *Relation de Chasles pour les angles :* $(uv) + (vw) = (uw) \mod[2\pi]$.
   En particulier, en faisant $u = w$, *on a :* $(uv) = -(vu)$

La première est une simple vérification. Pour la seconde, si U, V, W sont trois vecteurs unitaires, ona :

$$W = \mathcal{R}_{(vw)}(V) \quad \text{et} \quad V = \mathcal{R}_{(uv)}(U), \quad \text{donc :} \quad W = \mathcal{R}_{(vw)} \circ \mathcal{R}_{(uv)}(U)$$

D'autre part : $W = \mathcal{R}_{(u,w)}(U)$ donc $\mathcal{R}_{(vw)} \circ \mathcal{R}_{(uv)}(U) = \mathcal{R}_{(uw)}(U)$. Comme U est arbitraire :

$$\mathcal{R}_{(vw)+(uv)} = \mathcal{R}_{(uw)}$$

d'où, à cause de l'unicité de l'angle orienté, la relation de Chasles. $\square$

Angle orienté en dimension 3

Soit (π) un plan vectoriel dans un espace euclidien *orienté* de dimension 3, $\vec{E}_3$. On suppose (π) muni du produit scalaire induit et de l'orientation définie par le choix d'un vecteur normal $\vec{n}$ que l'on supposera unitaire. Puisque (π) est orienté, on peut définir l'angle orienté $\theta = (uv)$ entre deux vecteurs $u, v \in (\pi) \setminus \{0\}$. Soit $(\varepsilon_1, \varepsilon_2)$ une base directe de (π) (cela veut dire que $\mathcal{B}' = (\varepsilon_1, \varepsilon_2, \vec{n})$ est une base directe de $\vec{E}_3$). On a alors

$$\sin\theta = \frac{\text{dét}\,\|u, v, \vec{n}\,\|_{\mathcal{B}'}}{\|u\|\,\|v\|}$$

(si u, v sont indépendants ; si u, v sont liés on pose $\sin\theta = 0$). En effet, si $U = \dfrac{u}{\|u\|}$ et $V = \dfrac{v}{\|v\|}$, on a :

$$\frac{\text{dét}\,\|u, v, \vec{n}\,\|_{\mathcal{B}'}}{\|u\|\,\|v\|} = \text{dét}\,\|U, V, \vec{n}\|_{\mathcal{B}'} = \begin{vmatrix} U_1 & V_1 & 0 \\ U_2 & V_2 & 0 \\ 0 & 0 & 1 \end{vmatrix} = \text{dét}\,\|U, V\|_{\varepsilon_1, \varepsilon_2} = \sin\theta$$

On a en fait :

Proposition 7.31 *Soit $\vec{n}$ un vecteur unitaire normal au plan* $\text{Vect}\,\{u, v\}$ *et θ l'angle orienté (uv) pour l'orientation défine par $\vec{n}$. Pour toute b.o.n. directe $\mathcal{B}$ (et non seulement pour les bases dont deux vecteurs sont dans le plan de rotation), on a :*

$$\boxed{\sin\theta = \frac{\text{dét}\,\|u, v, \vec{n}\,\|_{\mathcal{B}}}{\|u\|\,\|v\|}}$$

En effet, si $v_1, ..., v_n$ sont n vecteurs d'un espace vectoriel E de dimension n et (e_i), (e'_i) sont deux bases, alors (cf. exercice 17, chapitre 4) :

$$\text{dét}\, \|v_1, ..., v_n\|_{e_i} = \text{dét}\, P_{e_i \to e'_i} \cdot \text{dét}\, \|v_1, ..., v_n\|_{e'_i}$$

Or la matrice de passage de la base orthonormée directe $\mathcal{B}' = (\varepsilon_1, \varepsilon_2, \vec{n})$ ci-dessus à la base orthonormée directe appartient à $\mathrm{SO}(n,3)$ donc son déterminant vaut 1. On en déduit que $\text{dét}\, \|u, v, \vec{n}\,\|_{\mathcal{B}'} = \text{dét}\, \|u, v, \vec{n}\,\|_{\mathcal{B}}$, d'où la formule.

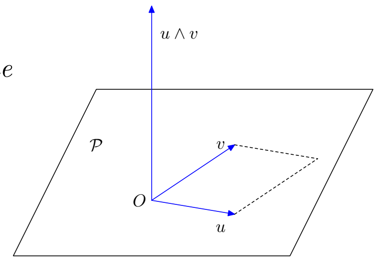
Exercice 30 .

7.12 Produit vectoriel

La notion de produit vectoriel est employée en Mécanique, notamment pour définir le moment d'une force. On peut en donner plusieurs définitions équivalentes : chacune présente son propre intérêt que nous mettrons en évidence à la fin de cette section.
Dans tout ce paragraphe, $\vec{E}_3$ est un espace vectoriel euclidien de dimension 3, *orienté*.

Définition 7.32 - *Soient u et v deux vecteurs indépendants de $\vec{E}_3$. Le* ***produit vectoriel*** *de u et v est le vecteur noté $u \wedge v$ défini par les conditions suivantes :*

- *il est orthogonal au plan* $\mathrm{Vect}\,\{u, v\}$ *;*
- $\|u \wedge v\| = \mathcal{A}ire(\mathcal{P})$ *, où $\mathcal{P}$ est le parallélogramme défini par les vecteurs u et v ;*
- *$(u, v, u \wedge v)$ est une base directe de $\vec{E}_3$.*



Si u et v sont liés, on pose : $u \wedge v = 0$.

Propriétés 7.33 .

1. *u et v sont liés si et seulement si $u \wedge v = 0$.*
2. *$u \wedge v = -v \wedge u$.*
3. *Soit $\vec{n}$ un vecteur unitaire normal au plan* $\mathrm{Vect}\,\{u, v\}$ *et θ l'angle orienté (uv) pour l'orientation défine par $\vec{n}$. Alors :*

$$u \wedge v = \|u\|\, \|v\|\, \sin\theta\, \vec{n}$$

4. *Si (e_i) est une b.o.n. directe et $\vec{n}$ un vecteur unitaire normal au plan* $\mathrm{Vect}\,\{u, v\}$, *alors :*

$$u \wedge v = \text{dét}\, \|u, v, \vec{n}\|_{e_i} \vec{n}$$

Démonstration :

1. et 2. sont triviales.
3. Soit $\vec{N}$ le vecteur unitaire de direction $u \wedge v$. D'après la définition : $u \wedge v = \mathcal{A}ire(\mathcal{P})\vec{N}$, donc, en utilisant la formule (3) page 220, $u \wedge v = \|u\|\, \|v\|\, |\sin\theta|\, \vec{N}$. Or $\sin\theta > 0$ si $\vec{n} = \vec{N}$ et $\sin\theta < 0$ si $\vec{n} = -\vec{N}$, d'où la formule.
4. est une conséquence immédiate de la proposition 7.31. $\square$

Corollaire 7.34 *L'application* $E \times E \longrightarrow E$ *définie par :* $(u, v) \longmapsto u \wedge v$ *est bilinéaire et alternée.*

Expression du produit vectoriel en fonction des coordonnées des vecteurs.

Soit (e_1, e_2, e_3) est une *b.o.n. directe*; on a, comme conséquence immédiate de la propriété 7.33 :

$$e_1 \wedge e_2 = e_3\,, \quad e_2 \wedge e_3 = e_1\,, \quad e_3 \wedge e_1 = e_2. \qquad (\star)$$

On déduit du corollaire que, si (e_1, e_2, e_3) est une ***b.o.n. directe*** et si $x = x_1e_1 + x_2e_2 + x_3e_3$, et $y = y_1e_1 + y_2e_2 + y_3e_3$, alors

$$\boxed{x \wedge y = (x_2y_3 - x_3y_2)\, e_1 + (x_3y_1 - x_1y_3)e_2 + (x_1y_2 - x_2y_1)e_3}$$

Propriétés 7.35 *1.* ***Identité de Lagrange :*** $\|x \wedge y\|^2 = \|x\|\,\|y\|^2 - \langle\, x\,,\, y\ \rangle^2$

2. ***Double produit vectoriel :*** $x \wedge (y \wedge z) = \langle\, x\,,\, z\,\rangle\, y - \langle\, x\,,\, y\rangle\, z$

3. ***Identité de Jacobi :*** $x \wedge (y \wedge z) + y \wedge (z \wedge x) + z \wedge (x \wedge y)$

Démonstration : Pour montrer l'identité de Lagrange, on utilise 7.33, 3. et on exprime $\sin^2\theta$ en fonction de $\cos^2\theta$.

La formule du double produit vectoriel est triviale si y et z sont liés. Si $y \wedge z \neq 0$, on prend une b.o.n. directe, (i, j, k) telle que i soit colinéaire à y et $j \in \mathrm{Vect}\{y, z\}$. On aura alors $x = x_1i + x_2j + x_3k$, $y = y_1i$ et $z = z_1i + z_2j$. La formule se vérifie alors facilement.

L'identité de Jacobi se montre facilement à l'aide de la formule du double produit vectoriel. □

Nota - L'identité de Jacobi montre que la loi de composition sur $\vec{E}_3$ définie par le produit vectoriel *n'est pas associative.*
Un espace vectoriel muni d'une loi de composition interne bilinéaire et alternée vérifiant l'identité de Jacobi est dit ***algèbre de Lie***[13]. On note habituellement $[\ ,\ ]$ la loi interne alternée d'une algèbre de Lie ; avec cette notation l'identité de Jacobi s'écrit :

$$[\,x, [y, z]\,] + [\,y, [z, x]\,] + [\,z, [x, y]\,] = 0$$

Le produit vectoriel définit donc sur $(\mathbb{R}^3, \langle\ ,\ \rangle_{can})$ une structure d'algèbre de Lie.

On a enfin la propriété suivante :

Propriété 7.36 *Soient* $x, y \in \vec{E}_3$ *et* (e_i) *étant une b.o.n directe.* $x \wedge y$ *est le seul vecteur de* $\vec{E}_3$ *tel que*

$$\langle\, x \wedge y\,,\, z\,\rangle = \text{dét}\, \|x, y, z\|_{e_i}, \qquad \text{pour tout } z \in \vec{E}_3.$$

[13]Il s'agit d'une notion particulièrement importante en Mathématiques

Démonstration : Si x et y sont liées, la propriété est triviale. S'il sont indépendants, prenons $\vec{n}$ de manière que $(x, y, \vec{n})$ soit une base directe. Soit $z = \lambda u + \mu v + \langle\, z, \vec{n}\,\rangle\, \vec{n}$, on a :

$$\text{dét}\, \|x, y, z\|_{e_i} = \text{dét}\, \|x, y, \langle\, z, \vec{n}\,\rangle\, \vec{n}\|_{e_i} = \langle\, z, \vec{n}\,\rangle\, \text{dét}\, \|x, y, \vec{n}\|_{e_i}$$

D'autre part,

$$\langle\, x \wedge y, z\,\rangle = \langle\, z, \vec{n}\,\rangle\, \langle\, x \wedge y, \vec{n}\,\rangle = \langle\, z, \vec{n}\,\rangle\, \text{dét}\, \|x, y, \vec{n}\|_{e_i}$$

d'après 4. de la Propriété 7.33, d'où l'identité.
D'autre part, s'il existe un autre vecteur u tel que pour tout $z \in \vec{E}_3$ on aurait aussi $\langle\, u, z\,\rangle = \text{dét}\, \|x, y, z\|_{e_i}$ et donc

$$\langle\, x \wedge y - u\,,\, z\,\rangle = 0 \qquad \text{pour tout } z \in \vec{E}_3$$

d'où $x \wedge y = u$, car le produit scalaire est une forme non dégénérée. $\square$

Remarques. –

1. La définition 7.32 donnée dans cette section, a l'avantage de mettre en évidence la signification géométrique du produit vectoriel : son rapport avec la notion d'aire et le fait qu'il permet, à partir de deux vecteurs indépendants, de construire une base directe.

2. La propriété (7.36) peut être prise comme définition du produit vectoriel, car elle est ***indépendante du choix de la b.o.n. directe***. En effet si on change de b.o.n. directe le déterminant dét $\|x, y, z\|_{e_i}$ est multiplié par le déterminant d'une matrice de $\mathrm{SO}(3, \mathbb{R})$ qui vaut 1. Cette définition a l'intérêt de mettre en évidence le fait que le produit vectoriel est une application multilinéaire alternée et de plus elle peut (dans un certain sens) être généralisée en dimension n. Cependant, cette généralisation n'a aucun intérêt dans le cadre d'un cours élémentaire d'algèbre linéaire : en revanche elle intervient assez naturellement dans le cadre de la théorie de l'*Algèbre extérieure* associée à un espace vectoriel, théorie qui a des nombreuses applications, mais qui dépasse le cadre de ce cours.

3. Une autre définition de produit vectoriel peut être donnée à partir des formules $(\star)$, page 250. On peut définir le produit vectoriel comme la seule application bilinéaire alternée $f : \vec{E}_3 \times \vec{E}_3 \longrightarrow \vec{E}_3$, qui, dans une b.o.n. directe vérifie [14] :

$$f(e_1, e_2) = e_3\,, \quad f(e_2, e_3) = e_1\,, \quad f(e_3, e_1) = e_2$$

   Ici aussi il faut vérifier l'indépendance du choix de la b.o.n. directe (ce qui cependant est moins facile que dans le cas de la définition par le déterminant).

Exercices 31. 32.

[14] cf. exercice 1

7.13 Diagonalisation des endomorphismes autoadjoints d'un espace euclidien

Nous terminons ce chapitre par une application importante du produit scalaire en montrant que *toute matrice symétrique* réelle *est diagonalisable dans* $\mathbb{R}$.

Définition 7.37 *Un endomorphisme f d'un espace euclidien est dit* ***autoadjoint*** *si*

$$\langle f(x),\, y\rangle \;=\; \langle x,\, f(y)\rangle\ , \quad \forall x,\, y \in E$$

En d'autres termes, f est autoadjoint si $f^* = f$.

La signification est claire. Soit $\{e_i\}$ une base orthonormée et $A = M(f)_{e_i}$; la condition de la définition s'exprime sous forme matricielle par

$$^t(AX)Y \;=\; {}^tX \cdot AY \quad , \quad \forall X, Y \in \mathcal{M}_{n,1}(\mathbb{R})$$

c'est-à-dire :

$$^tX\,{}^tAY \;=\; {}^tXAY \quad , \quad \forall X, Y \in \mathcal{M}_{n,1}(\mathbb{R})$$

ce qui est équivalent à ${}^tA = A$. Ainsi : f *est autoadjoint si et seulement si la matrice qui le représente* dans une base orthonormée *est symétrique.*

Théorème 7.38 *Soit f un endomorphisme autoadjoint d'un espace euclidien. Alors :*

1. *f est diagonalisable.*
2. *Les sous-espaces propres de f sont deux à deux orthogonaux.*
   En particulier, on peut construire une base orthonormée de vecteurs propres en choisissant une base orthonormée dans chaque espace propre.

NOTA - Soit $A \in \mathcal{M}_n(\mathbb{R})$ symétrique. On peut voir A comme la matrice d'un endomorphisme f de l'espace euclidien $(\mathbb{R}^n,\, \langle\, ,\, \rangle_{\text{can}})$ dans la base canonique $\{e_i\}$. Puisque la base canonique est orthonormée, f est autoadjoint. On peut donc énoncer le théorème en termes de matrices de la manière suivante :

Théorème 7.38' – *Toute matrice symétrique réelle est diagonalisable dans $\mathbb{R}$ et les espaces propres sont deux à deux orthogonaux (pour le produit scalaire canonique de $\mathbb{R}^n$).*

Par ailleurs, si $\{v_i\}$ est une base orthonormée de vecteurs propres, la matrice de passage $P = P_{(e_i)\to(v_i)}$ est orthonormée (cf. proposition 7.24) c'est-à-dire $P^{-1} ={}^t P$. On peut donc énoncer ce même résultat aussi de la manière suivante :

Théorème 7.38" – *Soit $A \in \mathcal{M}_n(\mathbb{R}^n)$ symétrique. Il existe alors $P \in \mathrm{O}(n, \mathbb{R})$ telle que la matrice $A' ={}^t P\, A\, P$ est diagonale.*

Notons que les théorèmes 7.38, 7.38' et 7.38" sont trois manières différentes d'énoncer le même théorème.

Démonstration du théorème :

(a) Montrons d'abord que le polynôme caractéristique de f est scindé, c'est-à-dire que toutes les valeurs propres de f sont réelles. Soit A la matrice qui représente f dans une base orthonormée et λ une valeur propre, réelle ou complexe, de A (λ existe car $P_A(X)$, considéré comme polynôme de $\mathbb{C}[X]$, admet, d'après le théorème de D'Alembert, au moins une racine)[15]. Montrons que $\lambda \in \mathbb{R}$. Soit $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathcal{M}_{n1}(\mathbb{C})$ la matrice d'un vecteur propre v correspondant à λ.
Puisque A est symétrique, on a :

$$^t(AX)\,\overline{X} = {}^tX\,A\,\overline{X} \tag{$*$}$$

D'autre part $AX = \lambda X$, d'où $\overline{A}\,\overline{X} = \overline{\lambda}\,\overline{X}$ et, comme A est réelle, $A\,\overline{X} = \overline{\lambda}\,\overline{X}$.
En reportant dans (*) :

$$^t(\lambda X)\,\overline{X} = {}^tX\,\overline{\lambda}\,\overline{X}$$

c'est-à-dire

$$\lambda\,(\,|x_1|^2 + \cdots + |x_n|^2\,) = \overline{\lambda}\,(\,|x_1|^2 + \cdots + |x_n|^2\,)$$

et donc $\lambda = \overline{\lambda}$, puisque $v \neq 0$. Les valeurs propres sont donc réelles.

(b) Montrons, par récurrence sur la dimension n de E, qu'il existe une base de vecteurs propres.
Pour $n = 1$, il n'y a rien à démontrer. Supposons la propriété vraie à l'ordre $n-1$. Soit λ une valeur propre de f, x un vecteur propre correspondant à λ et $H = \text{Vect}\{x\}^{\perp}$. Notons que : $\dim H = n-1$.

• Tout d'abord, H est stable par f, c'est-à-dire $f(H) \subset H$. En effet soit $y \in H$ (c'est-à-dire $y \perp x$) ; il s'agit de montrer que $f(y) \in H$, c'est-à-dire $f\,(y) \perp x$. On a :

$$\langle f(y),\, x\rangle \;=\; \langle y,\, f(x)\rangle \;=\; \lambda\,\langle y,\, x\rangle \;= 0, \text{ puisque } y \perp x;$$

donc $f(y) \perp x$. Ainsi la restriction $\tilde{f}$ de f à H est un endomorphisme de H.

• Par ailleurs $\tilde{f}$ est un endomorphisme autoadjoint de H (pour la restriction du produit scalaire à H), car si $v, w \in H$

$$\left\langle \tilde{f}(v),\, w\right\rangle_H \;=\; \langle f(v),\, w\rangle_E \;=\; \langle v,\, f(w)\rangle_E \;=\; \left\langle v,\, \tilde{f}(w)\right\rangle_H$$

Or $\dim H = n-1$, ainsi, d'après l'hypothèse de récurrence, il existe une base $\{e_2, \cdots, e_n\}$ formée de vecteurs propres de $\tilde{f}$. Il est clair que $\{x,\, e_2,\, \cdots,\, e_n\}$ est une base de E formée de vecteurs propres de f.

(c) Montrons que les sous-espaces propres sont deux à deux orthogonaux. Soient v_1 et v_2 vecteurs propres correspondant aux valeurs propres λ_1, λ_2, avec $\lambda_1 \neq \lambda_2$. Montrons que $v_1 \perp v_2$. On a :

[15] Ce passage aux complexes n'est peut-être pas élégant mais néanmoins correct. Si l'on restait dans le cadre des endomorphismes il faudrait utiliser le complexifié de l'espace vectoriel que nous n'avons pas introduit. Le lecteur qui connaît les espaces hermitiens, reconnaîtra dans la démonstration qui suit l'utilisation de l'expression matricielle du produit scalaire hermitien.

$$\langle f(v_1), \, v_2 \rangle = \lambda_1 \, \langle v_1, \, v_2 \rangle$$

or f est autoadjoint, donc :

$$\langle f(v_1), \, v_2 \rangle = \langle v_1, \, f(v_2) \rangle = \lambda_2 \, \langle v_1, \, v_2 \rangle \, .$$

Ainsi : $(\lambda_1 - \lambda_2) \langle v_1, \, v_2 \rangle = 0$ et puisque $\lambda_1 \neq \lambda_2$, on a : $\langle v_1, \, v_2 \rangle = 0$ □

REMARQUE. – *Les matrices symétriques* ***complexes*** *(non réelles) ne sont pas nécessairement diagonalisables (ni dans* $\mathbb{R}$*, ni dans* $\mathbb{C}$*).*

Par exemple, soit $A = \begin{pmatrix} 0 & \alpha \\ \alpha & \beta \end{pmatrix}$, $\alpha, \beta \in \mathbb{C}$. On a : $P_A(X) = X^2 - \beta X - \alpha^2$.
Si $\beta^2 + 4\,\alpha^2 = 0$ (dans $\mathbb{C}$ cela est possible sans que α et β soient tous deux nuls), on aura deux racines confondues, c'est-à-dire une valeur propre double $\lambda = \beta/2$ et donc $P_A(X) = (X - \beta/2)^2$. A serait alors diagonalisable si et seulement si $m_A(X) = X - \beta/2$ c'est-à-dire $A = (\beta/2)\, I$ ce qui évidemment est exclu.
Par exemple : $A = \begin{pmatrix} 0 & 1 \\ 1 & 2\,i \end{pmatrix}$ est symétrique et non diagonalisable.

Caractérisation du produit scalaire à l'aide de sa matrice.

Le théorème 7.38 a un corollaire important. Soit s une forme bilinéaire symétrique sur un espace vectoriel réel E (non nécessairement euclidien). Pour savoir si s définit un produit scalaire on peut effectuer la réduction de Gauss (cf. page 225). Mais on peut aussi utiliser le résultat suivant :

Proposition 7.39 – *Soit s une forme bilinéaire symétrique sur un espace vectoriel réel E de dimension finie, $\{e_i\}$ une base de E et $S = M(s)_{e_i}$. Alors s définit un produit scalaire si et seulement si la matrice S a toutes ses valeurs propres strictement positives.*

Démonstration :
Soit $\langle \ , \ \rangle_{e_i}$ le produit scalaire associé à cette base (cf. exemple 2. page 222), que nous noterons, pour simplifier $\langle \ , \ \rangle$.

(a) Montrons tout d'abord qu' il existe un et un seul endomorphisme f_s tel que :

$$\langle x, f_s(y) \rangle = s\,(x, y\,)\,, \qquad \forall x, y \in E.$$

Pour cela supposons que f_s existe. En notant $A = M(f_s)_{e_i}$, la relation de l'énoncé est équivalente à :

$${}^t X\, A\, Y = {}^t X\, S\, Y \ , \quad \forall X, Y \in \mathcal{M}_{n,1}(\mathbb{R}), \quad \text{c'est-à-dire :} \quad A = S,$$

ce qui montre que f_s est unique.
Réciproquement en définissant f_s par $M(f_s)_{e_i} = S$, on voit immédiatement que f_s vérifie l'identité de l'énoncé.

(b) D'autre part f_s est autoadjoint pour $\langle \ , \ \rangle$. En effet,

$$\langle x, f_s(y) \rangle = s(x,y) = s(y,x) = \langle y, f_s(x) \rangle = \langle f_s(x), y \rangle \ , \quad \forall x, y \in E.$$

Par conséquent, en prenant dans chaque espace propre une base orthonormée, on obtient une base orthonormée $\{v_i\}$ de vecteurs propres de f_s (notons qu'il s'agit de vecteurs propres de la matrice S).

Soient maintenant $x = \sum_{i=1}^n x_i\, v_i$ et $y = \sum_{j=1}^n y_j\, v_j$ deux vecteurs arbitraires de E. On a

$$\begin{aligned} s(x,y) &= \sum_{i,j=1}^n x_i\, y_j\, s(v_i, v_j) = \sum_{i,j=1}^n x_i\, y_j\, \langle v_i, f_s(v_j)\rangle \\ &= \sum_{i,j=1}^n x_i\, y_j\, \lambda_j\, \langle v_i, v_j\rangle = \sum_{i,j=1}^n x_i\, y_j\, \delta_{ij} = \lambda_1\, x_1 y_1 + \cdots \lambda_n\, x_n y_n \end{aligned}$$

d'où :

$$s(x,x) = \lambda_1\, {x_1}^2 + \cdots + \lambda_n\, {x_n}^2$$

ce qui montre que s est définie positive si et seulement si toutes les valeurs propres de S sont strictement positives. □

NOTA. – Nous verrons au chapitre 9 (cf. corollaire 9.25) que le résultat de ce théorème est, en fait, plus riche que celui que l'on vient d'énoncer.

Exemple – *Vérifier que la matrice* $A = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 2 & -1 \\ 0 & -1 & 3 \end{pmatrix}$ *définit un produit scalaire sur* $\mathbb{R}^3$.

On a : $P_A(\lambda) = -\lambda^3 + 6\,\lambda^2 - 9\,\lambda + 2$ et $\mathrm{Sp}_{\mathbb{R}}(A) = \{\,2,\, 2+\sqrt{3},\, 2-\sqrt{3}\,\}$.
Comme les valeurs propres sont strictement positives, A définit un produit scalaire.

Exercices 33. 34. 35. 36. 37.

EXERCICES

1 **Produit vectoriel dans $\mathbb{R}^3$**

1. Soit $\{e_1, e_2, e_3\}$ la base canonique de $\mathbb{R}^3$. Montrer qu'il existe une unique application $f : \mathbb{R}^3 \times \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ bilinéaire, alternée telle que
$$f(e_1, e_2) = e_3, \quad f(e_2, e_3) = e_1, \quad f(e_3, e_1) = e_2$$
$f(x, y)$ est dit ***produit vectoriel*** de x par y et est noté $x \wedge y$.
2. Calculer $x \wedge y$ en fonction des composantes de x et y dans la base canonique.
3. Montrer que :
$$\langle x \wedge y,\, z\rangle = \text{dét}\,\|x,\, y,\, z\|_{e_i}, \quad \forall x, y, z \in \mathbb{R}^3.$$
En déduire que $x \wedge y = 0$ si et seulement si x et y sont liés et que $x \wedge y$ est orthogonal à x et à y.

2 Comment peut-on définir l'angle entre deux plans de $\mathbb{R}^3$?
Exemple : calculer l'angle entre les plans $P\ :\ 2\,x+y-3z=0$ et $P'\ :\ x-3\,y+2\,z=0$.

3 Déterminer la matrice qui dans la base canonique de $\mathbb{R}^2$ représente la symétrie orthogonale par rapport à la droite d'angle polaire $\frac{\theta}{2}$.

4

1. Déterminer la matrice qui représente dans la base canonique de $\mathbb{R}^3$ la projection orthogonale sur le plan P d'équation $x + 2\,y - 3\,z = 0$.
2. En déduire la matrice qui représente la symétrie orthogonale par rapport à π.

5 Montrer que l'application : $\langle\,,\,\rangle \,:\, \mathbb{R}^3 \times \mathbb{R}^3 \longrightarrow \mathbb{R}$ définie par :

$$\langle x,\, y\rangle = \left(\frac{1}{\sqrt{3}}x_1 - x_2 + x_3\right)\left(\frac{1}{\sqrt{3}}y_1 - y_2 + y_3\right) + (x_2 - x_3)(y_2 - y_3) + 3x_3y_3$$

est un produit scalaire.

6 On définit sur $\mathbb{R}_2\,[x]$: $\langle P,\, Q\rangle := (a_0 + a_1)\,b_0 + (a_0 + 3\,a_1)b_1 + 3\,a_2\,b_2$ où $P = a_0 + a_1\,x + a_2\,x^2$, et $Q = b_0 + b_1\,x + b_2\,x^2$. Montrer qu'il s'agit d'un produit scalaire.

* **7** On considère l'intégrale :

$$I_n = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} x^n e^{-x^2/2}\,\mathrm{d}x \qquad (n \text{ entier, positif ou nul})$$

1. Montrer que cette intégrale converge et que $I_{2p+1} = 0$.
   Montrer la relation de récurrence $I_n = (n-1)I_{n-2}$ et calculer I_{2p}
   $\left(\text{rappel : } \int_{-\infty}^{+\infty} e^{-\frac{x^2}{2}}\,\mathrm{d}x = \sqrt{2\pi}\right)$.
2. On définit : $\langle\,,\,\rangle \,:\, \mathbb{R}\,[x] \times \mathbb{R}\,[x] \longrightarrow \mathbb{R}$ par :
   $$\langle P,\, Q\rangle := \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} e^{-x^2/2}P(x)Q(x)\,\mathrm{d}x$$
   Montrer qu'il s'agit d'un produit scalaire.

8 Pour quelles valeurs de $\lambda \in \mathbb{R}$ les formes bilinéaires sur $\mathbb{R}^3$ ci-dessous définissent-elles un produit scalaire ?

1. $b(x, y\,) = x_1\,y_1 + 6\,x_2\,y_2 + 3\,x_3\,y_3 + 2\,x_1\,y_2 + 2\,x_2\,y_1 + 3\,\lambda\,x_1\,y_3 + 3\,\lambda\,x_3\,y_1$
2. $b(x, y\,) = x_1\,y_1 + 10\,x_2\,y_2 + 6\,x_1\,y_2 + \lambda\,x_3\,y_3 - x_2\,y_3 - x_3\,y_2$
3. $b(x, y\,) = 2\,x_1\,y_1 + 7\,x_1\,y_2 + 7\,x_2\,y_1 + 8\,x_2\,y_2 - 3\,x_3\,y_3 + \lambda\,x_2\,y_3 + \lambda\,x_3\,y_2$

9 On munit $\mathbb{R}_2[x]$ du produit scalaire de l'exercice 7. Déterminer une base orthonormée.

10 **Distance associée à la norme** (cf. Appendice A.7, pages 384, 392)
Soit $\mathcal{E}$ un ensemble (non nécessairement un espace vectoriel).
Une application $\mathrm{d} \,:\, \mathcal{E} \times \mathcal{E} \longrightarrow \mathbb{R}^+$ est dite ***distance*** si elle vérifie les propriétés suivantes :

a) $\mathrm{d}(x, y\,) = 0 \Longleftrightarrow x = y$
b) $\mathrm{d}(x, y\,) = \mathrm{d}(y, x)$
c) $\mathrm{d}(x, z) \leq \mathrm{d}(x, y\,) + d(y, z)$

1. Montrer que si E est un espace euclidien, l'application $\mathrm{d} : E \times E \longrightarrow \mathbb{R}^+$ définie par
   $$\mathrm{d}(v, w) \;:= \|v - w\|$$
   définit une distance sur E dite *distance associée à la norme.*
2. Justifier cette définition en considérant $E = \mathbb{R}^3$ muni du produit scalaire canonique.

11 Déterminer l'angle (non orienté) entre les vecteurs e_1 et e_2 de la base canonique de $\mathbb{R}^3$, $\mathbb{R}^3$ étant muni du produit scalaire de l'exercice 5.

12 On munit $\mathbb{R}^4$ de la forme bilinéaire b définie, dans la base canonique, par :

$$\begin{aligned} b(x,y) := \; & x_1\,y_1 + 2\,x_2\,y_2 + 4\,x_3\,y_3 + 18\,x_4\,y_4 + x_1\,y_3 + x_3\,y_1 + 2\,x_2\,y_4 \\ & +2\,x_4\,y_2 + 6\,x_3\,y_4 + 6\,x_4\,y_3 \end{aligned}$$

1. Montrer qu'il s'agit d'un produit scalaire.
2. Écrire la matrice de b dans la base canonique de $\mathbb{R}^4$.

13 Soit $\mathcal{A} \in \mathcal{M}_n(K)$. Montrer que

$$\begin{array}{lll} a) & {}^tXAY = 0\,,\ \forall X, Y \in \mathcal{M}_{n,1}(K) & \Longleftrightarrow \quad A = 0 \\ b) & {}^tXAX = 0\,,\ ,\ \forall X \in \mathcal{M}_{n,1}(K) & \Longleftrightarrow \quad A \text{ est antisymétrique} \end{array}$$

14 Déterminer le sous-espace orthogonal au plan engendré par les vecteurs e_1 et e_2 de la base canonique de $\mathbb{R}^3$, pour le produit scalaire de l'exercice 5.

15 On considère l'espace euclidien $\mathbb{R}^4$ muni du produit scalaire de l'exercice 12. Soit F le sous-espace de $\mathbb{R}^4$ défini par :

$$\begin{cases} x_1 - x_2 + x_3 - \ x_4 & = 0 \\ \qquad x_2 \qquad - 2x_4 & = 0 \end{cases}$$

Déterminer $F^\perp$.

16 Soient E un espace euclidien et F un sous-espace de E. On appelle ***projecteur orthogonal*** sur F le projecteur sur F relativement à la somme directe $E = F \oplus F^\perp$. Soit p un projecteur ; montrer que les propriétés suivantes sont équivalentes :

1. p est un projecteur orthogonal.
2. $\mathrm{Im}(\mathrm{id} - p) = (\mathrm{Im}\, p)^\perp$.
3. $\mathrm{Im}(\mathrm{id} - p) \subset (\mathrm{Im}\, p)^\perp$.
4. $\langle p(x), p(y)\rangle = \langle x, p(y)\rangle$, $\quad \forall\, x, y \in E$.
5. $\langle p(x), y\rangle = \langle x, p(y)\rangle$, $\quad \forall\, x, y \in E$.
   (On dit que p est «auto-adjoint» : cf. définition 7.37)

* **17** Soit F un sous-espace vectoriel de E et $v \in E$; on appelle ***distance de*** v ***à*** F le scalaire

$$\mathrm{d}(v,\, F) := \operatorname*{Inf}_{v' \in F}\ \mathrm{d}(v,\, v')$$

d étant la distance associée à la norme (cf. exercice 10). Montrer que si p_F est le projecteur orthogonal sur F, alors $\mathrm{d}(v,\, F) = \|v - p_F(v)\|$. En d'autres termes :

$p_F(v)$ est le vecteur de F qui est à distance minimale de v.

18

1. Soit E un espace euclidien et p_F le projecteur orthogonal sur un sous-espace F. Montrer que si $\{e_1, \cdots, e_p\}$ est une base orthonormée de F
   $$p_F(v) = \ \langle v, e_1\rangle\ e_1 + \cdots + \langle v, e_p\rangle\ e_p$$
   Interpréter le procédé d'orthonormalisation de Schmidt.
2. On considère $\mathbb{R}[x]$ muni du produit scalaire de l'exercice 7. et $\mathbb{R}_3[x]$ et $\mathbb{R}_2[x]$ munis du produit scalaire induit. Calculer la distance du polynôme x^3 à $\mathbb{R}_2[x]$ (cf exercice 17.)

* **19** **Matrice de Gram**

1. Soit $(E, \langle\, , \,\rangle)$ un espace euclidien et $\{v_1, ..., v_p\}$ une famille ordonnée de vecteurs de E. On appelle *matrice de Gram* associée à la famille $(v_1, ..., v_p)$, la matrice

$$\mathrm{Gram}(v_1, ..., v_p) = \begin{pmatrix} \langle v_1, v_1\rangle & \cdots & \langle v_1, v_p\rangle \\ \cdots & \cdots & \cdots \\ \langle v_p, v_1\rangle & \cdots & \langle v_p, v_p\rangle \end{pmatrix}$$

On note $\mathrm{G}(v_1, ..., v_p) := \text{dét}\,\mathrm{Gram}(v_1, ..., v_p)$. Soit $\mathcal{B} = \{e_1, ..., e_p\}$ une base orthonormée de $\mathrm{Vect}\{v_1, ..., v_p\}$ et $A = \|v_1, \ldots, v_p\|_{\mathcal{B}}$.
Montrer que $\mathrm{Gram}(v_1, ..., v_p) = {}^tA\,A$.

2. En déduire que $\mathrm{G}(v_1, ..., v_p) \geq 0$ et que la famille $\{v_1, ..., v_p\}$ est libre si et seulement si $\mathrm{G}(v_1, ..., v_p) \neq 0$.

3. Soit $\{v_1, ..., v_p\}$ une famille libre, $F = \mathrm{Vect}\{v_1, ..., v_p\}$ et $x \in E$. Montrer que

$$\mathrm{d}(x, F)^2 = \frac{\mathrm{G}(x, v_1, ..., v_p)}{\mathrm{G}(v_1, ..., v_p)}$$

4. Retrouver par cette méthode le résultat de l'exercice 18.2. et comparer les deux méthodes.

* **20** Soit E un espace euclidien, F, G deux sous-espaces vectoriels de E. Montrer que :

$$\begin{array}{rcl} F \subset G & \Longrightarrow & F^\perp \supset G^\perp \\ (F+G)^\perp & = & F^\perp \cap G^\perp \\ (F\cap G)^\perp & = & F^\perp + G^\perp \end{array}$$

21 Soit f un endomorphisme d'un espace euclidien. Montrer que

$$(\mathrm{Ker}\, f)^\perp = \mathrm{Im}\, f^* \quad \text{et} \quad (\mathrm{Im}\, f)^\perp = \mathrm{Ker}\, f^*$$

22 Montrer que si $A \in \mathcal{M}_n(\mathbb{R})$ vérifie ${}^tA\,A = 0$, alors $A = 0$.

** **23** **Méthode des moindres carrés. Inverse généralisée d'une application linéaire injective** (cf. aussi Appendice A.5).

1. Soient $(E, \langle\, , \,\rangle_E)$ et $(E', \langle\, , \,\rangle_{E'})$ deux espaces euclidiens et $f : E \to E'$ une application linéaire. Montrer qu'il existe une et une seule application linéaire $f^* : E' \to E$ telle que

$$\langle f(x), y\rangle_E = \langle x, f^*(y)\rangle_{E'}\ , \qquad \forall\, x \in E,\, y \in E'.$$

2. On suppose dans la suite que $q = \dim E \leq \dim E' = n$ et que f *est injective.* Montrer que $\text{dét}\,(f^* \circ f) \neq 0$.

3. Soit $p : E' \to E'$ la projection orthogonale sur $\mathrm{Im}\, f$. Montrer que :

$$p = f \circ (f^* \circ f)^{-1} \circ f^*.$$

4. On note $f^\dagger : E' \to E$ l'application définie par

$$f^\dagger = (f^* \circ f)^{-1} \circ f^*;$$

$f^\dagger$ est dite *inverse généralisée* de f (noter que $f^\dagger \circ f = \mathrm{id}_E$ et $f \circ f^\dagger = p$, et que si f est bijective alors $f^\dagger = f^{-1}$).
Soit le système linéaire $f(x) = b$, avec f injective. On appelle *solution des moindres carrés* le vecteur $x_0 \in E$ tel que $\|(f(x_0) - b\| = \operatorname*{Inf}_{x\in E} \|f(x) - b\|$. Montrer que la solution des moindres carrés du système $f(x) = b$ est donnée par $x_0 = f^\dagger(b)$.

5. *Exemple.* a) Calculer l'inverse généralisée de la matrice $\begin{pmatrix} 2 & 1 \\ 1 & -1 \\ 1 & 3 \end{pmatrix}$.

   b) Pour quelles valeurs de $a \in \mathbb{R}$ le système

$$\begin{cases} 2\,x + \ y &= 1 \\ \ \ x - \ y &= 2 \\ \ \ x + 3\,y &= a \end{cases}$$

   admet-il une solution ? Calculer, dans le cas où le système est incompatible, la solution des moindres carrés.

* **24** Soit $A \in \mathrm{O}\,(n, \mathbb{R})$. Montrer que si $\det A = 1$ chaque terme de A est égal à son cofacteur et si $\det A = -1$ chaque terme de A est égal à l'opposé de son cofacteur.

* **25** **Décomposition de Householder** (cf. Appendice A.4).

A l'aide du procédé d'orthonormalisation de Schmidt, montrer que toute matrice $A \in \mathrm{GL}(n, \mathbb{R})$ peut s'écrire sous la forme

$$A = Q\,R$$

où $Q \in \mathrm{O}\,(n, \mathbb{R})$ et R est triangulaire supérieure (décomposition de Householder).

Exemple. Donner la décomposition de Householder de la matrice

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

26 Soit f une transformation orthogonale d'un espace euclidien. Montrer que

$$\mathrm{Ker}(f - \mathrm{id}) = \mathrm{Im}(f - \mathrm{id})^{\perp}$$

En déduire que si $(f - \mathrm{id})^2 = 0$, alors $f = \mathrm{id}$.

27 Préciser la nature des endomorphismes de $\mathbb{R}^3$ qui dans la base canonique sont représentés par les matrices

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

*** **28** **Expression canonique des transformations orthogonales** (cf. aussi exercice 16 du chapitre 8)

Soit f une transformation orthogonale d'un espace euclidien. On posera dans la suite $g = f + f^*$.

1. Montrer que g est autoadjoint (donc E est somme directe des espaces propres de g). On notera $\lambda_1, ..., \lambda_m$ les valeurs propres de g et $V_{\lambda_1}, ..., V_{\lambda_m}$ les sous-espaces propres correspondants.
2. Montrer que les V_{λ_i} sont stables par f.
3. Soit $f_{\lambda_i} := f\,|_{V_{\lambda_i}}$. Montrer que le polynôme $Q(X) = X^2 - \lambda_i X + 1$ est annulateur de f_{λ_i}.
4. On suppose $V_2 \neq \{0\}$. Montrer que $f_2 = \mathrm{id}$.
5. On suppose $V_{-2} \neq \{0\}$. Montrer que $f_{-2} = -\,\mathrm{id}$
6. Soit $\lambda_i \neq \pm 2$ tel que $V_{\lambda_i} \neq \{0\}$. Montrer que si $v \in V_{\lambda_i}$ alors v n'est pas vecteur propre de f. En déduire que l'espace $W = \mathrm{Vect}\{v, f(v)\}$ est de dimension 2.
7. Montrer que W est stable par f.

8. Soit $\tilde{f} := f\,|_W$. Montrer que $\tilde{f}$ est une rotation.
9. Soit $W^{\perp}$ l'orthogonal de W dans V_{λ_i}. Montrer que $W^{\perp}$ est stable par f.
10. Déduire des questions précédentes les résultats suivants :
Soit f une transformation orthogonale d'un espace euclidien. Il existe alors une base $\{e_i\}$ telle que :

$$M(f)_{e_i} = \begin{pmatrix} \boxed{\begin{matrix} 1 & & \\ & \ddots & \\ & & 1 \end{matrix}} & & & & \\ & \boxed{\begin{matrix} -1 & & \\ & \ddots & \\ & & -1 \end{matrix}} & & & \\ & & \boxed{\begin{matrix} \cos\theta_1 & -\sin\theta_1 \\ \sin\theta_1 & \cos\theta_1 \end{matrix}} & & \\ & & & \ddots & \\ & & & & \boxed{\begin{matrix} \cos\theta_r & -\sin\theta_r \\ \sin\theta_r & \cos\theta_r \end{matrix}} \end{pmatrix}$$

*** $\boxed{29}$ **Décomposition des transformations orthogonales en produit de réflexions**

Soit E un espace euclidien de dimension n. f étant un endomorphisme de E, on notera E_λ^f l'espace propre de f correspondant à la valeur propre λ. On appelle ***réflexion*** une transformation orthogonale σ telle que $\dim E_1^\sigma = n-1$. On se propose de démontrer la propriété suivante :

$(\mathcal{P})$ *Soit E un espace euclidien de dimension n. Toute $f \in \mathrm{O}(E)$ s'écrit sous la forme :*

$$f = \sigma_{a_1} \circ \cdots \circ \sigma_{a_r}\,, \quad \text{avec} \;:\quad n - \dim E_1^f \;\leq\; r \;\leq\; n. \qquad .$$

1. Montrer que si $f \in \mathrm{O}(E)$ est le produit de r réflexions, $f = \sigma_{a_1} \circ \cdots \circ \sigma_{a_r}$, alors : $r \geq n - \dim E_1^f$.
2. (a) Montrer que si σ est une réflexion, $E_1^{\sigma\,\perp}$ est espace propre de σ pour la valeur propre -1. En déduire que σ est diagonalisable et que $\sigma^2 = \mathrm{id}$.
   (b) Montrer qu'un endomorphisme σ est une réflexion si et seulement si il existe $a \in E$, $a \neq 0$, tel que $\sigma(x) = x - 2\,\dfrac{\langle x, a\rangle}{\|\,a\,\|^2}\,a$. On notera par la suite σ_a la réflexion associée au vecteur a.
   (c) Montrer que pour tout couple de vecteurs non nuls $y, z \in E$ tels que $y \neq z$ et $\|\,y\,\| = \|\,z\,\|$, il existe une réflexion σ_a qui échange y et z, c'est-à-dire telle que $\sigma_a(y) = z$.
3. (a) Soit $(\,E, \langle\,,\,\rangle\,)$ un espace euclidien de dimension $n \geq 2$ et soit $f \in \mathrm{O}\,(E)$. On suppose que $E_1^f \neq 0$. Soit $z \neq 0$ tel que $f(z) = z$ et $H = z^{\perp}$. Montrer que H est stable par f.
   (b) Soit $b \in H$, $b \neq 0$, et $\tilde{\sigma}_b$ la réflexion de l'espace euclidien H (muni du produit scalaire induit) définie par b (cf. 2.(b)). Notons σ_b la réflexion de E définie par le même vecteur b. Montrer que $\sigma_b\,|_H = \tilde{\sigma}_b$. En déduire par récurrence la propriété $(\mathcal{P})$ dans le cas où $E_1^f \neq 0$,
   (c) On suppose que $E_1^f = \{0\}$. Soit $z \neq 0$ et $a = z - f(z)$. Montrer que $\sigma_a \circ f\,(z) = z$. En déduire la propriété $(\mathcal{P})$ dans le cas où $E_1^f = \{0\}$.
4. Exemple. Soit $A = \dfrac{1}{3}\begin{pmatrix} 2 & 2 & -1 \\ -1 & 2 & 2 \\ 2 & -1 & 2 \end{pmatrix}$. Montrer que $A \in \mathrm{O}\,(3, \mathbb{R})$. Décomposer A en produit de réflexions.

(*cf. aussi la démonstration de ce résultat dans l'Appendice A.10, page 410*)

* **30** On appelle *similitude* un endomorphisme d'un espace euclidien E qui vérifie

$$\|f(x)\| = k\,\|x\| \;\text{ pour tout } x \in E, \text{ avec } k > 0.$$

1. Montrer que les similitudes conservent les angles.
2. Soit f un endomorphisme injectif d'un espace euclidien E. On dit que f *conserve l'orthogonalité* s'il transforme tout couple de vecteurs orthogonaux en un couple de vecteurs orthogonaux. Soit (e_i) une b.o.n. Montrer que si f conserve l'orthogonalité, alors $f(e_i)+f(e_j) \perp f(e_i)-f(e_j)$ pour tous $i \neq j$. En déduire que tous les vecteurs $f(e_i)$, $i = 1, \ldots, n$, ont même norme.
3. Montrer que si f est injective et conserve l'orthogonalité, alors elle est une *similitude*. En particulier, tout endomorphisme injectif qui conserve les angles est une similitude.

* **31** **Matrices des rotations dans la base canonique de $\mathbb{R}^3$.**

1. Soit $\vec{n}$ un vecteur unitaire de $\mathbb{R}^3$, $x \in \mathbb{R}^3$ et x' la projection orthogonale de x sur $\vec{n}^{\perp}$. Montrer que si f est la rotation d'angle θ autour de l'axe dirigé et orienté par $\vec{n}$, on a :

$$f(x') = (\cos\theta)\, x' + (\sin\theta)\, \vec{n} \wedge x$$

   En déduire que :

$$f(x) = (\cos\theta)\, x + (\sin\theta)\, \vec{n} \wedge x + (1 - \cos\theta)\ \langle x, \vec{n}\rangle\ \vec{n}$$

2. A l'aide de cette relation, donner la matrice de f dans la base canonique de $\mathbb{R}^3$, en fonction de θ et des composantes de $\vec{n}$.

* **32** Montrer que si $f \in \mathrm{End}(E)$, $f \neq 0$, alors

$$f(x \wedge y) = f(x) \wedge f(y),\ \forall x, y \in E, \quad \text{si et seulement si } f \in \mathrm{SO}(E).$$

33 Déterminer une base orthonormée formée de vecteurs propres de la matrice

$$A = \begin{pmatrix} 5 & -1 & 2 \\ -1 & 5 & 2 \\ 2 & 2 & 2 \end{pmatrix}$$

Interpréter géométriquement l'endomorphisme qui dans la base canonique de $\mathbb{R}^3$ est représenté par A.

*

34 Soit $A = (\, a_{ij}\,)$ une matrice symétrique réelle. Montrer que ses valeurs propres λ_i vérifient

$$\sum_{i=1}^{n} \lambda_i^2 = \sum_{i,j=1}^{n} a_{ij}^2$$

*

35 Soit $\mathcal{B} = \{v_1, \cdots, v_n\}$ une base orthogonale d'un espace euclidien E et $f \in \mathrm{End}\,(E)$. On dit que f *conserve l'orthogonalité de* $\mathcal{B}$ si la famille $\{f(v_1), \cdots, f(v_n)\}$ est orthogonale. Montrer que pour tout $f \in \mathrm{End}\,(E)$ il existe une base orthogonale dont f conserve l'orthogonalité.

*

36 Montrer que l'endomorphisme

$$\psi_M \;: \begin{array}{ccc} \mathcal{M}_n(\mathbb{R}) & \longrightarrow & \mathcal{M}_n(\mathbb{R}) \\ A & \longmapsto & {}^tM\,A\,M + M\,A\,{}^tM \end{array}$$

avec $M \in \mathcal{M}_n(\mathbb{R})$ fixée, est diagonalisable.

**

37 **Décomposition polaire**

1. Soit ρ un endomorphisme autoadjoint d'un espace euclidien E. On dit que ρ est *défini positif* si toutes ses valeurs propres sont strictement positives.
   Montrer que si ρ est défini positif, il existe un et un seul endomorphisme σ autoadjoint défini positif tel que $\rho = \sigma^2$ (σ est dit *racine carrée positive* de ρ).
2. Soit f un endomorphisme bijectif de E.
   Montrer que l'endomorphisme $f^* \circ f$ est autoadjoint et défini positif.
3. Soit σ la racine carrée positive de $f^* \circ f$.
   Montrer que $u := f \circ \sigma^{-1}$ est une transformation orthogonale. En déduire le résultat suivant :

   Tout endomorphisme inversible f de E s'écrit d'une manière unique *sous la forme :*

   $f = u \circ \sigma$ avec u orthogonal et σ autoadjoint défini positif

   (décomposition polaire de f).

INDICATIONS

1

1. Puisque f est bilinéaire, elle est déterminée par la connaissance des $f(e_i,\, e_j)$. Comme f est alternée, il suffit de connaître $f(e_i, e_j)$ pour $i < j$.
2. $x \wedge y = (x_2\, y_3 - x_3\, y_2)\, e_1 + (x_3\, y_1 - x_1\, y_3)\, e_2 + (x_1\, y_2 - x_2\, y_1)\, e_3$

3. Développer le déterminant selon les éléments de la troisième ligne.

2

On peut définir l'angle entre les deux plans comme étant l'angle entre les deux normales n et n'. Si les deux plans ne sont pas orientés, les normales ne le sont pas non plus ; cela définit deux angles θ_1 et θ_2 (cf figure 14). On peut prendre (par définition) comme angle entre les deux plans celui qui est compris entre 0 et $\pi/2$.

Dans l'exemple on trouve $\theta_1 = \pi/3$, $\theta_2 = \dfrac{2\pi}{3}$. L'angle entre P et P' est donc $\pi/3$.

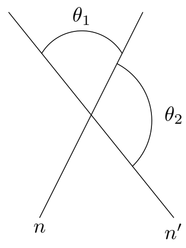


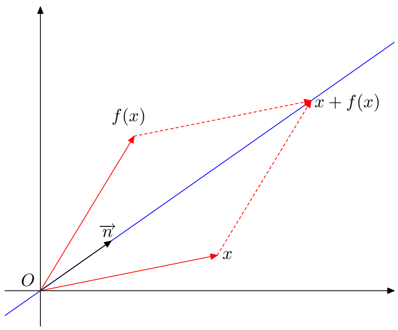
Figure 14

3

Soit $\vec{n} = \cos\dfrac{\theta}{2}\, e_1 + \sin\dfrac{\theta}{2}\, e_2$ un vecteur unitaire porté par la droite. On a : $f(x) + x = 2(x \cdot \vec{n})\vec{n} \quad \forall\, x \in \mathbb{R}^2$.

Calculer $f(e_1)$ et $f(e_2)$. On trouve :

$$M(f)_{e_i} = \begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix}.$$



4

1. – Soit $\vec{n}$ un vecteur unitaire normal à P. On a :

$$f(x) + \langle x, \vec{n}\rangle\, \vec{n} = x$$

$$\text{d'où :} \quad f(x) = x - \langle x, \vec{n}\rangle\, \vec{n}$$

Si $\vec{n} = \dfrac{1}{\sqrt{14}}\begin{pmatrix} 1 \\ 2 \\ -3 \end{pmatrix}$, en calculant $f(e_1), f(e_2), f(e_3)$ on trouve :

$$M(f)_{e_i} = \frac{1}{14}\begin{pmatrix} 13 & -2 & 3 \\ -2 & 10 & 6 \\ 3 & 6 & 5 \end{pmatrix}$$

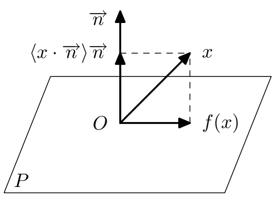


Figure 15

2. – Si g est la symétrie : $g(x) + x = 2\, f(x)$. On trouve :

$$M(g)_{e_i} = \frac{1}{7}\begin{pmatrix} 6 & -2 & 3 \\ -2 & 3 & 6 \\ 3 & 6 & -2 \end{pmatrix}$$

5 Vérifier les axiomes de la définition 7.1.

6 Vérifier les axiomes de la définition 7.1.
Pour iii) : $\langle P, P\rangle = a_0^2 + 2\, a_0\, a_1 + 3\, a_1^2 + 3\, a_2^2 = (a_0 + a_1)^2 + 2\, a_1^2 + 3\, a_2^2$

7

1. Montrer la convergence absolue des deux intégrales $\int_{-\infty}^{a} e^{-x^2/2}\, x^n \,\mathrm{d}x$ et $\int_{a}^{+\infty} e^{-x^2/2}\, x^n \,\mathrm{d}x$. I_{2p+1} est nul puisque c'est l'intégrale d'une fonction impaire sur $]-\infty, +\infty[$. En intégrant par parties, on trouve :
$$\int_{-X}^{X} e^{-x^2/2} x^n \mathrm{d}x = \left[-x^{n-1} e^{-x^2/2}\right]_{-X}^{X} + (n-1)\int_{-X}^{X} x^{n-2} e^{-x^2/2} \mathrm{d}x$$
d'où, en passant à la limite pour $X \to +\infty$: $I_n = (n-1) I_{n-2}$. On en déduit immédiatement, par récurrence :
$$I_{2p} = 1 \cdot 3 \cdot 5 \cdots (2p-1).$$
2. Tout d'abord l'intégrale converge, car elle est une somme finie d'intégrales convergentes. La bilinéarité et la symétrie sont immédiates. D'autre part, clairement $e^{-x^2/2} P(x)^2 \geq 0$, et donc $\langle P, P\rangle \geq 0$. Si $\langle P, P\rangle = 0$, puisque $e^{-x^2/2} P(x)^2 \geq 0$ et cette fonction est continue, on a nécessairement $e^{-x^2/2} P(x) = 0$ sur $]-\infty, +\infty[$, donc $P(x) = 0$.

8

1. $f(x,x) = (x_1 + 2x_2 + 3\lambda x_3)^2 + 2(x_2 - 3\lambda x_3)^2 + (3 - 27\lambda^2) x_3^2$. Il faut et il suffit que $-\dfrac{1}{3} < \lambda < \dfrac{1}{3}$.
2. Pour aucune valeur de λ (f n'est pas symétrique)
3. Pour aucune valeur de λ (le coefficient de x_3^2 dans $f(x,x)$ est négatif).

9 Appliquer la méthode de Schmidt à la base $\{1, x, x^2\}$. La base orthogonale obtenue de $\{1, x, x^2\}$ par le procédé de Schmidt est :

$\varepsilon_1 = 1$;

$\varepsilon_2 = x + \lambda 1$, avec λ tel que $\langle x + \lambda, 1\rangle = 0$, c'est-à-dire $I_1 + \lambda I_0 = 0$. Or $I_0 = 1$ et $I_1 = 0$, donc $\lambda = 0$, c'est-à-dire $\varepsilon_2 = x$;

$\varepsilon_3 = x^2 + \lambda x + \mu 1$. En imposant $\langle \varepsilon_3, \varepsilon_1\rangle = 0$ on trouve : $I_2 + \lambda I_1 + \mu I_0 = 0$. On a $I_2 = 1$, d'où : $\mu = -1$. En imposant $\langle \varepsilon_3, \varepsilon_2\rangle = 0$ on a $I_3 + \lambda I_2 + \mu I_1 = 0$. Comme $I_3 = 0$, on trouve $\lambda = 0$.

Donc $\varepsilon_1 = 1$, $\varepsilon_2 = x$, $\varepsilon_3 = x^2 - 1$. En normalisant on trouve :
$$e_1 = 1\,, \quad e_2 = x\,, \quad e_3 = \frac{1}{\sqrt{2}}\,(x^2 - 1).$$

10

1. – Simples vérifications.
2. – Il s'agit de la distance «naturelle» entre les points P et Q, extrémités de v et w (cf. figure 16)

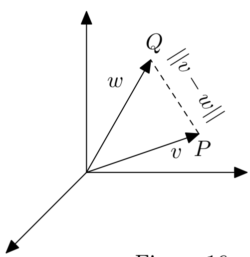


Figure 16

11 $\theta = \dfrac{2\pi}{3}$ (e_1 et e_2 ne sont pas orthogonaux pour ce produit scalaire).

12

1. $b(x,x) = (x_1 + x_3)^2 + 2\,(x_2 + x_4)^2 + 3\,(x_3 + 2\,x_4)^2 + 4\,x_4^2$

2. $A = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 2 & 0 & 2 \\ 1 & 0 & 4 & 6 \\ 0 & 2 & 6 & 18 \end{pmatrix}$

13

1. On a évidemment $AX = 0 \ \forall X \in \mathcal{M}_{n,1}(K) \iff A = 0$, (car, en termes d'applications, cela signifie : $f(x) = 0$, $\forall x \in E \iff f = 0$). Donc :

   ${}^tXAY = 0, \ \forall X, Y \in \mathcal{M}_{n,1}(K) \iff {}^tXA = 0, \ \forall X \in \mathcal{M}_{n,1}(K) \iff$

   (en prenant la transposée) ${}^tAX = 0, \ \forall X \in \mathcal{M}_{n,1}(K) \iff {}^tA = 0 \iff A = 0.$

2. ${}^tXAX = 0, \ \forall X \in \mathcal{M}_{n,1}(K) \iff$ en remplaçant X par $X + Y$: ${}^t(X+Y)A(X+Y) = 0, \ \forall X, Y \in \mathcal{M}_{n,1}(K)$; etc.

14 Soit $\vec{n} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$ une base de l'orthogonal ; imposer $\langle \vec{n}, e_1 \rangle = 0$ et $\langle \vec{n}, e_2 \rangle = 0$. On trouve $\vec{n} = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$.

15 $F^\perp = \{x \in \mathbb{R}^4 \mid b(x, y) = 0, \forall y \in F\}$, c'est-à-dire, sous forme matricielle :
$F^\perp = \{X \in \mathcal{M}_{n,1}(\mathbb{R}) \mid {}^tXAY = 0, \ \forall Y \in F\}$

Soit $Y = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix}$ $\qquad Y \in F \iff \begin{cases} y_1 - y_2 + y_3 - y_4 &= 0 \\ \quad y_2 \quad - 2\,y_4 &= 0 \end{cases}$

$\iff Y = \begin{pmatrix} 2\mu - \lambda \\ 2\mu \\ \lambda \\ \mu \end{pmatrix} \qquad Y = \lambda \begin{pmatrix} -1 \\ 0 \\ 1 \\ 0 \end{pmatrix} + \mu \begin{pmatrix} 2 \\ 2 \\ 0 \\ 1 \end{pmatrix} \qquad (\lambda, \mu \in \mathbb{R})$

${}^tXAY = 0$, $\forall Y \in F$ équivaut à :

$$(\,x_1\, x_2\, x_3\, x_4\,) \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 2 & 0 & 2 \\ 1 & 0 & 4 & 6 \\ 0 & 2 & 6 & 18 \end{pmatrix} \begin{pmatrix} 2\mu - \lambda \\ 2\mu \\ \lambda \\ \mu \end{pmatrix} = 0, \quad \forall \lambda, \, \mu \in \mathbb{R},$$

c'est-à-dire :

$$\lambda\,(3\,x_3 + 6\,x_4) + \mu\,(3\,x_1 + 6\,x_2 + 9\,x_3 + 22\,x_4) = 0, \quad \forall\, \lambda, \, \mu \in \mathbb{R}$$

D'où :

$$F^\perp \quad : \quad \begin{cases} x_3 \, + \, 2\,x_4 &= 0 \\ 3\,x_1 + 6\,x_2 + 9\,x_3 + 22\,x_4 &= 0 \end{cases}$$

16 1. $\Leftrightarrow$ 2. Si p est un projecteur alors $E = \operatorname{Im} p \oplus \operatorname{Im}(\operatorname{id} - p)$ (cf. exercice 9 chapitre 3). Donc p est orthogonal si et seulement si $\operatorname{Im}(\operatorname{id} - p) = (\operatorname{Im} p)^\perp$.

2. $\Leftrightarrow$ 3. Il est clair que 2. $\Rightarrow$ 3. Réciproquement, si $\mathrm{Im}(\mathrm{id}-p) \subset (\mathrm{Im}\, p)^{\perp}$, on a $\mathrm{Im}(\mathrm{id}-p) = (\mathrm{Im}\, p)^{\perp}$, puisque $\dim \mathrm{Im}(\mathrm{id}-p) = n - \dim \mathrm{Im}\, p = \dim(\mathrm{Im}\, p)^{\perp}$.

3. $\Leftrightarrow$ 4. On a : $\mathrm{Im}(\mathrm{id}-p) \subset (\mathrm{Im}\, p)^{\perp}$ si et seulement si $\langle x - p(x), p(y)\rangle = 0 \ \ \forall x, y \in E$, c'est-à dire : $\langle p(x), p(y)\rangle = \langle x, p(y)\rangle$, $\forall x, y \in E$.

4. $\Leftrightarrow$ 5. Si 4. est vérifiée, en échangeant le rôle de x et y, on voit que cela implique que $\langle p(x), y\rangle = \langle x, p(y)\rangle$, $\forall x, y \in E$. Réciproquement si cette identité est satisfaite, en remplaçant y par $p(y)$, on a : $\langle p(x), p(y)\rangle = \langle x, p(y)\rangle$, $\forall x, y \in E$.

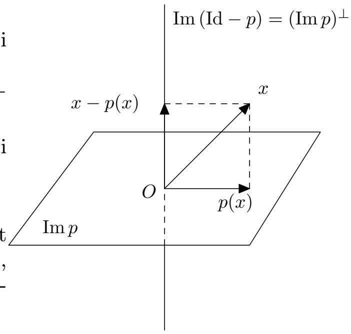


Figure 17

17

$\mathrm{d}(v - p_f(v))^2 = \|v - p_F(v)\|^2 = \|v\|^2 + \|p_F(v)\|^2 - 2\left\langle v, p_F(v)\right\rangle$. Or $v - p_F(v) \perp p_F(v)$, donc $\langle v - p_F(v), p_F(v)\rangle = 0$, c'est-à-dire $\langle v, p_F(v)\rangle = \|p_f(v)\|^2$.
Aussi $\mathrm{d}(v - p_f(v))^2 = \|v\|^2 - \|p_F(v)\|^{\cdot}$;

D'autre part, pour tout $v' \in F$, $\mathrm{d}'v, v')^2 = \|v - v'\|^2 = \|v\|^2 + \|v'\|^2 - 2\langle v, v'\rangle = \|v\|^2 + \|v'\|^2 - 2\left\langle p_F(v), v'\right\rangle$ d'où :

$$\begin{aligned} \mathrm{d}(v, v')^2 - d(v, p_F(v))^2 &= \|v'\|^2 + \|p_F(v)\|^2 - \left\langle p_F(v), v'\right\rangle \\ &= \|v' - p_F(v)\|^2 = d(\, v', p_F(v)\,)^2 \end{aligned}$$

c'est-à-dire :

$$(v, v')^2 = \mathrm{d}(v, p_F(v))^2 + \mathrm{d}(v', p_F(v)\,)^2$$

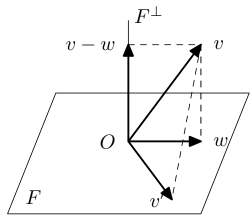


Figure 18

(théorème de Pythagore). Donc l' $\operatorname*{Inf}_{v' \in F}\ \mathrm{d}(v,\ v')$ s'obtient pour $v' = p_F(v)$.

18

1. $p_F(v) \in F$, donc $p_F(v) = a_1\, e_1 + \cdots + a_p\, e_p$. Imposer $v - p_F v \in F^{\perp}$ (cf. exercice 16).
   Soit $\{e_1, \cdots, e_n\}$ la base orthonormée associée à $\{v_1, \cdots, v_n\}$ par le procédé de Schmidt. On la construit en normant les vecteurs $\varepsilon_k = v_k - p_{F_{k-1}}(v_p)$, où $F_{k-1} = \mathrm{Vect}\{v_1, \cdots, v_{k-1}\}$, c'est-à-dire en normant les vecteurs
   $$\varepsilon_k = p_{\mathrm{Vect}\{v_1, \cdots, v_{k-1}\}^{\perp}}(v_k)$$
2. $\mathrm{d}(x^3, \mathbb{R}_2[x]) = \|x^3 - p_{\mathbb{R}_2[x]}(x^3)\|$
   $p_{\mathbb{R}_2[x]}(x^3) = \left\langle x^3, e_1\right\rangle e_1 + \left\langle x^3, e_2\right\rangle e_2 + \left\langle x^3, e_3\right\rangle e_3$ où $\{e_1,\, e_2,\, e_3\}$ est une base orthonormée de $\mathbb{R}_2[x]$, (par exemple la base obtenue de $\{1,\, x,\, x^2\}$ en appliquant le procédé de Schmidt). On trouve :
   $$\begin{aligned} p(x^3) &= I_3 + I_4\, x \ + \left\langle x^3, \frac{x^2-1}{\sqrt{2}}\right\rangle \frac{x^2-1}{\sqrt{2}} = I_3 + I_4\, x + \frac{1}{2}(I_5 - I_3)\,(x^2-1) \\ &= I_4\, x = 3\, x \end{aligned}.$$
   On a donc :
   $$\|p(x^3) - x^3\| = \|3\, x - x^3\| = \sqrt{9\, I_2 - 6\, I_4 + I_6} = \sqrt{6}.$$

19

1. Il suffit d'effectuer les calculs.
2. $\mathrm{G}(v_1, ..., v_p) = (\mathrm{dét}\, A)^2 \geq 0$ et par conséquent $\mathrm{G}(v_1, ..., v_p) = 0$ si et seulement si les vecteurs colonnes de A sont liés, c'est-à-dire $\{v_1, ..., v_p\}$ est une famille liée.
3. Notons x_F la projection orthogonale de x sur F et $x^{\perp} = x - x_F$ la projection orthogonale sur $F^{\perp}$. On a :

$$G(x, e_1, ..., e_p)$$

$$= \begin{vmatrix} \langle x, x\rangle & \langle x, v_1\rangle & \cdots & \langle x, v_p\rangle \\ \langle v_1, x\rangle & \langle v_1, v_1\rangle & \cdots & \langle v_1, v_p\rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle v_p, x\rangle & \langle v_p, v_1\rangle & \cdots & \langle v_p, v_p\rangle \end{vmatrix}$$

$$= \begin{vmatrix} \|x^{\perp}\|^2 & \langle x_F, v_1\rangle & \cdots & \langle x_F, v_p\rangle \\ 0 & \langle v_1, v_1\rangle & \cdots & \langle v_1, v_p\rangle \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \langle v_p, v_1\rangle & \cdots & \langle v_p, v_p\rangle \end{vmatrix} + \begin{vmatrix} \|x_F\|^2 & \langle x_F, v_1\rangle & \cdots & \langle x_F, v_p\rangle \\ \langle v_1, x_F\rangle & \langle v_1, v_1\rangle & \cdots & \langle v_1, v_p\rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle v_p, x_F\rangle & \langle v_p, v_1\rangle & \cdots & \langle v_p, v_p\rangle \end{vmatrix}$$

$$= \mathrm{d}(x, F)^2 \cdot G(e_1, ..., e_p) + G(x_F, e_1, ..., e_p) = \mathrm{d}(x, F)^2 \cdot G(e_1, ..., e_p)$$

puisque la famille $\{x_F, v_1, ..., v_p\}$ est liée.

NOTA. *Cette méthode permet d'éviter le calcul de la base orthonormée (cf. exercice 18).*

20 Se servir de l'exercice 36 du chapitre 3 et de l'isomorphisme j (cf. proposition 7.15).

21 On a : $\langle f(x), y\rangle = \langle x, f^*(y)\rangle$, pour tous $x, y \in E$. Si donc $x \in \operatorname{Ker} f$, alors $\langle x, f^*(y)\rangle = 0$, pour tout $y \in E$, c'est-à-dire $y \in (\operatorname{Im} f^*)^{\perp}$. Donc $\operatorname{Ker} f \subset (\operatorname{Im} f^*)^{\perp}$ et, en prenant l'orthogonal : $\operatorname{Im} f^* \subset (\operatorname{Ker} f)^{\perp}$. Vérifier ensuite que $\dim \operatorname{Im} f^* = \dim(\operatorname{Ker} f)^{\perp}$.
Pour la seconde égalité, remplacer f par f^* et prendre l'orthogonal des deux membres.

22 Si f est l'endomorphisme associé à A dans la base canonique, la condition $f^* \circ f = 0$ implique

$$0 = \langle f^* \circ f\,(x), x\rangle = \langle\, f(x)\,,\, f(x)\,\rangle \ ; \text{ donc : } \|\,f(x)\|^2 = 0, \ \ \forall x \in E\,.$$

23

1. Raisonner comme dans la démonstration de la proposition 7.16. Si $A = M(f)_{e_i, e'_i}$, où $\{e_i\}$ et $\{e'_i\}$ sont des bases orthonormées de E et de E', on a : $M(f^*)_{e'_i, e_i} = {}^tA$.
2. Soit x tel que $f^* \circ f(x) = 0$. On a :
$$0 = \langle f^* \circ f(x), x\rangle_E = \langle f(x), f(x)\rangle_{E'} \text{ , d'où } x = 0.$$
3. La projection orthogonale de $z \in E'$ sur $\operatorname{Im} f$ est caractérisée par
$$\begin{cases} p(z) & \in \operatorname{Im} f \\ z - p(z) & \in (\operatorname{Im} f)^{\perp} \end{cases} .$$
Or : $p(z) \in \operatorname{Im} f$ équivaut à $p(z) = f(u)$, avec $u \in E$.
$z - p(z) \perp \operatorname{Im} f$ équivaut à : $\langle z - f(u), f(x)\rangle = 0\,, \ \ \forall x \in E$
c'est-à-dire à : $\Big\langle f^*\Big(z - f(u)\Big), x\Big\rangle = 0\,, \ \ \forall x \in E$ et donc à $f^*(z) - f^* \circ f\,(u) = 0$;
d'où : $u = (f^* \circ \ f)^{-1} \circ f\ (z)$. On en déduit l'expression de p.
4. $\|\,b - f(x_0)\,\| = \|\,b - f \circ f^{\dagger}(b)\,\| = \|\,b - p(b)\,\| = \mathrm{d}\,(\,b, \operatorname{Im} f\,)$ (cf exercice 17). Donc :
$$\|\,b - f(x_0)\,\| = \operatorname*{Inf}_{x \in E}\ \|\,b - f(x)\,\|$$
5. a) $A^{\dagger} = (A^*\,A)^{-1}A^*$. On trouve
$$A^{\dagger} = \frac{1}{50}\begin{pmatrix} 18 & 15 & -1 \\ 2 & -10 & 14 \end{pmatrix}$$
b) Le système est compatible pour $a = -2$ et, dans ce cas, la solution est
$x = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$.
La solution des moindres carrés est
$$x_0 = A^{\dagger}\begin{pmatrix} 1 \\ 2 \\ a \end{pmatrix} = \frac{1}{50}\begin{pmatrix} 48 - a \\ -22 + 14\,a \end{pmatrix}$$
(Noter que pour $a = -2$, on retrouve la solution du cas compatible).

24 Utiliser le fait que $A\,{}^t\mathrm{Cof}\,(A) = (\mathrm{dét}\,A)\,I$, pour toute matrice A, et que $A\,{}^tA = I$ si A est orthogonale.

25

1. Soit $A = \|\, v_1, \ldots, v_n \|_{e_i}$, $\{e_i\}$ étant la base canonique de $\mathbb{R}^n$. On a : $A = P_{e_i \to v_i}$. Soit $\{e'_i\}$ la base orthonormée obtenue en appliquant le procédé de Schmidt à la base $\{v_i\}$, et $Q := P_{e_i \to e'_i}$. D'après la proposition 7.24, $Q \in \mathrm{O}(n, \mathbb{R})$. On a :
$$A = P_{e_i \to v_i} = P_{e_i \to e'_i}\, P_{e'_i \to v_i} = Q\,R$$
où $R = P_{e'_i \to v_i}$, qui est triangulaire supérieure (cf. démonstration du procédé de Schmidt).
2. Le procédé de Schmidt appliqué à la base $\{v_1, v_2, v_3\}$ donne la base orthonormée $\{e'_1, e'_2, e'_3\}$ et
$$Q = \|\, e'_1, e'_2, e'_3 \,\| = \frac{1}{\sqrt{6}} \begin{pmatrix} 0 & 2 & \sqrt{2} \\ \sqrt{3} & -1 & \sqrt{2} \\ \sqrt{3} & 1 & -\sqrt{2} \end{pmatrix}$$
$$R = Q^{-1}\,A = {}^tQ\,A = \frac{1}{\sqrt{6}} \begin{pmatrix} 2\sqrt{3} & \sqrt{3} & 1 \\ 0 & 3 & 1 \\ 0 & 0 & 2\sqrt{2} \end{pmatrix}.$$

26
$$\begin{aligned} x \in \mathrm{Im}(f - \mathrm{id})^{\perp} &\iff \langle x, f(y) - y \rangle = 0\,, \quad \forall y \in E \\ &\iff \langle f^*(x), y \rangle - \langle x, y \rangle = 0\,, \quad \forall y \in E \\ &\iff \langle\, (f^* - \mathrm{id})(x), y \rangle = 0\,, \quad \forall y \in E \iff x \in \mathrm{Ker}(f^* - \mathrm{id}) \end{aligned}.$$
Si $(f - \mathrm{id})^2 = 0$ alors $\mathrm{Im}(f - \mathrm{id}) \subset \mathrm{Ker}(f - \mathrm{id})$, donc : $\mathrm{Im}(f - \mathrm{id}) \subset \mathrm{Im}(f - \mathrm{id})^{\perp}$, d'où : $f - \mathrm{id} = 0$.

27 A : réflexion par rapport au plan d'équation $x + y + z = 0$.

B : rotation d'angle $\dfrac{5\pi}{3}$ autour de l'axe dirigé et orienté par le vecteur $\dfrac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$.

C : rotation d'angle $\dfrac{2\pi}{3}$ autour de l'axe dirigé et orienté par le vecteur $\dfrac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ -1 \\ -1 \end{pmatrix}$ suivie d'une réflexion par rapport au plan d'équation $x - y - z = 0$.

28

1. $g^* = f^* + f = g$
2. Utiliser le fait que g commute avec f.
3. $\forall x \in V_{\lambda_i}$ $g(x) = \lambda_i x$, c'est-à-dire $(f + f^* - \lambda_i\,\mathrm{id})\,(x) = 0$, d'où :
$$(f^2 + f \circ f^* - \lambda_i f)\,(x) = 0\,, \quad \forall x \in V_{\lambda_i}.$$
4. Pour $\lambda_i = 2$ $\quad Q(X) = (X - 1)^2$. Appliquer le résultat de l'exercice 26.
5. Même démonstration.
6. Si $v \in V_{\lambda_i}$ était vecteur propre de $f|_{\lambda_i} = f_{\lambda_i}$, λ_i devrait être racine de $Q(X) = X^2 - \lambda_i X + 1$, car ce polynôme annule f_{λ_i} (cf. proposition 6.18 page 176). Or f_{λ_i} est orthogonale, donc ses seules valeurs propres sont $+1$ et -1 et $Q(1) = 2 - \lambda_i \neq 0$, $Q(-1) = 2 + \lambda_i \neq 0$. W est de dimension 2 car v et $f(v)$ ne sont pas liés.
7. Soit $x \in W$, $x = \lambda v + \mu f(v)$. Puisque $(f^2 - \lambda_i\, f + \mathrm{id})\,|_{V_{\lambda_i}} = 0$, $f(x) = \lambda f(v) + \mu f^2(v) = \lambda f(v) + \mu(\lambda_i f(v) - v) \in W$. Donc W est stable par f.
8. $\tilde{f}$ est orthogonale. D'autre part $M(\tilde{f})_{\{v, f(v)\}} = \begin{pmatrix} 0 & -1 \\ 1 & \lambda \end{pmatrix}$. Donc $\mathrm{dét}\,\tilde{f} = 1$.

9. Soit $x \in W^{\perp}$, c'est à dire $\langle x, v\rangle = 0$ et $\langle x, f(v)\rangle = 0$; il s'agit de montrer que $\langle f(x), v\rangle = 0$ et $\langle f(x), f(v)\rangle = 0$. Puisque f est orthogonale, on a : $\langle f(x), f(v)\rangle = \langle x, v\rangle = 0$. D'autre part, $\langle f(x), v\rangle = \left\langle f^2(x), f(v)\right\rangle$; utiliser le fait que $(f^2 - \lambda_i\, f + \mathrm{id})\ |_{V_{\lambda_i}} = 0$.

10. Raisonner par récurrence, en répétant, dans $W^{\perp}$, la construction d'un sous-espace de dimension 2 stable par f. On décompose ainsi chaque V_{λ_i} en des sous-espaces stables de dimension 2 tels que les restrictions de f à chacun de ces sous-espaces soient des rotations. Écrire la matrice de f dans une base de E adaptée à la décomposition en somme directe $E = V_2 \oplus V_{-2} \oplus V_{\lambda_1} \oplus \cdots \oplus V_{\lambda_m}$, où chaque V_{λ_m} est décomposé en somme directe de sous-espaces de dimension 2, en restriction desquels f est une rotation.

29

1. Montrer par recurrence que si $F_1, ..., F_r$ sont r sous-espaces vectoriels d'un espace vectoriel E de dimension n, alors :
$$\dim(F_1 \cap \cdots \cap F_r) \ \geq\ \textstyle\sum_{i=1}^{r} \dim F_i \ - n\,(r-1).$$
En déduire l'inégalité en tenant compte du fait que
$$E_1^{\sigma_1} \cap \cdots \cap E_1^{\sigma_r} \subset E_1^{f}.$$

2. (a) Montrer que $E_1^{\sigma\,\perp}$ est stable par σ. Puisqu'il est de dimension 1, cela signifie que c'est un espace propre. Or les valeurs propres d'une transformation orthogonale sont nécessairement +1 ou –1. Donc $E_1^{\sigma\,\perp} = E_{-1}^{\sigma}$.
On en déduit que $E = E_1^{\sigma} \oplus E_1^{\sigma\,\perp} = E_1^{\sigma} \oplus E_{-1}^{\sigma}$; donc σ est diagonalisable. Pour montrer que $\sigma^2(x) = x$ pour tout x, décomposer x sur E_1^{σ} et E_{-1}^{σ}.
(b) On vérifie facilement que σ_a ainsi définie est une réflexion. Réciproquement, si σ est une réflexion, le projecteur (orthogonal) p_{-1} sur E_{-1}^{σ} est $p_{-1}(x) = \dfrac{\langle x, a\rangle}{\|a\|^2} a$, où a est un vecteur propre correspondant à -1. D'autre part, si $x = x_1 + x_{-1}$ avec $x_1 \in E_1^{\sigma}$ et $x_{-1} \in E_{-1}^{\sigma}$, on a : $\sigma(x) = x_1 - x_{-1}$. D'où l'expression de σ.
(c) Prendre $a = y - z$.

3. (a) Soit $y \in H$, c'est-à-dire $\langle y, z\rangle = 0$ pour tout $z \in E$. On a :
$$\langle f(y), z\rangle \ = \ \langle f(y), f(z)\rangle \ = \ \langle y, z\rangle \ = 0.$$
(b) On a $\sigma_b(x) \ = \ x - 2\,\dfrac{\langle x, b\rangle}{\|b\|^2}\,b$ pour tout $x \in E$. En particulier, pour tout $y \in H$: $\sigma_b(y) = y - 2\,\dfrac{\langle y, b\rangle}{\|b\|^2}\,b = \tilde{\sigma}_b(y)$.
La propriété $(\mathcal{P})$ est vraie pour $n = 1$, car dans ce cas $\mathrm{O}(E) = \{\mathrm{id}, \sigma = -\,\mathrm{id}\}$, et $\mathrm{id} = \sigma^0$. Supposons la propriété vraie jusqu'à l'ordre $n-1$ et soit $\tilde{f} = f\ |_H$. D'après l'hypothèse de récurrence, on peut écrire $\tilde{f} = \tilde{\sigma}_{b_1} \circ \cdots \circ \tilde{\sigma}_{b_r}$, avec $r \leq n-1$ et $b_1, ..., b_r \in H$. Or $\sigma_{b_i}(z) = z$, donc $f = \sigma_{b_1} \circ \cdots \circ \sigma_{b_r}$, car les deux membres sont égaux sur H et sur $\mathrm{Vect}\{z\}$.
(c) σ_a échange z et $f(z)$ (cf. 2.(c)), donc : $\sigma_a \circ f(z) = z$.
Ainsi l'application $\sigma_a \circ f$ est une application orthogonale dont l'espace propre correspondant à +1 n'est pas réduit à $\{0\}$. En utilisant le résultat de 3.(b), on peut écrire $\sigma_a \circ f = \sigma_{a_1} \circ \cdots \circ a_r$, avec $r \leq n-1$, d'où :
$$f = \sigma_a \circ \sigma_{a_1} \circ \cdots \circ a_r.$$
f est donc le produit de $r + 1 \leq n$ réflexions.

4. $z = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \in E_1^{A}$. On est donc dans le cas 3.(b) et A est produit d'au plus deux réflexions σ_b, σ_{b_1}. Prendre $b \in z^{\perp}$, par exemple $b = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$. D'après 2.(b) :

$$M(\sigma_b) = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Puisque $A = \sigma_b \circ \sigma_{b_1}$, on a : $\sigma_{b_1} = \sigma_b\, A = \dfrac{1}{3}\begin{pmatrix} -1 & 2 & 2 \\ 2 & 2 & -1 \\ 2 & -1 & 2 \end{pmatrix}.$

30

1. Utiliser la formule (*), page 246.
2. On voit facilement que $e_i + e_j \perp e_i - e_j$ donc $\langle f(e_i+e_j), f(e_i - e_j)\rangle = 0$, d'où : $\| f(e_i)\| = \|f(e_j)\|$. Notons μ cette valeur commune.
3. Soit $x = \sum_{i=1}^{n} x_i\, e_i$ un vecteur quelconque de E. On a : $\|x\|^2 = \sum_{i=1}^{n} x_i^2$ et $\|f(x)\|^2 = \mu^2 \sum_{i=1}^{n} x_i^2$. Conclure.

31

1. Si $x' = 0$ la relation est triviale.
   Soit $x' \neq 0$; le vecteur $\vec{n} \wedge x = \vec{n} \wedge x'$ appartient au plan $\pi = \mathrm{Vect}\{\vec{n}\}^{\perp}$ et il est orthogonal à x' et de même norme que x'. De plus la base $\{x', \vec{n}\wedge x'\}$ de π est orientée positivement par l'orientation induite par $\vec{n}$.
2. Si $\vec{n} = \begin{pmatrix} a \\ b \\ c \end{pmatrix}$ (avec $a^2+b^2+c^2 = 1$) on a :

$$M(f)_{e_i} = \begin{pmatrix} \cos\theta + a^2(1-\cos\theta) & -c\sin\theta + ab(1-\cos\theta) & b\sin\theta + ac(1-\cos\theta) \\ c\sin\theta + ab(1-\cos\theta) & \cos\theta + b^2(1-\cos\theta) & -a\sin\theta + bc(1-\cos\theta) \\ -b\sin\theta + ac(1-\cos\theta) & a\sin\theta + bc(1-\cos\theta) & \cos\theta + c^2(1-\cos\theta) \end{pmatrix}$$

32 Supposons $f \in \mathrm{SO}(E)$. Si x et y sont liés la relation est triviale.
Si x et y sont indépendants, alors $(x, y, x\wedge y)$ est une base directe. Puisque $f \in \mathrm{SO}(E)$, $(f(x), f(y), f(x\wedge y))$ est aussi une base directe. Par ailleurs f conserve l'orthogonalité, donc $f(x\wedge y) \perp \mathrm{Vect}\,\{f(x), f(y)\}$ (puisque $x\wedge y \perp \mathrm{Vect}\,\{x,y\}$). Aussi, existe-t-il $\lambda \in \mathbb{R}$ tel que :

$$f(x\wedge y) = \lambda f(x) \wedge f(y)$$

En utilisant l'identité de Lagrange et le fait que $f \in \mathrm{SO}(E)$, montrer que

$$\|f(x\wedge y)\|^2 = \|f(x)\wedge f(y)\|^2$$

d'où $|\lambda| = 1$. D'autre part

$$\text{dét}\,\|f(x), f(y), f(x\wedge y)\| = (\text{dét}\, f)\cdot \text{dét}\,\|x, y, x\wedge y\| = \text{dét}\,\|x,y,x\wedge y\| > 0$$

et

$$\text{dét}\,\|f(x), f(y), f(x\wedge y)\| = \lambda \cdot \text{dét}\,\|f(x), f(y), f(x)\wedge f(y)\|$$

d'où $\lambda = 1$.

Réciproquement, supposons que f conserve le produit vectoriel et soit (e_i) une b.o.n. directe. Il s'agit de montrer que $(f(e_i))$ est aussi une b.o.n. directe.
On a : $f(e_3) = f(e_1\wedge e_2) = f(e_1)\wedge f(e_2)$, donc $f(e_3) \perp \mathrm{Vect}\,\{f(e_1), f(e_2)\}$. En calculant de même $f(e_1)$ et $f(e_2)$, on voit que la famille $(f(e_i))$ est une famille orthogonale. D'autre part, $f(e_3) = f(e_1)\wedge f(e_2)$, d'où : $\|f(e_3)\| = \|f(e_1)\|\,\|f(e_2)\|$ (utiliser l'identité de Lagrange). En échangeant le rôle des e_i, montrer que les vecteurs $f(e_i)$ sont de norme 1.

33 $\mathrm{Sp}'A = \{0, 6, 6\}$ $\quad A' = P^{-1}AP = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 6 \end{pmatrix}$, avec :

$$P = \|u, v, w\|\,, \quad u = \frac{1}{\sqrt{6}}\begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix}, \quad v = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \quad w = \frac{1}{\sqrt{3}}\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

Projection orthogonale sur le plan $\mathrm{Vect}\{v, w\}$ suivie d'une homothétie de rapport 6.

34 On a $\mathrm{Tr}({}^tAA) = \sum_{i,j=1}^{n} a_{ij}^2$. Puisque A est symétrique, il existe P orthogonale telle que

$$A' = {}^tPAP = \begin{pmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{pmatrix}$$

$$\sum_{i=1}^{n} \lambda_i^2 = \mathrm{Tr}({}^tA'A') = \mathrm{Tr}^t({}^tPAP)({}^tPAP) = \mathrm{Tr}^tAA = \sum_{i=1}^{n} a_{ij}^2.$$

35 Montrer que f conserve l'orthogonalité de $\mathcal{B}$ si et seulement si $\mathcal{B}$ est une base orthogonale de vecteurs propres de $f^* \circ f$. Utiliser ensuite le fait que $f^* \circ f$ est autoadjoint.

36 On munit $\mathcal{M}_n(\mathbb{R})$ du produit scalaire canonique (cf. page 222) ; montrer d'abord que pour toutes matrices $A,\, B, H \in \mathcal{M}_n(\mathbb{R})$, on a :

$$\begin{aligned} \langle HA\,,\, B\rangle &= \big\langle A\,,\, {}^tH\,B\big\rangle \\ \langle AH\,,\, B\rangle &= \big\langle A\,,\, B\,{}^tH\big\rangle \end{aligned}$$

A l'aide de ces relations, montrer que ψ_M est autoadjoint.

37

1. Soit $\mathrm{Sp}'\rho = \{\lambda_1, ..., \lambda_n\}$. Si $\{e_i\}$ est une base orthonormé de vecteurs propres de ρ, on définit σ par :

$$M(\sigma)_{e_i} = \begin{pmatrix} \sqrt{\lambda_1} & & 0 \\ & \ddots & \\ 0 & & \sqrt{\lambda_n} \end{pmatrix} \tag{1}$$

   D'autre part σ est unique, parce que si σ est un endomorphisme autoadjoint et défini positif tel que $\rho = \sigma^2$, en considérant une base orthonormée $\{e_i\}$ de vecteurs propres de σ, on voit immédiatement que cette base est aussi une base de vecteurs propres de ρ et que $M(\sigma)_{e_i}$ a nécessairement l'expression (1).

2. $(f^* \circ f)^* = f^* \circ f$, donc $f^* \circ f$ est autoadjoint. D'autre part si v est vecteur propre pour la valeur propre λ, on a :

$$\langle f^* \circ f(v), v\rangle = \langle f(v), f(v)\rangle = \langle f^* \circ f(v), v\rangle = \lambda\ \langle v, v\rangle$$

   donc $\lambda \geq 0$ et, puisque f est bijectif, $\lambda > 0$.

3. $$\begin{aligned}\langle u(x), u(y)\rangle &= \big\langle f\circ\sigma^{-1}(x), f\circ\sigma^{-1}(y)\big\rangle = \big\langle f^*\circ f\circ\sigma^{-1}(x), \sigma^{-1}(y)\big\rangle \\ &= \big\langle \sigma(x), \sigma^{-1}(y)\big\rangle = \big\langle x, \sigma\circ\sigma^{-1}(y)\big\rangle = \langle x, y\rangle\end{aligned}$$

   u est donc orthogonal. Aussi $f = u \circ \sigma$, avec u orthogonal et σ est défini positif.
   Pour montrer l'unicité, on suppose que $f = u \circ \sigma$, avec u orthogonal et σ autoadjoint défini positif. On a $\sigma = u^* \circ f$, et $u^* \circ f = f^* \circ u$, car σ est autoadjoint. Montrer que $\sigma^2 = f^* \circ f$ et en déduire l'unicité de σ, en tenant compte de 1. Puisque f est bijectif, σ est bijectif. En déduire l'unicité de u.

Chapitre 8

Espaces hermitiens

8.1 Formes hermitiennes. Produit scalaire hermitien

Dans ce chapitre, nous nous proposons d'établir, dans le cadre des espaces vectoriels complexes, une théorie analogue à celle du produit scalaire [1]. Il ne s'agit pas d'une généralisation gratuite : les notions que nous introduirons interviennent en des nombreuses branches des mathématiques et de la physique mathématique, par exemple en théorie quantique des champs. D'où leur importance.

1. Formes hermitiennes

Soit E un espace vectoriel de dimension finie sur $\mathbb{C}$. Pour définir l'analogue du produit scalaire sur E, on serait tenté de considérer la forme bilinéaire symétrique

$$s : \quad \begin{array}{ccc} E \times E & \longrightarrow & \mathbb{C} \\ (x,y) & \longmapsto & x_1 y_1 + \cdots + x_n y_n \end{array}$$

($x = \sum_{i=1}^{n} x_i e_i, \; y = \sum_{i=1}^{n} y_i e_i, \; \{e_i\}$ base de E). Cependant la propriété *spécifique* du produit scalaire, à savoir que s est définie positive, n'est pas satisfaite. Ceci pour deux raisons :

- d'abord parce qu'il n'a pas de sens de dire que $s(x,x) \geq 0$; ($s(x,x)$ est un nombre complexe) ;
- d'autre part parce que si $s(x,x) = x_1^2 + \cdots + x_n^2$ est nul cela n'implique pas que $x = 0$.

Considérons, en revanche, l'application

$$s' : \quad \begin{array}{ccc} E \times E & \longrightarrow & \mathbb{C} \\ (x,y) & \longmapsto & \overline{x}_1 \, y_1 + \cdots + \overline{x}_n \, y_n \end{array}$$

On aura :

$$s'(x,x) = |\, x_1 \,|^{\,2} + \cdots + |\, x_n \,|^{\,2}$$

[1] On trouvera en Appendice A.13 une table de correspondance entre les définitions concernant les espaces euclidiens et la théorie que nous allons développer. Il est conseillé de se rapporter à cette table pour se familiariser avec la terminologie.

et dans ce cas on a bien :

$$s'(x,x) \geq 0\,,\ \forall\, x \in E \quad \text{et} \quad s'(x,x) = 0 \Longrightarrow x = 0.$$

C'est pourquoi pour les espaces vectoriels complexes on utilise, comme analogue du produit scalaire, l'application s' plutôt que s.
Cependant, s' n'est pas bilinéaire ; on a, en effet, $\forall x, y, z \in E$, $\forall \lambda \in K$:

$$\begin{aligned} s'(x+y,z) &= \ s'(x,z) + s'(y,z) \\ s'(x,y+z) &= \ s'(x,y) + s'(x,z) \\ s'(x,\lambda y) &= \ \lambda s'(x,y) \end{aligned}$$

mais :

$$s'(\lambda x, y) = \overline{\lambda}\, s'(x,y).$$

(on dit que s' est ***antilinéaire*** en x.)
De plus s' n'est pas symétrique, mais vérifie la propriété suivante :

$$s'(x,y) = \overline{s'(y,x)}$$

Cette propriété est dite ***symétrie hermitienne.***

Définition 8.1 – *Soit E un espace vectoriel su $\mathbb{C}$. On appelle* ***forme hermitienne*** *une application*

$$h : E \times E \longrightarrow \mathbb{C}$$

antilinéaire dans le premier argument, linéaire dans le second argument [2] *et à symétrie hermitienne, c'est-à-dire telle que, $\forall x, y, z \in E$, $\forall \lambda \in K$:*

$$\begin{array}{lrcl} a) & h\,(x+y,z) & = & h\,(x,z) + h\,(y,z) \\ b) & h\,(x,y+z) & = & h\,(x,y) + h\,(x,z) \\ c) & h\,(x,\lambda\, y) & = & \lambda\, h\,(x,y) \\ d) & h\,(\lambda\, x,y) & = & \overline{\lambda}\, h\,(x,y) \\ & & \text{et} & \\ e) & h\,(x,y) & = & \overline{h\,(y,x)}. \end{array}$$

2. Produit scalaire hermitien

Les formes hermitiennes jouent le même rôle que les formes bilinéaires symétriques du chapitre précédent. Notons, en effet, que si h est une forme hermitienne, on a, en vertu de la propriété *e)* :

$$h(x,x) \in \mathbb{R}\,, \quad \forall\, x\, \in E. \qquad (*)$$

On peut donc parler de «forme hermitienne définie positive» : une forme hermitienne h est dite ***définie positive*** si :

$$h\,(x,x) \geq 0\,,\ \forall x \in E \quad \text{et} \quad h\,(x,x) = 0 \Longleftrightarrow x = 0.$$

[2] Une telle application, antilinéaire dans le premier argument et linéaire dans le second argument, est dite ***sesquilinéaire*** (terme qui vient du grec et signifie : à moitié linéaire)

Définition 8.2 – *Soit E un espace vectoriel sur $\mathbb{C}$. On appelle* ***produit scalaire hermitien*** *une forme hermitienne*

$$\begin{array}{rccc} \langle \mid \rangle : & E \times E & \longrightarrow & \mathbb{C} \\ & (x,y) & \longmapsto & \langle x \mid y \rangle \end{array}$$

définie positive. En d'autres termes, $\langle \mid \rangle$ vérifie les propriétés suivantes :

$$\begin{array}{lrcl} a) & \langle x+y \mid z \rangle & = & \langle x \mid z \rangle + \langle y \mid z \rangle \\ b) & \langle x \mid y+z \rangle & = & \langle x \mid y \rangle + \langle x \mid z \rangle \\ c) & \langle x \mid \lambda y \rangle & = & \lambda \, \langle x \mid y \rangle \\ d) & \langle \lambda x \mid y \rangle & = & \overline{\lambda} \, \langle x \mid y \rangle \\ e) & \langle x \mid y \rangle & = & \overline{\langle y \mid x \rangle} \\ & & \text{et} & \\ f) & \langle x \mid x \rangle \geq 0 \,, \ \forall x \in E & & \text{et } \langle x \mid x \rangle = 0 \Longleftrightarrow x = 0 \end{array}.$$

Un espace vectoriel de dimension ***finie*** *sur $\mathbb{C}$ muni d'un produit scalaire hermitien est dit* ***espace hermitien.***

REMARQUE. – Un espace vectoriel complexe E *non nécessairement de dimension finie*, muni d'un produit scalaire hermitien est dit ***préhilbertien*** complexe. De même un espace vectoriel réel, non nécessairement de dimension finie, muni d'un produit scalaire est dit ***préhilbertien réel***. Ainsi les espaces préhilbertiens de dimension finie sont les espaces hermitiens et euclidiens. Le préfixe "pré" devant l'adjectif "hilbertien" tient au fait que l'on réserve le nom d'***espace hilbertien*** aux espaces préhilbertiens qui sont *complets* (c'est-à-dire : toute suite de Cauchy converge) pour la norme que nous allons définir. On démontre que tout espace préhilbertien de dimension finie (c'est-à-dire tout espace euclidien ou hermitien) est complet.

Exemple 1 – $E = \mathbb{C}^n$ avec $\langle \mid \rangle$ défini par :

$$\langle x \mid y \rangle := \overline{x}_1 \, y_1 + \cdots + \overline{x}_n \, y_n \qquad \text{où : } x = (x_1, \cdots, x_n), \, y = (y_1, \cdots, y_n)$$

Le produit scalaire hermitien ainsi défini est dit ***produit scalaire hermitien canonique***. Comme pour les espaces euclidiens (cf. corollaire 7.7), nous verrons que si E est un espace hermitien, moyennant le choix d'une base, on peut identifier E à $\mathbb{C}^n$ muni du produit scalaire hermitien canonique. En d'autres termes il s'agit, à un changement de base près, du seul exemple d'espace hermitien

Exemple 2 – Soit E un espace vectoriel complexe de dimension n et $\{e_i\}$ une base de E. On définit alors un produit scalaire hermitien sur E, en posant :

$$\langle x \mid y \rangle_{e_i} := \overline{x}_1 \, y_1 + \cdots + \overline{x}_n \, y_n$$

pour $x = \sum_{i=1}^n x_i \, e_i$, $y = \sum_{i=1}^n y_i \, e_i$. $\langle x \mid y \rangle_{e_i}$ est dit ***produit scalaire hermitien associé à la base*** $\{e_i\}$.

Exemple 3 – Soient a et b deux applications continues de $[0,1]$ dans $\mathbb{R}$. On pose :

$$\int_0^1 \Big[a(x) + i \, b(x) \Big] dx := \int_0^1 a(x) \, dx + i \int_0^1 b(x) \, \mathrm{d}x$$

On considère l'espace vectoriel : $E = \{f \ : \ [0,1] \longrightarrow \mathbb{C}\,, \ \text{ continues }\}$. Il est facile de voir que l'application $h \ : \ E \times E \longrightarrow \mathbb{C}$ définie par :

$$h(f,g) = \int_0^1 \overline{f(x)}\, g(x)\, \mathrm{d}x$$

est une forme hermitienne définie positive. La vérification est laissée en exercice.

Comme nous allons le voir, tous les résultats sur les espaces euclidiens se transportent sans difficultés aux espaces hermitiens.

3. Réduction de Gauss

Soit h une forme hermitienne sur un espace vectoriel (complexe) E de dimension finie et soit $\{e_i\}$ une base de E. Puisque h est anti-linéaire dans le premier argument et linéaire dans le second, on a, pour $x = \sum_{i=1}^n x_i e_i$ et $y = \sum_{j=1}^n y_j e_j$:

$$h(x,y) = h\,(\sum_{i=1}^n x_i e_i \ , \ \sum_{j=1}^n y_j e_j) = \sum_{i,j=1}^n \overline{x}_i\, y_j\, h(e_i, e_j)$$

Les $h(e_i, e_j)$ sont des éléments de $\mathbb{C}$. Si on note $a_{ij} := h(e_i, e_j)$, l'expression de h dans la base $\{e_i\}$ est :

$$\boxed{h(x,y) = \sum_{i,j=1}^n a_{ij}\, \overline{x}_i\, y_j}$$

c'est-à-dire h est du type :

$$h(x,y) = a_{11}\, \overline{x}_1\, y_1 + a_{12}\, \overline{x}_1\, y_2 + \cdots + a_{ij}\, \overline{x}_i\, y_j + \cdots + a_{nn}\, \overline{x}_n\, y_n$$

Ainsi, par exemple :

$$f\,(x,y\,) = \overline{x}_1\, y_1 + (1 - 3\,i\,)\, \overline{x}_2\, y_2 - 5\,i\, \overline{x}_3\, y_1 + \overline{x}_1\, y_3$$

est anti-linéaire dans le premier argument et linéaire dans le second.
La vérification de la symétrie hermitienne est tout aussi facile : le fait que la valeur de $h(x,y\,)$ change en sa conjuguée lorsqu'on échange les rôles de x et y, équivaut, bien entendu, à

$$a_{ij} = \overline{a_{ji}}$$

c'est-à-dire le coefficient de $\overline{x}_i\, y_j$ doit être égal au conjugué de celui de $\overline{x}_j\, y_i$. Ainsi, par exemple, l'expression $f(x,y)$ ci-dessus n'est pas une forme hermitienne, alors que :

$$\begin{aligned} h(x,y\,) = \ & \overline{x}_1\, y_1 + 5\, \overline{x}_2\, y_2 + 4\, \overline{x}_3\, y_3 - i\ \overline{x}_1\, y_2 + i\, \overline{x}_2\, y_1 + (3 + 2\,i\,)\, \overline{x}_1 y_3 \\ & +(3 - 2\,i\,)\, \overline{x}_3\, y_1 + (1 + i\,)\, \overline{x}_2\, y_3 + (1 - i\,)\, \overline{x}_3\, y_2 \end{aligned}$$

est hermitienne.

Pour reconnaître si une forme hermitienne est définie positive, on peut adapter sans difficultés la méthode de Gauss. Voici comment procède-t-on.
Notons tout d'abord que, comme il en est pour le formes bilinéaires, une forme hermitienne h est connue si l'application $\tilde{q} : E \longrightarrow \mathbb{R}$ définie par

$$\tilde{q}\,(x) = h\,(x,x)$$

est connue. Par exemple, soit

$$h(x,y\,) = \overline{x}_1\,y_1 + 5\,\overline{x}_2\,y_2 + (3+i)\,\overline{x}_1\,y_2 + (3-i)\,\overline{x}_2\,y_1;$$

on a

$$\tilde{q}(x) = |x_1|^2 + 5\,|x_2|^2 + (3+i)\,\overline{x}_1\,x_2 + (3-i)\,\overline{x}_2\,x_1.$$

Pour revenir à $h(x,y)$ il suffira de remplacer les x_i par le y_i (cf. aussi exercice 3).

Notons que, comme dans le cas euclidien, si h est définie positive tous les coefficients des termes $|x_i|^2$ dans $\tilde{q}\,(x)$ sont strictement positifs (car ces coefficients sont les valeurs de $h(e_i,e_i)$). Ainsi, si dans $\tilde{q}\,(x)$ il n'y a pas de carrés de modules, h n'est pas un produit scalaire hermitien. Supposons donc que $\tilde{q}\,(x)$ contient un terme en $|x_i|^2$ et, pour simplifier, raisonnons sur un exemple (le cas général se traite d'une manière analogue).

Soit h la forme hermitienne sur $\mathbb{C}^3$ définie dans la base canonique $\{e_i\}$ par :

$$\tilde{q}(x) = |x_1|^2 + 5\,|x_2|^2 + 3\,|x_3|^2 + 2\,i\,\overline{x}_1\,x_2 - 2\,i\,x_1\,\overline{x}_2 + i\,\overline{x}_2\,x_3 - i\,x_2\,\overline{x}_3$$

On choisit un terme en $|x_i|^2$ (par exemple $|x_1|^2$) et on cherche le coefficient de $\overline{x}_i$ (ici : $x_1 + 2\,ix_2$).
On écrit :

$$\tilde{q}\,(x) = |x_1 + 2\,i\,x_2|^2 + \text{ termes correctifs}$$

Dans les termes correctifs il n'y a ni $\overline{x}_i$ ni x_i, comme on le vérifie facilement. On itère ensuite le procédé. Dans notre cas :

$$\begin{aligned}
\tilde{q}(x) &= |x_1 + 2\,i\,x_2|^2 - 4\,|x_2|^2 + 5\,|x_2|^2 + 3\,|x_3|^2 + i\,\overline{x}_2\,x_3 - i\,\overline{x}_3 x_2\\
&= |x_1 + 2\,i\,x_2|^2 + |x_2|^2 + 3\,|x_3|^2 + i\,\overline{x}_2\,x_3 - \,i\,\overline{x}_3\,x_2\\
&= |x_1 + 2\,i\,x_2|^2 + |x_2 + i\,x_3|^2 - |x_3|^2 + 3\,|x_3|^2 = |x_1 + 2\,i\,x_2|^2 + |x_2 + i\,x_3|^2 + 2\,|x_3|^2
\end{aligned}$$

On voit immédiatement que $\tilde{q}(x) \geq 0$, $\forall x \in \mathbb{C}^3$, et que $\tilde{q}(x) = 0$ si et seulement si $x = 0$.
h est donc définie positive.

Exercices 1. 2. 3. 4. 5.

8.2 Inégalité de Cauchy-Schwarz. Norme

Soit $(E, \langle\ |\ \rangle)$ un espace vectoriel complexe (non nécessairement de dimension finie) muni d'un produit scalaire hermitien. On pose

$$\|\, x \,\| := \sqrt{\langle x\,|\,x\rangle}.$$

Inégalité de Cauchy-Schwarz . 8.3 –
Soit E préhilbertien. Alors pour tous $x, y \in E$:

$$| \langle x \,|\, y \rangle | \leq \| x \| \, \| y \|,$$

l'égalité ayant lieu si et seulement si x et y sont liés.

Démonstration : Si $\langle x \,|\, y \rangle = 0$ l'inégalité est évidente.
Supposons $\langle x \,|\, y \rangle \neq 0$ (en particulier, $y \neq 0$). Pour tout $\lambda \in \mathbb{C}$ on a :

$$\begin{aligned} 0 \leq \ \langle x + \lambda\, y \,|\, x + \lambda\, y \rangle \ &= \| x \|^2 + \overline{\lambda} \ \langle y \,|\, x \rangle \ + \ \lambda \, \langle x \,|\, y \rangle \ + \ |\lambda|^2 \ \| y \|^2 \\ &= \| x \|^2 + \overline{\lambda} \, \overline{\langle x \,|\, y \rangle} \ + \lambda \, \langle x \,|\, y \rangle \ + \ |\lambda|^2 \ \| y \|^2 \end{aligned}$$

Ceci étant vrai pour tout $\lambda \in \mathbb{C}$, posons $\lambda = \mu \, \dfrac{\overline{\langle x \,|\, y \rangle}}{| \langle x \,|\, y \rangle |}$, avec $\mu \in \mathbb{R}$.
On aura :

$$0 \leq \| x \|^2 + 2 \, \mu \, | \langle x \,|\, y \rangle | + \mu^2 \ \| y \|^2 \qquad \forall \mu \in \mathbb{R}.$$

Puisque $y \neq 0$, le second membre est un trinôme en μ ; son discriminant est donc ≤ 0, c'est-à-dire :

$$| \langle x \,|\, y \rangle |^2 \leq \| x \|^2 \| y \|^2 \qquad \forall x, y \in E.$$

Si $x + \lambda\, y = 0$ on a l'égalité, dans cette expression (il suffit de remplacer x par $-\lambda\, y$).
Réciproquement l'égalité implique l'existence d'une racine μ dans le trinôme et donc d'un $\lambda \in \mathbb{C}$ tel que $x + \lambda\, y = 0$. □

L'inégalité de Cauchy-Schwartz permet de montrer que tout espace préhilbertien est un espace vectoriel normé :

Proposition 8.4 – *L'application : $\| \ \ \| : \ E \longrightarrow \mathbb{R}^+$ définie par $\| x \| = \sqrt{\langle x \,|\, x \rangle}$ est une* ***norme*** *sur E, c'est-à-dire elle vérifie :*

1. $\| x \| \geq 0$ *et* $\| x \| = 0 \Longleftrightarrow x = 0$;
2. $\| \lambda\, x \| = |\lambda| \ \| x \|, \qquad \forall \lambda \in \mathbb{C}, \ \forall x \in E$;
3. $\| x + y \| \leq \| x \| + \| y \|, \quad \forall \, x, y \in E.$
   De plus l'égalité a lieu si et seulement si il existe $\lambda \in \mathbb{R}$, $\lambda \geq 0$ tel que $y = \lambda\, x$ (ou $x = \lambda\, y$).

Démonstration : Il suffira de démontrer 3., car 1. et 2. dérivent immédiatement de la définition du produit scalaire hermitien.
On a, en notant $\mathcal{R}e \ z$ la partie réelle de $z \in \mathbb{C}$:

$$\begin{aligned} \| x + y \|^2 \ &= \ \langle x + y \,|\, x + y \rangle \ = \| x \|^2 + \langle y \,|\, x \rangle + \langle x \,|\, y \rangle + \| y \|^2 \\ &= \| x \|^2 + 2 \, \mathcal{R}e \, \langle x \,|\, y \rangle + \| y \|^2 \ \leq \ \| x \|^2 + 2 \, | \langle x \,|\, y \rangle | + \| y \|^2 \\ &\leq \| x \|^2 + 2 \ \| x \| \, \| y \| + \| y \|^2 = (\| x \| + \| y \|)^2 \end{aligned}$$

d'où : $\| x + y \| \leq \| x \| + \| y \|$.
Si $y = \lambda\, x$ avec $\lambda \in \mathbb{R}_+$, on l'égalité. Réciproquement si on a l'égalité, en remontant les calculs on trouve :

– d'une part que x et y sont colinéaires en vertu de l'inégalité de Cauchy-Schwarz ; donc, si, par exemple $x \neq 0$, $y = \lambda x$ avec $\lambda \in \mathbb{C}$;
– d'autre part que $\mathcal{R}e \langle x \mid y \rangle = | \langle x \mid y \rangle |$.

Puisque $\langle x \mid y \rangle = \langle x \mid \lambda x \rangle = \lambda \| x \|^2$, on aura $\mathcal{R}e \langle x \mid y \rangle = | \langle x \mid y \rangle |$ si et seulement si $\mathcal{R}e\, \lambda = |\lambda|$, c'est-à-dire si et seulement si $\lambda \in \mathbb{R}$ et $\lambda \geq 0$. □

Exercice 6.

8.3 Matrices hermitiennes

Comme il en est des formes bilinéaires, il est utile de représenter les formes hermitiennes par des matrices.

Soit E un espace vectoriel de dimension finie n sur $\mathbb{C}$ et $h : E \times E \longrightarrow \mathbb{C}$ une forme hermitienne. Si $\{e_i\}$ est une base de E et $x = \sum x_i e_i$, $y = \sum y_j e_j$, on a :

$$h(x,\, y) = \sum \overline{x}_i\, y_j\, h(e_i,\, e_j)$$

(cf. page 276). h est donc déterminée par la connaissance des valeurs $h(e_i,\, e_j)$ sur une base.

Définition 8.5 – *Soient h une forme hermitienne sur E, et $\{e_i\}$ une base de E. On appelle* **matrice de h** *dans la base $\{e_i\}$ la matrice :*

$$M(h)_{e_i} = \begin{pmatrix} h(e_1,e_1) & h(e_1,e_2) & \cdots & h(e_1,e_n) \\ h(e_2,e_1) & \cdots & \cdots & h(e_2,e_n) \\ \cdots & \cdots & \cdots & \cdots \\ h(e_n,e_1) & \cdots & \cdots & h(e_n,e_n) \end{pmatrix}$$

$$= \big(h\,(\underset{\substack{\uparrow \\ \textit{ième} \\ \textit{ligne}}}{e_i},\ \ \underset{\substack{\uparrow \\ \textit{jème} \\ \textit{colonne}}}{e_j})\big)$$

Ainsi l'élément de la i^{eme} ligne et j^{eme} colonne est le coefficient de $\overline{x}_i\, y_j$.

Exemple – Soit h la forme hermitienne sur $\mathbb{C}^3$ qui dans la base canonique $\{e_i\}$ est représentée par :

$$\begin{aligned} h(x,y) = \quad & \overline{x}_1\, y_1 + 5\, \overline{x}_2\, y_2 + 4\, \overline{x}_3\, y_3 - i\, \overline{x}_1\, y_2 + i\, \overline{x}_2\, y_1 + (3 + 2\, i\,)\, \overline{x}_1\, y_3 \\ & + (3 - 2\, i\,)\, \overline{x}_3\, y_1 + (1 + i\,)\, \overline{x}_2\, y_3 + (1 - i\,)\, \overline{x}_3\, y_2 \end{aligned} .$$

On a :

$$M(h)_{e_i} = \begin{pmatrix} 1 & -i & 3 + 2\, i \\ i & 5 & 1 + i \\ 3 - 2\, i & 1 - i & 4 \end{pmatrix}$$

Notons que la matrice H d'une forme hermitienne vérifie la propriété ${}^t\overline{H} = H$. On pose donc la définition suivante :

Définition 8.6 – *Une matrice $H \in \mathcal{M}_n(\mathbb{C})$ est dite* **hermitienne** *si : ${}^t\overline{H} = H$.*
En particulier les matrices symétriques réelles sont les matrices hermitiennes réelles.

Les matrices hermitiennes sont donc les matrices qui représentent les formes hermitiennes.

Remarque. –

1. Les éléments de la diagonale d'une matrice hermitienne sont des réels.
2. Nous allons voir (cf. théorème 8.16 et proposition 8.19) que, tout comme pour les matrices symétriques réelles (cf. théorème 7.38 et proposition 7.39) :
   - les matrices hermitiennes *sont diagonalisables dans* $\mathbb{R}$;
   - une forme hermitienne définit un produit scalaire hermitien si et seulement si la matrice qui la représente dans une base quelconque – qui est justement une matrice hermitienne – a toutes ses valeurs propres *strictement positives.*

Soient h une forme hermitienne sur E, $\{e_i\}$ une base, et

$$H = M(h)_{e_i}, \quad X = M(x)_{e_i}, \quad Y = M(y)_{e_i} \qquad (x,\, y \in E)$$

On a :

$$\boxed{h(x,\, y) = {}^t\overline{X}\, H\, Y}.$$

Changement de base

Soit h une forme hermitienne, $H = M\,(h)_{e_i}$, $\{e'_i\}$ une nouvelle base et $P = P_{e_i \to e'_i}$ la matrice de passage. Si X' et Y' sont les matrices de x et y dans la base $\{e'_i\}$, on a :

$$X' = P^{-1}X \qquad \text{et} \qquad Y' = P^{-1}Y.$$

En notant $H' = M(h)_{e'_i}$, on a :

$$h(x, y) = {}^t\overline{X'}\, H'\, Y'.$$

D'autre part :

$$h(x, y) = {}^t\overline{X}\, H\, Y = {}^t\overline{(PX')}\, H\, P\, Y' = {}^t\overline{X'}({}^t\overline{P}\, H\, P)\, Y$$

pour tous $X, Y \in \mathcal{M}_{n,1}(\mathbb{C})$, d'où :

$$\boxed{H' = {}^t\overline{P}\, H\, P}$$

Exercice 7.

8.4 Bases orthonormées. Orthogonalité

Les résultats sur l'orthogonalité dans les espaces euclidiens se transportent facilement au cas hermitien.

Définition 8.7 – *Soit* $(E, \langle\ |\ \rangle)$ *un espace hermitien. Une base* $\{e_i\}$ *est dite* ***orthogonale*** *si* $\langle e_i\,|e_j\rangle = 0$ *pour* $i \neq j$*. Elle est dite* ***orthonormée*** *si* $\langle e_i\,|e_j\rangle = \delta_{ij}$*.*

Il est clair que $\{e_i\}$ est une base orthogonale si et seulement si :

$$M(\langle\ |\ \rangle)_{e_i} = \begin{pmatrix} a_1 & 0 & \cdots & 0 \\ 0 & a_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & a_n \end{pmatrix} \qquad (a_i \in \mathbb{R})$$

ou encore :

$$\tilde{q}(x) = a_1 \, |x_1|^2 + \cdots + a_n \, |x_n|^2 \qquad (\text{ où } \; x = \sum_{i=1}^{n} x_i e_i)$$

De même, $\{e_i\}$ est une base orthonormée si et seulement si :

$$M(\langle \ | \ \rangle)_{e_i} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & 1 \end{pmatrix}$$

ou encore :

$$\tilde{q}(x) = |x_1|^2 + \cdots + |x_n|^2 \qquad (\text{ où } : x = \textstyle\sum_{i=1}^{n} x_i e_i).$$

Théorème 8.8 – *Sur tout espace hermitien il existe des bases orthonormées.*
En particulier tout espace hermitien est isomorphe à $\mathbb{C}^n$ *muni du produit scalaire hermitien canonique.*

La démonstration est la même que dans le cas euclidien (cf. théorème 7.6 page 228).
Notons que si X et Y sont les matrices qui représentent *dans une base orthonormée* les vecteurs x, y, on a :

$$\langle x \,|\, y \rangle \; = \; {}^t\overline{X}\, Y$$

Les procédés d'orthonormalisation de Schmidt s'adapte facilement au cas hermitien, permettant d'associer à toute base une base orthonormée d'une façon canonique.

Comme dans le cas euclidien on pose :

Définition 8.9 – *Soit* $A \subset E$. *On note :* $A^{\perp} = \{x \in \,|\, \langle x \,|\, a \rangle = 0\,, \;\; \forall a \in E\}$.
$A^{\perp}$ *est un sous-espace vectoriel de* E, *dit* **orthogonal** *de* A.

NOTA – A cause de la symétrie hermitienne, $A^{\perp}$ est aussi l'ensemble

$$\{x \in E | \; \langle a \,|\, x \rangle = 0\,, \;\; \forall a \in E\}.$$

Proposition 8.10 – *Soit* F *un sous-espace vectoriel d'un espace vectoriel hermitien* E. *Alors :*

1. $\dim E = \dim F + \dim F^{\perp}$
2. $E = F \oplus F^{\perp}$
3. $F^{\perp\perp} = F$

Ces propriétés se démontrent comme les propriétés 7.14 page 236.

Exercices 8. 9.

8.5 Endomorphisme adjoint

Proposition 8.11 *Soit E un espace hermitien et $f \in \mathrm{End}(E)$. Il existe un et un seul endomorphisme f^* de E tel que*

$$\langle\, f(x) \,|\, y\rangle \;=\; \langle\, x \,|\, f^*(y)\rangle\ , \quad \forall x, y \in E.$$

*f^* est dit **adjoint** de f. Si $\{e_i\}$ est une base* orthonormée *et $A = M(f)_{e_i}$, alors la matrice $A^* = M(f^*)_{e_i}$ est : $A^* = {}^t\overline{A}$.*

Démonstration : La démonstration est la même que dans le cas euclidien (cf. proposition 7.16). Soit $\{e_i\}$ une base orthonormée de E et notons

$$A = M(f)_{e_i}, \quad A^* = M(f^*)_{e_i}, \quad X = M(x)_{e_i},\ Y = M(y)_{e_i}$$

Puisque la base $\{e_i\}$ est orthonormée, l'identité de l'énoncé s'écrit :

$${}^t\overline{(A\,X)}\,Y \;=\; {}^t\overline{X}\,A^*\,Y\ , \quad \forall\, X,\, Y \in \mathcal{M}_{n,1}(\mathbb{R})$$

ce qui est équivalent à ${}^t\overline{A} \,=\, A^*$ Ceci montre que A^* (donc f^*) est unique.
Réciproquement, si on définit $f^* \in \mathrm{End}(E)$ par $M(f^*) \,=\, {}^t\overline{A}$, on voit, en remontant les calculs, que f^* vérifie l'identité de l'énoncé. □

On a les propriétés suivantes, qui se démontrent immédiatement dans une base orthonormée, compte tenu du fait que dans une base orthonormée $A^* \,=\, {}^t\overline{A}$.

Proposition 8.12 – *Pour tout endomorphisme f et pour tout scalaire λ, on a :*

a) $f^{**} = f\ , \quad (\mathrm{id})^* = \mathrm{id}$

b) $(f+g)^* = f^* + g^*\ , \qquad (\lambda f)^* = \overline{\lambda}\, f^*\ , \qquad (f\circ g)^* = g^* \circ f^*$

c) $\mathrm{rg}\ f^* = \mathrm{rg}\ f\ , \qquad \mathrm{d\acute{e}t}\, f^* = \overline{\mathrm{d\acute{e}t}\, f}$

NOTA. – Le résultat ci-dessus (8.11) justifie la notation habituelle pour $A \in \mathcal{M}_n(\mathbb{C})$:

$$A^* \underset{def}{=} {}^t\overline{A}.$$

La matrice A^* est dite ***adjointe*** de A.

Exercice 10.

8.6 Groupe unitaire

Le but de ce paragraphe est d'étudier les endomorphismes qui conservent le produit scalaire hermitien. Il s'agit donc de l'analogue, dans le cas complexe, des transformations orthogonales.

Proposition 8.13 – *Soit $(E, \langle\ |\ \rangle)$ un espace hermitien et f un endomorphisme de E. Les propriétés suivantes sont équivalentes :*

1. $\|\, f(x)\,\| \;=\; \|\, x\,\|\ , \ \ \forall x \in E$;
2. $\langle f(x)\,|\,f(y)\rangle \;=\; \langle x\ |\,y\rangle\ , \ \ \forall x, y \in E$;

3. *Si* $\{e_i\}$ *est une* base orthonormée *de* E *et* $A = M(f)_{e_i}$, *alors :* ${}^t\overline{A}\,A = I$ *(ou, d'une manière équivalente :* $A\,{}^t\overline{A} = I$*).*

Un endomorphisme qui vérifie ces propriétés est dit ***unitaire****.*

La démonstration est analogue à celle de la proposition 7.19.

Notons que f est unitaire si et seulement si $f^* \circ f = \mathrm{id}$, ou, d'une manière équivalente : $f \circ f^* = \mathrm{id}$.

Définition 8.14 – *Une matrice* $A \in \mathcal{M}_n(\mathbb{C})$ *vérifiant* ${}^t\overline{A}A = I$ *(ou d'une manière équivalente* $A\,{}^t\overline{A} = 1$ *) est dite* ***unitaire****. L'ensemble des matrices unitaires est un groupe dit* ***groupe unitaire****, noté* $\mathrm{U}\,(n, \mathbb{C})$ *:*

$$\mathrm{U}\,(n, \mathbb{C}) := \{A \in \mathcal{M}_n(\mathbb{C}) \mid {}^t\overline{A}\,A = I\}$$

Les matrices unitaires sont les matrices qui représentent *dans une base orthonormée* les transformations unitaires d'un espace hermitien. Il s'agit donc de l'analogue, dans le cas complexe des matrices orthogonales et, plus précisément, les matrices orthogonales sont les matrices unitaires réelles :

$$\mathrm{O}\,(n, \mathbb{R}) = \mathrm{U}\,(n, \mathbb{C}) \cap \mathcal{M}_n\,(\mathbb{R}).$$

Si $A \in \mathrm{U}\,(n, \mathbb{C})$, puisque ${}^t\overline{A}A = I$ on a : $|\,\text{dét}\,A\,|^2 = 1$, donc :

$$|\text{dét}\ A\,| = 1.$$

Les matrices unitaires de déterminant égal à 1 forment un sous-groupe de $\mathrm{U}(n, \mathbb{C})$, dit ***groupe spécial unitaire***, noté $\mathrm{SU}\,(n, \mathbb{C})$:

$$\mathrm{SU}\,(n, \mathbb{C}) := \{A \in \mathrm{U}\,(n, \mathbb{C}) \mid \text{dét}\ A = 1\}.$$

Les matrices spéciales unitaires sont donc l'analogue, dans le cas complexe des matrices spéciales orthogonales et, plus précisément :

$$\mathrm{SO}\,(n, \mathbb{R}) = \mathrm{SU}\,(n, \mathbb{C}) \cap \mathcal{M}_n\,(\mathbb{R}).$$

Comme dans le cas euclidien (cf. proposition 7.21 page 240 et proposition 7.24 page 241) on voit facilement que :

1. *Les endomorphismes unitaires sont les endomorphismes qui transforment les bases orthonormées en bases orthonormées.*
2. *La matrice de passage d'une base orthonormée à une base orthonormée est unitaire.*

Exemple 1 – Soit la matrice :

$$A = \begin{pmatrix} e^{i\varphi_1} & \cdots & & 0 \\ \vdots & e^{i\varphi_2} & & \vdots \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & & e^{i\varphi_n} \end{pmatrix}$$

On a ${}^t\overline{A}\,A = I$; donc A est unitaire : $A \in \mathrm{U}\,(n, \mathbb{C})$. D'autre part :

$$\text{dét}\ A = e^{i(\varphi_1 + \varphi_2 + \cdots + \varphi_n)}$$

donc $A \in \mathrm{SU}\,(n, \mathbb{C}) \Longleftrightarrow \varphi_1 + \varphi_2 + \cdots + \varphi_n = 2k\pi$.

Exemple 2 – **Détermination de** $\mathrm{SU}(2, \mathbb{C})$.

Soit $A = \begin{pmatrix} a & c \\ b & d \end{pmatrix} \in \mathcal{M}_2(\mathbb{C})$; on a : ${}^t\overline{A} = \begin{pmatrix} \overline{a} & \overline{b} \\ \overline{c} & \overline{d} \end{pmatrix}$.

La condition ${}^t\overline{A}\,A = I$ est équivalente au système

$$\begin{array}{r} \text{(I)} \\ \text{(II)} \\ \text{(III)} \end{array} \quad \left\{ \begin{array}{l} a\,\overline{a} + b\,\overline{b} = 1 \\ c\,\overline{c} + d\,\overline{d} = 1 \\ a\,\overline{c} + b\,\overline{d} = 0 \end{array} \right.$$

La condition $\det A = 1$ donne l'équation supplémentaire

$$\text{(IV)} \qquad ad - bc = 1.$$

D'autre part :

$$\text{(I) - (IV)} \quad \Longrightarrow \quad a\,(\overline{a} - d) + b\,(\overline{b} + c) = 0.$$

Il existe donc $\lambda \in \mathbb{C}$ tel que

$$\overline{a} - d = \lambda\, b \quad \text{et} \quad \overline{b} + c = -\lambda\, a$$

c'est-à-dire :

$$\begin{array}{l} d = \ \ \overline{a} - \lambda\, b \\ c = -\overline{b} - \lambda\, a \end{array}$$

En reportant dans (III) :

$$-\overline{\lambda}\,(a\,\overline{a} + b\,\overline{b}) = 0$$

d'où, d'après (I) : $\lambda = 0$ et

$$d = \overline{a}\ , \qquad c = -\overline{b}$$

Donc A s'écrit :

$$A = \begin{pmatrix} a & b \\ -\overline{b} & \overline{a} \end{pmatrix}, \qquad \text{avec :} \qquad |a|^2 + |b|^2 = 1$$

(on vérifie facilement la réciproque). En posant $a = \rho_1\, e^{i\theta}$ et $b = \rho_2\, e^{i\varphi}$, on a

$$|a|^2 + |b|^2 = 1 \Longleftrightarrow \rho_1^2 + \rho_2^2 = 1.$$

Il existe donc $\alpha \in \mathbb{R}$ tel que $\rho_1 = \cos\alpha$ et $\rho_2 = -\sin\alpha$
On a ainsi :

$$A \in \mathrm{SU}(2, \mathbb{C}) \Longleftrightarrow A = \begin{pmatrix} \cos\alpha\, e^{i\theta} & -\sin\alpha\, e^{i\varphi} \\ & \\ \sin\alpha\, e^{-i\varphi} & \cos\alpha\, e^{-i\theta} \end{pmatrix}$$

A dépend donc de trois paramètres réels : α, θ, φ. Pour $\theta = 0$ et $\varphi = 0$ on retrouve les matrices de $\mathrm{SO}(2, \mathbb{R})$.

Remarque. – *Les valeurs propres d'une matrice unitaire sont toutes de module 1.*

En particulier, les valeurs propres d'une matrice orthogonale, considérée comme matrice de $\mathcal{M}_n(\mathbb{C})$, sont des complexes de module 1.

En effet, soit f l'endomorphisme unitaire de $\mathbb{C}^n$ qui dans la base canonique est représenté par la matrice unitaire A. On a : $\| f(x) \| = \| x \|$, $\forall x \in \mathbb{C}^n$. Si $f(x) = \lambda x$ avec $x \neq 0$, alors :

$$\| \lambda x \| = \| x \| \quad \text{d'où :} \quad |\lambda|\, \| x \| = \| x \| \qquad \text{et donc} \quad |\lambda| = 1.$$

Exercices 11. 12.

8.7 Diagonalisation des endomorphismes autoadjoints d'un espace hermitien - Endomorphismes normaux

Définition 8.15 – *Un endomorphisme f d'un espace hermitien est dit **autoadjoint**, ou aussi **hermitien**, si : $f^* = f$, c'est-à-dire si :*

$$\langle f(x) \mid y \rangle \;=\; \langle x \mid f(x) \rangle \;, \qquad \forall x, y \in E.$$

Si $\{e_i\}$ est une ***base orthonormée*** et $A = M(f)_{e_i}$, f est autoadjoint si et seulement si ${}^t\overline{A} = A$, c'est-à-dire :

> *f est autoadjoint si et seulement si la matrice qui le représente* dans une base orthonormée *est hermitienne.*

Le théorème suivant qui affirme que tout automorphisme autoadjoint (donc toute matrice hermitienne) est diagonalisable dans $\mathbb{R}$, est la généralisation du théorème 7.38 page 252 :

Théorème 8.16 – *Soit f un endomorphisme autoadjoint d'un espace hermitien.*

1. *Les valeurs propres de f sont toutes réelles.*
2. *f est diagonalisable.*
3. *Les sous-espaces propres de f sont deux à deux orthogonaux (on peut donc construire une base orthonormée de vecteurs propres en choisissant une base orthonormée dans chaque espace propre).*

NOTA - En termes de matrices ce théorème peut aussi s'énoncer ainsi (cf. théorème 7.38') :

Théorème 8.16'

Toute matrice hermitienne est diagonalisable **dans** $\mathbb{R}$ *et les espaces propres sont deux à deux orthogonaux pour le produit scalaire hermitien canonique de $\mathbb{C}^n$, $\langle x \mid y \rangle \;=\; \overline{x}_1\, y_1 + \cdots + \overline{x}_n\, y_n$.*

Ou encore, compte tenu du fait que la matrice de passage d'une base orthonormée (la base canonique de $\mathbb{C}^n$) à une base orthonormée (la base orthonormée de vecteurs propres) est unitaire (cf. théorème 7.38") :

Théorème 8.16"

Soit A une matrice hermitienne. Il existe alors une matrice unitaire U, telle que la matrice $A' = {}^t\,\overline{U}\, A\, U$ soit diagonale **réelle**.

Notons que les théorèmes 8.16, 8.16' et 8.16" sont trois manières différentes d'énoncer le même théorème

Démonstration : Il est facile de vérifier que les valeurs propres d'un endomorphisme autoadjoint sont toutes réelles. Soit $x \neq 0$ tel que $f(x) = \lambda\, x$. On a :

$$\langle\, f(x) \mid x\,\rangle = \langle \lambda\, x \mid x\,\rangle = \overline{\lambda}\, \|\, x\, \|^2$$

et d'autre part :

$$\langle\, f(x) \mid x\,\rangle = \langle x \mid f(x) \rangle = \langle\, x \mid \lambda\, x\,\rangle = \lambda\, \|\, x\, \|^2$$

d'où : $\lambda = \overline{\lambda}$.

2. et 3. viennent du théorème suivant, beaucoup plus général, concernant ce que l'on appelle les endormophismes «normaux».

Définition 8.17 – *Un endomorphisme f d'un espace hermitien est dit* ***normal*** *si :*

$$f^* \circ f = f \circ f^*$$

Une matrice $A \in \mathcal{M}_n(\mathbb{C})$ est dite normale si : ${}^t\overline{A}A = A\,{}^t\overline{A}$.

Il est clair que les matrices normales sont les matrices qui représentent les endomorphismes normaux dans une *base orthonormée.*
En particulier sont normales :
les matrices *symétriques réelles,*
les matrices *antisymétriques réelles,*
les matrices *hermitiennes,*
les matrices *orthogonales,*
les matrices *unitaires,*
mais aussi d'autres matrices, comme par exemple :

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

Remarque. – – La propriété $f^* \circ f = f \circ f^*$ est équivalente à

$$\langle\, f(x) \,|\, f(y)\,\rangle \;=\; \langle\, f^*(x) \,|\, f^*(y)\,\rangle\,, \quad \forall\, x, y \in E$$

Le théorème suivant affirme que tout endomorphisme normal (et donc toute matrice normale) est diagonalisable (les valeurs propres, sauf dans le cas hermitien, ne sont pas forcement réelles).

Théorème 8.18 – *Soit f un endomorphisme normal d'un espace hermitien.*
Alors f est diagonalisable et les sous-espaces propres sont deux à deux orthogonaux (en particulier on peut construire une base orthonormée de vecteurs propres en prenant une base orthonormée dans chaque espace propre).

Ou, en termes de matrices :

Toute matrice normale est diagonalisable dans $\mathbb{C}$ et les espaces propres sont deux à deux orthogonaux pour le produit scalaire hermitien canonique de $\mathbb{C}^n$.

Ou encore :

Soit A une matrice normale. Il existe alors une matrice unitaire U, telle que la matrice $A' = {}^t\overline{U}\,A\,U$ soit diagonale (non nécessairement réelle).

Démonstration : Par récurrence sur $n = \dim E$.
Pour $n = 1$ il n'y a rien à démontrer. Supposons le théorème vrai jusqu'à l'ordre $n-1$ et montrons qu'il est vrai à l'ordre n.
Soit λ une valeur propre, E_λ l'espace propre correspondant ; on a :

$$E = E_\lambda \oplus E_\lambda^\perp$$

Notons que $\dim E_\lambda^\perp \leq n-1$.
La démonstration suit les étapes de la démonstration du théorème 7.38. On montre d'abord que $E_\lambda^\perp$ est stable par f, puis que la restriction de f à $E_\lambda^\perp$ est un endomorphisme normal.

a) $E_\lambda^\perp$ est stable par f.

Notons d'abord que E_λ *est stable par* f^*. En effet, soit $x \in E_\lambda$. On a : $f \circ f^*(x) = f^* \circ f(x) = \lambda f^*(x)$, ce qui montre que $f^*(x) \in E_\lambda$.
Soit maintenant $y \in E_\lambda^\perp$. On a, pour tout $x \in E_\lambda$:

$$\langle f(y) \,|\, x\rangle \;=\; \langle x \,|\, f^*(y)\rangle = 0\,, \;\; \text{puisque } y \in E_\lambda^\perp \;\text{ et } f^*(x) \in E_\lambda\,;$$

ce qui veut dire que $f(y) \in E_\lambda^\perp$ et que donc $E_\lambda^\perp$ est stable par f.

b) La restriction de f à $E_\lambda^\perp$ est un endomorphisme normal.

Notons tout d'abord que $E_\lambda^\perp$ *est stable par* f^*. En effet, soit $x \in E_\lambda^\perp$; pour tout $y \in E_\lambda$, on a :

$$\langle\, f^*(x) \,|\, y\rangle \;=\; \langle x \,|\, f(y)\rangle \;=\; \lambda\, \langle x \,|\, y\rangle = 0\,;$$

ce qui montre que $f^*(x) \in E_\lambda^\perp$ et que donc $E_\lambda^\perp$ est stable par f^*.

Considérons maintenant l'identité $\langle f(x) \,|\, f(y)\rangle \;=\; \langle f^*(x) \,|\, f^*(y)\rangle$ (cf. remarque page 286). Puisque, comme on vient de le voir, $E_\lambda^\perp$ est stable par f et par f^*, on a :

$$\langle\, f(x) \,|\, f(y)\rangle \,\big|_{E_\lambda^\perp} = \langle\, f^*(x) \,|\, f^*(y)\rangle_{E_\lambda^\perp} \qquad \forall x, y \in E_\lambda^\perp.$$

ce qui montre que que la restriction de f à $E_\lambda^\perp$ est normale.

Le théorème se démontre maintenant facilement. $E_\lambda^\perp$ étant de dimension $\leq n-1$, la restriction $\tilde{f}$ de f à $E_\lambda^\perp$ est diagonalisable, d'après l'hypothèse de récurrence. Donc, $E_\lambda^\perp$ est somme directe de sous-espaces propres de $\tilde{f}$ qui sont, toujours d'après l'hypothèse de récurrence, deux à deux orthogonaux. Comme les sous-espaces propres de $\tilde{f}$ sont des sous-espaces propres de f, et, par ailleurs, $E = E_\lambda \oplus E_\lambda^\perp$, il est clair que E est somme directe de sous-espaces propres deux à deux orthogonaux. Le théorème est démontré. $\square$

Le théorème 8.16 permet de démontrer un résultat analogue à celui de 7.39 page 254 :

Proposition 8.19 – *Soit h une forme hermitienne sur un espace vectoriel de dimension finie sur $\mathbb{C}$, $\{e_i\}$ une base de E et $H = M(h)_{e_i}$. Alors h définit un produit scalaire hermitien si et seulement si la matrice H a toutes ses valeurs propres strictement positives.*

La démonstration est analogue à celle de la proposition 7.39 et elle est laissée en exercice.

Exercices 13. 14. 15.16.

EXERCICES

1 Vérifier que les applications suivantes sont antilinéaires :

a) $f : \begin{array}{ccc} \mathbb{C}^3 & \longrightarrow & \mathbb{C} \\ (x_1, x_2, x_3) & \longmapsto & \overline{x}_1 + 2\,\overline{x}_2 - \overline{x}_3 \end{array}$ *b)* $f : \begin{array}{ccc} \mathcal{M}_n(\mathbb{C}) & \longrightarrow & \mathbb{C} \\ A & \longmapsto & \operatorname{Tr} \overline{A} \end{array}$

2 Les applications suivantes sont-elles des formes hermitiennes ?

a) $h_1 \ : \ \mathbb{C}^3 \times \mathbb{C}^3 \to \mathbb{C}$ définie par :

$$\begin{aligned} h_1\,(x,y) = \ & \overline{x}_1\,y_1 + 3\,\overline{x}_2\,y_2 + 2\,i\,\overline{x}_3\,y_3 + (2+3\,i)\,\overline{x}_1\,y_2 + (2-3\,i)\,\overline{x}_2\,y_1 \\ & + (1-5\,i)\,\overline{x}_2\,y_3 + (1+5\,i)\,\overline{x}_3\,y_2 \end{aligned}$$

b) $h_2 \ : \underset{(A,B)}{\mathcal{M}_n(\mathbb{C}) \times \mathcal{M}_n(\mathbb{C})} \underset{\longmapsto}{\longrightarrow} \underset{\mathrm{Tr}\ \overline{A}\, B}{\mathbb{C}}$

3 Montrer que si h est une forme hermitienne et $\tilde{q}(x) := h(x,x)$, alors :

$$h(x,y) = \frac{1}{4}\Big[\,\tilde{q}(x+y) - \tilde{q}(x-y) - i\,\tilde{q}(x+iy) + i\,\tilde{q}(x-iy)\Big]$$

(ce qui montre que h est déterminée par la connaissance de $\tilde{q}$).

4 Soit h la forme hermitienne sur $\mathbb{C}^3$ définie dans la base canonique $\{e_i\}$ par :

$$\tilde{q}(x) = |x_1|^2 + 3\,|x_2|^2 + 6\,|x_3|^2 + i\,\overline{x}_1\,x_2 - i\,x_1\,\overline{x}_2 + 2\,i\,\overline{x}_2\,x_3 - 2\,i\,\overline{x}_3\,x_2 .$$

S'agit-il d'un produit scalaire hermitien ?

* **5** Soit f un endomorphisme d'un espace hermitien. Montrer que

$$\langle f(x)\,|\,x\rangle = 0\,,\ \forall\, x \in E \iff f = 0.$$

6 Montrer que la forme hermitienne de l'exercice 2 b) est un produit scalaire hermitien. Exprimer la norme de la matrice A en fonction de ses coefficients.

7 Soit $\mathcal{H} = \{$matrices hermitiennes d'ordre 2 à trace nulle$\}$. Montrer que $\mathcal{H}$ est un espace vectoriel sur $\mathbb{R}$. En donner une base.
Montrer que $\|\ A\ \| := \sqrt{-\mathrm{dét} A}$ est une norme sur $\mathcal{H}$.

8 On suppose $\mathbb{C}^3$ muni du produit scalaire hermitien de l'exercice 4.

1. Déterminer une base orthonormée.
2. Soit F le sous-espace vectoriel de $\mathbb{C}^3$ défini par $x_1 + i\,x_2 = 0$. Construire par la méthode de Schmidt une base orthonormée de F.

9 Soit E l'espace vectoriel des fonctions polynômes de la variable réelle x, à coefficients complexes ($P \ : \ \mathbb{R} \longrightarrow \mathbb{C}$), de dégré ≤ 2. On définit :

$$\tilde{q}(P) = \int_0^1 \ |P(x)|^2 \,\mathrm{d}x$$

1. Montrer que $\tilde{q}$ définit une structure d'espace hermitien.
2. Déterminer une base orthonormée du sous-espace vectoriel F formé par les polynômes P de la forme :
$$P = 2\,\mu - \lambda + 3\,i\,\mu\,x^2 \qquad (\lambda, \mu \in \mathbb{C})\ .$$

10 Soit $(E, \langle\,|\,\rangle)$ un espace hermitien et $f \in \mathrm{End}_{\,\mathbb{C}}(E)$.
Soit λ une valeur propre de f ; montrer que :

1. si $f^* = f^{-1}$ alors $|\lambda| = 1$;
2. si $f^* = f$ alors $\lambda \in \mathbb{R}$;
3. si $f^* = -f$ alors λ est un imaginaire pur ;
4. s'il existe un endomorphisme g de E tel que $f = g^* \circ g$, alors λ est un réel positif ou nul.

* **11** Déterminer toutes les matrices unitaires d'ordre 2 qui transforment le vecteur $v_1 = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ i \end{pmatrix}$ en le vecteur $w_1 = \frac{1}{2}\begin{pmatrix} 1+i \\ 1-i \end{pmatrix}$.

** **12** **Quaternions. Paramétrisation de $\mathbf{SU(2,\mathbb{C})}$ par la sphère unité de $\mathbb{R}^4$**

On munit $\mathbb{R}^3$ du produit scalaire canonique, noté $\langle\ ,\ \rangle$ et $\mathbb{C}^2$ du produit hermitien canonique $\langle z \mid w\rangle = \overline{z}_1\, w_1 + \overline{z}_2\, w_2$. On note :

$$\mathcal{A} = \{M \in \mathcal{M}_2(\mathbb{C}) \mid M^* = -M,\ \operatorname{Tr} M = 0\}$$

(matrices antihermitiennes à trace nulle).

1. Montrer que $\mathcal{A}$ est peut être vu comme espace vectoriel de dimension 3 sur $\mathbb{R}$ et que les trois matrices de $\mathcal{M}_2(\mathbb{C})$:

$$\mathcal{I} = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} \quad \mathcal{J} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \quad \mathcal{K} = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}$$

forment une base de cet espace.
Montrer que $\mathbb{R}^4$ est isomorphe à $\tilde{H} := \mathcal{A} \oplus \operatorname{Vect}_{\mathbb{R}}\{I\}$, $\operatorname{Vect}_{\mathbb{R}}\{I\}$ désignant l'espace vectoriel sur $\mathbb{R}$ engendré par la matrice $I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$. $\tilde{H}$ est dit représentation matricielle du corps des ***quaternions*** (cf. Appendice A.1).

2. Soit $A \in \operatorname{SU}(2,\mathbb{C})$.
   (a) Montrer que les valeurs propres λ_1, λ_2 de A sont complexes conjuguées et en déduire que $\operatorname{Tr} A \in \mathbb{R}$.
   (b) A l'aide du théorème de Cayley-Hamilton, montrer que $A^* = (\operatorname{Tr} A)\, I - A$. En déduire que la matrice $M = A - \frac{1}{2}(\operatorname{Tr} A)\, I$ est antihermitienne et à trace nulle, et que :
   $$\operatorname{SU}(2,\mathbb{C}) \ \subset\ \tilde{H} \simeq \mathbb{R}^4$$
   (c) Montrer que si $A \in \operatorname{SU}(2,\mathbb{C})$, alors A s'écrit d'une manière unique :
   $$A = x_1\, I + x_2\, \mathcal{I} + x_3\, J + x_4\, K$$
   avec $(x_1, x_2, x_3, x_4) \in \mathbb{R}^4$ tels que :
   $$x_1^2 + x_2^2 + x_3^2 + x_4^2 = 1$$

13 Soit E un espace hermitien. Montrer qu'un projecteur est orthogonal si et seulement si il est autoadjoint.

* **14**
1. On note $\mathcal{H}$ l'ensemble des endomorphismes *hermitiens* (c'est-à-dire autoadjoints) d'un espace hermitien et $\mathcal{H}^+$ l'ensemble des endomorphismes hermitiens dont les valeurs propres sont dans $\mathbb{R}^+$.
   (a) Montrer que : $f \in \mathcal{H} \iff \langle f(x) \mid x\rangle \in \mathbb{R}$, $\forall x \in E$.
   (b) Montrer que : $f \in \mathcal{H}^+ \iff \langle f(x) \mid x\rangle \in \mathbb{R}_+$, $\forall x \in E$.
   (c) Soit $f \in \mathcal{H}^+$. Montrer que : $\langle f(x) \mid x\rangle = 0 \implies x \in \operatorname{Ker} f$.
2. (a) Soient $f, g \in \mathcal{H}^+$. On dit que $f < g$ si $g - f \in \mathcal{H}^+$. Montrer que si $f < g$, alors :
   $$\operatorname{Ker} g \subset \operatorname{Ker} f, \qquad \operatorname{Im} f \subset \operatorname{Im} g.$$
   (b) Montrer que si $f,\ g \in \mathcal{H}^+$, alors : $f + g \in \mathcal{H}^+$ et $f < f + g$. En déduire que :
   $$\operatorname{Im}(f+g) = \operatorname{Im} f + \operatorname{Im} g \quad \text{et} \quad \operatorname{Ker}(f+g) = \operatorname{Ker} f \cap \operatorname{Ker} g.$$

15 Soit $A \in \mathcal{M}_n(\mathbb{R})$ antisymétrique.

1. Montrer que $B = iA$ est hermitienne ; en déduire que toutes les valeurs propres de A sont imaginaires pures.
2. En utilisant le fait que A est normale, montrer le résultat suivant :
   Toute matrice antisymétrique réelle est de rang pair.

* **16** (cf. exercice 28 du chapitre 7).
Soit $A \in O(n, \mathbb{R})$.

1. Montrer que A est diagonalisable dans une base orthonormé de $\mathbb{C}^n$ et que

$$\mathrm{Sp}'_A\,\mathbb{C} = \left\{ e^{i\theta_1}, e^{-i\theta_1}, \ldots, e^{i\theta_p}, e^{-i\theta_p}, \underbrace{1, \ldots, 1}_{r}, \underbrace{-1, \ldots, -1}_{s}, \right\}$$

2. Soit $\mathcal{B}' = \{w_1, \overline{w_1}, \ldots w_p, \overline{w_p}, u_1, \ldots, u_r, v_1, \ldots, v_s\}$ une base orthonormée de $\mathbb{C}^n$ formée de vecteurs propres correspondants au spectre ci-dessus. On considère la famille $\mathcal{B}' = \{w'_1, \ldots, w'_{2p}, u_1, \ldots, u_r, v_1, \ldots v_s\}$, où :

$$\left\{ \begin{array}{ll} w'_1 = \dfrac{1}{\sqrt{2}}(w_1 + \overline{w_1}), & w'_2 = \dfrac{i}{\sqrt{2}}(w_1 - \overline{w_1}) \\ w'_3 = \dfrac{1}{\sqrt{2}}(w_2 + \overline{w_2}), & w'_4 = \dfrac{i}{\sqrt{2}}(w_2 - \overline{w_2}) \\ \cdots\cdots & \cdots\cdots \\ w'_{2p-1} = \dfrac{1}{\sqrt{2}}(w_p + \overline{w_p}), & w'_{2p} = \dfrac{i}{\sqrt{2}}(w_p - \overline{w_p}) \end{array} \right.$$

Montrer qu'il s'agit d'une base orthonormée de $\mathbb{R}^n$ et que la matrice de f dans cette base $\mathcal{B}'$ est

$$M(f)_{\mathcal{B}'} = \begin{pmatrix} \boxed{\begin{matrix} \cos\theta_1 & -\sin\theta_1 \\ \sin\theta_1 & \cos\theta_1 \end{matrix}} & & & & \\ & \ddots & & & \\ & & \boxed{\begin{matrix} \cos\theta_p & -\sin\theta_p \\ \sin\theta_p & \cos\theta_p \end{matrix}} & & \\ & & & \boxed{\begin{matrix} 1 & & \\ & \ddots & \\ & & 1 \end{matrix}} & \\ & & & & \boxed{\begin{matrix} -1 & & \\ & \ddots & \\ & & -1 \end{matrix}} \end{pmatrix}$$

INDICATIONS

1 Simples vérifications

2 f_1 : non (elle ne vérifie pas la propriété de symétrie hermitienne.)
f_2 : oui

3 Exprimer le second membre à l'aide de h, par exemple :

$$\tilde{q}\,(x+y) = h(x+y\,,\,x+y\,) = h(x,x) + h(x,y\,) + h(y,x) + h(y,y\,)$$

etc.

4 $\tilde{q}\,(x) = |x_1 + i\,x_2|^2 + 2\,|x_2 - i\,x_3|^2 + 4\,|x_3|^2$.

5 Développer $\langle f(y+z)\,|\,y+z\rangle$ et $\langle f(y+iz)\,|\,y+iz\rangle$ et en déduire que :

$$\langle f(x)\,|\,x\rangle = 0 \ ,\forall x \in E \iff \begin{cases} \quad\ \langle f(y)\,|\,z\,\rangle + \ \langle f(z)\,|\,y\,\rangle &= 0 \\ -i\,\langle f(y)\,|\,z\,\rangle + i\,\langle f(z)\,|\,y\,\rangle &= 0 \ \ \forall y,z \in E \end{cases}$$

6 Soit $\{E_{ij}\}$ la base canonique de $\mathcal{M}_n(\mathbb{C})$. Si $A = \Big(x_{ij}\Big) = \sum_{i,j=1}^n x_{\,ij}\,E_{ij}$ et $B = \Big(y_{ij}\Big) = \sum_{i,j=1}^n y_{\,ij}\,E_{ij}$, on a :

$$h(A,B) = \overline{x}_{11}\,y_{11} + \overline{x}_{12}\,y_{12} + \cdots + \overline{x}_{nn}\,y_{nn} = \textstyle\sum_{i,j=1}^n \overline{x}_{ij}\,y_{ij}$$

Il s'agit donc du produit scalaire hermitien associé à la base $\{E_{ij}\}$ (cf. exemple 2. page 275)

$$\parallel A \parallel = \sum_{i,j=1}^{n} |\,x_{ij}|^2$$

7 $A \in \mathcal{H} \iff A = \begin{pmatrix} x_1 & x_2 + i\,x_3 \\ x_2 - i\,x_3 & -x_1 \end{pmatrix}$ avec $x_1, x_2, x_3 \in \mathbb{R}$

$$\iff A = x_1 \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} + x_2 \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + x_3 \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix}$$

Les trois matrices mises en évidence forment une base de $\mathcal{H}$. On a :

$$\parallel A \parallel = \sqrt{-\text{dét} A} = \sqrt{x_1^2 + x_2^2 + x_3^2}.$$

Il s'agit donc de la norme associée au produit scalaire euclidien : $(x,y) \mapsto x_1\,y_1 + x_2\,y_2 + x_3\,y_3$.

8

1. En appliquant la méthode de Schmidt à la base canonique de $\mathbb{C}^3$, on trouve la base orthonormée

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \quad e_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} -i \\ 1 \\ 0 \end{pmatrix} \quad e_3 = \frac{1}{2} \begin{pmatrix} -1 \\ -i \\ 1 \end{pmatrix}$$

2. Une base de F est, par exemple, $\{v_1, v_2\}$ avec $v_1 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$ et $v_2 = \begin{pmatrix} 1 \\ i \\ 0 \end{pmatrix}$. En appliquant le procédé de Schmidt, on trouve la base orthonormée de F :

$$e'_1 = \frac{1}{\sqrt{6}}\,v_1 \quad , \quad e'_2 = \frac{1}{\sqrt{24}} \begin{pmatrix} 3 \\ 3\,i \\ 1 \end{pmatrix}.$$

9

1. La forme hermitienne est :

$$h(P,Q) = \int_0^1 \overline{P(x)}\,Q(x)\,\mathrm{d}x$$

2. Une base de F est, par exemple : $\{v_1 = 1\ ,\ v_2 = 2 + 3i\,x^2\}$. En appliquant la méthode de Schmidt à cette base, on trouve la base orthonormée $\{\ e_1 = 1\ ,\ e_2 = \dfrac{\sqrt{5}}{2}\,(-1 + 3\,x^2)\ \}$.

10 1. $\| x \|^2 = \langle f^* \circ f(x) \mid x \rangle = \langle f(x) \mid f(x) \rangle = |\lambda|^2 \| x \|^2, \ldots$
2. 3. démonstrations analogues.
4. $\lambda = \dfrac{\| g(x) \|^2}{\| x \|^2}$.

11 Choisir un vecteur v_2 tel que $\{v_1, v_2\}$ soit une base orthonormée ; l'image de v_2 doit appartenir à $\mathrm{Vect}\{w_1\}^\perp$ et être normée. Revenir ensuite à la base canonique de $\mathbb{C}^2$.
Par exemple, en prenant $v_2 = \dfrac{1}{\sqrt{2}}\begin{pmatrix} i \\ 1 \end{pmatrix}$, on doit avoir $f(v_2) \equiv w_2 = \dfrac{e^{i\varphi}}{\sqrt{2}}\begin{pmatrix} i \\ -1 \end{pmatrix}$ avec $\varphi \in \mathbb{R}$. On trouve :

$$A = \frac{1}{2}\begin{pmatrix} \dfrac{1+i}{\sqrt{2}} + e^{i\varphi} & \dfrac{1-i}{\sqrt{2}} + i\, e^{i\varphi} \\ \dfrac{1-i}{\sqrt{2}} + i\, e^{i\varphi} & -\dfrac{(1+i)}{\sqrt{2}} - e^{i\varphi} \end{pmatrix}$$

12 1. Les matrices $A \in \mathcal{A}$ sont les matrices de la forme

$$A = \begin{pmatrix} i\,x, & y + i\,z \\ -y + i\,z & -i\,x \end{pmatrix} = x\,\mathcal{I} + y\,\mathcal{J} + z\,\mathcal{K}$$

avec x, y, $z \in \mathbb{R}$. Ce qui montre que la famille $\{\mathcal{I}, \mathcal{J}, \mathcal{K}\}$ engendre $\mathcal{A}$. Puisqu'elle est $\mathbb{R}-$libre elle est une base de $\mathcal{A}$.
On a $\mathrm{Vect}_{\mathbb{R}}\{I\} \cap \mathcal{A} = \{0\}$, donc $\mathrm{Vect}_{\mathbb{R}}\{I\}$ et $\mathcal{A}$ sont en somme directe et $\dim \mathcal{A} \oplus \mathrm{Vect}_{\mathbb{R}}\{I\} = 4$. Aussi $\mathbb{R}^4$ est isomorphe à $\tilde{\mathcal{H}} = \mathcal{A} \oplus \mathrm{Vect}_{\mathbb{R}}\{I\}$. On peut construire un isomorphisme canonique ψ, en imposant que l'image par ψ de la base canonique de $\mathbb{R}^4$ est $\{I, \mathcal{I}, \mathcal{J}, \mathcal{K}\}$ et en prolongeant par linéarité.

2. (a) Puisque A est unitaire, $|\lambda_1| = |\lambda_2| = 1$ (cf. exercice 10, 1). D'autre part $\det A = 1$, donc $\lambda_1 \lambda_2 = 1$, d'ou : $\lambda_2 = \overline{\lambda_1}$ et par conséquent : $\mathrm{Tr}\, A = \lambda_1 + \lambda_2 \in \mathbb{R}$.
   (b) On a : $P_A(X) = X^2 - \mathrm{Tr}(A)\, X + \det A = 1$, d'où : $A^2 - (\mathrm{Tr}\, A)\, A + I = 0$. Multiplier à gauche par A^*.
   On a : $\mathrm{Tr}(M) = 0$ et, par ailleurs, $M^* = A^* - \frac{\mathrm{Tr}\, A}{2}\, I$, car $\mathrm{Tr}\, A \in \mathbb{R}$, d'où $M^* = -M$.
   Pour toute $A \in \mathrm{SU}(2, \mathbb{C})$, on peut écrire : $A = (A - \frac{\mathrm{Tr}\, A}{2}) + \frac{\mathrm{Tr}\, A}{2}$.
   (c) Soit $A \in \mathrm{SU}(2, \mathbb{C})$. D'après 2. b) et 1., on peut écrire $A = x_1\, I + x_2\, \mathcal{I} + x_3\, \mathcal{J} + x_4\, \mathcal{K}$. La condition $A^* A = I$ est équivalente à $x_1^2 + x_2^2 + x_3^2 + x_4^2 = 1$.

13 Même démonstration que pour l'exercice 16 du chapitre 7.

14 1. (a) Si $f \in \mathcal{H}$, on a $\langle f(x) \mid x \rangle = \langle x \mid f(x) \rangle$. D'autre part $\langle x \mid f(x) \rangle = \overline{\langle f(x) \mid x \rangle}$, donc $\langle f(x) \mid x \rangle \in \mathbb{R}$ pour tout $x \in E$. Réciproquement, on suppose que $\langle f(x) \mid x \rangle \in \mathbb{R}$ pour tout $x \in E$. En développant, comme dans l'exercice 5, $\langle f(y+z) \mid y+z \rangle$ et $\langle f(y+iz) \mid y+iz \rangle$, montrer que :

$$\langle f(y) \mid z \rangle + \langle f(z) \mid y \rangle \in \mathbb{R} \quad \text{et} \quad \langle f(y) \mid z \rangle - \langle f(z) \mid y \rangle \in i\,\mathbb{R}$$

En déduire que

$$\begin{cases} \langle f(z) \mid y \rangle + \langle f(y) \mid z \rangle = \ \ \langle y \mid f(z) \rangle + \langle z \mid f(y) \rangle \\ \langle f(z) \mid y \rangle - \langle f(y) \mid z \rangle = -\,\langle y \mid f(z) \rangle + \langle z \mid f(y) \rangle \end{cases}$$

et que $f \in \mathcal{H}$.
   (b) Si $\langle f(x) \mid x \rangle \in R_+$ pour tout $x \in E$, alors pour v vecteur propre de f : $\langle f(v) \mid v \rangle \in R_+$, c'est-à-dire $\langle \lambda v \mid v \rangle \in \mathbb{R}_+$, d'où $\lambda \in \mathbb{R}_+$, compte tenu du fait que λ est réel.
   Si $f \in \mathcal{H}_+$, considérer une base orthonormée $\{v_1, \ldots, v_n\}$ de vecteurs propres. Calculer $\langle f(x) \mid x \rangle$ en décomposant x sur cette base.

(c) En décomposant x sur une base orthonormeé de vecteurs propres, $x = \sum_{i=1}^{n} x_i v_i$, on a : $\langle f(x) \,|\, x\rangle = \sum_{i=1}^{n} |x_i|^2 \lambda_i = 0$. Or $f \in \mathcal{H}_+$, donc, pour tout i, $\lambda_i \geq 0$ et par conséquent $\lambda_i x_i = 0$ pour tout i. On a donc : $f(x) = \sum_{i=1}^{n} x_i \lambda_i v_i = 0$.

2. (a) Soit $x \in \operatorname{Ker} g$. Puisque $g - f \in \mathcal{H}^+$, on a : $\langle (g-f)(x) \,|\, x\rangle \geq 0$, d'où : $0 = \langle g(x) \,|\, x\rangle \geq \langle f(x) \,|\, x\rangle \geq 0$ (car $f \in \mathcal{H}^+$), donc $\langle f(x) \,|\, x\rangle = 0$, ce qui implique $x \in \operatorname{Ker} f$, d'après 1.(c), donc $\operatorname{Ker} g \subset \operatorname{Ker} f$.
Pour montrer la propriété sur les images, utiliser l'identité $(\operatorname{Ker} f)^{\perp} = \operatorname{Im} f^*$, qui se démontre de la même manière que la propriété analogue de l'exercice 21 du chapitre 7.

(b) Utiliser 1.(b) pour montrer que $f + g \in \mathcal{H}^+$ et $f < f + g$.
Puisque $f < f + g$, $\operatorname{Im} f \subset \operatorname{Im}(f+g)$, d'après 2. (a). De même $\operatorname{Im} g \subset \operatorname{Im}(f+g)$, d'où $\operatorname{Im} f + \operatorname{Im} g \subset \operatorname{Im}(f+g)$. L'inclusion inverse est toujours vraie.
Même démonstration pour l'égalité sur les noyaux.

15

1. Utiliser le fait que valeurs propres d'une matrice hermitienne sont réelles.
2. Puisque A est réelle, si λ est valeur propre, $\overline{\lambda}$ est aussi valeur propre. Compte tenu de 1, les valeurs propres non nulles sont en nombre pair. Puisque A est diagonalisable (dans $\mathbb{C}$), $\operatorname{rg} A =$ nombre des valeurs propres non nulles de A.

16

1. A est normale, donc diagonalisable dans $\mathbb{C}$, dans une base orhonormée (cf. Théorème 8.18). Ses valeurs propres sont de module 1 (cf. Remarque page 284). Puisque A est réelle, ses valeurs propres non réelles sont complexes conjuguées (cf exercice 7, chapitre 6).
2. Simple vérification. Par exemple, calculons le premier bloc. On a :
$w_1 = \dfrac{w'_1 - i\, w'_2}{\sqrt{2}}$, $\quad \overline{w}_1 = \dfrac{w'_1 + i\, w'_2}{\sqrt{2}}$. Donc :
$$\begin{aligned} Aw'_1 &= \frac{1}{\sqrt{2}}(Aw_1 + A\overline{w_1}) = \frac{1}{\sqrt{2}}(e^{i\theta} w_1 + e^{-i\theta}\overline{w_1}) = \frac{1}{2}\Big[e^{i\theta}(w'_1 - iw'_2) + e^{-i\theta}(w'_1 + iw'_2)\Big] \\ &= \frac{e^{i\theta} + e^{-i\theta}}{2} - i\frac{e^{i\theta} - e^{-i\theta}}{2} = \cos\theta\, w'_1 + \sin\theta\, w'_2 \end{aligned}$$
ce qui donne la première colonne. De même on trouve $Aw'_2 = -\sin\theta\, w'_1 + \cos\theta\, w'_2$.

Chapitre 9

Formes bilinéaires et formes quadratiques

Le cadre de la théorie de la Relativité restreinte est l'espace vectoriel $\mathbb{R}^4$ dans lequel le produit scalaire est remplacé par la forme bilinéaire

$$f : \mathbb{R}^4 \times \mathbb{R}^4 \longrightarrow \mathbb{R}$$

définie par :

$$f(x,y) = x_1 y_1 + x_2 y_2 + x_3 y_3 - x_4 y_4$$

On voit immédiatement que f n'est pas définie positive, car l'expression

$$f(x,x) = x_1^2 + x_2^2 + x_3^2 - x_4^2$$

peut être positive, négative ou même nulle.
Il s'agit là d'une généralisation du produit scalaire et l'on doit s'attendre à ce que certaines propriétés soient conservées et d'autres non. Par exemple, il peut exister des vecteurs non nuls à "norme nulle" ; on ne peut pas, *a priori*, normaliser une base orthogonale (on risquerait de diviser par 0...). La notion d'orthogonalité devient, en général, plus délicate à manier (noter, par exemple, qu'un vecteur de "norme nulle", c'est-à-dire tel que $f(x,x) = 0$ est orthogonal à lui-même) ; le groupe orthogonal - c'est-à-dire l'anologue du groupe des isométries vectorielles - a, bien entendu un aspect différent (dans le cas de la Relativité on trouve, au lieu des rotations et des symétries, ce que l'on appelle les «transformations de Lorentz» (cf. exercice 26), etc.
Le but de ce chapitre est d'étudier les formes bilinéaires dans ce contexte général où elles ne sont pas nécessairement définies positives. L'intérêt ressort, entre autres, de l'exemple ci-dessus : l'application à la théorie de la Relativité.

9.1 Rang et noyau d'une forme bilinéaire

Rappelons (cf. Définition 7.2) qu'une forme bilinéaire sur un espace vectoriel E sur K est une application $b : E \times E \longrightarrow K$ qui est linéaire dans les deux arguments.

Commençons par considérer le cas où E est de dimension finie. Si $\{e_i\}$ est une base de E, on note :

$$B = M(b)_{e_i} = \begin{pmatrix} b(e_1,e_1) & \dots & b(e_1,e_n) \\ \vdots & & \vdots \\ b(e_n,e_1) & \dots & b(e_n,e_n) \end{pmatrix}$$

la matrice de b dans la base $\{e_i\}$ (cf. définition 7.11 page 233). Rappelons que si $X = M(x)_{e_i}$ et $Y = M(y)_{e_i}$, on a : $b(x,y) = {}^t X\,B\,Y$.
Comme nous l'avons vu (cf. page 235), si $\{e'_i\}$ est une autre base, $B' = M(b)_{e'_i}$ et $P = P_{e_i \to e'_i}$, on a : $B' = {}^tP\,B\,P$ et, puisque P est inversible : $\operatorname{rg} B' = \operatorname{rg} B$ (cf. chapitre 3, exercice 8). On peut donc poser la définition suivante :

Définition 9.1 – *Soit b une forme bilinéaire sur un espace vectoriel E de dimension finie. On appelle **rang de** b le rang de la matrice qui représente b dans une base (quelconque) :*

$$\operatorname{rg}(b) = \operatorname{rg} M(b)_{e_i}$$

*b est dite **non dégénérée** si le rang de b est maximum (c'est-à-dire : $\operatorname{rg}(b) = n = \dim E$). On a donc :*

$$\boxed{b \text{ est non dégénérée} \iff \begin{array}{c} \text{dét } M(b)_{e_i} \neq 0 \\ \{e_i\}\text{base quelconque.} \end{array}}$$

On notera que le déterminant de $M(b)_{e_i}$ dépend du choix de la base. On a en effet $\text{dét}\, B' = (\text{dét}\, P)^2 \text{dét}\, B$. Cependant, comme $\text{dét}\, P \neq 0$, on a $\text{dét}\, B' \neq 0$ si et seulement si $\text{dét}\, B \neq 0$; donc le fait que le déterminant soit nul ou non ne dépend pas du choix de la base.

Exemple 1 – Tout produit scalaire est non dégénéré. En effet, si $\{e_i\}$ est une base orthonormée :

$$M(\langle\ ,\ \rangle)_{e_i} = \begin{pmatrix} 1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1 \end{pmatrix} = I$$

donc $\operatorname{rg} \langle\ ,\ \rangle = n$.

Exemple 2 – La forme bilinéaire de la Relativité, dite *forme de Lorentz*, est aussi non dégénérée, car dans la base canonique de $\mathbb{R}^4$, on a :

$$M(f)_{e_i} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

donc $\operatorname{rg} f = 4$.

Exemple 3 – Soit $b : \mathbb{R}^3 \times \mathbb{R}^3 \longrightarrow \mathbb{R}$ qui dans la base canonique est définie par :

$$b(x,y) = x_1\,y_1 - 3\,x_3\,y_3 + x_1\,y_2 + x_2\,y_1 - x_1\,y_3 - x_3\,y_1 - 3\,x_3\,y_2 - 3\,x_2\,y_3$$

On a :

$$M(b)_{e_i} = \begin{pmatrix} \boxed{\begin{matrix} 1 & 1 \\ 1 & 0 \end{matrix}} & \begin{matrix} -1 \\ -3 \end{matrix} \\ \begin{matrix} -1 & -3 \end{matrix} & -3 \end{pmatrix}$$

$\text{dét}\ M(b)_{e_i} = 0$, donc s est dégénérée. D'autre part le mineur encadré est non nul, donc $\operatorname{rg} s = 2$.

NOTA. – *Le rang d'une forme bilinéaire $b : E \times E \longrightarrow K$* **n'est pas** *la dimension de* $\operatorname{Im} b$ *(qui d'ailleurs, étant incluse dans K ne pourrait être que de dimension 0 ou 1).*

On peut cependant interpréter le rang d'une forme bilinéaire comme la dimension de l'image d'une certaine application linéaire. Soit en effet $b : E \times E \longrightarrow K$ une forme bilinéaire et j l'application :

$$j : \begin{array}{ccc} E & \longrightarrow & E^* \\ y & \longmapsto & j(y) \end{array} \qquad \text{où} \quad j(y) : \begin{array}{ccc} E & \longrightarrow & \mathbb{R} \\ x & \longmapsto & b(x,y) \end{array}$$

En d'autres termes : $j(y)(x) := b(x,y)$. Nous avons vu (cf. proposition 7.15 page 237) que si b est un produit scalaire j est un isomorphisme d'espace vectoriel.

Proposition 9.2 – *Soit $\{e_i\}$ une base de E, $\{\varphi_i\}$ la base de E^* duale de $\{e_i\}$. On a :*

$$M(j)_{e_i,\varphi_j} = M(b)_{e_i}$$

En effet, soit

$$M(j)_{e_i,\varphi_j} = \| \, j(e_i), \dots, j(e_n) \, \|_{\varphi_i} = \begin{pmatrix} a_{11} & \cdots & a_{1i} & \cdots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \cdots & a_{ni} & \cdots & a_{nn} \end{pmatrix}$$

La i-ème colonne est costituée par les composantes de $j(e_i)$ dans la base $\{\varphi_{e_j}\}$. On a donc, d'après l'expression de la matrice : $j(e_i) = a_{1i}\varphi_1 + \cdots + a_{ni}\varphi_n$, donc :

$$j(e_i)(e_k) = a_{1i}\varphi_1(e_k) + \cdots + a_{ki}\varphi_k(e_k) + \cdots + a_{ni}\varphi_n(e_k) = a_{ki}$$

car $\varphi_j(e_k) = \delta_{jk}$. D'autre part, d'après la définition de b :

$$j(e_i)(e_k) = b(e_k, e_i) = b_{ki}$$

donc : $M(j)_{e_i,\varphi_j} = M(b)_{e_i}$ □

Ainsi le rang de b est égal au rang de j (au sens de la définition 3.3), c'est-à-dire égal à la dimension de $\operatorname{Im} j$.

La proposition 9.2 justifie la définition suivante, *qui est valable même si E n'est pas de dimension finie* :

Définition 9.3 – *Soit E un espace vectoriel sur K (non nécessairement de dimension finie) et $b : E \times E \longrightarrow K$ une forme bilinéaire.*

1. *On appelle* **rang de b** *le rang de l'application* $j : \begin{array}{ccc} E & \longrightarrow & E^* \\ y & \longmapsto & b(\cdot, y) \end{array}$.
2. *On appelle* **noyau de b** *- noté $N(b)$ - le noyau de l'application j, c'est-à-dire :*

$$N(b) := \big\{ y \in E \mid b(x,y) = 0 \ , \ \forall x \in E \big\}$$

3. *b est dite* **non dégénérée** *si j est injective, c'est-à-dire si : $N(b) = 0$, ou en d'autres termes, si :*

$$b(x,y) = 0 \, , \ \forall x \in E \implies y = 0$$

De la proposition 9.2 on déduit :

Proposition 9.4 – *Soit E un espace vectoriel* de dimension finie *et b une forme bilinéaire sur E. Alors :*

1. *Le noyau de b est le noyau de la matrice qui représente b (dans une base quelconque).*
2. $\boxed{\dim E = \operatorname{rg} b + \dim N(b).}$

Remarques. –

1. Bien que le noyau d'une forme bilinéaire se calcule comme le noyau de la matrice de b, la signification n'est pas la même. En particulier le noyau de b n'est pas l'ensemble $\{(x,y) \in E \times E \mid b(x,y) = 0\}$ (cet ensemble d'ailleurs n'est même pas un espace vectoriel).

2. Ne pas confondre $N(b)$ avec l'ensemble

$$\tilde{N}(b) = \Big\{x \in E \mid b(x,y) = 0\,,\ \forall\, y \in E\Big\}$$

qui est le noyau de la ***transposée*** de la matrice de b. En effet, en utilisant les notations matricielles :

$$\begin{aligned} x \in \tilde{N}(b) \quad &\Longleftrightarrow\ {}^tX\,B\,Y = 0\,,\ \forall Y \in \mathcal{M}_{n,1}(K) \Longleftrightarrow {}^tX\,B = 0 \Longleftrightarrow {}^tB\,X = 0 \\ &\Longleftrightarrow\ X \in \operatorname{Ker}{}^tB. \end{aligned}$$

$\tilde{N}(b)$ est dit *noyau à gauche de b*.

Il est clair que si b est symétrique, ce qui sera le cas dans la suite, alors $N(b) = \tilde{N}(b)$

Exemple 1 –

Bien entendu, pour un produit scalaire, comme pour toute forme non dégénérée, le noyau est réduit à $\{0\}$.

Exemple 2 –

Soit b la forme bilinéaire de l'exemple 3 ci-dessus (cf page 296). On peut écrire :

$$b(x,y) = (y_1 + y_2 - y_3)\,x_1 + (y_1 - 3y_3)\,x_2 + (-y_1 - 3y_2 - 3y_3)\,x_3$$

donc $y \in N(b)$ si et seulement si :

$$\left\{\begin{array}{rrrl} y_1 + & y_2 - & y_3 & = 0 \\ y_1 & & -\,3\,y_3 & = 0 \\ -y_1 & -\,3\,y_2 - & 3\,y_3 & = 0 \end{array}\right.$$

En résolvant, on trouve : $y_1 = 3\,\lambda$, $y_2 = -2\,\lambda$, $y_3 = \lambda$; donc $N(b)$ est engendré par le vecteur $y = \begin{pmatrix} 3 \\ -2 \\ 1 \end{pmatrix}$.

On aurait pu déterminer directement $N(b)$ à l'aide de la matrice de b :

$$B = \begin{pmatrix} 1 & 1 & -1 \\ 1 & 0 & -3 \\ -1 & -3 & -3 \end{pmatrix}$$

donc le noyau est donné justement par le système ci-dessus.

Exercices 1. 2. 3.

9.2 Formes bilinéaires symétriques et formes quadratiques en dimension finie

Par la suite nous nous intéressons aux formes bilinéaires s qui sont symétriques, c'est-à-dire telles que : $s(x,y) = s(y,x)\,,\ \forall\, x, y \in E$.
Il est clair que si E est de dimension finie, une forme bilinéaire s est symétrique si et seulement si la matrice de s (dans une base quelconque) est symétrique.
L'étude des formes bilinéaires symétriques est facilitée par l'introduction de la notion de forme quadratique.

Définition 9.5 – *Soit E un espace vectoriel de dimension* finie *sur K. Une application $q : E \longrightarrow K$ est dite* ***forme quadratique*** *sur E si, étant donnée une base $\{e_i\}$, $q(x)$ est un polynôme homogène de degré 2 en les composantes x_i de x dans la base $\{e_i\}$.*

Par exemple l'application $q : \mathbb{R}^3 \longrightarrow \mathbb{R}$ définie par
$$q(x) = 2\,x_1^2 - 3\,x_2^2 + 3\,x_3^2 + 2\,x_1\,x_2 - 3\,x_1\,x_3 + 5\,x_2\,x_3$$
(où les x_i sont les composantes de x dans la base canonique) est une forme quadratique.

Naturellement, pour que cette définition ait un sens, il faut vérifier que le fait d'être un polynôme homogène de degré 2 ne dépend pas du choix de la base. La vérification est facile. Soient $\{e_i\}$ et $\{e'_i\}$ deux bases de E et $x = \sum_{k=1}^n x_k e_k = \sum_{i=1}^n x'_i\, e'_i$ un vecteur de E. Supposons que :
$$q(x) = \sum_{i,j=1}^n a_{ij}\, x_i\, x_j .$$
Il s'agit de montrer qu'il existe des $a'_{ij} \in K$ tels que : $q(x) = \sum_{i,j=1}^n a'_{ij}\, x'_i\, x'_j$.
Soit $P = P_{e_i \to e'_i} = \Big(p_{ik}\Big)$ la matrice de passage : $e'_i = \sum_{k=1}^n p_{ki}\, e_k$. On a :
$$x = \sum_{i=1}^n x'_i\, e'_i = \sum_{i,k=1}^n x'_i\, p_{ki}\, x_k$$
Puisque $x = \sum_{k=1}^n x_k\, e_k$, on a $x_k = \sum_{i=1}^n p_{ki}\, x'_i$ (cf. proposition 3.25 page 77), donc
$$\begin{aligned} q(x) &= \sum_{k,l=1}^n a_{kl}\, x_k\, x_l = \sum_{i,j,k,l=1}^n a_{kl}\, p_{ki}\, x'_i\, p_{lj}\, x'_j \\ &= \textstyle\sum_{i,j=1}^n \underbrace{\Big(\displaystyle\sum_{k,l=1}^n a_{kl}\, p_{ki}\, p_{lj} \Big)}_{a'_{ij}} x'_i\, x'_j = \textstyle\sum_{i,j=1}^n a'_{ij}\, x'_i\, x'_j \end{aligned}$$
La définition de forme quadratique ne dépend donc pas du choix de la base.

Nous allons montrer maintenant qu'il existe une correspondance bijective entre les formes bilinéaires symétriques et les formes quadratiques.

Considérons une forme bilinéaire symétrique s sur E et soit l'application $q : E \longrightarrow K$ définie par $q(x) = s(x,x)$. Il est facile de voir que q est une forme quadratique.
En effet si $\{e_i\}$ est une base de E et $x = \sum_{i=1}^n x_i\, e_i$, $y = \sum_{j=1}^n y_j\, e_j$, on a :
$$\begin{aligned} s(x,y) = \ & \sum_{ij=1}^n a_{ij}\, x_i\, y_j = a_{11}\, x_1\, y_1 + \cdots + a_{nn}\, x_n\, y_n + a_{12}\, x_1\, y_2 + \\ & + a_{21}\, x_2\, y_1 + \cdots + a_{ij}\, x_i\, y_j + a_{ji}\, x_j\, y_i + \cdots \end{aligned}$$

(avec $a_{ij} = a_{ji}$) et donc :

$$q(x) = s(x,x) = a_{11}x_1^2 + \cdots + a_{nn}x_n^2 + 2\,a_{12}x_1x_2 + \cdots + 2\,a_{ij}x_ix_j + \cdots$$

qui est un polynôme homogène de degré 2 en les x_i. Donc q est une forme quadratique.

Par exemple, si :

$$s(x,y) = x_1\,y_1 - 3\,x_3\,y_3 + x_1\,y_2 + x_2\,y_1 - 5\,x_1\,y_3 - 5\,x_3\,y_1 + 3\,x_3\,y_2 + 3\,x_2\,y_3$$

On a :

$$q(x) = x_1^2 - 3\,x_3^2 + 2\,x_1\,x_2 - 10\,x_1\,x_3 + 6\,x_2\,x_3.$$

Réciproquement[1], à toute forme quadratique q est associée une *unique* forme bilinéaire symétrique s telle que $s(x,x) = q(x)\,,\ \forall x \in E$. Soit en effet $\{e_i\}$ une base de E et

$$q(x) = a_{11}x_1^2 + \cdots + a_{nn}x_n^2 + a_{12}x_1x_2 + \cdots + a_{ij}x_ix_j + \cdots$$

Il suffit d'appliquer ce que l'on appelle la ***règle de dédoublement des carrés*** (dite aussi ***opération de polarisation***) :

– aux termes "*carrés*" $a_{ii}\,x_i^2$ on associe $a_{ii}\,x_i\,y_i$

– aux termes "*rectangles*" $a_{ij}\,x_i\,x_j$; (avec $i \neq j$) on associe $\dfrac{1}{2}a_{ij}\,x_i\,y_j + \dfrac{1}{2}a_{ij}\,x_j\,y_i$.

Il est clair que l'on obtient ainsi une forme bilinéaire symétrique s telle que $s(x,x) = q(x)\ \forall x \in E$.
D'autre part, il ne peut exister qu'une seule forme bilinéaire symétrique s telle que $s(x,x) = q(x)\ \forall x \in E$. En effet, si s est une telle forme, on a :

$$\begin{aligned} q(x+y) &= s(x+y, x+y) = s(x,x) + s(x,y) + s(y,x) + s(y,y) \\ &= q(x) + 2\,s(x,y) + q(y) \end{aligned}$$

donc nécessairement :

$$s(x,y) = \frac{1}{2}\big[q(x+y) - q(x) - q(y)\big]\,,$$

ce qui montre que s est parfaitement déterminée par la connaissance de q.

Nous avons ainsi démontré :

Proposition 9.6 –

1. *Soit E un espace vectoriel de dimension finie sur K et $s : E \times E \longrightarrow K$ une forme bilinéaire symétrique. Alors l'application $q : E \longrightarrow K$ définie par :*

$$\boxed{q(x) = s(x,x)}$$

est une forme quadratique sur E.

[1] Les considérations qui suivent ne sont valables que pour les corps K *de caractéristique* $\neq 2$, c'est-à-dire $1_K + 1_K \neq 0$, 1_K étant le neutre de la multiplication C'est le cas bien entendu, de $\mathbb{R}$ et de $\mathbb{C}$. En revanche $\mathbb{Z}/2\mathbb{Z}$ est de caractéristique 2.

2. *Réciproquement, soit* $q : E \longrightarrow K$ *une forme quadratique sur* E *(*K *étant de caractéristique différente de 2), il existe alors une et une seule forme bilinéaire symétrique* $s : E \times E \longrightarrow K$ *telle que* $s(x,x) = q(x)$, $\forall x \in E$. s *est donnée par :*

$$\boxed{s(x,y) = \frac{1}{2}\big[q(x+y) - q(x) - q(y)\big]}$$

s *est dite* ***forme polaire de*** q.

Exemple –
Soit $q : \mathbb{R}^3 \longrightarrow \mathbb{R}$ définie dans la base canonique par :

$$q(x) = 3\,x_1^2 + 2\,x_2^2 - x_3^2 + 5\,x_1\,x_2 - 6\,x_1\,x_3 + 7\,x_2\,x_3$$

En polarisant on obtient :

$$s(x,y) = 3\,x_1\,y_1 + 2\,x_2\,y_2 - x_3\,y_3 + \frac{5}{2}\,x_1\,y_2 + \frac{5}{2}\,x_2\,y_1 - 3\,x_1\,y_3 - 3\,x_3\,y_1 + \frac{7}{2}\,x_2\,y_3 + \frac{7}{2}\,x_3\,y_2$$

Remarques

1. Si s est un produit scalaire, $q(x) = \| x \|^2$. $q(x)$ joue donc le même rôle que le carré de la norme pour le produit scalaire.
2. Puisqu'une forme quadratique est un polynôme homogène de degré 2 en les composantes de x, il est clair que l'on a :

$$q(\lambda x) = \lambda^2 q(x)\,, \quad \forall x \in E\,,\ \forall \lambda \in K$$

(on exprime cette propriété en disant que q est une application *homogène de degré 2*).

Il ne faut pas croire, cependant, que toute application homogène de degré 2 soit une forme quadratique. Par exemple l'application $q : \mathbb{R}^2 \longrightarrow \mathbb{R}$ définie par :

$$q(x) = \begin{cases} \dfrac{x_1^4 + x_2^4}{x_1^2 + x_2^2} & \text{si } x = (x_1, x_2) \neq (0,0) \\ 0 & \text{si } x = (0,0) \end{cases}$$

est homogène de degré 2, mais elle n'est pas un polynôme en les x_i (cf. exercice 6).

Exercices 4. 5. 6.

9.3 Définition de forme quadratique en dimension infinie

La définition 9.5 n'est valable qu'en dimension finie. Cependant la proposition 9.6 suggère la définition suivante :

Définition 9.7 *Soit* E *un espace vectoriel sur* K *(non nécessairement de dimension finie). Une application* $q : E \longrightarrow K$ *est dite* ***forme quadratique*** *s'il existe une application bilinéaire symétrique* $s : E \times E \longrightarrow K$, *telle que :* $s(x,x) = q(x)$.
Dans ce cas [2], s *est donnée par :*

$$s(x,y) = \frac{1}{2}\Big[q(x+y) - q(x) - q(y)\Big].$$

s *est dite forme polaire de* q.

[2] Ceci n'est valable que si la caractéristique de K est différente de 2.

Ainsi, pour montrer qu'une application $q : E \longrightarrow K$ est une forme quadratique il faut montrer que :

1. que l'application s définie par $s(x,y) := \frac{1}{2}\Big[q(x+y) - q(x) - q(y)\Big]$ est une forme bilinéaire ;
2. que $s(x,x) = q(x)$.

Exemple –

Soit $q : \mathbb{R}[x] \longrightarrow \mathbb{R}$ définie par : $q(P) = \int_0^1 P(x)^2\,\mathrm{d}x$. On a :

$$\begin{aligned} s(P,Q) &= \tfrac{1}{2}\Big[q(P+Q) - q(P) - q(Q)\Big]\,\mathrm{d}x \\ &= \tfrac{1}{2}\Big[\int_0^1 \Big(P(x)+Q(x)\Big)^2 \mathrm{d}x - \int_0^1 P(x)^2\,\mathrm{d}x - \int_0^1 Q(x)^2\,\mathrm{d}x\Big] \\ &= \int_0^1 P(x)Q(x)\,\mathrm{d}x \end{aligned}$$

En vertu de la linéarité de l'intégrale, s est une forme bilinéaire symétrique. Puisque $s(P,P) = q(P)$, q est une forme quadratique.

Exercice 7.

9.4 Rang, Noyau et vecteurs isotropes d'une forme quadratique

Puisque la donnée d'une forme bilinéaire symétrique est équivalente à la donnée d'une forme quadratique, les définitions sur les formes bilinéaires symétriques se transportent sur les formes quadratiques.

Définition 9.8 – *On appelle rang, noyau, matrice d'une forme quadratique q, le rang, le noyau, la matrice de la forme polaire associée à q.*
*q est dite **non dégénérée** si sa forme polaire s est non dégénérée, c'est-à-dire si $N(s) = \{0\}$, ce qui signifie :*

$$s(x,y) = 0\,,\ \forall y \in E \implies x = 0.$$

Si q est une forme quadratique à valeurs réelles*, q est dite **définie positive** si sa forme polaire est définie positive, c'est-à-dire si*

$$q(x) \geq 0\,,\ \forall x \in E \quad \text{et} \quad q(x) = 0 \Longleftrightarrow x = 0.$$ [3]

*Plus généralement, q est dité **définie** si*

$$q(x) = 0 \Longleftrightarrow x = 0$$

Remarque. –

1. Une forme quadratique définie positive est nécessairement non dégénérée.

   En effet si x est tel que $s(x,y) = 0$, $\forall y \in E$; on a en particulier $s(x,x) = 0$, c'est-à-dire $q(x) = 0$, donc $x = 0$.
2. Le noyau $N(q)$ de q n'est pas l'ensemble des vecteurs x tels que $q(x) = 0$, mais l'ensemble $N(q) := \{x \in E \mid s(x,y) = 0\,,\ \forall y \in E\}$. Les vecteurs x tels que $q(x) = 0$ sont dits ***isotropes***. Les formes quadratiques définies sont donc les formes quadratiques dont le seul vecteur isotrope est le vecteur nul.

[3] q est dite ***définie négative*** si $q(x) \leq 0\,, \forall x \in E$ et $q(x) = 0 \Longleftrightarrow x = 0$. Il est clair que q est définie négative si et seulement si $-q$ est définie positive

Exemple – Soit $q : \mathbb{R}^3 \longrightarrow \mathbb{R}$ la forme quadratique définie dans la base canonique par :

$$q(x) = 4\,x_1^2 + 3\,x_2^2 + 5\,x_1\,x_2 - 3\,x_1\,x_3 + 8\,x_2\,x_3$$

on a :

$$M(q)_{e_i} = \begin{pmatrix} 4 & 5/2 & -3/2 \\ 5/2 & +3 & 4 \\ -3/2 & 4 & 0 \end{pmatrix}$$

(attention à *diviser par 2 les coefficients des termes rectangles*). q est non dégénérée car dét $M(q)_{e_i} \neq 0$. D'autre part, $q(e_3) = s(e_3, e_3) = 0$, c'est-à-dire e_3 est isotrope (cf. dernier coefficient de la matrice), donc q n'est pas définie.

Définition 9.9 – *Soit q une forme quadratique sur E. On appelle* **cône isotrope** *l'ensemble $\mathcal{I}(q)$ défini par :*

$$\mathcal{I}(q) := \{\, x \in E \mid q(x) = 0 \,\}.$$

Notons que $\mathcal{I}(q)$ n'est pas un espace vectoriel, mais justement un «cône» c'est-à-dire un sous-ensemble de vecteurs $\mathcal{C}$ tel que si $x \in \mathcal{C}$, alors $\lambda\,x \in \mathcal{C}$, $\forall \lambda \in K$.

Exemple 1 – Soit $E = \mathbb{R}^2$ et $q(x) = x_1^2 - x_2^2$. On a :
$\mathcal{I}(q) = \Big\{ (x_1, x_2) \in \mathbb{R}^2 \mid x_1 = \pm\, x_2 \Big\}$ (cf. Fig. 1)

Exemple 2 – Soit $E = \mathbb{R}^3$, et $q(x) = x_1^2 + x_2^2 - x_3^2$. On a :
$\mathcal{I}(q) = \Big\{ (x_1, x_2, x_3) \in \mathbb{R}^3 \mid x_3 = \pm\,\sqrt{x_1^2 + x_2^2} \Big\}$ (cf. Fig. 2)
Notons que dans ces deux exemples $N(q) = 0$ (c'est-à-dire q est non dégénérée).

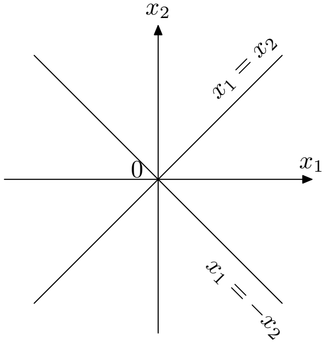


Figure 1

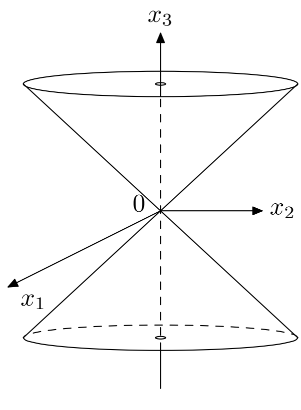


Figure 2

Propriété 9.10

$$\boxed{N(q) \subset \mathcal{I}(q)}$$

En effet si $x \in N(q)$ on a $s(x, y) = 0$, $\forall y \in E$; donc en particulier $s(x, x) = 0$ c'est-à-dire $q(x) = 0$.

Résumons ici, pour la commodité du lecteur, les définitions principales :

Si s est la ***forme polaire*** de q : $\begin{cases} q(x) &= s(x,x) \\ s(x,y) &= \frac{1}{2}[q(x+y) - q(x) - q(y)] \end{cases}$

Noyau : $N(q) = \{y \in E \mid s(x,y) = 0\,,\ \forall x \in E\}$

Cône isotrope : $\mathcal{I}(q) = \{x \in E \mid q(x) = 0\}$

$$N(q) \subset \mathcal{I}(q)$$

q est dite ***non dégénérée*** si $N(q) = 0$

q est ***définie*** si $q(x) = 0 \Longleftrightarrow x = 0$ c'est-à-dire si $\mathcal{I}(q) = \{0\}$

q (*à valeurs réelles*) est dite ***définie positive*** si : $q(x) \geq 0$ et $q(x) = 0 \Longrightarrow x = 0$

En dimension finie : $\dim E = \dim N(q) + \operatorname{rg}(q)$

Exercices 8. 9.

9.5 Bases orthogonales. Réduction des formes quadratiques.

Dans les paragraphes qui suivent nous allons nous occuper du problème de la ***réduction des formes quadratiques***, c'est-à-dire de la recherche de bases qui permettent d'en simplifier l'écriture.

Définition 9.11 – *Une base $\{e_i\}$ de E est dite* **orthogonale** *pour la forme bilinéaire symétrique s si : $s\,(e_i, e_j) = 0\,,;\ \forall i \neq j$. Elle est dite* **orthonormée** *si :* $s\,(e_i, e_j) = \delta_{ij}$.

Remarque. – En dimension finie, $\{e_i\}$ est une *base orthogonale* si et seulement si :

$$M(q)_{e_i} = \begin{pmatrix} a_1 & & 0 \\ & \ddots & \\ 0 & & a_n \end{pmatrix} \quad , \qquad \text{où : } a_i = s\,(e_i, e_i) = q(e_i')$$

ou encore, d'une manière équivalente, si et seulement si :

$$q(x) = a_1\, x_1^2 + \cdots + a_n\, x_n^2 \qquad (\text{où } x = \sum_{i=1}^{n} x_i e_i) \tag{$*$}$$

Notons que le rang de q est le nombre des a_i non nuls, ou encore le nombre des carrés dans l'expression (*).

De même $\{e_i\}$ est une *base orthonormée* si et seulement si :

$$M(q)_{e_i} = \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix}$$

ou aussi, si et seulement si :

$$q(x) = x_1^2 + \cdots + x_n^2 \qquad (\, x = \textstyle\sum_{i=1}^{n} x_i e_i)$$

Notons en particulier que s'il existe une base orthonormée, nécessairement q est de rang n (c'est-à-dire elle est non dégénérée). En particulier, *dans le cas où q est à valeurs réelles, il existe une base orthonormée si et seulement si q est définie positive, c'est-à-dire sa forme polaire est un produit scalaire.*

Chercher une base orthogonale revient donc à déterminer une base dans laquelle la matrice de q est diagonale, ou aussi, à écrire q sous la forme d'une somme de termes carrés.

Remarque. – Il ne faut pas confondre ce problème avec la diagonalisation des endomorphismes. Si A est la matrice d'un endomorphisme, la diagonaliser signifie chercher une matrice inversible P telle que $P^{-1}AP$ soit diagonale. Ici il s'agit de chercher une matrice P telle que ${}^tPAP$ soit diagonale.

Comme on vient de le voir, il n'existe pas forcément de bases orthonormée, par exemple si $\operatorname{rang} q \leq n$. Le théorème suivant affirme que, en revanche, *on peut toujours trouver des bases orthogonales.*

Théorème 9.12 – *Soit (E, q) un espace vectoriel de dimension finie, $E \neq \{0\}$, muni d'une forme quadratique q. Il existe alors sur E des* **bases orthogonales** *pour q.*
En d'autres termes, il existe toujours une base $\{e_i\}$ telle que, si $x = \sum_{i=1}^{n} x_i\, e_i$, alors

$$q(x) = a_1\, {x_1}^2 + \cdots + a_r\, {x_r}^2 \qquad (\text{avec } a_i \in K) \quad \text{où } r = \operatorname{rg} q$$

ou, d'une manière équivalente :

$$M(q)_{e_i} = \begin{pmatrix} a_1 & & & & & 0 \\ & \ddots & & & & \\ & & a_r & & & \\ & & & 0 & & \\ & & & & \ddots & \\ 0 & & & & & 0 \end{pmatrix}$$

Démonstration : La démonstration, qui suit de près celle du théorème 7.6 (cf. page 228), se fait par récurrence sur la dimension n de E.
Pour $n = 1$ il n'y a rien à démontrer. Supposons le théorème vrai à l'ordre $n-1$ et soit (E, q) un espace vectoriel de dimension n muni d'une forme quadratique.
– Si $q = 0$, le théorème est trivial : toutes les bases sont orthogonales.
– Supposons $q \neq 0$ et raisonnons comme dans le théorème 7.6. Soit $v \in E$, tel que $q(v) \neq 0$ et posons :

$$F = \{\, x \in E \mid s\,(x, v) = 0\,\}.$$

Il est facile de voir que F est un sous-espace vectoriel de E et que $v \notin F$. En effet, $F = \operatorname{Ker} \omega$, où ω est la forme linéaire définie par $\omega(x) = s(x, v)$. Comme $\omega(v) = s(v, v) \neq 0$, ω n'est pas la forme nulle et donc $\dim \operatorname{Ker} \omega \equiv \dim \operatorname{Ker} F = n-1$. D'autre part $v \notin F$, car $s(v, v) \neq 0$.
On en déduit que $E = \operatorname{Vect}\{v\} \oplus F$. Or, d'après l'hypothèse de récurrence, sur F il existe des bases orthogonales. Si donc $\{v_2, \cdots, v_n\}$ est une base orthogonale de F alors $\{v, v_2, \cdots, v_n\}$ est une base orthogonale de E . □

9.6 Recherche d'une base orthogonale par la méthode de Gauss

Pour chercher une base orthogonale, la méthode de Schmidt (cf. page 229) risque de ne pas pouvoir s'appliquer. En effet, lors des calculs on est amené à diviser par $q(\varepsilon_i)$ et ce scalaire peut être nul (c'est le cas où ε_i est isotrope).

Par exemple, soit q la forme quadratique sur $\mathbb{R}^2$ définie dans la base canonique $\{v_1, v_2\}$ par $q(x) = 2\,x_1\,x_2$. Si on pose :

$$\begin{cases} \varepsilon_1 = v_1 \\ \varepsilon_2 = v_2 + \lambda\, v_1 \end{cases}$$

et on cherche λ tel que $s(\varepsilon_1, \varepsilon_2) = 0$, c'est-à-dire $s(v_1, v_2) + \lambda\, s(v_1, v_1) = 0$, on trouve $1 + \lambda \cdot 0 = 0$

C'est pourquoi, lorsque l'espace n'est pas euclidien, pour chercher une base orthogonale on utilise une autre méthode fondée sur la réduction en carrés de Gauss.

a) Réduction en carrés d'une forme quadratique

Nous avons déjà appris (cf. page 225) à écrire une forme quadratique q comme somme de carrés de formes linéaires indépendantes, dans le cas où l'un au moins des termes carrés est non nul. Cela est suffisant pour les espaces euclidiens, car dans ce cas tous les termes carrés de q ont un coefficient strictement positif (cf. remarque page 224). Nous nous proposons maintenant de réduire en carrés une forme quadratique qui ne comporte que des termes rectangles, c'est-à-dire qui, dans une certaine base, s'écrit :

$$q(x) = a_{12}\, x_1\, x_2 + a_{13}\, x_1\, x_3 + \cdots + a_{ik}\, x_i\, x_k + \cdots + a_{n-1,n}\, x_{n-1}\, x_n.$$

Illustrons la méthode sur un exemple.

Soit $q : \mathbb{R}^3 \longrightarrow \mathbb{R}$ la forme quadratique qui s'écrit dans la base canonique :

$$q(x) = 5\,x_1\,x_2 + 6\,x_1\,x_3 + 3\,x_2\,x_3$$

– On choisit un terme rectangle à coefficient non nul $k x_1 x_j$, avec $k \neq 0$,

par exemple $5x_1x_2$

– On calcule les dérivées partielles q'_{x_i}, q'_{x_j}

$$\begin{aligned} q'_{x_1} &= 5\,x_2 + 6\,x_3 \\ q'_{x_2} &= 5\,x_1 + 3\,x_3 \end{aligned}$$

– On écrit :

$$q(x) = \frac{1}{k} q'_{x_i} \cdot q'_{x_j} + \text{termes correctifs :}$$

$$q(x) = \frac{1}{5}(5x_2 + 6x_3)(5x_1 + 3x_3) - \underbrace{\frac{18}{5}x_3^2}_{\text{terme correctif}}$$

– On aura une expression du type : $q(x) = \dfrac{1}{k}\varphi_1(x)\varphi_2(x)$ + des termes ne contenant ni x_i ni x_j, où φ_1 et φ_2 sont deux formes linéaires (en fait $\varphi_1 = q'_{x_i}$, $\varphi_2 = q'_{x_j}$).

$$q(x) = \frac{1}{5}\,\underbrace{(5x_2 + 6x_3)}_{\varphi_1}\,\underbrace{(5x_1 + 3x_3)}_{\varphi_2} - \underbrace{\frac{18}{5}x_3^2}_{\text{terme sans } x_1 \text{ et sans } x_2}$$

– On écrit le produit $\varphi_1\varphi_2$ sous la forme $\varphi_1\,\varphi_2 = \frac{1}{4}\underbrace{\left[(\varphi_1 + \varphi_2)^2 - (\varphi_1 - \varphi_2)^2\right]}$

$$\begin{aligned} q(x) = {} & \frac{1}{5 \times 4}(5x_1 + 5x_2 + 9x_3)^2 \\ & - \frac{1}{5 \times 4}(-5x_1 + 5x_2 - 3x_3)^2 - \frac{18}{5}x_3^2. \end{aligned}$$

– Si dans les termes correctifs on a un terme, carré, on continue comme à la page 225 ; s'il n'y a que des termes rectangles, on recommence l'opération jusqu'à écrire q comme une somme de carrés de formes linéaires.

Dans notre cas, q est la somme de carrés de formes linéaires :

$$q(x) = \frac{1}{20}(5x_1 + 5x_2 + 9x_3)^2 - \frac{1}{20}(5x_1 + 5x_2 - 3x_3)^2 - \frac{18}{5}x_3^2$$

Proposition 9.13 – *Les formes linéaires ainsi obtenues sont linéairement indépendantes.*

(Pour la démonstration : cf. exercice 11).

b) Recherche d'une base orthogonale

Pour chercher une base orthogonale on part de la remarque suivante. Soit

$$q(x) = a_{11}x_1^2 + \cdots + a_{nn}x_n^2 + \cdots + a_{ij}x_ix_j + \cdots$$

où $x = \sum_{i=1}^n x_ie_i$. D'après le théorème 9.12 on peut trouver une base orthogonale $\{v_i\}$. Si $\{v_i\}$ est une base orthogonale et $x = x'_1v_1 + \cdots + x'_nv_n$, on aura :

$$q(x) = b_1{x'}_1^2 + \cdots + b_n{x'}_n^2 \,, \qquad \text{avec} \quad b_i = s(v_i, v_i) = q(v_i),$$

c'est-à-dire :

$$M(q)_{v_i} = \begin{pmatrix} b_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & b_n \end{pmatrix}$$

Si $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = M(x)_{e_i}$ et $X' = \begin{pmatrix} x'_1 \\ \vdots \\ x'_n \end{pmatrix} = M(x)_{v_i}$, on a :

$$X' = P^{-1}X \qquad \text{avec} \qquad P = P_{e_i \to v_i}$$

c'est-à-dire, en notant l_{ij} les coefficients de la matrice P^{-1} :

$$\begin{aligned} x'_1 &= l_{11}x_1 + \cdots + l_{1n}\,x_n := \varphi_1(x) \\ &\cdots\cdots\cdots\cdots\cdots \\ x'_n &= l_{n1}x_1 + \cdots + l_{nn}x_n := \varphi_n(x). \end{aligned}$$

$\varphi_1, \ldots, \varphi_n$ sont de formes linéaires indépendantes (puisque dét $P \neq 0$).
On aura donc :

$$q(x) = b_1\,\varphi_1(x)^2 + \cdots + b_n\,\varphi_n(x)^2$$

Ainsi :

> *Si l'on écrit $q(x)$ sous la forme d'une combinaison linéaire de carrés de formes linéaires $\varphi_1, ..., \varphi_n$, les formes $\varphi_1, ..., \varphi_n$ sont les composantes de x dans une base orthogonale.*

Ceci suggère une méthode simple pour trouver une base orthogonale. Illustrons cela sur un exemple.

Exemple –
Soit $q : \mathbb{R}^3 \longrightarrow \mathbb{R}$ définie dans la base canonique $\{e_i\}$ par :

$$q(x) = x_1^2 + 3\,x_2^2 + 7\,x_3^2 + 2\,x_1\,x_2 + 8\,x_2\,x_3$$

La réduction de Gauss donne :

$$\begin{aligned} q(x) &= (x_1 + x_2)^2 + 2\,(x_2 + 2x_3)^2 - x_3^2 \\ &\equiv \varphi_1(x)^2 + 2\,\varphi_2(x)^2 - \varphi_3(x)^2 \end{aligned}$$

Posons :

$$(*) \quad \begin{cases} x_1' \equiv \varphi_1(x) = x_1 + x_2 \\ x_2' = \varphi_2(x) = x_2 + 2x_3 \\ x_3' = \varphi_3(x) = x_3 \end{cases}$$

x_1', x_2', x_3' sont les composantes du vecteur x dans une base orthogonale $\{v_i\}$.
Or $X' = P^{-1}X$ avec $P = P_{e_i \to v_i} = \| \, v_1, v_2, v_3 \, \|_{e_i}$ donc $X = PX'$, c'est-à-dire :

En résolvant le système (*) par rapport à x_1, x_2, x_3 on obtient un système du type $X = PX'$. Les colonnes de P sont les vecteurs de la base orthogonale.

Dans notre cas :

$$\begin{cases} x_1 = x_1' - x_2' + 2\,x_3' \\ x_2 = \phantom{x_1' -{}} x_2' - 2\,x_3' \\ x_3 = \phantom{x_1' - x_2' + 2\,{}} x_3' \end{cases} \qquad \text{donc :} \qquad P = \begin{array}{c} \begin{pmatrix} 1 & -1 & 2 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{pmatrix} \\ \begin{array}{ccc} \uparrow & \uparrow & \uparrow \\ v_1 & v_2 & v_3 \end{array} \end{array}$$

Ainsi les vecteurs $v_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$, $v_2 = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}$, $v_3 = \begin{pmatrix} 2 \\ -2 \\ 1 \end{pmatrix}$ forment une base orthogonale pour q.

Remarque - D'après la remarque page 304, on a :

$$\operatorname{rg} q = \textit{nombre de carrés dans la réduction de Gauss.}$$

Exercices 10. 11.

9.7 Classification des formes quadratiques sur un espace vectoriel complexe

Théorème 9.14 – *Soit E un espace vectoriel de dimension n sur $\mathbb{C}$ (E est donc isomorphe à $\mathbb{C}^n$ moyennant le choix d'une base) et q une forme quadratique sur E. Il existe alors une base $\{e_i\}$ telle que si $x = \sum_{i=1}^n x_i e_i$:*

$$q(x) = x_1^2 + \cdots + x_r^2 \qquad (\, r = \operatorname{rg} q \,)$$

c'est-à-dire :

$$M(q)_{e_i} = \left(\begin{array}{c|c} \boxed{\begin{matrix} 1 & & \\ & \ddots & \\ & & 1 \end{matrix}} & 0 \\ 0 & \begin{matrix} 0 & & \\ & \ddots & \\ & & 0 \end{matrix} \end{array} \right)$$

La démonstration est immédiate. Soit $\{\varepsilon_i\}$ une base orthogonale et

$$q(x) = a_1\, y_1^2 + \cdots + a_r\, y_r^2\,, \qquad \text{pour} \quad x = \textstyle\sum_{i=1}^n\, y_i\, \varepsilon_i.$$

Comme $a_i \in \mathbb{C}$, il existe $\lambda_i \in \mathbb{C}$ tel que $a_i = \lambda_i^2$; donc :

$$q(x) = (\lambda_1 y_1)^2 + \cdots + (\lambda_r y_r)^2.$$

En posant $x_i = \lambda_i\, y_i$, on aura $q(x) = x_1^2 + \cdots + x_r^2$. $\square$

Corollaire 9.15 – *Pour une forme quadratique q sur un espace vectoriel de dimension n sur $\mathbb{C}$ il existe une base orthonormée si et seulement si $\operatorname{rg} q = n$ (c'est-à-dire : si et seulement si q est non dégénérée).*

Exercice 12.(2.)

9.8 Classification des formes quadratiques sur un espace vectoriel réel. Théorème de Sylvester

Théorème de Sylvester . 9.16 *Soit E un espace vectoriel de dimension n sur $\mathbb{R}$ (E est donc isomorphe à $\mathbb{R}^n$ moyennant le choix d'une base), et q une forme quadratique sur E. Il existe alors une base $\{e_i\}$ de E telle que si $x = \sum_{i=1}^n x_i e_i$, on a :*

$$q(x) = x_1^2 + \cdots + x_p^2 - x_{p+1}^2 - \cdots - x_r^2$$

c'est-à-dire :

$$M(q)_{e_i} = \begin{pmatrix} \boxed{\begin{matrix}1 & & \\ & \ddots & \\ & & 1\end{matrix}} & & 0 \\ & \boxed{\begin{matrix}-1 & & \\ & \ddots & \\ & & -1\end{matrix}} & \\ 0 & & \begin{matrix}0 & & \\ & \ddots & \\ & & 0\end{matrix} \end{pmatrix}$$

où $r = \operatorname{rang}(q)$ et p est un entier qui ne dépend que de la forme quadratique q *(et non pas de la base). Le couple $(p, r-p)$, noté $\operatorname{sign}(q)$, est dit* ***signature*** *de q.*

Démonstration : Soit $\{e_i\}$ une base orthogonale. Si $x = \sum_{i=1}^n y_i e_i$ on aura :

$$q(x) = a_1\, y_1^2 + \cdots + a_r\, y_r^2 \qquad (\text{avec } a_i \in \mathbb{R}).$$

Supposons que $a_1, \ldots, a_p > 0$ et $a_{p+1}, \ldots, a_r < 0$.
On pourra écrire :

$$\begin{aligned} q(x) &= (\sqrt{a_1}\, y_1)^2 + \cdots + (\sqrt{a_p}\, y_p)^2 - (\sqrt{-a_{p+1}}\, y_{p+1})^2 - \cdots - (\sqrt{-a_r}\, y_r)^2 \\ &= x_1^2 + \cdots + x_p^2 - x_{p+1}^2 - \cdots - x_r^2 \end{aligned}$$

avec $x_i = \sqrt{a_i}\, y_i \quad (i = 1, \cdots p) \qquad$ et $\qquad x_j = \sqrt{-a_j}\, y_j \quad (j = p+1, \cdots, r)$.

Il reste à montrer que p ne dépend pas du choix de la base. Considérons deux bases $\{e_i\}$ et $\{e'_i\}$ telles que

$$q(x) = x_1^2 + \cdots + x_p^2 - x_{p+1}^2 - \cdots - x_r^2 \qquad \left(x = \textstyle\sum_{i=1}^n x_i e_i\right)$$

$$\text{et} \quad q(x) = y_1^2 + \cdots + y_{p'}^2 - y_{p'+1}^2 - \cdots - y_r^2 \qquad \left(x = \textstyle\sum_{i=1}^n y_i e'_i\right)$$

Soient

$$\begin{array}{ll} F = \text{Vect}\,\{e_1, \ldots, e_p\} & F' = \text{Vect}\{e'_1, \ldots, e'_{p'}\} \\ G = \text{Vect}\{e_{p+1}, \ldots, e_n\} & G' = \text{Vect}\{e'_{p'+1}, \ldots, e'_n\} \end{array}$$

On a :

$$\begin{array}{llll} x \in F \setminus \{0\} & \Longrightarrow q(x) > 0, & x \in F' \setminus \{0\} & \Longrightarrow q(x) > 0 \\ x \in G & \Longrightarrow q(x) \leq 0, & x \in G' & \Longrightarrow q(x) \leq 0 \end{array}$$

donc :

$$x \in F \cap G' \Longrightarrow x = 0 \ \text{ et, par conséquent, } \ F \cap G' = \{0\}$$

Ainsi F et G' sont en somme directe. Puisque $F \oplus G' \subset E$, on a :

$$\dim F + \dim G' \leq n$$

c'est-à-dire : $p + (n - p') \leq n$ et donc : $p \leq p'$.
De la même manière on voit que $F' \cap G = \{0\}$ et donc $p' \leq p$. $\square$

Corollaire 9.17 – *Soit q une forme quadratique sur un espace vectoriel **réel** E de dimension finie n. Alors :*

$$\begin{array}{lcl} q \text{ est définie positive} & \Longleftrightarrow & \text{sign}(q) = (n, 0) \\ \text{(c'est-à-dire } E \text{ est euclidien)} & & \Updownarrow \\ & & \text{il existe des bases orthonormées} \\ & & \\ q \text{ est définie négative} & \Longleftrightarrow & \text{sign}(q) = (0, n) \\ & & \\ q \text{ est non dégénérée} & \Longleftrightarrow & \text{sign}(q) = (p, n-p) \\ \text{(c'est-à-dire } N(q) = \{0\}) & & \end{array}$$

Exemple – *Déterminer la signature de la forme quadratique : $q : \mathbb{R}^3 \longrightarrow \mathbb{R}$ définie dans la base canonique, par :*

$$q(x) = x_1^2 + 2\,x_2^2 + 15\,x_3^2 - 4\,x_1\,x_2 + 6\,x_1\,x_3 - 8\,x_2\,x_3$$

En appliquant la méthode de Gauss on trouve :

$$\begin{aligned} q(x) \quad &= (x_1 - 2\,x_2 + 3\,x_3)^2 - 2\,(x_2 - x_3)^2 + 8\,x_3^2 \\ &= {x'}_1^2 - {x'}_2^2 + {x'}_3^2 \end{aligned}$$

(avec $x'_1 = x_1 - 2\,x_2 + 3\,x_3$, $x'_2 = \sqrt{2}\,(x_2 - x_3)$, $x'_3 = \sqrt{8}\,x_3$). Donc : $\text{sign}(q) = (2, 1)$.

Exercices 12.(1) 13. 14.

9.9 Sous-espaces orthogonaux

Comme dans le cas d'un produit scalaire, on dit que deux vecteurs $x, y \in E$ sont orthogonaux pour la forme bilinéaire s, si $s(x,y) = 0$. On note cela : $x \underset{s}{\perp} y$.

Définition 9.18 – *Soit A un sous-ensemble de E (non nécessairement un sous-espace vectoriel). On appelle* **orthogonal de** *A pour la forme bilinéaire s (ou, simplement, orthogonal de A), l'ensemble*

$$A^{\perp} := \{x \in E \mid s(x\,,\,a) = 0\,,\ \forall a \in A\}.$$

Les propriétés suivantes se vérifient facilement :

Propriétés 9.19 –

1. *$A^{\perp}$ est un sous-espace vectoriel de E (même si A ne l'est pas).*
2. $\{0\}^{\perp} = E$
3. $E^{\perp} = N(q)$
4. $\forall A \subset E\ :\ N(q) \subset A^{\perp}$.

Le lemme suivant généralise la proposition 7.15.

Lemme 9.20 – *Soit* $j : \begin{array}{ccc} E & \longrightarrow & E^* \\ y & \mapsto & s(\ ,y) \end{array}$. *Si E est de dimension finie, pour tout sous-espace vectoriel F de E on a :* $\Big(j(F)\Big)^0 = F^{\perp}$.[4]

En effet,
$$j(F)^0 = \{\, x \in E^{**} \simeq E \ \mid\ \varphi(x) = 0\,,\ \forall \varphi \in j(F)\,\}$$
et
$$\varphi \in j(F) \iff \exists\, y \in F \text{ tel que } \varphi = s(\cdot, y)$$
donc :
$$j(F)^0 = \{x \in E \mid s(x,y) = 0 \quad \forall y \in F\} = F^{\perp} \qquad \square$$

Proposition 9.21 – *Soit E de dimension finie et F un sous-espace vectoriel de E. Alors :*

1. $\dim E = \dim F + \dim F^{\perp} - \dim(F \cap N)$
2. $F^{\perp\perp} = F + N \qquad$ où $N := N(q)$.

En particulier si q ***est non dégénérée :***

$$\dim E = \dim F + \dim F^{\perp}$$
$$F^{\perp\perp} = F$$

Démonstration :

1. D'après le lemme : $\quad \dim F^{\perp} = \dim j(F)^0 = \dim E - \dim j(F)$

   d'autre part : $\quad \dim j(F) = \dim\ \mathrm{Im}\big(j|_F\big)$

   Or, d'après le théorème du rang :

[4] En identifiant E^{**} à E (cf. proposition 3.36) ; en fait : $\Big(j(F)\Big)^0 \subset E^{**}$.

$$\begin{aligned}\dim F &= \dim\ \mathrm{Im}\big(j|_F\big) + \dim \mathrm{Ker}\big(j|_F\big) = \\ &= \dim\big(\mathrm{Im}\, j|_F\big) + \dim(\mathrm{Ker}\, j\ \cap F) = \dim\big(j(F)\big) + \dim\,(F \cap N) \\ &= \dim E - \dim F^{\perp} + \dim\,(F\cap N)\end{aligned}$$

2. On a $F \subset F^{\perp\perp}$ (en effet, si $x \in F$: $s(x,z) = 0\ \forall z \in F^{\perp}$, c'est-à-dire $s(z,x) = 0\ \forall z \in F^{\perp}$, donc $x \in F^{\perp\perp}$).
   D'autre part, le noyau est inclus dans tout orthogonal (cf. Propriété 9.19, 4.) ; donc : $N \subset F^{\perp\perp}$. On a ainsi :
   $$F + N \subset F^{\perp\perp}$$
   Or, d'après 1 : $$\dim E = \dim F + \dim F^{\perp} - \dim\,(F\cap N)$$
   et, en remplaçant F par $F^{\perp}$:
   $$\dim E = \dim F^{\perp} + \dim F^{\perp\perp} - \dim\,(F^{\perp} \cap N)$$
   mais $N \subset F^{\perp}$, donc $\dim F^{\perp}\cap N = \dim N$.
   On en déduit, par soustraction :
   $$\dim F - \dim F^{\perp\perp} + \dim N - \dim F\cap N = 0$$
   Or : $$\dim(F+N) = \dim F + \dim N - \dim F \cap N$$
   et donc : $$\dim F^{\perp\perp} = \dim\,(F+N)$$
   Puisque $F+N$ est un sous-espace de $F^{\perp\perp}$, on a $F^{\perp\perp} = F + N$. $\square$

Sous-espaces isotropes

Nous nous proposons maintenant de regarder dans quel cas on a :

$$E = F \oplus F^{\perp}$$

comme il en est dans les espaces euclidiens.
Pour cela il est nécessaire tout d'abord que $F\cap F^{\perp} = \{0\}$. Or cette condition n'est pas toujours vérifiée, *même si q est non dégénérée.* Soit en effet v un vecteur isotrope non nul et $F = \mathrm{Vect}\{v\}$. On a $s(v,v) = 0$, c'est-à-dire $v \underset{s}{\perp} v$, donc $v \in F^{\perp}$ et par conséquent : $F\cap F^{\perp} \neq \{0\}$.

Définition 9.22 – *Un sous-espace vectoriel F de E est dit* ***isotrope*** *si*

$$F \cap F^{\perp} \neq \{0\}.$$

Comme on vient de le voir, si $\mathcal{I} \neq \{0\}$ il existe des sous-espaces isotropes : par exemple les droites vectorielles engendrées par les vecteurs isotropes. Réciproquement si F est isotrope, et $v \in F\cap F^{\perp}$, $v \neq 0$, on a $v \in \mathcal{I}$ et donc $\mathcal{I} \neq 0$. Ainsi :

Il existe des sous-espaces isotropes si et seulement si $\mathcal{I} \neq \{0\}$.

Proposition 9.23 – *Soit E de dimension finie et F un sous-espace de E. Alors :*

$$E = F\oplus F^{\perp} \iff F \textit{ \textbf{est non isotrope}}\ \ (\text{c'est-à-dire : } F\cap F^{\perp} = \{0\})$$

Démonstration : : La condition est évidemment nécessaire.
Supposons que $F\cap F^{\perp} = \{0\}$. Puisque $N \subset F^{\perp}$ (cf. propriété 9.19. *4.*), on a $F \cap N = \{0\}$; donc $\dim E = \dim F + \dim F^{\perp}$. Il s'ensuit que $E = F \oplus F^{\perp}$. $\square$

Exercices 15.16. 17. 18. 19. 20. 21.

9.10 Formes quadratiques dans un espace euclidien

Dans ce paragraphe on suppose que l'espace vectoriel E est muni de deux structures : une forme quadratique q et une structure euclidienne $\langle\, ,\, \rangle$ [5]. Nous allons démontrer que l'on peut toujours construire des bases qui sont orthogonales à la fois pour $\langle\, ,\rangle$ et pour q (cela permet de simplifier les calculs car dans de telles bases les matrices de $\langle\, ,\, \rangle$ et de q sont diagonales).

Théorème 9.24 – *Soit $(E, \langle\, ,\rangle)$ un espace euclidien et q une forme quadratique sur E. Il existe alors des bases qui sont orthogonales à la fois pour $\langle\, ,\, \rangle$ et pour q.*

La démonstration reprend en fait celle de la proposition 7.39 (cf. page 254) :

a) *Soit s la forme polaire de q. Il existe un et un seul endomorphisme f_s tel que :*

$$\langle x, f_s(y)\rangle = s(x,y), \quad \forall x,y \in E.$$

Pour cela, on procède comme dans la Proposition 7.39 : on considère une base $\{e_i\}$ orthonormée pour $\langle\, ,\, \rangle$. En écrivant la relation ci-dessus sous forme matricielle, on voit facilement que, si $S = M(q)_{e_i}$, f_s est l'endomorphisme qui dans la base $\{e_i\}$ est représenté par la matrice S.

b) *f_s est autoadjoint pour $\langle\, ,\, \rangle$.*

En effet : $\langle x, f_s(y)\rangle = s(x,y) = s(y,x) = \langle y, f_s(x)\rangle = \langle f_s(x), y\rangle$

D'après le théorème 7.38, les sous-espaces propres E_λ de f_s sont deux à deux orthogonaux pour $\langle\, ,\, \rangle$. En fait :

c) *Les espaces propres E_λ sont aussi deux à deux orthogonaux pour s.*

En effet, soient v_i et v_j tels que $f_s(v_i) = \lambda_i v_i$ et $f_s(v_j) = \lambda_j v_j$, avec $i \neq j$; on a :

$$s(v_i, v_j) = \langle v_i, f_s(v_j)\rangle = \lambda_j \langle v_i, v_j\rangle .$$

Puisque $\langle v_i, v_j\rangle = 0$, on a aussi $s(v_i, v_j) = 0$.

Il suffit maintenant de prendre dans chaque espace propre E_λ une base orthogonale de vecteurs propres (orthogonale pour $\langle\, ,\, \rangle$), pour obtenir une base qui est aussi orthogonale pour s. □

Le corollaire suivant permet de déterminer une base orthogonale et calculer la signature d'une forme quadratique en diagonalisant la matrice qui la représente dans une base quelconque.

Corollaire 9.25 – *Soit q une forme quadratique sur un espace vectoriel réel de dimension finie et $S = M(q)_{e_i}$ où e_i est une base quelconque. Alors :*

1. *On peut construire une base orthogonale de q formée de vecteurs propres de S.*
2. *De plus :* $\operatorname{sign}(q) = (n_+, n_-)$
   où n_+ = *nombre des valeurs propres strictement positives de* S
   n_- = *nombre des valeurs propres strictement négatives de* S.

[5] La situation habituelle est celle qui consiste à considérer une forme quadratique q sur $(\mathbb{R}^n, \langle\, ,\, \rangle)$, où $\langle\, ,\, \rangle$ est le produit scalaire canonique.

Démonstration : Soit $\{e_i\}$ une base de E et $\langle\, ,\, \rangle$ le produit scalaire associé à cette base (cf. exemple 2, page 222). On considère, comme dans le théorème ci-dessus, l'endomorphisme f_s défini par : $\langle x, f_s(y)\rangle = s(x,y)$, $\forall x, y \in E$. Comme on l'a vu, on a : $M(f_s)_{e_i} = S$.
Soit $\{v_1 \cdots v_n\}$ une base de vecteurs propres de f_s orthogonale à la fois pour $\langle\, ,\rangle$ et pour q, et normeé pour $\langle\, ,\, \rangle$, construite comme dans le théorème. On a :

$$\begin{aligned} s(x,y) &= \sum_{i,j=1}^{n} x_i\, y_j\, s(v_i, v_j) = \sum_{i,j=1}^{n} x_i\, y_j\, \langle v_i, f_s(v_j\rangle = \sum_{i,j=1}^{n} x_i\, y_j\, \lambda_j\, \langle v_i, v_j\rangle \\ &= \sum_{i,j=1}^{n} x_i\, y_j\, \delta_{ij} = \lambda_1\, x_1\, y_1 + \cdots + \lambda_n\, x_n\, y_n \end{aligned}$$

d'où :

$$s(x,x) = \lambda_1\, {x_1}^2 + \cdots + \lambda_n\, {x_n}^2$$

ce qui montre que la signature de s est donnée par les signes des valeurs propres de f_s (donc de S). $\square$

Exemple –
Soit $E = \mathbb{R}^3$ *et* q *la forme quadratique définie dans une base* $\{e_i\}$ *par :*

$$q(x) = -2\,x_1^2 - 2\,x_2^2 - 2\,x_3^2 + 6\,x_1x_2 + 6\,x_2x_3 + 6\,x_1x_3$$

Déterminer une base orthogonale et la signature de q *par la méthode indiquée dans le corollaire.*
On a :

$$S = M(s)_{e_i} = \begin{pmatrix} -2 & 3 & 3 \\ 3 & -2 & 3 \\ 3 & 3 & -2 \end{pmatrix} \qquad \text{et} \qquad P_S(\lambda) = -(\lambda - 4)(\lambda + 5)^2$$

donc $\mathrm{Sp}'(S) = \{4, -5, -5\}$ et par conséquent $\mathrm{sign}(s) = (1, 2)$.
Pour avoir une base orthogonale pour s, il suffit de déterminer une base de vecteurs propres de S qui est orthogonale pour le produit scalaire $\langle\, ,\rangle$ associé à la base $\{e_i\}$.

On trouve : E_4 : engendré par $v_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$, E_{-5} : $x + y + z = 0$.

Soit, par exemple, $v_2 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \in E_{-5}$ et prenons $v_3 = \begin{pmatrix} a \\ b \\ c \end{pmatrix} \in E_{-5}$ tel que $\langle v_3, v_2\rangle = 0$. Pour cela, a, b, c doivent vérifier le système :

$$\begin{cases} a + b + c = 0 \\ a \qquad - c = 0 \end{cases}$$

ce qui donne, par exemple, $v_3 = \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}$.

Ainsi les vecteurs v_1, v_2, v_3 qui dans la base $\{e_i\}$ sont donnés par

$$v_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \quad , \quad v_2 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \quad , \quad v_3 = \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}$$

forment une base orthogonale à la fois pour $\langle,\rangle$ et pour q.

NOTA. – La méthode fondée sur la réduction de Gauss (cf. page 306 et exemple page 310) permet d'obtenir plus facilement une base orthogonale et la signature de q. Cependant cette méthode est nécessaire si l'on cherche une base orthogonale pour q qui est aussi orthogonale pour le produit scalaire. En particulier elle est utilisée pour déterminer la forme réduite des quadriques dans une base orthonormée (cf. Appendice A.11).

Exercices 22. 23. 24.

9.11 Endomorphisme adjoint

La notion d'endomorphisme adjoint, que nous avons introduite dans le cadre des espaces euclidiens (cf. proposition 7.16), se généralise facilement aux espaces vectoriels de dimension finie munis d'une forme quadratique *non dégénérée.*

Proposition 9.26 – *Soient (E, q) un espace vectoriel de dimension finie muni d'une forme quadratique* non dégénérée, *et $f \in \mathrm{End}(E)$. Il existe alors un et un seul endomorphisme f^* de E tel que*

$$s\big(f(x), y\big) = s\big(x, f^*(y)\big)\,, \qquad \forall x, y \in E$$

(s étant la forme polaire de q). f^* est dit ***adjoint de*** *f relativement à s.*

En effet, soient $\{e_i\}$ une base de E, $S = M(s)_{e_i}$, $A = M(f)_{e_i}$ et $X = M(x)_{e_i}$, $Y = M(y)_{e_i}$. Supposons que f^* existe et soit $A^* = M(f^*)_{e_i}$. L'identité de l'énoncé s'écrit :

$$(A\,X)\,{}^tS\,Y = {}^tXSA^*Y\,, \qquad \forall X, Y \in \mathcal{M}_{n,1}(K)$$

c'est-à-dire

$${}^tX({}^tAS)Y = {}^tXSA^*Y\,, \qquad \forall X, Y \in \mathcal{M}_{n,1}(K)$$

ce qui est équivalent à

$${}^tAS = SA^*.$$

Comme S est non dégénérée, S est inversible et donc :

$$\boxed{A^* = S^{-1}\,{}^tA\,S} \qquad (*)$$

Ceci montre que A^* (donc f^*) est unique.
Réciproquement, si on définit $f^* : E \longrightarrow E$ par $M(f^*)_{e_i} = S^{-1}\,{}^tAS$, en remontant les calculs on voit que f^* vérifie l'identité de l'énoncé. □

REMARQUE. – Si $\{e_i\}$ est une *base orthonormée* (lorsqu'elle existe, par exemple, dans le cas euclidien ou dans le cas où q est à valeurs complexes et est non dégénérée (cf. corollaire 9.15)), on a $S = I$ et on retrouve donc la formule $A^* = {}^tA$ que l'a déjà vue dans le cas euclidien (cf. 7.16.).

Exemple – Soit $E = \mathbb{R}^2$ muni de la forme quadratique $q(x) = x_1^2 - x_2^2$.
Si $A = M(f)_{e_i} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, on a :

$$A^* = S^{-1}\begin{pmatrix} a & c \\ b & d \end{pmatrix} S = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} a & c \\ b & d \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} a & -c \\ -b & -d \end{pmatrix}$$

Proposition 9.27 – *Pour tous endomorphismes f et g et pour tout scalaire λ, on a :*

a) $f^{**} = f$, $(\mathrm{id})^* = \mathrm{id}$;

b) $(f+g)^* = f^* + g^*$, $(\lambda f)^* = \lambda f^*$, $(f \circ g)^* = g^* \circ f^*$;

c) $\mathrm{rg}\ f^* = \mathrm{rg}\ f$, $\mathrm{dét}\, f^* = \mathrm{dét}\, f$.

a) et b) se démontrent facilement à l'aide de la formule (*) ci-dessus ou, mieux, par exemple :

$$s\,(f(x)\,,\,y\,) = s\,(x,\ f^*(y)\,) = s\,(f^{**}(x)\,,\,y\,) = 0\,, \quad \forall y \in E,$$

donc $f(x) = f^{**}(x)$ pour tout $x \in E$, c'est-à-dire $f = f^{**}$ (cf. exercice 8).
Pour c), utiliser la formule (*). □

Exercice 25.

9.12 Groupe orthogonal associé à une forme quadratique

Comme au chapitre 7, nous allons étudier les endomorphismes f de E qui conservent une forme quadratique q, c'est-à-dire tels que $q\big(f(x)\big) = q(x)$, $\forall x \in E$. Il s'agit donc de la généralisation aux espaces non euclidiens des isométries vectorielles. Dans le cadre de la Relativité $\big(q(x) = x_1^2 + x_2^2 + x_3^2 - x_4^2\big)$ on obtient le «groupe de Lorentz» (cf. exercice 26) qui, comme on peut l'imaginer, joue un rôle important en Physique Mathématique.

Proposition 9.28 – *Soit (E,q) un espace vectoriel de dimension finie muni d'une forme quadratique q* non dégénérée *et soit f un endomorphisme de E. Les propriétés suivantes sont équivalentes :*

1. $q\big(f(x)\big) = q(x)\,, \qquad \forall x \in E$
2. $s\big(f(x), f(y)\big) = s(x,y)\,, \quad \forall x,y \in E$
3. $f^* \circ f = \mathrm{id}$ (*ou d'une manière équivalente :* $f \circ f^* = \mathrm{id}$).

   En particulier f est bijectif.

Un tel endormorphisme est dit ***orthogonal*** *relativement à q.*

Démonstration : L'équivalence de a) et b) se montre comme dans la proposition 7.19 page 239. Pour montrer l'équivalence de b) et c) on pourrait utiliser l'expression matricielle de f^* (cf. formule (*), page 315) et procéder comme en 7.19. Cependant il est plus simple de raisonner de la manière suivante :

$$s\big(f(x),f(y)\big)=s(x,y)\,,\ \forall x,y\in E \iff s\big(f^*\circ f(x),y\big)=s(x,y)\,,\ \forall x,y\in E$$
$$\iff f^*\circ f(x)=x\,,\ \forall x\in E \quad \text{(cf. exercice 8)} \iff f^*\circ f=\mathrm{id}\quad \square$$

Proposition 9.29 – *Soit* $\mathrm{O}(q):=\big\{f\in\mathrm{End}\,(E)\mid f^*\circ f=\mathrm{id}\big\}$. *On a :*

1. $\mathrm{id}\in\mathrm{O}(q)$;
2. *si* $f,g\in\mathrm{O}(q)$ *alors* $f\circ g\in\mathrm{O}(q)$;
3. *si* $f\in\mathrm{O}(q)$ *alors* $f^{-1}\in\mathrm{O}(q)$.

En particulier, $\mathrm{O}(q)$ *est un groupe pour la loi de composition des applications (cf. Appendice A.1), dit* ***groupe orthogonal de*** q.

La démonstration est une simple vérification. Par exemple, si $f,g\in\mathrm{O}(q)$, on a :
$$(f\circ g)^*\circ(f\circ g)=g^*\circ f^*\circ f\circ g=g^*\circ\mathrm{id}\circ g=g^*\circ g=\mathrm{id}$$
donc $f\circ g\in\mathrm{O}(q)$. $\square$

Proposition 9.30 – *Si* $f\in\mathrm{O}(q)$ *alors* $\mathrm{dét}f=\pm1$. *L'ensemble*
$$\mathrm{SO}(q):=\big\{f\in\mathrm{O}(q)\mid \mathrm{dét}f=1\big\}$$
est un sous-groupe de $\mathrm{O}(q)$*, dit* ***groupe spécial orthogonal*** *de* q.

En effet, si $f\in\mathrm{O}(q)$ on a $f^*\circ f=\mathrm{id}$, donc $\mathrm{dét}(f^*\circ f)=1$. Puisque $\mathrm{dét}f^*=\mathrm{dét}f$, on a $(\mathrm{dét}f)^2=1$, donc $\mathrm{dét}f=\pm1$.
La vérification de la seconde partie est laissée en exercice. $\square$

Proposition 9.31 – *Soit* $\{e_i\}$ *une base de* E, $S=M(q)_{e_i}$ *et* $A=M(f)_{e_i}$. *On a :*
$$f\in\mathrm{O}(q)\iff {}^tA\,SA=S$$

En effet

$$\begin{aligned} f\in\mathrm{O}(q) &\iff s\big(f(x),f(y)\big)=s(x,y)\,,\ \forall\,x,y\in E\\ &\iff {}^t(AX)S(AY)={}^tXSY\,,\ \forall\,X,Y\in\mathcal{M}_{n,1}(K)\\ &\iff {}^tX\ {}^tA\,SAY={}^tXSY\,,\ \forall\,X,Y\in\mathcal{M}_{n,1}(K)\iff {}^tA\,SA=S\quad\square\end{aligned}$$

NOTA. – En particulier, si $\{e_i\}$ est une *base orthonormée* (à supposer qu'elle existe), on a $S=I$, donc : $f\in\mathrm{O}(q)\iff {}^tAA=I$ (cf. 7.19).

Exercice 26.

EXERCICES

1 Soit la forme bilinéaire $b : \mathbb{R}^3 \times \mathbb{R}^3 \longrightarrow \mathbb{R}$ définie dans la base canonique par :

$$b(x,y) := x_1y_1 + x_2y_2 + 3\,x_3y_3 + 6\,x_1y_3 + 2\,x_2y_1 - 3\,x_2y_3 + 3\,x_3y_1 + x_3y_2$$

Ecrire la matrice de b dans la base canonique. Calculer $N(b)$, $\mathrm{rg}(b)$ et $b(z,w)$ où : $z = \begin{pmatrix} 2 \\ -1 \\ 0 \end{pmatrix}$, $w = \begin{pmatrix} 5 \\ 15 \\ 1 \end{pmatrix}$. Ecrire la matrice de b dans la base $\{v_1, v_2, v_3\}$ où $v_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$, $v_2 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$, $v_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$.

2 Déterminer le *noyau à gauche* $\tilde{N}(b)$ de la forme bilinéaire b de l'exercice 1.

** **3** **Représentation matricielle canonique d'une forme bilinéaire antisymétrique**

Soit E un espace vectoriel de dimension finie sur K et $\Omega : E \times E \longrightarrow K$ une forme bilinéaire antisymétrique, c'est-à-dire telle que :

$$\Omega(x,x) = 0\,, \quad \forall x \in E.$$

1. Montrer que $\forall x, y \in E \ : \ \Omega(x,y) = -\Omega(y,x)$.
2. Montrer que si $x, y \in E$ sont liés, alors $\Omega(x,y) = 0$. En déduire que si $\dim E = 1$, alors $\Omega = 0$.
3. Soit $\dim E \geq 2$ et $\Omega \neq 0$. Montrer qu'il existe deux vecteurs indépendants $u_1, u_2 \in E$ tels que $\Omega(u_1, u_2) = 1$.
4. On note $F = \mathrm{Vect}\{u_1, u_2\}$. Montrer que :
$$M(\Omega|_F)_{\{u_1,u_2\}} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$
5. Soit $W = F^{\perp} = \{w \in E \mid \Omega(u,w) = 0\,,\ \forall\, u \in F\}$. Montrer que $E = F \oplus W$.
6. En déduire le résultat suivant :
Soit Ω une forme bilinéaire antisymétrique sur un espace vectoriel E de dimension finie sur K. Il existe alors une base $\{e_i\}$ telle que :

$$M(\Omega)_{e_i} = \begin{pmatrix} \boxed{\begin{matrix} 0 & 1 \\ -1 & 0 \end{matrix}} & & & & & & \Large 0 \\ & \boxed{\begin{matrix} 0 & 1 \\ -1 & 0 \end{matrix}} & & & & & \\ & & \ddots & & & & \\ & & & \boxed{\begin{matrix} 0 & 1 \\ -1 & 0 \end{matrix}} & & & \\ & & & & 0 & & \\ \Large 0 & & & & & \ddots & \\ & & & & & & 0 \end{pmatrix}$$

En particulier : *toute forme bilinéaire antisymétrique est de rang pair.*

4 Déterminer la forme polaire de la forme quadratique $q : \mathbb{C}^3 \longrightarrow \mathbb{C}$ définie par :

$$q(x) = 3\,x_1^2 - 2\,(1+i\,)\,x_2^2 - 2\,i\,x_1x_2 + x_1x_3 + (5-i\,)\,x_2x_3$$

Ecrire la matrice de q dans la base canonique de $\mathbb{C}^3$.

5 Soit q une forme quadratique et s sa forme polaire. Montrer que :

$$s(x,y) = \frac{1}{4}\Big(q(x+y) - q(x-y)\Big)$$

** **6** **Normes et normes euclidiennes**

1. Soit E un espace vectoriel sur K (non nécessairement de dimension finie) et q une forme quadratique sur E. Montrer que :
$$q(\lambda x) = \lambda^2 q(x)\,, \quad \forall x \in E, \ \forall \lambda \in K.$$
2. Soit $f : \mathbb{R}^n \longrightarrow \mathbb{R}$. On dit que f est positivement homogène de degré r si :
$$f(\lambda\, x) = \lambda^r\, f(x)\,, \quad \forall x \in \mathbb{R}^n\,, \ \forall\, \lambda \in \mathbb{R}\,, \ \ \lambda > 0.$$
   (a) Montrer que si f est positivement homogène de degré 0 et si elle est continue à l'origine, alors f est constante.
   (b) On suppose que les dérivées partielles de f existent. Montrer que si f est positivement homogène de degré r, ses dérivées partielles sont positivement homogènes de degré $r-1$.
   (c) En déduire le résultat suivant :
   Soit $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ telle que $f(\lambda\, x) = \lambda^2\, f(x)\,, \ \forall x \in \mathbb{R}^n\,, \forall \lambda \in \mathbb{R}$, $\lambda > 0$. Si f est de classe $\mathcal{C}^2$ (c'est-à-dire continue avec toutes ses dérivées partielles continues, jusqu'à l'ordre 2 inclus), alors f est une forme quadratique (et en particulier elle est de classe $\mathcal{C}^\infty$).
   (d) En déduire aussi le résultat suivant :
   Soit $\| \ \| : E \longrightarrow \mathbb{R}$ une norme sur un espace vectoriel réel de dimension finie. Alors $\| \ \|$ est une ***norme euclidienne*** *(c'est-à-dire elle provient d'un produit scalaire) si et seulement si son carré est une fonction de classe $\mathcal{C}^2$.*

7 Montrer que l'application $q : \mathbb{R}[x] \longrightarrow \mathbb{R}$ définie par $q(P) = \int_0^1 P(x)P''(x)\mathrm{d}x$ est une forme quadratique sur $\mathbb{R}[x]$.

8 Soit q une forme quadratique non dégénérée et s sa forme polaire. Montrer que : $s(x,z) = s(y,z)\,, \ \forall z \in E \implies x = y$.

9 Déterminer les vecteurs isotropes des formes quadratiques $q : \mathbb{R}^3 \longrightarrow \mathbb{R}$ suivantes et vérifier que $N(q) \subset \mathcal{I}(q)$:

a) $q(x) = x_1^2 + 3\,x_2^2 - 8\,x_3^2 - 4\,x_1x_2 + 2\,x_1x_3 - 10\,x_2x_3$
b) $q(x) = x_1^2 + 2\,x_2^2 - 2\,x_3^2 + 2\,x_1x_2 + 2\,x_1x_3$

10 Déterminer une base orthogonale pour les formes quadratiques $q : \mathbb{R}^3 \longrightarrow \mathbb{R}$ suivantes :

a) $q(x) = x_1^2 + 4\,x_2^2 + 9\,x_3^2 + 2\,x_1x_2 + 6\,x_2x_3$
b) $q(x) = x_1^2 + 3\,x_2^2 + 8\,x_3^2 - 4\,x_1x_2 + 6\,x_1x_3 - 10\,x_2x_3$
c) $q(x) = 5\,x_1x_2 + 6\,x_1x_3 + 2\,x_2x_3$

Même question pour $q : \mathbb{R}^4 \longrightarrow \mathbb{R}$ définie par :

d) $q(x) = x_1^2 + x_2^2 + x_3^2 - 2\,x_4^2 - 2\,x_1x_2 - 2\,x_1x_3 - 2\,x_1x_4 + 2\,x_2x_3 - 4\,x_2x_4$

* **11** Démontrer la proposition 9.13 : *les formes linéaires obtenues par la réduction en carrés de Gauss sont indépendantes.*

12

1. Déterminer la signature des formes quadratiques de l'exercice 10., ainsi qu'une base orthonormée, si elle existe.

2. On considère les formes quadratiques de l'exercice 10. comme formes quadratiques sur $\mathbb{C}^3$ (respect. : $\mathbb{C}^4$). Dans quel cas a-t-on une base orthonormée ? En déterminer une, le cas échéant.

13 Soit :
$$q : \begin{array}{ccc} \mathcal{M}_2(\mathbb{R}) & \longrightarrow & \mathbb{R} \\ A & \longmapsto & \text{dét}\, A \end{array}$$

Montrer que q est une forme quadratique. Déterminer son rang, sa signature, une base orthogonale et les vecteurs isotropes.

14 On définit sur $\mathbb{R}_2[x]$ la forme quadratique :
$$q(P) = \int_0^1 P(x)\, P''(x) \mathrm{d}x$$
(cf. exercice 7.). Déterminer le rang, la signature, le noyau, les vecteurs isotropes et une base orthogonale.

15

1. Soit $q : \mathbb{R}^3 \longrightarrow \mathbb{R}$ définie dans la base canonique par :
$$q(x) = x_1^2 - x_3^2 + 2\, x_1 x_2 + 2\, x_2 x_3,$$
et F le sous-espace vectoriel de $\mathbb{R}^3$ engendré par $v = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$.
Déterminer $F^\perp$ et vérifier que $F^\perp \supset N$.
Déterminer $F^{\perp\perp}$ et vérifier que $F^{\perp\perp} = F + N$.
2. Mêmes questions pour $q : \mathbb{R}^4 \longrightarrow \mathbb{R}$ définie par :
$$q(x) = x_1^2 + 2\, x_2^2 + 2\, x_1 x_2 - 2\, x_2 x_3 - 2\, x_2 x_4 + 4\, x_3 x_4$$
et F défini par les équations
$$\begin{cases} x_1 + x_2 - x_3 + x_4 = 0 \\ x_2 - x_4 = 0 \end{cases}$$

16 Soit $\mathcal{M}_2(\mathbb{R})$ muni de la forme quadratique de l'exercice 13. et $F = \{A \in \mathcal{M}_2(\mathbb{R}) \mid \mathrm{Tr} A = 0\}$. Déterminer $F^\perp$. A-t-on $\mathcal{M}_2(\mathbb{R}) = F \oplus F^\perp$?

17 Soit $\mathbb{R}_2[x]$ muni de la forme quadratique de l'exercice 14. et F le sous-espace engendré par $Q(x) = 1 - 2\, x - 15\, x^2$. A-t-on $\mathbb{R}_2[x] = F \oplus F^\perp$?

* **18** Soit (E, q) un espace vectoriel muni d'une forme quadratique, et F, G deux sous-espaces vectoriels de E.

1. Montrer
   (a) $F \subset G \Rightarrow G^\perp \subset F^\perp$
   (b) $(F + G)^\perp = F^\perp \cap G^\perp$
   (c) $F^\perp + G^\perp \subset (F \cap G)^\perp$.
2. Montrer que si q est non dégénérée, dans (c) on a l'égalité.
3. Donner un exemple où, dans (c), on a l'inclusion stricte.

19 Soit $q : \mathbb{R}^4 \longrightarrow \mathbb{R}$ définie par :
$$q(x) = x_1^2 + 2\, x_1 x_2 + 2\, x_1 x_3 + 2\, x_1 x_4 + 4\, x_2 x_3 + 4\, x_3 x_4.$$
Montrer que tout hyperplan F d'équation $a\, x_1 + x_2 + b\, x_3 + x_4 = 0$ est isotrope.

* **20** Tout sous-espace isotrope contient des vecteurs isotropes non nuls (il suffit de prendre $v \in F \cap F^\perp$, $v \neq 0$). Construire un exemple de sous-espace non isotrope contenant des vecteurs isotropes non nuls.

* **21** **Sous-espaces totalement isotropes**

1. Soit (E, q) un espace vectoriel muni d'une forme quadratique. Un sous-espace vectoriel F de E est dit *totalement isotrope* si $s|_F = 0$, s étant la forme polaire de q. Montrer que les propriétés suivantes sont équivalentes :

   a) F est totalement isotrope
   b) $F \subset \mathcal{I}$, $\mathcal{I}$ étant le cône isotrope
   c) $F \subset F^{\perp}$

   (Noter que le noyau N est un sous-espace totalement isotrope, d'après c))

2. Donner l'exemple d'une forme quadratique non nulle dans $\mathbb{R}^3$ admettant un sous-espace totalement isotrope de dimension 2 non trivial, c'est-à-dire différent du noyau.

* **22** Construire une matrice symétrique non diagonale $A \in \mathcal{M}_3(\mathbb{R})$ ayant une valeur propre strictement positive, une valeur propre strictement négative et une valeur propre nulle.

* **23** On considère l'espace euclidien $\mathbb{R}^3$ muni du produit scalaire canonique et soit $f \in \operatorname{End}(\mathbb{R}^3)$ définie, dans la base canonique, par :

$$A = M(f)_{e_i} = \frac{1}{11}\begin{pmatrix} -9 & 6 & 2 \\ 6 & 7 & 6 \\ 2 & 6 & -9 \end{pmatrix}$$

Montrer que f est une transformation orthogonale. Déterminer les valeurs propres de f, à l'aide de la trace et en remarquant que A est symétrique. En déduire la signature de la forme quadratique $q : \mathbb{R}^3 \longrightarrow \mathbb{R}$ qui, dans la base canonique s'écrit :

$$q(x) = -9\,x_1^2 + 7\,x_2^2 - 9\,x_3^2 + 12\,x_1\,x_2 + 4\,x_1\,x_3 + 12\,x_2\,x_3$$

** **24** **Réduction simultanée de deux formes quadratiques**
(*Généralisation du théorème 9.24*)

1. Soient s et t deux formes bilinéaires symétriques sur un espace vectoriel E de dimension finie, *s étant supposée non dégénérée.* Montrer qu'il existe un unique endomorphisme f de E tel que

$$s\,(\,f(x), y\,) = t\,(x, y\,)\,, \quad \forall x, y \in E \tag{1}$$

   Montrer que l'on a :

$$s\,(\,f(x), y\,) = s(\,x, f(y)\,)\,, \quad \forall x, y \in E \tag{2}$$

2. Montrer que, si $\lambda \neq \mu$, les sous-espaces propres pour f, E_λ et E_μ, sont orthogonaux pour s et en déduire qu'ils le sont aussi pour t.
3. Montrer la propriété suivante :

   Il existe une base orthogonale à la fois pour s et pour t si et seulement si f est diagonalisable.

4. Application. Soient :

$$\begin{aligned} q_s(x) &= x_1^2 + x_2^2 - x_3^2 - 2\,x_1x_2 + 2\,x_2x_3 \\ q_t(x) &= (1+a)\,x_1^2 + x_2^2 + (1+a)\,x_3^2 - 2\,x_1x_2 - 2\,(1+a)\,x_1\,x_3 + 2\,x_2x_3 \end{aligned}$$

   Pour quelles valeurs de a existe-t-il une base orthogonale à la fois pour q_s et q_t ? Déterminer une telle base, lorsqu'elle existe.

25 Soit E un espace vectoriel de dimension finie muni d'une forme quadratique *non dégénérée.* Montrer que :

$$(\operatorname{Im} f)^{\perp} = \operatorname{Ker} f^* \quad \text{et} \quad (\operatorname{Ker} f)^{\perp} = \operatorname{Im} f^*$$

26

1. Déterminer $\mathrm{O}(q)$ où $q : \mathbb{R}^2 \longrightarrow \mathbb{R}$ est définie, dans la base canonique, par $q(x) = 2\,x_1x_2$.
2. (**Transformation de Lorentz en dimension 2**) Même question pour $q : \mathbb{R}^2 \longrightarrow \mathbb{R}$ définie par $q(x) = x_1^2 - x_2^2$

INDICATIONS

1 $B = M(b)_{e_i} = \begin{pmatrix} 1 & 0 & 6 \\ 2 & 1 & -3 \\ 3 & 1 & 3 \end{pmatrix}$. $N(b)$ est engendré par $\begin{pmatrix} -6 \\ 15 \\ 1 \end{pmatrix}$;

$\mathrm{rg}(b) = 2$; $b(z, w) = 0$.

$$B' := M(b)_{v_i} = {}^tPBP = \begin{pmatrix} 14 & 8 & 6 \\ 7 & 2 & 0 \\ 7 & 4 & 3 \end{pmatrix}.$$

2 $\tilde{N}(b)$ est engendré par $\begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}$.

3

1. $0 = \Omega(x + y, x + y) = \Omega(x, x) + \Omega(x, y) + \Omega(y, x) + \Omega(y, y)$.
2. Si $y = \lambda x$ $\quad \Omega(x, y) = \Omega(x, \lambda x) = \lambda\, \Omega(x, x) = 0$.
3. Il existe deux vecteurs v_1, v_2 tels que $\Omega(v_1, v_2) \neq 0$ (autrement $\Omega = 0$). Poser $u_1 = v_1$, $u_2 = \dfrac{v_2}{\Omega(v_1, v_2)}$.
4. On a $F \cap W = \{0\}$. Montrer que $E = F \oplus W$ en faisant une analyse et une synthèse. On trouve la décomposition de $x \in E$:
$$x = \underbrace{\Omega(x, u_2)u_1 - \Omega(x, u_1)\, u_2}_{\in F} + \underbrace{(x - \Omega(x, u_2)u_1 + \Omega(x, u_1)\, u_2)}_{\in W}$$
5. Utiliser le fait que Ω est antisymétrique.
6. On restreint Ω à W et on raisonne par récurrence.

4 $s(x, y) = 3\, x_1y_1 - 2\,(1 + i)\, x_2y_2 - i\, x_1y_2 - i\, x_2y_1 + \dfrac{1}{2}\, x_1\, y_3 + \dfrac{1}{2}\, x_3\, y_1$
$\qquad + \dfrac{5 - i}{2}\, x_2\, y_3 + \dfrac{5 - i}{2}\, x_3\, y_2.$

$$M(q)_{e_i} = \begin{pmatrix} 3 & -i & 1/2 \\ -i & -2\,(1 + i) & \dfrac{5 - i}{2} \\ 1/2 & \dfrac{5 - i}{2} & 0 \end{pmatrix}$$

5 $\frac{1}{4}\,[\, q(x + y) - q(x - y)\,] = \frac{1}{4}\,[\, s(x + y,\ x + y) - s(x - y,\ x - y)] = \cdots$

6

1. $q(\lambda x) = s(\lambda x, \lambda x) = \cdots$
2. (a) La condition d'homogénéité assure que f est constante sur toute demi-droite issue de l'origine. La continuité de f à l'origine implique que f prend la même valeur sur chacune de ces demi-droites (cf. Fig. 3).

(b) Si $f(\lambda\, x) = \lambda^r\, f(x)$, $\forall \lambda > 0$ on a, en dérivant par rapport à x_k : $\dfrac{\partial f}{\partial x^k}\,(\lambda x)\,\lambda = \lambda^2\, \dfrac{\partial f}{\partial x^k}(x)$.

(c) Si f est positivement homogène de degré 2, ses dérivées secondes sont positivement homogènes de degré 0 ; étant continues à l'origine elles sont des fonctions constantes. En intégrant, on voit que f est un polynôme de degré 2. Il s'agit d'un polynôme homogène puisque $f(\lambda\, x) = \lambda^2\, f(x)$, $\forall \lambda > 0$.

(d) Si $\|\ \|$ est une norme, son carré est une fonction positivement homogène de degré 2.

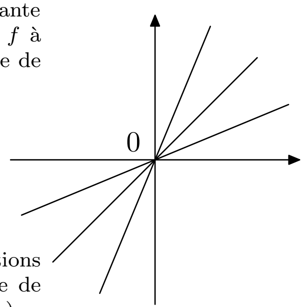


Figure 3

7 Poser $s(P,Q) = \dfrac{1}{2}\displaystyle\int_0^1 \Big(P(x)\,Q''(x) + P''(x)\,Q(x)\Big)\,dx.$

8 $s(x-y,z)=0\ \ \forall z\in E$ signifie que $x-y\in N=\{0\}$ et, par conséquent $x=y$.

9 a) $q(x)=(x_1-2\,x_2+x_3)^2-(x_2+3\,x_3)^2$. $\mathcal{I}(q)$ est la réunion des deux plans $x_1-2\,x_2+x_3=x_2+3\,x_3$ et $x_1-2\,x_2+x_3=-(x_2+3\,x_2)$ c'est-à-dire des plans $\pi_1\ :\ x_1-3\,x_2-2\,x_3\ =\ 0$ et $\pi_2\ :\ x_1-x_2+4\,x_3=0$.

$N(q)$ est engendré par : $\begin{pmatrix} 7\\ 3\\ -1\end{pmatrix}$.

b) $q(x)=(x_1+x_2+x_3)^2+(x_2-x_3)^2-4\,x_3^2$. $\mathcal{I}(q)$ est le cône défini par

$$x_3=\pm\frac{1}{2}\sqrt{(x_1+x_2+x_3)^2+(x_2-x_3)^2}$$

$N(q)=\{0\}$.

10 a) $q(x)=(x_1+x_2)^2+3\,(x_2+x_3)^2+6\,x_3^2$;

$v_1=\begin{pmatrix}1\\0\\0\end{pmatrix}$, $v_2=\begin{pmatrix}-1\\1\\0\end{pmatrix}$, $v_3=\begin{pmatrix}3\\-3\\1\end{pmatrix}$.

b) $q(x)=(x_1-2\,x_2+3\,x_3)^2-(x_2-x_3)^2+0\cdot x_3^2$.

Résoudre le système : $\begin{cases} x_1' = x_1-2\,x_2+3x_3\\ x_2' = x_2-x_3\\ x_3' = x_3\end{cases}$

c) On est amené à résoudre le système : $\begin{cases} x_1' = 5\,x_1+5\,x_2+8\,x_3\\ x_2' = -5\,x_1+5\,x_2+4\,x_3\\ x_3' = x_3\end{cases}$

d) $q(x)=(x_1-x_2-x_3-x_4)^2-3\,(x_2+\dfrac{1}{3}x_3+x_4)^2+3\,(x_2+\dfrac{1}{3}x_3)^2+0\cdot x_2^2.$

Résoudre le système :

$$\begin{cases} x_1' = x_1-x_2-x_3-x_4\\ x_2' = x_2+\frac{1}{3}x_3+x_4\\ x_3' = x_2+\frac{1}{3}x_3\\ x_4' = x_2\end{cases}$$

Attention : ne pas poser $x_4'=x_4$ (il faut que les quatre formes x_1', x_2', x_3', x_4' soient indépendantes).

11 Supposons (avec les notations de page 307) que :

$$q(x)=\tfrac{1}{4}\Big[\Big(\varphi_1(x)+\varphi_2(x)\Big)^2-\Big(\varphi_1(x)-\varphi_2(x)\Big)^2\Big]+\lambda_3\,\psi_3(x)^2+\cdots+\lambda_p\,\psi_p(x)^2$$

Puisque φ_1 ne contient pas x_i $(i\neq j)$ et contient x_j, alors que φ_2 ne contient pas x_j tout en contenant x_i, φ_1 et φ_2 sont indépendantes et $\varphi_1+\varphi_2$, $\varphi_1-\varphi_2$ engendrent un sous-espace F de E^* de dimension 2. Soit $G=\mathrm{Vect}\{\psi_3,\dots,\psi_p\}$; puisque les formes $\psi_2,\dots\psi_p$ ne contiennent pas les variables x_i, x_j, F et G sont en somme directe. Raisonner par récurrence.

12 1. a) : (3,0) b) : (1,1) c) : (1,2) d) : (2,1)
On a une base orthonormée uniquement dans le cas a). Si $\{v_1,v_2,v_3\}$ est la base orthogonale trouvée dans l'exercice 10. a) , on a :

$$M(q)_{v_i} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 6 \end{pmatrix}.$$

Donc $q(v_1) = 1$, $q(v_2) = 3$, $q(v_3) = 6$. Une base orthonormée est donc :

$$\varepsilon_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \varepsilon_2 = \frac{1}{\sqrt{3}} \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \quad \varepsilon_3 = \frac{1}{\sqrt{6}} \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}.$$

2. On a une base orthonormée dans les cas a) et c). Pour a) même résultat qu'en 1. Pour c) on trouve, par exemple, $\varepsilon_1 = \dfrac{1}{\sqrt{20}} v_1$, $\varepsilon_2 = \dfrac{1}{i\sqrt{20}} v_2$, $\varepsilon_3 = \dfrac{1}{i}\sqrt{\dfrac{5}{12}}\, v_3$, $\{v_1, v_2, v_3\}$ étant la base ci-dessus.

13 Si $A = \begin{pmatrix} x_1 & x_2 \\ x_3 & x_4 \end{pmatrix}$, $q(A) = x_1x_4 - x_2x_3$. q est donc une forme quadratique. Par la réduction en carrés, on trouve :

$$q(x) = \frac{1}{4}\Big[(x_1 + x_4)^2 - (x_1 - x_4)^2\Big] - \frac{1}{4}\Big[(x_2 + x_3)^2 - (x_2 - x_3)^2\Big].$$

Base orthogonale : $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$.

Vecteurs isotropes : les matrices A telles que $\text{dét} A = 0$.

14 Dans la base $e_1 = 1$, $e_2 = x$, $e_3 = x^2$, on a :

$$M(q)_{e_i} = \begin{pmatrix} 0 & 0 & 2 \\ 0 & 0 & 1 \\ 2 & 1 & 4/3 \end{pmatrix} \qquad \operatorname{rg} q = 2.$$

Si $P = a_1 + a_2\, x + a_3\, x^2$, alors : $q(P) = \dfrac{4}{3}(a_3^2 + 3\, a_1 a_3 + 6\, a_2\, a_3)$

$$= \frac{4}{3}\Big[\Big(a_3 + \frac{3}{2}a_1 + 3\, a_2\Big)^2 - \Big(\frac{3}{2}a_1 + 3\, a_2\Big)^2\Big].$$

Le noyau est engendré par $P = 2 - x$. Les vecteurs isotropes sont donnés par les deux familles de polynômes $P_1(x) = a_1 + a_2\, x$ et $P_2(x) = a_1 + a_2\, x - 3\,(a_1 + 2\, a_2)\, x^2$.
Une base orthogonale est donnée par :

$$\Big\{ x^2, \quad \frac{2}{3} - x^2, \quad -2 + x \Big\}.$$

15 1. $M(q)_{e_i} = S = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & -1 \end{pmatrix}$. On a :

$$Y \in F^\perp \iff {}^tX\, S\, Y \;=\; 0,\ \forall X \in F.$$

Or : $X \in F \iff X = \begin{pmatrix} \lambda \\ \lambda \\ \lambda \end{pmatrix}$, donc :

$${}^tX\, S\, Y = (\,\lambda\ \ \lambda\ \ \lambda\,) \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & -1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = 2\,\lambda\,(y_1 + y_2).$$

Ainsi

$${}^tXSY = 0 \iff \forall X \in F \iff 2\,\lambda\,(y_1 + y_2) = 0\ ,\ \forall \lambda \in \mathbb{R} \iff y_1 + y_2 = 0$$

ce qui définit l'équation de $F^\perp$.

N est engendré par $\begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}$; on a bien $N \subset F^\perp$.

$F^{\perp\perp} = \{\, Y \mid {}^tXSY = 0\ , \forall X \in F^\perp \,\}$ et $X \in F^\perp \iff X = \begin{pmatrix} -\lambda \\ -\lambda \\ \mu \end{pmatrix}$.

Donc : ${}^tXSY = 0\ ,\forall X \in F^{\perp} \Longleftrightarrow (\lambda+\mu)(y_2 - y_3) = 0\ ,\forall \lambda, \mu \in \mathbb{R}$.
Ainsi $F^{\perp\perp}$ est le plan défini par $y_2 - y_3 = 0$.

2. Même méthode. On trouve $X \in F \Longleftrightarrow X = \begin{pmatrix} \lambda - 2\mu \\ \mu \\ \lambda \\ \lambda \end{pmatrix}$ et :

$$F^{\perp} = \begin{cases} y_1 \qquad\qquad\quad - 2\,y_4 = 0 \\ y_1 + y_2 - y_3 + y_4 = 0 \end{cases}$$

16 $F^{\perp} = \left\{ A \in \mathcal{M}_2(\mathbb{R}) \mid A = \begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix},\ a \in \mathbb{R} \right\}$. F est non isotrope.

17 Q est isotrope.

18

1. Faire une démonstration directe. Par exemple pour (b) :
   $x \in (F+G)^{\perp} \Longrightarrow s(\,f+g\,,\,x) = 0\,,\ \forall f \in F,\ \forall g \in G$. En particulier pour $f = 0$ on a $x \in G^{\perp}$ et pour $g = 0$ on a $x \in F^{\perp}$.
   Réciproquement, si $x \in F^{\perp} \cap G^{\perp}$, on a $s(f\,,\,x) = 0\ ,\forall f \in F$, et $s(g\,,\,x) = 0\ ,\forall g \in G$. D'où $s(f+g\,,\,x) = 0\ \forall f \in F$ et $\forall g \in G$, c'est-à-dire $x \in (F+G)^{\perp}$.
2. Si q est non dégénérée, on a $F^{\perp\perp} = F$ pour tout sous-espace F. Remplacer dans (b) F par $F^{\perp}$ et G par $G^{\perp}$.
3. Il faut que q soit dégénérée. Le cas le plus simple non trivial est $q : \mathbb{R}^2 \to \mathbb{R}$ définie par $q(x) = x_1^2$. Prendre F et G tels que $F \cap G = \{0\}$ [d'où : $F \cap G)^{\perp} = \mathbb{R}^2$] et tels que $F^{\perp} = G^{\perp}$ et $\dim F^{\perp} = 1$ [d'où : $\dim(F^{\perp} + G^{\perp}) = 1$]. Par exemple :
   $$F = \mathrm{Vect}\{v\} \text{ et } G = \mathrm{Vect}\{w\}\ ,\ \text{avec } v = \begin{pmatrix} 1 \\ 0 \end{pmatrix},\ w = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

19 $F^{\perp}$ est défini par

$$\begin{cases} (1-a)\,x_1 + x_2 + (1-2\,a)\,x_3 + \ x_4 = 0 \\ (1-b)\,x_1 + 2\,x_2 \qquad\quad - 2\,b\,x_3 + 2\,x_4 = 0 \end{cases}$$

donc $\dim F^{\perp} \geq 2$. Or :

$$4 \geq \dim(F + F^{\perp}) = \dim F + \dim F^{\perp} - \dim(F \cap F^{\perp}),$$

donc : $\dim(F \cap F^{\perp}) \geq \dim F^{\perp} - 1 \geq 1$.

20 Il suffit de considérer un espace E muni d'une forme quadratique non dégénérée et non définie. E lui-même répond à la question. Si l'on veut un exemple moins trivial (un sous-espace $F \underset{\neq}{\subset} E$) il faut que $\dim F \geq 2$, car tout sous-espace de dimension 1 est isotrope si et seulement si il est engendré par un vecteur isotrope. Prendre, par exemple : $E = \mathbb{R}^3$, $q(x) = x_1^2 + x_2^2 - x_3^2$ et F le plan d'équation $x_1 = 0$.

21

1. a) et b) sont équivalentes car $s|_F = 0$ est équivalent à $q|_F = 0$ ce qui signifie que $F \subset \mathcal{I}$. D'autre part
   $$F \subset F^{\perp} \iff \forall x \in F : x \perp y \quad \forall y \in F \iff s(x,y) \ = \ 0\,,\ \ \forall x, y \in F.$$
2. Pour que $\mathcal{I}$ contienne un plan vectoriel, il est nécessaire que $\operatorname{rang} q \neq 3$ [bien vérifier que pour la signature (2,1) ou (1,2) $\mathcal{I}$ ne contient pas de plan vectoriel]. Si $\operatorname{rang} q = 1$, on peut écrire, dans une certaine base, $q(x) = \pm\, x_1^2$ et alors $\mathcal{I} = N = \{$ plan défini par $x_1 = 0\,\}$; aussi le seul plan totalement isotrope est le noyau. Pour qu'il existe un plan isotrope non trivial, il faut donc que $\operatorname{rang} q = 2$. Si sign $q = (2,0)$ ou $(0,2)$, à un choix de base près : $q(x) = \pm(x_1^2 + x_2^2)$ et $\mathcal{I} = N$ est de dimension 1 ; il n'existe donc pas de plan totalement isotrope. Pour avoir un plan isotrope non trivial, il faut donc et il suffit que sign $q = (1,1)$, c'est-à-dire : $q(x) = x_1^2 - x_2^2$. Le plan d'équation $x_1 = x_2$ est totalement isotrope.

22 Il suffit de considérer une forme quadratique $q : \mathbb{R}^3 \longrightarrow \mathbb{R}$ de signature (1,1). Par exemple : $q(x) = (2\,x_1 - 3\,x_2 + x_3)^2 - 2\,(x_2 + x_3)^2$. La matrice de q est :

$$A = \begin{pmatrix} 4 & -6 & 2 \\ -6 & 7 & -5 \\ 2 & -5 & -1 \end{pmatrix} \quad \text{(on a : } \operatorname{Sp} A = \{0,\ 5-\sqrt{73}\,,\ 5+\sqrt{73}).$$

23 Puisque A est orthogonale et symétrique, $A^2 = I$, c'est-à-dire $f^2 = \mathrm{id}$. Par conséquent, si v appartient au plan de rotation, l'angle de rotation θ vaut $k\pi$, puisque $f \circ f(v) = v$. D'autre part, $\mathrm{Tr} f = -1$; donc : soit $2\cos\theta - 1 = -1$, soit $2\cos\theta + 1 = -1$.
Le premier cas donne $\cos\theta = 0$, ce qui est exclu, car cela signifierait que $f = \mathrm{id}$. Il reste le second cas, $\cos\theta = -1$, c'est-à-dire $\theta = \pi$. f est donc une rotation d'angle π et, par des considérations géométriques, $\mathrm{Sp}' f = \{1, -1, -1\}$. On en déduit $\mathrm{sign}(q) = (1, 2)$.

24

1. En notant A, S, T les matrices de f, s, t dans une base $\{e_i\}$, la relation (1) s'écrit : $SA = T$. Puisque s est non dégénérée : $A = S^{-1}T$. Donc si f existe elle est unique car sa matrice est $S^{-1}T$. Réciproquement, en définissant f par $M(f)_{e_i} = S^{-1}T$ et en remontant les calculs, on vérifie (1). La relation (2) vient de la symétrie de s.
2. Écrire la relation (2) pour $x \in E_\lambda$, $y \in E_\mu$. Utiliser (1) pour montrer que $x \underset{t}{\perp} y$.
3. Supposons f diagonalisable et soit $s_\lambda := s|_{E_\lambda}$ (s_λ est une forme bilinéaire sur E_λ). Dans chaque E_λ on prend une base orthogonale pour s_λ (elle sera aussi orthogonale pour s). Puisque f est diagonalisable, E est somme directe des E_λ et donc la collection de toutes ces bases est une base de E orthogonale pour s. D'après le résultat ci-dessus, elle est aussi orthogonale pour t.
Réciproquement, supposons qu'il existe une base orthogonale pour s et t. Dans cette base, avec les notations ci-dessus, S et T sont diagonales, donc $A = S^{-1}T$ est diagonale.
4. Application. $A = S^{-1}T = \begin{pmatrix} a & 0 & -a \\ -1 & 1 & 1 \\ a & 0 & -a \end{pmatrix}$
Cette matrice est diagonalisable si et seulement si $a = 0$. Pour avoir une base orthogonale à la fois pour s et pour t il faut prendre une base de vecteurs propres de A qui est orthogonale pour s. On trouve, par exemple :
$$v_1 = \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix} \qquad v_2 = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} \qquad v_3 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$

25 Utiliser l'identité $s\,(\,f(x), y\,) = s\,(\,x, f^*(y)\,)\,, \forall x, y \in E$:

$$f^*(y) = 0 \implies s(\,f(x), y\,) = 0 \ \forall x \in E \implies y \in (\operatorname{Im} f)^\perp$$

c'est-à-dire : $\operatorname{Ker} f^* \subset (\operatorname{Im} f)^\perp$. Raisonner sur les dimensions pour montrer l'égalité. Remplacer ensuite f par f^* et prendre l'orthogonal pour montrer la seconde identité.

26

1. Les matrices orthogonales sont les matrices $A \in \mathcal{M}_2(\mathbb{R})$ telles que ${}^tASA = S$ avec $S = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$. On trouve :
$$A = \begin{pmatrix} a & 0 \\ 0 & 1/a \end{pmatrix} \ \text{avec } a \in \mathbb{R}\backslash\{0\} \ \text{ et } \ A = \begin{pmatrix} 0 & b \\ 1/b & 0 \end{pmatrix} \ \text{avec } b \in \mathbb{R}\backslash\{0\}$$
2. $A = \begin{pmatrix} \varepsilon \operatorname{ch} a & \operatorname{sh} a \\ \varepsilon\varepsilon' \operatorname{sh} a & \varepsilon' \operatorname{ch} a \end{pmatrix}$ avec $a \in \mathbb{R}$ et $\varepsilon, \varepsilon' = \pm 1$.

Chapitre 10

Formes hermitiennes

Dans ce chapitre nous établissons pour les formes hermitiennes la même théorie que nous avons développée pour les formes bilinéaires symétriques. Les résultats sont tout à fait analogues et les démonstrations se font de la même manière. Le lecteur est invité a se rapporter souvent à la table en Appendice A.13.

10.1 Rang et noyau d'une forme hermitienne

Rappelons (cf. définition 8.1) qu'une forme hermitienne sur un espace vectoriel complexe E est une application $h : E \times E \longrightarrow \mathbb{C}$ antilinéaire dans le premier argument et linéaire dans le second argument, vérifiant en plus la propriété dite de «symétrie hermitienne» :

$$h(x, y\,) = \overline{h(y, x\,)}$$

Commençons par considérer le cas où E est de dimension finie. Comme nous l'avons vu (cf. page 280), si $\{e_i\}$ et $\{e'_i\}$ sont deux bases de E et $H = M(h)_{\,e_i}\,, H' = M(h)_{\,e'_i}$, alors :

$$H' = {}^t\overline{P}\,H\,P$$

où $P = P_{e_i \to e'_i}$. Cette formule montre, comme dans le cas des applications bilinéaires, que :

Proposition 10.1 – *Le rang de la matrice $M(h)_{e_i}$ ne dépend pas du choix de la base $\{e_i\}$. Ce rang est dit **rang** de la forme hermitienne.*

En revanche, tout comme pour les formes bilinéaires, le déterminant de $M(h)_{e_i}$ dépend du choix de la base, mais s'il est non nul dans une base il est non nul en toute base.

Définition 10.2 – *Une forme hermitienne (sur un espace vectoriel E de dimension finie sur $\mathbb{C}$) est dite **non dégénérée** si elle est de rang maximum (égal à la dimension de E), c'est-à-dire si :* dét $M(h)_{e_i} \neq 0$ *($\{e_i\}$ étant une base arbitraire de E).*

Exemple – Soit $h : \mathbb{C}^2 \longrightarrow \mathbb{C}$ définie dans la base canonique $\{e_i\}$ par :

$$h\,(x, y) = 2\,\overline{x}_1\,y_1 - 3\,i\,\overline{x}_1 y_2 + 3\,i\,\overline{x}_2 y_1 + 3\,\overline{x}_2\,y_2$$

On a :

$$M(h)_{e_i} = \begin{pmatrix} 2 & -3\,i \\ 3\,i & 3 \end{pmatrix}$$

h est de rang 2 et donc non dégénérée.

Définition 10.3 – *On appelle* **noyau** *de la forme hermitienne h le sous-espace vectoriel $N(h)$ de E défini par :*

$$N(h) := \{\, x \in E \,|\, h(x, y) = 0\,, \ \ \forall y \in E\,\}$$

(cf. définition 9.3).

NOTA – A cause de la symétrie hermitienne, on a aussi

$$N(h) = \{\, y \in E \,|\, h(x, y\,) = 0\,, \ \ \forall x \in E\,\}.$$

Le noyau de h est en fait le noyau de la matrice qui représente h (dans une base quelconque).

En effet, soit $H = M(h)_{e_i}$ on a $h(x, y) = {}^t\overline{X}\,H\,Y$; donc :

$$y \in \operatorname{Ker} h \ \text{ si et seulement si } {}^t\overline{X}\,H\,Y = 0\,, \ \forall\, X \in \mathcal{M}_{1,n}(\mathbb{C})$$

c'est-à-dire : $y \in \operatorname{Ker} h$ si et seulement si $H\,Y = 0$.

On en déduit, comme pour la proposition 9.4 :

Proposition 10.4 – *Soit h une forme hermitienne sur un espace vectoriel E de dimension finie. Alors :*

1. *Le noyau de h est le noyau de la matrice qui représente h dans une base (quelconque).*
2. $\dim E = \operatorname{rg} h + \dim N(h)$

Interprétation du rang et du noyau. Rang et noyau en dimension infinie.

Comme pour les formes bilinéaires, (cf. définition 9.3) le rang et le noyau peuvent être interprétés comme le rang et le noyau d'une certaine application linéaire.
Notons

$$E^{\overline{*}} := \left\{ \varphi : E \to \mathbb{C} \ \text{ telles que : } \left\{ \begin{array}{rcll} \varphi(x+y) & = & \varphi(x) + \varphi(y) & \\ \varphi(\lambda x) & = & \overline{\lambda}\,\varphi(x)\,, & \forall x, y \in E. \end{array} \right. \right\}$$

Les éléments de $E^{\overline{*}}$ sont dits *formes antilinéaires.* On voit facilement que $E^{\overline{*}}$ est un espace vectoriel sur $\mathbb{C}$ pour les lois que l'on définit habituellement sur l'espace des applications (exemple 4. page 5)
Si E est de dimension n et $\{e_i\}$ est une base de E, on montre, comme pour le théorème 3.35, qu'il existe une unique base $\{\varphi_i\}_{i=1\cdots n}$ de $E^{\overline{*}}$, dite *base antiduale*, telle que :

$$\varphi_i(e_j) = \delta_{ij}.$$

Il s'agit des applications

$$\begin{array}{rccc} \varphi_k : & E & \longrightarrow & \mathbb{C} \\ & x = \sum_{i=1}^n x_i\, e_i & \longmapsto & \overline{x}_k \end{array}$$

(cf. exercice 2). On vérifie, exactement comme pour la proposition 9.2 page 297 :

Proposition 10.5 – *Soit h une forme hermitienne sur un espace vectoriel E de dimension finie sur $\mathbb{C}$. On associe à h l'application linéaire*

$$\begin{array}{rccc} j(h) : & E & \longrightarrow & E^{\overline{*}} \\ & y & \longmapsto & h(\cdot, y) \end{array} \qquad \text{où} \qquad \begin{array}{rccc} h(\cdot, y) : & E & \longrightarrow & \mathbb{C} \\ & x & \longmapsto & h(x, y) \end{array}$$

Si $\{e_i\}$ est une base de E et $\{\varphi_i\}$ sa base antiduale, on a : $M(h)_{e_i} = M\Big(j(h)\Big)_{e_i, \varphi_j}$.

Ces remarques justifient la définition suivante, valable en dimension infinie.

Définition 10.6 – *Soit E un espace vectoriel sur $\mathbb{C}$ (non nécessairement de dimension finie) et h une forme sesquilinéaire (cf. page 274). On appelle* **rang de** h *le rang de l'application*

$$j(h) : \begin{array}{ccc} E & \longrightarrow & \overline{E^*} \\ y & \longmapsto & h(\cdot, y) \end{array} \quad \text{où} \quad h(\cdot, y) : \begin{array}{ccc} E & \longrightarrow & \mathbb{C} \\ x & \longmapsto & h(x, y) \end{array}$$

On appelle **noyau de** h *le noyau de $j(h)$:*

$$N(h) := \{y \in E \mid h(x, y) = 0\,, \quad \forall x \in E\}$$

On dit que h est **non dégénérée** *si $j(h)$ est injective c'est-à-dire si $N(h) = 0$, ce qui signifie que : $h(x, y) = 0\,, \ \forall x \in E \implies y = 0$.*

Notons enfin que *si h est une forme hermitienne définie positive* [1]*, alors h est non dégénérée.*

En effet si x est tel que $h(x, y) = 0\ , \forall y \in E$, on a en particulier $h(x, x) = 0$, donc $x = 0$.

Exercices 1. 2. 3.

10.2 Orthogonalité. Vecteurs isotropes

Définition 10.7 – *Soit h une forme hermitienne sur E.*

1. *$x, y \in E$ sont dits* **orthogonaux** *pour h si $h(x, y) = 0$.*
2. *$x \in E$ est dit* **isotrope** *si $\tilde{q}(x) = 0$, où : $\tilde{q}(x) = h(x, x)$. On note :*
$$\mathcal{I} = \{x \in E \mid \tilde{q}(x) = 0\}$$
3. *Soit $A \subset E$; on pose : $A^\perp = \{\, x \in E \mid h(x, a) = 0\ , \quad \forall a \in A\,\}$. $A^\perp$ est un sous-espace vectoriel de E dit* **orthogonal** *de A par h.*

Comme pour les formes quadratiques (cf. 9.19, 9.21, 9.23) on montre :

Proposition 10.8 –

1. $$\{0\}^\perp = E \quad , \quad E^\perp = N(h) \ ;$$
$$N(h) \subset \mathcal{I} \quad , \quad A^\perp \subset N(h) \ , \forall A \subset E.$$
2. *Si E est de dimension finie, pour tout sous-espace vectoriel F de E, on a :*
   a) $\dim E = \dim F + \dim F^\perp - \dim(F \cap N)$ *;*
   b) $F^{\perp\perp} = F + N$ où $N = N(h)$.
3. *Un sous-espace vectoriel F de E est dit* **isotrope** *si $F \cap F^\perp \neq \{0\}$.*
Si E est de dimension finie, on a :
$$E = F \oplus F^\perp \iff F \text{ est non isotrope.}$$

(pour la démonstration de 2. a), cf. exercice 4.).

Exercice 4.

[1] c'est-à-dire : $h(x, x) \geq 0\,, \ \forall x \in E$ et $\Big(h(x, x) = 0\Big) \iff x = 0$ (c'est-à-dire, en fait, si h définit un produit scalaire hermitien, cf. définition 8.2).

10.3 Bases orthogonales et classification des formes hermitiennes

Comme pour le produit scalaire hermitien (cf. définition 8.7), on pose la définition :
Définition 10.9 – *Une base $\{e_i\}$ est dite* ***orthogonale*** *pour la forme hermitienne h si $h(e_i, e_j) = 0$ pour $i \neq j$. Elle est dite* ***orthonormée*** *si $h(e_i, e_j) = \delta_{ij}$.*
Il est clair que $\{e_i\}$ est une base orthogonale si et seulement si :

$$M(h)_{e_i} = \begin{pmatrix} a_1 & 0 & \cdots & 0 \\ 0 & a_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & a_n \end{pmatrix} \qquad (a_i \in \mathbb{R})$$

ou encore :

$$\tilde{q}(x) \equiv h(x,x) = a_1\,|x_1|^2 + \cdots + a_n\,|x_n|^2 \qquad (\text{ si } x = \textstyle\sum_{i=1}^n x_i e_i\ ).$$

De même, $\{e_i\}$ est une base orthonormée si et seulement si :

$$M(h)_{e_i} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & 1 \end{pmatrix}$$

ou encore :

$$\tilde{q}(x) = |x_1|^2 + \cdots + |x_n|^2 \qquad (\text{ si } x = \textstyle\sum_{i=1}^n x_i e_i).$$

On voit immédiatement que les bases orthonormées n'existent pas toujours : il existe une base orthonormée si et seulement si h est définie positive, c'est-à-dire si h définit un produit scalaire hermitien. En revanche, il existe toujours des bases orthogonales :

Théorème 10.10 – *Soit $E \neq \{0\}$ un espace vectoriel de dimension finie sur $\mathbb{C}$ muni d'une forme hermitienne h. Il existe alors toujours des bases orthogonales pour h.*

La démonstration est la même que pour les formes quadratiques (cf. théorème 9.12). De la même manière on montre aussi (cf. théorème 9.16) :

Théorème de Sylvester . 10.11 – *Soient E un espace vectoriel de dimension finie sur $\mathbb{C}$ et h une forme hermitienne sur E. Il existe alors une base (orthogonale) $\{e_i\}$ de E telle que :*

$$M(q)_{e_i} = \begin{pmatrix} \boxed{\begin{matrix} 1 & & \\ & \ddots & \\ & & 1 \end{matrix}} & & 0 \\ & \boxed{\begin{matrix} -1 & & \\ & \ddots & \\ & & -1 \end{matrix}} & \\ 0 & & \begin{matrix} 0 & & \\ & \ddots & \\ & & 0 \end{matrix} \end{pmatrix}$$

c'est-à-dire telle que, si $x = \sum_{i=1}^{n} x_i e_i$ *:*

$$\tilde{q}(x) = |x_1|^2 + \cdots + |x_p|^2 - |x_{p+1}|^2 - \cdots - |x_r|^2$$

où $r = \mathrm{rg}(h)$. *L'entier* p *ne dépend pas du choix de la base. Le couple* $(p, r-p)$ *est dit* ***signature*** *de* h *.*

On en déduit (cf. corollaire 9.17) :

Corollaire 10.12 *Soit* h *une forme hermitienne sur un espace vectoriel complexe* E *de dimension finie* n*. Alors :*

h est définie positive (c'est-à-dire E est hermitien)	$\Longleftrightarrow$	$\mathrm{sign}(h) = (n, 0)$ $\Updownarrow$ *il existe des bases orthonormées*
h est définie négative (c'est-à-dire $h(x) \leq 0$ *et* $h(x) = 0 \Longrightarrow x = 0$*)*	$\Longleftrightarrow$	$\mathrm{sign}(h) = (0, n)$
h est non dégénérée (c'est-à-dire $N(h) = \{0\}$*)*	$\Longleftrightarrow$	$\mathrm{sign}(h) = (p, n-p)$

Exercices 5. 6. 7.

10.4 Groupe unitaire associé à une forme hermitienne

Le but de ce paragraphe est d'étudier les endomorphismes qui conservent une forme hermitienne (non dégénérée). Il s'agit donc de l'analogue des transformations orthogonales étudiées au chapitre précédent (cf. page 316). Commençons par définir la notion d'endomorphisme adjoint.

Proposition 10.13 – *Soit* (E, h) *un espace vectoriel complexe de dimension finie muni d'une forme hermitienne* ***non dégénérée.*** *Pour tout endomorphisme* f *de* E *il existe un et un seul endomorphisme* f^* *de* E*, dit* ***adjoint*** *de* f *relativement à* h*, tel que :*

$$h\big(f(x), y\big) = h\big(x, f^*(y)\big), \quad \forall x, y \in E.$$

Si $\{e_i\}$ *est une base de* E*,* $H = M(h)_{e_i}$*,* $A = M(f)_{e_i}$ *et* $A^* = M(f^*)_{e_i}$*, on a :*

$$A^* \ = \ H^{-1}\, {}^t\overline{A}\ H$$

La démonstration est immédiate : la condition équivaut à ${}^t(\overline{A}X)\,H\,Y = {}^t\,\overline{X}\,H\,A^*\,Y$, $\forall X, Y \in \mathcal{M}_{n,1}(\mathbb{C})$, d'où on obtient facilement la relation entre A^* et A.
Les propriétés de la proposition 9.27 se généralisent facilement et se démontrent de la même manière.

Proposition 10.14 – *Pour tous endomorphismes* f *et* g *et pour tout* $\lambda \in \mathbb{C}$*, on a :*

a) $f^{**} = f$, $(\mathrm{id})^* = \mathrm{id}$;

b) $(f+g)^* = f^* + g^*$, $(\lambda f)^* = \overline{\lambda}\, f^*$, $(f \circ g)^* = g^* \circ f^*$;

c) $\mathrm{rg}\, f^* = \mathrm{rg}\, f$, $\mathrm{dét}\, f^* = \overline{\mathrm{dét}\, f}$.

Proposition 10.15 – *Soient (E,h) un espace vectoriel de dimension finie sur $\mathbb{C}$ muni d'une forme hermitienne h* ***non dégénérée*** *et f un endomorphisme de E.*
Les propriétés suivantes sont équivalentes :

1. $\tilde{q}\big(f(x)\big) = \tilde{q}(x)\,, \quad \forall x \in E$ *, où* $\tilde{q}(x) := h(x,x)$
2. $h\left(f(x), f(y)\right) = h\,(x,y)\,, \quad \forall x,y \in E$
3. $f^* \circ f = \mathrm{id}$ *(ou, d'une manière équivalente,* $f \circ f^* = \mathrm{id}$*).*
   En particulier f est bijective.

Un endomorphisme qui vérifie ces propriétés est dit ***unitaire*** *relativement à h.*

La démonstration est analogue à celle de la proposition 9.28. On vérifie facilement que :

Proposition 10.16 – *Les applications unitaires forment un groupe pour la loi des composition des applications, dit* ***groupe unitaire associé à*** *h, noté $\mathrm{U}(h)$. Si $f \in \mathrm{U}(h)$ alors $|\,\mathrm{dét}\, f\,| = 1$.*
L'ensemble $\mathrm{SU}(h) := \left\{\, f \in \mathrm{U}(h) \,|\, \mathrm{dét}\, f = 1 \,\right\}$ *est un sous-groupe de $\mathrm{U}(h)$ dit* ***groupe spécial unitaire associé à*** *h.*

Finalement, on trouve facilement la caractérisation des matrices qui représentent les transformations unitaires associées à h. Si $\{e_i\}$ est une base de E, $H = M\,(h)_{e_i}$ et $A = M\,(f)_{e_i}$, alors $f \in \mathrm{U}(h)$ si et seulement si

$$ {}^t\overline{A}\, H\, A \,=\, H $$

En effet : $$\begin{aligned} f \in \mathrm{U}(h) &\iff h\left(f(x), f(y)\right) = h\,(x,y)\,, \ \ \forall x,y \in E \\ &\iff {}^t\overline{(AX)}\, H\, A\, Y = {}^t\,\overline{X}\, H\, Y\,, \ \ \forall X,Y \in \mathcal{M}_{n,1}(K) \\ &\iff {}^t\overline{X}\, {}^t\overline{A}\, H\, A\, Y \,=\, {}^t\overline{X}\, H\, Y\,, \ \ \forall X,Y \in \mathcal{M}_{n,1}(K) \\ &\iff {}^t\overline{A}\, H\, A \,=\, H \end{aligned}$$

Exercice 8.

10.5 Formes hermitiennes dans un espace hermitien

Dans ce paragraphe on suppose, comme au chapitre 9 (cf. page 313), que l'espace vectoriel E est muni de deux structures : une forme hermitienne h et un produit scalaire hermitien $\langle\ |\ \rangle$. Comme dans le cas réel (cf. théorème 9.24 page 313), on montre qu'il existe des bases orthogonales à la fois pour h et pour $\langle\,|\,\rangle$.

Théorème 10.17 – *Soit h une forme hermitienne sur un espace hermitien $(E, \langle\ |\ \rangle)$. Il existe alors des bases orthogonales à la fois pour h et pour $\langle\ |\ \rangle$.*
De telles bases se construisent en prenant une base orthogonale (orthogonale pour $\langle\ |\ \rangle$) dans chaque espace propre de l'endomorphisme f_h défini par :

$$ \langle f_h\,(x) \,|\, y\,\rangle = h\,(x,y\,)\,, \qquad \forall x,y \in E $$

f_h est en effet autoadjoint donc diagonalisable et ses espaces propres sont deux à deux orthogonaux pour $\langle\ |\ \rangle$. [2]

[2]cf. théorème 8.16, page 285.

La démonstration est la même que celle du cas réel (théorème 9.24 page 313).
De la même manière que le corollaire 9.25, on démontre aussi le corollaire :

Corollaire 10.18 – *Soit h une forme hermitienne sur un espace vectoriel complexe de dimension finie et $H = M(h)_{e_i}$ où $\{e_i\}$ est une base quelconque. Alors :*

1. *On peut construire une base orthogonale de h formée par des vecteurs propres de H.*
2. *De plus :* $\operatorname{sign}(h) = (n_+, n_-)$, *où :*
   n_+ *= nombre des valeurs propres strictement positives de H,*
   n_- *= nombre des valeurs propres strictement négatives de H.*

Exercices 9. 10.

EXERCICES

1 Déterminer le rang et le noyau de forme hermitienne h sur $\mathbb{C}^3$ définie par

$$\tilde{q}(x) = h(x,x) = |x_1|^2 + 4\,|x_2|^2 + 2\,|x_3|^2 - 2\,i\,\overline{x}_1\,x_2 + 2\,i\,\overline{x}_2\,x_1 + (1-i)\,\overline{x}_1\,x_3 + (1+i)\,\overline{x}_3\,x_1 + 2\,(1+i)\,\overline{x}_2\,x_3 + 2\,(1-i)\,\overline{x}_3\,x_2$$

2

1. Soit E un espace vectoriel de dimension finie sur $\mathbb{C}$, $\{e_i\}$ une base de E. Montrer qu'il existe une et une seule base $\{\varphi_i\}$ de $E^{\overline{*}}$ telle que : $\varphi_i(e_j) = \delta_{ij}$ (base antiduale de $\{e_i\}$).
2. Déterminer la base de $(\mathbb{C}^3)^{\overline{*}}$ antiduale de la base :
   $$e_1 = (1, i, -2\,i), \quad e_2 = (0, 1+i, -1), \quad e_3 = (i, 0, 0).$$

* **3** Soit E un espace vectoriel de dimension finie sur $\mathbb{C}$. Montrer que $E^{\overline{**}}$ est canoniquement isomorphe à E.

* **4** Soit E un espace vectoriel de dimension finie sur $\mathbb{C}$ muni d'une forme hermitienne h. Si F est un sous-espace vectoriel de E, on pose :

$$F^{\overline{0}} = \{\omega \in E^{\overline{*}} \;:\; \omega(x) = 0\,, \quad \forall x \in F\}$$

1. Montrer que $\dim F^{\overline{0}} = \dim E - \dim F$.
2. Soit
   $$j : \begin{array}{ccc} E & \longrightarrow & E^{\overline{*}} \\ y & \longmapsto & h(\cdot\,, y) \end{array}$$
   En identifiant $E^{\overline{**}}$ avec E (cf. exercice 3.), montrer que $j(F)^{\overline{0}} = F^{\perp}$.
3. En déduire que pour tout sous-espace vectoriel F de E on a :
   $$\dim E = \dim F + \dim F^{\perp} - \dim (F \cap N).$$

5 Déterminer le rang et la signature de la forme hermitienne associée à la matrice hermitienne

$$\begin{pmatrix} 1 & 0 & -2\,i \\ 0 & -2 & 1+i \\ 2\,i & 1-i & 0 \end{pmatrix}$$

6 Soit la forme hermitienne sur $\mathbb{C}^2$ définie par :

$$\tilde{q}\,(x) = |\,x_1 - (2+i)\,x_2\,|^2$$

Déterminer $F^{\perp\perp}$, où $F = \text{Vect}\begin{pmatrix} 1 \\ 1 \end{pmatrix}$.

7 Construire une matrice hermitienne non nulle d'ordre 3 dont toutes les lignes sont proportionnelles.

8 On considère sur $\mathbb{C}^n$ la forme hermitienne définie par la matrice $H = \begin{pmatrix} I_p & 0 \\ 0 & -I_q \end{pmatrix}$ et on note $\mathrm{U}(p,q)$ le groupe unitaire associé à cette forme hermitienne. On note aussi $\mathrm{U}(n) = \mathrm{U}(n,\mathbb{C})$. Montrer que :

$$\mathrm{U}\,(n) \cap \mathrm{U}\,(p,q) = \mathrm{U}\,(p) \times \mathrm{U}\,(q)$$

9 **Réduction simultanée de deux formes hermitiennes**

1. Généraliser le résultat de l'exercice 24. du chapitre 9. au cas des formes hermitiennes :

   Soient h et h' deux formes hermitiennes sur un espace vectoriel complexe E de dimension finie. On suppose que h est non dégénérée. Il existe alors une base qui est orthogonale à la fois pour h et pour h' si et seulement si l'endomorphisme f défini par

   $$h\,(\,f(x),\,y\,) = h'\,(\,x,\,y\,)\,, \qquad \forall x, y \in E$$

   est diagonalisable.

2. Soient H et H' les matrices de deux formes hermitiennes h et h' dans une certaine base, h étant supposée non dégénérée. Montrer que si $H^2 = {H'}^2$, alors il existe une base qui est orthogonale à la fois pour h et pour h'.

10 Sans effectuer des calculs, expliquer pourquoi la matrice

$$\begin{pmatrix} \sqrt{2} & 1+i\sqrt{2} & -i \\ 1-i\sqrt{2} & \sqrt{2} & \sqrt{2} \\ i & \sqrt{2} & -\sqrt{2} \end{pmatrix}$$

admet au moins une valeur propre positive et au moins une valeur propre négative.

INDICATIONS

1 Déterminer le rang de la matrice associée à h à l'aide de la méthode du pivot. On trouve $\operatorname{rang} h = 1$. $N(h)$ est l'hyperplan $x_1 - 2\,i\,x_2 + (1-i\,)\,x_3 = 0$.

2

1. Si $\varphi_i(e_j) = \delta_{ij}$ et les φ_i sont antilinéaires, nécessairement : $\varphi_i(x) = \overline{x}_i$. Vérifier que la famille $\{\varphi_i\}$ est libre et génératrice, en notant que, si $\theta \in E^{\overline{*}}$, alors :

   $$\theta(x) = \sum_i \overline{x}_i\,\theta(e_i) = \sum_i \varphi_i(x)\,\theta(e_i) = \sum_i a_i\,\varphi_i(x) \quad \text{ où } a_i = \theta(e_i).$$

2. Soit $\{\varphi_1, \varphi_2, \varphi_3\}$ la base antiduale. Si $\varphi_1(x) = a\,\overline{x}_1 + b\,\overline{x}_2 + c\,\overline{x}_3$ on doit avoir : $\varphi_1(e_1) = 1$, $\varphi_1(e_2) = 0$, $\varphi_1(e_3) = 0$; ce qui donne le système :

   $$\begin{cases} a \quad\quad + i\,c = 1 \\ (1-i\,)\,b \ - c = 0 \\ -i\,a \quad\quad = 0 \end{cases} \qquad \text{D'où :} \quad a = 0\,, \quad b = \frac{-1+i}{2}\,, \quad c = -i$$

   et donc : $\varphi_1(x) = -\dfrac{1+i}{2}\,\overline{x_2} - i\,\overline{x_3}$.

   De même on trouve : $\varphi_2(x) = \dfrac{1+i}{2}\,\overline{x}_2\,,\ \varphi_3(x) = i\,\overline{x}_1 + \dfrac{1+i}{2}\,\overline{x}_2 - \overline{x}_3$

3 Considérer l'application :

$$\psi\ : \begin{array}{c} E \longrightarrow \overline{E^{*}}^{*} \\ x \longmapsto \psi(x) \end{array} \qquad \text{où :} \qquad \psi(x)\ : \begin{array}{c} E^{*} \longrightarrow \mathbb{C} \\ \omega \longmapsto \overline{\omega(x)} \end{array}$$

Montrer que $\psi(x)$ est bien antilinéaire, et que ψ est linéaire et injective.

4

1. Même démonstration que pour la proposition 3.39 page 87.
2. $j(F)^{\overline{0}} \subset \overline{E^{*}}^{*} \simeq E$ (cf. exercice 3).
   $$\begin{aligned} j(F)^{\overline{0}} &= \{z \in E \mid \varphi(z) = 0\,,\quad \forall \varphi \in j(F)\} = \{z \in E \mid h(z,x) = 0\,,\ \forall x \in F\} \\ &= F^{\perp} \end{aligned}$$
3. Mêmes démonstrations que pour le lemme 9.20 page 311 et la proposition 9.21.

5 $\tilde{q}(x) = \Big| x_1 + 2\,i\,x_3 \Big|^2 - 2\,\Big| x_2 + (1-i\,) \Big|^2$; rang $\tilde{q} = 2$, sign $\tilde{q} = (1,1)$.

6 Puisque rang $\tilde{q} = 1$, on a $\dim N = 1$. N est défini par $x_1 - (2-3\,i)\,x_2 = 0$. On voit immédiatement que $F \cap N = \{0\}$, donc $F^{\perp\perp} = F + N = \mathbb{C}^2$.

7 Il suffit de considérer une forme hermitienne de rang 1 et de prendre sa matrice. Par exemple : $\tilde{q}\,(x) = \Big| (1+3\,i\,)\,x_1 - i\,x_2 + (1+2\,i\,) \Big|^2$. On obtient la matrice

$$\begin{pmatrix} 10 & -3-i & 7-i \\ -3+i & 1 & -2+i \\ 7+i & 2+i & 5 \end{pmatrix}$$

dont les lignes sont bien proportionnelles, malgré les apparences ...

8 On a : $\begin{pmatrix} A & B \\ C & D \end{pmatrix} \in \mathrm{U}(\,n\,) \cap \mathrm{U}(\,p,q\,)$ si et seulement si $B = 0$, $C = 0$ et $A \in \mathrm{U}(p)$, $D \in \mathrm{U}(q)$ (cf. exercice 22. du chapitre 7).

9

1. Suivre l'exercice 24. du chapitre 9.
2. Il existe une base orthogonale commune si et seulement si la matrice $H^{-1}\,H'$ est diagonalisable. Ce sera le cas, en particulier, si $H^{-1}\,H'$ est unitaire (car toute matrice unitaire est normale donc diagonalisable dans $\mathbb{C}$) c'est-à-dire si $H^{-1}\,H'\,(H^{-1}\,H')^{*} = I$, ce qui équivaut à $H^2 = H'^2$, car H et H' sont hermitiennes.

10 Si h est la forme hermitienne sur $\mathbb{C}^3$ représentée par H dans la base canonique $\{e_i\}$, on a : $\tilde{q}\,(e_1) = \sqrt{2} > 0$ et $\tilde{q}\,(e_3) = -\sqrt{2} < 0$; donc la signature ne peut être ni $(r,0)$, ni $(0,r)$.

Appendice A.1

Vocabulaire de base

Applications

Soient E et E' deux ensembles. Une ***application*** f de E dans E' est un procédé qui à tout élément x de E associe un élément (unique) de E', noté $f(x)$. On note cela :

$$f\ :\begin{array}{ccc} E & \longrightarrow & E' \\ x & \longmapsto & f(x) \end{array}$$

Exemples :

$$f\ :\ \varphi\begin{array}{ccc} \mathbb{R} & \longrightarrow & \mathbb{R} \\ x & \longmapsto & \sin x \end{array} \qquad \text{est une application ;}$$

$$f\ :\begin{array}{ccc} \mathbb{R}^+ & \longrightarrow & \mathbb{R} \\ x & \longmapsto & \{\sqrt{x}\,,\,-\sqrt{x}\} \end{array} \qquad \text{n'est pas une application.}$$

Si $f\ :\ E \longrightarrow F$ et $g\ :\ F \longrightarrow G$ sont deux applications, on note $g \circ f$ l'application (dite composée) :

$$g \circ f\ :\begin{array}{ccc} E & \longrightarrow & G \\ x & \longmapsto & g(\,f(x)\,) \end{array}$$

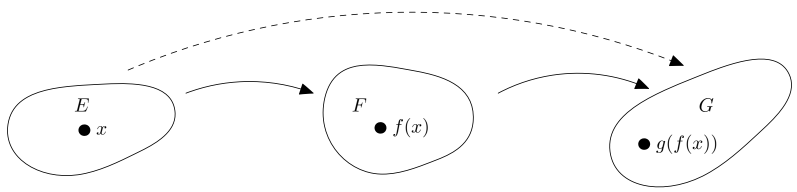


Figure 1

Définition A.1.1 . –
Soit $f\ :\ E \longrightarrow E'$ *une application. Si* $x \in E$, $f(x)$ *est dit* ***image de*** x *par* f *;* x *est dit* ***antécédent*** *de* $f(x)$*.*
Si $A \subset E$, *on appelle* ***image de*** A *par* f *le sous-ensemble de* E' *:*

$$f(A)\ := \Big\{\text{images des éléments de } A\Big\} = \Big\{y \in E' \,|\, \exists x \in A\ :\ y = f(x)\Big\}$$

$f(E)$ *est noté* $\mathrm{Im} f$ *et est dit* ***image de*** f*.*

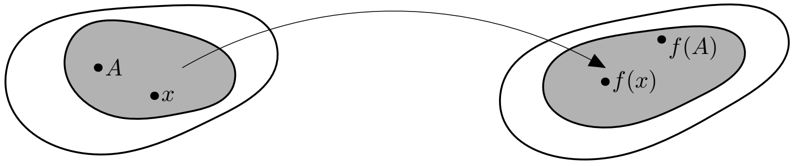


Figure 2

Exemple –

Soit $f\ : \begin{array}{ccc}\mathbb{R} & \longrightarrow & \mathbb{R} \\ x & \longmapsto & \sin x\end{array}$. On a : $f([0,\pi]) = [0,1]$ et $\operatorname{Im} f = [-1,+1]$.

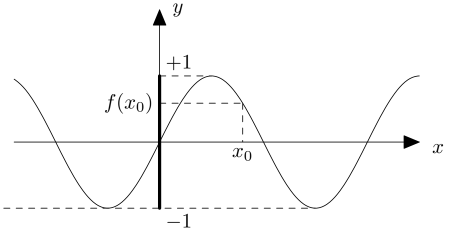


Figure 3

Définition A.1.2 –

- *Une application $f\ :\ E \longrightarrow E'$ est dite **surjective** si $f(E) = E'$, c'est-à-dire si tout élément de E' est l'image par f d'un élément de E : $\forall y \in E' \quad \exists x \in E\ :\ y = f(x)$*
- *L'application $f\ :\ E \longrightarrow E'$ est dite **injective** si : $x \neq x' \implies f(x) \neq f(x')$ c'est-à-dire : si deux éléments sont distincts alors leurs images sont distinctes. En prenant la contraposée, cette propriété est équivalente à :*

$$f(x) = f(x') \implies x = x'.$$

- *$f\ :\ E \longrightarrow E'$ est dite **bijective** si elle est injective et surjective, c'est-à-dire si : $\forall y \in E' \quad$ il existe un et un seul $x \in E \quad$ tel que $y = f(x)$.*

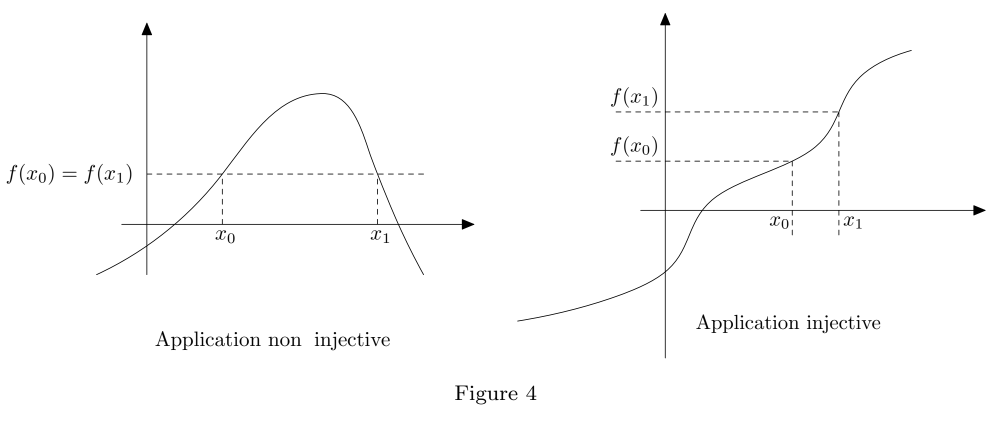


Figure 4

Si $f\ :\ E \longrightarrow E'$ est bijective, on peut définir une application $g\ :\ E' \longrightarrow E$ comme suit : à l'élément $y \in E'$ on fait correspondre son antécédent x dans E qui existe (puisque f est surjective) et est unique (puisque f est injective). g est dite ***application réciproque*** ou ***inverse*** de f et est notée f^{-1}. On a, d'après la définition :

$$f \circ f^{-1} = \mathrm{id}_{E'}\ , \qquad f^{-1} \circ f = \mathrm{id}_E,$$

où :

$$\mathrm{id}_E\ : \underset{x}{E} \underset{\longmapsto}{\longrightarrow} \underset{x}{E}\ .$$

Lois de composition

Une ***loi de composition interne*** sur un ensemble E est une application :

$$*\ :\ E \times E \longrightarrow E$$

L'image du couple (x, y) est notée $x * y$. On emploie aussi, selon les cas, les notations $x \cdot y$ ou $x + y$.
Si E et Ω sont deux ensembles, une ***loi de composition externe*** sur E de domaine d'opérateurs Ω est une application :

$$\Omega \times E \longrightarrow E$$

L'image du couple (λ, x) est notée $\lambda \cdot x$ ou aussi $\lambda\, x$.

Définition A.1.3 –

- *Une loi interne $*$ sur E est dite* ***associative*** *si :*

$$a * (b * c) = (a * b) * c \qquad \forall a, b, c \in E.$$
$$a * b = b * a \qquad \forall a, b \in E.$$

- *S'il existe un élément $e \in E$ tel que :*

$$e * x = x * e = x \qquad \forall x \in E$$

  e est dit ***élément neutre****. (On voit facilement que l'élément neutre, s'il existe, est unique).*

- *Si l'élément neutre existe, un élément $x \in E$ est dit* ***symétrisable*** *s'il existe $x' \in E$ tel que :*

$$x * x' = x' * x = e.$$

  x' est dit ***symétrique*** *de x.*

Exemple – La multiplication dans $\mathbb{R}$ est commutative et associative ; l'élément neutre est 1 ; tout $x \in \mathbb{R}$, $x \neq 0$, est symétrisable et son symétrique est $\frac{1}{x}$.
De même l'addition dans $\mathbb{R}$ est commutative et associative ; l'élément neutre est 0 ; tout $x \in \mathbb{R}$ est symétrisable et le symétrique de x est $-x$.

En général, si la loi est notée multiplicativement, le symétrique de x est noté x^{-1} ; si la loi est notée $+$, le symétrique de x est noté $-x$.

Définition A.1.4 – *Soient $(E, *)$ et $(E', *')$ deux ensembles munis de lois de composition internes. Une application $f\ :\ E \longrightarrow E'$ est dite* ***homomorphisme*** *si :*

$$f(x * y) = f(x) *' f(y).$$

Un homomorphisme bijectif est appelé ***isomorphisme****.*

Groupes, anneaux, corps

Groupes

Définition A.1.5 – *On appelle* **groupe** *un ensemble muni d'une loi interne qui vérifie les propriétés suivantes :*

- *la loi est associative,*
- *il existe un élément neutre,*
- *tout élément est symétrisable.*

Exemples : $(\mathbb{Z},+)$, $(\mathbb{Q},+)$, $(\mathbb{R},+)$, $(\mathbb{C},+)$, $(\mathbb{R}\setminus\{0\},\cdot)$, $(\mathbb{C}\setminus\{0\},\cdot)$.

L'ensemble des applications bijectives de E dans E est un groupe pour la loi de composition des applications.

Un ***sous-groupe*** d'un groupe G est un sous-ensemble non vide H de G auquel la loi induite (c'est-à-dire la restriction de la loi de G à H), confère une structure de groupe.

On voit facilement que :

$H \subset G$ est un sous-groupe de G si et seulement si :

1. $H \neq \emptyset$
2. $\forall\, a, b \in H,\ ab^{-1} \in H$.

Par exemple $(\mathbb{Z},+)$ est sous-groupe de $(\mathbb{Q},+)$. $(\mathbb{Q},+)$ est sous-groupe de $(\mathbb{R},+)$.

Anneaux

Définition A.1.6 – *On appelle* **anneau** *un ensemble A muni de deux lois de composition internes notées $+$ et $\cdot$, telles que :*

- *$(A,+)$ est un groupe commutatif ;*
- *La loi $\cdot$ est associative ;*
- $a\cdot(b+c) = a\cdot b + a\cdot c$
  $(a+b)\cdot c = a\cdot c + b\cdot c \qquad \forall a,b,c \in A.$

L'anneau est dit unitaire si la loi multiplicative admet un élément neutre, que l'on note 1_A ou plus simplement 1.
($a\cdot b$ est noté souvent, plus simplement, $a\,b$).

Exemples : $(\mathbb{Z},+,\cdot)$, $(\mathbb{R},+,\cdot)$, $(\mathbb{C},+,\cdot), \mathcal{M}_n(\mathbb{R})$, $\mathcal{M}_n(\mathbb{C})$;

$\mathbb{R}[x]$, $\mathbb{C}[x]$ (fonctions polynômes à cœfficients respectivement dans $\mathbb{R}$ et dans $\mathbb{C}$) avec les lois habituelles de somme et de produit de fonctions.

$\mathbb{Z}_2 := \{0,1\}$ avec les lois :

$0+0=0$	$0\cdot 0=0$
$0+1=1+0=1$	$0\cdot 1=1\cdot 0=0$
$1+1=0$	$1\cdot 1=1$

Plus généralement si n est un entier strictement positif :

$\mathbb{Z}_n := \{0,1,2,\ldots,n-1\}$ avec les lois : $a+b :=$ reste de la division de $a+b$ par n
$a\cdot b :=$ reste de la division de $a\cdot b$ par n

est un anneau dit *anneau des entiers modulo* n (cf. aussi exemple 2. page 353).

Un anneau est dit ***intègre*** si on a la propriété : $ab=0 \implies a=0$ ou $b=0$.

Par exemple : $\mathbb{Z}$, $\mathbb{R}$, $\mathbb{C}$ sont intègres.
$\mathcal{M}_n(\mathbb{R})$ n'est pas intègre
$\mathbb{Z}_n$ n'est pas intègre si n n'est pas premier (par exemple, dans $\mathbb{Z}_6$ on a $2\cdot 3=0$).

Soit A un anneau unitaire. S'il existe un entier $p > 0$ tel que

$$\underbrace{1 + \cdots + 1}_{\text{p fois}} = 0$$

on dit que l'anneau est de caractéristique non nulle et le plus petit entier p pour lequel cela se réalise est dit ***caractéristique*** de l'anneau . Si un tel entier n'existe pas l'anneau est dit de caractéristique nulle.

Par exemple : $\mathbb{Z}$, $\mathbb{R}$, $\mathbb{C}$ sont de caractéristique 0.
$\mathbb{Z}_n$ est de caractéristique n.

On peut montrer que la caractéristique d'un anneau intègre est soit 0, soit un nombre premier.

Un ***sous-anneau*** d'un anneau A est un sous-ensemble non vide de A auquel les lois induites confèrent une structure d'anneau.

$B \subset A$ est un sous-anneau de A si et seulement si

1. $B \neq \emptyset$
2. $\forall\, a, b \in B$: $a - b \in B$ et $ab \in B$.

Un ***idéal*** d'un anneau *commutatif* A est un sous-ensemble non vide I de A tel que :

1. $\forall\, x, y \in I\ : x - y \in I$
2. $\forall\, a \in A$ et $\forall\, x \in I$: $ax \in I$.

En particulier, tout idéal est un sous-anneau.

Exemples : $n\mathbb{Z}$ est un idéal de $\mathbb{Z}$.
Soit A un anneau commutatif, $a \in A$ et

$$(a) := \Big\{ x \in A \mid x = ab, \text{ avec } b \in A \Big\} \equiv aA$$

(a) est un idéal, dit *idéal engendré par a.*

Définition A.1.7 *Un anneau commutatif A est dit* **principal** *si ses seuls idéaux sont du type (a) avec $a \in A$.*

Nous verrons (cf. Appendice A.2.) que les anneaux des polynômes $\mathbb{R}[x]$ et $\mathbb{C}[x]$ sont principaux.

Corps

Définition A.1.8 – *On appelle* **corps** *un ensemble K muni de deux lois internes, notées $+$ et $\cdot$, telles que :*

- *$(K, +, \cdot)$ est un anneau*
- *$(K\backslash\{0\}, \cdot)$ est un groupe (0 étant le neutre de la loi $+$).*

Le corps est dit commutatif si la loi $\cdot$ est commutative.

Tout corps est un anneau intègre. Donc sa caractéristique est soit 0, soit un nombre premier.

Exemples :

$(\mathbb{Q}, +, \cdot)$, $(\mathbb{R}, +, \cdot)$, $(\mathbb{C}, +, \cdot)$

On montre que $\mathbb{Z}_p$ est un corps si et seulement si p est premier. Par exemple, dans $\mathbb{Z}_7$ le symétrique de 4 relativement à la loi + est 3, et relativement à la loi $\cdot$ est 2.

Modules. Algèbres

Définition A.1.9 – *Soit A un anneau commutatif unitaire. Un* ***A-module*** *est un ensemble M muni d'une loi interne notée $+$ et d'une loi externe de domaine d'opérateurs A, telles que :*

- $(M,+)$ *est un groupe abélien ,*
- $\lambda\,(\mu\,x) = (\lambda\,\mu)\,x$ *,*
- $(\lambda+\mu)\,x = \lambda\,x + \mu\,x$ *,*
- $\lambda\,(x+y) = \lambda\,x + \lambda\,y$ *,*
- $1_A x = x$ *,* $\qquad \forall\,\lambda,\ \mu \in A$ *et* $\forall\,x,\ y \in M$.

En d'autres termes, un module est une sorte d'espace vectoriel dont les scalaires sont les éléments d'un anneau (au lieu d'être les éléments d'un corps).

Exemple : $\mathcal{M}_n(\mathbb{Z}) = \Big\{$ensemble des matrices carrées d'ordre n à cœfficients dans $\mathbb{Z}\Big\}$

Définition A.1.10 – *Soit A un anneau commutatif unitaire. On appelle* ***A-algèbre*** *un module $\mathcal{M}$ sur A, muni d'une loi notée multiplicativement, telle que :*

1. $(\mathcal{M},+,\cdot)$ *est un anneau unitaire,*
2. $\lambda(x\,y) = (\lambda\,x)\,y = x\,(\lambda\,y)\,, \qquad \forall\,\lambda \in A,\ \forall\,x,y \in \mathcal{M}$

Le cas le plus fréquent est celui où $A = K$ est un corps commutatif. Dans ce cas, une algèbre est un espace vectoriel sur lequel est définie une loi multiplicative, vérifiant les propriétés 1. et 2.

Exemples : $\mathbb{C}$ en tant qu'espace vectoriel (sur $\mathbb{R}$ ou sur $\mathbb{C}$) avec la loi de produit habituelle ; $\mathcal{M}_n(\mathbb{Z})$, $\mathcal{M}_n(\mathbb{R})$, $\mathcal{M}_n(\mathbb{C})$, avec les lois habituelles de somme, produit par un scalaire et produit de matrices.
L'algèbre des *quaternions* est ainsi définie. On considère l'espace vectoriel $\mathbb{R}^4$ dont la base canonique est notée $(1,i,j,k)$. On définit une loi multiplicative par la table de multiplication

	1	i	j	k
1	1	i	j	k
i	i	-1	k	-j
j	j	-k	-1	i
k	k	j	-i	-1

que l'on prolonge linéairement à $\mathbb{R}^4$. L'algèbre des quaternions est notée habituellement H (du nom du mathématicien Hamilton qui l'a introduite). On voit facilement que H $(+,\cdot)$ est un *corps* non commutatif (cf. exercice 12 chapitre 8, où l'on a construit un corps de matrices $\tilde{H}$ isomorphe à H).

Groupes de transformations

Définition A.1.11 – *Soient G un groupe et E un ensemble. On dit que G est un* ***groupe de transformations*** *de E (ou aussi que* ***G opère sur E****)*[1] *si l'on se donne une application*

$$\begin{array}{ccc} G\times E & \longrightarrow & E \\ (a,x) & \longmapsto & a\cdot x \end{array}$$

telle que : 1. $\forall\,a,b \in G$: $\quad a\cdot(b\cdot x) = (a\cdot b)\cdot x$;
2. $\forall\,x \in E$: $\quad e\cdot x = x$.

[1]on dit encore que G *agit* sur E.

Soit $a \in G$ et φ_a l'application de E dans E définie par $x \longmapsto a \cdot x$; φ_a *est bijective.*
En effet, pour tous $x \in E$:

$$\varphi_a \circ \varphi_{a^{-1}}(x) = a \cdot (a^{-1}) \cdot x = (a\, a^{-1}) \cdot x = e \cdot x = x$$

donc $\varphi_a \circ \varphi_{a^{-1}} = \mathrm{id}_E$, c'est-à-dire : $\varphi_a^{-1} = \varphi_{a^{-1}}$.

Exemples :

1. $E = \mathbb{R}^2$ et G le groupe orthogonal : $G = \mathrm{O}(2,\mathbb{R})$ (cf. chapitre 7. page 240). G est une groupe de transformations de E :

$$\begin{array}{ccc} G \times \mathbb{R}^2 & \longrightarrow & \mathbb{R}^2 \\ (A\ ,\ x) & \longmapsto & A\,x \end{array}$$

2. $G = \mathrm{GL}\,(n,\mathbb{R})$, $E = \mathcal{M}_n(\mathbb{R})$ (cf. chapitre 3 page 76). On définit

$$\begin{array}{ccc} G \times \mathcal{M}_n(\mathbb{R}) & \longrightarrow & \mathcal{M}_n(\mathbb{R}) \\ (P\ ,\ A\ ) & \longmapsto & P^{-1}\,A\,P \end{array}$$

3. Soit E un espace vectoriel et $G = (E,+)$. On définit

$$\begin{array}{ccc} G \times E & \longrightarrow & E \\ (a,x) & \longmapsto & x + a \end{array}$$

Cette action est dite ***action des translations.***

Définition A.1.12 – *Soit G un groupe de transformations de E et $x \in E$. On appelle* ***orbite*** *de x, l'ensemble des transformés de x par les éléments de G :*

$$\mathrm{orbite}\{x\} = G \cdot x \equiv \{\, a \cdot x \ \mid a \in G\}$$

On dit que G ***opère transitivement*** *si :*

$$\forall\, x, y \in E\,,\ \exists\, a \in G \ \ \textit{tel que}\ \ :\ x = a \cdot y$$

c'est-à-dire, si pour tout couple d'éléments de E il existe une transformation qui fait passer de l'un à l'autre (en d'autres termes : il y a une seule orbite).

Par exemple, l'action de $\mathrm{O}(2,\mathbb{R})$ sur $\mathbb{R}^2$ (cf. exemple 1) n'est pas transitive : deux vecteurs de norme différente ne peuvent pas être transformés l'un dans l'autre par une transformation orthogonale. Les différentes orbites sont les cercles de centre 0 et le point 0.
L'action

$$\begin{array}{ccc} \mathrm{GL}\,(n,\mathbb{R})\ \times\ \mathbb{R}^n \setminus \{0\} & \longrightarrow & \mathbb{R}^n \\ (A\ ,\ x) & \longmapsto & A\,x \end{array}$$

est transitive (soient $x, y \in \mathbb{R} \setminus \{0\}$, compléter x en une base $\mathcal{B} = \{e_1 = x, e_2, ..., e_n\}$, y en une base $\mathcal{B}' = \{{e'}_1 = y, {e'}_2, ..., {e'}_n$ et considérer l'isomorphisme f défini par $f(e_i) = {e'}_i$).
L'action de $\mathrm{GL}\,(n,\mathbb{R})$ sur $\mathcal{M}_n(\mathbb{R})$ (exemple 2.) n'est pas transitive : les orbites sont constituées par les matrices semblables entre elles (il existe donc une bijection canonique entre l'ensemble des orbites et $\mathrm{End}\,(\mathbb{R}^n)$).
L'action des translations (exemple 3.) est transitive.

Définition A.1.13 – *On dit que G opère **effectivement** si e est l'unique élément de G tel que $\varphi_e = \mathrm{id}$, c'est-à-dire si :*

$$a \cdot x = x \ , \ \forall x \in E \implies a = e.$$

*On dit G opère **librement** si*

$$\exists x \in E \ : a \cdot x = x \implies a = e$$

(c'est-à-dire si la seule transformation qui a des "points fixes" est l'identité).

Par exemple, l'action de $\mathrm{O}(2, \mathbb{R})$ sur $\mathbb{R}^2$ (exemple 1.) est effective, mais elle n'est pas libre, car la symétrie par rapport à une droite vectorielle laisse invariants les vecteurs de la droite. L'action de $\mathrm{GL}\,(n, \mathbb{R})$ de l'exemple 2. n'est pas effective (et donc n'est pas libre) car $\varphi_{-I} = \mathrm{id}$.

G opère librement si et seulement si

$$a \cdot x = b \cdot x \implies a = b$$

En effet supposons que G opère librement et soit $a \cdot x = b \cdot x$. En faisant agir b^{-1}, on a $b^{-1}a \cdot x = x$, d'où $b^{-1}a = e$ et donc $a = b$. La réciproque est évidente.

En particulier, G opère *transitivement et librement* si et seulement si

$$\forall\, x, y \in E \, , \ \exists\, !\, a \in G \ \textit{ tel que : } x = a \cdot y.$$

Par exemple, l'action des translations sur un espace vectoriel E est libre et transitive.

Appendice A.2

Polynômes

Fonctions polynômes. Polynômes formels

Soit A un anneau commutatif unitaire (c'est-à-dire pour lequel il existe un élément neutre pour la multiplication, noté 1) : par exemple $A = \mathbb{Z},\ \mathbb{Q},\ \mathbb{R},\ \mathbb{C},\ \mathbb{Z}_n, \ldots$

Définition A.2.1 – *On appelle* ***fonction polynôme*** *sur A une application : $\alpha : A \longrightarrow A$ du type :*

$$\alpha(x) = a_0\, 1 + a_1\, x + a_2\, x^2 + \cdots + a_n\, x^n$$

où $a_0, a_1, \ldots, a_n \in A$ et $n \in \mathbb{N}$. L'ensemble des fonctions polynômes sur A est noté $A[x]$.

$A[x]$ est muni d'une structure d'anneau pour les lois :

$$\begin{aligned}(\alpha+\beta)(x) &:= \alpha(x)+\beta(x)\\ (\alpha\cdot\beta)(x) &:= \alpha(x)\cdot\beta(x)\,.\end{aligned}$$

Un calcul facile montre que si $\alpha(x) = \sum\limits_i a_i x^i$ et $\beta(x) = \sum\limits_j b_j x^j$ (où $x^0 := 1$), alors :

$$(\alpha\cdot\beta)(x) = \sum_n c_n\, x^n\,, \quad \text{avec} \quad c_n = \sum_{k=0}^{n} a_k\, b_{n-k}.$$

Si K est un corps (commutatif), $K[x]$ est muni d'une structure d'espace vectoriel sur K pour les lois (cf. page 5) :

$$\begin{aligned}(\alpha+\beta)(x) &:= \alpha(x)+\beta(x)\\ (\lambda\alpha)(x) &:= \lambda\alpha(x)\,.\end{aligned}$$

Si l'on ne considérait que le cas où $A = \mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}, \ldots$ l'on pourrait se limiter à étudier les fonctions polynômes. Cependant lorsque A ne contient qu'un nombre *fini* d'éléments (par exemple $A = \mathbb{Z}_n$) certaines difficultés apparaissent.
Considérons par exemple sur $\mathbb{Z}_2[x]$ les deux fonctions polynômes :

$$\alpha(x) = x^2 \qquad \text{et} \qquad \beta(x) = x^3$$

On a : $\quad \alpha(0) = 0 \quad \beta(0) = 0$
$\qquad\quad\ \ \alpha(1) = 1 \quad \beta(1) = 1$

Donc les applications α et β sont égales sur tous les éléments de $\mathbb{Z}_2$ et par conséquent, $\alpha = \beta$.
On voit de cet exemple que l'on ne peut pas définir la notion de degré d'une fonction polynôme comme on le fait classiquement dans $\mathbb{R}[x]$: α (qui est égal à β) est-il de degré 2 ou 3 ?
On préfère par conséquent adopter un autre point de vue et étudier, au lieu des fonctions polynômes, ce que l'on appelle les «polynômes formels».

Considérons une suite $(a_0,\, a_1, \ldots, a_n, \ldots)$ d'éléments de A nuls à partir d'un certain rang. On peut lui associer la fonction polynôme :

$$\alpha(x) = a_0 \cdot 1 + a_1 x + \cdots + a_n x^n + \cdots$$

L'exemple précédent montre qu'en général cette correspondance ***n'est pas injective***. En effet :

$$\begin{array}{lcl} (0,0,1,0 \cdots\cdots 0 \cdots) & \longmapsto & \alpha(x) = x^2 \\ (0,0,0,1,0 \cdots\cdot 0 \cdots) & \longmapsto & \beta(x) = x^3 \end{array} ;$$

puisque $\alpha = \beta$ on a deux suites différentes auxquelles correspond la même fonction polynôme. En d'autres termes, étant donnée une fonction polynôme, on ne peut pas parler d'une manière unique de la suite de ses cœfficients. En revanche si l'on parle d'une suite d'éléments de A (nuls à partir d'un certain rang) la fonction polynôme admettant cette suite comme suite de cœfficients est bien définie. Ceci montre qu'il est préférable de s'intéresser aux suites $(a_0,\, a_1, \ldots, a_n, \ldots)$ d'éléments de A (nuls à partir d'un certain rang).

Définition A.2.2 – *Soit A un anneau commutatif unitaire. On appelle* ***polynôme formel*** *sur A (ou simplement polynôme sur A) toute suite $(a_0,\, a_1, \ldots, a_n, \ldots)$ d'éléments de A nuls à partir d'un certain rang.*
On dit que deux polynômes formels $(a_0,\, a_1, \ldots,\, a_n, \ldots)$ et $(b_0,\, b_1, \ldots,\, b_n, \ldots)$ sont égaux si $a_i = b_i\,, \quad \forall i \in \mathbb{N}$.
L'ensemble des polynômes formels sur A est noté $A[X]$.

On a une application :

$$\varphi\ :\ \begin{array}{ccc} A[X] & \longrightarrow & A[x] \\ (a_0, a_1, \ldots, a_n, \ldots) & \longmapsto & a_0\, 1 + a_1\, x + \cdots + a_n\, x^n + \cdots \end{array}$$

On vient de voir que cette application (qui est, bien entendu, surjective) n'est pas, en général, injective. Nous verrons, cependant, que si K est un corps ayant une infinité d'éléments, alors $\varphi\,:\,K[X] \to K[x]$ est une bijection (cf. corollaire A.2.11) .

$A[X]$ est muni d'une structure d'anneau pour les lois :

$$\begin{array}{rl} (a_0,\, a_1, \ldots, a_n,\, \ldots) + (b_0,\, b_1, \ldots, b_n, \ldots) & := (a_0 + b_0,\, a_1 + b_1, \ldots, a_n + b_n, \ldots) \\ (a_0,\, a_1, \ldots, a_n,\, \ldots) \cdot (b_0,\, b_1, \ldots, b_n, \ldots) & := (c_0,\, c_1, \ldots, c_n, \ldots) \end{array}$$

$$\text{où :} \qquad c_n = \sum_{k=0}^{n} a_k\, b_{n-k}.$$

Si K est un corps commutatif, $K[X]$ est un espace vectoriel sur K pour les lois d'addition (comme ci-dessus) et de produit par un scalaire défini par :

$$\lambda \cdot (a_0,\, a_1, \ldots, a_n, \ldots)\ := (\lambda a_0,\, \lambda a_1,\, \ldots, \lambda a_n, \ldots)$$

L'application φ est un homomorphisme pour ces structures. Comme nous allons le voir (cf. ci-dessous, corollaire A.2.11), si K est un corps ayant une infinité d'éléments, $K[X]$ est isomorphe, en tant qu'anneau et en tant qu'espace vectoriel, à $K[x]$.

Ecriture à l'aide de l'indéterminée

On appelle ***indéterminée*** le polynôme formel $X = (0,1,0,\ldots 0,\ldots)$. On voit facilement que :

$$\begin{array}{l} X^2 = (0,0,1,0,\ldots\ldots 0 \ldots\ \ ) \\ X^3 = (0,0,0,1,0,\ldots\ 0,\ldots\ \ ) \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ X^n = (0,0,\ldots.., 0, \underset{\substack{\uparrow \\ (n+1)^{\grave{e}me} rang}}{1}, 0 \ldots\ \ ) \end{array}$$

Par conséquent, en posant :

$$X^0 \;:= 1 \;:= (1, 0, \dots 0, \dots)$$

on peut écrire le polynôme formel $P = (a_0,\, a_1,\, a_2, \dots,\, a_n, \dots)$ sous la forme :

$$P = a_0 1 + a_1 X + a_2 X^2 + \cdots + a_n X^n + \cdots$$

Avec cette notation, l'application φ s'écrit :

$$\begin{array}{rrcl} \varphi : & A[X] & \longrightarrow & A[x] \\ & \sum\limits_i a_i X^i & \longmapsto & \sum\limits_i a_i\, x^i \end{array}$$

On appelle ***valeur en*** $x_0 \in A$, du polynôme formel $P = \sum\limits_i a_i X^i$ la valeur en x_0, notée $P(x_0)$, de la fonction polynôme associée : $P(x_0) \;:= \sum_i a_i x_0^i$.

Soit $P = (a_0,\, a_1, \dots, a_n, \dots) = \sum\limits_i a_i X^i$. Si $P \neq 0$ on appelle ***degré de*** P le plus grand des entiers k tels que $a_k \neq 0$. Le degré de P est noté $\mathrm{d}^0 P$. Si $P = 0$, on pose par convention $\mathrm{d}^0 P = -\infty$.

On voit immédiatement que $\mathrm{d}^0(P+Q) \leq \mathrm{Sup}(\mathrm{d}^0 P, \mathrm{d}^0 Q)$ et que, si A est un corps[1], alors

$$\mathrm{d}^0(PQ) = \mathrm{d}^0 P + \mathrm{d}^0 Q.$$

Résultats principaux de la théorie des polynômes

Dans tout ce paragraphe, K est un corps commutatif. La théorie repose sur deux théorèmes importants, le théorème de la division euclidienne A.2.3 et le théorème qui affirme que $K[X]$ est un anneau principal.

Théorème de la division euclidienne. A.2.3 – *Soient* $A,\, B \in K[X]$, $B \neq 0$. *Il existe alors un et un seul couple* $Q,\, R \in K[X]$ *tels que :*

$$A = B\,Q + R \qquad \textit{et} \qquad d^0 R < d^0 B$$

Q est dit ***quotient*** *de la division euclidienne de A par B et R* ***reste***[2]*.*
Si $R = 0$ on dit que B ***divise*** *A et l'on écrit $B \,|\, A$.*

Démonstration - L'unicité est immédiate. En effet, soit $A = BQ + R$ et $A = BQ' + R'$ avec $d^0 R < d^0 B$ et $d^0 R' < d^0 B$. On a : $B(Q - Q') = R - R'$. Si $Q = Q'$, alors $R = R'$. Par ailleurs $Q = Q'$, car i on avait $Q - Q' \neq 0$, alors $\mathrm{d}^0 B + \mathrm{d}^0(Q - Q') = \mathrm{d}^0(R - R') \leq B$, ce qui est exclu.
Pour l'existence :
- Si $\mathrm{d}^0 A < \mathrm{d}^0 B$ il suffit de prendre $Q = 0$ et $R = A$.
- Si $\mathrm{d}^0 A > \mathrm{d}^0 B$, on raisonne par récurrence. Si $\mathrm{d}^0 A = 0$, le résultat et trivial. Supposons le théorème vrai à l'ordre n et soit $\mathrm{d}^0 A = n + 1$. Soient a_{n+1} et b_p les cœfficients dominants de A et B (c'est -à-dire les cœfficients des termes de plus haut degré) et posons $Q_1 := (a_{n+1}/b_p) X^{n+1-p}$ et $A_1 := A - BQ_1$. Puisque $\mathrm{d}^0 A_1 \leq n$, on peut écrire $A_1 = BQ_2 + R_2$, avec $\mathrm{d}^0 R_2 < \mathrm{d}^0 B$, d'où : $A - BQ_1 = BQ_2 + R_2$ et donc $A = B(Q_1 + Q_2) + R_2$ $\square$.

Théorème A.2.4 *Si K est un corps commutatif, $K[X]$ est anneau principal.*

[1] Cette propriété est vraie plus généralement si A est un anneau intègre c'est-à- dire vérifiant la propriété “ $ab = 0 \Longrightarrow a = 0$ ou $b = 0$ ” (pour $a, b \in A$). C'est le cas lorsque A est un corps.

[2] Compte tenu de la convention : $d^0 0 = -\infty$, la condition $d^0 R < d^0 B$ signifie : « soit $R = 0$, soit, si $R \neq 0$, $d^0 R < d^0 B$ »

Démonstration - Soit $I \neq \{0\}$ un idéal de $K[X]$ et d le plus petit des degrés des polynômes non nuls de I. Considérons un polynôme $P \in I$ de degré d et montrons que $I = (P)$.
On a : $(P) \subset I$, car $P \in I$, donc $QP \in I$, $\forall Q \in K[X]$, puisque I est un idéal. Réciproquement, si $A \in I$, en effectuant la division euclidienne de A par P, on a : $A = PQ + R$ avec $\mathrm{d}^0 R < \mathrm{d}^0 P$. Or $R = A - PQ \in I$, car $A \in I$ et $PQ \in I$; donc nécessairement $R = 0$. Ceci signifie que $A = PQ \in (P)$, donc $A \in (P)$. $\square$

On peut maintenant démontrer facilement les théorèmes suivants :

Théorème A.2.5 ***-(Existence du PGCD)-*** *Soient $\{A_1, \ldots, A_n\}$ n polynômes de $K[X]$; il existe alors un unique polynôme D, normalisé (c'est-à-dire de cœfficient dominant égal à 1), qui vérifie les deux propriétés suivantes :*

a) $D \mid A_i\,, \quad \forall i = 1, \ldots, n.$

b) Si P est un polynôme qui divise tous les A_i, alors $P \mid D$.

D est dit ***plus grand commun diviseur*** *(PGCD) des polynômes $A_1, \ldots, A_n$ et il s'écrit sous la forme*

$$D = U_1 A_1 + \cdots + U_n A_n\,, \quad \text{avec : } U_1, ..., U_n \in K[X]$$

Les polynômes $A_1, \ldots, A_n$ sont dits ***premiers entre eux*** *si le PGCD de $\{A_1 \cdots A_n\}$ est 1.*

Démonstration - Soit $\mathcal{I} = \Big\{ U_1 A_1 + \cdots + U_p A_p \mid U_1, ..., U_p \in K[X] \Big\}$. On voit facilement que $\mathcal{I}$ est un idéal. Puisque $K[X]$ est principal, il existe un polynôme D, que l'on peut supposer normalisé, tel que $\mathcal{I} = (D)$. Ceci signifie que, pour tous $U, ..., U_p$, D divise $A_1 U_1 + \cdots A_p U_p$. En particulier, en prenant $U_i = 1$ et $U_j = 0$ pour $j \neq i$, on voit que D divise tous les A_i. Réciproquement, soit Q un polynôme qui divise tous les A_i, alors $Q \mid A_1 U_1 + \cdots + A_p U_p$ quels que soient $U_1, ..., U_p$, c'est-à-dire Q divise tous les polynômes de $\mathcal{I}$ et, en particulier, $Q \mid D$, car $D \in (D) = \mathcal{I}$. L'unicité du PGCD est laissée en exercice. $\square$

On en déduit comme corollaire (cf. page 161) :

Théorème de Bezout . A.2.6 *Les polynômes $A_1, \ldots, A_n$ sont premiers entre eux si et seulement si il existe des polynômes $U_1, \ldots, U_n$ tels que : $A_1 U_1 + A_2 U_2 + \cdots + A_n U_n = 1$.*

On note $A_1 \wedge \cdots \wedge A_p$ le PGCD de $A_1, ..., A_p$.
On peut vérifier que $A \cdot (P_1 \wedge \cdots \wedge P_n) = (AP_1) \wedge \cdots \wedge (AP_n)$.

Théorème de Gauss. A.2.7 - *Soient A, B_1, B_2 trois polynômes. Si A divise $B_1 B_2$ et A est premier avec B_1, alors A divise B_2.*

En effet $A \wedge B = 1$ d'où : $(AB_1) \wedge (AB_2) = A$. Or $A \mid AB_2$ et $A \mid B_1 B_2$; donc $A \mid B_2$, car B_2 est le PGCD de AB_2 et $B_1 B_2$ $\square$

On peut maintenant démontrer d'une manière plus précise le résultat (i), page 159 :

Proposition A.2.8 *Si $P \in K[X]$ est divisible par des polynômes 2 à 2 premiers entre eux, alors il est divisible par leur produit. En particulier, si $a_1, \ldots, a_p$ sont p racines deux à deux distinctes de P d'ordre respectivement $\alpha_1, \ldots, \alpha_p$, alors P est divisible par $(X - a_1)^{\alpha_1} \cdots (X - a_p)^{\alpha_p}$, c'est-à-dire P s'écrit sous la forme :*

$$P(X) = (X - a_1)^{\alpha_1} \cdots (X - a_p)^{\alpha_p}\, Q(X).$$

Démonstration - Notons d'abord que si A est premier avec P_1 et P_2, alors A est premier avec P_1P_2. En effet de $AU_1 + P_1V_1 = 1$ et $AU_2 + P_2V_2 = 1$, en multipliant membre à membre, on a : $AU + (P_1P_2)V = 1$, avec $U = AU_1U_2 + U_1P_2V_2 + U_2P_1V_1$ et $V = V_1V_2$, ce qui montre que A est premier avec P_1P_2.
Soient maintenant trois polynômes P_1, P_2, P_3 deux à deux premiers entre eux et supposons qu'ils divisent le polynôme A. Puisque $P_1 \mid A$, il existe un polynôme Q_1 tel que $A = P_1Q_1$. D'autre part, $P_2 \mid A$, donc $P_2 \mid P_1Q_1$. Mais P_1 et P_2 sont premiers entre eux, donc, d'après le théorème de Gauss, $P_2 \mid Q_1$. Il existe donc Q_2 tel que $Q_1 = P_2Q_2$ et donc : $A = (P_1P_2)Q_2$. D'autre part, $P_3 \mid A$, donc $P_3 \mid (P_1P_2)Q_2$ et comme P_3 est premier avec P_1P_2, il divise Q_2. Il existe donc Q_3 tel que $Q_2 = P_3Q_3$ et, par conséquent $A = (P_1P_2P_3)Q_3$. La démonstration se généralise facilement au cas de p polynômes deux à deux premiers entre eux. □

Isomorphisme entre K[X] et K[x], si K est infini.

De la proposition A.2.8 il dérive un certain nombre de corollaires.

Corollaire A.2.9 – *Si $P \in K[X]$ admet n racines distinctes, alors :*
- *soit* $P = 0$
- *soit* $d^0 P \geq n$

Corollaire A.2.10 (***critère d'identité des polynômes***) –
Soient A, B de degré $\leq n$ et $a_1, \ldots, a_{n+1} \in K$ deux à deux distincts. Si $A(a_i) = B(a_i)$ pour $i = 1 \ldots n+1$ alors $A = B$.

En effet le polynôme $A - B$ aura $n+1$ racines distinctes ; puisque son degré est n, on a $A - B = 0$.

Corollaire A.2.11 – *Soit K est un corps ayant une infinité d'éléments (par exemple $K = \mathbb{R}$ ou $\mathbb{C}$) ; alors l'application $\varphi : K[X] \to K[x]$ qui à tout polynôme formel associe sa fonction polynôme est un isomorphisme (d'anneau et d'espace vectoriel).*

En effet deux polynômes qui définissent la même fonction polynôme prennent la même valeur sur tous les éléments de K. Leur différence s'annule donc une infinité de fois et, ne pouvant être de degré infini, elle est nulle.

Appendice A.3

Quotients

Ensemble quotient

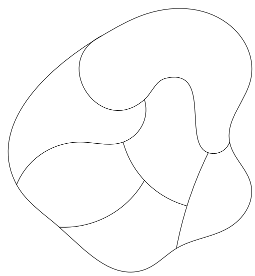


Figure 1

Intuitivement, la notion de quotient est très simple : il s'agit de partager un ensemble E en sous-ensembles disjoints. Ces sous-ensembles sont dits «classes» ; l'ensemble des classes est dit ensemble quotient. Par exemple $\mathbb{Z}$ peut être partagé en deux classes :

$$Z = P \cup I \qquad \text{où :}$$

$$P = \{x \in \mathbb{Z} \,|\, x = 2k \text{ avec } k \in \mathbb{Z}\} \qquad = \{\text{nombres pairs}\}$$

$$I = \{x \in \mathbb{Z} \,|\, x = 2k+1 \text{ avec } k \in \mathbb{Z}\} \quad = \{\text{nombres impairs}\}$$

L'ensemble quotient, noté $\mathbb{Z}_2$ a deux éléments : $\mathbb{Z}_2 = \{P, I\}$
On a vu au chapitre 4. (cf. proposition 4.32 page 132) que l'ensemble des bases d'un espace vectoriel de dimension finie peut se partager en deux classes, les *classes d'orientation.* Ici aussi l'ensemble quotient est formé de deux éléments.

Définition A.3.1 – *Soit E un ensemble ; on appelle **partition** de E une famille de sous-ensembles non vides de E $(X_i)_{i\in I}$ deux à deux disjoints, qui recouvre E, c'est-à-dire telle que :*

1. $\bigcup\limits_{i\in E} X_i = E \qquad$ *(où* $\bigcup\limits_{i\in I} X_i = \{x \in E \mid \exists\, i \in I : x \in X_i\}$
2. $X_i \cap X_j = \emptyset\,, \qquad \forall\, i, j \in I$ *tels que* $i \neq j$.

*On appelle **ensemble quotient** l'ensemble des éléments X_i de la partition.*

La notion de relation d'équivalence, que nous allons définir, permet de construire les partitions et aussi de vérifier facilement si une famille de sous-ensembles de E est une partition.

Définition A.3.2 – *On appelle **relation binaire** sur un ensemble E une assertion notée $\mathcal{R}$ qui porte sur les couples d'éléments de E.*

Par exemple :

$\mathcal{R}_1$: "x est inférieur ou égal à y" est une relation binaire sur $\mathbb{R}$
$\mathcal{R}_2$: "$x - y$ est un nombre pair" est une relation binaire sur $\mathbb{Z}$.
$\mathcal{R}_3$: "a est fils de b" est une relation binaire sur une population.

Si la relation $\mathcal{R}$ appliquée à x et y est vraie, on écrit $x\mathcal{R}y$.
En reprenant les exemples ci-dessus, on peut écrire :

$$\begin{array}{ll} 3 \;\; \mathcal{R}_1 \;\; 4 & \text{car } \; 3 \leq 4 \\ 12 \;\; \mathcal{R}_2 \;\; 16 & \text{car } \; 12-16 \text{ est un nombre pair.} \end{array}$$

Définition A.3.3 – *On appelle* ***relation d'équivalence*** *sur un ensemble E une relation binaire sur E qui est :*

- *réflexive, c'est-à-dire :* $\quad a\mathcal{R}\,a\,, \quad \forall a \in E;$
- *symétrique, c'est-à-dire :* $\quad a\mathcal{R}\,b \quad \Longrightarrow \quad b\mathcal{R}\,a;$
- *transitive, c'est-à-dire :* $\quad (a\mathcal{R}\,b\;,\; b\mathcal{R}\,c) \quad \Longrightarrow \quad a\mathcal{R}\,c.$

Par exemple la relation $\mathcal{R}_1$ ci-dessus est réflexive et transitive mais elle n'est pas symétrique ; donc elle n'est pas une relation d'équivalence. En revanche $\mathcal{R}_2$ est une relation d'équivalence. $\mathcal{R}_3$ n'est ni réflexive, ni symétrique, ni transitive.

NOTATION. - Si $\mathcal{R}$ est une relation d'équivalence, on écrit souvent $x \underset{\mathcal{R}}{\sim} y$, ou plus simplement $x \sim y$, au lieu de $x\,\mathcal{R}\,y$, et l'on dit que x et y sont *équivalents modulo* $\mathcal{R}$.

Définition A.3.4 – *Soit $(E, \mathcal{R})$ un ensemble muni d'une relation d'équivalence. Si $x \in E$, on appelle* ***classe d'équivalence*** *de x par $\mathcal{R}$ (ou simplement : classe de x) le sous-ensemble de E noté $\overline{x}$ de tous les éléments équivalents à x :*

$$\overline{x} = \{\, y \in E \mid y \sim x \,\}$$

Tout élément de la classe est dit ***représentant*** *de la classe. En particulier, x est représentant de $\overline{x}$ (d'après la propriété réflexive).*

Théorème A.3.5 – *Soit $\mathcal{R}$ une relation d'équivalence sur E. Alors les classes d'équivalence des différents éléments de E forment une partition de E.*
Réciproquement, si $(X_i)_{i\in I}$ est une partition de E, il existe sur E une relation d'équivalence dont les classes sont les éléments de la partition.

Démonstration : Montrons d'abord que les classes forment une partition de E.
Il est clair que :

$$\bigcup_{x\in E} \overline{x} = E$$

De plus, soient $x, y \in E$; supposons que $\overline{x} \cap \overline{y} \neq 0$ et soit $z \in \overline{x} \cap \overline{y}$. On a alors $z \sim x$ et $z \sim y$. A l'aide des propriétés de symétrie et transitivité on voit immédiatement que $x \sim y$ c'est-à-dire $\overline{x} = \overline{y}$. Ce qui montre que deux classes sont soit confondues, soit disjointes. L'ensemble des classes d'équivalence forme donc une partition.
Réciproquement, si $(X_i)_{i\in I}$ est une partition de E, on définit sur E la relation binaire $\mathcal{R}$:

$$x\,\mathcal{R}\,y \quad \Longleftrightarrow \quad \exists\, i \in I \;:\; x, y \in X_i;$$

On voit facilement qu'il s'agit d'une relation d'équivalence et que les classes d'équivalence sont les sous-ensembles X_i de la partition. $\square$

Ainsi une partition peut être définie par une relation d'équivalence.

Définition A.3.6 – *Soit E un ensemble et $\mathcal{R}$ une relation d'équivalence sur E. On appelle* ***ensemble quotient*** *de E par $\mathcal{R}$ l'ensemble, noté $E/_{\mathcal{R}}$, des classes d'équivalence pour la relation $\mathcal{R}$.*

Exemple 1. – Soit la relation d'équivalence $\mathcal{R}$ sur $\mathbb{Z}$ définie par :

$x \sim y \quad \Longleftrightarrow \quad x - y$ est un nombre pair.

On a $\quad \mathbb{Z}/_{\mathcal{R}} = \mathbb{Z}_2$ (cf. ci-dessus).

Exemple 2. – Plus généralement, soit $n \in \mathbb{N}$, $n > 0$; on définit sur $\mathbb{Z}$ la relation binaire :

$$\mathcal{R}_n \,:\, x \sim y \quad \Longleftrightarrow \quad \exists\, k \in \mathbb{Z} \,:\, x - y = n\,k$$

Il s'agit d'une relation d'équivalence. L'ensemble quotient $E/_{\mathcal{R}_n}$ est noté $\mathbb{Z}_n$ et est dit ***ensemble des entiers modulo* n** .
$\mathbb{Z}_3$, par exemple, a trois éléments :

$$\begin{array}{ll} \overline{0} & = \{x \in \mathbb{Z} \,:\, x = 3\,k, \qquad \text{avec } k \in \mathbb{Z}\} \\ \overline{1} & = \{x \in \mathbb{Z} \,:\, x = 3\,k + 1, \quad \text{avec } k \in \mathbb{Z}\} \\ \overline{2} & = \{x \in \mathbb{Z} \,:\, x = 3\,k + 2, \quad \text{avec } k \in \mathbb{Z}\} \end{array}$$

Proposition A.3.7 – *Soit E un espace vectoriel et F un sous-espace vectoriel de E. La relation*

$$x\,\mathcal{R}\,y \quad \Longleftrightarrow \quad x - y \in F$$

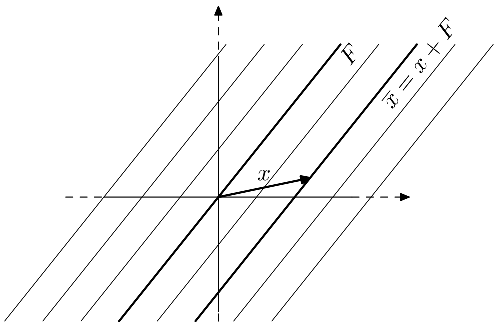
est une relation d'équivalence sur E. L'ensemble quotient par cette relation est noté $E/_F$.

La vérification est immédiate et elle est laissée en exercice.
Si $x \in E$, la classe de x est :

$$\overline{x} \,=\, \{\, y \in E \;|\; y = x + a\,,\ \text{ avec } a \in F\,\} \underset{notat.}{=} x + F$$

Notons que $x + F$ n'est pas un sous-espace vectoriel de E, mais un sous-espace affine (de direction F) : $x + F = \tau_x(F)$ (cf. Appendice A.7).

$E/_F$ est donc l'ensemble de tous les sous-espaces affines de direction F.



L'espace quotient $E/_F$

Figure 2

Lois de composition qui passent au quotient

Définition A.3.8 –

1. *Soit $(E, *, \mathcal{R})$ un ensemble muni d'une loi interne et d'une relation d'équivalence. On dit que la loi $*$ est **compatible** avec la relation d'équivalence $\mathcal{R}$ si :*

$$\left.\begin{array}{l} x \sim x' \\ y \sim y' \end{array}\right\} \quad \Longrightarrow \quad x * y \,\sim\, x' * y'$$

2. *Soit $(E, \Omega, \mathcal{R})$ un ensemble muni d'une loi de composition externe de domaine d'opérateurs Ω et d'une relation d'équivalence. On dit que la loi externe est **compatible** avec la relation d'équivalence si :*

$$x \sim x' \quad \Longrightarrow \quad \lambda\,x \,\sim\, \lambda\,x'\,, \quad \forall \lambda \in \Omega$$

Exemple –

Soit E un espace vectoriel et F un sous-espace vectoriel de E. La relation d'équivalence :

$$x\,\mathcal{R}\,y \quad \Longleftrightarrow \quad x - y \in F$$

est compatible avec les lois de E (vérification immédiate).

Proposition A.3.9 – (*Définition des lois quotient*)

1. *Soit $(E, *, \mathcal{R})$ un ensemble muni d'une loi interne et d'une relation d'équivalence. Si $*$ est compatible avec $\mathcal{R}$, on peut définir une loi interne sur $E/_{\mathcal{R}}$ (dite loi quotient) en posant :*

$$\overline{x} * \overline{y} = \overline{x * y}$$

2. *Soit $(E, \Omega, \mathcal{R})$ un ensemble muni d'une loi externe de domaine d'opérateurs Ω et d'une relation d'équivalence. Si la loi externe est compatible avec $\mathcal{R}$, on peut définir une loi externe sur $E/_{\mathcal{R}}$, de domaine d'opérateurs Ω, en posant :*

$$\lambda \cdot \overline{x} \; := \overline{\lambda\, x}$$

Démonstration : Il s'agit de montrer que les définitions ne dépendent pas du choix des représentants.
Plus précisément, la loi interne sur le quotient est définie de la manière suivante : soient C_1, C_2 deux classes d'équivalence (c'est-à-dire deux éléments de $E/_{\mathcal{R}}$) ; on choisit un représentant dans chaque classe (par exemple $x \in C_1$ et $y \in C_2$; donc $C_1 = \overline{x}$ et $C_2 = \overline{y}$) ; on pose alors :

$$C_1 * C_2 = \overline{x * y}$$

Il s'agit de montrer que si l'on change de représentant, le résultat ne change pas.
Soient $x' \in C_1$ et $y' \in C_2$ deux autres représentants ; on aura $x \sim x'$ et $y \sim y'$; comme la loi est compatible avec la relation d'équivalence, $x * y \sim x' * y'$, c'est-à-dire $\overline{x * y} = \overline{x' * y'}$.
De la même manière on montre que la loi externe sur le quotient est bien définie. □

Espace vectoriel quotient

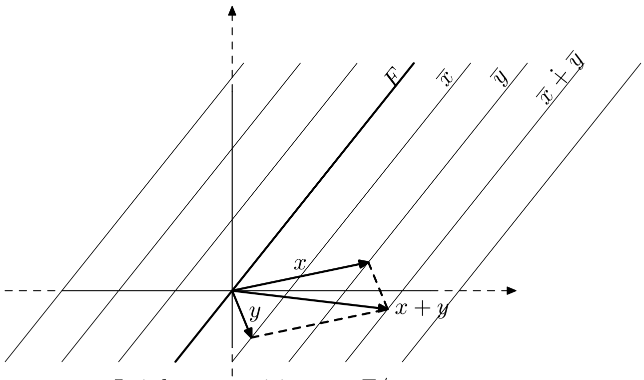
Soit E un espace vectoriel, F un sous-espace vectoriel de E et $E/_F$ l'espace quotient : $E/_F = \{\, x + F \,\}_{x \in E}$. Nous avons vu que la relation d'équivalence définie par F : $x \sim y \Longleftrightarrow x - y \in F$, est compatible avec les lois de E. Par conséquent les lois de E “passent au quotient”. On pose :

$$\begin{array}{rcl} \overline{x} \dot{+} \overline{y} & := & \overline{x + y} \\ \lambda \cdot \overline{x} & := & \overline{\lambda\, x} \end{array}$$

c'est-à-dire :

$$\begin{array}{rl} (x + F) \dot{+} (y + F) & = (x + y) + F \\ \lambda \cdot (x + F) & = \lambda\, x + F \end{array}$$

(notons qu'il s'agit d'égalités entre des sous-ensembles de E).



Loi de composition sur $E/_F$

Figure 3

Proposition A.3.10 – *$E/_F$ muni des lois quotient est un espace vectoriel.*

La démonstration est une simple vérification. Notons que l'élément neutre est $0_{E/F} = \overline{0} = F$ et $-\overline{x} = \overline{-x}$.

Théorème A.3.11 – *Soient E et E' deux espaces vectoriels et $f : E \to E'$ une application linéaire. Alors $E/_{\operatorname{Ker} f}$ est canoniquement isomorphe à $\operatorname{Im} f$. On note cela :*

$$E/_{\operatorname{Ker} f} \simeq \operatorname{Im} f$$

Démonstration : Soit

$$\begin{array}{rcl} i : E/_{\operatorname{Ker} f} & \longrightarrow & \operatorname{Im} f \\ \overline{x} & \longmapsto & f(x) \end{array}.$$

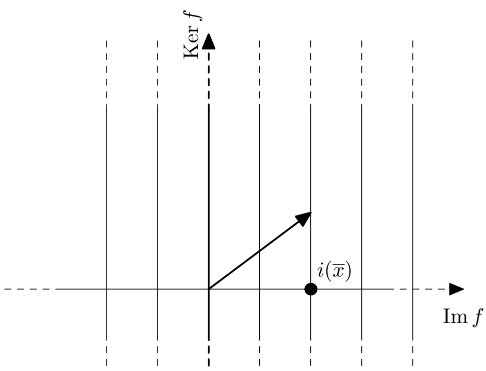
i est bien définie (c'est-à-dire elle ne dépend pas du choix du représentant dans $\overline{x}$). Soit en effet $x' \in \overline{x}$ un autre représentant. On a $x \sim x'$; donc $x - x' \in \operatorname{Ker} f$ et par conséquent $f(x - x') = 0$ c'est-à-dire $f(x) = f(x')$. Ce qui montre que i est bien définie.

D'autre part :

$$\begin{array}{rl} & i(\overline{x} + \overline{y}) = i(\overline{x+y}) = f(x+y) = f(x) + f(y) = i(\overline{x}) + i(\overline{y}) \\ \text{et} & i(\lambda\overline{x}) = i(\overline{\lambda x}) = f(\lambda x) = \lambda f(x) = \lambda i(\overline{x}) \end{array}$$

ce qui montre que i est linéaire.

Enfin, i est bijective. En effet, si $\overline{x} \in \operatorname{Ker} i$, on a : $i(\overline{x}) = 0$ c'est-à-dire $f(x) = 0$, donc $x \in \operatorname{Ker} f$. Mais $\operatorname{Ker} f = \overline{0}$; donc $x \in \overline{0}$, c'est-à-dire $\overline{x} = \overline{0}$. Ceci montre que i est injective. De plus, i est surjective, car si $y = f(x) \in \operatorname{Im} f$, on a : $y = f(x) = i(\overline{x})$. $\square$



$$f : \mathbb{R}^2 \to \mathbb{R}^2 : \quad (x, y) \mapsto (x, 0)$$

$E/_{\mathrm{Ker}\, f} = \{$droites parallèles à 0y$\}$; L'application i fait correspondre à la droite $\overline{x}$ son intersection avec la droite $\mathrm{Im}\, f$ (c'est-à-dire avec l'axe $0x$).

Figure 4

Corollaire A.3.12 – *Si E est de dimension finie,*

$$\boxed{\dim E/_F = \dim E - \dim F}$$

Soit en effet G un supplémentaire de F dans E, $E = F \oplus G$, et considérons la projection de E sur G parallèlement à F :

$$f \ : \ \begin{array}{ccc} E = F \oplus G & \to & G \\ x_1 + x_2 & \longmapsto & x_2 \end{array}$$

On a $F = \mathrm{Ker}\, f$, donc $E/_F = E/_{\mathrm{Ker}\, f} \simeq \mathrm{Im}\, f$. Donc :

$$\dim E/_F = \dim(\mathrm{Im}\, f) = \dim E - \dim(\mathrm{Ker}\, f) = \dim E - \dim F \quad \square$$

Corollaire A.3.13 – *Soient E_1 et E_2 deux sous-espaces vectoriels de E. Alors $E_1 \oplus E_2 \big/_{E_2}$ est canoniquement isomorphe à E_1 :*

$$E_1 \oplus E_2 \Big/_{E_2} \simeq E_1$$

Il suffit de considérer la projection sur E_1 parallèlement à E_2 :

$$f \ : \begin{array}{ccc} E_1 \oplus E_2 & \longrightarrow & E_1 \\ x_1 + x_2 & \longmapsto & x_1 \end{array}$$

On a $\mathrm{Ker}\, f = E_2$ et $\mathrm{Im}\, f = E_1$. On applique ensuite le théorème. $\square$

Exemple : ***cylindre et tore***.

On peut construire un ***cylindre*** en “enroulant” une feuille de papier (cf. figure 5.)

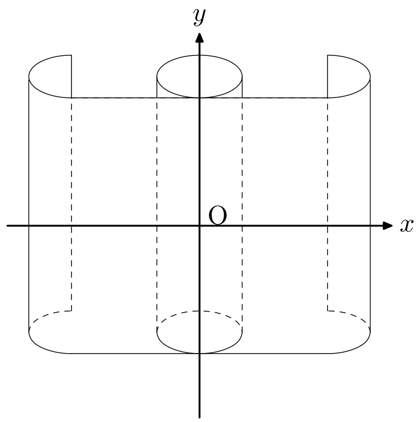


Figure 5

D'une manière plus précise, cela revient à identifier dans le plan $\mathbb{R}^2$ les points (x, y) avec les points $(x+2\,k\,\pi\,R)_{k\in\mathbb{Z}}$. En d'autres termes, le cylindre peut être défini comme l'espace quotient :

$$C = \mathbb{R}^2/\mathcal{R} \qquad \text{où} \quad (x_1, y_1) \underset{\mathcal{R}}{\sim} (x_2, y_2) \iff \begin{cases} \exists\, k \in \mathbb{Z} \ \text{ tel que :} \\ x_2 = x_1 + 2\,k\,\pi\,R \\ y_2 = y_1 \end{cases} .$$

De même, le ***tore*** est la surface que l'on obtient en “recollant” les bords d'un cylindre (cf. figure 6.)

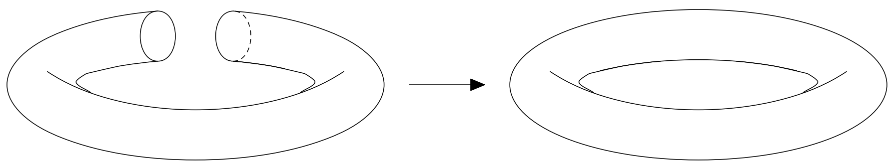


Figure 6

Plus précisément le tore peut être défini comme l'espace quotient :

$$T = \mathbb{R}^2/\mathcal{R} \qquad \text{où} \quad (x_1, y_1) \underset{\mathcal{R}}{\sim} (x_2, y_2) \iff \begin{cases} \exists\, k, h \in \mathbb{Z} \ \text{ tels que :} \\ x_2 = x_1 + 2\,k\,\pi\,R_1 \\ y_2 = y_1 + 2\,h\,\pi\,R_2 \end{cases} .$$

Il est facile de vérifier que la structure naturelle d'espace vectoriel sur $\mathbb{R}^2$ est compatible avec ces relations d'équivalence, ce qui permet de définir sur le cylindre et sur le tore une structure d'espace vectoriel par passage au quotient.

Appendice A.4

Compléments sur la méthode du pivot. Indications sur les méthodes directes

Compléments sur la méthode du pivot

Considérons un système d'équations linéaires, que, pour simplifier, nous supposerons de Cramer :

$$\begin{array}{c} L_1 \\ \dots \\ L_n \end{array} \left\{ \begin{array}{c} a_{11}\,x_1 + \dots\dots + x_{1n}\,x_n = b_1 \\ \dots\dots\dots\dots\dots\dots\dots \\ a_{n1}\,x_1 + \dots\dots + a_{nn}\,x_n = b_n \end{array} \right. \tag{1}$$

Le système (1) peut être écrit sous la forme $A\,X = B$, où $A = (\,a_{ij}\,) = \|\,c_1, \dots, c_n\,\|$, $X = {}^t(x_1, \dots, x_n)$, $B = {}^t(b_1, \dots, b_n)$.
On appelle *méthode directe* de résolution du système (1), une méthode qui aboutit à la solution du problème au bout d'un nombre fini d'opérations arithmétiques (additions, multiplications, divisions).
Dans la pratique, par exemple pour des problèmes d'économie, pour les approximations des équations différentielles, etc, on a à faire à des systèmes d'ordre très élevé, de 50, 100, voire des milliers d'équations. Comme on peut l'imaginer, il est très important de pouvoir disposer d'une méthode de calcul des solutions qui :

1. d'une part ne comporte qu'un nombre raisonnable d'opérations arithmétiques de manière à réduire le temps de calcul des ordinateurs ;
2. d'autre part qu'elle soit «stable» au sens suivant. Les ordinateurs n'ont pas une précision infinie : chaque opération entraîne nécessairement des erreurs d'arrondis. Il est important que les calculs soient menés de manière telle que, en s'accumulant, les erreurs n'entraînent pas des écarts importants entre la solution calculée et la solution théorique.

Estimation du nombre d'opérations

Estimons, par exemple, le nombre d'opérations nécessaires pour résoudre le système (1) par les formules de Cramer

$$x_i = \frac{\Delta_i}{\Delta}, \quad \text{où :}\ \ \Delta = \text{dét}\,A \quad \text{et} \quad \Delta_i = \text{dét}\,\|\,c_1, ..., B, ..., c_n\,\|.$$

Il faut calculer $n+1$ déterminants, puis effectuer n divisions. Chaque déterminant est la somme de $n\,!$ monômes , chacun desquels est le produit de n termes (cf. formule 4.10, page

113). Puisque pour chaque monôme il faut effectuer $n-1$ multiplications, pour calculer un déterminant d'ordre n, il faut :

– $(n-1)\,n!$ multiplications
– $n!-1$ additions

donc au total $(n-1)\,n!+n!-1 = n\,n!-1$ opérations. Pour calculer $n+1$ déterminants il faudra donc effectuer $(n+1)(n\,n!-1)$ opérations. Il faut effectuer ensuite n divisions. Aussi le nombre d'opérations nécessaires pour calculer la solution du système (1) par les formules de Cramer est

$$T_C = (n+1)(n\,n!-1)+n = n\,(n+1)!-n-1+n = n\,(n+1)!-1$$

Rien que pour $n=10$ on trouve $T_C \simeq 400.000.000$ (399.167.999 exactement). On comprend que lorsque n est très grand, il est impossible de résoudre le système par les formules de Cramer.

Évaluons, par contre, le nombre d'opérations nécessaires pour résoudre le système (1) par la méthode d'élimination de Gauss que nous avons vue au chapitre 2.
Il faut d'abord mettre le système sous forme échelonnée, puis le résoudre, par la méthode de la remontée. Soint $L_1, ..., L_n$ les n équations du système et supposons que a_{11} est le pivot. Il s'agit tout d'abord de remplacer chaque équation L_i $(i=2,...,n)$, par $L_i^{(1)} = L_i - \dfrac{a_{i1}}{a_{11}}\,L_1$ de manière à mettre le système sous la forme :

$$\begin{array}{l} L_1 \\ L_2^{(1)} \\ \dots \\ L_n^{(1)} \end{array} \left\{ \begin{array}{llllll} a_{11}\,x_1 + & a_{12}\,x_2 + \dots & \dots & + \,a_{1n}\,x_n & = b_1 \\ & a_{22}^{(1)}\;x_2 + & \dots & + \,a_{2n}^{(1)}\,x_n & = b_2^{(1)} \\ & \dots & \dots\dots & \dots\dots\dots\dots & \\ & a_{n2}^{(1)}\;x_2 + & \dots & + \,a_{nn}^{(1)}\,x_n & = b_n^{(1)} \end{array} \right. \tag{2}$$

Pour calculer les cœfficients de chaque équation $L_i^{(1)}$ $, i=2,...,n,$ il faut 1 division, n multiplications et n additions, soit $2\,n+1$ opérations. Puisqu'il y a $(n-1)$ équations, il faut au total $(2n+1)(n-1)$ opération pour passer du système (1) au système (2). Aussi pour mettre donc la matrice du système sous forme échelonnée il faut

$$\sum_{k=1}^{n}(2\,k+1)\,(k-1) = \frac{4\,n^3+3\,n^2-7n}{6} \quad \text{opérations.}$$

Enfin la solution du système échelonné demande :

1 division pour x_n,
1 division une multiplication et une addition pour x_{n-1},.......
1 division , $n-i$ multiplications et $n-i$ additions, soit $2\,(n-i)+1$ opérations pour x_i

donc au total :

$$1+3+5+\cdots+(2\,n-1) = n^2$$

opérations. Le nombre d'opérations nécessaires pour résoudre le système (1) par la méthode de Gauss est donc

$$T_G = \frac{4\,n^3+3\,n^2-7n}{6} + n^2 = \frac{4\,n^3+9\,n^2-7\,n}{6}$$

Pour $n=10$ on trouve $T_G = 790$. La comparaison de deux méthodes se passe de commentaire.

Erreurs d'arrondi. Choix du pivot

Rien ne s'oppose, a priori, à ce que n'importe quel cœfficient non nul soit choisi comme pivot dans la méthode de Gauss. Cependant, à cause des erreurs d'arrondi, cette manière de procéder peut amener à des résultats numériques très éloignés du résultat théorique. Voici un exemple instructif tiré de [1] page 35. Considérons le système

$$\begin{matrix} L_1 \\ L_2 \end{matrix} \begin{cases} a\,x_1 + x_2 &= 1 \\ \quad x_1 + x_2 &= 2 \end{cases} \qquad \text{avec } a = 10^{-4}$$

dont la solution exacte est :

$$x = \frac{1}{1-a}\,, \qquad y = \frac{1-2\,a}{1-a}$$

c'est-à-dire

$$x = 1 + a + o(a^2) = 1,00010.... \approx 1\,, \qquad y = 1 - a + o(a^2) = 0,99990.... \approx 1$$

Supposons maintenant que l'on résout le système par la méthode de Gauss, en arrondissant les calculs intermédiaires *aux trois premiers chiffres significatifs.*
Si l'on choisit a comme pivot, on a :

$$\begin{matrix} {\scriptstyle L_1} \\ {\scriptstyle L_2 - a\,L_1} \end{matrix} \begin{cases} a\,x + & y &= \quad 1 \\ & (1-10^4)\,y &= 2 - 10^4 \end{cases}$$

Puisque $1-10^4$ et $2-10^4$ sont arrondis tous deux au même nombre, la seconde équation donne $y=1$, d'où la solution

$$x = 0\,, \quad y = 1.$$

très éloignée de la vraie solution. Si en revanche on échange l'ordre des équations, en prenant comme pivot 1, on trouve :

$$\begin{matrix} {\scriptstyle L_2} \\ {\scriptstyle L_1} \end{matrix} \begin{cases} x + y &= 2 \\ a\,x + y &= 1 \end{cases} \;\longrightarrow\; \begin{matrix} {\scriptstyle L_2} \\ {\scriptstyle L_1 - a\,L_2} \end{matrix} \begin{cases} x + y &= 2 \\ (1-a)\,y &= 1 - 2\,a \end{cases}$$

Dans la seconde équation $1-10^{-4}$ et $1-2\cdot 10^{-4}$ sont tous deux arrondis au même nombre, ce qui donne $y=1$, d'où la solution très satisfaisante :

$$x = 1\,, \quad y = 1$$

Ceci montre que des erreurs d'approximation peuvent provenir de la division par des pivots "trop petits". Il est donc indispensable, dans les calculs numériques[2], par exemple par ordinateur, d'avoir recours à l'une des stratégies suivantes :

– *Méthode du pivot partiel :* on prend comme pivot dans le système (2) le plus grand des $|\,a^{(1)}_{22}\,|, \dots, |\,a^{(1)}_{n2}\,|$, en permutant éventuellement les $n-1$ dernières équations.

– *Méthode du pivot total :* on prend comme pivot le plus grand des $|\,a^1_{ij}\,|$, en permutant éventuellement les équations et aussi les "colonnes".

Évidemment ces stratégies doivent être appliquées à chaque étape de l'élimination.
On montre que aussi bien la méthode du pivot partiel que la méthode du pivot total sont *stables*, c'est-à-dire insensibles à la propagation des erreurs.

[1] G.E. Forsythe- C.B Moler, *Computer of Linear Algebraic Systems,* Prentice-Hall, 1967.
[2] Ce problème ne se pose pas, en revanche, dans les questions de calcul formel.

La factorisation LU d'une matrice

Il est clair que dans les programmes de calcul fondés sur la méthode de Gauss, il n'y a pas besoin d'écrire le système sous la forme (1) : ce qui compte ce sont uniquement les matrices A et B. La méthode de Gauss se ramène en fait à travailler non pas sur des équations, mais à faire des opérations sur des matrices. Par exemple, si $A^{(1)}$ est la matrice du système (2), on a $A^{(1)} = E_1\,A$ où E_1 est la matrice *triangulaire inférieure* :

$$E_1 = \begin{pmatrix} 1 & 0 & & 0 \\ -\dfrac{a_{21}}{a_{11}} & 1 & & \vdots \\ \vdots & & \ddots & 0 \\ -\dfrac{a_{n1}}{a_{11}} & 0 & \cdots\ 0 & 1 \end{pmatrix}$$

Si l'on continue à échelonner la matrice, en supposant que $a_{22}^{(1)} \neq 0$ et en le prenant comme pivot [3], on obtient un système dont la matrice $A^{(2)}$ est égale à $E_2\,A^{(1)}$, où :

$$E_2 = \begin{pmatrix} 1 & 0 \\ 0 & \boxed{\begin{matrix} 1 & 0 & & 0 \\ -\dfrac{a_{31}^{(1)}}{a_{22}^{(1)}} & 1 & & \vdots \\ \vdots & & \ddots & 0 \\ -\dfrac{a_{n1}^{(1)}}{a_{22}^{(1)}} & 0 & \cdots\ 0 & 1 \end{matrix}} \end{pmatrix}$$

On a donc : $A^{(2)} = E_2\,E_1\,A$. On aboutit ainsi, toujours à condition que dans la matrice où il faut effectuer l'élimination le premier terme est non nul, à une matrice *triangulaire supérieure* $U = (E_{n-1} \cdots E_2\,E_1)\,A$, où

$$E_k = \begin{pmatrix} 1 & & & & & 0 \\ & \ddots & & & & \\ & & 1 & & & \\ & & -\ell_{k+1,k} & 1 & & \\ & & \vdots & & \ddots & \\ & & -\ell_{nk} & & & 1 \end{pmatrix} \quad \text{avec} \quad \ell_{ik} := \frac{a_{ik}^{(k-1)}}{a_{kk}^{(k-1)}}$$

Si on pose $L = (E_{n-1} \cdots E_2\,E_1)^{-1}$, on pourra donc écrire :

$$A = L\,U$$

où L est une matrice triangulaire inférieure et U une matrice triangulaire supérieure[4].

On peut donner des conditions *suffisantes* pour que dans la matrice où il faut effectuer l'élimination le premier terme soit non nul de manière qu'on puisse le prendre comme pivot, c'est-à-dire pour que l'on puisse aboutir une à factorisation LU. On montre en effet :

[3]Nous ne nous soucions pas ici des questions de stratégie de pivot qui exigent des transpositions de lignes et de colonnes.

[4]Les notations "L" et "U" sont d'origine anglo-saxonne : L pour *lower* et U pour *upper*.

Théorème A.4.1 (***factorisation*** **LU** ***d'une matrice***) - *Soit* $A = (\, a_{ij}\,) \in \mathrm{GL}(n, K)$ *telle que :*

$$\begin{vmatrix} a_{11} & \dots & a_{1k} \\ \vdots & & \vdots \\ a_{k1} & \dots & a_{kk} \end{vmatrix} \neq 0\,, \qquad 1 \leq k \leq n \tag{$*$}$$

Alors A *se factorise* d'une manière unique *sous la forme* $A = L\,U$ *avec* L *triangulaire inférieure et* U *triangulaire supérieure.*
Si la condition (*) n'est pas satisfaite, on peut s'y ramener par une permutation préalable des lignes et des colonnes.

L'intérêt de la factorisation LU apparaît quand il s'agit de résoudre plusieurs systèmes ayant la *même* matrice A, par exemple dans des problèmes d'itération. Il suffit en effet de calculer une fois pour toutes les matrices L et U et ensuite résoudre chaque système $A\,v = c$, c'est-à-dire $L\,U\,v = c$, en résolvant deux systèmes *à matrices triangulaires* :

$$U\,v = w\,, \quad \text{puis} \quad L\,w = c$$

par la méthode de la remontée. Notons enfin que la matrice L s'exprime très simplement en fonction des cœfficients ℓ_{ij}. On peut vérifier en effet que :

$$L = \begin{pmatrix} 1 & & & & 0 \\ \ell_{21} & 1 & & & \\ \ell_{31} & \ell_{32} & 1 & & \\ \vdots & \vdots & \vdots & \ddots & \\ \ell_{n1} & \ell_{n2} & \ell_{n3} & \cdots & 1 \end{pmatrix}$$

alors que il n'y a pas d'écriture simple de la matrice $E_{n-1} \cdots E_2\, E_1$ en fonction des ℓ_{ij} (mais il n'y a pas besoin de calculer cette matrice et seule son inverse est nécessaire).

Quelques indications sur d'autres méthodes directes

Méthode de Cholesky

La méthode de Choleski s'applique au cas ou la matrice A est symétrique définie positive, c'est-à-dire admet n valeurs propres strictement positives (cf. exercice 37 chapitre 7). On peut montrer :

Théorème A.4.2 (***factorisation de Choleski***) – *Toute matrice* $A \in \mathrm{GL}(n, \mathbb{R})$ *définie positive se factorise sous la forme*

$$A = B\,{}^{t}B \qquad \text{avec } B \text{ triangulaire inférieure.}$$

De plus la décomposition est unique si l'on impose que les cœfficients de la diagonale de B *sont positifs.*

La démonstration n'est pas difficile : il suffit de montrer que le système $B\,{}^{t}B = A$ dont les inconnues sont les cœfficients b_{ij} de B peut être résolu explicitement en fonction des cœfficients de A.

L'avantage de la méthode est que la factorisation du type LU ne fait intervenir qu'une matrice. Dans la pratique, on calcule la matrice $B = (\, b_{ij}\,)$ en résolvant le système $A = B\,{}^{t}B$ (cf. exemple ci-dessous).

Exemple – On cherche la décomposition de Choleski de la matrice

$$A = \begin{pmatrix} 1 & 2 \\ 2 & 5 \end{pmatrix}$$

(on vérifie qu'effectivement A est définie positive). Soit $B = \begin{pmatrix} a & 0 \\ b & c \end{pmatrix}$. L'équation matricille $A = B\,{}^tB$ conduit au système

$$\begin{cases} a^2 &= 1 \\ a\,b &= 2 \\ b^2 + c^2 &= 5 \end{cases}$$

et à la solution $B = \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix}$.

Méthode de Householder

Une autre méthode directe est fondée sur la propriété suivante (cf. exercice 25 chapitre 7 pour le cas réel) :

Théorème A.4.3 (*factorisation de Householder*) – *Toute matrice* $A \in \mathrm{GL}(n\mathbb{C})$ *se factorise sous la forme* $A = Q\,R$, *avec* $Q \in \mathrm{U}(n, \mathbb{C})$ *et* R *triangulaire supérieure.*

Si l'on a trouvé une telle factorisation, le système $A\,X = B$, c'est-à-dire $Q\,R\,X = B$, est équivalent à

$$R\,X = {}^t\,\overline{Q}\,B$$

c'est-à-dire à un seul système à matrice triangulaire (alors que la factorisation LU exige la résolution de deux systèmes).

Dans l'exercice 25 du chapitre 7 nous avons vu comme la factorisation de Householder n'est pas autre chose, en fait, que la méthode d'orthonormalisation de Schmidt. Cependant ce ne pas ainsi que dans la pratique on calcule les matrice Q et R, car cette méthode conduit à des erreurs d'arrondis importants. On doit à Householder une méthode de calcul très ingénieuse et stable, fondée sur la factorisation de Q en produit de réflexions (cf. exercice 29, chapitre 7, pour le cas réel), pour laquelle nous renvoyons aux livres spécialisés.

Le lecteur intéressé aux méthodes numériques les plus couramment utilisées en Analyse Matricielle et en Optimisation, pourra consulter

P.G. Ciarlet *Introduction à l'Analyse Numérique Matricielle et l'Optimisation* Masson, 1982

Il s'agit d'un excellent livre, destiné aux étudiants en Master, qui allie la rigueur et la clarté.

Appendice A.5

Inverses généralisées

Le concept d'inverse généralisée

Une matrice a une inverse si et seulement si elle est carrée et son déterminant est non nul. Depuis quelques dizaines d'années on s'est attaché à définir une notion d'inverse pour les matrices à déterminant nul, ou même rectangulaires, qui, bien entendu, se réduit à l'inverse habituelle dans le cas des matrices inversibles. Ces notions se sont révélées très utiles en des diverses branches des mathématiques et des applications.
Pour illustrer l'intérêt de cette théorie, considérons la définition suivante :

Soit $A \in \mathcal{M}_{p,n}(K)$. On appelle inverse généralisée de A une matrice $\tilde{A} \in \mathcal{M}_{n,p}(K)$ telle que :

$$A\,\tilde{A}\,A = A$$

Notons que si A est inversible, alors nécessairement $\tilde{A} = A^{-1}$.
Nous verrons au paragraphe suivant que pour toute matrice A il existe une infinité d'inverses généralisées (sauf dans le cas où A est inversible, où l'inverse généralisée est unique) et qu'il existe des méthodes pour en construire explicitement.
Supposons avoir construit une inverse généralisée $\tilde{A}$ de A et soit :

$$A\,x = b$$

un système linéaire de p équations à n inconnues, avec $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ et $b = \begin{pmatrix} b_1 \\ \vdots \\ b_p \end{pmatrix}$.
Si A est inversible, la solution est unique et elle est donné par

$$x = \tilde{A}\,b\,, \qquad \Big(\text{c'est-à-dire on a : } A\,\tilde{A}\,b = b\Big).$$

Comme nous allons le voir (cf. théorème A.5.2), dans le cas général – même si A n'est pas inversible, ni même carrée – le système est compatible si et seulement si

$$A\,\tilde{A}\,b = b$$

et dans ce cas les solutions sont de la forme :

$$x = \tilde{A}\,b + (\tilde{A}\,A - I_n)y$$

où $y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$ est un vecteur arbitraire.
On voit de cet exemple l'intérêt que peut avoir la construction d'une inverse généralisée. Plus importante encore est l'application à la méthode des moindres carrés, que nous verrons ci-dessous.

Le concept d'inverse généralisée a été introduit pour la première fois par Fredholm dans l'étude des opérateurs intégraux. Pour les matrices les premiers résultats importants ont été obtenus par E. H. Moore en 1920 qui a pu définir une notion d'inverse généralisée *unique* pour toute matrice à coefficients complexes. Ces résultats, qui n'avaient pas été publiés et n'avaient fait l'objet que d'une conférence, ont été retrouvés plus tard, sous des formes différentes par divers mathématiciens. En particulier, Penrose en 1955 a pu caractériser l'inverse généralisée de Moore par un système d'axiomes, ce qui en a facilité l'application à divers domaines des mathématiques.

Inverses généralisées

La proposition suivante donne la méthode pour la construction des inverses généralisées.

Proposition A.5.1 – *Soient E et E' deux espaces vectoriels de dimension finie sur K et $f : E \longrightarrow E'$ une application linéaire. Il existe alors des applications linéaires $\tilde{f} : E' \longrightarrow E$ telles que :*

$$\boxed{f \circ \tilde{f} \circ f = f}$$

*Les applications $\tilde{f}$ sont dites **inverses généralisées** de f.*

Démonstration : Soient F un supplémentaire de $\operatorname{Ker} f$ dans E et G un supplémentaire de $\operatorname{Im} f$ dans E' :

$$E = \operatorname{Ker} f \oplus F \quad , \quad E' = \operatorname{Im} f \oplus G$$

Notons p_F et $p_{\operatorname{Im} f}$ les projecteurs sur F et sur $\operatorname{Im} f$ associés à ces sommes directes.
Soit $y \in E'$, $y = y_1 + y_2$, avec $y_1 \in \operatorname{Im} f$ et $y_2 \in G$. Puisque $y_1 \in \operatorname{Im} f$, il existe $x \in E$ tel que $y_1 = f(x)$. On pose :

$$\tilde{f}(y) := p_F(x)$$

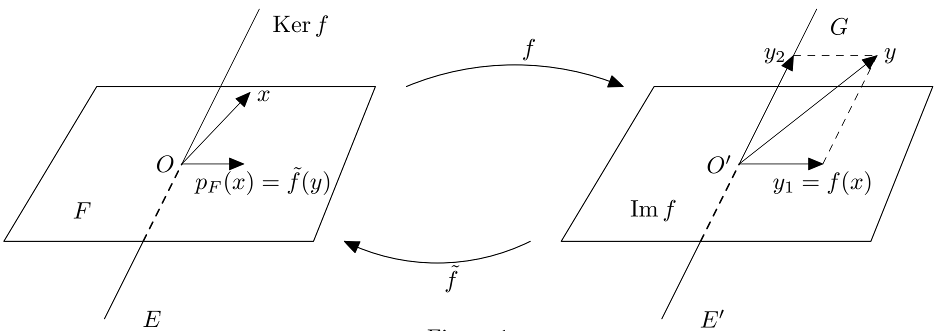


Figure 1

Montrons tout d'abord que cette définition ne dépend pas du choix de x. Soit $x' \in E$ tel que $y = f(x) = f(x')$; on a $f(x - x') = 0$, donc $x - x' \in \operatorname{Ker} f$ et par conséquent $p_F(x - x') = 0$, c'est-à-dire : $p_F(x) = p_F(x')$. Ainsi $\tilde{f}(y)$ ne dépend pas du choix de x.
On vérifie facilement que $\tilde{f}$ est linéaire. D'autre part si $y = f(x)$, alors $\tilde{f}\Big(f(x)\Big) = p_F(x)$ et donc $f \circ \tilde{f} \circ f(x) = f\Big(p_F(x)\Big)$. Or $x = p_F(x) + x_0$, avec $x_0 \in \operatorname{Ker} f$; donc $f\Big(p_F(x)\Big) = f(x)$, ce qui montre que $f \circ \tilde{f} \circ f = f$. $\square$

Remarque. – Les inverses généralisées vérifient aussi la propriété :

$$\boxed{\tilde{f} \circ f \circ \tilde{f} = \tilde{f}}$$

En effet, si $y \in E'$, $y = f(x) + y_2$ (avec $y_2 \in F$) on a, par définition, $\tilde{f}(y) = p_F(x)$. Soit $y' = f \circ \tilde{f}(y)$: on aura :

$$\tilde{f}(y') = p_F\Big(\tilde{f}(y)\Big) = p_F^2(x) = p_F(x) = \tilde{f}(y)$$

donc $\tilde{f} \circ f \circ \tilde{f}(y) = \tilde{f}(y)\,, \quad \forall y \in E'$.

Notons aussi que l'inverse généralisée de f n'est pas unique, car elle dépend du choix des supplémentaires F et G.

Théorème A.5.2 *– Soient E et E' deux espaces vectoriels sur le même corps K et $f : E \to E'$ une application linéaire. Considérons l'équation en $x \in E$ comme inconnue :*

$$f(x) = b\,, \qquad \textit{où} \quad b \in E' \quad \textit{est donné.}$$ [1]

Cette équation admet une solution si et seulement si :

$$f \circ \tilde{f}(b) = b$$

où $\tilde{f}$ est une inverse généralisée quelconque de b. De plus, toute solution est alors de la forme :

$$x = \tilde{f}(b) + (\tilde{f} \circ f - \mathrm{id}_E)\, y$$

où y est un vecteur arbitraire de E.

Démonstration :
La démonstration se fait par les lemmes suivants :

Lemme 1. - *$f \circ \tilde{f}$ et $\tilde{f} \circ f$ sont des projecteurs.*

En effet :

$$(f \circ \tilde{f})^2 = \underbrace{f \circ \tilde{f} \circ f}_{f} \circ \tilde{f} = f \circ \tilde{f} \quad \text{et} \quad (\tilde{f} \circ f)^2 = \tilde{f} \circ \underbrace{f \circ \tilde{f} \circ f}_{f} = \tilde{f} \circ f \qquad \diamondsuit$$

Lemme 2. -

$$\begin{aligned} \mathrm{Ker}(f \circ \tilde{f} - \mathrm{id}_{E'}) &= \mathrm{Im}\, f \\ \mathrm{Im}(\tilde{f} \circ f - \mathrm{id}_E) &= \mathrm{Ker.}\, f \end{aligned}$$

Pour montrer ce lemme, notons d'abord que pour tout projecteur p on a : $\mathrm{Ker}(p - \mathrm{id}) = \mathrm{Im}\, p$ (cf. exercice 9, chapitre 3). Puisque $f \circ \tilde{f}$ est un projecteur, on a :

$$\mathrm{Ker}\,(f \circ \tilde{f} - \mathrm{id}) = \mathrm{Im}\, f \circ \tilde{f} \;\subset \mathrm{Im}\, f.$$

D'autre part, de $f \circ \tilde{f} \circ f = f$ on en déduit : $(f \circ \tilde{f} - \mathrm{id}) \circ f = 0$, c'est-à-dire $\mathrm{Im}\, f \subset \mathrm{Ker}\,(f \circ \tilde{f} - \mathrm{id})$.
D'une manière semblable on montre que $\mathrm{Im}(\tilde{f} \circ f - \mathrm{id}) = \mathrm{Ker}\, f$, et donc $\mathrm{Im}\, f = \mathrm{Ker}(f \circ \tilde{f} - \mathrm{id})$. $\diamondsuit$

[1] Il s'agit, bien entendu de l'expression intrinsèque d'un système d'équations linéaires.

Lemme 3. - *L'équation* $f(x) = b$ *admet des solutions si et seulement si* $f \circ \tilde{f}(b) = b$

Cela est immédiat, puisque dire que $f(x) = b$ admet une solution signifie que $b \in \operatorname{Im} f$ et l'on vient de voir que $\operatorname{Im} f = \operatorname{Ker}(f \circ \tilde{f} - \mathrm{id})$. $\diamondsuit$

Lemme 4. - *La solution générale de l'équation* $f(x) = b$ *est donnée par :* $x = x_0 + z$ *, où* x_0 *est une solution particulière de la même équation et* z *la solution générale de* $f(x) = 0$ *(c'est-à-dire un élément quelconque de* $\operatorname{Ker} f$*).*

Pour la démonstration, cf. la démonstration de la proposition A.7.4 page 382.

Le théorème se déduit maintenant facilement. Si $f(x) = b$ a une solution, $\tilde{f}(b)$ est une solution particulière, d'après le lemme 3. La solution générale est donc :
$x = \tilde{f}(b) + (\tilde{f} \circ f - \mathrm{id})\, y$, avec $y \in E$, d'après les lemmes 2. et 4. $\square$

Inverse généralisée de Penrose-Moore

Nous avons vu que la construction d'une inverse généralisée dépend du choix des supplémentaires de $\operatorname{Ker} f$ et de $\operatorname{Im} f$. Si l'on suppose que E et E' sont des espaces hermitiens, il existe un choix canonique pour le supplémentaire : le supplémentaire orthogonal. On en déduit que si E et E' sont des espaces hermitiens, il existe une inverse généralisée canonique : c'est l'inverse généralisée de Penrose-Moore.

Théorème A.5.3 – *Soient* E *et* E' *deux espaces hermitiens (ou euclidiens) et* $f : E \longrightarrow E'$ *une application linéaire. L'inverse généralisée de Penrose-Moore est l'unique application* $f^\dagger$: $E' \longrightarrow E$ *qui vérifie les propriétés suivantes :*

1. $f \circ f^\dagger \circ f = f$
2. $f^\dagger \circ f \circ f^\dagger = f^\dagger$
3. $(f \circ f^\dagger)^* = f \circ f^\dagger$
4. $(f^\dagger \circ f)^* = f^\dagger \circ f$

On connaît déjà 1. et 2. ; pour montrer 3. et 4. il suffit de remarquer qu'un projecteur est orthogonal si et seulement si il est autoadjoint (cf. exercice 13, chapitre 8).

Notons en effet que :

$f \circ f^\dagger$ *est le projecteur orthogonal sur* $\operatorname{Im} f$

$f^\dagger \circ f$ *est le projecteur orthogonal sur* $(\operatorname{Ker} f)^\perp$

En effet $\operatorname{Im}(f \circ f^\dagger) \subset \operatorname{Im} f$; comme $f \circ f^\dagger$ est un projecteur orthogonal, cela suffit pour assurer que $\operatorname{Im}(f \circ f^\dagger) = \operatorname{Im} f$ (cf. exercice 16, chapitre 7).
De même, puisque $\operatorname{Im}(f^\dagger \circ f) \subset (\operatorname{Ker} f)^\perp$, $f^\dagger \circ f$ est le projecteur orthogonal sur $(\operatorname{Ker} f)^\perp$.

Penrose a démontré qu'il existe une et une seule application $f^\dagger : E' \longrightarrow E$ qui vérifie les propriétés 1, 2, 3. et 4. La démonstration de ce résultat est laissée en exercice et se fait sans difficultés par les étapes suivantes :

a) Soient E et E' deux espaces hermitiens et $f : E \longrightarrow E'$ une application linéaire. On considère une application $f^\dagger : E' \longrightarrow E$ telle que :

i) $f \circ f^\dagger \circ f = f$
ii) $f^\dagger \circ f \circ f^\dagger = f^\dagger$;

montrer que $E = \operatorname{Ker} f \oplus \operatorname{Im} f^\dagger$ et $E' = \operatorname{Ker} f^\dagger \oplus \operatorname{Im} f$ (cf. exercice 11, chapitre 3).

b) En déduire que si $f^\dagger$ vérifie les propriétés i) et ii), alors elle est une inverse généralisée de f.

c) Montrer qu'un projecteur d'un espace hermitien est un projecteur orthogonal si et seulement si il est autoadjoint (cf. exercice 13, chapitre 8).

En termes de matrices ce résultat s'énonce ainsi :

Théorème de Penrose-Moore . A.5.4 – *Soit $A \in \mathcal{M}_{p,n}(\mathbb{C})$. Il existe une et une seule matrice $A^\dagger \in \mathcal{M}_{n,p}(\mathbb{C})$ qui vérifie :*

1. $A\,A^\dagger A = A$
2. $A^\dagger A\,A^\dagger = A^\dagger$
3. $(A\,A^\dagger)^* = A\,A^\dagger$
4. $(A^\dagger A)^* = A^\dagger A$ *(où : $A^* := {}^t\overline{A}$).*

$A^\dagger$ est dite ***inverse généralisée de Penrose-Moore***

Formule de Mac Duffee pour le calcul explicite de $A^\dagger$.

Soit $A = \| c_1, \ldots, c_n \| \in \mathcal{M}_{p,n}(\mathbb{C})$. Supposons A de rang r et soient $V_1, \ldots, V_r$ r vecteurs indépendants choisis parmi les c_i. Puisque les c_i s'écrivent comme combinaison linéaire des $V_1, \ldots, V_r$, il existe une unique matrice $C \in \mathcal{M}_{r,n}(\mathbb{C})$ telle que :

$$A = B \cdot C$$

où $B = \| V_1, \ldots, V_r \| \in \mathcal{M}_{p,r}(\mathbb{C})$. Si $c_i = \sum_{k=1}^{r} \lambda_{ki}\, V_k$ pour $i = 1, \ldots, n$, alors :

$$C = \begin{pmatrix} \lambda_{11} & \cdots & \lambda_{1n} \\ \vdots & & \vdots \\ \lambda_{r1} & \cdots & \lambda_{rn} \end{pmatrix} = \| c_1, \ldots, c_n \|_{V_i}$$

La matrice $B^* A\, C^* \in \mathcal{M}_r(\mathbb{C})$ est inversible. En effet on peut écrire : $B^* A\, C^* = (B^* B)(C\, C^*)$; puisque B et C sont de rang r, il n'est pas difficile de montrer que $B^* B$ et $C\,C^*$ sont inversibles [2]. La formule de Mac Duffee s'écrit :

$$\boxed{A^\dagger = C^* \cdot (B^* A\, C^*)^{-1} \cdot B^*}$$

Cette formule se démontre facilement en écrivant le second membre sous la forme :

$$C^*(CC^*)^{-1}(B^*B)^{-1}B^*$$

et en montrant que cette matrice vérifie les axiomes de Penrose.

Exemple –

Calculer l'inverse généralisée de Penrose-Moore de la matrice :

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

[2] Pour montrer que $B^* B$ est inversible, cf. exercice 23 chapitre 7 ; pour montrer que $C\,C^*$ est inversible, remplacer B par C^*.

On a rang $A = 2$. Une base de l'espace engendré par les vecteurs colonnes est $\{V_1 = c_1\ , V_2 = c_2\}$, aussi :

$$B = \|\, V_1,\, V_2 \,\| = \begin{pmatrix} 2 & 1 \\ 1 & 1 \\ 3 & 2 \end{pmatrix}$$

D'autre part : $c_3 = -V_1 + V_2$, $c_4 = V_2$, donc :

$$C = \|\, c_1,\, c_2,\, c_3,\, c_4 \,\|_{V_i} = \begin{pmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 1 & 1 \end{pmatrix}$$

Le calcul donne :

$$A^\dagger = \frac{1}{15}\begin{pmatrix} 5 & -4 & 1 \\ -5 & 7 & 2 \\ -10 & 11 & 1 \\ -5 & 7 & 2 \end{pmatrix}$$

Application : Solution des moindres carrés

La notion d'inverse généralisée permet d'étendre la méthode des moindres carrés, que nous avons déjà vue dans un cas particulier (cf. exercice 23, chapitre 7).
Soit $A \in \mathcal{M}_{p,n}(\mathbb{C})$ et considérons le système linéaire

$$A\,x = b$$

Le théorème de Rouché-Fontené permet d'avoir les conditions de compatibilité sous forme explicite d'équations, par l'annulation des déterminants caractéristiques. Notons cependant que la notion de compatibilité est "instable" : une petite variation des coefficients peut faire en sorte que les égalités ne soient plus satisfaites et qu'un système compatible devienne incompatible. Or si l'on considère un système qui règle un phénomène physique, les coefficients de la matrice sont donnés par des mesures (par exemple de la température, de la pression, etc.) et donc sont connus avec une précision limitée, c'est-à-dire en fait, à une petite variation près. Ceci veut dire que les conditions de compatibilité n'ont pas, en fait, une vraie signification physique : les résultats de mesures physiques qui déterminent la matrice A pourraient donner un système incompatible, alors que, peut-être, pour des raisons physiques, il pourrait être clair que le système admet des solutions. Il est donc important d'avoir une valeur approchée des solutions, même dans les cas où le système apparaît comme incompatible.

Définition A.5.5 – *Soit $A\,x = b$ un système linéaire, avec $A \in \mathcal{M}_{p,n}(\mathbb{C})$, $x \in \mathbb{C}^n$ et $b \in \mathbb{C}^p$. On appelle* ***solution des moindres carrés*** *le vecteur $x_0 \in \mathbb{C}^n$ tel que*

$$\|\, A\,x_0 - b \,\| = \operatorname*{Inf}_{x \in \mathbb{C}^n} \|\, A\,x - b \,\|.$$

Théorème A.5.6 – *La solution des moindres carrés du système $A\,x = b$ est donnée par*

$$x_0 = A^\dagger b.$$

La démonstration est immédiate (cf. solution de l'exercice 23, chapitre 7). On a :

$$\|\, A\,x_0 - b \,\| = \|\, A\,A^\dagger\, b - b \,\| = \|\, p_{\operatorname{Im} f}(b) - b \,\| = \mathrm{d}(b, \operatorname{Im} f) = \operatorname*{Inf}_{x \in \mathbb{C}^n} \|\, A\,x - b \,\|$$

d étant la distance associée à la norme. □

Pour la théorie et les applications des inverses généralisées, on pourra consulter

A. Ben-Israel, T.N.E. Greville
Generalised inverses : theory and applications
A. Wiley - Interscience Publication
New York, 1974

Il s'agit d'un livre exigeant un niveau "avancé" de connaissance et de pratique de la théorie des matrices.

Appendice A.6

Exponentielle d'une matrice

Suites et séries de matrices

Comme nous l'avons vu (cf. exercice 5, chapitre 6.), $\mathcal{M}_n(\mathbb{C})$ est un espace vectoriel normé. Si $A = \left(a_{ij}\right)$, on pose (cf. exemple 4 page 222) :

$$\| A \| \ := \sqrt{\sum_{i,j=1}^{n} |a_{ij}|^2} = \mathrm{Tr}(A^* A)$$

On en déduit que $\mathcal{M}_n\,(\mathbb{C})$ est un espace métrique pour la distance définie par

$$\mathrm{d}(A, B) \ := \|A - B\|$$

(cf. exercice 10, chapitre 7). On montre facilement que

$$\|AB\| \leq \|A\| \cdot \|B\|$$

d'où :

$$\|A^n\| \leq \|A\|^n$$

La vérification de ces inégalités est laissée en exercice.

Les définitions et les propriétés des suites et séries numériques passent facilement à l'espace métrique $(\mathcal{M}_n\,(\mathbb{C}), \mathrm{d})$. On a par exemple :

1. On dit qu'une suite de matrices $(A_n)_{n\in\mathbb{N}}$ converge vers la matrice A si

$$\forall \varepsilon > 0 \ \exists n_0 \in \mathbb{N} \ \text{ tel que } \ \forall n \geq n_0 \ : \ \mathrm{d}(A_n, A) \leq \varepsilon$$

2. On dit qu'une série de matrices $\sum_{n=0}^{+\infty} A_n$ converge si la suite des sommes partielles $S_k = \sum_{n=0}^{k} A_n$ converge ; dans ce cas, la limite pour $k \to +\infty$ de S_k est dite somme de la série. On dit qu'une série de matrices est absolument convergente si la série numérique $\sum_{n=0}^{+\infty}\|A_n\|$ est convergente.

3. Soient $\sum_{n=0}^{+\infty} A_n$ et $\sum_{n=0}^{+\infty} \mathcal{B}_n$ deux séries de matrices absolument convergentes, de somme respectivement A et B. Alors les séries $\sum_{n=0}^{+\infty} (A_n + B_n)$ et $\sum_{n=0}^{+\infty} C_n$ (où $C_n = \sum_{k=0}^{n} A_k\, B_{n-k}$) sont convergentes et ont comme somme respectivement $A + B$ et $A\,B$.

La démonstration est analogue à celle que l'on donne pour les séries numériques, la valeur absolue étant remplacée par la norme.

Exponentielle d'une matrice

Proposition A.6.1 – *Soit $A \in \mathcal{M}_n(\mathbb{C})$. La série de matrices de terme général $\dfrac{A^n}{n!}$ est absolument convergente, quelle que soit la matrice A. Sa somme est notée e^A :*

$$\boxed{e^A = I_n + A + \frac{A^2}{2!} + \frac{A^3}{3!} + \cdots + \frac{A^n}{n!} + \cdots}$$

La démonstration est immédiate. On a : $\left\| \dfrac{A^k}{k!} \right\| \leq \dfrac{\|A\|^k}{k!}$. Puisque la série numérique $\displaystyle\sum_{k=0}^{+\infty} \frac{\|A\|^k}{k!}$ est convergente (sa somme est $e^{\|A\|}$) on en déduit que $\displaystyle\sum_{k=0}^{+\infty} \left\| \frac{A^k}{k!} \right\|$ est convergente. Donc $\displaystyle\sum_{n=0}^{+\infty} \frac{A^n}{n!}$ est une série de matrices absolument convergente. $\square$

Exemple : –

$$\boxed{e^{\lambda I} = e^{\lambda} I}$$

car $e^{\lambda I} = \lambda I + \lambda I + \dfrac{\lambda^2 I}{2!} + \cdots + \dfrac{\lambda^k I}{k!} + \cdots = e^{\lambda} I$

En particulier : $e^0 = I_n$ (0 étant ici la matrice nulle de $\mathcal{M}_n(\mathbb{C})$)

Proposition A.6.2 – *Si A et B **commutent**, alors :*

$$\boxed{e^{A+B} = e^A \cdot e^B}$$

En effet, puisque A et B commutent, on peut appliquer la formule du binôme (cf. exercice 22, chapitre 3). On a :

$$(A+B)^n = \sum_{k=0}^{n} C_n^k A^k B^{n-k}$$

donc :

$$\begin{aligned} e^{A+B} &= \sum_{n=0}^{+\infty} \frac{(A+B)^n}{n!} = \sum_{n=0}^{+\infty} \sum_{k=0}^{n} \frac{n!}{n!\,k!\,(n-k)!} A^k B^{n-k} \\ &= \sum_{n=0}^{+\infty} \sum_{k=0}^{n} \frac{A^k}{k!} \frac{B^{n-k}}{(n-k)!} = e^A \cdot e^B \end{aligned}$$

$\square$

Corollaire A.6.3 – *Pour toute matrice $A \in \mathcal{M}_n(\mathbb{C})$, même non inversible, la matrice e^A est inversible et son inverse est e^{-A}.*

En effet A et $-A$ commutent, donc $e^A e^{-A} = e^{A-A} = e^0 = I$.

Proposition A.6.4 – *Si $B = P^{-1} A P$, alors $e^B = P^{-1} e^A P$.*

En effet, $B^n = P^{-1} A^n P$; donc :

$$\begin{aligned} e^B &= \sum_{n=0}^{+\infty} \frac{B^n}{n!} = \sum_{n=0}^{+\infty} \frac{P^{-1} A^n P}{n!} = \lim_{n\to+\infty} \sum_{k=0}^{n} P^{-1} \frac{A^k}{k!} P = \\ &= \lim_{n\to+\infty} P^{-1} \left(\sum_{k=0}^{n} \frac{A^k}{k!} \right) P \overset{(1)}{=} P^{-1} \lim_{n\to+\infty} \sum_{k=0}^{n} \frac{A^k}{k!} P = P^{-1} e^A P. \end{aligned}$$

(Le passage (1) est justifié par le fait que l'application $\begin{array}{ccc}\mathcal{M}_n(\mathbb{C}) & \to & \mathcal{M}_n(\mathbb{C})\\ M & \longmapsto & P^{-1}MP\end{array}$ est continue). □

Corollaire A.6.5 – $$\boxed{\text{dét}\,(e^A) = e^{\operatorname{Tr} A}}$$

En effet, puisque $A \in \mathcal{M}_n(\mathbb{C})$, A est trigonalisable. On a donc $A = P^{-1}\,B\,P$ avec :

$$B = \begin{pmatrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}. \quad \text{Notons que : } B^k = \begin{pmatrix} \lambda_1^k & & * \\ & \ddots & \\ 0 & & \lambda_n^k \end{pmatrix} \text{ donc :}$$

$$\begin{aligned} e^A &= P^{-1}\,e^B\,P = P^{-1}\left(I + B + \frac{B^2}{2!} + \cdots + \frac{B^k}{k!} + \cdots\right)P = \\ &P^{-1}\begin{pmatrix} 1+\lambda_1+\cdots+\dfrac{\lambda_1^k}{k!}+\cdots & & * \\ & \ddots & \\ 0 & & 1+\lambda_n+\cdots+\dfrac{\lambda_n^k}{k!}+\cdots \end{pmatrix}P = \\ &= P^{-1}\begin{pmatrix} e^{\lambda_1} & & * \\ & \ddots & \\ 0 & & e^{\lambda_n} \end{pmatrix}P \quad , \quad \text{d'où : } \quad \text{dét}\, e^A = e^{\lambda_1+\cdots+\lambda_n} = e^{\operatorname{Tr} A} \quad \square \end{aligned}$$

Méthode de calcul de e^A par la réduction de Jordan

Lemme -

$$\textit{Soit} \quad A = \begin{pmatrix} \boxed{A_1} & & & 0 \\ & \boxed{A_2} & & \\ & & \ddots & \\ 0 & & & \boxed{A_r} \end{pmatrix} \qquad \textit{avec } A_i \textit{ matrices carrées complexes}$$

$$\textit{Alors :} \qquad e^A = \begin{pmatrix} \boxed{e^{A_1}} & & & 0 \\ & \boxed{e^{A_2}} & & \\ & & \ddots & \\ 0 & & & \boxed{e^{A_r}} \end{pmatrix}$$

Ce lemme se démontre facilement en utilisant le fait que :

$$A^k = \begin{pmatrix} \boxed{A_1^k} & & & 0 \\ & \boxed{A_2^k} & & \\ & & \ddots & \\ 0 & & & \boxed{A_r^k} \end{pmatrix}$$. Puisque le polynôme caractéristique de A est scindé, A admet une réduction de Jordan : il existe P inversible et une matrice

$$B = \begin{pmatrix} \boxed{J_1} & & & 0 \\ & \boxed{J_2} & & \\ & & \ddots & \\ 0 & & & \boxed{J_r} \end{pmatrix}, \quad \text{avec } J_\alpha = \begin{pmatrix} \lambda_1 & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda_\alpha \end{pmatrix}$$

telles que $A = P^{-1}\,B\,P$.

Or $J_\alpha = \lambda_\alpha I + V_\alpha$ et V_α est nilpotente. Par ailleurs $\lambda_\alpha I$ et V_α commutent, donc, si p est l'indice de nilpotence de V_α :

$$e^{J_\alpha} = e^{\lambda_\alpha I}\, e^{V_\alpha} = e^{\lambda_\alpha}\, I\, e^{V_\alpha} = e^{\lambda_\alpha}\left(I + V_\alpha + \frac{V_\alpha^2}{2\,!} + \cdots + \frac{V_\alpha^{p-1}}{(p-1)\,!}\right)$$

ce qui permet, compte tenu du lemme, le calcul effectif de e^A.

Application aux systèmes différentiels linéaires à cœfficients constants

Proposition A.6.6 – *Soit $t \in \mathbb{R}$; on a :*

$$\boxed{\frac{\mathrm{d}}{\mathrm{d}t}(e^{tA}) = A\, e^{tA}}$$

Démonstration :

$$\frac{\mathrm{d}e^{\,tA}}{\mathrm{d}t} = \lim_{h\to 0} \frac{e^{(t+h)A} - e^{tA}}{h} = \lim_{h\to 0}\left(\frac{e^{hA} - I}{h} \cdot e^{tA}\right)$$

Montrons que :

$$\lim_{h\to 0} \frac{e^{hA} - I}{h} = A$$

(ce qui montre la proposition).
On a :

$$e^{hA} = I + hA + \frac{h^2 A^2}{2\,!} + \cdots + \frac{h^n A^n}{n\,!} + \cdots$$

d'où :

$$e^{hA} - I - hA = \frac{h^2 A^2}{2\,!} + \cdots + \frac{h^n A^n}{n\,!} + \cdots$$

On en déduit que :

$$\|e^{hA} - I - hA\| \le \frac{\|hA\|^2}{2\,!} + \cdots + \frac{\|hA\|^n}{n\,!} + \cdots = e^{\|hA\|} - 1 - \|hA\|$$

et donc :

$$\|\frac{e^{hA} - I}{h} - A\| \le \frac{e^{\|hA\|} - 1}{|h|} - \|A\| = \frac{e^{|h|\,\|A\|} - 1}{|h|} - \|A\|$$

Lorsque $h \to 0$: $\dfrac{e^{|h|\,\|A\|} - 1}{|h|} \sim \|A\|$, d'où $\|\dfrac{e^{hA} - I}{h} - A\| \to 0.$ □

Proposition A.6.7 – *Soit $\frac{\mathrm{d}x}{\mathrm{d}t} = A\,x$ un système différentiel linéaire à cœfficients constants $(A \in \mathcal{M}_n(\mathbb{C}), x \in \mathcal{M}_{n,1}(\mathbb{C}))$. Alors la solution qui pour $t = 0$ prend la valeur x_0 $(x_0 \in \mathcal{M}_{n,1}(\mathbb{C}))$ est :*

$$x(t) = e^{tA}\, x_0$$

En effet soit x défini par $x := e^{tA}\, x_0$. On a : $\frac{\mathrm{d}x}{\mathrm{d}t} = A\, e^{tA}\, x_0 = A\, x$; donc $e^{tA}\, x_0$ est bien solution. D'autre part $x(0) = e^0\, x_0 = I\, x_0 = x_0$; ce qui montre que x vérifie la condition initiale. □

Exemple *– Soit le système* $\frac{\mathrm{d}X}{\mathrm{d}t} = A\,X$*, où :*

$$A = \begin{pmatrix} b & a & \cdots & \cdots & a \\ a & b & a & \cdots & a \\ a & a & b & \cdots & a \\ & & & \ddots & \\ a & \cdots & & & \\ a & \cdots & & a & b \end{pmatrix}$$

Déterminer la solution qui pour $t = 0$ *prend la valeur* x_0*, où* x_0 *appartient à l'hyperplan* F *défini par* $x_1 + \cdots + x_n = 0$*.*

On a $A = (b-a)\,I + a\,U$, où U est la matrice dont tous les cœfficients sont égaux à 1. Puisque I et U commutent,

$$\begin{aligned} e^{t\,A} &= e^{(b-a)t\,I}\,e^{at\,U} = e^{b-a}\left[I + \sum_{k=1}^{+\infty}\frac{(at\,U)^k}{k\,!}\right] = e^{(b-a)t}\left[I + \sum_{k=1}^{+\infty}\frac{(at)^k\,n^{k-1}}{k\,!}\,U\right] \\ &= e^{(b-a)t}\left[I + \frac{1}{n}(e^{atn} - 1)\,U\right] \end{aligned}$$

Or l'hyperplan $F\,:\,x_1 + \cdots + x_n = 0$ est le noyau de U, aussi, la solution cherchée est : $x = e^{(b-a)t}\,x_0$. C'est-à-dire : les solutions qui passent par l'hyperplan F demeurent dans F.

Plus généralement on a :

Corollaire A.6.8 *– Soit* $\dfrac{\mathrm{d}x}{\mathrm{d}t} = A\,x$ *un système différentiel linéaire à cœfficients constants. Si une solution passe à un instant donné par un sous-espace propre* E_λ*, elle reste dans* E_λ *(on dit que les sous-espaces propres sont «invariants» sous le flot des solutions).*

En effet si $x_0 \in E_\lambda$, on a :

$$e^{tA}\,x_0 = \sum_{n=0}^{\infty}\frac{A^n}{n\,!}\,x_0 = \sum_{n=0}^{\infty}\frac{\lambda^n}{n\,!}\,x_0 = e^{\lambda t}\,x_0 \in E_\lambda. \quad \square$$

L'exponentielle des matrices joue un rôle important dans la Géométrie Différentielle moderne depuis les travaux de Sophus Lie à la fin du XIXème siècle. Il est difficile d'en donner une idée même approximative, car les concepts qui rentrent en jeu sont assez élaborés. Disons, en simplifiant, que le problème abordé par Lie, problème encore largement ouvert aujourd'hui, était de chercher à classifier les équations différentielles qui se déduisent les unes des autres par changement de variables. Il était donc amené à étudier les groupes de transformations (cf. Appendice A.1. page 342) dependant d'un certain nombre (fini) de paramètres, ceux-ci étant des fonctions différentiables – c'est le cas par exemple de $\mathrm{O}(n,\mathbb{R})$, $\mathrm{U}(n,\mathbb{C})$, etc. – plus précisément ce que l'on appelle aujourd'hui les *groupes de Lie*, qui sont très souvent des groupes de matrices. On peut démontrer que si G est un groupe de Lie de matrices à cœfficients dans K, ($K = \mathbb{R}$ ou $\mathbb{C}$), toutes les informations essentielles sur G sont contenues, dans un certain sens, dans les propriétés de l'ensemble

$$\mathfrak{g} = \left\{\,A \in \mathcal{M}_n(K)\ \mid\ e^A \in G\,\right\}.$$ [1]

[1] Un peu comme les informations sur une fonction sont contenues dans sa dérivée. L'analogie n'est pas seulement formelle. En effet, si on considère une courbe $c :]-\varepsilon, +\varepsilon[\to G$, définie par $c(t) = e^{tA}$, on a, d'après la proposition A.6.6 : $\left[\frac{\mathrm{d}c(t)}{\mathrm{d}t}\right]_{t=0} = A$.

Cet ensemble, dit «algèbre de Lie» de G, est en fait beaucoup plus simple à manier que G. En effet on peut vérifier tout d'abord que $\mathfrak{g}$ est un sous-espace vectoriel de $\mathcal{M}_n(K)$ ce qui permet d'utiliser les techniques de l'algèbre linéaire. D'autre part le crochet défini par

$$[A, B] := AB - BA$$

permet de munir $\mathfrak{g}$ justement d'une structure d'*algèbre de Lie* (cf. page 250), structure très riche et bien étudiée, qui donne des informations importantes sur G.

Appendice A.7

Espaces affines

Comme nous l'avons vu au chapitre 1, la notion d'espace vectoriel a été introduite pour rendre compte des propriétés de certaines grandeurs physiques – comme la force, la vitesse, l'accélération, etc. – qui sont déterminées non seulement par leur intensité, mais aussi par la direction et le sens selon lesquels elles s'exercent. Ces grandeurs sont représentées graphiquement par des flèches, dites «vecteurs», ayant comme origine le point d'application.

Puisque la loi d'addition et la structure d'espace vectoriel ne peuvent être définies que sur le vecteurs de même origine (cf. chapitre 1), le modèle qui permet d'avoir une intuition géométrique de la notion d'espace vectoriel est l'espace (ou le plan) ordinaire *dans lequel on fixe un point* O : les points P du plan ou de l'espace sont en correspondance bijective avec les vecteurs issus de O et d'extrémité P (cf. figure 1.).

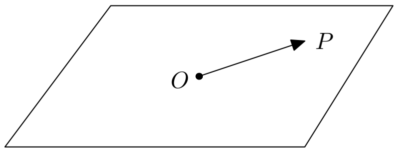


Figure 1

Le modèle naturel de l'espace affine est en revanche le plan ou l'espace ordinaire dans lequel *on ne choisit pas de point privilégié* et les éléments sont les points eux-mêmes. La notion d'espace affine est donc plus 'naturelle', plus 'primitive' dans un certain sens, que la notion d'espace vectoriel. Cependant, comme nous allons le voir, pour définir correctement les espaces affines et pour effectuer les calculs, on a besoin de la notion d'espace vectoriel. En fait les deux notions, d'espace vectoriel et d'espace affine, se trouvent imbriquées dans la pratique, c'est pourquoi, si l'on veut avoir une notion intuitive, il peut être utile d'avoir présent à l'esprit le modèle suivant où les deux structures se présentent simultanément.[1]

Considérons l'espace vectoriel $\mathbb{R}^3$. Un sous-ensemble $\mathcal{A}$ de $\mathbb{R}^3$ est dit sous-espace affine s'il est l'image d'un sous-espace vectoriel F par une translation. F sera dit «direction» de $\mathcal{A}$ (cf. figure 2.)

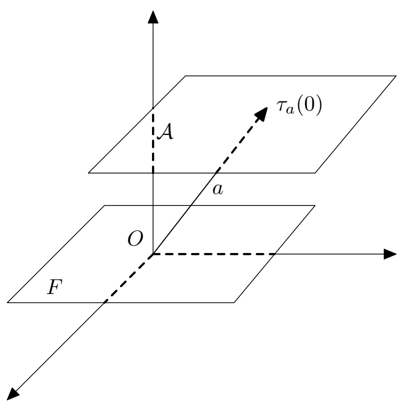


Figure 2

Notons que dans un sous-espace vectoriel F il y a un élément privilégié (l'élément 0), ce qui permet d'identifier les points P de F à des «vecteurs» (d'origine 0 et d'extrémité P).
Cela, en revanche, n'est plus le cas pour le sous-espace affine $\mathcal{A}$: dans un plan qui ne passe pas par l'origine, il n'y a pas d'élément privilégié et l'on ne peut pas mettre en correspondance bijective ses points avec des vecteurs qui sont dans $\mathcal{A}$.
En revanche, si l'on fixe un point O de $\mathcal{A}$, on peut identifier les points de $\mathcal{A}$ à des vecteurs d'origine O (cf. figure 3) et faire de $\mathcal{A}$ un espace vectoriel (dit *vectorialisé* de $\mathcal{A}$), comme nous le verrons d'une manière plus précise ci-dessous.

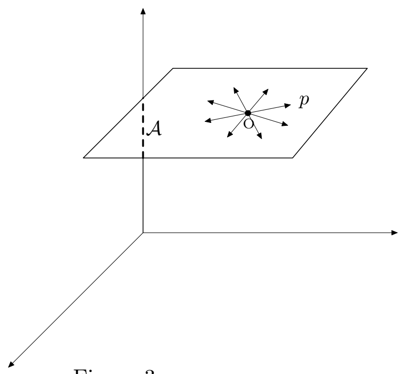


Figure 3

Le fait que dans un espace affine il n'y ait pas d'élément privilégié, implique la perte de la structure d'espace vectoriel. Cependant, c'est en cela entre autres que réside l'intérêt de la notion d'espace affine : car du moment qu'il n'y a pas un élément qui joue un rôle le privilégié, on a plus de souplesse pour choisir l'origine des coordonnées de manière à simplifier les problèmes (par exemple, l'équation d'un cercle est plus simple si on l'écrit dans un système de coordonnées dont l'origine est le centre du cercle).

I. Sous-variétés affines d'un espace vectoriel et applications affines

Sous-variétés affines d'un espace vectoriel

Soit E un espace vectoriel sur K et $\vec{v} \in E$. On note $t_{\vec{v}}$ l'application :

$$t_{\vec{v}} \ : \begin{array}{ccc} E & \longrightarrow & E \\ a & \longmapsto & a + \vec{v} \end{array}$$

$t_{\vec{v}}$ est appelée ***translation de vecteur directeur*** $\vec{v}$.

Notons que, sauf si $\vec{v} = 0$, $t_{\vec{v}}$ n'est pas une application linéaire.

Définition A.7.1 – *Soit F un sous-espace vectoriel de E et $\vec{v} \in E$; l'image de F par $t_{\vec{v}}$:*

$$t_{\vec{v}}(F) = \Big\{ b \in E \mid b = a + \vec{v}\,,\ a \in F \Big\} \underset{\text{notation}}{=} \vec{v} + F$$

*est dite **sous-variété affine** ou **sous-espace affine** de **direction** F.*

Si F est une droite vectorielle (resp. : plan vectoriel), $t_{\vec{v}}(F)$ est dit *droite affine* (resp. : *plan affine*). Bien entendu, E lui-même est un espace affine de direction E (car $E = t_{\vec{v}}(E)$).

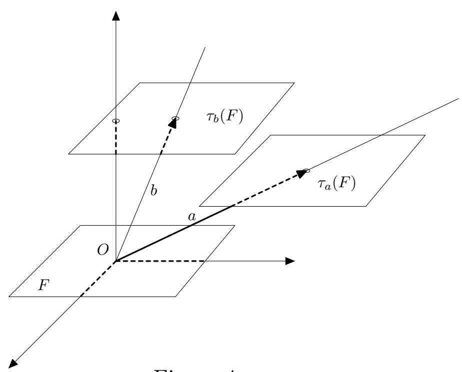


Figure 4

Un espace affine $\mathcal{A} = t_{\vec{v}}(F)$ n'est pas en général un sous-espace vectoriel, car il n'est pas stable pour les lois de somme et de produit par un scalaire, à moins qu'il passe par 0, ce qui se produit si $\vec{v} \in F$: on a alors $\mathcal{A} = F$.

On peut cependant mettre sur un espace affine une structure d'espace vectoriel de la manière suivante. Soit $\mathcal{A}$ un sous-espace affine de direction F. On choisit un point $a \in \mathcal{A}$; on a : $\mathcal{A} = t_a F$. Il suffit alors de transporter sur $\mathcal{A}$ la structure d'espace vectoriel de F par la bijection t_a (cf. exercice 4, chapitre 3). On obtient les lois suivantes :

si $b = a + x$ et $c = a + y$ (avec $x, y \in F$), on pose :

$$\begin{aligned} b \dot{+} c \ &:= (x + y) + a \\ \lambda \cdot b \ &:= \lambda\, x + a \end{aligned}$$

qui font de $\mathcal{A}$ un espace vectoriel, dit ***vectorialisé*** de $\mathcal{A}$ en a.

Notons que cette structure d'espace vectoriel n'est pas canonique, mais dépend du choix de a.
Dans la suite les éléments d'un sous-espace affine $\mathcal{A}$ seront notés, en général, par $a, b, c, \ldots$, et on notera $x, y, v, w, \ldots$ les éléments de la direction F de $\mathcal{A}$ (ou éventuellement $\vec{x}, \vec{y}, \vec{z}, \vec{v}, \vec{w}, \ldots$ pour mieux marquer qu'ils sont les éléments d'un espace vectoriel).

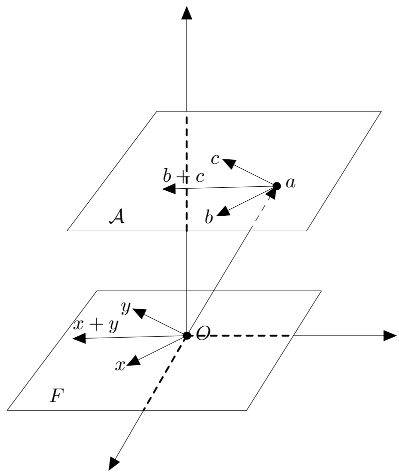


Figure 5

Applications affines

Définition A.7.2 – *Soient E et E' deux espaces vectoriels sur K. On appelle* ***application affine*** *toute application $g : E \to E'$ du type : $g = t_a \circ f$, où $f \in \mathcal{L}_K(E, E')$ et $a \in E$, c'est-à-dire :*

$$g(x) = a + f(x)\,, \quad \text{avec } f \text{ linéaire.}$$

Notons que si $a \neq 0$, g n'est pas linéaire.

Toute application affine s'écrit d'une manière *unique* sous la forme $g = t_a \circ f$, avec $a \in E'$ et $f \in \mathcal{L}_K(E, E')$. En effet, si $g = t_a \circ f = t_{a'} \circ f'$ (avec $a' \in E'$ et $f' \in \mathcal{L}_K(E, E')$), on a :

$$a + f(x) = a' + f'(x)\,, \qquad \forall x \in E.$$

En faisant $x = 0$, on trouve $a = a'$, d'où $f(x) = f'(x)$, $\quad \forall x \in E$.

Si $g = t_a \circ f$, f est dite *application linéaire associée à g* et a est dit *vecteur de translation.*

Par exemple :

$$\begin{aligned} g \ : \quad \mathbb{R}^3 \quad &\longrightarrow \quad \mathbb{R}^3 \\ (x_1, x_2, x_3) \quad &\longrightarrow \quad (2x_1 - x_2 + 3\,,\ -x_1 + x_3 + 2\,,\ x_2 - x_3) \end{aligned}$$

est une application affine de vecteur de translation $a = (3, 2, 0)$ et d'application linéaire associée :

$$\begin{aligned} f \ : \quad \mathbb{R}^3 \quad &\longrightarrow \quad \mathbb{R}^3 \\ (x_1, x_2, x_3) \quad &\longrightarrow \quad (2x_1 - x_2\,,\ -x_1 + x_3\,,\ x_2 - x_3) \end{aligned}$$

Il est facile de se convaincre que les applications affines se caractérisent par le fait que les composantes de l'image du vecteur x sont des polynômes, non homogènes en général, de degré 1 en les composantes de x.

Notons que puisque $f(0) = 0$, on a $g(x) = g(0) + f(x)$. On a donc :

* *Les applications affines sont les applications $g : E \longrightarrow E'$ caractérisées par le fait que l'application $x \longmapsto g(x) - g(0)$* est linéaire.

Proposition A.7.3 – *Les applications affines transforment les sous-espaces affines en sous-espaces affines.*

Soit en effet $g = t_a \circ f$ une application affine et $\mathcal{A} = t_b(F)$ un sous-espace affine de direction F. On a :

$$\begin{aligned} g(A) &= (t_a \circ f)\Big(t_b(F)\Big) = \{\, x \in E \,|\, x = a + f(b + y\,), \ \text{avec } y \in F \,\} \\ &= \{\, x \in E \,|\, x = a + f(b) + f(y\,),\ y \in F \,\} = t_{a+f(b)} f(F) = t_{g(b)} \circ f(F) \end{aligned}$$

Donc $g(A)$ est un sous-espace affine de direction $f(F)$ □

De même que les sous-espaces vectoriels peuvent être caractérisés par un système d'équations linéaires *homogènes*, les sous-espaces affines peuvent être caractérisés par un système d'équations linéaires *non homogènes* :

Proposition A.7.4 ***(Caractérisation des sous-espaces affines par des équations linéaires)*** *Les sous-espaces affines de E sont les ensembles du type :*

$$\mathcal{A} = \{x \in E \mid f(x) = a, \ \text{ avec } \ f \in \mathcal{L}(E, E') \text{ et } \ a \in E' \ \}$$

Démonstration : L'ensemble des solutions du système linéaire $f(x) = a$ est un sous-espace affine de direction $\operatorname{Ker} f$, plus précisément :

$$\mathcal{A} = x_0 + \operatorname{Ker} f \ , \quad \text{avec} \quad x_0 \text{ solution particulière de } \ f(x) = a.$$

En effet, si $x = x_0 + u$, avec $f(u) = 0$ et $f(x_0) = a$, alors $f(x) = a$; aussi : $x_0 + \operatorname{Ker} f \subset \mathcal{A}$. Et inversement, si $x \in \mathcal{A}$, en considérant une solution particulière x_0 de $f(x) = a$, alors $x - x_0 \in \operatorname{Ker} f$, c'est-à-dire $x \in x_0 + \operatorname{Ker} f$, et donc $\mathcal{A} \subset x_0 + \operatorname{Ker} f$.

Réciproquement, soit $\mathcal{A}$ sous-espace affine de direction F, $\mathcal{A} = t_b(F)$, et considérons une application linéaire f de noyau F. On a : $\mathcal{A} = b + F = b + \operatorname{Ker} f = \{\, x = b + y \ \mid \ y \in \operatorname{Ker} f \,\}$, donc : $x \in \mathcal{A}$ si et seulement si $f(x) = a$, où on a posé : $a = f(b)$. □

Ainsi, par exemple, le système

$$\begin{cases} 2\,x + y - z - 3\,t + w = 1 \\ \quad x - y + 2\,z - t - w = 2 \end{cases}$$

définit un sous-espace affine de dimension 2 de $\mathbb{R}^5$.

Proposition A.7.5 – *L'ensemble des applications affines bijectives de E dans lui-même est un groupe (pour la loi de composition des applications) dit* ***groupe affine*** *de E noté* $\mathrm{GA}(E)$.

La vérification est immédiate. Notons en effet qu'une application affine est bijective si et seulement si l'application linéaire associée est bijective. Soient h et h' deux applications affines bijectives : $h(x) = a + f(x)\,,\ h'(x) = b + g(x)$ avec f et g linéaires bijectives. On a :

$$h \circ h'(x) = a + f(b + g(x)) = a + f(b) + f \circ g(x)$$

Ce qui montre que $h \circ h'$ est affine bijective (car $f \circ g$ est bijective).
D'autre part :

$$\begin{aligned} h^{-1}(x) \ = y \iff x = h(y\,) \iff x = a + f(y\,) \iff y = f^{-1}(x - a) \\ \iff y = -f^{-1}(a) + f^{-1}(x) \end{aligned}$$

Donc $h^{-1} = t_{-f^{-1}(a)} \circ f^{-1}$, ce qui montre que h^{-1} est affine. □

Barycentre

Les sous-espaces affines ne sont pas stables, en général, pour la somme (cf. figure 6). Ils sont stables en revanche pour certains types de sommes, dites *sommes barycentriques.* Cette propriété permet d'ailleurs de caractériser les sous-espaces affines et les applications affines.

Définition A.7.6 – *Soit $\mathcal{A}$ un sous-espace affine de E, $a_1, \cdots, a_p \in \mathcal{A}$ et $t_1, \cdots, t_p \in K$, tels que $t_1 + \cdots + t_p = t \neq 0$. Le point :*

$$G = \frac{1}{t}\Big(t_1\, a_1 + \cdots + t_p\, a_p\Big)$$

est dit ***barycentre*** *des points $a_1, \cdots, a_p$ affectés des cœfficients $t_1, \cdots t_p$.*

On voit immédiatement que $G \in \mathcal{A}$. En effet, si $\mathcal{A} = t_a(F)$, on a : $a_i = x_i + a$ avec $x_i \in F$. Donc :

$$G \;=\; \frac{1}{t}\sum_{i=1}^{p} t_i\,(x_i + a) \;=\; a + \frac{1}{t}\Big(t_1\, a_1 + \cdots + t_p\, a_p\Big) \in \mathcal{A}$$

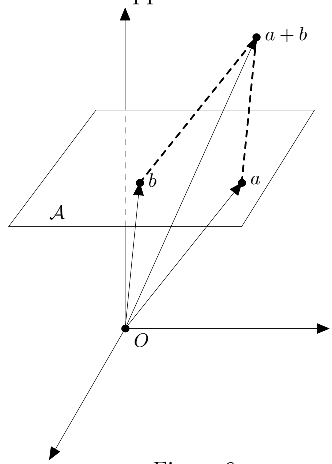


Figure 6

On a en fait la réciproque :

Proposition A.7.7 - ***Caractérisation des sous-espaces affines.***
Un sous-ensemble $\mathcal{A}$ de E est un sous-espace affine si et seulement si il est stable par composition barycentrique (c'est-à-dire le barycentre toute famille finie d'éléments de $\mathcal{A}$ est dans $\mathcal{A}$).

Il s'agit de montrer que si $\mathcal{A}$ est stable par composition barycentrique, alors il est le translaté d'un sous-espace vectoriel F. Fixons $a_0 \in \mathcal{A}$ et montrons que l'ensemble $F := t_{a_0}(\mathcal{A})$ est un sous-espace vectoriel de E. Soit $x = a - a_0 \in F$. On a :

$$\lambda\, x = \lambda\, a - \lambda\, a_0. \tag{i}$$

Or, par hypothèse, le barycentre des points a et a_0 de cœfficients respectivement λ et $(1-\lambda)$ est dans $\mathcal{A}$, c'est-à-dire : $b := \lambda\, a + (1-\lambda)\, a_0 \in \mathcal{A}$. Donc $\lambda\, a_0 = \lambda\, a + a_0 - b$, avec $b \in \mathcal{A}$. En reportant dans (i) on obtient

$$\lambda\, x = \lambda\, a - \lambda\, a - a_0 + b = b - a_0 \in F.$$

F est donc stable pour la loi externe.

De même, soient $x = a - a_0$ et $y = a' - a_0$ dans $\mathcal{A}$. On a :

$$x + y = a + a' - 2\, a_0 = (a + a' - a_0) - a_0.$$

Or $a + a' - a_0$ est le barycentre des points a, a', a_0 affectés des cœfficients $1, 1, -1$, donc il appartient à $\mathcal{A}$, ce qui montre que $x + y \in F$. $\square$

La notion de barycentre permet aussi de caractériser les applications affines :

Proposition A.7.8 - ***Caractérisation des applications affines.***
Soient E et E' deux espaces vectoriels. Une application $g :\, E \to E'$ est affine si et seulement si elle applique le barycentre de tout système de points en le barycentre des images affectées

des mêmes cœfficients (on dit qu'elle «conserve le barycentre»). Pour cela il faut et il suffit que :

$$g(t_1\, a_1 + \cdots + t_p\, a_n) \;=\; t_1\, g(a_1) + \cdots + t_p\, g(a_n)$$
$$\forall\, a_1, \cdots, a_p \in E \;\; et \;\; \forall\, t_1, \cdots, t_p \in K, \; tels\ que : \; t_1 + \cdots + t_p = 1.$$

Démonstration : On voit facilement que toute application affine conserve le barycentre : la vérification est laissée en exercice. Réciproquement supposons que g conserve le barycentre. Alors pour tous $a_1, \cdots, a_p \in E$ et pour tous $t_1, \cdots, t_p \in K$ tels que $t_1 + \cdots + t_p = 1$, on a :

$$g(t_1\, a_1 + \cdots + t_p\, a_p) = t_1\, g(a_1) + \cdots + t_p\, g(a_p)$$

Soit $c = g(0)$; d'après la remarque (*) page 382, g sera affine si et seulement si l'application f définie par $f(x) := g(x) - c$ est linéaire, c'est-à-dire si :

a) $g(\lambda\, x) = \lambda g(x) + (1 - \lambda)\, c$

b) $g(x + y\,) = g(x) + g(y\,) - c$

Or $\lambda\, x$ peut être écrit sous forme de somme barycentrique : $\lambda\, x = \lambda\, x - (1 - \lambda)\, 0$; donc

$$g(\,\lambda\, x) = g(\lambda\, x - (1 - \lambda)0) = \lambda\, g(x) - (1 - \lambda) g(0) = \lambda\, g(x) - (1 - \lambda)\, c$$

De même :

$$g(x + y) = g(-0 + x + y) = -g(0) + g(x) + g(y) = -c + g(x) + g(y). \qquad \square$$

Isométries d'un espace vectoriel euclidien

Soit E un espace euclidien. On définit une distance sur E en posant : $\mathrm{d}(a, b\,) = \|a - b\,\|$ (cf. exercice 10. chapitre 7).

Définition A.7.9 – *Soit E un espace euclidien. On appelle **isométrie** une application $g\ : E \to E$ (non nécessairement linéaire) telle que $\mathrm{d}\Big(g(a), g(b\,)\Big) = \mathrm{d}(a, b\,)$, $\forall a, b \in E$, c'est-à-dire telle que :*

$$\|g(a) - g(b\,)\| = \|a - b\,\|, \quad \forall\, a,\, b \in E$$

Il est clair que toute transformation orthogonale g est une isométrie, car, g étant linéaire, on a :

$$\|g(a) - g(a\,)\| = \|g(a - b)\| = \|a - b\|$$

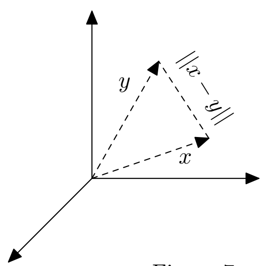


Figure 7

La réciproque cependant n'est pas vraie : les translations, par exemple, sont des isométries, mais elles ne sont par orthogonales, car elles ne sont pas linéaires. Nous allons voir que les isométries sont en fait les composées des transformations orthogonales par des translations.

Théorème A.7.10 – *Toute isométrie d'un espace vectoriel euclidien est nécessairement affine et plus précisément elle est du type :*

$$g(x) = f(x) + a \qquad \text{avec } a \in E \text{ et } f \in O(E)$$

En particulier, l'ensemble $\mathrm{Is}(E)$ des isométries de E est un sous-groupe du groupe affine $\mathrm{GA}(E)$.

Démonstration : Montrons d'abord le lemme suivant :

LEMME - *Soit f une isométrie qui laisse fixe* 0*. Alors* $f \in \mathrm{O}(E)$.

En effet, soit $f \in \mathrm{Is}(E)$ telle que $f(0) = 0$. On a :

$$\|f(a)\| = \|f(a) - 0\| = \|f(a) - f(0)\| = \|a - 0\| = \|a\|$$

D'autre part, de l'identité $\|x+y\|^2 + \|x-y\|^2 = 2\|x\|^2 + 2\|y\|^2$, on a :

$$\begin{aligned} \|f(a) + f(b)\|^2 &= -\|f(a) - f(b)\|^2 + 2\left(\|f(a)\|^2 + \|f(b)\|^2\right) \\ &= -\|a - b\|^2 + 2\left(\|a\|^2 + \|b\|^2\right) = \|a + b\|^2; \end{aligned}$$

d'où :

$$\begin{aligned} 2\,\langle f(a), f(b)\rangle &= \|f(a) + f(b)\|^2 - \|f(a)\|^2 - \|f(b)\|^2 \\ &= \|a + b\|^2 - \|a\|^2 - \|b\|^2 = 2\,\langle a, b\rangle \end{aligned} .$$

f conserve donc le produit scalaire. Il ne reste plus à montrer que f est linéaire. Soit $\{e_i\}$ une base orthonormée. Alors les $f(e_i)$, $i = 1, \ldots, n$, sont linéairement indépendants, car

$$\sum_{i=1}^n \lambda_i\, f(e_i) = 0 \implies \quad \begin{aligned} 0 \ &= \left\langle \sum_{i=1}^n \lambda_i\, f(e_i),\ f(e_j)\right\rangle \\ &= \sum_{i=1}^n \lambda_i\ \langle e_i, e_j\rangle \ = \ \lambda_j\ \langle e_j, e_j\rangle \ = \ \lambda_j \end{aligned}$$

Les $f(e_i)$ forment donc une base orthonormée. Enfin :

$$\begin{aligned} &\langle f(\lambda\, a + \mu\, b) - \lambda\, f(a) - \mu\, f(b),\ f(e_i)\rangle = \\ &\qquad = \langle f(\lambda\, a + \mu\, b), f(e_i)\rangle - \lambda\ \langle f(a), f(e_i)\rangle - \mu\ \langle f(b), f(e_i)\rangle \\ &\qquad = \ \langle \lambda\, a + \mu\, b, e_i\rangle - \lambda\ \langle a, e_i\rangle - \mu\ \langle b, e_i\rangle \ = \ 0 \end{aligned}$$

Comme les $f(e_i)$ forment une base, on en déduit que : $f(\lambda\, a + \mu\, b) = \lambda\, f(a) + \mu\, f(b)$. $\diamondsuit$

Le théorème s'en déduit maintenant facilement. En effet, soit $g \in \mathrm{Is}(E)$; on a :

$$t_{-g(0)} \circ g\,(0) = g(0) - g(0) = 0.$$

Donc $f := t_{-g(0)} \circ g$ est une isométrie qui laisse fixe 0 et, par conséquent $f \in \mathrm{O}(E)$. Il s'ensuit que : $g = t_a \circ f$, avec $a = g(0)$, c'est-à-dire : . $g(x) = f(x) + a$. $\square$

II. Notion abstraite d'espace affine

Espace affine

Il est important, aussi bien pour des raisons théoriques que pratiques, que de pouvoir disposer d'une notion d'espace affine en elle-même et non pas de le voir comme un sous-ensemble d'un espace vectoriel [2].

Pour comprendre les axiomes qui sont à la base de la notion d'espace affine, partons de la considération suivante. Soit $\mathcal{A}$ un sous-espace affine de direction F, F étant un sous-espace vectoriel de E. Comme on l'a vu, $\mathcal{A}$, en général, n'est pas stable pour la somme et donc la somme ne peut servir pour une définition axiomatique de $\mathcal{A}$. On a cependant une autre loi de composition sur $\mathcal{A}$, la loi :

$$\begin{array}{ccc} F \times \mathcal{A} & \longrightarrow & \mathcal{A} \\ (\vec{v}\,,\ a) & \longmapsto & b = t_{\vec{v}}\, a \equiv \vec{v} + a \end{array}$$

(en effet, si $\mathcal{A}$ est un sous-espace affine de direction F, alors pour tout $a \in \mathcal{A}$: $\mathcal{A} = a + F$; donc si $\vec{v} \in F$, on a bien : $b = a + \vec{v} \in \mathcal{A}$). Cette loi vérifie les propriétés suivantes :

1. $t_0 a = a$.
2. Pour tous $a, b \in \mathcal{A}$ il existe une unique translation $t_{\vec{v}}$ qui amène a sur b (il suffit de prendre $\vec{v} = b - a$).
3. $t_{\vec{v}} \circ (t_{\vec{w}}\, a) = (t_{\vec{v}} \circ t_{\vec{w}})\, a\,,\ \forall a \in \mathcal{A}\,,\ \forall\, \vec{v}, \vec{w} \in F.$

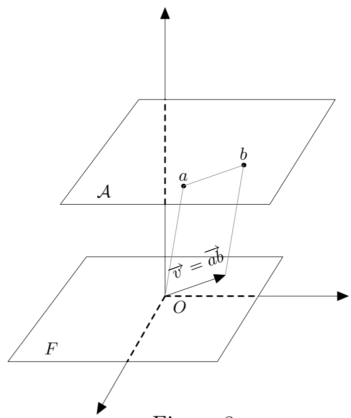


Figure 8

Avec la terminologie introduite dans l'Appendice A.1 (cf. page 344), on peut exprimer les propriétés 1. 2. en disant que $(F, +)$ ***agit transitivement et librement sur*** $\mathcal{A}$.

Remarque. – Puisque d'après 1. le vecteur de translation $\vec{v}$ qui amène a sur b est *unique*, $\vec{v}$ ne dépend que de a et b et il est noté $\overrightarrow{ab}$. On a donc :

$$1. \qquad \boxed{b = a + \overrightarrow{ab}}$$

Avec cette notation, la propriété 2 : $(a + \overrightarrow{v}) + \overrightarrow{w} = a + (\overrightarrow{v} + \overrightarrow{w})$, peut s'exprimer de la manière suivante. Notons que si $a + \overrightarrow{v} = b$, alors $\overrightarrow{v} = \overrightarrow{ab}$. On a donc :

$$(a + \overrightarrow{v}) + \overrightarrow{w} = (a + \overrightarrow{ab}) + \overrightarrow{w} = b + \overrightarrow{w}.$$

Si on pose $c = b + \overrightarrow{w}$, on a : $\overrightarrow{w} = \overrightarrow{bc}$ et donc

$$(a + \overrightarrow{v}) + \overrightarrow{w} = b + \overrightarrow{bc} = c. \qquad \text{(i)}$$

D'autre part :

$$a + (\overrightarrow{v} + \overrightarrow{w}) = a + (\overrightarrow{ab} + \overrightarrow{bc}). \qquad \text{(ii)}$$

[2] De même qu'il est intéressant d'avoir une notion "abstraite" d'espace vectoriel, même si par le choix d'une base on peut l'identifier à $\mathbb{R}^n$ ou, plus généralement, à K^n.

En comparant (i) et (ii), on voit que la propiété 2. est satisfaite si et seulement si

$$c = a + (\overrightarrow{ab} + \overrightarrow{bc}),$$

ce qui, d'après la propriété d'unicité 1. signifie que

$$2. \qquad \boxed{\overrightarrow{ac} = \overrightarrow{ab} + \overrightarrow{bc}}$$

Cette identité est dite ***relation de Chasles.***

Ceci nous amène à la définition suivante :

Définition A.7.11 – *Soit $\mathcal{A}$ un ensemble et E un espace vectoriel. On dit que $\mathcal{A}$ est muni d'une structure d'**espace affine** de direction E, si $(E, +)$ agit transitivement et librement sur A. Si E est de dimension n on dit que $\mathcal{A}$ est de dimension n.*

En d'autres termes, $\mathcal{A}$ est un espace affine de direction E si l'on se donne une application

$$t: \begin{array}{ccc} E \times \mathcal{A} & \longrightarrow & \mathcal{A} \\ (\vec{v}\,,\ P) & \longmapsto & t_{\vec{v}}(P) \underset{not}{=} P + \vec{v} \end{array}$$

vérifiant les propriétés suivantes :

1. $P + \overrightarrow{0} = P$

   [On notera $\vec{v}, \vec{0}, etc.$ les éléments de E pour mieux distinguer les points (de $\mathcal{A}$) et les vecteurs de E.]

2. Pour tous $P\,, Q \in \mathcal{A}$, il existe un ***unique*** vecteur $\vec{v} \in E$, tel que $Q = P + \vec{v}$. Le vecteur $\vec{v}$ sera noté $\overrightarrow{PQ}$, et l'on a donc :

$$Q = P + \overrightarrow{PQ} \tag{1}$$

3. Pour tous $P \in \mathcal{A}$ et $\vec{v}\,, \vec{w} \in E$, on a : $(P + \vec{v}) + \vec{w} = P + (\vec{v} + \vec{w})$, ce qui équivaut à la relation

$$\overrightarrow{PQ} = \overrightarrow{PR} + \overrightarrow{RQ} \tag{2}$$

L'équivalence de l'axiome 2. avec la relation de Chasles se démontre exactement comme ci-dessus.

L'application $t_{\vec{v}} : \mathcal{A} \longrightarrow \mathcal{A}$ est dite ***translation*** associée $\vec{v}$.

Notons que d'après l'axiome 1., le choix d'un point $P \in \mathcal{A}$ induit une bijection $V_P : \underset{Q\ \longmapsto\ \overrightarrow{PQ}}{\mathcal{A} \longrightarrow E}$.
On peut donc définir la notion abstraite d'espace affine aussi de la manière suivante :

Définition A.7.12 – *Soit $\mathcal{A}$ un ensemble et E un espace vectoriel. On dit que $\mathcal{A}$ est un espace affine de direction E si l'on a une application*

$$V : \underset{(P,Q)\ \longmapsto\ \overrightarrow{PQ}}{\mathcal{A} \times \mathcal{A} \longrightarrow E}$$

telle que :

1. *Pour tout $P \in \mathcal{A}$, l'application $V_P : \underset{Q\ \longmapsto\ \overrightarrow{PQ}}{\mathcal{A} \longrightarrow E}$ est bijective ;*
2. $\overrightarrow{PQ} = \overrightarrow{PR} + \overrightarrow{RQ}$.

Remarquons que, puisque $Q = V_P^{-1}(\overrightarrow{PQ})$, l'action de $(E, +)$ sur $\mathcal{A}$ est définie par : $P + \vec{v} := V_P^{-1}(\vec{v})$.

Les propriétés suivantes sont conséquence immédiate de la relation de Chasles :

i) $\overrightarrow{PP} = 0$ (qui s'obtient en faisant $P{=}Q{=}R$)
ii) $\overrightarrow{PQ} = -\overrightarrow{QP}$ (qui s'obtient en faisant $Q{=}P$ et en tenant compte de i))

Exemples

1. L'exemple le plus important est celui que l'on a vu en 1.1 : on définit une structure d'espace affine sur un espace vectoriel E en posant $\overrightarrow{ab} := b - a$. Cette structure sera dite ***standard*** (ou *canonique*).
2. Plus généralement, soit $\mathcal{A}$ un ensemble et $f : \mathcal{A} \longrightarrow E$ une bijection sur un espace vectoriel E. On définit une structure d'espace affine sur $\mathcal{A}$ par : $\overrightarrow{ab} := f(b) - f(a)$.

Remarque. – Dans le cas d'un sous-espace affine d'un espace vectoriel, si $b = a + \overrightarrow{ab}$ on peut écrire : $\overrightarrow{ab} = b - a$. C'est la raison pour laquelle, en plus de la notation $Q = P + \overrightarrow{PQ}$, on utilise parfois la notation $\overrightarrow{PQ} = Q - P$ pour un espace affine quelconque. Cependant cette manière d'écrire peut prêter à confusion : on ne peut pas "sommer" a priori deux éléments de $\mathcal{A}$: les éléments de $\mathcal{A}$ ne peuvent être "sommés" qu'avec les éléments de F.

Applications affines

Comme nous l'avons vu (cf. définition A.7.2), dans le cadre des sous-variétés affines d'un espace vectoriel E, une application $g : E \longrightarrow E$ est dite affine si elle est du type $g(x) = a + f(x)$ avec $a \in E$ et f linéaire. Notons qu'alors $a = g(0)$; aussi les applications affines sont les applications du type :

$$g(x) = g(0) + f(x)\,, \qquad \text{avec } f \text{ linéaire.}$$

Définition A.7.13 – *Soient $\mathcal{A}$ et $\mathcal{A}'$ deux espaces affines de directions E et E'. Une application $g : \mathcal{A} \longrightarrow \mathcal{A}'$ est dite* ***affine*** *s'il existe un point O tel que g s'écrive sous la forme :*

$$g(P) = g(O) + f(\overrightarrow{OP})\,, \quad \text{avec } f \in \mathcal{L}_K(E, E')\ ;$$

On vérifie facilement que si cette propriété est satisfaite, alors elle est satisfaite pour tout $O \in \mathcal{A}$. f est dite partie linéaire *de g. La partie linéaire de g est notée habituellement $\vec{g}$.*

Avec les notations ci-dessus (cf. (1) page 387) on peut dire aussi que g est affine si et seulement si l'application $\vec{g} : E \longrightarrow E$ définie par

$$\vec{g}(\overrightarrow{PQ}) := \overrightarrow{g(P)g(Q)}'$$

est linéaire.

L'ensemble des applications affines bijectives de $\mathcal{A}$ dans $\mathcal{A}$ forme un groupe, dit *groupe affine* de $\mathcal{A}$, noté $\mathrm{GA}(\mathcal{A})$.

Deux espaces affines $\mathcal{A}$ et $\mathcal{A}'$ sont dits isomorphes s'il existe une application affine bijective $g : \mathcal{A} \longrightarrow \mathcal{A}'$. Ceci veut dire que $\overrightarrow{g}$ est isomorphisme entre les directions de $\mathcal{A}$ et $\mathcal{A}'$.

Remarquons que si on note

$$V_P(Q) = \overrightarrow{PQ} \qquad \text{et} \qquad V'_{P'}(Q') = \overrightarrow{P'Q'}'$$

les bijections qui définissent les structures affines de $\mathcal{A}$ et $\mathcal{A}'$, la linéarité de $\vec{g}$, c'est-à-dire la condition $\vec{g}(\overrightarrow{PQ}) = \overrightarrow{g(P), g(Q)}'$, peut s'écrire

$$\vec{g}\,V_P(Q) = V'_{g(P)}g(Q)$$

et traduit donc le fait que g échange les structures affines de $\mathcal{A}$ et $\mathcal{A}'$

Vectorialisé

Comme on l'a vu, on ne peut pas définir une somme d'éléments de $\mathcal{A}$. Cependant, le choix d'un point dans $\mathcal{A}$ permet de définir sur $\mathcal{A}$ une structure d'espace vectoriel. En effet, d'après l'axiome 1., le choix d'un point $O \in \mathcal{A}$ induit une bijection

$$V_O : \underset{P}{\mathcal{A}} \underset{\longmapsto}{\longrightarrow} \underset{\overrightarrow{OP}}{E}$$

ce qui permet de transporter sur $\mathcal{A}$ la structure d'espace vectoriel de F. On a (cf. exercice 4 du chapitre 3) :

$$\begin{aligned} P \dot{+} Q \ &:= V_O^{-1}\Big(V_O(P) + V_O(Q)\Big) = V_O^{-1}(\overrightarrow{OP} + \overrightarrow{OQ}) = O + \overrightarrow{OP} + \overrightarrow{OQ} \\ \lambda \cdot P \ &:= V_O^{-1}\Big(\lambda\, V_O(P)\Big) = V_O^{-1}(\lambda\,\overrightarrow{OP}) = O + \lambda\,\overrightarrow{OP} \end{aligned}$$

c'est-à-dire, (cf. page 381) :

$$\begin{aligned} P \dot{+} Q &:= O + \overrightarrow{OP} + \overrightarrow{OQ} && \text{(a)} \\ \lambda \cdot P &:= O + \lambda\,\overrightarrow{OP} && \text{(b)} \end{aligned}$$

L'espace vectoriel ainsi obtenu est dit ***vectorialisé*** de $\mathcal{A}$ en O et il est noté $\overrightarrow{\mathcal{A}_O}$.

Notons que $P \dot{+} O = O + \overrightarrow{OP} + \overrightarrow{OO} = O + \overrightarrow{OP} = P$, donc O est le neutre.

Proposition A.7.14 – *Soit $\mathcal{A}$ un espace affine et $O \in \mathcal{A}$. Alors $\mathcal{A}$ est isomorphe à l'espace affine $\overrightarrow{\mathcal{A}_O}$ muni de la structure standard.*

Il suffit de considérer l'application $g_O : \overrightarrow{\mathcal{A}_O} \longrightarrow \mathcal{A}$ définie par $g_O(P) = O + V_O(P)$[3]. Comme V_O est un isomorphisme entre $\mathcal{A}_O$ et E, V_O est un isomorphisme d'espaces affines. Plus précisément on a : $V_O(Q \dot{-} P) = V_O(Q) - V_O(P) = \overrightarrow{OQ} - \overrightarrow{OP} = \overrightarrow{PQ}$, c'est-à-dire :

$$\overrightarrow{PQ} = V_O\,(Q \dot{-} P) \qquad (*)$$

Nota – Habituellement la loi $\dot{+}$ est notée plus simplement $+$. Par ailleurs, puisque V_O est un isomorphisme, on identifie $V_O(Q \dot{-} P)$ avec $Q - P$ si bien que la relation ci-dessus est écrite parfois (par abus de notation, car l'identification dépend du choix d'un point) : $\overrightarrow{PQ} = Q - P$.

Ceci veut dire que *modulo le choix d'un point on peut identifier l'espace affine $\mathcal{A}$ avec l'espace vectoriel $\overrightarrow{\mathcal{A}_O}$ muni de la structure affine standard et donc avec E muni de la structure standard* [4]. Ce qui veut dire aussi que *l'exemple standard est le seul exemple d'espace affine* (à un isomorphisme près) [5]. En particulier, *les propriétés d'un espace affine peuvent être démontrées dans le cas standard,* comme dans la section A.3.1*, sans perte de généralité.*

[3] Noter qu'en fait, au point de vue ensembliste, $\mathcal{A} = \overrightarrow{\mathcal{A}_O}$ et V_O est l'identité.

[4] De même que, modulo le choix d'une base, on identifie les vecteurs d'un espace vectoriel au n-uplet de ses composantes, ou les endomorphismes aux matrices.

[5] De même que K^n est le seul exemple d'espace vectoriel de dimension n sur K, à un isomorphisme près.

Sous-espaces affines. Barycentre

Définition A.7.15 – *Soient $P_1, \cdots, P_q$ des points de $\mathcal{A}$ et $t_1, \cdots, t_q$ des scalaires tels que $t_1 + \cdots + t_q = t \neq 0$. Soit O un point quelconque de $\mathcal{A}$. Alors le point*

$$G = O + \frac{1}{t}\Big(t_1\,\overrightarrow{O\,P_1} + \cdots + t_q\,\overrightarrow{O\,P_q}\Big)$$

est indépendant [6] *du choix de O et il est dit* ***barycentre*** *des $(P_i, t_i)_{i=1,\ldots,q}$.*

En utilisant la structure vectorielle du véctorialisé en O, on peut écrire (cf. formules (a) et (b) page 389) : $G = \frac{1}{t}\Big(t_1\,P_1 \dot{+} \cdots \dot{+} t_q\,P_q\Big)$. En fait, puisque G ne dépend pas du choix de O, on peut écrire

$$G = \frac{1}{t}\Big(t_1\,P_1 + \cdots + t_q\,P_q\Big)$$

cette somme est dite dite *somme barycentrique* [7].

Définition A.7.16 – *Soit $\mathcal{A}$ un espace affine de direction E. Un sous-ensemble $\mathcal{P} \subset \mathcal{A}$ est dit* ***sous-espace affine*** *de $\mathcal{A}$ s'il existe un sous-espace F de E tel que $\mathcal{P}$ soit un espace affine de direction F. On appelle dimension de $\mathcal{P}$ la dimension de F.*

Cela veut dire que si P_0 est un point de $\mathcal{P}$, alors

$$\mathcal{P} = \Big\{ P \in \mathcal{A} \,\Big|\, P = P_0 + \overrightarrow{P_0P}, \ \text{ avec } \overrightarrow{P_0P} \in F \Big\}$$

On notera aussi

$$\mathcal{P} = P_0 + F\,, \quad \text{ avec } \ P_0 \in \mathcal{P}.$$

Exemples– Soient $P, Q \in \mathcal{A}$, $P \neq Q$, et

$$\mathcal{D} = \Big\{ M \in \mathcal{A} \,\Big|\, M = P + \overrightarrow{PQ}, \quad t \in K \Big\}$$

est un sous-espace affine de $\mathcal{A}$ de direction Vect$\{\overrightarrow{P\,Q}\}$, dit ***droite affine*** engendrée par P et Q (cf. figure 9)

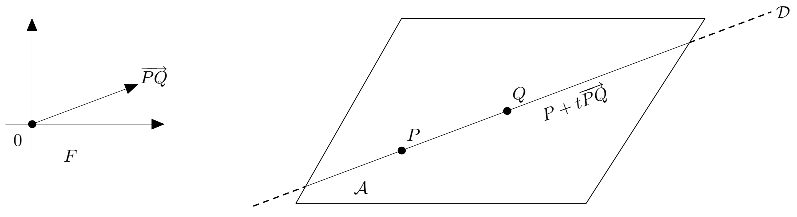


Figure 9

De même, étant donnés trois points P, Q, R qui ne sont pas sur une même droite affine, l'ensemble des points $P + s\,\overrightarrow{P\,Q} + t\,\overrightarrow{P\,R}$ est un sous-espace affine de direction Vect$\{\overrightarrow{P\,Q}, \overrightarrow{P\,R}\}$, dit ***plan affine*** défini par les points P, Q, R.

[6] Conséquence facile de la relation de Chasles.

[7] Comme on l'a dit, dans un espace affine il n'a pas de sens de "sommer" des points, sinon en vectorialisant en un point et avec les abus de notations signalés (la somme dépendra du choix du point). En revanche la somme barycentrique a un sens précis, car elle ne dépend pas du choix du point.

Plus généralement, étant donnés k points $P_1, ..., P_k$, on appelle sous-espace affine défini par ces points l'ensemble

$$\mathrm{Aff}\{P_1, ..., P_k\} := \Big\{ P \in \mathcal{A} \mid P = P_1 + t_2\overrightarrow{P_1P_2} + \cdots + t_k\overrightarrow{P_1P_k}, \text{ avec } t_1, ..., t_k \in K \Big\}.$$

Il est clair que la direction de ce sous-espace est $\mathrm{Vect}\{\overrightarrow{P_1P_2}, ..., \overrightarrow{P_1P_k}\}$.
Notons que la somme qui apparaît dans la définition de $\mathrm{Aff}\{P_1, ..., P_k\}$ peut être écrite comme somme baycentrique, car on peut l'écrire sous la forme :

$$P = \Big(1 - (t_2 + \cdot + t_k)\Big)\overrightarrow{P_1P_1} + t_2\overrightarrow{P_1P_2} + \cdots + t_k\overrightarrow{P_1P_k}.$$

Aussi peut-on écrire :

$$\mathrm{Aff}\{P_1, ..., P_k\} := \Big\{ P \in \mathcal{A} \mid P = t_1P_1 + t_2P_2 + \cdots + t_kP_k, \text{ avec } t_1 + \cdots + t_k = 1 \Big\}.$$

Le théorème A.7.7 se généralise immédiatement :

Proposition A.7.17 – *Une une partie non vide $\mathcal{P}$ d'un espace affine $\mathcal{A}$ est un sous-espace affine de $\mathcal{A}$ si et seulement si elle est stable par composition barycentrique.*

Il suffit de fixer un point $P \in \mathcal{P}$ et de considérer le vectorialisé $\overrightarrow{\mathcal{A}_P}$: on est alors ramené à la démonstration de la proposition A.7.7. $\square$

Comme on l'a vu dans l'exemple ci-dessus, un sous-espace affine de dimension k est déterminé par $k+1$ points qui ne sont pas contenus dans un sous-espace affine de dimension $(k-1)$. Ceci nous amène à la définition de repère affine.

Définition A.7.18 *Soit $\mathcal{A}$ un espace affine de dimension n. On appelle* **repère affine** *la donnée de $n+1$ points de $\mathcal{A}$ non contenus dans un sous-espace affine de dimension $n-1$.*

Proposition A.7.19 *Soit $\{P_0, P_1, ..., P_n\}$ un repère affine de $\mathcal{A}$. Alors tout point $P \in \mathcal{A}$ s'écrit d'une manière unique sous la forme*

$$P = \sum_{i=0}^{n} \alpha_i\, P_i\,, \qquad \textit{avec : } \alpha_0 + \alpha_1 + \cdots + \alpha_n = 1$$

Les α_i sont dites coordonnées barycentriques de P.

Démonstration : Choisissons un point du repère affine, par exemple P_0 et considérons le vectorialisé $\overrightarrow{\mathcal{A}_{P_0}}$ de $\mathcal{A}$ en P_0. Les vecteurs $\vec{v_1} = \overrightarrow{P_0\,P_1}$, ..., $\vec{v_n} = \overrightarrow{P_0\,P_n}$ forment une base de $\overrightarrow{\mathcal{A}_{P_0}}$, il existe un *unique* n-uplet de scalaires $(t_1, ..., t_n)$ tels que

$$\overrightarrow{P_0\,P} = t_1\,\overrightarrow{P_0\,P_1} + \cdots + t_n\,\overrightarrow{P_0\,P_n}$$

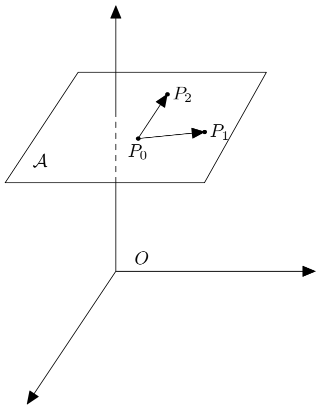


Figure 10

Or pour tout $P \in \mathcal{A}$ on peut écrire :

$$P = P_0 + \overrightarrow{P_0\,P} = P_0 + t_1\,\overrightarrow{P_0\,P_1} + \cdots + t_n\,\overrightarrow{P_0\,P_n} = \alpha_0\,P_0 + \alpha_1\,P_1 + \cdots + \alpha_n\,P_n$$

avec $\alpha_0 = 1 - (t_1 + \cdots + t_n)$, $\alpha_1 = t_1$, $\ldots, \alpha_n = t_n$ et les α_i sont uniques, puisque le n-uplet $(t_1, ..., t_n)$ est unique). $\square$

Toutes les propriétés que nous avons vues dans la section 1. (cas de standard) se transportent immédiatement dans les espaces affines abstraits, car tout espace affine est isomorphe, modulo le choix d'un point, à sa direction munie de la structure standard. En particulier :

Proposition A.7.20 – *Une application $g : \mathcal{A} \longrightarrow \mathcal{A}'$ est affine si et seulement si elle conserve le barycentre, c'est-à-dire si*

$$g\left(\sum_{i=1}^{r} \alpha_i\, P_i\right) = \sum_{i=1}^{r} \alpha_i\, g(P_i)$$

$\forall P_1, \ldots, P_r$ et $\forall \alpha_1, \ldots, \alpha_r \in K,$ telsque $\alpha_1 + \cdots + \alpha_r = 1.$

Corollaire A.7.21 *Deux applications affines qui coïncident sur un repère affine sont égales en tout point de $\mathcal{A}$.*

En effet, soit $\{P_1, ..., P_n\}$ un repère affine et f, g deux applications affines telles que $f(P_i) = g(P_i)$, pour $i = 1, ..., n$ et soit $P = \sum_{i=1}^{n} \alpha_i\, P_i \in \mathcal{A}$ (avec $\alpha_1 + \cdots + \alpha_n = 1$). Puisque f et g conservent le barycentre, on a :

$$f(P) = f\left(\sum_{i=1}^{n} \alpha_i\, P_i\right) = \sum_{i=1}^{n} \alpha_i\, f(P_i) = \sum_{i=1}^{n} \alpha_i\, g(P_i) = g(P). \quad \square$$

Isométries

Définition A.7.22 *On appelle **espace affine euclidien** un espace affine dont la direction est un espace euclidien.*

Dans la suite, on notera $\mathcal{E}$ un espace affine euclidien et E sa direction.

Sur un espace affine euclidien on définit une distance, on posant :

$$\mathrm{d}(P, Q) = \| \overrightarrow{PQ} \|$$

Une application f entre espaces affines euclidiens $\mathcal{E}$ et $\mathcal{E}'$ est dite ***isométrie*** si elle conserve la distance, c'est-à-dire si

$$\mathrm{d}(f(P), f(Q)) = \mathrm{d}(P, Q)\,, \quad \forall\ P, Q \in \mathcal{E}$$

On vérifie immédiatement que les translations sont des isométries.

Les isométries de $\mathcal{E}$ forment un groupe noté $\mathrm{Is}(\mathcal{E})$. Comme pour le théorème A.7.10, on a :

Théorème A.7.23 *Soit $\mathcal{E}$ un espace affine euclidien de direction E. Alors une application $f : \mathcal{E} \longrightarrow \mathcal{E}$ est une isométrie si et seulement si elle est affine et sa partie linéaire $\vec{f}$ est orthogonale, c'est-à-dire si f est de la forme :*

$$f(P) = f(O) + \vec{f}(\overrightarrow{OP})\,, \qquad \text{avec } \vec{f} \in O(E)$$

O étant un point quelconque de $\mathcal{E}$.

La démonstration est immédiate : en choisissant un point O de $\mathcal{E}$, $\mathcal{E}$ est isomorphe à E muni de la structure standard ; on transporte alors par isomorphisme le théorème A.7.10.

Plus précisément, supposons qu'il existe un point Ω fixe pour f. En prenant $O = \Omega$, on a : $f(P) = \Omega + \vec{f}(\overrightarrow{\Omega P})$. En vectorialisant en Ω, c'est-à-dire en posant $V_\Omega(P) = \overrightarrow{\Omega P}$, on a : $V_\Omega(f(P)) = \vec{f}(V_\Omega(P))$, donc

$$f = V_\Omega^{-1} \circ \vec{f} \circ V_\Omega$$

C'est-à-dire : l'étude de f se ramène à celui de l'application linéaire $\vec{f}$. En notant $\mathrm{Is}_\Omega(\mathcal{E})$ le sous-groupe des isométries de $\mathrm{Is}(\mathcal{E})$ qui laissent fixe Ω, on a :

$$\mathrm{Is}_\Omega(\mathcal{E}) = V_\Omega^{-1} \circ \mathrm{O}(E) \circ V_\Omega$$

Supposons maintenant qu'il n'y a pas de point fixe. Soit Ω un point *quelconque* ; alors $f(\Omega) \neq \Omega$, donc le vecteur $\vec{v} = \overrightarrow{f(\Omega)\Omega}$ est non nul. Considérons l'application $g = t_{\vec{v}} \circ f$. On a :

$$g(\Omega) = t_{\vec{v}}(f(\Omega)) = f(\Omega) + \vec{v} = f(\Omega) + \overrightarrow{f(\Omega)\Omega} = \Omega.$$

Donc g est une application affine qui admet Ω comme point fixe et l'on peut écrire :

$$f = t_{-\vec{v}} \circ g$$

Comme g admet un point fixe, l'étude de g est ramenée à celle de sa partie linéaire qui est une transformation orthogonale. On a ainsi :

(*)

Le groupe $\mathrm{Is}_\Omega(\mathcal{E})$ *des isométries qui laissent fixe un point* Ω *est isomorphe au groupe orthogonal de la direction* E *de* $\mathcal{E}$ *:*

$$\mathrm{Is}_\Omega(\mathcal{E}) = V_\Omega^{-1} \circ \mathrm{O}(E) \circ V_\Omega.$$

Toute isométrie est la composée d'une translation par une isométrie qui admet un point fixe. Elle peut donc être "identifiée" au produit d'une translation par une application orthogonale.

Appendice A.8

Sur les isométries dans le plan et dans l'espace

Dans cet appendice, on notera $\mathrm{Is}(n)$ le groupe des isométries d'un espace affine euclidien de dimension n.

Les isométries du plan

Soit $\mathcal{E}$ un espace affine euclidien de dimension 2. On notera E sa direction. Conformément au résultat (*) page 393 et à la Proposition 7.25, Is(2) est formé par

- l'identité $\mathrm{id}_{\mathcal{E}}$
- les rotations autour d'un point Ω, d'angle $\theta \neq 0 \bmod[2\pi]$ (c'est-à-dire différentes de $\mathrm{id}_{\mathcal{E}}$), que l'on notera $R_{\Omega,\,\theta}$
- les réflexion par rapport à une droite $\mathcal{D}$, que l'on notera $s_{\mathcal{D}}$.
- les translations $t_{\vec{v}}$

et les composées de ces applications.

Notons que

- les rotations différentes de $\mathrm{id}_{\mathcal{E}}$ sont caractérisées par avoir un seul point fixe ;
- les réflexions par le fait d'avoir une droite fixe.

Nous allons maintenant étudier les composées de ces isométries avec une translation. Pour cela on a besoin du Lemme suivant, qui est valable en dimension quelconque (finie) :

Lemme A.8.1 – *Soit $f : \mathcal{E} \longrightarrow \mathcal{E}$ une application affine et notons :* $\mathrm{Fix}(f) = \{\,\text{points fixes de} f\,\}$.

- *Si 1 n'est pas valeur propre de $\vec{f}$, alors f admet un unique point fixe.*
- *Si 1 est valeur propre de $\vec{f}$ alors : soit $\mathrm{Fix}(f) = \emptyset$, soit il existe un point fixe Ω et*

$$\mathrm{Fix}(f) = \Omega + E_1^{\vec{f}}$$

c'est-à-dire l'ensemble des points fixes, soit il est vide, soit il est un sous-espace affine dont la direction est l'espace propre correspondant à la valeur propre 1 de $\vec{f}$, $E_1^{\vec{f}}$

La démonstration est très simple si on vectorialise en un point O et on identifie $\mathcal{E}$ avec son vectorialisé muni de la structure standard, c'est-à-dire si on identifie les points P avec les vecteurs $\overrightarrow{OP}$ (cf. Nota page 389). Soit $f(P) = \vec{f}(P) + \vec{v}$, on a :

$$f(P) = P \Longleftrightarrow \vec{f}(P) = P + \vec{v} \Longleftrightarrow (\vec{f} - \vec{\mathrm{id}})(P) = -\vec{v}$$

- si dét $(\vec{f} - \vec{\mathrm{id}}) \neq 0$, il existe une et une seule solution qui s'obtient en résolvant le système :

$$\boxed{(\vec{f} - \vec{\mathrm{id}})(P) = -\vec{v}} \qquad (1)$$

- si dét $(\vec{f} - \vec{\mathrm{id}}) = 0$, supposons $\mathrm{Fix}\{f\} \neq \emptyset$ et soit Ω un point fixe. On a alors $(\vec{f} - \vec{\mathrm{id}})(\Omega) = -\vec{v}$. Si P est un autre point fixe, on aussi $(\vec{f} - \vec{\mathrm{id}})(P) = -\vec{v}$, d'où par soustraction

$$(\vec{f} - \vec{\mathrm{id}})(P - \Omega) = 0$$

ce qui veut dire que $P - \Omega \in E_1^{\vec{f}}$, c'est-à-dire : $P \in \Omega + E_1^{\vec{f}}$. □

Remarque. – A titre d'exemple voici la démonstration si on ne vectorialise pas :

Supposons que 1 n'es pas valeur propre, c'est-à-dire que$E_1^{\vec{f}} = \{\vec{0}\}$, ou : $\mathrm{Ker}(\vec{f} - \mathrm{id}_E) = \{\vec{0}\}$, ce qui en dimension finie est équivalent à : $\vec{f} - \mathrm{id}_E$ est bijective. Prenons un point O arbitraire et soit Ω un point quelconque. On a

$$\vec{f}\,(\overrightarrow{O\Omega}) = \overrightarrow{f(O)\,f(\Omega)} = \overrightarrow{f(O)O} + \overrightarrow{O\Omega} + \overrightarrow{\Omega f(\Omega)} = \overrightarrow{f(O)\,\Omega} = \overrightarrow{f(O)\,O} + \overrightarrow{O\Omega}$$

c'est-à-dire : $(\vec{f} - \mathrm{id}_E)(\overrightarrow{O\Omega}) = \overrightarrow{f(O)O} + \overrightarrow{\Omega f(\Omega)}$. Ainsi les points fixes Ω (c'est-à-dire qui vérifient $\overrightarrow{\Omega f(\Omega)} = \vec{0}$) sont les solutions du système : $(\vec{f} - \mathrm{id}_E)(\overrightarrow{O\Omega}) = \overrightarrow{f(O)O}$. Ce système a une solution unique (puisque $\vec{f} - \mathrm{id}_E$ est bijective), donnée par la formule $\overrightarrow{O\Omega} = (\vec{f} - \mathrm{id}_E)^{-1}\,\overrightarrow{f(O)\,O}$, O étant un point arbitraire.

Supposons que 1 est valeur propre et qu'il existe au moins un point fixe Ω. Alors on peut écrire $f(P) = f(\Omega) + \vec{f}(\overrightarrow{\Omega P}) = \Omega + \vec{f}(\overrightarrow{\Omega P})$. Donc P est fixe $\iff \vec{f}(\overrightarrow{\Omega P}) = \overrightarrow{\Omega P} \iff \overrightarrow{\Omega P} \in E_1^{\vec{f}} \iff P \in \Omega + E_1^{\vec{f}}$.

Comme on le voit, elle est laborieuse et au fond peu naturelle. Par conséquent dans la suite nous donnerons habituellement les démonstrations en vectorialisant en un point. Cependant le point de vue proprement affine peut aussi être plus clair dans certains cas, comme nous le verrons.

1er cas : $f = t_{\vec{v}} \circ R_{\Omega,\,\theta}$

Vectorialisons en Ω (c'est-à-dire prenons Ω comme origine). On a $\vec{f} = R_{\Omega,\,\theta}$ dont la matrice dans une b.o.n. est : $\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$. Cette matrice n'a pas 1 comme valeur propre (car $\theta \neq 0 \bmod [2\pi]$). Donc, d'après le Lemme, f *a un unique point fixe*, et par conséquent f est une *rotation* autour d'un point Ω' donné part la formule (1) ci-dessus : $\Omega' = \Omega - (R_{\Omega,\,\theta} - \mathrm{id}_E)^{-1}(\vec{v})$.

2ème cas : $f = t_{\vec{v}} \circ s_{\mathcal{D}}$

On a $\vec{f} = \vec{s}_{\mathcal{D}}$, donc 1 est valeur propre. Par conséquent, soit f n'a pas de points fixes, soit elle en a une infinité.

Pour chercher les points fixes il faut résoudre le système (1) : $(\vec{s}_{\mathcal{D}} - \mathrm{id})(P) = -\vec{v}$, En prenant l'origine O sur $\mathcal{D}$ et une b.o.n. (e_1, e_2) avec $e_1 \in \mathcal{D}$ [1], la matrice de $\vec{s}_{\mathcal{D}}$ s'écrit $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$.

Pour avoir les points fixes, il faut donc résoudre le système :

$$\begin{pmatrix} 0 & 0 \\ 0 & -2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = - \begin{pmatrix} a \\ b \end{pmatrix}$$

où a, b sont les coordonnées de $\vec{v}$. Ce système est compatible si et seulement si $a = 0$, c'est-à-dire $\vec{v} \perp \mathcal{D}$ et les solutions sont données par $y = b/2$, x arbitraire. L'ensemble des points fixes

[1] Plus précisément, $e_1 \in D$, D étant la direction de $\mathcal{D}$, mais justement en vectorialisant on peut identifier $\mathcal{D}$ et D.

est donc la droite $\mathcal{D}'$ parallèle à $\mathcal{D}$ et passant par le point $\frac{1}{2}\vec{v}$. Il s'agit donc d'une *réflexion* par rapport à cette droite.

Supposons maintenant que $\vec{v} \notin \mathcal{D}^{\perp}$. En posant $\vec{v} = \vec{v_1} + \lambda\vec{u}$, avec $\vec{v}_1 \perp \mathcal{D}$ et $\vec{u}//\mathcal{D}$, on a :

$$f = t_{\vec{v}} \circ s_{\mathcal{D}} = T_{\lambda\vec{u}} \circ \underbrace{T_{\vec{v_1}} \circ s_{\mathcal{D}}}_{\text{réflexion}}$$

f est donc la composée d'une réflexion par rapport à une droite parallèle à $\mathcal{D}$ avec une translation *dans la direction de* $\mathcal{D}$. Une telle application est dite ***glissement***. Notons que les glissements n'ont pas de points fixes, car $\vec{v}$ n'est pas orthogonal à $\mathcal{D}$.
On a ainsi :

Théorème A.8.2 - Is(2) *est composé de*

- *l'identité*
- *des rotations autour d'un point*
- *des réflexions par rapport à une droite*
- *des glissements*
- *et des translations.*

Les isométries de l'espace

Conformément au résultat (*) page 393 et à la Proposition 7.27, Is(3) est formé par

- l'identité $\mathrm{id}_{\mathcal{E}}$;
- les rotations $R_{\mathcal{D},\theta}$, d'angle $\theta \neq 0[\mathrm{mod} 2\pi[$, autour d'une droite affine $\mathcal{D}$, caractériseés par le fait d' avoir une droite fixe, l'axe de rotation ;
- les réflexions par rapport à un plan affine $s_{\mathcal{P}}$, caractérisées par le fait d'avoir un plan fixe.
- les "rotations-réflexions" $s_{\mathcal{P}} \circ R_{\mathcal{D},\theta}$, où $\mathcal{P}$ est un plan orthogonal à $\mathcal{D}$, caractérisées par le fait d'avoir un unique point fixe ;
- les translations $t_{\vec{v}}$;

et les composées de ces applications.

Nous allons étudier maintenant ces dernières.

1er cas : $f = t_{\vec{v}} \circ R_{\mathcal{D},\theta}$.

Notons, pour abréger, $R = R_{\mathcal{D},\theta}$ et cherchons les points fixes. On a $\vec{f} = R$. Puisque 1 est valeur propre de $\vec{f}$, alors f soit n'a pas de points fixes, soit en a une infinité. Pour chercher les points fixes, il faut résoudre l'équation $f(P) = P$ c'est-à-dire

$$R(P) - P = -\vec{v}. \qquad (*)$$

Soit $\vec{k}$ un vecteur unitaire directeur de $\mathcal{D}$; on a : $< R(P), \vec{k} > = < R(P), R(\vec{k}) > = < P, \vec{k} >$, puisque R est orthogonale. Donc une condition de compatibilité de l'équation (*) est que $< \vec{v}, \vec{k} > = 0$.

– *Supposons maintenant que* $\vec{v} \perp \vec{k}$ et prenons un repère affine orthonormé $(\Omega, \vec{v}/\|v\|, \vec{w}, \vec{k})$ avec $\Omega \in \mathcal{D}$. Cherchons les solutions de l'équation (*). Si (x_0, y_0, z_0) sont les coordonnées de P dans ce repère, on est amené à résoudre le système

$$\begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_0 \\ y_0 \\ z_0 \end{pmatrix} + \begin{pmatrix} \|v\| \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} x_0 \\ y_0 \\ z_0 \end{pmatrix}$$

ce qui donne

$$\begin{cases} (\cos\theta - 1)x_0 & - & \sin\theta\, y_0 & = -\|v\| \\ \sin\theta\, x_0 & + (\cos\theta - 1)y_0 & = & 0 \end{cases}$$

Ce système admet une solution unique (x_0, y_0) (puisque $\theta \neq 2k\pi$). Puisque z_0 est arbitraire, on donc une droite $\mathcal{D}'$ de points fixes (c'est une droite de direction $\vec{k}$). On a ainsi :

si $\vec{v} \perp \vec{k}$, f est une rotation autour d'un axe $\mathcal{D}'$ de direction $\vec{k}$, c'est-à-dire parallèle à $\mathcal{D}$.

–*Supposons que $\vec{v}$ n'est pas orthogonal à $\vec{k}$.* En posant $\vec{v} = \vec{v}_1 + \lambda\vec{k}$, avec $\vec{v}_1 \perp \vec{k}$, on a :

$$t_{\vec{v}} \circ R = t_{\lambda\vec{k}} \circ \underbrace{t_{\vec{v}_1} \circ R}_{\text{rotation}}$$

f est donc une rotation suivie d'une translation dans la direction de l'axe de rotation. Cette application est dite ***vissage.***

2ème cas : $f = t_{\vec{v}} \circ s_{\mathcal{P}}$

On a $\vec{f} = s_{\mathcal{P}}$, donc 1 est valeur propre de $\vec{f}$ et par conséquent, soit il n'y a pas de points fixes, soit il y en a une infinité. Les points fixes P doivent vérifier l'équation $f(P) - P = -\vec{v}$, Si $\vec{k}$ est un vecteur unitaire normal à $\mathcal{P}$, on voit, comme dans le cas 1. qu'une condition nécessaire pour que cette équation admette une solution est que $\vec{v}$ soit parallèle à $\vec{k}$. Un calcul analogue à celui effectué ci-dessus, montre que

si $\vec{v} \perp \vec{k}$, f est une *réflexion* par rapport à un plan $\mathcal{P}'$ parallèle à $\mathcal{P}$. [2]

Si $\vec{v}$ n'est pas parallèle à $\vec{k}$, on pose $\vec{v} = \lambda\vec{k} + \vec{v}_1$. On a

$$f = t_{\vec{v}} \circ s_{\mathcal{P}} = t_{\vec{v}_1} \circ \underbrace{t_{\lambda\vec{k}} \circ s_{\mathcal{P}}}_{\text{réflexion}}$$

f est donc *le produit d'une réflexion par une translation dans une direction parallèle au plan de réflexion.* Une telle application est dite ***glissement.***

3ème cas : $f = t_{\vec{v}} \circ (s_{\mathcal{P}} \circ R_{\mathcal{D}})$
où $s_{\mathcal{P}}$ est la réflexion par rapport à un plan affine $\mathcal{P}$ et et $R_{\mathcal{D}}$ la rotation autour d'un axe $\mathcal{D}$ orthogonal à $\mathcal{P}$.

On a $\vec{f} = \overrightarrow{s_{\mathcal{P}} \circ R_{\mathcal{D}}}$. Si Ω est l'intersection du plan de réflexion et de l'axe de rotation, la matrice de $\vec{f}$ dans un repère affine orthonormé $(\Omega, \vec{u}, \vec{w}, \vec{k})$, où $\vec{k}$ est un vecteur unitaire directeur de l'axe de rotation, est

$$\begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & -1 \end{pmatrix}, \quad \theta \neq 2k\pi$$

Puisque 1 n'est pas valeur propre de $\vec{f}$, f au un unique point fixe, le point Ω, et donc f est une rotation-reflexion.
On a ainsi :

Théorème A.8.3 - Is(3) *est composé de*

- *l'identité*
- *des rotations autour d'un axe*

[2] on trouve : $\mathcal{P}' = t_{\frac{1}{2}\vec{v}}\mathcal{P}$

- *des réflexions par rapport à un plan*
- *des rotations-reflexions*
- *des vissages*
- *des glissements*
- *et des translations*

Appendice A.9

Groupes de symétries

La notion de symétrie joue un rôle important en Mathématiques, en Physique ou en Chimie notamment, mais aussi en des nombreuses disciplines scientifiques. Il s'agit d'étudier les propriétés qui sont invariantes pour un certain type de transformations. Cela permet souvent de simplifier un problème (par exemple trouver des solutions non triviales à partir de solutions déjà connues), ou classifier les solutions, ou encore comprendre plus profondément la raison d'être de certaines propriétés (par exemple les propriétés remarquables des solutions de certains types d'équations différentielles).
A titre d'exemple, l'étude des symétries permet de décrire et classifier les polyèdres réguliers, ou d'expliquer pourquoi dans la nature on ne trouve que 32 types de cristaux. En cherchant des "symétries" on peut réduire l'ordre d'une équation différentielle, trouver des intégrales premières, etc. Et il est bien connu que, en étudiant les symétries des racines des équations algébriques, Évariste Galois a pu donner une réponse à l'ancien problème de leur résolubilité par radicaux en jetant les bases de la Théorie des Groupes. Pour terminer cette rapide introduction qui est bien loin d'être exhaustive, on peut encore mentionner l'importance de la notion de symétrie en Physique Théorique ou en Théorie quantique des champs.
Comme on peut l'imaginer, en Géométrie les transformations sous l'action desquelles il est intéressant d'étudier l'invariance d'une configuration de points ou d'une figure sont les isométries. Dans cette Appendice nous donnons un bref aperçu de ce type de problèmes.

Isométries laissant invariant un ensemble

Soit $\mathcal{E}$ un espace affine euclidien et $\mathrm{Is}(\mathcal{E})$ le groupe des isométries de $\mathcal{E}$.

Définition A.9.1 *Soit $S \in \mathcal{E}$ un sous-ensemble de $\mathcal{E}$. On dit qu'une isométrie f laisse $\mathcal{S}$ invariant si $f(\mathcal{S}) = \mathcal{S}$.*

Par exemple, si $\mathcal{S}$ est une pyramide dont la base est un polygône régulier n côtés, les rotations d'angle $\dfrac{2\pi}{n}$ autour de l'axe laissent $\mathcal{S}$ invariante.

Proposition A.9.2 - *L'ensemble* $\mathrm{Inv}(\mathcal{S})$ *des isométries qui laissent $\mathcal{S}$ invariant est un sous-groupe de* $\mathrm{Is}(\mathcal{E})$ *dit* ***groupe (complet) des symétries*** *de $\mathcal{S}$. Tout sous-groupe de* $\mathrm{Inv}(\mathcal{S})$ *est dit* ***groupe de symétries de $\mathcal{S}$****.*

La vérification est immédiate.

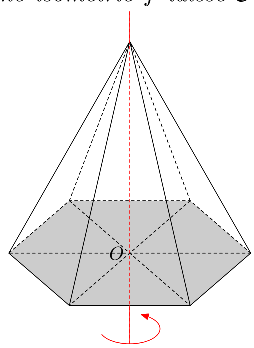


Figure 1

Exemple 1. - *Soit* $\mathcal{E}$ *de dimension* 2 *et* A, B *deux points distincts de* $\mathcal{E}$*. Déterminer* $\mathrm{Inv}(\{A, B\})$

Si A et B sont fixes, toute la droite $\mathcal{D} = (A, B)$ est fixe [1]. Donc, compte tenu du Théorème A.8.2, page 397 :

$$\mathrm{Inv}(\{A, B\}) = \{\mathrm{id}_E, s_D\}$$

Si A et B sont échangés entre eux, alors la point $\Omega = \dfrac{1}{2}A + \dfrac{1}{2}B$ est fixe pour toute $f \in \mathrm{Inv}(\{A, B\})$. Donc (cf. Théorème A.8.2) :

$$f \in \mathrm{Inv}(\{A, B\}) \iff \begin{cases} f = \mathrm{id}_{\mathcal{E}} & \\ f = s_\Delta & (\text{ réflexion par rapport à la droite } \Delta \\ & \quad \text{passant par } \Omega \text{ orthogonale à } \mathcal{D}) \\ f = R_\Omega & (\text{ rotation de centre } \Omega) \end{cases}$$

On a donc

$$\mathrm{Inv}(\{A, B\}) = \{\mathrm{id}_{\mathcal{E}}, R_\Omega, s_{\mathcal{D}}, s_\Delta\}$$

La table de multiplication de ce groupe est

	id	R_Ω	$s_{\mathcal{D}}$	s_Δ
id	id	R_Ω	$s_{\mathcal{D}}$	s_Δ
R_Ω	R_Ω	id	s_Δ	$s_{\mathcal{D}}$
$s_{\mathcal{D}}$	$s_{\mathcal{D}}$	s_Δ	id	R_Ω
s_Δ	s_Δ	$s_{\mathcal{D}}$	R_Ω	id

NOTA - Un groupe muni de cette table de multiplication est dit ***groupe de Klein***.

Exemple 2. - *Déterminer le sous-groupe* $\mathcal{I}$ *de* $\mathrm{Is}(2)$ *qui laisse invariant un triangle équilatère.*

Les sommets (A, B, C) du triangle forment un repère affine. Toute isométrie f est donc déterminée par ses valeurs sur A, B et C. Or, pour que le triangle soit invariant par f il faut que $f(A), f(B), f(C) \in \{A, B, C\}$ et puisque f est bijective, le triplet $\Big(f(A), f(B), f(C)\Big)$ est une permutation de (A, B, C). Il y a donc au plus $3!$ applications affines qui laissent invariant le triangle.
Soit $G = \dfrac{1}{3}(A + B + C)$. On a :

$$f(G) = \frac{1}{3}\Big(f(A) + f(B) + f(C)\Big) = \frac{1}{3}(A + B + C) = G$$

Donc G est invariant pour toutes les isométries de $\mathcal{I}$. $\mathcal{I}$ est donc un sous-groupe de $\mathrm{O}_G(E)$ et par conséquent il ne contient que
- l'identité
- les rotations autour de G
- les réflexions par rapport à une droite passant par G.

Les rotations autour de G sont de deux types :

$$R_1 = R_{G,2\pi/3} = \begin{cases} A & \longmapsto & B \\ B & \longmapsto & C \\ C & \longmapsto & A \end{cases} \qquad R_2 = R_{G,4\pi/3} = \begin{cases} A & \longmapsto & C \\ B & \longmapsto & A \\ C & \longmapsto & B \end{cases}$$

[1] Les points de $\mathcal{D}$ sont les barycentres de A et B et f conserve les barycentres, car elle est affine.

De symétries par rapport à des droites passant par G il y en au maximum 3, (car $\mathcal{I}$ comporte au plus 6 éléments et on en a déjà trouvé 3 : id, R_1, R_2). Il est clair que les symétries s_A, s_B, s_Cpar rapport aux médiatrices des côtés du triangle laissent le triangle invariant, c'est-à-dire elles appartiennent à $\mathcal{I}$. On donc

$$\mathcal{I} = \{\text{id}, R_1, R_2, s_A, s_B, s_C\}$$

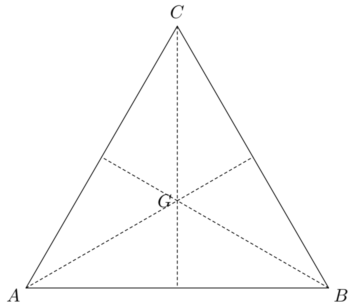


Figure 2

La table de multiplication de $\mathcal{I}$ est la suivante :

	Id	R_1	R_2	s_A	s_B	s_C
Id	Id	R_1	R_2	s_A	s_B	s_C
R_1	R_1	R_2	Id	s_C	s_A	s_B
R_2	R_2	Id	R_1	s_B	s_C	s_A
s_A	s_A	s_B	s_C	id	R_1	R_2
s_B	s_B	s_C	s_A	R_2	id	R_1
s_C	s_C	s_A	s_B	R_1	R_2	id

Groupes de symétries discrets d'un ensemble borné de $\mathbb{R}^3$

Dans cette section nous nous proposons d'étudier le groupe Inv($\mathcal{S}$), où $\mathcal{S}$ est un sous-ensemble ***borné*** d'un espace affine euclidien $\mathcal{E}$ de dimension 3. Sans perte de généralité (cf. Nota page 389) on peut supposer que $\mathcal{E}$ est l'espace $\mathbb{R}^3$ muni du produit scalaire canonique et de la structure affine standard.
Pour déterminer tous les groupes de symétries de $\mathcal{S}$, il faudrait classifier tous les sous-groupes de Is(3), ce qui est un problème très difficile. Heureusement, les groupes de symétries qui se présentent le plus souvent en Physique sont soit des groupes "discrets", soit des "groupes de Lie". L'étude de ces derniers n'étant pas du ressort d'un cours d'Algèbre Linéaire élémentaire, on s'intéressera aux *sous-groupes discrets* de Is(3).

Définition A.9.3 *Soit G un groupe de transformations qui agit sur un ensemble E (cf. définition A.1.11, page 342) et $G \cdot x$ l' orbite de x sous l'action de G. On dit que G est* ***discret*** *si :*

> $\forall\, x \in E$ *et pour toute boule $\mathcal{B}_r$ de centre x et de rayon r il existe au plus un nombre fini de points de l'orbite de x contenus dans $\mathcal{B}_r$, c'est-à-dire : $G{\cdot}x \cap \mathcal{B}_r$ est un ensemble fini.*

Il est clair que si G est fini, alors il est discret.
Un exemple de groupe infini et discret est le groupe engendré par les translations t_a, avec $a \in \mathbb{R}^3$, et $a \neq 0$. En revanche, par exemple, le groupe $G = \{t_q \mid q \in \mathbb{Q}^3\}$ n'est pas discret.

1. Une propriété de finitude

Théorème A.9.4 - *Tout groupe discret G de symétries d'un sous-ensemble borné de $\mathbb{R}^3$ est isomorphe à un sous-groupe* ***fini*** *du groupe orthogonal.*

Démonstration : Notons tout d'abord que ces groupes ne peuvent contenir ni les translations, ni les vissages, ni les glissements car l'une de ces transformations, suffisamment répétée, appliquera l'ensemble en dehors d'une sphère qui le contient. Par conséquent, G ne contient que des rotations et des rotations-réflexions (et bien-sûr l'identité). Il s'ensuit que chaque élément de G admet un point fixe. Il n'est pas évident, cependant, qu'ils admettent un point fixe *commun.* C'est en fait le cas :

Lemme – *Soit $\mathcal{S}$ un sous-ensemble borné de $\mathbb{R}^3$ et G un groupe discret de symétries de $\mathcal{S}$. Alors il existe un point $y \in \mathbb{R}^3$ qui est fixe pour tout $g \in G$.*

En effet, soit $x \in \mathcal{S}$. Tous les points de l'orbite de x sont dans $\mathcal{S}$, donc contenus dans une boule. Puisque G est discret l'orbite de x sera finie. Soit $G \cdot x = \{x_1, ..., x_p\}$ l'orbite de x et $g \in G$. Puisque g est une application affine, le barycentre y de l'orbite $G \cdot x$ sera transformé en le barycentre des points $\{g(x_1), ..., g(x_p)\}$: $g(y) = \sum_{i=1}^{p} g(x_i)/p$. Or tout élément $g \in G$ transforme l'orbite $G \cdot x$ en elle-même. Donc les points $\{g(x_1), ..., g(x_p)\}$ sont les points $\{x_1, ..., x_p\}$ dans un ordre différent. Il s'ensuit que

$$g(y) = \sum_{i=1}^{p} \frac{g(x_i)}{p} = \sum_{i=1}^{p} \frac{x_i}{p} = y$$

Aussi y est fixe pour tous les éléments de G. ♢

Il s'ensuit que G est un sous-groupe de $\mathrm{O}_y(3)$, lequel est isomorphe à $\mathrm{O}(3)$ (cf. (*) page 393). De plus, toutes les orbites sont finies. En effet, soit $x \in E$ et soit $R = \|x - y\|$. Pour toute $g \in \mathrm{Is}(E)$ on a :

$$\|g(x) - y\| = \|g(x) - g(y)\| = \|x - y\| = R.$$

Donc l'orbite de x est contenue dans la boule de centre y et rayon R et, comme G est discret, elle est finie.

On en déduit que G est fini. En effet, toute symétrie $g \in G$, étant une application affine, est déterminée par son action sur un repère affine, c'est-à-dire par son action sur 4 points a_1, a_2, a_3, a_4 non contenus dans un hyperplan. Comme les orbites de ces 4 points sont finies, les valeurs possibles des $g(a_i)$ sont en nombre fini, donc G est fini. □

2. Réduction à l'étude des sous-groupes finis de SO(3)[2]

Nous allons montrer maintenant que la détermination des sous-groupes finis de O(3) se ramène à la détermination des sous-groupes finis de SO(3).

Proposition A.9.5 *Soit G un sous-groupe fini de* $\mathrm{O}(3)$ *et* $K = G \cap \mathrm{SO}(3)$. *Alors*

1. *soit $G = K$, c'est-à-dire $G \subset \mathrm{SO}(3)$;*
2. *soit $G = K \cup j\,K$, où $j = -\,\mathrm{id}$;*
3. *soit G est isomorphe à un sous-groupe G^+ de* $\mathrm{SO}(3)$ *contenant K comme sous-groupe distingué d'indice 2 : $G^+ = K \cup K^+$, où $K^+ = \{g_0 = jg \mid g \in G, g \notin K\}$.*

Démonstration :
Supposons que $K \neq G$ et soit

$$\text{dét} \;: \begin{array}{ccc} G & \longrightarrow & \{-1, +1\} \\ g & \longmapsto & \text{dét g} \end{array}$$

[2]L'étude de cette section et de la suivante demande quelques connaissances élémentaires sur les groupes quotients.

l'application qui à $g \in G$ associe son déterminant. Il s'agit d'un homomorphisme de groupe, et $K = \text{Ker dét}$. Donc K est un sous-groupe distingué de G. Par ailleurs $G/K \simeq \text{Im dét} = \{-1, +1\}$. Donc K est d'indice 2 et par conséquent il est constitué par la moitié des éléments de G. Il y a deux possibilités :

1. $j \in G$. Alors $j \notin K$ (car $\text{dét}\, j = -1$). Donc : $G = K \cup j\,K$.
2. $j \notin G$. Considérons l'ensemble $K^+ = \{g_0 = jg \mid g \in G, g \notin K\}$. Si $g_0 \in K^+$, alors dét $g_0 = 1$ donc $K^+ \subset \mathrm{SO}(3)$. D'autre part :

$$K \cap K^+ = \emptyset$$

En effet, si $g_0 \in K \cap K^+$, on a $g_0 = jg$, avec $g \in G$ et $G \notin K$, d'où $j = g_0\, g^{-1}$. Puisque $g_0 \in K \subset G$ et $g^{-1} \in G$, il s'ensuivrait que $j \in G$, ce qui est exclu.
En outre

$$\text{Card} K = \text{Card} K^+$$

En effet, $G = K \cup K_1$, où $K_1 = \{g \in G \mid g \notin K\}$. Puisque K est un sous-groupe distingué d'indice 2, K et K_1 sont les deux classes de G (modulo K) et donc elles ont le même nombre d'éléments. D'autre part $K^+ = jK_1$. Il existe donc une bijection évidente de K^+ sur K_1 et donc $\text{Card}K^+ = \text{Card}K_1$, d'où $\text{Card}K = \text{Card}K^+$.

Soit $G^+ = K \cup K^+$. On voit facilement que G^+ est un groupe et tous ses éléments sont des rotations et que K est un sous-groupe distingué de G^+ d'indice 2. Il ne reste plus à démontrer que G est isomorphe à G^+. Soit

$$\varphi \ : \begin{array}{ccc} K \cup K^+ & \longrightarrow & G \\ g_0 & \longmapsto & \varphi(g_0) \end{array}$$

où $\varphi(g_0)$ est ainsi défini :

$$\varphi \ : \begin{cases} g_0 & \text{si } g_0 \in K \\ jg_0 & \text{si } g_0 \in K^+ \end{cases}$$

Cette application est évidemment bijective et un petit calcul montre que c'est un isomorphisme. □

Corollaire A.9.6 - *Les sous-groupes finis de* O(3) *peuvent être construits facilement à partir des sous-groupes finis de* SO(3)*. Aussi, la détermination des groupes de symétries discrets d'un sous-ensemble borné de* $\mathbb{R}^3$ *se ramène à la détermination des sous-groupes finis de* SO(3)*.*

3. Détermination des sous-groupes finis de SO(3)

Considérons la sphère S_r de centre 0de $\mathbb{R}^3$ et de rayon r : $\mathrm{S}_r = \{x \in \mathbb{R}^3 \mid \|x\| = r\}$. Il est clair que tout sous-groupe G de SO(3) laisse S_r invariante. En effet, pour toute $g \in \mathrm{SO}(3)$: $\|g(x)\| = \|x\| = r$, c'est-à-dire $g(x) \in \mathrm{S}_r$.

Définition A.9.7 *Soit G un sous-groupe de* SO(3)*. Un point $x \in S_r$ est dit **pôle** s'il existe une rotation $g \in G$, $g \neq \mathrm{id}$, telle que $g(x) = x$.*

Remarquons qu'un pôle est l'intersection de S_r avec l'axe d'une rotation non triviale (c'est-à-dire différente de id) ; pour chaque rotation non triviale il y a donc deux pôles sur S_r : $\pm x$.
On voit immédiatement que G *applique les pôles sur les pôles*. En effet, x est pôle de g_1, alors, pour toute $g \in G$, $g(x)$ est pôle de $g \circ g_1 \circ g^{-1}$. Donc l'ensemble des pôles est partagé en orbites par l'action de G.

Lemme –
Soit G un sous-groupe fini de SO(3)*. Si x est un pôle, on note G_x le* groupe d'isotropie *de x, c'est-à-dire l'ensemble des rotations de G qui laissent fixe x : $G_x = \{g \in G \mid g(x) = x\}$. On a alors les propriétés suivantes :*

1. Soit $\mathcal{O} = \{ x_1, ..., x_p \}$ orbite de pôles. Alors :
$$\operatorname{Card} G_{x_1} = \operatorname{Card} G_{x_2} = \cdots = \operatorname{Card} G_{x_p}.$$
2. Soient $\mathcal{O}_1, \dots \mathcal{O}_k$ les différentes orbites et notons
$$p_i = \operatorname{Card} \mathcal{O}_i \qquad n_i = \operatorname{Card} G_x \quad [\text{si } x \in \mathcal{O}_i] \quad \text{et} \quad n = \operatorname{Card} G$$
Alors pour chaque $i = 1, \dots, k$, on a : $n_i \geq 2$, et $n = p_i\, n_i$.
3. Le nombre de rotations non triviales qui laissent fixe un pôle quelconque est : $\sum_{i=1}^{k} p_i(n_i - 1)$.
4. On a la formule
$$\boxed{2\left(1 - \frac{1}{n}\right) = \sum_{i=1}^{k} \left(1 - \frac{1}{n_i}\right)} \qquad (*)$$

Démonstration

1. vient du fait que si $g \in G$ est tel que $x_j = g(x_i)$, l'application φ définie par : $\varphi : \begin{array}{ccc} G_{x_i} & \longrightarrow & G_{x_j} \\ g_0 & \longmapsto & g \circ g_0 \circ g^{-1} \end{array}$ est bijective.
2. $n_i \geq 2$, car il existe au moins deux rotations qui laissent fixe le pôle x_i : celle qui fait que x_i soit pôle et l'identité.
D'autre part G_{x_i} est un sous-groupe (non distingué) de G. En notant $[G \backslash G_{x_i}]_l$ l'ensemble des classes à gauche de G modulo G_{x_i}, on sait que que $[G \backslash G_{x_i}]_l$ est en bijection avec l'orbite de x_i, $G \cdot x_i$ (prendre l'application $\psi_{x_i}([g]) = g(x_i)$ et vérifier qu'elle est bien définie et que c'est une bijection). On en déeduit l'égalité : $\operatorname{Card}(G) = \operatorname{Card}(G_{x_i}) \cdot \operatorname{Card}(G \cdot x_i)$.
3. Le nombre des rotations *non triviales* qui laissent fixe le pôle x_i est $n_i - 1$. Par conséquent, le nombre des rotations non triviales qui laissent fixe un pôle quelconque de l'orbite de x_i est $p_i(n_i - 1)$ (car il y a p_i pôles sur cette orbite).
4. En effet, il y a $n - 1$ rotations non triviales dans G : chacune laisse fixes deux pôles et chaque rotation doit être comptée deux fois (une fois pour le pôle x et une fois pour le pôle $-x$). On a donc :
$$2(n-1) = \sum_{i=1}^{k} p_i(n_i - 1)$$
on en déduit facilement la formule. □

Théorème A.9.8 *Compte tenu de la formule (*), il n'y a que 5 sous-groupes finis de* SO(3) *:*

A) *Le groupe noté C_n, caractérisé par : $k = 2$, $n = n_1 = n_2$.*
Il s'agit du groupe des rotations laissant invariante la ***n-pyramide****.*

B) *Le groupe noté D_m, caractérisé par : $k = 3$, $n = 2m$, $m \geq 2$, $n_1 = n_2 = 2$, $n_3 = m$*
Il s'agit du groupe des rotations laissant invariant le ***m-prisme****, dit* ***groupe diédral****.*

C) *Le groupe noté $\mathcal{T}$, caractérisé par : $k = 3$, $n = 12$, $n_1 = 2$, $n_2 = n_3 = 3$.*
Il s'agit du groupe des rotations laissant invariant le ***tétraèdre****. .*

D) *Le groupe noté $\mathcal{O}$, caractérisé par :* $k=3,\ \ n=24;\ \ n_1=2,\, n_2=3,\, n_3=4.$
Il s'agit du groupe des rotations laissant invariant le **cube** *ainsi que l'***octaèdre**.

E) *Le groupe noté $\mathcal{Y}$, caractérisé par :* $k=3,\ \ n=60,\ \ n_1=2,\, n_2=3,\, n_3=5.$
*Il s'agit du groupe des rotations laissant invariant l'***icosaèdre** *ainsi que le* **dodécaèdre**.

Démonstration : En effet :

a) On ne peut avoir que $k=2$ où $k=3$. Car si on avait $k=1$, on aurait $2\left(1-\dfrac{1}{n}\right)=1-\dfrac{1}{n_1}$. Or : $n \geq n_1$, d'où $\dfrac{1}{n} \leq \dfrac{1}{n_1}$, donc $1-\dfrac{1}{n} \geq 1-\dfrac{1}{n_1}$ et par conséquent : $2\left(1-\dfrac{1}{n}\right) > 1-\dfrac{1}{n_1}$. Si on avait $k \geq 4$, on aurait $\displaystyle\sum_{i=1}^{k}\left(1-\frac{1}{n_i}\right) \geq 2$, d'où $2\left(1-\dfrac{1}{n}\right) \geq 2$, ce qui est exclu.

b) Si $k=2$ on a : $2-\dfrac{2}{n}=2-\dfrac{1}{n_1}-\dfrac{1}{n_2}$, d'où $\dfrac{2}{n}=\dfrac{1}{n_1}+\dfrac{1}{n_2}$. Or $\dfrac{1}{n_i} \geq \dfrac{1}{n}$. Si, par exemple, on avait $\dfrac{1}{n_1} > \dfrac{1}{n}$, on aurait $2-\dfrac{2}{n} < 2-\dfrac{1}{n_1}-\dfrac{1}{n_2}$. Donc $n_1=n$. De même : $n_2=n$.

c) Supposons $k=3$

i) Si on avait $n_1 \geq 3$, on aurait $\dfrac{1}{n_1}+\dfrac{1}{n_2}+\dfrac{1}{n_3} \leq 1$, tandis que on a : $1+\dfrac{2}{n} > 1$.
Donc $n_1 \geq 3$ est impossible et comme les $n_i \geq 2$, on a : $n_1=2$.

ii) Si on avait $n_2 \geq 4$, on aurait $\dfrac{1}{n_2} \leq \dfrac{1}{4}$ et (puisqe $n_3 \geq n_2$) $\dfrac{1}{n_3} \leq \dfrac{1}{4}$. D'ou :

$$\frac{1}{n_1}+\frac{1}{n_2}+\frac{1}{n_3} \leq \frac{1}{2}+\frac{1}{4}+\frac{1}{4}=1 \quad \text{tandis que : } 1+\frac{2}{n} > 1$$

iii) Si $n_2=2$, on a $1+\dfrac{2}{n}=1+\dfrac{1}{n_3}$. Donc $n_3=\dfrac{n}{2}$, ce qui implique que n est pair et $n \geq 4$ (car $n_3 \geq 2$).

iv) Si $n_2=3$, on a : $1+\dfrac{2}{n}=\dfrac{1}{2}+\dfrac{1}{3}+\dfrac{1}{n_3}$, d'où : $\dfrac{1}{6}+\dfrac{2}{n}=\dfrac{1}{n_3}$. Donc $\dfrac{1}{n_3} > \dfrac{1}{6}$, donc : $n_3 < 6$. Comme $n_3 \geq n_2 = 3$, on aura trois cas possibles : $n_3=3,\ ,\ n_3=4\,,\ n_3=5$. $\square$

La construction des solides invariants pour chacun des cas est plus compliquée. Regardons juste le cas A) : $k=2,\ \ n=n_1=n_2=n_3$.
Il y a deux orbites et sur chacune $p_i=\dfrac{n}{n_i}=1$ pôle. Chaque pôle est fixe pour $n_i=n$ rotations, c'est-à-dire tous les éléments de G laissent fixes les deux pôles. Aussi G *est constitué par n rotations autour d'un axe.*

Propriété A.9.9 -
Soit G un sous-groupe fini de $\mathrm{SO}(3)$ *d'ordre $n \geq 2$, constitué par des rotations autour d'un même axe. Alors G est le groupe cyclique C_n engendré par la rotations d'angle* $\theta=\dfrac{2\pi}{n}$.

En effet, soit $G=\{e, g_1, \ldots, g_{n-1}\}$, g_i étant des rotations d'angles $\theta_1, \ldots, \theta_{n-1}$, compris entre 0 et 2π. Supposons que θ_1 est le plus petit de ces angles. D'après la théorème de la division euclidienne, on a :

$$\theta_i = m_i\,\theta_1 + \varphi_i, \qquad \text{avec } \ 0 \leq \varphi_i \leq \theta_1.$$

Or $g_i g_1^{-m_i}$ est une rotation d'angle $\theta_i - m_i\theta_1 = \varphi_i$ et appartient à G, d'où : $\varphi_i = 0$. Donc tous les g_i sont des rotations d'angles multiples de θ_1, c'est-à-dire g_1 engendre G. D'autre part, g_1 est d'ordre n, car g_1 engendre G, donc n est le plus petit entier positif tel que $g_1^n = e$, d'où : $n\theta_1 = 0[\mathrm{mod}\, 2\pi]$, c'est-à-dire $\theta_1 = \dfrac{2\pi}{n}$. $\diamondsuit$

Considédrons la *n-pyramide*, c'est-à-dire la pyramide droite dont la base est un polygone régulier à n côtés, telle que la distance du sommet de la pyramide à l'un des sommets de la base n'est pas égale à la longueur des côtés de la base. Il est clair que le groupe des rotations qui laissent invariante la n-pyramide est C_n.

Pour $n = 2$ on considère la "pyramide" dont la base est la figure 3 ci-contre :

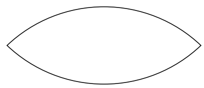


Figure 3

Enfin, signalons que l'on appelle *m-prisme* un cylindre droit dont la base est un polygone régulier à m côtés et la hauteur n'est pas égale à la longueur des côtés de la base (pour $m = 2$ on prend comme base la base de la 2-pyramide ci-dessus). Ci-contre, à la figure 4, un trois-prisme.

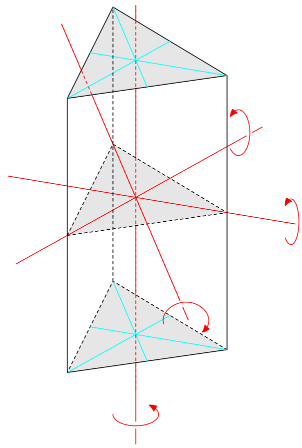


Figure 4

4. Sous-groupes finis de O(3)

Terminons cette section en donnons quelques indications sur les sous-groupes finis de O(3). D'après le Proposition A.9.5, on peut construire tous les sous-groupes finis de O(3) à partir des 5 sous-groupes finis de SO(3) :

- On trouve d'abord les 5 groupes notés

$$C_n \cup jC_n, \quad D_m \cup jD_m, \quad \mathcal{T} \cup j\mathcal{T}, \quad \mathcal{O} \cup j\mathcal{O}, \quad \mathcal{Y} \cup j\mathcal{Y}$$

  qui s'obtiennent en ajoutant aux rotations des sous-groupes finis de SO(3) leurs composéses par la symétrie par rapport 0.

- La construction des autres sous-groupes (ceux du point 3. de la Proposition A.9.5), est un peu plus compliquée. En principe on devrait en trouver 5 autres, mais en fait on n'en trouve que 3. On peut montrer en effet que dans les cas des groupes $\mathcal{T}$ et $\mathcal{Y}$ on ne peut pas construire un groupe de rotations (le G^+ de la Proposition) contenant un sous-groupe distingué d'indice deux.

En résumant : *il existe exactement 13 sous-groupes finis de* O(3).

Pour chacun de ces groupes on peut construire le solide qui l'admet comme groupe complet de symétries. Pour cela on pourra consulter, par exemple :

Paul B. Yale
Geometry and Symmetry
Dover Publications Inc. 1989

Pour la classification groupes cristallographiques on pourra consulter, par exemple :

Willard Miller
Symmetry Groups and their Applications
Academic Press, 1972

Appendice A.10

Sur la décomposition des transformations orthogonales

1. Symétries, réflexions, retournements.

Définition.

*On appelle **symétrie** un endomorphisme s d'un espace vectoriel E tel que $s^2 = \mathrm{id}$ et $s \neq \mathrm{id}$. Si E est euclidien, la symétrie est dite orthogonale si elle est une transformation orthogonale.*

Puisque $X^2 - 1$ est annulateur de s, s est diagonalisable et $E = E_1 \oplus E_{-1}$. Il existe donc une base (e_i) de E, telle que

$$M(s)_i = \begin{pmatrix} \boxed{\begin{matrix} 1 & & \\ & \ddots & \\ & & 1 \end{matrix}} & \\ & \boxed{\begin{matrix} -1 & & \\ & \ddots & \\ & & -1 \end{matrix}} \end{pmatrix}$$

le bloc des 1 étant d'ordre $n_1 = \dim E_1$ et celui des -1 d'ordre $n_2 = \dim E_{-1}$.
On dit que s est une symétrie *par rapport à E_1 parallèlement à E_{-1}*

Si s est orthogonale, on a $s^* \circ s = \mathrm{id}$, et comme $s^2 = \mathrm{id}$ et $\det s \neq 0$, on a $s^* = s$, c'est-à-dire s est autoadjointe et, par conséquent, les espaces propres E_1 et E_{-1} sont orthogonaux. On peut donc trouver une *base orthonormée* dans laquelle s a l'expression (1). Réciproquement, si pour un endomorphisme s il existe une base *orthonormée* dans laquelle s a la forme (1), alors s est une symétrie *orthogonale*.

Définition. – *On appelle **réflexion** une symétrie orthogonale par rapport à un sous-espace de dimension $n-1$ (c'est-à-dire l'espace propre correspondant à $+1$ est de dimension $n-1$). On appelle **retournement** une symétrie orthogonale par rapport àun sous-espace de dimension $n-2$.*

Noter que dans $\mathbb{R}^3$ les retournements sont les rotations d'angle π.

On a donc :

1. s est une réflexion si et seulement si il existe une base *orthonormée* (e_i) telle que

$$M(s)_{e_i} = \begin{pmatrix} \boxed{\begin{matrix} 1 & & \\ & \ddots & \\ & & 1 \end{matrix}} & \\ & -1 \end{pmatrix} \tag{2}$$

2. ρ est un retournement si et seulement si il existe une base *orthonormée* (e_i) telle que

$$M(\rho)_{e_i} = \begin{pmatrix} \boxed{\begin{matrix} 1 & & \\ & \ddots & \\ & & 1 \end{matrix}} & \\ & \boxed{\begin{matrix} -1 & \\ & -1 \end{matrix}} \end{pmatrix} \tag{3}$$

2. Décomposition des transformations orthogonales

Nous allons donner d'abord une autre démonstration simple du résultat de l'exercice 29 du chapitre 7, qui permet de mieux comprendre l'idée sous-jacente [1].

Théorème 1 - ***Décomposition des transformations orthogonales en produit de réflexions***

Soit f une transformation orthogonale d'un espace euclidien de dimension $n \geq 2$. Alors f peut s'écrire sous la forme :

$$f = s_1 \circ \cdots \circ s_r, \quad \text{avec } \ r \leq n$$

où les s_i sont des réflexions (la décomposition n'est pas unique).

Démonstration :

On considère une b.o.n. qui donne la forme réduite de l'exercice 28 du chapitre 7. En notant que (quitte changer l'ordre de la base) les -1 peuvent être groupés deux à deux, donnant des matrices de rotations dans un plan [2], on a deux cas possibles selon que les -1 sont en nombre pair (c'est-à-dire dét $f = 1$) ou impair (c'est-à-dire dét $f = -1$.

1. dét $\mathbf{f = 1}$ Dans ce cas :

$$M(f)_{e_i} = \begin{pmatrix} \boxed{R_1} & & & & & \\ & \ddots & & & & \\ & & \boxed{R_k} & & & \\ & & & 1 & & \\ & & & & \ddots & \\ & & & & & 1 \end{pmatrix} \text{où : } R_i = \begin{pmatrix} \cos\theta_i & -\sin\theta_i \\ \sin\theta_i & \cos\theta_i \end{pmatrix} \neq I, 2k \leq n \tag{4}$$

[1] Pour les calculs explicites des éléments de la décomposion de f, il est plus pratique, cépendant, d'utiliser la construction par récurrence (cf. exemple de l'exercice 29 du chapitre 7).

[2] car $\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} \cos\pi & -\sin\pi \\ \sin\pi & \cos\pi \end{pmatrix}$

Or dans le plan toute rotation est un produit de deux réflexions [3] on peut donc écrire $R_i = S_iT_i$ où S_i et T_i sont des matrices de reflexions dans le plan. On aura donc :

$$M(f)_{e_i}=\begin{pmatrix} \boxed{S_1} & & & & \\ & 1 & & & \\ & & \ddots & & \\ & & & \ddots & \\ & & & & 1 \end{pmatrix}\begin{pmatrix} \boxed{T_1} & & & & \\ & 1 & & & \\ & & \ddots & & \\ & & & \ddots & \\ & & & & 1 \end{pmatrix}\cdots\cdots$$

$$\cdots\cdots\begin{pmatrix} 1 & & & & \\ & \ddots & & & \\ & & \boxed{S_k} & & \\ & & & \ddots & \\ & & & & 1 \end{pmatrix}\begin{pmatrix} 1 & & & & \\ & \ddots & & & \\ & & \boxed{T_k} & & \\ & & & \ddots & \\ & & & & 1 \end{pmatrix}$$

Or toutes les matrices qui apparaîssent dans l'expression ci-dessus sont des matrices de reflexions.

En effet, par exemple soit s_1l'application qui dans la base (e_i) est représentée par la première de ces matrices. S_1 est la matrice d'une reflexion σ_1 dans le plan $F = \text{Vect}\,(e_1, e_2)$: $\sigma_1 = s_1|_F$. Il existe une b.o.n. $(\varepsilon_1, \varepsilon_2)$ telle que $M(\sigma_1)_{\varepsilon_i} = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$. il est clair que dans la base $(\varepsilon_1, \varepsilon_2, e_3, ..., e_n)$ la matrice de s_1a la forme (*) ci-dessus.

Donc f est produit de $2k \leq n$ réflexions.

2. dét f = -1 Dans ce cas

$$M(f)_{e_i} = \begin{pmatrix} \boxed{R_1} & & & \\ & \ddots & & \\ & & \boxed{R_k} & \\ & & & \boxed{\begin{matrix} 1 & & \\ & \ddots & \\ & & 1 \end{matrix}} & \\ & & & & -1 \end{pmatrix} \qquad (2k+1 \leq n)$$

On peut écrire

$$M(f)_{e_i} = \begin{pmatrix} \boxed{R_1} & & & & \\ & \ddots & & & \\ & & \boxed{R_k} & & \\ & & & \boxed{\begin{matrix} 1 & & \\ & \ddots & \\ & & 1 \end{matrix}} & \\ & & & & 1 \end{pmatrix}\begin{pmatrix} 1 & & & & \\ & \ddots & & & \\ & & 1 & & \\ & & & \boxed{\begin{matrix} 1 & & \\ & \ddots & \\ & & 1 \end{matrix}} & \\ & & & & -1 \end{pmatrix}$$

Or la première de ces matrices se décompose en produit de $2k$ réflexions (d'après (a)) et la seconde est une réflexion. Donc f est produit de $2k + 1 \leq n$ réflexions.

[3]car le produit de deux matrice de d'eterminant -1 a un déterminant égal à $+1$.

Théorème 2 - ***Décomposition des transformations orthogonales en produit de retournements***

Soit E une espace euclidien de dimension $n \geq 3$. Alors toute transformation orthogonale directe *est un produit de k retournements, avec $k \leq n-1$.*

NOTA.

1. Ce théorème montre qu'en dimension 3 toute rotation est produit d'au plus deux rotations d'angle π.
2. En dimension 2. si on désigne par R_θ la rotation d'angle θ, on a : $R_\theta \circ R_\varphi = R_{\theta+\varphi}$. Cette formule *est donc fausse en dimension* 3 (d'après 1).

Démonstration - Considérons une transformation orthogonales directe f. Puisque dét $f = 1$, il existe une b.o.n. (e_i) telle que $A = M(f)_{e_i}$ a l'expression (4).
On a deux cas possibles :

1. ***1 est valeur propre.*** Dans ce cas, il existe une b.o.n. dans laquelle la matrice de f a la forme (4) que l'on peut écrire :

$$A = \begin{pmatrix} \boxed{R_1} & & & \\ & 1 & & \\ & & \ddots & \\ & & & \ddots & \\ & & & & 1 \end{pmatrix} \cdots\cdots \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & & & \\ & & & \boxed{R_k} & & & \\ & & & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix}, \; 2k{+}1 \leq n$$

Or chacune de ces matrices est un produit de deux retournements et donc A est produit de $2k \leq n-1$ retournements. Vérifions-le, par exemple pour la première matrice. En écrivant $R_1 = S_1 T_1$, où S_1 et T_1 sont deux réflexions, on peut écrire :

$$\begin{pmatrix} \boxed{R_1} & & & & \\ & 1 & & & \\ & & \ddots & & \\ & & & \ddots & \\ & & & & 1 \end{pmatrix} = \begin{pmatrix} \boxed{S_1} & & & & \\ & 1 & & & \\ & & \ddots & & \\ & & & 1 & \\ & & & & -1 \end{pmatrix} \begin{pmatrix} \boxed{T_1} & & & & \\ & 1 & & & \\ & & \ddots & & \\ & & & 1 & \\ & & & & -1 \end{pmatrix}$$

Le même raisonnement que l'on a fait pour le théorème 1., permet de voir que les deux matrices à second membre représentent des retournements (S_1 est semblable à $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$, etc.

2. ***1 n'est pas valeur propre*** (notons que cela implique que E est de dimension paire). Il existe alors une base orthonormée (e_i) dans laquelle

$$A = \begin{pmatrix} \boxed{R_1} & & \\ & \ddots & \\ & & \boxed{R_k} \end{pmatrix}, \quad R_j \neq I, \qquad 2k = n$$

Le même raisonnement que ci-dessus montre que A est produit de $2k \leq n$ retournements. □

NOTA - *On peut en fait montrer que A est produit de $p \leq n-1$ retournements (la démonstration est plus compliqué).*

Appendice A.11

Coniques et quadriques

Coniques

Définition A.11.1 – *Soient q une forme quadratique non nulle et φ une forme linéaire sur l'espace euclidien $\mathbb{R}^2$. On supposera $\mathbb{R}^2$ muni du produit scalaire canonique $\langle\ ,\ \rangle$. On appelle* ***conique*** *l'ensemble $\mathcal{C}$ des $v \in \mathbb{R}^2$ vérifiant l'équation :*

$$q(v) + \varphi(v) = k \quad \text{où} \quad k \in \mathbb{R}.$$

Nous allons donner une classification des coniques selon la signature de q. En changeant éventuellement le signe des deux membres de l'équation, on peut supposer que la signature de q est $(2,0)$, $(1,1)$, ou $(1,0)$.

Si $\{e_1, e_2\}$ est la base canonique et $v = Xe_1 + Ye_2$, l'équation d'une conique est du type

$$\alpha\, X^2 + 2\,\beta\, X\, Y + \gamma\, Y^2 + \lambda\, X + \mu\, Y = k \tag{1}$$

où $\alpha, \beta, \gamma, \lambda, \mu \in \mathbb{R}$.

D'après le théorème 9.24, il existe une base $\{v_1, v_2\}$ orthogonale à la fois pour q et pour $\langle\ ,\ \rangle$: il suffit de prendre une base de vecteurs propres de la matrice

$$A = \begin{pmatrix} \alpha & \beta \\ \beta & \alpha \end{pmatrix}$$

de manière qu'elle soit orthonormée pour $\langle\ ,\ \rangle$. Les directions définies par v_1 et v_2 sont dites ***directions principales***. Si $v = X'\, v_1 + Y'\, v_2$, l'équation de la conique s'écrira dans la nouvelle base :

$$a\, {X'}^2 + b{Y'}^2 - 2\, r\, X' - 2\, s\, Y' = k \tag{2}$$

où $a = q(v_1)\,,\ b = q(v_2)\,,\ -2\, r = \varphi(v_1)\,,\ -2\, s = \varphi(v_2)$.

1. q est non dégénérée - Dans ce cas a et b sont non nuls. On peut écrire l'équation (2) sous la forme

$$a\left(X' - \frac{r}{a}\right)^2 + b\left(X' - \frac{s}{b}\right)^2 = h.$$

En posant

$$x = X' - \frac{r}{a}\,, \qquad y = X' - \frac{s}{b}$$

on obtient l'équation :

$$a\, x^2 + b\, y^2 = h \tag{3}$$

où $h = k - (r/a)^2 - (s/b)^2$. Notons que l'on passe de l'équation (1) à l'équation (3) par une isométrie affine : une transformation orthogonale pour passer de la base canonique à la base définie par les directions principales, suivie d'une translation.

`a)` `sign (q) = (2,0)` , c'est-à-dire $a > 0$ et $b > 0$. Si $h = 0$, $\mathcal{C}$ se réduit à un point ; si $h < 0$, $\mathcal{C} = \emptyset$. Supposons que $h > 0$. Dans ce cas $\mathcal{C}$ est une ***ellipse*** de centre $O = \left(\frac{r}{a}, \frac{s}{b}\right)$; elle peut se mettre sous la forme

$$\frac{x^2}{A^2} + \frac{y^2}{B^2} = 1$$

en posant $A = \sqrt{h/a}$ et $B = \sqrt{h/b}$ (cf. fig. 1)

`b)` `sign (q) = (1,1)` , cest-à- dire $ab < 0$. Si $h \neq 0$, $\mathcal{C}$ est une ***hyperbole*** : si, par exemple $a > 0$ et $b < 0$, elle peut se mettre sous la forme

$$\frac{x^2}{A^2} - \frac{y^2}{B^2} = 1$$

en posant $A = \sqrt{h/a}$ et $B = \sqrt{-h/b}$ (cf. fig. 2).

Si $h = 0$, $\mathcal{C}$ se réduit aux deux droites d'équation $y = \pm\sqrt{|\,a/b\,|}\,x$ (cf. fig. 3).

2. q est dégénérée - Dans ce cas $ab = 0$ et $\operatorname{sign} q = (1,0)$.

Supposons, par exemple que $a \neq 0$ et $b = 0$. Cela signifie que la direction principale v_2 est isotrope pour q. L'équation (2) peut s'écrire :

$$a\left(X' - \frac{r}{a}\right)^2 - 2\,s\,Y' = h \qquad (2')$$

`a)` Si $s \neq 0$, ce qui signifie que $v_2 \notin \operatorname{Ker}\varphi$. En effectuant le changement de coordonnées affine $x = X' - \frac{r}{a}$, $y = h + 2\,s\,Y'$, on obtient l'équation

$$y = a\,x^2$$

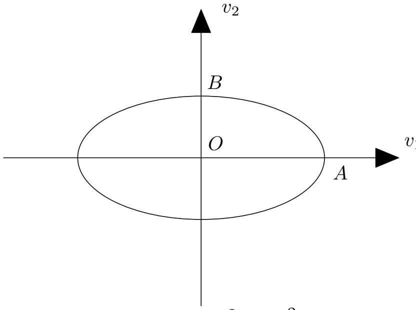
qui est l'équation d'une ***parabole*** (cf. fig 4.)

`b)` Si $s = 0$, c'est-à-dire si $v_2 \in \operatorname{Ker}\varphi$, l'équation (2') se réduit à $x^2 = h/a$.
Si $h < 0$, $\mathcal{C} = \emptyset$; si $h = 0$, $\mathcal{C}$ se réduit à la doite $x = 0$; si $h > 0$, $\mathcal{C}$ est constituée de deux droites parallèles (cf. fig. 5).

Nous avons donc :

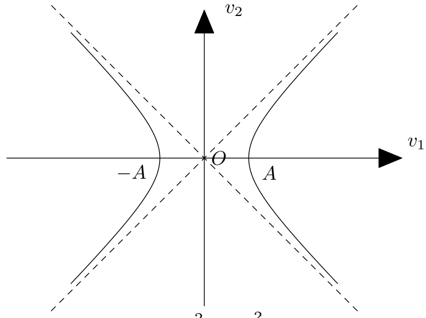
Théorème A.11.2 – *Soit $\mathcal{C}$ une conique définie par l'équation $q(v)+\varphi(v) = k$. On suppose que $\mathcal{C} \neq \emptyset$ et que $\mathcal{C}$ ne se réduit pas à un point. Alors :*

1. *Si $\operatorname{sign} q = (2,0)$ alors $\mathcal{C}$ est une* ***ellipse.***
2. *Si $\operatorname{sign} q = (1,1)$ alors $\mathcal{C}$ est une* ***hyperbole*** *qui éventuellement dégénère en deux droites non parallèles.*
3. *Si $\operatorname{sign} q = (1,0)$ alors $\mathcal{C}$ est une* ***parabole****, qui dégénère en une droite, ou en deux droites parallèles, si la direction principale isotrope est contenue dans $\operatorname{Ker}\varphi$.*



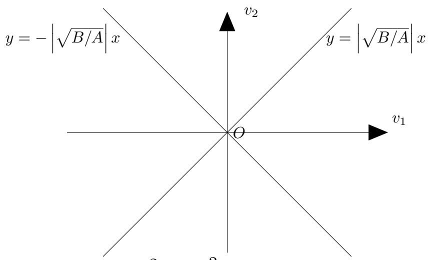
Ellipse : $\frac{x^2}{A^2} + \frac{y^2}{B^2} = 1$

Figure 1



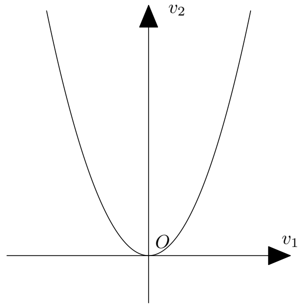
Hyperbole : $\frac{x^2}{A^2} - \frac{y^2}{B^2} = 1$

Figure 2



$\frac{x^2}{A^2} - \frac{y^2}{B^2} = 0$

Figure 3



Parabole : $y = ax^2$

Figure 4

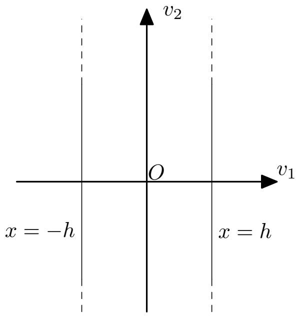


Figure 5

Quadriques

Définition A.11.3 – *Soient q une forme quadratique non nulle et φ une forme linéaire sur l'espace euclidien $\mathbb{R}^3$. On supposera $\mathbb{R}^3$ muni du produit scalaire canonique $\langle \, , \rangle$. On appelle* ***quadrique*** *l'ensemble $\mathcal{Q}$ des $v \in \mathbb{R}^3$ vérifiant l'équation :*

$$q(v) + \varphi(v) = k\,, \quad \text{où} \;\; k \in \mathbb{R}.$$

En échangeant éventuellement le signe des deux membres de l'équation, on peut supposer que $\operatorname{sign} q = (r, s)$ avec $r \geq s$.
Dans la base canonique $\{e_1, e_2, e_3\}$ de $\mathbb{R}^3$, l'équation de $\mathcal{Q}$ s'écrit

$$\alpha X^2 + \beta Y^2 + \gamma Z^2 + 2\lambda XY + 2\mu XZ + 2\nu YZ + \rho X + \sigma Y + \tau Z = k \tag{4}$$

Comme ci-dessus, on peut construire une base $\{v_1, v_2, v_3\}$ orthogonale pour q et orthonormée pour $\langle\, ,\, \rangle$, en prenant une base orthonormée de vecteurs propres de la matrice

$$A = \begin{pmatrix} \alpha & \lambda & \mu \\ \lambda & \beta & \nu \\ \mu & \nu & \gamma \end{pmatrix}$$

Dans cette base $\mathcal{Q}$ s'écrit :

$$a\, X'^2 + b\, Y'^2 + c\, Z'^2 - 2\, r\, X' - 2\, s\, Y' - 2\, t\, Z' = k \tag{6}$$

où $a = q(v_1)\,, b = q(v_2)\,, c = q(v_3)\,, -2\,r = \varphi(v_1)\,, -2\,s = \varphi(v_2)\,, -2\,t = \varphi(v_3)$. Les directions définies par $v_1,\, v_2,\, v_3$ sont dites *principales*.

1. rang q = 3 - Dans ce cas : $abc \neq 0$. En effectuant une translation de vecteur de translation $\overrightarrow{\omega} = (r/a,\, s/b,\, t/c)$, $\mathcal{Q}$ s'écrit :

$$a\, x^2 + b\, y^2 + c\, z^2 = h \tag{7}$$

où $x = X' - r/a\,,\ y = Y' - s/b\,,\ z = Z' - t/c\,.$ x, y, z sont donc les coordonnées dans le repère affine d'origine $\overrightarrow{\omega}$ et d'axes dirigés par les directions principales.

a) `sign (q) = (3,0).` Si $h < 0$, $\mathcal{Q} = \emptyset$; si $h = 0$, $\mathcal{Q}$ se réduit à un point ; si $h > 0$, $\mathcal{Q}$ est un ***ellipsoïde*** (cf. fig. 6).

b) `sign (q) = (2,1).` Quitte à changer le signe des deux membres, on peut supposer $a > 0$, $b > 0$, et $c < 0$.

i) Si $h > 0$, l'équation de $\mathcal{Q}$ s'écrit

$$a\, x^2 + b\, y^2 = h + |\, c\,|\, z^2$$

En coupant avec un plan d'équation $z = K$ on obtient des ellipses : $\mathcal{Q}$ est un ***hyperboloïde à une nappe*** (cf. fig. 7).

ii) Si $h = 0$, $\mathcal{Q}$ est ***cône*** d'axe dirigé par le vecteur v_3 et dont les sections par des plans orthogonaux à v_3 sont des ellipses (éventuellement des cercles, si $a = b$) (cf. fig. 8.)

iii) Si $h < 0$, les sections de $\mathcal{Q}$ avec des plans $z = K$ sont des ellipses si $K \geq \sqrt{|\, h/c\,|}$ qui se réduisent à un point pour $K = \sqrt{|\, h/c\,|}$; pour $K \leq \sqrt{|\, h/c\,|}$ ces plans n'intersectent pas $\mathcal{Q}$. $\mathcal{Q}$ est un ***hyperboloïde à deux nappes*** (cf. fig. 9).

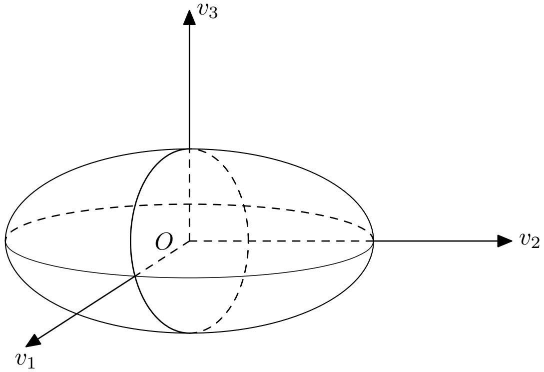
2. rang q = 2 - On peut supposer que $a\, b \neq 0$ et $c = 0$ (ceci signifie que la direction principale définie par v_3 est isotrope). L'équation (6) s'écrit alors :

$$a\, x^2 + b\, y^2 - 2\, t\, Z' = h,$$

où $x = X' - r/a\,,\ y = Y' - s/b.$

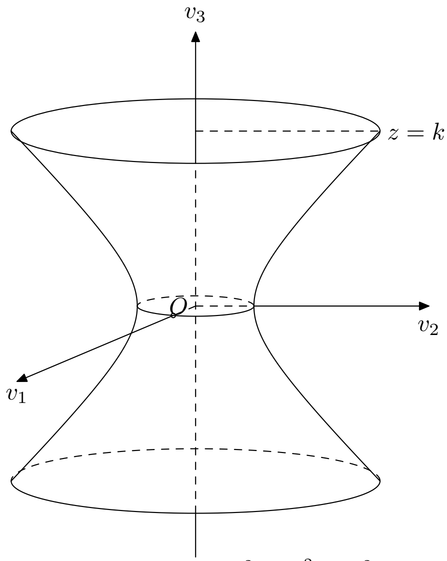
a) Si $t = 0$, c'est-à-dire si $v_3 \in \operatorname{Ker} \varphi$, alors $\mathcal{Q}$ est un cylindre d'axe dirigé par v_3 et dont les sections sont des ellipses ou des hyperboles[1] (cf. théorème A.11.2, *1. 2.*) : ***cylindre elliptique*** ou ***hyperbolique.*** (cf. fig. 10).

[1]éventuellement dégénérées en deux droites non parallèles ou en un point.



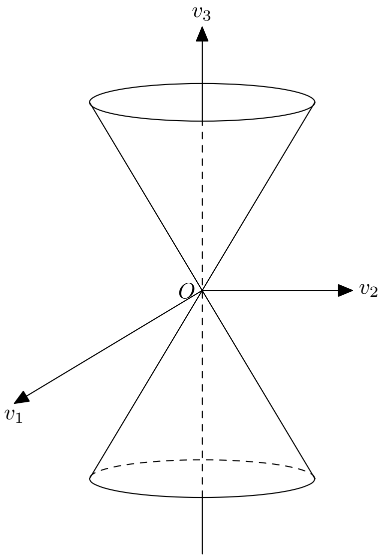
Elllipsoïde : $\frac{x^2}{A^2} + \frac{y^2}{B^2} + \frac{z^2}{C^2} = 1$

Figure 6



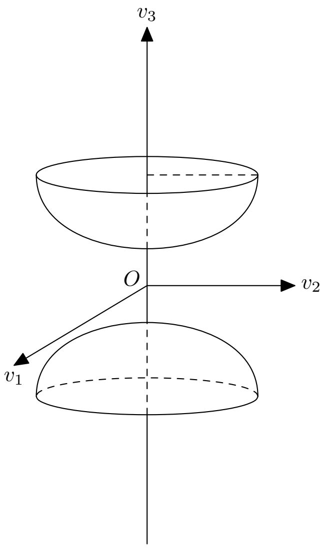
Hyperboloïde à une nappe : $\frac{x^2}{A^2} + \frac{y^2}{B^2} - \frac{z^2}{C^2} = 1$

Figure 7



Cône : $\frac{x^2}{A^2} + \frac{y^2}{B^2} - \frac{z^2}{C^2} = 0$

Figure 8



Hyperboloïde à deux nappes :
$\frac{x^2}{A^2} + \frac{y^2}{B^2} - \frac{z^2}{C^2} = -1$

Figure 9

b) Si $t \neq 0$, c'est-à-dire si $v_3 \notin \operatorname{Ker}\varphi$, en effectuant une nouvelle transformation affine : $x = x$, $y = y$, $z = h + 2\,t\,Z'$, l'équation de $\mathcal{Q}$ s'écrit

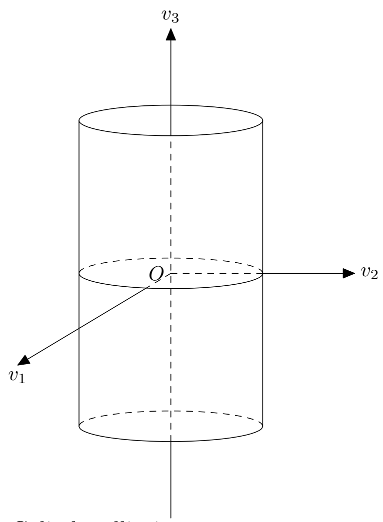
$$a\,x^2 + b\,y^2 = z \tag{8}$$

i) si sign (q) = (2,0) (c'est-à-dire si $a > 0$ et $b > 0$), en étudiant les sections par des plans d'équation $z = K$ on voit facilement que $\mathcal{Q}$ a la forme décrite en fig. 11 : il s'agit d'un ***paraboloïde elliptique*** (les sections par des plans orthogonaux aux directions principales sont soit des paraboles, soit des ellipses).

ii) si sign (q) = (1,1) , en supposant que $a > 0$ et $b < 0$ [2], l'équation (8) s'écrit :

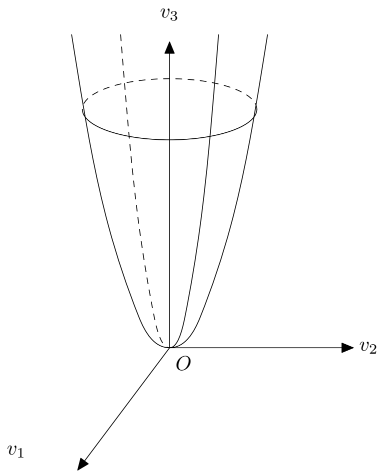
$$a\,x^2 - |\,b\,|\,y^2 = z \tag{8'}$$

Les sections par des plans orthogonaux à v_1 ou à v_2 sont des paraboles et les sections par des plans orthogonaux à v_3 sont des hyperboles : $\mathcal{Q}$ est un ***paraboloïde hyperbolique*** (cf. fig. 12).



Cylindre elliptique :
$\frac{x^2}{A^2} + \frac{y^2}{B^2} = 1$

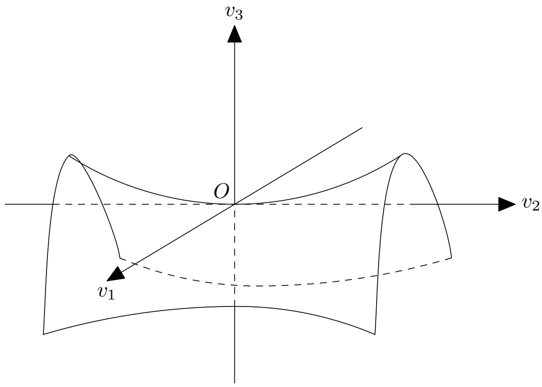
Figure 10



Paraboloïde elliptique :
$\frac{x^2}{A^2} + \frac{y^2}{B^2} = z$

Figure 11

[2]Le cas $a < 0$ et $b > 0$ se traite de la même manière.



Paraboloïde hyperbolique : $\frac{x^2}{A^2} - \frac{y^2}{B^2} = z$

Figure 12

3. rang q = 1 - On peut supposer, par exemple, que $a \neq 0$ et $b = c = 0$. L'équation (6) s'écrit alors :

$$a\,x^2 - 2\,s\,Y' - 2\,t\,Z' = h. \tag{9}$$

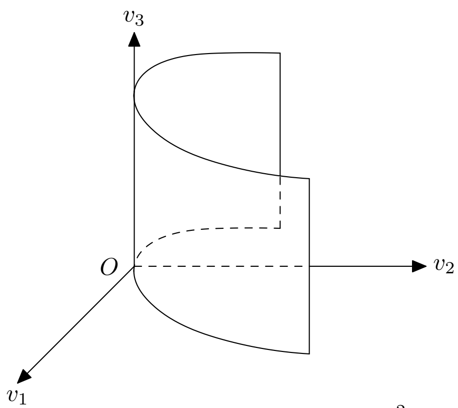
`a)` $s \neq 0$ `(ou` $t \neq 0$`).` Supposons, par exemple $s \neq 0$. En effectuant le changement de variables $x = x$, $y = 2\,s\,Y' + 2\,t\,Z'$, $z = Z'$ l'équation s'écrit

$$a\,x^2 - y = h \tag{10}$$

Il s'agit d'un ***cylindre parabolique*** : les sections par des plans orthogonaux à l'axe, qui est dirigé par v_3, sont des paraboles (cf. fig. 13).

Le cas où $t \neq 0$ se traite d'une manière analogue, par le changement de variables $x = x$, $y = Y'$, $z = 2\,s\,Y' + 2\,t\,Z'$.

`b)` $s = t = 0$. Dans ce cas l'équation (9) se réduit à $a\,x^2 = h$. $\mathcal{Q}$ dégénère alors deux plans parallèles (si $h > 0$) (cf. fig. 14) en un plan (si $h = 0$) et $\mathcal{Q} = \emptyset$ si $h < 0$.



Cylindre parabolique : $y = ax^2$

Figure 13

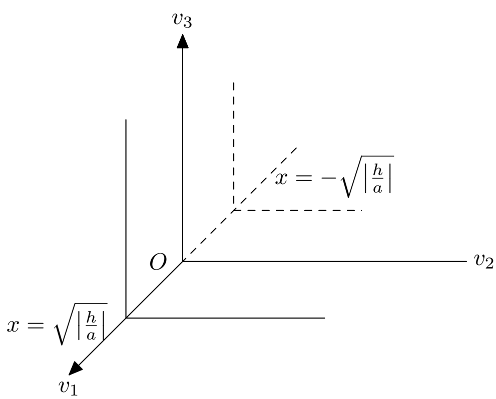


Figure 14

Remarquons que la condition pour qu'une direction principale appartienne à $\operatorname{Ker}\varphi$ est une condition "ouverte" (c'est-à-dire stable par petites variations des cœfficients de $\mathcal{Q}$) alors que la condition pour qu'elle appartienne à $\operatorname{Ker}\varphi$ est "fermée" (instable par petites variations des cœfficients). On dira dans le premier cas que la condition est *générique* et dans le second cas qu'elle est *exceptionnelle*[3].
En résumant, nous avons :

Théorème A.11.4 – *Soit $\mathcal{Q}$ une quadrique définie par l'équation $q(v) + \varphi(v) = k$. On suppose $\mathcal{Q} \neq \emptyset$ et $\mathcal{Q}$ non réduite à un point.*

1. rang q =3 *:*

   (a) si $\operatorname{sign} q = (3,0)$, *$\mathcal{Q}$ est un* ***ellipsoïde.***

   (b) si $\operatorname{sign} q = (2,1)$, *$\mathcal{Q}$ est* - *un* ***hyperboloïde à une nappe*** $(h > 0)$
   - *un* ***cône*** $(h = 0)$
   - *un* ***hyperboloïde à deux nappes*** $(h < 0)$.

2. rang q = 2 *:*

   (a) génériquement (c'est-à-dire si la direction principale qui est isotrope n'est pas contenue dans $\operatorname{Ker}\varphi$), *$\mathcal{Q}$ est un* ***paraboloïde elliptique****, si* sign q = (2,0), *ou un* ***paraboloïde hyperbolique****, si* sign (q) = (1,1).

   (b) exceptionnellement (c'est-à-dire si la direction principale isotrope est contenue dans $\operatorname{Ker}\varphi$), *$\mathcal{Q}$ est un cylindre dont les sections par des plans orthogonaux à l'axe sont des coniques de signature (2,0) ou (1,1) :* ***cylindre elliptique*** *ou* ***cylindre hyperbolique*** *pouvant éventuellement dégénérer en deux plans non parallèles.*

3. rang q = 1 *:*

   (a) génériquement (c'est-à-dire si l'une des deux directions principales isotropes n'est pas contenue dans $\operatorname{Ker}\varphi$), *$\mathcal{Q}$ est un* ***cylindre parabolique***

   (b) exceptionnellement (c'est-à-dire si les deux directions principales isotropes sont contenues dans $\operatorname{Ker}\varphi$), *$\mathcal{Q}$ dégénère en* ***deux plans parallèles****, éventuellement confondus.*

[3]Une définition précise de la notion de généricité dépasse le cadre de ce cours. On peut se contenter ici de cette notion intuitive)

Appendice A.12

Portrait de phase d'un système autonome

Systèmes autonomes

Considérons un système d'équations différentielles

$$\begin{cases} \dfrac{\mathrm{d}x_1}{\mathrm{d}t} & = X_1\,(t, x_1, \ldots, x_n) \\ \quad\vdots & \\ \dfrac{\mathrm{d}x_n}{\mathrm{d}t} & = X_n\,(t, x_1, \ldots x_n) \end{cases} \tag{1}$$

ou, sous forme abrégée :

$$\dot{x} = X(t, x) \tag{1'}$$

avec $x = (x_1, \ldots, x_n)$, $\dot{x} = \frac{\mathrm{d}x}{\mathrm{d}t}$ et X est un «champ de vecteurs», c'est-à-dire une application d'un ouvert $\Gamma \subset \mathbb{R} \times \mathbb{R}^n$ dans $\mathbb{R}^n$:

$$X \;:\; \Gamma \subset \underset{(t,x_1,\ldots,x_n)}{\mathbb{R} \times \mathbb{R}^n} \;\underset{\longmapsto}{\to}\; \underset{(X^1,\ldots,X^n)}{\mathbb{R}^n}$$

que l'on supposera de classe $\mathcal{C}^1$.

Si les fonctions X_i ne dépendent pas de t, le système est dit ***autonome***. Un système autonome est donc de la forme :

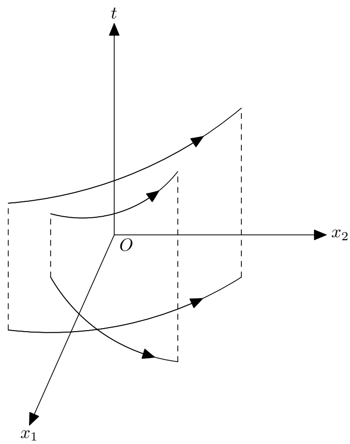
$$\dot{x} = X(x). \tag{2}$$

Le théorème d'existence et unicité affirme que $\forall (t_0, x_0) \in \Gamma$ il existe une et une seule solution "maximale" (c'est-à-dire non prolongeable) du système (1), $x = \varphi(t)$, telle que $\varphi(t_0) = x_0$.

Une solution $x = \varphi(t)$ peut être représentée géométriquement par une courbe dans l'espace $\mathbb{R} \times \mathbb{R}^n$ des variables $(t, x_1, \cdots, x_n)$. Ces courbes sont dites ***courbes intégrales*** du champ X. Le théorème d'existence et unicité affirme donc que par chaque point de Γ il passe une et une seule courbe intégrale maximale.

Il existe cependant une autre manière de représenter les solutions qui s'avère préférable dans le cas des systèmes autonomes. Si $x = \varphi(t)$ est une solution, lorsque t varie le point x décrit une courbe de l'espace $\mathbb{R}^n$ de coordonnées $(x_1, \cdots, x_n)$. Cet espace est dit ***espace des phases*** et la courbe décrite ***trajectoire*** ou ***orbite*** de la solution. L'espace $\mathbb{R} \times \mathbb{R}^n$ est dit espace des phases élargi.

Sur chaque trajectoire on peut préciser le sens de parcours : par convention, c'est le sens dans lequel le point décrit la trajectoire lorsque t augmente. Il est clair que les trajectoires sont les projections sur l'espace des phases des courbes intégrales tracées dans l'espace $\mathbb{R}^n \times \mathbb{R}$.
Notons que, alors que les courbes intégrales dans l'espace des phases élargi ne peuvent se rencontrer, d'après le théorème d'existence et d'unicité, il n'en est pas de même en général pour les trajectoires dans l'espace des phases (cf. fig. 1).
Cependant, lorsque le système est autonome, les trajectoires dans l'espace des phases ne se rencontrent pas. On peut démontrer en effet le théorème suivant :



$\{(x_1, x_2, t)\}$: espace des phases élargi
$\{(x_1, x_2)\}$: espace des phases
Figure 1

Théorème A.12.1 – *Soit $x = X(x)$ un système autonome, où X est un champ de vecteurs de classe $\mathcal{C}^1$ sur un ouvert Γ de l'espace des phases.*
Pour chaque point de Γ il passe une et une seule trajectoire. De plus, si une trajectoire se recoupe, alors elle est nécessairement une orbite fermée représentant une solution périodique. Plus précisément, il n'y a que trois types de trajectoires dans l'espace des phases pour un système autonome.

1. *Les trajectoires réduites à un point, dites «positions d'équilibre» qui correspondent aux solutions constantes : $x(t) = a \quad (a \in \mathbb{R}^n)$.*
2. *Les trajectoires fermées qui correspondent aux solutions périodiques.*
3. *Les trajectoires qui ne se recoupent pas.*

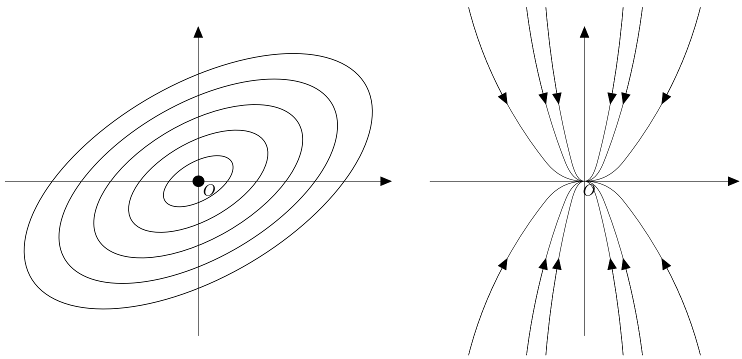
Trajectoires possibles dans l'espace des phases d'un système autonome.

Figure 2

Remarque - Il est clair que la fonction $x = a$ ($a =$ const.) est solution si et seulement si $X(a) = 0$. Donc, les positions d'équilibre s'obtiennent en cherchant les points a tels que $X(a) = 0$. Ces points sont dits ***points singuliers*** du champ X (il ne s'agit pas de singularité au sens de la différentiabilité, car X est toujours supposé $\mathcal{C}^1$).

Un cas particulièrement important est celui où la fonction $X(x)$ est *linéaire* en x : c'est le cas des systèmes linéaires à cœfficients constants que nous avons étudié au chapitre 6 (cf. page 170).

L'importance de tels systèmes vient, entre autres, d'un théorème de Liapounov concernant l'étude de la stabilité . Sans entrer dans les détails de la théorie de la stabilité, on peut dire qu'elle consiste à étudier les orbites des solutions dans le voisinage d'une position d'équilibre. Il s'agit de savoir si les trajectoires qui passent dans un voisinage suffisamment petit d'une position d'équilibre – on les obtient par exemple en faisant varier légèrement les conditions initiales qui caractérisent la position d'équilibre – restent dans ce voisinage (voire si elles tendent vers le point d'équilibre). Si c'est le cas, la position d'équilibre est dite stable (et asymptotiquement stable si les orbites tendent vers la position d'équilibre).



O stable $\qquad$ O asymptotiquement stable

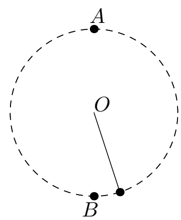
Figure 3

Le problème est important, parce que les mesures physiques n'ont qu'une précision finie, et de ce fait on ne peut jamais être sûr d'imposer les conditions initiales qui donnent comme solution la position d'équilibre. Si la position d'équilibre est stable, cela a peu d'importance ; si elle est instable, des petites variations peuvent entraîner des fortes perturbations du système (cf. fig. 4).

Liapounov a démontré que, sous des conditions assez générales, on peut étudier la stabilité de la position d'équilibre $x = x_0$ en remplaçant le système $\dot{x} = X(x)$ par le système linéaire à cœfficients constants

$$\dot{x} = A\,x$$

où $A = \Big(A_{ij}\Big) \in \mathcal{M}_n\,(\mathbb{R})$ et $A_{ij} = \left(\dfrac{\partial X_i}{\partial x_j}\right)_{x=x_0}$.



Oscillation du pendule

A: position d'équilibre instable

B: position d'équilibre stable

Figure 4

Ces quelques considérations, même si un peu approximatives, permettent de comprendre pourquoi il est important d'étudier le «portrait des phases» – c'est-à-dire l'allure des trajectoires dans l'espace des phases – d'un système linéaire à cœfficients constants.

Trajectoires d'un système linéaire autonome dans le plan

Considérons le système différentiel

$$\begin{cases} \dot{x}_1 &= a\,x_1 + b\,x_2 \\ \dot{x}_2 &= c\,x_1 + d\,x_2 \end{cases} \qquad a, b, c, d \in \mathbb{R}$$

ou, sous forme matricielle, $\dot{x} = A\,x$ avec : $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Nous allons étudier les trajectoires en fonction des valeurs propres de A. Notons que l'origine est toujours une position d'équilibre. Ce sera la seule si et seulement si dét $A \neq 0$.

Etudions d'abord ce cas. On a trois possibilités selon que les valeurs propres sont réelles et distinctes, complexes conjuguées ou confondues.

1^er cas **Les valeurs propres λ_1, λ_2 de A sont réelles et distinctes** (et non nulles car dét $A \neq 0$).

Soient $\vec{v}_1$ et $\vec{v}_2$ les vecteurs propres correspondants. La solution générale s'écrit :

$$x = c_1\,e^{\lambda_1 t}\,\vec{v}_1 + c_2\,e^{\lambda_2 t}\,\vec{v}_2$$

Les composants de x sur la base $\{\vec{v}_1, \vec{v}_2\}$ sont donc :

$$x_1 = c_1\,e^{\lambda_1 t} \quad , \quad x_2 = c_2\,e^{\lambda_2 t} \tag{1}$$

Notons que $\vec{v}_1$ et $\vec{v}_2$ ne sont pas nécessairement orthogonaux. Pour simplifier l'étude nous considérons une transformation linéaire qui amène $\{\vec{v}_1, \vec{v}_2\}$ sur la base canonique $\{e_1, e_2\}$. On tracera les orbites dans le plan $\{e_1, e_2\}$ et on reviendra par la transformation inverse au plan $\{\vec{v}_1, \vec{v}_2\}$.

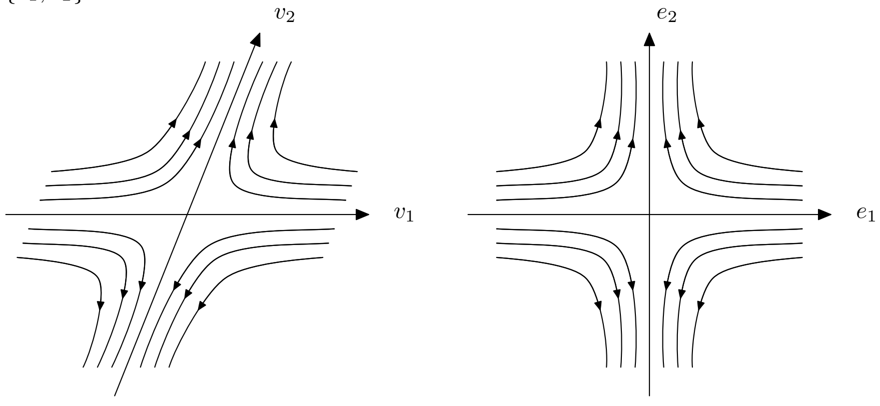


Figure 5

Remarquons que, puisque c_1 et c_2 sont des constantes arbitraires, en plus de la trajectoire (1) on aura aussi les trajectoires obtenues en changeant les signes de c_1 et c_2. Il suffira donc de tracer les trajectoires qui sont dans le premier quadrant (correspondant à $c_1 \geq 0$ et $c_2 \geq 0$) et compléter par des symétries par rapport aux axes. Pour c_1 ou c_2 nuls on obtient les demi-axes (par exemple le demi-axe $O\,x_1$ positif est obtenu pour $c_2 = 0$ et $c_1 > 0$).

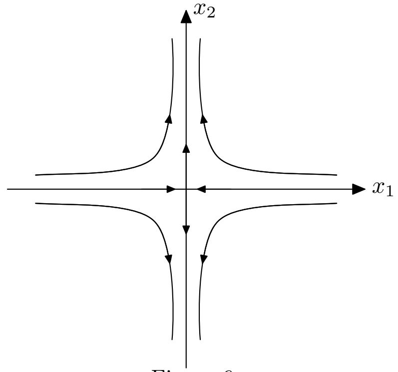


Figure 6

Si $\lambda_1 > 0$, x_1 croît et le point s'éloigne de l'origine ; si $\lambda_1 < 0$ le point se rapproche de l'origine.

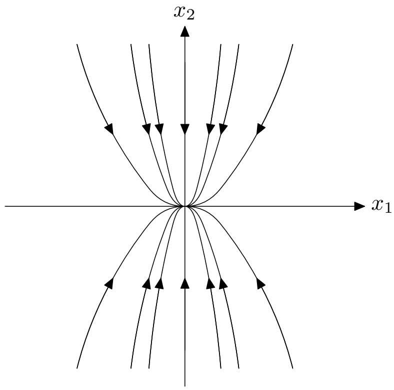
L'allure des trajectoires dépend donc du signe de λ_1 et λ_2.

a) λ_1 **et** λ_2 **sont de même signe.**

Supposons, pour fixer les idées, que $|\lambda_1| < |\lambda_2|$.

i) $\boxed{\lambda_1 < 0\,,\ \lambda_2 < 0}$

Le point tend vers l'origine des coordonnées. Puisque pour $t \to -\infty$ l'ordonnée croît plus vite que l'abscisse, l'axe $O\,x_2$ est direction parabolique. On obtient ce que l'on appelle un ***nœud stable*** (figure 7) (plus précisément : asymptotiquement stable) .

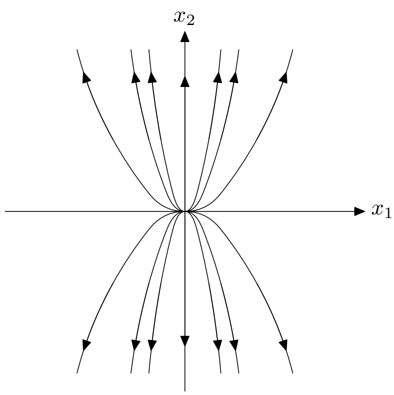


nœud stable

Figure 7

ii) $\boxed{\lambda_1 > 0\,,\ \lambda_2 > 0}$

Les trajectoires sont les mêmes mais le mouvement a lieu dans le sens inverse (on est ramené au cas précédent en changeant t en $-t$). On obtient un ***nœud instable*** (figure 8).



nœud instable

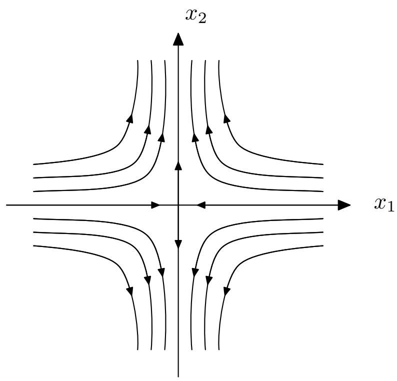
Figure 8

b) λ_1 et λ_2 sont de signe contraire.

Supposons, pour fixer les idées, que $\lambda_1 < 0$ et $\lambda_2 > 0$.

Dans ce cas le mouvement selon l'axe $O\,x_1$ s'effectue vers l'origine et suivant l'axe $O\,x_2$ en s'éloignant de l'origine. On obtient un **col** (figure 9).

Comme ci-dessus si $\lambda_1 > 0$ et $\lambda_2 < 0$, les trajectoires sont les mêmes mais parcourues dans le sens contraire.



col

Figure 9

2^{eme} cas	**Les valeurs propres de A sont des complexes conjugués λ et $\overline{\lambda}$.**

La solution s'écrit :

$$x = c\,e^{\lambda\,t}\,\vec{v}_1 + \overline{c}\,e^{\overline{\lambda}\,t}\,\vec{v}_2.$$

En posant $\vec{v}_2 = \dfrac{1}{2}(\vec{w}_1 + i\,\vec{w}_2)$ $(\vec{w}_1$ et $\vec{w}_2$ sont deux vecteurs réels indépendants), on peut écrire :

$$x = \frac{c\,e^{\lambda t} + \overline{c}\,e^{\overline{\lambda} t}}{2}\,\vec{w}_1 + \frac{c\,e^{\lambda\,t} - \overline{c}e^{\overline{\lambda}\,t}}{2\,i}\,\vec{w}_2.$$

Si on pose $\xi = c\,e^{\lambda\,t} = \xi_1 + i\,\xi_2$, on a :

$$x = \xi_1\,\vec{w}_1 + \xi_2\,\vec{w}_2.$$

On peut donc tracer les trajectoires dans le plan $(\vec{w}_1, \vec{w}_2)$. Comme ci-dessus, on peut se ramener, par une transformation linéaire, à tracer les orbites sur un plan rapporté à des axes orthogonaux.
En coordonnées polaires, la trajectoire est décrite par

$$\xi = c\,e^{\lambda\,t}.$$

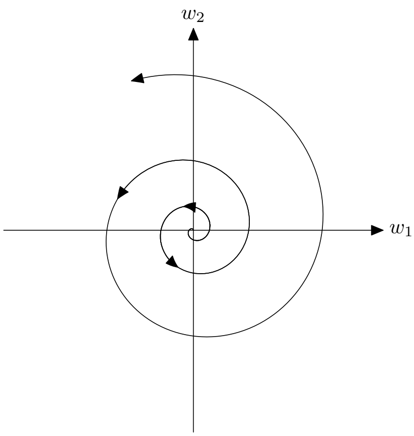
En posant $\xi = \rho\,e^{i\varphi}$, $c = R\,e^{i\,\alpha}$, $\lambda = \mu + i\,\nu$, on a facilement :

$$\begin{cases} \rho &= R\,e^{\mu\,t} \\ \varphi &= \nu\,t + \alpha \end{cases}$$

Si $\mu \neq 0$, lorsque t augmente ρ et φ augmentent ou diminuent, selon le signe de μ et ν. On a des spirales logarithmiques.

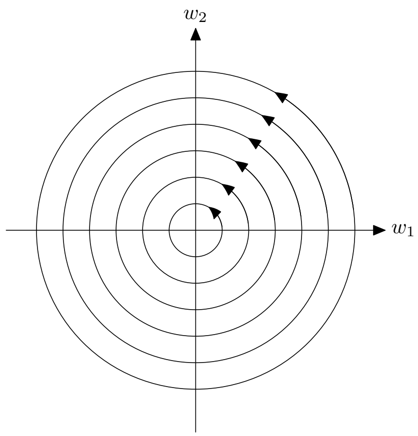
Si $\mu < 0$ le point se rapproche de O (***foyer stable***) ; si $\mu > 0$ il s'en éloigne (***foyer instable***) cf. figure 10.

Si $\mu = 0$ (c'est-à-dire λ est imaginaire pur) on obtient des cercles de centre O ; toute trajectoire est périodique à l'exception de l'origine qui est une position d'équilibre dite ***centre*** (figure 11).



$\mu \neq 0,\ \mu > 0$: foyer instable

Figure 10



$\mu = 0$: centre

Figure 11

$\boxed{3^{\text{eme}}\ \text{cas}}$ **A admet une valeur propre double** λ (non nulle car dét $A \neq 0$).

Deux cas se présentent selon que A est diagonalisable ou non.

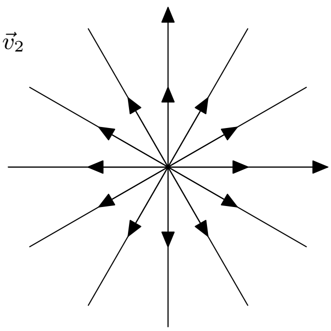
a) A est diagonalisable

Si $\vec{v}_1$ et $\vec{v}_2$ sont les vecteurs propres, la solution est :

$$x = c_1\, e^{\lambda t}\, \vec{v}_1 + c_2\, e^{\lambda t}\, \vec{v}_2$$

d'où $x_1 = c_1\, e^{\lambda t}$, $x_2 = c_2\, e^{\lambda t}$; donc :
– si $\lambda \neq 0$, on a $c_2\, x_1 + c_1\, x_2 = 0$; les trajectoires sont donc des demi-droites issues de l'origine ;

– si $\lambda = 0$, la solution s'écrit $x = x_0$ (x_0 est un vecteur constant arbitraire). Chaque point est donc une position d'équilibre.



$\lambda > 0$: nœud instable

Figure 12

b) A n'est pas diagonalisable

Il existe alors une base $\{\vec{v}_1, \vec{v}_2\}$ telle que :

$$\begin{cases} A\,\vec{v}_1 & = \lambda\,\vec{v}_1 \\ A\,\vec{v}_2 & = \vec{v}_1 + \lambda\,\vec{v}_2 \end{cases}$$

c'est-à-dire : A est semblable à la matrice $\begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}$.

Dans ce cas, la solution s'écrit :

$$x = c_1\, e^{\lambda t}\, \vec{v}_1 + c_2\, e^{\lambda t}\, (\vec{v}_1\, t + \vec{v}_2)$$

d'où les coordonnées de la trajectoire :

$$x_1 = (c_1 + c_2\, t)\, e^{\lambda t}\,, \quad x_2 = c_2\, e^{\lambda t}.$$

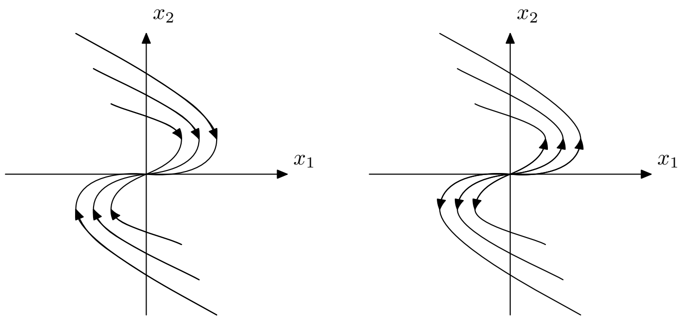
Supposons, pour fixer les idées, que $\lambda < 0$. Si l'on échange simultanément les signes de c_1 et c_2 les trajectoires se transforment en leurs symétriques par rapport à l'origine. Il suffira donc d'étudier le cas où $c_1 \geq 0\,, c_2 \geq 0$ (le cas où c_1 et c_2 sont de signe contraire se traite d'une manière analogue).

Si $c_2 = 0$ on a le demi-axe positif des abscisses.

Si $c_1 = 0, \quad x_1 = c_2\, t\, e^{\lambda t}\,, \quad x_2 = c_2\, e^{\lambda t}$

x_1 et x_2 tendent vers zéro lorsque $t \to +\infty$, x_2 gardant un signe constant positif, x_1 changeant de signe. On obtient un ***nœud dégénéré stable.***

Le cas $\lambda > 0$ se déduit du précédent en changeant t en $-t$: x_1 change de signe et x_2 ne change pas. Les trajectoires se déduisent donc par symétrie par rapport à l'axe $O\,x_2$ et sont décrites en sens inverse.



$\lambda < 0$: nœud dégénéré stable $\qquad$ $\lambda > 0$: nœud dégénéré instable

Figure 13

Si c_1 et c_2 sont nuls, en posant $\tau = c_1 + c_2\, t$, on a :

$$x_1 = \tau\, e^{\frac{\lambda \tau - c_1}{c_2}}\,, \quad x_2 = c_2\, e^{\frac{\lambda \tau - c_1}{c_2}}.$$

On obtient les mêmes trajectoires, mais parcourues avec un certain retard.

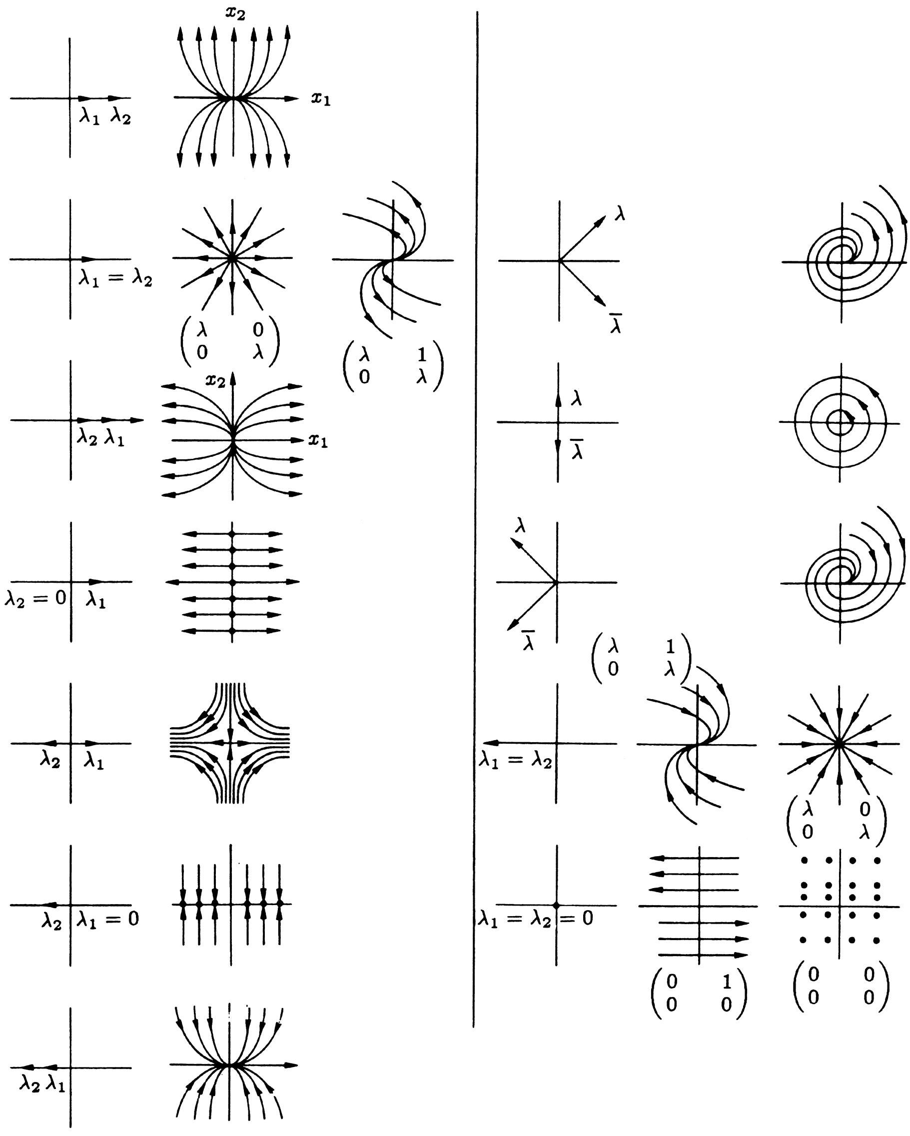
Il reste à étudier le cas où $\det A = 0$, qui est laissé en exercice. Dans le tableau à la page suivante on résume tous les résultats. A gauche sont représentées, dans le plan de coordonnées (α, β) les valeurs propres $\lambda = \alpha + i\,\beta$, et à droite les trajectoires dans le plan des phases de manière à pouvoir voir le changement des trajectoires lorsque les valeurs propres varient. Comme on le voit, il y a un changement de *type* de trajectoire lorsque l'une des valeurs propres traverse l'axe imaginaire (c'est ce qu'on appelle les ***bifurcations***).

La théorie qualitative des équations différentielles est l'une des branches les plus intéressantes des mathématiques contemporaines.
Un très beau livre sur ce sujet, accessible aux étudiants de fin de Licence et Master, est :

M.W. Hirsch – S. Smale

Differential equations, dynamical systems and linear algebra

Academic Press.



Plan des phases d'un système linéaire autonome

Figure 14

Appendice A.13

Formes bilinéaires et sesquilinéaires. Table de correspondance

<table>
<tr>
<td>

FORMES BILINÉAIRES

(E espace vectoriel sur K)

$b\,:\,E\times E\,\to\,K$ telle que :

$b(x+y,z)=b(x,z)+b(y,z)$
$b(x,y+z)=b(x,y)+b(x,z)$
$b(\lambda x,y)=\lambda\,b(x,y)=b(x,\lambda\,y)$

En termes de matrices :
$b(x,y)={}^tX\,B\,Y \quad (B=M(b)_{e_i})$

Changement de base :
$B'={}^tP\,B\,P \quad (P=P_{e_i\to e'_i})$

Noyau :
$N=\{y\in E\,|\,b(x,y)=0\,,\;\forall x\in E\}$

b ***non dégénérée*** $\iff N=\{0\}$

</td>
<td>

FORMES SESQUI-LINÉAIRES

(E espace vectoriel sur $\mathbb{C}$)

$f\,:\,E\times E\,\to\,\mathbb{C}$ telle que :

$f(x+y,z)=f(x,z)+f(y,z)$
$f(x,y+z)=f(x,y)+f(x,z)$
$f(\lambda x,y)=\overline{\lambda}\,f(x,y)\,;\;f(x,\lambda y)=\lambda\,f(x,y)$

En termes de matrices :
$f(x,y)={}^t\overline{X}\,M\,Y \quad (M=M(f)_{e_i})$

Changement de base :
$M'={}^t\,\overline{P}\,M\,P \quad (P=P_{e_i\to e'_i})$

Noyau :
$N=\{y\in E\,|\,f(x,y)=0\;\;\forall x\in E\}$

f ***non dégénérée*** $\iff N=\{0\}$

</td>
</tr>
<tr>
<td>

FORME BILINÉAIRE SYMÉTRIQUE

s bilinéaire et de plus :
$s(x,y)=s(y,z)$

s est représentée dans une base quelconque par une matrice symétrique S :
${}^tS=S$

$q(x)\;:=\;s(x,x)$
$s(x,y)\;=\;\frac{1}{2}(q(x+y)-q(x)-q(y))$

$\mathcal{I}\;:=\;\{x\in E\,|\,q(x)=0\}$

lorsque $K=\mathbb{R}$, q est dite ***définie*** si $\mathcal{I}=\{0\}$

Si E est de dimension finie sur $\mathbb{R}$ il existe une base dans laquelle :
$q(x)=x_1^2+\cdots+x_p^2-x_{p+1}^2-\cdots-x_r^2$

</td>
<td>

FORME HERMITIENNE

h sesquilinéaire et de plus :
$h(x,y)=\overline{h(y,x)}$

h est représentée dans une base quelconque par une matrice hermitienne H :
${}^t\overline{H}=H$

$\tilde{q}(x)\;:=\;h(x,x)\;\in\mathbb{R}$
$h(x,y)\;=\;\frac{1}{4}[\tilde{q}(x+y)-\tilde{q}(x-y)$
$-i\,\tilde{q}(x+i\,y)+i\,\tilde{q}(x-i\,y)]$

$\mathcal{I}\;:=\;\{x\in E\,|\,\tilde{q}(x)=0\}$

q est dite ***définie*** si $\mathcal{I}=\{0\}$

Si E est de dimension finie sur $\mathbb{C}$ il existe une base dans laquelle :
$\tilde{q}(x)=|x_1|^2+\cdots+|x_p|^2-|x_{p+1}|^2-\cdots-|x_r|^2$

</td>
</tr>
</table>

Endomorphisme adjoint

(pour s bilinéaire, symétrique, non dégénérée en dimension finie)

$$s(f(x), y) = s(x, f^*(y))$$

En termes de matrices :
si $s = M(s)_{e_i}$, $A = M(f)_{e_i}$, $A^* = M(f^*)_{e_i}$:

$$A^* = S^{-1}\, {}^tA\, S$$

Si $\{e_i\}$ est ***orthonormée*** : $A^* = {}^tA$

Endomorphisme adjoint

(pour h hermitienne non dégénérée en dimension finie)

$$h(f(x), y) = h(x, f^*(y))$$

En termes de matrices :
si $H = M(h)_{e_i}$, $A = M(f)_{e_i}$, $A^* = M(f^*)_{e_i}$:

$$A^* = H^{-1}\, {}^t\overline{A}\, H$$

Si $\{e_i\}$ est ***orthonormée*** : $A^* = {}^t\overline{A}$

Groupe orthogonal

q forme quadratique *non dégénérée* sur un espace vectoriel de dimension finie sur K :

$$\begin{aligned} f \in O(q) &\iff q(f(x)) = q(x) \quad \forall x \in E \\ &\iff s(f(x), f(y)) = s(x, y) \quad \forall x, y \in E \\ &\iff f^* \circ f = \mathrm{id} \iff f \circ f^* = \mathrm{id} \end{aligned}$$

Si $f \in O(q)$: $\det f = \pm 1$

En termes de matrices :

Si $A = M(f)_{e_i}$, $S = M(s)_{e_i}$

$$f \in O(q) \iff {}^tA\, S\, A = S$$

Dans une base ***orthonormée :***

$$f \in O(q) \iff {}^tA\, A = I$$

Groupe unitaire

h forme hermitienne *non dégénérée* sur un espace vectoriel de dimension finie sur $\mathbb{C}$:

$$\begin{aligned} f \in U(h) &\iff \tilde{q}(f(x)) = \tilde{q}(x) \quad \forall x \in E \\ &\iff h(f(x), f(y)) = h(x, y) \quad \forall x, y \in E \\ &\iff f^* \circ f = \mathrm{id} \iff f \circ f^* = \mathrm{id} \end{aligned}$$

Si $f \in U(h)$: $|\det f| = 1$

En termes de matrices :

Si $A = M(f)_{e_i}$, $H = M(h)_{e_i}$

$$f \in U(h) \iff \overline{{}^tA}\, H\, A = H$$

Dans une base ***orthonormée :***

$$f \in U(h) \iff \overline{{}^tA} A = I$$

Espace euclidien

Espace vectoriel de dimension finie sur $\mathbb{R}$
muni d'un ***produit scalaire***
c'est-à-dire d'une forme bilinéaire symétrique $<,>$ définie positive

Norme :

$$\|x\| := \sqrt{< x, x >}$$

Inégalité de Cauchy-Schwarz :

$$|\, \langle x, y \rangle \,| \leq \|x\|\, \|y\|$$

Espace préhilbertien réel :
Espace vectoriel réel
(éventuellement de dimension infinie)
muni d'un produit scalaire.

Espace hermitien

Espace vectoriel de dimension finie sur $\mathbb{C}$
muni d'un ***produit scalaire hermitien***
c'est-à-dire d'une forme hermitienne
$< \mid >$ définie positive

Norme :

$$\|x\| := \sqrt{< x \mid x >}$$

Inégalité de Cauchy-Schwarz :

$$|\, \langle x \mid y \rangle \,| \leq \|x\|\, \|y\|$$

Espace préhilbertien complexe :
Espace vectoriel complexe
(éventuellement de dimension infinie)
muni d'un produit scalaire hermitien.

<table>
<tr>
<td>
MATRICES ORTHOGONALES
$\mathrm{O}(n,\mathbb{R}) = \{A \in \mathcal{M}_n(\mathbb{R}) \mid {}^tA\,A = I\}$
$A \in \mathrm{O}(n,\mathbb{R}) \implies \text{dét}\, A = \pm 1$
$\mathrm{SO}(n,\mathbb{R}) = \{A \in 0(n,\mathbb{R}) \mid \text{dét}\, A = 1\}$
$\mathrm{O}(n,\mathbb{R}) = \mathrm{U}(n,\mathbb{C}) \cap \mathrm{GL}(n,\mathbb{R})$
$\mathrm{SO}\, n,\mathbb{R}) = \mathrm{SU}(n,\mathbb{C}) \cap \mathrm{GL}(n,\mathbb{R})$
$A \in \mathrm{O}(n,\mathbb{R}) \iff A$ représente dans une base orthonormée une transformation orthogonale d'un espace euclidien de dimension n.
Toute matrice orthogonale est diagonalisable dans $\mathbb{C}$ (non nécessairement dans $\mathbb{R}$)
et ses valeurs propres sont de module 1 .
</td>
<td>
MATRICES UNITAIRES
$\mathrm{U}(n,\mathbb{C}) = \{A \in \mathcal{M}_n(\mathbb{C}) \mid {}^t\overline{A}A = I\}$
$A \in \mathrm{U}(n,\mathbb{C}) \implies |\text{dét}\, A| = 1$
$\mathrm{SU}(n,\mathbb{C}) = \{A \in \mathrm{U}(n,\mathbb{C}) \mid \text{dét}\, A = 1\}$
$A \in \mathrm{U}(n,\mathbb{C}) \iff A$ représente dans une base orthonormée une transformation unitaire d'un espace hermitien de dimension n.
Toute matrice unitaire est diagonalisable (dans $\mathbb{C}$)
et ses valeurs propres sont de module 1 .
</td>
</tr>
<tr>
<td>
ENDOMORPHISME AUTOADJOINT
D'UN ESPACE EUCLIDIEN
f est dit autoadjoint si $f^* = f$
Dans une base orthonormée $\{e_i\}$
si $A = M(f)_{e_i}$
f autoadjoint $\iff {}^tA = A$
(c'est-à-dire A est symétrique)
Tout endormorphisme autoadjoint d'un espace euclidien est diagonalisable dans $\mathbb{R}$ et les espaces propres sont deux à deux orthogonaux. En particulier, on peut construire une base orthonormée de vecteurs propres.
En termes de matrices :
Toute matrice symétrique réelle est diagonalisable dans $\mathbb{R}$ et les espaces propres sont deux à deux orthogonaux pour le produit scalaire canonique de $\mathbb{R}^n$.
c'est-à-dire :
Pour toute matrice symétrique réelle A il existe une matrice orthogonale O telle que $A' ={}^t O\,A\,O$ soit diagonale.
</td>
<td>
ENDOMORPHISME AUTOADJOINT
D'UN ESPACE HERMITIEN
f est dit autoadjoint si $f^* = f$
Dans une base orthonormée
$\{e_i\}$ si $A = M(f)_{e_i}$
f autoadjoint $\iff {}^t\overline{A} = A$
(c'est-à-dire A est hermitienne)
Tout endormorphisme autoadjoint d'un espace hermitien est diagonalisable dans $\mathbb{R}$ et les espaces propres sont deux à deux orthogonaux. En particulier, on peut construire une baseorthonormée de vecteurs propres.
En termes de matrices :
Toute matrice hermitienne est diagonalisable diagonamlisable dans $\mathbb{R}$ et les espaces propres sont deux à deux orthognaux pour le produit scalaire hermitien canonique de $\mathbb{C}^n$.
c'est-à-dire :
Pour toute matrice hermitienne H il existe une matrice unitaire U telle que $H' = \overline{{}^tU}\,H\,U$ soit diagonale réelle.
</td>
</tr>
</table>

Index

Vous pouvez faire part de vos remarques,
critiques, suggestions
aux auteurs à cette adresse :

auteurs@cepadues.com

Achevé d'imprimer en France
en mai 2011 chez Messages SAS
111, rue Nicolas Vauquelin • 31100 Toulouse
Tél. : 05 61 41 24 14 • Fax : 05 61 19 00 43
imprimerie@messages.fr