Algèbre linéaire

Pierre de la Harpe

Section de Mathématiques
Université de Genève, C.P. 240
CH-1211 Genève 24, Suisse
Mel : Pierre.delaHarpe@math.unige.ch

Plan du cours (semestre d'hiver)

I. Introduction.
II. Espaces vectoriels.
III. Applications linéaires.
IV. Groupes symétriques et déterminants.
V. Vecteurs propres et valeurs propres.
VI. Espaces hermitiens et euclidiens, adjoints des opérateurs, théorème spectral.

Il est recommandé aux étudiants d'explorer les nombreux livres existant, par exemple :

S. Lang, Linear algebra (Springer, 1971) ;

S. Axler, Linear algebra done right (Springer, 1996).

Ces notes remplacent des notes analogues rédigées ces dernières années. La cuvée actuelle a largement bénéficié de plusieurs corrections et mises à jour suggérées par Goulnara Arzhantseva, Matthias Franz, Jacob Greenstein et Ghislain Jaudon, ainsi que des compétences TEXniques de Jacob Greenstein pour les figures et la mise en page.

Table des matières

Plan du cours (semestre d'hiver) 3

Chapitre I. Introduction 7
1. Exemples de systèmes linéaires. 7
2. La méthode de Gauss 13
3. Ensembles et applications : terminologie et notations 19
4. Exercices 24

Chapitre II. Espaces vectoriels 31
1. Définitions, premiers exemples, premières propriétés 31
2. Bases 34
3. Dimension d'un espace vectoriel 39
4. Sommes et sommes directes de sous-espaces vectoriels 43
5. Sommes directes d'espaces vectoriels 47
6. Quotients d'espaces vectoriels 48
7. Exercices 49

Chapitre III. Applications linéaires 53
1. Matrices 53
2. Applications linéaires : généralités 56
3. Applications linéaires et bases 59
4. Applications linéaires et matrices 60
5. Applications linéaires et changements de bases 65
6. Noyaux et images des applications linéaires 66
7. Applications linéaires et rang 68
8. Systèmes linéaires 72
9. Dual d'un espace vectoriel 72
10. Exercices 74

Chapitre IV. Groupes symétriques et déterminants 79
1. Groupes : généralités 79
2. Groupes symétriques 80
3. Déterminants : existence et unicité 84
4. Déterminants : calculs 86
5. Déterminants : compléments 91
6. Déterminant d'un endomorphisme 93
7. Résumé de quelques propriétés des déterminants 93
8. Exercices 94

Chapitre V. Vecteurs propres et valeurs propres 97
Prologue 97

1. Sous-espaces invariants, vecteurs propres, valeurs propres 99
2. Polynôme caractéristique 102
3. Endomorphismes linéaires des espaces vectoriels COMPLEXES. Triangulation 106
4. Exercices 108

Chapitre VI. Espaces hermitiens et euclidiens, adjoints des opérateurs, théorème spectral 113
1. Espaces hermitiens 113
2. Espaces euclidiens 121
3. Adjoints, opérateurs auto-adjoints, unitaires et normaux 122
4. Opérateurs linéaires agissant sur les espaces euclidiens 127
5. Le théorème spectral 127
6. Cas euclidien 129
7. Classification des endomorphismes linéaires d'un espace réel de dimension deux 134
8. Un exemple historique de calcul de valeurs propres et vecteurs propres pour une matrice symétrique réelle 135
9. Exercices 138

CHAPITRE I

Introduction

1. Exemples de systèmes linéaires.

Le programme du semestre d'hiver est centré sur les équations linéaires et les applications linéaires, avec un double objectif de *calcul* et de *compréhension géométrique.*

Un *système linéaire à m équations et n inconnues*, où m et n sont des entiers positifs, consiste en des équations de la forme

$$(*) \qquad \begin{array}{rcl} a_{1,1}x_1 + a_{1,2}x_2 + \cdots + a_{1,n}x_n &=& b_1 \\ a_{2,1}x_1 + a_{2,2}x_2 + \cdots + a_{2,n}x_n &=& b_2 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots & & \\ a_{m,1}x_1 + a_{m,2}x_2 + \cdots + a_{m,n}x_n &=& b_m \end{array}$$

où les $a_{i,j}$ $(1 \leq i \leq m,\ 1 \leq j \leq n)$ et les b_k $(1 \leq k \leq m)$ sont des nombres donnés (pour l'instant des nombres réels[1]).

Résoudre le système $(*)$, c'est trouver *tous* les n-uples de nombres réels qui donnent lieu à des égalités dans $(*)$ lorsqu'on les substitue aux «inconnues» $x_1, \ldots, x_n$.

Exemple 1. Le système à 2 équations et 2 inconnues

$$\begin{aligned} x_1 + x_2 &= 0 \\ x_1 - x_2 &= -2 \end{aligned}$$

est représenté géométriquement par deux droites dans le plan de coordonnées (x_1, x_2) (voir Fig. 1).

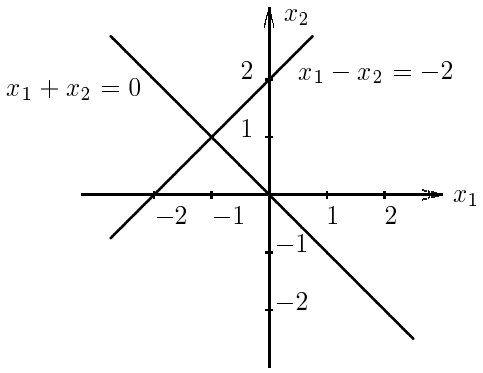

Fig. 1

[1] Il sera plus tard avantageux de permettre à ces «nombres» d'être dans un autre corps : celui des nombre complexes, celui des nombres rationnels, le corps à deux éléments, . . ., le corps à 128 éléments (voir la fin du chapitre VII pour l'intérêt de 128), etc.

Il y a une unique solution correspondant au point d'intersection des deux droites : $x_1 = -1$ et $x_2 = 1$.

EXEMPLE 2. Le système à 2 équations et 3 inconnues

$$\begin{aligned} x_1 + x_2 + x_3 &= 3 \\ x_3 &= 1 \end{aligned}$$

est représenté géométriquement par deux plans dans l'espace de coordonnées (x_1, x_2, x_3) (voir Fig. 2).

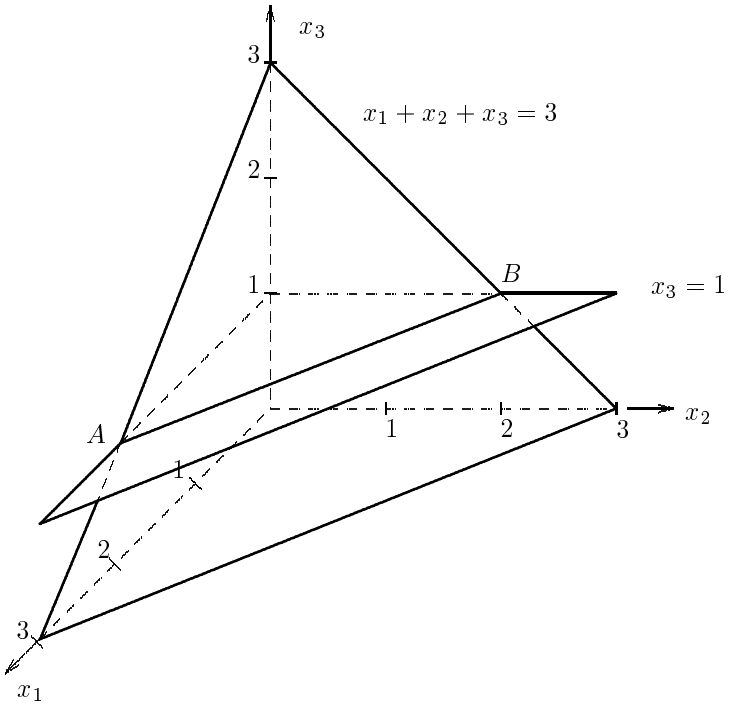

FIG. 2

Il y a une infinité de solutions correspondant à la droite d'intersection des deux plans, droite passant par les points A et B ; on peut paramétrer ces solutions par un nombre réel t :

$$x_1 = 2 - t \qquad x_2 = t \qquad x_3 = 1.$$

EXEMPLE 3. Le système à 3 équations et 2 inconnues

$$\begin{aligned} -x_1 + x_2 &= 1 \\ 2x_1 - x_2 &= 0 \\ x_1 + x_2 &= 1 \end{aligned}$$

est représenté géométriquement par trois droites dans le plan de coordonnées (x_1, x_2) (voir Fig. 3).

Ces trois droites n'ont aucun point commun, et le système *n'a aucune solution.*

EXEMPLE 4. Le système à 3 équations et 2 inconnues

$$\begin{aligned} x_1 - x_2 &= 0 \\ x_1 + 2x_2 &= 6 \\ 2x_1 + x_2 &= 6 \end{aligned}$$

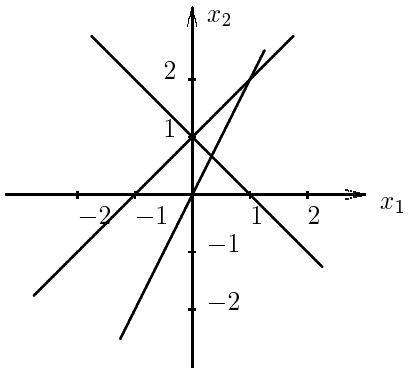

FIG. 3

est représenté géométriquement par trois droites concourrantes dans le plan de coordonnées (x_1, x_2) (voir Fig. 4).

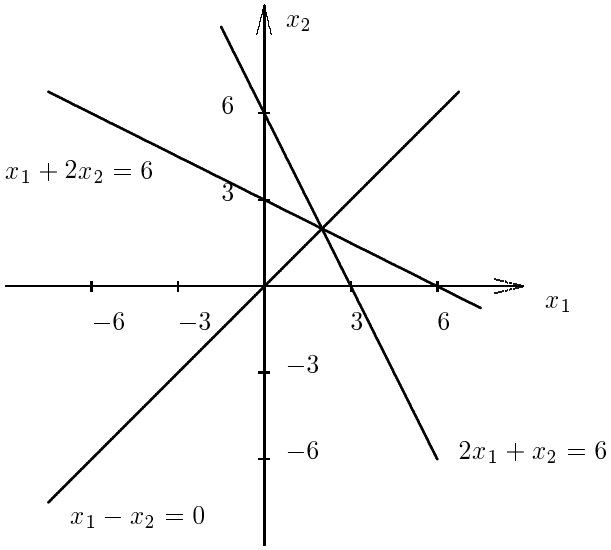

FIG. 4

Il y a une unique solution correspondant au point d'intersection de ces trois droites : $x_1 = 2$ et $x_2 = 2$.

EXEMPLE 5. On se donne trois points du plan, de coordonnées cartésiennes $(1, 7)$, $(6, 2)$ et $(4, 6)$. Quelle est l'équation du cercle passant par ces trois points ?

REMARQUE. Ces points n'étant pas alignés, ils déterminent un unique cercle qui les contient.

Un cercle dans le plan a une équation de la forme

$$x^2 + y^2 + ax + by + c = 0.$$

Dans cet exemple, les *inconnues* sont a, b, c ! Les conditions pour qu'un tel cercle passe par les points indiqués s'écrivent

$$\begin{array}{rcrcrcrcl} 1 & + & 49 & + & a & + & 7b & + & c & = & 0 \\ 36 & + & 4 & + & 6a & + & 2b & + & c & = & 0 \\ 16 & + & 36 & + & 4a & + & 6b & + & c & = & 0. \end{array}$$

On montre que ce système possède une unique solution : $a = -2$, $b = -4$, $c = -20$. Le cercle cherché a donc pour équation $x^2 - 2x + y^2 - 4y - 20 = 0$, ou encore

$$(x-1)^2 + (y-2)^2 = 5^2.$$

C'est donc le cercle de rayon 5 centré en $(1, 2)$.

Remarque. Le problème analogue posé avec trois points alignés, pour fixer les idées avec $(-1, 0)$, $(0, 0)$ et $(1, 0)$, conduit à un système

$$\begin{array}{rcrcrcl} 1 & + & a & + & c & = & 0 \\ & & & & c & = & 0 \\ 1 & - & a & + & c & = & 0 \end{array}$$

qui n'a *aucune solution.*

Exemple 6. Les systèmes

$$\begin{cases} x + 2y^2 = 5 \\ 3x + 4y = 0 \end{cases} \qquad \text{et} \qquad \begin{cases} x + yz = 0 \\ x - y = 0 \\ y + z = 1 \end{cases}$$

ne sont pas linéaires.

De nombreux problèmes *physiques* ou *industriels* donnent lieu à des systèmes linéaires.

Exemple 7. On considère un réseau électrique contenant deux sources (des piles, ou des centrales électriques) et trois utilisateurs schématisés ici par des résistances électriques (voir Fig. 5).

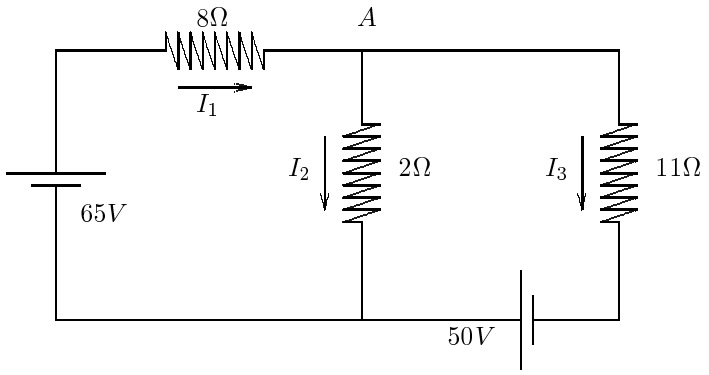

Fig. 5

On cherche les courants I_1, I_2, I_3 à disposition des utilisateurs. On obtient un système linéaire

$$(*) \qquad \begin{array}{rcrcrcl} I_1 & - & I_2 & - & I_3 & = & 0 \\ 8I_1 & + & 2I_2 & & & = & 65 \\ & - & 2I_2 & + & 11I_3 & = & 50 \end{array}$$

où la première équation exprime la conservation de la charge au point A, et les deux autres équations expriment la loi d'Ohm dans le circuit de gauche et celui de droite.

La solution de ce système est : $I_1 = 7,5$, $I_2 = 2,5$ et $I_3 = 5$ (exercice).

Le réseau suisse à haute tension correspond à un système linéaire du même type, avec sans doute n de l'ordre de 10^3 inconnues et équations.

Exemple 8. On cherche la distribution de température à l'équilibre dans une plaque rectangulaire métallique (= conductrice de la chaleur) lorsque la température au bord est donnée, et constante au cours du temps.

Envisageons un modèle de grille (Fig. 6).

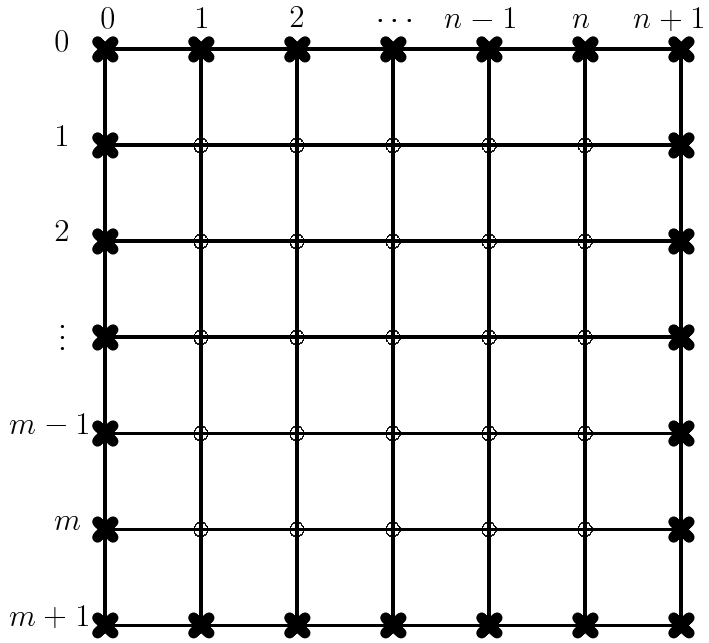

Fig. 6

Les températures des points de croisement sont notées $T_{i,j}$, avec $0 \leq i \leq m+1$ et $0 \leq j \leq n+1$. On suppose ces températures connues au bord de la grille, c'est-à-dire lorsque i est égal à 0 ou $m+1$, et lorsque j est égal à 0 ou $n+1$. Il reste donc mn températures à déterminer, qui sont les $T_{i,j}$, avec $1 \leq i \leq m$ et $1 \leq j \leq n$. Les équations à résoudre sont de la forme

$$T_{i,j} = \frac{1}{4}\left(T_{i-1,j} + T_{i+1,j} + T_{i,j-1} + T_{i,j+1}\right)$$

et sont au nombre de mn. Si $2 \leq i \leq m-1$ et $2 \leq j \leq n-1$, les cinq termes $T_{*,*}$ d'une telle équation sont des inconnues. Si l'un au moins des indices i, j prend une valeur

limite (1 ou m pour i, et 1 ou n pour j), certains des $T_{*,*}$ sont des inconnues et d'autres des données (le terme de gauche, $T_{i,j}$, est toujours une inconnue).

Variante : calcul de la température dans un volume dont la température en surface est connue. En recherche industrielle, le nombre d'équations est souvent de l'ordre de $(10^2)^3 = 10^6$ équations.

Exemple de résultat qui sera montré dans la suite du cours, lorsque $m = n$. *Pour le système $(*)$ de la page 7, les conditions suivantes sont équivalentes :*

(i) *Pour tout second membre $(b_1, \dots, b_n)$, le système $(*)$ possède une unique solution.*

(ii) *Pour tout second membre $(b_1, \dots, b_n)$, le système $(*)$ possède au moins une solution.*

(iii) *Pour tout second membre $(b_1, \dots, b_n)$, le système $(*)$ possède au plus une solution.*

(iv) *Le système*

$$\begin{array}{rcl} a_{1,1}x_1 + a_{1,2}x_2 + \cdots + a_{1,n}x_n &=& 0 \\ a_{2,1}x_1 + a_{2,2}x_2 + \cdots + a_{2,n}x_n &=& 0 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots & & \\ a_{n,1}x_1 + a_{n,2}x_2 + \cdots + a_{n,n}x_n &=& 0 \end{array}$$

possède une unique solution, qui est $x_1 = \cdots = x_n = 0$.

(v) *La matrice* $\begin{pmatrix} a_{1,1} & \dots & a_{1,n} \\ \dots\dots & \dots & \dots\dots \\ a_{n,1} & \dots & a_{n,n} \end{pmatrix}$ *est inversible.*

(vi) *Le déterminant* $\det \begin{pmatrix} a_{1,1} & \dots & a_{1,n} \\ \dots\dots & \dots & \dots\dots \\ a_{n,1} & \dots & a_{n,n} \end{pmatrix}$ *est différent de zéro.*

Note pour les lecteurs ayant des notions de calcul matriciel

L'étude du système $(*)$ de la page 7 est essentiellement équivalente à celle de l'application linéaire de $\mathbb{R}^n$ dans $\mathbb{R}^m$ donnée par

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \mapsto \begin{pmatrix} a_{1,1}x_1 + a_{1,2}x_2 + \cdots + a_{1,n}x_n \\ a_{2,1}x_1 + a_{2,2}x_2 + \cdots + a_{2,n}x_n \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{m,1}x_1 + a_{m,2}x_2 + \cdots + a_{m,n}x_n \end{pmatrix} = \begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} & \dots & a_{2,n} \\ \cdots\cdots & \cdots\cdots & \cdots & \cdots\cdots \\ a_{m,1} & a_{m,2} & \dots & a_{m,n} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}.$$

Ces applications interviennent dans de nombreux problèmes. Indiquons-en un, pour lequel $m = n = 2$, qui est lié au nom de Fibonacci et à un modèle grossier décrivant une population de lapines. Pour tout entier $n \geq 0$, on note x_n le nombre de lapines vivant au temps où naissent les individus de la n-ième génération. On convient qu'une lapine naissant au temps n_0 donne naissance à une lapine à chaque unité de temps $n \geq n_0 + 2$, et on néglige les décès. Le lecteur doit se convaincre que l'équation décrivant ce modèle est

$$x_n = x_{n-1} + x_{n-2} \qquad \text{pour } n \geq 2.$$

L'application linéaire correspondante[2] est alors

$$\mathbb{R}^2 \ni \begin{pmatrix} x_{n-1} \\ x_n \end{pmatrix} \longmapsto \begin{pmatrix} x_n \\ x_{n+1} \end{pmatrix} = \begin{pmatrix} x_n \\ x_{n-1} + x_n \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x_{n-1} \\ x_n \end{pmatrix} \in \mathbb{R}^2.$$

Le calcul des itérations fournit

$$\begin{pmatrix} x_{n-1} \\ x_n \end{pmatrix} \longmapsto \begin{pmatrix} x_n \\ x_{n-1} + x_n \end{pmatrix} \longmapsto \begin{pmatrix} x_{n-1} + x_n \\ x_{n-1} + 2x_n \end{pmatrix} \longmapsto \begin{pmatrix} x_{n-1} + 2x_n \\ 2x_{n-1} + 3x_n \end{pmatrix} \longmapsto \text{etc.}$$

[2]Les termes non expliqués ici seront repris dans la suite du cours.

Supposons de plus que $x_0 = 1$ et $x_1 = 1$, de sorte que le début de la suite des x_n est

$$1,\ 1,\ 2,\ 3,\ 5,\ 8,\ 13,\ 21,\ 34,\ 55,\ 89,\ \ldots$$

Voici quelques exemples de résultats concernant ce modèle :

(i) *Le n-ième terme de la suite est*

$$x_n \;=\; \frac{1}{\sqrt{5}}\left(\left(\frac{1+\sqrt{5}}{2}\right)^{n+1} \;-\; \left(\frac{1-\sqrt{5}}{2}\right)^{n+1}\right)$$

pour tout $n \geq 0$.

(ii) *Pour tout $n \geq 1$, les nombres x_{n-1} et x_n sont des entiers premiers entre eux.*

(iii) *Pour tout $n \geq 1$,*

$$\begin{aligned} x_{2n-1} &= x_{n-1}\,(x_{n-2} + x_n) \\ x_{2n} &= (x_{n-1})^2 + (x_n)^2. \end{aligned}$$

Pour (i), voir l'exercice I.4. Pour (ii) et (iii), voir par exemple C. Houzel, *Analyse mathématique* (Belin 1996), page 355 et la table de la page 360.

Pour (i), la formule approchée $x_n \approx \frac{1}{\sqrt{5}}\left(\frac{1+\sqrt{5}}{2}\right)^{n+1}$ est de plus en plus précise quand n devient grand, par exemple précise à 10^{-8} près si $n = 20$.

2. La méthode de Gauss

Deux systèmes linéaires

$$\begin{array}{l} a_{1,1}x_1 + \cdots + a_{1,n}x_n = b_1 \\ a_{2,1}x_1 + \cdots + a_{2,n}x_n = b_2 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{m,1}x_1 + \cdots + a_{m,n}x_n = b_m \end{array} \qquad \text{et} \qquad \begin{array}{l} \alpha_{1,1}x_1 + \cdots + \alpha_{1,\nu}x_\nu = \beta_1 \\ \alpha_{2,1}x_1 + \cdots + \alpha_{2,\nu}x_\nu = \beta_2 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \alpha_{\mu,1}x_1 + \cdots + \alpha_{\mu,\nu}x_\nu = \beta_\mu \end{array}$$

sont dits *équivalents* si

(i) ils ont le même nombre d'équations (c'est-à-dire $m = \mu$) et le même nombre d'inconnues (c'est-à-dire $n = \nu$),

(ii) ils ont les mêmes solutions.

Il résulte de cette définition qu'on transforme un système linéaire en un système linéaire équivalent par chacun des trois *opérations élémentaires* suivantes :

(E) échanger deux lignes,

(M) multiplier une ligne par un nombre *non nul*,

(T) ajouter un multiple d'une ligne à une autre ligne.

La stratégie de Gauss pour résoudre (= trouver toutes les solutions de) un système linéaire consiste à infliger au système une suite d'opérations élémentaires qui le transforment en un système équivalent aux solutions « évidentes » .

Illustrons d'abord la méthode par l'exemple du système

$$\begin{array}{rcrcrcl} & & y & + & z & = & 3 \\ 2x & & & + & 4z & = & 2 \\ 2x & + & 3y & + & 4z & = & 5 \end{array}$$

à $m = 3$ équations et $n = 3$ inconnues. On obtient une suite de systèmes équivalents, d'abord

$$\begin{array}{rcrcrcl} 2x & & & + & 4z & = & 2 \\ & & y & + & z & = & 3 \\ 2x & + & 3y & + & 4z & = & 5 \end{array}$$

par (E), puis

$$\begin{array}{rcrcrcl} x & & & + & 2z & = & 1 \\ & & y & + & z & = & 3 \\ 2x & + & 3y & + & 4z & = & 5 \end{array}$$

par (M), et enfin

$$\begin{array}{rcrcrcl} x & & & + & 2z & = & 1 \\ & & y & + & z & = & 3 \\ & & 3y & & & = & 3 \end{array}$$

par (T). Ce dernier système est «évident» à résoudre, puisque la troisième équation fournit $y = 1$, puis la deuxième $z = 2$, et enfin la première $x = -3$.

Plus généralement, pour résoudre un système linéaire

$$(*) \qquad \begin{array}{rcrcrcrcl} a_{1,1}x_1 & + & a_{1,2}x_2 & + & \cdots & + & a_{1,n}x_n & = & b_1 \\ a_{2,1}x_1 & + & a_{2,2}x_2 & + & \cdots & + & a_{2,n}x_n & = & b_2 \\ \multicolumn{7}{c}{\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots} & & \\ a_{m,1}x_1 & + & a_{m,2}x_2 & + & \cdots & + & a_{m,n}x_n & = & b_m \end{array}$$

on a un *algorithme*. Le voici, formulé en fonction de la *matrice*

$$a = \begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} & \dots & a_{2,n} \\ \cdots\cdots & \cdots\cdots & \cdots & \cdots\cdots \\ a_{m,1} & a_{m,2} & \dots & a_{m,n} \end{pmatrix},$$

autrement dit en fonction du tableau des coefficients du terme de gauche de $(*)$.

(∘) Soit j_1 l'indice[3] de la première colonne non nulle de a.

(E) Par un échange convenable de lignes, on se ramène au cas d'un système tel que $a_{1,j_1} \neq 0$.

(M) En multipliant la première ligne par une constante convenable, on se ramène à un système tel que $a_{1,j_1} = 1$.

(T) En ajoutant à chacune des lignes 2 à m un multiple convenable de la première ligne, on se ramène à un système du type

$$(**) \qquad \begin{array}{rcrcrcrcl} x_{j_1} & + & a'_{1,j_1+1}x_{j_1+1} & + & \cdots & + & a'_{1,n}x_n & = & b'_1 \\ & & a'_{2,j_1+1}x_{j_1+1} & + & \cdots & + & a'_{2,n}x_n & = & b'_2 \\ \multicolumn{7}{c}{\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots} & & \\ & & a'_{m,j_1+1}x_{j_1+1} & + & \cdots & + & a'_{m,n}x_n & = & b'_m. \end{array}$$

[3] Même s'il arrive souvent que $j_1 = 1$, il convient de ne pas oublier cette étape. Ainsi, elle joue un rôle clé dans les exemples 2 et 3 ci-dessous, ainsi que dans l'exercice corrigé en fin de chapitre pour passer de $(*)$ à $(**)$ lorsque $a = 1$.

Dans le système ainsi obtenu, la matrice des coefficients du terme de gauche est de la forme

$$\begin{pmatrix} 0 & 0 & \cdots & 0 & 1 & * & * & \cdots & * \\ & & & & 0 & & & & \\ & & 0 & & \vdots & & & a' & \\ & & & & 0 & & & & \end{pmatrix}$$

où a' désigne maintenant une matrice à $m-1$ lignes et $n-j_1$ colonnes, et où le 0 écrit au niveau de a' désigne une matrice à $m-1$ lignes et j_1-1 colonnes dont tous les coefficients sont nuls (remarque : si $j_1-1=0$, le coefficient en haut à gauche de la matrice est 1).

On répète alors les mêmes opérations pour le système « plus petit »

$$(***)\qquad \begin{array}{l} a'_{2,j_1+1}x_{j_1+1} + \cdots + a'_{2,n}x_n = b'_2 \\ a'_{3,j_1+1}x_{j_1+1} + \cdots + a'_{3,n}x_n = b'_3 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a'_{m,j_1+1}x_{j_1+1} + \cdots + a'_{m,n}x_n = b'_m \end{array}$$

et ainsi de suite.

Après un nombre fini de telles étapes, on obtient un système dont la solution est « évidente » . Les trois exemples qui suivent illustrent l'algorithme. (Le lecteur exigeant trouvera sans doute – à juste titre – notre description trop vague ; une formulation plus satisfaisante viendra en son temps, peut-être dans un cours de programmation.)

Exemple 1. (Système homogène, 4 équations à 3 inconnues). Pour le système

$$\begin{array}{rcrcrcl} 2x & + & y & + & 5z & = & 0 \\ x & - & 3y & + & 6z & = & 0 \\ 3x & + & 5y & + & 4z & = & 0 \\ & & 7y & - & 7z & = & 0 \end{array}$$

on obtient successivement

$$\begin{array}{rcrcrcl} x & - & 3y & + & 6z & = & 0 \\ 2x & + & y & + & 5z & = & 0 \\ 3x & + & 5y & + & 4z & = & 0 \\ & & 7y & - & 7z & = & 0 \end{array}$$

par (E), puis

$$\begin{array}{rcrcrcl} x & - & 3y & + & 6z & = & 0 \\ & & 7y & - & 7z & = & 0 \\ & & 14y & - & 14z & = & 0 \\ & & 7y & - & 7z & = & 0 \end{array}$$

par (T), deux fois, et enfin

$$\begin{array}{rcrcrcl} x & - & 3y & + & 6z & = & 0 \\ & & y & - & z & = & 0 \\ & & & & 0 & = & 0 \\ & & & & 0 & = & 0 \end{array}$$

par (M) et (T). Voici donc la solution générale du système :

z est arbitraire, notons $z=\lambda$,

$y=z=\lambda$,

$x = 3y - 6z = -3\lambda$.

Autrement dit : $\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -3 \\ 1 \\ 1 \end{pmatrix} \lambda$ avec λ arbitraire. [A la relecture et avec des notions introduites dans un chapitre ultérieur, le lecteur constatera que le *rang* du système, c'est-à-dire la différence entre le nombre de variables, 3, et la dimension du « noyau » , 1, est ici 2.]

Exemple 2. (Système homogène, 3 équations à 4 inconnues). Pour le système

$$\begin{array}{rcrcrcrcl} 4x & + & 12y & - & 7z & + & 6t & = & 0 \\ x & + & 3y & - & 2z & + & t & = & 0 \\ 3x & + & 9y & - & 2z & + & 11t & = & 0 \end{array}$$

on obtient successivement

$$\begin{array}{rcrcrcrcl} x & + & 3y & - & 2z & + & t & = & 0 \\ 4x & + & 12y & - & 7z & + & 6t & = & 0 \\ 3x & + & 9y & - & 2z & + & 11t & = & 0 \end{array}$$

puis

$$\begin{array}{rcrcrcrcl} x & + & 3y & - & 2z & + & t & = & 0 \\ & & & & z & + & 2t & = & 0 \\ & & & & 4z & + & 8t & = & 0 \end{array}$$

puis

$$\begin{array}{rcrcrcrcl} x & + & 3y & - & 2z & + & t & = & 0 \\ & & & & z & + & 2t & = & 0 \\ & & & & & & 0 & = & 0 \end{array}$$

et enfin la solution générale :

y et t sont arbitraires, notons $y = \lambda$ et $t = \mu$,

$z = -2t = -2\mu$,

$x = -3y - 5t = -3\lambda - 5\mu$.

Autrement dit : $\begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = \lambda \begin{pmatrix} -3 \\ 1 \\ 0 \\ 0 \end{pmatrix} + \mu \begin{pmatrix} -5 \\ 0 \\ -2 \\ 1 \end{pmatrix}$ avec λ et μ arbitraires. [A la relecture, le lecteur constatera que le rang du système est $4 - 2 = 2$.]

Exemple 3. (Système non homogène, 3 équations à 3 inconnues). Pour le système

$$\begin{array}{rcrcrcl} x & - & 2y & + & 5z & = & 1 \\ 2x & - & 4y & + & 8z & = & 2 \\ -3x & + & 6y & + & 7z & = & 1 \end{array}$$

on obtient

$$\begin{array}{rcrcrcl} x & - & 2y & + & 5z & = & 1 \\ & & & & -2z & = & 0 \\ & & & + & 22z & = & 4 \end{array}$$

et le système n'a pas de solution.

EXEMPLE 4. (Même type que l'exemple précédent). Pour le système

$$\begin{array}{rcrcrcr} x & - & 2y & + & 5z & = & 1 \\ 2x & - & 4y & + & 8z & = & 2 \\ -3x & + & 6y & + & 7z & = & -3 \end{array}$$

on obtient successivement

$$\begin{array}{rcrcrcr} x & - & 2y & + & 5z & = & 1 \\ & & & & -2z & = & 0 \\ & & & & 22z & = & 0 \end{array}$$

et

$$\begin{array}{rcrcrcr} x & - & 2y & + & 5z & = & 1 \\ & & & & z & = & 0 \\ & & & & 0 & = & 0. \end{array}$$

On trouve la solution générale avec y arbitraire, disons $y = \lambda$, et

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + \lambda \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}.$$

Remarque pour la relecture (comparer avec le § III.8) : les deux matrices

$$\begin{pmatrix} 1 & -2 & 5 \\ 2 & -4 & 8 \\ -3 & 6 & 7 \end{pmatrix} \qquad \text{et} \qquad \begin{pmatrix} 1 & -2 & 5 & 1 \\ 2 & -4 & 8 & 2 \\ -3 & 6 & 7 & -3 \end{pmatrix}$$

sont de rang 2.

REMARQUE. Les termes de gauche des exemples 3 et 4 dont identiques ; en pratique, c'est une situation courante que d'avoir a considérer de nombreux systèmes qui ne diffèrent ainsi que par leurs seconds membres.

Notice historique : Carl Friedrich Gauss, né à Brunswick en 1777, mort à Göttingen en 1855. Son importance dans le développement des mathématiques est de tout premier ordre, et ne peut se comparer qu'à bien peu d'autres — peut-être à un tout petit cercle qui comprendrait Archimède, Fermat, Newton, Euler, Riemann, Poincaré, Hilbert, von Neumann, Weil, Grothendieck, Serre, Gromov, ... ?

Il a révolutionné presque tous les domaines des mathématiques : l'arithmétique, la géométrie élémentaire, la géométrie différentielle (en particulier l'étude des surfaces), la théorie des erreurs (méthode des moindres carrés), ... Il a entrevu les notions de corps et de groupe, et ses écrits non publiés de son vivant contiennent d'importantes découvertes « ultérieures » . Ce fut également un physicien remarquable (étude du magnétisme), en particulier un astronome (il fut directeur de l'observatoire de Göttingen), ainsi qu'un calculateur extraordinnaire.

Il semble qu'il savait mener de front les activités les plus diverses, comme le montre l'exemple suivant qui illustre l'interdépendance profonde des recherches dites pures et appliquées : la période où Gauss trouvait certains de ses résultats fondamentaux sur les surfaces est précisément celle où il travaillait sur mandat gouvernemental à des relevés géodésiques dans le royaume de Hannovre.

Pour en savoir plus à propos de celui qu'on appelle souvent « le prince des mathématiciens » , lire par exemple le chapitre XIV de *Les grands mathématiciens* de E.T. Bell (Payot, 1939), ou plusieurs passages du *Mathematics and its history* de J. Stillwell (Springer, 1989).

Gauss avait un « Tagebuch » (Journal intime) dans lequel il notait ses trouvailles. C'est ainsi qu'on sait que sa méthode d'élimination, exposée ci-dessus, date de juin 1798 ; il avait donc 21 ans, et ce fut pour lui une période extraordinairment productrice. Voir[4] le volume X de ses Oeuvres (= Werke, X, page 533).

Ce n'est que bien plus tard qu'il publia des exposés détaillés de la méthode, par exemple dans *Supplementum theoriae combinationis observationum erroribus minimis obnoxiae = Théorie de la combinaison des observations sujette au minimum d'erreurs, supplément*, article paru dans un journal spécialisé de 1828 (Werke, IV, 55–93). Le sujet principal

[4]Merci à Gerhard Wanner pour ces indications bibliographiques.

du mémoire est la méthode des moindres carrés, elle-même motivée par l'utilisation des résultats de mesures astronomiques ou géodésiques.

Ses notations ne sont pas les nôtres ! par exemple, pour les coefficients $a_{i,j}$ introduits ci-dessus ($1 \leq i \leq m$, $1 \leq j \leq n$), Gauss écrit $[ij]$, ou plutôt

$$[aa], [ab], [ac], \ldots, [bb], [bc], \ldots, [cc], \ldots$$

(le cas traité étant symétrique, il lui est inutile d'introduire $[ba]$, $[ca]$, $[cb]$, etc). Gauss désigne ensuite les coefficients du système obtenu après une transformation élémentaire par

$$[aa, 1], [ab, 1], \ldots, [bb, 1], \ldots \ ;$$

après deux transformations élémentaires, il utilise $[aa, 2]$, $[ab, 2], \ldots$, $[bb, 2], \ldots$, et cetera.

Le premier exemple numérique choisi par Gauss pour illustrer sa méthode est un système de 13 équations à 13 inconnues, notées A, B, $\ldots$, N, tiré de mesures d'angles faites en Hollande par un nommé Krayenhof. Nous ne recopions ici que la première de ces 13 équations,

$$-2,197 \ = \ 5A + C + D + E + H + I + 5,917\, N,$$

et la valeur de la dernière inconnue,

$$N \ = \ 0,119\,681$$

(rappel : les calculettes n'existaient pas encore).

Liste sommaire de quelques contemporains de Gauss (1777–1855)

Adrien-Marie Legendre, 1752–1833

Wolfgang Amadeus Mozart, 1756–1791

Joseph Fourier 1768–1830

François René de Chateaubriand, 1768–1848

Georges Cuvier, 1769–1832

Ludwig van Beethoven, 1770–1827

Georg Wilhelm Friedrich Hegel, 1770–1831

Robert Brown, 1773–1858

Louis-Joseph Gay-Lussac, 1778–1850

Stendhal, 1783–1842

Augustin-Louis Cauchy, 1789–1857

Michael Faraday, 1791–1867

Franz Schubert, 1797–1828

Eugène Delacroix, 1798–1863

Niels Henrik Abel, 1802–1829

Alexandre Dumas, 1802–1870

Victor Hugo, 1802–1885

Carl Gustav Jacobi, 1804–1851

Gustav Lejeune Dirichlet, 1805–1859

Frédéric Chopin, 1810–1849

Evariste Galois, 1811–1832

Ferdinand Eisenstein, 1823–1852

Dates de quelques événements historiques

1778 Déclaration d'indépendance des Etats-Unis d'Amériques.

1789 Début de la Révolution française : réunion des Etats généraux et prise de la Bastille.

1793 Mise à mort de Louis XVI.

1798–1814 Genève fait partie de la France.

1799 Coup d'état du 18 brumaire : Bonaparte 1er Consul.

1814 Première abdication de Napoléon.

1815 Traité de Vienne. En Suisse : pacte fédéral.

1830 Révolutions libérales en France.

1847 Publication du Manifeste communiste de Marx et Engels.

1848 Révolutions en Europe, 2e république en France, Constitution fédérale suisse.

3. Ensembles et applications : terminologie et notations

Il s'agit d'indications extrêmement *sommaires*! Il y aura des compléments dans la suite du cours, aux exercices, au cours d'analyse I, ..., etc.

La théorie naïve des ensembles utilisée ici repose sur un *usage* fréquent, et non sur des définitions formelles. Notons toutefois qu'il serait *absolument illusoire* de «définir» un ensemble comme une «collection d'objets» ... !!!

Quelques exemples standard d'ensembles

L'ensemble $\{0, 1, 2, \ldots, 8, 9\}$ des chiffres de la numérotation usuelle.

L'ensemble $\mathbb{N} = \{0, 1, 2, \ldots\}$ des entiers naturels[5].

L'ensemble $\mathbb{Z} = \{\ldots, -2, -1, 0, 1, 2, 3, \ldots\}$ des entiers rationnels[6].

L'ensemble $\mathbb{Q}$ des nombres rationnels, contenant par exemple -2, $\frac{3}{4}$, $11\frac{1}{2}$.

L'ensemble $\mathbb{R}$ des nombre réels et l'ensemble $\mathbb{C}$ des nombres complexes.

Produit de deux ensembles X et Y

C'est l'ensemble noté $X \times Y$ des paires (x, y) avec l'élément x dans l'ensemble X, et de même y dans Y. L'exemple standard est $\mathbb{R}^2$, qui est l'ensemble des paires (x, y) de nombres réels, et auquel *il faut penser comme à un plan.*

Généralisations immédiates : $\mathbb{R}^3$, $\mathbb{R}^4$, ..., $\mathbb{R}^n$.

---

[5] In english books, $\mathbb{N}$ denotes most often the set $\{1, 2, \ldots\}$. Observe that «positive» is the translation of the french «strictement positif», in symbol > 0, and that «non-negative» is the translation of the french «positif», in symbol ≥ 0.

[6] Voici à ce sujet une citation extraite du *Cours d'algèbre* de Roger Godement (Hermann 1963), livre qui se veut être «un ouvrage de référence accessible au débutant». «Indiquons en passant que les entiers de signe quelconque (que nous appelons, comme le font tous les mathématiciens, des *entiers rationnels*) sont généralement appelés, dans l'enseignement secondaire français, des «entiers algébriques» (de même qu'on y appelle «nombres algébriques» les nombres de signe quelconque, entiers ou non, que nous appelons, à nouveau commes tous les mathématiciens, des *nombres réels*. Ces divergences de terminologie n'auraient aucune importance si les mathématiciens n'utilisaient déjà dans un tout autre sens les expressions «entier algébrique» et «nombre algébrique» que l'on définira plus loin.»

Piste pour une formalisation (ne faisant pas l'objet de ce cours)

Il faut d'abord s'entendre sur l'ensemble $\mathbb{N}$, qui est muni d'une structure grâce aux opérations usuelles (addition, comparaison, ...), régies par certains axiomes, puis définir successivement $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$, $\mathbb{R}^n$ pour tout entier $n \geq 1$, des ensembles de fonctions, etc.

Démonstration par récurrence

Pour démontrer une suite d'assertions $A(n)$, où n est dans $\mathbb{N}$, la méthode dite «par récurrence» consiste à démontrer d'abord l'assertion pour $n = 0$ [ou peut-être pour $n = 1$], puis à montrer que, si l'assertion $A(n)$ est vraie, alors $A(n+1)$ l'est aussi.

EXEMPLE. Montrons que $1 + 2 + \cdots + n = \dfrac{n(n+1)}{2}$ pour tout entier $n \geq 1$.

La vérification de l'assertion pour $n = 1$ est immédiate : $1 = \frac{1\times 2}{2}$. L'argument de récurrence consiste à supposer que l'assertion est vraie pour n, puis à vérifier que

$$1 + 2 + \cdots + n + (n+1) = \frac{n(n+1)}{2} + (n+1) = \frac{(n+1)(n+2)}{2}$$

(la première égalité étant vraie par hypothèse de récurrence), ce qui montre bien que l'assertion est également vraie pour $n + 1$.

Une variante (parmi d'autres) : pour démontrer une suite d'assertions $B(n)$, où n est dans $\mathbb{N}$, la variante consiste à démontrer d'abord l'assertion pour $n = 0$ *et* pour $n = 1$, puis à montrer que, si les assertions $B(n-1)$ *et* $B(n)$ sont vraies, alors $B(n+1)$ l'est aussi. (On peut se ramener au cas précédent en définissant une assertion $A(n)$ comme suit : «les assertions $B(k)$ sont vraies pour tout $k \in \{0, 1, \ldots, n\}$» .)

Signes d'appartenance

Elément *dans* un ensemble : par exemple $3 \in \mathbb{N}$ et $\pi \in \mathbb{R}$. On écrit le plus souvent $\pi \in \mathbb{R}$ mais parfois aussi $\mathbb{R} \ni \pi$.

Sous-ensemble *dans* un ensemble : par exemple $\mathbb{N} \subset \mathbb{Z}$ et $\mathbb{Q} \subset \mathbb{R}$.

Négation : $\mathbb{Z} \subset \mathbb{R}$ et $\frac{3}{5} \in \mathbb{R}$ mais $\frac{3}{5} \notin \mathbb{Z}$.

De même[7] : $[0,1] \subset \mathbb{R}$ et $\mathbb{Z} \subset \mathbb{R}$ mais $[0,1] \not\subset \mathbb{Z}$.

Définition d'un sous-ensemble par une ou plusieurs propriétés

Par exemple, $\{x \in \mathbb{R} \mid x \geq 0\}$ désigne l'ensemble des nombres réels positifs.

$\{n \in \mathbb{Z} \mid \text{il existe } a, b \in \mathbb{Z} \text{ tels que } n = a^2 + b^2\}$ désigne l'ensemble des entiers qui sont sommes de deux carrés d'entiers, c'est-à-dire l'ensemble

$$\{0, 1, 2, 4, 5, 8, 9, 10, 13, 16, 17, 18, 20, 25, \ldots\}\,.$$

$\{(x, y) \in \mathbb{R}^2 \mid y = 0\}$ désigne le premier axe d'un plan décrit par ses coordonnées cartésiennes.

[7] Ici comme souvent ailleurs, $[a, b]$ désigne un intervalle de $\mathbb{R}$, y compris ses extrémités a et b.

Attention : l'usage des différents types de parenthèses : (), [], {}.

Pour désigner une paire d'éléments, on écrit par exemple $(0,1) \in \mathbb{R}^2$, et il faut noter que $(0,1) \neq (1,0)$.

Pour désigner un intervalle de la droite, on utilise les parenthèses carrées : $[0,1] \subset \mathbb{R}$ est l'intervalle unité avec ses extrémités, $]0,1[\subset \mathbb{R}$ est l'intervalle unité sans ses extrémités, $[0,1[\subset \mathbb{R}$ est l'intervalle unité avec l'extrémité de gauche seulement.

Pour désigner un sous-ensemble en indiquant ses éléments, on utilise le troisième type de parenthèses : $\{0,1,2,3\} \subset \mathbb{Z}$; il faut noter que l'ordre n'importe pas : $\{0,1\} = \{1,0\}$.

Intersections et réunions de sous-ensembles

Soient A, B deux sous-ensembles d'un ensemble X.

Réunion : $A \cup B = \{x \in X \mid x \in A \quad \text{ou} \quad x \in B\}$.

Par exemple, $[0,7] \cup [2,10] = [0,10]$.

Intersection : $A \cap B = \{x \in X \mid x \in A \quad \text{et} \quad x \in B\}$.

Par exemple, $[0,7] \cap [2,10] = [2,7]$.

Exemples. Pour un entier $d \geq 0$, désignons par $d\mathbb{Z}$ le sous-ensemble des entiers rationnels qui sont des multiples entiers de d. Alors

$$2\mathbb{Z} \cap 5\mathbb{Z} = 10\mathbb{Z}$$

$$2\mathbb{Z} \cup 5\mathbb{Z} = \{\ldots, -10, -8, -6, -5, -4, -2, 0, 2, 4, 5, 6, 8, 10, 12, 14, 15, 16, 18, \ldots\}\,.$$

Exercice. Vérifier que $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$ et $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$.

Applications

Soient X, Y deux ensembles. Une *application* α de X dans Y fait correspondre à tout élément $x \in X$ un élément $\alpha(x) \in Y$. On écrit dans ce cours[8]

$$\alpha : X \longrightarrow Y$$

et

$$\alpha : X \longrightarrow Y, \quad x \longmapsto \alpha(x) \qquad \text{ou} \qquad \alpha : X \ni x \longmapsto \alpha(x) \in Y$$

(observer les nuances typographiques des débuts de flèches). La donnée de α est équivalente à celle de son *graphe*

$$\{(x,y) \in X \times Y \mid y = \alpha(x)\}$$

qui est un sous-ensemble du produit de X et Y. On définit alors

l'image de α, c'est-à-dire

$$\mathrm{Im}(\alpha) = \{y \in Y \mid \text{il existe } x \in X \text{ tel que } \alpha(x) = y\}$$

l'image d'un sous-ensemble A de X

$$\alpha(A) = \{y \in Y \mid \text{il existe } x \in A \text{ tel que } \alpha(x) = y\}$$

[8] L'usage des flèches dans ce contexte est récent. Il remonte au tout début des anées 1950. Voir B. Eckmann, *The exact sequence,* prépublication, automne 2002.

la préimage d'un sous-ensemble C de Y

$$\alpha^{-1}(C) \;=\; \{x \in X \,|\, \alpha(x) \in C\}$$

la préimage d'un point $y \in Y$

$$\alpha^{-1}(y) \;=\; \{x \in X \,|\, \alpha(x) = y\}$$

Attention : $\alpha^{-1}(y)$ n'est en général pas un point de X ! Considérez l'exemple où $\alpha : \mathbb{R} \longrightarrow \mathbb{R}$ est l'application donnée par $\alpha(t) = t^2$; alors $\alpha^{-1}(-1)$ est l'ensemble vide, noté $\varnothing$, alors que $\alpha^{-1}(1) = \{-1, 1\}$ est un sous-ensemble à deux points de $\mathbb{R}$.

PROPOSITION. *Soient X, Y deux ensembles, $\alpha : X \longrightarrow Y$ une application, A, B des sous-ensembles de X et C, D des sous-ensembles de Y. Alors*

$$\begin{aligned}
\alpha(A \cup B) &= \alpha(A) \cup \alpha(B) \\
\alpha(A \cap B) &\subset \alpha(A) \cap \alpha(B) \qquad\qquad \textit{sic !} \\
\alpha^{-1}(C \cup D) &= \alpha^{-1}(C) \cup \alpha^{-1}(D) \\
\alpha^{-1}(C \cap D) &= \alpha^{-1}(C) \cap \alpha^{-1}(D).
\end{aligned}$$

PREUVE. *Preuve de la première égalité.* Soit $y \in \alpha(A \cup B)$. Il existe $x \in A \cup B$ tel que $y = \alpha(x)$: donc $y \in \alpha(A) \cup \alpha(B)$. Ainsi $\alpha(A \cup B) \subset \alpha(A) \cup \alpha(B)$.

Soit $y \in \alpha(A) \cup \alpha(B)$. S'il existe $x \in A$ tel que $y = \alpha(x)$, alors $y \in \alpha(A) \subset \alpha(A \cup B)$; sinon, il existe $x' \in B$ tel que $y = \alpha(x')$, et de même $y \in \alpha(A \cup B)$. Ainsi $\alpha(A) \cup \alpha(B) \subset \alpha(A \cup B)$.

Preuve des autres égalités : exercice. □

A propos du « sic » . Soit $\alpha : \mathbb{R}^2 \ni (x, y) \longmapsto x \in \mathbb{R}$ la première projection du plan cartésien. Soient $A = \{(x, y) \in \mathbb{R}^2 \mid x - y = 0\}$ et $B = \{(x, y) \in \mathbb{R}^2 \mid x + y = 0\}$ les deux bissectrices. Alors

$$\alpha(A \cap B) \;=\; \{0\} \;\subsetneqq\; \alpha(A) \cap \alpha(B) \;=\; \mathbb{R}.$$

(La notation $\{0\}$ désigne le sous-ensemble de $\mathbb{R}$ réduit à un seul élément, 0.)

Composition d'applications

Soient X, Y, Z trois ensembles et $\alpha : X \longrightarrow Y$, $\beta : Y \longrightarrow Z$ deux applications. Noter que le but de α coïncide avec la source de β. On définit la *composition* de α et β

$$\beta\alpha \;:\; \begin{cases} X & \longrightarrow & Z \\ x & \longmapsto & \beta(\alpha(x)). \end{cases}$$

Ainsi, $\beta\alpha$ est bien la composition dans l'ordre indiqué par le diagramme

$$X \;\xrightarrow{\alpha}\; Y \;\xrightarrow{\beta}\; Z.$$

Noter l'ordre de l'écriture $\beta\alpha$! ; en général, « $\alpha\beta$ » n'a *aucun sens*. Noter aussi que plusieurs auteurs écrivent $\beta \circ \alpha$ au lieu de $\beta\alpha$, avec « $\circ$ » pour la composition.

EXEMPLE. Si

$$\alpha \ : \begin{cases} \mathbb{R} & \longrightarrow \ ]0,\infty[\\ x & \longmapsto \ x^2+1 \end{cases}$$

(où $]0,\infty[$ désigne $\{t \in \mathbb{R} \mid t > 0\}$) et

$$\beta \ : \begin{cases}]0,\infty[& \longrightarrow \ \mathbb{R} \\ x & \longmapsto \ \ln x \end{cases}$$

(où $\ln x$ désigne le logarithme naturel de x), alors

$$\beta\alpha \ : \begin{cases} \mathbb{R} & \longrightarrow \ \mathbb{R} \\ x & \longmapsto \ \ln(x^2+1). \end{cases}$$

Vocabulaire

Une application $\alpha : X \longrightarrow Y$ est dite

injective si $\alpha(x) \neq \alpha(x')$ pour tous $x, x' \in X$ tels que $x \neq x'$,

surjective si, pour tout $y \in Y$, il existe $x \in X$ tel que $\alpha(x) = y$,

bijective si α est à la fois injective et surjective.

EXEMPLES. (i) $\alpha : \mathbb{R} \ni t \longmapsto t^2 \in \mathbb{R}$ n'est ni injective ni surjective,

(ii) $\beta :]0,\infty[\ni t \longmapsto t^2 \in \mathbb{R}$ est injective non surjective,

(iii) $\gamma : \mathbb{R} \ni t \longmapsto t^2 \in [0,\infty[$ est surjective non injective,

(iv) $\delta : [0,\infty[\ni t \longmapsto t^2 \in [0,\infty[$ est bijective.

EXERCICE. Dessiner les graphes de ces quatre applications.

Inverse d'une application bijective

Soit $\alpha : X \longrightarrow Y$ une application bijective. Il existe une application bien définie de Y dans X, qu'on note $\alpha^{-1} : Y \longrightarrow X$, telle que $\alpha^{-1}(y) = x$ si $\alpha(x) = y$.

Attention : Pour une application $\alpha : X \longrightarrow Y$ qui n'est *PAS* bijective, l'écriture α^{-1} s'utilise aussi, mais jamais isolément et *DANS UN AUTRE SENS* : si $C \subset Y$ alors

$$\alpha^{-1}(C) = \{x \in X \mid \alpha(x) \in C\}$$

comme *DÉJÀ* défini plus haut.

Mémorisation de lettres grecques

Les lettres grecques (surtout minuscules) sont d'usage fréquent en mathématiques. En voici la liste, à mémoriser :

α = alpha	μ = mu
β = bêta	ν = nu
γ = gamma	ξ = ksi
δ = delta	π ou ϖ = pi
ϵ = epsilon	ρ = rhô
ζ = dzêta	σ = sigma
η = êta	τ = tau
θ ou ϑ = thêta	ϕ ou φ = phi
ι = iota	χ = khi
κ ou $\varkappa$ = kappa	ψ = psi
λ = lambda	ω = omega

(les lettres omicronn et upsilonn ont été omises à dessein, vu les risques de confusion typographique avec o et v).

4. Exercices

(I.1) Dans le plan $\mathbb{R}^3$ muni de ses coordonnées cartésiennes (x, y, z), dessiner les plans d'équations

$$x + y + z - 1 = 0 \qquad \text{et} \qquad x - 3y = 0,$$

ainsi que l'intersection des deux plans d'équations

$$3x + 2y + z = 6 \qquad \text{et} \qquad z = 1.$$

(I.2) Montrer par récurrence sur n l'identité

$$\sum_{j=1}^{n} j^3 = \frac{n^2(n+1)^2}{4}.$$

(I.3) Il existe une indentité de la forme

$$(*) \qquad \sum_{j=0}^{n} j^2 = an^3 + bn^2 + cn + d$$

valable pour tout entier positif ou nul n.

(i) Deviner les coefficients a, b, c.

(ii) Montrer par récurence sur n ce que devient l'identité $(*)$ quand on y remplace a, b, c par les valeurs trouvées en (i).

[Il y a de nombreuses manières de calculer $\sum_{j=0}^{n} j^2$. Voir le livre de R.L. Graham, D.E. Knuth et O. Patashnik, *Concrete mathematics,* Addison-Wesley 1989. L'index énumère les pages concernées, voir sous « Sums of consecutive squares » .]

(I.4) La *suite de Fibonacci* $(x_n)_{n\geq 0}$ est définie par

$$x_0 = 1\,, \quad x_1 = 1\,, \quad x_n = x_{n-1} + x_{n-2} \quad \text{pour tout } n \geq 2.$$

Montrer par récurrence que

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

pour tout $n \geq 0$.

[*Indication* : Vérifier la formule pour $n = 0$ *et* $n = 1$ puis montrer par récurrence que la formule pour x_{n-1} *et* x_n implique la formule pour x_{n+1}. A la relecture : comparer avec l'exercice (V.16)]

(I.5$^\sharp$) Dans le jeux des *tours de Hanoi*, le matériel consiste en 8 disques de carton de rayons distincts deux à deux et trois piquets verticaux A, B et C. Les disques sont troués en leurs centres et peuvent être empilés sur les piquets, toujours par ordre de rayons décroissants – le plus grand disque au bas de chaque *tour* et le plus petit en haut.

Le jeu commence avec les 8 disques empilés sur le piquet A, et le joueur doit les déplacer sur le piquet B en un minimum de coups. Chaque coup consiste en le mouvement d'un disque d'un piquet à un des deux autres ; il ne faut jamais poser un disque sur un disque plus petit.

Combien de coups faut-il ? (Le jeux est dû à Edouard Lucas et remonte à 1883).

[*Indication* : Considérer la solution T_n du jeux analogue pour $n = 1, 2, 3, \ldots, 8, \ldots$, deviner la forme générale de T_n, trouver une relation entre T_n et T_{n+1}, et achever par récurrence sur n. Variante : déduire de la relation de récurrence pour T_n une relation de récurrence pour $2^{-n}T_n$, et conclure.]

(I.6) Corriger les erreurs, y compris les fautes de frappe, dans les énoncés suivants.

(i) $2 \subset \mathbb{N}$.

(ii) $\frac{7}{2} \in \mathbb{Z}$.

(iii) Pour toute application α d'un ensemble X dans un ensemble Y et pour deux sous-ensembles $A, B \in X$, on a $\alpha(A \cap B) \supset \alpha(A) \cap \alpha(B)$.

(iv) Soient $\alpha : X \longrightarrow Y$ une application, A une partie de X et B une partie de Y ; alors $\alpha^{-1}(\alpha(A)) = A$ et $\alpha(\alpha^{-1}(B)) = B$.

(I.7) Pour tout entier $d \geq 0$, soit $d\mathbb{Z}$ le sous-ensemble des entiers rationnels formé des multiples de d. Enumérez les éléments de

$$\{n \in 6\mathbb{Z} \cap 10\mathbb{Z} \mid -100 \leq n \leq 100\}$$

et

$$\{n \in 6\mathbb{Z} \cup 10\mathbb{Z} \mid -20 \leq n \leq 20\}.$$

(I.8) Compter le nombre de bijections $\{1, 2, \ldots, n\} \longrightarrow \{1, 2, \ldots, n\}$ pour $n \leq 4$.

(I.9) Etant donnés deux entiers k, N tels que $1 \leq k \leq N$, quel est le rapport $p(k,N)$ entre le nombre de toutes les injections de $\{1,\ldots,k\}$ dans $\{1,\ldots,N\}$ et le nombre de toutes les applications de $\{1,\ldots,k\}$ dans $\{1,\ldots,N\}$?

En déduire une «justification» de la phrase suivante : sur 4 invitations à souper la même semaine, il y a 65 % de chances qu'il y en ait deux le même soir. Et, si vous disposez d'une calculette, de même pour : étant donné 23 personnes prises «au hasard», il y a plus d'une chance sur deux pour que deux d'entre elles aient leur anniversaire le même jour de l'année.

[*Indication* : Avant de recourir à la calculette, majorer $\ln\Big(\prod_{j=0}^{k}\big(1-\frac{j}{N}\big)\Big)$ par $\sum_{j=0}^{k}\big(-\frac{j}{N}\big)$.]

(I.10) Soient A, B, C des ensembles et $f : A \longrightarrow B$, $g : B \longrightarrow C$ des applications.

(i) Montrer que, si gf est surjective, alors g est surjective.

(ii) Quel est l'énoncé analogue pour l'injectivité ?

(iii) Donner un exemple tel que gf est bijective sans que ni f ni g ne le soit.

(iv) Montrer que, si gf est bijective et f est surjective, alors g et f sont bijectives.

(I.11$^\sharp$) On considère les applications $f, g : \mathbb{N} \times \mathbb{N} \longrightarrow \mathbb{N}$ définies par

$$f(x,y) = \frac{1}{2}(x+y)(x+y+1)+y \qquad \text{et} \qquad g(x,y) = 2^x(2y+1)-1.$$

(i) Reporter sur un dessin plan les valeurs de $f(x,y)$ lorsque $x+y \leq 6$.

(ii) Montrer que l'application f est injective.
[*Indication* : Soient (x,y) et (x',y') deux points distincts de $\mathbb{N}^2$. Vérifier que

$$f(x,y) < f(x',y')$$

d'abord lorsque $x+y < x'+y'$, ensuite lorsque $x+y = x'+y'$ et $y < y'$.]

(iii) Montrer que l'application f est surjective.
[*Indication* : Il y a plusieurs manières de procéder. Indication pour l'une d'entre elles : soit $k \in \mathbb{N}$; on définit $n, j \in \mathbb{N}$ par $\frac{1}{2}n(n+1) \leq k < \frac{1}{2}(n+1)(n+2)$ et $j = k - \frac{1}{2}n(n+1)$, et on vérifie que $f(n-j,j) = k$. Indication pour une autre manière : vérifies que $f(0,0) = 0$ et que

$$f(x,y+1) = f(x+1,y)+1 \text{ et } f(x+1,0) = f(0,x)+1,$$

pour tout $(x,y) \in \mathbb{N}^2$, de sorte que l'image de f contient 0 ainsi que le successeur immédiat de tout nombre qu'elle contient.]

(iv) Vérifier que l'application g est également une bijection.

(I.12) Résoudre le système linéaire

$$\begin{array}{rcrcrcl} I_1 & - & I_2 & - & I_3 & = & 0 \\ 8I_1 & + & 2I_2 & & & = & 65 \\ & - & 2I_2 & + & 11I_3 & = & 50 \end{array}$$

en trois inconnues I_1, I_2, I_3.

(I.13) Résoudre simultanément les systèmes linéaires

$$\begin{cases} x & & + z + t = 2 \\ x + y & & + t = 3 \\ & & z + t = 5 \end{cases} \qquad \text{et} \qquad \begin{cases} x & & + z + t = 3 \\ x + y & & + t = 5 \\ & & z + t = 2 \end{cases}$$

à 3 équations et 4 inconnues.

(I.14) La « formule de Simpson » est une égalité de la forme

$$\frac{1}{2}\int_{-1}^{1} \varphi(t)dt = a\varphi(-1) + b\varphi(0) + c\varphi(1)$$

pour toute fonction polynomiale de degré au plus 3, où a, b, c sont des constantes.

Déterminer ces constantes.

[Si nécessaire, attendre que l'expression $\int_{-1}^{1} \varphi(t)dt$ ait été expliquée au cours d'Analyse I avant de résoudre cet exercice.]

(I.15) Pour quelles valeurs du paramètre $a \in \mathbb{R}$ le système linéaire

$$\begin{array}{rcrcrcl} x & - & 3y & - & 2z & = & 4 \\ 3x & - & 4y & - & 9z & = & 5 \\ 4x & - & 7y & + & (a^2-20)z & = & a+6 \end{array}$$

a-t-il exactement une solution ? aucune solution ? une infinité de solutions ?

Lorsqu'il y a des solutions, les calculer.

(I.16) Pour quelles valeurs des paramètres s et t le système linéaire en les inconnues x, y, z

$$\begin{array}{rcrcrcl} sx & + & 2y & + & z & = & 1 \\ & & (t-1)y & + & sz & = & s \\ sx & + & 2ty & + & tz & = & 1 \end{array}$$

possède-t-il une unique solution, zéro solution, une famille à **un** paramètre libre de solutions, une famille à **deux** paramètres libres de solutions ?

Un exercice avec correction

Déterminer les valeurs des paramètres a et b pour lesquelles le système linéaire

$$\begin{array}{rcrcrcll} ax & + & by & + & z & = & 1 & (\ell_1) \\ x & + & aby & + & z & = & -4b & (\ell_2) \\ x & + & by & + & az & = & 3 & (\ell_3) \end{array}$$

en x, y, z possède des solutions, et déterminer la solution générale dans chacun de ces cas.

Solution

On commence par réduire le système à une forme triangulaire, par exemple[9] via les étapes suivantes :

$$\begin{array}{lrcrcrcl} & x & + & by & + & az & = & 3 & (\ell_3) \\ (*) & & & (a-1)by & + & (1-a)z & = & -4b-3 & (\ell_2-\ell_3) \\ & & & (1-a)by & + & (1-a^2)z & = & 1-3a & (\ell_1-a\ell_3) \end{array}$$

et

$$\begin{array}{lrcrcrcl} & x & + & by & + & az & = & 3 & \\ (**) & & & (a-1)by & + & (1-a)z & = & -4b-3 & . \\ & & & & & (1-a)(2+a)z & = & -4b-3+1-3a & \end{array}$$

Le **cas générique** est celui où $b \neq 0$ *et* $a \notin \{1,-2\}$ (cas d'un système dit «de Cramer»). Un tel système possède une unique solution, qu'on obtient immédiatement à partir de $(**)$:

$$\begin{aligned} z &= \frac{-2-3a-4b}{(1-a)(2+a)}\ , \\ y &= \frac{1}{(a-1)b}\Big(-4b-3+(a-1)z\Big) = 4\frac{ab+b+1}{(1-a)(2+a)b}\ , \\ x &= 3-by-az = \frac{2-a-4b}{(1-a)(2+a)}\ . \end{aligned}$$

[Il est *tout à fait indispensable* d'opérer de temps à autre quelques vérification. Ici, on vérifie facilement que les valeurs trouvées de x, y, z satisfont les equations (ℓ_1) à (ℓ_3) ou, si on manque de temps, au moins l'une de ces équations.]

Il reste trois cas particuliers à considérer.

Cas où $b = 0$**.** Les deux dernières équations du système $(**)$ donnent

$$(2+a)(-3) = -2-3a \qquad \text{ou encore} \qquad -6 = -2,$$

ce qui est impossible. *Il n'y a donc aucune solution.*

Cas où $a = 1$**.** On trouve de même

$$0 = -4b-3 \qquad \text{et} \qquad 0 = -4b-5,$$

ce qui est également impossible. *Il n'y a donc aucune solution.*

Cas où $a = -2$**.** Le système $(**)$ s'écrit

$$\begin{array}{lrcrcrcl} & x & + & by & - & 2z & = & 3 \\ (***) & & & -3by & + & 3z & = & -4b-3, \\ & & & & & 0 & = & -4b+4 \end{array}$$

et il faut distinguer deux sous-cas.

Si $b \neq 1$, *il n'y a aucune solution.*

---

[9] On peut bien remplacer comme ici (ℓ_1) par $(\ell_1) - a(\ell_3)$. Mais il est maladroit de remplacer (ℓ_3) par $a(\ell_3) - (\ell_1)$, car ceci pose un problème lorsque $a = 0$; il faudrait alors discuter dans la suite séparément le cas $a = 0$, ce qui entraîne des complications ; l'expérience montre que cette démarche est génératrice d'erreurs.

Si $b = 1$, le système se réduit à

$$\begin{array}{rcrcrcl} x & + & y & - & 2z & = & 3 \\ & - & 3y & + & 3z & = & -7. \end{array}$$

Il y a une infinité à un paramètre (noté ici λ) de solutions : $z = \lambda$, $y = \lambda + \frac{7}{3}$ et $x = \lambda + \frac{2}{3}$, ce qui s'écrit aussi

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \frac{2}{3} \\ \frac{7}{3} \\ 0 \end{pmatrix} + \lambda \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

CHAPITRE II

Espaces vectoriels

1. Définitions, premiers exemples, premières propriétés

Un *espace vectoriel réel* est un ensemble V dans lequel est spécifié un élément *zéro*, noté 0, et qui est muni de deux *opérations* appelées respectivement *addition*

$$V \times V \ni (x, y) \longmapsto x + y \in V$$

et *multiplication par les scalaires*

$$\mathbb{R} \times V \ni (\lambda, x) \longmapsto \lambda x \in V,$$

possédant les propriétés suivantes :

$$\begin{cases} (x+y)+z = x+(y+z) & \text{pour tous } x, y, z \in V\,; \\ 0+x = x & \text{pour tout } x \in V\,; \\ \text{pour tout } x \in V, \text{ il existe } y \in V \text{ tel que } x+y=0\,; & \\ x+y = y+x & \text{pour tout } x, y \in V\,; \end{cases}$$

$$\begin{cases} \lambda(x+y) = \lambda x + \lambda y & \text{pour tous } \lambda \in \mathbb{R} \text{ et } x, y \in V\,; \\ (\lambda+\mu)x = \lambda x + \mu x & \text{pour tous } \lambda, \mu \in \mathbb{R} \text{ et } x \in V\,; \\ (\lambda\mu)x = \lambda(\mu x) & \text{pour tous } \lambda, \mu \in \mathbb{R} \text{ et } x \in V\,; \\ 1x = x & \text{pour tout } x \in V. \end{cases}$$

L'usage est d'écrire

$x + y + z$ au lieu de $(x + y) + z$ ou de $x + (y + z)$,

$-x$ au lieu de y lorsque $x + y = 0$,

$x - y$ au lieu de $x + (-y)$,

$\lambda\mu x$ au lieu de $\lambda(\mu x)$ ou de $(\lambda\mu)x$,

et nous écrirons souvent « espace vectoriel » pour « espace vectoriel réel » . Répétons ici la note (1) du chapitre I : plus tard dans le cours, on verra qu'on peut systématiquement remplacer l'ensemble $\mathbb{R}$ des nombres réels par un autre corps, par exemple par le corps $\mathbb{Q}$ des nombres rationnels, le corps $\mathbb{C}$ des nombres complexes, un « corps fini premier » $\mathbb{F}_p$, ou plus généralement un corps arbitraire.

Soit V un espace vectoriel. Un *sous-espace vectoriel* de V est un sous-ensemble U de V tel que $0 \in U$ et

$$x + y \in U \text{ et } \lambda x \in U \text{ pour tous } x, y \in U \text{ et } \lambda \in \mathbb{R},$$

ou de manière équivalente tel que $0 \in U$ et

$$\lambda x + \mu y \in U \text{ pour tous } \lambda, \mu \in \mathbb{R} \text{ et } x, y \in U.$$

Notons qu'un sous-espace vectoriel U d'un espace vectoriel V est *en soi* un espace vectoriel.

EXEMPLES. Chacune des affirmations des exemples qui suivent nécessite évidemment une justification. Le lecteur est invité à formuler pour lui-même des arguments convenables dans tous leurs détails.

(i) $\mathbb{R}^n$ est un espace vectoriel pour les opérations définies par

$$\begin{aligned}(x_1,\dots,x_n)+(y_1,\dots,y_n) &= (x_1+y_1,\dots,x_n+y_n)\\ \lambda(x_1,\dots,x_n) &= (\lambda x_1,\dots,\lambda x_n)\end{aligned}$$

pour tous $x_1,\dots,x_n,y_1,\dots,y_n,\lambda\in\mathbb{R}$.

Attention : Dans ce cours, on *pense* le plus souvent aux vecteurs de $\mathbb{R}^n$ comme à des *colonnes* de n nombres réels. Toutefois, il arrive qu'on les *écrive* comme des *lignes*, pour des raisons typographiques.

Pour deux entiers p,n tels que $0\le p\le n$, l'ensemble

$$\{(x_1,\dots,x_n)\in\mathbb{R}^n \mid x_{p+1}=\cdots=x_n=0\}$$

est un sous-espace vectoriel de $\mathbb{R}^n$; on l'identifie parfois à $\mathbb{R}^p$. Les cas $p=0$ et $p=n$ se généralisent comme suit.

(ii) Si V est un espace vectoriel, $\{0\}$ et V lui-même sont des sous-espaces vectoriels de V, qu'on appelle parfois les *sous-espaces triviaux* ou *sous-espaces banals* de V.

(iii) Pour un ensemble X, l'ensemble $\mathcal{A}(X,\mathbb{R})$ de toutes les applications de X dans $\mathbb{R}$ est un espace vectoriel pour les opérations définies par

$$\begin{aligned}\alpha+\beta &: x\longmapsto \alpha(x)+\beta(x)\\ \lambda\alpha &: x\longmapsto \lambda\alpha(x)\end{aligned}$$

pour tous $\alpha,\beta\in\mathcal{A}(X,\mathbb{R})$, $\lambda\in\mathbb{R}$ et $x\in X$. [Au choix du lecteur : supposer X non vide, ou convenir qu'il existe une unique application de l'ensemble vide dans $\mathbb{R}$, et qu'elle constitue donc un espace vectoriel réduit à un seul élément.]

Dans le cas où X est un sous-ensemble de la droite réelle, l'espace $\mathcal{C}(X,\mathbb{R})$ de toutes les fonctions *continues* à valeurs réelles sur X est naturellement un sous-espace vectoriel de $\mathcal{A}(X,\mathbb{R})$. Si $X=\mathbb{R}$, l'espace $\mathcal{P}(\mathbb{R},\mathbb{R})$ des fonctions polynomiales et l'espace $\mathcal{P}_d(\mathbb{R},\mathbb{R})$ des fonctions polynomiales de degré au plus d (où d est un entier positif fixé) sont des sous-espaces vectoriels de $\mathcal{C}(\mathbb{R},\mathbb{R})$, et *a fortiori* de $\mathcal{A}(\mathbb{R},\mathbb{R})$.

L'ensemble des fonctions polynomiales de degré exactement d est un sous-ensemble de $\mathcal{A}(\mathbb{R},\mathbb{R})$ qui *n'est pas* un sous-espace vectoriel.

(iv) **Exemple fondamental.** On considère un système linéaire, par exemple

$$(*)\qquad \begin{cases} x_1 + 3x_2 - 10x_3 + 12x_4 = 0\\ 2x_1 - x_2 + x_3 = 0\end{cases}$$

qui est *homogène*, c'est-à-dire «sans second membre» , ou encore de la forme[1] $\sum_{j=1}^n a_{i,j}x_j=0$ pour $i\in\{1,\dots,m\}$. Alors

$$\big\{(x_1,x_2,x_3,x_4)\in\mathbb{R}^4 \mid x_1,x_2,x_3,x_4 \quad \text{sont des solutions de } (*)\big\}$$

est un sous-espace vectoriel de $\mathbb{R}^4$.

[1] Avec $m=2$ et $n=4$ pour l'exemple considéré en $(*)$.

(v) L'ensemble $\mathbb{R}^{\mathbb{N}}$ des suites infinies $(x_0, x_1, x_2, \ldots)$ de nombres réels est un espace vectoriel pour les opérations définies par

$$\begin{aligned}(x_0, x_1, \ldots) + (y_0, x_1, \ldots) &= (x_0 + y_0, x_1 + y_1, \ldots)\\ \lambda(x_0, x_1, \ldots) &= (\lambda x_0, \lambda x_1, \ldots)\end{aligned}$$

pour tous $(x_0, x_1, \ldots), (y_0, y_1, \ldots) \in \mathbb{R}^{\mathbb{N}}$ et $\lambda \in \mathbb{R}$. Le sous-ensemble des suites bornées en est un sous-espace vectoriel ; *idem* pour les suites convergentes.

Comparer avec l'exemple (i).

(vi) Soient V un espace vectoriel et $(U_i)_{i\in I}$ une famille de sous-espaces vectoriels de V. Alors $\bigcap_{i\in I} U_i$ est encore un sous-espace vectoriel de V.

Attention : $\bigcup_{i\in I} V_i$ n'est en général *pas* un sous-espace vectoriel de V. (Exercice : esquisser un dessin pour le cas de $V = \mathbb{R}^2$ et de deux droites U_1, U_2 dont la réunion n'est pas un sous-espace vectoriel de V).

NOTATION. Le symbole 0 désigne en général *à la fois* le nombre zéro dans $\mathbb{R}$ et le vecteur nul dans V. Ce n'est qu'exceptionnellement que, comme ci-dessous, on distingue $0 \in \mathbb{R}$ de $\underline{0} \in V$.

Quelques propriétés d'un espace vectoriel V

(i) *Le zéro de V est unique.*

En effet, si $\underline{0}$ et $\underline{0}'$ sont deux éléments de V tels que $\underline{0}+x = x+\underline{0} = x$ et $\underline{0}'+x = x+\underline{0}' = x$ pour tout $x \in V$, alors $\underline{0}' = \underline{0}' + \underline{0} = \underline{0}$.

(ii) *L'inverse $-x$ d'un vecteur $x \in V$ est unique.*

En effet, soient $y, z \in V$ deux candidats à s'écrire $-x$, c'est-à-dire deux vecteurs de V tels que $x+y = y+x = \underline{0}$ et $x+z = z+x = \underline{0}$. Alors $y = y+\underline{0} = y+x+z = \underline{0}+z = z$.

(iii) *Soient $x, y \in V$ tels que $x + y = x$; alors $y = \underline{0}$.*

En effet, $y = x + y - x = x - x = \underline{0}$.

(iv) *Pour tout $\lambda \in \mathbb{R}$, on a $\lambda\underline{0} = \underline{0}$.*

En effet, soit $x \in V$; on a $\lambda\underline{0} + \lambda x = \lambda(\underline{0} + x) = \lambda x$, d'où $\lambda\underline{0} = \underline{0}$ par (iii).

(v) *Pour tout $x \in V$, on a $0x = \underline{0}$.*

En effet : $0x + x = 0x + 1x = (0 + 1)x = x$, donc $0x = \underline{0}$ par (iii).

(vi) *Soient $\lambda \in \mathbb{R}$ et $x \in V$ tels que $\lambda x = \underline{0}$; alors $\lambda = 0$ ou $x = \underline{0}$.*

(Plus précisément, l'une au moins de ces égalités est vraie.)

En effet, si $\lambda \neq 0$ et $\lambda x = \underline{0}$, alors $x = 1x = (\lambda^{-1}\lambda)x = \lambda^{-1}(\lambda x) = \lambda^{-1}\underline{0} = \underline{0}$, la dernière de ces égalités résultant de (iv).

2. Bases

Dans $\mathbb{R}^n$, il existe une famille particulière de vecteurs

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}, \; e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}, \; \ldots, \; e_n = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$$

qui s'appelle la *base canonique.* Elle possède les deux propriétés suivantes :

(i) Tout élément de $\mathbb{R}^n$ est de la forme

$$(x_1, \ldots, x_n) = \sum_{j=1}^{n} x_j e_j,$$

(ii) $\sum_{j=1}^{n} x_j e_j = 0$ si et seulement si $x_1 = x_2 = \cdots = x_n = 0$.

Pour formaliser le cas général, on se propose de définir les notions de *combinaison linéaire, famille génératrice, indépendance linéaire* (et dépendance linéaire), et de *base.*

Combinaisons linéaires et familles génératrices

Soit V un espace vectoriel et $(a_\iota)_{\iota \in I}$ une famille[2] d'éléments de V. Une combinaison linéaire d'éléments de $(a_\iota)_{\iota \in I}$ est un vecteur de V du type $\sum_{\iota \in I} \lambda_\iota a_\iota$ où $(\lambda_\iota)_{\iota \in I}$ est une famille de nombres réels *presque tous nuls,* c'est-à-dire une famille dont au plus un nombre fini sont non nuls[3]. L'ensemble de toutes les combinaisons linéaires d'éléments de $(a_\iota)_{\iota \in I}$ est un sous-espace vectoriel de V (vérification à faire de tête !) que nous notons $\langle (a_\iota)_{\iota \in I} \rangle$.

En particulier, si $I = \{1, 2, \ldots, n\}$, une *combinaison linéaire* d'éléments d'une famille $(a_i)_{1 \leq i \leq n}$ est un vecteur de V du type $\lambda_1 a_1 + \ldots + \lambda_n a_n$. Dans une première approche, le lecteur aura sans doute avantage à supposer systématiquement ci-dessous que toutes les familles en jeux sont *finies* (et donc à ignorer les mots « presque tous nuls »).

On dit que la famille $(a_\iota)_{\iota \in I}$ *engendre* V, ou que cette famille est *génératrice* dans V, si $\langle (a_\iota)_{\iota \in I} \rangle = V$.

Exemples. (i) Dans $\mathbb{R}^2$, six vecteurs pointant vers les sommets d'un hexagone entourant l'origine engendrent $\mathbb{R}^2$ tout entier. Il en est de même de deux vecteurs qui ne sont pas sur une même droite. Pour $x \in \mathbb{R}^2$, la famille à deux vecteurs $(x, 2x)$ n'engendre pas $\mathbb{R}^2$, pas plus que les familles $(x, 0)$ et $(x, -x)$.

---

[2] On appelle *famille d'éléments* d'un ensemble V une application $i \longmapsto a_i$ d'un ensemble I dans V, et on désigne par $(a_i)_{i \in I}$ une telle famille. Pour une famille finie (c'est-à-dire telle que I est fini), le cas particulier le plus fréquent est celui où I est de la forme $\{1, \ldots, n\}$, et on écrit alors la famille $(x_i)_{1 \leq i \leq n}$. Noter que les éléments x_i *ne sont pas* nécessairement distincts : si $n \geq 2$, on peut avoir par exemple $x_1 = x_2$. Dans le cas où les éléments d'une famille sont distincts deux à deux, on identifie souvent la famille à un *sous-ensemble* de V ; d'où des écritures du type « soit $\{x_i\}_{i \in I} \subset V$ une famille finie de V » .

[3] C'est une différence importante entre le cours d'Algèbre I et le cours d'Analyse I : dans le premier, on considère des sommes *finies*, ou à la rigueur des sommes infinies *dont presque tous les éléments sont nuls ;* dans le second, on considère d'autres types de sommes, comme par exemple $\sum_{n=1}^{\infty} \frac{1}{n^2}$.

(ii) La base canonique $(e_1, \ldots, e_n)$ de $\mathbb{R}^n$ engendre $\mathbb{R}^n$.

(iii) L'espace $\mathcal{P}(\mathbb{R}, \mathbb{R})$ des fonctions polynomiales de $\mathbb{R}$ dans $\mathbb{R}$ est engendré par la famille $(t \longmapsto t^n)_{n\in\mathbb{N}}$.

(iv) **Exemple fondamental.** Considérons un système linéaire

$$(*) \qquad \begin{array}{ccccccccc} a_{1,1}x_1 & + & a_{1,2}x_2 & + & \cdots & + & a_{1,n}x_n & = & b_1 \\ a_{2,1}x_1 & + & a_{2,2}x_2 & + & \cdots & + & a_{2,n}x_n & = & b_2 \\ \multicolumn{9}{c}{\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots} \\ a_{m,1}x_1 & + & a_{m,2}x_2 & + & \cdots & + & a_{m,n}x_n & = & b_m \end{array}$$

comme au début du § I.1. Posons

$$a_1 = \begin{pmatrix} a_{1,1} \\ \vdots \\ a_{m,1} \end{pmatrix}, \ \ldots, \ a_n = \begin{pmatrix} a_{1,n} \\ \vdots \\ a_{m,n} \end{pmatrix}, \ b = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix} \in \mathbb{R}^m$$

de sorte que $(*)$ s'écrit

$$x_1a_1 + x_2a_2 + \cdots + x_na_n = b.$$

Alors le système $(*)$ possède une solution *si et seulement si*

$$b \in \langle a_1, \ldots, a_n \rangle.$$

Si cette dernière inclusion est satisfaite, il se peut que le système $(*)$ possède une unique solution, et il se peut tout aussi bien qu'il possède de nombreuses solutions.

(v) Soient V un espace vectoriel, $(b_1, \ldots, b_n)$ une famille génératrice et $a_1, \ldots, a_k$ des vecteurs de V. Si $b_1 \in \langle a_1, \ldots, a_k \rangle$, alors la famille $(a_1, \ldots, a_k, b_2, \ldots, b_n)$ est génératrice.

Dépendance et indépendance linéaire

Soit V un espace vectoriel. Une famille finie $(a_1, a_2, \ldots, a_n)$ d'éléments de V est dite *linéairement dépendante* s'il existe des nombres $\lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{R}$ *non tous nuls* tels que $\sum_{j=1}^n \lambda_j a_j = 0$. Plus généralement, une famille arbitraire $(a_\iota)_{\iota\in I}$ est dite linéairement dépendante si elle possède une sous-famille finie linéairement dépendante.

Reformulons cette définition comme suit.

Une famille finie $(a_1, a_2, \ldots, a_n)$ d'éléments de V est dite *linéairement indépendante* si, pour $\lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{R}$, l'équation $\sum_{j=1}^n \lambda_j a_j = 0$ implique $\lambda_1 = \lambda_2 = \cdots = \lambda_n = 0$. Plus généralement, une famille arbitraire $(a_\iota)_{\iota\in I}$ est dite linéairement indépendante si, pour une famille $(\lambda_\iota)_{\iota\in I}$ de nombres réels presque tous nuls, l'équation $\sum_{\iota\in I} \lambda_\iota a_\iota = 0$, implique $\lambda_\iota = 0$ pour tout $\iota \in I$.

Synonymes : une famille est *libre* si ses éléments sont linéairement indépendants, et *liée* dans le cas contraire.

Remarque. S'il y a des répétitions dans la famille $(a_\iota)_{\iota\in I}$, c'est-à-dire si au moins un élément apparaît au moins deux fois, comme dans (u, v, w, w, x, y, y, y), la famille est nécessairement liée.

Si le vecteur 0 apparaît dans la famille, alors celle-ci est liée.

EXEMPLES. (i) Soient V un espace vectoriel et $x \in V$. La famille réduite au seul élément x est libre si et seulement $x \neq 0$ (propriété (vi) à la fin du § II.1).

(ii) Soient x, y deux vecteurs non nuls et distincts d'un espace vectoriel. Alors la famille (x, y) est libre si et seulement si x et y ne sont pas sur une même droite.

(iii) Les vecteurs de la base canonique de $\mathbb{R}^n$ sont linéairement indépendants.

(iv) **Exemple fondamental.** Considérons un système linéaire homogène

$$(**) \qquad \begin{array}{ccccccccc} a_{1,1}x_1 & + & a_{1,2}x_2 & + & \cdots & + & a_{1,n}x_n & = & 0 \\ a_{2,1}x_1 & + & a_{2,2}x_2 & + & \cdots & + & a_{2,n}x_n & = & 0 \\ \multicolumn{9}{c}{\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots} \\ a_{m,1}x_1 & + & a_{m,2}x_2 & + & \cdots & + & a_{m,n}x_n & = & 0 \end{array}$$

et posons comme plus haut

$$a_1 = \begin{pmatrix} a_{1,1} \\ \vdots \\ a_{m,1} \end{pmatrix}, \ \ldots, \ a_n = \begin{pmatrix} a_{1,n} \\ \vdots \\ a_{m,n} \end{pmatrix} \in \mathbb{R}^m.$$

Alors le système $(**)$ possède une solution autre que $x_1 = \ldots = x_n = 0$ *si et seulement si* la famille $(a_1, \ldots, a_n)$ est linéairement dépendante.

(v) Pour $n \in \mathbb{N}$, les applications $t \longmapsto t^n$ sont linéairement indépendantes dans l'espace noté $\mathcal{P}(\mathbb{R}, \mathbb{R})$ plus haut.

Attention : ces applications ne sont pas linéaires (sauf pour $n = 1$), mais ce sont des éléments d'un espace vectoriel, d'où l'usage du vocabulaire de l'algèbre linéaire !

De même, dans l'espace $\mathcal{C}(\mathbb{R}_+^*, \mathbb{R})$ des fonctions à valeurs réelles sur $\mathbb{R}_+^* =]0, \infty[$, la famille $(t \longmapsto t^a)_{a \in \mathbb{R}}$ est libre.

EXERCICE. Le montrer ; si cela s'avère difficile, re-essayer après avoir progressé dans le cours d'analyse I.

REMARQUE. Cette famille n'est pas dénombrable.

Définition principale du paragraphe : base

Une *base* d'un espace vectoriel V est une famille libre et génératrice dans V.

EXEMPLES. (i) Dans $\mathbb{R}^2$, toute paire de vecteurs non nuls et non alignés est une base.

(ii) Dans $\mathbb{R}^n$, la base canonique est une base (!).

(iii) Dans $\mathcal{P}(\mathbb{R}, \mathbb{R})$, la famille $(t \longmapsto t^n)_{n \in \mathbb{N}}$ est une base.

PROPOSITION. *Soit $(b_\iota)_{\iota \in I}$ une famille d'un espace vectoriel V. Les deux conditions suivantes sont équivalentes :*

(i) *$(b_\iota)_{\iota \in I}$ est une base de V ;*

(ii) *tout vecteur $x \in V$ s'écrit de manière unique comme combinaison linéaire d'éléments de $(b_\iota)_{\iota \in I}$.*

PREUVE. (i) $\Longrightarrow$ (ii). Supposons que $(b_\iota)_{\iota\in I}$ est une base, et soit $x \in V$. Comme $(b_\iota)_{\iota\in I}$ engendre V, il existe une famille $(\lambda_\iota)_{\iota\in I}$ de nombres réels presque tous nuls telle que $x = \sum_{\iota\in I} \lambda_\iota b_\iota$, ce qui montre *l'existence* de l'écriture de (ii).

Supposons qu'il existe *deux* telles écritures : $x = \sum_{\iota\in I} \lambda_\iota b_\iota$ et $x = \sum_{\iota\in I} \mu_\iota b_\iota$. Alors

$$0 \;=\; x - x \;=\; \sum_{\iota\in I} (\lambda_\iota - \mu_\iota)\, b_\iota$$

et, comme les b_ι sont linéairement indépendants,

$$\lambda_\iota - \mu_\iota \;=\; 0 \qquad \text{pour tout} \qquad \iota \in I,$$

ce qui montre l'affirmation *d'unicité* de (ii).

(ii) $\Longrightarrow$ (i) Supposons que tout vecteur de V s'écrive comme combinaison linéaire d'éléments de $(b_\iota)_{\iota\in I}$. Alors $(b_\iota)_{\iota\in I}$ engendre évidemment V. D'autre part, soit $(\lambda_\iota)_{\iota\in I}$ une famille de nombres réels presque tous nuls telle que $\sum_{\iota\in I} \lambda_\iota b_\iota = 0$. On a évidemment aussi $\sum_{\iota\in I} 0 b_\iota = 0$, d'où $\lambda_\iota = 0$ pour tout $\iota \in I$ par unicité. En d'autres termes, les éléments de la famille $(b_\iota)_{\iota\in I}$ sont linéairement indépendants. □

COROLLAIRE. *Soit V un espace vectoriel possédant une base finie $(b_1, \ldots, b_n)$. L'application*

$$\begin{cases} V & \longrightarrow & \mathbb{R}^n \\ x = \displaystyle\sum_{i=1}^{n} x_i b_i & \longmapsto & \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \end{cases}$$

est bien définie, et c'est une bijection.

REMARQUE. Il est montré plus loin que tout espace vectoriel « de type fini » possède une base finie, et que le nombre n des éléments d'une telle base ne dépend d'aucun choix. Par ailleurs, dans un chapitre ultérieur, le terme « bijection » est précisé en « isomorphisme linéaire » .

Soient $(b_\iota)_{\iota\in I}$ une base d'un espace vectoriel V et x un vecteur de V ; soit $x = \sum_{\iota\in I} x_\iota b_\iota$ l'unique écriture de x comme combinaison linéaire d'éléments de $(b_\iota)_{\iota\in I}$. Les nombres x_ι sont les *coordonnées* du vecteur x dans la base $(b_\iota)_{\iota\in I}$.

EXEMPLES. (i) Les vecteurs $\begin{pmatrix} 2 \\ 1 \end{pmatrix}$ et $\begin{pmatrix} 1 \\ 3 \end{pmatrix}$ de $\mathbb{R}^2$ constituent une base de $\mathbb{R}^2$. Cherchons les coordonnées du vecteur $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ relativement à cette base.

L'égalité $\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \lambda \begin{pmatrix} 2 \\ 1 \end{pmatrix} + \mu \begin{pmatrix} 1 \\ 3 \end{pmatrix}$ équivaut aux deux équations $2\lambda + \mu = 1$ et $\lambda + 3\mu = 0$. La solution de ce système linéaire à deux équations et deux inconnues n'offre aucune difficulté, et l'unique solution est donnée par $\mu = -\frac{1}{5}$, $\lambda = \frac{3}{5}$. Par suite

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix} = 3/5 \begin{pmatrix} 2 \\ 1 \end{pmatrix} - 1/5 \begin{pmatrix} 1 \\ 3 \end{pmatrix}.$$

(ii) Revenons à l'exemple fondamental d'un système linéaire, écrit sous la forme

$$x_1 a_1 + \cdots + x_n a_n \ = \ b,$$

les lettres $a_1, \ldots, a_n, b$ désignant des vecteurs de $\mathbb{R}^m$ et $x_1, \ldots, x_n$ les inconnues. Supposons que ce système ait au moins une solution, c'est-à-dire que b soit dans le sous-espace $\langle a_1, \ldots, a_n \rangle$ de $\mathbb{R}^m$. Alors ce système possède une *unique* solution si et seulement si la famille $(a_1, \ldots, a_n)$ est libre.

Nota bene : lorsque $(a_1, \ldots, a_n)$ est libre, et que c'est donc une base du sous-espace vectoriel $\langle a_1, \ldots, a_n \rangle$ de $\mathbb{R}^n$, l'unique solution du système fournit les coordonnées $x_1, \ldots, x_n$ de b relativement à cette base.

PROPOSITION. *Soient V un espace vectoriel et $(b_\iota)_{\iota \in I}$ une famille d'éléments de V. Les conditions suivantes sont équivalentes :*

(i) *$(b_\iota)_{\iota \in I}$ est une base de V ;*

(ii) *$(b_\iota)_{\iota \in I}$ est une famille libre maximale de V ;*

(iii) *$(b_\iota)_{\iota \in I}$ est une famille génératrice minimale de V.*

PREUVE. (i) $\Longrightarrow$ (ii) & (iii). Supposons que $(b_\iota)_{\iota \in I}$ est une base, c'est-à-dire une famille libre et génératrice.

Alors $(b_\iota)_{\iota \in I}$ est libre maximale. En effet, pour toute famille contenant d'une part $(b_\iota)_{\iota \in I}$ et d'autre part un élément $x \in V$, il existe une famille $(\lambda_\iota)_{\iota \in I}$ de nombres réels presque tous nuls telle que $x = \sum_{\iota \in I} \lambda_\iota b_\iota$, de sorte que

$$x - \sum_{\iota \in I} \lambda_\iota b_\iota \ = \ 0$$

est une relation linéaire non banale entre éléments de la famille obtenue par réunion de (x) et $(b_\iota)_{\iota \in I}$. Donc $(b_\iota)_{\iota \in I}$ est bien maximale parmi les familles libres d'éléments de V.

Par ailleurs $(b_\iota)_{\iota \in I}$ est génératrice minimale. En effet, s'il existait un élément ι_0 de I tel que la famille $(b_\iota)_{\iota \in I \setminus \{\iota_0\}}$ soit encore une famille génératrice, il existerait une famille $(\lambda_\iota)_{\iota \in I \setminus \{\iota_0\}}$ de nombres réels presque tous nuls telle que

$$b_{\iota_0} \ = \ \sum_{\iota \in I \setminus \{\iota_0\}} \lambda_\iota b_\iota,$$

ce qui contredirait l'hypothèse selon laquelle $(b_\iota)_{\iota \in I}$ est libre. Donc $(b_\iota)_{\iota \in I}$ est bien minimale parmi les familles génératrices d'éléments de V.

(ii) $\Longrightarrow$ (i). Supposons $(b_\iota)_{\iota \in I}$ libre maximale. Il s'agit de montrer que $(b_\iota)_{\iota \in I}$ est génératrice.

Soit $x \in V$. Comme la famille obtenue par réunion de x et de $(b_\iota)_{\iota \in I}$ est liée, il existe $\lambda \in \mathbb{R}$ et une famille $(\lambda_\iota)_{\iota \in I}$ de nombres réels presque tous nuls, ces nombres n'étant pas tous nuls, tels que

$$\lambda x \ + \ \sum_{\iota \in I} \lambda_\iota b_\iota \ = \ 0.$$

Or $\lambda \neq 0$ (sinon on aurait $\sum_{\iota\in I} \lambda_\iota b_\iota = 0$ avec les nombres de $(\lambda_\iota)_{\iota\in I}$ non tous nuls, et la famille $(b_\iota)_{\iota\in I}$ serait donc liée, contrairement à l'hypothèse). Par suite

$$x = \sum_{\iota\in I} \left(-\frac{\lambda_\iota}{\lambda}\right) b_\iota$$

ou encore $x \in \langle (b_\iota)_{\iota\in I} \rangle$.

(iii) $\Longrightarrow$ (i). Supposons la famille $(b_\iota)_{\iota\in I}$ génératrice minimale. Soit $(\lambda_\iota)_{\iota\in I}$ une famille de nombres réels presque tous nuls telle que $\sum_{\iota\in I} \lambda_\iota b_\iota = 0$. S'il existait $\iota_0 \in I$ tel que $\lambda_{\iota_0} \neq 0$, alors on aurait $b_{\iota_0} = \sum_{\iota\in I, \iota\neq\iota_0} \left(-\frac{\lambda_\iota}{\lambda_{\iota_0}}\right) b_\iota$, de sorte que la famille $I \setminus \{\iota_0\}$ serait aussi génératrice, contrairement à l'hypothèse de minimalité portant sur I. La famille $(b_\iota)_{\iota\in I}$ est donc libre, de sorte que c'est une base. □

3. Dimension d'un espace vectoriel

Un espace vectoriel V est *de type fini* s'il possède une famille génératrice finie.

Par exemple, $\mathbb{R}^n$ est de type fini pour tout $n \geq 0$, alors que $\mathcal{P}(\mathbb{R}, \mathbb{R})$ ne l'est pas.

Théorème II.1. *Soient V un espace vectoriel de type fini, $\mathcal{L} \subset V$ une famille libre et $\mathcal{G} \subset V$ une famille génératrice finie.*

Il existe une base $\mathcal{B}$ de V telle que

$$\mathcal{L} \subset \mathcal{B} \subset \mathcal{L} \cup \mathcal{G}.$$

En particulier, pour toute famille libre $\mathcal{L}$ de V, il existe une base $\mathcal{B}$ de V telle que $\mathcal{L} \subset \mathcal{B}$.

Avant de montrer ce théorème, formulons deux lemmes. Ce sont des conséquences immédiates des définitions.

Lemme 1. *Soient V un espace vectoriel, $\mathcal{F}$ une famillle de V, et $\mathcal{G}$ une famillle génératrice de V. Pour que la famille $\mathcal{F}$ soit génératrice, il faut et il suffit que $g \in \langle \mathcal{F} \rangle$ pour tout $g \in \mathcal{G}$.*

Lemme 2. *Soient V un espace vectoriel, $\mathcal{L}$ une famillle libre de V, et $U = \langle \mathcal{L} \rangle$ le sous-espace vectoriel de V qu'elle engendre. On suppose que $U \subsetneqq V$ et on choisit $x \in V$, $x \notin U$. Alors la famille $\mathcal{L} \cup \{x\}$ est libre.*

Démonstration du théorème. Soit $\mathcal{E}$ l'ensemble des parties $\mathcal{A} \subset \mathcal{G}$ telles que

(i) $\mathcal{L} \cap \mathcal{A} = \emptyset$;

(ii) $\mathcal{L} \cup \mathcal{A}$ est libre.

On ordonne $\mathcal{E}$ par inclusion : $\mathcal{A}_1$ inférieur ou égal à $\mathcal{A}_2$ si (par définition) $\mathcal{A}_1 \subset \mathcal{A}_2$.

Notons g le nombre d'éléments de $\mathcal{G}$. Pour tout $\mathcal{A} \in \mathcal{E}$, la partie $\mathcal{A}$ possède au plus g éléments. Il existe donc une partie $\mathcal{A}_{\max} \in \mathcal{E}$ possédant le maximum d'éléments. Alors $\mathcal{L} \cup \mathcal{A}_{\max}$ est une famille génératrice (sinon il existerait $x \in \mathcal{G}$ tel que $x \notin \langle \mathcal{L} \cup \mathcal{A}_{\max} \rangle$, et $\mathcal{L} \cup (\mathcal{A}_{\max} \cup \{x\})$ serait encore libre, contrairement à la définition de $\mathcal{A}_{\max}$).

Il suffit donc de poser $\mathcal{B} = \mathcal{L} \cup \mathcal{A}_{\max}$ pour conclure. □

Corollaire. *Tout espace vectoriel de type fini possède une base finie.*

Point de la situation

Dans un espace vectoriel V, nous avons défini des sous-ensembles de *vecteurs linéairement indépendants*, dits aussi sous-ensembles (ou familles) *libres*, des sous-ensembles *générateurs*, et des *bases*, qui sont des sous-ensembles à la fois libres et générateurs. La proposition précédente montre en particulier qu'un espace vectoriel V *de type fini* possède une base finie.

Le pas suivant, obligé et fondamental, consiste à montrer que deux bases d'un même espace vectoriel ont le même nombre d'éléments, et c'est là que le «lemme» qui suit joue un rôle crucial ; c'est aussi le premier résultat du cours dont la preuve ne se résume pas à une simple vérification de définitions. Avant de lire la preuve du lemme d'échange, c'est un bon exercice de méditer une preuve de l'assertion suivante : «Soient V un espace vectoriel, $(g_1, \ldots, g_n)$ une famille génératrice de V et $a_1, \ldots, a_k \in V$. Si $g_1 \in \langle a_1, \ldots, a_k \rangle$, alors la famille $(a_1, \ldots, a_k, g_2, \ldots, g_n)$ est génératrice.»

THÉORÈME II.2 (Lemme d'échange de Grassmann[4]). *Soit V un espace vectoriel de type fini, $(a_1, \ldots, a_m)$ une famille libre de V et $(b_1, \ldots, b_n)$ une famille génératrice de V.*

Alors $m \leq n$; de plus, il existe une renumérotation des vecteurs b_j telle que la famille $(a_1, \ldots, a_m, b_{m+1}, \ldots, b_n)$ est génératrice.

PREUVE. Si $m = 0$ il n'y a rien à montrer. On peut donc supposer désormais que $m \geq 1$. Pour tout entier $r \in \{1, \ldots, m\}$, formulons l'assertion

$$A(r) \ : \begin{cases} r \leq n \text{ et, après renumérotation si nécessaire des} \\ \text{vecteurs } b_j, \text{ la famille } (a_1, \ldots, a_r, b_{r+1}, \ldots, b_n) \text{ est} \\ \text{génératrice.} \end{cases}$$

La preuve consiste à montrer $A(r)$ par récurrence sur l'entier r ; cette preuve implique que l'assertion $A(m)$ est vraie, qui est la conclusion du lemme.

Preuve de l'assertion $A(1)$. Comme $m \geq 1$, l'espace V n'est pas réduit à $\{0\}$, donc $r = 1 \leq n$. Par ailleurs, comme $(b_1, \ldots, b_n)$ est une famille génératrice, il existe des nombres réels $\mu_1, \ldots, \mu_n$ tels que $a_1 = \sum_{j=1}^{n} \mu_j b_j$; quitte à renuméroter les b_j, nous pouvons supposer que $\mu_1 \neq 0$. Il en résulte que

$$b_1 \ = \ \mu_1^{-1} a_1 \ - \ \sum_{j=2}^{n} \mu_1^{-1} \mu_j b_j \ \in \ \langle a_1, b_2, \ldots, b_n \rangle$$

(rappel : $\langle a_1, b_2, \ldots, b_n \rangle$ désigne le sous-espace de V engendré par $a_1, b_2, \ldots, b_n$) ; par suite

$$\langle a_1, b_2, \ldots, b_n \rangle \ = \ \langle b_1, b_2, \ldots, b_n \rangle \ = \ V.$$

(Remarque : si $n = 1$, la somme $\sum_{j=2}^{n} \mu_1^{-1} \mu_j b_j$ porte sur un ensemble vide d'indices, et sa valeur est donc le vecteur nul de V.)

Pas de récurrence : si $r \in \{1, \ldots, m-1\}$ et si l'assertion $A(r)$ est vraie, alors l'assertion $A(r+1)$ est aussi vraie. On sait que $r \leq n$ (c'est une partie de l'assertion $A(r)$). S'il y avait égalité $r = n$, la famille $(a_1, \ldots, a_r)$ serait génératrice (par l'autre

[4] Hermann Grassmann, né en 1809 et mort en 1877 à Stettin (Allemagne), entreprit des études de théologie, enseigna les mathématiques dans l'enseignement moyen, et gagna une réputation par ses travaux de linguistique et de sanscrit. Son ouvrage *Ausdehnungslehre* (1844) fut peu compris de ses contemporains, mais eut une influence considérable sur le développement ultérieur de l'algèbre.

partie de l'assertion $A(r)$), ce qui serait en contradiction avec le fait que la famille $(a_1, \ldots, a_r, a_{r+1})$ est libre. Donc $r+1 \leq n$.

Par ailleurs, comme $(a_1, \ldots, a_r, b_{r+1}, \ldots, b_n)$ est une famille génératrice, il existe des nombres réels $\lambda_1, \ldots, \lambda_r, \mu_{r+1}, \ldots, \mu_n$ tels que $a_{r+1} = \sum_{i=1}^{r} \lambda_i a_i + \sum_{j=r+1}^{n} \mu_j b_j$; de plus il n'est pas possible que $\mu_{r+1} = \cdots = \mu_n = 0$ (car la famille $\mathcal{A}$ est libre) ; quitte à renuméroter les vecteurs $b_{r+1}, \ldots, b_n$, nous pouvons donc supposer que $\mu_{r+1} \neq 0$. Il en résulte que

$$\begin{aligned} b_{r+1} &= \mu_{r+1}^{-1} \left(a_{r+1} - \lambda_1 a_1 - \cdots - \lambda_r a_r - \mu_{r+2} b_{r+2} - \cdots - \mu_n b_n\right) \\ &\in \langle a_1, \ldots, a_{r+1}, b_{r+2}, \ldots, b_n \rangle \end{aligned}$$

et par suite que

$$\langle a_1, \ldots, a_{r+1}, b_{r+2}, \ldots, b_n \rangle = \langle a_1, \ldots, a_r, b_{r+1}, \ldots, b_n \rangle = V.$$

Comme déjà noté en début de preuve, ces arguments montrent que l'assertion $A(m)$ est vraie, ce qui achève la preuve. □

THÉORÈME II.3. *Dans un espace vectoriel de type fini, toutes les bases ont le même nombre d'éléments.*

PREUVE. Nous savons que V possède au moins une base finie ; notons-la $\mathcal{B}$ et notons n le nombre de ses éléments. Soit $\mathcal{B}'$ une autre base de V. Le lemme d'échange implique d'abord que toute famille libre finie dans V a au plus n éléments, donc que toute famille libre de V est finie, et en particulier que $\mathcal{B}'$ est une base finie ; notons n' le nombre de ses éléments. Le même lemme d'échange impique ensuite que $n' \leq n$ et $n \leq n'$, d'où le résultat. □

Dimension

Le résultat précédent rend possible la défintion suivante. La dimension d'un espace vectoriel de type fini est le nombre d'éléments d'une quelconque de ses bases.

On note $\dim_{\mathbb{R}} V$ la dimension d'un tel espace V (ou bien sûr $\dim_{\mathbb{K}} V$ s'il s'agit d'un espace vectoriel sur un corps $\mathbb{K}$ qui n'est pas nécessairement celui des nombres réels).

EXEMPLE. Pour tout entier $n \geq 0$, l'espace $\mathbb{R}^n$ est de dimension n.

REMARQUE. Les résultats précédents se généralisent aux espaces vectoriels quelconques et familles génératrices quelconques. Les arguments se basent sur le « lemme de Zorn » . Voir le cours d'Algèbre II.

Désormais, les expressions « espace vectoriel de dimension finie » et « espace vectoriel de type fini » sont synonymes. D'un espace vectoriel qui n'est pas de type fini, on dit qu'il est « de dimension infinie » .

On dit désormais « V est un espace vectoriel de dimension n » pour « V est un espace vectoriel de dimension finie notée n » .

COROLLAIRE. *Soit V un espace vectoriel de dimension n.*

(i) *Toute famille libre à n éléments est une base.*

(ii) *Toute famille génératrice à n éléments est une base.*

COROLLAIRE. *Soient V un espace vectoriel de dimension n, soit U un sous-espace vectoriel de V et soit m la dimension de U. Alors toute base $(b_1, \ldots, b_m)$ de U peut être complétée en une base $(b_1, \ldots, b_m, b_{m+1}, \ldots, b_n)$ de V ; en particulier, $m \leq n$.*

PREUVES. Ces deux corollaires sont des conséquences immédiates du lemme d'échange. □

CONVENTION. L'espace vectoriel réduit à un élément $V = \{0\}$ possède une base qui est l'ensemble vide (!), et cet espace est donc de dimension 0. Voir à ce sujet la *remarque* du paragraphe suivant concernant les sommes directes $U_1 \oplus U_2$ telles que U_1 ou U_2 est réduit à $\{0\}$.

Il est important de réaliser qu'un espace vectoriel *n'a en général pas de base privilégiée.*

Soient par exemple $n \geq 0$ un entier et $\mathcal{P}_n$ l'espace des applications polynomiales de $\mathbb{R}$ dans $\mathbb{R}$ de degré au plus n. Voici un tout petit choix de bases de V, dont chacune joue un rôle bien utile dans divers problèmes :

(i) La base formée des monômes 1, t, t^2, …, t^n.

(ii) La base formée des monômes pondérés 1, t, $\frac{1}{2}t^2$, …, $\frac{1}{n!}t^n$.

(iii) La base $(P_j)_{0 \leq j \leq n}$ formée des «polynômes de Legendre» , définis par

$$P_j(t) = \frac{1}{2^j j!} \frac{d^j}{dt^j} \left((t^2 - 1)^j \right).$$

(iv) La base $(H_j)_{0 \leq j \leq n}$ formée des «polynômes de Hermite» , définis par

$$\exp(-\frac{t^2}{2}) H_j(t) = \frac{1}{j!} \frac{d^j}{dt^j} \left(\exp(-\frac{t^2}{2}) \right).$$

(v) La base formée des fonctions polynomiales $t \longmapsto \frac{1}{j!}\, t(t-1)\cdots(t-j+1)$, pour $j \in \{0, \ldots, n\}$, qui prennent des valeurs entières aux points entiers.

REMARQUE (à propos des bases mentionnées en (i) à (iv)). Plus généralement, soit Q_j une application polynomiale de degré exactement j, c'est-à-dire une application de la forme $t \longmapsto a_{j,0} + a_{j,1}t + \ldots + a_{j,j}t^j$, avec $a_{j,j} \neq 0$; alors la famille $(Q_j)_{0 \leq j \leq n}$ est une base de l'espace vectoriel $\mathcal{P}_n$.

EXERCICE (à propos de la base mentionnée ci-dessus en (v)). Pour tout $f \in \mathcal{P}_n$, on définit Δf, $\Delta^2 f$, … (où Δ est la lettre grecque majuscule *delta*) par

$$\begin{aligned} \Delta f(t) &= f(t+1) - f(t), \\ \Delta^j f(t) &= \Delta^{j-1} f(t+1) - \Delta^{j-1} f(t) \quad \text{si} \quad j \geq 2. \end{aligned}$$

Pour tout $f \in \mathcal{P}_n$, vérifier d'abord que $\Delta^{n+1} f = 0$. Vérifier ensuite que

$$f(t) = f(0) + \Delta f(0)\, t + \Delta^2 f(0)\, \frac{t(t-1)}{2} + \ldots + \Delta^n f(0)\, \frac{t(t-1)(t-2)\cdots(t-n+1)}{n!}.$$

[Vérifier d'abord ce point pour $n = 1$ et $n = 2$. Le cas général est nettement plus difficile ; voir par exemple le théorème 1.2 de la page 12 dans le livre de E. Hairer et G. Wanner, *Analysis by its history*, Springer, 1996.] Montrer enfin que $f(\mathbb{Z}) \subset \mathbb{Z}$ *si et seulement si* les coordonnées $\Delta^j f(0)$ de f dans la base de (v) sont des entiers pour tout j.

Exemple d'un système linéaire

Considérons un système linéaire avec le même nombre n d'équations et d'inconnues

$$(*) \qquad \begin{array}{rcl} a_{1,1}x_1 + a_{1,2}x_2 + \cdots + a_{1,n}x_n &=& b_1 \\ a_{2,1}x_1 + a_{2,2}x_2 + \cdots + a_{2,n}x_n &=& b_2 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots && \\ a_{n,1}x_1 + a_{n,2}x_2 + \cdots + a_{n,n}x_n &=& b_n \end{array}$$

qu'on écrit aussi

$$(**) \qquad x_1 a_1 + \cdots + x_n a_n = b$$

où $a_1, \ldots, a_n, b$ sont comme plus haut les vecteurs de $\mathbb{R}^n$ dont les composantes apparaissent dans l'écriture $(*)$. Alors les conditions suivantes sont équivalentes :

(i) pour tout $b \in \mathbb{R}^n$, le système possède *au moins* une solution ;

(ii) le système $x_1 a_1 + \cdots + x_n a_n = 0$ possède *l'unique solution* $x_1 = \cdots = x_n = 0$;

(iii) pour tout $b \in \mathbb{R}^n$, le système possède *au plus* une solution ;

(iv) pour tout $b \in \mathbb{R}^n$, le système possède *exactement* une solution.

Remarque. Il s'agit d'une partie des équivalences annoncées à la fin du § I.1.

Preuve. On a en effet les reformulations :

(i) $\Longleftrightarrow$ la famille $(a_1, \ldots, a_n)$ engendre $\mathbb{R}^n$;

(ii) $\Longleftrightarrow$ la famille $(a_1, \ldots, a_n)$ est libre ;

l'équivalence des conditions (ii) et (iii) est immédiate ;

(iv) $\Longleftrightarrow$ la famille $(a_1, \ldots, a_n)$ est une base de $\mathbb{R}^n$.

Par suite les conditions (i) à (iv) sont équivalentes (voir le corollaire qui suit la définition de la dimension). □

4. Sommes et sommes directes de sous-espaces vectoriels

Soient V un espace vectoriel et U_1, U_2 deux sous-espaces vectoriels de V. Le sous-ensemble de V constitué des éléments de la forme $u_1 + u_2$ avec $u_j \in U_j$ $(j = 1, 2)$ est un sous-espace vectoriel de V ; noter que c'est le sous-espace $\langle U_1 \cup U_2 \rangle$ *engendré* par la réunion de U_1 et U_2, ou encore que c'est le plus petit sous-espace vectoriel de V contenant[5] à la fois U_1 et U_2.

Par définition, ce sous-espace est la *somme* des sous-espaces U_1 et U_2 ; on le note $U_1 + U_2$.

[5] Rappel : Avec ces notations, $U_1 \cap U_2$ est toujours un sous-espace vectoriel de V, alors que $U_1 \cup U_2$ n'en est en général pas un.

PROPOSITION. *Avec les notations ci-dessus, et pour un espace V de dimension finie :*

$$\dim_{\mathbb{R}}(U_1 + U_2) \;=\; \dim_{\mathbb{R}}(U_1) + \dim_{\mathbb{R}}(U_2) - \dim_{\mathbb{R}}(U_1 \cap U_2).$$

PREUVE. Soit $(a_1, \ldots, a_l)$ une base de $U_1 \cap U_2$. On sait qu'il existe

- des vecteurs $b_{l+1}, \ldots, b_m \in U_1$ tels que $(a_1, \ldots, a_l, b_{l+1}, \ldots, b_m)$ est une base de U_1,
- des vecteurs $c_{l+1}, \ldots, c_n \in U_2$ tels que $(a_1, \ldots, a_l, c_{l+1}, \ldots, c_n)$ est une base de U_2.

Il reste à montrer l'assertion

$$(\sharp) \qquad (a_1, \ldots, a_l, b_{l+1}, \ldots, b_m, c_{l+1}, \ldots, c_n) \quad \textit{est une base de} \quad U_1 + U_2$$

dont la proposition résulte immédiatement.

D'une part, il est évident que la famille $(\sharp)$ est génératrice dans $U_1 + U_2$.

D'autre part, c'est une famille libre. En effet, une combinaison linéaire des éléments de $(\sharp)$ peut s'écrire

$$(\alpha_1 a_1 + \cdots + \alpha_l a_l) \;+\; (\beta_{l+1} b_{l+1} + \cdots + \beta_m b_m) \;=\; \gamma_{l+1} c_{l+1} + \cdots + \gamma_n c_n.$$

Le terme de gauche est dans U_1 et celui de droite dans U_2 ; comme ils sont égaux, ils sont chacun dans $U_1 \cap U_2$. En particulier, le terme de droite est une combinaison linéaire des vecteurs $a_1, \ldots, a_l$ de la base choisie de $U_1 \cap U_2$. Comme la famille $(a_1, \ldots, a_l, c_{l+1}, \ldots, c_n)$ est libre (c'est une base de U_2), il en résulte que $\gamma_{l+1} = \cdots = \gamma_n = 0$. Comme la famille $(a_1, \ldots, a_l, b_{l+1}, \ldots, b_m)$ est également libre, on obtient de même les égalités

$$\alpha_1 = \cdots = \alpha_l = \beta_{l+1} = \cdots = \beta_n = 0,$$

ce qui achève l'argument. □

EXEMPLE. Dans l'espace $V = \mathbb{R}^3$, notons U_{13} [respectivement U_{23}] le sous-espace des vecteurs de la forme $\begin{pmatrix} x \\ 0 \\ z \end{pmatrix}$ $\left[\text{resp. } \begin{pmatrix} 0 \\ y \\ z \end{pmatrix}\right]$. Alors $U_{13} + U_{23} = \mathbb{R}^3$, donc

$$\dim_{\mathbb{R}}(U_{13} + U_{23}) \;=\; 3 \;;$$

et on a bien

$$\dim_{\mathbb{R}}(U_{13}) \;+\; \dim_{\mathbb{R}}(U_{23}) \;-\; \dim_{\mathbb{R}}(U_{13} \cap U_{23}) \;=\; 2 + 2 - 1 \;=\; 3.$$

PROPOSITION. *Soient V un espace vectoriel et U_1, U_2 deux sous-espaces vectoriels de V. Les propriétés suivantes sont équivalentes :*

(i) *tout vecteur $v \in V$ s'écrit de manière unique sous la forme $v = u_1 + u_2$ avec $u_j \in U_j$ $(j = 1, 2)$;*

(ii) *$V = U_1 + U_2$ et $U_1 \cap U_2 = \{0\}$.*

Si de plus V est de dimension finie, ces propriétés sont encore équivalentes à

(iii) *$V = U_1 + U_2$ et $\dim_{\mathbb{R}}(V) = \dim_{\mathbb{R}}(U_1) + \dim_{\mathbb{R}}(U_2)$.*

PREUVE. *Premier pas.* Les assertions

tout $v \in V$ s'écrit $v = u_1 + u_2$ avec $u_j \in U_j$ $(j = 1, 2)$

et

$V = U_1 + U_2$

sont équivalentes, comme il résulte immédiatement des définitions.

Deuxième pas. Les assertions

tout $v \in V$ qui s'écrit $v = u_1 + u_2$ avec $u_j \in U_j$ $(j = 1, 2)$ s'écrit de manière unique sous cette forme

et

$$U_1 \cap U_2 \,=\, \{0\}$$

sont équivalentes.

En effet, supposons d'abord que $U_1 \cap U_2 = \{0\}$. Considérons $u_1, u_1' \in U_1$ et $u_2, u_2' \in U_2$ tels que $u_1 + u_2 = u_1' + u_2'$. Alors $u_1 - u_1' = u_2 - u_2' \in U_1 \cap U_2$, donc $u_1 - u_1' = 0$ et $u_2 - u_2' = 0$, d'où l'unicité. Réciproquement, supposons que $U_1 \cap U_2 \neq \{0\}$, et soient $w \in U_1 \cap U_2$, $w \neq 0$. En posant $u_1 = w$, $u_2 = w$, $u_1' = 0$ et $u_2' = 0$, on obtient deux écritures $u_1 - u_2 = u_1' - u_2' = 0$ de zéro, et l'unicité n'est pas vraie.

Fin de la preuve. Les deux pas précédents montrent l'équivalence des conditions (i) et (ii) de l'énoncé.

Lorque V est de dimension finie, l'équivalence de (i) et (ii) avec (iii) résulte de la proposition précédente. □

Sous les conditions de la proposition ci-dessus, on dit que l'espace V est *somme directe* des sous-espaces U_1 et U_2, et on écrit

$$V \,=\, U_1 \oplus U_2.$$

Remarque (concernant certaines conventions). Soit $V = U_1 \oplus U_2$ un espace vectoriel de dimension finie écrit comme une somme directe de deux sous-espaces. Pour que la formule $\dim_{\mathbb{R}}(V) = \dim_{\mathbb{R}}(U_1) + \dim_{\mathbb{R}}(U_2)$ soit vraie dans tous les cas, y compris celui où $U_1 = \{0\}$, il faut bien convenir que la dimension d'un espace vectoriel réduit à zéro est nulle. Pour que la définition de la dimension soit vraie dans tous les cas, il faut bien convenir que l'espace vectoriel réduit à zéro a pour base l'ensemble vide. Enfin, pour que tout vecteur de cet espace $\{0\}$ soit combinaison linéaire d'éléments de la base, il faut bien convenir que, dans un espace vectoriel, une somme de zéro vecteurs (= une somme portant sur une famille vide de vecteurs !) est le vecteur nul.

Toute autre choix de conventions serait sans doute logiquement possible, mais conduirait à des énoncés d'une complexité bien supérieure, avec une multitude de cas et sous-cas.

Exemples. (i) L'espace $V = \mathbb{R}^3$ est une somme directe

$$\mathbb{R}^3 \,=\, \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 \;\middle|\; z = 0 \right\} \oplus \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 \;\middle|\; x = y = 0 \right\}$$

(plan horizontal $\oplus$ axe vertical).

(ii) L'espace $\mathcal{P}$ des fonctions polynomiales de $\mathbb{R}$ dans $\mathbb{R}$ est somme directe

$$\mathcal{P} \,=\, \mathcal{P}_+ \oplus \mathcal{P}_-$$

où $\mathcal{P}_+$ et $\mathcal{P}_-$ désignent respectivement les sous-espaces des fonctions paires et impaires. En effet, toute fonction $f \in \mathcal{P}$ s'écrit[6] comme somme de la fonction paire $\frac{1}{2}(f(t)+f(-t))$ et de la fonction impaire $\frac{1}{2}(f(t) - f(-t))$. Par ailleurs, si $f = f_+ + f_-$ avec $f_+ \in \mathcal{P}_+$ et $f_- \in \mathcal{P}_-$, alors nécessairement $f_+(t) = \frac{1}{2}(f_+(t) + f_+(-t)) = \frac{1}{2}(f(t) + f(-t))$, et de même $f_-(t) = \frac{1}{2}(f(t) - f(-t))$.

PROPOSITION. *Soient V un espace vectoriel de dimension finie et U un sous-espace vectoriel de V. Il existe un sous-espace vectoriel U' de V tel que $V = U \oplus U'$.*

PREUVE. Soit $(u_1, \ldots, u_k)$ une base de U. On peut la compléter en une base $(u_1, \ldots, u_k, u_{k+1}, \ldots, u_n)$ de V. Il suffit alors de poser $U' = \langle u_{k+1}, \ldots, u_n \rangle$. □

Attention : le sous-espace U' *n'est en général pas déterminé uniquement* par U. Les seuls cas où il le soit sont celui où $U = \{0\}$ (et alors $U' = V$) et celui où $U = V$ (et alors $U' = \{0\}$). [Dans la situation d'un espace vectoriel réel de dimension finie *muni d'un produit scalaire*, nous verrons au chapitre VI qu'il y a un choix privilégié pour U', à savoir l'orthogonal de U ; mais, si $V \neq \{0\}$, le produit scalaire sur V n'est pas uniquement déterminé !]

Avec les notations de la proposition précédente, on dit que U' est un *sous-espace supplémentaire* de U dans V.

EXEMPLE. Si $V = \mathbb{R}^3$, et si U désigne le « plan horizontal » des vecteurs de troisième coordonnée nulle, alors *toute* droite non horizontale de $\mathbb{R}^3$ est un supplémentaire de U.

Considérons plus généralement une suite $(U_j)_{1 \leq j \leq k}$ de sous-espaces vectoriels d'un espace vectoriel V. Le sous-espace de V engendré par la réunion des U_j est la *somme* de ces sous-espaces ; on le note

$$U_1 + \cdots + U_k \qquad \text{ou} \qquad \sum_{j=1}^k V_j.$$

Tout vecteur de v s'écrit d'une manière au moins comme $v = \sum_{j=1}^k u_j$, avec $u_j \in U_j$ pour $j \in \{1, \ldots, k\}$. Lorsque cette écriture est unique pour tout $v \in V$, on dit que la somme est *directe* et on écrit

$$U_1 \oplus \cdots \oplus U_k \qquad \text{ou} \qquad \bigoplus_{j=1}^k U_j.$$

Quitte à présenter un exemple qui anticipe sur la matière d'un prochain chapitre, considérons un entier $n \geq 2$, l'espace $M_n(\mathbb{R})$ des matrices n-fois-n à coefficients réels et les trois sous-espaces vectoriels suivants :

le sous-espace $\mathbb{R}I_n$ des multiples scalaires de la matrice unité ;

le sous-espace $\mathrm{Sym}_n^{(0)}(\mathbb{R})$ des $(n \times n)$-matrices réelles symétriques à traces nulles ;

le sous-espace $\mathrm{Ant}_n(\mathbb{R})$ des $(n \times n)$-matrices réelles antisymétriques.

Alors

$$M_n(\mathbb{R}) \;=\; \mathbb{R}I_n \oplus \mathrm{Sym}_n^{(0)}(\mathbb{R}) \oplus \mathrm{Ant}_n(\mathbb{R}).$$

Voici un autre exemple, qui anticipe également sur la suite du cours. Soit a une matrice réelle carrée symétrique n-fois-n et $\lambda_1, \ldots, \lambda_k$ ses valeurs propres ; pour tout $j \in$

[6] Dans un premier temps, il est absolument essentiel de distinguer une fonction f de sa valeur $f(t)$ en un point t de l'espace source ; on suppose ici que le lecteur, suffisamment avisé, est prêt pour l'abus d'écriture systématiquement commis plus haut, consistant à écrire « la fonction $f(t)$ » au lieu, par exemple, de « la fonction $t \longmapsto f(t)$ ».

$\{1, \ldots, k\}$, soit $E_j = \mathrm{Ker}(\lambda_j - a) = \{x \in \mathbb{R}^n \mid ax = \lambda_j x\}$ l'espace propre de λ_j ; alors $\mathbb{R}^n = \bigoplus_{j=1}^n E_j$.

5. Sommes directes d'espaces vectoriels

REMARQUE. il est important que le titre du § II.4 contient le mot «SOUS-espace» et que le titre du présent paragraphe ne le contient pas.

Soient V, W deux espaces vectoriels. Sur l'ensemble produit $V \times W$, on définit une *somme* et un *produit par un nombre réel* par les formules

$$\begin{aligned}(v, w) + (v', w') &= (v + v', w + w') \\ \lambda(v, w) &= (\lambda v, \lambda w)\end{aligned}$$

pour tous $v, v' \in V$, $w, w' \in W$ et $\lambda \in \mathbb{R}$. On vérifie que $V \times W$ est un espace vectoriel pour ces opérations, qui s'appelle la *somme directe des espaces* V *et* W ; on le note $V \oplus W$ (ou parfois $V \times W$, voir plus bas).

Il est alors naturel d'identifier V au sous-espace de $V \oplus W$ des vecteurs de la forme $(v, 0)$, et de même W à celui des vecteurs de la forme $(0, w)$. *Une fois ces identifications faites,* l'espace $V \oplus W$ est somme directe de ses sous-espaces V et W au sens du paragraphe précédent.

Par exemple, $\mathbb{R}^n \times \mathbb{R}^n$ s'identifie à $\mathbb{R}^{2n}$.

De même, étant donné une famille finie $V_1, V_2, \ldots, V_k$ d'espaces vectoriels, on définit naturellement sur l'ensemble produit $V_1 \times \cdots \times V_k$ une addition et une multiplication par un nombre réel qui font de cet ensemble produit un espace vectoriel appelé la *somme directe* des V_j et noté

$$V_1 \oplus V_2 \oplus \cdots \oplus V_k.$$

PROPOSITION. *Avec les notations ci-dessus, la dimension de la somme directe*

$$V_1 \oplus V_2 \oplus \cdots \oplus V_k$$

est la somme des dimensions des espaces vectoriels $V_1, V_2, \ldots, V_k$.

PREUVE. Exercice. □

On peut s'étonner de ce que la *somme* directe des espaces vectoriels V et W soit un espace vectoriel dont l'ensemble sous-jacent est le *produit* des ensembles sous-jacents à V et W, mais c'est bien l'usage consacré ! En fait, les deux termes «somme directe» et «produit» (ou «produit direct») sont utilisés, indifféremment tant que les espaces en scène sont en nombre fini. Il existe des notions apparentées pour des familles infinies d'espaces vectoriels, et il s'agit alors de distinguer les sommes directes des produits directs – mais ceci ne concerne pas ce cours.

6. Quotients d'espaces vectoriels

Une *relation d'équivalence* sur un ensemble E est une partie R de $E \times E$ telle que, pour tous $x, y, z \in E$,

- $(x, x) \in R$ (réflexivité) ;
- si $(x, y) \in R$, alors $(y, x) \in R$ (symétrie) ;
- si (x, y) et (y, z) sont dans R, alors $(x, z) \in R$ (transitivité).

Pour $x \in E$, la *classe de x modulo R* est le sous-ensemble

$$[x]_R \;=\; \{y \in E \mid (x, y) \in R\}$$

de E. On vérifie que ces classes forment une *partition* de E, c'est-à-dire que

- tout élément de E appartient à une classe ;
- deux classes $[x]_R, [y]_R$ sont ou bien identiques ou bien *disjointes*.

On note E/R l'ensemble de ces classes, qui est un sous-ensemble de l'ensemble de toutes les parties de E (mais qui n'est en aucun cas un sous-ensemble de E !).

Prendre garde aux notations : pour $x \in E$, on a $[x]_R \subset E$ et $[x]_R \in E/R$. (Même distinction entre $\subset$ et $\in$ que, pour $I = \{0, 1, \ldots, 9\}$ et P l'ensemble des sous-ensembles de I, dans les deux écritures $\{0, 1\} \subset I$ et $\{0, 1\} \in P$.)

Soit V un espace vectoriel et W un sous-espace vectoriel de V. On définit une relation R_W sur V en posant, pour $x, y \in V$,

$$(x, y) \in R_W \quad \text{si} \quad y - x \in W,$$

et on vérifie que R_W est une relation d'équivalence sur W. On écrit V/W l'ensemble quotient (au lieu de V/R_W) et $[x]_W$ la classe dans V/W d'un vecteur $x \in V$ (au lieu de $[x]_{R_W}$). On définit alors sur l'ensemble V/W une *somme* et une *multiplication par un scalaire* en posant

$$\begin{aligned} [x]_W + [y]_W \;&=\; [x + y]_W \\ \lambda[x]_W \;&=\; [\lambda x]_W \end{aligned}$$

pour tous $[x]_W, [y]_W \in V/W$ et $\lambda \in \mathbb{R}$.

Exercice. Vérifier d'abord que ces définitions ont un sens, et ensuite qu'elles munissent l'ensemble V/W d'une structure d'espace vectoriel.

Définition. L'espace vectoriel V/W est le *quotient* de V par W.

Exemple. Soient p, q deux entiers positifs et $n = p + q$ leur somme. On identifie $\mathbb{R}^q$ au sous-espace de $\mathbb{R}^n$ des vecteurs dont les p dernières coordonnées (relativement à la base canonique) sont nulles. Alors le quotient $\mathbb{R}^n/\mathbb{R}^q$ s'identifie naturellement à $\mathbb{R}^p$.

Proposition. *Soient V un espace vectoriel de dimension finie et W son sous-espace vectoriel. Alors* $\dim_{\mathbb{R}} V/W = \dim_{\mathbb{R}} V - \dim_{\mathbb{R}} W$.

Preuve. Notons n (respectivement q) la dimension de V (resp. de W) et choisissons une base $\mathcal{B} = (b_1, \ldots, b_n)$ de V telle que $(b_1, \ldots, b_q)$ soit une base de W. Posons $c_i = [b_i]_W$ pour $i \in \{q + 1, \ldots, n\}$.

Affirmation : la suite $\mathcal{C} = (c_{q+1}, \ldots, c_n)$ est une base de V/W.

En effet, il est évident que $\mathcal{C}$ est une partie génératrice de V/W. Il reste donc à vérifier que c'est une partie libre. Soient $\lambda_{q+1}, \ldots, \lambda_n$ une suite de nombres réels tels

que $\sum_{i=q+1}^{n} \lambda_i c_i = 0$. Cette dernière égalité s'écrit $\left[\sum_{i=q+1}^{n} \lambda_i b_i\right]_W = 0_{V/W}$, ou encore $\sum_{i=q+1}^{n} \lambda_i b_i \in W$. Il existe donc des nombres $\lambda_1, \ldots, \lambda_q \in \mathbb{R}$ tels que $\sum_{i=q+1}^{n} \lambda_i b_i = \sum_{j=1}^{q} \lambda_j b_j$. Mais il résulte alors du fait que $\mathcal{B}$ est une base de V que les λ_i sont tous nuls. □

REMARQUE. Comparer cette preuve avec celle de la « formule de la dimension » au § III.6 ci-dessous.

7. Exercices

(II.1) Dans l'espace vectoriel $\mathcal{P}_3$ des fonctions $\mathbb{R} \longrightarrow \mathbb{R}$ de la forme $f : t \longmapsto a + bt + ct^2 + dt^3$, avec $a, b, c, d \in \mathbb{R}$, on considère les sous-ensembles définis par les conditions suivantes :

(i) f est de degré 3,

(ii) $f(2) = 3f(4)$,

(iii) $f(t) \geq 0$ pour tout $t \in [-1, 1]$,

(iv) la dérivée f' de f est paire,

(v) $f(1-t) = f(t)$ pour tout $t \in [0, 1]$,

(vi) $f(2) = 0$,

(vii) $f(2) = 1$,

(viii) $tf'(t) = f(t)$ pour tout $t \in \mathbb{R}$.

Sont-ils des sous-espaces vectoriels de $\mathcal{P}_3$? Lorsque *ce n'est pas le cas,* justifiez.

(II.2) Les ensembles de vecteurs suivants sont-ils linéairement indépendants ?

(i) $\begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ -1 \end{pmatrix} \in \mathbb{R}^4,$

(ii) $\begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ -1 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 0 \\ 1 \end{pmatrix} \in \mathbb{R}^4,$

(iii) $t \longmapsto t$, $t \longmapsto \frac{1}{t}$ dans l'espace $\mathcal{C}(\mathbb{R}^*, \mathbb{R})$ des fonctions continues à valeurs réelles sur $\mathbb{R}^*$.

(II.3) Dans l'espace vectoriel des fonctions continues de $\mathbb{R}$ dans $\mathbb{R}$, montrer que les fonctions $t \longmapsto e^{ct}$ $(c \in \mathbb{R})$ sont linéairement indépendantes.

(II.4) Quelles sont les nombres λ tels que $\{(1+\lambda, 1-\lambda), (1-\lambda, 1+\lambda)\}$ est une base de $\mathbb{R}^2$?

(II.5) Soit V le sous-espace vectoriel de $\mathbb{R}^n$ formé des vecteurs dont la somme des coordonnées relativement à la base canonique est nulle. Ecrire une base de V.

(II.6) Etant donné un espace vectoriel V de dimension finie et un sous-espace vectoriel U de V, on considère les assertions suivantes. Sont-elles correctes ?

(i) Toute famille génératrice de V peut être complétée en une base ;

(ii) toute famille libre de V peut être complétée en une base ;

(iii) de toute famille génératrice de V, on peut extraire une base ;

(iv) de toute famille libre de V, on peut extraire une base ;

(v) $V \setminus \{0\}$ est une famille génératrice ;

(vi) $V \setminus \{0\}$ est une famille libre ;

(vii) de toute famille génératrice de V, on peut extraire une famille génératrice de U ;

(viii) toute famille de U libre dans U est libre dans V ;

(ix) toute base de U peut être complétée en une base de V.

(II.7) Dans l'espace $\mathcal{P}_d$ des fonctions polynomiales $\mathbb{R} \longrightarrow \mathbb{R}$ de degré au plus d, montrer que les *polynômes de Legendre*

$$P_j(t) \;=\; \frac{1}{2^j j!}\frac{d^j}{dt^j}\left((t^2-1)^j\right) \qquad 0 \le j \le d$$

constituent une base.

(II.8) Pour $k \in \{-2,-1,0,1,2\}$, on note f_k la fonction $\mathbb{R} \longrightarrow \mathbb{R}$ définie par

$$f_k(t) \;=\; \left(\prod_{-2\le j\le 2\,,\, j\neq k} (k-j)\right)^{-1} \prod_{-2\le j\le 2\,,\, j\neq k} (t-j)\,;$$

par exemple $f_{-2}(t) = \frac{1}{24}(t+1)t(t-1)(t-2)$.

Calculer $f_k(l)$ pour $k, l \in \{-2,-1,0,1,2\}$ et montrer que

$$\{f_{-2}(t), f_{-1}(t), f_0(t), f_1(t), f_2(t)\}$$

est une base de l'espace vectoriel $\mathcal{P}_4$ des fonctions polynomiales de degrés au plus 4.

Relativement à cette base, écrire les coordonnées d'un vecteur $f \in \mathcal{P}_4$ en fonction des valeurs $f(-2)$, $f(-1)$, $f(0)$, $f(1)$ et $f(2)$.

(II.9) Dans $\mathbb{R}^3$, on considère le sous-espace vectoriel U_1 des vecteurs de la forme $(x,0,0)$, et de même les sous-espaces U_2, U_3, ainsi que le sous-espace $U_{1,2}$ des vecteurs de la forme $(x,y,0)$, et de même les sous-espaces $U_{1,3}$, $U_{2,3}$. Indiquer lesquelles des relations suivantes sont correctes.

(i) $U_1 + U_2 = U_{1,2}$

(ii) $U_1 + U_{2,3} = \mathbb{R}^3$

(iii) $U_{1,2} + U_{2,3} = \mathbb{R}^3$

(iv) $U_1 \oplus U_2 = U_{1,2}$

(v) $\{0\} \oplus \mathbb{R} = \mathbb{R}$

(vi) $U_1 + U_1 = U_1$

(vii) $U_1 + U_{1,2} = U_{1,2}$

(viii) $U_1 + U_2 + U_3 = \mathbb{R}^3$

(ix) $U_{1,2} \oplus U_{2,3} = \mathbb{R}^3$

(x) $\{x \in \mathbb{R} \mid x \leq 0\} \oplus \{x \in \mathbb{R} \mid x > 0\} = \mathbb{R}$

(II.10) Dans l'espace $\mathcal{P}_5$ des applications polynomiales $\mathbb{R} \longrightarrow \mathbb{R}$ de degrés au plus 5, définir *plusieurs* espaces supplémentaires du sous-espace vectoriel U formé des applications de la forme $t \longmapsto at^3 + bt^4$ (où $a, b \in \mathbb{R}$).

(II.11) Les assertions suivantes sont-elles vraies on non ?

(i) Si $(b_1, \ldots, b_n)$ est une base d'un espace vectoriel V (où $n \geq 2$) , alors $(b_1 - b_2, b_2 - b_3, \ldots, b_{n-1} - b_n, b_n)$ est aussi une base de V.

(ii) Il existe une base de $\mathcal{P}_3$ qui ne contienne aucune fonction polynomiale de degré 2.

(iii) Si V est un espace vectoriel de dimension n, il existe n sous-espaces vectoriels $U_1, \ldots, U_n$ de V, chacun de dimension 1, tels que $V = U_1 \oplus \cdots \oplus U_n$.

(iv) Il existe une base (f_0, f_1, f_2, f_3) de $\mathcal{P}_3$ telle que $f_j(2) = 0$ pour tout $j \in \{0, 1, 2, 3\}$.

CHAPITRE III

Applications linéaires

1. Matrices

Il s'agit essentiellement d'un paragraphe de définitions et de notations.

Soient X un ensemble et m, n deux entiers strictement positifs. Une *matrice à coefficients dans X à m lignes et n colonnes* est un tableau $(a_{i,j})_{1\leq i\leq m, 1\leq j\leq n}$ d'éléments de X. L'ensemble des matrices de ce type est noté $M_{m,n}(X)$; par exemple :

$$\begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \end{pmatrix} \in M_{2,3}(X).$$

A la notation près, $M_{m,n}(X)$ est donc l'ensemble X^{mn} produit de mn copies de l'ensemble X. Lorsque $m = n$, on écrit aussi $M_m(X)$ au lieu de $M_{m,m}(X)$. Les cas les plus importants sont $X = \mathbb{R}$ et $X = \mathbb{C}$.

Pour $m, n \geq 0$, l'ensemble $M_{m,n}(\mathbb{R})$ est un espace vectoriel pour les opérations définies par

$$\begin{aligned} (a_{i,j})_{1\leq i\leq m, 1\leq j\leq n} + (b_{i,j})_{1\leq i\leq m, 1\leq j\leq n} &= (a_{i,j} + b_{i,j})_{1\leq i\leq m, 1\leq j\leq n} \\ \lambda \left((a_{i,j})_{1\leq i\leq m, 1\leq j\leq n} \right) &= (\lambda a_{i,j})_{1\leq i\leq m, 1\leq j\leq n}. \end{aligned}$$

Pour toute paire d'entiers $k \in \{1, \ldots, m\}$ et $l \in \{1, \ldots, n\}$, on définit la *matrice élémentaire* $E_{k,l} \in M_{m,n}(\mathbb{R})$ comme étant la matrice ayant un 1 au croisement de la $k$ème ligne et de la $l$ème colonne, et des 0 partout ailleurs. (Attention, dans l'écriture $E_{k,l}$, les indices k et l ont pour rôle de distinguer les matrices entre elles, par exemple la matrice $E_{1,2}$ de la matrice $E_{2,2}$; au contraire de l'écriture $(a_{i,j})_{1\leq i\leq m, 1\leq j\leq n}$, où les indices ont pour rôle de distinguer divers coefficients d'une même matrice). Par exemple, dans $M_{3,4}(\mathbb{R})$, la matrice $E_{3,2}$ s'écrit

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

On vérifie que les mn matrices $E_{k,l}$ constituent une base de $M_{m,n}(\mathbb{R})$, de sorte que $M_{m,n}(\mathbb{R})$ *est un espace vectoriel de dimension mn.*

La *transposée* d'une matrice $a = (a_{i,j})_{1\leq i\leq m, 1\leq j\leq n} \in M_{m,n}(X)$ est la matrice ${}^t a$ de $M_{n,m}(X)$ définie par

$$({}^t a)_{j,i} = a_{i,j} \qquad \text{pour tous } j \in \{1, \ldots, n\} \text{ et } i \in \{1, \ldots, m\}.$$

Par exemple,

$$\text{la transposée de la matrice } \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \text{ est la matrice } \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix}.$$

Notons que, pour toute matrice a, la transposée de ${}^t a$ est la matrice a ; en formule : ${}^t({}^t a) = a$.

Une matrice carrée $a \in M_n(\mathbb{R})$ est *symétrique* si ${}^t a = a$, *antisymétrique* si ${}^t a = -a$, et *diagonale* si $a_{i,j} = 0$ lorsque $i \neq j$; noter qu'une matrice diagonale est nécessairement symétrique.

La *matrice unité* dans $M_n(\mathbb{R})$ est la matrice dont les coefficients diagonaux sont 1 et les autres coefficients 0. Elle est notée 1, ou I, ou I_n.

PROPOSITION. *Soit n un entier, $n \geq 1$. Dans l'espace vectoriel $M_n(\mathbb{R})$:*

(i) *le sous-ensemble des matrices symétriques constitue un sous-espace vectoriel de dimension $n(n+1)/2$;*

(ii) *le sous-ensemble des matrices antisymétriques constitue un sous-espace vectoriel de dimension $n(n-1)/2$;*

(iii) *le sous-ensemble des matrices diagonales constitue un sous-espace vectoriel de dimension n.*

PREUVE. Exercice. □

Produits de matrices

Le produit d'une matrice $a = (a_{i,j})_{1 \leq i \leq l, 1 \leq j \leq m} \in M_{l,m}(\mathbb{R})$ et d'une matrice $b = (b_{j,k})_{1 \leq j \leq m, 1 \leq k \leq n} \in M_{m,n}(\mathbb{R})$ est la matrice $c = (c_{i,k})_{1 \leq i \leq l, 1 \leq k \leq n} \in M_{l,n}(\mathbb{R})$ définie par

$$c_{i,k} = \sum_{j=1}^{m} a_{i,j} b_{j,k}.$$

Ainsi, le produit des matrices fournit une application

$$M_{l,m}(\mathbb{R}) \times M_{m,n}(\mathbb{R}) \longrightarrow M_{l,n}(\mathbb{R}) \qquad (a, b) \longmapsto ab.$$

Noter que le produit ab n'est défini *que si le nombre de colonnes de a est égal au nombre de lignes de b.*

La justification de cette définition du produit apparaît au théorème principal du § III.4.

Voici d'abord deux exemples numériques :

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 2 & -2 \\ 3 & -3 \end{pmatrix} = \begin{pmatrix} 14 & -14 \\ 32 & -32 \end{pmatrix}$$

et

$$\begin{pmatrix} 1 & -1 \\ 2 & -2 \\ 3 & -3 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} = \begin{pmatrix} -3 & -3 & -3 \\ -6 & -6 & -6 \\ -9 & -9 & -9 \end{pmatrix}.$$

Vu l'indentification de $M_{n,1}(\mathbb{R})$ avec $\mathbb{R}^n$, on a en particulier un produit

$$M_{m,n}(\mathbb{R}) \times \mathbb{R}^n \longrightarrow \mathbb{R}^m \qquad (a,x) \longmapsto ax.$$

Ceci justifie l'écriture

$$ax = b$$

d'un système linéaire à m équations et n inconnues (la matrice $a \in M_{m,n}(\mathbb{R})$ et le vecteur $b \in \mathbb{R}^m$ sont donnés, et les inconnues sont les composantes $x_1, \ldots, x_n$ du vecteur $x \in \mathbb{R}^n$).

Mentionnons enfin les exemples très particuliers de multiplications avec des matrices unité : si $a \in M_{m,n}(\mathbb{R})$, alors $I_m a = a = a I_n$.

Attention : si a, b sont deux matrices telles que les deux produits ab et ba sont définis, ces deux produits sont en général distincts, comme le montre un exemple ci-dessus. En voici un autre :

$$\begin{pmatrix} 2 & 3 \\ 4 & 6 \end{pmatrix} \begin{pmatrix} 3 & -3 \\ -2 & 2 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \neq \begin{pmatrix} -6 & -9 \\ 4 & 6 \end{pmatrix} = \begin{pmatrix} 3 & -3 \\ -2 & 2 \end{pmatrix} \begin{pmatrix} 2 & 3 \\ 4 & 6 \end{pmatrix}.$$

Exercice. Soient a, b deux matrices telles que le produit ab est défini ; montrer que ${}^t(ab) = {}^t b\, {}^t a$, et noter l'ordre des facteurs dans le membre de droite.

Il est élémentaire[1] de vérifier les identités « naturelles » . Par exemple, si a, b, c sont trois matrices telles que l'un des produits $(ab)c$, $a(bc)$ est défini, alors l'autre l'est aussi et ces deux produits sont égaux. De même pour les deux expressions $a(b+c)$ et $ab+ac$, ou $(a+b)c$ et $ac+bc$, ou encore pour les trois expressions $\lambda(ab)$, $(\lambda a)b$ et $a(\lambda b)$ lorsque λ est un nombre réel.

[Remarque pour une *deuxième lecture*. Pour une matrice $a \in M_{l,m}(\mathbb{R})$ fixée, l'application

$$M_{m,n}(\mathbb{R}) \longrightarrow M_{l,n}(\mathbb{R}), \qquad b \longmapsto ab$$

de prémultiplication par a est linéaire. De même, pour $b \in M_{m,n}(\mathbb{R})$ fixé, l'application

$$M_{l,m}(\mathbb{R}) \longrightarrow M_{l,n}(\mathbb{R}), \qquad a \longmapsto ab$$

de postmultiplication par b est linéaire.]

Une matrice carrée $a \in M_n(\mathbb{R})$ est dite *inversible* s'il existe des matrices $b, c \in M_n(\mathbb{R})$ telles que $ba = ac = I_n$. Dans ce cas, on a de plus $c = I_n c = (ba)c = b(ac) = bI_n = b$; il en résulte aussi qu'il y a au plus une telle matrice b (ou c !). On écrit a^{-1} au lieu de b ou c. [Dans la suite du cours, nous montrerons que, s'il existe une matrice $b \in M_n(\mathbb{R})$ telle que $ba = I_n$, alors a est inversible et $a^{-1} = b$; de même s'il existe une matrice $c \in M_n(\mathbb{R})$ telle que $ac = I_n$, alors a est inversible et $a^{-1} = c$.]

Exercices. Si $a, b \in M_n(\mathbb{R})$ sont deux matrices inversibles, montrer que ab est aussi inversible et que $(ab)^{-1} = b^{-1}a^{-1}$. Noter l'ordre des facteurs dans le membre de droite.

Soit $a = \begin{pmatrix} \lambda & \mu \\ 0 & \nu \end{pmatrix} \in M_2(\mathbb{R})$; montrer que a est inversible si et seulement si $\lambda \neq 0$ et $\nu \neq 0$. [En cas de difficulté, supposez d'abord que $\mu = 0$.]

[1] C'est aussi fastidieux. Il est avantageux de tirer parti du théorème du § III.4 qui suit.

REMARQUE. Considérons un système linéaire à n équations et n inconnues, écrit sous la forme $ax = b$. Si la matrice a est inversible, les solutions du système s'écrivent évidemment $x = a^{-1}b$. En pratique, le calcul de a^{-1} est parfois difficile.

2. Applications linéaires : généralités

Il s'agit à nouveau d'un paragraphe de définitions et de notations.

Soient U, V deux espaces vectoriels. Une *application linéaire* de U dans V est une application $\alpha : U \longrightarrow V$ telle que

$$\begin{aligned} \alpha(u+u') &= \alpha(u)+\alpha(u') && \text{pour tous } u, u' \in U, \\ \alpha(\lambda u) &= \lambda\alpha(u) && \text{pour tous } u \in U \text{ et } \lambda \in \mathbb{R}. \end{aligned}$$

Noter que, nécessairement, $\alpha(0) = 0$ pour une telle application. On note $\mathcal{L}(U, V)$ l'ensemble des applications linéaires de U dans V ; on écrit $\mathcal{L}(U)$ au lieu de $\mathcal{L}(U, U)$.

Pour U et V comme ci-dessus, $\mathcal{L}(U, V)$ est un espace vectoriel pour les opérations définies par

$$\begin{aligned} (\alpha+\alpha')(u) &= \alpha(u)+\alpha'(u) \\ (\lambda\alpha)(u) &= \lambda\alpha(u) \end{aligned}$$

pour tous $\alpha, \alpha' \in \mathcal{L}(U, V)$, $\lambda \in \mathbb{R}$ et $u \in U$.

Pour les applications linéaires, la terminologie suivante est courante. Si $\alpha \in \mathcal{L}(U, V)$, on dit aussi que α est un *homomorphisme linéaire* de U dans V ; si α est de plus bijective, c'est un *isomorphisme linéaire*. Lorsque la source U et le but V d'un homomorphisme linéaire coïncident, c'est à dire lorsque $U = V$, on dit que α est un *endomorphisme linéaire* de V ; un endomorphisme linéaire qui est bijectif est un *automorphisme linéaire.*

Les exemples qui suivent font appel à des notions qui n'ont pas encore été introduites dans le présent cours (angles, projections orthogonales, symétries orthogonales), mais qui font partie de la culture générale (au moins en petites dimensions) ; contrairement à tout le formalisme mis en place en ce début de cours, ces notions ne se généralisent pas aux cas d'espaces vectoriels et d'applications linéaires sur d'autres corps que $\mathbb{R}$ (ou, avec des modifications mineures, $\mathbb{C}$).

EXEMPLES. (i) Considérons d'abord des endomorphismes de $\mathbb{R}^2$, cas où $U = V = \mathbb{R}^2$. La projection sur le premier axe $\begin{pmatrix} x \\ y \end{pmatrix} \longmapsto \begin{pmatrix} x \\ 0 \end{pmatrix}$ est un endomorphisme linéaire de $\mathbb{R}^2$. La symétrie relativement au premier axe $\begin{pmatrix} x \\ y \end{pmatrix} \longmapsto \begin{pmatrix} x \\ -y \end{pmatrix}$ est un automorphisme linéaire de $\mathbb{R}^2$. Plus généralement, la projection orthogonale sur toute droite du plan passant par l'origine et la symétrie orthogonale relativement à une telle droite sont des endomorphismes de $\mathbb{R}^2$. Une rotation du plan autour de l'origine est un automorphisme de $\mathbb{R}^2$. Une translation d'amplitude non nulle *n'est pas* une application linéaire, par exemple parce que l'image de l'origine n'est pas l'origine.

(ii) Notons $\mathcal{P}$ l'espace des applications polynomiales de $\mathbb{R}$ dans $\mathbb{R}$. Les trois applications de $\mathcal{P}$ dans $\mathcal{P}$ définies par

$$\begin{aligned} f &\longmapsto f' && \text{dérivation} \\ f(t) &\longmapsto \int_0^t f(\tau)d\tau && \text{primitive nulle en 0} \\ f(t) &\longmapsto t^2 f(t) && \text{multiplication par } t^2 \end{aligned}$$

sont linéaires[2]. En revanche, les applications

$$\begin{aligned} f &\longmapsto |f(0)| \\ f &\longmapsto \max_{-1\leq t\leq 1} |f(t)| \\ f(t) &\longmapsto f(t)^2 \end{aligned}$$

ne sont pas linéaires.

(iii) Considérons le cas où $U = \mathbb{R}^m$ et $V = \mathbb{R}^n$; soit $a \in M_{m,n}(\mathbb{R})$. L'application

$$\alpha : \mathbb{R}^n \longrightarrow \mathbb{R}^m, \qquad x \longmapsto ax$$

est linéaire. Son étude est essentiellement équivalente à celle des systèmes linéaires de la forme $ax = b$.

(iv) Soit U un espace vectoriel de dimension finie m et $b_1, \ldots, b_m$ une base de U. Pour tout $i \in \{1, \ldots, m\}$, l'application

$$U \longrightarrow \mathbb{R}, \qquad u = \sum_{j=1}^{m} u_j b_j \longmapsto u_i$$

est linéaire. On appelle *forme linéaire* sur un espace vectoriel U une application linéaire de U dans $\mathbb{R}$.

Composition d'applications linéaires

Soient U, V, W trois espaces vectoriels et $\alpha \in \mathcal{L}(U, V)$, $\beta \in \mathcal{L}(V, W)$ deux applications linéaires (noter que l'espace but de α est identiques à l'espace source de β). On vérifie que la composition $\beta\alpha : U \longrightarrow W$ est une application linéaire. Ainsi, la composition fournit une application

$$\mathcal{L}(U, V) \times \mathcal{L}(V, W) \longrightarrow \mathcal{L}(U, W), \qquad (\alpha, \beta) \longmapsto \beta\alpha.$$

[Remarque pour une *deuxième lecture*. Pour une application linéaire $\alpha \in \mathcal{L}(U, V)$ fixée, l'application

$$\mathcal{L}(V, W) \longrightarrow \mathcal{L}(U, W), \qquad \beta \longmapsto \beta\alpha$$

de précomposition avec α est linéaire. De même, pour $\beta \in \mathcal{L}(V, W)$ fixé, l'application

$$\mathcal{L}(U, V) \longrightarrow \mathcal{L}(U, W), \qquad \alpha \longmapsto \beta\alpha$$

de postcomposition avec β est linéaire.]

Dans la composition, noter l'ordre d'écriture $\beta\alpha$, conforme à l'écriture déjà rencontrée au § I.3.

[2] On a commis ici l'abus de désigner une application $f \in \mathcal{P}$ par sa valeur $f(t)$ en un point $t \in \mathbb{R}$.

EXERCICES. Soient U, V, W trois espaces vectoriels et $\alpha \in \mathcal{L}(U,V)$, $\beta \in \mathcal{L}(V,W)$ deux applications linéaires.

Montrer que, si les applications α et β sont injectives, alors l'application $\beta\alpha$ l'est aussi. Exhiber un exemple montrant que $\beta\alpha$ peut être injective sans que β le soit.

Montrer que, si les applications α et β sont surjectives, alors l'application $\beta\alpha$ l'est aussi. Exhiber un exemple montrant que $\beta\alpha$ peut être surjective sans que α le soit.

Montrer que, si les applications α et β sont bijectives, alors l'application $\beta\alpha$ l'est aussi (cette assertion est utilisée plus bas dans la preuve du dernier corollaire du § III.3). Exhiber un exemple montrant que $\beta\alpha$ peut être bijective sans que ni α ni β le soit.

PROPRIÉTÉS DE LA COMPOSITION DES APPLICATIONS LINÉAIRES. Soient $\alpha, \alpha', \alpha'' \in \mathcal{L}(U,V)$, $\beta, \beta', \beta'' \in \mathcal{L}(V,W)$ et $\lambda \in \mathbb{R}$. Alors

$$\begin{aligned}(\beta' + \beta'')\alpha &= \beta'\alpha + \beta''\alpha \\ \beta(\alpha' + \alpha'') &= \beta\alpha' + \beta\alpha'' \\ (\lambda\beta)\alpha &= \beta(\lambda\alpha) = \lambda(\beta\alpha).\end{aligned}$$

PREUVE. Exercice. □

Sur l'ordre de composition des applications linéaires

Deux endomorphismes α, β d'un même espace vectoriel V peuvent être composés de deux manières ; en général, les compositions $\beta\alpha$ et $\alpha\beta$ ne coïncident pas, comme le montre l'exemple suivant. Soit $V = \mathbb{R}^2$; notons α la projection de $\mathbb{R}^2$ sur le premier axe et β la rotation d'un quart de tour dans le sens positif (contraire à celui des aiguilles d'une montre). Si (e_1, e_2) désigne la base canonique de $\mathbb{R}^2$, on a par exemple

$$\beta\alpha(e_1) = \beta(e_1) = e_2 \qquad \text{et} \qquad \alpha\beta(e_1) = \alpha(e_2) = 0,$$

ce qui montre bien que $\beta\alpha \neq \alpha\beta$ (l'inégalité a un sens dans $\mathcal{L}(\mathbb{R}^2)$).

Plus spectaculairement, soient U, V deux espaces vectoriels distincts et $\alpha \in \mathcal{L}(U,V)$, $\beta \in \mathcal{L}(V,U)$. Les compositions $\beta\alpha$ et $\alpha\beta$ sont toutes deux définies, mais la première appartient à l'espace $\mathcal{L}(U)$ alors que la seconde appartient à un autre espace, à savoir $\mathcal{L}(V)$; il n'y aurait donc aucun sens à écrire une égalité entre elles.

Inverses des isomorphismes linéaires

Soient U, V deux espaces vectoriels et $\alpha : U \longrightarrow V$ une application linéaire.

Supposons d'abord l'application α bijective, de sorte que l'application inverse $\alpha^{-1} : V \longrightarrow U$ est bien définie (voir le chapitre I). Alors l'application α^{-1} est linéaire. En effet, soient d'abord $v_1, v_2 \in V$. Par hypothèse de bijectivité, il existe deux vecteurs bien définis $u_1, u_2 \in U$ tels que $\alpha(u_1) = v_1$ et $\alpha(u_2) = v_2$; notons que $\alpha^{-1}(v_1) = u_1$ et $\alpha^{-1}(v_2) = u(_2)$, par définition de α^{-1}. Comme $\alpha(u_1 + u_2) = v_1 + v_2$ par linéarité de α, on a aussi

$$\alpha^{-1}(v_1 + v_2) = u_1 + u_2 = \alpha^{-1}(v_1) + \alpha^{-1}(v_2).$$

On montre de même que

$$\alpha^{-1}(\lambda v) = \lambda\alpha^{-1}(v)$$

pour tous $\lambda \in \mathbb{R}$ et $v \in V$. Ceci achève de montrer que α^{-1} est linéaire. Notons les compositions

$$\alpha^{-1}\alpha = \mathrm{id}_U \in \mathcal{L}(U) \qquad \text{et} \qquad \alpha\alpha^{-1} = \mathrm{id}_V \in \mathcal{L}(V)$$

où id_U désigne l'application «identique» $u \longmapsto u$ de U dans U, et de même pour id_V.

Réciproquement, s'il existe une application linéaire $\beta \in \mathcal{L}(V,U)$ telle que $\beta\alpha = \mathrm{id}_U$ et $\alpha\beta = \mathrm{id}_V$, alors l'application linéaire α est bijective[3] et $\beta = \alpha^{-1}$.

3. Applications linéaires et bases

Voici le résultat principal du paragraphe.

Proposition. *Soient U, V deux espaces vectoriels. On suppose U de dimension finie, notée m, et on choisit une base $(b_1, \ldots, b_m)$ de U.*

Pour toute famille $(v_1, \ldots, v_m)$ de vecteurs de V, il existe une unique application linéaire $\alpha \in \mathcal{L}(U,V)$ telle que $\alpha(b_i) = v_i$ pour tout $i \in \{1, \ldots, m\}$.

Preuve de l'existence. Tout vecteur $u \in U$ s'écrit de manière unique $u = \sum_{i=1}^m u_i b_i$. Il suffit de poser $\alpha(u) = \sum_{i=1}^m u_i v_i$; on vérifie sans peine que α est linéaire et applique b_i sur v_i pour tout $i \in \{1, \ldots, m\}$.

Preuve de l'unicité. Soit $\beta \in \mathcal{L}(U,V)$ une application linéaire telle que $\beta(b_i) = v_i$ pour tout $i \in \{1, \ldots, m\}$. Pour tous $u_1, \ldots, u_m \in \mathbb{R}$, on a nécessairement

$$\beta\left(\sum_{i=1}^m u_i b_i\right) = \sum_{i=1}^m u_i \beta(b_i) = \sum_{i=1}^m u_i v_i \,;$$

par suite, $\beta(u) = \alpha(u)$ pour tout $u \in U$, c'est-à-dire $\beta = \alpha$. □

Exemples (avec $U = V = \mathbb{R}^2$). Notons (e_1, e_2) la base canonique de $\mathbb{R}^2$.

Si $v_1 = e_1$ et $v_2 = 0$, alors α est la projection sur le premier axe.

Si $v_1 = e_1$ et $v_2 = e_2$, alors α est l'identité de $\mathbb{R}^2$.

Si $v_1 = v_2 = 0$, alors α applique tout vecteur de $\mathbb{R}^2$ sur l'origine ($\alpha = 0$).

Si $v_1 = e_2$ et $v_2 = -e_1$, alors α est la rotation d'un quart de tour dans le sens positif.

Si $v_1 = e_2$ et $v_2 = e_1$, alors α est la symétrie orthogonale fixant la première bissectrice.

Si $v_1 = v_2 = e_1$, alors α est la projection oblique sur le premier axe qui applique $e_2 - e_1$ sur 0.

Corollaire. *Les espaces vectoriels $\mathcal{L}(U,V)$ et $V \oplus \cdots \oplus V$ (m facteurs) sont isomorphes, et par suite*

$$\dim_{\mathbb{R}}(\mathcal{L}(U,V)) = \dim_{\mathbb{R}}(U)\dim_{\mathbb{R}}(V).$$

[3] Preuve de l'injectivité : si $u_1, u_2 \in U$ sont tels que $u_1 \neq u_2$, alors $\beta(\alpha(u_1)) \neq \beta(\alpha(u_2))$, de sorte que $\alpha(u_1) \neq \alpha(u_2)$. Preuve de la surjectivité : $\alpha(\beta(v)) = v$ pour tout $v \in V$.

PREUVE. Il est immédiat que l'application

$$\mathcal{L}(U,V) \longrightarrow V \oplus \cdots \oplus V, \qquad \alpha \longmapsto (\alpha(b_1), \ldots, \alpha(b_2))$$

est linéaire. La proposition précédente montre que c'est un isomorphisme linéaire.

L'assertion sur les dimensions résulte alors de ce que la dimension est additive par somme directe. □

REMARQUE. Il faut considérer $V \oplus \cdots \oplus V$ comme une somme directe de m espaces vectoriels distincts, copies de V (comme les différents facteurs $\mathbb{R}$ dans l'écriture $\mathbb{R}^m$). En notations plus explicites, on pourrait envisager chaque «copie» de V comme un espace vectoriel $V^{(i)}$ donné avec un isomorphisme $\varphi^{(i)} : V \longrightarrow V^{(i)}$, et l'application de la preuve précédente écrite sous la forme

$$\mathcal{L}(U,V) \longrightarrow V^{(1)} \oplus \cdots \oplus V^{(m)}, \qquad \alpha \longmapsto \left(\varphi^{(1)}(\alpha(b_1)), \ldots, \varphi^{(m)}(\alpha(b_2))\right).$$

COROLLAIRE. *Deux espaces vectoriels de dimension finie sont isomorphes si et seulement s'ils ont même dimension.*

En particulier, tout espace vectoriel de dimension finie n est isomorphe à $\mathbb{R}^n$.

PREUVE. Soient U, V deux espaces vectoriels de dimensions finies.

Supposons d'abord que les dimensions de U et V sont égales à un même nombre n. Choisissons une base $(a_1, \ldots, a_n)$ de U et une base $(b_1, \ldots, b_n)$ de V. En appliquant deux fois l'assertion d'existence de la proposition précédente, on définit deux applications linéaires $\varphi : U \longrightarrow V$ et $\psi : V \longrightarrow U$ telles que $\varphi(a_i) = b_i$ et $\psi(b_i) = a_i$ pour tout $i \in \{1, \ldots, n\}$. Vu que $\psi(\varphi(a_i)) = a_i$ et $\varphi(\psi(b_i)) = b_i$ pour tout $i \in \{1, \ldots, n\}$, l'assertion d'unicité de la proposition implique que $\psi\varphi = \mathrm{id}_U$ et $\varphi\psi = \mathrm{id}_V$. Par suite, φ est bien un isomorphisme de U sur V.

Réciproquement, supposons qu'il existe un isomorphisme $\varphi : U \longrightarrow V$. Si $(a_1, \ldots, a_n)$ est une base de U, on vérifie que $(\varphi(a_1), \ldots, \varphi(a_n))$ est une base de V, donc en particulier que U et V ont la même dimension.

Dans le cas où $V = \mathbb{R}^n$, l'application linéaire

$$U \longrightarrow \mathbb{R}^n, \qquad u = \sum_{i=1}^{n} u_i a_i \longmapsto (u_i)_{1 \le i \le n}$$

est un isomorphisme. C'est ici la base canonique de $\mathbb{R}^n$ qui joue le rôle de $(b_1, \ldots, b_n)$. □

4. Applications linéaires et matrices

Matrice d'une application linéaire relativement à des bases données

Soient V, W deux espaces vectoriels de dimension finie, $\mathcal{B} = (b_1, \ldots, b_m)$ une base de V et $\mathcal{C} = (c_1, \ldots, c_n)$ une base de W.

Si $\sigma \in \mathcal{L}(W,V)$ est une application linéaire (attention : c'est ici W qui est l'espace source et V l'espace but !!!),

$$\sigma(c_j) = \sum_{i=1}^{m} s_{i,j} b_i \qquad \text{(écriture unique)} \quad \text{pour tout } j \in \{1, \ldots, n\}.$$

La matrice $(s_{i,j})_{1\leq i\leq m, 1\leq j\leq n}$ est par définition la *matrice de σ relativement aux bases $\mathcal{B}$ et $\mathcal{C}$;* on écrit

$$(s_{i,j})_{1\leq i\leq m,1\leq j\leq n} \;=\; M_{\mathcal{B},\mathcal{C}}(\sigma) \;\in\; M_{m,n}(\mathbb{R}).$$

C'est donc **la matrice dont les colonnes sont les coordonnés dans la base au but des images par σ des vecteurs de la base à la source.**

Difficulté : ne pas confondre les indices lignes, correspondant à l'espace *but* V de σ, et les indices colonnes, correspondant à l'espace *source* W de σ. (Noter que W [respectivement $\mathcal{C}$, w, j, n] est dans l'alphabet le successeur de V [resp. $\mathcal{B}$, v, i, m].)

Exemples. (i) Considérons le cas où $V = W = \mathbb{R}^2$ et où $\mathcal{B} = \mathcal{C}$ est la base canonique $\mathcal{C}an = (e_1, e_2)$.

Notons ρ_θ la rotation d'angle θ dans le sens positif. Notons d_θ une droite passant par l'origine de $\mathbb{R}^2$ et faisant avec le premier axe un angle θ (c'est donc l'image du premier axe par ρ_θ) ; notons encore σ_θ la symétrie orthogonale relativement à θ et π_θ la projection orthogonale d'image d_θ.

Introduisons la base $\mathcal{C} = (\rho_\theta(e_1), \rho_\theta(e_2))$, image de la base canonique par la rotation ρ_θ.

Certaines des matrices associées aux applications $\rho_\theta, \sigma_\theta, \pi_\theta$ s'écrivent sans aucune peine. Ainsi :

$$M_{\mathcal{C},\mathcal{C}an}(\rho_\theta) \;=\; \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad M_{\mathcal{C},\mathcal{C}}(\sigma_\theta) \;=\; \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad \text{et} \quad M_{\mathcal{C},\mathcal{C}}(\pi_\theta) \;=\; \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$$

(la première de ces égalités montre que la matrice identité peut représenter autre chose que l'application identique). D'autre part, les propriétés des sinus et cosinus montrent que

$$M_{\mathcal{C},\mathcal{C}an}(\sigma_\theta) \;=\; \begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix} \quad \text{et} \quad M_{\mathcal{C},\mathcal{C}an}(\pi_\theta) = \begin{pmatrix} \cos\theta & \sin\theta \\ 0 & 0 \end{pmatrix}.$$

Si h_2 désigne l'homothétie centrée à l'origine de rapport 2, alors

$$M_{\mathcal{C}an,\mathcal{C}an}(h_2) \;=\; \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}.$$

En fait $M_{\mathcal{B},\mathcal{B}}(h_2) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$ pour *toute* base $\mathcal{B}$ de $\mathbb{R}^2$!

(ii) Soient maintenant $V = W$ l'espace $\mathcal{P}_n$ des fonctions polynomiales de degrés au plus n, et soit $\mathcal{B}$ la base[4] de $\mathcal{P}_n$ constituée des fonctions $t \longmapsto t^j$ $(0 \leq j \leq n)$. Si σ

[4]Dans un espace vectoriel de dimension n, les vecteurs d'une base sont souvent indexés par les entiers de 1 à n ; mais ceci n'est pas une obligation. L'espace $\mathcal{P}_n$ est de dimension $n+1$; il est souvent agréable d'indexer comme ici les vecteurs d'une de ses bases par les entiers de 0 à n (plutôt que de 1 à $n+1$).

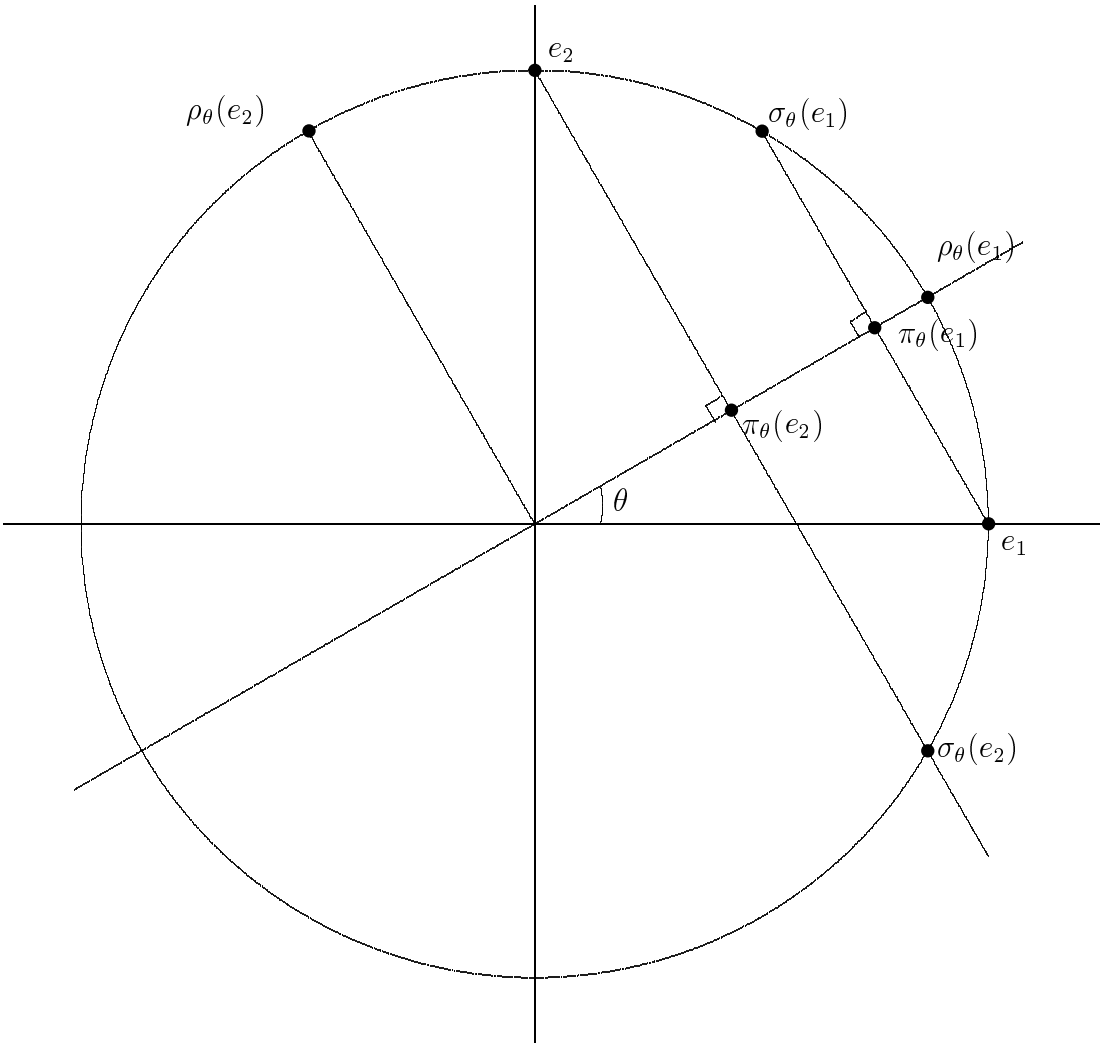

FIG. 1

est l'application de dérivation, alors

$$M_{\mathcal{B},\mathcal{B}}(\sigma) = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 2 & \cdots & 0 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & n-1 & 0 \\ 0 & 0 & 0 & \cdots & 0 & n \\ 0 & 0 & 0 & \cdots & 0 & 0 \end{pmatrix}$$

(iii) Soient W l'espace $\mathbb{R}^3$, muni de sa base canonique $\mathcal{C}an = (e_1, e_2, e_3)$. Soit V le sous-espace de W des vecteurs $\begin{pmatrix} x \\ y \\ z \end{pmatrix}$ tels que $x + y + z = 0$; on vérifie que $\mathcal{B} = (e_1 - e_2, e_2 - e_3)$ est une base de V. Considérons dans $\mathbb{R}^3$ la projection orthogonale p

d'image V, définie par

$$p\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x - \frac{1}{3}(x+y+z) \\ y - \frac{1}{3}(x+y+z) \\ z - \frac{1}{3}(x+y+z) \end{pmatrix}.$$

On calcule

$$p(e_1) = \begin{pmatrix} \frac{2}{3} \\ -\frac{1}{3} \\ -\frac{1}{3} \end{pmatrix} = \frac{2}{3}\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} + \frac{1}{3}\begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}$$

$$p(e_2) = \begin{pmatrix} -\frac{1}{3} \\ \frac{2}{3} \\ -\frac{1}{3} \end{pmatrix} = -\frac{1}{3}\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} + \frac{1}{3}\begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}$$

$$p(e_3) = \begin{pmatrix} -\frac{1}{3} \\ -\frac{1}{3} \\ \frac{2}{3} \end{pmatrix} = -\frac{1}{3}\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} - \frac{2}{3}\begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}$$

de sorte que

$$M_{\mathcal{B},\mathcal{C}an}(p) = \begin{pmatrix} \frac{2}{3} & -\frac{1}{3} & -\frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & -\frac{2}{3} \end{pmatrix}.$$

A titre de vérification, notons que

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

Dans cette égalité, la colonne $\begin{matrix} 1 \\ -1 \\ 0 \end{matrix}$ à gauche du signe «égal» a pour coefficients les coordonnées du vecteur $e_1 - e_2$ dans la base $\mathcal{C}an$, alors que la colonne $\begin{matrix} 1 \\ 0 \end{matrix}$ à droite du signe «égal» a pour coefficients les coordonnées du même vecteur dans la base $\mathcal{B}$. On vérifie de même

$$\begin{pmatrix} \frac{2}{3} & -\frac{1}{3} & -\frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & -\frac{2}{3} \end{pmatrix}\begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \qquad \text{et} \qquad \begin{pmatrix} \frac{2}{3} & -\frac{1}{3} & -\frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & -\frac{2}{3} \end{pmatrix}\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

(iv) Soit V un espace de dimension m muni d'une base $\mathcal{B} = (b_1, \ldots, b_m)$. Si α est un endomorphisme de V, alors $M_{\mathcal{B},\mathcal{B}}(\alpha) = I_m$ si et seulement si $\alpha = \mathrm{id}_V$.

En revanche, soient σ un automorphisme de V et $\mathcal{C}$ la base $(\sigma(b_1), \ldots, \sigma(b_m))$; alors $M_{\mathcal{C},\mathcal{B}}(\sigma) = I_m$, y compris dans le cas où σ n'est pas égal à l'indentité !

Application linéaire définie par une matrice, relativement à des bases données

Soient V, W deux espaces vectoriels de dimension finie, $\mathcal{B} = (b_1, \ldots, b_m)$ une base de V et $\mathcal{C} = (c_1, \ldots, c_n)$ une base de W, comme plus haut.

Pour toute matrice $s \in M_{m,n}(\mathbb{R})$, on définit l'application linéaire $\ell_{\mathcal{B},\mathcal{C}}(s) \in \mathcal{L}(W, V)$ de l'espace W (donné *avec* sa base $\mathcal{C}$) dans l'espace V (donné *avec* sa base $\mathcal{B}$) comme l'unique application linéaire de W dans V telle que

$$\big(\ell_{\mathcal{B},\mathcal{C}}(s)\big)(c_j) \;=\; \sum_{i=1}^{m} s_{i,j} b_i \qquad \text{pour tout } j, \quad 1 \leq j \leq n.$$

Il en résulte que, si $w = \sum_{j=1}^{m} w_j c_j \in W$ a pour image par $\ell_{\mathcal{B},\mathcal{C}}(s)$ un vecteur $v = \sum_{i=1}^{n} v_i b_i$, alors

$$v_i \;=\; \sum_{j=1}^{n} s_{i,j} w_j \qquad \text{pour tout } i, \quad 1 \leq i \leq m.$$

Exemples. (i) Si $W = \mathbb{R}^3$ et $V = \mathbb{R}^2$ sont munis des bases canoniques respectivement notées $\mathcal{C}an_3$ et $\mathcal{C}an_2$, la matrice $\begin{pmatrix} p & q & r \\ u & v & w \end{pmatrix}$ dans $M_{2,3}(\mathbb{R}^3)$ définit l'application linéaire

$$\ell_{\mathcal{C}an_2,\mathcal{C}an_3} \;:\; \mathbb{R}^3 \longrightarrow \mathbb{R}^2, \qquad \begin{pmatrix} x \\ y \\ z \end{pmatrix} \longmapsto \begin{pmatrix} px + qy + rz \\ ux + vy + wz \end{pmatrix}.$$

(ii) Si $W = \mathbb{R}^n$ est muni de la base canonique $\mathcal{C}an$, toute matrice $\begin{pmatrix} \lambda_1 & \cdots & \lambda_n \end{pmatrix}$ dans $M_{1,n}(\mathbb{R}^n)$ fournit une forme linéaire

$$\mathbb{R}^n \longrightarrow \mathbb{R}, \qquad \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \longmapsto \sum_{i=1}^{n} \lambda_i x_i$$

($\mathbb{R}$ est muni de la base à un élément, le nombre 1 !).

Théorème III.1 (résultat principal du paragraphe). *On conserve les notations précédentes. Les applications*

$$\begin{aligned} \mathcal{L}(W, V) &\longmapsto M_{m,n}(\mathbb{R}), \qquad \sigma \longmapsto M_{\mathcal{B},\mathcal{C}}(\sigma) \\ M_{m,n}(\mathbb{R}) &\longmapsto \mathcal{L}(W, V), \qquad s \longmapsto \ell_{\mathcal{B},\mathcal{C}}(s) \end{aligned}$$

sont d'une part des isomorphismes linéaires inverses l'un de l'autre, et sont d'autre part compatibles avec la composition des applications et la multiplication des matrices. Cette compatibilité se traduit par les formules suivantes, où U [respectivement V, W] est un espace vectoriel de dimension l [resp. m, n] donné avec une base $\mathcal{A}$ [resp. $\mathcal{B}$, $\mathcal{C}$] :

$$(*) \qquad M_{\mathcal{A},\mathcal{B}}(\tau) M_{\mathcal{B},\mathcal{C}}(\sigma) \;=\; M_{\mathcal{A},\mathcal{C}}(\tau\sigma) \qquad \textit{pour } \sigma \in \mathcal{L}(W, V) \textit{ et } \tau \in \mathcal{L}(V, U)\,;$$

$$(**) \qquad \ell_{\mathcal{A},\mathcal{B}}(t) \ell_{\mathcal{B},\mathcal{C}}(s) \;=\; \ell_{\mathcal{A},\mathcal{C}}(ts) \qquad \textit{pour } s \in M_{m,n}(\mathbb{R}) \textit{ et } t \in M_{l,m}(\mathbb{R}).$$

REMARQUE. Ce théorème peut être vu comme une ébauche de *dictionnaire* entre les langages « applications linéaires entre espaces vectoriels munis de bases » et « matrices » . Un dictionnaire est conçu pour être *utilisé* avant d'être *lu de part en part*. Nous utiliserons cette analogie pour justifier le fait de ne rédiger ici qu'une partie de la preuve ; ceci dit, une preuve complète ne nécessite aucune idée originale et consiste en vérifications de routine.

PREUVE D'UNE ASSERTION DU THÉORÈME. Posons

$$\begin{aligned} s &= M_{\mathcal{B},\mathcal{C}}(\sigma) \quad \in M_{m,n}(\mathbb{R}) \\ t &= M_{\mathcal{A},\mathcal{B}}(\tau) \quad \in M_{l,m}(\mathbb{R}) \\ x &= M_{\mathcal{A},\mathcal{C}}(\tau\sigma) \in M_{l,n}(\mathbb{R}); \end{aligned}$$

il s'agit de vérifier que $x = ts$.

Pour tout $k \in \{1, \ldots, n\}$, on a d'une part

$$(\tau\sigma)(c_k) = \tau\left(\sigma(c_k)\right) = \tau\left(\sum_{j=1}^{m} s_{j,k} b_j\right) = \sum_{j=1}^{m} s_{j,k}\tau(b_j) = \sum_{j=1}^{m} s_{j,k} \sum_{i=1}^{l} t_{i,j} a_i = \sum_{i=1}^{l} \left(\sum_{j=1}^{m} t_{i,j} s_{j,k}\right) a_i$$

et d'autre part

$$(\tau\sigma)(c_k) = \sum_{i=1}^{l} x_{i,k} a_i.$$

Par suite, on a

$$x_{i,k} = \sum_{j=1}^{m} t_{i,j} s_{j,k}$$

pour tous $i \in \{1, \ldots, l\}$ et $j \in \{1, \ldots, m\}$, c'est-à-dire $x = ts$. □

Il résulte par exemple du théorème que, pour un espace vectoriel V muni d'une base $\mathcal{B}$ et un automorphisme σ de V, les matrices relativement à $\mathcal{B}$ de σ et de σ^{-1} sont inverses l'une de l'autre :

$$M_{\mathcal{B},\mathcal{B}}(\sigma^{-1}) = M_{\mathcal{B},\mathcal{B}}(\sigma)^{-1}.$$

5. Applications linéaires et changements de bases

Au début du § III.4, nous avons montré comment associer une matrice à une application linéaire et des bases. A différents choix de bases correspondent différentes matrices, et il s'agit maintenant de comprendre la relations entre ces matrices.

Pour cela, considérons un espace vectoriel V de dimension finie donné avec deux bases $\mathcal{B} = (b_1, \ldots, b_m)$ et $\mathcal{B}' = (b'_1, \ldots, b'_m)$, ainsi qu'un espace vectoriel W de dimension finie donné avec deux bases $\mathcal{C} = (c_1, \ldots, c_n)$ et $\mathcal{C}' = (c'_1, \ldots, c'_n)$. C'est une conséquence importante de la formule $(*)$ du théorème précédent que, pour $\sigma \in \mathcal{L}(W, V)$, nous avons

$$M_{\mathcal{B}',\mathcal{C}'}(\sigma) = M_{\mathcal{B}',\mathcal{B}}(id_V)\, M_{\mathcal{B},\mathcal{C}}(\sigma)\, M_{\mathcal{C},\mathcal{C}'}(id_W).$$

La matrice $M_{\mathcal{B}',\mathcal{B}}(id_V) \in M_m(\mathbb{R})$ est la *matrice de passage* de $\mathcal{B}$ à $\mathcal{B}'$, et $M_{\mathcal{C},\mathcal{C}'}(id_W) \in M_n(\mathbb{R})$ est la matrice de passage de $\mathcal{C}'$ à $\mathcal{C}$.

Cas particulier lorsque $W = V$, $\mathcal{C} = \mathcal{B}$, $\mathcal{C}' = \mathcal{B}'$ et $\sigma \in \mathcal{L}(V)$: si on pose $t = M_{\mathcal{B}',\mathcal{B}}(id_V)$, alors

$$M_{\mathcal{B}',\mathcal{B}'}(\sigma) = tM_{\mathcal{B},\mathcal{B}}(\sigma)t^{-1}.$$

Deux matrices $s, s' \in M_m(\mathbb{R})$ sont dites *semblables* s'il existe une matrice inversible $t \in M_m(\mathbb{R})$ telle que $s' = tst^{-1}$. Dans la formule ci-dessus, les deux matrices $M_{\mathcal{B}',\mathcal{B}'}(\sigma)$ et $M_{\mathcal{B},\mathcal{B}}(\sigma)t^{-1}$ représentant σ sont donc semblables

Exemple. Soit $\pi_\theta \in \mathcal{L}(\mathbb{R}^2)$ la projection orthogonale de $\mathbb{R}^2$ sur une droite faisant un angle θ avec le premier axe. Soient $\mathcal{C}an = (e_1, e_2)$ la base canonique de $\mathbb{R}^2$ et $\mathcal{B} = (\rho_\theta(e_1), \rho_\theta(e_2))$ la base obtenue par une rotation d'angle θ. Comme au § III.4, les matrices de π_θ relativement à $\mathcal{C}an$ et $\mathcal{B}$ sont

$$M_{\mathcal{C}an,\mathcal{C}an}(\pi_\theta) = \begin{pmatrix} \cos^2\theta & \cos\theta\sin\theta \\ \cos\theta\sin\theta & \sin^2\theta \end{pmatrix} \qquad \text{et} \qquad M_{\mathcal{B},\mathcal{B}}(\pi_\theta) = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}.$$

La matrice de changement de base, notée t, et son inverse sont données par

$$t = M_{\mathcal{C}an,\mathcal{B}}(id) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \quad \text{et} \quad t^{-1} = M_{\mathcal{B},\mathcal{C}an}(id) = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}.$$

On vérifie que

$$t \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} t^{-1} = \begin{pmatrix} \cos^2\theta & \cos\theta\sin\theta \\ \cos\theta\sin\theta & \sin^2\theta \end{pmatrix}.$$

6. Noyaux et images des applications linéaires

Soient V, W deux espaces vectoriels et $\sigma \in \mathcal{L}(W, V)$ une application linéaire.

Notons que *l'image* $\mathrm{Im}(\sigma)$ de σ (définie comme au chapitre I) est un sous-espace vectoriel de V. Il est évident que l'application σ est surjective si et seulement si $\mathrm{Im}(\sigma) = V$.

Notons aussi que l'image inverse $\sigma^{-1}(0)$ de l'origine est un sous-espace vectoriel de W. Par définition, c'est le *noyau* de σ, noté $\mathrm{Ker}(\sigma)$.

Proposition. *Avec les notations ci-dessus, l'application σ est injective si et seulement si* $\mathrm{Ker}(\sigma) = \{0\}$.

Preuve. Supposons d'abord que le noyau de σ est réduit à zéro. Soient $w_1, w_2 \in W$ tels que $\sigma(w_1) = \sigma(w_2)$. Alors $\sigma(w_1 - w_2) = 0$, donc $w_1 - w_2 = 0$ par hypothèse, donc $w_1 = w_2$. Ainsi, σ est injective.

Réciproquement, si l'application σ est injective, alors il n'y a qu'un vecteur dans W d'image l'origine de V, donc $\mathrm{Ker}(\sigma) = \{0\}$. □

Remarques. (i) Soient V, W des espaces vectoriels, $\alpha \in \mathcal{L}(W, V)$ et $v \in V$. On a l'alternative suivante :

(ii) ou bien $v \notin \mathrm{Im}(\alpha)$, c'est-à-dire $\alpha^{-1}(v) = \emptyset$;

(iii) ou bien il existe $w \in W$ tel que $\alpha(w) = v$, et alors $\alpha^{-1}(v) = w + \mathrm{Ker}(\alpha)$,

où $w + \mathrm{Ker}(\alpha) = \{x \in W \mid \text{il existe } y \in \mathrm{Ker}(\alpha) \text{ tel que } x = w + y\}$, par définition. (Attention : $w + \mathrm{Ker}(\alpha)$ *n'est pas* en général un sous-espace vectoriel de W !)

(iv) Soient U, V, W, X des espaces vectoriels et $\rho \in \mathcal{L}(X, W)$, $\sigma \in \mathcal{L}(W, V)$ et $\tau \in \mathcal{L}(V, U)$ des applications linéaires.

On a $\sigma\,(\mathrm{Ker}(\tau\sigma)) \subset \mathrm{Ker}(\tau)$; si de plus σ est un isomorphisme de W sur V, alors l'inclusion est une égalité, de sorte qu'en particulier les espaces vectoriels $\mathrm{Ker}(\tau\sigma)$ et $\mathrm{Ker}(\tau)$ sont isomorphes.

On a $\mathrm{Im}(\sigma\rho) = \sigma\,(\mathrm{Im}(\rho))$; si de plus σ est une injection de W dans V, alors les espaces $\mathrm{Im}(\sigma\rho)$ et $\mathrm{Im}(\rho)$ sont isomorphes.

Théorème III.2 (formule de la dimension). *Soient V, W deux espaces vectoriels, avec W de dimension finie, et $\sigma \in \mathcal{L}(W, V)$. Alors les espaces* $\mathrm{Im}(\sigma)$ *et* $\mathrm{Ker}(\sigma)$ *sont de dimension finie, et*

$$\dim(W) \;=\; \dim\big(\mathrm{Im}(\sigma)\big) + \dim\big(\mathrm{Ker}(\sigma)\big).$$

Preuve. Notons d'abord que les sous-espaces $\mathrm{Ker}(\sigma)$ de W et $\mathrm{Im}(\sigma)$ de V sont de dimensions finies. Soit $(c_1, \ldots, c_n)$ une base de W obtenue par complétion d'une base $(c_1, \ldots, c_q)$ de $\mathrm{Ker}(\sigma)$.

Affirmation : la suite $\mathcal{B} = (\sigma(c_{q+1}), \ldots, \sigma(c_n))$ est une base de $\mathrm{Im}(\sigma)$.

En effet, il est évident que $\mathcal{B}$ est une partie génératrice de $\mathrm{Im}(\sigma)$. Il reste donc à vérifier que c'est une partie libre. Soient $\lambda_{q+1}, \ldots, \lambda_n$ une suite de nombres réels tels que $\sum_{j=q+1}^{n} \lambda_j \sigma(c_j) = 0$. Alors $\sum_{j=q+1}^{n} \lambda_j c_j$ est dans le noyau de σ, de sorte qu'il existe des nombres réels $\lambda_1, \ldots, \lambda_q$ tels que $\sum_{j=q+1}^{n} \lambda_j c_j = \sum_{i=1}^{q} \lambda_i c_i$. Il résulte alors du fait que $(c_1, \ldots, c_n)$ est une base de W que les λ_j sont tous nuls. □

Autre preuve (esquisse). Le sous-espace $\mathrm{Ker}(\sigma)$ de W donne lieu à un espace quotient $W/\mathrm{Ker}(\sigma)$ (voir le § II.6). Pour tout élément $x \in W/\mathrm{Ker}(\sigma)$, on définit un élément $\alpha(x) \in \mathrm{Im}(\sigma)$ en posant $\alpha(x) = \sigma(w)$, où w est un vecteur de W tel que $[w]_{\mathrm{Ker}(\sigma)} = x$. On vérifie que l'application

$$\alpha \;=\; W/\mathrm{Ker}(\sigma) \longrightarrow \mathrm{Im}(\sigma)$$

est ainsi bien définie, linéaire, et que c'est un isomorphisme. La formule de la dimension résulte donc de la dernière proposition du § II.6. □

Corollaire. *Soient V, W deux espaces vectoriels de dimension finie et σ une application linéaire de W dans V.*

(i) *Si* $\dim(W) < \dim(V)$, *alors σ n'est pas surjective ;*

(ii) *si* $\dim(W) > \dim(V)$, *alors σ n'est pas injective ;*

(iii) *si* $\dim(W) = \dim(V)$, *alors les trois propriétés suivantes sont équivalentes : σ est surjective, σ est injective, σ est bijective.*

Exemples pour montrer que l'hypothèse «de dimension finie» ne peut être omise dans le corollaire précédent. Soit $\mathcal{P}$ l'espace vectoriel de toutes les applications polynomiales de $\mathbb{R}$ dans $\mathbb{R}$, qui est de dimension infinie. Soit $\sigma \in \mathcal{L}(\mathcal{P})$ la dérivation ; alors σ est une application surjective non injective. Soit $\tau \in \mathcal{L}(\mathcal{P})$ le passage à une primitive

nulle en zéro, c'est-à-dire l'application linéaire définie par $\big(\tau(f)\big)(x) = \int_0^x f(y)\,dy$; alors ρ est une application injective non surjective.

Cas des matrices

Pour toute matrice $s \in M_{m,n}(\mathbb{R})$, on écrit $\mathrm{Im}(s)$ pour l'image de l'application $\ell_{\mathcal{C}an,\mathcal{C}an}(s) : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ (le premier «$\mathcal{C}an$» désigne la base canonique dans l'espace *but* $\mathbb{R}^m$, et le second la base canonique dans l'espace *source* $\mathbb{R}^n$). De même, $\mathrm{Ker}(s)$ désigne le noyau de cette application.

Corollaire. *Soit $s \in M_n(\mathbb{R})$ une matrice carrée. Les propriétés suivantes sont équivalentes :*

(i) $\mathrm{Ker}(s) = \{0\}$;

(ii) $\mathrm{Im}(s) = \mathbb{R}^n$;

(iii) $\ell_{\mathcal{C}an,\mathcal{C}an}(s)$ *est un isomorphisme ;*

(iv) s *est inversible ;*

(v) *les colonnes de s constituent une base de $\mathbb{R}^n$;*

(vi) *les colonnes de s constituent une famille libre de $\mathbb{R}^n$;*

(vii) *les colonnes de s constituent une famille génératrice de $\mathbb{R}^n$.*

Preuve. L'équivalence de (i), (ii) et (iii) est un cas particulier du corollaire précédent. L'équivalence de (iii), (iv) et (v) résulte immédiatement du théorème du § III.4. Les propriétés (v), (vi) et (vii) sont équivalentes par un corollaire du § II.3. □

7. Applications linéaires et rang

Le *rang* d'une application linéaire $\sigma \in \mathcal{L}(W,V)$ est la dimension $\mathrm{rang}(\sigma)$ de $\mathrm{Im}(\sigma)$; on remarque que $\mathrm{rang}(\sigma) \leq \max\{\dim(V), \dim(W)\}$. Le rang d'une matrice $s \in M_{m,n}(\mathbb{R})$ est le rang de l'application linéaire $\ell_{\mathcal{C}an,\mathcal{C}an}(s) \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$.

Remarques. (i) Soient U, V, W, X des espaces vectoriels de dimension finie et $\rho \in \mathcal{L}(X,W)$, $\sigma \in \mathcal{L}(W,V)$, $\tau \in \mathcal{L}(V,U)$ des applications linéaires. Si ρ et τ sont des isomorphismes, alors les applications σ et $\tau\sigma\rho$ ont même rang.

En effet, $\mathrm{Im}(\tau\sigma\rho) = \mathrm{Im}(\tau\sigma)$ par surjectivité de l'application ρ, et l'isomorphisme τ fournit une application linéaire

$$\mathrm{Im}(\sigma) \longrightarrow \mathrm{Im}(\tau\sigma), \qquad v \longmapsto \tau(v)$$

qui est un isomorphisme par injectivité de l'application τ. Par suite $\dim_{\mathbb{R}}(\mathrm{Im}(\sigma)) = \dim_{\mathbb{R}}(\mathrm{Im}(\tau\sigma)) = \dim_{\mathbb{R}}(\mathrm{Im}(\tau\sigma\rho))$.

(ii) Soient V un espace vectoriel de dimension finie, $v_1, \ldots, v_n$ des vecteurs de V et U le sous-espace vectoriel de V qu'ils engendrent. Si r désigne la dimension de U, il existe des entiers $i_1, i_2, \ldots, i_r$, avec $1 \leq i_1 < i_2 < \ldots < i_r \leq n$, tels que $(v_{i_1}, v_{i_2}, \ldots, v_{i_r})$ soit une base de U. (C'est un cas particulier du lemme d'échange.)

En particulier, si $s \in M_{m,n}(\mathbb{R})$ est une matrice de rang r, il existe des entiers $i_1, i_2, \ldots, i_r$ comme ci-dessus tels que les colonnes de s correspondant à ces entiers constituent une base du sous-espace $\mathrm{Im}(s)$ de $\mathbb{R}^m$. Cette suite d'entiers n'est pas univoquement

définie par s, mais sa longueur l'est : le rang d'une matrice est le nombre maximum de ses vecteurs colonnes qui soient linéairement indépendants. (En anticipant sur la fin du § III.7, on peut également affirmer que le rang d'une matrice est le nombre maximum de ses vecteurs lignes qui soient linéairement indépendants.)

PROPOSITION (= version linéaire du «théorème du rang»). *Soient V, W deux espaces vectoriels de dimension finie et $\sigma \in \mathcal{L}(W, V)$ une application linéaire de rang r. Alors il existe une base $\mathcal{B} = (b_1, \ldots, b_m)$ de V et une base $\mathcal{C} = (c_1, \ldots, c_n)$ de W telles que*

$$\sigma(c_j) \;=\; \begin{cases} b_j & \textit{pour} \quad j \in \{1, \ldots, r\}, \\ 0 & \textit{sinon,} \end{cases}$$

c'est-à-dire telles que

$$M_{\mathcal{B},\mathcal{C}}(\sigma) \;=\; \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix},$$

où I_r désigne une matrice unité de taille r et où les trois 0 désignent des matrices nulles de tailles appropriées (voir plus bas).

PREUVE. La formule de la dimension montre que

$$\dim_{\mathbb{R}}(\mathrm{Ker}(\sigma)) \;=\; \dim_{\mathbb{R}}(W) - \dim_{\mathbb{R}}(\mathrm{Im}(\sigma)) \;=\; n - r.$$

Il existe donc une base $(c_1, \ldots, c_n)$ de W telle que $(c_{r+1}, \ldots, c_n)$ est une base de $\mathrm{Ker}(\sigma)$. Posons $b_i \;=\; \sigma(c_i)$ pour $i \in \{1, \ldots, r\}$; il résulte de la preuve de la formule de la dimension que $(b_1, \ldots, b_r)$ est une base de $\mathrm{Im}(\sigma)$. On peut donc la compléter en une base $(b_1, \ldots, b_m)$ de V. On vérifie alors que la matrice $M_{\mathcal{B},\mathcal{C}}(\sigma)$ a la forme indiquée. □

REMARQUE. Cette proposition joue un rôle clé dans quelques-uns des théorèmes les plus importants de la théorie des fonctions de plusieurs variables (Analyse II), à savoir le théorème du rang et ses cas particuliers (théorème des fonctions implicites, théorème de la submersion, théorème de l'immersion).

EXEMPLE. Déterminons une base du noyau et une base de l'image de la matrice

$$s \;=\; \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 3 & 2 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix}.$$

Si on pose $\sigma = l_{\mathcal{C}an_3,\mathcal{C}an_4}(s)$, on a donc

$$\sigma \;:\; \mathbb{R}^4 \longrightarrow \mathbb{R}^3, \qquad \sigma \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} \;=\; \begin{pmatrix} x + 2y + 3z + 4t \\ 4x + 3y + 2z + t \\ x + y + z + t \end{pmatrix}$$

et $M_{\mathcal{C}an_3,\mathcal{C}an_4}(\sigma) = s$.

Les coordonnées d'un vecteur du noyau sont les solutions du système linéaire homogène

$$\begin{array}{rcrcrcl} x & + & 2y & + & 3z & + & 4t \;=\; 0 \\ 4x & + & 3y & + & 2z & + & t \;=\; 0. \\ x & + & y & + & z & + & t \;=\; 0 \end{array}$$

La méthode de Gauss fournit successivement

$$\begin{array}{rcrcrcr} x & + & 2y & + & 3z & + & 4t & = & 0 \\ & & -5y & + & -10z & + & -15t & = & 0 \\ & & -y & + & -2z & + & -3t & = & 0 \end{array}$$

et

$$\begin{array}{rcrcrcr} x & + & 2y & + & 3z & + & 4t & = & 0 \\ & & y & + & 2z & + & 3t & = & 0. \\ & & & & & & 0 & = & 0 \end{array}$$

Les solutions du système s'écrivent donc

$$\begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = \begin{pmatrix} \lambda + 2\mu \\ -2\lambda - 3\mu \\ \lambda \\ \mu \end{pmatrix} = \lambda \begin{pmatrix} 1 \\ -2 \\ 1 \\ 0 \end{pmatrix} + \mu \begin{pmatrix} 2 \\ -3 \\ 0 \\ 1 \end{pmatrix}$$

où λ et μ sont arbitraires. Par suite les vecteurs

$$k_3 = \begin{pmatrix} 1 \\ -2 \\ 1 \\ 0 \end{pmatrix} \qquad \text{et} \qquad k_4 = \begin{pmatrix} 2 \\ -3 \\ 0 \\ 1 \end{pmatrix}$$

constituent une base du noyau $\mathrm{Ker}(s)$.

La formule de la dimension montre que la matrice s est de rang 2. Si e_j $(1 \leq j \leq 4)$ désignent les vecteurs de la base canonique de $\mathbb{R}^4$, on vérifie que (e_1, e_2, k_3, k_4) est une base de $\mathbb{R}^4$. Les images par σ des vecteurs de cette base qui ne sont pas dans $\mathrm{Ker}(\sigma)$, c'est-à-dire les vecteurs

$$\sigma(e_1) = \begin{pmatrix} 1 \\ 4 \\ 1 \end{pmatrix} \qquad \text{et} \qquad \sigma(e_2) = \begin{pmatrix} 2 \\ 3 \\ 1 \end{pmatrix},$$

constituent une base de l'image $\mathrm{Im}(s)$.

Ecriture de matrices par blocs

Nous n'indiquons ici que des exemples, pour ne pas avoir recours à des notations plus lourdes.

Etant donné une matrice $s \in M_{m,n}(\mathbb{R})$ et des décompositions $m = m_1 + m_2$ et $n = n_1 + n_2$ en sommes d'entiers positifs, on a une écriture par blocs

$$s = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

avec $a \in M_{m_1,n_1}(\mathbb{R})$, $b \in M_{m_1,n_2}(\mathbb{R})$, etc.

On peut montrer (c'est élémentaire et fastidieux) que l'écriture par blocs est compatible avec le produit des matrices. Par exemple, pour des entiers donnés comme sommes d'entiers positifs $l = l_1 + l_2$, $m = m_1 + m_2 + m_3$ et $n = n_1 + n_2$, on a

$$\begin{pmatrix} a & b & c \\ d & e & f \end{pmatrix} \begin{pmatrix} a' & b' \\ c' & d' \\ e' & f' \end{pmatrix} = \begin{pmatrix} aa' + bc' + ce' & ab' + bd' + cf' \\ da' + ec' + fe' & db' + ed' + ff' \end{pmatrix},$$

où $a \in M_{l_1,m_1}(\mathbb{R})$, $b \in M_{l_1,m_2}(\mathbb{R})$, …, $f' \in M_{m_3,n_2}(\mathbb{R})$, …, $aa', bc', ce' \in M_{l_1,n_1}(\mathbb{R})$, ……, $db' + ed' + ff' \in M_{l_2,n_2}(\mathbb{R})$.

La proposition précédente implique l'énoncé suivant pour les matrices.

COROLLAIRE. *Soit $s \in M_{m,n}(\mathbb{R})$ une matrice de rang r. Il existe des matrices inversibles $g \in M_m(\mathbb{R})$, $h \in M_n(\mathbb{R})$ telles que*

$$gsh = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}.$$

Matrices équivalentes

Deux matrices $s', s'' \in M_{m,n}(\mathbb{R})$ sont *équivalentes* s'il existe des matrices inversibles $g \in M_m(\mathbb{R})$, $h \in M_n(\mathbb{R})$ telles que $s'' = gs'h$. [Ne pas confondre avec la définition de «semblable» , au début du § III.5.]

PROPOSITION. *Deux matrices de $M_{m,n}(\mathbb{R})$ sont équivalentes si et seulement si elles ont même rang.*

PREUVE. Soient $s', s'' \in M_{m,n}(\mathbb{R})$.

Si s' et s'' ont même rang, il existe des matrices inversibles $g', g'' \in M_m(\mathbb{R})$ et $h', h'' \in M_n(\mathbb{R})$ telles que

$$g's'h' = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} = g''s''h'',$$

de sorte que la matrice

$$s'' = \left((g'')^{-1}g'\right) s' \left(h'(h'')^{-1}\right)$$

est équivalente à s'.

Si les matrices s' et s'' sont équivalentes, c'est-à-dire s'il existe des matrices inversibles $g \in M_m(\mathbb{R})$ et $h \in M_n(\mathbb{R})$ telles que $s'' = gs'h$, alors

$$\dim_{\mathbb{R}}(\mathrm{Im}(s')) = \dim_{\mathbb{R}}(\mathrm{Im}(gs')) = \dim_{\mathbb{R}}\left(\mathrm{Im}(s''h^{-1})\right) = \dim_{\mathbb{R}}(\mathrm{Im}(s'')),$$

c'est-à-dire $\mathrm{rang}(s') = \mathrm{rang}(s'')$; comparer avec la première remarque du § III.7. □

PROPOSITION. *Une matrice et sa matrice transposée ont le même rang.*

REFORMULATION. Pour $s \in M_{m,n}(\mathbb{R})$, la dimension du sous-espace de $\mathbb{R}^m$ engendré par les *colonnes* de s est égale à la dimension du sous-espace de $\mathbb{R}^n$ engendré par les transposées des *lignes* de s. [Remarque : c'est *évident* lorsque $m = 1$ ou $n = 1$!]

PREUVE. Si $s = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$, alors ${}^t s = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$ et l'assertion de la proposition est évidente. [Noter toutefois que les «0» dans l'écriture par blocs de s représentent des matrices dont les tailles sont en général différentes de celles des matrices représentées par des «0» dans l'écriture par blocs de ${}^t s$!]

Dans le cas général, on peut écrire

$$s = g \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} h \quad \text{et} \quad {}^t s = {}^t h \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} {}^t g,$$

de sorte que

$$\operatorname{rang}\left({}^t s\right) \;=\; \operatorname{rang}\left({}^t h \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} {}^t g\right) \;=\; \operatorname{rang}\left(\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}\right) \;=\; r \;=\; \operatorname{rang}(s)\,. \qquad \square$$

8. Systèmes linéaires

Considérons une matrice $a \in M_{m,n}(\mathbb{R})$, un vecteur $b \in \mathbb{R}^m$, le système linéaire

$$(*) \qquad ax \;=\; b$$

et le système linéaire homogène

$$(**) \qquad ax \;=\; 0$$

correspondant. On définit la *matrice augmentée*[5] du système $(*)$.

Observation 1. *Si $x, y \in \mathbb{R}^n$ sont deux solutions de $(*)$, alors $x-y$ est une solution de $(**)$.*

Observation 2. *Si x est une solution de $(*)$ et z une solution de $(**)$, alors $x+z$ est une solution de $(*)$.*

Théorème III.3. *Soit r le rang de a. Les solutions du système homogène $(**)$ forment un sous-espace vectoriel de $\mathbb{R}^n$ de dimension $n-r$. En particulier, 0 est l'unique solution de $(**)$ si et seulement si $r = n$.*

Preuve. Voir la formule de la dimension. $\square$

Théorème III.4. *Le système à m équations et n inconnues $(*)$ possède une solution si et seulement si la matrice a et la matrice augmentée $\begin{pmatrix} a & b \end{pmatrix}$ ont même rang.*

Lorsque ces conditions sont satisfaites, la solution générale de $(*)$ contient $n-r$ paramètres arbitraires ; elle est obtenue en ajoutant une solution particulière de $(*)$ à la solution générale du système homogène $(**)$.

Preuve. Pour que le système $(*)$ ait une solution, il faut et il suffit que $b \in \operatorname{Im}(a)$, c'est-à-dire que $\langle \operatorname{Im}(a), b\rangle = \operatorname{Im}(a)$ [pour la notation $\langle \ldots \rangle$, voir le § II.2]. Le reste de l'énoncé résulte immédiatement de ce qui précède. $\square$

9. Dual d'un espace vectoriel

Le *dual* d'un espace vectoriel V est l'espace vectoriel $V' = \mathcal{L}(V, \mathbb{R})$.

Si V est de dimension finie, alors V et V' ont même dimension, et sont donc isomorphes (voir les corollaires du § III.3). Mais attention !!!!! il n'existe pas d'isomorphisme *canonique* entre V et V'. (De plus, si V est de dimension infinie, V et V' ne sont pas isomorphes.)

[5] Attention : $a \in M_{m,n}(\mathbb{R})$ et $b \in M_{m,1}(\mathbb{R})$ sont deux blocs de la matrice augmentée, qui est une matrice à m lignes et $n+1$ colonnes ; il ne s'agit aucunement d'un produit de a par b, produit qui n'est d'ailleurs pas défini lorsque $m \neq n$.

La *base duale* d'une base $\mathcal{B} = (b_1, \ldots, b_n)$ d'un espace vectoriel V de dimension finie est la base $\mathcal{B}' = (b'_1, \ldots, b'_n)$ de l'espace dual V' définie par

$$b'_j(b_i) \;=\; \begin{cases} 1 & \text{si} \quad i = j, \\ 0 & \text{sinon.} \end{cases}$$

Pour montrer que $\mathcal{B}'$ est une base de V', il suffit de montrer que ses vecteurs sont linéairement indépendants (car le nombre d'éléments de $\mathcal{B}'$ est égal à la dimension de V'), ce dont nous laissons la vérification au lecteur.

Exemples. (i) Soit (e_1, e_2, e_3) la base canonique de $\mathbb{R}^3$ et (e'_1, e'_2, e'_3) la base duale de l'espace vectoriel dual. Alors

$$e'_1 \begin{pmatrix} x \\ y \\ z \end{pmatrix} \;=\; x, \qquad e'_2 \begin{pmatrix} x \\ y \\ z \end{pmatrix} \;=\; y, \qquad \text{et} \qquad e'_3 \begin{pmatrix} x \\ y \\ z \end{pmatrix} \;=\; z.$$

(ii) Soit $\mathcal{P}_n$ l'espace des applications polynomiales de $\mathbb{R}$ dans $\mathbb{R}$ de degré au plus n et $\mathcal{B} = (f_0, \ldots, f_n)$ la base définie par $f_j(t) = \frac{1}{j!}t^j$. Notons $D \in \mathcal{L}(\mathcal{P}_n)$ l'application de dérivation, D^2 la dérivée seconde, etc . Notons par ailleurs $\delta_0 \in \mathcal{L}(\mathcal{P}_n, \mathbb{R})$ la forme linéaire «évaluation en zéro» définie par $\delta_0(f) = f(0)$. Alors $\delta_0 D^j \in (\mathcal{P}_n)'$ pour tout $j \in \{0, \ldots, n\}$.

On vérifie sans peine que $\mathcal{B}' = (\delta_0, \delta_0 D, \delta_0 D^2, \ldots, \delta_0 D^n)$ est la base duale de $\mathcal{B}$.

Soit alors f un élément de $\mathcal{P}_n$, et $f = \sum_{j=0}^n \lambda_j f_j$ son écriture dans la base $\mathcal{B}$. Alors $\delta_0 D^k(f) = \lambda_k$, ou encore $\lambda_k = \frac{d^k f}{dt^k}(0)$ pour tout $k \in \{0, \ldots, n\}$, de sorte que

$$f(t) \;=\; \sum_{k=0}^{n} \frac{d^k f}{dt^k}(0) f_k(t) \;=\; \sum_{k=0}^{n} \frac{1}{k!} \frac{d^k f}{dt^k}(0) t^k,$$

ce qui est un cas particulier de la formule de Taylor.

Définition. Soient V, W des espaces vectoriels et $\sigma : W \longrightarrow V$ une application linéaire. La *transposée* de σ est l'application linéaire σ' de V' dans W' définie par $\sigma'(\varphi) = \varphi\sigma$.

Proposition. *Soient U, V, W des espaces vectoriels et $\sigma \in \mathcal{L}(W, V)$, $\tau \in \mathcal{L}(V, U)$. Alors $(\tau\sigma)' = \sigma'\tau'$.*

Preuve. Il est immédiat que

$$(\tau\sigma)'(\varphi) \;=\; \varphi\tau\sigma \;=\; \sigma'(\varphi\tau) \;=\; \sigma'\left(\tau'(\varphi)\right)$$

pour tout $\varphi \in V'$. □

Proposition. *Soit V un espace vectoriel de dimension finie. L'application de V dans le dual du dual de V définie par*

$$V \longrightarrow V'', \qquad v \longmapsto \begin{cases} V' \longrightarrow \mathbb{R} \\ \varphi \longmapsto \varphi(v) \end{cases}$$

est un isomorphisme linéaire.

LEMME. *Soit V un espace vectoriel de dimension finie. Pour tout vecteur non nul $v \in V$, il existe une forme linéaire $\varphi \in V'$ telle que $\varphi(v) = 1$.*

PREUVE DU LEMME. Il existe une base de V dont v soit le premier vecteur. Il suffit de prendre pour φ le premier vecteur de la base duale. □

PREUVE DE LA PROPOSITION. Le lemme montre que l'application de la proposition est injective. La conclusion résulte du fait que V et le dual de son dual sont des espaces vectoriels de même dimension. □

PROPOSITION. *Soient V, W, σ, σ' comme dans la définition précédente. Soient $\mathcal{B}$ [respectivement $\mathcal{C}$] une base de V [respectivement W] et $\mathcal{B}'$ [respectivement $\mathcal{C}'$] sa base duale. Alors*

$$M_{\mathcal{C}',\mathcal{B}'}(\sigma') \;=\; {}^t\left(M_{\mathcal{B},\mathcal{C}}(\sigma)\right).$$

PREUVE. Convenons des notations

$$\mathcal{B} \;=\; (b_1, \ldots, b_m) \qquad\qquad \mathcal{B}' \;=\; (b'_1, \ldots, b'_m)$$
$$\mathcal{C} \;=\; (c_1, \ldots, c_n) \qquad\qquad \mathcal{C}' \;=\; (c'_1, \ldots, c'_n)$$

ainsi que $s = M_{\mathcal{B},\mathcal{C}}(\sigma)$ et $x = M_{\mathcal{C}',\mathcal{B}'}(\sigma')$. Pour tous $i \in \{1, \ldots, m\}$ et $j \in \{1, \ldots, n\}$, nous allons écrire $\left(\sigma'(b'_i)\right)(c_j)$ de deux manières ; notons que $\sigma'(b'_i)$ est l'élément $b'_i\sigma$ de W'.

D'abord par définition de la matrice s, nous avons

$$\left(\sigma'(b'_i)\right)(c_j) \;=\; b'_i\left(\sigma(c_j)\right) \;=\; b'_i\left(\sum_{k=1}^{m} s_{k,j} b_k\right) \;=\; \sum_{k=1}^{m} s_{k,j} b'_i(b_k) \;=\; s_{i,j}.$$

Ensuite, par définition de la matrice x, nous avons

$$\left(\sigma'(b'_i)\right)(c_j) \;=\; \left(\sum_{l=1}^{n} x_{l,i} c'_l\right)(c_j) \;=\; \sum_{l=1}^{n} x_{l,i} c'_l(c_j) \;=\; x_{j,i}.$$

Ceci montre bien que $x = {}^t s$. □

10. Exercices

(III.1) On pose $a = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ et $b = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$. Dans $M_2(\mathbb{R})$, calculer les trois matrices

$$a^2 + 2ab + b^2 \quad , \qquad a^2 + 2ba + b^2 \qquad \text{et} \qquad a^2 + ab + ba + b^2 \;=\; (a+b)^2.$$

(III.2) Soit a la matrice $\begin{pmatrix} 0 & 1 & 2 & 3 \\ 0 & 0 & 4 & 5 \\ 0 & 0 & 0 & 6 \\ 0 & 0 & 0 & 0 \end{pmatrix}$. Calculer les puissances a^2, a^3, a^4 et a^5.

(III.3) $a \;=\; \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \in M_2(\mathbb{R})$. Montrer par récurrence sur k que, pour tout entier $k \geq 2$, tous les coefficients de la matrice a^k sont strictement positifs.

(III.4) Pour tout entier $p \geq 1$ et pour toute matrice carrée $c \in M_p(\mathbb{R})$, on définit la *trace* de c comme étant le nombre $\mathrm{trace}(c) = \sum_{l=1}^{p} c_{l,l}$.

Pour deux entiers $m, n \geq 1$ et deux matrices $a \in M_{m,n}(\mathbb{R})$, $b \in M_{n,m}(\mathbb{R})$, montrer que $\mathrm{tr}(ab) = \mathrm{tr}(ba)$.

(III.5) Soient $n \geq 1$ un entier et $a \in M_n(\mathbb{R})$ une matrice.

(i) Si $a^2 = 0$, montrer que $I_n + a$ est inversible, et que $(I_n + a)^{-1} = I_n - a$.

(ii) Si $a^3 = 0$, montrer que $I_n + a$ est inversible, et trouver une formule analogue à celle de (i) pour $(I_n + a)^{-1}$.

(ii)$^\sharp$ Plus généralement, s'il existe un entier $k \geq 1$ tel que $a^k = 0$, montrer que $I_n + a$ est inversible et trouver une formule pour $(I_n + a)^{-1}$.

(iii) Ecrire une matrice non nulle $b \in M_2(\mathbb{R})$ telle que $b^2 = 0$.

(iv) Ecrire une matrice $c \in M_3(\mathbb{R})$ telle que $c^2 \neq 0$ et $c^3 = 0$.

(III.6) Soit a une matrice à 2 lignes et 3 colonnes. Trouver une matrice $t \in M_2(\mathbb{R})$ telle que la matrice ta soit obtenue à partir de a en échangeant les 2 lignes.

(III.7) Soient V, W deux espaces vectoriels, $\alpha : V \longrightarrow W$ une application linéaire et $(v_i)_{i \in I}$ une famille de vecteurs de V.

Si l'application α est injective, montrer que la famille $(v_i)_{i \in I}$ est libre dans V si et seulement si la famille $(\alpha(v_i))_{i \in I}$ est libre dans W.

Si l'application α est surjective, est-il vrai que la famille $(v_i)_{i \in I}$ engendre V si et seulement si la famille $(\alpha(v_i))_{i \in I}$ engendre W ?

(III.8) Soient V un espace vectoriel et e un endomorphisme de V tel que $e^2 = e$ (on désigne par e^2 la composition de e avec e).

(i) Montrer que V est la somme directe de $\mathrm{Im}(e)$ et $\mathrm{Ker}(e)$.

(ii) Vérifier que $\mathrm{Im}(1 - e) = \mathrm{Ker}(e)$ et $\mathrm{Ker}(1 - e) = \mathrm{Im}(e)$; le symbole 1 désigne ici l'automorphisme identique de l'espace vectoriel V (noté ailleurs id_V).

NOTE. Un endomorphisme e de V est dit *idempotent* s'il est tel que $e^2 = e$.

(III.9) Soient V un espace vectoriel et α un automorphisme de V tel que $\alpha^2 = 1$. On pose $e = \frac{1}{2}(1 - \alpha)$; noter que cette égalité a un sens dans $\mathcal{L}(V)$.

(i) Vérifier que l'endomorphisme e de V est idempotent.

(ii) Vérifier que $\alpha(x) = x$ pour tout $x \in \mathrm{Ker}(e)$ et $\alpha(x) = -x$ pour tout $x \in \mathrm{Im}(e)$.

(III.10) Soient $\mathcal{P}_n$ des fonctions polynomiales de degrés au plus n, et soit $\mathcal{B}$ la base de $\mathcal{P}_n$ constituée des fonctions $t \longmapsto t^j$ $(0 \leq j \leq n)$. Calculer la matrice $M_{\mathcal{B},\mathcal{B}}(\alpha)$ de l'endomorphisme linéaire α de $\mathcal{P}_n$ défini par

(i) $(\alpha(f))(t) = f(t + 1)$ pour tout $t \in \mathbb{R}$,

(ii) $\alpha(f) = f'$ (dérivée de f).

(III.11) On désigne par $\mathcal{C} = (e_1, e_2, e_3)$ la base canonique de $\mathbb{R}^3$. Ecrire les matrices relativement à $\mathcal{C}$ de

(i) la projection $(x, y, z) \longmapsto (x, y, 0)$ «sur le plan horizontal» ,

(ii) une rotation d'un tiers de tour autour de l'axe engendré par le vecteur $e_1 + e_2 + e_3$.

(III.12) Soit $\mathcal{P}_3$ l'espace vectoriel des fonctions polynomiales de $\mathbb{R}$ dans $\mathbb{R}$ de degrés au plus 3, muni de la base $(f_j)_{j=0,1,2,3}$ définie par $f_j(t) = t^j$. On considère l'application $\alpha : \mathcal{P}_3 \longrightarrow \mathcal{P}_3$ qui applique une fonction polynomiale $P(t)$ sur le reste de la division de $(1 + t + t^2 + t^3)P(t)$ par t^4.

(i) Vérifier que l'application α est linéaire.

(ii) Ecrire sa matrice relativement à la base $(f_j)_{j=0,1,2,3}$.

(III.13) Etant donné la matrice $a = \begin{pmatrix} 0 & -2 & 3 & 5 \\ 2 & 3 & 1 & -2 \\ 6 & 5 & 9 & 4 \end{pmatrix}$, déterminer une base de son noyau et une base de son image. [Indication. Une méthode consiste à déterminer d'abord une base de $\mathrm{Ker}(a)$, puis de la compléter en une base de $\mathbb{R}^4$, et enfin d'en déduire une base de $\mathrm{Im}(a)$.]

(III.14) On note x, y et z les coordonnées d'un point de $\mathbb{R}^3$ relativement à la base canonique. Ecrire une matrice $a \in M_3(\mathbb{R})$ telle que $\mathrm{Im}(a)$ soit le plan d'équation $x + 2y + 3z = 0$.

(III.15) On considère les vecteurs $a = \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix}$, $b = \begin{pmatrix} 0 \\ 2 \\ 1 \end{pmatrix}$ et $d = \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}$ de $V = \mathbb{R}^3$, ainsi que la base canonique $\mathcal{C} = (e_1, e_2, e_3)$.

(i) Vérifier que $\mathcal{B} = (a, b, d)$ est une base de V.

(ii) Ecrire la matrice de changement de base $M_{\mathcal{C},\mathcal{B}}(id_V)$.

(iii) Résoudre les trois systèmes linéaires

$$\begin{cases} x & & + z = 1 \\ 2x + 2y + z = 0 \\ -x + y - z = 0 \end{cases} \qquad \begin{cases} x & & + z = 0 \\ 2x + 2y + z = 1 \\ -x + y - z = 0 \end{cases}$$

et

$$\begin{cases} x & & + z = 0 \\ 2x + 2y + z = 0 \\ -x + y - z = 1 \end{cases}.$$

(iv) Ecrire la matrice de changement de base $M_{\mathcal{B},\mathcal{C}}(id_V)$.

(v) Vérifier que les matrices $M_{\mathcal{C},\mathcal{B}}(id_V)$ et $M_{\mathcal{B},\mathcal{C}}(id_V)$ sont inverses l'une de l'autre.

Soit E le sous-espace vectoriel de V engendré par a et b, soit D la droite engendrée par d, et soit $\pi : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ la projection sur E parallèlement à D (= de noyau D).

(vi) Ecrire la matrice de π relativement à la base $\mathcal{B}$.

(vii) Ecrire la matrice de π relativement à la base $\mathcal{C}$.

(III.16) Soit a la matrice $\begin{pmatrix} 9 & 12 & 15 \\ 19 & 26 & 33 \\ 29 & 40 & 51 \end{pmatrix}$. Sachant que $a = \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$, pouvez-vous décider *sans calcul* si a est de rang 3 ?

(III.17) Soient V un espace vectoriel réel, $\phi : V \longrightarrow \mathbb{R}$ une forme linéaire et $x \in V$ un vecteur tel que $\phi(x) \neq 0$. Montrer que V est somme directe du noyau $\operatorname{Ker}\phi$ et de la droite $\mathbb{R}x$.

(III.18) Existe-t-il une application $\alpha \in \mathcal{L}(\mathbb{R}^3)$ telle que $\operatorname{Ker}(\alpha) = \operatorname{Im}(\alpha)$? une application $\beta \in \mathcal{L}(\mathbb{R}^4)$ telle que $\operatorname{Ker}(\alpha) = \operatorname{Im}(\alpha)$?

(III.19) Soit V l'espace des formes linéaires sur $\mathbb{R}^3$ et soit W l'espace des fonctions $\phi : \mathbb{R}^3 \longrightarrow \mathbb{R}$ de la forme $\phi(x,y,z) = ax^2 + by^2 + cz^2 + 2dyz + 2exz + 2fxy$, avec $a, b, c, d, e, f \in \mathbb{R}$. On considère l'application $D : W \longrightarrow V$ définie par

$$(D\phi)(x,y,z) = \frac{\partial\phi}{\partial x}(x,y,z) + \frac{\partial\phi}{\partial y}(x,y,z) + \frac{\partial\phi}{\partial z}(x,y,z)\,.$$

(i) Déterminer l'image de D.

(ii) Déterminer une base du noyau de D.

(III.20) (i) Dans l'espace $\mathbb{R}^2$ muni de la base canonique (e_1, e_2), on considère trois droites L_1, L_2, L_3 distinctes deux à deux, ainsi que les trois droites $D_1 = \mathbb{R}e_1$, $D_2 = \mathbb{R}e_2$, $D_3 = \mathbb{R}(e_1 + e_2)$. Montrer qu'il existe un endomorphisme linéaire α de $\mathbb{R}^2$ tel que $\alpha(D_j) = L_j$ pour $j \in \{1,2,3\}$.

(ii) Dans l'espace $\mathbb{R}^3$ muni de la base canonique (e_1, e_2, e_3), on considère quatre droites L_1, L_2, L_3, L_4 telles qu'auncun plan de $\mathbb{R}^3$ ne contienne trois d'entre elles, ainsi que les quatre droites $D_1 = \mathbb{R}e_1$, $D_2 = \mathbb{R}e_2$, $D_3 = \mathbb{R}e_3$, $D_4 = \mathbb{R}(e_1 + e_2 + e_3)$. Montrer qu'il existe un endomorphisme linéaire α de $\mathbb{R}^3$ tel que $\alpha(D_j) = L_j$ pour $j \in \{1,2,3,4\}$.

(iii) Formuler un énoncé analogue pour $\mathbb{K}^n$, où $\mathbb{K}$ est un corps et $n \geq 2$ un entier.

(III.21$^\sharp$) On considère l'ensemble $\mathbb{R}$ des nombres réels comme un espace vectoriel sur $\mathbb{Q}$ (sic!!!), et on souhaite montrer qu'il est de *dimension infinie*. (Bien sûr, le cas le plus fréquent est celui où on considère $\mathbb{R}$ comme espace vectoriel *réel*, et sa dimension est alors 1.)

Pour cela, on désigne par

$$\begin{aligned}\mathbb{P} = \{&2, 3, 5, 7, 11, \ldots, 1949, 1951, 1973, 1979, 1987, 1993, 1997, 1999,\\ &2003, 2011, 2017, 2027, 2029, 2039, 2053, 2063, 2069, 2081, 2083, 2087,\\ &2089, 2099, \ldots\}\end{aligned}$$

l'ensemble des nombres premiers, et on demande de montrer que $(\ln p)_{p\in\mathbb{P}}$ est une famille libre de $\mathbb{R}$. En d'autres termes, si $p_1, \ldots, p_k$ sont des nombres premiers distincts et si $\lambda_1, \ldots, \lambda_k$ sont des nombres *rationnels* tels que

$$\sum_{j=1}^{k} \lambda_j \ln p_j = 0,$$

on demande de montrer que $\lambda_j = 0$ pour tout $j \in \{1, \ldots, k\}$.

[*Indication* : Se ramener au cas où $\lambda_j \in \mathbb{Z}$ pour tout j, puis utiliser le théorème d'arithmétique selon lequel tout nombre entier positif possède une unique décomposition en produit de nombres premiers.]

CHAPITRE IV

Groupes symétriques et déterminants

1. Groupes : généralités

DÉFINITION. Un *groupe* est un ensemble G muni d'une opération $\begin{cases} G \times G \longrightarrow G \\ (a,b) \longmapsto ab \end{cases}$ satisfaisant aux conditions suivantes :

- associativité : $(ab)c = a(bc)$ pour tous $a, b, c \in G$;
- il existe un élément neutre $e \in G$ tel que $ea = ae = a$ pour tout $a \in G$;
- pour tout $a \in G$, il existe un élément, appelé « inverse » et noté a^{-1}, tel que $aa^{-1} = a^{-1}a = e$.

EXERCICE. L'élément neutre est unique, et tout élément a un *unique* inverse.

Dans ce cours l'opération est le plus souvent notée multiplicativement : on écrit ab le produit de deux éléments $a, b \in G$. Lorsque le groupe G est *abélien*[1] ou *commutatif,* c'est-à-dire lorsque $ab = ba$ pour tous $a, b \in G$, on utilise parfois la notation additive : $a + b$ au lieu de ab ; la commutativité s'écrit alors : $a + b = b + a$ pour tous $a, b \in G$.

L'élément neutre d'un groupe G se note souvent 1 au lieu de e, ou 1_G ou e_G s'il convient de préciser ; ou encore 0 en notation additive.

L'*ordre* d'un groupe fini est le nombre de ses éléments.

EXEMPLES. Les ensembles suivants sont des groupes :

(i) $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$, un espace vectoriel, pour l'addition.

(ii) $\{1, -1\}$, $\mathbb{Q}^*$, $\mathbb{R}_+^*$, $\mathbb{R}^*$, $\mathbb{C}^*$, $\big\{z \in \mathbb{C} \big| |z| = 1\big\}$, pour la multiplication usuelle. Observer que ces groupes *abéliens* sont notés *multiplicativement* !

---

[1] *Notice biographique.* L'adjectif « abélien » se réfère à Niels Henrik Abel, mathématicien norvégien, 1802–1829. Il a montré que les équations de degrés 5 et plus sont fondamentalement différentes de celles des degrés inférieurs. Plus précisément, il existe des formules avec radicaux fournissant en termes des coefficients $a, b, c, \ldots$ les racines d'une équation de degré deux ($ax^2 + bx + c = 0$), trois ($ax^3 + bx^2 + cx + d = 0$) ou quatre, mais pas en général pour les degrés supérieurs ($a_0x^n + a_1x^{n-1} + \cdots + a_n = 0$). Les autres travaux d'Abel, tout à fait fondamentaux, sont plus difficiles à évoquer rapidement. Il a créé (avec Jacobi, son cadet de deux ans) la théorie des fonctions elliptiques, qui permet notamment de comprendre les intégrales du type

$$\int \frac{dx}{\sqrt{P(x)}}$$

où $P(x)$ désigne un polynôme de degré 3 ou 4. Ses résultats sur les fonctions appelées depuis « fonctions abéliennes » fournissent une relation entre *toutes* les intégrales de certaines expressions et un certain nombre g de telles expressions *particulières*, où le nombre g (= genre) est une caractéristique géométrique de la courbe associée (d'équation $y^2 = P(x)$). Ses oeuvres complètes se limitent à deux volumes, environ 600 pages publiées de son vivant et 300 pages posthumes. Ces pages sont d'une inusable richesse, et l'importance des idées d'Abel pour le développement ultérieur des mathématiques est immense.

(iii) L'ensemble $GL(n,\mathbb{R})$ des matrices inversibles n-fois-n à coefficients réels, pour la multiplication usuelle des matrices, l'ensemble $\mathrm{Aut}(V)$ des automorphismes linéaires d'un espace vectoriel V, avec pour produit la composition des applications.

(iv) L'ensemble des applications bijectives d'un ensemble dans lui-même, pour la composition. Lorsque l'ensemble a un nombre fini n d'éléments, ce groupe est le *groupe symétrique*, noté $\mathrm{Sym}(n)$, et ses éléments sont appelés des *permutations* (de l'ensemble à n éléments).

Définition. Un *homomorphisme* d'un groupe G dans un groupe G' est une application $\phi : G \longrightarrow G'$ telle que $\phi(ab) = \phi(a)\phi(b)$ pour tous $a, b \in G$.

Exercice. Il en résulte que $\phi(1) = 1$ et $\phi(a^{-1}) = \phi(a)^{-1}$ pour tout $a \in G$.

Définition. Un *sous-groupe* d'un groupe G est un sous-ensemble H de G tel que $1 \in H$ et $ab, a^{-1} \in H$ pour tous $a, b \in H$.

Exemples d'homomorphismes : $\begin{cases} \mathbb{C}^* \longrightarrow \mathbb{R}_+^* \\ z \longmapsto |z| \end{cases}$, $\begin{cases} \mathbb{R} \longrightarrow \mathbb{R}_+^* \\ t \longmapsto e^t \end{cases}$, $\begin{cases} \mathbb{R}^* \longrightarrow \{1,-1\} \\ t \longmapsto \mathrm{sign}(t) \end{cases}$, une application linéaire entre espaces vectoriels, l'application $\begin{cases} \mathbb{Z} \longrightarrow G \\ n \longmapsto g^n \end{cases}$ où G est un groupe (noté multiplicativement) et g l'un de ses éléments.

2. Groupes symétriques

Pour tout entier $n \geq 1$, nous notons J_n l'ensemble $\{1, 2, \ldots, n\}$; rappelons que le groupe $\mathrm{Sym}(n)$ est l'ensemble de toutes les applications bijectives de J_n dans J_n, aussi appelées *permutations* de J_n, avec pour produit la composition des applications. C'est un groupe d'ordre $n!$. Le *support* d'une permutation $\sigma \in \mathrm{Sym}(n)$ est le sous-ensemble $\{j \in J_n \mid \sigma(j) \neq j\}$ de J_n.

Définition. Soit k un entier tel que $2 \leq k \leq n$, et $a_1, \ldots, a_k$ des éléments distincts de J_n. La permutation $\sigma \in \mathrm{Sym}(n)$ définie par

$$\sigma(i) = \begin{cases} a_{j+1} & \text{si } i = a_j \text{ pour un } j \text{ tel que } 1 \leq j < k \\ a_1 & \text{si } i = a_k \\ i & \text{sinon} \end{cases}$$

est un *k-cycle* ; on écrit

$$\sigma = (a_1\,,\,\ldots\,,\,a_k)$$

et on note que le support d'un tel cycle est le sous-ensemble $\{a_1, \ldots, a_k\}$ de J_n. Une *transposition* est un 2-cycle.

Remarques. (i) L'écriture d'un k-cycle n'est pas unique ; par exemple

$$(2\,,\,3\,,\,5) = (3\,,\,5\,,\,2) = (5\,,\,2\,,\,3)$$

dans $\mathrm{Sym}(5)$, et plus généralement

$$(a_1\,,\,\ldots\,,\,a_{k-1}\,,\,a_k) = (a_2\,,\,\ldots\,,\,a_k\,,\,a_1).$$

(ii) Il faut toutefois noter que l'ordre d'écriture des entiers a_j n'est pas arbitraire ! par exemple $(2\,,\,3\,,\,5) \neq (2\,,\,5\,,\,3)$.

(iii) L'identité est l'unique permutation de support vide. Il faut bien observer que, avec les notations adoptées, nous avons $\{2\,,\,3\,,\,5\} = \{2\,,\,5\,,\,3\}$ (comparer avec la remarque précédente).

(iv) Le support d'une permutation distincte de l'identité contient au moins deux éléments.

(v) Les deux éléments de Sym(2) sont l'identité et la transposition $(1\,,\,2)$. Il y a trois transposition dans Sym(3) qui sont $(1\,,\,2)$, $(1\,,\,3)$ et $(2\,,\,3)$. Plus généralement, il y a $\frac{n(n-1)}{2}$ transpositions dans Sym(n).

(vi) Les éléments d'ordre[2] 2 ne sont pas tous des transpositions ! En fait, ce sont précisément les produits de transpositions à supports disjoints ; par exemple, $(1\,,\,3)(7\,,\,8)$ et $(1\,,\,2)(3\,,\,4)(5\,,\,6)$ sont des éléments d'ordre deux dans Sym(8).

(vii) Soit X un ensemble fini à n éléments. Quitte à re-nommer les éléments, on peut voir Sym(n) comme le groupe de toutes les applications bijectives de X dans X.

Sur l'ordre des produits.

Nous convenons de composer (= multiplier) les permutations comme des applications : $(\sigma\tau)(i) = \sigma(\tau(i))$. Il en résulte qu'on a par exemple

$$(1\,,\,2)(2\,,\,3) \;=\; (1\,,\,2\,,\,3) \;\neq\; (1\,,\,3\,,\,2) \;=\; (2\,,\,3)(1\,,\,2).$$

Toutefois, il faut aussi noter que des cycles à supports disjoints commutent. Par exemple

$$(1\,,\,3)(2\,,\,5)(4\,,\,7) \;=\; (4\,,\,7)(2\,,\,5)(1\,,\,3) \;=\; (4\,,\,7)(1\,,\,3)(2\,,\,5)$$

et

$$(1\,,\,2\,,\,3)(4\,,\,5) \;=\; (4\,,\,5)(1\,,\,2\,,\,3).$$

Plus généralement des permutations à supports disjoints commutent.

PROPOSITION. *Toute permutation de* Sym(n) *s'écrit comme produit de cycles à supports disjoints deux à deux, et ceci de manière unique à l'ordre près des facteurs.*

PREUVE. Soit $\sigma \in \mathrm{Sym}(n)$. On définit un graphe orienté de la manière suivante : ses sommets sont les entiers $1, \ldots, n$, et ses arêtes orientées sont les « flèches » $i \to \sigma(i)$, pour $i \in J_n$. Ainsi ce graphe a-t-il exactement n sommets et n arêtes orientées ; parmi ces dernières, certaines sont des *boucles* dont l'extrémité coïncide avec l'origine, et d'autres sont des arêtes ayant chacune a une extrémité et une origine distinctes.

Le support de σ est le complément dans J_n des sommets correspondant aux boucles. Les autres arêtes de ce graphe orienté forment des chemins fermés de longueurs ≥ 2 qui fournissent la décomposition de la permutation en cycles. □

EXERCICE. Montrer que l'ordre d'une permutation est donné par le plus petit commun multiple des ordres des cycles dans la décomposition de la proposition précédente.

[2] L'*ordre* d'un élément g d'un groupe G est le plus petit entier $k \geq 1$ tel que $g^k = 1$, s'il existe, et le symbole ∞ sinon.

Il est *essentiel* de se familiariser avec un grand nombre de cas particuliers. On laisse au lecteur le soin de dessiner le graphe orienté correspondant à la permutation $\sigma \in \mathrm{Sym}(5)$ définie par $\sigma(1) = 3$, $\sigma(2) = 4$, $\sigma(3) = 5$, $\sigma(4) = 2$, $\sigma(5) = 1$, *ainsi que les graphes correspondant à de nombreux autres exemples.*

PROPOSITION. *Toute permutation $\sigma \in \mathrm{Sym}(n)$ s'écrit au moins d'une manière comme un produit de transpositions.*

PREUVE. On procède par récurrence sur n. Si $n \leq 2$, il n'y a rien à montrer. Sinon, on suppose la proposition vraie pour toute permutation dans $\mathrm{Sym}(n-1)$, et on distingue deux cas.

Lorsque $\sigma(n) = n$, l'assertion pour σ résulte immédiatement de l'hypothèse de récurrence.

Lorsque $\sigma(n) = k < n$, on considère la transposition $\tau = (k\,,\,n)$ et le produit $\tau\sigma$. Comme $(\tau\sigma)(n) = n$, il résulte du premier cas que $\tau\sigma$ est un produit de transpositions $\tau_1, \ldots, \tau_p$. Par suite $\sigma = \tau\big(\tau\sigma\big) = \tau \prod_{1\leq j\leq p} \tau_j$ est encore un produit de transpositions. □

REMARQUE. Si σ s'écrit comme un produit $\tau_1\tau_2 \ldots \tau_k$ de transpositions, alors $\sigma^{-1} = \tau_k\tau_{k-1} \ldots \tau_1$.

EXERCICE. Vérifier que

$$\begin{aligned} (1\,,\,2) &= (3\,,\,4)(1\,,\,2)(3\,,\,4) \\ (1\,,\,2)(2\,,\,3)(1\,,\,2) &= (2\,,\,3)(1\,,\,2)(2\,,\,3) = (1\,,\,3) \\ id &= \Big((1\,,\,2)(2\,,\,3)\ldots(n-2\,,\,n-1)(n-1\,,\,n)\Big)^n. \end{aligned}$$

En particulier, il n'y a pas unicité de la décomposition en produit de transpositions dans la proposition précédente.

THÉORÈME IV.1 (résultat principal du paragraphe). (i) *Soient σ une permutation et $\tau_1, \ldots, \tau_p$, $\rho_1, \ldots, \rho_q$ des transpositions dans* $\mathrm{Sym}(n)$ *telles que*

$$\sigma = \prod_{1\leq j\leq p} \tau_j = \prod_{1\leq k\leq q} \rho_k.$$

Alors $p - q$ est pair.

(ii) *L'application*

$$\mathrm{sign} : \begin{cases} \mathrm{Sym}(n) & \longrightarrow \{\pm 1\} \\ \sigma & \longmapsto \mathrm{sign}(\sigma) \end{cases}$$

qui applique une permutation σ sur $+1$ si elle est produit d'un nombre pair de transpositions et sur -1 sinon est un homomorphisme de groupes.

DÉFINITION. On dit qu'une permutation est *paire* si elle s'écrit comme un produit d'un nombre pair de transpositions, et *impaire* sinon. Le nombre $\mathrm{sign}(\sigma) \in \{\pm 1\}$ s'appelle la *signature* de la permutation σ.

Les permutations paires de $\mathrm{Sym}(n)$ constituent un sous-groupe appelé le *groupe alterné* $\mathrm{Alt}(n)$.

Preuve du théorème IV.1. Etant donné une permutation $\sigma \in \mathrm{Sym}(n)$ et une fonction $F : \mathbb{R}^n \longrightarrow \mathbb{R}$, on définit une nouvelle fonction ${}^{\sigma}F : \mathbb{R}^n \longrightarrow \mathbb{R}$ en posant

$$ {}^{\sigma}F(x_1, \ldots, x_n) \;=\; F\left(x_{\sigma(1)}, \ldots, x_{\sigma(n)}\right). $$

Pour $\sigma, \sigma' \in \mathrm{Sym}(n)$, nous avons[3]

$$ \begin{aligned} {}^{\sigma'}\left({}^{\sigma}F\right)(x_1, \ldots, x_n) \;&=\; \left({}^{\sigma}F\right)\left(x_{\sigma'(1)}, \ldots, x_{\sigma'(n)}\right) \\ &=\; F\left(x_{\sigma'\sigma(1)}, \ldots, x_{\sigma'\sigma(n)}\right) \\ &=\; {}^{\sigma'\sigma}F(x_1, \ldots, x_n) \end{aligned} $$

pour tous $(x_1, \ldots, x_n) \in \mathbb{R}^n$, c'est-à-dire

$$ {}^{\sigma'}\left({}^{\sigma}F\right) \;=\; {}^{\sigma'\sigma}F. $$

Considérons alors la fonction $\Delta : \mathbb{R}^n \longrightarrow \mathbb{R}$ définie par

$$ \Delta(x_1, \ldots, x_n) \;=\; \prod_{i<j} (x_i - x_j) $$

où le produit porte sur les $\frac{n(n-1)}{2}$ couples d'entiers (i, j) tels que $1 \leq i < j \leq n$. Observons que $\Delta \neq 0$, puisque $\Delta(x_1, \ldots, x_n) \neq 0$ chaque fois que les nombres $x_1, \ldots, x_n$ sont distincts deux à deux. [Par exemple $\Delta(x_1, x_2, x_3) = (x_1 - x_2)(x_1 - x_3)(x_2 - x_3)$.]

Affirmation. Si $\tau = (r, s)$ est une transposition, alors ${}^{\tau}\Delta = -\Delta$. En effet, dans le produit définissant Δ, les facteurs ne contenant ni x_r ni x_s ne sont pas modifiés par l'action de τ. Si on suppose $r < s$, les autres s'écrivent

$$ (x_r - x_s) \prod_{i=1}^{r-1} \Big((x_i - x_r)(x_i - x_s)\Big) \prod_{i=r+1}^{s-1} \Big((x_r - x_i)(x_i - x_s)\Big) \prod_{i=s+1}^{n} \Big((x_r - x_i)(x_s - x_i)\Big). $$

Le premier facteur change de signe lorsqu'on échange x_r et x_s, et chaque autre paire de facteurs ne change visiblement pas. Ainsi, l'effet global de τ sur Δ est bien *un* changement de signe.

Par suite, si $\sigma = \prod_{1 \leq j \leq p} \tau_j$ est un produit de transpositions, nous avons

$$ {}^{\sigma}\Delta \;=\; {}^{\tau_1}\left(\cdots {}^{\tau_{p-1}}\left({}^{\tau_p}\Delta\right)\cdots\right) \;=\; (-1)^p \Delta $$

et la parité de p ne dépend que de σ, non du produit $\sigma = \prod_{1 \leq j \leq p} \tau_j$ lui-même.

La preuve de l'assertion (i) est ainsi achevée, et celle de (ii) est laissée au lecteur. □

Exemple. La signature d'un k-cycle est $(-1)^{k-1}$, car

$$ (a_1\,,\, a_2\,,\, \ldots\,,\, a_k) \;=\; (a_1\,,\, a_2)(a_2\,,\, a_3)\ldots(a_{k-1}\,,\, a_k). $$

[3] Voici une justification de la deuxième égalité (la difficulté etant de ne pas confondre $\sigma\sigma'$ et $\sigma'\sigma$). Posons $y_1 = x_{\sigma'(1)}, \ldots, y_n = x_{\sigma'(n)}$. Alors $\left({}^{\sigma}F\right)\left(x_{\sigma'(1)}, \ldots, x_{\sigma'(n)}\right) = \left({}^{\sigma}F\right)(y_1, \ldots, y_n) = F\left(y_{\sigma(1)}, \ldots, y_{\sigma(n)}\right) = F\left(x_{\sigma'\sigma(1)}, \ldots, x_{\sigma'\sigma(n)}\right)$.

3. Déterminants : existence et unicité

Ce paragraphe et le suivant suivent de près l'exposé du chapitre VI de S. Lang : *Linear algebra*, paru chez Springer en 1971 (troisième édition 1987).

DÉFINITION. Un *déterminant* sur $\mathbb{R}^n$ est une application

$$D : \mathbb{R}^n \times \cdots \times \mathbb{R}^n \longrightarrow \mathbb{R} \qquad \text{(avec } n \text{ copies de } \mathbb{R}^n\text{)}$$

qui vérifie :

(i) pour tous $k \in \{1, \ldots, n\}$ et $a^1, \ldots, a^{k-1}, a^{k+1}, \ldots, a^n \in \mathbb{R}^n$, l'application

$$\begin{aligned} \mathbb{R}^n &\longrightarrow \mathbb{R}, \\ x &\longmapsto D(a^1, \ldots, a^{k-1}, x, a^{k+1}, \ldots, a^n) \end{aligned}$$

est linéaire ;

(ii) pour tous $k \in \{1, \ldots, n-1\}$ et $a^1, \ldots, a^k, a^{k+2}, \ldots, a^n \in \mathbb{R}^n$,

$$D(a^1, \ldots, a^{k-1}, a^k, a^k, a^{k+2}, \ldots, a^n) = 0;$$

(iii) si $(e^1, \ldots, e^n)$ désigne la base canonique de $\mathbb{R}^n$, alors $D(e^1, \ldots, e^n) = 1$.

REFORMULATION. Un déterminant est une application $D : M_n(\mathbb{R}) \longrightarrow \mathbb{R}$ telle que

(i) $D(a)$ dépend linéairement de chaque colonne de a ;

(ii) $D(a) = 0$ si a possède 2 colonnes adjacentes égales ;

(iii) $D(I_n) = 1$.

PROPOSITION. *Un déterminant $D : \mathbb{R}^n \times \ldots \times \mathbb{R}^n \longrightarrow \mathbb{R}$ sur $\mathbb{R}^n$ possède les propriétés suivantes :*

(iv) $D(a^1, \ldots, a^j, \ldots, a^k, \ldots, a^n) = -D(a^1, \ldots, a^k, \ldots, a^j, \ldots, a^n)$
pour tous $j, k \in \{1, \ldots, n\}$ tels que $j < k$;

(v) $D(a^1, \ldots, a^j, \ldots, a^k, \ldots, a^n) = 0$ *si* $a^j = a^k$ *pour tous* $j, k \in \{1, \ldots, n\}$ *tels que* $j < k$;

(vi) $D(a^1, \ldots, a^{j-1}, a^j + \lambda a^k, a^{j+1}, \ldots, a^n) = D(a^1, \ldots, a^i, \ldots, a^n)$ *pour tous* $j, k \in \{1, \ldots, n\}$ *tels que* $j \neq k$ *et* $\lambda \in \mathbb{R}$.

(vii) $D\left(a^{\sigma(1)}, \ldots, a^{\sigma(n)}\right) = \operatorname{sign}(\sigma) D\left(a^1, \ldots, a^n\right)$ *pour toute permutation* $\sigma \in \operatorname{Sym}(n)$.

PREUVE (ESQUISSE). Pour montrer (iv) lorsque $k = j + 1$, on développe

$$D(a^1, \ldots, a^{j-1}, a^j + a^{j+1}, a^j + a^{j+1}, a^{j+2}, \ldots, a^n) = 0.$$

De ce cas particulier de (iv), et de (ii), on déduit (v). La propriété (iv) dans le cas général résulte alors de (v) comme (iv) pour $k = j+1$ résulte de (ii). On montre ensuite (vi) en utilisant (i) et (v).

La propriété (vii) résulte enfin de (iv) et de l'existence d'une suite de transpositions dont le produit est égal à σ. □

REMARQUE. L'existence et l'unicité de D ne sont nullement évidentes (si ce n'est pour $n = 1$), et font l'objet du théorème IV.2.

Dans le cas particulier où $n = 2$, on peut procéder comme suit. Pour toute application $D : M_2(\mathbb{R}) \longrightarrow \mathbb{R}$ possédant les propriétés (i) à (iii), nous avons

$$\begin{aligned}
D\begin{pmatrix} a & b \\ c & d \end{pmatrix} &= aD\begin{pmatrix} 1 & b \\ 0 & d \end{pmatrix} + cD\begin{pmatrix} 0 & b \\ 1 & d \end{pmatrix} && \text{par (i)} \\
&= abD\begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} + adD\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + cbD\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + cdD\begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix} && \text{par (i)} \\
&= adD\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + cbD\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} && \text{par (ii)} \\
&= (ad - bc)D\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} && \text{par (iv)} \\
&= ad - bc && \text{par (iii)},
\end{aligned}$$

d'où l'unicité. L'existence vient de ce que l'application $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \longmapsto ad - bc$ vérifie les propriétés (i) à (iii).

Le lecteur qui souhaite un exposé du cas $n = 3$ en trouvera un aux pages 143–147 du livre de S. Lang. (Par ailleurs, il existe une formule presque toujours inutile pour les calculs et hélas bien trop souvent utilisée pour le cas $n = 3$.)

NOTATION. Pour une matrice $a \in M_n(\mathbb{R})$ et deux indices $i, j \in \{1, \ldots, n\}$, on désigne par $A_{i,j} \in M_{n-1}(\mathbb{R})$ la matrice qu'on obtient à partir de a en supprimant la i-ème ligne et la j-ème colonne. Si par exemple $a = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$, alors $A_{1,3} = \begin{pmatrix} 4 & 5 \\ 7 & 8 \end{pmatrix}$.

PROPOSITION. *Il existe un déterminant $M_n(\mathbb{R}) \longrightarrow \mathbb{R}$.*

PREUVE (ESQUISSE). La proposition est évidente lorsque $n = 1$ (et presqu'immédiate – voir ci-dessus – lorsque $n = 2$). On procède par récurrence sur n en supposant $n \geq 2$ et en supposant qu'il existe un déterminant $D : M_{n-1}(\mathbb{R}) \longrightarrow \mathbb{R}$.

Choisissons (arbitrairement) un entier $i \in \{1, \ldots, n\}$ et définissons $F : M_n(\mathbb{R}) \longrightarrow \mathbb{R}$ par

$$F(a) = (-1)^{i+1} a_{i,1} D(A_{i,1}) + \cdots + (-1)^{i+n} a_{i,n} D(A_{i,n}).$$

Il reste à vérifier que l'application F possède les propriétés (i) à (iii), ce dont nous laissons le détail au lecteur. □

REMARQUE. Il n'est nullement évident jusqu'ici que l'application F de la preuve ci-dessus est indépendante du choix de i, mais c'est vrai vu ce qui suit.

PROPOSITION (Lemme clé). *On considère des vecteurs $b^1, \ldots, b^n \in \mathbb{R}^n$ et une matrice carrée $c = (c_{j,k})_{1 \leq j,k \leq n} \in M_n(\mathbb{R})$; on pose $a^k = \sum_{j=1}^n c_{j,k} b^j \in \mathbb{R}^n$ pour tout $k \in \{1, \ldots, n\}$. Alors*

$$D(a^1, \ldots, a^n) = \sum_{\sigma \in \mathrm{Sym}(n)} \mathrm{sign}(\sigma) c_{\sigma(1),1} \cdots c_{\sigma(n),n} D(b^1, \ldots, b^n).$$

PREUVE. Par linéarité, nous avons d'abord

$$D(a^1,\dots,a^n) \;=\; \sum_{\sigma} c_{\sigma(1),1}\cdots c_{\sigma(n),n} D(b^{\sigma(1)},\dots,b^{\sigma(n)}),$$

où la somme porte sur n^n termes, et plus précisément sur TOUTES les applications σ de $\{1,\dots,n\}$ dans $\{1,\dots,n\}$. Mais, par la propriété (v) ci-dessus, chaque terme correspondant à une application σ non injective est nul, de sorte que

$$D(a^1,\dots,a^n) \;=\; \sum_{\sigma\in \mathrm{Sym}(n)} c_{\sigma(1),1}\cdots c_{\sigma(n),n} D(b^{\sigma(1)},\dots,b^{\sigma(n)})$$

(sommes sur $n!$ termes, et plus précisément sur les permutations $\sigma \in \mathrm{Sym}(n)$). La propriété (vii) ci-dessus permet de conclure. □

THÉORÈME IV.2. *Pour tout $n \geq 1$, il existe un unique déterminant $D : M_n(\mathbb{R}) \longrightarrow \mathbb{R}$. De plus, si $a \in M_n(\mathbb{R})$, alors*

$$(\sharp) \qquad D(a) \;=\; \sum_{\sigma\in \mathrm{Sym}(n)} \mathrm{sign}(\sigma) a_{\sigma(1),1} a_{\sigma(2),2}\cdots a_{\sigma(n),n}.$$

PREUVE. Soit d'abord D *un* déterminant sur $\mathbb{R}^n$. On choisit pour $b^1,\dots,b^n$ les vecteurs $e^1,\dots,e^n$ de la base canonique de $\mathbb{R}^n$, de sorte que la k-ième colonne a^k de la matrice a s'écrit $a^k = \sum_{i=1}^n a_{i,k} e^i$. Le lemme-clé montre alors que $D(a)$ est donné par la formule $(\sharp)$. L'unicité du théorème en résulte. □

NOTATION. On écrit désormais $\det(a)$ le déterminant d'une matrice carrée a.

REMARQUE. Il est possible d'exposer la théorie des déterminants en les définissant par la formule $(\sharp)$.

COROLLAIRE (développement d'un déterminant relativement à la i-ème ligne). *Si $a \in M_n(\mathbb{R})$, alors*

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

pour tout $i \in \{1,\dots,n\}$.

PREUVE. Voir la preuve de l'existence du déterminant sur $\mathbb{R}^n$. □

4. Déterminants : calculs

Le théorème IV.2 fournit une formule pour le déterminant d'une matrice d'ordre n qui contient $n!$ termes. Or la fonction factorielle croit rapidement :

$$3! = 6 \quad , \qquad 4! = 24 \quad , \qquad 100! \approx 0,9 \times 10^{138} \quad , \qquad 1000! \approx 4 \times 10^{2567}.$$

A raison d'un million de termes par seconde, un calculateur opiniâtre peut traiter

$$365 \times 24 \times 60 \times 60 \times 10^6 \;\approx\; 3 \times 10^{13}$$

termes par an, donc au mieux environ 10^{25} termes pendant une période proche de l'âge estimé de l'univers. Il convient donc d'avoir à disposition d'autres méthodes de calculs de déterminants, Y COMPRIS POUR LE CAS $n = 3$.

Cas des matrices triangulaires. Pour une matrice triangulaire supérieure

$$a = \begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ 0 & a_{2,2} & \dots & a_{2,n} \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & a_{n,n} \end{pmatrix}$$

($a_{i,j} = 0$ si $i > j$), on a

$$(*) \qquad \det(a) = \prod_{i=1}^{n} a_{i,i}.$$

Ceci vaut pour le cas particulier important des matrices diagonales.

A propos des matrices triangulaires, voir le théorème du § V.3.

Preuve. Rappelons que, pour toute matrice $a \in M_n(\mathbb{R})$, nous avons

$$\det(a) = \sum_{j=1}^{n} (-1)^{n+j} a_{n,j} \det(A_{n,j}).$$

Si a est comme ci-dessus, c'est-à-dire triangulaire, il vient $\det(a) = a_{n,n} \det(A_{n,n})$, où $A_{n,n}$ est une matrice triangulaire dans $M_{n-1}(\mathbb{R})$. Un argument par récurrence standard achève de montrer la formule $(*)$. □

Remarque. On peut aussi montrer la formule $(*)$ à l'aide du théorème IV.2.

Exercice. Enoncer et montrer le résultat analogue pour les matrices triangulaires inférieures.

Proposition. *Si ${}^t a$ désigne la matrice transposée d'une matrice carrée a, alors*

$$\det({}^t a) = \det(a).$$

Preuve. La valeur d'un produit de n nombres réels étant indépendante de l'ordre des facteurs, nous avons pour toute permutation $\rho \in \mathrm{Sym}(n)$

$$a_{\rho(1),1} a_{\rho(2),2} \cdots a_{\rho(n),n} = a_{1,\rho^{-1}(1)} a_{2,\rho^{-1}(2)} \cdots a_{n,\rho^{-1}(n)} \ ;$$

de plus $\mathrm{sign}(\rho^{-1}) = \mathrm{sign}(\rho)$. Par suite

$$\begin{aligned} \det(a) &= \sum_{\rho \in \mathrm{Sym}(n)} \mathrm{sign}(\rho) a_{\rho(1),1} a_{\rho(2),2} \cdots a_{\rho(n),n} \\ &= \sum_{\rho \in \mathrm{Sym}(n)} \mathrm{sign}(\rho^{-1}) a_{1,\rho^{-1}(1)} a_{2,\rho^{-1}(2)} \cdots a_{n,\rho^{-1}(n)} \\ &\overset{***}{=} \sum_{\sigma \in \mathrm{Sym}(n)} \mathrm{sign}(\sigma) a_{1,\sigma(1)} a_{2,\sigma(2)} \cdots a_{n,\sigma(n)} \\ &= \sum_{\sigma \in \mathrm{Sym}(n)} \mathrm{sign}(\sigma) ({}^t a)_{\sigma(1),1} ({}^t a)_{\sigma(2),2} \cdots ({}^t a)_{\sigma(n),n} = \det({}^t a). \end{aligned}$$

(Comparer l'égalité $\overset{***}{=}$ avec un changement de variable sous un signe d'intégration, par exemple avec $\int_0^1 (1-s) ds = \int_0^1 t dt$.) □

COROLLAIRE. *Le déterminant d'une matrice $a \in M_n(\mathbb{R})$ dépend linéairement de chaque ligne de a.*

PREUVE. Le déterminant $\det(a) = \det({}^t a)$ dépend linéairement de chaque colonne de ${}^t a$, c'est-à-dire de chaque ligne de a. □

COROLLAIRE (développement d'un déterminant relativement à la j-ème colonne). *Si $a \in M_n(\mathbb{R})$, alors*

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

pour tout $j \in \{1, \ldots, n\}$.

PREUVE. Cela résulte de la proposition précédente et des formules de développement des déterminants relativement aux lignes ; voici l'argument en détail.

Pour toute paire (i, j) d'indices dans $\{1, \ldots, n\}$, désignons par ${}^t A_{j,i} \in M_{n-1}(\mathbb{R})$ la matrice qu'on obtient à partir de ${}^t a$ en supprimant la j-ème ligne et la i-ème colonne. Si par exemple $a = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$, alors ${}^t A_{1,3} = \begin{pmatrix} 2 & 5 \\ 3 & 6 \end{pmatrix}$.

Attention : distinguer ${}^t A_{j,i}$ de la transposée de $A_{j,i}$! En fait ${}^t A_{j,i}$ est la matrice transposée de $A_{i,j}$, et c'est la proposition précédente qui implique que

$$\det({}^t A_{j,i}) \;=\; \det({}^t(A_{i,j})) \;=\; \det(A_{i,j}).$$

En développant la déterminant de ${}^t a$ relativement à la j-ième ligne, on obtient

$$\det(a) \;=\; \det({}^t a) \;=\; \sum_{i=1}^{n} (-1)^{j+i} ({}^t a)_{j,i} \det({}^t A_{j,i}) \;=\; \sum_{i=1}^{n} (-1)^{i+j} a_{i,j} \det(A_{i,j}),$$

ceci pour tout $j \in \{1, \ldots, n\}$. □

EXEMPLE DE CALCUL. Soit $a = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 8 \end{pmatrix}$. Il y a de nombreuses manières de procéder.

En voici une première. Le déterminant de a est égal à celui de la matrice obtenue à partir de a en soustrayant 4 fois la 1ère ligne à la 2ème ligne et 7 fois la 1ère ligne à la 3ème ligne ; ainsi :

$$\det(a) \;=\; \begin{pmatrix} 1 & 2 & 3 \\ 0 & -3 & -6 \\ 0 & -6 & -13 \end{pmatrix}.$$

En développant par rapport à la 1ère colonne et en utilisant la linéarité du déterminant par rapport à chaque ligne, on trouve

$$\det(a) \;=\; \det \begin{pmatrix} -3 & -6 \\ -6 & -13 \end{pmatrix} \;=\; -3 \det \begin{pmatrix} 1 & 2 \\ -6 & -13 \end{pmatrix} \;=\; 3 \det \begin{pmatrix} 1 & 2 \\ 6 & 13 \end{pmatrix}$$

de sorte que $\det(a) = 3(13 - 12) = 3$.

En voici une autre. On remarque que la matrice $b = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$ a ses lignes linéairement dépendantes, donc que son déterminant est nul. Par linéarité du déterminant en la troisième ligne, on obtient donc

$$\det(a) = \det(b) - \det \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 0 & 0 & 1 \end{pmatrix} = 0 - \det \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 0 & 0 & 1 \end{pmatrix}.$$

On peut donc développer par rapport à la 3ème ligne :

$$\det(a) = -\det \begin{pmatrix} 1 & 2 \\ 4 & 5 \end{pmatrix} = -(5-8) = 3.$$

THÉORÈME IV.3 (multiplicativité du déterminant). *Pour $b, c \in M_n(\mathbb{R})$, on a*

$$\det(bc) = \det(b)\det(c).$$

PREUVE. Posons $a = bc$. Notons $a^1, \ldots, a^n$ les colonnes de a, et de même pour $b^1, \ldots, b^n$ et $c^1, \ldots, c^n$. Par définition du produit de deux matrices, la i-ème coordonnée du k-ième vecteur colonne de la matrice a est donné par

$$(a^k)_i = a_{i,k} = \sum_{j=1}^{n} b_{i,j}c_{j,k} = \sum_{j=1}^{n} c_{j,k}(b^j)_i = \left(\sum_{j=1}^{n} c_{j,k}b^j\right)_i$$

$(i, k \in \{1, \ldots, n\})$; en d'autres termes :

$$a^k = \sum_{j=1}^{n} c_{j,k}b^j \qquad \text{pour tout} \quad k \in \{1, \ldots, n\}.$$

En appliquant le lemme-clé, on obtient

$$\begin{aligned} \det(a) = \det(a^1, \ldots, a^n) &= \sum_{\sigma \in \mathrm{Sym}(n)} \mathrm{sign}(\sigma) c_{\sigma(1),1} \cdots c_{\sigma(n),n} \det(b^1, \ldots, b^n) \\ &= \det(b) \sum_{\sigma \in \mathrm{Sym}(n)} \mathrm{sign}(\sigma) c_{\sigma(1),1} \cdots c_{\sigma(n),n} \end{aligned}$$

c'est-à-dire $\det(a) = \det(b)\det(c)$. □

THÉORÈME IV.4. *Soit $a \in M_n(\mathbb{R})$. Alors la matrice a est inversible si et seulement si son déterminant* $\det(a)$ *est différent de* 0 *; dans ce cas,* $\det(a^{-1}) = \det(a)^{-1}$.

PREUVE. Si a est inversible, alors $aa^{-1} = I_n$, donc

$$\det(a)\det(a^{-1}) = \det(aa^{-1}) = \det(I_n) = 1,$$

de sorte que $\det(a) \neq 0$ et $\det(a^{-1}) = \det(a)^{-1}$.

Si a n'est pas inversible, ses colonnes sont linéairement dépendantes, de sorte qu'il existe des nombres réels $\lambda_1, \ldots, \lambda_n$ non tous nuls tels que $\lambda_1 a^1 + \cdots + \lambda_n a^n = 0$. Soit

$j \in \{1, \dots, n\}$ tel que $j \neq 0$; on ne restreint pas la généralité de l'argument en supposant $\lambda_j = 1$. Alors

$$\det(a) = \det(a^1, \dots, a^n) = \det\Big(a^1, \dots, a^{j-1}, a^j + \sum_{1 \leq i < j \leq n} \lambda_i a^i, a^{j+1}, \dots, a^n\Big)$$
$$= \det(a^1 \dots, a^{j-1}, 0, a^{j+1}, \dots, a^n) = 0. \qquad \square$$

Proposition (Reformulation d'une partie de ce qui précède). *La restriction aux matrices inversibles du déterminant fournit un homomorphisme de groupes*

$$\det : GL(n, \mathbb{R}) \longrightarrow \mathbb{R}^*.$$

La notation $GL(n, \mathbb{R})$ a été définie au § IV.1.

Exercice. Écrire quelques matrices du noyau de cet homomorphisme.

Corollaire. *Deux matrices semblables ont même déterminant.*

Preuve. Deux matrices a et b étant par définition semblables s'il existe une matrice inversible s telle que $a = s^{-1}bs$, ce corollaire est une conséquence immédiate de la proposition et du théorème qui précèdent. $\square$

Proposition. *Soient p, q deux entiers positifs, et $n = p + q$. Soit $s = \begin{pmatrix} a & b \\ 0 & d \end{pmatrix} \in M_n(\mathbb{R})$ une matrice écrite par blocs, avec $a \in M_p(\mathbb{R})$, $b \in M_{p,q}(\mathbb{R})$ et $d \in M_q(\mathbb{R})$. Alors*

$$\det(s) = \det(a) \det(d).$$

Preuve. Procédons par récurrence sur p. Supposons d'abord $p = 1$. Alors

$$\det(s) = \det \begin{pmatrix} a_{1,1} & b_* & \cdots & b_* \\ 0 & d_* & \cdots & d_* \\ \vdots & \vdots & \vdots & \vdots \\ 0 & d_* & \cdots & d_* \end{pmatrix} = a_{1,1} \det(d)$$

par développement du déterminant selon la première ligne. [La notation x_* indique une paire d'indices $*$ qu'il est inutile de préciser.]

Supposons ensuite $p > 1$ et la proposition démontrée jusqu'à $p - 1$. En développant selon la première colonne, nous obtenons

$$\det(s) = \sum_{i=1}^{n} (-1)^{i+1} s_{i,1} \det(S_{i,1}) = \sum_{i=1}^{p} (-1)^{i+1} a_{i,1} \det \begin{pmatrix} A_{i,1} & * \\ 0 & d \end{pmatrix}$$
$$\overset{\text{hyp. réc.}}{=} \sum_{i=1}^{p} (-1)^{i+1} a_{i,1} \det(A_{i,1}) \det(d) = \det(a) \det(d),$$

ce qui achève la preuve. $\square$

Un *système de Cramer*[4] est un système linéaire de n équations à n inconnues, de la forme

$$ax = b,$$

où $a \in M_n(\mathbb{R})$ est une matrice inversible et où $b \in \mathbb{R}^n$.

[4]Gabriel Cramer, 1704–1752, a donné son nom à une rue parallèle à la rue de la Servette.

PROPOSITION. *Pour une matrice inversible $a = (a^1, \dots, a^n) \in M_n(\mathbb{R})$ et un vecteur $b \in \mathbb{R}^n$, le système de Cramer $ax = b$ possède une unique solution, donnée par*

$$x_i \;=\; \frac{1}{\det(a)} \det\left(a^1, \dots, a^{i-1}, b, a^{i+1}, \dots, a^n\right)$$

pour $i \in \{1, \dots, n\}$.

REMARQUE. Ces formules ont un intérêt théorique, mais sont pratiquement presque toujours *inappliquables* pour les calculs pratiques (voir le début du § IV.4).

PREUVE. Nous savons déjà qu'un système de Cramer possède une unique solution donnée par $x = a^{-1}b$. Soient $x_1, \dots, x_n$ les coordonnées du vecteur x (c'est-à-dire les n inconnues du système) relativement à la base canonique $e^1, \dots, e^n$, de sorte que $s = \sum_{j=1}^n x_j e^j$. Alors l'égalité

$$b \;=\; ax \;=\; \sum_{j=1}^n x_j a e^j \;=\; \sum_{j=1}^n x_j a^j$$

implique

$$\begin{aligned} \det(a^1, \dots, a^{i-1}, b, a^{i+1}, \dots, a^n) \;&=\; \det(a^1, \dots, a^{i-1}, \sum_{j=1}^n x_j a^j, a^{i+1}, \dots, a^n) \\ &= \sum_{j=1}^n x_j \det(a^1, \dots, a^{i-1}, a^j, a^{i+1}, \dots, a^n) \;=\; x_i \det(a) \end{aligned}$$

car la dernière somme n'a qu'un terme non nul, celui pour lequel $j = i$. Les « formules de Cramer » en résultent. □

5. Déterminants : compléments

Voici quelques compléments, sans démonstration.

Soit $a \in M_{m,n}(\mathbb{R})$ une matrice, non nécessairement carrée. Pour tout entier $k \leq \min(m, n)$, on appelle *mineur d'ordre k de a* un nombre de la forme

$$\det \begin{pmatrix} a_{i_1,j_1} & a_{i_1,j_2} & \dots & a_{i_1,j_k} \\ a_{i_2,j_1} & a_{i_2,j_2} & \dots & a_{i_2,j_k} \\ \dots & \dots & \dots & \dots \\ a_{i_k,j_1} & a_{i_k,j_2} & \dots & a_{i_k,j_k} \end{pmatrix}$$

avec $1 \leq i_1 < i_2 < \dots < i_k \leq m$ et $1 \leq j_1 < j_2 \dots < j_k \leq n$.

EXEMPLES. Considérons la matrice

$$a \;=\; \begin{pmatrix} m & n & p & q \\ r & s & t & u \\ v & w & x & y \end{pmatrix}.$$

Il y a autant de 2-mineurs de a que de choix de 2 des 3 lignes et de 2 des 4 colonnes, c'est-à-dire $3 \times 6 = 24$ 2-mineurs. En voici quelques-uns :

$$\det \begin{pmatrix} m & p \\ r & t \end{pmatrix} = mt - pr \;\; (i_1 = 1,\, i_2 = 2,\, j_1 = 1,\, j_2 = 2)\,;$$

$$\det\begin{pmatrix} m & q \\ v & y \end{pmatrix} = my - qv \ (i_1 = 1,\ i_2 = 3,\ j_1 = 1,\ j_2 = 4)\,;$$

$$\det\begin{pmatrix} t & u \\ x & y \end{pmatrix} = ty - ux \ (i_1 = 2,\ i_2 = 3,\ j_1 = 3,\ j_2 = 4).$$

THÉORÈME IV.5. *Avec les notations précédentes, le rang de a est le maximum des entiers k tels qu'il existe un mineur d'ordre k qui est non nul.*

PREUVE. Elle ne fait pas partie du cours. Voir par exemple le numéro 9 du paragraphe 19 dans le livre de Roger Godement, *Cours d'algèbre,* Hermann, 1963. Noter toutefois que le cas d'une matrice carrée $a \in M_n(\mathbb{R})$ de rang n est couvert par le théorème précédent. □

Interprétation géométrique

Soient a^1, a^2 deux vecteurs du plan euclidien ; alors $|\det(a^1, a^2)|$ est l'aire du parallélogramme de sommets $0, a^1, a^2, a^1 + a^2$.

Soient a^1, a^2, a^3 trois vecteurs de l'espace euclidien de dimension 3 usuel ; alors $|\det(a^1, a^2, a^3)|$ est le volume du pavé

$$P = \left\{ x \in \mathbb{R}^3 \;\middle|\; \begin{array}{l} \text{il existe } x_1, x_2, x_3 \in [0,1] \\ \text{tels que } x = x_1 a^1 + x_2 a^2 + x_3 a^3 \end{array} \right\}$$

de côtés a^1, a^2, a^3.

Plus généralement, pour des vecteurs $a^1, \ldots, a^n$ de l'espace euclidien[5] $\mathbb{R}^n$, la valeur absolue du déterminant de $(a^1, \ldots, a^n)$ est le volume du pavé défini par ces vecteurs.

Les matrices des cofacteurs.

Soient $n \geq 2$ un entier et $a = (a_{j,k})_{1 \leq j,k \leq n} \in M_n(\mathbb{R})$. Pour $j, k \in \{1, \ldots, n\}$, on désigne comme plus haut par $A_{j,k} \in M_{n-1}(\mathbb{R})$ la matrice obtenue à partir de a par supression de la $j$ième ligne et de la $k$ième colonne. La *matrice des cofacteurs de a* est la matrice $\mathrm{co}(a) \in M_n(\mathbb{R})$ définie par

$$\mathrm{co}(a)_{j,k} = (-1)^{j+k} \det(A_{k,j}) \qquad (1 \leq j, k \leq n).$$

Par exemple,

$$\mathrm{co}\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

et

$$\mathrm{co}\begin{pmatrix} a & b & c \\ d & e & f \\ g & h & j \end{pmatrix} = \begin{pmatrix} ej - fh & -(bj - ch) & bf - ce \\ * & * & * \\ * & * & * \end{pmatrix}$$

(exercice : remplacer les $*$ par les expressions convenables).

Affirmation. Avec les notations ci-dessus :

$$a\,\mathrm{co}(a) = \mathrm{co}(a)a = \det(a) I_n.$$

[5] La définition de ce terme fera partie d'un chapitre ultérieur

Démonstration. Pour $j, l \in \{1, \ldots, n\}$, nous avons

$$(*) \qquad (a\,\mathrm{co}(a))_{j,l} = \sum_{k=1}^{n} a_{j,k}(-1)^{k+l} \det(A_{l,k})$$

par définition du produit des matrices a et $\mathrm{co}(a)$. Notons $a_{(1)}, \ldots, a_{(n)} \in M_{1,n}(\mathbb{R})$ les lignes de a. Le terme de droite dans $(*)$ est le développement relativement à la $l$ième ligne de la matrice dont les lignes sont $a_{(1)}, \ldots, a_{(l-1)}, a_{(j)}, a_{(l+1)}, \ldots, a_{(n)} \in M_{1,n}(\mathbb{R})$. Ce déterminant est donc $\det(a)$ si $j = l$ et 0 si $j \neq l$. Par suite $a\,\mathrm{co}(a) = \det(a) I_n$.

La vérification de $\mathrm{co}(a)a = \det(a) I_n$ est analogue.

6. Déterminant d'un endomorphisme

Soient V un espace vectoriel de dimension finie et $\alpha \in \mathcal{L}(V)$. Rappelons que, étant donné deux bases $\mathcal{B}, \mathcal{C}$ de V, les matrices $M_{\mathcal{B},\mathcal{B}}(\alpha)$ et $M_{\mathcal{C},\mathcal{C}}(\alpha)$ sont semblables ; elles ont donc même déterminant.

On définit le déterminant $\det(\alpha)$ de α comme le déterminant de ces matrices. Alors $\det(\alpha) \neq 0$ si et seulement si α est un automorphisme.

Propriétés. Pour $\alpha, \beta \in \mathcal{L}(V)$, on a $\det(\beta\alpha) = \det(\beta)\det(\alpha)$. En particulier, la restriction du déterminant aux automorphismes de V fournit un homomorphisme de groupes

$$\det : GL(V) \longrightarrow \mathbb{R}^*$$

où, comme d'habitude, $GL(V)$ désigne le *groupe général linéaire* de V (le groupe des automorphismes linéaires de l'espace vectoriel V) et $\mathbb{R}^*$ le groupe multiplicatif des nombres réels non nuls.

Exemples, pour V un espace euclidien de dimension 2 (voir le début du § III.4). Le déterminant d'une rotation est 1. Le déterminant d'une projection orthogonale sur une droite est 0. Le déterminant d'une symétrie relativement à une droite est -1.

7. Résumé de quelques propriétés des déterminants

La théorie s'étend sans aucune difficulté au cas de matrices à coefficients dans un corps $\mathbb{K}$ quelconque, par exemple $\mathbb{R}$, $\mathbb{C}$, ou $\mathbb{Q}$, et même à des cas plus généraux, par exemple aux matrices à coefficients dans $\mathbb{Z}$.

Les paragraphes précédents montrent que, pour tout entier $n \geq 1$, il existe une application

$$\det : M_n(\mathbb{K}) \longrightarrow \mathbb{K}\ ,\ a \longmapsto \det(a)$$

avec les propriétés suivantes :

(I) $\det(a)$ est linéaire en chaque colonne de la matrice a, ainsi qu'en chaque ligne de a.

(II) $\det(a^{\sigma(1)}, \ldots, a^{\sigma(n)}) = \mathrm{sign}(\sigma)\det(a^1, \ldots, a^n)$ pour tous $a^1, \ldots, a^n \in \mathbb{K}^n$ et $\sigma \in \mathrm{Sym}(n)$; en particulier, $det(a) = 0$ s'il existe deux colonnes distinctes de a qui sont égales. De même, $\det(a) = 0$ s'il existe deux lignes distinctes de a qui sont égales.

(III) $\det(I_n) = 1$.

(IV) $\det(ab) = \det(a)\det(b)$.

(V) $\det(a) = \det({}^t a)$.

(VI) $\det\left(a^1 + \sum_{k=2}^n \lambda_k a^k, a^2, \ldots, a^n\right) = \det(a^1, \ldots, a^n)$ pour tous $\lambda_2, \ldots, \lambda_n \in \mathbb{K}$; de même si l'on ajoute à la $j$ème colonne une combinaison linéaire des autres colonnes ; formules alanogues pour les lignes.

(VII) $\det(a) = \sum_{j=1}^n (-1)^{i+j} a_{i,j} \det(A_{i,j})$ pour tout $i \in \{1, \ldots, n\}$ (développement selon la i-ème ligne).

(VIII) $\det(a) = \sum_{i=1}^n (-1)^{i+j} a_{i,j} \det(A_{i,j})$ pour tout $j \in \{1, \ldots, n\}$ (développement selon la j-ème colonne).

(IX) $\det(a) = \sum_{\sigma \in \mathrm{Sym}(n)} \mathrm{sign}(\sigma) a_{\sigma(1),1} \cdots a_{\sigma(n),n}$.

(X) $\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \det(a)\det(d)$ pour une matrice $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ triangulaire par blocs.

8. Exercices

(IV.1) Vérifier que, pour tout entier $d \geq 1$, l'ensemble

$$\begin{aligned} \mu(d) &= \{z \in \mathbb{C} \mid z^d = 1\} \\ &= \left\{ z \in \mathbb{C} \;\middle|\; \text{il existe} \quad j \in \mathbb{Z} \quad \text{tel que} \quad z = \exp\left(\frac{i2\pi}{d} j\right) \right\} \end{aligned}$$

des racines d-ièmes de l'unité est un groupe abélien d'ordre d.

(IV.2) Quand $\begin{cases} G \longrightarrow G \\ g \longmapsto g^{-1} \end{cases}$ est-il un homomorphisme d'un groupe G dans lui-même ? Même question pour $\begin{cases} G \longrightarrow G \\ g \longmapsto g^2 \end{cases}$.

(IV.3) Pour un homomorphisme $\phi : G \longrightarrow G'$, vérifier que le *noyau*

$$\mathrm{Ker}(\phi) = \{g \in G \mid \phi(g) = 1\}$$

de ϕ est un sous-groupe de G et que *l'image*

$$\mathrm{Im}(\phi) = \{g' \in G' \mid \text{il existe } g \in G \text{ tel que } \phi(g) = g'\}$$

est un sous-groupe de G'.

(IV.4) Ecrire la décomposition en cycles du produit $(1\,,\,2\,,\,3)(2\,,\,3\,,\,4)$ dans $\mathrm{Sym}(4)$.

(IV.5) Ecrire les décompositions en cycles des permutations du groupe $\mathrm{Alt}(4)$.

(IV.6) Quel est l'ordre du groupe symétrique $\mathrm{Sym}(n)$?

(IV.7) Sur un «échiquier» de n-fois-n cases, compter le nombre de manières de disposer n tours de telle sorte qu'aucune tour ne mette en échec aucune autre.

(IV.8) Soit Sym(5) le groupe des permutations de l'ensemble $\{1,2,3,4,5\}$ et soit $\sigma \in \mathrm{Sym}(5)$ la permutation définie par $\sigma(1) = 2$, $\sigma(2) = 1$, $\sigma(3) = 4$, $\sigma(4) = 5$ et $\sigma(5) = 3$.

(i) Quel est le plus petit entier $k \geq 1$ tel que σ^k soit la transformation identique ?

(ii) Trouver tous les éléments $\tau \in \mathrm{Sym}(5)$ tels que $\tau\sigma = \sigma\tau$.

(IV.9) Calculer le déterminant de la matrice $\begin{pmatrix} 3 & 0 & 1 \\ 1 & 2 & 5 \\ -1 & 4 & 2 \end{pmatrix}$.

[*Indication* : ajouter d'abord à la troisième ligne de cette matrice un multiple convenable de la deuxième ligne.]

(IV.10) Calculer les déterminants des matrices

$$\begin{pmatrix} 3 & 1 & 2 \\ 4 & 5 & 1 \\ -1 & 2 & -3 \end{pmatrix} \qquad \text{et} \qquad \begin{pmatrix} 1 & 1 & -2 & 4 \\ 0 & 1 & 1 & 3 \\ 2 & -1 & 1 & 0 \\ 3 & 1 & 2 & 5 \end{pmatrix}.$$

(IV.11) Calculer les déterminants des matrices suivantes :

$$a = \begin{pmatrix} 30 & 31 & 32 \\ 33 & 34 & 35 \\ 36 & 37 & 38 \end{pmatrix}, \quad b = \begin{pmatrix} 53 & 59 & 61 \\ 67 & 71 & 73 \\ 79 & 83 & 89 \end{pmatrix}, \quad \text{et} \quad c = \begin{pmatrix} 5 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -4 \\ 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & -2 & 0 & 0 & 0 \end{pmatrix}.$$

(IV.12) Soit $a \in M_3(\mathbb{R})$ une matrice telle que $\det(a) = 3$. Calculer $\det(2a)$.

(IV.13) Calculer le déterminant de la matrice

$$\begin{pmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ \frac{1}{2} & \frac{1}{3} & 0 \\ \frac{1}{4} & \frac{1}{5} & \frac{1}{6} \end{pmatrix}.$$

(IV.14) Soient $z, w \in \mathbb{C}$ et $a \in GL_n(\mathbb{C})$. Vérifier que

$$\det\left(zI_n + wa^{-1}\right) = \frac{\det(wI_n + za)}{\det(a)}.$$

(IV.15) Pour tout entier $n \geq 1$, soit $b_n = (b_n(i,j))_{1 \leq i,j \leq n}$ la matrice définie par

$$b_n(i,j) = \begin{cases} 3 & \text{si} \quad j = i, \\ 2 & \text{si} \quad j = i+1, \\ 1 & \text{si} \quad j = i-1, \\ 0 & \text{dans les autres cas ;} \end{cases} \qquad \text{par exemple :} \quad b_4 = \begin{pmatrix} 3 & 2 & 0 & 0 \\ 1 & 3 & 2 & 0 \\ 0 & 1 & 3 & 2 \\ 0 & 0 & 1 & 3 \end{pmatrix}.$$

(i) Calculer $\det(b_n)$ pour $n \leq 2$.

(ii) Etablir une relation linéaire entre $\det(b_n)$, $\det(b_{n-1})$ et $\det(b_{n-2})$ pour $n \geq 3$.

(iii) Calculer $\det(b_n)$ pour $n = 3, 4, 5$.

(IV.16) Calculer le déterminant de la matrice

$$\begin{pmatrix} 101 & 102 & 103 & 104 & 105 \\ 732 & 85 & 84 & 985 & 373 \\ 176 & 177 & 178 & 179 & 180 \\ 359 & 360 & 361 & 362 & 363 \\ 423 & 975 & -58 & 199 & 199 \end{pmatrix}$$

et décider si le déterminant de

$$\begin{pmatrix} 97 & 81 & 101 \\ 1993 & 1200 & 1999 \\ 2003 & 15 & 2017 \end{pmatrix}$$

est divisible par 3. [Remarque : $\det(a) \in \mathbb{Z}$ si $a \in M_n(\mathbb{Z})$.]

(IV.17) **Exercice sur le déterminant de Vandermonde.** Il s'agit du déterminant

$$V_n(x_1, \ldots, x_n) = \det \begin{pmatrix} 1 & x_1 & x_1^2 & \ldots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \ldots & x_2^{n-1} \\ 1 & x_3 & x_3^2 & \ldots & x_3^{n-1} \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 1 & x_n & x_n^2 & \ldots & x_n^{n-1} \end{pmatrix}.$$

(i) Montrer que $V_n(x_1, \ldots, x_n) = \left(\prod_{j=2}^{n} (x_j - x_1) \right) V_{n-1}(x_2, \ldots, x_n)$.

[*Indication* : soustraire à la dernière colonne x_1 fois l'avant-dernière, puis à l'avant-dernière x_1 fois l'antépénultième, et ainsi de suite $n-1$ fois ; développer alors selon la première ligne.]

(ii) Montrer que

$$V_n(x_1, \ldots, x_n) = \prod_{1 \leq i < j \leq n} (x_j - x_i).$$

(iii) Soit $\mathcal{C}$ la courbe de $\mathbb{R}^3$ image de l'application $\mathbb{R} \ni t \longmapsto (t, t^2, t^3) \in \mathbb{R}^3$. Déduire de ce qui précède que trois points de $\mathcal{C}$ distincts deux à deux et distincts de l'origine sont linéairement indépendants.

(iv) Enoncer un analogue de (iii) pour $\mathbb{R}^n$.

CHAPITRE V

Vecteurs propres et valeurs propres

Prologue

Un *anneau avec unité* est un ensemble A muni de deux opérations, une *addition* et une *multiplication,* satisfaisant les propriétés suivantes :

(i) avec l'addition, A est un groupe commutatif (et on note toujours 0 l'élément neutre correspondant) ;

(ii) la multiplication est associative ($(ab)c = a(bc)\ \forall\ a, b, c \in A$), et possède un élément neutre noté 1 ($1a = a1 = a\ \forall\ a \in A$) ;

(iii) de plus $(a + b)c = ac + bc$ et $a(b + c) = ab + ac\ \forall\ a, b, c \in A$ (distributivité).

L'anneau est *commutatif* si de plus $ab = ba$ pour tous $a, b \in A$.

Par exemple, pour les opérations usuelles, $\mathbb{Z}$, $\mathbb{Q}$ et $\mathbb{R}$ sont des anneaux commutatifs. Pour tout entier $n \geq 1$, l'ensemble $M_n(\mathbb{R})$ muni de l'addition et de la multiplication usuelles est un anneau, qui n'est pas commutatif dès que $n \geq 2$.

Un *corps commutatif* est un anneau commutatif avec unité $\mathbb{K}$ tel que $0 \neq 1$ et tel que l'ensemble $\mathbb{K}^*$ soit un groupe pour la multiplication. En d'autres termes, un anneau commutatif $\mathbb{K}$ tel que $0 \neq 1$ est un corps si tout élément non nul z de $\mathbb{K}$ possède un inverse multiplicatif.

Tout corps apparaissant dans ce cours est commutatif ; nous dirons donc désormais « corps » au lieu de « corps commutatif » .

Jusqu'ici, le corps $\mathbb{R}$ des nombres réels a été utilisé dans un souci pédagogique de familiarité, mais sans raison logique contraignante. En fait, le début de ce cours vaut pour un espace vectoriel sur un corps $\mathbb{K}$ et une application linéaire relativement à $\mathbb{K}$. (Les seules et très rares exceptions concernent quelques exemples impliquant les notions d'angle et de longueur dans les espaces euclidiens, le plus souvent dans $\mathbb{R}^2$, et plus précisément des rotations, projections orthogonales et symétries.)

Plus bas, par souci de simplicité, nous allons devoir utiliser des *nombres complexes* plutôt que des nombres réels. En guise d'introduction[1] au corps $\mathbb{C}$ des nombres complexes, nous nous limiterons à rappeler les points suivants.

(i) L'ensemble $\mathbb{C}$ des nombres complexes est en bijection avec le plan $\mathbb{R}^2$. On écrit $x + iy$ l'élément de $\mathbb{C}$ qui correspond au point de coordonnées x et y dans $\mathbb{R}^2$.

[1] Voir aussi, par exemple, l'appendice au livre de S. Lang *Linear algebra* (Springer 1971, troisième édition 1987), déjà cité plusieurs fois.

(ii) L'addition de deux nombres complexes est définie par

$$(x+iy)+(x'+iy') \;=\; (x+x')+i(y+y').$$

Un nombre réel x est identifié au nombre complexe $x+i0$.

(iii) La multiplication de deux nombres complexes est définie par

$$(x+iy)(x'+iy') \;=\; (xx'-yy')+i(xy'+yx').$$

En particulier $i^2=-1$.

On vérifie que ces opérations font de $\mathbb{C}$ *un anneau commutatif.*

(iv) Le *conjugé* d'un nombre complexe $x+iy$ est défini comme étant le nombre complexe

$$\overline{x+iy} \;=\; x-iy.$$

On vérifie que $\overline{(x+iy)}(x+iy)=x^2+y^2$, donc en particulier que $\overline{(x+iy)}(x+iy)\neq 0$ si $(x,y)\neq(0,0)$.

(v) Un calcul direct montre que, pour tout $x+iy\in\mathbb{C}$ tel que $x+iy\neq 0$, le nombre $\frac{x}{x^2+y^2}-i\frac{y}{x^2+y^2}$ est l'inverse de $x+iy$. Ecrit différemment :

$$\text{si}\quad z\neq 0,\quad \text{alors}\quad z^{-1} \;=\; \frac{1}{\overline{z}z}\overline{z}.$$

Il résulte de (v) *que l'anneau* $\mathbb{C}$ **est un corps**.

(vi) Pour tout nombre réel θ, on peut écrire

$$e^{i\theta} \;=\; \cos\theta + i\sin\theta.$$

(Plus généralement, pour tout nombre complexe z, le nombre complexe e^z est défini par la série absolument convergente $\sum_{n=0}^{\infty} z^n/n!$, et on obtient la formule ci-dessus pour $z=i\theta$, c'est-à-dire pour z «imaginaire pur» . On peut aussi prendre la formule $e^{i\theta}=\cos\theta+i\sin\theta$ comme une définition de $e^{i\theta}$.) Un nombre complexe z s'écrit de deux manières :

$$z \;=\; x+iy \;=\; re^{i\theta},$$

où $x,y,r,\theta\in\mathbb{R}$, avec $r\geq 0$ et θ défini à 2π près, sont tels que

$$r^2 \;=\; x^2+y^2,\quad x \;=\; r\cos\theta,\quad \text{et}\quad y \;=\; r\sin\theta.$$

Le nombre r est le *module* de z et θ est son *argument.* Pour deux nombres complexes écrits en termes de leurs modules et arguments, la multiplication prend la forme

$$\left(re^{i\theta}\right)\left(r'e^{i\theta'}\right) \;=\; (rr')e^{i(\theta+\theta')}.$$

De plus $\overline{re^{i\theta}}=re^{-i\theta}$.

(vii) Soit $p(t)=a_nt^n+\cdots+a_1t+a_0$ un polynôme à coefficients dans un corps $\mathbb{K}$. On dit que $\lambda\in\mathbb{K}$ est une *racine* de $p(t)$ si $p(\lambda)=0$.

Dans ce cas, soit k le plus grand entier tel que $(t-\lambda)^k$ soit un facteur de $p(t)$; il existe donc un polynôme $q(t)$ de degré $n-k$ tel que $p(t)=(t-\lambda)^kq(t)$, et $q(\lambda)\neq 0$ [sinon $q(t)$ serait de la forme $(t-\lambda)r(t)$, contrairement à la définition de k]. L'entier k est la *multiplicité* de la racine λ de $p(t)$. Par exemple, si $p(t)=t^4-2t^3+2t-1$, alors 1 est une racine de multiplicité 3 et -1 une racine de multiplicité 1, car $p(t)=(t-1)^3(t+1)$.

Lorsque $\mathbb{K} = \mathbb{R}$, il est équivalent de définir la multiplicité d'une racine λ de $p(t)$ comme le plus grand entier k tel que

$$p(\lambda) = p'(\lambda) = p^{(2)}(\lambda) = \ldots = p^{(k-1)}(\lambda) = 0$$

où $p^{(j)}(t)$ désigne la j-ième dérivée de $p(t)$.

(viii) Tout polynôme à coefficients complexes non constant possède au moins une racine (on dit que $\mathbb{C}$ est «algébriquement clos»). C'est là l'énoncé du *théorème fondamental de l'algèbre*, que nous ne démontrons pas dans ce cours.

Dans le chapitre ci-desous, la conséquence la plus importante de ce fait est que tout endomorphisme d'un espace vectoriel *complexe* de dimension finie possède au moins une valeur propre. C'est ce qui fait que l'«algèbre linéaire complexe» (c'est-à-dire l'algèbre linéaire des espaces vectoriels complexes et des applications linéaires entre ces espaces) est PLUS SIMPLE que l'«algèbre linéaire réelle» .

(ix) On peut préciser l'énoncé précédent comme suit : pour un polynôme non nul p à coefficients complexes, le nombre des racines *comptées avec leurs multiplicités* est égal au degré de p. Nous verrons au § V.3 la conséquence suivante en algèbre linéaire : pour un endomorphisme α d'un espace vectoriel complexe V,

$$\sum_{\lambda \in Sp_{\mathbb{C}}(\alpha)} m_{\text{alg}}(\lambda, \alpha) = \dim_{\mathbb{C}}(V)$$

(les notations $Sp_{\mathbb{C}}(\alpha)$ et $m_{\text{alg}}(\lambda, \alpha)$ seront définies en temps voulu).

1. Sous-espaces invariants, vecteurs propres, valeurs propres

Le style de plusieurs passages ci-dessous est résolument plus concis que celui des chapitres précédents. Si nécessaire, prendre la suite comme des indications à compléter par votre *lecture* d'un des très nombreux livres existant sur le sujet.

Soit α un endomorphisme d'un espace vectoriel V de dimension finie sur un corps $\mathbb{K}$. Un sous-espace U de V est *invariant par* α si $\alpha(U) \subset U$.

Exemples. (i) $\{0\}$, V, $\text{Ker}(\alpha)$ et $\text{Im}(\alpha)$ sont des espaces invariants par $\alpha \in \mathcal{L}(V)$. De même, pour tout $\lambda \in \mathbb{K}$, les sous-espaces $\text{Ker}(\lambda \,\text{id}_V - \alpha)$ et $\text{Im}(\lambda \,\text{id}_V - \alpha)$ sont invariants par α.

(ii) Si $\alpha \in \mathcal{L}(\mathcal{P}_d)$ désigne la dérivation, où $\mathcal{P}_d$ désigne l'espace des fonctions polynomiales de degré au plus d, alors $\mathcal{P}_j$ est invariant par α pour tout $j \in \{0, 1, \ldots, d\}$.

Traduction matricielle. Soit $\mathcal{B} = (b_1, \ldots, b_n)$ une base de V telle que $(b_1, \ldots, b_p)$ soit une base de U. Alors U est invariant par α si et seulement si $M_{\mathcal{B}}(\alpha)$ est de la forme $\begin{pmatrix} x & y \\ 0 & z \end{pmatrix}$, avec $x \in M_p(\mathbb{K})$, $y \in M_{p,n-p}(\mathbb{K})$ et $z \in M_{n-p}(\mathbb{K})$; c'est une conséquence immédiate de la définition de la matrice $M_{\mathcal{B}}(\alpha)$.

De même, pour que chacun des sous-espaces $\langle b_1, \ldots, b_j \rangle$ de V soit invariant par α $(1 \leq j \leq n-1)$, il faut et il suffit que la matrice $M_{\mathcal{B}}(\alpha)$ soit triangulaire supérieure.

Un endomorphisme $\alpha \in \mathcal{L}(V)$ est *triangulable*[2] [respectivement *diagonalisable*] s'il existe une base $\mathcal{B}$ de V telle que la matrice $M_{\mathcal{B}}(\alpha)$ correspondante soit triangulaire [resp. diagonale].

Cas particulier important : le sous-espace U de V invariant par α est de dimension 1, de sorte qu'il existe un vecteur non nul u tel que $U = \mathbb{K}u$; la condition d'invariance s'écrit alors : il existe un nombre $\lambda \in \mathbb{K}$ tel que $\alpha(u) = \lambda u$. Formalisons comme suit.

Définition. Soient V un espace vectoriel de dimension finie sur un corps $\mathbb{K}$ et α un endomorphisme de V.

Une *valeur propre* de α dans $\mathbb{K}$ est un nombre $\lambda \in \mathbb{K}$ tel qu'il existe un vecteur *non nul* $u \in V$ pour lequel $\alpha(u) = \lambda u$. Un tel u est un *vecteur propre* de α de valeur propre λ ; l'ensemble de tous ces vecteurs propres et le vecteur nul constituent *l'espace propre*

$$\{u \in V \mid \alpha(u) = \lambda u\} \;=\; \mathrm{Ker}(\lambda\,\mathrm{id}_V - \alpha)$$

correspondant. On appelle *spectre de α dans $\mathbb{K}$* et on note $Sp_{\mathbb{K}}(\alpha)$ l'ensemble des valeurs propres de α dans $\mathbb{K}$.

Définitions analogues pour une *matrice* $a \in M_n(\mathbb{K})$.

Exemples. (i) Supposons que V est un espace euclidien de dimension 2. Une *projection orthogonale* de V sur une droite d de V a 1 et 0 comme valeurs propres ; l'espace propre correspondant à la valeur propre 1 est la droite d, et l'espace propre correspondant à la valeur propre 0 est la droite orthogonale. La *symétrie orthogonale* de V relativement à d a 1 et -1 comme valeurs propres ; les espaces propres sont à nouveau d et $d^{\perp}$. Une rotation d'un angle θ non multiple de π n'a *aucune valeur propre réelle.*

(ii) Le nombre 0 est valeur propre de α si et seulement si α n'est pas inversible. Lorsque 0 est valeur propre de α, l'espace propre correspondant est le noyau de α.

(iii) Si $\alpha = \lambda\,\mathrm{id}_V$ pour un nombre $\lambda \in \mathbb{K}$, alors $Sp_{\mathbb{K}}(\alpha) = \{\lambda\}$ et *tout* vecteur non nul de V est vecteur propre.

(iv) Soit $a = \mathrm{diag}(\lambda_1, \ldots, \lambda_n) \in M_n(\mathbb{K})$ une matrice diagonale. Alors le j-ème vecteur de la base canonique de $\mathbb{K}^n$ est un vecteur propre de a de valeur propre λ_j $(1 \leq j \leq n)$.

(v) Soit $a = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$ une matrice de rotation plane ; on suppose que θ n'est pas un multiple entier de π. Alors $Sp_{\mathbb{R}}(a) = \emptyset$ comme déjà noté en (i). En revanche, $Sp_{\mathbb{C}}(a)$ contient $\exp(i\theta)$ et $\exp(-i\theta)$, comme on le voit en appliquant a aux vecteurs $\begin{pmatrix} 1 \\ -i \end{pmatrix}$ et $\begin{pmatrix} 1 \\ i \end{pmatrix}$ de $\mathbb{C}^2$.

Proposition. (i) *Toute matrice triangulaire dans $M_n(\mathbb{K})$ possède au moins une valeur propre dans $\mathbb{K}$.*

(ii) *Deux matrices semblables*[3] *ont mêmes valeurs propres.*

[2] On dit aussi «triangularisable» ou «trigonalisable» .

[3] Rappel : cela veut dire deux matrices $a, b \in M_n(\mathbb{K})$ pour lesquelles il existe une matrice $s \in GL(n, \mathbb{K})$ telle que $b = sas^{-1}$.

PREUVE. (i) Pour une matrice triangulaire inférieure $\begin{pmatrix} a_{1,1} & 0 & \dots \\ a_{1,2} & a_{2,2} & \dots \\ \dots & \dots & \dots \end{pmatrix}$, le dernier vecteur de la base canonique est un vecteur propre. Pour une matrice triangulaire supérieure $\begin{pmatrix} a_{1,1} & a_{1,2} & \dots \\ 0 & a_{2,2} & \dots \\ \dots & \dots & \dots \end{pmatrix}$, le premier vecteur de la base canonique est un vecteur propre.

(ii) Soient a, b, s des matrices de $M_n(\mathbb{K})$, avec s inversible, telles que $b = sas^{-1}$. Supposons que a possède un vecteur propre u de valeur propre λ, c'est-à-dire que $au = \lambda u$. Alors $(sas^{-1})(su) = \lambda su$, de sorte que su est un vecteur propre de b de valeur propre λ. □

THÉORÈME V.1. *Soient V un espace vectoriel de dimension finie sur un corps $\mathbb{K}$ et α un endomorphisme de V. Soient $\lambda_1, \dots, \lambda_k \in \mathbb{K}$ des valeurs propres distinctes de α et $v_1, \dots, v_k \in V$ des vecteurs propres correspondants. Alors les vecteurs $v_1, \dots, v_k$ sont linéairement indépendants.*

PREUVE(PAR RÉCURRENCE SUR k). Le théorème est évident si $k = 1$. On suppose désormais $k \geq 2$ et le théorème démontré jusqu'à $k - 1$.

Soient $\mu_1, \dots, \mu_k \in \mathbb{K}$ tels que

$$(*) \qquad \mu_1 v_1 + \dots + \mu_k v_k = 0.$$

En appliquant α, on obtient

$$(**) \qquad \mu_1 \alpha(v_1) + \dots + \mu_k \alpha(v_k) = \mu_1 \lambda_1 v_1 + \dots + \mu_k \lambda_k v_k = 0.$$

En soustrayant λ_k fois $(*)$ à $(**)$, on obtient

$$\mu_1(\lambda_1 - \lambda_k)v_1 + \dots + \mu_{k-1}(\lambda_{k-1} - \lambda_k)v_{k-1} = 0.$$

L'hyphothèse de récurrence implique alors que

$$\mu_1 = \dots = \mu_{k-1} = 0.$$

Par suite, $(*)$ s'écrit $\mu_k v_k = 0$; comme $v_k \neq 0$, on a aussi $\mu_k = 0$. □

COROLLAIRE. *Si V est de dimension n, un endomorphisme linéaire α de V possède au plus n valeurs propres.*

COROLLAIRE. *Avec les mêmes notations, l'application*

$$\begin{array}{ccc} \operatorname{Ker}(\lambda_1 \operatorname{id}_V - \alpha) \oplus \dots \oplus \operatorname{Ker}(\lambda_k \operatorname{id}_V - \alpha) & \longrightarrow & V \\ (v_1, \dots, v_k) & \longmapsto & v_1 + \dots + v_k \end{array}$$

est injective.

Il résulte du premier corollaire que, dans les exemples de (i) ci-dessus, il n'y a pas de valeur propre autre que les valeurs indiquées. De même dans l'exemple de (iv) lorsque les λ_j sont distincts deux à deux (par le premier corollaire), ou même en général (par le second corollaire).

2. Polynôme caractéristique

Pour V et $\alpha \in \mathcal{L}(V)$ comme au début du § V.1, on définit le *polynôme caractéristique* de α comme étant

$$p_\alpha(t) \;=\; \det\left(t\,\mathrm{id}_V - \alpha\right).$$

C'est un polynôme à coefficients dans $\mathbb{K}$ dont le degré est la dimension de V, dont le coefficient dominant est 1 et dont le terme constant est $(-1)^{\dim(V)}\det(\alpha)$. On définit de même le *polynôme caractéristique* d'une matrice $a \in M_n(\mathbb{K})$.

NOTATION. On écrit souvent $t - \alpha$ au lieu de $t\,\mathrm{id}_V - \alpha$.

Par exemple, si

$$a \;=\; \begin{pmatrix} 2 & 1 & 0 \\ 0 & 1 & -1 \\ 0 & 2 & 4 \end{pmatrix},$$

alors

$$p_a(t) \;=\; \det\begin{pmatrix} t-2 & -1 & 0 \\ 0 & t-1 & 1 \\ 0 & -2 & t-4 \end{pmatrix} \;=\; (t-2)\Big((t-1)(t-4)+2\Big) \;=\; (t-2)^2(t-3).$$

Si

$$r_\theta \;=\; \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

est la matrice d'une rotation de $\mathbb{R}^2$ d'angle θ, alors

$$p_{r_\theta}(t) \;=\; t^2 - 2(\cos\theta)t + 1.$$

THÉORÈME V.2. *Pour $\alpha \in \mathcal{L}(V)$ comme ci-dessus, les valeurs propres de α dans $\mathbb{K}$ sont exactement les racines dans $\mathbb{K}$ du polynôme caractéristique p_α.*

PREUVE. Cela résulte du fait qu'un endomorphisme β de V est non inversible (ou encore non injectif) si et seulement si son déterminant est nul.

Plus précisément, soit $\lambda \in \mathbb{K}$. D'une part, l'endomorphisme $\lambda - \alpha$ de V est non inversible si et seulement s'il existe un vecteur propre v de α de valeur propre λ. D'autre part, l'endomorphisme $\lambda - \alpha$ de V est non inversible si et seulement si $\det(\lambda - \alpha) = 0$, c'est-à-dire si et seulement si λ est une racine du p_α. □

EXEMPLE. Pour chacune des matrices a et r_θ introduites juste avant le théorème, vérifions que les racines du polynôme caractéristique coincident avec les valeurs propres.

Le noyau de $2 - a$ contient le premier vecteur de la base canonique :

$$(2-a)\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \;=\; \begin{pmatrix} 0 & -1 & 0 \\ 0 & 1 & 1 \\ 0 & -2 & -2 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \;=\; \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

de sorte que 2 est bien une valeur propre de a. De même, le noyau de $3 - a$ n'est pas réduit à zéro :

$$(3-a)\begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix} = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 2 & 1 \\ 0 & -2 & -1 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

et 3 est une valeur propre de a. Plus précisément, l'étude des systèmes linéaires

$$\begin{pmatrix} 0 & -1 & 0 \\ 0 & 1 & 1 \\ 0 & -2 & -2 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \qquad \text{et} \qquad \begin{pmatrix} 1 & -1 & 0 \\ 0 & 2 & 1 \\ 0 & -2 & -1 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

montre que l'espace propre de a correspondant à la valeur propre 2 [respectivement 3] est de dimension 1, engendré par $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ $\left[\text{resp. } \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix}\right]$.

Si θ n'est pas un multiple entier de π, on sait que r_θ ne possède aucune valeur propre réelle et on vérifie que le polynôme p_{r_θ} ne possède aucune racine réelle :

$$p_{r_\theta}(t) = t^2 - 2(\cos\theta)t + 1 = (t - \cos\theta)^2 + (\sin\theta)^2 > 0 \qquad \text{pour tout} \quad t \in \mathbb{R}.$$

Propriétés des polynômes caractéristiques. (i) Soient $\alpha, \sigma \in \mathcal{L}(V)$ avec σ inversible. Alors α et $\sigma\alpha\sigma^{-1}$ ont même polynôme caractéristique. [Rappel : deux matrices semblables ont même déterminant.]

(ii) Si $a = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$, alors $p_a(t) = \prod_{j=1}^n (t - \lambda_j)$.

(iii) De même, si a est une matrice triangulaire de coefficients diagonaux $\lambda_1, \ldots, \lambda_n$, alors $p_a(t) = \prod_{j=1}^n (t - \lambda_j)$. Par suite,

les valeurs propres d'une matrice triangulaire sont ses coefficients diagonaux.

En particulier, si $\alpha \in \mathcal{L}(V)$ est un endomorphisme triangulable, alors les coefficients diagonaux de deux matrices triangulaires qui représentent α (relativement à deux bases de V) sont égaux, à l'ordre près.

(iv) Une matrice carrée et sa transposée ont même polynôme caractéristique.

(v) Soit $a = \begin{pmatrix} x & y \\ 0 & z \end{pmatrix}$ une matrice triangulaire par blocs, avec x, z des matrices carrées. Alors $p_a(t) = p_x(t)p_z(t)$. [Rappel : le déterminant d'une matrice triangulaire par blocs et le produit des déterminants des blocs diagonaux.]

Soit α un endomorphisme d'un espace vectoriel V de dimension finie sur un corps $\mathbb{K}$ et soit λ une valeur propre de α. La *multiplicté géométrique* $m_{\text{geo}}(\lambda, \alpha)$ de λ est la dimension de l'espace propre $\operatorname{Ker}(\lambda \operatorname{id}_V - \alpha)$. La *multiplicté algébrique* $m_{\text{alg}}(\lambda, \alpha)$ de λ est la multiplicité de la racine λ du polynôme caractéristique p_α. S'il n'y a pas de confusion possible, on écrit $m_{\text{geo}}(\lambda)$ et $m_{\text{alg}}(\lambda)$ au lieu de $m_{\text{geo}}(\lambda, \alpha)$ et $m_{\text{alg}}(\lambda, \alpha)$.

Les définitions pour une matrice $a \in M_n(\mathbb{K})$ sont analogues.

Pour un endomorphisme $\alpha \in \mathcal{L}(V)$ [ou une matrice $a \in M_n(\mathbb{K})$] et un nombre $\lambda \in \mathbb{K}$, on écrit parfois $m_{\text{geo}}(\lambda) = 0$ et $m_{\text{alg}}(\lambda) = 0$ pour signifier que λ *n'est pas* une valeur propre de α [ou de a].

Par exemple, pour la matrice $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, la valeur propre 0 a multiplicité géométrique 1, l'espace propre correspondant étant engendré par le premier vecteur de la base canonique de $\mathbb{K}^2$, et la multiplicité algébrique est 2, le polynôme caractéristique étant t^2. Pour la matrice $a = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 1 & -1 \\ 0 & 2 & 4 \end{pmatrix}$, les calculs déjà indiqués montrent que

$$m_{\text{geo}}(2, a) = 1 < 2 = m_{\text{alg}}(2, a).$$

A titre d'exercice, le lecteur vérifiera que, pour une matrice diagonale a de coefficients diagonaux $\lambda_1, \ldots, \lambda_n$, l'égalité $m_{\text{geo}}(\lambda_j, a) = m_{\text{alg}}(\lambda_j, a)$ est vraie pour tout $j \in \{1, \ldots, n\}$.

PROPOSITION. *Avec les notations ci-dessus, on a toujours* $m_{\text{geo}}(\lambda, \alpha) \leq m_{\text{alg}}(\lambda, \alpha)$.

PREUVE. Notons n la dimension de V ; posons $k = m_{\text{geo}}(\lambda, \alpha)$ et $l = n - k$. Soit $\mathcal{B} = (b_1, \ldots, b_n)$ une base de V telle que $(b_1, \ldots, b_k)$ soit une base de l'espace propre $\operatorname{Ker}(\lambda \operatorname{id}_V - \alpha)$. La matrice représentative de $\lambda \operatorname{id}_V - \alpha$ relativement à la base $\mathcal{B}$ est triangulaire par blocs de la forme

$$M_{\mathcal{B}}(\lambda \operatorname{id}_V - \alpha) = \begin{pmatrix} 0 & * \\ 0 & * \end{pmatrix}$$

de sorte qu'on peut écrire

$$M_{\mathcal{B}}(\alpha) = \begin{pmatrix} \lambda I_k & b \\ 0 & d \end{pmatrix}$$

où $b \in M_{k,l}(\mathbb{K})$ et $d \in M_l(\mathbb{K})$. Le polynôme caractéristique de α s'écrit donc

$$p_\alpha(t) = \det \begin{pmatrix} (t - \lambda) I_k & -b \\ 0 & t I_l - d \end{pmatrix} = (t - \lambda)^k p_d(t),$$

ce qui montre que λ est une racine de $p_\alpha(t)$ de multiplicité au moins k, en d'autres termes que $m_{\text{alg}}(\lambda, \alpha) \geq k$. □

CONSÉQUENCE DES DÉFINITIONS ET DE LA PROPOSITION.

$$1 \leq m_{\text{geo}}(\lambda, \alpha) \leq m_{\text{alg}}(\lambda, \alpha) \leq \dim(V)$$

pour toute valeur propre λ de α.

EXEMPLE. On appelle *bloc de Jordan* une matrice de la forme

$$J_n(\lambda) = \begin{pmatrix} \lambda & 1 & 0 & 0 & \ldots & 0 & 0 \\ 0 & \lambda & 1 & 0 & \ldots & 0 & 0 \\ 0 & 0 & \lambda & 1 & \ldots & 0 & 0 \\ 0 & 0 & 0 & \lambda & \ldots & 0 & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & 0 & \cdots & \lambda & 1 \\ 0 & 0 & 0 & 0 & \cdots & 0 & \lambda \end{pmatrix} \in M_n(\mathbb{K}),$$

où $\lambda \in \mathbb{K}$. Cette matrice étant triangulaire, il est immédiat que son polynôme caractérisitque est $(t - \lambda)^n$, donc qu'elle a une et une seule valeur propre, à savoir λ, de multiplicité algébrique $m_{\text{alg}}(\lambda) = n$.

Pour cette matrice et un vecteur $x \in \mathbb{K}^n$, l'équation $J_n(\lambda)x = \lambda x$ s'écrit comme le système linéaire

$$\begin{cases} \lambda x_1 + x_2 = \lambda x_1 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \lambda x_{n-1} + x_n = \lambda x_{n-1} \\ \lambda x_n = \lambda x_n \end{cases}.$$

Ses solutions sont : x_1 arbitraire et $x_2 = \cdots = x_n = 0$. Par suite λ est une valeur propre de $J_n(\lambda)$ de multiplicité géométrique $m_{\text{geo}}(\lambda) = 1$.

Exercice. Pour tout $\lambda \in \mathbb{K}$ et pour tous $k, l \in \{1, \ldots, n\}$ avec $1 \leq k \leq l \leq n$, écrire une matrice $a \in M_n(\mathbb{K})$ telle que $m_{\text{geo}}(\lambda) = k$ et $m_{\text{alg}}(\lambda) = l$.

[*Indication* : Pour une matrice $a = \begin{pmatrix} u & 0 \\ 0 & v \end{pmatrix}$ diagonale par blocs et un nombre $\lambda \in \mathbb{K}$, on vérifie que $m_{\text{geo}}(\lambda, a) = m_{\text{geo}}(\lambda, u) + m_{\text{geo}}(\lambda, v)$ et $m_{\text{alg}}(\lambda, a) = m_{\text{alg}}(\lambda, u) + m_{\text{alg}}(\lambda, v)$.]

Remarque. Soient a une matrice n-fois-n à coefficients réels possédant une valeur propre $\lambda \in \mathbb{R}$. D'une part, on peut bien sûr considérer a comme un élément de $M_n(\mathbb{R})$, donc l'espace propre correspondant $\text{Ker}(\lambda I_n - a)$ comme un sous-espace vectoriel de $\mathbb{R}^n$. La multiplicité géométrique $m = m_{\text{geo}}(\lambda, a)$ de λ est donc donnée par

$$m = \dim_{\mathbb{R}} (\text{Ker}(\lambda I_n - a)).$$

D'autre part, on peut aussi considérer a comme un élément de $M_n(\mathbb{C})$, donc l'espace propre correspondant $\text{Ker}(\lambda I_n - a)$ comme un sous-espace vectoriel de $\mathbb{C}^n$. La multiplicité géométrique vue ainsi est donc

$$m' = \dim_{\mathbb{C}} (\text{Ker}(\lambda I_n - a)).$$

En fait

$$m = m',$$

de sorte que l'entier $m_{\text{geo}}(\lambda, a)$ est défini sans ambiguïté. Voici une manière d'argumenter.

Soit n un entier, $n \geq 1$. Pour tout vecteur $z = \begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix} \in \mathbb{C}^n$, définissons le *vecteur conjugué* $\overline{z} = \begin{pmatrix} \overline{z_1} \\ \vdots \\ \overline{z_n} \end{pmatrix} \in \mathbb{C}^n$. Noter que $\overline{\overline{z}} = z$. Si on pose $x = \frac{1}{2}(z + \overline{z})$ et $y = \frac{1}{2i}(z - \overline{z})$, alors $z = x + iy$ et $x, y \in \mathbb{R}^n$; en particulier, $z \in \mathbb{R}^n$ si et seulement si $\overline{z} = z$.

Affirmation. *Soit E un sous-espace vectoriel* complexe *de $\mathbb{C}^n$ tel que*

$$(*) \qquad \overline{z} \in E \quad \textit{pour tout} \quad z \in E \ ;$$

notons k la dimension de E (dimension complexe *!). Il existe une base $(b_1, \ldots, b_k)$ de E telle que $\overline{b_j} = b_j$ pour $j = 1, \ldots, k$.*

Preuve. Soit $(z_1, \ldots, z_k)$ une base de E. Pour $j = 1, \ldots, k$, posons

$$x_j = \frac{1}{2}(z_j + \overline{z_j}) \qquad \text{et} \qquad y_j = \frac{1}{2i}(z_j - \overline{z_j}).$$

D'une part, chacun des vecteurs $x_1, \ldots, x_k, y_1, \ldots, y_k$ est égal à son propre vecteur conjugé. D'autre part, ces vecteurs engendrent E, car $z_j = x_j + iy_j$ $(1 \leq j \leq k)$. Il résulte du lemme d'échange de Grassmann que, parmi ces $2k$ vecteurs, on peut en choisir k qui constituent une base de E. □

Cas particulier. Soit $a \in M_n(\mathbb{R})$ une matrice n-fois-n à coefficients réels, qu'on peut aussi voir comme étant à coefficients complexes. Considérons les applications linéaires $\alpha : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ et $\alpha' : \mathbb{C}^n \longrightarrow \mathbb{C}^n$ définies par a. Le sous-espace $\mathrm{Ker}(\alpha')$ de $\mathbb{C}^n$ satisfait la condition $(*)$ ci-dessus (à vérifier !). Il en résulte que la dimension complexe du sous-espace vectoriel complexe $\mathrm{Ker}(\alpha')$ de $\mathbb{C}^n$ est égale à la dimension réelle du sous-espace vectoriel réel $\mathrm{Ker}(\alpha)$ de $\mathbb{R}^n$. De même, le rang de l'application $\mathbb{C}$-linéaire α' est égal au rang de l'application $\mathbb{R}$-linéaire α (comme on s'en convainc par exemple en invoquant la formule de la dimension). De même encore, pour toute valeur propre λ *réelle* de a :

$$\dim_{\mathbb{C}} \Big(\mathrm{Ker}(\lambda\,\mathrm{id}_V - \alpha') \Big) \;=\; \dim_{\mathbb{R}} \Big(\mathrm{Ker}(\lambda\,\mathrm{id}_V - \alpha) \Big)$$

(comme annoncé).

Plus généralement, étant donné une valeur propre $\lambda \in \mathbb{C}$ de a et l'espace propre correspondant $E = \{z \in C^n \,|\, az = \lambda z\}$, on vérifie que $\overline{\lambda}$ est aussi une valeur propre de a, d'espace propre $\{z \in \mathbb{C}^n \,|\, \overline{z} \in E\}$. En particulier, les multiplicités géométriques $\dim_{\mathbb{C}} \big(\mathrm{Ker}(\lambda\,\mathrm{id}_V - \alpha') \big)$ de λ et $\dim_{\mathbb{C}} \big(\mathrm{Ker}(\overline{\lambda}\,\mathrm{id}_V - \alpha') \big)$ de $\overline{\lambda}$ sont égales.

3. Endomorphismes linéaires des espaces vectoriels COMPLEXES. Triangulation

Au lieu de « triangulation » , on dit aussi « trigonalisation » .

Il résulte du théorème du paragraphe précédent et du théorème fondamental de l'algèbre que *tout endomorphisme d'un espace vectoriel complexe de dimension finie possède au moins une valeur propre,* et plus précisément que, pour tout endomorphisme α d'un espace vectoriel complexe V,

$$\sum_{\lambda \in Sp_{\mathbb{C}}(\alpha)} m_{\mathrm{alg}}(\lambda, \alpha) \;=\; \dim_{\mathbb{C}}(V).$$

Répétons que c'est à cause de cela que l'algèbre linéaire complexe est PLUS SIMPLE que l'algèbre linéaire réelle.

Digression. Pour montrer qu'un endomorphisme α d'un espace vectoriel complexe V de dimension $n < \infty$ possède au moins une valeur propre, il existe un argument qui, contrairement à celui du cours, ne fait usage ni du polynôme caractéristique de α ni de la théorie des déterminants. En voici une esquisse.

On choisit arbitrairement un vecteur $v \in V$, $v \neq 0$. Soit m le plus petit entier tel que $\alpha^m(v)$ soit dans le sous-espace de V engendré par $v, \alpha(v), \alpha^2(v), \ldots, \alpha^{m-1}(v)$ et soient $c_0, c_1, c_2, \ldots, c_{m-1} \in \mathbb{C}$ tels que

$$(*) \qquad \alpha^m(v) \;=\; c_0 v + c_1 \alpha(v) + c_2 \alpha^2(v) + \ldots + c_{m-1}\alpha^{m-1}(v)$$

(noter que $m \leq n = \dim_{\mathbb{C}}(V)$). On considère le polynôme

$$p(t) \;=\; t^m - c_{m-1}t^{m-1} - \ldots - c_1 t - c_0$$

et sa factorisation $p(t) = \prod_{j=1}^{m}(t - \lambda_j)$ en termes linéaires en t. On peut alors récrire $(*)$ sous la forme

$$(**) \qquad (\alpha - \lambda_1\,\mathrm{id}_V)\,(\alpha - \lambda_2\,\mathrm{id}_V) \cdots (\alpha - \lambda_m\,\mathrm{id}_V)\,v \;=\; 0.$$

Il résulte de $(**)$ que l'un au moins des endomorphismes $\alpha - \lambda_j\,\mathrm{id}_V$ n'est pas injectif, donc que l'un au moins des nombres λ_j est une valeur propre de α.

Un argument du même type montre le résultat suivant. Soit α un endomorphisme d'un espace vectoriel réel V de dimension finie ; alors il existe un sous-espace U de V de dimension 1 ou 2 qui est invariant par α. Pour les détails, voir S. Adler, *Linear Algebra Done Right,* Second Edition (Springer, 1997), Théorème 5.24.

Théorème V.3. *Tout endomorphisme d'un espace vectoriel complexe de dimension finie est triangulable.*

En particulier, pour toute matrice $a \in M_n(\mathbb{C})$ (a fortiori pour toute matrice $a \in M_n(\mathbb{R})$), il existe une matrice inversible s dans $GL(n, \mathbb{C})$ telle que la matrice sas^{-1} est triangulaire.

PREUVE. Notons V l'espace vectoriel, n sa dimension, et procédons par récurrence sur n. Si $n = 1$, il n'y a rien à montrer. On suppose donc désormais $n > 1$, et le théorème vrai pour les espaces de dimensions strictement inférieures à n.

Soit $\alpha \in \mathcal{L}(V)$. L'endomorphisme α possède une valeur propre $\lambda \in \mathbb{C}$. Notons U le sous-espace $\mathrm{Ker}(\lambda\,\mathrm{id}_V - \alpha)$ de V ; c'est un sous-espace de V distinct de $\{0\}$ qui est invariant par α [en effet, si $u \in U$, alors $\alpha(u) = \lambda u \in U$]. Notons k la dimension de U et posons $l = n - k$.

Il existe une base $\mathcal{B} = (b_1, \ldots, b_n)$ de V telle que $(b_1, \ldots, b_k)$ soit une base de U. La matrice de α relativement à cette base est triangulaire par blocs, de la forme

$$M_{\mathcal{B}}(\alpha) \;=\; \begin{pmatrix} \lambda I_k & y \\ 0 & z \end{pmatrix}$$

(où les blocs $I_k \in M_k(\mathbb{C})$ et $0 \in M_{l,k}(\mathbb{C})$ traduisent l'égalité $(\lambda\,\mathrm{id}_V - \alpha)(U) = \{0\}$).

Comme $l < n$, l'hypothèse de récurrence s'applique à z. Il existe donc une matrice inversible $t \in GL(l, \mathbb{C})$ telle que la matrice tzt^{-1} soit triangulaire supérieure. Par suite

$$\begin{pmatrix} I_k & 0 \\ 0 & t \end{pmatrix} \begin{pmatrix} \lambda I_k & y \\ 0 & z \end{pmatrix} \begin{pmatrix} I_k & 0 \\ 0 & t \end{pmatrix}^{-1} \;=\; \begin{pmatrix} \lambda I_k & yt^{-1} \\ 0 & tzt^{-1} \end{pmatrix}$$

est aussi triangulaire supérieure, ce qui achève la preuve. □

COMPLÉMENT. *Soit α un endomorphisme d'un espace vectoriel réel dont le polynôme caractéristique a toutes ses racines réelles ; alors α est triangularisable.*

En particulier, pour toute matrice carrée $a \in M_n(\mathbb{R})$ dont le polynôme caractéristique n'a que des racines réelles, il existe une matrice inversible s dans $GL(n, \mathbb{R})$ telle que la matrice (réelle!!!) sas^{-1} est triangulaire.

PREUVE. A nouveau par récurrence sur n, parallèle à la preuve du théorème précédent. □

Attention : l'exemple $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ montre qu'il existe des matrices complexes non diagonalisables. (Il s'agit bien sûr d'une matrice réelle, mais on souhaite la voir ici comme une matrice complexe et dire qu'elle n'est pas diagonalisable *comme matrice complexe.*)

PROPOSITION. (i) *Soit $a \in M_n(\mathbb{C})$. Pour que la matrice a soit diagonalisable, il faut et il suffit que*

$$(*) \qquad m_{\mathrm{geo}}(\lambda, a) \;=\; m_{\mathrm{alg}}(\lambda, a)$$

pour toute valeur propre λ de a.

En particulier, si le polynôme caractéristique de a possède n racines distinctes, alors a est diagonalisable.

(ii) *Soit $a \in M_n(\mathbb{R})$. Pour que la matrice a soit diagonalisable sur $\mathbb{R}$ c'est-à-dire pour qu'il existe une matrice $s \in GL(n, \mathbb{R})$ telle que la matrice réelle sas^{-1} soit diagonale, il faut et il suffit que le polynôme caractéristique de a ait toutes ses racines réelles et que*

$$m_{\mathrm{geo}}(\lambda, a) \;=\; m_{\mathrm{alg}}(\lambda, a)$$

pour toute valeur propre λ de a.

En particulier, si le polynôme caractéristique de a possède n racines réelles distinctes, alors a est diagonalisable sur $\mathbb{R}$.

PREUVE. (i) Soient $\lambda_1, \ldots, \lambda_k$ les valeurs propres de a ; notons E_j l'espace propre $\mathrm{Ker}(\lambda_j I_n - a)$ correspondant à λ_j. Par le second corollaire du théorème V.1, il y a une injection naturelle de la somme directe $E_1 \oplus \cdots \oplus E_k$ dans $\mathbb{C}^n$. Sous la condition $(*)$, on a de plus

$$\sum_{j=1}^{k} \dim(E_j) \;=\; \sum_{j=1}^{k} m_{\mathrm{geo}}(\lambda_j, a) \;=\; \sum_{j=1}^{k} m_{\mathrm{alg}}(\lambda_j, a) \;=\; n,$$

de sorte que cette injection est un isomorphisme. Soient $\mathcal{B}$ une base de $\mathbb{C}^n$ qui est réunion de bases des E_j et $\mathcal{C}$ la base canonique de $\mathbb{C}^n$. Si $s = M_{\mathcal{B},\mathcal{C}}(\mathrm{id}_{\mathbb{C}^n})$, alors sas^{-1} est une matrice diagonale.

La preuve de la réciproque (a diagonalisable implique $(*)$), est laissée en exercice au lecteur.

(ii) La preuve est analogue. □

THÉORÈME V.4 (théorème de la forme normale de Jordan[4]). *Pour tout endomorphisme linéaire α d'un espace vectoriel complexe V de dimension finie, il existe une base de V relativement à laquelle la matrice de α est diagonale par blocs, avec des blocs diagonaux qui sont des blocs de Jordan.*

PREUVE. Ne fait pas partie du cours. Voir un ouvrage «plus avancé» d'algèbre linéaire. □

EXEMPLE. La matrice $\begin{pmatrix} 2 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 2 & 1 & 0 \\ 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 3 \end{pmatrix}$ est diagonale par blocs de Jordan, avec trois blocs dans $M_1(\mathbb{C}) \approx \mathbb{C}$ et un bloc dans $M_2(\mathbb{C})$.

(Le théorème reste vrai pour un espace vectoriel sur un corps $\mathbb{K}$ et un endomorphisme dont toutes les valeurs propres sont dans $\mathbb{K}$.)

4. Exercices

(V.1) Calculer les nombres complexes $(1+2i)^3$, $\frac{5}{-3+4i}$ et $\left(\frac{2+i}{3-2i}\right)^2$.

(V.2) Soit $\mathcal{B} = (b_1, \ldots, b_n)$ une base de $\mathbb{R}^n$. Lorsqu'on considère les b_j comme des vecteurs à coordonnées complexes, montrer qu'ils forment encore une base de $\mathbb{C}^n$.

[4]Camille Jordan, mathématicien français, 1838–1922. Son «Traité des substitutions et des équations algébriques» est entre autres le premier traité de théorie des groupes. On retrouve son nom en topologie plane avec les «courbes de Jordan» , et le résultat important selon lequel elles «séparent le plan» .

(V.3) L'exercice se rapporte à un *modèle du corps des nombres complexes.*

Soit $\mathcal{C}$ le sous-ensemble de $M_2(\mathbb{R})$ des matrices de la forme $z = xI + yJ$; le symbole I désigne la matrice unité I_2 à deux lignes et deux colonnes, et $J = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$. Pour une matrice z de ce type, on pose $\overline{z} = xI - yJ \in \mathcal{C}$. Vérifier les assertions (i) à (v) ci-dessous.

(i) $\mathcal{C}$ est un sous-espace vectoriel de $M_2(\mathbb{R})$ de dimension réelle deux.

(ii) Le produit $z_1 z_2$ est dans $\mathcal{C}$ pour tous $z_1, z_2 \in \mathcal{C}$.

(iii) $z_1 z_2 = z_2 z_1$ pour tous $z_1, z_2 \in \mathcal{C}$.

(iv) $z\overline{z} = (x^2 + y^2)I$ pour tout $z = xI + yJ \in \mathcal{C}$.

(v) $z = xI + yJ$ est inversible dès que $z \neq 0$.
[*Indication* : deviner l'inverse, puis vérifier.]

(V.4) Dessiner dans $\mathbb{R}^2$ les espaces propres de la matrice $\begin{pmatrix} 2 & 1 \\ 1 & -3 \end{pmatrix}$.

(V.5) Déterminer dans $\mathbb{C}^2$ les espaces propres de la matrice $\begin{pmatrix} 1 & -5 \\ 2 & -1 \end{pmatrix}$.

(V.6) Décrire TOUS les vecteurs propres des matrices $\begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}$ et $\begin{pmatrix} 2 & 3 \\ 0 & 2 \end{pmatrix}$.

(V.7) Pour tout entier $n \geq 1$, soit $a^{(n)} \in M_n(\mathbb{C})$ la matrice définie par

$$a_{i,j}^{(n)} = \begin{cases} 1 & \text{si} \quad |i-j| = 1 \\ 0 & \text{dans les autres cas} \end{cases}$$

et soit $p_n(t)$ son polynôme caractéristique.

(i) Calculer $p_1(t)$ et $p_2(t)$.

(ii) Etablir une relation linéaire entre $p_n(t)$, $p_{n-1}(t)$ et $p_{n-2}(t)$ pour tout $n \geq 3$.

(iii) Calculer $p_3(t)$ et $p_4(t)$.

(V.8) Pour $j \in \{1, 2, 3, 4\}$, écrire une matrice $a_j \in M_4(\mathbb{C})$ pour laquelle le nombre 2 est une valeur propre de multiplicité algébrique 4 et de multiplicité géométrique j.

(V.9) Soit $\theta \in \mathbb{R}$. Trouver une matrice $s \in GL(2, \mathbb{C})$ telle que la matrice

$$s \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} s^{-1}$$

soit diagonale.

(V.10) Soient n un entier et $a \in M_n(\mathbb{C})$ une matrice telle que $a^4 = I_n$. Montrer que toute valeur propre de a est l'un des nombres $1, i, -1, -i$.

(V.11) Soit $a \in M_n(\mathbb{C})$ une matrice de valeurs propres $\lambda_1, \ldots, \lambda_n$ distinctes deux à deux. Ecrire la liste des valeurs propres de a^2, de $a^3 - a + 3I_n$ et (lorsque a est inversible) de a^{-1}.

(V.12) Soient α un endomorphisme d'un espace vectoriel complexe V tel que $\alpha^2 = \mathrm{id}_V$ et $\lambda \in \mathbb{C}$ un nombre complexe tel que $\lambda \neq \pm 1$. Trouver une combinaison linéaire de id_V et α égale à $(\lambda\,\mathrm{id}_V - \alpha)^{-1}$.

(V.13) On considère un entier $n \geq 1$, un nombre complexe λ, l'espace vectoriel complexe V des fonctions $f : \mathbb{R} \longrightarrow \mathbb{C}$ de la forme $f(t) = (a_0 + a_1 t + \cdots + a_n t^n)e^{\lambda t}$, avec $a_0, \ldots, a_n \in \mathbb{C}$, la suite $\mathcal{B} = (b_0, \ldots, b_n)$ d'éléments de V définie par $b_j(t) = \frac{1}{j!}t^j e^{\lambda t}$, et l'endomorphisme $\alpha \in \mathcal{L}(V)$ défini par $\alpha(f) = f'$ (dérivée).

(i) Vérifier que $\mathcal{B}$ est une base de V.

(ii) Ecrire la matrice $M_{\mathcal{B}}(\alpha)$.

Quelles sont les valeurs propres de cette matrice ? leurs multiplicités ?

(V.14) Dans l'espace $\mathbb{R}^3$, trouver une base formée de vecteurs propres de la matrice $a = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$.

(V.15) Soit $a = \begin{pmatrix} 0 & 1 \\ \frac{1}{4} & 1 \end{pmatrix}$. Déterminer les espaces propres de a dans $\mathbb{R}^2$, et les dessiner. Pour un vecteur non nul $u \in \mathbb{R}^2$, esquisser ses images u $\alpha(u)$, $\alpha^2(u)$, …, $\alpha^k(u)$, … dans les trois cas suivants :

u est dans l'un des espaces propres,

u est dans l'autre espace propre,

u n'est pas dans un espace propre.

(V.16$^\sharp$) (= exercice pour faire trébucher les calculettes) On considère l'équation aux différences $u_n = pu_{n-1} + qu_{n-2}$ où p et q sont deux constantes réelles telles que $p^2 + 4q > 0$.

(i) Vérifier que les solutions de cette équation sont de la forme $u_n = \alpha s^n + \beta t^n$, où α, β sont des constantes et où s, t sont les racines du polynôme $x^2 - px - q$.

(ii) Calculer $\lim_{n\to\infty} u_n$ pour la suite définie par $u_0 = 1$, $u_1 = \frac{1}{2}(1 - \sqrt{2})$ et $u_n = u_{n-1} + \frac{1}{4}u_{n-2}$.

(iii) Progammez votre calculette selon (ii) et contemplez la suite u_{25k} pour $1 \leq k \leq 20$.
Indications pour comprendre ces résultats.

(iv) Si $a = \begin{pmatrix} 0 & 1 \\ q & p \end{pmatrix}$, vérifier que $\begin{pmatrix} u_n \\ u_{n+1} \end{pmatrix} = a \begin{pmatrix} u_{n-1} \\ u_n \end{pmatrix}$ pour tout $n \geq 1$.

(v) On suppose p, q et les notations telles que $s > 1$ et $|t| < 1$. Calculer $\lim_{n\to\infty} u_n$ d'abord lorsque $\begin{pmatrix} u_0 \\ u_1 \end{pmatrix}$ est sur l'axe propre correspondant à s, ensuite sur l'axe propre correspondant à t.

Les données numériques de (ii) correspondent-elles à l'un de ces cas ?

(vi) Calculer $\lim_{n\to\infty} u_n$ lorsque $\begin{pmatrix} u_0 \\ u_1 \end{pmatrix}$ n'est pas sur l'un des axes propres de a.

[Pour (v) et (vi), voir l'exercice précédent.]

(vii) Dans la situation de (ii), observer que les pentes des axes propres de a sont irrationnelles, et expliquer le résultat de (iii) par une approximation rationnelle d'un nombre irrationnel.

(V.17) Soient $\mathbb{K}$ un corps, $n \geq 1$ un entier, $c_1, \ldots, c_n \in \mathbb{K}$ et

$$a = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 0 & 1 \\ -c_n & -c_{n-1} & -c_{n-2} & \cdots & -c_2 & -c_1 \end{pmatrix} \in M_n(\mathbb{K}).$$

Calculer le polynôme caractéristique de a.

(V.18) On considère deux entiers $n, k \geq 0$, l'espace $\mathcal{P}_n$ des fonctions polynomiales $\mathbb{R} \longrightarrow \mathbb{R}$ de degrés au plus n, et une fonction polynomiale Q de degré k à coefficient dominant 1. (Si Q est de la forme $x \longmapsto q_0 + q_1 x + \cdots + q_k x^k$, son *coefficient dominant* est q_k.)

(i) Calculer les valeurs propres de l'endomorphisme linéaire $\delta_Q \in \mathcal{L}(\mathcal{P}_n)$ défini par $(\delta_Q P)(x) = \frac{d^k}{dx^k}(Q(x)P(x))$. Pour tout $j \in \{0, 1, \ldots, h\}$, montrer que δ_Q possède un unique vecteur propre $P_j \in \mathcal{P}_n$ de degré j à coefficient dominant 1.

[*Indication* : Si $\mathcal{B} = (f_0, \ldots, f_n)$ est la base de $\mathcal{P}_n$ définie par $f_j(x) = x^j$, la matrice $M_{\mathcal{B}}(\delta_Q)$ est triangulaire.]

(ii) Dans le cas où $k = 2$ et $Q(x) = x^2 - 1$, calculer P_0, P_1, P_2.

[Pour en savoir plus sur les opérateurs δ_Q, voir par exemple l'article de G. Másson et B. Shapiro, *On polynomial eigenfunctions of a hypergeometric-type operator*, Experimental Mathematics **10 :4** (2001), pages 609–618.]

(V.19) Soient V, W deux espaces vectoriels complexes et $\alpha \in \mathcal{L}(V, W)$, $\beta \in \mathcal{L}(W, V)$ deux applications linéaires. Montrer que l'ensemble des valeurs propres *non nulles* de $\alpha\beta$ coïncide avec l'ensemble des valeurs propres non nulles de $\beta\alpha$.

Exhiber un exemple tel que 0 soit une valeur propre de $\alpha\beta$, mais pas de $\beta\alpha$.

[*Indication* : Si λ *n'est pas* une valeur propre de $\alpha\beta$ et si $\rho = (\alpha\beta - \lambda \mathrm{id}_W)^{-1}$, vérifier que $\beta\alpha - \lambda \mathrm{id}_V$ est inversible d'inverse $\lambda^{-1}(\beta\rho\alpha - \mathrm{id}_V)$.]

CHAPITRE VI

Espaces hermitiens et euclidiens, adjoints des opérateurs, théorème spectral

1. Espaces hermitiens

Lecture conseillée : S. Axler, *Linear algebra done right,* Second edition, Springer (1997), chapitre 5.

DÉFINITION 1.1. Une *forme hermitienne*[1] sur un espace vectoriel complexe V est une application $h : V \times V \longrightarrow \mathbb{C}$ telle que

(i) (= antilinéarité en la première variable) $h(v + v', w) = h(v, w) + h(v', w)$ et $h(\lambda v, w) = \overline{\lambda} h(v, w)$ pour tous $v, v', w \in V$ et $\lambda \in \mathbb{C}$;

(ii) (= linéarité en la seconde variable) $h(v, w + w') = h(v, w) + h(v, w')$ et $h(v, \lambda w) = \lambda h(v, w)$ pour tous $v, w, w' \in V$ et $\lambda \in \mathbb{C}$;

(iii) $h(w, v) = \overline{h(v, w)}$ pour tous $v, w \in V$.

REMARQUES. (a) Ces conditions ne sont pas indépendantes ; par exemple, (ii) résulte de (i) et (iii).

(b) De nombreux auteurs décident de définir les formes hermitiennes linéaires en la première variable et antilinéaires en la seconde, contrairement au choix adopté ici ! ! !

(c) Il résulte de (iii) que $h(v, v) \in \mathbb{R}$ pour tout $v \in V$.

DÉFINITION 1.2. On dit qu'une forme hermitienne h sur V est

non dégénérée si, pour $v \in V$, on a $v = 0 \Longleftrightarrow h(v, w) = 0$ pour tout $w \in V$;

positive si $h(v, v) \geq 0$ pour tout $v \in V$;

définie positive si $h(v, v) > 0$ pour tout $v \in V, v \neq 0$.

Il est évident qu'une forme définie positive est non dégénérée.

EXEMPLE. Sur $\mathbb{C}^3$, on définit trois formes hermitiennes par

$$\begin{aligned} h_I(v, w) &= \overline{v_1} w_1 + \overline{v_3} w_3, \\ h_{II}(v, w) &= \overline{v_1} w_1 + \overline{v_2} w_2 - \overline{v_3} w_3, \\ h_{III}(v, w) &= \overline{v_1} w_1 + \overline{v_2} w_2 + \overline{v_3} w_3. \end{aligned}$$

La forme h_I est positive dégénérée, la forme h_{II} est non dégénérée et n'est pas positive, la forme h_{III} est définie positive. Noter que l'application $h_{IV} : \mathbb{C}^3 \times \mathbb{C}^3 \longrightarrow \mathbb{C}$ définie par $h_{IV}(v, w) = \overline{v_1} w_1 + i\overline{v_2} w_2$ *n'est pas* une forme hermitienne ! (car $\overline{h_{IV}(v, v)}$ et $h_{IV}(v, v)$ sont en général des nombres complexes *différents*).

[1] En référence à Charles Hermite, mathématicien français, 1827–1901. Il a obtenu de nombreux résultats sur les fonctions elliptiques, et plus généralement sur les liens profonds qui existent entre l'arithmétique et l'analyse. On lui doit la preuve de la transcendance du nombre $e = \sum_{n=0}^{\infty} \frac{1}{n!} \approx 2,71828$.

THÉORÈME VI.1 (inégalité de Cauchy[2]-Schwarz[3]). *Soit V un espace vectoriel complexe muni d'une forme hermitienne positive h. Alors*

$$|h(v,w)|^2 \leq h(v,v)h(w,w)$$

pour tous $v, w \in V$.

PREUVE. Etant donné $v, w \in V$, posons $a = h(v,v)$, $b = h(v,w)$ et $c = h(w,w)$, et remarquons que a, c sont des nombres réels. Il s'agit de montrer que $|b|^2 \leq ac$.

Supposons d'abord que $c \neq 0$. Pour tout $\lambda \in \mathbb{C}$, on a

$$(*) \qquad h(v+\lambda w, v+\lambda w) = a + \overline{\lambda b} + \lambda b + |\lambda|^2 c \geq 0.$$

En particulier, pour $\lambda = -\overline{b}/c$, on trouve $a - b\overline{b}/c \geq 0$.

De même, si $a \neq 0$, on a pour tout $\lambda \in \mathbb{C}$

$$(**) \qquad h(\lambda v + w, \lambda v + w) = |\lambda|^2 a + \overline{\lambda} b + \lambda \overline{b} + c \geq 0.$$

En particulier, pour $\lambda = -b/a$, on trouve $c - |b|^2/a \geq 0$.

Si $a = c = 0$, on pose $\lambda = -\overline{b}$ dans $(*)$ ou $\lambda = -b$ dans $(**)$ et on trouve $-2|b|^2 \geq 0$, c'est-à-dire $b = 0$. □

PROPOSITION 1.1. *Soit V un espace vectoriel complexe de dimension finie et muni d'une forme hermitienne non dégénérée h.*

Pour toute forme linéaire $\xi : V \longrightarrow \mathbb{C}$, il existe un unique vecteur $v_\xi \in V$ tel que

$$\xi(w) = h(v_\xi, w)$$

pour tout $w \in V$.

PREUVE. Notons V' l'espace dual de V.

Soit d'abord h une forme hermitienne quelconque sur V. Pour tout $v \in V$, notons $\xi_v \in V'$ la forme linéaire définie par $\xi_v(w) = h(v,w)$. L'application

$$\phi : V \longrightarrow V', \quad v \longmapsto \xi_v$$

a les propriétés suivantes : $\phi(v+v') = \phi(v) + \phi(v')$ et $\phi(\lambda v) = \overline{\lambda}\phi(v)$ pour tous $v, v' \in V$ et $\lambda \in \mathbb{C}$. L'image de ϕ est donc un sous-espace vectoriel de V' (bien que ϕ ne soit PAS linéaire !).

Supposons désormais la forme hermitienne h non dégénérée, c'est-à-dire l'application ϕ injective. Soient $v_1, \ldots, v_k$ des vecteurs linéairement indépendants de V et $\xi_1, \ldots, \xi_k$ leurs images par ϕ ; alors ces images sont linéairement indépendantes. (Preuve : soient $\lambda_1, \ldots, \lambda_k \in \mathbb{C}$ tels que $\lambda_1\xi_1 + \cdots + \lambda_k\xi_k = 0$. Alors $\overline{\lambda_1}v_1 + \cdots + \overline{\lambda_k}v_k = 0$ [car l'application

[2]Augustin Louis Cauchy, mathématicien français, 1789–1857 ; il a joué un rôle important dans la création de la théorie des groupes, dans les fondements de l'analyse réelle (voir le cours d'Analyse I), et dans les débuts de l'analyse complexe (voir le cours d'Analyse II).

[3]Hermann Schwarz, mathématicien allemand, 1843–1921, professeur au Poly de Zurich de 1869 à 1875 ; il a tiré parti de l'inégalité dite de Cauchy-Schwarz pour des produits scalaires définis par des intégrales doubles du type

$$\langle f \mid g \rangle = \int\int_D f(x,y)g(x,y)dxdy$$

où f et g désignent des fonctions à valeurs réelles définies sur un ensemble D du plan et satisfaisant des conditions appropriées. A cette inégalité se rattache également le nom de V. Bunyakovski, qui l'a utilisée (avant Schwarz) pour des produits scalaires définis par des intégrales simples (voir les exemples (ii)– (iv) de la page 115).

ϕ est injective, et $\phi(\overline{\lambda_1}v_1 + \cdots + \overline{\lambda_k}v_k) = \lambda_1\xi_1 + \cdots + \lambda_k\xi_k]$, donc $\overline{\lambda_1} = \cdots = \overline{\lambda_k} = 0$, ou encore $\lambda_1 = \cdots = \lambda_k = 0$.) Il en résulte que la dimension du sous-espace $\phi(V)$ de V' est égale à celle de V, donc aussi à celle de V', de sorte que $\phi(V) = V'$.

Ceci montre l'assertion d'existence de la proposition. La preuve de l'unicité est laissée en exercice au lecteur. □

Remarque. Un argument tout à fait analogue montre que, pour toute forme antilinéaire $\eta : V \longrightarrow \mathbb{C}$, il existe un unique vecteur $w_\eta \in V$ tel que $\eta(v) = h(v, w_\eta)$ pour tout $v \in V$.

Définition 1.3. Un *produit scalaire hermitien* (on dit aussi plus simplement *produit scalaire*) sur un espace vectoriel complexe est une forme hermitienne définie positive sur cet espace.

On écrit souvent $\langle v \mid w\rangle$ le produit scalaire de deux vecteurs v, w d'un espace hermitien. Un *espace hermitien* est un espace vectoriel complexe de dimension finie muni d'un produit scalaire.

Dans un espace muni d'un produit scalaire, on définit la *norme* d'un vecteur v comme étant le nombre réel positif $\|v\| = \sqrt{\langle v \mid v\rangle}$.

Propriétés de la norme

Proposition 1.2. *Soit V un espace vectoriel muni d'un produit scalaire.*

(i) $\|v + w\| \leq \|v\| + \|w\|$ *pour tous* $v, w \in V$.

(ii) $\|\lambda v\| = |\lambda|\,\|v\|$ *pour tous* $\lambda \in \mathbb{C}$ *et* $v \in V$.

(iii) $\|v\| \geq 0$ *pour tout* $v \in V$, *avec égalité si et seulement si* $v = 0$.

Preuve. Pour $v, w \in V$, on a

$$\|v + w\|^2 \;=\; \|v\|^2 + 2\,\mathrm{Re}\,\big(\langle v \mid w\rangle\big) + \|w\|^2 \;\leq\; \|v\|^2 + 2\,\|v\|\,\|w\| + \|w\|^2$$

par l'inégalité de Cauchy-Schwarz, d'où la première assertion. Les deux autres assertions résultent immédiatement des définitions. □

Exemples de produits scalaires. (i) L'espace $\mathbb{C}^n$ muni du produit scalaire « canonique »

$$\langle v \mid w\rangle \;=\; \sum_{j=1}^{n} \overline{v_j} w_j.$$

Il existe d'autres produits scalaires sur $\mathbb{C}^n$! Par exemple, pour tout n–uple $(c_1, \ldots, c_n)$ de nombres réels strictement positifs, on obtient un produit scalaire h_c tel que $h_c(v, w) = \sum_{j=1}^{n} c_j \overline{v_j} w_j$.

(ii) L'espace $\mathcal{C}_{\mathbb{C}}[a, b]$ des fonctions continues à valeurs complexes sur un intervalle compact (= fermé borné) $[a, b]$ de l'axe réel, muni du produit scalaire

$$\langle f \mid g\rangle \;=\; \int_a^b \overline{f(t)}\, g(t)\, dt.$$

(iii) Le même espace muni du produit scalaire

$$\langle f \mid g \rangle_w \;=\; \int_a^b \overline{f(t)}\, g(t)\, w(t) dt$$

où $w : [a, b] \longrightarrow \mathbb{R}_+^*$ est une fonction continue par morceaux donnée.

(iv) L'espace des fonctions continues $\mathbb{R} \longrightarrow \mathbb{C}$ qui sont 2π-périodiques, muni du produit scalaire

$$\langle f \mid g \rangle \;=\; \int_{-\pi}^{\pi} \overline{f(t)}\, g(t)\, dt.$$

Comparer avec l'exemple (ii) !

(v) Tout sous-espace vectoriel d'un espace hermitien.

En analyse, il est important de distinguer les produits scalaires qui définissent des espaces *complets,* aussi nommés *espaces de Hilbert*[4]. Il se trouve qu'un espace hermitien de dimension finie est nécessairement complet. La question en dimension infinie est plus délicate, et n'est pas abordée ici.

Définition 1.4. Soit V un espace vectoriel complexe muni d'un produit scalaire. Deux vecteurs $v, w \in V$ sont *orthogonaux* si $\langle v \mid w \rangle = 0$.

Remarques. (i) Si v, w sont orthogonaux, alors

$$\|v + w\|^2 \;=\; \|v\|^2 + \|w\|^2 .$$

(ii) Des vecteurs $v_1, \ldots, v_k$ non nuls orthogonaux deux à deux sont linéairement indépendants.

En effet, soient $\lambda_1, \ldots, \lambda_k \in \mathbb{C}$ tels que $\lambda_1 v_1 + \cdots + \lambda_k v_k = 0$. Pour tout $j \in \{1, \ldots, k\}$, on obtient $\langle v_j \mid \lambda_1 v_1 + \cdots + \lambda_k v_k \rangle = \lambda_j \langle v_j \mid v_j \rangle = 0$, donc $\lambda_j = 0$.

Définition 1.5. Une *suite orthogonale* de vecteurs de V est une suite de vecteurs de V dont les vecteurs sont orthogonaux deux à deux. Une telle suite est *orthonormale* si, de plus, le produit scalaire de tout vecteur de la suite avec lui-même est égal à 1.

Pour toute partie S de V, *l'orthogonal de* S est le sous-ensemble

$$S^\perp \;=\; \{w \in V \mid \langle v \mid w \rangle = 0 \quad \text{pour tout} \quad v \in S\}.$$

Remarque. C'est un sous-espace vectoriel de V.

Théorème VI.2 (norme d'un vecteur et coordonnées selon une base orthonormale). *Soient V un espace hermitien et $(e_1, \ldots, e_n)$ une base orthonormale de V. Alors*

$$v \;=\; \sum_{j=1}^{n} \langle e_j \mid v \rangle e_j \qquad \textit{et} \qquad \|v\|^2 \;=\; \sum_{j=1}^{n} |\langle e_j \mid v \rangle|^2$$

pour tout $v \in V$.

[4] En référence à David Hilbert, mathématicien allemand, 1862–1943. Il est notamment célèbre pour sa liste de 23 problèmes, délivrée à l'occasion du Deuxième Congrès International des Mathématiciens, tenu à Paris du 6 au 12 août 1900.

PREUVE. Puisque $(e_1, \ldots, e_n)$ est une base de V, il existe $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$ (uniquement définis) tels que $v = \sum_{k=1}^{n} \lambda_k e_k$. Par suite

$$\langle e_j \mid v \rangle \;=\; \left\langle e_j \;\middle|\; \sum_{k=1}^{n} \lambda_k e_k \right\rangle \;=\; \sum_{k=1}^{n} \lambda_k \langle e_j \mid e_k \rangle \;=\; \lambda_j$$

pour tout $j \in \{1, \ldots, n\}$, d'où la première égalité. En développant

$$\|v\|^2 \;=\; \left\langle \sum_{j=1}^{n} \lambda_j e_j \;\middle|\; \sum_{k=1}^{n} \lambda_k e_k \right\rangle,$$

on obtient la seconde. □

THÉORÈME VI.3 (Procédé d'orthonormalisation de Gram-Schmidt[5]). *Soient V un espace vectoriel complexe muni d'un produit scalaire et $(v_1, \ldots, v_k)$ une suite de vecteurs de V linéairement indépendants. Alors il existe une suite orthonormale $(e_1, \ldots, e_k)$ dans V telle que*

$$\langle e_1, \ldots, e_j \rangle \;=\; \langle v_1, \ldots, v_j \rangle \qquad \textit{pour tout} \qquad j \in \{1, \ldots, k\}.$$

RAPPEL : $\langle e_1, \ldots, e_j \rangle$ désigne le sous-espace vectoriel de V engendré par $e_1, \ldots, e_j$. Dans cet énoncé, notons que l'ordre des vecteurs $v_1, \ldots, v_k$ est important ; par exemple, si on applique une première fois le procédé à la suite $(v_1, v_2, v_3, \ldots)$ et une seconde fois à la suite $(v_2, v_1, v_3, \ldots)$, on trouve deux résultats *différents*.

PREUVE, PAR RÉCURRENCE SUR k. Si $k = 1$, il suffit de poser $e_1 = v_1 \, \|v_1\|^{-1}$. On suppose désormais $k > 1$ et le procédé démontré jusqu'à $k - 1$. Il existe donc une suite $e_1, \ldots, e_{k-1}$ dans V telle que $\langle e_1, \ldots, e_j \rangle = \langle v_1, \ldots, v_j \rangle$ pour $j \leq k - 1$, et il s'agit de définir e_k.

Posons

$$w_k \;=\; v_k - \sum_{i=1}^{k-1} \langle e_i \mid v_k \rangle e_i.$$

Par hypothèse (indépendance linéaire et récurrence), v_k n'est pas dans le sous-espace $\langle e_1, \ldots, e_{k-1} \rangle = \langle v_1, \ldots, v_{k-1} \rangle$. En particulier, $w_k \neq 0$. Il suffit de poser $e_k = w_k \, \|w_k\|^{-1}$ et de vérifier que les conditions sont satisfaites. Par exemple :

$$\langle e_j \mid e_k \rangle \;=\; \frac{1}{\|w_k\|} \left(\langle e_j \mid v_k \rangle \;-\; \left\langle e_j \middle| \sum_{i=1}^{k-1} \langle e_i \mid v_k \rangle e_i \right\rangle \right) \;=\; 0$$

pour tout $j \in \{1, \ldots, k-1\}$. □

EXEMPLE. Soit $V = \mathcal{C}_{\mathbb{R}}[-1, 1]$ l'espace vectoriel réel (voir le § suivant) des fonctions continues à valeurs réelles sur l'intervalle $[-1, 1]$, muni du produit scalaire de l'exemple 1.9.ii, et soit $(v_0, v_1, \ldots)$ la suite des fonctions définies par $v_n(t) = t^n$. Soit $(e_0, e_1, \ldots)$ la suite de polynômes définis par le procédé de Gram-Schmidt. On définit le *n-ième polynôme de Legendre* P_n en posant

$$(*) \qquad P_n(t) \;=\; \frac{e_n(t)}{e_n(1)}.$$

[5] Jorgen Petersen Gram, mathématicien danois, 1850–1916. Erhard Schmidt, mathématicien allemand, 1876–1959.

(On admet ici que $e_n(1) \neq 0$; voir par exemple l'exercice qui suit).

Montrons par exemple que

$$P_0(t) = 1, \qquad P_1(t) = t, \qquad P_2(t) = \frac{3t^2 - 1}{2}.$$

En effet, on a successivement (en reprenant les notations du numéro précédent) :

$$\|v_0\|^2 = \int_{-1}^{1} dt = 2, \quad e_0(t) = \frac{1}{\sqrt{2}}, \quad P_0(t) = 1 \ ;$$

$$w_1(t) = t - \left(\int_{-1}^{1} \frac{1}{\sqrt{2}} s ds\right) \frac{1}{\sqrt{2}} = t, \quad \|w_1\|^2 = \int_{-1}^{1} t^2 dt = \frac{2}{3},$$

$$e_1(t) = \sqrt{\frac{3}{2}} t, \quad P_1(t) = t;$$

$$w_2(t) = t^2 - \left(\int_{-1}^{1} \sqrt{\frac{3}{2}} s s^2 ds\right) \sqrt{\frac{3}{2}} t - \left(\int_{-1}^{1} \frac{1}{\sqrt{2}} s^2 ds\right) \frac{1}{\sqrt{2}} = t^2 - \frac{1}{3},$$

$$P_2(t) = \frac{1}{2}\left(3t^2 - 1\right).$$

Exercice. On se propose de montrer que

$$(**) \qquad P_n(t) \ = \ \frac{1}{2^n n!} \frac{d^n}{dt^n}\left((t^2 - 1)^n\right).$$

Dans les points (i) à (vi) ci-dessous, $P_n(t)$ est défini par la formule $(**)$, et il s'agit de montrer que ces fonctions sont les mêmes que celles définies à l'exemple précédent.

(i) Vérifier que $P_n(t)$ est un polynôme de degré n et que $P_n(1) = 1$.

(ii) Vérifier que $\frac{d^j}{dt^j}\left((t^2 - 1)^n\right)$ s'annule en $t = \pm 1$ si $j < n$.

(iii) Montrer que $\int_{-1}^{1} P_n(t) t^m dt = 0$ si $m < n$.

[Indication : intégrer m fois par parties, en utilisant (ii).]

(iv) En déduire que $\int_{-1}^{1} P_n(t) P_m(t) dt = 0$ si $n \neq m$.

(v) Conclure : les fonctions définies par $(**)$ coïncident avec celles définies par $(*)$.

(vi) On peut montrer que $\int_{-1}^{1} \left(P_n(t)\right)^2 dt = \frac{2}{2n+1}$ pour tout $n \geq 0$. Vérifier cette égalité lorsque $0 \leq n \leq 2$.

Corollaire. (i) *Tout espace vectoriel hermitien possède une base orthonormale.*

(ii) *Dans un espace hermitien, toute suite orthonormale peut être complétée en une base orthonormale. En particulier :*

Soient V un espace hermitien de dimension n et U un sous-espace de V de dimension k ; alors il existe une base orthonormale $(e_1, \ldots, e_n)$ de V telle que $(e_1, \ldots, e_k)$ soit une base orthonormale de U.

THÉORÈME VI.4. *Soient V un espace hermitien et U un sous-espace vectoriel de V.*

(i) L'espace V est somme directe $U \oplus U^\perp$ de U et de l'orthogonal de U.

(ii) Soit $(e_1, \dots, e_n)$ une base orthonormale de V telle que $(e_1, \dots, e_m)$ est une base orthonormale de U, où $m = \dim U \leq n = \dim V$. Alors $(e_{m+1}, \dots, e_n)$ est une base orthonormale de $U^\perp$.

(iii) Soit $v \in V$; soient $v^\parallel \in U$ et $v^\perp \in U^\perp$ les vecteurs définis par $v = v^\parallel + v^\perp$. Alors :

$$\|v^\perp\| = \|v - v^\parallel\| \leq \|v - u\| \ \forall\ u \in U, \text{ avec égalité si et seulement si } u = v^\parallel ;$$
$$\|v^\parallel\| = \|v - v^\perp\| \leq \|v - w\| \ \forall\ w \in U^\perp, \text{ avec égalité si et seulement si } w = v^\perp.$$

PREUVE. Il est évident que $U \cap U^\perp = \{0\}$, donc que la somme de U et $U^\perp$ est directe.

Soit $(e_1, \dots, e_n)$ comme dans (ii). Pour $v \in V$, on pose

$$v^\parallel = \sum_{j=1}^{m} \langle e_j \mid v \rangle e_j \qquad \text{et} \qquad v^\perp = \sum_{j=m+1}^{n} \langle e_j \mid v \rangle e_j$$

de sorte que $v = v^\parallel + v^\perp$ avec $v^\parallel \in U$ et $v^\perp \in U^\perp$. La somme directe $U \oplus U^\perp$ est donc égale à V. La vérification de (ii) est alors immédiate.

Pour tout $u \in U$, nous avons

$$\|v^\perp\|^2 = \|v - v^\parallel\|^2 \leq \|v - v^\parallel\|^2 + \|v^\parallel - u\|^2 = \|v - u\|^2$$

(voir la remarque (i) suivant la définition 1.10), avec égalité si et seulement si $v^\parallel - u = 0$. La preuve de la dernière assertion du théorème est analogue. □ □

Digression. Voici l'esquisse d'un autre argument pour montrer, étant donné $v \in V$, l'existence et l'unicité du vecteur $v^\parallel \in U$ tel que $\|v - v^\parallel\|^2 \leq \|v - u\|^2$ pour tout $u \in U$.

On choisit arbitrairement $u_0 \in U$ et on pose $R = \|v - u_0\|$. On considère le sous-ensemble

$$K = \{u \in U \mid \|v - u\| \leq R\},$$

qui est manifestement fermé borné (= compact) dans U, et la fonction

$$\delta_L : \begin{cases} K & \longrightarrow & \mathbb{R} \\ u & \longmapsto & \|v - u\|^2 \end{cases}$$

qui est manifestement continue. Par un théorème bien connu (voir le cours d'Analyse I, semestre d'hiver si la dimension de U est 1, et semestre d'été en général), il existe un vecteur $v^\parallel \in K \subset U$ qui réalise le minimum de cette fonction.

Soient $v_1^\parallel, v_2^\parallel$ deux minima de cette fonction, et soit $v^\parallel = \frac{1}{2}\left(v_1^\parallel + v_2^\parallel\right)$. Si $v_1^\parallel$ et $v_2^\parallel$ étaient distincts, on montrerait sans peine que $\|v - v^\parallel\| < \left\|v - v_1^\parallel\right\|$, contrairement à la définition de $v_1^\parallel$. C'est ainsi qu'on s'assure de l'unicité du vecteur $v^\parallel$.

TERMINOLOGIE L'espace $U^\perp$ est le *supplémentaire orthogonal* de U dans V. Une telle somme $U \oplus U^\perp$ est une *somme directe orthogonale* (ou "somme orthogonale"). L'application

$$p_U : \begin{cases} V = U \oplus U^\perp & \longrightarrow & V \\ v = v^\parallel + v^\perp & \longmapsto & v^\parallel \end{cases}$$

est la *projection orthogonale de V sur U*. Noter que cette projection est linéaire : $p_U \in \mathcal{L}(V)$.

Propriétés de la projection orthogonale.

Soit $p_U \in \mathcal{L}(V)$ la projection orthogonale de V sur U (notations du théorème VI.4). Il résulte de l'assertion (i) du théorème VI.4 que

$p_U(v) = v$ si et seulement si $v \in U$,

$\mathrm{Im}(p_U) = U$,

$\mathrm{Ker}(p_U) = U^\perp$,

et de la remarque (i) suivant la définition de "vecteurs orthogonaux" que

$\|p_U(v)\| \leq \|v\|$ pour tout $v \in V$, avec égalité si et seulement si $v \in U$.

Notons de plus que

$p_U^2 = p_U$,

$p_{U^\perp} = \mathrm{id}_V - p_U$,

$p_U = \mathrm{id}_V$ si et seulement si $U = V$,

$p_U = 0$ si et seulement si $U = \{0\}$.

[Voir encore l'exercice (VI.2).]

Proposition 1.3. *Soient V un espace hermitien et U un sous-espace vectoriel de V de dimension m.*

(i) L'orthogonal de l'orthogonal de U coïncide avec U ; autrement dit : $(U^\perp)^\perp = U$.

(ii) Si $(e_1, \ldots, e_m)$ une base orthonormale de U, la projection orthogonale $p_U \in \mathcal{L}(V)$ de V sur U est donnée par

$$p_U(v) = \sum_{j=1}^{m} \langle e_j \mid v \rangle e_j \qquad \textit{pour tout} \quad v \in V.$$

Preuve. C'est une conséquence immédiate du théorème VI.2. □

Remarques. (i) La proposition vaut aussi pour un espace *euclidien* (voir le paragraphe suivant).

(ii) L'assertion (ii) de la proposition est encore vraie – et souvent appliquée – lorsque V est un espace vectoriel de dimension infinie muni d'un produit scalaire, c'est-à-dire d'une forme hermitienne (ou bilinéaire symétrique dans le cas d'un espace vectoriel réel) définie positive, et U un sous-espace vectoriel de U *de dimension finie.*

Exercice. Pour tout sous-ensemble non vide S d'un espace hermitien V, montrer que $\left(S^\perp\right)^\perp$ est le sous-espace vectoriel de V engendré par S.

Exemple. Soit $V = \mathcal{C}([0,1])$ l'espace des fonctions continues sur l'intervalle unité, à valeurs réelles, espace muni du produit scalaire défini par

$$\langle f_1 \mid f_2 \rangle = \int_0^1 f_1(t) f_2(t) dt.$$

Soit U le sous-espace de V des fonctions polynomiales de degrés au plus 1, c'est-à-dire le sous-espace de V engendré par les fonctions v_0, v_1, où $v_j(t) = t^j$. Soit par ailleurs $f \in V$ la fonction définie par $f(t) = e^t$. Calculons la projection orthogonale g de f dans U, c'est-à-dire la fonction $g \in U$ telle que $\|f - g\| \leq \|f - u\|$ pour tout $u \in U$.

Appliquons le procédé de Gram-Schmidt à la base (v_0, v_1) de U. Avec les notations de la preuve du théorème VI.3, on calcule successivement

$$\begin{aligned}
\|v_0\|^2 &= \int_0^1 dt = 1 \qquad \text{donc} \qquad e_0 = v_0 \\
\langle e_0 \mid v_1\rangle &= \int_0^1 t dt = \frac{1}{2} \\
w_1(t) &= v_1(t) - \langle e_0 \mid v_1\rangle e_0(t) = t - \frac{1}{2} \\
\|w_1\|^2 &= \int_0^1 \left(t - \frac{1}{2}\right)^2 dt = \frac{1}{3} - \frac{1}{2} + \frac{1}{4} = \frac{1}{12} \\
e_1(t) &= \frac{w_1(t)}{\|w_1\|} = 2\sqrt{3}\, t - \sqrt{3} \\
\langle e_0 \mid f\rangle &= \int_0^1 e^t dt = \left|e^t\right|_0^1 = e - 1 \\
\langle e_1 \mid f\rangle &= 2\sqrt{3}\int_0^1 te^t dt - \sqrt{3}\int_0^1 e^t dt \\
&= 2\sqrt{3}\left|te^t - e^t\right|_0^1 - \sqrt{3}\left|e^t\right|_0^1 = 2\sqrt{3} - \sqrt{3}(e-1) \\
&= 3\sqrt{3} - \sqrt{3}e
\end{aligned}$$

et enfin

$$\begin{aligned}
g(t) &= \langle e_0 \mid f\rangle e_0(t) + \langle e_1 \mid f\rangle e_1(t) \\
&= e - 1 + \sqrt{3}(3-e)(2\sqrt{3}\, t - \sqrt{3}) \\
&= 6(3-e)t + 4e - 10
\end{aligned}$$

pour tout $t \in [0, 1]$.

Le lecteur est invité à calculer la projection h de la fonction exponentielle f dans le sous-espace de V engendré par les fonctions polynomiales de degrés au plus 2.

Les problèmes de ce type, consistant à calculer la meilleure approximation (au sens convenable) d'une fonction, ci-dessus f, par une fonction d'un sous-espace donné, ci-dessus U, sont d'une grande importance pratique.

2. Espaces euclidiens

Soit V un espace vectoriel sur un corps $\mathbb{K}$. Une *forme bilinéaire symétrique* sur V est une application

$$\beta : V \times V \longrightarrow \mathbb{K}$$

telle que

(i) $\beta(v + v', w) = \beta(v, w) + \beta(v', w)$ et $\beta(\lambda v, w) = \lambda\beta(v, w)$

(ii) $\beta(v, w + w') = \beta(v, w) + \beta(v, w')$ et $\beta(v, \lambda w) = \lambda\beta(v, w)$

(iii) $\beta(w, v) = \beta(v, w)$

pour tous $v, v', w, w' \in V$ et $\lambda \in \mathbb{K}$.

Soit V un espace vectoriel *réel*. Un *produit scalaire euclidien* (on dit aussi *produit scalaire*) est une forme bilinéaire symétrique sur V telle que $\beta(v, v) > 0$ pour tout $v \in V$, $v \neq 0$.

Un *espace euclidien* est un espace vectoriel réel de dimension finie muni d'un produit scalaire. Par exemple, $\mathbb{R}^n$ muni du *produit scalaire canonique* défini par

$$\langle x \mid y \rangle \;=\; \sum_{j=1}^{n} x_j y_j$$

est un espace euclidien.

On définit la *norme* d'un vecteur dans un espace euclidien par $\|x\| = \sqrt{\langle x \mid x \rangle}$, et on montre les analogues des résultats du paragraphe précédent, en particulier :

(i) toute forme linéaire peut s'écrire $w \longmapsto \langle v \mid w \rangle$ pour un certain $v \in V$ (proposition 1.1) ;

(ii) l'inégalité de Cauchy-Schwarz (théorème VI.1) ;

(iii) les propriétés de la norme (proposition 1.2) ;

(iv) le procédé d'orthonormalisation de Gram-Schmidt (théorème VI.3) ;

(v) l'existence et les propriétés des projections orthogonales (page 120).

3. Adjoints, opérateurs auto-adjoints, unitaires et normaux

Définition 3.1. Soit V, W deux espaces hermitiens et $\alpha : V \longrightarrow W$ une application linéaire. On définit *l'application adjointe* $\alpha^* : W \longrightarrow V$ de α par

$$(\sharp) \qquad \langle \alpha^*(w) \mid v \rangle_V \;=\; \langle w \mid \alpha(v) \rangle_W \qquad \text{pour tous} \quad v \in V \quad \text{et} \quad w \in W.$$

Voici une manière moins elliptique de présenter cette définition.

Soit $w \in W$. Notons $\xi_w : V \longrightarrow \mathbb{C}$ la forme linéaire définie par $\xi_w(v) = \langle w \mid \alpha(v) \rangle_W$ pour tout $v \in V$. La proposition 1.1 montre qu'il existe un unique vecteur $u_w \in V$ tel que $\xi_w(v) = \langle u_w \mid v \rangle_V$ pour tout $v \in V$. Pour $w, w' \in W$ et $\lambda \in \mathbb{C}$, il est immédiat de vérifier que

$$\xi_{w+w'} \;=\; \xi_w + \xi_{w'} \qquad \text{et} \qquad \xi_{\lambda w} \;=\; \overline{\lambda}\xi_w$$

donc que

$$u_{w+w'} \;=\; u_w + u_{w'} \qquad \text{et} \qquad u_{\lambda w} \;=\; \lambda u_w.$$

En d'autres termes, l'application $w \longmapsto u_w$ de W dans V est linéaire ; c'est cette application qui est, par définition, l'adjointe α^* de l'application α, et qui est définie par les égalités de $(\sharp)$.

Dans le contexte du présent chapitre, on dit souvent « opérateur linéaire » ou même « opérateur » au lieu de « application linéaire » . Ainsi α^* est *l'opérateur adjoint* de α.

Propriétés immédiates

Soient $\alpha, \beta \in \mathcal{L}(V, W)$ et $\lambda \in \mathbb{C}$. Alors

$$\begin{aligned}(\alpha+\beta)^* &= \alpha^* + \beta^*;\\ (\lambda\alpha)^* &= \overline{\lambda}\alpha^*;\\ (\alpha^*)^* &= \alpha.\end{aligned}$$

Si $W = V$, alors

$$(\mathrm{id}_V)^* = \mathrm{id}_V\,.$$

Enfin, si X est un espace hermitien et si $\gamma \in \mathcal{L}(W, X)$, alors

$$(\gamma\alpha)^* = \alpha^*\gamma^*.$$

Conséquence. Si $\alpha : V \longrightarrow W$ est inversible, alors $(\alpha^{-1})^* = (\alpha^*)^{-1}$.

Proposition 3.1. *Dans la situation de la définition précédente, on suppose de plus V et W munis respectivement de bases orthonormales $\mathcal{E} = (e_1, \ldots, e_m)$ et $\mathcal{F} = (f_1, \ldots, f_n)$. Alors les matrices $M_{\mathcal{F},\mathcal{E}}(\alpha)$ et $M_{\mathcal{E},\mathcal{F}}(\alpha^*)$ relativement à ces bases satisfont la relation*

$$M_{\mathcal{E},\mathcal{F}}(\alpha^*)_{j,k} = \overline{M_{\mathcal{F},\mathcal{E}}(\alpha)_{k,j}}$$

pour tous $j \in \{1, \ldots, m\}$ et $k \in \{1, \ldots n\}$.

Preuve. On a

$$M_{\mathcal{E},\mathcal{F}}(\alpha^*)_{j,k} = \langle e_j \mid \alpha^* f_k\rangle = \langle \alpha e_j \mid f_k\rangle = \overline{\langle f_k \mid \alpha e_j\rangle} = \overline{M_{\mathcal{F},\mathcal{E}}(\alpha)_{k,j}}. \qquad \square$$

La proposition 3.1 «oblige» à la définition suivante : si $a \in M_{n,m}(\mathbb{C})$ est une matrice complexe, la *matrice adjointe* $a^* \in M_{m,n}(\mathbb{C})$ de a est la matrice transposée conjugée de a. Ainsi, la conclusion de la proposition 3.1 s'écrit

$$M_{\mathcal{E},\mathcal{F}}(\alpha^*) = (M_{\mathcal{F},\mathcal{E}}(\alpha))^*.$$

Proposition 3.2. *Soient V un espace hermitien et $\alpha \in \mathcal{L}(V)$. Les trois propriétés suivantes sont équivalentes :*

(i) *l'application* $h : \begin{cases} V \times V \longrightarrow \mathbb{C} \\ (v, w) \longmapsto \langle v \mid \alpha(w)\rangle \end{cases}$ *est une forme hermitienne ;*

(ii) $\langle u \mid \alpha(u)\rangle \in \mathbb{R}$ *pour tout $u \in V$;*

(iii) $\alpha^* = \alpha$.

Preuve. L'implication (i) $\Longrightarrow$ (ii) résulte des définitions. Réciproquement, supposons que la propriété (ii) est vérifiée, c'est-à-dire que $h(u, u) \in \mathbb{R}$ pour tout $u \in V$; la

seule condition à vérifier pour h qui ne soit pas immédiate est la condition "$h(v,w) = \overline{h(w,v)}$". Alors, pour tous $v, w \in V$, on a

$$\begin{aligned}
4h(v,w) &= h(v+w, v+w) - h(v-w, v-w) \\
&\qquad - i\,h(v+iw, v+iw) + i\,h(v-iw, v-iw) \\
&= h(v+w, v+w) - h(-(v-w), -(v-w)) \\
&\qquad - i\,h(-i(v+iw), -i(v+iw)) + i\,h(i(v-iw), i(v-iw)) \\
&= h(w+v, w+v) - h(w-v, w-v) \\
&\qquad + i\,h(w+iv, w+iv) - i\,h(w-iv, w-iv) \\
&= 4\overline{h(w,v)}
\end{aligned}$$

de sorte que la propriété (i) est aussi vérifiée.

L'équivalence (i) $\iff$ (iii) se montre comme suit :

$h(v,w) = \overline{h(w,v)}$ pour tous $v, w \in V$ si et seulement si

$\langle v \mid \alpha(w)\rangle = \overline{\langle w \mid \alpha(v)\rangle} = \overline{\langle \alpha^*(w) \mid v\rangle} = \langle v \mid \alpha^*(w)\rangle$ pour tous $v, w \in V$ si et seulement si

$\alpha w = \alpha^* w$ pour tous $w \in V$ si et seulement si $\alpha = \alpha^*$. □

Définition 3.2. Un opérateur satisfaisant les propriétés de la proposition précédente est dit *autoadjoint* ou *hermitien.*

Il y a une définition analogue pour les *matrices* complexes (voir la proposition 3.1) : la matrice $a \in M_n(\mathbb{C})$ est *autoadjointe* ou *hermitienne* si elle est égale à sa transposée conjuguée, ce qu'on écrit $a^* = a$. Ainsi

étant donné un espace hermitien V muni d'une base orthonormale $\mathcal{E}$, un opérateur $\alpha \in \mathcal{L}(V)$ est autoadjoint si et seulement si la matrice $M_{\mathcal{E}}(\alpha)$ est autoadjointe.

Les opérateurs autoadjoints (et leurs analogues en dimension infinie) jouent un rôle fondamental dans le formalisme de la physique quantique.

Remarque. Il résulte en particulier de la proposition 3.2 qu'un opérateur α sur un espace hermitien V tel que $\langle v \mid \alpha v\rangle \geq 0$ pour tout $v \in V$ (on dit alors que α est *positif*, voir l'exercice (VI.7)) est nécessairement autoadjoint.

L'analogue de ce fait pour un espace euclidien (espace vectoriel *réel* muni d'un produit scalaire) n'est *pas* vrai : vérifier par exemple que, sur l'espace $\mathbb{R}^2$ muni du produit scalaire standard, la matrice *non symétrique* $\begin{pmatrix} 10 & 1 \\ 3 & 10 \end{pmatrix}$ définit un opérateur $\alpha : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ tel que $\langle v \mid \alpha(v)\rangle \geq 0$ pour tout $v \in \mathbb{R}^2$.

Proposition 3.3. *Soient V, W deux espaces hermitiens et $\alpha \in \mathcal{L}(V,W)$. Les trois propriétés suivantes sont équivalentes :*

(i) $\|\alpha(v)\| = \|v\|$ *pour tout $v \in V$;*

(ii) $\langle \alpha(v) \mid \alpha(w)\rangle = \langle v \mid w\rangle$ *pour tous $v, w \in V$;*

(iii) $\alpha^*\alpha = \mathrm{id}_V$.

De plus, si $\dim(W) = \dim(V)$, *ces conditions sont équivalentes à*

(iv) $\alpha\alpha^* = \mathrm{id}_W$.

PREUVE. Les implications (ii) $\Longrightarrow$ (i) et (ii) $\Longleftrightarrow$ (iii) résultent immédiatement des définitions, et il s'agit de montrer que (i) $\Longrightarrow$ (ii).

On vérifie d'abord qu'on a

$$\begin{aligned} 4\langle v \mid w\rangle &= \langle v+w \mid v+w\rangle - \langle v-w \mid v-w\rangle - i\,\langle v+iw \mid v+iw\rangle + i\,\langle v-iw \mid v-iw\rangle \\ &= \|v+w\|^2 - \|v-w\|^2 - i\,\|v+iw\|^2 + i\,\|v-iw\|^2 \end{aligned}$$

pour tous $v, w \in V$ (comparer avec la preuve de la proposition 3.2), puis on établit que (i) implique (ii). La preuve de la dernière assertion est laissée au lecteur. □

DÉFINITION 3.3. Une application linéaire satisfaisant les propriétés de la proposition précédente est une *isométrie* de V dans W. Si l'espace but W coïncide avec l'espace source V, on dit aussi qu'une telle application est un *opérateur unitaire* sur V.

Une matrice $a \in M_n(\mathbb{C})$ est dite *unitaire* si $a^*a = I_n$. (Condition équivalente : $aa^* = I_n$.) Ainsi

> *étant donné un espace hermitien V muni d'une base orthonormale $\mathcal{E}$, un opérateur $\alpha \in \mathcal{L}(V)$ est unitaire si et seulement si la matrice $M_{\mathcal{E}}(\alpha)$ est unitaire.*

EXERCICE. Si α, β sont deux opérateurs unitaires sur V, alors les opérateurs $\alpha\beta$ et α^{-1} sur V sont aussi unitaires.

Le *groupe unitaire* $\mathcal{U}(V)$ d'un espace hermitien V est le groupe constitué des opérateurs unitaires sur V. Si V est l'espace $\mathbb{C}^n$ muni du produit scalaire canonique, on écrit $\mathcal{U}(n)$ au lieu de $\mathcal{U}(\mathbb{C}^n)$.

Pour tout $\alpha \in \mathcal{L}(V)$, la proposition 3.1 implique que $\det(\alpha^*) = \overline{\det(\alpha)}$. Par suite, si α est unitaire, alors $1 = \det(\alpha^*\alpha) = \overline{\det(\alpha)}\det(\alpha)$, et $|\det(\alpha)| = 1$. Le déterminant fournit donc un homomorphisme $\det : \mathcal{U}(V) \longrightarrow \{z \in \mathbb{C} \mid |z| = 1\}$ dont le noyau (l'image inverse de $\{1\}$) est par définition le *groupe unitaire spécial,* noté $\mathcal{SU}(V)$; si $V = \mathbb{C}^n$, on obtient ainsi le groupe noté $\mathcal{SU}(n)$.

EXERCICE. Montrer que le déterminant $\mathcal{U}(V) \longrightarrow \{z \in \mathbb{C} \mid |z| = 1\}$ est un homomorphisme de groupes qui est surjectif.

CARACTÉRISATION. Une matrice $a \in M_n(\mathbb{C})$ est unitaire si et seulement si ses colonnes $(a_1, \ldots, a_n)$ constituent une base orthonormale de l'espace hermitien $\mathbb{C}^n$ (muni du produit scalaire canonique).

En effet, notons $\alpha \in \mathcal{L}(\mathbb{C}^n)$ l'opérateur linéaire de matrice a relativement à la base canonique $(e_1, \ldots, e_n)$. Si a (c'est-à-dire α) est unitaire, alors $\langle a_j | a_k\rangle = \langle \alpha(e_j) | \alpha(e_k)\rangle = \langle e_j | e_k\rangle$, donc $\langle a_j | a_k\rangle$ vaut 1 si $j = k$ et 0 si $j \neq k$, ce qui veut bien dire que les vecteurs $a_1, \ldots, a_n \in \mathbb{C}^n$ sont orthonormaux (où $j, k = 1, \ldots, n$). Réciproquement, si les vecteurs $a_1, \ldots, a_n$ sont orthonormaux, alors $\langle (\alpha^*\alpha)(e_j) | e_k\rangle = \langle \alpha(e_j) | \alpha(e_k)\rangle = \langle e_j | e_k\rangle$ pour tout $j, k \in \{1, \ldots, n\}$, donc $(\alpha^*\alpha)(e_j) = e_j$ pour tout $j \in \{1, \ldots, n\}$, donc $\alpha^*\alpha = \mathrm{id}$, donc α (c'est-à-dire a) est unitaire.

Plus généralement, pour un espace hermitien V de dimension n muni d'une base orthonormale, un opérateur linéaire sur V est unitaire si et seulement si sa matrice

relativement à la base donnée a des colonnes qui constituent une base orthonormale de $\mathbb{C}^n$.

DÉFINITION 3.4. Un opérateur α sur un espace hermitien est *normal* s'il commute avec son adjoint, c'est-à-dire si $\alpha^*\alpha = \alpha\alpha^*$.

On définit de même les matrices complexes carrées qui sont normales.

EXEMPLES. (i) La matrice $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ n'est *pas* normale.

(ii) La matrice $\begin{pmatrix} 1 & \lambda \\ \lambda & 1 \end{pmatrix}$ est normale pour tout $\lambda \in \mathbb{C}$; elle est autoadjointe si et seulement si $\lambda \in \mathbb{R}$, unitaire si et seulement si $\lambda = 0$.

(iii) Tout opérateur unitaire est normal, tout opérateur autoadjoint est normal.

(iv) Toute matrice diagonale est normale.

LEMME. *Soient V un espace hermitien et $\alpha \in \mathcal{L}(V)$ un opérateur normal. Alors*

$$\mathrm{Ker}(\alpha^*) = \mathrm{Ker}(\alpha).$$

De plus, si $v \in V$ est un vecteur propre de α de valeur propre λ, alors v est aussi un vecteur propre de α^* de valeur propre $\overline{\lambda}$.

PREUVE. Soit $w \in V$. On a

$$\begin{aligned} \|\alpha(w)\|^2 &= \langle \alpha(w) \mid \alpha(w) \rangle = \langle \alpha^*\alpha(w) \mid w \rangle = \langle \alpha\alpha^*(w) \mid w \rangle = \langle \alpha^*(w) \mid \alpha^*(w) \rangle \\ &= \|\alpha^*(w)\|^2 . \end{aligned}$$

Il en résulte que $\mathrm{Ker}(\alpha) = \mathrm{Ker}(\alpha^*)$.

Soit v un vecteur propre comme dans l'énoncé, de sorte que $v \in \mathrm{Ker}(\lambda \mathrm{id}_V - \alpha)$. Vu ce qui précède, on a aussi $v \in \mathrm{Ker}((\lambda \mathrm{id}_V - \alpha)^*) = \mathrm{Ker}(\overline{\lambda} \mathrm{id}_V - \alpha^*)$, c'est-à-dire $\alpha^*(v) = \overline{\lambda} v$. □

EXERCICE. Si $a = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \in M_2(\mathbb{C})$, déterminer $\mathrm{Ker}(a)$, $\mathrm{Ker}(a^*)$, et constater que $\mathrm{Ker}(a) \neq \mathrm{Ker}(a^*)$.

THÉORÈME VI.5. *Soient V un espace hermitien et α un opérateur linéaire sur V.*

(i) *Si α est normal, deux vecteurs propres de α correspondant à des valeurs propres distinctes sont orthogonaux.*

(ii) *Si α est autoadjoint, toute valeur propre de α est un nombre réel.*

(iii) *Si α est unitaire, toute valeur propre de α est un nombre complexe de module* 1.

PREUVE. (i) Soient $v, w \in V \setminus \{0\}$ et $\lambda, \mu \in \mathbb{C}$ tels que $\alpha(v) = \lambda v$ et $\alpha(w) = \mu w$. On a

$$\mu\langle v \mid w \rangle = \langle v \mid \alpha(w) \rangle = \langle \alpha^*(v) \mid w \rangle = \langle \overline{\lambda} v \mid w \rangle = \lambda \langle v \mid w \rangle$$

(la troisième égalité en vertu du lemme précédent), donc $\langle v \mid w \rangle = 0$ si $\lambda \neq \mu$.

(ii) Pour α autoadjoint, si $\alpha(v) = \lambda v$ avec $v \neq 0$, alors $\alpha(v) = \overline{\lambda} v$ par le lemme.

(iii) Pour α unitaire, si $\alpha(v) = \lambda v$ avec $v \neq 0$, on a

$$\|v\| \;=\; \|\alpha(v)\| \;=\; \|\lambda v\| \;=\; |\lambda|\;\|v\|$$

et par suite $|\lambda| = 1$. □

4. Opérateurs linéaires agissant sur les espaces euclidiens

Le contenu du § VI.3 passe sans aucune difficulté aux opérateurs linéaires agissant sur les espaces euclidiens, c'est-à-dire sur les espaces vectoriels *réels* munis de produits scalaires. Notons en particulier les énoncés suivants.

Soient V, W deux espaces euclidiens, respectivement munis de bases orthonormales $\mathcal{E}, \mathcal{F}$, et $\alpha \in \mathcal{L}(V, W)$. La matrice $M_{\mathcal{E},\mathcal{F}}(\alpha^*)$ de l'opérateur $\alpha^* \in \mathcal{L}(W, V)$ adjoint de α est la matrice transposée de $M_{\mathcal{F},\mathcal{E}}(\alpha)$. Voir la proposition 3.1

Pour un opérateur $\alpha \in \mathcal{L}(V)$ et une base orthonormale $\mathcal{E}$ de V, les trois conditions suivantes sont équivalentes :

(i) *la forme bilinéaire $V \times V \longrightarrow \mathbb{R}$, $(v, w) \longmapsto \langle v \mid \alpha(w)\rangle$ est symétrique,*

(ii) $\alpha^* = \alpha$,

(iii) *la matrice $M_{\mathcal{E}}(\alpha)$ est symétrique.*

Voir la proposition 3.2 et l'énoncé suivant la définition 3.2. Un opérateur $\alpha \in \mathcal{L}(V)$ est autoadjoint si ces conditions sont satisfaites.

Toute valeur propre (a priori complexe) d'une matrice réelle symétrique est réelle.

Voir néanmoins la remarque de la page 124.

Il y a de minimes changements terminologiques. Par exemple, on dit parfois que α^* est l'opérateur « transposé » (au lieu de « adjoint ») de α. Un endomorphisme linéaire d'un espace euclidien qui préserve les normes[6] est dit *orthogonal* (au lieu de unitaire) ; on parle de même du *groupe orthogonal* $O(V)$ d'un espace euclidien V (analogue du « groupe unitaire » d'un espace hermitien). On laisse au lecteur le soin de vérifier que le déterminant fournit un homomorphisme de groupes

$$O(V) \;\longrightarrow \{1, -1\}$$

qui est surjectif ; son noyau est le *groupe spécial orthogonal* $SO(V)$ de l'espace euclidien V ; dans le cas où V est l'espace $\mathbb{R}^n$ muni du produit scalaire canonique, ces groupes sont le plus souvent notés $O(n)$ et $SO(n)$.

5. Le théorème spectral

Lemme. *Soient V un espace hermitien et $\alpha : V \longrightarrow V$ un opérateur linéaire.*

(i) *Soit U un sous-espace de V. Alors U est invariant par α si et seulement si son orthogonal $U^\perp$ est invariant par α^*.*

(ii) *On suppose α normal ; soit e_1 un vecteur propre de α. Alors la restriction de α à l'orthogonal V_1 de e_1 est un opérateur normal.*

[6] Pour un opérateur $\alpha \in \mathcal{L}(V)$, les quatre propriétés suivantes sont équivalentes : (i) α préserve les normes, (ii) α préserve les produits scalaires, (iii) $\alpha^*\alpha = \mathrm{id}_V$, (iv) $\alpha\alpha^* = \mathrm{id}_V$.

PREUVE. (i) Le sous-espace U est invariant par α, c'est-à-dire $\alpha(U) \subset U$, si et seulement si $\langle \alpha u \mid w \rangle = 0$ pour tous $u \in U$ et $w \in U^\perp$. Cette condition est évidemment équivalente à : $\langle u \mid \alpha^*(w) \rangle = 0$ pour tous $u \in U$ et $w \in U^\perp$, c'est-à-dire à : $U^\perp$ est invariant par α^*.

(ii) Il résulte du lemme précédant le théorème VI.5 que la droite $\mathbb{C}e_1$ est invariante par α et par α^*, et de l'assertion (i) que V_1 est également invariant par α et α^*. Notons $\alpha_1 \in \mathcal{L}(V_1)$ la restriction de α à V_1. Pour $v, v' \in V_1$, nous avons

$$\langle v \mid \alpha_1(v') \rangle_{V_1} = \langle v \mid \alpha(v') \rangle_V = \langle \alpha^*(v) \mid v' \rangle_V = \langle \alpha_1^*(v) \mid v' \rangle_{V_1}$$

de sorte que l'adjoint α_1^* de α_1 est la restriction de α^* à V_1. L'égalité $\alpha\alpha^* = \alpha^*\alpha$ implique donc $\alpha_1\alpha_1^* = \alpha_1^*\alpha_1$. □

Attention : En général, pour un opérateur non normal $\alpha \in \mathcal{L}(V)$ laissant invariant un sous-espace U de V, l'espace U *n'est pas* invariant par α^*. Exemple : $V = \mathbb{C}^2$, U le sous-espace des vecteurs dont la seconde coordonnée est nulle, est α l'opérateur de matrice $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ relativement à la base canonique.

THÉORÈME VI.6 (théorème spectral). *Tout opérateur normal α sur un espace hermitien V est diagonalisable relativement à une base orthonormale de V.*

PREUVE, PAR RÉCURRENCE SUR LA DIMENSION n DE V. Le théorème est évident si $n = 1$. On suppose désormais $n \geq 2$, et le théorème démontré pour tout espace de dimension strictement inférieure à n.

On peut choisir un vecteur propre e_1 de α de norme 1. L'espace $\mathbb{C}e_1$ est invariant par α^* (lemme du § VI.3), donc l'orthogonal $(\mathbb{C}e_1)^\perp$ est invariant par α (lemme précédent). On vérifie que la restriction de α à $(\mathbb{C}e_1)^\perp$ est un opérateur normal, et on achève en appliquant l'hypothèse de récurrence à cette restriction. □

REMARQUES. (i) Soit V un espace vectoriel complexe de dimension finie et soit α un opérateur linéaire sur V. Il a été vu au chapitre précédent que α est trigonalisable, c'est-à-dire qu'il existe une base de V relativement à laquelle la matrice de α est triangulaire.

EXERCICE. Vérifier que la même preuve montre que, si de plus V est muni d'un produit scalaire, alors il existe une base *orthonormale* de V relativement à laquelle la matrice de α est triangulaire.

(ii) En général, un opérateur linéaire α sur un espace vectoriel V *n'est pas* diagonalisable (exemple : un opérateur non nul de carré nul). Lorsque V est muni d'un produit scalaire, même si α est diagonalisable, il n'existe en général pas de base *orthonormale* relativement à laquelle la matrice de α est diagonale (exemple : une projection oblique).

Lorsque V possède un produit scalaire, un opérateur linéaire α sur V est diagonalisable relativement à une base orthonormale si et seulement s'il est normal ; l'une des implications est le théorème spectral énoncé ci-dessus, et la vérification de l'autre est un exercice facile (toute matrice diagonale est normale).

(iii) Dans la preuve ci-dessus du théorème spectral, par récurrence sur la dimension de l'espace, le début du raisonnement invoque l'existence d'un vecteur propre, justifiée dans un chapitre précédent par le "théorème fondamental de l'algèbre". Pour le cas particulier des opérateurs normaux, voici un autre argument montrant qu'il existe un vecteur propre.

PROPOSITION. *Soient V un espace hermitien et $\alpha \in \mathcal{L}(V)$ un opérateur normal. Soient $\mathbb{S}(V) = \{v \in V | \langle v|v\rangle = 1\}$ la sphère unité de V et λ_M le maximum de la fonction $\varphi : \mathbb{S}(V) \longrightarrow \mathbb{R}$, $\varphi(v) = \langle v|\alpha(v)\rangle$. Alors λ_M est une valeur propre de α et tout vecteur $v \in \mathbb{S}(V)$ où φ atteint son maximum est un vecteur propre de α correspondant à la valeuer propre λ_M.*

IDÉE DE LA PREUVE. La sphère $\mathbb{S}(V)$ étant compacte, la fonction continue φ y atteint son maximum. En appliquant la méthode des multiplicateurs de Lagrange, on vérifie que tout maximum v de φ sur $\mathbb{S}(V)$ est un vecteur propre de valeur propre λ_M. Pour une introduction à la méthode de Lagrange, voir par exemple le chapitre IV du livre de E. Hairer et G. Wanner, *Analysis by its history*, Springer, 1996. □

COROLLAIRE. *Soit a une matrice complexe à n lignes et n colonnes. Si a est normale, il existe une matrice unitaire $u \in \mathcal{U}(n)$ telle que la matrice u^*au est diagonale. Si la matrice a est de plus autoadjointe [respectivement unitaire], les coefficients diagonaux de uau^* sont réels [resp. de module 1].*

PREUVE. On peut considérer la matrice a comme un opérateur sur l'espace $\mathbb{C}^n$ muni de sa base canonique $\mathcal{C}an = (e_1, \ldots, e_n)$. Soit par ailleurs $\mathcal{B} = (b_1, \ldots, b_n)$ une base orthonormale de $\mathbb{C}^n$ constituée de vecteurs propres de a. [L'existence de $\mathcal{B}$ est assurée par le théorème spectral.] Soit u la matrice dont les colonnes sont les coordonnées relatives à $\mathcal{C}an$ des vecteurs de la base $\mathcal{B}$; ainsi u est unitaire, et c'est la matrice $M_{\mathcal{C}an,\mathcal{B}}(id)$ de changement de bases, de $\mathcal{B}$ vers $\mathcal{C}an$. Il résulte des définitions que la matrice $u^{-1}au$ est diagonale. Les assertions de la dernière phrase du corollaire sont alors immédiates. □

COROLLAIRE. *Soit λ une valeur propre d'un opérateur normal sur un espace hermitien. Alors la mutliplicité géométrique de λ coïncide avec sa multiplicité algébrique.*

PREUVE. C'est le cas des matrices diagonales. □

COROLLAIRE. *Soit α un opérateur normal.*

Pour que α soit autoadjoint, il faut et il suffit que toutes les valeurs propres de α soient réelles.

Pour que α soit unitaire, il faut et il suffit que toute les valeurs propres de α soient de module 1.

6. Cas euclidien

Nous retiendrons ici l'énoncé suivant, dont l'importance ne peut pas être surestimée !

THÉORÈME VI.7 (théorème spectral pour les opérateurs autoadjoints[7] réels et les matrices symétriques réelles). *Soient V un espace euclidien et α un opérateur autoadjoint sur V. Alors toutes les valeurs propres de α sont réelles et V possède une base orthonormale de vecteurs propres de α ; en d'autres termes, relativement à une telle base orthonormale $\mathcal{E}$, la matrice $M_{\mathcal{E}}(\alpha)$ de α est une matrice diagonale réelle.*

[7] Pour les opérateurs sur un espace euclidien, on dit aussi «symétrique» au lieu de «autoadjoint» .

Soient n un nombre entier strictement positif et $a \in M_n(\mathbb{R})$ une matrice symétrique *réelle. Alors il existe une matrice orthogonale $u \in O(n)$ telle que la matrice uau^{-1} est diagonale réelle.*

Preuve directe de l'assertion "toute valeur propre d'une matrice symétrique réelle est réelle". Soit $a \in M_n(\mathbb{R})$ une matrice symétrique, λ une valeur propre de a, a priori complexe, et $z \in M_{1,n}(\mathbb{C}) = \mathbb{C}^n$ un vecteur propre, $a(z) = \lambda z$. Notons ${}^t\overline{z} \in M_{n,1}(\mathbb{C})$ la matrice conjuguée transposée de z. Le produit matriciel ${}^t\overline{z}z$ est le nombre $\sum_{j=1}^n |z_j|^2$, et s'identifie donc au produit scalaire $\langle z|z\rangle$ pour la structure hermitienne canonique de $\mathbb{C}^n$; notons que ce nombre est strictement positif puisque $z \neq 0$. D'une part

$$ {}^t\overline{z}az \;=\; {}^t\overline{z}\lambda z \;=\; \lambda \sum_{j=1}^{n} |z|^2 \,. $$

D'autre part, comme ${}^t\overline{a} = a$ (car la matrice a est réelle symétrique),

$$ {}^t\overline{z}az \;=\; {}^t\overline{z}{}^t\overline{a}z \;=\; {}^t\overline{(az)}z \;=\; {}^t\overline{(\lambda z)}z \;=\; \overline{\lambda} \sum_{j=1}^{n} |z|^2 \,. $$

Il en résulte que $\overline{\lambda} = \lambda$. □

Exemple 1. Etant donné un entier $n \geq 1$, notons V_n l'espace vectoriel réel des fonctions $[-\pi, \pi] \longrightarrow \mathbb{R}$ du type $t \longmapsto \sum_{j=1}^n c_j \sin(jt)$, avec $c_1, \ldots, c_n \in \mathbb{R}$. Considérons l'opérateur linéaire $\alpha \in \mathcal{L}(V_n)$ défini par $\alpha(f) = -f''$ (moins la dérivée seconde de f).

Pour $j \in \{1, \ldots, n\}$, notons $e_j \in V_n$ la fonction définie par $e_j(t) = \sin(jt)$. Alors $\alpha(e_j) = j^2 e_j$ (règles de dérivation des sinus et cosinus), donc $e_1, \ldots, e_n$ sont des vecteurs propres de α de valeurs propres respectivement $1, 4, 9, \ldots, n^2$; en particulier, les vecteurs e_j sont linéairement indépendants. De plus, ces vecteurs engendrent V_n par définition, donc ils constituent une *base de V_n*. Définissons sur V_n une forme bilinéaire symétrique par

$$ \langle f \mid g \rangle \;=\; \frac{1}{\pi} \int_{-\pi}^{\pi} f(t)g(t)dt. $$

En utilisant une identité du type $\sin x \sin y = -\frac{1}{2}\cos(x+y) + \frac{1}{2}\cos(x-y)$ et les primitives de $\sin(\cdot)$ et $\cos(\cdot)$, on calcule sans peine

$$ \langle e_j \mid e_k \rangle \;=\; \begin{cases} 0 & \text{si} \quad j \neq k \\ 1 & \text{si} \quad j = k \end{cases} $$

(d'où une *autre* preuve de l'indépendance linéaire des vecteurs e_j). Il en résulte que, si $f(t) = \sum_{j=1}^n c_j \sin(jt)$, alors $\langle f \mid f \rangle = \sum_{j=1}^n (c_j)^2$; par suite, la forme bilinéaire symétrique sur V_n définie ci-dessus est *définie positive*, de sorte que V_n est un *espace euclidien.*

Notons que les fonctions de V_n sont toutes nulles en $-\pi$ et en π ; pour $f, g \in V_n$, deux intégrations par parties montrent donc que

$$ \begin{aligned} \langle \alpha(f) \mid g \rangle \;&=\; \frac{-1}{\pi} \int_{-\pi}^{\pi} f''(t)g(t)dt \;=\; \frac{-1}{\pi} \left\{ |f'g|_{-\pi}^{\pi} - \int_{-\pi}^{\pi} f'(t)g'(t)dt \right\} \\ &=\; \frac{1}{\pi} \int_{-\pi}^{\pi} f'(t)g'(t)dt \;=\; \frac{1}{\pi} \left\{ |fg'|_{-\pi}^{\pi} - \int_{-\pi}^{\pi} f(t)g''(t)dt \right\} \\ &=\; \langle f \mid \alpha(g) \rangle. \end{aligned} $$

Ainsi

> *α est un opérateur audoadjoint sur V_n, et les e_j constituent une base orthonormale de V_n formée de vecteurs propres de α.*

L'étude de l'opérateur[8] α et de ses vecteurs propres est essentiellement équivalente à celle des solutions f non identiquement nulles de l'équation différentielle

$$f'' + \lambda f \;=\; 0$$

pour lesquelles

$$f(-\pi) \;=\; f(\pi) \;=\; 0.$$

De telles solutions n'existent que si λ (a priori un nombre réel) est le carré d'un entier > 0.

Variante et généralisation. Soit $[a,b] \subset \mathbb{R}$ un intervalle compact $(-\infty < a < b < \infty)$; on considère des fonctions à valeurs réelles p, q, ρ sur $[a,b]$, avec p et ρ à valeurs strictement positives et q continûment différentiable. L'opérateur que nous allons définir permet d'étudier certains problèmes où interviennent l'équation différentielle

$$(*) \qquad \frac{d}{dt}\left[p(t)\frac{df}{dt}\right] - q(t)f(t) + \lambda\rho(t)f(t) \;=\; 0$$

et les conditions aux limites

$$(**) \qquad f(a) \;=\; f(b) \;=\; 0$$

(on retrouve le cas précédent lorsque $p(t) = \rho(t) = 1$ et $q(t) = 0$ pour tout $t \in [a,b]$). Les conditions $(*)$ et $(**)$ constituent un *problème de Sturm*[9]*-Liouville*[10], et on peut montrer qu'il n'existe de solutions non nulles que pour une suite de valeurs du paramètre λ, ces valeurs étant précisément les « valeurs propres » d'un opérateur α que nous allons définir.

Soit V l'espace vectoriel des fonctions de $[a,b]$ dans $\mathbb{R}$ qui sont deux fois continûment différentiables et dont les valeurs en a et b sont nulles. Cet espace est muni du produit scalaire défini par

$$\langle f \mid g \rangle \;=\; \int_a^b f(t)g(t)p(t)dt.$$

Soit $\alpha : V \longrightarrow V$ l'application linéaire définie par

$$\alpha(f) \;=\; -\frac{d}{dt}\left[p(t)\frac{df}{dt}\right] + q(t)f(t).$$

En utilisant l'identité

$$-\frac{d}{dt}\left\{\left[f(t)\frac{dg}{dt} - \frac{df}{dt}g(t)\right]p(t)\right\} \;=\; f(t)\big(\alpha(g)\big)(t) - g(t)\big(\alpha(f)\big)(t),$$

il est facile de vérifier que

$$\langle f \mid \alpha(g) \rangle \;=\; \langle g \mid \alpha(f) \rangle.$$

---

[8] Ou, plus précisément, d'une « limite » appropriée pour $n \to \infty$ des opérateurs $\alpha_n = \alpha \in \mathcal{L}(V_n)$.

[9] Charles Sturm, né à Genève en 1803, établi dès 1825 à Paris où il meurt en 1855 ; un congrès a été organisé à Genève du 15 au 19 septembre 2003 à l'occasion du 200ème anniversaire de sa naissance ; voir

`http ://theory.physics.unige.ch/sturm.html`

[10] Joseph Liouville, mathématicien français (1809–1882) ; on lui doit de nombreuses contributions à l'analyse. Le « théorème de Liouville » du cours d'Analyse II (une fonction entière bornée est constante) semble néanmoins dû à Cauchy, et la preuve qu'on en donne le plus souvent due à Jordan ! On doit aussi à Liouville la notion et la preuve de l'existence de *nombres transcendants*. Ainsi, Liouville a par exemple montré que le nombre $\sum_{n=1}^{\infty} 10^{-n!}$ est transcendant (1844). Pour les premières preuves de transcendance de nombres plus célèbres, il fallut attendre Hermite (preuve de la transcendance de e en 1873) et Lindemann (preuve de la transcendance de π en 1882, et par là la résolution d'un problème ouvert depuis plusieurs millénaires !).

On en déduit que les «valeurs propres» du problème de Sturm-Liouville sont toutes réelles.

De nombreux cas particuliers interviennent dans des problèmes physiques, par exemple dans l'étude des vibrations d'un tambour. (Pour une introduction aux «problèmes de Sturm-Liouville», voir par exemple le chapitre 10 de *Ordinary differential equations, 3rd,* de G. Birkhoff et G.-C. Rota, J. Wiley, 1978.)

EXEMPLE 2. Soient V un espace euclidien de dimension n et $\alpha \in \mathcal{L}(V)$ un opérateur autoadjoint inversible. On définit une «hypersurface»

$$S = \{v \in V \mid \langle v \mid \alpha(v)\rangle = 1\}$$

(courbe si $n = 2$, surface si $n = 3$). Notons que les hypersurfaces associées à α et à $\frac{1}{2}(\alpha + \alpha^*)$ son identiques. Sans restreindre la généralité, on suppose donc désormais que α est autoadjoint ; pour simplifier, supposons *de plus* que α est inversible. Le problème consiste à décrire qualitativement S.

Soit $\mathcal{E} = (e_1, \ldots, e_n)$ une base orthonormale de vecteurs propres de α, avec valeurs propres correspondantes $\lambda_1, \ldots, \lambda_n$ (et les λ_j sont non nuls, puisque α est inversible). Dans les coordonnées correspondant à $\mathcal{E}$, l'équation de la surface s'écrit

$$\sum_{j=1}^{n} \lambda_j x_j^2 = 1.$$

On peut de plus supposer la numérotation telle que

$$\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n.$$

De plus, on va supposer que $\lambda_1 > 0$, pour que S ne soit pas vide. Si $n = 2$, la courbe est alors une *cônique*, il y a deux cas à distinguer :

(i) si $\lambda_1 \geq \lambda_2 > 0$, alors S est une ellipse ;

(ii) si $\lambda_1 > 0 > \lambda_2$, alors S est une hyperbole.

[Le cas $\lambda_1 > \lambda_2 = 0$ fournirait une paire de droites parallèles.]

Si $n = 3$, la surface est alros une *quadrique*, il y a trois cas à distinguer :

(i) si $\lambda_1 \geq \lambda_2 \geq \lambda_3 > 0$, alors S est un ellipsoïde ;

(ii) si $\lambda_1 \geq \lambda_2 > 0 > \lambda_3$, alors S est un hyperboloïde à une nappe ;

(iii) si $\lambda_1 > 0 > \lambda_2 \geq \lambda_3$, alors S est un hyperboloïde à deux nappes.

EXERCICE. Dessiner un ellipsoïde, un hyperboloïde à une nappe, un hyperboloïde à deux nappes.

EXEMPLE 3. On considère le système différentiel

$$\begin{cases} \dfrac{dx}{dt} = -2x + y \\ \dfrac{dy}{dt} = x - 2y \end{cases} \qquad \text{c'est-à-dire} \qquad \begin{pmatrix} \frac{dx}{dt} \\ \frac{dy}{dt} \end{pmatrix} = \begin{pmatrix} -2 & 1 \\ 1 & -2 \end{pmatrix} \begin{pmatrix} x(t) \\ y(t) \end{pmatrix}$$

où $x(t)$, $y(t)$ sont des fonctions différentiables inconnues de $\mathbb{R}$ dans $\mathbb{R}$.

On calcule sans peine les valeurs et vecteurs propres de la matrice $a = \begin{pmatrix} -2 & 1 \\ 1 & -2 \end{pmatrix}$, donnés par

$$\begin{pmatrix} -2 & 1 \\ 1 & -2 \end{pmatrix} \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix} = - \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix} \text{ et } \begin{pmatrix} -2 & 1 \\ 1 & -2 \end{pmatrix} \begin{pmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{pmatrix} = -3 \begin{pmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{pmatrix}.$$

Si on pose

$$s = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{pmatrix},$$

(remarque : s est unitaire et $s^{-1} = s$), on calcule

$$s^{-1}as = \begin{pmatrix} -1 & 0 \\ 0 & -3 \end{pmatrix}.$$

Posons alors

$$\begin{pmatrix} u(t) \\ v(t) \end{pmatrix} = s^{-1} \begin{pmatrix} x(t) \\ y(t) \end{pmatrix}.$$

On trouve

$$\begin{pmatrix} \frac{du}{dt} \\ \frac{dv}{dt} \end{pmatrix} = s^{-1} \begin{pmatrix} \frac{dx}{dt} \\ \frac{dy}{dt} \end{pmatrix} = s^{-1}a \begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = s^{-1}as \begin{pmatrix} u(t) \\ v(t) \end{pmatrix}.$$

Ce dernier système s'écrit aussi

$$\begin{cases} \dfrac{du}{dt} = -u \\ \dfrac{dv}{dt} = -3v \end{cases}$$

et s'intègre donc à vue :

$$u(t) = C_1 e^{-t} \qquad \text{et} \qquad v(t) = C_2 e^{-3t}$$

où C_1, C_2 sont des constantes d'intégration. *Le point important ici est que, en diagonalisant la matrice a, on a réduit le système initial à deux équations indépendantes l'une de l'autre.* Les solutions du système initial sont donc

$$\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = s \begin{pmatrix} u(t) \\ v(t) \end{pmatrix} = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{pmatrix} \begin{pmatrix} C_1 e^{-t} \\ C_2 e^{-3t} \end{pmatrix}$$

ou encore

$$\begin{cases} x(t) = c_1 e^{-t} + c_2 e^{-3t} \\ y(t) = c_1 e^{-t} - c_2 e^{-3t} \end{cases}$$

avec $c_j = \frac{1}{\sqrt{2}} C_j$, $j = 1, 2$ des constantes d'intégration.

Il est *vital* de *vérifier* que cette solution satisfait bien le système original, ce dont nous laissons le soin au lecteur.

7. Classification des endomorphismes linéaires d'un espace réel de dimension deux

PROPOSITION. *Soit V un espace vectoriel réel de dimension 2 et $\alpha \in \mathcal{L}(V)$ un endomorphisme linéaire. Alors α est d'un et d'un seul des trois types suivants.*

(1) Il existe deux nombres réels λ, μ et une base $\mathcal{B}$ de V tels que

$$M_{\mathcal{B}}(\alpha) = \begin{pmatrix} \lambda & 0 \\ 0 & \mu \end{pmatrix}.$$

(2) Il existe un nombre réel λ et une base $\mathcal{B}$ de V tels que

$$M_{\mathcal{B}}(\alpha) = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}.$$

(3) Il existe deux nombres réels ρ, σ, $\sigma \neq 0$, et une base $\mathcal{B}$ de V tels que

$$M_{\mathcal{B}}(\alpha) = \begin{pmatrix} \rho & \sigma \\ -\sigma & \rho \end{pmatrix}.$$

Attention : Même si les matrices de la forme (2) sont symétriques, il n'y aurait aucun sens à écrire que α est hermitien (ou symétrique), car V n'est pas supposé possèder de structure euclidienne !

PREUVE. Soit p_α le polynôme caractéristique de α. Il y a trois cas à distinguer.

(a) Le polynôme p_α possède deux racines réelles distinctes λ, μ. Soit $\mathcal{B}$ une base de V constituée d'un vecteur propre de α pour la valeur propre λ et d'un vecteur propre de α pour la valeur propre μ. Alors $M_{\mathcal{B}}(\alpha)$ est de type (1), avec de plus $\lambda \neq \mu$.

(b) Le polynôme possède une racine rélle λ de multiplicité deux.

Premier sous-cas : la multiplicité géométrique de la valeur propre λ de α est deux. Alors $\alpha = \lambda \,\mathrm{id}_V$ et $M_{\mathcal{B}}(\alpha)$ est de type (1) pour toute base $\mathcal{B}$ de V, avec $\lambda = \mu$.

Second sous-cas : la multiplicité géométrique de la valeur propre λ de α est un. Soit $\mathcal{B}' = \{e, f'\}$ une base de V telle que le vecteur e soit un vecteur propre de α pour la valeur propre λ. L'équation $\alpha(e) = \lambda e$ se traduit en $M_{\mathcal{B}'}(\alpha) = \begin{pmatrix} \lambda & x \\ 0 & y \end{pmatrix}$, l'équation $p_\alpha(t) = \det \begin{pmatrix} t-\lambda & -x \\ 0 & t-y \end{pmatrix}$ montre que $y = \lambda$, et l'hypothèse sur la multiplicité géométrique de la valeur propre implique que $x \neq 0$. Notons que $\alpha(f') = xe + \lambda f'$.

Posons $f = \frac{1}{x} f'$ et $\mathcal{B} = (e, f)$. Alors $\alpha(e) = \lambda e$ et $\alpha(f) = e + \lambda f$, de sorte que $M_{\mathcal{B}}(\alpha)$ est bien comme dans (2).

(c) Le polynôme p_α possède une valeur propre complexe $\lambda = \rho + i\sigma \in \mathbb{C}$, avec $\rho, \sigma \in \mathbb{R}$ et $\sigma \neq 0$. Introduisons une base *arbitraire* $\mathcal{B}'$ de V qui nous permet d'identifier V à $\mathbb{R}^2$

et α à une matrice $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in M_2(\mathbb{R})$. Soit $x+iy \in \mathbb{C}^2$ un vecteur propre de valeur propre λ, avec $x, y \in \mathbb{R}^2$. On a donc

$$(*) \qquad \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x_1+iy_1 \\ x_2+iy_2 \end{pmatrix} = (\rho+i\sigma) \begin{pmatrix} x_1+iy_1 \\ x_2+iy_2 \end{pmatrix}.$$

En égalant les complexes conjugués des deux membres de (*), on obtient

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x_1-iy_1 \\ x_2-iy_2 \end{pmatrix} = (\rho-i\sigma) \begin{pmatrix} x_1-iy_1 \\ x_2-iy_2 \end{pmatrix},$$

ce qui montre que $\overline{\lambda} = \rho - \sigma$ est aussi une valeur propre de a. Comme λ et $\overline{\lambda}$ sont distincts ($\beta \neq 0$), les vecteurs $x+iy$ et $x-iy$ de $\mathbb{C}^2$ sont $\mathbb{C}$-linéairement indépendants. Il en résulte que les vecteurs x, y sont aussi $\mathbb{C}$-linéairement indépendants, et donc *a fortiori* $\mathbb{R}$-linéairement indépendants. Dans l'équation (*), l'égalité des parties réelles s'écrit

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \rho \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} - \sigma \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \qquad \text{ou} \qquad \alpha(x) = \rho x - \sigma y$$

et l'égalité des parties imaginaires

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \sigma \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \rho \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \qquad \text{ou} \qquad \alpha(y) = \sigma x + \rho y.$$

Il suffit donc de poser $\mathcal{B} = \{x, y\}$. □ □

8. Un exemple historique de calcul de valeurs propres et vecteurs propres pour une matrice symétrique réelle

L'exemple qui suit, dû à Lagrange, est une illustration du théorème spectral pour les matrices réelles symétriques. Historiquement, c'est sans doute même la première manifestation écrite du théorème spectral pour des matrices d'ordre quelconque (par opposition à des matrices 2-fois-2 ou 3-fois-3).

En 1759, Lagrange[11] s'intéresse à l'équation des ondes décrivant, par exemple, la propagation du son dans un tuyau d'orgue. Pour décrire les vibrations d'un milieu continu, il imagine un modèle faisant intervenir un *très grand nombre* n d'oscillateurs harmoniques couplés.

Pour chaque $j \in \{1, 2, \ldots, n\}$, la $j$ème masse est à distance $x_j(t)$ de sa position d'équilibre ; elle est soumise à la force du ressort de gauche, proportionnelle à $-(x_j(t)-x_{j-1}(t))$, et à celle du ressort de droite, proportionnelle à $+(x_{j+1}(t)-x_j(t))$. Les équations de Newton fournissent donc le système différentiel linéaire à coefficients constants

$$(*) \qquad \ddot{x}_j(t) = K\,(x_{j-1}(t) - 2x_j(t) + x_{j+1}(t)) \qquad j = 1, \ldots, n$$

où $x_0(t) = x_{n+1}(t) = 0$ pour tout t et où K est une constante positive convenable. Le modèle est suffisamment robuste pour décrire aussi bien des petits mouvements dans

[11] Voir *Recherches sur la nature et la propagation du son,* Oeuvres 1, pp. 39-148. En 1760-61, Lagrange publie aussi *Nouvelles recherches sur la nature et la propagation du son,* Ibid., pp. 151-332. Joseph-Louis de Lagrange, né à Turin en 1736 et mort à Paris en 1813 a travaillé notamment sur des problèmes de théorie des nombres (tout entier est somme de quatre carrés), de mécanique (fondements de la mécanique analytique), de calcul des variations (équations d'Euler-Lagrange), et d'équations différentielles. Depuis 1755 (sic), il est professeur de mathématiques à l'Ecole royale d'Artillerie de Turin. Il sera directeur de la section mathématique de l'Académie de Berlin en 1766 et pensionnaire de l'Académie des sciences de Paris en 1787.

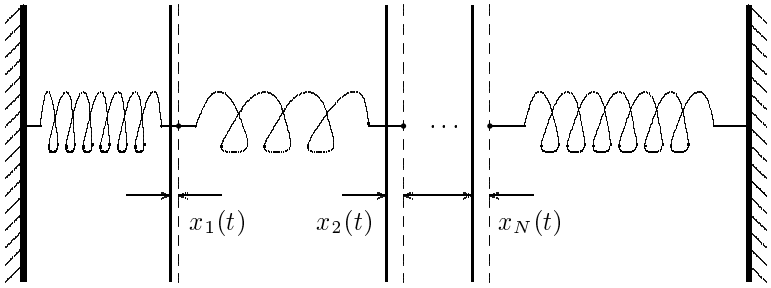

FIG. 1

la direction d'alignement des oscillateurs (molécules d'air dans un tuyau d'orgue) que dans la direction transverse (éléments de corde d'un violon).

Si $n = 1$, «on sait bien» que les solutions de l'équation différentielle $\ddot{y}(t) + 2Ky(t) = 0$ sont de la forme $y(t) = \text{cste} \cos\left(\sqrt{2K}t + t_0\right)$. Il est donc raisonnable de chercher des solutions de la forme

$$y_j(t) \;=\; x_j \cos\left(\omega t + t_0\right).$$

En insérant dans $(*)$ et en simplifiant par $\cos(\omega t + t_0)$, on trouve

$$-\omega^2 x_j \;=\; K\left(x_{j-1} - 2x_j + x_{j+1}\right).$$

En d'autres termes, si on définit la matrice

$$a \;=\; \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 & 0 \\ 1 & 0 & 1 & \cdots & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & 0 & \cdots & 1 & 0 \end{pmatrix}$$

alors le vecteur x de coordonnées $x_1, \ldots, x_n$ satisfait l'équation

$$(**) \qquad ax \;=\; \left(2 \,-\, \frac{\omega^2}{K}\right) x$$

et il s'agit de calculer les valeurs propres et vecteurs propres de a.

Rappelons les formules d'addition pour les fonctions trigonométriques :

$$\sin(\alpha + \beta) \;=\; \sin\alpha\,\cos\beta \;+\; \cos\alpha\,\sin\beta$$
$$\sin(\alpha - \beta) \;+\; \sin(\alpha + \beta) \;=\; 2\,\cos\beta\,\sin\alpha.$$

On a donc, pour tout $\ell \in \{1, \ldots, n\}$:

$$\begin{pmatrix} 0 & 1 & 0 & \cdots & 0 & 0 \\ 1 & 0 & 1 & \cdots & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & 0 & \cdots & 1 & 0 \end{pmatrix} \begin{pmatrix} \sin\frac{\ell\pi}{n+1} \\ \sin\frac{2\ell\pi}{n+1} \\ \sin\frac{3\ell\pi}{n+1} \\ \vdots \\ \sin\frac{(n-1)\ell\pi}{n+1} \\ \sin\frac{n\ell\pi}{n+1} \end{pmatrix} = 2\cos\left(\frac{\ell\pi}{n+1}\right) \begin{pmatrix} \sin\frac{\ell\pi}{n+1} \\ \sin\frac{2\ell\pi}{n+1} \\ \sin\frac{3\ell\pi}{n+1} \\ \vdots \\ \sin\frac{(n-1)\ell\pi}{n+1} \\ \sin\frac{n\ell\pi}{n+1} \end{pmatrix}.$$

Notons encore que $2 - \frac{\omega^2}{K}$ est l'une des valeurs propres $2\cos\left(\frac{\ell\pi}{n+1}\right)$, pour $\ell \in \{1, \dots, n\}$, si $\omega^2 = 2K\left(1 - \cos\left(\frac{\ell\pi}{n+1}\right)\right) = K\left(2\sin\left(\frac{\ell\pi}{2(n+1)}\right)\right)^2$, ou encore si

$$\omega \;=\; 2\sqrt{K}\sin\left(\frac{\ell\pi}{2(n+1)}\right).$$

On a donc les résultats suivants.

(i) *Les valeurs propres de la matrice a sont*

$$2\cos\frac{\pi}{n+1}\,,\; 2\,\cos\frac{2\pi}{n+1}\,,\;\dots\;,\; 2\,\cos\frac{n\pi}{n+1}$$

et sont en particulier toutes simples.

(ii) *Pour tout $\ell \in \{1, 2, \dots, n\}$, le système $(*)$ possède des solutions de la forme*

$$y_j(t) \;=\; \sin\left(\frac{j\ell\pi}{n+1}\right)\cos\left(2\sqrt{K}\sin\left(\frac{\ell\pi}{2(n+1)}\right)\,t \;+\; t_\ell\right).$$

Lorsque $\ell = 1$ (son fondamental), les amplitudes maximales $\sin\left(\frac{j\ell\pi}{n+1}\right)$ dessinent en fonction de j une courbe ayant un seul maximum, pour $j = \frac{n}{2}$. Pour $\ell = 2, 3, \dots$, on obtient des courbes avec 1, 2, ... zéros entre 1 et n ; ces courbes correspondent par ailleurs à des fréquences proportionnelles aux nombres $\sin\left(\frac{\ell\pi}{2(n+1)}\right)$, c'est-à-dire (puisque n est très grand !) à des nombres en progression arithmétique, c'est-à-dire encore aux harmoniques à l'octave, la quinte supérieure, la deuxième octave, ... du son fondamental.

(iii) *Par linéarité, les combinaisons linéaires des solutions ci-dessus sont encore des solutions du système $(*)$.* Il en résulte que *la solution générale du système $(*)$ est de la forme*

$$x_j(t) \;=\; \sum_{\ell=1}^{n} C_\ell \sin\left(\frac{j\ell\pi}{n+1}\right)\cos\left(2\sqrt{K}\sin\left(\frac{\ell\pi}{2(n+1)}\right)\,t \;+\; t_\ell\right)$$

où les C_ℓ et les t_ℓ sont $2n$ constantes d'intégration. La description d'un son musical met en jeux une grande quantité d'harmoniques.

Remarques. (i) On pourrait vérifier explicitement que les vecteurs propres de a écrits ci-dessus constituent une base orthogonale de $\mathbb{R}^n$. (Pour obtenir une base *orthonormale*, il faudrait encore les normaliser, ce qui n'offre guère d'intérêt.)

(ii) Pour trouver les valeurs propres de a (mais pas ses vecteurs propres), on peut aussi procéder comme suit. Si $P_n(\lambda) = \det(\lambda I_n - a)$ désigne le polynôme caractéristique de a pour n lignes et n colonnes, on peut développer le déterminant relativement à sa première ligne et on trouve

$$P_{n+1}(\lambda) \;=\; \lambda P_n(\lambda) - P_{n-1}(\lambda).$$

On a par ailleurs $P_1(\lambda) = \lambda$ et $P_2(\lambda) = \lambda^2 - 1$ (et encore, si on veut, $P_0(\lambda) = 1$). On vérifie alors par récurrence sur n que

$$(***) \qquad P_n(\lambda) \;=\; \frac{\tau^{n+1} - \tau^{-n-1}}{\tau - \tau^{-1}}$$

où

$$\tau \;=\; \frac{\lambda+\sqrt{\lambda^2-4}}{2} \quad \text{et} \quad \tau^{-1} \;=\; \frac{\lambda-\sqrt{\lambda^2-4}}{2}$$

et donc

$$\lambda \;=\; \tau+\tau^{-1}.$$

Le numérateur de $P_n(\lambda) = \frac{\tau^{-n-1}\left(\tau^{2(n+1)}-1\right)}{\tau-\tau^{-1}}$ s'annule lorsque $\tau^2 = \exp\left(\frac{2i\ell\pi}{n+1}\right)$ pour $\ell = 0, 1, \ldots, n$; la valeur $\ell = 0$ ne compte pas, car c'est un zéro du dénominateur. On a donc bien retrouvé les valeurs propres

$$\lambda \;=\; \tau+\tau^{-1} \;=\; \exp\left(\frac{i\ell\pi}{n+1}\right)+\exp\left(\frac{-i\ell\pi}{n+1}\right) \;=\; 2\cos\left(\frac{\ell\pi}{n+1}\right)$$

comme plus haut.

(iii) Autant que je sache et comme déjà mentionné plus haut, le calcul de Lagrange reproduit ci-dessus est, du point de vue historique, le premier exemple de diagonalisation d'une matrice carrée de taille quelconque.

9. Exercices

Ci-dessous, les lettres V, W désignent des espaces hermitiens.

(VI.1) Soient α, β deux opérateurs autoadjoints sur un espace hermitien. Montrer que $\alpha\beta$ est autoadjoint si et seulement si α et β commutent.

(VI.2) Soient V un espace hermitien, U un sous-espace et $U^\perp$ le supplémentaire orthogonal de U dans V. Soit

$$p \;:\; \begin{cases} V \longrightarrow V \\ v \longmapsto v^{\|} \end{cases}$$

la projection orthogonale de V sur U, définie comme après le théorème VI.4. Vérifier que

$$p^* \;=\; p \;=\; p^2.$$

Si $U \neq \{0\}$ et $U^\perp \neq \{0\}$, vérifier que les valeurs propres de p sont 0 et 1, que celles de $\mathrm{id}_V - 2p$ sont 1 et -1, et que $\mathrm{id}_V - 2p$ est à la fois autoadjoint et unitaire.

Pour $\lambda \in \mathbb{C} \setminus \{0, 1\}$, trouver $\mu, \nu \in \mathbb{C}$ tels que $(\lambda\,\mathrm{id}_V - p)^{-1} = \mu p + \nu(\mathrm{id}_V - p)$.

(VI.3) Pour tout opérateur $\beta : V \longrightarrow W$, vérifier que $\mathrm{Ker}(\beta^*) = (Im(\beta))^\perp$.

(VI.4) Vérifier que $\alpha \in \mathcal{L}(V)$ est normal si et *seulement si* $\|\alpha^* v\| = \|\alpha v\|$ pour tout $v \in V$.

[*Indication* : calculer $\langle(\alpha^*\alpha - \alpha\alpha^*)v \mid v\rangle$ et appliquer un argument de la preuve de la proposition 3.2.]

(VI.5) Soient $\alpha_1, \alpha_2 \in \mathcal{L}(V)$ deux opérateurs autoadjoints. Vérifier que $\alpha_1 + i\alpha_2$ est normal si et seulement α_1 et α_2 commutent.

(VI.6) Soient $\alpha_1, \alpha_2 \in \mathcal{L}(V)$ des opérateurs autoadjoints tels que $\alpha_1\alpha_2 = \alpha_2\alpha_1$. Montrer qu'il existe une base orthonormale de V formée de vecteurs propres *communs* à α_1 et α_2.

[*Indication* : vérifier que tout espace propre de α_1 est invariant par α_2.]

Généraliser à une suite $\alpha_1, \ldots, \alpha_k$ de k opérateurs autoadjoints commutant deux à deux.

(VI.7) Soit $\alpha \in \mathcal{L}(V)$ un opérateur autoadjoint. Montrer que les propriétés suivantes sont équivalentes.

(i) $\langle v \mid \alpha v\rangle \geq 0$ pour tout $v \in V$;

(ii) $\lambda \geq 0$ pour toute valeur propre λ de α ;

(iii) il existe un opérateur autoadjoint $\gamma \in \mathcal{L}(V)$ dont toutes les valeurs propres sont ≥ 0 tel que $\alpha = \gamma^2$;

(iv) il existe un opérateur $\beta \in \mathcal{L}(V)$ tel que $\alpha = \beta^*\beta$.

[*Indication* : pour (ii) $\Longrightarrow$ (iii), diagonaliser α.]

Les opérateurs satisfaisant ces conditions sont dits *positifs*. Montrer que la somme de deux opérateurs positifs est un opérateur positif.

(VI.8) Soit u_ϵ une matrice complexe de la forme $\begin{pmatrix} 1+i\epsilon x & i\epsilon y \\ i\epsilon z & 1+i\epsilon t \end{pmatrix}$, dépendant d'un paramètre ϵ auquel on pense comme étant *petit,* et soit a la matrice $\begin{pmatrix} x & y \\ z & t \end{pmatrix}$. Montrer que la relation « $u_\epsilon^* u_\epsilon = I +$ termes en ϵ^2 » équivaut à la relation « $a^* = a$ » .

(VI.9) Vérifier que l'application $h : M_n(\mathbb{C}) \times M_n(\mathbb{C}) \longrightarrow \mathbb{C}$ définie par $h(a,b) = \mathrm{trace}(a^*b)$ est un produit scalaire sur l'espace vectoriel complexe $M_n(\mathbb{C})$. [Rappel : la trace d'une matrice $c \in M_n(\mathbb{C})$ est le nombre $\mathrm{trace}(c) = \sum_{j=1}^n c_{j,j}$.]

(VI.10) On considère un entier $n \geq 1$, deux matrices à coefficients réels, à 1 ligne et n colonnes, notées $a = (a_1, \ldots, a_n)$ et $b = (b_1, \ldots, b_n)$, la matrice transposée ${}^t a$ de a, ainsi que la matrice

$$c = {}^t ab \in M_n(\mathbb{R}).$$

On suppose les matrices a et b non nulles.

(i) Décrire le noyau de c.

(ii) Montrer qu'il existe un nombre réel s tel que $c^2 = sc$.

(iii) Calculer les valeurs propres de c, avec leurs multiplicités.
[*Indication* : distinguer deux cas, selon que $\sum_{i=1}^n a_i b_i$ est nul ou non.]

(iv) Calculer le déterminant de $2I_n + c$, où I_n désigne la matrice unité dans $M_n(\mathbb{R})$.

(v) Pour que la matrice c soit symétrique, il suffit que a soit un multiple scalaire de b. Montrer que cette condition est aussi nécessaire.

(VI.11) Soit V l'espace des fonctions continues $f : \mathbb{R} \longrightarrow \mathbb{R}$ telles que

$$\int_{\mathbb{R}} |f(t)|^2 e^{-t^2} dt \, < \, \infty,$$

muni du produit scalaire défini par

$$\langle f_1 \mid f_2 \rangle \, = \, \int_{\mathbb{R}} f_1(t) f_2(t) e^{-t^2} dt.$$

Soient $f_0, \ldots, f_n \in V$ les fonctions définies par $f_j(t) = t^j$, et soient $h_0, \ldots, h_n$ les fonctions obtenues à partir des f_j par le procédé de Gram-Schmidt. Calculer h_j pour $j \leq 2$.

[*Indication* : Pour cet exercice, on admet que $\int_{-\infty}^{\infty} e^{-t^2} dt = \sqrt{\pi}$.]

(VI.12) Soit V l'espace des fonctions continues $f : [0, \infty[\longrightarrow \mathbb{R}$ telles que

$$\int_0^{\infty} |f(t)|^2 e^{-t} dt \, < \, \infty,$$

muni du produit scalaire défini par

$$\langle f_1 \mid f_2 \rangle \, = \, \int_0^{\infty} f_1(t) f_2(t) e^{-t} dt.$$

Soient $f_0, f_1, f_2, \ldots \in V$ les fonctions définies par $f_j(t) = t^j$ et soit U le sous-espace de V engendré par f_0 et f_1. Calculer la projection orthogonale de f_2 sur U.

[Rappel : en intégrant par parties, on calcule

$$\int_0^{\infty} t^n e^{-t} dt \, = \, n \int_0^{\infty} t^{n-1} e^{-t} dt \, = \, \cdots \, = \, n!$$

pour tout $n \geq 0$.]

(VI.13) Soient d, e deux droites passant par l'origine du plan euclidien $\mathbb{R}^2$. On note $\theta \in [0, \pi/2]$ l'angle de ces droites et p [respectivement q] la projection orthogonale de $\mathbb{R}^2$ sur d [resp. e].

Calculer le spectre de pq.

(VI.14) Soient V un espace hermitien, n un entier et $v_1, \ldots, v_n$ une suite de n vecteurs de V. La *matrice de Gram* de ces vecteurs est par définition la matrice

$$g \, = \, (g_{j,k} J)_{1 \leq j,k \leq n} \, \in \, M_n(\mathbb{C}) \, , \qquad g_{j,k} \, = \, \langle v_j \mid v_k \rangle.$$

(i) Utiliser une variante du procédé de Gram-Schmidt pour montrer qu'il existe des vecteurs $e_1, \ldots, e_n \in V$ tels que

(A) $\langle e_1, \ldots, e_j \rangle = \langle v_1, \ldots, v_j \rangle$ pour tout $j \in \{1, \ldots, n\}$;

(B) $\langle e_j \mid e_k \rangle = 0$ pour tous $j, k \in \{1, \ldots, n\}$ avec $j \neq k$;

(C) $\langle e_j \mid e_j \rangle = \begin{cases} 1 & \text{si } \langle v_1, \ldots, v_j \rangle \neq \langle v_1, \ldots, v_{j-1} \rangle \\ 0 & \text{si } \langle v_1, \ldots, v_j \rangle = \langle v_1, \ldots, v_{j-1} \rangle. \end{cases}$

(ii) Soit $c = (c_{j,k})_{1\leq j,k\leq n} \in M_n(\mathbb{C})$ la matrice telle que $e_k = \sum_{j=1}^{k} c_{j,k} v_k$ pour tout $k \in \{1, \ldots, n\}$. Ecrire d'abord les coefficients de la matrice gc en fonction des produits scalaires $\langle v_j \mid e_k \rangle$, ensuite les coefficients de la matrice $c^* gc$ en fonction des $\langle e_j \mid e_k \rangle$.

(iii) Déduire de ce qui précède que la matrice g est inversible si et seulement si les vecteurs $v_1, \ldots, v_n$ sont linéairement indépendants.

(VI.15) Soient V un espace vectoriel complexe de dimension finie et $a \in \mathcal{L}(V)$ un opérateur sur V tel que $a^3 = 1$.

(i) On pose $\omega = \exp(2i\pi/3)$ et $p_j = \frac{1}{3}(1 + \omega^j a + \omega^{2j} a^2)$ pour $j \in \{0, 1, 2\}$. Vérifier que

$$\begin{aligned} p_j^2 &= p_j \qquad \text{pour} \quad j \in \{0,1,2\}, \\ p_j p_k &= 0 \qquad \text{pour} \quad j,k \in \{0,1,2\} \quad \text{avec } j \neq k, \\ a &= p_0 + \omega^2 p_1 + \omega p_2, \\ ap_j &= p_j a = \omega^j p_j \qquad \text{pour } j \in \{0,1,2\}. \end{aligned}$$

(ii) Montrer qu'il existe une base $\{e_1, \ldots, e_l, e_{l+1}, \ldots, e_m, e_{m+1}, \ldots, e_n\}$ de V telle que

$\{e_1, \ldots, e_l\}$ soit une base de l'image de p_0,

$\{e_{l+1}, \ldots, e_m\}$ soit une base de l'image de p_1,

$\{e_{m+1}, \ldots, e_n\}$ soit une base de l'image de p_2.

Quelle est la matrice de a relativement à cette base ?

(iii) On suppose de plus que V est muni d'un produit scalaire et que $a^* a = 1$, c'est-à-dire que a est un opérateur *unitaire*. Vérifier que $p_j^* = p_j$ pour $j \in \{0, 1, 2\}$, et que les images des p_j sont des sous-espaces de V orthogonaux deux à deux.

(iv) [(iv)$^\sharp$] Formuler des énoncés analogues en remplaçant 3 par un nombre $N \geq 2$.

(VI.16) On appelle *réseau cubique à faces centrées* l'ensemble des points

$$\Lambda = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 \;\middle|\; x, y, z \in \mathbb{Z} \quad \text{et} J \quad x + y + z \equiv 0 \pmod 2 \right\}$$

de l'espace euclidien standard.

(i) Soit C le cube défini par les inégalités $0 \leq x \leq 2$, $0 \leq y \leq 2$, $0 \leq z \leq 2$. Décrire les points de $\Lambda \cap C$.

(ii) Soient $\lambda_j \in L$ deux points respectivement de coordonnées x_j, y_j, z_j $(j = 1, 2)$ et soit $d = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2}$ leur distance. Vérifier que $d \geq \sqrt{2}$, et que $\sqrt{2}$ est la plus grande valeur qui convient pour cette inégalité.

(iii) On place en chaque point de Λ une boule de rayon $\sqrt{2}/2$. Combien de ces boules touchent-elles la boule centrée à l'originie ?

(VI.17) Soient n un entier, $n \geq 2$, et $\mathbb{S}^{n-1} = \{v \in \mathbb{R}^n \mid \|v\| = 1\}$ la sphère unité de l'espace euclidien de dimension n. Soit $E^{(2)}$ l'espace vectoriel des fonctions de $\mathbb{R}^n$ dans $\mathbb{R}$ de la forme $v \longmapsto \sum_{1\leq i,j\leq n} a_{i,j} v_i v_j + \sum_{1\leq k\leq n} b_k v_k + c$, où $a_{i,j}$, b_k et c sont des constantes réelles, et soit F l'espace vectoriel des restrictions à $\mathbb{S}^{n-1}$ de fonctions dans $E^{(2)}$.

Soit X un sous-ensemble fini de $\mathbb{S}^{n-1}$ qui est un *ensemble à deux distances*, c'est-à-dire pour lequel il existe $\alpha, \beta \in [-1, 1[$ tels que $\langle x \mid y\rangle \in \{\alpha, \beta\}$ pour tous $x, y \in X$, $x \neq y$. L'objet de cet exercice est de majorer la cardinalité d'un tel X.

(i) Calculer la dimension de $E^{(2)}$ et vérifier que $\dim(F) \leq \frac{1}{2}n(n+3)$.

(ii) Pour tout $y \in X$, on définit $f_y \in E^{(2)}$ par

$$f_y(v) = \frac{(\langle y \mid v\rangle - \alpha)(\langle y \mid v\rangle - \beta)}{(1-\alpha)(1-\beta)} \qquad (v \in \mathbb{R}^n)$$

et on note g_y la restriction de f_y à $\mathbb{S}^{n-1}$. Montrer que $(g_y)_{y\in X}$ est une famille libre dans F.

(iii) Déduire de (ii) que $|X| \leq \frac{1}{2}n(n+3)$, et vérifier qu'il existe un exemple de X en dimension $n = 2$ pour lequel l'inégalité est une égalité.

[Les seuls autres dimensions où on connaisse des exemples avec égalité $|X| = \frac{1}{2}n(n+3)$ sont $n = 6$ et $n = 22$.]

(VI.18) On considère l'espace $\mathbb{R}^4$ muni du produit scalaire canonique et de la base canonique (e_1, e_2, e_3, e_4), le sous-espace de dimension 3

$$V = \left\{ \sum_{j=1}^{4} x_j e_j \in \mathbb{R}^4 \;\middle|\; \sum_{j=1}^{4} x_j = 0 \right\}$$

et la sphère unité

$$\mathbb{S}^2 = \{v \in V \mid \langle v \mid v\rangle = 1\}$$

de V.

(i) Vérifier que les 4 vecteurs $w_j = \frac{2}{\sqrt{3}} e_j - \frac{1}{2\sqrt{3}}(e_1 + e_2 + e_3 + e_4)$, $1 \leq j \leq 4$, sont les sommets d'un *tétraèdre régulier* inscrit dans $\mathbb{S}^2$, c'est-à-dire qu'il existe un nombre ρ_M, à déterminer, tel que $\|w_j - w_k\|^2 = \rho_M$ pour tous $j, k \in \{1, 2, 3, 4\}$ tels que $j \neq k$.

(ii) Soient $v_1, v_2, v_3, v_4 \in \mathbb{S}^2$; on pose

$$\rho = \inf_{\substack{1\leq j,k\leq 4 \\ j\neq k}} \|v_j - v_k\|^2 .$$

Montrer que $\rho \leq \rho_M$, avec égalité si et seulement si $\|v_j - v_k\|^2 = \rho_M$ pour tous $j, k \in \{1, 2, 3, 4\}$ tels que $j \neq k$.

[*Indication* : Si $\|v_j - v_k\|^2 \geq \rho$, alors $\langle v_j \mid v_k\rangle \leq 1 - \rho/2$ et $0 \leq \sum_{j,k} \langle v_j \mid v_k\rangle \leq 4 + 12(1 - \rho/2)$.]

REMARQUE. Le *problème de Tammes* (du nom d'un botaniste hollandais qui s'est intéressé vers 1930 à la géométrie des grains de pollen) consiste à déterminer, pour tout entier $n \geq 1$, les configurations de n points $w_1, \ldots, w_n$ sur la sphère $\mathbb{S}^2$ telles que $inf_{j \neq k} \|w_j - w_k\|^2$ soit maximum. La solution du problème est connue pour $n \leq 12$ et $n = 24$, et inconnue dans tous les autres cas. Le présent exercice fournit la solution du problème pour $n = 4$.

Le problème n'a guère de sens pour $n = 1$. Sa solution pour $n = 2$ est immédiate : on trouve les paires de points antipodaux. La solution pour $n = 3$ est assez facile : on trouve les triples de points disposés sur un même grand cercle de $\mathbb{S}^2$ (qu'on peut appeler l'équateur) et qui sont les sommets d'un triangle équilatéral. [L'indication fournie ci-desus pour quatre points est valable à un seul changement près : si $u_1, u_2, u_3 \in \mathbb{S}^2$ sont tels que $\|u_j - u_k\|^2 \geq \rho^2$ pour $j \neq k$, alors $0 \leq \sum_{j,k} \langle u_j \mid u_k \rangle \leq 3 + 6(1 - \rho/2)$.] Les solutions pour $n = 5, 7, 8, 9, 10, 11, 12$ et 24 sont trop longues à décrire ici.

(iii) Deviner quelle est la solution du problème de Tammes pour $n = 6$.

Pour en savoir plus.

M. Berger, *Le problème des dictateurs ennemis,* La Recherche, numéro "Spécial mathématiques", octobre 2001, pages 38–44.

M. Berger, *Les dictateurs, l'architecte et le golfeur,* Pour la Science, dossier "La sphère sous toutes ses formes", décembre 2003, pages 40–43.