

PLAN

I. MATRICES, ENDOMORPHISMES ET DÉTERMINANTS

1. Espaces vectoriels
2. Dimension des espaces vectoriels
3. Applications linéaires
4. Matrices
5. Polynômes d'endomorphismes
6. Permutations
7. Formes multilinéaires alternées
8. Déterminants
9. Déterminant d'un endomorphisme
10. Déterminant d'une matrice
11. Calcul et développements d'un déterminant
12. EXERCICES

II. RÉDUCTION DES ENDOMORPHISMES

13. Valeurs propres, vecteurs propres
14. Endomorphismes diagonalisables
15. Polynôme caractéristique d'une matrice carrée
16. Polynôme caractéristique d'un endomorphisme
17. Polynômes
18. Trigonalisation, théorème de Hamilton-Cayley
19. Polynôme minimal
20. Réduction des endomorphismes

21. EXERCICES

III. EXPONENTIELLES DE MATRICES

22. Normes sur un espace vectoriel

23. Suites et séries vectorielles

24. Suites et séries de matrices

25. Exponentielle de matrices

26. Continuité et dérivabilité des fonctions vectorielles

27. EXERCICES

IV. SYSTÈMES ET ÉQUATIONS DIFFÉRENTIELS LINÉAIRES HOMOGÈNES À COEFFICIENTS CONSTANTS

28. Systèmes différentiels linéaires homogènes à coefficients constants

29. Conseils pratiques

30. Allure des trajectoires

31. Equations différentielles linéaires homogènes à coefficients constants d'ordre supérieur

32. EXERCICES

V. DUALITÉ, FORMES BILINÉAIRES ET QUADRATIQUES

33. Dualité, bases duales, bidualité

34. Orthogonalité, transposée d'une application linéaire

35. Formes bilinéaires

36. Formes bilinéaires symétriques et formes quadratiques

37. Orthogonalité, vecteurs isotropes

38. Bases orthogonales

39. Classification des formes quadratiques sur \mathbb{R} et \mathbb{C}

40. Méthode de Gauss de réduction des formes quadratiques

41. Groupe orthogonal

42. EXERCICES

VI. ESPACES EUCLIDIENS

43. Produit scalaire

44. Orthogonalité

45. Adjoint d'un endomorphisme, endomorphismes symétriques

46. Isométries du plan

47. Isométries en dimension 3

48. Isométries en dimension quelconque

49. Réduction simultanée des formes quadratiques

50. Coniques

51. Surfaces quadriques

52. EXERCICES

Le but de ce chapitre est d'introduire la notion de *déterminant*. On commence par quelques rappels sans démonstration sur les espaces vectoriels et les applications linéaires, présentés de façon à mettre en valeur l'intérêt des concepts plutôt que les détails techniques. Il est indispensable de bien les maîtriser avant d'aborder la suite du cours. On introduit ensuite quelques notions de bases sur les permutations d'un ensemble fini, puis on définit le déterminant d'une famille de n vecteurs dans un espace vectoriel de dimension n , puis le déterminant d'un endomorphisme et le déterminant d'une matrice carrée, avec application au calcul de l'inverse.

1. Espaces vectoriels

Dans tout ce cours, on fixe un corps commutatif K contenant le corps \mathbf{Q} des rationnels; ce sera pour nous pratiquement toujours soit le corps \mathbf{R} des nombres réels, soit le corps \mathbf{C} des nombres complexes.

Après avoir « fait de la géométrie » dans le plan (c'est-à-dire \mathbf{R}^2) ou dans l'espace (c'est-à-dire \mathbf{R}^3) pendant des siècles, les mathématiciens ont eu l'idée d'essayer de dégager celles des propriétés de ces espaces qui permettaient en fait de « faire de la géométrie ». Ces propriétés de base se réduisent à ceci : on peut ajouter deux vecteurs et on peut multiplier un vecteur par un « scalaire », c'est-à-dire un élément du corps K , ces deux opérations obéissant à quelques règles simples de compatibilité. On définira donc un *espace vectoriel* comme un ensemble E muni de deux lois, une addition interne et une multiplication externe, qui vérifient

- l'ensemble E muni de la loi $+$ est un groupe commutatif;
- on a

$$(a + b)x = ax + bx \quad a(x + y) = ax + ay \quad a(bx) = (ab)x ,$$

où a et b sont des *scalaires* (c'est-à-dire des éléments de K) quelconques, et x et y des *vecteurs* (c'est-à-dire des éléments de E) quelconques.

Exemples 1.1. 1) L'ensemble K^n , muni des deux lois

$$(x_1, \dots, x_n) + (y_1, \dots, y_n) = (x_1 + y_1, \dots, x_n + y_n) \quad a(x_1, \dots, x_n) = (ax_1, \dots, ax_n)$$

est un espace vectoriel.

2) L'ensemble $K[T]$ des polynômes à une indéterminée T et à coefficients dans K , muni de l'addition des polynômes, et de la multiplication habituelle par un scalaire, est un espace vectoriel. La multiplication des polynômes entre eux est une structure supplémentaire qui n'intervient pas ici.

3) Un troisième exemple plus exotique est le suivant : si l'on part d'un espace vectoriel E et de n'importe quel ensemble X , l'ensemble de toutes les fonctions de X dans E est muni d'une structure d'espace vectoriel (exercice) : il suffit de remarquer que pour additionner deux fonctions par exemple, il suffit d'être capable d'additionner leurs images.

2. Dimension des espaces vectoriels

Vous avez tous, même de façon intuitive, une notion de la dimension ; on a envie par exemple de dire qu'une droite est de dimension 1, un plan de dimension 2, l'espace dans lequel nous vivons de dimension 3, et l'espace vectoriel K^n de dimension n . Le problème auquel nous sommes confrontés est d'arriver à définir la dimension d'un espace vectoriel de façon intrinsèque, c'est-à-dire en ne faisant appel qu'aux opérations définissant la structure.

La raison pour laquelle on dit que l'espace dans lequel nous vivons est de dimension 3 est que l'on peut repérer un point par 3 coordonnées. Plus généralement, dans K^n , on attribue à un vecteur (x_1, \dots, x_n) les « coordonnées » x_1, \dots, x_n parce que l'on peut écrire

$$x = x_1 e_1 + \dots + x_n e_n,$$

où $e_1 = (1, 0, \dots, 0)$, $e_2 = (0, 1, 0, \dots, 0)$, \dots , $e_n = (0, \dots, 0, 1)$. De plus, ces « coordonnées » sont uniquement définies.

C'est de cette façon que l'on va approcher la définition de la dimension d'un espace vectoriel. Les deux difficultés principales auxquelles on se heurte sont :

- dans un espace vectoriel quelconque, existe-t-il toujours un « repère » analogue à la famille (e_1, \dots, e_n) de vecteurs de K^n ?
- deux tels « repères » ont-ils toujours le même nombre de vecteurs (ce problème se pose déjà dans K^n) ?

Pour avancer un peu dans la théorie, il est utile à ce stade d'introduire un peu de terminologie, pour préciser les notions de « coordonnées » et de « repère » utilisées de façon intuitive plus haut.

Etant donné un espace vectoriel E , une *combinaison linéaire* de vecteurs e_1, \dots, e_r de E est une somme $a_1 e_1 + \dots + a_r e_r$, où a_1, \dots, a_r sont des scalaires. Pour pouvoir associer à tout vecteur des « coordonnées », il faut déjà que tout vecteur de E soit combinaison linéaire de e_1, \dots, e_r ; une famille qui a cette propriété est dite *génératrice*.

Pour que ces coordonnées soit uniques, il faut que tout vecteur de E qui est combinaison linéaire de e_1, \dots, e_r le soit *de façon unique*, c'est-à-dire que les scalaires a_1, \dots, a_r soient uniquement déterminés ; une famille qui a cette propriété est dite *libre*. De façon équivalente, une famille (e_1, \dots, e_r) est libre si aucun de ses vecteurs n'est combinaison linéaire des autres, ou encore si toute combinaison linéaire nulle de ces vecteurs a tous ses coefficients nuls.

Finalement, une *base* (plutôt qu'un repère) de E est une famille libre et génératrice : tout vecteur x de E est combinaison linéaire des vecteurs d'une base, et les coefficients de

cette combinaison linéaire sont uniquement déterminés par x ; on les appelle les *coordonnées* (ou les *composantes*) de x dans la base.

Malheureusement, nous n'avons pas beaucoup progressé vers une définition rigoureuse de la dimension : il reste encore à répondre aux deux questions ci-dessus. C'est là le résultat principal de la théorie ; sa démonstration est délicate, et je renvoie pour cela à votre cours de première année. Il s'agit du fait qu'étant donné un espace vectoriel E , *il existe toujours une base de E , et deux bases quelconques ont même nombre d'éléments*. Cela permet de définir la *dimension* de E comme le nombre d'éléments dans une base quelconque de E .

Exemples 2.1. 1) Il est clair que les vecteurs e_1, e_2, \dots, e_n de K^n définis plus haut forment une base de cet espace vectoriel ; on l'appelle la *base canonique* de K^n . Il en résulte que K^n est de dimension n .

2) Il n'existe pas de base de l'espace vectoriel $K[T]$ qui ait un nombre fini d'éléments. C'est une difficulté qui a été camouflée jusqu'à présent (où nous n'avons considéré que des familles libres ou génératrices *finies*), la raison en étant que la théorie marche tout aussi bien en général, mais que nous ne considérerons dans ce cours que des espaces vectoriels *de dimension finie*, c'est-à-dire dont toutes les bases ont un nombre fini d'éléments (on montre qu'il suffit pour cela qu'il existe une famille génératrice finie). Pour tout entier $d \geq 0$, l'ensemble des polynômes de $K[T]$ de degré $\leq d$ forment un espace vectoriel ; une base en est $(1, T, T^2, \dots, T^d)$, de sorte que sa dimension est $d + 1$.

3. Applications linéaires

(3.1) De nombreuses transformations courantes de l'espace (projections, symétries, rotations, ...) ont en commun une propriété importante, la *linéarité* : on dit qu'une application $u : E \rightarrow F$ entre des espaces vectoriels est *linéaire* si

$$u(x + y) = u(x) + u(y) \quad u(ax) = au(x),$$

pour tous x, y dans E et tout a dans K . On vérifie facilement (c'est le même principe que dans l'exemple 1.1.3)) que $\mathcal{L}(E, F)$, l'ensemble de toutes les applications linéaires de E dans F , est lui-même un espace vectoriel. Si E et F sont de dimension finie, $\mathcal{L}(E, F)$ est de dimension $(\dim E)(\dim F)$.

Un *endomorphisme* de E est une application linéaire de E dans lui-même. On note $\mathcal{L}(E)$ l'espace vectoriel de tous les endomorphismes de E .

Exemples 3.2. 1) L'application de $K[T]$ dans K qui à un polynôme P associe $P(1)$ est linéaire. En revanche, celle qui à P associe $P(1)^2$ (ou $P(0) + 1$) ne l'est pas.

2) L'application $u : K^3 \rightarrow K^2$ définie par $u(x, y, z) = (2x + 3y - z, x - z)$ est linéaire. Pas celle qui à (x, y, z) associe $(x + y^2, x - z - 1)$.

3) Si E est un espace vectoriel et a un élément de K fixé, l'application $u : E \rightarrow E$ définie par $u(x) = ax$ est un endomorphisme de E (appelé *homothétie de rapport a*).

4) Pour ceux qui connaissent un peu de géométrie : les projections, les symétries dans un espace vectoriel E sont des endomorphismes de E . Si l'espace vectoriel \mathbf{R}^2 est muni de sa structure euclidienne orientée canonique, la rotation r_θ d'angle θ est linéaire (cf. § 46).

L'espace vectoriel $\mathcal{L}(E)$ est aussi une K -algèbre, puisqu'on peut « multiplier » deux endomorphismes u et v en en prenant la composée. Les propriétés de compatibilité sont :

$$(au) \circ v = a(u \circ v) \quad (u + v) \circ w = (u \circ w) + (v \circ w),$$

et la même distributivité de l'autre côté.

4. Matrices

On rappelle qu'une matrice à $m \times n$ à coefficients dans K est un tableau à m lignes et n colonnes d'éléments de K . On peut additionner des matrices, les multiplier par un élément de K , et l'on vérifie que ces opérations satisfont les propriétés de compatibilité voulues. On obtient donc un K -espace vectoriel que l'on note $\mathcal{M}_{m,n}(K)$. Se donner une matrice revient à se donner mn éléments de K . En particulier, $\mathcal{M}_{m,n}(K)$ est de dimension mn .

Lorsque $m = n$, on dit que l'on a des matrices carrées d'ordre n . On peut les multiplier entre elles (parce qu'elles sont carrées), et l'on obtient ainsi une K -algèbre que l'on note $\mathcal{M}_n(K)$. L'élément neutre de la multiplication (qui est la matrice dont tous les coefficients diagonaux valent 1 et les autres 0) est noté I_n . Une matrice carrée M d'ordre n est inversible s'il existe une matrice carrée N de même ordre telle que $MN = NM = I_n$. La matrice N est alors unique; on la note M^{-1} et on l'appelle l'inverse de M .

Soient E un espace vectoriel de dimension n et $\mathcal{B} = (e_1, \dots, e_n)$ une base de E . On associe à tout vecteur x de E la matrice colonne $X_{\mathcal{B}}$ à n lignes dont la i ième ligne est la i ième composante de x dans la base \mathcal{B} .

Exemple 4.1. Dans la base canonique \mathcal{B} de \mathbf{R}^2 , la matrice des composantes du vecteur $(4, 7)$ est

$$X_{\mathcal{B}} = \begin{pmatrix} 4 \\ 7 \end{pmatrix};$$

dans la base $\mathcal{B}' = ((1, 2); (3, 5))$, c'est

$$X_{\mathcal{B}'} = \begin{pmatrix} 1 \\ 1 \end{pmatrix},$$

puisque $(4, 7) = (1, 2) + (3, 5)$.

Soit maintenant u un endomorphisme de E ; on peut décomposer chaque vecteur $u(e_j)$ comme combinaison linéaire des éléments de \mathcal{B} :

$$u(e_j) = a_{1j}e_1 + \dots + a_{nj}e_n$$

On obtient ainsi une matrice carrée d'ordre n , de coefficients $(a_{ij})_{1 \leq i, j \leq n}$, où l'on peut « lire » l'image de e_j sur la j ième colonne; c'est la matrice de f dans la base \mathcal{B} ; on la note $M_{\mathcal{B}}(u)$.

Exemple 4.2. 1) Si $\mathcal{B} = (e_1, e_2)$ est une base de \mathbf{R}^2 , que D_1 est la droite dirigée par e_1 et D_2 la droite dirigée par e_2 , la matrice dans la base \mathcal{B} de la projection sur D_1 parallèlement à D_2 est

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

La matrice de la symétrie par rapport à D_1 parallèlement à D_2 est

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

2) La rotation r_θ d'angle θ dans le plan euclidien orienté (cf. § 46) a pour matrice dans la base canonique

$$\begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} .$$

La relation fondamentale est la suivante : si $X_{\mathcal{B}}$ est la matrice des composantes d'un vecteur x de E dans la base \mathcal{B} , celle des composantes de $u(x)$ dans la même base est

$$(4.3) \quad \boxed{M_{\mathcal{B}}(u) X_{\mathcal{B}}}$$

En d'autres termes, prendre l'image d'un vecteur par un endomorphisme revient à multiplier la matrice de ses composantes par celle de u . Cette propriété caractérise la matrice de u .

Exemple 4.4. L'image du vecteur $(4, 7)$ par la rotation $r_{\pi/4}$ dans le plan euclidien orienté (cf. § 46) est le vecteur

$$\begin{pmatrix} \sqrt{2}/2 & -\sqrt{2}/2 \\ \sqrt{2}/2 & \sqrt{2}/2 \end{pmatrix} \begin{pmatrix} 4 \\ 7 \end{pmatrix} = \begin{pmatrix} -3\sqrt{2}/2 \\ 11\sqrt{2}/2 \end{pmatrix} .$$

En associant à un endomorphisme sa matrice dans la base \mathcal{B} , on définit une application $M_{\mathcal{B}}$ de $\mathcal{L}(E)$ dans $\mathcal{M}_n(K)$, dont il est facile de vérifier que c'est une application linéaire bijective. C'est en fait mieux : on a

$$\boxed{M_{\mathcal{B}}(v \circ u) = M_{\mathcal{B}}(v)M_{\mathcal{B}}(u)}$$

de sorte que $M_{\mathcal{B}}$ est un morphisme de K -algèbres. Cela peut se justifier de la façon suivante : pour tout vecteur x de matrice de composantes X , la matrice des composantes de $(v \circ u)(x) = v(u(x))$ est, par (4.3),

$$M(v)(M(u)X) = (M(v)M(u)) X ,$$

de sorte que $M(v)M(u)$ est bien la matrice de $v \circ u$. Il faut bien se dire que la multiplication des matrices, qui semble *a priori* tomber du ciel, est définie justement pour que l'égalité ci-dessus soit vérifiée.

(4.5) La matrice d'un endomorphisme dépend de la base que l'on a choisie. Nous allons maintenant étudier l'effet d'un changement de base. On se donne une autre base $\mathcal{B}' = (e'_1, \dots, e'_n)$ de E , que l'on exprime à partir de l'ancienne par sa *matrice de passage* $P_{\mathcal{B} \rightarrow \mathcal{B}'} = (p_{ij})_{1 \leq i, j \leq n}$ (ou simplement P), dont les colonnes sont les composantes des vecteurs de \mathcal{B}' dans l'ancienne base \mathcal{B} :

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

On montre facilement que $P_{\mathcal{B} \rightarrow \mathcal{B}'}$ est inversible, et que son inverse est $P_{\mathcal{B}' \rightarrow \mathcal{B}}$; cela signifie simplement

$$P_{\mathcal{B} \rightarrow \mathcal{B}'} P_{\mathcal{B}' \rightarrow \mathcal{B}} = P_{\mathcal{B}' \rightarrow \mathcal{B}} P_{\mathcal{B} \rightarrow \mathcal{B}'} = I_n .$$

On a la relation fondamentale suivante : pour tout vecteur x ,

$$(4.6) \quad \boxed{X_{\mathcal{B}} = P_{\mathcal{B} \rightarrow \mathcal{B}'} X_{\mathcal{B}'}}$$

On en déduit que, si $y = u(x)$, on a

$$M_{\mathcal{B}'}(u) X_{\mathcal{B}'} = Y_{\mathcal{B}'} = P_{\mathcal{B}' \rightarrow \mathcal{B}} Y_{\mathcal{B}} = P_{\mathcal{B}' \rightarrow \mathcal{B}} M_{\mathcal{B}}(u) X_{\mathcal{B}} = P_{\mathcal{B}' \rightarrow \mathcal{B}} M_{\mathcal{B}}(u) P_{\mathcal{B} \rightarrow \mathcal{B}'} X_{\mathcal{B}'}$$

pour tout $X_{\mathcal{B}'}$, de sorte que

$$\boxed{M_{\mathcal{B}'}(u) = P_{\mathcal{B}' \rightarrow \mathcal{B}} M_{\mathcal{B}}(u) P_{\mathcal{B} \rightarrow \mathcal{B}'}}$$

ou encore

$$(4.7) \quad \boxed{M_{\mathcal{B}'}(u) = P_{\mathcal{B} \rightarrow \mathcal{B}'}^{-1} M_{\mathcal{B}}(u) P_{\mathcal{B} \rightarrow \mathcal{B}'}}$$

Exemple 4.8. Soient \mathcal{B} la base canonique de \mathbf{R}^2 et \mathcal{B}' la base $((1, 2); (3, 5))$. On a

$$P_{\mathcal{B} \rightarrow \mathcal{B}'} = \begin{pmatrix} 1 & 3 \\ 2 & 5 \end{pmatrix} \quad P_{\mathcal{B}' \rightarrow \mathcal{B}} = P_{\mathcal{B} \rightarrow \mathcal{B}'}^{-1} = \begin{pmatrix} -5 & 3 \\ 2 & -1 \end{pmatrix} .$$

La matrice de la rotation $r_{\pi/4}$ dans la base \mathcal{B}' est (cf. § 46)

$$\begin{pmatrix} -5 & 3 \\ 2 & -1 \end{pmatrix} \begin{pmatrix} \sqrt{2}/2 & -\sqrt{2}/2 \\ \sqrt{2}/2 & \sqrt{2}/2 \end{pmatrix} \begin{pmatrix} 1 & 3 \\ 2 & 5 \end{pmatrix} = \begin{pmatrix} -5 & 3 \\ 2 & -1 \end{pmatrix} \begin{pmatrix} -\sqrt{2}/2 & -\sqrt{2} \\ 3\sqrt{2}/2 & 4\sqrt{2} \end{pmatrix} = \begin{pmatrix} 7\sqrt{2} & 17\sqrt{2} \\ -5\sqrt{2}/2 & -6\sqrt{2} \end{pmatrix} .$$

5. Polynômes d'endomorphismes

Soient E un espace vectoriel et u un endomorphisme de E . On introduit les notations commodes suivantes :

$$u^0 = \text{Id}_E \quad , \quad u^1 = u \quad , \quad u^2 = u \circ u \quad , \quad \dots \quad , \quad u^n = u \circ u \circ \dots \circ u \quad .$$

Plus généralement, si $P(T) = a_0 + a_1T + \dots + a_nT^n$ est un polynôme à coefficients dans K , on pose

$$P(u) = a_0\text{Id}_E + a_1u + \dots + a_nu^n \quad .$$

On vérifie que l'on définit ainsi un morphisme d'algèbres de $K[T]$ dans $\mathcal{L}(E)$ (exercice). Cela veut dire entre autres que l'image d'un produit est le produit des images. Par exemple, le polynôme $P(T) = T^4 - 1$ se factorise sur \mathbf{R} en $P(T) = (T - 1)(T + 1)(T^2 + 1)$; on a donc, pour tout endomorphisme u d'un espace vectoriel réel E ,

$$u^4 - \text{Id}_E = (u - \text{Id}_E)(u + \text{Id}_E)(u^2 + \text{Id}_E) \quad .$$

Dans le cas complexe :

$$u^4 - \text{Id}_E = (u - \text{Id}_E)(u + \text{Id}_E)(u - i\text{Id}_E)(u + i\text{Id}_E) \quad .$$

On remarque que l'on peut faire cette construction avec n'importe quelle K -algèbre à la place de $\mathcal{L}(E)$, comme par exemple avec $\mathcal{M}_n(K)$.

Proposition 5.1.— *Pour tout endomorphisme u de E , il existe un polynôme non nul P tel que $P(u) = 0$.*

Démonstration. Les propriétés de la dimension d'un espace vectoriel que l'on a rappelées plus haut entraînent que les $n^2 + 1$ vecteurs $\text{Id}_E, u, u^2, \dots, u^{n^2}$ de $\mathcal{L}(E)$ sont liés. Il existe donc des scalaires non tous nuls $a_0, a_1, a_2, \dots, a_{n^2}$ tels que

$$a_0\text{Id}_E + a_1u + a_2u^2 + \dots + a_{n^2}u^{n^2} = 0 \quad .$$

Le polynôme $P(T) = a_0 + a_1T + a_2T^2 + \dots + a_{n^2}T^{n^2}$ n'est pas nul, et il vérifie $P(u) = 0$. ■

On en déduit de la même façon que pour toute matrice carrée M , il existe un polynôme non nul P tel que $P(M) = 0$.

Exercice 5.2. Trouver un polynôme non nul P qui annule la matrice $\begin{pmatrix} 0 & 1 \\ 2 & 3 \end{pmatrix}$.

6. Permutations

Soit f une application de l'ensemble $\{1, \dots, n\}$ dans lui-même. On la note (provisoirement) sous forme de tableau

$$\begin{pmatrix} 1 & 2 & \dots & n \\ f(1) & f(2) & \dots & f(n) \end{pmatrix} \quad .$$

ou tout simplement en écrivant uniquement la deuxième ligne On peut composer deux telles applications : si f correspond au tableau

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 1 \end{pmatrix}$$

et g au tableau

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 4 & 3 \end{pmatrix},$$

alors $g \circ f$ correspond au tableau

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 4 & 3 & 2 \end{pmatrix}.$$

On rappelle qu'une telle application est *injective* s'il n'existe pas deux éléments distincts de $\{1, \dots, n\}$ qui sont envoyés sur le même élément, c'est-à-dire si tous les éléments de la deuxième ligne du tableau sont distincts, *surjective* si tous les éléments de $\{1, \dots, n\}$ sont atteints par f , c'est-à-dire si tous les entiers entre 1 et n apparaissent sur la deuxième ligne. On a donc

$$f \text{ injective} \iff f \text{ surjective} \iff f \text{ bijective}.$$

On appelle *permutation* de $\{1, \dots, n\}$ toute application bijective de $\{1, \dots, n\}$ dans lui-même. L'ensemble des permutations de $\{1, \dots, n\}$ forment un *groupe*, que l'on note \mathcal{S}_n ; il n'est pas commutatif pour $n > 2$.

Exemple 6.1. Les éléments de \mathcal{S}_3 sont

$$\begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}.$$

On montre en général que \mathcal{S}_n a $n!$ éléments.

Un exemple important de permutation est celui des *transpositions*, c'est-à-dire des permutations qui n'échangent que deux éléments de $\{1, \dots, n\}$, laissant les autres inchangés. Pour tous $i, j \in \{1, \dots, n\}$, on note (i, j) la transposition qui envoie i sur j et j sur i . On remarque que

$$(j, i) = (i, j) = (i, j)^{-1}.$$

Les quatre premiers des éléments de \mathcal{S}_3 ci-dessus s'écrivent par exemple Id , $(1, 2)$, $(2, 3)$ et $(1, 3)$.

Soient σ une permutation et i et j des entiers tels que $1 \leq i < j \leq n$; on dit que σ inverse i et j si $\sigma(i) > \sigma(j)$. On peut compter le nombre d'inversions d'une permutation pour tous les $n(n-1)/2$ couples d'entiers $1 \leq i < j \leq n$.

Exemples 6.2. 1) L'identité n'a aucune inversion. La transposition $(1, 2)$ n'a qu'une seule inversion : celle de 1 et de 2. La transposition (i, j) a $2|j-i|-1$ inversions.

3) Les permutations $\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}$ et $\begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}$ ont chacune 2 inversions.

Définition 6.3.— Soit σ un élément de \mathcal{S}_n . On appelle le nombre

$$(-1)^{\text{nombre d'inversions de } \sigma}$$

la signature de σ . On dira qu'une permutation est paire si sa signature est 1, impaire dans le cas contraire.

Exemples 6.4.1) L'identité est paire. Toute transposition est impaire.

2) Dans \mathcal{S}_3 , il y a trois permutations impaires (qui sont des transpositions), et trois permutations paires (dont l'identité).

Il est clair que le produit

$$\prod_{1 \leq i < j \leq n} \frac{\sigma(j) - \sigma(i)}{j - i}$$

est du signe de $\varepsilon(\sigma)$. Son carré vaut

$$\prod_{1 \leq i < j \leq n} \left(\frac{\sigma(j) - \sigma(i)}{j - i} \right)^2 = \frac{\prod_{i \neq j} (\sigma(j) - \sigma(i))}{\prod_{i \neq j} (j - i)},$$

qui vaut 1 car les termes du numérateur sont les mêmes que ceux du dénominateur, écrits dans un ordre différent. On en déduit

$$\varepsilon(\sigma) = \prod_{1 \leq i < j \leq n} \frac{\sigma(j) - \sigma(i)}{j - i}.$$

Proposition 6.5.— La signature $\varepsilon : \mathcal{S} \rightarrow \{-1, 1\}$ est un morphisme de groupes ; en d'autres termes, la signature d'un produit est le produit des signatures.

Démonstration. Soient σ et τ des permutations ; comme $\frac{\sigma(j) - \sigma(i)}{j - i} = \frac{\sigma(i) - \sigma(j)}{i - j}$, cette quantité ne dépend pas de l'ordre de i et j ; on a donc

$$\varepsilon(\sigma) = \prod_{1 \leq i < j \leq n} \frac{\sigma(\tau(j)) - \sigma(\tau(i))}{\tau(j) - \tau(i)},$$

d'où

$$\varepsilon(\sigma \circ \tau) = \prod_{1 \leq i < j \leq n} \frac{\sigma(\tau(j)) - \sigma(\tau(i))}{j - i} = \prod_{1 \leq i < j \leq n} \left(\frac{\sigma(\tau(j)) - \sigma(\tau(i))}{\tau(j) - \tau(i)} \right) \left(\frac{\tau(j) - \tau(i)}{j - i} \right) = \varepsilon(\sigma)\varepsilon(\tau),$$

ce qui montre la proposition. ■

On peut remarquer que les éléments de \mathcal{S}_3 peuvent s'écrire

$$\text{Id} \quad , \quad (1, 2) \quad , \quad (2, 3) \quad , \quad (1, 3) \quad , \quad (1, 2)(2, 3) \quad , \quad (1, 3)(2, 3) \quad ;$$

en d'autres termes, ils sont tous produits de transpositions. La proposition suivante généralise ce fait.

Proposition 6.6.— *Toute permutation se décompose en produit de transpositions. Cette décomposition n'est pas unique, mais toute décomposition d'une permutation paire (resp. impaire) a un nombre pair (resp. impair) de termes.*

Démonstration. Plutôt que de démontrer l'existence de la décomposition en général, montrons le principe de la démonstration sur l'exemple $\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 4 & 1 & 5 & 3 \end{pmatrix}$. Le principe est de multiplier σ à gauche par des transpositions de façon à avoir à chaque fois un élément fixe de plus. Ici $\sigma(1) = 2$, et on compose par $(1, 2)$: $(1, 2)\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 4 & 2 & 5 & 3 \end{pmatrix}$. On continue en oubliant le 1 ; puisque $\sigma(2) = 4$, on compose par $(2, 4)$:

$$(2, 4)(1, 2)\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 2 & 4 & 5 & 3 \end{pmatrix} \quad (3, 4)(2, 4)(1, 2)\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 2 & 3 & 5 & 4 \end{pmatrix} = (4, 5) ,$$

de sorte que $\sigma = (1, 2)(2, 4)(3, 4)(4, 5)$.

Cette décomposition n'est pas unique ; on a par exemple :

$$(2 \ 3 \ 1) = (1, 2)(2, 3) = (2, 3)(1, 3) = (2, 3)(1, 2)(1, 2)(1, 3) ;$$

cependant, pour toute décomposition $\sigma = \tau_1 \circ \dots \circ \tau_r$, on a $\varepsilon(\sigma) = \varepsilon(\tau_1) \cdots \varepsilon(\tau_r) = (-1)^r$, de sorte que r a même parité de σ . ■

7. Formes multilinéaires alternées

Nous avons associé à tout endomorphisme u d'un espace vectoriel une matrice, mais celle-ci dépend du choix d'une base. Notre but est d'associer à u un invariant, c'est-à-dire quelque chose qui ne dépende d'aucune autre donnée que celle de u .

Définition 7.1.— *Soit E un K -espace vectoriel. On appelle forme m -multilinéaire sur E une application $F : E^m \rightarrow K$ qui est linéaire en chaque variable. On dit qu'une telle forme est alternée si*

$$F(x_1, \dots, x_m) = 0$$

dès que deux des x_j sont égaux.

Exemple 7.2. Soit $F : \mathbf{R}^2 \times \mathbf{R}^2 \rightarrow \mathbf{R}$ une forme bilinéaire sur \mathbf{R}^2 . Posons $e_1 = (1, 0)$ et $e_2 = (0, 1)$; on a

$$\begin{aligned} F((x_{11}, x_{21}), (x_{12}, x_{22})) &= F(x_{11}e_1 + x_{21}e_2, x_{12}e_1 + x_{22}e_2) \\ &= ax_{11}x_{12} + bx_{11}x_{22} + cx_{21}x_{12} + dx_{21}x_{22} , \end{aligned}$$

avec $a = F(e_1, e_1)$, $b = F(e_1, e_2)$, $c = F(e_2, e_1)$ et $d = F(e_2, e_2)$. Inversement, toute application de cette forme est bilinéaire. Elle est alternée si et seulement si on a

$$0 = F((x_{11}, x_{21}), (x_{11}, x_{21})) = ax_{11}^2 + (b + c)x_{11}x_{21} + dx_{21}^2$$

pour tout (x_{11}, x_{21}) , c'est-à-dire si et seulement si $a = b + c = d = 0$. La forme F s'écrit alors

$$F((x_{11}, x_{21}), (x_{12}, x_{22})) = b(x_{11}x_{22} - x_{21}x_{12}) .$$

On reconnaît là le déterminant d'ordre 2 !

Proposition 7.3.— Soient E un K -espace vectoriel et $F : E^m \rightarrow K$ une application multilinéaire. Pour que F soit alternée, il faut et il suffit que l'on ait

$$F(x_1, \dots, x_i, \dots, x_j, \dots, x_m) = -F(x_1, \dots, x_j, \dots, x_i, \dots, x_m) .$$

pour tous x_1, \dots, x_m dans E et $1 \leq i < j \leq m$.

Démonstration. Si F vérifie la propriété de la proposition, on a

$$F(x_1, \dots, x, \dots, x, \dots, x_m) = -F(x_1, \dots, x, \dots, x, \dots, x_m) ,$$

c'est-à-dire $2F(x_1, \dots, x, \dots, x, \dots, x_m) = 0$. Comme K contient le corps \mathbf{Q} des rationnels, on peut multiplier cette égalité par $1/2$, et F est alternée. Réciproquement, si F est alternée, on a

$$\begin{aligned} 0 &= F(x_1, \dots, x_i + x_j, \dots, x_i + x_j, \dots, x_m) \\ &= F(x_1, \dots, x_i, \dots, x_i, \dots, x_m) + F(x_1, \dots, x_i, \dots, x_j, \dots, x_m) \\ &\quad + F(x_1, \dots, x_j, \dots, x_i, \dots, x_m) + F(x_1, \dots, x_j, \dots, x_j, \dots, x_m) \\ &= F(x_1, \dots, x_i, \dots, x_j, \dots, x_m) + F(x_1, \dots, x_j, \dots, x_i, \dots, x_m) , \end{aligned}$$

ce qui montre la propriété cherchée. ■

On peut exprimer la conclusion de cette proposition en disant que, pour toute transposition τ , on a

$$F(x_{\tau(1)}, \dots, x_{\tau(m)}) = -F(x_1, \dots, x_m) .$$

Cette propriété se généralise à toute permutation :

Proposition 7.4.— Soient E un K -espace vectoriel et $F : E^m \rightarrow K$ une application multilinéaire alternée.

a) Pour toute permutation σ dans \mathcal{S}_m , on a

$$F(x_{\sigma(1)}, \dots, x_{\sigma(m)}) = \varepsilon(\sigma)F(x_1, \dots, x_m) ;$$

b) $F(x_1, \dots, x_m)$ ne change pas lorsqu'on ajoute à l'un des x_i une combinaison linéaire des autres ;

c) $F(x_1, \dots, x_m)$ est nul si les x_i sont liés. En particulier, toute forme multilinéaire alternée est nulle si $m > \dim E$.

Démonstration. On décompose σ en produit de k transpositions $\sigma = \tau_k \circ \dots \circ \tau_1$, avec $\varepsilon(\sigma) = (-1)^k$; le point a) s'en déduit par récurrence sur k . On a

$$\begin{aligned} F(x_1, \dots, x_{i-1}, x_i + \sum_{j \neq i} \lambda_j x_j, x_{i+1}, \dots, x_n) = \\ F(x_1, \dots, x_n) + \sum_{j \neq i} \lambda_j F(x_1, \dots, x_j, \dots, x_j, \dots, x_n) = F(x_1, \dots, x_n), \end{aligned}$$

d'où b). Si les x_i sont liés, l'un d'eux est combinaison linéaire des autres. Par b), on peut ajouter à ce vecteur l'opposé de cette combinaison linéaire sans changer la valeur de F , qui est donc nulle. Ceci montre c). ■

8. Déterminants

Nous allons maintenant étudier les formes n -multilinéaires alternées sur un espace vectoriel E de dimension n . Soit F une telle forme; on veut développer » $F(x_1, \dots, x_n)$ comme dans l'exemple 7.2. On se donne pour cela une base $\mathcal{B} = (e_1, \dots, e_n)$, et on décompose chaque x_j dans cette base en

$$x_j = \sum_{i=1}^n x_{ij} e_i.$$

On a par multilinéarité

$$\begin{aligned} F(x_1, \dots, x_n) &= F\left(\sum_{i_1=1}^n x_{i_1 1} e_{i_1}, \sum_{i_2=1}^n x_{i_2 2} e_{i_2}, \dots, \sum_{i_n=1}^n x_{i_n n} e_{i_n}\right) \\ &= \sum_{1 \leq i_1, i_2, \dots, i_n \leq n} x_{i_1 1} x_{i_2 2} \cdots x_{i_n n} F(e_{i_1}, e_{i_2}, \dots, e_{i_n}). \end{aligned}$$

Puisque F est alternée, les termes pour lesquels deux des indices i_1, \dots, i_n sont les mêmes vont s'annuler. Resteront ceux pour lesquels ils sont tous distincts, c'est-à-dire ceux pour lesquels l'application $k \mapsto i_k$ est injective, donc une permutation. On en déduit

$$F(x_1, \dots, x_n) = \sum_{\sigma \in \mathcal{S}_n} x_{\sigma(1)1} \cdots x_{\sigma(n)n} F(e_{\sigma(1)}, e_{\sigma(2)}, \dots, e_{\sigma(n)}).$$

On utilise la proposition 7.4 pour conclure

$$F(x_1, \dots, x_n) = F(e_1, e_2, \dots, e_n) \sum_{\sigma \in \mathcal{S}_n} x_{\sigma(1)1} \cdots x_{\sigma(n)n} \varepsilon(\sigma).$$

Comment interpréter cette formule? Elle nous dit que toutes les formes n -multilinéaires alternées sur E sont proportionnelles à la forme $\det_{\mathcal{B}}$ définie par

$$(8.1) \quad \det_{\mathcal{B}}(x_1, \dots, x_n) = \sum_{\sigma \in \mathcal{S}_n} \varepsilon(\sigma) x_{\sigma(1)1} \cdots x_{\sigma(n)n}.$$

On l'appelle le *déterminant des vecteurs* x_1, \dots, x_n dans la base \mathcal{B} .

Exemple 8.2. Dans K^2 , le déterminant dans la base canonique \mathcal{B} est donné par la formule (on pourra comparer avec l'exemple 7.2)

$$\det_{\mathcal{B}}((x_{11}, x_{21}), (x_{12}, x_{22})) = x_{11}x_{22} - x_{21}x_{12} .$$

Proposition 8.3.— *L'application $\det_{\mathcal{B}}$ est une forme n -multilinéaire alternée qui vérifie $\det_{\mathcal{B}}(e_1, \dots, e_n) = 1$. Si $F : E^n \rightarrow K$ est une forme multilinéaire alternée, on a, pour tous x_1, \dots, x_n dans E ,*

$$F(x_1, \dots, x_n) = F(e_1, \dots, e_n) \det_{\mathcal{B}}(x_1, \dots, x_n) .$$

Démonstration. Il n'est pas difficile de montrer que chacune des expressions $\varepsilon(\sigma)x_{\sigma(1)1} \cdots x_{\sigma(n)n}$ qui intervient dans le déterminant est n -multilinéaire en (x_1, \dots, x_n) (exercice). Leur somme l'est donc aussi. Montrons maintenant que le déterminant est alterné. Calculons pour cela $\det_{\mathcal{B}}(x_1, \dots, x_j, \dots, x_i, \dots, x_n)$, où x_j est à la i -ième place et x_i à la j -ième. Cela vaut, en notant τ la transposition (i, j) ,

$$\begin{aligned} & \sum_{\sigma \in \mathcal{S}_n} \varepsilon(\sigma)x_{\sigma(1)1} \cdots x_{\sigma(i)j} \cdots x_{\sigma(j)i} \cdots x_{\sigma(n)n} \\ &= \sum_{\sigma \in \mathcal{S}_n} \varepsilon(\sigma)x_{\sigma \circ \tau(1)1} \cdots x_{\sigma \circ \tau(j)j} \cdots x_{\sigma \circ \tau(i)i} \cdots x_{\sigma \circ \tau(n)n} \\ &= \sum_{\sigma \in \mathcal{S}_n} -\varepsilon(\sigma \circ \tau)x_{\sigma \circ \tau(1)1} \cdots x_{\sigma \circ \tau(n)n} = -\det_{\mathcal{B}}(x_1, \dots, x_n) , \end{aligned}$$

ce qui montre bien, par la proposition 7.3, que $\det_{\mathcal{B}}$ est alternée.

Lorsqu'on évalue $\det_{\mathcal{B}}$ sur les vecteurs de \mathcal{B} , le coefficient x_{ij} vaut 1 si $i = j$, et 0 sinon. Un terme de la somme qui définit $\det_{\mathcal{B}}(e_1, \dots, e_n)$ ne peut être non nul que si $\sigma(i) = i$ pour tout i , auquel cas σ est l'identité et le terme vaut 1. Ce déterminant vaut donc 1. Le reste de la proposition a été montré avant son énoncé. ■

Le théorème suivant regroupe les propriétés essentielles des déterminants.

Théorème 8.4.— *Soit \mathcal{B} une base d'un espace vectoriel E de dimension n .*

- a) $\det_{\mathcal{B}}(x_1, \dots, x_n)$ est linéaire en chaque x_i ;
- b) $\det_{\mathcal{B}}(x_1, \dots, x_n)$ s'annule dès que deux des x_i sont égaux ;
- c) pour toute permutation σ , on a

$$\det_{\mathcal{B}}(x_{\sigma(1)}, \dots, x_{\sigma(n)}) = \varepsilon(\sigma) \det_{\mathcal{B}}(x_1, \dots, x_n) ;$$

en particulier, le déterminant change de signe lorsqu'on échange deux des x_i ;

- d) $\det_{\mathcal{B}}(x_1, \dots, x_n)$ ne change pas lorsqu'on ajoute à l'un des x_i une combinaison linéaire des autres ;

e) si \mathcal{B}' est une autre base de E , on a

$$\det_{\mathcal{B}'}(x_1, \dots, x_n) = \det_{\mathcal{B}'}(e_1, \dots, e_n) \det_{\mathcal{B}}(x_1, \dots, x_n) ;$$

f) pour que $\det_{\mathcal{B}}(x_1, \dots, x_n)$ soit non nul, il faut et il suffit que (x_1, \dots, x_n) soit une base de E .

Démonstration. Les propriétés a) et b) constituent la définition d'une application multilinéaire alternée; c) et d) ont été montrés en 7.4, et e) en 8.3. Montrons f) : si (x_1, \dots, x_n) est une base \mathcal{B}' , alors $\det_{\mathcal{B}}(x_1, \dots, x_n)$ n'est pas nul par e); réciproquement, si ces vecteurs sont liés, leur déterminant est nul par la proposition 7.4.c). ■

Remarque 8.5. On peut donner l'interprétation intuitive suivante du déterminant dans \mathbf{R}^3 , relativement à la base canonique \mathcal{B} . Si x_1, x_2, x_3 sont des vecteurs de \mathbf{R}^3 , le réel $|\det_{\mathcal{B}}(x_1, x_2, x_3)|$ est le *volume* du parallélépipède construit sur les vecteurs x_1, x_2, x_3 . On retrouve bien le fait que pour que le déterminant soit nul, il faut et il suffit que les vecteurs soient liés, c'est-à-dire contenus dans un plan (le parallélépipède est alors « plat »).

9. Déterminant d'un endomorphisme

Comme promis, on va associer à tout endomorphisme d'un espace vectoriel un nombre qui ne dépend d'aucune autre donnée.

Théorème 9.1.— Soit u un endomorphisme d'un espace vectoriel E .

a) Le scalaire $\det_{\mathcal{B}}(u(e_1), \dots, u(e_n))$ ne dépend pas du choix de la base $\mathcal{B} = (e_1, \dots, e_n)$. On l'appelle le déterminant de l'endomorphisme u , et on le note $\det(u)$.

b) Pour tous x_1, \dots, x_n dans E et toute base \mathcal{B} de E , on a

$$\det_{\mathcal{B}}(u(x_1), \dots, u(x_n)) = \det(u) \det_{\mathcal{B}}(x_1, \dots, x_n) .$$

Démonstration. Pour toute base $\mathcal{B} = (e_1, \dots, e_n)$ de E , nous poserons

$$\delta(\mathcal{B}) = \det_{\mathcal{B}}(u(e_1), \dots, u(e_n)) .$$

Il est clair que l'application $(x_1, \dots, x_n) \mapsto \det_{\mathcal{B}}(u(x_1), \dots, u(x_n))$ est n -multilinéaire alternée; on peut lui appliquer la proposition 8.3 :

$$\det_{\mathcal{B}}(u(x_1), \dots, u(x_n)) = \delta(\mathcal{B}) \det_{\mathcal{B}}(x_1, \dots, x_n) ,$$

Il suffit maintenant de démontrer que $\delta(\mathcal{B})$ ne dépend pas de \mathcal{B} . Prenons une autre base \mathcal{B}' et posons $\lambda = \det_{\mathcal{B}'}(e_1, \dots, e_n)$; le théorème 8.4.e) entraîne

$$1 = \det_{\mathcal{B}'}(e'_1, \dots, e'_n) = \lambda \det_{\mathcal{B}}(e'_1, \dots, e'_n) ,$$

et

$$\delta(\mathcal{B}') = \det_{\mathcal{B}'}(u(e'_1), \dots, u(e'_n)) = \lambda \det_{\mathcal{B}}(u(e'_1), \dots, u(e'_n)) = \lambda \delta(\mathcal{B}) \det_{\mathcal{B}}(e'_1, \dots, e'_n) = \delta(\mathcal{B}) ,$$

ce qui termine la démonstration. ■

Exemple 9.2. Le déterminant de l'identité est 1 ; plus généralement, le déterminant de l'homothétie de rapport a dans un espace vectoriel E de dimension n , c'est-à-dire de $a\text{Id}_E$, est a^n .

Théorème 9.3.— Soient u et v des endomorphismes d'un espace vectoriel E .

a) On a $\det(u \circ v) = \det(u) \det(v) = \det(v \circ u)$.;

b) Pour que u soit bijectif, il faut et il suffit que son déterminant soit non nul. On a alors $\det u^{-1} = \frac{1}{\det u}$.

Démonstration. Soit $\mathcal{B} = (e_1, \dots, e_n)$ une base de E . On a

$$\det(u \circ v) = \det_{\mathcal{B}}(u(v(e_1)), \dots, u(v(e_n))) = \det(u) \det_{\mathcal{B}}(v(e_1), \dots, v(e_n)) = \det(u) \det(v),$$

ce qui montre a).

Pour qu'un endomorphisme u soit bijectif, il faut et il suffit que $(u(e_1), \dots, u(e_n))$ soit encore une base de E . Comme $\det_{\mathcal{B}}(u(e_1), \dots, u(e_n)) = \det(u)$ par définition, cela a lieu par le théorème 8.4.f) exactement lorsque $\det(u)$ n'est pas nul. La formule donnant le déterminant de l'inverse résulte de a). ■

Remarque 9.4. On peut poursuivre l'interprétation intuitive du déterminant dans \mathbf{R}^3 de la remarque 8.5. La valeur absolue du déterminant d'un endomorphisme u de \mathbf{R}^3 est le volume du parallélépipède construit sur les vecteurs $u(\vec{i}), u(\vec{j}), u(\vec{k})$. Le théorème 9.1.b) dit que lorsqu'on prend l'image d'un parallélépipède par u , son volume est multiplié par $|\det(u)|$. Plus généralement, on peut montrer que u multiplie tous les volumes (des sous-ensembles raisonnables de \mathbf{R}^3) par $|\det(u)|$. Si u n'est pas bijectif, son image est contenue dans un plan, et le volume d'une image par u est nul; on retrouve une moitié de la proposition 9.3.

On notera que la fonction $\det : \mathcal{L}(E) \rightarrow K$ que l'on a définie n'est pas linéaire! En particulier, le déterminant d'une somme n'est pas en général la somme des déterminants. Le comportement par la multiplication par un scalaire n'est pas celui d'une application linéaire, puisque

$$\det(au) = a^n \det(u).$$

10. Déterminant d'une matrice

Soit $M = (a_{ij})_{1 \leq i, j \leq n}$ une matrice carrée. On pose, comme en (8.1),

$$(10.1) \quad \det M = \sum_{\sigma \in \mathcal{S}_n} \varepsilon(\sigma) a_{\sigma(1)1} \cdots a_{\sigma(n)n}.$$

On note aussi $|M|$ le déterminant de M . De façon équivalente, c'est le *déterminant des vecteurs colonnes de M* (c'est-à-dire les vecteurs de K^n dont les composantes dans la base canonique sont les coefficients des colonnes de M), ou encore le déterminant de tout endomorphisme dont la matrice dans une base est M . Cela permet de traduire dans le cadre des matrices les propriétés du déterminant que l'on a déjà démontrées.

Exemples 10.2. 1) Le déterminant de la matrice identité I_n est 1 (cf. ex. 9.2).

2) Pour des matrices d'ordre 2 et 3, on a

$$\det M = a_{11}a_{22} - a_{21}a_{12}$$

$$\det M = a_{11}a_{22}a_{33} - a_{21}a_{12}a_{33} - a_{11}a_{32}a_{23} - a_{31}a_{22}a_{13} + a_{21}a_{32}a_{13} + a_{31}a_{12}a_{23} .$$

En général, le déterminant d'une matrice d'ordre n comporte $n!$ termes (c'est-à-dire énormément).

Définition 10.3.— Soit M une matrice (carrée ou non). On appelle transposée de M , et l'on note tM , la matrice dont les lignes sont les colonnes de M .

Par exemple, si

$$M = \begin{pmatrix} 1 & 2 \\ -3 & 4 \\ 5 & 0 \end{pmatrix} ,$$

on a

$${}^tM = \begin{pmatrix} 1 & -3 & 5 \\ 2 & 4 & 0 \end{pmatrix} .$$

Proposition 10.4.— Le déterminant d'une matrice carrée est égal à celui de sa transposée.

C'est un bon exercice de vérifier la proposition directement sur les formules ci-dessus dans le cas $n = 2$ ou 3 ; on voit que les mêmes termes apparaissent dans le développement de chacun des deux déterminants, seul l'ordre des termes dans les produits change. C'est ce qui se passe en général.

Démonstration de la proposition. Notons $M = (a_{ij})_{1 \leq i, j \leq n}$ et ${}^tM = (b_{ij})_{1 \leq i, j \leq n}$, avec $b_{ij} = a_{ji}$; on a, pour toute permutation σ ,

$$\varepsilon(\sigma)b_{\sigma(1)1} \cdots b_{\sigma(n)n} = \varepsilon(\sigma)a_{1\sigma(1)} \cdots a_{n\sigma(n)} = \varepsilon(\sigma^{-1})a_{\sigma^{-1}(1)1} \cdots a_{\sigma^{-1}(n)n}$$

(les termes de ces deux derniers produits sont simplement réarrangés dans un ordre différent, et $\varepsilon(\sigma) = \varepsilon(\sigma^{-1})$). On a donc

$$\det {}^tM = \sum_{\sigma \in \mathcal{S}_n} \varepsilon(\sigma^{-1})a_{\sigma^{-1}(1)1} \cdots a_{\sigma^{-1}(n)n} = \det M ,$$

puisque l'on peut aussi bien indexer la somme par σ^{-1} . ■

Le théorème suivant regroupe les propriétés fondamentales des déterminants des matrices.

Théorème 10.5.— Soit M une matrice carrée d'ordre n .

a) $\det M$ dépend linéairement de chaque ligne et de chaque colonne;

b) $\det M$ s'annule dès que deux colonnes, ou deux lignes, sont égales ;

c) $\det M$ ne change pas lorsqu'on ajoute à l'une des colonnes, ou à l'une des lignes, une combinaison linéaire des autres ;

d) $\det M$ change de signe si l'on échange deux lignes ou deux colonnes.

Démonstration. On obtient toutes les assertions du théorème relatives aux colonnes en interprétant le déterminant de M comme le déterminant de ses vecteurs colonnes et en appliquant le théorème 8.4. Les assertions relatives aux lignes sont des assertions relatives aux colonnes de la transposée, donc se déduisent de la proposition 10.4. ■

Passons maintenant aux propriétés multiplicatives.

Théorème 10.6.— Soient M et N des matrices carrées d'ordre n .

a) $\det(MN) = \det M \det N$.

b) pour que M soit inversible, il faut et il suffit que son déterminant soit non nul.

On a alors $\det M^{-1} = \frac{1}{\det M}$.

Démonstration. Soit \mathcal{B} une base d'un espace vectoriel E de dimension n , et soient u et v des endomorphismes de E de matrices respectives M et N dans la base \mathcal{B} . On a vu que $\det M = \det(u)$ et $\det N = \det(v)$. La matrice dans la base \mathcal{B} de $u \circ v$ est MN , d'où

$$\det(MN) = \det(u \circ v) = \det(u) \det(v) = \det M \det N ,$$

grâce au théorème 9.3.a). Ceci montre a).

Si M est inversible, on a par a)

$$1 = \det I_n = \det(MM^{-1}) = \det M \det M^{-1} ,$$

ce qui montre que le déterminant de M n'est pas nul et que celui de M^{-1} est son inverse. Réciproquement, si $\det M$ n'est pas nul, l'endomorphisme u ci-dessus est bijectif par le théorème 9.3.b). On a $MM_{\mathcal{B}}(u^{-1}) = M_{\mathcal{B}}(u^{-1})M = I_n$, de sorte que la matrice $M_{\mathcal{B}}(u^{-1})$ est l'inverse de M ; ceci montre b). ■

C'est en fait cette proposition qui justifie la façon un peu théorique dont on a introduit le déterminant : essayez donc de démontrer directement le a) en partant de la définition !

Exemple 10.7. On vérifie que l'inverse d'une matrice $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ de déterminant non nul est donné par

$$M^{-1} = \frac{1}{\det M} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} .$$

Nous généraliserons plus bas cette formule (prop. 11.6).

Définition 10.8.— On dit que des matrices carrées M et N sont semblables s'il existe une matrice inversible P telle que $N = P^{-1}MP$.

(10.9) À cause de (4.7), pour que deux matrices soient semblables, il faut et il suffit qu'elles soient les matrices du même endomorphisme d'un espace vectoriel E dans des bases peut-être différentes.

Corollaire 10.10.— Des matrices semblables ont même déterminant.

Démonstration. On peut le voir soit en invoquant le théorème 10.6 :

$$\det N = \det(P^{-1}MP) = \det P^{-1} \det M \det P = \frac{1}{\det P} \det M \det P = \det M ,$$

soit en remarquant que M et N sont les matrices du même endomorphisme f d'un espace vectoriel E dans des bases peut-être différentes, et que leur déterminant vaut $\det(f)$. ■

Il n'est sans doute pas inutile de répéter que *le déterminant n'est pas linéaire !* On a $\det(\lambda M) = \lambda^n \det M$ et en général

$$\det(M + N) \neq \det M + \det N$$

11. Calcul et développements d'un déterminant

Dans la pratique, il est extrêmement rare que l'on calcule un déterminant par la formule (10.1), qui est bien trop compliquée. On a recours à diverses astuces. Considérons tout d'abord une matrice dont l'écriture *par blocs* est

$$M = \begin{pmatrix} M' & N \\ 0 & M'' \end{pmatrix} ,$$

où M' est une matrice carrée d'ordre p , M'' une matrice carrée d'ordre q (avec $p + q = n$), et N une matrice à p lignes et q colonnes.

Proposition 11.1.— Le déterminant de la matrice $M = \begin{pmatrix} M' & N \\ 0 & M'' \end{pmatrix}$ est $(\det M')(\det M'')$.

Démonstration. Notons a_{ij} les coefficients de M . On a

$$\det M = \sum_{\sigma \in \mathcal{S}_n} \varepsilon(\sigma) a_{\sigma(1)1} \cdots a_{\sigma(n)n} .$$

Pour qu'un terme soit non nul, il faut que $\sigma(1), \dots, \sigma(p) \in \{1, \dots, p\}$, d'où forcément $\sigma(p+1), \dots, \sigma(n) \in \{p+1, \dots, n\}$. Il existe alors des permutations $\sigma' \in \mathcal{S}_p$ et $\sigma'' \in \mathcal{S}_q$ telle que

$$\begin{aligned} \sigma(j) &= \sigma'(j) & \text{si } 1 \leq j \leq p , \\ \sigma(p+k) &= p + \sigma''(k) & \text{si } 1 \leq k \leq q . \end{aligned}$$

On a donc

$$\det M = \sum_{\sigma' \in \mathcal{S}_p, \sigma'' \in \mathcal{S}_q} \varepsilon(\sigma) a'_{\sigma'(1)1} \cdots a'_{\sigma'(p)p} a''_{\sigma''(1)1} \cdots a''_{\sigma''(q)q} .$$

Il reste à calculer la signature de σ . En décomposant σ' et σ'' en produit de transpositions, on voit que c'est $\varepsilon(\sigma')\varepsilon(\sigma'')$, d'où

$$\det M = \left(\sum_{\sigma' \in \mathcal{S}_p} \varepsilon(\sigma') a'_{\sigma'(1)1} \cdots a'_{\sigma'(p)p} \right) \left(\sum_{\sigma'' \in \mathcal{S}_q} \varepsilon(\sigma'') a''_{\sigma''(1)1} \cdots a''_{\sigma''(q)q} \right) = \det(M') \det(M'') ,$$

ce qui montre la proposition. ■

Exemples 11.2. 1) En combinant les propriétés du déterminant que l'on a vues, on obtient presque sans calcul

$$\begin{aligned} & \begin{vmatrix} 11 & 0 & 12 & 0 \\ 3 & 1 & 4 & 2 \\ 9 & 0 & 10 & 0 \\ 7 & 5 & 8 & 6 \end{vmatrix} = - \begin{vmatrix} 7 & 5 & 8 & 6 \\ 3 & 1 & 4 & 2 \\ 9 & 0 & 10 & 0 \\ 11 & 0 & 12 & 0 \end{vmatrix} = \begin{vmatrix} 6 & 5 & 8 & 7 \\ 2 & 1 & 4 & 3 \\ 0 & 0 & 10 & 9 \\ 0 & 0 & 12 & 11 \end{vmatrix} \\ & = \begin{vmatrix} 6 & 5 \\ 2 & 1 \end{vmatrix} \begin{vmatrix} 10 & 9 \\ 12 & 11 \end{vmatrix} = (6 \cdot 1 - 5 \cdot 2)(10 \cdot 11 - 9 \cdot 12) = -8 . \end{aligned}$$

2) On dit qu'une matrice $M = (a_{ij})$ est *triangulaire supérieure* si a_{ij} est nul pour $i > j$, c'est-à-dire si tous les coefficients en-dessous de la diagonale sont nuls. Si on découpe une telle matrice en blocs comme dans la proposition, ceux-ci sont encore triangulaires supérieurs. On en déduit, par récurrence sur la taille de la matrice, que *le déterminant d'une matrice triangulaire supérieure est le produit de ses termes diagonaux*. On peut aussi retrouver ce résultat directement à partir de la définition du déterminant. Les mêmes remarques s'appliquent aux matrices *triangulaires inférieures*, qui sont celles dont la transposée est triangulaire supérieure.

On s'intéresse toujours à une matrice $M = (a_{ij})_{1 \leq i, j \leq n}$. Notons M_{ij} la matrice carrée d'ordre $n - 1$ obtenue en supprimant la i ème ligne et la j ème colonne de M . On appelle *cofacteur d'indice (i, j)* le nombre $C_{ij} = (-1)^{i+j} \det M_{ij}$. Par exemple, lorsque $n = 2$, on a $C_{11} = a_{22}$, $C_{12} = -a_{21}$, $C_{21} = -a_{12}$ et $C_{22} = a_{11}$. Le théorème suivant permet de ramener le calcul d'un déterminant d'ordre n au calcul de n déterminants d'ordre $n - 1$. On peut ainsi, théoriquement parce que c'est en fait très long, calculer par cette méthode les déterminants d'ordre quelconque.

Théorème 11.3.— *Pour chaque $i, j = 1, \dots, n$, on a les formules*

$$\det M = a_{1j} C_{1j} + \cdots + a_{nj} C_{nj}$$

(développement suivant la j ème colonne) et

$$\det M = a_{i1} C_{i1} + \cdots + a_{in} C_{in}$$

(développement suivant la i ième ligne).

Démonstration. Démontrons la première formule (la seconde se démontre de la même façon ou s'obtient à partir de la première en considérant la transposée de M). Le déterminant est multilinéaire; en particulier, il est linéaire en la i ième colonne. Si N_{ij} est la matrice obtenue à partir de M en remplaçant tous les éléments de la j ième colonne, sauf celui de la i ième ligne, par 0, on obtient

$$\det M = \det N_{1j} + \cdots + \det N_{nj} .$$

Si l'on fait passer la j ième colonne de N_{ij} à la première place, le déterminant est multiplié par $(-1)^{j-1}$ (cela revient à faire $j-1$ transpositions de colonnes); de même, si l'on fait passer la i ième ligne de N_{ij} à la première place, le déterminant est multiplié par $(-1)^{i-1}$. On a donc

$$\det N_{ij} = (-1)^{(i-1)+(j-1)} \begin{vmatrix} a_{ij} & a_{i1} & \cdots & a_{in} \\ 0 & & & \\ \cdot & & M_{ij} & \\ \cdot & & & \\ \cdot & & & \\ 0 & & & \end{vmatrix} .$$

Par la proposition 11.1, cela vaut $a_{ij}C_{ij}$, et la proposition est démontrée. ■

Exemple 11.4. Il est avantageux de développer selon une ligne ou une colonne où il y a beaucoup de zéros :

$$\begin{vmatrix} 1 & 2 & 3 & 4 \\ 4 & 0 & 5 & 6 \\ 0 & 0 & 3 & 0 \\ 1 & 6 & 4 & 2 \end{vmatrix} = 3 \begin{vmatrix} 1 & 2 & 4 \\ 4 & 0 & 6 \\ 1 & 6 & 2 \end{vmatrix} = -3 \cdot 4 \begin{vmatrix} 2 & 4 \\ 6 & 2 \end{vmatrix} - 3 \cdot 6 \begin{vmatrix} 1 & 2 \\ 1 & 6 \end{vmatrix} = 240 - 72 = 168 .$$

(11.5) Puisqu'on a n^2 cofacteurs, on peut en faire une matrice \tilde{M} dont le coefficient d'ordre (i, j) est C_{ij} . On l'appelle la *comatrice* de M . La proposition suivante généralise la formule de l'exemple 10.7.

Proposition 11.6.— On a

$$M {}^t\tilde{M} = {}^t\tilde{M} M = (\det M) I_n ;$$

en particulier, si M est inversible, son inverse est donné par

$$M^{-1} = \frac{1}{\det M} {}^t\tilde{M} .$$

Démonstration. Le terme d'indice (i, j) du produit $M {}^t\tilde{M}$ est $\sum_{k=1}^n a_{ik}C_{jk}$. Lorsque $i = j$, on reconnaît la deuxième formule du théorème 11.3 : cela vaut $\det M$. Lorsque $i \neq j$, c'est aussi le développement suivant la j ième ligne du déterminant de la matrice obtenue à partir de M en recopiant la i ième ligne à la ligne j . Cette matrice a deux lignes égales, donc son déterminant est nul. On a donc $M {}^t\tilde{M} = (\det M)I_n$. L'autre égalité se démontre de la même façon en utilisant le développement suivant les colonnes. Cela démontre la proposition. ■

Il faut signaler que cette méthode de calcul de l'inverse d'une matrice n'a qu'un intérêt théorique : elle nécessite de calculer un déterminant d'ordre n et n^2 déterminants d'ordre $n - 1$. Dans la pratique, il existe des méthodes bien plus performantes.

12.

EXERCICES

(12.1) Soit u l'endomorphisme de \mathbf{R}^3 de matrice $\begin{pmatrix} 4 & 2 & -4 \\ -6 & -4 & 6 \\ -1 & -1 & 1 \end{pmatrix}$ dans la base canonique (e_1, e_2, e_3) .

a) Déterminer le noyau et l'image de u .

b) Soient $u_1 = e_1 - e_2$, $u_2 = 2e_2 + e_3$ et $u_3 = e_1 + e_3$. Montrer qu'il s'agit d'une base de \mathbf{R}^3 et déterminer la matrice de u dans cette base.

c) Soit n un entier naturel ; donner la matrice de u^n dans la base canonique.

(12.2) Soit E l'espace vectoriel des polynômes en T de degré 3. On considère les applications de E dans E suivantes :

$$u : P \mapsto P' + P'' + TP(0) \quad \text{et} \quad g : P \mapsto T^3P(0) - P'$$

Montrer qu'il s'agit d'applications linéaires. Déterminer leurs rang, noyau et image, ainsi que leur application inverse si elle existe.

(12.3) Dans $\mathcal{M}_3(\mathbf{R})$, on considère la matrice $M = \begin{pmatrix} -5 & 3 & 3 \\ -8 & 6 & 4 \\ -1 & 1 & 3 \end{pmatrix}$. Soit u l'endomorphisme de \mathbf{R}^3 de matrice M dans la base canonique (e_1, e_2, e_3) .

a) Déterminer les images par u des vecteurs $u = e_1 + e_2$, $v = e_2 - e_3$ et $w = e_1 + 2e_2 + e_3$.

b) Montrer que ces trois vecteurs forment une base de \mathbf{R}^3 et donner la matrice N associée à u dans cette base.

c) Calculer N^n , puis M^n , pour tout $n \in \mathbf{N}$.

(12.4) Calculer la signature de la permutation σ de $\{1, \dots, n\}$ définie par

$$\sigma(1) = 2, \quad \sigma(2) = 3, \quad \dots, \quad \sigma(n-1) = n, \quad \sigma(n) = 1.$$

(12.5) Calculez le déterminant

$$\begin{vmatrix} 5 & 4 & 2 & 1 \\ 2 & 3 & 1 & -2 \\ -5 & -7 & -3 & 9 \\ 1 & -2 & -1 & 4 \end{vmatrix}.$$

(12.6) Calculer les déterminants $n \times n$ suivants

$$D_n(x) = \begin{vmatrix} 1+x^2 & x & 0 & \dots & 0 \\ x & 1+x^2 & x & & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & & & & x \\ 0 & \dots & 0 & x & 1+x^2 \end{vmatrix} \quad \text{et} \quad V_n(a_1, \dots, a_n) = \begin{vmatrix} 1 & 1 & 1 \\ a_1 & a_2 & a_n \\ a_1^2 & a_2^2 & a_n^2 \\ \vdots & \vdots & \vdots \\ a_1^{n-1} & a_2^{n-1} & a_n^{n-1} \end{vmatrix}.$$

Pour le premier procédez par récurrence. Pour le second, remarquez que $V_n(a_1, \dots, a_{n-1}, x)$ est un polynôme de degré $n - 1$ en x . Quelles sont ses racines? Le coefficient du terme de plus haut degré? Déduisez-en la forme de $V_n(a_1, \dots, a_n)$.

(12.7) Développer le déterminant suivant sous la forme d'un produit de facteurs linéaires en x :

$$\begin{vmatrix} x + 2 & 2x + 3 & 3x + 4 \\ 2x + 3 & 3x + 4 & 4x + 5 \\ 3x + 5 & 5x + 8 & 10x + 17 \end{vmatrix}.$$

(12.8) Soit $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ une matrice fixée de $\mathcal{M}_2(\mathbf{R})$. On définit une application linéaire $T : \mathcal{M}_2(\mathbf{R}) \rightarrow \mathcal{M}_2(\mathbf{R})$ en posant $T(M) = AM$ pour toute matrice M dans $\mathcal{M}_2(\mathbf{R})$.

- a) Montrer que T est un endomorphisme de $\mathcal{M}_2(\mathbf{R})$.
- b) Déterminer la matrice de T dans la base canonique de $\mathcal{M}_2(\mathbf{R})$.
- c) Quel est le déterminant de T ?

(12.9) Soit M une matrice d'ordre n *antisymétrique*, c'est-à-dire telle que ${}^tM = -M$. Montrer que M n'est pas inversible si n est impair.

II. RÉDUCTION DES ENDOMORPHISMES

Pour décrire un endomorphisme u d'un espace vectoriel E , on cherche une base de E dans laquelle la matrice de u soit aussi simple que possible. Pour diverses raisons, on voudrait que cette matrice soit *diagonale*, c'est-à-dire que les coefficients en dehors de la diagonale soient nuls. En d'autres termes, les vecteurs de cette base sont tels que leur image par u leur est colinéaire. On les appelle des *vecteurs propres* de u . L'avantage d'avoir une matrice diagonale est qu'il est alors plus facile de faire des calculs faisant intervenir u (le calcul des itérés de u devient par exemple très facile).

Il n'existe pas toujours de base de E formée de vecteurs propres de u . Par exemple, la rotation r_θ dans le plan réel euclidien orienté (*cf.* § 46) n'a pas de vecteur propre si $\theta \not\equiv 0 \pmod{\pi}$ puisque aucun vecteur non nul x n'est colinéaire à $r_\theta(x)$. Le problème dans ce cas se résout en passant du corps \mathbf{R} au corps \mathbf{C} : nous verrons qu'un endomorphisme d'un espace vectoriel complexe a toujours un vecteur propre. Néanmoins, même sur \mathbf{C} , il peut ne pas y avoir « assez » de vecteurs propres pour former une base ; c'est le cas pour l'endomorphisme de \mathbf{C}^2 de matrice $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$: tous les vecteurs propres sont colinéaires à e_1 , donc ne peuvent former une base. On cherchera pour ces endomorphismes une base dans laquelle leur matrice est aussi simple que possible, c'est-à-dire aussi proche que possible d'une forme diagonale.

13. Valeurs propres, vecteurs propres

On travaille toujours avec un espace vectoriel E sur un corps K , qui est un corps commutatif contenant le corps \mathbf{Q} des rationnels (le plus souvent, K sera soit le corps \mathbf{R} des nombres réels, soit le corps \mathbf{C} des nombres complexes).

Définition 13.1.— Soit u un endomorphisme de E . S'il existe un scalaire λ et un vecteur non nul x de E vérifiant $u(x) = \lambda x$, on dit que λ est une valeur propre de u , et que x est un vecteur propre de u associé à λ .

On remarque que tout multiple non nul d'un vecteur propre de u est encore vecteur propre de u pour la même valeur propre.

Exemples 13.2. 1) Comme on l'a vu dans l'introduction, une rotation du plan réel euclidien orienté (*cf.* § 46) n'a en général pas de vecteur propre. En revanche, pour une symétrie orthogonale s_D par rapport à une droite D , il y a deux valeurs propres : les vecteurs non nuls de D sont inchangés par s_D donc sont propres pour la valeur propre 1, les vecteurs non nuls perpendiculaires à D sont changés par s_D en leur opposé donc sont propres pour la valeur propre -1 .

2) Soit E l'espace vectoriel des fonctions $\mathbf{R} \rightarrow \mathbf{R}$ indéfiniment dérivables. L'application $u : E \rightarrow E$ qui à une fonction associe sa dérivée est un endomorphisme de E . La fonction

$f : t \mapsto e^{\lambda t}$ est vecteur propre de valeur propre λ , puisque $u(f) = f' = \lambda f$; tout réel est donc valeur propre de u .

Définition 13.3.— Soient u un endomorphisme de E et λ une valeur propre de u . On appelle sous-espace propre associé à λ le sous-espace vectoriel $\text{Ker}(\lambda \text{Id}_E - u)$ de E ; on le note E_λ .

Remarquons que l'égalité $u(x) = \lambda x$ est équivalente à $(\lambda \text{Id}_E - u)(x) = 0$; le sous-espace propre E_λ est donc l'ensemble de tous les vecteurs propres associés à λ et du vecteur nul; il est en particulier non nul.

Nous supposons à partir de maintenant E de dimension finie (ce qui exclut l'exemple 13.2.2).

Proposition 13.4.— Soient u un endomorphisme de E et λ un scalaire. Pour que λ soit valeur propre de u , il faut et il suffit que $\det(\lambda \text{Id}_E - u) = 0$.

Démonstration. Si λ est valeur propre de u , l'endomorphisme $\lambda \text{Id}_E - u$ n'est pas injectif, donc pas bijectif; son déterminant est donc nul (th. 9.3.b)). Inversement, si ce déterminant est nul, $\lambda \text{Id}_E - u$ n'est pas bijectif, donc pas injectif puisque E est de dimension finie; son noyau n'est alors pas nul, donc il existe des vecteurs propres pour λ . ■

Si E_1 et E_2 sont des sous-espaces vectoriels de E , on dit que E est somme directe de E_1 et E_2 , et l'on note $E = E_1 \oplus E_2$, si tout élément de x de E peut s'écrire de façon *unique* $x = x_1 + x_2$, avec $x_1 \in E_1$ et $x_2 \in E_2$. On peut exprimer cela en disant que l'application $\rho : E_1 \times E_2 \rightarrow E$ définie par $\rho(x_1, x_2) = x_1 + x_2$ est bijective. On généralise cela au cas de plusieurs sous-espaces vectoriels dans la définition suivante.

Définition 13.5.— On dit que des sous-espaces vectoriels E_1, \dots, E_p d'un espace vectoriel E sont en somme directe si l'application

$$\begin{aligned} \rho : E_1 \times \dots \times E_p &\longrightarrow E \\ (x_1, \dots, x_p) &\longmapsto x_1 + \dots + x_p \end{aligned}$$

est injective. Si ρ est bijective, on dit que E est somme directe de E_1, \dots, E_p , et on écrit $E = E_1 \oplus \dots \oplus E_p$.

Proposition 13.6.— Soient E_1, \dots, E_p des sous-espaces vectoriels d'un espace vectoriel E . On suppose $E = E_1 \oplus \dots \oplus E_p$.

a) On a $\dim E = \dim E_1 + \dots + \dim E_p$.

b) Si \mathcal{B}_j est une base de E_j , pour $j = 1, \dots, p$, alors la famille obtenue en juxtaposant $\mathcal{B}_1, \dots, \mathcal{B}_p$ est une base de E .

Démonstration. L'application ρ définie ci-dessus est un isomorphisme, de sorte que E a même dimension que $E_1 \times \dots \times E_p$, c'est-à-dire $\dim E_1 + \dots + \dim E_p$, ce qui montre a).

La famille obtenue en juxtaposant $\mathcal{B}_1, \dots, \mathcal{B}_p$ est génératrice; c'est donc une base puisque son cardinal est la dimension de E . ■

Proposition 13.7.— *Les sous-espaces propres d'un endomorphisme associés à des valeurs propres deux à deux distinctes sont en somme directe.*

Démonstration. Procédons par récurrence sur le nombre p de valeurs propres. Si $p = 1$, il n'y a rien à démontrer. Soient $\lambda_1, \dots, \lambda_p$ des valeurs propres de u deux à deux distinctes; il s'agit de montrer que l'application ρ définie ci-dessus est injective, c'est-à-dire que si $x_1 + \dots + x_p = 0$, avec $x_j \in E_{\lambda_j}$, alors $x_1 = \dots = x_p = 0$. Appliquons u ; on obtient

$$0 = u(x_1 + \dots + x_p) = \lambda_1 x_1 + \dots + \lambda_p x_p .$$

Si on soustrait $\lambda_1(x_1 + \dots + x_p)$ (qui est nul), on obtient

$$(\lambda_2 - \lambda_1)x_2 + \dots + (\lambda_p - \lambda_1)x_p = 0 ;$$

par l'hypothèse de récurrence, $E_{\lambda_2}, \dots, E_{\lambda_p}$ sont en somme directe, donc $(\lambda_j - \lambda_1)x_j$ est nul pour $j = 2, \dots, p$. Comme λ_1 est différent de λ_j , on obtient $x_2 = \dots = x_p = 0$, donc aussi $x_1 = 0$. ■

Corollaire 13.8.— *Le nombre de valeurs propres d'un endomorphisme de E est au plus la dimension de E .*

Démonstration. Si $\lambda_1, \dots, \lambda_p$ sont les valeurs propres, $\dim E = \dim E_{\lambda_1} + \dots + \dim E_{\lambda_p}$ (prop. 13.7 et 13.6.a). Comme la dimension des espaces propres est au moins 1, on a $p \leq \dim E$. ■

14. Endomorphismes diagonalisables

Définition 14.1.— *On dit qu'un endomorphisme u de E est diagonalisable s'il existe une base de E formée de vecteurs propres de E .*

Si (e_1, \dots, e_n) est une base de E formée de vecteurs propres de u , la matrice de u dans cette base est

$$\begin{pmatrix} \lambda_1 & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2 & 0 & \cdots & 0 \\ 0 & 0 & \cdot & & \cdot \\ \cdot & \cdot & & & \cdot \\ 0 & 0 & & & \lambda_n \end{pmatrix} ,$$

où λ_j est la valeur propre associée au vecteur propre e_j . Réciproquement, s'il existe une base de E dans laquelle sa matrice est diagonale, cette base est automatiquement formée de vecteurs propres et u est diagonalisable.

Exemple 14.2. Dans le plan réel euclidien orienté, une symétrie orthogonale s_D par rapport à une droite D est diagonalisable : la base du plan qui consiste en un vecteur non nul de D et un vecteur non nul perpendiculaire est formée de vecteurs propres de s_D (cf. ex. 13.2.1).

Proposition 14.3.— Soient u un endomorphisme de E et $\lambda_1, \dots, \lambda_p$ ses valeurs propres distinctes. Les conditions suivantes sont équivalentes :

- (i) u est diagonalisable ;
- (ii) $E = E_{\lambda_1} \oplus \dots \oplus E_{\lambda_p}$;
- (iii) $\dim E_{\lambda_1} + \dots + \dim E_{\lambda_p} \geq \dim E$.

Démonstration. L'application ρ définie ci-dessus est injective par la proposition 13.7 ; si la propriété (iii) est vérifiée, ρ est bijective, ce qui montre (ii). Si (ii) est vérifiée, on obtient une base de E en juxtaposant une base de E_{λ_1} , une base de E_{λ_2} , etc. Cette base est formée de vecteurs propres de u , donc u est diagonalisable, ce qui prouve (i). Enfin, si (i) est vérifiée, il existe une base de E formée de vecteurs propres ; la dimension de E_{λ_j} est alors plus grande que le nombre de vecteurs de cette base dont λ_j est la valeur propre associée. On en déduit (iii), ce qui démontre la proposition. ■

Corollaire 14.4.— Soit E un espace vectoriel de dimension n . Tout endomorphisme de E qui a n valeurs propres distinctes est diagonalisable.

Démonstration. En effet, la propriété (iii) de la proposition est vérifiée puisque $p \geq n$ par hypothèse, et qu'un espace propre n'est jamais nul, donc est de dimension ≥ 1 . ■

Remarque 14.5. La réciproque du corollaire est évidemment fautive ; par exemple, une homothétie λId_E est diagonalisable puisque $E = E_\lambda$, mais n'a qu'une seule valeur propre. Il est utile de remarquer que pour qu'un endomorphisme u qui n'a qu'une seule valeur propre λ soit diagonalisable, il faut et il suffit que ce soit une homothétie (en effet, par la proposition 14.3, il faut que $E = E_\lambda$).

15. Polynôme caractéristique d'une matrice carrée

(15.1) Soit M une matrice carrée d'ordre n à coefficients dans K . Il lui est associé un endomorphisme u_M de l'espace vectoriel K^n de matrice M dans la base canonique ; en d'autres termes, on a, si $M = (a_{ij})_{1 \leq i, j \leq n}$,

$$u_M(x_1, \dots, x_n) = \left(\sum_{j=1}^n a_{1j} x_j, \dots, \sum_{j=1}^n a_{nj} x_j \right),$$

ou encore, si l'on représente un élément de K^n par un vecteur colonne X ,

$$u_M(X) = MX.$$

On appellera *valeurs propres*, *vecteurs propres*, *sous-espaces propres* de M les notions correspondantes pour l'endomorphisme u_M . Un scalaire λ est donc valeur propre de M s'il existe un vecteur colonne non nul X tel que $MX = \lambda X$; le vecteur X est vecteur propre de M .

Exemple 15.2. Le vecteur $\begin{pmatrix} 1 \\ 2 \end{pmatrix}$ est vecteur propre pour la matrice $M = \begin{pmatrix} 1 & 1 \\ -2 & 4 \end{pmatrix}$, associé à la valeur propre 3, puisque

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

De même, on dira que M est *diagonalisable* si u_M l'est; compte-tenu de (10.9), cela signifie que M est semblable à une matrice diagonale.

Exemples 15.3. 1) Dans le plan réel euclidien orienté \mathbf{R}^2 , la matrice de la symétrie orthogonale s par rapport à la droite d'équation $x = y$ est

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

dans la base canonique. Cette matrice est diagonalisable puisque s l'est (ex. 14.2).

2) Compte-tenu du corollaire 14.4, une matrice d'ordre n qui a n valeurs propres distinctes est diagonalisable.

Proposition 15.4.— Pour qu'un scalaire λ soit valeur propre de M , il faut et il suffit que $\det(\lambda I_n - M)$ soit nul.

Démonstration. Cela résulte de la proposition 14.3.(iii) puisque $\lambda I_n - M$ est la matrice dans la base canonique de l'endomorphisme $\lambda \text{Id}_{K^n} - u_M$. ■

Soit T une indéterminée; considérons la matrice

$$T I_n - M = \begin{pmatrix} T - a_{11} & -a_{12} & -a_{13} & \cdots & -a_{1n} \\ -a_{21} & T - a_{22} & -a_{23} & \cdots & -a_{2n} \\ \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot & \cdot \\ -a_{n1} & & & -a_{n,n-1} & T - a_{nn} \end{pmatrix} .$$

C'est une matrice dont les coefficients sont des polynômes en T , donc pas des réels ou des complexes comme d'habitude. La définition du déterminant que l'on a donnée se transpose sans aucun problème dans ce cadre un peu plus général, puisqu'il suffit de pouvoir multiplier et additionner les coefficients; mais comment être sûr que les propriétés du déterminant que l'on a démontrées sont encore valables? Pour s'en convaincre, il suffit de remarquer que nous avons développé la théorie du déterminant sur un corps quelconque, pourvu qu'il contienne \mathbf{Q} , et que la matrice ci-dessus est à coefficients dans le corps $K(T)$ des fractions rationnelles à coefficients dans K , qui contient K , donc \mathbf{Q} . On se rend compte sur cet exemple l'intérêt qu'il y a parfois à travailler dans un cadre général (surtout lorsque ça ne coûte pas plus cher...).

Le déterminant de cette matrice est une somme de produits de coefficients de cette matrice, donc un polynôme en T . Chacun de ces produits a n termes, qui sont des coefficients dans chacune des n colonnes; ce sont donc des polynômes de degré au plus n , et un seul de ces produits est de degré maximal n : c'est le produit des termes diagonaux. Il correspond à la permutation $\sigma = \text{Id}$, donc apparaît avec le signe $+1$. On a donc montré :

Théorème 15.5.— *Le déterminant de la matrice $TI_n - M$ est un polynôme en T unitaire de degré n , appelé polynôme caractéristique de M , et noté P_M .*

La proposition précédente dit que *les valeurs propres de M sont les racines de son polynôme caractéristique*. Il est utile de remarquer que son terme constant est $(-1)^n$ fois le déterminant de M .

Exemples 15.6. 1) Le polynôme caractéristique de la matrice $M = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ est

$$\begin{vmatrix} T & -1 \\ -1 & T \end{vmatrix} = T^2 - 1 .$$

Les valeurs propres de M sont donc ± 1 . Ce sont bien les valeurs propres de la symétrie plane orthogonale s par rapport à la droite d'équation $x = y$, dont M est la matrice dans la base canonique (*cf.* ex. 13.2 et 14.2). La matrice M est diagonalisable; on peut le voir de deux façons : soit parce que s l'est (ex. 14.2), soit parce que M a deux valeurs propres distinctes (cor. 14.4).

2) La rotation plane r_θ a pour matrice dans la base canonique (*cf.* § 46)

$$R_\theta = \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix} .$$

Son polynôme caractéristique est

$$\begin{vmatrix} T - \cos \theta & -\sin \theta \\ \sin \theta & T - \cos \theta \end{vmatrix} = (T - \cos \theta)^2 + \sin^2 \theta = T^2 - 2T \cos \theta + 1 .$$

Il n'a de racine réelle que si $\sin \theta$ est nul. On retrouve le fait que r_θ n'a de vecteur propre réel que si $\theta \equiv 0 \pmod{\pi}$ (*cf.* ex. 13.2). En revanche, R_θ a des valeurs propres complexes, qui sont $e^{\pm i\theta}$; on vérifie que des vecteurs propres (complexes) associés sont $\begin{pmatrix} 1 \\ \pm i \end{pmatrix}$.

3) Si $M = (a_{ij})$ est une matrice *triangulaire* (supérieure ou inférieure), il en est de même de la matrice $TI_n - M$; son déterminant est le produit de ses termes diagonaux (ex. 11.2.2) :

$$P_M(T) = (T - a_{11}) \cdots (T - a_{nn}) ,$$

et les valeurs propres de M sont ses coefficients diagonaux. En particulier si les coefficients diagonaux de M sont tous distincts, M est diagonalisable (cor. 14.4).

Corollaire 15.7.— *Toute matrice carrée à coefficients complexes a une valeur propre complexe.*

Démonstration. C'est le théorème de d'Alembert : tout polynôme de degré ≥ 1 à coefficients complexes a une racine complexe. ■

Tout endomorphisme d'un espace vectoriel *complexe* a donc aussi une valeur propre.

16. Polynôme caractéristique d'un endomorphisme

Soit u un endomorphisme d'un espace vectoriel E de dimension n . Soient \mathcal{B} une base de E et M la matrice de u dans la base \mathcal{B} . Pour qu'un scalaire λ soit valeur propre de u , il faut et il suffit que le déterminant de l'endomorphisme $\lambda \text{Id}_E - u$ soit nul (prop. 13.4). Ce dernier n'est autre que le déterminant de sa matrice $\lambda I_n - M$ dans la base \mathcal{B} . En d'autres termes, *les valeurs propres d'un endomorphisme sont les valeurs propres de sa matrice dans une base quelconque*. Comme des matrices sont semblables si et seulement si ce sont les matrices d'un même endomorphisme dans des bases peut-être différentes (10.9), on en déduit que *des matrices semblables ont mêmes valeurs propres*. On a en fait un résultat plus précis.

Théorème 16.1.— *Des matrices semblables ont même polynôme caractéristique.*

Démonstration. Nous allons montrer que si les matrices M et N sont semblables, il en est de même des matrices $TI_n - M$ et $TI_n - N$. Supposons donc qu'il existe une matrice carrée inversible P telle que $N = P^{-1}MP$; on a

$$P^{-1}(TI_n - M)P = P^{-1}(TI_n)P - P^{-1}MP = TI_n - N.$$

Ces deux matrices ont donc même déterminant (cor. 10.10), d'où le théorème. ■

(16.2) Cela permet de définir le polynôme caractéristique P_u d'un endomorphisme u : c'est le polynôme caractéristique de sa matrice dans une base quelconque. Ses racines sont les valeurs propres de u ; son terme constant est $(-1)^n$ fois le déterminant de u .

Exemple 16.3. Le polynôme caractéristique d'une symétrie par rapport à une droite dans le plan \mathbf{R}^2 est $T^2 - 1$ (cf. ex. 14.2); celui d'une rotation d'angle θ est $T^2 - 2T \cos \theta + 1$ (ex. 15.6.2).

17. Polynômes

On rappelle qu'un polynôme P à coefficients dans K est une suite (a_0, a_1, \dots) pour laquelle il existe un entier m tel que a_i soit nul pour $i > m$. On note un tel polynôme

$$P = \sum_{i \geq 0} a_i T^i,$$

où la somme est finie par hypothèse. Si P n'est pas nul, son degré est le plus grand entier i tel que a_i soit non nul; on pose souvent que le degré du polynôme nul est $-\infty$. On multiplie et on additionne les polynômes selon les règles que vous connaissez; le degré d'un produit est la somme des degrés.

On dit qu'un polynôme B *divise* un polynôme A , ou que A est un *multiple* de B , s'il existe un polynôme Q tel que $A = BQ$.

Théorème 17.1.— Soient A et B des polynômes. Si B n'est pas nul, il existe des polynômes Q et R , uniquement déterminés, tels que

$$A = BQ + R \quad \text{et} \quad \deg R < \deg B .$$

L'opération effectuée dans le théorème s'appelle la *division euclidienne* de A par B ; le polynôme Q est le *quotient*, R est le *reste*.

Démonstration. Commençons par l'unicité ; si $A = BQ + R = BQ' + R'$, on écrit $R - R' = B(Q' - Q)$, de sorte que B divise $R - R'$. Comme ce dernier est de degré $< \deg B$ par hypothèse, il est nul, et $R = R'$, d'où $Q = Q'$ puisque B n'est pas nul.

Montrons l'existence, en raisonnant par récurrence sur le degré n de A . Si $A = 0$, on prend $Q = R = 0$; on suppose donc A non nul. Si $n < \deg B$, on prend $Q = 0$ et $R = A$; sinon, on écrit

$$A = \sum_{i=0}^n a_i T^i \quad B = \sum_{i=0}^m b_i T^i ,$$

où a_n et b_m ne sont pas nuls. Le polynôme $A' = A - (a_n/b_m)T^{n-m}B$ est de degré $< n$; on lui applique l'hypothèse de récurrence : $A' = BQ' + R'$, avec $\deg R' < \deg B$. On en déduit

$$A = B((a_n/b_m)T^{n-m} + Q') + R' ;$$

il suffit de prendre $Q = (a_n/b_m)T^{n-m} + Q'$ et $R = R'$. ■

Exemple 17.2. La division euclidienne de $T^3 + T + 1$ par $T + 1$ est

$$T^3 + 2T + 1 = (T + 1)(T^2 - T + 3) - 2 .$$

Un élément α de K est *racine* d'un polynôme P si et seulement si $P(\alpha) = 0$.

Corollaire 17.3.— Pour qu'un élément α de K soit racine de P , il faut et il suffit que P soit divisible par $T - \alpha$.

Démonstration. La division euclidienne de P par $T - \alpha$ s'écrit $P = (T - \alpha)Q + R$, où le polynôme R est de degré < 1 , donc constant. On obtient sa valeur en substituant α à T : c'est $P(\alpha)$. Ceci démontre le corollaire. ■

Si α est racine de P , la *multiplicité* de cette racine est l'entier maximal m tel que $(T - \alpha)^m$ divise P .

(17.4) On appelle *idéal* tout sous-ensemble non vide I de $K[T]$ tel que, si A et B sont dans I , le polynôme $A - B$ soit dans I , ainsi que tout polynôme multiple de A .

Théorème 17.5.— Soit I un idéal; il existe un polynôme B tel que I soit l'ensemble des polynômes multiples de B .

On dit que B engendre I , ou que B est un générateur de I .

Démonstration du théorème. Si I ne contient que le polynôme nul, on prend $B = 0$. Sinon, soit B un polynôme dans I , non nul de degré minimal; comme I est un idéal, il contient tous les multiples de B . Inversement, si A est un élément quelconque de I , on fait la division euclidienne $A = BQ + R$. Comme I est un idéal, le polynôme $R = A - BQ$ est dans I , et son degré est strictement plus petit que celui de B ; il est donc nul, à cause du choix de B , et A est multiple de B . ■

Définition 17.6.— Soient A et B des polynômes. On dit qu'un polynôme D est un p.g.c.d. de A et B si D divise A et B , et si tout diviseur commun à A et B divise D . On dit que A et B sont premiers entre eux si 1 est un p.g.c.d. de A et B .

Le p.g.c.d. n'est pas unique : si D est un p.g.c.d., λD l'est aussi pour tout scalaire λ non nul. On vérifie facilement que l'on obtient ainsi tous les p.g.c.d. de A et B .

Proposition 17.7.— Deux polynômes A et B admettent toujours un p.g.c.d.; il peut s'écrire $AU + BV$, où U et V sont des polynômes.

Démonstration. On vérifie sans difficulté que l'ensemble

$$I = \{ AP + BQ \mid P, Q \in K[T] \}$$

est un idéal; si D engendre I (th. 17.5), I est l'ensemble des multiples de D , et on peut écrire $D = AU + BV$. Comme A et B sont dans I , ce sont des multiples de D ; inversement, tout polynôme qui divise A et B divise $AU + BV$, donc D . Ceci montre que D est un p.g.c.d. de A et B . ■

Théorème de Bézout 17.8.— Pour que des polynômes A et B soient premiers entre eux, il faut et il suffit qu'il existe des polynômes U et V tels que $AU + BV = 1$.

Démonstration. Si A et B sont premiers entre eux, 1 est un p.g.c.d. et la proposition 17.7 entraîne qu'il existe des polynômes U et V tels que $AU + BV = 1$. Inversement, si $AU + BV = 1$, tout p.g.c.d. de A et B divise A et B , donc 1 ; il est donc constant, ce qui signifie que A et B sont premiers entre eux. ■

Lemme de Gauss 17.9.— Soient A , B et C des polynômes.

a) Si A et B sont premiers entre eux et que A divise BC , alors A divise C .

b) Si A est premier avec B et avec C , il est premier avec BC .

Démonstration. Sous les hypothèses de a), on peut, grâce au théorème de Bézout, écrire $1 = AU + BV$. On peut aussi écrire $BC = AP$, d'où

$$C = C(AU + BV) = ACU + BCV = ACU + APV = A(CU + PV),$$

ce qui montre que A divise C et prouve a).

Plaçons-nous sous les hypothèses de b) et notons D le p.g.c.d. de A et BC . Tout diviseur commun à D et B divise A et B donc D est premier avec B . Comme D divise BC , le a) entraîne que D divise C . Comme il divise aussi A , il vaut 1 . ■

Corollaire 17.10.— Soient P, P_1, \dots, P_m des polynômes.

a) Si les polynômes P_1, \dots, P_m sont premiers entre eux deux à deux et divisent P , le produit $P_1 \cdots P_m$ divise P .

b) Si P est premier avec chacun des polynômes P_1, \dots, P_m , il est premier avec $P_1 \cdots P_m$.

Démonstration. Le a) se démontre par récurrence sur m : on écrit $P = P_1 Q$; pour tout $j \geq 2$, le polynôme P_j divise P et est premier avec P_1 donc divise Q par le point a) du lemme de Gauss. Par hypothèse de récurrence, Q est divisible par $P_2 \cdots P_m$, d'où a).

Le b) se démontre aussi par récurrence sur m , le cas $m = 2$ étant le b) du lemme de Gauss. ■

Exemple 17.11. Si Q est un polynôme dont α n'est pas racine, il est clair que Q et $T - \alpha$ sont premiers entre eux (leur p.g.c.d. est de degré au plus 1). Le corollaire 17.10.b) entraîne que Q est premier avec $(T - \alpha)^n$ pour tout entier n . En particulier, si $\alpha_1, \dots, \alpha_p$ sont des racines distinctes, de multiplicité n_1, \dots, n_p , d'un polynôme P , celui-ci est divisible par $\prod_{j=1}^p (T - \alpha_j)^{n_j}$ (cor. 17.10.a)).

(17.12) On dit qu'un polynôme P à coefficients dans K est *scindé* dans K , s'il se décompose en produit de facteurs du premier degré

$$P(T) = a \prod_{j=1}^p (T - \alpha_j)^{n_j} ,$$

avec $\alpha_j \in K$ et $a \in K$. A cause de l'exemple ci-dessus, cela signifie exactement que la somme $\sigma(P)$ des multiplicités de ses racines dans K est égale à son degré. Si P est scindé, tout polynôme qui divise P est aussi scindé : si on écrit $P = AB$, on $\deg P = \sigma(P) = \sigma(A) + \sigma(B)$ et $\sigma(A) \leq \deg A$, $\sigma(B) \leq \deg B$; cela force l'égalité $\sigma(A) = \deg A$, et A est scindé.

Exemple 17.13. Tout polynôme de $\mathbf{C}[T]$ est scindé : c'est le théorème de d'Alembert. Il existe dans $\mathbf{R}[T]$ des polynômes qui ne sont pas scindés, comme par exemple

$$T^3 - 1 = (T - 1)(T^2 + T + 1) ,$$

puisque $T^2 + T + 1$ n'a pas de racine réelle.

18. Trigonalisation, théorème de Hamilton-Cayley

Nous allons montrer notre premier résultat de réduction.

Définition 18.1.— On dit qu'une matrice est trigonalisable si elle est semblable à une matrice triangulaire supérieure. On dit qu'un endomorphisme est trigonalisable s'il existe une base dans laquelle sa matrice est triangulaire supérieure.

Un endomorphisme est donc trigonalisable si sa matrice dans une base quelconque est trigonalisable. Dans une base dans laquelle sa matrice est triangulaire supérieure, les coefficients diagonaux a_{ii} sont ses valeurs propres (ex. 15.6.3). Plus précisément, son polynôme caractéristique s'écrit $\prod_{i=1}^n (T - a_{ii})$; il est donc scindé dans K (cf. 17.12). Le théorème suivant dit que la réciproque est vraie.

Théorème 18.2.— Pour qu'un endomorphisme d'un K -espace vectoriel soit trigonalisable, il faut et il suffit que son polynôme caractéristique soit scindé dans K .

Démonstration. On procède par récurrence sur la dimension n de l'espace vectoriel E . Si $n = 1$, il n'y a rien à montrer : toute matrice carrée d'ordre 1 est triangulaire. Supposons $n > 1$; l'hypothèse nous dit que le polynôme caractéristique de l'endomorphisme u s'écrit $P_u(T) = \prod_{j=1}^n (T - \lambda_j)$, où $\lambda_1, \dots, \lambda_n$ sont des éléments de K , pas obligatoirement tous distincts. Le scalaire λ_n est alors valeur propre de u , de sorte que $u - \lambda_n \text{Id}_E$ n'est pas injectif donc pas surjectif, puisque c'est un endomorphisme d'un espace vectoriel de dimension finie. Son image est donc un sous-espace vectoriel de E de dimension $r < n$; on en prend une base (e'_1, \dots, e'_r) que l'on complète en une base $\mathcal{B}' = (e'_1, \dots, e'_n)$ de E . Soit F le sous-espace vectoriel de E engendré par e'_1, \dots, e'_{n-1} . Pour tout $j = 1, \dots, n-1$, le vecteur

$$u(e'_j) = (u - \lambda_n \text{Id}_E)(e'_j) + \lambda_n e'_j,$$

est dans F , donc se décompose en combinaison linéaire des vecteurs e'_1, \dots, e'_{n-1} . De plus, $u(e'_n) = (u - \lambda_n \text{Id}_E)(e'_n) + \lambda_n e'_n$, de sorte que la matrice de u dans la base \mathcal{B}' s'écrit par blocs

$$\begin{pmatrix} A & C \\ 0 & \lambda_n \end{pmatrix},$$

où A est une matrice carrée d'ordre $n-1$ et C une colonne. On a alors $P_u(T) = P_A(T)(T - \lambda_n)$, de sorte que le polynôme caractéristique de A est $\prod_{j=1}^{n-1} (T - \lambda_j)$; il est donc scindé. L'hypothèse de récurrence entraîne que l'endomorphisme de F induit par u (qui est de matrice A dans la base (e'_1, \dots, e'_{n-1}) de F) est trigonalisable : il existe une base (e_1, \dots, e_{n-1}) de F dans laquelle sa matrice est triangulaire supérieure. La matrice de u dans la base $(e_1, \dots, e_{n-1}, e'_n)$ de E est triangulaire supérieure, ce qui termine la démonstration. ■

Le théorème s'applique à tous les endomorphismes (ou matrices) complexes, et aux endomorphismes (ou matrices) réels dont le polynôme caractéristique a toutes ses racines réelles.

Soit u un endomorphisme d'un espace vectoriel E de dimension finie. On a vu en 5.1 qu'il existe toujours un polynôme non nul P tel que $P(u) = 0$. On peut utiliser le théorème pour préciser cet énoncé.

Théorème de Hamilton-Cayley 18.3. – Soit u un endomorphisme d'un espace vectoriel E ; on a $P_u(u) = 0$.

Il y a bien sûr un énoncé analogue pour les matrices.

Démonstration du théorème. Nous donnerons deux démonstrations différentes de ce théorème ; la première n'est valable que dans le cas où P_u est scindé (ce qui suffit dans les cas $K = \mathbf{R}$ ou \mathbf{C}) et utilise le théorème précédent. La seconde utilise les propriétés des comatrices.

Première démonstration. On suppose ici que le polynôme caractéristique de u s'écrit $P_u(T) = \prod_{i=1}^n (T - \lambda_i)$, où $\lambda_1, \dots, \lambda_n$ sont des éléments de K . D'après le théorème, il existe une base $\mathcal{B} = (e_1, \dots, e_n)$ de E dans laquelle la matrice de u est triangulaire supérieure, de coefficients diagonaux $\lambda_1, \dots, \lambda_n$. Montrons par récurrence sur $j = 1, \dots, n$ que

$$(u - \lambda_j \text{Id}_E) \circ \dots \circ (u - \lambda_1 \text{Id}_E)(e_i) = 0$$

pour $1 \leq i \leq j$. Pour $i = 1$, cela dit simplement que $(u - \lambda_1 \text{Id}_E)(e_1)$ est nul, ce qui se lit sur la matrice de u . Supposons la propriété vraie pour j . Pour montrer la propriété pour $j + 1$, il suffit grâce à l'hypothèse de récurrence de démontrer que le vecteur

$$x = (u - \lambda_{j+1} \text{Id}_E)(u - \lambda_j \text{Id}_E) \circ \dots \circ (u - \lambda_1 \text{Id}_E)(e_{j+1})$$

est nul. On peut aussi l'écrire

$$x = (u - \lambda_j \text{Id}_E) \circ \dots \circ (u - \lambda_1 \text{Id}_E)(u - \lambda_{j+1} \text{Id}_E)(e_{j+1}) .$$

Le fait que la matrice (a_{ij}) de u soit triangulaire supérieure dans la base \mathcal{B} entraîne que $(u - \lambda_{j+1} \text{Id}_E)(e_{j+1}) = \sum_{i=1}^j a_{i,j+1} e_i$ est combinaison linéaire de e_1, \dots, e_j . Il est donc dans le noyau de $(u - \lambda_j \text{Id}_E) \circ \dots \circ (u - \lambda_1 \text{Id}_E)$ par l'hypothèse de récurrence. La propriété est donc démontrée par récurrence ; pour $j = n$, elle dit que l'endomorphisme $(u - \lambda_n \text{Id}_E) \circ \dots \circ (u - \lambda_1 \text{Id}_E)$ s'annule sur tous les vecteurs de la base \mathcal{B} ; il est donc nul. ■

Seconde démonstration. Soit M une matrice carrée d'ordre n ; posons $N = T I_n - M$, et notons comme d'habitude \tilde{N} sa comatrice (cf. (11.5)). Il est clair que les coefficients de \tilde{N} sont des polynômes de degré $\leq n - 1$; on peut donc l'écrire

$$\tilde{N} = \sum_{j=0}^{n-1} T^j \tilde{N}_j ,$$

où les \tilde{N}_j sont des matrices carrées à coefficients dans K . La proposition 11.6 donne l'identité

$$N {}^t \tilde{N} = (\det N) I_n = P_M(T) I_n ,$$

qui s'écrit aussi

$$(\mathbb{T}I_n - M) \left(\sum_{j=0}^{n-1} \mathbb{T}^j \tilde{N}_j \right) = P_M(\mathbb{T}) I_n .$$

Si on note $P_M(\mathbb{T}) = \sum_{j=0}^n a_j \mathbb{T}^j$ (avec $a_n = 1$), on obtient en identifiant les coefficients

$$M\tilde{N}_0 = -a_0 I_n , \quad \tilde{N}_0 - M\tilde{N}_1 = a_1 I_n , \quad \dots , \quad \tilde{N}_{n-2} - M\tilde{N}_{n-1} = a_{n-1} I_n , \quad \tilde{N}_{n-1} = a_n I_n = I_n .$$

On en déduit

$$\begin{aligned} P_M(M) &= \sum_{j=0}^n a_j M^j = \sum_{j=0}^n M^j (a_j I_n) \\ &= -M\tilde{N}_0 + M(\tilde{N}_0 - M\tilde{N}_1) + \dots + M^{n-1}(\tilde{N}_{n-2} - M\tilde{N}_{n-1}) + M^n \tilde{N}_{n-1} = 0 \end{aligned}$$

(les termes se regroupent deux à deux et s'annulent). ■

Exemple 18.4. Reprenons l'exercice 5.2; on demande de trouver un polynôme qui annule la matrice $M = \begin{pmatrix} 0 & 1 \\ 2 & 3 \end{pmatrix}$. On a

$$P_M(\mathbb{T}) = \begin{vmatrix} \mathbb{T} & -1 \\ -2 & \mathbb{T} - 3 \end{vmatrix} = \mathbb{T}^2 - 3\mathbb{T} - 2 .$$

Le théorème de Hamilton-Cayley donne donc $M^2 - 3M - 2I_2 = 0$; on peut le vérifier :

$$M^2 - 3M - 2I_2 = \begin{pmatrix} 2 & 3 \\ 6 & 11 \end{pmatrix} - 3 \begin{pmatrix} 0 & 1 \\ 2 & 3 \end{pmatrix} - 2 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} .$$

19. Polynôme minimal

On introduit ici un second polynôme attaché à un endomorphisme d'un espace vectoriel.

Proposition 19.1.— *Soit u un endomorphisme d'un K -espace vectoriel E de dimension finie. Il existe un unique polynôme unitaire p_u à coefficients dans K tel qu'un polynôme P à coefficients dans K vérifie $P(u) = 0$ si et seulement si p_u divise P . On l'appelle le polynôme minimal de u .*

Démonstration. On vérifie facilement que l'ensemble des polynômes P tels que $P(u) = 0$ est un idéal (cf. 17.4), qui contient des polynômes non nuls par la proposition 5.1. Il est donc engendré par un polynôme p_u (th. 17.5) *non nul*, que l'on peut donc choisir unitaire en le multipliant par l'inverse de son coefficient dominant. Le lecteur vérifiera sans peine la propriété d'unicité. ■

Le polynôme minimal est toujours de degré ≥ 1 . Pour que le polynôme minimal de u soit T , il faut et il suffit que u soit nul. Pour que le polynôme minimal de u soit de degré 1 (égal à $T - \lambda$), il faut et il suffit que u soit une homothétie (égale à λId_E).

On définit le polynôme minimal p_M d'une matrice carrée M de façon analogue; c'est aussi le polynôme minimal de l'endomorphisme u_M de K^n défini en (15.1).

Le théorème de Hamilton-Cayley 18.3 s'énonce simplement sous la forme suivante :

le polynôme minimal divise le polynôme caractéristique.

Exemple 19.2. Reprenons la matrice $M = \begin{pmatrix} 0 & 1 \\ 2 & 3 \end{pmatrix}$ de l'exercice 18.4, de polynôme caractéristique $P_M(T) = T^2 - 3T - 2$; comme M n'est pas multiple de l'identité, p_M est de degré > 1 . Comme p_M est unitaire et divise P_M , on en déduit $p_M = P_M$.

Les racines de p_u sont donc des valeurs propres de u . Montrons la réciproque.

Théorème 19.3.— *Les valeurs propres d'un endomorphisme sont les racines de son polynôme minimal.*

Démonstration. Soient λ une valeur propre d'un endomorphisme u , et x un vecteur propre associé. On fait la division euclidienne de p_u par $T - \lambda$; elle s'écrit

$$p_u(T) = Q(T)(T - \lambda) + c,$$

où c est un polynôme de degré < 1 , donc une constante. On en déduit

$$0 = p_u(u) = Q(u)(u - \lambda \text{Id}_E) + c \text{Id}_E.$$

Si l'on applique cela au vecteur x , on obtient

$$0 = Q(u)(u - \lambda \text{Id}_E)(x) + c \text{Id}_E(x) = Q(u)(u(x) - \lambda x) + cx = Q(u)(0) + cx = cx;$$

comme x n'est pas nul, on obtient $c = 0$, de sorte que $p_u(T) = Q(T)(T - \lambda)$. Cela signifie que λ est racine de p_u . ■

Exemple 19.4. Supposons qu'un endomorphisme u d'un espace vectoriel de dimension n ait n valeurs propres distinctes. Celles-ci sont alors racines du polynôme minimal; comme il divise le polynôme caractéristique, il est de degré au plus n ; on en déduit $p_u = P_u$ (cela généralise l'exemple précédent).

20. Réduction des endomorphismes

Les résultats que l'on démontre ici reposent sur l'énoncé suivant.

Lemme des noyaux 20.1.— Soient u un endomorphisme d'un espace vectoriel E et P un polynôme.

a) Le sous-espace vectoriel $\text{Ker } P(u)$ de E est stable par u , c'est-à-dire

$$u(\text{Ker } P(u)) \subset \text{Ker } P(u) ;$$

b) si $P = AB$, avec A et B premiers entre eux, on a

$$\text{Ker } P(u) = \text{Ker } A(u) \oplus \text{Ker } B(u) .$$

Démonstration. Remarquons que u et $P(u)$ commutent. Si x est un élément du noyau de $P(u)$, on a donc $P(u)(u(x)) = u(P(u))(x) = 0$, de sorte que $u(x)$ est encore dans le noyau de $P(u)$; cela montre a). D'après le théorème de Bézout, il existe des polynômes U et V tels que $AU + BV = 1$, d'où $A(u)U(u) + B(u)V(u) = \text{Id}_E$. Si x est dans le noyau de $A(u)$ et dans le noyau de $B(u)$, on a

$$x = \text{Id}_E(x) = A(u)U(u)(x) + B(u)V(u)(x) = U(u)A(u)(x) + V(u)B(u)(x) = 0 ,$$

de sorte que les noyaux de $A(u)$ et de $B(u)$ sont en somme directe (déf. 13.5). Remarquons ensuite que le noyau de $A(u)$ est contenu dans celui de $P(u)$: si $A(u)(x) = 0$, on a $P(u)(x) = B(u)(A(u)(x)) = 0$. De même, le noyau de $B(u)$ est contenu dans celui de $P(u)$. Enfin, pour tout x dans le noyau de $P(u)$, on a d'autre part

$$x = \text{Id}_E(x) = A(u)U(u)(x) + B(u)V(u)(x) .$$

Comme $B(u)(A(u)U(u)(x)) = U(u)(P(u)(x)) = 0$, le vecteur $A(u)U(u)(x)$ est dans le noyau de $B(u)$; de même, $B(u)V(u)(x)$ est dans le noyau de $A(u)$. Cela montre b). ■

Corollaire 20.2.— Soient u un endomorphisme d'un espace vectoriel E , et P_1, \dots, P_m des polynômes premiers entre eux deux à deux; on pose $P = P_1 \cdots P_m$. On a

$$\text{Ker } P(u) = \text{Ker } P_1(u) \oplus \cdots \oplus \text{Ker } P_m(u) .$$

Démonstration. On procède par récurrence sur m , le cas $m = 2$ étant le lemme ci-dessus. Si P_1, \dots, P_m sont premiers entre eux deux à deux, le corollaire 17.10.a) entraîne que P_1 et $P_2 \cdots P_m$ sont premiers entre eux. Le lemme des noyaux entraîne

$$\text{Ker } P(u) = \text{Ker } P_1(u) \oplus \text{Ker}(P_2 \cdots P_m)(u) ;$$

on conclut par l'hypothèse de récurrence. ■

Soient u un endomorphisme d'un espace vectoriel E de dimension n , et $\lambda_1, \dots, \lambda_p$ ses valeurs propres distinctes ; on suppose que son polynôme minimal est scindé, c'est-à-dire qu'il se décompose en

$$p_u(T) = \prod_{j=1}^p (T - \lambda_j)^{m_j} ,$$

avec $m_j > 0$ pour tout j (th. 19.3).

Définition 20.3.— *Sous les hypothèses et avec les notations précédentes, on appelle sous-espace caractéristique associé à la valeur propre λ_j le noyau de l'endomorphisme $(u - \lambda_j \text{Id}_E)^{m_j}$; on le note E'_{λ_j} .*

Le lemme 20.1.a) entraîne que l'espace caractéristique E'_λ est stable par u . Il contient l'espace propre E_λ :

$$\boxed{E_\lambda \subset E'_\lambda}$$

Si le polynôme caractéristique de u est scindé, il se décompose en

$$P_u(T) = \prod_{j=1}^p (T - \lambda_j)^{n_j} .$$

(20.4) On appelle n_j la *multiplicité* de la valeur propre λ_j . Par (17.12), le polynôme minimal est aussi scindé, et l'on a $m_j \leq n_j$ pour tout j , par le théorème de Hamilton-Cayley.

Exemple 20.5. Soient E un espace vectoriel de dimension 3 et u l'endomorphisme de E de matrice $M = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}$ dans une base (e_1, e_2, e_3) de E . Son polynôme caractéristique est $P_u(T) = (T - 2)^2(T - 3)$, de sorte que son polynôme minimal est $p_u(T) = (T - 2)^m(T - 3)$ avec $m = 1$ ou 2 . On vérifie par le calcul que la matrice $(M - 2I_3)(M - 3I_3)$ n'est pas nulle, d'où $m = 2$. L'espace propre E_2 est la droite engendrée par le vecteur e_1 . Pour déterminer l'espace caractéristique E'_2 , il faut résoudre le système $(M - 2I_3)^2 X = 0$; on trouve que c'est le plan engendré par e_1 et e_2 . Par définition, les espaces E_3 et E'_3 sont égaux au noyau de $M - 3I_3$, c'est-à-dire à la droite engendrée par e_3 .

Théorème 20.6.— *Soit u un endomorphisme d'un espace vectoriel E dont le polynôme caractéristique est scindé.*

a) *L'espace vectoriel E est somme directe des espaces caractéristiques de u .*

b) *La dimension de E'_λ est la multiplicité n_λ de la valeur propre λ . En particulier, la dimension de l'espace propre E_λ est $\leq n_\lambda$.*

Démonstration. Il résulte de l'exemple 17.11 que les polynômes $(T - \lambda_j)^{m_j}$ sont premiers entre eux deux à deux, pour $j = 1, \dots, r$. Le corollaire 20.2 entraîne donc a), puisque E est

le noyau de leur produit p_u . Les espaces caractéristiques sont stables par u (lemme 20.1.a); cela signifie que si l'on juxtapose des bases de chaque E'_{λ_j} , on obtient une base de E dans laquelle la matrice de u est « diagonale par blocs » :

$$\begin{pmatrix} M_1 & 0 & 0 & \cdots & 0 \\ 0 & M_2 & 0 & \cdots & 0 \\ 0 & 0 & \cdot & & \cdot \\ \cdot & \cdot & & \cdot & \cdot \\ 0 & 0 & & \cdot & M_p \end{pmatrix},$$

où chaque M_j est carrée d'ordre la dimension d_j de E'_{λ_j} . La proposition 11.1 entraîne que le polynôme caractéristique de u est le produit des polynômes caractéristiques des M_j ; en particulier, ceux-ci sont scindés.

Comme $(M_j - \lambda_j I_{d_j})^{m_j} = 0$ par définition de E'_{λ_j} , la seule valeur propre possible de M_j est λ_j ; comme P_{M_j} est scindé, c'est $(T - \lambda_j)^{d_j}$. On en déduit

$$P_u(T) = P_{M_1}(T) \cdots P_{M_p}(T) = \prod_{j=1}^p (T - \lambda_j)^{d_j},$$

ce qui prouve $d_j = n_j$ pour tout j . ■

(20.7) La démonstration précédente montre que toute matrice dont le polynôme caractéristique est scindé est semblable à une matrice diagonale par bloc où chaque bloc a une seule valeur propre.

On dit qu'un endomorphisme w est *nilpotent* s'il existe un entier positif m tel que $w^m = 0$.

Théorème (Décomposition de Dunford) 20.8.— *Soit u un endomorphisme dont le polynôme caractéristique est scindé. On peut écrire $u = v + w$ où v est diagonalisable, w nilpotent et $vw = wv$.*

Démonstration. Soit u un endomorphisme d'un espace vectoriel E dont le polynôme caractéristique est scindé. La démonstration du théorème précédent montre qu'il existe une base \mathcal{B} de E dans laquelle la matrice de u est diagonale par blocs, chaque bloc n'ayant qu'une seule valeur propre. Soit v l'endomorphisme de E de matrice

$$\begin{pmatrix} \lambda_1 I_{n_1} & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2 I_{n_2} & 0 & \cdots & 0 \\ 0 & 0 & \cdot & & \cdot \\ \cdot & \cdot & & \cdot & \cdot \\ 0 & 0 & & \cdot & \lambda_p I_{n_p} \end{pmatrix}$$

dans la base \mathcal{B} . Il est clair que v est diagonalisable et commute avec $w = u - v$. On a d'autre part $w^{\max\{m_1, \dots, m_p\}} = 0$, ce qui montre le théorème. ■

Théorème 20.9.— Soit u un endomorphisme d'un espace vectoriel E . Les conditions suivantes sont équivalentes :

(i) u est diagonalisable ;

(ii) le polynôme caractéristique de u est scindé et la dimension de chaque espace propre E_λ est égale à la multiplicité de la valeur propre λ ;

(iii) le polynôme minimal de u est scindé et n'a que des racines simples.

Lorsque le polynôme caractéristique de u est scindé, une condition nécessaire et suffisante pour que u soit diagonalisable est donc $E_\lambda = E'_\lambda$ pour toute valeur propre λ .

Démonstration du théorème. Supposons u diagonalisable, de valeurs propres distinctes $\lambda_1, \dots, \lambda_p$. La propriété (ii) est évidente si l'on calcule le polynôme caractéristique de u dans une base où sa matrice est diagonale ; il est tout aussi clair que $p_u(T) = \prod_{j=1}^p (T - \lambda_j)$, d'où (iii).

Supposons (ii) vérifiée ; comme E_λ est inclus dans E'_λ et a même dimension (th. 20.6.b), ils sont égaux. Le théorème 20.6.a) entraîne que E est somme directe des espaces propres, d'où (i) (prop. 14.3).

Enfin, supposons (iii) vérifiée ; on a alors $E_\lambda = E'_\lambda$ par définition des espaces caractéristiques. Le théorème 20.6.a) entraîne que E est somme directe des espaces propres, d'où (i) (prop. 14.3). ■

Terminons par quelques remarques importantes sur la matrices réelles. Soit M une matrice carrée d'ordre n à coefficients réels ; les racines complexes de son polynôme caractéristique peuvent être classées d'une part en valeurs propres réelles, d'autre part en paires de valeurs propres complexes non réelles conjuguées $\mu_1, \bar{\mu}_1, \dots, \mu_s, \bar{\mu}_s$, les multiplicités de μ_j et de $\bar{\mu}_j$ étant les mêmes. Les espaces propres E_{μ_j} et $E_{\bar{\mu}_j}$ sont *conjugués*, ainsi que les espaces caractéristiques E'_{μ_j} et $E'_{\bar{\mu}_j}$; cela permet parfois de diminuer le volume des calculs (*cf.* exemple 2) ci-dessous).

(20.10) Soit M une matrice carrée réelle ; comparons son polynôme minimal *réel* p_M , c'est-à-dire le polynôme unitaire réel de plus petit degré qui annule M , avec son polynôme minimal *complexe* p_M , c'est-à-dire le polynôme unitaire complexe de plus petit degré qui annule M . On sait que q_M divise p_M . Ecrivons $q_M(T) = a_0 + a_1 T + \dots + a_{m-1} T^{m-1} + T^m$; comme $q_M(M)$ est nulle, il en est de même de sa partie réelle

$$(\operatorname{Re}(q_M))(M) = \operatorname{Re}(a_0)I_n + \operatorname{Re}(a_1)M + \dots + \operatorname{Re}(a_{m-1})M^{m-1} + M^m$$

Par définition de p_M , le degré m du polynôme réel $\operatorname{Re}(q_M)$ est supérieur à celui de p_M , donc $p_M = q_M$.

Exemples 20.11. 1) Le polynôme caractéristique de la matrice $M = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{pmatrix}$ est

$P_M(T) = T^4 - 1$; ses racines sont $1, -1, i$ et $-i$. Elle a 4 valeurs propres distinctes, donc est diagonalisable, et $p_M = P_M$. L'espace propre pour 1 est engendré par $\begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$, l'espace propre pour -1 est engendré par $\begin{pmatrix} 1 \\ -1 \\ 1 \\ -1 \end{pmatrix}$, l'espace propre pour i est engendré par $\begin{pmatrix} 1 \\ i \\ -1 \\ -i \end{pmatrix}$, donc l'espace propre pour $-i$ est engendré par le conjugué $\begin{pmatrix} 1 \\ -i \\ -1 \\ i \end{pmatrix}$.

2) On considère la matrice

$$M = \begin{pmatrix} 0 & & \cdots & 0 & a_1 \\ 1 & 0 & & \cdots & 0 & a_2 \\ 0 & 1 & 0 & \cdots & 0 & a_3 \\ \vdots & & \ddots & \ddots & \vdots & \vdots \\ 0 & & \ddots & 1 & 0 & a_{n-1} \\ 0 & & \cdots & 0 & 1 & a_n \end{pmatrix}.$$

Contrairement à l'habitude, non allons calculer d'abord son polynôme minimal et en déduire son polynôme caractéristique. Soit u l'endomorphisme de matrice M dans la base canonique (e_1, \dots, e_n) de K^n . On a

$$u(e_1) = e_2 \quad u(e_2) = e_3 \quad \cdots \quad u(e_{n-1}) = e_n \quad u(e_n) = a_1 e_1 + \cdots + a_n e_n.$$

Notons $p_M(T) = b_0 + b_1 T + \cdots + b_{m-1} T^{m-1} + T^m$. Si $m < n$, on a

$$(p_M(u))(e_1) = b_0 e_1 + b_1 e_2 + \cdots + b_{m-1} e_m + e_{m+1},$$

qui n'est pas nul car son coefficient sur e_{m+1} ne l'est pas, ce qui est absurde. On a donc $m = n$, et on en déduit déjà $P_M = p_M$. On a

$$(p_M(u))(e_1) = b_0 e_1 + b_1 e_2 + \cdots + b_{n-1} e_n + a_1 e_1 + \cdots + a_n e_n,$$

d'où on déduit $b_0 = -a_1, b_1 = -a_2, \dots, b_{n-1} = -a_n$, c'est-à-dire

$$p_M(T) = P_M(T) = -a_1 - a_2 T - \cdots - a_n T^{n-1} + T^n.$$

On retrouve (sans calcul!) le polynôme caractéristique de la matrice de l'exemple précédent. Pour que la matrice M soit diagonalisable, il faut et il suffit que les racines de ce polynôme soient toutes simples (th. 20.9), c'est-à-dire que M ait n valeurs propres distinctes.

3) On vérifie par le calcul que le polynôme caractéristique de la matrice $M = \begin{pmatrix} 2 & -1 & 2 \\ 10 & -5 & 7 \\ 4 & -2 & 2 \end{pmatrix}$ est $T^2(T + 1)$. Ses valeurs propres sont donc -1 (de multiplicité 1 ; on dit que c'est une valeur propre *simple*), et 0 (de multiplicité 2 ; on dit que c'est une valeur propre *double*). On cherche l'espace propre de la valeur propre -1 en cherchant le noyau de la matrice

$$M + I_3 = \begin{pmatrix} 3 & -1 & 2 \\ 10 & -4 & 7 \\ 4 & -2 & 3 \end{pmatrix} .$$

On sait d'avance que ce noyau n'est pas nul, et qu'il est de dimension au plus la multiplicité de la valeur propre (th. 20.6.b), c'est-à-dire ici 1. Il est donc de dimension exactement 1. Il faut résoudre le système

$$\begin{cases} 3x_1 - x_2 + 2x_3 = 0 \\ 10x_1 - 4x_2 + 7x_3 = 0 \\ 4x_1 - 2x_2 + 3x_3 = 0 \end{cases} .$$

On trouve que toute solution est proportionnelle au vecteur $e_1 = (1, -1, -2)$, qui engendre donc E_{-1} .

On cherche l'espace propre de la valeur propre 0 en cherchant le noyau de la matrice M . On sait d'avance que ce noyau n'est pas nul, et qu'il est de dimension au plus la multiplicité de la valeur propre, c'est-à-dire ici 2. Il est donc de dimension 1 ou 2. Il faut résoudre le système

$$\begin{cases} 2x_1 - x_2 + 2x_3 = 0 \\ 10x_1 - 5x_2 + 7x_3 = 0 \\ 4x_1 - 2x_2 + 2x_3 = 0 \end{cases} .$$

On trouve que toute solution est proportionnelle au vecteur $e_2 = (1, 2, 0)$, qui engendre donc E_0 . La somme des dimensions des espaces propres est 2, strictement moins que la dimension de l'espace \mathbf{R}^3 : la matrice M n'est donc pas diagonalisable (prop. 14.3). Le théorème 20.6 nous dit que l'espace caractéristique $E'_0 = \text{Ker } M^2$ est de dimension 2 ; il contient déjà le vecteur propre e_2 ; comme

$$M^2 = \begin{pmatrix} 2 & -1 & 1 \\ -2 & 1 & -1 \\ -4 & 2 & -2 \end{pmatrix} ,$$

on voit en résolvant le système

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

que E'_0 contient aussi le vecteur $e_3 = (0, 1, 1)$. La matrice de passage de la base canonique de \mathbf{R}^3 à la base (e_1, e_2, e_3) est la matrice

$$P = \begin{pmatrix} 1 & 1 & 0 \\ -1 & 2 & 1 \\ -2 & 0 & 1 \end{pmatrix},$$

dont le déterminant est 1, et son inverse est

$$P^{-1} = {}^t\tilde{P} = \begin{pmatrix} 2 & -1 & 1 \\ -1 & 1 & -1 \\ 4 & -2 & 3 \end{pmatrix}.$$

On a par construction $Me_1 = -e_1$ et $Me_2 = 0$, et on calcule $Me_3 = e_2$. On en déduit (formule (4.7))

$$M = \begin{pmatrix} 1 & 1 & 0 \\ -1 & 2 & 1 \\ -2 & 0 & 1 \end{pmatrix} \begin{pmatrix} -1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 2 & -1 & 1 \\ -1 & 1 & -1 \\ 4 & -2 & 3 \end{pmatrix}.$$

Pour trouver la décomposition de Dunford de M en somme d'une matrice diagonalisable D et d'une matrice N nilpotente qui commutent, on calcule

$$N = \begin{pmatrix} 1 & 1 & 0 \\ -1 & 2 & 1 \\ -2 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 2 & -1 & 1 \\ -1 & 1 & -1 \\ 4 & -2 & 3 \end{pmatrix} = \begin{pmatrix} 4 & -2 & 3 \\ 8 & -4 & 6 \\ 0 & 0 & 0 \end{pmatrix}$$

et

$$D = M - N = \begin{pmatrix} -2 & 1 & -1 \\ 2 & -1 & 1 \\ 4 & -2 & 2 \end{pmatrix}.$$

21.

EXERCICES

(21.1) Soient u et v des endomorphismes d'un espace vectoriel E . Montrer que $u \circ v$ et $v \circ u$ ont les mêmes valeurs propres.

(21.2) Soit u un endomorphisme inversible d'un espace vectoriel E . Déterminer les valeurs propres de u^{-1} en fonction de celles de u .

(21.3) Soit E l'espace vectoriel réel engendré par les fonctions \sin et \cos .

a) Quelle est la dimension de E ?

b) Montrer que la dérivation ∂ est un endomorphisme de E . Quel est son rang ?

c) Montrer que ∂ n'a pas de vecteur propre dans E .

d) On considère à présent l'espace vectoriel complexe F engendré par les fonctions \cos et \sin . Déterminer une base de F formée de vecteurs propres de ∂ .

(21.4) On considère les suites $(u_n)_{n \in \mathbf{N}}$, $(v_n)_{n \in \mathbf{N}}$ et $(w_n)_{n \in \mathbf{N}}$ définies par leurs premiers termes u_0, v_0 et w_0 et les relations de récurrence

$$u_{n+1} = u_n + 2v_n - 2w_n, \quad v_{n+1} = 2u_n + v_n - 2w_n \quad \text{et} \quad w_{n+1} = 2u_n + 2v_n - 3w_n.$$

a) Ecrire ces relations sous la forme $U_{n+1} = MU_n$ où M est une matrice que l'on déterminera, et (U_n) une suite de vecteurs de \mathbf{R}^3 .

b) Déterminer une base dans laquelle l'endomorphisme de \mathbf{R}^3 associé à M a une matrice diagonale.

c) Calculer M^n et en déduire la formule générale donnant u_n , v_n et w_n en fonction de n et de u_0 , v_0 et w_0 .

(21.5) Déterminer les valeurs propres et les vecteurs propres associés à la matrice $M = \begin{pmatrix} 3 & 1 \\ -2 & 0 \end{pmatrix}$. Calculer M^n pour tout $n \in \mathbf{N}$.

(21.6) Considérons la matrice $M = \begin{pmatrix} 2 & -1 & 1 \\ 1 & 0 & 1 \\ 1 & -1 & 2 \end{pmatrix}$. Quel est le polynôme minimal de M ? Montrer que, pour tout $n \in \mathbf{N}$, il existe deux réels α_n et β_n tels que l'on ait $M^n = \alpha_n M + \beta_n I_3$. Calculer α_n et β_n en fonction de n . On en déduit alors la forme explicite de M^n pour tout $n \in \mathbf{N}$.

(21.7) Déterminer le réel x pour que la matrice $\begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & x \\ 0 & 0 & 1 & -x \end{pmatrix}$ soit diagonalisable.

(21.8) Calculer les valeurs propres et les vecteurs propres des matrices suivantes. Sont-elles diagonalisables? Si oui, déterminer une base de vecteurs propres.

$$\begin{pmatrix} 4 & 1 \\ 0 & 3 \end{pmatrix}; \quad \begin{pmatrix} 2 & 4 \\ 1 & 1 \end{pmatrix}; \quad \begin{pmatrix} 2 & -1 \\ 1 & 3 \end{pmatrix}; \quad \begin{pmatrix} 1 & -1 & 1 \\ -1 & 1 & -3 \\ 1 & -3 & 1 \end{pmatrix}; \quad \begin{pmatrix} 1 & 2 & -2 \\ 2 & 1 & -2 \\ 2 & 2 & -3 \end{pmatrix}; \\ \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}; \quad \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}; \quad \begin{pmatrix} -7 & -2 & 1 \\ 28 & 8 & -4 \\ 31 & 10 & -5 \end{pmatrix}; \quad \begin{pmatrix} 7 & 4 & 0 & 0 \\ -12 & -7 & 0 & 0 \\ 20 & 11 & -6 & -7 \\ -12 & -6 & 6 & 6 \end{pmatrix}.$$

(21.9) Soit E l'espace vectoriel réel des polynômes de degré au plus n à coefficients réels.

a) Quelle est la dimension de E ?

b) Montrer que la dérivation est un endomorphisme de E . Quel est son noyau? Quel est son rang? Quelles sont ses valeurs propres? Est-il diagonalisable?

(21.10) On considère la matrice

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

a) Calculer toutes les puissances M^n de M , pour $n \in \mathbf{N}$.

b) Montrer que M est inversible; calculer toutes les puissances M^{-n} de M^{-1} , pour $n \in \mathbf{N}$.

(21.11) a) Soit M une matrice carrée complexe vérifiant $M^k = I$, pour un certain entier naturel k . Montrer que M est diagonalisable.

b) Soit u un endomorphisme diagonalisable d'un espace vectoriel réel vérifiant $u^k = \text{Id}_E$, pour un certain entier naturel k . Montrer que $u^2 = \text{Id}_E$.

(21.12) Soit u un endomorphisme diagonalisable d'un espace vectoriel E . Montrer que $E = \text{Ker}(u) \oplus u(E)$.

(21.13) Soit M une matrice carrée *réelle* telle qu'il existe une matrice inversible *complexe* P telle que PMP^{-1} soit diagonale à coefficients *réels*. Montrer qu'il existe une matrice inversible *réelle* Q telle que QMQ^{-1} soit diagonale (*Indication* : utiliser le théorème 20.9 et (20.10)).

(21.14) Dans chacun des cas suivants, trouver le p.g.c.d. R des polynômes P et Q , puis déterminer des polynômes A et B tels que $AP + BQ = R$.

a) $P(T) = T^4 + T^3 - T^2 - 2T - 2$ et $Q(T) = T^3 - 1$;

b) $P(T) = T^5 - T^4 - 2T^3 + T^2 - 3T + 2$ et $Q(T) = T^5 - 2T^4 - 3T + 6$;

(21.15) a) Montrer que tout polynôme de à coefficients réels est produit de polynômes réels de degré ≤ 2 . (*Indication* : grouper les racines complexes non réelles par paires de racines conjuguées).

b) Pour tout $n \in \mathbb{N}$, décomposer $T^n - 1$ en produit de polynômes réels de degré ≤ 2 . Grâce à la factorisation de $T^5 - 1$, calculer l'expression rationnelle de $\cos \frac{2\pi}{5}$.

(21.16) Trouver les polynômes minimaux des matrices

$$A = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \quad \text{et} \quad C = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix}.$$

(21.17) Montrer que le polynôme minimal de la matrice bloc $M = \begin{pmatrix} A & B \\ 0 & C \end{pmatrix}$ divise le produit $p_A p_C$ des polynômes minimaux de A et de C . Trouver un exemple où $p_M = p_A p_C$ et un exemple où le degré de p_M est strictement inférieur à celui de $p_A p_C$.

(21.18) On rappelle qu'un endomorphisme u d'un espace vectoriel est dit *nilpotent* s'il existe un entier N tel que $u^N = 0$. Soit u un endomorphisme nilpotent d'un espace vectoriel E de dimension n .

a) Montrer que $u^n = 0$.

b) Montrer que toutes les valeurs propres de u sont nulles; u peut-il être diagonalisable? Quel est l'espace caractéristique E'_0 ? Quel est le polynôme caractéristique de u ? Calculer $\det(\text{Id}_E + u)$.

c) Montrer qu'il existe une base de E dans laquelle la matrice $M = (a_{ij})$ de u vérifie $a_{ij} = 0$ si $i \geq j$.

d) Supposons $u^{n-1} \neq 0$; soit x un vecteur tel que $u^{n-1}(x) \neq 0$. Montrer que les vecteurs $x, u(x), u^2(x), \dots, u^{n-1}(x)$ forment une base de E . Quelle est la matrice de u dans cette base?

e) En vous servant de d), déterminer des bases dans lesquelles les matrices

$$A = \begin{pmatrix} 0 & 1 & 0 \\ -5 & 4 & 1 \\ -2 & 1 & 2 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 125 & -75 & 15 \end{pmatrix} \quad \text{et} \quad C = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 12 & -16 & 7 \end{pmatrix}$$

sont triangulaires. Déterminer des matrices R, S et T telles que $R^2 = A, S^2 = B$ et $T^2 = C$.

III. EXPONENTIELLES DE MATRICES

Le but de ce chapitre est de définir l'exponentielle e^M d'une matrice complexe carrée M par la même série qui définit l'exponentielle d'un nombre complexe, à savoir $e^M = \sum_{k=0}^{\infty} M^k/k!$. Pour cela, nous avons besoin de définir la somme d'une série de matrices ; cela nécessite l'introduction de la notion de norme sur un espace vectoriel réel ou complexe (pour mesurer les « distances »), puis de celle de convergence d'une suite ou d'une série de vecteurs. Nous montrons quelques propriétés de l'exponentielle, puis introduisons les notions de base sur la continuité et la dérivabilité des fonctions à valeurs vectorielles, en vue du chapitre suivant.

22. Normes sur un espace vectoriel

Commençons par quelques rappels sur la convergence des suites et séries de nombres réels.

Vous avez vu en première année qu'une suite (u_m) de nombres réels est dite convergente vers une limite ℓ si, pour tout $\varepsilon > 0$, il existe un entier m_0 tel que, pour tout $m \geq m_0$, on ait $|u_m - \ell| < \varepsilon$. On peut exprimer cette dernière inégalité en disant que la « distance » entre u_m et ℓ est moindre que ε .

Essayons de généraliser cette notion à des suites vectorielles (u_m) , où chaque terme de la suite est dans un espace vectoriel fixé E . La seule chose à adapter est la notion de « distance » entre deux vecteurs.

Prenons pour fixer les idées $E = \mathbf{R}^2$; on est habitué à considérer la distance dite euclidienne entre des vecteurs (x_1, y_1) et (x_2, y_2) , donnée par $\|(x_1, y_1) - (x_2, y_2)\|_2$, où l'on a posé

$$\|(x, y)\|_2 = \sqrt{x^2 + y^2}.$$

On peut alors copier la définition de la convergence : on dit qu'une suite (u_m) de vecteurs de \mathbf{R}^2 converge vers un vecteur ℓ si, pour tout $\varepsilon > 0$, il existe un entier m_0 tel que, pour tout $m \geq m_0$, on ait $\|u_m - \ell\|_2 < \varepsilon$.

On peut aussi considérer les composantes de u_m et de ℓ , c'est-à-dire poser $u_m = (v_m, w_m)$ et $\ell = (a, b)$, et demander que la suite réelle (v_m) converge vers a , tandis que la suite (w_m) converge vers b . Il n'est pas difficile de voir que cela revient à demander que, pour tout $\varepsilon > 0$, il existe un entier m_0 tel que, pour tout $m \geq m_0$, on ait $\|u_m - \ell\|_{\infty} < \varepsilon$, où l'on a posé

$$\|(x, y)\|_{\infty} = \max(|x|, |y|)$$

(ces notations $\|\cdot\|_2$ et $\|\cdot\|_{\infty}$ sont traditionnelles). Comment vérifier que ces deux notions de convergence coïncident ? Il n'est pas difficile de montrer (*cf.* ex. 22.4) que l'on a

$$(22.1) \quad \frac{1}{\sqrt{2}} \|(x, y)\|_2 \leq \|(x, y)\|_{\infty} \leq \|(x, y)\|_2$$

pour tout vecteur (x, y) .

(22.2) On en déduit que les deux notions de convergence sont équivalentes : supposons que la suite (u_m) converge dans le premier sens. Soit $\varepsilon > 0$; il existe un entier m_0 tel que, pour tout $m \geq m_0$, on ait $\|u_m - \ell\|_2 < \varepsilon$, ce qui entraîne

$$\|u_m - \ell\|_\infty \leq \|u_m - \ell\|_2 < \varepsilon ,$$

de sorte que la suite converge au deuxième sens. Inversement, supposons que la suite (u_m) converge dans le second sens. Soit $\varepsilon > 0$; il existe un entier m_0 tel que, pour tout $m \geq m_0$, on ait $\|u_m - \ell\|_\infty < \varepsilon/\sqrt{2}$, ce qui entraîne

$$\|u_m - \ell\|_2 \leq \sqrt{2} \|u_m - \ell\|_\infty < \varepsilon ,$$

de sorte que la suite converge au premier sens.

Déjà dans le cadre simple de cette discussion, on voit apparaître les points suivants :

1) il faut un moyen de mesurer la « distance » entre deux vecteurs, mais il ne semble pas y avoir de moyen unique ;

2) les notions de convergence définies à l'aide de ces diverses distances seront néanmoins les mêmes si l'on arrive à les comparer comme dans (22.1).

Pour répondre au point 1), on est amené à poser la définition suivante.

Définition 22.3.— Soit E un espace vectoriel réel ou complexe ; on appelle norme sur E toute application $\| \cdot \| : E \rightarrow \mathbf{R}^+$ telle que, pour tout x dans E et tout scalaire λ , on ait

- 1) $\|x\| = 0 \iff x = 0$;
- 2) $\|\lambda x\| = |\lambda| \|x\|$;
- 3) $\|x + y\| \leq \|x\| + \|y\|$ (inégalité triangulaire).

Ici, $|\lambda|$ désigne la valeur absolue de λ si λ est réel, son module s'il est complexe.

Exemple 22.4. Sur $E = \mathbf{R}^n$ ou \mathbf{C}^n , on utilisera souvent les normes suivantes (de nouveau, les notations sont traditionnelles) :

$$\begin{aligned} \|(x_1, \dots, x_n)\|_2 &= \sqrt{|x_1|^2 + \dots + |x_n|^2} , \\ \|(x_1, \dots, x_n)\|_1 &= |x_1| + \dots + |x_n| , \\ \|(x_1, \dots, x_n)\|_\infty &= \max(|x_1|, \dots, |x_n|) . \end{aligned}$$

Les propriétés 1) et 2) de la définition sont faciles à vérifier ; l'inégalité triangulaire est aussi facile pour $\| \cdot \|_1$ et $\| \cdot \|_\infty$, moins pour $\| \cdot \|_2$. Nous l'admettrons pour le moment. Montrons les inégalités suivantes, analogues à celles de (22.1) :

$$\|x\|_\infty \leq \|x\|_2 \leq \|x\|_1 \leq n\|x\|_\infty ,$$

pour tout vecteur x . La première et la troisième sont évidentes ; la seconde résulte de

$$\|x\|_1^2 = (|x_1| + \dots + |x_n|)^2 \geq x_1^2 + \dots + x_n^2 = \|x\|_2^2.$$

Les inégalités ci-dessus se généralisent : on peut toujours « comparer » deux normes $\|\cdot\|$ et $\|\cdot\|'$ sur un espace vectoriel E réel ou complexe de dimension finie, car elles sont *équivalentes*, c'est-à-dire qu'il existe des réels strictement positifs a et b tels que, pour tout x dans E , on ait

$$a \|x\| \leq \|x\|' \leq b \|x\|.$$

Nous ne démontrerons pas ce théorème.

23. Suites et séries vectorielles

Le raisonnement suivi en (22.2) montre que la définition suivante ne dépend pas du choix de la norme $\|\cdot\|$ sur E .

Définition 23.1.— Soient E un espace vectoriel réel ou complexe et (u_m) une suite d'éléments de E . On dit que la suite (u_m) converge vers un élément ℓ de E si, pour tout $\varepsilon > 0$, il existe un entier m_0 tel que, pour tout $m \geq m_0$, on ait $\|u_m - \ell\| < \varepsilon$.

Cette notion ne dépendant pas de la norme, on peut choisir sur \mathbf{R}^n la norme $\|\cdot\|_\infty$; on en déduit que pour qu'une suite $(u_m)_{m \in \mathbf{N}} = ((u_{m,1}, \dots, u_{m,n}))_{m \in \mathbf{N}}$ d'éléments de \mathbf{R}^n converge vers $\ell = (\ell_1, \dots, \ell_n)$, il faut et il suffit que pour chaque $j = 1, \dots, n$, la suite des j èmes composantes $(u_{m,j})_{m \in \mathbf{N}}$ converge vers ℓ_j . Plus généralement, on a la même assertion dans un espace vectoriel réel ou complexe quelconque E , une fois que l'on a choisi une base (pour pouvoir parler de composantes d'un vecteur). Cette remarque fait que la plupart des théorèmes que vous connaissez sur les limites de suites réelles se transportent sans modification au cas des suites vectorielles : par exemple, une suite convergente est bornée, la limite d'une somme est la somme des limites (mais pas le produit, car nous n'avons pas défini de produit de vecteurs), de même, toute suite de Cauchy, c'est-à-dire toute suite qui vérifie

$$\forall \varepsilon > 0 \quad \exists m_0 \quad \forall p \geq m_0 \quad \forall q \geq m_0 \quad \|u_p - u_q\| < \varepsilon$$

est convergente.

On peut aussi définir la convergence d'une série $\sum u_k$, où u_k est dans E , puisque chacun sait qu'une série n'est autre qu'une suite. Le critère de Cauchy s'écrit dans ce cas : pour qu'une série $\sum u_k$ soit convergente, il faut et il suffit que, pour tout $\varepsilon > 0$, il existe un entier m_0 tel que, pour tout $p > q \geq m_0$, on ait

$$\|u_{q+1} + u_{q+2} + \dots + u_p\| < \varepsilon.$$

Définition 23.2.— Soit E un espace vectoriel réel ou complexe. On dit qu'une série $\sum u_k$ d'éléments de E converge absolument si la série réelle $\sum \|u_k\|$ converge.

On vérifie que cette notion ne dépend pas de la norme choisie puisqu'elles sont toutes équivalentes (se rappeler qu'une série à termes positifs converge si et seulement si ses sommes partielles sont bornées). Le critère de Cauchy donne immédiatement :

Proposition 23.3.— *Soit E un espace vectoriel réel ou complexe. Toute série absolument convergente d'éléments de E est convergente.*

Démonstration. On vérifie le critère de Cauchy. On se donne $\varepsilon > 0$; comme la série $\sum \|u_k\|$ converge, elle est de Cauchy, et il existe un entier m_0 tel que, pour tout $p > q \geq m_0$, on ait $\sum_{k=q+1}^p \|u_k\| < \varepsilon$. Grâce à l'inégalité triangulaire, on en déduit

$$\left\| \sum_{k=q+1}^p u_k \right\| \leq \sum_{k=q+1}^p \|u_k\| < \varepsilon .$$

Cela montre que la suite $\left(\sum_{k=0}^m u_k \right)_m$ est de Cauchy, donc elle converge. ■

Pour tout m , on a par l'inégalité triangulaire

$$\left\| \sum_{k=0}^m u_k \right\| \leq \sum_{k=0}^m \|u_k\| \leq \sum_{k=0}^{\infty} \|u_k\| ;$$

en faisant tendre m vers l'infini, on obtient, si la série $\sum u_k$ converge absolument,

$$(23.4) \quad \left\| \sum_{k=0}^{\infty} u_k \right\| \leq \sum_{k=0}^{\infty} \|u_k\| .$$

24. Suites et séries de matrices

Les matrices carrées d'ordre n (à coefficients réels ou complexes) forment un espace vectoriel. On peut donc considérer des normes sur cet espace vectoriel, et des suites et des séries convergentes de matrices. Par exemple, la suite de matrices de terme général

$$U_m = \begin{pmatrix} m \sin(1/m) & \frac{m}{m^2+1} \\ \frac{2m+3}{m} & 4 \end{pmatrix}$$

converge vers la matrice $\begin{pmatrix} 1 & 0 \\ 2 & 4 \end{pmatrix}$ lorsque m tend vers l'infini.

Le fait que l'on puisse multiplier les matrices entre elles nous permet aussi d'étendre les théorèmes sur la limite d'un produit de suites ou de séries. Il faut quand même faire très attention que le produit des matrices n'est pas commutatif. Redémontrons donc ces résultats rapidement. Toutes les normes sont équivalentes : on peut donc utiliser n'importe laquelle. Dans les questions qui font intervenir des produits de matrices, il est souvent pratique d'utiliser la norme $\| \cdot \|_1$ introduite dans l'exemple 22.4, parce qu'elle vérifie la propriété utile suivante : pour toutes matrices carrées M et N , on a

$$\|MN\|_1 \leq \|M\|_1 \|N\|_1 .$$

En effet, si $M = (a_{ij})$ et $N = (b_{ij})$, on a

$$\|MN\|_1 = \sum_{i,j} \left| \sum_k a_{ik} b_{kj} \right| \leq \sum_{i,j,k} |a_{ik}| |b_{kj}| \leq \left(\sum_{i,k} |a_{ik}| \right) \left(\sum_{l,j} |b_{lj}| \right) = \|M\|_1 \|N\|_1 .$$

Proposition 24.1.— Soient (U_m) une suite de matrices qui converge vers U et (V_m) une suite de matrices qui converge vers V . La suite $(U_m V_m)$ converge vers la matrice UV .

Démonstration. On se donne comme d'habitude $\varepsilon > 0$; il existe un entier m_0 tel que, pour tout $m \geq m_0$, on ait

$$\|U_m - U\|_1 < \frac{\varepsilon}{2(\varepsilon + \|V\|_1)} \quad \text{et} \quad \|V_m - V\|_1 < \frac{\varepsilon}{2(\|U\|_1 + 1)} ,$$

qui entraîne en particulier $\|V_m\|_1 \leq \|V_m - V\|_1 + \|V\|_1 < \varepsilon + \|V\|_1$. Pour $m \geq m_0$, on a alors

$$\begin{aligned} \|U_m V_m - UV\|_1 &= \|(U_m V_m - UV_m) + (UV_m - UV)\|_1 \\ &\leq \|(U_m - U)V_m\|_1 + \|U(V_m - V)\|_1 \\ &\leq \|U_m - U\|_1 \|V_m\|_1 + \|U\|_1 \|V_m - V\|_1 \\ &\leq \|U_m - U\|_1 (\varepsilon + \|V\|_1) + \|U\|_1 \|V_m - V\|_1 \\ &< (\varepsilon + \|V\|_1) \frac{\varepsilon}{2(\varepsilon + \|V\|_1)} + \|U\|_1 \frac{\varepsilon}{2(\|U\|_1 + 1)} < \varepsilon , \end{aligned}$$

ce qui démontre la proposition. ■

Exemple 24.2. Si une suite (U_m) de matrices inversibles converge vers une matrice U et que la suite des inverses (U_m^{-1}) converge vers une matrice V , la matrice U est inversible, d'inverse V .

La construction d'une série qui converge vers le produit des sommes de deux séries convergentes $\sum U_k$ et $\sum V_k$ est plus délicate.

Proposition 24.3.— Soient $\sum U_k$ et $\sum V_k$ des séries de matrices absolument convergentes de sommes respectives U et V . La série de terme général

$$W_k = U_0 V_k + U_1 V_{k-1} + \cdots + U_k V_0$$

est absolument convergente de somme UV .

Démonstration. On remarque que

$$\begin{aligned} \sum_{k=0}^m \|W_k\|_1 &= \sum_{i,j \geq 0, i+j \leq m} \|U_i V_j\|_1 \leq \sum_{i,j \geq 0, i+j \leq m} \|U_i\|_1 \|V_j\|_1 \\ &\leq \left(\sum_{i=0}^m \|U_i\|_1 \right) \left(\sum_{j=0}^m \|V_j\|_1 \right) \leq \left(\sum_{i=0}^{\infty} \|U_i\|_1 \right) \left(\sum_{j=0}^{\infty} \|V_j\|_1 \right) , \end{aligned}$$

puisque les séries $\sum \|U_i\|_1$ et $\sum \|V_j\|_1$ sont convergentes. Les sommes partielles de la série à termes positifs $\sum \|W_k\|_1$ sont majorées ; elle est donc convergente, de sorte que la série $\sum W_k$ est absolument convergente. Pour calculer sa somme, on regarde la différence

$$\begin{aligned} \left\| \sum_{k=0}^{2m} W_k - \left(\sum_{i=0}^m U_i \right) \left(\sum_{j=0}^m V_j \right) \right\|_1 &= \left\| \sum_{\substack{i > m \text{ ou } j > m \\ i+j \leq 2m}} U_i V_j \right\|_1 \\ &\leq \sum_{\substack{i > m \text{ ou } j > m \\ i+j \leq 2m}} \|U_i\|_1 \|V_j\|_1 \\ &\leq \left(\sum_{i=0}^{2m} \|U_i\|_1 \right) \left(\sum_{j=0}^{2m} \|V_j\|_1 \right) - \left(\sum_{i=0}^m \|U_i\|_1 \right) \left(\sum_{j=0}^m \|V_j\|_1 \right). \end{aligned}$$

Comme le membre de droite tend vers 0 lorsque m vers $+\infty$, et que la limite de la suite $\left(\left(\sum_{i=0}^m U_i \right) \left(\sum_{j=0}^m V_j \right) \right)$ est UV par la proposition 24.1, on en déduit que la série de terme général W_k est convergente de somme UV . ■

25. Exponentielle de matrices

Nous pouvons maintenant définir l'exponentielle d'une matrice carrée M , réelle ou complexe. On se contente de copier la définition de l'exponentielle usuelle :

$$\exp M = e^M = \sum_{k=0}^{\infty} \frac{M^k}{k!}$$

Cela a un sens, puisque cette série est *absolument convergente* : on a pour tout k

$$\left\| \frac{M^k}{k!} \right\|_1 \leq \frac{\|M\|_1^k}{k!}$$

et la série de terme général le membre de droite est convergente, de somme $e^{\|M\|_1}$.

Exemples 25.1. 1) Si M est une matrice diagonale $\begin{pmatrix} \lambda_1 & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2 & 0 & \cdots & 0 \\ 0 & 0 & \cdot & & \cdot \\ \cdot & \cdot & & & \cdot \\ 0 & 0 & & & \lambda_n \end{pmatrix}$, sa puissance

k ième est $\begin{pmatrix} \lambda_1^k & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2^k & 0 & \cdots & 0 \\ 0 & 0 & \cdot & & \cdot \\ \cdot & \cdot & & & \cdot \\ 0 & 0 & & & \lambda_n^k \end{pmatrix}$ et l'on a

$$e^M = \begin{pmatrix} e^{\lambda_1} & 0 & 0 & \cdots & 0 \\ 0 & e^{\lambda_2} & 0 & \cdots & 0 \\ 0 & 0 & \cdot & & \cdot \\ \cdot & \cdot & & & \cdot \\ 0 & 0 & & & e^{\lambda_n} \end{pmatrix}.$$

Plus généralement, l'exponentielle d'une matrice $\begin{pmatrix} M_1 & 0 & 0 & \cdots & 0 \\ 0 & M_2 & 0 & \cdots & 0 \\ 0 & 0 & \cdot & & \cdot \\ \cdot & \cdot & & & \cdot \\ 0 & 0 & & & M_n \end{pmatrix}$ « dia-

gonale par blocs » est

$$e^M = \begin{pmatrix} e^{M_1} & 0 & 0 & \cdots & 0 \\ 0 & e^{M_2} & 0 & \cdots & 0 \\ 0 & 0 & \cdot & & \cdot \\ \cdot & \cdot & & & \cdot \\ 0 & 0 & & & e^{M_n} \end{pmatrix}.$$

2) Si $M = \begin{pmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$, on a $M^2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ et $M^3 = 0$, de sorte que

$$e^M = I_3 + M + \frac{M^2}{2!} = \begin{pmatrix} 1 & 1 & 3/2 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}.$$

La propriété fondamentale de l'exponentielle usuelle est qu'elle transforme sommes en produits. Ce n'est plus vrai en général pour les exponentielles de matrices : si on avait $e^{M+N} = e^M e^N$, cela entraîne en particulier que e^M et e^N commutent ; or on vérifie facilement que

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

ne commutent pas (on verra d'ailleurs dans l'exemple 25.5 que l'exponentielle de la matrice $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ est $\frac{1}{2} \begin{pmatrix} e + 1/e & e - 1/e \\ e - 1/e & e + 1/e \end{pmatrix}$). Cependant, on a le résultat fondamental suivant.

Proposition 25.2. — Si M et N sont des matrices carrées qui commutent (c'est-à-dire qui vérifient $MN = NM$), on a

$$e^{M+N} = e^M e^N = e^N e^M.$$

Démonstration. Le point essentiel est que la formule du binôme, qui donne le développement de $(M+N)^k$, est encore valable pour des matrices qui commutent (on le démontre par récurrence sur k). On a donc

$$(M+N)^k = \sum_{j=0}^k \mathbf{C}_k^j M^{k-j} N^j = \sum_{j=0}^k \frac{k!}{(k-j)!j!} M^{k-j} N^j,$$

de sorte que

$$e^{M+N} = \sum_{k=0}^{\infty} \frac{(M+N)^k}{k!} = \sum_{k=0}^{\infty} \left(\sum_{j=0}^k \frac{M^{k-j}}{(k-j)!} \frac{N^j}{j!} \right).$$

On retrouve la série dont la proposition 24.3 nous dit qu'elle converge vers

$$\left(\sum_{i=0}^{\infty} \frac{M^i}{i!} \right) \left(\sum_{j=0}^{\infty} \frac{N^j}{j!} \right),$$

c'est-à-dire $e^M e^N$. ■

Corollaire 25.3.— Si M est une matrice carrée, e^M est inversible, d'inverse e^{-M} .

Démonstration. Il suffit d'appliquer la proposition à M et $-M$, qui commutent, et d'utiliser $e^{0_n} = I_n$. ■

Proposition 25.4.— Si M et N sont des matrices carrées semblables, c'est-à-dire qu'il existe une matrice inversible P telle que $N = P^{-1}MP$, alors e^M et e^N sont semblables; plus précisément

$$e^{P^{-1}MP} = P^{-1}e^M P.$$

Démonstration. Il suffit de remarquer que pour tout $k \geq 0$, on a

$$(P^{-1}MP)^k = \overbrace{(P^{-1}MP)(P^{-1}MP) \cdots (P^{-1}MP)}^{k \text{ fois}} = P^{-1}M^k P,$$

de sorte que

$$\sum_{k=0}^m \frac{(P^{-1}MP)^k}{k!} = P^{-1} \left(\sum_{k=0}^m \frac{M^k}{k!} \right) P.$$

Il suffit alors de faire tendre m vers l'infini, en utilisant la proposition 24.1, pour conclure. ■

Cette proposition va nous permettre d'utiliser nos résultats sur la réduction des matrices, puisqu'on a vu dans l'exemple 25.1.1) comment calculer l'exponentielle d'une matrice diagonale.

Exemple 25.5. La matrice $M = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ est diagonalisable, de valeurs propres 1 et -1 , une base de vecteurs propres étant $(1, 1)$ et $(1, -1)$. On a donc

$$P = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \quad P^{-1} = \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & -1/2 \end{pmatrix} \quad M = P \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} P^{-1},$$

d'où

$$e^M = P \begin{pmatrix} e & 0 \\ 0 & e^{-1} \end{pmatrix} P^{-1} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} e & 0 \\ 0 & e^{-1} \end{pmatrix} \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & -1/2 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} e + e^{-1} & e - e^{-1} \\ e - e^{-1} & e + e^{-1} \end{pmatrix}.$$

On peut aussi, dans ce cas particulier, remarquer que $M^2 = I_2$ pour obtenir e^M :

$$\begin{aligned} e^M &= \sum_{k=0}^{\infty} \frac{M^k}{k!} = \sum_{k=0}^{\infty} \frac{M^{2k}}{(2k)!} + \sum_{k=0}^{\infty} \frac{M^{2k+1}}{(2k+1)!} \\ &= \left(\sum_{k=0}^{\infty} \frac{1}{(2k)!} \right) I_2 + \left(\sum_{k=0}^{\infty} \frac{1}{(2k+1)!} \right) M = \frac{1}{2} (e + e^{-1}) I_2 + \frac{1}{2} (e - e^{-1}) M . \end{aligned}$$

Lorsqu'une matrice complexe n'est pas diagonalisable, on a quand même à notre disposition sa décomposition de Dunford $M = D + N$, où D est diagonalisable, N nilpotente et $DN = ND$ (th. 20.8). On a

$$e^M = e^D e^N ,$$

où e^D se calcule comme ci-dessus et

$$e^N = I_n + N + \frac{N^2}{2!} + \cdots + \frac{N^{n-1}}{(n-1)!} .$$

Si par exemple M a une seule valeur propre λ , sa décomposition de Dunford est $M = \lambda I_n + (M - \lambda I_n)$ et

$$e^M = e^\lambda (I_n + (M - \lambda I_n) + \frac{(M - \lambda I_n)^2}{2!} + \cdots + \frac{(M - \lambda I_n)^{n-1}}{(n-1)!}) .$$

Dans le cas général, M est semblable à une matrice diagonale par blocs, où chaque bloc a une seule valeur propre, et on applique la méthode ci-dessus à chaque bloc.

Exemples 25.6. 1) On vérifie par le calcul que le polynôme caractéristique de la matrice

$$M = \begin{pmatrix} -4 & 1 & 1 \\ 1 & -1 & -2 \\ -2 & 1 & -1 \end{pmatrix} \text{ est } (T + 2)^3 ; \text{ elle a donc } -2 \text{ comme unique valeur propre et}$$

$$e^M = e^{-2I_3} e^{M+2I_3} = e^{-2} (I_3 + (M + 2I_3) + \frac{1}{2} (M + 2I_3)^2) .$$

On calcule

$$(M + 2I_3) = \begin{pmatrix} -2 & 1 & 1 \\ 1 & 1 & -2 \\ -2 & 1 & 1 \end{pmatrix} \quad (M + 2I_3)^2 = \begin{pmatrix} 3 & 0 & -3 \\ 3 & 0 & -3 \\ 3 & 0 & -3 \end{pmatrix} ,$$

d'où

$$e^M = e^{-2} \begin{pmatrix} 1/2 & 1 & -1/2 \\ 5/2 & 2 & -7/2 \\ -1/2 & 1 & 1/2 \end{pmatrix} .$$

2) On a vu dans l'exemple 20.11.3) que la matrice $M = \begin{pmatrix} 2 & -1 & 2 \\ 10 & -5 & 7 \\ 4 & -2 & 2 \end{pmatrix}$ s'écrit sous la forme

$$M = P M' P^{-1} ,$$

avec

$$P = \begin{pmatrix} 1 & 1 & 0 \\ -1 & 2 & 1 \\ -2 & 0 & 1 \end{pmatrix} \quad M' = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \quad P^{-1} = \begin{pmatrix} 2 & -1 & 1 \\ -1 & 1 & -1 \\ 4 & -2 & 3 \end{pmatrix} .$$

La matrice M' est « diagonale par blocs » ; on en déduit

$$e^{M'} = \begin{pmatrix} e^{-1} & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

et

$$\begin{aligned} e^M &= P e^{M'} P^{-1} = \begin{pmatrix} 1 & 1 & 0 \\ -1 & 2 & 1 \\ -2 & 0 & 1 \end{pmatrix} \begin{pmatrix} e^{-1} & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 2 & -1 & 1 \\ -1 & 1 & -1 \\ 4 & -2 & 3 \end{pmatrix} \\ &= \begin{pmatrix} 2e^{-1} + 3 & -e^{-1} - 1 & e^{-1} + 2 \\ -2e^{-1} + 10 & e^{-1} - 4 & -e^{-1} + 7 \\ -4e^{-1} + 4 & 2e^{-1} - 2 & -2e^{-1} + 3 \end{pmatrix} . \end{aligned}$$

Pour conclure, il n'est pas inutile de remarquer qu'en un certain sens, les séries (entières) de matrices n'existent pas (c'était bien la peine de se fatiguer...). Expliquons-nous : si M est une matrice carrée d'ordre n , le théorème de Hamilton-Cayley dit que $P_M(M)$ est nul. Ceci signifie que la matrice M^n est combinaison linéaire à coefficients scalaires des matrices I_n, M, \dots, M^{n-1} . Cela entraîne, par récurrence sur l'entier naturel k , que toutes les puissances M^k de M ont la même propriété, donc aussi les sommes partielles de la série qui définit e^M . On peut en déduire (exercice) qu'il existe des scalaires a_0, a_1, \dots, a_{n-1} tels que

$$e^M = a_0 I_n + a_1 M + \dots + a_{n-1} M^{n-1} .$$

Néanmoins, on n'a pas travaillé pour rien : le formalisme des séries nous a permis de démontrer facilement les propriétés essentielles de l'exponentielle (prop. 25.2 et 25.4).

26. Continuité et dérivabilité des fonctions vectorielles

Maintenant que l'on a introduit la notion de norme sur un espace vectoriel, il n'est pas difficile d'étendre les définitions que vous connaissez de la continuité et dérivabilité des fonctions réelles de variable réelle aux fonctions vectorielles.

Lorsque f est une fonction à valeurs réelles définie sur un sous-ensemble D de \mathbf{R} , rappelons que f est continue en un point t_0 de D si, pour tout $\varepsilon > 0$, la distance entre $f(t)$ et $f(t_0)$ peut être rendue moindre que ε pour tout t dans D assez proche de t_0 ; en d'autres termes, il existe $\delta > 0$ tel que, pour tout t dans D tel que $|t - t_0| < \delta$, on ait $|f(t) - f(t_0)| < \varepsilon$.

Si on considère maintenant une fonction f définie sur D , mais à valeurs dans un espace vectoriel E sur lequel on a choisi une norme, on peut adapter la définition de la

façon suivante : on dit que f est continue en un point t_0 de D si, pour tout $\varepsilon > 0$, il existe $\delta > 0$ tel que, pour tout t dans D tel que $|t - t_0| < \delta$, on ait $\|f(t) - f(t_0)\| < \varepsilon$.

Dans le cas $E = \mathbf{R}^n$, la donnée de f est équivalente à celle de ses n fonctions composantes $f_1, \dots, f_n : D \rightarrow \mathbf{R}$ définies, pour tout t dans D , par

$$f(t) = (f_1(t), \dots, f_n(t)) .$$

Si on choisit comme norme sur \mathbf{R}^n la norme $\| \cdot \|_\infty$, on voit tout de suite que la continuité de f en un point t_0 de D est équivalente à celle de chacune des fonctions f_1, \dots, f_n en ce même point.

Il faut noter que les fonctions à valeurs vectorielles se rencontrent couramment en physique par exemple : si l'on veut décrire le mouvement d'un mobile ponctuel qui se déplace dans l'espace \mathbf{R}^3 , on a recours à une fonction f à valeurs dans \mathbf{R}^3 qui décrit au temps t la position $f(t)$ du mobile. Ces fonctions sont en général continues.

Exemples 26.1. 1) La fonction $t \mapsto (t^2, t^3, \sin t)$ est continue sur \mathbf{R} , puisque chacune de ses composantes l'est.

2) La fonction $t \mapsto (1/t, \sqrt{1-t}, \sqrt{1+t})$ est définie sur $]-1, 0[\cup]0, 1]$ et y est continue en tout point.

3) La fonction $t \mapsto \begin{pmatrix} 1/t & t^2 + 1 \\ \sqrt{t} & 2 \end{pmatrix}$ est définie sur $]0, +\infty[$ et y est continue en tout point.

On définit de façon analogue la limite d'une fonction vectorielle en un point. De nouveau, si f est une fonction de D dans \mathbf{R}^n de composantes $f_1, \dots, f_n : D \rightarrow \mathbf{R}$, la fonction f a pour limite un vecteur $\ell = (\ell_1, \dots, \ell_n)$ quand t tend vers t_0 si et seulement si chaque f_j tend vers ℓ_j lorsque t tend vers t_0 . Cela nous permet de montrer que la limite d'une somme est la somme des limites et autres théorèmes généraux classiques. Lorsque les fonctions que l'on considère sont à valeurs matricielles, on peut aussi parler de produit.

Proposition 26.2.— Soit D un sous-ensemble de \mathbf{R} , soient $F : D \rightarrow \mathcal{M}_{m \times n}(\mathbf{C})$ et $G : D \rightarrow \mathcal{M}_{n \times p}(\mathbf{C})$ des fonctions à valeurs matricielles, admettant des limites en un point t_0 de D . Alors la fonction $FG : D \rightarrow \mathcal{M}_{m \times p}(\mathbf{C})$, qui à t associe $F(t)G(t)$, a une limite en t_0 et

$$\lim_{t \rightarrow t_0} (F(t)G(t)) = \left(\lim_{t \rightarrow t_0} F(t) \right) \left(\lim_{t \rightarrow t_0} G(t) \right) .$$

Démonstration. La démonstration est tout-à-fait analogue à la démonstration pour les fonctions à valeurs réelles (voir aussi la démonstration de la proposition 24.1); elle sera laissée au lecteur. ■

Passons à la dérivabilité des fonctions à valeurs vectorielles; pas de problème particulier non plus : on dit qu'une fonction f définie sur une partie D de \mathbf{R} (réunion d'intervalles

ouverts) et à valeurs dans un espace vectoriel E est dérivable en un point t_0 de D si la limite du vecteur

$$\frac{f(t) - f(t_0)}{t - t_0}$$

existe quand t tend vers t_0 . On notera que $f(t) - f(t_0)$ est un élément de E , on peut donc le diviser par le scalaire $t - t_0$ si celui-ci n'est pas nul. La limite de ce rapport, lorsqu'elle existe, s'appelle la dérivée de f en t_0 ; elle est notée $f'(t_0)$, c'est un élément de E .

Dans l'exemple d'une fonction f à valeurs dans \mathbf{R}^3 décrivant le mouvement d'un mobile ponctuel qui se déplace dans l'espace, le vecteur $f'(t_0)$ (lorsqu'il existe), s'appelle le « vecteur vitesse ». Sa norme est la vitesse du mobile à l'instant t_0 ; il est dirigé selon la tangente à la trajectoire du mobile.

Dans le cas $E = \mathbf{R}^n$, on vérifie sans mal que la dérivabilité de f en t_0 est équivalente à celle de chacune des fonctions f_1, \dots, f_n en ce même point. Le vecteur dérivé $f'(t_0)$ est alors le vecteur de composantes $(f'_1(t_0), \dots, f'_n(t_0))$.

Exemples 26.3. 1) La fonction $t \mapsto (t^2, t^3, \sin t)$ est dérivable sur \mathbf{R} , de dérivée en t_0 le vecteur $(2t_0, 3t_0^2, \cos t_0)$

2) La fonction $t \mapsto (1/t, \sqrt{1-t}, \sqrt{1+t})$ est dérivable sur $] -1, 0[\cup]0, 1[$, de dérivée en t_0 le vecteur $(-1/t_0^2, \frac{-1}{2\sqrt{1-t_0}}, \frac{1}{2\sqrt{1+t_0}})$.

3) La fonction $t \mapsto \begin{pmatrix} 1/t & t^2 + 1 \\ \sqrt{t} & 2 \end{pmatrix}$ est dérivable sur $]0, +\infty[$, de dérivée en t_0 le vecteur $\begin{pmatrix} -1/t_0^2 & 2t_0 \\ 1/(2\sqrt{t_0}) & 0 \end{pmatrix}$.

On montre sans difficulté que la dérivée d'une somme est la somme des dérivées et autres théorèmes généraux classiques. Lorsque les fonctions que l'on considère sont à valeurs matricielles, on peut aussi parler de produit. On montre une formule analogue à celle de la dérivée d'un produit de fonctions à valeurs réelles (il faut faire bien attention en appliquant la proposition qui suit de ne pas changer l'ordre dans les multiplications : le produit des matrices n'est pas commutatif!).

Proposition 26.4.— Soient $F : D \rightarrow \mathcal{M}_{m \times n}(\mathbf{C})$ et $G : D \rightarrow \mathcal{M}_{n \times p}(\mathbf{C})$ des fonctions à valeurs matricielles, définies au voisinage d'un point t_0 de \mathbf{R} et dérivables en ce point. Alors la fonction FG est dérivable en t_0 et

$$(FG)'(t_0) = F'(t_0)G(t_0) + F(t_0)G'(t_0) .$$

Démonstration. Il s'agit de trouver la limite du taux d'accroissement

$$\delta(t) = \frac{F(t)G(t) - F(t_0)G(t_0)}{t - t_0}$$

lorsque t tend vers t_0 . On l'écrit

$$\begin{aligned}\delta(t) &= \frac{F(t)G(t) - F(t_0)G(t) + F(t_0)G(t) - F(t_0)G(t_0)}{t - t_0} \\ &= \frac{F(t) - F(t_0)}{t - t_0} G(t) + F(t_0) \frac{G(t) - G(t_0)}{t - t_0},\end{aligned}$$

qui, grâce à la proposition 26.2, tend vers $F'(t_0)G(t_0) + F(t_0)G'(t_0)$ lorsque t tend vers t_0 . ■

Voici enfin l'exemple principal qui nous servira dans le chapitre suivant.

Proposition 26.5.— *Soit M une matrice carrée. La fonction $t \mapsto e^{tM}$ est dérivable sur \mathbf{R} , de dérivée en t_0 la matrice Me^{t_0M} (aussi égale à $e^{t_0M}M$).*

Démonstration. Commençons par démontrer le résultat préliminaire suivant.

Lemme 26.6.— *Soient M une matrice carrée d'ordre n ; on a*

$$\|e^M - I_n - M\|_1 \leq \|M\|_1^2 e^{\|M\|_1}.$$

Démonstration. On a $e^M - I_n - M = \sum_{k=2}^{\infty} \frac{M^k}{k!}$; cette série est bien sûr toujours absolument convergente. On en déduit (cf. (23.4))

$$\begin{aligned}\|e^M - I_n - M\|_1 &\leq \sum_{k=2}^{\infty} \left\| \frac{M^k}{k!} \right\|_1 = \sum_{k=0}^{\infty} \left\| \frac{M^{k+2}}{(k+2)!} \right\|_1 \\ &\leq \|M\|_1^2 \sum_{k=0}^{\infty} \frac{\|M\|_1^k}{(k+2)!} \\ &\leq \|M\|_1^2 \sum_{k=0}^{\infty} \frac{\|M\|_1^k}{k!} = \|M\|_1^2 e^{\|M\|_1},\end{aligned}$$

d'où le lemme. ■

Démontrons maintenant la proposition. Il s'agit de montrer que le rapport

$$\frac{e^{tM} - e^{t_0M}}{t - t_0} - Me^{t_0M} = \frac{e^{tM} - e^{t_0M} - (t - t_0)Me^{t_0M}}{t - t_0}$$

tend vers 0 quand t tend vers t_0 . Puisque t_0M et $(t - t_0)M$ commutent, on a par la proposition 25.2

$$e^{tM} = e^{(t-t_0)M + t_0M} = e^{(t-t_0)M} e^{t_0M}.$$

Posons $t - t_0 = \alpha$; on a

$$\begin{aligned}e^{tM} - e^{t_0M} - (t - t_0)Me^{t_0M} &= e^{(t-t_0)M} e^{t_0M} - e^{t_0M} - (t - t_0)Me^{t_0M} \\ &= (e^{\alpha M} - I_n - \alpha M) e^{t_0M}.\end{aligned}$$

Le lemme (appliqué à la matrice αM) entraîne

$$\begin{aligned} \|e^{tM} - e^{t_0M} - (t - t_0)Me^{t_0M}\|_1 &\leq \|e^{\alpha M} - I_n - \alpha M\|_1 \|e^{t_0M}\|_1 \\ &\leq \alpha^2 \|M\|_1^2 e^{|\alpha| \|M\|_1} \|e^{t_0M}\|_1 \end{aligned}$$

d'où

$$\left\| \frac{e^{tM} - e^{t_0M} - (t - t_0)Me^{t_0M}}{t - t_0} \right\|_1 \leq \alpha \|M\|_1^2 e^{|\alpha| \|M\|_1} \|e^{t_0M}\|_1 .$$

Or la fonction (réelle) de α au membre de droite tend vers 0 quand α tend vers 0. Cela démontre la proposition. ■

27.

EXERCICES

(27.1) Calculer l'exponentielle des matrices suivantes

a) $\begin{pmatrix} 1 & 1 \\ -4 & 3 \end{pmatrix} .$

b) $\begin{pmatrix} 5 & -1 & -3 \\ 1 & 1 & -1 \\ 1 & 0 & 0 \end{pmatrix} .$

c) $\begin{pmatrix} -1 & -1 \\ 1 & -3 \end{pmatrix} .$

d) $\begin{pmatrix} 3 & 0 & 1 \\ 2 & 1 & 1 \\ -1 & 1 & 1 \end{pmatrix} .$

e) $\begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$; généraliser à une matrice carrée d'ordre n du même type.

(27.2) Si M est une matrice carrée, on note $\text{Tr } M$ (la *trace* de M) la somme de ses coefficients diagonaux. Montrer que

$$\det e^M = e^{\text{Tr } M} .$$

(27.3) Si M est une matrice carrée, montrer que la transposée de e^M est l'exponentielle de tM .

(27.4) Si M est une matrice carrée *réelle* d'ordre n , on note $\cos M$ la partie réelle de e^{iM} et $\sin M$ sa partie imaginaire.

a) Montrer que $\cos M$ et $\sin M$ commutent, et que $(\cos M)^2 + (\sin M)^2 = I_n$.

b) Soit θ un réel; calculer $\cos \begin{pmatrix} \theta & 1 \\ 0 & \theta \end{pmatrix}$ et $\sin \begin{pmatrix} \theta & 1 \\ 0 & \theta \end{pmatrix}$.

(27.5) Soit $M = (a_{ij})_{1 \leq i, j \leq n}$ une matrice complexe carrée d'ordre n telle que pour tous $1 \leq i, j \leq n$ on ait $|a_{ij}| < 1/n$.

a) Montrer que la suite de matrices $(N_m)_{m \in \mathbf{N}}$ définie par $N_m = I + M + M^2 + \dots + M^m$ converge.

b) Montrer que $I_n - M$ est inversible d'inverse la matrice $\lim(N_m)$.

IV. SYSTÈMES ET ÉQUATIONS DIFFÉRENTIELS LINÉAIRES HOMOGÈNES À COEFFICIENTS CONSTANTS

Soit a un nombre complexe ; résoudre l'équation différentielle $x' = ax$, c'est trouver une fonction dérivable $f : \mathbf{R} \rightarrow \mathbf{C}$, qui vérifie, pour tout réel t , la relation $f'(t) = af(t)$. Il est clair que pour toute constante C , la fonction $t \mapsto Ce^{at}$ est solution de cette équation ; inversement, si f est solution, on calcule la dérivée de la fonction $g : t \mapsto f(t)e^{-at}$. On trouve

$$g'(t) = f'(t)e^{-at} - af(t)e^{-at} = 0 .$$

La fonction g est donc constante, et f est du type $f(t) = Ce^{at}$. On a donc trouvé toutes les solutions de l'équation différentielle. Dans ce chapitre, nous généralisons cette méthode à la résolution de systèmes linéaires d'équations différentielles. Il n'est pas étonnant que les exponentielles de matrices y jouent un rôle important.

28. Systèmes différentiels linéaires homogènes à coefficients constants

Définition 28.1.— *On appelle système différentiel linéaire homogène à coefficients constants un système de la forme*

$$(S) \quad \begin{cases} x'_1(t) = a_{11}x_1(t) + a_{12}x_2(t) + \cdots + a_{1n}x_n(t) \\ x'_2(t) = a_{21}x_1(t) + a_{22}x_2(t) + \cdots + a_{2n}x_n(t) \\ \quad \cdot \\ \quad \cdot \\ \quad \cdot \\ x'_n(t) = a_{n1}x_1(t) + a_{n2}x_2(t) + \cdots + a_{nn}x_n(t) , \end{cases}$$

où les a_{ij} sont des constantes complexes, et où les fonctions x_i sont des fonctions dérivables inconnues à valeurs complexes.

On utilise la notation abrégée

$$(S) \quad X'(t) = MX(t) ,$$

où $M = (a_{ij})$ est une matrice $n \times n$ complexe constante.

Les exponentielles de matrices permettent de résoudre, en théorie du moins, le système (S). On a en effet :

Théorème 28.2.— *Les solutions complexes de (S) sont les fonctions de la forme $t \mapsto e^{tM}X_0$, où $X_0 \in \mathbf{C}^n$; l'ensemble des solutions est un espace vectoriel complexe de dimension n .*

Démonstration. Montrons d'abord que toute fonction du type $X(t) = e^{tM}X_0$ est solution de (S₀) ; il s'agit de calculer la dérivée de X en tout point. On utilise les propositions 26.4 et 26.5, qui donnent

$$X'(t) = Me^{tM}X_0 = MX(t) .$$

Inversement, montrons que toute solution X de (S) est de cette forme. On procède comme dans l'introduction, en dérivant la fonction $Y : t \mapsto e^{-tM}X(t)$; on obtient (prop. 26.4)

$$Y'(t) = -e^{-tM}MX(t) + e^{-tM}X'(t)$$

qui est nul. La fonction Y est donc constante, égale à sa valeur en 0, soit $X(0)$, de sorte que $X(t) = e^{tM}X(0)$ pour tout t .

Il reste à montrer que l'ensemble des fonctions de ce type est un espace vectoriel de dimension n ; cela résulte du fait que l'application qui au vecteur X_0 de \mathbf{C}^n associe la fonction $t \mapsto e^{tM}X_0$ est linéaire injective (son application réciproque est donnée par $X \mapsto X(0)$). Ceci démontre le théorème. ■

Corollaire 28.3.— Soient X_0 un vecteur de \mathbf{C}^n et t_0 un réel. La fonction $X(t) = e^{(t-t_0)M}X_0$ est l'unique solution de (S) qui vérifie $X(t_0) = X_0$.

Démonstration. La fonction X peut aussi s'écrire $X(t) = e^{tM}(e^{-t_0M}X_0)$; elle est donc solution de (S) par le théorème, et elle vérifie bien sûr $X(t_0) = X_0$. Réciproquement, si Y est une solution vérifiant $Y(t_0) = X_0$, la fonction $t \mapsto Y(t+t_0)$ est encore solution et prend la valeur X_0 en 0 ; par le théorème, elle est égale à $t \mapsto e^{tM}X_0$, d'où le corollaire. ■

Cela signifie qu'une solution dépend de façon linéaire de sa valeur à un instant donné.

Passons maintenant à une description un peu plus explicite des solutions. Commençons par le cas simple où la matrice M est diagonalisable.

Théorème 28.4.— Supposons qu'il existe une base (X_1, \dots, X_n) formée de vecteurs propres de M correspondant aux valeurs propres $\lambda_1, \dots, \lambda_n$ (ce qui signifie exactement que la matrice M est diagonalisable). Les solutions de (S) sont les fonctions de la forme

$$t \longmapsto c_1 e^{\lambda_1 t} X_1 + c_2 e^{\lambda_2 t} X_2 + \dots + c_n e^{\lambda_n t} X_n ,$$

où c_1, c_2, \dots, c_n sont des constantes arbitraires.

Démonstration. Toute « position initiale » X_0 se décompose en $X_0 = \sum_{j=1}^n c_j X_j$, où si P est la matrice de passage de la base canonique à la base (X_1, \dots, X_n) , on a

$$X_0 = P \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} .$$

On a d'autre part $D = P^{-1}MP$, où D est la matrice diagonale de termes diagonaux $(\lambda_1, \dots, \lambda_n)$, d'où (prop. 25.4)

$$e^{tM}X_0 = Pe^{tD}P^{-1}X_0 = Pe^{tD} \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} = P \begin{pmatrix} c_1 e^{t\lambda_1} \\ \vdots \\ c_n e^{t\lambda_n} \end{pmatrix} = \sum_{j=1}^n c_j e^{t\lambda_j} X_j .$$

On peut d'ailleurs remarquer que dans une base où la matrice est diagonale, le système différentiel se réduit aux n équations différentielles scalaires $x'_j = \lambda_j x_j$, avec $j = 1, \dots, n$. ■

Dans ce cas, les solutions sont donc combinaisons linéaires d'exponentielles (réelles ou complexes).

Exemple 28.5. Résolvons le système

$$\begin{cases} x'_1 = & 2x_2 - 2x_3 \\ x'_2 = -2x_1 & + x_3 \\ x'_3 = 2x_1 - x_2 & . \end{cases}$$

La matrice M est $\begin{pmatrix} 0 & 2 & -2 \\ -2 & 0 & 1 \\ 2 & -1 & 0 \end{pmatrix}$ et son polynôme caractéristique est $T^3 + 9T = T(T + 3i)(T - 3i)$.

Une base de l'espace propre pour la valeur propre 0 est $X_1 = (1, 2, 2)$. Pour la valeur propre $3i$, c'est $X_2 = (4, -1 + 3i, -1 - 3i)$; pour la valeur propre $-3i$, c'est le conjugué $X_3 = (4, -1 - 3i, -1 + 3i)$. Les solutions complexes du système sont donc

$$\begin{cases} x_1 = c_1 + 4c_2 e^{3it} + 4c_3 e^{-3it} \\ x_2 = 2c_1 + (-1 + 3i)c_2 e^{3it} + (-1 - 3i)c_3 e^{-3it} \\ x_3 = 2c_1 + (-1 - 3i)c_2 e^{3it} + (-1 + 3i)c_3 e^{-3it} , \end{cases}$$

c'est-à-dire qu'une base de l'espace vectoriel des solutions est composée de

$$X_1(t) = \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix} , \quad X_2(t) = \begin{pmatrix} 4 \\ -1 + 3i \\ -1 - 3i \end{pmatrix} e^{3it} , \quad X_3(t) = \begin{pmatrix} 4 \\ -1 - 3i \\ -1 + 3i \end{pmatrix} e^{-3it} .$$

Pour trouver les solutions réelles, on constate que les solutions X_1 , $(X_2 + X_3)/2 = \operatorname{Re}(X_2)$ et $(X_2 - X_3)/2i = \operatorname{Im}(X_2)$ sont réelles, linéairement indépendantes. Toute solution réelle est donc combinaison linéaire à coefficients réels de ces solutions.

Passons maintenant au cas général. Si λ_j est une valeur propre de M , on note comme d'habitude n_j sa multiplicité comme racine du polynôme caractéristique et m_j sa multiplicité comme racine du polynôme minimal; on a $0 < m_j \leq n_j$ (cf. (20.4)).

Théorème 28.6.— Soient $\lambda_1, \dots, \lambda_p$ les valeurs propres distinctes de M , et m_1, \dots, m_p leurs multiplicités comme racines du polynôme minimal. Les solutions de (S) sont de la forme

$$t \mapsto e^{\lambda_1 t} P_1(t) + e^{\lambda_2 t} P_2(t) + \dots + e^{\lambda_p t} P_p(t) ,$$

où $P_j(t)$ est une fonction polynomiale de degré $< m_j$ à valeurs dans l'espace caractéristique associé à la valeur propre λ_j .

Pour que la matrice M soit diagonalisable, il faut et il suffit que tous les m_j soient égaux à 1. Les fonctions polynomiales P_j du théorème sont alors constantes, et les espaces caractéristiques sont les espaces propres ; on retrouve le résultat du théorème 28.4.

Démonstration du théorème. On rappelle que l'espace caractéristique E'_{λ_j} associé à la valeur propre λ_j est le noyau de $(M - \lambda_j I_n)^{m_j}$ (def. 20.3), et que \mathbf{C}^n est somme directe des espaces $E'_{\lambda_1}, \dots, E'_{\lambda_p}$ (th. 20.6). D'autre part, on sait que toute solution X de (S) est de la forme $t \mapsto e^{tM} X_0$; décomposons X_0 en $X_1 + \dots + X_p$, avec $x_j \in E'_{\lambda_j}$. On a $e^{tM} X_j = e^{\lambda_j t} e^{(M - \lambda_j I_n)t} X_j$. Puisque $(M - \lambda_j I)^{m_j}$ est nul sur E'_{λ_j} , on en déduit

$$e^{tM} X_j = e^{\lambda_j t} \sum_{k=0}^{m_j-1} \frac{((M - \lambda_j I_n)t)^k}{k!} X_j ,$$

de sorte que $X(t) = \sum_{j=1}^p e^{\lambda_j t} P_j(t)$ avec

$$P_j(t) = \sum_{k=0}^{m_j-1} \frac{t^k}{k!} (M - \lambda_j I_n)^k X_j .$$

Ceci termine la démonstration du théorème. ■

On prendra garde que toutes les fonctions du type ci-dessus ne sont pas solution du système (ce n'est le cas que si M est diagonalisable).

Exemples 28.7. 1) Considérons le système

$$\begin{cases} x'_1 = 2x_1 - x_2 + 2x_3 \\ x'_2 = 10x_1 - 5x_2 + 7x_3 \\ x'_3 = 4x_1 - 2x_2 + 2x_3 \end{cases} .$$

La matrice M est $\begin{pmatrix} 2 & -1 & 2 \\ 10 & -5 & 7 \\ 4 & -2 & 2 \end{pmatrix}$ et son polynôme caractéristique est

$T^2(T + 1)$. A la valeur propre simple $\lambda = -1$ correspond le vecteur propre $X_1 = (1, -1, -2)$ et la solution $X_1(t) = e^{-t} X_1$.

L'espace propre correspondant à la valeur propre double 0 est de dimension 1 : la matrice M n'est donc pas diagonalisable. On cherche les autres solutions sous la forme

$$X(t) = (\alpha_1 t + \beta_1, \alpha_2 t + \beta_2, \alpha_3 t + \beta_3) .$$

En procédant par identification, on obtient les systèmes

$$\begin{cases} 0 = 2\alpha_1 - \alpha_2 + 2\alpha_3 \\ 0 = 10\alpha_1 - 5\alpha_2 + 7\alpha_3 \\ 0 = 4\alpha_1 - 2\alpha_2 + 2\alpha_3 \end{cases} \quad \begin{cases} \alpha_1 = 2\beta_1 - \beta_2 + 2\beta_3 \\ \alpha_2 = 10\beta_1 - 5\beta_2 + 7\beta_3 \\ \alpha_3 = 4\beta_1 - 2\beta_2 + 2\beta_3 \end{cases} ,$$

dont la résolution ne présente pas de difficulté.

On peut aussi procéder la façon suivante : on vérifie que le vecteur $X_2 = (1, 2, 0)$ engendre l'espace propre pour la valeur propre 0 (c'est-à-dire le noyau de M) ; cela revient à résoudre le premier des deux systèmes ci-dessus. La fonction constante $X_2(t) = X_2$ est donc solution. D'autre part, l'espace caractéristique E'_0 est le noyau de M^2 , qui vaut $\begin{pmatrix} 2 & -1 & 1 \\ -2 & 1 & -1 \\ -4 & 2 & -2 \end{pmatrix}$. Il contient le vecteur X_2 , ainsi que le vecteur $X_3 = (0, 1, 1)$. La démonstration du théorème montre que la fonction $X_3(t) = e^{tM}X_3 = X_3 + tMX_3 = (t, 2t + 1, 1)$ est solution du système. La solution générale du système est alors

$$\begin{cases} x_1 = c_1 e^{-t} + c_2 t + c_3 \\ x_2 = -c_1 e^{-t} + c_2(2t + 1) + 2c_3 \\ x_3 = -2c_1 e^{-t} + c_2 \end{cases} .$$

On remarquera que la seconde méthode est un peu plus élégante et moins calculatoire. C'est particulièrement frappant dans l'exemple suivant.

2) Considérons le système

$$\begin{cases} x'_1 = -4x_1 + x_2 + x_3 \\ x'_2 = x_1 - x_2 - 2x_3 \\ x'_3 = -2x_1 + x_2 - x_3 \end{cases} .$$

La matrice M est $\begin{pmatrix} -4 & 1 & 1 \\ 1 & -1 & -2 \\ -2 & 1 & -1 \end{pmatrix}$, et son polynôme caractéristique est $(T + 2)^3$. Les solutions sont données par $t \mapsto e^{tM}X_0$. Comme $(M + 2I_3)^3 = 0$ (théorème de Cayley-Hamilton), on écrit :

$$e^{tM} = e^{-2I_3 t} e^{(M+2I_3)t} = e^{-2t} (I_3 + (M + 2I_3)t + \frac{1}{2}(M + 2I_3)^2 t^2) .$$

Il ne reste qu'à calculer

$$(M + 2I_3) = \begin{pmatrix} -2 & 1 & 1 \\ 1 & 1 & -2 \\ -2 & 1 & 1 \end{pmatrix} \quad (M + 2I_3)^2 = \begin{pmatrix} 3 & 0 & -3 \\ 3 & 0 & -3 \\ 3 & 0 & -3 \end{pmatrix},$$

d'où

$$\begin{cases} x_1(t) = e^{-2t} \left[\left(1 - 2t + \frac{3t^2}{2}\right) c_1 + tc_2 + \left(t - \frac{3t^2}{2}\right) c_3 \right] \\ x_2(t) = e^{-2t} \left[\left(t + \frac{3t^2}{2}\right) c_1 + (1+t)c_2 - \left(2t + \frac{3t^2}{2}\right) c_3 \right] \\ x_3(t) = e^{-2t} \left[\left(-2t + \frac{3t^2}{2}\right) c_1 + tc_2 + \left(1+t - \frac{3t^2}{2}\right) c_3 \right]. \end{cases}$$

29. Conseils pratiques

Pour résoudre le système $X' = MX$, on pourra procéder de la façon suivante :

1) Tout d'abord, on détermine les valeurs propres et vecteurs propres de M . Si ces vecteurs propres forment une base, c'est-à-dire si la matrice M est diagonalisable, le théorème 28.4 donne immédiatement la solution.

2) Si la matrice M n'est pas diagonalisable, mais qu'elle n'admet qu'une valeur propre λ , on utilisera la méthode de l'exemple 28.7.2) pour calculer e^{tM} en écrivant

$$e^{tM} = e^{\lambda t} e^{t(M-\lambda I)} = e^{\lambda t} \sum_{k=0}^{n-1} (M-\lambda I)^k \frac{t^k}{k!}.$$

3) Si la matrice M n'est pas diagonalisable mais admet plusieurs valeurs propres $\lambda_1, \dots, \lambda_p$ de multiplicité respectives n_1, \dots, n_p , on pourra procéder par identification en cherchant, pour chaque valeur propre λ_j un polynôme $P_j(t)$ de degré $< n_j$ tel que $e^{\lambda_j t} P_j(t)$ soit solution. On pourra aussi chercher une base (X_j^1, X_j^2, \dots) de l'espace caractéristique E'_{λ_j} satisfaisant $X_j^\ell \in \text{Ker}(M - \lambda_j I)^\ell$, et calculer $e^{tM} X_j^\ell$ comme en 2) : lorsque par exemple $n_j = 2$, cela revient à trouver un vecteur propre X_j et un vecteur X'_j de $\text{Ker}(M - \lambda_j I)^2$ non colinéaire à X_j . Les solutions correspondantes seront (cf. exemple 28.7.1)

$$e^{\lambda_j t} X_j \quad \text{et} \quad e^{\lambda_j t} [X'_j + t(M - \lambda_j I)X'_j].$$

4) Lorsque la matrice M est à coefficients réels, ses valeurs propres se répartissent en valeurs propres réelles $\lambda_1, \dots, \lambda_p$ et en paires de valeurs propres complexes conjuguées $\mu_1, \bar{\mu}_1, \dots, \mu_q, \bar{\mu}_q$. On rappelle que les espaces propres (ou caractéristiques) correspondant à des valeurs propres conjuguées sont aussi conjugués ; il est donc inutile de faire le calcul deux fois. La base de solutions (complexes) obtenue par les méthodes ci-dessus sera alors du type $(X_1(t), \dots, X_p(t), Y_1(t), \bar{Y}_1(t), \dots, Y_q(t), \bar{Y}_q(t))$, avec X_1, \dots, X_p réelles. On prendra comme base de solutions réelles

$$(X_1(t), \dots, X_p(t), \text{Re } Y_1(t), \text{Im } Y_1(t), \dots, \text{Re } Y_q(t), \text{Im } Y_q(t)).$$

30. Allure des trajectoires

Considérons le système différentiel homogène $X'(t) = MX(t)$, où M est une matrice réelle carrée d'ordre n . On peut considérer une solution $X : \mathbf{R} \rightarrow \mathbf{R}^n$ comme décrivant la position d'un mobile ponctuel à l'instant t ; son image est alors la *trajectoire* de ce point. Nous allons étudier ces trajectoires de façon qualitative, dans le cas $n = 2$ (pour pouvoir faire des dessins) et $\det M \neq 0$ (pour nous concentrer sur les cas intéressants).

Premier cas : $M = \lambda I_2$ ($\lambda \neq 0$). On a alors $X(t) = e^{\lambda t} X(0)$, de sorte que notre mobile décrit entièrement la demi-droite vectorielle passant par $X(0)$. Le sens de déplacement dépend du signe de λ . Pour $\lambda > 0$ par exemple, l'allure des trajectoires est la suivante :

Deuxième cas : M a deux valeurs propres réelles distinctes λ et μ (non nulles). Elle est alors diagonalisable dans une base \mathcal{B} , dans laquelle la loi du mouvement $X(t)$ s'écrit

$$x_1(t) = e^{\lambda t} x_1(0) \qquad x_2(t) = e^{\mu t} x_2(0) .$$

On remarque que la quantité $|x_1(t)|^\mu |x_2(t)|^{-\lambda}$ reste constante au cours du temps; cela signifie que le point se déplace sur une courbe d'équation $|x_1|^\mu |x_2|^{-\lambda} = C$ dans la base \mathcal{B} . Il y a principalement deux cas de figure (exercice) :

$$\mu > \lambda > 0$$

$$\lambda > 0 > \mu$$

Troisième cas : M a deux valeurs propres distinctes complexes conjuguées $\lambda = \alpha - i\beta$ et $\bar{\lambda}$, de vecteurs propres (complexes) associés X_λ et \bar{X}_λ . Elle est alors diagonalisable, et une base de solutions réelles est donnée par $(\operatorname{Re}(e^{(\alpha-i\beta)t}X_\lambda), \operatorname{Im}(e^{(\alpha-i\beta)t}X_\lambda))$. Une solution réelle générale est donc de la forme suivante, avec a et b réels,

$$\begin{aligned} & a \operatorname{Re}(e^{(\alpha-i\beta)t}X_\lambda) + b \operatorname{Im}(e^{(\alpha-i\beta)t}X_\lambda) \\ &= \frac{a}{2}(e^{(\alpha-i\beta)t}X_\lambda + e^{(\alpha+i\beta)t}\bar{X}_\lambda) + \frac{b}{2i}(e^{(\alpha-i\beta)t}X_\lambda - e^{(\alpha+i\beta)t}\bar{X}_\lambda) \\ &= \frac{1}{2}[(a-ib)e^{(\alpha-i\beta)t}X_\lambda + (a+ib)e^{(\alpha+i\beta)t}\bar{X}_\lambda] \\ &= \operatorname{Re}(ce^{(\alpha-i\beta)t}X_\lambda) , \end{aligned}$$

où $c = a - ib$ est un réel quelconque. Si on écrit $c = |c|e^{-i\beta\phi}$, on obtient, dans la base $(\operatorname{Re}(X_\lambda), \operatorname{Im}(X_\lambda))$,

$$X(t) = \operatorname{Re}(|c|e^{(\alpha-i\beta)t+i\beta\phi}(\operatorname{Re}(X_\lambda) + i \operatorname{Im}(X_\lambda))) = |c| e^{\alpha t} \begin{pmatrix} \cos \beta(t + \phi) \\ \sin \beta(t + \phi) \end{pmatrix} .$$

Les deux cas de figure principaux sont les suivants (exercice) :

$$\alpha > 0, \beta > 0$$

$$\alpha = 0, \beta > 0$$

On remarquera que dans ce dernier cas, toutes les solutions sont *bornées* sur \mathbf{R} (le point « ne part pas à l'infini »). L'exercice 32.12 propose une généralisation de ce résultat.

Quatrième cas : M n'est pas diagonalisable. La matrice M est alors semblable à $\begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}$, où λ est un réel non nul (exercice), et $\exp(tM)$ est semblable à

$$\exp\left(\begin{pmatrix} t\lambda & 0 \\ 0 & t\lambda \end{pmatrix} + \begin{pmatrix} 0 & t \\ 0 & 0 \end{pmatrix}\right) = e^{\lambda t} \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix},$$

On en déduit que dans une certaine base de \mathbf{R}^2 , on a

$$X(t) = e^{\lambda t} \begin{pmatrix} x_1(0) + tx_2(0) \\ x_2(0) \end{pmatrix}.$$

Pour $\lambda > 0$, les trajectoires ont l'allure suivante :

31. Equations différentielles linéaires homogènes à coefficients constants d'ordre supérieur

Définition 31.1.— On appelle *équation différentielle linéaire homogène scalaire d'ordre n à coefficients constants* une équation de la forme

$$(E) \quad x^{(n)}(t) + a_{n-1}x^{(n-1)}(t) + \cdots + a_1x'(t) + a_0x(t) = 0,$$

où les a_j sont des constantes complexes, et où la fonction x est une fonction complexe inconnue n fois dérivable.

L'introduction des inconnues auxiliaires $x_0 = x, x_1 = x', \dots, x_{n-1} = x^{(n-1)}$ nous ramène au système $X'(t) = MX(t)$, avec

$$X = \begin{pmatrix} x_0 \\ x_1 \\ \cdot \\ \cdot \\ x_{n-1} \end{pmatrix} \quad M = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ & & & & \cdot \\ & & & & \cdot \\ 0 & & & 0 & 1 \\ -a_0 & -a_1 & & \cdots & -a_{n-1} \end{pmatrix}.$$

On peut donc appliquer les résultats précédents sur les systèmes. En particulier, le théorème 28.2 entraîne que l'ensemble des solutions est un espace vectoriel de dimension n . Le polynôme caractéristique de la matrice M est le polynôme

$$P_M(T) = T^n + a_{n-1}T^{n-1} + \cdots + a_1T + a_0 ;$$

c'est aussi son polynôme minimal (*cf.* ex. 20.11.2); on l'appelle aussi le polynôme caractéristique de l'équation différentielle (E), et l'équation $P_M(\lambda) = 0$ son *équation caractéristique*.

Théorème 31.2.— Soient $\lambda_1, \dots, \lambda_p$ les racines distinctes de l'équation caractéristique, et n_1, \dots, n_p leurs multiplicités. Les solutions complexes de l'équation (E) sont toutes les fonctions du type

$$x(t) = Q_1(t)e^{\lambda_1 t} + \cdots + Q_p(t)e^{\lambda_p t},$$

où Q_j est un polynôme de degré $< n_j$, à coefficients complexes.

Démonstration. On ramène comme ci-dessus l'équation au système $X' = MX$; le théorème 28.6 nous dit que toute solution X de ce système est du type

$$X(t) = e^{\lambda_1 t}P_1(t) + e^{\lambda_2 t}P_2(t) + \cdots + e^{\lambda_p t}P_p(t),$$

où $P_j(t)$ est un polynôme de degré $< n_j$. En prenant la première composante de cette égalité, on voit que toute solution x de (E) est de la forme

$$x(t) = e^{\lambda_1 t}P_{11}(t) + e^{\lambda_2 t}P_{21}(t) + \cdots + e^{\lambda_p t}P_{p1}(t),$$

où $P_{j1}(t)$ est un polynôme de degré $< n_j$. En d'autres termes, l'espace des solutions est contenu dans l'espace vectoriel engendré par les n fonctions $t \mapsto t^k e^{\lambda_j t}$, pour $0 \leq k < n_j$. Comme l'espace des solutions est de dimension n ; ces deux espaces sont égaux. ■

Exemples 31.3. 1) Les solutions de l'équation différentielle $x''' + 3x'' + 3x' + x = 0$, dont l'équation caractéristique est $(\lambda + 1)^3 = 0$, sont les fonctions de la forme $x(t) = (at^2 + bt + c)e^{-t}$.

2) Les solutions de l'équation différentielle $x^{(4)} + 2x'' + x = 0$, d'équation caractéristique $(\lambda^2 + 1)^2 = 0$, sont les fonctions de la forme $x(t) = (at + b)e^{it} + (a't + b')e^{-it}$. Les solutions réelles sont $x(t) = (at + b)\cos t + (a't + b')\sin t$.

32.

EXERCICES

(32.1) Résoudre le système différentiel

$$\begin{cases} x' = x + y \\ y' = -4x - 3y \end{cases}$$

(32.2) Résoudre le système différentiel

$$\begin{cases} 2u' = 5u - v - 3w \\ 2v' = u + v - w \\ w' = u \end{cases}$$

(32.3) Résoudre le système différentiel

$$\begin{cases} x' = -x - y \\ y' = x - 3y \end{cases}$$

(32.4) Résoudre le système différentiel

$$\begin{cases} x' = 7x + 6y + 3z \\ y' = -x - y - z \\ z' = -4x - 2y \end{cases}$$

(32.5) Résoudre le système différentiel

$$\begin{cases} x' = 8x - y - 5z \\ y' = -2x + 3y + z \\ z' = 4x - y - z \end{cases}$$

(32.6) Résoudre le système différentiel

$$\begin{cases} x' = 4x - 6y + z \\ y' = 2x - 4y + z \\ z' = 4x - 8y + 2z \end{cases}$$

(32.7) Résoudre le système différentiel

$$\begin{cases} x' = y - 12z \\ y' = -x + 2y - 20z \\ z' = x - 5z \end{cases}$$

(32.8) Résoudre l'équation différentielle $x''' + 3x'' - 4x = 0$.

(32.9) Trouver toutes les solutions réelles de l'équation différentielle $x^{(n)} = x$ (on distinguera le cas n pair du cas n impair).

(32.10) Trouver une équation différentielle linéaire à coefficients constants réels dont les fonctions $t^2 e^{-3t} \cos(2t)$ et $t^3 e^{-t}$ soient solutions.

(32.11) Résoudre le système différentiel

$$\begin{cases} x'' - x + \frac{15}{4}y' = 0 \\ x' + y'' + y = 0 \end{cases}$$

a) en éliminant x et ses dérivées entre les deux équations pour obtenir une équation différentielle d'ordre 4 en y ;

b) en se ramenant à un système différentiel dans \mathbf{R}^4 ;

c) en cherchant des solutions sous la forme $\begin{pmatrix} x \\ y \end{pmatrix} = e^{rt} \begin{pmatrix} a \\ b \end{pmatrix}$, où a et b sont constantes.

(32.12) a) Soit P un polynôme à coefficients complexes. A quelles conditions nécessaires et suffisantes sur le complexe λ et le polynôme P la fonction $t \mapsto e^{\lambda t} P(t)$ est-elle bornée sur \mathbf{R} ?

b) Soit M une matrice carrée complexe; montrer que pour que toutes les solutions du système différentiel $X' = MX$ soient bornées sur tout \mathbf{R} , il faut et il suffit que M soit diagonalisable sur \mathbf{C} et n'admette que des valeurs propres de partie réelle nulle.

SOLUTIONS

(32.1) Une base de solutions est $e^{-t} \begin{pmatrix} 1 \\ -2 \end{pmatrix}$ et $e^{-t} \begin{pmatrix} t \\ -2t+1 \end{pmatrix}$.

(32.2) Les valeurs propres sont 1 (triple).

(32.3) Une base de solutions est $e^{-2t} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ et $e^{-2t} \begin{pmatrix} 2t+1 \\ 2t-1 \end{pmatrix}$.

(32.4) Une base de solutions est $e^{4t} \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$, $e^t \begin{pmatrix} 0 \\ 1 \\ -2 \end{pmatrix}$ et $e^t \begin{pmatrix} 1 \\ t \\ -2t-2 \end{pmatrix}$.

(32.5) Une base de solutions est $e^{2t} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$, $e^{4t} \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$ et $e^{4t} \begin{pmatrix} 3t+1 \\ -3t+1 \\ 3t \end{pmatrix}$.

(32.7) La seule valeur propre est -1 .

On fixe un corps K , égal à \mathbf{R} ou \mathbf{C} . Tous les espaces vectoriels considérés seront des espaces vectoriels sur K de dimension finie.

33. Dualité, bases duales, bidualité

Soit E un espace vectoriel. On rappelle qu'une forme linéaire sur E est une application linéaire de E dans K . Par exemple, si E est l'espace vectoriel des polynômes de degré $\leq n$ et a un élément fixé de K , l'application $P \mapsto P(a)$ est une forme linéaire sur E .

Définition 33.1.— Soit E un espace vectoriel. On appelle espace dual (ou simplement dual) de E l'espace vectoriel des formes linéaires sur E ; on le note E^* .

Soit $\mathcal{B} = (e_1, \dots, e_n)$ une base de E . Pour tout $j = 1, \dots, n$, on définit une forme linéaire e_j^* sur E en posant $e_j^*(e_k) = \delta_{jk}$, où δ_{jk} est le symbole de Kronecker (mathématicien allemand, 1823–1891) qui vaut 1 si $j = k$, et 0 sinon.

Vérifions que la famille de formes linéaires $\mathcal{B}^* = (e_1^*, \dots, e_n^*)$ ainsi construite est une base de E^* :

(33.2) si $\sum_{j=1}^n \lambda_j e_j^*$ est nulle, on obtient en l'appliquant à e_k

$$0 = \left(\sum_{j=1}^n \lambda_j e_j^* \right) (e_k) = \sum_{j=1}^n \lambda_j e_j^*(e_k) = \sum_{j=1}^n \lambda_j \delta_{jk} = \lambda_k,$$

ce qui prouve que \mathcal{B}^* est libre ;

(33.3) soit f une forme linéaire sur E ; on a

$$f = \sum_{j=1}^n f(e_j) e_j^* ;$$

en effet, ces formes linéaires prennent la même valeur sur chacun des e_k , donc sont égales. Ceci prouve que \mathcal{B}^* est génératrice.

(33.4) On appelle la base \mathcal{B}^* la *base duale* de la base \mathcal{B} . Remarquons que \mathcal{B} et \mathcal{B}^* ont le même nombre d'éléments, de sorte que E et E^* ont même dimension.

Attention : la notation e^* n'a de sens que si e fait partie d'une base !

Considérons maintenant le *bidual* $(E^*)^*$, noté encore E^{**} ; il existe une application canonique $\psi : E \rightarrow E^{**}$ définie par $\psi(x)(f) = f(x)$, pour tout x dans E et f dans E^* .

On vérifie qu'elle est linéaire : si x et y sont dans E et λ dans K , on a, pour tout f dans E^* ,

$$\psi(\lambda x + y)(f) = f(\lambda x + y) = \lambda f(x) + f(y) = \lambda \psi(x)(f) + \psi(y)(f) = (\lambda \psi(x) + \psi(y))(f),$$

de sorte que $\psi(\lambda x + y) = \lambda \psi(x) + \psi(y)$.

Théorème 33.5.— *L'application linéaire canonique $\psi : E \rightarrow E^{**}$ est bijective.*

Comme E et E^{**} ont même dimension (cf. 33.4), il suffit de montrer que ψ est injective. Soit x un élément non nul de E ; il existe une base (e_1, \dots, e_n) de E telle que $e_1 = x$ (théorème de la base incomplète). Notons (e_1^*, \dots, e_n^*) la base duale; on a $\psi(x)(e_1^*) = e_1^*(x) = 1$, de sorte que $\psi(x)$ n'est pas nul. Cela montre que ψ est injective. ■

34. Orthogonalité, transposée d'une application linéaire

Définition 34.1.— *Soient E un espace vectoriel et U une partie de E ; on appelle orthogonal de U dans E^* , et l'on note U° , l'ensemble des formes linéaires $f \in E^*$ telles que $f(x) = 0$ pour tout x dans U .*

(34.2) On vérifie que U° est un sous-espace vectoriel de E^* : si f et g sont dans U° et λ dans K , on a, pour tout x dans U ,

$$(\lambda f + g)(x) = \lambda f(x) + g(x) = 0,$$

de sorte que $\lambda f + g$ est dans U° . Si V est une partie de E^* , son orthogonal V° est, selon la définition, dans E^{**} . Soit x un élément de E ; pour que $\psi(x)$ (défini dans le théorème 33.5) soit dans V° , il faut et il suffit que, pour tout f dans V , on ait $0 = \psi(x)(f) = f(x)$, par définition de ψ . Si l'on identifie E^{**} à E par ψ , on a donc $V^\circ = \{x \in E \mid \forall f \in V \ f(x) = 0\}$.

Théorème 34.3.— *Soient E un espace vectoriel et F un sous-espace vectoriel de E . On a*

$$\dim(F) + \dim(F^\circ) = \dim(E).$$

Soient p la dimension de F et (e_1, \dots, e_p) une base de F ; on peut la compléter en une base (e_1, \dots, e_n) de E . Soit $\{e_1^*, \dots, e_n^*\}$ la base duale; pour qu'une forme linéaire f soit dans F^* , il faut et il suffit qu'elle s'annule sur e_1, \dots, e_p . Ecrivons $f = \sum_{j=1}^n \lambda_j e_j^*$; comme $\lambda_j = f(e_j)$ (cf. 33.3), cette condition s'écrit $\lambda_1 = \dots = \lambda_p = 0$. L'espace vectoriel F° est donc engendré par $\{e_{p+1}^*, \dots, e_n^*\}$: il est de dimension $n - p = \dim(E) - \dim(F)$. ■

Définition 34.4.— *Soient E et F des espaces vectoriels et $u : E \rightarrow F$ une application linéaire. On appelle transposée de u l'application ${}^t u : F^* \rightarrow E^*$ définie par*

$${}^t u(g) = g \circ u,$$

pour tout g dans F^* .

On vérifie que ${}^t u$ est une application *linéaire* de F^* dans E^* : si f et g sont dans F^* et λ dans K , on a pour tout x dans E

$$\begin{aligned} {}^t u(\lambda f + g)(x) &= (\lambda f + g)(u(x)) = \lambda f(u(x)) + g(u(x)) \\ &= \lambda {}^t u(f)(x) + {}^t u(g)(x) = (\lambda {}^t u(f) + {}^t u(g))(x), \end{aligned}$$

de sorte que ${}^t u(\lambda f + g) = \lambda {}^t u(f) + {}^t u(g)$. Considérons l'application ${}^t({}^t u)$; elle va de E^{**} dans F^{**} (on « renverse les flèches » quand on prend la transposée). On vérifie que, si l'on identifie E à E^{**} et F à F^{**} par les isomorphismes $\psi_E : E \rightarrow E^{**}$ et $\psi_F : F \rightarrow F^{**}$ du théorème 33.5, on a

$$(34.5) \quad {}^t({}^t u) = u.$$

En effet, on a pour tout x dans E et tout g dans F^* ,

$${}^t({}^t u)(\psi_E(x))(g) = \psi_E(x)({}^t u(g)) = \psi_E(x)(g \circ u) = (g \circ u)(x) = g(u(x)) = \psi_F(u(x))(g),$$

de sorte que ${}^t({}^t u)(\psi_E(x)) = \psi_F(u(x))$.

Théorème 34.6.— Soient E et F des espaces vectoriels et $u : E \rightarrow F$ une application linéaire. On a

$$\text{Ker}({}^t u) = (\text{Im}(u))^\circ,$$

et u et ${}^t u$ ont même rang.

Le noyau de ${}^t u$ est l'ensemble des formes linéaires g sur F telles que ${}^t u(g) = 0$, c'est-à-dire telles que $({}^t u(g))(x) = g(u(x))$ s'annule pour tout x dans E . Cela revient exactement à dire que g est orthogonal à tous les $u(x)$, c'est-à-dire à l'image de u . Ceci montre le premier point. On a d'autre part

$$\dim(\text{Ker}({}^t u)) + \dim(\text{Im}({}^t u)) = \dim(F^*)$$

et

$$\dim(\text{Im}(u)) + \dim(\text{Im}(u))^\circ = \dim(F)$$

par 34.3. Comme on vient de voir l'égalité $\dim(\text{Ker}({}^t u)) = \dim(\text{Im}(u))^\circ$, on en déduit $\dim(\text{Im}({}^t u)) = \dim(\text{Im}(u))$. ■

Corollaire 34.7.— On a les équivalences

$$\begin{aligned} {}^t u \text{ injective} &\iff u \text{ surjective}; \\ {}^t u \text{ surjective} &\iff u \text{ injective}. \end{aligned}$$

La première équivalence résulte du théorème. La seconde résulte de la première appliquée à ${}^t u$, en utilisant (34.5). ■

Proposition 34.8.— Soient E et F des espaces vectoriels, soit \mathcal{B} une base de E , soit \mathcal{B}' une base de F , et soit $u : E \rightarrow F$ une application linéaire. La matrice de ${}^t u$ dans les bases \mathcal{B}'^* et \mathcal{B}^* est la transposée (au sens habituel) de la matrice de u dans les bases \mathcal{B} et \mathcal{B}' .

Ecrivons $\mathcal{B} = (e_1, \dots, e_n)$ et $\mathcal{B}' = (e'_1, \dots, e'_m)$. Soit $M = (a_{jk})$ la matrice de u dans les bases \mathcal{B} et \mathcal{B}' ; elle est définie par

$$u(e_k) = \sum_{j=1}^m a_{jk} e'_j .$$

On a alors

$${}^t u(e'_j^*)(e_k) = e'_j^*(u(e_k)) = a_{jk} ,$$

de sorte que ${}^t u(e'_j^*) = \sum_{k=1}^n a_{jk} e_k^*$ par (33.3). La matrice (a_{jk}^*) de ${}^t u$ dans les bases \mathcal{B}'^* et \mathcal{B}^* , définie par ${}^t u(e'_j^*) = \sum_{k=1}^n a_{jk}^* e_k^*$, est donc ${}^t M$. ■

Corollaire 34.9.— Le rang d'une matrice est égal au rang de sa transposée.

Cela résulte de la proposition et du théorème 34.6. ■

35. Formes bilinéaires

Soit E un espace vectoriel.

Définition 35.1.— On appelle forme bilinéaire sur E toute application $B : E \times E \rightarrow K$ qui vérifie

- a) pour tout $y \in E$ fixé, l'application $B^y : x \mapsto B(x, y)$ est une forme linéaire sur E ;
- b) pour tout $x \in E$ fixé, l'application ${}^x B : y \mapsto B(x, y)$ est une forme linéaire sur E .

En d'autres termes, pour tous $x, y, z \in E$ et tout $\lambda \in K$, on a

$$\begin{aligned} B(\lambda x, y) &= B(x, \lambda y) = \lambda B(x, y) \\ B(x + y, z) &= B(x, z) + B(y, z) & B(x, y + z) &= B(x, y) + B(x, z) . \end{aligned}$$

On prendra garde qu'une application bilinéaire n'est en général pas linéaire!

Exemple 35.2. La relation $B((x_1, \dots, x_n), (y_1, \dots, y_n)) = x_1 y_1 + \dots + x_n y_n$ définit une forme bilinéaire sur K^n .

Si $\mathcal{B} = (e_1, \dots, e_n)$ est une base de E , et si $x = \sum_{j=1}^n x_j e_j$ et $y = \sum_{j=1}^n y_j e_j$ sont des éléments de E , on a

$$(35.3) \quad B(x, y) = B\left(\sum_{j=1}^n x_j e_j, \sum_{j=1}^n y_j e_j\right) = \sum_{1 \leq j, k \leq n} B(e_j, e_k) x_j y_k ,$$

de sorte que la forme B est déterminée de façon unique par la matrice $(B(e_j, e_k))_{1 \leq j, k \leq n}$, qui est appelée *matrice de la forme bilinéaire B dans la base \mathcal{B}* ; on la note $M_{B, \mathcal{B}}$, ou simplement $M_{\mathcal{B}}$. Si on associe au vecteur x la matrice colonne $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ et au vecteur y la matrice colonne Y analogue, on a

$$(35.4) \quad \boxed{B(x, y) = {}^t X M_{\mathcal{B}} Y .}$$

Soient maintenant \mathcal{B}' une autre base de E , et $P_{\mathcal{B} \rightarrow \mathcal{B}'}$ la matrice de passage de \mathcal{B} à \mathcal{B}' . Si X' (resp. Y') est la matrice colonne des composantes de x (resp. de y) dans la base \mathcal{B}' , on a $X = P_{\mathcal{B} \rightarrow \mathcal{B}'} X'$ (cf. (4.6)). On en déduit

$$B(x, y) = {}^t X M_{\mathcal{B}} Y = {}^t X' {}^t P_{\mathcal{B} \rightarrow \mathcal{B}'} M_{\mathcal{B}} P_{\mathcal{B} \rightarrow \mathcal{B}'} Y' ,$$

de sorte que

$$\boxed{M_{\mathcal{B}'} = {}^t P_{\mathcal{B} \rightarrow \mathcal{B}'} M_{\mathcal{B}} P_{\mathcal{B} \rightarrow \mathcal{B}'} .}$$

En particulier, le déterminant de la matrice $M_{\mathcal{B}}$ dépend de la base choisie.

(35.5) Soit B une forme bilinéaire sur E ; on définit des applications linéaires

$$\begin{array}{ll} \widehat{B} : E \longrightarrow E^* & \check{B} : E \longrightarrow E^* \\ y \longmapsto B^y & x \longmapsto {}^x B . \end{array}$$

La forme linéaire $\widehat{B}(y)$ est donc l'application $x \mapsto B(x, y)$, tandis que la forme linéaire $\check{B}(x)$ est l'application $y \mapsto B(x, y)$. On peut considérer la transposée ${}^t \check{B} : E^{**} \rightarrow E^*$ de l'application \check{B} ; si l'on identifie E^{**} à E par l'application linéaire ψ du théorème 33.5, on a ${}^t \check{B} = \widehat{B}$. En effet, pour tous x et y dans E , on a

$${}^t \check{B}(\psi(y))(x) = \psi(y)(\check{B}(x)) = \check{B}(x)(y) = B(x, y) = \widehat{B}(y)(x) .$$

Il en résulte en particulier que \widehat{B} et \check{B} ont même rang (th. 34.6).

Définition 35.6.— Soit B une forme bilinéaire. On appelle *rang de B* le rang commun des applications linéaires \widehat{B} et \check{B} associées. On dit que B est non dégénérée si ce rang est la dimension de E , c'est-à-dire si \widehat{B} et \check{B} sont bijectives.

La proposition suivante permet d'interpréter ces définitions en termes de matrices.

Proposition 35.7.— Soit B une forme bilinéaire sur un espace vectoriel E . Si \mathcal{B} est une base de E et \mathcal{B}^* sa base duale, la matrice de l'application linéaire \widehat{B} dans les bases \mathcal{B} et \mathcal{B}^* est la matrice $M_{\mathcal{B}}$ de B dans la base \mathcal{B} .

Ecrivons $\mathcal{B} = (e_1, \dots, e_n)$ et $\mathcal{B}^* = (e_1^*, \dots, e_n^*)$. Soit (b_{jk}) la matrice de \widehat{B} dans les bases \mathcal{B} et \mathcal{B}^* . Il s'agit de vérifier que $b_{jk} = B(e_j, e_k)$. On a

$$B(e_j, e_k) = \widehat{B}(e_k)(e_j) = \left(\sum_{l=1}^n b_{lk} e_l^* \right) (e_j) = \sum_{l=1}^n b_{lk} e_l^*(e_j) = \sum_{l=1}^n b_{lk} \delta_{lj} = b_{jk} ,$$

ce qui démontre la proposition. ■

En particulier, le rang de B est le rang de sa matrice dans une base quelconque. Pour que B soit non dégénérée, il faut et il suffit que sa matrice dans une base soit inversible (sa matrice dans une base quelconque est alors inversible) ou encore que, pour tout vecteur non nul y , il existe $x \in E$ tel que $B(x, y) \neq 0$.

36. Formes bilinéaires symétriques et formes quadratiques

Définition 36.1.— Une forme bilinéaire B sur E est dite *symétrique* si, pour tous éléments x et y de E , on a

$$B(y, x) = B(x, y) .$$

La forme bilinéaire de l'exemple 35.2 est symétrique. Si la forme bilinéaire B est symétrique, les applications linéaires \widehat{B} et \check{B} définies en (35.5) sont égales. On définit le noyau de B comme le noyau de \widehat{B} (ou celui de \check{B}). En d'autres termes,

$$\text{Ker } B = \{ x \in E \mid B(x, y) = 0 \text{ pour tout } y \text{ dans } E \}$$

Attention, le noyau de B n'est en général pas l'ensemble des vecteurs isotropes!

Proposition 36.2.— Pour qu'une forme bilinéaire soit symétrique, il faut et il suffit que sa matrice dans une base donnée soit symétrique. Sa matrice dans une base quelconque est alors symétrique.

Si une forme bilinéaire est symétrique, il est clair que sa matrice dans une base quelconque est symétrique. Inversement, si la matrice d'une forme bilinéaire B dans une base $\mathcal{B} = (e_1, \dots, e_n)$ est symétrique, c'est-à-dire si l'on a $B(e_j, e_k) = B(e_k, e_j)$ pour tous j et k , il résulte de la formule (35.3) que

$$B(x, y) = \sum_{1 \leq j, k \leq n} B(e_j, e_k) x_j y_k = \sum_{1 \leq j, k \leq n} B(e_k, e_j) y_k x_j = B(y, x) ,$$

ce qui prouve que B est symétrique. ■

Définition 36.3.— On appelle *forme quadratique* sur E toute application $Q : E \rightarrow K$ telle qu'il existe une forme bilinéaire B vérifiant $Q(x) = B(x, x)$ pour tout x dans E .

Soit $\mathcal{B} = (e_1, \dots, e_n)$ une base de E ; si $x = \sum_{j=1}^n x_j e_j$, la formule (35.3) montre que

$$Q(x) = \sum_{1 \leq j, k \leq n} B(e_j, e_k) x_j x_k ;$$

les formes quadratique sont donc les fonctions polynomiales *homogènes de degré 2* : elles vérifient, pour tout élément x de E et tout réel λ , l'égalité $Q(\lambda x) = \lambda^2 Q(x)$ (elle ne sont pas linéaires!).

Théorème 36.4.— *Etant donnée une forme quadratique Q , il existe une unique forme bilinéaire symétrique B vérifiant $B(x, x) = Q(x)$. Elle est donnée par la formule dite de polarisation*

$$(36.5) \quad B(x, y) = \frac{1}{2} \left(Q(x+y) - Q(x) - Q(y) \right) .$$

On l'appelle la forme bilinéaire symétrique associée à Q .

Soit B une forme bilinéaire vérifiant $B(x, x) = Q(x)$. On a

$$Q(x+y) = B(x+y, x+y) = B(x, x) + B(x, y) + B(y, x) + B(y, y) = Q(x) + B(x, y) + B(y, x) + Q(y) .$$

Si B est symétrique, on a $B(y, x) = B(x, y)$ et l'on obtient la formule du théorème. ■

Exemples 36.6. 1) Etant donnée une forme quadratique Q , il existe en général une infinité de formes bilinéaires B vérifiant $B(x, x) = Q(x)$. Par exemple, la relation $Q(x_1, x_2) = 2x_1 x_2$ définit une forme quadratique sur \mathbf{R}^2 : la forme bilinéaire *symétrique* associée est $B((x_1, x_2), (y_1, y_2)) = x_1 y_2 + x_2 y_1$. Soit λ un réel quelconque; la forme bilinéaire B_λ définie par $B_\lambda((x_1, x_2), (y_1, y_2)) = \lambda x_1 y_2 + (2 - \lambda) x_2 y_1$ vérifie aussi $B_\lambda(x, x) = Q(x)$, quel que soit λ , mais elle n'est symétrique que si $\lambda = 1$.

2) Soit Q la forme quadratique sur K^n donnée par

$$Q(x_1, \dots, x_n) = \sum_{1 \leq j \leq k \leq n} a_{jk} x_j x_k .$$

La matrice (b_{jk}) dans la base canonique de la forme bilinéaire associée à Q est donnée par

$$b_{jj} = a_{jj} \quad , \quad b_{jk} = \frac{1}{2} a_{jk} \quad \text{si } j < k \quad , \quad \text{et} \quad b_{jk} = \frac{1}{2} a_{kj} \quad \text{si } j > k .$$

3) Si f est une forme linéaire non nulle sur E , l'application $x \mapsto f(x)^2$ est une forme quadratique sur E (la forme bilinéaire symétrique associée est $B(x, y) = f(x)f(y)$). On a $\widehat{B}(y) = f(y)f$, de sorte que l'image de \widehat{B} est la droite de E^* engendrée par f ; le rang de Q est donc 1. On pourra montrer en exercice que toute forme quadratique de rang 1 est proportionnelle à une forme de ce type.

37. Orthogonalité, vecteurs isotropes

Définition 37.1.— Soit B une forme bilinéaire symétrique sur un espace vectoriel E .

1) Des vecteurs x et y de E sont dits orthogonaux si $B(x, y) = 0$.

2) Soit U une partie de E ; on appelle l'ensemble

$$U^\perp = \{y \in E \mid \forall x \in U \quad B(x, y) = 0\}$$

l'orthogonal de U .

3) Un vecteur x de E est dit isotrope si $Q(x) = B(x, x) = 0$.

Exemples 37.2. 1) La fonction $Q(x_1, x_2) = x_1^2 - x_2^2$ est une forme quadratique sur \mathbf{R}^2 , de forme bilinéaire associée $B((x_1, x_2), (y_1, y_2)) = x_1y_1 - x_2y_2$. Les vecteurs $(1, 0)$ et $(0, 1)$ sont orthogonaux. L'orthogonal du vecteur $(1, 0)$ est la droite d'équation $x_1 = 0$. Le vecteur $(1, 1)$ est isotrope.

2) Le noyau de B est l'orthogonal de E , c'est-à-dire $\text{Ker}(B) = E^\perp$.

On peut utiliser l'application linéaire \widehat{B} définie en (35.5) pour faire le lien entre la partie U° de E^* définie en 34.1 et la partie U^\perp de E . En effet,

$$\begin{aligned} U^\perp &= \{y \in E \mid \forall x \in U \quad B(x, y) = 0\} \\ &= \{y \in E \mid \forall x \in U \quad \widehat{B}(y)(x) = 0\} = \{y \in E \mid \widehat{B}(y) \in U^\circ\}, \end{aligned}$$

de sorte que

$$(37.3) \quad U^\perp = \widehat{B}^{-1}(U^\circ).$$

Cela entraîne en particulier par (34.2) que U^\perp est un sous-espace vectoriel de E (on peut aussi le vérifier directement!)

Théorème 37.4.— Soient B une forme bilinéaire symétrique sur un espace vectoriel E , et F un sous-espace vectoriel de E . On a

$$\dim(F) + \dim(F^\perp) \geq \dim(E) \quad \text{et} \quad F \subset (F^\perp)^\perp.$$

Lorsque B est non dégénérée, il y a égalité dans ces deux relations.

Choisissons une base (e_1, \dots, e_p) de F . Le sous-espace vectoriel F^\perp de E est défini par les p équations linéaires $B(e_1, y) = \dots = B(e_p, y) = 0$. Il est donc de dimension au moins $\dim(E) - p$. On sait aussi (37.3) que F^\perp est l'image inverse de F° par l'application linéaire \widehat{B} . Lorsque B est non dégénérée, \widehat{B} est bijective, de sorte que F^\perp et F° ont même dimension, c'est-à-dire $\dim(E) - \dim(F)$ (cf. th. 34.3).

Pour tout x dans F , on a par définition $B(x, y) = 0$ pour tout $y \in F^\perp$, soit encore $B(y, x) = 0$ puisque B est symétrique. Ceci montre que F est contenu dans $(F^\perp)^\perp$. Lorsque B est non dégénérée, il résulte de ce qui précède que ces espaces ont même dimension $\dim(E) - \dim(F^\perp)$, donc sont égaux. ■

Remarque 37.5. Même lorsque B est non dégénérée, E n'est pas en général somme directe de F et F^\perp . Il suffit de reprendre l'exemple 37.2 : la forme B sur \mathbf{R}^2 est non dégénérée et la droite $F = \{(x_1, x_2) \mid x_1 = x_2\}$ vérifie $F = F^\perp$; on a donc bien $\dim(F) + \dim(F^\perp) = 2$, mais $F \cap F^\perp \neq \{0\}$.

38. Bases orthogonales

Définition 38.1.— Soit B une forme bilinéaire symétrique sur un espace vectoriel E . Une base (e_1, \dots, e_n) de E est dite orthogonale (pour B) si l'on a $B(e_j, e_k) = 0$ pour $j \neq k$.

En d'autres termes, une base \mathcal{B} est orthogonale pour B si la matrice $M_{\mathcal{B}, \mathcal{B}}$ est diagonale. Il ressort de la formule (35.3) que s'il existe une base orthogonale pour B , la matrice de B dans une base quelconque de E est symétrique : la forme bilinéaire B est donc *symétrique* (cf. 36.2).

Théorème 38.2.— Toute forme bilinéaire symétrique admet une base orthogonale.

Soit B une forme bilinéaire symétrique sur un espace vectoriel E . On procède par récurrence sur la dimension n de E . Si $n = 1$, toute base est orthogonale. Supposons $n \geq 2$. Si B est nulle, toute base est orthogonale. On suppose donc B non nulle ; la formule (36.5) montre qu'il existe un vecteur e_1 non isotrope. Si F est la droite vectorielle engendrée par e_1 , on a donc $F \cap F^\perp = \{0\}$. Le théorème 37.4 entraîne que l'on a $\dim F + \dim F^\perp \geq n$, de sorte que $E = F \oplus F^\perp$. L'hypothèse de récurrence montre qu'il existe une base (e_2, \dots, e_n) de F^\perp orthogonale pour la restriction de B à ce sous-espace. Comme on a de plus $B(e_1, e_j) = 0$ pour $2 \leq j \leq n$, la famille (e_1, e_2, \dots, e_n) est une base de E orthogonale pour B . ■

Corollaire 38.3.— Toute forme quadratique sur un espace vectoriel de dimension n est combinaison linéaire de n carrés de formes linéaires.

Soit B une forme bilinéaire symétrique sur un espace vectoriel E de dimension n . Par le théorème 38.2, il existe une base $\mathcal{B} = (e_1, \dots, e_n)$ de E orthogonale pour B . Tout élément x de E s'écrit $\sum_{j=1}^n x_j e_j$, et

$$Q(x) = B(x, x) = \sum_{j,k} B(e_j, e_k) x_j x_k = \sum_j B(e_j, e_j) x_j^2.$$

Mais x_j s'écrit aussi $e_j^*(x)$, où (e_1^*, \dots, e_n^*) est la base duale de \mathcal{B} . On a donc $Q = \sum_j B(e_j, e_j) (e_j^*)^2$. ■

39. Classification des formes quadratiques sur \mathbf{R} et \mathbf{C}

Définition 39.1.— Des formes quadratiques Q et Q' sur un espace vectoriel E sont dites équivalentes s'il existe un automorphisme u de E tel que, pour tout x dans E , on ait

$$Q(x) = Q'(u(x)).$$

Il revient au même de dire que $B(x, y) = B'(u(x), u(y))$ pour tous x et y (utiliser la formule de polarisation). La relation ainsi définie est une relation d'équivalence. Des formes quadratiques Q et Q' sont équivalentes s'il existe des bases \mathcal{B} et \mathcal{B}' de E pour lesquelles on ait $M_{\mathcal{B}, \mathcal{B}} = M_{\mathcal{B}', \mathcal{B}'}$, ou encore s'il existe une base \mathcal{B} de E et une matrice inversible P telles que $M_{\mathcal{B}', \mathcal{B}} = {}^t P M_{\mathcal{B}, \mathcal{B}} P$. En particulier, des formes quadratiques équivalentes ont même rang. Nous allons voir que la réciproque est vraie sur \mathbf{C} (mais pas sur \mathbf{R}).

Théorème (classification des formes quadratiques sur \mathbf{C}) 39.2.— *Toute forme quadratique de rang r sur un espace vectoriel complexe E peut s'écrire sous la forme*

$$f_1^2 + \cdots + f_r^2 ,$$

où f_1, \dots, f_r sont des formes linéaires sur E indépendantes.

Soient Q une forme quadratique sur E et $\mathcal{B} = (e_1, \dots, e_n)$ une base orthogonale pour Q , de sorte qu'il existe des complexes $\lambda_1, \dots, \lambda_n$ tels que $Q(\sum_{j=1}^n x_j e_j) = \sum_{j=1}^n \lambda_j x_j^2$. La matrice de Q dans la base \mathcal{B} est la matrice diagonale de coefficients diagonaux $\lambda_1, \dots, \lambda_n$, de sorte que le rang r de Q est le nombre de λ_j non nuls; quitte à réordonner les éléments de \mathcal{B} , on peut supposer $\lambda_1, \dots, \lambda_r$ non nuls et λ_j nul pour $j > r$. Pour $1 \leq j \leq r$, soit μ_j un nombre complexe (non nul) dont le carré est λ_j (c'est cette étape qui ne serait pas toujours possible sur \mathbf{R}). Si (f_1, \dots, f_r) est la base duale de la base $(e_1/\mu_1, \dots, e_r/\mu_r, e_{r+1}, \dots, e_n)$ de E , la forme quadratique Q s'écrit alors $f_1^2 + \cdots + f_r^2$. ■

Les formes linéaires indépendantes qui apparaissent dans cette décomposition ne sont en général pas uniquement déterminées par Q : la forme quadratique Q sur \mathbf{C}^2 définie par $Q(x, y) = 2x^2 + 2y^2$ se décompose par exemple en

$$Q(x_1, x_2) = (x_1 \sqrt{2})^2 + (x_2 \sqrt{2})^2 = (x_1 + x_2)^2 + (x_1 - x_2)^2 .$$

Le théorème n'est plus valable pour les espaces vectoriels réels : les formes quadratiques Q sur \mathbf{R}^n de la forme du théorème doivent vérifier $Q(x) \geq 0$ pour tout x , ce qui n'est pas toujours réalisé en général (cf. exemple 37.2).

Corollaire 39.3.— *Pour que des formes quadratiques sur un espace vectoriel complexe soient équivalentes, il faut et il suffit qu'elles aient même rang.*

On a déjà vu que des formes quadratiques équivalentes ont même rang. Inversement, si Q et Q' sont des formes quadratiques de même rang sur un espace vectoriel complexe E , le théorème montre qu'elles ont même matrice $\begin{pmatrix} I_r & 0 \\ 0 & 0_{n-r} \end{pmatrix}$ dans des bases peut-être différentes de E ; elles sont donc équivalentes. ■

Passons maintenant aux formes quadratiques réelles.

Définition 39.4.— Une forme quadratique Q sur un espace vectoriel réel E est dite positive si $Q(x) \geq 0$ pour tout élément x de E . Elle est dite définie positive si $Q(x) > 0$ pour tout x non nul.

Théorème (classification des formes quadratiques sur \mathbf{R}) 39.5.— Toute forme quadratique de rang r sur un espace vectoriel réel E peut s'écrire sous la forme

$$f_1^2 + \cdots + f_p^2 - f_{p+1}^2 - \cdots - f_r^2,$$

où f_1, \dots, f_r sont des formes linéaires sur E indépendantes, et où p est un entier qui ne dépend que de Q . Le couple $(p, r - p)$ s'appelle la signature de Q .

Soient Q une forme quadratique sur E et $\mathcal{B} = (e_1, \dots, e_n)$ une base orthogonale pour Q , de sorte qu'il existe des réels $\lambda_1, \dots, \lambda_n$ tels que $Q(\sum_{j=1}^n x_j e_j) = \sum_{j=1}^n \lambda_j x_j^2$. Comme dans la démonstration du théorème 39.2, on peut supposer $\lambda_1, \dots, \lambda_r$ non nuls et λ_j nul pour $j > r$ puis, quitte à réordonner une nouvelle fois les éléments de \mathcal{B} , les p premiers λ_j strictement positifs, et les $r - p$ suivants strictement négatifs. On pose alors $\mu_j = \sqrt{|\lambda_j|}$. Si (f_1, \dots, f_n) est la base duale de la base $(e_1/\mu_1, \dots, e_r/\mu_r, e_{r+1}, \dots, e_n)$ de E , la forme quadratique Q s'écrit alors $f_1^2 + \cdots + f_p^2 - f_{p+1}^2 - \cdots - f_r^2$.

Il reste à montrer que l'entier p ne dépend pas de la base dans laquelle Q s'écrit sous la forme du théorème. Soient F le sous-espace vectoriel de E engendré par e_1, \dots, e_p , et G le sous-espace vectoriel engendré par e_{p+1}, \dots, e_n . La restriction de Q à F est définie positive, et sa restriction à G est négative. Si la restriction de Q à un sous-espace vectoriel H de E est définie positive, on a $H \cap G = \{0\}$, donc $\dim(H) + \dim(G) \leq n$, soit $\dim(H) \leq p$. Cela montre que p est la plus grande dimension d'un sous-espace vectoriel de E sur lequel Q est définie positive; il ne dépend donc que de Q . ■

Corollaire 39.6.— Pour que des formes quadratiques sur un espace vectoriel réel soient équivalentes, il faut et il suffit qu'elles aient même signature.

On a déjà vu que des formes quadratiques équivalentes ont même rang. L'entier p intervenant dans la signature a été caractérisé dans la démonstration précédente comme la plus grande dimension d'un sous-espace vectoriel sur lequel la forme quadratique est définie positive : il est donc le même pour des formes quadratiques équivalentes. Inversement, si Q et Q' sont des formes quadratiques de même signature sur un espace vectoriel réel E , le théorème montre qu'elles ont même matrice $\begin{pmatrix} I_p & 0 & 0 \\ 0 & -I_{r-p} & 0 \\ 0 & 0 & 0_{n-r} \end{pmatrix}$ dans des bases peut-être différentes de E ; elles sont donc équivalentes. ■

40. Méthode de Gauss de réduction des formes quadratiques

Pour trouver le rang d'une forme quadratique, on peut calculer celui de sa matrice dans une base. En revanche, les résultats que nous avons démontrés ne permettent pas en

général de déterminer pratiquement la signature d'une forme quadratique réelle. La méthode de Gauss va nous permettre de décomposer effectivement toute forme quadratique (réelle ou complexe) en combinaison linéaire de carrés de formes linéaires indépendantes, donc en particulier de déterminer la signature dans le cas réel.

Premier cas : Q contient un terme en x_1^2 (Q « contient un carré »); elle s'écrit alors

$$Q(x_1, \dots, x_n) = \lambda_1 x_1^2 + x_1 f(x_2, \dots, x_n) + R(x_2, \dots, x_n),$$

où λ_1 est non nul, f est une forme linéaire, et R une forme quadratique. On s'arrange pour que les deux premiers termes soient ceux du développement d'un binôme, en écrivant

$$Q(x) = \lambda_1 \left(x_1 + \frac{f}{2\lambda_1} \right)^2 - \frac{f^2}{4\lambda_1} + R.$$

Soit f_1 la forme linéaire $x_1 + \frac{f}{2\lambda_1}$; l'hypothèse de récurrence entraîne que l'on peut écrire la forme quadratique $-\frac{f^2}{4\lambda_1} + R$ comme $\sum_{j=2}^r \lambda_j f_j^2$, où f_2, \dots, f_r sont des formes linéaires en (x_2, \dots, x_n) qui sont indépendantes. Comme f_1 contient un terme en x_1 , elle n'est pas combinaison linéaire de f_2, \dots, f_r . Les formes linéaires f_1, f_2, \dots, f_r sont donc indépendantes, d'où $Q = \sum_{j=1}^r \lambda_j f_j^2$.

Deuxième cas : Q ne contient pas de carré; elle s'écrit alors (après renumérotage éventuel des x_j)

$$Q(x_1, \dots, x_n) = ax_1 x_2 + x_1 f'_1(x_3, \dots, x_n) + x_2 f'_2(x_3, \dots, x_n) + R(x_3, \dots, x_n),$$

où a est non nul, f'_1 et f'_2 sont des formes linéaires, et R une forme quadratique. On écrit

$$\begin{aligned} Q(x) &= a \left(x_1 + \frac{f'_2}{a} \right) \left(x_2 + \frac{f'_1}{a} \right) - \frac{f'_1 f'_2}{a} + R \\ &= \frac{a}{4} \left(x_1 + x_2 + \frac{f'_2 + f'_1}{a} \right)^2 - \frac{a}{4} \left(x_1 - x_2 + \frac{f'_2 - f'_1}{a} \right)^2 - \frac{f'_1 f'_2}{a} + R, \end{aligned}$$

où l'on a utilisé l'identité $uv = \frac{1}{4}(u+v)^2 - \frac{1}{4}(u-v)^2$. Il ne reste plus qu'à appliquer l'hypothèse de récurrence à la forme quadratique $-\frac{1}{a}f'_1 f'_2 + R$, qui s'écrit donc $\sum_{j=3}^r \lambda_j f_j^2$, où f_3, \dots, f_r sont des formes linéaires en (x_3, \dots, x_n) qui sont indépendantes. On vérifie sans mal que les formes linéaires f_1, f_2, \dots, f_r sont indépendantes, d'où $Q = \sum_{j=1}^r \lambda_j f_j^2$.

Exemples 40.1. 1) Considérons la forme quadratique $Q(x, y, z) = xy + yz + zx$ sur \mathbf{R}^3 . Elle est « sans terme carré »; on écrit

$$Q(x, y, z) = (x+z)(y+z) - z^2 = \frac{1}{4}(x+y+2z)^2 - \frac{1}{4}(x-y)^2 - z^2.$$

La forme Q est de signature $(1, 2)$, donc de rang 3 : elle est non dégénérée.

2) Considérons la forme quadratique $Q(x, y, z, t) = xy + yz + zt + tx$ sur \mathbf{R}^4 . Elle est « sans terme carré »; on écrit

$$Q(x, y, z, t) = (x+z)(y+t) = \frac{1}{4}(x+y+z+t)^2 - \frac{1}{4}(x-y+z-t)^2.$$

La forme Q est de signature $(1, 1)$, donc de rang 2 : elle est dégénérée.

3) Considérons la forme quadratique $Q(x, y, z, t) = x^2 + 2y^2 + 3z^2 - 2zt + tx + 3xy - yt$ sur \mathbf{R}^4 . Elle a plusieurs termes carrés, mais il est avantageux de commencer par le plus simple, c'est-à-dire celui qui apparaît le moins de fois, ici z . On écrit

$$Q(x, y, z, t) = 3z^2 - 2zt + R(x, y, t) = 3\left(z - \frac{t}{3}\right)^2 - \frac{t^2}{3} + R(x, y, t)$$

et

$$-\frac{t^2}{3} + R(x, y, t) = x^2 + 2y^2 + tx + 3xy - yt - \frac{t^2}{3}.$$

On choisit x :

$$-\frac{t^2}{3} + R(x, y, t) = \left(x + \frac{1}{2}(t + 3y)\right)^2 - \frac{1}{4}(t + 3y)^2 + 2y^2 - yt - \frac{t^2}{3}.$$

Il ne reste alors que y et t , et

$$-\frac{1}{4}(t + 3y)^2 + 2y^2 - yt - \frac{t^2}{3} = -\frac{1}{4}y^2 - \frac{7}{12}t^2 - \frac{5}{2}yt.$$

On choisit y :

$$-\frac{1}{4}y^2 - \frac{7}{12}t^2 - \frac{5}{2}yt = -\frac{1}{4}(y + 5t)^2 + \frac{17}{3}t^2.$$

Au total, on a obtenu

$$Q(x, y, z, t) = 3\left(z - \frac{t}{3}\right)^2 + \left(x + \frac{1}{2}(t + 3y)\right)^2 - \frac{1}{4}(y + 5t)^2 + \frac{17}{3}t^2.$$

La forme Q est de signature $(3, 1)$, donc de rang 4 : elle est non dégénérée.

4) Considérons la forme quadratique $Q(x, y, z) = x^2 + y^2 + 3z^2 + 4xy + 2xz + 2yz$ sur \mathbf{R}^3 .

On a

$$\begin{aligned} Q(x, y, z) &= x^2 + x(4y + 2z) + y^2 + 3z^2 + 2yz \\ &= (x + 2y + z)^2 - (2y + z)^2 + y^2 + 3z^2 + 2yz \\ &= (x + 2y + z)^2 - 3y^2 + 2z^2 - 2yz \\ &= (x + 2y + z)^2 - 3\left(y + \frac{1}{3}z\right)^2 + \frac{1}{3}z^2 + 2z^2 \\ &= (x + 2y + z)^2 - 3\left(y + \frac{z}{3}\right)^2 + \frac{7}{3}z^2. \end{aligned}$$

La forme Q est de signature $(2, 1)$, donc de rang 3 : elle est non dégénérée.

41. Groupe orthogonal

(41.1) Soient E un espace vectoriel, Q une forme quadratique sur E de forme bilinéaire associée B , et u un endomorphisme de E . Les propriétés suivantes sont équivalentes :

(i) l'endomorphisme u préserve la forme bilinéaire B , c'est-à-dire

$$\forall x, y \in E \quad B(u(x), u(y)) = B(x, y);$$

(ii) l'endomorphisme u préserve la forme quadratique Q , c'est-à-dire

$$\forall x \in E \quad Q(u(x)) = Q(x) .$$

(41.2) En effet, il est clair que (i) entraîne (ii). Soit u un endomorphisme qui préserve la forme quadratique Q ; la formule (36.5) entraîne, pour tous x, y dans E ,

$$\begin{aligned} 2B(u(x), u(y)) &= Q(u(x) + u(y)) - Q(u(x)) - Q(u(y)) \\ &= Q(u(x + y)) - Q(u(x)) - Q(u(y)) = Q(x + y) - Q(x) - Q(y) = 2B(x, y) , \end{aligned}$$

ce qui montre que u préserve B . On dit qu'un automorphisme de E qui vérifie l'une des propriétés (i) ou (ii) est *orthogonal pour Q* .

Proposition 41.3.— *Soit Q une forme quadratique sur un espace vectoriel E . Les automorphismes de E orthogonaux pour Q forment un sous-groupe de $GL(E)$ que l'on appelle le groupe orthogonal de Q , et que l'on note $O(Q)$.*

L'identité est un automorphisme de E qui préserve Q ; il est donc dans $O(Q)$. Il s'agit de montrer que si u et v préservent Q , il en est de même de $v^{-1}u$. Or, pour tout x dans E , on a

$$Q(v^{-1}u(x)) = Q(v(v^{-1}u(x))) = Q(u(x)) = Q(x) ,$$

ce qui montre la proposition. ■

Remarque 41.4. Lorsque Q est *non dégénérée*, tout endomorphisme u de E qui préserve Q est bijectif, donc est dans $O(Q)$. En effet, si $u(x) = 0$, on a, puisque u préserve B ,

$$B(x, y) = B(u(x), u(y)) = 0$$

pour tout y dans E , de sorte que x est orthogonal à E , donc nul. On en déduit que u est injectif, donc bijectif.

On peut caractériser les endomorphismes orthogonaux par leur matrice.

Proposition 41.5.— *Soit Q une forme quadratique sur un espace vectoriel E . Soient \mathcal{B} une base de E et $M_{\mathcal{B}}$ la matrice de Q dans cette base. Soient u un endomorphisme de E et M sa matrice dans la base \mathcal{B} .*

1) *Pour que l'endomorphisme u soit orthogonal pour Q , il faut et il suffit que ${}^tMM_{\mathcal{B}}M = M_{\mathcal{B}}$.*

2) *Si Q est non dégénérée, le déterminant d'un endomorphisme orthogonal pour Q est ± 1 . Les endomorphismes orthogonaux pour Q et de déterminant 1 forment un sous-groupe de $O(Q)$ que l'on appelle le groupe spécial orthogonal de Q , et que l'on note $SO(Q)$.*

Si l'on représente des vecteurs x et y de E par les matrices colonnes X et Y de leurs composantes dans la base \mathcal{B} , on a (formule (35.4))

$$B(x, y) = {}^tXM_{\mathcal{B}}Y$$

d'où,

$$B(u(x), u(y)) = {}^t(MX)(MY) = {}^tX{}^tMM_{\mathcal{B}}MY ,$$

c'est-à-dire que la matrice de la forme bilinéaire $(x, y) \mapsto B(u(x), u(y))$ dans la base \mathcal{B} est ${}^tMM_{\mathcal{B}}M$. Pour que l'endomorphisme u soit orthogonal, il faut et il suffit que cette forme bilinéaire soit égale à B , ou encore que sa matrice soit celle de B , c'est-à-dire ${}^tMM_{\mathcal{B}}M = M_{\mathcal{B}}$. Ceci prouve 1).

Si Q est non dégénérée, la matrice $M_{\mathcal{B}}$ est inversible; en prenant les déterminants dans l'égalité de 1), on obtient

$$\det({}^tM) \det(M_{\mathcal{B}}) \det(M) = \det(M_{\mathcal{B}}) ,$$

d'où $\det(M)^2 = 1$, puisque $\det({}^tM) = \det(M)$ et $\det(M_{\mathcal{B}}) \neq 0$. Le sous-ensemble $SO(Q)$ de $O(Q)$ est le noyau du morphisme de groupes $\det : O(Q) \rightarrow \{\pm 1\}$; c'est donc un sous-groupe. ■

Proposition 41.6.— *Les groupes orthogonaux de formes quadratiques équivalentes sont isomorphes.*

Soient Q et Q' des formes quadratiques équivalentes sur un espace vectoriel E ; il existe un automorphisme v de E tel que, pour tout x dans E , on ait $Q(x) = Q'(v(x))$. Soit $u \in O(Q)$; pour tout x dans E , on a

$$Q'(v \circ u \circ v^{-1}(x)) = Q(u \circ v^{-1}(x)) = Q(v^{-1}(x)) = Q'(v \circ v^{-1}(x)) = Q'(x) ,$$

de sorte que $vuv^{-1} \in O(Q')$. L'application $O(Q) \rightarrow O(Q')$, $u \mapsto v \circ u \circ v^{-1}$ est bijective (son inverse est donné par $u' \mapsto v^{-1} \circ u' \circ v$). On vérifie sans difficulté que c'est un morphisme de groupes. C'est donc un isomorphisme de groupes de $O(Q)$ sur $O(Q')$. ■

Cette proposition, jointe aux théorèmes 39.2 et 39.5, montre que l'étude des groupes orthogonaux des formes quadratiques *non dégénérées* se ramène

- sur le corps \mathbf{C} , au cas $Q(x) = x_1^2 + \cdots + x_n^2$; le groupe correspondant est noté $O(n, \mathbf{C})$;

- sur le corps \mathbf{R} , au cas $Q(x) = x_1^2 + \cdots + x_p^2 - x_{p+1}^2 - \cdots - x_{p+q}^2$; le groupe correspondant est noté $O(p, q, \mathbf{R})$. C'est le groupe orthogonal de toutes les formes quadratiques non dégénérées de signature (p, q) . On note $O(n, \mathbf{R})$ au lieu de $O(n, 0, \mathbf{R})$. Comme $O(Q) = O(-Q)$, les groupes $O(p, q, \mathbf{R})$ et $O(q, p, \mathbf{R})$ sont isomorphes.

La proposition 41.5 entraîne que le groupe $O(n, K)$ est isomorphe au groupe des matrices carrées $M \in \mathcal{M}_n(K)$ orthogonales, c'est-à-dire vérifiant ${}^tMM = I$.

(42.1) Soient f_1, \dots, f_n des formes linéaires sur un espace vectoriel E . Montrer que f_1, \dots, f_n engendrent E^* si et seulement si

$$\text{Ker}(f_1) \cap \dots \cap \text{Ker}(f_n) = \{0\}.$$

(42.2) Soit E l'espace vectoriel des polynômes de degré $\leq n$ à coefficients réels, et soient a_0, \dots, a_n des réels deux à deux distincts. On considère les formes linéaires f_0, \dots, f_n sur E définies par $f_j(P) = P(a_j)$ pour tout élément P de E .

a) Quelle est la dimension de E ?

b) Montrer que $\mathcal{B}^* = (f_0, \dots, f_n)$ est une base de E^* (on pourra utiliser l'exercice précédent).

c) En déduire que la matrice $(a_j^k)_{0 \leq j, k \leq n}$ est inversible.

d) Déterminer la base de E duale de \mathcal{B}^* .

e) (Interpolation de Lagrange) Etant donnés des réels b_0, \dots, b_n , déterminer un polynôme P de degré au plus n qui prend la valeur b_j en a_j , pour chaque $j = 0, \dots, n$.

(42.3) Soit E l'espace vectoriel des polynômes de degré ≤ 3 à coefficients réels. On considère les formes linéaires f_1, \dots, f_4 sur E définies par

$$f_1(P) = P(0), \quad f_2(P) = P(1), \quad f_3(P) = P'(0), \quad f_4(P) = P'(1)$$

pour tout élément P de E .

a) Montrer que $\mathcal{B}^* = (f_1, \dots, f_4)$ est une base de E^* .

b) Déterminer la base de E duale de \mathcal{B}^* .

c) Soit f la forme linéaire sur E définie par

$$f(P) = \int_0^1 P(t) dt.$$

Déterminer les composantes de f dans la base \mathcal{B}^* .

(42.4) Soient E un espace vectoriel, U et V des parties de E . Montrer les relations suivantes

$$\begin{aligned} U \subset V &\implies V^\circ \subset U^\circ \\ U \subset (U^\circ)^\circ &\qquad (U \cup V)^\circ = U^\circ \cap V^\circ & (U \cap V)^\circ \supset U^\circ + V^\circ. \end{aligned}$$

On suppose que U et V sont des sous-espaces vectoriels de E ; montrer que l'on a égalité dans les inclusions de la deuxième ligne.

(42.5) Soient E, F et G des espaces vectoriels, et $u : E \rightarrow F$ et $v : F \rightarrow G$ des applications linéaires; montrer que

$${}^t(v \circ u) = {}^t u \circ {}^t v \qquad {}^t(\text{Id}_E) = \text{Id}_{E^*}.$$

(42.6) Soient \mathcal{B} et \mathcal{B}' des bases d'un espace vectoriel E et soit P la matrice de passage de \mathcal{B} à \mathcal{B}' (cf. (4.5)). Déterminer la matrice de passage de la base \mathcal{B}^* à la base \mathcal{B}'^* de E^* (on pourra procéder directement ou utiliser la prop. 34.8).

(42.7) Soit B une forme bilinéaire symétrique sur un espace vectoriel E et soient U et V des parties de E . Montrer les relations suivantes

$$\begin{aligned} U \subset V &\implies V^\perp \subset U^\perp \\ U \subset (U^\perp)^\perp &\qquad (U \cup V)^\perp = U^\perp \cap V^\perp & (U \cap V)^\perp \supset U^\perp + V^\perp. \end{aligned}$$

On suppose que U et V sont des sous-espaces vectoriels de E ; montrer que l'on a égalité dans les deux dernières inclusions lorsque B est non dégénérée, mais pas en général.

(42.8) Soit Q une forme quadratique combinaison linéaire à coefficients non nuls des carrés des formes linéaires indépendantes f_1, \dots, f_r . Montrer que r est le rang de Q et que son noyau est l'intersection des noyaux des f_j .

(42.9) Soit B une forme bilinéaire sur un espace vectoriel E . On suppose que la relation $B(x, y) = 0$ entraîne $B(y, x) = 0$. Montrer que B est soit symétrique, soit alternée (cf. déf. 7.1) (on pourra appliquer l'hypothèse à $B(x, B(x, y)z - B(x, z)y)$).

(42.10) Soit M une matrice carrée symétrique inversible à coefficients complexes. Montrer qu'il existe une matrice P telle que $M = {}^t P P$.

(42.11) Montrer que l'application $(A, B) \mapsto \text{Tr}(AB)$ définit une forme bilinéaire symétrique non dégénérée sur l'espace vectoriel des matrices carrées. Lorsque le corps de base est \mathbf{R} , quelle est sa signature?

(42.12) Réduire en combinaison linéaire de carrés de formes linéaires indépendantes les formes quadratiques suivantes

$$Q(x, y, z, t) = 9x^2 - 6y^2 - 8z^2 + 6xy - 14xz + 18xt + 8yz + 12yt - 4zt$$

$$Q'(x, y, z, t) = x^2 + y^2 + z^2 - 2t^2 - 2xy + 2xz - 2xt + 2yz - 4yt.$$

(42.13) Soit Q la forme quadratique sur \mathbf{R}^3 définie par $Q(x, y, z) = x^2 + y^2 + z^2 - 4(xy + yz + zx)$.

a) Décomposer Q en combinaison linéaire de carrés de formes linéaires indépendantes. Déterminez son rang et sa signature.

b) Trouver une base de \mathbf{R}^3 orthogonale pour Q .

(42.14) Soit Q la forme quadratique sur \mathbf{R}^3 définie par $Q(x, y, z) = x^2 - 2yz - xz$.

a) Déterminez le rang et la signature de Q . Trouver une base de \mathbf{R}^3 orthogonale pour Q .

b) Pour chaque réel t , on note P_t le plan d'équation $3x + 2y + tz = 0$. Déterminez les valeurs de t pour lesquelles P_t contient un vecteur isotrope non nul, puis celles pour lesquelles la restriction de Q à P_t est dégénérée.

(42.15) Soit M une matrice réelle carrée d'ordre 2. On pose $Q(M) = \det(M) - \text{Tr}(M)^2/4$.

a) Montrer que Q est une forme quadratique non dégénérée. Quelle est sa signature? Exprimer $Q(M)$ à l'aide de $\text{Tr}(M)$ et de $\text{Tr}(M^2)$.

b) Déterminer la forme bilinéaire symétrique associée à Q , ainsi que son noyau. Etudiez les valeurs propres de M suivant le signe de $Q(M)$.

(42.16) **Formule de Pythagore.** Soient Q une forme quadratique sur un espace vectoriel E et e_1, \dots, e_r des vecteurs de E deux à deux orthogonaux pour Q . Montrer l'égalité

$$Q(e_1 + \dots + e_r) = Q(e_1) + \dots + Q(e_r).$$

(42.17) Soit Q la forme quadratique à n variables

$$Q(x_1, \dots, x_n) = \sum_{1 \leq j < k \leq n} x_j x_k.$$

Déterminer des formes linéaires indépendantes f_1, \dots, f_n à coefficients rationnels telles que

$$Q = f_1^2 + \frac{3}{4}f_2^2 + \frac{4}{6}f_3^2 + \dots + \frac{n+1}{2n}f_n^2.$$

(42.18) Déterminer la signature des formes quadratiques suivantes

$$Q(x_1, \dots, x_n) = \sum_{1 \leq j < k \leq n} x_j x_k$$

$$Q'(x_1, \dots, x_n) = x_1 x_2 + x_2 x_3 + \dots + x_{n-1} x_n + x_n x_1.$$

(42.19) Soit M une matrice réelle antisymétrique (c'est-à-dire vérifiant ${}^tM = -M$). Montrer que la matrice e^M est orthogonale.

SOLUTIONS

(42.6) C'est ${}^tP^{-1}$.

(42.7) Lorsque B est non dégénérée, la première égalité donne

$$(U^\perp + V^\perp)^\perp = (U^\perp \cup V^\perp)^\perp = (U^\perp)^\perp \cap (V^\perp)^\perp = U \cap V.$$

Il suffit de reprendre l'orthogonal.

(42.9) Comme $B(x, B(x, y)z - B(x, z)y) = 0$, on a $B(B(x, y)z - B(x, z)y, x) = 0$, c'est-à-dire $B(x, y)B(z, x) = B(x, z)B(y, x)$. Supposons qu'il existe x_0 tel que $B(x_0, x_0)$ soit non nul (dans le cas contraire, B est antisymétrique par 4.a)). On prend $x = z = x_0$ dans la formule, d'où $B(y, x_0) = B(x_0, y)$ pour tout y . Soit x quelconque; si $B(x, x_0) \neq 0$, on obtient $B(x, y) = B(y, x)$ en faisant $z = x_0$ dans la formule. Si $B(x, x_0) = 0$, on a $B(x + x_0, x_0) \neq 0$, d'où $B(x + x_0, y) = B(y, x + x_0)$ et B est symétrique.

(42.17) On obtient

$$\begin{aligned} Q(x) &= (x_1 + \frac{1}{2}(x_2 + \dots + x_n))^2 + \frac{3}{4}[\sum_{j \geq 2} x_j^2 + \frac{2}{3} \sum_{k > j \geq 2} x_j x_k] \\ &= f_1^2 + \frac{3}{4}[(x_2 + \frac{1}{3}(x_3 + \dots + x_n))^2 + \frac{8}{9} \sum_{j \geq 3} x_j^2 + \frac{4}{9} \sum_{k > j \geq 3} x_j x_k] \\ &= f_1^2 + \frac{3}{4}f_2^2 + \frac{4}{6}[\sum_{j \geq 3} x_j^2 + \frac{2}{4} \sum_{k > j \geq 3} x_j x_k] = \dots \end{aligned}$$

et on peut prendre $f_j = x_j + \frac{1}{j+1}(x_{j+1} + \dots + x_n)$.

(42.18) Pour la première, on écrit

$$\begin{aligned} Q(x) &= (x_1 + x_3 + \dots + x_n)(x_2 + x_3 + \dots + x_n) - (x_3 + \dots + x_n)^2 + \sum_{3 \leq j < k} x_j x_k \\ &= (x_1 + x_3 + \dots + x_n)(x_2 + x_3 + \dots + x_n) - \sum_{3 \leq j \leq k} x_j x_k. \end{aligned}$$

On peut utiliser l'exercice précédent : la signature est $(1, n-1)$.

Pour la seconde, on écrit

$$\begin{aligned} Q'(x) &= (x_1 + x_3)(x_2 + x_n) + x_3 x_4 + \dots + x_{n-1} x_n - x_3 x_n \\ &= (x_1 + x_3)(x_2 + x_n) + (x_3 + x_5)(x_4 - x_n) + x_5 x_6 + \dots + x_{n-1} x_n + x_n x_5. \end{aligned}$$

On est ramené à la forme du départ. Il suffit de traiter les cas $n = 1, 2, 3$ et 4 . On trouve

$$\begin{aligned} \text{si } n \equiv 0 \pmod{4}, \quad & \text{la signature est } \left(\frac{n}{2} - 1, \frac{n}{2} - 1\right); \\ \text{si } n \equiv 1 \pmod{4}, \quad & \text{la signature est } \left(\frac{n+1}{2}, \frac{n-1}{2}\right); \\ \text{si } n \equiv 2 \pmod{4}, \quad & \text{la signature est } \left(\frac{n}{2}, \frac{n}{2}\right); \\ \text{si } n \equiv 3 \pmod{4}, \quad & \text{la signature est } \left(\frac{n-1}{2}, \frac{n+1}{2}\right). \end{aligned}$$

La forme est dégénérée dans le premier cas, non dégénérée dans les autres.

Dans ce chapitre, on ne considérera que des espaces vectoriels *réels de dimension finie*.

43. Produit scalaire

Définition 43.1.— On appelle produit scalaire sur un espace vectoriel réel E une forme bilinéaire symétrique définie positive. On appelle espace vectoriel euclidien un espace vectoriel réel muni d'un produit scalaire.

On notera le produit scalaire $\langle x, y \rangle$. A part ses propriétés de bilinéarité, il vérifie

$$\langle x, y \rangle = \langle y, x \rangle \quad \langle x, x \rangle > 0 \quad \text{si } x \neq 0 .$$

La seconde relation dit que le seul vecteur isotrope est le vecteur nul. On posera $\|x\| = \sqrt{\langle x, x \rangle}$. La relation de polarisation s'écrit

$$\|x + y\|^2 = \|x\|^2 + \|y\|^2 + 2\langle x, y \rangle ;$$

lorsque x et y sont orthogonaux, elle se réduit à la *relation de Pythagore*

$$\|x + y\|^2 = \|x\|^2 + \|y\|^2 .$$

Exemples 43.2. 1) Les produits scalaires sur un espace vectoriel réel de dimension n sont les formes bilinéaires symétriques de signature $(n, 0)$.

2) La relation

$$\langle (x_1, \dots, x_n), (y_1, \dots, y_n) \rangle = x_1 y_1 + \dots + x_n y_n$$

définit un produit scalaire sur \mathbf{R}^n , donc le munit d'une structure d'espace euclidien dite canonique.

3) Soit E l'espace vectoriel des polynômes à coefficients réels de degré au plus n . La relation

$$\langle P, Q \rangle = \int_{-1}^1 P(t)Q(t)dt$$

définit un produit scalaire sur E (cf. exerc. 52.1).

Lemme (Inégalité de Cauchy-Schwarz) 43.3.— Soit E un espace vectoriel euclidien. Pour tous x et y dans E , on a

$$|\langle x, y \rangle| \leq \|x\| \cdot \|y\| .$$

Il n'y a égalité que si x et y sont colinéaires.

Si $x = 0$, le lemme est évident. On suppose donc $x \neq 0$. Pour tout réel t , la quantité $\langle tx + y, tx + y \rangle$ est positive. On la développe :

$$t^2 \langle x, x \rangle + 2t \langle x, y \rangle + \langle y, y \rangle \geq 0$$

pour tout t . C'est une fonction polynomiale en t de degré 2 ; elle ne garde un signe constant pour tout t que si son discriminant réduit $\langle x, y \rangle^2 - \langle x, x \rangle \langle y, y \rangle$ est négatif, ce qui montre l'inégalité cherchée. Si ce discriminant est nul, il y a une racine double t_0 , qui vérifie alors $\langle t_0 x + y, t_0 x + y \rangle = 0$, de sorte que $t_0 x + y = 0$. ■

(43.4) C'est cette inégalité qui permet de définir l'*angle non orienté* de deux vecteurs non nuls x et y d'un espace euclidien E . En effet, elle entraîne qu'il existe un unique réel θ compris entre 0 et π tel que

$$\cos \theta = \frac{\langle x, y \rangle}{\|x\| \cdot \|y\|} ;$$

on l'appelle l'angle non orienté de x et de y . Pour qu'il soit nul, il faut et il suffit que $y = \lambda x$ avec $\lambda > 0$; pour qu'il soit égal à π , il faut et il suffit que $y = \lambda x$ avec $\lambda < 0$; pour qu'il soit égal à $\pi/2$, il faut et il suffit que x et y soient orthogonaux.

Nous définirons en (46.6) l'angle *orienté* de deux vecteurs dans un *plan euclidien orienté*.

Théorème 43.5.— *L'application $\| \cdot \|$ est une norme (cf. déf. 22.3) : elle vérifie*

$$\begin{aligned} \|\lambda x\| &= |\lambda| \|x\| \\ \|x\| &> 0 \quad \text{si } x \neq 0 \\ \|x + y\| &\leq \|x\| + \|y\| \quad (\text{inégalité triangulaire}) . \end{aligned}$$

Seule l'inégalité triangulaire est à démontrer. Comme ses deux membres sont positifs, il suffit de montrer l'inégalité obtenue par élévation au carré, soit

$$\langle x + y, x + y \rangle \leq \langle x, x \rangle + \langle y, y \rangle + 2\|x\| \cdot \|y\|$$

ou encore, par bilinéarité du produit scalaire,

$$2\langle x, y \rangle \leq 2\|x\| \cdot \|y\| ,$$

qui découle de l'inégalité de Cauchy-Schwarz. ■

Pour qu'il y ait égalité dans l'inégalité triangulaire, il faut et il suffit qu'il y ait égalité dans l'inégalité de Cauchy-Schwarz, donc que x et y soient colinéaires.

Il existe des tas de normes sur \mathbf{R}^n qui ne sont pas euclidiennes (c'est-à-dire qui ne sont pas définies à partir d'un produit scalaire). Par exemple la norme

$$\|(x_1, \dots, x_n)\| = |x_1| + \dots + |x_n| .$$

On donne dans l'exercice 52.5 une caractérisation des normes dites euclidiennes, c'est-à-dire des normes qui sont définies à partir d'un produit scalaire.

44. Orthogonalité

Dans un espace vectoriel euclidien, le théorème 34.3 admet la forme plus précise suivante.

Théorème 44.1.— *Soit E un espace vectoriel euclidien. Pour tout sous-espace vectoriel F de E , on a*

$$E = F \oplus F^\perp .$$

Tout vecteur dans $F \cap F^\perp$ est isotrope, donc nul. On conclut avec le théorème 34.3. ■

(44.2) Par le théorème 39.5, tout produit scalaire admet une base orthogonale. En divisant chaque vecteur de cette base par sa norme, on obtient une base *orthonormée*, c'est-à-dire une base (e_1, \dots, e_n) qui vérifie $\langle e_j, e_k \rangle = \delta_{jk}$. Il est aussi utile d'introduire la notion suivante : on dit que des vecteurs x_1, \dots, x_p forment une *famille orthonormée* si $\langle x_j, x_k \rangle = \delta_{jk}$. Une telle famille est toujours libre (si on a une relation $\sum_{j=1}^p \lambda_j x_j = 0$, on obtient en prenant le produit scalaire par x_k l'égalité $\lambda_k = 0$).

Proposition 44.3.— *Soit (e_1, \dots, e_n) une base orthonormée d'un espace vectoriel euclidien E . Pour tout vecteur x de E , on a*

$$x = \sum_{j=1}^n \langle x, e_j \rangle e_j \qquad \|x\|^2 = \sum_{j=1}^n \langle x, e_j \rangle^2 .$$

On écrit $x = \sum_{j=1}^n x_j e_j$; en prenant le produit scalaire par e_k , on obtient $\langle x, e_k \rangle = x_k$. En prenant le produit scalaire par x , on obtient la seconde égalité. ■

(44.4) On a défini en § 41.1 les endomorphismes orthogonaux. Ce sont les endomorphismes u de E qui vérifient l'une des propriétés équivalentes suivantes :

(i) u préserve le produit scalaire : pour tous x et y dans E ,

$$\langle u(x), u(y) \rangle = \langle x, y \rangle ;$$

(ii) u est une *isométrie* : pour tout x dans E ,

$$\|u(x)\| = \|x\| ;$$

(iii) il existe une base orthonormée de E dont l'image par u est une base orthonormée ;

(iv) l'image par u de toute base orthonormée de E est une base orthonormée ;

(v) la matrice M de u dans une base orthonormée de E est *orthogonale*, c'est-à-dire qu'elle vérifie ${}^tMM = I$.

Vérifions que ces propriétés sont effectivement équivalentes. Il est clair que (i) entraîne (iv), qui entraîne (iii). D'autre part, on a vu en (41.2) que (i) est équivalent à (ii). Montrons que (iii) entraîne (ii) : soit (e_1, \dots, e_n) une base orthonormée telle que $(u(e_1), \dots, u(e_n))$ soit une base orthonormée. Soit x un vecteur de E ; la proposition 44.3 donne $x = \sum_{j=1}^n \langle x, e_j \rangle e_j$ mais aussi $u(x) = \sum_{j=1}^n \langle u(x), u(e_j) \rangle u(e_j)$. En appliquant u à la première de ces égalités, on obtient d'autre part $u(x) = \sum_{j=1}^n \langle x, e_j \rangle u(e_j)$, d'où $\langle u(x), u(e_j) \rangle = \langle x, e_j \rangle$ pour tout j , ce qui entraîne (ii) par la proposition 44.3. Enfin, l'équivalence de (i) et (v) résulte de la proposition 41.5.1).

(44.5) On peut définir la *distance* de deux points x et y d'un espace vectoriel euclidien par $d(x, y) = \|x - y\|$. Soit u un endomorphisme orthogonal ; la propriété (ii) et la linéarité de u entraînent

$$d(u(x), u(y)) = \|u(x) - u(y)\| = \|u(x - y)\| = \|x - y\| = d(x, y) ;$$

on dit que u préserve les distances, ou encore que u est une *isométrie*. C'est pourquoi on parlera dans le contexte des espaces euclidiens d'isométries plutôt que d'endomorphismes orthogonaux. Rappelons (prop. 41.5.2) que le déterminant d'une isométrie est ± 1 . On appelle *isométrie directe* une isométrie de déterminant 1, *isométrie indirecte* une isométrie de déterminant -1 .

Corollaire 44.6.— Soit \mathcal{B} une base orthonormée d'un espace vectoriel euclidien E . Pour qu'une base \mathcal{B}' de E soit orthonormée, il faut et il suffit que la matrice de passage $P_{\mathcal{B} \rightarrow \mathcal{B}'}$ soit orthogonale.

Posons $\mathcal{B} = (e_1, \dots, e_n)$ et $\mathcal{B}' = (e'_1, \dots, e'_n)$. Soit u l'endomorphisme de E défini par $u(e_j) = e'_j$ pour tout j . Le corollaire résulte de l'équivalence des propriétés (iii) et (v). On peut aussi le vérifier par un calcul direct. ■

Proposition 44.7.— Soient u une isométrie d'un espace vectoriel euclidien E et F un sous-espace vectoriel de E .

a) On a $u(F^\perp) = u(F)^\perp$.

b) Si F est stable par u , il en est de même de F^\perp .

Soit y un élément de F^\perp ; pour tout x dans F , on a

$$\langle u(x), u(y) \rangle = \langle x, y \rangle = 0,$$

ce qui signifie que $u(y)$ est orthogonal à tous les éléments de $u(F)$, c'est-à-dire $y \in u(F)^\perp$. Cela montre l'inclusion $u(F^\perp) \subset u(F)^\perp$. Mais u est un automorphisme, de sorte que l'on a, en utilisant le théorème 34.3,

$$\begin{aligned} \dim(u(F^\perp)) &= \dim(F^\perp) = \dim(E) - \dim(F) \\ \dim(u(F)^\perp) &= \dim(E) - \dim(u(F)) = \dim(E) - \dim(F). \end{aligned}$$

Les espaces vectoriels $u(F^\perp)$ et $u(F)^\perp$ ont donc même dimension. Comme l'un est contenu dans l'autre, ils sont égaux, ce qui prouve a). Si F est stable par u , on a $u(F) = F$, d'où $u(F^\perp) = F^\perp$; cela prouve b). ■

Nous allons décrire un algorithme permettant de construire explicitement une base orthonormée à partir de la donnée d'une base quelconque.

Théorème (Orthonormalisation de Schmidt) 44.8.— *Soit (e_1, \dots, e_n) une base d'un espace vectoriel euclidien E . Il existe une unique base orthonormée (e'_1, \dots, e'_n) de E telle que, pour tout $j \in \{1, \dots, n\}$, les familles $\{e_1, \dots, e_j\}$ et $\{e'_1, \dots, e'_j\}$ engendrent le même sous-espace vectoriel de E , et $\langle e_j, e'_j \rangle > 0$.*

Il est plus simple pour les applications pratiques de construire d'abord une base (e''_1, \dots, e''_n) seulement orthogonale qui satisfait aux propriétés demandées. On prend tout d'abord $e''_1 = e_1$. Supposons e''_1, \dots, e''_{j-1} construits ; on cherche e''_j sous la forme

$$e''_j = e_j + \lambda_1 e''_1 + \dots + \lambda_{j-1} e''_{j-1}.$$

Pour $1 \leq k \leq j-1$, la relation $\langle e''_j, e''_k \rangle = 0$ est équivalente à $\langle e_j, e''_k \rangle + \lambda_k \|e''_k\|^2 = 0$, ce qui détermine $\lambda_1, \dots, \lambda_{j-1}$. On construit ainsi par récurrence une base orthogonale (e''_1, \dots, e''_n) . On remarque que $\langle e_j, e''_j \rangle = \langle e''_j, e''_j \rangle > 0$; il faut donc prendre $e'_j = e''_j / \|e''_j\|$. ■

On remarquera que le même procédé permet d'orthonormaliser une famille libre de vecteurs de E .

Corollaire 44.9.— *Soit M une matrice carrée réelle inversible. Il existe une unique paire (O, T) , où O est une matrice orthogonale et T une matrice triangulaire supérieure à coefficients diagonaux strictement positifs telle que $M = OT$.*

Les vecteurs colonnes e_1, \dots, e_n de M forment une base \mathcal{B} de \mathbf{R}^n à laquelle on peut appliquer le procédé de Schmidt pour obtenir une base \mathcal{B}' orthonormée. Soit \mathcal{C} la base canonique de \mathbf{R}^n ; la matrice de passage $P_{\mathcal{C} \rightarrow \mathcal{B}}$ est M ; la matrice de passage $P_{\mathcal{B} \rightarrow \mathcal{B}'}$

est triangulaire supérieure la démonstration ci-dessus montre que ses coefficients diagonaux sont strictement positifs ; la matrice de passage $P_{\mathcal{C} \rightarrow \mathcal{B}'}$ est orthogonale (cor. 44.6). Comme

$$P_{\mathcal{C} \rightarrow \mathcal{B}} = P_{\mathcal{C} \rightarrow \mathcal{B}'} P_{\mathcal{B}' \rightarrow \mathcal{B}} = P_{\mathcal{C} \rightarrow \mathcal{B}'} P_{\mathcal{B} \rightarrow \mathcal{B}'}^{-1} ,$$

et que $P_{\mathcal{B} \rightarrow \mathcal{B}'}$ est aussi triangulaire supérieure à coefficients diagonaux strictement positifs, le corollaire en résulte. ■

Exemple 44.10. Appliquons le procédé de Schmidt à la base $((1, 1, 1), (1, 1, 0), (1, 0, 0))$ de l'espace euclidien \mathbf{R}^3 . On obtient successivement :

$$e_1'' = e_1 = (1, 1, 1)$$

$$e_2'' = e_2 - \frac{\langle e_2, e_1'' \rangle}{\|e_1''\|^2} e_1'' = (1, 1, 0) - \frac{2}{3}(1, 1, 1) = \left(\frac{1}{3}, \frac{1}{3}, -\frac{2}{3}\right)$$

$$e_3'' = e_3 - \frac{\langle e_3, e_1'' \rangle}{\|e_1''\|^2} e_1'' - \frac{\langle e_3, e_2'' \rangle}{\|e_2''\|^2} e_2'' = (1, 0, 0) - \frac{1}{3}(1, 1, 1) - \frac{1}{2}\left(\frac{1}{3}, \frac{1}{3}, -\frac{2}{3}\right) = \left(\frac{1}{2}, -\frac{1}{2}, 0\right) ;$$

on en déduit

$$e_1' = \frac{e_1''}{\|e_1''\|} = \left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right)$$

$$e_2' = \frac{e_2''}{\|e_2''\|} = \left(\frac{1}{\sqrt{6}}, \frac{1}{\sqrt{6}}, -\frac{2}{\sqrt{6}}\right)$$

$$e_3' = \frac{e_3''}{\|e_3''\|} = \left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}, 0\right) .$$

Il est clair que la base obtenue par ce procédé dépend de l'ordre dans lequel on prend les vecteurs de la base d'origine : si on était parti de la base $((1, 0, 0), (1, 1, 0), (1, 1, 1))$, on aurait obtenu la base orthonormée $((1, 0, 0), (0, 1, 0), (0, 0, 1))$.

Le corollaire se traduit de la façon suivante pour la matrice $M = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$: la

matrice $O = \begin{pmatrix} 1/\sqrt{3} & 1/\sqrt{6} & 1/\sqrt{2} \\ 1/\sqrt{3} & 1/\sqrt{6} & -1/\sqrt{2} \\ 1/\sqrt{3} & -2/\sqrt{6} & 0 \end{pmatrix}$ est orthogonale. On a $T = O^{-1}M = {}^tOM$, d'où

$$\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 1/\sqrt{3} & 1/\sqrt{6} & 1/\sqrt{2} \\ 1/\sqrt{3} & 1/\sqrt{6} & -1/\sqrt{2} \\ 1/\sqrt{3} & -2/\sqrt{6} & 0 \end{pmatrix} \begin{pmatrix} \sqrt{3} & 2/\sqrt{3} & 1/\sqrt{3} \\ 0 & 2/\sqrt{6} & 1/\sqrt{6} \\ 0 & 0 & 1/\sqrt{2} \end{pmatrix} .$$

45. Adjoint d'un endomorphisme, endomorphismes symétriques

Proposition 45.1.— Soit u un endomorphisme d'un espace vectoriel euclidien E . Il existe un unique endomorphisme u^* de E qui vérifie, pour tous x et y dans E ,

$$\langle u(x), y \rangle = \langle x, u^*(y) \rangle .$$

On l'appelle l'adjoint de u . Il a même rang que u .

Notons pour une fois S le produit scalaire de E . Avec les notations de (35.5), l'égalité demandée est équivalente à $\widehat{S}(y)(u(x)) = \widehat{S}(u^*(y))(x)$, soit encore à ${}^t u \circ \widehat{S} = \widehat{S} \circ u^*$. Comme S est non dégénéré, \widehat{S} est bijectif, et $u^* = \widehat{S}^{-1} \circ {}^t u \circ \widehat{S}$ est l'adjoint de u . Il a en particulier même rang. ■

(45.2) On remarquera par exemple que pour qu'un endomorphisme u de E soit une isométrie, il faut et il suffit que $u \circ u^* = u^{-1}$.

Proposition 45.3.— Soient \mathcal{B} une base orthonormée d'un espace vectoriel euclidien E et u un endomorphisme de E ; on a

$$M_{\mathcal{B}}(u^*) = {}^t M_{\mathcal{B}}(u) .$$

Si $\mathcal{B} = (e_1, \dots, e_n)$, on écrit $u(e_j) = \sum_{i=1}^n a_{ij} e_i$. Par la proposition 44.3, on a $a_{ij} = \langle u(e_j), e_i \rangle = \langle e_j, u^*(e_i) \rangle$. Par la même proposition, cela signifie $u^*(e_i) = \sum_{j=1}^n a_{ij} e_j$, d'où la proposition. ■

On montrera en exercice les identités suivantes

$$(u \circ v)^* = v^* \circ u^* \quad (u^*)^* = u .$$

Définition 45.4.— Soit u un endomorphisme d'un espace vectoriel euclidien E . On dit que u est symétrique si $u = u^*$. En d'autres termes, on a, pour tous x et y dans E ,

$$\langle u(x), y \rangle = \langle x, u(y) \rangle .$$

Par la proposition 45.3, les endomorphismes symétriques sont ceux dont la matrice dans une base orthonormée est symétrique. Si n est la dimension de E , ils forment un sous-espace vectoriel de dimension $n(n+1)/2$ de l'espace vectoriel des endomorphismes de E . Nous allons maintenant voir deux exemples importants d'endomorphismes symétriques d'un espace vectoriel euclidien.

Soit F un sous-espace vectoriel d'un espace vectoriel euclidien E ; on rappelle que $E = F \oplus F^\perp$ (th. 44.1), de sorte que tout élément x de E s'écrit de façon unique $x = a + b$, avec $a \in F$ et $b \in F^\perp$. On définit donc une application $p_F : E \rightarrow E$ en posant $p_F(x) = a$. On pourra remarquer que $p_F + p_{F^\perp} = \text{Id}_E$.

Théorème 45.5.— L'application p_F ainsi définie est linéaire et vérifie $p_F \circ p_F = p_F$. Son noyau est F^\perp et son image F . C'est un endomorphisme symétrique de E que l'on appelle la projection orthogonale sur F . Pour tout x dans E et tout y dans F , on a

$$\|x - p_F(x)\| \leq \|x - y\| .$$

En d'autres termes, la distance de x à $p_F(x)$ réalise la plus petite distance (cf. (44.5)) de x à un point de F . On la note $d(x, F)$.

Soient x et x' des points de E et λ un réel. On écrit $x = a + b$ et $x' = a' + b'$, avec $a, b \in F$ et $a', b' \in F^\perp$. On a alors $x + \lambda x' = (a + \lambda a') + (b + \lambda b')$, avec $a + \lambda a' \in F$ et $b + \lambda b' \in F^\perp$, d'où par définition de p_F :

$$p_F(x + \lambda x') = a + \lambda a' = p_F(x) + \lambda p_F(x').$$

Ceci prouve que p_F est linéaire. Il est clair que son noyau est F^\perp et que $p_F(x) = x$ si x est dans F ; en particulier, son image est F , et $p_F \circ p_F = p_F$. On a

$$\langle p_F(x), x' \rangle = \langle a, x' \rangle = \langle a, a' \rangle$$

et

$$\langle x, p_F(x') \rangle = \langle x, a' \rangle = \langle a, a' \rangle,$$

ce qui prouve que p_F est symétrique. Soit enfin y un point de F ; on a

$$\|x - y\|^2 = \|(a - y) + b\|^2 = \|a - y\|^2 + \|b\|^2 \geq \|a - y\|^2 = \|x - p_F(x)\|^2,$$

ce qui termine la démonstration. ■

Soit (e_1, \dots, e_p) une base orthonormée de F . Pour tout x dans E , on a

$$p_F(x) = \langle x, e_1 \rangle e_1 + \dots + \langle x, e_p \rangle e_p.$$

En effet, soit a le vecteur de F ainsi défini; on a $\langle x - a, e_j \rangle = 0$ pour $j = 1, \dots, p$, de sorte que $x - a \in F^\perp$. Si la base (e_1, \dots, e_p) est simplement orthogonale, la formule devient

$$(45.6) \quad p_F(x) = \langle x, e_1 \rangle \frac{e_1}{\|e_1\|^2} + \dots + \langle x, e_p \rangle \frac{e_p}{\|e_p\|^2}.$$

On reconnaît les formules qui ont apparu dans le procédé de Schmidt (th. 44.8). Si on y regarde de plus près, on s'aperçoit (avec les mêmes notations que dans la démonstration de ce théorème), que $e_j'' = e_j - p_{F_j}(e_j)$, où F_j est le sous-espace vectoriel de E engendré par e_1, \dots, e_j (ou par e_1'', \dots, e_j''), ou encore $e_j'' = p_{F_j^\perp}(e_j)$ (qui est bien orthogonal à F_j , comme on le souhaite!).

Définition 45.7.— Soit s_F l'endomorphisme de E défini par $s_F(x) = 2p_F(x) - x$. C'est un automorphisme symétrique et orthogonal de E qui vérifie $s_F \circ s_F = \text{Id}_E$; on l'appelle la symétrie orthogonale par rapport à F . Lorsque F est un hyperplan de E , on dit que s_F est une réflexion.

L'endomorphisme s_F est symétrique car combinaison linéaire d'endomorphismes symétriques. On a

$$s_F \circ s_F = (2p_F - \text{Id}_E) \circ (2p_F - \text{Id}_E) = 4p_F \circ p_F - 4p_F + \text{Id}_E = \text{Id}_E .$$

Enfin, comme s_F est symétrique, on a $s_F \circ s_F^* = s_F \circ s_F = \text{Id}_E$, de sorte que s_F est une isométrie par (45.2). D'autre part, on remarquera que $s_{F^\perp} = -s_F$.

Du point de vue géométrique, $s_F(x)$ est caractérisé par les deux propriétés suivantes : $x - s_F(x)$ est orthogonal à F , et le milieu du segment joignant x à $s_F(x)$ est sur F .

Exemples 45.8. 1) Soit F la droite dirigée par le vecteur $e = (1, 2)$ dans l'espace vectoriel euclidien \mathbf{R}^2 . On a par (45.6)

$$p_F((1, 0)) = \langle (1, 0), e \rangle \frac{e}{\|e\|^2} = \frac{1}{5} e \quad p_F((0, 1)) = \langle (0, 1), e \rangle \frac{e}{\|e\|^2} = \frac{2}{5} e$$

et la matrice de p_F dans la base canonique est $\begin{pmatrix} 1/5 & 2/5 \\ 2/5 & 4/5 \end{pmatrix}$. Celle de s_F est

$$2 \begin{pmatrix} 1/5 & 2/5 \\ 2/5 & 4/5 \end{pmatrix} - \text{I}_2 = \begin{pmatrix} -3/5 & 4/5 \\ 4/5 & 3/5 \end{pmatrix} .$$

Son déterminant est -1 ; ce n'est donc pas un élément de $\text{SO}(2)$ (*cf.* prop. 41.5.2).

2) Soit F le plan d'équation $3x - y + 2z = 0$ dans \mathbf{R}^3 . Pour trouver la matrice de p_F , il est plus simple de commencer par celle de p_{F^\perp} ; la droite F^\perp est dirigée par le vecteur $e = (3, -1, 2)$. La formule (45.6) donne

$$p_{F^\perp}((1, 0, 0)) = \frac{3}{14} e \quad p_{F^\perp}((0, 1, 0)) = \frac{-1}{14} e \quad p_{F^\perp}((0, 0, 1)) = \frac{2}{14} e .$$

La matrice de p_F dans la base canonique est

$$\text{I}_3 - \begin{pmatrix} 9/14 & -3/14 & 6/14 \\ -3/14 & 1/14 & -2/14 \\ 6/14 & -2/14 & 4/14 \end{pmatrix} = \begin{pmatrix} 5/14 & 3/14 & -3/7 \\ 3/14 & 13/14 & 1/7 \\ -3/7 & 1/7 & 5/7 \end{pmatrix} ,$$

celle de s_F

$$\begin{pmatrix} -2/7 & 3/7 & -6/7 \\ 3/7 & 6/7 & 2/7 \\ -6/7 & 2/7 & 3/7 \end{pmatrix} .$$

Revenons à la situation générale. On remarque que p_F est *diagonalisable*, puisque E est somme directe de son espace propre pour la valeur propre 0 , c'est-à-dire son noyau F^\perp , et de son espace propre pour la valeur propre 1 , c'est-à-dire F . De la même façon, s_F est diagonalisable : F est son espace propre pour la valeur propre 1 , et F^\perp est son espace propre pour la valeur propre -1 (on peut d'ailleurs montrer que les symétries orthogonales sont les seuls endomorphismes orthogonaux diagonalisables (exerc. 52.20)). C'est une propriété commune à tous les endomorphismes symétriques, comme le montre le théorème suivant.

Théorème 45.9.— *Tout endomorphisme symétrique d'un espace vectoriel euclidien est diagonalisable dans une base orthonormée.*

On raisonne par récurrence sur la dimension de l'espace. Soit u un endomorphisme symétrique d'un espace vectoriel euclidien E . Soit \mathcal{B} une base orthonormée de E ; la matrice M de u dans la base \mathcal{B} est symétrique (prop. 45.3). Soient λ une valeur propre (complexe) de M et X un vecteur colonne (complexe) non nul tel que $MX = \lambda X$. Puisque M est réelle, on a $M\bar{X} = \bar{\lambda}\bar{X}$ et

$$\bar{\lambda}'X\bar{X} = {}'XM\bar{X} = {}'({}'XM\bar{X}) = \bar{\lambda}'\bar{X}MX = \lambda'{}'\bar{X}X$$

Si $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$, on a $'X\bar{X} = {}'\bar{X}X = \sum_{j=1}^n |x_j|^2 \neq 0$, de sorte que $\bar{\lambda} = \lambda$. Il s'ensuit que λ est réel.

L'endomorphisme $(u - \lambda \text{Id}_E)$ n'est donc pas inversible; en d'autres termes, λ est valeur propre de u . Soient x un vecteur propre associé, F la droite vectorielle engendrée par x , et y un vecteur orthogonal à x . On a

$$\langle u(y), x \rangle = \langle y, u(x) \rangle = \lambda \langle y, x \rangle = 0,$$

de sorte que $u(y) \in F^\perp$. Ceci montre que F^\perp est stable par u ; l'espace vectoriel F^\perp est euclidien et la restriction u' de u à cet espace est symétrique. L'hypothèse de récurrence entraîne que u' est diagonalisable dans une base orthonormée (e_2, \dots, e_n) de F^\perp . Il est clair que u est diagonalisable dans la base orthonormée $(x/\|x\|, e_2, \dots, e_n)$ de E . ■

(45.10) Soit u un endomorphisme symétrique d'un espace vectoriel euclidien E . Le théorème montre que les espaces propres de u sont deux à deux orthogonaux et que E est leur somme directe (on dit parfois que E est somme directe orthogonale des espaces propres).

Corollaire 45.11.— *Soit M une matrice symétrique réelle. Il existe une matrice orthogonale réelle P et une matrice diagonale réelle D telle que $M = PDP^{-1}$.*

L'endomorphisme u de \mathbf{R}^n de matrice M dans la base canonique \mathcal{C} est symétrique par la proposition 45.3. Il est diagonalisable dans une base orthonormée \mathcal{B} de \mathbf{R}^n . La matrice de passage $P_{\mathcal{C} \rightarrow \mathcal{B}}$ est orthogonale (cor. 44.6), et $P_{\mathcal{C} \rightarrow \mathcal{B}}^{-1}MP_{\mathcal{C} \rightarrow \mathcal{B}}$ est diagonale. ■

Exemple 45.12. La matrice $M = \begin{pmatrix} 1 & 0 & 2 \\ 0 & 3 & 0 \\ 2 & 0 & 1 \end{pmatrix}$ est symétrique; c'est donc la matrice d'un endomorphisme symétrique de l'espace vectoriel euclidien \mathbf{R}^3 . Son polynôme caractéristique est $(T+1)(T-3)^2$; l'espace propre E_{-1} est dirigé par le vecteur unitaire $(1/\sqrt{2}, 0, -1/\sqrt{2})$. L'espace propre E_3 est par (45.10) l'orthogonal de E_{-1} , c'est-à-dire le plan d'équation $x - z = 0$. Il reste à trouver une base orthonormée de E_3 . Il est facile de voir

que $((1/\sqrt{2}, 0, 1/\sqrt{2}), (0, 1, 0))$ convient (en général, on peut partir d'une base quelconque et lui appliquer le procédé de Schmidt). On obtient donc

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

46. Isométries du plan

Dans toute cette section, E est un plan vectoriel euclidien, c'est-à-dire un espace vectoriel euclidien de dimension 2. On va identifier les isométries de E à l'aide de leur matrice $M = \begin{pmatrix} a & c \\ b & d \end{pmatrix}$ dans une base orthonormée. Rappelons que le déterminant de M vaut $\delta = \pm 1$ (prop. 41.5.2) et que

$$M = {}^t M^{-1} = \frac{1}{\delta} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix},$$

ce qui se traduit par les égalités $d = \delta a$, $c = -\delta b$ et $\delta = ad - bc = \delta a^2 + \delta b^2$.

Les matrices orthogonales de déterminant -1 sont donc de la forme

$$M = \begin{pmatrix} a & b \\ b & -a \end{pmatrix} \quad \text{avec } a^2 + b^2 = 1.$$

Une telle matrice vérifie $M^2 = I_2$; ses valeurs propres sont 1 et -1 , les vecteurs propres associés e_1 et e_{-1} sont orthogonaux par le théorème 45.9, de sorte que M est la matrice de la réflexion par rapport à la droite engendrée par e_1 . *Les isométries indirectes d'un plan euclidien sont les réflexions.* Il en résulte que toute isométrie u d'un plan euclidien est produit d'au plus deux réflexions : si u est une réflexion, c'est évident; sinon, pour toute réflexion s , le produit $s \circ u$ est une isométrie indirecte, donc une réflexion, et $u = s \circ (s \circ u)$. On notera s_D la symétrie orthogonale par rapport à une droite D .

Les matrices orthogonales de déterminant 1 (c'est-à-dire les éléments de $SO(2)$) sont de la forme

$$(46.1) \quad M = \begin{pmatrix} a & -b \\ b & a \end{pmatrix} \quad \text{avec } a^2 + b^2 = 1.$$

Proposition 46.2.— *L'application qui à un nombre complexe $a + ib$ de module 1 associe la matrice $\begin{pmatrix} a & -b \\ b & a \end{pmatrix}$ définit un isomorphisme du groupe U des nombres complexes de module 1 sur le groupe $SO(2)$. Celui-ci est en particulier un groupe commutatif.*

Il faut vérifier que l'application ϕ ainsi définie envoie zz' sur $\phi(z)\phi(z')$. On écrit $z = a + ib$ et $z' = a' + ib'$; on a alors

$$\begin{pmatrix} a & -b \\ b & a \end{pmatrix} \begin{pmatrix} a' & -b' \\ b' & a' \end{pmatrix} = \begin{pmatrix} aa' - bb' & -ab' - a'b \\ ab' + a'b & aa' - bb' \end{pmatrix},$$

tandis que

$$zz' = (a + ib)(a' + ib') = (aa' - bb') + i(ab' + a'b),$$

ce qui démontre la proposition. ■

On appelle *rotations* les isométries directes d'un plan euclidien. Elles forment un groupe commutatif isomorphe à U .

Si u est une rotation, sa matrice M dans une base orthonormée \mathcal{B} est dans $SO(2)$, donc du type (46.1). Le coefficient a est indépendant du choix de \mathcal{B} puisque c'est la moitié de la trace de u . Le coefficient b est déterminé au signe près par la relation $a^2 + b^2 = 1$. Si P est la matrice de passage de \mathcal{B} à une base orthonormée \mathcal{B}' , la matrice P est elle-même orthogonale. De deux choses l'une : soit $\det(P) = 1$ et $PM = MP$ par la proposition, de sorte que $M_{\mathcal{B}'} = P^{-1}MP = M$, soit $\det(P) = -1$ et $\det(PM) = -1$, de sorte que $P^2 = (PM)^2 = I_2$ et

$$P^{-1}MP = PMP = (PM)^2M^{-1} = M^{-1} = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}.$$

Il est indispensable ici d'introduire la notion *d'orientation* d'un espace vectoriel. Se donner une orientation d'un espace vectoriel E , c'est se donner une partition de l'ensemble des bases de E en deux sous-ensembles : les bases « directes » et les bases « indirectes », de façon que

- a) si \mathcal{B} est directe et que $\det_{\mathcal{B}} \mathcal{B}' > 0$, alors \mathcal{B}' est directe ;
- b) si \mathcal{B} est directe et que $\det_{\mathcal{B}} \mathcal{B}' < 0$, alors \mathcal{B}' est indirecte.

En termes plus savants, on dit que des bases \mathcal{B} et \mathcal{B}' sont en relation si $\det_{\mathcal{B}} \mathcal{B}' > 0$; c'est une relation d'équivalence sur l'ensemble des bases de E et il y a deux classes d'équivalence. Orienter E , c'est choisir une de ces classes. Il y a donc deux orientations possibles. Si on permute deux vecteurs d'une base directe, on obtient une base indirecte.

On oriente canoniquement \mathbf{R}^n en décrétant que la base canonique est directe.

Revenons à notre plan euclidien. Orientons-le; il résulte de ce qui précède que la matrice d'une rotation dans une base *directe* de E est indépendante du choix de cette base. Il existe un réel θ , uniquement déterminé modulo $2\pi\mathbf{Z}$, tel qu'elle s'écrive

$$(46.3) \quad \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}.$$

On l'appelle θ *l'angle* de la rotation (on devrait plutôt dire la « mesure de l'angle de la rotation » !); il vérifie $\text{tr}(u) = 2 \cos \theta$ (on remarquera que le membre de droite de cette égalité ne dépend pas de l'orientation choisie). Pour tout réel θ , on notera r_θ la rotation d'angle θ . On a les relations (*cf.* prop. 46.2)

$$(46.4) \quad r_\theta \circ r_\theta = r_{\theta+\theta'} \quad r_0 = \text{Id}_E \quad r_\pi = s_{\{0\}} = -\text{Id}_E.$$

Nous allons maintenant définir la mesure de l'angle orienté de vecteurs *non nuls* du plan euclidien orienté.

Lemme 46.5.— Soient x et y des vecteurs non nuls de même norme du plan vectoriel euclidien E . Il existe une unique rotation r telle que $r(x) = y$.

Soit e_1 le vecteur unitaire $x/\|x\|$; il existe un vecteur e_2 tel que $\mathcal{B} = (e_1, e_2)$ soit une base orthonormée directe. On écrit $y/\|y\| = ae_1 + be_2$, avec $a^2 + b^2 = 1$. Soit r la rotation de matrice $\begin{pmatrix} a & -b \\ b & a \end{pmatrix}$ dans la base \mathcal{B} . On a $r(x/\|x\|) = y/\|y\|$, d'où $r(x) = y$ puisque $\|x\| = \|y\|$. Si r' est une autre rotation telle que $r'(x) = y$, on a $r'^{-1} \circ r(x) = x$: 1 est valeur propre de la rotation $r'^{-1} \circ r$, qui est donc l'identité (le produit des valeurs propres de $r'^{-1} \circ r$ est son déterminant, c'est-à-dire 1); on a bien $r = r'$. ■

(46.6) Soient x et y des vecteurs *non nuls* du plan euclidien orienté E . Il existe par le lemme une unique rotation r qui envoie $x/\|x\|$ sur $y/\|y\|$; on définit (la mesure de) l'angle (orienté) $\sphericalangle(x, y)$ entre x et y comme celui de r ; c'est un réel défini modulo $2\pi\mathbf{Z}$. Il ressort de la démonstration du lemme que

$$\cos \sphericalangle(x, y) = \frac{\langle x, y \rangle}{\|x\| \cdot \|y\|} \qquad \sin \sphericalangle(x, y) = \frac{\det_{\mathcal{B}}(x, y)}{\|x\| \cdot \|y\|},$$

où \mathcal{B} est n'importe quelle base orthonormée *directe* du plan. Si $-\pi < \sphericalangle(x, y) \leq \pi$, l'angle non orienté entre x et y défini en (43.4) est donc $|\sphericalangle(x, y)|$. Si r est une rotation, on a $\sphericalangle(r(x), r(y)) = \sphericalangle(x, y)$; si s est une réflexion, on a $\sphericalangle(s(x), s(y)) = -\sphericalangle(x, y)$.

La formule (46.4) entraîne

$$\sphericalangle(x, y) + \sphericalangle(y, z) = \sphericalangle(x, z) \qquad \sphericalangle(y, x) = \sphericalangle(x, -y) \qquad \sphericalangle(x, -y) = \sphericalangle(x, y) + \pi.$$

Soient D (resp. D') une droite de E et e (resp. e') un vecteur unitaire sur D (resp. sur D'). Il résulte de la dernière relation que $\sphericalangle(e, e')$ *modulo* $\pi\mathbf{Z}$, ne dépend que de D et D' , pas du choix des vecteurs unitaires e et e' . On l'appelle (la mesure de) l'angle (orienté) entre D et D' ; on le note $\sphericalangle(D, D')$.

Proposition 46.7.— Soient D et D' des droites d'un plan vectoriel euclidien orienté. On a

$$s_{D'} \circ s_D = r_{2\sphericalangle(D, D')}.$$

Soient e un vecteur unitaire de D et e' un vecteur unitaire de D' . On sait que $s_{D'} \circ s_D$ est une rotation; il s'agit d'en déterminer l'angle, qui vaut

$$\begin{aligned} \sphericalangle(e, s_{D'} \circ s_D(e)) &= \sphericalangle(e, s_{D'}(e)) = \sphericalangle(e, e') + \sphericalangle(e', s_{D'}(e)) \\ &= \sphericalangle(e, e') + \sphericalangle(s_{D'}(e'), s_{D'}(e)) = \sphericalangle(e, e') - \sphericalangle(e', e) \\ &= 2\sphericalangle(e, e'), \end{aligned}$$

ce qui prouve la proposition. ■

47. Isométries en dimension 3

Dans toute cette section, E est un espace vectoriel euclidien de dimension 3. Contrairement au cas de la dimension 2, il n'existe pas de forme simple pour les matrices des isométries de E . Nous allons utiliser des arguments plus géométriques.

Théorème 47.1.— *Soit u une isométrie d'un espace vectoriel euclidien de dimension 3. Il existe une base orthonormée de E dans laquelle la matrice de u est*

$$\begin{pmatrix} \delta & 0 & 0 \\ 0 & a & -b \\ 0 & b & a \end{pmatrix},$$

où $\delta = \pm 1$ est le déterminant de u .

Le polynôme caractéristique de u est de degré 3, donc admet une racine réelle λ , qui est donc une valeur propre de u . Soit x un vecteur propre associé. On a

$$\|x\| = \|u(x)\| = \|\lambda x\| = |\lambda| \|x\|,$$

de sorte que $\lambda = \pm 1$.

Supposons tout d'abord que 1 soit valeur propre : il existe un vecteur non nul x , que l'on peut supposer unitaire, tel que $u(x) = x$. La droite D engendrée par x est stable par u ; il en est de même par la proposition 44.7.b) du plan $P = D^\perp$. Les résultats du § 46 entraînent que la matrice de la restriction de u à P est, dans une base orthonormale convenable (e_1, e_2) de P ,

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad \text{si c'est une réflexion, c'est-à-dire si } \delta = -1,$$

$$\begin{pmatrix} a & -b \\ b & a \end{pmatrix} \quad \text{si c'est une rotation, c'est-à-dire si } \delta = 1.$$

Dans le premier cas, la matrice de u dans la base orthonormée (e_1, e_2, x) est

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

qui est de la forme demandée. Dans le second cas, la matrice de u dans la base orthonormée (x, e_1, e_2) est

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & a & -b \\ 0 & b & a \end{pmatrix}.$$

Si -1 est valeur propre de u , alors 1 est valeur propre de $-u$, et l'on peut appliquer à l'isométrie $-u$ (dont le déterminant est $-\delta$) ce qui précède; ceci termine la démonstration du théorème. ■

De nouveau, a ne dépend que de u , puisque $\text{tr}(u) = 2a + \det(u)$, mais b n'est déterminé qu'au signe près (il suffit de changer le deuxième vecteur de la base en son opposé pour changer le signe de b).

Isométries directes. Soit u une isométrie directe de E . On remarque que $-1 \leq \text{tr}(u) \leq 3$. Si sa trace est 3, il résulte du théorème que u est l'identité. Si $\text{tr}(u) < 3$, l'espace propre E_1 est de dimension 1 et la restriction de u à son orthogonal est une rotation plane; on dit que u est une *rotation*, et l'on appelle E_1 son *axe*. On peut définir l'angle orienté θ de u si l'on a orienté $(E_1)^\perp$; il vérifie $\text{tr}(u) = 2 \cos \theta + 1$ (on remarquera que le membre de droite de cette égalité ne dépend pas de l'orientation choisie). Orienter E ne suffit pas; il faut encore orienter l'axe en choisissant un vecteur non nul e dessus. On convient alors de dire qu'une base (e_1, e_2) de $(E_1)^\perp$ est directe si la base (e, e_1, e_2) de E est directe.

Un cas particulier important est celui des rotations d'angle π . Leur matrice dans une base convenable est

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix};$$

il s'agit donc des symétries orthogonales par rapport à des droites. Elles sont caractérisées par les égalités $\det(u) = 1$, $\text{tr}(u) = -1$.

(47.2) Enfin, on a vu dans § 46 (prop. 46.7) que toute rotation plane est produit de deux réflexions planes. Cela entraîne immédiatement que toute rotation r de l'espace est produit de deux réflexions; plus exactement, si P est un plan contenant l'axe de r , la composée $s_P \circ r$ est une réflexion par rapport à un plan P' contenant D , de sorte que $r = s_P \circ s_{P'}$.

Isométries indirectes. Soit u une isométrie indirecte de E . Les égalités

$$\begin{pmatrix} -1 & 0 & 0 \\ 0 & a & -b \\ 0 & b & a \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & a & -b \\ 0 & b & a \end{pmatrix} \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & a & -b \\ 0 & b & a \end{pmatrix}$$

montrent avec le théorème 47.1 que u est le produit (commutatif) d'une rotation d'axe D et d'une réflexion par rapport à D^\perp . En particulier, u est le produit de 3 réflexions.

Produit vectoriel. Soient x et y des vecteurs d'un espace vectoriel euclidien orienté E de dimension 3 et \mathcal{B} une base orthonormée directe de E ; l'application

$$z \mapsto \det_{\mathcal{B}}(x, y, z)$$

est une forme linéaire f sur E qui ne dépend pas du choix de \mathcal{B} , puisque, pour toute autre base orthonormée directe \mathcal{B}' de E , on a $\det_{\mathcal{B}'} \mathcal{B} = 1$. Notons provisoirement S le produit scalaire; l'application $\hat{S}: E \rightarrow E^*$ associée (cf. (35.5)) est bijective. L'élément $(\hat{S})^{-1}(f)$ de E , que l'on notera $x \wedge y$, est l'unique vecteur qui vérifie, pour tout z dans E ,

$$(47.3) \quad \det_{\mathcal{B}}(x, y, z) = \langle x \wedge y, z \rangle.$$

Proposition 47.4.— Soit E un espace vectoriel euclidien orienté de dimension 3.

a) L'application

$$\begin{aligned} E \times E &\longrightarrow E \\ (x, y) &\longmapsto x \wedge y \end{aligned}$$

est bilinéaire alternée.

b) Le vecteur $x \wedge y$ est orthogonal à x et à y ; pour qu'il soit nul, il faut et il suffit que x et y soient colinéaires. Si x et y ne sont pas colinéaires, $(x, y, x \wedge y)$ est une base directe de E .

c) Si θ est l'angle non orienté de x et y , on a

$$\|x \wedge y\| = \|x\| \|y\| |\sin \theta| .$$

Le a) résulte du fait que le déterminant est bilinéaire alterné en ses deux premières variables : pour tout z , on a par la formule (47.3)

$$\langle y \wedge x, z \rangle = \det_{\mathcal{B}}(y, x, z) = -\det_{\mathcal{B}}(x, y, z) = -\langle x \wedge y, z \rangle ,$$

de sorte que $y \wedge x = -x \wedge y$. La même formule entraîne

$$\langle x \wedge y, x \rangle = \det_{\mathcal{B}}(x, y, x) = 0 \quad \text{et} \quad \langle x \wedge y, y \rangle = \det_{\mathcal{B}}(x, y, y) = 0 ,$$

de sorte que $x \wedge y$ est orthogonal à x et à y .

Si x et y ne sont pas colinéaires, on peut compléter (x, y) en une base (x, y, z) de E et la formule (47.3) montre que $\langle x \wedge y, z \rangle$, donc aussi $x \wedge y$, ne sont pas nuls. On a aussi

$$\det_{\mathcal{B}}(x, y, x \wedge y) = \|x \wedge y\|^2 ,$$

ce qui montre d'une part que si x et y sont colinéaires, $x \wedge y = 0$, d'autre part que si x et y ne sont pas colinéaires, la base $(x, y, x \wedge y)$ est directe. Cela montre b).

Pour c), supposons d'abord x et y unitaires et orthogonaux et complétons (x, y) en une base orthonormée directe $\mathcal{B} = (x, y, e)$. Pour tout z , $\det_{\mathcal{B}}(x, y, z)$ est alors la composante de z sur e , c'est-à-dire $\langle e, z \rangle$ (prop. 44.3), de sorte que $x \wedge y = e$; c'est un vecteur de norme 1, ce qui démontre c) dans ce cas. Le cas où x et y sont seulement orthogonaux s'en déduit par bilinéarité du produit vectoriel. Supposons maintenant x et y seulement unitaires. Les vecteurs x et $x\langle x, y \rangle - y$ sont orthogonaux, donc par le cas déjà traité,

$$\|x \wedge y\| = \|x \wedge (x\langle x, y \rangle - y)\| = \|x\| \|x\langle x, y \rangle - y\| = \|x\langle x, y \rangle - y\|$$

et

$$\|x\langle x, y \rangle - y\|^2 = \langle x, y \rangle^2 - 2\langle x, y \rangle^2 + 1 = 1 - \langle x, y \rangle^2 = 1 - \cos^2 \theta = \sin^2 \theta$$

par la définition même de θ (cf. (43.4)). Ceci démontre c) dans ce cas, et le cas général s'en déduit par bilinéarité du produit vectoriel. ■

Enfin, si (x_1, x_2, x_3) sont les coordonnées de x dans une base orthonormée directe \mathcal{B} , et (y_1, y_2, y_3) celles de y , celles de $x \wedge y$ sont

$$(47.5) \quad \left(\begin{vmatrix} x_2 & y_2 \\ x_3 & y_3 \end{vmatrix}, - \begin{vmatrix} x_1 & y_1 \\ x_3 & y_3 \end{vmatrix}, \begin{vmatrix} x_1 & y_1 \\ x_2 & y_2 \end{vmatrix} \right)$$

comme le montre le développement du déterminant

$$\det_{\mathcal{B}}(x, y, z) = \begin{vmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ x_3 & y_3 & z_3 \end{vmatrix}$$

par rapport à la dernière colonne.

Remarques 47.6. 1) Soient r une rotation et r_D une rotation d'axe D . La composée $r \circ r_D \circ r^{-1}$ est encore une isométrie directe, donc une rotation. Soit x un vecteur de D ; on a

$$r \circ r_D \circ r^{-1}(r(x)) = r(r_D(x)) = r(x),$$

de sorte que $r(x)$ est invariant par la rotation $r \circ r_D \circ r^{-1}$, dont l'axe est donc $r(D)$. Comme $r \circ r_D \circ r^{-1}$ a même trace que r_D , c'est une rotation d'axe $r(D)$ et de même angle que r_D . En particulier, ce n'est en général pas r_D , et $\text{SO}(3)$ n'est pas commutatif.

2) Soient D et D' des droites distinctes, r une rotation d'axe D et r' une rotation d'axe D' . Etudions la rotation $r' \circ r$. Soit P'' le plan contenant D et D' ; par (47.2), on peut écrire $r = s_{P''} \circ s_P$ et $r' = s_{P'} \circ s_{P''}$, où P (resp. P') est un plan contenant D (resp. D'). On a

$$r' \circ r = (s_{P'} \circ s_{P''}) \circ (s_{P''} \circ s_P) = s_{P'} \circ s_P$$

et l'axe de $r' \circ r$ est donc la droite $P \cap P'$.

3) Si P est un plan, $-s_P$ est un demi-tour d'axe P^\perp . Il ressort de (47.2) que *toute rotation est produit de deux demi-tours*.

4) Soit r une rotation d'axe D orienté par un vecteur directeur e et d'angle θ . Soit y un vecteur non nul orthogonal à D ; posons $y' = y/\|y\|$ et $e' = e/\|e\|$. La proposition 47.4.c) entraîne que $(e', y', e' \wedge y')$ est une base orthonormée directe de E ; la forme de la matrice de r dans cette base montre

$$\sin \theta = \langle e' \wedge y', r(y') \rangle = \det_{\mathcal{B}}(e', y', r(y')) = \frac{1}{\|y\|^2 \|e\|} \det_{\mathcal{B}}(e, y, r(y)),$$

où \mathcal{B} est une base orthonormée directe quelconque. Soit maintenant x un vecteur quelconque non colinéaire à e' . Le vecteur $x - \langle x, e' \rangle e'$ est non nul et orthogonal à e ; la formule ci-dessus montre que $\sin \theta$ est du signe de

$$\det_{\mathcal{B}}(e, x - \langle x, e' \rangle e', r(x - \langle x, e' \rangle e')) = \det_{\mathcal{B}}(e, x, r(x) - \langle x, e' \rangle e') = \det_{\mathcal{B}}(e, x, r(x)),$$

donc aussi du signe de $\det_{\mathcal{B}'}(e, x, r(x))$, où \mathcal{B}' est une base directe quelconque.

Exemples 47.7. 1) La matrice $M = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$ est orthogonale (elle vérifie ${}^tMM = I_3$)

de déterminant 1. C'est donc la matrice dans la base canonique d'une rotation r de \mathbf{R}^3 dont l'axe D est dirigé par un vecteur propre pour la valeur propre 1, comme par exemple $e = (1, 1, 1)$, dont on peut se servir pour orienter D . La trace de M est 0, donc l'angle θ de r vérifie $\cos \theta = -1/2$. Pour savoir si c'est $2\pi/3$ ou $-2\pi/3$, il suffit par la remarque 47.6.4) de calculer le signe du déterminant $\det_{\mathcal{C}}(e, x, r(x))$, où x est un vecteur non colinéaire à e , comme par exemple $x = (1, 0, 0)$. Comme $r(x) = (0, 1, 0)$ et que

$$\begin{vmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 0 \end{vmatrix} = 1,$$

$\sin \theta$ est positif et $\theta = 2\pi/3$.

2) Considérons la rotation r d'axe dirigé (et orienté) par le vecteur $e' = (36, -48, 25)$ dans \mathbf{R}^3 et d'angle $\pi/2$. Le plan $(e')^\perp$ est d'équation $36x - 48y + 25z = 0$, donc contient le vecteur $e'_1 = (4, 3, 0)$. On calcule $e'_2 = e' \wedge e'_1 = (-75, 100, 300)$ par la formule (47.5). On obtient une base orthonormée directe $\mathcal{B} = (e, e_1, e_2)$ en posant

$$e = e'/\|e'\| = (36/65, -48/65, 25/65)$$

$$e_1 = e'_1/\|e'_1\| = (4/5, 3/5, 0) \quad e_2 = e'_2/\|e'_2\| = (-3/13, 4/13, 12/13).$$

La matrice de r dans cette base est

$$M = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}.$$

La matrice de passage de la base canonique à la base orthonormée \mathcal{B} est la matrice orthogonale

$$P = \begin{pmatrix} 36/65 & 4/5 & -3/13 \\ -48/65 & 3/5 & 4/13 \\ 25/65 & 0 & 12/13 \end{pmatrix} = \frac{1}{65} \begin{pmatrix} 36 & 52 & -15 \\ -48 & 39 & 20 \\ 25 & 0 & 60 \end{pmatrix}.$$

Finalement, la matrice de r dans la base canonique est

$$PMP^{-1} = PM{}^tP = \frac{1}{4225} \begin{pmatrix} 1296 & -3353 & -2220 \\ -103 & 2304 & -3540 \\ 4020 & 1140 & 625 \end{pmatrix}.$$

On vérifie que sa trace est bien $(1296 + 2304 + 625)/4225 = 1 = 1 + 2 \cos \pi/2$.

48. Isométries en dimension quelconque

On a vu que toute isométrie d'un plan euclidien peut s'écrire comme la composée d'au plus deux réflexions (prop. 46.7). De même, toute isométrie d'un espace euclidien de dimension 3 peut s'écrire comme la composée d'au plus trois réflexions (cf. (47.2)). Nous allons généraliser ces résultats aux espaces vectoriels euclidiens de dimension (finie) quelconque.

Lemme 48.1.— Soient x et y des vecteurs distincts de même norme dans un espace vectoriel euclidien E . Si H est l'hyperplan $(x - y)^\perp$ de E , on a $s_H(x) = y$.

Comme $\langle x + y, x - y \rangle = \|x\|^2 - \|y\|^2 = 0$, on a $(x + y)/2 \in H$. Comme

$$x = (x + y)/2 + (x - y)/2,$$

avec $(x + y)/2 \in H$ et $(x - y)/2 \in H^\perp$, on a $s_H(x) = (x + y)/2 - (x - y)/2 = y$. ■

Théorème 48.2.— Toute isométrie d'un espace vectoriel euclidien de dimension n est produit d'au plus n réflexions.

On procède par récurrence sur n . Si $n = 1$, les seules isométries sont Id_E et $-\text{Id}_E$, qui est une réflexion (par rapport à 0). Supposons le théorème vrai pour les isométries des espaces vectoriels euclidiens de dimension $< n$.

Soit u une isométrie d'un espace vectoriel euclidien de dimension n . Si u est l'identité, c'est évident. Sinon, il existe un vecteur x (non nul) tel que $u(x) \neq x$. Comme u est une isométrie, x et $u(x)$ ont même norme. Le lemme entraîne qu'il existe une réflexion s telle que $s(x) = u(x)$. On a alors $s \circ u(x) = x$, ce qui entraîne (prop. 44.7.b) que l'isométrie $s \circ u$ laisse stable l'hyperplan $H = x^\perp$. L'hypothèse de récurrence entraîne que la restriction de l'isométrie $s \circ u$ à l'espace vectoriel euclidien H peut s'écrire $s_{H'_1} \circ \cdots \circ s_{H'_m}$, avec $m \leq n - 1$, où H'_1, \dots, H'_m sont des hyperplans de H . Pour $1 \leq j \leq m$, soit H_j l'hyperplan de E engendré par H'_j et le vecteur x . La composée $s_{H_1} \circ \cdots \circ s_{H_m}$ coïncide avec $s \circ u$ sur H , et aussi sur la droite D engendrée par x , puisqu'elle laisse x invariant. Comme $E = H \oplus D$, elle coïncide avec $s \circ u$ sur tout E ; on a donc $s \circ u = s_{H_1} \circ \cdots \circ s_{H_m}$, d'où $u = s \circ s_{H_1} \circ \cdots \circ s_{H_m}$, ce qui prouve le théorème. ■

Pour terminer, nous allons nous intéresser aux applications ϕ d'un espace vectoriel euclidien E dans lui-même qui préservent la distance (cf. (44.5)), c'est-à-dire qui vérifient

$$d(\phi(x), \phi(y)) = d(x, y)$$

pour tous x et y ; on dit qu'une telle application est isométrique. Les isométries (linéaires) en sont un exemple; un autre est fourni par les translations $t_a : E \rightarrow E$ (où a est un vecteur fixé dans E) définies par

$$t_a(x) = x + a.$$

Théorème 48.3.— Soient E un espace vectoriel euclidien et $\phi : E \rightarrow E$ une application isométrique. Il existe un unique vecteur a et une unique isométrie (linéaire) u tels que $\phi = t_a \circ u$.

L'unicité est facile : si $\phi = t_a \circ u = t_b \circ v$, on a en prenant l'image de 0 l'égalité

$$\phi(0) = t_a \circ u(0) = t_a(0) = a = t_b \circ v(0) = t_b(0) = b,$$

d'où $a = b = \phi(0)$, $t_a = t_b$ et $u = t_{-a} \circ \phi = v$.

Pour l'existence, il s'agit de montrer que l'application $\psi = t_{-\phi(0)} \circ \phi$, qui vérifie $\psi(0) = 0$ et préserve les distances, donc qui conserve la norme, est linéaire. Soit (e_1, \dots, e_n) une base orthonormée de E . On a pour $j \neq k$

$$\|\psi(e_j) - \psi(e_k)\| = d(\psi(e_j), \psi(e_k)) = d(e_j, e_k) = \|e_j - e_k\| = \sqrt{2},$$

ce qui montre par Pythagore que les vecteurs unitaires $\psi(e_j)$ et $\psi(e_k)$ sont orthogonaux; ils forment donc une base orthonormée $(\psi(e_1), \dots, \psi(e_n))$. Soit x un vecteur quelconque de E , que l'on écrit $x = \sum_{j=1}^n x_j e_j$. On décompose aussi $\psi(x)$ sur la base $(\psi(e_1), \dots, \psi(e_n))$ en $\psi(x) = \sum_{j=1}^n y_j \psi(e_j)$. On a alors

$$d(x, e_k)^2 = \|x - e_k\|^2 = \|x\|^2 - 2x_k + 1$$

et

$$d(\psi(x), \psi(e_k))^2 = \|\psi(x) - \psi(e_k)\|^2 = \|\psi(x)\|^2 - 2y_k + 1 = \|x\|^2 - 2y_k + 1;$$

comme $d(\psi(x), \psi(e_k)) = d(x, e_k)$, on en déduit $x_k = y_k$ pour tout k . On a donc

$$\psi\left(\sum_{j=1}^n x_j e_j\right) = \sum_{j=1}^n x_j \psi(e_j),$$

et cela entraîne que ψ est linéaire. ■

La décomposition $\phi = t_a \circ u$ n'est en fait pas « la bonne ». Par exemple, la rotation d'angle θ « autour d'un point a » d'un plan euclidien (qui peut être définie comme la composée $t_a \circ r_\theta \circ t_{-a}$) se décompose selon le théorème en $t_{a-r_\theta(a)} \circ r_\theta$ (exercice), ce qui en est une description beaucoup moins parlante. On dira qu'une application isométrique ϕ a un point fixe s'il existe a dans E tel que $\phi(a) = a$.

Théorème 48.4.— Soient E un espace vectoriel euclidien et $\phi : E \rightarrow E$ une application isométrique. Il existe un unique vecteur c et une unique application isométrique ϕ_0 avec un point fixe tels que $\phi = t_c \circ \phi_0 = \phi_0 \circ t_c$.

D'après le théorème précédent, on peut écrire $\phi = t_a \circ u$ (avec $a = \phi(0)$), où u a bien un point fixe, 0 , mais ne commute en général pas avec t_a . Soit c un vecteur de E ; on pose $\phi_0 = t_{-c} \circ \phi$, de sorte que $\phi = t_c \circ \phi_0$. Pour que t_c et ϕ_0 commutent, il faut et il suffit que $\phi = \phi_0 \circ t_c$, soit encore

$$t_a \circ u = \phi = \phi_0 \circ t_c = t_{-c} \circ \phi \circ t_c = t_{a-c} \circ u \circ t_c,$$

c'est-à-dire $t_{-c} \circ u \circ t_c = u$. Il faut donc que pour tout x dans E , on ait $u(x+c) - c = u(x)$, soit encore $u(c) = c$, ou $c \in \text{Ker}(\text{Id}_E - u)$. Pour que ϕ_0 ait un point fixe x_0 , il faut et il suffit que

$$x_0 = \phi_0(x_0) = u(x_0) + a - c,$$

c'est-à-dire $a = (\text{Id}_E - u)(x_0) + c$; tout revient donc à montrer l'existence et l'unicité d'une décomposition $a = b + c$ avec b dans l'image de $\text{Id}_E - u$ et c dans son noyau. Cela résulte du lemme suivant, qui termine la démonstration du théorème. ■

Lemme 48.5.— Soit u un endomorphisme orthogonal d'un espace vectoriel euclidien E . On a $\text{Im}(\text{Id}_E - u) = \text{Ker}(\text{Id}_E - u)^\perp$.

Soient x un élément de $\text{Ker}(\text{Id}_E - u)$ et $z = y - u(y)$ un élément de $\text{Im}(\text{Id}_E - u)$. On a

$$\langle x, z \rangle = \langle x, y - u(y) \rangle = \langle x, y \rangle - \langle x, u(y) \rangle = \langle u(x), u(y) \rangle - \langle x, u(y) \rangle = \langle u(x) - x, u(y) \rangle = 0,$$

de sorte que $\text{Im}(\text{Id}_E - u)$ est contenu dans l'orthogonal de $\text{Ker}(\text{Id}_E - u)$. Comme ces espaces ont même dimension (th. 34.3), ils sont égaux. ■

49. Réduction simultanée des formes quadratiques

Soit Q une forme quadratique quelconque sur l'espace vectoriel euclidien E . On sait qu'il existe une base orthogonale pour Q (th. 38.2), mais on va faire mieux.

Théorème 49.1.— Soit Q une forme quadratique d'un espace vectoriel euclidien E . Il existe une base orthonormée de E qui est orthogonale pour Q .

Notons pour une fois S le produit scalaire de E , et B la forme bilinéaire symétrique associée à Q . On cherche un endomorphisme u de E tel que $B(x, y) = \langle x, u(y) \rangle$. Avec les notations de 35.5, on écrit $B(x, y) = \widehat{B}(y)(x)$ et $\langle x, u(y) \rangle = \widehat{S}(u(y))(x)$; il suffit alors de prendre $u = \widehat{S}^{-1} \circ \widehat{B}$. On a alors

$$\langle x, u(y) \rangle = B(x, y) = B(y, x) = \langle y, u(x) \rangle,$$

de sorte que u est *symétrique*, donc diagonalisable dans une base orthonormée $\mathcal{B} = (e_1, \dots, e_n)$ de E (th. 45.9). Cela signifie que l'on peut écrire $u(e_j) = \lambda_j e_j$, d'où

$$B(e_j, e_k) = \langle e_j, u(e_k) \rangle = \lambda_k \langle e_j, e_k \rangle = \lambda_k \delta_{jk},$$

ce qui entraîne que la base \mathcal{B} est orthogonale pour Q . ■

A quoi sert ce théorème? Entre autres à l'étude des quadriques. Nous commencerons par définir une quadrique dans \mathbf{R}^n comme l'ensemble des points (x_1, \dots, x_n) qui vérifient une équation du type $P(x_1, \dots, x_n) = 0$, où P est un polynôme en n variables de degré au plus 2. Le polynôme P se décompose en la somme d'un polynôme homogène de degré 2, d'un polynôme homogène de degré 1, et d'une constante. Dans un espace vectoriel quelconque, un polynôme homogène de degré 2 en les coordonnées d'un vecteur correspond à une forme quadratique, et un polynôme homogène de degré 1 à une forme linéaire.

Définition 49.2.— Soient E un espace vectoriel, Q une forme quadratique, f une forme linéaire sur E , et c un réel. On appelle quadrique d'équation $Q(x) + f(x) + c = 0$ le lieu \mathcal{Q} des points x de E tels que $Q(x) + f(x) + c = 0$.

(49.3) On dit qu'une quadrique d'équation $Q(x) + f(x) + c = 0$ est *propre* si la forme quadratique

$$\tilde{Q}(x, t) = Q(x) + f(x)t + ct^2$$

sur $E \times \mathbf{R}$ est non dégénérée.

Nous allons tout d'abord chercher à simplifier l'équation de \mathcal{Q} , et dans un premier temps, à nous débarrasser si possible du terme linéaire. Soit p un point de E ; la translatée $t_{-p}(\mathcal{Q})$ est l'ensemble des points x de E qui vérifient l'équation

$$\begin{aligned} 0 &= Q(x + p) + f(x + p) + c \\ &= Q(x) + 2\widehat{B}(x, p) + f(x) + Q(p) + f(p) + c \\ &= Q(x) + (2\widehat{B}(p) + f)(x) + (Q(p) + f(p) + c), \end{aligned}$$

où B est la forme bilinéaire symétrique associée à Q . C'est donc encore une quadrique : la forme quadratique est Q , la forme linéaire $2\widehat{B}(p) + f$, et la constante $Q(p) + f(p) + c$. Si la forme quadratique Q est *non dégénérée*, l'application $\widehat{B} : E \rightarrow E^*$ est bijective; en prenant $p = -\widehat{B}^{-1}(f/2)$, le terme linéaire disparaît dans l'équation de la quadrique $t_{-p}(\mathcal{Q})$, qui est alors du type $Q(x) + c' = 0$. Ceci entraîne qu'elle est symétrique par rapport à l'origine, de sorte que \mathcal{Q} est symétrique par rapport au point p . C'est pour cela que l'on dit que les quadriques dont la forme quadratique est non dégénérée sont des quadriques *à centre*.

(49.4) Supposons \mathcal{Q} à centre l'origine (c'est-à-dire sans terme linéaire et avec Q non dégénérée). Le théorème 49.1 entraîne qu'il existe une base orthonormée \mathcal{B} de E telle que les points de \mathcal{Q} soient les points x de E qui vérifient

$$\sum_{j=1}^n \lambda_j x_j^2 + c = 0,$$

où x_1, \dots, x_n sont les coordonnées de x dans \mathcal{B} , et où $\lambda_1, \dots, \lambda_n$ ne sont pas nuls (puisque Q est non dégénérée). On dit que c'est une *équation réduite* de \mathcal{Q} . Pour que la quadrique \mathcal{Q} soit propre, il faut et il suffit que c ne soit pas nul.

Soit u l'endomorphisme symétrique de E associé à Q comme dans la démonstration du théorème 49.1; on appelle *axe* de \mathcal{Q} toute droite engendrée par un vecteur propre de u , et *sommet* de \mathcal{Q} tout point de \mathcal{Q} qui est sur un axe. Les axes correspondants à des valeurs propres distinctes sont orthogonaux par (45.10).

A chaque valeur propre λ telle que $c\lambda < 0$ correspondent des sommets de longueur $\sqrt{-c/\lambda}$; il y en a deux si λ est valeur propre simple, une infinité sinon. Pour tout j , la symétrie orthogonale par rapport à l'hyperplan d'équation $x_j = 0$ (c'est-à-dire par rapport à l'orthogonal d'un axe) laisse stable \mathcal{Q} .

Exemples 49.5. 1) La quadrique plane \mathcal{Q} (on dit alors la *conique*) d'équation $y = x^2 + 1$ n'est pas une conique à centre puisque la forme quadratique $Q(x, y) = x^2$ est dégénérée. En revanche, la forme quadratique $\tilde{Q}(x, y, t) = x^2 - yt + t^2$ est non dégénérée, de sorte que \mathcal{Q} est *propre*.

2) On considère la conique \mathcal{Q} d'équation $3x^2 + 2xy + 3y^2 + 4x - 4y = 0$. La matrice de la forme quadratique Q est $M = \begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix}$. Elle est inversible, donc \mathcal{Q} a un centre (a, b) , que l'on détermine en annulant le terme linéaire de

$$3(x+a)^2 + 2(x+a)(y+b) + 3(y+b)^2 + 4(x+a) - 4(y+b) = 0,$$

c'est-à-dire $6ax + 2ay + 2bx + 6by + 4x - 4y$. On obtient $6a + 2b + 4 = 2a + 6b - 4 = 0$, d'où $a = -1$ et $b = 1$: le centre de \mathcal{Q} est le point $p = (-1, 1)$. Soit \mathcal{Q}' la translatée de \mathcal{Q} par $(1, -1)$; elle est symétrique par rapport à l'origine, d'équation

$$3x^2 + 2xy + 3y^2 - 4 = 0.$$

Le polynôme caractéristique de M est $T^2 - 6T + 8$. Ses racines sont 2 et 4 ; elles sont bien réelles comme le prédit la théorie. Des vecteurs propres unitaires associés sont $e_2 = \begin{pmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{pmatrix}$ et $e_4 = \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix}$; ils sont bien orthogonaux comme le prédit la théorie. L'équation réduite de \mathcal{Q}' est, dans la base (e_2, e_4) ,

$$2X^2 + 4Y^2 = 4.$$

Les axes sont les droites (passant par p) dirigées par e_2 et e_4 , et la conique \mathcal{Q}' est symétrique par rapport à ces axes. Les sommets sont $\pm e_2\sqrt{2}$ et $\pm e_4$ (l'un d'eux est l'origine).

3) On considère la quadrique d'équation $x^2 + 3y^2 + z^2 + 4xz = 3$ dans \mathbf{R}^3 . Sa matrice est la matrice M de l'exemple 45.12, dont les valeurs propres sont -1 et 3 (double) ; dans la base orthonormée formée des vecteurs $(1/\sqrt{2}, 0, -1/\sqrt{2})$, $(1/\sqrt{2}, 0, 1/\sqrt{2})$ et $(0, 1, 0)$, l'équation réduite est

$$-X^2 + 3Y^2 + 3Z^2 = 3.$$

Les axes sont d'une part la droite dirigée par $(1, 0, -1)$ (qui engendre E_{-1}), d'autre part toutes les droites contenues dans $E_3 = E_{-1}^\perp$, c'est-à-dire toutes les droites orthogonales à E_{-1} . La quadrique \mathcal{Q} est symétrique par rapport au plan E_3 et par rapport à tout plan contenant E_{-1} ; (47.2) entraîne qu'elle est stable par toute rotation d'axe E_{-1} (on dit que \mathcal{Q} a une symétrie de révolution par rapport à cet axe).

Soit F un sous-espace vectoriel de E ; on appellera *sous-espace affine de E dirigé par F* tout sous-ensemble de E qui est le translaté de F par un vecteur p de E , c'est-à-dire $\{x \in E \mid x - p \in F\}$.

Soient \mathcal{Q} une quadrique d'équation $Q(x) + f(x) + c = 0$, et \mathcal{D} une droite affine passant par p et dirigée par un vecteur e ; regardons l'intersection de \mathcal{Q} et de \mathcal{D} . Il s'agit de résoudre l'équation

$$0 = Q(p + te) + f(p + te) + c = Q(e)t^2 + (2B(e, p) + f(e))t + Q(p) + f(p) + c,$$

de degré au plus 2 et de discriminant $\Delta = (2B(e, p) + f(e))^2 - 4Q(e)(Q(p) + f(p) + c)$. Ainsi, soit \mathcal{D} est contenue dans \mathcal{Q} , soit elle la rencontre en au plus 2 points. Lorsque $Q(e) \neq 0$, on est dans l'un des cas suivants :

- a) soit $\Delta < 0$ et \mathcal{D} ne rencontre pas \mathcal{Q} ;
- b) soit $\Delta > 0$ et \mathcal{D} rencontre \mathcal{Q} en deux points ;
- c) soit $\Delta = 0$ et \mathcal{D} rencontre \mathcal{Q} en exactement un point.

En général (même si $Q(e)$ est nul), on dira que \mathcal{D} est *tangente* à \mathcal{Q} si $\Delta = 0$.

Proposition 49.6.— Soient \mathcal{Q} une quadrique dans un espace vectoriel euclidien E et p un point de \mathcal{Q} . On appelle espace tangent à \mathcal{Q} en p la réunion des droites (affines) tangentes à \mathcal{Q} en p ; on le note $T_p\mathcal{Q}$. Soit il est égal à E et la quadrique \mathcal{Q} n'est pas propre, soit c'est un hyperplan affine.

Comme p est sur \mathcal{Q} , la condition $\Delta = 0$ se réduit à

$$0 = 2B(e, p) + f(e) = (2\widehat{B}(p) + f)(e) .$$

Il s'ensuit que la réunion des droites tangentes à \mathcal{Q} en p est le sous-espace affine de E dirigé par $\text{Ker}(2\widehat{B}(p) + f)$ et passant par p . C'est un hyperplan sauf si $2\widehat{B}(p) + f = 0$; on vérifie que cette condition est équivalente à dire que le vecteur $(p, 1)$ de $E \times \mathbf{R}$ est orthogonal à $E \times \{0\}$ pour la forme quadratique \tilde{Q} . Comme $\tilde{Q}(p, 1) = 0$, cela entraîne que $(p, 1)$ est dans le noyau de \tilde{Q} , ce qui ne peut arriver que si \tilde{Q} est dégénérée. ■

(49.7) Lorsque \mathcal{Q} est à centre l'origine, $T_p\mathcal{Q}$ est dirigé par l'orthogonal de p pour \mathcal{Q} .

50. Coniques

Par (49.4), l'équation d'une conique \mathcal{Q} propre à centre l'origine s'écrit, dans un repère orthormé convenable, $\lambda_1 x^2 + \lambda_2 y^2 + c' = 0$, avec λ_1, λ_2 et c' non nuls. En divisant par $-c'$, on peut supposer $c' = -1$. On a alors trois cas :

a) soit \mathcal{Q} est de signature $(2, 0)$, c'est-à-dire $\lambda_1 > 0$ et $\lambda_2 > 0$, et on écrit l'équation sous la forme

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 ,$$

avec $a \geq b > 0$. On dit que \mathcal{Q} est une *ellipse* (ou un *cercle* si $a = b$). Elle a 4 sommets.

b) Soit \mathcal{Q} est de signature $(1, 1)$, c'est-à-dire $\lambda_1 > 0$ et $\lambda_2 < 0$, et on écrit l'équation sous la forme

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1 ,$$

avec $a, b > 0$. On dit que \mathcal{Q} est une *hyperbole*. Elle a 2 sommets.

c) Soit \mathcal{Q} est de signature $(0, 2)$ et \mathcal{Q} est vide.

Soit maintenant \mathcal{Q} une conique propre dégénérée ; il existe une base orthonormée dans laquelle son équation s'écrit $\alpha y^2 + \beta x + \gamma y + c = 0$. Comme la forme quadratique

$(x, y, t) \mapsto \alpha y^2 + \beta xt + \gamma yt + ct^2$ est non dégénérée, on a $\alpha\beta \neq 0$. En traduisant par $(c/\beta - \gamma^2/2\alpha, \gamma/2\alpha)$, on obtient une équation de la forme $y^2 = 2qx$. On dit que \mathcal{Q} est une *parabole*.

(50.1) La forme quadratique \mathcal{Q} se lit sur n'importe quelle équation de conique propre. Si \mathcal{Q} est de signature $(1, 1)$, la conique est une hyperbole; si \mathcal{Q} est de signature $(2, 0)$ ou $(0, 2)$, c'est une ellipse ou l'ensemble vide; si \mathcal{Q} est dégénérée, c'est une parabole.

Théorème 50.2.— a) *A toute conique propre qui n'est pas un cercle sont associés un point F , appelé foyer, une droite affine \mathcal{D} ne contenant pas F , appelée directrice, et un nombre réel $e > 0$, son excentricité, tels que la conique soit l'ensemble des points p du plan vérifiant $d(p, F) = ed(p, \mathcal{D})$.*

b) *Inversement, étant donné un point F , une droite affine \mathcal{D} ne contenant pas F et un nombre réel $e > 0$, l'ensemble des points p du plan tels que $d(p, F) = ed(p, \mathcal{D})$ est une conique propre, une ellipse si $e < 1$ une parabole si $e = 1$, une hyperbole si $e > 1$.*

Donnons-nous un point F et une droite \mathcal{D} ne passant pas par F . Dans une base orthonormée dont le second vecteur dirige \mathcal{D} , celle-ci a pour équation $x = h$. Si F a pour coordonnées (x_0, y_0) dans cette base, le point p de coordonnées (x, y) vérifie $d(p, F) = ed(p, \mathcal{D})$ si et seulement si

$$(50.3) \quad (x - x_0)^2 + (y - y_0)^2 = e^2(x - h)^2,$$

qui est de degré 2, de formes quadratiques associées $\mathcal{Q}(x, y) = (1 - e^2)x^2 + y^2$ et

$$\tilde{\mathcal{Q}}(x, y, t) = (x - x_0t)^2 + (y - y_0t)^2 - e^2(x - ht)^2,$$

qui est non dégénérée puisque x_0 étant différent de h , les formes linéaires $x - x_0t$, $y - y_0t$ et $x - ht$ sont indépendantes. L'ensemble \mathcal{E} des points p vérifiant $d(p, F) = ed(p, \mathcal{D})$ est donc une conique propre. Si $e < 1$, \mathcal{Q} est définie positive; comme \mathcal{E} n'est pas vide, c'est une ellipse par (50.1); si $e > 1$, la signature de \mathcal{Q} est $(1, 1)$ et \mathcal{E} est une hyperbole; si $e = 1$, \mathcal{Q} est dégénérée et \mathcal{E} est une parabole. Ceci montre b).

Inversement, soit $\frac{x^2}{a^2} \pm \frac{y^2}{b^2} = 1$ ou $y^2 = 2qx$ une équation réduite de conique qui n'est pas un cercle. Elle est équivalente à une formule (50.3) avec dans le premier cas $y_0 = 0$, $e = \sqrt{1 \mp b^2/a^2}$ (qui est strictement positif puisqu'on n'a pas affaire à un cercle), $x_0 = ae$ et $h = a/e$; dans le second $y_0 = 0$, $e = 1$, et $x_0 = -h = -q/2$. Ceci montre a). ■

La fin de la démonstration montre que le foyer est sur un axe de symétrie de la conique, que l'on appelle *axe focal* et qui est orthogonal à \mathcal{D} . Par symétrie, les coniques propres à centre qui ne sont pas des cercles ont deux foyers (situés sur le grand axe pour une ellipse, sur l'axe joignant les deux sommets pour une hyperbole). Une ellipse est contenue dans la bande comprise entre ses deux directrices, tandis qu'une hyperbole est toute entière à l'extérieur de cette bande. On peut en fait décrire ces coniques uniquement à l'aide de ces foyers.

Théorème 50.4.— a) Une ellipse de foyers F et F' est l'ensemble des points p du plan tels que $d(p, F) + d(p, F') = 2a$, pour un certain réel a tel que $2a > d(F, F')$.

b) Une hyperbole de foyers F et F' est l'ensemble des points p du plan tels que $|d(p, F) - d(p, F')| = 2a$, pour un certain réel a tel que $0 < 2a < d(F, F')$.

Soit \mathcal{Q} une ellipse de foyers F et F' et de directrices \mathcal{D} et \mathcal{D}' . Pour tout point p de \mathcal{Q} , on a

$$d(p, F) + d(p, F') = e(d(p, \mathcal{D}) + d(p, \mathcal{D}')) = ed(\mathcal{D}, \mathcal{D}'),$$

puisque p est entre \mathcal{D} et \mathcal{D}' . Ceci montre que \mathcal{Q} est contenue dans l'ensemble

$$\mathcal{E} = \{ p \mid d(p, F) + d(p, F') = 2a \},$$

avec $a = ed(\mathcal{D}, \mathcal{D}')/2$, demi-grand axe. Inversement, choisissons une base orthonormée (e_1, e_2) dans laquelle $F' - F = 2ce_1$, avec $c = d(F, F')/2 < a$. Soit p un point de \mathcal{E} ; on a

$$\|p - F\|^2 - \|p - F'\|^2 = \langle 2p - F - F', F' - F \rangle = 2c \langle 2p - F - F', e_1 \rangle$$

et

$$\|p - F\| + \|p - F'\| = 2a.$$

En divisant la première égalité par la seconde, on obtient

$$\|p - F\| - \|p - F'\| = \frac{c}{a} \langle 2p - F - F', e_1 \rangle,$$

puis

$$\|p - F\| = a + \frac{c}{2a} \langle 2p - F - F', e_1 \rangle,$$

et enfin, en élevant au carré,

$$(50.5) \quad \langle p - F, p - F \rangle - \left(a + \frac{c}{2a} \langle 2p - F - F', e_1 \rangle \right)^2 = 0,$$

qui est bien un polynôme de degré 2 en les coordonnées (x, y) de p dont le terme homogène de degré 2 est $x^2 + y^2 - \frac{c^2}{a^2}x^2$. Puisque $c < a$, cette forme quadratique est définie positive; comme \mathcal{E} n'est pas vide puisqu'il contient \mathcal{Q} , c'est une ellipse (cf. (50.1)), donc c'est \mathcal{Q} . Ceci prouve a).

De la même façon, si \mathcal{Q} une hyperbole de foyers F et F' et de directrices \mathcal{D} et \mathcal{D}' , on a pour tout point p de \mathcal{Q}

$$d(p, F) - d(p, F') = e(d(p, \mathcal{D}) - d(p, \mathcal{D}')) = \pm ed(\mathcal{D}, \mathcal{D}'),$$

puisque p n'est pas entre \mathcal{D} et \mathcal{D}' . Ceci montre que \mathcal{Q} est contenue dans

$$\mathcal{E} = \{ p \mid |d(p, F) - d(p, F')| = 2a \}.$$

Inversement, on montre comme ci-dessus que \mathcal{E} est défini par la même équation (50.5). La signature de la forme quadratique $x^2 + y^2 - \frac{c^2}{a^2}x^2$ est maintenant $(1, 1)$, de sorte que \mathcal{E} est une hyperbole (cf. (50.1)). Ceci prouve b). ■

51. Surfaces quadriques

Nous ne considérerons que les quadriques propres \mathcal{Q} centrées à l'origine. Comme on l'a vu en (49.4), leur équation s'écrit, dans un repère orthormé convenable, soit

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1 ,$$

auquel cas on dit que c'est un *ellipsoïde* (ou une *sphère* si $a = b = c$), soit

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 1 ,$$

auquel cas on dit que c'est un *hyperboloïde à une nappe*, soit

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} - \frac{z^2}{c^2} = 1 ,$$

auquel cas on dit que c'est un *hyperboloïde à deux nappes* (le plan d'équation $x = 0$ est un plan de symétrie qui ne rencontre pas \mathcal{Q} , qui est formé de deux « morceaux » disjoints séparés par ce plan).

Théorème 51.1.— *Un hyperboloïde à une nappe \mathcal{H} est une surface réglée : par chaque point p de \mathcal{H} passent deux droites contenues dans \mathcal{H} , qui sont l'intersection de \mathcal{H} et du plan $T_p\mathcal{H}$.*

Comme \mathcal{H} est à centre l'origine, on a vu en (49.7) que $T_p\mathcal{H}$ est un plan affine dirigé par l'orthogonal $p^{\perp\mathcal{Q}}$ de p pour \mathcal{Q} . Notons

$$Q(x, y, z) = \frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} .$$

Soit $p + p'$ (avec p' dans $p^{\perp\mathcal{Q}}$) un point de $T_p\mathcal{H}$; pour qu'il soit dans \mathcal{H} , il faut et il suffit que

$$0 = Q(p + p') - 1 = Q(p) + 2B(p, p') + Q(p') - 1 = Q(p') ,$$

puisque $Q(p) = 1$ et $B(p, p') = 0$. L'intersection $T_p\mathcal{H} \cap \mathcal{H}$ est donc le translaté par p de l'ensemble des vecteurs isotropes pour la restriction Q' de Q à $p^{\perp\mathcal{Q}}$. Comme $Q(p) = 1 > 0$ et que la signature de Q est $(2, 1)$, la signature de Q' est $(1, 1)$; elle s'écrit $Q'(u, v) = u^2 - v^2$ dans une base convenable de $p^{\perp\mathcal{Q}}$, c'est donc la réunion de deux droites. ■

Remarque 51.2. Dans le cas d'un ellipsoïde ou d'un hyperboloïde à deux nappes, la même démonstration montre que $T_p\mathcal{H} \cap \mathcal{H}$ est réduit au point p (avec les notations ci-dessus, la restriction Q' est respectivement définie positive et définie négative).

(52.1) **Polynômes de Legendre** : soit E l'espace vectoriel des polynômes à coefficients réels de degré au plus n .

a) Montrer que la relation

$$\langle P, Q \rangle = \int_{-1}^1 P(t)Q(t)dt$$

définit un produit scalaire sur E .

b) Soit

$$L_m(x) = \frac{1}{2^m m!} \frac{d^m}{dx^m} ((x^2 - 1)^m).$$

Montrer que L_m est un polynôme de degré m et que (L_0, \dots, L_n) est une base de E . Quel est le coefficient dominant de L_m ? Calculer $L_m(1)$.

c) Calculer $\langle L_j, X^k \rangle$ pour $k < j$. En déduire que (L_0, \dots, L_n) est une base orthogonale de E .

d) Calculer $\|L_m\|$ (on pourra utiliser l'égalité $\int_0^{\pi/2} \sin^{2m+1} \theta d\theta = \frac{2^{2m} (m!)^2}{(2m+1)!}$).

e) Montrer pour tout élément P de E l'inégalité

$$|P(x)| \leq \|P\| \sqrt{\sum_{j=0}^n \frac{2j+1}{2} L_j(x)^2}$$

pour tout x . On peut d'autre part démontrer $|L_j(x)| \leq 1$ pour tout j et $|x| \leq 1$; en déduire

$$|P(x)| \leq \frac{n+1}{\sqrt{2}} \|P\|$$

pour $|x| \leq 1$.

(52.2) **Polynômes de Tchebycheff** : a) Soit m un entier positif. Montrer qu'il existe un unique polynôme T_m tel que l'on ait $T_m(\cos \theta) = \cos m\theta$ pour tout réel θ . Calculer T_0 , T_1 , T_2 et T_3 . Quels sont le degré et le coefficient dominant de T_m ?

b) Soit E l'espace vectoriel des polynômes à coefficients réels de degré au plus n . Montrer que la relation

$$\langle P, Q \rangle = \int_{-1}^1 \frac{P(t)Q(t)}{\sqrt{1-t^2}} dt$$

définit un produit scalaire sur E .

c) Montrer que (T_0, \dots, T_n) est une base orthogonale de E . Calculer la norme de T_m .

d) Soit P un élément de E unitaire de degré n . Montrer l'inégalité

$$\max_{|x| \leq 1} |P(x)| \geq \frac{1}{2^{n-1}},$$

avec égalité si et seulement si $P = T_n/2^{n-1}$ (on pourra considérer les racines du polynôme $P - T_n/2^{n-1}$).

(52.3) **Polynômes d'Hermite** : soit E l'espace vectoriel des polynômes à coefficients réels de degré au plus n .

a) Montrer que la relation

$$\langle P, Q \rangle = \int_{-\infty}^{+\infty} P(t)Q(t)e^{-t^2/2} dt$$

définit un produit scalaire sur E .

b) Soit

$$H_m(x) = \frac{e^{x^2/2}}{m!} \frac{d^m}{dx^m} (e^{-x^2/2}) .$$

Montrer que H_m est un polynôme de degré m et que (H_0, \dots, H_n) est une base de E . Quel est le coefficient dominant de H_m ?

c) Calculer $\langle H_j, X^k \rangle$ pour $k < j$. En déduire que (H_0, \dots, H_n) est une base orthogonale de E .

(52.4) Soient f et g des fonctions continues de $[a, b]$ dans \mathbf{R} . Montrer l'inégalité

$$\int_a^b |f(t)g(t)| dt \leq \sqrt{\int_a^b f(t)^2 dt} \sqrt{\int_a^b g(t)^2 dt} .$$

(52.5) Soient E un espace vectoriel réel et $Q : E \rightarrow \mathbf{R}^+$ une application qui vérifie, pour tous x et y dans E , la relation

$$(*) \quad Q(x+y) + Q(x-y) = 2(Q(x) + Q(y)) .$$

a) Montrer la relation

$$Q(x+y+z) + Q(x) + Q(y) + Q(z) = Q(x+y) + Q(y+z) + Q(z+x)$$

(on pourra utiliser $(*)$ pour déterminer les sommes deux à deux de $Q(x+y+z)$, $Q(x+y-z)$ et $Q(x-y-z)$).

b) Montrer que l'application $B : E \times E \rightarrow \mathbf{R}$ définie par

$$B(x, y) = Q(x+y) - Q(x) - Q(y)$$

vérifie $B(x+y, z) = B(x, z) + B(y, z)$.

c) Soit N une norme sur un espace vectoriel réel E . Montrer que N est euclidienne, c'est-à-dire qu'elle peut être définie à partir d'un produit scalaire, si et seulement si on a pour tous x et y dans E l'égalité

$$N(x+y)^2 + N(x-y)^2 = 2(N(x)^2 + N(y)^2) .$$

(52.6) On dit qu'une matrice symétrique réelle est *définie positive* si la forme quadratique de matrice M dans la base canonique de \mathbf{R}^n est définie positive.

a) Pour qu'une matrice symétrique réelle M soit définie positive, il faut et il suffit que ${}^t X M X > 0$ pour toute matrice colonne réelle non nulle X .

b) La somme de deux matrices symétriques réelles définies positives est définie positive.

c) Pour qu'une matrice réelle M soit définie positive, il faut et il suffit qu'il existe une matrice P inversible telle que $M = {}^t P P$.

d) Pour qu'une matrice symétrique réelle M soit définie positive, il faut et il suffit que ses valeurs propres soient strictement positives.

e) L'inverse d'une matrice symétrique réelle définie positive est définie positive.

f) Soit M une matrice symétrique réelle définie positive; montrer qu'il existe une unique matrice symétrique réelle définie positive N telle que $M = N^2$.

(52.7) Soit Q une forme quadratique sur \mathbf{R}^n de matrice M dans la base canonique; soit λ la plus grande valeur propre de M . Montrer l'inégalité

$$Q(x_1, \dots, x_n) \leq \lambda(x_1^2 + \dots + x_n^2) ,$$

pour tous réels x_1, \dots, x_n .

(52.8) Soient θ un réel non multiple de π et n un entier positif.

a) On pose $A_n = \frac{\sin(n+1)\theta}{\sin \theta}$. Calculer A_0 , A_1 , A_2 et A_3 en fonction de $\cos \theta$.

b) Exprimer $A_{n+2} + A_n$ en fonction de A_{n+1} .

c) On pose

$$D_n = \begin{vmatrix} 2 \cos \theta & 1 & 0 & \cdots & \cdots & 0 \\ 1 & 2 \cos \theta & 1 & 0 & \cdots & 0 \\ 0 & 1 & 2 \cos \theta & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 & 2 \cos \theta & 1 \\ 0 & \cdots & \cdots & 0 & 1 & 2 \cos \theta \end{vmatrix},$$

de sorte que $D_0 = 1$ et $D_1 = 2 \cos \theta$. Calculer D_2 et D_3 . Etablir une relation entre D_{n+2} , D_{n+1} et D_n . En déduire $D_n = A_n$ pour tout n .

d) En déduire les valeurs propres de la matrice carrée d'ordre n

$$M = \begin{pmatrix} 2 & -1 & 0 & \cdots & \cdots & 0 \\ -1 & 2 & -1 & 0 & \cdots & 0 \\ 0 & -1 & 2 & -1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & -1 & 2 & -1 \\ 0 & \cdots & \cdots & 0 & -1 & 2 \end{pmatrix}.$$

e) Montrer que M est définie positive et que, pour tous réels x_1, \dots, x_n , on a

$$x_1^2 + (x_1 - x_2)^2 + (x_2 - x_3)^2 + \cdots + (x_{n-1} - x_n)^2 + x_n^2 \geq 4 \sin^2 \frac{\pi}{2(n+1)} (x_1^2 + \cdots + x_n^2).$$

(52.9) a) Montrer l'inégalité

$$\sum_{1 \leq i < j \leq n} x_i x_j \leq \frac{n-1}{2} \sum_{i=1}^n x_i^2,$$

pour tous réels x_1, \dots, x_n (utiliser l'exercice (52.7)).

b) Soient $M = (a_{ij})_{1 \leq i, j \leq n}$ une matrice réelle et λ une valeur propre de M . Montrer l'inégalité

$$|\operatorname{Im}(\lambda)| \leq \frac{n-1}{2} \max_{1 \leq i < j \leq n} |a_{ij} - a_{ji}|$$

(on pourra introduire un vecteur propre X associé à λ et calculer ${}^t X(M - {}^t M)\bar{X}$ comme dans la démonstration du th. 45.9).

(52.10) Soit Q une forme quadratique définie positive sur \mathbf{R}^n .

a) Montrer que l'intégrale

$$\int_{\mathbf{R}^n} e^{-Q(x)} dx$$

converge et calculer sa valeur (on rappelle l'égalité $\int_{-\infty}^{+\infty} e^{-x^2} dx = \sqrt{\pi}$).

b) Calculer le volume de $\{x \in \mathbf{R}^n \mid Q(x) \leq 1\}$.

(52.11) Soit $M = (a_{ij})_{1 \leq i, j \leq n}$ une matrice réelle symétrique. Pour $1 \leq k \leq n$, on note M_k la matrice $(a_{ij})_{1 \leq i, j \leq k}$.

a) On suppose M définie positive; montrer que les déterminants de M_1, \dots, M_n sont tous strictement positifs.

b) On suppose que $d_j = \det(M_j)$ est strictement positif pour $1 \leq j \leq n$; on veut montrer que M est définie positive. Soient Q la forme quadratique de matrice M dans la base canonique de \mathbf{R}^n et F_k le

sous-espace vectoriel de \mathbf{R}^n engendré par les k premiers vecteurs de la base canonique. On raisonne par l'absurde; soit k le plus petit entier tel que la restriction de Q à F_k ne soit pas définie positive. Montrer que la signature de cette restriction est $(k-1, 1)$; en déduire une contradiction.

c) On suppose plus généralement que tous les d_j sont non nuls. Montrer qu'il existe une base de \mathbf{R}^n dans laquelle la matrice de Q est diagonale de coefficients diagonaux $(d_1, d_2/d_1, d_3/d_2, \dots, d_n/d_{n-1})$. On retrouve ainsi le résultat du b).

(52.12) Soit M une matrice réelle symétrique. Montrer que $\exp(M)$ est définie positive.

(52.13) Soient $M = (a_{ij})_{1 \leq i, j \leq n}$ et $N = (b_{ij})_{1 \leq i, j \leq n}$ des matrices réelles symétriques définies positives. Montrer que la matrice $(a_{ij} \bar{b}_{ij})_{1 \leq i, j \leq n}$ est symétrique définie positive. (montrer qu'il existe des réels p_{ij} et $d_k > 0$ tels que $a_{ij} = \sum_k p_{ki} d_k p_{kj}$).

(52.14) Soient M et N des matrices réelles symétriques. On suppose que N et $M - N$ sont définies positives; montrer que $M^{-1} - N^{-1}$ est définie positive (on pourra diagonaliser simultanément M et N).

(52.15) a) Montrer que la matrice $\left(\frac{1}{i+j-1} \right)_{1 \leq i, j \leq n}$ est définie positive (on pourra utiliser un analogue de l'exercice 52.1).

b) **Inégalité de Hilbert : pour tous réels x_1, \dots, x_n , montrer l'inégalité

$$\sum_{1 \leq i, j \leq n} \frac{x_i x_j}{i+j-1} < \pi \sum_{i=1}^n x_i^2.$$

(52.16) Soient u et v des endomorphismes symétriques d'un espace vectoriel euclidien E et k un entier strictement positif tel que $u^k = v^k$. Si k est pair, on suppose de plus que les valeurs propres de u et de v sont positives.

a) Montrer que tout vecteur propre de u^k est vecteur propre de u .

b) En déduire $u = v$.

(52.17) Soit M une matrice inversible réelle.

a) Montrer qu'il existe une unique matrice symétrique réelle définie positive N telle que $N^2 = {}^t M M$ (utiliser l'exercice 52.6).

b) Montrer que M peut s'écrire $M = O N$, avec O orthogonale, et que cette décomposition de M en produit d'une matrice orthogonale et d'une matrice symétrique réelle définie positive est unique (on pourra utiliser l'exercice précédent).

c) **Décomposition de Cartan** : montrer que M peut s'écrire $M = O D O'$, avec O et O' orthogonales et D diagonale à coefficients diagonaux strictement positifs. Cette décomposition est-elle unique?

(52.18) **Inégalité de Hadamard** : soit $M = (a_{ij})$ une matrice réelle carrée d'ordre n . Nous allons montrer l'inégalité

$$\det(M)^2 \leq \prod_{j=1}^n \left(\sum_{i=1}^n a_{ij}^2 \right).$$

a) Montrer qu'il y a égalité lorsque les vecteurs colonnes de M sont orthogonaux deux à deux.

b) Démontrer l'inégalité (on pourra appliquer le procédé d'orthonormalisation de Schmidt aux vecteurs colonnes e_1, \dots, e_n de M et montrer, avec les notations de la démonstration du théorème 44.8, les inégalités $\|e'_j\| \leq \|e_j\|$).

c) A quelle condition nécessaire et suffisante y a-t-il égalité?

(52.19) Soit F un sous-espace vectoriel d'un espace vectoriel euclidien E . Montrer que le déterminant de s_F est $(-1)^{\dim(F^\perp)}$ et que la trace de p_F est $\dim(F)$.

(52.20) Montrer qu'un endomorphisme orthogonal diagonalisable d'un espace vectoriel euclidien est une symétrie orthogonale.

(52.21) Exhiber une matrice symétrique *complexe* qui n'est pas diagonalisable.

(52.22) Soient M une matrice symétrique réelle définie positive et M' une matrice symétrique réelle. On va montrer de deux façons différentes que MM' est diagonalisable.

a) Ecrivons (exerc. 52.6.e) $M = N^2$, avec N symétrique réelle. Montrer que MM' est semblable à $NM'N$. Conclure.

b) Soient Q la forme quadratique sur \mathbf{R}^n de matrice M^{-1} , et Q' la forme quadratique de matrice M' . Conclure en appliquant le théorème 49.1.

(52.23) Soient x_1, \dots, x_p des vecteurs non nuls d'un espace vectoriel euclidien de dimension n , qui forment deux à deux des angles obtus, c'est-à-dire qui vérifient $\langle x_j, x_k \rangle \leq 0$ pour tous $j \neq k$.

a) Soient J une partie minimale de $\{1, \dots, p\}$ telle que les vecteurs $(x_j)_{j \in J}$ soient liés, et $\sum_{j \in J} \lambda_j x_j$ une combinaison linéaire nulle telle que l'ensemble J' des indices $j \in J$ tels que $\lambda_j > 0$ soit non vide. Montrer $\|\sum_{j \in J'} \lambda_j x_j\|^2 = 0$; en déduire $J' = J$.

b) Montrer $p \leq 2n$ (on pourra procéder par récurrence sur n , considérer une partie J comme en a) et montrer que pour tout j dans J et tout k non dans J , x_j est orthogonal à x_k).

c) Caractériser les ensembles de $2n$ vecteurs qui vérifient l'hypothèse.

d) On suppose maintenant $\langle x_j, x_k \rangle < 0$ pour $j \neq k$. Montrer $p \leq n + 1$. Montrer qu'il existe effectivement des ensembles de $n + 1$ vecteurs vérifiant l'hypothèse.

(52.24) Soit E un espace vectoriel euclidien de dimension n . A tous éléments x_1, \dots, x_m de E , on associe le réel

$$G(x_1, \dots, x_m) = \det(\langle x_j, x_k \rangle)_{1 \leq j, k \leq m},$$

appelé *déterminant de Gram* de (x_1, \dots, x_m) .

a) Montrer $G(x_1, \dots, x_m) = 0$ si x_1, \dots, x_m sont liés.

b) Montrer $G(x_1, \dots, x_m) > 0$ si x_1, \dots, x_m sont libres (on pourra interpréter la matrice définissant $G(x_1, \dots, x_m)$ comme la matrice de la restriction du produit scalaire au sous-espace vectoriel de E engendré par x_1, \dots, x_m). Pour $m = 2$, montrer que l'on retrouve l'inégalité de Cauchy-Schwarz.

c) Soient $\lambda_1, \dots, \lambda_m$ des réels; montrer

$$G(\lambda_1 x_1 + \dots + \lambda_m x_m, x_2, \dots, x_m) = \lambda_1^2 G(x_1, \dots, x_m)$$

(on pourra interpréter les matrices définissant $G(\lambda_1 x_1 + \dots + \lambda_m x_m, x_2, \dots, x_m)$ et $G(x_1, \dots, x_m)$ comme les matrices dans des bases différentes de la restriction du produit scalaire au sous-espace vectoriel de E engendré par x_1, \dots, x_m et utiliser la formule de changement de base).

d) On suppose (x_1, \dots, x_m) libres. Soient F le sous-espace vectoriel de E engendré par x_1, \dots, x_m et x un point de E . Montrer la formule (cf. th. 45.5)

$$d(x, F)^2 = \frac{G(x, x_1, \dots, x_m)}{G(x_1, \dots, x_m)}.$$

(52.25) Montrer que

$$\frac{1}{3} \begin{pmatrix} -2 & -1 & 2 \\ 2 & -2 & 1 \\ 1 & 2 & 2 \end{pmatrix}$$

est une matrice de rotation. En déterminer l'angle et l'axe.

(52.26) Déterminer tous les éléments du groupe $O(1, 1, \mathbf{R})$ (défini à la fin du chapitre précédent).

(52.27) a) Soient x, y et z des vecteurs de \mathbf{R}^3 et \mathcal{B} une base orthonormée directe. Montrer que le réel $\det_{\mathcal{B}}(x, y, z)$ ne dépend pas du choix de \mathcal{B} . On le note $[x, y, z]$.

b) Soit u un endomorphisme de \mathbf{R}^3 ; montrer l'égalité

$$[u(x), y, z] + [x, u(y), w] + [x, y, u(z)] = \text{tr}(u) [x, y, z].$$

En déduire qu'il existe un endomorphisme v de \mathbf{R}^3 tel que

$$u(x) \wedge y + x \wedge u(y) = v(x \wedge y)$$

pour tous x et y (on pourra faire intervenir l'adjoint de u).

(52.28) Déterminer la nature de l'endomorphisme u de \mathbf{R}^3 dont la matrice relativement à la base canonique de \mathbf{R}^3 est donnée ci-après, et préciser les éléments caractéristiques de u :

$$\begin{pmatrix} -2 & -1 & 2 \\ 2 & -2 & 1 \\ 1 & 2 & 2 \end{pmatrix}, \quad \begin{pmatrix} 3 & 1 & \sqrt{6} \\ 1 & 3 & -\sqrt{6} \\ -\sqrt{6} & \sqrt{6} & 2 \end{pmatrix}, \quad \begin{pmatrix} x^2 & x\sqrt{1-x^2} & \sqrt{1-x^2} \\ x\sqrt{1-x^2} & 1-x^2 & -x \\ \sqrt{1-x^2} & -x & 0 \end{pmatrix}.$$

(52.29) Soit r la rotation d'axe dirigé et orienté par un vecteur unitaire e et d'angle θ .

a) Montrer

$$r(x) = x + \sin\theta(e \wedge x) + (1 - \cos\theta)e \wedge (e \wedge x)$$

pour tout x dans \mathbf{R}^3 .

b) On suppose que r n'est pas une symétrie, et l'on note $t = \tan \frac{\theta}{2}$ et $e' = te$. Montrer l'égalité

$$r(x) = x + \frac{2}{1 + \|e'\|^2} e' \wedge (x + e' \wedge x)$$

pour tout x dans \mathbf{R}^3 (formule de Rodrigues).

(52.30) Pour chacune des quadriques \mathcal{Q} suivantes, préciser :

- un repère orthonormé dans lequel \mathcal{Q} admet une équation réduite ;
- l'équation réduite de \mathcal{Q} ;
- la nature de \mathcal{Q} .

a) $x^2 + y^2 + z^2 - 2yz - 4x + 4y - 1 = 0$.

b) $x^2 + y^2 + z^2 - 2xy - 1 = 0$.

c) $x^2 + y^2 + z^2 - 2xy + 2xz + 3x - y + z + 1 = 0$.

d) $x^2 + 4y^2 + 5z^2 - 4xy - 2x + 4y = 0$.

e) $x^2 - 4x - 3y + 4z - 2 = 0$.

f) $7x^2 - 2y^2 + 4z^2 + 4xy + 20xz + 16yz - 36x + 72y - 108z + 36 = 0$.

g) $(x - y)(y - z) + (y - z)(z - x) + (z - x)(x - y) + (x - y) = 0$.

h) $xy + yz = 1$.

i) $xy + xz + yz + 2y + 1 = 0$.

(52.31) Déterminer la matrice d'un demi-tour autour de la droite de vecteur unitaire (α, β, γ) .