

ESTIA 1^eAnnée - Mathématiques
Cours d'algèbre linéaire
Edition 2008

Xavier Dussau, Jean Esterle, Fouad Zarouf et Rachid Zarouf ¹

26 novembre 2008

¹I.Harlouchet-en eskuhartzearekin

Introduction

Ce cours d'algèbre linéaire se compose de 9 Chapitres. Dans le premier Chapitre on rappelle la définition et on donne sans démonstration les résultats classiques sur les espaces vectoriels et les applications linéaires (théorème de base incomplète, théorème de la dimension, etc...), avec en annexe une discussion de la notion d'application injective, surjective, bijective, illustrée par l'introduction des fonctions inverses des fonctions trigonométriques.

Au Chapitre 2 on introduit le calcul matriciel et on traite en détail des formules classiques concernant les changements de base. Au Chapitre 3 on rappelle les principales propriétés des déterminants, et on donne en annexe une introduction aux **notations de la Physique** (convention de sommation sur l'indice répété, etc...). Au Chapitre 4 on introduit le polynôme caractéristique p_A d'une matrice carrée A , et on étudie la diagonalisation des matrices et des endomorphismes.

On démontre directement au Chapitre 5 que $p_A(A) = 0$ (théorème de Cayley-Hamilton) par la méthode des déterminants, on introduit le polynôme minimal q_A d'une matrice carrée A , et on démontre le **théorème de décomposition de Jordan** : toute matrice carrée A dont le polynôme caractéristique est scindé s'écrit de manière unique sous la forme $A = D + N$, avec D diagonalisable, N nilpotente et $DN = ND$. Un calcul explicite de D et N est obtenu grâce au théorème chinois pour les polynômes.

Ces résultats sont appliqués au Chapitre 6 à l'itération des matrices et à la théorie des systèmes différentiels linéaires. On présente au Chapitre 7 la théorie des espaces vectoriels euclidiens (projections orthogonales, procédé d'orthogonalisation de Gram-Schmidt), et le Chapitre 8 est consacré aux matrices carrées symétriques et orthogonales, avec applications à la Géométrie. Le dernier Chapitre est consacré à la notion de **tenseur** en Physique.

De même que dans le cours d'algèbre, on a mis l'accent sur la possibilité de faire des **calculs effectifs**, et quand c'était possible on a systématiquement utilisé le logiciel de calcul formel MUPAD. Ceci dit l'effectivité des calculs trouve vite ses limites en algèbre linéaire dès qu'on s'écarte de problèmes relevant de la méthode du pivot de Gauss : l'impossibilité de trouver des formules algébriques exactes pour résoudre les équations de degré supérieur à 4, et la complexité des formules de Cardan-Tartaglia, font qu'en pratique on s'occupe essentiellement des matrices carrées à 3 lignes et 3 colonnes ayant $-2, -1, 0, 1$ ou 2 comme valeur propre évidente.

Aitzin solasa

Algebra linealaren ikastaldi hau bederatzi kapituluz osatua da. Lehen kapituluaren definizioa oroitarazi eta frogapenik gabe espazio bektorialeiei eta aplikazio linealei buruzko emaitza klasikoak (oinarri ezosoaren teorema, dimentsioaren teorema, etab.) emanak dira. Aplikazio injektibo, suprajektibo eta bijektiboen

nozioak eranskinean eztabaidatuak dira, funtzio trigonometrikoen alderantzizko funtzioen aurkezpenaz ilustraturik.

Bigarren kapituluan matrize-kalkulia aurkeztu eta oinarri-aldaketen formula klasiko batzu xeheki landuak dira. Hirugarren kapituluan, determinanteen propietate nagusiak oroitaraziak dira, eta eranskinean **Fisikako idazkeren** sarrera bat egina da (indize errepikatuarekiko batuketaren hitzarmena, etab.). Laugarren kapituluan, A matrize karratuaren p_A polinomio karakteristikoa aurkeztua da, eta matrizeen eta endomorfismoen diagonalizazioa ikertua.

Boskarren kapituluan $p_A(A) = 0$ (Cayley-Hamiltonen teorema) zuzenean determinanteen metodoaren bidez frogatua da, A matrize karratu baten q_A polinomio minimala aurkeztua da, eta **Jordan-en deskonposaketa teorema** frogatua : polinomio karakteristiko zatitua duen A matrize karratu oro era bakarrean $A = D + N$ forman idazten da, non D diagonalizagarria den, N nilpotentea eta $DN = ND$. D eta N -ren kalkulu esplizitua polinomioendako teorema txinoari esker lortua da.

Emaitza hauek matrizeen iterazioari eta sistema diferentzial linealen teoriari aplikatuak zaizkie seigarren kapituluan. Zazpigarren kapituluan, espazio bektorial euklidearren teoria aurkeztua da (projekzio ortogonalak, Gram-Schmidt-en ortonormalketa prozedura), eta zortzigarren kapitulua matrize karratu simetrikoen eta ortogonalei buruzkoa da, Geometriarako aplikazioez horniturik. Azken kapitulua Fisikako **tentsore** nozioari eskainia zaio.

Algebra ikastaldian bezala, **kalkulu eraginkorrak** egiteko aukera azpimarratu da, eta ahal zenean MUPAD kalkulu formalaren programa sistematikoki erabili da. Dena den, algebra linealean, Gauss-en ezabapen-metodoari dagozkion problemetarik urrunduz gero, kalkulu eraginkortasuna laster bere mугetara heltzen da : 4. mailatik gorako ekuazioen ebazteko formula algebraiko zehatzen aurkitzeko ezintasunak, eta Cartan-Tartaglia-ren formulen konplexutasunak, funtsean $-2, -1, 0, 1$ edo 2 balore propio nabaria duten 3 errenkadako eta 3 zutabeko matrize karratuez arduratzea ekartzen

Table des matières

1	Espaces vectoriels, applications linéaires	1
1.1	Indépendance linéaire, bases	1
1.2	Théorème de la dimension pour la somme de deux espaces vectoriels	3
1.3	Applications linéaires	4
1.4	Théorème de la dimension pour les applications linéaires	7
1.5	Annexe au Chapitre 1 : Application réciproque d'une bijection	10
1.6	Les fonctions Arctg, Arccos et Arcsin	13
1.7	Chapitre 1 sous MUPAD	15
1.8	Exercices sur le Chapitre 1	17
2	Matrices, changements de base	21
2.1	Calcul matriciel	21
2.2	Matrice associée à une application linéaire	22
2.3	Formule de changement de base pour les coordonnées de vecteurs	25
2.4	Formule de changement de base pour les applications linéaires	27
2.5	Chapitre 2 sous MUPAD	29
2.6	Exercices sur le Chapitre 2	31
3	Déterminants, notations indicielles de la Physique	33
3.1	Propriétés des déterminants	33
3.2	Matrices carrées inversibles	37
3.3	Formules de Cramer	39
3.4	Méthode du pivot de Gauss	40
3.5	Déterminant d'un endomorphisme et d'une famille de vecteurs	42
3.6	Rang d'une matrice	43
3.7	Rang d'une famille de vecteurs, rang d'une application linéaire	45
3.8	Annexe au Chapitre 3 : Introduction aux notations indicielles de la Physique	46
3.9	Chapitre 3 sous MUPAD	48
3.10	Exercices sur le Chapitre 3	54
4	Valeurs propres, vecteurs propres, diagonalisation	59
4.1	Introduction à la diagonalisation des matrices	59
4.2	Polynôme caractéristique d'une matrice	61

4.3	Matrices semblables	62
4.4	Dimension d'un sous-espace propre et ordre de multiplicité d'une valeur propre	63
4.5	Matrices diagonalisables	63
4.6	Endomorphismes diagonalisables	69
4.7	Chapitre 4 sous MUPAD	71
4.8	Exercices sur le Chapitre 4	78
5	Polynôme minimal, décomposition de Jordan	81
5.1	Théorème de Cayley-Hamilton	81
5.2	Polynôme minimal	83
5.3	Nouvelle caractérisation des matrices diagonalisables	85
5.4	Matrices nilpotentes	88
5.5	Théorème de décomposition de Jordan	90
5.6	Un algorithme de calcul rapide pour la décomposition de Jordan	94
5.7	Chapitre 5 sous Mupad	99
5.8	Exercices sur le Chapitre 5	105
6	Itération, systèmes différentiels linéaires	109
6.1	Calcul des puissances d'une matrice	109
6.2	Rappel sur les équations différentielles linéaires du premier ordre	116
6.3	Systèmes différentiels linéaires et exponentielles de matrices	118
6.4	Chapitre 6 sous Mupad	127
6.5	Exercices sur le Chapitre 6	130
7	Espaces vectoriels euclidiens	135
7.1	Produit scalaire, inégalité de Cauchy-Schwartz	135
7.2	Procédé d'orthonormalisation de Gram-Schmidt	137
7.3	Projections orthogonales	140
7.4	Chapitre 7 sous Mupad	142
7.5	Exercices sur le Chapitre 7	144
8	Matrices symétriques, matrices orthogonales	147
8.1	Matrices symétriques	147
8.2	Matrices orthogonales	149
8.3	Formes quadratiques	154
8.4	Matrices définies positives	163
8.5	Applications aux coniques et quadriques	167
8.6	Chapitre 8 sous Mupad	173
8.7	Tracé de coniques sous Matlab	175
8.8	Tracé de quadriques sous Matlab	184
8.9	Exercices sur le Chapitre 8	189

9 Les tenseurs en Physique (en préparation)	191
9.1 Coordonnées covariantes et contravariantes	191
9.2 A suivre...	194
9.3 Exercices sur le Chapitre 9	194

Chapitre 1

Espaces vectoriels, applications linéaires

1.1 Indépendance linéaire, bases

Définition 1.1.1 *Un espace vectoriel E sur un corps K (on dit aussi un K -espace vectoriel) est un groupe abélien additif muni d'une loi de composition externe $(\lambda, x) \mapsto \lambda.x$ de $K \times E$ dans E possédant les propriétés suivantes*

- (i) $(\lambda + \mu).x = \lambda.x + \mu.x \quad \forall x \in E, \forall \lambda \in K, \forall \mu \in K,$
- (ii) $\lambda.(\mu.x) = (\lambda\mu).x \quad \forall x \in E, \forall \lambda \in K, \forall \mu \in K,$
- (iii) $\lambda.(x + y) = \lambda.x + \lambda.y \quad \forall x \in E, \forall y \in E, \forall \lambda \in K,$
- (iv) $1.x = x \quad \forall x \in E.$

On écrira λx au lieu de $\lambda.x$ si aucune confusion n'est à craindre. On a l'importante notion de *sous-espace vectoriel*.

Définition 1.1.2 *On dit qu'une partie non vide F d'un K -espace vectoriel E est un sous-espace vectoriel de E si on a les deux propriétés suivantes*

- (i) $x + y \in F \quad \forall x \in F, \forall y \in F.$
- (ii) $\lambda x \in F \quad \forall \lambda \in K, \forall x \in F.$

Il est clair que si F est un sous-espace vectoriel de E alors F est lui-même un K -espace vectoriel pour la restriction à F de l'addition de E et du produit d'un élément de E par un élément de K . De plus pour qu'une partie non vide F de E soit un sous-espace vectoriel de E il faut et il suffit que l'on ait :

$$(1.1) \quad \lambda x + \mu y \in F \quad \forall \lambda \in K, \forall \mu \in K, \forall x \in F, \forall y \in F.$$

Une récurrence immédiate montre que l'on a alors la propriété suivante

$$(1.2) \quad \sum_{j=0}^n \lambda_j x_j \in F \quad \forall n \geq 1, \quad \forall x_1, \dots, x_n \in F, \quad \forall \lambda_1, \dots, \lambda_n \in K.$$

D'autre part il est clair que l'on a la propriété suivante

Proposition 1.1.3 *Soit $(E_i)_{i \in I}$ une famille quelconque de sous-espaces vectoriels d'un espace vectoriel E .*

Alors $\bigcap_{i \in I} E_i$ est un sous-espace vectoriel de E .

Corollaire 1.1.4 *Soit E un espace vectoriel sur un corps K et soit A une partie non vide de E . Alors l'intersection de tous les sous-espaces vectoriels de E contenant A est un sous-espace vectoriel de E contenant A , qui est appelé le sous-espace vectoriel de E engendré par A , et qui est noté $\text{Vect}(A)$.*

Définition 1.1.5 *1) On dit qu'une famille (e_1, \dots, e_k) d'éléments d'un K -espace vectoriel E est libre (ou que les vecteurs e_1, \dots, e_k sont linéairement indépendants) si on a la condition suivante*

$$(1.3) \quad \lambda_1 e_1 + \dots + \lambda_k e_k = 0 \implies \lambda_i = 0 \quad \forall i \leq k.$$

2) On dit qu'une famille (e_1, \dots, e_k) d'éléments d'un K -espace vectoriel E est génératrice si pour tout $x \in E$ il existe $\lambda_1, \dots, \lambda_k \in K$ tels que $x = \lambda_1 e_1 + \dots + \lambda_k e_k$.

3) On dit qu'une famille (e_1, \dots, e_k) d'éléments d'un K -espace vectoriel E est une base de E si elle est à la fois libre et génératrice.

On dit qu'un K -espace vectoriel $E \neq \{0\}$ est de dimension finie s'il existe une famille finie d'éléments de E qui est une base de E . On a le résultat fondamental suivant.

Théorème 1.1.6 *1) Pour qu'un K -espace vectoriel $E \neq \{0\}$ soit de dimension finie il faut et il suffit qu'il possède une famille génératrice finie. Dans ce cas toutes les bases de E ont le même nombre d'éléments, et ce nombre est appelé la dimension de E .*

2) Si $\dim(E) = n$, et si (f_1, \dots, f_k) est une famille génératrice d'éléments de E alors $k \geq n$. Si $k = n$, alors (f_1, \dots, f_k) est une base de E . Si $k > n$, alors il existe une suite strictement croissante i_1, \dots, i_n d'entiers inférieurs ou égaux à k tels que $(f_{i_1}, \dots, f_{i_n})$ soit une base de E .

3) Si $\dim(E) = n$, et si (e_1, \dots, e_p) est une famille libre d'éléments de E alors $p \leq n$. Si $p = n$, alors (e_1, \dots, e_p) est une base de E . Si $p < n$, alors il existe une famille (e_{p+1}, \dots, e_n) de $n - p$ éléments de E telle que $(e_1, \dots, e_p, e_{p+1}, \dots, e_n)$ soit une base de E ("théorème de la base incomplète").

On dira par convention que l'espace vectoriel réduit à $\{0\}$ est de dimension nulle.

1.2. THÉORÈME DE LA DIMENSION POUR LA SOMME DE DEUX ESPACES VECTORIELS 3

Corollaire 1.1.7 Soit E un espace vectoriel de dimension finie, et soit F un sous-espace vectoriel de E . Alors F est de dimension finie, et $\dim F \leq \dim E$. Si on a $\dim F = \dim E$, alors $E = F$.

Démonstration : Si $F = \{0\}$ il n'y a rien à démontrer. Sinon toute famille libre d'éléments de F possède au plus n éléments, où $n = \dim(E)$. Soit p le plus grand entier pour lequel il existe une famille libre (e_1, \dots, e_p) de p éléments de F . On a $p \leq n$. Soit $x \in F$. Comme la famille (e_1, \dots, e_p, x) n'est pas libre, il existe une famille $(\lambda_1, \dots, \lambda_{p+1})$ d'éléments non tous nuls de K telle que $\lambda_1 e_1 + \dots + \lambda_p e_p + \lambda_{p+1} x = 0$. On a $\lambda_{p+1} \neq 0$, car sinon $\lambda_1, \dots, \lambda_p$ seraient également tous nuls puisque (e_1, \dots, e_p) est libre. On obtient

$$x = \sum_{j=1}^p -\frac{\lambda_j}{\lambda_{p+1}} e_j.$$

Ceci montre que (e_1, \dots, e_p) est génératrice. C'est donc une base de F . Donc F est de dimension finie, et $\dim(F) = p \leq n = \dim(E)$.

Si $\dim(F) = \dim(E)$, soit \mathcal{B} une base de F . Alors \mathcal{B} est une base de E , donc $E = F$. ♣

Exemple 1.1.8 1) Soit K un corps. On munit l'ensemble K^n des familles (x_1, \dots, x_n) de n éléments de K des opérations suivantes

$$(x_1, \dots, x_n) + (y_1, \dots, y_n) = (x_1 + y_1, \dots, x_n + y_n) \quad \forall (x_1, \dots, x_n) \in K^n, \\ \forall (y_1, \dots, y_n) \in K^n,$$

$$\lambda(x_1, \dots, x_n) = (\lambda x_1, \dots, \lambda x_n) \quad \forall \lambda \in K, \quad \forall (x_1, \dots, x_n) \in K^n.$$

Posons $e_i = (\delta_{i,j})_{1 \leq j \leq n}$ pour $1 \leq i \leq n$. On vérifie facilement que K^n est un K -espace vectoriel de dimension n et que (e_1, \dots, e_n) est une base de K^n .

2) Le plan vectoriel du Lycée est un espace vectoriel réel de dimension 2, et l'espace vectoriel du Lycée est un espace vectoriel réel de dimension 3.

3) Muni des opérations usuelles, \mathbb{C} est un espace vectoriel réel de dimension 2, et $(1, i)$ est une base de \mathbb{C} .

4) L'espace $K[x]$ des polynômes sur un corps K est un K -espace vectoriel pour les opérations usuelles. Il n'est pas de dimension finie.

5) Pour $n \geq 0$, l'espace $K_n[x]$ des polynômes de degré inférieur ou égal à n à coefficients dans K est un K -espace vectoriel de dimension $n+1$, et la famille $(1, x, \dots, x^n)$ est une base de $K_n[x]$.

1.2 Théorème de la dimension pour la somme de deux espaces vectoriels

Définition 1.2.1 Soient F et G deux sous-espaces vectoriels d'un espace vectoriel E . On pose $F + G = \{x + y\}_{x \in F, y \in G}$.

On a alors le théorème suivant.

Théorème 1.2.2 *Soient F et G deux sous-espaces vectoriels de dimension finie d'un espace vectoriel E . Alors $\dim(F + G) = \dim(F) + \dim(G) - \dim(F \cap G)$.*

On dit que F et G sont en *somme directe* si $F \cap G = \{0\}$. Dans ce cas tout $x \in F + G$ s'écrit de manière unique sous la forme $x = y + z$, avec $y \in F, z \in G$. Dans cette situation on écrit $F \oplus G$ au lieu de $F + G$.

Exemple 1.2.3 *On considère l'espace vectoriel $\mathbb{R}_2[X]$ des polynômes à coefficients réels de degré inférieur ou égal à 2. On pose $E = \{p \in \mathbb{R}_2[x] \mid p(2) = 0\}$ et $F = \{p \in \mathbb{R}_2[x] \mid p(1) = p(-1) = 0\}$. Alors E et F sont deux sous-espaces vectoriels de $\mathbb{R}_2[x]$, et $\mathbb{R}_2[x] = E \oplus F$.*

En effet pour $p, q \in E, \lambda, \mu \in \mathbb{R}$ on a $(\lambda p + \mu q)(2) = \lambda p(2) + \mu q(2) = 0$, donc E est un sous-espace vectoriel de $\mathbb{R}_2[x]$. De même pour $p, q \in F, \lambda, \mu \in \mathbb{R}$ on a $(\lambda p + \mu q)(1) = \lambda p(1) + \mu q(1) = 0$, et $(\lambda p + \mu q)(-1) = \lambda p(-1) + \mu q(-1) = 0$, donc F est un sous-espace vectoriel de $\mathbb{R}_2[X]$.

Soit $p \in E \cap F$. On a $p(-1) = p(1) = p(2) = 0$. Comme un polynôme non nul de degré inférieur ou égal à 2 possède au plus deux racines distinctes, on a $p = 0$ et $E \cap F = \{0\}$, et on peut considérer la somme directe $E \oplus F$.

Posons $u = x - 2, v = x(x - 2)$. On a $u \in E$ et $v \in E$. Si $\lambda \in \mathbb{R}, \mu \in \mathbb{R}$, et si $\lambda u + \mu v = 0$ alors $(\lambda + \mu x)(x - 2) = 0$, donc $\lambda + \mu x = 0$ et $\lambda = \mu = 0$. Donc la famille $\{u, v\}$ est libre.

Si $p \in E$, alors p est divisible par $x - 2$ et il existe $q \in \mathbb{R}[x]$ tel que $p = q(x - 2)$. On a $d^\circ(q) + d^\circ(x - 2) = d^\circ(p) \leq 2$ donc $d^\circ(q) \leq 1$, et q est de la forme $q = \lambda + \mu x$ avec $\lambda, \mu \in \mathbb{R}$. On a alors $p = \lambda(x - 2) + \mu x(x - 2) = \lambda u + \mu v$ et on voit que $\{u, v\}$ est une famille génératrice de E . Par conséquent $\{u, v\}$ est une base de E et $\dim(E) = 2$.

Posons $w = x^2 - 1 \in F$. Si $p \in F$ alors p est divisible par $x - 1$ et $x + 1$ qui sont premiers entre eux. D'après le théorème de Gauss, p est divisible par $(x - 1)(x + 1) = w$. On a donc $p = qw$, avec $w \in \mathbb{R}[x]$. On a $2 \geq d^\circ(p) = d^\circ(q) + d^\circ(w) = d^\circ(q) + 2$ donc $d^\circ(q) = 0$ ou $q = 0$, et le polynôme q est constant. Donc $p = \lambda w$, avec $\lambda \in \mathbb{R}$, et F est le sous-espace vectoriel de dimension 1 de $\mathbb{R}_2[x]$ engendré par w .

On a $E \oplus F \subset \mathbb{R}_2[x]$, et $\dim(E \oplus F) = \dim(E + F) = \dim(E) + \dim(F) - \dim(E \cap F) = \dim(E) + \dim(F) = 3 = \dim(\mathbb{R}_2[x])$. Donc $\mathbb{R}_2[x] = E \oplus F$.

1.3 Applications linéaires

On va maintenant introduire l'importante notion d'**application linéaire**.

Définition 1.3.1 *Soient E et F deux espaces vectoriels sur un corps K . On dit qu'une application $u : E \rightarrow F$ est **linéaire** si les deux propriétés suivantes sont vérifiées*

- 1) $u(x + y) = u(x) + u(y) \quad \forall x \in E, \forall y \in F.$
 2) $u(\lambda x) = \lambda u(x) \quad \forall x \in E, \forall \lambda \in K.$

Si $u : E \rightarrow F$ est linéaire, on appelle **noyau de u** l'ensemble $\text{Ker}(u) := \{x \in E \mid u(x) = 0\}$, et on appelle **image de u** l'ensemble $\text{Im}(u) := u(E) = \{u(x)\}_{x \in E}.$

Une application linéaire $u : E \rightarrow E$ est appelée un **endomorphisme** de $E.$

Il est clair qu'une application $u : E \rightarrow F$ est linéaire si et seulement si la condition suivante est vérifiée

$$(1.4) \quad u(\lambda x + \mu y) = \lambda u(x) + \mu u(y) \quad \forall x, y \in E, \forall \lambda, \mu \in K.$$

Si A, B, C sont trois ensembles et si $\phi : A \rightarrow B$ et $\psi : B \rightarrow C$ sont deux applications, la composée $\psi \circ \phi : A \rightarrow C$ est l'application définie par la formule

$$(1.5) \quad (\psi \circ \phi)(x) = \psi(\phi(x)) \quad \forall x \in A.$$

On a alors le résultat évident suivant :

Proposition 1.3.2 Soient E_1, \dots, E_k des espaces vectoriels sur un corps K , avec $k \geq 3$, et pour $1 \leq i \leq k - 1$ soit $u_i : E_{i+1} \rightarrow E_i$ une application linéaire. Alors l'application composée $u_{k-1} \circ \dots \circ u_1 : E_k \rightarrow E_1$ est linéaire.

Proposition 1.3.3 Soient E et F deux espaces vectoriels sur un corps K et soit $u : E \rightarrow F$ une application linéaire.

(i) $\text{Ker}(u)$ est un sous-espace vectoriel de E , et $\text{Im}(u)$ est un sous-espace vectoriel de $F.$

(ii) Si E est de dimension finie, et si $\mathcal{B} = (e_1, \dots, e_n)$ est une base de E , alors $u(\mathcal{B}) := (u(e_1), \dots, u(e_n))$ est une famille génératrice de $\text{Im}(u).$ De plus $u(\mathcal{B})$ est une base de $\text{Im}(u)$ si $\text{Ker}(u) = \{0\}.$

Démonstration (i) Si $x_1, x_2 \in \text{Ker}(u)$, $\lambda_1, \lambda_2 \in K$, on a $u(\lambda_1 x_1 + \lambda_2 x_2) = \lambda_1 u(x_1) + \lambda_2 u(x_2) = 0$, donc $\lambda_1 x_1 + \lambda_2 x_2 \in \text{Ker}(u)$, et $\text{Ker}(u)$ est un sous-espace vectoriel de $E.$

Soient maintenant $y_1, y_2 \in \text{Im}(u)$ et $\lambda_1, \lambda_2 \in K.$ Il existe $x_1, x_2 \in E$ tels que $u(x_1) = y_1, u(x_2) = y_2$, et on a $\lambda_1 y_1 + \lambda_2 y_2 = \lambda_1 u(x_1) + \lambda_2 u(x_2) = u(\lambda_1 x_1 + \lambda_2 x_2) \in \text{Im}(u).$ Donc $\text{Im}(u)$ est un sous-espace vectoriel de $F.$

(ii) Si E est de dimension finie, et si $\mathcal{B} = (e_1, \dots, e_n)$ est une base de E , soit $y \in \text{Im}(u).$ Il existe $x \in E$ tel que $u(x) = y$, et il existe $\alpha_1, \dots, \alpha_n \in K$ tels que $x = \sum_{k=1}^n \alpha_k e_k.$ On a alors $y = u(x) = u(\sum_{k=1}^n \alpha_k e_k) = \sum_{k=1}^n \alpha_k u(e_k).$ Donc $u(\mathcal{B})$ est une famille génératrice de $\text{Im}(u).$

Si on suppose de plus que $\text{Ker}(u) = \{0\}$, soient $\lambda_1, \dots, \lambda_n \in K$ tels que $\lambda_1 u(e_1) + \dots + \lambda_n u(e_n) = 0.$ On a $u(\lambda_1 e_1 + \dots + \lambda_n e_n) = \lambda_1 u(e_1) + \dots + \lambda_n u(e_n) = 0$, donc $\lambda_1 e_1 + \dots + \lambda_n e_n = 0.$

Comme $\text{Ker}(u) = \{0\}$, on obtient $\lambda_1 e_1 + \dots + \lambda_n e_n = 0.$ Comme \mathcal{B} est libre, on a $\lambda_k = 0$ pour $1 \leq k \leq n$, ce qui prouve que $u(\mathcal{B})$ est libre, et $u(\mathcal{B})$ est dans ce cas une base de $\text{Im}(u).$ ♣

Exemple 1.3.4 1) L'application $D : p \mapsto p'$ est un endomorphisme de $\mathbb{C}_n[x]$ pour $n \geq 1$. Le noyau de D est l'espace vectoriel de dimension 1 formé des polynômes constants, et l'image de D est égale à $\mathbb{C}_{n-1}[x]$.

2) Les homothéties vectorielles et les symétries vectorielles sont des endomorphismes du plan vectoriel et de l'espace vectoriel de dimension 3 vus au Lycée. Les rotations vectorielles sont des endomorphismes du plan vectoriel euclidien.

3) Soit E un espace vectoriel sur un corps K et soit $\mathcal{B} = (e_1, \dots, e_p)$ une famille finie d'éléments de E . On note $F_{\mathcal{B}} = \{\lambda_1 e_1 + \dots + \lambda_p e_p\}_{(\lambda_1, \dots, \lambda_p) \in K^p}$ l'ensemble des combinaisons linéaires d'éléments de \mathcal{B} . Alors $F_{\mathcal{B}}$ est un sous-espace vectoriel de E , et $F_{\mathcal{B}} = \text{Vect}(\mathcal{B})$.

4) Si E est un espace vectoriel et si F et G sont des sous-espaces vectoriels de E tels que $E = F \oplus G$, l'application $P = P_{F,G}$ qui à $x \in E$ associe l'unique $y \in F$ tel que $x - y \in G$ est un endomorphisme de E tel que $\text{Im}(P) = F$ et $\text{Ker}(P) = G$. Cette application linéaire est appelée projection de E sur F parallèlement à G , et on a $P \circ P = P$.

En effet il est clair que D est linéaire et que $\text{Ker}(D) = \{p \in \mathbb{C}_n[x] \mid d^{\circ}(p) = 0\}$, qui est l'espace vectoriel de dimension 1 formé des polynômes constants. D'autre part $\text{Im}(D) \subset \mathbb{C}_{n-1}[x]$. Si $p = \alpha_0 + \dots + \alpha_{n-1}x^{n-1} \in \mathbb{C}_{n-1}[x]$, on a

$$p = D(\alpha_0 x + \frac{\alpha_1}{2}x^2 + \dots + \frac{\alpha_{n-1}}{n}x^n) \in \text{Im}(D),$$

et par conséquent on a bien $\text{Im}(D) = \mathbb{C}_{n-1}[x]$.

Considérons maintenant une famille finie $\mathcal{B} = (e_1, \dots, e_p)$ d'éléments d'un espace vectoriel E sur un corps K . Pour $\Lambda = (\lambda_1, \dots, \lambda_p) \in K^p$, posons $\phi(\Lambda) = \sum_{k=1}^p \lambda_k e_k$. Une vérification immédiate montre que $\phi : K^p \rightarrow E$ est linéaire, et par conséquent $F_{\mathcal{B}} = \text{Im}(\phi)$ est un sous-espace vectoriel de E . D'autre part il est clair que $F_{\mathcal{B}} \subset G$ si G est un sous-espace vectoriel de E contenant \mathcal{B} . Par conséquent $F_{\mathcal{B}}$ coïncide avec le sous-espace vectoriel $\text{Vect}(\mathcal{B})$ de E engendré par \mathcal{B} .

Considérons maintenant deux sous-espaces vectoriels F et G d'un espace vectoriel E tels que $E = F \oplus G$. Pour $x \in E$, il existe un unique couple $(y, z) \in F \times G$ tel que $x = y + z$, et par conséquent l'application $P = P_{F,G}$ introduite ci-dessus est bien définie. Soient $x_1, x_2 \in E$ et soient $\lambda_1, \lambda_2 \in K$. On a $\lambda_1 x_1 + \lambda_2 x_2 - \lambda_1 P(x_1) - \lambda_2 P(x_2) = \lambda_1(x_1 - P(x_1)) + \lambda_2(x_2 - P(x_2)) \in G$, et on a donc $P(\lambda_1 x_1 + \lambda_2 x_2) = \lambda_1 P(x_1) + \lambda_2 P(x_2)$, ce qui montre que P est linéaire. On a $\text{Im}(P) \subset F$. Si $y \in F$ on a par définition $P(y) = y$, car $y - 0 = y \in F$, et $y \in \text{Im}(P)$, ce qui montre que $\text{Im}(P) = F$. Donc si $x \in E$, on a $P(x) \in F$ et $P(P(x)) = P(x)$, ce qui montre que $P \circ P = P$.

Si $z \in G$ on a $z - 0 = z \in G$, donc $P(z) = 0$. Réciproquement si $P(z) = 0$, on a $z = z - P(z) \in F$, et par conséquent $\text{Ker}(P) = G$.

Notons que si $E = F \oplus G$, et si (y, z) est l'unique élément de $F \times G$ tel que $x = y + z$, on a $y = P_{F,G}(x)$ et $z = P_{G,F}(x)$. Avec les notations ci-dessus, on obtient, si $E = F \oplus G$,

$$(1.6) \quad P_{F,G} + P_{G,F} = \mathcal{I}_E,$$

où $\mathcal{I}_E : x \mapsto x$ est l'application identité de E .

1.4 Théorème de la dimension pour les applications linéaires

On va maintenant démontrer un "théorème de la dimension" pour les applications linéaires.

Théorème 1.4.1 *Soient E et F deux espaces vectoriels sur un corps K et soit $u : E \rightarrow F$ une application linéaire.*

Si E est de dimension finie, alors $\dim(\text{Ker}(u)) + \dim(\text{Im}(u)) = \dim(E)$.

Démonstration : Si $\text{Ker}(u) = E$, le résultat est évident. Si $\text{Ker}(u) = \{0\}$, il résulte de la proposition 1.13 (ii) que $\dim(E) = \dim(\text{Im}(u))$, et le résultat est encore vérifié.

Supposons maintenant que $\text{Ker}(u) \neq \{0\}$ et $\text{Ker}(u) \neq E$. Soit (e_1, \dots, e_p) une base de $\text{Ker}(u)$. D'après le théorème de la base incomplète, il existe $e_{p+1}, \dots, e_n \in E$ tels que (e_1, \dots, e_n) soit une base de E . Comme $u(e_1) = \dots = u(e_p) = 0$, il résulte de la proposition 1.13 (ii) que $(u(e_{p+1}), \dots, u(e_n))$ est une famille génératrice de $\text{Im}(u)$. Soit G le de E engendré par e_{p+1}, \dots, e_n .

Si $x \in G \cap \text{Ker}(u)$, il existe $\alpha_1, \dots, \alpha_p$ et $\beta_1, \dots, \beta_{n-p} \in K$ tels que $x = \alpha_1 e_1 + \dots + \alpha_p e_p = \beta_1 e_{p+1} + \dots + \beta_{n-p} e_n$. On a alors

$$\alpha_1 e_1 + \dots + \alpha_p e_p - \beta_1 e_{p+1} - \dots - \beta_{n-p} e_n = 0.$$

Comme (e_1, \dots, e_n) est une base de E , on a $\alpha_1 = \dots = \alpha_p = \beta_1 = \dots = \beta_{n-p} = 0$, et $x = 0$. Soit maintenant v la restriction de u à G , c'est à dire l'application $v : G \rightarrow F$ définie par la formule $v(x) = u(x)$ pour $x \in G$. On a $\text{Ker}(v) = \text{Ker}(u) \cap G = \{0\}$. Il résulte alors de la proposition 1.13 (ii) que $(u(e_{p+1}), \dots, u(e_n)) = (v(e_{p+1}), \dots, v(e_n))$ est libre. Donc $(u(e_{p+1}), \dots, u(e_n))$ est une base de $\text{Im}(u)$ et $\dim(\text{Im}(u)) = n - p$.

On a alors $\dim(\text{Ker}(u)) + \dim(\text{Im}(u)) = n - p + p = n = \dim(E)$, ce qui achève la démonstration. ♣

Exemple 1.4.2 *Posons $F := \{(x, y, z) \in \mathbb{R}^3 \mid x - 7y + 8z = 0\}$. Alors F est un sous-espace vectoriel de \mathbb{R}^3 , et $\dim(F) = 2$.*

En effet posons $u(x, y, z) = x - 7y + 8z$ pour $(x, y, z) \in \mathbb{R}^3$. Il est clair que u est une application linéaire de \mathbb{R}^3 dans \mathbb{R} . Comme $u \neq 0$ on a $\dim(\text{Im}(u)) \geq 1$, donc $\dim(\text{Im}(u)) = 1$ puisque $\text{Im}(u) \subset \mathbb{R}$. Comme $F = \text{Ker}(u)$, F est un sous-espace vectoriel de \mathbb{R}^3 et $\dim(F) = \dim(\mathbb{R}^3) - \dim(\text{Im}(u)) = 2$.

Soient A et B deux ensembles, soit $\phi : A \rightarrow B$ une application, et soit $y \in B$. On dit qu'un élément x de A est un **antécédent** de y pour ϕ quand $\phi(x) = y$. On dit que ϕ est **injective** quand tout $y \in B$ possède **au plus un antécédent** pour ϕ , et on dit que ϕ est **surjective** quand tout $y \in B$ possède **au moins un antécédent** pour ϕ . On dit enfin que ϕ est **bijective** quand ϕ est à la fois **injective** et **bijective**.

Si ϕ est bijective, on appelle **application réciproque** de ϕ l'application $\phi^{-1} : B \rightarrow A$ qui à $y \in B$ associe l'unique $x \in A$ tel que $\phi(x) = y$. Cette application réciproque est caractérisée par les propriétés suivantes

$$(1.7) \quad \phi^{-1} \circ \phi = \mathcal{I}_A \quad \text{et} \quad \phi \circ \phi^{-1} = \mathcal{I}_B$$

où \mathcal{I}_A et \mathcal{I}_B désignent l'application identité $x \mapsto x$ sur A et B .

On donnera en annexe divers exemples illustrant ces notions. En ce qui concerne les applications linéaires, on a les deux propriétés suivantes

(1.8) *Une application linéaire u est injective si et seulement si $\text{Ker}(u) = \{0\}$.*

(1.9) *Si $u : E \rightarrow F$ est linéaire et bijective, alors $u^{-1} : F \rightarrow E$ est linéaire.*

En effet $\text{Ker}(u)$ est l'ensemble des antécédents de 0. Comme $u(0) = 0$, on voit que $\text{Ker}(u) = \{0\}$ si u est injective. Supposons maintenant que $u : E \rightarrow F$ est linéaire et non injective. Il existe alors $y \in F$ possédant deux antécédents distincts x_1 et x_2 pour u . On a alors $u(x_1 - x_2) = y - y = 0$, donc $x_1 - x_2 \in \text{Ker}(u)$ et $\text{Ker}(u) \neq \{0\}$, ce qui établit (1.8).

Supposons maintenant que $u : E \rightarrow F$ est linéaire et bijective, et soient $y_1, y_2 \in F$ et $\lambda_1, \lambda_2 \in K$. On a $u(\lambda_1 u^{-1}(y_1) + \lambda_2 u^{-1}(y_2)) = \lambda_1 (u \circ u^{-1})(y_1) + \lambda_2 (u \circ u^{-1})(y_2) = \lambda_1 y_1 + \lambda_2 y_2$. Donc $\lambda_1 u^{-1}(y_1) + \lambda_2 u^{-1}(y_2) = (u^{-1} \circ u)(\lambda_1 u^{-1}(y_1) + \lambda_2 u^{-1}(y_2)) = u^{-1}[u(\lambda_1 u^{-1}(y_1) + \lambda_2 u^{-1}(y_2))] = u^{-1}(\lambda_1 y_1 + \lambda_2 y_2)$, ce qui prouve que u^{-1} est linéaire.

On a alors la conséquence suivante du théorème de la dimension.

Corollaire 1.4.3 *Soient E et F deux espaces vectoriels sur un corps K , et soit $u : E \rightarrow F$ une application linéaire. Si E et F sont de dimension finie, et si $\dim(E) = \dim(F)$, alors les conditions suivantes sont équivalentes*

- (i) u est injective
- (ii) u est surjective
- (iii) u est bijective

Démonstration : Si u est injective on a $\dim(\text{Ker}(u)) = 0$, donc $\dim(\text{Im}(u)) = \dim(E) - \dim(\text{Ker}(u)) = \dim(E) = \dim(F)$. Comme $\text{Im}(u) \subset F$, on a alors $\text{Im}(u) = F$ et u est surjective. Si u est surjective le même calcul montre

1.4. THÉORÈME DE LA DIMENSION POUR LES APPLICATIONS LINÉAIRES 9

que $\dim(\text{Ker}(u)) = \dim(E) - \dim(\text{Im}(u)) = \dim(E) - \dim(F) = 0$, donc $\text{Ker}(u) = \{0\}$ et u est injective d'après la propriété (1.8). Donc (i) et (ii) sont équivalents, ce qui montre que (i) \Rightarrow (iii). D'autre part il est évident que (iii) \Rightarrow (i). ♣

Exemple 1.4.4 Il existe une application linéaire $v : \mathbb{R}_n[x] \rightarrow \mathbb{R}_n[x]$ telle que $v(p) + v(p)' = p$ pour $p \in \mathbb{R}_n[x]$.

En effet posons $u(p) = p + p'$ pour $p \in \mathbb{R}_n[x]$. Il est clair que $u : \mathbb{R}_n[x] \rightarrow \mathbb{R}_n[x]$ est linéaire. Si $p \neq 0$ on a $d^\circ(p') < d^\circ(p)$ donc $d^\circ(u(p)) = d^\circ(p + p') = d^\circ(p) \geq 0$ et $u(p) \neq 0$. Par conséquent $\text{Ker}(u) = \{0\}$, u est injective et donc u est bijective. Posons $v = u^{-1}$. Alors v est linéaire. Si $p \in \mathbb{R}_n[x]$ on a $u(v(p)) = p$, c'est à dire que $v(p) + v(p)' = p$.

Pour conclure ce Chapitre on va introduire une structure vectorielle sur l'ensemble $\mathcal{L}(E, F)$ des applications linéaires d'un espace vectoriel E dans un espace vectoriel F , et une structure d'anneau sur l'ensemble des endomorphismes d'un espace vectoriel E .

Pour $\lambda \in K, u \in \mathcal{L}(E, F)$, on définit $\lambda u \in \mathcal{L}(E, F)$ par la formule

$$(1.10) \quad (\lambda u)(x) = \lambda u(x) \quad \forall x \in E.$$

De même pour $u \in \mathcal{L}(E, F), v \in \mathcal{L}(E, F)$, on définit $u + v \in \mathcal{L}(E, F)$ par la formule

$$(1.11) \quad (u + v)(x) = u(x) + v(x) \quad \forall x \in E.$$

Une vérification immédiate montre que les application λu et $u + v$ définies ci-dessus sont bien des applications linéaires de E dans F .

Des vérifications de routine permettent alors d'obtenir le résultat suivant

Proposition 1.4.5 (i) Soient E et F deux espaces vectoriels sur un corps K et soit $\mathcal{L}(E, F)$ l'ensemble des applications linéaires de E dans F . Alors $\mathcal{L}(E, F)$ est un espace vectoriel sur K pour les lois définies par les formules (1.10) et (1.11).

(ii) Soit E un espace vectoriel sur un corps K . Alors $\mathcal{L}(E, E)$ est un anneau unitaire pour l'addition définie par la formule (1.11) et pour la composition des applications.

On verra au Chapitre suivant que si $\dim(E) = m$ et $\dim(F) = n$ alors $\dim(\mathcal{L}(E, F)) = mn$.

1.5 Annexe au Chapitre 1 : Application réciproque d'une bijection

Soient U_1, U_2, U_3 trois ensembles et soient $\phi_1 : U_2 \rightarrow U_1$ et $\phi_2 : U_3 \rightarrow U_2$ deux applications. On définit l'application composée $\phi_2 \circ \phi_1 : U_1 \rightarrow U_3$ par la formule

$$(1.12) \quad (\phi_2 \circ \phi_1)(u) = \phi_2(\phi_1(u)) \quad \forall u \in U_1.$$

Soient maintenant U_1, U_2, U_3, U_4 quatre ensembles et soient $\phi_1 : U_2 \rightarrow U_1$, $\phi_2 : U_3 \rightarrow U_2$ et $\phi_3 : U_4 \rightarrow U_3$ trois applications. On déduit immédiatement de la formule (1.12) que l'on a

$$(1.13) \quad \phi_3 \circ (\phi_2 \circ \phi_1) = (\phi_3 \circ \phi_2) \circ \phi_1.$$

Supposons maintenant que $\phi : U \rightarrow V$ est bijective. Dans ce cas tout élément v de V possède un unique antécédent u pour ϕ . Autrement dit il existe un unique $u \in U$ tel que $\phi(u) = v$.

Définition 1.5.1 Soit $\phi : U \rightarrow V$ une application bijective. On appelle **application réciproque** de ϕ l'application $\phi^{-1} : V \rightarrow U$ qui à $v \in V$ associe l'unique élément de U tel que $\phi(u) = v$. Autrement dit si $u \in U, v \in V$, alors

$$u = \phi^{-1}(v) \iff v = \phi(u).$$

Notons $\mathcal{I}_U : u \mapsto u$ l'application identité sur U et notons $\mathcal{I}_V : v \mapsto v$ l'application identité sur V . Si $\phi : U \rightarrow V$ est bijective, on obtient les formules

$$(1.14) \quad \phi^{-1} \circ \phi = \mathcal{I}_U, \quad \phi \circ \phi^{-1} = \mathcal{I}_V.$$

En effet comme $\phi(u) = \phi(u)$ on a $(\phi^{-1} \circ \phi)(u) = \phi^{-1}(\phi(u)) = u$ pour $u \in U$. D'autre part on a par définition $(\phi \circ \phi^{-1})(v) = \phi(\phi^{-1}(v)) = v$ pour $v \in V$.

Proposition 1.5.2 Soit U un ensemble, et soit $\mathcal{S}(U)$ l'ensemble des bijections de U sur lui-même. Alors $(\mathcal{S}(U), \circ)$ est un groupe dont l'unité est \mathcal{I}_U , et l'inverse de $\phi \in \mathcal{S}(U)$ est égal à l'application réciproque ϕ^{-1} . De plus si U est un ensemble fini possédant n éléments alors $\mathcal{S}(U)$ possède $n!$ éléments.

Démonstration : Le nombre des bijections d'un ensemble de n éléments dans lui-même est égal au nombre d'arrangements de cet ensemble, et on a vu en Terminale que ce nombre est égal à $n!$. Les autres propriétés résultent des formules (1.13) et (1.14).

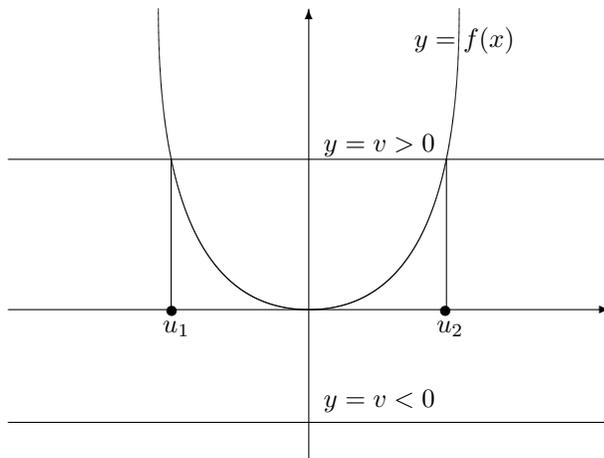
Nous utiliserons au Chapitre 3 un certain nombre de propriétés du groupe \mathcal{S}_n des permutations d'ordre n , qui est le groupe des bijections de l'ensemble $\{1, \dots, n\}$ sur lui-même. Nous allons maintenant illustrer les notions d'application injective, surjective, bijective et la notion d'antécédent dans le cadre des fonctions continues d'une variable réelle. Si I et J sont deux intervalles de \mathbb{R}

1.5. ANNEXE AU CHAPITRE 1 : APPLICATION RÉCIPROQUE D'UNE BIJECTION 11

et si $f : I \rightarrow J$ est une fonction, le ou les antécédents pour f d'un réel $v \in J$ sont donnés par les abscisses des points d'intersection de la droite horizontale d'équation $y = v$ avec la courbe représentative \mathcal{F} de f . Ce type de discussion graphique est abordé dès la classe de seconde. On va maintenant prendre divers exemples.

Exemple 1 : On considère

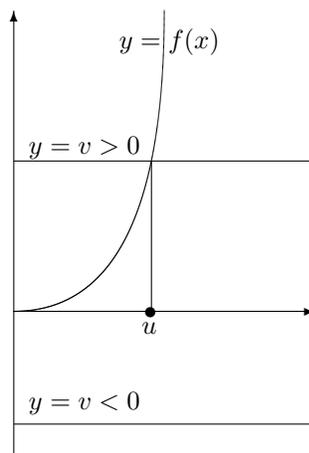
$$\begin{aligned} f : \mathbb{R} &\rightarrow \mathbb{R} \\ x &\mapsto x^2 \end{aligned}$$



Si $v > 0$ la droite d'équation $y = v$ rencontre le graphe de f en deux points distincts. Dans ce cas v possède deux antécédents distincts x_1 et x_2 pour f . Donc f **n'est pas injective**. Si $v < 0$ la droite d'équation $y = v$ ne rencontre pas le graphe de f . Dans ce cas v n'a pas d'antécédent pour f donc f **n'est pas surjective**.

Exemple 2 : On considère

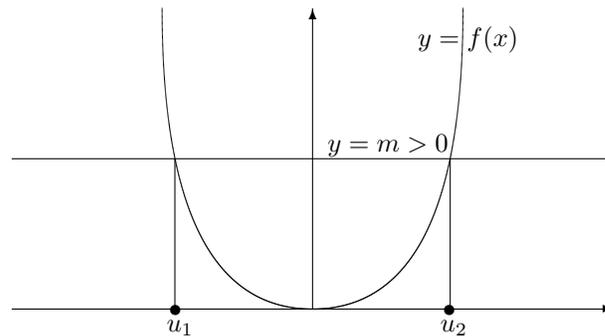
$$\begin{aligned} f : [0, +\infty[&\rightarrow \mathbb{R} \\ x &\mapsto x^2 \end{aligned}$$



Si $v \geq 0$ la droite d'équation $y = v$ rencontre le graphe de f en un seul point. Dans ce cas v possède exactement un antécédent u pour f . Si $v < 0$ la droite d'équation $y = v$ ne rencontre pas le graphe de f . Dans ce cas v n'a pas d'antécédent pour f . Donc f **est injective**, mais f **n'est pas surjective**.

Exemple 3 : On considère

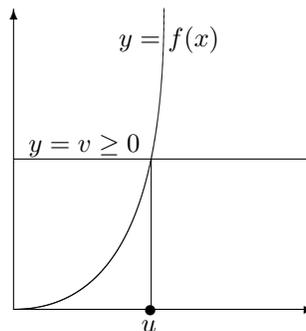
$$f : \mathbb{R} \rightarrow [0, +\infty[\\ x \mapsto x^2$$



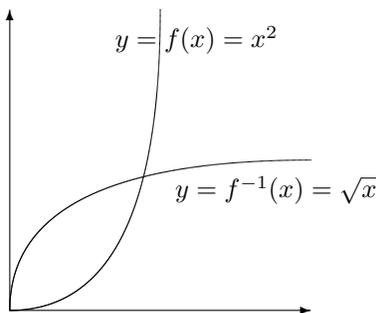
Si $v > 0$ la droite d'équation $y = v$ rencontre le graphe de f en deux points distincts. Dans ce cas v possède deux antécédents distincts u_1 et u_2 pour f . Donc f **n'est pas injective**. Si $v = 0$ l'intersection de la droite d'équation $y = 0$ et du graphe de f est réduite au point $(0, 0)$. Dans ce cas v a au moins un d'antécédent pour f pour tout $v \geq 0$ donc f **est surjective**.

Exemple 4 : On considère

$$f : [0, +\infty[\rightarrow [0, +\infty[\\ x \mapsto x^2$$



Si $v \geq 0$ la droite d'équation $y = v$ rencontre le graphe de f en un point et un seul, donc v possède un unique antécédent pour tout $v \geq 0$, et f est bijective. Comme on l'a vu en Terminale, le graphe de $f^{-1} : [0, +\infty[\rightarrow [0, +\infty[$ est le symétrique du graphe de f par rapport à la droite d'équation $y = x$. On obtient la figure suivante



On sait depuis la Terminale que si I est un intervalle de \mathbb{R} et si $f : I \rightarrow \mathbb{R}$ est continue et strictement monotone alors $J = f(I) = \{f(x)\}_{x \in I}$ est un intervalle de \mathbb{R} et que f , considérée comme application de I dans J , est une bijection de I sur J . De plus si f est dérivable en $f^{-1}(x)$, et si $f'(f^{-1}(x)) \neq 0$, alors f^{-1} est dérivable en x et $(f^{-1})'(x) = \frac{1}{f'(f^{-1}(x))}$ (géométriquement, ceci résulte immédiatement du fait que la dérivée de f en x représente la pente de la tangente au graphe de f au point de coordonnées $(x, f(x))$ et du fait que le graphe de f^{-1} est le symétrique du graphe de f par rapport à la droite d'équation $y = x$).

Si $f : [0, +\infty[\rightarrow [0, +\infty[$ est la fonction de l'exemple 4 il est clair que $f^{-1} : [0, +\infty[\rightarrow [0, +\infty[$ est définie par la formule $f^{-1}(x) = \sqrt{x}$ pour $x \geq 0$. Comme $f'(\sqrt{x}) = 2\sqrt{x} \neq 0$ pour $x \neq 0$, on retrouve le fait que si $g(x) = \sqrt{x}$, alors $g'(x) = \frac{1}{2\sqrt{x}}$ pour $x > 0$.

1.6 Les fonctions Arctg, Arccos et Arcsin

Ces considérations permettent d'introduire de nouvelles fonctions utiles pour le calcul intégral.

Proposition 1.6.1 Soit $f :]-\frac{\pi}{2}, \frac{\pi}{2}[\rightarrow]-\infty, +\infty[$
 $x \mapsto \operatorname{tg}(x)$.

La fonction réciproque $f^{-1} :]-\infty, +\infty[\rightarrow]-\frac{\pi}{2}, \frac{\pi}{2}[$ est une primitive de $\frac{1}{1+x^2}$ sur \mathbb{R} . Cette fonction est appelée "Arc tangente" et notée Arctg.

Démonstration : La fonction tangente est strictement croissante sur $] -\frac{\pi}{2}, \frac{\pi}{2}[$, donc la fonction $Arctg$ est bien définie sur \mathbb{R} . On a $(Arctg)'(x) = \frac{1}{f'(Arctg(x))} = \frac{1}{1+tg^2(Arctg(x))} = \frac{1}{1+x^2}$ pour $x \in \mathbb{R}$. ♣

Il résulte de la définition ci-dessus que si u et v sont deux réels alors $u = Arctg(v) \iff tg(u) = v$ et $u \in] -\frac{\pi}{2}, \frac{\pi}{2}[$. On retiendra les valeurs suivantes :

$$Arctg(0) = 0, Arctg(1/\sqrt{3}) = \pi/6, Arctg(1) = \pi/4, Arctg(\sqrt{3}) = \pi/3,$$

$$\lim_{x \rightarrow +\infty} Arctg(x) = \pi/2.$$

D'autre part $Arctg(-x) = -Arctg(x)$ pour $x \in \mathbb{R}$, c'est à dire que la fonction $Arctg$ est **impaire**. Cette fonction est évidemment croissante sur \mathbb{R} .

La définition de la fonction $Arctg$ peut causer quelques difficultés. Par exemple $tg(\frac{\pi}{2} - Arctg(v)) = \frac{1}{tg(Arctg(v))} = \frac{1}{v} = tg(Arctg(\frac{1}{v}))$ pour $v \neq 0$, mais $\frac{\pi}{2} - Arctg(v) \neq Arctg(\frac{1}{v})$ pour $v < 0$. On renvoie aux exercices 10 et 11 pour ce type de questions. On retiendra de la fonction $Arctg$ la formule suivante (nous renvoyons au Chapitre 4 du *Cours d'algèbre* pour une étude générale de l'intégration des fractions rationnelles).

$$(1.15) \quad \int_a^b \frac{dx}{1+x^2} = Arctg(b) - Arctg(a) \quad \forall a \in \mathbb{R}, \forall b \in \mathbb{R}.$$

Les fonctions $x \mapsto \sin x$ et $x \mapsto \cos x$ ne sont évidemment pas monotones sur \mathbb{R} , mais leurs restrictions à des intervalles convenables le sont. Ceci permet d'introduire les fonctions $Arccos$ et $Arcsin$, également utiles pour le calcul intégral.

Proposition 1.6.2 Soient $g : [-\frac{\pi}{2}, \frac{\pi}{2}] \longrightarrow [-1, 1]$ et $h : [0, \pi] \longrightarrow [-1, 1]$
 $x \mapsto \sin x$ $x \mapsto \cos x$.

La fonction réciproque $g^{-1} : [-1, +1] \rightarrow [-\frac{\pi}{2}, \frac{\pi}{2}]$ est une primitive de $\frac{1}{\sqrt{1-x^2}}$ sur $] -1, 1[$. Cette fonction est appelée "Arc sinus" et notée $Arcsin$.

De même la fonction réciproque $h^{-1} : [-1, +1] \rightarrow [0, \pi]$ est une primitive de $-\frac{1}{\sqrt{1-x^2}}$ sur $] -1, 1[$. Cette fonction est appelée "Arc cosinus" et notée $Arccos$.

Démonstration : Il est clair que la restriction de la fonction sinus à l'intervalle $[-\frac{\pi}{2}, \frac{\pi}{2}]$ est strictement croissante et que la restriction de la fonction cosinus à l'intervalle $[0, \pi]$ est strictement décroissante, donc les fonctions $Arcsin$ et $Arccos$ sont bien définies sur $[-1, 1]$. Pour $x \in [-1, 1]$ on a $\sin(Arccos(x)) \geq 0$ et $\cos(Arcsin(x)) \geq 0$, donc $\sin(Arccos(x)) = \sqrt{1 - \cos^2(Arccos(x))}$ et $\cos(Arcsin(x)) = \sqrt{1 - \sin^2(Arcsin(x))}$. On obtient

$$(1.16) \quad \sin(Arccos(x)) = \sqrt{1-x^2}, \cos(Arcsin(x)) = \sqrt{1-x^2} \quad \forall x \in [-1, 1].$$

Pour $x \in] -1, 1[$ on a alors $(Arcsin)'(x) = \frac{1}{g'(Arcsin(x))} = \frac{1}{\cos(Arcsin(x))} = \frac{1}{\sqrt{1-x^2}}$, et $(Arccos)'(x) = \frac{1}{h'(Arccos(x))} = -\frac{1}{\sin(Arccos(x))} = -\frac{1}{\sqrt{1-x^2}}$. ♣

La fonction Arcsin est impaire et croissante sur $[-1, 1]$, et la fonction Arccos est décroissante sur $[0, \pi]$. D'autre part on a $\cos(\text{Arccos}(-x)) = -x = \cos(\pi - \text{Arccos}(x))$. Comme $\text{Arccos}(-x)$ et $\pi - \text{Arccos}(x)$ appartiennent à $[0, \pi]$ on obtient

$$(1.17) \quad \text{Arccos}(-x) = \pi - \text{Arccos}(x) \quad \forall x \in [0, 1].$$

Il résulte de la définition de $\text{Arccos}(x)$ et $\text{Arcsin}(x)$ que l'on a

$$\text{Arccos}(1) = 0, \text{Arccos}\left(\frac{\sqrt{3}}{2}\right) = \frac{\pi}{6}, \text{Arccos}\left(\frac{\sqrt{2}}{2}\right) = \frac{\pi}{4}, \text{Arccos}\left(\frac{1}{2}\right) = \frac{\pi}{3}, \text{Arccos}(0) = \frac{\pi}{2}, \text{Arccos}\left(-\frac{1}{2}\right) = \frac{2\pi}{3}, \text{Arccos}\left(-\frac{\sqrt{2}}{2}\right) = \frac{3\pi}{4}, \text{Arccos}\left(-\frac{\sqrt{3}}{2}\right) = \frac{5\pi}{6}, \text{Arccos}(-1) = \pi,$$

$$\text{Arcsin}(-1) = -\frac{\pi}{2}, \text{Arcsin}\left(-\frac{\sqrt{3}}{2}\right) = -\frac{\pi}{3}, \text{Arcsin}\left(-\frac{\sqrt{2}}{2}\right) = -\frac{\pi}{4}, \text{Arcsin}\left(-\frac{1}{2}\right) = -\frac{\pi}{6}, \text{Arcsin}(0) = 0, \text{Arcsin}\left(\frac{1}{2}\right) = \frac{\pi}{6}, \text{Arcsin}\left(\frac{\sqrt{2}}{2}\right) = \frac{\pi}{4}, \text{Arcsin}\left(\frac{\sqrt{3}}{2}\right) = \frac{\pi}{3}, \text{Arcsin}(1) = \frac{\pi}{2}.$$

La définition des fonctions Arccos et Arcsin présente encore un peu plus de difficultés que celle de la fonction Arctg : le fait que $\cos(u) = v$ est équivalent au fait que $u = \text{Arccos}(v) + 2k\pi$ ou $u = -\text{Arccos}(v) + 2k\pi$, avec $k \in \mathbf{Z}$, et le fait que $\sin(u) = v$ est équivalent au fait que $u = \text{Arcsin}(v) + 2k\pi$ ou $u = \pi - \text{Arcsin}(v) + 2k\pi$, avec $k \in \mathbf{Z}$. Ceci est la base de nombreux exercices, d'un intérêt souvent relatif (voir par exemple l'exercice 13). On retiendra la formule suivante, qui résulte immédiatement de la proposition 1.6.2.

$$(1.18) \quad \int_a^b \frac{dx}{\sqrt{1-x^2}} = \text{Arcsin}(b) - \text{Arcsin}(a) \\ = \text{Arccos}(a) - \text{Arccos}(b), \quad \forall a \in [-1, 1], \forall b \in [-1, 1].$$

$$\text{Par exemple } \int_{-\frac{1}{2}}^{\frac{1}{2}} \frac{dx}{\sqrt{1-x^2}} = \text{Arcsin}\left(\frac{1}{2}\right) - \text{Arcsin}\left(-\frac{1}{2}\right) = \frac{\pi}{3}.$$

1.7 Chapitre 1 sous MUPAD

MUPAD est programmé pour calculer une base de l'espace vectoriel engendré par une famille finie de vecteurs. par exemple voici comment on calcule une base

du sous-espace vectoriel F de \mathbf{R}^3 engendré par $\mathbf{v}_1, \mathbf{v}_2$ et \mathbf{v}_3 , où $\mathbf{v}_1 = \begin{pmatrix} -4 \\ 3 \\ 7 \\ 8 \\ 1 \end{pmatrix}$,

$$\mathbf{v}_2 = \begin{pmatrix} 1 \\ 0 \\ 9 \\ 5 \\ 6 \end{pmatrix}, \mathbf{v}_3 = \begin{pmatrix} -3 \\ 3 \\ 16 \\ 13 \\ 7 \end{pmatrix},$$

```
M:=Dom::Matrix ();
v1:= M([-4,3,7,8,1]): v2:= M([1,0,9,5,6]):v3:=M([-3,3,16,13,7]):
linalg::basis([v1,v2,v3]);
```

```
Dom::Matrix(Dom::ExpressionField(id, iszero))
```

```
-- +-      +- +-      +- --
| | -4 | | 1 | |
| |      | |      | |
| | 3 | | 0 | |
| |      | |      | |
| | 7 | | 9 | |
| |      | |      | |
| | 8 | | 5 | |
| |      | |      | |
| | 1 | | 6 | |
-- +-      +- +-      +- --
```

On voit donc que $(\mathbf{v}_1, \mathbf{v}_2)$ est une base de F .

On peut également utiliser Mupad pour calculer une base de l'intersection de deux sous-espaces vectoriels. Par exemple on va calculer ci-dessous une base de $G \cap H$, où G est le sous-espace vectoriel de \mathbf{R}^3 engendré par $\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$ et $\begin{pmatrix} 2 \\ -1 \\ 0 \end{pmatrix}$, et où H est le sous-espace vectoriel de \mathbf{R}^3 engendré par $\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$ et $\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$.

```
M:=Dom::Matrix() :
u1:= M([1,0,-1]): u2:=M([2,-1,0]) : u3:=M([1,0,1]) : u4:=M([1,1,0]):
linalg::intBasis([u1,u2],[u3,u4]);
```

```
{ +-      +- }
{ | -1/2 | }
{ |      | }
```

$$\begin{aligned} & \{ \mid \quad 1 \quad \mid \} \\ & \{ \mid \quad \quad \quad \mid \} \\ & \{ \mid -3/2 \quad \mid \} \\ & \{ +- \quad \quad \quad -+ \} \end{aligned}$$

On voit donc que $G \cap H$ est la droite vectorielle de vecteur directeur $\mathbf{u} = \begin{pmatrix} -1/2 \\ 1 \\ -3/2 \end{pmatrix}$.

1.8 Exercices sur le Chapitre 1

exercice 1

Parmi les sous-ensembles de \mathbf{R}^2 suivants, identifier ceux qui sont stables pour l'addition, pour la multiplication externe, et enfin ceux qui sont des sous-espaces vectoriels de \mathbf{R}^2 :

$$\mathbf{Z}^2, A = \{(x, y) \mid |x| = |y|\}, B = \{(x, y) \mid x + 2y = 0\}.$$

exercice 2

Montrer que $\{(1+x)^k\}_{0 \leq k \leq n}$ est une famille libre de $\mathbf{R}_n[x]$. En déduire que c'est une base de $\mathbf{R}_n[x]$.

exercice 3

On pose $E := \{(x, y, z) \in \mathbf{R}^3 \mid 2236x + 3458y + 5256z + 1438 = 0\}$ et $F := \{(x, y, z) \in \mathbf{R}^3 \mid 4x + 2y + 3z = 0\}$. Ces ensembles sont-ils des sous-espaces vectoriels de \mathbf{R}^3 ? Si oui, calculer leur dimension et leur trouver une base.

exercice 4

(a) Soit $\mathcal{F}(\mathbf{R}, \mathbf{R})$ l'espace vectoriel sur \mathbf{R} constitué des fonctions de \mathbf{R} dans \mathbf{R} . Montrer que $\mathcal{F}(\mathbf{R}, \mathbf{R})$ n'est pas de dimension finie.

(b) Les ensembles suivants sont-ils des espaces vectoriels sur \mathbf{R} ?

$$V_1 = \{f \in \mathcal{F}(\mathbf{R}, \mathbf{R}) \mid f(0) = 0\}, V_2 = \{f \in \mathcal{F}(\mathbf{R}, \mathbf{R}) \mid f(3) > 0\},$$

$$V_3 = \{f \in \mathcal{F}(\mathbf{R}, \mathbf{R}) \mid f \text{ est paire } \}.$$

exercice 5

Dans \mathbf{R}^3 , déterminer $F := \text{Vect}\{(0, 0, 1), (1, 0, 0)\}$. Comparer F et $G := \text{Vect}\{(1, 0, 1), (-1, 0, 2)\}$.

exercice 6

Soient $U := \{(x, x, x)\}_{x \in \mathbf{R}}$ et $V := \{(0, y, z)\}_{y, z \in \mathbf{R}}$.
 Montrer que $\mathbf{R}^3 = U \oplus V$.

exercice 7

Soit E un espace vectoriel de dimension finie sur un corps K . En utilisant le théorème de la base incomplète, montrer que si F est un sous-espace vectoriel de E il existe un sous-espace vectoriel G de E tel que $E = F \oplus G$.

exercice 8

On considère l'application

$$\begin{aligned} \phi : \mathbf{R}_2[x] &\rightarrow \mathbf{R}_2[x] \\ p &\longmapsto 2p - p'. \end{aligned}$$

Déterminer l'image et le noyau de ϕ .

exercice 9

Soit E un ensemble possédant au moins 3 éléments. Montrer que le groupe $S(E)$ des bijections de E dans lui-même n'est pas commutatif.

exercice 10

On pose $g(x) = \operatorname{Arctg}(x) + \operatorname{Arctg}(\frac{1}{x})$ pour $x \neq 0$. Montrer que $g'(x) = 0$ pour $x \neq 0$. En déduire que $g(x) = -\frac{\pi}{2}$ pour $x < 0$ et que $g(x) = \frac{\pi}{2}$ pour $x > 0$.

exercice 11

Comparer $\operatorname{Arctg}(v) + \operatorname{Arctg}(w)$ et $\operatorname{Arctg}(\frac{v+w}{1-vw})$ pour $v, w \in \mathbf{R}$, $vw \neq -1$.

exercice 12

On pose $\int_0^{+\infty} f(x)dx = \lim_{L \rightarrow +\infty} \int_0^L f(x)dx$ pour f continue sur $[0, +\infty[$ quand cette limite existe.

Calculer $\int_0^{+\infty} \frac{dx}{4+x^2}$.

exercice 13

Soit $x \in [-1, 1]$.

- Comparer $2\operatorname{Arccos}(x)$ et $\operatorname{Arccos}(2x^2 - 1)$.
- Comparer $2\operatorname{Arcsin}(x)$ et $\operatorname{Arcsin}(2x\sqrt{1-x^2})$.

exercice 14

On pose $\int_0^a f(x)dx = \lim_{\epsilon \rightarrow 0^+} \int_0^{a-\epsilon} f(x)dx$ pour f continue sur $[0, a[$, $a > 0$, quand cette limite existe.

Calculer $\int_0^2 \frac{dx}{\sqrt{4-x^2}}$.

exercice 15(sous Mupad)

Déterminer une base du sous-espace vectoriel de \mathbf{R}^5 engendré par $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ et \mathbf{v}_4 , où $\mathbf{v}_1 = \begin{pmatrix} 1 \\ 2 \\ 3 \\ -1 \\ 2 \end{pmatrix}$, $\mathbf{v}_2 = \begin{pmatrix} 2 \\ 3 \\ 4 \\ -4 \\ 1 \end{pmatrix}$, $\mathbf{v}_3 = \begin{pmatrix} 3 \\ 5 \\ 7 \\ -5 \\ 3 \end{pmatrix}$, $\mathbf{v}_4 = \begin{pmatrix} 1 \\ -1 \\ 1 \\ -1 \\ 7 \end{pmatrix}$.

exercice 16

1) Trouver une base du plan vectoriel P_1 d'équation $x + y + z = 0$ et une base du plan vectoriel P_2 d'équation $-x + 7y + 3z = 0$.

2) En utilisant Mupad, trouver un vecteur directeur de la droite vectorielle $D = P_1 \cap P_2$.

Chapitre 2

Matrices, changements de base

2.1 Calcul matriciel

Soit K un corps. Une *matrice* à coefficients dans K est un tableau rectangulaire d'éléments de K .

On les note sous la forme $A = \begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} & \dots & a_{2,n} \\ \dots & \dots & \dots & \dots \\ a_{m,1} & a_{m,2} & \dots & a_{m,n} \end{pmatrix}$ ou sous la forme

synthétique $A = (a_{i,j})_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}}$ (on notera que l'indice correspondant à la ligne est écrit avant celui correspondant à la colonne). L'ensemble des matrices à m lignes et n colonnes à coefficients dans K est noté $\mathcal{M}_{m,n}(K)$.

Pour $A = (a_{i,j})_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}} \in \mathcal{M}_{m,n}(K)$, $B = (b_{i,j})_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}} \in \mathcal{M}_{m,n}(K)$, $\lambda \in K$, on pose

$$A + B = (a_{i,j} + b_{i,j})_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}}$$

$$\lambda A = (\lambda a_{i,j})_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}}$$

D'autre part pour $A = (a_{i,j})_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}} \in \mathcal{M}_{m,n}(K)$, $B = (b_{i,j})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq p}} \in \mathcal{M}_{n,p}(K)$, on pose

$$AB = \left(\sum_{1 \leq k \leq n} a_{i,k} b_{k,j} \right)_{\substack{1 \leq i \leq m \\ 1 \leq j \leq p}}$$

En effet $D(1) = 0, D(x) = 1$ et $D(x^2) = 2x$.

Exemple 2.2.3 Pour $n \geq 1$ soit $\mathcal{B}_n = (e_1, \dots, e_n)$ la "base canonique" de K^n ,

définie par la formule $e_j = \begin{bmatrix} \delta_{1,i} \\ \vdots \\ \delta_{n,j} \end{bmatrix}$ pour $1 \leq j \leq n$, soient m et n deux entiers

et soit $A \in \mathcal{M}_{m,n}(K)$. Soit $u_A : K^n \rightarrow K^m$ l'application linéaire définie par la formule $u_A(X) = AX$ pour $X \in K^n$. Alors $M_{u_A, \mathcal{B}_n, \mathcal{B}_m} = A$.

En effet on a $u_A(e_j) = Ae_j = \begin{bmatrix} a_{1,j} \\ \vdots \\ a_{m,j} \end{bmatrix} = \sum_{i=1}^m a_{i,j} e_i$ pour $1 \leq j \leq n$. Par

conséquent toutes les colonnes de $M_{u_A, \mathcal{B}_n, \mathcal{B}_m}$ et A coïncident, et $M_{u_A, \mathcal{B}_n, \mathcal{B}_m} = A$.

Le résultat simple suivant permet de ramener les calculs concernant les applications linéaires à des calculs matriciels.

Proposition 2.2.4 Soient E et F deux espaces vectoriels non nuls de dimension finie, soit \mathcal{B}_0 une base de E , soit \mathcal{B}_1 une base de F et soit $u \in \mathcal{L}(E, F)$. On a la formule

$$(2.1) \quad Y = M_{u, \mathcal{B}_0, \mathcal{B}_1} X$$

où X désigne les coordonnées d'un élément quelconque x de E dans la base \mathcal{B}_0 et où Y désigne les coordonnées de $u(x)$ dans la base \mathcal{B}_1 .

Démonstration : Posons $\mathcal{B}_0 = (e_1, \dots, e_n)$, $\mathcal{B}_1 = (f_1, \dots, f_m)$, $M_{u, \mathcal{B}_0, \mathcal{B}_1} :=$

$$(m_{i,j})_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}}, X = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, Y = \begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix}.$$

On a $M_{u, \mathcal{B}_0, \mathcal{B}_1} X = \begin{bmatrix} \sum_{j=1}^n m_{1,j} x_j \\ \vdots \\ \sum_{j=1}^n m_{m,j} x_j \end{bmatrix}$. D'autre part $u(x) = \sum_{j=1}^n x_j u(e_j) =$

$$\sum_{j=1}^n x_j (\sum_{i=1}^m m_{i,j} f_i) = \sum_{i=1}^m (\sum_{j=1}^n m_{i,j} x_j) f_i.$$

Par conséquent $\begin{bmatrix} \sum_{j=1}^n m_{1,j} x_j \\ \vdots \\ \sum_{j=1}^n m_{m,j} x_j \end{bmatrix}$ est la matrice colonne formée des coor-

données de $u(x)$ dans la base \mathcal{B}_1 , et on a bien

$$Y = M_{u, \mathcal{B}_0, \mathcal{B}_1} X. \clubsuit$$

Notons que la relation $Y = M_{u, \mathcal{B}_0, \mathcal{B}_1} X$ caractérise la matrice $M_{u, \mathcal{B}_0, \mathcal{B}_1}$. En effet si une matrice $N = (n_{i,j})_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}}$ vérifie la même relation, on a, en appliquant

cette formule avec $x = e_j$ pour $1 \leq j \leq n$, ce qui donne $X = \begin{bmatrix} \delta_{1,j} \\ \cdot \\ \cdot \\ \cdot \\ \delta_{n,j} \end{bmatrix}$,

$$\begin{bmatrix} m_{1,j} \\ \cdot \\ \cdot \\ \cdot \\ m_{m,j} \end{bmatrix} = Y = NX = N \begin{bmatrix} \delta_{1,j} \\ \cdot \\ \cdot \\ \cdot \\ \delta_{n,j} \end{bmatrix} = \begin{bmatrix} n_{1,j} \\ \cdot \\ \cdot \\ \cdot \\ n_{m,j} \end{bmatrix}.$$

Par conséquent toutes les colonnes de $M_{u, \mathcal{B}_0, \mathcal{B}_1}$ et N sont égales, et $N = M_{u, \mathcal{B}_0, \mathcal{B}_1}$.

Soient maintenant E, F, G trois espaces vectoriels non nuls de dimension finie, soit \mathcal{B}_0 une base de E , soit \mathcal{B}_1 une base de F et soit \mathcal{B}_2 une base de G . Pour $x \in E$ notons X le vecteur colonne formé des coordonnées de x dans \mathcal{B}_0 , Y le vecteur colonne formé des coordonnées de $u(x)$ dans \mathcal{B}_1 et Z le vecteur colonne formé des colonnes de $(u \circ v)(x) = u(v(x))$ dans \mathcal{B}_2 . On a $Z = M_{v, \mathcal{B}_1, \mathcal{B}_2} Y = M_{v, \mathcal{B}_1, \mathcal{B}_2} M_{u, \mathcal{B}_0, \mathcal{B}_1} X$ et on obtient

$$(2.2) \quad M_{v \circ u, \mathcal{B}_0, \mathcal{B}_2} = M_{v, \mathcal{B}_1, \mathcal{B}_2} M_{u, \mathcal{B}_0, \mathcal{B}_1}.$$

Soient E et F deux espaces vectoriels sur un corps K . On dira qu'une application $\phi : E \rightarrow F$ est un *isomorphisme linéaire* si ϕ est linéaire et bijective, et dans ce cas l'application réciproque $\phi^{-1} : F \rightarrow E$ est également un isomorphisme linéaire. Si \mathcal{A} et \mathcal{B} sont deux anneaux unitaires d'unités $1_{\mathcal{A}}$ et $1_{\mathcal{B}}$, on dira qu'une application $\psi : \mathcal{A} \rightarrow \mathcal{B}$ est un *homomorphisme d'anneaux* si les conditions suivantes sont vérifiées

$$(2.3) \quad \begin{cases} \psi(x+y) = \psi(x) + \psi(y) & \forall x \in \mathcal{A}, \quad \forall y \in \mathcal{B} \\ \psi(xy) = \psi(x)\psi(y) & \forall x \in \mathcal{A}, \quad \forall y \in \mathcal{B} \\ \psi(1_{\mathcal{A}}) = 1_{\mathcal{B}} \end{cases}$$

On dira que $\psi : \mathcal{A} \rightarrow \mathcal{B}$ est un *isomorphisme d'anneaux* si ψ est un homomorphisme d'anneaux bijectif. Dans ce cas on voit également que l'application réciproque $\psi^{-1} : \mathcal{B} \rightarrow \mathcal{A}$ est un isomorphisme d'anneaux. On a le résultat suivant.

Proposition 2.2.5 (i) Soient E et F deux espaces vectoriels de dimensions finies $n \geq 1$ et $m \geq 1$ sur un corps K , soit \mathcal{B}_0 une base de E et soit \mathcal{B}_1 une base de F . Alors l'application $u \rightarrow M_{u, \mathcal{B}_0, \mathcal{B}_1}$ est un isomorphisme linéaire de $\mathcal{L}(E, F)$ sur $\mathcal{M}_{m,n}(K)$.

(ii) Soit E un espace vectoriel de dimension $n \geq 1$ sur un corps K et soit \mathcal{B}_0 une base de E . Alors l'application $u \rightarrow M_{u, \mathcal{B}_0, \mathcal{B}_0}$ est un isomorphisme d'anneaux de $\mathcal{L}(E, E)$ sur $\mathcal{M}_{n, n}(K)$.

Démonstration : (i) Posons $\phi_{\mathcal{B}_0, \mathcal{B}_1}(u) = M_{u, \mathcal{B}_0, \mathcal{B}_1}$ pour $u \in \mathcal{L}(E, F)$. Comme $(\lambda u + \mu v)(x) = \lambda u(x) + \mu v(x)$ pour $\lambda, \mu \in K, u, v \in \mathcal{L}(E, F), x \in E$, il est clair que $\phi_{\mathcal{B}_0, \mathcal{B}_1}$ est linéaire. Si $M_{u, \mathcal{B}_0, \mathcal{B}_1} = 0$, soit $x \in E$, et soit Y le vecteur colonne formé des coordonnées de $u(x)$ dans la base \mathcal{B}_1 . Il résulte de la formule (2.1) que $Y = 0$, donc $u(x) = 0$ pour tout $x \in E$ et $u = 0$, ce qui prouve que $\phi_{\mathcal{B}_0, \mathcal{B}_1}$ est injective.

Notons e_1, \dots, e_n les éléments de la base \mathcal{B}_0 , notons f_1, \dots, f_m les éléments de la base \mathcal{B}_1 et soit $M = (m_{i,j})_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}} \in \mathcal{M}_{m, n}(K)$. Pour $x = \sum_{j=1}^n \alpha_j e_j \in E$, posons $u(x) = \sum_{i=1}^m \left(\sum_{j=1}^n \alpha_j m_{i,j} \right) f_i$. Comme l'application qui à $x \in E$ associe sa j^e coordonnée α_j dans la base \mathcal{B}_0 est linéaire pour $1 \leq j \leq n$, on voit que u est une application linéaire de E dans F . On a $u(e_k) = \sum_{i=1}^m \left(\sum_{j=1}^n \delta_{k,j} m_{i,j} \right) f_i = \sum_{i=1}^m m_{i,k} f_i$ pour $1 \leq k \leq m$. Donc toutes les colonnes de $M_{u, \mathcal{B}_0, \mathcal{B}_1}$ et M coïncident, et $M_{u, \mathcal{B}_0, \mathcal{B}_1} = M$, ce qui prouve que $\phi_{\mathcal{B}_0, \mathcal{B}_1}$ est surjective. Donc $\phi_{\mathcal{B}_0, \mathcal{B}_1}$ est bien un isomorphisme linéaire de $\mathcal{L}(E, F)$ sur $\mathcal{M}_{m, n}(K)$.

(ii) On sait déjà que $\phi_{\mathcal{B}_0, \mathcal{B}_0}$ est une application bijective de $\mathcal{L}(E, E)$ sur $\mathcal{M}_{n, n}(K)$, et que $\phi_{\mathcal{B}_0, \mathcal{B}_0}(u + v) = \phi_{\mathcal{B}_0, \mathcal{B}_0}(u) + \phi_{\mathcal{B}_0, \mathcal{B}_0}(v)$ pour $u, v \in \mathcal{L}(E, E)$. Soit $\mathcal{I} : x \mapsto x$ l'application identité sur E , qui est l'élément unité de $\mathcal{L}(E, E)$, et soit I_n la matrice unité introduite à la proposition 2.1.1. On a $\phi_{\mathcal{B}_0, \mathcal{B}_0}(\mathcal{I}) = I_n$, et il résulte de la formule (2.2) que $\phi_{\mathcal{B}_0, \mathcal{B}_0}(u \circ v) = \phi_{\mathcal{B}_0, \mathcal{B}_0}(u) \phi_{\mathcal{B}_0, \mathcal{B}_0}(v)$ pour $u, v \in \mathcal{L}(E, E)$. ♣

Corollaire 2.2.6 Si $\dim(E) = m$ et $\dim(F) = n$, alors $\dim(\mathcal{L}(E, F)) = mn$.

Corollaire 2.2.7 Soit E un espace vectoriel de dimension $n \geq 1$ et soit \mathcal{B} une base de E .

(i) Pour que $u \in \mathcal{L}(E, E)$ soit bijective, il faut et il suffit que la matrice $M_{u, \mathcal{B}, \mathcal{B}}$ soit inversible, et dans ce cas $M_{u^{-1}, \mathcal{B}, \mathcal{B}} = M_{u, \mathcal{B}, \mathcal{B}}^{-1}$.

(ii) On a $M_{u^n, \mathcal{B}, \mathcal{B}} = M_{u, \mathcal{B}, \mathcal{B}}^n$ pour $u \in \mathcal{L}(E, E), n \geq 1$.

2.3 Formule de changement de base pour les coordonnées de vecteurs

On va maintenant étudier la question du *changement de base* sur un espace vectoriel de dimension finie.

Définition 2.3.1 Soit E un espace vectoriel de dimension $n \geq 1$, soit $\mathcal{B} = (e_1, \dots, e_n)$ une base de E et soit $\mathcal{B}' = (e'_1, \dots, e'_p)$ une famille de p éléments de E . On appelle matrice représentant \mathcal{B}' dans la base \mathcal{B} la matrice $P_{\mathcal{B}, \mathcal{B}'} := (p_{i,j})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq p}}$, où $e'_j = \sum_{i=1}^n p_{i,j} e_i$ pour $1 \leq i \leq n, 1 \leq j \leq p$.

tout $x \in E$, en appliquant cette formule pour $x = e'_j$ on obtient
$$\begin{bmatrix} p_{1,j} \\ \cdot \\ \cdot \\ \cdot \\ p_{n,j} \end{bmatrix} =$$

$$A \begin{bmatrix} \delta_{1,j} \\ \cdot \\ \cdot \\ \cdot \\ \delta_{n,j} \end{bmatrix} = \begin{bmatrix} a_{1,j} \\ \cdot \\ \cdot \\ \cdot \\ a_{n,j} \end{bmatrix} \text{ pour } 1 \leq j \leq n, \text{ et par conséquent } A = P_{\mathcal{B},\mathcal{B}'}$$

Supposons de nouveau que \mathcal{B}' est une base de E . Avec les notations précédentes on a $X = P_{\mathcal{B},\mathcal{B}'}X' = P_{\mathcal{B},\mathcal{B}'}P_{\mathcal{B}',\mathcal{B}}X$ et $X' = P_{\mathcal{B}',\mathcal{B}}X = P_{\mathcal{B}',\mathcal{B}}P_{\mathcal{B},\mathcal{B}'}X'$. Donc $P_{\mathcal{B},\mathcal{B}'}P_{\mathcal{B}',\mathcal{B}} = P_{\mathcal{B},\mathcal{B}} = I_n$ et $P_{\mathcal{B}',\mathcal{B}}P_{\mathcal{B},\mathcal{B}'} = P_{\mathcal{B}',\mathcal{B}'} = I_n$. Ceci montre que $P_{\mathcal{B},\mathcal{B}'}$ est inversible et que $P_{\mathcal{B}',\mathcal{B}} = P_{\mathcal{B},\mathcal{B}'}^{-1}$. ♣

Exemple 2.3.3 1) Soit $\mathcal{B} = \left(\begin{bmatrix} a_{1,1} \\ \cdot \\ \cdot \\ \cdot \\ a_{m,1} \end{bmatrix}, \dots, \begin{bmatrix} a_{1,n} \\ \cdot \\ \cdot \\ \cdot \\ a_{m,n} \end{bmatrix} \right)$ une famille de n éléments de K^m , et soit \mathcal{B}_0 la base canonique de K^m . Alors $P_{\mathcal{B}_0,\mathcal{B}} = (a_{i,j})_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}}$.

2) La famille $\mathcal{B} = (x(x-1), x(x+1), (x+1)(x-1))$ est une base de $\mathbf{R}_2[x]$, et si on pose $\mathcal{B}_0 = (1, x, x^2)$ on a

$$P_{\mathcal{B}_0,\mathcal{B}} = \begin{bmatrix} 0 & 0 & -1 \\ -1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix}, P_{\mathcal{B},\mathcal{B}_0} = \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ -1 & 0 & 0 \end{bmatrix}.$$

En effet si $X \in K^m$, le vecteur colonne formé des coordonnées de X dans la base canonique de K^m est égal à X , ce qui donne la formule du premier exemple. Dans le deuxième exemple le calcul de $P_{\mathcal{B}_0,\mathcal{B}}$ est immédiat. Si $\beta_1x(x-1) + \beta_2x(x-1) + \beta_3(x+1)(x-1) = 0$ on obtient $\beta_1 = 0$ en remplaçant x par -1 , $\beta_2 = 0$ en remplaçant x par 1 et $\beta_3 = 0$ en remplaçant x par 0 , ce qui montre que \mathcal{B} est libre. Comme $\mathbf{R}_2(x)$ est de dimension 3, \mathcal{B} est une base de $\mathbf{R}_2(x)$.

Pour calculer $P_{\mathcal{B},\mathcal{B}_0}$ le plus simple est de résoudre l'équation $\beta_1x(x-1) + \beta_2x(x+1) + \beta_3(x+1)(x-1) = \alpha_0 + \alpha_1x + \alpha_2x^2$. On obtient $\beta_1 = \frac{1}{2}(\alpha_0 - \alpha_1 + \alpha_2)$ en remplaçant x par -1 , $\beta_2 = \frac{1}{2}(\alpha_0 + \alpha_1 + \alpha_2)$ en remplaçant x par 1 et $\beta_3 = -\alpha_0$

en remplaçant x par 0 , ce qui donne bien $P_{\mathcal{B},\mathcal{B}_0} = \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ -1 & 0 & 0 \end{bmatrix}$.

2.4 Formule de changement de base pour les applications linéaires

On va maintenant étudier l'effet d'un changement de bases sur la matrice représentant une application linéaire.

Proposition 2.4.1 Soient E et F deux espaces vectoriels non nuls de dimension finie, soient \mathcal{B}_0 et \mathcal{B}'_0 deux bases de E , soient \mathcal{B}_1 et \mathcal{B}'_1 deux bases de F et soit $u : E \rightarrow F$ une application linéaire. On a alors

$$(2.6) \quad M_{u, \mathcal{B}'_0, \mathcal{B}'_1} = P_{\mathcal{B}'_1, \mathcal{B}'_0}^{-1} M_{u, \mathcal{B}_0, \mathcal{B}_1} P_{\mathcal{B}_0, \mathcal{B}'_0}$$

Démonstration : Soit $x \in E$, soit X le vecteur colonne formé des coordonnées de x dans la base \mathcal{B}_0 , soit X' le vecteur colonne formé des coordonnées de x dans la base \mathcal{B}'_0 , soit Y le vecteur colonne formé des coordonnées de $u(x)$ dans la base \mathcal{B}_1 , et soit Y' le vecteur colonne formé des coordonnées de $u(x)$ dans la base \mathcal{B}'_1 .

Il résulte des formules (2.1), (2.4) et (2.5) que $Y = M_{u, \mathcal{B}_0, \mathcal{B}_1} X$, $X = P_{\mathcal{B}_0, \mathcal{B}'_0} X'$, $Y' = P_{\mathcal{B}'_1, \mathcal{B}_1}^{-1} Y$, ce qui donne $Y' = P_{\mathcal{B}'_1, \mathcal{B}'_0}^{-1} M_{u, \mathcal{B}_0, \mathcal{B}_1} P_{\mathcal{B}_0, \mathcal{B}'_0} X'$, relation qui caractérise la matrice $M_{u, \mathcal{B}'_0, \mathcal{B}'_1}$. ♣

Corollaire 2.4.2 Soit E un espace vectoriel non nul de dimension finie, soient \mathcal{B} et \mathcal{B}' deux bases de E et soit $u : E \rightarrow E$ un endomorphisme de E . On a alors

$$(2.7) \quad M_{u, \mathcal{B}', \mathcal{B}'} = P_{\mathcal{B}', \mathcal{B}}^{-1} M_{u, \mathcal{B}, \mathcal{B}} P_{\mathcal{B}, \mathcal{B}'}$$

Exemple 2.4.3 Soit $\mathcal{B} = (x(x-1), x(x+1), (x+1)(x-1))$ la base de $\mathbf{R}_2(x)$ introduite à l'exemple 2.12 et soit $D : p \mapsto p'$ l'opérateur de dérivation sur $\mathbf{R}_2[x]$. On a alors

$$M_{D, \mathcal{B}, \mathcal{B}} = \begin{bmatrix} -\frac{3}{2} & -\frac{1}{2} & -1 \\ \frac{1}{2} & \frac{3}{2} & 1 \\ 1 & -1 & 0 \end{bmatrix}$$

En effet posons $\mathcal{B}_0 = (1, x, x^2)$. On a vu que $M_{D, \mathcal{B}_0, \mathcal{B}_0} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix}$, que

$$P_{\mathcal{B}_0, \mathcal{B}} = \begin{bmatrix} 0 & 0 & -1 \\ -1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix}, \text{ et que } P_{\mathcal{B}_0, \mathcal{B}}^{-1} = P_{\mathcal{B}, \mathcal{B}_0} = \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ -1 & 0 & 0 \end{bmatrix}.$$

On a donc

$$\begin{aligned} M_{D, \mathcal{B}, \mathcal{B}} &= \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ -1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & -1 \\ -1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix} \\ &= \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ -1 & 0 & 0 \end{bmatrix} \begin{bmatrix} -1 & 1 & 0 \\ 2 & 2 & 2 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} -\frac{3}{2} & -\frac{1}{2} & -1 \\ \frac{1}{2} & \frac{3}{2} & 1 \\ 1 & -1 & 0 \end{bmatrix} \end{aligned}$$

2.5 Chapitre 2 sous MUPAD

Le logiciel Mupad est bien adapté aux opérations sur les matrices abordées dans ce chapitre. On notera qu'on commence par la commande $M := \text{Dom} :: \text{Matrix}()$; On définit ensuite les matrices utilisées en déclarant $A := M([\dots], \dots, [\dots])$, $[\dots]$ désignant les lignes de la matrice.

A titre d'exemple on va calculer la matrice $M_{D, \mathcal{B}, \mathcal{B}}$ en utilisant la formule

$$2.7, \text{ avec } P_{\mathcal{B}, \mathcal{B}_0} = \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ -1 & 0 & 0 \end{bmatrix} \text{ et } M_{D, \mathcal{B}_0, \mathcal{B}_0} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix}.$$

```
M:=Dom::Matrix();
P:=M([ [0, 0, -1], [-1, 1, 0], [1, 1, 1]]);
A:=M([ [0, 1, 0], [0, 0, 2], [0, 0, 0]]);
P^(-1)*A*P;
```

$$\begin{array}{ccc} +- & & -+ \\ | & 0, 0, -1 & | \\ | & & | \\ | & -1, 1, 0 & | \\ | & & | \\ | & 1, 1, 1 & | \\ +- & & -+ \end{array}$$

$$\begin{array}{ccc} +- & & -+ \\ | & 0, 1, 0 & | \\ | & & | \\ | & 0, 0, 2 & | \\ | & & | \\ | & 0, 0, 0 & | \end{array}$$

$$\begin{array}{ccc} +- & & -+ \\ | & -3/2, -1/2, -1 & | \\ | & & | \\ | & 1/2, 3/2, 1 & | \\ | & & | \\ | & 1, -1, 0 & | \end{array}$$

On retrouve donc sans surprise le fait que $M_{D,\mathcal{B},\mathcal{B}} = \begin{bmatrix} -\frac{3}{2} & -\frac{1}{2} & -1 \\ \frac{1}{2} & \frac{3}{2} & 1 \\ 1 & -1 & 0 \end{bmatrix}$.

Nous concluons ce Chapitre en montrant comment on peut utiliser Mupad pour calculer le carré et l'inverse d'une matrice un peu plus grande que les matrices ci-dessus.

```
M:=Dom::Matrix();
P:= M([ [1,6,8,9, 13], [11,4,77,9,2], [12,8,9,0,5], [6,8,9,11,0], [5,4,3,3,1]]);
P^2;
P^(-1);
```

```
Dom::Matrix(Dom::ExpressionField(id, iszero))
```

```
+ - - +
| 1, 6, 8, 9, 13 |
|
| 11, 4, 77, 9, 2 |
|
| 12, 8, 9, 0, 5 |
|
| 6, 8, 9, 11, 0 |
|
| 5, 4, 3, 3, 1 |
+ - - +
```

$$\begin{array}{cccccc}
 +- & & & & & -+ \\
 | & 282, & 218, & 662, & 201, & 78 & | \\
 | & & & & & & | \\
 | & 1043, & 778, & 1176, & 240, & 538 & | \\
 | & & & & & & | \\
 | & 233, & 196, & 808, & 195, & 222 & | \\
 | & & & & & & | \\
 | & 268, & 228, & 844, & 247, & 139 & | \\
 | & & & & & & | \\
 | & 108, & 98, & 405, & 117, & 89 & | \\
 +- & & & & & & -+
 \end{array}$$

$$\begin{array}{cccccc}
 +- & & & & & & -+ \\
 | & 66/11665, & 291/11665, & -3112/11665, & -4143/11665, & 2824/2333 & | \\
 | & & & & & & | \\
 | & -3589/69990, & -3629/69990, & 35963/69990, & 6707/11665, & -12590/6999 & | \\
 | & & & & & & | \\
 | & -32/6999, & 71/6999, & 307/6999, & 104/2333, & -1261/6999 & | \\
 | & & & & & & | \\
 | & 1328/34995, & 553/34995, & -9241/34995, & -1983/11665, & 5567/6999 & | \\
 | & & & & & & | \\
 | & 2684/34995, & 169/34995, & -2128/34995, & -1724/11665, & 2081/6999 & |
 \end{array}$$

2.6 Exercices sur le Chapitre 2

exercice 1

On considère l'espace vectoriel \mathbf{R}^2 muni de sa base canonique $\mathcal{B} = (e_1, e_2)$. Soit u l'endomorphisme de \mathbf{R}^2 défini par $u(e_1) = 2e_1 + e_2$ et $u(e_2) = e_1 + e_2$.

- Déterminer la matrice de u dans la base \mathcal{B} .
- Soient $v_1 = e_1 - e_2$ et $v_2 = e_1 + e_2$. Montrer que (v_1, v_2) est une base de \mathbf{R}^2 et déterminer la matrice de u dans cette nouvelle base.

exercice 2

On considère l'application $f : \mathbf{R}^3 \rightarrow \mathbf{R}^3$ définie par la formule

$$f(x, y, z) = (-z + x + y, x - y, z - 2x) \quad \forall (x, y, z) \in \mathbf{R}^3.$$

- Montrer que f est un endomorphisme de \mathbf{R}^3 .
- Déterminer le noyau et calculer le rang de f . L'application f est-elle injective? surjective?

- c) Déterminer la matrice de f dans la base canonique $\mathcal{B} = (e_1, e_2, e_3)$ de \mathbf{R}^3 .
 d) On pose $u_1 = (1, 1, 2)$, $u_2 = (-1, 1, 0)$, $u_3 = (0, 1, -1)$. Montrer que $\mathcal{B}_1 = (u_1, u_2, u_3)$ est une base de \mathbf{R}^3 .
 Déterminer la matrice de f dans la base \mathcal{B}_1 .
 e) Déterminer $w \in \mathbf{R}^3$ tel que $\mathcal{B}_2 = (f(u_1), f(u_2), w)$ soit une base de \mathbf{R}^3 .
 Déterminer la matrice de f dans la base \mathcal{B}_2 .

exercice 3

On munit \mathbf{R}^3 de la base canonique $\mathcal{B} = (e_1, e_2, e_3)$.
 Pour $a \in \mathbf{R}$, on définit $f : \mathbf{R}^3 \rightarrow \mathbf{R}^3$ par les formules

$$f_a(e_1) = ae_1 + e_2 + e_3, f_a(e_2) = e_1 + ae_2 + e_3, f_a(e_3) = e_1 + e_2 + ae_3.$$

- a) Donner une base de $\text{Ker}(f_{-2})$ et une base de $\text{Im}(f_{-2})$.
 b) Donner une base de $\text{Ker}(f_1)$ et une base de $\text{Im}(f_1)$.
 c) Montrer que l'application f_0 est bijective.
 d) Montrer que $\mathbf{R}^3 = \text{Ker}(f_{-2}) \oplus \text{Ker}(f_1)$.
 e) Ecrire la matrice A de f_a dans la base \mathcal{B} . Appelons \mathcal{B}_1 la base de $\text{Ker}(f_{-2})$ trouvée au a), et \mathcal{B}_2 la base de $\text{Ker}(f_1)$ trouvée au b). Ecrire la matrice B de f_a dans la base $\mathcal{B}' = \mathcal{B}_1 \cup \mathcal{B}_2$. Donner une formule matricielle exprimant B en fonction de A .
 f) Calculer B^n pour $n \geq 0$.
 g) Montrer que $A^2 - (2a + 1)A + (a - 1)(a + 2)I = 0$, I désignant la matrice unité de $\mathcal{M}_3(\mathbf{R})$.
 h) Pour quelles valeurs de a la matrice A est-elle inversible ? Dans le cas où A est inversible, calculer A^{-1} .

exercice 4(sous Mupad)

- 1) Soit $\mathcal{B} = (1 + x - x^2 + x^3 - x^4, 7 - 4x + x^2 + x^3 + 2x^4, 8 - x + 3x^2 + 7x^3 + 4x^4, 2 + 3x + 7x^2 - 5x^3 + x^4, 4 + 7x - 5x^2 + 8x^3 + x^4)$.
 Montrer que \mathcal{B} est une base de $\mathbf{R}_4[x]$.
 2) Donner les coordonnées de $q = 1 + x + x^2 + x^3 + x^4$ dans la base \mathcal{B} .
 3) Soit $D : p \mapsto p'$ l'opérateur de dérivation sur $\mathbf{R}_4[x]$. Calculer $M_{D, \mathcal{B}, \mathcal{B}}$.

Chapitre 3

Déterminants, notations indicielles de la Physique

3.1 Propriétés des déterminants

Dans toute la suite du cours on notera $\mathcal{M}_n(K)$ l'anneau des matrices à n lignes et n colonnes à coefficients dans un corps K , on notera $\mathcal{L}(E)$ l'anneau des endomorphismes d'un K -espace vectoriel E de dimension finie, et on notera $M_{u,\mathcal{B}} = M_{u,\mathcal{B},\mathcal{B}}$ la matrice représentant $u \in L(E)$ dans une base \mathcal{B} de E .

On note \mathcal{S}_n le groupe des permutations d'ordre n , c'est à dire l'ensemble des bijections de l'ensemble $\{1, \dots, n\}$ sur lui même muni de la composition des applications. On adopte la notation suivante, pour $\sigma \in \mathcal{S}_n$:

$$\sigma = \begin{array}{|c|c|c|c|} \hline 1 & 2 & \dots & n \\ \hline \sigma(1) & \sigma(2) & \dots & \sigma(n) \\ \hline \end{array}$$

La *signature* d'une permutation σ est définie par la formule

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

On vérifie que $\epsilon(\sigma) \in \{-1, 1\}$ et que $\epsilon(\sigma \circ \tau) = \epsilon(\sigma)\epsilon(\tau)$ pour $\sigma \in \mathcal{S}_n, \tau \in \mathcal{S}_n$. Ceci permet d'introduire la définition suivante :

Définition 3.1.1 Soit $A = (a_{i,j})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq n}} \in \mathcal{M}_n(K)$. On définit le **déterminant** de A par la formule

$$\det(A) = \sum_{\sigma \in \mathcal{S}_n} \epsilon(\sigma) a_{\sigma(1),1} \dots a_{\sigma(n),n}.$$

On utilisera également la notation

$$\det \left(\begin{pmatrix} a_{1,1} & \cdots & a_{1,n} \\ \cdot & \cdots & \cdot \\ \cdot & \cdots & \cdot \\ \cdot & \cdots & \cdot \\ a_{n,1} & \cdots & a_{n,n} \end{pmatrix} \right) = \begin{vmatrix} a_{1,1} & \cdots & a_{1,n} \\ \cdot & \cdots & \cdot \\ \cdot & \cdots & \cdot \\ \cdot & \cdots & \cdot \\ a_{n,1} & \cdots & a_{n,n} \end{vmatrix}.$$

Les déterminants possèdent la propriété fondamentale suivante, que nous admettrons sans démonstration.

Théorème 3.1.2 : Soient $A, B \in \mathcal{M}_n(K)$. Alors $\det(AB) = \det(A)\det(B)$.

Le groupe \mathcal{S}_2 se réduit à deux éléments σ_1 et σ_2 , avec

$$\sigma_1 = \begin{array}{|c|c|} \hline 1 & 2 \\ \hline 1 & 2 \\ \hline \end{array} \quad \text{et} \quad \sigma_2 = \begin{array}{|c|c|} \hline 1 & 2 \\ \hline 2 & 1 \\ \hline \end{array}.$$

On a $\epsilon(\sigma_1) = 1$ et $\epsilon(\sigma_2) = -1$, donc

$\sum_{i=1}^2 a_{\sigma_i(1),1} a_{\sigma_i(2),2} = a_{1,1}a_{2,2} - a_{2,1}a_{1,2}$ et on obtient

$$(3.1) \quad \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc.$$

Le groupe \mathcal{S}_3 possède six éléments $\sigma_1, \sigma_2, \sigma_3, \sigma_4, \sigma_5$ et σ_6 , avec

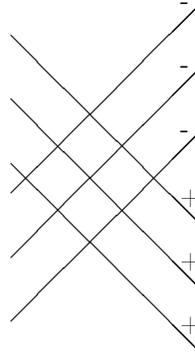
$$\begin{aligned} \sigma_1 &= \begin{array}{|c|c|c|} \hline 1 & 2 & 3 \\ \hline 1 & 2 & 3 \\ \hline \end{array}, & \sigma_2 &= \begin{array}{|c|c|c|} \hline 1 & 2 & 3 \\ \hline 2 & 3 & 1 \\ \hline \end{array}, & \sigma_3 &= \begin{array}{|c|c|c|} \hline 1 & 2 & 3 \\ \hline 3 & 1 & 2 \\ \hline \end{array}, \\ \sigma_4 &= \begin{array}{|c|c|c|} \hline 1 & 2 & 3 \\ \hline 3 & 2 & 1 \\ \hline \end{array}, & \sigma_5 &= \begin{array}{|c|c|c|} \hline 1 & 2 & 3 \\ \hline 1 & 3 & 2 \\ \hline \end{array}, & \sigma_6 &= \begin{array}{|c|c|c|} \hline 1 & 2 & 3 \\ \hline 2 & 1 & 3 \\ \hline \end{array}. \end{aligned}$$

On a $\epsilon(\sigma_1) = 1$, $\epsilon(\sigma_2) = \frac{(3-2)(1-2)(1-3)}{2} = 1$, $\epsilon(\sigma_3) = \frac{(1-3)(2-3)(2-1)}{2} = 1$, $\epsilon(\sigma_4) = \frac{(2-3)(1-3)(1-2)}{2} = -1$, $\epsilon(\sigma_5) = \frac{(3-1)(2-1)(2-3)}{2} = -1$, et $\epsilon(\sigma_6) = \frac{(1-2)(3-2)(3-1)}{2} = -1$. On obtient, si $A = (a_{i,j})_{\substack{1 \leq i \leq 3 \\ 1 \leq j \leq 3}} \in \mathcal{M}_3(K)$:

$$\det(A) = \sum_{i=1}^6 \epsilon(\sigma_i) a_{\sigma_i(1),1} a_{\sigma_i(2),2} a_{\sigma_i(3),3} = a_{1,1}a_{2,2}a_{3,3} + a_{2,1}a_{3,2}a_{1,3} + a_{3,1}a_{1,2}a_{2,3} - a_{3,1}a_{2,2}a_{1,3} - a_{1,1}a_{3,2}a_{2,3} - a_{2,1}a_{1,2}a_{3,3}.$$

On en déduit la règle de Sarrus

$$(3.2) \quad \begin{array}{c} a_{1,1} \quad a_{1,2} \quad a_{1,3} \\ \left| \begin{array}{ccc} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{array} \right| = \begin{array}{ccc} a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \\ a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \end{array} \end{array},$$



Autrement dit on recopie en dessous de la matrice ses deux premières lignes et on trouve le déterminant en calculant la somme obtenue en affectant les trois produits obtenus en descendant à 45° de gauche à droite du signe + et en affectant les trois produits obtenus en montant à 45° de gauche à droite du signe -. On obtient par exemple

$$\begin{vmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \\ 1 & 1 & 1 \end{vmatrix} = \begin{matrix} 1 & 2 & 3 & - \\ 3 & 2 & 1 & - \\ 1 & 1 & 1 & - \\ 1 & 2 & 3 & + \\ 3 & 2 & 1 & + \\ 1 & 1 & 1 & + \end{matrix}$$

Donc $\begin{vmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \\ 1 & 1 & 1 \end{vmatrix} = 2 + 9 + 2 - 6 - 1 - 6 = 0.$

On peut également *développer un déterminant suivant une ligne ou une colonne*. Si $A \in \mathcal{M}_n(K)$ on appellera *cofacteur d'indice (i, j)*, et on notera $C_{i,j}(A)$, le déterminant de la matrice à $(n-1)$ lignes et $(n-1)$ colonnes obtenu en retirant à A sa i^e ligne et sa j^e colonne. On a alors le théorème suivant.

Théorème 3.1.3 Soit $A = (a_{i,j})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq n}} \in \mathcal{M}_n(K).$

- 1) (Développement selon les lignes)
On a $\det(A) = \sum_{1 \leq j \leq n} a_{i,j} (-1)^{i+j} C_{i,j}(A)$ pour $1 \leq i \leq n.$
- 2) (Développement selon les colonnes)
On a $\det(A) = \sum_{1 \leq i \leq n} a_{i,j} (-1)^{i+j} C_{i,j}(A)$ pour $1 \leq j \leq n.$

Par exemple on a, en développant par rapport à la première colonne, $\begin{vmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \\ 1 & 1 & 1 \end{vmatrix} =$

$$\begin{vmatrix} 2 & 1 \\ 1 & 1 \end{vmatrix} - 3 \begin{vmatrix} 2 & 3 \\ 1 & 1 \end{vmatrix} + \begin{vmatrix} 2 & 3 \\ 2 & 1 \end{vmatrix} = 1 + 3 - 4 = 0.$$

On dira que $A = (a_{i,j})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq n}} \in \mathcal{M}_n(K)$ est *triangulaire supérieure* si $a_{i,j} = 0$ pour $i > j$, on dira que A est *triangulaire inférieure* si $a_{i,j} = 0$ pour $i < j$, et on dira que A est *triangulaire* si elle est triangulaire supérieure ou inférieure. On déduit du théorème 3.3 le résultat suivant.

Corollaire 3.1.4 Si $A = (a_{i,j})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq n}} \in \mathcal{M}_n(K)$ est triangulaire, $\det(A) = a_{1,1} \dots a_{n,n}.$

Démonstration : C'est évident si $n = 1$. Supposons la propriété vérifiée pour n , et soit $A = (a_{i,j})_{\substack{1 \leq i \leq n+1 \\ 1 \leq j \leq n+1}} \in \mathcal{M}_{n+1}(K)$ une matrice triangulaire. En développant par rapport à la première ligne ou la première colonne de A , on obtient $\det(A) = a_{1,1}C_{1,1}(A)$, et $C_{1,1}(A) = a_{2,2} \dots a_{n,n}$ d'après l'hypothèse de récurrence. Donc la propriété est vérifiée pour tout $n \geq 1$. ♣

La Proposition suivante résume quelques propriétés classiques des déterminants.

Proposition 3.1.5 Soit $A \in \mathcal{M}_n(K)$.

1) Si on échange deux colonnes de A , ou si on échange deux lignes de A , le déterminant de A est multiplié par -1 .

2) L'application $X \mapsto \det(A^{(i)}(X))$, où $A^{(i)}(X)$ désigne la matrice obtenue en remplaçant la i^{e} ligne de A par un vecteur ligne X , est une application linéaire de K^n dans K pour $1 \leq i \leq n$.

3) L'application $Y \mapsto \det(A_{(j)}(Y))$, où $A_{(j)}(Y)$ désigne la matrice obtenue en remplaçant la j^{e} colonne de A par Y , est une application linéaire de K^n dans K pour $1 \leq j \leq n$.

4) Si deux lignes de A sont égales, ou si deux colonnes de A sont égales, le déterminant de A est nul.

5) Le déterminant ne change pas si on ajoute à une ligne de A une combinaison linéaire des autres, ou si on ajoute à une colonne de A une combinaison linéaire des autres.

Ces résultats sont d'un grand intérêt pratique pour le calcul des déterminants.

Considérons par exemple le déterminant suivant :

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

On retranche la première ligne à la 2^e et à la 3^e, on retranche Deux fois la première ligne à la quatrième et 5 fois la première ligne à la cinquième.

$$L_2 \longrightarrow L_2 - L_1, L_3 \longrightarrow L_3 - L_1, L_4 \longrightarrow L_4 - 2L_1, L_5 \longrightarrow L_5 - 5L_1$$

$$\begin{aligned} D &= \begin{vmatrix} 1 & 2 & 3 & 4 & 5 \\ 0 & 0 & -2 & -2 & -4 \\ 0 & -3 & 1 & -1 & -6 \\ 0 & -5 & -3 & -7 & -6 \\ 0 & -12 & -12 & -21 & -22 \end{vmatrix} = \begin{vmatrix} 0 & -2 & -2 & -4 \\ -3 & 1 & -1 & -6 \\ -5 & -3 & -7 & -6 \\ -12 & -12 & -21 & -22 \end{vmatrix} \\ &= -2 \begin{vmatrix} 0 & 1 & 1 & 2 \\ -3 & 1 & -1 & -6 \\ -5 & -3 & -7 & -6 \\ -12 & -12 & -21 & -22 \end{vmatrix} = -4 \begin{vmatrix} 0 & 1 & 1 & 1 \\ -3 & 1 & -1 & -3 \\ -5 & -3 & -7 & -3 \\ -12 & -12 & -21 & -11 \end{vmatrix} \end{aligned}$$

On retranche la dernière colonne à la deuxième et à la troisième.

$$C_2 \longrightarrow C_2 - C_4, C_3 \longrightarrow C_3 - C_4.$$

$$D = -4 \begin{vmatrix} 0 & 0 & 0 & 1 \\ -3 & 4 & 2 & -3 \\ -5 & 0 & -4 & -3 \\ -12 & -1 & -10 & -11 \end{vmatrix} = 4 \begin{vmatrix} -3 & 4 & 2 \\ -5 & 0 & -4 \\ -12 & -1 & -10 \end{vmatrix} = 8 \begin{vmatrix} -3 & 4 & 1 \\ -5 & 0 & -2 \\ -12 & -1 & -5 \end{vmatrix}.$$

$$C_1 \longrightarrow C_1 + 3C_3, C_2 \longrightarrow C_2 - 4C_3.$$

$$D = 8 \begin{vmatrix} 0 & 0 & 1 \\ -11 & 8 & -2 \\ -27 & 19 & -5 \end{vmatrix} = 8 \begin{vmatrix} -11 & 8 \\ -27 & 19 \end{vmatrix} = 8(-209 + 216) = 56.$$

3.2 Matrices carrées inversibles

Le théorème 3.3 permet d'obtenir des formules pour calculer l'inverse d'une matrice. Rappelons qu'on dit qu'un élément b d'un anneau B est *inversible à gauche* s'il existe $c \in B$ tel que $cb = e$, e désignant l'élément unité de B pour la multiplication. De même on dit que b est *inversible à droite* s'il existe $c \in B$ tel que $bc = e$.

Définition 3.2.1 Soit $B = (b_{i,j})_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}} \in \mathcal{M}_{m,n}(K)$. On appelle **transposée de B** , la matrice $B^t = (b_{j,i})_{\substack{1 \leq j \leq n \\ 1 \leq i \leq m}}$ obtenue en échangeant les lignes et les colonnes de B .

On a les propriétés suivantes

$$(3.3) \quad (B^t)^t = B \quad \forall B \in \mathcal{M}_{m,n}(K).$$

$$(3.4) \quad (\beta B + \gamma C)^t = \beta B^t + \gamma C^t \quad \forall \beta, \gamma \in K, \forall B, C \in \mathcal{M}_n(K).$$

$$(3.5) \quad (BC)^t = C^t B^t \quad \forall B \in \mathcal{M}_{m,n}(K), \forall C \in \mathcal{M}_{n,p}(K).$$

$$(3.6) \quad \det(B^t) = \det(B) \quad \forall B \in \mathcal{M}_n(K).$$

On a le résultat suivant.

Théorème 3.2.2 Soit $A \in \mathcal{M}_n(K)$. Les conditions suivantes sont équivalentes

- 1) A est inversible à droite
- 2) A est inversible à gauche
- 3) A est inversible
- 4) $\det(A) \neq 0$.

D'autre part on a dans ce cas

$$A^{-1} = \frac{1}{\det(A)} (C(A))^t,$$

où $C(A)$ est la **matrice des cofacteurs de A** définie par la formule

$$C(A) = ((-1)^{i+j} C_{i,j}(A))_{\substack{1 \leq i \leq n \\ 1 \leq j \leq n}}.$$

Démonstration : On a $\det(I_n) = 1$, donc si A possède un inverse à gauche B on a $\det(B)\det(A) = \det(I_n) = 1$ et $\det(A) \neq 0$. De même $\det(A) \neq 0$ si A est inversible à gauche. Supposons maintenant que $\det(A) \neq 0$. On a

$$\sum_{k=1}^n a_{i,k} (-1)^{k+j} C_{j,k}(A) = \det(A^{(i,j)}),$$

où $A^{(i,j)}$ est la matrice obtenue en remplaçant la j^e ligne de A par la i^e , car ce changement n'affecte pas les cofacteurs $C_{j,k}$ pour $1 \leq k \leq n$. Si $i \neq j$, les i^e et j^e lignes de la matrice $A^{(i,j)}$ sont égales, donc $\det(A^{(i,j)}) = 0$. Si $i = j$, alors $A^{(i,j)} = A^{(i,i)} = A$. On voit donc que $A.(C(A))^t = \det(A)I_n$.

D'autre part

$$\sum_{k=1}^n a_{k,j} (-1)^{k+i} C_{k,i}(A) = \det(A_{(i,j)}),$$

où $A_{(i,j)}$ désigne la matrice obtenue en remplaçant la i^e colonne de A par la j^e . On a donc $\det(A_{(i,j)}) = 0$ pour $i \neq j$, $\det(A_{(i,j)}) = \det(A)$ pour $i = j$, et on obtient

$$(3.7) \quad A.(C(A))^t = (C(A))^t.A = \det(A)I_n \quad \forall A \in \mathcal{M}_n(K).$$

Donc si $\det(A) \neq 0$ alors A est inversible, et $A^{-1} = \frac{1}{\det(A)} (C(A))^t A$. Comme toute matrice inversible est à fortiori inversible à droite et à gauche, ceci achève la démonstration. ♣

Pour les matrices à deux lignes et deux colonnes on obtient

$$(3.8) \quad \begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad-bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} \quad \text{si } ad-bc \neq 0.$$

3.3 Formules de Cramer

On obtient aussi les formules de Cramer.

Théorème 3.3.1 *On considère le système linéaire de n équations à n inconnues*

$$\begin{cases} a_{1,1}x_1 + \dots + a_{1,n}x_n = b_1 \\ \vdots \\ a_{n,1}x_1 + \dots + a_{n,n}x_n = b_n \end{cases}$$

Si $\begin{vmatrix} a_{1,1} & \dots & a_{1,n} \\ \vdots & \dots & \vdots \\ a_{n,1} & \dots & a_{n,n} \end{vmatrix} \neq 0$, le système admet une solution unique (x_1, \dots, x_n)

donnée par les formules

$$x_j = \frac{\det A_{j,b_1,\dots,b_n}}{\det(A)} \text{ pour } 1 \leq j \leq n,$$

où $A = (a_{i,j})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq n}}$, et où A_{j,b_1,\dots,b_n} est la matrice obtenue en remplaçant

la j^{e} colonne de A par $\begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}$.

D'autre part si $\begin{vmatrix} a_{1,1} & \dots & a_{1,n} \\ \vdots & \dots & \vdots \\ a_{n,1} & \dots & a_{n,n} \end{vmatrix} = 0$, alors ou bien le système n'admet

aucune solution, ou bien le système admet plusieurs solutions (en fait une infinité de solutions si le corps K est infini).

Démonstration : Posons $X = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$, $B = \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}$. Le système est équivalent

à l'équation $AX = B$. Si $\det(A) \neq 0$, A est inversible et cette équation admet pour solution unique $X = A^{-1}B$.

On a alors $x_j = \frac{1}{\det(A)} \sum_{1 \leq k \leq n} (1)^{j+k} C_{k,j} b_k$, et $\sum_{1 \leq k \leq n} (1)^{j+k} C_{k,j} b_k$ coïncide avec le déterminant de la matrice obtenue en remplaçant la j^{e} colonne de A par B .

Considérons maintenant l'application linéaire $U : K^n \rightarrow K^n$ définie par l'équation $U(X) = AX$. Si $\det(A) = 0$, A n'est pas inversible. Il résulte alors de la proposition 2.6 que U n'est pas bijective, et il résulte du corollaire 1.18 que U n'est ni bijective, ni surjective. Si $B \notin \text{Im}(U)$, le système n'admet aucune solution. Si $B \in \text{Im}(U)$, il existe $X \in K^n$ tel que $AX = U(X) = B$. Soit Y un

élément non nul de $\text{Ker}(U)$. On a $A(X + \lambda Y) = U(X) + \lambda U(Y) = U(X) = B$ pour tout $\lambda \in K$, ce qui fait que le système admet plusieurs solutions (et une infinité de solutions si le corps K est infini). ♣

Considérons par exemple la matrice $A = \begin{pmatrix} 10 & 5 & 4 \\ 5 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}$.

On a par la règle de Sarrus $\det(A) = 40 + 20 + 45 - 24 - 30 - 50 = 1$.

$$\begin{aligned} \text{D'autre part } C_{1,1}(A) &= \begin{vmatrix} 2 & 3 \\ 1 & 2 \end{vmatrix} = 1, & C_{1,2}(A) &= \begin{vmatrix} 5 & 3 \\ 3 & 2 \end{vmatrix} = 1, & C_{1,3}(A) &= \begin{vmatrix} 5 & 2 \\ 3 & 1 \end{vmatrix} = -1, \\ C_{2,1}(A) &= \begin{vmatrix} 5 & 4 \\ 1 & 2 \end{vmatrix} = 6, & C_{2,2}(A) &= \begin{vmatrix} 10 & 4 \\ 3 & 2 \end{vmatrix} = 8, & C_{2,3}(A) &= \begin{vmatrix} 10 & 5 \\ 3 & 1 \end{vmatrix} = -5, \\ C_{3,1}(A) &= \begin{vmatrix} 5 & 4 \\ 2 & 3 \end{vmatrix} = 7, & C_{3,2}(A) &= \begin{vmatrix} 10 & 4 \\ 5 & 3 \end{vmatrix} = 10, & C_{3,3}(A) &= \begin{vmatrix} 10 & 5 \\ 5 & 2 \end{vmatrix} = -5. \end{aligned}$$

$$\text{Donc } C(A) = \begin{pmatrix} 1 & -1 & -1 \\ -6 & 8 & -5 \\ 7 & -10 & -5 \end{pmatrix} \text{ et } A^{-1} = (C(A))^t = \begin{pmatrix} 1 & -6 & 7 \\ -1 & 8 & -10 \\ -1 & 5 & -5 \end{pmatrix}.$$

Considérons maintenant le système

$$\begin{cases} 10x + 5y + 4z = 1 \\ 5x + 2y + 3z = 2 \\ 3x + y + 2z = -1 \end{cases}$$

Comme $\det(A) = 1$, on peut appliquer les formules de Cramer, et le système possède une solution unique donnée par les formules

$$x = \begin{vmatrix} 1 & 5 & 4 \\ 2 & 2 & 3 \\ -1 & 1 & 2 \end{vmatrix}, \quad y = \begin{vmatrix} 10 & 1 & 4 \\ 5 & 2 & 3 \\ 3 & -1 & 2 \end{vmatrix}, \quad z = \begin{vmatrix} 10 & 5 & 1 \\ 5 & 2 & 2 \\ 3 & 1 & -1 \end{vmatrix}.$$

En appliquant la règle de Sarrus on obtient $x = 4 + 8 - 15 + 8 - 3 - 20 = -18$, $y = 40 - 20 + 9 - 24 + 30 - 10 = 25$, $z = -20 + 5 + 30 - 6 - 20 + 25 = 14$.

En fait comme A^{-1} a déjà été calculé, on pouvait écrire directement

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = A^{-1} \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix} = \begin{pmatrix} 1 & -6 & 7 \\ -1 & 8 & -10 \\ -1 & 5 & -5 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix} = \begin{pmatrix} -18 \\ 25 \\ 14 \end{pmatrix}.$$

3.4 Méthode du pivot de Gauss

En pratique le calcul des cofacteurs pour obtenir l'inverse d'une matrice est long et fastidieux (25 déterminants d'ordre 4 si $A \in \mathcal{M}_5(K)$), et les formules

de Cramer peuvent aussi mener à des calculs assez longs pour résoudre des systèmes linéaires. On est plutôt amené à utiliser la **méthode du pivot de Gauss**, décrite dans le **Cours d'Analyse numérique** de X. Fischer.[5] Cette méthode, qui consiste à "triangler le système", permet aussi de résoudre des systèmes où le nombre d'équations est différent du nombre d'inconnues. Nous décrivons cette méthode en reprenant la matrice précédente.

On va résoudre directement l'équation $AX = B$.

$$\begin{cases} 10x + 5y + 4z = a \\ 5x + 2y + 3z = b \\ 3x + y + 2z = c \end{cases}$$

Pour simplifier les calculs on écrit en premier la variable y et on échange les premières et troisièmes équations.

$$\begin{cases} y + 3x + 2z = c \\ 2y + 5x + 3z = b \\ 5y + 10x + 4z = a \end{cases}$$

$$L_2 \longrightarrow L_2 - 2L_1, L_3 \longrightarrow L_3 - 5L_1.$$

$$\begin{cases} y + 3x + 2z = c \\ -x - z = b - 2c \\ -5x - 6z = a - 5c \end{cases}$$

$$L_3 \longrightarrow L_3 - 5L_2.$$

$$\begin{cases} y + 3x + 2z = c \\ -x - z = b - 2c \\ -z = a - 5b + 5c \end{cases}$$

La dernière équation donne $z = -a + 5b - 5c$. En reportant dans la deuxième on obtient $x = -z - b + 2c = a - 5b + 5c - b + 2c = a - 6b + 7c$. Enfin en reportant dans la première équation on obtient $y = -3x - 2z + c = -3a + 18b - 21c + 2a - 10b + 10c + c = -a + 8b - 10c$.

Le système admet donc une solution unique :

$$\begin{cases} x = a - 6b + 7c \\ y = -a + 8b - 10c \\ z = -a + 5b - 5c \end{cases}$$

Pour $a = 1$, $b = 2$, $c = 3$ on retrouve le fait que $x = 1 - 12 - 7 = -18$, $y = -1 + 16 + 10 = 25$, $z = -1 + 10 + 5 = 14$.

D'autre part on a $\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 & -6 & 7 \\ -1 & 8 & -10 \\ -1 & 5 & -5 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix}$, et on retrouve le fait

$$\text{que } \begin{pmatrix} 10 & 5 & 4 \\ 5 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & -6 & 7 \\ -1 & 8 & -10 \\ -1 & 5 & -5 \end{pmatrix}.$$

3.5 Déterminant d'un endomorphisme et d'une famille de vecteurs

On peut également définir la notion de déterminant d'un endomorphisme d'un espace vectoriel de dimension finie.

Proposition 3.5.1 *Soit E un espace vectoriel non nul de dimension finie, et soit \mathcal{B} une base de E . Pour $u \in \mathcal{L}(E)$, on pose*

$$\det(u) = \det(M_{u,\mathcal{B}}).$$

Cette définition est indépendante du choix de la base \mathcal{B} , et u est inversible si et seulement si $\det(u) \neq 0$. D'autre part $\det(u \circ v) = \det(u)\det(v)$ pour $u, v \in \mathcal{L}(E)$.

Démonstration : Si \mathcal{B}_1 est une autre base de E , on a $M_{u,\mathcal{B}_1} = P^{-1}M_{u,\mathcal{B}}P$, où $P = P_{\mathcal{B},\mathcal{B}_1}$ est la matrice de passage de \mathcal{B} à \mathcal{B}_1 . Donc $\det(M_{u,\mathcal{B}_1}) = \det(P^{-1})\det(M_{u,\mathcal{B}})\det(P) = \det(P)^{-1}\det(P)\det(M_{u,\mathcal{B}}) = \det(I_n)\det(M_{u,\mathcal{B}}) = \det(M_{u,\mathcal{B}})$, et la définition de $\det(u)$ est indépendante du choix de la base \mathcal{B} . Le fait que $\det(u \circ v) = \det(u)\det(v)$ pour $u, v \in \mathcal{L}(E)$ est alors une conséquence immédiate de la proposition 2.6 et du théorème 3.2. ♣

D'autre part on peut définir le déterminant d'une famille \mathcal{F} de n éléments d'un espace vectoriel E de dimension finie $n \geq 1$ dans une base donnée \mathcal{B}_0 de E .

Proposition 3.5.2 *Soit E un espace vectoriel de dimension finie $n \geq 1$ sur un corps K , soit \mathcal{B}_0 une base de E , et soit $\mathcal{F} = \{f_1, \dots, f_n\}$ une famille de n éléments de E . On définit le déterminant $\det_{\mathcal{B}_0}(\mathcal{F})$ de \mathcal{F} dans la base \mathcal{B}_0 par la formule*

$$\det_{\mathcal{B}_0}(\mathcal{F}) = \det(P_{\mathcal{B}_0,\mathcal{F}}).$$

Alors \mathcal{F} est une base de E si et seulement si $\det_{\mathcal{B}_0}(\mathcal{F}) \neq 0$.

De plus l'application $\phi : E^n \rightarrow K$ définie par la formule $\phi(\mathcal{F}) = \det_{\mathcal{B}_0}(\mathcal{F})$ est une "forme multilinéaire alternée" de E^n dans K :

si $\mathcal{F}_{i,j}$ désigne la famille obtenue en échangeant f_i et f_j , avec $1 \leq i < j \leq n$, alors $\phi(\mathcal{F}_{i,j}) = -\phi(\mathcal{F})$, et si $\mathcal{F}(i,x)$ désigne la famille obtenue en remplaçant f_i par $x \in E$, alors l'application $x \rightarrow \phi(\mathcal{F}_{i,x})$ est une application linéaire de E dans K pour $1 \leq i \leq n$.

Démonstration : Ceci résulte de la définition de la matrice $P_{\mathcal{B}_0,\mathcal{F}}$, de la proposition 2.10 et de la proposition 3.5. ♣

3.6 Rang d'une matrice

Les déterminants sont également utiles pour calculer le *rang* d'une matrice, qui se définit de la manière suivante.

Définition 3.6.1 Soit $B = (b_{i,j})_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}} \in \mathcal{M}_n(K)$. On appelle **rang** de B , et on note $rg(B)$, la dimension du sous-espace vectoriel de K^n engendré par les colonnes de B .

Soient p, m, n trois entiers, avec $1 \leq p \leq \inf(m, n)$. On dit qu'une matrice $A \in \mathcal{M}_p(K)$ est une matrice extraite d'une matrice $B \in \mathcal{M}_{m,n}(K)$ si A est une matrice obtenue en retirant à B $m - p$ de ses lignes et $n - p$ de ses colonnes. On a alors le résultat suivant.

Proposition 3.6.2 Soit $B \in \mathcal{M}_{m,n}(K)$ une matrice non nulle. Alors le rang de B est égal au plus grand entier $p \geq 1$ pour lequel il existe une matrice à p lignes et p colonnes extraite de B dont le déterminant soit non nul. En particulier le rang de B est égal à celui de sa transposée, c'est à dire à la dimension du sous-espace vectoriel de K^m engendré par les lignes de B .

Démonstration : Notons B_1, \dots, B_n les colonnes de B , soit E le sous-espace vectoriel de K^n engendré par les colonnes de B et soit p le rang de B . Il existe des entiers $i_1 < i_2 \dots < i_p$ tels que $(B_{i_1}, \dots, B_{i_p})$ soit une base de E . Soit $C = (c_{i,j})_{\substack{1 \leq i \leq m \\ 1 \leq j \leq p}} \in \mathcal{M}_{m,p}(K)$ la matrice dont les colonnes sont B_{i_1}, \dots, B_{i_p} , soit F le sous-espace vectoriel de K^p engendré par les lignes L_1, \dots, L_m de C , et soit $q \leq p$ la dimension de F . Il existe un sous-ensemble S de $\{1, \dots, m\}$ possédant q éléments tel que $(L_j)_{j \in S}$ soit une base de F .

Soit $X = \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix} \in K^p$ tel que $L_j X = 0$ pour $j \in S$, et soit $k \leq m$. Il

existe une famille $(\lambda_j)_{j \in S}$ d'éléments de K telle que $L_k = \sum_{j \in S} \lambda_j L_j$. On a alors $L_k X = \sum_{j \in S} \lambda_j L_j X = 0$, ce qui prouve que $CX = 0$. Comme les colonnes de C forment une famille libre, on a $X = 0$. Soient j_1, \dots, j_q les éléments de S , et soit $D \in \mathcal{M}_{q,p}(K)$ la matrice dont les lignes sont L_{j_1}, \dots, L_{j_q} . Il résulte de ce qui précède que les p colonnes de D forment une famille libre de K^q , ce qui prouve que $q \geq p$. Donc $q = p$. Comme l'équation $DX = 0$ admet 0 pour unique solution dans K^p , on a $\det(D) \neq 0$, et D est une matrice à p lignes et p colonnes extraite de B dont le déterminant est non nul.

Supposons maintenant qu'il existe une matrice U à r lignes et r colonnes extraite de B telle que $\det(U) \neq 0$. Il existe un sous-ensemble R de $\{1, \dots, m\}$ possédant r éléments et un sous-ensemble T de $\{1, \dots, n\}$ possédant aussi r éléments tels que $U = (b_{i,j})_{\substack{i \in R \\ j \in T}}$. Soient $j_1 < \dots < j_r$ les éléments de T , et soit $V \in \mathcal{M}_{m,r}$ la matrice dont les colonnes sont B_{j_1}, \dots, B_{j_r} . Si $X \in K^r$ vérifie $VX = 0$, on a a fortiori $UX = 0$, donc $X = 0$ et $(B_j)_{j \in T}$ est une famille libre de K^m . Donc $r \leq rg(B)$, et le rang de B est égal au plus grand entier p pour lequel

il existe une matrice à p lignes et p colonnes extraite de B dont le déterminant soit non nul. Comme le déterminant d'une matrice carrée est égal à celui de sa transposée, ceci montre que $rg(B) = rg(B^t)$, c'est à dire que le rang de B est égal à la dimension du sous-espace vectoriel de K^m engendré par les lignes de B . ♣

En fait on utilise rarement en pratique les déterminants pour calculer le rang d'une matrice. En effet le rang d'une matrice B ne change pas si on ajoute à une ligne de B une combinaison linéaire des autres, ou si on ajoute à une colonne de B une combinaison linéaire des autres. Il ne change pas non plus si on intervertit l'ordre des lignes ou l'ordre des colonnes de B , ou si on multiplie une ligne ou une colonne de B par un coefficient non nul. On va calculer à titre d'exemple le rang de la matrice suivante

$$B = \begin{pmatrix} 5 & 6 & 7 & 11 & 4 & 18 \\ 2 & 1 & 3 & 3 & 0 & 6 \\ 1 & 3 & 2 & 4 & 2 & 6 \\ -1 & 1 & 2 & 0 & -2 & 2 \end{pmatrix}.$$

$$L_1 \longrightarrow L_1 - 5L_3, L_2 \longrightarrow L_2 - 2L_3, L_4 \longrightarrow L_4 + L_3.$$

$$\begin{pmatrix} 0 & -9 & -3 & -9 & -6 & -12 \\ 0 & -5 & -1 & -5 & -4 & -6 \\ 1 & 3 & 2 & 4 & 2 & 6 \\ 0 & 4 & 4 & 4 & 0 & 8 \end{pmatrix}.$$

$$C_2 \longrightarrow C_2 - 3C_1, \text{ etc. } \dots$$

$$\begin{pmatrix} 0 & -9 & -3 & -9 & -6 & -12 \\ 0 & -5 & -1 & -5 & -4 & -6 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 4 & 4 & 4 & 0 & 8 \end{pmatrix}.$$

$$L_1 \longrightarrow -\frac{1}{3}L_1, L_2 \longrightarrow -L_2, L_4 \longrightarrow \frac{1}{4}L_4.$$

$$\begin{pmatrix} 0 & 3 & 1 & 3 & 2 & 4 \\ 0 & 5 & 1 & 5 & 4 & 6 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 2 \end{pmatrix}.$$

$$L_1 \longrightarrow L_1 - 3L_4, L_2 \longrightarrow L_2 - 5L_4.$$

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

$$C_3 \longrightarrow C_3 - C_2, \text{ etc. } \dots$$

3.7. RANG D'UNE FAMILLE DE VECTEURS, RANG D'UNE APPLICATION LINÉAIRE 45

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

$$L_2 \mapsto L_2 - 2L_1.$$

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

$$L_1 \mapsto -\frac{1}{2}L_1.$$

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

$$C_5 \mapsto C_5 + C_3, C_6 \mapsto C_6 - C_3.$$

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

Donc le rang de $\begin{pmatrix} 5 & 6 & 7 & 11 & 4 & 18 \\ 2 & 1 & 3 & 3 & 0 & 6 \\ 1 & 3 & 2 & 4 & 2 & 6 \\ -1 & 1 & 2 & 0 & -2 & 2 \end{pmatrix}$ est égal à 3 (on peut montrer

que les opérations ci-dessus permettent toujours d'aboutir à une matrice dont tous les coefficients sont égaux à 0 ou 1 qui possède au plus un coefficient égal à 1 sur chaque ligne et chaque colonne).

On a, d'après la règle de Sarrus, $\begin{vmatrix} 5 & 6 & 7 \\ 2 & 1 & 3 \\ 1 & 3 & 2 \end{vmatrix} = 10 + 42 + 18 - 7 - 45 - 24 =$

$70 - 76 = -6$. Donc le déterminant de la matrice $B \in \mathcal{M}_3(K)$ obtenue en retirant à A ses deux dernières lignes et ses trois dernières colonnes est non nul. Pour montrer que le rang de A est égal à 3 par cette méthode il faudrait encore vérifier que le déterminant de toutes les matrices obtenues en retirant à A une ligne et deux colonnes est nul, ce que nous ferons plus loin en utilisant Mupad.

3.7 Rang d'une famille de vecteurs, rang d'une application linéaire

On peut également définir le rang d'une famille de vecteurs ou d'une application linéaire.

Définition 3.7.1 1) Soit E un espace vectoriel de dimension finie, et soit \mathcal{F} une famille finie d'éléments de E . Le **rang** de \mathcal{F} est la dimension du sous-espace vectoriel de E engendré par \mathcal{F} .

2) Soient E et F deux espaces vectoriels de dimension finie et soit $u \in \mathcal{L}(E, F)$. Le **rang** de u est égal à la dimension de $u(E)$.

Proposition 3.7.2 1) Soit E un espace vectoriel de dimension finie $m \geq 1$, soit \mathcal{B} une base de E et soit \mathcal{F} une famille finie d'éléments de E . Alors le rang de \mathcal{F} est égal au rang de la matrice $P_{\mathcal{B}, \mathcal{F}}$.

2) Soient E et F deux espaces vectoriels non nuls de dimension finie, soit \mathcal{B} une base de E , soit \mathcal{B}' une base de F et soit $u \in \mathcal{L}(E, F)$. Alors le rang de u est égal à celui de la matrice $M_{u, \mathcal{B}, \mathcal{B}'}$.

Démonstration : 1) Pour $x \in E$ notons $[x]_{\mathcal{B}}$ le vecteur colonne formé des coordonnées de x dans la base \mathcal{B} . L'application $\theta : x \mapsto [x]_{\mathcal{B}}$ est une bijection linéaire de E sur K^m .

Soit F le sous-espace vectoriel de E engendré par \mathcal{F} . Alors F et $\theta(F)$ ont la même dimension. Comme $\theta(F)$ est le sous-espace vectoriel de K^m engendré par les colonnes de $P_{\mathcal{B}, \mathcal{F}}$, on voit que le rang de \mathcal{F} est égal au rang de $P_{\mathcal{B}, \mathcal{F}}$.

2) Soient e_1, \dots, e_n les éléments de la base \mathcal{B} . Posons $u(\mathcal{B}) = (u(e_1), \dots, u(e_n))$. L'espace vectoriel $u(E)$ est le sous-espace vectoriel de F engendré par $u(\mathcal{B})$. Donc $rg(u) = rg(P_{\mathcal{B}', u(\mathcal{B})}) = rg(M_{u, \mathcal{B}, \mathcal{B}'})$.

3.8 Annexe au Chapitre 3 : Introduction aux notations indicielles de la Physique

On utilise souvent en Physique la **convention de sommation de l'indice répété** : par exemple si i varie de 1 à n , alors la somme $\sum_{1 \leq i \leq n} x_{i,i}$ sera notée $x_{i,i}$.

Dans le cas d'un indice répété j coexistant avec un indice non répété i , l'indice i est appelé *indice libre* et l'indice répété est appelé *indice muet*. On peut changer le nom de l'indice muet sans changer la valeur de l'expression : si j varie entre 1 et n , on a $a_{i,j}x_j = a_{i,k}x_k$ car $\sum_{1 \leq j \leq n} a_{i,j}x_j = \sum_{1 \leq k \leq n} a_{i,k}x_k$.

Sauf mention expresse du contraire, l'indice libre prend les mêmes valeurs que l'indice muet. Par exemple si $n = 4$, écrire $a_{i,j}x_j = b_i$ est une façon concentrée d'écrire le système

$$\begin{cases} a_{1,1}x_1 + a_{1,2}x_2 + a_{1,3}x_3 + a_{1,4}x_4 = b_1 \\ a_{2,1}x_1 + a_{2,2}x_2 + a_{2,3}x_3 + a_{2,4}x_4 = b_2 \\ a_{3,1}x_1 + a_{3,2}x_2 + a_{3,3}x_3 + a_{3,4}x_4 = b_3 \\ a_{4,1}x_1 + a_{4,2}x_2 + a_{4,3}x_3 + a_{4,4}x_4 = b_4 \end{cases}$$

Pour n quelconque, en introduisant les cofacteurs de la matrice $A = (a_{i,j})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq n}}$, les formules de Cramer donnent pour $\det(A) \neq 0$

$$a_{i,j}x_j = b_i \iff b_i = (-1)^{i+j}C_{j,i}(A)b_j.$$

La convention de l'indice répété s'étend pour plusieurs indices. Par exemple si les indices varient de 1 à n ,

$$\begin{aligned} a_{i,k}b_{k,l}c_{l,j} &= \sum_{1 \leq k \leq n} \sum_{1 \leq l \leq n} a_{i,k}b_{k,l}c_{l,j} = \sum_{1 \leq k \leq n} a_{i,k} \left(\sum_{1 \leq l \leq n} b_{k,l}c_{l,j} \right) \\ &= \sum_{1 \leq l \leq n} \left(\sum_{1 \leq k \leq n} a_{i,k}b_{k,l} \right) c_{l,j}. \end{aligned}$$

Notons que ceci montre au passage que $a_{i,k}(b_{k,l}c_{l,j}) = (a_{i,k}b_{k,l})c_{l,j}$. Compte tenu de la définition du produit matriciel, on en déduit notamment que $(AB)C = A(BC)$ pour $A, B, C \in \mathcal{M}_n(K)$.

Ces conventions posent des problèmes de *substitution*. Par exemple si on considère l'expression $A = a_{i,j}x_i y_j$, avec $x_i = b_{i,j}z_j$, les indices variant de 1 à n , l'indice i est muet dans le premier cas, mais pas dans le second, tandis que l'indice j est muet dans les deux cas. Il faut donc dans la substitution remplacer j par un autre symbole, tout en gardant i . On écrira donc $x_i = b_{i,k}z_k$, ce qui donne

$$A = a_{i,j}b_{i,k}z_k y_j = \sum_{i=1}^n \left(\sum_{j=1}^n a_{i,j} \left(\sum_{k=1}^n b_{i,k}z_k \right) y_j \right).$$

On place souvent en Physique des indices placés plus haut que la variable concernée. C'est le cas des coordonnées d'un vecteur dans une base. On **note souvent en caractères gras les vecteurs**. Par exemple si $(\mathbf{e}_1, \dots, \mathbf{e}_n)$ est une base d'un espace vectoriel E , on notera souvent x^1, \dots, x^n les coordonnées d'un vecteur \mathbf{x} dans cette base. Ceci se traduit par la formule

$$\mathbf{x} = x^i \mathbf{e}_i.$$

Les exposants seront alors notés avec des parenthèses. Par exemple $(x^3)^2$ représentera le carré de la troisième coordonnée du vecteur \mathbf{x} . Si E est un espace vectoriel euclidien, et si $(\mathbf{e}_1, \dots, \mathbf{e}_n)$ est une base orthonormale de E (ces notions seront détaillées au Chapitre 7), la norme euclidienne d'un vecteur \mathbf{x} est donnée par la formule $\|\mathbf{x}\| = \sqrt{\sum_{1 \leq i \leq n} (x^i)^2}$ (on pourrait aussi écrire $\|\mathbf{x}\|^2 = x^i x^i$).

Nous terminons cette brève présentation en introduisant les symboles d'antisymétrie. Soient i_1, \dots, i_n des entiers compris entre 0 et n . Si ces entiers ne sont pas tous distincts le symbole d'antisymétrie $\epsilon_{i_1, \dots, i_n}$ est nul. Si les entiers i_1, \dots, i_n sont tous distincts le symbole d'antisymétrie $\epsilon_{i_1, \dots, i_n}$ est égal à la signature $\epsilon(\sigma)$ de la permutation

$$\sigma = \begin{array}{|c|c|c|c|} \hline 1 & 2 & \dots & n \\ \hline i_1 & i_2 & \dots & i_n \\ \hline \end{array}$$

En particulier le symbole d'antisymétrie est égal à 1, 0 ou -1 , et il change de signe si on permute i_p et i_q avec $p \neq q$.

Quand on a deux indices variant entre 1 et 2, on obtient

$$\epsilon^{1,1} = \epsilon^{2,2} = 0, \quad \epsilon^{1,2} = 1, \quad \epsilon^{2,1} = -1. \quad \text{On a alors, si } A = \begin{pmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{pmatrix}$$

$$\det(A) = \epsilon^{i,j} a_{i,1} a_{j,2} = \epsilon^{i,j} a_{1,i} a_{2,j}$$

que l'on peut aussi écrire sous la forme

$$\det(a_{ij}) = \epsilon^{ij} a_{i1} a_{j2} = \epsilon^{ij} a_{1i} a_{2j}$$

3.9 Chapitre 3 sous MUPAD

On a vu au chapitre précédent comment calculer l'inverse d'une matrice avec MUPAD. On peut évidemment aussi calculer les déterminants avec MUPAD, en utilisant la commande `linalg : :det()` ;

```
M:=Dom::Matrix();
A:=M([[1,2,3,4,5],[1,2,1,2,1],[1,-1,4,3,-1],[2,-1,3,1,4],[5,-2,3,-1,3]]);
linalg::det(A);
```

```
Dom::Matrix(Dom::ExpressionField(id, iszero))
```

$$\begin{array}{rccccc} +- & & & & -+ \\ | & 1, & 2, & 3, & 4, & 5 & | \\ | & & & & & & | \\ | & 1, & 2, & 1, & 2, & 1 & | \\ | & & & & & & | \\ | & 1, & -1, & 4, & 3, & -1 & | \\ | & & & & & & | \\ | & 2, & -1, & 3, & 1, & 4 & | \end{array}$$

$$\begin{array}{c} | \\ | \quad 5, -2, 3, -1, \quad 3 \quad | \\ +- \qquad \qquad \qquad -+ \end{array}$$

56

Donc $\begin{vmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 2 & 1 & 2 & 1 \\ 1 & -1 & 4 & 3 & -1 \\ 2 & -1 & 3 & 1 & 4 \\ 5 & -2 & 3 & -1 & 3 \end{vmatrix} = 56.$

```
M:=Dom::Matrix();
A:=M([[5,6,7,11,4,18],[2,1,3,3,0,6],[1,3,2,4,2,6],[-1,1,2,0,-2,2]]);
linalg::rank(A);
```

```
Dom::Matrix(Dom::ExpressionField(id, iszero))
```

$$\begin{array}{c} +- \qquad \qquad \qquad -+ \\ | \quad 5, 6, 7, 11, \quad 4, 18 \quad | \\ | \\ | \quad 2, 1, 3, \quad 3, \quad 0, \quad 6 \quad | \\ | \\ | \quad 1, 3, 2, \quad 4, \quad 2, \quad 6 \quad | \\ | \\ | \quad -1, 1, 2, \quad 0, -2, \quad 2 \quad | \\ +- \qquad \qquad \qquad -+ \end{array}$$

3

On peut également utiliser Mupad pour ajouter a une ligne d'une matrice un multiple d'une autre. Par exemple $L_1 \rightarrow L_1 - 5L_3$ (retrancher 5 fois la troisième ligne à la première) se traduit pour la matrice A par la commande `linalg : :addRow(A,3,1,-5)`, et $C_3 \rightarrow C_2 - 3C_1$ (Retrancher deux fois la première colonne à la deuxième) se traduit pour la matrice A par la commande `linalg : :addCol(A,1,2,-3)`. On peut ainsi effectuer sous Mupad le début des calculs

faits plus haut pour calculer le rang de la matrice $A = \begin{pmatrix} 5 & 6 & 7 & 11 & 4 & 18 \\ 2 & 1 & 3 & 3 & 0 & 6 \\ 1 & 3 & 2 & 4 & 2 & 6 \\ -1 & 1 & 2 & 0 & -2 & 2 \end{pmatrix}.$

50 CHAPITRE 3. DÉTERMINANTS, NOTATIONS INDICIELLES DE LA PHYSIQUE

```

M:=Dom::Matrix ();
A:=M([[5,6,7,11,4,18],[2,1,3,3,0,6],[1,3,2,4,2,6],[-1,1,2,0,-2,2]]);
A1:=linalg::addRow(A,3,1,-5);
A2:=linalg::addRow(A1,3,2,-2);
A3:=linalg::addRow(A2,3,4,1);
A4:=linalg::addCol(A3,1,2,-3);
A5:=linalg::addCol(A4,1,3,-2);
A6:=linalg::addCol(A5,1,4,-4);
A7:=linalg::addCol(A6,1,5,-2);
A8:=linalg::addCol(A7,1,6,-6);

```

```

+- -+
| 5, 6, 7, 11, 4, 18 |
|                       |
| 2, 1, 3, 3, 0, 6 |
|                       |
| 1, 3, 2, 4, 2, 6 |
|                       |
| -1, 1, 2, 0, -2, 2 |
+- -+

```

```

+- -+
| 0, -9, -3, -9, -6, -12 |
|                       |
| 0, -5, -1, -5, -4, -6 |
|                       |
| 1, 3, 2, 4, 2, 6 |
|                       |
| 0, 4, 4, 4, 0, 8 |
+- -+

```

```

+- -+
| 0, -9, -3, -9, -6, -12 |
|                       |
| 0, -5, -1, -5, -4, -6 |
|                       |
| 1, 0, 0, 0, 0, 0 |
|                       |
| 0, 4, 4, 4, 0, 8 |
+- -+

```

D'autre part on peut utiliser Mupad pour extraire des matrices d'une matrice donnée. On utilise la commande `linalg::submatrix(A, [i1, ..., ip], [j1, ..., jq])` ; qui permet de faire apparaître la matrice obtenue en supprimant les lignes autres que i_1, \dots, i_p et les colonnes autres que j_1, \dots, j_q . A titre d'exemple on calcule ci-dessous les déterminants de la matrice carrée à 4 lignes et 4 colonnes extraite

de la matrice $A = \begin{pmatrix} 5 & 6 & 7 & 11 & 4 & 18 \\ 2 & 1 & 3 & 3 & 0 & 6 \\ 1 & 3 & 2 & 4 & 2 & 6 \\ -1 & 1 & 2 & 0 & -2 & 2 \end{pmatrix}$ en gardant les coefficients ap-

partenant à la fois à une des quatre premières lignes et à une des quatre dernières colonnes de A .

```
M:=Dom::Matrix():
A:= M([[5,6,7,11,4,18],[2,1,3,3,0,6],[1,3,2,4,2,6],[-1,1,2,0,-2,2]]);
A1:=linalg::submatrix(A,[1,2,3,4],[3,4,5,6]);
```

```
+ - - +
| 5, 6, 7, 11, 4, 18 |
|                       |
| 2, 1, 3, 3, 0, 6 |
|                       |
| 1, 3, 2, 4, 2, 6 |
|                       |
| -1, 1, 2, 0, -2, 2 |
+ - - +
```

```
+ - - +
| 7, 11, 4, 18 |
|               |
| 3, 3, 0, 6 |
|               |
| 2, 4, 2, 6 |
|               |
| 2, 0, -2, 2 |
+ - - +
```

On peut enfin utiliser Mupad pour résoudre des systèmes linéaires. On doit ramener le système à un système de la forme $AX = B$, avec $A \in \mathcal{M}_{m,n}(K)$, $B \in K^n$.

On traite ci-dessous l'exemple du système

$$\begin{cases} 5x + 6y + 7z = 4 \\ 2x + y + 3z = 5 \\ x + 3y + 2z = 1 \end{cases}$$

```
M:=Dom::Matrix():
A:=M([[5,6,7],[2,1,3],[1,3,2]]);
B:=M([4,5,1]);
linalg::matlinsolve(A,B);
```

$$\begin{array}{ccc} +- & & -+ \\ | & 5, 6, 7 & | \\ | & & | \\ | & 2, 1, 3 & | \\ | & & | \\ | & 1, 3, 2 & | \\ +- & & -+ \\ & & | & 4 & | \\ & & | & & | \\ & & | & 5 & | \\ & & | & & | \\ & & | & 1 & | \\ +- & & -+ \end{array}$$

$$\begin{array}{ccc} +- & & -+ \\ | & -14/3 & | \\ | & & | \\ | & -5/3 & | \\ | & & | \\ | & 16/3 & | \\ +- & & -+ \end{array}$$

On a donc une solution unique $x = -14/3$, $y = -5/3$, $z = 16/3$.

On considère maintenant le système

$$\begin{cases} 5x + 6y + 7z + 11t = 1 \\ 2x + y + 3z + 3t = 1 \\ x + 3y + 2z + 4t = 1 \\ -x + y + 2z = 1 \end{cases}$$

```
A1:=M([[5,6,7,11],[2,1,3,3],[1,3,2,4],[-1,1,2,0]]);
B1:=M([1,1,1,1]);
linalg::matlinsolve(A1,B1);
```

$$\begin{array}{cccc|c} +- & & & & -+ \\ | & 5, & 6, & 7, & 11 & | \\ | & & & & & | \\ | & 2, & 1, & 3, & 3 & | \\ | & & & & & | \\ | & 1, & 3, & 2, & 4 & | \\ | & & & & & | \\ | & -1, & 1, & 2, & 0 & | \\ +- & & & & & -+ \end{array}$$

$$\begin{array}{ccc|c} +- & & -+ \\ | & 1 & | \\ | & & | \\ | & 1 & | \\ | & & | \\ | & 1 & | \\ | & & | \\ | & 1 & | \\ +- & & -+ \end{array}$$

□

Ceci veut dire que le système **n'a pas de solution**.

Considérons maintenant le système

$$\begin{cases} 5x + 6y + 7z + 11t = 29 \\ 2x + y + 3z + 3t = 9 \\ x + 3y + 2z + 4t = 10 \\ -x + y + 2z = 2 \end{cases}$$

```
B2:=M([29,9,10,2]);
linalg::matlinsolve(A1,B2);
```

$$\begin{array}{cc} + - & - + \\ | & 29 & | \\ | & & | \\ | & 9 & | \\ | & & | \\ | & 10 & | \\ | & & | \\ | & 2 & | \\ + - & - + \end{array}$$

$$\begin{array}{cccccc} - - & + - & - + & - - & + - & - + & - - & - - \\ | & | & 2 & | & | & | & -1 & | & | & | \\ | & | & & | & | & | & & | & | & | \\ | & | & 2 & | & | & | & -1 & | & | & | \\ | & | & & | & | & | & & | & | & | \\ | & | & 1 & | & | & | & 0 & | & | & | \\ | & | & & | & | & | & & | & | & | \\ | & | & 0 & | & | & | & 1 & | & | & | \\ - - & + - & - + & - - & + - & - + & - - & - - \end{array}$$

Ceci veut dire que le système **a une infinité de solutions**. Mupad donne d'abord une solution particulière, puis une base du sous-espace vectoriel de K^n défini par l'équation $AX = 0$. La solution générale du système est donc

$$\begin{cases} x = 2 - \lambda \\ y = 2 - \lambda \\ z = 1 \\ t = \lambda \end{cases}, \text{ où } \lambda \in \mathbf{R}.$$

3.10 Exercices sur le Chapitre 3

exercice 1

On considère la matrice

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

Calculer $\det(A)$. La matrice A est-elle inversible ?

En déduire que le système ci-dessous admet une solution unique, et trouver cette solution en utilisant les formules de Cramer.

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

exercice 2 (Matrices nilpotentes)

Soit $A \in \mathcal{M}_n(\mathbf{R})$ telle que $A^n = 0$, avec $n \geq 2$. Montrer que $\det(A) = 0$.

exercice 3 (Matrices antisymétriques)

Soit n un entier positif impair et soit $A \in \mathcal{M}(\mathbf{R})$ telle que ${}^t A = -A$. Montrer que $\det(A) = 0$.

exercice 4

Calculer les déterminants suivants :

$$\begin{aligned} & \begin{vmatrix} 4 & -1 & 10 \\ 2 & 5 & -3 \\ 1 & 6 & 2 \end{vmatrix}, \begin{vmatrix} 1 & 2 & 3 & -5 \\ 0 & 1 & 0 & 4 \\ 7 & 3 & -2 & 1 \\ -3 & -5 & 1 & 2 \end{vmatrix}, \begin{vmatrix} a & c+id \\ c-id & b \end{vmatrix}, \begin{vmatrix} a-1 & 1 \\ a^3 & a^2+a+1 \end{vmatrix}, \\ & \begin{vmatrix} a & b & c \\ b & c & a \\ c & b & a \end{vmatrix}, \begin{vmatrix} 1 & a & b+c \\ 1 & b & c+a \\ 1 & c & a+b \end{vmatrix}, \begin{vmatrix} 1 & 1 & 1 & 1 \\ 1 & i & -1 & -i \\ 1 & -1 & 1 & -1 \\ 1 & -i & -1 & i \end{vmatrix}, \begin{vmatrix} a & 1 & 1 & -1 \\ 1 & a & 1 & 1 \\ 1 & 1 & a & 1 \\ 1 & 1 & 1 & a \end{vmatrix}. \end{aligned}$$

exercice 5

Déterminer les racines du polynôme $p(x) = \begin{vmatrix} x & 2 & 3 & 4 \\ 2 & x & 3 & 4 \\ 3 & 4 & x & 2 \\ 4 & 3 & 2 & x \end{vmatrix}$.

exercice 6

a) Déterminer en fonction de a et b les coefficients de la matrice

$$A = \begin{pmatrix} 1 & a & b \\ -1 & a & b \\ a+b+1 & 1 & 1 \end{pmatrix}.$$

b) Résoudre dans \mathbf{R}^3 le système $AX = B$, avec $X = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$ et $B = \begin{pmatrix} u \\ v \\ w \end{pmatrix}$.

exercice 7

Soit ω une racine cubique de l'unité. Résoudre dans \mathbf{C} le système

$$\begin{cases} x + y + z = a \\ x + \omega y + \omega^2 z = b \\ x + \omega^2 y + \omega z = c \end{cases}$$

exercice 8

a) Résoudre le système

$$\begin{cases} x + y + z + t = 3 \\ x + 2y + z + t = 1 \\ x + y + 2z + t = 2 \\ x + y + z + 2t = 4 \\ x - y + z - t = 0 \end{cases}$$

b) Discuter selon les valeurs de a l'ensemble des solutions du système

$$\begin{cases} 2(a+1)x + 3y + az = a+4 \\ (4a-1)x + (a+1)y + (2a-1)z = 2a+4 \\ (5a-4)x + (a+1)y + (3a-4)z = a-1 \end{cases}$$

exercice 9

Utiliser la convention de sommation pour écrire les relations suivantes

a) $a_{1,1}x_1 + a_{2,2}x_2 + a_{3,3}x_3 + a_{4,4}x_4 = 8.$

b) $B_{1,1}y_{1,1} + B_{2,1}y_{1,2} + B_{3,1}y_{1,3} = A.$

exercice 10

Exprimer le produit de deux matrices $A = (a_{ij}) \in \mathcal{M}_n(K)$ et $B = (b_{ij}) \in \mathcal{M}_n(K)$ en utilisant la convention de l'indice répété.

exercice 11

Ecrire la définition de l'indépendance linéaire d'une famille $(\mathbf{e}_1, \dots, \mathbf{e}_n)$ de vecteurs en utilisant la convention de l'indice répété.

exercice 12

Développer l'expression $a_{i,j}b_{k,j,i}$, avec $n = 3$.

exercice 13

Calculer $a_{j,i}x_i y_j$, avec $x_i = b_{i,j}u_j$, $y_j = c_{i,j}v_i$, $n = 2$.

exercice 14

a) Calculer tous les symboles d'antisymétrie ϵ^{ijk} pour $i, j, k \in \{1, 2, 3\}$.

b) Soit $A = (a_{ij}) \in \mathcal{M}_3(K)$. Exprimer $\det(A)$ en utilisant les symboles d'antisymétrie et la convention de l'indice répété.

exercice 15(sous Mupad)

1) Calculer le déterminant de la matrice $\begin{pmatrix} 2 & 4 & -1 & 7 & 1 & 3 & 4 \\ 5 & -1 & 6 & 9 & -2 & 6 & 2 \\ 7 & -5 & -2 & 6 & 1 & 5 & 1 \\ 5 & 1 & 1 & -6 & 2 & -9 & 4 \\ 1 & 4 & 6 & 7 & 2 & 3 & 2 \\ 3 & 2 & 1 & 1 & -1 & 4 & 2 \\ 5 & 6 & 2 & 3 & -6 & -1 & 3 \end{pmatrix}.$

2) Résoudre le système

$$\begin{cases} 2x_1 + 4x_2 - x_3 + 7x_4 + x_5 + 3x_6 + 4x_7 = 2 \\ 5x_1 - x_2 + 6x_3 + 9x_4 - 2x_5 + 6x_6 + 2x_7 = -3 \\ 7x_1 - 5x_2 - 2x_3 + 6x_4 + x_5 + 5x_6 + x_7 = 12 \\ 5x_1 + x_2 + x_3 - 6x_4 + 2x_5 - 9x_6 + 4x_7 = -1 \\ x_1 + 4x_2 + 6x_3 + 7x_4 + 2x_5 + 3x_6 + 2x_7 = -4 \\ 3x_1 + 2x_2 + x_3 + x_4 - x_5 + 4x_6 + 2x_7 = 2 \\ 5x_1 + 6x_2 + 2x_3 + 3x_4 - 6x_5 - x_6 + 3x_7 = 6 \end{cases}$$

exercice 16(sous Mupad)

1) Déterminer le rang de la matrice $A = \begin{pmatrix} 2 & 4 & -1 & 7 & 1 & 3 & 16 \\ 5 & -1 & 6 & 9 & -2 & 6 & 23 \\ 7 & -5 & -2 & 6 & 1 & 6 & 13 \\ 5 & 1 & 1 & -6 & 2 & -9 & -6 \\ 1 & 4 & 6 & 7 & 2 & 3 & 23 \\ 20 & 3 & 10 & 23 & 4 & 9 & 69 \end{pmatrix}$.

2) Retrouver ce résultat en calculant le déterminant des 7 matrices à 6 lignes et 6 colonnes extraites de A et en calculant le déterminant d'une des matrices à 5 lignes et 5 colonnes extraites de A .

3) Résoudre le système

$$\begin{cases} 2x_1 + 4x_2 - x_3 + 7x_4 + x_5 + 3x_6 + 16x_7 = 2 \\ 5x_1 - x_2 + 6x_3 + 9x_4 - 2x_5 + 6x_6 + 23x_7 = -3 \\ 7x_1 - 5x_2 - 2x_3 + 6x_4 + x_5 + 6x_6 + 13x_7 = 12 \\ 5x_1 + x_2 + x_3 - 6x_4 + 2x_5 - 9x_6 - 6x_7 = -1 \\ x_1 + 4x_2 + 6x_3 + 7x_4 + 2x_5 + 3x_6 + 23x_7 = -4 \\ 20x_1 + 3x_2 + 10x_3 + 23x_4 + 4x_5 + 9x_6 + 69x_7 = 2 \end{cases}$$

4) Résoudre le système

$$\begin{cases} 2x_1 + 4x_2 - x_3 + 7x_4 + x_5 + 3x_6 + 16x_7 = 1 \\ 5x_1 - x_2 + 6x_3 + 9x_4 - 2x_5 + 6x_6 + 23x_7 = 1 \\ 7x_1 - 5x_2 - 2x_3 + 6x_4 + x_5 + 6x_6 + 13x_7 = 1 \\ 5x_1 + x_2 + x_3 - 6x_4 + 2x_5 - 9x_6 - 6x_7 = 1 \\ x_1 + 4x_2 + 6x_3 + 7x_4 + 2x_5 + 3x_6 + 23x_7 = 1 \\ 20x_1 + 3x_2 + 10x_3 + 23x_4 + 4x_5 + 9x_6 + 69x_7 = 1 \end{cases}$$

Chapitre 4

Valeurs propres, vecteurs propres, diagonalisation

4.1 Introduction à la diagonalisation des matrices

Dans tout ce chapitre K désigne un corps quelconque, et $\mathcal{M}_n(K)$ désigne l'ensemble des matrices carrées à n lignes et n colonnes à coefficients dans K . L'espace vectoriel K^n est identifié à l'espace vectoriel $\mathcal{M}_{n,1}$ des matrices unicolonnes à n lignes.

Définition 4.1.1 Soit $A \in \mathcal{M}_n(K)$, et soit $\lambda \in K$. On dit que λ est une **valeur propre** de A s'il existe un élément non nul X de K^n tel que $AX = \lambda X$. Dans ce cas on dit que X est un **vecteur propre** associé à la valeur propre λ .

Si λ est une valeur propre de A , l'ensemble $E_\lambda(A) = \{X \in K^n \mid AX = \lambda X\}$ est appelé le **sous-espace propre** associé à la valeur propre λ .

Afin d'alléger les notations, on écrira E_λ au lieu de $E_\lambda(A)$ si aucune confusion n'est à craindre. Soit $u : K^n \rightarrow K^n$ l'application linéaire définie par la formule $u(X) = AX$ et soit \mathcal{I}_n l'application identité sur K^n . Il est clair que si λ est une valeur propre de A alors $E_\lambda = \text{Ker}(u - \lambda\mathcal{I}_n)$, et par conséquent le sous-espace propre E_λ est un sous-espace vectoriel non nul de K^n .

Définition 4.1.2 Soit $p = a_0 + a_1x + \dots + a_mx^m \in K[x]$, et soit $A \in \mathcal{M}_n(K)$. On pose $p(A) = a_0I_n + a_1A + \dots + a_mA^m$.

Autrement dit $p(A)$ est la matrice obtenue en remplaçant x^k par A^k dans l'expression de p , avec la convention $A^0 = I_n$. On a les propriétés évidentes suivantes.

$$(4.1) \quad (p+q)(A) = p(A) + q(A) \quad \forall p \in K[x], \forall q \in K[x], \forall A \in \mathcal{M}_n(K).$$

$$(4.2) \quad (\lambda p)(A) = \lambda p(A) \quad \forall \lambda \in K, \forall p \in K[x], \forall A \in \mathcal{M}_n(K).$$

$$(4.3) \quad (pq)(A) = p(A)q(A) \quad \forall p \in K[x], \forall q \in K[x], \forall A \in \mathcal{M}_n(K).$$

Autrement dit si on fixe $A \in \mathcal{M}_n(K)$, alors l'application $p \mapsto p(A)$ est à la fois une *application linéaire* et un *homomorphisme d'anneaux* de $K[x]$ dans $\mathcal{M}_n(K)$. Notons que l'on déduit de (4.3) la propriété suivante.

$$(4.4) \quad q(A)p(A) = p(A)q(A) \quad \forall p \in K[x], \forall q \in K[x], \forall A \in \mathcal{M}_n(K).$$

Si E_1, \dots, E_p sont des sous-espaces vectoriels d'un espace vectoriel E on définit *l'espace somme* $E_1 + \dots + E_p$ comme étant l'ensemble des $\mathbf{x} \in E$ qui peuvent s'écrire sous la forme $\mathbf{x} = \sum_{i=1}^p \mathbf{x}_i$ avec $\mathbf{x}_i \in E_i$ pour $1 \leq i \leq p$. Il est clair que $E_1 + \dots + E_p$ est un sous-espace vectoriel de E .

Lemme 4.1.3 *Soit $A \in \mathcal{M}_n(K)$, soit λ une valeur propre de A et soit $p \in K[x]$. Alors $p(A)X = p(\lambda)X$ pour tout $X \in E_\lambda$.*

Démonstration : Soit $X \in E_\lambda$. On a $IX = X$, et une récurrence immédiate montre que $A^k X = \lambda^k X$ pour $k \geq 1$. Soit maintenant $p = a_0 + a_1 x + \dots + a_m x^m \in K[x]$. On a $p(A)X = a_0 X + a_1 \lambda X + \dots + a_m \lambda^m X = p(\lambda)X$. ♣

Proposition 4.1.4 *Soit $A \in \mathcal{M}_n(K)$, soient $\lambda_1, \dots, \lambda_p$ des valeurs propres distinctes de A et soit \mathcal{B}_i une base de E_{λ_i} pour $1 \leq i \leq p$. Alors $\mathcal{B} = \cup_{1 \leq i \leq p} \mathcal{B}_i$ est une base de $E_{\lambda_1} + \dots + E_{\lambda_p}$. En particulier $\dim(E_{\lambda_1} + \dots + E_{\lambda_p}) = \dim(E_{\lambda_1}) + \dots + \dim(E_{\lambda_p})$.*

Démonstration : Il est clair que \mathcal{B} est une famille génératrice de $E_{\lambda_1} + \dots + E_{\lambda_p}$. Posons $\mathcal{B}_i = (\mathbf{f}_{i,1}, \dots, \mathbf{f}_{i,q(i)})$, et soit $(\alpha_{i,j})_{\substack{1 \leq j \leq q(i) \\ 1 \leq i \leq p}}$ une famille d'éléments de K telle que $\sum_{\substack{1 \leq j \leq q(i) \\ 1 \leq i \leq p}} \alpha_{i,j} \mathbf{f}_{i,j} = 0$.

Posons $\mathbf{g}_i = \sum_{1 \leq j \leq q(i)} \alpha_{i,j} \mathbf{f}_{i,j}$, de sorte que $\mathbf{f}_i \in E_{\lambda_i}$ pour $1 \leq i \leq p$.

On a $\sum_{1 \leq k \leq p} \mathbf{g}_k = 0$. Posons $p_i = \prod_{\substack{1 \leq k \leq p \\ k \neq i}} (x - \lambda_k)$ pour $1 \leq i \leq p$.

On a $p_i(A)(\mathbf{g}_k) = p_i(\lambda_k) \mathbf{g}_k$ pour $1 \leq k \leq n$, donc $\sum_{1 \leq k \leq n} p_i(\lambda_k) \mathbf{g}_k = p_i(A)(0) = 0$. Comme $p_i(\lambda_k) = 0$ pour $k \neq i$, et comme $p_i(\lambda_i) \neq 0$, on a $\sum_{1 \leq j \leq q(i)} \alpha_{i,j} \mathbf{f}_{i,j} = 0$ pour $1 \leq i \leq p$. Comme \mathcal{B}_i est libre, on obtient $\alpha_{i,j} = 0$ pour $1 \leq j \leq q(i)$, $1 \leq i \leq p$, ce qui prouve que \mathcal{B} est libre. Donc \mathcal{B} est une base de $E_{\lambda_1} + \dots + E_{\lambda_p}$, et $\dim(E_{\lambda_1} + \dots + E_{\lambda_p}) = \dim(E_{\lambda_1}) + \dots + \dim(E_{\lambda_p})$. ♣

4.2 Polynôme caractéristique d'une matrice

On va maintenant donner un moyen pratique de trouver les valeurs propres de A .

Définition 4.2.1 Soit $A \in \mathcal{M}_n(K)$. Le polynôme $p_A = \det(xI_n - A) = (-1)^n \det(A - xI_n)$ est appelé le **polynôme caractéristique** de A .

Soit $p \in K[x]$ un polynôme non nul. Rappelons qu'on dit que $\lambda \in K$ est une **racine** de p quand $p(\lambda) = 0$. On sait que λ est une racine de p si et seulement si p est divisible par $x - \lambda$. Dans ce cas le plus grand entier k tel que p soit divisible par $(x - \lambda)^k$ est appelé l'**ordre de multiplicité** de la racine λ .

Proposition 4.2.2 Soit $A \in \mathcal{M}_n(K)$ et soit $\lambda \in K$. Alors λ est valeur propre de A si et seulement si $\det(\lambda I_n - A) = 0$. Autrement dit l'ensemble des valeurs propres de A coïncide avec l'ensemble des racines de son polynôme caractéristique p_A .

Démonstration : Il est clair que λ est valeur propre de A si et seulement si il existe un élément non nul $X \in K^n$ tel que $(\lambda I_n - A)X = 0$. Il résulte alors du corollaire 3.8 que cette condition est vérifiée si et seulement si $\det(\lambda I_n - A) = 0$.

♣

Exemple 4.2.3 Soit $A = \begin{pmatrix} 1/2 & 1/2 & 1/2 \\ 1/4 & 1/2 & 0 \\ 1/4 & 0 & 1/2 \end{pmatrix}$. Alors $p_A = x(x - \frac{1}{2})(x - 1)$, et

les valeurs propres de A sont $0, \frac{1}{2}$ et 1 .

$$\text{En effet } p_A = - \begin{vmatrix} 1/2 - x & 1/2 & 1/2 \\ 1/4 & 1/2 - x & 0 \\ 1/4 & 0 & 1/2 - x \end{vmatrix}.$$

En développant par rapport à la dernière colonne on obtient

$$\begin{aligned} -p_A &= -\frac{1}{2} \begin{vmatrix} 1/4 & 1/2 - x \\ 1/4 & 0 \end{vmatrix} - \left(\frac{1}{2} - x\right) \begin{vmatrix} 1/2 - x & 1/2 \\ 1/4 & 1/2 - x \end{vmatrix} = -\frac{1}{2} \left[-\frac{1}{4} \left(\frac{1}{2} - x\right)\right] \\ &- \left(\frac{1}{2} - x\right) \left[\left(\frac{1}{2} - x\right)^2 - \frac{1}{8}\right] = \frac{1}{4} \left(\frac{1}{2} - x\right) - \left(\frac{1}{2} - x\right)^3 = -\left(\frac{1}{2} - x\right) \left[\left(\frac{1}{2} - x\right)^2 - \frac{1}{4}\right] \\ &= x \left(\frac{1}{2} - x\right) (1 - x) = x \left(x - \frac{1}{2}\right) (x - 1). \end{aligned}$$

Si $A \in \mathcal{M}_n(K)$, on définit la **trace** de A , notée $Tr(A)$, comme étant la somme des coefficients situés sur la diagonale. Autrement dit on a la formule

$$(4.5) \quad Tr \left((a_{i,j})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq n}} \right) = \sum_{i=1}^n a_{i,i}.$$

Il résulte de la définition des déterminants que le terme de plus haut degré du polynôme caractéristique de $A \in \mathcal{M}_n(K)$ est égal à 1, que le coefficient de

son monôme de degré $n - 1$ est égal à $-Tr(A)$, et que son terme constant est égal à $(-1)^n det(A)$. On a alors le résultat suivant.

Proposition 4.2.4 *Soit $A \in \mathcal{M}_n(K)$. Alors la somme des valeurs propres de A , répétées selon leurs multiplicités, est égale à $Tr(A)$, et le produit des valeurs propres de A , répétées selon leurs multiplicités, est égal à $det(A)$.*

4.3 Matrices semblables

On va maintenant introduire la notion de similarité.

Définition 4.3.1 *Soient $A, B \in \mathcal{M}_n(K)$. On dit que B est **semblable** à A s'il existe une matrice inversible P telle que $B = P^{-1}AP$.*

Cette relation est ce qu'on appelle une *relation d'équivalence*, car on a les trois propriétés suivantes :

- 1) A est toujours semblable à A , car $A = I_n^{-1}AI_n$.
- 2) Si $B = P^{-1}AP$ est semblable à A , alors $A = PBP^{-1}$ est semblable à B , car $P = (P^{-1})^{-1}$.
- 3) Si $B = P^{-1}AP$ est semblable à A , et si $C = Q^{-1}BQ$ est semblable à B , alors $C = Q^{-1}P^{-1}APQ$ est semblable à A , car $Q^{-1}P^{-1} = (PQ)^{-1}$.

On a les propriétés suivantes.

Proposition 4.3.2 *Soit $A \in \mathcal{M}_n(K)$, et soit $B = P^{-1}AP \in \mathcal{M}_n(K)$ une matrice semblable à A . Alors A et B ont le même polynôme caractéristique. De plus $q(A) = Pq(B)P^{-1}$ pour tout $q \in K[x]$, et en particulier $A^m = PB^mP^{-1}$ pour $m \geq 1$.*

Démonstration : On a $p_B = det(xI_n - B) = det(P^{-1}(xI_n)P - (P^{-1}AP)) = det(P^{-1}(xI_n - A)P) = det(P^{-1})det(xI_n - A)det(P) = det(P)^{-1}p_A det(P) = det(I_n)p_A = p_A$.

D'autre part on a $A = PBP^{-1}$. Supposons que $A^m = PB^mP^{-1}$, avec $m \geq 1$. Alors $A^{m+1} = (PB^mP^{-1})(PBP^{-1}) = PB^m(P^{-1}P)BP^{-1} = PB^{m+1}P^{-1}$. On voit donc par récurrence que $A^m = PB^mP^{-1}$ pour $m \geq 1$. Comme $I_n = PP^{-1}$, on voit que si $q = b_0 + \dots + b_k x^k \in K[x]$, alors $q(A) = b_0 PI_n P^{-1} + \dots + b_m P B^m P^{-1} = P(b_0 I_n)P^{-1} + \dots + P(b_m B^m)P^{-1} = Pq(B)P^{-1}$. ♣

Corollaire 4.3.3 *Soient $A, B \in \mathcal{M}_n(K)$. Si A et B sont semblables, alors $Tr(A) = Tr(B)$.*

$$(4.6) \text{ Si } D = \begin{pmatrix} d_{1,1} & 0 & \dots & \dots & \dots & 0 \\ 0 & d_{2,2} & 0 & \dots & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & d_{j,j} & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \dots & 0 & d_{n,n} \end{pmatrix} \text{ est diagonale, alors } p_D = (x - d_{1,1}) \dots (x - d_{n,n}).$$

$$(4.7) \text{ Si } D = \begin{pmatrix} d_{1,1} & 0 & \dots & \dots & \dots & 0 \\ 0 & d_{2,2} & 0 & \dots & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & d_{j,j} & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \dots & 0 & d_{n,n} \end{pmatrix} \text{ est diagonale, et si } p \in K[x], \text{ alors } p(D) \text{ est diagonale, et}$$

$$p(D) = \begin{pmatrix} p(d_{1,1}) & 0 & \dots & \dots & \dots & 0 \\ 0 & p(d_{2,2}) & 0 & \dots & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & p(d_{j,j}) & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \dots & 0 & p(d_{n,n}) \end{pmatrix}.$$

On voit en particulier que si $D = \begin{pmatrix} d_{1,1} & 0 & \dots & \dots & \dots & 0 \\ 0 & d_{2,2} & 0 & \dots & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & d_{j,j} & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \dots & 0 & d_{n,n} \end{pmatrix}$ est diagonale, alors $D^p = \begin{pmatrix} d_{1,1}^p & 0 & \dots & \dots & \dots & 0 \\ 0 & d_{2,2}^p & 0 & \dots & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & d_{j,j}^p & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \dots & 0 & d_{n,n}^p \end{pmatrix}$ pour $p \geq 1$.

Définition 4.5.1 Soit $A \in \mathcal{M}_n(K)$. On dit que A est **diagonalisable** quand A est semblable à une matrice diagonale.

Proposition 4.5.2 Soit $A \in \mathcal{M}_n(K)$, et soit $P \in \mathcal{M}_n(K)$ une matrice inversible. Alors $P^{-1}AP$ est diagonale si et seulement si les colonnes $\mathbf{p}_1, \dots, \mathbf{p}_n$ de P sont des vecteurs propres de A .

Démonstration : Soient $\mathbf{p}_1, \dots, \mathbf{p}_n$ les colonnes de P . Comme P est de rang n , $\mathcal{B} = (\mathbf{p}_1, \dots, \mathbf{p}_n)$ est une base de K^n . Soit \mathcal{B}_0 la base canonique de K^n , et soit u l'endomorphisme de K^n défini par la formule $u(X) = AX$. On a $P_{\mathcal{B}_0, \mathcal{B}} = P$,

donc d'après la formule de changement de bases on a $\mathcal{M}_{u,\mathcal{B}} = P_{\mathcal{B}_0,\mathcal{B}}^{-1}AP_{u,\mathcal{B}_0,\mathcal{B}} = P^{-1}AP$. Comme la j^e colonne de $\mathcal{M}_{u,\mathcal{B}}$ est formée des coordonnées de \mathbf{p}_j dans la base $\mathcal{B} = (\mathbf{p}_1, \dots, \mathbf{p}_n)$, on voit que $P^{-1}AP$ est diagonale si et seulement si \mathbf{p}_j est un vecteur propre de A pour $1 \leq j \leq n$.

Corollaire 4.5.3 *Soit $A \in \mathcal{M}_n(K)$. Alors A est diagonalisable si et seulement si il existe une base de K^n formée de vecteurs propres de A .*

Démonstration : Si A est diagonalisable, il existe une matrice inversible $P \in \mathcal{M}_n(K)$ telle que $P^{-1}AP$ est diagonale, et les colonnes de P forment une base de K^n formée de vecteurs propres de A . Réciproquement soit $\mathcal{B} = (\mathbf{f}_1, \dots, \mathbf{f}_n)$ une base de K^n formée de vecteurs propres de A , et soit $P = P_{\mathcal{B}_0,\mathcal{B}}$, \mathcal{B}_0 désignant la base canonique de K^n . Alors P est inversible, et les colonnes de P sont égales à $\mathbf{f}_1, \dots, \mathbf{f}_n$. Il résulte alors de la proposition que $P^{-1}AP$ est diagonale, donc A est diagonalisable.

On dira qu'un polynôme $p \in K[x]$ de degré supérieur ou égal à 1 est **scindé** s'il se factorise dans $K[x]$ en produit de polynômes de degré 1. On a alors le théorème suivant.

Théorème 4.5.4 *Soit $A \in \mathcal{M}_n(K)$. Alors A est diagonalisable si et seulement si les deux conditions suivantes sont vérifiées*

(i) *Le polynôme caractéristique p_A de A est scindé.*

(ii) *L'ordre de multiplicité de chaque valeur propre de A est égal à la dimension du sous-espace propre correspondant.*

Démonstration : Supposons que A est diagonalisable, et soit D une matrice diagonale semblable à A . Il résulte de la proposition 4.3.2 appliquée à D que $p_A = p_D$ est scindé. Soient $\lambda_1, \dots, \lambda_p$ les valeurs propres distinctes de A et pour $1 \leq j \leq p$ soit n_j l'ordre de multiplicité de λ_j . Comme p_A est scindé, on a $p_A = (x - \lambda_1)^{n_1} \dots (x - \lambda_p)^{n_p}$, et $\sum_{1 \leq j \leq p} n_j = n$. Puisqu'il existe une base de K^n formée de vecteurs propres de A , on a $K^n = E_{\lambda_1} + \dots + E_{\lambda_p}$, donc $n = \dim(E_{\lambda_1} + \dots + E_{\lambda_p}) = \dim(E_{\lambda_1}) + \dots + \dim(E_{\lambda_p}) \leq n_1 + \dots + n_p = n$. Donc $\sum_{1 \leq j \leq p} (n_j - \dim(E_{\lambda_j})) = 0$. Comme $n_j \geq \dim(E_{\lambda_j})$ on voit que $n_j = \dim(E_{\lambda_j})$ pour $1 \leq j \leq p$, et les conditions 1) et 2) sont vérifiées.

Réciproquement supposons que A satisfait 1) et 2). Soient $\lambda_1, \dots, \lambda_p$ les valeurs propres distinctes de A et pour $1 \leq j \leq p$ soit n_j l'ordre de multiplicité de λ_j .

Comme p_A est scindé, on a de nouveau $\sum_{1 \leq j \leq p} n_j = n$. Soit \mathcal{B}_j une base de E_{λ_j} pour $1 \leq j \leq p$. Alors $\mathcal{B} = \cup_{1 \leq j \leq p} \mathcal{B}_j$ est libre d'après la proposition 4.1. Comme \mathcal{B}_j possède n_j éléments pour $1 \leq j \leq p$, \mathcal{B} possède n éléments et c'est donc une base de K^n formée de vecteurs propres de A , ce qui montre que A est diagonalisable. ♣

On dit qu'une valeur propre λ de $A \in \mathcal{M}_n(K)$ est **simple** quand elle est sa multiplicité $n(\lambda)$ est égale à 1. Comme $1 \leq \dim(E_\lambda) \leq n(\lambda)$, on a dans ce cas

$\dim(E_\lambda) = 1 = n(\lambda)$. Comme tout polynôme $p \in \mathbf{C}[x]$ est scindé dans $\mathbf{C}[x]$, on a alors le corollaire suivant.

Corollaire 4.5.5 *Soit $A \in \mathcal{M}_n(K)$. Si le polynôme caractéristique de A est scindé, et si toutes les valeurs propres de A sont simples, alors A est diagonalisable. En particulier une matrice $A \in \mathcal{M}_n(\mathbf{C})$ dont toutes les valeurs propres sont simples est diagonalisable.*

Exemple 4.5.6 *La matrice $A = \begin{pmatrix} 1/2 & 1/2 & 1/2 \\ 1/4 & 1/2 & 0 \\ 1/4 & 0 & 1/2 \end{pmatrix}$ est diagonalisable sur \mathbf{R} .*

En effet on a vu que $p_A = x(x - \frac{1}{2})(x - 1)$. Donc p_A est scindé dans $\mathbf{R}[x]$, et les valeurs propres $0, \frac{1}{2}$ et 1 de A sont toutes des valeurs propres simples.

On va maintenant **diagonaliser** A , cest à dire trouver une matrice inversible P telle que $P^{-1}AP$ soit diagonale. Pour cela il faut trouver une base de K^3 formée de vecteurs propres de A .

Les éléments du sous-espace propre E_0 sont les solutions du système

$$\begin{cases} x/2 + y/2 + z/2 = 0 \\ x/4 + y/2 = 0 \\ x/4 + y/2 = 0 \end{cases}$$

En additionnant membre à membre les deux dernières équations on obtient la première. On peut donc retirer la première équation et on est ramené au système

$$\begin{cases} x = -2y \\ x = -2z \end{cases}$$

On voit donc que E_0 est l'espace vectoriel de dimension 1 engendré par $\mathbf{f}_1 = \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix}$.

Les éléments du sous-espace $E_{1/2}$ sont les solutions du système

$$\begin{cases} x/2 + y/2 + z/2 = x/2 \\ x/4 + y/2 = y/2 \\ x/4 + z/2 = z/2 \end{cases}$$

La deuxième et la troisième équation donnent $x = 0$, et on est ramené au système

$$\begin{cases} x = 0 \\ z = -y \end{cases}$$

On voit donc que $E_{1/2}$ est l'espace vectoriel de dimension 1 engendré par

$$\mathbf{f}_2 = \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}.$$

Les éléments du sous-espace vectoriel E_1 sont les solutions du système

$$\begin{cases} x/2 + y/2 + z/2 = x \\ x/4 + y/2 = y \\ x/4 + z/2 = z \end{cases}$$

La deuxième équation donne $x = 2y$, la troisième équation donne $x = 2z$, ce qui implique que $y = z$, $x/2 + y/2 + z/2 = 2y = x$. On est donc ramené au système

$$\begin{cases} x = 2y \\ y = z \end{cases}$$

On voit donc que E_1 est l'espace vectoriel de dimension 1 engendré par

$$\mathbf{f}_3 = \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}.$$

Donc $(\mathbf{f}_1, \mathbf{f}_2, \mathbf{f}_3)$ est une base de K^3 formée de vecteurs propres de A et on peut prendre

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

Pour les applications, on aura besoin de calculer P^{-1} . Pour cela on résout le système suivant.

$$\begin{cases} -2x + 2z = a \\ x + y + z = b \\ x - y + z = c \end{cases}$$

$$L_3 \mapsto L_3 + L_2 + L_1.$$

$$\begin{cases} -2x + 2z = a \\ x + y + z = b \\ +4z = a + b + c \end{cases}$$

On obtient $z = \frac{a}{4} + \frac{b}{4} + \frac{c}{4}$, $x = z - \frac{a}{2} = -\frac{a}{4} + \frac{b}{4} + \frac{c}{4}$, et $y = x + z - c = \frac{b}{2} - \frac{c}{2}$.

Comme le système ci-dessus a pour solution $\begin{pmatrix} x \\ y \\ z \end{pmatrix} = P^{-1} \begin{pmatrix} a \\ b \\ c \end{pmatrix}$, on obtient

$$P^{-1} = \begin{pmatrix} -1/4 & 1/4 & 1/4 \\ 0 & 1/2 & -1/2 \\ 1/4 & 1/4 & 1/4 \end{pmatrix}.$$

Posons $D = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1/2 & 0 \\ 0 & 0 & 1 \end{pmatrix}$. On a $P^{-1}AP = D$, et on obtient pour $n \geq 1$

$$\begin{aligned} A^n &= PD^nP^{-1} = P \begin{pmatrix} -2 & 0 & 2 \\ 1 & 1 & 1 \\ 1 & -1 & 1 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1/2^n & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} -1/4 & 1/4 & 1/4 \\ 0 & 1/2 & -1/2 \\ 1/4 & 1/4 & 1/4 \end{pmatrix} \\ &= \begin{pmatrix} -2 & 0 & 2 \\ 1 & 1 & 1 \\ 1 & -1 & 1 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1/2^{n+1} & -1/2^{n+1} \\ 1/4 & 1/4 & 1/4 \end{pmatrix} = \begin{pmatrix} 1/2 & 1/2 & 1/2 \\ 1/4 & 1/4 + 1/2^{n+1} & 1/4 - 1/2^{n+1} \\ 1/4 & 1/4 - 1/2^{n+1} & 1/4 + 1/2^{n+1} \end{pmatrix} \end{aligned}$$

Exemple 4.5.7 Soit $B = \begin{pmatrix} 2 & 1 & 1 \\ 2 & 1 & -2 \\ -1 & 0 & -2 \end{pmatrix}$. Alors $p_B = (x-3)(x+1)^2$, et B n'est diagonalisable ni sur \mathbf{R} , ni sur \mathbf{C} .

En ajoutant aux deux premières colonnes un multiple de la troisième on obtient

$$\begin{aligned} p_B &= - \begin{vmatrix} 2-x & 1 & 1 \\ 2 & 1-x & -2 \\ -1 & 0 & -2-x \end{vmatrix} = - \begin{vmatrix} 0 & 0 & 1 \\ 2+2(2-x) & 3-x & -2 \\ -1+(2-x)(2+x) & 2+x & -2-x \end{vmatrix} \\ &= (x-3) \begin{vmatrix} 2 & 1 \\ 3-x^2 & 2+x \end{vmatrix} = (x-3)(x^2+2x+1) = (x-3)(x+1)^2. \end{aligned}$$

Les éléments du sous-espace propre E_{-1} sont les solutions du système

$$\begin{aligned} \begin{cases} 2x+y+z = -x \\ 2x+y-2z = -y \\ -x-2z = -z \end{cases} &\iff \begin{cases} 2x+y+z = -x \\ 2x+y-2z = -y \\ -x-2z = -z \end{cases} \iff \begin{cases} 3x+y+z = 0 \\ x+y-z = 0 \\ x+z = 0 \end{cases} \\ &\iff \begin{cases} 2x+y = 0 \\ x+z = 0 \end{cases} \end{aligned}$$

Dans \mathbf{R}^3 ou \mathbf{C}^3 , on obtient l'espace vectoriel de dimension 1 engendré par $\begin{pmatrix} 1 \\ -2 \\ -1 \end{pmatrix}$. Donc $\dim(E_{-1}) = 1$. Comme -1 est une valeur propre de B de multiplicité 2, B n'est pas diagonalisable.

Exemple 4.5.8 Soit $C = \begin{pmatrix} 1 & 2 & -2 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}$. Alors C est diagonalisable sur \mathbf{C} , mais C n'est pas diagonalisable sur \mathbf{R} .

$$\text{En effet } p_C = - \begin{vmatrix} 1-x & 2 & -2 \\ 0 & -x & -1 \\ 0 & 1 & -x \end{vmatrix} = (x-1)((-x)^2+1) = (x-1)(x^2+1).$$

Le polynôme x^2+1 n'a pas de racines réelles, donc il est irréductible dans $\mathbf{R}[x]$,

p_C n'est pas scindé dans $\mathbf{R}[x]$, et C n'est pas diagonalisable sur \mathbf{R} . Par contre $x^2 + 1 = (x - i)(x + i)$ et les trois valeurs propres 1, i et $-i$ de $C \in \mathcal{M}_3(\mathbf{C})$ sont simples, donc C est diagonalisable sur \mathbf{C} .

4.6 Endomorphismes diagonalisables

On va maintenant considérer le cas des endomorphismes sur un espace vectoriel de dimension finie. On introduit les notions suivantes.

Définition 4.6.1 Soit E un espace vectoriel de dimension finie sur un corps K , soit $u \in \mathcal{L}(E)$, et soit $\lambda \in K$.

On dit que λ est une valeur propre de u s'il existe $\mathbf{x} \in E \setminus \{0\}$ tel que $u(\mathbf{x}) = \lambda\mathbf{x}$. Dans ce cas on dit que \mathbf{x} est un vecteur propre de u associé à la valeur propre λ , et l'ensemble $E_\lambda(u) = \{\mathbf{x} \in E \mid u(\mathbf{x}) = \lambda\mathbf{x}\}$ est appelé le sous-espace propre associé à la valeur propre λ .

Le résultat simple suivant permet de se ramener au cas des matrices.

Proposition 4.6.2 Soit E un espace vectoriel de dimension finie sur un corps K , soit \mathcal{B} une base de E , soit $u \in \mathcal{L}(E)$, et soit $\lambda \in K$.

Pour que λ soit une valeur propre de u , il faut et il suffit que λ soit une valeur propre de la matrice $M_{u,\mathcal{B}}$ représentant u dans la base \mathcal{B} . Le sous-espace propre $E_\lambda(u)$ est alors l'ensemble des éléments \mathbf{x} de E pour lesquels le vecteur colonne $[\mathbf{x}]_{\mathcal{B}}$ formé des coordonnées de \mathbf{x} dans la base \mathcal{B} appartient au sous-espace propre $E_\lambda(M_{u,\mathcal{B}})$ associé à λ .

Démonstration : Soit $\mathbf{x} \in E$. Posons $X = [\mathbf{x}]_{\mathcal{B}}$, et soit Y le vecteur colonne formé des coordonnées de $u(\mathbf{x})$ dans la base \mathcal{B} . Il résulte de la formule (2.1) que $Y = M_{u,\mathcal{B}}X$. Par conséquent $u(\mathbf{x}) = \lambda\mathbf{x}$ si et seulement si $M_{u,\mathcal{B}}X = \lambda X$. La proposition résulte alors immédiatement de cette observation. ♣

Soit $\mathcal{I}_E : \mathbf{x} \mapsto \mathbf{x}$ l'application identité sur E . On définit le **polynôme caractéristique** p_u de $u \in \mathcal{L}(E)$ par la formule

$$(4.6) \quad p_u = \det(x\mathcal{I}_E - u).$$

Il résulte de la Proposition 3.9 que le polynôme caractéristique de u coïncide avec le polynôme caractéristique de la matrice $M_{u,\mathcal{B}}$ pour toute base \mathcal{B} de E , et il résulte de la Proposition 4.22 que l'ensemble des valeurs propres de u coïncide avec l'ensemble des racines du polynôme caractéristique p_u .

Définition 4.6.3 Soit E un espace vectoriel de dimension finie sur un corps K , et soit $u \in \mathcal{L}(E)$. On dit que u est diagonalisable s'il existe une base de E formée de vecteurs propres de u .

Proposition 4.6.4 Soit E un espace vectoriel de dimension finie sur un corps K , soient \mathcal{B} et \mathcal{B}' deux bases de E et soit $u \in \mathcal{L}(E)$.

Alors la base \mathcal{B}' est formée de vecteurs propres de u si et seulement la matrice de passage $P_{\mathcal{B},\mathcal{B}'}$ est telle que $P_{\mathcal{B},\mathcal{B}'}^{-1}M_{u,\mathcal{B}}P_{\mathcal{B},\mathcal{B}'}$ est diagonale.

En particulier que u soit diagonalisable, il faut et il suffit que la matrice $M_{u,\mathcal{B}}$ représentant u dans la base \mathcal{B} soit diagonalisable.

Démonstration : Il résulte de la formule (2.6) que $M_{u,\mathcal{B}'} = P_{\mathcal{B},\mathcal{B}'}^{-1}M_{u,\mathcal{B}}P_{\mathcal{B},\mathcal{B}'}$. Donc pour que \mathcal{B}' soit une base de E formée de vecteurs propres de u il faut et il suffit que $P_{\mathcal{B},\mathcal{B}'}^{-1}M_{u,\mathcal{B}}P_{\mathcal{B},\mathcal{B}'}$ soit diagonale. Ceci prouve que $M_{u,\mathcal{B}}$ est diagonalisable si u est diagonalisable. Réciproquement supposons que $M_{u,\mathcal{B}}$ est diagonalisable, et soit $P \in \mathcal{M}_n(K)$ une matrice inversible telle que $P^{-1}M_{u,\mathcal{B}}P$ soit diagonale. Soient $\mathbf{p}_1, \dots, \mathbf{p}_n$ les colonnes de P et pour $1 \leq j \leq n$ soit \mathbf{f}_j l'élément de E tel que le vecteur colonne formé des coordonnées de \mathbf{f}_j dans la base \mathcal{B} soit égal à \mathbf{p}_j . Posons $\mathcal{B}' = (\mathbf{f}_1, \dots, \mathbf{f}_n)$. On a $P_{\mathcal{B},\mathcal{B}'} = P$. Comme P est inversible, il résulte de la Proposition 2.10 que \mathcal{B}' est une base de E , et il résulte de ce qui précède que \mathbf{f}_j est un vecteur propre de u pour $1 \leq j \leq n$, ce qui montre que u est diagonalisable. ♣

Rappelons (voir l'exemple 1.15) que si $E = F \oplus G$ on note $P_{F,G}$ l'endomorphisme de E qui à $\mathbf{x} \in E$ associe l'unique $\mathbf{y} \in F$ tel que $\mathbf{x} - \mathbf{y} \in G$.

Exemple 4.6.5 Soit E un espace vectoriel de dimension finie, et soient F et G deux sous-espaces vectoriels de E tels que $E = F \oplus G$. Alors la projection $P_{F,G}$ est diagonalisable.

Si $F = \{0\}$ alors $P_{F,G}(\mathbf{x}) = 0$ pour tout $\mathbf{x} \in G$, et le résultat est évident. De même si $G = \{0\}$ alors $P_{F,G}(\mathbf{x}) = \mathbf{x}$ pour tout $\mathbf{x} \in G$, et le résultat est aussi évident. Si $F \neq \{0\}$ et si $G \neq \{0\}$, soit \mathcal{B}_1 une base de F et soit \mathcal{B}_2 une base de G . On a vu au Chapitre 1 que $\mathcal{B} = \mathcal{B}_1 \cup \mathcal{B}_2$ est une base de E . Comme $P_{F,G}(\mathbf{x}) = \mathbf{x}$ pour $\mathbf{x} \in \mathcal{B}_1$, et comme $P_{F,G}(\mathbf{x}) = 0$ pour $\mathbf{x} \in \mathcal{B}_2$ on voit que \mathcal{B} est une base de E formée de vecteurs propres de $P_{F,G}$, ce qui montre que $P_{F,G}$ est diagonalisable (on voit également que le polynôme caractéristique de $P_{F,G}$ est égal à $(-1)^{p+q}(x-1)^p x^q$, où $p = \dim(F)$ et $q = \dim(G)$).

Exemple 4.6.6 Posons $u(p) = (x-1)p' + p$ pour $p \in \mathbf{R}_2[x]$. Alors u est un endomorphisme diagonalisable de $\mathbf{R}_2[x]$ qui a pour valeurs propres 1, 2 et 3, et $(1, x-1, x^2-2x+1)$ est une base de $\mathbf{R}_2[x]$ formée de vecteurs propres de u .

En effet posons $\mathcal{B}_0 = (1, x, x^2)$. Alors \mathcal{B}_0 est une base de $\mathbf{R}_2[x]$. On a $u(1) = 1$, $u(x) = x - 1 + x = 2x - 1$, $u(x^2) = (x-1)2x + x = 3x^2 - 2x$, donc

$$M_{u,\mathcal{B}_0} = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 2 & -2 \\ 0 & 0 & 3 \end{pmatrix}, \text{ et } p_u = (x-1)(x-2)(x-3). \text{ Comme } u(1) =$$

$$1, u \in E_1. \text{ L'équation } \begin{pmatrix} 1 & -1 & 0 \\ 0 & 2 & -2 \\ 0 & 0 & 3 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = 2 \begin{pmatrix} a \\ b \\ c \end{pmatrix} \text{ donne le système}$$

$$\begin{cases} a - b = 2a \\ 2b - 2c = 2b \text{ soit } c = 0, a = -b. \text{ Donc } \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} \text{ convient, et } x-1 \in E_2. \text{ L'équa-} \\ 3c = 2c \end{cases}$$

tion $\begin{pmatrix} 1 & -1 & 0 \\ 0 & 2 & -2 \\ & 0 & 3 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = 3 \begin{pmatrix} a \\ b \\ c \end{pmatrix}$ donne le système $\begin{cases} a - b = 3a \\ 2b - 2c = 3b \text{ soit} \\ 3c = 3c \end{cases}$

$b = -2a = -2c$. Donc $\begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}$ convient, $x^2 - 2x + 1 \in E_2$, et $(1, x-1, x^2 - 2x + 1)$ est une base de $\mathbf{R}_2[x]$ formée de vecteurs propres de u .

Exemple 4.6.7 Posons $v(p) = 2p' + p$ pour $p \in \mathbf{R}_2[x]$. Alors v est un endomorphisme non diagonalisable de $\mathbf{R}_2[x]$.

En effet considérons de nouveau la base $\mathcal{B}_0 = (1, x, x^2)$ de $\mathbf{R}_2[x]$. On a $v(1) = 1$, $v(x) = x + 2$, $v(x^2) = x^2 + 4x$, donc $M_{v, \mathcal{B}_0} = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & 4 \\ 0 & 0 & 1 \end{pmatrix}$, donc $p_v = -(x-1)^3$. Si v était diagonalisable, il existerait $P \in \mathcal{M}_3(\mathbf{R})$ inversible telle que $P^{-1} \begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & 4 \\ 0 & 0 & 1 \end{pmatrix} P = I_3$. On aurait alors $\begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & 4 \\ 0 & 0 & 1 \end{pmatrix} = P I_3 P^{-1} = I_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$, ce qui est évidemment absurde. Donc v n'est pas diagonalisable.

4.7 Chapitre 4 sous MUPAD

On peut demander à Mupad de calculer le polynôme caractéristique d'une matrice en utilisant la commande `linalg : : charpoly(A,x)`;

```

M:=Dom::Matrix();
A:=M([[1,2,3,4,5], [1,2,1,2,1], [1,-1,4,3,-1], [2,-1,3,1,4], [5,-2,3,-1,3]]);
linalg::charpoly(A,x);

```

```

Dom::Matrix(Dom::ExpressionField(id, iszero))

```

```

+-          +-
|  1,  2,  3,  4,  5  |

```

$$\begin{array}{cccccc}
 | & & & & & | \\
 | & 1, & 2, & 1, & 2, & 1 & | \\
 | & & & & & & | \\
 | & 1, & -1, & 4, & 3, & -1 & | \\
 | & & & & & & | \\
 | & 2, & -1, & 3, & 1, & 4 & | \\
 | & & & & & & | \\
 | & 5, & -2, & 3, & -1, & 3 & | \\
 +- & & & & & & +-
 \end{array}$$

$$8x^5 + 60x^4 + 10x^3 - 11x^2 + x - 56$$

$$\text{Si } A = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 2 & 1 & 2 & 1 \\ 1 & -1 & 4 & 3 & -1 \\ 2 & -1 & 3 & 1 & 4 \\ 5 & -2 & 3 & -1 & 3 \end{pmatrix}, \text{ on a donc } p_A = -x^5 + 11x^4 - 10x^3 -$$

$$60x^2 - 8x + 56.$$

On peut également utiliser Mupad pour calculer les valeurs propres, avec la commande `linalg : eigenvalues(A)`; Mais le calcul symbolique bute sur les équations algébriques de degré supérieur ou égal à 5.

```
linalg::eigenValues(A);
```

$$\{\text{RootOf}(8 \text{ ______eigenValue_x} + 60 \text{ ______eigenValue_x}^2 + 10 \text{ ______eigenValue_x}^3 - 11 \text{ ______eigenValue_x}^4 + \text{ ______eigenValue_x}^5 - 56, \text{ ______eigenValue_x})\}$$

Le logiciel ne sait que répéter que les valeurs propres sont les racines du polynôme caractéristique. Il faudrait faire appel aux possibilités de **calcul numérique sous Mupad** pour obtenir des valeurs approchées des valeurs propres de la matrice ci-dessus, ce que nous ferons un peu plus loin.

On a plus de succès avec les matrices à 4 lignes et 4 colonnes, mais on se heurte à la complexité des formules de Cardan-Tartaglia (voir le *Cours d'algèbre*).

```
M:=Dom::Matrix();
```

```
A:=M([[1,2,3,4],[1,2,1,2],[1,-1,4,3],[2,-1,3,1]]);
linalg::charpoly(A,x);
linalg::eigenvalues(A);
```

```
Dom::Matrix(Dom::ExpressionField(id, iszero))
```

```
+ - - +
| 1, 2, 3, 4 |
|           |
| 1, 2, 1, 2 |
|           |
| 1, -1, 4, 3 |
|           |
| 2, -1, 3, 1 |
+ - - +
```

$$36 x^2 + 2 x^3 - 8 x^4 + x^4 + 4$$

```
{
{ 2 - (132 (40/9 I 3 5263 1/2 + 23264/27) 1/2 + 1/3
{
```

$$9 (40/9 I 3 5263 1/2 + 23264/27) 1/2 + 916)^{(1/2)} /$$

$$(6 (40/9 I 3 5263 1/2 + 23264/27) 1/2 - (264$$

... suivent 3 pages de formules, dont voici la fin

$$132 (40/9 I 3 5263 1/2 + 23264/27) 1/2 +$$

$$9 (40/9 I 3 5263 1/2 + 23264/27) 1/2 + 916)^{(1/2)} / (6 2/3$$

$$(40/9 I 3 5263 1/2 + 23264/27) 1/2 ($$

$$\begin{aligned}
& 132 \left(\frac{40}{9} I^3 \right)^{\frac{1}{2}} \left(\frac{5263}{27} \right)^{\frac{1}{2}} + \frac{23264}{27} \left(\frac{1}{3} \right)^{\frac{1}{3}} + \\
& 9 \left(\frac{40}{9} I^3 \right)^{\frac{1}{2}} \left(\frac{5263}{27} \right)^{\frac{1}{2}} + \frac{23264}{27} \left(\frac{2}{3} \right)^{\frac{2}{3}} + 916 \sqrt[4]{} + 2 \} \\
& \phantom{9 \left(\frac{40}{9} I^3 \right)^{\frac{1}{2}} \left(\frac{5263}{27} \right)^{\frac{1}{2}} + \frac{23264}{27} \left(\frac{2}{3} \right)^{\frac{2}{3}} + 916 \sqrt[4]{} + 2 } \}
\end{aligned}$$

On voit que dans ce cas Mupad a donné en plus de 3 pages des formules certes exactes, mais illisibles et inutilisables.

On a dans même dans les cas où le calcul symbolique fonctionne intérêt à tenter sa chance avec les possibilités de calcul numérique de Mupad, en utilisant la commande `numeric::eigenvalues(A)`; On applique également cette méthode à la première des matrices étudiées plus haut

```

M:=Dom::Matrix();
A:=M([[1,2,3,4,5],[1,2,1,2,1],[1,-1,4,3,-1],[2,-1,3,1,4],[5,-2,3,-1,3]]);
linalg::charpoly(A,x);
numeric::eigenvalues(A);

```

```
Dom::Matrix(Dom::ExpressionField(id, iszero))
```

```

+-          +-
|  1,  2,  3,  4,  5 |
|                      |
|  1,  2,  1,  2,  1 |
|                      |
|  1, -1,  4,  3, -1 |
|                      |
|  2, -1,  3,  1,  4 |
|                      |
|  5, -2,  3, -1,  3 |
+-          +-

```

$$8x^5 + 60x^4 + 10x^3 - 11x^2 + x - 56$$

```
[- 1.313899679 + 0.4452307358 I, - 1.313899679 - 0.4452307358 I,
```

```
0.8958378463, 3.529623176, 9.202338336]
```

```
M:=Dom::Matrix();
B:=M([[1,2,3,4],[1,2,1,2],[1,-1,4,3],[2,-1,3,1]]);
numeric::eigenvalues(B);
```

```
Dom::Matrix(Dom::ExpressionField(id, iszero))
```

```
+-          +-
|  1,  2,  3,  4 |
|                |
|  1,  2,  1,  2 |
|                |
|  1, -1,  4,  3 |
|                |
|  2, -1,  3,  1 |
+-          +-
```

```
[-1.75933678, -0.1121272467, 2.914685186, 6.95677884]
```

On voit donc que la matrice $A = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 2 & 1 & 2 & 1 \\ 1 & -1 & 4 & 3 & -1 \\ 2 & -1 & 3 & 1 & 4 \\ 5 & -2 & 3 & -1 & 3 \end{pmatrix}$ possède 2 valeurs

propres non réelles, approximativement égales à $-1.313899679 + 0.4452307358i$ et $-1.313899679 - 0.4452307358i$, et trois valeurs propres réelles, approximativement égales à 0.8958378463, 3.529623176, et 9.202338336.

De même la matrice $B = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 1 & 2 \\ 1 & -1 & 4 & 3 \\ 2 & -1 & 3 & 1 \end{pmatrix}$ possède 4 valeurs propres réelles,

approximativement égales à $-1.75933678, -0.1121272467, 2.914685186$ et 6.95677884 .

On revient maintenant à des matrices accessibles au calcul symbolique. Avec la commande `linalg :eigenvectors(A)`, on obtient pour chaque valeur propre l'indication de l'ordre de multiplicité et une base du sous-espace propre correspondant.

```
M:=Dom::Matrix();
A:=M([ [2,1,1],[2,1,-2],[-1,0,-2]]);
linalg::eigenvalues(A);
linalg::eigenvectors(A);
```

```
Dom::Matrix(Dom::ExpressionField(id, iszero))
```

```
+ -      - +
|  2, 1,  1 |
|           |
|  2, 1, -2 |
|           |
| -1, 0, -2 |
+ -      - +
```

```
{-1, 3}
```

```
-- --      -- +-      -+ -- -- --      -- +-      -+ -- -- --
| |      | | -1 | | | |      | | -5 | | | | |
| |      | |   | | | |      | |   | | | |
| | -1, 2, | |  2 | | | |, |  3, 1, | | -6 | | | |
| |      | |   | | | |      | |   | | | |
| |      | |  1 | | | |      | |  1 | | | |
```

Ceci signifie que la matrice $A = \begin{pmatrix} 2 & 1 & 1 \\ 2 & 1 & -2 \\ -1 & 0 & -2 \end{pmatrix}$ a -1 comme valeur propre de multiplicité 2 et 3 comme valeur propre simple, et que le sous-espace propre E_{-1} est le sous-espace vectoriel de dimension 1 de \mathbf{R}^3 engendré par $\begin{pmatrix} -1 \\ 2 \\ 1 \end{pmatrix}$, tandis que le sous-espace propre E_3 est le sous-espace vectoriel de dimension 1 de \mathbf{R}^3 engendré par $\begin{pmatrix} -5 \\ -6 \\ 1 \end{pmatrix}$. On voit donc que $\begin{pmatrix} 2 & 1 & 1 \\ 2 & 1 & -2 \\ -1 & 0 & -2 \end{pmatrix}$ **n'est pas diagonalisable.**

On peut également diagonaliser des matrices réelles ou complexes diagonalisable s sous Mupad. On utilise pour cela la commande `linalg : :jordanform(A,All)`; On obtient deux matrices. La première matrice Δ est triangulaire inférieure, et A est diagonalisable uniquement dans le cas où Δ est diagonale

(nous expliquerons ce phénomène au Chapitre suivant). La deuxième matrice est une matrice inversible P telle que $P^{-1}AP = \Delta$. Avec la commande `linalg : jordanForm(A)`, on obtient seulement la matrice Δ . Nous donnons deux exemples.

```
M:=Dom::Matrix();
A:=M([[1,-1,0],[0,2,-2],[0,0,3]]);
linalg::jordanForm(A,All);
```

```
Dom::Matrix(Dom::ExpressionField(id, iszero))
```

```
+-      +-
|  1, -1,  0 |
|          |
|  0,  2, -2 |
|          |
|  0,  0,  3 |
+-      +-
```

```
-- +-      -+ +-      -+ --
| |  1, 0, 0 | |  1, -1,  1 | |
| |          | |          | |
| |  0, 2, 0 | |  0,  1, -2 | |
| |          | |          | |
| |  0, 0, 3 | |  0,  0,  1 | |
-- +-      -+ +-      -+ --
```

Ceci signifie que la matrice $\begin{pmatrix} 1 & -1 & 0 \\ 0 & 2 & -2 \\ 0 & 0 & 3 \end{pmatrix}$ est diagonalisable, et que

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

On reprend un autre exemple où la matrice n'est pas diagonalisable.

```
M:=Dom::Matrix();
A:=M([ [2,1,1],[2,1,-2],[-1,0,-2]]);
linalg::jordanForm(A,All);
```

```
Dom::Matrix(Dom::ExpressionField(id, iszero))
```

$$\begin{array}{c}
 \begin{array}{c}
 +- \\
 | \\
 | \\
 | \\
 | \\
 +-
 \end{array}
 \begin{array}{c}
 2, 1, 1 \\
 2, 1, -2 \\
 -1, 0, -2
 \end{array}
 \begin{array}{c}
 -+ \\
 | \\
 | \\
 | \\
 -+
 \end{array}
 \end{array}$$

$$\begin{array}{c}
 \begin{array}{c}
 -- \quad +- \\
 | \quad | \\
 | \quad | \\
 | \quad | \\
 | \quad | \\
 -- \quad +-
 \end{array}
 \begin{array}{c}
 -1, 0, 0 \\
 1, -1, 0 \\
 0, 0, 3
 \end{array}
 \begin{array}{c}
 -+ \quad +- \\
 | \quad | \\
 | \quad | \\
 | \quad | \\
 -+ \quad +-
 \end{array}
 \begin{array}{c}
 -1/2, -1/2, -5 \\
 1, 1, -6 \\
 0, 1/2, 1
 \end{array}
 \begin{array}{c}
 -+ \quad -- \\
 | \quad | \\
 | \quad | \\
 | \quad | \\
 -+ \quad --
 \end{array}
 \end{array}$$

La matrice triangulaire inférieure qui apparait en premier n'est pas diagonale, et on retrouve le fait que $\begin{pmatrix} 2 & 1 & 1 \\ 2 & 1 & -2 \\ -1 & 0 & -2 \end{pmatrix}$ n'est pas diagonalisable.

4.8 Exercices sur le Chapitre 4

exercice 1

Dans chacun des cas suivants, déterminer si la matrice $A \in \mathcal{M}_3(\mathbf{C})$ est diagonalisable. Si oui, trouver une base de \mathbf{C}^3 formée de vecteurs propres de A et calculer A^n pour $n \in \mathbf{N}$.

$$a) A = \begin{pmatrix} 5 & -3 & 2 \\ 6 & -4 & 4 \\ 4 & -4 & 5 \end{pmatrix}.$$

$$b) A = \begin{pmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \\ 1 & 0 & -1 \end{pmatrix}.$$

$$c) A = \begin{pmatrix} -2 & -2 & 1 \\ -2 & 1 & -2 \\ 1 & -2 & -2 \end{pmatrix}.$$

$$d) A = \begin{pmatrix} 0 & 1 & 0 \\ -4 & 4 & 0 \\ -2 & 1 & 2 \end{pmatrix}.$$

exercice 2

Soit $A \in \mathcal{M}_n(K)$.

a) Montrer que si λ est valeur propre de A alors λ^k est valeur propre de A^k pour $k \geq 1$. b) Soit $p = a_0 + a_1x + \dots + a_px^p \in K[x]$. On pose $p(A) = a_0I_n + a_1A + \dots + a_pA^p$. Montrer que si λ est valeur propre de A alors $p(\lambda)$ est valeur propre de $p(A)$.

c) On suppose de plus que A est inversible. On pose $A^{-n} = (A^{-1})^n$ pour $n > 0$. Montrer que si λ est valeur propre de A , alors λ^n est valeur propre de A^n pour $n \in \mathbf{Z} \setminus \{0\}$.

exercice 3

Soit $A \in \mathcal{M}_n(K)$ une matrice inversible.

a) Trouver une relation entre le polynôme caractéristique de A et celui de A^{-1} .

b) En déduire que si λ est valeur propre de A alors λ^{-1} est valeur propre de A^{-1} avec le même ordre de multiplicité.

exercice 4

a) Soit $D : p \mapsto p'$ l'opérateur de dérivation sur $\mathbf{R}_n[x]$. Déterminer les valeurs propres de D et les sous-espaces propres associés. L'endomorphisme D est-il diagonalisable ?

b) Soit $E = \mathcal{C}^\infty(\mathbf{R})$ l'espace vectoriel des fonctions indéfiniment dérivables de \mathbf{R} dans \mathbf{R} . On pose $d(f) = f'$ pour $f \in \mathcal{C}^\infty(\mathbf{R})$. Montrer que d est un endomorphisme de E . Déterminer les valeurs propres de d et les sous-espaces propres correspondants.

exercice 5

Soit F le sous-espace vectoriel de \mathbf{R}^3 défini par l'équation $x + y - 7z = 0$ et soit G le sous-espace vectoriel de dimension 1 de \mathbf{R}^3 engendré par $\begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix}$.

- 1) Montrer que $\mathbf{R}^3 = F \oplus G$.
- 2) Donner la matrice A représentant u dans la base canonique de \mathbf{R}^3 .
- 3) En étudiant A , retrouver le fait que la projection $P_{F,G}$ est diagonalisable

exercice 6

Résoudre l'exercice 1 ainsi que la troisième question de l'exercice 5 en utilisant Mupad.

exercice 7 (sous Mupad)

On considère $A = \begin{pmatrix} 1 & 9 & -9 & 6 & 7 & 8 & -2 & 0 & 6 & 12 \\ -5 & 6 & -7 & 6 & 1 & 4 & -12 & 7 & 16 & 2 \\ 4 & 3 & 2 & 1 & 9 & 8 & 7 & 6 & 5 & 4 \\ 1 & 2 & 3 & 4 & 5 & 6 & 5 & 4 & 3 & 2 \\ 1 & 9 & -9 & 6 & 7 & 8 & -2 & 0 & 6 & 12 \\ 2 & 4 & 6 & 8 & 10 & 12 & 14 & 16 & 18 & 20 \\ 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 \\ 2 & 3 & 3 & 2 & 2 & 3 & 4 & 5 & 6 & 7 \\ 2 & 1 & 0 & 2 & 1 & 0 & -1 & -2 & -3 & -2 \\ 5 & 3 & 1 & -1 & -3 & -5 & -7 & -9 & -11 & 8 \end{pmatrix}$.

- 1) Déterminer le polynôme caractéristique de A .
- 2) Déterminer des valeurs approchées des valeurs propres de A .
- 3) En utilisant la commande `numeric : eigenvectors(A)` ; donner pour chaque valeur propre λ de A une valeur approchée d'un vecteur propre \mathbf{x}_λ ainsi qu'une estimation de la différence $A\mathbf{x}_\lambda - \lambda\mathbf{x}_\lambda$.

Chapitre 5

Polynôme minimal, décomposition de Jordan

Le résultat principal de ce Chapitre est le *théorème de décomposition de Jordan* : si le polynôme caractéristique de $A \in \mathcal{M}_n(K)$ est scindé alors A possède une unique décomposition de la forme $A = D + N$, avec D diagonalisable, N nilpotente vérifiant $DN = ND$.

5.1 Théorème de Cayley-Hamilton

Nous allons commencer par démontrer l'important théorème de Cayley-Hamilton.

Théorème 5.1.1 *Soit K un corps, et soit $A \in \mathcal{M}_n(K)$. Alors $p_A(A) = 0$.*

Démonstration : Soit \mathcal{B} un anneau commutatif. On peut munir l'ensemble $\mathcal{M}_n(\mathcal{B})$ des matrices carrées à n lignes et n colonnes à coefficients dans \mathcal{B} de l'addition et de la multiplication usuelles des matrices, et on obtient ainsi un anneau.

Pour $B = (b_{i,j})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq n}} \in \mathcal{M}_n(\mathcal{B})$, on définit le **déterminant** de B par la formule

$$\det(B) = \sum_{\sigma \in \mathcal{S}_n} \epsilon(\sigma) b_{\sigma(1),1} \cdots b_{\sigma(n),n}.$$

On a des propriétés analogues à celles des déterminants d'une matrice à coefficients dans un corps. En particulier si on note $C_{i,j}(B)$ le déterminant de la matrice obtenue en retirant à B sa i^e ligne et sa j^e colonne, on a pour $1 \leq i \leq n$ la formule

$$(5.1) \quad \det(B) = \sum_{j=1}^n (-1)^{i+j} b_{i,j} C_{i,j}(B) b_{i,j} \quad .$$

En appliquant cette formule à la matrice obtenue en remplaçant la k^e colonne de B par la i^e , qui est de déterminant nul, on obtient, pour $1 \leq i \leq n$, $1 \leq k \leq n$, $k \neq i$

$$(5.2) \quad \det(B) = \sum_{j=1}^n (-1)^{i+j} C_{k,j}(B) b_{i,j} \quad .$$

Soit maintenant $A = (a_{i,j})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq n}} \in \mathcal{M}_n(K)$. On a $p_A = \det(xI_n - A) = \alpha_0 + a_1x + \dots + \alpha_{n-1}x^{n-1} + x^n$, avec $\alpha_0, \dots, \alpha_{n-1} \in K$ (on notera que $\alpha_0 = (-1)^n \det(A)$ et que $\alpha_{n-1} = -\text{Tr}(A)$). Soit $\mathcal{B} = \{q(A)\}_{q \in K[x]}$. Il est clair que \mathcal{B} est un anneau commutatif. Comme $p_A(A) = \alpha_0 I_n + \alpha_1 A + \dots + \alpha_{n-1} x^{n-1} + x^n$, il résulte de la définition des déterminants que si on pose $\mathbf{B} = (B_{i,j})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq n}}$, avec $B_{i,j} = -a_{i,j} I_n$ pour $1 \leq i \leq n$, $1 \leq j \leq n$, $i \neq j$, et $B_{i,i} = A - a_{i,i} I_n$ pour $1 \leq i \leq n$, alors $p_A(A) = \det(\mathbf{B})$.

Soit (e_1, \dots, e_n) la base canonique de K^n . On sait que la i^e colonne de A est égale à Ae_i pour $1 \leq i \leq n$. On a $B_{i,j}e_i = -a_{i,j}I_n e_i = -a_{i,j}e_i$ pour $i \neq k$, et $B_{j,j}e_j = Ae_j - a_{j,j}I_n e_j = Ae_j - a_{j,j}e_j$. On obtient, pour $1 \leq j \leq n$

$$\sum_{i=1}^n B_{i,j}e_i = Ae_j - \sum_{i=1}^n a_{i,j}e_i = \begin{pmatrix} a_{1,j} \\ \cdot \\ \cdot \\ a_{n,j} \end{pmatrix} - \begin{pmatrix} a_{1,j} \\ \cdot \\ \cdot \\ a_{n,j} \end{pmatrix} = \begin{pmatrix} 0 \\ \cdot \\ \cdot \\ 0 \end{pmatrix}.$$

On obtient donc le système d'équations

$$\begin{cases} B_{1,1}e_1 + \dots + B_{1,n}e_n = 0 \\ \dots \dots \dots \\ B_{i,1}e_1 + \dots + B_{i,n}e_n = 0 \\ \dots \dots \dots \\ B_{n,1}e_1 + \dots + B_{n,n}e_n = 0 \end{cases}$$

Ceci peut s'écrire sous la forme matricielle $\mathbf{B} \begin{pmatrix} e_1 \\ \cdot \\ \cdot \\ e_n \end{pmatrix} = \begin{pmatrix} 0 \\ \cdot \\ \cdot \\ 0 \end{pmatrix}$, où 0 désigne

dans le second membre la matrice unicolonne dont tous les coefficients sont nuls.

Soit maintenant $\mathbf{C} \in \mathcal{M}_n(\mathcal{B})$. L'associativité du produit matriciel reste valable dans ce contexte, et on a

$$\mathbf{C}\mathbf{B} \begin{pmatrix} e_1 \\ \cdot \\ \cdot \\ e_n \end{pmatrix} = \mathbf{C} \left(\mathbf{B} \begin{pmatrix} e_1 \\ \cdot \\ \cdot \\ e_n \end{pmatrix} \right) = \mathbf{C} \begin{pmatrix} 0 \\ \cdot \\ \cdot \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ \cdot \\ \cdot \\ 0 \end{pmatrix}.$$

Posons maintenant $\mathbf{C} = ((-1)^{i+j}C_{j,i}(\mathbf{B}))_{\substack{1 \leq i \leq n \\ 1 \leq j \leq n}} \in \mathcal{M}_n(\mathcal{B})$, et soit $\mathbf{D} = \mathbf{C}\mathbf{B}$.

Il résulte des formules (5.1) et (5.2) que, de même que dans le cas des matrices à coefficients dans K , la matrice D est une matrice *diagonale* : si on pose $\mathbf{D} = (D_{i,j})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq n}}$, alors $D_{i,i} = \det(\mathbf{B})$ pour $1 \leq i \leq n$, et $D_{i,j} = 0$ pour $1 \leq i \leq n$, $1 \leq j \leq n$, $i \neq j$.

$$\text{On a alors } \begin{pmatrix} \det(\mathbf{B})e_1 \\ \vdots \\ \det(\mathbf{B})e_n \end{pmatrix} = \mathbf{C} \left(\mathbf{B} \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix} \right) = \mathbf{C} \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}. \text{ Donc}$$

$\det(\mathbf{B})e_j = 0$ pour $1 \leq j \leq n$, et $p_A(A) = \det(\mathbf{B}) = 0$. ♣

La théorème de Cayley-Hamilton permet notamment de calculer simplement l'inverse d'une matrice inversible $A \in \mathcal{M}_n(K)$: en écrivant $p_A = (-1)^n \det(A) + \alpha_1 x + \dots + \alpha_{n-1} x^{n-1} + x^n$, on a $0 = p_A(A) = (-1)^n \det(A) I_n + A(\alpha_1 I_n + \dots + \alpha_{n-1} A^{n-2} + A^{n-1})$, donc $A^{-1} = \frac{(-1)^{n+1}}{\det(A)} (\alpha_1 I_n + \dots + \alpha_{n-1} A^{n-2} + A^{n-1})$ si $\det(A) \neq 0$.

5.2 Polynôme minimal

Rappelons (voir la Définition 3.7 du *Cours d'Algèbre*) qu'on dit que $I \subset K[x]$ est un *idéal* de $K[x]$ quand les deux conditions suivantes sont vérifiées

- (i) I est un sous-groupe de $(K[x], +)$,
- (ii) $PQ \in I \quad \forall P \in I, \forall Q \in K[x]$.

D'autre part on sait d'après le théorème 3.8 du **Cours d'algèbre**[?] que pour tout idéal I de $K[x]$ non réduit à $\{0\}$ il existe un unique polynôme unitaire $B \in K[x]$ tel que $I = BK[x]$. Ceci permet d'introduire la notion de polynôme minimal.

Proposition 5.2.1 *Soit $A \in \mathcal{M}_n(K)$. L'ensemble $\mathcal{I}_A = \{q \in K[x] \mid q(A) = 0\}$ est un idéal de $K[x]$ non réduit à $\{0\}$, appelé **idéal annulateur** de A . Le polynôme unitaire $q_A \in K[x]$ tel que $\mathcal{I}_A = q_A \cdot K[x]$ est appelé le **polynôme minimal** de A .*

Démonstration : Comme $p_A \in \mathcal{I}_A$ on a $\mathcal{I}_A \neq \{0\}$. Si $p \in \mathcal{I}_A$, $q \in \mathcal{I}_A$ on a $(p - q)(A) = p(A) - q(A) = 0$, et $p - q \in \mathcal{I}_A$, ce qui prouve que \mathcal{I}_A est un sous-groupe de $(K[x], +)$. D'autre part si $p \in \mathcal{I}_A$, $q \in K[x]$ alors $(pq)(A) = p(A)q(A) = 0$, et $pq \in \mathcal{I}_A$. ♣

On notera que $q_A = q_A \cdot 1 \in \mathcal{I}_A$, donc $q_A(A) = 0$.

Proposition 5.2.2 *Soit $A \in \mathcal{M}_n(K)$. Alors le polynôme minimal q_A de A divise le polynôme caractéristique p_A de A , et toute racine de p_A est racine de q_A .*

Démonstration : Il résulte du théorème de Cayley-Hamilton que $p_A \in \mathcal{I}_A = q_A K[x]$, donc q_A divise p_A .

Soit λ une valeur propre de A , et soit $X \neq 0$ un élément de K^n tel que $AX = \lambda X$. D'après le lemme 4.3 on a $q_A(\lambda)X = q_A(A)X = 0$, donc $q_A(\lambda) = 0$.

♣

Exemple 5.2.3 Soit $A \in \mathcal{M}_4(\mathbf{R})$ telle que $p_A = (x^2 - 1)^2$. Alors q_A est l'un des 4 polynômes $(x-1)(x+1)$, $(x-1)^2(x+1)$, $(x-1)(x+1)^2$, $(x-1)^2(x+1)^2$, et il existe des matrices correspondant à chacun des quatre cas.

En effet il résulte de la proposition que $q_A = (x-1)^a(x+1)^b$, avec $1 \leq a \leq 2$, $1 \leq b \leq 2$, ce qui correspond bien aux quatre cas ci-dessus. Pour trouver des exemples correspondant à chacun des cas, on va considérer des matrices de la forme $A = \begin{pmatrix} B & 0_2 \\ 0_2 & C \end{pmatrix}$ avec $B \in \mathcal{M}_2(\mathbf{R})$ triangulaire supérieure de termes diagonaux égaux à 1, $C \in \mathcal{M}_2(\mathbf{R})$ triangulaire supérieure de termes diagonaux égaux à -1, le symbole 0_2 désignant la matrice $\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$, ce qui garantit que $p_A = (x-1)^2(x+1)^2$, $p_B = (x-1)^2$, $p_C = (x+1)^2$. Dans cette situation on a un "produit par blocs" qui donne $A^m = \begin{pmatrix} B^m & 0_2 \\ 0_2 & C^m \end{pmatrix}$ pour $m \geq 0$, avec la convention $A^0 = I_4$, $B^0 = C^0 = I_2$, et par conséquent $p(A) = \begin{pmatrix} p(B) & 0_2 \\ 0_2 & p(C) \end{pmatrix}$ pour $p \in \mathbf{R}[x]$. Donc $\mathcal{I}_A = \mathcal{I}_B \cap \mathcal{I}_C$, ce qui montre que q_A est égal au p.p.c.m de q_B et q_C . Notons que q_B divise $(x-1)^2$, donc $q_B = (x-1)^a$, avec $1 \leq a \leq 2$, et que q_C divise $(x+1)^2$, donc $q_C = (x+1)^b$, avec $1 \leq b \leq 2$. Comme $(x-1)^a$ et $(x+1)^b$ sont premiers entre eux, leur p.p.c.m est égal à leur produit, et $q_A = q_B q_C$.

Si $B = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, on a évidemment $q_B = x-1$, et si $B = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ on a $q_B = (x-1)^2$, puisque $B - I_2 \neq 0$. De même si $C = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$, on a $q_C = x+1$, et si $C = \begin{pmatrix} -1 & 1 \\ 0 & -1 \end{pmatrix}$ on a $q_C = (x+1)^2$, puisque $C + I_2 \neq 0$. On obtient ainsi les exemples cherchés :

$$\text{Si } A = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}, \text{ alors } p_A = (x-1)^2(x+1)^2 \text{ et } q_A = (x-1)(x+1).$$

$$\text{Si } A = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}, \text{ alors } p_A = (x-1)^2(x+1)^2 \text{ et } q_A = (x-1)^2(x+1).$$

$$\text{Si } A = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & -1 \end{pmatrix}, \text{ alors } p_A = (x-1)^2(x+1)^2 \text{ et } q_A = (x-1)(x+1)^2.$$

$$\text{Si } A = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & -1 \end{pmatrix}, \text{ alors } p_A = (x-1)^2(x+1)^2.$$

5.3 Nouvelle caractérisation des matrices diagonalisables

On va maintenant caractériser les matrices diagonalisables s par une condition portant sur le polynôme minimal.

Théorème 5.3.1 *Soit $A \in \mathcal{M}_n(K)$. Alors A est diagonalisable si et seulement si le polynôme minimal q_A de A est un polynôme scindé dont toutes les racines sont simples, et dans ce cas $q_A = (x - \mu_1) \dots (x - \mu_k)$ où μ_1, \dots, μ_k désignent les valeurs propres distinctes de A .*

Démonstration : Supposons que A est diagonalisable. Il existe une matrice

$$\text{diagonale } D = \begin{pmatrix} \lambda_1 & 0 & \dots & \dots & \dots & 0 \\ 0 & \lambda_2 & 0 & \dots & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \lambda_j & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \dots & 0 & \lambda_n \end{pmatrix} \text{ et une matrice inversible } P \in$$

$\mathcal{M}_n(K)$ telles que $A = PDP^{-1}$ soit diagonale. Soient μ_1, \dots, μ_k les valeurs propres distinctes de A , et soit $q = (x - \mu_1) \dots (x - \mu_k)$.

$$\text{On a } q(D) = \begin{pmatrix} q(\lambda_1) & 0 & \dots & \dots & \dots & 0 \\ 0 & q(\lambda_2) & 0 & \dots & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & q(\lambda_j) & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \dots & 0 & q(\lambda_n) \end{pmatrix}. \text{ Pour } 1 \leq j \leq n,$$

λ_j est une valeur propre de A donc il existe $i \leq k$ tel que $\lambda_j = \mu_i$, et $q(\lambda_j) = 0$. On a donc $q(D) = 0$, et $q(A) = Pq(D)P^{-1} = 0$, ce qui montre que q_A divise q . Comme toute valeur propre de A est racine de q_A , $q_A = q = (x - \mu_1) \dots (x - \mu_k)$ est un polynôme scindé dont toutes les racines sont simples. Supposons maintenant que le polynôme minimal de A est scindé et que toutes ses racines sont simples. On a $q_A = (x - \mu_1) \dots (x - \mu_k)$, avec μ_1, \dots, μ_k distincts. Posons $p_i = \prod_{j \neq i} (x - \mu_j)$. Comme les polynômes p_1, \dots, p_k n'ont pas de racine commune, leur p.g.c.d. est égal à 1. Il résulte alors du théorème de Bezout

(Théorème 3.9 du *Cours d'algèbre*) qu'il existe des polynômes u_1, \dots, u_k tels que $\sum_{j=1}^k p_j u_j = 1$. On a alors

$$I_n = \sum_{j=1}^k p_j(A) u_j(A).$$

Soit $X \in K^n$, et posons $X_j = p_j(A) u_j(A) X$ pour $1 \leq j \leq n$. On a $(A - \lambda_j I) p_j(A) = q_A(A) = 0$. Par conséquent $(A - \lambda_j I) X_j = q_A(A) u_j(A) X = 0$, et X_j appartient au sous-espace propre E_{μ_j} associé à μ_j pour $1 \leq j \leq n$. Comme $X = \sum_{i=1}^k X_j$, ceci montre que $K^n = E_{\mu_1} + \dots + E_{\mu_k}$. Soit \mathcal{B}_j une base de E_{μ_j} , et soit $\mathcal{B} = \cup_{1 \leq j \leq n} \mathcal{B}_j$. Alors \mathcal{B} est une partie génératrice de K^n . Comme \mathcal{B} est libre d'après la Proposition 4.4, on voit que \mathcal{B} est une base de K^n formée de vecteurs propres de A , ce qui montre que A est diagonalisable. ♣

Supposons que A est diagonalisable. Avec les notations ci-dessus on a $p_j(A) u_j(A) = v_{i,j}(A)(A - \mu_i I_n)$, où $v_{i,j} \in K[x]$ pour $i \neq j$. Donc si $Y \in E_{\mu_i}$ on a $p_j(A) u_j(A) Y = 0$ pour $j \neq i$ et $X = \sum_{j=1}^k p_j(A) u_j(A) Y = p_i(A) u_i(A) Y$. Soit $X \in K^n$. En appliquant ce résultat à $Y = p_i(A) u_i(A) X \in E_{\mu_i}$, on voit que $[p_i(A) u_i(A)]^2 X = p_i(A) u_i(A) X$, et par conséquent $[p_i(A) u_i(A)]^2 = p_i(A) u_i(A)$ pour $1 \leq i \leq n$.

Soit $U_i : X \mapsto p_i(A) X$ l'endomorphisme de K^n associé à $p_i(A)$ et soit $\Delta_i : X \mapsto p_i(A) u_i(A) X$ l'endomorphisme de K^n associé à $p_i(A) u_i(A)$. Il résulte de ce qui précède que $\text{Im}(\Delta_i) = E_{\mu_i}$. Si $X \in E_{\mu_j}$ avec $j \neq i$ on a $\Delta_i(X) = v_{j,i}(A)(A - \mu_j I_n) X = 0$, donc $\sum_{j \neq i} E_{\mu_j} \subset \text{Ker}(\Delta_i)$. Réciproquement si $X \in \text{Ker}(\Delta_i)$ on a $X = \sum_{j \neq i} p_j(A) u_j(A) X \in \sum_{j \neq i} E_{\mu_j}$, et on voit que $\text{Ker}(\Delta_i) = \sum_{j \neq i} E_{\mu_j}$. Il est clair que $\text{Im}(U_i) \subset \text{Im}(\Delta_i) = E_{\mu_i}$. Réciproquement si $X \in \text{Im}(U_i)$ on a $X = p_j(A) Y$, avec $Y \in K^n$, et $(A - \lambda_i) X = (A - \lambda_i) p_i(A) Y = q_A(A) Y = 0$, donc $X \in E_{\mu_i}$, ce qui montre que $E_{\mu_i} = \text{Im}(U_i)$ coïncide avec le sous-espace vectoriel de K^n engendré par les colonnes de A . On a donc le résultat suivant.

Théorème 5.3.2 Soit $A \in \mathcal{M}_n(K)$ une matrice diagonalisable, et soient μ_1, \dots, μ_k les valeurs propres distinctes de A . On pose $p_i = \prod_{\substack{1 \leq j \leq k \\ i \neq j}} (x - \mu_j)$ pour $1 \leq i \leq k$.

(i) Pour $1 \leq i \leq k$, le sous-espace propre associé à μ_i est le sous-espace vectoriel de K^n engendré par les colonnes de $p_i(A)$.

(ii) Si u_1, \dots, u_k est une famille de polynômes à coefficients dans K telle que $\sum_{1 \leq i \leq k} u_i p_i = 1$, alors l'application $\Delta_i : X \mapsto p_i(A) u_i(A) X$ est la projection

de K^n sur E_{μ_i} parallèlement à $\sum_{\substack{1 \leq j \leq k \\ j \neq i}} E_{\mu_j}$.

Ce résultat permet de simplifier les calculs permettant de trouver une base de K^n formée de vecteurs propres de A . Par exemple si $A = \begin{pmatrix} 1/2 & 1/2 & 1/2 \\ 1/4 & 1/2 & 0 \\ 1/4 & 0 & 1/2 \end{pmatrix}$ on

a vu au Chapitre précédent que $p_A = -x(x - \frac{1}{2})(x - 1)$, donc A est diagonalisable

$$\cdot \text{ On a } A - \frac{1}{2}I_3 = \begin{pmatrix} 0 & 1/2 & 1/2 \\ 1/4 & 0 & 0 \\ 1/4 & 0 & 0 \end{pmatrix}, A - I_3 = \begin{pmatrix} -1/2 & 1/2 & 1/2 \\ 1/4 & -1/2 & 0 \\ 1/4 & 0 & -1/2 \end{pmatrix}.$$

$$\text{ On obtient } A(A - I_3) = \begin{pmatrix} 1/2 & 1/2 & 1/2 \\ 1/4 & 1/2 & 0 \\ 1/4 & 0 & 1/2 \end{pmatrix} \begin{pmatrix} -1/2 & 1/2 & 1/2 \\ 1/4 & -1/2 & 0 \\ 1/4 & 0 & -1/2 \end{pmatrix} =$$

$$\begin{pmatrix} 0 & 0 & 0 \\ 0 & -1/8 & 1/8 \\ 0 & 1/8 & 1/8 \end{pmatrix}. \text{ Donc le vecteur } \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix} \text{ engendre } E_{1/2}.$$

$$\text{ On a aussi } A(A - \frac{1}{2}I_3) = \begin{pmatrix} 1/2 & 1/2 & 1/2 \\ 1/4 & 1/2 & 0 \\ 1/4 & 0 & 1/2 \end{pmatrix} \begin{pmatrix} -1/2 & 1/2 & 1/2 \\ 1/4 & -1/2 & 0 \\ 1/4 & 0 & -1/2 \end{pmatrix} =$$

$$\begin{pmatrix} 1/4 & 1/4 & 1/4 \\ 1/8 & 1/8 & 1/8 \\ 1/8 & 1/8 & 1/8 \end{pmatrix}. \text{ Donc le vecteur } \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix} \text{ engendre } E_1.$$

$$\text{ Enfin } (A - \frac{1}{2}I_3)(A - I_3) = \begin{pmatrix} 0 & 1/2 & 1/2 \\ 1/4 & 0 & 0 \\ 1/4 & 0 & 0 \end{pmatrix} \begin{pmatrix} -1/2 & 1/2 & 1/2 \\ 1/4 & -1/2 & 0 \\ 1/4 & 0 & -1/2 \end{pmatrix} =$$

$$\begin{pmatrix} 1/4 & -1/4 & -1/4 \\ -1/8 & 1/8 & 1/8 \\ -1/8 & 1/8 & 1/8 \end{pmatrix}. \text{ Donc le vecteur } \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix} \text{ engendre } E_0, \text{ et on a obtenu}$$

une base de K^n formée de vecteurs propres de A .

Soit E est un espace vectoriel non nul de dimension finie sur un corps K et soit $u \in \mathcal{L}(E)$. Pour $p = \alpha_0 + \dots + \alpha_n x^n \in K[x]$ on pose $p(u) = \alpha_0 \mathcal{I}_E + \dots + \alpha_n u^n$, où la puissance n^e u^n est calculée au sens de la composition des applications. Il est clair que l'application $p \mapsto p(u)$ est un homomorphisme d'anneaux de $K[x]$ dans $\mathcal{L}(E)$. On définit alors le **polynôme minimal** q_u de u comme étant l'unique polynôme unitaire q tel que $\mathcal{I}_u = qK[x]$, où $\mathcal{I}_u = \{p \in K[x] \mid p(u) = 0\}$.

Soit \mathcal{B} une base de E . Il résulte de la Proposition 3.9 que le polynôme minimal de u coïncide avec celui de la matrice $\mathcal{M}_{u, \mathcal{B}}$ représentant u dans la base \mathcal{B} . On déduit alors de la Proposition 4.2.4 et du Théorème 5.5 le résultat suivant.

Théorème 5.3.3 *Soit $E \neq \{0\}$ un espace vectoriel de dimension finie, et soit $u \in \mathcal{L}(E)$. Alors u est diagonalisable si et seulement si le polynôme minimal de u est un polynôme scindé dont toutes les racines sont simples.*

On dit qu'un sous-espace vectoriel F d'un espace vectoriel E est **invariant** pour $u \in \mathcal{L}(E)$ si $u(\mathbf{x}) \in F$ pour tout $\mathbf{x} \in F$. Dans ce cas on appelle **restriction de u à F** l'endomorphisme v de F défini par la formule $v(\mathbf{x}) = u(\mathbf{x})$ pour $\mathbf{x} \in F$. On a alors le corollaire suivant.

Corollaire 5.3.4 *Soit $E \neq \{0\}$ un espace vectoriel de dimension finie, soit $u \in \mathcal{L}(E)$, et soit $F \neq \{0\}$ un sous-espace vectoriel de E invariant pour u . Si u est diagonalisable, alors la restriction de u à F est diagonalisable.*

Démonstration : Soit v la restriction de u à F . Pour $\mathbf{x} \in F$ on a $q_u(v)(\mathbf{x}) = q_u(u)(\mathbf{x}) = 0$. Donc $q_u(v) = 0$, et q_v divise q_u . Par conséquent q_v est un polynôme scindé dont toutes les racines sont simples, et v est diagonalisable. ♣

Corollaire 5.3.5 Soient $A \in \mathcal{M}_n(K)$ et $B \in \mathcal{M}_n(K)$ deux matrices diagonalisables. Si $AB = BA$, alors il existe une base de K^n formée de vecteurs propres communs à A et B . En particulier $A - B$ est diagonalisable.

Démonstration : Soit μ une valeur propre de A , soit E_μ le sous-espace propre correspondant, et soit $v : X \mapsto BX$ l'endomorphisme de K^n associé à B . Pour $X \in E_\mu$, on a $Av(X) = ABX = BAX = v(\mu X) = \mu v(X)$. Donc $v(X) \in E_\mu$, et E_μ est invariant pour V . Comme B est diagonalisable, la restriction de V à E_μ est diagonalisable, ce qui signifie qu'il existe une base de E_μ formée de vecteurs propres de B . Soient maintenant μ_1, \dots, μ_k les valeurs propres distinctes de A . D'après ce qui précède on peut trouver pour $i \leq k$ une base \mathcal{B}_i de E_{μ_i} formée de vecteurs propres de B . Posons $\mathcal{B} = \cup_{1 \leq i \leq k} \mathcal{B}_i$. Puisque A est diagonalisable, \mathcal{B} est une base de K^n , et tous les éléments de \mathcal{B} sont des vecteurs propres communs à A et B , c'est à dire aussi des vecteurs propres de $A - B$. En particulier $A - B$ est diagonalisable.

5.4 Matrices nilpotentes

On va maintenant s'intéresser à une classe particulière de matrices.

Définition 5.4.1 Soit $A \in \mathcal{M}_n(K)$. On dit que A est nilpotente s'il existe $p \geq 1$ telle que $A^p = 0$.

Par exemple la matrice $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ est une matrice nilpotente non nulle, car $A^2 = 0$. Une matrice nilpotente non nulle **n'est pas diagonalisable**. En effet si A est nilpotente et diagonalisable, il existe $p \geq 1$ tel que $A^p = 0$, et il existe P inversible telle que $D = P^{-1}AP$ est diagonale. On a alors $D^p = P^{-1}A^pP = 0$. Comme D est diagonale, on a $D = 0$, et $A = PDP^{-1} = 0$.

Proposition 5.4.2 Soient $A \in \mathcal{M}_n(K)$ et $B \in \mathcal{M}_n(K)$ deux matrices nilpotentes. Si $AB = BA$, alors $A - B$ est nilpotente.

Démonstration : Soient $p \geq 1$ et $q \geq 1$ deux entiers tels que $A^p = B^q = 0$. Comme $AB = BA$, on peut calculer $(A - B)^{p+q}$ en utilisant le binôme de Newton (voir la formule (1.7) du *Cours d'Algèbre*) et on obtient

$$(A - B)^{p+q} = \sum_{k=0}^{p+q} (-1)^{p+q-k} C_{p+q}^k A^k B^{p+q-k}.$$

Si $p \leq k \leq p+q$, on a $A^k = 0$. Si $0 \leq k \leq p$, on a $p+q-k \geq q$, et $B^{p+q-k} = 0$. Donc $(A - B)^{p+q} = 0$. ♣

Si A est nilpotente, il est clair que toutes les valeurs propres de A sont nulles. En particulier si $A \in \mathcal{M}_n(\mathbf{C})$ est nilpotente, alors $p_A = x^n$. Ce résultat est en fait vrai pour un corps quelconque.

Proposition 5.4.3 *Soit $A \in \mathcal{M}_n(K)$ une matrice nilpotente. Alors $p_A = x^n$. En particulier $A^n = 0$.*

Démonstration : Le résultat est trivial si $n = 1$. Supposons que ce résultat est vrai pour $n \geq 1$, et soit $A \in \mathcal{M}_{n+1}(K)$ une matrice nilpotente. Si $A = 0$, $p_A = (-1)^{n+1}x^{n+1}$. Si $A \neq 0$, soit $p \geq 2$ le plus petit entier positif tel que $A^p = 0$, et soit $X_1 \in K^{n+1}$ tel que $A^{p-1}X_1 \neq 0$. Posons $Y_1 = A^{p-1}X_1$, et soit (Y_2, \dots, Y_{n+1}) une famille d'éléments de K^n telle que $\mathcal{B} = (Y_1, \dots, Y_{n+1})$ soit une base de K^{n+1} . Soit $u : X \mapsto AX$ l'endomorphisme de K^{n+1} associé à A , et soit $B = \mathcal{M}_{u, \mathcal{B}}$. Comme $A = \mathcal{M}_{u, \mathcal{B}_0}$, \mathcal{B}_0 désignant la base canonique de K^{n+1} , B est semblable à A et $p_A = p_B$. Comme $u(Y_1) = 0$, la première colonne de B est nulle, et on peut écrire

$$B = \begin{pmatrix} 0 & L_1 \\ 0_n & C \end{pmatrix},$$

avec $L_1 \in \mathcal{M}_{1,n}(K)$ et $C \in \mathcal{M}_n(K)$, 0_n désignant l'élément nul de $\mathcal{M}_{n,1}(K)$.

Une récurrence immédiate montre que pour $p \geq 1$ la matrice B^p est de la forme

$$B^p = \begin{pmatrix} 0 & L_p \\ 0_n & C^p \end{pmatrix},$$

avec $L_p \in \mathcal{M}_{1,n}(K)$. Donc C est nilpotente, et $p_C = x^n$. En développant par rapport à la première colonne on obtient $p_B = xp_C = x^{n+1}$, et on voit par récurrence sur n que $p_A = x^n$ si $A \in \mathcal{M}_n(K)$ est nilpotente. Le fait que dans ce cas $A^n = 0$ résulte alors du théorème de Cayley-Hamilton. ♣

Corollaire 5.4.4 *Soit $D \in \mathcal{M}_n(K)$ une matrice diagonalisable, et soit $N \in \mathcal{M}_n(K)$ une matrice nilpotente telle que $DN = ND$. Alors D et $D + N$ ont même polynôme caractéristique.*

Démonstration : On va procéder par récurrence sur le nombre de valeurs propres de D . Si D possède une seule valeur propre μ , il existe $P \in \mathcal{M}_n(K)$ inversible telle que $D = P(\mu I_n)P^{-1}$. Donc $D = \mu I_n$ et $p_{D+N} = \det((x - \mu)I_n - N) = (x - \mu)^n = p_{\mu I_n} = p_D$. Supposons maintenant que le résultat est vrai quand D possède k valeurs propres distinctes, soit $D \in \mathcal{M}_n(K)$ une matrice diagonalisable possédant $k + 1$ valeurs propres distinctes μ_1, \dots, μ_{k+1} de multiplicités n_1, \dots, n_k . Soit $u : X \mapsto DX$ l'endomorphisme de K^n associé à D et soit $v : X \mapsto NX$ l'endomorphisme de K^n associé à N . Notons que si A est de la forme $A = \begin{pmatrix} B & 0_{m, n-m} \\ 0_{n-m, m} & C \end{pmatrix}$, avec $B \in \mathcal{M}_m(K)$, $C \in \mathcal{M}_{n-m}(K)$, $0_{p,q}$ désignant la matrice à p lignes et q colonnes dont tous les coefficients sont

nuls, on a $\det(A) = \det(B)\det(C)$. Ceci montre en particulier que dans ce cas on a $p_A = p_B p_C$.

Soit \mathcal{B}_i une base du sous-espace propre E_{μ_i} associé à μ_i et posons $\mathcal{B} = \cup_{1 \leq j \leq k+1} \mathcal{B}_j$, $\mathcal{C} = \cup_{2 \leq j \leq k+1} \mathcal{B}_j$. La matrice Δ représentant u dans la base \mathcal{B} est une matrice diagonale de la forme $\Delta = \begin{pmatrix} \Delta_1 & 0_{m_1, n-m_1} \\ 0_{n-m_1, m_1} & \Delta_2 \end{pmatrix}$, avec $\Delta_1 = \mu_1 I_{m_1}$, $\Delta_2 \in \mathcal{M}_{n-m_1}$ diagonale, et on a $p_{\Delta_1} = (x - \mu_1)^{m_1}$ et $p_{\Delta_2} = (x - \mu_2)^{m_2} \dots (x - \mu_{k+1})^{m_{k+1}}$. Si $X \in E_{\mu_j}$ on a de même que plus haut $DNX = NDX = \mu_j NX$, donc $NX \in E_{\mu_j}$. Donc la matrice R représentant v dans la base \mathcal{B} est de la forme $R = \begin{pmatrix} R_1 & 0_{m_1, n-m_1} \\ 0_{n-m_1, m_1} & R_2 \end{pmatrix}$, avec $R_1 \in \mathcal{M}_{m_1}(K)$ et $R_2 \in \mathcal{M}_{n-m_1}(K)$. Comme N_1 est la matrice représentant la restriction de v à E_{μ_1} , N_1 est nilpotente, et N_1 commute évidemment avec Δ_1 . D'autre part comme N_2 est la matrice représentant la restriction de v à E_{μ_2} , N_2 est nilpotente, et comme Δ_2 est la matrice représentant la restriction de u à Δ_2 , N_2 commute avec Δ_2 .

On sait que $p_{\Delta_1} = p_{\Delta_1 + N_1}$, et d'après l'hypothèse de récurrence on a $p_{\Delta_2} = p_{\Delta_2 + N_2}$. Donc $p_{D+N} = p_{\Delta+R} = p_{\Delta_1 + N_1} p_{\Delta_2 + N_2} = (x - \mu_1)^{n_1} \dots (x - \mu_k)^{n_k} = p_D$, et le résultat est démontré par récurrence. ♣

5.5 Théorème de décomposition de Jordan

Nous sommes maintenant en mesure de démontrer le théorème de décomposition de Jordan.

Théorème 5.5.1 (Théorème de décomposition de Jordan)

Soit $A \in \mathcal{M}_n(K)$ une matrice dont le polynôme caractéristique est scindé. Alors il existe un unique couple (D, N) d'éléments de $\mathcal{M}_n(K)$ vérifiant les propriétés suivantes

- (1) $A = D + N$.
- (2) D est diagonalisable.
- (3) N est nilpotente.
- (4) $DN = ND$.

De plus A et D ont même polynôme caractéristique, et si μ_1, \dots, μ_k désignent les valeurs propres distinctes de A , de sorte que $p_A = (x - \mu_1)^{n_1} \dots (x - \mu_k)^{n_k}$, avec $n_1 \geq 1, \dots, n_k \geq 1$, on a $D = p(A)$, où $p \in K[x]$ est une solution quelconque du système d'équations de congruence

$$\left\{ \begin{array}{l} p \equiv \mu_1 \pmod{(x - \mu_1)^{n_1}} \\ \dots\dots\dots \\ p \equiv \mu_j \pmod{(x - \mu_j)^{n_j}} \\ \dots\dots\dots \\ p \equiv \mu_k \pmod{(x - \mu_k)^{n_k}} \end{array} \right.$$

Démonstration : Comme μ_1, \dots, μ_k sont distincts, $(x - \mu_i)^{n_i}$ et $(x - \mu_j)^{n_j}$ sont premiers entre eux pour $i \neq j$. Il résulte alors du théorème chinois (Théorème 3.17 du *Cours d'Algèbre*) que le système d'équations de congruence ci-dessus admet bien une solution $p \in K[x]$.

Pour $1 \leq j \leq k$, on a $p - \mu_j = (x - \mu_j)^{n_j} u_j$, avec $u_j \in K[x]$. Donc $(p - \mu_1) \dots (p - \mu_k) = p_A u$, avec $u = u_1 \dots u_k \in K[x]$. Posons $q = (x - \mu_1) \dots (x - \mu_k)$, et posons $D = p(A)$. On a $q(D) = (p(A) - \mu_1 I_n) \dots (p(A) - \mu_k I_n) = [(p - \mu_1) \dots (p - \mu_k)](A) = p_A(A)u(A) = 0$. Donc le polynôme minimal q_D de D divise q , et q_D est un polynôme scindé dont toutes les racines sont simples. Il résulte alors du théorème 5.1 que D est diagonalisable .

Posons $N = A - p(A)$. On a $ND = DN, A = D + N$, et si on pose $v = x - p$ on a $N = v(A)$. Pour $1 \leq j \leq k$ on a $v = x - p = x - \mu_j + (x - \mu_j)^{n_j} u_j = (x - \mu_j)(1 + (x - \mu_j)^{n_j - 1} u_j) = (x - \mu_j)w_j$, avec $w_j = (1 + (x - \mu_j)^{n_j - 1} u_j) \in K[x]$.

Comme $n = n_1 + \dots + n_k$, on a $q^n = (x - \mu_1)^{n_1} \dots (x - \mu_k)^{n_k} w_1^{n_1} \dots w_k^{n_k} = p_A w$, avec $w = w_1^{n_1} \dots w_k^{n_k} \in K[x]$. On a donc $N^n = q^n(A) = p_A(A)w(A) = 0$, et N est nilpotente, ce qui prouve l'existence de la décomposition cherchée.

Soit (D_1, N_1) un autre couple de matrices vérifiant les quatre conditions du théorème. On a $D_1 A = D_1^2 + D_1 N_1 = D_1^2 + N_1 D_1 = A D_1$ et $N_1 A = N_1 D_1 + D_1^2 = D_1 N_1 + D_1^2 = A N_1$. Donc $D_1 q(A) = q(A) D_1$ et $N_1 q(A) = q(A) N_1$ pour tout polynôme $q \in K[x]$. En particulier D_1 commute avec $D = p(A)$ et N_1 commute avec $N = v(A)$.

On a $D_1 - D = N - N_1$. Il résulte alors du Corollaire 5.9 et de la Proposition 5.11 que $D - D_1$ est à la fois diagonalisable et nilpotente. Donc $D - D_1 = 0$, $D = D_1, N = N_1$ et la décomposition de Jordan est unique. Le fait que $p_A = p_D$ résulte du corollaire 5.4.4. ♣

Notons que le théorème ci-dessus donne un **moyen effectif** de calculer la décomposition de Jordan d'une matrice quand on a pu calculer effectivement ses valeurs propres. La situation est particulièrement simple dans $\mathcal{M}_2(\mathbf{C})$. En effet dans ce cas ou bien A possède deux valeurs propres distinctes, ce qui implique que A est diagonalisable, et la décomposition de Jordan est alors la décomposition triviale $D = A, N = 0$, ou bien A possède une valeur propre double λ_1 . Dans ce cas $D = p(A)$, où $p \in \mathbf{C}[X]$ est une solution quelconque de l'équation de congruence

$$p \equiv \lambda_1 \pmod{(x - \lambda_1)^2} \tag{5.1}$$

on peut donc prendre $p = \lambda_1$, ce qui donne

$$D = \lambda_1 I_2 = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_1 \end{bmatrix}, N = A - \lambda_1 I_2 = A - \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_1 \end{bmatrix} \quad (5.2)$$

Exemple 5.5.2 Décomposition de Jordan de $A = \begin{bmatrix} 3 & 1 \\ -1 & 1 \end{bmatrix}$.

On a $\det(A) = 4$, $\text{Tr}(A) = 4$, donc $p_A = x^2 - 4x + 4 = (x - 2)^2$. Donc A possède 2 comme valeur propre double.

On a donc d'après la formule ci-dessus

$$D = 2I_2 = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}, N = A - 2I_2 = \begin{bmatrix} 3 & 1 \\ -1 & 1 \end{bmatrix} - \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ -1 & -1 \end{bmatrix}.$$

Pour les matrices 3×3 , on déduit du théorème 5-5-1 le résultat suivant.

Corollaire 5.5.3 Soit $A \in \mathcal{M}_3(\mathbf{C})$, et soit $p_A = (x - \lambda_1)(x - \lambda_2)(x - \lambda_3)$ le polynôme caractéristique de A , et soit $A = D + N$, avec D diagonalisable, N nilpotente, $ND = DN$ la décomposition de Jordan de A .

(i) Si les trois valeurs propres $\lambda_1, \lambda_2, \lambda_3$ sont distinctes, alors $D = A$ et $N = 0$.

(ii) Si les trois valeurs propres $\lambda_1, \lambda_2, \lambda_3$ sont égales, alors $D = \lambda_1 I$, et $N = A - \lambda_1 I$.

(iii) Si $\lambda_3 = \lambda_2 \neq \lambda_1$, alors $D = \lambda_2 I + \frac{1}{\lambda_1 - \lambda_2}(A - \lambda_2 I)^2$, et $N = A - D = \frac{1}{\lambda_2 - \lambda_1}(A - \lambda_1 I)(A - \lambda_2 I)$.

Démonstration : (i) Il est clair que si A admet trois valeurs propres distinctes, alors A est diagonalisable, ce qui donne $D = A$ et $N = A - A = 0$.

(ii) Si λ_1 est valeur propre de A de multiplicité 3, alors le polynôme constant $p = \lambda_1$ vérifie trivialement $p \equiv \lambda_1 \pmod{(x - \lambda_1)^3}$.

D'après le théorème 5-5-1, on a alors $D = \lambda_1 I$, $N = A - \lambda_1 I$.

(iii) Supposons maintenant que $\lambda_3 = \lambda_2 \neq \lambda_1$. Il résulte du théorème 5-5-1 que $N = p(A)$, où $p \in \mathbf{C}[x]$ est une solution quelconque du système d'équations de congruence

$$\begin{cases} p \equiv \lambda_1 & \pmod{(x - \lambda_1)} \\ p \equiv \lambda_2 & \pmod{(x - \lambda_2)^2} \end{cases}$$

Ce type d'équations est étudié dans le Chapitre 3 du *Cours d'algèbre* [?]. Il est plus simple ici de partir de la seconde équation, dont la solution générale est

$$p = \lambda_2 + q(x - \lambda_2)^2, \quad \text{avec } q \in \mathbf{C}[x].$$

La première équation est alors vérifiée si et seulement si $p - \lambda_1 = \lambda_2 - \lambda_1 + q(x - \lambda_2)^2$ est divisible par $x - \lambda_1$, ce qui équivaut au fait que ce polynôme admet λ_1 pour racine. On obtient

$$0 = p(\lambda_1) - \lambda_1 = \lambda_2 - \lambda_1 + q(\lambda_1)(\lambda_1 - \lambda_2)^2.$$

En simplifiant par $\lambda_1 - \lambda_2$, on obtient $q(\lambda_1) = \frac{1}{\lambda_1 - \lambda_2}$, et il suffit de prendre q constant, $q = \frac{1}{\lambda_1 - \lambda_2}$, ce qui donne $p = \lambda_2 + \frac{(x - \lambda_2)^2}{\lambda_1 - \lambda_2}$. On obtient

$$\begin{aligned} D = p(A) &= \lambda_2 I + \frac{1}{\lambda_1 - \lambda_2} (A - \lambda_2 I)^2, \quad N = A - D = A - \lambda_2 I - \frac{1}{\lambda_1 - \lambda_2} (A - \lambda_2 I)^2 \\ &= \frac{1}{\lambda_2 - \lambda_1} (A - \lambda_1 I)(A - \lambda_2 I). \end{aligned}$$

♣

Exemple 5.5.4 *Décomposition de Jordan de* $A = \begin{bmatrix} 0 & 2 & -1 \\ -1 & 4 & -2 \\ -1 & 4 & -2 \end{bmatrix}$.

On voit tout de suite que $\det(A) = 0$, puisque les deux dernières lignes de A sont égales. On obtient, en retirant la seconde ligne à la troisième dans le calcul du déterminant ci-dessous, puis en ajoutant la troisième colonne à la seconde

$$\begin{aligned} p_A &= - \begin{vmatrix} -x & 2 & -1 \\ -1 & 4-x & -2 \\ -1 & 4 & -2-x \end{vmatrix} = - \begin{vmatrix} -x & 2 & -1 \\ -1 & 4-x & -2 \\ 0 & x & -x \end{vmatrix} \\ &= - \begin{vmatrix} -x & 1 & -1 \\ -1 & 2-x & -2 \\ 0 & 0 & -x \end{vmatrix} = x(x(x-2) + 1) = x(x-1)^2. \end{aligned}$$

On voit donc que A admet 0 pour valeur propre simple et 1 pour valeur propre double. D'après le corollaire précédent, on a

$$\begin{aligned} D &= I - (A - I)^2, \quad N = A(A - I). \quad \text{Ici } A - I = \begin{bmatrix} -1 & 2 & -1 \\ -1 & 3 & -2 \\ -1 & 4 & -3 \end{bmatrix}, \quad (A - I)^2 = \\ \begin{bmatrix} -1 & 2 & -1 \\ -1 & 3 & -2 \\ -1 & 4 & -3 \end{bmatrix} \begin{bmatrix} -1 & 2 & -1 \\ -1 & 3 & -2 \\ -1 & 4 & -3 \end{bmatrix} &= \begin{bmatrix} 0 & 0 & 0 \\ 0 & -1 & 1 \\ 0 & -2 & 2 \end{bmatrix}. \\ \text{On obtient} \end{aligned}$$

$$\begin{aligned} D &= I - \begin{bmatrix} 0 & 0 & 0 \\ 0 & -1 & 1 \\ 0 & -2 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & -1 \\ 0 & 2 & -1 \end{bmatrix}, \\ N &= A - D = \begin{bmatrix} 0 & 2 & -1 \\ -1 & 4 & -2 \\ -1 & 4 & -2 \end{bmatrix} - \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & -1 \\ 0 & 2 & -1 \end{bmatrix} = \begin{bmatrix} -1 & 2 & -1 \\ -1 & 2 & -1 \\ -1 & 2 & -1 \end{bmatrix}. \end{aligned}$$

5.6 Un algorithme de calcul rapide pour la décomposition de Jordan

Soit K un corps, et soit 1 l'élément unité de K . On définit $n.1$ pour n entier, $n \geq 1$, par la formule $n.1 := 1 + \dots + 1$ (n fois) et on dira que le corps K est de **caractéristique nulle** quand $n.1 \neq 0$ pour tout $n \geq 1$. Si K n'est pas de caractéristique nulle on appelle caractéristique de K le plus petit entier $n \geq 1$ tel que $n.1 = 0$.

Définition 5.6.1 Soit K un corps, et soit $p \in K[x]$ un polynôme non nul. On pose

$$\tilde{p} = \frac{p}{p.g.c.d.(p,p')}. \quad (5.3)$$

Nous commençons par l'observation suivante.

Proposition 5.6.2 Soit K un corps de caractéristique nulle, soit $p = (x - \lambda_1)^{n_1} \dots (x - \lambda_k)^{n_k}$ un polynôme scindé sur K , avec $n_j \geq 1$ pour $1 \leq j \leq k$ et $\lambda_i \neq \lambda_j$ pour $i \neq j$. Alors $\tilde{p} = (x - \lambda_1) \dots (x - \lambda_k)$.

Démonstration : On a $p' = \sum_{j=0}^k n_j (x - \lambda_j)^{n_j-1} \prod_{i \neq j} (x - \lambda_i)^{n_i}$. Par conséquent $(x - \lambda_j)^{n_j-1}$ divise p' . Comme $(x - \lambda_i)^{n_i-1}$ et $(x - \lambda_j)^{n_j-1}$ sont premiers entre eux deux à deux pour $i \neq j$, p' est divisible par le produit $(x - \lambda_1)^{n_1-1} \dots (x - \lambda_k)^{n_k-1}$, et il existe $u \in K[x]$ tel que l'on ait

$$p' = (x - \lambda_1)^{n_1-1} \dots (x - \lambda_k)^{n_k-1} u.$$

Comme $n_j.1 \neq 0$, on a $u(\lambda_j) \neq 0$ et u est premier avec $(x - \lambda_j)$ pour $1 \leq j \leq k$. Donc u est premier avec $q := (x - \lambda_1) \dots (x - \lambda_k)$. On a alors

$$p = (x - \lambda_1)^{n_1-1} \dots (x - \lambda_k)^{n_k-1} q,$$

avec q premier avec u , ce qui montre que $p.g.c.d.(p,p') = (x - \lambda_1)^{n_1-1} \dots (x - \lambda_k)^{n_k-1}$. Donc $\tilde{p} = \frac{p}{p.g.c.d.(p,p')} = (x - \lambda_1) \dots (x - \lambda_k)$. ♣

Ce résultat a une première conséquence importante : dans la situation ci-dessus, le calcul de $(x - \lambda_1) \dots (x - \lambda_k)$ ne nécessite pas de connaître les valeurs propres, il suffit d'appliquer l'algorithme d'Euclide pour obtenir le p.g.c.d. $q(p)$ de p et p' et de diviser p par $q(p)$.

On admettra que pour tout corps K il existe un corps \tilde{K} contenant K et possédant les deux propriétés suivantes

- (i) Tout élément de \tilde{K} est racine d'un polynôme à coefficients dans K .
- (ii) Tout polynôme à coefficients dans \tilde{K} est scindé sur \tilde{K} .

On notera que $\tilde{\mathbf{R}} = \mathbf{C}$, d'après le théorème de d'Alembert.

On a le corollaire suivant.

Corollaire 5.6.3 *Soit K un corps de caractéristique nulle, et soit $p \in K[x]$ un polynôme non constant. Alors p et \tilde{p}' sont premiers entre eux, et il existe un polynôme $u_p \in K[x]$ tel que $\tilde{p}'u_p \equiv 1 \pmod{p}$.*

Démonstration : Il résulte de la proposition que $\tilde{p} = (x - \lambda_1)\dots(x - \lambda_k)$, $\lambda_1, \dots, \lambda_k$ désignant les racines distinctes de p dans \tilde{K} . Comme \tilde{p} est scindé à racines simples sur \tilde{K} , on a $\tilde{p}'(\lambda_j) \neq 0$ pour $1 \leq j \leq k$. Comme tout polynôme non constant à coefficients dans K est scindé sur \tilde{K} , ceci montre que p et \tilde{p}' sont premiers entre eux, et il résulte du théorème de Bezout qu'il existe deux polynômes u_p et v_p à coefficients dans K tels que $\tilde{p}'u_p + pv_p = 1$. ♣

On va maintenant donner un algorithme efficace pour obtenir la décomposition de Jordan dans \tilde{K} d'une matrice $A \in \mathcal{M}_n(K)$ quand K est de caractéristique nulle. On notera que cette méthode est directement inspirée de la méthode de Newton, qui fournit un schéma d'approximation d'une racine de l'équation $f(x) = 0$. Cette méthode consiste à choisir convenablement x_0 et à poser, pour $m \geq 0$,

$$x_{m+1} = x_m - \frac{f(x_m)}{f'(x_m)}.$$

Certaines hypothèses, que nous ne détaillerons pas ici, permettent de garantir que la suite $(x_m)_{m \geq 0}$ converge vers une solution x de l'équation $f(x) = 0$.

L'idée qui sous-tend l'algorithme ci-dessous est de trouver une solution non triviale D de l'équation matricielle $\tilde{p}_A(D) = 0$, qui se trouve être la partie diagonalisable (sur \tilde{K}) d'une matrice donnée $A \in \mathcal{M}_n(K)$. Il n'y a aucun problème de convergence ici, car la suite obtenue est **stationnaire** : on a $D_m = D$ pour m assez grand.

Cette utilisation la méthode de Newton est due au grand mathématicien contemporain Alexandre Grothendieck, qui l'a introduite dans les années 60, et elle joue un rôle important dans la théorie des *groupes algébriques*. Le choix pédagogique consistant à considérer la décomposition de Jordan d'une matrice carrée comme le résultat central de l'algèbre linéaire classique été suggéré au second auteur par les travaux de son élève Danielle Couty [2], où la décomposition de Jordan joue un rôle essentiel. Le caractère effectif de la procédure présentée ci-dessous conforte l'idée que la décomposition de Jordan est bien un résultat central. On peut toujours la calculer, alors que le calcul de la forme de Jordan (de même que la trigonalisation) d'une matrice carrée nécessite la connaissance des ses valeurs propres, ce qui mène à un calcul impossible à réaliser sauf valeur propre évidente pour $A \in \mathcal{M}_n(K)$, $n \geq 5$. Notons que cet algorithme, s'il est bien connu (on le trouvera par exemple proposé en exercice p.62-63 de l'ouvrage récent de A.Boyer et J.Risler [1]), reste curieusement très peu enseigné.

Nous utiliserons le résultat suivant, qui est une conséquence immédiate de la formule du binôme de Newton (voir le Chapitre 1 du *Cours d'algèbre* [?]).

Lemme 5.6.4 *Soit K un corps, soit $q \in K[x]$, et soit \mathcal{A} un anneau commutatif. On a alors, pour $U, V \in \mathcal{A}$,*

$$q(U + V) - q(U) \in V\mathcal{A}, \quad q(U + V) - q'(U)V \in V^2\mathcal{A}.$$

Démonstration : Pour $j \geq 1$, on a

$$\begin{aligned} (U + V)^j - U^j - jU^{j-1}V &= \sum_{i=0}^j C_j^i U^i V^{j-i} - U^j - jU^{j-1}V \\ &= \sum_{i=2}^j C_j^i U^{j-i} V^i \in V^2\mathcal{A}, \end{aligned}$$

ce qui implique par linéarité la seconde relation, et la première s'en déduit.

♣

Théorème 5.6.5 *Soit K un corps de caractéristique nulle, soit $A \in \mathcal{M}_n(K)$, soit $p = p_A$ le polynôme caractéristique de A , soit $n(p)$ un entier positif tel que p divise $\tilde{p}^{n(p)}$, et soit $u_p \in K[x]$ tel que $u_p \tilde{p}' \equiv 1 \pmod{p}$. On définit par récurrence une suite $(D_m)_{p \geq 0}$ d'éléments de $\mathcal{M}_n(K)$ en posant*

$$\begin{cases} D_0 = A, \\ D_{m+1} = D_m - \tilde{p}(D_m)u_p(D_m) \quad \text{pour } m \geq 0. \end{cases}$$

Alors $p_{D_m} = p, p(D_m) = 0$, et $\tilde{p}'(D_m)$ est inversible d'inverse égal à $u_p(D_m)$ pour $m \geq 0$. De plus $\tilde{p}(D_m) = 0$ et D_m coïncide avec la partie diagonalisable sur \tilde{K} de A pour $2^{2^m} \geq n(p)$.

Démonstration : Soit $\mathcal{A} := \{h(A)\}_{h \in K[x]}$, de sorte que \mathcal{A} est commutative. Il est clair que $D_m \in \mathcal{A}$ pour $m \geq 0$. Il résulte du théorème de Cayley-Hamilton que $p(A) = 0$. On a $\tilde{p}(D_{m+1}) = \tilde{p}(D_m - \tilde{p}(D_m)u_p(D_m)) = \tilde{p}(D_m) - \tilde{p}(D_m)U_m = \tilde{p}(D_m)(I + U_m)$, avec $U_m \in \mathcal{A}$. Une récurrence immédiate montre alors que $\tilde{p}(D_m) \in \tilde{p}(A)\mathcal{A}$, et $\tilde{p}(D_m)^{n(p)} = 0$ pour $m \geq 0$. Donc $\tilde{p}(D_m)u_p(D_m)$ est nilpotent et il résulte du corollaire 5.4.4 et de l'existence de la décomposition de Jordan de D_m sur \tilde{K} que D_m et D_{m+1} ont le même polynôme caractéristique. Donc $p_{D_m} = p_A = p, p(D_m) = 0$, et $u_p(D_m)\tilde{p}'(D_m) = \tilde{p}'(D_m)u_p(D_m) = I_n$, $\tilde{p}'(D_m)$ est inversible, et $[\tilde{p}'(D_m)]^{-1} = u_p(D_m)$. On voit donc que la suite $(D_m)_{m \geq 0}$ vérifie la relation

$$D_{m+1} = D_m - \tilde{p}(D_m)[\tilde{p}'(D_m)]^{-1},$$

ce qui est la formule de l'algorithme de Newton appliquée à \tilde{p} , considérée comme une fonction opérant sur $\mathcal{M}_n(K)$.

On a $D_0 = A$ et, pour $m \geq 1$,

$$A - D_m = \sum_{j=0}^{m-1} \tilde{p}(D_j)u(D_j),$$

ce qui prouve que $(A - D_m)^{n(p)} = 0$ pour $m \geq 0$. D'autre part, comme $\tilde{p}(U+V) - \tilde{p}(U) - \tilde{p}'(U)(U-V) \in (U-V)^2 \mathcal{A}$ pour $U, V \in \mathcal{A}$, il existe $W_m \in \mathcal{A}$ vérifiant, pour $m \geq 0$,

$$\begin{aligned} \tilde{p}(D_{m+1}) &= \tilde{p}(D_m - \tilde{p}(D_m)u(D_m)) \\ &= \tilde{p}(D_m) - \tilde{p}(D_m)u(D_m)\tilde{p}'(D_m) + \tilde{p}(D_m)^2u(D_m)^2W_m \\ &= \tilde{p}(D_m)^2u(D_m)^2W_m \in [\tilde{p}(D_m)]^2 \mathcal{A}. \end{aligned}$$

Une récurrence immédiate montre alors que l'on a, pour $m \geq 0$,

$$\tilde{p}(D_m) \in [\tilde{p}(A)]^{2^m} \mathcal{A}. \quad (5.4)$$

Donc $\tilde{p}(D_m) = 0$ si $2^m \geq n(p)$. Dans ce cas D_m est diagonalisable sur \tilde{K} , puisque \tilde{p} est scindé à racines simples sur \tilde{K} . Comme $A - D_m$ est nilpotent et commute avec D_m , on voit que $D_m = D$, où D est la partie diagonalisable sur \tilde{K} de A . ♣

Notons que la démonstration ci-dessus fournit une preuve directe de l'existence d'une décomposition de Jordan (avec la partie diagonalisable D diagonalisable sur \tilde{K}) pour $A \in \mathcal{M}_n(K)$ quand le corps K est de caractéristique nulle. On a au début de la démonstration utilisé l'existence de la décomposition de Jordan, mais cette référence à la décomposition de Jordan peut être évitée en remplaçant u_p par un polynôme u tel que $u\tilde{p}' - 1$ soit divisible par $\tilde{p}^{n(p)}$. D'autre part cet algorithme fonctionne sans modification pour un corps de caractéristique $\alpha > 0$ quand aucune des racines dans \tilde{K} du polynôme caractéristique p_A n'est de multiplicité de la forme $m\alpha$ avec $m \in \mathbf{N}$.

Notons également que le plus petit entier $n(p)$ tel que p divise $\tilde{p}^{n(p)}$ peut se calculer sans connaître les racines de p . Ce nombre est égal au plus grand ordre de multiplicité des racines de p dans \tilde{K} , voir l'exercice 10. On voit donc qu'une seule itération va suffire si toutes les valeurs propres de A dans \tilde{K} sont de multiplicité au plus 2, que deux itérations vont suffire si toutes les valeurs propres de A dans \tilde{K} sont de multiplicité au plus 4, et que trois itérations vont suffire si toutes les valeurs propres de A sont de multiplicité au plus 8. La formule 5.4 est donc une très bonne version algébrique des majorations d'erreurs obtenues en Analyse pour la méthode de Newton pour les fonctions dont on contrôle les deux premières dérivées.

Exemple 5.6.6 *Décomposition de Jordan de $A = \begin{bmatrix} 3 & 1 \\ -1 & 1 \end{bmatrix}$ par l'algorithme rapide.*

Posons $p = p_A = x^2 - 4x + 2$. On a $p' = 2x - 4$, et $p.g.c.d.(p, p') = x - 2$. Donc $\tilde{p} = \frac{p}{x-2} = x - 2$, et $\tilde{p}' = 1$. Il est clair que $n(p) = 2$, et on a $\tilde{p} \times 0 + \tilde{p}' = 1$. Pour appliquer l'algorithme on peut donc prendre $u = 1$, et on a

$$D_0 = A, \quad D_1 = A - \tilde{p}(A) = A - (A - 2I) = 2I.$$

La décomposition de Jordan de A est donc $A = D + N$, avec $D = 2I$ diagonalisable (en fait diagonale), et $N = A - 2I = \begin{bmatrix} 1 & 1 \\ -1 & -1 \end{bmatrix}$ nilpotente.

Exemple 5.6.7 Décomposition de Jordan de $A = \begin{bmatrix} 0 & 2 & -1 \\ -1 & 4 & -2 \\ -1 & 4 & -2 \end{bmatrix}$ par l'algorithme rapide.

Posons $p = p_A = x^3 - 2x^2 + x$. On a $p' = 3x^2 - 4x + 1$, donc $p.g.c.d.(p, p') = x - 1$, et $\tilde{p} = \frac{p}{x-1} = x^2 - x$. Donc $\tilde{p}' = 2x - 1$, et on vérifie bien que \tilde{p} et \tilde{p}' sont premiers entre eux. Comme $p'' = 6x - 4$ est premier avec p , on peut prendre $n(p) = 2$. La partie diagonalisable D de A est donc donnée par la formule

$$D = D_1 = A - \tilde{p}(A)[\tilde{p}'(A)^{-1}] = A - (A^2 - A)(2A - I)^{-1}.$$

On a

$$A - I = \begin{bmatrix} -1 & 2 & -1 \\ -1 & 3 & -2 \\ -1 & 4 & -3 \end{bmatrix},$$

$$A(A - I) = \begin{bmatrix} 0 & 2 & -1 \\ -1 & 4 & -2 \\ -1 & 4 & -2 \end{bmatrix} \begin{bmatrix} -1 & 2 & -1 \\ -1 & 3 & -2 \\ -1 & 4 & -3 \end{bmatrix} = \begin{bmatrix} -1 & 2 & -1 \\ -1 & 2 & -1 \\ -1 & 2 & -1 \end{bmatrix},$$

$$2A - I = \begin{bmatrix} -1 & 4 & -2 \\ -2 & 7 & -4 \\ -2 & 8 & -5 \end{bmatrix}, (2A - I)^{-1} = \begin{bmatrix} -3 & 4 & -2 \\ 2 & -1 & 0 \\ 2 & 0 & -1 \end{bmatrix},$$

$$\begin{aligned} A(A - I)(2A - I)^{-1} &= \begin{bmatrix} -1 & 2 & -1 \\ -1 & 2 & -1 \\ -1 & 2 & -1 \end{bmatrix} \begin{bmatrix} -3 & 4 & -2 \\ 2 & -1 & 0 \\ 2 & 0 & -1 \end{bmatrix} \\ &= \begin{bmatrix} -1 & 2 & -1 \\ -1 & 2 & -1 \\ -1 & 2 & -1 \end{bmatrix}, \end{aligned}$$

$$D = A - A(A - I)(2A - I)^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & -1 \\ 0 & 2 & -1 \end{bmatrix},$$

$$N = A - D = \begin{bmatrix} -1 & 2 & -1 \\ -1 & 2 & -1 \\ -1 & 2 & -1 \end{bmatrix}.$$

L'algorithme rapide est évidemment très performant pour les grosses matrices, et il présente le très gros intérêt de ne pas demander le calcul des valeurs propres. On peut le programmer sous Mupad (c'est proposé à l'exercice 12).

5.7 Chapitre 5 sous Mupad

La commande "linalg : minpoly(A,x)" permet d'obtenir le polynôme minimal, exprimé en fonction de l'indéterminée x , d'une matrice carrée A donnée.

Par exemple si $A = \begin{bmatrix} 0 & 2 & -1 \\ -1 & 4 & -2 \\ -1 & 4 & -2 \end{bmatrix}$ est la matrice étudiée ci-dessus, on obtient $q_A = x^3 - 2x^2 + x = x(x-1)^2$. Par contre Mupad ne donne pas spontanément la décomposition de Jordan, comme on va le voir ci-dessous.

Avec la commande "linalg : jordanForm(A,All)," Mupad donne deux matrices. La première est une "forme de Jordan" B de A c'est à dire une matrice semblable à A qui est somme d'une matrice diagonale D_B et d'une matrice N_B dont tous les termes sont nuls, sauf ceux situés au dessus de la diagonale qui sont égaux à 0 ou 1, et qui vérifient $D_B \cdot N_B = N_B \cdot D_B$. Il est clair que N_B est nilpotente, donc la décomposition de Jordan de B est la décomposition triviale $B = D_B + N_B$. La deuxième matrice est une matrice P telle que $A = PBP^{-1}$. Pour récupérer la décomposition de Jordan de A il suffit de calculer $D = PD_B P^{-1}$ et $N = PN_B P^{-1}$, que l'on a notés ci-dessus $D1$ et $N1$ car le symbole D est déjà réservé sous Mupad.

On va maintenant utiliser l'algorithme rapide pour calculer la décomposition de Jordan de la matrice

$$B := \begin{bmatrix} -1 & -2 & 5 & -2 & 0 & 1 \\ -5 & -2 & 10 & -4 & 0 & 2 \\ -6 & -7 & 17 & -6 & 0 & 3 \\ -11 & -19 & 36 & -13 & 0 & 8 \\ -15 & -23 & 47 & -21 & 3 & 10 \\ -17 & -25 & 52 & -23 & 1 & 13 \end{bmatrix}.$$

Il s'agit certes d'un exemple quelque peu académique, puisque cette matrice admet 3 et 4 comme racines doubles et 1 et 2 comme racines simples (on peut voir que les racines doubles sont 3 et 4 en observant dans le calcul ci-dessous que $\text{pgcd}(p_B, p'_B) = x^2 - 7x + 12$). Ceci dit il s'agit d'une matrice 6×6 pour laquelle le calcul des valeurs propres sous Mupad n'est pas possible. On voit que $n(p) = 2$, donc il suffit d'une seule itération pour obtenir la décomposition de Jordan. Le résultat est vérifié à la fin : la matrice D_1 obtenue est bien diagonalisable, puisque $\tilde{p}(D_1) = 0$, avec \tilde{p} à racines simples, elle commute avec B , et si on pose $N = B - D_1$ on a $N^2 = 0$.

On notera le calcul d'expressions du type $q(B)$, où p est un polynôme et B une matrice carrée, est effectué ci-dessous de manière très naïve : on calcule $q(B)$ en posant

$$q(B) = \text{coeff}(q, x, 0) * B^0 + \text{coeff}(q, x, 1) * B + \text{coeff}(q, x, 2) * B^2 \dots + \text{coeff}(q, x, n) * B^n,$$

ce qui impose d'écrire manuellement une somme de $n + 1$ termes, n désignant le degré de q . Une méthode plus performante sera proposée dans la prochaine

édition de ce cours. Nous proposons à l'exercice 11 un calcul de décomposition de Jordan pour une matrice non diagonalisable sans valeur propres calculables.

```
M:=Dom::Matrix();
A:=M([[0,2,-1],[-1,4,-2],[-1,4,-2]]);
q:=linalg::minpoly(A,x);
```

```
Dom::Matrix()
```

$$\begin{pmatrix} 0 & 2 & -1 \\ -1 & 4 & -2 \\ -1 & 4 & -2 \end{pmatrix}$$

$$x^3 - 2 \cdot x^2 + x$$

```
u:=linalg::jordanForm(A,All);
```

$$\left[\begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & -2 & 0 \\ 1 & -1 & 1 \\ 1 & -1 & 2 \end{pmatrix} \right]$$

```
P:=op(u,2);D1:=P*M([[1,0,0],[0,1,0],[0,0,0]])*(1/P);
N1:=P*M([[0,1,0],[0,0,0],[0,0,0]])*(1/P);
```

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

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

$$\begin{pmatrix} -1 & 2 & -1 \\ -1 & 2 & -1 \\ -1 & 2 & -1 \end{pmatrix}$$

```
M:=Dom::Matrix();
B:=M([[-1,-2,5,-2,0,1],[-5,-2,10,-4,0,2],
[-6,-7,17,-6,0,3],[-11,-19,36,-13,0,8],
[-15,-23,47,-21,3,10],[-17,-25,52,-23,1,13]]);
```

```
Dom::Matrix()
```

$$\begin{pmatrix} -1 & -2 & 5 & -2 & 0 & 1 \\ -5 & -2 & 10 & -4 & 0 & 2 \\ -6 & -7 & 17 & -6 & 0 & 3 \\ -11 & -19 & 36 & -13 & 0 & 8 \\ -15 & -23 & 47 & -21 & 3 & 10 \\ -17 & -25 & 52 & -23 & 1 & 13 \end{pmatrix}$$

```
linalg::eigenValues(B);
```

Fail

```
p:=linalg::charpoly(B,x);
```

$$x^6 - 17 \cdot x^5 + 117 \cdot x^4 - 415 \cdot x^3 + 794 \cdot x^2 - 768 \cdot x + 288$$

```
pprime:=diff(p,x);
```

$$6 \cdot x^5 - 85 \cdot x^4 + 468 \cdot x^3 - 1245 \cdot x^2 + 1588 \cdot x - 768$$

```
p1:=gcd(p,pprime);
tildep:=p/p1;tildepprime:=diff(tildep,x);
```

$$x^2 - 7 \cdot x + 12$$

$$x^4 - 10 \cdot x^3 + 35 \cdot x^2 - 50 \cdot x + 24$$

$$4 \cdot x^3 - 30 \cdot x^2 + 70 \cdot x - 50$$

```
tildep^2/p;
C:=coeff(tildep,x,0)*B^0+coeff(tildep,x,1)*B^1
+coeff(tildep,x,2)*B^2+coeff(tildep,x,3)*B^3
+coeff(tildep,x,4)*B^4;
```

$$x^2 - 3 \cdot x + 2$$

$$\begin{pmatrix} -26 & -11 & 52 & -23 & 4 & 4 \\ -64 & -27 & 128 & -57 & 10 & 10 \\ -85 & -36 & 170 & -75 & 13 & 13 \\ -224 & -95 & 448 & -197 & 34 & 34 \\ -245 & -104 & 490 & -215 & 37 & 37 \\ -283 & -120 & 566 & -249 & 43 & 43 \end{pmatrix}$$

`u:=gcdex(p,tildeprime,x);`

$$1, -\frac{8 \cdot x^2}{27} + \frac{94 \cdot x}{27} - \frac{121}{27}, \frac{2 \cdot x^5}{27} - \frac{85 \cdot x^4}{54} + \frac{310 \cdot x^3}{27} - \frac{1000 \cdot x^2}{27} + \frac{949 \cdot x}{18} - \frac{155}{6}$$

`v:=op(u,3);`

$$\frac{2 \cdot x^5}{27} - \frac{85 \cdot x^4}{54} + \frac{310 \cdot x^3}{27} - \frac{1000 \cdot x^2}{27} + \frac{949 \cdot x}{18} - \frac{155}{6}$$

`C1:=_plus(coeff(v,x,i)*B^i $ i=0..5);`

$$\begin{pmatrix} \frac{145}{27} & \frac{7}{6} & -\frac{517}{54} & 4 & -\frac{41}{54} & -\frac{10}{27} \\ \frac{391}{27} & \frac{8}{3} & -\frac{1339}{54} & \frac{21}{2} & -\frac{55}{27} & -\frac{53}{54} \\ \frac{965}{54} & \frac{25}{6} & -\frac{1705}{54} & 13 & -\frac{61}{27} & -\frac{37}{27} \\ \frac{1222}{27} & \frac{67}{6} & -\frac{4339}{54} & 33 & -\frac{148}{27} & -\frac{203}{54} \\ \frac{1237}{27} & \frac{35}{3} & -\frac{2213}{27} & 34 & -\frac{145}{27} & -\frac{233}{54} \\ \frac{2993}{54} & \frac{38}{3} & -\frac{2624}{27} & \frac{81}{2} & -\frac{193}{27} & -\frac{239}{54} \end{pmatrix}$$

`D1:=B-C*C1;`

$$\begin{pmatrix} \frac{14}{3} & \frac{1}{2} & -\frac{19}{3} & \frac{5}{2} & -\frac{2}{3} & \frac{1}{3} \\ \frac{29}{3} & \frac{9}{2} & -\frac{58}{3} & \frac{15}{2} & -\frac{5}{3} & \frac{1}{3} \\ \frac{73}{6} & 1 & -\frac{58}{3} & \frac{17}{2} & -\frac{13}{6} & \frac{5}{6} \\ \frac{107}{3} & \frac{3}{2} & -\frac{172}{3} & \frac{49}{2} & -\frac{17}{3} & \frac{7}{3} \\ \frac{211}{6} & -1 & -\frac{160}{3} & \frac{39}{2} & -\frac{19}{6} & \frac{23}{6} \\ \frac{253}{6} & 1 & -\frac{199}{3} & \frac{49}{2} & -\frac{37}{6} & \frac{35}{6} \end{pmatrix}$$

`N:=B-D1;`

5.8 Exercices sur le Chapitre 5

exercice 1

Soit $A \in \mathcal{M}_3(\mathbf{R})$ définie par

$$A = \begin{pmatrix} 3 & 2 & -2 \\ -1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$$

- Déterminer le polynôme caractéristique de A .
- Déterminer le polynôme minimal de A .
- La matrice A est-elle diagonalisable ?

exercice 2

Soit $P \in \mathcal{M}_n(\mathbf{R})$ telle que $P^2 = P$; montrer que P est une matrice diagonalisable .

exercice 3

Calculer $p(A) = 2A^8 - 3A^5 + A^4 + A^2 - 4I_3$, où A est l'élément de $\mathcal{M}_3(\mathbf{R})$ défini par

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

exercice 4

Pour $A \in \mathcal{M}_k(\mathbf{C})$ on note $\bar{A} \in \mathcal{M}_k(\mathbf{C})$ la matrice dont les coefficients sont les conjugués des coefficients de A .

- Vérifier que $\overline{AB} = \bar{A}.\bar{B}$ pour $A, B \in \mathcal{M}_k(\mathbf{C})$.
- En déduire que si $A \in \mathcal{M}_k(\mathbf{C})$ est diagonalisable sur \mathbf{C} , alors \bar{A} l'est aussi.
- Montrer également que si $A \in \mathcal{M}_k(\mathbf{C})$ est nilpotente, alors \bar{A} l'est aussi.
- Soit $A \in \mathcal{M}_k(\mathbf{R})$ et soit $A = D + N$, avec $D \in \mathcal{M}_k(\mathbf{C})$ diagonalisable sur \mathbf{C} , $N \in \mathcal{M}_k(\mathbf{C})$ nilpotente, $ND = DN$ la décomposition de Jordan de A dans $\mathcal{M}_k(\mathbf{C})$. Montrer que $D \in \mathcal{M}_k(\mathbf{R})$ et $N \in \mathcal{M}_k(\mathbf{R})$.

exercice 5

Caractériser toutes les matrices $A \in \mathcal{M}_2(\mathbf{C})$ dont le polynôme minimal est $x^2 + 1$.

exercice 6 Soit E un espace vectoriel de dimension finie sur \mathbf{C} ; on considère un endomorphisme f de E pour lequel il existe un entier $n \geq 1$ vérifiant $f^n = I_E$. Montrer que f est diagonalisable .

exercice 7

On considère la matrice $A = \begin{pmatrix} 3 & 1 & -1 \\ 2 & 2 & -1 \\ 2 & 2 & 0 \end{pmatrix}$

- Calculer le polynôme caractéristique de A .
- Déterminer la dimension du sous-espace propre associé à la valeur propre 2. La matrice A est-elle diagonalisable ?
- Quel est le polynôme minimal de A ?
- Calculer la partie diagonalisable $D(A)$ et la partie nilpotente $N(A)$ de A .
- Déterminer une matrice P inversible telle que

$$P^{-1}D(A)P = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

exercice 8

On considère les matrices réelles suivantes :

$$B = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 2 & 0 \\ -1 & -1 & 2 \end{pmatrix} ; C = \begin{pmatrix} 3 & 3 & 2 & 1 \\ 0 & 1 & 0 & 0 \\ -2 & 3 & -1 & 2 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Pour chacune d'elles, on répondra aux deux questions suivantes :

- Est-elle diagonalisable ?
- Si la réponse est non donner sa décomposition de Jordan .

exercice 9

a) Montrer que les matrices A et B de $\mathcal{M}_3(\mathbf{C})$ définies ci-dessous sont semblables.

$$A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} ; B = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

- En déduire que toute matrice appartenant à $\mathcal{M}_3(\mathbf{C})$ est semblable à sa transposée.
- Généraliser ce résultat pour des matrices carrées à coefficients complexes quelconques.

exercice 10

Soit K un corps de caractéristique nulle, et soit $p \in K[x]$ un polynôme non nul. Montrer que le plus petit entier $m \geq 1$ tel que p divise \tilde{p}^m est égal à la plus grande multiplicité des racines de p dans \tilde{K} .

exercice 11 (sous Mupad, exclusivité ESTIA)

- Déterminer le polynôme caractéristique $p := p_B$ de la matrice

$$\begin{bmatrix} -240 & -220 & -202 & -183 & -165 & -150 & -135 & -120 & -105 & -90 & -75 & -60 & -45 & -30 & -15 \\ 22 & 21 & 21 & 17 & 11 & 10 & 9 & 8 & 7 & 6 & 5 & 4 & 3 & 2 & 1 \\ 612 & 560 & 518 & 478 & 440 & 400 & 360 & 320 & 280 & 240 & 200 & 160 & 120 & 80 & 40 \\ -416 & -392 & -379 & -356 & -330 & -300 & -270 & -240 & -210 & -180 & -150 & -120 & -90 & -60 & -30 \\ 183 & 194 & 201 & 190 & 176 & 150 & 125 & 102 & 82 & 66 & 55 & 44 & 33 & 22 & 11 \\ -326 & -328 & -324 & -311 & -297 & -275 & -234 & -197 & -162 & -132 & -110 & -88 & -66 & -44 & -22 \\ 17 & 17 & 17 & 17 & 17 & 17 & 16 & 16 & 12 & 6 & 5 & 4 & 3 & 2 & 1 \\ 412 & 412 & 412 & 412 & 412 & 412 & 360 & 318 & 278 & 240 & 200 & 160 & 120 & 80 & 40 \\ -266 & -266 & -266 & -266 & -266 & -266 & -242 & -229 & -206 & -180 & -150 & -120 & -90 & -60 & -30 \\ 128 & 128 & 128 & 128 & 128 & 128 & 139 & 146 & 135 & 121 & 95 & 70 & 47 & 27 & 11 \\ -216 & -216 & -216 & -216 & -216 & -216 & -218 & -214 & -201 & -187 & -165 & -124 & -87 & -52 & -22 \\ 12 & 12 & 12 & 12 & 12 & 12 & 12 & 12 & 12 & 12 & 12 & 11 & 11 & 7 & 1 \\ 212 & 212 & 212 & 212 & 212 & 212 & 212 & 212 & 212 & 212 & 212 & 160 & 118 & 78 & 40 \\ -116 & -116 & -116 & -116 & -116 & -116 & -116 & -116 & -116 & -116 & -116 & -92 & -79 & -56 & -30 \\ 33 & 33 & 33 & 33 & 33 & 33 & 33 & 33 & 33 & 33 & 33 & 44 & 51 & 40 & 26 \end{bmatrix}.$$

b) Calculer le pgcd q de p et p' ainsi que le polynôme $\tilde{p} := p/q$. Vérifier que $\tilde{p}^3 = p$.

c) En suivant l'algorithme rapide proposé dans le cours, déterminer la décomposition de Jordan de la matrice B .

exercice 12

Ecrire un programme permettant de calculer sous Mupad la décomposition de Jordan d'une matrice carrée à coefficients complexes quelconque.

Chapitre 6

Itération, systèmes différentiels linéaires

6.1 Calcul des puissances d'une matrice

Le résultat suivant donne un moyen effectif de calculer les puissances d'une matrice dont le polynôme caractéristique est scindé quand on connaît ses valeurs propres.

Proposition 6.1.1 *Soit $A \in \mathcal{M}_n(K)$ une matrice dont le polynôme caractéristique est scindé, et soit $A = D + N$, avec D diagonalisable, N nilpotente, $ND = DN$ la décomposition de Jordan de A . On a alors, pour $p \geq 1$,*

$$A^p = \sum_{k=0}^p C_p^k D^{p-k} N^k = \sum_{k=0}^{\min(m-1,p)} C_p^k D^{p-k} N^k$$

où $C_p^k = \frac{p(p-1)\dots(p-k+1)!}{k!} = \frac{p!}{k!(p-k)!}$ pour $0 \leq k \leq p$, avec les conventions $0! = 1! = 1$, et où m est le plus petit entier ≥ 1 tel que $N^m = 0$.

Démonstration : Comme $ND = DN$, ceci résulte immédiatement du binôme de Newton, voir par exemple le Chapitre 1 du *Cours d'algèbre* [?]. Le fait que l'on peut arrêter la sommation à $m-1$ si $m-1 \leq p$ provient du fait que $N^k = 0$ pour $k \geq m$. ♣

Le calcul de D^p se fait en diagonalisant D : il existe Δ diagonale et $P \in \mathcal{M}_n(K)$ inversibles telles que $\Delta = P^{-1}DP$, soit $D = P\Delta P^{-1}$. Une récurrence immédiate montre que l'on a, pour $p \geq 1$,

$$D^p = P\Delta^p P^{-1} \tag{6.1}$$

D'autre part si $\Delta = \begin{pmatrix} \delta_{1,1} & 0 & \dots & \dots & \dots & 0 \\ 0 & \delta_{2,2} & 0 & \dots & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \delta_{j,j} & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \dots & 0 & \delta_{n,n} \end{pmatrix}$ est diagonale, on a

pour $p \geq 1$

$$\Delta^p = \begin{pmatrix} \delta_{1,1}^p & 0 & \dots & \dots & \dots & 0 \\ 0 & \delta_{2,2}^p & 0 & \dots & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \delta_{j,j}^p & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \dots & 0 & \delta_{n,n}^p \end{pmatrix},$$

ce qui permet de calculer effectivement D^p .

Exemple 6.1.2 On a, pour $p \geq 0$,

$$\begin{bmatrix} 1 & -1 \\ 1 & 3 \end{bmatrix}^p = \begin{bmatrix} 2^p - p2^{p-1} & -p2^{p-1} \\ p2^{p-1} & 2^p + p2^{p-1} \end{bmatrix}.$$

En effet posons $A = \begin{bmatrix} 1 & -1 \\ 1 & 3 \end{bmatrix}$. On a $p_A = x^2 - \text{Tr}(A)x + \det(A) = x^2 - 4x + 4 = (x-2)^2$, donc A possède une valeur propre double égale à 2. Comme le polynôme constant $p = 2$ vérifie trivialement l'équation $p \equiv 2 \pmod{(x-2)^2}$, la décomposition de Jordan $A = D + N$ de A donne $D = p(A) = 2I$, $N = A - 2I = \begin{bmatrix} -1 & -1 \\ 1 & 1 \end{bmatrix}$. On a $N^2 = 0$, et on obtient, pour $p \geq 0$,

$$A^p = 2^p I + p2^{p-1} N = \begin{bmatrix} 2^p - p2^{p-1} & -p2^{p-1} \\ p2^{p-1} & 2^p + p2^{p-1} \end{bmatrix}.$$

Exemple 6.1.3 On a, pour $p \geq 1$,

$$\begin{bmatrix} 0 & 2 & -1 \\ -1 & 4 & -2 \\ -1 & 4 & -2 \end{bmatrix}^p = \begin{bmatrix} 1-p & 2p & -p \\ -p & 2+2p & -1-p \\ -p & 2+2p & -1-p \end{bmatrix}.$$

En effet on connaît la décomposition de Jordan $A = D + N$ de $A = \begin{bmatrix} 0 & 2 & -1 \\ -1 & 4 & -2 \\ -1 & 4 & -2 \end{bmatrix}$, qui est

$$D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & -1 \\ 0 & 2 & -1 \end{bmatrix}, \quad N = \begin{bmatrix} -1 & 2 & -1 \\ -1 & 2 & -1 \\ -1 & 2 & -1 \end{bmatrix}.$$

Les polynômes caractéristiques p_A et p_D de A coïncident, et d'après un calcul effectué au chapitre précédent on a $p_D = p_A = x(x-1)^2$. Comme D est

diagonalisable le sous-espace E_0 associé à la valeur propre 0 de D est engendré par les colonnes de $D - I = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & 2 & -2 \end{bmatrix}$ et le sous espace propre associé à la

valeur propre 1 de D est engendré par les colonnes de $D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & -1 \\ 0 & 2 & -1 \end{bmatrix}$. Donc

si on pose $P = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 2 \end{bmatrix}$, on obtient une matrice inversible dont les colonnes sont des vecteurs propres de D associés aux valeurs propres 1, 1 et 0 de D . On obtient $P^{-1}DP = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$, soit pour $p \geq 1$,

$$D^p = P \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}^p P^{-1} = P \begin{bmatrix} 1^p & 0 & 0 \\ 0 & 1^p & 0 \\ 0 & 0 & 0^p \end{bmatrix} P^{-1} = P \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} P^{-1} = D.$$

D'autre part on a

$$N^2 = \begin{bmatrix} -1 & 2 & -1 \\ -1 & 2 & -1 \\ -1 & 2 & -1 \end{bmatrix} \begin{bmatrix} -1 & 2 & -1 \\ -1 & 2 & -1 \\ -1 & 2 & -1 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

$$ND = DN = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & -1 \\ 0 & 2 & -1 \end{bmatrix} \begin{bmatrix} -1 & 2 & -1 \\ -1 & 2 & -1 \\ -1 & 2 & -1 \end{bmatrix} = \begin{bmatrix} -1 & 2 & -1 \\ -1 & 2 & -1 \\ -1 & 2 & -1 \end{bmatrix} = N,$$

et on obtient, pour $p \geq 1$,

$$A^p = D^p + pD^{p-1}N = D + pN = \begin{bmatrix} 1-p & 2p & -p \\ -p & 2+2p & -1-p \\ -p & 2+2p & -1-p \end{bmatrix}.$$

Un exemple classique d'application de l'itération des matrices est donné par le calcul des suites récurrentes, que nous décrivons ci-dessous.

Exemple 6.1.4 On considère une suite récurrente $(u_n)_{n \geq 0}$, vérifiant pour $n \geq k$ la formule

$$u_n = a_1 u_{n-1} + a_2 u_{n-2} + \dots + a_k u_{n-k} \quad (6.2)$$

Alors si on pose $U_n = \begin{bmatrix} u_n \\ u_{n+1} \\ \vdots \\ u_{n+k-1} \end{bmatrix}$ on a $U_n = A^n U_0$, où

$$A = \begin{bmatrix} 0 & 1 & 0 & \dots & \dots & \dots & 0 \\ 0 & 0 & 1 & 0 & \dots & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \dots & 0 & 1 & 0 \\ 0 & 0 & \dots & \dots & \dots & 0 & 1 \\ a_k & a_{k-1} & \dots & \dots & \dots & a_2 & a_1 \end{bmatrix},$$

En effet il résulte de la formule 6.2 que $U_{n+1} = AU_n$ pour $n \geq 0$, et la formule $U_n = A^n U_0$ s'obtient par une récurrence immédiate.

Considérons par exemple la *suite de Fibonacci* $(u_n)_{n \geq 0}$, avec $u_0 = 0$, $u_1 = 1$, qui vérifie pour $n \geq 2$ la relation

$$u_n = u_{n-1} + u_{n-2}. \quad (6.3)$$

On obtient, pour $n \geq 0$,

$$\begin{bmatrix} u_n \\ u_{n+1} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}^n \begin{bmatrix} u_0 \\ u_1 \end{bmatrix}.$$

Posons $A = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}$. On a

$p_A = x^2 - \text{Tr}(A)x + \det A = x^2 - x - 1$, donc A possède deux valeurs propres distinctes $\lambda_1 = \frac{1-\sqrt{5}}{2}$ et $\lambda_2 = \frac{1+\sqrt{5}}{2}$. On a

$$A - \lambda_2 I = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} - \begin{bmatrix} \frac{1+\sqrt{5}}{2} & 0 \\ 0 & \frac{1+\sqrt{5}}{2} \end{bmatrix} = \begin{bmatrix} -\frac{1+\sqrt{5}}{2} & 1 \\ 1 & \frac{1-\sqrt{5}}{2} \end{bmatrix},$$

$$A - \lambda_1 I = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} - \begin{bmatrix} \frac{1-\sqrt{5}}{2} & 0 \\ 0 & \frac{1-\sqrt{5}}{2} \end{bmatrix} = \begin{bmatrix} -\frac{1-\sqrt{5}}{2} & 1 \\ 1 & \frac{1+\sqrt{5}}{2} \end{bmatrix}.$$

Les matrices $A - \lambda_2 I$ et $A - \lambda_1 I$ sont de rang 1, et les colonnes de $A - \lambda_2 I$ engendrent le sous-espace propre E_{λ_1} de A associé à λ_1 tandis que les colonnes de $A - \lambda_1 I$ engendrent le sous-espace propre E_{λ_2} de A associé à λ_2 . Donc si on pose

$$P = \begin{bmatrix} \frac{1+\sqrt{5}}{2} & 1 \\ -1 & \frac{1+\sqrt{5}}{2} \end{bmatrix},$$

on voit que les colonnes de P donnent une base de \mathbf{R}^2 formée de vecteurs propres de A et on a

$$P^{-1}AP = \begin{bmatrix} \frac{1-\sqrt{5}}{2} & 0 \\ 0 & \frac{1+\sqrt{5}}{2} \end{bmatrix}, \quad A = P \begin{bmatrix} \frac{1-\sqrt{5}}{2} & 0 \\ 0 & \frac{1+\sqrt{5}}{2} \end{bmatrix} P^{-1}.$$

On obtient, pour $p \geq 0$,

$$A^p = P \begin{bmatrix} \left(\frac{1-\sqrt{5}}{2}\right)^p & 0 \\ 0 & \left(\frac{1+\sqrt{5}}{2}\right)^p \end{bmatrix} P^{-1}.$$

On a $\det(P) = \frac{(1+\sqrt{5})^2}{4} + 1 = \frac{3+\sqrt{5}}{2} + 1 = \frac{5+\sqrt{5}}{2}$, $\frac{1}{\det(P)} = \frac{2}{5+\sqrt{5}} = \frac{2(5-\sqrt{5})}{25-5} = \frac{\sqrt{5}-1}{2\sqrt{5}}$. On a donc

$$P^{-1} = \frac{\sqrt{5}-1}{2\sqrt{5}} \begin{bmatrix} \frac{1+\sqrt{5}}{2} & -1 \\ 1 & \frac{1+\sqrt{5}}{2} \end{bmatrix} = \frac{1}{\sqrt{5}} \begin{bmatrix} 1 & \frac{1-\sqrt{5}}{2} \\ \frac{\sqrt{5}-1}{2} & 1 \end{bmatrix}.$$

On obtient, comme $\frac{\sqrt{5}-1}{2} \frac{\sqrt{5}+1}{2} = 1$,

$$\begin{aligned} A^p &= \frac{1}{\sqrt{5}} \begin{bmatrix} \frac{1+\sqrt{5}}{2} & 1 \\ -1 & \frac{1+\sqrt{5}}{2} \end{bmatrix} \begin{bmatrix} \left(\frac{1-\sqrt{5}}{2}\right)^p & 0 \\ 0 & \left(\frac{1+\sqrt{5}}{2}\right)^p \end{bmatrix} \begin{bmatrix} 1 & \frac{1-\sqrt{5}}{2} \\ \frac{\sqrt{5}-1}{2} & 1 \end{bmatrix} \\ &= \frac{1}{\sqrt{5}} \begin{bmatrix} \frac{1+\sqrt{5}}{2} & 1 \\ -1 & \frac{1+\sqrt{5}}{2} \end{bmatrix} \begin{bmatrix} \left(\frac{1-\sqrt{5}}{2}\right)^p & \left(\frac{1-\sqrt{5}}{2}\right)^{p+1} \\ \left(\frac{1+\sqrt{5}}{2}\right)^{p-1} & \left(\frac{1+\sqrt{5}}{2}\right)^p \end{bmatrix} \\ &= \frac{1}{\sqrt{5}} \begin{bmatrix} \left(\frac{1+\sqrt{5}}{2}\right)^{p-1} - \left(\frac{1-\sqrt{5}}{2}\right)^{p-1} & \left(\frac{1+\sqrt{5}}{2}\right)^p - \left(\frac{1-\sqrt{5}}{2}\right)^p \\ \left(\frac{1+\sqrt{5}}{2}\right)^p - \left(\frac{1-\sqrt{5}}{2}\right)^p & \left(\frac{1+\sqrt{5}}{2}\right)^{p+1} - \left(\frac{1-\sqrt{5}}{2}\right)^{p+1} \end{bmatrix}. \end{aligned}$$

En revenant à la suite de Fibonacci, on obtient, pour $n \geq 0$,

$$\begin{bmatrix} u_n \\ u_{n+1} \end{bmatrix} = \frac{1}{\sqrt{5}} \begin{bmatrix} \left(\frac{1+\sqrt{5}}{2}\right)^{n-1} - \left(\frac{1-\sqrt{5}}{2}\right)^{n-1} & \left(\frac{1+\sqrt{5}}{2}\right)^n - \left(\frac{1-\sqrt{5}}{2}\right)^n \\ \left(\frac{1+\sqrt{5}}{2}\right)^n - \left(\frac{1-\sqrt{5}}{2}\right)^n & \left(\frac{1+\sqrt{5}}{2}\right)^{n+1} - \left(\frac{1-\sqrt{5}}{2}\right)^{n+1} \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix},$$

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

Notons qu'il existe une méthode bien connue plus simple pour calculer les termes de la suite de Fibonacci : si a et b sont deux réels, on vérifie immédiatement que si on pose $v_n = a\lambda_1^n + b\lambda_2^n$ pour $n \geq 0$, alors la suite $(v_n)_{n \geq 0}$ vérifie la relation de récurrence 6-3. Comme une suite vérifiant cette relation est déterminée de manière unique par la donnée de ses deux premiers termes, on a, pour $n \geq 0$,

$$u_n = a\lambda_1^n + b\lambda_2^n,$$

où a et b vérifient le système

$$\begin{cases} a + b = 0 \\ a\lambda_1 + b\lambda_2 = 1 \end{cases}$$

On obtient $a = \frac{\begin{vmatrix} 0 & 1 \\ 1 & \lambda_2 \end{vmatrix}}{\begin{vmatrix} 1 & 1 \\ \lambda_1 & \lambda_2 \end{vmatrix}} = \frac{-1}{\lambda_2 - \lambda_1} = -\frac{1}{\sqrt{5}}$ et $b = \frac{\begin{vmatrix} 1 & 0 \\ \lambda_1 & 1 \end{vmatrix}}{\begin{vmatrix} 1 & 1 \\ \lambda_1 & \lambda_2 \end{vmatrix}} = \frac{1}{\lambda_2 - \lambda_1} = \frac{1}{\sqrt{5}}$.

Donc $u_n = \frac{\lambda_2^n - \lambda_1^n}{\sqrt{5}}$ pour $n \geq 0$, et on retrouve le résultat précédent. On peut en fait appliquer la méthode des suites récurrentes pour calculer les puissances successives d'une matrice carrée A . En effet soit $q_A = a_0 + a_1x + \dots + a_{n-1}x^{n-1} + x^n$ le polynôme minimal de A . Comme $q_A(A) = 0$, on a pour $n \geq m$,

$$A^n = -a_{m-1}A^{n-1} - a_{m-2}A^{n-2} - \dots - a_0A^{n-m},$$

et on obtient une suite récurrente à valeurs matricielles.

Soient $\lambda_1, \dots, \lambda_k$ les racines distinctes de q_A et soient p_1, \dots, p_k leurs ordres de multiplicité. On vérifie, avec la convention $0^0 = 1$, que la suite $(n^p \lambda_j^n)_{n \geq 0}$ vérifie la forme scalaire de la relation de récurrence ci-dessus pour $1 \leq j \leq k$, $0 \leq p \leq p_j$. On a donc, pour $n \geq 0$,

$$A^n = \sum_{j=1}^k \left(\sum_{p=0}^{p_j-1} n^p \lambda_j^n B_{p,j} \right),$$

à condition que les matrices $B_{p,j}$ vérifient le système

$$\left\{ \begin{array}{l} \sum_{j=1}^k B_{0,j} = I \\ \sum_{j=1}^k \left(\sum_{p=0}^{p_j-1} \lambda_j B_{p,j} \right) = A \\ \dots \dots \dots \dots \dots \\ \dots \dots \dots \dots \dots \\ \sum_{j=1}^k \left(\sum_{p=0}^{p_j-1} s^p \lambda_j^s B_{p,j} \right) = A^s \\ \dots \dots \dots \dots \dots \\ \dots \dots \dots \dots \dots \\ \sum_{j=1}^k \left(\sum_{p=0}^{p_j-1} (m-1)^p \lambda_j^{m-1} B_{p,j} \right) = A^{m-1} \end{array} \right.$$

On obtient un système de Cramer à inconnues matricielles, qui admet une solution unique permettant d'obtenir une formule explicite pour le calcul de A^n . Nous illustrons ceci sur les exemples étudiés plus haut.

Si $A = \begin{bmatrix} 1 & -1 \\ 1 & 3 \end{bmatrix}$, qui possède une valeur propre double égale à 2, on résout le système

$$\begin{cases} B = I \\ 2B + 2C = A \end{cases}$$

qui a pour solution $B = I, C = \frac{A-2I}{2}$, et on obtient, pour $p \geq 0$,

$$A^p = 2^p I + p2^p \frac{A - 2I}{2} = 2^p I + 2^{p-1} \begin{bmatrix} -1 & -1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 2^p - p2^{p-1} & -p2^{p-1} \\ p2^{p-1} & 2^p + p2^{p-1} \end{bmatrix}.$$

Soit maintenant $A = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}$, dont les valeurs propres sont $\lambda_1 = \frac{1-\sqrt{5}}{2}$ et $\lambda_2 = \frac{1+\sqrt{5}}{2}$. On résout le système

$$\begin{cases} B + C = I \\ \lambda_1 B + \lambda_2 C = A \end{cases}$$

On obtient $\lambda_1 B + \lambda_2(I - B) = A$, soit $B = -\frac{1}{\sqrt{5}}(A - \lambda_2 I)$, et $\lambda_1(I - C) + \lambda_2 C = A$, soit $C = \frac{1}{\sqrt{5}}(A - \lambda_1 I)$. On obtient, pour $p \geq 0$, en remarquant que $\lambda_1 \lambda_2 = 1$, et que λ_1 et λ_2 sont solutions de l'équation $x^{p+1} = x^p + x^{p-1}$.

$$\begin{aligned} \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}^p &= -\frac{\lambda_1^p}{\sqrt{5}}(A - \lambda_2 I) + \frac{\lambda_2^p}{\sqrt{5}}(A - \lambda_1 I) = \frac{\lambda_2^p - \lambda_1^p}{\sqrt{5}}A - \frac{\lambda_2^{p-1} - \lambda_1^{p-1}}{\sqrt{5}}I = \\ &= \frac{1}{\sqrt{5}} \begin{bmatrix} \left(\frac{1+\sqrt{5}}{2}\right)^{p-1} - \left(\frac{1-\sqrt{5}}{2}\right)^{p-1} & \left(\frac{1+\sqrt{5}}{2}\right)^p - \left(\frac{1-\sqrt{5}}{2}\right)^p \\ \left(\frac{1+\sqrt{5}}{2}\right)^p - \left(\frac{1-\sqrt{5}}{2}\right)^p & \left(\frac{1+\sqrt{5}}{2}\right)^{p+1} - \left(\frac{1-\sqrt{5}}{2}\right)^{p+1} \end{bmatrix} \\ &= \frac{1}{\sqrt{5}} \begin{bmatrix} \left(\frac{1+\sqrt{5}}{2}\right)^{p-1} - \left(\frac{1-\sqrt{5}}{2}\right)^{p-1} & \left(\frac{1+\sqrt{5}}{2}\right)^p - \left(\frac{1-\sqrt{5}}{2}\right)^p \\ \left(\frac{1+\sqrt{5}}{2}\right)^p - \left(\frac{1-\sqrt{5}}{2}\right)^p & \left(\frac{1+\sqrt{5}}{2}\right)^{p+1} - \left(\frac{1-\sqrt{5}}{2}\right)^{p+1} \end{bmatrix}. \end{aligned}$$

Posons maintenant $A = \begin{bmatrix} 0 & 2 & -1 \\ -1 & 4 & -2 \\ -1 & 4 & -2 \end{bmatrix}$. On a vu que $p_A = x(x-1)^2 = p_D$, où $A = D + N$ est la décomposition de Jordan de A . On a donc $q_D = x(x-1)$ et $D^2 - D = q_D(D) = 0$. Donc $D^p = D$ pour tout $p \geq 1$, ce qui donne

$$A^p = D + pDN = D + pN = \begin{bmatrix} 1-p & 2p & -p \\ -p & 2+2p & -1-p \\ -p & 2+2p & -1-p \end{bmatrix}.$$

En règle générale quand une matrice carrée à 3 lignes et 3 colonnes possède une valeur propre simple λ_1 et une valeur propre double λ_2 , le polynôme minimal q_D est égal à $(x - \lambda_1)(x - \lambda_2)$, $A = D + N$ désignant la décomposition de Jordan de A , et on a $D^p = \lambda_1^p B + \lambda_2^p C$, où B et C sont les solutions du système

$$\begin{cases} B + C = I \\ \lambda_1 B + \lambda_2 C = D \end{cases}$$

On voit donc que dans ce cas il est inutile de diagonaliser D pour calculer D^p .

6.2 Rappel sur les équations différentielles linéaires du premier ordre

On étudie en Terminale les équations différentielles du type

$$y'(t) = ay(t),$$

où $a \in \mathbf{R}$. Il est bien connu que la solution générale d'une telle équation est de la forme

$$y(t) = \lambda e^{ta},$$

où λ est une constante réelle, évidemment égale à $y(0)$.

On s'intéresse plus généralement aux équations différentielles du type

$$y'(t) = ay(t) + g(t),$$

où g est une fonction continue sur un intervalle ouvert I de \mathbf{R} . On dispose de deux méthodes. La première, basée sur le principe de la "variation de la constante", donne une formule très générale.

Théorème 6.2.1 *Soit I un intervalle de \mathbf{R} , soit $g : I \rightarrow \mathbf{R}$ une fonction continue et soit $a \in \mathbf{R}$. On considère l'équation différentielle*

$$(*) \quad y'(t) = ay(t) + g(t).$$

Alors pour tout $t_0 \in I$ et pour tout $b \in \mathbf{R}$ il existe sur I une unique solution y de l'équation () vérifiant $y(t_0) = c$, qui est donnée par la formule*

$$(6.1) \quad y(t) = e^{(t-t_0)a}c + \int_{t_0}^t e^{(t-s)a}g(s)ds = e^{(t-t_0)a}y(t_0) + \int_{t_0}^t e^{(t-s)a}g(s)ds.$$

Preuve : Soit $y : t \mapsto y(t)$ une fonction dérivable sur I , et posons $\lambda(t) = y(t)e^{-ta}$. Alors $y(t) = \lambda(t)e^{ta}$ (d'où le nom de "variation de la constante"), $y'(t) = \lambda'(t)e^{ta} + a\lambda(t)e^{ta} = \lambda'(t)e^{ta} + ay(t)$, donc y est solution de l'équation (*) si et seulement si on a, pour tout $t \in I$,

$$\lambda'(t)e^{ta} = g(t), \quad \lambda'(t) = e^{-ta}g(t).$$

D'autre part $y(t_0) = c$ si et seulement si $\lambda(t_0) = e^{-t_0a}c$. Les deux conditions ci-dessus caractérisent l'unique primitive sur I de la fonction $t \mapsto e^{-ta}g(t)$ qui prend la valeur $e^{-t_0a}c$ en t_0 . Cette fonction est définie sur I par la formule

$$\lambda(t) = e^{-t_0a}c + \int_{t_0}^t e^{-sa}g(s)ds.$$

On obtient une unique solution y , définie sur I par la formule

$$y(t) = e^{ta}\lambda(t) = e^{(t-t_0)a}c + \int_{t_0}^t e^{(t-s)a}g(s)ds = e^{t-t_0)a}y(t_0) + \int_{t_0}^t e^{(t-s)a}g(s)ds.$$

♣

Exemple 6.2.2 Résoudre l'équation différentielle

$$y'(t) = 2y(t) + 1,$$

avec la condition initiale $y(0) = 1$.

On a la condition initiale $y(0) = 1$, et ici $g(t) \equiv 1, a = 2$. On obtient, en appliquant la formule (6.1) :

$$\begin{aligned} y(t) &= e^{2t} + \int_0^t e^{2(t-s)}ds = e^{2t} + e^{2t} \left[-\frac{e^{-2s}}{2} \right]_0^t = e^{2t} \left[1 - \frac{e^{-2t}}{2} + \frac{1}{2} \right] \\ &= \frac{3}{2}e^{2t} - \frac{1}{2}. \end{aligned}$$

La deuxième méthode, plus empirique, est basée sur le "principe de superposition" suivant, qui permet de simplifier les calculs quand on a un moyen simple de trouver une solution particulière.

Proposition 6.2.3 Soit I un intervalle de \mathbf{R} , soit $g : I \rightarrow \mathbf{R}$ une fonction continue et soit $a \in \mathbf{R}$. On considère l'équation différentielle

$$(*) \quad y'(t) = ay(t) + g(t).$$

Si $y_0 : I \rightarrow \mathbf{R}$ est une solution particulière de l'équation (*), alors la solution générale de cette équation est donnée par la formule

$$y(t) = y_0(t) + \lambda e^{ta},$$

où λ est une constante réelle.

Preuve : Soit $t \mapsto y(t)$ une fonction dérivable sur I , et posons $u(t) = y(t) - y_0(t)$.

On a $y'(t) - ay(t) - g(t) = u'(t) - au(t) + y_0'(t) - ay_0(t) - g(t) = u'(t) - au(t)$. Par conséquent y est solution de l'équation (*) si et seulement si u est solution de l'équation homogène associée

$$(**) \quad u'(t) = au(t),$$

ce qui donne $u(t) = \lambda e^{ta}$, $y(t) = y_0(t) + \lambda e^{ta}$, où λ est une constante réelle. ♣

et dans ce cas on pose

$$\sum_{n=0}^{+\infty} U_n = \left(\sum_{n=0}^{+\infty} u_{i,j,n} \right)_{\substack{1 \leq i \leq p \\ 1 \leq j \leq q}}.$$

3) On dit qu'une fonction $F : t \mapsto (f_{i,j}(t))_{\substack{1 \leq i \leq p \\ 1 \leq j \leq q}}$ définie sur un intervalle ouvert I de \mathbf{R} et à valeurs dans $\mathcal{M}_{p,q}(\mathbf{C})$ est continue sur I quand $f_{i,j}$ est continue sur I pour $1 \leq i \leq p, 1 \leq j \leq q$, et dans ce cas on pose, pour $a, b \in I$,

$$\int_a^b F(t) dt = \left(\int_a^b f_{i,j}(t) dt \right)_{\substack{1 \leq i \leq p \\ 1 \leq j \leq q}}.$$

4) On dit qu'une fonction $F : t \mapsto (f_{i,j}(t))_{\substack{1 \leq i \leq p \\ 1 \leq j \leq q}}$ définie sur un intervalle ouvert I de \mathbf{R} et à valeurs dans $\mathcal{M}_{p,q}(\mathbf{C})$ est dérivable sur I quand $f_{i,j}$ est dérivable sur I pour $1 \leq i \leq p, 1 \leq j \leq q$, et dans ce cas on pose, pour $t \in I$,

$$F'(t) = (f'_{i,j}(t))_{\substack{1 \leq i \leq p \\ 1 \leq j \leq q}}.$$

Revenons à un système différentiel linéaire à coefficients constants de la forme (6.1) et posons

$$A = (a_{i,j})_{\substack{1 \leq i \leq k \\ 1 \leq j \leq k}}, Y(t) = \begin{bmatrix} y_1(t) \\ \cdot \\ \cdot \\ \cdot \\ y_k(t) \end{bmatrix}, G(t) = \begin{bmatrix} g_1(t) \\ \cdot \\ \cdot \\ \cdot \\ g_k(t) \end{bmatrix}.$$

Compte tenu de la définition ci-dessus, on a

$$Y'(t) = \begin{bmatrix} y'_1(t) \\ \cdot \\ \cdot \\ \cdot \\ y'_k(t) \end{bmatrix}.$$

Le système prend la forme

$$Y'(t) = AY(t) + G(t).$$

On va maintenant résoudre ce système en utilisant une forme appropriée de la méthode de la variation de la constante. On va utiliser le résultat simple suivant.

Proposition 6.3.2 Soit I un intervalle ouvert de \mathbf{R} .

(i) Soit $t_0 \in I$, et soit $F : I \rightarrow \mathcal{M}_{p,q}(\mathbf{C})$ une application continue. Posons, pour $t \in I$,

$$G(t) = \int_{t_0}^t F(s) ds.$$

Alors $G(t_0)$ est la matrice nulle, G est dérivable sur I et $G'(t) = F(t)$ pour tout $t \in I$.

(ii) Soient $F : I \rightarrow \mathcal{M}_{p,q}(\mathbf{C})$ et $G : I \rightarrow \mathcal{M}_{q,r}(\mathbf{C})$ deux applications dérivables sur I . Alors $FG : t \mapsto F(t)G(t)$ est dérivable sur I , et on a pour $t \in I$

$$(FG)'(t) = F(t)G'(t) + F'(t)G(t). \quad (6.5)$$

Preuve : (i) résulte immédiatement du résultat analogue pour les fonctions continues à valeurs réelles ou complexes et de la définition 6.2.1. Pour démontrer (ii) notons $f_{i,k}(t)$ le coefficient d'indice i, k de $F(t)$ pour $1 \leq i \leq p$, $1 \leq k \leq q$, notons $g_{k,j}(t)$ le coefficient d'indice k, j de $G(t)$ pour $1 \leq k \leq q$, $1 \leq j \leq r$, et notons $(fg)_{i,j}(t)$ le coefficient d'indice i, j de $F(t)G(t)$ pour $1 \leq i \leq p$, $1 \leq j \leq r$. On a

$$(fg)_{i,j}(t) = \sum_{k=1}^q f_{i,k}(t)g_{k,j}(t).$$

Par conséquent $(fg)_{i,j}$ est dérivable sur I pour $1 \leq i \leq p$, $1 \leq j \leq q$, et on a

$$(fg)'_{i,j}(t) = \sum_{k=1}^q f_{i,k}(t)g'_{k,j}(t) + \sum_{k=1}^q f'_{i,k}(t)g_{k,j}(t).$$

Par conséquent FG est dérivable sur I , et on a

$$(FG)'(t) = F(t)G'(t) + F'(t)G(t). \clubsuit$$

Pour $a \in \mathbf{R}$, (voir par exemple le Chapitre 1 du *Cours d'Analyse*) [3] on a

$$e^a = \sum_{n=0}^{+\infty} \frac{a^n}{n!}.$$

Soit maintenant $A \in \mathcal{M}_k(\mathbf{C})$ une matrice carrée à k lignes et k colonnes à coefficients réels ou complexes. On a vu au Chapitre 4 du cours d'analyse [3] que la série $\sum_{m=0}^{+\infty} \frac{A^m}{m!}$ converge dans $\mathcal{M}_k(\mathbf{C})$. On définit alors l'exponentielle de la matrice A par la formule

$$e^A = \sum_{m=0}^{+\infty} \frac{A^m}{m!}. \quad (6.6)$$

On a le résultat suivant (Proposition 4.5.3 de [3])

Proposition 6.3.3 Soient $A, B \in \mathcal{M}_k(\mathbf{C})$. Si $AB = BA$, alors $e^{A+B} = e^A e^B$. En particulier e^A est inversible pour toute matrice $A \in \mathcal{M}_k(\mathbf{C})$, et on a

$$(e^A)^{-1} = e^{-A}. \quad (6.7)$$

Il résulte également du théorème 4.5.4 de [3] que si on pose $F(t) = e^{tA}$ pour $t \in \mathbf{R}$, avec $A \in \mathcal{M}_k(\mathbf{R})$, alors F est dérivable sur \mathbf{R} , et que l'on a pour $t \in \mathbf{R}$

$$F'(t) = AF(t) = Ae^{tA},$$

ce qui peut s'écrire sous la forme

$$d(e^{tA}) = Ae^{tA} dt \quad (6.8)$$

Dans la suite de ce Chapitre on identifiera \mathbf{R}^k à l'espace vectoriel $\mathcal{M}_{k,1}(\mathbf{R})$ des matrices unicolonnes à k coefficients réels. Nous sommes maintenant en mesure d'adapter aux systèmes différentiels linéaires à coefficients constants la méthode de la variation de la constante, ce qui amène au résultat suivant.

Théorème 6.3.4 Soit I un intervalle de \mathbf{R} , soit $A \in \mathcal{M}_k(\mathbf{R})$, et soit $G : I \rightarrow \mathbf{R}^k$ une fonction continue. On considère le système différentiel

$$Y'(t) = AY(t) + G(t) \quad (6.9)$$

Alors pour tout $t_0 \in I$ et pour tout $C \in \mathbf{R}^k$ il existe sur I une unique solution Y du système différentiel vérifiant $Y(t_0) = C$, qui est donnée par la formule

$$Y(t) = e^{(t-t_0)A}C + \int_{t_0}^t e^{(t-s)A}G(s)ds = e^{(t-t_0)A}Y(t_0) + \int_{t_0}^t e^{(t-s)A}G(s)ds. \quad (6.10)$$

Preuve : Soit $Y : t \mapsto Y(t)$ une fonction dérivable sur I , et posons $\Lambda(t) = e^{-tA}Y(t)$. Alors $Y(t) = e^{tA}\Lambda(t)$, $Y'(t) = e^{tA}\Lambda'(t) + Ae^{tA}\Lambda(t) = e^{tA}\Lambda'(t) + AY(t)$, donc Y est solution du système proposé si et seulement si on a, pour tout $t \in I$,

$$e^{tA}\Lambda'(t) = G(t), \quad \Lambda'(t) = e^{-tA}G(t).$$

D'autre part $Y(t_0) = C$ si et seulement si $\Lambda(t_0) = e^{-t_0A}C$. Les deux conditions ci-dessus caractérisent l'unique primitive sur I de la fonction $t \mapsto e^{-tA}G(t)$ qui prend la valeur $e^{-t_0A}C$ en t_0 . Cette fonction est définie sur I par la formule

$$\Lambda(t) = e^{-t_0A}C + \int_{t_0}^t e^{-sA}G(s)ds.$$

On obtient une unique solution $Y : I \rightarrow \mathbf{R}^k$, définie sur I par la formule

$$Y(t) = e^{tA} \Lambda(t) = e^{(t-t_0)A} C + \int_{t_0}^t e^{(t-s)A} G(s) ds = e^{(t-t_0)A} Y(t_0) + \int_{t_0}^t e^{(t-s)A} G(s) ds.$$

♣

On a vu au chapitre 4 de [3] que si $B \in \mathcal{M}_k(\mathbf{C})$, et si $P \in \mathcal{M}_k(\mathbf{C})$ est inversible, on a

$$e^{PBP^{-1}} = Pe^B P^{-1}, \quad (6.11)$$

$$\text{D'autre part soit } B = \begin{pmatrix} b_{1,1} & 0 & \dots & \dots & \dots & 0 \\ 0 & b_{2,2} & 0 & \dots & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & b_{j,j} & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \dots & 0 & b_{k,k} \end{pmatrix} \text{ une matrice dia-}$$

gonale. On a, d'après les formules 4.7 et 6.3

$$e^B = \lim_{n \rightarrow +\infty} \begin{pmatrix} \sum_{m=0}^n \frac{b_{1,1}^m}{m!} & 0 & \dots & \dots & \dots & 0 \\ 0 & \sum_{m=0}^n \frac{b_{2,2}^m}{m!} & 0 & \dots & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \sum_{m=0}^n \frac{b_{j,j}^m}{m!} & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \dots & 0 & \sum_{m=0}^n \frac{b_{k,k}^m}{m!} \end{pmatrix}.$$

On obtient

$$e^B = \begin{pmatrix} e^{b_{1,1}} & 0 & \dots & \dots & \dots & 0 \\ 0 & e^{b_{2,2}} & 0 & \dots & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & e^{b_{j,j}} & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \dots & 0 & e^{b_{k,k}} \end{pmatrix},$$

d'où, pour $t \in \mathbf{R}$,

$$e^{tB} = \begin{pmatrix} e^{tb_{1,1}} & 0 & \dots & \dots & \dots & 0 \\ 0 & e^{tb_{2,2}} & 0 & \dots & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & e^{tb_{j,j}} & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \dots & 0 & e^{tb_{k,k}} \end{pmatrix} \quad (6.12)$$

Si $N \in \mathcal{M}_k(\mathbf{C})$ est nilpotente, on a $N^p = 0$, avec $1 \leq p \leq k$, et on a

$$e^N = \sum_{m=0}^{+\infty} \frac{N^m}{m!} = \sum_{m=0}^{p-1} \frac{N^m}{m!},$$

d'où, pour $t \in \mathbf{R}$,

$$e^{tN} = \sum_{m=0}^{p-1} \frac{t^m N^m}{m!}. \quad (6.13)$$

Soit maintenant $A \in \mathcal{M}_k(\mathbf{R})$, et soit $A = D + N$, avec D diagonalisable sur \mathbf{C} , N nilpotente, $ND = DN$, la décomposition de Jordan de A vue au Chapitre précédent (théorème 5.5.1). On a alors, pour $t \in \mathbf{R}$,

$$e^{tA} = e^{tD} e^{tN}, \quad (6.14)$$

et on calcule séparément e^{tD} et e^{tN} comme indiqué ci-dessus.

On notera que même dans le cas où $A \in \mathcal{M}_k(\mathbf{R})$ possède des valeurs propres complexes non réelles, la partie diagonalisable D et la partie nilpotente N de la décomposition de Jordan de A dans $\mathcal{M}_k(\mathbf{C})$ appartiennent à $\mathcal{M}_k(\mathbf{R})$, voir l'exercice 4 du Chapitre 5.

Exemple 6.3.5 Résoudre le système différentiel

$$\begin{cases} x'(t) = 3x(t) + y(t) + 1 \\ y'(t) = -x(t) + y(t) - 1 \end{cases}$$

avec la condition initiale $x(0) = 1, y(0) = 0$.

Si on pose $X(t) = \begin{bmatrix} x(t) \\ y(t) \end{bmatrix}$, $A = \begin{bmatrix} 3 & 1 \\ -1 & 1 \end{bmatrix}$, le système devient

$$X'(t) = AX(t) + \begin{bmatrix} 1 \\ -1 \end{bmatrix}.$$

Soit $A = D + N$ la décomposition de Jordan de A . On a vu au Chapitre précédent (exemple 5.5.2) que l'on a $D = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$, $N = \begin{bmatrix} 1 & 1 \\ -1 & -1 \end{bmatrix}$.

On obtient

$$e^{tD} = \begin{bmatrix} e^{2t} & 0 \\ 0 & e^{2t} \end{bmatrix} = e^{2t} I_2,$$

$$e^{tN} = I_2 + tN = \begin{bmatrix} 1+t & t \\ -t & 1-t \end{bmatrix},$$

$$e^{tA} = e^{tD} e^{tN} = \begin{bmatrix} (1+t)e^{2t} & te^{2t} \\ -te^{2t} & (1-t)e^{2t} \end{bmatrix},$$

$$X(t) = e^{tA} \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \int_0^t e^{(t-s)A} \begin{bmatrix} 1 \\ -1 \end{bmatrix} ds$$

$$\begin{aligned}
&= \begin{bmatrix} (1+t)e^{2t} \\ -te^{2t} \end{bmatrix} + \int_0^t e^{2(t-s)} \begin{bmatrix} 1+(t-s) & (t-s) \\ -(t-s) & 1-(t-s) \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix} ds \\
&= \begin{bmatrix} (1+t)e^{2t} \\ -te^{2t} \end{bmatrix} + e^{2t} \int_0^t e^{-2s} \begin{bmatrix} 1 \\ -1 \end{bmatrix} ds.
\end{aligned}$$

Comme $\int_0^t e^{-2s} ds = \left[-\frac{e^{-2s}}{2} \right]_0^t = -\frac{e^{-2t}}{2} + \frac{1}{2}$, on obtient

$$\begin{cases} x(t) = (1+t)e^{2t} + e^{2t} \int_0^t e^{-2s} ds = \left(\frac{3}{2} + t\right)e^{2t} - \frac{1}{2} \\ y(t) = -te^{2t} - e^{-2t} \int_0^t e^{2s} ds = -\left(\frac{1}{2} + t\right)e^{2t} + \frac{1}{2} \end{cases}$$

Le fait que certaines matrices carrées réelles soient diagonalisables sur \mathbf{C} mais pas sur \mathbf{R} peut causer quelques complications dans le calcul de leur exponentielle, comme le montre l'exemple suivant.

Exemple 6.3.6 Résoudre le système différentiel

$$\begin{cases} x'(t) = x(t) + y(t) + 1 \\ y'(t) = -x(t) + y(t) - 1 \end{cases}$$

avec la condition initiale $x(0) = y(0) = 0$.

Si on pose $X(t) = \begin{bmatrix} x(t) \\ y(t) \end{bmatrix}$, $A = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$, le système devient

$$X'(t) = AX(t) + \begin{bmatrix} 1 \\ -1 \end{bmatrix}.$$

On a $\text{Tr}(A) = \det(A) = 2$, donc $p_A = x^2 - 2x + 2$. On obtient deux racines complexes $\lambda_1 = 1 + i$ et $\lambda_2 = 1 - i$, donc A possède deux valeurs propres distinctes et est diagonalisable sur \mathbf{C} .

On a $A \begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} u+v \\ -u+v \end{bmatrix}$. L'équation $A \begin{bmatrix} u \\ v \end{bmatrix} = (1+i) \begin{bmatrix} u \\ v \end{bmatrix}$ donne

$$\begin{cases} v = iu \\ -u = iv \end{cases}$$

qui se réduit à l'équation $y = ix$. Donc $V_1 = \begin{bmatrix} 1 \\ i \end{bmatrix}$ est un vecteur propre de A associé à la valeur propre $1+i$. Comme les coefficients de A sont réels il est clair que $V_2 = \bar{V}_1 := \begin{bmatrix} \bar{1} \\ \bar{1+i} \end{bmatrix} = \begin{bmatrix} 1 \\ 1-i \end{bmatrix}$ est un vecteur propre de A associé à la valeur propre $\overline{1+i} = 1-i$.

Posons $P = [V_1, V_2] = \begin{bmatrix} 1 & 1 \\ i & -i \end{bmatrix}$. Alors $P^{-1}AP = \begin{bmatrix} 1+i & 0 \\ 0 & 1-i \end{bmatrix}$ est diagonale, et on obtient, pour $t \in \mathbf{R}$,

$$e^{tA} = P \begin{bmatrix} e^{t(1+i)} & 0 \\ 0 & e^{t(1-i)} \end{bmatrix} P^{-1}.$$

Comme $\det(P) = -2i$, on a $\frac{1}{\det(P)} = \frac{i}{2}$, $P^{-1} = \frac{i}{2} \begin{bmatrix} -i & -1 \\ -i & 1 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 & -i \\ 1 & i \end{bmatrix}$,
 et on obtient

$$\begin{aligned} e^{tA} &= \frac{1}{2} \begin{bmatrix} 1 & 1 \\ i & -i \end{bmatrix} \begin{bmatrix} e^{t(1+i)} & 0 \\ 0 & e^{t(1-i)} \end{bmatrix} \begin{bmatrix} 1 & -i \\ 1 & i \end{bmatrix} = \frac{e^t}{2} \begin{bmatrix} 1 & 1 \\ i & -i \end{bmatrix} \begin{bmatrix} e^{it} & -ie^{it} \\ e^{-it} & ie^{-it} \end{bmatrix} \\ &= \frac{e^t}{2} \begin{bmatrix} e^{it} + e^{-it} & -ie^{it} + ie^{-it} \\ ie^{it} - ie^{-it} & e^{it} + e^{-it} \end{bmatrix} = \frac{e^t}{2} \begin{bmatrix} 2\cos(t) & -i(2i\sin(t)) \\ i(2i\sin(t)) & 2\cos(t) \end{bmatrix} \\ &= e^t \begin{bmatrix} \cos(t) & \sin(t) \\ -\sin(t) & \cos(t) \end{bmatrix}. \end{aligned}$$

La condition initiale donnée est $x(0) = y(0) = 0$, donc la solution cherchée $X(t) = \begin{bmatrix} x(t) \\ y(t) \end{bmatrix}$ est donnée par la formule

$$X(t) = \int_0^t e^{(t-s)A} \begin{bmatrix} 1 \\ -1 \end{bmatrix} ds.$$

En posant $r = t - s$, on obtient $dr = -ds$, donc on a

$$X(t) = - \int_t^0 e^{rA} \begin{bmatrix} 1 \\ -1 \end{bmatrix} dr = \int_0^t e^{rA} \begin{bmatrix} 1 \\ -1 \end{bmatrix} dr.$$

On peut alors procéder à un calcul direct : comme $\det(A) = 2$, A est inversible et $A^{-1} = \frac{1}{2} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$. Posons $F(r) = A^{-1}e^{rA} \begin{bmatrix} 1 \\ -1 \end{bmatrix}$. On a $F'(r) = A^{-1}Ae^{rA} \begin{bmatrix} 1 \\ -1 \end{bmatrix} = e^{rA} \begin{bmatrix} 1 \\ -1 \end{bmatrix}$, et on obtient

$$\begin{aligned} X(t) &= \int_0^t e^{rA} \begin{bmatrix} 1 \\ -1 \end{bmatrix} dr = [F(r)]_0^t = A^{-1} \left[e^{tA} - \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right] \begin{bmatrix} 1 \\ -1 \end{bmatrix} \\ &= \frac{1}{2} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} e^t \cos(t) - 1 & e^t \sin(t) \\ -e^t \sin(t) & e^t \cos(t) - 1 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix} \\ &= \frac{1}{2} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} e^t \cos(t) - e^t \sin(t) - 1 \\ -e^t \cos(t) - e^t \sin(t) + 1 \end{bmatrix} \\ &= \begin{bmatrix} e^t \cos(t) - 1 \\ -e^t \sin(t) \end{bmatrix}. \end{aligned}$$

Ceci donne

$$\begin{cases} x(t) = e^t \cos(t) - 1 \\ y(t) = -e^t \sin(t) \end{cases}$$

Exemple 6.3.7 Résoudre le système différentiel

$$\begin{cases} x'(t) = x(t) + 2y(t) - z(t) \\ y'(t) = -x(t) + 5y(t) - 2z(t) \\ z'(t) = -x(t) + 4y(t) - z(t) + t \end{cases}$$

avec la condition initiale $x(0) = y(0) = z(0) = 0$.

Posons $X(t) = \begin{bmatrix} x(t) \\ y(t) \\ z(t) \end{bmatrix}$, $A = \begin{bmatrix} 1 & 2 & -1 \\ -1 & 5 & -2 \\ -1 & 4 & -1 \end{bmatrix}$, et $B = \begin{bmatrix} 0 & 2 & -1 \\ -1 & 4 & -2 \\ -1 & 4 & -2 \end{bmatrix}$. Le système s'écrit

$$X'(t) = AX(t) + \begin{bmatrix} 0 \\ 0 \\ t \end{bmatrix}.$$

On a $A = I + B$ et on a vu plus haut que l'on a, pour $p \geq 1$,

$$B^p = D + pN, \text{ où } D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & -1 \\ 0 & 2 & -1 \end{bmatrix} \text{ et } N = \begin{bmatrix} -1 & 2 & -1 \\ -1 & 2 & -1 \\ -1 & 2 & -1 \end{bmatrix}. \text{ On a donc}$$

$$e^{tB} = I + \sum_{p=1}^{+\infty} \frac{t^p}{p!} D + \sum_{p=1}^{+\infty} \frac{pt^p}{p!} N = I - D + e^t D + te^t N.$$

On obtient

$$e^{tA} = e^t e^{tB} = e^t \begin{bmatrix} 0 & 0 & 0 \\ 0 & -1 & 1 \\ 0 & -2 & 2 \end{bmatrix} + e^{2t} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & -1 \\ 0 & 2 & -1 \end{bmatrix} + te^{2t} \begin{bmatrix} -1 & 2 & -1 \\ -1 & 2 & -1 \\ -1 & 2 & -1 \end{bmatrix}.$$

La solution cherchée est donc donnée par la formule

$$\begin{bmatrix} x(t) \\ y(t) \\ z(t) \end{bmatrix} = e^{tA} \int_0^t e^{-sA} \begin{bmatrix} 0 \\ 0 \\ s \end{bmatrix} ds = \int_0^t e^{sA} \begin{bmatrix} 0 \\ 0 \\ t-s \end{bmatrix} ds.$$

On a

$$\begin{aligned} e^{sA} \begin{bmatrix} 0 \\ 0 \\ t-s \end{bmatrix} &= e^s \begin{bmatrix} 0 \\ t-s \\ 2t-2s \end{bmatrix} + e^{2s} \begin{bmatrix} 0 \\ s-t \\ s-t \end{bmatrix} + se^{2s} \begin{bmatrix} s-t \\ s-t \\ s-t \end{bmatrix} \\ &= (t-s)e^s \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix} + (s-t)e^{2s} \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} + s(s-t)e^{2s} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}. \end{aligned}$$

Par intégration par parties, on obtient

$$\int_0^t (t-s)e^s ds = [(t-s)e^s]_0^t + \int_0^t e^s ds = [(t-s)e^s]_0^t + [e^s]_0^t = -1 - t + e^t.$$

$$\int_0^t (s-t)e^{2s} ds = \left[\frac{(s-t)e^{2s}}{2} \right]_0^t - \int_0^t \frac{e^{2s}}{2} ds = \left[\frac{(s-t)e^{2s}}{2} \right]_0^t - \left[\frac{e^{2s}}{4} \right]_0^t = \frac{1}{4} + \frac{t}{2} - \frac{e^{2t}}{4}.$$

$$\begin{aligned} \int_0^t s(s-t)e^{2s} ds &= \left[\frac{s(s-t)e^{2s}}{2} \right]_0^t - \int_0^t se^{2s} ds + \frac{t}{2} \int_0^t e^{2s} ds \\ &= - \int_0^t (s-t)e^{2s} dt - \frac{t}{2} \int_0^t e^{2s} ds = -\frac{1}{4} - \frac{t}{2} + \frac{e^{2t}}{4} - \frac{te^{2t}}{4} + \frac{t}{4} = -\frac{1}{4} - \frac{t}{4} + \frac{e^{2t}}{4} - \frac{te^{2t}}{4} \end{aligned}$$

On obtient alors

$$\begin{cases} x(t) = -\frac{1}{4} - \frac{t}{4} + \frac{e^{2t}}{4} - \frac{te^{2t}}{4} \\ y(t) = -1 - t + e^t + \frac{1}{4} + \frac{t}{2} - \frac{e^{2t}}{4} - \frac{1}{4} - \frac{t}{4} + \frac{e^{2t}}{4} - \frac{te^{2t}}{4} = -1 - \frac{3t}{4} + e^t - \frac{te^{2t}}{4} \\ z(t) = -2 - 2t + 2e^t + \frac{1}{4} + \frac{t}{2} - \frac{e^{2t}}{4} - \frac{1}{4} - \frac{t}{4} + \frac{e^{2t}}{4} - \frac{te^{2t}}{4} = -2 - \frac{7t}{4} + 2e^t - \frac{te^{2t}}{4} \end{cases}$$

On verra dans [4] que l'on peut également résoudre les systèmes différentiels linéaires en utilisant la *transformée de Laplace*.

6.4 Chapitre 6 sous Mupad

Il n'existe pas de commande directe permettant de calculer sous Mupad les puissances d'une matrice dont on connaît les valeurs propres, et Mupad refuse d'entrer comme données initiales d'une suite récurrente des matrices (voir ci-dessous). Par contre Mupad accepte de résoudre les relations de récurrence scalaires associées au polynôme caractéristique d'une matrice. On peut calculer

ainsi $\begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}^n$ et $\begin{bmatrix} 0 & 2 & -1 \\ -1 & 4 & -2 \\ -1 & 4 & -2 \end{bmatrix}^n$ (on notera que Mupad ne simplifie pas le premier résultat).

En ce qui concerne les systèmes différentiels à coefficients constants, on dispose d'un moyen de les résoudre directement sous Mupad (pourvu que l'on puisse calculer les valeurs propres de la matrice associée au système), en utilisant la commande `solve(ode...)`. Par exemple pour résoudre le système différentiel de l'exemple 6.3.7 on peut utiliser la commande `solve(ode...)`, et résoudre directement le système, où utiliser la commande `exp(-x * A)`, pour calculer e^{-xA} ,

effectuer le produit $W = e^{-xA} \begin{bmatrix} 0 \\ 0 \\ x \end{bmatrix}$, puis la commande `map(W, int, x = 0..t)`,

pour calculer $\int_0^t e^{-xA} \begin{bmatrix} 0 \\ 0 \\ x \end{bmatrix} dx$. Il suffit alors de multiplier la matrice colonne

obtenue par e^{tA} pour obtenir la solution du système.

128 CHAPITRE 6. ITÉRATION, SYSTÈMES DIFFÉRENTIELS LINÉAIRES

```
M:=Dom::Matrix();
I1:=M([[1,0],[0,1]]);
A:=M([[0,1],[1,1]]);
solve((rec(x(n+2)=x(n+1)+x(n), x(n),{x(0)=I1,x(1)=A})));
```

Dom::Matrix()

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

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

$$\text{solve}\left(\text{rec}\left(-x(n) - x(n+1) + x(n+2), x(n), \left\{x(0) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, x(1) = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}\right\}\right)\right)$$

```
u:=solve((rec(x(n+2)=x(n+1)+x(n), x(n),{x(0)=a,x(1)=b})));
```

$$\left\{ \left(\frac{1}{2} - \frac{\sqrt{5}}{2}\right)^n \cdot \left(\frac{a}{2} + \frac{a \cdot \sqrt{5}}{10} - \frac{b \cdot \sqrt{5}}{5}\right) + \left(\frac{\sqrt{5}}{2} + \frac{1}{2}\right)^n \cdot \left(\frac{a}{2} - \frac{a \cdot \sqrt{5}}{10} + \frac{b \cdot \sqrt{5}}{5}\right) \right\}$$

```
u1:=1/2 +sqrt(5)/2;u2:=1/2 -sqrt(5)/2;v1:=1/2 +sqrt(5)/10;v2:=1/2 -sqrt(5)/10;
An:=(u2^n*v1+u1^n*v2)*I1 +(u1^n-u2^n)*(sqrt(5)/5)*A;
```

$$\frac{\sqrt{5}}{2} + \frac{1}{2}$$

$$\frac{1}{2} - \frac{\sqrt{5}}{2}$$

$$\frac{\sqrt{5}}{10} + \frac{1}{2}$$

$$\frac{1}{2} - \frac{\sqrt{5}}{10}$$

$$\left(\left(-\frac{\sqrt{5}}{10} + \frac{1}{2}\right) \cdot \frac{\left(\frac{\sqrt{5}}{2} + \frac{1}{2}\right)^n + \left(\frac{\sqrt{5}}{10} + \frac{1}{2}\right) \cdot \left(-\frac{\sqrt{5}}{2} + \frac{1}{2}\right)^n}{\sqrt{5} \cdot \left(-\frac{\sqrt{5}}{2} + \frac{1}{2}\right)^n + \left(\frac{\sqrt{5}}{2} + \frac{1}{2}\right)^n} \quad \frac{\sqrt{5} \cdot \left(-\frac{\sqrt{5}}{2} + \frac{1}{2}\right)^n + \left(\frac{\sqrt{5}}{2} + \frac{1}{2}\right)^n}{5} \right. \\ \left. \left(-\frac{\sqrt{5}}{10} + \frac{1}{2}\right) \cdot \left(\frac{\sqrt{5}}{2} + \frac{1}{2}\right)^n + \left(\frac{\sqrt{5}}{10} + \frac{1}{2}\right) \cdot \left(-\frac{\sqrt{5}}{2} + \frac{1}{2}\right)^n + \frac{\sqrt{5} \cdot \left(-\frac{\sqrt{5}}{2} + \frac{1}{2}\right)^n + \left(\frac{\sqrt{5}}{2} + \frac{1}{2}\right)^n}{5} \right)$$

```
B:=M([[0,2,-1],[-1,4,-2],[-1,4,-2]]);
v:=solve((rec(x(n+2)=2*x(n+1)-x(n), x(n),{x(1)=a,x(2)=b})));
```

$$\begin{pmatrix} 0 & 2 & -1 \\ -1 & 4 & -2 \\ -1 & 4 & -2 \end{pmatrix}$$

$$\{2 \cdot a - b + n \cdot (b - a)\}$$

```
Bn:=2*B-B^2 +n*B^2-n*B;
```

$$\begin{pmatrix} -n+1 & 2 \cdot n & -n \\ -n & 2 \cdot n+2 & -n-1 \\ -n & 2 \cdot n+2 & -n-1 \end{pmatrix}$$

```
M:=Dom::Matrix();
A:=M([[1,2,-1],[-1,5,-2],[-1,4,-1]]);
solve(ode({x'(t)=x(t)+2*y(t)-z(t),y'(t)=-x(t)+5*y(t)-2*z(t),z'(t)=-x(t)+4*y(t)-z(t)+t,x(0)=0,y(0)=0,z(0)=0},{x(t),y(t),z(t)}));
```

```
Dom::Matrix()
```

$$\begin{pmatrix} 1 & 2 & -1 \\ -1 & 5 & -2 \\ -1 & 4 & -1 \end{pmatrix}$$

$$\left\{ \left[\begin{array}{l} y(t) = e^t - \frac{3 \cdot t}{4} - \frac{t \cdot e^{-2 \cdot t}}{4} - 1, z(t) = 2 \cdot e^t - \frac{7 \cdot t}{4} - \frac{t \cdot e^{-2 \cdot t}}{4} - 2, x(t) = \frac{e^{2 \cdot t}}{4} - \frac{t}{4} - \frac{t \cdot e^{-2 \cdot t}}{4} - \frac{1}{4} \end{array} \right] \right\}$$

```
U:=exp(-x*A);
simplify(U);
V:=M([[0],[0],[x]]);
W:=U*V;
simplify(W);
```

$$\begin{pmatrix} 2 \cdot e^{-2 \cdot x} - x \cdot \left(-e^{-2 \cdot x} + \frac{e^{-2 \cdot x}}{x} \right) & -4 \cdot e^{-2 \cdot x} - x \cdot \left(2 \cdot e^{-2 \cdot x} - \frac{4 \cdot e^{-2 \cdot x}}{x} \right) & 2 \cdot e^{-2 \cdot x} - x \cdot \left(-e^{-2 \cdot x} + \frac{2 \cdot e^{-2 \cdot x}}{x} \right) \\ e^{-2 \cdot x} - x \cdot \left(-e^{-2 \cdot x} + \frac{e^{-2 \cdot x}}{x} \right) & -e^{-x} - 2 \cdot e^{-2 \cdot x} - x \cdot \left(2 \cdot e^{-2 \cdot x} - \frac{4 \cdot e^{-2 \cdot x}}{x} \right) & e^{-x} + e^{-2 \cdot x} - x \cdot \left(-e^{-2 \cdot x} + \frac{2 \cdot e^{-2 \cdot x}}{x} \right) \\ e^{-2 \cdot x} - x \cdot \left(-e^{-2 \cdot x} + \frac{e^{-2 \cdot x}}{x} \right) & -2 \cdot e^{-x} - 2 \cdot e^{-2 \cdot x} - x \cdot \left(2 \cdot e^{-2 \cdot x} - \frac{4 \cdot e^{-2 \cdot x}}{x} \right) & 2 \cdot e^{-x} + e^{-2 \cdot x} - x \cdot \left(-e^{-2 \cdot x} + \frac{2 \cdot e^{-2 \cdot x}}{x} \right) \end{pmatrix}$$

$$\begin{pmatrix} e^{-2 \cdot x} + x \cdot e^{-2 \cdot x} & -2 \cdot x \cdot e^{-2 \cdot x} & x \cdot e^{-2 \cdot x} \\ x \cdot e^{-2 \cdot x} & -e^{-x} + 2 \cdot e^{-2 \cdot x} - 2 \cdot x \cdot e^{-2 \cdot x} & e^{-x} - e^{-2 \cdot x} + x \cdot e^{-2 \cdot x} \\ x \cdot e^{-2 \cdot x} & -2 \cdot e^{-x} + 2 \cdot e^{-2 \cdot x} - 2 \cdot x \cdot e^{-2 \cdot x} & 2 \cdot e^{-x} - e^{-2 \cdot x} + x \cdot e^{-2 \cdot x} \end{pmatrix}$$

$$\begin{pmatrix} 0 \\ 0 \\ x \end{pmatrix}$$

$$\begin{pmatrix} x \cdot \left(2 \cdot e^{-2 \cdot x} - x \cdot \left(-e^{-2 \cdot x} + \frac{2 \cdot e^{-2 \cdot x}}{x} \right) \right) \\ x \cdot \left(e^{-x} + e^{-2 \cdot x} - x \cdot \left(-e^{-2 \cdot x} + \frac{2 \cdot e^{-2 \cdot x}}{x} \right) \right) \\ x \cdot \left(2 \cdot e^{-x} + e^{-2 \cdot x} - x \cdot \left(-e^{-2 \cdot x} + \frac{2 \cdot e^{-2 \cdot x}}{x} \right) \right) \end{pmatrix}$$

$$\begin{pmatrix} x^2 \cdot e^{-2 \cdot x} \\ x \cdot e^{-x} - x \cdot e^{-2 \cdot x} + x^2 \cdot e^{-2 \cdot x} \\ 2 \cdot x \cdot e^{-x} - x \cdot e^{-2 \cdot x} + x^2 \cdot e^{-2 \cdot x} \end{pmatrix}$$

```
H:=map(W,int,x=0..t);
```

$$\begin{pmatrix} -\frac{t}{2} \frac{t^2}{e^t} + \frac{1}{4} \\ -\frac{t^2}{2 \cdot (e^t)^2} + \frac{-t-1}{e^t} + 1 \\ -\frac{t^2}{2 \cdot (e^t)^2} + \frac{-2 \cdot t-2}{e^t} + 2 \end{pmatrix}$$

```
X:=exp(t*A)*H;
simplify(X)
```

```
array(1..3, 1..1,
(1, 1) = (2*exp(2*t) + t*(- exp(2*t) - 2/t*exp(2*t)))*(- 1/2*t^2/exp(t)^2 + 1/exp(t)*(- 2*t - 2) + 2) + (- 4*exp(2*t) + t*(2*exp(2*t) + 4/t*exp(2*t)))*(- 1/2*t^2/exp(t)^2 + 1/exp(t)*(- t - 1) + 1) + (2*exp(2*t) + t*(- exp(2*t) - 1/t*exp(2*t)))*(1/exp(t)^2*(- 1/2*t - 1/2*t^2 - 1/4) + 1/4),
(2, 1) = (exp(t) + exp(2*t) + t*(- exp(2*t) - 2/t*exp(2*t)))*(- 1/2*t^2/\exp(t)^2 + 1/exp(t)*(- 2*t - 2) + 2) + (- exp(t) - 2*exp(2*t) + t*(2*exp(2*t) + 4/t*exp(2*t)))*(- 1/2*t^2/exp(t)^2 + 1/exp(t)*(- t - 1) + 1) + (exp(\2*t) + t*(- exp(2*t) - 1/t*exp(2*t)))*(1/exp(t)^2*(- 1/2*t - 1/2*t^2 - 1/4) + 1/4),
(3, 1) = (2*exp(t) + exp(2*t) + t*(- exp(2*t) - 2/t*exp(2*t)))*(- 1/2*t^2/\2/exp(t)^2 + 1/exp(t)*(- 2*t - 2) + 2) + (- 2*exp(t) - 2*exp(2*t) + t*(2*exp(2*t) + 4/t*exp(2*t)))*(- 1/2*t^2/exp(t)^2 + 1/exp(t)*(- t - 1) + 1) + (\exp(2*t) + t*(- exp(2*t) - 1/t*exp(2*t)))*(1/exp(t)^2*(- 1/2*t - 1/2*t^2 - 1/4) + 1/4)
```

$$\begin{pmatrix} -\frac{t}{4} + \frac{e^{2 \cdot t}}{4} - \frac{t \cdot e^{-2 \cdot t}}{4} - \frac{1}{4} \\ -\frac{3 \cdot t}{4} + e^t - \frac{t \cdot e^{-2 \cdot t}}{4} - 1 \\ -\frac{7 \cdot t}{4} + 2 \cdot e^t - \frac{t \cdot e^{-2 \cdot t}}{4} - 2 \end{pmatrix}$$

6.5 Exercices sur le Chapitre 6

exercice 1

Soit $A = \begin{pmatrix} -6 & 1 \\ -1 & -8 \end{pmatrix}$.

1) Déterminer la décomposition de Jordan de A et calculer A^n pour $n \geq 0$.

2) Calculer e^{tA} pour $t \in \mathbf{R}$.

3) Résoudre le système différentiel

$$\begin{cases} x'(t) = -6x(t) - y(t) + 1 \\ y'(t) = x(t) - 8y(t) \end{cases},$$

avec la condition initiale $x(0) = 0, y(0) = 1$.

exercice 2

Soit $A = \begin{pmatrix} 2 & 1 & 1 \\ 2 & 1 & -2 \\ -1 & 0 & -2 \end{pmatrix}$.

1) Vérifier que $p_A = -(X - 3)(X + 1)^2$, où p_A désigne le polynôme caractéristique de A .

2) Calculer $(A - 3I)(A + I)$. En déduire que A n'est pas diagonalisable.

3) Déterminer la décomposition de Jordan de A et calculer A^n pour $n \geq 0$.

4) Interpréter le résultat obtenu en appliquant la formule de la question précédente pour $n < 0$.

exercice 3

Très déçu par les résultats de l'élection présidentielle américaine, le milliardaire G. Soros décide de placer 10^8 euros en France, en Espagne et en Allemagne. Chaque année il modifie son placement de la façon suivante : si x_n est la partie du capital placée en France à l'année n , y_n la partie du capital placée en Espagne à l'année n et z_n la partie du capital placée en Allemagne à l'année n , on a

$$x_{n+1} = \frac{1}{2}(x_n + y_n + z_n)$$

$$y_{n+1} = \frac{1}{4}x_n + \frac{1}{2}y_n$$

$$z_{n+1} = \frac{1}{4}x_n + \frac{1}{2}z_n.$$

On a évidemment $x_0 \geq 0$, $y_0 \geq 0$, $z_0 \geq 0$, et $x_0 + y_0 + z_0 = 10^8$.

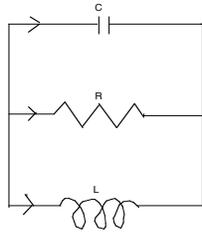
1) Vérifier que $x_n + y_n + z_n = 10^8$ pour tout $n \geq 0$.

2) Montrer que les trois suites $(x_n)_{n \geq 0}$, $(y_n)_{n \geq 0}$ et $(z_n)_{n \geq 0}$ sont convergentes, et déterminer leurs limites.

exercice 4 (Archives ESTIA, examen d'avril 2003)

On pose $A = \begin{bmatrix} 0 & 1 \\ -4 & -4 \end{bmatrix}$.

1) Déterminer le polynôme caractéristique et le polynôme minimal de A .



Circuit électrique 1

1.pdf

2) Déterminer la décomposition de Jordan de A et calculer A^n pour $n \geq 0$.

3) Calculer e^{tA} pour $t \in \mathbf{R}$.

4) On considère l'équation différentielle

$$y''(x) + 4y'(x) + 4y(x) = 4x.$$

avec la condition initiale $y(0) = y'(0) = 0$.

a) On pose $Z(x) = \begin{bmatrix} y(x) \\ y'(x) \end{bmatrix}$. Montrer que $Z(x)$ est solution du système différentiel

$$Z'(x) = AZ(x) + \begin{bmatrix} 0 \\ 4x \end{bmatrix},$$

avec $Z(0) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$.

b) Exprimer $Z(x)$ au moyen de la formule intégrale du cours.

c) Calculer $Z(x)$ par la méthode de votre choix et en déduire la solution de l'équation différentielle.

exercice 5

On considère le circuit électrique suivant (circuit 1) où la résistance R est mesurée en ohms, la capacité du condensateur C est mesurée en farads, et le coefficient d'induction du solénoïde L est mesuré en henrys. On rappelle les deux lois de Kirchhoff :

- 1) La perte de voltage dans tout sous-circuit est nulle
- 2) La somme des intensités des courants en chaque noeud est nulle.

On note V la perte de voltage au travers du condensateur, I l'intensité du courant au travers de L . On note I_1 l'intensité du courant au niveau du condensateur, I_2 l'intensité du courant au niveau de la résistance, I l'intensité du courant au travers du solénoïde, V la perte de voltage au niveau du condensateur, V_2 la perte de voltage au niveau de la résistance et V_3 la perte de voltage au travers du solénoïde. Les voltages sont exprimés en volts, et les intensités en ampères. On rappelle que l'on a

$$V_2(t) = RI_2(t), \quad CV_1'(t) = I_1(t), \quad LI'(t) = V_3(t).$$

- 1) En utilisant les lois de physique ci-dessus, montrer que $I_1(t) + I_2(t) + I(t) = 0$, que $V(t) = V_2(t) = V_3(t)$, et que l'on a

$$\begin{bmatrix} I'(t) \\ V'(t) \end{bmatrix} = \begin{bmatrix} 0 & \frac{1}{L} \\ -\frac{1}{C} & -\frac{1}{RC} \end{bmatrix} \begin{bmatrix} I(t) \\ V(t) \end{bmatrix} \quad (6.15)$$

- 2) Vérifier que les deux racines de la matrice ci-dessus sont réelles et distinctes si $L > 4R^2C$, et qu'elles sont imaginaires conjuguées si $L < 4R^2C$.

3) On suppose que $R = 1$ ohm, $C = 2$ farad, et $L = 1$ henry. Déterminer la solution générale du système, et déterminer la solution correspondant aux conditions initiales $I(0) = 2$ ampères et $V(0) = 1$ volt.

4) Les hypothèses étant celles de la question 2 étudier le comportement de $I(t)$ et $V(t)$ quand $t \rightarrow +\infty$. Les limites obtenues dépendent-elles de $I(0)$ et $V(0)$?

exercice 6

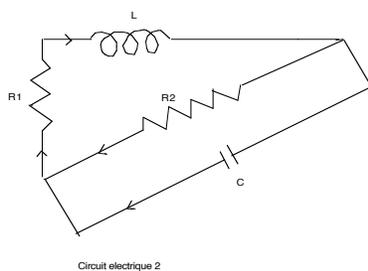
On considère maintenant le circuit électrique 2, et on note I l'intensité du courant au travers de la résistance R_1 et V la perte de voltage au travers du condensateur.

- 1) En s'inspirant de l'exercice précédent, montrer que l'on a

$$\begin{bmatrix} I'(t) \\ V'(t) \end{bmatrix} = \begin{bmatrix} -\frac{R_1}{L} & -\frac{1}{L} \\ \frac{1}{C} & -\frac{1}{CR_2} \end{bmatrix} \begin{bmatrix} I(t) \\ V(t) \end{bmatrix} \quad (6.16)$$

2) Donner la solution générale de ce système quand $R_1 = R_2 = 1$ ohm, $L = 1$ henry et $C = 2$ farad, ainsi que la solution correspondant aux conditions initiales $I(0) = 2$ ampères et $V(0) = 5$ volts.

3) Discuter en fonction des données et des conditions initiales le comportement de $I(t)$ et $V(t)$ quand $t \rightarrow +\infty$.

**exercice 7**

Reprendre quand c'est possible sous Mupad les calculs demandés dans les exercices précédents.

Chapitre 7

Espaces vectoriels euclidiens

7.1 Produit scalaire, inégalité de Cauchy-Schwartz

Tous les espaces vectoriels considérés dans ce Chapitre sont des espaces vectoriels sur \mathbf{R} .

Définition 7.1.1 Soit E un espace vectoriel sur \mathbf{R} .

1) On dit qu'une application $\phi : E \times E \longrightarrow \mathbf{R}$ est une forme bilinéaire symétrique quand les deux conditions suivantes sont vérifiées

$$(i) \phi(\mathbf{u}, \mathbf{v}) = \phi(\mathbf{v}, \mathbf{u}) \quad \forall \mathbf{u} \in E, \forall \mathbf{v} \in E,$$

$$(ii) \phi(\lambda_1 \mathbf{u}_1 + \lambda_2 \mathbf{u}_2, \mathbf{v}) = \lambda_1 \phi(\mathbf{u}_1, \mathbf{v}) + \lambda_2 \phi(\mathbf{u}_2, \mathbf{v}) \quad \forall \lambda_1 \in \mathbf{R}, \forall \lambda_2 \in \mathbf{R}, \forall \mathbf{u}_1 \in E, \forall \mathbf{u}_2 \in E, \forall \mathbf{v} \in E.$$

2) On dit qu'une forme bilinéaire symétrique ϕ est un produit scalaire sur E (ou que ϕ est définie positive) si $\phi(\mathbf{x}, \mathbf{x}) > 0$ pour $\mathbf{x} \in E, \mathbf{x} \neq 0$.

3) On dit qu'une application $\mathbf{x} \longmapsto \|\mathbf{x}\|$ de E dans \mathbf{R} est une norme sur E quand les trois conditions suivantes sont vérifiées

$$(i) \|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\| \quad \forall \mathbf{x} \in E, \forall \mathbf{y} \in E.$$

$$(ii) \|\lambda \mathbf{x}\| = |\lambda| \|\mathbf{x}\| \quad \forall \lambda \in \mathbf{R}, \forall \mathbf{x} \in E,$$

$$(iii) \|\mathbf{x}\| > 0 \text{ pour } \mathbf{x} \in E, \mathbf{x} \neq 0.$$

L'ensemble des formes bilinéaires symétriques sur E sera noté $L_2(E)$.

Exemple 7.1.2 1) La formule $\theta(f, g) = \int_0^{1/2} f(t)g(t)dt - \int_{1/2}^1 f(t)g(t)dt$ définit une forme bilinéaire symétrique sur l'espace vectoriel $\mathcal{C}(0, 1)$ des fonctions continues sur $[0, 1]$. Ce n'est pas un produit scalaire car $\theta(1, 1) = 0$.

2) Les produits scalaires du lycée dans le plan vectoriel ou l'espace vectoriel de dimension 3 usuels sont des produits scalaires.

3) Pour $X = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \in \mathbf{R}^3$, $Y = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} \in \mathbf{R}^3$, posons $\langle X, Y \rangle = {}^t X \cdot Y =$

$\sum_{1 \leq i \leq n} x_i y_i$. On obtient un produit scalaire sur \mathbf{R}^3 , appelé le produit scalaire canonique de \mathbf{R}^3 (notons que l'on identifie la matrice ${}^t X \cdot Y \in \mathcal{M}_{1,1}(\mathbf{R})$ à son unique coefficient).

4) La formule $\phi(f, g) = \int_a^b f(t)g(t)dt$ définit un produit scalaire sur l'espace vectoriel $\mathcal{C}(a, b)$ des fonctions continues sur un intervalle fermé borné $[a, b]$.

5) La valeur absolue est une norme sur \mathbf{R} , et le module est une norme sur \mathbf{C} (considéré comme espace vectoriel sur \mathbf{R}).

6) Les formules $\|(x, y, z)\|_1 = |x| + |y| + |z|$ et $\|(x, y, z)\|_\infty = \sup(|x|, |y|, |z|)$ définissent des normes sur \mathbf{R}^3 .

7) La formule $\|f\|_\infty = \max_{a \leq t \leq b} |f(t)|$ définit une norme sur $\mathcal{C}(a, b)$.

Notons que si ϕ est une forme bilinéaire symétrique sur E alors l'application $\phi_{\mathbf{x}} : \mathbf{y} \mapsto \phi(\mathbf{x}, \mathbf{y})$ est une application linéaire de E dans \mathbf{R} pour tout $\mathbf{x} \in E$. On voit par symétrie que l'application $\phi^{\mathbf{y}} : \mathbf{x} \mapsto \phi(\mathbf{x}, \mathbf{y})$ est linéaire pour tout $\mathbf{y} \in E$. On obtient alors, pour $\lambda_1, \dots, \lambda_m \in \mathbf{R}$, $\mu_1, \dots, \mu_n \in \mathbf{R}$, $\mathbf{x}_1, \dots, \mathbf{x}_m \in E$, $\mathbf{y}_1, \dots, \mathbf{y}_n \in E$

$$(7.1) \quad \phi\left(\sum_{i=1}^m \lambda_i \mathbf{x}_i, \sum_{j=1}^n \mu_j \mathbf{y}_j\right) = \sum_{i=1}^m \sum_{j=1}^n \lambda_i \mu_j \phi(\mathbf{x}_i, \mathbf{y}_j).$$

On va maintenant introduire l'inégalité de Cauchy-Schwartz.

Théorème 7.1.3 Soit E un espace vectoriel sur \mathbf{R} muni d'un produit scalaire $(\mathbf{u}, \mathbf{v}) \mapsto \langle \mathbf{u}, \mathbf{v} \rangle$. Pour tout couple (\mathbf{u}, \mathbf{v}) d'éléments de E , on a l'inégalité

$$|\langle \mathbf{u}, \mathbf{v} \rangle| \leq \sqrt{\langle \mathbf{u}, \mathbf{u} \rangle} \sqrt{\langle \mathbf{v}, \mathbf{v} \rangle}$$

Démonstration : Si $\mathbf{v} = 0$ alors $\langle \mathbf{u}, \mathbf{v} \rangle = 0 = \langle \mathbf{v}, \mathbf{v} \rangle$, et l'inégalité est vérifiée. Supposons maintenant que $\mathbf{v} \neq 0$. Pour $x \in \mathbf{R}$, on a

$$\begin{aligned} 0 &\leq \langle \mathbf{u} + x\mathbf{v}, \mathbf{u} + x\mathbf{v} \rangle = \langle \mathbf{u}, \mathbf{u} \rangle + x \langle \mathbf{v}, \mathbf{u} \rangle + x \langle \mathbf{u}, \mathbf{v} \rangle + x^2 \langle \mathbf{v}, \mathbf{v} \rangle \\ &= \langle \mathbf{u}, \mathbf{u} \rangle + 2x \langle \mathbf{u}, \mathbf{v} \rangle + x^2 \langle \mathbf{v}, \mathbf{v} \rangle. \end{aligned}$$

On reconnaît un trinôme de signe constant, dont le discriminant est nécessairement négatif ou nul. Ceci donne $4 \langle \mathbf{u}, \mathbf{v} \rangle^2 \leq 4 \langle \mathbf{u}, \mathbf{u} \rangle \langle \mathbf{v}, \mathbf{v} \rangle$, et en simplifiant par 4 et en passant aux racines carrées on obtient l'inégalité cherchée.

♣

En appliquant cette inégalité au produit scalaire introduit à l'exemple 7.2 (4) on obtient une inégalité classique du calcul intégral.

Corollaire 7.1.4 Soit $I = [a, b]$ un intervalle fermé borné de \mathbf{R} , et soient f et g deux fonctions continues sur I à valeurs réelles. On a alors

$$\left| \int_a^b f(t)g(t)dt \right| \leq \left[\int_a^b f^2(t)dt \right]^{1/2} \left[\int_a^b g^2(t)dt \right]^{1/2}.$$

Dans toute la suite on appellera *espace euclidien* un espace vectoriel réel muni d'un produit scalaire. On a l'important résultat suivant.

Corollaire 7.1.5 Soit E un espace euclidien. L'application $\mathbf{u} \mapsto \|\mathbf{u}\| = \sqrt{\langle \mathbf{u}, \mathbf{u} \rangle}$ est une norme sur E , appelée *norme euclidienne*.

Démonstration : On a évidemment $\|\mathbf{u}\| > 0$ pour $\mathbf{u} \in E \setminus \{0\}$, et $\|\lambda\mathbf{u}\| = |\lambda|\|\mathbf{u}\|$ pour $\lambda \in \mathbf{R}$, $\mathbf{u} \in E$. Soient $\mathbf{u}, \mathbf{v} \in E$. On a $\|\mathbf{u} + \mathbf{v}\|^2 = \langle \mathbf{u} + \mathbf{v}, \mathbf{u} + \mathbf{v} \rangle = \langle \mathbf{u}, \mathbf{u} \rangle + 2\langle \mathbf{u}, \mathbf{v} \rangle + \langle \mathbf{v}, \mathbf{v} \rangle \leq \langle \mathbf{u}, \mathbf{u} \rangle + 2|\langle \mathbf{u}, \mathbf{v} \rangle| + \langle \mathbf{v}, \mathbf{v} \rangle = \langle \mathbf{u}, \mathbf{u} \rangle + 2\sqrt{\langle \mathbf{u}, \mathbf{u} \rangle}\sqrt{\langle \mathbf{v}, \mathbf{v} \rangle} + \langle \mathbf{v}, \mathbf{v} \rangle = (\sqrt{\langle \mathbf{u}, \mathbf{u} \rangle} + \sqrt{\langle \mathbf{v}, \mathbf{v} \rangle})^2 = (\|\mathbf{u}\| + \|\mathbf{v}\|)^2$.

Donc $\|\mathbf{u} + \mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|$, et les trois conditions de la définition 7.1 (3) sont vérifiées. ♣

Un exemple important de norme euclidienne est donné par la norme $\|\cdot\|_2$ associée au produit scalaire canonique sur \mathbf{R}^k , qui est donnée par la formule

$$\|(x_1, \dots, x_k)\|_2 = \sum_{j=1}^k x_j^2. \quad (7.1)$$

7.2 Procédé d'orthonormalisation de Gram-Schmidt

Dans toute la suite on munit les espaces euclidiens de leur norme euclidienne.

Définition 7.2.1 On dit que deux éléments \mathbf{u} et \mathbf{v} d'un espace euclidien E sont orthogonaux, et on écrit $\mathbf{u} \perp \mathbf{v}$ quand $\langle \mathbf{u}, \mathbf{v} \rangle = 0$. On dit qu'une base $\mathcal{B} = (\mathbf{e}_1, \dots, \mathbf{e}_n)$ de E est orthogonale si ses éléments sont orthogonaux deux à deux, et on dit qu'une base de E est orthonormale si elle est orthogonale et si $\|\mathbf{e}_i\| = 1$ pour $1 \leq i \leq n$.

On a le "théorème de Pythagore".

Proposition 7.2.2 Soit $\mathbf{u}_1, \dots, \mathbf{u}_n$ une famille finie d'éléments d'un espace euclidien E orthogonaux deux à deux. On a alors

$$\left\| \sum_{i=1}^n \mathbf{u}_i \right\|^2 = \sum_{i=1}^n \|\mathbf{u}_i\|^2.$$

Démonstration : Comme $\mathbf{u}_i \perp \mathbf{u}_j$ pour $j \neq i$, on a d'après la formule 7.1

$$\left\| \sum_{i=1}^n \mathbf{u}_i \right\|^2 = \sum_{i=1}^n \sum_{j=1}^n \langle \mathbf{u}_i, \mathbf{u}_j \rangle = \sum_{i=1}^n \langle \mathbf{u}_i, \mathbf{u}_i \rangle = \sum_{i=1}^n \|\mathbf{u}_i\|^2. \clubsuit$$

Notons également que l'on a l' "identité du parallélogramme".

Proposition 7.2.3 Soit E un espace euclidien, et soient $\mathbf{u}, \mathbf{v} \in E$. On a

$$\|\mathbf{u} + \mathbf{v}\|^2 + \|\mathbf{u} - \mathbf{v}\|^2 = 2(\|\mathbf{u}\|^2 + \|\mathbf{v}\|^2). \quad (7.2)$$

Démonstration : On a $\|\mathbf{u} + \mathbf{v}\|^2 + \|\mathbf{u} - \mathbf{v}\|^2 = \langle \mathbf{u} + \mathbf{v}, \mathbf{u} + \mathbf{v} \rangle + \langle \mathbf{u} - \mathbf{v}, \mathbf{u} - \mathbf{v} \rangle = \langle \mathbf{u}, \mathbf{u} \rangle + 2\langle \mathbf{u}, \mathbf{v} \rangle + \langle \mathbf{v}, \mathbf{v} \rangle + \langle \mathbf{u}, \mathbf{u} \rangle - 2\langle \mathbf{u}, \mathbf{v} \rangle + \langle \mathbf{v}, \mathbf{v} \rangle = 2(\|\mathbf{u}\|^2 + \|\mathbf{v}\|^2). \clubsuit$

On va maintenant donner un procédé concret permettant de fabriquer des bases orthonormales.

Théorème 7.2.4 (Algorithme de Gram-Schmidt)

Soit E un espace euclidien, soit $\mathbf{e}_1, \dots, \mathbf{e}_p$ une famille libre d'éléments de E , et soit F le sous-espace vectoriel de E engendré par $(\mathbf{e}_1, \dots, \mathbf{e}_p)$. On pose $\mathbf{f}_1 = \mathbf{e}_1$ et $\mathbf{f}_k = \mathbf{e}_k - \sum_{1 \leq j \leq k-1} \frac{\langle \mathbf{e}_k, \mathbf{f}_j \rangle}{\|\mathbf{f}_j\|^2} \mathbf{f}_j$ pour $2 \leq k \leq p$. Alors $(\mathbf{f}_1, \dots, \mathbf{f}_p)$ est une base orthogonale de F , de sorte que $(\frac{\mathbf{f}_1}{\|\mathbf{f}_1\|}, \dots, \frac{\mathbf{f}_p}{\|\mathbf{f}_p\|})$ est une base orthonormale de F .

On va démontrer par récurrence finie que la famille $(\mathbf{f}_1, \dots, \mathbf{f}_k)$ est orthogonale, que $\mathbf{f}_k \neq 0$, et que \mathbf{f}_k appartient au sous-espace vectoriel de E engendré par $(\mathbf{e}_1, \dots, \mathbf{e}_k)$ pour $1 \leq k \leq p$. C'est évident si $k = 1$. Supposons maintenant que ces propriétés sont vérifiées pour un entier k tel que $1 \leq k \leq p - 1$. Il est clair que $\mathbf{f}_{k+1} \neq 0$ et que \mathbf{f}_{k+1} appartient alors au sous-espace vectoriel de E engendré par $(\mathbf{e}_1, \dots, \mathbf{e}_{k+1})$. D'autre part on a, pour $i \leq k$,

$$\begin{aligned} \langle \mathbf{f}_{k+1}, \mathbf{f}_i \rangle &= \langle \mathbf{e}_{k+1}, \mathbf{f}_i \rangle - \left\langle \sum_{1 \leq j \leq k} \frac{\langle \mathbf{e}_{k+1}, \mathbf{f}_j \rangle}{\|\mathbf{f}_j\|^2} \mathbf{f}_j, \mathbf{f}_i \right\rangle = \langle \mathbf{e}_{k+1}, \mathbf{f}_i \rangle \\ &- \sum_{1 \leq j \leq k} \langle \mathbf{e}_{k+1}, \mathbf{f}_j \rangle \frac{\langle \mathbf{f}_j, \mathbf{f}_i \rangle}{\|\mathbf{f}_j\|^2} = \langle \mathbf{e}_{k+1}, \mathbf{f}_i \rangle - \langle \mathbf{e}_{k+1}, \mathbf{f}_i \rangle = 0. \end{aligned}$$

Donc $(\mathbf{f}_1, \dots, \mathbf{f}_{k+1})$ est orthogonale, ce qui prouve les deux propriétés par récurrence finie. On voit donc que $(\mathbf{f}_1, \dots, \mathbf{f}_p)$ est une famille orthogonale d'éléments de F .

On a $\mathbf{f}_k = \mathbf{e}_k - \sum_{1 \leq j \leq k-1} \alpha_j \mathbf{e}_j$, avec $\alpha_1, \dots, \alpha_{j-1} \in \mathbf{R}$. On voit donc que la matrice P représentant $(\mathbf{f}_1, \dots, \mathbf{f}_p)$ dans la base $(\mathbf{e}_1, \dots, \mathbf{e}_p)$ de F est une matrice triangulaire supérieure dont tous les termes diagonaux sont égaux à 1. Donc $\det(P) = 1$, P est inversible et il résulte de la Proposition 2.10 que $(\mathbf{f}_1, \dots, \mathbf{f}_p)$ est une base de F . Donc $(\frac{\mathbf{f}_1}{\|\mathbf{f}_1\|}, \dots, \frac{\mathbf{f}_p}{\|\mathbf{f}_p\|})$ est une base orthonormale de F . \clubsuit

Exemple 7.2.5 On munit l'espace vectoriel E des polynômes de degré inférieur ou égal à 2 de la base $\mathcal{B}_0 = \{1, x, x^2\}$. En appliquant le procédé d'orthogonalisation de Gram-Schmidt la base \mathcal{B}_0 , trouver une base de E orthonormée pour le produit scalaire défini pour $f, g \in E$ par la formule

$$\langle f, g \rangle = \int_0^1 f(t)g(t)dt.$$

Posons $e_0 = 1$, $e_1 = x$ et $e_2 = x^2$. On a $f_0 = e_0$, et $\|f_0\|^2 = \int_0^1 dt = 1$. D'autre part on a

$$\langle e_1, f_0 \rangle = \int_0^1 t dt = \frac{1}{2},$$

donc $f_1 = e_1 - \frac{\langle e_1, f_0 \rangle}{\|f_0\|^2} f_0 = x - \frac{1}{2}$. On a alors

$$\|f_1\|^2 = \int_0^1 \left(t^2 - t + \frac{1}{4}\right) dt = \frac{1}{3} - \frac{1}{2} + \frac{1}{4} = \frac{1}{12} = \left(\frac{1}{2\sqrt{3}}\right)^2,$$

$$\langle e_2, f_0 \rangle = \int_0^1 t^2 dt = \frac{1}{3}, \quad \langle e_2, f_1 \rangle = \int_0^1 \left(t^3 - \frac{t^2}{2}\right) dt = \frac{1}{4} - \frac{1}{6} = \frac{1}{12}.$$

On obtient

$$f_2 = e_2 - \frac{\langle e_2, f_0 \rangle}{\|f_0\|^2} f_0 - \frac{\langle e_2, f_1 \rangle}{\|f_1\|^2} f_1 = x^2 - \frac{1}{3} - x + \frac{1}{2} = x^2 - x + \frac{1}{6}.$$

On a alors

$$\begin{aligned} \|f_2\|^2 &= \int_0^1 \left(t^2 - t + \frac{1}{6}\right)^2 dt = \int_0^1 \left(t^4 + t^2 + \frac{1}{36} - 2t^3 + \frac{t^2}{3} - \frac{t}{3}\right) dt \\ &= \frac{1}{5} + \frac{1}{3} + \frac{1}{36} - \frac{1}{2} + \frac{1}{9} - \frac{1}{6} = \frac{36 + 60 + 5 - 90 + 20 - 30}{180} = \frac{1}{180} = \left(\frac{1}{6\sqrt{5}}\right)^2. \end{aligned}$$

La base orthonormale cherchée est donc

$$\mathcal{B} = \left\{ \frac{f_0}{\|f_0\|}, \frac{f_1}{\|f_1\|}, \frac{f_2}{\|f_2\|} \right\} = \left\{ 1, \sqrt{3}(2x - 1), \sqrt{5}(6x^2 - 6x + 1) \right\}.$$

7.3 Projections orthogonales

Définition 7.3.1 Soit $(E, \langle \cdot, \cdot \rangle)$ un espace vectoriel euclidien, et soit A une partie non vide de E .

On pose $A^\perp := \{\mathbf{x} \in E \mid \langle \mathbf{x}, \mathbf{y} \rangle = 0 \ \forall \mathbf{y} \in A\}$.

Autrement dit A^\perp est l'ensemble des $\mathbf{y} \in E$ qui sont orthogonaux à tous les éléments de A . L'ensemble A^\perp est appelé l'orthogonal de A .

Proposition 7.3.2 1) $\{0\}^\perp = E$, et $E^\perp = \{0\}$.

2) A^\perp est un sous-espace vectoriel de E , $A \subset [A^\perp]^\perp$, et ou bien $A \cap A^\perp = \emptyset$, ou bien $A \cap A^\perp = \{0\}$.

Démonstration : 1) Il est clair que $\{0\}^\perp = E$, et que $0 \in E^\perp$. D'autre part si $\mathbf{x} \in E^\perp$, alors $\mathbf{x} \perp \mathbf{x}$, donc $\|\mathbf{x}\|^2 = 0$, et $\mathbf{x} = 0$.

2) Soit $\mathbf{x} \in E$. Posons $\phi_{\mathbf{x}}(\mathbf{y}) = \langle \mathbf{x}, \mathbf{y} \rangle$ pour $\mathbf{y} \in E$. Il résulte de la définition des produits scalaires que $\phi_{\mathbf{x}} : E \rightarrow \mathbb{R}$ est linéaire, et par conséquent $A^\perp = \bigcap_{\mathbf{x} \in A} \text{Ker}(\phi_{\mathbf{x}})$ est un sous-espace vectoriel de E . Il est clair que si $\mathbf{x} \in A$ on a $\mathbf{x} \perp \mathbf{y}$ pour tout $\mathbf{y} \in A^\perp$, et par conséquent $A \subset [A^\perp]^\perp$.

Si $\mathbf{x} \in A \cap A^\perp$, on a $\langle \mathbf{x}, \mathbf{x} \rangle = 0$, donc $\mathbf{x} = 0$. ♣

L'intérêt des bases orthonormales (BON) vient du fait qu'elles permettent de ramener les calculs de produit scalaire dans un espace euclidien quelconque à des calculs dans \mathbf{R}_n , identifié à $\mathcal{M}_{n,1}(\mathbf{R})$, muni de son produit scalaire canonique,

$$\text{défini pour } X = \begin{bmatrix} x_1 \\ \cdot \\ \cdot \\ \cdot \\ x_n \end{bmatrix} \in \mathbf{R}^n \text{ et } Y = \begin{bmatrix} y_1 \\ \cdot \\ \cdot \\ \cdot \\ y_n \end{bmatrix} \in \mathbf{R}^n \text{ par la formule}$$

$$\langle X, Y \rangle = \sum_{j=1}^n x_j y_j = {}^t X \cdot Y \quad (7.3)$$

On a le résultat suivant.

Proposition 7.3.3 Soit $E \neq \{0\}$ un espace euclidien de dimension finie et soit $\mathcal{B} = \{e_1, \dots, e_n\}$ une base orthonormale de E . On a alors, pour $\mathbf{x} \in E$

$$\mathbf{x} = \sum_{j=1}^n \langle \mathbf{x}, e_j \rangle e_j. \quad (7.4)$$

$$\text{En particulier si } \mathbf{x}, \mathbf{y} \in E, \text{ et si } \mathbf{x}_{\mathcal{B}} = \begin{bmatrix} \langle \mathbf{x}, e_1 \rangle \\ \cdot \\ \cdot \\ \cdot \\ \langle \mathbf{x}, e_n \rangle \end{bmatrix} \text{ et } \mathbf{y}_{\mathcal{B}} = \begin{bmatrix} \langle \mathbf{y}, e_1 \rangle \\ \cdot \\ \cdot \\ \cdot \\ \langle \mathbf{y}, e_n \rangle \end{bmatrix}$$

désignent les vecteurs colonnes formés des coordonnées de \mathbf{x} et \mathbf{y} dans la base \mathcal{B} , on a

$$\langle \mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{x}_B, \mathbf{y}_B \rangle = \sum_{j=1}^n \langle \mathbf{x}, e_j \rangle \langle \mathbf{y}, e_j \rangle = {}^t \mathbf{x}_B \mathbf{y}_B \quad (7.5)$$

Démonstration : Soit $\mathbf{x} \in E$, et posons $\mathbf{u} = \sum_{j=1}^n \langle \mathbf{x}, e_j \rangle e_j$. Pour $1 \leq p \leq n$, on a $\langle \mathbf{x} - \mathbf{u}, e_p \rangle = \langle \mathbf{x}, e_p \rangle - \sum_{j=1}^n \langle \mathbf{x}, e_j \rangle \langle e_j, e_p \rangle = \langle \mathbf{x}, e_p \rangle - \langle \mathbf{x}, e_p \rangle = 0$. Donc $\mathbf{x} - \mathbf{u} \in E^\perp = \{0\}$, et $\mathbf{x} = \mathbf{u}$.

On a alors, pour $\mathbf{x}, \mathbf{y} \in E$,

$$\begin{aligned} \langle \mathbf{x}, \mathbf{y} \rangle &= \left[\sum_{j=1}^n \langle \mathbf{x}, e_j \rangle e_j \right] \left[\sum_{k=1}^n \langle \mathbf{y}, e_k \rangle e_k \right] \\ &= \sum_{1 \leq j \leq n, 1 \leq k \leq n} \langle \mathbf{x}, e_j \rangle \langle \mathbf{y}, e_k \rangle \langle e_j, e_k \rangle = \sum_{j=1}^n \langle \mathbf{x}, e_j \rangle \langle \mathbf{y}, e_j \rangle. \end{aligned}$$

♣

Rappelons que si F et G sont deux sous-espaces vectoriels d'un espace vectoriel E , on pose $F + G = \{\mathbf{y} + \mathbf{z}\}_{\mathbf{y} \in F, \mathbf{z} \in G}$. Si $F \cap G = \{0\}$, tout élément \mathbf{x} de $F + G$ s'écrit de manière unique sous la forme $\mathbf{x} = \mathbf{x}_F + \mathbf{x}_G$, avec $\mathbf{x}_F \in F$, $\mathbf{x}_G \in G$. Dans ce cas on écrit $F \oplus G$ au lieu de $F + G$. Les applications $P_{F,G} : \mathbf{x} \mapsto \mathbf{x}_F$ et $P_{G,F} : \mathbf{x} \mapsto \mathbf{x}_G$ sont linéaires et sont appelées projection de $F \oplus G$ sur F parallèlement à G et projection de $F \oplus G$ sur G parallèlement à F . On a alors (voir le Chapitre 1 du présent cours)

$$P_{F,G} + P_{G,F} = I_{F \oplus G}, P_{F,G} \circ P_{F,G} = P_{F,G}, P_{G,F} \circ P_{G,F} = P_{G,F}, \text{Im}(P_{F,G}) = \text{Ker}(P_{G,F}) = F, \text{Im}(P_{G,F}) = \text{Ker}(P_{F,G}) = G.$$

On a l'important résultat suivant

Théorème 7.3.4 *Soit E un espace vectoriel euclidien, et soit F un sous-espace vectoriel de dimension finie de E .*

- 1) *On a $E = F \oplus F^\perp$, et $[F^\perp]^\perp = F$.*
- 2) *Posons $P_F = P_{F, F^\perp}$, $P_{F^\perp} = P_{F^\perp, F}$. On a les propriétés suivantes, pour $\mathbf{x} \in E$,*

$$(i) \|\mathbf{x}\|^2 = \|P_F(\mathbf{x})\|^2 + \|P_{F^\perp}(\mathbf{x})\|^2.$$

(ii) *Si $F \neq \{0\}$, on a*

$$P_F(\mathbf{x}) = \sum_{j=1}^p \langle \mathbf{x}, e_j \rangle e_j, \quad (7.6)$$

où $\{e_1, \dots, e_p\}$ est une base orthonormale quelconque de F .

Démonstration : Si $F = \{0\}$, $F^\perp = E$, $P_F = 0$, $P_{F^\perp} = I_E$ et il n'y a rien à démontrer. Supposons maintenant que $F \neq \{0\}$, et soit $\{e_1, \dots, e_p\}$ une base orthonormale de E . Soit $\mathbf{x} \in E$, et posons $\mathbf{u} = \sum_{j=1}^p \langle \mathbf{x}, e_j \rangle e_j$. De même que plus haut on voit que $\langle \mathbf{x} - \mathbf{u}, e_k \rangle = 0$ pour $1 \leq k \leq n$, ce qui montre que $\mathbf{x} - \mathbf{u} \in F^\perp$. Comme $\mathbf{u} \in F$, on voit que $\mathbf{x} = \mathbf{u} + (\mathbf{x} - \mathbf{u}) \in F + F^\perp$. Donc

$F + F^\perp = E$. Comme $F \cap F^\perp = \{0\}$, on a $E = F \oplus F^\perp$, et on vient de montrer que $P_F(\mathbf{x}) = \mathbf{u} = \sum_{j=1}^p \langle \mathbf{x}, e_j \rangle e_j$, et que $P_{F^\perp}(\mathbf{x}) = \mathbf{x} - \mathbf{u}$. Comme $\mathbf{u} \perp \mathbf{x} - \mathbf{u}$, on a bien $\|\mathbf{x}\|^2 = \|P_F(\mathbf{x})\|^2 + \|P_{F^\perp}(\mathbf{x})\|^2$, et la condition 2) est vérifiée.



7.4 Chapitre 7 sous Mupad

On peut évidemment effectuer sous Mupad le procédé d'orthogonalisation de Gram-Schmidt dans \mathbf{R}^n . On procède en deux temps

1) Après avoir établi la liste d'éléments de \mathbf{R}^n concernée, en déclarant ces éléments comme matrices unilignes ou comme matrices unicolonnes, on applique la commande

```
linalg : :orthog(.)
```

pour obtenir une famille orthogonale.

On applique ensuite la commande `map(.,linalg : :normalize)` pour remplacer les vecteurs orthogonaux de la première famille par des vecteurs de norme 1 qui leur sont proportionnels. On peut effectuer les deux opérations en même temps grâce à la commande `map(linalg : :orthog(.),linalg : :normalize)`

```
M:=Dom::Matrix();
S:=[M([1,1,1,1,1,1,1,1]),M([1,2,3,4,5,6,7,8]),M([1,12,33,44,55,66,77,88])];
U:=linalg::orthog(S);
```

Dom::Matrix()

$$\begin{bmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \\ 8 \end{pmatrix}, \begin{pmatrix} 1 \\ 12 \\ 33 \\ 44 \\ 55 \\ 66 \\ 77 \\ 88 \end{pmatrix} \\ \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} -\frac{7}{2} \\ -\frac{5}{2} \\ -\frac{3}{2} \\ -\frac{1}{2} \\ \frac{1}{2} \\ \frac{3}{2} \\ \frac{5}{2} \\ \frac{7}{2} \end{pmatrix}, \begin{pmatrix} -\frac{5}{2} \\ -\frac{55}{14} \\ \frac{65}{14} \\ \frac{45}{14} \\ \frac{25}{14} \\ \frac{5}{14} \\ -\frac{15}{14} \\ -\frac{5}{2} \end{pmatrix} \end{bmatrix}$$

```
V:=map(U,linalg::normalize);
```

$$\begin{bmatrix} \frac{1/2}{8} \\ \frac{1/2}{8} \end{bmatrix}, \begin{bmatrix} \frac{1/2}{42} \\ \frac{1/2}{5 \cdot 42} \\ \frac{1/2}{42} \\ \frac{1/2}{42} \\ \frac{1/2}{42} \\ \frac{1/2}{42} \\ \frac{1/2}{42} \\ \frac{1/2}{5 \cdot 42} \\ \frac{1/2}{42} \end{bmatrix}, \begin{bmatrix} \frac{1/2}{7} \\ \frac{1/2}{11 \cdot 7} \\ \frac{1/2}{13 \cdot 7} \\ \frac{1/2}{7} \\ \frac{1/2}{7} \\ \frac{1/2}{7} \\ \frac{1/2}{7} \\ \frac{1/2}{7} \\ \frac{1/2}{7} \end{bmatrix}, \begin{bmatrix} \frac{1/2}{450} \\ \frac{1/2}{450} \end{bmatrix}, \begin{bmatrix} \frac{1/2}{180} \\ \frac{1/2}{1260} \\ \frac{1/2}{1260} \\ \frac{1/2}{140} \\ \frac{1/2}{252} \\ \frac{1/2}{1260} \\ \frac{1/2}{420} \\ \frac{1/2}{180} \end{bmatrix}$$

7.5 Exercices sur le Chapitre 7

exercice 1

Soit $ABCD$ un parallélogramme du plan usuel. En utilisant la Proposition 7.2.3, montrer que $AB^2 + BC^2 + CD^2 + DA^2 = AC^2 + BD^2$.

exercice 2

Soit \mathcal{E} l'espace vectoriel des fonctions continues et périodiques de période 2π sur \mathbf{R} , que l'on munit du produit scalaire défini pour $f, g \in \mathcal{E}$ par la formule

$$\langle f, g \rangle = \frac{1}{2\pi} \int_0^{2\pi} f(t)g(t)dt.$$

On définit $e_1, e_2, e_3 \in \mathcal{E}$ par les formules

$$e_1(t) = 1, \quad e_2(t) = \cos(t), \quad e_3(t) = \cos^2(t).$$

Déterminer la famille orthonormale obtenue en appliquant à $\{e_1, e_2, e_3\}$ le procédé d'orthonormalisation de Gram-Schmidt.

exercice 3

On considère l'espace vectoriel \mathcal{E} des fonctions continues et périodiques de période 2π sur \mathbf{R} , muni du produit scalaire introduit à la question précédente. On pose $u_0(t) = 1$, et, pour $n \geq 1$,

$$u_n(t) = \cos(nt), \quad v_n(t) = \sin(nt).$$

1) Montrer que la famille $\mathcal{B} := (\cup_{n \geq 0} u_n) \cup (\cup_{n \geq 1} v_n)$ est une famille orthonormale d'éléments de \mathcal{E} .

2) Pour $p \geq 1$, on note F_p le sous-espace vectoriel de \mathcal{E} engendré par $\{u_0, \dots, u_p\} \cup \{v_1, \dots, v_p\}$. Soit P la projection orthogonale de \mathcal{E} sur F_p . En utilisant une formule du cours, expliciter $P(f)(x)$ pour $f \in \mathcal{E}$, $x \in \mathbf{R}$. Quel est le lien entre le polynôme trigonométrique $P(f)$ et la série de Fourier de f ?

exercice 4

Soit F l'image de \mathbf{R}^3 par l'application $u : X \mapsto \begin{bmatrix} 1 & 1 & 1 \\ 2 & -1 & 2 \\ 1 & 0 & 1 \end{bmatrix} X$.

Trouver une base orthogonale de F , et calculer la matrice représentant la projection orthogonale de \mathbf{R}^3 sur F par rapport à la base canonique de \mathbf{R}^3 .

exercice 5

En s'aidant de Mupad, répondre aux mêmes questions que dans l'exercice précédent pour $G := \{AX\}_{X \in \mathbb{R}^8}$, avec

$$A = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 8 & 7 & 6 & 5 & 4 & 3 & 2 & 1 \\ 1 & 3 & 5 & 7 & 9 & 11 & 13 & 15 \\ 15 & 13 & 11 & 9 & 7 & 5 & 3 & 1 \\ 1 & 4 & 7 & 10 & 13 & 16 & 19 & 22 \\ 22 & 19 & 16 & 13 & 10 & 7 & 4 & 1 \\ 1 & 5 & 9 & 13 & 17 & 21 & 25 & 29 \end{bmatrix}.$$

exercice 6

Dans le plan affine euclidien rapporté à un repère orthonormé (O, \vec{i}, \vec{j}) on considère la droite D d'équation

$$ax + by + c = 0,$$

avec $(a, b) \neq (0, 0)$.

1) Ecrire l'équation de la droite perpendiculaire à D et passant par le point M_0 de coordonnées (x_0, y_0) .

2) En déduire que la distance de M_0 à la droite D est donnée par la formule

$$d(M_0, D) = \frac{|ax_0 + by_0 + c|}{\sqrt{a^2 + b^2}}.$$

exercice 7

Dans l'espace affine euclidien usuel rapporté à un repère orthonormé $(O, \vec{i}, \vec{j}, \vec{k})$, on considère le plan P d'équation

$$ax + by + cz + d = 0,$$

avec $(a, b, c) \neq (0, 0, 0)$.

1) Ecrire l'équation de la droite perpendiculaire à P et passant par le point M_0 de coordonnées (x_0, y_0, z_0) .

2) En déduire que la distance de M_0 au plan P est donnée par la formule

$$d(M_0, P) = \frac{|ax_0 + by_0 + cz_0 + d|}{\sqrt{a^2 + b^2 + c^2}}.$$

exercice 8

Dans l'espace affine euclidien usuel rapporté à un repère orthonormé $(O, \vec{i}, \vec{j}, \vec{k})$, on considère les droites D_1 et D_2 d'équations paramétrées

$$\begin{cases} x = x_0 + \lambda u_0 \\ y = y_0 + \lambda v_0 \\ z = z_0 + \lambda w_0 \end{cases}$$

$$\begin{cases} x = x_1 + \lambda u_1 \\ y = y_1 + \lambda v_1 \\ z = z_1 + \lambda w_1 \end{cases}$$

On suppose D_1 et D_2 non parallèles. Donner l'équation de la perpendiculaire commune à D_1 et D_2 .

Chapitre 8

Matrices symétriques, matrices orthogonales

8.1 Matrices symétriques

Définition 8.1.1 (i) Si $A = (a_{i,j})_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}} \in \mathcal{M}_{m,n}(K)$, on appelle **transposée** de A la matrice

$${}^t A = (a_{j,i})_{\substack{1 \leq j \leq n \\ 1 \leq i \leq m}} \in \mathcal{M}_{n,m}(K)$$

obtenue en échangeant les lignes et les colonnes de A .

(ii) On dit que $B \in \mathcal{M}_n(K)$ est **symétrique** si $B = {}^t B$.

Il est clair que l'application $A \mapsto {}^t A$ est une application linéaire de $\mathcal{M}_{m,n}(K)$ sur $\mathcal{M}_{n,m}(K)$, et que si $A \in \mathcal{M}_{m,n}(K)$, on a

$${}^t({}^t A) = A. \quad (8.1)$$

D'autre part si $\mathcal{M}_{m,n}(K)$, et si $B \in \mathcal{M}_{n,p}(K)$, on a

$${}^t(AB) = ({}^t B)({}^t A) \in \mathcal{M}_{p,m}(K). \quad (8.2)$$

Enfin si $P \in \mathcal{M}_n(K)$, alors $\det(P) = \det({}^t P)$. En particulier **P et sa transposée ont le même polynôme caractéristique.**

Dans toute la suite on munit \mathbf{R}^n , identifié à $\mathcal{M}_{n,1}(\mathbf{R})$, du produit scalaire canonique défini pour $X, Y \in \mathbf{R}^n$ par la formule

$$\langle X, Y \rangle = {}^t X \cdot Y,$$

et on note $\|X\|_2 := \sqrt{\langle X, X \rangle}$ la norme euclidienne associée sur \mathbf{R}^n . On peut étendre ce produit scalaire en une "forme hermitienne" sur \mathbf{C}^n identifié à $\mathcal{M}_{n,1}(\mathbf{C})$ en posant, pour $X, Y \in \mathbf{C}^n$,

$$\langle X, Y \rangle = {}^t X \cdot \bar{Y},$$

où $\bar{A} = (\bar{a}_{i,j})_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}} \in \mathcal{M}_{m,n}(K)$ pour $A = (a_{i,j})_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}} \in \mathcal{M}_{m,n}(K)$.

Il est clair que $\langle X, X \rangle = \sum_{k=1}^n |x_k|^2 > 0$ si $X = \begin{bmatrix} x_1 \\ \cdot \\ \cdot \\ \cdot \\ x_n \end{bmatrix} \in \mathbf{C}^n \setminus \{0\}$.

On va montrer que toute matrice symétrique réelle est diagonalisable, et plus précisément que si $A \in \mathcal{M}_n(\mathbf{R})$ est symétrique il existe une base orthonormale de \mathbf{R}^n formée de vecteurs propres de A . Nous aurons besoin de trois lemmes.

Lemme 8.1.2 *Soit $A \in \mathcal{M}_n(\mathbf{R})$ une matrice symétrique réelle. Alors toutes les valeurs propres de A sont réelles (autrement dit le polyôme caractéristique de A est scindé sur \mathbf{R}).*

Démonstration : Soit $\lambda \in \mathbf{C}$ une valeur propre de A et soit $X \in \mathbf{C}^n \setminus \{0\}$ tel que $AX = \lambda X$. On a $\lambda^t X = {}^t(\lambda X) = ({}^t X)({}^t A) = ({}^t X)A$, et $A\bar{X} = \bar{A}\bar{X} = \overline{(AX)} = \overline{\lambda X} = \bar{\lambda}\bar{X}$. On obtient

$$\lambda \langle X, X \rangle = (\lambda^t X)\bar{X} = ({}^t X A)\bar{X} = ({}^t X)(A\bar{X}) = ({}^t X)(\bar{\lambda}\bar{X}) = \bar{\lambda} \langle X, X \rangle.$$

Comme $X \neq 0$, on a $\langle X, X \rangle \neq 0$, et $\lambda = \bar{\lambda}$, ce qui montre que $\lambda \in \mathbf{R}$. ♣

On dira que deux sous-espaces E et F de \mathbf{R}^n sont **orthogonaux**, et on note $E \perp F$, si $\langle x, y \rangle = 0$ pour tout $x \in E$ et pour tout $y \in F$, ce qui est équivalent à l'une ou l'autre des conditions $E \subset F^\perp$ ou $F \subset E^\perp$. On a le résultat suivant.

Lemme 8.1.3 *Soit $A \in \mathcal{M}_n(\mathbf{R})$, une matrice symétrique, Si λ et μ sont deux valeurs propres réelles distinctes de A , alors les sous-espaces propres E_λ et E_μ sont orthogonaux.*

Démonstration : Soit $X \in E_\lambda$, et soit $Y \in E_\mu$. On a

$$(\mu - \lambda) \langle X, Y \rangle = {}^t X(\mu Y) - {}^t(\lambda X)Y = {}^t XAY - {}^t X^t AY = 0,$$

ce qui donne $\langle X, Y \rangle = 0$, puisque $\lambda \neq \mu$. ♣

Lemme 8.1.4 *Soit $A \in \mathcal{M}_n(\mathbf{R})$ une matrice symétrique. Si A est nilpotente, alors $A = 0$.*

Démonstration : Supposons que A est symétrique. On a, pour $X \in \mathbf{R}^n$,

$$\|AX\|_2^2 = \langle AX, AX \rangle = {}^t X^t AAX = {}^t XA^2X = \langle X, A^2x \rangle \leq \|X\|_2 \|A^2X\|_2.$$

Donc si on pose $\text{Ker}(B) = \{X \in \mathbf{R}^n \mid BX = 0\}$ pour $B \in \mathcal{M}_n(\mathbf{R})$, on voit que $\text{Ker}(A^2) \subset \text{Ker}(A)$. Comme l'inclusion $\text{Ker}(A) \subset \text{Ker}(A^2)$ est toujours vérifiée, on a en fait $\text{Ker}(A^2) = \text{Ker}(A)$. Soit maintenant $m \geq 1$. On a $\text{Ker}(A^{m+1}) =$

$\{X \in \mathbf{R}^n \mid A^{m-1}X \in \text{Ker}(A^2)\} = \{X \in \mathbf{R}^n \mid A^{m-1}X \in \text{Ker}(A)\} = \text{Ker}(A^m)$, et on voit que $\text{Ker}(A^m) = \text{Ker}(A)$ pour tout $m \geq 1$. En particulier si A est à la fois symétrique et nilpotente on a $\text{Ker}(A) = \mathbf{R}^n$, de sorte que $A = 0$. ♣

On a alors le théorème suivant.

Théorème 8.1.5 *Soit $A \in \mathcal{M}_n(\mathbf{R})$ une matrice symétrique. Alors A est diagonalisable sur \mathbf{R} , et il existe une base orthonormale de \mathbf{R}^n formée de vecteurs propres de A .*

Démonstration : Comme le polynôme caractéristique de A est scindé sur \mathbf{R} , A admet une décomposition de Jordan sur \mathbf{R} , et il existe un unique couple (D, N) de matrices carrées réelles à n lignes et n colonnes vérifiant

- (i) $A = D + N$,
- (ii) D est diagonalisable sur \mathbf{R} , et N est nilpotente,
- (iii) $ND = DN$.

Il résulte de la formule 8.2 que tD est diagonalisable sur \mathbf{R} , que tN est nilpotente, et que ${}^tN{}^tD = {}^t(DN) = {}^t(ND) = {}^tD{}^tN$. Comme $A = {}^tA = {}^tD + {}^tN$, il résulte de l'unicité de la décomposition de Jordan que $D = {}^tD$ et $N = {}^tN$. Il résulte alors du lemme 8.1.4 que $N = 0$, et $A = D$ est diagonalisable sur \mathbf{R} . Par conséquent si $\lambda_1, \dots, \lambda_k$ sont les valeurs propres distinctes de A , et si \mathcal{B}_j est une base de E_{λ_j} pour $1 \leq j \leq k$, alors $\cup_{1 \leq j \leq k} \mathcal{B}_j$ est une base de \mathbf{R}^n formée de vecteurs propres de A . En utilisant le procédé d'orthonormalisation de Gram-Schmidt on peut trouver pour tout $j \leq k$ une base orthonormale \mathcal{B}_j de E_{λ_j} , et il résulte alors du lemme 8.1.3 que $\mathcal{B} := \cup_{1 \leq j \leq k} \mathcal{B}_j$, qui est une base de E formée de vecteurs propres de A , est orthonormale. ♣

8.2 Matrices orthogonales

On vient de voir que si $A \in \mathcal{M}_n(\mathbf{R})$ est symétrique, il existe une matrice inversible $P \in \mathcal{M}_n(\mathbf{R})$ telle que $P^{-1}AP$ soit diagonale dont les colonnes forment une base orthonormale de \mathbf{R}^n . Le résultat suivant donne une description complète de ce type de matrices.

Théorème 8.2.1 *Soit $B \in \mathcal{M}_n(\mathbf{R})$. Alors les conditions suivantes sont équivalentes.*

- (1) Les colonnes de B forment une base orthonormale de \mathbf{R}^n .
 - (2) Les lignes de B forment une base orthonormale de \mathbf{R}^n .
 - (3) ${}^tBB = I_n$.
 - (4) $B{}^tB = I_n$.
 - (5) B est inversible, et $B^{-1} = {}^tB$.
 - (6) $\langle BX, BY \rangle = \langle X, Y \rangle$ pour tout couple (X, Y) d'éléments de \mathbf{R}^n .
 - (7) $\|BX\|_2 = \|X\|_2$ pour tout $X \in \mathbf{R}^n$.
- Dans ce cas on dit que la matrice B est orthogonale, et $\det(B) \in \{-1, 1\}$.

Démonstration : Soient L_1, \dots, L_n les lignes de B , indexées de manière croissante de haut en bas, et soient C_1, \dots, C_n les colonnes de B , indexées de manière croissante de gauche à droite. On a

$${}^tBB = (\langle C_i, C_j \rangle)_{\substack{1 \leq i \leq n \\ 1 \leq j \leq n}}, \quad B^tB = (\langle L_i, L_j \rangle)_{\substack{1 \leq i \leq n \\ 1 \leq j \leq n}},$$

ce qui prouve que (1) est équivalent à (3) et que (2) est équivalent à (4). D'autre part si (3) ou (4) est vérifié alors $\det(B) \neq 0$, donc B est inversible et dans ce cas B^{-1} coïncide avec l'inverse à gauche ou à droite de B , c'est à dire que $B^{-1} = {}^tB$. Donc (3) ou (4) impliquent (5). Comme (5) implique trivialement (3) et (4) on voit que les conditions (1), (2), (3), (4) et (5) sont équivalentes. D'autre part si (3) est vérifié on a $\det(B)^2 = \det({}^tB)\det(B) = \det(I_n) = 1$, donc $\det(B) \in \{-1, 1\}$.

Supposons maintenant que (3) est vérifié, et soient $X, Y \in \mathbf{R}^n$. On a

$$\langle BX, BY \rangle = {}^tX^tBBY = {}^tXY = \langle X, Y \rangle,$$

donc (6) est vérifié. Supposons maintenant que (6) est vérifié, et soit $\{e_1, \dots, e_n\}$ la base canonique de \mathbf{R}^n . On a $C_j = Be_j$, donc on a $\langle C_i, C_j \rangle = \langle Be_i, Be_j \rangle = \delta_{i,j}$ pour $1 \leq i \leq n, 1 \leq j \leq n$, $\delta_{i,j}$ désignant le symbole de Kronecker. Ceci montre que la condition (1) est vérifiée, et les conditions (1), (2), (3), (4), (5) et (6) sont équivalentes.

Il est clair que (6) implique (7). Supposons maintenant que (7) est vérifié, et soient $X, Y \in \mathbf{R}^n$. On a

$$\begin{aligned} & \langle BX, BY \rangle \\ &= \frac{1}{4} (\langle B(X+Y), B(X+Y) \rangle - \langle B(X-Y), B(X-Y) \rangle) \\ &= \frac{1}{4} (\|B(X+Y)\|_2^2 - \|B(X-Y)\|_2^2) \\ &= \frac{1}{4} (\|X+Y\|_2^2 - \|X-Y\|_2^2) = \frac{1}{4} (\langle X+Y, X+Y \rangle - \langle X-Y, X-Y \rangle) \\ &= \langle X, Y \rangle, \end{aligned}$$

et on voit que les conditions (6) et (7) sont équivalentes, ce qui achève la démonstration. ♣

On a introduit au Chapitre 2 la matrice $P_{\mathcal{B}, \mathcal{B}'}$ représentant une famille \mathcal{B}' de n éléments d'un espace vectoriel de dimension $n \geq 1$ muni d'une base \mathcal{B} . On sait que \mathcal{B}' est une base de E si et seulement si $P_{\mathcal{B}, \mathcal{B}'}$ est inversible, et dans ce cas $P_{\mathcal{B}, \mathcal{B}'}$ est appelée matrice de passage de \mathcal{B} à \mathcal{B}' . Le corollaire suivant montre que si E est euclidien, et si \mathcal{B} est une base orthonormale de E , alors une autre base \mathcal{B}' de E est orthonormale si et seulement si la matrice de passage de \mathcal{B} à \mathcal{B}' est orthogonale.

Corollaire 8.2.2 *Soit E un espace euclidien de dimension finie, soit $\mathcal{B} = \{e_1, \dots, e_n\}$ une base orthonormale de E et soit $\mathcal{B}' = \{e'_1, \dots, e'_n\}$ une famille de n éléments de E . Alors les conditions suivantes sont équivalentes*

- (i) \mathcal{B}' est une base orthonormale de E .
- (ii) La matrice $P_{\mathcal{B}, \mathcal{B}'}$ représentant \mathcal{B}' dans la base \mathcal{B} est orthogonale.

Démonstration : Soit $P = P_{\mathcal{B}, \mathcal{B}'}$, et soient P_1, \dots, P_n les colonnes de P , de sorte que P_j est formée des coordonnées de e'_j dans la base \mathcal{B} . Comme \mathcal{B} est orthonormale, il résulte de la proposition 7.3.3 que $\langle e'_i, e'_j \rangle = \langle P_i, P_j \rangle$ pour $1 \leq i \leq n$, $1 \leq j \leq n$. Autrement dit la famille \mathcal{B}' est orthonormale si et seulement si la matrice P est orthogonale, et dans ce cas \mathcal{B}' est automatiquement une base de E . ♣

Notons que le résultat ci-dessus permet de simplifier notablement les calculs concernant la diagonalisation d'une matrice symétrique A : en construisant une base orthonormale de \mathbf{R}^n formée de vecteurs propres de A , on obtient une matrice orthogonale P telle que $P^{-1}AP$ soit diagonale, et le calcul de P^{-1} est immédiat puisque $P^{-1} = {}^t P$.

Exemple 8.2.3 Diagonaliser $A = \begin{bmatrix} 1 & -2 & 1 \\ -2 & 1 & 1 \\ 1 & 1 & -2 \end{bmatrix}$.

On a,

$$\begin{aligned} p_A = - \begin{vmatrix} 1-x & -2 & 1 \\ -2 & 1-x & 1 \\ 1 & 1 & -2-x \end{vmatrix} &= - \begin{vmatrix} -x & -2 & 1 \\ -x & 1-x & 1 \\ -x & 1 & -2-x \end{vmatrix} = x \begin{vmatrix} 1 & -2 & 1 \\ 1 & 1-x & 1 \\ 1 & 1 & -2-x \end{vmatrix} \\ &= x \begin{vmatrix} 1 & -2 & 1 \\ 0 & 3-x & 0 \\ 0 & 3 & -3-x \end{vmatrix} = x(x-3)(x+3). \end{aligned}$$

Donc A admet 0, 3 et -3 comme valeurs propres simples. On a

$$A^2 = \begin{bmatrix} 1 & -2 & 1 \\ -2 & 1 & 1 \\ 1 & 1 & -2 \end{bmatrix} \begin{bmatrix} 1 & -2 & 1 \\ -2 & 1 & 1 \\ 1 & 1 & -2 \end{bmatrix} = \begin{bmatrix} 6 & * & * \\ -3 & * & * \\ -3 & * & * \end{bmatrix},$$

$$(A+3I)(A-3I) = A^2 - 9I = \begin{bmatrix} -3 & * & * \\ -3 & * & * \\ -3 & * & * \end{bmatrix},$$

$$A(A+3I) = A^2 + 3A = \begin{bmatrix} 9 & * & * \\ -9 & * & * \\ 0 & * & * \end{bmatrix},$$

$$A(A-3I) = A^2 - 3A = \begin{bmatrix} 3 & * & * \\ 3 & * & * \\ -6 & * & * \end{bmatrix}.$$

On sait que les sous-espaces propres E_0 , E_3 et E_{-3} sont respectivement engendrés par les colonnes des matrices $(A+3I)(A-3I)$, $A(A+3I)$ et $A(A-3I)$, qui sont évidemment ici de rang 1 puisque les trois valeurs propres de A sont

simples. Donc $\left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix} \right\}$, qui est évidemment orthogonale, est formée de vecteurs propres de A , et on obtient une base orthonormale de \mathbf{R}^n formée de vecteurs propres de A en posant

$$\mathcal{B} := \left\{ \begin{bmatrix} \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{bmatrix}, \begin{bmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \\ 0 \end{bmatrix}, \begin{bmatrix} \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} \\ -\frac{2}{\sqrt{6}} \end{bmatrix} \right\}.$$

Posons

$$P = \begin{bmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & 0 & -\frac{2}{\sqrt{6}} \end{bmatrix} \quad \text{de sorte que } P^{-1} = {}^t P = \begin{bmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & -\frac{2}{\sqrt{6}} \end{bmatrix}.$$

$$\text{On a alors } P^{-1}AP = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & -3 \end{bmatrix}, \text{ et } A = P \begin{bmatrix} 0 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & -3 \end{bmatrix} P^{-1}.$$

Le corollaire précédent permet également d'orienter un espace vectoriel euclidien E de dimension finie à partir d'une base orthonormale \mathcal{B}_0 de E : on dira qu'une base orthonormale de E est directe si le déterminant de la matrice de passage $P_{\mathcal{B}_0, \mathcal{B}}$ est égal à 1, et on dira que la base orthonormale \mathcal{B} est indirecte si $\det P_{\mathcal{B}_0, \mathcal{B}} = -1$.

On déduit également du théorème 8.2.1 le résultat suivant, où l'espace euclidien considéré est muni de la norme euclidienne associée à son produit scalaire.

Corollaire 8.2.4 *Soient E un espace euclidiens de dimension finie $n \geq 1$, soit \mathcal{B} une base orthonormale de E , et soit $u : E \rightarrow E$ une application linéaire. Alors les trois conditions suivantes sont équivalentes*

- (i) $\|u(\vec{x})\| = \|\vec{x}\|$ pour tout $\vec{x} \in E$.
- (ii) $\langle u(\vec{x}), u(\vec{y}) \rangle = \langle \vec{x}, \vec{y} \rangle$ pour tout couple (\vec{x}, \vec{y}) éléments de E .
- (iii) La matrice $\mathcal{M}_{u, \mathcal{B}}$ représentant u par rapport à la base \mathcal{B} est orthogonale.

Démonstration : Posons $A := \mathcal{M}_{u, \mathcal{B}}$, soient $\vec{x}, \vec{y} \in E$, et soient X et Y les matrices unicolonnes formées des coordonnées de \vec{x} et \vec{y} dans la base \mathcal{B} . Alors AX et AY sont les matrices unicolonnes formées des coordonnées de $u(\vec{x})$ et $u(\vec{y})$ dans la base \mathcal{B} , et il résulte de la proposition 7.3.3 que l'on a

$$\langle \vec{x}, \vec{y} \rangle = \langle X, Y \rangle, \quad \langle u(\vec{x}), u(\vec{y}) \rangle = \langle AX, AY \rangle,$$

et le corollaire résulte immédiatement du théorème 8.2.1. ♣

On a plus généralement le résultat suivant.

Proposition 8.2.5 Soient $E \neq 0$ et $F \neq 0$ deux espaces euclidiens de dimension finie $n \geq 1$, soit \mathcal{B}_0 une base orthonormale de E , soit \mathcal{B}_1 une base orthonormale de F , soit $u : E \rightarrow F$ une application linéaire, et soit $A := \mathcal{M}_{u, \mathcal{B}_0, \mathcal{B}_1}$ la matrice représentant u par rapport aux bases \mathcal{B}_0 et \mathcal{B}_1 . Alors les trois conditions suivantes sont équivalentes

- (i) $\|u(\vec{x})\| = \|\vec{x}\|$ pour tout $\vec{x} \in E$.
- (ii) $\langle u(\vec{x}), u(\vec{y}) \rangle = \langle \vec{x}, \vec{y} \rangle$ pour tout couple (\vec{x}, \vec{y}) d'éléments de E .
- (iii) ${}^t AA = I_n$, où $n = \dim(E)$.

Dans ce cas on dira que u est une **isométrie vectorielle** de E dans F .

Démonstration : Il est clair que (ii) implique (i), et on a pour $\vec{x}, \vec{y} \in E$,

$$\langle \vec{x}, \vec{y} \rangle = \frac{1}{4} (\|\vec{x} + \vec{y}\|^2 + \|\vec{x} - \vec{y}\|^2),$$

$$\langle u(\vec{x}), u(\vec{y}) \rangle = \frac{1}{4} (\|u(\vec{x}) + u(\vec{y})\|^2 + \|u(\vec{x}) - u(\vec{y})\|^2),$$

ce qui de même que dans la démonstration du théorème 8.2.1 montre que (i) et (ii) sont en fait équivalents.

Soient maintenant $\vec{x}, \vec{y} \in E$, et soient X et Y les matrices unicolonnes formées des coordonnées de \vec{x} et \vec{y} dans la base \mathcal{B}_0 , de sorte que AX et AY sont les matrices unicolonnes formées des coordonnées de $u(\vec{x})$ et $u(\vec{y})$ dans la base \mathcal{B}_1 . Il résulte du lemme 7.3.3 que l'on a

$$\langle \vec{x}, \vec{y} \rangle = {}^t XY, \quad \langle u(\vec{x}), u(\vec{y}) \rangle = {}^t (AX)AY = {}^t X{}^t AA Y.$$

Donc (iii) implique (ii). D'autre part si (ii) est vérifié on a ${}^t XY = {}^t X({}^t AA)Y$ pour tout couple (X, Y) d'éléments de R^n . On voit alors que ${}^t AA Y - Y \in (R^n)^\perp = \{0\}$ pour tout $Y \in R^n$. En appliquant ceci à la base canonique de R^n on voit que toutes les colonnes de ${}^t AA$ et I_n coïncident et on voit que (ii) implique (iii), ce qui achève la démonstration. ♣

Exemple 8.2.6 Matrice représentant le retournement R (rotation vectorielle d'angle π) ayant pour axe la droite vectorielle engendrée par $\vec{i} - \vec{j} + \vec{k}$ dans l'espace vectoriel euclidien usuel E de dimension 3 usuel rapporté à une base orthonormée $\mathcal{B} = \{\vec{i}, \vec{j}, \vec{k}\}$.

L'espace vectoriel E est l'espace vectoriel associé à l'espace affine \mathcal{E} euclidien de Terminale, que l'on rapporte à un repère $(O, \vec{i}, \vec{j}, \vec{k})$. On note D la droite issue de O de vecteur directeur $\vec{i} - \vec{j} + \vec{k}$. On a alors $R(\overrightarrow{OM}) = \overrightarrow{OM'}$, où M' vérifie les deux conditions suivantes

- 1) $MM' \perp D$,
- 2) $I \in D$, où I désigne le milieu de MM' .

On a, en désignant par (x, y, z) les coordonnées de M et par (x', y', z') celles de M' ,

$\overrightarrow{MM'} = \overrightarrow{MO} + \overrightarrow{OM'} = (x' - x)\vec{i} + (y' - y)\vec{j} + (z' - z)\vec{k}$, et la condition 1) donne

$$1') x' - x - y' + y + z' - z = 0.$$

D'autre part $\vec{OI} = \frac{\vec{OM} + \vec{OM'}}{2} = \frac{x+x'}{2}\vec{i} + \frac{y+y'}{2}\vec{j} + \frac{z+z'}{2}\vec{k}$. La condition 2) donne

$$\begin{cases} \frac{x+x'}{2} = \lambda \\ \frac{y+y'}{2} = -\lambda \\ \frac{z+z'}{2} = \lambda \end{cases}$$

avec $\lambda \in \mathbf{R}$.

On obtient alors la condition suivante, équivalente à la condition 2)

$$2') x + x' = z + z' = -y - y'.$$

En appliquant ceci à \vec{i} , on a $x = 1$, $y = z = 0$, d'où

$$z' = x' + 1 = -y', \quad x' - 1 + x' + 1 + x' + 1 = 0, \quad x' = -\frac{1}{3}, \quad y' = -\frac{2}{3}, \quad z' = \frac{2}{3},$$

$R(\vec{i}) = -\frac{1}{3}\vec{i} - \frac{2}{3}\vec{j} + \frac{2}{3}\vec{k}$. Comme les rôles de \vec{i} et \vec{k} sont symétriques, on obtient

$$R(\vec{k}) = \frac{2}{3}\vec{i} - \frac{2}{3}\vec{j} - \frac{1}{3}\vec{k}. \text{ Dans le cas de } \vec{j}, \text{ on a } y = 1, x = z = 0, \text{ et on obtient}$$

$$z' = x' = -1 - y', \quad -1 - y' - y' + 1 - 1 - y' = 0, \quad y' = -\frac{1}{3}, \quad x' = z' = -\frac{2}{3},$$

$$R(\vec{j}) = -\frac{2}{3}\vec{i} - \frac{1}{3}\vec{j} - \frac{2}{3}\vec{k}.$$

On obtient

$$\mathcal{M}_{R,B} = \begin{matrix} & & R(\vec{i}) & R(\vec{j}) & R(\vec{k}) \\ \begin{matrix} \vec{i} \\ \vec{j} \\ \vec{k} \end{matrix} & \left[\begin{array}{ccc} -\frac{1}{3} & -\frac{2}{3} & \frac{2}{3} \\ -\frac{2}{3} & -\frac{1}{3} & -\frac{2}{3} \\ \frac{2}{3} & -\frac{2}{3} & -\frac{1}{3} \end{array} \right] \end{matrix}$$

qui est bien une matrice orthogonale.

8.3 Formes quadratiques

Rappelons que les formes bilinéaires symétriques ont été introduites à la définition 7.1.1.

Nous allons maintenant introduire la notion de *forme quadratique* sur un espace vectoriel réel. Comme on le verra plus loin, les formes quadratiques sur \mathbf{R}^n sont tout simplement les polynômes homogènes de degré 2 de n variables à coefficients réels, et les formes quadratiques sur un espace vectoriel réel E de dimension finie se décrivent de manière analogue en fonction de leurs coordonnées par rapport à une base quelconque de E .

Définition 8.3.1 Soit E un espace vectoriel réel. On dit qu'une application $q : E \rightarrow \mathbf{R}$ est une forme quadratique sur E si $q(\vec{0}) = 0$, et si l'application

$$(\vec{x}, \vec{y}) \mapsto q(\vec{x} + \vec{y}) - q(\vec{x} - \vec{y})$$

est une forme bilinéaire symétrique sur E .

Dans ce cas on pose, pour $\vec{x}, \vec{y} \in E$,

$$\phi_q(\vec{x}, \vec{y}) = \frac{1}{4} (q(\vec{x} + \vec{y}) - q(\vec{x} - \vec{y})), \quad (8.3)$$

de sorte que ϕ_q est également une forme bilinéaire symétrique sur E .

Soit q est une forme quadratique sur E . Comme $q(\vec{0}) = 0$, en remplaçant \vec{x} et \vec{y} par $\frac{\vec{x}}{2}$ dans l'identité 8.3 on obtient, pour $\vec{x} \in E$,

$$q(\vec{x}) = q\left(\frac{\vec{x}}{2} + \frac{\vec{x}}{2}\right) = 4\phi_q\left(\frac{\vec{x}}{2}, \frac{\vec{x}}{2}\right) = \phi_q(\vec{x}, \vec{x}). \quad (8.4)$$

Il est clair réciproquement que si $\phi : E \times E \rightarrow \mathbf{R}$ est une forme bilinéaire symétrique sur E , alors $q_\phi : \vec{x} \mapsto \phi(\vec{x}, \vec{x})$ est une forme quadratique sur E , puisque $\phi_{q_\phi}(\vec{x}, \vec{y}) = \frac{1}{4} (q_\phi(\vec{x} + \vec{y}) - q_\phi(\vec{x} - \vec{y})) = \frac{1}{4} (\phi(\vec{x} + \vec{y}, \vec{x} + \vec{y}) - \phi(\vec{x} - \vec{y}, \vec{x} - \vec{y})) = \phi(\vec{x}, \vec{y})$ pour $\vec{x}, \vec{y} \in E$. L'application $q \mapsto \phi_q$ est donc une bijection de l'ensemble $\mathcal{Q}(E)$ des formes quadratiques sur l'ensemble $L_2(E, \mathbf{R})$ des formes bilinéaires sur E .

Définition 8.3.2 Soit $E \neq \{0\}$ un espace vectoriel réel de dimension finie, soit $\phi : E \times E \rightarrow \mathbf{R}$ une forme bilinéaire symétrique sur E , et soit $\mathcal{B} = \{\vec{e}_1, \dots, \vec{e}_n\}$ une base de E .

On appelle matrice représentant ϕ dans la base \mathcal{B} la matrice

$$\mathcal{M}_{[\phi, \mathcal{B}]} = (\phi(\vec{e}_i, \vec{e}_j))_{\substack{1 \leq i \leq n \\ 1 \leq j \leq n}}. \quad (8.5)$$

Si q est une forme quadratique sur E , on pose

$$\mathcal{M}_{[q, \mathcal{B}]} := \mathcal{M}_{[\phi_q, \mathcal{B}]} \quad (8.6)$$

et cette matrice est également appelée la matrice représentant la forme quadratique q dans la base \mathcal{B} .

Les notations étant celles de la définition ci-dessus, on a les résultats élémentaires suivants.

Proposition 8.3.3 (i) La matrice $\mathcal{M}_{[\phi, \mathcal{B}]}$ est l'unique matrice carrée M vérifiant, pour tout couple (\vec{x}, \vec{y}) d'éléments de E ,

$$\phi(\vec{x}, \vec{y}) = {}^t X_{\mathcal{B}} M Y_{\mathcal{B}}, \quad (8.7)$$

où $X_{\mathcal{B}}$ et $Y_{\mathcal{B}}$ désignent les matrices unicolonnes formées des coordonnées de \vec{x} et \vec{y} dans la base \mathcal{B} .

(ii) La matrice $\mathcal{M}_{[q, \mathcal{B}]}$ est l'unique matrice carrée symétrique M vérifiant, pour tout $\vec{x} \in E$,

$$q(\vec{x}) = {}^t X_{\mathcal{B}} M X_{\mathcal{B}}, \quad (8.8)$$

(iii) Si \mathcal{B}_1 est une autre base de E , on a pour les formes bilinéaires symétriques et les formes quadratiques les "formules de changement de base"

$$\mathcal{M}_{[\phi, \mathcal{B}_1]} = {}^t P_{\mathcal{B}, \mathcal{B}_1} \mathcal{M}_{[\phi, \mathcal{B}]} P_{\mathcal{B}, \mathcal{B}_1}. \quad (8.9)$$

$$\mathcal{M}_{[q, \mathcal{B}_1]} = {}^t P_{\mathcal{B}, \mathcal{B}_1} \mathcal{M}_{[q, \mathcal{B}]} P_{\mathcal{B}, \mathcal{B}_1}. \quad (8.10)$$

Démonstration : (i) Soient $\vec{x} = \sum_{i=1}^n x_i \vec{e}_i$ et $\vec{y} = \sum_{i=1}^n y_i \vec{e}_i$ deux éléments de E . On a

$$\begin{aligned} \phi(\vec{x}, \vec{y}) &= \phi \left(\sum_{i=1}^n x_i \vec{e}_i, \sum_{j=1}^n y_j \vec{e}_j \right) = \sum_{\substack{1 \leq i \leq n \\ 1 \leq j \leq n}} x_i y_j \phi(\vec{e}_i, \vec{e}_j) \\ &= \sum_{i=1}^n x_i \left(\sum_{j=1}^n \phi(\vec{e}_i, \vec{e}_j) y_j \right) = {}^t X_{\mathcal{B}} \mathcal{M}_{[\phi, \mathcal{B}]} Y_{\mathcal{B}}. \end{aligned}$$

Si $M = (m_{i,j})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq n}}$ est une matrice vérifiant la condition 8.7, on a, $\{U_1, \dots, U_n\}$ désignant la base canonique de \mathbf{R}^n , pour $1 \leq i \leq n$, $1 \leq j \leq n$,

$$\phi(\vec{e}_i, \vec{e}_j) = {}^t U_i M U_j = m_{i,j},$$

et on voit que $M = \mathcal{M}_{[\phi, \mathcal{B}]}$.

(ii) Il résulte de la définition de $\mathcal{M}_{[q, \mathcal{B}]}$ que $\mathcal{M}_{[q, \mathcal{B}]}$ vérifie la condition 8.8. Réciproquement si une matrice symétrique $M \in \mathcal{M}_n(\mathbf{R})$ vérifie la condition 8.8, posons, pour $\vec{x}, \vec{y} \in \mathbf{R}^n$,

$$\phi(\vec{x}, \vec{y}) = {}^t X_{\mathcal{B}} M Y_{\mathcal{B}}.$$

Comme $\phi(\vec{y}, \vec{x}) = {}^t Y_{\mathcal{B}} M X_{\mathcal{B}} = {}^t ({}^t X_{\mathcal{B}} ({}^t M) Y_{\mathcal{B}}) = {}^t X_{\mathcal{B}} M Y_{\mathcal{B}}$, il est clair que ϕ est une forme bilinéaire symétrique sur E , et on a par construction $q(\vec{x}) = \phi(\vec{x}, \vec{x})$ pour tout $\vec{x} \in E$. Donc $\phi = \phi_q$, et il résulte alors de (i) que $M = \mathcal{M}_{[\phi, \mathcal{B}]} = \mathcal{M}_{[q, \mathcal{B}]}$, ce qui achève la démonstration de (ii).

(iii) Soit maintenant \mathcal{B}_1 une autre base de E . Il résulte des propriétés des matrices de changement de base rappelées au Chapitre 2 que l'on a, pour $\vec{x}, \vec{y} \in E$, en posant $P = P_{\mathcal{B}, \mathcal{B}_1}$, $M = \mathcal{M}_{[\phi, \mathcal{B}]}$,

$$X_{\mathcal{B}} = P X_{\mathcal{B}_1}, \quad {}^t (X_{\mathcal{B}}) = {}^t (X_{\mathcal{B}_1}) {}^t P, \quad Y_{\mathcal{B}} = P Y_{\mathcal{B}_1},$$

$$\phi(\vec{x}, \vec{y}) = {}^t X_{\mathcal{B}} M Y_{\mathcal{B}} = {}^t (X_{\mathcal{B}_1}) {}^t P M P Y_{\mathcal{B}_1},$$

et la formule 8.9 résulte de (i). La formule 8.10 est alors une conséquence immédiate de la formule 8.9 ♣

On déduit de la formule 8.7 que si q est une forme quadratique sur E , et si $(q_{i,j})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq n}}$ est la matrice représentant q dans une base $\mathcal{B} = \{\vec{e}_1, \dots, \vec{e}_n\}$ de E , on a, pour $\vec{x} = \sum_{i=1}^n x_i \vec{e}_i \in E$,

$$q(\vec{x}) = \sum_{\substack{1 \leq i \leq n \\ 1 \leq j \leq n}} q_{i,j} x_i x_j = \sum_{i=1}^n q_{i,i} x_i^2 + 2 \sum_{\substack{1 \leq i \leq n \\ 1 \leq j < i}} q_{i,j} x_i x_j$$

Réciproquement si $q : E \rightarrow \mathbf{R}$ est définie par une formule du type

$$q\left(\sum_{i=1}^n x_i \vec{e}_i\right) = \sum_{i=1}^n a_{i,i} x_i^2 + \sum_{\substack{1 \leq i \leq n \\ 1 \leq j < i}} a_{i,j} x_i x_j, \quad (8.11)$$

alors q est une forme quadratique sur E , et la matrice $\mathcal{M}_{[q,\mathcal{B}]}$ associée à q dans la base \mathcal{B} est définie par la formule

$$\mathcal{M}_{[q,\mathcal{B}]} = (q_{i,j})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq n}},$$

avec $q_{i,i} = a_{i,i}$, $q_{j,i} = q_{i,j} = \frac{a_{i,j}}{2}$ pour $1 \leq i \leq n, 1 \leq j < i$. (8.12)

Autrement dit la forme bilinéaire symétrique ϕ_q associée à q est donnée par la formule

$$\phi\left(\sum_{i=1}^n x_i \vec{e}_i, \sum_{i=1}^n y_i \vec{e}_i\right) = \frac{1}{2} \sum_{\substack{1 \leq i \leq n \\ 1 \leq j \leq n}} a_{i,j} x_i y_j + \frac{1}{2} \sum_{i=1}^n a_{i,i} x_i y_i, \quad (8.13)$$

avec $a_{i,j} = a_{j,i}$ pour $1 \leq i \leq n, 1 \leq j \leq n, i < j$.

On voit donc que, comme on l'avait indiqué plus haut, les formes quadratiques sur \mathbf{R}^n sont tout simplement les polynômes homogènes de degré 2 de n variables à coefficients réels.

Nous illustrons les formules ci-dessus avec l'exemple suivant.

Exemple 8.3.4 La fonction $q : (x, y, z) \mapsto x^2 + y^2 + z^2 - xy - yz - zx$ est une forme quadratique sur \mathbf{R}^3 , et sa matrice associée dans la base canonique de \mathbf{R}^3

est la matrice $\frac{1}{2} \begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix}$.

On va maintenant introduire la notion de rang d'une forme quadratique sur un espace vectoriel réel de dimension finie.

Proposition 8.3.5 Soit q une forme quadratique sur un espace vectoriel réel de dimension finie $n \geq 1$, et soit ϕ_q la forme bilinéaire symétrique sur E associée à q . Alors l'ensemble $N(q) := \{\vec{x} \in E \mid \phi_q(\vec{x}, \vec{y}) = 0 \ \forall \vec{y} \in E\}$ est un sous-espace vectoriel de E . De plus si \mathcal{B} est une base de E , et si $M = \mathcal{M}_{[q,\mathcal{B}]}$ est la

matrice représentant q dans la base \mathcal{B} , alors le rang de la matrice M est égal à $n - \dim(N(q))$.

Le nombre $n - \dim(N(q))$ est appelé le **rang** de la forme quadratique q , et ce nombre est noté $\text{rang}(q)$.

Démonstration : Pour $\vec{y} \in E$, posons $F_{\vec{y}} = \{\vec{x} \in E \mid \phi_q(\vec{x}, \vec{y}) = 0\}$. Alors $F_{\vec{y}}$ est un sous-espace vectoriel de E pour tout $\vec{y} \in E$, et $N(q) = \bigcap_{\vec{y} \in E} F_{\vec{y}}$ est également un sous-espace vectoriel de E .

Soit maintenant \mathcal{B} une base de E et soit $M = \mathcal{M}_{[q, \mathcal{B}]}$ la matrice représentant q dans la base \mathcal{B} . Pour $\vec{x} \in E$ notons $X_{\mathcal{B}}$ la matrice unicolonne formée des coordonnées de \vec{x} dans la base \mathcal{B} . On a pour $\vec{x}, \vec{y} \in E$,

$$\phi_q(\vec{x}, \vec{y}) = \phi_q(\vec{y}, \vec{x}) = {}^t Y_{\mathcal{B}} M X_{\mathcal{B}}.$$

On voit donc que $\vec{x} \in N(q)$ si et seulement si $\langle Y, M X_{\mathcal{B}} \rangle = 0 \quad \forall Y \in \mathbf{R}^n$, l'application $(X, Y) \mapsto \langle X, Y \rangle$ désignant le produit scalaire usuel sur \mathbf{R}^n . Comme $[\mathbf{R}^n]^{\perp} = \{0\}$ on voit que $\vec{x} \in N(q)$ si et seulement si $X_{\mathcal{B}} \in \text{Ker}(M) := \{X \in \mathbf{R}^n \mid M X = 0\}$. Il résulte du théorème de la dimension pour les applications linéaires du Chapitre 2, appliqué à l'application $X \mapsto M X$ sur \mathbf{R}^n , que $\dim(\text{Ker}(M)) + \text{rang}(M) = \dim(\mathbf{R}^n) = n$. Comme l'application $\vec{x} \mapsto X_{\mathcal{B}}$ est un isomorphisme linéaire de E sur \mathbf{R}^n , on obtient $\dim(N(q)) = \dim(\text{Ker}(M)) = n - \text{rang}(M)$, ce qui achève la démonstration. ♣

Exemple 8.3.6 La forme quadratique q de l'exemple 8.3.4 est de rang 2.

En effet on sait que $\text{rang}(q) = \text{rang}(A)$, où $A = \frac{1}{2} \begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix}$. Comme la somme des trois colonnes de A est nulle, $\det(A) = 0$, et $\text{rang}(A) \leq 2$. D'autre part $B := \frac{1}{2} \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}$ est une matrice carrée extraite de A , et $\det(B) = \frac{3}{2} \neq 0$, donc $\text{rang}(A) \geq 2$. On voit donc que $\text{rang}(q) = \text{rang}(A) = 2$.

On s'intéresse maintenant à la diagonalisation des formes quadratiques sur un espace vectoriel réel E de dimension finie. Plus précisément on introduit la notion suivante.

Définition 8.3.7 Soit Q une forme quadratique sur un espace vectoriel réel $E \neq \{0\}$ de dimension finie. On dit qu'une base \mathcal{B} de E est **diagonalisante** pour q quand la matrice $\mathcal{M}_{[q, \mathcal{B}]}$ représentant q dans la base \mathcal{B} est diagonale.

On sait que de telles bases existent : si $\mathcal{B}_0 = \{\vec{e}_1, \dots, \vec{e}_n\}$ est une base quelconque de E , la matrice $M_0 := \mathcal{M}_{[q, \mathcal{B}_0]}$ représentant q dans la base \mathcal{B}_0 est symétrique, donc il existe une matrice orthogonale $P = (p_{i,j})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq n}} \in \mathcal{M}_n(\mathbf{R})$ telle que ${}^t P M_0 P = P^{-1} M_0 P$ soit diagonale. Posons $\vec{f}_j = \sum_{i=1}^n p_{i,j} \vec{e}_i$ pour $1 \leq j \leq n$,

et soit $\mathcal{B} := \{\vec{f}_1, \dots, \vec{f}_n\}$. On a alors $P_{\mathcal{B}_0, \mathcal{B}} = P$, donc \mathcal{B} est une base de E , et on a

$$\mathcal{M}_{[q, \mathcal{B}]} = {}^t P M_0 P = P^{-1} M_0 P,$$

qui est diagonale, donc la base \mathcal{B} est diagonalisante pour q .

La diagonalisation d'une matrice symétrique nécessite le calcul exact des valeurs propres de cette matrice, ce qui n'est pas possible dans $\mathcal{M}_n(\mathbf{R})$ pour $n \geq 5$ (et en général fort compliqué dans $\mathcal{M}_3(\mathbf{R})$ et $\mathcal{M}_4(\mathbf{R})$). Nous verrons plus loin un procédé plus simple, qui utilise une idée de Gauss, et qui permet d'obtenir des bases diagonalisantes associées à des matrices de passage non nécessairement orthogonales. Nous allons tout d'abord dégager une propriété commune à toutes les bases de E diagonalisantes pour une forme quadratique donnée q sur E .

Théorème 8.3.8 (théorème d'inertie de Sylvester) *Soit $E \neq \{0\}$ un espace vectoriel réel de dimension finie, et soit q une forme quadratique sur E . Si \mathcal{B} est une base de E diagonalisante pour q , on note respectivement $n^+(q)_{\mathcal{B}}$ et $n^-(q)_{\mathcal{B}}$ le nombre de coefficients strictement positifs et strictement négatifs de la matrice $\mathcal{M}_{[q, \mathcal{B}]}$ représentant q dans la base \mathcal{B} .*

Alors $n^+(q)_{\mathcal{B}}$ et $n^-(q)_{\mathcal{B}}$ sont indépendants de la base diagonalisante \mathcal{B} .

*Le couple $(n^+(q), n^-(q)) := (n^+(q)_{\mathcal{B}}, n^-(q)_{\mathcal{B}})$ est appelé la **signature** de la forme quadratique q , et on a $\text{rang}(q) := n^+(q) + n^-(q)$.*

Démonstration : Soit $n = \dim(E)$, soit \mathcal{F} l'ensemble des sous-espaces vectoriels F de E tels que $q(\vec{x}) > 0$ pour tout $\vec{x} \in F \setminus \{0\}$, avec par convention $\{0\} \in \mathcal{F}$, et soit \mathcal{G} l'ensemble des sous-espaces vectoriels G de E tels que $q(\vec{x}) \leq 0$ pour tout $x \in G$. Soit $F \in \mathcal{F}$, et soit $G \in \mathcal{G}$. Si $\vec{x} \in F \cap G$, on a $q(\vec{x}) \geq 0$ et $q(\vec{x}) \leq 0$, donc $q(\vec{x}) = 0$, et $\vec{x} = 0$ puisque $\vec{x} \in F$. Donc $F \cap G = \{0\}$.

On a alors, d'après le "théorème de la dimension" vu au Chapitre 1

$$\dim(E) \geq \dim(F + G) = \dim(F) + \dim(G) - \dim(F \cap G) = \dim(F) + \dim(G). \quad (8.14)$$

Si \mathcal{B} est une base de E diagonalisante pour E , on pose $\mathcal{B}_+ = \{\vec{x} \in \mathcal{B} \mid q(\vec{x}) > 0\}$, $\mathcal{B}_- = \{\vec{x} \in \mathcal{B} \mid q(\vec{x}) \leq 0\}$, et on note respectivement $F_{\mathcal{B}}$ et $G_{\mathcal{B}}$ les sous-espaces vectoriels de E engendrés par \mathcal{B}_+ et \mathcal{B}_- , avec la convention que l'espace vectoriel engendré par l'ensemble vide est l'espace vectoriel réduit à $\{0\}$. Si $\mathcal{B}_+ \neq \emptyset$, la matrice associée à la restriction de q à $F_{\mathcal{B}}$ pour la base \mathcal{B}_+ de $F_{\mathcal{B}}$ est diagonale à coefficients diagonaux strictement positifs, on voit que $F_{\mathcal{B}} \in \mathcal{F}$, et $F_{\mathcal{B}} = \{0\} \in \mathcal{F}$ si \mathcal{B}_+ est vide. On voit de même que $G_{\mathcal{B}} \in \mathcal{G}$, et il est clair que $E = F_{\mathcal{B}} \oplus G_{\mathcal{B}}$, et $\dim(F_{\mathcal{B}}) + \dim(G_{\mathcal{B}}) = n$.

Soient maintenant \mathcal{B} et \mathcal{B}' deux bases de E diagonalisantes pour q . Il résulte de la formule 8.14 que l'on a

$$\dim(F_{\mathcal{B}}) \leq n - \dim(G'_{\mathcal{B}}) = \dim(F'_{\mathcal{B}}),$$

et en inversant les rôles de \mathcal{B} et \mathcal{B}' on voit qu'en fait $n^+(q)_{\mathcal{B}} = \dim(F_{\mathcal{B}}) = \dim(F'_{\mathcal{B}}) = n^+(q)_{\mathcal{B}'}$. On a alors

$$n^-(q)_B = n^+(-q)_B = n^+(-q)_{B'} = n^-(q)_B.$$

Le fait que $\text{rang}(q) = n^+(q) + n^-(q)$ résulte alors du fait que le rang d'une matrice diagonale ayant exactement p coefficients diagonaux non nuls est égal à p . ♣

On va maintenant introduire l'algorithme de Gauss, qui permet d'obtenir facilement des bases diagonalisantes pour les formes quadratiques définies sur un espace vectoriel réel de dimension finie. Nous allons introduire tout d'abord la notion de forme linéaire.

Définition 8.3.9 Une forme linéaire sur un espace vectoriel réel est une application linéaire de E dans \mathbf{R} .

L'ensemble $L(E, \mathbf{R})$ des formes linéaires sur E est un espace vectoriel pour les lois usuelles sur les applications linéaires, introduites au Chapitre 1, et sa dimension est égale à $\dim(E)\dim(\mathbf{R}) = \dim(E)$ si E est de dimension finie. Le produit fg de deux formes linéaires f et g sur E est défini pour $\vec{x} \in E$ par la formule

$$(fg)(\vec{x}) = f(\vec{x})g(\vec{x}).$$

Remarquons qu'une somme finie de produits de formes linéaires sur E définit une forme quadratique sur E , voir l'exercice 3.

L'algorithme de Gauss est basé sur les observations suivantes.

Proposition 8.3.10 Soient f_1, \dots, f_n des formes linéaires linéairement indépendantes sur un espace vectoriel E , et soit $(a_{i,j})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq i}}$ une famille de nombres réels. On pose,

$$q = \sum_{\substack{1 \leq i \leq n \\ 1 \leq j \leq i}} a_{i,j} f_i f_j.$$

(i) Si $a_{k,k} = a_{l,l} = 0$, et si $a_{k,l} \neq 0$, avec $l > k$, posons $g_k = \frac{f_k - a_{k,l} f_l}{2}$, $g_l = \frac{f_k + a_{k,l} f_l}{2}$, $g_i = f_i$ pour $1 \leq i \leq n$, $i \neq k, i \neq l$. Alors les formes linéaires g_1, \dots, g_n sont linéairement indépendantes, et on a

$$q = \sum_{\substack{1 \leq i \leq n \\ 1 \leq j \leq i}} b_{i,j} g_i g_j,$$

avec $b_{i,j} \in \mathbf{R}$ pour $1 \leq i \leq n$, $b_{k,k} = -1$, $b_{l,l} = 1$.

(ii) Si $a_{k,k} \neq 0$, posons $\epsilon = \frac{a_{k,k}}{|a_{k,k}|}$,

$$h_k = \sqrt{|a_{k,k}|} f_k + \frac{\epsilon}{2\sqrt{|a_{k,k}|}} \left(\sum_{1 \leq i < k} a_{k,i} f_i + \sum_{k < i \leq n} a_{i,k} f_i \right),$$

$h_i = f_i$ pour $1 \leq i \leq n, i \neq k$. Alors les formes linéaires h_1, \dots, h_n sont linéairement indépendantes, et on a

$$q = \sum_{\substack{1 \leq i \leq n \\ 1 \leq j \leq i}} c_{i,j} h_i h_j,$$

avec $c_{i,j} \in \mathbf{R}$ pour $1 \leq i \leq n, c_{k,k} = \epsilon, c_{k,i} = c_{j,k} = 0$ pour $1 \leq i < k, k < j \leq n$.

Démonstration : (i) Comme $a_{k,l} \neq 0$ les espaces vectoriels engendrés par $\{f_k, f_l\}$ et $\{g_k, g_l\}$ coïncident, donc l'espace vectoriel engendré par la famille $\{g_1, \dots, g_n\}$ est de dimension n , ce qui montre que cette famille est libre. Le fait que $b_{k,k} = 1$ et que $b_{l,l} = -1$ résulte du fait que $a_{k,k} = a_{l,l} = 0$ et de l'identité

$$g_l^2 - g_k^2 = \left(\frac{f_k + a_{k,l} f_l}{2} \right)^2 - \left(\frac{f_k - a_{k,l} f_l}{2} \right)^2 = a_{k,l} f_k f_l.$$

(ii) Il est clair que f_k appartient à l'espace vectoriel engendré par $\{h_1, \dots, h_n\} = \{h_k\} \cup \{f_i\}_{\substack{1 \leq i \leq n \\ i \neq k}}$, ce qui entraîne de même que plus haut que la famille $\{h_1, \dots, h_n\}$ est libre. Le fait que $c_{k,k} = \epsilon$ et que $b_{k,i} = b_{j,k} = 0$ pour $1 \leq i < k, k < j \leq n$ provient de l'identité

$$a_{k,k} f_k^2 + \sum_{1 \leq i < k} a_{k,i} f_k f_i + \sum_{k < i < n} a_{i,k} f_k f_i = \epsilon h_k^2 - \frac{\epsilon}{4|a_{k,k}|} \left(\sum_{1 \leq i < k} a_{k,i} f_j + \sum_{k < i \leq n} a_{i,k} f_i \right)^2.$$



Soit E un espace vectoriel réel de dimension $n \geq 1$, soit $\mathcal{B} = \{\vec{e}_1, \dots, \vec{e}_n\}$ une base de E et pour $1 \leq i \leq n$ soit e_i^* la forme linéaire sur E qui à $\vec{x} \in E$ associe sa i^e coordonnée dans la base \mathcal{B} . Alors $\mathcal{B}^* := \{e_1^*, \dots, e_n^*\}$ est une base de $L(E, \mathbf{R})$, appelée **base duale** de la base \mathcal{B} . Si \mathcal{B}_0 et \mathcal{B}_1 sont deux bases de E , on a alors un lien simple entre les matrices de passage $P_{\mathcal{B}_0, \mathcal{B}_1}$ et $P_{\mathcal{B}_0^*, \mathcal{B}_1^*} = P_{\mathcal{B}_1^*, \mathcal{B}_0^*}^{-1}$.

Proposition 8.3.11 *Soit E un espace vectoriel réel de dimension finie $n \geq 1$, soient \mathcal{B}_0 et \mathcal{B}_1 deux bases de E , et soient \mathcal{B}_0^* et \mathcal{B}_1^* les bases duales associées à \mathcal{B}_0 et \mathcal{B}_1 . On a alors*

$$P_{\mathcal{B}_1^*, \mathcal{B}_0^*} = {}^t P_{\mathcal{B}_0, \mathcal{B}_1} = {}^t P_{\mathcal{B}_1, \mathcal{B}_0}^{-1}. \quad (8.15)$$

Démonstration : Soit $\vec{x} \in E$. On a, en posant $\mathcal{B}_0 = \{\vec{e}_1, \dots, \vec{e}_n\}, \mathcal{B}_1 = \{\vec{f}_1, \dots, \vec{f}_n\}$,

$$\vec{x} = \sum_{j=1}^n e_j^*(\vec{x}) \vec{e}_j = \sum_{j=1}^n f_j^*(\vec{x}) \vec{f}_j.$$

Donc si on note X_0 la matrice unicolonne formée des coordonnées de \vec{x} dans la base \mathcal{B}_0 et X_1 la matrice unicolonne formée des coordonnées de \vec{x} dans la

base \mathcal{B}_1 , il résulte de la Proposition 2.3.2 du Chapitre 2 que l'on a, en posant

$$P_{\mathcal{B}_0, \mathcal{B}_1} = (p_{i,j})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq n}},$$

$$\begin{bmatrix} e_1^*(\vec{x}) \\ \cdot \\ \cdot \\ e_n^*(\vec{x}) \end{bmatrix} = X_0 = P_{\mathcal{B}_0, \mathcal{B}_1} X_1 = P_{\mathcal{B}_0, \mathcal{B}_1} \begin{bmatrix} f_1^*(\vec{x}) \\ \cdot \\ \cdot \\ f_n^*(\vec{x}) \end{bmatrix}.$$

On voit donc que $e_i^* = \sum_{j=1}^n p_{i,j} f_j^*$, et la i^e ligne de $P_{\mathcal{B}_0, \mathcal{B}_1}$ est formée des coordonnées de e_i^* dans la base \mathcal{B}_1^* . Par conséquent $P_{\mathcal{B}_1^*, \mathcal{B}_0^*} = {}^t P_{\mathcal{B}_0, \mathcal{B}_1} = {}^t P_{\mathcal{B}_1, \mathcal{B}_0}^{-1}$.



Considérons de nouveau une forme quadratique q sur un espace vectoriel réel E de dimension finie $n \geq 1$. Si $(q_{i,j})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq n}}$ est la matrice représentant q dans une base \mathcal{B} , la formule 8.8 prend la forme

$$q = \sum_{\substack{1 \leq i \leq n \\ 1 \leq j \leq n}} q_{i,j} e_i^* e_j^*,$$

et on voit qu'en appliquant successivement au plus $2n$ fois la proposition, on obtient une base $\{f_1, \dots, f_n\}$ de $L(E, \mathbf{R})$ et une famille $\epsilon_1, \dots, \epsilon_n$ de réels telles que l'on ait

$$q = \sum_{i=1}^n \epsilon_i f_i^2,$$

avec $\epsilon_i \in \{-1, 0, 1\}$ pour $1 \leq i \leq n$.

Soit $A = \{i \in \{1, \dots, n\} \mid \epsilon_i = 1\}$, soit $B = \{i \in \{1, \dots, n\} \mid \epsilon_i = -1\}$, soit a le nombre d'éléments de A et soit b le nombre d'éléments de B . Alors la signature de q est égale à (a, b) . On a en effet le résultat suivant.

Proposition 8.3.12 *Soit E un espace vectoriel réel de dimension finie $n \geq 1$, soit $\mathcal{B} := \{\vec{e}_1, \dots, \vec{e}_n\}$ une base de E , et soit $\mathcal{F} = \{f_1, \dots, f_n\}$ une base de $L(E, \mathbf{R})$. Alors il existe pour $1 \leq i \leq n$ un unique élément \vec{u}_i de E tel que l'on ait*

$$f_i(\vec{u}_i) = 1 \quad f_j(\vec{u}_i) = 0 \quad \text{pour } 1 \leq j \leq n, j \neq i,$$

et $\mathcal{U} := \{\vec{u}_1, \dots, \vec{u}_n\}$ est une base de E vérifiant $\mathcal{U}^* = \mathcal{F}$.

De plus si $q = \sum \epsilon_i f_i^2$, la matrice $M_{[q, \mathcal{U}]}$ représentant q dans la base \mathcal{U} est la matrice diagonale $(d_{i,j})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq n}}$ vérifiant $d_{i,i} = \epsilon_i$ pour $1 \leq i \leq n$.

Démonstration : Posons $P := P_{\mathcal{B}^*, \mathcal{F}}$, $P^{-1} = Q = (q_{i,j})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq n}}$, et pour $1 \leq i \leq n$ posons $\vec{u}_i = \sum_{j=1}^n q_{i,j} \vec{e}_j$. Si $\mathcal{U} := \{\vec{u}_1, \dots, \vec{u}_n\}$, alors $P_{\mathcal{B}, \mathcal{U}} = Q = P^{-1}$ qui est inversible, donc \mathcal{U} est une base de E , et on a $P_{\mathcal{B}^*, \mathcal{U}^*} = Q^{-1} = P = P_{\mathcal{B}^*, \mathcal{F}}$, donc $\mathcal{U}^* = \mathcal{F}$. D'autre part si $\vec{x} \in E$, alors $\vec{x} = \sum_{j=1}^n u_j^*(\vec{x}) \vec{u}_j = \sum_{j=1}^n f_j(\vec{x}) \vec{u}_j$, donc \vec{u}_i est l'unique $\vec{x} \in E$ tel que $f_i(\vec{u}_i) = 1$, $f_j(\vec{u}_i) = 0$ pour $1 \leq j \leq n, j \neq i$.

Soit $D = (d_{i,j})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq n}}$ la matrice diagonale vérifiant $d_{i,i} = 1$ pour $1 \leq i \leq n$. On a $q = \sum_{j=1}^n \epsilon_j f_j^2 = \sum_{j=1}^n \epsilon_j u_j^{*2}$. Il résulte alors de la Proposition 8.3.3 que la matrice $\mathcal{M}[q, \mathcal{U}]$ représentant q dans la base \mathcal{U} est égale à D . ♣

Exemple 8.3.13 *Calcul de la signature de la forme quadratique $q : (x, y, z) \mapsto xy + yz + zx$ par la méthode de Gauss.*

On est dans le second cas de la proposition 8.3.8 et on pose $f(x, y, z) = \frac{x+y}{4}$, $g(x, y, z) = \frac{x-y}{4}$. on obtient

$$q(x, y) = f^2(x, y) - g^2(x, y) + 4zf(x, y).$$

On est alors dans le premier cas de la proposition 8.3.8 et on pose $h(x, y, z) = f(x, y) + 2z$, $l(x, y, z) = 2z$, et on obtient la décomposition cherchée

$$q = h^2 - g^2 - l^2,$$

avec $h(x, y, z) = \frac{x+y}{4} + 2z$, $h(x, y, z) = \frac{x+y}{4} - 2z$, $l(x, y, z) = 2z$.

On sait d'après la proposition 8.2.9 que les formes linéaires h, k et l sont linéairement indépendantes (ceci peut aussi se vérifier très facilement directement) et il résulte alors du théorème 8.3.8 que la signature de q est égale à $(1, 2)$. En particulier la forme quadratique q est de rang 3.

8.4 Matrices définies positives

On va maintenant s'intéresser à un type particulier de matrices symétriques.

Proposition 8.4.1 *Soient $A \in \mathcal{M}_n(\mathbf{R})$ une matrice symétrique. Alors les conditions suivantes sont équivalentes.*

- (i) L'application $\phi_A : (X, Y) \mapsto {}^t XAY$ est un produit scalaire sur \mathbf{R}^n .
- (ii) ${}^t XAX > 0$ pour tout $X \in \mathbf{R}^n \setminus \{0\}$.
- (iii) Toutes les valeurs propres de A sont strictement positives.
- (iv) La signature de la forme quadratique $q_A : X \mapsto {}^t XAX$ est égale à $(n, 0)$.

Dans ce cas on dit que la matrice A est une **matrice définie positive**.

Démonstration : On sait que ϕ_A est une forme bilinéaire symétrique sur \mathbf{R}^n , que toutes les valeurs propres de A sont réelles, et qu'il existe une matrice orthogonale P telle que ${}^t PAP = P^{-1}AP$ soit diagonale. Comme l'application $X \mapsto {}^t XAX$ est la forme quadratique associée à ϕ_A , (i) et (ii) sont équivalents. Comme la matrice ${}^t PAP$ est semblable à A , il résulte de la définition de la signature d'une forme quadratique que (iii) et (iv) sont équivalents.

Posons maintenant $X' = \begin{bmatrix} x'_1 \\ \cdot \\ \cdot \\ \cdot \\ x'_n \end{bmatrix} := PX$ pour $X = \begin{bmatrix} x_1 \\ \cdot \\ \cdot \\ \cdot \\ x_n \end{bmatrix} \in \mathbf{R}^n$. L'application $X \mapsto X'$ est une isométrie de \mathbf{R}^n sur lui-même, et si $\lambda_1, \dots, \lambda_n$ désignent les valeurs propres de A , on a, pour $X \in \mathbf{R}^n$,

$$\phi_A(X, X) = \sum_{j=1}^n \lambda_j x_j'^2$$

et on voit que $\phi_A(X, X) > 0$ pour tout $X \neq 0$ si et seulement si toutes les valeurs propres de A sont strictement positives. ♣

On va maintenant introduire la **décomposition de Cholesky** d'une matrice définie positive, fort utile en Analyse numérique. Rappelons qu'on dit qu'une matrice carrée $A = (a_{i,j})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq n}}$ est *triangulaire supérieure* si $a_{i,j} = 0$ pour $i > j$. On dit de même qu'une matrice carrée $A = (a_{i,j})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq n}}$ est *triangulaire inférieure* si $a_{i,j} = 0$ pour $i < j$.

Théorème 8.4.2 (*décomposition de Cholesky*)

Toute matrice définie positive $A \in \mathcal{M}_n(\mathbf{R})$ peut s'écrire sous la forme

$$A = {}^t T T,$$

où $T = (t_{i,j})$ est une matrice triangulaire supérieure telle que $t_{i,i} > 0$ pour $1 \leq i \leq n$.

Réciproquement si $S \in \mathcal{M}_n(\mathbf{R})$ est une matrice triangulaire dont tous les termes diagonaux sont non nuls, ou plus généralement si $S \in \mathcal{M}_n(\mathbf{R})$ est inversible, alors ${}^t S S$ et $S {}^t S$ sont définies positives.

Démonstration : Supposons que A est définie positive, et posons, pour $X \in \mathbf{R}^n$,

$$q(X) = {}^t X A X.$$

Soient f_1, \dots, f_n des formes linéaires linéairement indépendantes sur \mathbf{R}^n telles que

$$q = \sum_{\substack{1 \leq i \leq n \\ 1 \leq j \leq i}} b_{i,j} f_i f_j.$$

On a vu à la proposition 3.8.9 qu'il existe une base $\mathcal{U} = \{u_1, \dots, u_n\}$ de \mathbf{R}^n telle que $f_i = u_i^*$ pour $1 \leq i \leq n$. Donc $b_{i,i} > 0$ pour $1 \leq i \leq n$.

$$[e_1^* \quad \dots \quad e_n^*]^t T = [f_1 \quad \dots \quad f_n],$$

$$q = \sum_{j=1}^n f_j^2 = [f_1 \quad \dots \quad f_n] \begin{bmatrix} f_1 \\ \vdots \\ f_n \end{bmatrix} = [e_1^* \quad \dots \quad e_n^*]^t T T \begin{bmatrix} e_1^* \\ \vdots \\ e_n^* \end{bmatrix}.$$

Autrement dit on a, pour $X \in \mathbf{R}^n$,

$$q(X) = {}^t X ({}^t T T) X = {}^t X A X.$$

Comme A et ${}^t T T$ sont symétriques, on a $A = {}^t T T$, ce qui donne la décomposition cherchée.

Soit maintenant $M \in \mathcal{M}_n(\mathbf{R})$ une matrice inversible. On a, pour $X \in \mathbf{R}^n \setminus \{0\}$, puisque $MX \neq 0$,

$${}^t X ({}^t M M) X = {}^t (M X) M X = \|M X\|_2 > 0,$$

et il résulte de la proposition 8.4.2 que la matrice symétrique ${}^t M M$ est définie positive.



Exemple 8.4.3 Décomposition de Cholesky de la matrice $A = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix}$.

On considère la forme quadratique $q(x, y, z) = 2(x^2 + y^2 + z^2 + xy + yz + zx)$, à laquelle on applique l'algorithme de Gauss (il est inutile de vérifier que la matrice symétrique en question est définie positive car l'algorithme aboutira si et seulement si elle l'est).

$$\begin{aligned} q(x, y, z) &= 2\left(x + \frac{y}{2} + \frac{z}{2}\right)^2 - \frac{y^2}{2} - \frac{z^2}{2} - yz + 2y^2 + 2z^2 + 2yz \\ &= 2\left(x + \frac{y}{2} + \frac{z}{2}\right)^2 + \frac{3y^2}{2} + \frac{3z^2}{2} + yz = 2\left(x + \frac{y}{2} + \frac{z}{2}\right)^2 + \frac{3}{2}\left(y + \frac{z}{3}\right)^2 - \frac{z^2}{6} + \frac{3z^2}{2} \\ &= 2\left(x + \frac{y}{2} + \frac{z}{2}\right)^2 + \frac{3}{2}\left(y + \frac{z}{3}\right)^2 + \frac{4}{3}z^2 \\ &= \left(\sqrt{2}x + \frac{y}{\sqrt{2}} + \frac{z}{\sqrt{2}}\right)^2 + \left(\sqrt{\frac{3}{2}}y + \frac{z}{\sqrt{6}}\right)^2 + \left(\frac{2z}{\sqrt{3}}\right)^2. \end{aligned}$$

On obtient alors

$$\begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix} = \begin{bmatrix} \sqrt{2} & 0 & 0 \\ \frac{1}{\sqrt{2}} & \sqrt{\frac{3}{2}} & 0 \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{2}{\sqrt{3}} \end{bmatrix} \begin{bmatrix} \sqrt{2} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 0 & \sqrt{\frac{3}{2}} & \frac{1}{\sqrt{6}} \\ 0 & 0 & \frac{2}{\sqrt{3}} \end{bmatrix}.$$

La décomposition de Cholesky est très intéressante pour les calculs effectifs de solutions de systèmes linéaires de n équations à n variables. En effet si $A \in \mathcal{M}_n(\mathbf{R})$ est définie positive, et si $A = {}^tMM$ est une décomposition de Cholesky de A , alors étant donné $Y \in \mathbf{R}^n$, pour résoudre le système linéaire $AX = Y$ on résout successivement les systèmes ${}^tMZ = Y$ et $MX = Z$, ce qui est économique au niveau des calculs puisque ces deux systèmes sont associés à des matrices triangulaires.

8.5 Applications aux coniques et quadriques

On va maintenant s'intéresser aux applications de l'étude des matrices symétriques à la géométrie. On va introduire les notions suivantes.

Définition 8.5.1 (i) Soit \mathcal{P} le plan affine euclidien usuel, rapporté à un repère orthonormé $(0, \vec{i}, \vec{j})$. On appelle coniques les courbes de \mathcal{P} qui ont une équation de la forme

$$ax^2 + bxy + cy^2 + dx + ey + f = 0,$$

où $(a, b, c) \neq (0, 0, 0)$.

(ii) Soit \mathcal{E} l'espace affine euclidien usuel, rapporté à un repère orthonormé $(0, \vec{i}, \vec{j}, \vec{k})$. On appelle quadriques les surfaces de \mathcal{P} qui ont une équation de la forme

$$ax^2 + by^2 + cz^2 + dxy + eyz + fzx + gx + hy + kz + l = 0,$$

avec $(a, b, c, d, e, f) \neq (0, 0, 0, 0, 0, 0)$.

On rappelle que si $f : \mathbf{R}^n \rightarrow \mathbf{R}$ est une fonction de n variables, la dérivée partielle $\frac{\partial f}{\partial x_j}(U)$ en $U = (u_1, \dots, u_n)$ pour $1 \leq j \leq n$ est quand elle existe la dérivée en u_j de la fonction $x \mapsto f(u_1, \dots, u_{j-1}, x, u_{j+1}, \dots, u_n)$. On définit de même par récurrence quand elles existent les dérivées partielles successives $\frac{\partial^k f}{\partial x_{i_1} \dots \partial x_{i_k}}(U)$. On sait en particulier que si f admet des dérivées partielles d'ordre 2 continues, alors $\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i}$ pour $1 \leq i \leq n, 1 \leq j \leq n$, ce qui se vérifie facilement directement dans le cas des polynômes. On introduit alors le gradient

$$\nabla(f)(U) := \left[\frac{\partial f}{\partial x_1}(U), \dots, \frac{\partial f}{\partial x_n}(U) \right],$$

ainsi que la hessienne

$$\text{Hess}(f)(U) = \left(\frac{\partial^2 f}{\partial x_i \partial x_j}(U) \right)_{\substack{1 \leq i \leq n \\ 1 \leq j \leq n}},$$

qui est toujours une matrice symétrique quand f admet des dérivées partielles d'ordre ≤ 2 continues.

Considérons maintenant une fonction polynômiale f de n variables de degré inférieur ou égal à 2 par rapport à l'ensemble des variables. Alors f est la somme

$$\text{d'une forme quadratique } q : X = \begin{bmatrix} x_1 \\ \cdot \\ \cdot \\ \cdot \\ x_n \end{bmatrix} \mapsto {}^t X A X, \text{ où } A = (a_{i,j})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq n}} \in$$

$\mathcal{M}_n(\mathbf{R})$ est symétrique, d'une forme linéaire $l : X \mapsto b_1 x_1 + \dots + b_n x_n = \langle$

$$x, b \rangle = X^t B = {}^t B X, \text{ où } B := \begin{bmatrix} b_1 \\ \cdot \\ \cdot \\ \cdot \\ b_n \end{bmatrix} \in \mathbf{R}^n, \text{ et d'une constante } c \in \mathbf{R}. \text{ On a}$$

dans ce cas une version très simple de la formule de Taylor en plusieurs variables.

Proposition 8.5.2 *Soit $f : X \mapsto X A^t X + X^t B + c$, avec $A \in \mathcal{M}_2(\mathbf{R})$ symétrique, $B \in \mathbf{R}^n$, $c \in \mathbf{R}$ un polynôme de n variables de degré inférieur ou égal à 2. On a alors, pour $U = (u_1, \dots, u_n)$ et $H = (h_1, \dots, h_n) \in \mathbf{R}^n$,*

$$\begin{aligned} f(U + H) &= c + H^t B + 2 H A^t U + H A^t H \\ &= f(U) + \langle H, \nabla(f)(U) \rangle + \frac{H \text{Hess}(f)(U)^t H}{2}. \end{aligned} \quad (8.16)$$

Démonstration : Comme A est symétrique, on a $U A^t H = {}^t (U A^t H) = H^t A^t U = H A^t U$, et on obtient,

$$\begin{aligned} f(U + H) &= c + (U + H)^t B + (U + H) A^t (U + H) = c + U^t B + H^t B + U A^t U + H A^t U + U A^t H + H A^t H \\ &= f(U) + H^t B + 2 H A^t U + H A^t H. \end{aligned}$$

D'autre part en posant $A = (a_{i,j})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq n}}$, $B = (b_1, \dots, b_n)$, on obtient

$$f(x_1, \dots, x_n) = c + \sum_{i=1}^n x_i b_i + \sum_{\substack{1 \leq i \leq n \\ 1 \leq j \leq n}} a_{i,j} x_i x_j.$$

Dans la somme ci-dessus, la somme des termes faisant intervenir x_i est égale à $x_i b_i + a_{i,i} x_i^2 + 2 x_i \sum_{\substack{1 \leq j \leq n \\ j \neq i}} a_{i,j} x_j$. on obtient, pour $1 \leq i \leq n$

$$\frac{\partial f}{\partial x_i}(x_1, \dots, x_n) = b_i + 2 a_{i,i} x_i + 2 \sum_{\substack{1 \leq j \leq n \\ j \neq i}} a_{i,j} x_j = b_i + 2 x_i \sum_{j=1}^n a_{i,j} x_j.$$

On a alors, pour $1 \leq i \leq n$, $1 \leq j \leq n$,

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(x_1, \dots, x_n) = 2a_{i,j}.$$

Donc $\nabla(f)(U) = B + 2AU = B + 2^t AU$, $Hess(f)(U) = 2A$, ce qui achève la démonstration. ♣

On va maintenant se restreindre aux dimensions 2 et 3, qui sont les seules où on dispose d'une vision concrète géométrique. Les notations étant les mêmes que plus haut, la forme quadratique sur \mathbf{R}^2 associée à une conique \mathcal{C} d'équation $XA^tX + X^tB + c = 0$ est la forme quadratique $q_{\mathcal{C}} : X \mapsto XA^tX$. De même la forme quadratique $q_{\mathcal{Q}}$ associée à une quadrique \mathcal{Q} d'équation $XA^tX + X^tB + c = 0$ est la forme quadratique $q_{\mathcal{Q}} : X \mapsto XA^tX$. Ces formes quadratiques sont bien entendu définies à une constante multiplicative réelle non nulle près, et cette constante n'affecte pas leur rang. Rappelons qu'on dit que $U \in \mathbf{R}^n$ est un **point critique** d'une application $f : \mathbf{R}^n \rightarrow \mathbf{R}$ si f possède des dérivées partielles d'ordre 1 en U , et si $\nabla(f)(U) = (0, \dots, 0)$.

Proposition 8.5.3 *Soit f un polynôme de degré 2 sur \mathbf{R}^n avec $n \in \{2, 3\}$, et soit \mathcal{S} la conique d'équation $f(X) = 0$ dans le plan ou la quadrique d'équation $f(X) = 0$ dans l'espace affine euclidien rapportés à un repère orthonormé. Alors tout point critique de f est un centre de symétrie pour \mathcal{S} .*

Démonstration : On effectue une translation des axes en remplaçant l'origine du repère par U . On a $f(X) = q(X) + l(X) + c$, où q est une forme quadratique sur \mathbf{R}^n , l une forme linéaire sur \mathbf{R}^n et c un réel. Posons $g(X) = f(U + X)$. L'équation de \mathcal{S} dans le nouveau repère devient $g(X) = 0$, et on a, d'après la proposition, pour $X \in \mathbf{R}^n$,

$$g(X) = f(U + X) = f(U) + \langle X, \nabla(f)(U) \rangle + q(X) = f(U) + q(X).$$

En particulier $g(X) = g(-X)$ pour $X \in \mathbf{R}^n$, ce qui montre que U est un centre de symétrie pour \mathcal{S} . ♣

On peut écrire l'équation d'une conique ou d'une quadrique sous la forme $XA^tX + X^tB + c = 0$, avec $A \in \mathcal{M}_n(\mathbf{R})$, symétrique, $B \in \mathbf{R}^n$, $c \in \mathbf{R}$, $n = 2$ ou $n = 3$. Si on pose $f(X) = XA^tX + X^tB + c$, les points critiques U de f sont donnés par l'équation

$$2A^tU + {}^tB = 0, \tag{8.17}$$

qui a toujours une solution unique $U = -\frac{{}^tBA^{-1}}{2}$ si A est inversible.

D'autre part comme A est symétrique, il existe une matrice orthogonale P telle que ${}^tPAP = P^{-1}AP$ soit diagonale. Ceci permet de construire un repère orthonormal dans lequel la forme quadratique associée l'équation de f est représentée par une matrice diagonale.

Ces deux observations permettent de parvenir à une classification complète des coniques.

Théorème 8.5.4 Dans le plan affine euclidien rapporté à un repère orthonormé $(0, \vec{i}, \vec{j})$, soit \mathcal{C} une conique d'équation

$$f(X) := XA^tX + X^tB + c = 0,$$

avec $A \in \mathcal{M}_2(\mathbf{R}^2) \neq 0$, symétrique, $B \in \mathbf{R}^2$, $c \in \mathbf{R}$, soient λ_1 et λ_2 les deux valeurs propres de A , répétées si nécessaire selon leurs ordres de multiplicité,

soit $P = \begin{bmatrix} p_{1,1} & p_{1,2} \\ p_{2,1} & p_{2,2} \end{bmatrix}$ une matrice orthogonale telle que $P^{-1}AP = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$,

et soit $\vec{u} = p_{1,1}\vec{i} + p_{2,1}\vec{j}$, $\vec{v} = p_{1,2}\vec{i} + p_{2,2}\vec{j}$.

(1) Si $\det(A) = \lambda_1\lambda_2 > 0$, alors f admet un unique point critique U , et l'équation de \mathcal{C} dans le repère (U, \vec{u}, \vec{v}) est de la forme

$$\lambda_1x^2 + \lambda_2y^2 = d,$$

ce qui donne 3 possibilités

$$d = 0, \quad \text{et} \quad \mathcal{C} = \{U\}. \quad (8.18)$$

$$d \neq 0 \text{ est de signe opposé à celui de } \lambda_1 \text{ et } \mathcal{C} = \emptyset \quad (8.19)$$

$$d \neq 0 \text{ est de même signe que } \lambda_1 \text{ et } \mathcal{C} \text{ est une ellipse} \quad (8.20)$$

(2) Si $\det(A) = \lambda_1\lambda_2 < 0$, alors f admet un unique point critique U , et l'équation de \mathcal{C} dans le repère (U, \vec{u}, \vec{v}) est de la forme

$$\lambda_1x^2 + \lambda_2y^2 = d,$$

ce qui donne 2 possibilités

$$d = 0, \quad \text{et} \quad \mathcal{C} \text{ est la réunion de deux droites} \quad (8.21)$$

$$d \neq 0, \quad \text{et} \quad \mathcal{C} \text{ est une hyperbole.} \quad (8.22)$$

(3) Si $\det(A) = \lambda_1\lambda_2 = 0$, alors on peut supposer que $\lambda_1 > 0$ et $\lambda_2 = 0$, et l'équation de \mathcal{C} dans le repère $(0, \vec{u}, \vec{v})$ est de la forme

$$\lambda_1x^2 + mx + ny + d = 0,$$

ce qui donne quatre possibilités

$$n \neq 0, \quad \text{et} \quad \mathcal{C} \text{ est une parabole} \quad (8.23)$$

$$n = 0, \quad m^2 - 4\lambda_1d < 0, \quad \text{et} \quad \mathcal{C} = \emptyset \quad (8.24)$$

$$n = 0, \quad m^2 - 4\lambda_1d > 0, \quad \text{et} \quad \mathcal{C} \text{ est la réunion de deux droites parallèles} \quad (8.25)$$

$$n = 0, \quad m^2 - 4\lambda_1d = 0, \quad \text{et} \quad \mathcal{C} \text{ est une droite} \quad (8.26)$$

En ce qui concerne les quadriques, la situation est un peu plus compliquée, mais on obtient aussi une classification complète.

Théorème 8.5.5 *Dans l'espace affine euclidien rapporté à un repère ortho-normé $(0, \vec{i}, \vec{j}, \vec{k})$, soit \mathcal{Q} une quadrique d'équation*

$$f(X) := XA^tX + X^tB + c = 0,$$

avec $A \in \mathcal{M}_3(\mathbf{R}) \neq 0$, symétrique, $B \in \mathbf{R}^3$, $c \in \mathbf{R}$, soient λ_1, λ_2 et λ_3 les trois valeurs propres de A , répétées si nécessaire selon leurs ordres de multipli-

cité, soit $P = \begin{bmatrix} p_{1,1} & p_{1,2} & p_{1,3} \\ p_{2,1} & p_{2,2} & p_{2,3} \\ p_{3,1} & p_{3,2} & p_{3,3} \end{bmatrix}$ une matrice orthogonale telle que $P^{-1}AP =$

$\begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix}$, et soit $\vec{u} = p_{1,1}\vec{i} + p_{2,1}\vec{j} + p_{3,1}\vec{k}$, $\vec{v} = p_{1,2}\vec{i} + p_{2,2}\vec{j} + p_{3,2}\vec{k}$,

$\vec{w} = p_{1,3}\vec{i} + p_{2,3}\vec{j} + p_{3,3}\vec{k}$.

(1) Si $\det(A) \neq 0$, et si les trois valeurs propres de A sont de même signe, alors f possède un unique point critique U , et l'équation de \mathcal{Q} dans le repère $(U, \vec{u}, \vec{v}, \vec{w})$ est de la forme

$$\lambda_1 x^2 + \lambda_2 y^2 + \lambda_3 z^2 + d = 0.$$

Dans ce cas on a 3 possibilités

$$d = 0, \text{ et } \mathcal{Q} = \{U\} \quad (8.27)$$

$d \neq 0$, et d et les valeurs propres de A sont de même signe, et dans ce cas $\mathcal{Q} = \emptyset$ (8.28)

$d \neq 0$, et d et les valeurs propres de A sont de même signe, et dans ce cas \mathcal{Q} est un ellipsoïde (8.29)

(2) Si $\det(A) \neq 0$, et si les valeurs propres de A ne sont pas de même signe, alors f possède un unique point critique U , et on peut supposer que $\lambda_1 > 0$, $\lambda_2 > 0$, $\lambda_3 < 0$, et l'équation de \mathcal{Q} dans le repère $(U, \vec{u}, \vec{v}, \vec{w})$ est de la forme

$$\lambda_1 x^2 + \lambda_2 y^2 + \lambda_3 z^2 + d = 0.$$

Dans ce cas on a 3 possibilités

$$d = 0, \text{ et } \mathcal{Q} \text{ est un cône de sommet } U \quad (8.30)$$

$d > 0$ et dans ce cas \mathcal{Q} est un hyperboloïde elliptique à deux nappes (8.31)

$d < 0$ et dans ce cas \mathcal{Q} est un hyperboloïde elliptique à une nappe (8.32)

(3) Si $\det(A) = 0$, A est de rang 2, et dans ce cas on peut supposer que $\lambda_3 = 0$. Dans ce cas l'équation de \mathcal{Q} dans le repère $(0, \vec{u}, \vec{v}, \vec{w})$ est de la forme

$$\lambda_1 x^2 + \lambda_2 y^2 + cx + dy + ez + f = 0,$$

Dans ce cas si U est le point de coordonnées $(-\frac{c}{2\lambda_1}, -\frac{d}{2\lambda_2}, 0)$ dans le repère $(0, \vec{u}, \vec{v}, \vec{k})$, l'équation de \mathcal{Q} dans $(U, \vec{u}, \vec{v}, \vec{k})$ est de la forme

$$\lambda_1 x^2 + \lambda_2 y^2 + ez + g = 0,$$

avec $g = f - \frac{c^2}{4\lambda_1} - \frac{d^2}{4\lambda_2}$, et dans ce cas on a les possibilités suivantes

$$\lambda_1 \lambda_2 > 0, e = g = 0, \text{ et } \mathcal{Q} = \{U\}. \quad (8.33)$$

$$\lambda_1 \lambda_2 > 0, e = 0, g \lambda_1 > 0, \text{ et } \mathcal{Q} = \emptyset. \quad (8.34)$$

$$\lambda_1 \lambda_2 > 0, e = 0, g \lambda_1 < 0, \text{ et } \mathcal{Q} \text{ est un cylindre à base elliptique.} \quad (8.35)$$

$$\lambda_1 \lambda_2 < 0, e = g = 0, \text{ et } \mathcal{Q} \text{ est la réunion de deux plans non parallèles.} \quad (8.36)$$

$$\lambda_1 \lambda_2 > 0, e \neq 0, \text{ et } \mathcal{Q} \text{ est un paraboloïde elliptique} \quad (8.37)$$

$$\lambda_1 \lambda_2 < 0, e \neq 0, \text{ et } \mathcal{Q} \text{ est un paraboloïde hyperbolique} \quad (8.38)$$

Si la matrice A est de rang 1, on peut supposer que $\lambda_1 > 0$, $\lambda_2 = \lambda_3 = 0$, et dans ce cas l'équation de \mathcal{Q} dans le repère $(0, \vec{u}, \vec{v}, \vec{w})$ est de la forme

$$\lambda_1 x^2 + bx + cy + dz + e = 0.$$

En notant U le point de coordonnées $(-\frac{b}{2\lambda_1}, 0, 0)$ dans le repère $(0, \vec{u}, \vec{v}, \vec{w})$, l'équation de \mathcal{Q} dans le repère $(U, \vec{u}, \vec{v}, \vec{w})$ devient

$$\lambda_1 x^2 + cy + dz + f = 0, (*)$$

avec $f = e - \frac{b^2}{4\lambda_1}$.

On a alors les possibilités suivantes

$$c = d = 0, f > 0, \text{ et } \mathcal{Q} = \emptyset. \quad (8.39)$$

$$c = d = 0, f = 0, \text{ et } \mathcal{Q} \text{ est un plan.} \quad (8.40)$$

$c = d = 0, f < 0$, et \mathcal{Q} est la réunion de deux plans parallèles. (8.41)

Reste la situation où $(c, d) \neq (0, 0)$. Dans ce cas si on pose $r := \sqrt{c^2 + d^2}$, $\vec{v}_1 = \frac{c}{r}\vec{v} + \frac{d}{r}\vec{w}$, $\vec{w}_1 = -\frac{d}{r}\vec{v} + \frac{c}{r}\vec{w}$, l'équation de \mathcal{Q} dans le repère orthonormé $(U, \vec{u}, \vec{v}_1, \vec{w}_1)$ devient

$$\lambda_1 x^2 + ry + f = 0,$$

et si on note V le point de coordonnées $(0, -\frac{f}{r}, 0)$ dans le repère orthonormé $(V, \vec{u}, \vec{v}_1, \vec{w}_1)$ devient

$$\lambda_1 x^2 + ry = 0.$$

Donc on a la dernière possibilité suivante

$\lambda_1 > 0, \lambda_2 = \lambda_3 = 0, (c, d) \neq (0, 0)$ dans l'équation (*), et dans ce cas \mathcal{Q} est un cylindre à base parabolique (8.42)

8.6 Chapitre 8 sous Mupad

La vérification du fait qu'une matrice symétrique est ou non définie positive, et le calcul de la décomposition de Cholesky d'une matrice définie positive sont accessibles au calcul formel. Nous prenons pour exemple la matrice

$$A = \begin{bmatrix} 8 & 7 & 6 & 5 & 4 & 3 \\ 7 & 7 & 6 & 5 & 4 & 3 \\ 6 & 6 & 6 & 5 & 4 & 3 \\ 5 & 5 & 5 & 5 & 4 & 3 \\ 4 & 4 & 4 & 4 & 4 & 3 \\ 3 & 3 & 3 & 3 & 3 & 3 \end{bmatrix}.$$

Après avoir vérifié au départ que A était bien définie positive avec la commande `linalg : :isPosDef(A)`, Mupad peut calculer exactement la décomposition de Cholesky de A , avec la commande `linalg : :factorCholeski(A)`. Par contre la tentative de calculer les valeurs propres de A (`linalg : :eigenvalues(A)`) ne peut aboutir car Mupad ne peut évidemment pas calculer les racines d'un polynôme de degré 6 (la réponse renvoie simplement aux racines du polynôme caractéristique).

```
M:=Dom::Matrix();
```

```
Dom::Matrix()
```

```
A:=M([[8,7,6,5,4,3],[7,7,6,5,4,3],[6,6,6,5,4,3],
[5,5,5,5,4,3],[4,4,4,4,4,3],[3,3,3,3,3,3]]);
linalg::isPosDef(A);
linalg::factorCholesky(A);
```

$$\begin{pmatrix} 8 & 7 & 6 & 5 & 4 & 3 \\ 7 & 7 & 6 & 5 & 4 & 3 \\ 6 & 6 & 6 & 5 & 4 & 3 \\ 5 & 5 & 5 & 5 & 4 & 3 \\ 4 & 4 & 4 & 4 & 4 & 3 \\ 3 & 3 & 3 & 3 & 3 & 3 \end{pmatrix}$$

```
true
```

$$\begin{pmatrix} 2 \cdot \sqrt{2} & 0 & 0 & 0 & 0 & 0 \\ \frac{7 \cdot \sqrt{2}}{4} & \frac{\sqrt{2} \cdot \sqrt{7}}{4} & 0 & 0 & 0 & 0 \\ \frac{3 \cdot \sqrt{2}}{2} & \frac{3 \cdot \sqrt{2} \cdot \sqrt{7}}{14} & \frac{\sqrt{6} \cdot \sqrt{7}}{7} & 0 & 0 & 0 \\ \frac{5 \cdot \sqrt{2}}{4} & \frac{5 \cdot \sqrt{2} \cdot \sqrt{7}}{28} & \frac{5 \cdot \sqrt{6} \cdot \sqrt{7}}{42} & \frac{\sqrt{5} \cdot \sqrt{6}}{6} & 0 & 0 \\ \sqrt{2} & \frac{\sqrt{2} \cdot \sqrt{7}}{7} & \frac{2 \cdot \sqrt{6} \cdot \sqrt{7}}{21} & \frac{2 \cdot \sqrt{5} \cdot \sqrt{6}}{15} & \frac{2 \cdot \sqrt{5}}{5} & 0 \\ \frac{3 \cdot \sqrt{2}}{4} & \frac{3 \cdot \sqrt{2} \cdot \sqrt{7}}{28} & \frac{\sqrt{6} \cdot \sqrt{7}}{14} & \frac{\sqrt{5} \cdot \sqrt{6}}{10} & \frac{3 \cdot \sqrt{5}}{10} & \frac{\sqrt{3}}{2} \end{pmatrix}$$

```
linalg::eigenvalues(A);
```

$$\text{RootOf}\left(117 \cdot X1^2 - 31 \cdot X1 - 196 \cdot X1^3 + 140 \cdot X1^4 - 33 \cdot X1^5 + X1^6 + 3, X1\right)$$

8.7 Tracé de coniques sous Matlab

On commence par s'intéresser à une courbe du type ellipse. On commence par la courbe d'équation $x^2 + xy + y^2 + 2x + y + 5 = 0$. Si on pose $f(x, y) = x^2 + xy + y^2 + 2x + y + 5$, les points critiques de f sont donnés par le système

$$\begin{cases} 2x + y + 2 = 0 \\ x + 2y + 1 = 0 \end{cases}$$

On résout ce système sous Matlab, et on voit que le centre de symétrie est le point de coordonnées $(0, 1)$. On reformule cette équation sous forme d'une équation du second degré en y , à savoir

$$y^2 + (x + 1)y + x^2 + 2x + 5 = 0.$$

On calcule les racines de $\Delta = (x + 1)^2 - 4(x^2 + 2x + 5)$. Elles ne sont pas réelles, donc Δ est toujours négatif et la courbe est vide.

```
>> A=[2 1];[1 2]]
```

```
A =
```

```
    2    1
    1    2
```

```
>> B=[-2;-1]
```

```
B =
```

```
   -2
   -1
```

```
>> X=A\B
```

```
X =
```

```
  -1.0000
   0.0000
```

```
>> p=[1 1]
```

```
p =
```

```

      1      1

>> q=conv(p,p)

q =

      1      2      1

>> r=[4,8,20]

r =

      4      8      20

>> roots(q-r)

ans =

-1.0000 + 2.3094i
-1.0000 - 2.3094i

```

Afin d'éviter l'ensemble vide, on considère la courbe d'équation $x^2 + xy + y^2 + 2x + y = 0$, qui contient le point de coordonnées $(0, 0)$. Il suffit de reprendre les mêmes polynômes p et q et de changer r en $r1 = 4x^2 + 8x$.

```

>> p = [1 1];
q=conv(p,p);
r1=[4 8 0];
roots(q-r1)
ans =

-2.1547
 0.1547

>>x=[-2.1547:0.01:0.1547];
y=ones(size(x));
z1=-(x+y)/2 +(sqrt((x+y).^2 -4*x.^2 -8*x))/2;
z2=-(x+y)/2 -(sqrt((x+y).^2 -4*x.^2 -8*x))/2;
plot(x,z1,'g');
hold on
plot(x,z2,'g')

```

```

hold on
axis equal
hold on
x0=ones(801);
x1=0*x0;
y1=[-4:0.01:4];
plot(x1,y1)
hold on
x2=[-5.05:0.01:5.05];
y2=ones(size(x2));
y3=0*y2;
plot(x2,y3);
title('ellipse d equation x^2 +xy +y^2 +2x+y=0');
print -depsc ellipse

```

On continue en traçant des courbes d'équation $2x^2 + xy - 0.129y^2 + x + y - a = 0$. Dans tous les cas si on pose $f(x, y) = 2x^2 + xy - 0.129y^2 + x + y - a$, les points critiques de f sont donnés par les solutions du système

$$\begin{cases} 4x + y + 1 = 0 \\ x - 0.258y + 1 = 0 \end{cases}$$

On résout ce système sous Matlab.

```

>> A=[[4 1];[1 -0.129]];
>> B=[-1;-1];
>> X=A\B

```

X =

```

-0.7447
 1.9789

```

On voit donc que le centre de symétrie de ces courbes est le point de coordonnées $(-0.7447, 1.9789)$. On a sous Matlab

```

>> 2*(-0.7447)^2 + (-0.7447)* 1.9789 -0.129*(1.9789)^2 -0.7447+ 1.9789

```

ans =

```

0.3645

```

On voit donc que ces hyperboles dégèrent en la réunion de deux droites si a est (à peu près) égal à 0.3645, et dans ce cas les droites obtenues sont les asymptotes des autres hyperboles. On va donc procéder au tracé de la courbe dans les cas $a = -1$, $a = 0.3645$, $a = 1$.

On aboutit à une équation du second degré permettant d'obtenir y en fonction de x .

$$-0.129y^2 + (x+1)y + 2x^2 + x - a = 0.$$

On obtient $\Delta = (x+1)^2 + 4(0.129)(2x^2 + x - a)$, qui correspond aux polynômes r_1, r_2, r_3 introduits ci-dessous. Dans les cas

```
>> p=[1 1];
>> q1=[2 1 1];
>> r1=conv(p,p)+4*0.129*q1
>> roots(r1)

ans =

    -0.6191 + 0.6023i
    -0.6191 - 0.6023i

>> q2=[2 1 -0.3645]
>> r2=conv(p,p)+4*0.129*q2;
>> roots(r2)

ans =

    -0.6191 + 0.1276i
    -0.6191 - 0.1276i

>> q3=[2 1 -1];
>> r3=conv(p,p)+4*0.129*q3;
>> roots(r3)

ans =

    -1.0000
    -0.2382
```

On voit donc que l'on peut faire varier x sur \mathbf{R} si $a = -1$ ou $a = 0.3645$, mais qu'il faut se limiter à $x < -1$ ou $x > -0.2382$ si $a = 1$. On procède maintenant au tracé des courbes.

```

>> x1=[-5:0.01:5];
y1=ones(size(x1));
z1=-(x1+y1)/2 +(1/2)*sqrt((x1+y1).^2 +4*(0.129)*(2*x1.^2 +x1+y1));
u1=-z1/(0.129);
plot(x1,u1,'r');
hold on
z2=-(x1+y1)/2 -(1/2)*sqrt((x1+y1).^2 +4*(0.129)*(2*x1.^2 +x1+y1));
u2=-z2/(0.129);
plot(x1,u2,'r');
hold on
z2=-(x1+y1)/2 +(1/2)*sqrt((x1+y1).^2 +4*(0.129)*(2*x1.^2 +x1-(0.3645)*y1));
u2=-z2/(0.129);
plot(x1,u2);
hold on
z3=-(x1+y1)/2 -(1/2)*sqrt((x1+y1).^2 +4*(0.129)*(2*x1.^2 +x1-(0.3645)*y1));
u3=-z3/(0.129);
plot(x1,u3);
hold on
x2=[-5:0.01:-1];
y2=ones(size(x2));
z2=-(x2+y2)/2 +(1/2)*sqrt((x2+y2).^2 +4*(0.129)*(2*x2.^2 +x2-y2));
u2=-z2/(0.129);
plot(x2,u2,'g');
hold on
z3=-(x2+y2)/2 -(1/2)*sqrt((x2+y2).^2 +4*(0.129)*(2*x2.^2 +x2-y2));
u3=-z3/(0.129);
plot(x2,u3,'g');
hold on
x3=[-0.2382:0.01:5];
y3=ones(size(x3));
z4=-(x3+y3)/2 +(1/2)*sqrt((x3+y3).^2 +4*(0.129)*(2*x3.^2 +x3-y3));
u4=-z4/(0.129);
plot(x3,u4,'g');
hold on
z5=-(x3+y3)/2 -(1/2)*sqrt((x3+y3).^2 +4*(0.129)*(2*x3.^2 +x3-y3));
u5=-z5/(0.129);
plot(x3,u5,'g');
title('Les courbes d equation 2x^2+xy-0.129y+x+y=a,a=-1,a=0.3645,a=1');
print -depsc hyperboles

```

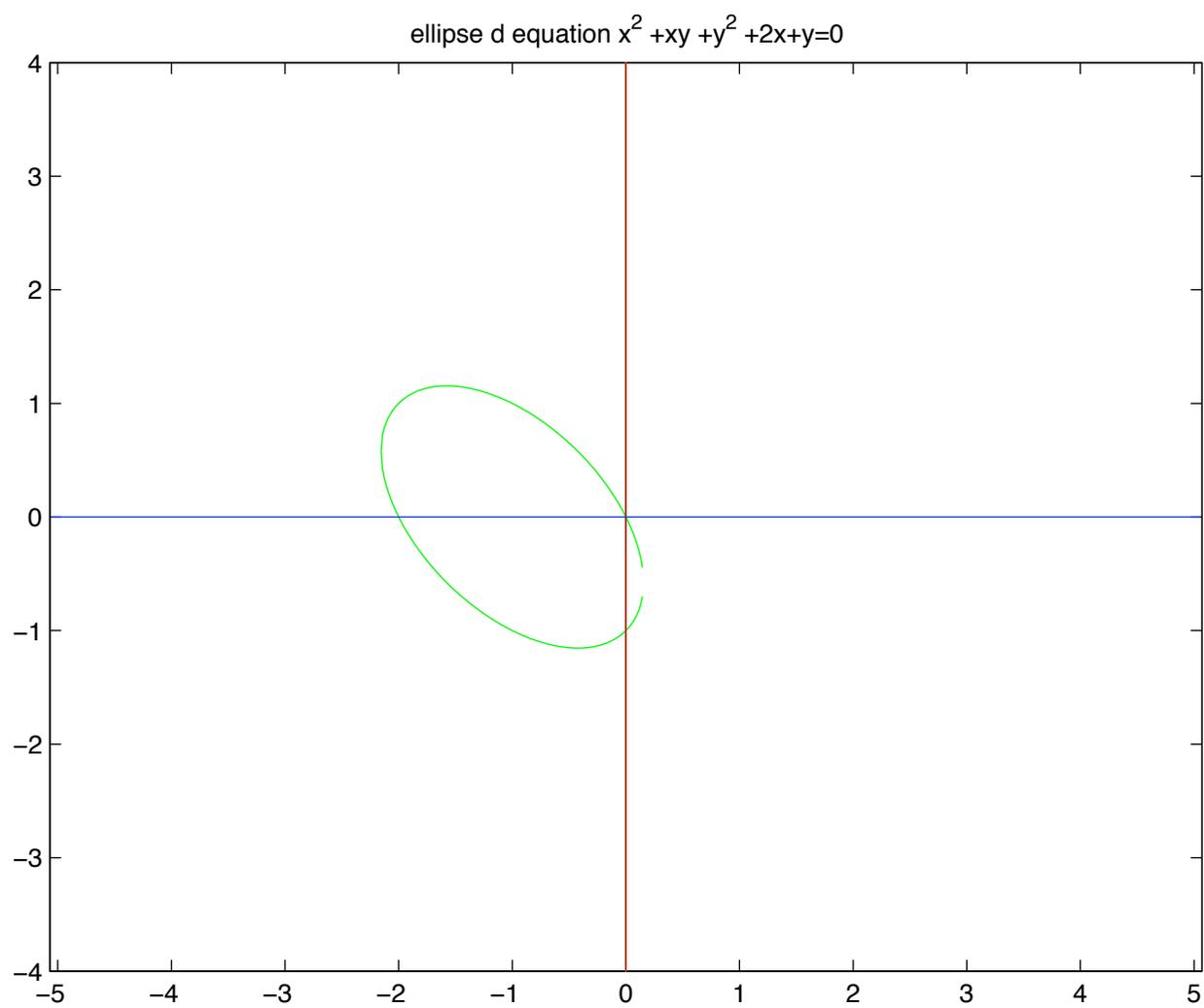
On va maintenant illustrer le cas des courbes du type parabole. On considère la famille de coniques d'équation $x^2 + 2xy + y^2 + ay + b = 0$. On va traiter les cas $a = 0, b = 0, a = 0, b = -1, a = 1, b = 0$. Si on veut exprimer x en fonction

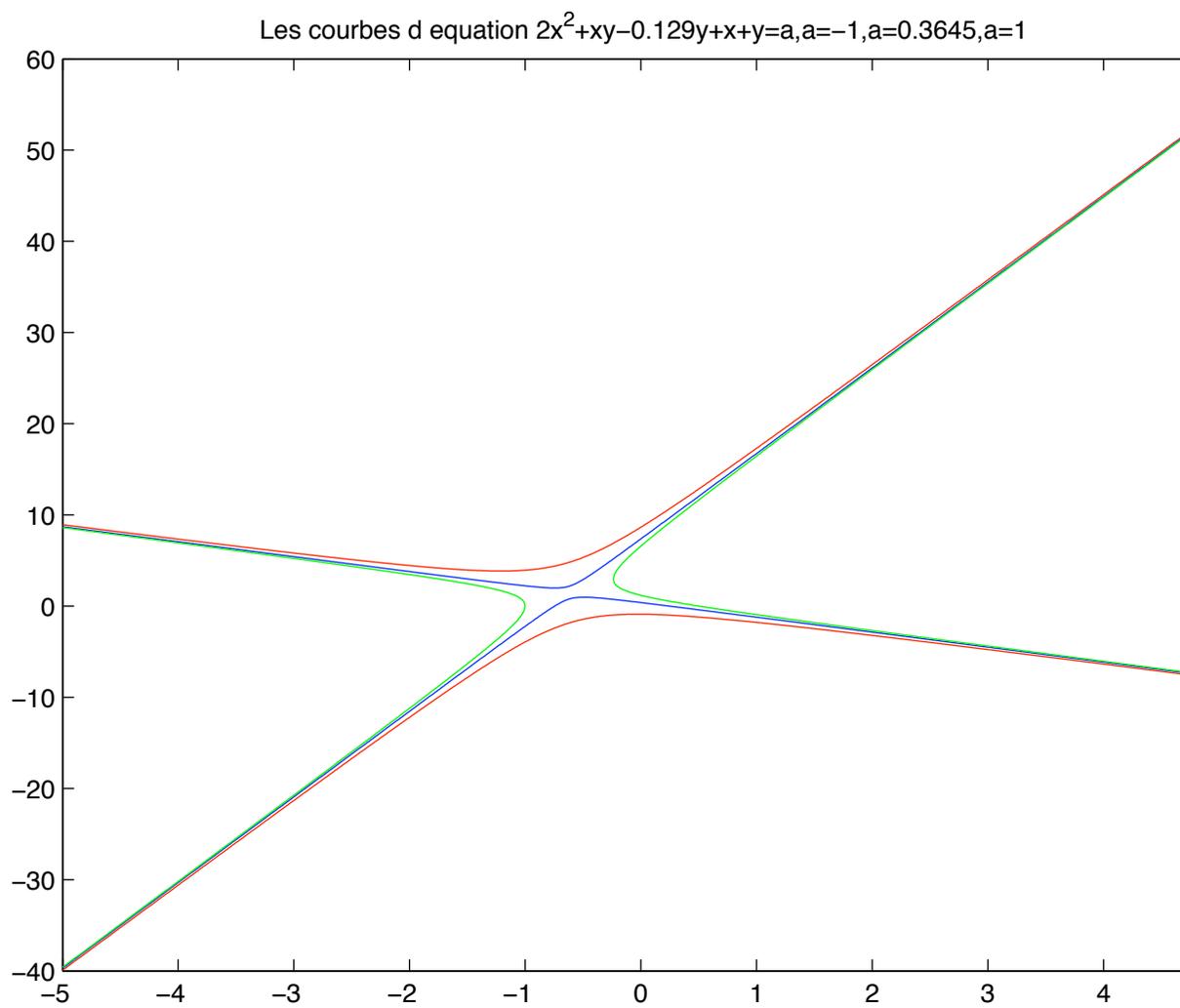
de y , on obtient une équation du second degré, dont le discriminant vaut

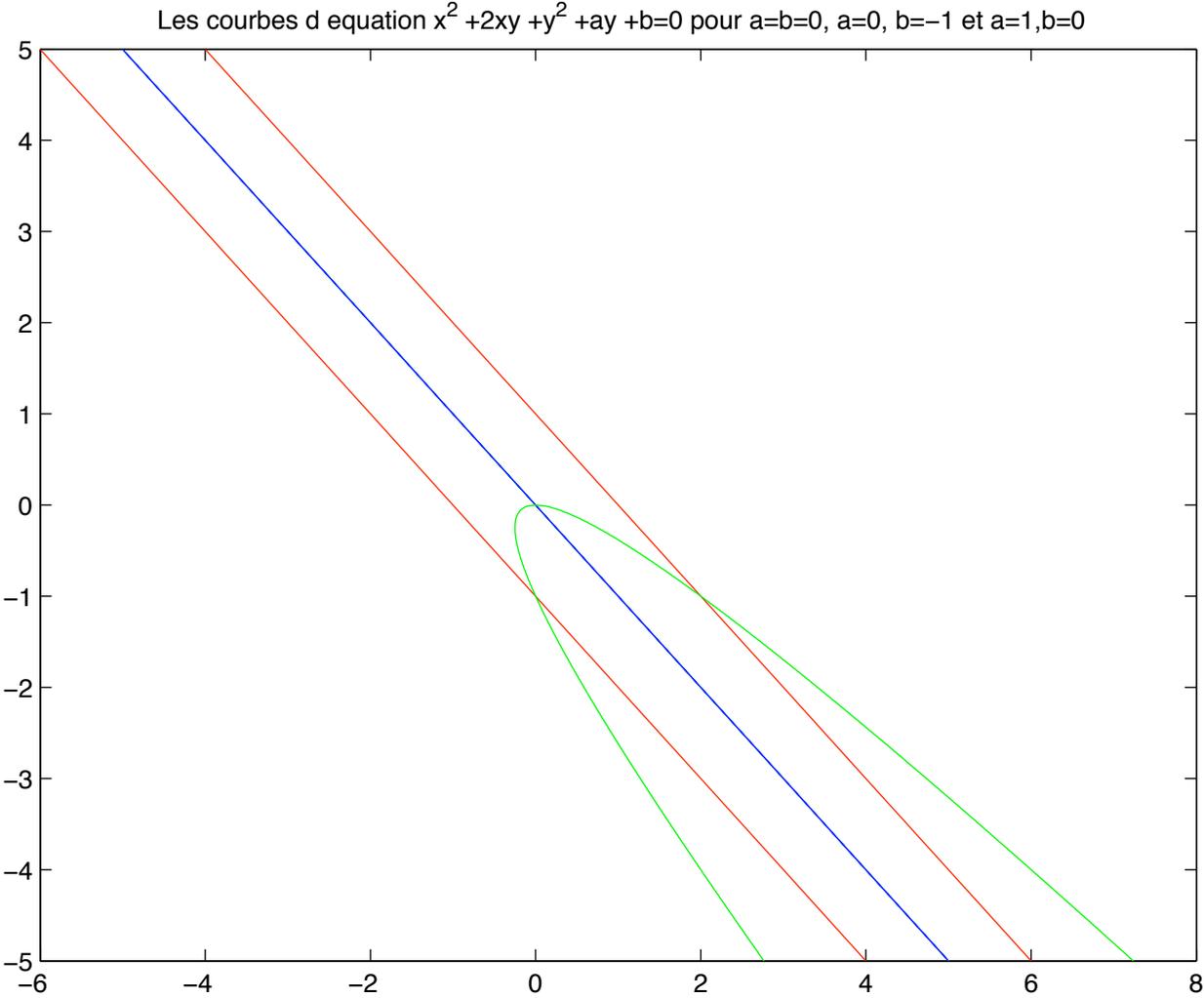
$$\Delta = 4y^2 - 4y^2 - 4ay - 4b = -4(ay + b).$$

Le cas $a = 0$ donne l'ensemble vide si $b > 0$. On va vérifier numériquement qu'on obtient une droite si $b = 0$, et la réunion de deux droites parallèles si $b < 0$ (pour nous dans le cas où $b = -1$). Le cas $a = 1, b = 0$ donne $\Delta \geq 0$ pour $y < 0$, et on va faire apparaître une parabole.

```
>> y=[-5:0,01:5];
x1=-y +sqrt(y.^2-y.^2);
x2=-y -sqrt(y.^2-y.^2);
plot(x1,y);
hold on
plot(x2,y);
u=ones(size(y));
x3=-y+sqrt(y.^2-(y.^2-u));
plot(x3,y,'r');
hold on
x4=-y-sqrt(y.^2-(y.^2-u));
plot(x4,y,'r');
hold on
y1=[-5:0.01:0];
x5=-y1+sqrt(y1.^2 -(y1.^2+y1));
x6=-y1-sqrt(y1.^2 -(y1.^2+y1));
plot(x5,y1,'g');
hold on
plot(x6,y1,'g');title('Les courbes d equation x^2 +2xy +y^2 +ay +b=0 pour a=b=0, a=0, 1');
print -depsc parabole
>> axis equal
```







8.8 Tracé de quadriques sous Matlab

On s'intéresse aux quadriques dont l'équation dans l'espace euclidien usuel rapporté à un repère orthonormé $(0, \vec{i}, \vec{j}, \vec{k})$ est de la forme $x^2 + y^2 + z^2 + xy + yz + zx + d = 0$.

On rentre sous Matlab la matrice A associée à la forme quadratique $q(x, y, z) = x^2 + y^2 + z^2 + xy + yz + zx$.

```
>> A=[1 0.5 0.5;0.5 1 0.5 ; 0.5 0.5 1]
```

A =

```

1.0000    0.5000    0.5000
0.5000    1.0000    0.5000
0.5000    0.5000    1.0000
```

On cherche maintenant sous Matlab une matrice orthogonale Q telle que $D = QAQ^{-1}$ soit diagonale. On commence par diagonaliser A .

```
>> [P,D]=eig(A)
```

P =

```

0.7309    0.3639    0.5774
-0.0503   -0.8149    0.5774
-0.6806    0.4511    0.5774
```

D =

```

0.5000         0         0
         0    0.5000         0
         0         0    2.0000
```

On calcule $P *^t P$

```
>> P*P'
```

ans =

```

1.0000    0.0000    0.0000
0.0000    1.0000   -0.0000
0.0000   -0.0000    1.0000
```

et on constate avec joie que Matlab a donné directement une matrice orthogonale P . On vérifie que $P * D * P^{-1}$ est bien égale à A .

```
>> P*D*P^(-1)
```

```
ans =
```

```
1.0000    0.5000    0.5000
0.5000    1.0000    0.5000
0.5000    0.5000    1.0000
```

Posons $\vec{u} = 0.7309\vec{i} - 0.0503\vec{j} - 0.6806\vec{k}$, $\vec{v} = 0.3639\vec{i} - 0.8149\vec{j} + 0.4511\vec{k}$, $\vec{w} = 0.5774\vec{i} + 0.5774\vec{j} + 0.5774\vec{k}$.

La matrice représentant la forme q dans la base orthonormale $\mathcal{B} = \{\vec{i}, \vec{j}, \vec{k}\}$ est égale à ${}^tPAP = P^{-1}AP = D = \begin{bmatrix} 0.5 & 0 & 0 \\ 0 & 0.5 & 0 \\ 0 & 0 & 2 \end{bmatrix}$. L'équation de la quadrique dans le repère $(O, \vec{i}, \vec{j}, \vec{k})$ est donc

$$\frac{x'^2}{2} + \frac{y'^2}{2} - 2z'^2 + d = 0,$$

que l'on réécrit sous la forme

$$x'^2 + y'^2 - 4z'^2 = -2d.$$

Il est clair que l'on obtient l'ensemble vide pour $d > 0$, le point O pour $d = 0$, et un ellipsoïde de révolution d'axe Oz' pour $d < 0$. On va représenter graphiquement la surface obtenue dans le repère initial $(0, \vec{i}, \vec{j}, \vec{k})$ en utilisant une paramétrisation par coordonnées sphériques, dans le cas où $d = -2.163$. On obtient

$x' = \sqrt{4.326}\cos\theta\cos\phi$, $y' = \sqrt{4.326}\sin\theta\cos\phi$, $z' = \frac{\sqrt{4.326}}{2}\sin\phi$, avec $0 \leq \theta \leq 2\pi$, $-\frac{\pi}{2} \leq \phi \leq \frac{\pi}{2}$. Ceci donne, compte tenu du fait que l'on a

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = P \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix},$$

$$\begin{cases} x = 0.7309x' + 0.3639y' + 0.5774z' \\ y = -0.0503x' - 0.8149y' + 0.5774z' \\ z = -0.6806x' + 0.4511y' + 0.5774z' \end{cases}$$

On obtient

$$\begin{aligned} x &= 0.7309\sqrt{4.326}\cos\theta\cos\phi + 0.3639\frac{\sqrt{4.326}}{2}\sin\theta\cos\phi + 0.5774\frac{\sqrt{4.326}}{2}\sin\phi \\ y &= -0.0503\sqrt{4.326}\cos\theta\cos\phi - 0.8149\sqrt{4.326}\sin\theta\cos\phi + 0.5774\frac{\sqrt{4.326}}{2}\sin\phi \\ z &= -0.6806\sqrt{4.326}\cos\theta\cos\phi + 0.4511\sqrt{4.326}\sin\theta\cos\phi + 0.5774\frac{\sqrt{4.326}}{2}\sin\phi \end{aligned}$$

On procède maintenant à la visualisation de l'ellipsoïde dans le cas où $d = -2.163$.

```
>> [theta, phi]= meshgrid (-pi:pi/100:pi);
    x= 0.7309*sqrt(4.326)*cos(theta).* cos(phi/2)    +0.3639 *sqrt(4.326/ 2)*sin(theta)
    y=-0.0503*sqrt(4.326)*cos(theta).* cos(phi/2)    -0.8149*sqrt(4.326)*sin(theta ).*cos
    z=-0.6806*sqrt(4.326)*cos(theta).* cos(phi/2)    +0.4511*sqrt(4.326)*sin(theta)* cos
plot3(x,y,z)
hold on
title('ellipsoïde d equation x^2+y^2+z^2 +xy +yz +zx -2.163=0');
print -depsc ellipsoïde
```

On va maintenant s'intéresser à des quadriques associées à une matrice symétrique inversible ayant des valeurs propres de signe opposé. On considère la quadrique d'équation

$$z^2 + 2xy + 2yz + 2zx - 2x + a = 0,$$

où $a \in \mathbf{R}$. On pose

$$A = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 1 \end{bmatrix}, b = [-2 \quad 0 \quad 0].$$

Les coordonnées x_0, y_0, z_0 du centre de symétrie S de la quadrique sont données par la formule

$$\begin{bmatrix} x_0 \\ y_0 \\ z_0 \end{bmatrix} = -\frac{Ab}{2}.$$

Un calcul direct sous Mupad donne $x_0 = -1, y_0 = 0, z_0 = 1$. Si x', y', z' désignent les coordonnées dans $(S, \vec{i}, \vec{j}, \vec{k})$ d'un point M ayant pour coordonnées x, y, z dans le repère $(S, \vec{i}, \vec{j}, \vec{k})$, on a $x = x' - 1, y = y', z = z' + 1$. En substituant dans l'équation (on peut s'aider de Mupad), on trouve que l'équation de la quadrique dans le repère $(S, \vec{i}, \vec{j}, \vec{k})$ est de la forme

$$z'^2 + 2x'y' + 2y'z' + 2z'x' + 1 + a = 0.$$

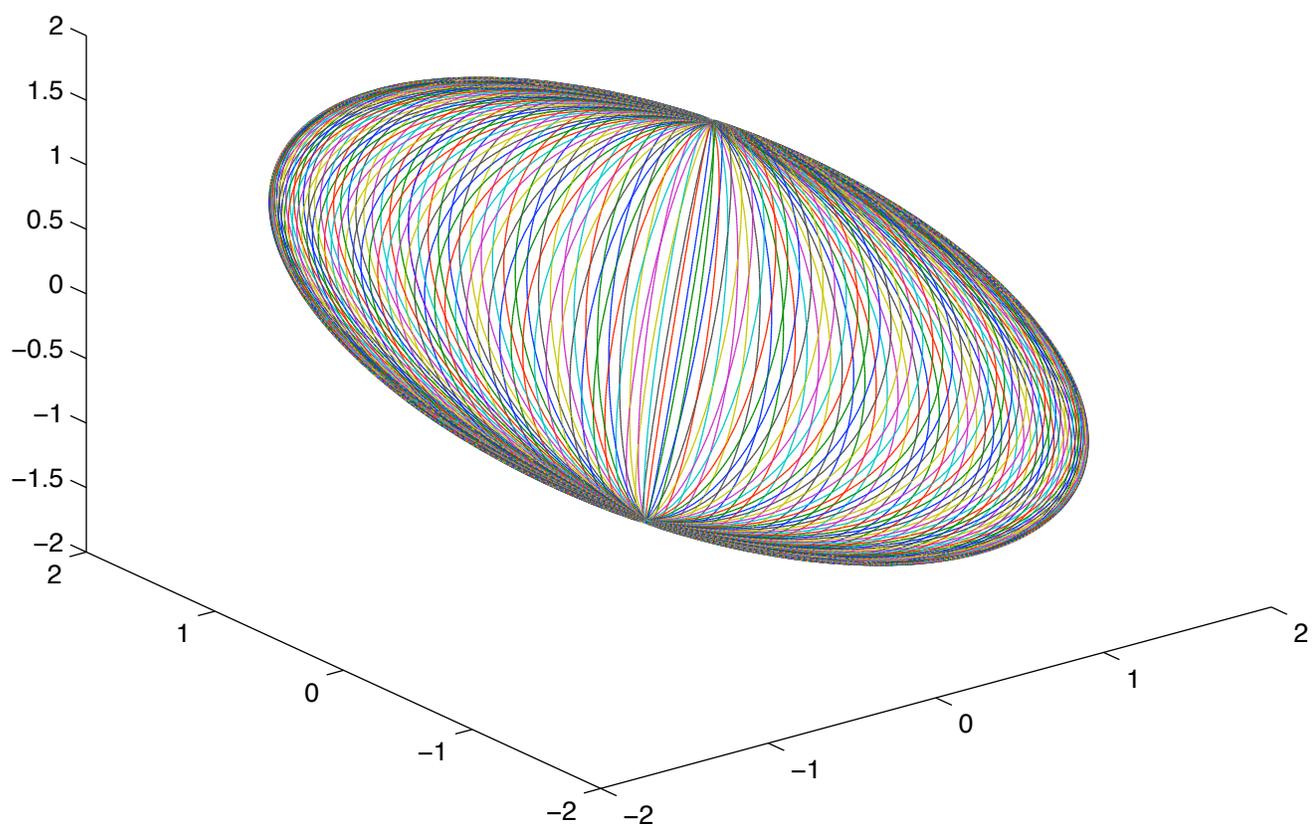
On introduit sous Matlab la matrice A associée à la forme quadratique $q(x', y', z') = z'^2 + 2x'y' + 2y'z' + 2z'x'$, et on la diagonalise.

```
>> A=[ [0 1 1],[1 0 1], [1 1 1]]
```

A =

```
    0    1    1    1    0    1    1    1    1
```

ellipsoïde d'équation $x^2+y^2+z^2+xy+yz+zx-2.163=0$



```
>> A=[0 1 1;1 0 1;1 1 1]
```

```
A =
```

```

0     1     1
1     0     1
1     1     1
```

```
>> [P,D]=eig(A)
```

```
P =
```

```

0.7071    0.5000    0.5000
-0.7071    0.5000    0.5000
-0.0000   -0.7071    0.7071
```

```
D =
```

```

-1.0000         0         0
         0   -0.4142         0
         0         0    2.4142
```

On voit donc que si on pose $\vec{u} = 0.7071\vec{i} - 0.7071\vec{j}$, $\vec{v} = 0.5\vec{i} + 0.5\vec{j} - 0.7071\vec{k}$, $\vec{w} = 0.5\vec{i} + 0.5\vec{j} + 0.7071\vec{k}$, l'équation de la quadrique dans le repère $(S, \vec{u}, \vec{v}, \vec{w})$ est de la forme

$$X^2 + -0.4142Y^2 - 2.4142Z^2 = 1 + a.$$

On obtient un cône à base elliptique d'axe OZ si $a = -1$, un hyperboloïde elliptique à une nappe d'axe OZ si $a > -1$, un hyperboloïde elliptique à deux nappes d'axe OZ si $a < -1$.

On essaie d'obtenir directement une représentation de la quadrique sous Matlab en exprimant y en fonction de x et z dans le repère original, dans les cas $a = -1$ (cône), $a = 0,8$ (hyperboloïde à une nappe), $a = -2,8$ (hyperboloïde à deux nappes).

```

u=[-4:0.1:2];v=[-2:0.1:4];
[x,z]=meshgrid(u,v);
y1=(-ones(size(x)).*ones(size(z))+2*(x.*ones(size(z)))-2*x.*z-z.^2)
./(x.*ones(size(z)) +ones(size(x)).*z);
subplot(221),plot3(x,y1,z),title(' a=-1,cone elliptique ');
hold on
```

```

y2=(0.8*ones(size(x)).*ones(size(z))+2*(x.*ones(size(z)))-2*x.*z-z.^2)
./(x.*ones(size(z)) +ones(size(x)).*z);
subplot(222),plot3(x,y2,z),title('a=0.8,hyperboloide a une nappe ');
hold on
y3=(-2.8*ones(size(x)).*ones(size(z))+2*(x.*ones(size(z)))-2*x.*z-z.^2)
./(x.*ones(size(z)) +ones(size(x)).*z);
subplot(223),plot3(x,y3,z),title('a=-2, hyperboloide a deux nappes'); hold on
print -depsc hyperboloides

```

8.9 Exercices sur le Chapitre 8

exercice 1

On pose $A = \begin{bmatrix} -\frac{1}{3} & -\frac{2}{3} & \frac{2}{3} \\ -\frac{2}{3} & -\frac{1}{3} & -\frac{2}{3} \\ \frac{2}{3} & -\frac{2}{3} & -\frac{1}{3} \end{bmatrix}$,

de sorte que A est à la fois symétrique et orthogonale.

1) Trouver une matrice orthogonale P telle que $P^{-1}AP$ soit diagonale.

2) Interpréter géométriquement les résultats de la question 1 en utilisant l'exemple 8.2.6.

exercice 2

Montrer que les matrices orthogonales de $\mathcal{M}_2(\mathbf{R})$ sont les matrices de la forme $R_\phi = \begin{bmatrix} \cos(\phi) & -\sin(\phi) \\ \sin(\phi) & \cos(\phi) \end{bmatrix}$ ou de la forme $S_\phi = \begin{bmatrix} \cos(\phi) & \sin(\phi) \\ \sin(\phi) & -\cos(\phi) \end{bmatrix}$, avec $\phi \in \mathbf{R}$. Montrer que dans le plan euclidien muni d'une base orthonormée (\vec{i}, \vec{j}) , R_ϕ représente une rotation vectorielle dont on précisera l'angle tandis que S_ϕ représente une symétrie orthogonale vectorielle dont on précisera l'axe.

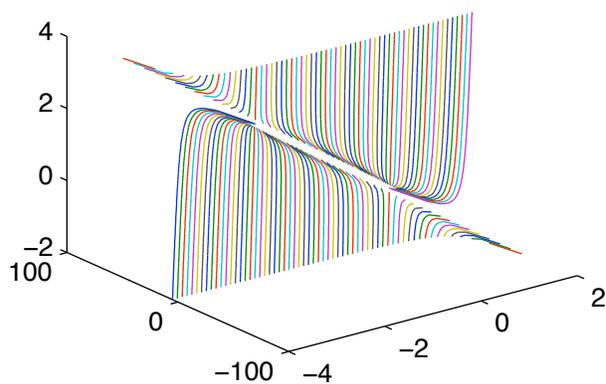
exercice 3

Soient f_1, \dots, f_k et g_1, \dots, g_k des formes linéaires sur un espace vectoriel réel E , et soient $a_1, \dots, a_k \in \mathbf{R}$. On pose, pour $\vec{x} \in E$,

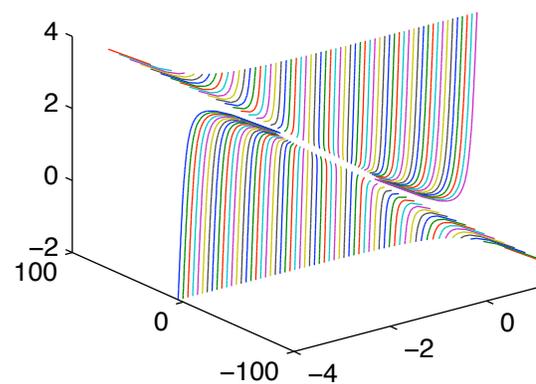
$$q(\vec{x}) = \sum_{j=1}^k a_j f_j(\vec{x}) g_j(\vec{x}).$$

Montrer que q est une forme quadratique sur E .

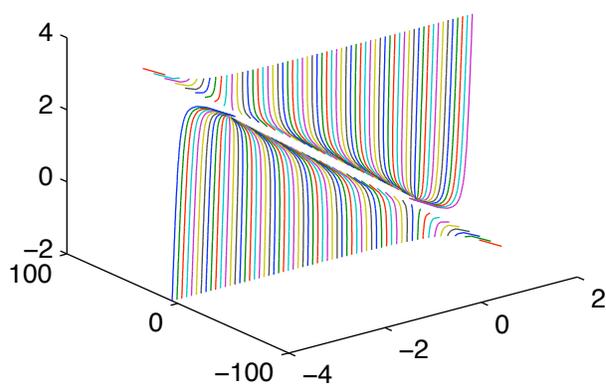
$a=-1$, cône elliptique



$a=0.8$, hyperboloïde à une nappe



$a=-2$, hyperboloïde à deux nappes



Chapitre 9

Les tenseurs en Physique (en préparation)

9.1 Coordonnées covariantes et contravariantes

Dans tout ce chapitre on considère un espace euclidien E de dimension n , muni d'un produit scalaire $(\mathbf{x}, \mathbf{y}) \rightarrow \mathbf{x} \cdot \mathbf{y}$. et rapporté à une base **non nécessairement orthonormale** $\mathcal{B} := \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$.

Définition 9.1.1 Soit $\mathbf{x} \in E$. Les **coordonnées covariantes** x^1, \dots, x^n de \mathbf{x} dans la base \mathcal{B} sont définies par les formules

$$x^j := \mathbf{x} \cdot \mathbf{e}_j \quad j = 1, \dots, n. \quad (9.1)$$

Les **coordonnées usuelles** x_1, \dots, x_n de \mathbf{x} , définies par la formule

$$\mathbf{x} = \sum_{i=1}^n x_i \mathbf{e}_i,$$

sont appelées les **coordonnées contravariantes** de \mathbf{x} dans la base \mathcal{B} .

Notons qu'en notation indicielle (voir le Chapitre 3), les coordonnées contravariantes sont définies par la formule

$$\mathbf{x} = x_i \mathbf{e}_i \quad (9.2)$$

On rappelle que si $\mathcal{B} := \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ et $\mathcal{B}' := \{\mathbf{e}'_1, \dots, \mathbf{e}'_n\}$ sont deux bases de E , la matrice de passage de \mathcal{B} à \mathcal{B}' est la matrice $P_{\mathcal{B}, \mathcal{B}'} = (p_{i,j})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq n}}$ dont la j^e colonne est formée des coordonnées de \mathbf{e}'_j dans la base \mathcal{B} . Autrement dit on a

$$\mathbf{e}'_j = \sum_{i=1}^n p_{i,j} \mathbf{e}_i, \quad 1 \leq j \leq n,$$

ou plus simplement, en notation indicielle,

$$\mathbf{e}'_j = p_{i,j} \mathbf{e}_i. \quad (9.3)$$

Si on note X_{contr} et X'_{contr} les vecteurs colonnes formés des coordonnées contravariantes de \mathbf{x} dans \mathcal{B} et \mathcal{B}' , on a, d'après les formules de changement de base du Chapitre 2

$$X_{contr} = P X'_{contr},$$

soit, en notation indicielle

$$x_i = p_{i,j} x'_j. \quad (9.4)$$

De même notons X_{cov} et X'_{cov} les vecteurs colonnes formés des coordonnées covariantes de \mathbf{x} dans \mathcal{B} et \mathcal{B}' . On a $x'^j = \mathbf{x} \cdot \mathbf{e}'_j = \mathbf{x} \cdot \left(\sum_{i=1}^n p_{i,j} \mathbf{e}_i \right) = \sum_{i=1}^n p_{i,j} x^i$, soit, en notation indicielle

$$x'^j = p_{i,j} x^i, \quad (9.5)$$

ce qui correspond à la formule matricielle

$$X'_{cov} = (P_{\mathcal{B},\mathcal{B}'})^t X_{cov}.$$

Soit $P' = P_{\mathcal{B}',\mathcal{B}} := (p'_{i,j})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq n}}$ la matrice de passage de \mathcal{B}' à \mathcal{B} . On obtient

$$X'_{contr} = P' X_{contr}, \quad X_{cov} = (P')^t X'_{cov},$$

et on retrouve le fait que $P' = P^{-1}$.

En adoptant des notations analogues aux notations ci-dessus pour $\mathbf{y} \in E$, on obtient

$$\mathbf{x} \cdot \mathbf{y} = \sum_{i=1}^n y_i \mathbf{x} \cdot \mathbf{e}_i = \sum_{i=1}^n x^i y_i.$$

Compte tenu du fait que $\mathbf{x} \cdot \mathbf{y} = \mathbf{y} \cdot \mathbf{x}$, on obtient, en notation indicielle

$$\mathbf{x} \cdot \mathbf{y} = x^i y_i = x_i y^i. \quad (9.6)$$

Définition 9.1.2 La matrice associée à une base $\mathcal{B} = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ de E est la matrice $A(\mathcal{B}) := (a_{i,j})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq n}}$, où $a_{i,j} = \mathbf{e}_i \cdot \mathbf{e}_j$ pour $1 \leq i \leq n, 1 \leq j \leq n$.

On a

$$x^i = \mathbf{x} \cdot \mathbf{e}_i = \left(\sum_{j=1}^n x_j \mathbf{e}_j \right) \cdot \mathbf{e}_i = \sum_{j=1}^n x_j \mathbf{e}_j \cdot \mathbf{e}_i = \sum_{j=1}^n a_{j,i} x_j = \sum_{j=1}^n a_{j,i} x_j$$

On obtient, en notation indicielle

$$x^i = a_{i,j}x_j = a_{j,i}x_j, \quad (9.7)$$

ce qui correspond à la relation matricielle

$$X_{cov} = A(\mathcal{B})X_{contr}.$$

Proposition 9.1.3 *La matrice $A(\mathcal{B})$ est définie positive.*

Démonstration : Il est clair que $A(\mathcal{B})$ est symétrique. Soit λ une valeur propre de $A(\mathcal{B})$, et soit \mathbf{x} un vecteur propre de norme 1 associé à la valeur propre λ . On a

$$\begin{aligned} 1 = \|\mathbf{x}\|^2 &= \sum_{i=1}^n x^i x_i = {}^t X_{cov} X_{contr} = {}^t X_{contr} {}^t (A(\mathcal{B})) X_{contr} = {}^t X_{contr} A(\mathcal{B}) X_{contr} \\ &= \lambda {}^t X_{contr} X_{contr} = \lambda \left(\sum_{i=1}^n (x^i)^2 \right). \end{aligned}$$

Comme $\sum_{i=1}^n (x^i)^2 > 0$, on voit que $\lambda > 0$, et $A(\mathcal{B})$ est définie positive. ♣
Dans la suite on adopte la notation

$$A(\mathcal{B})^{-1} = (a^{i,j})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq n}} = (a^{j,i})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq n}}. \quad (9.8)$$

On rappelle que les symboles de Kronecker $\delta_{i,j}$ sont définis par les formules $\delta_{i,j} = 0$ si $i \neq j$ et $\delta_{i,j} = 1$ si $i = j$.

Proposition 9.1.4 *Soit $\mathcal{B} = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ une base de E . Il existe alors une unique base $\mathcal{B}^* = \{\mathbf{e}_1^*, \dots, \mathbf{e}_n^*\}$, appelée **base duale** de la base \mathcal{B} , telle que l'on ait, pour $1 \leq i \leq n, 1 \leq j \leq n$,*

$$\mathbf{e}_i \cdot \mathbf{e}_j^* = \delta_{i,j}.$$

On a $A(\mathcal{B}^*) = A(\mathcal{B})^{-1}$. De plus si X_{cov}^* et X_{contr}^* désignent les matrices colonnes formées des coordonnées covariantes et contravariantes de $\mathbf{x} \in E$ dans la base \mathcal{B}^* , on a

$$X_{cov} = X_{contr}^*, \quad X_{contr} = X_{cov}^*. \quad (9.9)$$

Démonstration : On pourrait déduire l'existence de \mathcal{B}^* de l'existence d'une base duale de \mathcal{B} dans $L(E, \mathbf{R})$ au sens du Chapitre 8 (voir la proposition 8.3.11), mais on va procéder directement.

Posons $\mathbf{e}_i^* = \sum_{j=1}^n a^{j,i} \mathbf{e}_j$, $\mathcal{B}^* = \{\mathbf{e}_1^*, \dots, \mathbf{e}_n^*\}$. La matrice de passage $P_{\mathcal{B}, \mathcal{B}^*}$ est égale à la matrice inversible $A(\mathcal{B})^{-1}$, donc \mathcal{B}^* est une base de E .

On a $\mathbf{e}_i \cdot \mathbf{e}_j^* = \sum_{k=1}^n a^{k,j} \mathbf{e}_i \cdot \mathbf{e}_k = \sum_{k=1}^n a^{k,j} \mathbf{e}_k \cdot \mathbf{e}_i = \sum_{k=1}^n a^{k,j} a_{k,i} = \sum_{k=1}^n a^{j,k} a_{k,i}$. Il résulte alors de la formule 9.8 que l'on a bien $\mathbf{e}_i \cdot \mathbf{e}_j^* = \delta_{i,j}$. L'unicité de \mathcal{B}^*

provient du fait que si $\mathbf{e}_i \cdot \mathbf{x} = \mathbf{e}_i \cdot \mathbf{y}$ pour $1 \leq i \leq n$, alors $x - y \in E^\perp = \{0\}$, donc $x = y$.

Soit maintenant $\mathbf{x} \in E$, et soient x_1, \dots, x_n les coordonnées contravariantes de \mathbf{x} dans la base \mathcal{B} . On a, pour $1 \leq i \leq n$,

$$\mathbf{x} \cdot \mathbf{e}_i^* = \left(\sum_{j=1}^n x_j e_j \right) \cdot \mathbf{e}_i^* = \sum_{j=1}^n x_j \delta_{i,j} = x_i.$$

Donc $X_{contr} = X_{cov}^*$. Comme $(\mathcal{B}^*)^* = \mathcal{B}$, on en déduit que $X_{contr}^* = (X^*)_{cov}^* = X_{cov}$. ♣

9.2 A suivre...

9.3 Exercices sur le Chapitre 9

exercice 1

On munit \mathbf{R}^4 du produit scalaire usuel.

a) Vérifier que $\mathcal{B} = \{(-1, 1, 0, 0), (1, 1, 0, 0), (1, 1, 1, 0), (1, 1, 1, 1)\}$ est une base de \mathbf{R}^4 .

b) Calculer la matrice de passage de \mathcal{B}_0 à \mathcal{B} et la matrice de passage de \mathcal{B} à \mathcal{B}_0 , \mathcal{B}_0 désignant la base canonique de \mathbf{R}^4 .

c) Déterminer les coordonnées covariantes et contravariantes des éléments de \mathcal{B}_0 dans la base \mathcal{B} .

d) Même question avec le vecteur $(1, 2, 3, 4)$.

e) Déterminer la matrice $A(\mathcal{B})$ associée à la base \mathcal{B} .

exercice 2

Soit E un espace euclidien de dimension $n \geq 2$, et soit $A \in \mathcal{M}_n(\mathbf{R})$ une matrice symétrique définie positive. Existe-t'il une base \mathcal{B} de E pour laquelle la matrice associée $A(\mathcal{B})$ est égale à A ?

Si oui, cette base est-elle unique?

exercice 3

Soit E un espace euclidien de dimension $n \geq 2$. Pour $\mathbf{x} \in E$, on définit $\phi : E \rightarrow \mathbf{R}$ par la formule

$$\phi(\mathbf{x})(y) = \mathbf{x} \cdot \mathbf{y} \quad \forall y \in E.$$

a) Vérifier que $\phi(\mathbf{x}) \in L(E, \mathbf{R})$ et que $\phi : E \rightarrow L(E, \mathbf{R})$ est une application linéaire bijective.

b) Voyez-vous un lien entre les bases duales du Chapitre 8 et celles du Chapitre 9?

Bibliographie

- [1] A. Boyer et J.Risler, Mathématiques pour la licence : Groupes, anneaux, corps, Dunod, 2006.
- [2] D.Couty, *Formes normales d'automorphismes de \mathbf{C}^n à variété linéaire fixe et répulsive*, Séminaire ...
bibitemdez X.Dussau, J.Esterle, F.Zarouf et R.Zarouf, Algèbre, Cours ESTIA, Edition 2007.
- [3] X.Dussau, J.Esterle, O.Rejasse et F.Zarouf, Analyse élémentaire, Cours ESTIA, Edition 2005.
- [4] J.Esterle, Transformées, Cours ESTIA, Edition 2007.
- [5] X.Fischr, Analyse numérique, Cours ESTIA, 1999.