

Université Pierre et Marie Curie  
Licence de Sciences et Technologies  
Mention Mathématiques  
Année 2012–2013

**Espaces vectoriels euclidiens et hermitiens,  
isométries affines (2M271)**

par Laurent KOELBLEN et Patrick POLO  
(repris et revu par Marco MACULAN)

Je tiens à remercier Laurent Koelblen et Patrick Polo pour m'avoir généreusement fourni les fichiers du désormais ancien LM270, le présent texte est à 99 % le leur.

Marco Maculan  
Université Pierre et Marie Curie (Paris 6)  
Institut de Mathématiques de Jussieu (UMR 7586 du CNRS)  
<http://webusers.imj-prg.fr/~marco.maculan/>  
[marco.maculan@imj-prg.fr](mailto:marco.maculan@imj-prg.fr)

## TABLE DES MATIÈRES

<b>1. Formes bilinéaires symétriques et formes quadratiques</b> .....	1
1.0. Introduction.....	1
1.1. Formes bilinéaires.....	1
1.2. Formes quadratiques.....	7
1.3. Réduction d'une forme quadratique en « somme de carrés ».....	10
1.4. Appendice (†) : formes bilinéaires alternées.....	13
<b>2. Espaces euclidiens et groupes orthogonaux <math>O(n)</math></b> .....	15
2.1. Espaces euclidiens. Inégalité de Cauchy-Schwarz. Isométries.....	15
2.2. Endomorphismes auto-adjoints et théorème de diagonalisation simultanée.....	19
2.3. Orthogonalité. Orthonormalisation de Gram-Schmidt.....	23
2.4. Bases directes ou indirectes. Groupes $O(n)$ et $SO(n)$ . Étude de $O(2)$ et $O(3)$ .....	26
2.5. Appendice (†) : mesure des angles dans $\mathbb{R}^2$ .....	38
2.6. Appendice (†) : décomposition d'Iwasawa de $GL_n(\mathbb{R})$ .....	39
<b>3. Espaces affines euclidiens</b> .....	41
3.1. Définitions.....	41
3.2. Classification des isométries en petite dimension.....	44
3.3. Appendice : Espaces affines réels.....	45
<b>4. Formes hermitiennes, espaces hilbertiens et groupes unitaires</b> .....	53
4.0. Rappels sur les nombres complexes.....	53
4.1. Formes hermitiennes.....	53
4.2. Réduction en sommes de carrés de modules.....	58
4.3. Espaces hilbertiens. Inégalité de Cauchy-Schwarz. Isométries.....	61
4.4. Diagonalisation des endomorphismes auto-adjoints et normaux.....	64
4.5. Forme normale des éléments de $O(n)$ .....	66
4.6. Appendice (†) : espaces préhilbertiens réels ou complexes.....	69



## CHAPITRE 1

# FORMES BILINÉAIRES SYMÉTRIQUES ET FORMES QUADRATIQUES

**Résumé :** Dans ce chapitre, on commence l'étude des formes bilinéaires symétriques et des formes quadratiques, le cas le plus important étant sans doute celui du produit scalaire euclidien sur  $\mathbb{R}^n$ , qui sera étudié en détail dans le chapitre suivant. Dans le présent chapitre, on introduit les généralités concernant les formes bilinéaires symétriques : matrice dans une base, formule de changement de base, notion d'orthogonalité. Les résultats fondamentaux à retenir sont 1.2.5, le théorème de Sylvester 1.2.7, et l'algorithme de réduction en « somme de carrés » 1.3.2.

On a indiqué par des symboles  $\Downarrow$  les définitions, exemples et résultats fondamentaux. Par ailleurs, dans un appendice à la fin du chapitre on étudie les formes bilinéaires alternées.

### 1.0. Introduction

Soit  $V$  un  $\mathbb{R}$ -espace vectoriel de dimension  $n$ , une forme bilinéaire  $\phi$  sur  $V$  est un moyen d'associer à tout couple de vecteurs  $(v, w)$  un scalaire  $\phi(v, w) \in \mathbb{R}$ , qui dépend linéairement de  $v$  et de  $w$ . Un exemple typique est le produit scalaire usuel sur  $V = \mathbb{R}^3$ , dans ce cas, la fonction

$$v \mapsto \Phi(v) = \phi(v, v)$$

mesure la longueur du vecteur  $v$ ; on dit que  $\Phi$  est une forme quadratique car son expression en fonction des coordonnées  $(x_1, x_2, x_3)$  de  $v$  dans une base arbitraire  $\mathcal{B}$  de  $V$  est un polynôme homogène de degré 2 en  $x_1, x_2, x_3$ .

D'autres formes quadratiques apparaissent de façon naturelle. Par exemple, dans la théorie de la relativité on remplace l'espace  $\mathbb{R}^3$  par l'espace-temps  $\mathbb{R}^4$  muni de la forme quadratique de Lorentz et Minkowski :

$$\Psi(x, y, z, t) = x^2 + y^2 + z^2 - t^2$$

(noter que  $\Psi$  prend des valeurs  $> 0$  et  $< 0$  et qu'il existe des  $v \in \mathbb{R}^4 - \{0\}$  tels que  $\Psi(v) = 0$ ), elle correspond à la forme bilinéaire symétrique  $\psi$  sur  $\mathbb{R}^4$  définie par

$$\psi \left( \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix}, \begin{pmatrix} x' \\ y' \\ z' \\ t' \end{pmatrix} \right) = xx' + yy' + zz' - tt'.$$

Ceci conduit à étudier de façon générale, pour un espace vectoriel  $V$  de dimension finie  $n$  sur le corps  $k = \mathbb{R}$  ou  $\mathbb{C}$ , les formes bilinéaires symétriques  $V \times V \rightarrow k$  et les formes quadratiques associées.

### 1.1. Formes bilinéaires

Dans ce chapitre, on suppose que le corps de base  $k$  est  $\mathbb{Q}$ ,  $\mathbb{R}$  ou  $\mathbb{C}$ .

**Définition 1.1.1 (Formes bilinéaires).** — Soit  $E$  un  $k$ -espace vectoriel (pas nécessairement de dimension finie).



(1) Une **forme bilinéaire** sur  $E$  est une application  $\phi : E \times E \rightarrow k$  qui est linéaire en chaque variable, c.-à-d., pour tout  $x, x', y, y' \in E$  et  $\lambda, \mu \in k$ , on a :

$$\phi(\lambda x + x', y) = \lambda\phi(x, y) + \phi(x', y)$$

$$\phi(x, \mu y + y') = \mu\phi(x, y) + \phi(x, y')$$

et donc  $\phi(\lambda x + x', \mu y + y') = \lambda\mu\phi(x, y) + \lambda\phi(x, y') + \mu\phi(x', y) + \phi(x', y')$ . On rappelle que les formes bilinéaires sur  $E$  forment un espace vectoriel, noté  $\mathcal{L}_2(E, k)$ .

(2) Une forme bilinéaire  $\phi$  sur  $E$  est dite **symétrique** si pour tout  $x, y \in E$  on a  $\phi(x, y) = \phi(y, x)$ . Les formes bilinéaires symétriques (en abrégé : **fs**) sur  $E$  forment un sous-espace vectoriel de  $\mathcal{L}_2(E, k)$ , qu'on notera  $\mathcal{S}_2(E, k)$ .

(3) Une forme bilinéaire  $\phi$  sur  $E$  est dite **alternée** si elle vérifie  $\phi(x, x) = 0$  pour tout  $x \in E$ . Ceci entraîne que  $\phi$  est *antisymétrique*, i.e.  $\phi(y, x) = -\phi(x, y)$  : en effet, on a

$$0 = \phi(x + y, x + y) = \phi(x, x) + \phi(y, y) + \phi(x, y) + \phi(y, x)$$

et comme les deux premiers terme à droite sont nuls, on a  $\phi(x, y) + \phi(y, x) = 0$ . <sup>(1)</sup>

Les formes bilinéaires alternées (en abrégé : **fa**) sur  $E$  forment un sous-espace vectoriel de  $\mathcal{L}_2(E, k)$ , noté  $\mathcal{A}_2(E, k)$ .

**Remarque 1.1.2.** — Remarquons que, pour vérifier qu'une application  $\phi : E \times E \rightarrow k$  est une forme bilinéaire symétrique, il suffit de voir que  $\phi$  est linéaire en la 1ère variable et vérifie  $\phi(x, y) = \phi(y, x)$  (ces deux conditions entraînant la linéarité en la 2ème variable).

**Exemple 1.1.3.** — Considérons l'espace vectoriel  $E = \mathbb{R}^2$  et notons  $e_1, e_2$  la base canonique. Les applications

$$\phi_1(x_1e_2 + x_2e_2, y_1e_1 + y_2e_2) = x_1y_1 + x_2y_2,$$

$$\phi_2(x_1e_2 + x_2e_2, y_1e_1 + y_2e_2) = x_1y_1 - x_2y_2,$$

$$\phi_3(x_1e_2 + x_2e_2, y_1e_1 + y_2e_2) = x_1y_2 - x_2y_1,$$

$$\phi_4(x_1e_2 + x_2e_2, y_1e_1 + y_2e_2) = x_1y_2,$$

sont des formes bilinéaires. Les formes  $\phi_1$  et  $\phi_2$  sont symétriques, par contre  $\phi_3$  est alternée et, quant à  $\phi_4$ , on a  $\phi_4(e_1, e_2) = 1$  mais  $\phi_4(e_2, e_1) = 0$ .

**Exemple 1.1.4.** — Considérons  $I = [0, 1]$  l'intervalle unitaire et  $E = \mathcal{C}^0(I, \mathbb{R})$  le  $\mathbb{R}$ -espace vectoriel des fonctions continue sur  $I$  à valeurs réelles. Pour toute fonction  $f, g \in E$  on pose

$$\phi(f, g) := \int_0^1 f(x)g(x)dx.$$

L'application  $\phi : E \times E \rightarrow \mathbb{R}$  ainsi définie est bilinéaire.

**Proposition 1.1.5.** — Supposons que  $k$  soit de caractéristique différente de 2.

(1) Considérons une forme bilinéaire  $\phi$  sur  $E$  et posons, pour tout  $x, y \in E$ ,

$$\phi_s(x, y) = \frac{\phi(x, y) + \phi(y, x)}{2}, \quad \phi_a(x, y) = \frac{\phi(x, y) - \phi(y, x)}{2}.$$

Alors,  $\phi_s$  est une forme bilinéaire symétrique,  $\phi_a$  est une forme bilinéaire alternée, et on a  $\phi = \phi_s + \phi_a$ .

(2)  $\mathcal{L}_2(E, k) = \mathcal{S}_2(E, k) \oplus \mathcal{A}_2(E, k)$ .

*Démonstration.* — (1) C'est évident. (2) D'après (1),  $\mathcal{S}_2(E, k)$  et  $\mathcal{A}_2(E, k)$  engendrent  $\mathcal{L}_2(E, k)$  donc il s'agit de vérifier que l'intersection de ces deux sous-espace est nulle. En fait, si  $\phi$  est une forme bilinéaire symétrique et alternée, alors pour tout  $x, y \in E$  on a

$$\phi(x, y) = \phi(y, x) = -\phi(x, y),$$

et donc  $\phi(x, y) = 0$ . (On a utilisé la symétrie dans la première égalité et l'anti-symétrie dans la deuxième.  $\square$ )

---

Désormais, on suppose  $E$  de dimension finie  $n$ .

**Définition 1.1.6 (Matrices symétriques et alternées).** — Soit  $A \in M_n(k)$  une matrice.

<sup>(1)</sup> Si  $k$  est de caractéristique  $\neq 2$ , alors alternée  $\Leftrightarrow$  antisymétrique.



(1) On dit que  $A$  est **symétrique** si  ${}^tA = A$ . Ces matrices forment un sous-espace vectoriel de  $M_n(k)$ , qu'on note  $\text{MS}_n(k)$ .

(2) On dit que  $A$  est **alternée** si  $a_{ii} = 0$  pour tout  $i$  et  $a_{ji} = -a_{ij}$  pour tout  $i \neq j$ . Ceci équivaut à dire que  $A$  est *antisymétrique*, i.e.  ${}^tA = -A$  et que, de plus,  $a_{ii} = 0$  pour tout  $i$ .<sup>(2)</sup> Ces matrices forment un sous-espace vectoriel de  $M_n(k)$ , qu'on notera  $\text{MA}_n(k)$ .

**Lemme 1.1.7.** — On a :

$$(1) \quad \dim \text{MS}_n(k) = n(n+1)/2;$$

$$(2) \quad \dim \text{MA}_n(k) = n(n-1)/2.$$

*Démonstration.* — Notons  $N$  le nombre de coefficients qui sont strictement au-dessus de la diagonale ; c'est aussi le nombre de coefficients qui sont strictement en-dessous de la diagonale, et il y a  $n$  coefficients diagonaux. Donc  $2N + n = n^2$ , d'où  $N = n(n-1)/2$ .

(1) Une matrice symétrique est déterminée par le choix de ses  $n$  coefficients diagonaux et de ses  $N$  coefficients au-dessus de la diagonale, d'où  $\dim \text{MS}_n(k) = n + N = n(n+1)/2$ .

(2) Une matrice alternée est déterminée par le choix de ses  $N$  coefficients au-dessus de la diagonale, d'où  $\dim \text{MA}_n(k) = N = n(n-1)/2$ .  $\square$

**Définition et théorème 1.1.8 (Matrice d'une fbs et changement de base)**

Soit  $E$  est un  $k$ -espace vectoriel de dimension finie  $n$  et soit  $\mathcal{B} = (e_1, \dots, e_n)$  une base de  $E$ .

(1) Soit  $\phi$  une forme bilinéaire sur  $E$ . Pour tout  $i, j$  on pose  $a_{ij} = \phi(e_i, e_j)$  et on considère la matrice  $A = (a_{ij})_{i,j=1}^n$ . Alors, pour tout  $x_1, \dots, x_n \in k$  et tout  $y_1, \dots, y_n \in k$  on a

$$(*) \quad \phi \left( \sum_{i=1}^n x_i e_i, \sum_{j=1}^n y_j e_j \right) = \sum_{i,j=1}^n x_i y_j \phi(e_i, e_j) = \sum_{i,j=1}^n a_{ij} x_i y_j.$$

Donc, si l'on pose  $X = {}^t(x_1, \dots, x_n)$  et  $Y = {}^t(y_1, \dots, y_n)$  les vecteurs colonnes ci-dessus, on a la formule matricielle :

$$\phi \left( \sum_{i=1}^n x_i e_i, \sum_{j=1}^n y_j e_j \right) = {}^tX A Y.$$

En particulier, la matrice  $A$  détermine complètement la forme bilinéaire  $\phi$ . On dit que  $A$  est la matrice de  $\phi$  dans la base  $\mathcal{B}$  et on la note  $\text{Mat}_{\mathcal{B}}(\phi)$ .

(2) Réciproquement soit  $B \in M_n(k)$  une matrice. Alors l'application  $\psi: E \times E \rightarrow k$  définie pour tout  $X = {}^t(x_1, \dots, x_n) \in k^n$  et tout  $Y = {}^t(y_1, \dots, y_n) \in k^n$ , par

$$\psi \left( \sum_{i=1}^n x_i e_i, \sum_{j=1}^n y_j e_j \right) = {}^tX B Y,$$

est bilinéaire et vérifie  $B = \text{Mat}_{\mathcal{B}}(\psi)$ .

(3) L'application  $\mu_{\mathcal{B}}: \mathcal{L}_2(E, k) \rightarrow M_n(k)$ ,  $\phi \mapsto \text{Mat}_{\mathcal{B}}(\phi)$  est un isomorphisme de  $k$ -espaces vectoriels. De plus le sous-espace des fbs  $\mathcal{S}_2(E, k)$  (resp. le sous-espaces des fba  $\mathcal{A}_2(E, k)$ ) s'identifie à travers  $\mu_{\mathcal{B}}$  au sous-espaces des matrices symétriques  $\text{MS}_n(k)$  (resp. au sous-espace des matrices).

(4) Soient  $\mathcal{B}'$  une autre base de  $E$  et  $P$  la matrice de passage  $\text{Mat}_{\mathcal{B}}(\mathcal{B}')$ . Alors

$$(**) \quad A' = \text{Mat}_{\mathcal{B}'}(\phi) = {}^tP A P.$$

**Terminologie 1.1.8.1.** — On dit que  $A, A' \in M_n(k)$  sont **congruentes** s'il existe  $P \in \text{GL}_n(k)$  telle que  $A' = {}^tP A P$ .

*Démonstration.* — (1) Comme  $\phi$  est bilinéaire, on a bien l'égalité (\*), qui montre que  $\phi$  est déterminée par sa matrice, donc que l'application  $\mu_{\mathcal{B}}: \mathcal{L}_2(E) \rightarrow M_n(k)$ ,  $\phi \mapsto \text{Mat}_{\mathcal{B}}(\phi)$  est *injective*. D'autre part, on voit que le scalaire  $\sum_{i,j=1}^n a_{ij} x_i y_j \in k$  est égal au produit matriciel

$${}^tX A Y = (x_1, \dots, x_n) A \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}.$$

<sup>(2)</sup>Cette seconde condition étant automatiquement vérifiée si  $k$  est de caractéristique  $\neq 2$ .

Démontrons que  $A$  est bien définie. En fait, supposons qu'on ait  $\phi(X, Y) = {}^tX A' Y$  pour une certaine matrice  $A' = (a'_{ij}) \in M_n(k)$ . Comme  $e_i$  (resp.  $e_j$ ) est le vecteur colonne dont toutes les coordonnées sont nulles sauf la  $i$ -ème (resp.  $j$ -ème) qui vaut 1, on obtient que

$$a_{ij} = \phi(e_i, e_j) = (0, \dots, 1, \dots, 0) B \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} = a'_{ij},$$

d'où  $A' = A$ .

Démontrons que  $\phi$  est complètement déterminée par  $A$ . Supposons que  $\phi'$  soit une forme bilinéaire telle que  $A' := \text{Mat}_{\mathcal{B}}(\phi') = A$ . En particulier, pour tout  $X = {}^t(x_1, \dots, x_n)$  et  $Y = {}^t(x_1, \dots, x_n)$  on a

$$\phi \left( \sum_{i=1}^n x_i e_i, \sum_{j=1}^n y_j e_j \right) = {}^tX A Y = {}^tX A' Y = \phi' \left( \sum_{i=1}^n x_i e_i, \sum_{j=1}^n y_j e_j \right).$$

(2) Pour tout  $B = (B_{ij})_{i,j=1}^n \in \text{MS}_n(k)$ , l'application  $\psi$  est linéaire en les  $x_i$ , d'une part, et en les  $y_j$ , d'autre part, et elle donc bilinéaire. De plus, prenant  $x_{i_0} = 1 = y_{j_0}$  et  $x_i = 0 = y_j$  pour  $i \neq i_0$  et  $j \neq j_0$ , on obtient  $\phi_A(e_{i_0}, e_{j_0}) = a_{i_0, j_0}$  pour tout  $i_0, j_0 = 1, \dots, n$ , d'où  $\text{Mat}_{\mathcal{B}}(\phi_A) = A$ .

(3) Démontrons que l'application  $\mu_{\mathcal{B}}$  est linéaire. En effet, si  $\phi, \psi \in \mathcal{S}_2(E, k)$  et  $s \in k$ , alors  $s\phi + \psi$  est la forme bilinéaire symétrique définie par  $(s\phi + \psi)(u, v) = s\phi(u, v) + \psi(u, v)$  pour toute  $u, v \in E$ , donc *a fortiori* on a  $(s\phi + \psi)(e_i, e_j) = s\phi(e_i, e_j) + \psi(e_i, e_j)$  pour tout  $i, j$ , d'où  $\mu_{\mathcal{B}}(s\phi + \psi) = s\mu_{\mathcal{B}}(\phi) + \mu_{\mathcal{B}}(\psi)$ . L'injectivité de  $\mu_{\mathcal{B}}$  suit (1) et la surjectivité de (2) : l'application linéaire  $\mu_{\mathcal{B}}$  est donc un isomorphisme.

Il est en outre clair que si  $\phi$  est une fbs alors la matrice  $\text{Mat}_{\mathcal{B}}(\phi)$  est symétrique. D'autre part, si  $B = (b_{ij})$  est une matrice symétrique, pour tout  $x_1, \dots, x_n \in k$  et tout  $y_1, \dots, y_n \in k$  on a

$$\psi(y, x) = \psi \left( \sum_{j=1}^n y_j e_j, \sum_{i=1}^n x_i e_i \right) = \sum_{i,j=1}^n \underbrace{a_{ji}}_{=a_{ij}} y_j x_i = \sum_{i,j=1}^n a_{ij} x_i y_j = \psi(x, y),$$

et donc  $\psi$  est une forme bilinéaire symétrique. La vérification dans le cas des formes alternées est similaire et laissée au lecteur.

(4) Soient  $x, y \in E$ , ils correspondent dans la base  $\mathcal{B}$  (resp.  $\mathcal{B}'$ ) à des vecteurs colonnes  $X, Y$  (resp.  $X', Y'$ ). D'après la formule de changement de coordonnées, on a  $X = PX'$  et  $Y = PY'$ , d'où  ${}^tX = {}^tX' {}^tP$ , et donc :

$$\phi(x, y) = {}^tX A Y = {}^tX' {}^tP A P Y'$$

ce qui entraîne  $A' = {}^tP A P$ . Le théorème est démontré.  $\square$

**Remarque 1.1.9.** — On peut interpréter le fait que  $\mu_{\mathcal{B}}$  est un isomorphisme en disant que *se donner une forme bilinéaire symétrique sur  $E$  « est la même chose » que se donner une matrice symétrique.*

**Exemple 1.1.10.** — Reprenons les notations de l'exemple 1.1.3. Si on désigne par  $\mathcal{B}$  la base canonique de  $\mathbb{R}^2$ , on a

$$\begin{aligned} \text{Mat}_{\mathcal{B}}(\phi_1) &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} & \text{Mat}_{\mathcal{B}}(\phi_2) &= \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \\ \text{Mat}_{\mathcal{B}}(\phi_3) &= \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} & \text{Mat}_{\mathcal{B}}(\phi_4) &= \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}. \end{aligned}$$

### Définition et proposition 1.1.11 (Rang et noyau d'une forme bilinéaire)

Soit  $\phi$  une forme bilinéaire sur un  $k$ -espace vectoriel  $E$  de dimension  $n$  et soient  $\mathcal{B} = (e_1, \dots, e_n)$  une base de  $E$  et  $A = \text{Mat}_{\mathcal{B}}(\phi)$ .

- (1) On définit le **rang** de  $\phi$  par  $\text{rang}(\phi) = \text{rang}(A)$  ; ceci ne dépend pas du choix de la base  $\mathcal{B}$ .
- (2) Le **noyau**  $N(\phi)$  de  $\phi$  est défini par :

$$N(\phi) = \{y \in E \mid \forall x \in E, \phi(x, y) = 0\}.$$

C'est un sous-espace vectoriel de  $E$ , égal à  $\text{Ker}(A)$ . <sup>(3)</sup>

<sup>(3)</sup> Il est parfois noté  $\text{Ker}(\phi)$ , mais on évitera cette notation, pour éviter la confusion avec le noyau d'une application linéaire.



(3) On dit que  $\phi$  est **non dégénérée** si  $\text{rang}(\phi) = \dim E$ , i.e. si sa matrice dans une (et donc dans toute) base de  $E$  est inversible. Ceci équivaut à dire que  $N(\phi) = \{0\}$ .

*Démonstration.* — (1) Soient  $\mathcal{B}'$  une autre base de  $E$  et  $P = \text{Mat}_{\mathcal{B}}(\mathcal{B}')$ . Comme  $P$  et  ${}^tP$  sont inversibles (on a  $({}^tP)^{-1} = {}^t(P^{-1})$ ), alors la matrice  $A' = \text{Mat}_{\mathcal{B}'}(\phi) = {}^tPAP$  a même rang que  $A$ . Ceci prouve (1).

Prouvons (2). Soit  $y \in E$ , représenté dans la base  $\mathcal{B}$  par le vecteur colonne  $Y$ . Si  $AY = 0$ , alors pour  $x \in E$ , représenté par le vecteur colonne  $X$ , on a

$$\phi(x, y) = {}^tXAY = 0$$

et donc  $y \in N(\phi)$ . Réciproquement, supposons que  $y \in N(\phi)$  et notons  $Z$  le vecteur colonne  $AY$ . Alors pour tout vecteur colonne  $X$ , on a  $0 = {}^tXZ$ . En prenant pour  $X$  le vecteur dont toutes les coordonnées sont nulles sauf la  $i$ -ème qui vaut 1, on obtient que la  $i$ -ème coordonnée de  $Z$  est nulle. Donc  $AY = Z$  est nul, i.e.  $Y \in \text{Ker}(A)$ . Ceci prouve l'égalité  $N(\phi) = \text{Ker}(A)$ ; en particulier,  $N(\phi)$  est un sous-espace vectoriel de  $E$ . Ceci redémontre aussi que  $\text{rang}(A) = n - \dim N(\phi)$  ne dépend pas de la base  $\mathcal{B}$  choisie. Enfin, on a  $\text{rang}(A) = n$  si et seulement si  $N(\phi) = \{0\}$ , ce qui prouve l'équivalence énoncée dans (3).  $\square$

Attention, la matrice  $\text{Mat}_{\mathcal{B}}(\phi)$  n'est **pas** la matrice d'un endomorphisme de  $E$ ; toutefois on a la proposition suivante.

**Exemple 1.1.12.** — Reprenons les notations de l'exemple 1.1.3. Le noyau de  $\phi_1$ ,  $\phi_2$  et  $\phi_3$  est trivial et donc elles sont non-dégénérées. Par contre,  $\phi_4$  est dégénérée : en effet, on a  $N(\phi_4) = \text{Vect}(e_1)$ .

**Proposition 1.1.13.** — Soient  $E$  un  $k$ -espace vectoriel de dimension  $n$  et  $\phi$  une forme bilinéaire sur  $E$ .

(1) Pour tout  $y \in E$ , l'application  $\theta(y) : E \rightarrow k$  définie par  $\theta(y)(x) = \phi(x, y)$  est linéaire, donc appartient à l'espace dual  $E^*$ .

(2) L'application  $\theta : E \rightarrow E^*$ ,  $y \mapsto \theta(y)$ , est linéaire.

(3) Pour toute base  $\mathcal{B}$  de  $E$ , notant  $\mathcal{B}^*$  la base duale de  $E^*$ , la matrice  $\text{Mat}_{\mathcal{B}}(\phi)$  définie plus haut n'est autre que  $\text{Mat}_{(\mathcal{B}^*, \mathcal{B})}(\theta)$ .

(4) On a  $N(\phi) = \text{Ker}(\theta)$ . Par conséquent,  $\phi$  est non dégénérée si et seulement si  $\theta$  est un isomorphisme  $E \xrightarrow{\sim} E^*$ .

*Démonstration.* — (1) résulte de la linéarité de  $\phi$  en la 1ère variable : pour  $y \in E$  fixé, l'application  $x \mapsto \phi(x, y)$  est linéaire en  $x$ .

Montrons (2) : il faut voir que pour tout  $y, y' \in E$  et  $\mu \in k$ , les formes linéaires  $\theta(\mu y + y')$  et  $\mu\theta(y) + \theta(y')$  sont égales, c.-à-d., prennent la même valeur sur tout  $x \in E$ . Or, par définition, pour tout  $x \in E$  on a :

$$\theta(\mu y + y')(x) \stackrel{\text{déf}}{=} \phi(x, \mu y + y') \quad (\mu\theta(y) + \theta(y'))(x) \stackrel{\text{déf}}{=} \mu\theta(y)(x) + \theta(y')(x) \stackrel{\text{déf}}{=} \mu\phi(x, y) + \phi(x, y')$$

donc l'égalité à vérifier se ramène à l'égalité :

$$\forall x \in E, \quad \phi(x, \mu y + y') = \mu\phi(x, y) + \phi(x, y'),$$

qui est bien vérifiée, puisque  $\phi$  est linéaire en la 2ème variable. Ceci prouve (2). Alors, on a

$$\text{Ker}(\theta) = \{y_0 \in E \mid \theta(y_0) = 0\} = \{y_0 \in E \mid \phi(x, y_0) = 0, \quad \forall x \in E\} = N(\phi).$$

Comme  $\dim E^* = n = \dim E$ , on obtient donc les équivalences :

$$\phi \text{ non dégénérée} \Leftrightarrow N(\phi) = (0) \Leftrightarrow \theta \text{ est injectif} \Leftrightarrow \theta \text{ est un isomorphisme,}$$

d'où (4). Enfin, soient  $\mathcal{B}$  une base de  $E$ ,  $\mathcal{B}^*$  la base duale de  $E^*$ . Comme  $\theta(e_j)(e_i) = \phi(e_i, e_j)$ , pour tout  $i = 1, \dots, n$ , on a  $\theta(e_j) = \sum_{i=1}^n \phi(e_i, e_j)e_i^*$  et ceci prouve que  $\text{Mat}_{(\mathcal{B}^*, \mathcal{B})}(\theta) = \text{Mat}_{\mathcal{B}}(\phi)$ , d'où (3).  $\square$

**Exemple 1.1.14.** — Reprenons les notations de l'exemple 1.1.3. Pour tout  $i = 1, \dots, 4$  on note par  $\theta_i : E \rightarrow E^*$  l'application associée par la proposition précédente à la forme bilinéaire  $\phi_i$ . Si  $e_1^*, e_2^* \in E^*$  est la base duale à la base canonique, on a

$$\begin{cases} \theta_1(e_1) = e_1^* \\ \theta_1(e_2) = e_2^* \end{cases} \quad \begin{cases} \theta_2(e_1) = e_1^* \\ \theta_2(e_2) = -e_2^* \end{cases}$$

$$\begin{cases} \theta_3(e_1) = e_2^* \\ \theta_3(e_2) = -e_1^* \end{cases} \quad \begin{cases} \theta_4(e_1) = 0 \\ \theta_4(e_2) = e_1^*. \end{cases}$$

**Définition 1.1.15 (Termes carrés et doubles produits).** — Soit  $\phi$  une forme bilinéaire symétrique sur  $E$ . En séparant, d'une part, les « termes carrés »  $x_i y_i$  et, d'autre part, les « doubles produits »  $x_i y_j$  avec  $i \neq j$ , la formule (\*) de 1.1.8 se réécrit de la façon suivante (puisque  $a_{ji} = a_{ij}$  pour tout  $i \neq j$ ) :

$$(*) \quad \forall \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \in k^n, \quad \phi \left( \sum_{i=1}^n x_i e_i, \sum_{j=1}^n y_j e_j \right) = \sum_{i=1}^n a_{ii} x_i y_i + \sum_{1 \leq i < j \leq n} a_{ij} (x_i y_j + x_j y_i).$$

La terminologie « termes carrés » et « doubles produits » est justifiée par le fait que si les deux vecteurs sont égaux, i.e. si  $y_i = x_i$  pour tout  $i$ , le terme de droite ci-dessus devient :

$$(*) \quad \sum_{i=1}^n a_{ii} x_i^2 + \sum_{1 \leq i < j \leq n} 2a_{ij} x_i x_j.$$

(Noter le coefficient 2 qui apparaît avant les doubles produits.)

**Définition 1.1.16.** — Soit  $\phi$  une forme bilinéaire sur un  $k$ -espace vectoriel  $E$ .

(1) On dit que deux vecteurs  $x, y \in E$  sont **orthogonaux** (pour  $\phi$ ) si  $\phi(x, y) = 0$ . Plus généralement, on dit que deux sous-ensembles  $X, Y$  de  $E$  sont orthogonaux si l'on a  $\phi(x, y) = 0$  pour tout  $x \in X$  et  $y \in Y$ . On notera  $X \perp Y$  pour signifier que  $X$  et  $Y$  sont orthogonaux.

(2) Pour tout sous-ensemble  $Y$  de  $E$ , on définit son **orthogonal** (relativement à  $\phi$ ), noté  $Y^\perp$  ou simplement  $Y^\perp$ , par :

$$Y^\perp = \{x \in E \mid \phi(x, y) = 0, \quad \forall y \in Y\}$$

**Exemple 1.1.17.** — Reprenons les notations de l'exemple 1.1.3 et considérons  $F = \text{Vect}(e_1)$ . Pour tout  $i = 1, \dots, 4$  désignons par  $F_i^\perp$  l'orthogonal à  $F$  par rapport à la forme bilinéaire  $\phi_i$ . Avec ces notations, on a  $F_1 = F_2 = \text{Vect}(e_2)$ ,  $F_3 = \text{Vect}(e_1)$  et  $F_4 = E$ .

**Définition et proposition 1.1.18 (Orthogonalité).** — Soit  $\phi$  une forme bilinéaire sur un  $k$ -espace vectoriel  $E$ . Pour tout sous-ensembles  $Y \subset E$ , l'orthogonal  $Y^\perp$  est un **sous-espace vectoriel** de  $E$  (même si  $Y$  n'en est pas un). De plus, on a les propriétés suivantes :

$$(1) \quad Y \subseteq Z \implies Z^\perp \subseteq Y^\perp;$$

$$(2) \quad Y^\perp = \text{Vect}(Y)^\perp;$$

(3) Si  $Y$  est un sous-espace vectoriel  $F$  de  $E$  et si  $(f_1, \dots, f_p)$  est une famille génératrice de  $F$ , alors

$$F^\perp = \{f_1, \dots, f_p\}^\perp = \{x \in E \mid \phi(x, f_i) = 0, \quad \forall i = 1, \dots, p\}.$$

En particulier, on a  $N(\phi) = E^\perp = \{x \in E \mid \phi(x, y) = 0 \text{ pour tout } y \in Y\}$ .

*Démonstration.* — Soient  $x, x' \in Y^\perp$  et  $\lambda \in k$ , alors on a, pour tout  $y \in Y$ ,  $\phi(\lambda x + x', y) = \lambda \phi(x, y) + \phi(x', y) = 0$ , ce qui montre que  $\lambda x + x' \in Y^\perp$ . Donc  $Y^\perp$  est un sous-espace vectoriel de  $E$ .

Il est immédiat que si  $Y \subseteq Z$ , alors  $Z^\perp \subseteq Y^\perp$  car si  $x \in Z^\perp$  alors  $x$  est orthogonal à tout élément de  $Z$ , donc  $x$  est a fortiori orthogonal à tout élément de  $Y$  (puisque  $Y \subseteq Z$ ), donc  $x \in Y^\perp$ .

Comme  $Y \subseteq \text{Vect}(Y)$ , ceci donne déjà l'inclusion  $\text{Vect}(Y)^\perp \subseteq Y^\perp$ . Montrons l'inclusion réciproque. Soit  $x \in Y^\perp$  et soit  $v$  un élément arbitraire de  $\text{Vect}(Y)$ , par définition,  $v$  s'écrit comme une combinaison linéaire finie  $v = \lambda_1 y_1 + \dots + \lambda_r y_r$ , avec  $y_i \in Y$  et  $\lambda_i \in k$ ; alors on a

$$\phi(x, v) = \sum_{i=1}^r \lambda_i \underbrace{\phi(x, y_i)}_{=0} = 0$$

et donc  $x \in \text{Vect}(Y)^\perp$ . Ceci montre l'inclusion  $Y^\perp \subseteq \text{Vect}(Y)^\perp$ , d'où l'égalité  $\text{Vect}(Y)^\perp = Y^\perp$ . L'assertion (2) est démontrée.  $\square$

**Théorème 1.1.19 (Orthogonal d'un sous-espace).** — Soit  $\phi$  une forme bilinéaire symétrique ou alternée sur un  $k$ -espace vectoriel  $E$  de dimension finie. Soit  $F$  un sous-espace vectoriel de  $E$ . Alors,

$$(1) \quad F \subseteq (F^\perp)^\perp \quad \text{et} \quad \dim F^\perp \geq \dim E - \dim F.$$

$$(2) \quad N(\phi) = \{0\} \Leftrightarrow \phi \text{ est non dégénérée.}$$

$$(3) \quad \text{Si } \phi \text{ est non dégénérée, on a } \dim F^\perp = \dim E - \dim F \quad \text{et} \quad F = (F^\perp)^\perp.$$

$$(4) \quad \text{Si } F \cap F^\perp = \{0\}, \text{ alors } E = F \oplus F^\perp.$$

*Démonstration.* — Posons  $r = \dim F$  et  $n = \dim E$ .

(1) Soit  $f \in F$ , pour tout  $x \in F^\perp$  on a  $\phi(f, x) = 0$ , d'où  $f \in (F^\perp)^\perp$ . Ceci montre la première assertion. Prouvons la seconde.

Soit  $(f_1, \dots, f_r)$  une base de  $F$ , complétons-la en une base  $\mathcal{B} = (f_1, \dots, f_n)$  de  $E$ , et soit  $A = (a_{ij})_{1 \leq i, j \leq n}$  la matrice de  $\phi$  dans la base  $\mathcal{B}$ , i.e.  $a_{ij} = \phi(f_i, f_j)$  pour  $i, j = 1, \dots, n$ .

D'après le point (2) de 1.1.18,  $F^\perp$  est formé des vecteurs  $v = x_1 f_1 + \dots + x_n f_n \in E$  tels que  $\phi(f_i, v) = 0$  pour  $i = 1, \dots, r$ . Comme  $\phi(f_i, x_1 f_1 + \dots + x_n f_n) = \sum_{j=1}^n x_j \phi(f_i, f_j) = \sum_{j=1}^n a_{ij} x_j$ , ceci équivaut à dire

que le vecteur colonne  $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$  est solution du système linéaire homogène :

$$(\Sigma) \quad \begin{cases} a_{11} x_1 + \dots + a_{1n} x_n = 0 \\ \vdots \\ a_{r1} x_1 + \dots + a_{rn} x_n = 0 \end{cases}$$

dont la matrice  $B$  est formée des  $r$  premières lignes de  $A$ . Comme l'espace des solutions du système est de dimension  $n - \text{rang}(B)$ , on obtient :

$$\dim F^\perp = n - \text{rang}(B) \geq n - r,$$

ce qui prouve la seconde assertion de (1).

(2) D'après (1), dans le cas particulier où  $F = E$ , on a  $B = A$  et l'on obtient que  $\dim E^\perp = n - \text{rang}(A)$ . Donc  $N(\phi) = E^\perp$  est nul si et seulement si  $\text{rang}(A) = n$ .

(3) Supposons  $\phi$  non dégénérée. Alors  $A$  est de rang  $n$ , i.e. ses lignes sont linéairement indépendantes, en particulier les  $r$  premières lignes le sont, donc la matrice  $B$  est de rang  $r$ , et donc  $\dim F^\perp = n - r$ . Remplaçant alors  $F$  par  $F^\perp$ , on obtient l'égalité  $\dim(F^\perp)^\perp = n - (n - r) = r$ , et par conséquent l'inclusion  $F \subseteq (F^\perp)^\perp$  est une égalité.

(4) Supposons  $F \cap F^\perp = \{0\}$  (sans supposer  $\phi$  non dégénérée). Alors  $F$  et  $F^\perp$  sont en somme directe, et le sous-espace  $F \oplus F^\perp$  de  $E$  est de dimension  $d = r + \dim F^\perp$ . D'après (1), on a  $d \geq n$ , d'où  $E = F \oplus F^\perp$  (et  $\dim F^\perp = n - r$ ). Ceci termine la preuve de (4) et du théorème.  $\square$

**Définition 1.1.20 (Restriction à un sous-espace).** — Soit  $\phi$  une forme bilinéaire sur  $E$ , et soit  $F$  un sous-espace de  $E$ .

(1) On note  $\phi_F$  la forme bilinéaire sur  $F$  obtenue en restreignant  $\phi$  à  $F \times F$ , i.e.  $\phi_F(x, y) = \phi(x, y)$  pour tout  $x, y \in F$ ; on l'appelle la *restriction* de  $\phi$  à  $F$ .

(2) On a  $F \cap F^\perp = \{x \in F \mid \phi(x, y) = 0 \text{ pour tout } y \in F\} = N(\phi_F)$ . Donc l'assertion (4) de 1.1.19 peut se récrire comme suit : « si  $\phi_F$  est non dégénérée, alors  $E = F \oplus F^\perp$  ».

**Remarque 1.1.20.1.** — Attention, même si  $\phi$  est non dégénérée,  $\phi_F$  ne l'est pas nécessairement. Par exemple, soit  $\psi$  la forme bilinéaire symétrique sur  $\mathbb{R}^2$  définie par

$$\psi(u, v) = x_1 y_2 + x_2 y_1 \quad \text{si } u = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \text{ et } v = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix};$$

sa matrice dans la base canonique  $\mathcal{B} = (e_1, e_2)$  de  $\mathbb{R}^2$  est  $A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ , donc  $\psi$  est non dégénérée. Cependant, on a  $\psi(e_1, e_1) = 0 = \psi(e_2, e_2)$  donc pour  $F = \mathbb{R}e_1$  (ou  $\mathbb{R}e_2$ ), on a  $\psi_F = 0$ .

## 1.2. Formes quadratiques

**Définition 1.2.1 (Formes quadratiques).** — Soit  $E$  un  $k$ -espace vectoriel de dimension  $n$ .

(1) Une **forme quadratique**  $Q$  sur  $E$  est une fonction  $E \rightarrow k$  de la forme  $x \mapsto \phi(x, x)$ , où  $\phi$  est une forme bilinéaire symétrique sur  $E$ .

(2) Comme  $\phi(x + y, x + y) = \phi(x, x) + \phi(y, y) + 2\phi(x, y) = Q(x) + Q(y) + 2\phi(x, y)$ , on a les égalités :

$$(*) \quad \boxed{\phi(x, y) = \frac{1}{2}(Q(x + y) - Q(x) - Q(y)) = \frac{1}{4}(Q(x + y) - Q(x - y))}$$

qui montrent que :  $\phi$  est entièrement déterminée par  $Q$ ; on dit que  $\phi$  est la **forme polaire** de  $Q$  (et les égalités (\*) ci-dessus sont appelées « égalités de polarisation »).



(3) Soient  $\mathcal{B} = (e_1, \dots, e_n)$  une base de  $E$  et  $A = (a_{ij})_{i,j=1}^n$  la matrice de  $\phi$  dans la base  $\mathcal{B}$  (i.e.  $a_{ij} = \phi(e_i, e_j)$ ), on a  $a_{ji} = a_{ij}$  puisque  $\phi$  est symétrique. Alors pour tout  $u = x_1 e_1 + \dots + x_n e_n$  dans  $E$ , on a :

$$Q(u) = \phi\left(\sum_i x_i e_i, \sum_j x_j e_j\right) = \sum_{i=1}^n a_{ii} x_i^2 + \sum_{i < j} 2a_{ij} x_i x_j$$

donc  $Q(u)$ , considéré comme une fonction  $Q(x_1, \dots, x_n)$  des coordonnées  $(x_1, \dots, x_n)$  de  $u$  dans la base  $\mathcal{B}$ , est donné par :

$$(\star) \quad \boxed{Q(x_1, \dots, x_n) = \sum_{i=1}^n a_{ii} x_i^2 + \sum_{1 \leq i < j \leq n} 2a_{ij} x_i x_j}$$

on prendra garde au facteur 2 qui apparaît avant les « doubles produits »  $x_i x_j$  !

(4) Réciproquement, si c'est  $Q$  qui est donnée par une formule explicite  $Q(x_1, \dots, x_n) = \sum_{i=1}^n b_i x_i^2 + \sum_{1 \leq i < j \leq n} b_{ij} x_i x_j$  (les  $x_i$  étant les coordonnées dans la base  $\mathcal{B}$ ), alors sa forme polaire  $\phi$  est donnée par :

$$(\star') \quad \forall \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \in k^n, \quad \phi\left(\sum_{i=1}^n x_i e_i, \sum_{j=1}^n y_j e_j\right) = \sum_{i=1}^n b_i x_i y_i + \sum_{1 \leq i < j \leq n} \frac{b_{ij}}{2} (x_i y_j + x_j y_i)$$

et donc  $A = \text{Mat}_{\mathcal{B}}(\phi)$  a pour coefficients diagonaux les  $b_i$ , mais ses coefficients non diagonaux sont les  $b_{ij}/2$ .

**Définitions 1.2.2 (Rang et noyau d'une forme quadratique).** — Soient  $E$  un  $k$ -espace vectoriel de dimension finie,  $Q$  une forme quadratique sur  $E$ , et  $\phi$  sa forme polaire.

(1) On définit le rang et le noyau de  $Q$  par :

$$\boxed{\text{rang}(Q) = \text{rang}(\phi)} \quad \boxed{N(Q) = N(\phi)}$$

(2) Attention !  $N(Q)$  n'est pas égal, en général, à l'ensemble  $C(Q) = \{x \in E \mid Q(x) = 0\}$ . On dit qu'un vecteur  $x \in E$  est **isotrope** (pour  $Q$ ) si  $Q(x) = 0$ , et l'ensemble  $C(Q)$  des vecteurs isotropes s'appelle le **cône isotrope** de  $Q$ . Ce n'est pas, en général, un sous-espace vectoriel de  $E$ , mais c'est un *cône*, i.e. il vérifie la propriété suivante : si  $x \in C(Q)$  et  $\lambda \in k$ , alors  $\lambda x \in C(Q)$ .

**Remarque 1.2.3.** — On a toujours  $\boxed{N(Q) \subseteq C(Q)}$  mais l'inclusion est en général stricte. Par exemple, dans le cas considéré en 1.1.20.1, on a  $N(Q) = N(\psi) = \{0\}$  mais le cône isotrope est la réunion des deux droites  $\mathbb{R}e_1$  et  $\mathbb{R}e_2$ .

**Définition 1.2.4 (Bases orthogonales).** — Soit  $E$  un  $k$ -espace vectoriel de dimension  $n$  et soient  $\phi$  une forme bilinéaire symétrique sur  $E$ , et  $Q$  la forme quadratique associée. Soit  $\mathcal{B} = (e_1, \dots, e_n)$  une base de  $E$

a) On dit que  $\mathcal{B}$  est une base **orthogonale** pour  $\phi$  (ou pour  $Q$ ) si l'on  $\phi(e_i, e_j) = 0$  pour  $i \neq j$ .

b) Ceci équivaut à dire que la matrice  $A = \text{Mat}_{\mathcal{B}}(\phi)$  est **diagonale**; si l'on note  $\lambda_1, \dots, \lambda_n$  ses coefficients diagonaux et  $(x_1, \dots, x_n)$  les coordonnées dans la base  $\mathcal{B}$ , ceci équivaut encore à dire que  $Q(x_1, \dots, x_n) = \lambda_1 x_1^2 + \dots + \lambda_n x_n^2$ .

**Théorème 1.2.5 (Existence de bases orthogonales pour une fbs).** — Soit  $\phi$  une forme bilinéaire symétrique sur un  $k$ -espace vectoriel  $E$  de dimension  $n$ , et soit  $Q$  la forme quadratique associée.

(1) Il existe une base  $\mathcal{B}$  de  $E$  orthogonale pour  $\phi$ .

(2) Soient  $\mathcal{B} = (e_1, \dots, e_n)$  une base orthogonale pour  $\phi$  et  $D$  la matrice diagonale  $\text{Mat}_{\mathcal{B}}(\phi)$ . Quitte à renuméroter les  $e_i$ , on peut supposer que les coefficients diagonaux  $\lambda_1, \dots, \lambda_r$  sont non nuls, et que  $\lambda_i = 0$  pour  $i > r$ . Notons  $(x_1, \dots, x_n)$  les coordonnées dans la base  $\mathcal{B}$ , alors :

$$(a) \text{ On a } \boxed{Q(x_1, \dots, x_n) = \lambda_1 x_1^2 + \dots + \lambda_r x_r^2} \quad (*)$$

(b) On a  $\boxed{r = \text{rang}(\phi)}$ , plus précisément,  $N(\phi)$  est le sous-espace  $\text{Vect}(e_{r+1}, \dots, e_n)$ , donné par les équations  $x_1 = 0 = \dots = x_r$ .

*Démonstration.* — (1) Montrons l'existence d'une base orthogonale en procédant par récurrence sur  $n = \dim E$ . Il n'y a rien à montrer si  $n = 0$  ou si  $\phi = 0$ . On peut donc supposer  $n \geq 1$  et le résultat établi pour  $n - 1$ , et  $\phi \neq 0$ . Alors, d'après le point (2) de 1.2.1, la forme quadratique  $Q$  est non nulle, donc il existe  $e_1 \in E$  tel que  $Q(e_1) \neq 0$ . Posons  $F = ke_1$ , comme  $\phi(e_1, e_1) \neq 0$ , alors  $F \cap F^\perp = \{0\}$  donc, d'après le théorème 1.1.19, on a

$$E = F \oplus F^\perp.$$

Par hypothèse de récurrence, il existe une base  $(e_2, \dots, e_n)$  de  $F^\perp$  telle que  $\phi(e_i, e_j) = 0$  pour  $i \neq j$ . Alors  $(e_1, e_2, \dots, e_n)$  est une base de  $E$  orthogonale pour  $\phi$ . Ceci prouve l'assertion (1).

Puis, (2.a) et la première assertion de (2.b) découlent aussitôt des définitions; prouvons la dernière assertion. D'après la démonstration du théorème 1.1.19, on sait que  $N(\phi)$  est égal au noyau de  $D$ , qui est bien le sous-espace  $F = \text{Vect}(e_{r+1}, \dots, e_n)$ , donné par les équations  $x_1 = 0 = \dots = x_r$ . Mais ceci peut se voir directement ici, de la façon suivante. D'après (\*),  $\phi$  est donnée dans la base  $\mathcal{B}$  par :

$$(*) \quad \forall u = \sum_{i=1}^n x_i e_i, \quad \forall v = \sum_{j=1}^n y_j e_j, \quad \phi(u, v) = \lambda_1 x_1 y_1 + \dots + \lambda_r x_r y_r.$$

Supposons  $u \in N(\phi)$ , alors pour tout  $i = 1, \dots, r$ , prenant  $v = e_i$  (c.-à-d.,  $y_i = 1$  et  $y_j = 0$  pour  $j \neq i$ ), on obtient  $x_i = 0$ , d'où  $u \in F = \text{Vect}(e_{r+1}, \dots, e_n)$ . Réciproquement, (\*) montre aussi que tout  $u \in F$  (i.e. tel que  $x_1 = 0 = \dots = x_r$ ) appartient à  $N(\phi)$ , d'où l'égalité désirée. Le théorème est démontré.  $\square$

Le théorème précédent est valable pour tout corps  $k$  de caractéristique  $\neq 2$ . La possibilité d'effectuer des réductions supplémentaires dépend de propriétés « arithmétiques » de  $k$ , c.-à-d., de quels éléments de  $k$  sont des carrés. Lorsque  $k = \mathbb{C}$  ou  $\mathbb{R}$ , on peut donner des versions plus précises.



**Théorème 1.2.6 (Formes quadratiques sur  $\mathbb{C}$ ).** — Soient  $E$  un  $\mathbb{C}$ -espace vectoriel de dimension finie,  $Q$  une forme quadratique sur  $E$  et  $\phi$  sa forme polaire. Il existe une base  $\mathcal{B}$  de  $E$  pour laquelle  $\text{Mat}_{\mathcal{B}}(\phi)$  est la matrice diagonale de termes diagonaux  $(1, \dots, 1, 0, \dots, 0)$ , le nombre de 1 étant égal à  $r = \text{rang}(\phi)$ ; si l'on note  $(x_1, \dots, x_n)$  les coordonnées dans la base  $\mathcal{B}$ , on a alors  $Q(x_1, \dots, x_n) = x_1^2 + \dots + x_r^2$ .

*Démonstration.* — Soit  $r = \text{rang}(\phi)$ . D'après le théorème 1.2.5, il existe une base orthogonale  $(e_1, \dots, e_n)$  telle que  $Q(e_i) \neq 0$  pour  $i \leq r$ , et  $Q(e_i) = 0$  pour  $i > r$ . Pour tout  $i = 1, \dots, r$ , soit  $\lambda_i \in \mathbb{C}$  tel que  $\lambda_i^2 = Q(e_i)$ . Remplaçant  $e_i$  par  $e_i/\lambda_i$ , pour  $i \leq r$ , on obtient une base  $\mathcal{B}$  ayant la propriété énoncée dans le théorème.  $\square$



**Théorème 1.2.7 (Théorème d'inertie de Sylvester).** — Soient  $E$  un  $\mathbb{R}$ -espace vectoriel de dimension  $n$ ,  $Q$  une forme quadratique sur  $E$  et  $\phi$  sa forme polaire.

- (1) Soit  $\mathcal{B} = (e_1, \dots, e_n)$  une base orthogonale pour  $\phi$  et soient  $p$  (resp.  $q$ ) le nombre d'indices  $i$  tels que  $Q(e_i) > 0$  (resp.  $< 0$ ). Alors  $p$  et  $q$  ne dépendent pas de la base orthogonale choisie.
- (2) Le couple  $(p, q)$  s'appelle la **signature** de  $Q$  (ou de  $\phi$ ); on a  $\text{rang}(\phi) = p + q$ .
- (3) De plus, on peut choisir  $\mathcal{B}$  de sorte que la matrice diagonale  $D = \text{Mat}_{\mathcal{B}}(\phi)$  ait pour termes diagonaux  $(1, \dots, 1, -1, \dots, -1, 0, \dots, 0)$ , le nombre de 1 (resp. -1) étant  $p$  (resp.  $q$ ).

*Démonstration.* — Posons  $r = \text{rang}(\phi)$ . Soient  $\mathcal{B} = (e_1, \dots, e_n)$  et  $\mathcal{C} = (f_1, \dots, f_n)$  deux bases de  $E$  orthogonales pour  $\phi$ . Notons  $p$  (resp.  $p'$ ) le nombre d'indices  $i$  tels que  $Q(e_i) > 0$  (resp.  $Q(f_i) > 0$ ) et  $q$  (resp.  $q'$ ) le nombre d'indices  $i$  tels que  $Q(e_i) < 0$  (resp.  $Q(f_i) < 0$ ). Alors

$$r = p + q = p' + q'$$

et il s'agit de montrer que  $q = q'$  et  $p = p'$ . Quitte à renuméroter les éléments de  $\mathcal{B}$  et  $\mathcal{C}$ , on peut supposer que

$$(*) \quad \begin{cases} Q(e_i) > 0 & \text{pour } i = 1, \dots, p \\ Q(e_i) < 0 & \text{pour } i = p + 1, \dots, p + q \\ Q(e_i) = 0 & \text{pour } i > p + q = r; \end{cases} \quad \begin{cases} Q(f_i) > 0 & \text{pour } i = 1, \dots, p' \\ Q(f_i) < 0 & \text{pour } i = p' + 1, \dots, p' + q' \\ Q(f_i) = 0 & \text{pour } i > p' + q' = r. \end{cases}$$

Notons  $P_+$  le sous-espace de  $E$  engendré par les vecteurs  $e_i$  tels que  $Q(e_i) \geq 0$ . Ces vecteurs sont au nombre de  $n - q$ , donc  $\dim P_+ = n - q$ . Soit  $x$  un élément arbitraire de  $P_+$ , écrivons  $x = \sum_{i \in I} x_i e_i$ , avec  $I = \{1, \dots, p\} \cup \{r + 1, \dots, n\}$ ; alors, d'après (\*), on obtient

$$(1) \quad Q(x) = \sum_{i=1}^p x_i^2 Q(e_i) \geq 0.$$

D'autre part, soit  $P'_-$  le sous-espace de  $E$  engendré par les vecteurs  $f_j$  tels que  $Q(f_j) < 0$ . Ces vecteurs sont au nombre de  $q'$ , donc  $\dim P'_- = q'$ . Soit  $y$  un élément non nul de  $P'_-$ , on peut écrire  $y = \sum_{j=p'+1}^{p'+q'} y_j f_j$ ,

avec au moins l'un des  $y_j$  non nul (car  $y \neq 0$ ). Alors, d'après  $(\star)$  à nouveau, on obtient

$$(2) \quad Q(y) = \sum_{j=p'+1}^{p'+q'} y_j^2 Q(f_j) < 0.$$

Par conséquent, on a  $P_+ \cap P'_- = \{0\}$  et donc

$$n = \dim E \geq \dim P_+ + \dim P'_- = n - q + q'$$

d'où  $q \geq q'$ . Échangeant les rôles des bases  $\mathcal{B}$  et  $\mathcal{C}$ , on obtient de même  $q' \geq q$ , d'où  $q = q'$ , et de même  $p = p'$ . Ceci prouve la première assertion du théorème.

Voyons la deuxième assertion. Soit  $\mathcal{B} = (e_1, \dots, e_n)$  comme ci-dessus; pour  $i = 1, \dots, p + q$ , notons  $|Q(e_i)| > 0$  la valeur absolue de  $Q(e_i)$ . En remplaçant  $e_i$  par  $e_i / \sqrt{|Q(e_i)|}$ , pour  $i = 1, \dots, p + q$ , on obtient une base orthogonale ayant la propriété énoncée dans le théorème.  $\square$

### 1.3. Réduction d'une forme quadratique en « somme de carrés »

À nouveau, on suppose que le corps de base est  $\mathbb{Q}$ ,  $\mathbb{R}$  ou  $\mathbb{C}$ .

**Remarque 1.3.0.** — Soit  $E$  un  $k$ -espace vectoriel de dimension  $n$ . D'après 1.2.1, il revient au même de se donner sur  $E$  une forme quadratique  $Q$  ou sa forme polaire  $\phi$ .

Le langage des formes quadratiques permet d'être plus concis : si  $E$  est muni d'une base  $\mathcal{B}$ , et donc de coordonnées  $(x_1, \dots, x_n)$  relativement à cette base (par exemple, si  $E = k^n$ ), on dira simplement, disons pour  $n = 3$  : « soit  $Q$  la forme quadratique  $ax_1^2 + bx_2^2 + cx_3^2 + dx_1x_2 + ex_1x_3 + fx_2x_3$  », ce qui est plus rapide que d'écrire : soit  $\phi$  la forme bilinéaire symétrique définie par :

$$\begin{aligned} \phi(x_1e_1 + x_2e_2 + x_3e_3, y_1e_1 + y_2e_2 + y_3e_3) = \\ ax_1y_1 + bx_2y_2 + cx_3y_3 + \frac{d}{2}(x_1y_2 + x_2y_1) + \frac{e}{2}(x_1y_3 + x_3y_1) + \frac{f}{2}(x_2y_3 + x_3y_2). \end{aligned}$$

De même, le fait d'écrire une forme quadratique comme un polynôme (homogène) de degré 2 en les coordonnées  $x_i$ , i.e.

$$Q(x_1, \dots, x_n) = \sum_{i=1}^n a_i x_i^2 + \sum_{1 \leq i < j \leq n} a_{ij} x_i x_j$$

permet d'effectuer sur ce polynôme des opérations algébriques simples, qui équivalent à trouver une base orthogonale pour  $\phi$  : c'est ce qu'on explique ci-dessous.

**Définition 1.3.1.** — Soient  $E$  un  $k$ -espace vectoriel de dimension  $n$ ,  $Q$  une forme quadratique sur  $E$  et  $\phi$  sa forme polaire. Soit  $\mathcal{B} = (e_1, \dots, e_n)$  une base de  $E$ , notons  $(x_1, \dots, x_n)$  les coordonnées dans cette base, i.e.  $x_i$  désigne en fait la forme linéaire  $f_i = e_i^*$  sur  $E$ .

(1) On dit que  $Q$  s'écrit dans la base  $\mathcal{B}$  comme **somme de carrés de formes linéaires indépendantes** si l'expression de  $Q$  en fonction des coordonnées  $x_i$  est de la forme

$$Q = q_1 x_1^2 + \dots + q_n x_n^2.$$

D'après 1.2.1, ceci équivaut à dire que la matrice de  $\phi$  dans la base  $\mathcal{B}$  est **diagonale**, avec les  $q_i$  pour coefficients diagonaux.

(2) Les formes linéaires  $f_i = e_i^*$  sont linéairement indépendantes ( $\mathcal{B}^* = (e_1^*, \dots, e_n^*)$  est la base duale de  $\mathcal{B}$ ), d'où la terminologie « somme de carrés de **formes linéaires indépendantes** ». En pratique, pour abrégé on écrira souvent « somme de carrés », mais il est essentiel de s'assurer que les formes linéaires en question sont bien linéairement indépendantes (voir plus bas).

#### **Théorème 1.3.2 (Réduction d'une forme quadratique en somme de carrés)**

Soient  $E$  un  $k$ -espace vectoriel de dimension  $n$ , et  $Q$  une forme quadratique sur  $E$ , donnée dans une base  $\mathcal{B} = (e_1, \dots, e_n)$  par

$$(*) \quad \forall (x_1, \dots, x_n) \in k^n, \quad Q(x_1e_1 + \dots + x_ne_n) = \sum_i b_i x_i^2 + \sum_{i < j} b_{ij} x_i x_j.$$

(1) Par une suite d'opérations « élémentaires » (décrites dans la démonstration), on peut trouver un nouveau système de coordonnées  $(y_1, \dots, y_n)$  sur  $E$ , dans lequel  $Q$  s'écrit comme une somme de carrés, i.e. :

$$(**) \quad Q(y_1, \dots, y_n) = a_1 y_1^2 + \dots + a_n y_n^2.$$



(2) Le nombre de coefficients  $a_i$  non nuls est égal à  $r = \text{rang}(Q)$ , et si  $k = \mathbb{R}$ , la signature de  $Q$  est  $(p, q)$ , où  $p$  (resp.  $q$ ) est le nombre de coefficients  $a_i$  qui sont  $> 0$  (resp.  $< 0$ ).

(3) De plus,  $N(\phi)$  est le sous-espace vectoriel de  $E$  défini par les équations  $y_i = 0$ , pour  $i$  parcourant l'ensemble des  $i \in \{1, \dots, n\}$  tels que  $a_i \neq 0$ .

*Démonstration.* — Remarquons d'abord que si  $Q$  s'écrit sous la forme (\*\*) dans une base  $\mathcal{B}'$ , alors la matrice de sa forme polaire  $y$  est diagonale, avec les  $a_i$  pour coefficients diagonaux, d'où les assertions (2) et (3) du théorème, compte-tenu des théorèmes 1.2.5 et 1.2.7.

Il reste à donner une démonstration « algorithmique » de l'assertion (1). On procède par récurrence sur le nombre  $n$  de variables. Si  $n = 1$  on a  $Q(x_1 e_1) = b_1 x_1^2$ , et (\*\*) est vérifié. On peut donc supposer  $n > 1$  et le résultat démontré pour  $n - 1$ . Distinguons deux cas.

(a) Si dans l'écriture (\*) plus haut, il existe un coefficient « diagonal »  $b_i$  non nul, on peut supposer, quitte à changer l'ordre des coordonnées, que  $b_1 \neq 0$ . On considère alors la somme de **tous** les termes contenant la variable  $x_1$  et on l'écrit comme suit :

$$S = b_1 x_1^2 + \sum_{j=2}^n b_{1j} x_1 x_j = b_1 \left( x_1^2 + 2x_1 \underbrace{\left( \sum_{j=2}^n \frac{b_{1j}}{2b_1} x_j \right)}_{=L(x_2, \dots, x_n)} \right)$$

alors  $L$  est une forme linéaire ne contenant plus la variable  $x_1$  (i.e.  $L$  est une combinaison linéaire des formes linéaires  $e_2^*, \dots, e_n^*$ ). Puis, en utilisant que

$$(x_1 + L)^2 = x_1^2 + 2x_1 L + L^2, \quad \text{d'où} \quad x_1^2 + 2x_1 L = (x_1 + L)^2 - L^2,$$

on réécrit ceci sous la forme :

$$S = b_1 (x_1 + L)^2 - b_1 L^2 = b_1 \left( x_1 + \sum_{j=2}^n \frac{b_{1j}}{2b_1} x_j \right)^2 - \sum_{j=2}^n \frac{b_{1j}^2}{4b_1} x_j^2 - \sum_{2 \leq i < j \leq n} \frac{b_{1i} b_{1j}}{2b_1} x_i x_j.$$

Donc, en posant  $y_1 = x_1 + \sum_{j=2}^n \frac{b_{1j}}{2b_1} x_j$  (et  $b'_j = b_j - b_{1j}^2/4b_1$  pour  $j = 2, \dots, n$ , et  $b'_{ij} = b_{ij} - b_{1i} b_{1j}/2b_1$  pour  $2 \leq i < j \leq n$ ), on obtient une écriture :

$$(\dagger) \quad Q(y_1, x_2, \dots, x_n) = b_1 y_1^2 + \underbrace{\sum_{j=2}^n b'_j x_j^2 + \sum_{2 \leq i < j \leq n} b'_{ij} x_i x_j}_{Q_1(x_2, \dots, x_n)}$$

où la forme quadratique  $Q_1(x_2, \dots, x_n)$  ne dépend que des variables  $x_2, \dots, x_n$ .

L'opération  $y_1 = x_1 + L(x_2, \dots, x_n)$  et  $x_j = x_j$  pour  $j \geq 2$ , est bien un changement de coordonnées, car la matrice exprimant  $(y_1, x_2, \dots, x_n)$  en fonction de  $(x_1, \dots, x_n)$  est triangulaire avec des 1 sur la diagonale, donc inversible; explicitement le changement de coordonnées inverse est donné par  $x_j = x_j$  pour  $j \geq 2$  et  $x_1 = y_1 - L(x_2, \dots, x_n)$ .

Par hypothèse de récurrence on peut faire un changement de coordonnées  $(x_2, \dots, x_n) \rightarrow (y_2, \dots, y_n)$  tel que  $Q_1(x_2, \dots, x_n) = a_2 y_2^2 + \dots + a_n y_n^2$  d'où, d'après (†) :

$$Q(y_1, \dots, y_n) = a_1 y_1^2 + \dots + a_n y_n^2$$

(avec  $a_1 = b_1$ ), ce qui prouve le résultat voulu dans ce cas.

(b) Supposons au contraire que **tous** les coefficients « diagonaux »  $b_i$  soient nuls. Si  $Q = 0$ , il n'y a rien à montrer; sinon on peut supposer, quitte à changer l'ordre des coordonnées, que  $b_{12} \neq 0$ . Le plus simple est alors de procéder comme suit : faisons le changement de coordonnées

$$x_1 = x'_1 + x'_2, \quad x_2 = x'_1 - x'_2, \quad x_j = x'_j \quad \text{pour } j > 2$$

(c'est bien un changement de coordonnées, dont l'inverse est donné par  $x'_1 = (x_1 + x_2)/2$ ,  $x'_2 = (x_1 - x_2)/2$ ,  $x'_j = x_j$  pour  $j > 2$ ). Alors les termes  $b_{ij} x_i x_j$  (avec  $i < j$ ) se transforment comme suit :

$$\begin{cases} b_{12} x_1 x_2 \rightarrow b_{12} (x_1'^2 - x_2'^2) \\ b_{ij} x_i x_j \rightarrow b_{ij} x_i x_j \quad \text{si } i, j \geq 3 \end{cases} \quad \begin{cases} b_{1j} x_1 x_j \rightarrow b_{1j} (x_1' + x_2') x_j \quad \text{pour } j \geq 3 \\ b_{2j} x_2 x_j \rightarrow b_{2j} (x_1' - x_2') x_j \quad \text{pour } j \geq 3 \end{cases}$$

donc on obtient

$$Q(x'_1, \dots, x'_n) = b_{12} (x_1'^2 - x_2'^2) + \sum_{j=3}^n \left( (b_{1j} + b_{2j}) x_1' x_j + (b_{1j} - b_{2j}) x_2' x_j \right) + \sum_{3 \leq i < j \leq n} b_{ij} x_i' x_j'$$

et l'on est ramené au cas **(a)**, c.-à-d., on peut éliminer la variable  $x'_1$  et se ramener, à nouveau, au cas de  $n - 1$  variables. Le théorème est démontré.  $\square$

**Remarque 1.3.3.** — Dans le cas **(b)**, une méthode plus sophistiquée, qui permet d'éliminer en même temps les variables  $x_1$  et  $x_2$ , est la suivante. On la désignera par **(b')**. On considère la somme de **tous** les termes contenant  $x_1$  **ou**  $x_2$  et on l'écrit comme suit :

$$\begin{aligned} S &= b_{12} x_1 x_2 + \sum_{j=3}^n b_{1j} x_1 x_j + \sum_{j=3}^n b_{2j} x_2 x_j = b_{12} \left( x_1 + \sum_{j=3}^n \frac{b_{2j}}{b_{12}} x_j \right) \left( x_2 + \sum_{j=3}^n \frac{b_{1j}}{b_{12}} x_j \right) - \sum_{3 \leq j, \ell \leq n} \frac{b_{1j} b_{2\ell}}{b_{12}} x_j x_\ell \\ &= b_{12} \underbrace{\left( x_1 + \sum_{j=3}^n \frac{b_{2j}}{b_{12}} x_j \right)}_{=X} \underbrace{\left( x_2 + \sum_{j=3}^n \frac{b_{1j}}{b_{12}} x_j \right)}_{=Y} - \sum_{j=3}^n \frac{b_{1j} b_{2j}}{b_{12}} x_j^2 - \sum_{3 \leq j < \ell \leq n} \frac{2b_{1j} b_{2\ell}}{b_{12}} x_j x_\ell \end{aligned}$$

Puis, en utilisant l'égalité

$$(\star) \quad \boxed{XY = \frac{1}{4} \left( (X+Y)^2 - (X-Y)^2 \right)}$$

et en posant

$$x'_1 = \frac{1}{2}(X+Y) = \frac{1}{2} \left( x_1 + x_2 + \sum_{j=3}^n \frac{b_{1j} + b_{2j}}{b_{12}} x_j \right), \quad x'_2 = \frac{1}{2}(X-Y) = \frac{1}{2} \left( x_1 - x_2 + \sum_{j=3}^n \frac{b_{1j} - b_{2j}}{b_{12}} x_j \right),$$

on obtient :

$$Q(x'_1, x'_2, x_3, \dots, x_n) = b_{12} (x_1'^2 - x_2'^2) - \sum_{j=3}^n \frac{b_{1j} b_{2j}}{b_{12}} x_j^2 + \sum_{3 \leq j < \ell \leq n} c_{j\ell} x_j x_\ell$$

où  $c_{j\ell} = b_{j\ell} - 2b_{1j}b_{2\ell}/b_{12}$  pour tout  $j < \ell$  dans  $\{3, \dots, n\}$ .

Illustrons ceci par deux exemples : dans le premier n'apparaissent que des changements de coordonnées du type **(a)**.

**Exemple 1.3.4.** — Considérons dans  $\mathbb{R}^4$  la forme quadratique

$$Q = x_1^2 - 2x_1x_2 + 4x_1x_3 + 2x_1x_4 + x_2^2 + 4x_3^2 + 5x_4^2 - 4x_2x_3 + 6x_2x_4 - 4x_3x_4.$$

Considérant les termes contenant  $x_1$ , on écrit d'abord :

$$Q = \underbrace{(x_1 - x_2 + 2x_3 + x_4)^2}_{y_1} + 4x_4^2 + 8x_2x_4 - 8x_3x_4$$

puis

$$4x_4^2 + 8x_2x_4 - 8x_3x_4 = 4 \underbrace{(x_4 + x_2 - x_3)^2}_{y_4} - 4x_2^2 + 8x_2x_3 - 4x_3^2$$

puis  $-4x_2^2 + 8x_2x_3 - 4x_3^2 = -4 \underbrace{(x_2 - x_3)^2}_{y_2}$ . Donc en faisant successivement les changements de coordonnées :

$$y_1 = x_1 - x_2 + 2x_3 + x_4, \quad y_4 = x_4 + x_2 - x_3, \quad y_2 = x_2 - x_3, \quad y_3 = x_3,$$

on obtient que  $Q(y_1, y_4, y_2, y_3) = y_1^2 + 4y_4^2 - 4y_2^2$ . Donc  $Q$  est de signature  $(2, 1)$  et de rang  $2 + 1 = 3$ . Son noyau  $N(Q)$  est la droite définie par les équations  $y_2 = 0 = y_4 = y_1$ , donc, dans les coordonnées initiales, par les équations  $x_2 = x_3, x_4 = 0, x_1 + x_3 = 0$ .

**Exemple 1.3.5.** — Considérons dans  $\mathbb{R}^4$  la forme quadratique :

$$Q(x_1, x_2, x_3, x_4) = x_1^2 + x_2^2 + x_3^2 + 2x_1x_2 + 2x_1x_3 - 3x_2x_4 - 4x_3x_4 + 5x_3x_4.$$

Considérant les termes contenant  $x_1$ , écrivons d'abord :

$$Q(x_1, x_2, x_3, x_4) = (x_1 + x_2 + x_3)^2 - 5x_2x_3 - 4x_2x_4 + 5x_3x_4 = y_1^2 - 5x_2x_3 - 4x_2x_4 + 5x_3x_4.$$

1ère méthode : transformons le terme  $x_2x_3$  en posant  $x_2 = y_2 + y_3$  et  $x_3 = y_2 - y_3$ , on obtient :

$$Q(y_1, y_2, y_3, x_4) = y_1^2 - 5(y_2^2 - y_3^2) + y_2x_4 - 9y_3x_4 = y_1^2 - 5y_2^2 + y_2x_4 + 5y_3^2 - 9y_3x_4.$$

Puis  $-5y_2^2 + y_2x_4 = -5(y_2 - \frac{1}{10}x_4)^2 + \frac{1}{20}x_4^2$  donne, en posant  $z_2 = y_2 - \frac{1}{10}x_4$  :

$$Q(y_1, z_2, y_3, x_4) = y_1^2 - 5z_2^2 + 5y_3^2 - 9y_3x_4 + \frac{1}{20}x_4^2$$

Puis  $5y_3^2 - 9y_3x_4 = 5(y_3 - \frac{9}{10}x_4)^2 - \frac{81}{20}x_4^2$  donne, en posant  $z_3 = y_3 - \frac{9}{10}x_4$  :

$$Q(y_1, z_2, z_3, x_4) = y_1^2 - 5z_2^2 + 5z_3^2 - 4x_4^2.$$

Donc la signature de  $Q$  est  $(2, 2)$  et son rang est  $2 + 2 = 4$ , i.e.  $Q$  est non dégénérée.

2ème méthode : considérons tous les termes contenant  $x_2$  ou  $x_3$  et écrivons :

$$-5x_2x_3 - 4x_2x_4 + 5x_3x_4 = -5\underbrace{(x_2 - x_4)}_{=X} \underbrace{\left(x_3 + \frac{4}{5}x_4\right)}_{=Y} - 4x_4^2$$

alors, posant  $z_2 = \frac{1}{2}(X + Y) = \frac{1}{2}(x_2 + x_3 - \frac{4}{5}x_4)$  et  $z_3 = \frac{1}{2}(X - Y) = \frac{1}{2}(x_2 - x_3 - \frac{9}{5}x_4)$ , on obtient que

$$Q(y_1, z_2, z_3, x_4) = y_1^2 - 5z_2^2 + 5z_3^2 - 4x_4^2$$

et l'on retrouve le résultat précédent.

**Exemple 1.3.6 (Erreur à ne pas commettre!).** — Reprenons le calcul précédent, au moment où l'on obtient les termes  $Q_1(x_2, x_3, x_4) = -5x_2x_3 - 4x_2x_4 + 5x_3x_4$ . Il ne faut **pas** écrire que

$$\begin{aligned} -5x_2x_3 - 4x_2x_4 + 5x_3x_4 &= -\frac{5}{4}\left((x_2+x_3)^2 - (x_2-x_3)^2\right) - \left((x_2+x_4)^2 - (x_2-x_4)^2\right) + \frac{5}{4}\left((x_3+x_4)^2 - (x_3-x_4)^2\right) \\ &= -\frac{5}{4}y_1^2 + \frac{5}{4}y_2^2 - y_3^2 + y_4^2 + \frac{5}{4}y_5^2 - \frac{5}{4}y_6^2 \end{aligned}$$

où l'on a posé  $y_1 = x_2 + x_3$ ,  $y_2 = x_2 - x_3$ ,  $y_3 = x_2 + x_4$ ,  $y_4 = x_2 - x_4$ ,  $y_5 = x_3 + x_4$ ,  $y_6 = x_3 - x_4$ , et conclure que la signature est  $(3, 3)$  et le rang 6. Ceci est erroné (et la conclusion absurde!) : comme on part ici d'une forme quadratique  $Q_1$  en 3 variables, son rang est  $r \leq 3$  et donc on doit obtenir à la fin une somme ayant **au plus** 3 termes, or ici on en a écrit 6. L'erreur est que l'on n'a pas fait un vrai « changement de coordonnées », car on a introduit trop de formes linéaires, qui ne sont plus linéairement indépendantes : par exemple, on a  $y_4 = -y_3 + y_1 + y_2$ ,  $y_5 = y_3 - y_2$ ,  $y_6 = y_1 - y_3$ .

Donc, pour ne pas se tromper dans ces calculs, il vaut mieux procéder pas à pas, en effectuant à chaque pas une transformation de type **(a)**, **(b)** ou **(b')**. Il ne faut pas effectuer plusieurs opérations en même temps!

#### 1.4. Appendice (†) : formes bilinéaires alternées

Dans ce chapitre,  $k$  est un corps arbitraire et  $E$  est un  $k$ -espace vectoriel de dimension finie  $n$ .



**Théorème 1.4.1 (de la base symplectique).** — Soient  $E$  un  $k$ -espace vectoriel de dimension  $m$  et  $\phi$  une forme **symplectique** sur  $E$ . Alors  $m = 2n$  est pair, et il existe une base  $\mathcal{B} = (e_1, f_1, \dots, e_n, f_n)$  de  $E$  telle que, pour tout  $i, j$ ,

$$\phi(e_i, e_j) = 0 = \phi(f_i, f_j) \quad \text{et} \quad \phi(e_i, f_j) = \begin{cases} 1 & \text{si } i = j, \\ 0 & \text{sinon.} \end{cases}$$

Alors  $\text{Mat}_{\mathcal{B}}(\phi)$  est une matrice diagonale par blocs :

$$\text{Mat}_{\mathcal{B}}(\phi) = \begin{pmatrix} A_1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & A_n \end{pmatrix} \quad \text{où chaque } A_i \text{ égale } \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

*Démonstration.* — Par récurrence sur  $m$ . Il n'y a rien à montrer si  $m = 0$ . Supposons  $m > 0$  et soit  $e_1 \in E - \{0\}$ . Comme  $\phi$  est non dégénérée, il existe  $f_1 \in E$  tel que  $\phi(e_1, f_1) = t \neq 0$ , et remplaçant  $f_1$  par  $t^{-1}f_1$  on se ramène à  $t = 1$ . Alors la famille  $(e_1, f_1)$  est libre (car si  $v = se_1 + s'f_1 = 0$ , alors  $0 = \phi(e_1, v) = s'$  et  $0 = \phi(f_1, v) = -s$ ), donc forme une base  $\mathcal{B}_1$  du sous-espace  $P = ke_1 + kf_1$ . On a

$$\text{Mat}_{\mathcal{B}_1}(\phi_P) = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

donc  $\phi_P$  est non dégénérée, et donc d'après le théorème 1.1.19, on a

$$(*) \quad E = P \oplus P^\perp.$$

De plus, comme  $\phi$  est non dégénérée, on a  $(P^\perp)^\perp = P$ , et donc

$$N(\phi_{P^\perp}) = P^\perp \cap (P^\perp)^\perp = P^\perp \cap P = \{0\}.$$

Donc la forme alternée  $\phi_{P^\perp}$  est non dégénérée, i.e. c'est une forme symplectique sur  $P^\perp$ .

Alors, par hypothèse de récurrence,  $P^\perp$  est de dimension paire  $2(n-1)$  et admet une base  $\mathcal{C} = (e_2, f_2, \dots, e_n, f_n)$  vérifiant les conditions du théorème. Alors  $\mathcal{B} = \mathcal{B}_1 \cup \mathcal{C}$  est une base de  $E$  vérifiant les conditions du théorème; en particulier  $\dim E = 2n$ . Le théorème est démontré.  $\square$

La proposition qui suit est utile pour se ramener au cas d'une forme non dégénérée :

**Proposition 1.4.2 (Réduction au cas non dégénéré).** — Soient  $E$  un  $k$ -espace vectoriel de dimension  $n$ ,  $\phi \in \mathcal{L}_2(E, k)$  symétrique ou alternée,  $r = \text{rang}(\phi)$ ,  $S$  un supplémentaire de  $N(\phi)$  dans  $E$ , et  $\mathcal{B}$  (resp.  $\mathcal{C}$ ) une base de  $S$  (resp. de  $N(\phi)$ ).

(1) Alors la matrice de  $\phi$  dans la base  $\mathcal{B} \cup \mathcal{C}$  de  $E$  est

$$A = \left( \begin{array}{c|c} \text{Mat}_{\mathcal{B}}(\phi_S) & \mathbf{0}_{n-r,r} \\ \hline \mathbf{0}_{n-r,r} & \mathbf{0}_{n-r,n-r} \end{array} \right)$$

par conséquent,  $\phi_S$  est non dégénérée.

(2) D'autre part,  $\phi$  induit sur l'espace quotient  $\bar{E} = E/N(\phi)$  une forme bilinéaire  $\bar{\phi}$  telle que

$$(\dagger) \quad \bar{\phi}(\bar{x}, \bar{y}) = \phi(x, y), \quad \forall x, y \in E$$

et  $\bar{\phi}$  est non dégénérée.

*Démonstration.* — Il est clair que  $\text{Mat}_{\mathcal{B} \cup \mathcal{C}}(\phi)$  a la forme indiquée; alors  $\text{rang}(\phi_S) = \text{rang} \text{Mat}_{\mathcal{B}}(\phi_S)$  égale  $\text{rang}(A)$ , qui égale  $n - \dim N(\phi) = \dim S$ . Ceci montre que  $\phi_S$  est non dégénérée, d'où le point (1).

(2) Montrons que  $\bar{\phi}$  est bien définie, i.e. que la formule (†) fait sens. Soient  $x', y' \in E$  tels que  $\bar{x}' = \bar{x}$  et  $\bar{y}' = \bar{y}$ , alors  $u = x' - x$  et  $v = y' - y$  appartiennent à  $N(\phi)$ , d'où

$$\phi(x', y') = \phi(x + u, y + v) = \phi(x, y) + \underbrace{\phi(u, y) + \phi(x, v) + \phi(u, v)}_{=0 \text{ puisque } u, v \in N(\phi)}.$$

Ceci montre que  $\bar{\phi}$  est bien définie, et la formule (†) montre alors que  $\bar{\phi}$  est une forme bilinéaire (symétrique ou alternée) sur  $\bar{E}$ . De plus, si  $\bar{y} \in N(\phi)$ , alors pour tout  $x \in E$  on a  $\phi(x, y) = \bar{\phi}(\bar{x}, \bar{y}) = 0$ , d'où  $y \in N(\phi)$  et donc  $\bar{y} = 0$ . Ceci prouve que  $\bar{\phi}$  est non dégénérée.  $\square$

**Corollaire 1.4.3 (Bases standard pour une forme bilinéaire alternée)**

Soient  $E$  un  $k$ -espace vectoriel de dimension  $m$  et  $\phi$  une forme bilinéaire alternée sur  $E$ . Alors il existe une base  $(v_1, \dots, v_r, e_1, f_1, \dots, e_n, f_n)$  de  $E$ , où  $(v_1, \dots, v_r)$  est une base de  $N(\phi)$  et les  $e_i, f_i$  vérifient les conditions du théorème 1.4.1.

*Démonstration.* — Ceci résulte de la proposition 1.4.2 et du théorème 1.4.1.  $\square$

## CHAPITRE 2

### ESPACES EUCLIDIENS ET GROUPES ORTHOGONAUX $O(n)$

**Résumé :** Ce chapitre constitue le coeur de la partie « géométrique » du cours. Dans la section 1, on introduit les espaces vectoriels euclidiens, dont le prototype est  $\mathbb{R}^n$  muni du produit scalaire standard, puis on démontre l'inégalité de Cauchy-Schwarz, qui assure que le produit scalaire donne naissance à une « norme euclidienne » (d'où une notion de « distance », cf. 2.1.5). On introduit ensuite le groupe  $O(n)$  des isométries de l'espace euclidien  $\mathbb{R}^n$ . Dans la section 2, on démontre l'important théorème de diagonalisation simultanée 2.2.9. Dans la section 3, on introduit les projections et symétries orthogonales, ainsi que le procédé d'orthonormalisation de Gram-Schmidt. Enfin, dans la section 4, on étudie en profondeur les isométries de  $\mathbb{R}^2$  et  $\mathbb{R}^3$ . En résumé, ce chapitre contient beaucoup de résultats nouveaux et importants, qu'il faut essayer d'assimiler !

On a indiqué par des symboles  les définitions, exemples et résultats fondamentaux. Par ailleurs, des *compléments de cours*, pour les étudiants intéressés, sont donnés dans des appendices à la fin du chapitre ; ces passages n'interviendront pas dans les évaluations.

#### 2.1. Espaces euclidiens. Inégalité de Cauchy-Schwarz. Isométries

**Définitions 2.1.1 (Produits scalaires et espaces euclidiens).** — Soit  $E$  un  $\mathbb{R}$ -espace vectoriel, pas nécessairement de dimension finie.

(1) Soient  $\phi$  une forme bilinéaire symétrique sur  $E$  et  $Q$  la forme quadratique associée (i.e.  $Q(x) = \phi(x, x)$  pour tout  $x \in E$ ). On dit que  $Q$  (ou  $\phi$ ) est **définie positive** si l'on a :

(Déf. Pos.) 
$$\forall x \in E - \{0\}, \quad Q(x) = \phi(x, x) > 0.$$

Dans ce cas, on dit que  $\phi$  est un **produit scalaire** et on note souvent  $\phi(x, y) = (x | y)$ .

Remarquons que si  $Q$  (ou  $\phi$ ) est définie positive, elle est non-dégénérée : en effet, si  $x \in N(\phi)$ , on a  $0 = \phi(x, y)$  pour tout  $y \in E$ , en particulier  $\phi(x, x) = 0$ , d'où  $x = 0$ .

(2) Dans ce cas, on dit que : «  $E$ , muni de  $( | )$  » (ou que : « le couple  $(E, \phi)$  ») est un **espace euclidien**.

(1) Pour abrégé, on écrira souvent : « Soit  $E$  un espace euclidien », sans préciser le produit scalaire  $( | )$ , celui-ci étant sous-entendu.

**Exemples 2.1.2.** — (1)  $\mathbb{R}^n$  muni du produit scalaire euclidien standard :

$$(x | y) = x_1 y_1 + \cdots + x_n y_n \quad \text{si } x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \quad y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

et de la forme quadratique associée  $Q(x) = x_1^2 + \cdots + x_n^2$ , est un espace euclidien de dimension  $n$ . Pour ce produit scalaire, la base canonique  $(e_1, \dots, e_n)$  de  $\mathbb{R}^n$  est *orthonormée*, i.e. on a  $(e_i | e_j) = 1$  si  $i = j$  et  $= 0$  sinon.

<sup>(0)</sup>version du 13/5/2013

<sup>(1)</sup>En fait, on réserve d'habitude cette terminologie au cas où  $E$  est de dimension finie ; sinon on dit que  $E$  est un espace *préhilbertien réel* (voir l'explication de cette terminologie dans l'Appendice 4.6 à la fin du dernier chapitre). Nous n'utiliserons pas cette terminologie.

(2) L'espace vectoriel  $E = \mathcal{C}^0([0, 1], \mathbb{R})$  des fonctions continues  $f : [0, 1] \rightarrow \mathbb{R}$ , muni du produit scalaire

$$(f | g) = \int_0^1 f(t)g(t)dt,$$

est un espace euclidien, qui n'est pas de dimension finie.

### Définition et proposition 2.1.3 (Familles et bases orthonormées)

Soit  $E$ , muni de  $( | )$ , un espace euclidien.

(1) Une famille  $(e_i)_{i \in I}$  de vecteurs est dite **orthonormée** si  $(e_i | e_i) = 1$  et  $(e_i | e_j) = 0$  pour tout  $i \neq j$ .

(2) Supposons  $E$  de dimension  $n$ . Une **base orthonormée** est une base  $(e_1, \dots, e_n)$  de  $E$  qui est une famille orthonormée, i.e. qui vérifie  $(e_i | e_i) = 1$  et  $(e_i | e_j) = 0$  pour tout  $i \neq j$ .

(3) Toute famille orthonormée est libre. En particulier, si  $\dim E = n$ , toute famille orthonormée  $(f_1, \dots, f_n)$  de cardinal  $n$  est une base orthonormée de  $E$ .

(4) Dans la suite, on abrégera souvent « base orthonormée » en : *b.o.n.* ou *BON*.

*Démonstration.* — Prouvons (3). Supposons qu'on ait une relation  $0 = t_1 e_{i_1} + \dots + t_p e_{i_p}$ , avec  $i_1, \dots, i_p \in I$  deux à deux distincts, et  $t_1, \dots, t_p \in \mathbb{R}$ . Fixons un indice  $r \in \{1, \dots, p\}$  et appliquons  $(e_{i_r} | \cdot)$  à l'égalité précédente. Comme  $(e_{i_r} | e_{i_s}) = 0$  pour  $s \neq r$ , on obtient  $0 = t_r (e_{i_r} | e_{i_r}) = t_r$ , d'où  $t_r = 0$ . Ceci prouve que la famille  $(e_i)_{i \in I}$  est libre.  $\square$

**Théorème 2.1.4 (Existence de b.o.n.).** — Soit  $E$  un espace euclidien de dimension  $n$ . Alors  $E$  admet une base orthonormée.

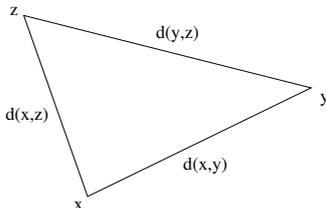
*Démonstration.* — D'après le théorème d'inertie de Sylvester 1.2.7, il existe une base  $(e_1, \dots, e_n)$  orthogonale (i.e.  $(e_i | e_j) = 0$  pour  $i \neq j$ ) et telle que  $(e_i | e_i) \in \{1, -1, 0\}$ ; or comme  $( | )$  est défini positif on a nécessairement  $(e_i | e_i) = 1$ , donc  $(e_1, \dots, e_n)$  est une b.o.n.  $\square$

**Définition 2.1.5 (Normes).** — Soit  $E$  un  $\mathbb{R}$ -espace vectoriel. Une **norme**  $\| \cdot \|$  sur  $E$  est une application  $E \rightarrow \mathbb{R}_+$ ,  $x \mapsto \|x\|$  vérifiant les trois propriétés suivantes :

- (1)  $\|x\| = 0 \Leftrightarrow x = 0$ .
- (2) Pour tout  $t \in \mathbb{R}$ ,  $x \in E$ , on a  $\|tx\| = |t| \cdot \|x\|$  (où  $|t|$  est la valeur absolue de  $t$ ).
- (3)  $\|u + v\| \leq \|u\| + \|v\|$ , pour tout  $u, v \in E$ .

Remarque. L'inégalité précédente est nommée **Inégalité triangulaire**, pour la raison suivante. Si on pose  $d(x, y) = \|y - x\|$ , pour tout  $x, y \in E$ , alors, compte-tenu de (1) et (2) ci-dessus, (3) équivaut à dire (en posant  $u = y - x$ ,  $v = z - y$ ) que l'application  $d : E \times E \rightarrow \mathbb{R}_+$  est une **distance** sur  $E$ , i.e. vérifie :

- (1')  $d(x, y) = 0 \Leftrightarrow x = y$ .
- (2')  $d(x, y) = d(y, x)$ .
- (3') Inégalité triangulaire : pour tout  $x, y, z \in E$ , on a :  $d(x, z) \leq d(x, y) + d(y, z)$



### Théorème 2.1.6 (Inégalité de Cauchy-Schwarz et norme euclidienne)

Soit  $E$ , muni de  $( | )$ , un espace euclidien et soit  $Q(x) = (x | x)$  la forme quadratique associée.

(1) On a l'inégalité de Cauchy-Schwarz :

$$(CS) \quad \forall x, y \in E \quad (x | y)^2 \leq Q(x)Q(y)$$

avec égalité si et seulement si  $x$  et  $y$  sont liés.

(2) Par conséquent, l'application  $x \mapsto \|x\| = \sqrt{(x | x)}$  est une norme sur  $E$ , appelée la **norme euclidienne** associée à  $( | )$ , et l'inégalité de Cauchy-Schwarz se réécrit comme suit (où dans le terme de gauche  $|\cdot|$  désigne la valeur absolue dans  $\mathbb{R}$ ) :

$$(CS) \quad \forall x, y \in E \quad |(x | y)| \leq \|x\| \cdot \|y\|.$$

*Démonstration.* — Si  $y = \lambda x$ , on a  $Q(y) = \lambda^2 Q(x)$  et  $(x | y)^2 = \lambda^2 (x | x)^2 = Q(y)Q(x)$ , et de même si  $x = \lambda y$ . Donc on a l'égalité si  $x, y$  sont liés, en particulier si  $x = 0$  ou  $y = 0$ . Supposons donc  $x$  et  $y$  non nuls ; pour tout  $t \in \mathbb{R}$ , on a :

$$0 \leq Q(tx + y) = t^2 Q(x) + 2t(x | y) + Q(y)$$

donc le discriminant réduit  $\Delta' = (x | y)^2 - Q(x)Q(y)$  de ce trinôme <sup>(2)</sup> en  $t$  est  $\leq 0$ , ce qui prouve l'inégalité (CS). De plus, si  $\Delta' = 0$  le trinôme ci-dessus a une racine double réelle  $t_0 = -(x | y)/Q(x)$ , et l'égalité  $Q(t_0 x + y) = 0$  entraîne, puisque  $Q$  est définie positive,  $t_0 x + y = 0$ , i.e.

$$y = \frac{(x | y)}{(x | x)} x.$$

Ceci prouve (1).

Prouvons que  $x \mapsto \|x\| = \sqrt{(x | x)}$  est une norme sur  $E$ . Comme  $( | )$  est défini positif, on a  $\|x\| = 0 \Leftrightarrow x = 0$ . D'autre part, pour tout  $t \in \mathbb{R}$  et  $x \in E$ , on a  $|t| = \sqrt{t^2}$  et donc

$$\|tx\| = \sqrt{t^2(x | x)} = |t| \cdot \|x\|.$$

Enfin, soient  $x, y \in E$ . D'abord, l'inégalité de Cauchy-Schwarz équivaut (en prenant la racine carrée) à :

$$|(x | y)| \leq \|x\| \cdot \|y\| ;$$

alors, multipliant par 2 et ajoutant  $\|x\|^2 + \|y\|^2$  aux deux membres, on obtient

$$\begin{aligned} \|x + y\|^2 &= \|x\|^2 + \|y\|^2 + 2(x | y) \leq \|x\|^2 + \|y\|^2 + 2\|x\| \|y\| \\ &\leq \|x\|^2 + \|y\|^2 + 2\|x\| \cdot \|y\| = (\|x\| + \|y\|)^2. \end{aligned}$$

Prenant la racine carrée, ceci entraîne (et équivaut à) l'inégalité triangulaire. Le théorème est démontré.  $\square$

Récrivons certaines conséquences de l'égalité  $(x + y | x + y) = (x | x) + (y | y) + 2(x | y)$  en utilisant la norme  $\| \cdot \|$  (ou plutôt son carré) :

**Proposition 2.1.7 (Pythagore, parallélogramme et médiane, polarisation)**

Soit  $E$  un espace euclidien, et soit  $\| \cdot \|$  la norme associée au produit scalaire  $( | )$ . On a les égalités suivantes :

(Pythagore)  $\|x_1 + \dots + x_n\|^2 = \|x_1\|^2 + \dots + \|x_n\|^2$  si  $x_1, \dots, x_n$  sont orthogonaux

(Parallélogramme/Médiane)  $\|x + y\|^2 + \|x - y\|^2 = 2\|x\|^2 + 2\|y\|^2$

(Polarisation)  $4(x | y) = \|x + y\|^2 - \|x - y\|^2$

*Démonstration.* — L'égalité de Pythagore est immédiate si  $n = 2$ , et dans ce cas on a même la réciproque : si  $\|x_1 + x_2\|^2 = \|x_1\|^2 + \|x_2\|^2$  alors  $(x_1 | x_2) = 0$ . L'égalité pour  $n$  vecteurs orthogonaux s'obtient par récurrence sur  $n$ . On prendra garde que la réciproque est fautive pour  $n \geq 3$  : prendre par exemple dans  $\mathbb{R}^2$  euclidien les vecteurs  $x_1 = e_1$ ,  $x_2 = e_1 + e_2$ ,  $x_3 = e_2 - e_1$ .

Les deux autres égalités s'obtiennent en ajoutant (resp. soustrayant) les égalités :

$$\|x + y\|^2 = (x + y | x + y) = \|x\|^2 + \|y\|^2 + 2(x | y)$$

$$\|x - y\|^2 = (x - y | x - y) = \|x\|^2 + \|y\|^2 - 2(x | y)$$

$\square$

**Remarques 2.1.7.1.** — La deuxième égalité s'appelle « identité du parallélogramme », car elle exprime que dans le parallélogramme construit sur les vecteurs  $x$  et  $y$ , la somme des carrés des longueurs des quatre côtés égale la somme des carrés des longueurs des deux diagonales (qui sont  $x + y$  et  $x - y$ ). Elle s'appelle aussi « identité de la médiane », car dans le triangle construit sur les vecteurs  $x$  et  $y$ , la « médiane » joignant 0 au milieu du côté  $x - y$  est  $(x + y)/2$ , et l'on a donc une formule exprimant (le carré de) la longueur de la médiane en fonction de la longueur des côtés :

$$\left\| \frac{x + y}{2} \right\|^2 = \frac{\|x\|^2}{2} + \frac{\|y\|^2}{2} - \frac{\|x - y\|^2}{4}.$$

<sup>(2)</sup>Pour un trinôme  $aX^2 + 2bX + c$  dont le coefficient de  $X$  est pair, il est commode de considérer le discriminant réduit  $\Delta' = b^2 - ac$  (au lieu du discriminant usuel  $\Delta = (2b)^2 - 4ac = 4\Delta'$ ).



Enfin, la dernière égalité est appelée « identité de polarisation », car elle exprime en fonction de la forme quadratique  $Q(x) = \|x\|^2$  le produit scalaire, qui est la « forme polaire » de  $Q$ . On l'a déjà rencontrée dans le Chap. 4 sous la forme  $4\phi(x, y) = Q(x + y) - Q(x - y)$ .

Avant d'introduire la définition suivante, rappelons que la fonction cosinus induit une **bijection** de  $[0, \pi]$  sur  $[-1, 1]$  (on a  $\cos(0) = 1$ ,  $\cos(\pi) = -1$ , et  $\cos$  est strictement décroissante sur l'intervalle  $[0, \pi]$ ).



**Définition 2.1.8 (Angle non orienté de deux vecteurs non nuls).** — Soit  $E$ , muni de  $(\cdot | \cdot)$ , un espace euclidien et soit  $\|\cdot\|$  la norme euclidienne. Soient  $u, v$  deux vecteurs non nuls. D'après l'inégalité de Cauchy-Schwarz, on a

$$|(u | v)| \leq \|u\| \cdot \|v\| \quad \text{d'où} \quad -1 \leq \frac{(u | v)}{\|u\| \cdot \|v\|} \leq 1$$

donc il existe un unique  $\theta \in [0, \pi]$  tel que  $\cos(\theta) = \frac{(u | v)}{\|u\| \cdot \|v\|}$  i.e.  $(u | v) = \cos(\theta) \|u\| \cdot \|v\|$ . On appelle  $\theta$  l'**angle non-orienté** des vecteurs  $u$  et  $v$ , il ne change pas si l'on échange  $u$  et  $v$ .



**Définition et proposition 2.1.9 (Isométries vectorielles).** — Soient  $E, F$  deux espaces euclidiens de même dimension  $n$ , notons  $(\cdot | \cdot)_E$  et  $\|\cdot\|_E$  (resp.  $(\cdot | \cdot)_F$  et  $\|\cdot\|_F$ ) le produit scalaire et la norme euclidienne sur  $E$  (resp.  $F$ ). Soit  $f : E \rightarrow F$  une application linéaire.

(1) Les conditions suivantes sont équivalentes :

(a)  $f$  préserve la norme :  $\forall x \in E, \quad \|x\|_E = \|f(x)\|_F$

(b)  $f$  préserve le produit scalaire :  $\forall x, y \in E, \quad (x | y)_E = (f(x) | f(y))_F$

(c) Pour toute b.o.n.  $\mathcal{B} = (e_1, \dots, e_n)$  de  $E$ , la famille  $(f(e_1), \dots, f(e_n))$  est une b.o.n. de  $F$ .

(d) Il existe une b.o.n.  $\mathcal{B} = (e_1, \dots, e_n)$  de  $E$  telle que  $(f(e_1), \dots, f(e_n))$  soit une b.o.n. de  $F$ .

(2) Sous ces conditions, on dit que  $f$  est une **isométrie vectorielle** de  $E$  sur  $F$

(3) Dans ce cas,  $f$  est bijective, et son inverse  $f^{-1}$  est aussi une isométrie.

*Démonstration.* — Supposons que  $f$  préserve la norme, et soient  $x, y \in E$ . Alors  $\|x + y\|_E^2 = \|f(x + y)\|_F^2 = \|f(x) + f(y)\|_F^2$ , et le premier (resp. dernier) membre égale :

$$\|x\|_E^2 + \|y\|_E^2 + 2(x | y)_E, \quad \text{resp.} \quad \|f(x)\|_F^2 + \|f(y)\|_F^2 + 2(f(x) | f(y))_F$$

et comme  $\|x\|_E^2 = \|f(x)\|_F^2$  et  $\|y\|_E^2 = \|f(y)\|_F^2$ , on obtient que  $(x | y)_E = (f(x) | f(y))_F$ . Ceci prouve que (a)  $\Rightarrow$  (b).

Les implications (b)  $\Rightarrow$  (c)  $\Rightarrow$  (d) sont évidentes, montrons que (d)  $\Rightarrow$  (a). Supposons (d) vérifiée. Pour tout  $x = x_1 e_1 + \dots + x_n e_n$  dans  $E$ , on a  $f(x) = \sum_i x_i f(e_i)$  et, comme  $(e_1, \dots, e_n)$  et  $(f(e_1), \dots, f(e_n))$  sont des b.o.n., on obtient

$$\|x\|_E^2 = \sum_{i=1}^n x_i^2 = \|f(x)\|_F^2$$

donc (a) est vérifiée. Ceci prouve l'assertion (1).

Prouvons (3). Soit  $f : E \rightarrow F$  une isométrie, et soit  $\mathcal{B} = (e_1, \dots, e_n)$  de  $E$ . Comme  $f(\mathcal{B})$  est une b.o.n. (donc une base) de  $F$ , alors  $f$  est bijective. Son inverse  $f^{-1}$  envoie la b.o.n.  $f(\mathcal{B}) = (f(e_1), \dots, f(e_n))$  de  $F$  sur la b.o.n.  $\mathcal{B}$  de  $E$ , donc  $f^{-1}$  est une isométrie. Ceci prouve (3). La proposition est démontrée.  $\square$

**Terminologie 2.1.9.1.** — On a introduit la terminologie isométrie « vectorielle » pour pouvoir faire plus tard la distinction avec la notion d'isométrie « affine », qu'on introduira lorsqu'on étudiera les espaces et applications affines.

Dans la suite de ce chapitre, comme on ne considère que des applications linéaires, on dira simplement « isométrie » au lieu de « isométrie vectorielle ».

**Définition et corollaire 2.1.10.** — (1) On dit que deux espaces euclidiens  $E$  et  $E'$  sont **isométriques** s'il existe une isométrie  $f : E \xrightarrow{\sim} E'$ .

(2) Tout espace euclidien  $E$  de dimension  $n$  est isométrique à  $\mathbb{R}^n$  muni du produit scalaire euclidien standard.

*Démonstration.* — Soit  $\mathcal{B} = (e_1, \dots, e_n)$  la base canonique de  $\mathbb{R}^n$ , qui est orthonormée pour le produit scalaire standard. D'après le théorème 2.1.4,  $E$  admet une b.o.n.  $\mathcal{C} = (f_1, \dots, f_n)$ . Alors l'application linéaire  $u : \mathbb{R}^n \rightarrow E$  définie par  $u(e_i) = f_i$ , pour  $i = 1, \dots, n$ , est une isométrie de  $\mathbb{R}^n$  sur  $E$ .  $\square$

**Définition 2.1.11.** — On note  $O(n) = \{A \in M_n(\mathbb{R}) \mid {}^tAA = I_n\}$ . Rappelons (cf. ??) que l'égalité  ${}^tAA = I_n$  entraîne que  $A$  est inversible et  $A^{-1} = {}^tA$ . Donc  $O(n) \subset GL_n(\mathbb{R})$  et, si  $A \in O(n)$ , son inverse  $B = A^{-1} = {}^tA$  vérifie  $B^{-1} = A = {}^tB$ , donc appartient aussi à  $O(n)$ . De plus, pour tout  $A, B \in O(n)$ , on a l'égalité  ${}^t(AB)AB = {}^tB{}^tAAB = {}^tBB = I_n$ , donc  $AB \in O(n)$ . Donc  $O(n)$  est un *sous-groupe* de  $GL_n(\mathbb{R})$ , appelé le **groupe orthogonal**.

Munissons  $\mathbb{R}^n$  du produit scalaire euclidien standard  $(\mid)$ . Pour tout  $X, Y \in \mathbb{R}^n$  on a  $(X \mid Y) = {}^tXY$ , i.e. la matrice de  $(\mid)$  dans la base canonique  $\mathcal{B}_0 = (e_1, \dots, e_n)$  est la matrice identité  $I_n$ . Donc une matrice arbitraire  $A \in M_n(\mathbb{R})$  préserve le produit scalaire si et seulement si, on a, pour tout  $X, Y \in \mathbb{R}^n$  :

$${}^tXY = (X \mid Y) = (AX \mid AY) = {}^tX({}^tAA)Y$$

ce qui équivaut à dire que  ${}^tAA = I_n$  (cf. 1.1.8). Ceci montre que  $O(n)$  est le groupe des isométries de  $\mathbb{R}^n$  muni du produit scalaire euclidien standard  $(\mid)$ .

De plus, notons  $C_1, \dots, C_n$  les colonnes de  $A$  (i.e.  $C_i$  est le vecteur  $Ae_i \in \mathbb{R}^n$ ). Remarquons que, pour tout  $i, j \in \{1, \dots, n\}$ , le coefficient d'indice  $(i, j)$  de  ${}^tAA$  est le produit matriciel de la  $i$ -ème ligne de  ${}^tA$ , i.e. de  ${}^tC_i$ , par la colonne  $C_j$ , c.-à-d., on a  $({}^tAA)_{ij} = (Ae_i \mid Ae_j)$ , donc la condition  ${}^tAA = I_n$  équivaut aussi à dire que les colonnes de  $A$  sont de norme 1 et deux à deux orthogonales. Tenant compte de la proposition 2.1.9, on obtient donc les caractérisations suivantes de  $O(n)$ , chacune étant utile :



**Proposition 2.1.12 (Groupe orthogonal  $O(n)$ ).** — On munit  $\mathbb{R}^n$  du produit scalaire euclidien standard  $(\mid)$  et l'on note  $\|\cdot\|$  la norme euclidienne associée. Alors  $O(n)$  est le groupe des isométries de  $\mathbb{R}^n$  ; il est caractérisé par chacune des égalités suivantes :

$$\begin{aligned} O(n) &= \{A \in M_n(\mathbb{R}) \mid {}^tAA = I_n\} \\ &= \{A \in GL_n(\mathbb{R}) \mid A^{-1} = {}^tA\} \\ &= \{A \in M_n(\mathbb{R}) \mid (AX \mid AY) = (X \mid Y), \quad \forall X, Y \in \mathbb{R}^n\} \\ &= \{A \in M_n(\mathbb{R}) \mid \|AX\| = \|X\|, \quad \forall X \in \mathbb{R}^n\} \\ &= \{A \in M_n(\mathbb{R}) \mid (Af_1, \dots, Af_n) \text{ est une b.o.n., pour toute b.o.n. } (f_1, \dots, f_n)\} \\ &= \{A \in M_n(\mathbb{R}) \mid (Ae_1, \dots, Ae_n) \text{ est une b.o.n., où } (e_1, \dots, e_n) \text{ est la base canonique de } \mathbb{R}^n\} \\ &= \{A \in M_n(\mathbb{R}) \mid \text{les colonnes de } A \text{ sont de norme 1 et deux à deux orthogonales}\} \end{aligned}$$

Les éléments de  $O(n)$  sont parfois appelés « endomorphismes orthogonaux » (mais voir la remarque 2.3.2.2 plus bas).

**Remarque 2.1.13.** — Il existe d'autres groupes orthogonaux (qui ne sont isomorphes à aucun  $O(n)$ ). Soient  $p, q$  des entiers  $\geq 1$  et soit  $\phi$  la forme bilinéaire symétrique sur  $\mathbb{R}^{p+q}$  définie par  $\phi(X, Y) = \sum_{i=1}^p x_i y_i - \sum_{i=p+1}^q x_i y_i$ , i.e. la matrice de  $\phi$  dans la base canonique de  $\mathbb{R}^{p+q}$  est  $J = \left( \begin{array}{c|c} I_p & \mathbf{0}_{p,q} \\ \hline \mathbf{0}_{q,p} & -I_q \end{array} \right)$ . Alors

$$\{A \in M_n(\mathbb{R}) \mid {}^tAJA = J\} = \{A \in M_n(\mathbb{R}) \mid \phi(AX, AY) = \phi(X, Y), \quad \forall X, Y \in \mathbb{R}^n\}$$

est un sous-groupe de  $GL_n(\mathbb{R})$ , noté  $O(p, q)$ . On ne considérera pas ces groupes dans ce cours.

## 2.2. Endomorphismes auto-adjoints et théorème de diagonalisation simultanée

Commençons par introduire l'adjoint dans le cas général d'une forme bilinéaire symétrique non dégénérée, même si on se limitera dans la suite au cas euclidien.

**Théorème et définition 2.2.1 (Adjoint d'un endomorphisme).** — Soient  $E$  un  $\mathbb{R}$ -espace vectoriel de dimension  $n$ ,  $\phi$  une forme bilinéaire symétrique sur  $E$ , **non dégénérée**. Pour tout  $u \in \text{End}(E)$ , il existe un unique endomorphisme  $u^*$  de  $E$ , appelé **l'adjoint** de  $u$ , vérifiant :

$$(1) \quad \forall x, y \in E, \quad \boxed{\phi(u(x), y) = \phi(x, u^*(y))}$$

Pour toute base  $\mathcal{B}$  de  $E$ , si l'on note  $J = \text{Mat}_{\mathcal{B}}(\phi)$  et  $A = \text{Mat}_{\mathcal{B}}(u)$ , on a

$$(2) \quad \boxed{A^* = \text{Mat}_{\mathcal{B}}(u^*) = J^{-1} {}^tA J}$$

*Démonstration.* — Supposons qu'il existe  $u^*$  vérifiant (1) et soient  $\mathcal{B}$  une base de  $E$ ,  $J = \text{Mat}_{\mathcal{B}}(\phi)$ ,  $A = \text{Mat}_{\mathcal{B}}(u)$  et  $A^* = \text{Mat}_{\mathcal{B}}(u^*)$ . Soient  $x, y \in E$  arbitraires, et notons  $X, Y \in \mathbb{R}^n$  les vecteurs colonnes des coordonnées dans la base  $\mathcal{B}$ . Alors on a

$${}^tX {}^tA J Y = \phi(u(x), y) = \phi(x, u^*(y)) = {}^tX J A^* Y$$

d'où  ${}^tAJ = JA^*$  et donc, puisque  $J$  est inversible (car  $\phi$  non-dégénérée),  $A^* = J^{-1}{}^tAJ$ . Ceci montre que  $u^*$ , s'il existe, vérifie (2) et est donc unique.

Réciproquement, si l'on note  $u^*$  l'endomorphisme de  $E$  dont la matrice dans la base  $\mathcal{B}$  est  $A^* = J^{-1}{}^tAJ$ , alors pour tout  $x, y$  on a :

$$\phi(x, u^*(y)) = {}^tXJA^*Y = {}^tX{}^tAJY = \phi(u(x), y)$$

donc  $u^*$  vérifie (1). Ceci prouve l'existence, et le théorème est démontré.  $\square$

**Remarque 2.2.2.** — Il résulte de la formule (2) (ou directement de la définition (1)) que, pour tout  $u, v \in \text{End}(E)$  et  $s, t \in \mathbb{R}$ , on a  $(su + tv)^* = su^* + tv^*$ , i.e. l'application  $\text{End}(E) \rightarrow \text{End}(E)$ ,  $u \mapsto u^*$  est linéaire.

Remarquons aussi que si  $\phi$  est un produit scalaire et si  $\mathcal{B}$  est une b.o.n., alors la matrice de  $\phi$  dans  $\mathcal{B}$  est  $J = I_n$ . On peut donc énoncer le théorème dans le cas euclidien sous la forme suivante.

### **Théorème 2.2.3 (Adjoint d'un endomorphisme dans le cas euclidien)**

Soit  $E$  muni de  $(\mid)$  un espace euclidien de dimension  $n$ . Pour tout  $u \in \text{End}(E)$ , il existe un unique endomorphisme  $u^*$  de  $E$ , appelé l'adjoint de  $u$ , vérifiant :

$$(*) \quad \forall x, y \in E, \quad \boxed{(u(x) \mid y) = (x \mid u^*(y))}.$$

Pour toute b.o.n.  $\mathcal{B}$  de  $E$ , si l'on note  $A = \text{Mat}_{\mathcal{B}}(u)$ , on a

$$(**) \quad \boxed{A^* = \text{Mat}_{\mathcal{B}}(u^*) = {}^tA}.$$

**Définition 2.2.4 (Endomorphismes auto-adjoints).** — Soit  $E$  un espace euclidien de dimension  $n$ . On dit qu'un endomorphisme  $u \in \text{End}(E)$  est **auto-adjoint** (ou symétrique) s'il vérifie  $u^* = u$ . Ceci équivaut à dire que, pour toute b.o.n.  $\mathcal{B}$  de  $E$ , la matrice  $S = \text{Mat}_{\mathcal{B}}(u)$  est **symétrique**.

### **Proposition 2.2.5 (Endomorphismes auto-adjoints et formes bilinéaires symétriques)**

Soit  $E$  muni de  $(\mid)$  un espace euclidien de dimension  $n$  et soit  $\phi$  une autre forme bilinéaire symétrique (arbitraire) sur  $E$ . Alors il existe un unique  $u \in \text{End}(E)$  auto-adjoint pour  $(\mid)$  tel que :

$$(\dagger) \quad \forall x, y \in E, \quad \boxed{\phi(x, y) = (u(x) \mid y) = (x \mid u(y))}.$$

Pour toute b.o.n.  $\mathcal{B}$  de  $E$ , on a

$$(\ddagger) \quad \boxed{\text{Mat}_{\mathcal{B}}(u) = \text{Mat}_{\mathcal{B}}(\phi)}.$$

*Démonstration.* — Soient  $\mathcal{B}$  une b.o.n. de  $E$  et  $S = \text{Mat}_{\mathcal{B}}(\phi)$ , on a  ${}^tS = S$ . Pour  $x, y \in E$ , notons  $X, Y \in \mathbb{R}^n$  les coordonnées dans la base  $\mathcal{B}$ . S'il existe  $u$  vérifiant  $(\dagger)$ , soit  $A = \text{Mat}_{\mathcal{B}}(u)$ , alors l'égalité

$${}^tXSY = \phi(x, y) = (x \mid u(y)) = {}^tXAY$$

entraîne  $A = S$ . Ceci montre que  $u$ , s'il existe, vérifie  $(\ddagger)$  et est donc unique.

Réciproquement, si l'on note  $u$  l'endomorphisme de  $E$  dont la matrice dans la base  $\mathcal{B}$  est  $S$ , alors pour tout  $x, y$  on a :

$$\begin{aligned} \phi(x, y) &= {}^tXSY = (x \mid u(y)) \\ &= {}^tX{}^tSY = (u(x) \mid y) \end{aligned}$$

donc  $u$  vérifie  $(\dagger)$ . Ceci prouve l'existence, et la proposition est démontrée.  $\square$

On a maintenant le théorème important et utile suivant.

### **Théorème 2.2.6 (Diagonalisation des endomorphismes auto-adjoints)**

Soient  $E$  muni de  $(\mid)$  un espace euclidien de dimension  $n$ , et  $u$  un endomorphisme auto-adjoint. Alors,  $u$  est diagonalisable et ses espaces propres sont deux à deux orthogonaux. Par conséquent, il existe une b.o.n. de  $E$  formée de vecteurs propres de  $u$ .

### **Corollaire 2.2.7 (Diagonalisation des matrices symétriques réelles)**

Soit  $S \in M_n(\mathbb{R})$  une matrice symétrique réelle. Alors  $S$  est diagonalisable dans une base orthonormée : il existe  $P \in O(n)$  telle que  $P^{-1}SP$  soit diagonale.

Le point le plus difficile de la démonstration est la proposition suivante :

**Proposition 2.2.8 (Existence d'une valeur propre réelle).** — Soit  $A \in M_n(\mathbb{R})$  symétrique. Alors  $A$  admet au moins une valeur propre réelle.

Admettons pour le moment cette proposition et démontrons le théorème, par récurrence sur  $n = \dim E$ . C'est ok si  $n = 1$ , donc on peut supposer  $n \geq 2$  et le résultat établi pour  $n - 1$ . D'après la proposition,  $u$  admet au moins une valeur propre réelle  $\lambda_1$ , soit  $f_1$  un vecteur propre associé, qu'on peut supposer de norme 1 (quitte à remplacer  $f_1$  par  $\frac{1}{\|f_1\|}f_1$ ). Montrons que  $E_1 = (\mathbb{R}f_1)^\perp$  est stable par  $u$  : pour tout  $x \in E_1$ , on a :

$$(u(x) | f_1) = (x | u^*(f_1)) = (x | u(f_1)) = (x | \lambda_1 f_1) = \lambda_1(x | f_1) = 0,$$

donc  $u(x) \in E_1$ . La restriction  $u_1$  de  $u$  à  $E_1$  est encore auto-adjointe, puisque pour tout  $x, y \in E_1$  on a :

$$(u_1(x) | y) = (u(x) | y) = (x | u(y)) = (x | u_1(y)).$$

Donc, par hypothèse de récurrence, il existe une b.o.n.  $\mathcal{C} = (f_2, \dots, f_n)$  de  $E_1$  formée de vecteurs propres de  $u_1$ , donc de  $u$ . Alors,  $\mathcal{B} = \{f_1\} \cup \mathcal{C}$  est une b.o.n. de  $E$  formée de vecteurs propres de  $u$ . Ceci prouve la première assertion du théorème.

Le fait que les espaces propres soient deux à deux orthogonaux peut se déduire de la démonstration précédente, mais il est plus simple de le voir directement. Soient  $\lambda \neq \mu$  deux valeurs propres distinctes de  $u$  et soient  $x \in V_\lambda$  et  $y \in V_\mu$  ; alors

$$\lambda(x | y) = (u(x) | y) = (x | u(y)) = \mu(x | y)$$

et comme  $\lambda \neq \mu$  ceci entraîne  $(x | y) = 0$ . Ceci prouve le théorème, modulo la démonstration de la proposition 2.2.8.  $\square$

*Démonstration de la proposition 2.2.8.* — On munit  $\mathbb{R}^n$  de la norme euclidienne usuelle et l'on considère la **sphère unité** :

$$S^{n-1} = \{x \in \mathbb{R}^n \mid \|x\|^2 = x_1^2 + \dots + x_n^2 = 1\};$$

celle-ci est **compacte**. D'autre part, la fonction

$$f : \mathbb{R}^n \rightarrow \mathbb{R}, \quad x \mapsto (Ax | x)$$

est **continue**, car c'est un polynôme de degré 2 en les coordonnées  $x_1, \dots, x_n$ . Par conséquent,  $f$  atteint un maximum  $\lambda$  en un point  $x_0$  de  $S^{n-1}$ , i.e. on a :

$$\forall x \in S^{n-1}, \quad (Ax | x) \leq \lambda = (Ax_0 | x_0).$$

Alors, pour tout  $x \neq 0$  dans  $\mathbb{R}^n$ , on a  $\frac{1}{\|x\|}x \in S^{n-1}$ , d'où

$$\left( \frac{Ax}{\|x\|} \mid \frac{x}{\|x\|} \right) \leq \lambda$$

et donc :

$$(1) \quad \forall x \in \mathbb{R}^n - \{0\}, \quad (Ax | x) \leq \lambda(x | x).$$

Fixons  $v \in \mathbb{R}^n$  et soit  $t \in \mathbb{R}$  variable. On a, d'une part :

$$f(x_0 + tv) = (A(x_0 + tv) | A(x_0 + tv)) = (Ax_0 | x_0) + t(Ax_0 | v) + t(Av | x_0) + t^2(Av | v)$$

et comme  $(Av | x_0) = (v | {}^tAx_0) = (v | Ax_0) = (Ax_0 | v)$ , ceci se récrit :

$$(2) \quad f(x_0 + tv) = (Ax_0 | x_0) + 2t(Ax_0 | v) + t^2(Av | v).$$

D'autre part, on a :

$$\lambda(x_0 + tv | x_0 + tv) = \lambda \underbrace{(x_0 | x_0)}_{=1} + 2t(\lambda x_0 | v) + t^2(\lambda v | v).$$

D'après (1), et tenant compte de l'égalité  $\lambda = (Ax_0 | x_0)$ , on obtient :

$$\forall t \in \mathbb{R}, \quad 2t(Ax_0 - \lambda x_0 | v) + t^2(Av - \lambda v | v) \leq 0.$$

On a donc un trinôme du second degré en  $t$ , toujours négatif et qui s'annule pour  $t = 0$ . On en déduit que son discriminant réduit  $\Delta' = (Ax_0 - \lambda x_0 | v)^2$  est nul, donc :

$$\forall v \in \mathbb{R}^n, \quad (Ax_0 - \lambda x_0 | v) = 0$$

et donc  $Ax_0 - \lambda x_0 = 0$ , i.e.  $Ax_0 = \lambda x_0$ . Ceci prouve que  $x_0$  est un vecteur propre pour  $\lambda$ . Ceci achève la démonstration de la proposition 2.2.8 et du théorème 2.2.6.  $\square$



**Théorème 2.2.9 (Réduction simultanée).** — Soient  $E$  muni de  $(\mid)$  un espace euclidien de dimension  $n$ ,  $Q$  une forme quadratique arbitraire sur  $E$ ,  $\phi$  sa forme polaire,  $\mathcal{B}_0 = (e_1, \dots, e_n)$  une base orthonormée de  $E$ , et  $u$  l'endomorphisme de  $E$  tel que  $\text{Mat}_{\mathcal{B}_0}(u) = \text{Mat}_{\mathcal{B}_0}(\phi) = A$ .

Alors il existe une base  $\mathcal{B} = (f_1, \dots, f_n)$  **orthonormée** pour  $(\mid)$  et formée de vecteurs propres de  $u$ , i.e.  $u(f_i) = \lambda_i f_i$  pour  $i = 1, \dots, n$ , et l'on a

$$\text{Mat}_{\mathcal{B}}(\phi) = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{pmatrix}$$

où  $\lambda_1, \dots, \lambda_n$  sont les valeurs propres de  $u$ ; plus précisément, la matrice de passage  $P = \text{Mat}_{\mathcal{B}_0}(\mathcal{B})$  est orthogonale, i.e.  ${}^tP = P^{-1}$ , donc la matrice ci-dessus égale à la fois  ${}^tPAP = \text{Mat}_{\mathcal{B}}(\phi)$  et  $P^{-1}AP = \text{Mat}_{\mathcal{B}}(u)$ .

**Remarque 2.2.9.1.** — Ce théorème est appelé « théorème de **réduction simultanée** » ou « de **diagonalisation simultanée** » car la base  $\mathcal{B}$  donnée par l'énoncé est à la fois **orthonormée** pour  $(\mid)$  et **orthogonale** pour  $\phi$ . En d'autres termes, si l'on note  $(x_1, \dots, x_n)$  les coordonnées dans  $\mathcal{B}$  d'un vecteur  $x$  arbitraire, la base  $\mathcal{B}$  **réduit simultanément** la forme  $x \mapsto (x \mid x)$  à la forme standard  $x_1^2 + \cdots + x_n^2$ , et la forme  $Q$  en la somme de carrés  $\lambda_1 x_1^2 + \cdots + \lambda_n x_n^2$ .

*Démonstration.* — Notons  $u$  l'endomorphisme auto-adjoint tel que

$$\forall x, y \in E, \quad \phi(x, y) = (u(x) \mid y) = (x \mid (u(y))),$$

cf. Proposition 2.2.5. D'après le théorème 2.2.6, il existe une base  $\mathcal{B} = (f_1, \dots, f_n)$  **orthonormée** pour  $(\mid)$  et formée de vecteurs propres de  $u$ , i.e.  $u(f_i) = \lambda_i f_i$ , pour tout  $i$ . Alors, pour tout  $i, j$  on a :

$$\phi(f_i, f_j) = \lambda_i (f_i \mid f_j) = \lambda_j (f_i \mid f_j) = \begin{cases} 0 & \text{si } i \neq j, \\ \lambda_i & \text{si } i = j, \end{cases}$$

ce qui montre que  $\mathcal{B}$  est une base **orthogonale** pour  $\phi$ . De plus, comme  $\mathcal{B}_0$  et  $\mathcal{B}$  sont orthonormées, la matrice de passage  $P = \text{Mat}_{\mathcal{B}_0}(\mathcal{B})$  est orthogonale, i.e.  ${}^tP = P^{-1}$ , donc la matrice diagonale de l'énoncé égale à la fois  ${}^tPAP = \text{Mat}_{\mathcal{B}}(\phi)$  et  $P^{-1}AP = \text{Mat}_{\mathcal{B}}(u)$ .  $\square$

Répetons la version matricielle du théorème précédent :

**Corollaire 2.2.10 (Réduction simultanée des matrices symétriques réelles)**

Soit  $S \in M_n(\mathbb{R})$  telle que  ${}^tS = S$ . Il existe  $P \in O(n)$  telle que  $P^{-1}SP = {}^tPSP$  soit diagonale.



**Corollaire 2.2.11 (Calculs de signature).** — Soient  $Q$  une forme quadratique sur  $\mathbb{R}^n$ ,  $\phi$  sa forme polaire,  $A$  la matrice de  $\phi$  dans la base canonique. Alors la signature de  $Q$  est donnée par le nombre de valeurs propres de  $A$  qui sont  $> 0$  (resp.  $< 0$ ).

**Exemple 2.2.12.** — Illustrons ce qui précède par l'exemple suivant. Soient  $n \geq 2$  et  $Q$  la forme quadratique sur  $\mathbb{R}^n$  définie par

$$Q(x_1, \dots, x_n) = 2 \sum_{i < j} x_i x_j = \sum_{i \neq j} x_i x_j.$$

La matrice de sa forme polaire est

$$A = \begin{pmatrix} 0 & 1 & \cdots & 1 \\ 1 & 0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 1 \\ 1 & \cdots & 1 & 0 \end{pmatrix}$$

(tous les coefficients valent 1 sauf ceux de la diagonale qui sont nuls). On remarque que la matrice  $A + I_n$  est de rang 1, donc l'espace propre  $V_{-1} = \text{Ker}(A + I_n)$  est de dimension  $n - 1$ . Donc  $-1$  est une racine de multiplicité  $\geq n - 1$  du polynôme caractéristique  $P_A(X)$ . Comme  $0 = \text{Tr}(A)$  est la somme des racines

(dans  $\mathbb{C}$ ) de  $P_A(X)$ , la dernière racine  $\lambda$  vérifie  $\lambda + (n-1)(-1) = 0$ , d'où  $\lambda = n-1$ . Donc, d'après le théorème 2.2.9, il existe  $P \in O(n)$  tel que

$$P^{-1}AP = {}^tPAP = \begin{pmatrix} -1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & -1 & 0 \\ 0 & \cdots & 0 & n-1 \end{pmatrix}$$

il y a donc  $n-1$  valeurs propres égales à  $-1$ , et une seule valeur propre  $> 0$  (égale à  $n-1$ ), donc la signature de  $Q$  est  $(1, n-1)$ .

**Remarque 2.2.13.** — Pour des applications géométriques du théorème 2.2.9, voir l'étude des coniques et quadriques dans le chapitre suivant.

### 2.3. Orthogonalité. Orthonormalisation de Gram-Schmidt



**Définition 2.3.1 (Sous-espaces d'un espace euclidien).** — Soit  $E$ , muni de  $(\mid)$ , un espace euclidien et soit  $F$  un sous-espace vectoriel de  $E$ . Alors la restriction  $(\mid)_F$  de  $(\mid)$  à  $F$  (cf. 1.1.20) est un produit scalaire sur  $F$ , puisque  $(x \mid x)_F = (x \mid x) > 0$  pour tout  $x \in F - \{0\}$ . Donc  $F$  muni de  $(\mid)_F$  est un espace euclidien.



#### **Théorème et définition 2.3.2 (Projection orthogonale sur un sous-espace)**

Soit  $E$ , muni de  $(\mid)$ , un espace euclidien (pas nécessairement de dimension finie). Soit  $F$  un sous-espace de dimension finie  $r$  et soit  $F^\perp$  son orthogonal pour  $(\mid)$ .

(1) On a  $E = F \oplus F^\perp$ . Cette décomposition en somme directe permet de définir le projecteur  $\pi_F : E \rightarrow E$ , d'image  $F$  et de noyau  $F^\perp$ , et le projecteur  $\pi_{F^\perp} : E \rightarrow E$ , d'image  $F^\perp$  et de noyau  $F$ . Alors, pour tout  $x \in E$ , on a

$$x = \pi_F(x) + \pi_{F^\perp}(x)$$

et  $\pi_F$  (resp.  $\pi_{F^\perp}$ ) s'appelle la **projection orthogonale** sur  $F$  (resp. sur  $F^\perp$ ).

(2) Soit  $(e_1, \dots, e_r)$  une base orthonormée de  $F$ . Alors  $\pi_F(v) = (v \mid e_1)e_1 + \cdots + (v \mid e_r)e_r$  pour tout  $v \in E$ .

(3) On a  $(F^\perp)^\perp = F$ .

**Remarques 2.3.2.1.** — (a) On a toujours  $F \cap F^\perp = \{0\}$  car si  $x \in F \cap F^\perp$  alors  $(x \mid x) = 0$  d'où  $x = 0$ .

(b) Par conséquent, si  $E$  est de dimension finie  $n$ , l'égalité  $E = F \oplus F^\perp$  découle de 1.1.19. Toutefois, la démonstration donnée ci-dessous ne suppose pas  $E$  de dimension finie et permet, par exemple, de démontrer l'inégalité de Bessel pour les coefficients de Fourier d'une fonction continue  $f : [0, 2\pi] \rightarrow \mathbb{R}$  (voir par exemple le Devoir du 6/4/2012 et son corrigé).

*Démonstration.* — On va démontrer en même temps les assertions (1) et (2). Comme  $(\mid)_F$  est définie positive et comme  $\dim(F) = r < \infty$ , alors  $F$  possède une base orthonormée  $(e_1, \dots, e_r)$ , d'après le théorème 2.1.4. Pour tout  $v \in E$ , posons provisoirement

$$p(v) = (v \mid e_1)e_1 + \cdots + (v \mid e_r)e_r \in F$$

et  $q(v) = v - p(v)$ . Alors,  $v = p(v) + q(v)$ . D'autre part, pour  $j = 1, \dots, r$ , on a

$$(q(v) \mid e_j) = (v \mid e_j) - \sum_{i=1}^r (v \mid e_i) \underbrace{(e_i \mid e_j)}_{\substack{=1 \text{ si } i=j \\ =0 \text{ si } i \neq j}} = 0,$$

d'où  $q(v) \in F^\perp$ . Comme  $v = p(v) + q(v)$ , ceci montre que  $E = F + F^\perp$ .

De plus, comme on l'a déjà remarqué plus haut, si  $x \in F \cap F^\perp$ , alors  $(x \mid x) = 0$ , d'où  $x = 0$ . On a donc  $F \cap F^\perp = \{0\}$  et  $E = F + F^\perp$ , d'où  $E = F \oplus F^\perp$ . Ceci prouve (1).

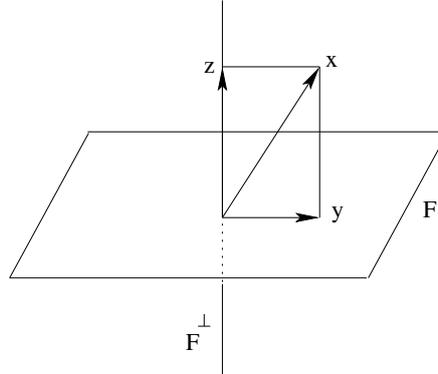
Alors tout  $v \in E$  s'écrit de façon unique  $v = x + y$  avec  $x \in F$  et  $y \in F^\perp$ , et l'on a  $x = \pi_F(v)$  et  $y = \pi_{F^\perp}(v)$ . Comme  $v = p(v) + q(v)$ , on a donc

$$\pi_F(v) = p(v) = \sum_{i=1}^r (v \mid e_i)e_i,$$

ce qui prouve (2).

Prouvons (3). Pour tout  $x \in F$  et  $y \in F^\perp$ , on a  $0 = (x | y) = (y | x)$ . Fixant  $x \in F$  et faisant varier  $y$  dans  $G = F^\perp$ , ceci montre que  $x \in G^\perp$ , d'où l'inclusion  $F \subset G^\perp$ . Réciproquement, soit  $v \in G^\perp$ . Comme  $E = F \oplus F^\perp$ , on peut écrire  $v = x + y$  avec  $y \in G = F^\perp$  et  $x \in F$ . Alors  $y = v - x \in G \cap G^\perp$ , donc  $0 = (y | y)$  d'où  $y = 0$  et donc  $v = x \in F$ . Ceci montre que  $F = G^\perp = (F^\perp)^\perp$ . Le théorème est démontré.  $\square$

Dans la figure qui suit, on a  $y = \pi_F(x)$  et  $z = \pi_{F^\perp}(x)$  :



**Remarque 2.3.2.2.** — Attention à la terminologie! Si  $F \neq E$ , la projection orthogonale  $\pi_F$  n'est **pas une isométrie** (car une isométrie est injective, or  $\text{Ker}(\pi_F) = F^\perp$  est non nul, sauf si  $F = E$ ), donc n'est **pas** un « endomorphisme orthogonal » de  $E$  (cf. 2.1.12).



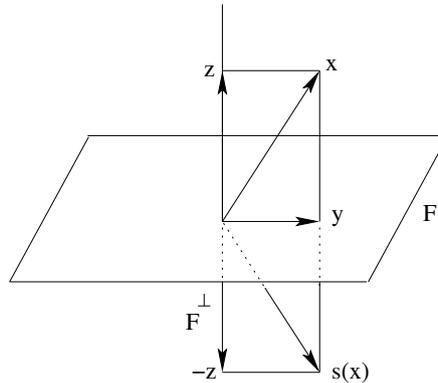
**Définition et proposition 2.3.3 (Symétries orthogonales).** — Soient  $E$  un espace euclidien de dimension  $n$  et  $F$  un sous-espace de dimension  $r$ .

(1) La **symétrie orthogonale**  $s_F$  par rapport à  $F$  est définie comme suit : pour tout  $v \in E$ , on a  $v = \pi_F(v) + \pi_{F^\perp}(v)$  et l'on pose :

$$(*) \quad s_F(v) = \pi_F(v) - \pi_{F^\perp}(v) = v - 2\pi_{F^\perp}(v).$$

Alors  $s_F^2 = \text{id}_E$  et  $s_F$  est une isométrie de  $E$ .

(2) Si  $\mathcal{C}_+$  est une base de  $F$  et  $\mathcal{C}_-$  une base de  $F^\perp$ , la matrice de  $s_F$  dans la base  $\mathcal{B} = \mathcal{C}_+ \cup \mathcal{C}_-$  de  $E$  est  $\text{Mat}_{\mathcal{B}}(s_F) = \left( \begin{array}{c|c} I_r & \mathbf{0}_{r, n-r} \\ \hline \mathbf{0}_{n-r, r} & -I_{n-r} \end{array} \right)$ . En particulier, on a  $\det s_F = (-1)^{n-r}$ .



**Démonstration.** — D'après la définition, il est clair que  $s_F^2 = \text{id}_E$ , donc  $s_F$  est bijective et égale à son inverse (i.e.  $s_F$  est **involutive**). Montrons que  $s_F$  est une isométrie. Comme  $(\pi_F(v) | \pi_{F^\perp}(v)) = 0$ , on a d'après l'égalité de Pythagore (cf. 2.1.7) :

$$\|v\|^2 = \|\pi_F(v)\|^2 + \|\pi_{F^\perp}(v)\|^2 = \|s_F(v)\|^2$$

et ceci prouve que  $s_F$  est une isométrie. Enfin, si  $\mathcal{B} = \mathcal{C}_+ \cup \mathcal{C}_-$  est comme dans la proposition, il est clair que  $\text{Mat}_{\mathcal{B}}(s_F)$  est comme indiquée.  $\square$



**Définition 2.3.4 (Réflexions orthogonales).** — Un cas particulier important de symétrie orthogonale est le suivant. Soit  $v_0 \in E$ ,  $v_0 \neq 0$ , alors  $H = (\mathbb{R}v_0)^\perp$  est appelé un **hyperplan** de  $E$ ; d'après le théorème 2.3.2 on a

$$E = \mathbb{R}v_0 \oplus H,$$

explicitement, si l'on pose  $u_0 = \frac{1}{\|v_0\|} v_0$  alors  $\|u_0\| = 1$  et tout  $x \in E$  s'écrit de façon unique

$$x = (x | u_0)u_0 + \pi_H(x), \quad \text{où} \quad \pi_H(x) = x - (x | u_0)u_0,$$

donc

$$\pi_{\mathbb{R}v_0}(x) = (x | u_0)u_0 = \frac{(x | v_0)}{(v_0 | v_0)} v_0.$$

La symétrie orthogonale  $s_H$  par rapport à  $H$  s'appelle la **réflexion orthogonale** par rapport à l'hyperplan  $H$ ; d'après ce qui précède elle est donnée par la formule :

$$(2.3.4.1) \quad \forall x \in E, \quad s_H(x) = x - 2 \frac{(x | v_0)}{(v_0 | v_0)} v_0.$$

Si  $\dim E = n$  alors  $\dim H = n - 1$ , et si  $\mathcal{C}_+$  est une base de  $H$ , alors  $\mathcal{B} = \mathcal{C}_+ \cup \{v_0\}$  est une base de  $E$  et  $\text{Mat}_{\mathcal{B}}(s_H) = \left( \begin{array}{c|c} I_{n-1} & \mathbf{0}_{n-1,1} \\ \hline \mathbf{0}_{1,n-1} & -1 \end{array} \right)$ , d'où en particulier  $\boxed{\det s_H = -1}$ .



**Théorème 2.3.5 (Orthonormalisation de Gram-Schmidt).** — Soit  $E$  un espace euclidien et soient  $v_1, \dots, v_n$  linéairement indépendants dans  $E$ . Alors il existe une unique famille  $(e_1, \dots, e_n)$  vérifiant les deux propriétés suivantes :

- (1) Pour tout  $i = 1, \dots, n$ ,  $(e_1, \dots, e_i)$  est une base orthonormée de  $V_i = \text{Vect}(v_1, \dots, v_i)$ .
- (2) Pour tout  $j = 1, \dots, n$ , on a  $(e_j | v_j) > 0$ .

*Démonstration.* — Pour  $j = 1$ , on cherche  $e_1 = t_1 v_1$  tel que  $1 = (e_1 | e_1) = t_1^2 (v_1 | v_1)$  et  $0 < (e_1 | v_1) = t_1 (v_1 | v_1)$ ; la 1ère condition donne  $t_1^2 = 1/(v_1 | v_1)$ , et la 2ème condition, qui implique  $t_1 > 0$ , donne alors :

$$(1) \quad t_1 = \frac{1}{\|v_1\|} \quad \text{d'où} \quad \boxed{e_1 = \frac{1}{\|v_1\|} v_1.}$$

Pour  $j = 2$ , on cherche d'abord un vecteur  $e'_2 \in \text{Vect}(v_1, v_2) = \text{Vect}(e_1, v_2)$ , donc de la forme  $e'_2 = v_2 + \lambda e_1$ , vérifiant la condition :

$$0 = (e'_2 | e_1) = (v_2 | e_1) + \lambda \underbrace{(e_1 | e_1)}_{=1} = (v_2 | e_1) + \lambda,$$

ce qui impose  $\lambda = -(v_2 | e_1)$ . Alors le vecteur

$$\boxed{e'_2 = v_2 - (v_2 | e_1)e_1}$$

est orthogonal à  $e_1$ , et est  $\neq 0$  puisque  $v_2 \notin \mathbb{R}v_1 = \mathbb{R}e_1$ , donc la famille  $(e_1, e'_2)$  est libre et forme une base de  $V_2 = \text{Vect}(v_1, v_2)$ .

On a  $v_2 = e'_2 + (v_2 | e_1)e_1$  et, puisque  $(e'_2 | e_1) = 0$ , l'égalité de Pythagore donne

$$\|v_2\|^2 = \|e'_2\|^2 + (v_2 | e_1)^2 \quad \text{d'où} \quad \boxed{\|e'_2\|^2 = \|v_2\|^2 - (v_2 | e_1)^2.}$$

Pour rendre  $e'_2$  unitaire (i.e. de norme 1), on le divise par sa norme, c.-à-d., on pose

$$(2) \quad \boxed{e_2 = \frac{1}{\|e'_2\|} e'_2 = \frac{1}{\|e'_2\|} (v_2 - (v_2 | e_1)e_1)}$$

alors  $(e_1, e_2)$  est une b.o.n. de  $V_2$ , et d'après (2) ci-dessus on a  $1 = (e_2 | e_2) = (e_2 | v_2)/\|e'_2\|$  donc  $(e_2 | v_2) = \|e'_2\| > 0$ . C'est bien le seul choix possible, car si  $f_2 \in V_2$  est orthogonal à  $e_1$  et unitaire, alors  $f_2 = \pm e_2$ , et la condition  $(f_2 | v_2) > 0$  entraîne  $f_2 = e_2$ .

Pour  $j = 3$ , on cherche d'abord un vecteur  $e'_3 \in V_3 = \text{Vect}(e_1, e_2, v_3)$ , donc de la forme  $e'_3 = v_3 + \mu_2 e_2 + \mu_1 e_1$ , vérifiant les relations linéaires :

$$\begin{cases} 0 = (e'_3 | e_1) = (v_3 | e_1) + \mu_1, \\ 0 = (e'_3 | e_2) = (v_3 | e_2) + \mu_2 \end{cases}$$

(on a utilisé le fait que  $e_1, e_2$  sont orthogonaux et unitaires), qui donnent  $\mu_i = -(v_3 | e_i)$  pour  $i = 1, 2$ . Alors le vecteur

$$\boxed{e'_3 = v_3 - (v_3 | e_2)e_2 - (v_3 | e_1)e_1}$$

est orthogonal à  $V_2 = \text{Vect}(e_1, e_2) = \text{Vect}(v_1, v_2)$ , et est  $\neq 0$  puisque  $v_3 \notin V_2$ , donc la famille  $(e_1, e_2, e'_3)$  est libre et forme une base de  $V_3$ . Comme  $e_1, e_2$  et  $e'_3$  sont orthogonaux, l'égalité de Pythagore donne

$$\|v_3\|^2 = \|e'_3\|^2 + (v_3 | e_2)^2 + (v_3 | e_1)^2 \quad \text{d'où} \quad \boxed{\|e'_3\|^2 = \|v_3\|^2 - (v_3 | e_2)^2 - (v_3 | e_1)^2.}$$

Pour rendre  $e'_3$  unitaire, on le divise par sa norme, c.-à-d., on pose

$$(3) \quad \boxed{e_3 = \frac{1}{\|e'_3\|} e'_3 = \frac{1}{\|e'_3\|} (v_3 - (v_3 | e_2)e_2 - (v_3 | e_1)e_1)}$$

alors  $(e_1, e_2, e_3)$  est une b.o.n. de  $V_3$ , et d'après (3) ci-dessus on a  $1 = (e_3 | e_3) = (e_3 | v_3)/\|e'_3\|$  donc  $(e_3 | v_3) = \|e'_3\| > 0$ . C'est bien le seul choix possible, car si  $f_3 \in V_3$  est orthogonal à  $V_2$  et unitaire, alors  $f_3 = \pm e_3$ , et la condition  $(f_3 | v_3) > 0$  entraîne  $f_3 = e_3$ .

En répétant ce processus on construit par récurrence, de façon unique, la famille  $(e_1, \dots, e_n)$ ; les formules explicites pour  $e'_n$  et  $e_n$  étant :

$$\boxed{e'_n = v_n - \sum_{i=1}^{n-1} (v_n | e_i) e_i} \quad \boxed{\|e'_n\|^2 = \|v_n\|^2 - \sum_{i=1}^{n-1} (v_n | e_i)^2} \quad \text{et} \quad \boxed{e_n = \frac{1}{\|e'_n\|} e'_n}$$

□

**Remarques 2.3.5.1.** — (1) Ce qui précède peut aussi s'exprimer, de façon abstraite, comme suit : l'orthogonal  $G_n$  de  $V_{n-1}$  dans  $V_n$ , i.e.  $G_n = \{x \in V_n \mid (x | e_i) = 0 \text{ pour } i = 1, \dots, n-1\}$ , est de dimension  $n - (n-1) = 1$ , et  $\sum_{i=1}^{n-1} (v_n | e_i) e_i$  est la projection orthogonale de  $v_n$  sur  $V_{n-1}$  tandis que  $e'_n$  est la projection orthogonale de  $v_n$  sur  $G_n$ ; la droite  $G_n = \mathbb{R}e'_n$  contient deux vecteurs de norme 1, à savoir  $\pm e_n$ , et  $e_n$  est déterminé par la condition  $(e_n | v_n) = \|e'_n\| > 0$ .

(2) La démonstration précédente fournit un algorithme pour calculer explicitement  $e_1, \dots, e_n$ . Illustrons ceci par l'exemple suivant.

**Exemple 2.3.5.2.** — On munit  $E = \mathbb{R}[X]$  du produit scalaire défini par  $(P | Q) = \int_{-1}^1 P(t)Q(t)dt$ . Appliquons le procédé d'orthonormalisation de Gram-Schmidt aux vecteurs  $1, X, X^2 \in \mathbb{R}_2[X]$ . On va noter  $(e_0, e_1, e_2)$  au lieu de  $(e_1, e_2, e_3)$  la base orthonormée obtenue, afin d'avoir l'égalité  $\deg(e_i) = i$ . On a  $(1 | 1) = 2$  donc on prend  $e_0 = \frac{1}{\sqrt{2}}$ . Alors  $(X | e_0) = \frac{1}{\sqrt{2}} \int_{-1}^1 t dt = 0$  et  $(X | X) = \int_{-1}^1 t^2 dt = \frac{2}{3}$ , d'où  $e_1 = \frac{\sqrt{3}}{\sqrt{2}} X$ .

Puis  $(X^2 | e_0) = \frac{1}{\sqrt{2}} \int_{-1}^1 t^2 dt = \frac{\sqrt{2}}{3}$  et  $(X^2 | e_1) = \frac{\sqrt{3}}{\sqrt{2}} \int_{-1}^1 t^3 dt = 0$ , d'où  $e'_2 = X^2 - \frac{\sqrt{2}}{3} e_0$ ,

$$\|e'_2\|^2 = \int_{-1}^1 t^4 dt - \frac{2}{9} = \frac{2 \cdot 4}{5 \cdot 9} \quad \text{et} \quad e_2 = \frac{3\sqrt{5}}{2\sqrt{2}} \left( X^2 - \frac{1}{3} \right) = \frac{\sqrt{5}}{\sqrt{2}} \left( \frac{3X^2 - 1}{2} \right).$$

## 2.4. Bases directes ou indirectes. Groupes $O(n)$ et $SO(n)$ . Étude de $O(2)$ et $O(3)$

**Notation 2.4.0.** — Soit  $E$  un  $\mathbb{R}$ -espace vectoriel de dimension  $n$ . Si  $\mathcal{B}, \mathcal{B}'$  sont deux bases de  $E$ , on note  $\boxed{\det_{\mathcal{B}}(\mathcal{B}'')}$  le déterminant de la matrice de passage  $\text{Mat}_{\mathcal{B}}(\mathcal{B}')$ .

### Définition et proposition 2.4.1 (Orientations d'un $\mathbb{R}$ -espace vectoriel de dimension finie)

Soit  $E$  un  $\mathbb{R}$ -espace vectoriel de dimension  $n$ .

(1) On définit une relation d'équivalence  $\sim$  dans l'ensemble des bases de  $E$  en posant :

$$\mathcal{B} \sim \mathcal{B}' \quad \text{si} \quad \det_{\mathcal{B}}(\mathcal{B}') > 0.$$

(2) Il y a exactement deux classes d'équivalence.

(3) Chacune de ces classes est appelée une **orientation** de  $E$ , et « choisir une orientation de  $E$  », c'est choisir une base  $\mathcal{B}_0$  et l'orientation « qui va avec », c.-à-d., toutes les bases  $\mathcal{B}$  de  $E$  telles que  $\det_{\mathcal{B}_0}(\mathcal{B}) > 0$ . Celles-ci sont appelées les bases (orientées) **directes**, les autres sont appelées les bases (orientées) **indirectes**.



*Démonstration.* — (1) D'abord, on a  $\det_{\mathcal{B}}(\mathcal{B}) = 1$ , donc  $\mathcal{B} \sim \mathcal{B}$  (i.e. la relation  $\sim$  est réflexive). Soient  $\mathcal{B}, \mathcal{B}'$  deux bases de  $E$ . Si  $\mathcal{B}''$  est une troisième base de  $E$ , on a

$$(\dagger) \quad \text{Mat}_{\mathcal{B}}(\mathcal{B}'') = \text{Mat}_{\mathcal{B}}(\mathcal{B}') \cdot \text{Mat}_{\mathcal{B}'}(\mathcal{B}'') \quad \text{et donc} \quad \det_{\mathcal{B}}(\mathcal{B}'') = \det_{\mathcal{B}}(\mathcal{B}') \cdot \det_{\mathcal{B}'}(\mathcal{B}'').$$

Donc si  $\det_{\mathcal{B}}(\mathcal{B}')$  et  $\det_{\mathcal{B}'}(\mathcal{B}'')$  sont  $> 0$ , il en est de même de  $\det_{\mathcal{B}}(\mathcal{B}'')$ ; ceci montre que la relation  $\sim$  est transitive. De plus, prenant  $\mathcal{B}'' = \mathcal{B}$ , on obtient  $\det_{\mathcal{B}}(\mathcal{B}') \det_{\mathcal{B}'}(\mathcal{B}) = \det_{\mathcal{B}}(\mathcal{B}) = 1$ , d'où  $\det_{\mathcal{B}'}(\mathcal{B}) = \det_{\mathcal{B}}(\mathcal{B}')^{-1}$ , donc si  $\det_{\mathcal{B}}(\mathcal{B}') > 0$ , il en est de même de  $\det_{\mathcal{B}'}(\mathcal{B})$ ; ceci montre que la relation  $\sim$  est symétrique. Donc  $\sim$  est bien un relation d'équivalence, ce qui prouve (1).

Prouvons (2). Soit  $\mathcal{B}_0 = (e_1, \dots, e_n)$  une base de  $E$ , notons  $\mathcal{C}_0 = (-e_1, e_2, \dots, e_n)$ , c'est aussi une base de  $E$ , et  $\det_{\mathcal{B}_0}(\mathcal{C}_0) = -1$ , donc  $\mathcal{B}_0 \not\sim \mathcal{C}_0$  donc il y a au moins deux classes d'équivalence.

En fait, ce sont les deux seules classes. En effet, soit  $\mathcal{B}$  une base arbitraire; si  $\det_{\mathcal{B}_0}(\mathcal{B}) > 0$  alors  $\mathcal{B}$  est dans la classe de  $\mathcal{B}_0$ , et si  $\det_{\mathcal{B}_0}(\mathcal{B}) < 0$  alors

$$\det_{\mathcal{C}_0}(\mathcal{B}) = \underbrace{\det_{\mathcal{C}_0}(\mathcal{B}_0)}_{=-1} \cdot \underbrace{\det_{\mathcal{B}_0}(\mathcal{B})}_{<0} > 0$$

donc  $\mathcal{B}$  est dans la classe de  $\mathcal{C}_0$ . La proposition est démontrée.  $\square$



**Exemple 2.4.2 (Orientation canonique de  $\mathbb{R}^n$ ).** —  $\mathbb{R}^n$  est orienté par le choix de la base canonique  $\mathcal{B}_0 = (e_1, \dots, e_n)$ ; on dit que c'est l'**orientation canonique** de  $\mathbb{R}^n$ .

Dans toute la suite de cette section, on munit  $\mathbb{R}^n$  du produit scalaire usuel  $(x | y) = x_1 y_1 + \dots + x_n y_n$  et l'on note  $O(n)$  le groupe des isométries :

$$\begin{aligned} O(n) &= \{A \in M_n(\mathbb{R}) \mid {}^t A A = I_n\} \\ &= \{A \in M_n(\mathbb{R}) \mid (Ax \mid Ay) = (x \mid y) \quad \forall x, y \in \mathbb{R}^n\} \end{aligned}$$

(voir 2.1.12 pour d'autres caractérisations de  $O(n)$ ).



**Proposition 2.4.3.** — Soit  $A \in O(n)$ . Alors :

$$(1) \quad \boxed{\det(A) = \pm 1}$$

(2) Si  $A$  admet une valeur propre  $\lambda \in \mathbb{R}$ , alors  $\lambda = \pm 1$ .

(3) Les espaces propres  $V_+ = \text{Ker}(A - I_n)$  et  $V_- = \text{Ker}(A + I_n)$  sont **orthogonaux**, c.-à-d.,  $(v_+ \mid v_-) = 0$  pour tout  $v_+ \in V_+, v_- \in V_-$ .

*Démonstration.* — (1) On a  $1 = \det(I_n) = \det({}^t A A) = \det({}^t A) \cdot \det(A)$ , or on sait (cf. 2.1.1) que  $\det({}^t A) = \det(A)$  d'où  $\det(A)^2 = 1$  et donc  $\det(A) = \pm 1$ .

(2) Si  $A$  admet une valeur propre  $\lambda \in \mathbb{R}$ , soit  $v \neq 0$  un vecteur propre associé, alors

$$(v \mid v) = (Av \mid Av) = (\lambda v \mid \lambda v) = \lambda^2 (v \mid v).$$

Comme  $(v \mid v) \neq 0$  (car  $> 0$ ), ceci entraîne  $\lambda^2 = 1$ , d'où  $\lambda = \pm 1$ .

(3) Soient  $v_+ \in V_+$  et  $v_- \in V_-$ , alors

$$(v_+ \mid v_-) = (Av_+ \mid Av_-) = (v_+ \mid -v_-) = -(v_+ \mid v_-)$$

d'où  $(v_+ \mid v_-) = 0$ .  $\square$

**Définition 2.4.4 (Groupe  $SL_n$ ).** — On rappelle qu'on note

$$GL_n(\mathbb{R}) = \{A \in M_n(\mathbb{R}) \mid \det(A) \neq 0\},$$

on l'appelle le **G**roupe **L**inéaire (en anglais : **G**eneral **L**inear group). On appelle groupe **S**pécial **L**inéaire (en anglais : **S**pecial **L**inear group) le sous-groupe

$$SL_n(\mathbb{R}) = \{A \in M_n(\mathbb{R}) \mid \det(A) = 1\}.$$

(Ceci explique le S dans la notation  $SO(n) = \{A \in O(n) \mid \det(A) = 1\}$  introduite ci-dessous.)



**Définition 2.4.5 (Groupe  $SO(n)$ ).** — (1) On pose  $SO(n) = \{A \in O(n) \mid \det(A) = 1\}$ , c'est un sous-groupe de  $O(n)$ , appelé le groupe **spécial orthogonal**. On le note aussi parfois  $O^+(n)$ . Les éléments de  $SO(n)$  s'appellent les isométries **directes** de  $\mathbb{R}^n$ .

(2) On pose aussi :  $O^-(n) = \{A \in O(n) \mid \det(A) = -1\}$ , ce n'est **pas** un sous-groupe de  $O(n)$  (car si  $A, B \in O^-(n)$  alors  $\det(AB) = 1$  donc  $AB \in SO(n)$ ), mais d'après la proposition précédente, on a

$$O(n) = SO(n) \cup O^-(n). \quad (\text{réunion disjointe}).$$

(3) Pour tout  $\sigma \in O^-(n)$ , l'application  $f \mapsto f\sigma$  est une **bijection** de  $SO(n)$  sur  $O^-(n)$ , dont l'inverse est l'application  $g \mapsto g\sigma^{-1}$  (en effet,  $\det(\sigma^{-1}) = \det(\sigma)^{-1} = -1$  donc pour tout  $g \in O^-(n)$  on a  $\det(g\sigma^{-1}) = 1$  d'où  $g\sigma^{-1} \in SO(n)$ ).



**Proposition 2.4.6.** — Soit  $\mathcal{B}$  une base orthonormée de  $\mathbb{R}^n$  et soit  $P$  la matrice de passage  $\text{Mat}_{\mathcal{B}_0}(\mathcal{B}) \in O(n)$ . On a les équivalences suivantes :

$$\begin{cases} \mathcal{B} \text{ est une b.o.n. directe (i.e. } \det(P) > 0) & \iff P \in SO(n) \\ \mathcal{B} \text{ est une b.o.n. indirecte (i.e. } \det(P) < 0) & \iff P \in O^-(n). \end{cases}$$

*Démonstration.* — On sait, d'après 2.1.12, que  $P \in O(n)$ , d'où  $\det(P) = \pm 1$ , d'après 2.4.3. Donc  $\det(P)$  est  $> 0$  (resp.  $< 0$ ) si et seulement si il égale 1 (resp.  $-1$ ). La proposition en découle.  $\square$

On rappelle que les fonctions cosinus et sinus sont de période  $2\pi$  et que :

$$\forall t \in \mathbb{R}, \quad \boxed{\cos^2(t) + \sin^2(t) = 1} \quad \boxed{\cos(-t) = \cos(t)} \quad \boxed{\sin(-t) = -\sin(t)}.$$

On note  $\mathbb{R}/2\pi\mathbb{Z}$  le groupe abélien quotient de  $\mathbb{R}$  par le sous-groupe  $2\pi\mathbb{Z}$  : deux réels  $t, t'$  définissent le même élément de  $\mathbb{R}/2\pi\mathbb{Z}$  si et seulement si il existe  $k \in \mathbb{Z}$  tel que  $t' - t = 2k\pi$ . Rappelons le lemme suivant.

**Lemme 2.4.7.** — Soient  $a, b \in \mathbb{R}$  tels que  $a^2 + b^2 = 1$ . Il existe un unique  $\theta \in \mathbb{R}/2\pi\mathbb{Z}$  tel que  $\cos(\theta) = a$  et  $\sin(\theta) = b$ .

*Démonstration.* — Comme cosinus et sinus sont de période  $2\pi$ , il suffit de montrer l'existence et l'unicité modulo  $2\pi$  dans l'intervalle  $[-\pi, \pi]$ . Rappelons que la fonction cosinus est paire et induit une bijection décroissante de  $[0, \pi]$  sur  $[-1, 1]$ . Distinguons les trois cas suivants.

(1) Si  $a = 1$  alors  $b = 0$ ; dans ce cas, 0 est l'unique élément  $\theta$  de  $[-\pi, \pi]$  tel que  $\cos(\theta) = 1$  et  $\sin(\theta) = 0$ .

(2) Si  $a = -1$  alors  $b = 0$ ; dans ce cas, les seuls éléments  $\theta$  de  $[-\pi, \pi]$  tels que  $\cos(\theta) = -1$  et  $\sin(\theta) = 0$  sont  $\pm\pi$ , qui sont égaux modulo  $2\pi$ .

(3) Enfin, si  $a \neq \pm 1$  alors  $b = \pm\sqrt{1 - a^2} \neq 0$ . Dans ce cas, il existe dans  $[-\pi, \pi]$  deux éléments, opposés,  $\theta$  et  $-\theta$ , tels que  $\cos(\pm\theta) = a$ , et comme la fonction sinus vérifie  $\sin^2 = 1 - \cos^2$  et est impaire, alors  $\theta$  est uniquement déterminé par la condition  $\sin(\theta) = b$ .  $\square$

Dans ce qui suit, on munit  $\mathbb{R}^2$  du produit scalaire usuel  $(\mid)$ , de la norme euclidienne associée  $\|\cdot\|$ , et de l'orientation définie par la base canonique  $\mathcal{B}_0 = (e_1, e_2)$ .

**Définition et proposition 2.4.8 (Angle orienté de deux vecteurs non nuls de  $\mathbb{R}^2$ )**

Soient  $w, w' \in \mathbb{R}^2 - \{0\}$ . Il existe un unique  $\theta \in \mathbb{R}/2\pi\mathbb{Z}$  vérifiant les deux conditions suivantes :

(1)  $(w \mid w') = \cos(\theta) \|w\| \cdot \|w'\|$

(2)  $\sin(\theta)$  est du même signe ( $> 0$ ,  $= 0$  ou  $< 0$ ) que  $\det_{\mathcal{B}_0}(w, w')$ .

On appelle  $\theta$  l'**angle orienté** de  $w$  à  $w'$  et on le note  $\widehat{ww'}$ . On a  $\widehat{w'w} = -\widehat{ww'}$ .

*Démonstration.* — Remplaçons d'abord  $w$  et  $w'$  par les vecteurs unitaires  $u = \frac{1}{\|w\|}w$  et  $v = \frac{1}{\|w'\|}w'$ , alors la condition (1) devient :  $\cos(\theta) = (u \mid v)$ . D'après l'inégalité de Cauchy-Schwarz, on sait que  $(u \mid v) \in [-1, 1]$  et que :

$$(u \mid v) = \pm 1 \iff u \text{ et } v \text{ sont liés} \iff \det_{\mathcal{B}_0}(u, v) = 0.$$

Par conséquent, notant  $\varepsilon \in \{-1, 0, 1\}$  le « signe » de  $\det_{\mathcal{B}_0}(u, v)$ , on voit que les conditions (1) et (2) équivalent à dire que  $\cos(\theta) = (u \mid v) = a \in [-1, 1]$  et  $\sin(\theta) = \varepsilon\sqrt{1 - a^2} = b$ ; d'après le lemme précédent, ceci détermine un unique  $\theta \in \mathbb{R}$  modulo  $2\pi$ .



Enfin, si  $\theta = \widehat{w'w}$ , alors  $-\theta$  vérifie (1) et est du même signe que  $\det_{\mathcal{B}_0}(w', w) = -\det_{\mathcal{B}_0}(w, w')$ , d'où  $-\theta = \widehat{w'w}$ . La proposition est démontrée.  $\square$

Rappelons les formules trigonométriques :

$$\forall \theta, \theta' \in \mathbb{R}, \quad \boxed{\cos(\theta + \theta') = \cos(\theta)\cos(\theta') - \sin(\theta)\sin(\theta')} \quad \boxed{\sin(\theta + \theta') = \cos(\theta)\sin(\theta') + \sin(\theta)\cos(\theta')}$$

En particulier, comme  $\cos(\pi/2) = 0$  et  $\sin(\pi/2) = 1$ , on a  $\cos(\theta + \frac{\pi}{2}) = -\sin(\theta)$  et  $\sin(\theta + \frac{\pi}{2}) = \cos(\theta)$ .

**Notation 2.4.9.** — Pour tout  $\theta \in \mathbb{R}$ , introduisons le vecteur  $\boxed{u_\theta = \cos(\theta)e_1 + \sin(\theta)e_2}$ . C'est l'unique vecteur unitaire  $u$  tel que  $\widehat{e_1u} = \theta$ . En effet, soit  $u = ae_1 + be_2$  un tel vecteur, alors  $a = (u | e_1) = \cos(\theta)$ , d'où  $b = \pm \sin(\theta)$ ; de plus,

$$\det_{\mathcal{B}_0}(e_1, u) = \det \begin{pmatrix} 1 & \cos(\theta) \\ 0 & b \end{pmatrix} = b$$

est du signe de  $\sin(\theta)$ , d'où  $b = \sin(\theta)$ .

**Isométries de  $\mathbb{R}^2$  euclidien.** — On peut maintenant aborder l'étude du groupe  $O(2)$  des isométries de l'espace euclidien  $\mathbb{R}^2$ . Soit  $f \in O(2)$ , écrivons  $f(e_1) = ae_1 + be_2$ , alors  $a^2 + b^2 = \|f(e_1)\| = 1$ . D'autre part, comme  $f$  préserve le produit scalaire, on a

$$(f(e_2) | f(e_1)) = (e_2 | e_1) = 0$$

donc  $f(e_2)$  appartient à la droite  $(\mathbb{R}f(e_1))^\perp$ , qui est engendrée par le vecteur unitaire  $\begin{pmatrix} -b \\ a \end{pmatrix}$ , et comme  $f(e_2)$  est aussi un vecteur unitaire de cette même droite, on a nécessairement  $f(e_2) = \pm \begin{pmatrix} -b \\ a \end{pmatrix}$ , donc deux cas sont possibles :

**1er cas.**  $f(e_2) = \begin{pmatrix} -b \\ a \end{pmatrix}$ , dans ce cas, la matrice de  $f$  dans la base canonique  $\mathcal{B}_0 = (e_1, e_2)$  est  $\begin{pmatrix} a & -b \\ b & a \end{pmatrix}$ , et son déterminant est  $a^2 + b^2 = 1$ .

**2ème cas.**  $f(e_2) = \begin{pmatrix} b \\ -a \end{pmatrix}$ , dans ce cas, la matrice de  $f$  dans la base canonique  $\mathcal{B}_0 = (e_1, e_2)$  est  $\begin{pmatrix} a & b \\ b & -a \end{pmatrix}$ , et son déterminant est  $-a^2 - b^2 = 1$ .

On voit que ces deux cas sont exclusifs l'un de l'autre :  $f$  appartient à  $SO(2)$  (resp. à  $O^-(2)$ ) si et seulement si on est dans le 1er cas (resp. 2ème cas). Étudions séparément ces deux cas.

 **Définition 2.4.10 (Rotations dans  $\mathbb{R}^2$ ).** — On munit  $\mathbb{R}^2$  du produit scalaire standard  $( | )$  et de l'orientation canonique. Soit  $\theta \in \mathbb{R}$ .

(1) On appelle **rotation d'angle  $\theta$** , et l'on note  $r_\theta$ , l'endomorphisme de  $\mathbb{R}^2$  dont la matrice dans la base canonique  $\mathcal{B}_0 = (e_1, e_2)$  est

$$(*) \quad R_\theta = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}$$

(2) On a  $\boxed{\det(r_\theta) = 1}$  et  $\boxed{\text{Tr}(r_\theta) = 2\cos(\theta)}$  d'où  $P_{r_\theta}(X) = X^2 - 2\cos(\theta)X + 1$ . Le discriminant réduit de ce trinôme est  $\Delta' = \cos^2(\theta) - 1$  qui est  $< 0$  sauf si  $\theta = 0$  ou  $\pi$  (modulo  $2\pi$ ), c.-à-d., sauf dans le cas de  $r_0 = \text{id}$  et de  $r_\pi = -\text{id}$ . En dehors de ces cas,  $r_\theta$  n'a pas de valeurs propres réelles.

 **Proposition 2.4.11 (Le groupe  $SO(2)$ ).** — (1) Pour tout  $\theta, \theta' \in \mathbb{R}$ , on a  $\boxed{r_\theta \circ r_{\theta'} = r_{\theta+\theta'} = r_{\theta'} \circ r_\theta}$ . Par conséquent, le groupe  $SO(2) = \{r_\theta \mid \theta \in \mathbb{R}\}$  est commutatif.

(2) Plus précisément, l'application  $\rho : \mathbb{R} \rightarrow SO(2)$ ,  $\theta \mapsto r_\theta$ , est un morphisme de groupes, et son noyau est le sous-groupe  $2\pi\mathbb{Z}$  de  $\mathbb{R}$ .

*Démonstration.* — D'après les deux cas étudiés plus haut, on sait que les éléments de  $SO(2)$  sont les matrices de la forme  $\begin{pmatrix} a & -b \\ b & a \end{pmatrix}$ , avec  $a^2 + b^2 = 1$ . Or, d'après le lemme 2.4.7, pour chaque telle matrice il existe  $\theta \in \mathbb{R}$ , unique modulo  $2\pi$ , tel que  $\cos(\theta) = a$  et  $\sin(\theta) = b$ . D'autre part, on a :

$$\begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} \begin{pmatrix} \cos(\theta') & -\sin(\theta') \\ \sin(\theta') & \cos(\theta') \end{pmatrix} = \begin{pmatrix} \cos(\theta)\cos(\theta') - \sin(\theta)\sin(\theta') & -\cos(\theta)\sin(\theta') - \sin(\theta)\cos(\theta') \\ \sin(\theta)\cos(\theta') + \cos(\theta)\sin(\theta') & \cos(\theta)\cos(\theta') - \sin(\theta)\sin(\theta') \end{pmatrix} = R_{\theta+\theta'},$$

et l'assertion (1) en résulte.

D'autre part, dire que  $\rho$  est un morphisme de groupes équivaut à dire que  $r_{\theta+\theta'} = r_\theta \circ r_{\theta'}$ , ce qu'on vient de vérifier. Enfin, le noyau de  $\rho$  est formé des  $\theta \in \mathbb{R}$  tels que  $R_\theta = I_2$ , i.e. tels que  $\cos(\theta) = 1$  et  $\sin(\theta) = 0$ , ce qui équivaut à  $\theta \in 2\pi\mathbb{Z}$ .  $\square$

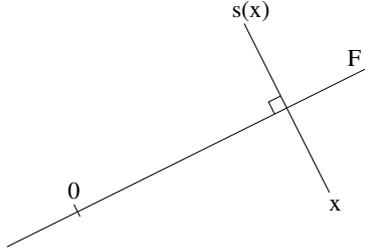
Étudions maintenant le cas des éléments de  $O^-(2)$ , i.e. des matrices de la forme

$$(\dagger) \quad A = \begin{pmatrix} a & b \\ b & -a \end{pmatrix}, \quad \text{avec } a^2 + b^2 = 1$$

c.-à-d., d'après le lemme 2.4.7, des matrices

$$(\ddagger) \quad J_\theta = \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ \sin(\theta) & -\cos(\theta) \end{pmatrix}.$$

Une telle matrice est de déterminant  $-1$  et de trace nulle ; son polynôme caractéristique est donc  $X^2 - 1 = (X-1)(X+1)$ , donc  $\mathbb{R}^2$  est la somme directe des espaces propres  $D_+ = \text{Ker}(A - I_2)$  et  $D_- = \text{Ker}(A + I_2)$ , chacun d'eux étant de dimension 1 (i.e. une droite vectorielle). De plus, d'après la proposition 2.4.3,  $D_+$  et  $D_-$  sont orthogonaux, donc  $A$  est la symétrie orthogonale par rapport à la droite  $F = D_+ = \text{Ker}(A - I_2)$  :



Réciproquement, pour  $c, d \in \mathbb{R}$  avec  $(c, d) \neq (0, 0)$ , déterminons la matrice dans la base canonique  $\mathcal{B}_0$  de la symétrie orthogonale  $s_\Delta$  par rapport à la droite  $\Delta$  engendré par le vecteur  $u = ce_1 + de_2$ . Alors  $\Delta$  est la droite orthogonale au vecteur  $v = \begin{pmatrix} -d \\ c \end{pmatrix}$  donc, d'après 2.3.4,  $s_\Delta$  est définie par la formule :

$$\forall x \in \mathbb{R}^2, \quad s_\Delta(x) = x - 2 \frac{(x | v)}{(v | v)} v.$$

On a  $(v | v) = c^2 + d^2$ ,  $(e_1 | v) = -d$  et  $(e_2 | v) = c$  donc, appliquant la formule ci-dessus à  $x = e_1$  puis  $x = e_2$ , on obtient

$$s_\Delta(e_1) = e_1 + 2 \frac{d}{c^2 + d^2} (-de_1 + ce_2), \quad s_\Delta(e_2) = e_2 - 2 \frac{c}{c^2 + d^2} (-de_1 + ce_2)$$

$$\text{d'où } \text{Mat}_{\mathcal{B}_0}(s_\Delta) = \frac{1}{c^2 + d^2} \begin{pmatrix} c^2 - d^2 & 2cd \\ 2cd & d^2 - c^2 \end{pmatrix}.$$

Appliquons ce qui précède dans le cas suivant. Notons  $s_\varphi$  la symétrie orthogonale par rapport à la droite engendrée par le vecteur  $u_\varphi = \cos(\varphi)e_1 + \sin(\varphi)e_2$ . (Notons que  $u_\varphi$  et  $u_{\varphi+\pi} = -u_\varphi$  engendrent la même droite, donc  $\varphi$  n'est déterminé que modulo  $\pi\mathbb{Z}$ .) D'après ce qui précède, la matrice de  $s_\varphi$  dans la base canonique  $\mathcal{B}_0$  est :

$$\begin{pmatrix} \cos^2(\varphi) - \sin^2(\varphi) & 2 \sin(\varphi) \cos(\varphi) \\ 2 \sin(\varphi) \cos(\varphi) & \sin^2(\varphi) - \cos^2(\varphi) \end{pmatrix} = \begin{pmatrix} \cos(2\varphi) & \sin(2\varphi) \\ \sin(2\varphi) & -\cos(2\varphi) \end{pmatrix} = J_{2\varphi}$$

On peut donc résumer ce qui précède dans la :



**Proposition 2.4.12 (Description de  $O^-(2)$ ).** — L'ensemble  $O^-(2) = \{A \in O(2) \mid \det(A) = -1\}$  est formé des matrices  $A = \begin{pmatrix} a & b \\ b & -a \end{pmatrix}$ , avec  $a^2 + b^2 = 1$ ; dans ce cas,  $A$  est la symétrie orthogonale par rapport à la droite  $D_+ = \text{Ker}(A - I_2)$ . D'autre part, il existe un unique  $\theta \in [0, 2\pi[$  tel que

$$A = \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ \sin(\theta) & -\cos(\theta) \end{pmatrix} = J_\theta$$

et  $A$  est la symétrie orthogonale  $s_{\theta/2}$  par rapport à la droite  $\mathbb{R}u_{\theta/2}$ .

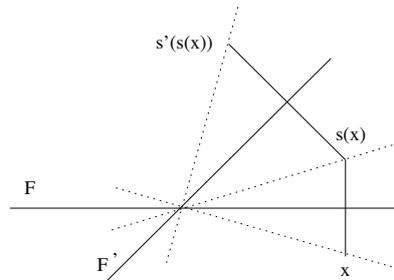
Enfin, on peut vérifier par un calcul matriciel (que le lecteur est invité à faire!) que  $J_{\theta'} J_\theta = R_{\theta' - \theta}$  d'où

$$(*) \quad \forall \varphi, \varphi' \in \mathbb{R}, \quad s_{\varphi'} \circ s_\varphi = r_{2(\varphi' - \varphi)}.$$

Ceci peut aussi se voir géométriquement comme suit. Posons  $f = s_{\varphi'} \circ s_\varphi$ ; Comme  $\det(f) = 1$ , alors  $f$  appartient à  $SO(2)$  donc est une rotation  $r_\theta$ . Pour déterminer l'angle  $\theta$ , il suffit de voir de combien « tourne » un vecteur  $x$  arbitraire. Or si on prend  $x = u_\varphi$ , alors  $s_\varphi(x) = x$  et donc  $f(x) = s_{\varphi'}(x)$  est le symétrique de  $x$  par rapport à la droite  $\mathbb{R}u_{\varphi'}$ , d'où

$$\varphi' - \varphi = x \widehat{u}_{\varphi'} = u_{\varphi'} \widehat{f}(x) \quad \text{et donc} \quad x \widehat{f}(x) = 2(\varphi' - \varphi).$$

(Dans la figure ci-dessous, prendre  $x$  sur la droite  $F$ .)



Observons que dans le terme de droite de (\*), la rotation obtenue ne dépend que de la différence  $\varphi' - \varphi$  et donc, pour obtenir  $r_\theta$ , on peut choisir arbitrairement  $\varphi$  ou bien  $\varphi'$ , c.-à-d., on a, pour tout  $\varphi, \theta \in \mathbb{R}$  :

$$(**) \quad r_\theta = s_{\varphi + \frac{\theta}{2}} \circ s_\varphi = s_\varphi \circ s_{\varphi - \frac{\theta}{2}}$$

Comme  $s_\varphi^2 = \text{id}$ , ceci permet de décrire complètement la « table de multiplication » dans  $O(2)$  (noter que  $O(2)$  n'est pas commutatif) :



**Proposition 2.4.13 (Structure de  $O(2)$ ).** — Pour tout  $\theta, \varphi \in \mathbb{R}$ , on a  $r_\theta \circ r_\varphi = r_{\theta + \varphi} = r_\varphi \circ r_\theta$

$$s_\theta \circ s_\varphi = r_{2(\theta - \varphi)} \quad \text{et}$$

$$r_\theta \circ s_\varphi = s_{\varphi + \frac{\theta}{2}}$$

$$s_\varphi \circ r_\theta = s_{\varphi - \frac{\theta}{2}}$$

Terminons ce paragraphe avec la proposition suivante. Par définition, la base canonique  $\mathcal{B}_0 = (e_1, e_2)$  est orthonormée directe. La base  $\mathcal{C}_0 = (e_2, e_1)$  est orthonormée indirecte puisque  $\text{Mat}_{\mathcal{B}_0}(\mathcal{C}_0) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$  est de déterminant  $-1$ .



**Proposition 2.4.14.** — Soit  $\theta \in \mathbb{R}$  et soit  $r_\theta$  la rotation d'angle  $\theta$  dans  $\mathbb{R}^2$  (cf. 2.4.10).

$$(1) \quad \text{La matrice de } r_\theta \text{ dans toute b.o.n. directe est } R_\theta = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}.$$

$$(2) \quad \text{La matrice de } r_\theta \text{ dans la base } \mathcal{C}_0 = (e_2, e_1) \text{ et dans toute b.o.n. indirecte est } R_{-\theta}.$$

*Démonstration.* — (1) On a  $\text{Mat}_{\mathcal{B}_0}(r_\theta) = R_\theta$ . Si  $\mathcal{B}$  est une b.o.n. directe, la matrice de passage  $P = \text{Mat}_{\mathcal{B}_0}(\mathcal{B})$  appartient à  $SO(2)$ , et comme  $SO(2)$  est commutatif,  $\text{Mat}_{\mathcal{B}}(r_\theta) = P^{-1}R_\theta P$  égale  $R_\theta$ .

(2) On a  $\text{Mat}_{\mathcal{C}_0}(r_\theta) = R_{-\theta}$ . Si  $\mathcal{C}$  est une b.o.n. indirecte, la matrice de passage  $P = \text{Mat}_{\mathcal{C}_0}(\mathcal{C})$  appartient à  $SO(2)$ , et comme  $SO(2)$  est commutatif,  $\text{Mat}_{\mathcal{C}}(r_\theta) = P^{-1}R_{-\theta}P$  égale  $R_{-\theta}$ .  $\square$

**Description des éléments de  $O(3)$ .** — Commençons par une remarque sur  $M_3(\mathbb{R})$ . Soit  $A \in M_3(\mathbb{R})$ , alors  $P_A(X) \in \mathbb{R}[X]$  est de degré 3 donc a dans  $\mathbb{C}$  :

- (a) ou bien une racine réelle  $\lambda_1$  et deux racines  $z, \bar{z} \in \mathbb{C} - \mathbb{R}$ ,
- (b) ou bien trois racines réelles  $\lambda_1, \lambda_2, \lambda_3$  (pas nécessairement distinctes).

Comme  $A$  est trigonalisable dans  $M_3(\mathbb{C})$ , on a

$$\det(A) = \begin{cases} \lambda_1 z \bar{z} & \text{dans le cas (a)} \\ \lambda_1 \lambda_2 \lambda_3 & \text{dans le cas (b)}. \end{cases}$$

Supposons maintenant  $A \in O(3)$  et notons  $u$  l'endomorphisme de  $\mathbb{R}^3$  correspondant à  $A$ . Alors on sait que  $\det(A) = \pm 1$ , et que si  $\lambda \in \mathbb{R}$  est valeur propre de  $A$ , alors  $\lambda = \pm 1$ . Dans le cas (a), on a  $z = x + iy$  avec  $y \neq 0$ , donc  $z\bar{z} = x^2 + y^2$  est un réel  $> 0$ , et donc l'égalité  $\det(A) = \lambda_1 z\bar{z}$  entraîne que  $\det(A)$  et  $\lambda_1$  (tous deux  $\in \{-1, 1\}$ ) sont de même signe, donc sont égaux. Dans le cas (b), on a  $\lambda_i = \pm 1$ , et les  $\lambda_i$  ne peuvent être tous les trois égaux à  $-\det(A)$ , car sinon leur produit vaudrait  $-\det(A)$  au lieu de  $\det(A)$ . Ceci montre déjà que, dans les deux cas,  $\det(A) = \pm 1$  est une valeur propre de  $A$ . Étudions les cas selon que  $\det(A) = 1$  ou  $-1$ .

**1er cas :**  $\det(A) = 1$ , i.e.  $A \in SO(3)$ . Le cas  $A = I_3$  étant trivial, supposons de plus que  $A \neq I_3$ . On a vu que 1 est valeur propre de  $A$ ; soit  $f_3$  un vecteur propre associé, quitte à diviser  $f_3$  par sa norme, on peut supposer que  $\|f_3\| = 1$ .

Soit  $P = (\mathbb{R}f_3)^\perp$ , c'est un sous-espace de  $\mathbb{R}^3$  de dimension 2, i.e. un plan. Pour tout  $x \in P$ , on a (puisque  $Af_3 = f_3$  et que  $A$  préserve le produit scalaire) :

$$(Ax \mid f_3) = (Ax \mid Af_3) = (x \mid f_3) = 0.$$

Ceci montre que  $Ax \in P$ , i.e.  $P$  est stable par  $A$ , donc  $A$  induit une isométrie  $u_P$  de  $P$ . Soit  $\mathcal{C} = (f_1, f_2)$  une base de  $P$ , alors  $\mathcal{B} = (f_1, f_2, f_3)$  est une base de  $\mathbb{R}^3$  et

$$(\star) \quad \text{Mat}_{\mathcal{B}}(u) = \left( \begin{array}{cc|c} \text{Mat}_{\mathcal{C}}(u_P) & & 0 \\ \hline 0 & & 1 \end{array} \right)$$

d'où  $\det(u_P) = \det(A) = 1$ , donc  $u_P$  est une rotation de  $P$ . Notons  $\theta$  son angle. On a  $\theta \neq 0$ , car sinon on aurait  $u_P = \text{id}_P$  et donc, d'après  $(\star)$ ,  $u = \text{id}$  d'où  $A = I_3$ , cas trivial qu'on a exclu. De plus, comme  $u_P \neq \text{id}_P$  alors 1 n'est pas valeur propre de  $u_P$  (cf. 2.4.10), et donc  $u_P - \text{id}_P$  et  $A - I_3$  sont de rang 2. Par conséquent,  $\text{Ker}(A - I_3)$  est de dimension 1, donc engendré par  $f_3$ . On dit que  $\mathbb{R}f_3 = \text{Ker}(A - I_3)$  est **l'axe** de la rotation  $A \neq I_3$ .

D'autre part, on peut déterminer  $\theta \in [-\pi, \pi]$  au signe près en prenant la trace. En effet, l'égalité précédente nous montre que  $\text{Tr}(A) = \text{Tr}(u) = 1 + \text{Tr}(u_P)$ ; d'autre part, on sait que  $\text{Tr}(u_P) = 2 \cos(\theta)$  (cf. 2.4.10) donc, combinant ces deux égalités, on obtient que :

$$\boxed{\text{Tr}(A) = 1 + 2 \cos(\theta)}, \quad \text{d'où} \quad \boxed{\cos(\theta) = \frac{\text{Tr}(A) - 1}{2}}$$

ce qui détermine  $\cos(\theta)$ , et détermine donc  $\theta$  au signe près.

Pour fixer le signe de  $\theta$ , il faut choisir une **orientation** de  $P$  (cf. 2.4.14). On la choisit en disant qu'une b.o.n.  $(f_1, f_2)$  de  $P$  est **directe** si la b.o.n.  $(f_1, f_2, f_3)$  de  $\mathbb{R}^3$  (muni de l'orientation canonique) est directe, c.-à-d., si  $\det_{\mathcal{B}_0}(f_1, f_2, f_3) = 1$ .

Soit  $(f_1, f_2)$  une telle b.o.n. directe de  $P$ . Alors on sait qu'il existe  $\theta \in \mathbb{R}$ , unique modulo  $2\pi\mathbb{Z}$ , tel que

$$\text{Mat}_{(f_1, f_2, f_3)}(u) = \begin{pmatrix} \cos(\theta) & -\sin(\theta) & 0 \\ \sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

et l'on dira alors que  $A$  (ou  $u$ ) est la **rotation d'axe orienté par  $f_3$  et d'angle  $\theta$** .

Remarquons tout de suite que si on change  $f_3$  en  $-f_3$ , alors la base  $(f_2, f_1, -f_3)$  est directe et l'on a

$$\text{Mat}_{(f_2, f_1, -f_3)}(u) = \begin{pmatrix} \cos(\theta) & \sin(\theta) & 0 \\ -\sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

donc  $A$  est aussi la **rotation d'axe orienté par  $-f_3$  et d'angle  $-\theta$** .



**Définition 2.4.15.** — Soit  $A \neq I_3$  une rotation d'axe  $\mathbb{R}f_3$  et d'angle  $\theta$ . Lorsque  $\theta = \pi$  (i.e. lorsque  $\cos(\theta) = -1$ , ce qui équivaut à  $\text{Tr}(A) = -1$ ), l'orientation de l'axe n'a pas d'importance, puisque  $-\pi = \pi$  modulo  $2\pi$ ; dans ce cas on dit que  $A$  est le **demi-tour d'axe**  $\mathbb{R}f_3$ .

Pour déterminer explicitement le signe de  $\sin(\theta)$  (lorsque  $\theta \neq \pi$ ), on dispose du lemme suivant.

**Lemme 2.4.16.** — Soit  $A \neq I_3$  une rotation d'axe  $D$ , orienté par un vecteur  $v_3 \neq 0$ , et d'angle  $\theta$ . Soit  $x \in \mathbb{R}^3 - D$ . Alors le signe de  $\sin(\theta)$  est le même que celui du déterminant  $\det_{\mathcal{B}_0}(x, Ax, v_3)$ .

*Démonstration.* — Notons  $f_3$  le vecteur unitaire  $\frac{1}{\|v_3\|}v_3$  et écrivons  $x = y + \mu f_3$ , où  $y$  est la projection orthogonale de  $x$  sur  $P = D^\perp$  (et  $\mu = (x | f_3)$ ). Comme  $x \notin D$ , on a  $y \neq 0$ , posons  $\rho = \|y\|$ , alors  $f_1 = \frac{1}{\rho}y$  est un vecteur unitaire de  $P$ . Il existe dans le plan  $P$  deux vecteurs unitaires orthogonaux à  $f_1$  et seul l'un d'eux,  $f_2$ , est tel que la matrice  $Q = \text{Mat}_{\mathcal{B}_0}(f_1, f_2, f_3)$  vérifie  $\det(Q) = 1$ , i.e. tel que  $(f_1, f_2)$  soit une b.o.n. directe de  $P$ . Alors, notant  $u$  l'endomorphisme de  $\mathbb{R}^3$  défini par  $A$ , on a :

$$\text{Mat}_{(f_1, f_2, f_3)}(u) = \begin{pmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

donc  $u(x) = u(\rho f_1 + \mu f_3)$  égale  $(\rho \cos \theta)f_1 + (\rho \sin \theta)f_2 + \mu f_3$ , donc la matrice

$$M = \text{Mat}_{(f_1, f_2, f_3)}(x, u(x), f_3) = \begin{pmatrix} \rho & \rho \cos \theta & 0 \\ 0 & \rho \sin \theta & 0 \\ \mu & \mu & 1 \end{pmatrix}$$

a pour déterminant  $\rho^2 \sin \theta$ . Comme  $M' = \text{Mat}_{\mathcal{B}_0}(x, Ax, f_3)$  égale  $QM$ , et  $\det(Q) = 1$ , on obtient que

$$\det(M') = \det(Q) \det(M) = \det(M) = \rho^2 \sin \theta$$

est du même signe que  $\sin \theta$ , et il en est de même pour  $\det_{\mathcal{B}_0}(x, Ax, v_3) = \|v_3\| \cdot \det_{\mathcal{B}_0}(x, Ax, f_3)$ .  $\square$

Remarquons que puisque le résultat ci-dessus est valable pour *tout*  $x \in \mathbb{R}^3 - D$ , on aura intérêt à choisir, en pratique, un tel  $x$  de façon à ce que le calcul du déterminant  $\det_{\mathcal{B}_0}(x, Ax, v_3)$  soit « le plus simple possible ». On récapitulera plus loin les résultats obtenus dans le cas où  $A \in SO(3)$ ; traitons maintenant le cas où  $A \in O^-(3)$ .

**2ème cas :**  $\det(A) = -1$ , i.e.  $A \in O^-(3)$ . Le cas  $A = -I_3$  étant trivial, supposons de plus que  $A \neq -I_3$ . On a vu que  $-1 = \det(A)$  est valeur propre de  $A$ ; soit  $f_3$  un vecteur propre associé, quitte à diviser  $f_3$  par sa norme, on peut supposer que  $\|f_3\| = 1$ .

On montre comme précédemment que le plan  $P = (\mathbb{R}f_3)^\perp$  est stable par  $u$ , et que la restriction  $u_P$  de  $u$  à  $P$  est de déterminant 1, donc une rotation. On fait le même choix d'orientation de  $P$ , c.-à-d., on choisit une b.o.n.  $(f_1, f_2)$  de  $P$  telle que  $(f_1, f_2, f_3)$  soit une b.o.n. directe de  $\mathbb{R}^3$ . Alors il existe  $\theta \in \mathbb{R}$ , unique modulo  $2\pi\mathbb{Z}$ , tel que

$$(\dagger) \quad \text{Mat}_{(f_1, f_2, f_3)}(u) = \begin{pmatrix} \cos(\theta) & -\sin(\theta) & 0 \\ \sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

On a  $\theta \neq \pi$ , car sinon on aurait  $u = -\text{id}$  d'où  $A = -I_3$ , cas trivial qu'on a exclu. De plus, comme  $u_P \neq -\text{id}_P$  alors  $-1$  n'est pas valeur propre de  $u_P$  (cf. 2.4.10), et donc  $u_P + \text{id}_P$  et  $A + I_3$  sont de rang 2. Par conséquent,  $\text{Ker}(A + I_3)$  est de dimension 1, donc engendré par  $f_3$ . On dit que  $D = \text{Ker}(A + I_3)$  est la droite des **anti-invariants** de  $A$ .

D'autre part, on peut déterminer  $\theta \in [-\pi, \pi]$  au signe près en prenant la trace. En effet,  $(\dagger)$  nous montre que  $\text{Tr}(A) = \text{Tr}(u)$  égale  $-1 + 2 \cos(\theta)$ , donc :

$$\boxed{\text{Tr}(A) = -1 + 2 \cos(\theta)}, \quad \text{d'où} \quad \boxed{\cos(\theta) = \frac{\text{Tr}(A) + 1}{2}}$$

ce qui détermine  $\cos(\theta)$ , et détermine donc  $\theta$  au signe près. Dans le cas particulier où  $\theta = 0$ , i.e.  $u_P = \text{id}_P$ ,  $A$  est la **symétrie orthogonale**  $\sigma_P$  par rapport au plan  $P$  (cf. 2.3.3 et 2.3.4).

Dans le cas général (i.e. lorsque  $\theta \neq 0$ ), on voit que

$$\text{Mat}_{(f_1, f_2, f_3)}(u\sigma_P) = \begin{pmatrix} \cos(\theta) & -\sin(\theta) & 0 \\ \sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{pmatrix} = \text{Mat}_{(f_1, f_2, f_3)}(\sigma_P u)$$

donc  $u\sigma_P = \sigma_P u$  est la rotation  $R$  d'axe  $D$  orienté par  $f_3$  et d'angle  $\theta$ , et donc  $u = R\sigma_P = \sigma_P R$  est la composée commutative de  $\sigma_P$  et de  $R$ . On dit alors que  $A$  est la **rotation gauche d'axe  $D$  orienté par  $f_3$  et d'angle  $\theta$** .

Remarquons tout de suite que si on change  $f_3$  en  $-f_3$ , alors la base  $(f_2, f_1, -f_3)$  est directe et l'on a

$$\text{Mat}_{(f_2, f_1, -f_3)}(u) = \begin{pmatrix} \cos(\theta) & \sin(\theta) & 0 \\ -\sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

donc  $A$  est aussi la **rotation gauche d'axe  $D$  orienté par  $-f_3$  et d'angle  $-\theta$** .

Pour déterminer explicitement le signe de  $\sin(\theta)$  (lorsque  $\theta \neq 0$ ), on dispose du lemme suivant.

**Lemme 2.4.17.** — Soit  $A \neq -I_3$  une rotation gauche d'axe  $D$ , orienté par un vecteur  $v_3 \neq 0$ , et d'angle  $\theta \neq 0$ . Soit  $x \in \mathbb{R}^3 - D$ . Alors le déterminant  $\det_{\mathcal{B}_0}(x, Ax, v_3)$  est de **même signe** que  $\sin(\theta)$ .

*Démonstration.* — Notons  $f_3$  le vecteur unitaire  $\frac{1}{\|v_3\|}v_3$  et écrivons  $x = y + \mu f_3$ , où  $y$  est la projection orthogonale de  $x$  sur  $P = D^\perp$ . Comme  $x \notin D$ , on a  $y \neq 0$ , posons  $\rho = \|y\|$ , alors  $f_1 = \frac{1}{\rho}y$  est un vecteur unitaire de  $P$ . Il existe dans le plan  $P$  deux vecteurs unitaires orthogonaux à  $f_1$  et seul l'un d'eux,  $f_2$ , est tel que la matrice  $Q = \text{Mat}_{\mathcal{B}_0}(f_1, f_2, f_3)$  vérifie  $\det(Q) = 1$ . Alors, notant  $u$  l'endomorphisme de  $\mathbb{R}^3$  défini par  $A$ , on a

$$\text{Mat}_{(f_1, f_2, f_3)}(u) = \begin{pmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

donc  $u(x) = u(\rho f_1 + \mu f_3)$  égale  $(\rho \cos \theta)f_1 + (\rho \sin \theta)f_2 - \mu f_3$ , donc la matrice

$$M = \text{Mat}_{(f_1, f_2, f_3)}(x, u(x), f_3) = \begin{pmatrix} \rho & \rho \cos \theta & 0 \\ 0 & \rho \sin \theta & 0 \\ \mu & -\mu & 1 \end{pmatrix}$$

a pour déterminant  $\rho^2 \sin \theta$ . Comme  $M' = \text{Mat}_{\mathcal{B}_0}(x, Ax, f_3)$  égale  $QM$ , et  $\det(Q) = 1$ , alors

$$\det(M') = \det(Q) \det(M) = \det(M) = \rho^2 \sin \theta$$

est de même signe que  $\sin \theta$ , et il en est de même pour  $\det_{\mathcal{B}_0}(x, Ax, v_3) = \|v_3\| \cdot \det_{\mathcal{B}_0}(x, Ax, f_3)$ .  $\square$

---

En résumé, on a obtenu le théorème suivant.

**Théorème 2.4.18 (Classification des éléments de  $O(3)$ ).** — On a  $O(3) = \text{SO}(3) \cup O^-(3)$  (réunion disjointe).

(1) Si  $A \in \text{SO}(3)$ , alors  $A = I_3$  ou bien  $D = \text{Ker}(A - I_3)$  est de dimension 1 et  $A$  est une rotation d'axe  $D$ . Dans le second cas, l'angle de rotation  $\theta$  est déterminé au signe près par l'égalité  $2 \cos(\theta) = \text{Tr}(A) - 1$ . Pour fixer le signe, on choisit un générateur  $v_3$  de  $D$ , ce qui détermine une orientation du plan  $P = D^\perp$ . Alors, pour  $x \in \mathbb{R}^3 - D$  arbitraire,  $\det_{\mathcal{B}_0}(x, Ax, v_3)$  est du même signe que  $\sin(\theta)$ , et l'on dit que  $A$  est la **rotation d'axe orienté par  $v_3$  et d'angle  $\theta$** . Dans le cas particulier où  $\theta = \pi$ , on dit que  $A$  est le **demi-tour d'axe  $D$** .

(2) Si  $A \in O^-(3)$ , alors  $A = -I_3$  ou bien  $D = \text{Ker}(A + I_3)$  est de dimension 1 et  $A$  est une rotation gauche d'axe  $D$ . Dans le second cas, l'angle de rotation  $\theta$  est déterminé au signe près par l'égalité  $2 \cos(\theta) = \text{Tr}(A) + 1$ . Pour fixer le signe, on choisit un générateur  $v_3$  de  $D$ , ce qui détermine une orientation du plan  $P = D^\perp$ . Alors, pour  $x \in \mathbb{R}^3 - D$  arbitraire,  $\det_{\mathcal{B}_0}(x, Ax, v_3)$  est de même signe que  $\sin(\theta)$ , et l'on dit que  $A$  est la **rotation gauche d'axe orienté par  $v_3$  et d'angle  $\theta$** . Dans le cas particulier où  $\theta = 0$ ,  $A$  est la **symétrie orthogonale par rapport au plan  $P$** .

**2.4.19. Produit vectoriel et déterminant des éléments de  $O(3)$ .** — Dans ce paragraphe, on donne une définition « intrinsèque » du produit vectoriel  $\wedge : \mathbb{R}^3 \times \mathbb{R}^3 \rightarrow \mathbb{R}^3$ , qui utilise la structure euclidienne de  $\mathbb{R}^3$ . Ceci a pour conséquence qu'on peut calculer le produit vectoriel dans n'importe quelle base orthonormée *directe*. De plus, ceci donne un moyen de calcul très simple du déterminant (égal à  $\pm 1$ ) d'un élément  $A$  de  $O(3)$  : il suffit pour cela de calculer un mineur (= sous-déterminant) de taille 2. On présente ces résultats sous la forme d'un exercice corrigé.

On munit  $E = \mathbb{R}^3$  du produit scalaire standard  $(\mid)$  et de la norme euclidienne associée  $\|x\| = \sqrt{(x \mid x)}$ . Soit  $\mathcal{B}_0 = (e_1, e_2, e_3)$  la base canonique (qui est orthonormée) et soit  $E^* = \text{Hom}_{\mathbb{R}}(E, \mathbb{R})$  l'espace dual de  $E$ . Pour tout  $x \in E$ , soit  $\phi_x \in E^*$  l'application  $w \mapsto (x \mid w)$ .

(1) Montrer que l'application  $\theta : E \rightarrow E^*$ ,  $x \mapsto \phi_x$  est linéaire et bijective.

Solution : D'abord, il résulte de la bilinéarité de  $(\mid)$  que, pour tout  $x \in E$ , l'application  $\phi_x : w \mapsto (x \mid w)$  est une forme linéaire sur  $E$ , i.e. un élément de  $E^*$ , et que l'application  $\theta : E \rightarrow E^*$ ,  $x \mapsto \phi_x$  est linéaire. Son noyau est

$$\text{Ker}(\theta) = \{x \in E \mid \phi_x = 0\} = \{x \in E \mid \forall y \in E, (x \mid y) = 0\}$$

qui est le noyau de  $(\mid)$ . Or ce noyau est nul, puisque  $(x \mid x) = 0$  entraîne  $x = 0$ . Ceci montre que  $\theta$  est injective. Comme  $\dim E^* = \dim E = 3$ , il résulte du théorème du rang que  $\theta$  est aussi surjective, donc bijective.

(2) Pour tout  $u, v \in E$ , montrer qu'il existe un unique vecteur  $f(u, v) \in E$  tel que  $\det_{\mathcal{B}_0}(u, v, w) = (f(u, v) \mid w)$  pour tout  $w \in E$ . On note  $f(u, v) = u \wedge v$  et on l'appelle le produit vectoriel de  $u$  et  $v$ .

Solution :  $u, v$  étant fixés, l'application  $\gamma_{u,v} : E \rightarrow \mathbb{R}$ ,  $w \mapsto \det_{\mathcal{B}_0}(u, v, w)$  est linéaire, i.e. est un élément de  $E^*$ . Donc, d'après la question précédente, il existe un unique vecteur  $f(u, v) \in E$  tel que  $\gamma_{u,v} = \phi_{f(u,v)}$ , i.e. tel que

$$\forall w \in E, \quad \det_{\mathcal{B}_0}(u, v, w) = (f(u, v) \mid w).$$

Désormais, on note  $f(u, v) = u \wedge v$ .

(3) Écrivant  $u = \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix}$ ,  $v = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}$  et prenant  $w = e_1$ , puis  $w = e_2$  et  $w = e_3$ , déterminer les coordonnées  $(f_1, f_2, f_3)$  de  $u \wedge v$  dans la base  $\mathcal{B}_0$ .

Solution : On a  $f_1 = (u \wedge v \mid e_1) = \det_{\mathcal{B}_0}(u, v, e_1) = \begin{vmatrix} u_1 & v_1 & 1 \\ u_2 & v_2 & 0 \\ u_3 & v_3 & 0 \end{vmatrix} = u_2 v_3 - u_3 v_2$ . On montre de même que  $f_2 = (u \wedge v \mid e_2) = u_3 v_1 - u_1 v_3$  et  $f_3 = (u \wedge v \mid e_3) = u_1 v_2 - u_2 v_1$ . On peut aussi procéder « par identification », i.e. pour  $u, v$  fixés et pour un vecteur  $w = \begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix}$  variable, on a :

$$(f(u, v) \mid w) = \begin{vmatrix} u_1 & v_1 & w_1 \\ u_2 & v_2 & w_2 \\ u_3 & v_3 & w_3 \end{vmatrix} = (u_2 v_3 - u_3 v_2) w_1 - (u_1 v_3 - u_3 v_1) w_2 + (u_1 v_2 - u_2 v_1) w_3$$

(en développant par rapport à la 3e colonne), d'où par identification :

$$\begin{cases} f_1 = u_2 v_3 - u_3 v_2, \\ f_2 = -(u_1 v_3 - u_3 v_1) = u_3 v_1 - u_1 v_3, \\ f_3 = u_1 v_2 - u_2 v_1. \end{cases}$$

(4) Montrer que l'application  $E \times E \rightarrow E$ ,  $(u, v) \mapsto u \wedge v$  est bilinéaire, et qu'elle est alternée (i.e.  $u \wedge u = 0$  pour tout  $u \in E$ ).

Solution : La bilinéarité de l'application  $(u, v) \mapsto u \wedge v$  ainsi que l'égalité  $u \wedge u = 0$  peuvent se déduire de l'égalité

$$\begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} \wedge \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = \begin{pmatrix} u_2 v_3 - u_3 v_2 \\ u_3 v_1 - u_1 v_3 \\ u_1 v_2 - u_2 v_1 \end{pmatrix}$$

établie dans la question précédente. Ceci peut aussi se voir directement, comme suit. Soient  $u, u', v, v' \in E$  et  $t \in \mathbb{R}$ . Pour tout  $w \in E$ , on a

$$\begin{aligned} (tu \wedge v + u' \wedge v \mid w) &= t(u \wedge v \mid w) + (u' \wedge v \mid w) = t \det_{\mathcal{B}_0}(u, v, w) + \det_{\mathcal{B}_0}(u', v, w) \\ &= \det_{\mathcal{B}_0}(tu + u', v, w) = ((tu + u') \wedge v \mid w). \end{aligned}$$

Comme  $(\mid)$  est non dégénéré (i.e. de noyau nul), ceci entraîne  $tu \wedge v + u' \wedge v = (tu + u') \wedge v$ . On montre de même que  $u \wedge (tv + v') = tu \wedge v + u \wedge v'$ . Donc l'application  $E \times E \rightarrow E$ ,  $(u, v) \mapsto u \wedge v$  est bilinéaire. De plus, pour tout  $u \in E$ , on a  $0 = \det_{\mathcal{B}_0}(u, u, u \wedge u) = (u \wedge u \mid u \wedge u)$ , d'où  $u \wedge u = 0$ .

(5) Soient  $u, v, w \in E$ . Pour toute base orthonormée directe  $\mathcal{B}$  de  $E$ , montrer que  $\det_{\mathcal{B}}(u, v, w) = \det_{\mathcal{B}_0}(u, v, w)$ .

**Solution** : Si  $\mathcal{C} = \{f_1, f_2, f_3\}$  et  $\mathcal{D}$  sont deux bases de  $E$  et si  $g \in \text{End}(E)$ , on note  $\text{Mat}_{\mathcal{D}, \mathcal{C}}(g)$  la matrice de  $g$  en prenant  $\mathcal{C}$  comme « base de départ » et  $\mathcal{D}$  comme « base d'arrivée », i.e. la matrice qui exprime  $g(f_i)$  ( $i = 1, 2, 3$ ) dans la base  $\mathcal{D}$ . Soit  $P = \text{Mat}_{\mathcal{B}_0}(\mathcal{B})$  et soit  $g$  l'endomorphisme de  $\mathbb{R}^3$  défini par  $g(e_1) = u$ ,  $g(e_2) = v$  et  $g(e_3) = w$ . On a

$$\text{Mat}_{\mathcal{B}_0}(u, v, w) = \text{Mat}_{\mathcal{B}_0, \mathcal{B}_0}(g) = \text{Mat}_{\mathcal{B}_0, \mathcal{B}}(\text{id}) \cdot \text{Mat}_{\mathcal{B}, \mathcal{B}_0}(g) = P \cdot \text{Mat}_{\mathcal{B}}(u, v, w)$$

d'où  $\det_{\mathcal{B}_0}(u, v, w) = \det(P) \cdot \det_{\mathcal{B}}(u, v, w)$ . Or par hypothèse  $P \in \text{SO}(3)$ , d'où  $\det(P) = 1$  et donc  $\det_{\mathcal{B}_0}(u, v, w) = \det_{\mathcal{B}}(u, v, w)$ .

(6) Soient  $u, v \in E$  deux vecteurs unitaires orthogonaux et soit  $p \in E$  l'unique vecteur tel que  $\mathcal{B} = (u, v, p)$  soit une base orthonormée directe de  $E$ . En utilisant la question précédente montrer que, pour tout  $w \in E$ , on a  $(u \wedge v \mid w) = (p \mid w)$ . Que peut-on en conclure ?

**Solution** : Tout  $w \in E$  s'écrit de façon unique  $w = au + bv + cp$ , avec  $c = (p \mid w)$  (et de même  $a = (u \mid w)$ , etc.) donc

$$(u \wedge v \mid w) = \det_{\mathcal{B}}(u, v, w) = \begin{vmatrix} 1 & 0 & a \\ 0 & 1 & b \\ 0 & 0 & (p \mid w) \end{vmatrix} = (p \mid w).$$

Comme  $(\mid)$  est non dégénéré (car défini positif), il en résulte que  $p = f(u, v)$ . Donc : *étant donnés deux vecteurs unitaires orthogonaux  $u, v$ , le produit vectoriel  $u \wedge v$  est l'unique vecteur  $p$  tel que  $(u, v, p)$  soit une base orthonormée directe.* Et donc si  $\mathcal{B}' = (u, v, p')$  est une base orthonormée, on a les équivalences :

$$\mathcal{B}' \text{ directe} \Leftrightarrow p' = p, \quad \mathcal{B}' \text{ indirecte} \Leftrightarrow p' = -p.$$

(7) Soit  $A = \begin{pmatrix} u_1 & v_1 & t_1 \\ u_2 & v_2 & t_2 \\ u_3 & v_3 & t_3 \end{pmatrix} \in O(3)$  et soient  $C_1, C_2, C_3$  les colonnes de  $A$ . Déduire de la question précédente que  $A \in \text{SO}(3) \Leftrightarrow C_3 = C_1 \wedge C_2$ , et aussi que  $A \in O^-(3) \Leftrightarrow C_3 = -C_1 \wedge C_2$ . (Remarque : on a donc  $C_3 = \det(A) C_1 \wedge C_2$ .)

Si par exemple  $t_3 \neq 0$ , en déduire, en utilisant la formule explicite pour  $C_1 \wedge C_2$  obtenue dans la question 3, que  $A \in \text{SO}(3) \Leftrightarrow u_1 v_2 - u_2 v_1 = t_3$  (et donc  $A \in O^-(3) \Leftrightarrow u_1 v_2 - u_2 v_1 = -t_3$ ).

**Solution** :  $A$  est la matrice de passage de la base canonique  $\mathcal{B}_0$  à la base orthonormée  $\mathcal{C} = (C_1, C_2, C_3)$ . Donc, d'après la question précédente, on sait que :

$$\begin{cases} A \in \text{SO}(3) \Leftrightarrow \mathcal{C} \text{ est directe} \Leftrightarrow C_3 = C_1 \wedge C_2, \\ A \in O^-(3) \Leftrightarrow \mathcal{C} \text{ est indirecte} \Leftrightarrow C_3 = -C_1 \wedge C_2. \end{cases}$$

En particulier, si  $t_3 \neq 0$  alors on a les équivalences :

$$\begin{cases} t_3 = u_1 v_2 - u_2 v_1 \Leftrightarrow C_3 = C_1 \wedge C_2 \Leftrightarrow \det(A) = 1, \\ t_3 = -(u_1 v_2 - u_2 v_1) \Leftrightarrow C_3 = -C_1 \wedge C_2 \Leftrightarrow \det(A) = -1. \end{cases}$$

On a bien sûr des résultats analogues si  $t_1 \neq 0$  ou si  $t_2 \neq 0$ , mais **attention**, pour  $t_2$  le « mineur » correspondant est  $u_3 v_1 - u_1 v_3 = -(u_1 v_3 - u_3 v_1)$ , voir le calcul fait dans la question 3.

(8) Soient  $x, y \in E$  deux vecteurs linéairement indépendants, et soient  $r = \|x\|$  et  $r' = \|y\|$ . Soit  $P$  le plan engendré par  $x$  et  $y$ , soit  $(u, v)$  une base orthonormée de  $P$ , où  $u = \frac{1}{r}x$  et soit  $\theta$  l'unique élément de  $\mathbb{R}/2\pi\mathbb{Z}$  tel que  $y = r'(\cos(\theta)u + \sin(\theta)v)$ . Montrer que  $\|x \wedge y\| = rr' |\sin(\theta)|$ . Indication : Utiliser la question 4 pour exprimer  $x \wedge y$  en fonction de  $u \wedge v$  puis, notant  $\mathcal{B}$  la base orthonormée directe  $(u, v, u \wedge v)$ , calculer  $\det_{\mathcal{B}}(x, y, x \wedge y)$  et utiliser la question 5.

**Solution** : Par hypothèse, on a  $x = ru$  et  $y = r'(\cos(\theta)u + \sin(\theta)v)$  donc, comme  $\wedge$  est bilinéaire et alterné, on a :

$$x \wedge y = ru \wedge r'(\cos(\theta)u + \sin(\theta)v) = rr' \sin(\theta)u \wedge v.$$

Par conséquent, notant  $\mathcal{B}$  la base orthonormée directe  $(u, v, u \wedge v)$ , on a

$$\text{Mat}_{\mathcal{B}}(x, y, x \wedge y) = \begin{pmatrix} r & r' \cos(\theta) & 0 \\ 0 & r' \sin(\theta) & 0 \\ 0 & 0 & rr' \sin(\theta) \end{pmatrix},$$

donc  $\det_{\mathcal{B}}(x, y, x \wedge y) = (rr')^2 \sin(\theta)^2$ . Or, d'après la question 5, ceci égale  $\det_{\mathcal{B}_0}(x, y, x \wedge y) = (x \wedge y | x \wedge y) = \|x \wedge y\|^2$ . On en déduit que  $\|x \wedge y\| = rr' |\sin(\theta)|$ .

(9) Montrer qu'une base orthonormée  $\mathcal{C} = (u, v, f)$  est directe si et seulement si la base  $\mathcal{D} = (f, u, v)$  est directe.

**Solution** : La matrice de passage  $\text{Mat}_{\mathcal{C}}(\mathcal{D}) = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$  est de déterminant 1, donc  $\mathcal{C}$  est directe si

et seulement si  $\mathcal{D}$  l'est. On pouvait aussi dire que le déterminant change de signe lorsqu'on échange deux colonnes, d'où :

$$\det_{\mathcal{B}_0}(u, v, f) = -\det_{\mathcal{B}_0}(u, f, v) = \det_{\mathcal{B}_0}(f, u, v)$$

et donc  $(u, v, f)$  est directe si et seulement si  $(f, u, v)$  l'est.

(10) Soient  $f \in E$  un vecteur unitaire et  $P$  le plan orthogonal à  $f$ . Soient  $a, b \in \mathbb{R}$  tels que  $a^2 + b^2 = 1$  et soit  $\theta$  l'unique élément de  $\mathbb{R}/2\pi\mathbb{Z}$  tel que  $\cos \theta = a$  et  $\sin \theta = b$ . Soit  $R$  la rotation d'axe  $\mathbb{R}f$  orienté par  $f$  et d'angle  $\theta$ . Montrer que :

$$(*) \quad \forall y \in P, \quad R(y) = ay + b(f \wedge y),$$

(si  $y = 0$  c'est clair, et si  $y \neq 0$  écrire  $y = ru$  avec  $u$  unitaire et  $r = \|y\|$ ), puis que :

$$(**) \quad \forall x \in E, \quad R(x) = (x | f)f + a\pi(x) + bf \wedge \pi(x) = (x | f)f + a\pi(x) + bf \wedge x,$$

où  $\pi(x) = x - (x | f)f$  est la projection orthogonale de  $x$  sur  $P$ . Enfin, si  $f = \begin{pmatrix} p \\ q \\ t \end{pmatrix}$ , écrire  $\text{Mat}_{\mathcal{B}_0}(R)$  sous

la forme  $aI_3 + (1 - a)S + bA$ , pour deux matrices  $S, A$  à déterminer.

**Solution** : Si  $y = 0$ , l'égalité (\*) est  $0 = 0$ , donc on peut supposer  $y \neq 0$  et poser  $y = ru$ , où  $u$  est unitaire et  $r = \|y\|$ . Les deux membres de (\*) étant linéaires en  $y$ , il suffit d'établir (\*) lorsque  $y = u$ . Soit alors  $v$  l'unique vecteur de  $P$  tel que  $(u, v, f)$  soit une base orthonormée directe de  $E$ . D'après la définition

de  $R$ , on a  $\text{Mat}_{(u,v,f)}(R) = \begin{pmatrix} \cos(\theta) & -\sin(\theta) & 0 \\ \sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{pmatrix}$  et donc  $R(u) = \cos(\theta)u + \sin(\theta)v$ . Or, d'après la

question précédente, la base  $(f, u, v)$  est directe, donc d'après la question 6, on a  $v = f \wedge u$ . On a donc :  $R(u) = \cos(\theta)u + \sin(\theta)f \wedge u$ , et revenant au cas de  $y \in P$  arbitraire, on a donc obtenu que :

$$(*) \quad \forall y \in P, \quad R(y) = \cos(\theta)u + \sin(\theta)f \wedge y = ay + b(f \wedge y).$$

Soit maintenant  $x \in E$ , sa projection orthogonale sur la droite  $\mathbb{R}f$  est  $(x | f)f$ , et sa projection orthogonale sur  $P$  est  $\pi(x) = x - (x | f)f$ . On a bien sûr  $R(\lambda f) = \lambda f$  pour tout  $\lambda \in \mathbb{R}$ , et donc on déduit de la question précédente que :

$$R(x) = (x | f)f + R(\pi(x)) = (x | f)f + a\pi(x) + bf \wedge \pi(x),$$

De plus, comme  $x = (x | f)f + \pi(x)$  et que  $\wedge$  est bilinéaire et alterné, on a  $f \wedge x = f \wedge \pi(x)$ , et l'on obtient l'égalité désirée :

$$(**) \quad \forall x \in E, \quad R(x) = (x | f)f + a\pi(x) + bf \wedge x.$$

Enfin, écrivons  $f = \begin{pmatrix} p \\ q \\ t \end{pmatrix}$ , avec  $\|f\|^2 = p^2 + q^2 + t^2 = 1$ . Alors, on a

$$R(e_1) = p \begin{pmatrix} p \\ q \\ t \end{pmatrix} + a \left( \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} - p \begin{pmatrix} p \\ q \\ t \end{pmatrix} \right) + b \begin{pmatrix} p \\ q \\ t \end{pmatrix} \wedge \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = a \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + (1 - a) \begin{pmatrix} p^2 \\ pq \\ pt \end{pmatrix} + b \begin{pmatrix} 0 \\ t \\ -q \end{pmatrix}$$

$$R(e_2) = q \begin{pmatrix} p \\ q \\ t \end{pmatrix} + a \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} - q \begin{pmatrix} p \\ q \\ t \end{pmatrix} + b \begin{pmatrix} p \\ q \\ t \end{pmatrix} \wedge \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = a \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + (1-a) \begin{pmatrix} pq \\ q^2 \\ qt \end{pmatrix} + b \begin{pmatrix} -t \\ 0 \\ p \end{pmatrix}$$

$$R(e_3) = t \begin{pmatrix} p \\ q \\ t \end{pmatrix} + a \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} - t \begin{pmatrix} p \\ q \\ t \end{pmatrix} + b \begin{pmatrix} p \\ q \\ t \end{pmatrix} \wedge \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = a \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} + (1-a) \begin{pmatrix} pt \\ qt \\ t^2 \end{pmatrix} + b \begin{pmatrix} q \\ -p \\ 0 \end{pmatrix}$$

donc

$$\text{Mat}_{\mathcal{B}_0} = aI_3 + (1-a) \begin{pmatrix} p^2 & pq & pt \\ pq & q^2 & qt \\ pt & qt & t^2 \end{pmatrix} + b \begin{pmatrix} 0 & -t & q \\ t & 0 & -p \\ -q & p & 0 \end{pmatrix}.$$

## 2.5. Appendice (†) : mesure des angles dans $\mathbb{R}^2$

Dans cet appendice, on donne une démonstration (parmi d'autres) de l'existence et des propriétés des fonctions cosinus et sinus.

On munit  $\mathbb{R}^2$  du produit scalaire usuel et on note  $(e_1, e_2)$  la base canonique. Tout nombre complexe  $z$  s'écrit de façon unique  $z = x + iy$ , où  $i$  est une racine carrée de  $-1$ , fixée une fois pour toutes. Ceci permet d'identifier  $\mathbb{C}$  à  $\mathbb{R}^2$ , en identifiant  $z = x + iy$  au vecteur de coordonnées  $(x, y)$ . Si  $z = x + iy$ , son **conjugué** est  $\bar{z} = x - iy$ . Soit

$$S^1 = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1\} = \{z = x + iy \in \mathbb{C} \mid z\bar{z} = 1\},$$

$S^1$  s'appelle le **cercle unité**.

On « rappelle » (ou l'on admet) les faits suivants : l'application  $f : \mathbb{R} \rightarrow \mathbb{C}$ ,  $t \mapsto e^{it} = \exp(it)$  vérifie les propriétés suivantes :

(a) c'est un morphisme de groupes :

$$(1) \quad \forall t, t' \in \mathbb{R}, \quad \boxed{e^{i(t+t')} = e^{it} e^{it'}}$$

(b) pour tout  $t \in \mathbb{R}$ , on a  $\overline{\exp(it)} = \exp(-it) = \exp(it)^{-1}$ , par conséquent,  $f$  est à valeurs dans  $S^1$ .

(c)  $f$  est dérivable et sa dérivée est l'application  $f' : t \mapsto ie^{it}$ .

On note alors  $\cos(t)$ , resp.  $\sin(t)$ , la partie réelle, resp. imaginaire, de  $e^{it}$ , c.-à-d., on pose

$$(2) \quad \forall t \in \mathbb{R}, \quad \boxed{e^{it} = \cos(t) + i \sin(t)}.$$

Alors, (a), (b) et (c) se récrivent :

$$(3) \quad \forall t, t' \in \mathbb{R}, \quad \cos(t+t') = \cos(t)\cos(t') - \sin(t)\sin(t'), \quad \sin(t+t') = \sin(t)\cos(t') + \cos(t)\sin(t'),$$

$$(4) \quad \forall t \in \mathbb{R}, \quad \cos(-t) = \cos(t), \quad \sin(-t) = -\sin(t), \quad \cos^2(t) + \sin^2(t) = 1,$$

$$(5) \quad \forall t \in \mathbb{R}, \quad \cos'(t) = -\sin(t), \quad \sin'(t) = \cos(t).$$

On montre alors le :

**Théorème 2.5.1.** — (1) *Le morphisme de groupes  $f : \mathbb{R} \rightarrow S^1$ ,  $t \mapsto e^{it}$  est surjectif.*

(2)  *$\text{Ker}(f)$  est le sous-groupe  $2\pi\mathbb{Z}$  de  $\mathbb{R}$  et  $\cos, \sin$  sont périodiques de période  $2\pi$ .*

*Démonstration.* — On a  $\cos(0) = 1$  donc  $\cos(t) > 0$  pour  $t$  voisin de 0, donc  $\sin$  est strictement croissante au voisinage de 0, et comme  $\sin(0) = 0$  on a  $\sin(t) > 0$  pour  $t > 0$  voisin de 0, et donc  $\cos(t) < 1$  pour  $t > 0$  voisin de 0.

Montrons d'abord qu'il existe  $t_0 > 0$  tel que  $\cos(t_0) = 0$ . Supposons au contraire que  $\cos(t) > 0$  pour tout  $t \geq 0$ . Alors  $\sin(t)$  serait croissante sur  $\mathbb{R}_+$ , donc  $\geq 0$  sur  $\mathbb{R}_+$ , et comme  $\cos'(t) = -\sin(t)$ , alors  $\cos(t)$  serait décroissante et  $\geq 0$  sur  $\mathbb{R}_+$  donc tendrait en  $+\infty$  vers une limite  $\ell \in [0, 1[$  ( $\ell < 1$  car  $\cos(t) < 1$  pour  $t > 0$  voisin de 0). Mais alors  $\cos(2t) = 2\cos^2(t) - 1$  tendrait aussi vers  $\ell$ , donc  $\ell$  serait racine du polynôme

$$2X^2 - X - 1 = 2(X-1)\left(X + \frac{1}{2}\right),$$

ce qui est impossible puisque  $\ell \in [0, 1[$ . Cette contradiction montre qu'il existe  $t_0 > 0$  tel que  $\cos(t_0) = 0$ . On note alors  $\frac{\pi}{2}$  le plus petit réel  $t_0 > 0$  tel que  $\cos(t_0) = 0$ . Tenant compte du fait que  $\cos(-t) = \cos(t)$ , on a donc :

$$(6) \quad \cos\left(\pm \frac{\pi}{2}\right) = 0 \quad \text{et} \quad \forall t \in \left] \frac{-\pi}{2}, \frac{\pi}{2} \right[ , \quad \cos(t) > 0.$$

Donc  $\sin$  est strictement croissante sur  $[0, \pi/2]$ , d'où  $\sin(\pi/2) = 1$  (étant  $\geq 0$  et de carré = 1), donc  $\sin$  est une bijection strictement croissante de  $[0, \pi/2]$  sur  $[0, 1]$ , et l'on a

$$(7) \quad e^{i\pi/2} = i,$$

et  $t \mapsto e^{it}$  est une bijection de  $[0, \pi/2]$  sur le quart de cercle  $\{(x, y) \in S^1 \mid x \geq 0, y \geq 0\}$ . D'après (1) ou (3), il résulte de (7) que :

$$(8) \quad \forall t \in \mathbb{R}, \quad \exp\left(i\left(t + \frac{\pi}{2}\right)\right) = ie^{it}, \quad \text{d'où} \quad \cos\left(t + \frac{\pi}{2}\right) = -\sin(t), \quad \sin\left(t + \frac{\pi}{2}\right) = \cos(t).$$

En particulier, on a  $e^{i\pi} = -1$ , et  $\sin(t) > 0$  pour tout  $t \in ]0, \pi[$ , donc  $\cos$  est une bijection strictement décroissante de  $[0, \pi]$  sur  $[-1, 1]$ .

Il en résulte que  $t \mapsto e^{it}$  est une bijection de  $[0, \pi]$  sur le demi-cercle supérieur  $\{(x, y) \in S^1 \mid y \geq 0\}$ , et de  $[-\pi, 0]$  sur le demi-cercle inférieur  $\{(x, y) \in S^1 \mid y \leq 0\}$ . Ceci prouve la surjectivité.

D'autre part, on a l'inclusion et les égalités suivantes :

$$(†) \quad 2\pi\mathbb{Z} \subseteq \text{Ker}(f) = \{T \in \mathbb{R} \mid e^{i(T+t)} = e^{it}, \quad \forall t \in \mathbb{R}\} = \{T \in \mathbb{R} \mid \cos(T+t) = \cos(t), \quad \forall t \in \mathbb{R}\} \\ = \{T \in \mathbb{R} \mid \sin((T+t) - t) = \sin(t), \quad \forall t \in \mathbb{R}\}.$$

En effet, comme  $e^{2i\pi} = 1$  on a  $2\pi\mathbb{Z} \subseteq \text{Ker}(f)$ . D'autre part, si  $T \in \text{Ker}(f)$  i.e. si  $e^{iT} = 1$  alors, d'après (1), on a  $e^{i(T+t)} = e^{it}$  pour tout  $t \in \mathbb{R}$ , d'où la première égalité.

D'autre part, comme on a  $\cos(t) = \sin(t + \frac{\pi}{2})$ , pour tout  $t \in \mathbb{R}$ , on obtient que

$$\{T \in \mathbb{R} \mid \cos(T+t) = \cos(t), \quad \forall t \in \mathbb{R}\} = \{T \in \mathbb{R} \mid \sin(T+t) = \sin(t), \quad \forall t \in \mathbb{R}\}$$

et cet ensemble égale donc  $\{T \in \mathbb{R} \mid e^{i(T+t)} = e^{it}, \quad \forall t \in \mathbb{R}\}$ .

Enfin, soit  $T \in \mathbb{R}$  tel que  $\cos(T+t) = \cos(t)$  pour tout  $t \in \mathbb{R}$ ; montrons que  $T \in 2\pi\mathbb{Z}$ . Quitte à changer  $T$  en  $-T$ , on peut supposer  $T \geq 0$ , soit alors  $n$  le plus grand entier  $\geq 0$  tel que  $2n\pi \leq T$ , alors  $T - 2n\pi$  vérifie la même propriété que  $T$  et appartient à  $[0, 2\pi[$ . Donc, pour obtenir que  $T = 2n\pi$ , il suffit de montrer que  $2\pi$  est le plus petit réel  $s > 0$  tel que  $\cos(s) = 1$ .

Or, d'après ce qui précède, on sait que  $1 > \cos(t) > -1$  pour tout  $t \in ]0, \pi[$ . Supposons que  $s \in ]0, 2\pi[$  vérifie  $\cos(s) = 1$ , alors  $\sin(s) = 0$  donc  $e^{is} = 1$ , donc  $x = e^{is/2}$  vérifie  $x^2 = 1$ , d'où  $x = \pm 1$  et donc  $\cos(s/2) = \pm 1$ , impossible puisque  $s/2 \in ]0, \pi[$ . Ceci montre que  $2\pi$  est le plus petit réel  $s > 0$  tel que  $\cos(s) = 1$ .

Il en résulte que l'inclusion dans (†) plus haut est une égalité. Ceci montre que  $\cos$  et  $\sin$  sont périodiques de période  $2\pi$ , et que  $\text{Ker}(f) = 2\pi\mathbb{Z}$ . Le théorème est démontré.  $\square$

**Remarque 2.5.2.** — Dans ce qui précède, on a défini  $2\pi$  comme le générateur positif du groupe  $\text{Ker}(f)$  (i.e. le plus petit élément  $> 0$  de  $\text{Ker}(f)$ ). On retrouve que  $2\pi$  est la longueur du cercle de rayon 1 comme suit. Pour toute fonction  $g : [a, b] \rightarrow \mathbb{R}^2$ ,  $t \mapsto g(t) = (x(t), y(t))$  continûment dérivable, on peut montrer que la longueur dans  $\mathbb{R}^2$  euclidien de la courbe  $\gamma = \{g(t) \mid t \in [a, b]\}$  est donnée par la formule :

$$L(\gamma) = \int_a^b \sqrt{(g'(t) \mid g'(t))} dt = \int_a^b \sqrt{x'(t)^2 + y'(t)^2} dt.$$

Appliquant ceci à  $f : [0, 2\pi] \rightarrow \mathbb{R}^2$ ,  $t \mapsto e^{it} = (\cos(t), \sin(t))$ , on obtient

$$L(S^1) = \int_0^{2\pi} \sqrt{1} dt = 2\pi.$$

## 2.6. Appendice (†) : décomposition d'Iwasawa de $GL_n(\mathbb{R})$

Afin d'énoncer un corollaire matriciel au théorème d'orthonormalisation de Gram-Schmidt, introduisons les notations suivantes. D'abord, notons  $\mathcal{B}_0 = (e_1, \dots, e_n)$  la base canonique de  $\mathbb{R}^n$ . Soit  $\text{TS}_n^+(\mathbb{R})$  l'ensemble des matrices triangulaires supérieures dont les coefficients diagonaux sont  $> 0$ .

**Lemme 2.6.1.** —  $\text{TS}_n^+(\mathbb{R})$  est un sous-groupe de  $GL_n(\mathbb{R})$ .

*Démonstration.* — Soient  $N, N' \in \text{TS}_n^+(\mathbb{R})$ , notons  $t_1, \dots, t_n$  et  $t'_1, \dots, t'_n$  leurs coefficients diagonaux. Alors  $NN'$  est triangulaire supérieure, de coefficients diagonaux  $t_i t'_i > 0$ , d'où  $NN' \in \text{TS}_n^+(\mathbb{R})$ . Donc il suffit de montrer que  $N^{-1} \in \text{TS}_n^+(\mathbb{R})$ . Soient  $C_1, \dots, C_n$  les vecteurs colonnes de  $N$ . Montrons par récurrence sur  $r \leq n$  que  $e_r = (1/t_r)C_r + \sum_{i < r} b_{ir}C_i$ . C'est OK pour  $r = 1$ , car  $C_1 = t_1 e_1$  d'où  $e_1 = (1/t_1)C_1$ . On peut donc supposer  $r \geq 2$  et l'assertion établie pour  $r - 1$ . En particulier, on a  $\text{Vect}(e_1, \dots, e_{r-1}) =$

$\text{Vect}(C_1, \dots, C_{r-1})$ . Comme  $C_r - t_r e_r \in \text{Vect}(e_1, \dots, e_{r-1})$ , on en déduit qu'il existe des coefficients  $b'_{ir}$ , pour  $i = 1, \dots, r-1$ , tels que

$$t_r e_r = C_r + \sum_{i=1}^{r-1} b'_{ir} C_i,$$

d'où le résultat voulu au cran  $r$ , en divisant par  $t_r$  et en posant  $b_{ir} = b'_{ir}/t_r$ . Il en résulte que  $N^{-1}$  est triangulaire supérieure, de termes diagonaux les  $1/t_i$ , qui sont  $> 0$ . Ceci prouve le lemme.  $\square$

On déduit alors du théorème d'orthonormalisation de Gram-Schmidt le :

**Corollaire 2.6.2 (Décomposition d'Iwasawa).** — *L'application ci-dessous est bijection :*

$$O(n) \times \text{TS}_n^+(\mathbb{R}) \rightarrow \text{GL}_n(\mathbb{R}), \quad (K, N) \mapsto KN$$

(Attention, ce n'est pas un morphisme de groupes, i.e.  $(KN)(K'N') \neq KK'NN'$  en général.)

*Démonstration.* — Soit  $P \in \text{GL}_n(\mathbb{R})$ , notons  $v_1, \dots, v_n$  les vecteurs colonnes de  $P$ . Alors  $\mathcal{B} = (v_1, \dots, v_n)$  est une base de  $\mathbb{R}^n$  et  $P = \text{Mat}_{\mathcal{B}_0}(\mathcal{B})$ , où  $\mathcal{B}_0$  est la base canonique de  $\mathbb{R}^n$ . D'après le théorème de Gram-Schmidt, il existe une base orthonormée  $\mathcal{C} = (f_1, \dots, f_n)$  de  $\mathbb{R}^n$  telle que, pour tout  $r = 1, \dots, n$ , on ait  $\text{Vect}(v_1, \dots, v_r) = \text{Vect}(f_1, \dots, f_r)$ , et  $(v_r | f_r) > 0$ , et ceci équivaut à dire que la matrice de passage  $N = \text{Mat}_{\mathcal{C}}(\mathcal{B})$  appartient à  $\text{TS}_n^+(\mathbb{R})$ . D'autre part, comme  $\mathcal{C}$  est orthonormée, la matrice de passage  $K = \text{Mat}_{\mathcal{B}_0}(\mathcal{C})$  appartient à  $O(n)$ . D'après la formule de changement de base, on a

$$P = \text{Mat}_{\mathcal{B}_0}(\mathcal{B}) = \text{Mat}_{\mathcal{B}_0}(\mathcal{C}) \cdot \text{Mat}_{\mathcal{C}}(\mathcal{B}) = KN$$

ce qui prouve l'existence.

Montrons l'unicité : supposons qu'on ait  $P = K' M$  avec  $K' \in O(n)$  et  $M = (m_{ij})_{i,j=1}^n \in \text{TS}_n^+(\mathbb{R})$ , et notons  $f_1, \dots, f_n$  les colonnes de la matrice  $K' = PM^{-1}$ . Alors  $\mathcal{C} = (f_1, \dots, f_n)$  est une b.o.n. de  $\mathbb{R}$ , on a  $\text{Mat}_{\mathcal{B}_0}(\mathcal{C}) = K'$  et la matrice exprimant  $\mathcal{B}$  dans la base  $\mathcal{C}$  est

$$\text{Mat}_{\mathcal{C}}(\mathcal{B}) = \text{Mat}_{\mathcal{C}}(\mathcal{B}_0) \cdot \text{Mat}_{\mathcal{B}_0}(\mathcal{B}) = K'^{-1} P = M.$$

Alors, pour tout  $r = 1, \dots, n$ , on a  $v_r = \sum_{i=1}^r m_{ir} f_i$ , avec  $(v_r | f_r) = m_{rr} > 0$ . De plus, d'après le lemme précédent (ou par un calcul direct), pour tout  $i = 1, \dots, n$ , les vecteurs  $f_1, \dots, f_i$  appartiennent à  $V_i = \text{Vect}(v_1, \dots, v_i)$ , donc en forment une base orthonormée. Donc la base  $\mathcal{C} = (f_1, \dots, f_n)$  vérifie les conditions du théorème de Gram-Schmidt, d'où par unicité  $M = N$ , puis  $K' = PN^{-1} = K$ .  $\square$

**Remarques 2.6.3.** — (1) En appliquant ce qui précède à  $P^{-1}$ , on voit qu'il existe un couple unique  $(K, N) \in O(n) \times \text{TS}_n^+(\mathbb{R})$  tel que  $P^{-1} = KN$ , i.e.  $P = N^{-1}K^{-1}$ . Comme l'application  $g \mapsto g^{-1}$  est une bijection sur chaque groupe, il en résulte que l'application  $\text{TS}_n^+(\mathbb{R}) \times O(n) \rightarrow \text{GL}_n(\mathbb{R})$ ,  $(N, K) \mapsto NK$  est aussi une bijection.

(2) L'espace vectoriel  $M_n(\mathbb{R}) \simeq \mathbb{R}^{n^2}$  a une structure naturelle d'espace topologique, définie par la distance associée à n'importe quelle norme sur  $M_n(\mathbb{R})$ , par exemple  $\|(a_{ij})\|_\infty = \text{Max}_{i,j} |a_{ij}|$  ou  $\|(a_{ij})\|_2 = \sqrt{\sum_{i,j} a_{ij}^2}$ , et donc les sous-ensembles  $O(n)$ ,  $\text{TS}_n^+(\mathbb{R})$  et  $\text{GL}_n(\mathbb{R})$  héritent d'une structure d'espace topologique. Il est facile de voir que l'application (de multiplication)  $(K, N) \mapsto KN$  est continue, et on peut montrer que c'est un homéomorphisme. D'abord, l'application  $\text{GL}_n(\mathbb{R}) \rightarrow \text{GL}_n(\mathbb{R})$ ,  $A \mapsto A^{-1}$  est continue, car  $A^{-1} = (1/\det(A)) {}^t\tilde{A}$ , où  $\tilde{A}$  désigne la matrice des cofacteurs de  $A$  (cf. ??). Il suffit donc de montrer que, si  $A = KN$ , l'application  $f : A \mapsto N^{-1}$  est continue, car alors les applications  $A \mapsto N$  et  $A \mapsto K = AN^{-1}$  le seront aussi. Or, si l'on note  $v_1, \dots, v_n$  les colonnes de  $A$ , la continuité de  $f : A \mapsto N^{-1}$  résulte de la démonstration du théorème de Gram-Schmidt, car les coefficients de  $N^{-1} = \text{Mat}_{(v_1, \dots, v_n)}(e_1, \dots, e_n)$  sont des fonctions continues des produits scalaires  $(v_i | v_j)$ .

## CHAPITRE 3

### ESPACES AFFINES EUCLIDIENS

**Résumé :** Dans ce chapitre, on étudie les « *espaces affines euclidiens* », intégration de la notion d'espace affine avec d'espace euclidien. On étudiera les isométries entre ces espaces et on les classifera notamment en dimension 2 et 3. À la fin du chapitre on a rassemblé un « formulaire » des espaces affines contenant les définitions et tous les résultats (mais pas les preuves !) qui sont nécessaires pour lire ce chapitre. Comme le précédent, ce chapitre contient beaucoup de résultats nouveaux et importants, qu'il faut essayer d'assimiler !

On a indiqué par des symboles  les définitions, exemples et résultats fondamentaux.

#### 3.1. Définitions

 **Définition 3.1.1.** — On dit qu'un espace affine  $(\mathcal{E}, E)$  est **euclidien** si  $E$  est de dimension finie et est muni d'un produit scalaire  $(\cdot | \cdot)$ . Dans ce cas,  $\mathcal{E}$  est muni de la **distance euclidienne**, définie par  $d(x, y) = \|\vec{xy}\| = \sqrt{(\vec{xy} | \vec{xy})}$ .

C'est bien une distance, car : (a)  $d(x, y) \geq 0$  et  $d(x, y) = 0 \Leftrightarrow x = y$ , (b)  $d(y, x) = d(x, y)$  puisque  $\|\vec{yx}\| = \|-\vec{xy}\| = |-1| \cdot \|\vec{xy}\| = \|\vec{xy}\|$ , et (c)  $d$  vérifie l'inégalité triangulaire :

$$d(x, y) + d(y, z) = \|\vec{xy}\| + \|\vec{yz}\| \geq \|\vec{xz}\| = d(x, z).$$

 **Définition 3.1.2 (Isométries).** — Soit  $\mathcal{E}$  un espace affine euclidien. Une **isométrie** de  $\mathcal{E}$  est une application  $f : \mathcal{E} \rightarrow \mathcal{E}$  qui conserve les distances, i.e. telle que :

$$\forall x, y \in \mathcal{E}, \quad \boxed{d(f(x), f(y)) = d(x, y).}$$

On notera  $\text{Is}(\mathcal{E})$  l'ensemble des isométries de  $\mathcal{E}$ .

On va voir qu'alors  $f$  est nécessairement affine et bijective. Commençons par quelques résultats préliminaires.

**Lemme 3.1.3.** — Soit  $f : \mathcal{E} \rightarrow \mathcal{E}$  une isométrie.

- (1)  $f$  est injective.
- (2) Si  $f$  est bijective,  $f^{-1}$  est une isométrie.
- (3) Si  $g : \mathcal{E} \rightarrow \mathcal{E}$  est une seconde isométrie,  $g \circ f$  est aussi une isométrie.

*Démonstration.* — (1) Si  $f(x) = f(y)$  alors  $0 = d(f(x), f(y)) = d(x, y)$  donc  $x = y$ . Ceci prouve que  $f$  est injective.

- (2) Soient  $x, y \in \mathcal{E}$ , posons  $x' = f^{-1}(x)$  et  $y' = f^{-1}(y)$ . Alors  $x = f(x')$  et  $y = f(y')$ , d'où

$$d(x, y) = d(f(x'), f(y')) = d(x', y') = d(f^{-1}(x), f^{-1}(y))$$

(la seconde égalité car  $f$  est une isométrie). Ceci montre que  $f^{-1}$  est une isométrie.

- (3) Pour tout  $x, y \in \mathcal{E}$ , on a  $d(x, y) = d(f(x), f(y)) = d(g(f(x)), g(f(y)))$ , donc  $g \circ f$  est une isométrie.  $\square$

 **Proposition 3.1.4 (Isométries affines).** — Soit  $f : \mathcal{E} \rightarrow \mathcal{E}$  une transformation affine.

(1)  $f$  est une isométrie si et seulement si sa partie linéaire  $\vec{f}$  est une isométrie vectorielle (i.e.  $\vec{f}$  préserve la norme).

(2) Dans ce cas,  $\vec{f}$  et  $f$  sont **bijectives**.

(3) En particulier, toute translation  $f = t_{\vec{u}}$  est une isométrie affine de  $\mathcal{E}$ .

*Démonstration.* — (1) Pour tout  $x, y \in \mathcal{E}$ , on a  $d(f(x), f(y)) = \|\overrightarrow{f(x)f(y)}\| = \|\vec{f}(\overrightarrow{xy})\|$  et ceci égale  $d(x, y) = \|\overrightarrow{xy}\|$  si et seulement si  $\|\vec{f}(\overrightarrow{xy})\| = \|\overrightarrow{xy}\|$ . Donc  $f$  est une isométrie si  $\vec{f}$  en est une. Réciproquement, si  $f$  est une isométrie, alors  $\vec{f}$  préserve la norme, car l'application  $\mathcal{E} \times \mathcal{E} \rightarrow E$ ,  $(x, y) \mapsto \overrightarrow{xy}$  est surjective (puisque pour  $x$  fixé,  $y \mapsto \overrightarrow{xy}$  est une bijection). Ceci prouve (1).

De plus, dans ce cas  $\vec{f}$  est **bijective** (car injective et  $E$  de dimension finie), donc  $f$  l'est aussi, d'après le point (3) de 3.3.12. Enfin, toute translation  $f = t_{\vec{u}}$  est une isométrie de  $\mathcal{E}$ , puisque sa partie linéaire égale  $\text{id}_E$ . La proposition est démontrée.  $\square$

**Théorème 3.1.5.** — Soit  $f : \mathcal{E} \rightarrow \mathcal{E}$  une isométrie. Alors  $f$  est **affine** et **bijective**.

Avant de démontrer le théorème, énonçons tout de suite un corollaire. On note  $O(E)$  le groupe des isométries vectorielles de  $E$ . Rappelons que pour tout  $\phi \in O(E)$ , on a  $\det(\phi) = \pm 1$ , et que l'on pose  $SO(E) = O^+(E) = \{\phi \in O(E) \mid \det(\phi) = 1\}$  et  $O^-(E) = \{\phi \in O(E) \mid \det(\phi) = -1\}$ ; de plus  $SO(E)$  est un sous-groupe de  $O(E)$  (appelé le groupe « spécial orthogonal » de  $E$ ).

**Corollaire 3.1.6 (Les groupes  $\text{Is}(\mathcal{E})$  et  $\text{Is}^+(\mathcal{E})$ ).** — Soit  $(\mathcal{E}, E)$  un espace affine euclidien. On a

$$(*) \quad \text{Is}(\mathcal{E}) = \{f \in \text{GA}(\mathcal{E}) \mid \vec{f} \in O(E)\}$$

et c'est un sous-groupe de  $\text{GA}(\mathcal{E})$ . De plus,  $\text{Is}(\mathcal{E})$  est la réunion de  $\text{Is}^+(\mathcal{E})$  (isométries **directes**) et de  $\text{Is}^-(\mathcal{E})$  (isométries **indirectes**), où  $\text{Is}^+(\mathcal{E})$  (resp.  $\text{Is}^-(\mathcal{E})$ ) désigne l'ensemble des  $f \in \text{Is}(\mathcal{E})$  tels que  $\vec{f} \in O^+(E)$  (resp.  $\vec{f} \in O^-(E)$ ). Enfin,  $\text{Is}^+(\mathcal{E})$  est un sous-groupe de  $\text{Is}(\mathcal{E})$ .

**Terminologie 3.1.7.** — Les isométries directes sont aussi appelées « **déplacements** », et les indirectes « **antidéplacements** ».

*Démonstration du corollaire.* — D'après le théorème, toute isométrie de  $\mathcal{E}$  est affine et bijective, donc appartient à  $\text{GA}(\mathcal{E})$ , et d'après la proposition 3.1.4, un élément  $f$  de  $\text{GA}(\mathcal{E})$  appartient à  $\text{Is}(\mathcal{E})$  si et seulement si  $\vec{f} \in O(E)$ . Ceci prouve l'égalité (\*).

Le fait que  $\text{Is}(\mathcal{E})$  soit un groupe découle du lemme 3.1.3. Enfin, la décomposition  $\text{Is}(\mathcal{E}) = \text{Is}^+(\mathcal{E}) \cup \text{Is}^-(\mathcal{E})$  et le fait que  $\text{Is}^+(\mathcal{E})$  soit un groupe, découlent des propriétés analogues de  $O(E)$ .  $\square$

*Démonstration du théorème.* — Il suffit de montrer que  $f$  est **affine**, elle sera alors **bijective** d'après la proposition 3.1.4. Fixons une origine arbitraire  $O \in \mathcal{E}$  et posons  $\vec{w} = \overrightarrow{f(O)O}$ . Alors  $\phi = t_{\vec{w}} \circ f$  est une isométrie, vérifiant  $\phi(O) = O$ . Il suffit de montrer que  $\phi$  est affine, car alors  $f = t_{-\vec{w}} \circ \phi$  le sera aussi. Comme  $\phi(O) = O$ , pour montrer que  $\phi$  est affine, il suffit de montrer que l'application  $\psi : E \rightarrow E$  définie par

$$\forall P \in \mathcal{E}, \quad \psi(\overrightarrow{OP}) = \overrightarrow{\phi(P)O} \quad \text{i.e.} \quad \forall \vec{u} \in E, \quad O + \psi(\vec{u}) = \phi(O + \vec{u})$$

est **linéaire**.

Soit  $\vec{u} \in E$ , posons  $P = O + \vec{u}$  et  $P' = \phi(P)$ , alors  $\psi(\vec{u}) = \overrightarrow{OP'}$  et donc :

$$\|\psi(\vec{u})\| = \|\overrightarrow{OP'}\| = d(O, P') \underset{(*)}{=} d(O, P) = \|\overrightarrow{OP}\| = \|\vec{u}\|$$

(l'égalité (\*) car  $\phi$  est une isométrie). Ceci prouve déjà que  $\psi$  préserve la norme, montrons de plus que  $\psi$  préserve le produit scalaire. Considérons un second vecteur  $\vec{v} \in E$  et posons  $Q = O + \vec{v}$  et  $Q' = \phi(Q)$ . Alors l'on a :

$$\text{d'une part,} \quad \overrightarrow{P'Q'} = \overrightarrow{OQ'} - \overrightarrow{OP'} = \vec{v} - \vec{u}$$

$$\text{d'autre part,} \quad \overrightarrow{P'Q'} = \overrightarrow{OQ'} - \overrightarrow{OP'} = \psi(\vec{v}) - \psi(\vec{u})$$

et

$$\|\psi(\vec{v}) - \psi(\vec{u})\| = \|\overrightarrow{P'Q'}\| = d(\phi(P), \phi(Q)) = d(P, Q) = \|\overrightarrow{PQ}\| = \|\vec{v} - \vec{u}\|.$$

Élevant cette égalité au carré, on obtient :

$$\|\psi(\vec{v})\|^2 + \|\psi(\vec{u})\|^2 - 2(\psi(\vec{v}) \mid \psi(\vec{u})) = \|\vec{v}\|^2 + \|\vec{u}\|^2 - 2(\vec{v} \mid \vec{u})$$

et comme  $\|\psi(\vec{v})\| = \|\vec{v}\|$  et  $\|\psi(\vec{u})\| = \|\vec{u}\|$ , ceci donne :

$$(\psi(\vec{v}) \mid \psi(\vec{u})) = (\vec{v} \mid \vec{u})$$

donc  $\psi$  préserve le produit scalaire.

On peut maintenant montrer que  $\psi$  est linéaire. Soit  $\mathcal{B} = (e_1, \dots, e_n)$  une base orthonormée de  $E$ , alors  $\psi(\mathcal{B}) = (\psi(e_1), \dots, \psi(e_n))$  est une famille orthonormée, donc libre, de cardinal  $n$ , donc c'est une base orthonormée de  $E$ . Soit  $\vec{u} \in E$  arbitraire, on peut écrire de façon unique

$$\vec{u} = \sum_{j=1}^n x_j e_j, \quad \psi(\vec{u}) = \sum_{j=1}^n y_j \psi(e_j)$$

mais alors, pour tout  $i$  on a  $y_i = (\psi(e_i) | \psi(\vec{u})) = (e_i | \vec{u}) = x_i$  et donc :

$$\forall \vec{u} = \sum_{j=1}^n x_j e_j, \quad \psi(\vec{u}) = \sum_{j=1}^n x_j \psi(e_j).$$

Ceci montre que  $\psi$  est linéaire. Ceci achève la preuve du théorème 3.1.5.  $\square$

On sait donc maintenant que toute isométrie de  $\mathcal{E}$  euclidien est nécessairement affine. Dans la suite, on écrira souvent : « Soit  $f$  une isométrie affine » (au lieu de simplement « isométrie ») pour insister sur le fait que  $f$  est affine (et, si l'on n'a pas vu le théorème 3.1.5, on prend comme hypothèse que  $f$  est une isométrie affine...). On veut maintenant établir que toute isométrie affine  $f$  s'écrit de façon unique  $f = t_{\vec{u}} \circ \phi$ , où  $\phi$  est une isométrie (affine) ayant un point fixe, et  $\phi$  et  $t_{\vec{u}}$  commutent :  $\phi \circ t_{\vec{u}} = t_{\vec{u}} \circ \phi$ . Commençons par le :



**Lemme 3.1.8.** — Soient  $(\mathcal{E}, E)$  un espace affine,  $f : \mathcal{E} \rightarrow \mathcal{E}$  une transformation affine, et  $\vec{u} \in E$ . Alors  $f \circ t_{\vec{u}} = t_{\vec{u}} \circ f$  si et seulement si  $\vec{f}(\vec{u}) = \vec{u}$  (i.e.  $\vec{u}$  est un point fixe de  $\vec{f}$ ).

*Démonstration.* — Pour tout point  $O \in \mathcal{E}$ , on a :

$$(f \circ t_{\vec{u}})(O) = f(O) + \vec{f}(\vec{u}), \quad (t_{\vec{u}} \circ f)(O) = f(O) + \vec{u}$$

et ces deux expressions coïncident si et seulement si  $\vec{f}(\vec{u}) = \vec{u}$ .  $\square$

**Théorème 3.1.9 (Décomposition canonique d'une isométrie affine)**

Soient  $(\mathcal{E}, E)$  un espace affine euclidien,  $f$  une isométrie affine de  $\mathcal{E}$ ,  $\vec{f}$  sa partie linéaire,  $F = \text{Fix}(\vec{f}) = \text{Ker}(\vec{f} - \text{id}_E)$ . Alors :

(1) Il existe une unique sea  $\mathcal{F}$  de direction  $F$  stable par  $f$ . La restriction  $f|_{\mathcal{F}}$  de  $f$  à  $\mathcal{F}$  est une translation  $t_{\vec{u}}$ , pour un certain  $\vec{u} \in F$ .

(2)  $g = t_{-\vec{u}} \circ f$  vérifie  $\text{Fix}(g) = \mathcal{F}$ , et l'on a  $f = t_{\vec{u}} \circ g = g \circ t_{\vec{u}}$ .

(3) L'écriture précédente est unique : si  $f = t_{\vec{u}'} \circ g'$ , avec  $\vec{u}' \in F$  et  $g'$  ayant un point fixe, alors  $\vec{u}' = \vec{u}$  et  $g' = g$ .

Par conséquent : toute isométrie affine  $f$  s'écrit de façon unique comme produit d'une isométrie ayant un point fixe et d'une translation de vecteur  $\vec{u} \in \text{Fix}(\vec{f})$ .

*Démonstration.* — Posons  $G = F^\perp$ , il est stable par  $\vec{f}$ . En effet, soit  $x \in G$ , pour tout  $y \in F$ , on a  $y = \vec{f}(y)$  et donc :

$$(\vec{f}(x) | y) = (\vec{f}(x) | \vec{f}(y)) = (x | y) = 0$$

ce qui prouve que  $\vec{f}(x) \in G$ . Choisissons un sea  $\mathcal{G}$  de direction  $G$ , et un point  $O \in \mathcal{G}$ . Comme  $E = F \oplus G$ , on peut écrire de façon unique

$$\overrightarrow{Of(O)} = u + v, \quad \text{avec } u \in F, v \in G.$$

Posons  $g = t_{-u} \circ f$ . Alors  $g(O) = O + v \in \mathcal{G}$ , et comme  $\vec{g} = \vec{f}$  envoie  $G$  dans  $G$ , alors pour tout  $P \in \mathcal{G}$  on a :

$$\vec{g}(\overrightarrow{OP}) \in G \quad \text{donc} \quad g(P) = g(O) + \vec{g}(\overrightarrow{OP}) \in \mathcal{G},$$

d'où  $g(\mathcal{G}) \subseteq \mathcal{G}$ . Notons alors  $h$  la restriction  $g|_{\mathcal{G}}$ , sa partie linéaire  $\vec{h}$  est la restriction à  $G$  de  $\vec{f}$ , d'où

$$\{w \in G \mid \vec{h}(w) = w\} = G \cap \text{Ker}(\vec{f} - \text{id}_E) = G \cap F = \{0\}$$

(la dernière égalité puisque  $G = F^\perp$ ). Donc  $\text{Fix}(\vec{h}) = \{0\}$  et donc, d'après la proposition 3.3.42,  $h$  admet un unique point fixe  $I_0$ , i.e.  $I_0$  est l'unique point fixe de  $g$  dans  $\mathcal{G}$ .

Posons maintenant  $\mathcal{F} = I_0 + F$ . Alors,  $f(I_0) = (t_{\vec{u}} \circ g)(I_0) = I_0 + \vec{u}$  et pour tout  $P \in \mathcal{F}$ , on a  $\overrightarrow{I_0 P} \in F = \text{Fix}(\vec{f}) = \text{Fix}(\vec{g})$  et donc :

$$\forall P \in \mathcal{F}, \quad g(P) = P \quad \text{et} \quad f(P) = P + \vec{u}.$$

Donc  $\mathcal{F} \subseteq \text{Fix}(g)$ . Réciproquement, si  $Q \in \text{Fix}(g)$ , alors  $\overrightarrow{I_0Q} = \overrightarrow{g(I_0)g(Q)} = \overrightarrow{f(I_0Q)}$ , donc  $\overrightarrow{I_0Q} \in \text{Fix}(\overrightarrow{f}) = F$  et  $Q \in \mathcal{F}$ . Ceci montre que  $\text{Fix}(g) = \mathcal{F}$ , et comme  $f = t_{\vec{u}} \circ g = g \circ t_{\vec{u}}$ , on a obtenu l'existence de  $\mathcal{F}, \vec{u}, g$  vérifiant les conditions (1) et (2) du théorème.

Démontrons maintenant l'unicité. Soit  $\mathcal{F}'$  un sea de  $\mathcal{E}$  de direction  $F$ , stable par  $f$ . Il est alors stable par  $g = t_{\vec{u}} \circ f$ . D'autre part, on a vu plus haut que  $g(\mathcal{G}) \subseteq \mathcal{G}$ ; de plus  $\mathcal{F}' \cap \mathcal{G}$  est un singleton  $\{J\}$ , puisque  $F \oplus G = E$  (cf. 3.3.39). Donc  $g(J) = J$ , or  $I_0$  est l'unique point fixe de  $g$  dans  $\mathcal{G}$ , d'où  $J = I_0$  et donc  $\mathcal{F}' = I_0 + F$  égale  $\mathcal{F}$ . Ceci prouve l'unicité de  $\mathcal{F}$ .

Enfin, supposons que  $f = t_{\vec{u}'} \circ g'$ , avec  $\vec{u}' \in F$  et  $g'$  ayant un point fixe  $A$ . Alors  $f(A) = (t_{\vec{u}'} \circ g')(A) = A + \vec{u}'$  et donc pour tout  $\vec{w} \in F$ , on a :

$$f(A + \vec{w}) = f(A) + \vec{w} = A + \vec{u}' + \vec{w} = A + \vec{w} + \vec{u}' \in A + F = \mathcal{F}'.$$

Ceci montre que  $\mathcal{F}'$  est stable par  $f$  et que la restriction de  $f$  à  $\mathcal{F}'$  est la translation  $t_{\vec{u}'}$ . Or, d'après ce qui précède, on a  $\mathcal{F}' = \mathcal{F}$  et donc  $\vec{u}' = \vec{u}$ , puis  $g' = t_{\vec{u}} \circ f = g$ . Le théorème est démontré.  $\square$



**Définition 3.1.10 (Repères orthonormés).** — Soit  $(\mathcal{E}, E)$  un espace affine euclidien. Un repère  $\mathcal{R} = (O, \mathcal{B})$  de  $\mathcal{E}$  est dit **orthonormé** si  $\mathcal{B}$  est une base orthonormée de  $E$ .

### 3.2. Classification des isométries en petite dimension



**3.2.1. Classification des isométries affines du plan.** — Soit  $(\mathcal{P}, E)$  un plan affine euclidien, on oriente  $E$  en choisissant une base orthonormée  $\mathcal{B} = (e_1, e_2)$ . En utilisant le théorème précédent, on va déterminer toutes les isométries affines  $f$  de  $\mathcal{P}$ . On rappelle que si  $f$  admet un point fixe  $O$ , alors via la bijection  $\mathcal{P} \xrightarrow{\sim} E, P \mapsto \overrightarrow{OP}$ ,  $f$  s'identifie à sa partie vectorielle  $\overrightarrow{f}$ , cf. le point (2) du théorème 3.3.12 et la discussion précédant 3.3.42.

Soit d'abord  $f \in \text{Is}^+(\mathcal{P})$ , alors il existe un unique  $\theta \in [0, 2\pi[$  tel que  $\overrightarrow{f}$  soit la rotation vectorielle d'angle  $\theta$ , i.e.

$$\text{Mat}_{\mathcal{B}}(\overrightarrow{f}) = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}.$$

Si  $\overrightarrow{f} = \text{id}_E$ , i.e. si  $\theta = 0$ , alors  $f$  est une **translation**  $t_{\vec{u}}$ , avec  $\vec{u} \in E$ . Sinon, si  $\theta \in ]0, 2\pi[$ , alors  $\text{Fix}(\overrightarrow{f}) = \{0\}$  donc  $f$  admet un point fixe  $I$  unique, alors  $f$  est la **rotation de centre  $I$  et d'angle  $\theta$** .

Soit maintenant  $f \in \text{Is}^-(\mathcal{P})$ , alors  $\overrightarrow{f}$  est la symétrie orthogonale par rapport à une droite vectorielle  $D$ . Si  $f$  possède un point fixe  $I$ , soit  $\mathcal{D}$  la droite affine  $I + D$ , alors  $f$  est la symétrie orthogonale  $\sigma_{\mathcal{D}}$  par rapport à  $\mathcal{D}$ . Si  $f$  ne possède pas de point fixe alors, d'après le théorème précédent, il existe une unique droite affine  $\mathcal{D}$  de direction  $D$  stable par  $f$ , et pour tout  $I \in \mathcal{D}$  on a  $f(I) = I + \vec{u}$  pour un certain  $\vec{u} \in D$ ,  $\vec{u} \neq 0$ . Alors  $f = t_{\vec{u}} \circ \sigma_{\mathcal{D}} = \sigma_{\mathcal{D}} \circ t_{\vec{u}}$ , et l'on dit que  $f$  est la **symétrie (orthogonale) glissée** par rapport à  $\mathcal{D}$ , de vecteur  $\vec{u}$  : pour tout point  $M \in \mathcal{P}$ , son image  $M' = f(M)$  s'obtient en effectuant la symétrie  $M \mapsto \sigma_{\mathcal{D}}(M)$ , puis en faisant un « glissement » de vecteur  $\vec{u}$  (dans la direction de la droite  $\mathcal{D}$ ), ou bien en faisant d'abord le « glissement »  $M \mapsto M + \vec{u}$ , puis en prenant le symétrique  $\sigma_{\mathcal{D}}(M + \vec{u})$ .



On peut résumer la classification par le tableau suivant :

Isométries $f$ de $\mathcal{P}$	directes	indirectes
$\text{Fix}(f) = \mathcal{P}$	$\text{id}_{\mathcal{P}}$	
$\text{Fix}(f) = \text{droite } \mathcal{D}$		symétrie orthogonale $\sigma_{\mathcal{D}}$
$\text{Fix}(f) = \text{point } I$	rotations de centre $I$ et d'angle $\theta \notin 2\pi\mathbb{Z}$	
$\text{Fix}(f) = \emptyset$	translations	symétries orthogonales glissées



**3.2.2. Classification des isométries affines de l'espace.** — Soit  $(\mathcal{E}, E)$  un espace affine euclidien de dimension 3, on oriente  $E$  en choisissant une base orthonormée  $\mathcal{B}_0 = (e_1, e_2, e_3)$ . En utilisant le théorème 3.1.9 et la classification des isométries vectorielles de  $E$  (cf. 2.4.18), on va déterminer toutes les isométries affines  $f$  de  $\mathcal{E}$ .

**1er cas.** Soit d'abord  $f \in \text{Is}^+(\mathcal{E})$ . Si  $\overrightarrow{f} = \text{id}_E$ , alors  $f$  est une **translation**  $t_{\vec{u}}$ , avec  $\vec{u} \in E$ . Sinon,  $\overrightarrow{f}$  est une rotation :  $\text{Fix}(f)$  est une droite vectorielle  $D$ , choisissons un vecteur directeur  $\vec{w}$  de  $D$ , alors il existe un unique  $\theta \in ]0, 2\pi[$  tel que  $\overrightarrow{f}$  soit la **rotation d'axe orienté par  $\vec{w}$  et d'angle  $\theta$** , c.-à-d.,

pour toute base orthonormée  $(\vec{v}_1, \vec{v}_2)$  du plan  $D^\perp$  telle que la base  $(\vec{v}_1, \vec{v}_2, \vec{w})$  soit directe (i.e. telle que  $\det_{\mathcal{B}_0}(\vec{v}_1, \vec{v}_2, \vec{w}) > 0$ ), on a :

$$\text{Mat}_{(\vec{v}_1, \vec{v}_2, \vec{w})}(\vec{f}) = \begin{pmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Si  $f$  a un point fixe  $I$ , soit  $\mathcal{D}$  la droite affine  $I + D$ , alors  $f$  est la **rotation d'axe  $\mathcal{D}$  orienté par  $\vec{w}$  et d'angle  $\theta$**  (et l'on a  $\text{Fix}(f) = \mathcal{D}$ ).

Si  $f$  n'a pas de point fixe alors, d'après le théorème 3.1.9, il existe une unique droite affine  $\mathcal{D}$  de direction  $D$  stable par  $f$ , et pour tout  $I \in \mathcal{D}$  on a  $f(I) = I + \vec{u}$  pour un certain  $\vec{u} \in D$ ,  $\vec{u} \neq 0$ . On peut alors prendre  $\vec{u}$  comme vecteur directeur de  $D$ , alors  $\vec{f}$  est la rotation d'axe orienté par  $\vec{u}$  et d'angle  $\theta'$ , où  $\theta' = \theta$  si  $\vec{u}$  et le vecteur  $\vec{w}$  précédent sont « de même sens », i.e. si  $\vec{u} = \lambda \vec{w}$  avec  $\lambda > 0$ , et  $\theta' = -\theta$  si  $\vec{u}$  et  $\vec{w}$  sont « de sens opposé », i.e. si  $\vec{u} = -\lambda \vec{w}$  avec  $\lambda > 0$ . On dit alors que  $f$  est le **vissage de vecteur  $\vec{u}$ , d'axe  $\mathcal{D}$  orienté par  $\vec{u}$ , et d'angle  $\theta'$** .

**2ème cas.** Soit maintenant  $f \in \text{Is}^-(\mathcal{E})$ , alors  $\vec{f}$  est de l'un des types suivants :

- (1)  $\vec{f} = -\text{id}_E$ ,
- (2)  $\vec{f}$  est une **rotation gauche**, c.-à-d.,  $D = \{\vec{v} \in E \mid \vec{f}(\vec{v}) = -\vec{v}\}$  est une droite vectorielle, et la restriction de  $\vec{f}$  au plan  $D^\perp$  est une rotation distincte de l'identité.
- (3)  $\vec{f}$  est la **symétrie orthogonale**  $\sigma_P$  par rapport à un plan  $P$ .

Dans les cas (1) et (2),  $\text{Fix}(\vec{f}) = \{0\}$  donc  $f$  a un point fixe  $I$  unique ; dans le cas (1),  $f$  est la **symétrie centrale**  $\sigma_I$  de centre  $I$ , dans le cas (2),  $f$  est une **rotation gauche de centre  $I$  et d'axe  $\mathcal{D} = I + D$** .

Supposons que  $\vec{f} = \sigma_P$ . Si  $f$  a un point fixe  $I$ , alors  $\text{Fix}(f)$  est le plan  $\mathcal{P} = I + P$  et  $f$  est la **symétrie orthogonale  $\sigma_{\mathcal{P}}$  par rapport à  $\mathcal{P}$** . Enfin, si  $f$  n'a pas de point fixe alors, d'après le théorème 3.1.9, il existe un unique plan affine  $\mathcal{P}$  de direction  $P$  stable par  $f$ , et pour tout  $I \in \mathcal{P}$  on a  $f(I) = I + \vec{u}$  pour un certain  $\vec{u} \in P$ ,  $\vec{u} \neq 0$ . Alors  $f = t_{\vec{u}} \circ \sigma_{\mathcal{P}} = \sigma_{\mathcal{P}} \circ t_{-\vec{u}}$ , et l'on dit que  $f$  est la **symétrie (orthogonale) glissée** par rapport à  $\mathcal{P}$ , de vecteur  $\vec{u}$ .

On peut résumer la classification par le tableau suivant :

Isométries $f$ de $\mathcal{E}$ ( $\dim \mathcal{E} = 3$ )	directes	indirectes
$\text{Fix}(f) = \mathcal{E}$	$\text{id}_{\mathcal{E}}$	
$\text{Fix}(f) = \text{plan } \mathcal{P}$		symétrie orthogonale $\sigma_{\mathcal{P}}$
$\text{Fix}(f) = \text{droite } \mathcal{D}$	rotations d'axe $\mathcal{D}$ et d'angle $\theta \notin 2\pi\mathbb{Z}$	
$\text{Fix}(f) = \text{point}$		symétries centrales et rotations gauches d'angle $\theta \notin \pi\mathbb{Z}$
$\text{Fix}(f) = \emptyset$	translations ( $\text{Fix}(\vec{f}) = E$ ) vissages ( $\dim \text{Fix}(\vec{f}) = 1$ )	symétries orthogonales glissées

### 3.3. Appendice : Espaces affines réels

Cette appendice rassemble en manière synthétique les définitions et les résultats sur les espaces affines dont on a besoin. Attention : elle n'est pas faite pour apprendre les espaces affines, mais pour s'y référer en cas de besoin d'une définition ou de l'énoncé d'un théorème. Par exemple, cette section ne contient aucune preuve. On conseil au lecteur voulant apprendre la théorie des espaces affines ou cherchant les détails d'une preuve de se référer au poly du cours 2M270.

**Définition 3.3.1 (Espaces affines réels).** — Soit  $E$  un  $\mathbb{R}$ -espace vectoriel. Un **espace affine  $\mathcal{E}$  de direction  $E$**  est un ensemble non-vide  $\mathcal{E}$  muni d'une application

$$\mathcal{E} \times \mathcal{E} \rightarrow E, \quad (x, y) \mapsto \vec{xy}$$

vérifiant les deux propriétés suivantes :

$$(1) \quad \boxed{\text{Relation de Chasles : } \vec{xz} = \vec{xy} + \vec{yz}} \quad \forall x, y, z \in \mathcal{E}.$$

$$(2) \quad \text{L'application } \mathcal{E} \times \mathcal{E} \rightarrow \mathcal{E} \times E, \quad (x, y) \mapsto (x, \vec{xy}) \text{ est bijective.}$$

On notera  $(x, \vec{u}) \mapsto (x, x + \vec{u})$  la **bijection inverse**, c.-à-d., pour tout  $x \in \mathcal{E}$  et  $\vec{u} \in E$ ,  $x + \vec{u}$  désigne l'unique  $y \in \mathcal{E}$  tel que  $\vec{xy} = \vec{u}$ .

**Vocabulaire** : les éléments de  $\mathcal{E}$  sont appelés « points », ceux de  $E$  sont appelés « vecteurs ». Si  $E$  est de dimension finie  $n$ , ce que nous supposons par la suite, on pose  $\dim \mathcal{E} = \dim E$ .

**Exemple 3.3.2.** — Tout  $\mathbb{R}$ -espace vectoriel  $E$  est un espace affine, pour l'application  $E \times E \rightarrow E$ ,  $(x, y) \mapsto \overrightarrow{xy} = y - x$ .

**Définition 3.3.3 (Repères cartésiens d'un espace affine de dimension finie)**

Soit  $(\mathcal{E}, E)$  un espace affine, avec  $\dim E = n$ . Un **repère** (cartésien)  $\mathcal{R}$  de  $\mathcal{E}$  est un couple  $(O, \mathcal{B})$ , où  $O$  est un point de  $\mathcal{E}$  et  $\mathcal{B} = (e_1, \dots, e_n)$  est une base de  $E$ ; on dira aussi que le  $(n+1)$ -uplet  $(O, e_1, \dots, e_n)$  est un repère de  $\mathcal{E}$ . Dans ce cas, pour tout point  $P$  de  $\mathcal{E}$ , il existe une unique  $n$ -uplet  $(x_1, \dots, x_n) \in \mathbb{R}^n$  tel que

$$\overrightarrow{OP} = x_1 e_1 + \dots + x_n e_n$$

et l'on dit que  $(x_1, \dots, x_n)$  sont les **coordonnées** de  $P$  dans le repère  $\mathcal{R}$ .

**Exemple 3.3.4.** — Soit  $\mathcal{B} = (e_1, \dots, e_n)$  la base canonique de l'espace vectoriel  $E = \mathbb{R}^n$ , posons

$$\mathcal{E} = \mathbb{R}^n = \{(x_1, \dots, x_n) \mid \forall i = 1, \dots, n, x_i \in \mathbb{R}\}$$

considéré comme espace affine de direction  $E = \mathbb{R}^n$ , c.-à-d., muni de l'application  $\mathcal{E} \times \mathcal{E} \rightarrow E$ ,  $(x, y) \mapsto \overrightarrow{xy} = y - x = \sum_{i=1}^n (y_i - x_i) e_i$ . Notons  $O$  le point  $(0, \dots, 0)$  de  $\mathcal{E}$  et  $\mathcal{R}$  le repère  $(O, \mathcal{B})$ . Alors pour tout point  $P = (x_1, \dots, x_n)$  de  $\mathcal{E}$ , ses coordonnées dans le repère  $\mathcal{R}$  sont  $(x_1, \dots, x_n)$ . Si on fixe un point arbitraire  $A = (a_1, \dots, a_n)$  de  $\mathcal{E}$  et qu'on considère le repère  $\mathcal{R}' = (A, \mathcal{B})$ , alors les coordonnées  $(x'_1, \dots, x'_n)$  de  $P$  dans  $\mathcal{R}'$  sont données par l'égalité  $\overrightarrow{AP} = x'_1 e_1 + \dots + x'_n e_n$ , et comme par définition  $\overrightarrow{AP} = \sum_{i=1}^n (x_i - a_i) e_i$ , on obtient  $x'_i = x_i - a_i$  pour tout  $i$ .

Ce qui précède est un cas particulier du théorème suivant :

**Théorème 3.3.5 (Changement de repère).** — Soit  $(\mathcal{E}, E)$  un espace affine de dimension  $n$ ,  $\mathcal{B} = (e_1, \dots, e_n)$  et  $\mathcal{C} = (f_1, \dots, f_n)$  deux bases de  $E$ , soit  $P = \text{Mat}_{\mathcal{B}}(\mathcal{C})$  la matrice de passage, et soient  $O, O' \in \mathcal{E}$ . Posons  $\overrightarrow{OO'} = t_1 e_1 + \dots + t_n e_n$ . Pour  $M \in \mathcal{E}$  arbitraire, notons  $(x_1, \dots, x_n)$  (resp.  $(x'_1, \dots, x'_n)$ ) ses coordonnées dans le repère  $\mathcal{R} = (O, \mathcal{B})$  (resp.  $\mathcal{R}' = (O', \mathcal{C})$ ). Alors on a

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = P \begin{pmatrix} x'_1 \\ \vdots \\ x'_n \end{pmatrix} + \begin{pmatrix} t_1 \\ \vdots \\ t_n \end{pmatrix}, \quad \text{et donc} \quad \begin{pmatrix} x'_1 \\ \vdots \\ x'_n \end{pmatrix} = P^{-1} \begin{pmatrix} x_1 - t_1 \\ \vdots \\ x_n - t_n \end{pmatrix}.$$

De façon abrégée, si l'on note  $X, X'$  et  $T$  les vecteurs colonnes ci-dessus, on a

$$\boxed{X = PX' + T} \quad \boxed{X' = P^{-1}(X - T)}$$

**Définition et proposition 3.3.6 (Applications affines).** — Soient  $(\mathcal{E}, E)$  et  $(\mathcal{E}', E')$  deux espaces affines. Une application  $f : \mathcal{E} \rightarrow \mathcal{E}'$  est une **application affine** s'il existe un point  $O \in \mathcal{E}$  tel que l'application  $\phi : E \rightarrow E'$  définie par :

$$\forall P \in \mathcal{E}, \quad \phi(\overrightarrow{OP}) = \overrightarrow{f(O)f(P)}$$

soit **linéaire**. Dans ce cas, on note  $\phi = \overrightarrow{f}$ , et les égalités ci-dessus sont vraies pour tout couple de points  $(I, P)$ , c.-à-d., on a **pour tout**  $I \in \mathcal{E}$  :

$$\boxed{\forall P \in \mathcal{E}, \quad \overrightarrow{f(I)f(P)} = \overrightarrow{f}(\overrightarrow{IP})} \quad \text{et} \quad \boxed{\forall \vec{u} \in E, \quad f(I + \vec{u}) = f(I) + \overrightarrow{f}(\vec{u}).}$$

On dit alors que  $\overrightarrow{f}$  est la **partie linéaire** de  $f$ .

**Remarque 3.3.7.** — Pour  $f : \mathcal{E} \rightarrow \mathcal{E}'$  arbitraire, et  $O \in \mathcal{E}$ , l'application  $\phi : E \rightarrow E'$  définie plus haut n'a aucune raison d'être linéaire : prendre par exemple  $\mathcal{E} = E = \mathbb{R}$ ,  $O = 0$  et  $f(t) = t^2$  pour tout  $t \in \mathbb{R}$ , alors  $\phi(t) = f(t)$  n'est pas linéaire.

**Proposition 3.3.8 (Écriture d'une application affine dans des repères cartésiens)**

Soient  $(\mathcal{E}, E)$  et  $(\mathcal{E}', E')$  deux espaces affines, avec  $\dim E = m$  et  $\dim E' = n$ , et soit  $f : \mathcal{E} \rightarrow \mathcal{E}'$  une application affine. Soient  $\mathcal{R} = (O, \mathcal{B})$  et  $\mathcal{R}' = (O', \mathcal{C})$  des repères de  $\mathcal{E}$  et  $\mathcal{E}'$ , soit  $A = \text{Mat}_{\mathcal{C}}(\overrightarrow{f}) \in M_{n,m}(\mathbb{R})$  et soient  $(b'_1, \dots, b'_n)$  les coordonnées de  $f(O)$  dans  $\mathcal{R}'$ .

Alors, pour tout  $P \in \mathcal{E}$ , de coordonnées  $(x_1, \dots, x_m)$  dans  $\mathcal{R}$ , les coordonnées  $(y_1, \dots, y_n)$  de  $f(P)$  dans  $\mathcal{R}'$  sont données par :

$$\boxed{\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = A \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix} + \begin{pmatrix} b'_1 \\ \vdots \\ b'_n \end{pmatrix}.$$

**Définition et proposition 3.3.9 (Translations de  $\mathcal{E}$ ).** — Pour tout  $\vec{u} \in E$ , on appelle **translation de vecteur**  $\vec{u}$  l'application  $t_{\vec{u}} : \mathcal{E} \rightarrow \mathcal{E}$ , qui à tout  $x \in \mathcal{E}$  associe le point  $x + \vec{u}$ . C'est une application **affine**, dont la partie linéaire est  $\text{id}_E$ .

**Lemme 3.3.10.** — Pour tout  $\vec{u}, \vec{v} \in E$ , on a  $\boxed{t_{\vec{u}} \circ t_{\vec{v}} = t_{\vec{u}+\vec{v}}}$ . Donc l'ensemble des translations est un groupe commutatif; en particulier, toute translation  $t_{\vec{u}}$  est **bijective, d'inverse**  $t_{-\vec{u}}$ .

**Proposition 3.3.11 (Composée d'applications affines).** — Soient  $\mathcal{E} \xrightarrow{f} \mathcal{E}' \xrightarrow{g} \mathcal{E}''$  deux applications affines, alors la composée  $g \circ f : \mathcal{E} \rightarrow \mathcal{E}''$  est affine, et sa partie linéaire est  $\boxed{\overrightarrow{(g \circ f)} = \overrightarrow{g} \circ \overrightarrow{f}}$ .

**Théorème et définition 3.3.12 (Transformations affines et groupe  $\text{GA}(\mathcal{E})$ )**

Soit  $(\mathcal{E}, E)$  un espace affine.

(1) On note  $\text{TAff}(\mathcal{E})$  l'ensemble des transformations affines de  $\mathcal{E}$ , i.e. applications affines  $\mathcal{E} \rightarrow \mathcal{E}$ ; il est stable par composition, c.-à-d., si  $f, g \in \text{TAff}(\mathcal{E})$  alors  $g \circ f \in \text{TAff}(\mathcal{E})$ .

(2) Pour tout  $O \in \mathcal{E}$ , tout  $f \in \text{TAff}(\mathcal{E})$  s'écrit  $f = t_{\overrightarrow{O f(O)}} \circ \phi_O$ , où  $\phi_O$  est l'application affine définie par  $\phi_O(P) = O + \overrightarrow{f}(\overrightarrow{OP})$ .

(3) Soit  $f \in \text{TAff}(\mathcal{E})$ , alors  $\boxed{f \text{ bijective} \iff \overrightarrow{f} \text{ bijective}}$ .

(4) On note  $\text{GA}(\mathcal{E}) = \{f \in \text{TAff}(\mathcal{E}) \mid f \text{ bijective}\}$ ; c'est un **groupe**, et l'application  $\text{GA}(\mathcal{E}) \rightarrow \text{GL}(E)$ ,  $f \mapsto \overrightarrow{f}$  est un morphisme de groupes.

(5) Pour tout  $f \in \text{TAff}(\mathcal{E})$  et  $\vec{u} \in E$ , on a  $f \circ t_{\vec{u}} = t_{\overrightarrow{f}(\vec{u})} \circ f$ ; en particulier si  $f \in \text{GA}(\mathcal{E})$  alors  $f \circ t_{\vec{u}} \circ f^{-1} = t_{\overrightarrow{f}(\vec{u})}$ .

**Remarque 3.3.13.** — Soient  $(\mathcal{E}, E)$  et  $(\mathcal{E}', E')$  deux espaces affines de dimension finie, et soit  $f : \mathcal{E} \rightarrow \mathcal{E}'$  une application affine bijective. Alors  $\overrightarrow{f} : E \rightarrow E'$  est une application linéaire bijective, donc  $E$  et  $E'$  sont isomorphes; en particulier ils ont même dimension, donc  $\dim \mathcal{E} = \dim \mathcal{E}'$ .

**3.3.14. Barycentres et sous-espaces affines.** —

**Théorème et définition 3.3.15 (Barycentres).** — Soit  $\mathcal{E}$  un espace affine, soient  $P_1, \dots, P_n \in \mathcal{E}$  et  $t_1, \dots, t_n$  des réels tels que  $\boxed{t_1 + \dots + t_n = 1}$ . Alors il existe un unique point  $G \in \mathcal{E}$  tel que :

$$(*) \quad \boxed{\forall I \in \mathcal{E}, \quad \overrightarrow{IG} = t_1 \overrightarrow{IP_1} + \dots + t_n \overrightarrow{IP_n}}$$

et  $G$  s'appelle le **barycentre** des points pondérés  $(P_i, t_i)$  (i.e. chaque point  $P_i$  est affecté du « poids »  $t_i$ ). On notera :

$$(**) \quad \boxed{G = t_1 P_1 + \dots + t_n P_n.}$$

Lorsque  $t_1 = \dots = t_n$ , d'où  $t_1 = \dots = t_n = 1/n$ , on dit que  $G$  est le **centre de gravité** ou **isobarycentre** des points  $P_1, \dots, P_n$ .

**Remarque 3.3.15.1.** — La démonstration du théorème montre que si un point  $G'$  vérifie l'égalité (\*) pour un point  $I$ , alors  $G'$  égale  $G$  et vérifie l'égalité (\*) pour tout point  $I$ .

**Définition 3.3.16 (Segments).** — Soient  $A, B \in \mathcal{E}$ , on définit le **segment**  $[A, B]$  comme l'ensemble des points qui sont « entre »  $A$  et  $B$ , i.e. l'ensemble des points  $P$  de la forme

$$P = A + t \overrightarrow{AB} = B + (1-t) \overrightarrow{BA}, \quad \text{avec } t \in [0, 1],$$

c'est donc aussi l'ensemble des barycentres  $P = (1-t)A + tB$ , avec  $t \in [0, 1]$ . En particulier, le *centre de gravité* de  $A, B$  (qui correspond à  $t = 1/2$ ) est le **milieu** du segment  $[A, B]$ .

**Proposition 3.3.17 (Associativité des barycentres).** — Soient  $A_1, \dots, A_n \in \mathcal{E}$ ,  $t_1, \dots, t_n \in \mathbb{R}$  tels que  $t_1 + \dots + t_n = 1$ , et  $G$  le barycentre des points pondérés  $(A_1, t_1), \dots, (A_n, t_n)$ .

(1) On suppose que  $\mu = t_1 + \dots + t_{n-1} = 1 - t_n$  est non nul. Soit  $H$  le barycentre des points pondérés  $(A_1, t_1/\mu), \dots, (A_{n-1}, t_{n-1}/\mu)$ , alors  $\boxed{G \text{ est le barycentre de } (H, \mu) \text{ et de } (A_n, t_n)}$ .

(2) En particulier, si  $t_1 = \dots = t_n = 1/n$ , soit  $G$  (resp.  $H$ ) le centre de gravité  $G$  de  $A_1, \dots, A_n$  (resp. de  $A_1, \dots, A_{n-1}$ ), alors  $G$  est le barycentre de  $(H, (n-1)/n)$  et de  $(A_n, 1/n)$ .

**Exemple 3.3.18 (Centre de gravité d'un triangle).** — Dans le plan affine  $\mathcal{P}$ , soient  $A, B, C$  trois points non alignés, soit  $G$  le centre de gravité des points  $A, B, C$ , et soit  $O \in \mathcal{P}$  arbitraire. Notons  $I$  le milieu du segment  $[BC]$ , alors la droite  $(AI)$  s'appelle la **médiane** du triangle **issue de**  $A$ . Appliquons la proposition précédente à  $A_1 = B, A_2 = C, A_3 = A$ , alors on a  $\mu = 2/3$  et  $t_i/\mu = 1/2$  pour  $i = 1, 2$ , donc  $H = I$  et l'on a :

$$\overrightarrow{OG} = \frac{2}{3} \overrightarrow{OI} + \frac{1}{3} \overrightarrow{OA}.$$

Prenant  $O = G$ , on obtient  $\overrightarrow{GA} = -2\overrightarrow{GI}$  : ceci montre que  $G$  est situé sur le segment  $[AI]$ , aux deux-tiers du segment en partant du sommet  $A$ . On a le même résultat si l'on introduit les milieux  $J$  et  $K$  de  $[BC]$  et  $[CA]$ , donc on obtient le résultat suivant : « les médianes d'un triangle se coupent aux deux-tiers de leur longueur (en partant des sommets), et le point d'intersection est le centre de gravité du triangle ».

**Théorème 3.3.19 (Applications affines et barycentres).** — Soit  $f : \mathcal{E} \rightarrow \mathcal{E}'$  une application affine. Alors  $f$  préserve les barycentres : si  $G \in \mathcal{E}$  est le barycentre des points  $(A_i, t_i)$  alors  $f(G) \in \mathcal{E}'$  est le barycentre des points  $(f(A_i), t_i)$ .

**Définition et proposition 3.3.20 (Sous-espaces affines).** — Soient  $(\mathcal{E}, E)$  un espace affine et  $\mathcal{F}$  un sous-ensemble non-vidé. Les conditions suivantes sont équivalentes :

(1) Pour tous  $M_0, M_1, \dots, M_p \in \mathcal{F}$  et tous  $t_0, t_1, \dots, t_p \in \mathbb{R}$  tels que  $t_0 + \dots + t_p = 1$ , le barycentre  $G = t_0 M_0 + \dots + t_p M_p$  appartient à  $\mathcal{F}$ .

(2) Pour tous  $M_0, M_1, \dots, M_p \in \mathcal{F}$  et tous  $\lambda_1, \dots, \lambda_p \in \mathbb{R}$ , le point  $M_0 + \lambda_1 \overrightarrow{M_0 M_1} + \dots + \lambda_p \overrightarrow{M_0 M_p}$  appartient à  $\mathcal{F}$ .

(3) Il existe  $M_0 \in \mathcal{F}$  tel que  $F = \{\overrightarrow{M_0 M} \mid M \in \mathcal{F}\}$  soit un sous-espace vectoriel de  $E$ .

(4) Pour tout  $M_0 \in \mathcal{F}$ ,  $F = \{\overrightarrow{M_0 M} \mid M \in \mathcal{F}\}$  est un sous-espace vectoriel de  $E$ , indépendant du choix de  $M_0 \in \mathcal{F}$ .

Lorsque ces conditions sont vérifiées, on dit que  $\mathcal{F}$  est un **sous-espace affine** (en abrégé : **sea**), de direction  $F$ , et l'on pose  $\dim \mathcal{F} = \dim F$ .

**Remarque 3.3.21.** — Soient  $(\mathcal{F}, F)$  comme dans la proposition précédente, alors  $\mathcal{F}$  est un espace affine de direction  $F$ , ce qui justifie la terminologie « sous-espace affine de  $\mathcal{E}$  de direction  $F$  ». En effet, l'application

$$\mathcal{F} \times \mathcal{F} \rightarrow F, \quad (M_0, M) \mapsto \overrightarrow{M_0 M}$$

est la restriction à  $\mathcal{F} \times \mathcal{F}$  de l'application  $\mathcal{E} \times \mathcal{E} \rightarrow E$ , donc vérifie la relation de Chasles. D'autre part, d'après le point (4) de la proposition, pour tout  $M_0 \in \mathcal{F}$ , l'application  $\mathcal{F} \rightarrow F, M \mapsto \overrightarrow{M_0 M}$  est bijective. Ceci prouve que  $\mathcal{F}$  est bien un espace affine de direction  $F$  (cf. définition 3.3.1).

**Définition 3.3.22 (Sea de direction  $F$  passant par un point  $A$ ).** — Soient  $(\mathcal{E}, E)$  un espace affine et  $F$  un sev de  $E$ . Pour tout  $A \in \mathcal{E}$ , il existe un unique sea de direction  $F$  passant par  $A$ , c'est :

$$\mathcal{F} = \{P \in \mathcal{E} \mid \overrightarrow{AP} \in F\} = \{A + \vec{u} \mid \vec{u} \in F\}$$

on le note  $A + F$ . De plus, pour tout point  $B \in \mathcal{F}$ , on a aussi  $\mathcal{F} = B + F$ .

**Exemples 3.3.23.** — (1) Pour tout  $P \in \mathcal{E}$ , le singleton  $\{P\}$  est un sea de  $\mathcal{E}$  de dimension 0, et réciproquement.

(2) Soit  $\mathcal{D}$  un sea de  $\mathcal{E}$  de dimension 1, et soit  $D = \mathbb{R}\vec{u}$  sa direction (une droite vectorielle dans  $E$ ). Pour tout  $A$  fixé dans  $\mathcal{D}$ , on a :

$$\mathcal{D} = A + \mathbb{R}\vec{u} = \{A + t\vec{u} \mid t \in \mathbb{R}\} = \{(1-t)A + tB \mid t \in \mathbb{R}\}$$

où l'on a posé  $B = A + \vec{u}$ . Réciproquement, pour tout  $A \neq B$  dans  $\mathcal{E}$ , l'ensemble des barycentres :

$$\{(1-t)A + tB \mid t \in \mathbb{R}\}$$

est un sea de  $\mathcal{E}$  de direction  $\mathbb{R}\overrightarrow{AB}$ , qu'on appelle la « droite affine »  $(AB)$ .

**Proposition 3.3.24 (Sous-espaces affines définis par des équations).** — Soit  $(\mathcal{E}, E)$  un espace affine. On fixe  $O \in \mathcal{E}$ . Soient  $f_1, \dots, f_p \in E^*$  des formes linéaires sur  $E$  et  $c_1, \dots, c_p \in \mathbb{R}$ . On pose

$$\mathcal{F} = \{P \in \mathcal{E} \mid f_i(\overrightarrow{OP}) = c_i, \quad \forall i = 1, \dots, p\}.$$

Alors, si  $\mathcal{F} \neq \emptyset$ , c'est un sea de  $\mathcal{E}$ , de direction l'espace vectoriel

$$F = \{\vec{u} \in E \mid f_i(\vec{u}) = 0, \quad \forall i = 1, \dots, p\} = \bigcap_{i=1}^p \text{Ker}(f_i).$$

**Exemples 3.3.25.** — Soit  $(\mathcal{E}, E)$  un espace affine de dimension  $n$ . Fixons un repère  $(O, e_1, \dots, e_n)$  de  $\mathcal{E}$ , d'où des coordonnées  $(x_1, \dots, x_n)$ . Alors toute forme linéaire sur  $E$  est de la forme  $x \mapsto a_1 x_1 + \dots + a_n x_n$  et donc se donner des équations  $f_i(\overrightarrow{OM}) = c_i$  pour  $i = 1, \dots, p$  équivaut à se donner un système :

$$\begin{cases} a_{11}x_1 + \dots + a_{1n}x_n & = & c_1 \\ \dots & & \dots \\ a_{p1}x_1 + \dots + a_{pn}x_n & = & c_p \end{cases}$$

et la proposition précédente signifie la chose suivante : **si l'ensemble  $\mathcal{F}$  des solutions de ce système est non-vidé, alors c'est un sea de  $\mathcal{E}$  de direction l'espace vectoriel  $F$  formé des solutions du système homogène :**

$$\begin{cases} a_{11}x_1 + \dots + a_{1n}x_n & = & 0 \\ \dots & & \dots \\ a_{p1}x_1 + \dots + a_{pn}x_n & = & 0. \end{cases}$$

Illustrons ceci par les deux exemples suivants :

(1)  $\mathcal{F} = \{(x_1, x_2, x_3) \in \mathbb{R}^3 \mid x_1 + x_2 + x_3 = 1, \quad x_1 + 2x_2 = 2\} \neq \emptyset$  (il contient, par exemple, le point  $(0, 1, 0)$ ), c'est un sea de  $\mathbb{R}^3$  de direction la droite vectorielle

$$F = \{(x_1, x_2, x_3) \in \mathbb{R}^3 \mid x_1 + x_2 + x_3 = 0 = x_1 + 2x_2\}.$$

(2) Par contre,

$$\mathcal{F} = \{(x_1, x_2) \in \mathbb{R}^2 \mid x_1 + x_2 = 1, \quad 2x_1 + 2x_2 = 3\}$$

est vide!

De plus, dans un espace affine  $\mathcal{E}$  de dimension  $n$ , tout sous-espace affine  $\mathcal{F}$  de dimension  $r$  peut être défini par exactement  $n - r$  équations :

**Proposition 3.3.26.** — Soient  $(\mathcal{E}, E)$  un espace affine de dimension  $n$  et  $\mathcal{F}$  un sea de direction  $F$ ,  $\dim F = r$ . Soit  $O \in \mathcal{E}$  arbitraire et  $(e_1, \dots, e_r)$  une base de  $F$ , complétons-la en une base  $\mathcal{B} = (e_1, \dots, e_n)$  de  $E$  et notons  $(x_1, \dots, x_n)$  les coordonnées dans le repère  $\mathcal{R} = (O, \mathcal{B})$ . Alors il existe  $c_{r+1}, \dots, c_n \in \mathbb{R}$  tels que

$$\mathcal{F} = \{M(x_1, \dots, x_n) \in \mathcal{E} \mid x_i = c_i, \quad \forall i = r+1, \dots, n\}.$$

**Définition et proposition 3.3.27 (Sous-espace affine engendré par une partie non-vide)**

Soit  $(\mathcal{E}, E)$  un espace affine,  $X$  une partie non vide de  $\mathcal{E}$ . Alors l'ensemble  $\mathcal{F}$  de tous les barycentres

$$G = t_0 A_0 + \dots + t_p A_p, \quad \text{où } A_0, \dots, A_p \in X, \quad t_0, \dots, t_p \in \mathbb{R}, \quad t_0 + \dots + t_p = 1$$

est un sous-espace affine de  $\mathcal{E}$ , et c'est le **plus petit sea de  $\mathcal{E}$  contenant  $X$** . On l'appelle le **sea engendré par  $X$**  et on le note  $\text{Aff}\langle X \rangle$ . De plus, pour tout choix d'un point  $A_0 \in X$ , on a :

$$\begin{aligned} \text{Aff}\langle X \rangle &= \{A_0 + \lambda_1 \overrightarrow{A_0 A_1} + \dots + \lambda_p \overrightarrow{A_0 A_p} \mid A_1, \dots, A_p \in X, \lambda_1, \dots, \lambda_p \in \mathbb{R}\} \\ &= A_0 + \text{Vect}\{\overrightarrow{A_0 B} \mid B \in X\} \end{aligned}$$

**Corollaire 3.3.28 (Sea engendré par des points  $A_0, \dots, A_p$ ).** — Soit  $(\mathcal{E}, E)$  un espace affine, le sous-espace affine engendré par des points  $A_0, \dots, A_p \in \mathcal{E}$  est

$$\begin{aligned} \text{Aff}\langle A_0, \dots, A_p \rangle &= \{G = t_0 A_0 + \dots + t_p A_p \mid t_0, \dots, t_p \in \mathbb{R}, \quad t_0 + \dots + t_p = 1\} \\ &= A_0 + \text{Vect}\{\overrightarrow{A_0 A_1}, \dots, \overrightarrow{A_0 A_p}\} \end{aligned}$$

sa dimension est celle de l'espace vectoriel  $F = \text{Vect}\{\overrightarrow{A_0 A_1}, \dots, \overrightarrow{A_0 A_p}\}$ ; en particulier,  $\dim \text{Aff}\langle A_0, \dots, A_p \rangle = p \Leftrightarrow$  les vecteurs  $\overrightarrow{A_0 A_1}, \dots, \overrightarrow{A_0 A_p}$  sont linéairement indépendants.

**Remarque 3.3.29.** — Dans le corollaire précédent, on a choisi le point  $A_0$  comme « origine », mais bien sûr le même résultat est valable pour tout point  $A_i$ , i.e. on a aussi  $F = \text{Vect}\{\overrightarrow{A_i A_j} \mid j \neq i\}$  et  $\text{Aff}\langle A_0, \dots, A_p \rangle = A_i + F$ .

**Définition 3.3.30 (Points affinement indépendants ou liés).** — Soit  $(\mathcal{E}, E)$  un espace affine et soit  $p \in \mathbb{N}^*$ . On dit que  $(p+1)$  points  $A_0, \dots, A_p \in \mathcal{E}$  sont **affinement indépendants** si le sea de  $\mathcal{E}$  qu'ils engendrent (i.e. le plus petit sea de  $\mathcal{E}$  qui les contient) est de dimension  $p$ , c.-à-d., si les vecteurs  $\overrightarrow{A_0 A_1}, \dots, \overrightarrow{A_0 A_p}$  sont linéairement indépendants (ceci équivaut à dire que, pour  $i$  fixé, les  $p$  vecteurs  $\overrightarrow{A_i A_j}$ , pour  $j \neq i$ , sont linéairement indépendants). Dans le cas contraire, on dit que  $A_0, \dots, A_p \in \mathcal{E}$  sont **affinement liés**.

Si  $p = 2$ , les trois points  $A_0, A_1, A_2$  sont affinement liés  $\Leftrightarrow$  ils sont alignés. Donc  $A_0, A_1, A_2$  sont affinement indépendants  $\Leftrightarrow$  ils sont **non alignés**.

On dit que des points  $A_0, \dots, A_p \in \mathcal{E}$  sont **coplanaires** s'ils sont contenus dans un sea  $\mathcal{P}$  de dimension 2 (un plan affine), i.e. si  $\dim \text{Aff}\langle A_0, \dots, A_p \rangle \leq 2$ . Donc quatre points  $A_0, \dots, A_3$  sont affinement indépendants  $\Leftrightarrow$  ils sont **non coplanaires**.

**Définition 3.3.31 (Retour sur les repères).** — Soit  $(\mathcal{E}, E)$  un espace affine de dimension  $n$ . Se donner un repère  $\mathcal{R} = (A_0, \mathcal{B})$  de  $\mathcal{E}$  (où  $\mathcal{B}$  est une base de  $E$ ) équivaut à se donner un  $(n+1)$ -uplet de points  $A_0, A_1, \dots, A_n$  **affinement indépendants** : dans ce cas les  $n$  vecteurs  $\overrightarrow{A_0 A_1}, \dots, \overrightarrow{A_0 A_n}$  sont linéairement indépendants donc forment une base  $\mathcal{B}$  de  $E$ , et la donnée de  $(A_0, \mathcal{B})$  équivaut bien sûr à celle du  $(n+1)$ -uplet  $(A_0, A_1, \dots, A_n)$ . On dira aussi qu'un tel  $(n+1)$ -uplet de points affinement indépendants forme un repère de  $\mathcal{E}$ .

Attention, cette fois l'ordre des points est important :  $A_0$  est l'origine et la base  $\mathcal{B} = (\overrightarrow{A_0 A_1}, \dots, \overrightarrow{A_0 A_n})$  est ordonnée, i.e. les coordonnées  $(x_1, \dots, x_n)$  dans ce repère correspondent au point

$$A_0 + x_1 \overrightarrow{A_0 A_1} + \dots + x_n \overrightarrow{A_0 A_n} = \left(1 - \sum_{i=1}^n x_i\right) A_0 + x_1 A_1 + \dots + x_n A_n.$$

**Définitions 3.3.32 (Sea parallèles).** — Soit  $(\mathcal{E}, E)$  un espace affine et  $\mathcal{F}_1, \mathcal{F}_2$  deux sea, de directions respectives  $F_1$  et  $F_2$ .

(1) On dit que  $\mathcal{F}_1$  est *faiblement parallèle* à  $\mathcal{F}_2$  si l'on a  $F_1 \subseteq F_2$ . Dans ce cas, si  $A_0 \in \mathcal{F}_1 \cap \mathcal{F}_2$ , alors  $\mathcal{F}_1 = A_0 + F_1 \subseteq A_0 + F_2 = \mathcal{F}_2$ , donc on a l'alternative suivante : ou bien

$$\mathcal{F}_1 \subseteq \mathcal{F}_2 \quad \text{ou bien} \quad \mathcal{F}_1 \cap \mathcal{F}_2 = \emptyset.$$

(2) On dit que  $\mathcal{F}_1$  et  $\mathcal{F}_2$  sont **parallèles** si l'on a  $F_1 = F_2$ . Dans ce cas, on a l'alternative suivante : ou bien  $\mathcal{F}_1 = \mathcal{F}_2$  ou bien  $\mathcal{F}_1 \cap \mathcal{F}_2 = \emptyset$ .

**Remarque 3.3.32.1.** — Attention, deux sea  $\mathcal{F}_1, \mathcal{F}_2$  qui vérifient  $\mathcal{F}_1 \cap \mathcal{F}_2 = \emptyset$  ne sont **pas nécessairement** parallèles! Par exemple, dans  $\mathbb{R}^3$  les droites affines  $\mathcal{D}_1$  et  $\mathcal{D}_2$  d'équations respectives  $x = y = 0$  et  $x - 1 = 0 = z$  vérifient  $\mathcal{D}_1 \cap \mathcal{D}_2 = \emptyset$  mais ne sont pas parallèles : la direction  $D_1$  de  $\mathcal{D}_1$  (resp.  $D_2$  de  $\mathcal{D}_2$ ) est la droite vectorielle d'équation  $x = 0 = y$  (resp.  $x = 0 = z$ ) et l'on a  $D_1 \neq D_2$ .

**Proposition 3.3.33 (Sea engendré par deux sea  $\mathcal{F}$  et  $\mathcal{F}'$ ).** — Soit  $(\mathcal{E}, E)$  un espace affine, soient  $(\mathcal{F}, F), (\mathcal{F}', F')$  deux sous-espaces affines,  $P \in \mathcal{F}$  et  $P' \in \mathcal{F}'$ .

- (1) Le sous-espace vectoriel  $V = F + F' + \overrightarrow{\mathbb{R}PP'}$  est indépendant du choix de  $P \in \mathcal{F}$  et  $P' \in \mathcal{F}'$ .
- (2) On a  $V = F + F' \iff \mathcal{F} \cap \mathcal{F}' \neq \emptyset$ .
- (3) Le sous-espace affine engendré par  $\mathcal{F} \cup \mathcal{F}'$  est  $P + V = P' + V$ , on le notera  $\mathcal{F} + \mathcal{F}'$ .<sup>(1)</sup>
- (4) Par conséquent, si  $\mathcal{F}$  et  $\mathcal{F}'$  sont de dimension finie (par exemple, si  $\mathcal{E}$  l'est), on a :

$$\dim(\mathcal{F} + \mathcal{F}') = \begin{cases} \dim(F + F') & \text{si } \mathcal{F} \cap \mathcal{F}' \neq \emptyset, \\ \dim(F + F') + 1 & \text{si } \mathcal{F} \cap \mathcal{F}' = \emptyset. \end{cases}$$

**Remarque 3.3.34.** — Dans la proposition précédente, si  $\dim \mathcal{F} = p$ ,  $\dim \mathcal{F}' = r$  et si  $(P, A_1, \dots, A_p)$  et  $(P', B_1, \dots, B_r)$  sont des repères de  $\mathcal{F}$  et  $\mathcal{F}'$ , alors  $\mathcal{F} + \mathcal{F}' = \text{Aff}\langle P, A_1, \dots, A_p, P', B_1, \dots, B_r \rangle$ .

**Proposition 3.3.35 (Intersection de deux sea  $\mathcal{F}$  et  $\mathcal{F}'$ ).** — Soit  $(\mathcal{E}, E)$  un espace affine et soient  $(\mathcal{F}, F), (\mathcal{F}', F')$  deux sous-espaces affines. Si  $\mathcal{F} \cap \mathcal{F}'$  est non vide, c'est un sea de direction  $F \cap F'$ .

**Proposition 3.3.36 (Applications affines et sous-espaces).** — Soit  $f : \mathcal{E} \rightarrow \mathcal{E}'$  une application affine, notons  $\phi$  sa partie linéaire.

- (1) Soit  $(\mathcal{F}, F)$  un sea de  $\mathcal{E}$ , alors  $f(\mathcal{F})$  est un sea de  $\mathcal{E}'$ ; plus précisément, si  $P \in \mathcal{F}$  alors  $f(\mathcal{F}) = f(P) + \phi(F)$ .
- (2) Soit  $(\mathcal{F}', F')$  un sea de  $\mathcal{E}'$ . On a  $f^{-1}(\mathcal{F}') = \emptyset$  si  $f(\mathcal{E}) \cap \mathcal{F}' = \emptyset$ , sinon  $f^{-1}(\mathcal{F}')$  est un sea de  $\mathcal{E}$  de direction  $F \cap F'$ ; plus précisément, si  $P \in \mathcal{F}$  est tel que  $f(P) \in \mathcal{F}'$  alors

$$\begin{aligned} f^{-1}(\mathcal{F}') &= \{Q = P + \vec{u} \in \mathcal{E} \mid f(Q) = f(P) + \phi(\vec{u}) \in \mathcal{F}'\} \\ &= \{Q = P + \vec{u} \in \mathcal{E} \mid \overrightarrow{f(P)f(Q)} = \phi(\vec{u}) \in F'\} = P + \phi^{-1}(F'). \end{aligned}$$

### 3.3.37. Projections, symétries, points fixes. —

**Proposition 3.3.38.** — Soit  $(\mathcal{E}, E)$  un espace affine et soient  $(\mathcal{F}, F), (\mathcal{F}', F')$  deux sous-espaces affines tels que  $F + F' = E$ . Alors  $\mathcal{F} \cap \mathcal{F}'$  est non vide et est un sous-espace affine de direction  $F \cap F'$ .

**Corollaire 3.3.39 (Sous-espaces supplémentaires).** — Si  $E = F \oplus F'$ , i.e. si  $F$  et  $F'$  sont supplémentaires, alors pour tous sea  $\mathcal{F}$  de direction  $F$  et  $\mathcal{F}'$  de direction  $F'$ , l'intersection  $\mathcal{F} \cap \mathcal{F}'$  est un singleton  $\{P\}$ .

*Démonstration.* — D'après la proposition précédente,  $\mathcal{F} \cap \mathcal{F}'$  est non vide et est un sea de direction  $F \cap F' = \{0\}$ , donc c'est un singleton  $\{P\}$ .  $\square$

**Définition 3.3.40.** — Soient  $(\mathcal{E}, E)$  un espace affine et  $(\mathcal{F}, F)$  et  $(\mathcal{F}', F')$  deux sea tels que  $E = F \oplus F'$ . D'après le corollaire précédent,  $\mathcal{F} \cap \mathcal{F}'$  est un singleton  $\{O\}$  et, comme  $E = F \oplus F'$ , on a :

$$\forall M \in \mathcal{E}, \quad \exists!(P, P') \in \mathcal{F} \times \mathcal{F}' \quad \text{tel que} \quad \overrightarrow{OM} = \overrightarrow{OP} + \overrightarrow{OP'}.$$

On définit alors :

- (1) la **projection**  $\pi = \pi_{\mathcal{F}, \mathcal{F}'}$  **sur  $\mathcal{F}$  parallèlement à  $\mathcal{F}'$**  par  $\pi(M) = P$ ,
- (2) la **symétrie**  $s = s_{\mathcal{F}, \mathcal{F}'}$  **par rapport à  $\mathcal{F}$  parallèlement à  $\mathcal{F}'$**  par  $s(M) = M + 2\overrightarrow{MP} = M - 2\overrightarrow{OP'}$ , i.e.  $\overrightarrow{Ms(M)} = 2\overrightarrow{MP} = 2\overrightarrow{OP'}$ .

(Faire un dessin dans le plan affine  $\mathcal{P}$ , pour deux droites sécantes  $\mathcal{F}$  et  $\mathcal{F}'$ , non nécessairement orthogonales, par exemple  $\mathcal{F}$  d'équation  $y = 1$  et  $\mathcal{F}'$  d'équation  $y - x = 1$ .)

Si de plus  $\dim \mathcal{F} = \dim \mathcal{E} - 1$  (de sorte que  $\dim \mathcal{F}' = 1$ , i.e.  $\mathcal{F}'$  est une droite affine), on dit que  $\mathcal{F}$  est un **hyperplan** affine de  $\mathcal{E}$  et que  $s_{\mathcal{F}, \mathcal{F}'}$  est la **réflexion par rapport à l'hyperplan  $\mathcal{F}$  parallèlement à la droite  $\mathcal{F}'$** .

**Définition 3.3.41 (Points fixes).** — Si  $X$  est un ensemble et  $f$  une application  $X \rightarrow X$ , on dit que  $x \in X$  est un **point fixe** de  $f$  si  $f(x) = x$ . On notera  $\text{Fix}(f)$  l'ensemble des points fixes de  $f$ ; il peut être vide.

Soit  $(\mathcal{E}, E)$  un espace affine et  $f : \mathcal{E} \rightarrow \mathcal{E}$  une transformation affine. Si  $f$  possède un point fixe  $O$ , on a vu que, via la bijection  $\mathcal{E} \xrightarrow{\sim} E, P \mapsto \overrightarrow{OP}$ ,  $f$  s'identifie à sa partie linéaire  $\vec{f}$  (cf. point (2) du théorème 3.3.12). Donc, lorsque  $f$  possède un point fixe, l'étude de la transformation  $f$  se ramène à l'étude de l'endomorphisme  $\vec{f}$  de  $E$ , et l'on peut appliquer les résultats connus sur les endomorphismes. Ceci explique l'intérêt d'étudier l'ensemble des points fixes de  $f$ . On a la proposition suivante.

<sup>(1)</sup> Attention, cette notation n'est peut-être pas standard, peut-être les puristes s'en tiennent-ils à la notation  $\text{Aff}\langle \mathcal{F} \cup \mathcal{F}' \rangle \dots$

**Proposition 3.3.42 (Points fixes de  $f$ ).** — Soient  $(\mathcal{E}, E)$  un espace affine de dimension finie,  $f : \mathcal{E} \rightarrow \mathcal{E}$  une application affine, et  $\vec{f}$  sa partie linéaire.

- (1) Si  $\text{Fix}(f)$  est non vide, c'est un sea de direction  $\text{Fix}(\vec{f}) = \text{Ker}(\vec{f} - \text{id}_E)$ .
- (2) Si  $\text{Ker}(\vec{f} - \text{id}_E) = \{0\}$ , alors  $f$  a un point fixe unique  $I$ .



## CHAPITRE 4

### FORMES HERMITIENNES, ESPACES HILBERTIENS ET GROUPES UNITAIRES

**Résumé :** Ce chapitre est l'analogie du Chap. 6 quand on remplace le corps  $\mathbb{R}$  par  $\mathbb{C}$ . De façon simplifiée : « on remplace le carré  $x^2$  d'un nombre réel par le module au carré  $|z|^2 = z\bar{z}$  d'un nombre complexe ». Ceci conduit à la notion de *forme hermitienne* (analogue de la notion de forme bilinéaire symétrique), puis de *produit scalaire hilbertien* sur  $\mathbb{C}^n$  (analogue du produit scalaire euclidien sur  $\mathbb{R}^n$ ). On introduit alors le groupe  $U(n)$  des isométries de l'espace hilbertien  $\mathbb{C}^n$ , puis la notion d'endomorphisme *auto-adjoint* et, plus généralement, d'endomorphisme *normal*. Un des avantages de se placer sur le corps  $\mathbb{C}$  est l'existence de valeurs propres et vecteurs propres ; on obtient ainsi les importants théorèmes de diagonalisation 4.4.7 et 4.4.8.

On a indiqué par des symboles  $\diamond$  les définitions, exemples et résultats fondamentaux. Par ailleurs, des *compléments de cours*, pour les étudiants intéressés, sont donnés dans un appendice à la fin du chapitre (on y esquisse brièvement l'utilisation des « espaces de Hilbert » (de dimension infinie) en Analyse). Ces passages n'interviendront pas dans les évaluations.

#### 4.0. Rappels sur les nombres complexes

Dans tout ce chapitre, le corps de base est  $\mathbb{C}$ . On note  $i$  une racine carrée de  $-1$ , choisie une fois pour toutes. On rappelle les points suivants.

##### Définitions 4.0.1 (Parties réelle et imaginaire, conjugaison complexe, module et argument)

(1) Tout  $z \in \mathbb{C}$  s'écrit de façon unique  $z = a + ib$ , avec  $a, b \in \mathbb{R}$  ;  $a$  s'appelle la *partie réelle* de  $z$  et se note  $\mathcal{R}(z)$ ,  $b$  s'appelle la *partie imaginaire* de  $z$  et se note  $\mathcal{I}(z)$ . Remarquons que, comme  $-iz = b - ia$ , on a  $\mathcal{R}(-iz) = \mathcal{I}(z)$  et  $\mathcal{I}(iz) = \mathcal{R}(z)$ .

(2) La *conjugaison complexe* est l'application qui à tout  $z = a + ib$  associe  $\bar{z} = a - ib$ . Remarquons que :  $z + \bar{z} = 2\mathcal{R}(z)$ ,  $z - \bar{z} = 2i\mathcal{I}(z)$  et  $z = \bar{z} \Leftrightarrow z \in \mathbb{R}$ . D'autre part, on a  $\overline{z + z'} = \bar{z} + \bar{z}'$  et  $\overline{zz'} = \bar{z}\bar{z}'$ .

(3) Si  $z = a + ib$ , on a  $z\bar{z} = a^2 + b^2$  et  $|z| = \sqrt{z\bar{z}}$  s'appelle le *module* (ou la *norme*) de  $z$ . D'après la dernière égalité ci-dessus, la norme est *multiplicative*, i.e. on a  $|z_1 z_2| = |z_1| \cdot |z_2|$ .

(4) Enfin, si  $z \neq 0$ , alors  $z/|z|$  est de module 1, donc de la forme  $e^{i\theta}$ , avec  $\theta \in \mathbb{R}/2\pi\mathbb{Z}$  (cf. Chap. 5, Appendice 2.5). Donc tout  $z \neq 0$  s'écrit de façon unique :

$$z = \rho e^{i\theta}, \quad \text{avec } \rho = |z| \in \mathbb{R}_+^* \text{ et } \theta \in \mathbb{R} \text{ défini modulo } 2\pi\mathbb{Z}$$

(par exemple, on peut prendre  $\theta$  dans  $[0, 2\pi[$  ou bien dans  $]-\pi, \pi]$ ); on dit que  $\theta$  est l'*argument* de  $z$ .

#### 4.1. Formes hermitiennes

**Exemple 4.1.0.** — Le carré de la norme d'un nombre complexe  $z$  est la valeur en  $x = y = z$  de la fonction de deux variables  $\phi(x, y) = x\bar{y}$ . Cette fonction  $\phi : \mathbb{C} \times \mathbb{C} \rightarrow \mathbb{C}$  est linéaire en la 1ère variable :

$$\phi(\lambda x + \mu x', y) = (\lambda x + \mu x')\bar{y} = \lambda x\bar{y} + \mu x'\bar{y} = \lambda\phi(x, y) + \mu\phi(x', y)$$

mais pas tout-à-fait linéaire en la 2ème variable, puisqu'on a :

$$\phi(x, \lambda y + \mu y') = x\overline{\lambda y + \mu y'} = x\bar{\lambda y} + x\bar{\mu y'} = \bar{\lambda}\phi(x, y) + \bar{\mu}\phi(x', y).$$

D'autre part, on a  $\phi(y, x) = y\bar{x} = \overline{x\bar{y}} = \overline{\phi(x, y)}$ . Ceci conduit aux définitions suivantes.

**Définition 4.1.1 (Applications semi-linéaires).** — Soient  $E, F$  deux  $\mathbb{C}$ -espaces vectoriels. Une application  $f : E \rightarrow F$  est dite **semi-linéaire** si elle vérifie :

$$\forall u, v \in E, \quad \forall z \in \mathbb{C}, \quad \boxed{f(u+v) = f(u) + f(v)} \quad \boxed{f(zu) = \bar{z}f(u)}.$$

Ces deux conditions équivalent bien sûr à la condition :  $\boxed{f(zu+v) = \bar{z}f(u) + f(v)}$ .



**Définition 4.1.2 (Formes hermitiennes).** — Soit  $E$  un  $\mathbb{C}$ -espace vectoriel.

(1) Une **forme hermitienne** sur  $E$  est une application  $\phi : E \times E \rightarrow \mathbb{C}$  qui vérifie les deux conditions suivantes :

(a)  $\phi$  est linéaire en la 1ère variable et semi-linéaire en la 2ème variable, i.e. :

$$\forall x, x', y, y' \in E, \quad \forall \lambda \in \mathbb{C}, \quad \begin{cases} \phi(\lambda x + x', y) = \lambda\phi(x, y) + \phi(x', y), \\ \phi(x, \lambda y + y') = \bar{\lambda}\phi(x, y) + \phi(x, y') \end{cases}$$

(b)  $\phi$  a la propriété de « symétrie hermitienne » ci-dessous :

$$(*) \quad \forall x, y \in E, \quad \phi(y, x) = \overline{\phi(x, y)}.$$

(2) Observons que, pour tout  $x \in E$ , (\*) entraîne  $\phi(x, x) = \overline{\phi(x, x)}$  d'où  $\boxed{\phi(x, x) \in \mathbb{R}}$ .

(3) On note  $\boxed{\mathcal{H}(E)}$  l'ensemble des formes hermitiennes sur  $E$ ; si  $\phi, \psi \in \mathcal{H}(E)$  et  $s, t \in \mathbb{R}$ , on voit facilement que l'application  $s\phi + t\psi : E \times E \rightarrow \mathbb{C}$  définie par  $(s\phi + t\psi)(u, v) = s\phi(u, v) + t\psi(u, v)$  est encore une forme hermitienne. Par conséquent,  $\mathcal{H}(E)$  est un  $\boxed{\mathbb{R}$ -espace vectoriel (mais **pas** un  $\mathbb{C}$ -espace vectoriel, car si  $\lambda \in \mathbb{C} - \mathbb{R}$ , alors  $\lambda\phi$  ne vérifie plus (\*)).

(4) Remarquons enfin que, pour vérifier qu'une application  $\phi : E \times E \rightarrow \mathbb{C}$  est une forme hermitienne, il suffit de voir que  $\phi$  vérifie (\*) et est linéaire en la 1ère variable; ces deux conditions impliquent en effet la semi-linéarité en la 2ème variable, car :

$$\phi(x, \lambda y + y') = \overline{\phi(\lambda y + y', x)} = \overline{\lambda\phi(y, x) + \phi(y', x)} = \bar{\lambda}\overline{\phi(y, x)} + \overline{\phi(y', x)} = \bar{\lambda}\phi(x, y) + \phi(x, y').$$

**Remarque 4.1.2.1.** — La notion de forme hermitienne sur un  $\mathbb{C}$ -espace vectoriel est une « variante » de la notion de forme bilinéaire sur un  $\mathbb{R}$ -espace vectoriel.

**Définition 4.1.3 (Formes quadratiques hermitiennes).** — Soit  $\phi$  une forme hermitienne sur un  $\mathbb{C}$ -espace vectoriel  $E$ . On dit que l'application  $Q : E \rightarrow \mathbb{R}, x \mapsto \phi(x, x)$  est une **forme quadratique hermitienne** sur  $E$ . D'après le lemme qui suit,  $\phi$  est entièrement déterminée par  $Q$ , et l'on dit que  $\phi$  est la **forme polaire** de  $Q$ . Notons aussi que pour tout  $\lambda \in \mathbb{C}$ , on a

$$Q(\lambda x) = \phi(\lambda x, \lambda x) = \lambda\bar{\lambda}\phi(x, x) = \lambda\bar{\lambda}Q(x) = |\lambda|^2 Q(x).$$

**Lemme 4.1.4 (Polarisation).** — Soient  $E$  un  $\mathbb{C}$ -espace vectoriel,  $\phi \in \mathcal{H}(E)$  et  $Q$  l'application  $E \rightarrow \mathbb{R}, x \mapsto \phi(x, x)$ . Alors, pour tout  $x, y \in E$ , on a :

$$(1) \quad \mathcal{R}(\phi(x, y)) = \frac{1}{2}(Q(x+y) - Q(x) - Q(y)) = \frac{1}{4}(Q(x+y) - Q(x-y))$$

$$(2) \quad \mathcal{I}(\phi(x, y)) = \frac{1}{2}(Q(x+iy) - Q(x) - Q(y)) = \frac{1}{4}(Q(x+iy) - Q(x-iy))$$

$$(3) \quad 4\phi(x, y) = Q(x+y) - Q(x-y) + iQ(x+iy) - iQ(x-iy).$$

*Démonstration.* — On a

$$Q(x+y) = \phi(x+y, x+y) = Q(x) + Q(y) + \phi(x, y) + \phi(y, x)$$

$$Q(x-y) = \phi(x-y, x-y) = Q(x) + Q(y) - \phi(x, y) - \phi(y, x)$$

et comme  $\phi(x, y) + \phi(y, x) = \phi(x, y) + \overline{\phi(x, y)} = 2\mathcal{R}(\phi(x, y))$ , on obtient (1). Comme  $\mathcal{I}(z) = \mathcal{R}(-iz)$  pour tout  $z \in \mathbb{C}$ , on obtient

$$\mathcal{I}(\phi(x, y)) = \mathcal{R}(-i\phi(x, y)) = \mathcal{R}(\phi(x, iy))$$

et donc (2) s'obtient en remplaçant  $y$  par  $iy$  dans (1) et en utilisant que  $Q(iy) = |i|^2 Q(y) = Q(y)$ . Enfin, (3) découle de (1) et (2).  $\square$

Désormais, on suppose  $E$  de dimension finie  $n$ .

**Définition 4.1.5 (Matrices hermitiennes).** — Une matrice  $A \in M_n(\mathbb{C})$  est **hermitienne** si  $\boxed{{}^t A = \bar{A}}$ . On note  $\text{MH}_n(\mathbb{C})$  l'ensemble de ces matrices, si  $A, B \in \text{MH}_n(\mathbb{C})$  et  $s, t \in \mathbb{R}$ , alors  $sA + tB \in \text{MH}_n(\mathbb{C})$ , donc  $\text{MH}_n(\mathbb{C})$  est un  $\mathbb{R}$ -espace vectoriel (mais pas un  $\mathbb{C}$ -espace vectoriel). Observons que si  $A \in \text{MH}_n(\mathbb{C})$ , ses coefficients diagonaux  $a_{ii}$  vérifient  $a_{ii} = \bar{a}_{ii}$  donc  $\boxed{a_{ii} \in \mathbb{R}}$ .

**Remarque 4.1.5.1.** — On a  $\dim_{\mathbb{R}} \text{MH}_n(\mathbb{C}) = n^2$ . En effet, notons  $N$  le nombre de coefficients qui sont strictement au-dessus de la diagonale. C'est aussi le nombre de coefficients qui sont strictement en-dessous de la diagonale, et il y a  $n$  coefficients diagonaux. Donc  $2N + n = n^2$ , d'où  $2N = n^2 - n = n(n-1)$ . Puis, une matrice hermitienne est déterminée par le choix de  $n$  coefficients réels sur la diagonale et de  $N$  coefficients complexes au-dessus (ceux en-dessous en étant les conjugués), pour chaque coefficient complexe, il faut choisir sa partie réelle et sa partie imaginaire, d'où au total  $n + 2N = n^2$  coefficients réels.

**Définition et théorème 4.1.6 (Matrice d'une forme hermitienne et changement de base)**

Soit  $\phi$  une forme hermitienne sur un  $\mathbb{C}$ -espace vectoriel  $E$  de dimension  $n$  et soit  $\mathcal{B} = (e_1, \dots, e_n)$  une base de  $E$ .

(1) La matrice  $\text{Mat}_{\mathcal{B}}(\phi)$  de  $\phi$  dans la base  $\mathcal{B}$  est la matrice  $A = (a_{ij})_{i,j=1}^n \in M_n(\mathbb{C})$ , où  $a_{ij} = \phi(e_i, e_j)$ . Comme  $\phi(e_j, e_i) = \overline{\phi(e_i, e_j)}$ , on a  $a_{ji} = \overline{a_{ij}}$ , donc  ${}^tA = \overline{A}$ , i.e.  $A \in \text{MH}_n(\mathbb{C})$ .

(2)  $\phi$  est entièrement déterminée par sa matrice  $A$  : en effet, d'après la linéarité (resp. semi-linéarité) en la 1ère (resp. 2ème) variable, on a l'égalité :

$$(*) \quad \forall \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \in \mathbb{C}^n, \quad \phi \left( \sum_{i=1}^n x_i e_i, \sum_{j=1}^n y_j e_j \right) = \sum_{i,j=1}^n x_i \overline{y_j} \phi(e_i, e_j) = \sum_{i,j=1}^n a_{ij} x_i \overline{y_j}.$$

Donc, si l'on note  $X, Y$  les vecteurs colonnes ci-dessus, on a la formule matricielle  $\phi(X, Y) = {}^tX A \overline{Y}$ .

(3) Réciproquement, pour tout  $A = (a_{ij})_{i,j=1}^n \in \text{MH}_n(\mathbb{C})$ , l'application  $\phi_A : E \times E \rightarrow \mathbb{C}$  définie par  $\phi_A(\sum_{i=1}^n x_i e_i, \sum_{j=1}^n y_j e_j) = \sum_{i,j=1}^n a_{ij} x_i \overline{y_j}$  est une forme hermitienne sur  $E$ , et  $\text{Mat}_{\mathcal{B}}(\phi_A) = A$ . Donc, se donner une forme hermitienne sur  $E$  « est la même chose » que se donner une matrice hermitienne : de façon précise, l'application  $\mu_{\mathcal{B}} : \mathcal{H}(E) \rightarrow \text{MH}_n(\mathbb{C})$ ,  $\phi \mapsto \text{Mat}_{\mathcal{B}}(\phi)$  est un isomorphisme de  $\mathbb{R}$ -espaces vectoriels.

(4) Soient  $\mathcal{B}'$  une autre base de  $E$  et  $P$  la matrice de passage  $\text{Mat}_{\mathcal{B}}(\mathcal{B}')$ . Alors

$$(**) \quad A' = \text{Mat}_{\mathcal{B}'}(\phi) = {}^tP A \overline{P}.$$

*Démonstration.* — (2) Comme  $\phi$  est linéaire (resp. semi-linéaire) en la 1ère (resp. 2ème) variable, on a bien l'égalité (\*), qui montre que  $\phi$  est déterminée par sa matrice, donc que l'application  $\mu_{\mathcal{B}} : \mathcal{H}(E) \rightarrow \text{MH}_n(\mathbb{C})$ ,  $\phi \mapsto \text{Mat}_{\mathcal{B}}(\phi)$  est *injective*. D'autre part, on voit que le scalaire  $\sum_{i,j=1}^n a_{ij} x_i \overline{y_j} \in \mathbb{C}$  est égal au produit matriciel

$${}^tX A \overline{Y} = (x_1, \dots, x_n) A \begin{pmatrix} \overline{y_1} \\ \vdots \\ \overline{y_n} \end{pmatrix}.$$

Ceci prouve (2). Avant de prouver (3), remarquons déjà que l'application  $\mu_{\mathcal{B}}$  est  $\mathbb{R}$ -linéaire. En effet, si  $\phi, \psi \in \mathcal{H}(E)$  et  $s \in \mathbb{R}$ , alors  $s\phi + \psi$  est la forme hermitienne définie par  $(s\phi + \psi)(u, v) = s\phi(u, v) + \psi(u, v)$  pour tout  $u, v \in E$ , donc *a fortiori* on a  $(s\phi + \psi)(e_i, e_j) = s\phi(e_i, e_j) + \psi(e_i, e_j)$  pour tout  $i, j$ , d'où  $\mu_{\mathcal{B}}(s\phi + \psi) = s\mu_{\mathcal{B}}(\phi) + \mu_{\mathcal{B}}(\psi)$ .

Prouvons (3). Pour tout  $A = (a_{ij})_{i,j=1}^n \in \text{MH}_n(\mathbb{C})$ , l'application  $\phi_A : E \times E \rightarrow \mathbb{C}$  définie par

$$\phi_A \left( \sum_{i=1}^n x_i e_i, \sum_{j=1}^n y_j e_j \right) = \sum_{i,j=1}^n a_{ij} x_i \overline{y_j}$$

est linéaire en les  $x_i$ , et vérifie :

$$\phi_A(y, x) = \phi_A \left( \sum_{j=1}^n y_j e_j, \sum_{i=1}^n x_i e_i \right) = \sum_{i,j=1}^n \underbrace{a_{ji}}_{=\overline{a_{ij}}} y_j \overline{x_i} = \sum_{i,j=1}^n \overline{a_{ij} x_i \overline{y_j}} = \overline{\phi_A(x, y)}$$

donc est une forme hermitienne sur  $E$ ; de plus, prenant  $x_{i_0} = 1 = y_{j_0}$  et  $x_i = 0 = y_j$  pour  $i \neq i_0$  et  $j \neq j_0$ , on obtient que  $\phi_A(e_{i_0}, e_{j_0}) = a_{i_0, j_0}$  pour tout  $i_0, j_0 = 1, \dots, n$ , d'où  $\text{Mat}_{\mathcal{B}}(\phi_A) = A$ . Ceci montre que l'application  $\mathbb{R}$ -linéaire injective  $\mu_{\mathcal{B}} : \mathcal{H}(E) \rightarrow \text{MH}_n(\mathbb{C})$ ,  $\phi \mapsto \text{Mat}_{\mathcal{B}}(\phi)$  est aussi *surjective*, donc c'est un isomorphisme de  $\mathbb{R}$ -espaces vectoriels. En particulier, *se donner une forme hermitienne sur  $E$  « est la même chose » que se donner une matrice hermitienne.*

Enfin, démontrons (4). Soient  $x, y \in E$ , ils correspondent dans la base  $\mathcal{B}$  (resp.  $\mathcal{B}'$ ) à des vecteurs colonnes  $X, Y$  (resp.  $X', Y'$ ). D'après la formule de changement de coordonnées, on a  $X = PX'$  et  $Y = PY'$ , d'où  ${}^tX = {}^tX' {}^tP$  et  $\overline{Y} = \overline{P} \overline{Y'}$ , et donc :

$$\phi(x, y) = {}^tX A \overline{Y} = {}^tX' {}^tP A \overline{P} \overline{Y'}$$

ce qui entraîne  $A' = {}^tP A \overline{P}$ . Le théorème est démontré.  $\square$

**Définition 4.1.7 (Carrés de modules et « doubles produits »).** — En séparant, d'une part, les termes  $x_i \overline{y_i}$  et, d'autre part, les termes  $x_i \overline{y_j}$  avec  $i \neq j$ , la formule (\*) de 4.1.6 se réécrit de la façon suivante (puisque  $a_{ji} = \overline{a_{ij}}$  pour tout  $i \neq j$ ) :

$$(*) \quad \forall \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \in k^n, \quad \phi \left( \sum_{i=1}^n x_i e_i, \sum_{j=1}^n y_j e_j \right) = \sum_{i=1}^n a_{ii} x_i \overline{y_i} + \sum_{1 \leq i < j \leq n} (a_{ij} x_i \overline{y_j} + \overline{a_{ij}} x_j \overline{y_i}).$$

En particulier, prenant  $Y = X$  (i.e.  $y_i = x_i$  pour tout  $i$ ), on voit que la forme quadratique hermitienne  $Q$  associée à  $\phi$  est donnée par la formule suivante (noter que  $x_i \overline{x_i} = |x_i|^2$ ) :

$$(*) \quad Q(x_1, \dots, x_n) = \sum_{i=1}^n a_{ii} |x_i|^2 + \sum_{1 \leq i < j \leq n} (a_{ij} x_i \overline{x_j} + \overline{a_{ij}} \overline{x_i} x_j) = \sum_{i=1}^n a_{ii} |x_i|^2 + \sum_{1 \leq i < j \leq n} 2\Re(a_{ij} x_i \overline{x_j}).$$

On voit donc apparaître les carrés des modules des  $x_i$ , et les parties réelles des doubles produits  $x_i \overline{x_j}$ . Pour abrégé, on parlera de « carrés de modules » et de « doubles produits ».

**Définition et proposition 4.1.8 (Rang d'une forme hermitienne).** — Soit  $\phi$  une forme hermitienne sur un  $\mathbb{C}$ -espace vectoriel  $E$  de dimension  $n$  et soit  $\mathcal{B} = (e_1, \dots, e_n)$  une base de  $E$ .

(1) On définit le **rang** de  $\phi$  par  $\text{rang}(\phi) = \text{rang}(A)$ , où  $A = \text{Mat}_{\mathcal{B}}(\phi)$ ; ceci ne dépend pas du choix de la base  $\mathcal{B}$ .

(2) On dit que  $\phi$  est **non-dégénérée** si  $\text{rang}(\phi) = \dim E$ , i.e. si sa matrice dans une (et donc dans toute) base de  $E$  est inversible.

*Démonstration.* — Soient  $\mathcal{B}'$  une autre base de  $E$  et  $P = \text{Mat}_{\mathcal{B}}(\mathcal{B}')$ . Comme  ${}^t P$  et  $\overline{P}$  sont inversibles (on a  $({}^t P)^{-1} = {}^t(P^{-1})$  et  $(\overline{P})^{-1} = \overline{P^{-1}}$ ), alors la matrice  $A' = \text{Mat}_{\mathcal{B}'}(\phi) = {}^t P A \overline{P}$  a même rang que  $A$ .  $\square$

**Définition et proposition 4.1.9 (Orthogonalité).** — Soit  $\phi$  une forme hermitienne sur un  $\mathbb{C}$ -espace vectoriel  $E$ .

(1) On dit que deux vecteurs  $x, y \in E$  sont **orthogonaux** (pour  $\phi$ ) si  $\phi(x, y) = 0$ ; ceci équivaut à dire que  $\phi(y, x) = 0$  (puisque  $\phi(y, x) = \overline{\phi(x, y)}$  et vice-versa). Plus généralement, on dit que deux sous-ensembles  $X, Y$  de  $E$  sont **orthogonaux** si l'on a  $\phi(x, y) = 0$  pour tout  $x \in X$  et  $y \in Y$ . On notera  $X \perp Y$  pour signifier que  $X$  et  $Y$  sont orthogonaux.

(2) Pour tout sous-ensemble  $Y$  de  $E$ , on définit son **orthogonal** (relativement à  $\phi$ ), noté  $Y^{\perp}$  ou simplement  $Y^{\perp}$ , par :

$$(*) \quad Y^{\perp} = \{x \in E \mid \phi(x, y) = 0, \quad \forall y \in Y\}$$

c'est un **sous-espace vectoriel** de  $E$  (même si  $Y$  n'en est pas un); de plus, on a les propriétés suivantes :

$$(**) \quad Y \subseteq Z \implies Z^{\perp} \subseteq Y^{\perp} \quad Y^{\perp} = \text{Vect}(Y)^{\perp}$$

en particulier, si  $Y$  est un sous-espace vectoriel  $F$  de  $E$  et si  $(f_1, \dots, f_p)$  est une famille génératrice de  $F$ , alors

$$F^{\perp} = \{f_1, \dots, f_p\}^{\perp} = \{x \in E \mid \phi(x, f_i) = 0, \quad \forall i = 1, \dots, p\}.$$

(3) On pose  $N(\phi) = E^{\perp} = \{x \in E \mid \phi(x, y) = 0, \quad \forall y \in Y\}$  et on l'appelle le **noyau** de  $\phi$ .

*Démonstration.* — Soient  $x, x' \in Y^{\perp}$  et  $\lambda \in \mathbb{C}$ , alors on a, pour tout  $y \in Y$ ,  $\phi(\lambda x + x', y) = \lambda \phi(x, y) + \phi(x', y) = 0$ , ce qui montre que  $\lambda x + x' \in Y^{\perp}$ . Donc  $Y^{\perp}$  est un sous-espace vectoriel de  $E$ .

Il est immédiat que si  $Y \subseteq Z$ , alors  $Z^{\perp} \subseteq Y^{\perp}$  car si  $x \in Z^{\perp}$  alors  $x$  est orthogonal à tout élément de  $Z$ , donc  $x$  est a fortiori orthogonal à tout élément de  $Y$  (puisque  $Y \subseteq Z$ ), donc  $x \in Y^{\perp}$ .

Comme  $Y \subseteq \text{Vect}(Y)$ , ceci donne déjà l'inclusion  $\text{Vect}(Y)^{\perp} \subseteq Y^{\perp}$ . Montrons l'inclusion réciproque. Soit  $x \in Y^{\perp}$  et soit  $v$  un élément arbitraire de  $\text{Vect}(Y)$ , par définition,  $v$  s'écrit comme une combinaison linéaire finie  $v = \lambda_1 y_1 + \dots + \lambda_r y_r$ , avec  $y_i \in Y$  et  $\lambda_i \in \mathbb{C}$ ; alors on a

$$\phi(x, v) = \sum_{i=1}^r \lambda_i \underbrace{\phi(x, y_i)}_{=0} = 0$$

et donc  $x \in \text{Vect}(Y)^{\perp}$ . Ceci montre l'inclusion  $Y^{\perp} \subseteq \text{Vect}(Y)^{\perp}$ , d'où l'égalité  $\text{Vect}(Y)^{\perp} = Y^{\perp}$ . L'assertion (2) est démontrée.  $\square$

**Théorème 4.1.10 (Orthogonal d'un sous-espace).** — Soit  $\phi$  une forme hermitienne sur un  $\mathbb{C}$ -espace vectoriel  $E$  de dimension  $n$  et soit  $F$  un sous-espace vectoriel de  $E$ , de dimension  $r$ .

(1) On a  $F \subseteq (F^{\perp})^{\perp}$  et  $\dim F^{\perp} \geq \dim E - \dim F$ .

(2)  $N(\phi) = \{0\} \iff \phi$  est non-dégénérée.

(3) Si  $\phi$  est non-dégénérée, on a  $\dim F^\perp = \dim E - \dim F$  et  $F = (F^\perp)^\perp$ .

(4) Si  $F \cap F^\perp = \{0\}$ , alors  $E = F \oplus F^\perp$ .

*Démonstration.* — Soit  $f \in F$ , pour tout  $x \in F^\perp$  on a  $\phi(f, x) = \overline{\phi(x, f)} = 0$ , d'où  $f \in (F^\perp)^\perp$ . Ceci montre la première assertion de (1). Prouvons la seconde.

Soit  $(f_1, \dots, f_r)$  une base de  $F$ , complétons-la en une base  $\mathcal{B} = (f_1, \dots, f_n)$  de  $E$ , et soit  $A = (a_{ij})_{1 \leq i, j \leq n}$  la matrice de  $\phi$  dans la base  $\mathcal{B}$ , i.e.  $a_{ij} = \phi(f_i, f_j)$  pour  $i, j = 1, \dots, n$ .

D'après le point (2) de 4.1.9,  $F^\perp$  est formé des vecteurs  $v = x_1 f_1 + \dots + x_n f_n \in E$  tels que  $\phi(v, f_i) = 0$  pour  $i = 1, \dots, r$ . Comme  $\phi(x_1 f_1 + \dots + x_n f_n, f_i) = \sum_{j=1}^n x_j \phi(f_j, f_i) = \sum_{j=1}^n a_{ji} x_j$ , ceci équivaut à dire que le vecteur

colonne  $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$  est solution du système linéaire homogène :

$$(\Sigma) \quad \begin{cases} a_{11} x_1 + \dots + a_{n1} x_n = 0 \\ \vdots \\ a_{1r} x_1 + \dots + a_{nr} x_n = 0 \end{cases}$$

dont la matrice  $B$  est formée des  $r$  premières lignes de la matrice  ${}^t A$ . Comme l'espace des solutions du système est de dimension  $n - \text{rang}(B)$ , on obtient :

$$\dim F^\perp = n - \text{rang}(B) \geq n - r,$$

ce qui prouve la seconde assertion de (1). De plus, dans le cas particulier où  $F = E$ , on a  $B = {}^t A$  et, comme  $\text{rang}({}^t A) = \text{rang}(A)$ , on obtient que  $\dim E^\perp = n - \text{rang}(A)$ . Donc  $N(\phi) = E^\perp$  est nul si et seulement si  $\text{rang}(A) = n$ . Ceci prouve (2).

Supposons  $\phi$  non-dégénérée. Alors  $A$  est de rang  $n$ , i.e. ses colonnes sont linéairement indépendantes, en particulier les  $r$  premières colonnes le sont, donc la matrice  $B$  est de rang  $r$ , et donc  $\dim F^\perp = n - r$ . Remplaçant alors  $F$  par  $F^\perp$ , on obtient l'égalité  $\dim (F^\perp)^\perp = n - (n - r) = r$ , et par conséquent l'inclusion  $F \subseteq (F^\perp)^\perp$  est une égalité. Ceci prouve (3).

Enfin, supposons  $F \cap F^\perp = \{0\}$  (sans supposer  $\phi$  non-dégénérée). Alors  $F$  et  $F^\perp$  sont en somme directe, et le sous-espace  $F \oplus F^\perp$  de  $E$  est de dimension  $d = r + \dim F^\perp$ . D'après (1), on a  $d \geq n$ , d'où  $E = F \oplus F^\perp$  (et  $\dim F^\perp = n - r$ ). Ceci prouve (4). Le théorème est démontré.  $\square$



**Définition 4.1.11 (Bases orthogonales).** — Soit  $E$  un  $\mathbb{C}$ -espace vectoriel de dimension  $n$  et soient  $\phi$  une forme hermitienne sur  $E$ , et  $Q$  la forme quadratique hermitienne associée. Soit  $\mathcal{B} = (e_1, \dots, e_n)$  une base de  $E$

a) On dit que  $\mathcal{B}$  est une base **orthogonale** pour  $\phi$  (ou pour  $Q$ ) si l'on  $\phi(e_i, e_j) = 0$  pour  $i \neq j$ .

b) Ceci équivaut à dire que la matrice  $A = \text{Mat}_{\mathcal{B}}(\phi)$  est **diagonale**; si l'on note  $\lambda_1, \dots, \lambda_n$  ses coefficients diagonaux (qui sont **réels**) et  $(z_1, \dots, z_n)$  les coordonnées dans la base  $\mathcal{B}$ , ceci équivaut encore à dire que  $Q(z_1, \dots, z_n) = \lambda_1 |z_1|^2 + \dots + \lambda_n |z_n|^2$ .



**Théorème 4.1.12 (de Sylvester dans le cas hermitien).** — Soit  $\phi$  une forme hermitienne sur un  $\mathbb{C}$ -espace vectoriel  $E$  de dimension  $n$ , et soit  $Q$  la forme quadratique hermitienne associée.

(1) Il existe une base  $\mathcal{B}$  de  $E$  orthogonale pour  $\phi$ .

(2) Soient  $\mathcal{B} = (e_1, \dots, e_n)$  une base orthogonale pour  $\phi$  et  $D$  la matrice diagonale  $\text{Mat}_{\mathcal{B}}(\phi)$ . Quitte à renommer les  $e_i$ , on peut supposer que les coefficients diagonaux  $\lambda_1, \dots, \lambda_r \in \mathbb{R}$  sont  $\neq 0$ , et que  $\lambda_i = 0$  pour  $i > r$ . Notons  $(z_1, \dots, z_n)$  les coordonnées dans la base  $\mathcal{B}$ , alors :

(a) On a  $Q(z_1, \dots, z_n) = \lambda_1 |z_1|^2 + \dots + \lambda_r |z_r|^2$ . (\*)

(b) Soit  $p$  (resp.  $q$ ) le nombre d'indices  $i$  tels que  $Q(e_i) > 0$  (resp.  $< 0$ ). Alors  $p$  et  $q$  ne dépendent pas de la base orthogonale choisie.

(c) Le couple  $(p, q)$  s'appelle la **signature** de  $\phi$ ; on a  $p + q = r = \text{rang}(\phi)$ .

(d)  $N(\phi)$  est le sous-espace  $\text{Vect}(e_{r+1}, \dots, e_n)$ , donné par les équations  $z_1 = 0 = \dots = z_r$ .

(3) De plus, on peut choisir  $\mathcal{B}$  de sorte que la matrice diagonale  $D = \text{Mat}_{\mathcal{B}}(\phi)$  ait pour termes diagonaux  $(1, \dots, 1, -1, \dots, -1, 0, \dots, 0)$ , le nombre de 1 (resp. -1) étant  $p$  (resp.  $q$ ).

*Démonstration.* — (1) Montrons l'existence d'une base orthogonale en procédant par récurrence sur  $n = \dim E$ . Il n'y a rien à montrer si  $n = 0$  ou si  $\phi = 0$ . On peut donc supposer  $n \geq 1$ , le résultat établi pour  $n - 1$ , et  $\phi \neq 0$ . Alors, d'après 4.1.4, la forme quadratique hermitienne  $Q$  est non nulle, donc il existe  $e_1 \in E$  tel que  $Q(e_1) \neq 0$ . Posons  $F = \mathbb{C}e_1$ , comme  $\phi(e_1, e_1) \neq 0$ , alors  $F \cap F^\perp = \{0\}$  donc, d'après le théorème 4.1.10, on a

$$E = F \oplus F^\perp.$$

Par hypothèse de récurrence, il existe une base  $(e_2, \dots, e_n)$  de  $F^\perp$  telle que  $\phi(e_i, e_j) = 0$  pour  $i \neq j$ . Alors  $(e_1, e_2, \dots, e_n)$  est une base de  $E$  orthogonale pour  $\phi$ . Ceci prouve l'assertion (1).

Puis, (2.a) et l'égalité  $p + q = r = \text{rang}(\phi)$  dans (2.c) découlent aussitôt des définitions. Prouvons maintenant (2.d). D'après (\*),  $\phi$  est donnée dans la base  $\mathcal{B}$  par :

$$(*) \quad \forall u = \sum_{i=1}^n x_i e_i, \quad \forall v = \sum_{j=1}^n y_j e_j, \quad \phi(u, v) = \lambda_1 x_1 \bar{y}_1 + \cdots + \lambda_r x_r \bar{y}_r.$$

Supposons  $u \in N(\phi)$ , alors pour tout  $i = 1, \dots, r$ , prenant  $v = e_i$  (c.-à-d.,  $y_i = 1$  et  $y_j = 0$  pour  $j \neq i$ ), on obtient  $x_i = 0$ , d'où  $u \in F = \text{Vect}(e_{r+1}, \dots, e_n)$ . Réciproquement, (\*) montre aussi que tout  $u \in F$  (i.e. tel que  $x_1 = 0 = \cdots = x_r$ ) appartient à  $N(\phi)$ , d'où l'égalité désirée. Ceci prouve (2.d).

Prouvons (2.b). On note  $r = \text{rang}(\phi)$ . Soient  $\mathcal{B} = (e_1, \dots, e_n)$  et  $\mathcal{C} = (f_1, \dots, f_n)$  deux bases de  $E$  orthogonales pour  $\phi$ . Notons  $p$  (resp.  $p'$ ) le nombre d'indices  $i$  tels que  $Q(e_i) > 0$  (resp.  $Q(f_i) > 0$ ) et  $q$  (resp.  $q'$ ) le nombre d'indices  $i$  tels que  $Q(e_i) < 0$  (resp.  $Q(f_i) < 0$ ). Alors

$$p + q = r = p' + q'$$

et il s'agit de montrer que  $q = q'$  et  $p = p'$ . Quitte à renuméroter les éléments de  $\mathcal{B}$  et  $\mathcal{C}$ , on peut supposer que

$$(*) \quad \begin{cases} Q(e_i) > 0 & \text{pour } i = 1, \dots, p \\ Q(e_i) < 0 & \text{pour } i = p + 1, \dots, p + q \\ Q(e_i) = 0 & \text{pour } i > p + q = r; \end{cases} \quad \begin{cases} Q(f_i) > 0 & \text{pour } i = 1, \dots, p' \\ Q(f_i) < 0 & \text{pour } i = p' + 1, \dots, p' + q' \\ Q(f_i) = 0 & \text{pour } i > p' + q' = r. \end{cases}$$

Notons  $P_+$  le sous-espace de  $E$  engendré par les vecteurs  $e_i$  tels que  $Q(e_i) \geq 0$ . Ces vecteurs sont au nombre de  $n - q$ , donc  $\dim P_+ = n - q$ . Soit  $x$  un élément arbitraire de  $P_+$ , écrivons  $x = \sum_{i \in I} x_i e_i$ , avec  $I = \{1, \dots, p\} \cup \{r + 1, \dots, n\}$ ; alors, d'après (\*), on obtient

$$(1) \quad Q(x) = \sum_{i=1}^p |x_i|^2 Q(e_i) \geq 0.$$

D'autre part, soit  $P'_-$  le sous-espace de  $E$  engendré par les vecteurs  $f_j$  tels que  $Q(f_j) < 0$ . Ces vecteurs sont au nombre de  $q'$ , donc  $\dim P'_- = q'$ . Soit  $y$  un élément non nul de  $P'_-$ , on peut écrire  $y = \sum_{j=p'+1}^{p'+q'} y_j f_j$ , avec au moins l'un des  $y_j$  non nul (car  $y \neq 0$ ). Alors, d'après (\*) à nouveau, on obtient

$$(2) \quad Q(y) = \sum_{j=p'+1}^{p'+q'} |y_j|^2 Q(f_j) < 0.$$

Par conséquent, on a  $P_+ \cap P'_- = \{0\}$  et donc

$$n = \dim E \geq \dim P_+ + \dim P'_- = n - q + q'$$

d'où  $q \geq q'$ . Échangeant les rôles des bases  $\mathcal{B}$  et  $\mathcal{C}$ , on obtient de même  $q' \geq q$ , d'où  $q = q'$ , et de même  $p = p'$ . Ceci prouve (2.b).

Voyons l'assertion (3). Soit  $\mathcal{B} = (e_1, \dots, e_n)$  comme ci-dessus; pour  $i = 1, \dots, p + q$ , notons  $|Q(e_i)| > 0$  la valeur absolue du réel  $Q(e_i) \neq 0$ . En remplaçant  $e_i$  par  $e_i / \sqrt{|Q(e_i)|}$ , pour  $i = 1, \dots, p + q$ , on obtient une base orthogonale ayant la propriété énoncée dans (3). Ceci achève la démonstration du théorème.  $\square$

## 4.2. Réduction en sommes de carrés de modules

**Définition 4.2.1.** — Soient  $E$  un  $\mathbb{C}$ -espace vectoriel de dimension  $n$ ,  $Q$  une forme quadratique hermitienne sur  $E$  et  $\phi$  sa forme polaire. Soit  $\mathcal{B} = (e_1, \dots, e_n)$  une base de  $E$ , notons  $(x_1, \dots, x_n)$  les coordonnées dans cette base, i.e.  $x_i$  désigne en fait la forme linéaire  $f_i = e_i^*$  sur  $E$ .

(1) On dit que  $Q$  s'écrit dans la base  $\mathcal{B}$  comme **somme de carrés de modules de formes linéaires indépendantes** si l'expression de  $Q$  en fonction des coordonnées  $x_i$  est de la forme

$$Q = q_1 |x_1|^2 + \cdots + q_n |x_n|^2, \quad (\text{avec } q_1, \dots, q_n \in \mathbb{R}).$$

D'après 4.1.11, ceci équivaut à dire que la matrice de  $\phi$  dans la base  $\mathcal{B}$  est **diagonale**, avec les  $q_i$  pour coefficients diagonaux.

(2) Les formes linéaires  $f_i = e_i^*$  sont linéairement indépendantes ( $\mathcal{B}^* = (e_1^*, \dots, e_n^*)$  est la base duale de  $\mathcal{B}$ ), d'où la terminologie « somme de carrés de modules de **formes linéaires indépendantes** ». En pratique, pour abrégé on écrira souvent « somme de carrés de modules », mais il est essentiel de s'assurer que les formes linéaires en question sont bien linéairement indépendantes (cf. l'exemple 1.3.6 pour les formes bilinéaires symétriques).

Comme dans le cas des formes bilinéaires symétriques, on dispose d'un procédé algorithmique simple pour réduire une forme quadratique hermitienne en « somme de carrés de modules » (de formes linéaires indépendantes); au lieu des égalités  $x^2 + 2xL = (x + L)^2 - L^2$  et  $4x_1x_2 = (x_1 + x_2)^2 - (x_1 - x_2)^2$ , on va utiliser les égalités :

$$|z|^2 + 2\Re(z\bar{L}) = |z + L|^2 - |L|^2, \quad 4\Re(z_1\bar{z}_2) = |z_1 + z_2|^2 - |z_1 - z_2|^2.$$


**Théorème 4.2.2 (Réduction d'une forme hermitienne en somme de carrés de modules)**

Soient  $E$  un  $\mathbb{C}$ -espace vectoriel de dimension  $n$ , et  $Q$  une forme quadratique hermitienne sur  $E$ , donnée en fonctions des coordonnées  $(x_1, \dots, x_n)$  dans une base  $\mathcal{B}$  par :

$$(*) \quad Q(x_1, \dots, x_n) = \sum_i b_i |x_i|^2 + \sum_{1 \leq i < j \leq n} (b_{ij} x_i \bar{x}_j + \overline{b_{ij}} x_j \bar{x}_i) \quad (b_i \in \mathbb{R}, \quad b_{ij} \in \mathbb{C}).$$

(1) Par une suite d'opérations « élémentaires » (décrites dans la démonstration), on peut trouver un nouveau système de coordonnées  $(y_1, \dots, y_n)$  sur  $E$ , dans lequel  $Q$  s'écrit comme une somme de carrés de modules, i.e. :

$$(**) \quad Q(y_1, \dots, y_n) = a_1 |y_1|^2 + \dots + a_n |y_n|^2.$$

(2) La signature de  $Q$  est  $(p, q)$ , où  $p$  (resp.  $q$ ) est le nombre de coefficients  $a_i$  qui sont  $> 0$  (resp.  $< 0$ ), et  $\text{rang}(Q) = p + q$ .

(3) De plus,  $N(Q)$  est le sous-espace vectoriel de  $E$  défini par les équations  $y_i = 0$ , pour  $i$  parcourant l'ensemble des  $i \in \{1, \dots, n\}$  tels que  $a_i \neq 0$ .

*Démonstration.* — Remarquons d'abord que si  $Q$  s'écrit sous la forme  $(**)$  dans une base  $\mathcal{B}'$ , alors la matrice de sa forme polaire  $y$  est diagonale, avec les  $a_i$  pour coefficients diagonaux, d'où les assertions (2) et (3) du théorème, compte-tenu du théorème 4.1.12.

Il reste à donner une démonstration « algorithmique » de l'assertion (1). On procède par récurrence sur le nombre  $n$  de variables. Si  $n = 1$  on a  $Q(x_1) = b_1 |x_1|^2$ , et  $(**)$  est vérifié. On peut donc supposer  $n > 1$  et le résultat démontré pour  $n - 1$ . Distinguons deux cas.

(a) Si dans l'écriture  $(*)$  plus haut, il existe un coefficient « diagonal »  $b_i$  non nul, on peut supposer, quitte à changer l'ordre des coordonnées, que  $b_1 \neq 0$ . On considère alors la somme de **tous** les termes contenant  $x_1$  ou  $\bar{x}_1$  et on l'écrit comme suit :

$$S = b_1 |x_1|^2 + \sum_{j=2}^n 2\Re(\overline{b_{j1}} x_1 \bar{x}_j) = b_1 \left( |x_1|^2 + 2\Re \left( x_1 \overline{L(x_2, \dots, x_n)} \right) \right)$$

où  $L(x_2, \dots, x_n) = \sum_{j=2}^n (b_{j1}/b_1) x_j$ . Puis, en utilisant que

$$|x_1 + L|^2 = |x_1|^2 + 2\Re(x_1 \bar{L}) + |L|^2, \quad \text{d'où} \quad |x_1|^2 + 2\Re(x_1 \bar{L}) = |x_1 + L|^2 - |L|^2,$$

on réécrit ceci sous la forme :

$$S = b_1 |x_1 + L|^2 - b_1 |L|^2 = b_1 \left| x_1 + \sum_{j=2}^n \frac{b_{j1}}{b_1} x_j \right|^2 - \frac{1}{b_1} \sum_{j=2}^n |b_{j1}|^2 |x_j|^2 - \frac{2}{b_1} \sum_{2 \leq i < j \leq n} \Re(b_{i1} \overline{b_{j1}} x_i \bar{x}_j).$$

Donc, en posant  $y_1 = x_1 + \sum_{j=2}^n \frac{b_{j1}}{b_1} x_j$  (et  $b'_j = b_j - |b_{j1}|^2/b_1$  pour  $j = 2, \dots, n$ , et  $b'_{ij} = b_{ij} - b_{i1} \overline{b_{j1}}/b_1$  pour  $2 \leq i < j \leq n$ ), on obtient une écriture :

$$(\dagger) \quad Q(y_1, x_2, \dots, x_n) = b_1 |y_1|^2 + \underbrace{\sum_{j=2}^n b'_j |x_j|^2 + \sum_{2 \leq i < j \leq n} 2\Re(b'_{ij} x_i \bar{x}_j)}_{Q_1(x_2, \dots, x_n)}$$

où la forme quadratique hermitienne  $Q_1(x_2, \dots, x_n)$  ne dépend que des variables  $x_2, \dots, x_n$ .

L'opération  $y_1 = x_1 + L(x_2, \dots, x_n)$  et  $x_j = x_j$  pour  $j \geq 2$ , est bien un changement de coordonnées, car la matrice exprimant  $(y_1, x_2, \dots, x_n)$  en fonction de  $(x_1, \dots, x_n)$  est triangulaire avec des 1 sur la diagonale, donc inversible; explicitement le changement de coordonnées inverse est donné par  $x_j = x_j$  pour  $j \geq 2$  et  $x_1 = y_1 - L(x_2, \dots, x_n)$ .

Par hypothèse de récurrence on peut faire un changement de coordonnées  $(x_2, \dots, x_n) \rightarrow (y_2, \dots, y_n)$  tel que  $Q_1(x_2, \dots, x_n) = a_2 |y_2|^2 + \dots + a_n |y_n|^2$  d'où, d'après  $(\dagger)$  :

$$Q(y_1, \dots, y_n) = a_1 |y_1|^2 + \dots + a_n |y_n|^2$$

(avec  $a_1 = b_1$ ), ce qui prouve le résultat voulu dans ce cas.

(b) Supposons au contraire que **tous** les coefficients « diagonaux »  $b_i$  soient nuls. Si  $Q = 0$ , il n'y a rien à montrer; sinon on peut supposer, quitte à changer l'ordre des coordonnées, que  $b_{12} \neq 0$ . Comme

$$\Re(b_{12} x_1 \bar{x}_2) = \frac{1}{4} (|b_{12} x_1 + x_2|^2 - |b_{12} x_1 - x_2|^2),$$

posons  $y_1 = \frac{1}{2}(b_{12} x_1 + x_2)$  et  $y_2 = \frac{1}{2}(b_{12} x_1 - x_2)$ ; c'est bien un changement de variable, dont l'inverse est donné par

$$x_1 = b_{12}^{-1} (y_1 + y_2) \quad x_2 = y_1 - y_2.$$

Alors :  $2\Re(b_{12} x_1 \bar{x}_2) = 2(|y_1|^2 - |y_2|^2)$ , les termes  $2\Re(b_{ij} x_i \bar{x}_j)$  sont inchangés pour  $i < j$  dans  $\{3, \dots, n\}$ , et l'on a :

$$\begin{cases} 2\Re(b_{1j} x_1 \bar{x}_j) = 2\Re(b_{1j} b_{12}^{-1} (y_1 + y_2) \bar{x}_j) & \text{pour } j \geq 3 \\ 2\Re(b_{2j} x_2 \bar{x}_j) = 2\Re(b_{2j} (y_1 - y_2) \bar{x}_j) & \text{pour } j \geq 3 \end{cases}$$

donc  $Q(y_1, y_2, x_3, \dots, x_n)$  égale :

$$2|y_1|^2 - 2|y_2|^2 + \sum_{j=3}^n 2\Re\left((b_{1j}b_{12}^{-1} + b_{2j})y_1\bar{x}_j + (b_{1j}b_{12}^{-1} - b_{2j})y_2\bar{x}_j\right) + \sum_{3 \leq i < j \leq n} 2\Re(b_{ij}x_i\bar{x}_j)$$

et l'on est ramené au cas **(a)**, c.-à-d., on peut maintenant éliminer la variable  $y_1$  et se ramener, à nouveau, au cas de  $n - 1$  variables. Le théorème est démontré.  $\square$

Illustrons ceci par deux exemples : dans le premier n'apparaissent que des changements de coordonnées du type **(a)**.

**Exemple 4.2.3.** — Considérons dans  $\mathbb{C}^3$  la forme quadratique hermitienne

$$q(x_1, x_2, x_3) = x_1\bar{x}_1 - 2ix_1\bar{x}_2 + 2ix_2\bar{x}_1 + ix_1\bar{x}_3 - ix_3\bar{x}_1 + 2x_2\bar{x}_2 + 2x_3\bar{x}_3 - 2ix_2\bar{x}_3 + 2ix_3\bar{x}_2.$$

Alors  $q$  contient le terme  $x_1\bar{x}_1 = |x_1|^2$ , et les termes contenant  $x_1$  ou  $\bar{x}_1$  sont :

$$\begin{aligned} x_1\bar{x}_1 - 2ix_1\bar{x}_2 + 2ix_2\bar{x}_1 + ix_1\bar{x}_3 - ix_3\bar{x}_1 &= |x_1|^2 + 2\Re(x_1\overline{(2ix_2 - ix_3)}) \\ &= |x_1 + 2ix_2 - ix_3|^2 - |2ix_2 - ix_3|^2 \\ &= |x_1 + 2ix_2 - ix_3|^2 - 4|x_2|^2 - |x_3|^2 + 4\Re(x_2\bar{x}_3) \end{aligned}$$

donc, posant  $y_1 = x_1 + 2ix_2 - ix_3$ , on obtient

$$\begin{aligned} q(y_1, x_2, x_3) &= |y_1|^2 - 4|x_2|^2 - |x_3|^2 + 4\Re(x_2\bar{x}_3) + 2|x_2|^2 + 2|x_3|^2 - 4\Re(ix_2\bar{x}_3) \\ &= |y_1|^2 - 2|x_2|^2 + |x_3|^2 + 4\Re(x_2\overline{(1+i)x_3}). \end{aligned}$$

Puis

$$\begin{aligned} -2\left(|x_2|^2 - 2\Re(x_2\overline{(1+i)x_3})\right) &= -2\left(|x_2 - (1+i)x_3|^2 - |(1+i)x_3|^2\right) \\ &= -2\left(|x_2 - (1+i)x_3|^2 - 2|x_3|^2\right) \end{aligned}$$

donc, posant  $y_2 = x_2 - (1+i)x_3$  on obtient :  $q(y_1, y_2, x_3) = |y_1|^2 - 2|y_2|^2 + 5|x_3|^2$ . Donc la signature de  $q$  est  $(2, 1)$  et son rang est  $2 + 1 = 3$ , i.e.  $h$  est non-dégénérée.

**Exemple 4.2.4.** — Considérons dans  $\mathbb{C}^3$  la forme quadratique hermitienne

$$Q(x_1, x_2, x_3) = x_1\bar{x}_2 + x_2\bar{x}_1 + ix_1\bar{x}_3 - ix_3\bar{x}_1 + (1+i)x_2\bar{x}_3 + (1-i)x_3\bar{x}_2.$$

Ici, il n'y a pas de « termes carrés »  $|x_i|^2$ , donc on va considérer le terme  $x_1\bar{x}_2$ . On a  $b_{12} = 1$ , faisons le changement de coordonnées :

$$y_1 = \frac{1}{2}(x_1 + x_2), \quad y_2 = \frac{1}{2}(x_1 - x_2), \quad \text{d'où} \quad x_1 = y_1 + y_2, \quad x_2 = y_1 - y_2.$$

On a  $2\Re(x_1\bar{x}_2) = 2(|y_1|^2 - |y_2|^2)$ , et

$$\begin{cases} 2\Re(ix_1\bar{x}_3) &= 2\Re(iy_1\bar{x}_3) + 2\Re(iy_2\bar{x}_3) \\ 2\Re((1+i)x_2\bar{x}_3) &= 2\Re((1+i)y_1\bar{x}_3) - 2\Re((1+i)y_2\bar{x}_3) \end{cases}$$

d'où

$$Q(y_1, y_2, x_3) = 2|y_1|^2 - 2|y_2|^2 + 2\Re((1+2i)y_1\bar{x}_3) - 2\Re(y_2\bar{x}_3).$$

Puis  $2\left(|y_1|^2 + 2\Re\left(\frac{1+2i}{2}y_1\bar{x}_3\right)\right) = 2\left(|y_1 + \frac{1-2i}{2}x_3|^2 - \left|\frac{1-2i}{2}x_3\right|^2\right)$  donc, posant  $z_1 = y_1 + \frac{1-2i}{2}x_3$ , on obtient

$$Q(z_1, y_2, x_3) = 2|z_1|^2 - \frac{5}{2}|x_3|^2 - 2|y_2|^2 - 2\Re(y_2\bar{x}_3).$$

Puis  $-2\left(|y_2|^2 + \Re(y_2\bar{x}_3)\right) = -2\left(|y_2 + \frac{1}{2}x_3|^2 - \frac{1}{4}|x_3|^2\right)$  donc, posant  $z_2 = y_2 + \frac{1}{2}x_3$ , on obtient

$$Q(z_1, z_2, x_3) = 2|z_1|^2 - 2|z_2|^2 - 2|x_3|^2.$$

Donc la signature de  $Q$  est  $(1, 2)$  et son rang est  $1 + 2 = 3$ , i.e.  $Q$  est non-dégénérée.

### 4.3. Espaces hilbertiens. Inégalité de Cauchy-Schwarz. Isométries

**Définitions 4.3.1 (Produits scalaires et espaces hilbertiens).** — Soit  $E$  un  $\mathbb{C}$ -espace vectoriel de dimension finie.

(1) Soient  $\phi$  une forme hermitienne sur  $E$  et  $Q$  la forme quadratique hermitienne associée (i.e.  $Q(x) = \phi(x, x)$  pour tout  $x \in E$ ). On dit que  $Q$  (ou  $\phi$ ) est **définie positive** si l'on a :

$$\text{(Déf. Pos.)} \quad \boxed{\forall x \in E - \{0\}, \quad Q(x) = \phi(x, x) > 0.}$$

Dans ce cas, on dit que  $\phi$  est un produit scalaire **hilbertien** et on note souvent  $\phi(x, y) = (x | y)$ .

Remarquons que si  $Q$  (ou  $\phi$ ) est définie positive, elle est non-dégénérée : en effet, si  $x \in N(\phi)$ , on a  $0 = \phi(x, y)$  pour tout  $y \in E$ , en particulier  $\phi(x, x) = 0$ , d'où  $x = 0$ .

(2) Dans ce cas, on dit que : «  $E$ , muni de  $( | )$  » (ou que : « le couple  $(E, \phi)$  ») est un **espace hilbertien**. Pour abrégé, on écrira souvent : « Soit  $E$  un espace hilbertien », sans préciser le produit scalaire  $( | )$ , celui-ci étant sous-entendu.

**Exemple 4.3.2.** — Le produit scalaire hilbertien standard sur  $\mathbb{C}^n$  est défini par :

$$(x | y) = x_1 \bar{y}_1 + \cdots + x_n \bar{y}_n \quad \text{si } x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \quad y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}.$$

Pour ce produit scalaire, la base canonique  $(e_1, \dots, e_n)$  de  $\mathbb{R}^n$  est *orthonormée*, i.e. on a  $(e_i | e_j) = 1$  si  $i = j$  et  $= 0$  sinon, et la forme quadratique hermitienne associée est  $Q(x) = |x_1|^2 + \cdots + |x_n|^2$ .

**Définition et proposition 4.3.3 (Familles et bases orthonormées).** — Soit  $E$ , muni de  $( | )$ , un espace hilbertien.

- (1) Une famille  $(e_i)_{i \in I}$  de vecteurs est dite **orthonormée** si  $(e_i | e_i) = 1$  et  $(e_i | e_j) = 0$  pour tout  $i \neq j$ .
- (2) Supposons  $E$  de dimension  $n$ . Une **base orthonormée** est une base  $(e_1, \dots, e_n)$  de  $E$  qui est une famille orthonormée, i.e. qui vérifie  $(e_i | e_i) = 1$  et  $(e_i | e_j) = 0$  pour tout  $i \neq j$ .
- (3) Toute famille orthonormée est libre. En particulier, si  $\dim E = n$ , toute famille orthonormée  $(f_1, \dots, f_n)$  de cardinal  $n$  est une base orthonormée de  $E$ .
- (4) Dans la suite, on abrégera souvent « base orthonormée » en : *b.o.n.*

*Démonstration.* — Prouvons (3). Supposons qu'on ait une relation  $0 = t_1 e_{i_1} + \cdots + t_p e_{i_p}$ , avec  $i_1, \dots, i_p \in I$  deux à deux distincts, et  $t_1, \dots, t_p \in \mathbb{R}$ . Fixons un indice  $r \in \{1, \dots, p\}$  et appliquons  $(e_{i_r} | \cdot)$  à l'égalité précédente. Comme  $(e_{i_r} | e_{i_s}) = 0$  pour  $s \neq r$ , on obtient  $0 = t_r (e_{i_r} | e_{i_r}) = t_r$ , d'où  $t_r = 0$ . Ceci prouve que la famille  $(e_i)_{i \in I}$  est libre.  $\square$

**Théorème 4.3.4 (Existence de b.o.n.).** — Soit  $E$  un espace hilbertien de dimension  $n$ . Alors  $E$  admet une base orthonormée

*Démonstration.* — D'après le théorème 4.1.12, il existe une base  $(e_1, \dots, e_n)$  orthogonale (i.e.  $(e_i | e_j) = 0$  pour  $i \neq j$ ) et telle que  $(e_i | e_i) \in \{1, -1, 0\}$ ; or comme  $( | )$  est défini positif on a nécessairement  $(e_i | e_i) = 1$ , donc  $(e_1, \dots, e_n)$  est une b.o.n.  $\square$

**Définition 4.3.5 (Sous-espaces d'un espace hilbertien).** — Soit  $E$ , muni de  $( | )$ , un espace hilbertien et soit  $F$  un sous-espace vectoriel de  $E$ . Alors la restriction  $( | )_F$  de  $( | )$  à  $F$  est un produit scalaire hilbertien sur  $F$ , puisque  $(x | x)_F = (x | x) > 0$  pour tout  $x \in F - \{0\}$ . Donc  $F$  muni de  $( | )_F$  est un espace hilbertien.

**Théorème et définition 4.3.6 (Projection orthogonale sur un sous-espace)**

Soit  $E$ , muni de  $( | )$ , un espace hilbertien et soient  $F$  un sous-espace et  $F^\perp$  son orthogonal pour  $( | )$ .

(1) On a  $\boxed{E = F \oplus F^\perp}$ . Le projecteur  $\pi_F : E \rightarrow E$ , d'image  $F$  et de noyau  $F^\perp$ , défini par cette décomposition s'appelle la **projection orthogonale** sur  $F$ .

(2) Soit  $(e_1, \dots, e_r)$  une base **orthonormée** de  $F$ . Alors  $\boxed{\pi_F(v) = (v | e_1)e_1 + \cdots + (v | e_r)e_r}$  pour tout  $v \in E$ .

(3) On a  $\boxed{(F^\perp)^\perp = F}$  donc la projection orthogonale  $\pi_{F^\perp}$  sur  $F^\perp$  n'est autre que  $\text{id}_E - \pi_F$ , i.e. on a  $\boxed{\text{id}_E = \pi_F + \pi_{F^\perp}}$ .

*Démonstration.* — Comme les formes hermitiennes  $( | )$  et  $( | )_F$  sont définies positives, donc non-dégénérées, on a  $(F^\perp)^\perp = F$  et  $E = F \oplus F^\perp$  d'après 4.1.10. Alors, tout  $x \in E$  s'écrit de façon unique  $x = f + g$  avec  $f \in F$  et  $g \in F^\perp$ , et le projecteur  $\pi_F$  sur  $F$  parallèlement à  $F^\perp$  (i.e. de noyau  $F^\perp$ ) est défini par  $\pi_F(x) = f$ . De plus, comme  $(F^\perp)^\perp = F$ , alors le projecteur  $\pi_{F^\perp}$  sur  $F^\perp$  parallèlement à  $(F^\perp)^\perp = F$  (i.e. de noyau  $F$ ) est défini par  $\pi_{F^\perp}(x) = g$ , donc on a bien  $\text{id}_E = \pi_F + \pi_{F^\perp}$ . Ceci prouve (1) et (3).

Prouvons (2). Soit  $r = \dim F$  et soit  $(e_1, \dots, e_r)$  une b.o.n. de  $F$ . Pour tout  $v \in E$ , notons provisoirement

$$\pi(v) = (v | e_1)e_1 + \cdots + (v | e_r)e_r \in F.$$

Alors, pour  $j = 1, \dots, r$ , on a  $(v - \pi(v) | e_j) = (v | e_j) - \sum_{i=1}^r (v | e_i) \underbrace{(e_i | e_j)}_{\substack{=1 \text{ si } i=j \\ =0 \text{ si } i \neq j}} = 0$ , d'où  $v - \pi(v) \in F^\perp$ , et donc

$v = \pi(v) + v - \pi(v)$ , avec  $\pi(v) \in F$  et  $v - \pi(v) \in F^\perp$ . Comme  $E = F \oplus F^\perp$ , ceci entraîne que  $\pi(v) = \pi_F(v)$ , d'où l'assertion (2).  $\square$



**Définition 4.3.7 (Normes).** — Soit  $E$  un  $\mathbb{C}$ -espace vectoriel. Une **norme**  $\|\cdot\|$  sur  $E$  est une application  $E \rightarrow \mathbb{R}_+$ ,  $v \mapsto \|v\|$  vérifiant les trois propriétés suivantes :

- (1)  $\|v\| = 0 \Leftrightarrow v = 0$ .
- (2) Pour tout  $z \in \mathbb{C}$ ,  $v \in E$ , on a  $\|zv\| = |z| \cdot \|v\|$  (où  $|z|$  est le module de  $z$ ).
- (3) (Inégalité triangulaire)  $\|u + v\| \leq \|u\| + \|v\|$ , pour tout  $u, v \in E$ .



**Théorème 4.3.8 (Inégalité de Cauchy-Schwarz et norme hilbertienne).** — Soit  $E$ , muni de  $( | )$ , un espace hilbertien et soit  $Q(x) = (x | x)$  la forme quadratique hermitienne associée.

- (1) On a l'inégalité de Cauchy-Schwarz :

$$(CS) \quad \boxed{\forall x, y \in E \quad |(x | y)|^2 \leq Q(x)Q(y)}$$

avec égalité si et seulement si  $x$  et  $y$  sont liés.

(2) Par conséquent, l'application  $x \mapsto \|x\| = \sqrt{(x | x)}$  est une norme sur  $E$ , appelée la **norme hilbertienne** associée à  $( | )$ , et l'inégalité de Cauchy-Schwarz se réécrit comme suit :

$$(CS) \quad \boxed{\forall x, y \in E \quad |(x | y)| \leq \|x\| \cdot \|y\|}$$

*Démonstration.* — (1) Soient  $x, y \in E$ . L'inégalité est trivialement vérifiée si  $x = 0$ , donc on peut supposer  $x \neq 0$ . Considérons alors le vecteur

$$v = y - \frac{(y | x)}{(x | x)}x$$

(c'est la projection orthogonale de  $y$  sur l'hyperplan  $(\mathbb{C}x)^\perp$ . Comme  $\overline{(y | x)} = (x | y)$ , on a :

$$0 \leq (v | v) = (y | y) - \frac{(y | x)}{(x | x)}(x | y) - \underbrace{\frac{(x | y)}{(x | x)}(y | x) + \frac{(y | x)(x | y)}{(x | x)^2}(x | x)}_{=0}$$

donc, multipliant par  $(x | x) > 0$ , on obtient que  $0 \leq (x | x)(v | v) = (y | y)(x | x) - |(x | y)|^2$ . Ceci prouve l'inégalité voulue, et montre que l'on a égalité si et seulement si  $(v | v) = 0$ , c.-à-d., si  $v = 0$ , i.e. si  $y = \frac{(y | x)}{(x | x)}x$ . Ceci prouve l'assertion (1).

Prouvons que  $v \mapsto \|v\| = \sqrt{(v | v)}$  est une norme sur  $E$ . Comme  $( | )$  est défini positif, on a  $\|v\| = 0 \Leftrightarrow v = 0$ . D'autre part, pour tout  $z \in \mathbb{C}$  et  $x \in E$ , on a  $|z| = \sqrt{z\bar{z}}$  et donc

$$\|zv\| = \sqrt{z\bar{z}(v | v)} = |z| \cdot \|v\|.$$

Enfin, soient  $x, y \in E$ . D'abord, l'inégalité de Cauchy-Schwarz équivaut (en prenant la racine carrée) à :

$$|(x | y)| \leq \|x\| \cdot \|y\|;$$

alors, multipliant par 2 et ajoutant  $\|x\|^2 + \|y\|^2$  aux deux membres, on obtient

$$(†) \quad \|x\|^2 + \|y\|^2 + 2|(x | y)| \leq \|x\|^2 + \|y\|^2 + 2\|x\| \cdot \|y\| = (\|x\| + \|y\|)^2.$$

D'autre part, d'après les égalités de polarisations 4.1.4, on a :

$$\|x + y\|^2 = \|x\|^2 + \|y\|^2 + 2\Re((x | y)).$$

Or, pour tout nombre complexe  $z = a + ib$ , on a :  $\Re(z) = a \leq \sqrt{a^2 + b^2} = |z|$ . On obtient donc

$$\|x + y\|^2 \leq \|x\|^2 + \|y\|^2 + 2|(x | y)|$$

ce qui, combiné avec (†), entraîne :

$$\|x + y\|^2 \leq (\|x\| + \|y\|)^2.$$

Prenant la racine carrée, ceci entraîne (et équivaut à) l'inégalité triangulaire. Le théorème est démontré.  $\square$

Récrivons les égalités de polarisation 4.1.4 en utilisant la norme  $\|\cdot\|$ , et ajoutons-y l'égalité de Pythagore :

**Proposition 4.3.9 (Polarisation).** — Soit  $E$ , muni de  $( | )$ , un espace hilbertien et soit  $\|\cdot\|$  la norme hilbertienne associée.

a) Pour tout  $x, y \in E$ , on a :

$$(1) \quad \mathcal{R}((x | y)) = \frac{1}{2}(\|x + y\|^2 - \|x\|^2 - \|y\|^2) = \frac{1}{4}(\|x + y\|^2 - \|x - y\|^2)$$

$$(2) \quad \mathcal{I}((x | y)) = \frac{1}{2}(\|x + iy\|^2 - \|x\|^2 - \|y\|^2) = \frac{1}{4}(\|x + iy\|^2 - \|x - iy\|^2)$$

$$(3) \quad 4(x | y) = \|x + y\|^2 - \|x - y\|^2 + i\|x + iy\|^2 - i\|x - iy\|^2.$$

b) Égalité de Pythagore) Si  $x_1, \dots, x_n$  sont orthogonaux, on a  $\|x_1 + \dots + x_n\|^2 = \|x_1\|^2 + \dots + \|x_n\|^2$ .

*Démonstration.* — La première assertion n'est qu'une reformulation de 4.1.4. L'égalité de Pythagore est immédiate si  $n = 2$ , et dans ce cas on a même la réciproque : si  $\|x_1 + x_2\|^2 = \|x_1\|^2 + \|x_2\|^2$  alors  $(x_1 | x_2) = 0$ . L'égalité pour  $n$  vecteurs orthogonaux s'obtient par récurrence sur  $n$ . On prendra garde que la réciproque est fautive pour  $n \geq 3$  : prendre par exemple dans  $\mathbb{C}^2$  hilbertien les vecteurs  $x_1 = e_1, x_2 = e_1 + e_2, x_3 = e_2 - e_1$ .  $\square$

**Définition et proposition 4.3.10 (Isométries vectorielles).** — Soient  $E, F$  deux espaces hilbertiens de même dimension  $n$ , notons  $( | )_E$  et  $\| \cdot \|_E$  (resp.  $( | )_F$  et  $\| \cdot \|_F$ ) le produit scalaire et la norme hilbertienne sur  $E$  (resp.  $F$ ). Soit  $f : E \rightarrow F$  une application linéaire.

(1) Les conditions suivantes sont équivalentes :

(a)  $f$  préserve la norme :  $\forall x \in E, \quad \|x\|_E = \|f(x)\|_F$

(b)  $f$  préserve le produit scalaire :  $\forall x, y \in E, \quad (x | y)_E = (f(x) | f(y))_F$

(c) Pour toute b.o.n.  $\mathcal{B} = (e_1, \dots, e_n)$  de  $E$ , la famille  $(f(e_1), \dots, f(e_n))$  est une b.o.n. de  $F$ .

(d) Il existe une b.o.n.  $\mathcal{B} = (e_1, \dots, e_n)$  de  $E$  telle que  $(f(e_1), \dots, f(e_n))$  soit une b.o.n. de  $F$ .

(2) Sous ces conditions, on dit que  $f$  est une **isométrie** vectorielle de  $E$  sur  $F$

(3) Dans ce cas,  $f$  est bijective, et son inverse  $f^{-1}$  est aussi une isométrie.

*Démonstration.* — Supposons que  $f$  préserve la norme, et soient  $x, y \in E$ . Alors  $\|x + y\|_E^2 = \|f(x + y)\|_F^2 = \|f(x) + f(y)\|_F^2$ , et le premier (resp. dernier) membre égale :

$$\|x\|_E^2 + \|y\|_E^2 + 2\mathcal{R}((x | y)_E), \quad \text{resp.} \quad \|f(x)\|_F^2 + \|f(y)\|_F^2 + 2\mathcal{R}((f(x) | f(y))_F)$$

et comme  $\|x\|_E^2 = \|f(x)\|_F^2$  et  $\|y\|_E^2 = \|f(y)\|_F^2$ , on obtient que  $\mathcal{R}((x | y)_E) = \mathcal{R}((f(x) | f(y))_F)$ .

D'autre part, pour tout  $z \in \mathbb{C}$ , on a  $\mathcal{I}(z) = \mathcal{R}(-iz)$ . Donc, appliquant ce qui précède à  $iy$  au lieu de  $y$ , on obtient aussi l'égalité :

$$\mathcal{I}((x | y)_E) = \mathcal{R}((x | iy)_E) = \mathcal{R}((f(x) | f(iy))_F) = \mathcal{R}((f(x) | if(y))_F) = \mathcal{I}((f(x) | f(y))_F),$$

d'où finalement  $(x | y)_E = (f(x) | f(y))_F$ . Ceci prouve que (a)  $\Rightarrow$  (b).

Les implications (b)  $\Rightarrow$  (c)  $\Rightarrow$  (d) sont évidentes, montrons que (d)  $\Rightarrow$  (a). Supposons (d) vérifiée. Pour tout  $x = x_1 e_1 + \dots + x_n e_n$  dans  $E$ , on a  $f(x) = \sum_i x_i f(e_i)$  et, comme  $(e_1, \dots, e_n)$  et  $(f(e_1), \dots, f(e_n))$  sont des b.o.n., on obtient

$$\|x\|_E^2 = \sum_{i=1}^n |x_i|^2 = \|f(x)\|_F^2$$

donc (a) est vérifiée. Ceci prouve l'assertion (1).

Prouvons (3). Soit  $f : E \rightarrow F$  une isométrie, et soit  $\mathcal{B} = (e_1, \dots, e_n)$  de  $E$ . Comme  $f(\mathcal{B})$  est un b.o.n. (donc une base) de  $F$ , alors  $f$  est bijective. Son inverse  $f^{-1}$  envoie la b.o.n.  $f(\mathcal{B}) = (f(e_1), \dots, f(e_n))$  de  $F$  sur la b.o.n.  $\mathcal{B}$  de  $E$ , donc  $f^{-1}$  est une isométrie. Ceci prouve (3). La proposition est démontrée.  $\square$

**Terminologie 4.3.10.1.** — On a introduit la terminologie isométrie « vectorielle » pour faire la distinction avec la notion d'isométrie « affine », étudiée au Chap. 6. Dans la suite de ce chapitre, comme on ne considère que des applications linéaires, on dira simplement « isométrie » au lieu de « isométrie vectorielle ».

**Définition et corollaire 4.3.11.** — (1) On dit que deux espaces hilbertiens  $E$  et  $E'$  sont **isométriques** s'il existe une isométrie  $f : E \xrightarrow{\sim} E'$ .

(2) Tout espace hilbertien  $E$  de dimension  $n$  est isométrique à  $\mathbb{C}^n$  muni du produit scalaire hilbertien standard.

*Démonstration.* — Soit  $\mathcal{B} = (e_1, \dots, e_n)$  la base canonique de  $\mathbb{C}^n$ , qui est orthonormée pour le produit scalaire standard. D'après le théorème 4.3.4,  $E$  admet une b.o.n.  $\mathcal{C} = (f_1, \dots, f_n)$ . Alors l'application linéaire  $u : \mathbb{C}^n \rightarrow E$  définie par  $u(e_i) = f_i$ , pour  $i = 1, \dots, n$ , est une isométrie de  $\mathbb{C}^n$  sur  $E$ .  $\square$

**Définition 4.3.12.** — On note  $U(n) = \{A \in M_n(\mathbb{C}) \mid {}^t A \bar{A} = I_n\}$ . Remarquons que l'égalité  ${}^t A \bar{A} = I_n$  équivaut à l'égalité  ${}^t \bar{A} A = I_n$ , qui entraîne que  $A$  est inversible et  $A^{-1} = {}^t \bar{A}$ . Donc  $U(n) \subset GL_n(\mathbb{C})$  et, si  $A \in U(n)$ , son inverse  $B = A^{-1} = {}^t \bar{A}$  vérifie  $B^{-1} = A = {}^t \bar{B}$ , donc appartient aussi à  $U(n)$ . De plus, pour tout  $A, B \in U(n)$ , on a l'égalité  ${}^t (AB) \overline{AB} = {}^t B {}^t A \bar{B} = {}^t B \bar{B} = I_n$ , donc  $AB \in U(n)$ . Donc  $U(n)$  est un sous-groupe de  $GL_n(\mathbb{C})$ , appelé le **groupe unitaire**.

Munissons  $\mathbb{C}^n$  du produit scalaire hilbertien standard  $(\mid)$ . Pour tout  $X, Y \in \mathbb{C}^n$  on a  $(X \mid Y) = {}^t X \bar{Y}$ , i.e. la matrice de  $(\mid)$  dans la base canonique  $\mathcal{B}_0 = (e_1, \dots, e_n)$  est la matrice identité  $I_n$ . Donc une matrice arbitraire  $A \in M_n(\mathbb{C})$  préserve le produit scalaire si et seulement si, on a, pour tout  $X, Y \in \mathbb{C}^n$  :

$${}^t X \bar{Y} = (X \mid Y) = (AX \mid AY) = {}^t X ({}^t A \bar{A}) Y$$

ce qui équivaut à dire que  ${}^t A \bar{A} = I_n$  (cf. 4.1.6). Ceci montre que  $U(n)$  est le groupe des isométries de  $\mathbb{C}^n$  muni produit scalaire hilbertien standard  $(\mid)$ .

De plus, notons  $C_1, \dots, C_n$  les colonnes de  $A$  (i.e.  $C_i$  est le vecteur  $Ae_i \in \mathbb{C}^n$ ). Remarquons que, pour tout  $i, j \in \{1, \dots, n\}$ , le coefficient d'indice  $(i, j)$  de  ${}^t A \bar{A}$  est le produit matriciel de la  $i$ -ème ligne de  ${}^t A$ , i.e. de  ${}^t C_i$ , par la colonne  $\bar{C}_j$ , c.-à-d., on a  $({}^t A \bar{A})_{ij} = (Ae_i \mid Ae_j)$ , donc la condition  ${}^t A \bar{A} = I_n$  équivaut aussi à dire que les colonnes de  $A$  sont de norme 1 et deux à deux orthogonales. Tenant compte de la proposition 4.3.10, on obtient donc les caractérisations suivantes de  $U(n)$ , chacune étant utile :



**Proposition 4.3.13 (Groupe unitaire  $U(n)$ ).** — On munit  $\mathbb{C}^n$  du produit scalaire hilbertien standard  $(\mid)$  et l'on note  $\|\cdot\|$  la norme hilbertienne associée. Alors  $U(n)$  est le groupe des isométries de  $\mathbb{C}^n$  ; il est caractérisé par chacune des égalités suivantes :

$$\begin{aligned} U(n) &= \{A \in M_n(\mathbb{C}) \mid {}^t A \bar{A} = I_n\} \\ &= \{A \in GL_n(\mathbb{C}) \mid A^{-1} = {}^t \bar{A}\} \\ &= \{A \in M_n(\mathbb{C}) \mid (AX \mid AY) = (X \mid Y), \quad \forall X, Y \in \mathbb{C}^n\} \\ &= \{A \in M_n(\mathbb{C}) \mid \|AX\| = \|X\|, \quad \forall X \in \mathbb{C}^n\} \\ &= \{A \in M_n(\mathbb{C}) \mid (Af_1, \dots, Af_n) \text{ est une b.o.n., pour toute b.o.n. } (f_1, \dots, f_n)\} \\ &= \{A \in M_n(\mathbb{C}) \mid (Ae_1, \dots, Ae_n) \text{ est une b.o.n., où } (e_1, \dots, e_n) \text{ est la base canonique de } \mathbb{C}^n\} \\ &= \{A \in M_n(\mathbb{C}) \mid \text{les colonnes de } A \text{ sont de norme 1 et deux à deux orthogonales}\} \end{aligned}$$

Les éléments de  $U(n)$  sont appelés « endomorphismes unitaires ».

**Remarque 4.3.14.** — Il existe d'autres groupes unitaires (qui ne sont isomorphes à aucun  $U(n)$ ). Soient  $p, q$  des entiers  $\geq 1$  et soit  $\phi$  la forme hermitienne sur  $\mathbb{C}^{p+q}$  définie par  $\phi(X, Y) = \sum_{i=1}^p x_i y_i - \sum_{i=p+1}^q x_i y_i$ , i.e. la matrice

de  $\phi$  dans la base canonique de  $\mathbb{C}^{p+q}$  est  $J = \begin{pmatrix} I_p & \mathbf{0}_{p,q} \\ \mathbf{0}_{q,p} & -I_q \end{pmatrix}$ . Alors

$$\{A \in M_n(\mathbb{C}) \mid {}^t A J A = J\} = \{A \in M_n(\mathbb{C}) \mid \phi(AX, AY) = \phi(X, Y), \quad \forall X, Y \in \mathbb{C}^n\}$$

est un sous-groupe de  $GL_n(\mathbb{C})$ , noté  $U(p, q)$ . On ne considérera pas ces groupes dans ce cours.

#### 4.4. Diagonalisation des endomorphismes auto-adjoints et normaux

Commençons par introduire l'adjoint dans le cas général d'une forme hermitienne non dégénérée, même si on se limitera dans la suite au cas hilbertien.

**Théorème et définition 4.4.1 (Adjoint d'un endomorphisme).** — Soient  $E$  un  $\mathbb{C}$ -espace vectoriel de dimension  $n$ , et  $\phi$  une forme hermitienne sur  $E$ , non dégénérée.

(1) Pour tout  $u \in \text{End}(E)$ , il existe un unique endomorphisme  $u^*$  de  $E$ , appelé l'adjoint de  $u$ , vérifiant :

$$(*) \quad \forall x, y \in E, \quad \boxed{\phi(u(x), y) = \phi(x, u^*(y))}.$$

(2) Pour toute base  $\mathcal{B}$  de  $E$ , si l'on note  $J = \text{Mat}_{\mathcal{B}}(\phi)$  et  $A = \text{Mat}_{\mathcal{B}}(u)$ , on a

$$(**) \quad \boxed{A^* = \text{Mat}_{\mathcal{B}}(u^*) = \overline{J^{-1} {}^t A J}}.$$

(3) On a  $(u^*)^* = u$ .

*Démonstration.* — Supposons qu'il existe  $u^*$  vérifiant  $(*)$  et soient  $\mathcal{B}$  une base de  $E$ ,  $J = \text{Mat}_{\mathcal{B}}(\phi)$ ,  $A = \text{Mat}_{\mathcal{B}}(u)$  et  $A^* = \text{Mat}_{\mathcal{B}}(u^*)$ . Soient  $x, y \in E$  arbitraires, et notons  $X, Y \in \mathbb{C}^n$  les vecteurs colonnes associés (coordonnées dans la base  $\mathcal{B}$ ). Alors on a

$${}^t X {}^t A J \bar{Y} = \phi(u(x), y) = \phi(x, u^*(y)) = {}^t X \overline{J A^* Y} = {}^t X \overline{J A^* Y}$$

d'où  ${}^t A J = \overline{J A^*}$  et donc, puisque  $J$  est inversible (car  $\phi$  non-dégénérée),  $A^* = \overline{J^{-1} {}^t A J}$ . Ceci montre que  $u^*$ , s'il existe, vérifie  $(**)$  et est donc unique.

Réciproquement, si l'on note  $u^*$  l'endomorphisme de  $E$  dont la matrice dans la base  $\mathcal{B}$  est  $A^* = \overline{J^{-1} {}^t A J}$ , alors pour tout  $x, y$  on a :

$$\phi(x, u^*(y)) = {}^t X \overline{J A^* Y} = {}^t X {}^t A J \bar{Y} = \phi(u(x), y)$$

donc  $u^*$  vérifie  $(*)$ . Ceci prouve les assertions (1) et (2).

Prouvons l'assertion (3). Pour tout  $x, y \in E$ , on a :

$$(u^*(x) \mid y) = \overline{(y \mid u^*(x))} = \overline{(u(y) \mid x)} = (x \mid u(y))$$

et ceci montre que  $u$  est l'adjoint de  $u^*$ , i.e.  $u = (u^*)^*$ . Le théorème est démontré.  $\square$

**Remarque 4.4.2.** — Il résulte de la formule (\*\*) (ou directement de la définition (\*)) que, pour tout  $u, v \in \text{End}(E)$  et  $s, t \in \mathbb{C}$ , on a  $(su + tv)^* = \bar{s}u^* + \bar{t}v^*$ , i.e. l'application  $\text{End}(E) \rightarrow \text{End}(E)$ ,  $u \mapsto u^*$  est semi-linéaire.

Remarquons aussi que si  $\phi$  est un produit scalaire hilbertien et si  $\mathcal{B}$  est une b.o.n., alors la matrice de  $\phi$  dans  $\mathcal{B}$  est  $J = I_n$ . On peut donc énoncer le théorème dans le cas hilbertien sous la forme suivante.

 **Théorème 4.4.3 (Adjoint d'un endomorphisme dans le cas hilbertien).** — Soit  $E$  muni de  $(\mid)$  un espace hilbertien de dimension  $n$ . Pour tout  $u \in \text{End}(E)$ , il existe un unique endomorphisme  $u^*$  de  $E$ , appelé l'adjoint de  $u$ , vérifiant :

$$(*) \quad \forall x, y \in E, \quad \boxed{(u(x) \mid y) = (x \mid u^*(y))}.$$

Pour toute b.o.n.  $\mathcal{B}$  de  $E$ , si l'on note  $A = \text{Mat}_{\mathcal{B}}(u)$ , on a

$$(**) \quad \boxed{A^* = \text{Mat}_{\mathcal{B}}(u^*) = {}^t\bar{A}}.$$

**Lemme 4.4.4 (Stabilité par  $u$  ou  $u^*$ ).** — Soient  $E$  muni de  $(\mid)$  un espace hilbertien de dimension  $n$ ,  $u \in \text{End}(E)$ ,  $F$  un sous-espace vectoriel de  $E$ , et  $F^\perp$  son orthogonal pour  $(\mid)$ . Alors :  $F$  est stable par  $u$  (i.e.  $u(F) \subseteq F$ ) si et seulement si  $F^\perp$  est stable par  $u^*$ .

*Démonstration.* — Supposons  $u(F) \subseteq F$ , et soit  $y \in F^\perp$ . Pour tout  $x \in F$ , on a :

$$(x \mid u^*(y)) = (u(x) \mid y) = 0$$

(la dernière égalité puisque  $u(x) \in F$  et  $y \in F^\perp$ ), et ceci montre que  $u^*(y) \in F^\perp$ . On a donc  $u^*(F^\perp) \subseteq F^\perp$ .

Réciproquement, supposons  $u^*(F^\perp) \subseteq F^\perp$ . Comme  $(F^\perp)^\perp = F$  et  $(u^*)^* = u$ , d'après 4.3.6 et 4.4.1, on obtient que  $u(F) \subseteq F$ , d'après ce qui précède.  $\square$

 **Définitions 4.4.5 (Endomorphismes normaux et auto-adjoints).** — Soit  $E$  un espace hilbertien de dimension  $n$  et soit  $u \in \text{End}(E)$ .

- (1) On dit que  $u$  est **auto-adjoint** (ou *hermitien*) si  $u^* = u$ .
- (2) On dit que  $u$  est un endomorphisme **unitaire** s'il est inversible et  $u^{-1} = u^*$  (cf. 4.3.13).
- (3) On dit que  $u$  est un endomorphisme **anti-hermitien** si  $u^* = -u$ .
- (4) Enfin, on dit que  $u$  est un endomorphisme **normal** s'il **commute à son adjoint**  $u^*$ , i.e. si l'on a  $uu^* = u^*u$ .

Ceci englobe les trois cas précédents.

Rappelons et complétons la définition 4.1.5 :

**Définition 4.4.6 (Matrices hermitiennes ou anti-hermitiennes).** — Une matrice  $A \in M_n(\mathbb{C})$  est dite **anti-hermitienne** (resp. *hermitienne*, cf. 4.1.5) si  $\boxed{{}^t\bar{A} = -A}$  (resp. si  $\bar{A} = A$ ).

Observons que si  $A$  est une matrice *anti-hermitienne* (resp. *hermitienne*), ses coefficients diagonaux  $a_{ii}$  vérifient  $\bar{a}_{ii} = -a_{ii}$  (resp.  $\bar{a}_{ii} = a_{ii}$ ) donc sont *imaginaires purs* (resp. réels).

 **Théorème 4.4.7 (Diagonalisation des endomorphismes normaux).** — Soient  $E$  muni de  $(\mid)$  un espace hilbertien de dimension  $n$ , et  $u$  un endomorphisme normal. Alors,  $u$  est diagonalisable et les espaces propres sont deux à deux orthogonaux. Par conséquent, il existe une b.o.n. de  $E$  formée de vecteurs propres de  $u$ .

*Démonstration.* — On procède par récurrence sur  $n = \dim E$ . C'est ok si  $n = 1$ , donc on peut supposer  $n \geq 2$  et le résultat établi pour  $n - 1$ . Comme  $\mathbb{C}$  est algébriquement clos, le polynôme caractéristique  $P_u(X)$  admet au moins une racine  $\lambda$  dans  $\mathbb{C}$ , et  $\lambda$  est valeur propre de  $u$ . Soit  $V_\lambda$  l'espace propre associé, il est **stable par  $u^*$**  : en effet, comme  $u$  et  $u^*$  commutent, on a, pour tout  $x \in V_\lambda$  :

$$u(u^*(x)) = u^*(u(x)) = u^*(\lambda x) = \lambda u^*(x), \quad \text{d'où} \quad u^*(x) \in V_\lambda.$$

Donc, d'après le lemme 4.4.4, l'orthogonal  $G = V_\lambda^\perp$  est stable par  $u$  et par  $u^*$ . D'autre part, d'après 4.1.10, on a  $E = V_\lambda \oplus G$  et  $\dim G = \dim E - \dim V_\lambda < \dim E$ .

Notons  $u_G$  (resp.  $u_G^*$ ) la restriction de  $u$  (resp.  $u^*$ ) à  $G$ , alors pour tout  $x, y \in G$ , on a  $u_G(x) = u(x)$  et  $u_G^*(y) = u^*(y)$  et donc :

$$(u_G(x) \mid y) = (u(x) \mid y) = (x \mid u^*(y)) = (x \mid u_G^*(y))$$

ce qui montre que l'adjoint de  $u_G$  est la restriction de  $u^*$  à  $G$ . Comme  $u$  et  $u^*$  commutent, il en est de même de leurs restrictions à  $G$ , i.e.  $u_G$  est un endomorphisme normal de  $G = V_\lambda^\perp$ . Alors, par hypothèse de récurrence,  $u_{V_\lambda^\perp}$  est diagonalisable, et ses espaces propres sont deux à deux orthogonaux. Comme  $E = V_\lambda \oplus V_\lambda^\perp$ , on obtient la même conclusion pour  $u$ . Ceci prouve qu'il existe une b.o.n.  $\mathcal{B} = (e_1, \dots, e_n)$  de  $E$  formée de vecteurs propres de  $u$ .

Il en résulte que les espaces propres de  $u$  sont deux à deux orthogonaux : en effet, soient  $\mu_1, \dots, \mu_p$  les valeurs propres, deux à deux distinctes, de  $u$  et  $V_1, \dots, V_p$  les espaces propres associés. Alors  $E = V_1 \oplus \dots \oplus V_p$  et donc, notant  $d_q = \dim V_q$  pour  $q = 1, \dots, p$ , on a :

$$(1) \quad n = d_1 + \dots + d_p.$$

Pour chaque  $q = 1, \dots, p$ , soit  $d'_q$  le nombre d'indices  $i \in \{1, \dots, n\}$  tels que le coefficient diagonal  $\lambda_i$  de  $D = \text{Mat}_{\mathcal{B}}(u)$  égale  $\mu_q$  (i.e.  $u(e_i) = \mu_q e_i$ ) et soit  $V'_q$  le sous-espace de  $V_q$  engendré ces  $e_i$ ; comme les  $e_i$  sont linéairement indépendants et comme  $V'_q$  est un sous-espace de  $V_q$ , on a :

$$(2) \quad d'_q = \dim V'_q \leq d_q.$$

D'une part, comme chaque  $e_i$  appartient à un  $V'_q$  et à un seul, on a :

$$(3) \quad n = d'_1 + \dots + d'_p.$$

Il en résulte que, pour chaque  $q$ , l'inégalité  $d'_q \leq d_q$  est une égalité, d'où  $V'_q = V_q$ . Ceci montre que chaque  $V_q$  est engendré par les éléments  $e_i \in \mathcal{B}$  qu'il contient. Comme les  $e_i$  sont deux à deux orthogonaux, on en déduit que  $V_q$  et  $V_{q'}$  sont orthogonaux si  $q \neq q'$ . Ceci achève la démonstration du théorème.  $\square$

On en déduit, en particulier, le théorème suivant.

#### **Théorème 4.4.8 (Diagonalisation des matrices hermitiennes, unitaires, ou anti-hermitiennes)**

On munit  $\mathbb{C}^n$  du produit scalaire hermitien standard. Soit  $A \in M_n(\mathbb{C})$ .

(1) Si  $A$  est **hermitienne** (i.e.  ${}^t\bar{A} = A$ , alors  $A$  est diagonalisable dans une base orthonormée  $\mathcal{B}$ , i.e. il existe  $P \in U(n)$  telle que  $P^{-1}AP = D$  soit une matrice diagonale. De plus, les valeurs propres de  $A$  sont **réelles**.

(2) Si  $A$  est **unitaire** (i.e. si  $A \in U(n)$ ), alors  $A$  est diagonalisable dans une base orthonormée  $\mathcal{B}$ , i.e. il existe  $P \in U(n)$  telle que  $P^{-1}AP = D$  soit une matrice diagonale. De plus, les valeurs propres de  $A$  sont des nombres complexes **de module 1**.

(3) Si  $A$  est **anti-hermitienne** (i.e.  ${}^t\bar{A} = -A$ ), alors  $A$  est diagonalisable dans une base orthonormée  $\mathcal{B}$ , i.e. il existe  $P \in U(n)$  telle que  $P^{-1}AP = D$  soit une matrice diagonale. De plus, les valeurs propres de  $A$  sont **imaginaires pures**.

*Démonstration.* — Dans chaque cas, l'assertion «  $A$  est diagonalisable dans une base orthonormée  $\mathcal{B}$  » découle du théorème précédent. Comme la base canonique  $\mathcal{B}_0$  est elle-même orthonormée, la matrice de passage  $P = \text{Mat}_{\mathcal{B}_0}(\mathcal{B})$  appartient à  $U(n)$ , d'où la 2ème assertion. Reste à voir l'assertion concernant les valeurs propres. On a

$$\bar{D} = {}^t\bar{D} = {}^t\bar{P} {}^t\bar{A} {}^t\bar{P}^{-1} = P^{-1} {}^t\bar{A} P = \begin{cases} P^{-1}AP = D & \text{si } A \text{ est hermitienne} \\ P^{-1}A^{-1}P = D^{-1} & \text{si } A \text{ est unitaire} \\ -P^{-1}AP = -D & \text{si } A \text{ est anti-hermitienne.} \end{cases}$$

Il en résulte que les termes diagonaux  $\lambda_i$  de  $D$  vérifient, respectivement,  $\bar{\lambda}_i = \lambda_i$  (resp.  $= \lambda_i^{-1}$ , resp.  $= -\lambda_i$ ), donc sont réels (resp. de module 1, resp. imaginaires purs). Le théorème est démontré.  $\square$

**Remarque 4.4.9.** — Le point (1) du théorème précédent fournit une autre démonstration de la proposition 2.2.8 et donc du théorème 2.2.6.

#### **4.5. Forme normale des éléments de $O(n)$**

On a décrit au Chap. 6 les éléments de  $O(2)$  et  $O(3)$ . On va voir maintenant la « forme normale » (= une matrice aussi simple que possible) des éléments de  $O(n)$ , pour  $n \geq 3$  arbitraire. On note  $\mathcal{B}_0$  la base canonique de  $\mathbb{R}^n$ .

**Théorème 4.5.1 (Forme normale des éléments de  $O(n)$ ).** — Soient  $A \in O(n)$  et  $f$  l'endomorphisme de  $V = \mathbb{R}^n$  associé à  $A$ . Notons  $V_+ = \text{Ker}(A - I_n)$  (resp.  $V_- = \text{Ker}(A + I_n)$ ) l'espace propre associé à la valeur propre 1 (resp.  $-1$ ) et  $p = \dim V_+$ ,  $q = \dim V_-$ . Alors il existe des bases orthonormées  $\mathcal{B}_+ = (x_1, \dots, x_p)$  de  $V_+$ ,  $\mathcal{B}_- = (y_1, \dots, y_q)$  de  $V_-$ , et  $\mathcal{C} = (v_1, u_1, \dots, v_r, u_r)$  de  $E = (V_+ \oplus V_-)^\perp$ , et  $\theta_1, \dots, \theta_r \in ]-\pi, \pi[-\{0\}$  telles que, notant  $\mathcal{B}$  la base orthonormée  $\mathcal{B}_+ \cup \mathcal{B}_- \cup \mathcal{C}$  de  $V$  et  $P$  la matrice de passage  $\text{Mat}_{\mathcal{B}_0}(\mathcal{B}) \in O(n)$ , on ait

$$P^{-1}AP = \text{Mat}_{\mathcal{B}}(f) = \begin{pmatrix} I_p & 0 & 0 & \cdots & 0 \\ 0 & -I_q & 0 & \cdots & 0 \\ 0 & 0 & R(\theta_1) & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 0 & R(\theta_r) \end{pmatrix}.$$

où  $R(\theta)$  désigne la matrice  $\begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \in O(2)$ . (En particulier,  $\dim E = 2r$  est paire). De plus,  $\theta_1, \dots, \theta_r \in ]-\pi, \pi[-\{0\}$  sont uniques au signe près.

Pour la démonstration, on a besoin du lemme suivant.

**Lemme 4.5.2.** — Soit  $u$  une isométrie de l'espace euclidien  $V$ , et soit  $F$  un sous-espace vectoriel stable par  $u$ , i.e.  $u(F) \subseteq F$ . Alors :

- (1) On a  $u(F) = F$  et donc  $u^{-1}(F) = F$ .
- (2) On a  $u(F^\perp) = F^\perp = u^{-1}(F^\perp)$ .

*Démonstration.* — D'abord l'isométrie  $u$  est bijective (cf. 2.1.9), donc en particulier injective, donc  $u(F)$  a même dimension que  $F$ . Par conséquent, l'inclusion  $u(F) \subseteq F$  entraîne  $u(F) = F$ , d'où aussi  $F = u^{-1}(F)$ . Ceci prouve (1). Le même argument montre que, pour prouver (2), il suffit de prouver que  $u(F^\perp) \subseteq F^\perp$ . Soient  $y \in F^\perp$  et  $x \in F$ , comme  $u$  est une isométrie, on a

$$(u(y) | x) = (y | u^{-1}(x)) = 0,$$

la 2ème égalité puisque  $u^{-1}(x) \in F$  d'après (1). Ceci montre que  $u(y) \in F^\perp$ , d'où l'assertion (2).  $\square$

*Démonstration.* — Commençons maintenant la démonstration du théorème 4.5.1. Comme  $V_+ \oplus V_-$  est stable par  $f$  alors, d'après le lemme, il en est de même de  $E = (V_+ \oplus V_-)^\perp$ . Notons  $f_E$  la restriction de  $f$  à  $E$ . Alors 1 et  $-1$  ne sont pas valeurs propres de  $f_E$ , puisque  $\text{Ker}(f_E - \text{id}_E) = \text{Ker}(f - \text{id}_V) \cap E = V_+ \cap E = \{0\}$  et de même  $\text{Ker}(f_E + \text{id}_E) = V_- \cap E = \{0\}$ .

Soit  $\mathcal{B}_+$  (resp.  $\mathcal{B}_-$ ) une b.o.n. de  $V_+$  (resp.  $V_-$ ). D'après 2.4.3, on sait que  $V_+$  et  $V_-$  sont orthogonaux, et que les valeurs propres réelles de  $f$  ne peuvent être que 1 et  $-1$ . Donc, d'une part,  $\mathcal{B}_+ \cup \mathcal{B}_-$  est une b.o.n. de  $V_+ \oplus V_-$  et, d'autre part,  $f_E$  n'a pas de valeurs propres réelles. Or, on a le lemme suivant :

**Lemme 4.5.3.** — *Soit  $W$  un espace euclidien de dimension  $m > 0$ , et soit  $f$  une isométrie de  $W$  n'ayant pas de valeurs propres réelles (i.e. telle que 1 et  $-1$  ne soient pas valeurs propres de  $f$ ). Alors il existe deux vecteurs unitaires et orthogonaux  $u$  et  $v$  et  $\theta \in ]-\pi, \pi[ - \{0\}$  tels que :*

$$f(u) = \cos(\theta)u - \sin(\theta)v, \quad f(v) = \sin(\theta)u + \cos(\theta)v$$

i.e. le plan  $P = \text{Vect}(u, v)$  est stable par  $f$  et l'on a

$$\text{Mat}_{(u,v)}(f_P) = \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix} \quad \text{Mat}_{(v,u)}(f_P) = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}.$$

En particulier, on a  $\dim W \geq 2$ .

Admettons pour le moment ce lemme, et achevons la démonstration du théorème 4.5.1. D'après le lemme précédent, il existe dans  $E$  un plan  $P_1$  stable par  $f$ , une b.o.n.  $\mathcal{C}_1 = (v_1, u_1)$  de  $P_1$  et  $\theta_1 \in ]-\pi, \pi[ - \{0\}$  tels que  $\text{Mat}_{\mathcal{C}_1}(f) = R(\theta_1)$ . Notons  $E_1$  l'orthogonal de  $P_1$  dans  $E$ , i.e. :

$$E_1 = \{x \in E \mid (x | y) = 0, \quad \forall y \in P_1\}.$$

D'après le lemme 4.5.2,  $E_1$  est stable par  $f$ . Bien sûr, la restriction  $f_{E_1}$  de  $f$  à  $E_1$  n'a pas de valeurs propres réelles (puisque  $f$  n'en avait pas) donc on peut à nouveau appliquer le lemme 4.5.3 : il existe dans  $E_1$  un plan  $P_2$  stable par  $f$ , une b.o.n.  $\mathcal{C}_2 = (v_2, u_2)$  de  $P_2$  et  $\theta_2 \in ]-\pi, \pi[ - \{0\}$  tels que  $\text{Mat}_{\mathcal{C}_2}(f) = R(\theta_2)$ . Notons  $E_2$  l'orthogonal de  $P_2$  dans  $E_1$ . Si  $E_2 \neq 0$ , on peut recommencer... On obtient ainsi qu'il existe une b.o.n.

$$\mathcal{C} = \mathcal{C}_1 \cup \dots \cup \mathcal{C}_r = (v_1, u_1, \dots, v_r, u_r)$$

de  $E$  (en particulier,  $\dim E = 2r$  est pair) et  $\theta_1, \dots, \theta_r \in ]-\pi, \pi[ - \{0\}$  tels que

$$(*) \quad \text{Mat}_{\mathcal{C}}(f_E) = \left( \begin{array}{c|c|c} R(\theta_1) & 0 & 0 \\ \hline 0 & \ddots & 0 \\ \hline 0 & 0 & R(\theta_r) \end{array} \right)$$

et alors  $\mathcal{B} = \mathcal{B}_+ \cup \mathcal{B}_- \cup \mathcal{C}$  est un b.o.n. de  $V$  telle que  $\text{Mat}_{\mathcal{B}}(f)$  ait la forme indiquée. Ceci prouve l'existence.

Montrons l'unicité au signe près de  $\theta_1, \dots, \theta_r$ , i.e. montrons l'unicité des paires  $\pm\theta_1, \dots, \pm\theta_r$ . Comme le polynôme caractéristique de  $R(\theta)$  est

$$X^2 - 2\cos(\theta)X + 1 = (X - e^{i\theta})(X - e^{-i\theta})$$

alors (\*) ci-dessus montre que le polynôme caractéristique de  $f_E$  est  $\prod_{s=1}^r ((X - e^{i\theta_s})(X - e^{-i\theta_s}))$  et que ses racines dans  $\mathbb{C}$  sont :

$$e^{i\theta_1}, e^{-i\theta_1}, \dots, e^{i\theta_r}, e^{-i\theta_r},$$

et donc  $\pm\theta_1, \dots, \pm\theta_r$  sont uniquement déterminés. Enfin, en général on ne peut pas faire mieux que de déterminer les  $\theta_s$  au signe près, puisque dans la base  $(u_s, v_s)$  la matrice de  $f_{P_s}$  est  $R(-\theta_s)$ . Ceci achève la démonstration du théorème 4.5.1, modulo la démonstration du lemme 4.5.3.  $\square$

Avant de démontrer le lemme 4.5.3, faisons les remarques suivantes.

**Remarques 4.5.4.** — (1) En dimension 2, on détermine le signe de  $\theta$  en choisissant une orientation de  $E$ , donnée par le choix d'une b.o.n.  $\mathcal{B}_0$  de  $E$ . Alors pour toute b.o.n.  $\mathcal{B}$  directe (i.e. telle que  $\det_{\mathcal{B}_0}(\mathcal{B}) = 1$ ), on a  $\text{Mat}_{\mathcal{B}}(f) = R(\theta)$  (cf. 2.4.14).

(2) De même, en dimension 3, on choisit l'orientation de  $\mathbb{R}^3$  donnée par la base canonique  $\mathcal{B}_0$ . Si  $f \in \text{SO}(3)$  et  $f \neq \text{id}$ , alors  $f$  possède un « axe de rotation »  $D = \text{Ker}(f - \text{id})$  qui est une droite vectorielle; on oriente  $D$  en choisissant un vecteur unitaire  $u \in D$ . Ayant fait ces choix, « l'angle de rotation »  $\theta$  est uniquement déterminé par la condition que pour toute b.o.n.  $(v_1, v_2)$  du plan  $P = D^\perp$  telle que la b.o.n.  $(v_1, v_2, u)$  de  $\mathbb{R}^3$  soit directe, on a  $\text{Mat}_{(v_1, v_2)}(f_P) = R(\theta)$  (cf. 2.4.18).

(3) *Attention!* En dimension paire  $\geq 4$ , une rotation ne possède pas nécessairement d'« axe de rotation », i.e. on peut avoir  $\text{Ker}(f - \text{id}) = \{0\}$ , c'est le cas par exemple pour

$$f = \left( \begin{array}{c|c} R(\theta_1) & 0 \\ \hline 0 & R(\theta_2) \end{array} \right) \in O(4)$$

avec  $\theta_1, \theta_2 \in ]-\pi, \pi[ - \{0\}$ .

*Démonstration.* — Démontrons maintenant le lemme 4.5.3. Fixons une base  $\mathcal{B} = (e_1, \dots, e_m)$  de  $W$ , ce qui permet d'identifier  $W$  à  $\mathbb{R}^m$  et  $f$  à la matrice  $A = \text{Mat}_{\mathcal{B}}(f) \in M_m(\mathbb{R})$ . On plonge  $\mathbb{R}^m$  dans  $\mathbb{C}^m$ , c.-à-d., on écrit :

$$\mathbb{R}^m = \{(x_1, \dots, x_m) \in \mathbb{C}^m \mid x_i \in \mathbb{R}, \quad \forall i = 1, \dots, m\}.$$

Alors, tout  $w = (z_1, \dots, z_m) \in \mathbb{C}^m$  s'écrit de façon unique

$$w = u + iv \quad \text{avec} \quad u, v \in \mathbb{R}^m : \quad \text{on a} \quad \begin{cases} u = (x_1, \dots, x_m) & \text{avec } x_j = \mathcal{R}(z_j) \\ v = (y_1, \dots, y_m) & \text{avec } y_j = \mathcal{I}(z_j). \end{cases}$$

On notera  $\boxed{u = \mathcal{R}(w)}$  et  $\boxed{v = \mathcal{I}(w)}$ . Si  $\lambda = a + ib \in \mathbb{C}$  (avec  $a, b \in \mathbb{R}$ ), alors  $\lambda w$  est le vecteur :

$$(1) \quad (a + ib)(u + iv) = \underbrace{(au - bv)}_{\in \mathbb{R}^m} + i \underbrace{(bu + av)}_{\in \mathbb{R}^m}.$$

Si l'on note  $\bar{w}$  le vecteur  $u - iv$ , on a donc

$$(2) \quad \bar{\lambda w} = (a - ib)(u - iv) = \underbrace{(au - bv)}_{\in \mathbb{R}^m} - i \underbrace{(bu + av)}_{\in \mathbb{R}^m} = \overline{\lambda w}.$$

Plus généralement, si  $B \in M_m(\mathbb{R})$ , alors les vecteurs  $Bu$  et  $Bv$  appartiennent encore à  $\mathbb{R}^m$ , et l'on a :

$$(3) \quad Bw = B(u + iv) = Bu + iBv \quad \text{et} \quad B\bar{w} = B(u - iv) = Bu - iBv = \overline{Bw}.$$

Appliquons ce qui précède dans le cas suivant. Soit  $\lambda = a + ib \in \mathbb{C} - \mathbb{R}$  une valeur propre de  $A$ , et soit  $w \in \mathbb{C}^m$  un vecteur propre associé. Écrivons  $w = u + iv$ , avec  $u, v \in \mathbb{R}^m$ . Alors

$$(4) \quad Au + iAv = Aw = \lambda w = (au - bv) + i(bu + av) \quad \text{d'où} \quad \begin{cases} Au = au - bv \\ Av = bu + av. \end{cases}$$

D'autre part, d'après (2) et (3) on a :

$$(5) \quad A\bar{w} = \overline{Aw} = \overline{\lambda w} = \bar{\lambda} \bar{w}$$

donc  $\bar{w}$  est vecteur propre de  $A$  pour la valeur propre  $\bar{\lambda}$ . Donc, puisque  $\bar{\lambda} \neq \lambda$  (car  $\lambda \notin \mathbb{R}$ ), les vecteurs propres  $w$  et  $\bar{w}$  sont linéairement indépendants sur  $\mathbb{C}$ . Comme  $w = u + iv$  et  $\bar{w} = u - iv$ , on en déduit que  $u, v$  sont linéairement indépendants sur  $\mathbb{C}$  (sinon  $w$  et  $\bar{w}$  seraient liés), donc a fortiori sur  $\mathbb{R}$ .

Il en résulte que le sous-espace vectoriel  $P = \mathbb{R}u + \mathbb{R}v$  de  $\mathbb{R}^m$  est de dimension 2, et d'après (4) il est stable par  $f$  et l'on a :

$$(6) \quad \text{Mat}_{(u,v)}(f_P) = \begin{pmatrix} a & b \\ -b & a \end{pmatrix} \quad \text{Mat}_{(v,u)}(f_P) = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}.$$

Tout ce qui précède est valable pour une matrice  $A \in M_m(\mathbb{R})$  arbitraire, une valeur propre complexe non réelle  $\lambda = a + ib$ , et un vecteur propre  $w = u + iv \in \mathbb{C}^m$  associé à  $\lambda$ .

Utilisons maintenant l'hypothèse additionnelle  $A \in O(m)$ , i.e.  ${}^tAA = I_m$ . On étend le produit scalaire euclidien  $\langle | \rangle$  sur  $\mathbb{R}^m$  en le produit scalaire hermitien standard sur  $\mathbb{C}^m$ , qu'on notera  $\langle | \rangle$ , i.e.

$$\forall W = \begin{pmatrix} w_1 \\ \vdots \\ w_m \end{pmatrix}, \quad \forall Z = \begin{pmatrix} z_1 \\ \vdots \\ z_m \end{pmatrix} \in \mathbb{C}^m, \quad \langle W | Z \rangle = {}^tW\bar{Z} = w_1 \bar{z}_1 + \dots + w_m \bar{z}_m.$$

Remarquons d'abord que si  $W, Z \in \mathbb{R}^m$ , alors  $\langle W | Z \rangle = (W | Z)$ , i.e. la restriction à  $\mathbb{R}^m$  de  $\langle | \rangle$  coïncide avec le produit scalaire euclidien. De plus, si l'on écrit  $W = U + iV$  et  $Z = X + iY$ , avec  $U, V, X, Y \in \mathbb{R}^m$ , alors  $\langle W | Z \rangle$  égale :

$$\langle U + iV | X + iY \rangle = \langle U | X \rangle + \langle V | Y \rangle + i(\langle V | X \rangle - \langle U | Y \rangle) = (U | X) + (V | Y) + i[(V | X) - (U | Y)].$$

D'autre part, comme  $A = \bar{A}$  et  ${}^tAA = I_m$ , on a  $A \in U(m)$  et donc, pour tout  $W, Z \in \mathbb{C}^m$  :

$$(7) \quad \langle AW | AZ \rangle = \langle W | Z \rangle.$$

Soient  $\lambda, w$  comme plus haut, avec  $Aw = \lambda w$ . On a vu qu'on a aussi  $A\bar{w} = \bar{\lambda} \bar{w}$ , donc  $w$  et  $\bar{w}$  appartiennent aux espaces propres  $V_\lambda$  et  $V_{\bar{\lambda}}$ , qui sont orthogonaux, d'après 4.4.8. Écrivant  $w = u + iv$ , avec  $u, v \in \mathbb{R}^m$ , on a donc :

$$0 = \langle u + iv | u - iv \rangle = (u | u) - (v | v) + 2i(u | v),$$

d'où  $(u | v) = 0$  et  $(u | u) = (v | v)$ , donc  $u, v \in \mathbb{R}^m$  sont orthogonaux et de même norme pour le produit scalaire euclidien. Remplaçant  $w$  par  $\frac{1}{\|u\|}w$ , on se ramène alors au cas où  $u, v$  sont orthogonaux et unitaires. Alors

$\mathcal{E} = (v, u)$  est une base orthonormée du plan  $P$ , et d'après (6) on a  $\text{Mat}_{\mathcal{E}}(f_P) = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$ , avec  $a, b \in \mathbb{R}$ .

Enfin, comme  $f$  est une isométrie et n'a pas  $\pm 1$  comme valeurs propres, il en est de même de  $f_P$ , et donc  $a^2 + b^2 = 1$  et il existe un unique  $\theta \in ]-\pi, \pi[ - \{0\}$  tel que  $a = \cos \theta$  et  $b = \sin \theta$  (d'où  $\lambda = e^{i\theta}$ ). Ceci achève la preuve du lemme 4.5.3 et donc du théorème 4.5.1.  $\square$

#### 4.6. Appendice (†) : espaces préhilbertiens réels ou complexes

Si  $E$ , muni de  $( | )$ , est un  $\mathbb{C}$ -espace vectoriel (resp.  $\mathbb{R}$ -espace vectoriel) de dimension infinie muni d'un produit scalaire hilbertien (resp. euclidien), on dit que  $E$  est un espace **préhilbertien** complexe (resp. réel). Dans ce cas on sait, d'après l'inégalité de Cauchy-Schwarz, que  $\|x\| = \sqrt{(x | x)}$  est une *norme* sur  $E$ . On dit alors que  $E$  est un espace hilbertien complexe (resp. réel) s'il est **complet** pour cette norme, i.e. si toute suite de Cauchy converge (ceci est automatiquement vérifié lorsque  $E$  est de dimension finie).

Ces espaces jouent un rôle important en Analyse. Par exemple, le  $\mathbb{R}$ -espace vectoriel  $E = \mathcal{C}^0([0, 1], \mathbb{R})$  des fonctions continues  $f : [0, 1] \rightarrow \mathbb{R}$ , muni du produit scalaire euclidien

$$(f | g) = \int_0^1 f(t)g(t)dt,$$

est un espace préhilbertien réel. Il n'est pas complet pour la norme euclidienne  $\|f\|_2 = \int_0^1 f^2(t)dt$ , mais il se plonge dans l'espace  $L^2([0, 1], \mathbb{R})$  des (classes de) fonctions  $f : [0, 1] \rightarrow \mathbb{R}$  qui sont de carré intégrable au sens de Lebesgue (i.e.  $f$  est mesurable et  $\int_0^1 f^2(t)dt < \infty$ ), et  $L^2([0, 1], \mathbb{R})$  muni du produit scalaire euclidien

$$(f | g) = \int_0^1 f(t)g(t)dt,$$

est un espace de Hilbert réel, i.e. il est complet pour la norme  $\|f\|_2 = \int_0^1 f^2(t)dt$ . (On parle ici de classes de fonctions, car on identifie deux fonctions  $f$  et  $g$  si elles coïncident en dehors d'un ensemble de mesure nulle, i.e. si  $f - g$  est nulle presque partout.)

De même, le  $\mathbb{C}$ -espace vectoriel  $E = \mathcal{C}^0([0, 1], \mathbb{C})$  des fonctions continues  $f : [0, 1] \rightarrow \mathbb{C}$ , muni du produit scalaire hilbertien

$$(f | g) = \int_0^1 f(t)\overline{g(t)} dt,$$

est un espace préhilbertien complexe. Il n'est pas complet pour la norme hilbertienne  $\|f\|_2 = \int_0^1 |f(t)|^2 dt$ , mais il se plonge dans l'espace  $L^2([0, 1], \mathbb{C})$  des (classes de) fonctions  $f : [0, 1] \rightarrow \mathbb{C}$  qui sont de carré intégrable au sens de Lebesgue, et  $L^2([0, 1], \mathbb{C})$  muni du produit scalaire hilbertien

$$(f | g) = \int_0^1 f(t)\overline{g(t)} dt,$$

est un espace de Hilbert complexe, i.e. il est complet pour la norme  $\|f\|_2 = \int_0^1 |f(t)|^2 dt$ .

