

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

**LM270 Algèbre et géométrie**

par Laurent KOELBLEN et Patrick POLO

Laurent Koelblen et Patrick Polo  
Université P. & M. Curie – Paris 6  
Institut de Mathématiques de Jussieu (UMR 7586 du CNRS)  
*URL* : <http://people.math.jussieu.fr/~koelblen>  
<http://people.math.jussieu.fr/~polo>  
*E-méls* : [koelblen@math.jussieu.fr](mailto:koelblen@math.jussieu.fr), [polo@math.jussieu.fr](mailto:polo@math.jussieu.fr)

# GUIDE D'UTILISATION DU POLYCOPIÉ À L'INTENTION DES ÉTUDIANTS DE L2

<sup>(1)</sup> L'assimilation d'un nouveau cours nécessite un travail personnel important. Cela passe par l'assiduité en cours (écouter un exposé oral stimule d'autres canaux d'apprentissage que la lecture...!), la lecture et la relecture de ses notes de cours et du polycopié, ainsi que la résolution des exercices proposés.

Dans ce polycopié, on a indiqué par des symboles  les définitions, exemples et résultats fondamentaux ; il faut essayer de se les approprier !

Pour cela, il peut être utile de chercher par soi-même des situations (aussi simples que possibles) où un énoncé nouvellement introduit s'applique, de chercher à voir si on pourrait arriver à la même conclusion sans utiliser ledit énoncé nouveau, de chercher à voir si la conclusion reste valable si l'on affaiblit les hypothèses, éventuellement de chercher des contre-exemples, etc. Les exercices proposés dans le polycopié, les séances de travaux dirigés et les sujets de contrôle continu, ont pour but de vous aider à effectuer ce cheminement.

Le jeu est à la fois plus difficile et plus intéressant que s'il s'agissait de régurgiter des données apprises par coeur et des exercices formatés. Si l'on n'y prend pas goût et si l'on n'y perçoit qu'élucubrations abstraites et inutiles, mieux vaut changer de cours ! Nous espérons au contraire que ce cours suscitera de l'intérêt, voire même parfois du plaisir ? Bon travail !

REMERCIEMENTS. Pour cette version 2012-2013, nous avons bénéficié des corrections et suggestions d'Ahmed Moussaoui et Thomas de La Rochefoucauld. Qu'ils en soient ici remerciés.

---

<sup>(1)</sup>Ce « Guide d'utilisation » tire son origine de l'« Avertissement au lecteur étudiant » rédigé par Jacques Féjoz dans le polycopié de son cours LM121 2007-2008 ; avec son aimable autorisation, nous l'avons repris à notre compte !



# TABLE DES MATIÈRES


<b>0. Rappels : espaces vectoriels et applications linéaires</b> .....	1
0.1. Espaces vectoriels : définition et exemples .....	1
0.2. Familles génératrices, familles libres, bases et dimension .....	3
0.3. Noyau, image, et théorème du rang .....	6
0.4. Applications linéaires et matrices .....	7
0.5. Changements de base .....	10
0.6. Appendice (†) : compléments sur les familles génératrices ou libres .....	14
<b>1. Dual, opérations sur les colonnes ou les lignes, déterminants, valeurs et vecteurs propres</b> .....	19
1.1. Formes linéaires, espace dual .....	19
1.2. Opérations élémentaires sur les colonnes (ou les lignes) .....	21
1.3. Codimension et équations d'un sous-espace .....	29
1.4. Déterminants .....	31
1.5. Endomorphismes : déterminant, trace, valeurs propres, etc. ....	38
1.6. Appendice (†) : Transposée d'une application linéaire .....	39
1.7. Appendice (†) : Bidual .....	40
<b>2. Rappels sur les endomorphismes diagonalisables, trigonalisation, Cayley-Hamilton, espaces caractéristiques</b> .....	43
2.1. Espaces propres et critères de diagonalisabilité .....	43
2.2. Trigonalisation, théorème de Cayley-Hamilton et espaces caractéristiques .....	47
2.3. Appendice (†) : somme directe externe d'espaces vectoriels .....	52
2.4. Appendice (†) : division euclidienne dans $\mathbb{C}[X]$ et théorème de Bézout .....	53
2.5. Appendice (†) : $\mathbb{C}$ est algébriquement clos .....	54
<b>3. Décomposition de Jordan, exponentielles de matrices, espaces quotients</b> .....	57
3.1. Endomorphismes nilpotents, partitions et formes normales de Jordan .....	57
3.2. Décomposition de Dunford .....	67
3.3. Exponentielles de matrices .....	69
3.4. Exponentielles de matrices et équations différentielles linéaires .....	73
3.5. Espaces quotients .....	75
3.6. Appendice (†) : Normes sur $\mathbb{K}^n$ et produits de séries absolument convergentes .....	78
<b>4. Applications multilinéaires et formes <math>p</math>-linéaires alternées</b> .....	81
4.1. Applications multilinéaires .....	81
4.2. Anneaux de polynômes .....	84
4.3. Groupes symétriques : quelques propriétés .....	87
4.4. Applications multilinéaires alternées .....	90
4.5. Appendice (†) : Bases et dimension de $\mathcal{L}(E_1, \dots, E_p; F)$ .....	93
4.6. Appendice (†) : Produit tensoriel .....	95
<b>5. Formes bilinéaires symétriques et formes quadratiques</b> .....	97
5.0. Introduction .....	97
5.1. Formes bilinéaires symétriques et formes quadratiques .....	97

5.2. Réduction d'une forme quadratique en « somme de carrés »	104
5.3. Appendice (†) : formes bilinéaires alternées	107
<b>6. Espaces euclidiens et groupes orthogonaux <math>O(n)</math></b>	111
6.1. Espaces euclidiens. Inégalité de Cauchy-Schwarz. Isométries	111
6.2. Endomorphismes auto-adjoints et théorème de diagonalisation simultanée	115
6.3. Orthogonalité. Orthonormalisation de Gram-Schmidt	119
6.4. Bases directes ou indirectes. Groupes $O(n)$ et $SO(n)$ . Étude de $O(2)$ et $O(3)$	122
6.5. Appendice (†) : mesure des angles dans $\mathbb{R}^2$	134
6.6. Appendice (†) : décomposition d'Iwasawa de $GL_n(\mathbb{R})$	135
<b>7. Espaces affines, coniques et quadriques</b>	137
7.1. Espaces affines réels	137
7.2. Barycentres et sous-espaces affines	141
7.3. Projections, symétries, points fixes	146
7.4. Espaces affines euclidiens	147
7.5. Coniques	151
7.6. Quadriques en dimension 3	157
<b>8. Formes hermitiennes, espaces hilbertiens et groupes unitaires <math>U(n)</math></b>	161
8.0. Rappels sur les nombres complexes	161
8.1. Formes hermitiennes	161
8.2. Réduction en sommes de carrés de modules	167
8.3. Espaces hilbertiens. Inégalité de Cauchy-Schwarz. Isométries	169
8.4. Diagonalisation des endomorphismes auto-adjoints et normaux	173
8.5. Forme normale des éléments de $O(n)$	176
8.6. Appendice (†) : espaces préhilbertiens réels ou complexes	179
<b>Index</b>	181

## CHAPITRE 0

### RAPPELS : ESPACES VECTORIELS ET APPLICATIONS LINÉAIRES

Ce chapitre 0 est constitué de **rappels** : espaces vectoriels, familles génératrices, familles libres, bases et dimension, applications linéaires et matrices, transposée d'une matrice, théorème du rang, formules de changement de base, matrices équivalentes, matrices semblables. Ces notions ont été introduites en L1 et ne seront pas traitées en cours, mais le lecteur pourra se reporter, si nécessaire, à ce chapitre 0 pour un rappel de ces notions ou résultats. On suppose également connue la théorie du pivot pour la résolution des systèmes linéaires  $AX = Y$  ; ceci équivaut à faire des opérations sur les lignes de la matrice  $A$  et on verra dans le chapitre 1 comment faire des opérations sur les *colonnes* (ce qui est plus pratique pour la recherche de vecteurs propres).

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 un appendice à la fin du chapitre ; ces passages n'interviendront pas dans les évaluations.

#### 0.1. Espaces vectoriels : définition et exemples

**0.1.1. Trois exemples importants.** — (1) Un exemple d'espace vectoriel sur  $\mathbb{R}$  est l'espace de dimension 3

$$\mathbb{R}^3 = \{(x, y, z) \mid x, y, z \in \mathbb{R}\}.$$

Dans ce cas, l'espace vectoriel nous est donné comme un ensemble de  $n$ -uplets (ici  $n = 3$ ) de « coordonnées ».

Deux autres exemples sont les suivants.

(2) Soient  $a, b \in \mathbb{R}$  ; l'ensemble  $\mathcal{S}(a, b)$  des suites  $(u_n)_{n \in \mathbb{N}}$  de nombres réels vérifiant la relation de récurrence linéaire

$$(*) \quad u_{n+2} = au_{n+1} + bu_n$$

est un  $\mathbb{R}$ -espace vectoriel, qui est de dimension 2, car toute suite vérifiant (\*) est déterminée par ses termes initiaux  $u_0$  et  $u_1$ , qui peuvent être choisis arbitrairement.

(3) Soient  $a, b \in \mathbb{R}$  et  $t_0 \in \mathbb{R}$  ; l'ensemble  $\mathcal{S}(a, b)$  des fonctions  $f : \mathbb{R} \rightarrow \mathbb{R}$  de classe  $C^\infty$ , vérifiant l'équation différentielle linéaire :

$$(E) \quad \forall t \in \mathbb{R}, \quad f''(t) = af'(t) + bf(t)$$

est un  $\mathbb{R}$ -espace vectoriel, qui est de dimension 2, car toute solution  $f$  de (E) est déterminée par les « conditions initiales »  $f(t_0)$  et  $f'(t_0)$ , qui peuvent être choisies arbitrairement.

Dans ces deux cas, le choix de « coordonnées » sur l'espace  $\mathcal{S}(a, b)$  n'est pas évident . . . Une des forces de l'algèbre linéaire est qu'elle permet de décrire simplement tous les éléments de  $\mathcal{S}(a, b)$ , une fois qu'on a choisi une base appropriée de cet espace . . .

---

Rappelons que la notion d'espace vectoriel sur un corps  $k$  est définie pour tout corps  $k$ , par exemple,  $k = \mathbb{Q}$  ou  $\mathbb{C}$ , ou un corps fini  $\mathbb{F}_q$  à  $q$  éléments ( $q$  étant une puissance d'un nombre premier  $p$ ), ou le corps des fractions rationnelles  $\mathbb{R}(X)$ , etc.

**Définition 0.1.2 (Espaces vectoriels).** — Soit  $k$  un corps. Un  $k$ -espace vectoriel  $V$  est un *groupe abélien*  $(V, +)$  (c.-à-d., un ensemble muni d'une loi de groupe  $+$  commutative) muni d'une « opération »  $(t, v) \mapsto t \cdot v$  de  $k$  sur  $V$  vérifiant les deux conditions suivantes :

- (i)  $1 \cdot v = v$  et  $t \cdot (t' \cdot v) = (tt') \cdot v$
- (ii)  $(t + t') \cdot v = t \cdot v + t' \cdot v$ ,  $t \cdot (v + v') = t \cdot v + t \cdot v'$ .

On peut mémoriser la condition (i) en disant que 1 agit par l'identité et que l'opération est « associative », et la condition (ii) en disant que l'action de  $k$  sur  $V$  est « compatible avec l'addition » (dans  $k$  et dans  $V$ ).

**Remarque 0.1.3 (Vecteur nul).** — Étant un groupe abélien,  $V$  est muni d'un élément zéro, qu'on notera provisoirement  $0_V$  ou  $\vec{0}$ . Par exemple, dans  $V = \mathbb{R}^3$ ,  $0_V$  est le vecteur nul

$$\vec{0} = (0, 0, 0).$$

Notons 0 l'élément zéro du corps  $k$ . Alors la condition (ii) entraîne, pour tout  $v \in V$ ,

$$0 \cdot v = (0 + 0) \cdot v = 0 \cdot v + 0 \cdot v, \quad \text{d'où } 0 \cdot v = 0_V = \vec{0}.$$

Par conséquent, le vecteur nul  $0_V = \vec{0}$  sera noté simplement (par abus de notation) 0. Ainsi, on note  $\{0\}$  l'espace vectoriel nul. Par exemple, dans  $\mathbb{R}^2$  l'espace des solutions du système

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

est l'espace vectoriel nul  $\{0\} = \{(0, 0)\}$ .

**Terminologie 0.1.4.** — On dira que  $k$  est (relativement à  $V$ ) le « corps des scalaires », et que les éléments de  $k$  sont les scalaires (ceux de  $V$  étant les vecteurs).

**Exemples 0.1.5.** — 1) Pour tout  $n \in \mathbb{N}^*$ ,  $k^n = \{(x_1, \dots, x_n) \mid x_i \in k\}$  est un  $k$ -espace vectoriel.

2) L'ensemble  $M_{m,n}(k)$  des matrices à  $m$  lignes et  $n$  colonnes à coefficients dans  $k$  est un  $k$ -espace vectoriel. Lorsque  $m = n$ , on le note simplement  $M_n(k)$ .

3) L'anneau de polynômes  $k[X]$  est un  $k$ -espace vectoriel.

**Définition 0.1.6 (Sous-espaces vectoriels).** — Soit  $V$  un  $k$ -espace vectoriel. Un *sous-espace vectoriel* (pour abrégé, on dira « sev »)  $W$  de  $V$  est un sous-ensemble de  $V$  qui est un sous-groupe (en particulier,  $0 \in W$ ) et qui est stable par l'opération de  $k$ . Ceci équivaut à dire que  $W \neq \emptyset$  et que, pour tous  $w, w' \in W$  et  $t \in k$ , on a  $t \cdot w + w' \in W$ .

**Exemples 0.1.7.** — (1) Le plan « horizontal »

$$P = \{(x, y, 0) \mid x, y \in \mathbb{R}\},$$

donné par l'équation  $z = 0$ , est un sous-espace vectoriel de  $\mathbb{R}^3$ .

(2) L'ensemble des matrices  $(a_{ij})_{1 \leq i, j \leq 3} \in M_3(\mathbb{R})$  qui sont triangulaires supérieures, i.e. telles que  $a_{21} = a_{31} = a_{32} = 0$ , est un sous-espace vectoriel de  $M_3(\mathbb{R})$ .

(3) L'ensemble  $\mathbb{R}_n[X]$  des polynômes en  $X$  de degré  $\leq n$  est un sous-espace vectoriel de  $\mathbb{R}[X]$ .

**Définition 0.1.8 (Applications linéaires et endomorphismes).** — Soient  $k$  un corps,  $V, W$  deux  $k$ -espaces vectoriels. On dit qu'une application  $\phi : V \rightarrow W$  est une *application linéaire* (ou « homomorphisme d'espaces vectoriels ») si elle préserve l'addition et l'opération des scalaires, c.-à-d., si pour tout  $v, v' \in V$  et  $t \in k$  on a :

$$\phi(v + v') = \phi(v) + \phi(v'), \quad \phi(t \cdot v) = t \cdot \phi(v).$$

Notons qu'on peut regrouper ces deux conditions en une seule condition :

$$(AL) \quad \phi(t \cdot v + v') = t \cdot \phi(v) + \phi(v')$$

et bien sûr cette condition implique (et est impliquée par) la suivante :

$$\phi(t \cdot v + t' \cdot v') = t \cdot \phi(v) + t' \cdot \phi(v').$$

Si  $W = V$ , on dit alors que  $\phi$  est un *endomorphisme* de  $V$ .



**Exemples 0.1.9.** — a) Soit  $V = \mathbb{R}[X]$ ; l'application  $d$  qui à tout polynôme  $P = a_n X^n + \dots + a_1 X + a_0$  associe le polynôme dérivé  $P' = na_n X^{n-1} + \dots + a_1$  est une application linéaire de  $V$  dans  $V$ .

b) Soit  $V = \mathcal{C}([0, 1], \mathbb{R})$  l'espace des fonctions continues  $f : [0, 1] \rightarrow \mathbb{R}$ . Alors l'application

$$V \rightarrow \mathbb{R}, \quad f \mapsto \int_0^1 f(x) dx$$

est linéaire, mais l'application  $f \mapsto \int_0^1 f^2(x) dx$  ne l'est pas.

**Définition 0.1.10 (Isomorphismes).** — Soient  $V, W$  deux  $k$ -espaces vectoriels, et  $\phi : V \rightarrow W$  une application linéaire. Suivant des principes généraux, on dit que  $\phi$  est un *isomorphisme* (d'espaces vectoriels) si elle est *bijective* et si l'application inverse  $\psi = \phi^{-1}$  est linéaire.

En fait, la seconde condition est automatiquement vérifiée. En effet, soient  $w, w' \in W$  et  $t \in k$ ; comme  $\phi$  est bijective il existe  $v, v' \in V$  uniques tels que  $\phi(v) = w$  et  $\phi(v') = w'$ . Alors

$$\phi(t \cdot v + v') = t \cdot \phi(v) + \phi(v') = t \cdot w + w',$$

donc appliquant  $\psi$  à cette égalité on obtient

$$\psi(t \cdot w + w') = t \cdot v + v' = t \cdot \psi(w) + \psi(w').$$

Donc : toute application linéaire bijective est un isomorphisme.

**Exemple 0.1.11.** — L'ensemble  $k^{\mathbb{N}}$  de toutes les suites  $(u_n)_{n \in \mathbb{N}}$  d'éléments de  $k$  muni de l'addition et de l'opération de  $k$  définies « terme à terme », c.-à-d.,

$$(u_n)_{n \in \mathbb{N}} + (v_n)_{n \in \mathbb{N}} = (u_n + v_n)_{n \in \mathbb{N}}, \quad t \cdot (u_n)_{n \in \mathbb{N}} = (t \cdot u_n)_{n \in \mathbb{N}},$$

est un  $k$ -espace vectoriel. Soient  $a, b \in k$  et soit  $\mathcal{S}(a, b)$  le sous-ensemble de  $k^{\mathbb{N}}$  formé des suites  $(u_n)_{n \in \mathbb{N}}$  vérifiant la relation de récurrence linéaire :

$$(*) \quad u_{n+2} = au_{n+1} + bu_n.$$

Alors  $\mathcal{S}(a, b)$  est un sous-espace vectoriel de  $k^{\mathbb{N}}$ . De plus, l'application  $\phi : \mathcal{S}(a, b) \rightarrow k^2$  qui à toute suite  $(u_n)_{n \in \mathbb{N}}$  associe le couple  $(u_0, u_1)$ , est linéaire (le vérifier!); elle est surjective (car on peut choisir arbitrairement  $u_0$  et  $u_1$ ), et injective (car les  $u_n$  sont déterminés à partir de  $u_0$  et  $u_1$  par la formule (\*)). Donc  $\phi$  est bijective, donc c'est un isomorphisme d'espaces vectoriels :

$$\phi : \mathcal{S}(a, b) \xrightarrow{\sim} k^2.$$

**Exercices 0.1.12.** — 1) Soit  $k = \mathbb{R}$ . Est-ce que l'ensemble des suites  $(u_n)_{n \in \mathbb{N}}$  vérifiant la relation de récurrence

$$u_{n+2} = u_{n+1} + u_n^2$$

est un sous-espace vectoriel de  $\mathbb{R}^{\mathbb{N}}$  ?

2) Soient  $k = \mathbb{C}$  et  $\mathcal{S}(-1, -1)$  l'espace des suites de nombres complexes  $(u_n)_{n \in \mathbb{N}}$  vérifiant la relation de récurrence linéaire

$$u_{n+2} = -u_{n+1} - u_n.$$

Soit  $(w_n)_{n \in \mathbb{N}}$  l'élément de  $\mathcal{S}(-1, -1)$  défini par  $w_0 = 2$  et  $w_1 = 1$ . Pouvez-vous calculer  $w_{2010}$  et  $w_{2011}$  ? (Cf. la Feuille d'exercices 1.)

## 0.2. Familles génératrices, familles libres, bases et dimension

**Définitions 0.2.1 (Sous-espace engendré. Familles génératrices).** — Soit  $V$  un  $k$ -espace vectoriel.

(1) Soit  $S$  un sous-ensemble (fini ou infini) de  $V$ . On note  $\boxed{\text{Vect}(S)}$  l'ensemble de toutes les combinaisons linéaires *finies*

$$\sigma = t_1 v_1 + \dots + t_r v_r, \quad \text{pour } r \in \mathbb{N}^* \text{ (variable), } v_i \in S, t_i \in k,$$

c'est un sous-espace vectoriel de  $V$ , car si  $\sigma' = t'_1 v'_1 + \dots + t'_p v'_p$  est une autre combinaison linéaire de ce type et si  $\lambda \in k$ , alors

$$\lambda \sigma + \sigma' = \lambda t_1 v_1 + \dots + \lambda t_r v_r + t'_1 v'_1 + \dots + t'_p v'_p$$

est encore une combinaison linéaire du même type. De plus, si  $E$  est un sous-espace vectoriel de  $V$  contenant  $S$ , alors il contient toute combinaison linéaire  $\sigma$  comme ci-dessus, i.e. il contient  $\text{Vect}(S)$ . Donc :

$\boxed{\text{Vect}(S)}$  est le plus petit sous-espace vectoriel de  $V$  contenant  $S$ . On l'appelle le sous-espace vectoriel *engendré* par  $S$ .



Dans la suite, on utilisera principalement ceci dans le cas où  $S$  est une famille *finie* de vecteurs  $v_1, \dots, v_n$  ; dans ce cas, on a simplement :

$$\text{Vect}(v_1, \dots, v_n) = \{t_1 v_1 + \dots + t_n v_n \mid t_1, \dots, t_n \in k\}.$$

(2) On dit que la famille  $\mathcal{F} = (v_1, \dots, v_n)$  est une *famille génératrice* de  $V$  si  $\text{Vect}(\mathcal{F}) = V$ , c.-à-d., si tout élément de  $V$  s'écrit comme combinaison linéaire de  $v_1, \dots, v_n$ . Dans ce cas, on dit aussi que les vecteurs  $v_1, \dots, v_n$  engendrent  $V$ .

Il résulte de la définition que : *toute famille contenant une famille génératrice est génératrice.*

**Exemple 0.2.2.** — Pour tout  $n \in \mathbb{N}$ , l'espace  $k^n$  est engendré par les vecteurs

$$e_1 = (1, 0, \dots, 0), \quad e_2 = (0, 1, 0, \dots, 0), \quad \dots, \quad e_n = (0, \dots, 0, 1).$$



**Définition 0.2.3 (Familles libres ou liées).** — Soient  $V$  un  $k$ -espace vectoriel et  $\mathcal{F} = (v_1, \dots, v_n)$  une famille d'éléments de  $V$ . On dit  $v_1, \dots, v_n$  sont *linéairement indépendants*, et que  $\mathcal{F}$  est une *famille libre*, s'il n'existe pas de relation linéaire non triviale entre les  $v_i$ , c.-à-d., si la condition suivante est vérifiée :

(FL) pour tous  $t_1, \dots, t_n \in k$ , si  $t_1 v_1 + \dots + t_n v_n = 0$ , alors  $t_1 = 0 = \dots = t_n$ .

Il résulte de la définition que : *toute sous-famille d'une famille libre est libre.*

Au contraire, on dit que  $v_1, \dots, v_n$  sont *linéairement dépendants*, et que  $\mathcal{F}$  est une *famille liée*, s'il existe une relation linéaire non triviale entre les  $v_i$ , c.-à-d., s'il existe des scalaires non tous nuls  $t_1, \dots, t_n$ , tels que  $t_1 v_1 + \dots + t_n v_n = 0$ . Dans ce cas, si par exemple  $t_i \neq 0$ , on peut exprimer  $v_i$  en fonction des  $v_j$ , pour  $j \neq i$  :

$$v_i = - \sum_{j \neq i} \frac{t_j}{t_i} v_j.$$

Il résulte de la définition que : *toute famille contenant une famille liée est liée.*

**Exemple 0.2.4.** — Dans  $k^n$ , la famille  $(e_1, \dots, e_n)$  (cf. 0.2.2) est libre. En effet, pour tous  $t_1, \dots, t_n \in k$  on a

$$t_1 e_1 + \dots + t_n e_n = (t_1, \dots, t_n),$$

donc si la somme de gauche est nulle, alors  $t_1 = 0 = \dots = t_n$ .



**Définition 0.2.5 (Bases).** — Soient  $V$  un  $k$ -espace vectoriel. On dit qu'une famille  $\mathcal{B} = (v_1, \dots, v_n)$  est une *base* de  $V$  si tout élément  $v$  de  $V$  s'écrit de façon unique comme combinaison linéaire des  $v_i$ , c.-à-d., si pour tout  $v \in V$ , il existe un unique  $n$ -uplet  $(t_1, \dots, t_n) \in k^n$  tel que  $v = t_1 v_1 + \dots + t_n v_n$ . Ceci équivaut à dire que la famille  $\mathcal{B}$  est à la fois génératrice et libre.

**Exemples 0.2.6.** — 1) Lorsque  $V = k^n$ , la famille  $(e_1, \dots, e_n)$  (cf. 0.2.2) est une base, appelée la *base canonique de  $k^n$* .



2) Soit  $M_{m,n}(k)$  le  $k$ -espace vectoriel des matrices à  $m$  lignes et  $n$  colonnes à coefficients dans  $k$ . On note  $E_{ij}$  la « *matrice élémentaire* » dont tous les coefficients sont nuls, sauf celui d'indice  $(i, j)$  (c.-à-d., celui situé sur la ligne  $i$  et la colonne  $j$ ), qui vaut 1. Alors, toute matrice  $A \in M_{m,n}(k)$  s'écrit de façon unique comme combinaison linéaire des  $E_{ij}$  :

$$A = \sum_{\substack{i=1, \dots, m \\ j=1, \dots, n}} a_{ij} E_{ij},$$

où  $a_{ij}$  est le coefficient d'indice  $(i, j)$  de  $A$ . Donc la famille  $(E_{ij})_{\substack{i=1, \dots, m \\ j=1, \dots, n}}$  est une *base* de  $M_{m,n}(k)$ .

3) La famille  $(1, X, \dots, X^d)$  est une base de l'espace vectoriel  $k_d[X]$  des polynômes de degré  $\leq d$ . En effet, tout polynôme  $P \in k_d[X]$  s'écrit de façon unique

$$P = a_0 + \dots + a_d X^d, \quad \text{avec } a_i \in k.$$

**Définition 0.2.7.** — Soit  $V$  un  $k$ -espace vectoriel. Disons provisoirement que  $V$  est *finiment engendré* (ou « de type fini », cf. le cours LM 125) s'il est engendré par un nombre fini de vecteurs  $v_1, \dots, v_p$ .

Remarque : il existe des  $k$ -espaces vectoriels qui ne sont pas finiment engendrés, par exemple, l'espace vectoriel  $k[X]$  (cf. Exercice 19 de la Feuille 1), mais dans ce cours on s'intéressera à ceux qui le sont.

Rappelons le résultat suivant, déjà vu en LM 125 :



**Théorème 0.2.8 (Dimension d'un espace vectoriel).** — Soit  $V$  un  $k$ -espace vectoriel finiment engendré.

(i) Il existe des bases de  $V$ , et toutes ont même cardinal  $n$ ; cet entier s'appelle la dimension de  $V$  sur  $k$  et se note  $\dim_k V$  ou simplement  $\dim V$ .

(ii) De toute famille génératrice  $\mathcal{F}$  on peut extraire une base, en particulier  $\mathcal{F}$  est de cardinal  $\geq n$ ; de plus si  $\text{card}(\mathcal{F}) = n$  alors  $\mathcal{F}$  est une base de  $V$ .

(iii) Toute famille libre est de cardinal  $\leq n$ , et toute famille libre de cardinal  $n$  est une base de  $V$ .

(iv) « Théorème de la base incomplète » : Toute famille libre peut être complétée en une base de  $V$ .

(v) Tout sous-espace  $W$  de  $V$  est de dimension finie  $\leq \dim_k V$ ; de plus si  $\dim_k W = \dim_k V$ , alors  $W = V$ . En d'autres termes, tout sous-espace vectoriel distinct de  $V$  est de dimension  $< \dim_k V$ .

*Démonstration.* Ceci a été vu en L1. Pour être complet, on redonne la démonstration dans un appendice à la fin de ce chapitre, où on introduira aussi les notions de familles génératrices ou libres dans un espace vectoriel arbitraire (i.e. qui n'est pas nécessairement finiment engendré).



**Terminologie 0.2.9.** — En raison du théorème précédent, on dira désormais «  $k$ -espace vectoriel de dimension finie » au lieu de «  $k$ -espace vectoriel finiment engendré », et si  $n = \dim_k V$ , on dira que  $V$  est de dimension  $n$ .

D'après les exemples de 0.2.6,  $k^n$  est de dimension  $n$ ,  $M_{m,n}(k)$  de dimension  $mn$ , et  $k_d[X]$  de dimension  $d + 1$ .



**Définition 0.2.10 (Coordonnées relativement à une base).** — Soit  $V$  un  $k$ -espace vectoriel de dimension  $n$  et soit  $\mathcal{B} = (v_1, \dots, v_n)$  une base de  $V$ . Alors tout  $v \in V$  s'écrit de façon unique

$$v = x_1 v_1 + \dots + x_n v_n;$$

on dit que  $(x_1, \dots, x_n)$  sont les *coordonnées de  $v$  par rapport à la base  $\mathcal{B} = (v_1, \dots, v_n)$* . Donc la donnée de  $\mathcal{B}$  fournit un isomorphisme de  $k$ -espaces vectoriels

$$\phi_{\mathcal{B}} : k^n \xrightarrow{\sim} V, \quad (x_1, \dots, x_n) \mapsto x_1 v_1 + \dots + x_n v_n.$$

**Remarque 0.2.11.** — Remarquons qu'une base de  $V$  est un  $n$ -uplet ordonné; par exemple, si  $\mathcal{B} = (v_1, v_2)$  est une base de  $V$ , alors  $\mathcal{C} = (v_2, v_1)$  est une base de  $V$  distincte de  $\mathcal{B}$ : l'image de  $(1, 2) \in \mathbb{R}^2$  par  $\phi_{\mathcal{B}}$  est le vecteur  $v_1 + 2v_2$ , tandis que son image par  $\phi_{\mathcal{C}}$  est le vecteur  $v_2 + 2v_1 \neq v_1 + 2v_2$ .

**Proposition 0.2.12.** — Soit  $f : V \rightarrow W$  une application linéaire.

- Si  $f$  est injective et si  $\mathcal{F} = (v_1, \dots, v_n)$  est une famille libre de  $V$ , alors  $f(\mathcal{F})$  est libre.
- Si  $f$  est surjective et si  $\mathcal{F} = (v_1, \dots, v_n)$  est une famille génératrice de  $V$ , alors  $f(\mathcal{F})$  engendre  $W$ .
- Si  $f$  est bijective et si  $\mathcal{B} = (v_1, \dots, v_n)$  est une base de  $V$ , alors  $f(\mathcal{B})$  est une base de  $W$ , d'où  $\dim W = n = \dim V$ .

*Démonstration.* a) Supposons  $f$  injective et soit  $\mathcal{F} = (v_1, \dots, v_n)$  une famille libre de  $V$ . Supposons qu'il existe une relation linéaire dans  $W$  :

$$t_1 f(v_1) + \dots + t_n f(v_n) = 0,$$

avec  $t_i \in k$ . Alors  $0 = f(t_1 v_1 + \dots + t_n v_n)$  d'où, puisque  $f$  est injective,  $t_1 v_1 + \dots + t_n v_n = 0$ , donc comme  $\mathcal{F}$  est libre,  $t_i = 0$  pour  $i = 1, \dots, n$ . Ceci montre que  $f(\mathcal{F})$  est libre.

b) Supposons  $f$  surjective et soit  $\mathcal{F} = (v_1, \dots, v_n)$  une famille génératrice de  $V$ . Soit  $w \in W$ ; comme  $f$  est surjective, il existe  $v \in V$  tel que  $f(v) = w$ . Comme  $\mathcal{F}$  engendre  $V$ , il existe  $t_1, \dots, t_n \in k$  tels que  $v = t_1 v_1 + \dots + t_n v_n$ , d'où

$$w = t_1 f(v_1) + \dots + t_n f(v_n).$$

Ceci montre que  $f(\mathcal{F})$  engendre  $W$ .

c) Supposons  $f$  bijective et soit  $\mathcal{B} = (v_1, \dots, v_n)$  une base de  $V$ . D'après a) et b),  $f(\mathcal{B})$  est une famille libre et génératrice de  $W$ , donc une base de  $W$ ; alors  $\dim W = n = \dim V$ .



**Corollaire 0.2.13.** — (i) Tout  $k$ -espace vectoriel  $V$  de dimension finie est isomorphe (de façon non canonique) à  $k^n$ , pour un unique  $n$ , égal à  $\dim_k V$ .

(ii) Deux  $k$ -espaces vectoriels de dimension finie  $V$  et  $W$  sont isomorphes si et seulement si ils ont la même dimension.

*Démonstration.* (i) Si  $V$  est de dimension  $n$  alors, par le choix d'une base, il est isomorphe à  $k^n$ . (Cet isomorphisme est « non canonique » car il dépend du choix de la base.) Réciproquement, si  $V \simeq k^m$ , alors d'après la proposition précédente, on a  $\dim V = \dim k^m = m$ , donc  $m = n$ . En particulier,  $k^m \not\simeq k^n$  si  $m \neq n$ .

(ii) Si  $V \simeq W$ , alors  $\dim_k V = \dim_k W$ , d'après la proposition précédente. Réciproquement, si  $\dim_k V = \dim_k W = n$ , alors  $V$  et  $W$  sont tous deux isomorphes à  $k^n$ .

**Exemples 0.2.14.** — 1) La droite d'équation  $x_1 + x_2 = 0$  dans  $\mathbb{R}^2$  admet pour base le vecteur  $e_1 - e_2$ , mais on peut tout aussi bien choisir le vecteur  $e_2 - e_1$ .

2) Le plan d'équation  $x_1 + x_2 + x_3 = 0$  dans  $\mathbb{R}^3$  admet pour base  $(e_1 - e_2, e_2 - e_3)$ , mais on peut aussi choisir  $(e_1 - e_3, e_1 - e_3)$ , ou  $(e_1 - e_2, e_1 + e_2 - 2e_3)$ , ou  $(e_1 - (e_1 + e_2 + e_3)/3, e_1 + e_2 - 2(e_1 + e_2 + e_3)/3)$ , etc.

**Exemple 0.2.15.** — Reprenons l'espace  $\mathcal{S}(a, b)$  des suites  $(u_n)_{n \in \mathbb{N}}$  d'éléments de  $k$  vérifiant la relation de récurrence linéaire  $u_{n+2} = au_{n+1} + bu_n$ . On a vu (cf. 0.1.11) qu'il est isomorphe à  $k^2$ , par l'application  $(u_n)_{n \in \mathbb{N}} \mapsto (u_0, u_1)$ , donc est de dimension 2. Supposons que le polynôme  $P = X^2 - aX - b$  ait deux racines distinctes  $\lambda \neq \mu$  dans  $k$ . Considérons les éléments  $\mathbf{u} = (u_n)_{n \in \mathbb{N}}$  et  $\mathbf{v} = (v_n)_{n \in \mathbb{N}}$  de  $\mathcal{S}(a, b)$  définis par

$$u_0 = 1 = v_0, \quad u_1 = \lambda, \quad v_1 = \mu.$$

Alors la famille  $(\mathbf{u}, \mathbf{v})$  est libre (car si  $s\mathbf{u} + t\mathbf{v} = 0$ , on obtient  $s + t = 0 = s\lambda + t\mu$ , donc  $t = -s$  et  $s(\lambda - \mu) = 0$ , d'où  $s = 0$ ), donc est une base de  $\mathcal{S}(a, b)$ . Par conséquent, tout élément  $\mathbf{w} = (w_n)_{n \in \mathbb{N}}$  de  $\mathcal{S}(a, b)$  s'écrit de façon unique

$$\mathbf{w} = s\mathbf{u} + t\mathbf{v},$$

et  $s, t$  sont déterminés par les conditions  $s + t = w_0$  et  $s\lambda + t\mu = w_1$ . Ceci permet-il de calculer  $w_{2010}$  et  $w_{2011}$  dans l'exercice 0.1.12?

### 0.3. Noyau, image, et théorème du rang

**Définition 0.3.1 (Noyau, image et rang).** — Soit  $f : V \rightarrow W$  une application linéaire. D'une part, on définit son noyau

$$\text{Ker}(f) = \{v \in V \mid f(v) = 0\},$$

c'est un sous-espace vectoriel de  $V$ . Noter que  $f$  est *injective* si et seulement si  $\text{Ker}(f) = \{0\}$  : en effet, on a  $f(v) = f(v') \iff f(v - v') = 0$ .

D'autre part, on définit son image

$$\text{Im}(f) = f(V) = \{f(v) \mid v \in V\} \subseteq W,$$

c'est un sous-espace vectoriel de  $W$ . Alors  $f$  est *surjective* si et seulement si  $\text{Im}(f) = W$ .

Lorsque  $\text{Im}(f)$  est de dimension finie  $r$  (ce qui est le cas si  $W$  ou  $V$  est de dimension finie, cf. ci-dessous), l'entier  $r = \dim \text{Im}(f)$  est appelé le **rang** de  $f$  et est noté  $\text{rg}(f)$  ou  $\text{rang}(f)$ .

**Théorème 0.3.2 (Théorème du rang).** — Soit  $f : V \rightarrow W$  une application linéaire, avec  $V$  de dimension finie  $n$ . Alors

$$n = \dim V = \dim \text{Ker}(f) + \text{rg}(f).$$

(En particulier,  $f$  est surjective si et seulement si  $W$  est de dimension  $\text{rg}(f) = n - \dim \text{Ker}(f)$ .)

*Démonstration.* Comme  $V$  est de dimension finie  $n$ , alors  $\text{Ker}(f)$  est de dimension finie  $d \leq n$ ; soit  $\mathcal{K} = (e_1, \dots, e_d)$  une base de  $\text{Ker}(f)$ .

1ère méthode (la plus courte). Complétons  $\mathcal{K}$  en une base  $\mathcal{B} = (e_1, \dots, e_d, e_{d+1}, \dots, e_n)$  de  $V$ . Alors  $\text{Im}(f) = f(V)$  est engendré par  $f(\mathcal{B})$  (cf. Prop. 0.2.12), donc par  $f(e_{d+1}), \dots, f(e_n)$  puisque  $f(e_i) = 0$  pour  $i \leq d$ . Montrons que ces vecteurs sont linéairement indépendants : supposons qu'on ait une relation de dépendance linéaire

$$0 = t_1 f(e_{d+1}) + \dots + t_{n-d} f(e_n) = f(t_1 e_{d+1} + \dots + t_{n-d} e_n)$$

alors le vecteur  $t_1 e_{d+1} + \dots + t_{n-d} e_n$  appartient à  $\text{Ker}(f)$ , donc est combinaison linéaire de  $e_1, \dots, e_d$ , d'où une égalité

$$t_1 e_{d+1} + \dots + t_{n-d} e_n - s_1 e_1 - \dots - s_d e_d = 0.$$

Comme  $(e_1, \dots, e_n)$  est une base de  $V$ , ceci implique  $t_i = 0 = s_j$  pour tous  $i, j$ . Ceci montre que les vecteurs  $f(e_{d+1}), \dots, f(e_n)$  sont linéairement indépendants, donc forment une base de  $f(V)$ , d'où  $\dim f(V) = n - d$  et donc  $\text{rg}(f) = \dim f(V) = n - \dim \text{Ker}(f)$ .



2ème méthode. Soit toujours  $\mathcal{K} = (e_1, \dots, e_d)$  une base de  $\text{Ker}(f)$ , donnons une autre démonstration, qui sera utile plus loin (cf. 0.5.10).<sup>(1)</sup> Comme  $\text{Im}(f) = f(V)$  est engendré par  $n$  éléments (les images d'une base de  $V$ ),  $\text{Im}(f)$  est de dimension finie  $r \leq n$ . Soit  $(w_1, \dots, w_r)$  une base de  $\text{Im}(f)$  et pour  $i = 1, \dots, r$ , soit  $v_i$  un élément de  $V$  tel que  $f(v_i) = w_i$ . Alors la famille

$$\mathcal{B} = (e_1, \dots, e_d, v_1, \dots, v_r)$$

est une base de  $V$ . En effet, elle est *génératrice* : soit  $v \in V$  arbitraire, son image  $f(v)$  s'écrit  $f(v) = t_1 w_1 + \dots + t_r w_r$ , d'où  $f(v - t_1 v_1 - \dots - t_r v_r) = 0$ , donc  $v - t_1 v_1 - \dots - t_r v_r$  appartient à  $\text{Ker}(f)$ , donc s'écrit  $s_1 e_1 + \dots + s_p e_p$ , d'où

$$v = t_1 v_1 + \dots + t_r v_r + s_1 e_1 + \dots + s_p e_p.$$

Ceci montre que  $\mathcal{B}$  est génératrice. Elle est aussi *libre* : si

$$0 = t_1 v_1 + \dots + t_r v_r + s_1 e_1 + \dots + s_p e_p$$

alors  $0 = f(0) = t_1 w_1 + \dots + t_r w_r$ , donc chaque  $t_i$  est nul (puisque  $(w_1, \dots, w_r)$  est libre), d'où  $0 = s_1 e_1 + \dots + s_p e_p$ , donc chaque  $s_j$  est nul (puisque  $(e_1, \dots, e_p)$  est libre). Ceci montre que  $\mathcal{B}$  est aussi libre, donc est une base de  $V$ . Donc  $\dim V = d + r = \dim \text{Ker}(f) + \text{rg}(f)$ .



**Proposition 0.3.3.** — Soit  $V$  un  $k$ -espace vectoriel de dimension  $n$  et soit  $u \in \text{End}_k(V)$  ou, plus généralement, soit  $u : V \rightarrow W$  une application linéaire, où  $\dim W = n = \dim V$ . Alors les conditions suivantes sont équivalentes :

- (i)  $u$  est bijectif;
- (ii)  $u$  est injectif;
- (iii)  $u$  est surjectif.

En effet, il est clair que (i) implique (ii) et (iii). Réciproquement, si  $u$  est injectif, i.e.  $\text{Ker}(u) = \{0\}$  (resp. surjectif, i.e.  $\text{Im}(u) = W$ ), il résulte du théorème du rang que  $u$  est aussi surjectif (resp. injectif), donc bijectif.

#### 0.4. Applications linéaires et matrices

**Définition 0.4.1 (Espaces d'applications linéaires).** — Soient  $V, W$  deux  $k$ -espaces vectoriels. On note  $\text{Hom}_k(V, W)$  ou  $\mathcal{L}(V, W)$  l'ensemble des applications linéaires de  $V$  dans  $W$ . Si  $\phi, \psi$  sont deux telles applications, et si  $t \in k$ , on définit les applications  $\phi + \psi$  et  $t \cdot \phi$  comme suit : pour tout  $v \in V$ ,

$$(*) \quad (\phi + \psi)(v) = \phi(v) + \psi(v) \quad \text{et} \quad (t \cdot \phi)(v) = t \cdot \phi(v).$$

Ce sont encore des applications *linéaires*  $V \rightarrow W$  : en effet, si  $v, v' \in V$  et  $s \in k$ , alors

$$\begin{aligned} (\phi + \psi)(s \cdot v + v') &= \phi(s \cdot v + v') + \psi(s \cdot v + v') && \text{(par définition)} \\ &= s \cdot \phi(v) + \phi(v') + s \cdot \psi(v) + \psi(v') && \text{(car } \phi, \psi \text{ linéaires)} \\ &= s \cdot (\phi + \psi)(v) + (\phi + \psi)(v') && \text{(par définition)} \end{aligned}$$

et de même

$$(t \cdot \phi)(s \cdot v + v') = t \cdot \phi(s \cdot v + v') = ts \cdot \phi(v) + t \cdot \phi(v') = s \cdot (t \cdot \phi)(v) + (t \cdot \phi)(v').$$

Ainsi, (\*) munit  $\text{Hom}_k(V, W) = \mathcal{L}(V, W)$  d'une structure de  $k$ -espace vectoriel; on dit que c'est l'espace des applications linéaires de  $V$  dans  $W$ .

Supposons  $V$  de dimension finie  $n$  et soit  $(e_1, \dots, e_n)$  une base de  $V$  (par exemple,  $V = k^n$  et  $(e_1, \dots, e_n)$  la base canonique). Soit  $\phi \in \text{Hom}_k(V, W)$ , posons  $w_i = \phi(e_i)$  pour  $i = 1, \dots, n$ ; alors tout  $v \in V$  s'écrit de façon unique  $v = x_1 e_1 + \dots + x_n e_n$  et l'on a

$$(*) \quad \phi(v) = x_1 w_1 + \dots + x_n w_n$$

donc  $\phi$  est déterminée par la donnée des  $n$  vecteurs  $w_1, \dots, w_n \in W$ . Réciproquement, pour tout  $n$ -uplet  $(w_1, \dots, w_n) \in W^n$ , l'application  $\phi : V \rightarrow W$  définie par la formule (\*) est linéaire. On a donc obtenu la

**Proposition 0.4.2.** — Si  $(e_1, \dots, e_n)$  est une base de  $V$ , se donner une application linéaire  $\phi : V \rightarrow W$  « est la même chose » que se donner un  $n$ -uplet  $(w_1, \dots, w_n) \in W^n$ .

<sup>(1)</sup>Cette démonstration a l'avantage suivant. Elle montre que si une base  $(w_1, \dots, w_r)$  de  $\text{Im}(f)$  est donnée à l'avance et si  $v_1, \dots, v_r \in V$  vérifient  $f(v_i) = w_i$  pour  $i = 1, \dots, r$ , alors  $(e_1, \dots, e_d, v_1, \dots, v_r)$  est une base de  $V$ .

Supposons de plus que  $W$  soit de dimension finie  $m$  et soit  $(f_1, \dots, f_m)$  une base de  $W$ . Alors, chaque  $w_j = \phi(e_j)$  s'écrit de façon unique

$$w_j = a_{1j}f_1 + \dots + a_{mj}f_m,$$

ce qu'on représente par le vecteur colonne (d'où le choix de l'indice  $j$  pour paramétrer ces colonnes) :

$$w_j = \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix}$$

et donc  $\phi$  est déterminée par la matrice suivante :

$$\text{Mat}_{(f_i), (e_j)}(\phi) = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

qui exprime les vecteurs  $\phi(e_j)$  (les colonnes) en fonction de  $f_1, \dots, f_m$ . Noter que la dimension  $n$  de l'espace de départ  $V$  est le nombre de colonnes, et la dimension  $m$  de l'espace d'arrivée  $W$  est le nombre de lignes.

Réciproquement, pour toute matrice  $A$  comme ci-dessus, ses colonnes définissent de façon unique  $n$  vecteurs  $w_1, \dots, w_n$  de  $W$ , à savoir

$$w_j = a_{1j}f_1 + \dots + a_{mj}f_m,$$

et ce  $n$ -uplet  $(w_1, \dots, w_n) \in W^n$  définit une application linéaire  $\phi : V \rightarrow W$  dont la matrice associée est  $A$ . On obtient donc une bijection :

$$\text{Hom}_k(V, W) \longleftrightarrow M_{m,n}(k).$$

De plus, on voit facilement que si  $A$  (resp.  $B$ ) est la matrice associée à  $\phi$  (resp.  $\psi$ ), alors  $tA + B$  est la matrice associée à  $t\phi + \psi$ , donc la bijection ci-dessus est un isomorphisme d'espaces vectoriels.

**Théorème 0.4.3 (Applications linéaires et matrices).** — Soient  $\mathcal{B} = (e_1, \dots, e_n)$  une base de  $V$ , et  $\mathcal{C} = (f_1, \dots, f_m)$  une base de  $W$ .

(1) Une application linéaire  $V \rightarrow W$  « est la même chose » qu'une matrice à  $m$  lignes et  $n$  colonnes, c.-à-d., l'application

$$\text{Hom}_k(V, W) \simeq M_{m,n}(k), \quad \phi \mapsto \text{Mat}_{\mathcal{C}, \mathcal{B}}(\phi).$$

est un isomorphisme d'espaces vectoriels.

(2) Ceci transforme la composition des applications linéaires en le produit des matrices : si  $U$  est un  $k$ -espace vectoriel de base  $\mathcal{A} = (d_1, \dots, d_p)$  et si  $\theta \in \text{Hom}_k(U, V)$ , alors

$$\text{Mat}_{\mathcal{C}, \mathcal{A}}(\phi \circ \theta) = \text{Mat}_{\mathcal{C}, \mathcal{B}}(\phi) \cdot \text{Mat}_{\mathcal{B}, \mathcal{A}}(\theta).$$

*Démonstration.* — On a déjà vu l'assertion (1), montrons l'assertion (2). Notons

$$A = \text{Mat}_{\mathcal{C}, \mathcal{B}}(\phi) = (a_{ij})_{\substack{j=1, \dots, n \\ i=1, \dots, m}} \quad B = \text{Mat}_{\mathcal{B}, \mathcal{A}}(\theta) = (b_{j\ell})_{\substack{\ell=1, \dots, p \\ j=1, \dots, n}}$$

alors, pour tout  $\ell = 1, \dots, p$ , on a :

$$(\phi \circ \theta)(d_\ell) = \phi\left(\sum_{j=1}^n b_{j\ell} e_j\right) = \sum_{j=1}^n b_{j\ell} \phi(e_j) = \sum_{j=1}^n \sum_{i=1}^m b_{j\ell} a_{ij} f_i = \sum_{i=1}^m \left(\sum_{j=1}^n a_{ij} b_{j\ell}\right) f_i.$$

Donc le coefficient d'indice  $(i, \ell)$  de  $M = \text{Mat}_{(f_i), (d_\ell)}(\phi \circ \theta)$  est  $\sum_{j=1}^n a_{ij} b_{j\ell}$ ; ceci montre que  $M = AB$ .  $\square$

**Remarque 0.4.4.** — Soit  $A = (a_{ij}) \in M_{m,n}(k)$  et soient  $(e_1, \dots, e_n)$  et  $(f_1, \dots, f_m)$  les bases canoniques de  $k^n$  et  $k^m$ . Alors par l'isomorphisme précédent,  $A$  correspond à l'application linéaire  $u : k^n \rightarrow k^m$  telle que, pour tout  $j = 1, \dots, n$ ,

$$u(e_j) = \sum_{i=1}^m a_{ij} f_i.$$

On identifiera, chaque fois que ce sera utile,  $A$  à l'application linéaire  $u : k^n \rightarrow k^m$  ainsi définie (i.e. la  $i$ -ième colonne de  $A$  est l'image du  $i$ -ième vecteur de la base canonique de  $k^n$ ).

**Corollaire 0.4.5.** — Soient  $A \in M_{m,n}(k)$ ,  $B \in M_{n,p}(k)$  et  $u : k^n \rightarrow k^m$ ,  $v : k^p \rightarrow k^n$  les applications linéaires associées. Alors  $AB$  est la matrice de  $u \circ v : k^p \rightarrow k^m$ .



**Remarque 0.4.6.** — Soient  $B \in M_{m,n}(k)$  et  $A \in M_{n,p}(k)$ . Si l'on note  $A_1, \dots, A_p \in k^n$  les colonnes de  $A$ , alors les colonnes de  $BA$  sont les vecteurs  $BA_1, \dots, BA_p \in k^m$ .<sup>(2)</sup>

En effet, ceci se voit directement sur la formule du produit matriciel :  $(BA)_{ij} = \sum_{\ell=1}^n b_{i\ell} a_{\ell j}$ . Ou bien, on peut raisonner comme suit : soit  $(e_1, \dots, e_p)$  la base canonique de  $k^p$ , alors  $A$  correspond à l'application linéaire qui envoie chaque  $e_j$  sur le vecteur  $Ae_j = A_j \in k^n$ , et  $BA$  correspond à l'application linéaire qui envoie chaque  $e_j$  sur le vecteur  $B(Ae_j) = BA_j \in k^m$ .

**Remarque 0.4.7.** — On ne peut effectuer le produit  $AB$  de deux matrices  $A \in M_{m,n}(k)$  et  $B \in M_{q,p}(k)$  que si  $n = q$ , c.-à-d., si le nombre de colonnes de  $A$  égale le nombre de lignes de  $B$ .

En raison de son importance, répétons le théorème 0.4.3 et le corollaire 0.4.5 dans le cas particulier où l'espace de départ est le même que celui d'arrivée, c.-à-d., le cas où l'on considère des endomorphismes d'un espace  $V$  de dimension finie  $n$ , ou des matrices carrées de taille  $n$ .

**Proposition 0.4.8 (Endomorphismes de  $V \simeq k^n$ ).** — Le  $k$ -espace vectoriel  $M_n(k)$  est un anneau (non commutatif si  $n \geq 2$ ), c.-à-d., la multiplication des matrices carrées est associative :  $A(BC) = (AB)C$ , distributive à gauche et à droite par rapport à l'addition :  $(A+B)C = AC+BC$  et  $A(B+C) = AB+AC$ , et la matrice identité  $I_n$  est élément neutre. De plus,  $M_n(k)$  est une  $k$ -algèbre, c.-à-d.,<sup>(3)</sup> pour  $A, B \in M_n(k)$  et  $t \in k$ , on a  $t \cdot (AB) = (t \cdot A)B = A(t \cdot B)$ , où  $\cdot$  désigne la loi externe.

De même, si  $V$  est un  $k$ -espace vectoriel de dimension  $n$ , l'espace des endomorphismes  $\text{End}_k(V)$  est une  $k$ -algèbre (la multiplication étant la composition des endomorphismes). De plus, si l'on choisit une base  $\mathcal{B} = (e_1, \dots, e_n)$  de  $V$ , l'application

$$\text{End}_k(V) \rightarrow M_n(k), \quad u \mapsto \text{Mat}_{\mathcal{B}}(u)$$

est un isomorphisme d'anneaux et de  $k$ -espaces vectoriels, c.-à-d., un isomorphisme de  $k$ -algèbres.



**Définition 0.4.9 (Noyau, image et rang d'une matrice).** — Soit  $A \in M_{m,n}(k)$ , on définit son noyau  $\text{Ker}(A)$ , son image  $\text{Im}(A)$ , et son rang, noté  $\text{rang}(A)$  ou  $\text{rg}(A)$ , comme le noyau, l'image et le rang de l'application linéaire  $u : k^n \rightarrow k^m$  associée. On a  $\text{rang}(u) \leq n$  (d'après le théorème du rang), et  $\text{rang}(u) \leq m$  (puisque  $\text{Im}(u)$  est un sous-espace de  $k^m$ ), donc  $\text{rang}(u) \leq \text{Min}(m, n)$ .

Or, l'image de  $u$  est le sous-espace de  $k^m$  engendré par les vecteurs colonnes  $C_1, \dots, C_n$  de  $A$ , donc par définition  $\text{rang}(A)$  est le nombre maximum de colonnes de  $A$  linéairement indépendantes, et l'on a  $\text{rang}(A) \leq \text{Min}(m, n)$ .

On verra plus bas que  $\text{rang}(A)$  est aussi le nombre maximum de lignes linéairement indépendantes, et l'on donnera des moyens algorithmiques pour calculer  $\text{rang}(A)$ .



**Définition 0.4.10 (Transposée d'une matrice).** — Soit dans  $M_{m,n}(k)$  :

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

sa transposée, notée  ${}^tA$ , est la matrice de  $M_{n,m}(k)$  suivante :

$${}^tA = \begin{pmatrix} a_{11} & a_{21} & \cdots & a_{m1} \\ a_{12} & a_{22} & \cdots & a_{m2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{mn} \end{pmatrix}$$

c.-à-d., la  $j$ -ème colonne de  $A$  est devenue la  $j$ -ème ligne de  ${}^tA$ , c.-à-d.,  $({}^tA)_{ji} = a_{ij}$ . On a évidemment

$$(\star) \quad \boxed{{}^t({}^tA) = A.}$$

<sup>(2)</sup>Cette remarque sera utile pour la construction du déterminant, voir Chap. 2.

<sup>(3)</sup>cf. Définition 4.2.1 dans le Chap. 3.

**Proposition 0.4.11.** — L'application  $A \mapsto {}^t A$  est linéaire : si  $A, A' \in M_{m,n}(k)$  et  $s \in k$ ,

$${}^t(s \cdot A + A') = s \cdot {}^t A + {}^t A'.$$

De plus, si  $B \in M_{n,p}(k)$ , alors  $BA \in M_{m,p}(k)$  et l'on a dans  $M_{p,m}(k)$  l'égalité :

$$(\star\star) \quad \boxed{{}^t(AB) = {}^t B {}^t A.}$$

*Démonstration.* Écrivons  $A = (a_{ij})$  et  $A' = (a'_{ij})$ . Alors  $sA + A' \in M_{m,n}(k)$  est la matrice  $(sa_{ij} + a'_{ij})$ , et sa transposée est la matrice  $C \in M_{n,m}(k)$  telle que, pour tout  $(j, i)$ ,

$$C_{ji} = (sA + A')_{ij} = sa_{ij} + a'_{ij} = s({}^t A)_{ji} + ({}^t A')_{ji},$$

donc  $C = s \cdot {}^t A + {}^t A'$ . Puis, si l'on pose  $B = (b_{j\ell}) \in M_{n,p}(k)$ , alors pour tout couple  $(i, \ell)$  on a

$$({}^t(AB))_{\ell i} = (AB)_{i\ell} = \sum_{r=1}^n a_{ir} b_{r\ell} = \sum_{r=1}^n ({}^t B)_{\ell r} \cdot ({}^t A)_{ri} = ({}^t B {}^t A)_{\ell i}$$

ce qui montre que  ${}^t(AB) = {}^t B {}^t A$ . □

## 0.5. Changements de base

**Définition 0.5.1 (Automorphismes et matrices inversibles).** — 1) Soit  $V$  un  $k$ -espace vectoriel. On dit qu'un endomorphisme  $f$  de  $V$  est un **automorphisme** s'il possède un inverse dans  $\text{End}_k(V)$ , i.e. s'il existe un endomorphisme  $g$  de  $V$  tel que  $f \circ g = \text{id}_V = g \circ f$ . Ceci équivaut à dire que  $f$  est bijectif, car on a vu (cf. 0.1.10) que dans ce cas l'application inverse  $g$  est automatiquement linéaire.

2) On note  $\text{GL}(V)$  l'ensemble des automorphismes de  $V$ ; c'est un groupe pour la composition des endomorphismes : en effet, la composition des endomorphismes est associative, l'application identique est élément neutre, et si  $f, g$  sont inversibles, alors  $f \circ g$  l'est aussi, son inverse étant  $g^{-1} \circ f^{-1}$  (puisque  $f \circ g \circ g^{-1} \circ f^{-1} = \text{id}_V = g^{-1} \circ f^{-1} \circ f \circ g$ ). <sup>(4)</sup>

3) De même, on dit qu'une matrice  $A \in M_n(k)$  est **inversible** s'il existe  $B \in M_n(k)$  vérifiant  $AB = I_n = BA$ , où  $I_n$  désigne la matrice identité de taille  $n$ ; dans ce cas  $B$  est notée  $A^{-1}$ . On note  $\text{GL}_n(k)$  l'ensemble des matrices inversibles.

Comme la correspondance bijective  $\text{End}_k(k^n) \longleftrightarrow M_n(k)$  transforme la composition des endomorphismes en le produit des matrices, on voit qu'une matrice  $A$  est inversible si et seulement si l'endomorphisme correspondant  $u$  de  $k^n$  est bijectif, et dans ce cas  $A^{-1}$  est la matrice de  $u^{-1}$ ; de plus  $\text{GL}_n(k)$  est un groupe pour la multiplication des matrices : la matrice  $I_n$  est élément neutre, et si  $A, B$  sont inversibles, alors  $AB$  l'est aussi, son inverse étant  $B^{-1}A^{-1}$ .

La proposition suivante est importante et très utile :

**Proposition 0.5.2.** — (i) Soit  $V$  un  $k$ -espace vectoriel de dimension finie et soient  $u, v \in \text{End}_k(V)$  tels que  $u \circ v = \text{id}_V$ . Alors  $u$  et  $v$  sont bijectifs et  $u = v^{-1}$ .

(i') Soient  $A, B \in M_n(k)$  telles que  $AB = I_n$ . Alors on a aussi  $BA = I_n$  et donc  $A$  et  $B$  sont inversibles et inverses l'une de l'autre.

(ii) Si  $A$  est inversible alors  ${}^t A$  est inversible et l'on a  $({}^t A)^{-1} = {}^t(A^{-1})$ .

*Démonstration.* (i) Pour tout  $x \in V$ , on a  $x = u(v(x))$ , donc  $u$  est surjectif, et  $v$  est injectif. Donc, d'après la proposition 0.3.3,  $u$  et  $v$  sont bijectifs; alors en multipliant l'égalité  $u \circ v = \text{id}_V$  à gauche par  $u^{-1}$  (ou bien à droite par  $v^{-1}$ ), on obtient que  $v = u^{-1}$ .

(i') Notons  $u$  (resp.  $v$ ) l'endomorphisme de  $k^n$  correspondant à  $A$  (resp.  $B$ ). Comme  $AB = I_n$  équivaut à  $u \circ v = \text{id}_V$  alors, d'après (i),  $u$  et  $v$  sont bijectifs et inverses l'un de l'autre, donc il en est de même de  $A$  et  $B$  et l'on a  $B = A^{-1}$  et  $BA = I_n$ .

(ii) Supposons  $A$  inversible, alors il existe  $B \in M_n(k)$  telle que  $AB = I_n = BA$ . Prenant la transposée de ces matrices, on obtient, puisque  ${}^t I_n = I_n$  :

$${}^t B {}^t A = {}^t(AB) = I_n = {}^t(BA) = {}^t A {}^t B.$$

Ceci montre que  ${}^t A$  est inversible, d'inverse  ${}^t B = {}^t(A^{-1})$ .

<sup>(4)</sup> Attention ! Noter l'inversion de l'ordre des facteurs :  $(f \circ g)^{-1}$  égale  $g^{-1} \circ f^{-1}$ , tandis que  $f^{-1} \circ g^{-1} = (g \circ f)^{-1}$  !



**Remarque 0.5.3.** — 1) Soit  $V = \mathbb{R}[X]$ , soit  $I$  l'opérateur « d'intégration », qui envoie chaque monôme  $X^n$  sur  $X^{n+1}/(n+1)$ , et soit  $D$  l'opérateur de dérivation, qui envoie chaque polynôme  $P$  sur le polynôme dérivé  $P'$ . Alors  $D \circ I = \text{id}_V$ , donc  $D$  est surjectif et  $I$  injectif, mais  $D$  n'est pas injectif car  $D(1) = 0$ , et  $I$  n'est pas surjectif car son image est formée des polynômes de terme constant nul.

2) De même, soit  $V$  l'espace des suites  $(u_n)_{n \in \mathbb{N}}$  et soient  $I$  (resp.  $D$ ) l'opérateur qui envoie toute suite  $(u_n)_{n \in \mathbb{N}}$  sur

$$(0, u_0, u_1, u_2, \dots) \quad \text{resp.} \quad (u_1, u_2, u_3, \dots).$$

Alors  $D \circ I = \text{id}_V$ , donc  $D$  est surjectif et  $I$  injectif, mais  $D$  n'est pas injectif car  $D$  annule la suite telle que  $u_0 = 1$  et  $u_i = 0$  pour  $i \geq 1$ , et  $I$  n'est pas surjectif car son image est formée des suites de terme  $u_0$  nul.

Ces deux exemples montrent que si  $V$  est un espace vectoriel qui n'est pas de dimension finie et si  $u, v \in \text{End}_k(V)$  vérifient  $u \circ v = \text{id}_V$ , alors  $u$  et  $v$  ne sont pas nécessairement bijectifs.

**Lemme 0.5.4.** — Soient  $f$  un endomorphisme d'un espace vectoriel  $V$  de dimension finie,  $\mathcal{B} = (v_1, \dots, v_n)$  une base de  $V$ , posons  $w_i = f(v_i)$ . Si  $(w_1, \dots, w_n)$  est une base de  $V$ , alors  $f$  est bijective, et son inverse est l'endomorphisme  $g$  de  $V$  défini par  $g(w_i) = v_i$  pour  $i = 1, \dots, n$ .

*Démonstration.* On suppose que  $\mathcal{C} = (w_1, \dots, w_n)$  est une base de  $V$ . Alors  $f$  est surjectif, car pour tout  $w \in V$  il existe  $t_1, \dots, t_n \in k$  tels que

$$w = t_1 w_1 + \dots + t_n w_n = f(t_1 v_1 + \dots + t_n v_n).$$

Donc, d'après la proposition 0.3.3,  $f$  est bijectif. (On peut aussi voir directement que  $f$  est injectif : soit  $v \in \text{Ker}(f)$ , il existe  $t_1, \dots, t_n \in k$  tels que  $v = x_1 v_1 + \dots + x_n v_n$ , alors  $0 = f(v) = x_1 w_1 + \dots + x_n w_n$  donc, puisque  $(w_1, \dots, w_n)$  est libre,  $t_1 = 0 = \dots = t_n$ , d'où  $v = 0$ .)

Soit  $g$  l'endomorphisme de  $V$  défini par  $g(w_i) = v_i$ , pour  $i = 1, \dots, n$ . Alors on a, d'une part,  $(g \circ f)(v_i) = v_i$  pour tout  $i$ , d'où  $g \circ f = \text{id}_V$ , et, d'autre part,  $(f \circ g)(w_i) = w_i$  pour tout  $i$ , d'où  $f \circ g = \text{id}_V$ .



**Définition 0.5.5 (Matrice de passage).** — Soient  $V$  un espace vectoriel de dimension finie,  $\mathcal{B} = (e_1, \dots, e_n)$  une base de  $V$ , et  $\mathcal{B}' = (v_1, \dots, v_n)$  une seconde base de  $V$ . Soit  $P$  la matrice  $n \times n$  exprimant la seconde base en fonction de la première, c.-à-d., chaque  $v_j$  s'écrit de façon unique

$$v_j = p_{1j} e_1 + \dots + p_{nj} e_n$$

et l'on forme la matrice

$$P = \begin{pmatrix} p_{11} & \dots & p_{1j} & \dots & p_{1n} \\ \vdots & & \vdots & & \vdots \\ p_{i1} & \dots & p_{ij} & \dots & p_{in} \\ \vdots & & \vdots & & \vdots \\ p_{n1} & \dots & p_{nj} & \dots & p_{nn} \end{pmatrix}$$

dont les colonnes sont les vecteurs  $v_1, \dots, v_n$  exprimés dans la base  $(e_1, \dots, e_n)$ . Alors  $P$  s'appelle la **matrice de passage** de la base  $\mathcal{B} = (e_1, \dots, e_n)$  à la base  $\mathcal{B}' = (v_1, \dots, v_n)$  et se note  $\boxed{\text{Mat}_{\mathcal{B}}(\mathcal{B}')}$ .

C'est une matrice **inversible** : son inverse  $P^{-1}$  est la matrice exprimant  $e_1, \dots, e_n$  dans la base  $(v_1, \dots, v_n)$ .

Remarquons que  $P$  peut être vue comme la matrice de l'application identité  $\text{id}_V$ , exprimée dans les bases :  $\mathcal{B}' = (v_1, \dots, v_n)$  au départ, et  $\mathcal{B} = (e_1, \dots, e_n)$  à l'arrivée, c.-à-d.,

$$\boxed{P = \text{Mat}_{\mathcal{B}, \mathcal{B}'}(\text{id}_V)} \quad \text{et de même} \quad \boxed{P^{-1} = \text{Mat}_{\mathcal{B}', \mathcal{B}}(\text{id}_V)}.$$

Conservons les notations précédentes. Tout  $v \in V$  s'écrit alors de façon unique

$$v = x_1 e_1 + \dots + x_n e_n \quad \text{et} \quad v = x'_1 v_1 + \dots + x'_n v_n$$

et les  $x_i$  (resp.  $x'_i$ ) s'appellent les **coordonnées** de  $v$  relativement à la base  $\mathcal{B}$  (resp.  $\mathcal{B}'$ ), cf. 0.2.10. Relativement à la base  $\mathcal{B}$  (resp.  $\mathcal{B}'$ ), on peut représenter  $v$  comme le vecteur colonne

$$X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \quad \text{resp.} \quad X' = \begin{pmatrix} x'_1 \\ \vdots \\ x'_n \end{pmatrix}$$



**Proposition 0.5.6 (Changement de coordonnées).** — La formule de changement de coordonnées, pour le changement de base  $\mathcal{B} \rightarrow \mathcal{B}'$  donné par la matrice de passage  $P$ , est donnée par :

$$X = PX'$$

(noter que cette formule exprime les anciennes coordonnées  $X$  en fonction des nouvelles  $X'$ ).

*Démonstration.* En effet, écrivant  $v_j = \sum_{i=1}^n p_{ij}e_i$  puis

$$v = \sum_{j=1}^n x'_j v_j = \sum_{j=1}^n \sum_{i=1}^n x'_j p_{ij} e_i = \sum_{i=1}^n \left( \sum_{j=1}^n p_{ij} x'_j \right) e_i$$

et comparant avec  $v = \sum_{i=1}^n x_i e_i$ , on obtient qu'on a  $x_i = \sum_{j=1}^n p_{ij} x'_j$  pour  $i = 1, \dots, n$ , d'où  $X = PX'$ .

Soient  $V, \mathcal{B}, \mathcal{B}'$  et  $P$  comme plus haut, et considérons maintenant une application linéaire  $u : V \rightarrow W$ . Soient  $\mathcal{C} = (f_1, \dots, f_m)$  une base de  $W$ , et  $A \in M_{m,n}(k)$  la matrice de  $u$  dans les bases  $\mathcal{B}$  et  $\mathcal{C}$ . Alors, d'après le théorème 0.4.3, on a :

$$\text{Mat}_{\mathcal{C}, \mathcal{B}'}(u) = \text{Mat}_{\mathcal{C}, \mathcal{B}}(u) \cdot \text{Mat}_{\mathcal{B}, \mathcal{B}'}(\text{id}_V)$$

donc la matrice de  $u$  dans les bases  $\mathcal{B}'$  et  $\mathcal{C}$  est  $AP$ .

Enfin, soit  $\mathcal{C}' = (w_1, \dots, w_m)$  une seconde base de  $W$  et soit  $Q$  la matrice de passage de  $\mathcal{C}$  à  $\mathcal{C}'$  ; alors  $Q = \text{Mat}_{\mathcal{C}', \mathcal{C}}(\text{id}_W)$  et  $Q^{-1} = \text{Mat}_{\mathcal{C}, \mathcal{C}'}(\text{id}_W)$ , d'où :

$$\text{Mat}_{\mathcal{C}', \mathcal{B}'}(u) = \text{Mat}_{\mathcal{C}', \mathcal{C}}(\text{id}_W) \cdot \text{Mat}_{\mathcal{C}, \mathcal{B}'}(u) = Q^{-1}AP.$$

On a donc obtenu le théorème suivant :

**Théorème 0.5.7 (Changement de bases pour une application linéaire)**

Soit  $A \in M_{m,n}(k)$  la matrice d'une application linéaire  $u : V \rightarrow W$ , relativement à des bases  $\mathcal{B} = (e_1, \dots, e_n)$  de  $V$  et  $\mathcal{C} = (f_1, \dots, f_m)$  de  $W$ . Soit  $\mathcal{B}' = (v_1, \dots, v_n)$  (resp.  $\mathcal{C}' = (w_1, \dots, w_m)$ ) une seconde base de  $V$  (resp. de  $W$ ) et soit  $P$  (resp.  $Q$ ) la matrice de passage correspondante. Alors la matrice de  $u$  dans les bases  $\mathcal{B}'$  et  $\mathcal{C}'$  est :

$$\text{Mat}_{\mathcal{C}', \mathcal{B}'}(u) = Q^{-1}AP.$$

**Remarque 0.5.8.** — Le théorème précédent est (évidemment) compatible avec la formule de changement de coordonnées 0.5.6 : si l'on désigne par  $X$  (resp.  $X'$ ) les coordonnées d'un vecteur  $v \in V$  relativement à  $\mathcal{B}$  (resp.  $\mathcal{B}'$ ), et  $Y$  (resp.  $Y'$ ) les coordonnées du vecteur  $u(v)$  dans la base  $\mathcal{C}$  (resp.  $\mathcal{C}'$ ), alors on a :

$$Y = AX, \quad X = PX', \quad Y = QY'$$

d'où  $Y' = Q^{-1}Y = Q^{-1}AX = Q^{-1}APX'$ . Compte-tenu de cette « compatibilité », on peut utiliser (comme moyen mnémotechnique) l'une de ces formules pour retrouver l'autre...

**Remarque 0.5.9.** — Même si l'on s'intéresse au départ à une matrice  $A \in M_{m,n}(k)$ , il est souvent utile de considérer  $A$  comme une application linéaire  $u : k^n \rightarrow k^m$  (définie par  $u(e_j) = a_{1j}f_1 + \dots + a_{mj}f_m$ , où  $(e_1, \dots, e_n)$ , resp.  $(f_1, \dots, f_m)$ , est la base canonique de  $k^n$ , resp.  $k^m$ ). Par exemple, le théorème précédent donne alors le corollaire suivant :

**Corollaire 0.5.10.** — Soit  $A \in M_{m,n}(k)$  et soit  $r = \text{rang}(A)$ .

1) Il existe des matrices inversibles  $P \in \text{GL}_n(k)$  et  $Q \in \text{GL}_m(k)$  telles que

$$Q^{-1}AP = \begin{pmatrix} I_r & \mathbf{0}_{r, n-r} \\ \mathbf{0}_{m-r, r} & \mathbf{0}_{m-r, n-r} \end{pmatrix}$$

où  $I_r$  est la matrice identité de taille  $r$  et où  $\mathbf{0}_{p,q}$  désigne la matrice nulle à  $p$  lignes et  $q$  colonnes.

2) Réciproquement, s'il existe des matrices inversibles  $P \in \text{GL}_n(k)$  et  $Q \in \text{GL}_m(k)$  et un entier  $s \in \mathbb{N}$  tels que

$$Q^{-1}AP = \begin{pmatrix} I_s & \mathbf{0}_{s, n-s} \\ \mathbf{0}_{m-s, s} & \mathbf{0}_{m-s, n-s} \end{pmatrix}$$

alors  $s = \text{rang}(A)$ .



*Démonstration.* Soient  $(e_1, \dots, e_n)$  et  $(f_1, \dots, f_m)$  les bases canoniques de  $k^n$  et  $k^m$  et soit  $u$  l'application linéaire  $k^n \rightarrow k^m$  correspondant à  $A$ . Par définition,  $r = \text{rang}(A)$  est la dimension de  $\text{Im}(u)$ . Soit donc  $(w_1, \dots, w_r)$  une base de  $\text{Im}(u)$ , on peut la compléter en une base  $\mathcal{C} = (w_1, \dots, w_m)$  de  $k^m$ ; notons  $Q$  la matrice de passage de  $(f_1, \dots, f_m)$  à  $\mathcal{C}$ .

Soient  $v_1, \dots, v_r$  des éléments de  $k^n$  tels que  $u(v_j) = w_j$ , pour  $j = 1, \dots, r$ , et soit  $(e_1, \dots, e_d)$  une base de  $\text{Ker}(u)$ . D'après la démonstration (2ème méthode) du théorème du rang 0.3.2,  $\mathcal{B} = (v_1, \dots, v_r, e_1, \dots, e_d)$  est une base de  $k^n$ . Alors, la matrice de  $u$  dans les bases  $\mathcal{B}$  et  $\mathcal{C}$  est

$$\text{Mat}_{\mathcal{C}, \mathcal{B}}(u) = \begin{pmatrix} I_r & \mathbf{0}_{r, n-r} \\ \mathbf{0}_{m-r, r} & \mathbf{0}_{m-r, n-r} \end{pmatrix}.$$

Or, si  $P$  désigne la matrice de passage de  $(e_1, \dots, e_n)$  à  $\mathcal{B}$ , on a

$$\text{Mat}_{\mathcal{C}, \mathcal{B}}(u) = Q^{-1} \cdot \text{Mat}_{(f_i), (e_j)}(u) \cdot P = Q^{-1}AP,$$

d'où l'assertion 1) du corollaire.

Réciproquement, supposons qu'il existe des matrices inversibles  $P \in \text{GL}_n(k)$  et  $Q \in \text{GL}_m(k)$  et un entier  $s \in \mathbb{N}$  tels que

$$Q^{-1}AP = \begin{pmatrix} I_s & \mathbf{0}_{s, n-s} \\ \mathbf{0}_{m-s, s} & \mathbf{0}_{m-s, n-s} \end{pmatrix}.$$

Ceci signifie qu'il existe des bases  $(v_1, \dots, v_n)$  de  $k^n$  et  $(w_1, \dots, w_m)$  de  $k^m$  telles que  $u(v_i) = w_i$  pour  $i = 1, \dots, s$ , et  $u(v_j) = 0$  pour  $j = s + 1, \dots, n$ . Alors  $\text{Im}(f) = \text{Vect}(w_1, \dots, w_s)$  est de dimension  $s$ , d'où  $s = \text{rang}(A)$ .

On déduit du corollaire 0.5.10 la proposition suivante.



**Proposition 0.5.11.** — Soit  $A \in M_{m,n}(k)$ . Alors

$$\boxed{\text{rang}(A) = \text{rang}({}^tA)}$$

par conséquent  $\text{rang}(A)$  est aussi le nombre maximum de lignes de  $A$  qui sont linéairement indépendantes.

*Démonstration.* D'après ce qui précède, il existe  $P \in \text{GL}_n(k)$  et  $Q \in \text{GL}_m(k)$  telles que

$$Q^{-1}AP = \begin{pmatrix} I_r & \mathbf{0}_{r, n-r} \\ \mathbf{0}_{m-r, r} & \mathbf{0}_{m-r, n-r} \end{pmatrix}.$$

où  $r = \text{rang}(A)$ . Alors

$${}^tP {}^tA {}^tQ^{-1} = \begin{pmatrix} I_r & \mathbf{0}_{r, m-r} \\ \mathbf{0}_{n-r, r} & \mathbf{0}_{n-r, m-r} \end{pmatrix}.$$

Or, d'après la proposition 0.5.2,  ${}^tP$  et  ${}^tQ^{-1}$  sont inversibles, donc l'égalité ci-dessus entraîne, d'après le corollaire précédent, que  $r = \text{rang}({}^tA)$ .



**Définition 0.5.12 (Matrices équivalentes).** — Soient  $A, B \in M_{m,n}(k)$ ; on dit que  $A$  et  $B$  sont *équivalentes* s'il existe des matrices inversibles  $P \in \text{GL}_n(k)$  et  $Q \in \text{GL}_m(k)$  telles que  $Q^{-1}AP = B$ . (D'après le corollaire 0.5.10, ceci équivaut à dire que  $A$  et  $B$  ont même rang).

**Cas des endomorphismes.** — Le théorème 0.5.7 traite le cas d'une application linéaire  $u : V \rightarrow W$ , où  $V, W$  sont *a priori* distincts. Dans ce cas, lorsque qu'on s'autorise des changements de bases arbitraires dans  $V$  et dans  $W$ , le corollaire 0.5.10 montre que le seul invariant de  $u$  est son rang, qui est un entier compris entre 0 et  $\text{Min}(\dim V, \dim W)$ .

Mais, lorsque  $V = W$  et qu'on s'intéresse à la nature géométrique d'un endomorphisme  $u$  de  $V$ , c.-à-d., lorsqu'on veut comparer  $u(x)$  et  $x$ , pour  $x$  variant dans  $V$ , pour pouvoir faire la comparaison on veut exprimer  $x$  et  $u(x)$  dans la même base, et c'est la raison pour laquelle, dans ce cas, on écrit la matrice de  $u$  dans une *même* base  $\mathcal{B}$  au départ et à l'arrivée.

(Par exemple, si  $V = W = k$ , les automorphismes de  $k$  comme  $k$ -espace vectoriel sont les homothéties  $h_\lambda : x \mapsto \lambda x$  avec  $\lambda \neq 0$ , si on prend  $\{1\}$  comme base de départ et  $\{1\}$  comme base d'arrivée, la matrice est (1) donc on a « perdu » le rapport  $\lambda$  de l'homothétie; mais si on impose de garder la même base au départ et à l'arrivée, la matrice est  $(\lambda)$ ..)

Alors, le théorème 0.5.7 donne dans ce cas :



**Théorème 0.5.13 (Changement de base pour un endomorphisme)**

Soit  $A$  la matrice d'un endomorphisme  $u$  de  $V$  relativement à une base  $\mathcal{B}$  de  $V$ . Si  $\mathcal{B}'$  est une seconde base, et si  $P$  est la matrice de passage de  $\mathcal{B}$  à  $\mathcal{B}'$ , alors la matrice de  $u$  dans la base  $\mathcal{B}'$  est :

$$\text{Mat}_{\mathcal{B}'}(u) = P^{-1}AP.$$



**Définition 0.5.14 (Matrices semblables).** — Soit  $A, B \in M_n(k)$  des matrices carrées de taille  $n$ . On dit que  $A, B$  sont *semblables* s'il existe une matrice inversible  $P \in \text{GL}_n(k)$  telle que  $P^{-1}AP = B$ . Dans ce cas, on dit que  $A, B$  sont dans la même *classe de similitude*.

- Remarques 0.5.15.** — 1) On notera que cette définition ne fait sens que pour des matrices carrées.  
 2)  $A$  et  $B$  sont semblables si et seulement si elles représentent, dans des bases différentes, le même endomorphisme de  $k^n$ .  
 3) Si  $A, B \in M_n(k)$  sont semblables, elles sont évidemment équivalentes, mais la réciproque est loin d'être vraie : les classes de similitude forment une partition de  $M_n(k)$  beaucoup plus raffinée que celle donnée par le rang (cf. le cas  $n = 1$ , et voir plus loin pour le cas  $n$  arbitraire).

**0.6. Appendice (†) : compléments sur les familles génératrices ou libres**

Dans cet appendice, on donne la définition des familles génératrices ou libres éventuellement infinies, ainsi qu'une démonstration du théorème sur la dimension (0.2.8).

**Définitions 0.6.1 (Familles génératrices).** — Soit  $V$  un  $k$ -espace vectoriel.

- 1) Soit  $S$  un sous-ensemble (fini ou infini) de  $V$ . L'ensemble de toutes les combinaisons linéaires *finies*

$$t_1v_1 + \dots + t_rv_r, \quad \text{pour } r \in \mathbb{N}^* \text{ (variable), } v_i \in S, t_i \in k,$$

forme un sous-espace vectoriel de  $V$ , et c'est le plus petit sous-espace vectoriel de  $V$  contenant  $S$ . On l'appelle le sous-espace *engendré* par  $S$  et on le note  $\text{Vect}(S)$ .

En particulier, si  $E$  est un sous-ensemble fini  $\{v_1, \dots, v_n\}$ , alors  $\text{Vect}(E)$  est l'ensemble des combinaisons linéaires

$$t_1v_1 + \dots + t_nv_n, \quad \text{où } t_1, \dots, t_n \in k.$$

- 2) On dit que  $S$  est un ensemble de générateurs (ou une *famille génératrice*) de  $V$  si le sous-espace engendré  $\text{Vect}(S)$  égale  $V$ , c.-à-d., si tout élément de  $V$  s'écrit comme combinaison linéaire des éléments de  $S$ .

Il résulte de la définition que : *toute famille contenant une famille génératrice est génératrice.*

**Exemple 0.6.2.** — Les monômes  $X^n$ , pour  $n \in \mathbb{N}$ , engendrent l'espace vectoriel  $k[X]$ . En effet, tout polynôme  $P \in k[X]$  s'écrit comme une combinaison linéaire finie :  $P = a_0 + a_1X + \dots + a_dX^d$ .

**Exercices 0.6.3.** — 1) Soit  $\mathbb{R}[[X]]$  l'espace des séries formelles  $\sum_{i=0}^{\infty} a_iX^i$ , avec  $a_i \in \mathbb{R}$ . Quel est le sous-espace de  $\mathbb{R}[[X]]$  engendré par les monômes  $X^n$ , pour  $n \in \mathbb{N}$ ? (Réponse : c'est  $\mathbb{R}[[X]]$ .)

- 2) Pouvez-vous montrer que l'espace vectoriel  $\mathbb{R}[X]$  n'est pas finiment engendré? (Voir Exercice 19 de la Feuille 1.)

**Définitions 0.6.4 (Familles libres ou liées).** — Soit  $V$  un  $k$ -espace vectoriel et soit  $S$  un sous-ensemble (fini ou infini) de  $V$ . On dit que les éléments de  $S$  sont *linéairement indépendants* (ou que  $S$  est une *famille libre*) s'il n'existe pas de relation linéaire non triviale entre les éléments de  $S$ , c.-à-d., si la condition suivante est vérifiée :

$$(FL) \quad \left\{ \begin{array}{l} \text{pour tous } r \in \mathbb{N}^*, v_1, \dots, v_r \in S \text{ deux-à-deux distincts, et } t_1, \dots, t_r \in k, \text{ si l'on a} \\ \text{une relation } t_1v_1 + \dots + t_nv_n = 0, \text{ alors } t_1 = 0 = \dots = t_r. \end{array} \right.$$

(Ceci prend une forme plus simple si  $S$  est un ensemble fini  $\{v_1, \dots, v_n\}$ ; dans ce cas la condition s'écrit plus simplement : pour tous  $t_1, \dots, t_n \in k$ , si  $t_1v_1 + \dots + t_nv_n = 0$ , alors  $t_1 = 0 = \dots = t_n$ .)

Il résulte de la définition que : *toute sous-famille d'une famille libre est libre.*

Au contraire, on dit que les éléments de  $S$  sont *linéairement dépendants* (ou que  $S$  est une *famille liée*) s'il existe une relation linéaire non triviale entre les éléments de  $S$ , c.-à-d., s'il existe un entier  $r \geq 1$ , des éléments  $v_1, \dots, v_r \in S$  deux-à-deux distincts, et des scalaires  $t_1, \dots, t_r$ , tous  $\neq 0$ , tels que  $t_1v_1 + \dots + t_nv_n = 0$ .

Il résulte de la définition que : *toute famille contenant une famille liée est liée.*

**Exemple 0.6.5.** — Dans  $k[X]$ , la famille des monômes  $(X^n)_{n \in \mathbb{N}}$  est libre. En effet, soient  $i_1 < \dots < i_r$  dans  $\mathbb{N}$  et soient  $t_1, \dots, t_r \in k$  tous non nuls. Alors le polynôme

$$t_1 X^{i_1} + \dots + t_r X^{i_r}$$

est non nul. Ceci montre que  $(X^n)_{n \in \mathbb{N}}$  est une famille libre.

**Remarque 0.6.6.** — Soient  $V$  un  $k$ -espace vectoriel et  $S = (v_i)_{i \in I}$  une famille de vecteurs indexée par un ensemble d'indice  $I$  infini. Alors  $S$  est libre si et seulement si la condition d'unicité suivante est vérifiée : (U) « si l'on a une égalité

$$\sum_{j \in J} t_j v_j = \sum_{p \in P} s_p v_p \quad t_j \in k, \quad s_p \in k,$$

où  $J, P$  sont deux sous-ensembles finis de  $I$ , alors  $\{j \in J \mid t_j \neq 0\}$  égale  $\{p \in P \mid s_p \neq 0\}$  et, notant  $L$  cet ensemble, on a  $t_\ell = s_\ell$  pour tout  $\ell \in L$  ».

En effet, supposons cette condition vérifiée, si l'on a une égalité  $\sum_{j \in J} t_j v_j = 0$  (le terme de droite correspond à  $P = \emptyset$  : une somme indexée par  $\emptyset$  vaut 0), alors  $\{j \in J \mid t_j \neq 0\}$  égale  $\emptyset$ , i.e. tous les  $t_j$  sont nuls ; ceci montre que  $S$  est une famille libre.

Réciproquement, supposons que  $S$  soit une famille libre, et qu'on ait une égalité  $\sum_{j \in J} t_j v_j = \sum_{p \in P} s_p v_p$  comme plus haut, alors on a :

$$0 = \sum_{j \in J-P} t_j v_j + \sum_{i \in J \cap P} (t_i - s_i) v_i - \sum_{p \in P-J} s_p v_p$$

et comme  $S$  est libre, ceci entraîne que  $t_j = 0 = s_p$  pour  $j \in J - P$  et  $p \in P - J$ , et que  $t_i = s_i$  pour tout  $i \in J \cap P$ , donc la condition (U) est vérifiée.

**Définition 0.6.7 (Bases).** — Soient  $V$  un  $k$ -espace vectoriel et  $(v_i)_{i \in I}$  une famille (finie ou infinie) de vecteurs de  $V$ . On dit que  $\mathcal{B} = (v_i)_{i \in I}$  est une *base* de  $V$  si tout élément  $v$  de  $V$  s'écrit de façon unique comme combinaison linéaire des  $v_i$ , c.-à-d., si :

(1)  $\mathcal{B}$  est une famille génératrice, i.e. pour tout  $v \in V$ , il existe un sous-ensemble fini  $J$  de  $I$  (dépendant de  $v$ ) et des scalaires  $t_j \in k$ , pour  $j \in J$ , tels que  $v = \sum_{j \in J} t_j v_j$ .

(2)  $\mathcal{B}$  vérifie la condition (U), i.e. si l'on a un second sous-ensemble fini  $P$  de  $I$  et des scalaires  $s_p$ , pour  $p \in P$ , tels que

$$v = \sum_{j \in J} t_j v_j = \sum_{p \in P} s_p v_p,$$

alors le sous-ensemble  $L = \{j \in J \mid t_j \neq 0\}$  égale  $\{p \in P \mid s_p \neq 0\}$  et pour tout  $\ell \in L$ , on a  $t_\ell = s_\ell$ .

D'après la remarque précédente, ceci équivaut à dire que  $\mathcal{B}$  est une famille *génératrice et libre*.

**Exemple 0.6.8.** — La famille des monômes  $(X^n)_{n \in \mathbb{N}}$  est une base de  $k[X]$  : en effet, tout polynôme non nul  $P \in k[X]$  s'écrit de façon unique

$$P = a_0 + \dots + a_d X^d, \quad \text{où } a_i \in k, a_d \neq 0, d = \deg P.$$

**Définitions 0.6.9.** — Soient  $V$  un  $k$ -espace vectoriel, et  $\mathcal{F} = (v_i)_{i \in I}$  une famille de vecteurs de  $V$ .

(i) On dit que  $\mathcal{F}$  est une *famille génératrice minimale* si c'est une famille génératrice, et si « on ne peut pas la rendre plus petite », c.-à-d., si pour tout sous-ensemble  $I' \neq I$ , la famille  $(v_i)_{i \in I'}$  n'est pas génératrice.

(ii) On dit que  $\mathcal{F}$  est une *famille libre maximale* si c'est une famille libre, et si « on ne peut pas la rendre plus grande », c.-à-d., si pour tout  $v \in V - \mathcal{F}$ , la famille  $\mathcal{F} \cup \{v\}$  est liée.

**Proposition 0.6.10.** — Soient  $V$  un  $k$ -espace vectoriel, et  $\mathcal{F} = (v_i)_{i \in I}$  une famille non vide de vecteurs de  $V$ . Les conditions suivantes sont équivalentes :

- a)  $\mathcal{F}$  est une base de  $V$  ;
- b)  $\mathcal{F}$  est une famille génératrice minimale ;
- c)  $\mathcal{F}$  est une famille libre maximale.

*Démonstration.* Supposons que  $(v_i)_{i \in I}$  soit une base de  $V$ , c.-à-d., une famille *génératrice et libre*. Alors, pour tout  $i \in I$ , la famille  $(v_j)_{j \in I - \{i\}}$  n'est pas génératrice : sinon  $v_i$  s'exprimerait comme combinaison linéaire des  $v_j$ , pour  $j \neq i$ , d'où une relation linéaire non triviale  $v_i - (t_1 v_{j_1} + \dots + t_r v_{j_r}) = 0$ . Ceci montre que  $(v_i)_{i \in I}$  est une famille génératrice minimale.

De plus, tout  $v \in V$  s'écrit comme combinaison linéaire (d'un nombre fini) des  $v_i$ , donc si  $v \notin \mathcal{F}$ , la famille strictement plus grande  $\mathcal{F} \cup \{v\}$  est liée. Ceci montre que  $\mathcal{F}$  est une famille libre maximale. On a donc prouvé que a) implique b) et c).

b)  $\Rightarrow$  a) Supposons que  $(v_i)_{i \in I}$  soit une famille génératrice minimale et montrons qu'elle est libre. Sinon, on aurait une relation linéaire non triviale

$$t_1 v_{i_1} + \dots + t_r v_{i_r} = 0$$

avec  $r \geq 1$  et les  $t_p \neq 0$ , d'où  $v_{i_r} = -t_r^{-1} \sum_{p \neq r} t_p v_{i_p}$  et donc la famille  $(v_i)_{i \in I - \{i_r\}}$  serait déjà génératrice, contredisant la minimalité. Ceci prouve b)  $\Rightarrow$  a).

c)  $\Rightarrow$  a) Supposons que  $\mathcal{F}$  soit une famille libre maximale et montrons qu'elle est génératrice. Soit  $v \in V - \mathcal{F}$ , alors la famille  $\mathcal{F} \cup \{v\}$  est liée, donc on a une relation linéaire

$$s v + t_1 v_{i_1} + \dots + t_r v_{i_r} = 0$$

non triviale (c.-à-d.,  $s$  et les  $t_p$  non tous nuls). On ne peut avoir  $s = 0$  car sinon, les  $v_i$  étant linéairement indépendants, on aurait  $t_1 = 0 = \dots = t_r$ . Donc  $s \neq 0$ , d'où  $v = -s^{-1}(t_1 v_{i_1} + \dots + t_r v_{i_r})$ . Donc  $\mathcal{F}$  est une famille libre et génératrice, donc une base de  $V$ . Ceci prouve c)  $\Rightarrow$  a).

**Corollaire 0.6.11 (Existence de bases).** — *Tout  $k$ -espace vectoriel finiment engendré  $V$  possède une base. (On convient que l'ensemble vide  $\emptyset$  est une base de l'espace nul  $\{0\}$ .)*

En effet, par hypothèse  $V$  est engendré par une famille finie  $\{v_1, \dots, v_r\}$ . Celle-ci contient une sous-famille génératrice  $\mathcal{B}$  de cardinal minimal, donc minimale, et d'après la proposition,  $\mathcal{B}$  est une base de  $V$ .

On va voir dans un instant (cf. 0.6.13 ci-dessous) que si  $(v_1, \dots, v_n)$  est une base de  $V$ , alors toutes les bases de  $V$  ont  $n$  éléments. Commençons par le lemme important suivant.

**Lemme 0.6.12.** — *Soient  $V$  un  $k$ -espace vectoriel,  $v_1, \dots, v_n \in V$ , et soient  $m > n$  et  $u_1, \dots, u_m$  des éléments de  $V$  qui sont combinaison linéaire de  $v_1, \dots, v_n$ . Alors  $u_1, \dots, u_m$  sont liés.*

*Démonstration.* On procède par récurrence sur  $n$ ; le résultat est clair si  $n = 1$ , supposons donc  $n \geq 2$  et le résultat établi pour  $n - 1$ . Écrivons :

$$\begin{cases} u_1 & = & a_{11}v_1 + \dots + a_{1n}v_n \\ u_2 & = & a_{21}v_1 + \dots + a_{2n}v_n \\ \vdots & & \vdots \\ u_m & = & a_{m1}v_1 + \dots + a_{mn}v_n. \end{cases}$$

Si tous les  $u_i$  sont combinaison linéaire des  $n - 1$  vecteurs  $v_2, \dots, v_n$ , alors les  $u_i$  sont liés, d'après l'hypothèse de récurrence. Sinon, quitte à renuméroter les  $u_i$ , on peut supposer que  $a_{11} \neq 0$ . Alors, pour  $i = 2, \dots, m$ , les  $m - 1$  vecteurs

$$u'_i = u_i - a_{i1}a_{11}^{-1}u_1$$

sont combinaison linéaire des  $n - 1$  vecteurs  $v_2, \dots, v_n$ , donc sont liés d'après l'hypothèse de récurrence. Donc il existe des scalaires non tous nuls  $t_2, \dots, t_m$  tels que

$$0 = t_2 u'_2 + \dots + t_m u'_m = t_2 u_2 + \dots + t_m u_m - \left( \sum_{i=2}^m t_i a_{i1} a_{11}^{-1} \right) u_1$$

et ceci montre que les  $u_i$  sont liés. Le lemme est démontré.

Le lemme précédent a les conséquences très importantes suivantes. Soit  $V$  un  $k$ -espace vectoriel engendré par un nombre fini d'éléments  $x_1, \dots, x_N$ . D'après le corollaire 0.6.11,  $V$  possède une base  $\mathcal{B} = (v_1, \dots, v_n)$  formée de  $n \leq N$  éléments. Comme  $\mathcal{B}$  est une famille libre de  $V$ , il résulte du lemme précédent que :

- a) toute famille génératrice de  $V$  a au moins  $n$  éléments ;
- b) toute famille génératrice de  $V$  ayant  $n$  éléments est *minimale*, donc d'après 0.6.10 est une base de  $V$ .

D'autre part, comme  $\mathcal{B}$  est une famille génératrice de  $V$ , il résulte du lemme précédent que :

- a') toute famille libre dans  $V$  a au plus  $n$  éléments ;
- b') toute famille libre de  $V$  ayant  $n$  éléments est *maximale*, donc d'après 0.6.10 est une base de  $V$ .

Enfin, en combinant a) et b), on voit que : *toute base de  $V$ , étant une famille à la fois génératrice et libre, a exactement  $n$  éléments.* On obtient donc le théorème fondamental suivant.

**Théorème 0.6.13 (Dimension d'un espace vectoriel).** — *Soit  $V$  un  $k$ -espace vectoriel finiment engendré.*

(i) *Il existe des bases de  $V$ , et toutes ont même cardinal  $n$ ; cet entier s'appelle la dimension de  $V$  sur  $k$  et se note  $\dim_k V$  ou simplement  $\dim V$ .*

(ii) *De toute famille génératrice  $\mathcal{F}$  on peut extraire une base, en particulier  $\mathcal{F}$  est de cardinal  $\geq n$ ; de plus si  $\text{card}(\mathcal{F}) = n$  alors  $\mathcal{F}$  est une base de  $V$ .*

(iii) *Toute famille libre est de cardinal  $\leq n$ , et toute famille libre de cardinal  $n$  est une base de  $V$ .*

(iv) « Théorème de la base incomplète » : *Toute famille libre peut être complétée en une base de  $V$ .*

(v) *Tout sous-espace  $W$  de  $V$  est finiment engendré, et  $\dim_k W \leq \dim_k V$ ; de plus si  $\dim_k W = \dim_k V$ , alors  $W = V$ . En d'autres termes, tout sous-espace vectoriel distinct de  $V$  est de dimension  $< \dim_k V$ .*

*Démonstration.* On a déjà vu les assertions (i), (ii) et (iii). L'assertion (iv) résulte du fait que toute famille libre peut être agrandie en une famille libre maximale, c.-à-d., en une base de  $V$  (cf. la proposition 0.6.10).

Démontrons (v). Soit  $W$  un sous-espace vectoriel de  $V$ . D'après (iii), toute famille libre d'éléments de  $W$  est de cardinal  $\leq n = \dim_k V$ , donc  $W$  possède une famille *libre maximale*  $\mathcal{C}$ , de cardinal  $m \leq n$ . Alors  $\mathcal{C}$  est une *base* de  $W$ , d'après la proposition 0.6.10, donc  $W$  est finiment engendré, et de dimension  $m \leq n$ . Si de plus  $m = n$  alors, d'après (iii),  $\mathcal{C}$  est une base de  $V$  (donc engendre  $V$ ), d'où  $W = V$ . Le théorème est démontré.

**Proposition 0.6.14.** — *Soit  $f : V \rightarrow W$  une application linéaire.*

a) *Si  $f$  est injective et si  $\mathcal{F}$  est une famille libre de  $V$ , alors  $f(\mathcal{F})$  est libre.*

b) *Si  $f$  est surjective et si  $\mathcal{F}$  est une famille génératrice de  $V$ , alors  $f(\mathcal{F})$  engendre  $W$ .*

c) *Si  $f$  est bijective et si  $\mathcal{B}$  est une base de  $V$ , alors  $f(\mathcal{B})$  est une base de  $W$ .*

*Démonstration.* a) Supposons  $f$  injective et soit  $\mathcal{F}$  une famille libre de  $V$ . S'il existe une relation linéaire dans  $W$  :

$$t_1 f(x_1) + \cdots + t_n f(x_n) = 0, \quad t_i \in k, \quad x_i \in \mathcal{F},$$

alors  $0 = f(t_1 x_1 + \cdots + t_n x_n)$ , donc comme  $f$  est injective  $t_1 x_1 + \cdots + t_n x_n = 0$ , donc comme  $\mathcal{F}$  est libre,  $t_i = 0$  pour  $i = 1, \dots, n$ . Ceci montre que  $f(\mathcal{F})$  est libre.

b) Supposons  $f$  surjective et soit  $\mathcal{F}$  une famille génératrice de  $V$ . Soit  $w \in W$ ; comme  $f$  est surjective, il existe  $v \in V$  tel que  $f(v) = w$ . Comme  $\mathcal{F}$  engendre  $V$ , il existe  $x_1, \dots, x_n \in \mathcal{F}$  et  $t_1, \dots, t_n \in k$  tels que  $v = t_1 x_1 + \cdots + t_n x_n$ , d'où  $w = t_1 f(x_1) + \cdots + t_n f(x_n)$ . Ceci montre que  $f(\mathcal{F})$  engendre  $W$ .

c) Supposons  $f$  bijective et soit  $\mathcal{B}$  une base de  $V$ . Alors, d'après a) et b),  $f(\mathcal{B})$  est une famille libre et génératrice de  $W$ , donc une base de  $W$ .





# CHAPITRE 1

## DUAL, OPÉRATIONS SUR LES COLONNES OU LES LIGNES, DÉTERMINANTS, VALEURS ET VECTEURS PROPRES

**Résumé :** Soient  $V$  un  $k$ -espace vectoriel de dimension  $n$  et  $\mathcal{B}$  une base de  $V$ . Dans ce chapitre, on commence par introduire l'espace dual  $V^*$  et la base duale  $\mathcal{B}^*$  de  $\mathcal{B}$ , qui est formée des « fonctions coordonnées » relativement à la base  $\mathcal{B}$ . Ensuite, on introduit les opérations élémentaires sur les *colonnes* d'une matrice  $A \in M_{m,n}(k)$ , ce qui est utile pour déterminer une *base* de  $\text{Ker}(A)$  et de  $\text{Im}(A)$ . Puis on revoit les opérations sur les lignes (vues en L1) et l'on fait le lien, d'une part, avec la théorie des systèmes linéaires (vue en L1) et, d'autre part, avec les formes linéaires sur  $V$ , i.e. les éléments de  $V^*$ , et l'on applique ceci aux équations qui définissent un sous-espace  $E$  de  $V$  (cf. §1.3).

On montre aussi comment les opérations sur les colonnes (ou bien les lignes) permettent de déterminer si une matrice carrée  $A \in M_n(k)$  est inversible et de calculer, dans ce cas, son inverse.

Enfin, on rappelle la construction et les propriétés du déterminant, et plus généralement du polynôme caractéristique, d'une matrice  $A \in M_n(k)$  ou d'un endomorphisme  $u$  de  $V$ . On peut laisser de côté la démonstration, par contre il faut connaître les propriétés du déterminant et du polynôme caractéristique énoncées dans 1.4.1, 1.4.6 et 1.5.2, ainsi que les formules de développement par rapport à une ligne ou une colonne. Bien entendu, la réduction des colonnes s'applique aussi au calcul du polynôme caractéristique et des vecteurs propres, cf. la feuille d'exercices n°1.

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 des appendices à la fin du chapitre ; ces passages n'interviendront pas dans les évaluations.

### 1.1. Formes linéaires, espace dual

**Définition 1.1.1 (Espace dual).** — Un cas particulier très important d'espace  $\mathcal{L}(V, W)$  est le cas où  $W = k$ . Dans ce cas, une application  $k$ -linéaire  $V \rightarrow k$  s'appelle une *forme linéaire* sur  $V$ , et  $\mathcal{L}(V, k)$  s'appelle *l'espace dual* de  $V$  et se note  $V^*$ .

Supposons  $V$  de dimension finie  $n$  et soit  $\mathcal{B} = (e_1, \dots, e_n)$  une base de  $V$ . On sait, d'après 0.4.2, que se donner une forme linéaire  $\phi : V \rightarrow k$  équivaut à se donner ses valeurs  $\phi(e_j) \in k$  sur les  $e_j$ . Pour  $i = 1, \dots, n$ , notons alors  $e_i^*$  la forme linéaire sur  $V$  définie par

$$e_i^*(e_i) = 1 \quad \text{et} \quad e_i^*(e_j) = 0 \quad \text{pour } j \neq i.$$

Alors, pour tout  $n$ -uplet  $(t_1, \dots, t_n) \in k^n$ , la forme linéaire  $t_1 e_1^* + \dots + t_n e_n^*$  vérifie :

$$(\star) \quad \forall j = 1, \dots, n, \quad (t_1 e_1^* + \dots + t_n e_n^*)(e_j) = \sum_{i=1}^n t_i \underbrace{e_i^*(e_j)}_{\substack{=1 \text{ si } i=j \\ =0 \text{ sinon}}} = t_j.$$

Il en résulte que pour tout  $\phi \in V^*$ , on a  $\phi = \phi(e_1)e_1^* + \dots + \phi(e_n)e_n^*$ , puisque ces deux applications linéaires  $V \rightarrow k$  coïncident sur la base  $\mathcal{B}$ . De plus, si l'on a une égalité  $\phi = t_1 e_1^* + \dots + t_n e_n^*$ , avec les  $t_i \in k$ , alors d'après  $(\star)$  on a nécessairement  $t_j = \phi(e_j)$  pour tout  $j$ . Ceci montre que tout  $\phi \in V^*$  s'écrit de façon unique comme combinaison des formes linéaires  $e_1^*, \dots, e_n^*$ , donc  $\mathcal{B}^* = (e_1^*, \dots, e_n^*)$  est une *base* de  $V^*$ , appelée la *base duale* de la base  $\mathcal{B} = (e_1, \dots, e_n)$ . En particulier,  $V^*$  est de dimension  $n$ . On a donc démontré le :

<sup>(0)</sup>version du 7/7/2012



**Théorème 1.1.2 (Base duale).** — Soient  $V$  un  $k$ -espace vectoriel de dimension  $n$  et  $\mathcal{B} = (e_1, \dots, e_n)$  une base de  $V$ . Pour  $i = 1, \dots, n$ , on note  $e_i^*$  la forme linéaire sur  $V$  définie par  $e_i^*(e_i) = 1$  et  $e_i^*(e_j) = 0$  pour  $j \neq i$ . Alors  $\mathcal{B}^* = (e_1^*, \dots, e_n^*)$  est une base de  $V^*$ , appelée la base duale de la base  $\mathcal{B}$ ; en effet toute forme linéaire  $\phi \in V^*$  s'écrit de façon unique comme combinaison linéaire des  $e_i^*$ , i.e. on a :

$$\phi = t_1 e_1^* + \dots + t_n e_n^*, \quad \text{avec} \quad t_j = \phi(e_j).$$

En particulier,  $V^*$  est de dimension  $n$ .

De plus,  $e_1^*, \dots, e_n^*$  sont les formes linéaires « coordonnées » par rapport à la base  $(e_1, \dots, e_n)$ , c.-à-d., tout  $v \in V$  s'écrit de façon unique  $v = x_1 e_1 + \dots + x_n e_n$ , et l'on a  $e_i^*(v) = x_i$  pour tout  $i$ .

**1.1.3. Matrices colonnes et matrices lignes.** — Soient  $V$  un  $k$ -espace vectoriel de dimension  $n$  et  $\mathcal{B} = (e_1, \dots, e_n)$  une base de  $V$ . Alors, tout vecteur  $v = x_1 e_1 + \dots + x_n e_n$  est représenté par le vecteur colonne

$$v = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in M_{n,1}(k).$$

D'autre part, toute forme linéaire  $\phi \in V^*$  est un élément de  $\mathcal{L}(V, k)$  donc est représentée par une matrice  $L \in M_{1,n}(k)$ , i.e. par une matrice ligne

$$L = (t_1 \quad \dots \quad t_n), \quad \text{où} \quad t_j = \phi(e_j).$$

Alors on a  $\phi(v) = \sum_{j=1}^n \phi(x_j e_j) = \sum_{j=1}^n x_j \underbrace{\phi(e_j)}_{=t_j} = \sum_{j=1}^n x_j t_j = t_1 x_1 + \dots + t_n x_n$  i.e.  $\phi(v)$  est donné

par le produit matriciel suivant :  $\phi(v) = L \cdot v = (t_1 \quad \dots \quad t_n) \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ . Donc, pour des entiers  $p, r \geq 1$

arbitraires, si l'on a  $p$  formes linéaires  $\phi_1, \dots, \phi_p \in V^*$ , représentées par des matrices lignes (à  $n$  colonnes)  $L_1 = (t_{11} \quad \dots \quad t_{1n}), \dots, L_p = (t_{p1} \quad \dots \quad t_{pn})$  et  $r$  vecteurs

$$v_1 = \begin{pmatrix} x_{11} \\ \vdots \\ x_{n1} \end{pmatrix}, \quad \dots, \quad v_r = \begin{pmatrix} x_{1r} \\ \vdots \\ x_{nr} \end{pmatrix}$$

alors, notant  $\begin{pmatrix} L_1 \\ \vdots \\ L_p \end{pmatrix}$  l'élément de  $M_{p,n}(k)$  dont les lignes sont  $L_1, \dots, L_p$  et  $(v_1 \dots v_r)$  l'élément de  $M_{n,r}(k)$

dont les colonnes sont  $v_1, \dots, v_r$ , on a l'égalité suivante de matrices à  $p$  lignes et  $r$  colonnes :

$$(\star) \quad \boxed{\begin{pmatrix} \phi_1(v_1) & \dots & \phi_1(v_r) \\ \vdots & & \vdots \\ \phi_p(v_1) & \dots & \phi_p(v_r) \end{pmatrix} = \begin{pmatrix} L_1 \\ \vdots \\ L_p \end{pmatrix} (v_1 \quad \dots \quad v_r).$$

On en déduit la proposition suivante :

**Proposition 1.1.4 (Matrice de passage pour les bases duales. Base « préduale »)**

Soient  $V$  un  $k$ -espace vectoriel de dimension  $n$ ,  $\mathcal{B}$  une base de  $V$ , et  $\mathcal{B}^*$  la base duale de  $V^*$ .

(1) Soient  $\mathcal{C}$  une seconde base de  $V$ ,  $\mathcal{C}^*$  la base duale, et  $P = \text{Mat}_{\mathcal{B}}(\mathcal{C})$ . Alors la matrice de passage  $\text{Mat}_{\mathcal{B}^*}(\mathcal{C}^*)$  égale  ${}^t P^{-1}$ .

(2) Pour toute base  $\Delta$  de  $V^*$ , il existe une unique base  $\mathcal{D}$  de  $V$  telle que  $\mathcal{D}^* = \Delta$ . On l'appellera la base « préduale » de  $\Delta$ .

*Démonstration.* Notons  $\mathcal{B} = (e_1, \dots, e_n)$ ,  $\mathcal{B}^* = (e_1^*, \dots, e_n^*)$  et  $\mathcal{C} = (v_1, \dots, v_n)$ ,  $\mathcal{C}^* = (v_1^*, \dots, v_n^*)$ . Notons  $Q = (q_{ij})$  la matrice de passage  $\text{Mat}_{\mathcal{B}^*}(\mathcal{C}^*)$ , i.e. pour  $j = 1, \dots, n$ , on a  $v_j^* = q_{1j} e_1^* + \dots + q_{nj} e_n^*$ . Donc  $v_j^*$  correspond à la matrice ligne

$$L_j = (q_{1j}, \dots, q_{nj}) = {}^t C_j,$$



où  $C_j$  est la  $j$ -ième colonne de  $Q$ . Donc la matrice  $\begin{pmatrix} L_1 \\ \vdots \\ L_n \end{pmatrix}$  égale  ${}^tQ$ . D'autre part, chaque  $v_j$  est la  $j$ -ième colonne de  $P = \text{Mat}_{\mathcal{B}}(\mathcal{C})$ . Donc, d'après  $(\star)$  on obtient :

$${}^tQP = \begin{pmatrix} L_1 \\ \vdots \\ L_n \end{pmatrix} (v_1 \ \cdots \ v_n) = \begin{pmatrix} v_1^*(v_1) & \cdots & v_1^*(v_n) \\ \vdots & & \vdots \\ v_n^*(v_1) & \cdots & v_n^*(v_n) \end{pmatrix} = I_n$$

(la dernière égalité découlant du fait que  $v_i^*(v_j) = 1$  si  $i = j$  et  $= 0$  sinon). Ceci prouve que  ${}^tQ = P^{-1}$ , i.e.  $Q = {}^tP^{-1}$ . L'assertion (1) est démontrée.

L'assertion (2) en découle. En effet, notons encore  $Q$  la matrice de passage  $\text{Mat}_{\mathcal{B}^*}(\Delta)$ . S'il existe une base  $\mathcal{D}$  de  $V$  telle que  $\mathcal{D}^* = \Delta$  alors, d'après l'assertion (1), la matrice de passage  $P = \text{Mat}_{\mathcal{B}}(\mathcal{D})$  vérifie  ${}^tP^{-1} = Q$ , i.e.  $P = {}^tQ^{-1}$ . Donc  $\mathcal{D}$  est déterminée (si elle existe) par la condition  $\text{Mat}_{\mathcal{B}}(\mathcal{D}) = {}^tQ^{-1}$ ; or cette condition définit bien une base de  $V$  : comme la matrice  $M = {}^tQ^{-1}$  est inversible, ses colonnes  $v_1, \dots, v_n$  forment une base  $\mathcal{D}$  de  $V$ , telle que  $\text{Mat}_{\mathcal{B}}(\mathcal{D}) = M$ . Ceci prouve l'assertion (2).

## 1.2. Opérations élémentaires sur les colonnes (ou les lignes)

**1.2.1. Réductions des colonnes.** — Revenons au calcul du rang d'une application linéaire  $k^n \rightarrow k^m$ . Soit  $A \in M_{m,n}(k)$ , notons  $(e_1, \dots, e_n)$  la base canonique de  $k^n$ . On peut calculer le rang de  $A$  et, plus précisément, une *base* de  $\text{Im}(A)$  et de  $\text{Ker}(A)$  de la façon suivante.

Observons qu'échanger les colonnes  $C_i$  et  $C_j$  de  $A$  revient à permuter, avant d'appliquer  $A$ , les vecteurs  $e_i$  et  $e_j$  de la base canonique de  $k^n$ , ce qui revient aussi à multiplier  $A$  à *droite* par la matrice  $P(i, j)$  de l'automorphisme de  $k^n$  qui échange  $e_i$  et  $e_j$ , et laisse fixe chaque  $e_\ell$  pour  $\ell \neq i, j$ . Par exemple, pour  $n = 4$

et  $(i, j) = (1, 4)$  on a  $P(1, 4) = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}$ .

De même, multiplier une colonne  $C_j$  par un scalaire  $\lambda_j \neq 0$  revient à remplacer le vecteur  $e_j$  par  $\lambda_j e_j$ , c.-à-d., à multiplier  $A$  à *droite* par la matrice diagonale dont tous les termes diagonaux valent 1, sauf celui d'indice  $j$  qui vaut  $\lambda_j$ . D'autre part, pour tout  $t \in k$  et  $i \neq j$ , ajouter  $tC_i$  à  $C_j$  revient à remplacer, avant d'appliquer  $A$ , le vecteur  $e_j$  par le vecteur  $e_j + te_i$ , et ceci revient encore à multiplier  $A$  à *droite* par la matrice  $B_{ij}(t)$  de l'automorphisme de  $k^n$  qui laisse fixe  $e_\ell$  pour  $\ell \neq j$ , et envoie  $e_j$  sur  $e_j + te_i$ , c.-à-d.,

$$B_{ij}(t) = I_n + tE_{ij}.$$



**Définition 1.2.2.** — On appellera *opérations élémentaires sur les colonnes* les opérations précédentes : échange de colonnes, multiplication d'une colonne par un scalaire  $\neq 0$ , ou ajout de  $tC_i$  à  $C_j$  avec  $j \neq i$ . D'après ce qui précède, on voit qu'effectuer ces opérations sur les colonnes revient à appliquer des automorphismes sur l'espace *de départ*  $k^n$ , donc ne change pas l'image de  $A$ , ni son rang.

On peut alors calculer  $\text{Im}(A)$  de la façon suivante. Soit  $i_1$  l'indice de la première ligne non nulle. Alors en permutant les colonnes, on peut supposer que  $a_{i_1,1} \neq 0$  puis, en multipliant la première colonne par  $a_{i_1,1}^{-1}$  on se ramène au cas où  $a_{i_1,1} = 1$ . Ensuite, en soustrayant  $a_{i_1,j}C_1$  de  $C_j$ , on se ramène au cas où  $a_{i_1,j} = 0$  pour  $j \geq 2$ . Soit alors  $i_2$  le plus petit indice tel qu'il existe  $j \geq 2$  tel que  $a_{i_2,j} \neq 0$ ; en procédant comme plus haut, on se ramène au cas où  $a_{i_2,2} = 1$  et  $a_{i_2,j} = 0$  pour  $j \geq 3$ . Soit alors  $i_3$  le plus petit indice tel qu'il existe  $j \geq 3$  tel que  $a_{i_3,j} \neq 0$ , en répétant le processus précédent, on obtient une matrice de la forme suivante :

$$(\dagger) \quad A' = \begin{pmatrix} 0 & 0 & 0 & 0 & \cdots & 0 & \cdots & 0 \\ 1 & 0 & 0 & 0 & \cdots & 0 & \cdots & 0 \\ a_{2,1} & 0 & 0 & 0 & \cdots & 0 & \cdots & 0 \\ \vdots & 0 & 0 & 0 & \cdots & 0 & \cdots & 0 \\ a_{i_2,1} & 1 & 0 & 0 & \cdots & 0 & \cdots & 0 \\ * & * & 0 & 0 & \cdots & 0 & \cdots & 0 \\ \vdots & * & 0 & 0 & \cdots & 0 & \cdots & 0 \\ a_{i_3,1} & a_{i_3,2} & 1 & 0 & \cdots & 0 & \cdots & 0 \\ * & * & * & 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ a_{i_r,1} & a_{i_r,2} & a_{i_r,3} & \cdots & 1 & 0 & \cdots & 0 \\ * & * & * & \cdots & * & 0 & \cdots & 0 \\ * & * & * & \cdots & * & 0 & \cdots & 0 \end{pmatrix}$$

c.-à-d., les colonnes d'indice  $> r$  sont nulles et, pour  $j = 1, \dots, r$ , les colonnes  $C'_j$  vérifient

$$C'_j = e_{i_j} + \sum_{\ell > i_j} a_{\ell,j} e_\ell,$$

avec  $i_1 < i_2 < \dots < i_r$ . Il en résulte que les vecteurs  $C'_1, \dots, C'_r \in k^m$  sont linéairement indépendants, donc forment une base de  $\text{Im}(u)$ ; en particulier,  $r = \text{rang}(A)$ .

Le procédé précédent s'appelle « réduction des colonnes », la matrice  $A'$  ainsi obtenue vérifie

$$(1) \quad A' = AP$$

où  $P$  est une matrice inversible correspondant aux opérations élémentaires sur les colonnes qu'on a effectuées. On a vu que  $\text{Im}(A') = \text{Im}(A)$ ; d'autre part, le noyau de  $A'$  est le sous-espace  $V$  de  $k^n$  engendré par les vecteurs  $e_{r+1}, \dots, e_n$  de la base canonique. Or, il résulte de (1) que

$$(2) \quad \text{Ker}(A) = P(V) = \text{Vect}(P_{r+1}, \dots, P_n),$$

où  $P_{r+1}, \dots, P_n$  désignent les  $(n-r)$  dernières colonnes de  $P$ . En effet, si  $v \in V$ , alors  $0 = A'v = APv$  donc  $Pv \in \text{Ker}(A)$ . Réciproquement, comme  $P$  est inversible, on a  $A = A'P^{-1}$  donc si  $x \in \text{Ker}(A)$  alors  $0 = Ax = A'P^{-1}x$ , d'où  $P^{-1}x \in \text{Ker}(A') = V$ , donc  $x \in P(V)$ .

L'égalité (2) permet de déterminer explicitement une base de  $\text{Ker}(A)$  de la façon suivante. Les opérations sur les colonnes qu'on a faites correspondent à multiplier  $A$  à droite par certaines matrices inversibles d'un type particulier, et  $P$  est le produit de ces matrices; alors la même suite d'opérations sur les colonnes de la matrice identité  $I_n$  conduit à la matrice  $P$ .

Donc, en pratique, on écrit en-dessous de  $A$  la matrice  $I_n$ , où  $n =$  nombre de colonnes de  $A =$  dimension de l'espace de départ, et à chaque étape on fait les mêmes opérations élémentaires sur les colonnes des deux matrices, on obtient alors à la fin du processus la matrice  $A' = AP$  ainsi que la matrice  $P = I_n P$ .<sup>(1)</sup> Illustrons ceci dans le cas de l'exemple suivant :

$$A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 3 & 6 & 6 \\ 3 & 8 & 15 & 16 \end{pmatrix}.$$

Écrivons  $I_4$  en-dessous de  $A$  et, notant  $C_1, \dots, C_4$  les colonnes de  $A$ , faisons les opérations  $C_i \rightarrow C_i - iC_1$  pour  $i = 2, 3, 4$  :

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 3 & 6 & 6 \\ 3 & 8 & 15 & 16 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \xrightarrow[\begin{smallmatrix} C_3 \rightarrow C_3 - 3C_1 \\ C_4 \rightarrow C_4 - 4C_1 \end{smallmatrix}]{\begin{smallmatrix} C_2 \rightarrow C_2 - 2C_1 \\ C_3 \rightarrow C_3 - 3C_1 \\ C_4 \rightarrow C_4 - 4C_1 \end{smallmatrix}} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 3 & 2 \\ 3 & 2 & 6 & 4 \\ 1 & -2 & -3 & -4 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \xrightarrow[\begin{smallmatrix} C_4 \rightarrow C_4 - 2C_2 \end{smallmatrix}]{\begin{smallmatrix} C_3 \rightarrow C_3 - 3C_2 \\ C_4 \rightarrow C_4 - 2C_2 \end{smallmatrix}} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 3 & 2 & 0 & 0 \\ 1 & -2 & 3 & 0 \\ 0 & 1 & -3 & -2 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

<sup>(1)</sup>Lorsque  $m = n$ , c.-à-d., lorsqu'on considère des matrices carrées, on peut utiliser ce processus pour déterminer si  $A$  est inversible et calculer son inverse; voir 1.2.8 plus bas.

Donc  $\text{Im}(A)$  et  $\text{Ker}(A)$  ont respectivement pour base les vecteurs

$$\begin{pmatrix} 1 \\ 1 \\ 3 \end{pmatrix} \text{ et } \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix} \quad \text{resp.} \quad \begin{pmatrix} 3 \\ -3 \\ 1 \\ 0 \end{pmatrix} \text{ et } \begin{pmatrix} 0 \\ -2 \\ 0 \\ 1 \end{pmatrix}.$$

En résumé, la « réduction des colonnes » d'une matrice  $A \in M_{m,n}(k)$  fournit des bases de  $\text{Im}(A)$  et de  $\text{Ker}(A)$ . Mais, étant donné un vecteur  $Y = (y_1, \dots, y_m)$  de  $k^m$  (resp.  $X = (x_1, \dots, x_n)$  de  $k^n$ ), il n'est pas forcément immédiat de déterminer si  $Y \in \text{Im}(A)$  (resp. si  $X \in \text{Ker}(A)$ ), et il est parfois plus commode de disposer d'équations définissant  $\text{Im}(A)$  et  $\text{Ker}(A)$ . On va voir que celles-ci sont obtenues par réduction des lignes de  $A$ .

**1.2.3. Remarque importante.** — Dans l'exemple précédent, le premier pivot était en position (1,1), puis le second en position (2,2). Bien sûr, il n'en est pas toujours ainsi; en théorie on peut s'y ramener par des échanges de colonnes, mais avec un peu de pratique on n'a pas besoin de faire ces échanges de colonnes : il suffit de se ramener à une matrice  $A'$  dont les colonnes soient échelonnées « à permutation près des colonnes »; alors les colonnes non nulles de  $A'$  forment une base de  $\text{Im}(A)$ , et dans la matrice du dessous (obtenue à partir de  $I_n$ ), les colonnes en-dessous des colonnes nulles de  $A'$  forment une base de  $\text{Ker}(A)$ . Illustrons ceci dans le cas suivant :



$$\begin{pmatrix} 3 & 2 & 4 & 1 \\ 6 & 3 & 6 & 1 \\ 15 & 8 & 16 & 3 \\ \hline 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \xrightarrow[\begin{smallmatrix} C_2 \rightarrow C_2 - 2C_4 \\ C_1 \rightarrow C_1 - 3C_4 \end{smallmatrix}]{\begin{smallmatrix} C_3 \rightarrow C_3 - 4C_4 \\ C_2 \rightarrow C_2 - 2C_4 \end{smallmatrix}} \begin{pmatrix} 0 & 0 & 0 & 1 \\ 3 & 1 & 2 & 1 \\ 6 & 2 & 4 & 3 \\ \hline 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -3 & -2 & -4 & 1 \end{pmatrix} \xrightarrow[\begin{smallmatrix} C_3 \rightarrow C_3 - 2C_2 \end{smallmatrix}]{\begin{smallmatrix} C_1 \rightarrow C_1 - 3C_2 \\ C_3 \rightarrow C_3 - 2C_2 \end{smallmatrix}} \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 2 & 0 & 3 \\ \hline 1 & 0 & 0 & 0 \\ -3 & 1 & -2 & 0 \\ 0 & 0 & 1 & 0 \\ 3 & -2 & 0 & 1 \end{pmatrix}.$$

Donc  $\text{Im}(A)$  et  $\text{Ker}(A)$  ont respectivement pour base les vecteurs

$$\begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix} \text{ et } \begin{pmatrix} 1 \\ 1 \\ 3 \end{pmatrix} \quad \text{resp.} \quad \begin{pmatrix} 1 \\ -3 \\ 0 \\ 3 \end{pmatrix} \text{ et } \begin{pmatrix} 0 \\ -2 \\ 1 \\ 0 \end{pmatrix}.$$

**1.2.4. Réduction des lignes.** — Soit toujours  $A \in M_{m,n}(k)$ . On peut calculer le rang de  $A$  et, plus précisément, des équations de  $\text{Ker}(A)$  et de  $\text{Im}(A)$  en faisant des opérations sur les lignes de  $A$ .

Observons que multiplier une ligne  $L_i$  de  $A$  par un scalaire  $\lambda_i \neq 0$  revient à multiplier  $A$  à gauche par la matrice diagonale inversible dont tous les termes diagonaux valent 1, sauf celui d'indice  $i$  qui vaut  $\lambda_i$ . De même, pour tout  $t \in k$  et  $i \neq j$ , ajouter  $tL_i$  à  $L_j$  revient à multiplier  $A$  à gauche par la matrice inversible

$$B_{ji}(t) = I_m + tE_{ji}$$

(dont l'inverse est  $B_{ji}(-t)$ ). Enfin, échanger les lignes  $L_i$  et  $L_j$  de  $A$  revient à multiplier  $A$  à gauche par la matrice de l'automorphisme de  $k^m$  qui échange les vecteurs  $f_i$  et  $f_j$  de la base canonique  $(f_1, \dots, f_m)$ , et laisse fixe chaque  $f_\ell$  pour  $\ell \neq i, j$  (cette matrice est égale à son inverse).



**Définition 1.2.5.** — On appellera opérations élémentaires sur les lignes les opérations précédentes : échange de lignes, multiplication d'une ligne par un scalaire  $\neq 0$ , ou ajout de  $tL_i$  à  $L_j$  avec  $j \neq i$ . D'après ce qui précède, on voit qu'effectuer ces opérations revient à appliquer des automorphismes sur l'espace d'arrivée  $k^m$ , donc ne change pas le noyau de  $A$ , ni son rang.

On peut alors calculer des équations de  $\text{Ker}(A)$  et  $\text{Im}(A)$  de la façon suivante. Soit  $j_1$  l'indice de la première colonne non nulle. En permutant les lignes, on peut supposer que  $a_{1,j_1} \neq 0$  puis, en multipliant la première ligne par  $a_{1,j_1}^{-1}$ , on se ramène au cas où  $a_{1,j_1} = 1$ . Ensuite, en soustrayant  $a_{i,j_1}L_1$  de  $L_i$  pour tout  $i \geq 2$ , on se ramène au cas où  $a_{i,j_1} = 0$  pour  $i \geq 2$ . Soit alors  $j_2$  le plus petit indice tel qu'il existe  $i \geq 2$  tel que  $a_{i,j_2} \neq 0$ ; en procédant comme plus haut, on se ramène au cas où  $a_{2,j_2} = 1$  et  $a_{i,j_2} = 0$  pour

$i \geq 3$ . En répétant ce processus, on obtient une matrice de la forme suivante :

$$(\ddagger) \quad A'' = \begin{pmatrix} 0 & 1 & * & a_{1,j_2} & * & a_{1,j_3} & \cdots & a_{1,j_r} & * & * \\ 0 & 0 & 0 & 1 & * & a_{2,j_3} & \cdots & a_{2,j_r} & * & * \\ 0 & 0 & 0 & 0 & 0 & 1 & * & a_{3,j_r} & * & * \\ \vdots & \vdots & \vdots & \vdots & \vdots & 0 & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 \end{pmatrix}$$

On a  $A'' = QA$ , où  $Q \in \text{GL}_m(k)$  est inversible, donc  $\text{Ker}(A) = \text{Ker}(A'')$ , a fortiori  $\text{rg}(A) = \text{rg}(A'') = r$ . De plus,  $A''$  donne directement les équations de  $\text{Ker}(A'')$ . En effet, considérant le système

$$A'' \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = 0$$

on voit qu'on peut choisir arbitrairement  $x_i$  pour  $i \notin \{j_1, \dots, j_r\}$ , et que chaque  $x_{j_\ell}$  s'exprime en fonction des  $x_i$  pour  $i > j_\ell$ .

D'autre part,  $\text{Im}(A'')$  est le sous-espace  $W$  de  $k^m$  engendré par les vecteurs  $f_1, \dots, f_r$ ; on en déduit que  $\text{Im}(A) = Q^{-1}(W)$ . Mais il n'est pas nécessaire d'inverser la matrice  $Q$  pour avoir les équations de  $\text{Im}(A)$ .

En effet, si l'on écrit côte à côte  $A$  et un vecteur colonne

$$Y = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} \in k^m$$

et qu'on applique à  $A$  et  $Y$  les mêmes opérations sur les lignes, on obtient le couple

$$(A'' = QA \mid Y'' = QY);$$

alors  $Y \in \text{Im}(A)$  si et seulement si  $Y'' \in \text{Im}(A'')$  et, puisque  $\text{Im}(A'') = \text{Vect}(f_1, \dots, f_r)$ , on a  $Y'' \in \text{Im}(A'')$  si et seulement si les  $(n-r)$  dernières coordonnées  $y''_{r+1}, \dots, y''_n$  de  $Y''$  sont nulles. Donc  $\text{Im}(A)$  est déterminée par les équations  $y''_{r+1} = 0, \dots, y''_n = 0$ .

Illustrons ceci sur le premier exemple considéré au paragraphe précédent, c.-à-d.,

$$A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 3 & 6 & 6 \\ 3 & 8 & 15 & 16 \end{pmatrix}.$$

Appliquons la réduction des lignes de  $A$  au couple :

$$\left( \begin{array}{cccc|c} 1 & 2 & 3 & 4 & y_1 \\ 1 & 3 & 6 & 6 & y_2 \\ 3 & 8 & 15 & 16 & y_3 \end{array} \right).$$

D'abord,  $L_2 \rightarrow L_2 - L_1$  et  $L_3 \rightarrow L_3 - 3L_1$  donnent :

$$\left( \begin{array}{cccc|c} 1 & 2 & 3 & 4 & y_1 \\ 0 & 1 & 3 & 2 & y_2 - y_1 \\ 0 & 2 & 6 & 4 & y_3 - 3y_1 \end{array} \right)$$

puis  $L_3 \rightarrow L_3 - 2L_2$  donne :

$$\left( \begin{array}{cccc|c} 1 & 2 & 3 & 4 & y_1 \\ 0 & 1 & 3 & 2 & y_2 - y_1 \\ 0 & 0 & 0 & 0 & y_3 - 3y_1 - 2(y_2 - y_1) = y_3 - y_1 - 2y_2 \end{array} \right).$$

On obtient ainsi que des équations de  $\text{Ker}(A)$  sont

$$\begin{cases} x_1 + 2x_2 + 3x_3 + 4x_4 = 0 \\ x_2 + 3x_3 + 2x_4 = 0; \end{cases}$$

ces équations expriment  $x_2$  puis  $x_1$  en fonction de  $x_3$  et  $x_4$ , qu'on peut choisir arbitrairement. En choisissant  $x_3 = 1$  et  $x_4 = 0$  (resp.  $x_3 = 0$  et  $x_4 = 1$ ), on obtient les vecteurs

$$u = \begin{pmatrix} 3 \\ -3 \\ 1 \\ 0 \end{pmatrix} \quad \text{resp.} \quad v = \begin{pmatrix} 0 \\ -2 \\ 0 \\ 1 \end{pmatrix}$$

qui forment donc une base de  $\text{Ker}(A)$ .

D'autre part, comme  $Q$  est inversible, les solutions de l'équation  $AX = Y$  sont les mêmes que celles de l'équation  $QAX = QY$ , c.-à-d.,  $A''X = Y''$ . Or, le système

$$\begin{cases} x_1 + 2x_2 + 3x_3 + 4x_4 = y_1 \\ x_2 + 3x_3 + 2x_4 = y_2 - y_1 \\ 0 = y_3 - y_1 - 2y_2 \end{cases}$$

admet des solutions si et seulement si

$$(*) \quad y_3 = y_1 + 2y_2.$$

On voit donc ainsi que  $(*)$  est une équation de  $\text{Im}(A)$ . Par exemple, prenant  $y_1 = 1$  et  $y_2 = 0$  (resp.  $y_1 = 0$  et  $y_2 = 1$ ), on obtient que les vecteurs

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

forment une base de  $\text{Im}(A)$ .

**1.2.6. Remarque importante.** — Dans l'exemple précédent, le premier pivot était en position (1,1), puis le second en position (2,2). Bien sûr, il n'en est pas toujours ainsi; en théorie on peut s'y ramener par des échanges de lignes, mais en pratique on n'a pas besoin de faire ces échanges de lignes : il suffit de se ramener à une matrice  $A'$  dont les lignes soient échelonnées « à permutation près des lignes »; alors les lignes non nulles de  $A'$  donnent des équations de  $\text{Ker}(A)$ , et les formes linéaires en face des lignes nulles de  $A'$  donnent des équations de  $\text{Im}(A)$ . Illustrons ceci dans le cas suivant :



$$\begin{pmatrix} 3 & 15 & 8 & 16 & | & y_1 \\ 1 & 6 & 3 & 6 & | & y_2 \\ 1 & 3 & 2 & 4 & | & y_3 \end{pmatrix} \xrightarrow{\substack{L_1 \rightarrow L_1 - 3L_3 \\ L_2 \rightarrow L_2 - L_3}} \begin{pmatrix} 0 & 6 & 2 & 4 & | & y_1 - 3y_3 \\ 0 & 3 & 1 & 2 & | & y_2 - y_3 \\ 1 & 3 & 2 & 4 & | & y_3 \end{pmatrix} \xrightarrow{L_1 \rightarrow L_1 - 2L_2} \begin{pmatrix} 0 & 0 & 0 & 0 & | & y_1 - 2y_2 - y_3 \\ 0 & 3 & 1 & 2 & | & y_2 - y_3 \\ 1 & 3 & 2 & 4 & | & y_3 \end{pmatrix}.$$

Donc des équations de  $\text{Ker}(A)$  sont

$$\begin{cases} 3x_2 + x_3 + 2x_4 = 0 \\ x_1 + 3x_2 + 2x_3 + 4x_4 = 0 \end{cases}$$

et une équation de  $\text{Im}(A)$  est  $y_1 - 2y_2 - y_3 = 0$ .

**1.2.7. Lien avec les systèmes linéaires.** — Ce qui précède équivaut à la théorie des systèmes linéaires

vue en L1. En effet, notant  $X$  le vecteur colonne  $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n$ , on associe à la matrice  $A \in M_{m,n}(k)$  le

système linéaire  $AX = 0$ , c.-à-d., les  $m$  équations

$$(L_i) \quad a_{i1}x_1 + \cdots + a_{in}x_n = 0$$

données par les lignes  $L_1, \dots, L_m$  de  $A$ . Chaque  $L_i$  peut aussi être vue comme une forme linéaire sur  $V = \mathbb{R}^n$ , i.e. la forme linéaire qui à tout  $v = x_1e_1 + \cdots + x_n e_n$  (où  $\mathcal{B} = (e_1, \dots, e_n)$  est la base canonique de  $\mathbb{R}^n$ ) associe le scalaire  $L_i(v) = a_{i1}x_1 + \cdots + a_{in}x_n$ .

Soit  $s$  le rang de ce système, c.-à-d., le nombre maximum de lignes de  $A$  qui sont linéairement indépendantes (donc  $s = \text{rang}(A)$ ). Le paragraphe précédent montre, en faisant des opérations élémentaires sur les lignes de  $A$ , que ce système a les mêmes solutions que le système échelonné  $A''X = 0$ , où  $A''$  est la matrice à lignes échelonnées obtenue en (§) plus haut : si  $j_1 < \cdots < j_s$  désignent les colonnes où se trouvent les pivots sur les lignes  $1, \dots, s$  de  $A''$ , on peut choisir arbitrairement  $x_i$  pour  $i \notin \{j_1, \dots, j_s\}$ , et

chaque  $x_{j_\ell}$  s'exprime en fonction des  $x_i$  pour  $i > j_\ell$ . Donc l'espace des solutions de ce système, qui n'est autre que  $\text{Ker}(A)$ , est de dimension égale à  $n - s$ . On retrouve ainsi la théorie des systèmes linéaires, vue en L1. D'autre part, comme  $\dim \text{Ker}(A) = n - \text{rg}(A)$  d'après le théorème du rang, on obtient le corollaire suivant (dont une autre démonstration a déjà été donnée en 0.5.10) :



**Corollaire 1.2.7.1.** — Pour tout  $A \in M_{m,n}(k)$ , on a  $\text{rg}({}^t A) = \text{rg}(A)$ .

Remarquons d'autre part que faire des opérations sur les lignes de  $A$  revient à faire des opérations sur les colonnes de  ${}^t A \dots$

**Remarque 1.2.7.2.** — Notons  $\mathcal{B} = (e_1, \dots, e_n)$  (resp.  $\mathcal{C} = (f_1, \dots, f_m)$ ) la base canonique de  $V = \mathbb{R}^n$  (resp. de  $W = \mathbb{R}^m$ ) et  $\mathcal{B}^* = (e_1^*, \dots, e_n^*)$  (resp.  $\mathcal{C}^* = (f_1^*, \dots, f_m^*)$ ) la base duale de  $V^*$  (resp. de  $W^*$ ). Alors les formes linéaires  $L_1, \dots, L_m$  données par les lignes de  $A$  ont les deux interprétations suivantes. Fixons un indice de ligne  $i \in \{1, \dots, m\}$ . D'une part, pour tout  $j \in \{1, \dots, n\}$ , on a  $L_i(e_j) = a_{ij}$  et donc

$$L_i = a_{i1}e_1^* + \dots + a_{in}e_n^*.$$

D'autre part,  $L_i(e_j)$  est la  $i$ -ième coordonnée dans la base  $\mathcal{C}$  du vecteur  $Ae_j$ , c.-à-d.,  $L_i(e_j) = f_i^*(Ae_j) = (f_i^* \circ A)(e_j)$ . Comme les formes linéaires  $L_i$  et  $f_i^* \circ A$  coïncident sur la base  $\mathcal{B} = (e_1, \dots, e_n)$  de  $\mathbb{R}^n$ , elles sont égales, et l'on a donc

$$(\star) \quad \forall i = 1, \dots, m, \quad f_i^* \circ A = L_i = a_{i1}e_1^* + \dots + a_{in}e_n^*.$$

Si l'on note  $u$  l'application linéaire  $V = k^n \rightarrow k^m = W$  définie par  $A$ , alors l'application  $W^* \rightarrow V^*$  qui à tout  $f \in W^*$  associe  $f \circ u \in V^*$  est linéaire : on dit que c'est la *transposée* de  $u$  et on la note  ${}^t u$ ; il résulte de  $(\star)$  ci-dessus que l'on a :

$$\text{Mat}_{\mathcal{B}^*, \mathcal{C}^*}({}^t u) = {}^t A$$

(les lecteurs intéressés pourront consulter l'appendice 1.6 à la fin de ce chapitre).

**1.2.8. Calcul de l'inverse d'une matrice carrée.** — Soit  $A \in M_n(k)$ . On peut utiliser l'algorithme de réduction des colonnes pour déterminer si  $A$  est inversible et calculer son inverse, de la façon suivante. Notons  $(e_1, \dots, e_n)$  la base canonique de  $k^n$ .

Si la première ligne de  $A$  est nulle, alors évidemment  $A$  n'est pas inversible, car alors l'image de  $A$  est contenue dans le sous-espace engendré par  $e_2, \dots, e_n$ . On peut donc supposer que la première ligne de  $A$  est non nulle; alors la première étape de l'algorithme de réduction des colonnes nous fournit une matrice  $A_1 = AP_1$  dont la première ligne est :

$$(1 \quad 0 \quad \dots \quad 0)$$

Soit  $t \in \{2, \dots, n\}$  et supposons qu'après  $t - 1$  étapes on ait obtenu une matrice

$$A_{t-1} = AP_1 \dots P_{t-1}$$

dont les  $t - 1$  premières lignes sont de la forme

$$\begin{pmatrix} 1 & 0 & 0 & 0 & \dots & 0 \\ * & 1 & 0 & 0 & \dots & 0 \\ * & * & 1 & 0 & \dots & 0 \end{pmatrix};$$

si la  $t$ -ème ligne a tous ses coefficients d'indice  $\geq t$  nuls, alors  $\text{Im}(A)$  est contenue dans le sous-espace engendré par les colonnes  $A_1, \dots, A_{t-1}$  et par les vecteurs  $e_{t+1}, \dots, e_n$  de la base canonique, donc  $A$  n'est pas inversible.

On obtient donc l'alternative suivante : ou bien au cours du processus on obtient une matrice de la forme

$$A_t = \left( \begin{array}{ccc|ccc} 1 & 0 & 0 & 0 & \dots & 0 \\ * & 1 & 0 & 0 & \dots & 0 \\ * & * & 1 & 0 & \dots & 0 \\ \hline * & * & * & 0 & \dots & 0 \\ * & * & * & * & \dots & * \\ * & * & * & * & \dots & * \end{array} \right)$$





auquel cas  $A$  n'est pas inversible (d'après ce qui précède), ou bien on arrive à une matrice triangulaire inférieure avec des 1 sur la diagonale :

$$A_n = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ * & 1 & 0 & \cdots & 0 \\ * & * & 1 & \ddots & 0 \\ \vdots & \vdots & \ddots & \ddots & 0 \\ a_{n1} & a_{n2} & \cdots & a_{n,n-1} & 1 \end{pmatrix};$$

alors soustrayant  $a_{nj}C_n$  à la  $j$ -ème colonne, pour  $j = 1, \dots, n-1$ , on obtient une matrice

$$A_{n+1} = A_n P_{n+1} = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ * & 1 & 0 & \cdots & 0 \\ * & * & 1 & \ddots & 0 \\ a_{n-1,1} & a_{n-1,2} & \cdots & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix},$$

retranchant alors  $a_{n-1,j}C_{n-1}$  à la  $j$ -ème colonne, pour  $j = 1, \dots, n-2$ , on obtient une matrice

$$A_{n+2} = A_n P_{n+1} P_{n+2} = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ * & 1 & 0 & \cdots & 0 \\ * & * & 1 & \ddots & 0 \\ 0 & \cdots & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix},$$

etc. En continuant ainsi, on arrive à une forme réduite  $A' = AP$  qui égale la matrice identité  $I_n$ ; alors, d'après la proposition 0.5.2, on a  $P = A^{-1}$ . De plus, comme indiqué dans le paragraphe 1.2.1, la matrice  $P$  est obtenue en écrivant au début du processus la matrice  $I_n$  en-dessous de  $A$ , et en effectuant à chaque étape les *mêmes* opérations élémentaires sur les colonnes des deux matrices; on obtient ainsi à la fin du processus, en haut la matrice réduite  $A' = AP = I_n$ , et en bas la matrice  $I_n P = P = A^{-1}$ .



Illustrons ceci par un exemple : soit  $A = \begin{pmatrix} 2 & -1 & 1 \\ 1 & 0 & 3 \\ 1 & 2 & 0 \end{pmatrix} \in M_3(\mathbb{Q})$ ; on a :

$$\begin{pmatrix} 2 & -1 & 1 \\ 1 & 0 & 3 \\ 1 & 2 & 0 \end{pmatrix} \xrightarrow[\begin{matrix} C_1 \rightarrow C_1/2 \\ C_2 \rightarrow C_2 + C_1/2 \\ C_3 \rightarrow C_3 - C_1/2 \end{matrix}]{\begin{matrix} C_2 \rightarrow 2C_2 \\ C_3 \rightarrow C_3 - 5C_2 \end{matrix}} \begin{pmatrix} 1 & 0 & 0 \\ 1/2 & 1/2 & 5/2 \\ 1/2 & 5/2 & -1/2 \end{pmatrix} \xrightarrow[\begin{matrix} C_1 \rightarrow C_1 + C_3/26 \\ C_2 \rightarrow C_2 + 5C_3/13 \\ C_3 \rightarrow -C_3/13 \end{matrix}]{\begin{matrix} C_1 \rightarrow C_1 - C_2/2 \end{matrix}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$



Notons que, au lieu de la suite d'opérations sur les colonnes utilisée plus haut, on peut choisir n'importe quelle suite d'opérations sur les colonnes de façon à obtenir des calculs les plus simples possibles. Ainsi, dans l'exemple considéré, il apparaît plus avantageux de procéder comme suit :

$$\begin{pmatrix} 2 & -1 & 1 \\ 1 & 0 & 3 \\ 1 & 2 & 0 \end{pmatrix} \xrightarrow[\begin{matrix} C_2 \rightarrow C_1 + 2C_2 \\ C_3 \rightarrow C_3 + C_2 \end{matrix}]{\begin{matrix} C_1 \rightarrow -C_2 \end{matrix}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 3 \\ -2 & 5 & 2 \end{pmatrix} \xrightarrow{C_3 \rightarrow C_3 - 3C_2} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -2 & 5 & -13 \end{pmatrix}.$$

$$\begin{array}{l} C_1 \rightarrow C_1 - 2C_3/13 \\ C_2 \rightarrow C_2 + 5C_3/13 \\ C_3 \rightarrow -C_3/13 \end{array} \rightarrow \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{array}{l} \\ \\ \\ \hline 6/13 & -2/13 & 3/13 \\ -3/13 & 1/13 & 5/13 \\ -2/13 & 5/13 & -1/13 \end{array}.$$



Bien sûr, on peut aussi procéder par réduction des lignes : partons de  $A \in M_n(k)$ , si la première colonne de  $A$  est nulle, alors évidemment  $A$  n'est pas inversible. On peut donc supposer que la première colonne de  $A$  est non nulle ; alors la première étape de l'algorithme de réduction des lignes nous fournit une matrice  $A_1 = Q_1 A$  dont la première colonne est :

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

Soit  $t \in \{2, \dots, n\}$  et supposons qu'après  $t - 1$  étapes on ait obtenu une matrice

$$A_{t-1} = Q_{t-1} \cdots Q_1 A$$

dont les  $t - 1$  premières colonnes sont de la forme

$$\begin{pmatrix} 1 & * & * \\ 0 & 1 & * \\ 0 & 0 & 1 \\ \vdots & \vdots & 0 \\ 0 & 0 & 0 \end{pmatrix};$$

si la  $t$ -ième colonne a tous ses coefficients d'indice  $\geq t$  nuls, alors  $\text{Im}(A)$  est contenue dans le sous-espace engendré par les vecteurs  $e_1, \dots, e_{t-1}$  de la base canonique et par les colonnes  $A_{t+1}, \dots, A_n$ , donc  $A$  n'est pas inversible.

On obtient donc l'alternative suivante : ou bien au cours du processus on obtient une matrice de la forme

$$A_t = \left( \begin{array}{ccc|ccc} 1 & * & * & * & \cdots & * \\ 0 & 1 & * & * & \cdots & * \\ 0 & 0 & 1 & * & \cdots & * \\ \hline 0 & 0 & 0 & 0 & * & * \\ \vdots & \vdots & \vdots & 0 & \vdots & * \\ 0 & 0 & 0 & 0 & * & * \end{array} \right)$$



auquel cas  $A$  n'est pas inversible (d'après ce qui précède), ou bien on arrive à une matrice triangulaire supérieure avec des 1 sur la diagonale :

$$A_n = \begin{pmatrix} 1 & * & * & \cdots & a_{1n} \\ 0 & 1 & * & \cdots & a_{2n} \\ 0 & 0 & 1 & * & * \\ \vdots & \ddots & \ddots & \ddots & a_{n-1,n} \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix};$$

alors soustrayant  $a_{in} L_n$  à la  $i$ -ième ligne, pour  $i = 1, \dots, n - 1$ , on obtient une matrice

$$A_{n+1} = Q_{n+1} A_n = \begin{pmatrix} 1 & * & * & a_{1,n-1} & 0 \\ 0 & 1 & * & a_{2,n-1} & 0 \\ 0 & 0 & 1 & \vdots & 0 \\ 0 & \ddots & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix};$$

retranchant alors  $a_{i,n-1}L_{n-1}$  à la  $i$ -ième ligne, pour  $i = 1, \dots, n-2$ , on obtient une matrice

$$A_{n+2} = Q_{n+2}Q_{n+1}A_n = \begin{pmatrix} 1 & * & * & 0 & 0 \\ 0 & 1 & * & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & \ddots & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix},$$

etc. En continuant ainsi, on arrive à une forme réduite  $A' = QA$  qui égale la matrice identité  $I_n$ ; alors, d'après la proposition 0.5.2, on a  $Q = A^{-1}$ . De plus, la matrice  $Q$  s'obtient en écrivant au début du processus les matrices  $A$  et  $I_n$  côte à côte, et en effectuant à chaque étape les *mêmes* opérations élémentaires sur les lignes des deux matrices; on obtient ainsi à la fin du processus, d'un côté la matrice réduite  $A' = QA = I_n$ , et de l'autre la matrice  $QI_n = Q = A^{-1}$ .

Illustrons ceci en reprenant l'exemple précédent :

$$\begin{aligned} & \left( \begin{array}{ccc|ccc} 2 & -1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 3 & 0 & 1 & 0 \\ 1 & 2 & 0 & 0 & 0 & 1 \end{array} \right) \xrightarrow[\begin{array}{l} L_1 \rightarrow L_1/2 \\ L_2 \rightarrow L_2 - L_1/2 \\ L_3 \rightarrow L_3 - L_1/2 \end{array}]{L_1 \rightarrow L_1/2} \left( \begin{array}{ccc|ccc} 1 & -1/2 & 1/2 & 1/2 & 0 & 0 \\ 0 & 1/2 & 5/2 & -1/2 & 1 & 0 \\ 0 & 5/2 & -1/2 & -1/2 & 0 & 1 \end{array} \right) \xrightarrow[\begin{array}{l} L_2 \rightarrow 2L_2 \\ L_3 \rightarrow L_3 - 5L_2 \end{array}]{L_2 \rightarrow 2L_2} \\ & \left( \begin{array}{ccc|ccc} 1 & -1/2 & 1/2 & 1/2 & 0 & 0 \\ 0 & 1 & 5 & -1 & 2 & 0 \\ 0 & 0 & -13 & 2 & -5 & 1 \end{array} \right) \xrightarrow[\begin{array}{l} L_1 \rightarrow L_1 + L_3/26 \\ L_2 \rightarrow L_2 + 5L_3/13 \\ L_3 \rightarrow -L_3/13 \end{array}]{L_1 \rightarrow L_1 + L_3/26} \left( \begin{array}{ccc|ccc} 1 & -1/2 & 0 & 15/26 & -5/26 & 1/26 \\ 0 & 1 & 0 & -3/13 & 1/13 & 5/13 \\ 0 & 0 & 1 & -2/13 & 5/13 & -1/13 \end{array} \right) \\ & \xrightarrow{L_1 \rightarrow L_1 + L_2/2} \left( \begin{array}{ccc|ccc} 1 & 0 & 0 & 6/13 & -2/13 & 3/13 \\ 0 & 1 & 0 & -3/13 & 1/13 & 5/13 \\ 0 & 0 & 1 & -2/13 & 5/13 & -1/13 \end{array} \right). \end{aligned}$$

Remarquons à nouveau que, au lieu de la suite d'opérations sur les lignes utilisée plus haut, on peut choisir n'importe quelle suite d'opérations sur les lignes de façon à obtenir des calculs les plus simples possibles. Ainsi, dans cet exemple, il apparaît plus avantageux de procéder comme suit :

$$\begin{aligned} & \left( \begin{array}{ccc|ccc} 2 & -1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 3 & 0 & 1 & 0 \\ 1 & 2 & 0 & 0 & 0 & 1 \end{array} \right) \xrightarrow[\begin{array}{l} L_2 \rightarrow L_1 - 2L_2 \\ L_3 \rightarrow L_3 - L_2 \end{array}]{L_1 \rightarrow L_2} \left( \begin{array}{ccc|ccc} 1 & 0 & 3 & 0 & 1 & 0 \\ 0 & -1 & -5 & 1 & -2 & 0 \\ 0 & 2 & -3 & 0 & -1 & 1 \end{array} \right) \xrightarrow[\begin{array}{l} L_2 \rightarrow -L_2 \\ L_3 \rightarrow L_3 + 2L_2 \end{array}]{L_2 \rightarrow -L_2} \\ & \left( \begin{array}{ccc|ccc} 1 & 0 & 3 & 0 & 1 & 0 \\ 0 & 1 & 5 & -1 & 2 & 0 \\ 0 & 0 & -13 & 2 & -5 & 1 \end{array} \right) \xrightarrow[\begin{array}{l} L_2 \rightarrow L_2 + 5L_3/13 \\ L_3 \rightarrow -L_3/13 \end{array}]{L_1 \rightarrow L_1 + 3L_3/13} \left( \begin{array}{ccc|ccc} 1 & 0 & 0 & 6/13 & -2/13 & 3/13 \\ 0 & 1 & 0 & -3/13 & 1/13 & 5/13 \\ 0 & 0 & 1 & -2/13 & 5/13 & -1/13 \end{array} \right). \end{aligned}$$

**Remarque 1.2.8.1.** — Attention, dans cet algorithme pour calculer  $A^{-1}$  il faut choisir de faire des opérations sur les colonnes *ou bien* sur les lignes, mais il ne faut *pas mélanger les deux*!

En effet, si l'on fait à la fois des opérations sur les lignes et sur les colonnes de  $A$ , et qu'on fait les mêmes opérations sur la matrice  $I_n$ , on arrive au bout du processus à une paire de matrices :

$$(QAP = I_n, \quad QI_nP = QP);$$

alors la première égalité donne  $A = Q^{-1}P^{-1}$  d'où  $A^{-1} = PQ$ , mais de l'autre côté c'est  $QP$  que l'on a calculé...

### 1.3. Codimension et équations d'un sous-espace

**Définition 1.3.1 (Codimension).** — Soient  $V$  un  $k$ -espace vectoriel de dimension  $n$  et  $E$  un sous-espace vectoriel de  $V$ , on appelle *codimension* de  $E$  dans  $V$  et l'on note  $\text{codim}_V(E)$  l'entier  $\dim V - \dim E$ .

Soit  $\mathcal{B} = (e_1, \dots, e_n)$  une base de  $V$ . Notons  $(x_1, \dots, x_n)$  les « fonctions coordonnées » par rapport à  $\mathcal{B}$ , i.e. tout  $v \in V$  s'écrit de façon unique

$$v = x_1(v)e_1 + \dots + x_n(v)e_n.$$

Si  $\mathcal{B}^* = (e_1^*, \dots, e_n^*)$  est la base duale de  $V^*$ , alors pour tout  $v \in V$  on a  $e_j^*(v) = x_j(v)$ , donc la fonction  $x_j : V \rightarrow k, v \mapsto x_j(v)$  n'est autre que la forme linéaire  $e_j^*$ . Considérons maintenant des « équations linéaires »

$$(L_i) \quad a_{i1}x_1 + \dots + a_{in}x_n = 0$$

pour  $i = 1, \dots, m$  et soit  $E$  le sous-espace de  $V$  défini par ces équations, i.e.  $E$  est l'ensemble des vecteurs  $v = x_1e_1 + \dots + x_n e_n$  dont les coordonnées  $(x_1, \dots, x_n)$  vérifient les équations précédentes. Alors  $E$  est le

noyau de la matrice  $A \in M_{m,n}(k)$  dont les lignes sont  $L_1, \dots, L_m$ , donc  $\dim E = n - s$  où  $s$  est le rang du système, i.e. le nombre maximum d'équations linéairement indépendantes.

Comme dans le paragraphe 1.2.7, ces équations peuvent aussi être vues comme des formes linéaires sur  $V$ , i.e. chaque  $L_i$  est la forme linéaire  $a_{i1}e_1^* + \dots + a_{in}e_n^*$ ; alors  $E$  est l'ensemble des  $v \in V$  tels que  $L_i(v) = 0$  pour  $i = 1, \dots, m$ . Quitte à renuméroter les  $L_i$ , on peut supposer que  $L_1, \dots, L_s$  sont linéairement indépendantes; alors pour  $i > s$  chaque  $L_i$  est combinaison linéaire de  $L_1, \dots, L_s$  donc l'égalité  $L_i(v) = 0$  est conséquence des égalités  $L_t(v) = 0$  pour  $t = 1, \dots, s$ , et donc  $E$  est le sous-espace de  $V$  défini par les équations  $L_t(v) = 0$  pour  $t = 1, \dots, s$ , et  $\dim E = n - s$ . Pour résumer, on a obtenu la



**Proposition 1.3.2.** — Soit  $V$  un espace vectoriel de dimension  $n$  et soient  $L_1, \dots, L_m \in V^*$ . Le sous-espace  $E$  de  $V$  défini par ces équations (i.e.  $E = \{v \in V \mid L_t(v) = 0, \forall t = 1, \dots, m\}$ ) est de dimension  $n - s$ , où  $s = \dim \text{Vect}(L_1, \dots, L_m)$ . En particulier, lorsque  $L_1, \dots, L_m$  sont linéairement indépendantes, on a  $\dim E = n - m$ .

**Remarque 1.3.3.** — Remarquons que  $E$  ne dépend que du sous-espace  $F = \text{Vect}(L_1, \dots, L_m)$  de  $V^*$  engendré par les  $L_t$ . En effet, d'une part, tout  $\phi \in F$  s'annule sur  $E$  (car  $\phi$  est combinaison linéaire de  $L_1, \dots, L_m$ ). D'autre part, si l'on pose

$$F^\circ = \{v \in V \mid \phi(v) = 0, \quad \forall \phi \in F\}$$

on a  $E \subseteq F^\circ$  d'après ce qui précède, et  $F^\circ \subseteq \{v \in V \mid L_t(v) = 0, \forall t = 1, \dots, m\} = E$ , d'où  $E = F^\circ$ . On appelle  $F^\circ$  l'orthogonal de  $F$  dans  $V$ ; alors la proposition précédente peut se récrire en disant que  $\dim F^\circ = n - \dim F$ .

Réciproquement, si  $E$  est un sous-espace vectoriel de  $V$  de dimension  $r$ , il peut être défini par  $n - r$  équations linéairement indépendantes. En effet, soit  $(e_1, \dots, e_r)$  une base de  $E$ ; complétons-la en une base  $\mathcal{B} = (e_1, \dots, e_n)$  de  $V$  et soit  $\mathcal{B}^* = (e_1^*, \dots, e_n^*)$  la base duale de  $V^*$ . Alors  $E$  est défini par les formes linéaires  $e_{r+1}^*, \dots, e_n^*$ , c.-à-d., par les équations  $x_i = 0$  pour  $i = r + 1, \dots, n$ . De plus, une forme linéaire  $\phi = t_1 e_1^* + \dots + t_n e_n^*$  s'annule sur  $E$  si et seulement si  $t_1 = 0 = \dots = t_r$ , donc, si l'on définit l'orthogonal  $E^\perp$  de  $E$  dans  $V^*$  par :

$$E^\perp = \{\phi \in V^* \mid \phi(x) = 0 \text{ pour tout } x \in E\},$$

alors  $E^\perp$  égale  $\text{Vect}(e_{r+1}^*, \dots, e_n^*)$  donc est de dimension  $n - r$ . Donc, si  $L_1, \dots, L_m$  sont des formes linéaires qui définissent  $E$ , c.-à-d., telles que

$$E = \{v \in V \mid L_i(v) = 0, \quad \forall i = 1, \dots, m\}$$

alors  $L_1, \dots, L_m \in E^\perp$ , et d'après la proposition précédente, on a :

$$n - \dim \text{Vect}(L_1, \dots, L_m) = \dim E = r$$

et donc  $\dim \text{Vect}(L_1, \dots, L_m) = n - r = \dim E^\perp$ , donc  $(L_1, \dots, L_m)$  est une famille génératrice de  $E^\perp$  et l'on peut en extraire un système de  $n - r$  équations qui définissent  $E$ . Pour résumer, on a obtenu la



**Proposition 1.3.4.** — Soient  $V$  un espace vectoriel de dimension  $n$  et  $E$  un sous-espace vectoriel de dimension  $r$ . Alors  $E$  peut être défini par  $n - r$  équations linéairement indépendantes; de plus de tout système d'équations  $L_1, \dots, L_m$  définissant  $E$ , on peut extraire  $n - r$  équations linéairement indépendantes définissant  $E$  (en particulier, on a  $m \geq n - r$ ).

**Exemples 1.3.5.** — 1) Dans  $\mathbb{R}^2$ , la droite  $D$  engendrée par le vecteur  $\begin{pmatrix} a \\ b \end{pmatrix} \neq 0$  est de codimension  $2 - 1 = 1$ . L'orthogonal  $D^\perp \subseteq (\mathbb{R}^2)^*$  est engendré par la forme linéaire  $\phi = be_1^* - ae_2^* \neq 0$ , i.e.  $D$  est définie par l'équation  $bx - ay = 0$ . Cette équation est unique à un scalaire près, i.e. toute autre équation de  $D$  est de la forme  $\lambda(bx - ay) = 0$ , avec  $\lambda \in \mathbb{R}^*$ .



2) Dans  $\mathbb{R}^3$ , le plan  $P$  engendré par les vecteurs  $e_1 - e_2$  et  $e_2 - e_3$  est de codimension  $3 - 2 = 1$ . Son orthogonal  $P^\perp \subseteq (\mathbb{R}^3)^*$  est engendré par la forme linéaire  $\phi = e_1^* + e_2^* + e_3^* \neq 0$ , i.e.  $P$  est défini par l'équation  $x + y + z = 0$ . Cette équation est unique à un scalaire près, i.e. toute autre équation de  $P$  est de la forme  $\lambda(x + y + z) = 0$ , avec  $\lambda \in \mathbb{R}^*$ .

3) Dans  $\mathbb{R}^3$ , la droite  $D$  engendrée par le vecteur  $ae_1 + be_2 - e_3$ , avec  $(a, b) \neq (0, 0)$ , est de codimension  $3 - 1 = 2$ , donc son orthogonal  $D^\perp \subseteq (\mathbb{R}^3)^*$  est de dimension 2, il admet pour base, par exemple,  $(e_1^* + ae_3^*, e_2^* + be_3^*)$ , i.e.  $D$  est définie par les équations  $x + az = 0$  et  $y + bz = 0$ . Une autre base de  $D^\perp$  est, par exemple,  $(be_1^* - ae_2^*, e_1^* + e_2^* + (a + b)e_3^*)$ , donc  $D$  est aussi définie par les équations  $bx - ay = 0$  et  $x + y + (a + b)z = 0$ .

**Remarques 1.3.6.** — Soient  $V$  un espace vectoriel de dimension  $n$  et  $V^*$  son dual.

(1) On voit facilement que pour tout sous-espace  $E$  de  $V$  (resp.  $F$  de  $V^*$ ), on a  $E \subseteq (E^\perp)^\circ$  (resp.  $F \subseteq (F^\circ)^\perp$ ), et il résulte des deux propositions précédentes que  $\dim(E^\perp)^\circ = n - \dim E^\perp = \dim E$  et  $\dim(F^\circ)^\perp = n - \dim F^\circ = \dim F$ . On a donc les égalités :  $E = (E^\perp)^\circ$  et  $F = (F^\circ)^\perp$ .

(2) On a introduit temporairement la notation  $F^\circ$  pour l'orthogonal de  $F$  dans  $V$ , pour le distinguer a priori de son orthogonal  $F^\perp$  dans le dual  $V^{**} = (V^*)^*$  de  $V^*$ . Mais en fait on a un isomorphisme canonique  $V = V^{**}$ , i.e.  $V$  est le dual de son dual (voir l'appendice 1.7), et les deux notions d'orthogonalité coïncident, donc pour un sous-espace  $F$  de  $V^*$  on peut aussi noter  $F^\perp$  son orthogonal dans  $V$ ...

#### 1.4. Déterminants

Soit  $k$  un corps, on rappelle que le déterminant d'une matrice  $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in M_2(k)$  est  $\det(A) = ad - bc$ ; on sait que  $A$  est inversible si et seulement si  $\det(A) \neq 0$ . En effet, on a

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} = \begin{pmatrix} ad - bc & 0 \\ 0 & ad - bc \end{pmatrix} = \det(A) \cdot I_2;$$

ceci montre que si  $\det(A) \neq 0$ , alors  $\frac{1}{\det(A)} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$  est l'inverse de  $A$ .

Réciproquement, supposons  $A$  inversible, alors le système

$$(S) \quad \begin{cases} ax + by = 0 \\ cx + dy = 0 \end{cases}$$

a  $(0, 0)$  comme unique solution. En particulier,  $a, b$  ne sont pas simultanément nuls. Si  $a \neq 0$  (resp. si  $b \neq 0$ ), (S) est équivalent au système obtenu en remplaçant la ligne  $L_2$  par  $aL_2 - cL_1$  (resp. par  $bL_2 - dL_1$ ) :

$$\begin{cases} ax + by = 0 \\ (ad - bc)y = 0 \end{cases} \quad \text{resp.} \quad \begin{cases} ax + by = 0 \\ (bc - ad)x = 0; \end{cases}$$

on en déduit que  $ad - bc = 0$  (sinon, l'espace des solutions serait la droite d'équation  $ax + by = 0$ ).

D'autre part, soit  $X$  une indéterminée; on définit le « polynôme caractéristique » de  $A$  comme

$$P_A(X) = X^2 - (a + d)X + (ad - bc)$$

c.-à-d., c'est le « déterminant » de la matrice  $\begin{pmatrix} a - X & b \\ c & d - X \end{pmatrix}$  à coefficients dans l'anneau de polynômes

$k[X]$ . Ceci conduit à définir, pour tout anneau commutatif  $R$  le déterminant d'une matrice  $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in M_2(R)$  comme  $\det(A) = ad - bc$ . Notons que le déterminant  $2 \times 2$  ainsi défini vérifie les trois propriétés suivantes :

(1) C'est une fonction  $R$ -linéaire de chacune des colonnes  $\begin{pmatrix} a \\ c \end{pmatrix}$  et  $\begin{pmatrix} b \\ d \end{pmatrix}$  de la matrice  $A$ , c.-à-d., pour tous  $t, a, b, c, d \in R$ , on a :

$$\det \begin{pmatrix} ta + a' & b \\ tc + c' & d \end{pmatrix} = (ta + a')d - b(tc + c') = t(ad - bc) + a'd - bc' = t \det \begin{pmatrix} a & b \\ c & d \end{pmatrix} + \det \begin{pmatrix} a' & b \\ c' & d \end{pmatrix}$$

et

$$\det \begin{pmatrix} a & tb + b' \\ c & td + d' \end{pmatrix} = a(td + d') - (tb + b')c = t(ad - bc) + ad' - b'c = t \det \begin{pmatrix} a & b \\ c & d \end{pmatrix} + \det \begin{pmatrix} a & b' \\ c & d' \end{pmatrix}$$

(2) Lorsque les deux colonnes sont égales, on a  $\det(A) = 0$ .

(3)  $\det(I_2) = 1$ .

On va voir que ces conditions s'étendent de façon naturelle pour les matrices  $n \times n$ , c.-à-d., on a le théorème suivant, où  $k$  désigne un anneau commutatif arbitraire. (Dans la suite, on s'intéressera uniquement au cas où  $k$  est un corps ou un anneau de polynômes sur un corps, mais ça ne coûte pas plus cher de le faire pour un anneau commutatif arbitraire, par exemple  $\mathbb{Z}$ , ou  $\mathbb{Z}/n\mathbb{Z}$ , ou  $\mathbb{Z}[X]$ , etc.)

**Théorème 1.4.1 (Existence et propriétés du déterminant).** — Soit  $k$  un anneau commutatif et soit  $n \in \mathbb{N}^*$ .

(a) Il existe une unique fonction  $\det : M_n(k) \rightarrow k$  vérifiant les trois propriétés suivantes :



(1) C'est une fonction  $k$ -linéaire de chacune des colonnes  $A_1, \dots, A_n$  de la matrice  $A$ , c.-à-d., pour tout  $i = 1, \dots, n$ , si  $A'_i \in M_{n,1}(k)$  est une autre matrice colonne à coefficients dans  $k$  et si  $t \in k$ , on a :

$$\det(A_1, \dots, tA_i + A'_i, \dots, A_n) = t \det(A_1, \dots, A_i, \dots, A_n) + \det(A_1, \dots, A'_i, \dots, A_n).$$

(2) Si deux colonnes sont égales, i.e. s'il existe  $i \neq j$  tels que  $A_i = A_j$ , alors  $\det(A) = 0$ .

(3)  $\det(I_n) = 1$ .

Plus précisément, pour toute fonction  $D : M_n(k) \rightarrow k$  vérifiant les propriétés (1) et (2), on a  $D = D(I_n) \cdot \det$ .

(b) Pour tout  $A, B \in M_n(k)$  on a

(\*)  $\det(BA) = \det(B) \cdot \det(A)$

(\*\*)  $\det({}^t A) = \det(A)$ .

(c) Enfin, il existe une matrice  $\tilde{A}$  (appelée la matrice des cofacteurs de  $A$ ) telle que

(\*\*\*)  $A \cdot {}^t \tilde{A} = \det(A) I_n = {}^t \tilde{A} \cdot A$ .

Par conséquent  $A$  est inversible si et seulement si  $\det(A)$  est un élément inversible de  $k$ ; dans ce cas, on a  $\det(A^{-1}) = \det(A)^{-1}$ .

*Démonstration.* — On va montrer d'abord que si une fonction  $D : M_n(k) \rightarrow k$  vérifie les propriétés (1) et (2), elle est entièrement déterminée par le scalaire  $D(I_n)$ . Ceci prouvera l'unicité, et permettra ensuite de construire par récurrence une fonction  $\det : M_n(k) \rightarrow k$  vérifiant toutes les propriétés ci-dessus.

Soit donc  $D : M_n(k) \rightarrow k$  vérifiant les propriétés (1) et (2). Notons d'abord que ces conditions entraînent les conditions (2') et (2<sup>-</sup>) qui suivent.

a) Pour  $i \neq j$  dans  $\{1, \dots, n\}$ ,  $D$  ne change pas si l'on ajoute à la colonne  $A_j$  un multiple  $tA_i$  de la colonne  $A_i$ , c.-à-d., on a :

(2')  $D(A_1, \dots, A_i, \dots, A_j + tA_i, \dots, A_n) = D(A_1, \dots, A_i, \dots, A_j, \dots, A_n)$ .

Par conséquent, si une colonne, disons  $A_j$ , est combinaison  $k$ -linéaire des autres colonnes, c.-à-d., s'il existe des  $t_i \in k$ , pour  $i \neq j$ , tels que  $A_j = \sum_{i \neq j} t_i A_i$  alors  $D(A) = 0$ . En effet, d'après (1),  $D(A)$  est la somme pour  $i \neq j$  des termes

$$D(A_1, \dots, A_i, \dots, t_i A_i, \dots, A_n)$$

(où  $t_i A_i$  est à la  $j$ -ème place), et d'après (2') chacun de ces termes est nul, d'où  $D(A) = 0$ .

b) Si l'on échange les colonnes  $i$  et  $j$ , la valeur de  $D$  est multipliée par  $-1$ , c.-à-d.,

(2<sup>-</sup>)  $D(A_1, \dots, A_j, \dots, A_i, \dots, A_n) = -D(A_1, \dots, A_i, \dots, A_j, \dots, A_n)$

En effet, si l'on place  $A_i + A_j$  dans les colonnes  $i$  et  $j$ , alors les conditions (2) et (1) entraînent que

$$0 = D(\dots, A_i + A_j, \dots, A_i + A_j, \dots) = D(\dots, A_i, \dots, A_i, \dots) + D(\dots, A_j, \dots, A_j, \dots) \\ + D(\dots, A_i, \dots, A_j, \dots) + D(\dots, A_j, \dots, A_i, \dots)$$

or dans le dernier membre les deux premiers termes sont nuls, par (2) à nouveau, d'où

$$D(\dots, A_i, \dots, A_j, \dots) = -D(\dots, A_j, \dots, A_i, \dots),$$

ce qui prouve (2<sup>-</sup>).

**Unicité.** Montrons maintenant que les conditions (1) et (2) permettent de calculer  $D$  en fonction du scalaire  $\lambda = D(I_n)$ . On procède par récurrence sur  $n$ . Si  $n = 1$ , alors  $M_1(k) = \{(a) \mid a \in k\}$  et  $I_1 = (1)$ , donc toute application linéaire  $D : M_1(k) \rightarrow k$  est de la forme  $(a) \mapsto \lambda a$ , où  $\lambda = D(I_1)$ . Soit donc  $n \geq 2$  et supposons avoir établi qu'il existe une application

$$\det_{n-1} : M_{n-1}(k) \rightarrow k$$

vérifiant les conditions (1), (2), (3), et que pour toute application  $\psi : M_{n-1}(k) \rightarrow k$  vérifiant les conditions (1) et (2), on a

$$\psi = \psi(I_{n-1}) \cdot \det_{n-1}$$

(ceci implique, en particulier, que  $\det_{n-1}$  est uniquement déterminé).

Soit  $A = (a_{ij})$  un élément arbitraire de  $M_n(k)$ . Considérons les matrices colonnes suivantes :

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \quad e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \quad \dots \quad e_n = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}$$

alors la première colonne  $A_1$  de  $A$  s'écrit  $A_1 = a_{11}e_1 + \dots + a_{n1}e_n$  donc, d'après (1), on a :

$$(\dagger) \quad D(A) = \sum_{i=1}^n a_{i1} D(A'(i, 1)),$$

où l'on désigne par  $A'(i, 1)$  la matrice dont les colonnes sont  $e_i, A_2, \dots, A_n$ , i.e.

$$A'(i, 1) = \begin{pmatrix} 0 & a_{12} & \dots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & a_{i2} & \dots & a_{in} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & a_{n2} & \dots & a_{nn} \end{pmatrix}.$$

Soit alors  $A''(i, 1)$  la matrice dont les colonnes sont :  $e_i, A_2 - a_{i2}e_i, \dots, A_n - a_{in}e_i$ , c.-à-d.,

$$A''(i, 1) = \begin{pmatrix} 0 & a_{12} & \dots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & a_{i-1,2} & \dots & a_{i-1,n} \\ 1 & 0 & \dots & 0 \\ 0 & a_{i+1,2} & \dots & a_{i+1,n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & a_{n2} & \dots & a_{nn} \end{pmatrix}.$$

On a  $D(A''(i, 1)) = D(A'(i, 1))$  d'après (2'), et donc (†') devient :

$$(\dagger'') \quad D(A) = \sum_{i=1}^n a_{i1} D(A''(i, 1)).$$

Pour  $i = 1, \dots, n$ , notons  $\phi_i$  la fonction  $M_{n-1}(k) \rightarrow k$  qui à tout  $B \in M_{n-1}(k)$  associe  $D(\widetilde{B}(i, 1))$ , où  $\widetilde{B}(i, 1)$  désigne la matrice :

$$\begin{pmatrix} 0 & b_{11} & \dots & b_{1,n-1} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & b_{i-1,1} & \dots & b_{i-1,n-1} \\ 1 & 0 & \dots & 0 \\ 0 & b_{i,1} & \dots & b_{i,n-1} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & b_{n-1,1} & \dots & b_{n-1,n-1} \end{pmatrix}.$$

Alors,  $\phi_i$  vérifie les conditions (1) et (2). De plus, on a

$$\widetilde{I_{n-1}}(i, 1) = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 & \dots & 0 \\ \vdots & 0 & \ddots & \ddots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 & 0 & \dots & 0 \\ 1 & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & 0 & \ddots & \vdots \\ 0 & 0 & \dots & 0 & 0 & \dots & 1 \end{pmatrix}$$

et cette matrice est déduite de la matrice identité  $I_n$  en faisant glisser la  $i$ -ème colonne à la première place, c.-à-d., en faisant  $i - 1$  échanges de colonnes, d'où

$$\phi_i(I_{n-1}) = D(\widetilde{I_{n-1}}(i, 1)) = (-1)^{i-1} D(I_n) = (-1)^{i+1} D(I_n)$$

donc, d'après l'hypothèse de récurrence, on obtient :

$$(\ddagger_i) \quad \phi_i = D(I_n)(-1)^{i+1} \det_{n-1}.$$

Notons  $A(i, 1) = A - L_i - C_1$  l'élément de  $M_{n-1}(k)$  obtenu à partir de  $A$  en supprimant la  $i$ -ème ligne et la première colonne, alors d'après  $(\ddagger_i)$  on a

$$D(A''(i, 1)) = D(I_n)(-1)^{i+1} \det_{n-1}(A - L_i - C_1)$$

et donc  $(\dagger'')$  donne :

$$(\star^1) \quad D(A) = D(I_n) \sum_{i=1}^n (-1)^{i+1} a_{i1} \det_{n-1}(A - L_i - C_1).$$

Donc, si  $D$  vérifie (1) et (2), elle est déterminé par le scalaire  $D(I_n) \in k$ , d'après la formule  $(\star^1)$  ci-dessus.

De plus, soit  $j$  un indice de colonne arbitraire, faisant glisser la colonne  $A_j$  à la première place, on obtient

$$D(A) = (-1)^{j-1} D(A_j, A_1, \dots, A_{j-1}, A_{j+1}, \dots, A_n),$$

puis appliquant  $(\star^1)$  à la matrice  $(A_j, A_1, \dots, A_{j-1}, A_{j+1}, \dots, A_n)$ , on obtient :

$$(\star^j) \quad D(A) = D(I_n) \sum_{i=1}^n (-1)^{i+j} a_{ij} \det_{n-1}(A - L_i - C_j),$$

ce qui montre que si  $D$  vérifie (1) et (2), alors elle vérifie les égalités  $(\star^j)$  pour tout  $j = 1, \dots, n$ .

**Remarques 1.4.2.** — 1) Pour ceux qui ont déjà rencontré les déterminants, on reconnaît dans  $(\star^j)$  le développement d'un déterminant selon la  $j$ -ème colonne.

2) En répétant le calcul précédent pour  $\det_{n-1}$ , puis  $\det_{n-2}$ , etc., on obtiendrait que si  $D$  vérifie (1) et (2), elle est *nécessairement donnée* par une certaine formule

$$D(A) = D(I_n) \sum_{\sigma \in S_n} \varepsilon(\sigma) a_{1\sigma(1)} \cdots a_{n\sigma(n)},$$

où  $S_n$  désigne le groupe des permutations (c.-à-d., bijections) de l'ensemble  $\{1, \dots, n\}$ , et où  $\varepsilon(\sigma)$  est un signe  $\pm 1$ , explicitement déterminé en fonction de  $\sigma$ . On pourrait alors montrer l'existence de  $\det_n$  en montrant que la fonction

$$A \mapsto \sum_{\sigma \in S_n} \varepsilon(\sigma) a_{1\sigma(1)} \cdots a_{n\sigma(n)}$$

vérifie (1), (2) et (3). Toutefois, il est plus simple de montrer l'existence de  $\det_n$  comme suit.

**Existence.** Fixons un indice de ligne  $i$  arbitraire, et pour  $j = 1, \dots, n$  notons

$$\Delta_{ij}(A) = \det_{n-1}(A - L_i - C_j),$$

où  $A - L_i - C_j$  désigne l'élément de  $M_{n-1}(k)$  obtenu à partir de  $A$  en supprimant la  $i$ -ème ligne et la  $j$ -ème colonne. Considérons la fonction  $D_i : M_n(k) \rightarrow k$  définie par

$$(\star_i) \quad D_i(A) = \sum_{j=1}^n (-1)^{j+i} a_{ij} \Delta_{ij}(A),$$

et montrons que  $D_i$  vérifie les conditions (1), (2) et (3).

Fixons un indice de colonne  $\ell$ . Alors les fonctions  $A \mapsto a_{i\ell}$  et  $A \mapsto \Delta_{ij}$  pour  $j \neq \ell$  sont des fonctions linéaires de  $C_\ell$ , tandis que les fonctions  $A \mapsto \Delta_{i\ell}$  et  $A \mapsto a_{ij}$  pour  $j \neq \ell$  ne dépendent pas de  $C_\ell$ , donc la somme dans  $(\star_i)$  est bien une fonction linéaire de chaque colonne  $C_\ell$ , i.e. (1) est vérifiée.

Lorsque  $A = I_n$ , la matrice  $I_n - L_i - C_j$  a sa  $i$ -ème colonne nulle si  $j \neq i$ , et égale  $I_{n-1}$  si  $j = i$ , donc la somme dans  $(\star_i)$  égale  $\det_{n-1}(I_{n-1}) = 1$ , donc (3) est vérifiée.

Enfin, supposons qu'il existe  $p < q$  tels que les colonnes  $C_p$  et  $C_q$  de  $A$  soient égales. Alors, d'une part on a  $\Delta_{ij}(A) = 0$  pour  $j \neq p, q$ . D'autre part, les matrices  $A - L_i - C_p$  et  $A - L_i - C_q$  se déduisent l'une de l'autre par  $q - 1 - p$  échanges de colonnes, car la colonne  $C = C_p = C_q$  est à la place  $p$  dans  $A - L_i - C_q$  et à la place  $q - 1$  dans  $A - L_i - C_p$ . Donc

$$\Delta_{ip}(A) = (-1)^{q-1-p} \Delta_{iq}(A) \quad \text{i.e.} \quad \Delta_{iq}(A) = (-1)^{p+1-q} \Delta_{ip}(A).$$

Posant  $\alpha = a_{ip} = a_{iq}$ , l'égalité  $(\star_i)$  donne alors :

$$D_i(A) = \alpha \Delta_{ip}(A) ((-1)^{i+p} + (-1)^{i+q+p+1-q}) = (-1)^{i+p} \alpha \Delta_{ip}(A) (1 - 1) = 0,$$



donc (2) est vérifiée. Compte-tenu de ce qui précède, ceci prouve que  $\det_n(A) = D_i(A)$  ne dépend pas de  $i$  et est l'unique application  $M_n(k) \rightarrow k$  vérifiant (1), (2) et (3); de plus toute application  $D : M_n(k) \rightarrow k$  vérifiant (1) et (2) égale  $D(I_n) \cdot \det_n$ .

On a donc démontré le point (a) du théorème 1.4.1, et l'on a obtenu au passage les formules de développement suivant une ligne  $(\star_i)$  ou suivant une colonne  $(\star^j)$ .

Prouvons le point (b). Fixons  $B \in M_n(k)$ . Alors, pour tout  $A \in M_n(k)$ , les colonnes de la matrice  $BA$  sont  $BA_1, \dots, BA_n$ , où  $A_1, \dots, A_n$  désignent les colonnes de  $A$  (cf. 0.4.6). Par conséquent, comme chaque application  $A_i \mapsto BA_i$  est linéaire, l'application

$$\phi_B : A \mapsto \det(BA) = \det(BA_1, \dots, BA_n)$$

vérifie (1) et (2), donc est égale à  $\phi_B(I_n) \cdot \det$ ; comme  $\phi_B(I_n) = \det(B)$  on obtient

$$(*) \quad \det(BA) = \det(B) \cdot \det(A).$$

Montrons que  $\det({}^t A) = \det(A)$ , en procédant par récurrence sur  $n$ . Il n'y a rien à montrer si  $n = 1$ , donc on peut supposer  $n \geq 2$  et le résultat établi pour  $n - 1$ . Soient  $A \in M_n(k)$ . D'après  $(\star_1)$  appliqué à  ${}^t A$ , on a

$$\det({}^t A) = \sum_{j=1}^n (-1)^{j+1} ({}^t A)_{1j} \Delta_{1j}({}^t A).$$

Or,  $({}^t A)_{1j} = a_{j1}$  et

$$\begin{aligned} \Delta_{1j}({}^t A) &= \det_{n-1} \left( {}^t A - L_1({}^t A) - C_j({}^t A) \right) = \det_{n-1} \left( {}^t (A - C_1(A) - L_j(A)) \right) \\ &= \det_{n-1} \left( A - C_1(A) - L_j(A) \right) = \Delta_{j1}(A) \end{aligned}$$

(l'avant-dernière égalité d'après l'hypothèse de récurrence). On obtient donc

$$\det({}^t A) = \sum_{j=1}^n (-1)^{j+1} a_{j1} \Delta_{j1}(A) = \det(A),$$

la seconde égalité étant  $(\star^1)$ . On a donc prouvé le point (b) du théorème 1.4.1.



**Définition 1.4.3 (Matrice des cofacteurs).** — Pour  $i, j \in \{1, \dots, n\}$ ,  $(-1)^{i+j} \tilde{\Delta}_{ij}(A)$  est appelé le *cofacteur* de  $A$  d'indice  $(i, j)$ . On appelle *matrice des cofacteurs* de  $A$  la matrice  $\tilde{A}$  dont le coefficient d'indice  $(i, j)$  est  $\tilde{A}_{ij} = (-1)^{i+j} \tilde{\Delta}_{ij}(A)$ .

Démontrons maintenant le point (c) du théorème 1.4.1. Pour tout  $i, \ell \in \{1, \dots, n\}$ , on a :

$$(A {}^t \tilde{A})_{i\ell} = \sum_{j=1}^n a_{ij} ({}^t \tilde{A})_{j\ell} = \sum_{j=1}^n (-1)^{j+\ell} a_{ij} \Delta_{\ell j}(A)$$

et l'on reconnaît là le développement suivant la ligne  $\ell$  de la matrice  $B(\ell, i)$  déduite de  $A$  en remplaçant la ligne d'indice  $\ell$  par celle d'indice  $i$ . Donc  $(A {}^t \tilde{A})_{i\ell} = 0$  si  $\ell \neq i$  (car  $B(\ell, i)$  a alors deux lignes égales), et  $(A {}^t \tilde{A})_{ii} = \det(A)$  si  $\ell = i$ . Ceci montre que

$$A {}^t \tilde{A} = \det(A) \cdot I_n.$$

De même,

$$({}^t \tilde{A} A)_{i\ell} = \sum_{j=1}^n (-1)^{i+j} \tilde{\Delta}_{ji}(A) a_{j\ell}$$

et l'on reconnaît là le développement suivant la colonne  $i$  de la matrice  $B'(i, \ell)$  déduite de  $A$  en remplaçant la colonne d'indice  $i$  par celle d'indice  $\ell$ . Donc, à nouveau,  $({}^t \tilde{A} A)_{i\ell} = 0$  si  $\ell \neq i$ , et  $= \det(A)$  si  $\ell = i$ , d'où

$${}^t \tilde{A} A = \det(A) \cdot I_n.$$

On a ainsi montré les égalités (\*\*\*) de 1.4.1; de plus, ceci montre que si  $\det(A)$  est inversible dans  $k$  (i.e. s'il existe  $\alpha \in k$  tel que  $\alpha \cdot \det(A) = 1$ ), alors  $\det(A)^{-1} {}^t \tilde{A}$  est l'inverse de  $A$ . Réciproquement, si  $A$  est inversible, l'égalité  $AA^{-1} = I_n$  et la *multiplicativité du déterminant* (\*) entraînent  $\det(A) \det(A^{-1}) = 1$ , donc  $\det(A)$  est inversible dans  $k$ , son inverse étant

$$\det(A^{-1}) = \det(A)^{-1}.$$

Ceci achève la preuve du théorème 1.4.1. □



**Corollaire 1.4.4.** — Soient  $k$  un anneau commutatif,  $A, P \in M_n(k)$  avec  $P$  inversible, alors

$$\boxed{\det(P^{-1}AP) = \det(P^{-1}) \det(A) \det(P) = \det(A).}$$

**Remarque 1.4.5 (Forme explicite du déterminant).** — Soit  $S_n$  le groupe des permutations (c.-à-d., bijections) de  $\{1, \dots, n\}$ . En procédant par récurrence sur  $n$ , on déduit de la formule de développement suivant une ligne (ou bien une colonne) que, pour toute matrice  $A = (a_{ij})_{1 \leq i, j \leq n}$  on a une formule explicite :

$$\det(A) = \sum_{\sigma \in S_n} \varepsilon(\sigma) a_{1\sigma(1)} \cdots a_{n\sigma(n)}, \quad \text{où } \varepsilon(\sigma) = \pm 1$$

c.-à-d., c'est la somme, avec certains signes  $+$  ou  $-$ , de tous les produits de  $n$  coefficients de  $A$ , en ne prenant dans chaque produit qu'un seul coefficient par ligne et par colonne ; de plus le terme « diagonal »  $a_{11}a_{22} \cdots a_{nn}$  (qui correspond à  $\sigma = \text{id}$ ) est précédé du signe «  $+$  ». (On verra plus loin comment calculer explicitement le signe  $\varepsilon(\sigma)$  ; par exemple si  $n = 3$  on a  $\det(A) = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{32}a_{21} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31} - a_{11}a_{23}a_{32}$ .)



**Proposition 1.4.6 (Déterminant de matrices triangulaires).** — 1) Si  $A$  est une matrice triangulaire de termes diagonaux  $\lambda_1, \dots, \lambda_n$ , alors

$$\det(A) = \lambda_1 \cdots \lambda_n.$$

2) Plus généralement, si  $A$  est une matrice triangulaire par blocs, c.-à-d., de la forme

$$A = \begin{pmatrix} A_1 & * & \cdots & * \\ 0 & A_2 & \ddots & * \\ 0 & \ddots & \ddots & * \\ 0 & \cdots & 0 & A_n \end{pmatrix}$$

(où les  $*$  désignent des coefficients arbitraires), alors

$$\det(A) = \det(A_1) \cdots \det(A_n).$$

*Démonstration.* — 1) Supposons que  $A$  soit une matrice triangulaire supérieure :

$$A = \begin{pmatrix} \lambda_1 & * & \cdots & * \\ 0 & \lambda_2 & \ddots & \vdots \\ 0 & \ddots & \ddots & * \\ 0 & \cdots & 0 & \lambda_n \end{pmatrix}$$

alors en développant par rapport à la première colonne, on obtient que

$$\det(A) = \lambda_1 \det \begin{pmatrix} \lambda_2 & * & * \\ 0 & \ddots & * \\ 0 & 0 & \lambda_n \end{pmatrix};$$

le résultat en découle, en répétant l'opération ou en procédant par récurrence sur  $n$ .

2) En procédant par récurrence sur  $n$ , il suffit de montrer que si

$$A = \begin{pmatrix} B & C \\ 0 & D \end{pmatrix}$$

où  $B$  (resp.  $D$ ) est une matrice carrée de taille  $p$  (resp.  $q$ ),  $C$  est une matrice à  $p$  lignes et  $q$  colonnes, et  $0$  désigne la matrice nulle à  $q$  lignes et  $p$  colonnes, alors

$$\det(A) = \det(B) \cdot \det(D).$$

Fixons  $C, D$  et considérons l'application  $\phi$  qui à une matrice  $B' \in M_p(k)$  associe

$$\phi(B') = \det \begin{pmatrix} B' & C \\ 0 & D \end{pmatrix}.$$

Alors on voit aussitôt que  $\phi$  vérifie les propriétés (1) et (2), donc d'après le théorème 1.4.1, on a

$$\phi(B') = \phi(I_p) \cdot \det(B') = \det \begin{pmatrix} I_p & C \\ 0 & D \end{pmatrix} \cdot \det(B');$$

de plus, en développant suivant la première colonne, puis suivant la seconde, etc. jusqu'à la  $p$ -ème colonne, on obtient que

$$\det \begin{pmatrix} I_p & C \\ 0 & D \end{pmatrix} = \det D,$$

d'où finalement  $\det(A) = \phi(B) = \det(B) \cdot \det(D)$ . □

**Remarque 1.4.6.1.** — Attention ! Si on décompose une matrice  $M$  sous la forme

$$M = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$$

où  $A, B, C, D$  sont toutes des matrices carrées (nécessairement de même taille  $p$ , de sorte que  $M$  est de taille  $2p$ ), il n'est pas vrai en général que  $\det(M)$  égale  $\det(A)\det(D) - \det(B)\det(C)$ . Par exemple si

$$M = \left( \begin{array}{cc|cc} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \hline 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{array} \right)$$

alors  $\det(A) = 0 = \det(B) = \det(C) = \det(D)$  mais  $\det(M) = -1$ .

**1.4.7. Cas d'un corps.** — Supposons dans ce paragraphe que l'anneau de base  $k$  soit un corps. Alors, en plus des propriétés (1), (2) et (2'), (2<sup>-</sup>) de 1.4.1, le déterminant vérifie la propriété suivante :

(2'') Si les colonnes  $A_1, \dots, A_n$  sont liées, alors  $\det(A_1, \dots, A_n) = 0$ .

En effet, si les colonnes  $A_1, \dots, A_n$  sont liées, il existe  $t_1, \dots, t_n \in k$  non tous nuls tels que  $t_1 A_1 + \dots + t_n A_n = 0$ . Soit  $j$  tel que  $t_j \neq 0$ , alors

$$A_j = - \sum_{i \neq j} t_i t_j^{-1} A_i.$$

Donc, d'après (1),  $\det(A)$  est la somme pour  $i \neq j$  des termes

$$-t_i t_j^{-1} \det(\dots, A_i, \dots, A_i, \dots)$$

(où  $A_i$  apparaît aux places  $i$  et  $j$ ), et d'après (2) chacun de ces termes est nul, d'où  $\det(A) = 0$ . Ceci prouve (2'').

D'autre part, et c'est le plus important, pour calculer pratiquement un déterminant, on ne procède pas véritablement en développant suivant les lignes ou les colonnes (ce qui serait trop fastidieux, et coûteux en temps de calcul), mais on essaie de faire des opérations élémentaires sur les lignes ou les colonnes, pour rendre la matrice triangulaire.



C'est-à-dire, soit  $A = (a_{ij})$  un élément arbitraire de  $M_n(k)$ . Si la première ligne de  $A$  est nulle, alors  $\det(A) = 0$  (en développant suivant la première ligne, ou bien en utilisant  $\det(A) = \det({}^t A)$ ), et il n'y a rien à calculer. Supposons donc qu'il existe au moins une colonne  $C_j$  telle que  $a_{1j} \neq 0$ . Faisons glisser la colonne  $C_j$  au-dessus de  $C_1, \dots, C_{j-1}$  pour l'amener à la première place; ceci introduit le signe  $(-1)^{j-1}$ , puis mettons  $a_{1j}$  en facteur, ceci donne :

$$\det(A) = (-1)^{j-1} a_{1j} \det(a_{1j}^{-1} C_j, C_1, \dots, C_{j-1}, C_{j+1}, \dots, C_n).$$

Soustrayons alors  $a_{1\ell} a_{1j}^{-1} C_j$  de  $C_\ell$ , pour  $\ell \neq j$ , pour annuler les coefficients de la première ligne autres que le premier; ceci ne change pas la valeur du déterminant et l'on obtient donc que

$$(\dagger) \quad \det(A) = (-1)^{j-1} a_{1j} \det \begin{pmatrix} 1 & 0 \\ v & A' \end{pmatrix} = (-1)^{j-1} a_{1j} \det(A')$$

où  $A'$  est la matrice carrée de taille  $n-1$  formée des lignes d'indice 2 à  $n$  des vecteurs  $C_\ell - a_{1\ell} a_{1j}^{-1} C_j$ , pour  $\ell \in \{1, \dots, n\} - \{j\}$  (ces vecteurs ont tous leur première ligne nulle). On répète ensuite l'opération pour  $A'$ , etc. Illustrons ceci par un :



**Exemple 1.4.7.1.** — Calculons le déterminant suivant (où  $n \in \mathbb{Z}$ ) :

$$D = \begin{vmatrix} 2 & 3 & 4 & 5 \\ 7 & 6 & 9 & 8 \\ 10 & 11 & 12 & 13 \\ 14 & 15 & 16 & n \end{vmatrix}$$

Mettant  $a_{11} = 2$  en facteur dans la première colonne et remplaçant  $C_2, C_3$  et  $C_4$  par  $C_2 - (3/2)C_1, C_3 - 2C_1$  et  $C_4 - (5/2)C_1$  on obtient :

$$D = 2 \begin{vmatrix} 1 & 0 & 0 & 0 \\ 7/2 & -9/2 & -5 & -19/2 \\ 5 & -4 & -8 & -12 \\ 7 & -6 & -12 & n-35 \end{vmatrix} = 2 \begin{vmatrix} -9/2 & -5 & -19/2 \\ -4 & -8 & -12 \\ -6 & -12 & n-35 \end{vmatrix}$$

puis remplaçant chaque colonne par son opposé (ce qui introduit un signe  $(-1)^3$ ), mettant 4 en facteur dans la 2ème ligne, puis échangeant  $L_1$  et  $L_2$  (ce qui introduit un signe  $-1$ ), on obtient :

$$D = 2 \cdot 4 \cdot (-1)^4 \begin{vmatrix} 1 & 2 & 3 \\ 9/2 & 5 & 19/2 \\ 6 & 12 & 35-n \end{vmatrix} = 8 \begin{vmatrix} 1 & 2 & 3 \\ 9/2 & 5 & 19/2 \\ 6 & 12 & 35-n \end{vmatrix}$$

puis remplaçant  $C_2$  et  $C_3$  par  $C_2 - 2C_1$  et  $C_3 - 3C_1$  on obtient :

$$D = 8 \begin{vmatrix} 1 & 0 & 0 \\ 9/2 & -4 & -4 \\ 6 & 0 & 17-n \end{vmatrix} = 8 \begin{vmatrix} -4 & -4 \\ 0 & 17-n \end{vmatrix} = 32(n-17).$$

### 1.5. Endomorphismes : déterminant, trace, valeurs propres, etc.

Soit  $V$  un espace vectoriel de dimension  $n$ . En raison de son importance, répétons encore ici le théorème de changement de base pour un endomorphisme  $u$  de  $V$  (étant entendu qu'on exprime la matrice de  $u$  dans une même base au départ et à l'arrivée) (cf. 0.5.13) :

#### **Théorème 1.5.1 (Changement de base pour un endomorphisme)**

Soit  $A$  la matrice d'un endomorphisme  $u$  de  $V$  relativement à une base  $\mathcal{B}$  de  $V$ . Si  $\mathcal{B}'$  est une seconde base, et si  $P$  est la matrice de passage de  $\mathcal{B}$  à  $\mathcal{B}'$ , alors la matrice de  $u$  dans la base  $\mathcal{B}'$  est :

$$\text{Mat}_{\mathcal{B}'}(u) = P^{-1}AP.$$

Ceci permet de définir le déterminant et le polynôme caractéristique de  $u$  comme suit.

#### **Théorème et définition 1.5.2 (Déterminant, polynôme caractéristique et trace d'un endomorphisme)**

Soit  $u \in \text{End}_k(V)$  et soit  $A$  sa matrice dans une base  $\mathcal{B}$  de  $V$ . On définit le déterminant et le polynôme caractéristique de  $u$  par :

$$\det(u) = \det(A), \quad P_u(X) = \det(A - XI_n) \in k[X];$$

ceci ne dépend pas du choix de la base  $\mathcal{B}$ . De plus,  $P_u(X)$  est de degré  $n = \dim_k(V)$  et est de la forme

$$P_u(X) = (-1)^n X^n + (-1)^{n-1} \text{Tr}(u) X^{n-1} + \dots + \det(u).$$

Le coefficient  $\text{Tr}(u)$  s'appelle la trace de  $u$ , il est égal à la somme des coefficients diagonaux  $a_{ii}$  de  $A$  (et ceci ne dépend pas de la base choisie).

*Démonstration.* — En effet, soient  $\mathcal{B}'$  une autre base,  $P$  la matrice de passage, et  $A'$  la matrice de  $u$  dans la base  $\mathcal{B}'$ . Alors on a  $A' = P^{-1}AP$  et donc aussi

$$A' - XI_n = P^{-1}(A - XI_n)P$$

(égalité dans l'anneau  $M_n(k[X])$  des matrices à coefficients dans  $k[X]$ ). Donc, d'après la multiplicativité du déterminant (cf. 1.4.4), on a

$$\det(A') = \det(A), \quad \det(A' - XI_n) = \det(A - XI_n)$$

ce qui prouve que  $\det(u)$  et  $P_u(X)$  sont bien définis. De plus, notons  $b_{ij}$  les coefficients de la matrice  $A - XI_n$ , i.e.  $b_{ij} = a_{ij}$  si  $i \neq j$  et  $b_{ii} = a_{ii} - X$ . D'après la formule 1.4.5, on a

$$P_u(X) = \sum_{\sigma \in S_n} \varepsilon(\sigma) b_{1\sigma(1)} \cdots b_{n\sigma(n)}$$

c.-à-d., c'est la somme, avec certains signes  $+$  ou  $-$ , de tous les produits de  $n$  coefficients de  $B = A - XI_n$ , en ne prenant dans chaque produit qu'un seul coefficient dans chaque ligne et chaque colonne, et de plus on a  $\varepsilon(\text{id}) = 1$ , i.e. le terme  $b_{11}b_{22} \cdots b_{nn}$  apparaît avec le signe  $+$ .

Donc, le terme de degré maximal en  $X$  est  $(-X)^n$ , qui s'obtient en développant le produit des termes diagonaux :

$$b_{11}b_{22} \cdots b_{nn} = (a_{11} - X) \cdots (a_{nn} - X).$$



Pour avoir un terme en  $X^{n-1}$ , il faut prendre  $n-1$  fois  $(-X)$  sur la diagonale, mais alors, comme chaque produit de  $n$  coefficients n'a qu'un seul coefficient par ligne et par colonne, le dernier terme du produit est le coefficient diagonal restant, dans lequel on prend le terme  $a_{ii}$ . Ceci montre que le coefficient de  $(-X)^{n-1}$  est

$$a_{11} + \cdots + a_{nn},$$

qu'on appelle la trace de  $A$ . Comme  $P_u(X)$  ne dépend pas de la base choisie, ce coefficient n'en dépend pas non plus ; on l'appelle la trace de  $u$  et on le note  $\text{Tr}(u)$ .  $\square$

**Remarque 1.5.3.** — On peut aussi montrer directement (cf. Feuille d'exercices n°1) que, pour tout  $A, B \in M_n(k)$ , on a  $\text{Tr}(AB) = \text{Tr}(BA)$ , d'où  $\text{Tr}(P^{-1}AP) = \text{Tr}(APP^{-1}) = \text{Tr}(A)$ .

D'après le théorème 1.4.1, on peut alors compléter la proposition 0.3.3 comme suit :



**Proposition 1.5.4.** — Soit  $u \in \text{End}_k(V)$  ( $\dim_k(V) = n$ ). Les conditions suivantes sont équivalentes :

- (i)  $u$  est injectif ;
- (ii)  $u$  est surjectif ;
- (iii)  $u$  est bijectif ;
- (iv)  $\det(u) \neq 0$ .

**Définition 1.5.5 (Valeurs, vecteurs et sous-espaces propres).** — 1) Soit  $u \in \text{End}_k(V)$ . On dit que  $\lambda \in k$  est une *valeur propre* de  $u$  si  $\text{Ker}(u - \lambda \text{id}_V) \neq \{0\}$  (i.e. s'il existe  $v \in V - \{0\}$  tel que  $u(v) = \lambda v$ ). Dans ce cas,  $V_\lambda = \text{Ker}(u - \lambda \text{id}_V)$  est appelé *l'espace propre* associé à  $\lambda$ , et tout vecteur  $v \in V_\lambda - \{0\}$  est appelé un *vecteur propre* de  $u$ , associé à la valeur propre  $\lambda$ .

2) Pour  $A \in M_n(k)$ , on définit ses valeurs, vecteurs et sous-espaces propres comme étant ceux de l'endomorphisme  $u$  de  $k^n$  défini par  $A$ , c.-à-d.,  $\lambda$  est valeur propre de  $A$  si et seulement si  $V_\lambda = \text{Ker}(A - \lambda I_n)$  est non nul.



**Proposition 1.5.6.** — Soient  $u \in \text{End}_k(V)$  et  $\lambda \in k$ , alors  $\lambda$  est une valeur propre de  $u \Leftrightarrow P_u(\lambda) = 0$ .

*Démonstration.* — D'après la proposition 1.5.4,  $\lambda$  est valeur propre de  $u$  si et seulement si  $\det(u - \lambda \text{id}_V) = 0$ . D'autre part, d'après la formule explicite exprimant le déterminant d'une matrice  $A$  en fonction des coefficients  $a_{ij}$  de  $A$  (cf. Remarque 1.4.5), on voit que :

$$\det(u - \lambda \text{id}_V) = P_u(\lambda)$$

c.-à-d., calculer le polynôme  $P_u(X) = \det(A - XI_n)$ , où  $A = \text{Mat}_{\mathcal{B}}(u)$ , puis l'évaluer en  $X = \lambda$  « est la même chose » que de calculer  $\det(A - \lambda I_n)$ . Ceci démontre la proposition.  $\square$

Signalons aussi le lemme utile suivant :



**Lemme 1.5.7.** — Soient  $u \in \text{End}_k(V)$  et  $v \in V$  un vecteur propre pour une valeur propre  $\lambda \in k$ . Alors, pour tout  $Q \in k[X]$  on a  $Q(u)(v) = Q(\lambda)v$ .

*Démonstration.* — En effet, on a  $u(v) = \lambda v$ , donc  $u^2(v) = u(u(v)) = u(\lambda v) = \lambda u(v) = \lambda^2 v$ , et l'on montre par récurrence sur  $n$  que  $u^n(v) = \lambda^n v$  pour tout  $n \in \mathbb{N}$ . Alors, pour tout polynôme  $Q = a_0 + a_1 X + \cdots + a_d X^d$ , on a :

$$Q(u)(v) = (a_0 \text{id}_V + a_1 u + \cdots + a_d u^d)(v) = a_0 v + a_1 u(v) + \cdots + a_d u^d(v) = a_0 v + a_1 \lambda v + \cdots + a_d \lambda^d v = Q(\lambda)v.$$

$\square$

## 1.6. Appendice (†) : Transposée d'une application linéaire

Soit  $A = (a_{ij}) \in M_{m,n}(k)$  et soit  $u$  l'endomorphisme  $k^n \rightarrow k^m$  associé à  $A$ . Notant  $\mathcal{B} = (e_1, \dots, e_n)$  (resp.  $\mathcal{C} = (f_1, \dots, f_m)$ ) la base canonique de  $k^n$  (resp.  $k^m$ ) et  $\mathcal{B}^* = (e_1^*, \dots, e_n^*)$  (resp.  $\mathcal{C}^* = (f_1^*, \dots, f_m^*)$ ) la base duale,  $u$  est défini par  $u(e_j) = \sum_{i=1}^m a_{ij} f_i$  pour  $j = 1, \dots, n$ , donc aussi par :

$$f_i^*(u(e_j)) = a_{ij} \quad \forall j = 1, \dots, n, \quad \forall i = 1, \dots, m.$$

Donc, pour  $i = 1, \dots, m$ , la forme linéaire  $f_i^* \circ u : k^n \rightarrow k$  s'écrit :

$$(\dagger) \quad f_i^* \circ u = \sum_{j=1}^n a_{ij} e_j^*.$$

Considérons alors l'application  $\theta : (k^m)^* \rightarrow (k^n)^*$  qui à toute forme linéaire  $\phi : k^m \rightarrow k$  associe la forme linéaire  $\phi \circ u : k^n \rightarrow k$ . Alors  $\theta$  est linéaire, car

$$\theta(t \cdot \phi + \psi) = (t \cdot \phi + \psi) \circ u = t \cdot \phi \circ u + \psi \circ u = t \cdot \theta(\phi) + \theta(\psi),$$

et (†) montre que  $\theta(f_i^*) = \sum_{j=1}^n a_{ij} e_j^*$ , d'où :  $\text{Mat}_{\mathcal{B}^*, \mathcal{C}^*}(\theta) = {}^t A$ . Ceci conduit à la définition et au théorème suivants.

**Définition 1.6.1 (Transposée d'une application linéaire).** — Soient  $E, F$  deux  $k$ -espaces vectoriels et  $u : E \rightarrow F$  une application linéaire. On appelle *transposée de  $u$*  et l'on note  ${}^t u$  l'application linéaire  $F^* \rightarrow E^*$  définie par  ${}^t u(\phi) = \phi \circ u$ , pour tout  $\phi \in F^*$ . On peut visualiser ceci par le diagramme :

$$\begin{array}{ccc} E & \xrightarrow{u} & F \\ & \searrow \text{---} & \downarrow \phi \\ & & k. \end{array}$$

${}^t u(\phi) = \phi \circ u$

Remarquons que si  $v$  est une application linéaire  $F \rightarrow G$ , alors  $\boxed{{}^t(v \circ u) = {}^t u \circ {}^t v}$ . En effet, pour tout  $\psi \in G^*$  on a

$${}^t(v \circ u)(\psi) = \psi \circ v \circ u = {}^t u(\psi \circ v) = {}^t u({}^t v(\psi)) = ({}^t u \circ {}^t v)(\psi).$$

**Lemme 1.6.2.** — Soit  $u : E \rightarrow F$  une application linéaire. Alors  $\text{Ker}({}^t u) = \text{Im}(u)^\perp = \{\phi \in F^* \mid \phi(u(E)) = 0\}$ .

*Démonstration.* On a  $\text{Ker}({}^t u) = \{\phi \in F^* \mid \phi \circ u = 0\} = \{\phi \in F^* \mid \phi(u(E)) = 0\} = \text{Im}(u)^\perp$ .

**Théorème 1.6.3.** — Soient  $E, F$  deux  $k$ -espaces vectoriels de dimension finie,  $u : E \rightarrow F$  une application linéaire,  $\mathcal{B}$  (resp.  $\mathcal{C}$ ) une base de  $E$  (resp. de  $F$ ) et  $A = \text{Mat}_{\mathcal{C}, \mathcal{B}}(u)$ .

(i) On a  $\text{Mat}_{\mathcal{B}^*, \mathcal{C}^*}({}^t u) = {}^t A$ .

(ii) On a :  $\text{rang}({}^t u) = \text{rang}(u)^\perp$  et  $\text{rang}(A) = \text{rang}({}^t A)$ . Par conséquent,  $\text{rang}(A)$  est aussi le nombre maximum de lignes de  $A$  linéairement indépendantes.

*Démonstration.* On a déjà vu le point (i) au début de ce paragraphe. D'après le lemme, on a  $\text{Ker}({}^t u) = \text{dim Im}(u)^\perp$  donc, d'après le théorème du rang, appliqué à  ${}^t u$ , et la proposition 1.3.4, on a

$$n - \text{rg}({}^t u) = \text{dim Ker}({}^t u) = \text{dim Im}(u)^\perp = n - \text{rg}(u),$$

d'où  $\text{rg}({}^t u) = \text{rg}(u)$ . Le point (ii) en découle.

## 1.7. Appendice (†) : Bidual

Soient  $V$  un  $k$ -espace vectoriel et  $V^*$  son dual. Puisque  $V^*$  est un  $k$ -espace vectoriel, on peut aussi former son dual  $(V^*)^*$ , qu'on note  $V^{**}$  et qu'on appelle le *bidual* de  $V$ . Remarquons que tout vecteur  $v \in V$  définit une forme linéaire  $\varepsilon_v$  sur  $V^*$ , l'application « d'évaluation en  $v$  » :

$$\varepsilon_v : V^* \rightarrow k, \quad f \mapsto f(v).$$

On a bien  $\varepsilon_v(t \cdot f + f') = (t \cdot f + f')(v) = t \cdot f(v) + f'(v) = t \cdot \varepsilon_v(f) + \varepsilon_v(f')$ , donc  $\varepsilon_v$  est un élément du bidual  $V^{**}$ .

De plus, l'application  $v \mapsto \varepsilon_v$  est linéaire. En effet, pour tout  $f \in V^*$ , on a

$$\varepsilon_{t \cdot v + v'}(f) = f(t \cdot v + v') = t \cdot f(v) + f(v') = t \cdot \varepsilon_v(f) + \varepsilon_{v'}(f),$$

ce qui montre que  $\varepsilon_{t \cdot v + v'} = t \cdot \varepsilon_v + \varepsilon_{v'}$ . On obtient donc une application linéaire « canonique » (c.-à-d., qui ne dépend d'aucun choix) :

$$\varepsilon : V \rightarrow V^{**}, \quad v \mapsto \varepsilon_v.$$

**Proposition 1.7.1.** — Supposons  $V$  de dimension finie  $n$ . Alors l'application canonique  $\varepsilon : V \rightarrow V^{**}$  est un isomorphisme, donc permet d'identifier canoniquement  $V$  à son bidual  $V^{**}$ .

*Démonstration.* D'abord, on a  $\text{dim}_k(V^{**}) = \text{dim}_k(V^*) = \text{dim}_k(V)$ . D'autre part, soit  $v_1 \in V$  non nul, on peut compléter la famille libre  $\{v_1\}$  en une base  $(v_1, \dots, v_n)$  de  $V$ , et si  $v_1^*$  est la forme linéaire définie par  $v_1^*(v_1) = 1$  et  $v_1^*(v_i) = 0$  pour  $i = 2, \dots, n$ , alors  $\varepsilon_{v_1}(v_1^*) = 1$ , donc  $\varepsilon_{v_1} \neq 0$ . Ceci montre que  $\varepsilon$  est injective ; comme  $\text{dim}_k(V^{**}) = \text{dim}_k(V)$  alors  $\varepsilon$  est un isomorphisme.

**Remarque 1.7.1.1.** — Attention ! Si  $V$  n'est pas de dimension finie, on peut montrer que  $\varepsilon : V \rightarrow V^{**}$  est injective mais pas surjective. (C'est facile dès qu'on dispose de l'existence de bases de  $V$  et de  $V^*$ ). (On peut même montrer (mais c'est plus difficile) qu'il n'existe aucun isomorphisme de  $V$  sur  $V^{**}$ .)

**Remarque 1.7.2.** — Soient  $V$  un  $k$ -espace vectoriel de dimension  $n$ . L'isomorphisme  $\varepsilon : V \xrightarrow{\sim} V^{**}$  permet, d'une part, de démontrer le point (2) de la proposition 1.1.4 comme suit. Soit  $\Delta = (f_1, \dots, f_n)$  une base de  $V^*$ , et soit  $\Delta^* = (f_1^*, \dots, f_n^*)$  la base duale de  $V^{**}$ . Via l'isomorphisme  $\varepsilon : V \xrightarrow{\sim} V^{**}$ ,  $\Delta^*$  est l'image d'une base  $\mathcal{B} = (e_1, \dots, e_n)$  de  $V$ ; alors, pour tous  $i, j$ , on a :

$$f_i(e_j) = \varepsilon_{e_j}(f_i) = f_j^*(f_i) = \begin{cases} 1 & \text{si } i = j, \\ 0 & \text{sinon,} \end{cases}$$

et donc  $\Delta = (f_1, \dots, f_n)$  est la base duale de  $\mathcal{B} = (e_1, \dots, e_n)$ .

D'autre part, soient  $F$  un sous-espace vectoriel de  $V^*$ , et  $F^\perp$  son orthogonal dans  $V^{**}$ . Alors, via l'isomorphisme  $\varepsilon : V \xrightarrow{\sim} V^{**}$ ,  $F^\perp$  s'identifie à

$$\{x \in V \mid \forall f \in F, \quad 0 = \varepsilon_x(f) = f(x)\} = F^\circ$$

(cf. la remarque 1.3.6).






## CHAPITRE 2

### RAPPELS SUR LES ENDOMORPHISMES DIAGONALISABLES, TRIGONALISATION, CAYLEY-HAMILTON, ESPACES CARACTÉRISTIQUES

**Résumé :** Dans ce chapitre, on commence par des rappels sur les espaces propres et des critères de diagonalisabilité (2.1.8, 2.1.11 et 2.1.12). Puis, prenant  $\mathbb{C}$  comme corps de base, on démontre les théorèmes de trigonalisation, de Cayley-Hamilton, et de décomposition en espaces caractéristiques. Il faut essayer de bien assimiler ces notions nouvelles.

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 (on y démontre le théorème de Bézout et le fait que le corps  $\mathbb{C}$  est algébriquement clos). Ces passages n'interviendront pas dans les évaluations, toutefois on mentionnera en cours le fait que  $k[X]$  est principal, la définition du polynôme minimal, et le théorème de Bézout.

#### 2.1. Espaces propres et critères de diagonalisabilité

**Définition 2.1.1 (Sous-espaces en somme directe).** — Soient  $V$  un  $k$ -espace vectoriel,  $E_1, \dots, E_n$  des sous-espaces de  $V$ . (Ni  $V$  ni les  $E_i$  ne sont supposés de dimension finie.)

(1) D'abord, on note  $E_1 + \dots + E_n$  (ou  $\sum_{i=1}^n E_i$ ) le sous-espace de  $V$  engendré par  $E_1 \cup \dots \cup E_n$ ; c'est l'ensemble de toutes les sommes

$$(*) \quad x_1 + \dots + x_n, \quad \text{avec } x_i \in E_i.$$

(2) On dit que les  $E_i$  sont en somme directe si pour tous  $x_1 \in E_1, \dots, x_n \in E_n$ , l'égalité  $x_1 + \dots + x_n = 0$  entraîne  $x_1 = 0 = \dots = x_n$ . Ceci équivaut à dire que tout élément  $x$  de  $E_1 + \dots + E_n$  s'écrit de façon unique  $x = x_1 + \dots + x_n$  avec  $x_i \in E_i$ . Dans ce cas,  $E_1 + \dots + E_n$  est noté  $E_1 \oplus \dots \oplus E_n$  ou  $\bigoplus_{i=1}^n E_i$ .

(3) Si chaque  $E_i$  est de dimension finie  $d_i$ , et si  $\mathcal{B}_1 = (e_1, \dots, e_{d_1})$  est une base de  $E_1$ , et  $\mathcal{B}_2 = (e_{d_1+1}, \dots, e_{d_1+d_2})$  une base de  $E_2$ , ... puis  $\mathcal{B}_n = (e_{d_1+\dots+d_{n-1}+1}, \dots, e_{d_1+\dots+d_n})$  une base de  $E_n$ , la condition précédente équivaut à dire que la famille  $\mathcal{F} = (e_1, \dots, e_{d_1+\dots+d_n})$  est une base de  $E_1 + \dots + E_n$ , et comme  $\mathcal{F}$  engendre de toute façon  $E_1 + \dots + E_n$ , ceci équivaut aussi à dire que

$$(*) \quad \dim(E_1 + \dots + E_n) = \dim(E_1) + \dots + \dim(E_n).$$

**Terminologie 2.1.2.** — Si  $E_1, \dots, E_n$  sont en somme directe et si de plus  $E_1 \oplus \dots \oplus E_n$  égale  $V$ , alors on dit que  $V$  est la somme directe des  $E_i$ .

**Remarques 2.1.3.** — (1) Il résulte de la définition que  $E_1, \dots, E_n$  sont en somme directe si et seulement si, pour tout  $i = 1, \dots, n$ , on a :  $E_i \cap \sum_{j \neq i} E_j = \{0\}$ .

(2) En particulier, si  $n = 2$ , alors  $E_1$  et  $E_2$  sont en somme directe si et seulement si  $E_1 \cap E_2 = \{0\}$ .

(3) **Attention !** Si des sous-espaces sont en somme directe, leur somme n'est pas nécessairement égale à l'espace tout entier : par exemple si  $E_1, E_2$  sont deux droites distinctes dans  $\mathbb{R}^3$ , leur somme est directe, et c'est un plan de  $\mathbb{R}^3$ , et non  $\mathbb{R}^3$  tout entier !

(4) **Attention !** Si  $n \geq 3$ , la condition  $E_i \cap E_j = \{0\}$  pour  $i \neq j$  n'entraîne pas que la somme des  $E_i$  soit directe : par exemple si  $E_1, E_2, E_3$  sont trois droites distinctes dans  $\mathbb{R}^2$ , elles vérifient  $E_i \cap E_j = \{0\}$  pour  $i \neq j$ , mais leur somme n'est pas directe (car  $E_1 + E_2$  égale  $\mathbb{R}^2$  donc contient  $E_3$ ).

**Définition 2.1.4 (Sous-espaces supplémentaires).** — Soient  $V$  un espace vectoriel,  $E, F$  deux sous-espaces de  $V$ . On dit que  $E$  et  $F$  sont des sous-espaces *supplémentaires* si  $V = E \oplus F$ , c.-à-d., si  $E \cap F = (0)$  et  $E + F = V$ .

Si  $V$  est de dimension finie, ceci équivaut à dire que  $E \cap F = (0)$  et  $\dim(E) + \dim(F) = \dim(V)$ .



**Proposition 2.1.5.** — Soit  $V$  un  $k$ -espace vectoriel de dimension finie  $n$ . Tout sous-espace  $E$  de  $V$  admet un supplémentaire.

*Démonstration.* Soit  $(e_1, \dots, e_r)$  une base de  $E$ , complétons-la en une base  $(e_1, \dots, e_n)$  de  $V$  et soit  $F$  le sous-espace de  $V$  engendré par  $e_{r+1}, \dots, e_n$ . Alors  $E \cap F = \{0\}$  et  $E + F = V$ , donc  $V = E \oplus F$ .

**Exercice 2.1.6.** — Soient  $V$  un espace vectoriel,  $E$  (resp.  $F$ ) un sous-espace de dimension finie  $m$  (resp.  $n$ ). Soit  $(v_1, \dots, v_r)$  une base de  $E \cap F$ , complétons-la en une base  $(e_1, \dots, e_{m-r}, v_1, \dots, v_r)$  de  $E$  (resp.  $(v_1, \dots, v_r, f_1, \dots, f_{n-r})$  de  $F$ ). Montrer que  $(e_1, \dots, e_{m-r}, v_1, \dots, v_r, f_1, \dots, f_{n-r})$  est une base de  $E + F$ . En déduire l'égalité  $\dim(E + F) = \dim(E) + \dim(F) - \dim(E \cap F)$ .

**Remarque 2.1.7.** — Soient  $V$  un  $k$ -espace vectoriel et  $f$  une forme linéaire sur  $V$  (cf. 1.1.1). Supposons  $f \neq 0$  et notons  $H = \text{Ker}(f)$ . Comme  $f \neq 0$ , il existe  $v \in V$  tel que  $f(v) \neq 0$ , et remplaçant  $v$  par  $f(v)^{-1}v$ , on peut supposer  $f(v) = 1$ . Alors, pour tout  $w \in V$ , on a  $f(w - f(w)v) = 0$ , donc  $w - f(w)v \in \text{Ker}(f) = H$  et donc  $w$  s'écrit

$$w = h + f(w)v \quad \text{avec} \quad h = w - f(w)v \in H,$$

d'où  $V = H + kv$ . D'autre part, si  $tv \in H$  alors  $0 = f(tv) = t$ ; on a donc  $H \cap kv = (0)$  et donc  $V = H \oplus kv$ , i.e.  $H$  et la droite  $kv$  sont supplémentaires. On dit que  $H$  est un *hyperplan* de  $V$ ; si  $V$  est de dimension finie  $n$ , alors  $H$  est de dimension  $n - 1$ .

Un exemple très important de sous-espaces en somme directe est celui des sous-espaces propres d'un endomorphisme de  $V$ ; rappelons-le ici.



**Théorème 2.1.8.** — Soient  $V$  un  $k$ -espace vectoriel (pas nécessairement de dimension finie),  $u$  un  $k$ -endomorphisme de  $V$ , et  $\lambda_1, \dots, \lambda_r$  des valeurs propres, deux à deux distinctes, de  $u$ . Pour  $i = 1, \dots, r$ , on note

$$E_i = V_{\lambda_i} = \{v \in V \mid u(v) = \lambda_i v\}$$

le sous-espace propre associé. Alors les  $V_{\lambda_i}$  sont en somme directe. (Mais bien sûr, leur somme  $\bigoplus_{i=1}^r V_{\lambda_i}$  n'est pas nécessairement égale à  $V$ ; si  $\dim(V) < \infty$ , c'est le cas si et seulement si  $u$  est diagonalisable.)

*Démonstration.* On va montrer par récurrence sur  $r$  l'assertion :  $(\star_r)$  si l'on a une égalité  $x_1 + \dots + x_r = 0$ , avec  $x_i \in V_{\lambda_i}$ , alors  $x_1 = 0 = \dots = x_r$ . C'est évident si  $r = 1$ , donc on peut supposer  $r \geq 2$  et l'assertion établie pour  $r - 1$ . Supposons qu'on ait une égalité  $x_1 + \dots + x_r = 0$ , avec  $x_i \in V_{\lambda_i}$ . En appliquant l'endomorphisme  $u$ , d'une part, et en multipliant par  $\lambda_r$ , d'autre part, on obtient les égalités :

$$\begin{cases} \lambda_1 x_1 + \dots + \lambda_{r-1} x_{r-1} + \lambda_r x_r = 0 \\ \lambda_r x_1 + \dots + \lambda_r x_{r-1} + \lambda_r x_r = 0 \end{cases}$$

d'où par soustraction l'égalité

$$(*) \quad (\lambda_1 - \lambda_r)x_1 + \dots + (\lambda_{r-1} - \lambda_r)x_{r-1} = 0.$$

Chaque vecteur  $y_i = (\lambda_i - \lambda_r)x_i$  appartient à  $V_{\lambda_i}$  donc, d'après l'hypothèse de récurrence  $(\star_{r-1})$ , l'égalité  $(*)$  entraîne  $y_i = 0$  pour tout  $i = 1, \dots, r - 1$ , et comme  $\lambda_i - \lambda_r \neq 0$  ceci entraîne  $x_i = 0$  pour tout  $i = 1, \dots, r - 1$ . Enfin, reportant ceci dans l'égalité initiale  $x_1 + \dots + x_r = 0$ , on obtient  $x_r = 0$ . Ceci montre que  $(\star_r)$  est vérifiée, et la proposition est démontrée.

**Exemples 2.1.9.** — (1) Soit  $V = \mathcal{C}^\infty(\mathbb{R}, \mathbb{R})$  le  $\mathbb{R}$ -espace vectoriel des fonctions  $f : \mathbb{R} \rightarrow \mathbb{R}$  de classe  $C^\infty$ . Alors les fonctions  $f_\lambda : t \mapsto \exp(\lambda t)$ , pour  $\lambda \in \mathbb{R}$  sont linéairement indépendantes, c.-à-d., quelques soient  $n \in \mathbb{N}^*$  et  $\lambda_1, \dots, \lambda_n$  des réels deux à deux distincts, les fonctions  $f_{\lambda_1}, \dots, f_{\lambda_n}$  sont linéairement indépendantes. En effet, l'opérateur de dérivation  $d : f \mapsto f'$  est un endomorphisme de  $V$  (car  $f'$  est  $C^\infty$  si  $f$  l'est), et chaque  $f_\lambda$  est un vecteur propre de  $d$  pour la valeur propre  $\lambda$ .

(2) Soit  $V = \mathbb{R}^{\mathbb{N}}$  le  $\mathbb{R}$ -espace vectoriel des suites réelles  $(u_n)_{n \in \mathbb{N}}$ . Alors, pour  $\lambda \in \mathbb{R}$ , les suites géométrique  $u(\lambda)$ , définies par  $u(\lambda)_n = \lambda^n$ , sont linéairement indépendantes, c.-à-d., quelques soient  $n \in \mathbb{N}^*$  et  $\lambda_1, \dots, \lambda_n$  des réels deux à deux distincts, les suites  $u(\lambda_1), \dots, u(\lambda_n)$  sont linéairement indépendantes. En effet, soit  $D : V \rightarrow V$  l'opérateur de décalage, défini par  $(D(u))_n = u_{n+1}$  (i.e. l'image par  $D$  de la suite  $(u_0, u_1, u_2, \dots)$  est la suite  $(u_1, u_2, u_3, \dots)$ ); alors chaque  $u(\lambda)$  est un vecteur propre de  $D$  pour la valeur propre  $\lambda$ .

**Définition 2.1.10 (Endomorphismes diagonalisables).** — Soient  $V$  un  $k$ -espace vectoriel de dimension finie et  $u \in \text{End}_k(V)$ . Les conditions suivantes sont équivalentes :

- (i)  $V$  admet une base formée de vecteurs propres de  $u$  ;
- (ii) les vecteurs propres de  $u$  engendrent  $V$  ;
- (iii) la somme des espaces propres de  $u$  égale  $V$  ;
- (iv)  $V$  est la somme directe des espaces propres de  $u$ .

Si ces conditions sont vérifiées, on dit que  $u$  est *diagonalisable*.

*Démonstration.* — En effet, il est clair que (iv)  $\Rightarrow$  (iii)  $\Leftrightarrow$  (ii)  $\Leftarrow$  (i), et (ii)  $\Rightarrow$  (i) car d'un système de générateurs on peut extraire une base. Enfin (iii)  $\Rightarrow$  (iv) d'après le théorème précédent.  $\square$

Une condition *suffisante* de diagonalisabilité est donnée par la proposition ci-dessous. (Bien entendu, cette condition n'est **pas nécessaire** : par exemple la matrice identité  $I_n$  ( $\geq 2$ ) est diagonale et a toutes ses valeurs propres égales à 1 !)



**Proposition 2.1.11 (Valeurs propres distinctes).** — Soit  $u \in \text{End}_k(V)$  ( $\dim_k(V) = n$ ). Si  $P_u(X)$  a  $n$  valeurs propres distinctes, alors  $u$  est diagonalisable.

*Démonstration.* — En effet,  $u$  possède alors  $n$  espaces propres distincts  $V_1, \dots, V_n$ , qui sont en somme directe d'après le théorème précédent. Alors le sous-espace  $E = V_1 \oplus \dots \oplus V_n$  de  $V$  est de dimension

$$\sum_{i=1}^n \dim V_i \geq n = \dim V$$

donc égale  $V$  (et de plus chaque  $V_i$  est de dimension 1).  $\square$

Enfin, une CNS (condition nécessaire et suffisante) de diagonalisabilité est donnée par la :



**Définition et proposition 2.1.12 (Multiplicités algébrique et géométrique d'une valeur propre)**

Soient  $V$  un  $\mathbb{C}$ -espace vectoriel de dimension  $n$ ,  $u \in \text{End}_{\mathbb{C}}(V)$ ,  $\lambda_1, \dots, \lambda_r$  les racines (deux à deux distinctes) de  $P_u(X)$  dans  $\mathbb{C}$ . D'une part,  $P_u(X)$  se factorise

$$P_u(X) = (-1)^n (X - \lambda_1)^{m_1} \dots (X - \lambda_r)^{m_r}$$

où  $m_i$  est la multiplicité de  $\lambda_i$  comme racine de  $P_u(X)$ . D'autre part, d'après 1.5.6,  $\lambda_1, \dots, \lambda_r$  sont les valeurs propres de  $u$ .

(1) On appelle *multiplicité algébrique* (resp. *géométrique*) de la valeur propre  $\lambda_i$  sa multiplicité  $m_i$  comme racine de  $P_u(X)$  (resp. la dimension  $n_i$  de l'espace propre  $V_{\lambda_i}$ ).

(2) On a  $\dim V_{\lambda_i} \leq m_i$  pour tout  $i$ .

(3)  $u$  est diagonalisable  $\Leftrightarrow \dim V_{\lambda_i} = m_i$  pour tout  $i$ .

*Démonstration.* — (2) Pour tout  $i$ , soit  $\mathcal{C}^i$  une base de  $V_{\lambda_i}$ . Comme les espaces propres sont en somme directe, la famille  $\mathcal{C} = \mathcal{C}^1 \cup \dots \cup \mathcal{C}^r$  est une famille libre, donc on peut la compléter en une base  $\mathcal{B}$  de  $V$ . Alors  $A = \text{Mat}_{\mathcal{B}}(u)$  est de la forme suivante :

$$A = \left( \begin{array}{c|c|c|c|c} \lambda_1 I_{n_1} & 0 & \cdots & 0 & * \\ \hline 0 & \lambda_2 I_{n_2} & \ddots & \vdots & * \\ \hline 0 & \ddots & \ddots & 0 & * \\ \hline \vdots & \ddots & 0 & \lambda_r I_{n_r} & * \\ \hline 0 & \cdots & 0 & 0 & B \end{array} \right)$$

où  $B$  est une matrice carrée de taille  $p = n - (n_1 + \dots + n_r)$ . En particulier,  $A$  est triangulaire par blocs. Donc, d'après 1.4.6, on a

$$P_u(X) = \det(A - XI_n) = P_B(X) \prod_{i=1}^r (\lambda_i - X)^{n_i}.$$

Donc  $\prod_{i=1}^r (\lambda_i - X)^{n_i}$  divise  $P_u(X)$ , d'où  $n_i \leq m_i$  pour tout  $i$ , ce qui prouve (2).

Si  $n_i = m_i$  pour tout  $i$ , alors le sous-espace  $E = \bigoplus_{i=1}^r V_{\lambda_i}$  est de dimension  $\sum_{i=1}^r m_i = n$ , donc égale  $V$ , donc  $u$  est diagonalisable. Réciproquement, si  $u$  est diagonalisable, il existe une base  $\mathcal{B}$  de  $V$  telle que

$$A = \text{Mat}_{\mathcal{B}}(u) = \begin{pmatrix} \lambda_1 I_{n_1} & 0 & \cdots & 0 \\ 0 & \lambda_2 I_{n_2} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & \lambda_r I_{n_r} \end{pmatrix}$$

alors  $P_u(X) = \det(A - XI_n) = \prod_{i=1}^r (\lambda_i - X)^{n_i} = (-1)^n \prod_{i=1}^r (X - \lambda_i)^{n_i}$ , d'où  $n_i = m_i$  pour tout  $i$ .  $\square$

Donnons de plus une propriété remarquable des endomorphismes diagonalisables, qui sera utile dans des chapitres ultérieurs. Commençons par une définition :

**Définition 2.1.13 (Restriction de  $u$  à un sous-espace stable).** — Soit  $u \in \text{End}_k(V)$ . On dit qu'un sous-espace  $E$  de  $V$  est stable par  $u$  si  $u(E) \subseteq E$ . Dans ce cas, la restriction de  $u$  à  $E$  induit un endomorphisme de  $E$ , que l'on notera  $u_E$ .

**Théorème 2.1.14 (Restriction d'un endomorphisme diagonalisable)**

Soient  $V$  un  $k$ -espace vectoriel de dimension finie,  $u$  un endomorphisme diagonalisable de  $V$ , et  $E$  un sous-espace de  $V$  stable par  $u$ . Alors  $E$  admet une base formée de vecteurs propres de  $u$ , i.e. la restriction  $u_E$  de  $u$  à  $E$  est diagonalisable.



*Démonstration.* — D'après la proposition précédente, il suffit de montrer que  $E$  est engendré par des vecteurs propres de  $u$ . Comme  $u$  est diagonalisable, tout  $x \in E$  s'écrit dans  $V$  comme une somme de vecteurs propres :

$$(\dagger) \quad x = x_1 + \cdots + x_r, \quad \text{avec } x_i \in V_{\mu_i} \text{ et } \mu_i \neq \mu_j \text{ si } i \neq j.$$

Montrons par récurrence sur  $r$  que pour tout  $x \in E$  et toute écriture  $(\dagger)$  comme ci-dessus, chaque  $x_i$  appartient à  $E$  (ce qui prouvera le théorème). C'est OK pour  $r = 1$ , donc on peut supposer  $r \geq 2$  et le résultat démontré pour  $r - 1$ . Appliquant  $u - \mu_r \text{id}_V$  à  $(\dagger)$  on obtient

$$x' = (u - \mu_r \text{id}_V)(x) = \sum_{i=1}^{r-1} (\mu_i - \mu_r)x_i$$

et  $x' \in E$  puisque  $E$  est stable par  $u$ . Donc par hypothèse de récurrence,  $(\mu_i - \mu_r)x_i$  appartient à  $E$  pour  $i = 1, \dots, r - 1$ , donc  $x_i$  y appartient aussi (puisque  $\mu_i - \mu_r \neq 0$ ), et reportant ceci dans  $(\dagger)$  on obtient aussi  $x_r \in E$ . Ceci prouve le théorème.  $\square$

Pour terminer ce paragraphe, donnons encore l'exemple ci-dessous d'endomorphismes diagonalisables.

**Rappels 2.1.15.** — Soit  $k$  un corps. Si  $n \cdot 1_k = 1_k + \cdots + 1_k$  ( $n$  termes) est  $\neq 0$  pour tout entier  $n > 0$ , on dit que  $k$  est de caractéristique 0 ; c'est le cas par exemple pour  $\mathbb{Q}, \mathbb{R}, \mathbb{C}$ . Sinon, le plus petit entier  $p > 0$  tel que  $p \cdot 1_k = 0$  est nécessairement un nombre premier (car si  $p = rs$  avec  $r, s \geq 1$ , l'égalité  $0 = (r \cdot 1_k)(s \cdot 1_k)$  entraîne que  $r \cdot 1_k = 0$  ou  $s \cdot 1_k = 0$ , disons  $r \cdot 1_k = 0$ , mais alors la minimalité de  $p$  entraîne que  $r = p$ ) ; dans ce cas on dit que  $k$  est de caractéristique  $p$ . D'autre part, si  $V$  est un  $k$ -espace vectoriel et  $p \in \text{End}_k(V)$ , rappelons qu'on dit que  $p$  est un **projecteur** si  $p^2 = p \circ p$  est égal à  $p$ .

**Proposition 2.1.16 (Symétries).** — Soient  $k$  un corps de caractéristique  $\neq 2$  (par exemple,  $k = \mathbb{Q}, \mathbb{R}$  ou  $\mathbb{C}$ ),  $V$  un  $k$ -espace vectoriel de dimension  $n$ , et  $s \in \text{End}_k(V)$  tel que  $s^2 = \text{id}_V$ . Alors  $s$  est diagonalisable ; plus précisément, soient

$$p_+ = \frac{\text{id}_V + s}{2}, \quad p_- = \frac{\text{id}_V - s}{2}, \quad V_{\pm} = \text{Im}(p_{\pm}).$$

Alors  $p_+$  et  $p_-$  sont des projecteurs et l'on a :

$$V = V_+ \oplus V_- \quad \text{et} \quad \forall x \in V_{\pm}, \quad s(x) = \pm x.$$

Donc, si  $s \neq \pm \text{id}_V$ , alors  $V_+$  et  $V_-$  sont non nuls, et  $V_{\pm}$  est l'espace propre associé à la valeur propre  $\pm 1$  ; dans ce cas,  $s$  est la symétrie par rapport à  $V_+$  parallèlement à  $V_-$ .

*Démonstration.* — On a  $p_+^2 = p_+$ ,  $p_-^2 = p_-$  et  $p_- = \text{id}_V - p_+$  d'où  $p_+ p_- = 0 = p_- p_+$ , donc  $p_+$  et  $p_-$  sont des projecteurs et l'on a  $V = V_+ \oplus V_-$ . De plus, si  $x \in V_{\pm}$ , on voit aussitôt que  $s(x) = \pm x$ , d'où la proposition.  $\square$

**Remarque 2.1.16.1.** — Attention, si  $k$  est de caractéristique 2, c.-à-d., si  $2 = 0$  dans  $k$  (par exemple, si  $k$  est le corps à deux éléments  $\mathbb{F}_2 = \mathbb{Z}/2\mathbb{Z}$ ), la matrice  $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \in M_2(k)$  vérifie  $A^2 = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} = I_2$  mais  $A$  n'est pas diagonalisable : en effet sa seule valeur propre est 1, donc si  $A$  était diagonalisable on aurait  $A = I_2$ , ce qui n'est pas le cas.

## 2.2. Trigonalisation, théorème de Cayley-Hamilton et espaces caractéristiques



**Définition 2.2.1 (Endomorphismes trigonalisables).** — Soit  $u \in \text{End}_k(V)$  ( $\dim_k(V) = n$ ). On dit que  $u$  est *trigonalisable* s'il existe une base  $\mathcal{B}$  de  $V$  dans laquelle la matrice  $A = \text{Mat}_{\mathcal{B}}(u)$  est triangulaire, disons supérieure. <sup>(1)</sup>

Dans ce cas, soient  $\lambda_1, \dots, \lambda_n$  les coefficients diagonaux, et soit  $X$  une indéterminée. Alors, d'après 1.4.6, on a

$$P_u(X) = \det(A - XI_n) = \prod_{i=1}^n (\lambda_i - X)$$

donc  $\lambda_1, \dots, \lambda_n$  sont les  $n$  racines (comptées avec multiplicités) du polynôme caractéristique  $P_u(X)$ . On voit donc qu'une condition *nécessaire* pour que  $u$  soit trigonalisable est que  $P_u(X)$  ait toutes ses racines dans  $k$ . Ceci conduit à la définition suivante :

**Définition 2.2.2 (Polynômes scindés, corps algébriquement clos)**

Soient  $k$  un corps et  $P \in k[X]$  un polynôme de degré  $n \geq 1$ .

(1) On dit que  $P$  est *scindé* dans  $k[X]$  s'il admet  $n$  racines  $\lambda_1, \dots, \lambda_n$  dans  $k$  (comptées avec multiplicités), c.-à-d., si  $P$  se factorise dans  $k[X]$  en produit de facteurs de degré 1 :

$$P = a(X - \lambda_1) \cdots (X - \lambda_n)$$

(où  $a$  est le coefficient dominant de  $P$ ).

(2) On dit que  $k$  est *algébriquement clos* si tout polynôme  $P \in k[X]$  de degré  $\geq 1$  est scindé. Par exemple, on sait que  $\mathbb{C}$  est algébriquement clos (une démonstration est donnée dans un appendice à ce chapitre).

Pour simplifier l'exposition, on suppose dans la suite de ce paragraphe 2.2 que  $k = \mathbb{C}$ . (Mais tous les résultats, à l'exception du corollaire 2.2.14, sont valables sur un corps algébriquement clos arbitraire [et le corollaire 2.2.14 est valable pour tout corps algébriquement clos de caractéristique 0]).



**Théorème 2.2.3 (Trigonalisation).** — (1) Soient  $V$  un  $\mathbb{C}$ -espace vectoriel de dimension  $n$  et  $u \in \text{End}_{\mathbb{C}}(V)$ . Il existe une base  $\mathcal{C}$  de  $V$  telle que  $\text{Mat}_{\mathcal{C}}(u)$  soit triangulaire supérieure (les coefficients diagonaux  $\lambda_1, \dots, \lambda_n$  sont alors les racines, comptées avec multiplicité, de  $P_u(X)$  dans  $\mathbb{C}$ ).

En termes matriciels : toute matrice  $A \in M_n(\mathbb{C})$  est semblable à une matrice triangulaire, i.e. il existe  $P \in \text{GL}_n(\mathbb{C})$  telle que  $P^{-1}AP$  soit triangulaire.

(2) Plus généralement, soient  $k$  un corps,  $V$  un  $k$ -espace vectoriel de dimension  $n$  et  $u \in \text{End}_k(V)$  ; on suppose que  $P_u(X)$  est scindé, i.e. qu'il a  $n$  racines  $\lambda_1, \dots, \lambda_n$  (pas nécessairement distinctes) dans  $k$ . Alors il existe une base  $\mathcal{B}$  de  $V$  dans laquelle la matrice de  $u$  est triangulaire supérieure, les coefficients diagonaux étant alors  $\lambda_1, \dots, \lambda_n$ .

En termes matriciels : si le polynôme caractéristique de  $A \in M_n(k)$  a toutes ses racines dans  $k$ , alors  $A$  est semblable à une matrice triangulaire, i.e. il existe  $P \in \text{GL}_n(k)$  telle que  $P^{-1}AP$  soit triangulaire.

*Démonstration.* — (1) On procède par récurrence sur  $n = \dim(V)$ . Il n'y a rien à montrer si  $n = 1$ , donc on peut supposer  $n \geq 2$  et le théorème démontré pour  $n-1$ . Soient  $\mathcal{B}$  une base arbitraire de  $V$  et  $A = \text{Mat}_{\mathcal{B}}(u)$ . Considérons la matrice  $B = {}^tA$ . Comme  $\mathbb{C}$  est algébriquement clos, le polynôme caractéristique de  $B$  a au

moins une racine  $\lambda$  dans  $\mathbb{C}$  donc, d'après 1.5.6, il existe un vecteur  $T = \begin{pmatrix} t_1 \\ \vdots \\ t_n \end{pmatrix} \neq 0$  tel que  ${}^tAT = BT = \lambda T$ .

Prenant la transposée de cette égalité, on obtient :

$$(*) \quad \lambda(t_1, \dots, t_n) = (t_1, \dots, t_n) {}^t({}^tA) = (t_1, \dots, t_n)A.$$

<sup>(1)</sup>Si la matrice de  $u$  dans une base  $(v_1, \dots, v_n)$  est triangulaire supérieure, alors la matrice dans la base  $(v_n, \dots, v_1)$  est triangulaire inférieure, et vice-versa, donc on pourrait dans la définition remplacer le mot « supérieure » par « inférieure ».

D'après 1.3.2, le sous-espace  $H$  de  $V$  défini par l'équation  $t_1x_1 + \dots + t_nx_n = 0$ , i.e. le sous-espace de  $V$  formé des vecteurs colonnes  $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$  tels que  $(t_1, \dots, t_n)X = (t_1, \dots, t_n) \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = 0$ , est de dimension  $n - 1$ . Pour tout  $X \in H$ , on a, d'après  $(\star)$  :

$$(t_1, \dots, t_n)AX = \lambda \underbrace{(t_1, \dots, t_n)X}_{=0} = 0,$$

donc  $AX \in H$ . Comme  $AX$  représente dans la base  $\mathcal{B}$  le vecteur  $u(X)$ , ceci montre que  $H$  est stable par  $u$ . Comme  $\dim H = n - 1$ , alors par hypothèse de récurrence il existe une base  $\mathcal{C}'$  de  $H$  telle  $\text{Mat}_{\mathcal{C}'}(u_H)$  soit triangulaire supérieure. Choisissons arbitrairement  $w \in V - H$ , alors la famille  $\mathcal{C} = \mathcal{C}' \cup \{w\}$  est libre, donc est une base de  $V$ , et  $\text{Mat}_{\mathcal{C}}(u)$  est de la forme :

$$A' = \left( \begin{array}{cccc|c} a_{1,1} & a_{1,2} & \cdots & a_{1,n-1} & a_{1,n} \\ 0 & a_{2,2} & \cdots & a_{2,n-1} & a_{2,n} \\ \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & \cdots & 0 & a_{n-1,n-1} & a_{n-1,n} \\ \hline 0 & \cdots & 0 & 0 & a_{n,n} \end{array} \right) \quad \text{où} \quad \text{Mat}_{\mathcal{C}'}(u_H) = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n-1} \\ 0 & a_{2,2} & \cdots & a_{2,n-1} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & a_{n-1,n-1} \end{pmatrix}$$

donc triangulaire supérieure. Enfin, on a déjà noté en 2.2.1 que, dans ce cas, les coefficients diagonaux sont les racines (comptées avec multiplicité) de  $P_u(X)$  dans  $\mathbb{C}$ . Ceci prouve (1).

Prouvons l'assertion plus générale (2), i.e.  $k$  est un corps arbitraire et l'on suppose que  $P = P_u(X)$  a  $n$  racines  $\lambda_1, \dots, \lambda_n$  (pas nécessairement distinctes) dans  $k$ . On procède par récurrence sur  $n = \dim(V)$ . Il n'y a rien à montrer si  $n = 1$ , donc on peut supposer  $n \geq 2$  et le théorème démontré pour  $n - 1$ . Soit  $\mathcal{B}$  une base arbitraire de  $V$  et soient  $A = \text{Mat}_{\mathcal{B}}(u)$  et  $B = {}^tA$ . Comme

$${}^t(A - XI_n) = B - XI_n$$

alors  $P_B(X) = P_A(X) = P_u(X) = P$ , donc  $\lambda = \lambda_1$  est racine de  $P_B(X)$ , donc valeur propre de  $B$  (cf. 1.5.6).

Il existe donc un vecteur  $T = \begin{pmatrix} t_1 \\ \vdots \\ t_n \end{pmatrix} \neq 0$  tel que  ${}^tAT = BT = \lambda T$ . Prenant la transposée de cette égalité,

on obtient :

$$(\star) \quad \lambda(t_1, \dots, t_n) = (t_1, \dots, t_n) {}^t({}^tA) = (t_1, \dots, t_n)A.$$

D'après 1.3.2, le sous-espace  $H$  de  $V$  défini par l'équation  $t_1x_1 + \dots + t_nx_n = 0$ , i.e. le sous-espace de  $V$  formé des vecteurs colonnes  $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$  tels que  $(t_1, \dots, t_n)X = (t_1, \dots, t_n) \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = 0$ , est de dimension  $n - 1$ . Pour tout  $X \in H$ , on a, d'après  $(\star)$  :

$$(t_1, \dots, t_n)AX = \lambda \underbrace{(t_1, \dots, t_n)X}_{=0} = 0,$$

donc  $AX \in H$ . Comme  $AX$  représente dans la base  $\mathcal{B}$  le vecteur  $u(X)$ , ceci montre que  $H$  est stable par  $u$ . Soient  $\mathcal{B}_H$  une base de  $H$  et  $w \in V - H$ , alors la famille  $\tilde{\mathcal{B}} = \mathcal{B}_H \cup \{w\}$  est libre, donc est une base de  $V$ . Soient  $u_H$  la restriction de  $u$  à  $H$  et  $A' = \text{Mat}_{\mathcal{B}_H}(u_H) \in M_{n-1}(k)$ ; alors  $\text{Mat}_{\tilde{\mathcal{B}}}(u)$  est de la forme :

$$(\star) \quad \tilde{A} = \left( \begin{array}{c|c} A' & Y \\ \hline \mathbf{0}_{1,n-1} & \mu \end{array} \right)$$

où  $\mu \in k$ ,  $Y \in M_{n-1,1}(k)$  est une matrice colonne, et  $\mathbf{0}_{1,n-1}$  est la matrice ligne  $(0, \dots, 0)$ .

Ceci entraîne que  $P_{A'}(X)(\mu - X) = P_{\tilde{A}}(X) = P_u(X)$ . Il en résulte que  $\mu$  est l'une des racines de  $P$ , disons  $\lambda_i$ , et que  $P_{A'}(X)$  est le produit, pour  $j \neq i$ , des  $(\lambda_j - X)$ ; en particulier,  $P_{A'}(X)$  a toutes ses racines dans  $k$ . (En fait, on peut montrer que  $\mu = \lambda = \lambda_1$ , voir plus bas). Comme  $\dim H = n - 1$ , alors par hypothèse de récurrence il existe une base  $\mathcal{C}'$  de  $H$  telle  $\text{Mat}_{\mathcal{C}'}(u_H)$  soit triangulaire supérieure (avec les  $\lambda_j$ , pour  $j \neq i$ , comme termes diagonaux). Alors  $\mathcal{C} = \mathcal{C}' \cup \{w\}$  est une base de  $V$ , et  $\text{Mat}_{\mathcal{C}}(u)$  est de

la forme :

$$A' = \left( \begin{array}{cccc|c} a_{1,1} & a_{1,2} & \cdots & a_{1,n-1} & a_{1,n} \\ 0 & a_{2,2} & \cdots & a_{2,n-1} & a_{2,n} \\ \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & \cdots & 0 & a_{n-1,n-1} & a_{n-1,n} \\ \hline 0 & \cdots & 0 & 0 & \mu \end{array} \right) \quad \text{où} \quad \left( \begin{array}{cccc} a_{1,1} & a_{1,2} & \cdots & a_{1,n-1} \\ 0 & a_{2,2} & \cdots & a_{2,n-1} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & a_{n-1,n-1} \end{array} \right) = \text{Mat}_{\mathcal{E}'}(u_H),$$

donc triangulaire supérieure. Enfin, on a déjà noté en 2.2.1 que, dans ce cas, les coefficients diagonaux sont les racines (comptées avec multiplicité) de  $P_u(X)$  dans  $k$ . Le théorème est démontré.

Remarquons enfin que  $\mu = \lambda$ . En effet,  ${}^t A$  est la matrice de l'endomorphisme  ${}^t u : V^* \rightarrow V^*$ ,  $\phi \mapsto \phi \circ u$ , et le vecteur  $T$  obtenu au début de la démonstration correspond donc à une forme linéaire  $f \in V^*$  telle que

$$f \circ u = ({}^t u)(f) = \lambda f.$$

De plus, on a  $H = \text{Ker}(f)$  et puisque  $w \notin H$  alors le scalaire  $z = f(w)$  est  $\neq 0$ . On a  $u(w) = \mu w + y$ , avec  $y \in H$  et donc

$$(\dagger) \quad f(u(w)) = f(\mu w + y) = \mu f(w).$$

D'autre part,

$$(\ddagger) \quad f(u(w)) = (f \circ u)(w) = (\lambda f)(w) = \lambda f(w),$$

et comme  $f(w) \neq 0$  la comparaison de  $(\dagger)$  et  $(\ddagger)$  donne  $\mu = \lambda$ . □

**Corollaire 2.2.4 (Déterminant (resp. trace) = produit (resp. somme) des valeurs propres)**

Soient  $A \in M_n(\mathbb{C})$  et  $\lambda_1, \dots, \lambda_n$  les  $n$  racines dans  $\mathbb{C}$  (comptées avec multiplicité) du polynôme caractéristique  $P_A(X)$ . Alors  $\boxed{\det(A) = \lambda_1 \cdots \lambda_n}$  et  $\boxed{\text{Tr}(A) = \lambda_1 + \cdots + \lambda_n}$ .

*Démonstration.* — D'après le théorème 2.2.3, il existe  $P \in \text{GL}_n(\mathbb{C})$  telle que  $A' = P^{-1}AP$  soit triangulaire; notons  $\lambda_1, \dots, \lambda_n$  ses coefficients diagonaux, alors

$$P_{A'}(X) = \prod_{i=1}^n (\lambda_i - X), \quad \det(A') = \lambda_1 \cdots \lambda_n, \quad \text{Tr}(A') = \lambda_1 + \cdots + \lambda_n.$$

D'autre part, d'après le théorème 1.5.2,  $A'$  et  $A$  ont même polynôme caractéristique, même déterminant et même trace. Alors  $\lambda_1, \dots, \lambda_n$  sont les racines de  $P_{A'}(X) = P_A(X)$  et l'on a  $\det(A) = \det(A') = \lambda_1 \cdots \lambda_n$  et  $\text{Tr}(A) = \text{Tr}(A') = \lambda_1 + \cdots + \lambda_n$ . □

**Définition 2.2.5 (Polynômes d'endomorphismes ou de matrices).** — Soit  $Q = a_0 + a_1X + \cdots + a_dX^d$  un polynôme à coefficients dans  $k$ .

(1) Soient  $V$  un  $k$ -espace vectoriel et  $u \in \text{End}_k(V)$ . On pose

$$Q(u) = a_0 \text{id}_V + a_1 u + \cdots + a_d u^d$$

où, bien sûr,  $u^i$  désigne  $u \circ \cdots \circ u$  ( $i$  fois); alors  $Q(u)$  est un élément de  $\text{End}_k(V)$ , qu'on appelle un « polynôme en  $u$  ». On vérifie aussitôt que si  $R \in k[X]$  est un second polynôme, alors

$$Q(u) \circ R(u) = (QR)(u) = (RQ)(u) = R(u) \circ Q(u);$$

en particulier, « les polynômes en  $u$  commutent entre eux ». Donc, l'application  $\phi : k[X] \rightarrow \text{End}_k(V)$ ,  $Q \mapsto Q(u)$ , est un *homomorphisme de  $k$ -algèbres*, c.-à-d.,  $\phi$  est linéaire, envoie  $1 \in k[X]$  sur  $\text{id}_V$  et vérifie  $\phi(QR) = \phi(Q)\phi(R)$ .

(2) De même, si  $A \in M_n(k)$ , on note  $Q(A)$  la matrice  $a_0 I_n + a_1 A + \cdots + a_d A^d$ . Si  $R \in k[X]$  est un second polynôme, alors

$$Q(A)R(A) = (QR)(A) = (RQ)(A) = R(A)Q(A);$$

en particulier, « les polynômes en  $A$  commutent entre eux ». De plus, si  $\dim(V) = n$  et si  $A$  est la matrice de  $u$  dans une base  $\mathcal{B}$ , alors la matrice dans  $\mathcal{B}$  de  $Q(u)$  est  $Q(A)$ .

**Remarque 2.2.5.1.** — Soit  $u \in \text{End}_k(V)$ . Si  $v \in V$  est un vecteur propre de  $u$  pour la valeur propre  $\lambda$ , alors

$$u^2(v) = u(u(v)) = u(\lambda v) = \lambda u(v) = \lambda^2 v$$

et l'on montre ainsi, par récurrence sur  $i$ , que  $u^i(v) = \lambda^i v$  pour tout  $i \in \mathbb{N}^*$  (et aussi pour  $i = 0$ , avec la convention  $u^0 = \text{id}_V$  et  $\lambda^0 = 1$ ). On en déduit que pour tout  $Q \in k[X]$ , on a  $Q(u)(v) = Q(\lambda)v$ .



**Théorème 2.2.6 (Théorème de Cayley-Hamilton).** — Soit  $V$  un  $\mathbb{C}$ -espace vectoriel de dimension  $n$ , et soient  $u \in \text{End}_{\mathbb{C}}(V)$  et  $P_u(X)$  son polynôme caractéristique. Alors l'endomorphisme  $P_u(u)$  est nul, c.-à-d. : «  $u$  est annulé par son polynôme caractéristique ».

*Démonstration.* — On a  $P_u(X) = (-1)^n \prod_{i=1}^n (X - \lambda_i)$ , où  $\lambda_1, \dots, \lambda_n$  sont les  $n$  racines (comptées avec multiplicité) de  $P_u(X)$  dans  $\mathbb{C}$ . D'après le théorème 2.2.3, il existe une base  $\mathcal{B} = (f_1, \dots, f_n)$  de  $V$  dans laquelle la matrice de  $u$  est triangulaire supérieure, avec  $\lambda_1, \dots, \lambda_n$  sur la diagonale. Ceci équivaut à dire que, pour tout  $i = 1, \dots, n$ , le sous-espace  $F_i$  de  $V$  engendré par  $f_1, \dots, f_i$  est stable par  $u$  et, plus précisément, que l'on a, pour  $i = 1, \dots, n$  :

$$(u - \lambda_i \text{id}_V)(F_i) \subseteq F_{i-1},$$

avec la convention  $F_0 = \{0\}$ . Comme  $F_n = V$ , on déduit des inclusions ci-dessus que  $(u - \lambda_n \text{id}_V)(V) \subseteq F_{n-1}$ , puis que  $(u - \lambda_{n-1} \text{id}_V)(u - \lambda_n \text{id}_V)(V) \subseteq F_{n-2}$ , etc., d'où finalement :

$$(u - \lambda_1 \text{id}_V) \cdots (u - \lambda_n \text{id}_V)(V) \subseteq F_0 = \{0\}.$$

Ceci montre que  $(-1)^n P_u(u) = 0$ , d'où  $P_u(u) = 0$ . □



**Corollaire 2.2.7 (Cayley-Hamilton pour  $k \subseteq \mathbb{C}$ ).** — Soient  $k$  un sous-corps de  $\mathbb{C}$  (par exemple  $k = \mathbb{R}$ ) et  $V$  un  $k$ -espace vectoriel de dimension  $n$ , et soient  $u \in \text{End}_k(V)$  et  $P_u(X)$  son polynôme caractéristique. Alors  $P_u(u) = 0$

*Démonstration.* — Soient  $\mathcal{B}$  une base de  $V$  et  $A = \text{Mat}_{\mathcal{B}}(u) \in M_n(k)$ . D'une part,  $P_u(X) = P_A(X)$ , notons  $P$  ce polynôme. D'autre part, comme  $k \subseteq \mathbb{C}$ , on peut considérer  $A$  comme élément de  $M_n(\mathbb{C})$ , donc d'après Cayley-Hamilton on a  $P(A) = 0$ . Or  $P(A)$  est la matrice dans la base  $\mathcal{B}$  de  $P(u)$ , d'où  $P(u) = 0$ . □



**Définition 2.2.8 (Espaces caractéristiques).** — Soient  $V$  un  $\mathbb{C}$ -espace vectoriel de dimension  $n$ ,  $u \in \text{End}_{\mathbb{C}}(V)$ ,  $\lambda \in \mathbb{C}$  une valeur propre de  $u$ , et  $m$  sa multiplicité algébrique (i.e. sa multiplicité comme racine de  $P_u(X)$ ). On pose

$$V_{(\mu)} = \text{Ker}((u - \lambda \text{id}_V)^m)$$

et l'appelle l'espace caractéristique associé à  $\mu$ ; il est stable par  $u$  (car il est stable par  $u - \lambda \text{id}_V$  donc aussi par  $u = (u - \lambda \text{id}_V) + \lambda \text{id}_V$ ). Notons aussi que  $V_{(\mu)}$  contient l'espace propre  $V_{\mu} = \text{Ker}(u - \lambda \text{id}_V)$ .

**Exemple 2.2.9.** — Si  $u$  est l'endomorphisme de  $\mathbb{C}^2$  de matrice  $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$  (i.e. ,  $u(e_2) = e_1$  et  $u(e_1) = 0$ ), alors  $P_u(X) = X^2$  a 0 comme unique racine. D'après Cayley-Hamilton (ou par un calcul direct), on a  $u^2 = 0$ , donc  $V_{(0)} = \mathbb{C}^2$  tandis que  $V_0 = \text{Ker}(u) = \mathbb{C}e_1$ .

Afin de démontrer plus bas un théorème sur les espaces caractéristiques, on aura besoin du théorème ci-dessous, que les étudiants de LM220 ont déjà vu (une démonstration est donnée en appendice à la fin de ce chapitre).

**Théorème 2.2.10 (Théorème de Bézout).** — Soient  $P_1, \dots, P_r \in \mathbb{C}[X]$  des polynômes sans racine commune. Alors il existe des polynômes  $S_1, \dots, S_r \in \mathbb{C}[X]$  tels que  $P_1 S_1 + \dots + P_r S_r = 1$ .



**Théorème 2.2.11 (Décomposition en espaces caractéristiques).** — Soient  $V$  un  $\mathbb{C}$ -espace vectoriel de dimension  $n$  et  $u \in \text{End}_{\mathbb{C}}(V)$ . Écrivons  $P_u(X) = (-1)^n \prod_{i=1}^r (X - \lambda_i)^{m_i}$ , où  $\lambda_1, \dots, \lambda_r$  sont les valeurs propres, deux à deux distinctes, de  $u$  et  $m_i$  est la multiplicité algébrique de  $\lambda_i$ . Alors :

- (1)  $V$  est la somme directe des espaces caractéristiques  $V_{(\lambda_1)}, \dots, V_{(\lambda_r)}$ .
- (2) On a  $\dim V_{(\lambda_i)} = m_i$  pour tout  $i$ .

*Démonstration.* — Pour tout  $i$ , posons

$$P_i = (-1)^n \prod_{j \neq i} (X - \lambda_j)^{m_j}.$$

Alors  $P_1, \dots, P_r \in \mathbb{C}[X]$  sont sans racines communes, donc d'après le théorème de Bézout il existe des polynômes  $S_1, \dots, S_r \in \mathbb{C}[X]$  tels que  $P_1 S_1 + \dots + P_r S_r = 1$ . Prenant les polynômes en  $u$  correspondants, on obtient :

$$P_1(u)S_1(u) + \dots + P_r(u)S_r(u) = \text{id}_V$$

Donc, pour tout  $v \in V$ , on a :

$$(*) \quad v = P_1(u)(x_1) + \dots + P_r(u)(x_r), \quad \text{où } x_i = S_i(u)(v).$$



D'autre part, d'après Cayley-Hamilton,  $P_u(u) = 0$ ; comme de plus  $P_u(X) = (X - \lambda_i)^{m_i} P_i(X)$  alors

$$(u - \lambda_i \text{id}_V)^{m_i} (P_i(u)(x_i)) = P_u(u)(x_i) = 0, \quad \text{pour tout } i.$$

Ceci montre que chaque  $v_i = P_i(u)(x_i)$  appartient à l'espace caractéristique  $V_{(\lambda_i)} = \text{Ker}((u - \lambda_i \text{id}_V)^{m_i})$ ; par conséquent l'égalité  $(\star)$  montre déjà que  $V = \sum_{i=1}^r V_{(\lambda_i)}$ .

Reste à montrer que la somme des  $V_{(\lambda_i)}$  est directe : supposons qu'on ait une égalité

$$v_1 + \dots + v_r = 0, \quad \text{avec } v_i \in V_{(\lambda_i)}$$

et montrons que chaque  $v_i$  est nul. Fixons un indice  $i$ , alors  $v_i = -\sum_{j \neq i} v_j$  appartient à  $V_{(\lambda_i)} \cap \sum_{j \neq i} V_{(\lambda_j)}$  donc est annulé par  $(u - \lambda_i \text{id}_V)^{m_i}$  et par  $P_i(u) = (-1)^n \prod_{j \neq i} (u - \lambda_j \text{id}_V)^{m_j}$ . Or, les polynômes  $Q_i(X) = (X - \lambda_i)^{m_i}$  et  $P_i(X)$  n'ont pas de racines communes donc, d'après Bézout, il existe  $R, T \in \mathbb{C}[X]$  tels que  $RP_i + TQ_i = 1$ . Alors  $\text{id}_V = R(u)P_i(u) + T(u)Q_i(u)$ , d'où

$$v_i = \text{id}_V(v_i) = R(u) \underbrace{P_i(u)(v_i)}_{=0} + T(u) \underbrace{Q_i(u)(v_i)}_{=0} = 0.$$

Ceci montre que la somme des  $V_{(\lambda_i)}$  est directe, et achève la démonstration de l'assertion (1).

Enfin, pour tout  $i$ , soit  $d_i = \dim V_{(\lambda_i)}$ , soit  $u_i$  la restriction de  $u$  à  $V_{(\lambda_i)}$  et soit  $\mathcal{B}_i$  une base de  $\dim V_{(\lambda_i)}$ . Alors  $\mathcal{B} = \mathcal{B}_1 \cup \dots \cup \mathcal{B}_r$  est une base de  $V$  et, notant  $A_i = \text{Mat}_{\mathcal{B}_i}(u_i)$ , on a :

$$(\dagger) \quad \text{Mat}_{\mathcal{B}}(u) = \begin{pmatrix} A_1 & 0 & \dots & 0 \\ 0 & A_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & A_r \end{pmatrix} \quad \text{d'où} \quad P_u(X) = \prod_{i=1}^r P_{u_i}(X).$$

D'autre part, si  $\mu$  est une valeur propre de  $u_i$  et  $x \neq 0$  un vecteur propre associé, on a  $(u_i - \lambda_i \text{id}_V)(x) = (\mu - \lambda_i)x$ , d'où  $(u_i - \lambda_i \text{id}_V)^p(x) = (\mu - \lambda_i)^p x$  pour tout  $p \in \mathbb{N}^*$ . Or  $(u_i - \lambda_i \text{id}_V)^{m_i} = 0$ , d'où  $(\mu - \lambda_i)^{m_i} x = 0$  et donc  $\mu = \lambda_i$ . Ceci montre que  $\lambda_i$  est la seule valeur propre de  $u_i$ , d'où  $P_{u_i}(X) = (-1)^{d_i} (X - \lambda_i)^{d_i}$ . D'après  $(\dagger)$ , on obtient donc que

$$P_u(X) = (-1)^n \prod_{i=1}^r (X - \lambda_i)^{d_i}$$

d'où l'égalité  $d_i = m_i$  pour tout  $i$ . Le théorème est démontré.  $\square$

**Remarque 2.2.12.** — On conserve les notations précédentes. En fait, chaque espace caractéristique  $V_{(\lambda_i)}$  est l'ensemble de tous les  $v \in V$  qui sont annulés par une certaine puissance de  $(u - \lambda_i \text{id}_V)$ , i.e. on a l'égalité

$$V_{(\lambda_i)} = \{v \in V \mid \exists p \in \mathbb{N}^* \text{ tel que } (u - \lambda_i \text{id}_V)^p(v) = 0\}.$$

En effet, l'inclusion  $\subseteq$  est claire, prouvons la réciproque. Quitte à renuméroter les  $\lambda_i$ , il suffit de le faire pour  $\lambda_1$ . Supposons que  $(u - \lambda_1 \text{id}_V)^p(v) = 0$ . On peut écrire  $v = v_1 + v_2 + \dots + v_r$ , avec  $v_i \in V_{(\lambda_i)}$ . Comme chaque  $V_i = V_{(\lambda_i)}$  est stable par  $u$  donc aussi par  $u - \lambda_1 \text{id}_V$ , l'égalité

$$0 = (u - \lambda_1 \text{id}_V)^p(v) = (u - \lambda_1 \text{id}_V)^p(v_1) + \dots + (u - \lambda_1 \text{id}_V)^p(v_r)$$

jointe au fait que les  $V_i$  sont en somme directe, entraîne que  $(u - \lambda_1 \text{id}_V)^p(v_i) = 0$  pour tout  $i$ . Or on a vu que la restriction  $u_i$  de  $u$  à  $V_i$  a pour polynôme caractéristique  $(\lambda_i - X)^{m_i}$ , donc  $\det(u_i - \lambda_1 \text{id}_{V_i}) = (\lambda_i - \lambda_1)^{m_i}$  est non nul lorsque  $i \neq 1$ , donc l'égalité  $(u - \lambda_1 \text{id}_V)^p(v_i) = 0$  entraîne  $v_i = 0$  pour  $i \neq 1$ . Donc  $v$  égale  $v_1$  et appartient donc à  $V_{(\lambda_1)}$ .

Remarquons aussi qu'une démonstration analogue (et en fait, plus simple) donne la :



**Proposition 2.2.13 (Polynômes sans racines multiples).** — Soient  $V$  un  $\mathbb{C}$ -espace vectoriel de dimension  $n$  et  $u \in \text{End}_{\mathbb{C}}(V)$ . On suppose que  $Q(u) = 0$ , où  $Q \in \mathbb{C}[X]$  un polynôme de degré  $d \geq 1$  ayant  $d$  racines distinctes dans  $\mathbb{C}$ . Alors  $u$  est diagonalisable.

*Démonstration.* — Par hypothèse, on a  $Q(X) = c \prod_{i=1}^d (X - \alpha_i)$ , où  $c \in \mathbb{C}^*$  est le coefficient dominant de  $Q$ , et  $\alpha_1, \dots, \alpha_d \in \mathbb{C}$  sont deux à deux distincts. Pour tout  $i$ , posons

$$P_i = c \prod_{j \neq i} (X - \alpha_j).$$

Alors  $P_1, \dots, P_d \in \mathbb{C}[X]$  sont sans racines communes donc, d'après le théorème de Bézout il existe des polynômes  $S_1, \dots, S_d \in \mathbb{C}[X]$  tels que  $P_1 S_1 + \dots + P_d S_d = 1$ . Prenant les polynômes en  $u$  correspondants, on obtient :

$$P_1(u)S_1(u) + \dots + P_d(u)S_d(u) = \text{id}_V$$

Donc, pour tout  $v \in V$ , on a :

$$(\star) \quad v = P_1(u)(x_1) + \dots + P_d(u)(x_d), \quad \text{où } x_i = S_i(u)(v).$$

Par hypothèse,  $Q(u)$  est l'endomorphisme nul; comme de plus  $Q(X) = (X - \alpha_i)P_i(X)$  alors

$$(u - \alpha_i \text{id}_V)(P_i(u)(x_i)) = Q(u)(x_i) = 0,$$

pour tout  $i$ . Ceci montre que chaque  $v_i = P_i(u)(x_i)$  appartient à  $\text{Ker}(u - \alpha_i \text{id}_V)$ ; par conséquent l'égalité  $(\star)$  montre que  $V$  est la somme des espaces  $\text{Ker}(u - \alpha_i \text{id}_V)$ .

Or, ceux de ces espaces qui sont non nuls sont des espaces propres de  $u$ , donc sont en somme directe (cf. 2.1.8). Donc, notant  $I$  le sous-ensemble de  $\{1, \dots, d\}$  formé des  $i$  tels que  $\text{Ker}(u - \alpha_i \text{id}_V) \neq \{0\}$ , on obtient que  $V = \bigoplus_{i \in I} V_{\alpha_i}$ , ce qui montre que  $u$  est diagonalisable.  $\square$



**Corollaire 2.2.14 (Automorphismes d'ordre fini de  $\mathbb{C}^n$ ).** — Soient  $V$  un  $\mathbb{C}$ -espace vectoriel de dimension  $n$  et  $u \in \text{End}_{\mathbb{C}}(V)$  tel que  $u^d = \text{id}_V$  pour un entier  $d \geq 1$  (dans ce cas, on dit que  $u$  est un automorphisme d'ordre fini). Alors  $u$  est diagonalisable (et ses valeurs propres sont des racines  $d$ -èmes de l'unité).

*Démonstration.* — En effet, le polynôme  $X^d - 1$  a  $d$  racines distinctes dans  $\mathbb{C}$ , à savoir  $\exp(\frac{2ir\pi}{d}) = \cos(\frac{2ir\pi}{d}) + i \sin(\frac{2ir\pi}{d})$  pour  $r = 0, 1, \dots, d-1$ .  $\square$

### 2.3. Appendice (†) : somme directe externe d'espaces vectoriels

**Définition 2.3.1 (Sommes directes).** — Soient  $V_1, \dots, V_n$  des  $k$ -espaces vectoriels. L'ensemble produit

$$V_1 \times \dots \times V_n = \{(v_1, \dots, v_n) \mid v_i \in V_i\}$$

est muni d'une structure d'espace vectoriel définie « composante par composante, c.-à-d.,  $t \cdot (v_1, \dots, v_n) = (t \cdot v_1, \dots, t \cdot v_n)$  et  $(v_1, \dots, v_n) + (v'_1, \dots, v'_n) = (v_1 + v'_1, \dots, v_n + v'_n)$ .

On l'appelle la *somme directe* (externe) des  $V_i$  et on le note

$$V_1 \oplus \dots \oplus V_n \quad \text{ou} \quad \bigoplus_{i=1}^n V_i.$$

De même, un  $n$ -uplet  $(v_1, \dots, v_n)$  (avec  $v_i \in V_i$ ) est aussi noté  $v_1 + \dots + v_n$  ou  $\sum_{i=1}^n v_i$ , c.-à-d., on identifie l'élément  $v_i \in V_i$  au  $n$ -uplet  $(0, \dots, 0, v_i, 0, \dots, 0)$  (où  $v_i$  est à la  $i$ -ème place).

Supposons que  $\mathcal{B}_1 = (e_1, \dots, e_{d_1})$  soit une base de  $V_1$ , puis  $\mathcal{B}_2 = (e_{d_1+1}, \dots, e_{d_1+d_2})$  une base de  $V_2$ , ... puis  $\mathcal{B}_n = (e_{d_1+\dots+d_{n-1}+1}, \dots, e_{d_1+\dots+d_n})$  une base de  $V_n$ . Alors tout élément  $v$  de  $\bigoplus_{i=1}^n V_i$  s'écrit de façon unique

$$v = \underbrace{t_1 e_1 + \dots + t_{d_1} e_{d_1}}_{v_1} + \dots + \underbrace{t_{d_1+\dots+d_{n-1}+1} e_{d_1+\dots+d_{n-1}+1} + \dots + t_{d_1+\dots+d_n} e_{d_1+\dots+d_n}}_{v_n}$$

avec les  $t_i$  dans  $k$ , d'où l'on déduit que la réunion disjointe des  $\mathcal{B}_i$  est une base de  $\bigoplus_{i=1}^n V_i$ . Par conséquent, on a la formule :

$$(*) \quad \boxed{\dim(\bigoplus_{i=1}^n V_i) = d_1 + \dots + d_n = \sum_{i=1}^n \dim(V_i).}$$

**Remarque 2.3.2.** — Supposons maintenant que  $E_1, \dots, E_n$  soient des sous-espaces d'un  $k$ -espace vectoriel  $V$ . D'une part, on note  $E_1 + \dots + E_n$  le sous-espace de  $V$  engendré par  $E_1 \cup \dots \cup E_n$ ; c'est l'ensemble de toutes les sommes

$$(*) \quad x_1 + \dots + x_n, \quad \text{avec } x_i \in E_i.$$

D'autre part, on peut former, la somme directe externe  $S = E_1 \oplus \dots \oplus E_n$  des  $E_i$ ; ce n'est pas un sous-espace de  $V$ , mais on a une application linéaire naturelle

$$\sigma : E_1 \oplus \dots \oplus E_n \rightarrow V, \quad (x_1, \dots, x_n) \mapsto x_1 + \dots + x_n$$

dont l'image est le sous-espace  $E_1 + \dots + E_n$  de  $V$ , et le noyau est le sous-espace de  $S$  formé des  $n$ -uplets  $(x_1, \dots, x_n)$  tels que  $x_1 + \dots + x_n = 0$ .

On voit donc que  $\text{Ker}(\sigma) = (0)$  si et seulement si les sous-espaces  $E_1, \dots, E_n$  sont en somme directe dans  $V$ , et dans ce cas  $\sigma$  est un isomorphisme de la somme directe externe  $S$  sur le sous-espace de  $V$  noté  $E_1 \oplus \dots \oplus E_n$  en 2.1.1. Ceci justifie l'usage de la notation  $E_1 \oplus \dots \oplus E_n$  dans les deux cas. D'autre part, pour des espaces vectoriels arbitraires  $E_1, \dots, E_n$ , la somme directe « externe »  $E_1 \oplus \dots \oplus E_n$  sera appelée simplement « somme directe ».

## 2.4. Appendice (†) : division euclidienne dans $\mathbb{C}[X]$ et théorème de Bézout

**Théorème 2.4.1 (Division euclidienne dans  $k[X]$ ).** — Soit  $k$  un corps et soit  $P \in k[X]$  un polynôme de degré  $d \geq 1$ . Pour tout  $F \in k[X]$ , il existe un unique couple  $(Q, R)$  d'éléments de  $k[X]$  tel que

$$F = PQ + R, \quad \text{et} \quad R = 0 \text{ ou bien } \deg(R) < \deg(P).$$

On appelle  $Q$  (resp.  $R$ ) le quotient (resp. le reste) de la division euclidienne de  $F$  par  $P$ .

*Démonstration.* — Montrons l'existence de  $(Q, R)$  en procédant par récurrence sur  $\deg(F)$ . Si  $F = 0$  ou si  $\deg(F) < d = \deg(P)$ , on prend  $Q = 0$  et  $R = F$ . Soit  $n \geq d$  et supposons l'existence établie pour les degrés  $< n$ . Soit  $F$  de degré  $n$ . Notons  $a$  le coefficient dominant de  $F$  et  $c$  celui de  $P$ . Alors  $ac^{-1}X^{n-d}P$  est de degré  $n$  et de coefficient dominant  $a$ , donc  $F - ac^{-1}X^{n-d}P$  est de degré  $< n$ . Par hypothèse de récurrence, il existe  $Q, R \in k[X]$  tels que

$$F - ac^{-1}X^{n-d}P = PQ + R, \quad \text{et} \quad R = 0 \text{ ou bien } \deg(R) < \deg(P).$$

Alors  $F = P(Q + ac^{-1}X^{n-d}) + R$ , ce qui prouve le résultat d'existence.

Montrons l'unicité : si  $Q_1, R_1$  vérifie les mêmes conditions, les égalités  $PQ + R = F = PQ_1 + R_1$  donnent

$$P(Q - Q_1) = R_1 - R.$$

Si  $Q - Q_1$  était  $\neq 0$  alors  $P(Q - Q_1)$  serait de degré  $d + \deg(Q - Q_1) \geq d$ . Or,  $R_1 - R$  est nul ou de degré  $< d$ . Donc nécessairement  $Q - Q_1 = 0$ , d'où  $R_1 - R = 0$ , d'où  $Q = Q_1$  et  $R = R_1$ . Ceci prouve l'unicité.  $\square$

**Définitions 2.4.2.** — 1) Soit  $I$  un sous-ensemble de  $k[X]$ . On dit que  $I$  est un idéal de  $k[X]$  si c'est un sous-espace vectoriel et si, pour tout  $P \in I$  et  $S \in k[X]$ , on a  $SP \in I$ .

2) Soient  $P_1, \dots, P_r \in k[X]$ . Alors l'ensemble  $I$  des sommes  $S_1P_1 + \dots + S_rP_r$  est un idéal de  $k[X]$ , et c'est le plus petit idéal de  $k[X]$  contenant  $P_1, \dots, P_r$ . En effet, toute combinaison linéaire de telles sommes est encore une somme de ce type, donc  $I$  est un sous-espace vectoriel de  $k[X]$ ; de plus, pour tout  $S \in k[X]$ ,  $S(S_1P_1 + \dots + S_rP_r) = SS_1P_1 + \dots + SS_rP_r$  est encore une somme de ce type, donc  $I$  est un idéal. Réciproquement, tout idéal  $J$  contenant  $P_1, \dots, P_r$  contient toute les sommes  $S_1P_1 + \dots + S_rP_r$ , donc contient  $I$ .

On dit que  $I$  est l'idéal engendré par  $P_1, \dots, P_r$  et on le note  $(P_1, \dots, P_r)$ .

3) On dit qu'un idéal  $I$  de  $k[X]$  est principal s'il peut être engendré par un seul élément, c.-à-d., s'il existe  $P \in I$  tel que  $I = \{SP \mid S \in k[X]\} = (P)$ .

**Théorème 2.4.3.** — Soit  $k$  un corps. Tout idéal  $I$  de  $k[X]$  est principal. Plus précisément, si  $I$  est un idéal non nul de  $k[X]$ , il existe un unique polynôme unitaire  $P \in I$  tel que  $I = (P)$ .

*Démonstration.* — Si  $I$  est l'idéal nul  $\{0\}$ , il est engendré par le polynôme nul 0. Donc on peut supposer  $I \neq \{0\}$ . Dans ce cas, l'ensemble  $\{\deg(Q) \mid Q \in I - \{0\}\}$  est un sous-ensemble non-vide de  $\mathbb{N}$ , donc admet un plus petit élément  $d$ . Soit  $P \in I - \{0\}$  tel que  $\deg(P) = d$ ; quitte à remplacer  $P$  par  $a^{-1}P$ , où  $a$  est le coefficient dominant de  $P$ , on peut supposer  $P$  unitaire.

Soit  $F$  un élément arbitraire de  $I$ , d'après le théorème 2.4.1, on peut écrire  $F = PQ + R$ , avec  $R = 0$  ou bien  $\deg(R) < \deg(P) = d$ . Comme  $I$  est un idéal, alors  $PQ \in I$  et donc  $R = F - PQ$  appartient à  $I$ . Si on avait  $R \neq 0$ , ce serait un élément non nul de  $I$  de degré  $< d$ , contredisant la minimalité de  $d$ . Donc  $R = 0$  et donc  $F = PQ$ . Il en résulte que  $I = \{PQ \mid Q \in k[X]\} = (P)$ , i.e.  $I$  est principal, engendré par le polynôme unitaire  $P$ . De plus,  $P$  est unique. En effet, si  $P_1$  est un second polynôme unitaire tel que  $I = (P_1)$ , alors il existe  $Q, Q_1 \in k[X]$  tels que  $P_1 = PQ$  et  $P = P_1Q_1$ . Il en résulte que  $Q$  et  $Q_1$  sont de degré zéro, donc des éléments de  $k$ , et comme  $P$  et  $P_1$  sont unitaires, l'égalité  $P_1 = PQ$  entraîne  $Q = 1$  d'où  $P_1 = P$ .  $\square$

**Théorème 2.4.4 (Théorème de Bézout).** — Soient  $P_1, \dots, P_r \in \mathbb{C}[X]$  des polynômes non nuls, sans racine commune. Alors il existe  $S_1, \dots, S_r \in \mathbb{C}[X]$  tels que  $S_1P_1 + \dots + S_rP_r = 1$ .

*Démonstration.* — Soit  $I$  l'idéal de  $\mathbb{C}[X]$  engendré par  $P_1, \dots, P_r$ . D'une part, c'est l'ensemble des sommes  $S_1P_1 + \dots + S_rP_r = 1$ , avec  $S_1, \dots, S_r \in \mathbb{C}[X]$ . D'autre part, on sait que c'est un idéal principal non nul, engendré par un certain polynôme  $D \neq 0$ .

D'une part, comme  $D \in I = (P_1, \dots, P_r)$ , il existe  $T_1, \dots, T_r \in \mathbb{C}[X]$  tels que  $T_1P_1 + \dots + T_rP_r = D$ . D'autre part, comme  $P_i \in I = (D)$ , alors chaque  $P_i$  égale  $DQ_i$ , pour un certain  $Q_i \in k[X]$ . Or,  $\mathbb{C}$  est algébriquement clos (cf. paragraphe suivant), donc si  $D$  était non constant il aurait une racine  $\alpha \in \mathbb{C}$ , et alors  $\alpha$  serait une racine commune à tous les  $P_i$  (puisque  $P_i = DQ_i$ ), contredisant l'hypothèse. Donc  $D$  est un polynôme constant non nul, i.e. un élément  $z \in \mathbb{C}^*$ . Alors remplaçant  $T_i$  par  $S_i = z^{-1}T_i$ , on obtient l'égalité voulue  $S_1P_1 + \dots + S_rP_r = 1$ .  $\square$

**Définition 2.4.5 (Polynôme minimal d'un endomorphisme).** — Soient  $k$  un sous-corps de  $\mathbb{C}$  (par exemple  $k = \mathbb{R}$ ),  $V$  un  $k$ -espace vectoriel de dimension finie et  $u \in \text{End}_k(V)$ . Alors  $I_u = \{Q \in k[X] \mid Q(u) = 0\}$  est un idéal de  $k[X]$ , non nul puisque  $P_u(X) \in I_u$  d'après le théorème de Cayley-Hamilton. Donc, d'après le théorème 2.4.3, il existe un unique polynôme unitaire  $M_u$  tel que  $I_u = (M_u)$ , c.-à-d.,  $M_u$  annule  $u$  et tout polynôme  $Q$  annihilant  $u$  est un multiple de  $M_u$ . On dit que  $M_u$  est le *polynôme minimal* de  $u$ .

**Proposition 2.4.6.** — Soit  $V$  un  $\mathbb{C}$ -espace vectoriel de dimension finie et soit  $u \in \text{End}_{\mathbb{C}}(V)$ . Alors  $u$  est diagonalisable  $\Leftrightarrow M_u$  a des racines simples.

*Démonstration.* —  $\Rightarrow$  Si  $u$  est diagonalisable, alors  $V$  est la somme directe des espaces propres  $V_i = V_{\lambda_i}$ , pour  $i = 1, \dots, r$ . Comme  $u - \lambda_i \text{id}_V$  est nul sur  $V_i$ , alors  $\prod_{i=1}^r (u - \lambda_i \text{id}_V)$  est nul, donc  $M_u$  divise le polynôme  $P = \prod_{i=1}^r (X - \lambda_i \text{id}_V)$  qui n'a que des racines simples, donc il en est de même de  $M_u$ . (En fait, on a  $M_u = P$  car chaque produit partiel  $\prod_{j \neq i} (u - \lambda_j \text{id}_V)$  est non nul sur  $V_i$ .)

$\Leftarrow$  Si  $M_u$  a des racines simples, alors  $u$  est diagonalisable d'après la proposition 2.2.13 (qui utilise le théorème de Bézout).  $\square$

## 2.5. Appendice (†) : $\mathbb{C}$ est algébriquement clos

**Théorème 2.5.1.** —  $\mathbb{C}$  est algébriquement clos, c.-à-d., tout polynôme  $P \in \mathbb{C}[X]$  de degré  $n \geq 1$  se factorise en produit de facteurs de degré 1, i.e.  $P = a(X - \lambda_1) \dots (X - \lambda_n)$ , où  $a$  est le coefficient dominant de  $P$ .

*Démonstration.* — Remarquons d'abord qu'il suffit de montrer l'assertion suivante :

(★) Tout  $P \in \mathbb{C}[X]$  non constant admet une racine dans  $\mathbb{C}$ .

En effet, supposons (★) établie et montrons par récurrence sur  $n$  que tout  $P \in \mathbb{C}[X]$  de degré  $n \geq 1$  se factorise comme indiqué dans le théorème. C'est évident pour  $n = 1$ , donc on peut supposer  $n \geq 2$  et le résultat établi pour  $n - 1$ . Soit  $P$  de degré  $n$  et de coefficient dominant  $a$ . D'après (★),  $P$  a au moins une racine  $\lambda_1$  dans  $\mathbb{C}$ . Faisant la division euclidienne de  $P$  par  $X - \lambda_1$ , on peut écrire

$$P = (X - \lambda_1)P_1 + R, \quad \text{avec } R = 0 \text{ ou bien } \deg(R) < 1.$$

Donc  $R = 0$  ou bien  $R$  est une constante  $c \neq 0$ . Or, évaluant l'égalité ci-dessus en  $X = \lambda_1$ , on trouve  $R(\lambda_1) = P(\lambda_1) = 0$ , donc nécessairement  $R = 0$ . Donc  $P = (X - \lambda_1)P_1$ , avec  $P_1$  non nul, de degré  $n - 1$  et de coefficient dominant  $a$ . Par hypothèse de récurrence,  $P_1$  se factorise en  $P_1 = a(X - \lambda_2) \dots (X - \lambda_n)$ , et donc  $P = (X - \lambda_1)P_1$  égale  $a(X - \lambda_1) \dots (X - \lambda_n)$ . Ceci montre que le théorème découle de l'assertion (★).

Démontrons maintenant l'assertion (★). Soit  $P \in \mathbb{C}[X]$  un polynôme de degré  $n \geq 1$ . Sans perte de généralité, on peut supposer  $P$  unitaire, i.e. de coefficient dominant égal à 1. Écrivons

$$P = X^n + a_1X^{n-1} + \dots + a_n.$$

Raisonnons par l'absurde et supposons que  $P$  ne s'annule pas sur  $\mathbb{C}$ . Alors, en particulier,  $a_n \neq 0$ . Notons  $|\cdot|$  la norme usuelle sur  $\mathbb{C}$ , c.-à-d., si  $z = x + iy$  alors  $|z| = \sqrt{z\bar{z}} = \sqrt{x^2 + y^2}$ .

Comme  $\lim_{|z| \rightarrow +\infty} |P(z)| = +\infty$ , il existe  $R > 0$  tel que

$$(1) \quad |z| \geq R \implies |P(z)| \geq |a_n|.$$

Explicitement, on peut prendre  $R = R_0 = \max\{1, 2na\}$ , où  $a = \max_{i=1}^n |a_i|$ . En effet, pour  $|z| \geq R_0$  et  $d = 1, \dots, n$ , on a  $|z^d| \geq |z| \geq 2na$  d'où

$$\left| \sum_{d=1}^n \frac{a_d}{z^d} \right| \leq \sum_{d=1}^n \frac{|a_d|}{2na} \leq \frac{1}{2}.$$

Comme  $|u + v| \geq |u| - |v|$ , on obtient que, pour  $|z| \geq R_0$ , on a

$$|P(z)| = |z^n| \cdot \left| 1 + \sum_{d=1}^n \frac{a_d}{z^d} \right| \geq 2na \left(1 - \frac{1}{2}\right) = na \geq n|a_n|.$$

Comme le disque  $D$  de centre 0 et de rayon  $R$  est compact, la fonction continue  $f : z \mapsto |P(z)|$  y atteint son minimum  $r_0$ , et  $r_0 > 0$  puisqu'on a supposé que  $P$  ne s'annule pas. Comme de plus

$$(2) \quad \forall z \notin D, \quad f(z) = |P(z)| \geq |a_n| = |P(0)| \geq r_0$$

alors  $r_0$  est le minimum de  $f$  sur  $\mathbb{C}$  tout entier.

Soit  $z_0 \in D$  tel que  $f(z_0) = r_0$ . En remplaçant  $z$  par  $z + z_0$  et  $P(z)$  par  $Q(z) := P(z_0)^{-1}P(z + z_0)$ , on se ramène au cas où  $z_0 = 0$  et où  $Q(0) = 1$  est le minimum de  $g = |Q|$  sur  $\mathbb{C}$ .

Observons que  $Q$  est, comme  $P$ , de degré  $n$ . Notons  $k$  l'ordre d'annulation en 0 de  $Q - 1$ . On peut alors écrire

$$Q(X) = 1 + b_k X^k + \dots + b_n X^n.$$

avec  $b_k$  et  $b_n$  tous deux  $\neq 0$ . Écrivons  $b_k = r e^{i\theta}$ , avec  $r > 0$  et  $\theta \in [0, 2\pi[$  et posons  $z_\varepsilon = \varepsilon e^{i(\pi-\theta)/k}$ , pour tout  $\varepsilon \in \mathbb{R}_+^*$ . Comme  $z_\varepsilon^k = \varepsilon^k e^{i(\pi-\theta)} = -\varepsilon^k e^{-i\theta}$ , alors

$$Q(z_\varepsilon) = 1 - r\varepsilon^k + \varepsilon^k h(\varepsilon), \quad \text{où } h(\varepsilon) = \sum_{j=1}^{n-k} b_{k+j} z_\varepsilon^j.$$

Comme  $\lim_{\varepsilon \rightarrow 0} \varepsilon^k = 0$  et  $\lim_{\varepsilon \rightarrow 0} h(\varepsilon) = 0$ , il existe  $\varepsilon_0 \in ]0, 1[$  tel que

$$\forall \varepsilon \leq \varepsilon_0, \quad r\varepsilon^k < 1 \quad \text{et} \quad |h(\varepsilon)| \leq \frac{r}{2}.$$

On a alors

$$|Q(z_{\varepsilon_0})| = |1 - r\varepsilon_0^k + \varepsilon_0^k h(\varepsilon_0)| \leq |1 - r\varepsilon_0^k| + \frac{r}{2}\varepsilon_0^k = 1 - r\varepsilon_0^k + \frac{r}{2}\varepsilon_0^k = 1 - \frac{r}{2}\varepsilon_0^k < 1.$$


Ceci contredit l'hypothèse que  $1 = Q(0)$  était le minimum de  $g = |Q|$  sur  $\mathbb{C}$ . Cette contradiction montre que l'hypothèse que  $P$  ne s'annule pas sur  $\mathbb{C}$  est impossible. Ceci achève la démonstration du théorème 2.5.1.  $\square$



## CHAPITRE 3

### DÉCOMPOSITION DE JORDAN, EXPONENTIELLES DE MATRICES, ESPACES QUOTIENTS

**Résumé :** Dans les sections 1 et 2 de ce chapitre, on raffine la décomposition en espaces caractéristiques en étudiant la suite des noyaux et en introduisant la forme normale de Jordan (ceci donne lieu à une jolie interprétation « graphique » en termes de *partitions*). Puis, dans les sections 3 et 4, on se place sur  $\mathbb{K} = \mathbb{R}$  ou  $\mathbb{C}$  et on introduit les exponentielles de matrices et leur utilisation pour l'étude des équations différentielles linéaires (à coefficients constants). Enfin, dans la section 5, on introduit la notion d'espace vectoriel quotient (cette section est indépendante des sections 1 à 4 et aurait pu figurer dans le chapitre 2). Cette construction peut sembler difficile à absorber ; elle ne sera guère utilisée dans la suite du cours, mais il est utile de la voir brièvement.

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 un appendice à la fin du chapitre ; ces passages n'interviendront pas dans les évaluations.

#### 3.1. Endomorphismes nilpotents, partitions et formes normales de Jordan

Revenons pour un moment à un corps  $k$  arbitraire.



**Définition 3.1.1 (Endomorphismes nilpotents).** — Soit  $V$  un  $k$ -espace vectoriel. On dit que  $u \in \text{End}_k(V)$  est *nilpotent* s'il existe un entier  $r \geq 1$  tel que  $u^r = 0$ . Dans ce cas, le plus petit entier  $r$  ayant cette propriété s'appelle l'*indice de nilpotence* de  $u$ . Par exemple, si  $A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$  alors  $A^2 = 0$  donc  $A$

est d'indice de nilpotence 2 ; si  $B = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$ , on a  $B^2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$  et  $B^3 = 0$ , donc  $B$  est d'indice de nilpotence 3.

Un exemple très important d'endomorphismes nilpotents est fourni par les matrices triangulaires strictes, i.e. les matrices triangulaires avec des 0 sur la diagonale :



**Proposition 3.1.2 (Matrices triangulaires strictes).** — Soit  $A \in M_n(k)$  une matrice triangulaire stricte. Alors  $A^n = 0$ , donc  $A$  est nilpotente.

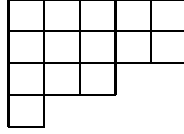
*Démonstration.* — Comme  $A$  est triangulaire stricte, on a  $P_A(X) = (-1)^n X^n$ , d'où aussitôt le résultat si  $k = \mathbb{C}$ , d'après Cayley-Hamilton. Pour un corps arbitraire  $k$ , reprenons la démonstration du théorème de Cayley-Hamilton. Notons  $\mathcal{B} = (e_1, \dots, e_n)$  la base canonique de  $k^n$  et  $F_i$  le sous-espace de  $k^n$  engendré par  $e_1, \dots, e_i$ . Si  $A$  est triangulaire supérieure stricte on a, pour  $i = 1, \dots, n$  :

$$A(F_i) \subseteq F_{i-1},$$

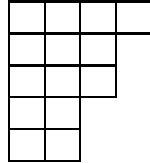
avec la convention  $F_0 = \{0\}$ . Comme  $F_n = k^n$ , on déduit des inclusions ci-dessus que  $A(k^n) \subseteq F_{n-1}$ , puis  $A^2(k^n) \subseteq F_{n-2}$ , etc., d'où finalement  $A^n(k^n) \subseteq F_0 = \{0\}$ , donc  $A^n = 0$ . Enfin, si  $A$  est triangulaire inférieure stricte, en appliquant ce qui précède à  $B = {}^t A$ , on obtient que  $B^n = {}^t(A^n)$  est nulle, d'où aussi  $A^n = 0$ .  $\square$



**Définition 3.1.3 (Partitions d'un entier).** — Soit  $n$  un entier  $\geq 1$ . On appelle *partition* de  $n$  une suite décroissante  $\mathbf{p} = (p_1, \dots, p_r)$  d'entiers  $\geq 1$  tels que  $p_1 + \dots + p_r = n$ . Les  $p_i$  s'appellent les *parts* de la partition,  $r$  le nombre de parts, et  $p_1$  la plus grande part. On peut représenter une telle partition par un diagramme formé de « boîtes », chaque boîte étant un carré de côté 1, la première ligne contenant  $p_1$  boîtes, la seconde  $p_2$  boîtes, etc., toutes les lignes étant alignées à gauche; par exemple, la partition  $\mathbf{p} = (5, 5, 3, 1)$  de  $5 + 5 + 3 + 1 = 14$  est représentée par le diagramme :



On voit alors que l'ensemble des partitions de  $n$  est muni d'une involution  $\mathbf{p} \mapsto \tilde{\mathbf{p}}$ , où le diagramme de  $\tilde{\mathbf{p}}$  est obtenu en « transposant » celui de  $\mathbf{p}$ , i.e. les lignes de  $\tilde{\mathbf{p}}$  sont les colonnes de  $\mathbf{p}$ , et vice-versa (on a donc  $\tilde{\tilde{\mathbf{p}}} = \mathbf{p}$ ). Ainsi, dans l'exemple précédent, le diagramme de  $\tilde{\mathbf{p}}$  est :



i.e.  $\tilde{\mathbf{p}} = (4, 3, 3, 2, 2)$ .

**Proposition 3.1.4 (La suite des noyaux).** — Soient  $V$  un  $k$ -espace vectoriel de dimension  $n$ ,  $u$  un endomorphisme nilpotent de  $V$ , et  $d$  l'indice de nilpotence de  $u$ .

(1) On a une suite croissante :

$$(*) \quad \{0\} = \text{Ker}(\text{id}_V) \subseteq \text{Ker}(u) \subseteq \text{Ker}(u^2) \subseteq \dots \subseteq \text{Ker}(u^d) = V.$$

(2) Pour tout  $i = 1, \dots, d$ , posons  $K_i = \dim \text{Ker}(u^i)$  et  $q_i = K_i - K_{i-1}$ . (On a  $u^0 = \text{id}_V$  et  $K_0 = 0$ .)

Alors on a  $q_{i-1} \geq q_i$  pour  $i = 2, \dots, d$ .

(3) La suite  $(*)$  ci-dessus est strictement croissante, i.e.  $\text{Ker}(u^{i-1}) \neq \text{Ker}(u^i)$  pour tout  $i = 1, \dots, d$ . En particulier, on a  $d \leq n$ .

(4) La suite  $\mathbf{q} = (q_1, \dots, q_d)$  est une partition de  $n$ , et la suite des dimensions des noyaux,  $\mathbf{K} = (K_1, K_2, \dots, K_d)$ , est la suite  $\sigma(\mathbf{q})$  des sommes partielles de  $\mathbf{q}$ .

*Démonstration.* — L'assertion (1) est immédiate : si  $x \in V$  vérifie  $u^i(x) = 0$  alors a fortiori  $u^{i+1}(x) = u(u^i(x)) = 0$ , d'où  $\text{Ker}(u^i) \subseteq \text{Ker}(u^{i+1})$  pour  $i = 0, \dots, d-1$ . De plus,  $\text{Ker}(u^d) = V$  puisque  $u^d = 0$ .

Pour montrer (2), fixons un indice  $i \geq 2$  et posons  $q = q_i$ . Soit  $\mathcal{B} = (e_1, \dots, e_q)$  une base d'un supplémentaire de  $\text{Ker}(u^{i-1})$  dans  $\text{Ker}(u^i)$ . Supposons qu'on ait une égalité

$$(\dagger) \quad x + t_1 u(e_1) + \dots + t_q u(e_q) = 0, \quad \text{avec } x \in \text{Ker}(u^{i-2}), \quad t_1, \dots, t_q \in k$$

alors, appliquant  $u^{i-2}$  à cette égalité, on obtient  $u^{i-1}(t_1 e_1 + \dots + t_q e_q) = 0$ , d'où  $t_1 e_1 + \dots + t_q e_q \in \text{Ker}(u^{i-1})$ . Comme  $\mathcal{B}$  est une base d'un supplémentaire de  $\text{Ker}(u^{i-1})$  dans  $\text{Ker}(u^i)$ , ceci entraîne  $t_1 = 0 = \dots = t_q$ , puis reportant ceci dans  $(\dagger)$  on obtient aussi que  $x = 0$ . Ceci montre que les sous-espaces  $\text{Ker}(u^{i-2})$  et  $ku(e_1), \dots, ku(e_q)$  de  $\text{Ker}(u^{i-1})$  sont en somme directe, d'où

$$K_{i-2} + q = \dim \left( \text{Ker}(u^{i-2}) \oplus \underbrace{ku(e_1) \oplus \dots \oplus ku(e_q)}_{=\text{Vect}(u(\mathcal{B}))} \right) \leq K_{i-1}$$

donc  $q_i = q \leq K_{i-1} - K_{i-2} = q_{i-1}$ , ce qui prouve l'assertion (2). Signalons aussi le point suivant, dont on aura besoin plus loin : si  $\mathcal{B}' = (f_1, \dots, f_t)$  est une base d'un supplémentaire de  $\text{Ker}(u^{i-2}) \oplus \text{Vect}(u(\mathcal{B}))$  dans  $\text{Ker}(u^{i-1})$ , alors  $t = K_{i-1} - (K_{i-2} + q_i) = q_{i-1} - q_i$ , et  $u(\mathcal{B}) \cup \mathcal{B}'$  est une base d'un supplémentaire de  $\text{Ker}(u^{i-2})$  dans  $\text{Ker}(u^{i-1})$ . On a donc obtenu le résultat suivant :

(\*)  $\left\{ \begin{array}{l} \text{si } \mathcal{B}_i \text{ est une base d'un supplémentaire de } \text{Ker}(u^{i-1}) \text{ dans } \text{Ker}(u^i), \text{ alors la famille } u(\mathcal{B}_i) \text{ est} \\ \text{libre et peut se compléter en une base d'un supplémentaire de } \text{Ker}(u^{i-2}) \text{ dans } \text{Ker}(u^{i-1}). \end{array} \right.$

L'assertion (3) en découle. En effet, si pour un certain  $i \leq d$  on avait  $\text{Ker}(u^{i-1}) = \text{Ker}(u^i)$ , i.e.  $q_i = 0$ , on aurait  $q_{i+1} = 0 = \dots = q_d$ , d'où  $\text{Ker}(u^{i-1}) = \text{Ker}(u^d) = V$ , donc  $u^{i-1} = 0$ , contredisant le fait que l'indice de nilpotence est  $d$ . Donc la suite  $(*)$  est strictement croissante. Comme la dimension croît d'au moins 1 à chaque cran, on a donc  $\dim \text{Ker}(u^i) \geq i$  pour  $i = 1, \dots, d$ , et comme d'autre part  $\dim V = n$ , on conclut que  $d \leq n$ .



Prouvons l'assertion (4). Par définition des  $q_i$ , on a :

$$K_1 = q_1, \quad K_2 = q_2 + K_1 = q_2 + q_1, \quad \dots \quad K_i = q_i + K_{i-1} = q_i + \dots + q_1, \quad \dots$$

$$K_d = q_d + K_{d-1} = q_d + \dots + q_1.$$

D'autre part, d'après (2), on a  $q_i \geq q_{i+1}$ . Il en résulte que  $\mathbf{q} = (q_1, \dots, q_d)$  est une partition de  $K_d = n$ , et que la suite  $\mathbf{K} = (K_1, K_2, \dots, K_d)$  est la suite  $\sigma(\mathbf{q})$  des sommes partielles de  $\mathbf{q}$ .  $\square$

### Définitions 3.1.5 (Blocs de Jordan nilpotents et matrices de Jordan nilpotentes)

1) Pour tout entier  $n \geq 1$ , on appelle « bloc de Jordan nilpotent » de taille  $n$ , et l'on note  $J_n$ , la matrice carrée de taille  $n$  ayant des 1 sur la diagonale juste au-dessus de la diagonale principale et des 0 partout ailleurs. Ainsi,  $J_1$  est la matrice nulle (0), et

$$J_2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad J_3 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \quad J_4 = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad \text{etc.}$$

Notant  $\mathcal{B} = (e_1, \dots, e_n)$  la base canonique de  $k^n$ , on a donc  $J_n e_i = e_{i-1}$  pour  $i = n, \dots, 2$  et  $J_n e_1 = 0$ . Il en résulte que  $\mathcal{B}$  égale  $(u^{n-1}(e_n), \dots, u(e_n), e_n)$  et que  $\text{Ker}(J_n)$  est la droite  $ke_1 = ku^{n-1}(e_n)$ .

2) On appelle « matrice de Jordan nilpotente » de taille  $n$  toute matrice  $A \in M_n(k)$  diagonale par blocs, dont les blocs diagonaux  $A_1, \dots, A_d$  sont des blocs de Jordan nilpotents  $J_{p_1}, \dots, J_{p_d}$ , rangés par ordre de taille décroissante, i.e. toute matrice  $A \in M_n(k)$  de la forme suivante :

$$A = \left( \begin{array}{c|c|c|c} J_{p_1} & 0 & \cdots & 0 \\ \hline 0 & J_{p_2} & \ddots & \vdots \\ \hline \vdots & \ddots & \ddots & 0 \\ \hline 0 & \cdots & 0 & J_{p_d} \end{array} \right) \quad \text{où } p_1 \geq \dots \geq p_d \geq 1.$$

Comme la somme des tailles des blocs diagonaux égale  $n$ , on a  $p_1 + \dots + p_d = n$ , i.e.  $\mathbf{p} = (p_1, \dots, p_d)$  est une partition de  $n$ , et l'on notera  $J_{\mathbf{p}}$  la matrice ci-dessus. On voit ainsi que les matrices de Jordan nilpotentes de taille  $n$  sont en bijection avec l'ensemble  $\mathcal{P}(n)$  des partitions de  $n$ .

### Théorème et définition 3.1.6 (Forme normale de Jordan d'un endomorphisme nilpotent)

Soient  $V$  un  $k$ -espace vectoriel de dimension  $n$ , et  $u$  un endomorphisme nilpotent de  $V$ , d'indice de nilpotence  $d$ .

(1) Il existe une base  $\mathcal{B}$  de  $V$  telle que  $\text{Mat}_{\mathcal{B}}(u)$  soit une matrice de Jordan nilpotente  $J_{\mathbf{p}}$ , pour une certaine partition  $\mathbf{p}$  de  $n$ .

(2) De plus,  $\mathbf{p}$  est uniquement déterminée par  $u$  : en effet,  $\tilde{\mathbf{p}}$  est la partition  $\mathbf{q}$  associée à la suite des noyaux  $\text{Ker}(u) \subseteq \dots \subseteq \text{Ker}(u^d)$ . En particulier,  $d$  est la plus grande part  $p_1$  de  $\mathbf{p}$ , et le nombre de parts de  $\mathbf{p}$  est  $q_1 = \dim \text{Ker}(u)$ .

(3) En termes matriciels : toute matrice nilpotente  $A \in M_n(k)$  est semblable à une unique matrice de Jordan nilpotente de taille  $n$ , et donc l'ensemble des classes de similitude de matrices nilpotentes  $A \in M_n(k)$  est en bijection avec l'ensemble  $\mathcal{P}(n)$  des partitions de  $n$ .

(4) On dit que la matrice  $J_{\mathbf{p}}$  obtenue est la **forme normale de Jordan** ou la **réduction de Jordan** de  $u$  (ou de  $A$ ).

*Démonstration.* — Soit  $\mathbf{q} = (q_1, \dots, q_d)$  la partition  $\mathbf{q}$  associée à la suite des noyaux de  $u$ , i.e.  $q_i = K_i - K_{i-1}$ , où  $K_i = \dim \text{Ker}(u^i)$ , pour  $i = 1, \dots, d$  (et  $K_0 = 0$ ). Montrons l'existence d'une base  $\mathcal{B}$  de  $V$  telle que  $\text{Mat}_{\mathcal{B}} = J_{\tilde{\mathbf{q}}}$ . Rappelons le résultat suivant, qu'on a obtenu dans la démonstration du point (2) de 3.1.4 :

(\*)  $\left\{ \begin{array}{l} \text{si } \mathcal{C}_i \text{ est une base d'un supplémentaire de } \text{Ker}(u^{i-1}) \text{ dans } \text{Ker}(u^i), \text{ alors la famille } u(\mathcal{C}_i) \text{ est} \\ \text{libre et peut se compléter en une base d'un supplémentaire de } \text{Ker}(u^{i-2}) \text{ dans } \text{Ker}(u^{i-1}) \end{array} \right.$

Soit  $\mathcal{B}_d = (e_1, \dots, e_{q_d})$  une base d'un supplémentaire de  $\text{Ker}(u^{d-1})$  dans  $V = \text{Ker}(u^d)$ . D'après (\*), la famille  $u(\mathcal{B}_d)$  est libre et se complète, par ajout d'une famille libre  $\mathcal{B}_{d-1} = (e_{q_d+1}, \dots, e_{q_d-1})$ , en une base  $\mathcal{L}_{d-1} = u(\mathcal{B}_d) \cup \mathcal{B}_{d-1}$  d'un supplémentaire de  $\text{Ker}(u^{d-2})$  dans  $\text{Ker}(u^{d-1})$ . (Et par conséquent,  $\mathcal{B}_d \cup \mathcal{L}_{d-1}$  est une base d'un supplémentaire de  $\text{Ker}(u^{d-2})$  dans  $V$ .)

Pour aider le lecteur, indiquons que la démonstration peut se « visualiser » comme suit : on dessine le diagramme de la partition  $\mathbf{q}$ , disons pour  $\mathbf{q} = (5, 5, 3, 1)$ , et l'on écrit les vecteurs  $e_1, \dots, e_{q_d}$  dans les



dernières cases des colonnes 1 à  $q_d$  (i.e. dans la ligne  $d$ ), puis on écrit au-dessus les vecteurs  $u(e_1), \dots, u(e_{q_d})$  et l'on complète la ligne  $d-1$  en écrivant les vecteurs  $e_{q_d+1}, \dots, e_{q_d-1}$ , dans les dernières cases des colonnes  $q_d+1$  à  $q_{d-1}$  :

base d'un supplém. de $\text{Ker}(u^{d-2})$ dans $\text{Ker}(u^{d-1})$ :	$u(e_1)$	$e_2$	$e_3$	
base d'un supplém. de $\text{Ker}(u^{d-1})$ dans $\text{Ker}(u^d)$ :	$e_1$			

D'après ce qui précède, les vecteurs qu'on a écrits forment une base d'un supplémentaire de  $\text{Ker}(u^{d-2})$  dans  $V$ .

Puis, d'après  $(\star)$  à nouveau, il existe une famille libre  $\mathcal{B}_{d-2} = (e_{q_{d-1}+1}, \dots, e_{q_{d-2}})$  dans  $\text{Ker}(u^{d-2})$  telle que

$$\mathcal{L}_{d-2} = u(\mathcal{L}_{d-1}) \cup \mathcal{B}_{d-2} = u^2(\mathcal{B}_d) \cup u(\mathcal{B}_{d-1}) \cup \mathcal{B}_{d-2}$$

soit une base d'un supplémentaire de  $\text{Ker}(u^{d-3})$  dans  $\text{Ker}(u^{d-2})$ ; alors  $\mathcal{B}_d \cup \mathcal{L}_{d-1} \cup \mathcal{L}_{d-2}$  est une base d'un supplémentaire de  $\text{Ker}(u^{d-3})$  dans  $V$ . Dans l'exemple précédent, ceci donne :

base d'un supplém. de $\text{Ker}(u)$ dans $\text{Ker}(u^2)$ :	$u^2(e_1)$	$u(e_2)$	$u(e_3)$	$e_4$	$e_5$
base d'un supplém. de $\text{Ker}(u^2)$ dans $\text{Ker}(u^3)$ :	$u(e_1)$	$e_2$	$e_3$		
base d'un supplém. de $\text{Ker}(u^3)$ dans $\text{Ker}(u^4)$ :	$e_1$				

Ainsi, après  $p \leq d$  étapes, on a rempli les  $p$  lignes du bas du diagramme par des vecteurs qui forment une base d'un supplémentaire de  $\text{Ker}(u^{d-p})$  dans  $V$  (plus précisément, les vecteurs de la  $i$ -ème ligne en partant du bas forment une base d'un supplémentaire de  $\text{Ker}(u^{d-i})$  dans  $\text{Ker}(u^{d-i+1})$ ). Continuant ainsi, on remplit le diagramme de la partition  $\mathbf{q}$  par des vecteurs qui forment une base  $\mathcal{L}$  de  $V$  :

base de $\text{Ker}(u)$ :	$u^3(e_1)$	$u^2(e_2)$	$u^2(e_3)$	$u(e_4)$	$u(e_5)$
base d'un supplém. de $\text{Ker}(u)$ dans $\text{Ker}(u^2)$ :	$u^2(e_1)$	$u(e_2)$	$u(e_3)$	$e_4$	$e_5$
base d'un supplém. de $\text{Ker}(u^2)$ dans $\text{Ker}(u^3)$ :	$u(e_1)$	$e_2$	$e_3$		
base d'un supplém. de $\text{Ker}(u^3)$ dans $\text{Ker}(u^4)$ :	$e_1$				

(Noter que les vecteurs  $e_1, \dots, e_{q_1}$  qui ont été introduits à chaque étape du processus sont situés dans la dernière case de chaque colonne.)

Lisons maintenant le diagramme colonne par colonne, de haut en bas : la 1ère colonne est formée des vecteurs  $v_1 = u^{d-1}(e_1), v_2 = u^{d-2}(e_1), \dots, v_d = e_1$ ; ils forment une base  $\mathcal{E}_1$  d'un sous-espace  $E_1$  stable par  $u$ , dans laquelle la restriction  $u_{E_1}$  de  $u$  à  $E_1$  a pour matrice le bloc de Jordan nilpotent  $J_d$ . De même, pour chaque  $j = 1, \dots, q_d$ , la  $j$ -ème colonne est formée des vecteurs  $u^{d-i}(e_j)$ , pour  $i = 1, \dots, d$ , qui forment une base  $\mathcal{E}_j$  d'un sous-espace  $E_j$  de dimension  $d$ , telle que  $\text{Mat}_{\mathcal{E}_j}(u_{E_j}) = J_d$ . Ensuite, pour  $j$  variant de  $q_d+1$  à  $q_{d-1}$ , la  $j$ -ème colonne est formée des vecteurs  $u^{d-1-i}(e_j)$ , pour  $i = 1, \dots, d-1$ , qui forment une base  $\mathcal{E}_j$  d'un sous-espace  $E_j$  stable par  $u$ , de dimension  $d-1$ , telle que  $\text{Mat}_{\mathcal{E}_j}(u_{E_j}) = J_{d-1}$ , etc.

Donc chaque *colonne*  $C_j$  du diagramme correspond à une base  $\mathcal{E}_j$  d'un sous-espace  $E_j$  stable par  $u$ , de dimension la *hauteur*  $h_j$  de la colonne  $C_j$ , telle que  $\text{Mat}_{\mathcal{E}_j}(u_{E_j})$  soit le bloc de Jordan nilpotent  $J_{h_j}$ . On voit donc que les blocs de Jordan nilpotents qui apparaissent correspondent aux colonnes du diagramme de  $\mathbf{q}$ , c.-à-d., aux lignes du diagramme de  $\tilde{\mathbf{q}}$ . On obtient donc que dans la base  $\mathcal{C} = \mathcal{E}_1 \cup \dots \cup \mathcal{E}_{q_1}$  (ses éléments sont les mêmes que ceux de  $\mathcal{L}$ , mais dans un ordre différent), la matrice de  $u$  est la matrice de Jordan nilpotente  $J_{\tilde{\mathbf{q}}}$ . Par exemple, pour le diagramme précédent, on obtient la matrice

$$J_{(4,3,3,2,2)} = \left( \begin{array}{c|c|c|c|c} J_4 & 0 & 0 & 0 & 0 \\ \hline 0 & J_3 & 0 & 0 & 0 \\ \hline 0 & 0 & J_3 & 0 & 0 \\ \hline 0 & 0 & 0 & J_2 & 0 \\ \hline 0 & 0 & 0 & 0 & J_2 \end{array} \right)$$

Il reste à montrer l'assertion d'unicité : supposons que dans une certaine base  $\mathcal{B} = (v_1, \dots, v_n)$  de  $V$ ,  $u$  ait pour matrice la matrice de Jordan nilpotente  $J_{\mathbf{p}}$ , pour une certaine partition  $\mathbf{p} = (p_1, \dots, p_s)$ . Considérons alors le diagramme de  $\mathbf{p}$  et écrivons les vecteurs  $v_1, \dots, v_n$  de la base  $\mathcal{B}$  dans ce diagramme, de gauche à droite dans chaque ligne, en commençant par la ligne du haut. Alors, par définition de la

matrice  $J_{\mathbf{p}}$ , si l'on note  $e_j$  le vecteur qui est dans la dernière case de la  $j$ -ème ligne en partant du haut, alors les vecteurs de cette ligne, lus de gauche à droite, sont :

$$(*_j) \quad u^{p_j-1}(e_j), \quad u^{p_j-2}(e_j), \quad \dots \quad u(e_j), \quad e_j.$$

Donc, en transposant ce diagramme, on obtient le diagramme de  $\tilde{\mathbf{p}}$ , contenant dans sa  $j$ -ème colonne les vecteurs  $(*_j)$  précédents :

$u^{p_1-1}(e_1)$	$u^{p_2-1}(e_2)$	$\dots$	$u^{p_s-1}(e_s)$
$u^{p_1-2}(e_1)$	$u^{p_2-2}(e_2)$	$\dots$	$u^{p_s-2}(e_s)$
$\vdots$	$\vdots$	$\dots$	$\vdots$
$\vdots$	$\vdots$	$\dots$	$e_s$
$\vdots$	$\vdots$	$\dots$	
$u(e_1)$	$e_2$		
$e_1$			

Comme ces vecteurs forment une base de  $V$ , on voit facilement que les vecteurs de la ligne 1 du haut forment une base de  $\text{Ker}(u)$ , ceux des lignes 1 et 2 du haut forment une base de  $\text{Ker}(u^2)$ , etc. On obtient donc que  $\tilde{\mathbf{p}}$  est la partition associée à la suite des noyaux de  $u$ , d'où  $\tilde{\mathbf{p}} = \mathbf{q}$  et donc  $\mathbf{p} = \tilde{\mathbf{q}}$ . Ceci achève la démonstration des assertions (1) et (2) du théorème 3.1.6. Enfin, l'assertion (3) est une conséquence immédiate des assertions (1) et (2). Le théorème est démontré.  $\square$

**Remarque 3.1.7.** — Dans la démonstration précédente, on a utilisé de façon répétée l'existence d'une base d'un supplémentaire de  $\text{Ker}(u^{i-1})$  dans  $\text{Ker}(u^i)$ . Pour construire explicitement une base de Jordan, on peut utiliser la méthode de réduction des colonnes pour construire successivement une base de  $\text{Ker}(u)$ , puis d'un supplémentaire de  $\text{Ker}(u)$  dans  $\text{Ker}(u^2)$ , puis d'un supplémentaire de  $\text{Ker}(u^2)$  dans  $\text{Ker}(u^3)$ , etc. Voir le paragraphe 3.1.14 plus bas.

**Exercice 3.1.7.1.** — Quel est le nombre de classes de similitude de matrices nilpotentes dans  $M_6(k)$ ? Donner un représentant de chaque classe.

Revenons maintenant au cas où  $k = \mathbb{C}$  et combinons les résultats obtenus dans les théorèmes 2.2.11 et 3.1.6. Commençons par la définition suivante :

**Définitions 3.1.8 (Blocs de Jordan et matrices de Jordan).** — (1) Pour  $\lambda \in \mathbb{C}$  et  $n \in \mathbb{N}^*$ , on appelle « bloc de Jordan » de taille  $n$  associé à  $\lambda$ , et l'on note  $J_n(\lambda)$ , la matrice carrée de taille  $n$  égale à  $\lambda I_n + J_n$ . (Donc  $J_n(0) = J_n$ .) Ainsi  $J_1(\lambda) = (\lambda)$ , et

$$J_2(\lambda) = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}, \quad J_3(\lambda) = \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix}, \quad \text{etc.}$$

(2) Pour  $\lambda \in \mathbb{C}$ ,  $n \in \mathbb{N}^*$ , et  $\mathbf{p}$  une partition de  $n$ , on note  $J_{\mathbf{p}}(\lambda) = \lambda I_n + J_{\mathbf{p}}(\lambda)$  et on dit que c'est une « matrice de Jordan » de taille  $n$  associé à  $\lambda$ .

(3) Enfin, étant donnés des  $r$ -uplets  $(\lambda_1, \dots, \lambda_r) \in \mathbb{C}^r$  et  $(m_1, \dots, m_r) \in (\mathbb{N}^*)^r$ , ainsi qu'une partition  $\mathbf{p}_i$  de chaque  $m_i$ , on note  $J_{\mathbf{p}_1, \dots, \mathbf{p}_r}(\lambda_1, \dots, \lambda_r)$  la matrice carrée de taille  $n = m_1 + \dots + m_r$ , diagonale par blocs, dont les blocs diagonaux sont  $J_{\mathbf{p}_1}(\lambda_1), \dots, J_{\mathbf{p}_r}(\lambda_r)$ . On dira que c'est une « matrice de Jordan associée au  $r$ -uplet  $((\lambda_1, m_1), \dots, (\lambda_r, m_r))$  ».

**Remarque 3.1.9.** — Soit  $J_{\mathbf{p}_1, \dots, \mathbf{p}_r}(\lambda_1, \dots, \lambda_r)$  comme ci-dessus; pour  $i = 1, \dots, r$ , notons  $\mathcal{C}_i$  l'ensemble des vecteurs de la base canonique de  $\mathbb{C}^n$  qui correspondent au bloc diagonal  $J_{\mathbf{p}_i}(\lambda_i)$ , alors chaque  $\mathcal{C}_i$  est une base d'un sous-espace  $W_i$  de  $V$  stable par  $u$ , telle que  $\text{Mat}_{\mathcal{C}_i}(u|_{W_i}) = J_{\mathbf{p}_i}(\lambda_i)$ , et  $J_{\mathbf{p}_1, \dots, \mathbf{p}_r}(\lambda_1, \dots, \lambda_r)$  est la matrice de  $u$  dans la base  $\mathcal{C} = \mathcal{C}_1 \cup \dots \cup \mathcal{C}_r$ .

Remarquons tout de suite que si l'on permute les  $\mathcal{C}_i$  entre elles, i.e. si  $\sigma$  est une permutation de  $\{1, \dots, r\}$  <sup>(1)</sup> et qu'on considère la base  $\mathcal{C}_\sigma = \mathcal{C}_{\sigma(1)} \cup \dots \cup \mathcal{C}_{\sigma(r)}$ , alors la matrice de  $u$  dans la base  $\mathcal{C}_\sigma$  est  $J_{\mathbf{p}_{\sigma(1)}, \dots, \mathbf{p}_{\sigma(r)}}(\lambda_{\sigma(1)}, \dots, \lambda_{\sigma(r)})$ . Pour toute permutation  $\sigma$  de  $\{1, \dots, r\}$ , les matrices  $J_{\mathbf{p}_1, \dots, \mathbf{p}_r}(\lambda_1, \dots, \lambda_r)$  et  $J_{\mathbf{p}_{\sigma(1)}, \dots, \mathbf{p}_{\sigma(r)}}(\lambda_{\sigma(1)}, \dots, \lambda_{\sigma(r)})$  sont donc semblables.

<sup>(1)</sup>c.-à-d., une bijection de l'ensemble  $\{1, \dots, r\}$  dans lui-même



**Théorème 3.1.10 (Forme normale de Jordan d'un endomorphisme de  $\mathbb{C}^n$ )**

Soient  $V$  un  $\mathbb{C}$ -espace vectoriel de dimension  $n$ ,  $u \in \text{End}_{\mathbb{C}}(V)$ ,  $\lambda_1, \dots, \lambda_r$  les valeurs propres, deux à deux distinctes, de  $u$  et pour tout  $i$ , soient  $u_i$  la restriction de  $u$  à l'espace caractéristique  $V_{(\lambda_i)}$  et  $\mathbf{q}_i$  la partition de  $m_i = \dim V_{(\lambda_i)}$  associée à la suite des noyaux de  $u_i$ .

(1) Il existe une base  $\mathcal{B}$  de  $V$  telle que  $\text{Mat}_{\mathcal{B}}(u)$  soit la matrice de Jordan  $J_{\mathbf{p}_1, \dots, \mathbf{p}_r}(\lambda_1, \dots, \lambda_r)$ , où  $\mathbf{p}_i = \tilde{\mathbf{q}}_i$ .

(2) Le  $r$ -uplet  $((\lambda_1, \mathbf{p}_1), \dots, (\lambda_r, \mathbf{p}_r))$  est uniquement déterminé, à l'ordre près, c.-à-d., si pour une base  $\mathcal{B}'$  on a  $\text{Mat}_{\mathcal{B}'}(u) = J_{\mathbf{p}'_1, \dots, \mathbf{p}'_s}(\mu_1, \dots, \mu_s)$ , avec  $\mu_i \neq \mu_j$  si  $i \neq j$ , alors  $s = r$  et il existe une permutation  $\sigma$  de  $\{1, \dots, r\}$  telle que  $\mu_i = \lambda_{\sigma(i)}$  et  $\mathbf{p}'_i = \mathbf{p}_{\sigma(i)}$ .

(3) En termes matriciels : toute  $A \in M_n(\mathbb{C})$  est semblable à une matrice de Jordan  $J_{\mathbf{p}_1, \dots, \mathbf{p}_r}(\lambda_1, \dots, \lambda_r)$ , qui est unique à permutation près des blocs diagonaux entre eux.

(4) On dit que la matrice  $J_{\mathbf{p}_1, \dots, \mathbf{p}_r}(\lambda_1, \dots, \lambda_r)$  est la **forme normale de Jordan** ou la **réduction de Jordan** de  $u$  (ou de  $A$ ).

*Démonstration.* — L'existence découle immédiatement des théorèmes 2.2.11 et 3.1.6. En effet, pour tout  $i$ , notons  $N_i = V_{(\lambda_i)}$ . Comme chaque  $u_i - \lambda_i \text{id}_{N_i}$  est nilpotent, alors, d'après le théorème 3.1.6, il existe une base  $\mathcal{B}_i$  de  $N_i$  telle que  $\text{Mat}_{\mathcal{B}_i}(u_i - \lambda_i \text{id}_{N_i}) = J_{\mathbf{p}_i}$ , où  $\mathbf{p}_i = \tilde{\mathbf{q}}_i$ , d'où  $\text{Mat}_{\mathcal{B}_i}(u_i) = J_{\mathbf{p}_i} + \lambda_i I_{m_i} = J_{\mathbf{p}_i}(\lambda_i)$ .

D'autre part, d'après le théorème 2.2.11, on a

$$(*) \quad V = N_1 \oplus \dots \oplus N_r \quad \text{et} \quad P_u(X) = (-1)^n \prod_{i=1}^r (X - \lambda_i)^{m_i}.$$

Donc  $\mathcal{B} = \mathcal{B}_1 \cup \dots \cup \mathcal{B}_r$  est une base de  $V$ , et dans cette base la matrice de  $u$  est  $J_{\mathbf{p}_1, \dots, \mathbf{p}_r}(\lambda_1, \dots, \lambda_r)$ . Ceci prouve l'existence.

Montrons l'unicité. Supposons que pour une base  $\mathcal{B}'$  on ait  $\text{Mat}_{\mathcal{B}'}(u) = J_{\mathbf{p}'_1, \dots, \mathbf{p}'_s}(\mu_1, \dots, \mu_s)$ , avec  $\mu_i \neq \mu_j$  si  $i \neq j$ . Comme cette matrice est triangulaire supérieure, on obtient que

$$P_u(X) = (-1)^n \prod_{i=1}^s (X - \mu_i)^{d_i}$$

où chaque  $d_i$  est la somme des parts de  $\mathbf{p}'_i$  (i.e.  $\mathbf{p}'_i$  est une partition de  $d_i$ ). Comme  $\mu_i \neq \mu_j$  si  $i \neq j$  alors,  $\mu_1, \dots, \mu_s$  sont les racines, deux à deux distinctes, de  $P_u$  et donc, comparant avec (\*) plus haut, on obtient déjà que  $s = r$  et qu'il existe une permutation  $\sigma$  de  $\{1, \dots, r\}$  telle que  $\mu_i = \lambda_{\sigma(i)}$  et  $d_i = m_{\sigma(i)}$  pour tout  $i = 1, \dots, r$ .

Reste à montrer que  $\mathbf{p}'_i = \mathbf{p}_{\sigma(i)}$ . En renumérotant les  $\mu_i$ , on se ramène d'abord au cas où  $\sigma = \text{id}$  (i.e.  $\mu_i = \lambda_i$  et  $d_i = m_i$  pour tout  $i$ ), et il faut montrer que  $\mathbf{p}'_i = \mathbf{p}_i$ . Pour tout  $i$ , notons  $\mathcal{B}'_i$  l'ensemble des vecteurs de la base  $\mathcal{B}'$  qui correspondent au bloc diagonal  $J_{\mathbf{p}'_i}(\lambda_i)$ , alors chaque  $\mathcal{B}'_i$  est une base d'un sous-espace  $W_i$  de dimension  $m_i$  de  $V$  stable par  $u$ , et l'on a  $\text{Mat}_{\mathcal{B}'_i}(u_{W_i}) = J_{\mathbf{p}'_i}(\lambda_i)$ . Comme  $J_{\mathbf{p}'_i}$  est triangulaire stricte, de taille  $m_i$ , on a  $0 = J_{\mathbf{p}'_i}^{m_i} = (J_{\mathbf{p}'_i}(\lambda_i) - \lambda_i \text{id}_{W_i})^{m_i}$  et donc  $(u_{W_i} - \lambda_i \text{id}_{W_i})^{m_i} = 0$ . Ceci montre que  $W_i$  est contenu dans  $\text{Ker}((u - \lambda_i)^{m_i}) = V_{(\lambda_i)} = N_i$ , et comme tous deux sont de dimension  $m_i$ , on obtient que  $W_i = N_i$  pour tout  $i$ .

Donc,  $\mathcal{B}'_i$  et  $\mathcal{B}_i$  sont deux bases du même espace,  $N_i$ , telles que la matrice de l'endomorphisme nilpotent  $u_i - \lambda_i \text{id}_{N_i}$  de  $N_i$  dans  $\mathcal{B}'_i$  (resp.  $\mathcal{B}_i$ ) est  $J_{\mathbf{p}'_i}$  (resp.  $J_{\mathbf{p}_i}$ ). D'après le résultat d'unicité dans le cas nilpotent, on conclut que  $\mathbf{p}'_i = \mathbf{p}_i$ . Ceci achève la preuve des assertions (1) et (2) du théorème 3.1.10. Enfin, l'assertion (3) est une conséquence immédiate des assertions (1) et (2). Le théorème est démontré.  $\square$



**Corollaire 3.1.11.** — Soient  $A, A' \in M_n(\mathbb{C})$ . Les conditions suivantes sont équivalentes :

- (1)  $A$  et  $A'$  sont semblables (i.e. sont dans la même classe de similitude).
- (2)  $A$  et  $A'$  ont même polynôme caractéristique, disons  $\prod_{i=1}^r (\lambda_i - X)^{m_i}$  avec  $\lambda_i \neq \lambda_j$  si  $i \neq j$ , et pour tout  $i = 1, \dots, r$ ,  $A - \lambda_i I_n$  et  $A' - \lambda_i I_n$  ont la même partition associée à la suite des noyaux.
- (3)  $A$  et  $A'$  ont la même forme normale de Jordan.



**Remarque 3.1.12.** — **En résumé**, pour déterminer la forme normale de Jordan d'une matrice  $A \in M_n(\mathbb{C})$ , on calcule d'abord son polynôme caractéristique  $P_A(X) = (-1)^n \prod_{i=1}^r (X - \lambda_i)^{m_i}$  puis, pour chaque  $i$ , on calcule la dimension  $K_{i,s} = \dim \text{Ker}((A - \lambda_i I_n)^s)$  (en déterminant le rang de  $(A - \lambda_i I_n)^s$ ), pour  $s = 1, 2, \dots$ , en s'arrêtant au cran  $t$  tel que  $K_{i,t} = m_i$ . Posant  $q_{i,s} = K_{i,s} - K_{i,s-1}$ , on obtient la partition  $\mathbf{q}_i = (q_{i,1}, \dots, q_{i,t})$  associée à la suite des noyaux de  $A - \lambda_i I_n$ ; notant  $\mathbf{p}_i$  sa transposée  $\tilde{\mathbf{q}}_i$ , on obtient alors la matrice de Jordan  $J_{\mathbf{p}_i}(\lambda_i)$ , dont le nombre de blocs de Jordan est  $q_{i,1} = K_{i,1} = \dim \text{Ker}(A - \lambda_i I_n)$ . En particulier, si  $K_{i,1} = 1$ , alors  $J_{\mathbf{p}_i}(\lambda_i)$  est le bloc de Jordan  $J_{m_i}(\lambda_i)$ .

**Exercice 3.1.13.** — Soient  $\lambda_1, \dots, \lambda_r \in \mathbb{C}$ , deux à deux distincts,  $m_1, \dots, m_r$  des entiers  $\geq 1$ , et soient  $n = m_1 + \dots + m_r$  et  $P$  le polynôme  $(-1)^n \prod_{i=1}^r (X - \lambda_i)^{m_i}$ . Alors, d'après le théorème précédent, le nombre de classes de similitude de matrices  $A \in M_n(\mathbb{C})$  dont le polynôme caractéristique est égal à  $P$ , est égal à  $p(m_1) \cdots p(m_r)$ , où  $p(m_i)$  désigne le nombre de partitions de  $m_i$ .

(1) Quel est le nombre de classes de similitudes de matrices dans  $M_7(\mathbb{C})$  dont le polynôme caractéristique est  $-X^7 + 3X^6 - 3X^5 + X^4$ ?

(2) À quelle condition sur  $P = (-1)^n \prod_{i=1}^r (X - \lambda_i)^{m_i} \in \mathbb{C}[X]$ , l'ensemble  $\mathcal{C}(P) = \{A \in M_n(\mathbb{C}) \mid P_A(X) = P\}$  est-il formé d'une seule classe de similitude?

**3.1.14. Opérations sur les colonnes et bases de Jordan.** — Étant donné  $A \in M_n(\mathbb{C})$  ou, plus généralement  $A \in M_n(k)$  avec  $P_A(X)$  scindé, la méthode de réduction des colonnes s'applique très bien au calcul explicite d'une « base de Jordan » de l'endomorphisme  $u$  défini par  $A$ , i.e. d'une base  $\mathcal{C}$  dans laquelle la matrice de  $u$  est la forme normale de Jordan  $J_A$  de  $A$ . Illustrons ceci par deux exemples.

**Exemple 1.** Soit  $A = \begin{pmatrix} 0 & 1 & -1 & 0 \\ -3 & -1 & 3 & 1 \\ -2 & 1 & 1 & 1 \\ 0 & 0 & 0 & 2 \end{pmatrix} \in M_4(\mathbb{C})$  et soit  $u$  l'endomorphisme de  $\mathbb{C}^4$  défini par

$A$ . Calculer le polynôme caractéristique  $P_A(X)$  et déterminer ses racines et leur multiplicité. Puis, pour chaque racine  $\lambda$ , déterminer une base de  $\text{Ker}(A - \lambda I_4)$ , puis de  $\text{Ker}((A - \lambda I_4)^2)$ , etc. jusqu'à obtenir l'espace caractéristique  $V_{(\lambda)}$ . Enfin, donner une base  $\mathcal{C}$  de  $\mathbb{C}^4$  telle que  $\text{Mat}_{\mathcal{C}}(u)$  soit la forme normale de Jordan  $J_A$  de  $A$ .

En développant par rapport à la dernière ligne, on obtient que  $P_A(X)$  égale :

$$\begin{aligned} (2-X) \begin{vmatrix} -X & 1 & -1 & 0 \\ -3 & -1-X & 3 & 1 \\ -2 & 1 & 1 & 1-X \end{vmatrix} &\xrightarrow{C_3 \rightarrow C_3 + C_2} (2-X) \begin{vmatrix} -X & 1 & 0 & 0 \\ -3 & -1-X & 2-X & 1 \\ -2 & 1 & 2-X & 1-X \end{vmatrix} \\ &= (2-X)^2 \begin{vmatrix} -X & 1 & 0 & 0 \\ -3 & -1-X & 1 & 1 \\ -2 & 1 & 1 & 1 \end{vmatrix} \xrightarrow{L_2 \rightarrow L_2 - L_3} (2-X)^2 \begin{vmatrix} -X & 1 & 0 & 0 \\ -1 & -2-X & 0 & 0 \\ -2 & 1 & 1 & 1 \end{vmatrix} = (2-X)^2 (X^2 + 2X + 1) \\ &= (X-2)^2 (X+1)^2. \end{aligned}$$

Pour la valeur propre  $\lambda = -1$ , posons  $B = A + I_4$  et faisons des opérations sur les colonnes de  $B = A + I_4$  :

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ \hline 1 & 1 & -1 & 0 \\ -3 & 0 & 3 & 1 \\ -2 & 1 & 2 & 1 \\ 0 & 0 & 0 & 3 \end{pmatrix} \xrightarrow[\begin{smallmatrix} C_3 \rightarrow C_3 + C_1 \\ C_1 \rightarrow C_1 - C_2 \end{smallmatrix}]{\begin{smallmatrix} C_3 \rightarrow C_3 + C_2 \\ C_1 \rightarrow C_1 + 3C_4 \end{smallmatrix}} \begin{pmatrix} 1 & 0 & 1 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ \hline 0 & 1 & 0 & 0 \\ -3 & 0 & 0 & 1 \\ -3 & 1 & 0 & 1 \\ 0 & 0 & 0 & 3 \end{pmatrix} \xrightarrow{C_1 \rightarrow C_1 + 3C_4} \begin{pmatrix} 1 & 0 & 1 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 3 & 0 & 0 & 1 \\ \hline 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 9 & 0 & 0 & 3 \end{pmatrix} = \begin{pmatrix} P \\ B' \end{pmatrix},$$

où  $B' = BP$ . Donc, comme les colonnes non nulles de  $B'$  sont échelonnées,  $\text{Ker}(B)$  est engendré par le vecteur  $v_1 = e_1 + e_3$ . Pour la suite, on a deux méthodes.

1ère méthode. Calculons  $B^2 = (A + I_4)^2$  :

$$\begin{pmatrix} 1 & 1 & -1 & 0 \\ -3 & 0 & 3 & 1 \\ -2 & 1 & 2 & 1 \\ 0 & 0 & 0 & 3 \end{pmatrix} \begin{pmatrix} 1 & 1 & -1 & 0 \\ -3 & 0 & 3 & 1 \\ -2 & 1 & 2 & 1 \\ 0 & 0 & 0 & 3 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ -9 & 0 & 9 & 6 \\ -9 & 0 & 9 & 6 \\ 0 & 0 & 0 & 9 \end{pmatrix}$$

Donc on voit que  $B^2 = (A + I_4)^2$  est de rang 2, et  $\text{Ker}(B^2)$  contient les vecteurs  $e_2$  et  $e_1 + e_3$ , qui forment donc une base de  $\text{Ker}(B^2)$ . De plus, comme  $\text{Ker}(B)$  est engendré par  $e_1 + e_3$ , alors  $e_2$  engendre un supplémentaire de  $\text{Ker}(B)$  dans  $\text{Ker}(B^2)$ .

2ème méthode. On a aussi la méthode plus courte qui suit, suggérée par A. Moussaoui (en réponse à une question de T. de La Rochefoucauld). Les colonnes de la matrice inversible  $P$  forment une base  $(f_1, f_2, f_3, f_4)$  de  $\mathbb{C}^4$ , et les colonnes de  $B' = BP$  sont les images par  $B$  des  $f_i$ . Il suffit donc de multiplier  $B'$  par  $B$  pour avoir la matrice  $BB' = B^2P$  dont les colonnes donnent les images par  $B^2$  des  $f_i$ . Comme ici  $Bf_3 = 0$ , on a bien sûr  $B^2f_3 = 0$  et il suffira de faire les mêmes opérations sur  $BB'$  et  $P$  pour créer une

autre colonne nulle de  $BB'$ . Dans le cas présent, le calcul est particulièrement simple (voir plus bas pour un exemple plus élaboré). On a :

$$BB' = \begin{pmatrix} 1 & 1 & -1 & 0 \\ -3 & 0 & 3 & 1 \\ -2 & 1 & 2 & 1 \\ 0 & 0 & 0 & 3 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 9 & 0 & 0 & 3 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 9 & 0 & 0 & 6 \\ 9 & 0 & 0 & 6 \\ 27 & 0 & 0 & 9 \end{pmatrix}$$

d'où

$$\begin{pmatrix} P \\ BB' \end{pmatrix} = \begin{pmatrix} 1 & 0 & 1 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 3 & 0 & 0 & 1 \\ \hline 0 & 0 & 0 & 0 \\ 9 & 0 & 0 & 6 \\ 9 & 0 & 0 & 6 \\ 27 & 0 & 0 & 9 \end{pmatrix}$$

donc on voit que, outre le vecteur  $e_1 + e_3$  qui appartenait déjà à  $\text{Ker}(B)$ ,  $\text{Ker}(B^2)$  contient le vecteur  $e_2$ , qui engendre donc un supplémentaire de  $\text{Ker}(B)$  dans  $\text{Ker}(B^2)$ .

Finalement, utilisant l'une ou l'autre méthode, comme la dimension de l'espace caractéristique  $V_{(-1)}$  est la multiplicité algébrique de la valeur propre  $-1$ , à savoir 2, on conclut que

$$V_{(-1)} = \text{Ker}((A + I_4)^2) = \text{Vect}(e_1 + e_3, e_2).$$

Considérons maintenant la valeur propre  $\lambda = 2$  et faisons des opérations sur les colonnes de la matrice  $C = A - 2I_4$  :

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ \hline -2 & 1 & -1 & 0 \\ -3 & -3 & 3 & 1 \\ -2 & 1 & -1 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \xrightarrow[\substack{C_3 \rightarrow C_3 + C_2 \\ C_1 \rightarrow C_1 + 2C_2}]{\substack{C_3 \rightarrow C_3 + C_2 \\ C_1 \rightarrow C_1 + 2C_2}} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 2 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ \hline 0 & 1 & 0 & 0 \\ -9 & -3 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \xrightarrow{C_1 \rightarrow C_1 + 9C_4} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 2 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 9 & 0 & 0 & 1 \\ \hline 0 & 1 & 0 & 0 \\ 0 & -3 & 0 & 1 \\ 9 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} Q \\ C' \end{pmatrix},$$

où  $C' = CQ$ . Donc, comme les colonnes non nulles de  $C'$  sont échelonnées,  $\text{Ker}(C)$  est engendré par le vecteur  $v_1 = e_2 + e_3$ . Pour la suite, on a comme avant deux méthodes.

1ère méthode. Calculons  $C^2 = (A - 2I_4)^2$  :

$$\begin{pmatrix} -2 & 1 & -1 & 0 \\ -3 & -3 & 3 & 1 \\ -2 & 1 & -1 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} -2 & 1 & -1 & 0 \\ -3 & -3 & 3 & 1 \\ -2 & 1 & -1 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 3 & -6 & 6 & 0 \\ 9 & 9 & -9 & 0 \\ 3 & -6 & 6 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

On voit donc que  $C^2 = (A - 2I_4)^2$  est de rang 2 (car les deux premières colonnes sont linéairement indépendantes), et que  $\text{Ker}(C^2)$  contient les vecteurs  $e_4$  et  $e_2 + e_3$ , qui forment donc une base de  $\text{Ker}(C^2)$ . De plus, comme  $\text{Ker}(C)$  est engendré par  $e_2 + e_3$ , alors  $e_4$  engendre un supplémentaire de  $\text{Ker}(C)$  dans  $\text{Ker}(C^2)$ .

2ème méthode. On a :

$$CC' = \begin{pmatrix} -2 & 1 & -1 & 0 \\ -3 & -3 & 3 & 1 \\ -2 & 1 & -1 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & -3 & 0 & 1 \\ 9 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} -9 & -6 & 0 & 0 \\ 27 & 9 & 0 & 0 \\ -9 & -6 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

d'où

$$\begin{pmatrix} Q \\ CC' \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 2 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 9 & 0 & 0 & 1 \\ \hline -9 & -6 & 0 & 0 \\ 27 & 9 & 0 & 0 \\ -9 & -6 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

donc on voit que, outre le vecteur  $e_2 + e_3$  qui appartenait déjà à  $\text{Ker}(C)$ ,  $\text{Ker}(C^2)$  contient le vecteur  $e_4$ , qui engendre donc un supplémentaire de  $\text{Ker}(C)$  dans  $\text{Ker}(C^2)$ .

Finalement, utilisant l'une ou l'autre méthode, comme la dimension de l'espace caractéristique  $V_{(2)}$  est la multiplicité algébrique de la valeur propre 2, à savoir 2, on conclut que

$$V_{(2)} = \text{Ker}((A - 2I_4)^2) = \text{Vect}(e_2 + e_3, e_4).$$

Il résulte de ce qui précède que la forme normale de Jordan de  $A$  est la matrice

$$J_A = \left( \begin{array}{cc|cc} -1 & 1 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ \hline 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 2 \end{array} \right).$$

Plus précisément, soit  $u$  l'endomorphisme de  $\mathbb{C}^4$  défini par  $A$ . Comme  $(u + \text{id})(e_2) = e_1 + e_3 \in V_{-1}$  et  $(u - 2 \text{id})(e_4) = e_2 + e_3 \in V_2$ , alors  $\mathcal{C} = (e_1 + e_3, e_2, e_2 + e_3, e_4)$  est une base de  $V = \mathbb{C}^4$  dans laquelle la matrice de  $u$  est la matrice  $J_A$  ci-dessus.

**Exemple 2.** Soit  $A = \begin{pmatrix} 1 & 1 & 1 & 1 \\ -1 & 1 & 1 & 1 \\ 1 & -1 & 3 & -1 \\ -1 & 1 & 1 & 3 \end{pmatrix} \in M_4(\mathbb{R})$  et soit  $u$  l'endomorphisme de  $\mathbb{R}^4$  défini par  $A$ .

Calculer  $P_A(X)$  et déterminer ses racines. Pour chaque racine  $\lambda$ , déterminer une base de chaque  $\text{Ker}((A - \lambda I_4)^i)$ , pour  $i = 1, 2, \dots$  et en déduire la forme normale de Jordan  $J$  de  $A$ , ainsi qu'une base  $\mathcal{C}$  de  $\mathbb{R}^4$  telle que  $\text{Mat}_{\mathcal{C}}(u) = J$ .

On a :

$$\begin{aligned} & \left| \begin{array}{cccc} 1-X & 1 & 1 & 1 \\ -1 & 1-X & 1 & 1 \\ 1 & -1 & 3-X & -1 \\ -1 & 1 & 1 & 3-X \end{array} \right| \xrightarrow{C_1 \rightarrow C_1 + C_4} \left| \begin{array}{cccc} 2-X & 1 & 1 & 1 \\ 0 & 1-X & 1 & 1 \\ 0 & -1 & 3-X & -1 \\ 2-X & 1 & 1 & 3-X \end{array} \right| = \\ & (2-X) \left| \begin{array}{cccc} 1 & 1 & 1 & 1 \\ 0 & 1-X & 1 & 1 \\ 0 & -1 & 3-X & -1 \\ 1 & 1 & 1 & 3-X \end{array} \right| \xrightarrow{L_4 \rightarrow L_4 - L_1} (2-X) \left| \begin{array}{cccc} 1 & 1 & 1 & 1 \\ 0 & 1-X & 1 & 1 \\ 0 & -1 & 3-X & -1 \\ 0 & 0 & 0 & 2-X \end{array} \right| = \\ & (2-X)^2 \left| \begin{array}{cc} 1-X & 1 \\ -1 & 3-X \end{array} \right| = (2-X)^2 (X^2 - 4X + 4) = (X-2)^4 \end{aligned}$$

donc  $\lambda = 2$  est la seule valeur propre de  $A$ . Posons  $B = A - 2I_4$  et déterminons une base de  $\text{Ker}(B^i)$ , pour  $i = 1, 2, \dots$  en faisant des opérations sur les colonnes.

Notons  $(e_1, e_2, e_3, e_4)$  la base canonique de  $\mathbb{R}^4$ . Le calcul précédent montre déjà que  $e_1 + e_4 \in \text{Ker}(B)$  :

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ \hline -1 & 1 & 1 & 1 \\ -1 & -1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ -1 & 1 & 1 & 1 \end{pmatrix} \xrightarrow{\begin{array}{l} C_1 \rightarrow C_1 + C_4 \\ C_3 \rightarrow C_3 - C_2 \\ C_4 \rightarrow C_4 - C_2 \end{array}} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & -1 & -1 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ \hline 0 & 1 & 0 & 0 \\ 0 & -1 & 2 & 2 \\ 0 & -1 & 2 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} = \begin{pmatrix} P \\ B' \end{pmatrix}, \quad \text{où } B' = BP.$$

Comme les colonnes non nulles de  $B'$  sont échelonnées, on obtient que  $\text{Ker}(B)$  est de dimension 1, engendré par  $e_1 + e_4$ . Comme le nombre de blocs de Jordan pour la valeur propre  $\lambda = 2$  est la dimension de  $\text{Ker}(A - 2I_4) = \text{Ker}(B)$ , on sait donc déjà que la forme normale de Jordan de  $A$  est

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

Calculons maintenant :

$$BB' = \begin{pmatrix} -1 & 1 & 1 & 1 \\ -1 & -1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ -1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & -1 & 2 & 2 \\ 0 & -1 & 2 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & -2 & 4 & 2 \\ 0 & 0 & 0 & -2 \\ 0 & 0 & 0 & -2 \\ 0 & -2 & 4 & 2 \end{pmatrix}$$

d'où

$$\begin{pmatrix} P \\ BB' \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & -1 & -1 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & -2 & 4 & 2 \\ 0 & 0 & 0 & -2 \\ 0 & 0 & 0 & -2 \\ 0 & -2 & 4 & 2 \end{pmatrix} \xrightarrow[\substack{C_3 \rightarrow C_3 + 2C_2 \\ C_4 \rightarrow C_4 + C_2}]{} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & -2 & 0 & 0 \\ 0 & 0 & 0 & -2 \\ 0 & 0 & 0 & -2 \\ 0 & -2 & 0 & 0 \end{pmatrix} = \begin{pmatrix} PQ \\ B'' \end{pmatrix}, \quad \text{où } B'' = B^2 PQ.$$

Comme les colonnes non nulles de  $B''$  sont échelonnées, on obtient que  $\text{Ker}(B^2)$  est de dimension 2, et que  $e_2 + e_3$  engendre un supplémentaire de  $\text{Ker}(B) = \mathbb{R}(e_1 + e_4)$  dans  $\text{Ker}(B^2)$ . Calculons maintenant :

$$BB'' = \begin{pmatrix} -1 & 1 & 1 & 1 \\ -1 & -1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ -1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 0 & -2 & 0 & 0 \\ 0 & 0 & 0 & -2 \\ 0 & 0 & 0 & -2 \\ 0 & -2 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & -4 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -4 \end{pmatrix}$$

d'où

$$\begin{pmatrix} PQ \\ BB'' \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & -4 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -4 \end{pmatrix} = \begin{pmatrix} PQ \\ B''' \end{pmatrix}, \quad \text{où } B''' = B^3 PQ$$

donc  $\text{Ker}(B^3)$  est de dimension 3, et  $e_2$  engendre un supplémentaire de  $\text{Ker}(B^2)$  dans  $\text{Ker}(B^3)$ .

Enfin, on voit que  $B^3(e_4) = -4(e_1 + e_4) \neq 0$ , donc  $e_4$  engendre un supplémentaire de  $\text{Ker}(B^3)$  dans  $\text{Ker}(B^4) = \mathbb{R}^4$ .

$$\text{Donc les vecteurs } v_4 = e_4, v_3 = (u - 2 \text{id})(e_4) = \begin{pmatrix} 1 \\ 1 \\ -1 \\ 1 \end{pmatrix},$$

$$v_2 = (u - 2 \text{id})(v_3) = \begin{pmatrix} -1 & 1 & 1 & 1 \\ -1 & -1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ -1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ -1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ -2 \\ -2 \\ 0 \end{pmatrix} = B^2 e_4$$

et

$$v_1 = (u - 2 \text{id})(v_2) = B^3 e_4 = \begin{pmatrix} -4 \\ 0 \\ 0 \\ -4 \end{pmatrix}$$



forment une base  $\mathcal{C}$  de  $\mathbb{R}^4$  telle que

$$\text{Mat}_{\mathcal{C}}(u) = J = \begin{pmatrix} 2 & 1 & 0 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 2 \end{pmatrix}.$$

**Remarque.** Les matrices échelonnées  $B'$ ,  $B''$  et  $B'''$  montrent que

$$\text{Im}(u) = \text{Vect}(e_1 + e_4, e_2, e_3), \quad \text{Im}(u^2) = \text{Vect}(e_1 + e_4, e_2 + e_3), \quad \text{Im}(u^3) = \text{Vect}(e_1 + e_4).$$

### 3.2. Décomposition de Dunford

Soient  $k$  un corps et  $V$  un  $k$ -espace vectoriel de dimension finie.



**Définition 3.2.1 (Endomorphismes qui commutent).** — On dit que deux endomorphismes  $u, v$  de  $V$  commutent si  $u \circ v = v \circ u$ . Dans ce cas, pour tout entier  $n \geq 1$  on peut calculer  $(u + v)^n$  par la formule du binôme :

$$(u + v)^n = \sum_{i=0}^n \binom{n}{i} u^{n-i} v^i, \quad \text{où} \quad \binom{n}{i} = \frac{n!}{i!(n-i)!} = \frac{n(n-1)\cdots(n-i+1)}{i!}.$$

(Le coefficient binomial  $\binom{n}{i}$  est aussi noté  $C_n^i$ .)

Attention, cette formule est fautive si  $u$  et  $v$  ne commutent pas ! Par exemple, on a  $(u + v)^2 = u^2 + uv + vu + v^2$  et ceci est  $\neq u^2 + 2uv + v^2$  si  $vu \neq uv$ , par exemple si  $u = E_{12}$  et  $v = E_{21}$  dans  $M_2(k)$ .

**Lemme 3.2.2.** — Soient  $u, v$  deux endomorphismes de  $V$  qui commutent.

(1) Soient  $\lambda$  une valeur propre de  $u$ ,  $m$  sa multiplicité algébrique, et  $V_\lambda = \text{Ker}(u - \lambda)$  (resp.  $V_{(\lambda)} = \text{Ker}((u - \lambda \text{id}_V)^m)$ ) l'espace propre (resp. caractéristique) associé. Alors  $V_\lambda$  et  $V_{(\lambda)}$  sont stables par  $v$ .

(2) Si  $u$  et  $v$  sont diagonalisables, alors  $V$  possède une base formée de vecteurs propres communs à  $u$  et  $v$ . Par conséquent,  $u + v$  et  $uv$  sont diagonalisables.

(3) Si  $u$  et  $v$  sont nilpotents, il en est de même de  $uv$  et de  $u + v$ .

**Démonstration.** — (1) Soit  $x \in V_\lambda$ , alors  $u(v(x)) = v(u(x)) = \lambda v(x)$ , donc  $v(x) \in V_\lambda$ . De même, comme  $v$  commute à  $u$ , il commute aussi à  $U = (u - \lambda \text{id}_V)^m$ , donc si  $x \in V_{(\lambda)}$ , alors  $U(v(x)) = v(U(x)) = 0$ , donc  $v(x) \in \text{Ker}(U) = V_{(\lambda)}$ . Ceci montre que  $V_\lambda$  et  $V_{(\lambda)}$  sont stables par  $v$ .

(2) Soit  $V = V_{\lambda_1} \oplus \cdots \oplus V_{\lambda_r}$  la décomposition de  $V$  en espaces propres de  $u$ . Fixons un indice  $i$ . D'après (1),  $V_{\lambda_i}$  est stable par  $v$  et d'après le théorème 2.1.14,  $V_{\lambda_i}$  admet une base  $\mathcal{B}_i$  formée de vecteurs propres de  $v$ , qui sont tous des vecteurs propres de  $u$  pour la valeur propre  $\lambda_i$ . Alors  $\mathcal{B} = \mathcal{B}_1 \cup \cdots \cup \mathcal{B}_r$  est une base de  $V$  formée de vecteurs propres communs à  $u$  et  $v$ ; et dans cette base les matrices de  $u + v$  et  $uv$  sont diagonales.

(3) Supposons  $u^r = 0 = v^s$ . Comme  $u$  et  $v$  commutent, on a  $(uv)^n = u^n v^n = 0$  si  $n \geq \max(r, s)$ . D'autre part, pour tout  $n \in \mathbb{N}$  on a

$$(u + v)^n = \sum_{i=0}^n \binom{n}{i} u^i v^{n-i};$$

le terme  $u^i$  (resp.  $v^{n-i}$ ) est nul si  $i \geq r$  (resp.  $n - i \geq s$ ), donc pour que le terme  $u^i v^{n-i}$  soit  $\neq 0$ , il faut que  $i \leq r - 1$  et  $n - i \leq s - 1$ , d'où  $n \leq r + s - 2$  : ceci montre que  $(u + v)^{r+s-1} = 0$ .  $\square$

**Remarques 3.2.3.** — 1) Attention, si  $u$  et  $v$  sont diagonalisables mais ne commutent pas, alors en général  $u + v$  et  $uv$  ne sont pas diagonalisables ! Par exemple, les matrices  $A = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$  et  $B = \begin{pmatrix} 1 & 1 \\ 0 & -1 \end{pmatrix}$  dans  $M_2(\mathbb{R})$  sont diagonalisables, mais ni  $A + B$  ni  $AB$  ne l'est. (Exercice : vérifier ces assertions.)

2) Attention, si  $u$  et  $v$  sont des endomorphismes nilpotents, alors en général ni  $uv$  ni  $u + v$  ne sont nilpotents ! Par exemple, dans  $M_2(k)$  les matrices élémentaires  $E_{12}$  et  $E_{21}$  sont de carré nul, mais  $E_{12}E_{21} = E_{11}$  n'est pas nilpotente, et  $S = E_{21} + E_{12}$  non plus (car  $S^2 = I_2$ ).

**Lemme 3.2.4.** — Soit  $u \in \text{End}_k(V)$ .

(1) Si  $u$  est nilpotent, 0 est valeur propre de  $u$ , et c'est la seule valeur propre.

(2) En particulier, si  $u$  est diagonalisable et nilpotent, alors  $u = 0$ .

*Démonstration.* — (1) Supposons  $u$  nilpotent et soit  $r$  son indice de nilpotence, c.-à-d.,  $u^r = 0$  mais  $u^{r-1} \neq 0$ . Soit  $x \in V$  tel que  $u^{r-1}(x) \neq 0$ , alors  $u^{r-1}(x)$  appartient à  $\text{Ker}(u)$  donc est vecteur propre pour la valeur propre 0.

Réciproquement, si  $\mu$  est une valeur propre de  $u$  et  $x \neq 0$  un vecteur propre associé, alors  $u(x) = \mu x$  entraîne  $0 = u^r(x) = \mu^r x$ , d'où  $\mu = 0$ . Ceci prouve (1).

(2) en découle, car si  $V$  admet une base  $(e_1, \dots, e_n)$  formée de vecteurs propres de  $u$ , chacun associé à la valeur propre 0, alors  $u(e_i) = 0$  pour tout  $i$ , donc  $u = 0$ .  $\square$



**Théorème et définition 3.2.5 (Décomposition de Dunford).** — Soient  $V$  un  $\mathbb{C}$ -espace vectoriel de dimension finie et  $u \in \text{End}_{\mathbb{C}}(V)$ . Alors  $u$  se décompose de façon unique sous la forme

$$(\dagger) \quad u = s + n, \quad \text{avec} \quad \begin{cases} s \text{ diagonalisable et } n \text{ nilpotent,} \\ s \text{ et } n \text{ qui commutent, i.e. } sn = ns. \end{cases}$$

On dit que  $s$  (resp.  $u$ ) est la partie semi-simple (resp. partie nilpotente) de  $u$ .

*Démonstration.* — Écrivons  $P_u(X) = (-1)^n \prod_{i=1}^r (X - \lambda_i)^{m_i}$ , où  $\lambda_1, \dots, \lambda_r$  sont les valeurs propres, deux à deux distinctes, de  $u$ . Pour tout  $i$ , soit  $u_i$  la restriction de  $u$  à l'espace caractéristique  $N_i = V_{(\lambda_i)}$  et soit  $\mathcal{B}_i$  une base de  $N_i$  telle que  $\text{Mat}_{\mathcal{B}_i}(u_i)$  soit une matrice de Jordan  $J_{\mathbf{p}_i}(\lambda_i)$  (alors, nécessairement,  $\tilde{\mathbf{p}}_i$  correspond à la suite des noyaux de  $u_i$ ).

Soit  $s$  l'endomorphisme de  $V$  qui égale  $\lambda_i \text{id}_{N_i}$  sur chaque  $N_i$ . Alors, pour tout  $x \in V$ , on a :

$$(s \circ u)(x) = (u \circ s)(x).$$

En effet, comme les deux membres sont linéaires en  $x$ , il suffit de vérifier cette égalité lorsque  $x \in N_i$ ; dans ce cas les deux membres égalent  $\lambda_i u(x)$ . Donc  $s$  et  $u$  commutent.

D'autre part, la matrice de  $n = u - s$  est triangulaire stricte, donc nilpotente (cf. 3.1.2). On a donc décomposé  $u$  sous la forme :  $u = s + n$  avec  $s$  et  $n$  qui commutent,  $s$  diagonalisable et  $n$  nilpotent. Ceci prouve l'existence.

Montrons l'unicité. Soit  $u = s' + n'$  une autre décomposition ayant les mêmes propriétés. Alors on a :

$$(*) \quad s - s' = n' - n.$$

D'autre part, comme  $s'$  et  $n'$  commutent, ils commutent avec leur somme  $s' + n' = u$ , donc, d'après le lemme 3.2.2, ils préservent chaque espace caractéristique  $N_i$  de  $u$ , donc commutent avec  $s$  qui est une homothétie sur chacun de ces espaces. Ils commutent donc aussi avec  $n = u - s$ .

Alors, comme  $s$  et  $s'$  (resp.  $n$  et  $n'$ ) commutent et sont diagonalisables (resp. nilpotents),  $s - s'$  est diagonalisable et  $n' - n$  est nilpotent, d'après 3.2.2. Donc,  $s - s' = n' - n$  est à la fois diagonalisable et nilpotent, donc nul d'après 3.2.4, d'où  $s = s'$  et  $n = n'$ . Ceci prouve l'unicité de la décomposition ( $\dagger$ ).  $\square$

**Terminologie 3.2.5.1.** — La décomposition de Dunford est aussi appelé *décomposition de Jordan*. On a évité cette terminologie, pour éviter une confusion avec la « réduction à la forme normale de Jordan ».



**Remarque 3.2.6.** — Attention! Si  $A$  est une matrice triangulaire supérieure, et si l'on note  $D$  la « partie diagonale » de  $A$  et  $T$  la partie de  $A$  « au-dessus de la diagonale » alors l'écriture

$$A = D + T$$

*n'est pas* en général la décomposition de Dunford, car  $D$  et  $T$  ne commutent pas nécessairement! (Le théorème précédent dit juste qu'on peut faire un changement de base, c.-à-d., remplacer  $A$  par une matrice conjuguée  $A' = P^{-1}AP$ , de sorte que  $A' = D' + T'$  soit la décomposition de Dunford de l'endomorphisme  $u$  défini par  $A$ .) Par exemple, dans  $M_2(\mathbb{C})$ ,

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

*n'est pas* la décomposition de Dunford-Jordan de la matrice  $A$  de gauche : celle-ci est diagonalisable ( $Ae_1 = e_1$  et  $A(e_2 - e_1) = -(e_2 + e_1)$ ) donc égale à sa partie semi-simple! (Et bien sûr, si l'on note  $u$  l'endomorphisme de  $\mathbb{R}^2$  défini par  $A$ , la matrice de  $u$  dans la base  $\mathcal{B}' = (e_1, e_2 - e_1)$  est  $A' = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ .)

### 3.3. Exponentielles de matrices

Dans cette section, on désigne par  $\mathbb{K}$  le corps  $\mathbb{R}$  des réels ou le corps  $\mathbb{C}$  des nombres complexes et l'on note  $|\cdot|$  la valeur absolue usuelle sur  $\mathbb{K}$ , c.-à-d., si  $x \in \mathbb{R}$ ,  $|x| = \sqrt{x^2}$  et si  $z = a + ib \in \mathbb{C}$  (où  $a, b \in \mathbb{R}$  et  $i^2 = -1$ ),  $|z| = \sqrt{z\bar{z}} = \sqrt{a^2 + b^2}$ .

**Définition 3.3.1 (Normes).** — Soit  $E$  un  $\mathbb{K}$ -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  $t \in \mathbb{K}$ ,  $x \in E$ , on a  $\|tx\| = |t| \cdot \|x\|$  (où  $|t|$  est la valeur absolue de  $t$ ).
- (3)  $\|u + v\| \leq \|u\| + \|v\|$ , pour tout  $u, v \in E$ .

Dans ce cas, on dit que  $E$  est un  $\mathbb{K}$ -*espace vectoriel normé* (en abrégé : evn).

**Définitions 3.3.2 (Suites de Cauchy).** — Soit  $(E, \|\cdot\|)$  un  $\mathbb{K}$ -espace vectoriel normé.

(1) On dit qu'une suite  $(u_n)_{n \in \mathbb{N}}$  d'éléments de  $E$  est une suite *de Cauchy* si la propriété suivante est vérifiée :

$$\text{(Cauchy)} \quad \forall \varepsilon > 0, \quad \exists n_0 \quad \text{tel que} \quad \forall m, n \geq n_0, \quad \|u_m - u_n\| < \varepsilon.$$

(On peut mémoriser ceci en disant que la « suite des différences  $u_m - u_n$  » tend vers 0 quand  $m, n \rightarrow +\infty$ .)

(2) On dit que  $(E, \|\cdot\|)$  est *complet* si toute suite de Cauchy  $(u_n)_{n \in \mathbb{N}}$  est convergente (i.e. il existe  $\ell \in E$  tel que  $\lim_{n \rightarrow +\infty} \|u_n - \ell\| = 0$ , un tel  $\ell$  étant alors unique).

On admet la proposition ci-dessous (une démonstration est donnée dans un appendice à la fin de ce chapitre).

**Proposition 3.3.3.** — Soit  $E$  un  $\mathbb{K}$ -espace vectoriel de dimension finie. Alors toutes les normes sur  $E$  sont équivalentes, c.-à-d., si  $\|\cdot\|$  et  $\|\cdot\|'$  sont deux normes sur  $E$ , il existe des constantes  $c, C \in \mathbb{R}_+^*$  telles que :

$$(\dagger) \quad \forall x \in E, \quad c \cdot \|x\| \leq \|x\|' \leq C \cdot \|x\|.$$

On en déduit les deux théorèmes suivants.

**Théorème 3.3.4.** — Soit  $E$  un  $\mathbb{K}$ -espace vectoriel de dimension finie, muni d'une norme  $\|\cdot\|$ .

(i) La notion de suite de Cauchy dans  $E$  ne dépend pas de la norme choisie. C'est-à-dire, si  $(u_n)_{n \in \mathbb{N}}$  est une suite d'éléments de  $E$  et si  $\|\cdot\|'$  est une seconde norme sur  $E$ , alors  $(u_n)_{n \in \mathbb{N}}$  est de Cauchy pour  $\|\cdot\|$  si et seulement si elle l'est pour  $\|\cdot\|'$ .

(ii)  $E$  est complet pour la norme  $\|\cdot\|$ .

*Démonstration.* — Le point (i) est une conséquence immédiate de la proposition précédente. Prouvons le point (ii). Soit  $\mathcal{B} = (e_1, \dots, e_d)$  une base de  $E$ . D'après le point (i), il suffit de montrer que  $E$  est complet pour la norme  $\|\cdot\| = \|\cdot\|_{\mathcal{B}, \infty}$  définie par :

$$\forall (x_1, \dots, x_d) \in \mathbb{K}^d, \quad \|x_1 e_1 + \dots + x_d e_d\| = \text{Max}(|x_1|, \dots, |x_d|).$$

Soit alors  $(u_n)_{n \in \mathbb{N}}$  une suite de Cauchy pour cette norme. Pour tout  $i = 1, \dots, d$ , notons  $(u_n^i)_{n \in \mathbb{N}}$  la suite à valeurs dans  $\mathbb{K}$  formée par les  $i$ -ièmes coordonnées des vecteurs  $u_n$  (c.-à-d.,  $u_n = u_n^1 e_1 + \dots + u_n^d e_d$ ). Alors pour tout  $i$  et tous  $m, n$  on a

$$|u_m^i - u_n^i| \leq \|u_m - u_n\|$$

et donc la suite  $(u_n^i)_{n \in \mathbb{N}}$  est de Cauchy, donc converge vers une limite  $\ell_i$  (puisque  $\mathbb{K} = \mathbb{R}$  ou  $\mathbb{C}$  est complet). Alors on voit facilement que la suite  $(u_n)_{n \in \mathbb{N}}$  converge vers l'élément  $\ell = \ell_1 e_1 + \dots + \ell_d e_d$  de  $E$ . Ceci prouve le théorème.  $\square$

**Théorème 3.3.5.** — Soient  $E, F$  deux  $\mathbb{K}$ -espaces vectoriels de dimension finie, munis de normes  $\|\cdot\|_E$  et  $\|\cdot\|_F$ , et soit  $f : E \rightarrow F$  une application linéaire. Alors il existe une constante  $k \in \mathbb{R}_+^*$  telle que :

$$(\ddagger) \quad \forall x \in E, \quad \|f(x)\|_F \leq k \cdot \|x\|_E.$$

En particulier,  $f$  est continue.

*Démonstration.* — Soit  $\mathcal{B} = (e_1, \dots, e_d)$  une base de  $E$  et soit  $M = \sum_{i=1}^d \|f(e_i)\|_F$ . Notons  $\|\cdot\|_{\mathcal{B}, \infty}$  la norme sur  $E$  définie plus haut. D'après la proposition 3.3.3, il existe  $C \in \mathbb{R}_+^*$  tel que  $\|v\|_\infty \leq C \cdot \|v\|_E$ , pour tout  $v \in E$ . Écrivant  $v = x_1 e_1 + \dots + x_d e_d$ , on a alors

$$\|f(v)\|_F \leq \sum_{i=1}^d |x_i| \cdot \|f(e_i)\|_F \leq \|v\|_\infty \cdot M \leq CM \cdot \|v\|_E,$$

d'où le théorème. □

Dans la suite, on munit  $\mathbb{K}^d$  de la norme  $\|\cdot\|_\infty$ , définie par

$$\|(x_1, \dots, x_d)\|_\infty = \text{Max}(|x_1|, \dots, |x_d|), \quad \forall (x_1, \dots, x_d) \in \mathbb{K}^d. \quad (2)$$

Alors la sphère unité  $S^{d-1} = \{x \in \mathbb{K}^d \mid \|x\|_\infty = 1\}$  est compacte (étant un fermé borné de  $\mathbb{K}^d$ ).

**Définition 3.3.6 (Normes matricielles sur  $M_d(\mathbb{K})$ ).** — Pour tout  $A \in M_d(\mathbb{K})$ , l'application  $x \mapsto \|Ax\|_\infty$  est continue, donc est bornée sur le compact  $S^{d-1}$ . On pose alors

$$\| \|A\| \| = \text{Max}_{x \in S^{d-1}} \|Ax\|_\infty.$$

On vérifie facilement que  $\| \| \cdot \| \|$  est une norme sur  $M_d(\mathbb{K})$ , appelée *norme matricielle* associée à la norme  $\|\cdot\|_\infty$  donnée sur  $\mathbb{K}^d$ .

**Remarque 3.3.7.** — Pour tout  $x \in \mathbb{K}^d - \{0\}$ ,  $x' = \frac{1}{\|x\|_\infty} x$  appartient à  $S^{d-1}$  et comme  $x = \|x\|_\infty \cdot x'$  on a :

$$\|Ax\|_\infty = \|x\|_\infty \cdot \|Ax'\|_\infty \leq \| \|A\| \| \cdot \|x\|_\infty,$$

inégalité qui est aussi vérifiée pour  $x = 0$ . Pour tout  $A, B \in M_d(\mathbb{K})$ , et  $x \in S^{d-1}$ , on a donc

$$\|ABx\|_\infty \leq \| \|A\| \| \cdot \|Bx\|_\infty \leq \| \|A\| \| \cdot \| \|B\| \|$$

d'où

$$(*) \quad \| \|AB\| \| \leq \| \|A\| \| \cdot \| \|B\| \|.$$

Par conséquent, pour tout  $n \in \mathbb{N}^*$ , on a

$$(*) \quad \| \|A^n\| \| \leq \| \|A\| \|^n.$$

(Par ailleurs, la norme matricielle de  $A^0 = I_d$  est égale à 1.)

**Proposition et définition 3.3.8.** — Pour tout  $A \in M_d(\mathbb{K})$ , la suite de matrices

$$S_n = \sum_{i=0}^n \frac{A^i}{i!} = I_d + A + \frac{A^2}{2} + \dots + \frac{A^n}{n!}$$

est de Cauchy donc converge. Sa limite est notée  $\exp(A)$ , et l'on a donc  $\exp(A) = \sum_{i=0}^{\infty} \frac{A^i}{i!}$ .

*Démonstration.* — En effet, d'après le théorème 3.3.4, le  $\mathbb{K}$ -espace vectoriel  $M_d(\mathbb{K}) \simeq \mathbb{K}^{d^2}$ , muni de la norme matricielle, est complet. D'autre part, pour tout  $q > p$ , on a

$$\| \|S_q - S_p\| \| = \| \| \sum_{i=p+1}^q \frac{A^i}{i!} \| \| \leq \sum_{i=p+1}^q \frac{\| \|A^i\| \|}{i!} \leq \sum_{i=p+1}^q \frac{\| \|A\| \|^i}{i!}$$

et comme la suite réelle  $\sum_{i=0}^n \frac{\| \|A\| \|^i}{i!}$  converge vers  $\exp(\| \|A\| \|)$ , alors pour tout  $\varepsilon > 0$  il existe  $p_0$  tel que

pour tous  $q > p \geq p_0$  on ait  $\sum_{i=p+1}^q \frac{\| \|A\| \|^i}{i!} < \varepsilon$ , et donc la suite  $(S_n)$  est de Cauchy, donc converge vers une matrice qu'on note  $\exp(A)$ . □



**Remarque 3.3.9.** — Si  $A$  est *nilpotente*, i.e. s'il existe  $n \in \mathbb{N}^*$  tel que  $A^n = 0$ , alors  $\exp(A)$  égale la somme finie  $I_d + A + \dots + A^{n-1}/(n-1)!$ ; plus généralement, pour tout  $t \in \mathbb{K}$  on a alors :

$$\exp(tA) = I_d + tA + \frac{t^2}{2}A^2 + \dots + \frac{t^{n-1}}{(n-1)!}A^{n-1} \quad (\text{si } A^n = 0).$$

<sup>(2)</sup>D'autre choix possibles sont les normes  $\|(x_1, \dots, x_d)\|_1 = |x_1| + \dots + |x_d|$  ou  $\|(x_1, \dots, x_d)\|_2 = \sqrt{|x_1|^2 + \dots + |x_d|^2}$ ; tous ces choix sont équivalents, d'après la proposition 3.3.3.



**Proposition 3.3.10.** — Soit  $A \in M_d(\mathbb{K})$  une matrice triangulaire, de termes diagonaux  $\lambda_1, \dots, \lambda_d$ . Alors  $\exp(A)$  est une matrice triangulaire, de termes diagonaux  $\exp(\lambda_1), \dots, \exp(\lambda_d)$ .

*Démonstration.* — Pour tout  $i, j$ , la forme linéaire  $\phi_{i,j} : M_d(\mathbb{K}) \rightarrow \mathbb{K}$ , qui à toute matrice  $A = (a_{ij})$  associe son coefficient  $a_{ij}$ , est continue. En effet, c'est un cas particulier du théorème 3.3.5 en prenant  $F = \mathbb{K}$  muni de la norme définie par la valeur absolue mais, plus simplement, ceci se voit directement comme suit. Comme  $a_{ij}$  est la  $i$ -ième coordonnée du vecteur  $Ae_j$  alors :

$$|a_{ij}| \leq \|Ae_j\|_\infty \leq \|A\| \cdot \underbrace{\|e_j\|_\infty}_{=1} = \|A\|.$$

Il en résulte que pour toute suite de matrices  $(B_n)_{n \in \mathbb{N}}$  convergeant vers une matrice  $B \in M_d(\mathbb{K})$ , on a  $\phi_{ij}(B) = \lim_{n \rightarrow +\infty} \phi_{ij}(B_n)$ . Appliquons ceci à la suite  $B_n = \sum_{p=0}^n A^p/p!$  des sommes partielles de  $\exp(A)$ . Comme  $A$  est supposée triangulaire, disons supérieure, alors chaque produit  $A^p/p!$  est une matrice triangulaire, de termes diagonaux les  $\lambda_i^p/p!$ , pour  $i = 1, \dots, d$ , et donc  $B_n$  est aussi triangulaire, de termes diagonaux les sommes partielles  $\sum_{p=0}^n \lambda_i^p/p!$ , pour  $i = 1, \dots, d$ . Donc, pour tout  $i, j = 1, \dots, d$ , on a

$$\phi_{ij}(\exp(A)) = \lim_{n \rightarrow +\infty} \phi_{ij}(B_n) = \begin{cases} 0 & \text{si } i > j \\ \exp(\lambda_i) & \text{si } i = j \end{cases}$$

d'où la proposition. □



**Proposition 3.3.11.** — Soient  $A \in M_d(\mathbb{K})$  et  $P \in GL_d(\mathbb{K})$ . Alors  $\exp({}^t A) = {}^t \exp(A)$  et  $\exp(P^{-1}AP) = P^{-1} \exp(A)P$ .

*Démonstration.* — Les applications  $A \mapsto {}^t A$  et  $A \mapsto P^{-1}AP$  sont linéaires, donc continues d'après le théorème 3.3.5. Par conséquent, pour toute suite de matrices  $(B_n)_{n \in \mathbb{N}}$  convergeant vers une matrice  $B \in M_d(\mathbb{K})$ , on a

$${}^t B = \lim_{n \rightarrow +\infty} {}^t B_n \quad \text{et} \quad P^{-1}BP = \lim_{n \rightarrow +\infty} P^{-1}B_nP.$$

Appliquons ceci à la suite  $B_n = \sum_{p=0}^n A^p/p!$  des sommes partielles de  $\exp(A)$ . Comme  ${}^t(A^n) = ({}^t A)^n$  et  $P^{-1}A^nP = (P^{-1}AP)^n$ , on obtient que

$${}^t \exp(A) = \lim_{n \rightarrow +\infty} \sum_{i=0}^n \frac{({}^t A)^n}{n!} = \exp({}^t A) \quad \text{et} \quad P^{-1}BP = \lim_{n \rightarrow +\infty} \sum_{i=0}^n \frac{(P^{-1}AP)^n}{n!} = \exp(P^{-1}AP).$$

□

**Corollaire 3.3.12.** — Pour tout  $A \in M_d(\mathbb{K})$  on a  $\det(\exp(A)) = \exp(\text{Tr}(A))$ .

*Démonstration.* — Comme  $M_d(\mathbb{R}) \subset M_d(\mathbb{C})$ , il suffit d'établir la formule lorsque  $\mathbb{K} = \mathbb{C}$ . Dans ce cas, d'après le théorème de trigonalisation 2.2.3, il existe  $P \in GL_d(\mathbb{C})$  telle que  $T = P^{-1}AP$  soit une matrice triangulaire supérieure. Notons  $\lambda_1, \dots, \lambda_d$  ses termes diagonaux. Alors, d'une part,

$$\lambda_1 + \dots + \lambda_d = \text{Tr}(T) = \text{Tr}(A).$$

D'autre part, comme  $P^{-1} \exp(A)P = \exp(T)$ , on a  $\det(\exp(A)) = \det(\exp(T))$ . Or, d'après la proposition 3.3.10,  $\exp(T)$  est une matrice triangulaire, de termes diagonaux les  $\exp(\lambda_i)$ . On a donc

$$\det(\exp(T)) = \prod_{i=1}^d \exp(\lambda_i) = \exp(\lambda_1 + \dots + \lambda_d) = \exp(\text{Tr}(A)).$$

□



**Proposition 3.3.13.** — Si  $A, B \in M_d(\mathbb{K})$  commutent (i.e. vérifient  $AB = BA$ ), alors  $\exp(A + B) = \exp(A) \exp(B)$ . En particulier, pour tout  $t, t' \in \mathbb{K}$ , on a

$$(*) \quad \exp((t + t')A) = \exp(tA) \exp(t'A).$$

En particulier, on a  $\exp(-A) \exp(A) = \exp(-A + A) = \exp(0) = I_d$  (où  $0$  désigne la matrice nulle de  $M_d(\mathbb{K})$ ) : ceci montre que  $\exp(A)$  est inversible, d'inverse  $\exp(-A)$ .

*Démonstration.* — Comme  $A$  et  $B$  commutent, on a pour tout  $n \in \mathbb{N}$  la formule du binôme :

$$(A + B)^n = \sum_{p=0}^n \binom{n}{p} A^p B^{n-p} = \sum_{\substack{p, q \geq 0 \\ p+q=n}} \frac{n!}{p!q!} A^p B^q.$$

Donc (renvoyant à l'appendice en fin de chapitre pour la justification de l'égalité (\*) ci-dessous) on a :

$$\begin{aligned} \exp(A) \exp(B) &= \left( \sum_{p \geq 0} \frac{A^p}{p!} \right) \left( \sum_{q \geq 0} \frac{B^q}{q!} \right) \stackrel{(*)}{=} \sum_{n \geq 0} \left( \sum_{\substack{p, q \geq 0 \\ p+q=n}} \frac{A^p B^q}{p! q!} \right) \\ &= \sum_{n \geq 0} \frac{1}{n!} \left( \sum_{\substack{p, q \geq 0 \\ p+q=n}} \frac{n!}{p! q!} A^p B^q \right) = \sum_{n \geq 0} \frac{(A+B)^n}{n!} = \exp(A+B). \end{aligned}$$

La formule (\*) en découle, puisque  $tA$  et  $t'A$  commutent.  $\square$

**Remarque 3.3.14.** — Attention! Si  $AB \neq BA$ , alors  $\exp(A+B) \neq \exp(A)\exp(B)$  en général. Par exemple, soient  $A = \begin{pmatrix} 0 & -1 \\ 0 & 0 \end{pmatrix}$  et  $B = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$ . Alors  $AB = \begin{pmatrix} -1 & 0 \\ 0 & 0 \end{pmatrix} \neq BA = \begin{pmatrix} 0 & 0 \\ 0 & -1 \end{pmatrix}$ ; d'autre part, comme  $A^2 = 0 = B^2$  on a  $\exp(A) = I_2 + A$  et  $\exp(B) = I_2 + B$  et donc

$$\exp(A) \exp(B) = I_2 + (A+B) + AB = \begin{pmatrix} 0 & -1 \\ 1 & 1 \end{pmatrix} \neq \exp(B) \exp(A) = I_2 + (A+B) + BA = \begin{pmatrix} 1 & -1 \\ 1 & 0 \end{pmatrix}.$$

Donc  $\exp(A)\exp(B)$  et  $\exp(B)\exp(A)$  ne peuvent pas tous les deux être égaux à  $\exp(A+B)$ . En fait on peut montrer (cf. TE3a 2010-11) que  $\exp(A+B)$  égale  $\begin{pmatrix} \cos(1) & -\sin(1) \\ \sin(1) & \cos(1) \end{pmatrix}$ , donc est différent de  $\exp(A)\exp(B)$  et de  $\exp(B)\exp(A)$ .

**Proposition 3.3.15.** — Soit  $J$  le bloc de Jordan nilpotent  $J_d(0)$ . Pour tout bloc de Jordan  $J_d(\lambda) = \lambda I_d + J$  et tout  $t \in \mathbb{K}$ , on a

$$\exp(tJ_d(\lambda)) = e^{\lambda t} \left( I_d + tJ + \frac{t^2}{2} J^2 + \dots + \frac{t^{d-1}}{(d-1)!} J^{d-1} \right) = e^{\lambda t} \begin{pmatrix} 1 & t & \frac{t^2}{2} & \dots & \frac{t^{d-1}}{(d-1)!} \\ 0 & 1 & t & \ddots & \vdots \\ 0 & 0 & 1 & \ddots & \frac{t^2}{2} \\ \vdots & \ddots & \ddots & \ddots & t \\ 0 & \dots & \dots & 0 & 1 \end{pmatrix}$$

et pour toute matrice de Jordan  $A = \left( \begin{array}{c|c|c|c} J_{d_1}(\lambda_1) & 0 & \dots & 0 \\ \hline 0 & J_{d_2}(\lambda_2) & \ddots & \vdots \\ \hline \vdots & \ddots & \ddots & 0 \\ \hline 0 & \dots & 0 & J_{d_r}(\lambda_r) \end{array} \right)$ , on a

$$\exp(tA) = \left( \begin{array}{c|c|c|c} \exp(tJ_{d_1}(\lambda_1)) & 0 & \dots & 0 \\ \hline 0 & \exp(tJ_{d_2}(\lambda_2)) & \ddots & \vdots \\ \hline \vdots & \ddots & \ddots & 0 \\ \hline 0 & \dots & 0 & \exp(tJ_{d_r}(\lambda_r)) \end{array} \right).$$

*Démonstration.* — En effet, soit  $\mathcal{B} = (e_1, \dots, e_d)$  la base canonique de  $\mathbb{K}^d$ . Alors  $Je_1 = 0$  et  $Je_i = e_{i-1}$  pour  $i = 2, \dots, d$ , et l'on en déduit, par récurrence sur  $i = 1, \dots, d$ , que l'on a

$$J^i e_j = \begin{cases} 0 & \text{si } j = 1, \dots, i, \\ e_{j-i} & \text{si } j = i+1, \dots, d, \end{cases}$$

i.e. la matrice de  $J^i$  a tous ses coefficients nuls sauf ceux de la  $i$ -ième diagonale au-dessus de la diagonale principale, sur laquelle les coefficients valent 1. Ceci donne la forme explicite de  $\exp(tJ)$  donnée dans la proposition. De plus, comme  $tJ$  et  $\lambda t I_d$  commutent, on a

$$\exp(\lambda t I_d + tJ) = \exp(\lambda t I_d) \cdot \exp(tJ) = \exp(\lambda t) I_d \cdot \exp(tJ) = e^{\lambda t} \exp(tJ),$$

ce qui prouve la première égalité.

Lorsque  $A$  est une matrice de Jordan ayant  $r$  blocs, notons  $\mathcal{B}_i$  le sous-ensemble de  $\mathcal{B}$  correspondant au bloc  $J_{d_i}(\lambda_i)$  et  $E_i$  le sous-espace vectoriel de  $\mathbb{K}^d$  de base  $\mathcal{B}_i$ , et soit  $u_i$  l'endomorphisme de  $\mathbb{K}^d$  qui est nul

sur  $E_j$  pour  $j \neq i$  et tel que  $\text{Mat}_{\mathcal{B}_i}(u_i) = J_{d_i}(\lambda_i)$ . Alors on obtient comme plus haut que  $\text{Mat}_{\mathcal{B}}(\exp(tu_i))$  est la matrice diagonale par blocs dont tous les blocs diagonaux sont l'identité, sauf le  $i$ -ième qui est  $\exp(tJ_{d_i}(\lambda_i))$ . De plus, comme  $A = u_1 + \dots + u_r$  et que les  $u_i$  commutent deux à deux, on déduit de 3.3.13 par récurrence sur  $r$  que

$$\exp(tA) = \exp(tu_1) \exp(tu_2) \cdots \exp(tu_r)$$

et en effectuant le produit de ces matrices diagonales par blocs, on obtient bien la matrice  $\exp(tA)$  indiquée dans la proposition.  $\square$



**Théorème 3.3.16.** — Soit  $A \in M_d(\mathbb{K})$ . La fonction  $\mathbb{R} \rightarrow M_d(\mathbb{K})$ ,  $t \mapsto \exp(tA)$  est dérivable, de dérivée la fonction  $t \mapsto A \exp(tA) = \exp(tA)A$ . Par conséquent, la fonction  $t \mapsto \exp(tA)$  est de classe  $C^\infty$ , sa dérivée  $n$ -ième étant la fonction  $t \mapsto A^n \exp(tA) = \exp(tA)A^n$ .

*Démonstration.* — Pour tout  $t \in \mathbb{R}$ , notons  $S_n(t)$  la somme partielle  $\sum_{i=0}^n t^i \frac{A^i}{i!}$ . Soit  $t \in \mathbb{R}$  tel que  $|t| < 1$ , alors  $|t|^i \leq t^2$  pour tout  $i \geq 2$  et donc

$$\|S_n(t) - I_d - tA\| \leq \sum_{i=2}^n |t|^i \cdot \frac{\|A\|^i}{i!} \leq t^2 \sum_{i=2}^n \frac{\|A\|^i}{i!} \leq t^2 \exp(\|A\|).$$

Ceci étant vrai pour tout  $n$ , on en déduit que

$$(\dagger) \quad \|\exp(tA) - I_d - tA\| \leq t^2 \exp(\|A\|).$$

Ceci montre que la fonction  $t \mapsto \exp(tA)$  est dérivable (donc a fortiori continue) en  $t = 0$ , de dérivée  $A$ .

Puis, pour  $t_0 \in \mathbb{R}$  arbitraire et  $t$  comme ci-dessus (i.e.  $|t| < 1$ ), on a

$$\exp((t_0 + t)A) - \exp(t_0A) - tA \exp(t_0A) = \exp(t_0A)(\exp(tA) - I_d - tA)$$

et donc

$$(\ddagger) \quad \|\exp((t_0 + t)A) - \exp(t_0A) - tA \exp(t_0A)\| \leq t^2 \|\exp(t_0A)\| \cdot \exp(\|A\|),$$

et ceci montre que  $t \mapsto \exp(tA)$  est dérivable (donc a fortiori continue) en  $t = t_0$ , de dérivée  $A \exp(t_0A)$ . On a donc montré que la fonction  $t \mapsto \exp(tA)$  est dérivable, de dérivée la fonction  $t \mapsto A \exp(tA)$ .

De plus, pour tout  $B \in M_d(\mathbb{K})$  on obtient, en reprenant la démonstration précédente, que la fonction  $t \mapsto B \exp(tA)$  est dérivable, de dérivée  $t \mapsto BA \exp(tA)$ ; elle est donc indéfiniment dérivable, sa dérivée  $n$ -ième étant la fonction  $t \mapsto BA^n \exp(tA)$ . Le théorème est démontré.  $\square$

### 3.4. Exponentielles de matrices et équations différentielles linéaires



**Théorème 3.4.1.** — Soit toujours  $\mathbb{K} = \mathbb{R}$  ou  $\mathbb{C}$  et soient  $A \in M_n(\mathbb{K})$  et  $X : \mathbb{R} \rightarrow \mathbb{K}^n$ ,  $t \mapsto X(t) = \begin{pmatrix} x_1(t) \\ \vdots \\ x_n(t) \end{pmatrix}$  une fonction de classe  $C^\infty$  vérifiant l'équation différentielle

$$X'(t) = A \cdot X(t), \quad \text{c.-à-d.,} \quad \begin{pmatrix} x_1'(t) \\ \vdots \\ x_n'(t) \end{pmatrix} = A \begin{pmatrix} x_1(t) \\ \vdots \\ x_n(t) \end{pmatrix}.$$

Alors on a  $X(t) = \exp(tA) \cdot X(0)$  pour tout  $t \in \mathbb{R}$ .

*Démonstration.* — Pour démontrer ce théorème, on a besoin du :

**Lemme 3.4.2.** — Soient  $X : \mathbb{R} \rightarrow \mathbb{K}^n$ ,  $t \mapsto X(t)$  et  $B : \mathbb{R} \rightarrow M_n(\mathbb{K})$ ,  $t \mapsto B(t)$  des fonctions dérivables. Alors la fonction  $F : \mathbb{R} \rightarrow \mathbb{K}^n$ ,  $t \mapsto F(t) = B(t) \cdot X(t)$  est dérivable et pour tout  $t \in \mathbb{R}$  on a

$$(\dagger) \quad F'(t) = B'(t) \cdot X(t) + B(t) \cdot X'(t).$$

*Démonstration du lemme.* — En effet, on a  $F(t) = \begin{pmatrix} F_1(t) \\ \vdots \\ F_n(t) \end{pmatrix}$ , où pour tout  $i = 1, \dots, n$ ,

$$F_i(t) = \sum_{j=1}^n B(t)_{ij} X_j(t) \in \mathbb{K}.$$

Alors chaque  $F_i$  est dérivable, de dérivée  $F'_i(t) = \sum_{j=1}^n (B'(t)_{ij}X_j(t) + B(t)_{ij}X'_j(t))$  et il en résulte que  $F$  est dérivable et qu'on a bien l'égalité de vecteurs colonnes :

$$F'(t) = B'(t) \cdot X(t) + B(t) \cdot X'(t).$$

□

La démonstration du théorème est maintenant facile : considérons la fonction  $Z : \mathbb{R} \rightarrow \mathbb{K}^n$  définie par  $Z(t) = \exp(-tA) \cdot X(t)$ . D'après le lemme précédent et le théorème 3.3.16,  $Z$  est dérivable et pour tout  $t \in \mathbb{R}$  on a

$$Z'(t) = -\exp(-tA)A \cdot X(t) + \exp(-tA) \cdot X'(t) = \exp(-tA) \cdot (X'(t) - A \cdot X(t)) = 0.$$

Donc  $Z$  est constante, d'où  $Z(t) = Z(0) = X(0)$ , et donc  $X(t) = \exp(tA) \cdot X(0)$  pour tout  $t \in \mathbb{R}$ . □

**Corollaire 3.4.3.** — Soit  $A \in M_n(\mathbb{K})$ . Alors l'ensemble  $E$  des fonctions  $X : \mathbb{R} \rightarrow \mathbb{K}^n$ , de classe  $C^\infty$ , qui sont solutions de l'équation différentielle  $X' = A \cdot X$ , est un  $\mathbb{K}$ -espace vectoriel de dimension  $n$ , isomorphe à  $\mathbb{K}^n$  par l'application  $E \rightarrow \mathbb{K}^n$ ,  $X \mapsto X(0)$ .

*Démonstration.* — D'abord, il est clair que si  $X_1, X_2 \in E$  et  $\lambda \in \mathbb{K}$ , alors  $\lambda X_1 + X_2$  est encore solution de l'équation différentielle, i.e. appartient à  $E$ . Ceci montre que  $E$  est un  $\mathbb{K}$ -espace vectoriel. Il est clair que l'application  $\phi : E \rightarrow \mathbb{K}^n$ ,  $X \mapsto X(0)$  est linéaire, et elle est injective puisque  $X$  est déterminé par la « condition initiale »  $X(0)$  d'après le théorème précédent.

Réciproquement, soit  $Y \in \mathbb{K}^n$ . D'après le lemme précédent et le théorème 3.3.16, la fonction  $X : t \mapsto \exp(tA) \cdot Y$  est dérivable, de dérivée  $X'(t) = A \exp(tA) \cdot Y = A \cdot X(t)$ , donc  $X$  appartient à  $E$  et vérifie  $X(0) = Y$ . Ceci montre que  $\phi$  est aussi surjectif ; c'est donc un isomorphisme, d'où le corollaire. □

Considérons maintenant une équation différentielle linéaire :

$$(\star) \quad f^{(n)} = a_{n-1}f^{(n-1)} + \dots + a_2f'' + a_1f' + a_0f,$$

à coefficients constants  $a_i \in \mathbb{K}$  et notons  $E$  le  $\mathbb{K}$ -espace vectoriel des fonctions  $f : \mathbb{R} \rightarrow \mathbb{K}$ , de classe  $C^\infty$ , qui sont solutions de  $(\star)$ . (Lorsque les  $a_i$  sont dans  $\mathbb{R}$ , même si l'on ne s'intéresse qu'aux solutions  $f : \mathbb{R} \rightarrow \mathbb{R}$ , il peut être utile de considérer aussi les solutions à valeurs dans  $\mathbb{C}$ .)

Ceci devient un cas particulier du cas étudié plus haut, en posant pour tout  $f \in E$  :

$$X(t) = \begin{pmatrix} f(t) \\ f'(t) \\ \vdots \\ f^{(n-1)}(t) \end{pmatrix}.$$

Alors  $X : \mathbb{R} \rightarrow \mathbb{K}^n$  est de classe  $C^\infty$ , et vérifie l'équation différentielle :

$$X' = \begin{pmatrix} f'(t) \\ f''(t) \\ \vdots \\ f^{(n)}(t) \end{pmatrix} = \underbrace{\begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & 0 & 1 \\ a_0 & a_1 & \dots & a_{n-2} & a_{n-1} \end{pmatrix}}_{=A} \begin{pmatrix} f(t) \\ f'(t) \\ \vdots \\ f^{(n-1)}(t) \end{pmatrix}.$$

Réciproquement, si  $X(t) = \begin{pmatrix} x_0(t) \\ x_1(t) \\ \vdots \\ x_{n-1}(t) \end{pmatrix}$  est une fonction  $\mathbb{R} \rightarrow \mathbb{K}^n$ , de classe  $C^\infty$ , solution de l'équation

différentielle  $X' = AX$ , alors on a

$$\begin{pmatrix} x'_0(t) \\ x'_1(t) \\ \vdots \\ x'_{n-2}(t) \\ x'_{n-1}(t) \end{pmatrix} = A \begin{pmatrix} x_0(t) \\ x_1(t) \\ \vdots \\ x_{n-1}(t) \\ x_n(t) \end{pmatrix} = \begin{pmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_{n-1}(t) \\ a_0x_0(t) + \dots + a_{n-1}x_{n-1}(t) \end{pmatrix}$$

d'où  $x_1 = x'_0$ , puis  $x_2 = x'_1 = x''_0$ , etc., et donc  $x_i = x_0^{(i)}$  pour  $i = 1, \dots, n-1$ , et

$$x_0^{(n)} = x'_{n-1} = a_{n-1}x_{n-1} + \dots + a_0x_0 = a_{n-1}x_0^{(n-1)} + \dots + a_0x_0$$



donc  $f = x_0$  est solution de  $(\star)$  et  $X = \begin{pmatrix} f(t) \\ f'(t) \\ \vdots \\ f^{(n-1)}(t) \end{pmatrix}$ . On a ainsi obtenu que l'application linéaire

$$\{\text{solutions } f \text{ de } (\star)\} \rightarrow \{\text{solutions } X \text{ de } X' = AX\}, \quad f \mapsto \begin{pmatrix} f \\ f' \\ \vdots \\ f^{(n-1)} \end{pmatrix}$$

est bijective, donc un isomorphisme de  $\mathbb{K}$ -espaces vectoriels. Combinant ceci avec le théorème précédent, on obtient le :

**Théorème 3.4.4.** — Soient  $a_0, \dots, a_{n-1} \in \mathbb{K}$ . Alors le  $\mathbb{K}$ -espace vectoriel  $E$  des fonctions  $f : \mathbb{R} \rightarrow \mathbb{K}$ , de classe  $C^\infty$ , vérifiant l'équation différentielle linéaire  $(\star)$  est de dimension  $n$ . Plus précisément, l'application

$$E \rightarrow \mathbb{K}^n, \quad f \mapsto \begin{pmatrix} f(0) \\ f'(0) \\ \vdots \\ f^{(n-1)}(0) \end{pmatrix}$$
 est un isomorphisme de  $\mathbb{K}$ -espaces vectoriels. En d'autres termes : toute

solution  $f$  de  $(\star)$  est entièrement déterminée par les « données initiales »  $(f(0), f'(0), \dots, f^{(n-1)}(0))$ , et celles-ci peuvent être choisies arbitrairement.

**Exemple 3.4.5.** — Soit  $\omega \in \mathbb{R}_+^*$  et soit  $E$  l'ensemble des fonctions  $f : \mathbb{R} \rightarrow \mathbb{R}$ , de classe  $C^\infty$ , vérifiant l'équation différentielle

$$(\dagger) \quad f'' + \omega^2 f = 0.$$

D'après le théorème précédent,  $E$  est un  $\mathbb{R}$ -espace vectoriel de dimension 2. D'autre part, les fonctions  $f_1 : t \mapsto \cos(\omega t)$  et  $f_2 : t \mapsto \sin(\omega t)$  appartiennent à  $E$  (car  $f_1'(t) = -\omega \sin(\omega t)$  et  $f_2'(t) = \omega \cos(\omega t)$ ); de plus elles sont linéairement indépendantes car  $f_1(0) = 1 = f_2'(0)$  et  $f_1'(0) = 0 = f_2(0)$ , donc elles forment une base de  $E$ . Donc toute solution  $f$  de  $(\dagger)$  s'écrit de façon unique  $f(t) = a \cos(\omega t) + b \sin(\omega t)$ , avec  $a, b \in \mathbb{R}$ . De plus, posons  $A = \sqrt{a^2 + b^2}$ ; si  $(a, b) \neq (0, 0)$ , il existe  $\varphi \in \mathbb{R}$ , unique modulo  $2\pi$ , tel que  $\sin(\varphi) = a/A$  et  $\cos(\varphi) = b/A$ , et l'on a donc  $f(t) = A \sin(\omega t + \varphi)$ .

### 3.5. Espaces quotients

**3.5.0. Introduction.** — Soit  $f : V \rightarrow W$  une application linéaire *surjective*, et soit  $K = \text{Ker}(f)$ . Alors, pour tout  $w \in W$ , il existe  $v \in V$  tel que  $f(v) = w$ , et si  $v'$  a la même propriété, on a  $v' - v \in K$ , c.-à-d.,  $v' \in v + K$ . On peut donc identifier les éléments de  $W$  aux classes :

$$v + K = \{v + x \mid x \in K\} = \{v' \in V \mid v' - v \in K\}$$

Cette construction se généralise : pour tout sous-espace vectoriel  $E$  de  $V$ , on peut construire de façon canonique un espace vectoriel noté  $V/E$  et appelé « quotient de  $V$  par  $E$  », et une application linéaire surjective  $\pi : V \rightarrow V/E$  dont le noyau est  $E$ .

Cette construction a d'innombrables applications, dont certaines sont indiquées plus bas. En termes imagés, on peut dire que travailler dans  $V/E$  permet de « négliger » les éléments de  $E$ , et donc dans certains cas de rendre la situation plus simple. (Une analogie est qu'il est souvent plus facile de calculer « modulo  $n$  » que dans  $\mathbb{Z}$ ; par exemple on voit facilement que  $3^{2010} \equiv 1 \pmod{7}$  sans avoir à calculer (heureusement!)  $3^{2010}$ .)

Soit  $A$  un groupe *abélien* (c.-à-d., commutatif) et  $E$  un sous-groupe. On considère sur  $A$  la relation d'équivalence  $\sim_E$  définie par  $E$ , c.-à-d., telle que :

$$x \sim_E y \Leftrightarrow x - y \in E,$$

pour tout  $a \in A$  on note

$$a + E = \{b \in A \mid b - a \in E\} = \{a + x \mid x \in E\}$$

sa classe d'équivalence, on note  $A/E$  l'ensemble de ces classes (c.-à-d., l'ensemble *quotient de  $A$  par la relation d'équivalence  $\sim_E$* ), et  $\pi$  la projection  $A \rightarrow A/E$  qui à tout  $a \in A$  associe sa classe  $a + E$ .

On va munir  $A/E$  d'une loi de groupe abélien, telle que l'application  $\pi : A \rightarrow A/E$  soit un morphisme de groupes abéliens. Cette condition impose que l'on ait :

$$(*) \quad (a + E) + (b + E) = \pi(a) + \pi(b) = \pi(a + b) = a + b + E.$$

On voudrait donc prendre ceci comme définition de l'addition dans  $A/E$ , mais il faut vérifier que ceci a un sens. En effet, la classe  $a + E$  ne détermine pas uniquement l'élément  $a$ ; c'est aussi la classe  $a' + E$  pour tout  $a' \sim_E a$ . Il faut donc vérifier que pour tous  $a' \sim_E a$  et  $b' \sim_E b$ , le « résultat de l'addition »

$$a' + b' + E$$

est le même. Or, il existe  $x, y \in E$  tels que  $a' = a + x$  et  $b' = b + y$ , et comme l'addition dans  $A$  est commutative, on a :

$$a' + b' = a + x + b + y = a + b + (x + y),$$

et  $x + y \in E$ , ce qui montre que  $a' + b' + E = a + b + E$ . On peut donc définir une addition dans  $A/E$  par la formule (\*), et d'après cette formule, il est clair que l'addition est associative et commutative, et admet pour élément neutre la classe  $0 + E$ . Enfin, pour tout  $a \in A$ , la classe  $-a + E$  est l'opposée de  $a + E$ , puisque

$$(a + E) + (-a + E) = a - a + E = 0 + E.$$

On a donc obtenu la

**Proposition 3.5.1 (Groupes abéliens quotients).** — Soient  $A$  un groupe abélien et  $E$  un sous-groupe de  $A$ . Il existe sur l'ensemble  $A/E$  une unique loi de groupe abélien telle que la projection  $\pi : A \rightarrow A/E$  soit un morphisme de groupes abéliens; si on note  $\bar{a} = \pi(a) = a + E$ , cette loi est définie par

$$\bar{a} + \bar{b} = \overline{a + b}.$$

**Remarque 3.5.1.1.** — C'est ainsi qu'on construit les « entiers modulo  $n$  »  $\mathbb{Z}/n\mathbb{Z}$ , pour tout  $n \in \mathbb{Z}$ .

Soient maintenant  $k$  un corps,  $V$  un  $k$ -espace vectoriel,  $E$  un sous-espace vectoriel. D'après ce qui précède, l'ensemble quotient  $V/E$  est muni d'une structure de groupe abélien. On va le munir d'une structure de  $k$ -espace vectoriel, telle que l'application  $\pi : V \rightarrow V/E$  soit linéaire. Cette condition impose que l'on ait, pour tous  $t \in k$  et  $v \in V$  :

$$(**) \quad t \cdot (v + E) = t \cdot \pi(v) = \pi(t \cdot v) = t \cdot v + E.$$

On voudrait donc prendre ceci comme définition de la loi externe, mais à nouveau il faut vérifier que ceci a un sens, puisque la classe  $v + E$  ne détermine pas uniquement l'élément  $v$ ; c'est aussi la classe  $v' + E$  pour tout  $v' \sim_E v$ . Il faut donc vérifier que pour tout  $v' \sim_E v$ , on a

$$t \cdot (v' + E) = t \cdot (v + E).$$

Or il existe  $x \in E$  tel que  $v' = v + x$ , d'où  $t \cdot v' = t \cdot v + t \cdot x$ , et  $t \cdot x \in E$  puisque  $E$  est un sous-espace vectoriel de  $V$ . On a donc obtenu l'assertion (1) du théorème suivant.

**Théorème 3.5.2 (Espaces vectoriels quotients).** — Soient  $k$  un corps,  $V$  un  $k$ -espace vectoriel et  $E$  un sous-espace vectoriel de  $V$ .

(1) Il existe sur le groupe abélien  $V/E$  une unique structure d'espace vectoriel telle que la projection  $\pi : V \rightarrow V/E$  soit linéaire : pour tout  $t \in k$  et  $v \in V$ , on a  $t \cdot \bar{v} = \overline{t \cdot v}$ . On dit que  $V/E$  est l'espace vectoriel quotient de  $V$  par  $E$ .

(2) Si  $V$  est de dimension finie,  $\boxed{\dim(V) = \dim(E) + \dim(V/E)}$  i.e.  $\boxed{\dim(V/E) = \dim(V) - \dim(E)}$ .

(3) Supposons  $V$  de dimension finie, et soit  $F$  un supplémentaire de  $E$  dans  $V$  (cf. 2.1.5). Alors  $\pi$  induit un isomorphisme  $F \xrightarrow{\sim} V/E$ . En particulier, si  $\mathcal{C} = (f_1, \dots, f_d)$  est une base de  $F$ , où  $d = \dim(V) - \dim(E)$ , alors  $\pi(\mathcal{C}) = (\pi(f_1), \dots, \pi(f_d))$  est une base de  $V/E$ .

*Démonstration.* — On a déjà vu l'assertion (1), démontrons (2). L'application linéaire  $\pi : V \rightarrow V/E$  est surjective, et son noyau est  $E$ . Donc, si  $V$  est de dimension finie on a, d'après le théorème du rang :

$$\dim(V) = \dim(E) + \dim(V/E).$$

(3) Notons  $\pi_F : F \rightarrow V/E$  la restriction de  $\pi$  à  $F$ . D'abord,  $\text{Ker}(\pi_F) = F \cap E = \{0\}$  donc  $\pi_F$  est injectif. D'autre part, tout  $x \in V/E$  égale  $\pi(v)$ , pour un certain  $v \in V$ . Comme  $V = E \oplus F$ , on peut écrire  $x = e + f$  avec  $e \in E$  et  $f \in F$ , et comme  $\pi(e) = 0$  on obtient que  $x = \pi(f) = \pi_F(f)$ . Ceci montre que  $\pi_F : F \rightarrow V/E$  est aussi surjectif, donc c'est un isomorphisme. Par conséquent, l'image par  $\pi$  de toute base de  $F$  est une base de  $V/E$ .  $\square$

Le quotient  $V/E$  a la « propriété universelle » suivante, relativement aux applications linéaires  $f : V \rightarrow W$  qui s'annulent sur  $E$ , c.-à-d., pour une telle application  $f$ , on peut remplacer  $V$  par l'espace « plus simple » (car de dimension plus petite)  $V/E$  :

**Théorème 3.5.3 (Passage au quotient d'un homomorphisme).** — Soient  $V, W$  des espaces vectoriels,  $E$  un sous-espace vectoriel de  $V$ ,  $\pi$  la projection  $V \rightarrow V/E$ , et  $f : V \rightarrow W$  une application linéaire telle que  $E \subseteq \text{Ker}(f)$ . Alors l'application

$$V/E \rightarrow W, \quad \pi(v) \mapsto f(v)$$

est bien définie et est l'unique application linéaire  $\bar{f} : V/E \rightarrow W$  telle que  $\bar{f} \circ \pi = f$ .

*Démonstration.* — Si l'on a une application linéaire  $\bar{f} : V/E \rightarrow W$  telle que  $\bar{f} \circ \pi = f$  alors pour tout  $v \in V$  on a nécessairement

$$(*) \quad \bar{f}(\pi(v)) = f(v).$$

Réciproquement, montrons que cette formule définit bien une application de  $V/E$  dans  $W$ . Il faut montrer que si  $\pi(v) = \pi(v')$ , c.-à-d., si  $v - v' \in E$ , alors  $f(v) = f(v')$ . Mais ceci est clair puisque  $E \subseteq \text{Ker}(f)$ . Donc  $\bar{f}$  est bien définie. Alors, pour tout  $v, v' \in V$  et  $t \in k$ , on a :

$$\bar{f}(t \cdot \pi(v) + \pi(v')) = \bar{f}(\pi(t \cdot v + v')) = f(t \cdot v + v') = t \cdot f(v) + f(v') = t \cdot \bar{f}(v) + \bar{f}(v')$$

ce qui montre que  $\bar{f}$  est linéaire. □



**Théorème 3.5.4 (Théorème d'isomorphisme de Noether).** — Soient  $V, W$  des espaces vectoriels,  $f : V \rightarrow W$  une application linéaire. Alors  $f$  induit un isomorphisme d'espaces vectoriels

$$\bar{f} : V/\text{Ker}(f) \xrightarrow{\sim} \text{Im}(f).$$

*Démonstration.* — On peut considérer  $f$  comme un homomorphisme  $f : V \rightarrow \text{Im}(f)$ , qui est alors surjectif. D'après le théorème 3.5.3, il existe donc un (unique) homomorphisme  $\bar{f} : V/\text{Ker}(f) \rightarrow \text{Im}(f)$  tel que  $\bar{f}(\bar{v}) = f(v)$  pour tout  $v \in V$ , et  $\bar{f}$  est, comme  $f$ , surjectif.

De plus,  $\bar{f}$  est injectif : en effet, soit  $x \in \text{Ker}(\bar{f})$  et soit  $v \in V$  tel que  $\bar{v} = x$ , alors  $0 = \bar{f}(\bar{v}) = f(v)$ , donc  $v \in \text{Ker}(f)$ , d'où  $x = 0$ . Donc  $\bar{f} : V/\text{Ker}(f) \rightarrow \text{Im}(f)$  est injectif et surjectif, donc c'est un isomorphisme. □

Pour illustrer l'usage des espaces quotients, terminons ce paragraphe 3.5 en donnant une autre démonstration <sup>(3)</sup> du théorème de trigonalisation ; on y remplace aussi l'hypothèse que le corps de base soit  $\mathbb{C}$  par l'hypothèse que le polynôme caractéristique  $P_u(X)$  soit scindé. Commençons par la proposition suivante, qui établit un lien entre espaces quotients et matrices triangulaires par blocs.

**Proposition 3.5.5 (Quotients et matrices triangulaires par blocs)**

Soient  $V$  un  $k$ -espace vectoriel de dimension  $n$ ,  $u \in \text{End}_k(V)$ , et  $E$  un sous-espace de  $V$  stable par  $u$ . Soit  $u_E : E \rightarrow E$  la restriction de  $u$  à  $E$  et soit  $\pi$  la projection  $V \rightarrow V/E$ .

(1)  $u$  induit un endomorphisme  $\bar{u} = u_{V/E}$  de  $V/E$ , tel que  $u_{V/E}(\pi(v)) = \pi(u(v))$  pour tout  $v \in V$ .

(2) Soit  $\mathcal{B} = (e_1, \dots, e_n)$  une base de  $V$  telle que  $\mathcal{C} = (e_1, \dots, e_r)$  soit une base de  $E$ . Alors  $\text{Mat}_{\mathcal{B}}(u)$  est triangulaire par blocs :

$$\text{Mat}_{\mathcal{B}}(u) = \left( \begin{array}{c|c} A & B \\ \hline \mathbf{0}_{n-r,r} & D \end{array} \right) \quad \text{avec} \quad \begin{cases} A = \text{Mat}_{\mathcal{C}}(u_E) \\ B \in M_{r,n-r}(k), \quad D \in M_{n-r}(k). \end{cases}$$

De plus,  $D$  est la matrice de  $u_{V/E}$  dans la base  $\mathcal{D} = (\pi(e_{r+1}), \dots, \pi(e_n))$  de  $V/E$ .

(3) Par conséquent, on a

$$(*) \quad P_u(X) = P_{u_E}(X) \cdot P_{u_{V/E}}(X).$$

<sup>(3)</sup> en quelque sorte « duale » de celle de 2.2.3

*Démonstration.* — (1) Comme  $u(E) \subseteq E$ , on a  $(\pi \circ u)(E) = \{0\}$  donc l'existence de  $\bar{u} = u_{V/E}$  résulte du théorème 3.5.3 appliqué à  $f = \pi \circ u : V \rightarrow V/E$ .

(2) Comme  $E = \text{Vect}(e_1, \dots, e_r)$  est stable par  $u$ , il est clair que  $M = \text{Mat}_{\mathcal{B}}(u)$  a la forme indiquée. De plus, écrivant  $B = (b_{ij})$  et  $D = (d_{\ell j})$  pour  $i = 1, \dots, r$  et  $j, \ell = 1, \dots, n - r$ , on a :

$$u(e_{r+j}) = \sum_{i=1}^r b_{ij} e_i + \sum_{\ell=1}^r d_{\ell j} e_{r+\ell}$$

donc  $\bar{u}(\pi(e_{r+j})) = \pi(u(e_{r+j}))$  égale  $\sum_{\ell=1}^r d_{\ell j} \pi(e_{r+\ell})$  pour  $j = 1, \dots, n - r$ , ce qui prouve que  $D$  est la matrice de  $\bar{u}$  dans la base  $\mathcal{D} = (\pi(e_{r+1}), \dots, \pi(e_n))$  de  $V/E$  (cf. point (3) de 3.5.2). Alors, comme

$$M - XI_n = \left( \begin{array}{c|c} A - XI_r & B \\ \hline \mathbf{0}_{n-r,r} & D - XI_{n-r} \end{array} \right)$$

on déduit de 1.4.6 que  $P_u(X) = \det(A - XI_r) \cdot \det(D - XI_{n-r})$  d'où  $P_u(X) = P_{u_E}(X) \cdot P_{u_{V/E}}(X)$ . La proposition est démontrée.  $\square$



**Théorème 3.5.6 (Trigonalisation lorsque  $P_u(X)$  est scindé).** — Soient  $V$  un  $k$ -espace vectoriel de dimension  $n$  et  $u \in \text{End}_k(V)$ , on suppose que  $P_u(X)$  est scindé, i.e. qu'il a  $n$  racines  $\lambda_1, \dots, \lambda_n$  (pas nécessairement distinctes) dans  $k$ . Alors il existe une base  $\mathcal{B}$  de  $V$  dans laquelle la matrice de  $u$  est triangulaire supérieure, les coefficients diagonaux étant alors  $\lambda_1, \dots, \lambda_n$ .

En termes matriciels : si le polynôme caractéristique de  $A \in M_n(k)$  a toutes ses racines dans  $k$ , alors  $A$  est semblable à une matrice triangulaire, i.e. il existe  $P \in \text{GL}_n(k)$  telle que  $P^{-1}AP$  soit triangulaire.

*Démonstration.* — On procède par récurrence sur  $n$ , il n'y a rien à montrer si  $n = 1$ . Supposons donc  $n \geq 2$  et le théorème établi pour  $n - 1$ . Puisque, par hypothèse,  $P_u(X)$  est scindé, alors  $u$  possède au moins un vecteur propre  $e_1$  associé à la valeur propre  $\lambda_1$ . La droite  $ke_1$  est stable par  $u$ , donc d'après le corollaire 3.5.5,  $u$  induit un endomorphisme  $\bar{u}$  du quotient  $\bar{V} = V/ke_1$ , qui est de dimension  $n - 1$ . D'après la formule (\*) de 3.5.5, on a

$$P_u(X) = (\lambda_1 - X)P_{\bar{u}}(X)$$

et donc  $P_{\bar{u}}(X)$  égale  $(-1)^{n-1} \prod_{i=2}^n (X - \lambda_i)$ , donc est scindé.

Par hypothèse de récurrence, il existe une base  $\mathcal{D} = (v_2, \dots, v_n)$  de  $\bar{V}$  telle que  $D = \text{Mat}_{\mathcal{D}}(\bar{u})$  soit triangulaire supérieure. Soient  $e_2, \dots, e_n \in V$  dont les images dans  $V/ke_1$  sont  $v_2, \dots, v_n$ . Alors  $\mathcal{B} = (e_1, e_2, \dots, e_n)$  est une base de  $V$  (cf. la preuve du théorème du rang 0.3.2) et, d'après 3.5.5, la matrice  $M$  de  $u$  dans  $\mathcal{B}$  est de la forme :

$$\left( \begin{array}{c|c} \lambda_1 & B \\ \hline \mathbf{0}_{n-1,1} & D \end{array} \right)$$

avec  $B \in M_{1,n-1}(k)$ , donc est triangulaire supérieure. Le théorème est démontré.  $\square$

### 3.6. Appendice (†) : Normes sur $\mathbb{K}^n$ et produits de séries absolument convergentes

Soit  $\mathbb{K} = \mathbb{R}$  ou  $\mathbb{C}$ . Commençons par démontrer la proposition 3.3.3 :

**Proposition 3.6.1.** — Soit  $E$  un  $\mathbb{K}$ -espace vectoriel de dimension finie. Alors toutes les normes sur  $E$  sont équivalentes, c.-à-d., si  $\|\cdot\|$  et  $\|\cdot\|'$  sont deux normes sur  $E$ , il existe des constantes  $c, C \in \mathbb{R}_+^*$  telles que :

$$(\dagger) \quad \forall v \in E, \quad c \cdot \|v\| \leq \|v\|' \leq C \cdot \|v\|.$$

*Démonstration.* — Soit  $\mathcal{B} = (e_1, \dots, e_d)$  une base de  $E$  et soit  $\|\cdot\|_\infty$  la norme sur  $E$  définie par  $\|v\|_\infty = \text{Max}_i |x_i|$  si  $v = \sum_i x_i e_i$ . Il suffit de montrer qu'il existe des constantes  $c, C \in \mathbb{R}_+^*$  telles que :

$$(\star) \quad \forall v \in E, \quad c \cdot \|v\|_\infty \leq \|v\| \leq C \cdot \|v\|_\infty.$$

En effet, on aura de même  $c' \cdot \|v\|_\infty \leq \|v\|' \leq C' \cdot \|v\|_\infty$  pour des constantes  $c', C' > 0$ , et donc :

$$\forall v \in E, \quad \frac{c'}{C} \cdot \|v\| \leq \|v\|' \leq \frac{C'}{c} \cdot \|v\|.$$

La seconde inégalité de  $(\star)$  s'obtient facilement : posant  $C = \sum_{i=1}^d \|e_i\|$ , on a

$$(*) \quad \left\| \sum_{i=1}^d x_i e_i \right\| \leq \sum_{i=1}^d |x_i| \cdot \|e_i\| \leq C \cdot \|v\|_\infty.$$

Montrons maintenant qu'il existe  $m > 0$  tel que, pour tout  $v \in E$ , on ait

$$(\ddagger) \quad \|v\| = 1 \quad \Rightarrow \quad \|v\|_\infty \leq m.$$

Dans le cas contraire, il existerait une suite  $(v_p)_{p \in \mathbb{N}}$  de vecteurs tels que  $\|v_p\| = 1$  et que  $t_p = \|v_p\|_\infty$  tende vers  $+\infty$ . Quitte à remplacer  $(v_p)_{p \in \mathbb{N}}$  par une sous-suite et à permuter les coordonnées  $(x_1, \dots, x_d)$ , on peut supposer que

$$\frac{1}{t_p} v_p = (1, x_{p,2}, \dots, x_{p,d})$$

avec  $|x_{p,i}| \leq 1$ . Comme le « cube unité » de  $E$  pour la norme  $\|\cdot\|_\infty$  est compact, on peut supposer, quitte à remplacer  $(v_p)_{p \in \mathbb{N}}$  par une sous-suite, que la suite  $(v_p)_{p \in \mathbb{N}}$  converge pour la norme  $\|\cdot\|_\infty$  vers une limite  $\ell$ . Comme la fonction « 1ère coordonnée » est continue pour la norme  $\|\cdot\|_\infty$  (c'est clair), et comme  $\|\cdot\|$  l'est aussi (d'après (\*)), alors la 1ère coordonnée de  $\ell$  est 1 et d'autre part on a  $\|\ell\| = \lim_{p \rightarrow +\infty} \frac{1}{t_p} = 0$ , donc  $\ell = 0$ , d'où une contradiction. Ceci prouve  $(\ddagger)$ . Alors, pour tout  $v \neq 0$ , appliquant  $(\ddagger)$  à  $v' = \frac{1}{\|v\|} v$ , on obtient que  $\|v\|_\infty \leq m \cdot \|v\|$ , inégalité qui est encore valable pour  $v = 0$ . On a donc :

$$\forall v \in E, \quad \frac{1}{m} \cdot \|v\|_\infty \leq \|v\|,$$

ce qui achève la démonstration.  $\square$

Démontrons maintenant la proposition suivante, utilisée dans la démonstration de la proposition 3.3.13.

**Proposition 3.6.2.** — Soient  $\sum_{i=0}^\infty A_i$  et  $\sum_{j=0}^\infty B_j$  deux séries convergentes dans  $M_d(\mathbb{K})$ . Supposons de plus que  $\sum_{i=0}^\infty A_i$  soit absolument convergente, i.e. que la série  $\sum_{i=0}^\infty \|A_i\|$  converge.

Alors, posant  $C_p = \sum_{i=0}^p A_i B_{p-i}$ , la série  $\sum_{p=0}^\infty C_p$  converge vers le produit  $(\sum_{i=0}^\infty A_i)(\sum_{j=0}^\infty B_j)$ .

*Démonstration.* — Notons  $(S_n)$  (resp.  $(T_n)$ ) la suite des sommes partielles  $\sum_{i=0}^n A_i$  (resp.  $\sum_{j=0}^n B_j$ ) et  $S$  (resp.  $T$ ) sa limite. Comme

$$\| \|ST - S_n T\| \| \leq \| \|S - S_n\| \| \cdot \| \|T\| \|,$$

la suite  $S_n T$  converge vers  $ST$ . Posant

$$U_n = \sum_{p=0}^{2n} C_p = \sum_{p=0}^{2n} \sum_{i=0}^p A_i B_{p-i} = \sum_{i=0}^{2n} A_i \left( \sum_{j=0}^{2n-i} B_j \right),$$

il suffit donc de montrer que la suite  $S_n T - U_n$  tend vers 0. Or on a

$$\begin{aligned} \| \|S_n T - U_n\| \| &\leq \sum_{i=0}^n \| \|A_i (T - \sum_{j=0}^{2n-i} B_j)\| \| + \sum_{i=n+1}^{2n} \| \|A_i (\sum_{j=0}^{2n-i} B_j)\| \| \\ &\leq \underbrace{\sum_{i=0}^n \| \|A_i\| \| \cdot \| \|T - \sum_{j=0}^{2n-i} B_j\| \|}_{M_n} + \underbrace{\sum_{i=n+1}^{2n} \| \|A_i\| \| \cdot \| \|\sum_{j=0}^{2n-i} B_j\| \|}_{M'_n}. \end{aligned}$$


Posons  $a = \sum_{i=0}^\infty \| \|A_i\| \|$  et remarquons que comme la suite  $T_n$  converge vers  $T$ , la suite des normes  $\| \|T_n\| \|$  est bornée par un réel  $C > 0$ . Alors, pour tout  $\varepsilon > 0$ , il existe  $n_0$  tel que  $\| \|T - \sum_{j=0}^n B_j\| \| < \frac{\varepsilon}{2a}$  et  $\sum_{i=n+1}^{2n} \| \|A_i\| \| < \frac{\varepsilon}{2C}$  pour tout  $n \geq n_0$ , d'où  $M_n + M'_n < \varepsilon$  pour tout  $n \geq n_0$ . Ceci prouve que  $\| \|S_n T - U_n\| \|$  tend vers 0, d'où la proposition.  $\square$



## CHAPITRE 4

### APPLICATIONS MULTILINÉAIRES ET FORMES $p$ -LINÉAIRES ALTERNÉES

**Résumé :** Ce chapitre est probablement le plus difficile du cours. On y introduit les formes  $p$ -linéaires, et notamment les formes  $p$ -linéaires alternées, dont le déterminant est un exemple. Il faut retenir de ce chapitre : la section 4 sur le groupe symétrique (décomposition en cycles et signature d'une permutation), la formule explicite 4.4.6 pour le déterminant, et la proposition 4.4.9 exprimant le rang d'une matrice en fonction de ses mineurs.

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.

#### 4.1. Applications multilinéaires

##### **Définition 4.1.1 (Espace des applications à valeurs dans un espace vectoriel)**

Soient  $X$  un ensemble et  $F$  un  $k$ -espace vectoriel, on note  $\text{Applic}(X, F)$  l'ensemble de toutes les applications  $X \rightarrow F$ . Si  $f, g \in \text{Applic}(X, F)$  et  $\lambda \in k$ , on définit l'application  $\lambda \cdot f + g$  par

$$(\lambda \cdot f + g)(x) = \lambda \cdot f(x) + g(x)$$


pour tout  $x \in X$  ; on vérifie facilement que ceci munit  $\text{Applic}(X, F)$  d'une structure de  $k$ -espace vectoriel.

##### **Définition 4.1.2 (Applications et formes $p$ -linéaires).** — Soit $p$ un entier $\geq 2$ .

(1) Soient  $E_1, \dots, E_p$  et  $F$  des  $k$ -espaces vectoriels ; une application

$$f : E_1 \times \dots \times E_p \rightarrow F$$


est dite  $p$ -linéaire ou *multilinéaire* (si  $p = 2$ , on dit *bilinéaire*) si elle est *linéaire par rapport à chacune des variables* (les autres variables étant fixées), c.-à-d., si pour tout  $i = 1, \dots, p$ , on a :

 (\*<sub>i</sub>)  $f(x_1, \dots, t_i \cdot x_i + x'_i, \dots, x_p) = t_i \cdot f(x_1, \dots, x_i, \dots, x_p) + f(x_1, \dots, x'_i, \dots, x_p)$

où chaque  $x_j$  est dans  $E_j$  (et  $x'_i \in E_i$ ), et  $t_i \in k$ . On notera  $\mathcal{L}(E_1, \dots, E_p; F)$  l'ensemble de ces applications, et si tous les  $E_i$  sont égaux à un même espace  $E$ , on le notera simplement  $\mathcal{L}_p(E, F)$ .

Si  $f, g \in \mathcal{L}(E_1, \dots, E_p; F)$  et  $\lambda \in k$ , on voit aisément que l'application  $\lambda \cdot f + g : E_1 \times \dots \times E_p \rightarrow F$  vérifie (\*<sub>i</sub>) pour tout  $i$  ; ceci montre que  $\mathcal{L}(E_1, \dots, E_p; F)$  est un  $k$ -espace vectoriel (c'est un sous-espace vectoriel de  $\text{Applic}(E_1 \times \dots \times E_p, F)$ ).

(2) Lorsque l'espace d'arrivée  $F$  est le corps  $k$ , un élément de  $\mathcal{L}_p(E, k)$ , i.e. une application  $p$ -linéaire  $\underbrace{E \times \dots \times E}_p \text{ facteurs} \rightarrow k$ , s'appelle une *forme  $p$ -linéaire* sur  $E$  (et *forme bilinéaire* si  $p = 2$ ).

 **Remarque 4.1.3.** — Il résulte de la définition que  $f(x_1, \dots, x_n) = 0$  s'il existe un indice  $i$  tel que  $x_i = 0$  (puisque  $f$  est linéaire par rapport à la  $i$ -ème variable).

**Remarque 4.1.4.** — Attention, si l'on considère, pour  $p \geq 2$ , l'espace vectoriel

$$E_1 \times \dots \times E_p = E_1 \oplus \dots \oplus E_p$$



une application multilinéaire  $E_1 \times \cdots \times E_p \rightarrow F$  n'est pas la même chose qu'une application linéaire  $E_1 \oplus \cdots \oplus E_p \rightarrow F$ . Par exemple, si  $t \in k$  on a :

$$f(tx_1, \dots, tx_p) = \begin{cases} t f(x_1, \dots, x_p) & \text{si } f \text{ est linéaire;} \\ t^p f(x_1, \dots, x_p) & \text{si } f \text{ est multilinéaire;} \end{cases}$$

d'autre part, prenant  $p = 2$  pour simplifier,  $f(x_1 + y_1, x_2 + y_2)$  égale :

$$\begin{cases} f(x_1, x_2) + f(y_1, y_2) & \text{si } f \text{ est linéaire;} \\ f(x_1, x_2) + f(x_1, y_2) + f(y_1, x_2) + f(y_1, y_2) & \text{si } f \text{ est multilinéaire.} \end{cases}$$



**Exemple 4.1.5 (Déterminant).** — Soient  $V$  un  $k$ -espace vectoriel de dimension  $n$  et  $\mathcal{B} = (e_1, \dots, e_n)$  une base de  $V$ . Considérons l'application « déterminant relativement à la base  $\mathcal{B}$  »

$$\det_{\mathcal{B}} : \underbrace{V \times \cdots \times V}_{n \text{ facteurs}} \rightarrow k, \quad (v_1, \dots, v_n) \mapsto \det_{\mathcal{B}}(v_1, \dots, v_n)$$

c.-à-d., on écrit chaque  $v_j$  dans la base  $\mathcal{B}$  :  $v_j = \sum_{i=1}^n a_{ij}e_i$ , d'où une matrice  $A = (a_{ij})_{1 \leq i, j \leq n}$ , alors  $\det_{\mathcal{B}}(v_1, \dots, v_n) = \det(A)$ . Alors, l'application

$$\det_{\mathcal{B}} : V \times \cdots \times V \rightarrow k, \quad (v_1, \dots, v_n) \mapsto \det_{\mathcal{B}}(v_1, \dots, v_n)$$

est  $n$ -linéaire. De plus, elle est alternée, c.-à-d., elle vérifie  $\det_{\mathcal{B}}(v_1, \dots, v_n) = 0$  si deux des  $v_j$  sont égaux. On a vu au Chap. 2 que ceci entraîne que  $\det_{\mathcal{B}}$  est une application antisymétrique, c.-à-d.,  $\det_{\mathcal{B}}(v_1, \dots, v_n)$  change de signe si l'on échange les places de deux vecteurs :

$$\det_{\mathcal{B}}(v_1, \dots, v_i, \dots, v_j, \dots, v_n) = -\det_{\mathcal{B}}(v_1, \dots, v_j, \dots, v_i, \dots, v_n).$$

(voir aussi l'exemple 4.1.7 plus bas).



**Exemple 4.1.6 (Produit scalaire sur  $\mathbb{R}^3$ ).** — Soit  $\phi : \mathbb{R}^3 \times \mathbb{R}^3 \rightarrow \mathbb{R}$  le produit scalaire « euclidien » usuel :

$$\phi(X, Y) = x_1y_1 + x_2y_2 + x_3y_3 \quad \text{si} \quad X = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}, \quad Y = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}$$

alors  $\phi : \mathbb{R}^3 \times \mathbb{R}^3 \rightarrow \mathbb{R}$  est une application bilinéaire, c.-à-d.,

$$\phi(X + sX', Y + tY') = \phi(X, Y) + s\phi(X', Y) + t\phi(X, Y') + st\phi(X', Y').$$

Comme  $\phi$  est à valeurs dans le corps des scalaires (ici  $\mathbb{R}$ ), on dit que  $\phi$  est une forme bilinéaire sur  $\mathbb{R}^3$ . De plus, on dit que  $\phi$  est symétrique car

$$\phi(X, Y) = \phi(Y, X), \quad \forall X, Y \in \mathbb{R}^3.$$

Plus généralement, pour tout  $n \geq 1$ , le produit scalaire « euclidien » sur  $\mathbb{R}^n$  défini par

$$\phi(X, Y) = x_1y_1 + \cdots + x_ny_n \quad \text{si} \quad X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \quad Y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

est une forme bilinéaire symétrique sur  $\mathbb{R}^n$ .

**Exemple 4.1.7 (Produit vectoriel dans  $\mathbb{R}^3$  « euclidien »).** — Le produit vectoriel  $\wedge : \mathbb{R}^3 \times \mathbb{R}^3 \rightarrow \mathbb{R}^3$  défini par

$$(*) \quad X \wedge Y = \begin{pmatrix} x_2y_3 - x_3y_2 \\ x_3y_1 - x_1y_3 \\ x_1y_2 - x_2y_1 \end{pmatrix} \quad \text{si} \quad X = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}, \quad Y = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}$$

est une application bilinéaire  $\mathbb{R}^3 \times \mathbb{R}^3 \rightarrow \mathbb{R}^3$  alternée, c.-à-d.,  $X \wedge X = 0$  pour tout  $X$ . Pour  $X, Y \in \mathbb{R}^3$  arbitraires, ceci entraîne :  $0 = (X + Y) \wedge (X + Y) = X \wedge X + Y \wedge Y + X \wedge Y + Y \wedge X$  d'où

$$X \wedge Y = -Y \wedge X, \quad \forall X, Y \in \mathbb{R}^3$$

donc le produit vectoriel est antisymétrique (ce qu'on voit aussi directement d'après la définition (\*)).

*Remarque.* Cette définition du produit vectoriel  $\wedge$  ne fait pas apparaître le produit scalaire euclidien, mais utilise les coordonnées de  $X, Y$  dans la base canonique; on verra plus loin une définition de  $\wedge$  en termes du produit scalaire, qui n'utilise pas les coordonnées.

**Exemple 4.1.8 (Couplage  $V^* \times V \rightarrow k$ ).** — Soient  $V$  un  $k$ -espace vectoriel,  $V^*$  l'espace dual. Alors l'application  $V^* \times V \rightarrow k$ ,  $(f, v) \mapsto f(v)$  est bilinéaire.





**Théorème 4.1.9 (Bases et dimension de  $\mathcal{L}_p(E, k)$ ).** — Soient  $p$  un entier  $\geq 2$  et  $E$  un  $k$ -espace vectoriel de dimension  $d$ . Notons  $P$  l'ensemble produit  $\{1, \dots, d\} \times \dots \times \{1, \dots, d\}$  ( $p$  facteurs), et soit  $\mathcal{B} = (e_1, \dots, e_d)$  une base de  $E$ .

(i) Tout élément  $\phi \in \mathcal{L}_p(E, k)$  est entièrement déterminé par ses valeurs  $\lambda_{(i_1, \dots, i_p)} = \phi(e_{i_1}, \dots, e_{i_p}) \in k$ , et ces éléments peuvent être choisis arbitrairement.

(ii) Par conséquent, les éléments  $\phi^{a_1, \dots, a_p}$ , pour  $(a_1, \dots, a_p) \in P$ , définis par :

$$\forall (i_1, \dots, i_p), \quad \phi^{a_1, \dots, a_p}(e_{i_1}, \dots, e_{i_p}) = \begin{cases} 1 & \text{si } i_1 = a_1, \dots, i_p = a_p \\ 0 & \text{sinon} \end{cases}$$

forment une base de  $\mathcal{L}_p(E, k)$ ; on a donc  $\dim \mathcal{L}_p(E, k) = d^p = (\dim E)^p$ .

*Démonstration.* — Soit  $\phi \in \mathcal{L}_p(E, k)$ . Pour tout  $(v_1, \dots, v_p) \in E \times \dots \times E$ , chaque  $v_q$  s'écrit de façon unique

$$(0) \quad v_q = t_{q,1} e_1 + \dots + t_{q,d} e_d, \quad \text{avec } t_{q,i} \in k,$$

donc, en appliquant les égalités  $(*_1), \dots, (*_p)$  de 4.1.2, on obtient la formule :

$$(1) \quad \begin{aligned} \phi(v_1, \dots, v_p) &= \phi(t_{1,1} e_1 + \dots + t_{1,d} e_d, \dots, t_{p,1} e_1 + \dots + t_{p,d} e_d) \\ &= \sum_{(i_1, \dots, i_p) \in P} t_{1,i_1} \dots t_{p,i_p} \phi(e_{i_1}, \dots, e_{i_p}). \end{aligned}$$

Ceci montre que  $\phi$  est entièrement déterminée par les  $D = d^p$  scalaires  $\phi(e_{i_1}, \dots, e_{i_p}) \in k$ ; en d'autres termes, l'application « d'évaluation sur les  $p$ -uplets  $(e_{i_1}, \dots, e_{i_p})$  » :

$$\text{ev} : \mathcal{L}_p(E, k) \rightarrow k^D, \quad \phi \mapsto \left( \phi(e_{i_1}, \dots, e_{i_p}) \right)_{(i_1, \dots, i_p) \in P}$$

est *injective*. D'autre part, cette application est *linéaire* : si  $\mu \in k$  et  $\phi, \psi \in \mathcal{L}_p(E, k)$ , on a

$$(\mu\phi + \psi)(v_1, \dots, v_p) = \mu\phi(v_1, \dots, v_p) + \psi(v_1, \dots, v_p)$$

pour tout  $(v_1, \dots, v_p) \in E \times \dots \times E$ , donc *a fortiori* pour tout  $p$ -uplet  $(e_{i_1}, \dots, e_{i_p})$ .

Réciproquement, pour tout  $(a_1, \dots, a_p) \in P$ , notons  $\phi^{a_1, \dots, a_p}$  l'application  $E \times \dots \times E \rightarrow k$  définie comme suit : pour tout  $p$ -uplet  $(v_1, \dots, v_p)$  comme en (0) ci-dessus,

$$(2) \quad \phi^{a_1, \dots, a_p}(v_1, \dots, v_p) = t_{1,a_1} \dots t_{p,a_p} \in k.$$

Fixons un indice  $q \in \{1, \dots, p\}$ , et montrons que  $\phi^{a_1, \dots, a_p}$  est linéaire par rapport à la  $q$ -ème variable, i.e. que, pour tout  $\lambda \in k$  et  $w \in E$ , on a

$$(*_q) \quad \phi^{a_1, \dots, a_p}(v_1, \dots, \lambda v_q + w, \dots, v_p) = \lambda \phi^{a_1, \dots, a_p}(v_1, \dots, v_q, \dots, v_p) + \phi^{a_1, \dots, a_p}(v_1, \dots, w, \dots, v_p).$$

Écrivons  $w = s_1 e_1 + \dots + s_d e_d$ , avec  $s_i \in k$ . Alors  $\lambda v_q + w = t'_{q,1} e_1 + \dots + t'_{q,d} e_d$ , avec  $t'_{q,i} = \lambda t_{q,i} + s_i$ , et donc la formule (2) donne :

$$\begin{aligned} \phi^{a_1, \dots, a_p}(v_1, \dots, \lambda v_q + w, \dots, v_p) &= t_{1,a_1} \dots t'_{q,a_q} \dots t_{p,a_p} = t_{1,a_1} \dots (\lambda t_{q,a_q} + s_{a_q}) \dots t_{p,a_p} \\ &= \lambda t_{1,a_1} \dots t_{q,a_q} \dots t_{p,a_p} + t_{1,a_1} \dots s_{a_q} \dots t_{p,a_p} \\ &= \lambda \phi^{a_1, \dots, a_p}(v_1, \dots, v_q, \dots, v_p) + \phi^{a_1, \dots, a_p}(v_1, \dots, w, \dots, v_p) \end{aligned}$$

ce qui montre que  $(*_q)$  est vérifiée, d'où  $\phi^{a_1, \dots, a_p} \in \mathcal{L}_p(E, k)$ . Remarquons de plus que, pour tout  $(i_1, \dots, i_p) \in P$ , l'élément  $(v_1, \dots, v_p) = (e_{i_1}, \dots, e_{i_p})$  de  $E \times \dots \times E$  correspond dans l'écriture (0) à :

$$\forall q = 1, \dots, p, \quad t_{q,j_q} = \begin{cases} 1 & \text{si } j_q = i_q \\ 0 & \text{sinon} \end{cases}$$

donc (2) donne, en particulier :

$$(2') \quad \forall (i_1, \dots, i_p) \in P, \quad \phi^{a_1, \dots, a_p}(e_{i_1}, \dots, e_{i_p}) = \begin{cases} 1 & \text{si } i_1 = a_1, \dots, i_p = a_p \\ 0 & \text{sinon.} \end{cases}$$

Maintenant, pour tout  $\Lambda = (\lambda_{(j_1, \dots, j_p)})_{(j_1, \dots, j_p) \in P} \in k^D$ , l'élément  $\phi_\Lambda = \sum_{(j_1, \dots, j_p) \in P} \lambda_{(j_1, \dots, j_p)} \phi^{(j_1, \dots, j_p)}$  appartient à  $\mathcal{L}_p(E, k)$  et, d'après (2') il vérifie, pour tout  $(i_1, \dots, i_p) \in P$  :

$$(3) \quad \phi_\Lambda(e_{i_1}, \dots, e_{i_p}) = \sum_{(j_1, \dots, j_p) \in P} \lambda_{(j_1, \dots, j_p)} \underbrace{\phi^{(j_1, \dots, j_p)}(e_{i_1}, \dots, e_{i_p})}_{\substack{=1 \text{ si } (j_1, \dots, j_p) = (i_1, \dots, i_p) \\ =0 \text{ sinon}}} = \lambda_{(i_1, \dots, i_p)}.$$

Ceci prouve que l'application linéaire injective  $\text{ev} : \mathcal{L}_p(E, k) \rightarrow k^D$  est aussi *surjective*, donc est un isomorphisme; en particulier,  $\dim \mathcal{L}_p(E, k) = D = d^p$ . Ceci prouve déjà l'assertion (i).

De plus, (3) implique que si l'on a une relation  $\sum_{(a_1, \dots, a_p) \in P} \lambda_{(a_1, \dots, a_p)} \phi^{(a_1, \dots, a_p)} = 0$ , alors chaque  $\lambda_{(a_1, \dots, a_p)}$  est nul. La famille  $\mathcal{F} = (\phi^{(a_1, \dots, a_p)})_{(a_1, \dots, a_p) \in P}$  est donc *libre*, et comme elle est de cardinal  $d^p$ , c'est une base de  $\mathcal{L}_p(E, k)$ . Ceci achève la preuve de l'assertion (ii). Le théorème est démontré.  $\square$

**Remarque 4.1.9.1.** — On conserve les notations de 4.1.9. Attention! Pour  $p \geq 2$ , l'ensemble produit  $\mathcal{B} \times \dots \times \mathcal{B}$  ( $p$  facteurs) n'a rien à voir avec l'espace vectoriel  $V = E \times \dots \times E = E \oplus \dots \oplus E$  ( $p$  facteurs)! En effet, d'un côté,  $V$  est de dimension  $dp$  et une base de  $V$  est la *réunion disjointe* de  $p$  copies de  $\mathcal{B}$  : l'élément  $e_i$  de la  $q$ -ème copie de  $\mathcal{B}$  correspond à l'élément  $v_{i,q} = (0, \dots, 0, e^i, 0, \dots, 0)$  de  $V$ , où  $e_i$  est à la  $q$ -ème place, et tout  $v \in V$  s'écrit de façon unique  $v = \sum_{i=1}^d \sum_{q=1}^p x_{i,q} v_{i,q}$ , avec  $t_{i,q} \in k$ .

Mais, de l'autre côté, l'ensemble produit  $\mathcal{B} \times \dots \times \mathcal{B}$  est formé des  $p$ -uplets  $(e_{i_1}, \dots, e_{i_p})$  et est de cardinal  $d^p$ .

*Remarque :*  $\mathcal{B} \times \dots \times \mathcal{B}$  est une base de l'espace « produit tensoriel »  $E \otimes \dots \otimes E$ , cf. l'appendice 4.6.

## 4.2. Anneaux de polynômes

Un autre exemple important d'application bilinéaire est la multiplication dans une  $k$ -algèbre  $A$ , par exemple  $A = k[X]$  ou  $A = M_n(k)$ . Rappelons d'abord les définitions suivantes :

**Définition 4.2.0 (Anneaux).** — Un anneau  $(A, +, \times)$  est un groupe abélien  $(A, +)$  muni d'une loi de multiplication  $\times$  vérifiant les conditions suivantes (pour alléger l'écriture, on note  $ab$  au lieu de  $a \times b$ ) :

(i)  $A$  possède un élément unité  $1 \neq 0$  tel que  $1 \times a = a = a \times 1$  pour tout  $a \in A$ .

(ii)  $\times$  est associative :  $a(bc) = (ab)c$ ,

(iii)  $\times$  est distributive à gauche et à droite par rapport à l'addition. Ceci équivaut à dire que  $\times$  est « *bi-additive* », c.-à-d., additive en chacune des variables :

$$(\star) \quad \forall a, a', b, b' \in A, \quad (a + a')b = ab + a'b \quad \text{et} \quad a(b + b') = ab + ab'.$$

On ne demande pas que  $\times$  soit commutative; si elle l'est on dit que  $A$  est un anneau commutatif. Par exemple, l'anneau de polynômes  $k[X]$  est commutatif, mais l'anneau de matrices  $M_n(k)$  ne l'est pas si  $n \geq 2$ .

**Définition 4.2.1 ( $k$ -algèbres).** — Soit  $k$  un corps. Une  $k$ -algèbre est un  $k$ -espace vectoriel  $A$  muni d'une multiplication  $\times$  qui fait de  $A$  un anneau (conditions (i-iii) plus haut) et qui de plus est  *$k$ -bilinéaire*. Comme (iii) plus haut exprime déjà la bi-additivité, ceci revient à ajouter la condition suivante, où  $\cdot$  désigne la loi externe :

$$(\dagger) \quad \forall a, b \in A, \quad \forall \lambda \in k, \quad (\lambda \cdot a)b = \lambda \cdot (ab) = a(\lambda \cdot b).$$

En particulier, prenant  $a = 1$  et  $b = \mu \cdot 1$ , avec  $\mu \in k$ , on obtient  $(\lambda \cdot 1)(\mu \cdot 1) = \lambda \cdot (\mu \cdot 1) = (\lambda\mu) \cdot 1$ , donc l'application linéaire  $k \rightarrow A$ ,  $\lambda \mapsto \lambda \cdot 1$  est un morphisme d'anneaux, qui de plus est *injectif* puisque  $k$  est un corps (si  $\lambda \neq 0$ , on a  $\lambda^{-1} \cdot (\lambda \cdot 1) = 1 \neq 0$ , donc  $\lambda \cdot 1 \neq 0$ ).

**Exemples 4.2.2.** — Des exemples importants (et à connaître!) de  $k$ -algèbres sont les algèbres  $M_n(k)$  et  $\text{End}_k(V)$  déjà rencontrées (cf. 0.4.8), l'algèbre de polynômes  $k[X]$  et ses généralisations, les algèbres de polynômes à plusieurs variables (par exemple,  $k[X, Y]$ ,  $k[X, Y, Z]$ , etc.) que l'on va introduire plus bas. Pour définir et étudier ces algèbres, la bilinéarité sera un outil très utile.

**Définition 4.2.3 (Morphismes de  $k$ -algèbres).** — Si  $A, B$  sont deux  $k$ -algèbres, une application  $f : A \rightarrow B$  est un *morphisme de  $k$ -algèbres* si elle est linéaire et si c'est un morphisme d'anneaux, c.-à-d., si l'on a :

$$\boxed{f(1_A) = 1_B} \quad \text{et} \quad \boxed{f(a + \lambda b) = f(a) + \lambda f(b)} \quad \boxed{f(ab) = f(a)f(b)} \quad \forall a, b \in A, \quad \forall \lambda \in k.$$

**Théorème 4.2.4 (Propriété universelle de  $k[X]$ ).** — Soient  $B$  une  $k$ -algèbre (pas nécessairement commutative) et  $u \in B$ . Il existe un unique morphisme de  $k$ -algèbre  $\text{ev}_u : k[X] \rightarrow B$  tel que  $\text{ev}_u(X) = u$ ; explicitement, si  $P = a_0 + a_1X + \dots + a_dX^d$ , on a

$$(*) \quad \text{ev}_u(P) = a_0 1_B + a_1 u + \dots + a_d u^d = P(u).$$

Le morphisme  $\text{ev}_u$  s'appelle le morphisme « d'évaluation en  $u$  ».



*Démonstration.* — L'unicité est claire car  $ev_u$ , s'il existe, doit vérifier la formule (\*). Réciproquement, comme  $(X^n)_{n \in \mathbb{N}}$  est une base du  $k$ -espace vectoriel  $k[X]$ , on peut définir une application  $k$ -linéaire  $ev_u : k[X] \rightarrow B$  par la formule (\*). Alors  $ev_u(1) = 1_B$  et il reste à vérifier que, pour tout  $P, Q \in k[X]$ ,

$$(\dagger) \quad ev_u(PQ) = ev_u(P) ev_u(Q).$$

Comme les deux membres de l'égalité sont bilinéaires en  $P, Q$ , il suffit de vérifier cette égalité lorsque  $P, Q$  parcourent la base  $(X^n)_{n \in \mathbb{N}}$  de  $k[X]$ , c.-à-d., lorsque  $P = X^r$  et  $Q = X^s$  pour des entiers  $r, s \in \mathbb{N}$ . Mais alors  $PQ = X^{r+s}$  et le résultat est évident :

$$ev_u(X^{r+s}) = u^{r+s} = u^r u^s = ev_u(X^r) ev_u(X^s).$$

De façon plus détaillée, la bilinéarité équivaut à dire ceci :  $P$ , resp.  $Q$ , est une somme finie de termes  $a_i X^i$ , resp.  $b_j X^j$ ; comme le produit dans  $k[X]$  est bilinéaire et que  $ev_u$  est linéaire,  $ev_u(PQ)$  est la somme des termes  $a_i b_j ev_u(X^{i+j})$ , d'autre part, comme  $ev_u$  est linéaire et que le produit dans  $B$  est bilinéaire,  $ev_u(P) ev_u(Q)$  est la somme des termes  $a_i b_j ev_u(X^i) ev_u(X^j)$ , donc pour vérifier l'égalité ( $\dagger$ ), il suffit de voir que  $ev_u(X^{i+j}) = ev_u(X^i) ev_u(X^j)$ , et c'est ce qu'on a fait plus haut !

Ceci est plus commode et plus rapide que d'écrire explicitement les formules :  $P = a_0 + a_1 X + \dots + a_d X^d$  et  $Q = b_0 + b_1 X + \dots + b_f X^f$  d'où

$$ev_u(PQ) = ev_u \left( \sum_{n=0}^{d+f} \left( \sum_{i+j=n} a_i b_j \right) X^n \right) = \sum_{n=0}^{d+f} \left( \sum_{i+j=n} a_i b_j \right) ev_u(X^n)$$

et, d'autre part,  $ev_u(P) = a_0 1_B + a_1 u + \dots + a_d u^d$  et  $Q = b_0 1_B + b_1 u + \dots + b_f u^f$  d'où

$$ev_u(P) ev_u(Q) = \sum_{n=0}^{d+f} \left( \sum_{i+j=n} a_i b_j \right) u^n.$$

□

#### Définition et proposition 4.2.5 (L'opérateur de dérivation dans $k[X]$ )

Soit  $k$  un corps arbitraire, on définit un endomorphisme  $D$  de  $k[X]$  par :

$$(1) \quad D(X^n) = nX^{n-1}, \quad \forall n \in \mathbb{N}.$$

(Lorsque  $k = \mathbb{R}$ , ceci coïncide avec la définition usuelle du polynôme dérivé.) Alors, pour tout  $P, Q \in k[X]$ , on a

$$(2) \quad D(PQ) = D(P)Q + PD(Q).$$

*Démonstration.* Comme  $(X^n)_{n \in \mathbb{N}}$  est une base de  $k[X]$ , l'endomorphisme  $D$  est uniquement déterminé par (1). Montrons (2). Comme les deux termes de (2) sont bilinéaires en  $P, Q$ , il suffit de vérifier cette égalité lorsque  $P, Q$  parcourent la base  $(X^n)_{n \in \mathbb{N}}$  de  $k[X]$ , c.-à-d., lorsque  $P = X^r$  et  $Q = X^s$  pour des entiers  $r, s \in \mathbb{N}$ . Mais alors  $PQ = X^{r+s}$  et le résultat est évident :

$$D(X^{r+s}) = (r+s)X^{r+s-1} = rX^{r-1}X^s + sX^r X^{s-1} = D(X^r)X^s + X^r D(X^s).$$

**Polynômes à  $n$  variables.** — On va généraliser la construction de l'anneau de polynômes  $k[X]$  au cas de  $n$  indéterminées  $X_1, \dots, X_n$ . Commençons par le cas  $n = 2$ , i.e. le cas de deux indéterminées  $X$  et  $Y$ .

#### Définition 4.2.6 (La $k$ -algèbre des polynômes en deux indéterminées)

Soit  $k[X, Y]$  le  $k$ -espace vectoriel de base les monômes  $X^r Y^s$ , pour  $(r, s) \in \mathbb{N}^2$ . On définit le degré d'un tel monôme comme étant  $r + s$ .

Tout élément non nul  $P \in k[X, Y]$  est une somme finie de termes  $a_{r,s} X^r Y^s$ , avec  $a_{r,s} \in k$ , et le plus grand des degrés  $r + s$  tel que  $a_{r,s} \neq 0$  s'appelle le degré de  $P$  et se note  $\deg(P)$ ; ainsi  $P$  peut s'écrire comme une somme finie

$$P = \sum_{\substack{(r,s) \in \mathbb{N}^2 \\ r+s \leq n}} a_{r,s} X^r Y^s,$$

où  $n = \deg(P)$ . On munit  $k[X, Y]$  de la multiplication bilinéaire définie par

$$(X^r Y^s) \times (X^t Y^u) = X^{r+t} Y^{s+u}$$

c.-à-d., de façon explicite :

$$\left( \sum_{\substack{(r,s) \in \mathbb{N}^2 \\ r+s \leq m}} a_{r,s} X^r Y^s \right) \left( \sum_{\substack{(t,u) \in \mathbb{N}^2 \\ t+u \leq n}} b_{t,u} X^t Y^u \right) = \sum_{\substack{(p,q) \in \mathbb{N}^2 \\ p+q \leq m+n}} \left( \sum_{\substack{(r,s),(t,u) \in \mathbb{N}^2 \\ r+t=p, s+u=q}} a_{r,s} b_{t,u} \right) X^p Y^q.$$

Alors, la multiplication est bilinéaire, donc en particulier distributive à gauche et à droite sur l'addition ; d'après la définition, le monôme  $1 = X^0 Y^0$  est élément unité, et il reste à vérifier que la multiplication est associative et commutative : pour ceci, voir ci-dessous le cas de  $n$  variables  $X_1, \dots, X_n$ .

Ce qui précède se généralise de façon évidente au cas de  $n$  variables. Toutefois, pour alléger l'écriture, il est utile d'observer que  $\mathbb{N}^n$  est muni de l'addition définie composante par composante par :

$$(r_1, \dots, r_n) + (s_1, \dots, s_n) = (r_1 + s_1, \dots, r_n + s_n).$$

Pour tout  $\mathbf{r} = (r_1, \dots, r_n)$  dans  $\mathbb{N}^n$ , on pose  $|\mathbf{r}| = r_1 + \dots + r_n$  et l'on note  $X^{\mathbf{r}}$  le monôme  $X_1^{r_1} \dots X_n^{r_n}$  ; il est de degré  $|\mathbf{r}|$ . On peut alors définir l'anneau de polynômes  $k[X_1, \dots, X_n]$  comme suit.

**Définition 4.2.7 (La  $k$ -algèbre des polynômes en  $n$  indéterminées)**

Soit  $k[X_1, \dots, X_n]$  le  $k$ -espace vectoriel de base les monômes

$$X^{\mathbf{r}} := X_1^{r_1} \dots X_n^{r_n},$$

pour  $\mathbf{r} \in \mathbb{N}^n$ , un tel monôme étant de degré  $|\mathbf{r}|$ . Tout élément non nul  $P \in k[X_1, \dots, X_n]$  est une somme finie de termes  $a_{\mathbf{r}} X^{\mathbf{r}}$ , avec  $a_{\mathbf{r}} \in k$ , et le plus grand des degrés  $|\mathbf{r}|$  tels que  $a_{\mathbf{r}} \neq 0$  s'appelle le degré de  $P$  et se note  $\deg(P)$  ; ainsi  $P$  peut s'écrire comme une somme finie

$$P = \sum_{\substack{\mathbf{r} \in \mathbb{N}^n \\ |\mathbf{r}| \leq n}} a_{\mathbf{r}} X^{\mathbf{r}},$$

où  $n = \deg(P)$ . On munit  $k[X_1, \dots, X_n]$  de la multiplication bilinéaire définie par

$$(X_1^{r_1} \dots X_n^{r_n}) \times (X_1^{s_1} \dots X_n^{s_n}) = X_1^{r_1+s_1} \dots X_n^{r_n+s_n}$$

c.-à-d., de façon explicite :

$$\left( \sum_{\substack{\mathbf{r} \in \mathbb{N}^n \\ |\mathbf{r}| \leq p}} a_{\mathbf{r}} X^{\mathbf{r}} \right) \left( \sum_{\substack{\mathbf{s} \in \mathbb{N}^n \\ |\mathbf{s}| \leq q}} b_{\mathbf{s}} X^{\mathbf{s}} \right) = \sum_{\substack{\mathbf{t} \in \mathbb{N}^n \\ |\mathbf{t}| \leq p+q}} \left( \sum_{\substack{\mathbf{r}, \mathbf{s} \in \mathbb{N}^n \\ \mathbf{r}+\mathbf{s}=\mathbf{t}}} a_{\mathbf{r}} b_{\mathbf{s}} \right) X^{\mathbf{t}}.$$

Alors, la multiplication est bilinéaire, donc en particulier distributive à gauche et à droite sur l'addition ; d'après la définition, le monôme  $1 = X_1^0 \dots X_n^0$  est élément unité. Vérifions que la multiplication est associative, i.e. que pour tous  $P, Q, R \in k[X_1, \dots, X_n]$  on a

$$(PQ)R = P(QR).$$



Les deux membres sont des fonctions trilineaires de  $P, Q, R$  donc, par trilinearité, il suffit de vérifier cette égalité lorsque  $P, Q, R$  sont des éléments de la base  $(X^{\mathbf{r}})_{\mathbf{r} \in \mathbb{N}^n}$ , c.-à-d., lorsque  $P = X^{\mathbf{r}}, Q = X^{\mathbf{s}}, R = X^{\mathbf{t}}$  avec  $\mathbf{r}, \mathbf{s}, \mathbf{t} \in \mathbb{N}^n$ . Mais alors le résultat est évident :

$$(X^{\mathbf{r}} X^{\mathbf{s}}) X^{\mathbf{t}} = X^{\mathbf{r}+\mathbf{s}} X^{\mathbf{t}} = X^{\mathbf{r}+\mathbf{s}+\mathbf{t}} = X^{\mathbf{r}} X^{\mathbf{s}+\mathbf{t}} = X^{\mathbf{r}} (X^{\mathbf{s}} X^{\mathbf{t}}).$$

De même, vérifions que la multiplication est commutative, i.e. que  $PQ = QP$  pour tous  $P, Q$  ; par bilinéarité, il suffit de le vérifier lorsque  $P = X^{\mathbf{r}}$  et  $Q = X^{\mathbf{s}}$ , auquel cas c'est évident :

$$X^{\mathbf{r}} X^{\mathbf{s}} = X^{\mathbf{r}+\mathbf{s}} = X^{\mathbf{s}} X^{\mathbf{r}}.$$

On a donc bien muni  $k[X_1, \dots, X_n]$  d'une structure de  $k$ -algèbre commutative.

Si  $B$  est une  $k$ -algèbre commutative, si  $u_1, \dots, u_n \in B$ , et si  $\mathbf{r} = (r_1, \dots, r_n) \in \mathbb{N}^n$ , on notera  $\mathbf{u}^{\mathbf{r}}$  l'élément  $u_1^{r_1} \dots u_n^{r_n}$  de  $B$ .

**Théorème 4.2.8 (Propriété universelle de  $k[X_1, \dots, X_n]$ ).** — Soient  $B$  une  $k$ -algèbre commutative. Pour tout  $n$ -uplet  $(u_1, \dots, u_n)$  d'éléments de  $B$ , il existe un unique morphisme de  $k$ -algèbres  $\text{ev}_{u_1, \dots, u_n} : k[X_1, \dots, X_n] \rightarrow B$  tel que  $\text{ev}_{u_1, \dots, u_n}(X_i) = u_i$ , pour  $i = 1, \dots, n$  ; si  $P = \sum_{\substack{\mathbf{r} \in \mathbb{N}^n \\ |\mathbf{r}| \leq p}} a_{\mathbf{r}} X^{\mathbf{r}}$ , on a

$$(*) \quad \text{ev}_{u_1, \dots, u_n}(P) = \sum_{\substack{\mathbf{r} \in \mathbb{N}^n \\ |\mathbf{r}| \leq p}} a_{\mathbf{r}} \mathbf{u}^{\mathbf{r}} = P(u_1, \dots, u_n).$$



Le morphisme  $\text{ev}_{u_1, \dots, u_n}$  s'appelle le morphisme « d'évaluation en  $(u_1, \dots, u_n)$  ».

*Démonstration.* — La démonstration est analogue à celle du théorème 4.2.4 : l'unicité est claire car  $\text{ev}_{u_1, \dots, u_n}$ , s'il existe, doit vérifier la formule (\*). Réciproquement, comme  $(X^r)_{r \in \mathbb{N}^n}$  est une base de  $k[X_1, \dots, X_n]$ , on peut définir une application  $k$ -linéaire  $\text{ev}_{u_1, \dots, u_n} : k[X_1, \dots, X_n] \rightarrow B$  par la formule (\*). Alors  $\text{ev}_{u_1, \dots, u_n}(1) = 1_B$  et il reste à vérifier que

$$\text{ev}_{u_1, \dots, u_n}(PQ) = \text{ev}_{u_1, \dots, u_n}(P) \text{ev}_{u_1, \dots, u_n}(Q),$$

pour tout  $P, Q \in k[X_1, \dots, X_n]$ . Par bilinéarité, il suffit de le vérifier lorsque  $P = X_1^{r_1} \cdots X_n^{r_n}$  et  $Q = X_1^{s_1} \cdots X_n^{s_n}$  sont des monômes. Alors  $PQ = X_1^{r_1+s_1} \cdots X_n^{r_n+s_n}$  et l'on a, d'une part,

$$\text{ev}_{u_1, \dots, u_n}(X_1^{r_1+s_1} \cdots X_n^{r_n+s_n}) = u_1^{r_1+s_1} \cdots u_n^{r_n+s_n}$$

et, d'autre part, puisque  $B$  est commutative :

$$\text{ev}_{u_1, \dots, u_n}(X_1^{r_1} \cdots X_n^{r_n}) \text{ev}_{u_1, \dots, u_n}(X_1^{s_1} \cdots X_n^{s_n}) = (u_1^{r_1} \cdots u_n^{r_n}) \times (u_1^{s_1} \cdots u_n^{s_n}) = u_1^{r_1+s_1} \cdots u_n^{r_n+s_n}.$$

Ceci prouve le théorème. □

### 4.3. Groupes symétriques : quelques propriétés



**Définition 4.3.1 (Groupe symétrique  $S_n$ ).** — Soit  $S_n$  le groupe des permutations de  $\{1, \dots, n\}$ , i.e. des bijections de  $\{1, \dots, n\}$  sur lui-même <sup>(1)</sup>. C'est un groupe pour la composition des applications, et il est de cardinal  $n!$ , car une permutation  $\sigma$  est déterminée par la donnée de  $\sigma(1)$ , pour lequel il y a  $n$  choix, puis de  $\sigma(2)$ , pour lequel il reste  $n-1$  choix, etc.

**Notation 4.3.2.** — On représente en général un élément  $\sigma$  de  $S_n$  par son écriture « à deux lignes » : sur la première ligne, on écrit  $1, 2, 3, \dots, n$ , dans cet ordre, et sur la seconde on écrit les nombres  $\sigma(1), \sigma(2), \sigma(3), \dots, \sigma(n)$ . Ainsi, par exemple,

$$\sigma = \begin{pmatrix} 123456 \\ 356124 \end{pmatrix} \quad \text{et} \quad \tau = \begin{pmatrix} 123456 \\ 615234 \end{pmatrix}$$

désignent les éléments de  $S_6$  tels que  $\sigma(1) = 3, \sigma(2) = 5, \sigma(3) = 6, \sigma(4) = 1, \sigma(5) = 2, \sigma(6) = 4$  et, respectivement,  $\tau(1) = 6, \tau(2) = 1, \tau(3) = 5, \tau(4) = 2, \tau(5) = 3, \tau(6) = 4$ . Par définition, le produit  $\sigma\tau$  est la bijection composée  $\sigma \circ \tau$ , i.e. on fait d'abord  $\tau$  puis  $\sigma$ , c.-à-d.,  $\sigma\tau(1) = \sigma(\tau(1)) = \sigma(6) = 4$ , etc. On trouve alors :

$$\sigma\tau = \begin{pmatrix} 123456 \\ 432561 \end{pmatrix} \quad \text{et} \quad \tau\sigma = \begin{pmatrix} 123456 \\ 534612 \end{pmatrix}.$$

On notera que le groupe  $S_n$  n'est pas commutatif en général (il l'est seulement pour  $n = 1$  ou  $2$ ).

Pour certaines permutations, on utilise une écriture plus condensée, introduite ci-dessous.



**Définition 4.3.3 (Transpositions et cycles).** — Pour  $i \neq j$ , on note  $(ij)$  la permutation qui échange  $i$  et  $j$  et laisse les autres éléments inchangés ; on dit que  $(ij)$  est une **transposition**. Remarquons que si  $\tau$  est une transposition, alors  $\tau^2 = \text{id}$  et donc  $\tau = \tau^{-1}$ .

Plus généralement, pour  $r \geq 2$ , on dit que  $c \in S_n$  est un  **$r$ -cycle** s'il existe  $i_1, \dots, i_r$ , deux à deux distincts, tels que  $c(j) = j$  pour  $j \notin \{i_1, \dots, i_r\}$  et

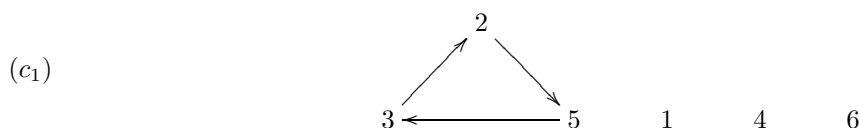
$$c(i_1) = i_2, \quad c(i_2) = i_3, \quad \dots, \quad c(i_{r-1}) = i_r, \quad c(i_r) = i_1.$$

Dans ce cas, on note  $c = (i_1 i_2 \cdots i_r)$ , et l'on dit que l'ensemble  $\{i_1, \dots, i_r\}$  est le **support** du cycle  $c$ .

Par exemple, dans  $S_6$ , les cycles  $c_1 = (253)$  et  $c_2 = (1635)$  désignent, respectivement, les permutations suivantes :

$$\begin{pmatrix} 123456 \\ 152436 \end{pmatrix} \quad \text{et} \quad \begin{pmatrix} 123456 \\ 625413 \end{pmatrix}.$$

Le terme de « cycle » s'explique si l'on représente graphiquement l'effet du cycle  $c$  sur les éléments  $1, \dots, n$  : dans l'exemple précédent, pour  $c_1$  et  $c_2$  on a respectivement :



<sup>(1)</sup>Il est parfois noté  $\Sigma_n$  (= « Sigma »  $n$ )

$$(c_2) \quad \begin{array}{ccc} 1 & \longrightarrow & 6 \\ \uparrow & & \downarrow \\ 5 & \longleftarrow & 3 \end{array} \quad 2 \quad 4$$

i.e. les éléments hors du support de  $c$  sont inchangés, et  $c$  permute « cycliquement » les éléments de son support.

On voit ainsi que l'inverse de  $c$  est le  $r$ -cycle de même support, que l'on parcourt en sens inverse : par exemple,  $c_1^{-1} = (235)$  et  $c_2^{-1} = (1536)$ . On voit aussi que  $c$  est un élément d'ordre  $r$ , c.-à-d.,  $c^r = \text{id}$  mais  $c^s \neq \text{id}$  pour  $s = 1, \dots, r-1$ . En particulier, on a donc  $c^{-1} = c^{r-1}$ .

**Remarque 4.3.4.** — Soit  $X$  un ensemble, réunion disjointe de sous-ensembles  $A_1, \dots, A_d$ , et pour chaque  $i = 1, \dots, d$ , soit  $c_i$  une permutation de  $A_i$ , i.e. une bijection de  $A_i$  dans lui-même ; on étend  $c_i$  en une bijection  $\sigma_i$  de  $X$  en prolongeant  $c_i$  par l'identité hors de  $A_i$ , i.e. on pose  $\sigma_i(x) = c_i(x)$  si  $x \in A_i$  et  $\sigma_i(x) = x$  si  $x \notin A_i$ . Alors on voit facilement que la composée  $\sigma = \sigma_1 \circ \dots \circ \sigma_d$  ne dépend pas de l'ordre des facteurs, c.-à-d.,  $\sigma = \sigma_d \circ \dots \circ \sigma_1$ , etc. En effet, quel que soit l'ordre des facteurs, le produit est la permutation  $\sigma$  de  $X$  qui coïncide avec  $c_i$  sur chaque  $A_i$ , i.e.  $\sigma(x) = c_i(x)$  si  $x \in A_i$  (ceci est bien défini, puisque les  $A_i$  sont disjoints et leur réunion est  $X$ ). On a donc obtenu le résultat suivant :

$$(*) \quad \boxed{\text{Dans } S_n, \text{ des cycles de supports disjoints commutent.}}$$

Par exemple, dans  $S_7$ , si  $\sigma = (257)$  et  $\tau = (1364)$  alors

$$\sigma\tau = \tau\sigma = \begin{array}{ccc} & 2 & \\ \nearrow & & \searrow \\ 7 & \longleftarrow & 5 \end{array} \quad \begin{array}{ccc} 1 & \longrightarrow & 3 \\ \uparrow & & \downarrow \\ 4 & \longleftarrow & 6 \end{array}$$

### **Théorème 4.3.5 (Décomposition en produit de cycles de supports disjoints)**

Soit  $\sigma \in S_n$ . Alors il existe des cycles  $c_1, \dots, c_d$  de supports disjoints, uniquement déterminés, tels que  $\sigma = c_1 \cdots c_n$ . (Et donc l'écriture précédente est unique, à l'ordre des facteurs près).

On admettra ce résultat (la démonstration n'est pas difficile, mais un peu pénible à écrire). Illustrons ceci par un exemple : soit

$$\sigma = \begin{pmatrix} 123456789 \\ 356184927 \end{pmatrix} \in S_9.$$

Choisissons un élément, par exemple 1, et regardons la suite de ses transformés par  $\sigma$  :  $1 \rightarrow 3 \rightarrow 6 \rightarrow 4 \rightarrow 1$ , ceci donne le cycle (1364) auquel appartient l'élément 1. Puis choisissons un élément parmi les éléments restants, par exemple 2, et regardons à nouveau la suite de ses transformés par  $\sigma$  :  $2 \rightarrow 5 \rightarrow 8 \rightarrow 2$ , ceci donne le cycle (258) auquel appartient l'élément 2. Puis on recommence ... Ici, la dernière étape consiste à choisir, disons, l'élément 7, et l'on obtient la transposition, i.e. le 2-cycle, (79), d'où la décomposition

$$\sigma = (1364)(258)(79) = \begin{array}{ccc} & 2 & \\ \nearrow & & \searrow \\ 8 & \longleftarrow & 5 \end{array} \quad \begin{array}{ccc} 1 & \longrightarrow & 3 \\ \uparrow & & \downarrow \\ 4 & \longleftarrow & 6 \end{array} \quad \begin{array}{c} 7 \\ \updownarrow \\ 9 \end{array}$$

**Théorème 4.3.6 ( $S_n$  est engendré par les transpositions).** — Les transpositions  $s_i = (i, i+1)$ , pour  $i = 1, \dots, n-1$ , engendrent  $S_n$ .

*Démonstration.* — On procède par récurrence sur  $n$ . C'est clair pour  $n = 2$  car  $S_2 = \{\text{id}, (12)\}$ . Supposons  $n \geq 3$  et le résultat établi pour  $n-1$ . On identifie  $S_{n-1}$  au sous-groupe de  $S_n$  formé des permutations  $\tau$  telles que  $\tau(n) = n$ . Posons  $s_i = (i, i+1)$ , pour  $i = 1, \dots, n-1$ , et notons  $H$  le sous-groupe de  $S_n$  engendré par les  $s_i$ .

Soit  $\sigma \in S_n$ . Si  $\sigma(n) = n$ , alors  $\sigma \in S_{n-1}$  et donc, par hypothèse de récurrence,  $\sigma$  appartient au sous-groupe engendré par les  $s_i$ , pour  $i \leq n-2$ , donc a fortiori  $\sigma \in H$ . On peut donc supposer que  $\sigma(n) = i < n$ . Mais alors,  $s_i\sigma(n) = i+1$ , et si  $i+1 < n$  alors  $s_{i+1}s_i\sigma(n) = i+2$ , etc. En continuant ainsi, on arrive à l'égalité  $s_{n-1} \cdots s_i\sigma(n) = n$ . Alors, d'après ce qui précède,  $\tau = s_{n-1} \cdots s_i\sigma$  appartient à  $H$ .

Multipliant l'égalité  $\tau = s_{n-1} \cdots s_i\sigma$  à gauche par  $s_{n-1}^{-1} = s_{n-1}$ , puis par  $s_{n-2}^{-1} = s_{n-2}$ , etc., on arrive à l'égalité  $\sigma = s_i \cdots s_{n-1}\tau$ . Comme  $\tau$  et les  $s_j$  appartiennent à  $H$ , il en est de même de  $\sigma$ . Ceci prouve que  $H = S_n$ , d'où le théorème.  $\square$

**Lemme 4.3.7.** — Soit  $k$  un corps. Tout élément  $\sigma \in S_n$  induit un automorphisme  $\phi_\sigma$  de la  $k$ -algèbre  $k[X_1, \dots, X_n]$ , défini par

$$(*) \quad \phi_\sigma(X_i) = X_{\sigma(i)}, \quad \forall i = 1, \dots, n.$$

$$(**) \quad \phi_\sigma(P)(X_1, \dots, X_n) = P(X_{\sigma(1)}, \dots, X_{\sigma(n)}), \quad \forall P \in k[X_1, \dots, X_n].$$

De plus, l'application  $\sigma \mapsto \phi_\sigma$  est un morphisme de groupes.

*Démonstration.* — D'après la propriété universelle de  $A = k[X_1, \dots, X_n]$ , il existe, pour tout  $\sigma \in S_n$ , un unique morphisme de  $k$ -algèbres  $\phi_\sigma : A \rightarrow A$  vérifiant (\*) et (\*\*). De plus, il résulte de (\*) que  $\phi_{\text{id}} = \text{id}_A$  et que  $\phi_\sigma \circ \phi_\tau = \phi_{\sigma\tau}$ .

Ceci entraîne, d'une part, que chaque  $\phi_\sigma$  est un automorphisme de  $A$ , d'inverse  $\phi_{\sigma^{-1}}$ , et, d'autre part, que l'application  $\sigma \mapsto \phi_\sigma$  est un morphisme de groupes, de  $S_n$  dans le groupe des automorphismes de  $k$ -algèbre de  $A$ .  $\square$

**Notation.** Pour tout  $P \in k[X_1, \dots, X_n]$ , on écrira simplement  $\sigma(P)$  au lieu de  $\phi_\sigma(P)$ .

**Théorème et définition 4.3.8 (Signature d'une permutation).** — Soit  $n$  un entier  $\geq 2$ .

(i) Il existe un unique morphisme de groupes  $\varepsilon : S_n \rightarrow \{\pm 1\}$  tel que  $\varepsilon(\tau) = -1$  pour toute transposition  $\tau = (ij)$ . On l'appelle la **signature**.

(ii) Par conséquent, si l'on écrit un élément  $\sigma \in S_n$  comme produit de transpositions de deux façons différentes :  $\sigma = s_1 \cdots s_p = t_1 \cdots t_q$  où les  $s_i$  et  $t_j$  sont des transpositions, alors  $p$  et  $q$  ont la même parité, puisque  $\varepsilon(\sigma) = (-1)^p = (-1)^q$ .

(iii) Enfin, pour tout  $\sigma \in S_n$ , on a  $\varepsilon(\sigma^{-1}) = \varepsilon(\sigma)$ .

*Démonstration.* — Remarquons d'abord que  $\varepsilon$ , s'il existe, est nécessairement unique, puisque les transpositions engendrent  $S_n$ , d'après le théorème 4.3.6. Il s'agit donc de montrer l'existence. (Ceci équivaut à montrer que, si  $s_1, \dots, s_p$  et  $t_1, \dots, t_q$  sont des transpositions et si  $s_1 \cdots s_p = t_1 \cdots t_q$ , alors  $p$  et  $q$  ont la même parité, ce qui n'est pas évident a priori.)

D'après le lemme précédent,  $S_n$  opère par automorphismes d'algèbre sur  $A = \mathbb{Q}[X_1, \dots, X_n]$ . Considérons le polynôme

$$V_n = \prod_{1 \leq i < j \leq n} (X_i - X_j).$$

Soit  $\sigma \in S_n$ . Alors  $\sigma(V_n) = \prod_{1 \leq i < j \leq n} (X_{\sigma(i)} - X_{\sigma(j)})$  et, pour tout  $i < j$ ,

$$(1) \quad \sigma(X_i - X_j) = \begin{cases} X_{\sigma(i)} - X_{\sigma(j)}, & \text{si } \sigma(i) < \sigma(j); \\ -(X_{\sigma(j)} - X_{\sigma(i)}), & \text{si } \sigma(i) > \sigma(j). \end{cases}$$

On dit qu'un couple  $(i, j)$  avec  $i < j$  est une *inversion* de  $\sigma$  si  $\sigma(i) > \sigma(j)$ ; on introduit le **nombre d'inversions** de  $\sigma$  :

$$(2) \quad \ell(\sigma) = |\{i < j \mid \sigma(i) > \sigma(j)\}|$$

et l'on définit la signature de  $\sigma$  par :

$$(3) \quad \varepsilon(\sigma) = (-1)^{\ell(\sigma)}.$$

On déduit alors de (1) que, pour tout  $\sigma \in S_n$ ,

$$(4) \quad \sigma(V_n) = (-1)^{\ell(\sigma)} V_n.$$

Ceci entraîne que  $\varepsilon : S_n \rightarrow \{\pm 1\}$  est un morphisme de groupes : en effet, pour  $\sigma, \tau \in S_n$  on a

$$\varepsilon(\sigma\tau)V_n = (\sigma\tau)(V_n) = \sigma(\varepsilon(\tau)V_n) = \varepsilon(\tau)\varepsilon(\sigma)V_n,$$

d'où  $\varepsilon(\sigma\tau) = \varepsilon(\sigma)\varepsilon(\tau)$ . En particulier, pour tout  $\sigma$  on a  $\varepsilon(\sigma^{-1}) = \varepsilon(\sigma)^{-1}$ , et comme  $\varepsilon(\sigma) \in \{\pm 1\}$ , on a  $\varepsilon(\sigma)^{-1} = \varepsilon(\sigma)$ . Ceci prouve déjà l'assertion (iii).

Enfin, pour  $i < j$  notons  $\tau_{ij}$  la transposition qui échange  $i$  et  $j$ . On vérifie facilement que les inversions de  $\tau_{ij}$  sont les couples  $(i, j)$  et  $(i, k)$ ,  $(k, j)$  pour  $i < k < j$ ; leur nombre est  $1 + 2(j - i - 1)$ , d'où  $\varepsilon(\tau_{ij}) = -1$ . Ceci prouve l'assertion (i) et, comme noté dans le théorème, l'assertion (ii) en découle aussitôt. Le théorème est démontré.  $\square$





**Définitions 4.3.9.** — 1) On dit qu'une permutation  $\sigma \in S_n$  est **paire**, resp. **impaire**, si  $\varepsilon(\sigma) = 1$ , resp.  $-1$ . Ceci équivaut à dire que  $\sigma$  s'écrit comme produit d'un nombre pair (resp. impair) de transpositions.

2)  $\text{Ker}(\varepsilon) = \{\sigma \in S_n \mid \varepsilon(\sigma) = 1\}$  est appelé **groupe alterné** d'ordre  $n$ , et noté  $A_n$ . Il est formé des permutations paires, et est de cardinal  $n!/2$ . (En effet, si  $\tau$  est une transposition (par exemple,  $\tau = \tau_{12}$ ), alors  $S_n$  est la réunion disjointe de  $A_n$  et de  $\tau A_n = \{\tau\sigma \mid \sigma \in A_n\}$ .)



**Proposition 4.3.10 (Signature d'un cycle).** — *Tout  $r$ -cycle est de signature  $(-1)^{r-1}$ .*

*Démonstration.* — Soit  $c = (i_1 \cdots i_r)$  un  $r$ -cycle. On voit facilement que le produit ci-dessous de  $(r-1)$  transpositions :

$$(i_1 i_2)(i_2 i_3) \cdots (i_{r-1} i_r)$$

envoie  $i_1$  sur  $i_2$ ,  $i_2$  sur  $i_3$ , ...,  $i_{r-1}$  sur  $i_r$ , et  $i_r$  sur  $i_1$ , donc égale  $c$ . Par conséquent,  $\varepsilon(c) = (-1)^{r-1}$ .  $\square$

#### 4.4. Applications multilinéaires alternées

Désormais, on fixe un corps  $k$ .

##### **Définition et proposition 4.4.1 (Applications $p$ -linéaires alternées et antisymétriques)**

Soient  $E, F$  deux  $k$ -espaces vectoriels et  $f \in \mathcal{L}_p(E, F)$  (i.e.  $f$  est une application  $p$ -linéaire  $E^p \rightarrow F$ ).

(1) On dit que  $f$  est **alternée** si elle vérifie :

$$(*) \quad f(x_1, \dots, x_p) = 0 \quad \text{si } x_1, \dots, x_p \text{ ne sont pas tous distincts.}$$

(2) On dit que  $f$  est **antisymétrique** si elle vérifie les conditions équivalentes suivantes :

$$(\dagger) \quad \forall i \neq j, \quad f(x_1, \dots, x_j, \dots, x_i, \dots, x_p) = -f(x_1, \dots, x_i, \dots, x_j, \dots, x_p)$$

$$(\dagger') \quad \forall \sigma \in S_p, \quad f(x_{\sigma(1)}, \dots, x_{\sigma(p)}) = \varepsilon(\sigma) f(x_1, \dots, x_p)$$

(3) Si  $f$  est alternée, elle est antisymétrique.

(4) L'ensemble des applications  $p$ -linéaires alternées  $f : E^p \rightarrow F$  un sous-espace vectoriel de  $\mathcal{L}_p(E, F)$ , on le notera  $\mathcal{A}_p(E, F)$ .

*Démonstration.* — L'équivalence de  $(\dagger)$  et  $(\dagger')$  résulte du théorème 4.3.8. Démontrons (3) (cf. le point (2<sup>-</sup>) de 1.4.1). Soient  $x_1, \dots, x_p \in E$ , fixons  $i < j$  dans  $\{1, \dots, p\}$ , et considérons l'application  $g : E \times E \rightarrow F$  définie par

$$g(X, Y) = f(x_1, \dots, x_{i-1}, X, x_{i+1}, \dots, x_{j-1}, Y, x_{j+1}, \dots, x_p),$$

i.e.  $X$  (resp.  $Y$ ) est à la  $i$ -ème (resp.  $j$ -ème) place). Comme  $f$  est  $p$ -linéaire et alternée, alors  $g$  est bilinéaire et alternée. Donc

$$0 = g(x_i + x_j, x_i + x_j) = g(x_i, x_i) + g(x_j, x_j) + g(x_i, x_j) + g(x_j, x_i),$$

et les deux premiers termes sont nuls, à nouveau puisque  $g$  est alternée. Donc

$$f(x_1, \dots, x_i, \dots, x_j, \dots, x_p) = g(x_i, x_j) = -g(x_j, x_i) = -f(x_1, \dots, x_j, \dots, x_i, \dots, x_p)$$

et ceci montre que  $f$  est antisymétrique.

D'autre part, on voit facilement que si  $\lambda \in k$  et si  $f, f' \in \mathcal{L}_p(E, F)$  sont alternées, alors  $\lambda f + f'$  l'est aussi, d'où (4).  $\square$

**Remarque 4.4.1.1.** — Soit  $f \in \mathcal{L}_p(E, k)$  antisymétrique et soit  $(x_1, \dots, x_p) \in E^p$ . S'il existe  $i \neq j$  tels que  $x_i = x_j$ , on a, d'après  $(\dagger)$  :

$$-f(x_1, \dots, x_i, \dots, x_j, \dots, x_p) = f(x_1, \dots, x_j, \dots, x_i, \dots, x_p) = f(x_1, \dots, x_i, \dots, x_j, \dots, x_p)$$

(la seconde égalité puisque  $x_i = x_j$ ), d'où  $2f(x_1, \dots, x_p) = 0$ . Donc, si  $k$  est de caractéristique  $\neq 2$  (c.-à-d., si  $2 \neq 0$  dans  $k$ ), alors :  $f$  antisymétrique  $\iff f$  alternée.



**Proposition 4.4.2 (Antisymétrisée d'une application  $p$ -linéaire).** — Soient  $E, F$  deux  $k$ -espaces vectoriels et  $f : E^p \rightarrow F$  une application  $p$ -linéaire. On définit son antisymétrisée  $A(f) : E^p \rightarrow F$  comme suit : pour tout  $(x_1, \dots, x_p) \in X^p$ ,

$$A(f)(x_1, \dots, x_p) = \sum_{\sigma \in S_p} \varepsilon(\sigma) f(x_{\sigma(1)}, \dots, x_{\sigma(p)}).$$

Alors  $A(f) : E^p \rightarrow A$  est une application  $p$ -linéaire alternée (donc antisymétrique).



*Démonstration.* — D'abord, comme  $f$  est  $p$ -linéaire alors, pour tout  $\sigma \in S_p$ , l'application  $\sigma(f) : E^p \rightarrow F$  définie par  $\sigma(f)(x_1, \dots, x_p) = f(x_{\sigma(1)}, \dots, x_{\sigma(p)})$  est  $p$ -linéaire, et donc  $A(f) = \sum_{\sigma \in S_p} \varepsilon(\sigma) \sigma(f)$  l'est aussi.

Montrons que  $A(f)$  est alternée. Soit  $(x_1, \dots, x_p) \in E^p$ , supposons qu'il existe  $i \neq j$  tels que  $x_i = x_j$ . Notons  $\tau$  la transposition qui échange  $i$  et  $j$ , alors le groupe symétrique  $S_p$  est la réunion disjointe du groupe alterné  $A_p$  et de  $\tau A_p = \{\tau\sigma \mid \sigma \in A_p\}$ ; en effet

$$\begin{cases} \sigma \in A_p \Leftrightarrow \varepsilon(\sigma) = 1 \\ \sigma \notin A_p \Leftrightarrow \varepsilon(\sigma) = -1 \Leftrightarrow \varepsilon(\tau\sigma) = 1 \Leftrightarrow \tau\sigma \in A_p \Leftrightarrow \sigma \in \tau A_p. \end{cases}$$

(la dernière équivalence de la seconde ligne résultant du fait que  $\tau^2 = 1$ ). Il en résulte :

$$(\dagger) \quad A(f)(x_1, \dots, x_n) = \sum_{\sigma \in A_p} \underbrace{\varepsilon(\sigma)}_{=1} \left[ f(x_{\sigma(1)}, \dots, x_{\sigma(p)}) + \underbrace{\varepsilon(\tau)}_{=-1} f(x_{\tau\sigma(1)}, \dots, x_{\tau\sigma(p)}) \right].$$

Fixons  $\sigma \in A_p$ . Pour tout  $s \in \{1, \dots, p\}$  tel que  $\sigma(s) \neq i, j$ , on a  $\tau(\sigma(s)) = \sigma(s)$ . Comme  $x_i = x_j$ , on a lorsque  $\sigma(s) = i$  les égalités :

$$x_{\tau(\sigma(s))} = x_{\tau(i)} = x_j = x_i = x_{\sigma(s)},$$

et de même  $x_{\tau(\sigma(s))} = x_{\sigma(s)}$  lorsque  $\sigma(s) = j$ . Donc  $\tau\sigma(s) = \sigma(s)$  pour tout  $s$ . Alors, comme  $\varepsilon(\tau) = -1$ , la somme entre crochets dans  $(\dagger)$  est nulle, pour tout  $\sigma \in A_p$ . Ceci prouve que  $A(f)(x_1, \dots, x_n) = 0$  si deux des  $x_i$  sont égaux. La proposition est démontrée.  $\square$

**Notation 4.4.3.** — Soient  $E$  un  $k$ -espace vectoriel et  $f_1, \dots, f_p \in E^*$  des formes linéaires sur  $E$ . On note  $f_1 \otimes \dots \otimes f_p$  l'application  $E^p \rightarrow k$  définie par

$$(f_1 \otimes \dots \otimes f_p)(x_1, \dots, x_p) = f_1(x_1) \cdots f_p(x_p).$$

Comme le produit  $k \times \dots \times k \rightarrow k$ ,  $(t_1, \dots, t_p) \mapsto t_1 \cdots t_p$  est  $p$ -linéaire, et que chaque  $f_i : E \rightarrow k$  est linéaire, on voit facilement que  $f_1 \otimes \dots \otimes f_p$  est  $p$ -linéaire.

#### Définition et proposition 4.4.4 (Produit extérieur de formes linéaires)

Soient  $E$  un  $k$ -espace vectoriel et  $f_1, \dots, f_p \in E^*$  des formes linéaires sur  $E$ . On appelle **produit extérieur** de  $f_1, \dots, f_p$  et on note  $f_1 \wedge \dots \wedge f_p$  l'antisymétrisée de  $f_1 \otimes \dots \otimes f_p$ , i.e. l'application  $E^p \rightarrow k$  définie par

$$(1) \quad (f_1 \wedge \dots \wedge f_p)(x_1, \dots, x_p) = \sum_{\sigma \in S_p} \varepsilon(\sigma) f_1(x_{\sigma(1)}) \cdots f_p(x_{\sigma(p)}).$$

C'est une application  $p$ -linéaire alternée  $E^p \rightarrow k$ . De plus, pour tout  $(x_1, \dots, x_p) \in E^p$ , on a aussi :

$$(2) \quad (f_1 \wedge \dots \wedge f_p)(x_1, \dots, x_p) = \sum_{\sigma \in S_p} \varepsilon(\sigma) f_{\sigma(1)}(x_1) \cdots f_{\sigma(p)}(x_p).$$

*Démonstration.* — Compte-tenu de 4.4.3, la première assertion découle de la proposition 4.4.2. Montrons la formule (2).

Fixons  $\sigma \in S_p$ . On peut réordonner le produit  $\prod_{i=1}^p f_i(x_{\sigma(i)})$  en fonction des indices  $\sigma(i)$ , qui parcourent aussi  $\{1, \dots, p\}$ , i.e. on a :

$$f_1(x_{\sigma(1)}) \cdots f_p(x_{\sigma(p)}) = \prod_{i=1}^p f_{\sigma^{-1}(i)}(x_i).$$

Comme  $\sigma \mapsto \sigma^{-1}$  est une bijection de  $S_p$  sur lui-même, et  $\varepsilon(\sigma^{-1}) = \varepsilon(\sigma)$ , on obtient la formule (2).  $\square$

#### Théorème 4.4.5 (Bases et dimension de $\mathcal{A}_p(E, k)$ , lien avec le déterminant)

Soient  $E$  un  $k$ -espace vectoriel de dimension  $n$ ,  $\mathcal{B} = (e_1, \dots, e_n)$  une base de  $E$ ,  $\mathcal{B}^* = (e_1^*, \dots, e_n^*)$  la base duale. Alors  $\mathcal{A}_p(E, k) = 0$  pour  $p > n$ , et pour  $p = 1, \dots, n$ , les produits extérieurs

$$e_{i_1}^* \wedge \dots \wedge e_{i_p}^* \quad \text{pour } i_1 < \dots < i_p \text{ dans } \{1, \dots, n\}$$

forment une base de  $\mathcal{A}_p(E, k)$ . Donc, pour  $p = 1, \dots, n$ ,

$$\dim_k \mathcal{A}_p(E, k) = \binom{n}{p} = \frac{n!}{p!(n-p)!}.$$

En particulier,  $\mathcal{A}_n(E, k)$  est de dimension 1 et, pour tout  $v_1, \dots, v_n \in E$ , on a :

$$(*) \quad (e_1^* \wedge \dots \wedge e_n^*)(v_1, \dots, v_n) = \det_{\mathcal{B}}(v_1, \dots, v_n)$$



et donc

$$(**) \quad \det_{\mathcal{B}}(v_1, \dots, v_n) = \sum_{\sigma \in S_n} \varepsilon(\sigma) e_1^*(v_{\sigma(1)}) \cdots e_n^*(v_{\sigma(n)}) = \sum_{\sigma \in S_n} \varepsilon(\sigma) e_{\sigma(1)}^*(v_1) \cdots e_{\sigma(n)}^*(v_n).$$

*Démonstration.* — D'après le théorème 4.1.9, une base de  $\mathcal{L}_p(E, k)$  est donnée par les éléments  $\phi^{j_1, \dots, j_p}$ , pour  $(j_1, \dots, j_p) \in \{1, \dots, n\}^p$ , définis par

$$\phi^{j_1, \dots, j_p}(e_{i_1}, \dots, e_{i_p}) = \begin{cases} 1 & \text{si } i_1 = j_1, \dots, i_p = j_p; \\ 0 & \text{sinon.} \end{cases}$$

On voit d'ailleurs que  $\phi^{j_1, \dots, j_p} = e_{j_1}^* \otimes \cdots \otimes e_{j_p}^*$ , car ces deux éléments de  $\mathcal{L}_p(E, k)$  coïncident sur tous les  $p$ -uplets  $(e_{i_1}, \dots, e_{i_p})$ , donc sur  $E^p$  tout entier (par  $p$ -linéarité).

Comme  $\mathcal{A}_p(E, k) \subset \mathcal{L}_p(E, k)$ , on peut écrire tout  $f \in \mathcal{A}_p(E, k)$ , sous la forme :

$$f = \sum_{(j_1, \dots, j_p) \in \{1, \dots, n\}^p} c_{(j_1, \dots, j_p)}^f e_{j_1}^* \otimes \cdots \otimes e_{j_p}^*$$

et l'on a  $c_{(j_1, \dots, j_p)}^f = f(e_{i_1}, \dots, e_{i_p})$ . Comme  $f$  est alternée (donc antisymétrique), on obtient :

$$(1) \quad c_{(j_1, \dots, j_p)}^f = 0 \quad \text{si } j_1, \dots, j_p \text{ ne sont pas tous distincts,}$$

et 4.4.1 (2) (†') donne

$$(2) \quad \text{si } j_1 < \cdots < j_p \text{ et } \sigma \in S_p \text{ alors } c_{(j_{\sigma(1)}, \dots, j_{\sigma(p)})}^f = \varepsilon(\sigma) c_{(j_1, \dots, j_p)}^f.$$

Alors, (1) entraîne déjà que  $\mathcal{A}_p(E, k) = 0$  si  $p > n$ , et (1) et (2) entraînent que, si  $p \leq n$ , tout  $f \in \mathcal{A}_p(E, k)$  s'écrit

$$f = \sum_{j_1 < \cdots < j_p} c_{(j_1, \dots, j_p)}^f \left( \sum_{\sigma \in S_p} \varepsilon(\sigma) e_{j_{\sigma(1)}}^* \otimes \cdots \otimes e_{j_{\sigma(p)}}^* \right) = \sum_{j_1 < \cdots < j_p} c_{(j_1, \dots, j_p)}^f e_{j_1}^* \wedge \cdots \wedge e_{j_p}^*.$$

D'autre part, les produits extérieurs  $e_{j_1}^* \wedge \cdots \wedge e_{j_p}^*$ , pour  $j_1 < \cdots < j_p$ , appartiennent à  $\mathcal{A}_p(E, k)$ , d'après la proposition précédente, et ils sont linéairement indépendants, car pour tout  $i_1 < \cdots < i_p$  dans  $\{1, \dots, n\}$ , on a

$$(e_{j_1}^* \wedge \cdots \wedge e_{j_p}^*)(e_{i_1}, \dots, e_{i_p}) = \begin{cases} 1 & \text{si } i_1 = j_1, \dots, i_p = j_p; \\ 0 & \text{sinon.} \end{cases}$$

Ceci montre que les  $e_{i_1}^* \wedge \cdots \wedge e_{i_p}^*$ , pour  $i_1 < \cdots < i_p$  dans  $\{1, \dots, n\}$ , forment une base de  $\mathcal{A}_p(E, k)$  et donc, pour  $p = 1, \dots, n$ , on a

$$\dim_k \mathcal{A}_p(E, k) = \binom{n}{p} = \frac{n!}{p!(n-p)}.$$

En particulier,  $\mathcal{A}_n(E, k)$  est de dimension 1, engendré par  $e_1^* \wedge \cdots \wedge e_n^*$ . Or, d'après l'exemple 4.1.5, l'application « déterminant relativement à la base  $\mathcal{B}$  »

$$\det_{\mathcal{B}} : E^n \rightarrow k, \quad (v_1, \dots, v_n) \mapsto \det_{\mathcal{B}}(v_1, \dots, v_n)$$

est également un élément de  $\mathcal{A}_n(E, k)$ , donc il existe  $\lambda \in k$  tel que  $\det_{\mathcal{B}} = \lambda e_1^* \wedge \cdots \wedge e_n^*$ ; appliquant les deux membres au  $n$ -uplet  $(e_1, \dots, e_n)$  on trouve  $\lambda = 1$ . Ceci prouve l'égalité (\*), et (\*\*) découle alors des formules (1) et (2) de 4.4.4. Le théorème est démontré.  $\square$

La formule 4.4.5 (\*) donne donc la formule explicite pour le déterminant d'une matrice  $A \in M_n(k)$ , évoquée en 2.1.5 dans le Chap. 2 :

**Proposition 4.4.6 (Formule explicite pour le déterminant).** — Soient  $V$  un  $k$ -espace vectoriel de dimension  $n$ ,  $\mathcal{B} = (e_1, \dots, e_n)$  une base de  $V$ , et  $v_1, \dots, v_n \in V$ . Écrivons, pour  $j = 1, \dots, n$ ,  $v_j = \sum_{i=1}^n a_{ij} e_i$ , d'où la matrice

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} = \text{Mat}_{\mathcal{B}}(v_1, \dots, v_n)$$

et remarquons que  $a_{ij} = e_i^*(v_j)$ , où  $(e_1^*, \dots, e_n^*)$  désigne la base duale de  $\mathcal{B}$ . Alors, on a :

$$\det_{\mathcal{B}}(v_1, \dots, v_n) = \det(A) = \sum_{\sigma \in S_n} \varepsilon(\sigma) a_{1\sigma(1)} \cdots a_{n\sigma(n)} = \sum_{\sigma \in S_n} \varepsilon(\sigma) a_{\sigma(1)1} \cdots a_{\sigma(n)n}$$

et ceci fournit une autre démonstration de l'égalité  $\det(A) = \det({}^t A)$ .



**Définition 4.4.7 (Matrices extraites et mineurs).** — 1) Soit  $A \in M_n(k)$ . Pour  $I, J$  deux parties non-vides de  $\{1, \dots, n\}$ , on note  $A_{I,J}$  la matrice à  $|I|$  lignes et  $|J|$  colonnes suivante :

$$A_{I,J} = (a_{ij})_{i \in I, j \in J}$$

et l'on dit que  $A_{I,J}$  est une « matrice extraite » de  $A$ .

2) Si  $I$  et  $J$  ont même cardinal, disons  $|I| = |J| = s$ , on note

$$\Delta_{I,J}(A) = \det(A_{I,J})$$

et l'on dit que  $\Delta_{I,J}(A)$  est un « mineur de taille  $s$  » de  $A$ . Si  $I = J$ , on note  $\Delta_I(A) = \Delta_{I,I}(A)$  et l'on dit que c'est un « mineur principal de taille  $s$  » de  $A$ .

**Lemme 4.4.8.** — Soient  $A \in M_n(k)$ ,  $I = \{i_1 < \dots < i_p\}$  et  $J = \{j_1 < \dots < j_p\}$  deux sous-ensembles non-vides de  $\{1, \dots, n\}$  de même cardinal  $p$ . Notons  $A_1, \dots, A_n$  les colonnes de  $A$ ,  $(e_1, \dots, e_n)$  la base canonique de  $k^n$ , et  $(e_1^*, \dots, e_n^*)$  la base duale. Alors

$$\Delta_{I,J}(A) = (e_{i_1}^* \wedge \dots \wedge e_{i_p}^*)(A_{j_1}, \dots, A_{j_p}).$$

*Démonstration.* — D'après 4.4.6 et 4.4.5 (\*), on a :

$$\Delta_{I,J}(A) = \det(A_{I,J}) = \sum_{\sigma \in S_p} \varepsilon(\sigma) a_{i_1, \sigma(j_1)} \cdots a_{i_p, \sigma(j_p)} = (e_{i_1}^* \wedge \dots \wedge e_{i_p}^*)(A_{j_1}, \dots, A_{j_p}).$$

□

On en déduit la proposition suivante.

**Proposition 4.4.9.** — Soit  $A \in M_n(k)$ . Alors

$$\text{rang}(A) = \sup\{s \mid \text{il existe un mineur } \Delta_{I,J}(A) \text{ de taille } s \text{ non nul}\}.$$

*Démonstration.* — Soit  $r = \text{rang}(A)$  et soit  $r'$  le terme de droite ci-dessus. Montrons d'abord que  $r' \leq r$ .

Soient  $s > r$  et  $I = \{i_1 < \dots < i_s\}$  et  $J = \{j_1 < \dots < j_s\}$  deux sous-ensembles de  $\{1, \dots, n\}$  de cardinal  $s$ . Comme  $s > r$ , les colonnes  $A_{j_1}, \dots, A_{j_s}$  de  $A$  sont liées, donc *a fortiori* les colonnes de la matrice extraite  $A_{I,J}$  le sont, donc  $\Delta_{I,J}(A) = 0$ . Ceci montre que tout mineur de taille  $> r$  est nul, d'où  $r' \leq r$ .

Réciproquement, comme  $\text{rang}(A) = r$ , on peut supposer, quitte à permuter les colonnes de  $A$ , que les colonnes  $A_1, \dots, A_r$  sont linéairement indépendantes. Complétons-les en une base  $(v_1, \dots, v_n)$  de  $k^n$  (où  $v_i = A_i$  pour  $i = 1, \dots, r$ ) et soit  $(v_1^*, \dots, v_n^*)$  la base duale. Alors

$$(\dagger) \quad (v_1^* \wedge \dots \wedge v_r^*)(A_1, \dots, A_r) = 1.$$

D'autre part, d'après le théorème 4.4.5, on peut écrire

$$v_1^* \wedge \dots \wedge v_r^* = \sum_{\substack{i_1, \dots, i_r \in \{1, \dots, n\} \\ i_1 < \dots < i_r}} c_{i_1, \dots, i_r} e_{i_1}^* \wedge \dots \wedge e_{i_r}^*$$

donc  $(\dagger)$  et le lemme précédent donnent l'égalité :

$$1 = \sum_{\substack{i_1, \dots, i_r \in \{1, \dots, n\} \\ i_1 < \dots < i_r}} c_{i_1, \dots, i_r} (e_{i_1}^* \wedge \dots \wedge e_{i_r}^*)(A_1, \dots, A_r) = \sum_{\substack{i_1, \dots, i_r \in \{1, \dots, n\} \\ i_1 < \dots < i_r}} c_{i_1, \dots, i_r} \Delta_{\{i_1, \dots, i_r\}, \{1, \dots, r\}}(A)$$

qui entraîne que l'un au moins des  $\Delta_{\{i_1, \dots, i_r\}, \{1, \dots, r\}}(A)$  est non nul. Ceci montre que  $r' \geq r$ , d'où  $r' = r$ . □

**Remarque 4.4.10.** — Compte-tenu de l'égalité  $\det(B) = \det({}^t B)$  pour tout  $B \in M_s(k)$ , le corollaire précédent donne une autre démonstration de l'égalité  $\text{rang}(A) = \text{rang}({}^t A)$ , cf. 0.5.11.

#### 4.5. Appendice $(\dagger)$ : Bases et dimension de $\mathcal{L}(E_1, \dots, E_p; F)$

Soit  $p$  un entier  $\geq 2$  et soient  $E_1, \dots, E_p$  et  $F$  des  $k$ -espaces vectoriels de dimension finie, soit  $\mathcal{C} = (f_1, \dots, f_r)$  une base de  $F$  et, pour  $q = 1, \dots, p$ , soient  $d_q = \dim E_q$  et

$$\mathcal{B}_q = (e_1^q, \dots, e_{d_q}^q)$$

une base de  $E_i$ . Notons  $P$  l'ensemble produit  $P = \{1, \dots, d_1\} \times \dots \times \{1, \dots, d_p\}$ , il est de cardinal  $D = d_1 \cdots d_p$ .

**Théorème 4.5.1 (Bases et dimension de  $\mathcal{L}(E_1, \dots, E_p; F)$ ).** — Avec les notations précédentes, on a :

(i) Tout  $\phi \in \mathcal{L}(E_1, \dots, E_p; F)$  est entièrement déterminé par ses valeurs  $u_{(i_1, \dots, i_p)} = \phi(e_{i_1}^1, \dots, e_{i_p}^p) \in F$  pour  $(i_1, \dots, i_p) \in P$ , et ces éléments peuvent être choisis arbitrairement.

(ii) Par conséquent, les éléments  $\phi_s^{a_1, \dots, a_p}$ , pour  $(a_1, \dots, a_p) \in P$  et  $s = 1, \dots, r$ , définis par :

$$\forall (i_1, \dots, i_p), \quad \phi_s^{a_1, \dots, a_p}(e_{i_1}, \dots, e_{i_p}) = \begin{cases} f_s \in F & \text{si } i_1 = a_1, \dots, i_p = a_p \\ 0 & \text{sinon} \end{cases}$$

forment une base de  $\mathcal{L}(E_1, \dots, E_p; F)$ ; celui-ci est donc de dimension  $rD = rd_1 \cdots d_p$ .

*Démonstration.* — Soit  $\phi \in \mathcal{L}(E_1, \dots, E_p; F)$ . Pour tout  $(v_1, \dots, v_p) \in E_1 \times \cdots \times E_p$ , chaque  $v_q$  s'écrit de façon unique

$$(0) \quad v_q = t_{q,1} e_1^q + \cdots + t_{q,d} e_{d_q}^q, \quad \text{avec } t_{q,i} \in k,$$

donc, en appliquant les égalités  $(*_1), \dots, (*_p)$  de 4.1.2, on obtient la formule :

$$(1) \quad \phi(v_1, \dots, v_p) = \phi\left(t_{1,1} e_1^1 + \cdots + t_{1,d} e_{d_1}^1, \dots, t_{p,1} e_1^p + \cdots + t_{p,d} e_{d_p}^p\right) \\ = \sum_{(i_1, \dots, i_p) \in P} t_{1,i_1} \cdots t_{p,i_p} \phi(e_{i_1}^1, \dots, e_{i_p}^p).$$

Ceci montre que  $\phi$  est entièrement déterminée par les  $D$  éléments  $\phi(e_{i_1}, \dots, e_{i_p}) \in F$ ; en d'autres termes, l'application « d'évaluation sur les  $p$ -uplets  $(e_{i_1}, \dots, e_{i_p})$  » :

$$\text{ev} : \mathcal{L}(E_1, \dots, E_p; k) \rightarrow F^D \simeq k^{rD}, \quad \phi \mapsto \left( \phi(e_{i_1}, \dots, e_{i_p}) \right)_{(i_1, \dots, i_p) \in P}$$

est *injective*. D'autre part, cette application est *linéaire* : si  $\mu \in k$  et  $\phi, \psi \in \mathcal{L}(E_1, \dots, E_p; k)$ , on a

$$(\mu\phi + \psi)(v_1, \dots, v_p) = \mu\phi(v_1, \dots, v_p) + \psi(v_1, \dots, v_p)$$

pour *tout*  $(v_1, \dots, v_p) \in E_1 \times \cdots \times E_p$ , donc *a fortiori* pour tout  $p$ -uplet  $(e_{i_1}, \dots, e_{i_p})$ .

Réciproquement, pour tout  $(a_1, \dots, a_p) \in P$ , notons  $\phi_s^{a_1, \dots, a_p}$  l'application  $E_1 \times \cdots \times E_p \rightarrow F$  définie comme suit : pour tout  $p$ -uplet  $(v_1, \dots, v_p)$  comme en (0) ci-dessus,

$$(2) \quad \phi_s^{a_1, \dots, a_p}(v_1, \dots, v_p) = \underbrace{(t_{1,a_1} \cdots t_{p,a_p})}_{\in k} \cdot f_s \in F.$$

Fixons un indice  $q \in \{1, \dots, p\}$ , et montrons que  $\phi_s^{a_1, \dots, a_p}$  est linéaire par rapport à la  $q$ -ème variable, i.e. que, pour tout  $\lambda \in k$  et  $w \in E_q$ , on a

$$(*_q) \quad \phi_s^{a_1, \dots, a_p}(v_1, \dots, \lambda v_q + w, \dots, v_p) = \lambda \phi_s^{a_1, \dots, a_p}(v_1, \dots, v_q, \dots, v_p) + \phi_s^{a_1, \dots, a_p}(v_1, \dots, w, \dots, v_p).$$

Écrivons  $w = s_1 e_1^q + \cdots + s_{d_q} e_{d_q}^q$ , avec  $s_i \in k$ . Alors  $\lambda v_q + w = t'_{q,1} e_1^q + \cdots + t'_{q,d} e_{d_q}^q$ , avec  $t'_{q,i} = \lambda t_{q,i} + s_i$ , et donc la formule (2) donne :

$$\phi_s^{a_1, \dots, a_p}(v_1, \dots, \lambda v_q + w, \dots, v_p) = (t_{1,a_1} \cdots t'_{q,a_q} \cdots t_{p,a_p}) \cdot f_s = (t_{1,a_1} \cdots (\lambda t_{q,a_q} + s_{a_q}) \cdots t_{p,a_p}) \cdot f_s \\ = \lambda t_{1,a_1} \cdots t_{q,a_q} \cdots t_{p,a_p} + t_{1,a_1} \cdots s_{a_q} \cdots t_{p,a_p} \\ = \lambda \phi_s^{a_1, \dots, a_p}(v_1, \dots, v_q, \dots, v_p) + \phi_s^{a_1, \dots, a_p}(v_1, \dots, w, \dots, v_p)$$

ce qui montre que  $(*_q)$  est vérifiée, d'où  $\phi_s^{a_1, \dots, a_p} \in \mathcal{L}(E_1, \dots, E_p; k)$ . Remarquons de plus que, pour tout  $(i_1, \dots, i_p) \in P$ , l'élément  $(v_1, \dots, v_p) = (e_{i_1}, \dots, e_{i_p})$  de  $E_1 \times \cdots \times E_p$  correspond dans l'écriture (0) à :

$$\forall q = 1, \dots, p, \quad t_{q,j_q} = \begin{cases} 1 & \text{si } j_q = i_q \\ 0 & \text{sinon} \end{cases}$$

donc (2) donne, en particulier :

$$(2') \quad \forall (i_1, \dots, i_p) \in P, \quad \phi_s^{a_1, \dots, a_p}(e_{i_1}, \dots, e_{i_p}) = \begin{cases} f_s & \text{si } i_1 = a_1, \dots, i_p = a_p \\ 0 & \text{sinon.} \end{cases}$$

Maintenant, pour tout élément  $U = \left( \sum_{s=1}^r \lambda_{(j_1, \dots, j_p)}^s f_s \right)_{(j_1, \dots, j_p) \in P} \in F^D$ , l'élément

$$\phi_U = \sum_{(j_1, \dots, j_p) \in P} \sum_{s=1}^r \lambda_{(j_1, \dots, j_p)}^s \phi_s^{(j_1, \dots, j_p)}$$

appartient à  $\mathcal{L}(E_1, \dots, E_p; k)$  et, d'après (2') il vérifie, pour tout  $(i_1, \dots, i_p) \in P$  :

$$(3) \quad \phi_U(e_{i_1}, \dots, e_{i_p}) = \sum_{(j_1, \dots, j_p) \in P} \sum_{s=1}^r \lambda_{(j_1, \dots, j_p)}^s \underbrace{\phi_s^{(j_1, \dots, j_p)}(e_{i_1}, \dots, e_{i_p})}_{=f_s \text{ si } (j_1, \dots, j_p)=(i_1, \dots, i_p) \\ =0 \text{ sinon}} = \sum_{s=1}^r \lambda_{(i_1, \dots, i_p)} f_s.$$

Ceci prouve que l'application linéaire injective  $\text{ev} : \mathcal{L}(E_1, \dots, E_p; F) \rightarrow F^D$  est aussi *surjective*, donc est un isomorphisme; en particulier,  $\dim \mathcal{L}(E_1, \dots, E_p; k) = rD = rd_1 \cdots d_p$ . Ceci prouve déjà l'assertion (i).

De plus, (3) implique que si l'on a une relation  $\sum_{(j_1, \dots, j_p) \in P} \sum_{s=1}^r \lambda_{(j_1, \dots, j_p)}^s \phi_s^{(j_1, \dots, j_p)} = 0$ , alors pour tout  $(i_1, \dots, i_p) \in P$  on a :

$$\sum_{s=1}^r \lambda_{(i_1, \dots, i_p)} f_s = 0$$

donc chaque  $\lambda_{(i_1, \dots, i_p)}^s$  est nul (puisque  $(f_1, \dots, f_r)$  est une base de  $F$ ). La famille

$$\mathcal{F} = \left( \phi_s^{(a_1, \dots, a_p)} \right)_{\substack{(a_1, \dots, a_p) \in P \\ s=1, \dots, r}}$$

est donc *libre*, et comme elle est de cardinal  $rD = rd_1 \cdots d_p$ , c'est une base de  $\mathcal{L}(E_1, \dots, E_p; k)$ . Ceci achève la preuve de l'assertion (ii). Le théorème est démontré.  $\square$

#### 4.6. Appendice (†) : Produit tensoriel

Soient  $E_1, \dots, E_p$  et  $F$  des  $k$ -espaces vectoriels, chaque  $E_q$  étant de dimension finie  $d_q$ . Pour  $q = 1, \dots, p$ , soit  $\mathcal{B}_q = (e_1^q, \dots, e_{d_q}^q)$  une base de  $E_q$ . D'après le théorème 4.5.1, on voit que l'espace des applications  $p$ -linéaires  $E_1 \times \cdots \times E_p \rightarrow F$  est la même chose que l'espace des applications *linéaires* de «  $E_1 \otimes \cdots \otimes E_p$  » vers  $F$ , où «  $E_1 \otimes \cdots \otimes E_p$  » désigne un espace vectoriel de base les  $p$ -uplets

$$(e_{j_1}^1, \dots, e_{j_p}^p) \in \mathcal{B}_1 \times \cdots \times \mathcal{B}_p.$$

On pourrait *définir* de la sorte l'espace «  $E_1 \otimes \cdots \otimes E_p$  », mais il faudrait alors démontrer qu'il ne dépend pas des bases choisies. En fait, on définit le *produit tensoriel*

$$E_1 \otimes \cdots \otimes E_p$$

comme suit. On considère d'abord le  $k$ -espace vectoriel  $V$  ayant pour base  $\mathcal{B}$  l'ensemble  $E_1 \times \cdots \times E_p$ , c.-à-d.,  $V$  admet pour base des vecteurs

$$b_{x_1, \dots, x_p}$$

indexés par  $E_1 \times \cdots \times E_p$ . Alors, se donner une application linéaire  $\phi : V \rightarrow F$  est la même chose que se donner pour tout vecteur de base  $b_{x_1, \dots, x_p}$  un élément  $f(x_1, \dots, x_p)$  de  $F$ ; on a donc une bijection naturelle

$$\begin{aligned} \mathcal{L}(V, F) &\xrightarrow{\sim} \text{Applic}(E_1 \times \cdots \times E_p, F) \\ \phi &\longmapsto \left( (x_1, \dots, x_p) \mapsto \phi(b_{x_1, \dots, x_p}) \right) \end{aligned}$$

où dans le terme de droite de la première ligne,  $E_1 \times \cdots \times E_p$  est regardé juste comme un ensemble.

Considérons maintenant le sous-espace  $W$  de  $V$  engendré par tous les éléments de l'une des formes suivantes, pour  $q = 1, \dots, p$  :

$$(*q) \quad \begin{cases} b_{x_1, \dots, x_q + y_q, \dots, x_p} - b_{x_1, \dots, x_q, \dots, x_p} - b_{x_1, \dots, y_q, \dots, x_p} \\ b_{x_1, \dots, \lambda \cdot x_q, \dots, x_p} - \lambda \cdot b_{x_1, \dots, x_q, \dots, x_p} \end{cases}$$

on pose

$$E_1 \otimes \cdots \otimes E_p = V/W$$

et pour  $x_1 \in E_1, \dots, x_p \in E_p$ , on note  $x_1 \otimes \cdots \otimes x_p$  l'image de  $b_{x_1, \dots, x_p}$  dans  $V/W = E_1 \otimes \cdots \otimes E_p$ .

Alors, d'après le théorème 3.5.3, une application linéaire  $E_1 \otimes \cdots \otimes E_p \rightarrow F$  est la même chose qu'une application linéaire  $\phi : V \rightarrow F$  qui s'annule sur les générateurs de  $W$ ; si l'on note  $f$  l'application d'ensembles  $E_1 \times \cdots \times E_p \rightarrow F$  correspondant à  $\phi$  (c.-à-d.,  $f(x_1, \dots, x_p) = \phi(b_{x_1, \dots, x_p})$ ), la condition que  $\phi$  s'annule sur les générateurs  $(*q)$  de  $W$  équivaut alors à dire que  $f$  est linéaire par rapport à la  $q$ -ème variable. On obtient ainsi que l'on a un isomorphisme d'espaces vectoriels :

$$(1) \quad \begin{aligned} \mathcal{L}(E_1 \otimes \cdots \otimes E_p, F) &\xrightarrow{\sim} \mathcal{L}(E_1, \dots, E_p; F) \\ \phi &\longmapsto \left( (x_1, \dots, x_p) \mapsto \phi(b_{x_1, \dots, x_p}) \right). \end{aligned}$$

D'autre part, tout  $x_1 \in E_1$  s'écrit de façon unique  $x_1 = \sum_{j_1=1}^{d_1} t_{1,j_1} e_{j_1}^1$  et donc, notant  $\pi$  la projection  $V \rightarrow E_1 \otimes \cdots \otimes E_p$ , on a, d'après  $(*_1)$  :

$$x_1 \otimes \cdots \otimes x_p = \pi(b_{x_1 \otimes \cdots \otimes x_p}) = \sum_{j_1=1}^{d_1} t_{1,j_1} \pi(b_{e_{j_1}, x_2, \dots, x_p}) = \sum_{j_1=1}^{d_1} t_{1,j_1} e_{j_1}^1 \otimes x_2 \otimes \cdots \otimes x_p$$

donc écrivant  $x_q = \sum_{j_q=1}^{d_q} t_{q,j_q} e_{j_q}^q$  pour  $q = 2, \dots, p$  et appliquant  $(*_2), \dots, (*_p)$ , on obtient que :

$$x_1 \otimes \cdots \otimes x_p = \sum_{j_1, \dots, j_p} t_{1,j_1} \cdots t_{p,j_p} e_{j_1}^1 \otimes \cdots \otimes e_{j_p}^p,$$

ce qui montre que  $E = E_1 \otimes \cdots \otimes E_p$  est engendré par les  $d_1 \cdots d_p$  vecteurs

$$e_{j_1}^1 \otimes \cdots \otimes e_{j_p}^p$$

où  $j_q \in \{1, \dots, d_q\}$  pour  $q = 1, \dots, p$ , donc  $E$  est de dimension finie  $n \leq d_1 \cdots d_p$ . Alors  $n$  est aussi la dimension de l'espace dual  $E^* = \mathcal{L}(E, k)$ ; or d'après (1) appliqué à  $F = k$ , on a un isomorphisme de  $k$ -espaces vectoriels

$$\mathcal{L}(E, k) \simeq \mathcal{L}(E_1, \dots, E_p; k)$$

et l'on a vu en 4.5.1 que  $\dim \mathcal{L}(E_1, \dots, E_p; k) = d_1 \cdots d_p$ . On obtient donc que

$$\dim(E_1 \otimes \cdots \otimes E_p) = d_1 \cdots d_p$$

et donc les  $d_1 \cdots d_p$  vecteurs  $e_{j_1}^1 \otimes \cdots \otimes e_{j_p}^p$  forment une base de  $E_1 \otimes \cdots \otimes E_p$ . On a donc obtenu le :

**Théorème 4.6.1.** — Soient  $E_1, \dots, E_p$  des  $k$ -espaces vectoriels de dimension finie et, pour  $q = 1, \dots, p$ , soit  $(e_{j_q}^q)$  une base de  $E_q$ . Alors, pour tout  $k$ -espace vectoriel  $F$ , on a un isomorphisme de  $k$ -espaces vectoriels

$$\boxed{\mathcal{L}(E_1 \otimes \cdots \otimes E_p, F) \simeq \mathcal{L}(E_1, \dots, E_p; F)}.$$

De plus,  $E_1 \otimes \cdots \otimes E_p$  est de dimension  $d_1 \cdots d_p$  et les vecteurs  $e_{j_1}^1 \otimes \cdots \otimes e_{j_p}^p$ , où chaque  $j_q$  parcourt  $\{1, \dots, d_q\}$ , en forment une base.

## CHAPITRE 5

# 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 5.1.13, le théorème de Sylvester 5.1.15, et l'algorithme de réduction en « somme de carrés » 5.2.2.

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

### 5.0. 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.

### 5.1. Formes bilinéaires symétriques et formes quadratiques

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

**Définition 5.1.1 (Formes bilinéaires symétriques).** — 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 :

$$\boxed{\phi(\lambda x + x', y) = \lambda\phi(x, y) + \phi(x', y)} \quad \boxed{\phi(x, \mu y + y') = \mu\phi(x, y) + \phi(x, y')}$$

et donc  $\boxed{\phi(\lambda x + x', \mu y + y') = \lambda\mu\phi(x, y) + \lambda\phi(x, y') + \mu\phi(x', y) + \phi(x', y')}$ . On rappelle (cf. 4.1.2) que les formes bilinéaires sur  $E$  forment un espace vectoriel, noté  $\mathcal{L}_2(E, k)$ .

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

(3) 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).

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



**Définition 5.1.2 (Matrices symétriques).** — (1) Une matrice  $A \in M_n(k)$  est dite **symétrique** si  $\boxed{{}^tA = A}$ . Ces matrices forment un sous-espace vectoriel de  $M_n(k)$ , qu'on notera  $\boxed{\text{MS}_n(k)}$ .

(2) On a  $\boxed{\dim \text{MS}_n(k) = n(n+1)/2}$ . 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  $\boxed{N = n(n-1)/2}$ . Puis, 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  $\dim \text{MS}_n(k) = n + N = n(n+1)/2$ .

**Définition et théorème 5.1.3 (Matrice d'une fbs et changement de base)**

Soit  $\phi$  une forme bilinéaire symétrique sur un  $k$ -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  $A = (a_{ij})_{i,j=1}^n \in M_n(k)$ , où  $a_{ij} = \phi(e_i, e_j)$ . Comme  $\phi(e_j, e_i) = \phi(e_i, e_j)$ , on a  $a_{ji} = a_{ij}$ , donc  ${}^tA = A$ , i.e.  $A \in \text{MS}_n(k)$ .

(2)  $\phi$  est entièrement déterminée par sa matrice  $A$  : en effet, par bilinéarité on a l'égalité :

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

Donc, si l'on note  $X, Y$  les vecteurs colonnes ci-dessus, on a la formule matricielle  $\boxed{\phi(X, Y) = {}^tXAY}$ .

(3) Réciproquement, pour tout  $A = (a_{ij})_{i,j=1}^n \in \text{MS}_n(k)$ , l'application  $\phi_A : E \times E \rightarrow k$  définie par  $\phi_A(\sum_{i=1}^n x_i e_i, \sum_{j=1}^n y_j e_j) = \sum_{i,j=1}^n a_{ij} x_i y_j$  est une forme bilinéaire symétrique sur  $E$ , et  $\text{Mat}_{\mathcal{B}}(\phi_A) = A$ . Donc, se donner une forme bilinéaire symétrique sur  $E$  « est la même chose » que se donner une matrice symétrique : de façon précise, l'application  $\mu_{\mathcal{B}} : \mathcal{S}_2(E) \rightarrow \text{MS}_n(k)$ ,  $\phi \mapsto \text{Mat}_{\mathcal{B}}(\phi)$  est un isomorphisme de  $k$ -espaces vectoriels.

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

$$(**) \quad \boxed{A' = \text{Mat}_{\mathcal{B}'}(\phi) = {}^tPAP}$$

**Terminologie 5.1.3.1.** — On dit que  $A, A' \in M_n(k)$  sont **congruentes** s'il existe  $P \in \text{GL}_n(k)$  telle que  $A' = {}^tPAP$ .

*Démonstration.* — (2) Comme  $\phi$  est bilinéaire, on a bien l'égalité (\*), qui montre que  $\phi$  est déterminée par sa matrice, donc que l'application  $\mu_{\mathcal{B}} : \mathcal{S}_2(E) \rightarrow \text{MS}_n(k)$ ,  $\phi \mapsto \text{Mat}_{\mathcal{B}}(\phi)$  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

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



Ceci prouve (2). Avant de prouver (3), remarquons déjà 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  $\mu_{\mathcal{B}}(s\phi + \psi) = s\mu_{\mathcal{B}}(\phi) + \mu_{\mathcal{B}}(\psi)$ .

Prouvons (3). Pour tout  $A = (a_{ij})_{i,j=1}^n \in \text{MS}_n(k)$ , l'application  $\phi_A : E \times E \rightarrow k$  définie par

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

est linéaire en les  $x_i$ , d'une part, et en les  $y_j$ , d'autre part, et elle vérifie :

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

donc c'est une forme bilinéaire symétrique 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 linéaire injective  $\mu_{\mathcal{B}} : \mathcal{S}_2(E, k) \rightarrow \text{MS}_n(k)$ ,  $\phi \mapsto \text{Mat}_{\mathcal{B}}(\phi)$  est aussi *surjective*, donc c'est un isomorphisme de  $k$ -espaces vectoriels. En particulier, *se donner une forme bilinéaire symétrique sur  $E$  est la même chose que se donner une matrice symétrique*.

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 donc :

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

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

#### Définition et proposition 5.1.4 (Rang d'une forme bilinéaire symétrique)

Soit  $\phi$  une forme bilinéaire symétrique sur un  $k$ -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  $P$  et  ${}^tP$  sont inversibles (on a  $({}^tP)^{-1} = {}^t(P^{-1})$ ), alors la matrice  $A' = \text{Mat}_{\mathcal{B}'}(\phi) = {}^tP A P$  a même rang que  $A$ .  $\square$

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

**Proposition 5.1.5.** — Soient  $E$  un  $k$ -espace vectoriel de dimension  $n$  et  $\phi$  une forme bilinéaire symétrique 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).  $\square$

**Définition 5.1.6 (Termes carrés et doubles produits).** — 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 5.1.3 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 et proposition 5.1.7 (Orthogonalité).** — Soit  $\phi$  une forme bilinéaire symétrique 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 :

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

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  $\lambda \in k$ , 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 5.1.8 (Orthogonal d'un sous-espace).** — Soit  $\phi$  une forme bilinéaire symétrique sur un  $k$ -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  $\dim F^\perp \geq \dim E - \dim F$ .



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

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

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

*Démonstration.* — 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 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 5.1.7,  $F^\perp$  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 :

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

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). De plus, 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  $\text{rang}(A) = n$ . Ceci prouve (2).

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

(1) On note  $\phi_F$  la forme bilinéaire symétrique 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  $F \cap F^\perp = \{x \in F \mid \phi(x, y) = 0, \forall y \in F\} = N(\phi_F)$ . Donc l'assertion (4) de 5.1.8 peut se récrire comme suit : « si  $\phi_F$  est non dégénérée, alors  $E = F \oplus F^\perp$  ».

**Remarque 5.1.9.1.** — 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  $\psi_F = 0$ .

**Définition 5.1.10 (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 l'égalité :

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

qui montre que :  $\phi$  est entièrement déterminée par  $Q$ ; on dit que  $\phi$  est la **forme polaire** de  $Q$  (et l'égalité (\*) ci-dessus est appelée « égalité 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 5.1.11 (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  $\lambda x \in C(Q)$ .

(3) On a toujours  $\boxed{N(Q) \subseteq C(Q)}$  mais l'inclusion est en général stricte. Par exemple, dans le cas considéré en 5.1.9.1, 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 5.1.12 (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 5.1.13 (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 5.1.10, 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 5.1.8, 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 5.1.8, 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 5.1.14 (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 5.1.13, 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 5.1.15 (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.  $\square$

## 5.2. 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 5.2.0.** — Soit  $E$  un  $k$ -espace vectoriel de dimension  $n$ . D'après 5.1.10, 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 5.2.1.** — 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 5.1.10, 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 5.2.2 (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

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



(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 5.1.13 et 5.1.15.

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 :

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

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 récrit 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  $(\dagger)$  :

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

(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  $b_{12} \neq 0$ . 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 :

$$\begin{cases} b_{12} x_1 x_2 \rightarrow b_{12} (x_1'^2 - x_2'^2) \\ b_{ij} x_i x_j \rightarrow b_{ij} x_i x_j & \text{si } i, j \geq 3 \end{cases} \quad \begin{cases} b_{1j} x_1 x_j \rightarrow b_{1j} (x'_1 + x'_2) x_j & \text{pour } j \geq 3 \\ b_{2j} x_2 x_j \rightarrow b_{2j} (x'_1 - x'_2) x_j & \text{pour } j \geq 3 \end{cases}$$

donc on obtient

$$Q(x'_1, \dots, x'_n) = b_{12}(x_1'^2 - x_2'^2) + \sum_{j=3}^n \left( (b_{1j} + b_{2j})x'_1x_j + (b_{1j} - b_{2j})x'_2x_j \right) + \sum_{3 \leq i < j \leq n} b_{ij}x'_ix'_j$$

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 5.2.3.** — 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 :

$$\begin{aligned} S &= b_{12}x_1x_2 + \sum_{j=3}^n b_{1j}x_1x_j + \sum_{j=3}^n b_{2j}x_2x_j = b_{12} \left( x_1 + \sum_{j=3}^n \frac{b_{2j}}{b_{12}}x_j \right) \left( x_2 + \sum_{j=3}^n \frac{b_{1j}}{b_{12}}x_j \right) - \sum_{3 \leq j, \ell \leq n} \frac{b_{1j}b_{2\ell}}{b_{12}}x_jx_\ell \\ &= b_{12} \underbrace{\left( x_1 + \sum_{j=3}^n \frac{b_{2j}}{b_{12}}x_j \right)}_{=X} \underbrace{\left( x_2 + \sum_{j=3}^n \frac{b_{1j}}{b_{12}}x_j \right)}_{=Y} - \sum_{j=3}^n \frac{b_{1j}b_{2j}}{b_{12}}x_j^2 - \sum_{3 \leq j < \ell \leq n} \frac{2b_{1j}b_{2\ell}}{b_{12}}x_jx_\ell \end{aligned}$$

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 :

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

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 5.2.4.** — 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 5.2.5.** — 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 :

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

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 5.2.6 (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 !

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

Dans ce chapitre,  $k$  est un corps arbitraire.

**Définition 5.3.1 (Formes bilinéaires alternées).** — Soit  $E$  un  $k$ -espace vectoriel (pas nécessairement de dimension finie). Une forme bilinéaire  $\phi$  sur  $E$  est dite **alternée** si elle vérifie  $\phi(x, x) = 0$  pour tout  $x \in E$ . On a vu (cf. 4.4.1) que 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  $\mathcal{A}_2(E, k)$ .

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

**Définition 5.3.2 (Matrices alternées).** — (1) Une matrice  $A \in M_n(k)$  est dite **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  $\mathbf{MA}_n(k)$ .

(2) On a  $\dim \mathbf{MA}_n(k) = n(n-1)/2$ . 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

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

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

diagonale, et il y a  $n$  coefficients diagonaux. Donc  $2N + n = n^2$ , d'où  $N = n(n-1)/2$ . Puis, une matrice alternée est déterminée par le choix de ses  $N$  coefficients au-dessus de la diagonale, d'où  $\dim \text{MA}_n(k) = N = n(n-1)/2$ .



### Définition et théorème 5.3.3 (Matrice d'une fba et changement de base)

Soit  $\phi$  une forme bilinéaire alternée sur un  $k$ -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  $A = (a_{ij})_{i,j=1}^n \in M_n(k)$ , où  $a_{ij} = \phi(e_i, e_j)$ . Comme  $\phi$  est alternée, on a  $a_{ii} = 0$  et  $a_{ji} = -a_{ij}$ , pour tout  $i \neq j$  donc  $A \in \text{MA}_n(k)$ .

(2)  $\phi$  est entièrement déterminée par sa matrice  $A$  : en effet, par bilinéarité on a l'égalité :

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

Donc, si l'on note  $X, Y$  les vecteurs colonnes ci-dessus, on a la formule matricielle  $\phi(X, Y) = {}^t X A Y$ .

(3) Réciproquement, pour tout  $A = (a_{ij})_{i,j=1}^n \in \text{MA}_n(k)$ , l'application  $\phi_A : E \times E \rightarrow k$  définie par  $\phi_A(\sum_{i=1}^n x_i e_i, \sum_{j=1}^n y_j e_j) = \sum_{i,j=1}^n a_{ij} x_i y_j$  est une forme bilinéaire alternée sur  $E$ , et  $\text{Mat}_{\mathcal{B}}(\phi_A) = A$ . Donc, se donner une forme bilinéaire alternée sur  $E$  « est la même chose » que se donner une matrice alternée : de façon précise, l'application  $\mu_{\mathcal{B}} : \mathcal{S}_2(E) \rightarrow \text{MA}_n(k)$ ,  $\phi \mapsto \text{Mat}_{\mathcal{B}}(\phi)$  est un isomorphisme de  $k$ -espaces vectoriels.

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

$$(**) \quad A' = \text{Mat}_{\mathcal{B}'}(\phi) = {}^t P A P.$$

La démonstration est identique à celle de 5.1.3.

### Définition et proposition 5.3.4 (Rang d'une fba. Formes symplectiques)

Soit  $\phi$  une forme bilinéaire alternée sur un  $k$ -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. Dans ce cas, on dit que  $\phi$  est une forme **symplectique** sur  $E$ .

La démonstration est identique à celle de 5.1.4

**Définition et proposition 5.3.5 (Orthogonalité).** — Soit  $\phi$  une forme bilinéaire alternée sur un  $k$ -espace vectoriel  $E$ .

(1) On dit que deux vecteurs  $x, y \in E$  sont **orthogonaux** (pour  $\phi$ ) si  $\phi(x, y) = 0$  ; ceci équivaut à dire que  $\phi(y, x) = 0$  (puisque  $\phi(y, x) = -\phi(x, y)$ ). 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  $\phi$ .

La démonstration est identique à celle de 5.1.7.





**Théorème 5.3.6 (Orthogonal d'un sous-espace).** — Soit  $\phi$  une forme bilinéaire alternée sur un  $k$ -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)  $N(\phi) = \{0\} \Leftrightarrow \phi$  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  $E = F \oplus F^\perp$ .

La démonstration est identique à celle de 5.3.6.

**Définition 5.3.7 (Restriction à un sous-espace).** — Soit  $\phi$  une forme bilinéaire alternée sur  $E$ , et soit  $F$  un sous-espace de  $E$ .

(1) On note  $\phi_F$  la forme bilinéaire alternée 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  $F \cap F^\perp = \{x \in F \mid \phi(x, y) = 0, \forall y \in F\} = N(\phi_F)$ . Donc l'assertion (4) de 5.1.8 peut se récrire comme suit : « si  $\phi_F$  est non dégénérée, alors  $E = F \oplus F^\perp$  ».



**Théorème 5.3.8 (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 5.3.6, 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 5.3.9 (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  $\overline{E} = E/N(\phi)$  une forme bilinéaire  $\overline{\phi}$  telle que

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

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

**Corollaire 5.3.10 (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 5.3.8.

*Démonstration.* — Ceci résulte de la proposition 5.3.9 et du théorème 5.3.8.  $\square$

## CHAPITRE 6

### ESPACES EUCLIDIENS ET GROUPE 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. 6.1.5). 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 6.2.9. 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.

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

**Définitions 6.1.1 (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 6.1.2.** — (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 10/7/2012

<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 8.6 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 6.1.3 (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 6.1.4 (Existence de b.o.n.).** — Soit  $E$  un espace euclidien de dimension  $n$ . Alors  $E$  admet une base orthonormée.

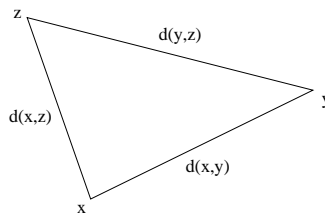
*Démonstration.* — D'après le théorème d'inertie de Sylvester 5.1.15, il existe une base  $(e_1, \dots, e_n)$  orthogonale (i.e.  $(e_i | e_j) = 0$  pour  $i \neq j$ ) et telle que  $(e_i | e_i) \in \{1, -1, 0\}$ ; or comme  $( | )$  est défini positif on a nécessairement  $(e_i | e_i) = 1$ , donc  $(e_1, \dots, e_n)$  est une b.o.n.  $\square$

**Définition 6.1.5 (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 6.1.6 (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)  $\forall x, y \in E \quad (x | y)^2 \leq Q(x)Q(y)$

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

*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  $Q(t_0x + y) = 0$  entraîne, puisque  $Q$  est définie positive,  $t_0x + y = 0$ , 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 6.1.7 (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 6.1.7.1.** — 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 6.1.8 (Angle non orienté de deux vecteurs non nuls).** — Soit  $E$ , muni de  $(\mid)$ , 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

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

donc il existe un unique  $\theta \in [0, \pi]$  tel que  $\cos(\theta) = \frac{(u \mid v)}{\|u\| \cdot \|v\|}$  i.e.  $(u \mid v) = \cos(\theta) \|u\| \cdot \|v\|$ . 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 6.1.9 (Isométries vectorielles).** — Soient  $E, F$  deux espaces euclidiens de même dimension  $n$ , notons  $(\mid)_E$  et  $\|\cdot\|_E$  (resp.  $(\mid)_F$  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 :  $\forall x \in E, \|x\|_E = \|f(x)\|_F$

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

(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 :

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

et comme  $\|x\|_E^2 = \|f(x)\|_F^2$  et  $\|y\|_E^2 = \|f(y)\|_F^2$ , on obtient que  $(x \mid y)_E = (f(x) \mid f(y))_F$ . 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 6.1.9.1.** — 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 6.1.10.** — (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 6.1.4,  $E$  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 6.1.11.** — On note  $O(n) = \{A \in M_n(\mathbb{R}) \mid {}^tAA = I_n\}$ . Rappelons (cf. 0.5.2) 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. 5.1.3). 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 6.1.9, on obtient donc les caractérisations suivantes de  $O(n)$ , chacune étant utile :

**Proposition 6.1.12 (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 6.3.2.1 plus bas).

**Remarque 6.1.13.** — 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.

## 6.2. 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 6.2.1 (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

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

*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

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

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 6.2.2.** — 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 6.2.3 (Adjoint d'un endomorphisme dans le cas euclidien)**

Soit  $E$  muni de  $(|)$  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 :

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

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 6.2.4 (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 6.2.5 (Endomorphismes auto-adjoints et formes bilinéaires symétriques)**

Soit  $E$  muni de  $(|)$  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  $(|)$  tel que :

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

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

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

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 :

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

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

On a maintenant le théorème important et utile suivant.

### **Théorème 6.2.6 (Diagonalisation des endomorphismes auto-adjoints)**

Soient  $E$  muni de  $(|)$  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 6.2.7 (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 6.2.8 (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 6.2.8.  $\square$

*Démonstration de la proposition 6.2.8.* — 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 6.2.8 et du théorème 6.2.6.  $\square$



**Théorème 6.2.9 (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 6.2.9.1.** — 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 6.2.5. D'après le théorème 6.2.6, 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 6.2.10 (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 6.2.11 (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 6.2.12.** — 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 6.2.9, 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 6.2.13.** — Pour des applications géométriques du théorème 6.2.9, voir l'étude des coniques et quadriques dans le chapitre suivant.

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



**Définition 6.3.1 (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. 5.1.9) 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 6.3.2 (Projection orthogonale sur un sous-espace)**

Soit  $E$ , muni de  $(\mid)$ , un espace euclidien et soient  $F$  un sous-espace et  $F^\perp$  son orthogonal pour  $(\mid)$ .

(1) On a  $E = F \oplus F^\perp$ . 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  $\pi_F(v) = (v \mid e_1)e_1 + \cdots + (v \mid e_r)e_r$  pour tout  $v \in E$ .

(3) On a  $(F^\perp)^\perp = F$  donc la projection orthogonale  $\pi_{F^\perp}$  sur  $F^\perp$  n'est autre que  $\text{id}_E - \pi_F$ , i.e. on a  $\text{id}_E = \pi_F + \pi_{F^\perp}$ .

*Démonstration.* — Comme les formes bilinéaires symétriques  $(\mid)$  et  $(\mid)_F$  sont définies positives, donc non-dégénérées, on a  $(F^\perp)^\perp = F$  et  $E = F \oplus F^\perp$  d'après 5.1.8. 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).

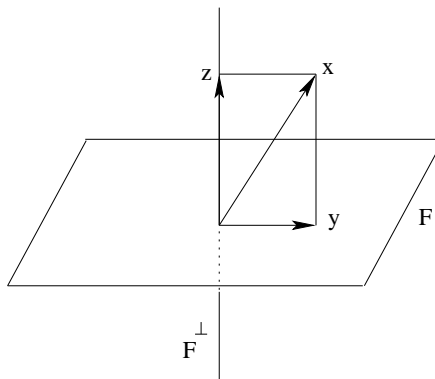
Prouvons (2). Soit  $r = \dim F$  et soit  $(e_1, \dots, e_r)$  une b.o.n. de  $F$ . Pour tout  $v \in E$ , notons provisoirement

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

Alors, pour  $j = 1, \dots, r$ , on a  $(v - \pi(v) \mid e_j) = (v \mid e_j) - \sum_{i=1}^r (v \mid e_i) \underbrace{(e_i \mid e_j)}_{\substack{=1 \text{ si } i=j \\ =0 \text{ si } i \neq j}} = 0$ , d'où  $v - \pi(v) \in F^\perp$ ,

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).  $\square$

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



**Remarque 6.3.2.1.** — 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. 6.1.12).



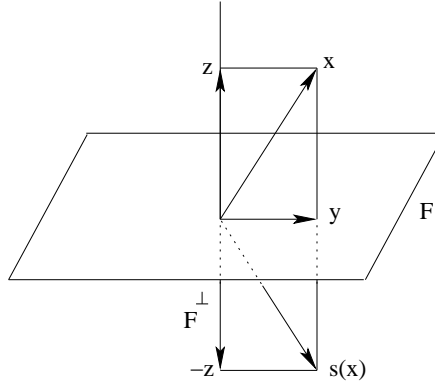
**Définition et proposition 6.3.3 (Symétries orthogonales).** — Soient  $E$  un espace euclidien de dimension  $n$  et  $F$  un sous-espace de dimension  $r$ .

(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 :

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

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  $\boxed{\det s_F = (-1)^{n-r}}$ .



**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  $(\pi_F(v) \mid \pi_{F^\perp}(v)) = 0$ , on a d'après l'égalité de Pythagore (cf. 6.1.7) :

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

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 6.3.4 (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 6.3.2 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

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

donc

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

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 :

$$(6.3.4.1) \quad \boxed{\forall x \in E, \quad s_H(x) = x - 2 \frac{(x \mid v_0)}{(v_0 \mid v_0)} v_0.}$$

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 6.3.5 (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 6.3.5.1.** — (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.  $G_n = \{x \in V_n \mid (x \mid e_i) = 0 \text{ pour } i = 1, \dots, n-1\}$ , est de dimension  $n - (n-1) = 1$ , et  $\sum_{i=1}^{n-1} (v_n \mid e_i) e_i$  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  $(e_n \mid v_n) = \|e'_n\| > 0$ .

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

**Exemple 6.3.5.2.** — On munit  $E = \mathbb{R}[X]$  du produit scalaire défini par  $(P \mid Q) = \int_{-1}^1 P(t)Q(t)dt$ . 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  $(1 \mid 1) = 2$  donc on prend  $e_0 = \frac{1}{\sqrt{2}}$ . Alors  $(X \mid e_0) = \frac{1}{\sqrt{2}} \int_{-1}^1 t dt = 0$  et  $(X \mid X) = \int_{-1}^1 t^2 dt = \frac{2}{3}$ , d'où

$$e_1 = \frac{\sqrt{3}}{\sqrt{2}} X.$$

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

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

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

**Notation 6.4.0.** — 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 6.4.1 (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 6.4.2 (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  $(x | y) = x_1y_1 + \dots + x_ny_n$  et l'on note  $O(n)$  le groupe des isométries :

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

(voir 6.1.12 pour d'autres caractérisations de  $O(n)$ ).



**Proposition 6.4.3.** — Soit  $A \in O(n)$ . Alors :

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

(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.,  $(v_+ | v_-) = 0$  pour tout  $v_+ \in V_+$ ,  $v_- \in V_-$ .

*Démonstration.* — (1) On a  $1 = \det(I_n) = \det({}^tAA) = \det({}^tA) \cdot \det(A)$ , or on sait (cf. 2.1.1) que  $\det({}^tA) = \det(A)$  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

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

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

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

$$(v_+ | v_-) = (Av_+ | Av_-) = (v_+ | -v_-) = -(v_+ | v_-)$$

d'où  $(v_+ | v_-) = 0$ . □

**Définition 6.4.4 (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**écial **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 6.4.5 (Groupe  $SO(n)$ ).** — (1) On pose  $\boxed{SO(n) = \{A \in O(n) \mid \det(A) = 1\}}$ , c'est un sous-groupe de  $O(n)$ , appelé le groupe **s**écial **o**rtogonal. On le note aussi parfois  $O^+(n)$ . Les éléments de  $SO(n)$  s'appellent les isométries **d**irectes de  $\mathbb{R}^n$ .

(2) On pose aussi :  $\boxed{O^-(n) = \{A \in O(n) \mid \det(A) = -1\}}$ , 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 **b**ijection 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 6.4.6.** — 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 6.1.12, que  $P \in O(n)$ , d'où  $\det(P) = \pm 1$ , d'après 6.4.3. Donc  $\det(P)$  est  $> 0$  (resp.  $< 0$ ) si et seulement si il égale 1 (resp.  $-1$ ). La proposition en découle. □

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 6.4.7.** — 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 6.4.8 (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  $\theta = \widehat{ww'}$ , 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 6.4.9.** — 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  $\widehat{e_1 u} = \theta$ . En effet, soit  $u = ae_1 + be_2$  un tel vecteur, alors  $a = (u \mid e_1) = \cos(\theta)$ , 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 6.4.10 (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 6.4.11 (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 6.4.7, 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  $\theta \in 2\pi\mathbb{Z}$ .  $\square$

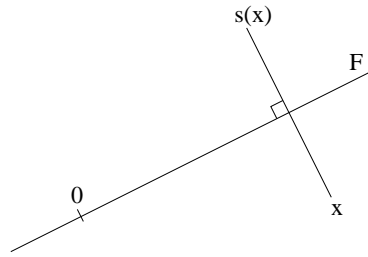
Étudions maintenant le cas des éléments de  $O^-(2)$ , i.e. des matrices de la forme

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

c.-à-d., d'après le lemme 6.4.7, 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 6.4.3,  $D_+$  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 6.3.4,  $s_\Delta$  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

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

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 6.4.12 (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  $\mathbb{R}u_{\theta/2}$ .

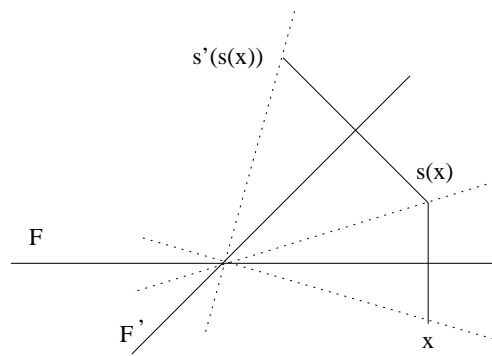
Enfin, on peut vérifier par un calcul matriciel (que le lecteur est invité à faire!) que  $J_{\theta'} J_\theta = R_{\theta' - \theta}$  d'où

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

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  $\varphi, \theta \in \mathbb{R}$  :

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

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

$s_\theta \circ s_\varphi = r_{2(\theta-\varphi)}$  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 6.4.14.** — Soit  $\theta \in \mathbb{R}$  et soit  $r_\theta$  la rotation d'angle  $\theta$  dans  $\mathbb{R}^2$  (cf. 6.4.10).

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

(2) La matrice de  $r_\theta$  dans la base  $\mathcal{C}_0 = (e_2, e_1)$  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.  $A \in \text{SO}(3)$ . 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

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

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 (\*),  $u = \text{id}$  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. 6.4.10), 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. 6.4.10) 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. 6.4.14). 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 6.4.15.** — 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 6.4.16.** — 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  $\mu = (x \mid f_3)$ ). 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  $A \in \text{SO}(3)$ ; traitons maintenant le cas où  $A \in \text{O}^-(3)$ .

**2ème cas :**  $\det(A) = -1$ , i.e.  $A \in \text{O}^-(3)$ . 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. 6.4.10), 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. 6.3.3 et 6.3.4).

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 6.4.17.** — 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 6.4.18 (Classification des éléments de  $O(3)$ ).** — On a  $O(3) = SO(3) \cup O^-(3)$  (réunion disjointe).

(1) Si  $A \in SO(3)$ , 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  $P$** .

**6.4.19. 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  $(|)$  et de la norme euclidienne associée  $\|x\| = \sqrt{(x|x)}$ . 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  $w \mapsto (x|w)$ .

(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  $(|)$  que, pour tout  $x \in E$ , l'application  $\phi_x : w \mapsto (x|w)$  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

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

qui est le noyau de  $(|)$ . Or ce noyau est nul, puisque  $(x|x) = 0$  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  $\det_{\mathcal{B}_0}(u, v, w) = (f(u, v) | w)$  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

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

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  $f_1 = (u \wedge v | e_1) = \det_{\mathcal{B}_0}(u, v, e_1) = \begin{vmatrix} u_1 & v_1 & 1 \\ u_2 & v_2 & 0 \\ u_3 & v_3 & 0 \end{vmatrix} = u_2 v_3 - u_3 v_2$ . On montre de même que  $f_2 = (u \wedge v | e_2) = u_3 v_1 - u_1 v_3$  et  $f_3 = (u \wedge v | e_3) = u_1 v_2 - u_2 v_1$ . 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 :

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

(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

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

Comme  $(|)$  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  $0 = \det_{\mathcal{B}_0}(u, u, u \wedge u) = (u \wedge u | u \wedge u)$ , 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  $A$ . Dédurre 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  $\det_{\mathcal{B}_0}(x, y, x \wedge y) = (x \wedge y \mid x \wedge y) = \|x \wedge y\|^2$ . 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

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

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

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

donc

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

## 6.5. 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 6.5.1.** — (1) *Le morphisme de groupes  $f : \mathbb{R} \rightarrow S^1$ ,  $t \mapsto e^{it}$  est surjectif.*

(2) *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  $\cos(t) \geq 0$  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 :

$$\begin{aligned} (\dagger) \quad 2\pi\mathbb{Z} \subseteq \text{Ker}(f) &= \{T \in \mathbb{R} \mid e^{i(T+t)} = e^{it}, \quad \forall t \in \mathbb{R}\} = \{T \in \mathbb{R} \mid \cos(T+t) = \cos(t), \quad \forall t \in \mathbb{R}\} \\ &= \{T \in \mathbb{R} \mid \sin((T+t)) = \sin(t), \quad \forall t \in \mathbb{R}\}. \end{aligned}$$

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  $(\dagger)$  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 6.5.2.** — 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

$$L(S^1) = \int_0^{2\pi} \sqrt{1} dt = 2\pi.$$

## 6.6. Appendice $(\dagger)$ : décomposition d'Iwasawa de $\text{GL}_n(\mathbb{R})$

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 6.6.1.** —  $\text{TS}_n^+(\mathbb{R})$  est un sous-groupe de  $\text{GL}_n(\mathbb{R})$ .

*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 6.6.2 (Décomposition d'Iwasawa).** — *L'application ci-dessous est bijective :*

$$O(n) \times \mathrm{TS}_n^+(\mathbb{R}) \rightarrow \mathrm{GL}_n(\mathbb{R}), \quad (K, N) \mapsto KN$$

(Attention, ce n'est pas un morphisme de groupes, i.e.  $(KN)(K'N') \neq KK'NN'$  en général.)

*Démonstration.* — Soit  $P \in \mathrm{GL}_n(\mathbb{R})$ , 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  $P = \mathrm{Mat}_{\mathcal{B}_0}(\mathcal{B})$ , 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  $\mathrm{Vect}(v_1, \dots, v_r) = \mathrm{Vect}(f_1, \dots, f_r)$ , et  $(v_r \mid f_r) > 0$ , et ceci équivaut à dire que la matrice de passage  $N = \mathrm{Mat}_{\mathcal{C}}(\mathcal{B})$  appartient à  $\mathrm{TS}_n^+(\mathbb{R})$ . D'autre part, comme  $\mathcal{C}$  est orthonormée, la matrice de passage  $K = \mathrm{Mat}_{\mathcal{B}_0}(\mathcal{C})$  appartient à  $O(n)$ . D'après la formule de changement de base, on a

$$P = \mathrm{Mat}_{\mathcal{B}_0}(\mathcal{B}) = \mathrm{Mat}_{\mathcal{B}_0}(\mathcal{C}) \cdot \mathrm{Mat}_{\mathcal{C}}(\mathcal{B}) = KN$$

ce qui prouve l'existence.

Montrons l'unicité : supposons qu'on ait  $P = K' M$  avec  $K' \in O(n)$  et  $M = (m_{ij})_{i,j=1}^n \in \mathrm{TS}_n^+(\mathbb{R})$ , 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  $\mathrm{Mat}_{\mathcal{B}_0}(\mathcal{C}) = K'$  et la matrice exprimant  $\mathcal{B}$  dans la base  $\mathcal{C}$  est

$$\mathrm{Mat}_{\mathcal{C}}(\mathcal{B}) = \mathrm{Mat}_{\mathcal{C}}(\mathcal{B}_0) \cdot \mathrm{Mat}_{\mathcal{B}_0}(\mathcal{B}) = K'^{-1}P = M.$$

Alors, pour tout  $r = 1, \dots, n$ , on a  $v_r = \sum_{i=1}^r m_{ir} f_i$ , avec  $(v_r \mid f_r) = m_{rr} > 0$ . 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  $V_i = \mathrm{Vect}(v_1, \dots, v_i)$ , 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 6.6.3.** — (1) En appliquant ce qui précède à  $P^{-1}$ , on voit qu'il existe un couple unique  $(K, N) \in O(n) \times \mathrm{TS}_n^+(\mathbb{R})$  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  $\mathrm{TS}_n^+(\mathbb{R}) \times O(n) \rightarrow \mathrm{GL}_n(\mathbb{R})$ ,  $(N, K) \mapsto NK$  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  $\|(a_{ij})\|_\infty = \mathrm{Max}_{i,j} |a_{ij}|$  ou  $\|(a_{ij})\|_2 = \sqrt{\sum_{i,j} a_{ij}^2}$ , et donc les sous-ensembles  $O(n)$ ,  $\mathrm{TS}_n^+(\mathbb{R})$  et  $\mathrm{GL}_n(\mathbb{R})$  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  $\mathrm{GL}_n(\mathbb{R}) \rightarrow \mathrm{GL}_n(\mathbb{R})$ ,  $A \mapsto A^{-1}$  est continue, car  $A^{-1} = (1/\det(A)) {}^t\tilde{A}$ , où  $\tilde{A}$  désigne la matrice des cofacteurs de  $A$  (cf. 1.4.3). 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  $N^{-1} = \mathrm{Mat}_{(v_1, \dots, v_n)}(e_1, \dots, e_n)$  sont des fonctions continues des produits scalaires  $(v_i \mid v_j)$ .

## CHAPITRE 7

### ESPACES AFFINES, CONIQUES ET QUADRIQUES

**Résumé :** Dans ce chapitre, on introduit la notion importante d'espace affine. De façon imagée, un espace affine  $\mathcal{E}$  est « la même chose » qu'un espace vectoriel  $E$  dont on a oublié l'origine ; c'est le cadre naturel pour faire de la géométrie. Dans la section 1, on commence par introduire les notions de repère, de changement de repère, et d'application affine. Dans la section 2, on introduit les barycentres, et la notion de sous-espace affine. Dans la section 3, on incorpore les résultats du Chap. 6 en introduisant les « *espaces affines euclidiens* » et leurs isométries ; en particulier on étudie en profondeur les isométries du plan affine et de l'espace affine de dimension 3. On termine le chapitre par l'étude des coniques (dans le plan) et des quadriques (dans l'espace). Comme le précédent, ce chapitre contient beaucoup de résultats nouveaux et importants, qu'il faut essayer d'assimiler !

On a indiqué par des symboles  $\diamond$  les définitions, exemples et résultats fondamentaux.

#### 7.1. Espaces affines réels

**Exemple 7.1.1.** — Dans  $\mathbb{R}^2$ , soit  $\mathcal{D}_1$  la « droite affine » d'équation  $x_1 + x_2 = 1$ , ce n'est **pas un sous-espace vectoriel** de  $\mathbb{R}^2$  : si  $x = (x_1, x_2)$  et  $y = (y_1, y_2)$  appartiennent à  $\mathcal{D}_1$ , alors  $x + y \notin \mathcal{D}_1$  (car  $x_1 + y_1 + x_2 + y_2 = 2$ ), et si  $t \in \mathbb{R}$ ,  $t \neq 1$ , alors  $tx = (tx_1, tx_2) \notin \mathcal{D}_1$ . Mais le vecteur  $y - x = \begin{pmatrix} y_1 - x_1 \\ y_2 - x_2 \end{pmatrix}$  appartient à la droite vectorielle  $D$  d'équation  $z_1 + z_2 = 0$ , engendrée par le vecteur  $\vec{u} = e_1 - e_2$ . On a donc une application

$$\mathcal{D}_1 \times \mathcal{D}_1 \rightarrow D, \quad (x, y) \mapsto y - x.$$

Notons  $\overrightarrow{xy}$  le vecteur  $y - x \in \mathbb{R}^2$ , alors pour  $x, y, z \in \mathcal{D}_1$  on a

$$(1) \quad \overrightarrow{xz} = z - x = (y - x) + (z - y) = \overrightarrow{xy} + \overrightarrow{yz}.$$

D'autre part, pour tout  $x = (x_1, x_2)$  appartenant à  $\mathcal{D}_1$ , et pour tout vecteur  $\vec{v} = t\vec{u} \in D$ , il existe un unique  $y \in \mathcal{D}_1$  tel que  $\overrightarrow{xy} = \vec{v}$ , à savoir  $y = (x_1 + t, x_2 - t)$ , et l'on notera  $y = x + \vec{v}$ . Donc :

$$(2) \quad \text{pour tout } x \in \mathcal{D}_1, \text{ l'application } \{x\} \times \mathcal{D}_1 \rightarrow D, \quad (x, y) \mapsto \overrightarrow{xy} \text{ est bijective}$$

et ceci équivaut aussi à dire que :

$$(2') \quad \text{l'application } \mathcal{D}_1 \times \mathcal{D}_1 \rightarrow \mathcal{D}_1 \times D, \quad (x, y) \mapsto (x, \overrightarrow{xy}) \text{ est bijective ;}$$

son inverse est notée  $(x, \vec{v}) \mapsto (x, x + \vec{v})$ .

**Remarque 7.1.2.** — Ce qui précède est une propriété intrinsèque de  $\mathcal{D}_1$ , et ne dépend pas des coordonnées  $(x_1, x_2)$  choisies. Par exemple, soient  $f_1 = (e_1 - e_2)/2$  et  $f_2 = (e_1 + e_2)/2$ , alors

$$\text{Mat}_{(e_1, e_2)}(f_1, f_2) = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$$


est inversible, donc  $\mathcal{C} = (f_1, f_2)$  est une base de  $\mathbb{R}^2$  et les coordonnées  $(y_1, y_2)$  dans la base  $\mathcal{C}$  sont données par

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = P \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} y_1 + y_2 \\ -y_1 + y_2 \end{pmatrix} \quad \text{d'où} \quad \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} x_1 - x_2 \\ x_1 + x_2 \end{pmatrix}.$$

Avec ces nouvelles coordonnées,  $\mathcal{D}_1$  est la « droite affine » d'équation  $y_2 = 1$  et  $D$  est la droite vectorielle d'équation  $y_2 = 0$  ; si  $P$  et  $Q$ , de coordonnées  $(p, 1)$  et  $(q, 1)$ , appartiennent à  $\mathcal{D}_1$ , le vecteur  $\overrightarrow{PQ} = (q - p)f_1$

appartient à  $D$ , et réciproquement, pour tout  $\vec{v} = tf_1 \in D$ , le point  $P + tf_1 = (p + t, 1)$  est l'unique point  $Q$  de  $\mathcal{D}_1$  tel que  $\overrightarrow{PQ} = tf_1$ .

Ce qui précède est un cas particulier de, et conduit à, la définition suivante :

 **Définition 7.1.3 (Espaces affines réels).** — Soit  $E$  un  $\mathbb{R}$ -espace vectoriel. Un **espace affine**  $\mathcal{E}$  de **direction**  $E$  est un ensemble non-vidé  $\mathcal{E}$  muni d'une application

$$\mathcal{E} \times \mathcal{E} \rightarrow E, \quad (x, y) \mapsto \overrightarrow{xy}$$

vérifiant les deux propriétés suivantes :


(1) Relation de Chasles :  $\overrightarrow{xz} = \overrightarrow{xy} + \overrightarrow{yz}$   $\forall x, y, z \in \mathcal{E}$ .

(2) L'application  $\mathcal{E} \times \mathcal{E} \rightarrow \mathcal{E} \times E, (x, y) \mapsto (x, \overrightarrow{xy})$  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  $\overrightarrow{xy} = \vec{u}$ .

**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  $\dim \mathcal{E} = \dim E$ .

**Remarque 7.1.4.** — Dans la définition précédente, on peut remplacer  $\mathbb{R}$  par un corps arbitraire  $k$ , on obtient ainsi la notion d'espace affine sur  $k$ . Comme nous ne considérerons que des espaces affines sur  $\mathbb{R}$ , nous dirons simplement « espaces affines », sans préciser « réels ». De plus, pour abrégé, on écrira souvent : « soit  $(\mathcal{E}, E)$  un espace affine », ou même : « soit  $\mathcal{E}$  un espace affine (de dimension  $n$ ) », sans préciser l'espace vectoriel  $E$  (appelé la « direction » de  $\mathcal{E}$ ).

 **Exemple 7.1.5.** — 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 7.1.6 (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 un 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 7.1.7.** — 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 7.1.8 (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

$$\boxed{X = PX' + T} \quad \boxed{X' = P^{-1}(X - T)}$$



*Démonstration.* — Soient  $X, X'$  comme ci-dessus et notons  $Y$  le vecteur colonne  $PX'$ . Alors, d'après la formule de changement de base dans  $E$ , on a :

$$\overrightarrow{O'M} = x'_1 f_1 + \cdots + x'_n f_n = y_1 e_1 + \cdots + y_n e_n,$$

et comme  $\overrightarrow{OM} = \overrightarrow{OO'} + \overrightarrow{O'M}$  on obtient

$$X = \overrightarrow{OM} = \sum_{i=1}^n (t_i + y_i) e_i = T + PX'$$

et l'on a donc aussi  $X' = P^{-1}(X - T)$ , ce que l'on peut aussi retrouver en écrivant que

$$\sum_{i=1}^n x'_i f_i = \overrightarrow{O'M} = \overrightarrow{OM} - \overrightarrow{OO'} = \sum_{i=1}^n (x_i - t_i) e_i$$

d'où  $X' = P^{-1}(X - T)$ . □



**Définition et proposition 7.1.9 (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  $f$ .

*Démonstration.* — Fixons un point  $O \in \mathcal{E}$ . Comme l'application  $\theta : E \rightarrow \mathcal{E}$ ,  $\vec{u} \mapsto O + \vec{u}$  est une bijection, d'inverse  $P \mapsto \overrightarrow{OP}$  alors l'application  $\theta^{-1} \circ f \circ \theta$  est l'unique application  $\phi : E \rightarrow E$  telle que

$$\forall \vec{u} \in E, \quad f(O + \vec{u}) = f(O) + \phi(\vec{u}) \quad (\text{i.e. } \overrightarrow{f(O)f(P)} = \phi(\overrightarrow{OP}), \text{ pour tout } P \in \mathcal{E}).$$

**On suppose que**, pour ce choix de  $O$ , l'application  $\phi : E \rightarrow E$  est **linéaire**. Soit  $I \in \mathcal{E}$  arbitraire. D'après la relation de Chasles, et la définition de  $\phi$ , on a, pour tout  $P \in \mathcal{E}$  :

$$\overrightarrow{f(I)f(P)} = \overrightarrow{f(O)f(P)} - \overrightarrow{f(O)f(I)} = \phi(\overrightarrow{OP}) - \phi(\overrightarrow{OI})$$

et comme  $\phi$  est supposée linéaire, ceci égale  $\phi(\overrightarrow{OP} - \overrightarrow{OI})$ . D'après la relation de Chasles, à nouveau, on a  $\overrightarrow{OP} - \overrightarrow{OI} = \overrightarrow{IP}$ . Posant  $\overrightarrow{f} = \phi$ , on a donc obtenu :

$$\forall I, P \in \mathcal{E}, \quad \overrightarrow{f(I)f(P)} = \overrightarrow{f}(\overrightarrow{IP}).$$

La proposition est démontrée. □

**Remarque 7.1.10.** — 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 7.1.11 (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  $A = \text{Mat}_{\mathcal{C}, \mathcal{B}}(\overrightarrow{f}) \in M_{n,m}(\mathbb{R})$  et soient  $(b'_1, \dots, b'_n)$  les coordonnées de  $f(O)$  dans  $\mathcal{R}'$ .

Alors, pour tout  $P \in \mathcal{P}$ , 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 :

$$\boxed{\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = A \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix} + \begin{pmatrix} b'_1 \\ \vdots \\ b'_n \end{pmatrix}.}$$

*Démonstration.* — En effet, on a  $\begin{pmatrix} y_1 - b'_1 \\ \vdots \\ y_n - b'_n \end{pmatrix} = \overrightarrow{f(O)f(P)} = \overrightarrow{f}(\overrightarrow{OP}) = A \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix}$ . □



**Définition et proposition 7.1.12 (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  $\text{id}_E$ .

*Démonstration.* — On fixe  $O \in \mathcal{E}$  et pour tout  $P \in \mathcal{E}$  on pose  $P' = P + \vec{u}$ . Alors

$$\overrightarrow{O'P'} = \underbrace{\overrightarrow{O'O}}_{=-\vec{u}} + \overrightarrow{OP} + \underbrace{\overrightarrow{PP'}}_{=\vec{u}} = \overrightarrow{OP}.$$

Avec les notations de 7.1.9, ceci montre que l'application  $\phi$  égale  $\text{id}_E$ , qui est linéaire. Donc  $t_{-\vec{u}}$  est une application affine, de partie linéaire  $\text{id}_E$ .  $\square$



**Lemme 7.1.13.** — 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  $t_{-\vec{u}}$ .

*Démonstration.* — Pour tout  $P \in \mathcal{E}$ , on a  $(t_{-\vec{u}} \circ t_{-\vec{v}})(P) = t_{-\vec{u}}(P + \vec{v}) = P + \vec{v} + \vec{u} = t_{-\vec{u}+\vec{v}}(P)$ . Ceci prouve la première assertion, et la seconde en découle facilement.  $\square$



**Proposition 7.1.14 (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  $\boxed{\overrightarrow{(g \circ f)} = \overrightarrow{g} \circ \overrightarrow{f}}$ .

*Démonstration.* — Fixons  $O \in \mathcal{E}$  et posons  $O' = f(O)$ . Pour tout  $M \in \mathcal{E}$  et  $P \in \mathcal{E}'$ , on a :

$$\overrightarrow{f(O)f(M)} = \overrightarrow{f}(\overrightarrow{OM}) \quad \text{et} \quad \overrightarrow{g(O')g(P)} = \overrightarrow{g}(\overrightarrow{O'P}).$$

En particulier, appliquant la deuxième égalité au cas où  $P = f(M)$ , on obtient :

$$\overrightarrow{(g \circ f)(O)(g \circ f)(M)} = \overrightarrow{g(f(O))g(f(M))} = \overrightarrow{g}(\overrightarrow{f(O)f(M)}) = \overrightarrow{g}(\overrightarrow{f}(\overrightarrow{OM})) = (\overrightarrow{g} \circ \overrightarrow{f})(\overrightarrow{OM}).$$

D'après 7.1.9, ceci prouve que  $g \circ f$  est affine, de partie linéaire  $\overrightarrow{g} \circ \overrightarrow{f}$ .  $\square$

**Théorème et définition 7.1.15 (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  $f = t_{\overrightarrow{Of(O)}} \circ \phi_O$ , 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  $f \circ t_{-\vec{u}} \circ f^{-1} = t_{\overrightarrow{f}(\vec{u})}$ .

*Démonstration.* — (1) découle de 7.1.14. Prouvons (2). Soient  $O \in \mathcal{E}$  et  $f \in \text{TAff}(\mathcal{E})$ , posons  $\vec{u} = \overrightarrow{Of(O)}$  et  $g = t_{-\vec{u}} \circ f \in \text{TAff}(\mathcal{E})$ . Alors  $g$  est affine, sa partie linéaire est  $\text{id}_E \circ \overrightarrow{f} = \overrightarrow{f}$ , et l'on a  $g(O) = O$ . Donc, pour tout  $P \in \mathcal{E}$ , on a  $\overrightarrow{Og(P)} = \overrightarrow{f}(\overrightarrow{OP})$  donc  $g$  égale  $\phi_O$ , et l'on a  $f = t_{-\vec{u}} \circ \phi_O$ , ce qui prouve (2). De plus, comme  $t_{-\vec{u}}$  est bijective et  $f = t_{-\vec{u}} \circ \phi_O$ , on a :

$$f \text{ bijective} \iff \phi_O \text{ bijective} \iff \overrightarrow{f} \text{ bijective}$$

ce qui prouve (3).

Prouvons (4). Soit  $f \in \text{GA}(\mathcal{E})$  et soient  $O, P \in \mathcal{E}$ . Posons  $O' = f^{-1}(O)$  et  $P' = f^{-1}(P)$ , comme  $f$  est affine et  $f(O') = O$ ,  $f(P') = P$ , on a :

$$\overrightarrow{f}(O'P') = \overrightarrow{f(O')f(P')} = \overrightarrow{OP}$$

et comme  $\overrightarrow{f}$  est bijective, d'après (3), on obtient que

$$\overrightarrow{f^{-1}(O)f^{-1}(P)} = \overrightarrow{O'P'} = (\overrightarrow{f})^{-1}(\overrightarrow{OP})$$

et ceci prouve que  $f^{-1}$  est affine, de partie linéaire  $(\overrightarrow{f})^{-1}$ . Le reste du point (4) découle alors de 7.1.14.

Prouvons (5). Pour tout  $P \in \mathcal{E}$ , on a

$$(f \circ t_{\vec{u}})(P) = f(P + \vec{u}) = f(P) + \vec{f}(\vec{u}) = (t_{\vec{f}(\vec{u})} \circ f)(P)$$

ce qui prouve le premier point de (5), et le second en découle.  $\square$

**Remarque 7.1.16.** — 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  $\vec{f} : E \rightarrow E'$  est une application linéaire bijective, donc  $E$  et  $E'$  sont isomorphes; en particulier ils ont même dimension, donc  $\dim \mathcal{E} = \dim \mathcal{E}'$ .

## 7.2. Barycentres et sous-espaces affines

**Théorème et définition 7.2.1 (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  $t_1 + \dots + t_n = 1$ . Alors il existe un unique point  $G \in \mathcal{E}$  tel que :

$$(*) \quad \forall I \in \mathcal{E}, \quad \vec{IG} = t_1 \vec{IP}_1 + \dots + t_n \vec{IP}_n$$

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 :

$$(**) \quad G = t_1 P_1 + \dots + t_n P_n.$$

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  $P_1, \dots, P_n$ .

*Démonstration.* — Fixons une origine  $O \in \mathcal{E}$ . Si  $G$  existe, il vérifie nécessairement

$$(\dagger) \quad \vec{OG} = t_1 \vec{OP}_1 + \dots + t_n \vec{OP}_n$$

d'où l'unicité de  $G$  (s'il existe). Réciproquement, définissons  $G$  par l'égalité ci-dessus et montrons que  $G$  vérifie (\*) pour tout  $I \in \mathcal{E}$ . D'après la relation de Chasles et (\dagger), on a

$$\sum_{i=1}^n t_i \vec{IP}_i = \sum_{i=1}^n t_i (\vec{IO} + \vec{OP}_i) = \underbrace{\left( \sum_{i=1}^n t_i \right)}_{=1} \vec{IO} + \sum_{i=1}^n t_i \vec{OP}_i = \vec{IO} + \vec{OG} = \vec{IG}.$$

Ceci prouve le théorème.  $\square$

**Remarque 7.2.1.1.** — 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 7.2.2 (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

$$P = A + t\vec{AB} = B + (1-t)\vec{BA}, \quad \text{avec } t \in [0, 1],$$

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 7.2.3 (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  $G$  est le barycentre de  $(H, \mu)$  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  $(A_n, 1/n)$ .

*Démonstration.* — Soit  $O \in \mathcal{E}$ . Alors  $H$  est défini par  $\vec{OH} = \frac{1}{\mu} \sum_{i=1}^{n-1} t_i \vec{OA}_i$  et le barycentre  $G'$  de  $(H, \mu)$  et de  $(A_n, t_n)$  est défini par

$$\vec{OG}' = \mu \vec{OH} + t_n \vec{OA}_n = \sum_{i=1}^{n-1} t_i \vec{OA}_i + t_n \vec{OA}_n = \vec{OG}$$

d'où  $G' = G$ .  $\square$



**Exemple 7.2.4 (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 7.2.5 (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  $(f(A_i), t_i)$ .

*Démonstration.* — Fixons  $O \in \mathcal{E}$  et pour tout  $P \in \mathcal{E}$  notons  $P' = f(P)$ . Par définition, on a  $\overrightarrow{OG} = \sum_{i=1}^n t_i \overrightarrow{OA_i}$  d'où, puisque  $\overrightarrow{f}$  est linéaire :

$$\overrightarrow{O'G'} = \overrightarrow{f}\left(\sum_{i=1}^n t_i \overrightarrow{OA_i}\right) = \sum_{i=1}^n t_i \overrightarrow{f}(\overrightarrow{OA_i}) = \sum_{i=1}^n t_i \overrightarrow{O'A'_i}.$$

Ceci montre que  $G'$  est le barycentre des points  $(A'_i, t_i)$ . □



**Définition et proposition 7.2.6 (Sous-espaces affines).** — Soient  $(\mathcal{E}, E)$  un espace affine et  $\mathcal{F}$  un sous-ensemble **non-vide**. 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  $\dim \mathcal{F} = \dim F$ .

*Démonstration.* — Montrons que (1)  $\Leftrightarrow$  (2) : en prenant  $M_0$  comme origine,  $G$  est défini par l'égalité

$$\overrightarrow{M_0 G} = \sum_{i=1}^p t_i \overrightarrow{M_0 M_i} \quad \text{i.e.} \quad G = M_0 + \sum_{i=1}^p t_i \overrightarrow{M_0 M_i}$$

et l'équivalence (1)  $\Leftrightarrow$  (2) en découle (si  $\lambda_1, \dots, \lambda_p$  sont donnés, on prend  $t_i = \lambda_i$  pour  $i = 1, \dots, p$  et  $t_0 = 1 - \sum_{i=1}^p \lambda_i$ ).

Il est clair que (2)  $\Rightarrow$  (3). Supposons (3) vérifié pour un point  $M_0$  fixé et soit  $M'_0 \in \mathcal{F}$ , notons provisoirement

$$F' = \{\overrightarrow{M'_0 M} \mid M \in \mathcal{F}\}.$$

Posons  $\vec{u} = \overrightarrow{M_0 M'_0}$ . Comme  $M'_0 \in \mathcal{F}$ , on a  $\vec{u} \in F$ , et comme  $F$  est un sous-espace vectoriel de  $E$ , l'application  $\vec{v} \mapsto \vec{v} - \vec{u}$  est une bijection de  $F$  sur lui-même (d'inverse  $\vec{w} \mapsto \vec{w} + \vec{u}$ ). D'après la relation de Chasles, pour tout  $M \in \mathcal{F}$  on a

$$\overrightarrow{M'_0 M} = \overrightarrow{M_0 M} - \overrightarrow{M_0 M'_0} = \overrightarrow{M_0 M} - \vec{u}$$

et donc

$$F' = \{\vec{v} - \vec{u} \mid \vec{v} \in F\} = F.$$

Ceci prouve que (3)  $\Rightarrow$  (4). Enfin, il est clair que (4)  $\Rightarrow$  (2). La proposition est démontrée. □

**Remarque 7.2.7.** — 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

$$\mathcal{F} \times \mathcal{F} \rightarrow F, \quad (M_0, M) \mapsto \overrightarrow{M_0M}$$

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  $\mathcal{F} \rightarrow F, M \mapsto \overrightarrow{M_0M}$  est bijective. Ceci prouve que  $\mathcal{F}$  est bien un espace affine de direction  $F$  (cf. définition 7.1.3).



**Définition 7.2.8 (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 7.2.9.** — (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 7.2.10 (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

$$F = \{\vec{u} \in E \mid f_i(\vec{u}) = 0, \quad \forall i = 1, \dots, p\} = \bigcap_{i=1}^p \text{Ker}(f_i).$$

*Démonstration.* — Supposons  $\mathcal{F} \neq \emptyset$  et soit  $M_0 \in \mathcal{F}$ . Alors un point arbitraire  $M \in \mathcal{E}$  appartient à  $\mathcal{F}$  si et seulement si on a :

$$\forall i = 1, \dots, p, \quad f_i(\overrightarrow{OM}) = c_i = f_i(\overrightarrow{OM_0})$$

ce qui équivaut à :

$$\forall i = 1, \dots, p, \quad f_i(\overrightarrow{M_0M}) = f_i(\overrightarrow{OM} - \overrightarrow{OM_0}) = 0$$

et encore à :  $\overrightarrow{M_0M} \in F = \bigcap_{i=1}^p \text{Ker}(f_i)$ . D'après la proposition 7.2.6, ceci montre que  $\mathcal{F}$ , s'il est non-vide, est un sea de direction  $F$ .  $\square$



**Exemples 7.2.11.** — 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  $x \mapsto a_1x_1 + \dots + a_nx_n$  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-vide, 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 7.2.12.** — 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

$$\mathcal{F} = \{M(x_1, \dots, x_n) \in \mathcal{E} \mid x_i = c_i, \quad \forall i = r+1, \dots, n\}.$$

*Démonstration.* — Soit  $M_0 \in \mathcal{F}$ , posons  $c_i = x_i(M_0)$  pour  $i = r+1, \dots, n$ . Pour tout  $M \in \mathcal{F}$ , on a

$$\overrightarrow{M_0M} \in F = \text{Vect}(e_1, \dots, e_r)$$

donc  $\overrightarrow{M_0M}$  a toutes ses coordonnées d'indice  $i > r$  nulles, et donc  $M = M_0 + \overrightarrow{M_0M}$  vérifie  $x_i(M) = x_i(M_0) = c_i$  pour  $i > r$ .

Réciproquement, si  $M \in \mathcal{E}$  vérifie  $x_i(M) = x_i(M_0) = c_i$  pour  $i > r$ , alors le vecteur  $\overrightarrow{M_0M}$  a toutes ses coordonnées d'indice  $i > r$  nulles, donc appartient à  $F$ , et donc  $M = M_0 + \overrightarrow{M_0M}$  appartient à  $\mathcal{F}$ . Ceci prouve la proposition.  $\square$

**Définition et proposition 7.2.13 (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

$$G = t_0A_0 + \dots + t_pA_p, \quad \text{où } A_0, \dots, A_p \in X, \quad t_0, \dots, t_p \in \mathbb{R}, \quad t_0 + \dots + t_p = 1$$

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 :

$$\begin{aligned} \text{Aff}\langle X \rangle &= \{A_0 + \lambda_1 \overrightarrow{A_0A_1} + \dots + \lambda_p \overrightarrow{A_0A_p} \mid A_1, \dots, A_p \in X, \lambda_1, \dots, \lambda_p \in \mathbb{R}\} \\ &= A_0 + \text{Vect}\{\overrightarrow{A_0B} \mid B \in X\} \end{aligned}$$

*Démonstration.* — On définit  $\mathcal{F}$  par :

$$\mathcal{F} = \{G = t_0A_0 + \dots + t_pA_p \mid A_0, \dots, A_p \in X, \quad t_0, \dots, t_p \in \mathbb{R}, \quad t_0 + \dots + t_p = 1\},$$

cette définition ne dépend d'aucun choix et tous les points de  $X$  y jouent le même rôle. On pourrait montrer directement, avec cette définition, que  $\mathcal{F}$  vérifie la propriété (1) de la proposition 7.2.6, mais il est plus commode de procéder comme suit. Fixons un point  $A_0 \in X$ , alors, comme dans la démonstration de (1)  $\Leftrightarrow$  (2) dans 7.2.6, on voit que

$$\begin{aligned} \text{Aff}\langle X \rangle &= \{A_0 + \lambda_1 \overrightarrow{A_0A_1} + \dots + \lambda_p \overrightarrow{A_0A_p} \mid A_1, \dots, A_p \in X, \lambda_1, \dots, \lambda_p \in \mathbb{R}\} \\ &= A_0 + \text{Vect}\{\overrightarrow{A_0B} \mid B \in X\}. \end{aligned}$$

Ceci montre que  $\mathcal{F}$  est un sea de direction  $F = \text{Vect}\{\overrightarrow{A_0B} \mid B \in X\}$  (qui ne dépend pas du choix de  $A_0$ , cf. 7.2.6). De plus,  $\mathcal{F}$  contient  $X$ , et si  $\mathcal{F}'$  est un sea contenant  $X$ , alors sa direction  $F'$  contient tous les vecteurs  $\overrightarrow{A_0B}$ , pour  $B \in \mathcal{F}$ , donc  $F'$  contient  $F$ , et  $\mathcal{F}' = A_0 + F'$  contient  $\mathcal{F}$ . Ceci montre que  $\mathcal{F}$  est le **plus petit sea de  $\mathcal{E}$  contenant  $X$** , ce qui justifie la notation  $\text{Aff}\langle X \rangle$ .  $\square$

**Corollaire 7.2.14 (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

$$\begin{aligned} \text{Aff}\langle A_0, \dots, A_p \rangle &= \{G = t_0A_0 + \dots + t_pA_p \mid t_0, \dots, t_p \in \mathbb{R}, \quad t_0 + \dots + t_p = 1\} \\ &= A_0 + \text{Vect}\{\overrightarrow{A_0A_1}, \dots, \overrightarrow{A_0A_p}\} \end{aligned}$$

sa dimension est celle de l'espace vectoriel  $F = \text{Vect}\{\overrightarrow{A_0A_1}, \dots, \overrightarrow{A_0A_p}\}$ ; en particulier,  $\dim \text{Aff}\langle A_0, \dots, A_p \rangle = p \Leftrightarrow$  les vecteurs  $\overrightarrow{A_0A_1}, \dots, \overrightarrow{A_0A_p}$  sont linéairement indépendants.

**Remarque 7.2.15.** — 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  $F = \text{Vect}\{\overrightarrow{A_iA_j} \mid j \neq i\}$  et  $\text{Aff}\langle A_0, \dots, A_p \rangle = A_i + F$ .



**Définition 7.2.16 (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  $\overrightarrow{A_0A_1}, \dots, \overrightarrow{A_0A_p}$  sont linéairement indépendants (ceci équivaut à dire que, pour  $i$  fixé, les  $p$  vecteurs  $\overrightarrow{A_iA_j}$ , 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  $\iff$  ils sont alignés. Donc  $A_0, A_1, A_2$  sont affinement indépendants  $\iff$  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  $\iff$  ils sont **non coplanaires**.



**Définition 7.2.17 (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  $\overrightarrow{A_0A_1}, \dots, \overrightarrow{A_0A_n}$  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  $\mathcal{B} = (\overrightarrow{A_0A_1}, \dots, \overrightarrow{A_0A_n})$  est ordonnée, i.e. les coordonnées  $(x_1, \dots, x_n)$  dans ce repère correspondent au point

$$A_0 + x_1 \overrightarrow{A_0A_1} + \dots + x_n \overrightarrow{A_0A_n} = \left(1 - \sum_{i=1}^n x_i\right) A_0 + x_1 A_1 + \dots + x_n A_n.$$



**Définitions 7.2.18 (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 7.2.18.1.** — 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 7.2.19 (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  $V = F + F' + \mathbb{R}PP'$  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 :

$$\dim(\mathcal{F} + \mathcal{F}') = \begin{cases} \dim(F + F') & \text{si } \mathcal{F} \cap \mathcal{F}' \neq \emptyset, \\ \dim(F + F') + 1 & \text{si } \mathcal{F} \cap \mathcal{F}' = \emptyset. \end{cases}$$

**Démonstration.** — (1) Si  $Q \in \mathcal{F}$  et  $Q' \in \mathcal{F}'$ , alors  $\overrightarrow{QQ'} = \overbrace{\overrightarrow{QP}}^{\in F} + \overrightarrow{PP'} + \overbrace{\overrightarrow{P'Q'}}^{\in F'}$  et donc  $F + F' + \overrightarrow{QQ'} = F + F' + \overrightarrow{PP'}$ .

Prouvons (2). L'implication  $\Leftarrow$  est évidente, car si  $Q \in \mathcal{F} \cap \mathcal{F}'$ , on peut prendre  $P = Q = P'$ , d'où  $V = F + F'$ . Réciproquement, supposons que  $\overrightarrow{PP'} = \vec{u} + \vec{u}'$  avec  $\vec{u} \in F$  et  $\vec{u}' \in F'$ . Alors le point  $Q = P + \vec{u}$  appartient à  $\mathcal{F}$ , et l'on a aussi

$$\overrightarrow{P'Q} = \overrightarrow{PQ} - \overrightarrow{PP'} = \vec{u} - (\vec{u} + \vec{u}') = -\vec{u}' \in F'$$

<sup>(1)</sup>Attention, cette notation n'est peut-être pas standard, peut-être les puristes s'en tiennent-ils à la notation  $\text{Aff}(\mathcal{F} \cup \mathcal{F}')$ ...

donc  $Q \in \mathcal{F}' \cap \mathcal{F}$ . Ceci prouve (2).

Prouvons (3). Par définition,  $\text{Aff}\langle \mathcal{F} \cup \mathcal{F}' \rangle = P + W$ , où  $W$  est le sous-espace vectoriel engendré par les vecteurs  $\overrightarrow{PM}$  et  $\overrightarrow{PN} = \overrightarrow{PP'} + \overrightarrow{P'N}$ , où  $M$  (resp.  $N$ ) parcourt  $\mathcal{F}$  (resp.  $\mathcal{F}'$ ). Donc  $W$  est le sev engendré par  $F$ ,  $F'$  et  $\overrightarrow{PP'}$ , i.e.  $W = V$ . Ceci prouve (3). Enfin, (4) découle de (2), car  $\dim V \leq \dim(F + F') + 1$ , avec égalité si et seulement si  $\overrightarrow{PP'} \notin F + F'$ .  $\square$

**Remarque 7.2.20.** — 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 7.2.21 (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  $F \cap F'$ .

*Démonstration.* — Supposons que  $A \in \mathcal{F} \cap \mathcal{F}'$ . Alors un point arbitraire  $P \in \mathcal{E}$  appartient à  $\mathcal{F} \cap \mathcal{F}'$  si et seulement si  $\overrightarrow{AP}$  appartient à  $F$  et à  $F'$ , i.e. à  $F \cap F'$ . D'après la proposition 7.2.6, ceci montre que  $\mathcal{F} \cap \mathcal{F}'$  est un sea de direction  $F \cap F'$ .  $\square$

**Proposition 7.2.22 (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  $\phi^{-1}(F')$ ; plus précisément, si  $P \in \mathcal{F}$  est tel que  $f(P) \in \mathcal{F}'$  alors

$$\begin{aligned} f^{-1}(\mathcal{F}') &= \{Q = P + \vec{u} \in \mathcal{E} \mid f(Q) = f(P) + \phi(\vec{u}) \in \mathcal{F}'\} \\ &= \{Q = P + \vec{u} \in \mathcal{E} \mid \overrightarrow{f(P)f(Q)} = \phi(\vec{u}) \in F'\} = P + \phi^{-1}(F'). \end{aligned}$$

*Démonstration.* — Laissée au lecteur comme exercice.  $\square$

### 7.3. Projections, symétries, points fixes

**Proposition 7.3.1.** — 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  $F \cap F'$ .

*Démonstration.* — Ceci résulte des propositions 7.2.19 et 7.2.21. En effet, soient  $P \in \mathcal{F}$  et  $P' \in \mathcal{F}'$ . Comme  $F + F' = E$ , on a nécessairement  $\overrightarrow{PP'} \in F + F'$  donc, d'après la proposition 7.2.19, on a  $\mathcal{F} \cap \mathcal{F}' \neq \emptyset$ . Donc, d'après la proposition 7.2.21,  $\mathcal{F} \cap \mathcal{F}'$  est un sea de direction  $F \cap F'$ .  $\square$

**Corollaire 7.3.2 (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 7.3.3.** — 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 7.3.4 (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 7.1.15). 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.

 **Proposition 7.3.5 (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  $I$ .

*Démonstration.* — (1) Si  $\text{Fix}(f) \neq \emptyset$ , soit  $I \in \text{Fix}(f)$ . Alors, pour un point arbitraire  $P \in \mathcal{E}$  on a :

$$P \in \text{Fix}(f) \iff f(P) = P \iff \overrightarrow{IP} = \overrightarrow{If(P)} = \overrightarrow{f(I)f(P)} = \vec{f}(\overrightarrow{IP}) \iff \overrightarrow{IP} \in \text{Fix}(\vec{f}).$$

Ceci montre qu'alors  $\text{Fix}(f)$  est un sea de direction  $\text{Fix}(\vec{f}) = \text{Ker}(\vec{f} - \text{id}_E)$ .

(2) Supposons  $\text{Ker}(\vec{f} - \text{id}_E) = \{0\}$  et fixons une origine  $O \in \mathcal{E}$  arbitraire. Pour tout point  $P \in \mathcal{E}$ , on a


$$\overrightarrow{Of(P)} = \overrightarrow{Of(O)} + \overrightarrow{f(O)f(P)} = \overrightarrow{Of(O)} + \vec{f}(\overrightarrow{OP})$$

et donc on a :

$$P = f(P) \iff \overrightarrow{OP} = \overrightarrow{Of(P)} \iff (\vec{f} - \text{id}_E)(\overrightarrow{OP}) = -\overrightarrow{Of(O)}.$$


Or, d'après hypothèse,  $\vec{f} - \text{id}_E$  est injective, donc **bijective** puisque  $E$  est de dimension finie. Donc il existe un unique vecteur  $\vec{u} \in E$  tel que  $(\vec{f} - \text{id}_E)(\vec{u}) = -\overrightarrow{Of(O)}$ , et il existe donc un unique point  $I \in \mathcal{E}$  tel que  $(\vec{f} - \text{id}_E)(\overrightarrow{OI}) = -\overrightarrow{Of(O)}$ , i.e.  $I$  est l'unique point fixe de  $f$ . La proposition est démontrée.  $\square$

## 7.4. Espaces affines euclidiens

 **Définition 7.4.1.** — 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  $d(x, y) = 0 \iff x = y$ , (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 7.4.2 (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 7.4.3.** — 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 7.4.4 (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 **bijectif** (car injective et  $E$  de dimension finie), donc  $f$  l'est aussi, d'après le point (3) de 7.1.15. 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 7.4.5.** — 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 7.4.6 (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 7.4.7.** — 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 7.4.4, 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 7.4.3. 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 7.4.4. 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

$$\forall P \in \mathcal{E}, \quad \psi(\overrightarrow{OP}) = \overrightarrow{O\phi(P)} \quad \text{i.e.} \quad \forall \vec{u} \in E, \quad O + \psi(\vec{u}) = \phi(O + \vec{u})$$

est **linéaire**.

Soit  $\vec{u} \in E$ , posons  $P = O + \vec{u}$  et  $P' = \phi(P)$ , alors  $\psi(\vec{u}) = \overrightarrow{OP'}$  et donc :

$$\|\psi(\vec{u})\| = \|\overrightarrow{OP'}\| = d(O, P') \underbrace{=}_{(*)} d(O, P) = \|\overrightarrow{OP}\| = \|\vec{u}\|$$

(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 :

$$\begin{aligned} \text{d'une part,} \quad & \overrightarrow{P'Q'} = \overrightarrow{OQ'} - \overrightarrow{OP'} = \vec{v} - \vec{u} \\ \text{d'autre part,} \quad & \overrightarrow{P'Q'} = \overrightarrow{OQ'} - \overrightarrow{OP'} = \psi(\vec{v}) - \psi(\vec{u}) \end{aligned}$$

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 :

$$\|\psi(\vec{v})\|^2 + \|\psi(\vec{u})\|^2 - 2(\psi(\vec{v}) | \psi(\vec{u})) = \|\vec{v}\|^2 + \|\vec{u}\|^2 - 2(\vec{v} | \vec{u})$$

et comme  $\|\psi(\vec{v})\| = \|\vec{v}\|$  et  $\|\psi(\vec{u})\| = \|\vec{u}\|$ , ceci donne :

$$(\psi(\vec{v}) | \psi(\vec{u})) = (\vec{v} | \vec{u})$$

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  $\psi(\mathcal{B}) = (\phi(e_1), \dots, \phi(e_n))$  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 7.4.5.  $\square$

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 7.4.5, 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 7.4.8.** — 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 7.4.9 (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 un 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

$$\{w \in G | \vec{h}(w) = w\} = G \cap \text{Ker}(\vec{f} - \text{id}_E) = G \cap F = \{0\}$$

(la dernière égalité puisque  $G = F^\perp$ ). Donc  $\text{Fix}(\vec{h}) = \{0\}$  et donc, d'après la proposition 7.3.5,  $h$  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  $\overrightarrow{I_0 Q} = \overrightarrow{g(I_0)g(Q)} = \vec{f}(\overrightarrow{I_0 Q})$ , donc  $\overrightarrow{I_0 Q} \in \text{Fix}(\vec{f}) = F$  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. 7.3.2). 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 7.4.10 (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  $E$ .



**7.4.11. 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  $\vec{f}$ , cf. le point (2) du théorème 7.1.15 et la discussion précédant 7.3.5.

Soit d'abord  $f \in \text{Is}^+(\mathcal{P})$ , alors il existe un unique  $\theta \in [0, 2\pi[$  tel que  $\vec{f}$  soit la rotation vectorielle d'angle  $\theta$ , i.e.

$$\text{Mat}_{\mathcal{B}}(\vec{f}) = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}.$$

Si  $\vec{f} = \text{id}_E$ , 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  $\text{Fix}(\vec{f}) = \{0\}$  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  $\vec{f}$  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



**7.4.12. 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 7.4.9 et la classification des isométries vectorielles de  $E$  (cf. 6.4.18), on va déterminer toutes les isométries affines  $f$  de  $\mathcal{E}$ .

**1er cas.** Soit d'abord  $f \in \text{Is}^+(\mathcal{E})$ . Si  $\vec{f} = \text{id}_E$ , alors  $f$  est une **translation**  $t_{\vec{u}}$ , avec  $\vec{u} \in E$ . Sinon,  $\vec{f}$  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  $\vec{f}$  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 7.4.9, 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 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 7.4.9, 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

## 7.5. Coniques

On se place dans le plan affine euclidien  $\mathcal{P}$ . Si  $M, N \in \mathcal{P}$ , on notera  $MN$  la longueur du segment  $[M, N]$ , i.e. la norme euclidienne du vecteur  $\overrightarrow{MN}$ .



**Lemme 7.5.1 (Distance d'un point à une droite).** — Soient  $\mathcal{D}$  une droite affine,  $M \in \mathcal{P}$  et  $H$  la projection orthogonale de  $M$  sur  $\mathcal{D}$ . Pour tout  $P \in \mathcal{D}$ , on a  $MH \leq MP$ , avec égalité si et seulement si  $P = H$ . Donc  $MH$  est la distance minimale de  $M$  à un point de  $\mathcal{D}$ , et l'on pose  $MH = d(M, \mathcal{D})$ .

*Démonstration.* — Soit  $P \in \mathcal{P}$ , on a  $\overrightarrow{MP} = \overrightarrow{MH} + \overrightarrow{HP}$  et les vecteurs  $\overrightarrow{MH}$  et  $\overrightarrow{HP}$  sont orthogonaux, donc d'après le théorème de Pythagore on a :

$$MP^2 = \|\overrightarrow{MP}\|^2 = \|\overrightarrow{MH}\|^2 + \|\overrightarrow{HP}\|^2 = MH^2 + HP^2 \geq MH^2,$$

avec égalité si et seulement si  $P = H$ . □

### Définition 7.5.2 (Définition des coniques par foyer et directrice)

Soient  $\mathcal{D}$  une droite affine,  $F$  un point de  $\mathcal{P}$  n'appartenant pas à  $\mathcal{D}$ , et  $e$  un réel strictement positif. La **conique**  $\mathcal{C}$  de **directrice**  $\mathcal{D}$ , de **foyer**  $F$  et d'**excentricité**  $e$  est :

$$\mathcal{C} = \{M \in \mathcal{P} \mid FM = e \cdot d(M, \mathcal{D})\}.$$

Notons  $K$  la projection orthogonale de  $F$  sur  $\mathcal{D}$  et  $h = KF$ . Soit  $\Delta$  la droite  $(KF)$ , i.e. la perpendiculaire à  $\mathcal{D}$  passant par  $F$ . Soit  $(\vec{i}, \vec{j})$  une base orthonormée de  $\mathbb{R}^2$ , telle que  $\overrightarrow{FK} = -h\vec{i}$ , i.e. dans le repère  $\mathcal{R} = (F, \vec{i}, \vec{j})$ ,  $K$  a pour coordonnées  $(-h, 0)$ . Enfin, on note  $\sigma_\Delta$  la symétrie orthogonale par rapport à  $\Delta$ .

Soit  $M(x, y)$  un point arbitraire de  $\mathcal{P}$ , de coordonnées  $(x, y)$  dans  $\mathcal{R}$ ; sa projection orthogonale sur  $\mathcal{D}$  est le point  $H(-h, y)$ , donc  $d(M, \mathcal{D}) = |x + h|$  et  $M$  appartient à  $\mathcal{C}$  si et seulement si

$$x^2 + y^2 = e^2(x + h)^2 \quad \text{i.e.} \quad y^2 + (1 - e^2)x^2 - 2e^2hx - e^2h^2 = 0.$$

Notons que l'équation ci-dessus est inchangée si l'on change  $y$  en  $-y$ , par conséquent

$$M(x, y) \in \mathcal{C} \iff M(x, -y) \in \mathcal{C}$$

donc  $\sigma_\Delta(\mathcal{C}) = \mathcal{C}$ , i.e.  $\Delta$  est un **axe de symétrie** de  $\mathcal{C}$ .

On pose  $p = eh > 0$  et on l'appelle le **paramètre** de la conique  $\mathcal{C}$ . Tous les points  $M$  de la droite d'équation  $x = 0$  (la parallèle à  $\mathcal{D}$  passant par  $F$ ) vérifient  $d(M, \mathcal{D}) = h$ , donc les points d'intersection de cette droite avec  $\mathcal{C}$  sont les deux points de coordonnées  $(0, p)$  et  $(0, -p)$ . D'autre part, les points d'intersection de  $\mathcal{C}$  avec la droite  $\Delta$  (d'équation  $y = 0$ ) sont les points  $(\lambda, 0)$ , où  $\lambda$  est solution de l'équation :

$$(\dagger) \quad (1 - e^2)\lambda^2 - 2ep\lambda - p^2 = 0.$$

**7.5.3. Paraboles.** — Supposons d'abord  $e = 1$ . Dans ce cas, on trouve  $\lambda = -p/2$  et l'équation de  $\mathcal{C}$  est

$$y^2 = 2px + p^2 = 2p\left(x + \frac{p}{2}\right)$$

Soit alors  $O$  le point de coordonnées  $(-p/2, 0)$  dans le repère  $(F, \vec{i}, \vec{j})$ , c'est le milieu du segment  $[KF]$ . Les nouvelles coordonnées dans le repère  $\mathcal{R}_0 = (O, \vec{i}, \vec{j})$  sont  $X = x + (p/2)$  et  $Y = y$ , donc l'équation de  $\mathcal{C}$  devient :

$$\mathcal{C} = \{M(X, Y) \in \mathcal{P} \mid Y^2 = 2pX\}$$

et l'on dit que  $\mathcal{C}$  est une **parabole** de paramètre  $p$ . Dans  $\mathcal{R}_0$ , la directrice  $\mathcal{D}$  a pour équation  $X = -p/2$  et le foyer  $F$  a pour coordonnées  $(p/2, 0)$ .

Supposons maintenant  $e \neq 1$ . Alors le discriminant réduit du trinôme  $(\dagger)$  est  $e^2p^2 + (1 - e^2)p^2 = p^2$  donc les solutions sont :

$$\lambda_1 = \frac{ep - p}{1 - e^2} = \frac{-p}{1 + e}, \quad \lambda_2 = \frac{ep + p}{1 - e^2} = \frac{p}{1 - e} \quad \text{et l'on a} \quad \left| \frac{\lambda_2 - \lambda_1}{2} \right| = \frac{p}{|1 - e^2|}.$$

Notons  $A_1, A_2$  les points correspondants de  $\Delta \cap \mathcal{C}$ , et  $O$  leur milieu. L'abscisse  $x_O$  de  $O$  dans le repère  $(F, \vec{i}, \vec{j})$  est  $x_O = \frac{\lambda_1 + \lambda_2}{2} = \frac{ep}{1 - e^2}$  et le trinôme s'écrit :

$$(1 - e^2)(x - x_O)^2 - \frac{e^2p^2}{1 - e^2} - p^2 = (1 - e^2)(x - x_O)^2 - \frac{p^2}{1 - e^2}.$$

Notons  $X = x - x_O$  et  $Y = y$  les coordonnées dans le repère  $\mathcal{R}_0 = (O, \vec{i}, \vec{j})$ , alors l'équation de  $\mathcal{C}$  devient

$$\mathcal{C} = \{M(X, Y) \mid (1 - e^2)X^2 + Y^2 = \frac{p^2}{1 - e^2}\}.$$

L'équation est inchangée si l'on change  $X$  en  $-X$  et l'on voit donc que la droite  $D_0$  d'équation  $X = 0$  est un second **axe de symétrie** de  $\mathcal{C}$  (et donc le point  $O$  est un **centre de symétrie** de  $\mathcal{C}$ ). Par conséquent,

$\mathcal{C}$  possède un second couple (foyer, directrice)  $(F', \mathcal{D}')$ , symétrique de  $(F, \mathcal{D})$  par rapport à la droite  $D_0$ . On dit alors que la droite  $\Delta$ , qui contient les foyers, est l'**axe focal**.

D'autre part, les points  $A_1, A_2 \in \Delta$  s'appellent les **sommets** de la conique; dans  $\mathcal{R}_0$  ils ont pour abscisse :

$$\pm a, \quad \text{où} \quad a = \left| \frac{\lambda_2 - \lambda_1}{2} \right| = \frac{p}{|1 - e^2|}.$$

Posons également  $b = \frac{p}{\sqrt{|1 - e^2|}}$ , d'où  $b^2 = \frac{p^2}{|1 - e^2|} = \varepsilon \frac{p^2}{1 - e^2}$ , où  $\varepsilon = \text{signe de } 1 - e^2$ . Alors l'équation de  $\mathcal{C}$  se récrit :

$$\mathcal{C} = \{M(X, Y) \mid \frac{X^2}{a^2} + \varepsilon \frac{Y^2}{b^2} = 1\}.$$

Distinguons maintenant les cas  $0 < e < 1$  et  $1 < e$ .



**7.5.4. Ellipses.** — Supposons maintenant  $0 < e < 1$ . Dans ce cas,  $\varepsilon = 1$ , l'équation de  $\mathcal{C}$  est :

$$(*) \quad \mathcal{C} = \{M(X, Y) \mid \frac{X^2}{a^2} + \frac{Y^2}{b^2} = 1\}$$

et  $\mathcal{C}$  est une **ellipse**. Dans le repère  $\mathcal{R} = (F, \vec{i}, \vec{j})$ ,  $K = \Delta \cap \mathcal{D}$  est d'abscisse  $-h$ , le sommet  $A = A_1$  est d'abscisse  $-p/(1 + e) = -h/(e^{-1} + 1)$  donc est compris entre  $K$  et  $F$ , de plus  $AF = e \cdot AK < AK$ , donc  $A$  est plus près de  $F$  que de  $K$ . Le centre de symétrie  $O$  est d'abscisse  $ep/(1 - e^2) > 0$ , et enfin  $F', A' = A_2$  et  $K'$  sont symétriques de  $F, A, K$  par rapport à  $O$ . Ces points sont donc positionnés comme suit :

points	$K$	$A$	$F$	$O$	$F'$	$A'$	$K'$
abscisses dans $\mathcal{R}$	$-h$	$\frac{-h}{1 + e^{-1}}$	$0$	$\frac{ep}{1 - e^2}$	$\dots$	$\dots$	$\dots$
abscisses dans $\mathcal{R}_0$	$-h - c$	$-a$	$-c$	$0$	$c$	$a$	$h + c$

où  $a = \frac{p}{1 - e^2}$  et  $c = \frac{ep}{1 - e^2} = ea$  d'où  $e = c/a$ . Comme  $b^2 = \frac{p^2}{1 - e^2}$ , on a  $c^2 = \frac{e^2 p^2}{(1 - e^2)^2} = a^2 - b^2$ , d'une part, et  $p = b^2/a$ , d'autre part. Enfin,  $p = eh$  donne  $h = p/e = b^2/c$ , d'où :

$$c = \sqrt{a^2 - b^2}, \quad e = \frac{c}{a}, \quad h = \frac{b^2}{c}, \quad p = \frac{b^2}{a}, \quad KO = h + c = \frac{b^2 + c^2}{c} = \frac{a^2}{c}.$$

(Le lecteur est invité à faire la construction pour  $a = 5, b = 4$  (et/ou  $a = 5, b = 3$ ) et à dessiner approximativement les ellipses correspondantes.)

Dans l'équation (\*) de l'ellipse, on peut faire  $b = a$ , on obtient alors le **cercle** de centre  $F = O$  et de rayon  $a$ . Remarquons que  $b^2 = (1 - e^2)a^2$  et  $c = \sqrt{a^2 - b^2} = ea$ , d'où

$$h = \frac{b^2}{c} = \frac{1 - e^2}{e} a.$$

Donc, en gardant  $F$  et  $a$  fixés, on obtient le cercle précédent en faisant tendre  $e$  vers 0, et dans ce cas la directrice  $\mathcal{D}$ , qui est la droite d'équation  $x = -\frac{1 - e^2}{e} a$  dans le repère  $\mathcal{R}$ , « tend vers l'infini ». On verra plus bas que si on prend la définition « bifocale » de l'ellipse, alors le cas du cercle apparaît de façon naturelle, comme le cas où les deux foyers sont égaux.



**7.5.5. Hyperboles.** — Supposons enfin  $e > 1$ . Dans ce cas,  $\varepsilon = -1$ , l'équation de  $\mathcal{C}$  est :

$$\mathcal{C} = \{M(X, Y) \mid \frac{X^2}{a^2} - \frac{Y^2}{b^2} = 1\}$$

et  $\mathcal{C}$  est une **hyperbole**. Elle admet pour asymptotes les droites  $Y = \frac{b}{a}X$  et  $Y = -\frac{b}{a}X$ .

Dans le repère  $\mathcal{R} = (F, \vec{i}, \vec{j})$ ,  $K = \Delta \cap \mathcal{D}$  est d'abscisse  $-h$ , le sommet  $A = A_1$  est d'abscisse  $-p/(1 + e) = -h/(e^{-1} + 1)$  donc est compris entre  $K$  et  $F$ , de plus  $AF = e \cdot AK > AK$ , donc ici  $A$  est plus près de  $K$  que de  $F$ .

Le centre de symétrie  $O$  est d'abscisse  $\frac{ep}{1-e^2} = \frac{e^2h}{1-e^2} = -h - \frac{h}{e^2-1} < -h$ , et enfin  $F', A' = A_2$  et  $K'$  sont symétriques de  $F, A, K$  par rapport à  $O$ . Ces points sont donc positionnés comme suit :

points	$F'$	$A'$	$K'$	$O$	$K$	$A$	$F$
abscisses dans $\mathcal{R}$	...	...	...	$\frac{ep}{1-e^2}$	$-h$	$\frac{-h}{1+e^{-1}}$	$0$
abscisses dans $\mathcal{R}_0$	$-c$	$-a$	$-(c-h)$	$0$	$c-h$	$a$	$c$

où  $a = \frac{p}{e^2-1}$  et  $c = \frac{ep}{e^2-1} = ea$  d'où  $e = c/a$ . Comme  $b^2 = \frac{p^2}{e^2-1}$ , on a  $c^2 = \frac{e^2p^2}{(e^2-1)^2} = a^2 + b^2$ , d'une part, et  $p = b^2/a$ , d'autre part. Enfin,  $p = eh$  donne  $h = p/e = b^2/c$ , d'où :

$$c = \sqrt{a^2 + b^2}, \quad e = \frac{c}{a}, \quad h = \frac{b^2}{c}, \quad p = \frac{b^2}{a}, \quad OK = c - h = \frac{c^2 - b^2}{c} = \frac{a^2}{c}.$$

(Le lecteur est invité à construire ces points, ainsi que les asymptotes, pour  $a = 5, b = 4$  (et/ou  $a = 5, b = 3$ ) et à dessiner approximativement les hyperboles correspondantes.)



**Proposition 7.5.6 (Définition bifocale de l'ellipse).** — Soient  $F, F'$  deux points de  $\mathcal{P}$  et  $O$  leur milieu. Soit  $\mathcal{R} = (O, \vec{i}, \vec{j})$  un repère orthonormé de  $\mathcal{P}$  tel que  $\overrightarrow{OF'} = c\vec{i}$  avec  $c = \frac{1}{2}FF' \geq 0$  (et donc  $\overrightarrow{OF} = -c\vec{i}$ ). Notons  $(x, y)$  les coordonnées dans ce repère. Soient  $a, b \in \mathbb{R}_+^*$  tels que  $a > c$  et  $b^2 = a^2 - c^2$ . Alors l'ensemble

$$\mathcal{C} = \{M(x, y) \in \mathcal{P} \mid FM + F'M = 2a\}$$

est l'ellipse d'équation  $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$  si  $F \neq F'$ , et c'est le cercle de centre  $F$  et de rayon  $a$  si  $F' = F$ .

*Démonstration.* — Comme  $2a > 0$ , l'égalité  $FM + F'M = 2a$  équivaut (en l'élevant au carré) à :

$$(0) \quad 2(x^2 + y^2 + c^2) + 2\sqrt{((x-c)^2 + y^2)((x+c)^2 + y^2)} = 4a^2$$

qui équivaut à :

$$x^2 + y^2 + c^2 - 2a^2 = -\sqrt{((x-c)^2 + y^2)((x+c)^2 + y^2)}.$$

Ceci équivaut à :

$$(1) \quad x^2 + y^2 \leq 2a^2 - c^2 = a^2 + b^2 \quad \text{et}$$

$$(2) \quad \underbrace{(x^2 + y^2 + c^2 - 2a^2)^2}_{\beta} = ((x-c)^2 + y^2)((x+c)^2 + y^2) = \underbrace{(x^2 + c^2 + y^2)^2}_{\alpha} - 4c^2x^2,$$

puis (2) équivaut à :

$$4c^2x^2 = \alpha^2 - \beta^2 = 2a^2 \cdot 2(x^2 + y^2 + c^2 - a^2)$$

qui équivaut à

$$(3) \quad (a^2 - c^2)x^2 + a^2y^2 = a^2(a^2 - c^2) \quad \text{soit} \quad \frac{x^2}{a^2} + \frac{y^2}{b^2} = 1.$$

Donc (0) équivaut à :

$$(\dagger) \quad x^2 + y^2 \leq a^2 + b^2 \quad \text{et} \quad \frac{x^2}{a^2} + \frac{y^2}{b^2} = 1.$$

Mais  $\frac{x^2}{a^2} + \frac{y^2}{b^2} \geq \frac{x^2}{a^2 + b^2} + \frac{y^2}{a^2 + b^2}$ , donc si (3) est satisfaite, on a automatiquement  $x^2 + y^2 \leq a^2 + b^2$ . Par conséquent, (0) et ( $\dagger$ ) sont équivalents à la seule condition :

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1.$$

Ceci prouve la proposition. □

**Proposition 7.5.7 (Définition bifocale de l'hyperbole).** — Soient  $(x, y)$  les coordonnées dans un repère orthonormé  $\mathcal{R} = (O, \vec{i}, \vec{j})$  de  $\mathcal{P}$ , soient  $a, b \in \mathbb{R}_+^*$  et soit  $\mathcal{C}$  l'hyperbole d'équation

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1.$$

Ses foyers  $F$  et  $F'$  ont pour coordonnées  $(-c, 0)$  et  $(c, 0)$ , où  $c = \sqrt{a^2 + b^2}$ . Alors

$$\mathcal{C} = \{M(x, y) \in \mathcal{P} \mid |FM - F'M| = 2a\}$$



et, plus précisément, la branche de l'hyperbole la plus proche de  $F$  (resp.  $F'$ ) est formée des points  $M$  tels que  $F'M - FM = 2a$  (resp.  $= -2a$ ).

La démonstration est laissée comme exercice pour le lecteur.

Soit maintenant  $\mathcal{R}_0 = (O, \vec{u}, \vec{v})$  un repère **orthonormé** de  $\mathcal{P}$ . Notons  $(x, y)$  les coordonnées dans ce repère, et étudions la nature géométrique de l'ensemble

$$\mathcal{C} = \{M(x, y) \mid ax^2 + 2bxy + cy^2 + 2dx + 2ey + f = 0\}$$

où  $a, b, c, d, e, f$  sont des réels fixés, avec  $(a, b, c) \neq (0, 0, 0)$ . En d'autres termes, soit  $\phi$  la forme bilinéaire symétrique dont la matrice dans la base  $\mathcal{B} = (\vec{u}, \vec{v})$  est :

$$S = \text{Mat}_{\mathcal{B}}(\phi) = \begin{pmatrix} a & b \\ b & c \end{pmatrix} \quad \text{i.e.} \quad \phi\left(\begin{pmatrix} x \\ y \end{pmatrix}, \begin{pmatrix} x' \\ y' \end{pmatrix}\right) = axx' + b(xy' + yx') + cyy'$$

et soit  $Q$  la forme quadratique associée, i.e.  $Q(x, y) = ax^2 + 2bxy + cy^2$ . Soit  $L$  la forme linéaire sur  $\mathbb{R}^2$  définie par :

$$L(\vec{u}) = 2d, \quad L(\vec{v}) = 2e, \quad \text{d'où} \quad L(x\vec{u} + y\vec{v}) = 2dx + 2ey.$$

Enfin, soit  $h : \mathcal{P} \rightarrow \mathbb{R}$  la fonction définie par  $h(M) = Q(x, y) + L(x, y) + f$  si  $\overrightarrow{OM} = x\vec{u} + y\vec{v}$ . On s'intéresse donc à l'ensemble

$$\mathcal{C} = \{M(x, y) \mid h(x, y) = 0\}.$$

Par hypothèse,  $(a, b, c) \neq (0, 0, 0)$  i.e. la forme quadratique  $Q$  est non nulle, i.e. la matrice  $S$  est non nulle. Donc  $S$  est de rang 1 ou 2, selon que son déterminant  $\delta = ac - b^2$  est nul ou non. On a le :



**Théorème 7.5.8.** — Soit  $\mathcal{R}_0 = (O, \vec{u}, \vec{v})$  un repère orthonormé de  $\mathcal{P}$ , soient  $(x, y)$  les coordonnées correspondantes, soient  $a, b, c, d, e, f \in \mathbb{R}$ , avec  $(a, b, c) \neq (0, 0, 0)$ , et soit

$$\mathcal{C} = \{M(x, y) \mid ax^2 + 2bxy + cy^2 + 2dx + 2ey + f = 0\}.$$

On suppose  $\mathcal{C} \neq \emptyset$ . Alors :

- (1) Si  $\delta = 0$ ,  $\mathcal{C}$  est une parabole ou bien la réunion de deux droites parallèles ou « confondues » (i.e. égales).
- (2) Si  $\delta = ac - b^2 > 0$ ,  $\mathcal{C}$  est une ellipse ou un point.
- (3) Si  $\delta < 0$ ,  $\mathcal{C}$  est une hyperbole ou bien la réunion de deux droites sécantes.

Afin de démontrer ceci, rappelons d'abord les points suivants. Soit  $E$  muni de  $(\mid)$  un espace euclidien,  $Q$  une forme quadratique arbitraire sur  $E$ ,  $\phi$  sa forme polaire,  $\mathcal{B}_0 = (e_1, \dots, e_n)$  une base orthonormée de  $E$ ,  $A = (a_{ij})_{i,j=1}^n = \text{Mat}_{\mathcal{B}_0}(\phi)$ , et  $u$  l'endomorphisme de  $E$  tel que  $\text{Mat}_{\mathcal{B}_0}(u) = A$ , i.e.  $u(e_j) = \sum_{k=1}^n a_{kj}e_k$  pour tout  $j$ . Alors, pour tout  $i, j$  on a :

$$(e_i \mid u(e_j)) = a_{ij} = \phi(e_i, e_j) = a_{ji} = (u(e_i) \mid e_j)$$

donc, par bilinéarité, on a

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

Par conséquent, si l'on montre qu'il existe une base  $\mathcal{B} = (f_1, \dots, f_n)$  **orthonormée** pour  $(\mid)$  et formée de vecteurs propres de  $u$ , i.e.  $u(f_i) = \lambda_i f_i$ , on aura pour tout  $i, j$  :

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

donc  $\mathcal{B}$  sera aussi une base **orthogonale** pour  $\phi$ . Donc, 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  $x_1^2 + \dots + x_n^2$ , et la forme  $Q$  en la somme de carrés  $\lambda_1 x_1^2 + \dots + \lambda_n x_n^2$ ; pour cette raison, le théorème ci-dessous (cf. 6.2.9) est appelé « théorème de **réduction simultanée** » ou « de **diagonalisation simultanée** ».



**Théorème 7.5.9 (Réduction simultanée).** — Soient  $E$  muni de  $(\mid)$  un espace euclidien,  $Q$  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  $(|)$  et formée de vecteurs propres de  $u$ , i.e.  $u(f_i) = \lambda_i f_i$ , et l'on a

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

et  $\lambda_1, \dots, \lambda_n$  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  $P^{-1}AP = \text{Mat}_{\mathcal{B}}(u)$ .

Démontrons maintenant le théorème 7.5.8. D'après le théorème 7.5.9, il existe une base orthonormée  $\mathcal{B} = (\vec{f}_1, \vec{f}_2)$  de  $\mathbb{R}^2$  telle que, notant  $(X, Y)$  les coordonnées dans le repère orthonormé  $\mathcal{R} = (O, \mathcal{B})$ , on ait

$$Q(X, Y) = \lambda X^2 + \mu Y^2, \quad \text{avec } \lambda\mu = \det S = ac - b^2 = \delta.$$

D'autre part, si l'on pose  $2\rho = L(\vec{f}_1)$  et  $2\sigma = L(\vec{f}_2)$ , alors la forme linéaire  $L$  s'exprime, dans les coordonnées  $(X, Y)$ , par  $L(X, Y) = 2\rho X + 2\sigma Y$  (on n'a pas besoin de calculer explicitement les coefficients  $\rho$  et  $\sigma$ ). Donc, avec les coordonnées  $(X, Y)$ , on obtient que :

$$\mathcal{C} = \{M(X, Y) \mid \lambda X^2 + \mu Y^2 + 2\rho X + 2\sigma Y + f = 0\}.$$

Distinguons maintenant les cas suivants.

(1) Si  $0 = \delta = \lambda\mu$ , on peut supposer, quitte à échanger  $X$  et  $Y$  (i.e. à remplacer  $(\vec{f}_1, \vec{f}_2)$  par  $(\vec{f}_2, \vec{f}_1)$ ) que  $\lambda = 0 \neq \mu$  ( $\lambda, \mu$  ne sont pas tous deux nuls, puisque la matrice  $S$  n'est pas nulle). Dans ce cas, on obtient l'équation

$$\mu \left( Y + \frac{\sigma}{\mu} \right)^2 = -2\rho X - f + \frac{\sigma^2}{\mu}, \quad \text{soit} \quad \left( Y + \frac{\sigma}{\mu} \right)^2 = -2\frac{\rho}{\mu} X - \frac{f\mu - \sigma^2}{\mu}.$$

Si  $\rho = 0$ , on obtient  $\emptyset$  si  $K = -\frac{f\mu - \sigma^2}{\mu}$  est  $< 0$ , et si  $K \geq 0$  on obtient les droites parallèles d'équation

$$Y = -\frac{\sigma}{\mu} \pm \sqrt{K},$$

celles-ci étant confondues (i.e. égales) si  $K = 0$ . Si  $\rho \neq 0$ , on obtient l'équation

$$\left( Y + \frac{\sigma}{\mu} \right)^2 = -2\frac{\rho}{\mu} \left( X + \frac{f\mu - \sigma^2}{2\rho} \right).$$

Soit alors  $\Omega$  le point de coordonnées  $\left( -\frac{f\mu - \sigma^2}{2\rho}, -\frac{\sigma}{\mu} \right)$ . Notons  $(\tilde{X}, \tilde{Y})$  les coordonnées dans le repère orthonormé  $(\Omega, \vec{f}_1, \vec{f}_2)$  et  $p = -\frac{\rho}{\mu}$ , alors

$$\mathcal{C} = \{M(\tilde{X}, \tilde{Y}) \mid \tilde{Y}^2 = 2p\tilde{X}\}$$

est une parabole d'axe  $\Delta = \Omega + \mathbb{R}\vec{f}_1$ , de sommet  $\Omega$  et de paramètre  $p$ .

Supposons  $\delta = \lambda\mu \neq 0$ . Alors l'équation de  $\mathcal{C}$  est :

$$\lambda \left( X + \frac{\rho}{\lambda} \right)^2 + \mu \left( Y + \frac{\sigma}{\mu} \right)^2 = -f + \frac{\rho^2}{\lambda} + \frac{\sigma^2}{\mu}.$$

Soit  $\Omega$  le point de coordonnées  $\left( -\frac{\rho}{\lambda}, -\frac{\sigma}{\mu} \right)$ . Notant  $(\tilde{X}, \tilde{Y})$  les coordonnées dans le repère orthonormé  $(\Omega, \vec{f}_1, \vec{f}_2)$ , on obtient l'équation

$$\lambda\tilde{X}^2 + \mu\tilde{Y}^2 = K = -f + \frac{\rho^2}{\lambda} + \frac{\sigma^2}{\mu}.$$

(2) Si  $\delta = \lambda\mu > 0$ , alors  $\lambda$  et  $\mu$  sont du même signe. Si  $K$  est du signe opposé, alors  $\mathcal{C} = \emptyset$ , tandis que  $\mathcal{C} = \{\Omega\}$  si  $K = 0$ . Enfin, si  $K$  est du même signe que  $\lambda$  et  $\mu$ , alors  $K/\lambda$  et  $K/\mu$  sont tous deux  $> 0$ ; quitte à échanger  $\tilde{X}$  et  $\tilde{Y}$ , on peut supposer que  $a^2 = K/\lambda > K/\mu = b^2$ , alors

$$\mathcal{C} = \{M(\tilde{X}, \tilde{Y}) \mid \frac{\tilde{X}^2}{a^2} + \frac{\tilde{Y}^2}{b^2} = 1\}$$

est une ellipse, d'axe focal la droite  $\Omega\tilde{X}$ .

(3) Si  $\delta = \lambda\mu < 0$ , alors  $\lambda$  et  $\mu$  sont de signe opposé. Posons  $-\lambda/\mu = t^2$ . Si  $K = 0$ ,  $\mathcal{C}$  est la réunion des droites  $\tilde{Y} = t\tilde{X}$  et  $\tilde{Y} = -t\tilde{X}$ . Enfin, si  $K \neq 0$ , alors

$$\mathcal{C} = \left\{ M(\tilde{X}, \tilde{Y}) \mid \frac{\tilde{X}^2}{K/\lambda} + \frac{\tilde{Y}^2}{K/\mu} = 1 \right\}$$

est une hyperbole, d'asymptotes les droites précédentes, et d'axe focal  $\Omega\tilde{X}$  si  $K/\lambda > 0 > K/\mu$  (et d'axe focal  $\Omega\tilde{Y}$  si  $K/\lambda < 0 < K/\mu$ ). Ceci achève la démonstration du théorème 7.5.8.  $\square$

On peut énoncer le théorème 7.5.8 sous une forme un peu plus générale, de la façon suivante. Soit  $\mathcal{R}' = (O, \vec{u}', \vec{v}')$  un repère arbitraire de  $\mathcal{P}$ , i.e. les vecteurs  $\vec{u}', \vec{v}'$  ne sont pas nécessairement unitaires ni orthogonaux. Notons  $(x', y')$  les coordonnées dans ce repère, soient  $a', b', c', d', e', f' \in \mathbb{R}$ , avec  $(a', b', c') \neq (0, 0, 0)$ , et soit

$$\mathcal{C} = \{M(x', y') \mid a'x'^2 + 2b'x'y' + c'y'^2 + 2d'x' + 2e'y' + f' = 0\}.$$

Soit  $\phi$  la forme bilinéaire symétrique dont la matrice dans la base  $\mathcal{B}' = (\vec{u}', \vec{v}')$  est :

$$A' = \text{Mat}_{\mathcal{B}'}(\phi) = \begin{pmatrix} a' & b' \\ b' & c' \end{pmatrix} \quad \text{i.e.} \quad \phi\left(\begin{pmatrix} x'_1 \\ y'_1 \end{pmatrix}, \begin{pmatrix} x'_2 \\ y'_2 \end{pmatrix}\right) = a'x'_1x'_2 + b'(x'_1y'_2 + y'_1x'_2) + c'y'_1y'_2$$

et soit  $Q$  la forme quadratique associée, i.e.  $Q(x', y') = a'x'^2 + 2b'x'y' + c'y'^2$ . Soient  $\mathcal{B}_0 = (\vec{u}, \vec{v})$ ,  $P$  la matrice de passage  $\text{Mat}_{\mathcal{B}'}(\mathcal{B}_0)$ , et  $(x, y)$  les coordonnées dans le repère  $\mathcal{R}_0 = (O, \mathcal{B}_0)$ . La matrice de  $\phi$  dans  $\mathcal{B}_0$  est :

$$A = \begin{pmatrix} a & b \\ b & c \end{pmatrix} = {}^t P A' P,$$

donc  $ac - b^2 = \det A = (\det P)^2 \cdot \det A' = (\det P)^2 \cdot (a'c' - b'^2)$  a même signe que  $a'c' - b'^2$ . Par conséquent, on déduit du théorème 7.5.8 le :



**Corollaire 7.5.10.** — Soit  $(O, \vec{u}, \vec{v})$  un repère arbitraire de  $\mathcal{P}$ , soient  $(x, y)$  les coordonnées correspondantes, soient  $a, b, c, d, e, f \in \mathbb{R}$ , avec  $(a, b, c) \neq (0, 0, 0)$ , et soit

$$\mathcal{C} = \{M(x, y) \mid ax^2 + 2bxy + cy^2 + 2dx + 2ey + f = 0\}.$$

On suppose  $\mathcal{C} \neq \emptyset$ . Alors :

(1) Si  $\delta = 0$ ,  $\mathcal{C}$  est une parabole ou bien la réunion de deux droites parallèles ou « confondues » (i.e. égales).

(2) Si  $\delta = ac - b^2 > 0$ ,  $\mathcal{C}$  est une ellipse ou un point.

(3) Si  $\delta < 0$ ,  $\mathcal{C}$  est une hyperbole ou bien la réunion de deux droites sécantes.

## 7.6. Quadriques en dimension 3

On a des résultats analogues dans l'espace affine euclidien  $\mathcal{E}$  de dimension 3. Fixons un repère orthonormé  $\mathcal{R}_0 = (O, \mathcal{B}_0)$  et soient  $(x, y, z)$  les coordonnées dans ce repère. Soit  $Q(x, y, z)$  une forme quadratique non nulle,  $L(x, y, z)$  une forme linéaire, et  $c$  une constante. On considère

$$\mathcal{C} = \{M(x, y, z) \in \mathcal{E} \mid Q(x, y, z) + L(x, y, z) + c = 0\}.$$

D'abord, d'après le théorème 7.5.9, il existe une base orthonormée  $\mathcal{B}$  de  $E = \mathbb{R}^3$  telle que, notant  $(X, Y, Z)$  les coordonnées dans le repère  $\mathcal{R} = (O, \mathcal{B})$ , on ait

$$Q(X, Y, Z) = \lambda X^2 + \mu Y^2 + \nu Z^2.$$

Alors  $L(X, Y, Z)$  est encore une forme linéaire en  $X, Y, Z$ , donc de la forme  $2\rho X + 2\sigma Y + 2\tau Z$ , et l'on obtient que :

$$(*) \quad \mathcal{C} = \{M(X, Y, Z) \in \mathcal{E} \mid \lambda X^2 + \mu Y^2 + \nu Z^2 + 2\rho X + 2\sigma Y + 2\tau Z + c = 0\}.$$

Distinguons les cas en fonction du rang de  $Q$ , qui peut être 3, 2 ou 1 (car on a supposé  $Q \neq 0$ ).

Supposons d'abord  $\boxed{\text{rang}(Q) = 3}$ , i.e.  $\lambda\mu\nu \neq 0$ . Dans ce cas, on peut faire disparaître tous les termes linéaires, en remplaçant  $X$  par  $X + (\rho/\lambda)$ , etc. Notons alors

$$\Omega = \left( -\frac{\rho}{\lambda}, -\frac{\sigma}{\mu}, -\frac{\tau}{\nu} \right)$$

et, pour alléger l'écriture, notons encore  $(X, Y, Z)$  les coordonnées dans le repère  $(\Omega, \mathcal{B})$  (au lieu de les noter  $\tilde{X}$ , etc.). On obtient alors une équation

$$\lambda X^2 + \mu Y^2 + \nu Z^2 = K.$$

De plus, quitte à changer  $K$  en  $-K$ , on peut supposer que  $Q$  est de signature  $(3, 0)$  ou bien  $(2, 1)$ .

**7.6.1. Ellipsoïdes.** — Supposons  $Q$  de signature  $(3, 0)$ . Alors  $\mathcal{C}$  est vide si  $K < 0$ , et  $\mathcal{C} = \{\Omega\}$  si  $K = 0$ . Si  $K > 0$ , posons  $K/\lambda = a^2$ ,  $K/\mu = b^2$ ,  $K/\nu = c^2$ , avec  $a, b, c > 0$ , alors

$$\mathcal{C} = \{M(X, Y, Z) \in \mathcal{E} \mid \frac{X^2}{a^2} + \frac{Y^2}{b^2} + \frac{Z^2}{c^2} = 1\}$$

et l'on dit que  $\mathcal{C}$  est un **ellipsoïde**. Si  $a = b = c$ , c'est une **sphère** de rayon  $a$ . Si  $a = b \neq c$ , l'ellipsoïde est invariant par rotation autour de l'axe  $\Omega Z$ , on dit alors que c'est un **ellipsoïde de révolution**.

**7.6.2. Cônes et hyperboloïdes.** — Supposons  $Q$  de signature  $(2, 1)$ . Quitte à permuter  $X, Y, Z$ , on peut supposer que  $\nu < 0 < \lambda, \mu$ . Posons  $\nu' = -\nu > 0$ .

(a) Considérons d'abord le cas où  $K = 0$ . Dans ce cas,  $\mathcal{C}$  est le **cône quadratique**

$$\mathcal{C}_0 = \{M(X, Y, Z) \in \mathcal{E} \mid \lambda X^2 + \mu Y^2 = \nu' Z^2\}.$$

On voit que le complémentaire dans  $\mathcal{E}$  de ce cône a 3 composantes connexes; dans chacune de ces composantes connexes,  $Q(X, Y, Z) = \lambda X^2 + \mu Y^2 - \nu' Z^2$  ne s'annule pas, donc garde un signe constant, et ce signe est :

(1) à « l'extérieur » du cône, qui est la composante connexe contenant, par exemple, le point  $(1, 1, 0)$ , le signe est  $> 0$ ,

(2) à « l'intérieur du demi-cône supérieur », qui est la composante connexe contenant, par exemple, le point  $(0, 0, 1)$ , le signe est  $< 0$ ,

(3) de même, à « l'intérieur du demi-cône inférieur », qui est la composante connexe contenant, par exemple, le point  $(0, 0, -1)$ , le signe est  $< 0$ .

(b) On en déduit que lorsque  $K > 0$ ,

$$\mathcal{C} = \{M(X, Y, Z) \in \mathcal{E} \mid \lambda X^2 + \mu Y^2 - \nu' Z^2 = K > 0\}$$

est situé « à l'extérieur » du cône  $\mathcal{C}_0$ , et l'on peut montrer que dans ce cas  $\mathcal{C}$  est **connexe**. Posant  $K/\lambda = a^2$ ,  $K/\mu = b^2$ ,  $K/\nu' = c^2$ , on obtient que

$$\mathcal{C} = \{M(X, Y, Z) \in \mathcal{E} \mid \frac{X^2}{a^2} + \frac{Y^2}{b^2} - \frac{Z^2}{c^2} = 1\}$$

est connexe; on dit que c'est un **hyperboloïde à une nappe**.

(c) Si  $K < 0$ , alors

$$\mathcal{C} = \{M(X, Y, Z) \in \mathcal{E} \mid \lambda X^2 + \mu Y^2 - \nu' Z^2 = K < 0\}$$

est situé « à l'intérieur » du cône  $\mathcal{C}_0$ , et est symétrique par rapport au plan  $Z = 0$ . Donc, dans ce cas,  $\mathcal{C}$  est la réunion disjointe des deux ouverts  $\mathcal{C}_+ = \mathcal{C} \cap \{Z > 0\}$  et  $\mathcal{C}_- = \mathcal{C} \cap \{Z < 0\}$ , donc  $\mathcal{C}$  n'est **pas connexe**. On peut montrer que  $\mathcal{C}_+$  et  $\mathcal{C}_-$  sont connexes, donc

$$\mathcal{C} = \{M(X, Y, Z) \in \mathcal{E} \mid \frac{\lambda}{K} X^2 + \frac{\mu}{K} Y^2 - \frac{\nu'}{K} Z^2 = 1\}$$

(où ici  $K < 0$ ) a deux composantes connexes. Posant alors  $-K/\lambda = a^2$ ,  $-K/\mu = b^2$ ,  $-K/\nu' = c^2$ , on obtient que

$$\mathcal{C} = \{M(X, Y, Z) \in \mathcal{E} \mid \frac{Z^2}{c^2} - \frac{X^2}{a^2} - \frac{Y^2}{b^2} = 1\}$$

a deux composantes connexes; on dit que c'est un **hyperboloïde à deux nappes**.

**Définition 7.6.3.** — Si  $a = b$  dans ce qui précède, on obtient :

$$\text{le cône } \mathcal{C}_0 = \{M(X, Y, Z) \mid X^2 + Y^2 - \frac{a^2}{c^2} Z^2 = 0\}$$

$$\text{l'hyperboloïde à une nappe } \mathcal{C}_1 = \{M(X, Y, Z) \mid X^2 + Y^2 - \frac{a^2}{c^2} Z^2 = 1\}$$

$$\text{l'hyperboloïde à deux nappes } \mathcal{C}_2 = \{M(X, Y, Z) \mid \frac{a^2}{c^2} Z^2 - X^2 - Y^2 = 1\};$$

ces surfaces sont invariantes par rotation autour de l'axe  $\Omega Z$  ; on dit que  $\mathcal{C}_0$  est un **cône de révolution**, et  $\mathcal{C}_1$  (resp.  $\mathcal{C}_2$ ) un **hyperboloïde de révolution** à une nappe (resp. deux nappes).

**7.6.4. Paraboloïdes et cylindres.** — Supposons maintenant  $\boxed{\text{rang}(Q) = 2}$ . Quitte à permuter  $X, Y, Z$ , on peut supposer que  $\nu = 0 \neq \lambda\mu$ . Alors, écrivant  $X$  et  $Y$  au lieu de  $X + \frac{\rho}{\lambda}$  et  $Y + \frac{\sigma}{\mu}$ , on obtient une équation de la forme :

$$(\dagger) \quad \lambda X^2 + \mu Y^2 = pZ + t.$$

(a) Supposons, pour commencer  $p \neq 0$ . Alors, notant  $Z$  au lieu de  $Z + \frac{t}{p}$ , on se ramène à l'équation

$$\lambda X^2 + \mu Y^2 = pZ.$$

De plus, quitte à changer  $p$  en  $-p$ , on peut supposer que  $Q = \lambda X^2 + \mu Y^2$  est de signature  $(2, 0)$  ou bien  $(1, 1)$ . Dans le premier cas, posons  $1/\lambda = a^2$  et  $1/\mu = b^2$ , alors

$$\mathcal{C} = \{M(X, Y, Z) \in \mathcal{E} \mid \frac{X^2}{a^2} + \frac{Y^2}{b^2} = pZ\}$$

et l'on dit que c'est un **paraboloïde elliptique**. Dans le second cas, quitte à échanger  $X$  et  $Y$ , on peut supposer que  $1/\lambda = a^2$  et  $-1/\mu = b^2$ , alors

$$\mathcal{C} = \{M(X, Y, Z) \in \mathcal{E} \mid \frac{X^2}{a^2} - \frac{Y^2}{b^2} = pZ\}$$

et l'on dit que c'est un **paraboloïde hyperbolique**.

(b) Considérons maintenant le cas où  $p = 0$  (dans ce cas, la coordonnée  $Z$  n'intervient pas dans l'équation  $(\dagger)$  plus haut). Quitte à changer  $t$  en  $-t$ , on peut supposer que  $\lambda, \mu > 0$  ou bien que  $\lambda > 0 > \mu$ . Dans le premier cas, si  $t < 0$  on a  $\mathcal{C} = \emptyset$ , et si  $t = 0$  alors  $\mathcal{C}$  est la droite  $X = Y = 0$ ; enfin si  $t > 0$ , alors  $\mathcal{C}$  est un **cylindre elliptique**, de section l'ellipse du plan  $Z = 0$  définie par l'équation  $(\dagger)$ .

Dans le second cas, si  $t = 0$  alors  $\mathcal{C}$  est la réunion de deux plans sécants, et si  $t \neq 0$  alors  $\mathcal{C}$  est un **cylindre hyperbolique**, de section l'hyperbole du plan  $Z = 0$  définie par l'équation  $(\dagger)$ .

Supposons enfin  $\boxed{\text{rang}(Q) = 1}$ . Quitte à permuter  $X, Y, Z$ , on peut supposer que  $\nu = \mu = 0 \neq \lambda$ . Alors, écrivant  $X$  au lieu de  $X + \frac{\rho}{\lambda}$ , on obtient une équation de la forme :

$$X^2 = pZ + qY + t.$$

Supposons, pour simplifier  $(p, q) \neq (0, 0)$  (on laisse au lecteur le soin de traiter le cas  $p = q = 0$ ). Posons  $r = \sqrt{p^2 + q^2}$ , alors la matrice

$$\frac{1}{r} \begin{pmatrix} p & q \\ q & -p \end{pmatrix}$$

est orthogonale, donc les coordonnées  $\tilde{Z} = \frac{1}{r}(pZ + qY)$ ,  $\tilde{Y} = \frac{1}{r}(qZ - pY)$  correspondent à une nouvelle base orthonormée. Posant  $Z' = \tilde{Z} + (t/r)$  et  $Y' = \tilde{Y}$ , on obtient que

$$\mathcal{C} = \{M(X, Y', Z') \in \mathcal{E} \mid X^2 = rZ'\}$$

et l'on dit alors que  $\mathcal{C}$  est un **cylindre parabolique**.

(Pour voir des figures d'ellipsoïde, d'hyperboloïde à une ou deux nappes, de paraboloïde elliptique ou hyperbolique, et de cylindre parabolique, voir par exemple : J. Lelong-Ferrand & J.-M. Arnaudiès, Cours de mathématiques, Tome 3, Géométrie et cinématique, (2ème édition), Chap. III, §§7–9, ou : C. Tisseron, Géométries affine, projective et euclidienne, p. 112.)



## CHAPITRE 8

### FORMES HERMITIENNES, ESPACES HILBERTIENS ET GROUPES UNITAIRES $U(n)$

**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 8.4.7 et 8.4.8.

On a indiqué par des symboles  $\Downarrow$  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.

#### 8.0. 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 8.0.1 (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 6.5). 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  $z$ .

#### 8.1. Formes hermitiennes

**Exemple 8.1.0.** — 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 :

$$\phi(\lambda x + \mu x', y) = (\lambda x + \mu x')\bar{y} = \lambda x\bar{y} + \mu x'\bar{y} = \lambda\phi(x, y) + \mu\phi(x', y)$$

<sup>(0)</sup>version du 10/7/2012

mais pas tout-à-fait linéaire en la 2ème variable, puisqu'on a :

$$\phi(x, \lambda y + \mu y') = x \overline{\lambda y + \mu y'} = x \overline{\lambda y} + x \overline{\mu y'} = \overline{\lambda} \phi(x, y) + \overline{\mu} \phi(x', y).$$

D'autre part, on a  $\phi(y, x) = y \overline{x} = \overline{x \overline{y}} = \overline{\phi(x, y)}$ . Ceci conduit aux définitions suivantes.

**Définition 8.1.1 (Applications semi-linéaires).** — Soient  $E, F$  deux  $\mathbb{C}$ -espaces vectoriels. Une application  $f : E \rightarrow F$  est dite **semi-linéaire** si elle vérifie :

$$\forall u, v \in E, \quad \forall z \in \mathbb{C}, \quad \boxed{f(u+v) = f(u) + f(v)} \quad \boxed{f(zu) = \overline{z}f(u)}.$$

Ces deux conditions équivalent bien sûr à la condition :  $\boxed{f(zu+v) = \overline{z}f(u) + f(v)}$ .



**Définition 8.1.2 (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. :

$$\forall x, x', y, y' \in E, \quad \forall \lambda \in \mathbb{C}, \quad \begin{cases} \phi(\lambda x + x', y) = \lambda \phi(x, y) + \phi(x', y), \\ \phi(x, \lambda y + y') = \overline{\lambda} \phi(x, y) + \phi(x, y') \end{cases}$$

(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 :

$$\phi(x, \lambda y + y') = \overline{\phi(\lambda y + y', x)} = \overline{\lambda \phi(y, x) + \phi(y', x)} = \overline{\lambda} \overline{\phi(y, x)} + \overline{\phi(y', x)} = \overline{\lambda} \phi(x, y) + \phi(x, y').$$

**Remarque 8.1.2.1.** — 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 8.1.3 (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

$$Q(\lambda x) = \phi(\lambda x, \lambda x) = \lambda \overline{\lambda} \phi(x, x) = \lambda \overline{\lambda} Q(x) = |\lambda|^2 Q(x).$$

**Lemme 8.1.4 (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 8.1.5 (Matrices hermitiennes).** — Une matrice  $A \in M_n(\mathbb{C})$  est **hermitienne** si  ${}^tA = \overline{A}$ . 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  $a_{ii} = \overline{a_{ii}}$  donc  $a_{ii} \in \mathbb{R}$ .

**Remarque 8.1.5.1.** — 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 8.1.6 (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 :

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

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 8.1.7 (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 8.1.6 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 8.1.8 (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  ${}^tP$  et  $\overline{P}$  sont inversibles (on a  $({}^tP)^{-1} = {}^t(P^{-1})$  et  $(\overline{P})^{-1} = \overline{P^{-1}}$ ), alors la matrice  $A' = \text{Mat}_{\mathcal{B}'}(\phi) = {}^tP A \overline{P}$  a même rang que  $A$ .  $\square$

**Définition et proposition 8.1.9 (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  $Y^{\perp\phi}$  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  $N(\phi) = E^\perp = \{x \in E \mid \phi(x, y) = 0, \forall y \in Y\}$  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 8.1.10 (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)  $N(\phi) = \{0\} \Leftrightarrow \phi$  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 8.1.9,  $F^\perp$  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 8.1.11 (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 8.1.12 (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 renuméroter 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 8.1.4, la forme quadratique hermitienne  $Q$  est non nulle, donc il existe  $e_1 \in E$  tel que  $Q(e_1) \neq 0$ . Posons  $F = \text{Vect}(e_1)$ , comme  $\phi(e_1, e_1) \neq 0$ , alors  $F \cap F^\perp = \{0\}$  donc, d'après le théorème 8.1.10, 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 :

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

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  $u \in F^\perp = \text{Vect}(e_{r+1}, \dots, e_n)$ . Réciproquement, (\*) montre aussi que tout  $u \in F^\perp$  (i.e. tel que  $x_1 = 0 = \dots = x_r$ ) 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  $(\star)$  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 (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.  $\square$

## 8.2. Réduction en sommes de carrés de modules

**Définition 8.2.1.** — 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

$$Q = q_1 |x_1|^2 + \dots + q_n |x_n|^2, \quad (\text{avec } q_1, \dots, q_n \in \mathbb{R}).$$

D'après 8.1.11, 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 5.2.6 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 :

$$|z|^2 + 2\Re(z\bar{L}) = |z + L|^2 - |L|^2, \quad 4\Re(z_1\bar{z}_2) = |z_1 + \bar{z}_2|^2 - |z_1 - \bar{z}_2|^2.$$

### **Théorème 8.2.2 (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 :

$$(*) \quad Q(x_1, \dots, x_n) = \sum_i b_i |x_i|^2 + \sum_{1 \leq i < j \leq n} (b_{ij} x_i \bar{x}_j + \bar{b}_{ij} x_j \bar{x}_i) \quad (b_i \in \mathbb{R}, \quad b_{ij} \in \mathbb{C}).$$

(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 8.1.12.

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  $\overline{x_1}$  et on l'écrit comme suit :

$$S = b_1 |x_1|^2 + \sum_{j=2}^n 2\mathcal{R}(\overline{b_{j1}} x_1 \overline{x_j}) = b_1 \left( |x_1|^2 + 2\mathcal{R} \left( x_1 \overline{L(x_2, \dots, x_n)} \right) \right)$$

où  $L(x_2, \dots, x_n) = \sum_{j=2}^n (b_{j1}/b_1)x_j$ . Puis, en utilisant que

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

on réécrit ceci sous la forme :

$$S = b_1 |x_1 + L|^2 - b_1 |L|^2 = b_1 \left| x_1 + \sum_{j=2}^n \frac{b_{j1}}{b_1} x_j \right|^2 - \frac{1}{b_1} \sum_{j=2}^n |b_{j1}|^2 |x_j|^2 - \frac{2}{b_1} \sum_{2 \leq i < j \leq n} \mathcal{R}(b_{i1} \overline{b_{j1}} x_i \overline{x_j}).$$

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 :

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

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 (†) :

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

(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

$$\mathcal{R}(b_{12} x_1 \overline{x_2}) = \frac{1}{4} (|b_{12} x_1 + x_2|^2 - |b_{12} x_1 - x_2|^2),$$

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 :  $2\mathcal{R}(b_{12} x_1 \overline{x_2}) = 2(|y_1|^2 - |y_2|^2)$ , les termes  $2\mathcal{R}(b_{ij} x_i \overline{x_j})$  sont inchangés pour  $i < j$  dans  $\{3, \dots, n\}$ , et l'on a :

$$\begin{cases} 2\mathcal{R}(b_{1j} x_1 \overline{x_j}) = 2\mathcal{R}(b_{1j} b_{12}^{-1} (y_1 + y_2) \overline{x_j}) & \text{pour } j \geq 3 \\ 2\mathcal{R}(b_{2j} x_2 \overline{x_j}) = 2\mathcal{R}(b_{2j} (y_1 - y_2) \overline{x_j}) & \text{pour } j \geq 3 \end{cases}$$

donc  $Q(y_1, y_2, x_3, \dots, x_n)$  égale :

$$2|y_1|^2 - 2|y_2|^2 + \sum_{j=3}^n 2\mathcal{R} \left( (b_{1j} b_{12}^{-1} + b_{2j}) y_1 \overline{x_j} + (b_{1j} b_{12}^{-1} - b_{2j}) y_2 \overline{x_j} \right) + \sum_{3 \leq i < j \leq n} 2\mathcal{R}(b_{ij} x_i \overline{x_j})$$

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 8.2.3.** — 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 8.2.4.** — 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.

### 8.3. Espaces hilbertiens. Inégalité de Cauchy-Schwarz. Isométries

**Définitions 8.3.1 (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 :

(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 **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  $(\mid)$  » (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  $(\mid)$ , celui-ci étant sous-entendu.



**Exemple 8.3.2.** — Le produit scalaire hilbertien standard sur  $\mathbb{C}^n$  est défini par :

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

Pour ce produit scalaire, la base canonique  $(e_1, \dots, e_n)$  de  $\mathbb{R}^n$  est *orthonormée*, i.e. on a  $(e_i \mid e_j) = 1$  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 8.3.3 (Familles et bases orthonormées)**

Soit  $E$ , muni de  $(\mid)$ , un espace hilbertien.

(1) Une famille  $(e_i)_{i \in I}$  de vecteurs est dite **orthonormée** si  $(e_i \mid e_i) = 1$  et  $(e_i \mid e_j) = 0$  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  $(e_i \mid e_i) = 1$  et  $(e_i \mid e_j) = 0$  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  $(e_{i_r} \mid \mid)$  à l'égalité précédente. Comme  $(e_{i_r} \mid e_{i_s}) = 0$  pour  $s \neq r$ , on obtient  $0 = t_r (e_{i_r} \mid e_{i_r}) = t_r$ , d'où  $t_r = 0$ . Ceci prouve que la famille  $(e_i)_{i \in I}$  est libre.  $\square$



**Théorème 8.3.4 (Existence de b.o.n.).** — Soit  $E$  un espace hilbertien de dimension  $n$ . Alors  $E$  admet une base orthonormée

*Démonstration.* — D'après le théorème 8.1.12, il existe une base  $(e_1, \dots, e_n)$  orthogonale (i.e.  $(e_i \mid e_j) = 0$  pour  $i \neq j$ ) et telle que  $(e_i \mid e_i) \in \{1, -1, 0\}$ ; or comme  $(\mid)$  est défini positif on a nécessairement  $(e_i \mid e_i) = 1$ , donc  $(e_1, \dots, e_n)$  est une b.o.n.  $\square$

**Définition 8.3.5 (Sous-espaces d'un espace hilbertien).** — Soit  $E$ , muni de  $(\mid)$ , un espace hilbertien et soit  $F$  un sous-espace vectoriel de  $E$ . Alors la restriction  $(\mid)_F$  de  $(\mid)$  à  $F$  est un produit scalaire hilbertien 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 hilbertien.



**Théorème et définition 8.3.6 (Projection orthogonale sur un sous-espace)**

Soit  $E$ , muni de  $(\mid)$ , un espace hilbertien et soient  $F$  un sous-espace et  $F^\perp$  son orthogonal pour  $(\mid)$ .

(1) On a  $E = F \oplus F^\perp$ . 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  $\pi_F(v) = (v \mid e_1)e_1 + \cdots + (v \mid e_r)e_r$  pour tout  $v \in E$ .

(3) On a  $(F^\perp)^\perp = F$  donc la projection orthogonale  $\pi_{F^\perp}$  sur  $F^\perp$  n'est autre que  $\text{id}_E - \pi_F$ , i.e. on a  $\text{id}_E = \pi_F + \pi_{F^\perp}$ .

*Démonstration.* — Comme les formes hermitiennes  $(\mid)$  et  $(\mid)_F$  sont définies positives, donc non-dégénérées, on a  $(F^\perp)^\perp = F$  et  $E = F \oplus F^\perp$  d'après 8.1.10. 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

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



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 8.3.7 (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 8.3.8 (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

$$(\dagger) \quad \|x\|^2 + \|y\|^2 + 2|(x | y)| \leq \|x\|^2 + \|y\|^2 + 2\|x\| \cdot \|y\| = (\|x\| + \|y\|)^2.$$

D'autre part, d'après les égalités de polarisations 8.1.4, 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  $(\dagger)$ , 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 8.1.4 en utilisant la norme  $\|\cdot\|$ , et ajoutons-y l'égalité de Pythagore :

**Proposition 8.3.9 (Polarisation).** — Soit  $E$ , muni de  $(\cdot | \cdot)$ , 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 8.1.4. 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 8.3.10 (Isométries vectorielles).** — Soient  $E, F$  deux espaces hilbertiens de même dimension  $n$ , notons  $(\cdot | \cdot)_E$  et  $\|\cdot\|_E$  (resp.  $(\cdot | \cdot)_F$  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 8.3.10.1.** — 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 8.3.11.** — (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 8.3.4,  $E$  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 8.3.12.** — On note  $U(n) = \{A \in M_n(\mathbb{C}) \mid {}^t A \overline{A} = I_n\}$ . Remarquons que l'égalité  ${}^t A \overline{A} = I_n$  équivaut à l'égalité  $\overline{A} A = I_n$ , qui entraîne que  $A$  est inversible et  $A^{-1} = \overline{A}$ . Donc  $U(n) \subset GL_n(\mathbb{C})$  et, si  $A \in U(n)$ , son inverse  $B = A^{-1} = \overline{A}$  vérifie  $B^{-1} = A = \overline{B}$ , donc appartient aussi à  $U(n)$ . De plus, pour tout  $A, B \in U(n)$ , on a l'égalité  ${}^t(AB) \overline{AB} = {}^t B {}^t A \overline{A} \overline{B} = {}^t B \overline{B} = I_n$ , 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  $(X \mid Y) = {}^t X \overline{Y}$ , 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  $X, Y \in \mathbb{C}^n$  :

$${}^t X \overline{Y} = (X \mid Y) = (AX \mid AY) = {}^t X ({}^t A \overline{A}) Y$$

ce qui équivaut à dire que  ${}^t A \overline{A} = I_n$  (cf. 8.1.6). 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  ${}^t A \overline{A}$  est le produit matriciel de la  $i$ -ème ligne de  ${}^t A$ , i.e. de  ${}^t C_i$ , par la colonne  $\overline{C_j}$ , c.-à-d., on a  $({}^t A \overline{A})_{ij} = (Ae_i \mid Ae_j)$ , donc la condition  ${}^t A \overline{A} = I_n$  équivaut aussi à dire que les colonnes de  $A$  sont de norme 1 et deux à deux orthogonales. Tenant compte de la proposition 8.3.10, on obtient donc les caractérisations suivantes de  $U(n)$ , chacune étant utile :



**Proposition 8.3.13 (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 :

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

Les éléments de  $U(n)$  sont appelés « endomorphismes unitaires ».

**Remarque 8.3.14.** — 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.

## 8.4. 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 8.4.1 (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

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

(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

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

d'où  ${}^tA J = \overline{J A^*}$  et donc, puisque  $J$  est inversible (car  $\phi$  non-dégénérée),  $A^* = \overline{J^{-1} {}^tA J}$ . 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  $A^* = \overline{J^{-1} {}^tA J}$ , alors pour tout  $x, y$  on a :

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

donc  $u^*$  vérifie (\*). Ceci prouve les assertions (1) et (2).

Prouvons l'assertion (3). Pour tout  $x, y \in E$ , on a :

$$(u^*(x) | y) = \overline{(y | u^*(x))} = \overline{(u(y) | x)} = (x | u(y))$$

et ceci montre que  $u$  est l'adjoint de  $u^*$ , i.e.  $u = (u^*)^*$ . Le théorème est démontré.  $\square$

**Remarque 8.4.2.** — 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 8.4.3 (Adjoint d'un endomorphisme dans le cas hilbertien)**

Soit  $E$  muni de  $(|)$  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 :

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

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

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

**Lemme 8.4.4 (Stabilité par  $u$  ou  $u^*$ ).** — Soient  $E$  muni de  $(|)$  un espace hilbertien de dimension  $n$ ,  $u \in \text{End}(E)$ ,  $F$  un sous-espace vectoriel de  $E$ , et  $F^\perp$  son orthogonal pour  $(|)$ . 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 :

$$(x | u^*(y)) = (u(x) | y) = 0$$

(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 8.3.6 et 8.4.1, on obtient que  $u(F) \subseteq F$ , d'après ce qui précède.  $\square$

**Définitions 8.4.5 (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. 8.3.13).
- (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  $u u^* = u^* u$ . Ceci englobe les trois cas précédents.

Rappelons et complétons la définition 8.1.5 :

**Définition 8.4.6 (Matrices hermitiennes ou anti-hermitiennes).** — Une matrice  $A \in M_n(\mathbb{C})$  est dite **anti-hermitienne** (resp. *hermitienne*, cf. 8.1.5) si  $\boxed{{}^t\bar{A} = -A}$  (resp. si  $\boxed{{}^t\bar{A} = A}$ ).

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 8.4.7 (Diagonalisation des endomorphismes normaux)**

Soient  $E$  muni de  $(\cdot | \cdot)$  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 8.4.4, l'orthogonal  $G = V_\lambda^\perp$  est stable par  $u$  et par  $u^*$ . D'autre part, d'après 8.1.10, 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 :

$$(u_G(x) | y) = (u(x) | y) = (x | u^*(y)) = (x | u_G^*(y))$$

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 8.4.8 (Diagonalisation des matrices hermitiennes, unitaires, ou anti-hermitiennes)**

On munit  $\mathbb{C}^n$  du produit scalaire hilbertien standard. Soit  $A \in M_n(\mathbb{C})$ .

(1) Si  $A$  est **hermitienne** (i.e.  $\overline{A}^t = A$ ), 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.  $\overline{A}^t = -A$ ), 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

$$\overline{D} = {}^t\overline{D} = {}^t\overline{P}^t \overline{A}^t \overline{P}^{-1} = P^{-1} {}^t\overline{A} P = \begin{cases} P^{-1}AP = D & \text{si } A \text{ est hermitienne} \\ P^{-1}A^{-1}P = D^{-1} & \text{si } A \text{ est unitaire} \\ -P^{-1}AP = -D & \text{si } A \text{ est anti-hermitienne.} \end{cases}$$



Il en résulte que les termes diagonaux  $\lambda_i$  de  $D$  vérifient, respectivement,  $\overline{\lambda_i} = \lambda_i$  (resp.  $= \lambda_i^{-1}$ , resp.  $= -\lambda_i$ ), donc sont réels (resp. de module 1, resp. imaginaires purs). Le théorème est démontré.  $\square$

**Remarque 8.4.9.** — Le point (1) du théorème précédent fournit une autre démonstration de la proposition 6.2.8 et donc du théorème 6.2.6.

### 8.5. 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 8.5.1 (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

$$P^{-1}AP = \text{Mat}_{\mathcal{B}}(f) = \begin{pmatrix} I_p & 0 & 0 & \cdots & 0 \\ 0 & -I_q & 0 & \cdots & 0 \\ \hline 0 & 0 & R(\theta_1) & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \hline 0 & \cdots & 0 & 0 & R(\theta_r) \end{pmatrix}.$$

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 8.5.2.** — 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. 6.1.9), 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 8.5.1. 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 6.4.3, 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 8.5.3.** — 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 8.5.1. 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. :

$$E_1 = \{x \in E \mid (x \mid y) = 0, \quad \forall y \in P_1\}.$$

D'après le lemme 8.5.2,  $E_1$  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 8.5.3 : 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 8.5.1, modulo la démonstration du lemme 8.5.3.  $\square$

Avant de démontrer le lemme 8.5.3, faisons les remarques suivantes.

**Remarques 8.5.4.** — (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. 6.4.14).

(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. 6.4.18).

(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 8.5.3. 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

$$(2) \quad \bar{\lambda}\bar{w} = (a - ib)(u - iv) = \underbrace{(au - bv)}_{\in \mathbb{R}^m} - i \underbrace{(bu + av)}_{\in \mathbb{R}^m} = \overline{\lambda w}.$$

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 hilbertien standard sur  $\mathbb{C}^m$ , qu'on notera  $\langle | \rangle$ , i.e.

$$\forall W = \begin{pmatrix} w_1 \\ \vdots \\ w_m \end{pmatrix}, \forall Z = \begin{pmatrix} z_1 \\ \vdots \\ z_m \end{pmatrix} \in \mathbb{C}^m, \quad \langle W | Z \rangle = {}^tW\bar{Z} = w_1\bar{z}_1 + \cdots + w_m\bar{z}_m.$$

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 8.4.8. 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  $\mathcal{C} = (v, u)$  est une base orthonormée du plan  $P$ , et d'après (6) on a  $\text{Mat}_{\mathcal{C}}(f_P) = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$ , 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 8.5.3 et donc du théorème 8.5.1.  $\square$



## 8.6. Appendice (†) : espaces préhilbertiens réels ou complexes

Si  $E$ , muni de  $(\mid)$ , 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  $\|x\| = \sqrt{(x \mid x)}$  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

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

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

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

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

$$(f \mid g) = \int_0^1 f(t)\overline{g(t)} dt,$$

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

$$(f \mid g) = \int_0^1 f(t)\overline{g(t)} dt,$$

est un espace de Hilbert complexe, i.e. il est complet pour la norme  $\|f\|_2 = \int_0^1 |f(t)|^2 dt$ .



# INDEX

- Absolument convergentes (séries), 79
- Adjoint d'un endomorphisme, 116, 174
- Affinement indépendants (points), 145
- Algèbres  $M_n(k)$  et  $\text{End}_k(V)$ , 9
- Algèbres, 84
- Algébriquement clos (corps), 47
- Alterné (groupe), 90
- Alternées
  - applications  $p$ -linéaires, 90
  - formes bilinéaires, 107
- Angle non orienté de deux vecteurs, 114
- Angles orientés dans  $\mathbb{R}^2$ , 124
- Anti-hermitiennes (matrices), 174
- Antidéplacements, 148
- Applications affines, 139
- Applications linéaires, 2
- Applications linéaires et matrices, 8
- Auto-adjoint (endomorphisme), 116, 174
- Axe focal, 153
- Barycentres, 141
- Bases, 4
- Base duale, 20
- Base préduale, 20
- Bases orthogonales
  - pour une forme bilinéaire symétrique, 102
  - pour une forme hermitienne, 166
- Bases orthonormées, 112, 170
- Bases orthonormées (existence de), 112, 170
- Bézout (théorème de), 50, 53
- Bidual, 40
- Binôme (formule du), 67
- Binomial (coefficient), 67
- Blocs de Jordan nilpotents, 59
- Blocs de Jordan, 61
- $\mathbb{C}$  est algébriquement clos, 54
- Cauchy (suites de), 69
- Cauchy-Schwarz (inégalité de), 112, 171
- Cayley-Hamilton (théorème de), 50
- Centre de gravité, 141
- Changement de coordonnées (formule de), 12
- Changement de bases pour une application linéaire, 12
- Changement de base
  - pour un endomorphisme, 14
  - pour une forme bilinéaire symétrique, 98
  - pour une forme bilinéaire alternée, 108
  - pour une forme hermitienne, 163
- Changement de repère, 138
- Codimension, 29
- Cofacteurs (matrice des), 35
- Complets (esp. vect. normés), 69
- Congruentes (matrices), 98
- Coniques, 151–157
- Conjugaison complexe, 161
- Continues (applications linéaires), 69
- Coordonnées, 5
- Cycles, 87
- Cylindres, 159
- Cônes quadratiques, 158
- Décomposition canonique d'une isométrie affine, 149
- Déplacements, 148
- Dérivée de  $t \mapsto \exp(tA)$ , 73
- Diagonalisables (endomorphismes), 45
- Diagonalisation des endomorphismes
  - auto-adjoints, 116, 175
  - normaux, 175
  - unitaires, 175
- Diagonalisation simultanée (théorème de), 118
- Dimension d'un espace vectoriel, 4
- Dimension de l'orthogonal
  - pour une forme bilinéaire symétrique, 100
  - pour une forme bilinéaire alternée, 109
- Directes (isométries), 148
- Direction (d'un espace affine), 138
- Directrice d'une conique, 152
- Discriminant réduit d'un trinôme, 113
- Divison euclidienne, 53
- Droites affines, 143
- Dual (espace), 19
- Dunford (décomposition de), 68
- Décomposition en produit de cycles, 88
- Définition bifocale
  - de l'ellipse, 154
  - de l'hyperbole, 154
- Déterminant d'une matrice, 31
- Déterminant d'un endomorphisme, 38
- Déterminant (formule explicite pour le), 92
- Ellipses, 153
- Ellipsoïdes, 158
- Endomorphismes, 2
- Équations différentielles, 73, 75
- Équivalentes (matrices), 13
- Équivalentes (normes), 69
- Espace vectoriel, 2
- Espaces affines, 138
- Espaces caractéristiques, 50
- Espaces caractéristiques (décomposition en), 50
- Espaces propres en somme directe, 44
- Euclidiens
  - espaces vectoriels, 111
  - espaces affines, 147
- Exponentielle d'une matrice, 70
- Familles génératrices, 3

- Familles libres, 4
- Familles liées, 4
- Forme linéaire, 19
- Forme normale des éléments de  $O(n)$ , 176
- Forme polaire
  - d'une forme quadratique, 101
  - d'une forme quadratique hermitienne, 162
- Formes bilinéaires
  - symétriques, 97
  - alternées, 107
- Formes hermitiennes, 162
- Formes quadratiques, 101
- Formes quadratiques hermitiennes, 162
- Formes symplectiques, 108
- Foyer d'une conique, 152
- $GA(\mathcal{E})$ , 140
- $GL_n(k)$ , 10
- $GL(V)$ , 10
- Gram-Schmidt (orthonormalisation de), 120
- Hermitiennes
  - formes, 162
  - matrices, 163
- Hilbertiens (espaces), 169
- Hyperboles, 153
- Hyperboloïdes, 158
- Hyperplan, 44
- Indirectes (isométries), 148
- Inverse (calcul de l'inverse d'une matrice), 26
- Isobarycentre, 141
- Isométries vectorielles, 114, 172
- Isométries d'un espace affine, 147, 148
- Isométries affines
  - du plan, 150
  - de l'espace, 151
- Isomorphismes, 3
- Isotrope (vecteur, cône), 102
- Iwasawa (décomposition d'), 136
- Jordan (matrices de Jordan nilpotentes), 59
- Jordan (matrices de), 61
- Jordan (forme normale de), 59, 62
- Matrice d'une forme bilinéaire
  - symétrique, 98
  - alternée, 108
- Matrice d'une forme hermitienne, 163
- Matrices élémentaires  $E_{ij}$ , 4
- Matrices lignes (et formes linéaires), 20
- Matrice triangulaire (déterminant d'une), 36
- Mineurs d'une matrice, 93, 130
- Module et argument d'un nombre complexe, 161
- Multiplicité algébrique d'une valeur propre, 45
- Multiplicité géométrique d'une valeur propre, 45
- Médiane (égalité de la), 113
- Noether (théorème d'isomorphisme de), 77
- Non dégénérée
  - forme bilinéaire symétrique, 99
  - forme bilinéaire alternée, 108
  - forme hermitienne, 164
- Normal (endomorphisme), 174
- Norme
  - sur un  $\mathbb{K}$ -espace vectoriel ( $\mathbb{K} = \mathbb{R}$  ou  $\mathbb{C}$ ), 69
  - sur un  $\mathbb{R}$ -espace vectoriel, 112
  - sur un  $\mathbb{C}$ -espace vectoriel, 171
  - matricielle, 70
  - euclidienne, 112
  - hilbertienne, 171
- Noyau
  - d'une application linéaire, 6
  - d'une matrice, 9
  - d'une forme quadratique, 102
- $O(n)$ , 115
- Opérations (élémentaires) sur les colonnes, 21, 63
- Opérations (élémentaires) sur les lignes, 23
- Orientation d'un  $\mathbb{R}$ -espace vectoriel de dimension finie, 122
- Orthogonal  $E^\perp$  de  $E$  dans  $V^*$ , 30
- Orthogonal d'un sous-ensemble de  $E$ 
  - pour une forme bilinéaire symétrique, 100
  - pour une forme bilinéaire alternée, 108
- Orthonormées (bases), 112, 170
- $p$ -linéaires (applications et formes), 81
- Paraboles, 152
- Paraboloïdes, 159
- Parallèles (sous-espaces affines), 145
- Parallélogramme (égalité du), 113
- Partie linéaire d'une application affine, 139
- Parties réelle et imaginaire, 161
- Partitions, 58
- Points fixes d'une application affine, 147
- Polarisation (égalité de), 101, 162, 172
- Polynôme caractéristique, 38
- Polynômes d'endomorphismes, 49
- Polynômes en  $n$  indéterminées, 86
- Produit extérieur de formes linéaires, 91
- Produit scalaire euclidien, 111
- Produit scalaire hilbertien, 169
- Produit tensoriel, 96
- Produit vectoriel, 130
- Projection orthogonale, 119, 170
- Projection sur  $\mathcal{F}$  parallèlement à  $\mathcal{F}'$ , 146
- Propres (valeurs, vecteurs, et sous-espaces), 39
- Propriété universelle de  $k[X]$ , 84
- Propriété universelle de  $k[X_1, \dots, X_n]$ , 86
- Propriété universelle des quotients, 77
- Pythagore (égalité de), 113, 172
- Quadriques (en dimension 3), 157–159
- Quotients (espaces vectoriels), 76
- Rang
  - d'une application linéaire, 6
  - (théorème du), 6
  - d'une matrice, 9
  - de  ${}^t A = \text{rang}(A)$ , 13
  - d'une forme bilinéaire symétrique, 99
  - d'une forme quadratique, 102
  - d'une forme bilinéaire alternée, 108
  - d'une forme hermitienne, 164
- Réduction d'une forme quadratique en somme de carrés, 104–107
- Réduction d'une forme hermitienne en somme de carrés de modules, 167–169
- Repère (cartésien) d'un espace affine, 138
- Restriction
  - d'un endomorphisme à un sous-espace stable, 46
  - d'une forme bilinéaire à un sous-espace, 101, 109
- Rotations gauches, 151
- Scindés (polynômes), 47
- Segments, 141
- Semblables (matrices), 14
- Semi-linéaire (application), 162
- Signature d'une permutation, 89
- Signature
  - d'une forme quadratique sur  $\mathbb{R}^n$ , 103
  - d'une forme hermitienne sur  $\mathbb{C}^n$ , 166
- Similitude (classes de), 14
- $SL_n$ , 123
- $SO(n)$ , 123
- Somme directe ( $n$  sous-espaces en), 43
- Somme directe (externe), 52
- Sous-espaces vectoriels, 2
- Sous-espace vectoriel engendré par une partie de  $V$ , 3
- Sous-espaces affines, 142
- Sous-espace affine engendré par une partie de  $\mathcal{E}$ , 144

Stable (sous-espace stable par un endomorphisme), 46  
 Supplémentaires (sous-espaces), 44  
 Supplémentaire (existence d'un), 44  
 Support d'un cycle, 87  
 Sylvester (théorème de), 103, 166  
 Symplectique (forme), 108  
 Symétrie orthogonale, 120  
 Symétrie orthogonale glissée, 150  
 Symétrie par rapport à  $\mathcal{F}$  parallèlement à  $\mathcal{F}'$ , 146  
 Symétries tournées, 151  
 Symétrique (groupe), 87  
 Symétriques  
   formes bilinéaires, 97  
   matrices, 98  
 Trace d'un endomorphisme, 38  
 Translations, 140  
 Transpositions, 87  
 Transposée d'une matrice, 9  
 Transposée d'une application linéaire, 40  
 Triangulaire (inégalité), 112, 171  
 Triangulaire par blocs (matrice), 36  
 Triangulaires strictes (matrices), 57  
 Trigonalisables (endomorphismes), 47  
 Trigonalisation, 47  
 $U(n)$ , 173  
 Vissages, 151