

Algèbre linéaire : un tour d'horizon en 100 pages

Boyer Pascal

3 septembre 2024

Table des matières

1	Espaces vectoriels	2
1.1	Généralités	2
1.2	Théorie de la dimension	3
1.3	Application linéaires.	7
1.4	Rappels sur la dualité	8
2	Matrices	9
2.1	Généralités	9
2.2	Systèmes linéaires.	11
2.3	Décompositions	13
2.4	Calculs matriciels dans un anneau principal	15
2.5	Généralités sur les modules	17
2.6	Théorème de la base adaptée	18
2.7	Exercices	21
3	Réduction des endomorphismes	26
3.1	Matrices équivalentes	27
3.2	Vecteurs propres et espaces propres	30
3.3	Polynôme minimal	33
3.4	Trigonalisation	33
3.5	Noyaux emboîtés	41
3.6	Endomorphismes cycliques	43
3.7	Invariants de similitude	44
3.8	Sous-espaces stables	46
3.9	Exercices	49
4	Algèbre bilinéaire	68
4.1	Formes sesquilinéaires : généralités	68
4.2	Endomorphismes remarquables	71
4.3	Formes quadratiques	72
4.4	Le cas réel	77
4.5	Le cas hermitien	82

4.6	Classes de congruences	84
4.7	Classes de similitudes unitaires	84
4.8	Valeurs propres de matrices hermitiennes	85
4.9	Exercices	91
5	Liste des projets possibles	99

1 Espaces vectoriels

Dans ce qui suit $\mathbb{K}$ est un corps que l'on pourra supposer dans un premier temps égal à $\mathbb{Q}$, $\mathbb{R}$ ou $\mathbb{C}$.

1.1 Généralités

Définition 1. *Un $\mathbb{K}$ -espace vectoriel est un triplet $(E, +, \cdot)$ où*

- *$(E, +)$ est un groupe commutatif,*
- *muni d'une loi externe $(\lambda, e) \in \mathbb{K} \times E \mapsto \lambda \cdot e \in E$ qui vérifie les propriétés suivantes :*
 - *pour tout $\lambda, \mu \in \mathbb{K}$ et pour tout $e \in E$, on a $(\lambda + \mu) \cdot e = \lambda \cdot e + \mu \cdot e$;*
 - *pour tout $\lambda \in \mathbb{K}$ et $e, f \in E$, on a $\lambda \cdot (e + f) = \lambda \cdot e + \lambda \cdot f$;*
 - *pour tout $\lambda, \mu \in \mathbb{K}$ et $e \in E$, on a $(\lambda \mu) \cdot e = \lambda \cdot (\mu \cdot e)$;*
 - *pour tout $e \in E$ on a $1 \cdot e = e$.*

Remarque: ces définitions prennent aussi sens dans le cas où $\mathbb{K}$ est simplement un anneau unitaire A , on parle alors de A -module, cf. le §??.

Exemples :

- $\mathbb{K}$ et plus généralement $\mathbb{K}^n$, $\mathbb{K}^{\mathbb{N}}$ ou $\mathbb{K}^{(\mathbb{N})}$;
- $M_{m,n}(\mathbb{K})$ et $\mathbb{K}[X]$;
- les fonctions de X dans $\mathbb{K}$ où X est un ensemble quelconque ;
- le produit quelconque d'une famille d'espaces vectoriels est un espace vectoriel.

Définition 2. *Soit E un $\mathbb{K}$ -espace vectoriel ; un sous-ensemble $F \subset E$ est un sous-espace vectoriel si et seulement si c'est un sous-groupe stable par la loi externe, i.e. si et seulement si F est non vide et pour tout $f_1, f_2 \in F$ et pour tout $\lambda \in \mathbb{K}$, on a : $f_1 + \lambda f_2 \in F$.*

Exemples :

- $\mathbb{K}_n[X] \subset \mathbb{K}[X]$ le sous-ensemble des polynômes de degré $\leq n$;
- l'ensemble des suites convergentes de $\mathbb{K}^{\mathbb{N}}$;
- $\mathbb{R} \subset \mathbb{C}$ est un sous- $\mathbb{R}$ -espace vectoriel mais n'est pas un sous- $\mathbb{C}$ -espace vectoriel.

Remarque: comme précédemment un sous-espace vectoriel est un espace vectoriel et habituellement on se sert de cette remarque pour tester si on est en présence d'un espace vectoriel.

Remarque: l'intersection quelconque d'une famille de sous-espaces vectoriels est un espace vectoriel ce qui permet de définir le sous-espace vectoriel engendré par un sous-ensemble $A \subset E$ que l'on note $\langle A \rangle$.

Remarque: si $\mathbb{K}$ est un corps infini, toute réunion finie de sous-espaces vectoriels est un sous-espace vectoriel si et seulement s'ils sont tous contenus dans un seul. En particulier une réunion finie d'hyperplans distincts n'est pas un sous-espace vectoriel.

Exemple fondamental : soit $(e_i)_{i \in I}$ une famille *quelconque* d'éléments de E alors $\langle \{e_i : i \in I\} \rangle$ est l'ensemble des combinaisons linéaires $\sum_{i \in I} \lambda_i e_i$ à support fini.

Définition 3. Soient F, G des sous-espaces vectoriels d'un espace vectoriel E . La somme $F + G$ est le sous-espace $\langle F \cup G \rangle$ engendré par F et G . On dit que F et G sont en somme directe et on écrit $F \oplus G$ si $F \cap G = \{0\}$.

Remarque: on vérifie aisément que $F + G = \{f + g : f \in F \text{ et } g \in G\}$; en outre F et G sont en somme directe si et seulement si l'écriture d'un élément $e \in F + G$ sous-la forme $f + g$ est unique.

Définition 4. On dit que F et G sont supplémentaires si $E = F \oplus G$, i.e. si la somme $F + G$ est tout l'espace et qu'ils sont en somme directe.

Remarque: on veillera bien à ne pas confondre *supplémentaires* et *complémentaires*; rappelons que le complémentaire d'un sous-espace vectoriel n'est jamais un sous-espace puisqu'il ne contient pas le vecteur nul!

Exercice : montrer que des sous-espaces $E_1, \dots, E_n$ sont en somme directe si et seulement si pour tout $i = 1, \dots, n$

$$E_i \cap \left(\sum_{1 \leq k \neq i \leq n} E_k \right) = \{0\}.$$

1.2 Théorie de la dimension

Définition 5. Une famille $\{(e_i)_{i \in I}\}$ de vecteurs d'un espace vectoriel E est dite libre si pour tout famille $(\lambda_i)_{i \in I} \in \mathbb{K}^I$

$$\sum_{i \in I} \lambda_i e_i = 0 \Rightarrow \forall i \in I, \lambda_i = 0.$$

Elle est dite génératrice si $\langle \{e_i : i \in I\} \rangle = E$, i.e. si tout vecteur de E peut s'écrire comme une combinaison linéaire à support fini des e_i .

Remarque: la famille $(X^i)_{i \in \mathbb{N}} \in \mathbb{K}[X]$ est libre et génératrice.

Remarque: la famille $(e_i)_{i \in I}$ est dite liée si elle n'est pas libre, i.e. s'il existe une famille $(\lambda_i)_{i \in I} \in \mathbb{K}^{(I)}$ non nulle telle que $\sum_{i \in I} \lambda_i e_i = 0$.

Définition 6. Une famille $(e_i)_{i \in I}$ de vecteurs de E est une base si elle est libre et génératrice.

Théorème 7. *dit de la base incomplète.*

Soient $\{f_1, \dots, f_p\}$ une famille libre de vecteurs et $\{g_1, \dots, g_q\}$ une famille génératrice de E . Il existe alors un entier $n \geq p$ et une base $\{e_1, \dots, e_n\}$ de E telle que $e_i = f_i$ pour $1 \leq i \leq p$ et $e_j \in \{g_1, \dots, g_q\}$ pour $p+1 \leq j \leq n$.

Preuve : Considérons une famille de cardinal maximal de la forme

$$(f_1, \dots, f_p, g_{i_1}, \dots, g_{i_r})$$

qui soit libre. Montrons alors qu'elle est aussi génératrice en vérifiant que pour tout $1 \leq j \leq q$, le vecteur g_j appartient à l'espace vectoriel engendré par cette famille. Si j est l'un des i_k pour $1 \leq k \leq r$, c'est clair sinon, comme par maximalité, la famille $(f_1, \dots, f_p, g_{i_1}, \dots, g_{i_r}, g_j)$ est liée, il existe une famille $(\lambda_i)_{1 \leq i \leq p+r+1}$ non nulle telle que

$$\lambda_1 f_1 + \dots + \lambda_p f_p + \lambda_{p+1} g_{i_1} + \dots + \lambda_{p+r} g_{i_r} + \lambda_{p+r+1} g_j = 0.$$

Comme la famille $(f_1, \dots, f_p, g_{i_1}, \dots, g_{i_r})$ est libre, nécessairement $\lambda_{p+r+1} \neq 0$ et donc $g_j \in \langle f_1, \dots, f_p, g_{i_1}, \dots, g_{i_r} \rangle$. Ainsi donc $(f_1, \dots, f_p, g_{i_1}, \dots, g_{i_r})$ est à la fois libre et génératrice, c'est donc une base.

Remarque: ainsi tout espace contenant une famille génératrice finie admet une base. Autrement dit tout espace vectoriel de type fini admet des bases.

Lemme 8. *Dans un espace possédant une famille génératrice de cardinal n , toute famille de vecteurs non nuls de cardinal $\geq n+1$ est nécessairement liée.*

Preuve : On raisonne par récurrence sur n . Pour $n = 1$ et v une base, pour f, g deux vecteurs non nuls, il existe λ, μ non nuls tels que $f = \lambda v$ et $g = \mu v$, de sorte $\mu f - \lambda g = 0$ et donc (f, g) est liée.

Supposons donc le résultat acquis jusqu'au rang $n-1$ et considérons une famille de $n+1$ vecteurs non nuls $(f_1, \dots, f_{n+1})$ d'un espace vectoriel admettant $(g_1, \dots, g_n)$ comme famille génératrice. Pour tout $i = 1, \dots, n+1$, on écrit alors

$$f_i = \lambda_{i,1} g_1 + \dots + \lambda_{i,n} g_n.$$

Quitte à modifier l'ordre des g_i , supposons $\lambda_{n+1,n} \neq 0$ et on définit pour tout $i = 1, \dots, n$

$$\tilde{f}_i = \lambda_{n+1,n} f_i - \lambda_{i,n} f_{n+1} \in \langle g_1, \dots, g_{n-1} \rangle.$$

D'après l'hypothèse de récurrence la famille $(\tilde{f}_1, \dots, \tilde{f}_n)$ est liée, i.e. il existe une famille non nulle $(\mu_1, \dots, \mu_n)$ telle que

$$\mu_1 \tilde{f}_1 + \dots + \mu_n \tilde{f}_n = 0$$

ce qui fournit la relation

$$\mu_1 \lambda_{n+1,n} f_1 + \dots + \mu_n \lambda_{n+1,n} f_n - \left(\sum_{i=1}^n \mu_i \lambda_{i,n} \right) f_{n+1} = 0$$

prouvant, comme $(\mu_1\lambda_{n+1,n}, \dots, \mu_n\lambda_{n+1,n}, -\sum_{i=1}^n \mu_i\lambda_{i,n})$ n'est pas nulle, que la famille $(f_1, \dots, f_{n+1})$ est liée.

Corollaire 9. *Soit E un espace vectoriel muni d'une base de cardinal n . Alors toute les bases de E ont pour cardinal n .*

Preuve : Soient $(e_1, \dots, e_n)$ et $(f_1, \dots, f_m)$ deux bases de E . En appliquant le lemme précédent à la famille génératrice $(e_1, \dots, e_n)$ (resp. $(f_1, \dots, f_m)$) et à la famille libre $(f_1, \dots, f_m)$ (resp. $(e_1, \dots, e_n)$), on en déduit $m \leq n$ (resp. $n \leq m$) et donc finalement $n = m$.

Définition 10. *Le cardinal d'une base (et donc de toute base) d'un espace vectoriel E est appelé sa dimension ; elle est finie ou infinie.*

Corollaire 11. *Tout sous-espace vectoriel F de E est de dimension inférieure ou égale à celle de E avec égalité si et seulement si $F = E$.*

Preuve : Il suffit de remarquer qu'un base de F est une famille libre de E et d'appliquer le corollaire précédent.

Définition 12. *On appelle hyperplan d'un espace vectoriel E de dimension finie, tout sous-espace de dimension $n - 1$.*

Remarque: en dimension infinie, un hyperplan est un sous-espace tel que E/F est de dimension 1. La dimension de l'espace quotient E/F s'appelle la *codimension* de F dans E .

Remarque: la dimension de $E \times F$ est la somme des dimensions de E et F . L'espace vectoriel des applications linéaires $\mathcal{L}(E, F)$ de E vers F est de dimension $\dim E \cdot \dim F$.

Remarque: toute famille libre est de cardinal $\leq n$ avec égalité si et seulement si c'est une base.

Remarque: la dimension de $F + G$ est inférieure ou égale à $\dim F + \dim G$ avec égalité si et seulement si F et G sont en somme directe. Plus précisément on a la formule du rang.

Théorème 13. *Soient F, G deux sous-espace de E alors*

$$\dim(F + G) = \dim F + \dim G - \dim(F \cap G).$$

Preuve : Soit $(e_1, \dots, e_r)$ une base de $F \cap G$ que l'on complète en une base $(e_1, \dots, e_r, f_1, \dots, f_p)$ (resp. $(e_1, \dots, e_r, g_1, \dots, g_q)$) de F (resp. de G). On vérifie alors aisément que $(e_1, \dots, e_r, f_1, \dots, f_p, g_1, \dots, g_q)$ est une base de $F + G$ ce qui donne la formule de l'énoncé.

Finissons ce paragraphe par un court mot sur la dimension infinie.

Proposition 14. *Soit E un $\mathbb{K}$ -espace vectoriel et V, W_1, W_2 des sous-espaces tels que $V \cap W_1 = \{0\}$ et $V + W_2 = E$. Il existe alors un supplémentaire W de V contenu dans W_2 et contenant W_1 .*

Preuve : Considérons l'ensemble $\mathcal{E}$ des sous-espaces de E contenant W_1 et contenus dans W_2 ; $\mathcal{E}$ n'est pas vide car $W_1 \in \mathcal{E}$. En outre $\mathcal{E}$ est partiellement ordonné par la relation d'inclusion et est inductif. Rappelons que cela signifie que toute chaîne totalement ordonnée admet un majorant : ici pour une telle chaîne, un majorant est simplement donné par la réunion qui est clairement un sous-espace.

D'après le lemme de Zorn, $\mathcal{E}$ admet un élément maximal, notons le W . Par définition on a donc $W \cap V = \{0\}$ et $W_1 \subset W \subset W_2$. Il reste alors à prouver que $V + W = E$; tout élément $x \in E$ s'écrit $x = v + w_2$ avec $v \in V$ et $w_2 \in W_2$. Si $w_2 \in W$ alors c'est gagné, sinon on considère le sous-espace engendré X par W et w_2 . Par maximalité de W , $X \notin \mathcal{E}$ de sorte qu'il existe $0 \neq y \in X \cap V$; ainsi $y = w + \lambda w_2 \in V$ et donc $y \in W \cap V$ ce qui n'est pas. *Remarque:* le lecteur notera bien l'utilisation essentielle du lemme de Zorn qui, rappelons le, est équivalent à l'axiome du choix. Ainsi notre preuve n'est pas du tout constructive.

Corollaire 15. *Tout sous-espace V de E admet un supplémentaire.*

Corollaire 16. *Tout espace vectoriel non nul admet une base.*

Preuve : Considérons l'ensemble $\mathcal{A}$ des familles libres de E ; c'est clairement un ensemble non vide, partiellement ordonné par l'inclusion et inductif. D'après le lemme de Zorn, il possède un élément maximal qui est donc une famille libre maximale c'est donc nécessairement une famille génératrice et donc une base.

Remarque: le lecteur pourra s'exercer sur $\mathbb{K}^{\mathbb{N}}$ en vérifiant que toute base est nécessairement non dénombrable.

Corollaire 17. (Théorème de la base incomplète)

Soit $(e_i)_{i \in I}$ une partie génératrice de E . Soit $J \subset I$ tel que $(e_i)_{i \in J}$ est libre, il existe alors $J \subset K \subset I$ tel que $(e_i)_{i \in K}$ soit une base.

Preuve : On considère l'ensemble $\mathcal{A}$ des familles libres $(e_i)_{i \in A}$ pour $A \subset I$. C'est un ensemble non vide partiellement ordonné par l'inclusion et clairement inductif. D'après le lemme de Zorn, $\mathcal{A}$ possède un élément maximal K ; comme précédemment $(e_i)_{i \in K}$ est libre et génératrice par maximalité de K .

Remarque: citons enfin le cas des espaces de Hilbert, i.e. des espaces hermitiens, au sens du paragraphe sur l'algèbre bilinéaire, qui sont complets, i.e. toutes les suites de Cauchy sont convergentes.

Définition 18. *On dit que $(e_i)_{i \in I}$ est une base de Hilbert d'un espace de Hilbert H si et seulement si :*

- *c'est une base orthonormée, i.e. $\langle e_i, e_j \rangle = \delta_{i,j}$;*
- *la famille est complète au sens que pour tout $x \in H$ il existe $(\lambda_i)_{i \in I}$ telle que $\sum_{i \in I} \lambda_i e_i = x$, i.e. la série correspondante dans H est convergente de limite x .*

Remarque: le lecteur vérifiera aisément qu'une base au sens de Hilbert n'est pas une base au sens classique, cf. par exemple les espaces L^2 .

1.3 Application linéaires.

Définition 19. Une application linéaire ou un morphisme f d'un espace vectoriel E dans un espace F est une application telle que pour tous $\lambda \in \mathbb{K}$ et $x, y \in E$ on a $f(x + \lambda y) = f(x) + \lambda f(y)$.

Remarque: une application linéaire de E dans E est appelée un endomorphisme. Dans le cas où $F = \mathbb{K}$, on parle de *forme linéaire*.

Remarque: pour toute application linéaire $f : E \rightarrow F$ vérifie $f(0) = 0$ et

$$f\left(\sum_{i=1}^n \lambda_i e_i\right) = \sum_{i=1}^n \lambda_i f(e_i).$$

Notation 1. On note $\mathcal{L}(E, F)$ (resp. $\mathcal{L}(E) = \mathcal{L}(E, E)$), l'ensemble des morphismes de E dans F (resp. des endomorphismes de E); c'est un espace vectoriel de dimension $\dim E \cdot \dim F$.

En ce qui concerne l'existence des applications linéaires, on a le résultat suivant.

Proposition 20. Soit $(e_i)_{1 \leq i \leq n}$ une base de E . Pour n'importe quel ensemble de n vecteurs $\{f_1, \dots, f_n\}$ de F , il existe une unique application linéaire telle que pour tout $i = 1, \dots, n$, on ait $f(e_i) = f_i$.

Remarque: ainsi deux applications linéaires sont égales si et seulement si elles coïncident sur une base.

Définition 21. Pour $f \in \mathcal{L}(E, F)$, on note $\text{Ker } f$ l'ensemble des $e \in E$ tels que $f(e) = 0$; c'est un sous-espace vectoriel de E que l'on appelle le noyau de f .

Remarque: l'image de f est aussi un sous-espace de F que l'on note $\text{Im } f$. Plus généralement l'image directe ou réciproque d'un sous-espace est un sous-espace vectoriel.

Proposition 22. Une application linéaire f est injective si et seulement si $\text{Ker } f = \{0\}$.

Remarque: f est surjective si et seulement si l'image d'une base de E est une famille génératrice de F . Ainsi f est bijective, et on dit que f est un *isomorphisme*, si l'image d'une base est une base : c'est alors vrai pour toute base.

Remarque: une application linéaire $f : E \rightarrow F$ où $\dim E = \dim F$ est injective si et seulement si elle est surjective.

Notation 2. On note $GL(E)$ l'ensemble des isomorphismes de E , on dit aussi automorphisme. C'est un groupe pour la composition.

Remarque: pour $E = \mathbb{K}^n$ les vecteurs $e_i = (0, \dots, 0, 1, 0, \dots, 0)$ définissent une base dite *canonique*. Tout espace vectoriel muni d'une base $(e_i)_{1 \leq i \leq n}$ de cardinal n est isomorphe à $\mathbb{K}^n$ où $f : \mathbb{K}^n \rightarrow E$ est défini par $f(x_1, \dots, x_n) = \sum_{i=1}^n x_i e_i$. En particulier deux espaces vectoriels de même dimension sont toujours isomorphes.

Théorème 23. Soit $f \in \mathcal{L}(E, F)$ alors

$$\dim E = \dim \text{Ker } f + \dim \text{Im}(f).$$

Définition 24. La dimension de $\text{Im } f$ s'appelle le rang de f ; on le note $\text{rg } f$.

1.4 Rappels sur la dualité

Définition 25. Étant donné un espace vectoriel E , l'ensemble des formes linéaires sur E est un espace vectoriel noté E^* et dit le dual de E .

Remarque: une base $(e_i)_{1 \leq i \leq n}$ de E étant fixée, l'application linéaire $e_i^* \in E^*$ définie par $e_i^*(e_j) = \delta_{i,j}$ est une base de E^* dite la base duale de $(e_i)_i$. On se méfiera de la notation car e_i^* dépend de toute la base $(e_i)_i$ et pas seulement du seul vecteur e_i .

Proposition 26. Étant donné un sous-espace $F \subset E$, le sous-ensemble $F^\perp \subset E^*$ des formes linéaires s'annulant sur F est un sous-espace de dimension $\dim E - \dim F$, i.e. la dimension de $F^\perp$ est égale à la codimension de F .

Définition 27. Soit $f \in \mathcal{L}(E, F)$, on lui associe alors son application adjointe notée $f^* \in \mathcal{L}(F^*, E^*)$ définie par la formule

$$y^* \in F^* \mapsto f^*(y^*) = y^* \circ f$$

au sens où $f^*(y^*)$ est la forme linéaire sur E définie par $x \mapsto y^*(f(x))$.

Remarque: on notera en particulier que f et f^* ont le même rang.

Proposition 28. Soit E un espace vectoriel de dimension finie; alors le bidual $(E^*)^*$ s'identifie canoniquement à E .

Remarque: l'application $E \rightarrow (E^*)^*$ est donnée par $x \mapsto (f \mapsto f(x))$.

2 Matrices

2.1 Généralités

Définition 29. Une matrice à coefficients dans $\mathbb{K}$ de taille $m \times n$ est un tableau $(a_{i,j})_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}}$ de scalaires $a_{i,j} \in \mathbb{K}$ placés sur la i -ème ligne et la j -ème colonne. On note $\mathbb{M}_{m,n}(\mathbb{K})$ l'ensemble de ces matrices que l'on muni d'une structure d'espace vectoriel en l'identifiant avec $\mathbb{K}^{nm}$, i.e. coefficient par coefficient.

Remarque: une matrice ligne (resp. colonne) correspond au cas où $n = 1$ (resp. $m = 1$); on dit aussi vecteur ligne (resp. colonne). Les lignes (resp. les colonnes) d'une matrice sont appelées ses vecteurs lignes (resp. colonnes).

Remarque: les matrices $E_{i,j}$ dont les coefficients sont tous nuls sauf celui d'indice (i,j) égal à 1, forment une base de $\mathbb{M}_{m,n}(\mathbb{K})$.

Remarque: la matrice $(b_{i,j} = a_{j,i})_{i,j} \in \mathbb{M}_{n,m}(\mathbb{K})$ s'appelle la matrice transposée, on la note $B = {}^tA$ si $A = (a_{i,j})_{i,j}$.

Remarque: dans le cas où $m = n$, on parle de matrices carrées et on note $\mathbb{M}_n(\mathbb{K})$ pour $\mathbb{M}_{n,n}(\mathbb{K})$. Les éléments $a_{i,i}$ de $A = (a_{i,j})_{i,j}$ sont dits *diagonaux*. Ainsi une matrice est dite :

- *diagonale* si tous ses coefficients diagonaux sont nuls; on parle aussi de matrice *antidiagonale* si $a_{i,j} = 0$ sauf pour $i + j = n + 1$.
- *triangulaire supérieure* (resp. *inférieure*) si tous les $a_{i,j}$ sont nuls pour $i > j$ (resp. $j > i$).
- *tridiagonale* si $a_{i,j} = 0$ pour tout $|j - i| > 1$.

Les matrices ne sont pas de simples tableaux de chiffres mais doivent leur introduction en ce qu'ils permettent :

- d'étudier les systèmes linéaires;
- de manipuler les endomorphismes des espaces vectoriels.

Ainsi pour $f : E \rightarrow F$ un endomorphisme entre deux espaces vectoriels munis des bases respectives $(e_i)_{1 \leq i \leq n}$ et $(f_j)_{1 \leq j \leq m}$, on lui associe la matrice $A(f) = (a_{i,j})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq m}}$ telle que pour tout $1 \leq i \leq n$ on a

$$f(e_i) = \sum_{j=1}^m a_{i,j} f_j.$$

Autrement dit les vecteurs colonnes de A sont les $f(e_i)$ exprimés dans la base $(f_j)_j$.

Remarque: d'après ce qui précède, f est déterminé par sa matrice $A(f)$ de sorte que l'on doit pouvoir exprimer l'image $f(x)$ de tout vecteur $x = \sum_{i=1}^n x_i e_i$.

Définition 30. Pour toute matrice $A \in \mathbb{M}_{m,n}(\mathbb{K})$ et tout vecteur colonne $X \in \mathbb{M}_{n,1}(\mathbb{K})$ on définit le vecteur colonne $Y = AX \in \mathbb{M}_{m,1}(\mathbb{K})$ par la

formule :

$$y_j = \sum_{k=1}^n a_{j,k} x_k.$$

Pour une matrice $B \in \mathbb{M}_{n,r}$ dont on note $C_1, \dots, C_r$ les vecteurs colonnes, on définit la matrice $M = AB \in \mathbb{M}_{m,r}(\mathbb{K})$ dont les vecteurs colonnes sont les AC_i pour $i = 1, \dots, r$.

Proposition 31. Soit $f : E \rightarrow F$ et $A(f)$ sa matrice relativement à des bases $(e_i)_{1 \leq i \leq n}$ et $(f_j)_{1 \leq j \leq m}$ de respectivement E et F . Pour tout $x = \sum_{i=1}^n x_i e_i$, on note X le vecteur colonne $(x_i)_{1 \leq i \leq n}$. Alors les coordonnées de $f(x)$ dans la base $(f_j)_{1 \leq j \leq m}$ sont les coordonnées $(y_j)_{1 \leq j \leq m}$ du vecteur colonne $A(f)X$, i.e. $f(x) = \sum_{j=1}^m y_j f_j$.

Corollaire 32. Pour tout $f : E \rightarrow F$ et $g : F \rightarrow G$ des endomorphismes ; on suppose E, F, G munis de base $(e_i)_{1 \leq i \leq n}$, $(f_j)_{1 \leq j \leq m}$ et $(g_k)_{1 \leq k \leq r}$. On note $A(f), A(g)$ et $A(g \circ f)$ les matrices associées à f, g et $g \circ f$ relativement à ces bases. On a alors

$$A(g \circ f) = A(g)A(f).$$

Proposition 33. Si E, F sont munies de bases respectives $(e_i)_i$ et $(f_j)_j$ alors la matrice de f^* dans les bases duales associées de F^* et E^* est la transposée de la matrice de f dans les bases $(e_i)_i$ et $(f_j)_j$.

En particulier $\mathcal{L}(E)$ étant une algèbre on en déduit le corollaire suivant.

Corollaire 34. La multiplication des matrices définie plus haut, munit $\mathbb{M}_n(\mathbb{K})$ d'une structure d'algèbre.

Définition 35. Les matrices de $\mathbb{M}_n(\mathbb{K})$ qui s'identifient aux automorphismes de E sont dites inversibles ; l'ensemble de ces matrices inversibles est noté $GL_n(\mathbb{K})$.

Remarque: ainsi une matrice est inversible si et seulement si ses vecteurs colonnes forment une base.

Définition 36. Étant donné un espace vectoriel E muni de deux bases $(e_i)_{1 \leq i \leq n}$ et $(e'_i)_{1 \leq i \leq n}$, on appelle matrice de passage de $(e_i)_i$ à $(e'_i)_i$ et on la note $P_{e_i \leftarrow e'_i}$, la matrice de $\mathbb{M}_n(\mathbb{K})$ dont la j -ème colonne est donnée par les coordonnées de e'_j dans la base $(e_i)_i$, i.e. $e'_j = \sum_{i=1}^n p_{i,j} e_i$.

Remarque: la matrice $P_{e_i \leftarrow e'_i}$ peut aussi se voir comme la matrice de l'application de l'identité de $E \rightarrow E$ où l'espace de départ est muni de la base $(e_i)_i$ et l'espace d'arrivée de la base $(e'_i)_i$. On en déduit alors que :

- $P_{e_i \leftarrow e'_i}$ est inversible, d'inverse $P_{e'_i \leftarrow e_i}$;
- si X' est le vecteur colonne des coordonnées d'un vecteur e de E dans la base $(e'_i)_i$, alors $X = P_{e_i \leftarrow e'_i} X'$ est celui de e dans la base $(e_i)_i$;

- si $A(f)$ est la matrice de $f : E \rightarrow F$ muni des bases $(e_i)_i$ et $(f_j)_j$ de respectivement E et F alors, pour des bases $(e'_i)_i$ et $(f'_j)_j$, la matrice $A'(f)$ relativement à ces bases est $P_{e_i \leftarrow e'_i}^{-1} A(f) P_{f_j \leftarrow f'_j}$. Dans le cas particulier où $E = F$ et où $A(f)$ et $A'(f)$ représentent la matrice de f dans respectivement les bases $(e_i)_i$ et $(e'_i)_i$ alors $A'(f) = P_{e_i \leftarrow e'_i}^{-1} A(f) P_{e_i \leftarrow e'_i}$.

Exemples : étant donnée une matrice $A \in \mathbb{M}_{m,n}(\mathbb{K})$, la multiplication à gauche (resp. à droite) par la matrice

- $T_{i,j}(\lambda)$ dont les coefficients diagonaux sont égaux à 1, tous les autres étant nuls sauf $t_{i,j} = \lambda$, correspond à modifier les lignes (resp. les colonnes) de A selon la règle $L_i \rightarrow L_i + \lambda L_j$ (resp. $C_i \rightarrow C_i + \lambda C_j$);
- $D_i(\lambda)$ matrice diagonale dont les coefficients diagonaux sont égaux à 1 sauf $d_{i,i} = \lambda$ correspond à modifier les lignes (resp. les colonnes) de A selon la règle $L_i \rightarrow \lambda L_i$ (resp. $C_i \rightarrow \lambda C_i$);
- $P_{i,j} = I - E_{i,i} - E_{j,j} + E_{i,j} + E_{j,i}$, correspond à modifier les lignes (resp. les colonnes) de A selon la règle $L_i \leftrightarrow L_j$ (resp. $C_i \leftrightarrow C_j$).

Remarque: pour $i \neq j$, les matrices $T_{i,j}(\lambda)$ (resp. $D_i(\lambda)$) sont des matrices dites de transvections (resp. de dilatations) élémentaires relativement à la base canonique. Les matrices $P_{i,j}$ sont des cas particuliers des matrices de permutation. Ces trois types de matrices permettent d'effectuer les opérations élémentaires sur les lignes et les colonnes d'une matrice. Nous reviendrons sur ce point lors de l'étude des systèmes linéaires.

2.2 Systèmes linéaires.

Définition 37. Une équation linéaire en les variables $x_1, \dots, x_n$ est une expression de la forme

$$L(x_1, \dots, x_n) = b \text{ où } L(x_1, \dots, x_n) = a_1 x_1 + \dots + a_n x_n$$

dont on cherche les solutions dans $\mathbb{K}^n$. On dit que l'équation est homogène lorsque $b = 0$.

Remarque: on peut et on doit interpréter $L(x_1, \dots, x_n)$ comme une forme linéaire sur $\mathbb{K}^n$ écrite dans la base canonique.

Définition 38. Un système linéaire de m équations à n variables est une collection de m équations linéaires

$$(S) = \begin{cases} a_{1,1}x_1 + \dots + a_{1,n}x_n = b_1 \\ a_{2,1}x_1 + \dots + a_{2,n}x_n = b_2 \\ \vdots \\ a_{m,1}x_1 + \dots + a_{m,n}x_n = b_m \end{cases}$$

que l'on cherche à résoudre simultanément. Il est dit incompatible s'il ne possède pas de solutions, compatible sinon.

Remarque: comme suggéré par les notations, on introduit la matrice $A_S = (a_{i,j}) \in \mathbb{M}_{m,n}(\mathbb{K})$ et on écrit le système précédent sous la forme $A_S X = B$ où X (resp. B) est le vecteur colonne de coordonnées les x_i (resp. les b_i).

Définition 39. Deux systèmes linéaires (S) et (S') sont dit équivalents s'ils ont le même ensemble de solutions.

Proposition 40. Deux systèmes linéaires (S) et (S') sont équivalents si et seulement s'il existe une matrice inversible $P \in GL_m(\mathbb{K})$ telle que $A_S = PA_{S'}$ et $B = PB'$.

Remarque: en utilisant que $GL_n(\mathbb{K})$ est engendré par les matrices de transvections et de dilatations (en général, par commodité, on rajoute aussi les matrices de permutation $P_{i,j}$), on doit pouvoir manipuler le système (S) pour arriver au système (S') qui lui est équivalent, encore faut-il que ce processus soit constructif, ce qui est assuré par l'algorithme de Gauss.

Définition 41. Soit (S) un système linéaire non nécessairement homogène que l'on écrit sous forme matricielle $A_S X = B$. On introduit alors la matrice $\tilde{A}_S$ en rajoutant la colonne B à la matrice A_S .

Définition 42. Une matrice $M \in \mathbb{M}_{m,n}(\mathbb{K})$ est dite échelonnée si en dessous du premier élément non nul de chaque ligne, il n'y a que des zéros. Elle est dite en outre échelonnée réduite si tout premier élément non nul de chaque ligne, appelé pivot, est égal à 1, et que chaque pivot est le seul élément non nul de sa colonne.

Proposition 43. Pour toute matrice $M \in \mathbb{M}_{m,n}(\mathbb{K})$, il existe une unique matrice $P \in GL_m(\mathbb{K})$ telle que PM est échelonnée réduite.

Remarque: la mise en place pratique de ce résultat est ce que l'on appelle, **l'algorithme de Gauss**.

Ainsi étant donné un système linéaire (S) de matrice augmentée $\tilde{A}_S$, on lui applique l'algorithme de Gauss pour obtenir l'échelonnée réduite associée, par exemple de la forme

$$\begin{pmatrix} 0 & \dots & 1 & \bullet & \dots & \bullet & \dots & \bullet & \bullet & \bullet & \bullet & \bullet \\ 0 & \dots & 0 & 0 & \dots & 1 & \bullet & 0 & 0 & \bullet & 0 & \bullet \\ 0 & \dots & 0 & 0 & \dots & 0 & 0 & 1 & 0 & \bullet & 0 & \bullet \\ 0 & \dots & 0 & 0 & \dots & 0 & 0 & 0 & 1 & \bullet & 0 & \bullet \\ 0 & \dots & 0 & 0 & \dots & 0 & 0 & 0 & 0 & 0 & 1 & \bullet \\ 0 & \dots & 0 & 0 & \dots & 0 & 0 & 0 & 0 & 0 & 0 & \alpha \\ 0 & \dots & 0 & 0 & \dots & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

- Si la dernière colonne contient un pivot, alors le système est incompatible, ce qui dans l'exemple précédent correspond au cas $\alpha \neq 0$.

- Sinon, les positions des pivots fournissent les indices des variables dites *principales* alors que les autres sont dites *secondaires*. L'ensemble des solutions est alors un sous-espace affine de dimension le nombre de variables secondaires ; autrement dit quelles que soient les valeurs prises par ces variables secondaires, on obtient une unique solution pour les variables principales que l'on obtient en résolvant ce système du bas vers le haut.

Définition 44. *Le système (S) est dit de Cramer s'il possède une unique solution.*

Remarque: autrement dit, (S) est de Cramer s'il est compatible sans variable secondaire, ce qui permet de prouver la proposition suivante.

Proposition 45. *(S) est de Cramer si et seulement si A_S est une matrice inversible, ce qui impose en particulier que $m = n$.*

Remarque: on utilise des systèmes linéaires et leur résolution via l'algorithme de Gauss, par exemple pour trouver l'inverse d'une matrice, pour donner des équations linéaires d'un sous-espace dont on connaît une famille génératrice...

2.3 Décompositions

L'algorithme de Gauss fournit les résultats suivants :

- (a) $SL_n(\mathbb{K})$ est engendré par les matrices de transvections élémentaires.

En dimension supérieure ou égale à 3, les transvections sont toutes conjuguées ; on en déduit la simplicité de $SL_n(K)$. En dimension 2, pour (e_1, e_2) une base on note τ_λ la transvection définie par $\tau_\lambda(x_1e_1 + x_2e_2) = (x_1e_1 + x_2e_2) + \lambda x_2e_1$; toute transvection est conjuguée à un τ_λ et τ_λ, τ_μ sont conjuguées si et seulement si $\frac{\lambda}{\mu} \in (K^\times)^2$.

On en déduit alors que $GL_n(\mathbb{K})$ est engendré par les matrices de transvections élémentaires et les matrices de dilatation $D_n(\lambda) = \text{diag}(1, \dots, 1, \lambda)$ où $\lambda \in \mathbb{K}^\times$.

- (b) **Décomposition de Bruhat** : $T(n, \mathbb{C}) \backslash GL(n, \mathbb{C}) / T(n, \mathbb{C}) \simeq \mathfrak{S}_n$ où $T(n, \mathbb{C})$ désigne l'ensemble des matrices triangulaires supérieures. On interprète ce résultat en termes de drapeaux : l'ensemble des classes de paires de drapeaux complets sous l'action de $GL(n, \mathbb{C})$, est en bijection avec $\mathfrak{S}_n$.

- (c) **Décomposition LU** : on se demande, dans la méthode du pivot, si pour pivoter on a besoin d'échanger des lignes (en général oui surtout pour minimiser les erreurs d'arrondis) : la condition pour que ce ne soit pas nécessaire est que $\det M^{(k)}$ soit inversible pour tout $1 \leq k \leq n$ où $M^{(k)}$ désigne le mineur principal d'ordre k : la factorisation est alors unique si on impose à la diagonale de L d'être égale à 1. Pour la résolution des systèmes linéaires : on résoud $LY = B$ puis $UX = Y$ (on notera que le calcul de L

est particulièrement simple à partir de matrices $T_{i,j}$ cf. Ciarlet p.83). On pourra mentionner le cas particulièrement simple des matrices tridiagonales (Ciarlet p.85) (qui nécessite $8n - 6$ opérations) ou plus généralement des matrices bandes, dont la décomposition LU est encore de la même forme (notion de structure de données statique)

(d) Soit B une matrice de $\mathbb{M}_{m,n}(\mathbb{R})$ et $c \in \mathbb{R}^m$; on cherche u tel que la norme euclidienne de $Bu - c$ soit minimale (méthode des moindres carrés). On introduit la fonctionnelle

$$2J(v) := \|Bv - c\|^2 - \|c\|^2 = ({}^tBBv, v) - 2({}^tBc, v)$$

que l'on cherche à minimiser. La matrice symétrique tBB étant positive, la fonction J est convexe de sorte que l'ensemble des solutions coïncide avec celui de l'équation

$$J'(u) = {}^tBBu - {}^tBc = 0$$

On est ainsi naturellement amené à résoudre des équations $AX = B$ avec A est symétrique définie positive où on applique **la décomposition de Cholesky** : $A = B^tB$ avec B triangulaire inférieure (La décomposition LU donne $A = LDV$ avec, par unicité, $V = {}^tL$ et on prend la racine carrée de D .)

(e) Toute matrice est algorithmiquement semblable à **une matrice de Hessenberg** : dans la méthode QR si A est de Hessenberg alors tous les A_k aussi ce qui raccourcit les temps de calcul.

(f) **La factorisation QR** où Q est unitaire et R triangulaire supérieure. Si on impose aux coefficients diagonaux de R d'être positifs ou nuls et si A est inversible, alors la factorisation $A = QR$ est unique. (sur $\mathbb{R}$, Q est orthogonale). Pour les construire on peut utiliser le procédé d'orthonormalisation de Gram-Schmidt, mais il est à éviter car il engendre des erreurs d'arrondis. L'utilisation des matrices de Householder $H_u = I_n - 2 \frac{u {}^tu}{\|u\|^2}$ est bien plus efficace.

(g) Etant donnée une matrice symétrique A , il existe une matrice P produit de $(n - 2)$ matrices de Householder, telle que tPAP soit tridiagonale (Ciarlet p.120). En appliquant alors la méthode de Givens, on obtient des valeurs approchées des valeurs propres : les polynômes caractéristiques des mineurs principaux forment une suite de Sturm ce qui permet de localiser les racines aussi précisément que l'on veut par exemple par dichotomie (Ciarlet p.123).

(h) Méthode QR de localisation des racines : A inversible à valeurs propres toutes de modules distinctes telle qu'il existe une matrice de passage à une base de diagonalisation admettant une décomposition LU , on construit alors une suite A_k définie par récurrence $A_k = Q_k R_k = R_{k-1} Q_{k-1}$. Alors A_k converge vers la matrice diagonale des valeurs propres de A rangées par ordre décroissant de leur module.

2.4 Calculs matriciels dans un anneau principal

Rappelons que $\mathbb{M}_n(A)$ désigne l'ensemble des matrices carré de taille n à coefficients dans A . Dans ce paragraphe A est un anneau principal (idéalement euclidien comme $\mathbb{Z}$ ou $\mathbb{K}[X]$).

Lemme 46. *La matrice $M \in \mathbb{M}_n(A)$ est inversible si et seulement si $\det M \in A^\times$.*

Preuve : Si la matrice M est inversible, alors en appliquant le déterminant à l'égalité $M.M^{-1} = I_n$ on obtient que l'inverse de $\det M$ est $\det(M^{-1})$. Inversement, notons $\tilde{M}$ la transposée de la matrice des cofacteurs de M . De l'égalité $\tilde{M}.M = \det M$ on en déduit que si $\det M \in A^\times$ alors $(\det M)^{-1}\tilde{M}$ est l'inverse de M .

Le lemme suivant est l'ingrédient calculatoire essentiel pour les calculs matriciels de la suite.

Lemme 47. *Soient $x, y \in A$ non tous deux nuls, z un pgcd de x et $ux + vy = z$ une relation de Bézout. Alors la matrice $M := \begin{pmatrix} u & v \\ -y/z & x/z \end{pmatrix}$ de déterminant 1 vérifie l'équation*

$$M \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} z \\ 0 \end{pmatrix}.$$

Nous utiliserons la matrice de taille (n, n) suivante :

$$L_{j,k}(x, y) = \begin{pmatrix} 1 & 0 & \dots & & & & \dots & 0 \\ 0 & \ddots & & & & & 0 & \vdots \\ & & 1 & & & & & \\ \vdots & 0 & \dots & u & 0 & \dots & v & \vdots \\ & \vdots & & 1 & & & & \\ & & & & \ddots & & & \\ & & & & & 1 & & \\ & & & -y/z & 0 & \dots & x/z & \\ & & & & & & 1 & \\ & & & & & & & \ddots \\ 0 & \dots & & & & & & 1 \end{pmatrix}$$

u étant en (j, j) , v en (j, k) , $-y/z$ en (k, j) et x/z en (k, k) .

Remarque: comme d'habitude, en notant l_i la ligne d'indice i d'une matrice M , la multiplication à gauche d'une matrice $M \in \mathbb{M}_{n,m}(A)$ par la matrice $L_{j,k}(x, y)$ remplace l_j et l_k par respectivement $\alpha l_j + \beta l_k$ et $\gamma l_j + \delta l_k$.

Proposition 48. *Soit $M \in \mathbb{M}_{n,m}(A)$. Il existe alors une matrice $L \in SL_n(A)$ telle que LM soit triangulaire supérieure.*

Preuve : La démonstration consiste à appliquer plusieurs fois le lemme 47 pour faire apparaître des zéros sous la diagonale principale.

a) Soit $M = (a_{ij})$ ($1 \leq i \leq n$, $1 \leq j \leq m$). Multiplions M à gauche par la matrice $L_1 = L_{1,2}(a_{1,1}, a_{2,1})$ de sorte que la première colonne de $M_1 = L_1 M$ commence par $\begin{pmatrix} d \\ 0 \end{pmatrix}$ avec $d = a_{11} \wedge a_{21}$.

b) On multiplie ensuite M_1 à gauche par une matrice $L_2 = L_{1,3}(d, a_{3,1})$ de façon à faire apparaître un zéro à la place $(3, 1)$ et à remplacer d par $d_1 = a_{1,1} \wedge a_{2,1} \wedge a_{3,1}$. On continue ainsi jusqu'à ce que l'on obtienne la matrice $M_{n-1} = L_{n-1} \cdots L_1 M$ dont la première colonne est $(d_{n-1}, 0 \dots 0)$ avec d_{n-1} le pgcd des éléments de la première colonne de M .

c) On continue de même avec la seconde colonne, en commençant par multiplier à gauche par une matrice $L_{2,3}$ de façon à laisser la première ligne inchangée et à ne manipuler que les lignes $l_2, \dots, l_n$. En procédant ainsi on obtient une deuxième colonne de la forme $(a, b, 0, \dots, 0)$. En continuant ainsi pour toutes les colonnes, on obtient une matrice triangulaire supérieure.

Définition 49. Une matrice $M \in \mathbb{M}_{n,m}(A)$ est dite réduite si

$$M = \begin{pmatrix} a_{1,1} & 0 & \dots & & 0 \\ 0 & a_{2,2} & 0 & \dots & \vdots \\ \vdots & & \ddots & & \\ & & & a_{n,n} & \dots & 0 \end{pmatrix}$$

avec

$$a_{i,i} \mid a_{i+1,i+1}, \quad 1 \leq i \leq \inf(n, m) - 1.$$

Remarque: on a représenté une matrice M avec $n < m$. Il est à noter que les derniers $a_{i,i}$ peuvent être nuls et que tous les éléments non sur la diagonale sont nuls.

Théorème 50. Soit $M \in \mathbb{M}_{n,m}(A)$. Il existe alors $L \in SL_n A$ et $R \in SL_m(A)$ telles que $M' = LMR$ soit réduite.

Remarque: l'énoncé analogue sur un corps K est que toute matrice $M \in \mathbb{M}_{n,m}(K)$ est équivalente à une matrice de la forme $M' = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$.

Preuve : Nous avons vu comment en manipulant les lignes, on faisait remonter le pgcd de chaque colonne sur la première ligne. On commence donc par faire cela pour chacune des colonnes. Puis, comme on a opéré sur les lignes, on opère sur les colonnes de sorte à ramener le pgcd de la première ligne, et donc celui de tous les coefficients de la matrice, en position $(1, 1)$.

À ce stade, le coefficient $a_{1,1}$ est le pgcd de tous les $a_{i,j}$. On recommence alors à opérer sur les lignes de façon à obtenir la première colonne égale à $(a_{1,1}, 0, \dots, 0)$: on note en particulier que la première ligne n'est pas

modifiée. On opère ensuite de même sur les colonnes pour que la première ligne soit égale à $(a_{1,1}, 0, \dots, 0)$. Comme précédemment, on ne modifie pas la première ligne de sorte qu'à ce stade la matrice est diagonale par blocs avec un premier bloc de taille 1 et le deuxième de taille $(n-1, m-1)$.

On conclut alors par récurrence.

2.5 Généralités sur les modules

Dans ce paragraphe A est un anneau commutatif quelconque.

Définition 51. - *Un A -module est un groupe commutatif $(M, +)$ muni d'une application $A \times M \rightarrow M$, où l'on note ax l'image de (a, x) , telle que :*

1. $\forall a \in A$ et $x, y \in M$, $a(x + y) = ax + ay$;
2. $\forall a, b \in A$ et $x \in M$, $(a + b)x = ax + bx$;
3. $\forall a, b \in A$ et $x \in M$, $1x = x$ et $a(bx) = (ab)x$.

- *Un sous-module d'un A -module M est un sous-groupe N de M stable par l'action de A .*

- *Un morphisme de A -modules $M \rightarrow N$ est un morphisme des groupes additifs $(M, +) \rightarrow (N, +)$ qui est de plus A -linéaire.*

Remarque: la notion de A -module est formellement identique à celle de K -espace vectoriel sauf que l'action externe est relative à un anneau A plutôt qu'à un corps K .

Exemples : les constructions habituelles sur les espaces vectoriels se généralisent au cas des A -modules (quotient, somme, intersection, A -module engendré...).

On utilisera dans la suite plus spécifiquement les exemples suivants.

- Si $(G, +)$ est un groupe commutatif, il est canoniquement muni d'une structure de $\mathbb{Z}$ -module, en définissant, pour $n \geq 0$, ng comme $\overbrace{g + g + \dots + g}^n$, et $(-1)g$ comme $-g$.

- Si V est un K -espace vectoriel et $u \in \mathcal{L}(V)$ un endomorphisme de V , on munit V d'une structure de $K[X]$ -module en posant pour tout $P \in K[X]$ et pour tout $\vec{v} \in V$, $P \cdot \vec{v} := P(u)(\vec{v})$.

- L'anneau A est lui-même un A -module. Les sous- A -modules de A sont ses idéaux.

Définition 52. *Un sous-ensemble S d'un A -module M est dit*

- *libre si l'égalité $\sum_{s \in S} a_s s = 0$ où la famille $(a_s)_{s \in S}$ est supposée à support fini, implique que pour tout $s \in S$, on a $a_s = 0$.*
- *générateur si tout élément $m \in M$ peut s'écrire sous la forme $\sum_{s \in S} a_s s$ où la famille $(a_s)_{s \in S}$ est à support fini.*
- *une base si S est à la fois libre et générateur.*

On dit que M est

- *libre si M admet une base.*

- de type fini s'il admet un sous-ensemble fini S générateur.
- de torsion si l'ensemble des éléments $\lambda \in A$ qui annulent M i.e. tels que $\forall m \in M$ on ait $\lambda m = 0$, est un idéal non nul de A . Cet idéal est appelé annulateur de M et noté $\text{Ann}(M)$.

Exemple : l'annulateur de $M = A/I$ est l'idéal I .

Remarque : la donnée d'une base d'un A -module libre de type fini est équivalent à la donnée d'un isomorphisme $A^n \longrightarrow M$.

Proposition 53. *Soit M un A -module libre de type fini. Alors toutes ses bases ont le même cardinal.*

Preuve : Soient $(e_1, \dots, e_n)$ une base de M et $\mathcal{M} \in A$ un idéal maximal de sorte que le quotient $A/\mathcal{M}$ est un corps k . On note $\mathcal{M}M = \{\sum_{i=1}^n m_i e_i : m_i \in \mathcal{M}\}$ de sorte que $V := M/\mathcal{M}M$ est un k -espace vectoriel, i.e. un groupe muni d'une action externe de $A/\mathcal{M}$. Notons par ailleurs que $(\bar{e}_i)_{i=1, \dots, n}$ est une base de V . C'est clairement une famille génératrice. Pour la liberté, $\sum_{i=1}^n \lambda_i \bar{e}_i = 0$ s'écrit aussi $\sum_{i=1}^n \mu_i e_i = \sum_{i=1}^n m_i e_i$ où $\bar{\mu}_i = \lambda_i$ et les $m_i \in \mathcal{M}$. La famille $(e_i)_{i=1, \dots, n}$ étant libre, on en déduit que $\mu_i = m_i$ et donc $\lambda_i = \bar{\mu}_i = 0$.

Ainsi n est la dimension de l'espace vectoriel $M/\mathcal{M}M$ et est donc le cardinal de toute base du A -module M .

Proposition 54. *Un sous-module d'un module de type fini est de type fini.*

Preuve : Soit $f : A^n \twoheadrightarrow M$ définie par la donnée d'une famille génératrice de cardinal n de M . Pour un sous-module N de M , il suffit de montrer que $f^{-1}(N)$ est de type fini, i.e. on est ramené au cas où $M = A^n$. On raisonne alors par récurrence sur n : dans le cas $n = 1$, un sous-module de A est un idéal qui est donc principal et donc libre de rang 1.

Supposons alors le résultat acquis jusqu'au rang $n - 1$ et traitons le cas de n . Considérons alors l'application $g : N \hookrightarrow A^n \twoheadrightarrow A$ où la deuxième flèche est donnée par la première projection $(a_1, \dots, a_n) \mapsto a_1$. L'image de g est de la forme $a_1 A$ et notons $n_1 \in N$ un antécédent puis $N' = N \cap A^{n-1}$ où A^{n-1} est le noyau de la première projection. Notons que $n_1 A \cap N' = (0)$ puisque si $g(\lambda n_1) = 0$ alors $\lambda = 0$. En outre pour $n \in N$, on peut écrire $n = \lambda n_1 + (n - \lambda n_1)$ où $f(n) = \lambda a_1 = f(\lambda n_1)$ et donc $n - \lambda n_1 \in N'$. Autrement dit on a $N = A n_1 \oplus N'$ avec $N' \subset A^{n-1}$. Par récurrence N' est de type fini et donc N aussi.

2.6 Théorème de la base adaptée

Théorème 55. *Soit N un sous-module d'un A -module L libre de type fini. Alors N est un sous-module libre de type fini et il existe une base, dite adaptée, $(f_1, \dots, f_n)$ de L ainsi que des éléments $a_i \in A$, $1 \leq i \leq n$ tels*

que :

$$\begin{cases} a_1 \mid a_2 \mid \cdots \mid a_n, \\ \text{les } (a_i f_i) \text{ tels que } a_i \neq 0 \text{ forment une base de } N. \end{cases}$$

De plus, la suite des idéaux (a_i) satisfaisant ces conditions est unique.

Preuve : D'après la proposition 54, N est de type fini, notons alors $(g_1, \dots, g_m)$ une famille génératrice de N et écrivons la matrice de passage M des g_i pour $1 \leq i \leq m$ dans une base $(e_1, \dots, e_n)$ de L . D'après 50, il existe $P \in SL_n A$ et $Q \in SL_m(A)$ telles que $M' = PMQ$ soit réduite avec des éléments $a_{i,i}$ sur la diagonale que l'on note simplement a_i .

La matrice P (resp. Q) s'interprète comme une matrice de changement de base de L (resp. de changement de famille génératrice de N). Notons $(f_1, \dots, f_n)$ la nouvelle base de L ; l'écriture matricielle de M' s'interprète alors en disant que $(a_1 f_1, \dots, a_r f_r)$ est une famille génératrice de N , où on a noté a_r le dernier des a_i non nuls. On note alors que cette nouvelle famille génératrice de N est libre, i.e. N est aussi libre avec $M/N \simeq A/(a_1) \times \cdots \times A/(a_r) \times A^{n-r}$. Montrons alors l'unicité des (a_i) . Notons tout d'abord que $n - r$ ne dépend que de N . Pour ce faire considérons un irréductible p ne divisant pas a_r de sorte que pour tout $1 \leq i \leq n$, a_i est inversible modulo p et donc pour $M_i = A/(a_i)$ on a M_i/pM_i est nul. On voit ainsi que $n - r$ est la dimension du $A/(p)$ espace vectoriel $(M/N)/p(M/N)$. Considérons alors

$$\frac{A}{(a_1)} \times \cdots \times \frac{A}{(a_q)} \simeq \frac{A}{(a'_1)} \times \cdots \times \frac{A}{(a'_s)},$$

les (a_i) et les (a'_i) vérifiant les propriétés de divisibilité de l'énoncé et sont tous non nuls. Alors $(a_r) = \text{Ann}(M/N) = (a'_s)$ et donc $\frac{A}{(a_1)} \times \cdots \times \frac{A}{(a_{q-1})} \simeq \frac{A}{(a'_1)} \times \cdots \times \frac{A}{(a'_{s-1})}$. En procédant de même de manière, on identifie de proche en proche les a_{q-i} avec les a'_{s-i} jusqu'à obtenir $s = q$.

Définition 56. On dit que $m \in M$ est un élément de torsion si $m \neq 0$ et s'il existe $\lambda \in A, \lambda \neq 0$, tel que $\lambda m = 0$. L'ensemble des éléments de torsion de M est noté M_t . Si $M_t = \{0\}$, on dit que M est sans torsion.

Remarque: M_t est un sous-module de M . Par ailleurs M est de torsion si et seulement si $M = M_t$. Avec ces notions la preuve du théorème précédent se réécrit.

Théorème 57. Soit M un A -module de type fini, M_t son sous-module de torsion. Alors il existe un sous-module $L \subset M$ libre de rang r tel que $M = M_t \oplus L$ ainsi que des éléments $a_1 \mid a_2 \mid \cdots \mid a_q$ de A tels que

$$M_t \simeq A/(a_1) \oplus A/(a_2) \oplus \cdots \oplus A/(a_q).$$

Définition 58. Le rang de la partie libre de M s'appelle le rang de M . Les idéaux non nuls (a_i) pour $1 \leq i \leq q$ sont les facteurs invariants de M .

Remarque: le corps des rationnels $\mathbb{Q}$ est un exemple de $\mathbb{Z}$ -module sans torsion et non libre (si $q_1 = a/b$ et $q_2 = c/d$, sont deux rationnels non nuls, on a la relation $bcq_1 - adq_2 = 0$). Ainsi $\mathbb{Q}$ n'est pas un $\mathbb{Z}$ -module de type fini.

Proposition 59. Soit M un A -module de type fini. Les conditions suivantes sont équivalentes :

1. le module M est indécomposable ;
2. $M \simeq A$, ou il existe un élément irréductible $p \in A$, un entier $\alpha > 0$ tels que $M \simeq A/(p^\alpha)$.

Preuve : 1. $\Rightarrow$ 2.

D'après le théorème 57 on peut supposer $M = A/(a)$. Si l'élément a a au moins deux facteurs irréductibles, il résulte du lemme chinois que M n'est pas indécomposable.

2. $\Rightarrow$ 1.

L'anneau A étant intègre, il est clair que le A -module A est indécomposable. Si $\alpha > 0$, les sous-modules de $\tilde{M} = A/(p^\alpha)$ sont engendrés par les images dans $\tilde{M}$ des éléments p^γ pour $\gamma \leq \alpha$. Si M_1 et M_2 sont deux tels sous-modules, on a toujours $M_1 \subset M_2$ ou $M_2 \subset M_1$; ils ne peuvent donc pas être en somme directe.

Définition 60. Soient M un A -module, $p \in \mathcal{P}$ un élément irréductible. On note $M(p)$ l'ensemble des éléments $x \in M$ de p -torsion, i.e. annulés par une puissance de p .

Remarque: en appliquant le théorème chinois au théorème 57, on obtient la décomposition canonique en indécomposable donnée par le théorème suivant.

Théorème 61. Soit M un A -module de torsion de type fini, $(a) = \text{Ann}(M)$ son annulateur. Alors :

1. $M = \bigoplus_{p_i \in \mathcal{P}, p_i | a} M(p_i)$ et $M(p_i) \neq (0)$ pour chaque élément irréductible p_i tel que $p_i | a$
2. pour chaque élément irréductible $p_i \in \mathcal{P}$, $p_i | a$, il existe une suite d'entiers $\nu_{i1} \leq \nu_{i2} \leq \dots \leq \nu_{ik}$ unique telle que :

$$M(p_i) \simeq \prod_{j=1}^k A/(p_i^{\nu_{ij}}).$$

3. la décomposition $M \simeq \prod_{i,j} A/(p_i^{\nu_{ij}})$ est l'unique décomposition de M en produit de modules indécomposables.

Exemple : prenons $A = \mathbb{Z}$, $M = M_t = \mathbb{Z}/96\mathbb{Z} \times \mathbb{Z}/72\mathbb{Z} \times \mathbb{Z}/10\mathbb{Z}$. On a $96 = 2^5 \times 3$, $72 = 2^3 \times 3^2$, d'où :

$$M \simeq \mathbb{Z}/32\mathbb{Z} \times \mathbb{Z}/3\mathbb{Z} \times \mathbb{Z}/8\mathbb{Z} \times \mathbb{Z}/9\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/5\mathbb{Z}$$

par le théorème chinois. On a donc

$$\begin{aligned} M(2) &\simeq \mathbb{Z}/32\mathbb{Z} \times \mathbb{Z}/8\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z} \\ M(3) &\simeq \mathbb{Z}/9\mathbb{Z} \times \mathbb{Z}/3\mathbb{Z} \\ M(5) &\simeq \mathbb{Z}/5\mathbb{Z}. \end{aligned} \tag{1}$$

Pour trouver la décomposition en indécomposable (resp. en facteurs invariants), on lit le tableau ci-dessus en lignes (resp. en colonnes), et on trouve $M = M(2) \oplus M(3) \oplus M(5)$ (resp. $a_3 = 32 \times 9 \times 5 = 1440$, $a_2 = 8 \times 3 = 24$, $a_1 = 2$, d'où la décomposition $M \simeq \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/24\mathbb{Z} \times \mathbb{Z}/1440\mathbb{Z}$).

Remarque: en considérant un groupe abélien comme un $\mathbb{Z}$ -module, on peut donner la classification des groupes abéliens finis d'ordre n donné. On procède de la manière suivante. Soit G un groupe d'ordre n .

1. On écrit $n = p_1^{\nu_1} \dots p_s^{\nu_s}$ avec p_i premiers, ν_i entiers ;
2. on a alors $G \simeq G(p_1) \oplus \dots \oplus G(p_s)$;
3. pour chaque entier i , $1 \leq i \leq s$ il existe une suite (ν_{ij}) unique d'entiers > 0 tels que $\sum_j \nu_{ij} = \nu_i$ et $G(p_i) \simeq \bigoplus_j \mathbb{Z}/p_i^{\nu_{ij}}\mathbb{Z}$;
4. deux groupes d'ordre n sont isomorphes si et seulement si tous les p_i et les entiers ν_{ij} sont les mêmes.

Exemple : donnons à isomorphisme près tous les groupes abéliens d'ordre $108 = 2^2 \times 3^3$. Soit $G = G(2) \oplus G(3)$ avec $G(2)$ d'ordre 4 et $G(3)$ d'ordre 27. Il y a à isomorphisme près deux possibilités pour $G(2)$, $\mathbb{Z}/4\mathbb{Z}$ et $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$, et trois pour $G(3)$, $\mathbb{Z}/27\mathbb{Z}$, $\mathbb{Z}/9\mathbb{Z} \times \mathbb{Z}/3\mathbb{Z}$ et $(\mathbb{Z}/3\mathbb{Z})^3$. Il y a donc à isomorphisme près six groupes abéliens d'ordre 108.

2.7 Exercices

Exercice 1. (*Fresnel p.127*) Soit A une matrice de $\mathbb{M}_n(K)$. On introduit la matrice N de $\mathbb{M}_{n^2, n+1}(K)$ où pour $1 \leq k \leq n+1$, son k -ième vecteur colonne est constitué des éléments de la matrice A^{k-1} pris dans un ordre prescrit une fois pour toutes. Montrer qu'il existe une matrice P de $SL_{n^2}(K)$ telle que PN soit de la forme $\begin{pmatrix} M_1 & M_2 \\ 0 & M_3 \end{pmatrix}$, où M_1 est un élément de $\mathbb{M}_{d, d+1}(K)$ "triangulaire supérieure" dont les termes "diagonaux" sont tous non nuls et donner un moyen de calculer le polynôme minimal

Preuve : L'existence de S découle du pivot de Gauss par opérations sur les lignes. Ainsi les $d+1$ - premières colonnes de $N' = SN$ sont liées soit ${}^t(u_0, \dots, u_{d-1}, -1, 0, \dots, 0)$ un vecteur du noyau de N' et donc aussi de

N car S est inversible. En notant C_i les colonnes de N , on a donc $C_d = u_0 C_0 + \dots + u_{d-1} C_{d-1}$ et donc $A^d - u_{d-1} A^{d-1} - \dots - u_0 \text{Id} = 0$. Supposons qu'il existe un polynôme annulateur de degré inférieur ou égal à $d-1$, de sorte que les $d-1$ premières colonnes de N sont liées et donc aussi celles de N' ce qui n'est pas, de sorte que $X^d - u_{d-1} X^{d-1} - \dots - u_0$ est bien le polynôme minimal de A .

Exercice 2. Soit M une matrice de $\mathbb{M}_n(\mathbb{R})$; les lettres L, L' (resp. U, U') désignent des matrices triangulaires inférieures (resp. supérieures) avec des coefficients diagonaux égaux à 1. La lettre D désigne une matrice diagonale.

- (1) Montrer que l'on peut mettre M sous la forme LU (resp. LDU) si et seulement si $\det M^{(k)} = 1$ (resp. $\det M^{(k)} \neq 0$) pour tout $1 \leq k \leq n$ où $M^{(k)}$ désigne le mineur principal d'ordre k .
- (2) Ecrire la matrice $\begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}$ sous la forme $U'LU$. Trouver une matrice de $\mathbb{M}_2(\mathbb{R})$ qui ne peut pas s'écrire sous la forme LDU où D désigne une matrice diagonale.
- (3) On note M_k la matrice obtenue à partir de M en substituant sa dernière ligne à sa ligne d'indice k . On suppose que $\det M_k^{(k)} \neq 0$ pour tout $1 \leq k \leq n$. Montrer que l'on peut mettre, de manière unique, M sous la forme $U'LU$ où U, U' ont des coefficients nuls en dehors de sa diagonale et de la dernière colonne.
- (4) Soit $M \in GL_n(\mathbb{R})$, montrer qu'il existe U' avec $n-1$ coefficients nuls sur sa dernière colonne telle que les k premières lignes de $(U'M)^{(k+1)}$ sont indépendantes pour tout $k < n$.
- (5) Montrer que toute matrice M de déterminant n peut se mettre sous la forme $L'U'LU$ où L' a ses coefficients nuls hors de sa diagonale et de sa dernière ligne.

Preuve : (1) La preuve est identique dans le cas respé, on traite la décomposition LU . Supposons par récurrence que $M^{(n-1)} = LU$. On a

$$M = \begin{pmatrix} LU & A \\ B & m \end{pmatrix} = \begin{pmatrix} L & 0 \\ BU^{-1} & 1 \end{pmatrix} \begin{pmatrix} U & L^{-1}A \\ 0 & m' \end{pmatrix}$$

où $m' = m - BU^{-1}L^{-1}A$. Par passage au déterminant on a $m' = 1$ d'où le résultat. Pour l'unicité si $LU = L'U'$ alors $L'^{-1}L = U'U^{-1}$ qui est forcément égale à l'identité.

(2) La décomposition

$$\begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} = \begin{pmatrix} 1 & -\tan(\theta/2) \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ \sin \theta & 1 \end{pmatrix} \begin{pmatrix} 1 & -\tan(\theta/2) \\ 0 & 1 \end{pmatrix}$$

est utilisée pour implémenter la rotation des images numériques. La matrice $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ ne peut pas s'écrire sous la forme LDU .

(3) On note $(c_1, \dots, c_{n-1}, 1)$ les coefficients de la dernière colonne de U'^{-1} . Par multi-linéarité du déterminant on a

$$\det(U'^{-1}M)^{(k)} = \det M^{(k)} + \sum_{j=1}^k c_j M_j^{(k)}$$

Les équations $\det(U'^{-1}M)^{(k)} = 1$ pour $1 \leq k \leq n-1$ forment un système triangulaire inversible car $\det M_k^{(k)} \neq 0$, il a donc une unique solution et $(U'^{-1}M)^{(k)} = LU$ d'après (1).

(4) On démontre cette propriété pour tout $k \leq r$ par récurrence sur $r < n$ en utilisant que les matrices U' forment un groupe. On note $U_{k,k}$ la matrice identité et $U_{k,l}$ ($k < l < n$) la matrice qui par multiplication à gauche ajoute la ligne l à la ligne k . On extrait de M les matrices supérieures gauche $M^{(i,j)} \in \mathbb{M}_{i,j}(\mathbb{R})$.

Pour $r = 1$, $\text{rg} M^{(n,2)} = 2$ garantit $\text{rg} M^{(n-1,2)} \geq 1$. Si la première ligne de $M^{(n-1,2)}$ est nulle on note $l > 1$ l'indice d'une ligne non nulle, sinon on pose $l = 1$. La matrice $U' = U_{1,l}$ convient, i.e. la première ligne de $(U_{1,l}M)^{(2)}$ n'est pas nulle.

On suppose que M satisfait la propriété pour tout $k \leq r-1$ si bien que les $r-1$ premières lignes de $M^{(n-1,r+1)}$ sont indépendantes. Par ailleurs, $\text{rg} M^{(n-1,r+1)} \geq \text{rg} M^{(n,r+1)} - 1 = r$. On choisit $U' = U_{r,l}$ où $l = r$ si la r -ième ligne de $M^{(n-1,r+1)}$ est indépendante des précédentes et $l > r$ est l'indice d'une ligne indépendante sinon. Les r premières lignes de $(U'M)^{(k+1)}$ sont alors indépendantes pour tout $k \leq r$.

(5) Soit U_0 la matrice obtenue en appliquant la question précédente à M . On remarque que $L_0 = L'^{-1}$ est de la même forme que L' et donc que U_0 et L_0 commutent. Pour tout vecteur v indépendant des $(n-1)$ -premières lignes de M , on peut choisir L_0 telle que la multiplication à gauche de L_0 sur M change la dernière ligne de M en v . Par construction de U_0 , il est possible de choisir les coefficients de v successivement pour que $L_0 U_0 M$ satisfasse les hypothèses de la question (3). En notant U_1 la matrice U' obtenue en appliquant la question (3) à $L_0 U_0 M$ on obtient $U_0 L_0 M = L_0 U_0 M = U_1 L U$. On conclut en posant $U' = U_0^{-1} U_1$.

Exercice 3. Soit $x = (n_1, \dots, n_p) \in \mathbb{Z}^p$.

a) Montrer que les conditions suivantes sont équivalentes :

1. $\text{PGCD}(n_1, \dots, n_p) = 1$
2. Il existe $A \in \text{SL}_p(\mathbb{Z})$ telle que $A^t x = {}^t(1, 0, \dots, 0)$
3. Le vecteur x fait partie d'une base de $\mathbb{Z}^p$.

b) On pose $p = 4$ et $x = (10, 6, 7, 11)$. Compléter x en une base de $\mathbb{Z}^4$.

Preuve : (i) implique (ii) : le résultat se démontre par récurrence ; la

première étape du calcul est la suivante :

$$\begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \cdots & u & v \\ 0 & \cdots & -n_r/(n_r, n_{r-1}) & n_{r-1}/(n_r, n_{r-1}) \end{pmatrix} \begin{pmatrix} n_1 \\ \vdots \\ n_r \end{pmatrix} = \begin{pmatrix} n_1 \\ \vdots \\ (n_r, n_{r-1}) \\ 0 \end{pmatrix}$$

où u et v sont les coefficients de Bezout entre n_{r-1} et n_r : $un_{r-1} + vn_r = (n_r, n_{r-1})$.

On considère $f : \mathbb{Z}^r \rightarrow \mathbb{Z}$ défini par $f(a_1, \dots, a_r) = a_1 n_1 + \dots + a_r n_r$. Le théorème de la base adaptée, nous assure l'existence d'une base $(e_1, \dots, e_n)$ de $\mathbb{Z}^n$ tel que $\text{Ker } \phi = \mathbb{Z}a_1 e_1 \oplus \dots \oplus \mathbb{Z}a_r e_r$ avec $a_i | a_{i+1}$ dans $\mathbb{Z}$ que l'on appelle les facteurs invariants de $\mathbb{Z}^n / \text{Ker } \phi$; on a en outre que les a_i sont tous égaux à 1 pour $1 \leq i < r$ et $a_r = 0$ (par un argument de dimension). Ainsi si on note A transposée de la matrice de passage de la base $(e_r, e_{r-1}, \dots, e_1)$ dans

la base canonique, on a $A \in SL_r(\mathbb{Z})$ et $A \begin{pmatrix} n_1 \\ \vdots \\ n_r \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$.

(ii) implique (iii) : il est évident que la famille $\begin{pmatrix} n_1 \\ \vdots \\ n_r \end{pmatrix} = A^{-1} \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$,

$A^{-1} \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \dots, A^{-1} \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$ forme une base de $\mathbb{Z}^n$.

(iii) implique (i) : soit $e_2, \dots, e_r$ une famille qui complète $\begin{pmatrix} n_1 \\ \vdots \\ n_r \end{pmatrix}$ en

une base et on note A la matrice de passage de cette base dans la base canonique ; on calcule le déterminant de A en le développant par rapport à la première colonne de sorte que celui-ci est divisible par le pgcd des n_i qui est donc égal à 1 car $\det A = \pm 1$.

Exemples : le premier cas est simple car on a la relation $7 - 6 = 1$ de sorte que la matrice suivante est de déterminant -11

$$\begin{pmatrix} 10 & 6 & 7 & 11 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

de sorte que les 4 vecteurs colonnes de la transposée de la matrice ci-dessus constituent une base de $\mathbb{Z}^4$.

Dans le deuxième exemple on a la relation de Bezout : $1 = 6 \cdot 6 - 2 \cdot 10 - 15$. On cherche donc 6 coefficients a, b, c, d, e, f tels que $cf - de = 6$, $af - be = 2$ et $ad - bc = -1$. Par exemple la matrice suivante convient

$$\begin{pmatrix} 6 & 1 & 0 \\ 10 & 0 & -1 \\ 15 & 6 & 2 \end{pmatrix}$$

Exercice 4. Soient n un entier positif et $G \subset \mathbb{Z}^n$ un sous-groupe de rang n . Soit $(g_1, \dots, g_n)$ une base de G ; on note M la matrice de passage de cette base dans la base canonique de $\mathbb{Z}^n$.

- (i) Montrer que le groupe $\mathbb{Z}^n/G$ est fini.
- (ii) Montrer que $\text{card}(\mathbb{Z}^n/G) = |\det M|$.
- (iii) Soit H un groupe abélien engendré par trois éléments h_1, h_2, h_3 soumis aux relations

$$3h_1 + h_2 + h_3 = 0$$

$$25h_1 + 8h_2 + 10h_3 = 0$$

$$46h_1 + 20h_2 + 11h_3 = 0$$

Montrer que $\text{card}(H) = 19$ puis que $H \simeq \mathbb{Z}/19\mathbb{Z}$. Quelle généralisation cela suggère-t-il ?

- (iv) Triangulariser la matrice

$$\begin{pmatrix} 3 & 1 & 1 \\ 25 & 8 & 10 \\ 46 & 20 & 11 \end{pmatrix}$$

en multipliant à droite par une matrice de $SL_3(\mathbb{Z})$ que l'on précisera.

- (v) En déduire un isomorphisme $\varphi : H \simeq \mathbb{Z}/19\mathbb{Z}$ et préciser les valeurs de $\varphi(h_1), \varphi(h_2), \varphi(h_3)$.

Preuve :

(i) Le théorème de la base adaptée fournit une base $(f_1, \dots, f_n)$ de $\mathbb{Z}^n$ ainsi que des entiers $1 < a_1 | \dots | a_n \neq 0$ tels que $(a_1 f_1, \dots, a_n f_n)$ soit une base de G . On obtient alors $\mathbb{Z}^n/G \simeq \mathbb{Z}/a_1\mathbb{Z} \times \dots \times \mathbb{Z}/a_n\mathbb{Z}$.

(ii) D'après ce qui précède, on a donc $\text{card}(\mathbb{Z}^n/G) = \prod_{i=1}^n a_i$ qui est donc égal à $|\det M|$.

(iii) Soit $M = \begin{pmatrix} 3 & 1 & 1 \\ 25 & 8 & 10 \\ 46 & 20 & 11 \end{pmatrix}$ de sorte que $M \begin{pmatrix} h_1 \\ h_2 \\ h_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$ Il

existe alors des matrices $L, R \in GL_3(\mathbb{Z})$ telles que $M = L \text{diag}(a_1, a_2, a_3) R$

avec $a_1|a_2|a_3$. En outre si on pose $\begin{pmatrix} h'_1 \\ h'_2 \\ h'_3 \end{pmatrix} := R \begin{pmatrix} h_1 \\ h_2 \\ h_3 \end{pmatrix}$, H est aussi engendré par h'_1, h'_2, h'_3 et l'équation $L \text{diag}(a_1, a_2, a_3) \begin{pmatrix} h'_1 \\ h'_2 \\ h'_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$ est équivalente à

$$\begin{cases} a_1 h'_1 = 0 \\ a_2 h'_2 = 0 \\ a_3 h'_3 = 0 \end{cases}$$

et donc $H \simeq \mathbb{Z}/a_1\mathbb{Z} \times \mathbb{Z}/a_2\mathbb{Z} \times \mathbb{Z}/a_3\mathbb{Z}$, avec $a_1.a_2.a_3 = \det M$. L'énoncé nous suggère de simplement calculer $\det M$; on vérifie aisément qu'il est égal à -19 (cf. (iv) ci-après) comme annoncé. On obtient alors $a_1 = a_2 = 1$ et $a_3 = 19$.

De manière générale si la décomposition en facteurs premiers de $\det M$ ne fait apparaître aucune multiplicité (i.e. $p^2 \nmid \det M$ pour tout premier p), alors tous les a_i sont égaux à 1 sauf le dernier égal à $\det M$ et le groupe quotient est alors cyclique.

(iv) On calcule

$$\begin{pmatrix} 3 & 1 & 1 \\ 25 & 8 & 10 \\ 46 & 20 & 11 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & -1 & 0 \\ 1 & 3 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & -2 \\ 0 & 0 & -1 \end{pmatrix} \\ \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 3 & 1 & 0 \\ 25 & 8 & 2 \\ 46 & 20 & -9 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 8 & -1 & 2 \\ 20 & 14 & -9 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 8 & 1 & 0 \\ 20 & -14 & -19 \end{pmatrix} \\ \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & -1 & 0 \\ 1 & 3 & -1 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 & 2 \\ 1 & -3 & -5 \\ 0 & 0 & -1 \end{pmatrix}$$

(v) On a $\begin{pmatrix} h'_1 \\ h'_2 \\ h'_3 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 2 \\ 1 & -3 & -5 \\ 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} h_1 \\ h_2 \\ h_3 \end{pmatrix}$ ce qui s'inverse faci-

lement (la matrice est "triangulaire") soit $h_3 = h'_3$, $h_2 + 2h_3 = h'_1$ soit $h_2 = h'_1 - 2h'_3$ et $h_1 - 3h_2 - 5h_3 = h'_2$ soit $h_1 = 3h'_1 + h'_2 - h'_3$. Comme $\varphi(h'_1) = \varphi(h'_2) = 0$ et $\varphi(h'_3) = 1$, on obtient $\varphi(h_1) = -1$, $\varphi(h_2) = -2$ et $\varphi(h_3) = 1$.

3 Réduction des endomorphismes

Comme on l'a vu plus haut, à chaque endomorphisme f , on associe des matrices qui dépendent du choix des bases. On cherche alors à trouver des bases pour que la matrice soit la plus simple possible.

Remarque: on utilisera le langage de la section suivante avec les endomorphismes orthogonaux (resp. unitaires), symétriques (resp. hermitiens) dans le cadre réel (resp. complexe). On espère que la perte de logique de présentation des objets, sera compensée par l'aspect pratique de rerouter en une seule section, un ensemble de résultat portant sur la réduction.

3.1 Matrices équivalentes

Définition 62. Deux matrices $A, A' \in \mathbb{M}_{m,n}(\mathbb{K})$ (resp. de $\mathbb{M}_n(\mathbb{K})$) sont dites équivalentes (resp. semblables) s'il existe deux matrices $P \in GL_m(\mathbb{K})$ et $Q \in GL_n(\mathbb{K})$ (resp. $P \in GL_n(\mathbb{K})$) telles que $A' = PAQ$ (resp. $A' = P^{-1}AP$).

Remarque: A et A' sont équivalentes (resp. semblables) si elles représentent le même morphisme $f : E \rightarrow F$ (resp. $f : E \rightarrow E$) relativement à des bases différentes au départ et à l'arrivée (resp. des bases différentes mais les mêmes au départ et à l'arrivée).

Proposition 63. Toute matrice $A \in \mathbb{M}_{m,n}(\mathbb{K})$ est équivalente à une matrice de la forme

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

où le nombre de 1 est égal au rang de f .

Remarque: ainsi les classes d'équivalence dans $\mathcal{L}(E, F)$ sont données par le rang.

Considérons à présent le cas $\mathbb{K} = \mathbb{R}$ (resp. $\mathbb{K} = \mathbb{C}$) et où les espaces considérés sont munis d'un produit scalaire (resp. hermitien). On ne s'autorise alors à ne considérer que des bases orthonormées et donc des matrices de changement de base orthogonale (resp. unitaire).

Proposition 64. (Décomposition polaire) Soit $A \in \mathbb{M}_m(\mathbb{R})$ (resp. $\mathbb{M}_m(\mathbb{C})$), on peut alors écrire A sous la forme $A = PU$ où $P \in \mathbb{M}_m(\mathbb{R})$ (resp. $P \in \mathbb{M}_m(\mathbb{C})$) est positive semi-définie de même rang que A et $U \in \mathbb{M}_m(\mathbb{R})$ (resp. $\mathbb{M}_m(\mathbb{C})$) dont les vecteurs colonnes forment une famille orthonormale, i.e. $U^t U = I_m$ (resp. $U U^* = I_m$). La matrice P est uniquement déterminée comme l'unique racine carrée positive de $A^t A$ (resp. de AA^*) alors que U n'est uniquement déterminée que si A est de rang m .

Remarque: il faut voir cette décomposition comme la généralisation dans le cas $n = 1$ de l'écriture d'un nombre complexe sous la forme $\rho e^{i\theta}$.

Preuve : On traite le cas de $\mathbb{R}$, le cas de $\mathbb{C}$ se traitant de manière similaire. La matrice $A^t A$ est symétrique positive et donc diagonalisable en base ortho-normée, i.e. il existe O orthogonale telle que $OA^t A(^t O) = \text{diag}(\lambda_1, \dots, \lambda_m)$ avec les λ_i positifs ou nuls. On note alors

$$P = {}^t O \text{diag}(\sqrt{\lambda_1}, \dots, \sqrt{\lambda_m}) O$$

qui est donc positive semi-définie. Supposons que P est inversible, i.e. que A est de rang m , alors $U = P^{-1}A$ vérifie

$$U^t U = P^{-1}(A^t A)^t P^{-1} = P^{-1}(P^2)P^{-1} = I_m.$$

Si A n'est pas de rang m , soit $(A_i)_{i \in \mathbb{N}}$ une suite de matrices de $\mathbb{M}_m(\mathbb{R})$ de rang m convergeant vers A . D'après ce qui précède, on peut écrire de manière unique $A_i = P_i U_i$: comme $(U_i)_{i \in \mathbb{N}}$ appartient au compact $O_m(\mathbb{R})$ et admet donc une valeur d'adhérence $U \in O_m(\mathbb{R})$ de sorte que la suite extraite correspondante des P_i converge vers $P := AU^{-1}$ nécessairement symétrique positive semi-définie puisque l'ensemble des matrices symétriques positives semi-définies est clairement fermé.

Remarque: en raisonnant avec des suites extraites comme dans la preuve ci-dessus, il est aisé de démontrer que la décomposition polaire dans le cas inversible, est un homéomorphisme.

Remarque: dans la preuve précédente on construit assez naturellement la matrice hermitienne positive P alors que U n'est donné qu'artificiellement par la formule $P^{-1}A$. Pour corriger cette injustice montrons le corollaire suivant.

Corollaire 65. *L'application $M \mapsto \sqrt{\frac{1}{n} \text{tr}(MM^*)}$ définit une norme $\| - \|$ sur $\mathbb{M}_n(\mathbb{C})$ telle que $\|U\| = 1$ pour toute matrice unitaire. Pour $A \in GL_n(\mathbb{C})$ de décomposition polaire $A = SU$, la matrice unitaire est unique-ment déterminée par la propriété suivante :*

$$\|A - U\| = \min_{U' \in \mathbb{U}_n(\mathbb{C})} \|A - U'\|,$$

où $\mathbb{U}_n(\mathbb{C})$ désigne l'ensemble des matrices unitaires de $\mathbb{M}_n(\mathbb{C})$.

Remarque: une autre façon d'énoncer le résultat est de dire que U est la projection orthogonale de A sur la boule unité des matrices M telles que $\|M\| \leq 1$ qui est un compact convexe.

Preuve : Calculons tout d'abord

$$\begin{aligned} \|A - U\| &= \|S - I_n\| = \|PDP^* - I_n\| \\ &= \|D - I_n\| = \sqrt{\frac{1}{n} \sum_{i=1}^n (\sigma_i - 1)^2} \end{aligned}$$

où les σ_i sont les valeurs propres de S (on les appellera bientôt les valeurs singulières de A). Calculons alors

$$\begin{aligned}\|A - U'\| &= \|SU - U'\| = \|SU(U')^{-1} - I_n\| \\ &= \|PDP^*U(U')^{-1} - I_n\| = \|D(P^*U(U')^{-1}P)\|.\end{aligned}$$

Notons alors $V = P^*U(U')^{-1}P$ est qui est unitaire dont on note $v_{i,j}$ les coefficients. On a alors

$$\begin{aligned}\|A - U'\| &= \|DV - I_n\| = \sqrt{\frac{1}{n} \sum_{i=1}^n \sigma_i^2 + n - \sigma_i(v_{i,i} + \overline{v_{i,i}})} \\ &\leq \sqrt{\frac{1}{n} \sum_{i=1}^n (\sigma_i^2 + 1 - 2\sigma_i)} = \|A - U\|\end{aligned}$$

où on a utilisé la majoration $v_{i,i} + \overline{v_{i,i}} = 2\operatorname{Re}(v_{i,i}) \leq |v_{i,i}| \leq 1$, puisque les vecteurs colonnes de V sont de norme 1.

Pour le cas d'égalité on doit avoir $|v_{i,i}| = 1$ et $2\operatorname{Re}(v_{i,i}) = 1$ soit $v_{i,i} = 1$. Or la matrice V étant unitaire, on a $v_{i,j} = \delta_{i,j}$, i.e. $V = I_n$ soit $U(U')^{-1} = I_n$ et donc $U = U'$.

Corollaire 66. (valeurs singulières)

Soit $A \in \mathbb{M}_m(\mathbb{C})$ de rang k , on peut alors l'écrire sous la forme

$$A = VDW^*,$$

où $V \in \mathbb{M}_m(\mathbb{C})$ et $W \in \mathbb{M}_m(\mathbb{C})$ sont unitaires et $D = \operatorname{diag}(d_1, \dots, d_m)$ avec $d_1 \geq d_2 \geq \dots \geq d_k > d_{k+1} = \dots = d_m = 0$. Les nombres $d_{i,i}$ sont les racines carrées positives des valeurs propres de AA^* et sont donc uniquement déterminées : on les appelle les valeurs singulières de A .

Remarque: V et W ne sont pas uniquement déterminées, on peut juste dire que les colonnes de V (resp. W) sont des vecteurs propres de AA^* (resp. A^*A).

Preuve : On part de la décomposition polaire $A = PU$ et on diagonalise $P = VDV^*$ en base orthonormée où D est comme dans l'énoncé. On pose alors $W = U^*V$ de sorte que $A = VDW^*$.

Remarque: les valeurs singulières apparaissent dans le conditionnement d'une matrice en analyse numérique. Rappelons brièvement de quoi il s'agit. Soit A une matrice inversible et B une matrice colonne, on cherche à résoudre l'équation $AX = B$ d'inconnue X . D'un point de vue pratique, B peut subir une petite perturbation δ_B due par exemple à des arrondis et on cherche à contrôler la perturbation δ_X induite sur X , i.e. $A(X + \delta_X) = B + \delta_B$, soit $A\delta_X = \delta_B$. On choisit une norme subordonnée par exemple à la classique norme euclidienne $\|\cdot\|_2$ de sorte que $\|A\|_2$ est égal à la plus grande valeur propre de A^*A , i.e. à la plus grande valeur singulière. On a alors, en utilisant

$$\|A\delta_X\|_2 \leq \|A\|_2 \|\delta_X\|_2, \quad \|A^{-1}B\|_2 \leq \|A^{-1}\|_2 \|B\|_2 \quad (2)$$

on en déduit

$$\frac{|||\delta_X|||_2}{|||X|||_2} \leq |||A|||_2 \cdot |||A^{-1}|||_2 \frac{|||\delta_B|||_2}{|||B|||_2}.$$

Le conditionnement de A relativement à la norme $\| - \|_2$ est alors la quantité $|||A|||_2 \cdot |||A^{-1}|||_2$ qui est donc le quotient $\frac{\mu_n}{\mu_1}$ de la plus grande valeur singulière par la plus petite. En utilisant qu'en dimension finie, la sphère unité est compacte, les inégalités de (2) sont optimales, i.e. les cas d'égalités existent, de sorte que la majoration précédente est optimale.

Illustrons le phénomène :

- on prend $A \in GL_2(\mathbb{R})$ symétrique définie positive de sorte que les valeurs singulières sont égales aux valeurs propres.
- Soient e_1 et e_2 les vecteurs propres associés à $\lambda_1 \leq \lambda_2$ et on suppose λ_2 grand et λ_1 petit.
- On pose $B = e_2$ de sorte que $X = \frac{1}{\lambda_2} e_2$ qui est petit, et
- $\delta_B = \lambda_1 e_1$, qui est donc petit,

de sorte que $\delta_X = e_1$ est grand.

Remarque: on peut aussi perturber la matrice A en gardant B : à nouveau la perturbation de X est contrôlé par le conditionnement de A .

3.2 Vecteurs propres et espaces propres

Définition 67. Un vecteur $v \in E$ est dit propre par un endomorphisme f s'il est non nul et s'il existe un scalaire $\lambda \in \mathbb{K}$ tel que $f(v) = \lambda v$; le scalaire λ est alors appelé une valeur propre. On note $\text{Spec} f$ l'ensemble des valeurs propres de f : c'est le spectre de f .

Remarque: un vecteur propre définit une droite stable par f ; plus généralement un sous-espace F de E est dit stable par f si $f(F) \subset F$. Si on prend une base de F que l'on complète par en une base de E , la matrice de f dans cette base sera triangulaire par blocs, i.e. de la forme $\begin{pmatrix} A & C \\ 0 & B \end{pmatrix}$ avec $A \in \mathbb{M}_{n_1}(K)$ et $B \in \mathbb{M}_{n_2}(K)$ avec $n_1 + n_2 = n$.

Exemples : le noyau $\text{Ker } f$ et l'image $\text{Im } f$ sont des sous-espaces stables par f .

Proposition 68. Soit E un espace vectoriel muni d'une base $(e_i)_{1 \leq i \leq n}$. Il existe alors une unique application $\det_{(e_i)_i} : E^n \rightarrow \mathbb{K}$ qui soit multilinéaire alternée et telle que $\det_{(e_i)_i}(e_1, \dots, e_n) = 1$.

Définition 69. Pour $E = \mathbb{K}^n$ muni de la base canonique et $\mathbb{M}_n(\mathbb{K})$ identifié via ses vecteurs colonnes à E^n , l'application de la proposition précédente définit $\det : \mathbb{M}_n(\mathbb{K}) \rightarrow \mathbb{K}$ et s'appelle le déterminant.

Remarque: par opération élémentaires sur les vecteurs colonnes d'une matrice, on montre que $\det A \neq 0$ si et seulement si $A \in GL_n(\mathbb{K})$ ainsi que le corollaire suivant.

Corollaire 70. Pour $A, B \in \mathbb{M}_n(\mathbb{K})$ on a $\det(AB) = \det A \cdot \det B$.

Définition 71. Le polynôme caractéristique d'un endomorphisme $f \in \mathcal{L}(E)$ est le déterminant $\chi_A(X) := \det(A(f) - XI_n) \in \mathbb{K}[X]$ où $A(f)$ est la matrice de f dans une base de E quelconque.

Remarque: le fait que χ_A soit indépendant de la base provient du fait que $\det(P^{-1}AP) = \det A$ d'après le corollaire précédent.

Remarque: dans le cas où χ_A est totalement décomposé (par exemple si $\mathbb{K} = \mathbb{C}$), le produit des valeurs propres est égal à $(-1)^n$ fois le coefficient constant de χ_A et donc au déterminant de A .

Lemme 72. Sur $\mathbb{C}$, la norme du produit des valeurs propres d'une matrice $A \in \mathbb{M}_n(\mathbb{C})$ est égal au produit des valeurs absolues de ses valeurs singulières.

Remarque: au corollaire 173, on donnera des raffinements de cette égalité.

Preuve : Il suffit de noter que le déterminant d'une matrice unitaire est nécessairement de module 1, puisque $U^*U = I_n$ implique que $|\det U| = 1$.

Remarque: la formation du polynôme caractéristique $A \mapsto \chi_A(X)$ est clairement continue puisque polynomiale. Rappelons par ailleurs que, cf. le corollaire ??, les racines dépendent continûment de leur polynôme, de sorte que les valeurs propres dépendent continûment de la matrice. La proposition suivante quantifie cette propriété.

Proposition 73. (Disques de Gershgorin) Les valeurs propres de $A = (a_{i,j})_{1 \leq i,j \leq n} \in \mathbb{M}_n(\mathbb{C})$ appartiennent à la réunion des disques fermés centrés en $a_{i,i}$ et de rayon $\sum_{j \neq i} |a_{i,j}|$.

Preuve : Le résultat découle directement du lemme d'Hadamard appliqué à $A - \lambda \text{Id}$. Rappelons que ce lemme dit que si pour tout $1 \leq i \leq n$, on a $|a_{i,i}| > \sum_{j \neq i} |a_{i,j}|$ alors A est inversible. En effet soit X de coordonnées $(x_i)_{1 \leq i \leq n}$ dans le noyau de A et soit i_0 tel que $|x_{i_0}|$ soit maximal parmi les $|x_i|$. De l'égalité $a_{i_0,i_0}x_{i_0} = -\sum_{j \neq i_0} a_{i_0,j}x_j$ on en déduit la majoration $|a_{i_0,i_0}x_{i_0}| \leq |x_{i_0}| \sum_{j \neq i_0} |a_{i_0,j}|$ et donc $x_{i_0} = 0$ soit $X = 0$.

D'un point de vue pratique, examinons la perturbation subie par les valeurs propres lorsqu'on perturbe la matrice. Commençons par l'exemple simple donné par la matrice compagnon de $X^{100} - 10^{-100}$ dont les valeurs propres sont de module égal à 1. Cette matrice est ainsi une très faible perturbation du bloc de Jordan de taille 100 dont les valeurs propres sont toutes nulles. En conclusion une perturbation d'ordre 10^{-100} nous conduit à une perturbation de l'ordre 0, 1 ce qui est énorme. Essayons de quantifier ce phénomène : pour ce faire on considère une norme matricielle $\|\cdot\|$ telle que pour toute matrice diagonale $D = \text{diag}(\lambda_1, \dots, \lambda_n)$, on ait $\|D\| = \max_i |\lambda_i|$. On peut par exemple prendre les normes $\|\cdot\|_1$, $\|\cdot\|_2$ ou $\|\cdot\|_\infty$.

Proposition 74. Soit A une matrice diagonalisable avec $\text{Spec} A = \{\lambda_1, \dots, \lambda_n\}$. Alors

$$\text{Spec}(A + \delta_A) \subset \bigcup_{i=1}^n \{z \in \mathbb{C} : |z - \lambda_i| \leq \gamma(A) \|\delta_A\|\},$$

où

$$\gamma(A) = \inf\{\|P\| \cdot \|P^{-1}\| : P^{-1}AP = \text{diag}(\lambda_1, \dots, \lambda_n)\}.$$

Remarque: Ainsi donc le contrôle des valeurs propres est donné par le conditionnement des matrices de passage et pas par le conditionnement de la matrice de départ.

Preuve : Considérons P diagonalisant A , i.e.

$$P^{-1}AP = D := \text{diag}(\lambda_1, \dots, \lambda_n)$$

et soit λ une valeur propre de $A + \delta_A$. Si $\lambda \in \{\lambda_1, \dots, \lambda_n\}$, il n'y a rien à montrer, sinon $(D - \lambda I_n)$ est inversible et on peut écrire

$$\begin{aligned} P^{-1}(A + \delta_A - \lambda I_n)P &= D - \lambda I_n + P^{-1}\delta_A P \\ &= (D - \lambda I_n)(I_n + (D - \lambda I_n)^{-1}P^{-1}\delta_A P). \end{aligned}$$

La matrice $(I_n + (D - \lambda I_n)^{-1}P^{-1}\delta_A P)$ n'étant pas inversible, -1 est donc une valeur propre de $(D - \lambda I_n)^{-1}P^{-1}\delta_A P$, de sorte que d'après ?? sa norme est ≥ 1 , ce qui donne :

$$1 \leq \|(D - \lambda I_n)^{-1}P^{-1}\delta_A P\| \leq \|(D - \lambda I_n)^{-1}\| \cdot \|P^{-1}\| \cdot \|\delta_A\| \cdot \|P\|.$$

Comme par hypothèse $\|(D - \lambda I_n)^{-1}\| = \frac{1}{\min |\lambda_i - \lambda|}$, il existe au moins un indice i tel que

$$|\lambda_1 - \lambda| \leq \|P\| \cdot \|P^{-1}\| \cdot \|\delta_A\|.$$

Remarque: En particulier si A est normale, la matrice de passage est orthogonale et son conditionnement est égal à 1. Autrement dit lorsqu'on perturbe une matrice normale, ses valeurs propres sont perturbées dans la même proportion. Le cas le plus intéressant est certainement celui où A et δ_A sont toutes deux symétriques, on renvoie au §4.8 pour une étude plus précise dans cette situation.

Définition 75. Le sous-espace propre E_λ (resp. caractéristique $E(\lambda)$) associée à une valeur propre λ est $\text{Ker}(f - \lambda \text{Id})$ (resp. $\text{Ker}(f - \lambda \text{Id})^n$ où n est la dimension de E , ou plus simplement la multiplicité de λ dans le polynôme minimal μ_f).

Remarque: la dimension du sous-espace caractéristique est égale à la multiplicité de λ dans le polynôme caractéristique.

Lemme 76. (dit des noyaux)

Si $P = P_1 P_2$ est un polynôme annulateur de f avec $P_1 \wedge P_2 = 1$ alors $E = \text{Ker } P_1(f) \oplus \text{Ker } P_2(f)$.

Preuve : On part d'une relation de Bezout $UP_1 + VP_2 = 1$ de sorte que pour tout $e \in E$, on a $e = e_1 + e_2$ avec $e_1 = UP_1(f)(e)$ et $e_2 = VP_2(f)(e)$. On a alors

$$P_2(f)(e_1) = U(f) \circ P_1P_2(f)(e_1) = 0 \text{ et } P_1(f)(e_2) = V(f) \circ P_1P_2(f)(e_2) = 0$$

et donc $e_1 \in \text{Ker } P_2(f)$ et $e_2 \in \text{Ker } P_1(f)$ et $E = \text{Ker } P_1(f) + \text{Ker } P_2(f)$. En outre si $e \in \text{Ker } P_1(f) \cap \text{Ker } P_2(f)$ alors $e_1 = e_2 = 0$ et $e = e_1 + e_2 = 0$.

Remarque: il est important de noter que les projecteurs sur chacun de ces sous-espaces stables parallèlement à l'autre, sont des polynômes en f .

3.3 Polynôme minimal

La théorie de la réduction d'un endomorphisme est totalement contrôlée par une série de polynômes qu'on lui associe, appelés *ses invariants de similitude*. Nous verrons que le polynôme caractéristique est le produit des invariants de similitude. En général, la réponse à une question sur un endomorphisme s'exprime en utilisant toute la puissance des invariants de similitude. Le plus gros des invariants de similitude est donné par la définition suivante où l'on rappelle qu'étant donné un endomorphisme f et un polynôme $P(X) = \sum_i a_i X^i \in \mathbb{K}[X]$, $P(f)$ désigne l'endomorphisme $\sum_i a_i f^i$ où f^i désigne $f \circ \dots \circ f$.

Définition 77. Pour $f \in \mathcal{L}(E)$, l'ensemble I_f des polynôme $P \in \mathbb{K}[X]$ tels que $P(f)$ est l'endomorphisme nul, est un idéal de $\mathbb{K}[X]$; cet anneau étant principal, il existe un unique polynôme unitaire μ_f , appelé polynôme minimal de f , tel que $I_f = \langle \mu_f \rangle$.

Remarque: comme E est de dimension finie, la famille $\text{Id}, f, f^2, \dots, f^{n^2+1}$ est liée de sorte que I_f n'est pas l'idéal nul et donc μ_f n'est pas le polynôme nul.

Théorème 78. (de Cayley-Hamilton)

Le polynôme caractéristique χ_f appartient à I_f , i.e. $\chi_f(f)$ est l'endomorphisme nul.

3.4 Trigonalisation

Définition 79. Un endomorphisme est dit trigonalisable s'il existe une base dans laquelle sa matrice est triangulaire supérieure.

Remarque: une autre façon d'énoncer cette propriété est de demander qu'il existe une drapeau complet

$$\{0\} = F_0 \subset F_1 \subset \dots \subset F_n = E$$

avec $\dim F_i = i$, stable par f , i.e. pour tout $i = 0, \dots, n$ on a $f(F_i) \subset F_i$.

Théorème 80. *Un endomorphisme f est trigonalisable si et seulement si χ_f est scindé sur $\mathbb{K}$.*

Remarque: ainsi si K est algébriquement clos tous les endomorphismes sont trigonalisables.

On s'intéresse à présent à la question de la trigonalisation simultanée, i.e. étant donné un sous-ensemble $\mathcal{E}$ de $\mathcal{L}(E)$, on cherche des drapeaux, si possible complets, stables par tous les $u \in \mathcal{E}$. On rappelle que si u est un endomorphisme d'un espace vectoriel E et si $F \subset E$ est un sous-espace stable par u alors u induit un endomorphisme $\bar{u}$ de E/F .

Étant donné un sous-ensemble $\mathcal{E}$ de $\mathcal{L}(E)$ et $G \subset F \subset E$ des sous-espaces stables par tous les éléments u de $\mathcal{E}$, l'ensemble des quotients de $\mathcal{E}$ pour $\{G \subset F\}$ est par définition $\{\bar{u} \in \mathcal{L}(F/G) : u \in \mathcal{E}\}$.

Définition 81. *Une propriété P sera dite stable par quotients si pour tout ensemble $\mathcal{E} \subset \mathcal{L}(E)$ constitués d'éléments satisfaisant P alors l'ensemble quotient de $\mathcal{E}$ pour $\{G \subset F\}$ est aussi constitué d'éléments de $\mathcal{L}(F/G)$ satisfaisant P .*

Principe général : soit $\mathcal{P}$ est un ensemble de propriétés stables par quotients et vérifiant la propriété suivante : pour tout $\mathcal{E} \subset \mathcal{L}(E)$ constitué d'éléments vérifiant $\mathcal{P}$ avec $\dim E > 1$, $\mathcal{E}$ est réductible i.e. il existe un sous-espace vectoriel non trivial F de E stable par tous les éléments de $\mathcal{E}$. Alors $\mathcal{E}$ est triangularisable.

Exemple : pour E un $\mathbb{C}$ -espace vectoriel de dimension finie, tout sous-ensemble commutatif de $\mathcal{L}(E)$ est triangularisable. En effet la commutativité est clairement une propriété stable par quotient. La propriété de réductibilité découlera alors des faits suivants :

- tout endomorphisme admet une valeur propre et
- tout sous-espace propre de A est stable par toute matrice B commutant avec A .

Remarque 82. en utilisant les faits suivants :

- les valeurs propres d'une matrice triangulaires sont ses termes diagonaux ;
- les termes diagonaux d'un produit (resp. somme) de matrice triangulaire sont les produits $t_{i,i}^{(1)} t_{i,i}^{(2)}$ (resp. $t_{i,i}^{(1)} + t_{i,i}^{(2)}$) de leurs termes diagonaux.

on en déduit que si $\{A_1, \dots, A_k\}$ sont simultanément diagonalisables et si p est un polynôme non commutatif (i.e. une combinaison linéaire de mots) en $\{A_1, \dots, A_k\}$ alors

$$\sigma(p(A_1, \dots, A_k)) \subset p(\sigma(A_1), \dots, \sigma(A_k))$$

où $p(\sigma(A_1), \dots, \sigma(A_k))$ désigne l'ensemble des $p(\lambda_1, \dots, \lambda_k)$ où pour tout $i = 1, \dots, k$, $\lambda_i \in \sigma(A_i)$. Dans la suite nous allons donner la réciproque à ce résultat.

Si $\mathcal{A}$ est une sous-algèbre de $\mathcal{L}(E)$ l'ensemble $\mathcal{A}.x := \{Ax : A \in \mathcal{A}\}$, où $x \in E$, est un sous-espace stable par $\mathcal{A}$. Si $\mathcal{A}.x = E$, on dit que x est un *vecteur cyclique* pour $\mathcal{A}$. La détermination des sous-algèbres de $\mathcal{L}(E)$ qui possèdent des sous-espaces invariants non triviaux est réglée par le théorème suivant qui s'occupe de la partie réductibilité du principe général énoncé plus haut dans le cas des sous-algèbres de $\mathcal{L}(E)$.

Théorème 83. (de Burnside)

Toute sous-algèbre propre de $\mathcal{L}(E)$ est réductible.

Preuve : Soit $\mathcal{A}$ une sous-algèbre irréductible de $\mathcal{L}(E)$; comme tout endomorphisme est une somme d'endomorphisme de rang 1, nous allons montrer que tout endomorphisme de rang 1 appartient à $\mathcal{A}$.

Montrons tout d'abord que $\mathcal{A}$ contient un élément de rang 1. Soit $u_0 \in \mathcal{A}$ non nul de rang minimal ; si ce rang est strictement plus grand que 1, alors il existe des vecteurs x_1 et x_2 tels que $(u_0(x_1), u_0(x_2))$ est linéairement indépendant. Comme $\{u \circ u_0(x_1) : u \in \mathcal{A}\} = E$, il existe $u_1 \in \mathcal{A}$ tel que $u_1 \circ u_0(x_1) = x_2$ et donc $(u_0 \circ u_1 \circ u_0(x_1), u_0(x_1))$ est libre. Soit alors λ tel que la restriction de $u_1 \circ u_0 - \lambda \text{Id}$ à $u_0(E)$ n'est pas inversible ; $(u_0 \circ u_1 - \lambda \text{Id})u_0$ est non nul car l'image de x_1 est non nulle, et $(u_0 \circ u_1 - \lambda \text{Id}) \circ u_0$ est de rang strictement plus petit que celui de u_0 , d'où la contradiction et donc u_0 est de rang 1.

Pour y_0 dans l'image de u_0 , on considère la forme linéaire φ_0 définie par $u_0(x) = \varphi_0(x)y_0$. Soit alors $u \in \mathcal{L}(E)$ défini par $u(x) = \varphi(x)y$ où $y \in E$ et $\varphi \in E^*$. Montrons alors que u appartient à $\mathcal{A}$. Pour $v \in \mathcal{A}$, on a $u_0 \circ v \in \mathcal{A}$ et $u_0 \circ v(x) = \varphi_0(v(x))y_0$. Soit alors $F' \subset E^*$, l'ensemble des formes linéaires φ telles que $x \mapsto \varphi(x)y_0$ appartienne à $\mathcal{A}$: F' est clairement un sous-espace de E^* . Si ce sous-espace était strict, il existerait $x_0 \neq 0$ tel que $\varphi(x_0) = 0$ pour tout $\varphi \in F'$ (un espace vectoriel de dimension finie est réflexif). La contradiction découle alors du fait que $\varphi_0(v(x_0)) = 0$ pour tout $v \in \mathcal{A}$ implique que x_0 est nul car $\{v(x_0) : v \in \mathcal{A}\} = E$. Soit donc $v_1 \in \mathcal{A}$ tel que $\varphi = \varphi_0 \circ v_1$.

De même comme $y_0 \neq 0$, alors $\{v(x_0) : v \in \mathcal{A}\} = E$ et donc pour tout $y \in E$ soit $v_2 \in \mathcal{A}$ tel que $v_2(y_0) = y$ et donc $u = v_2 \circ u_0 \circ v_1$.

Corollaire 84. *Les seuls idéaux bilatères de $\mathcal{L}(E)$ sont $\{0\}$ et $\mathcal{L}(E)$.*

Preuve : Soit $\mathcal{I}$ un idéal bilatère de $\mathcal{L}(E)$ non réduit à 0. Il suffit alors de montrer que $\mathcal{I}$ est irréductible. Si $u \neq 0$ appartient à $\mathcal{I}$, pour tout $0 \neq x \in E$, il existe $v \in \mathcal{L}(E)$ tel que $u \circ v(x) \neq 0$. Soit $y \in E$ et $w \in \mathcal{L}(E)$ tel que $w \circ u \circ v(x) = y$. On a $w \circ u \circ v \in \mathcal{I}$ de sorte que tout vecteur $x \neq 0$ est cyclique pour $\mathcal{I}$ et donc $\mathcal{I}$ est irréductible.

Corollaire 85. *Soit E est $\mathbb{C}$ -espace vectoriel de dimension finie alors tout automorphisme d'algèbre ϕ de $\mathcal{L}(E)$ est intérieur, i.e. il existe $P \in GL(E)$ tel que pour tout $A \in \mathcal{L}(E)$, $\phi(A) = PAP^{-1}$.*

Preuve : Soit $A_0 \in \mathcal{L}(E)$ un idempotent de rang 1, $\phi(A_0)$ est alors un idempotent, montrons qu'il est aussi de rang 1. L'ensemble $\{A_0 B A_0 : B \in \mathcal{L}(E)\}$ est un sous-espace vectoriel de $\mathcal{L}(E)$ de dimension 1 : on peut l'identifier avec $\mathcal{L}(\text{Im } A_0)$. Son image par ϕ , $\{\phi(A_0) C \phi(A_0) : C \in \mathcal{L}(E)\}$ est donc aussi un sous-espace de $\mathcal{L}(E)$ de dimension 1 identifié à $\mathcal{L}(\text{Im } \phi(A_0))$ de sorte que $\phi(A_0)$ est de rang 1. Comme tous les idempotents de rang 1 sont semblables à $\text{diag}(1, 0, \dots, 0)$, quitte à composer ϕ par $A \mapsto P A P^{-1}$, on peut supposer que $\phi(A_0) = A_0$.

Notons x_0 un vecteur directeur de $\text{Im } A_0$ et soit $P \in \mathcal{L}(E)$ défini par $P(Bx_0) = \phi(B)x_0$: si $B_1 x_0 = B_2 x_0$ alors comme $A_0 x_0 = x_0$, on a $(B_1 - B_2)A_0 = 0$ et donc $(\phi(B_1) - \phi(B_2))A_0 = 0$ de sorte que $\phi(B_1)x_0 = \phi(B_2)x_0$ et P est bien définie et évidemment linéaire. Supposons que $\phi(B)x_0 = 0$ de sorte que $\phi(B)\phi(A_0) = \phi(BA_0) = 0$ et donc $BA_0 = 0$ soit $Bx_0 = 0$ ce qui prouve l'injectivité de P et comme on est en dimension finie $P \in GL(E)$.

Soit alors $A \in \mathcal{L}(E)$, on a $P(AB)x_0 = \phi(AB)x_0 = \phi(A)\phi(B)x_0 = \phi(A)PBx_0$ et donc $PAy = \phi(A)Py$ pour tout $y = Bx_0$. Quand B décrit $\mathcal{L}(E)$, y décrit E et donc $PA = \phi(A)P$ pour tout $A \in \mathcal{L}(E)$ d'où le résultat.

Corollaire 86. *Toute algèbre d'endomorphismes nilpotents est triangularisable.*

Preuve : La propriété d'être nilpotent est stable par quotient comme en outre il existe des éléments de $\mathcal{L}(E)$ qui ne sont pas nilpotents, toute algèbre constituée d'endomorphismes nilpotents est, d'après le théorème de Burnside, réductible. La triangularisation découle alors du principe général énoncé plus haut.

Théorème 87. *Si $\mathcal{A}$ est une sous-algèbre de $\mathcal{L}(E)$ alors $\mathcal{A}$ est triangularisable si et seulement si tout commutateur $BC - CB$ avec $B, C \in \mathcal{A}$ est nilpotent.*

Preuve : Si $\mathcal{A}$ est triangularisable alors d'après la remarque 82, on a $\sigma(BC - CB) = \{0\}$ et donc $BC - CB$ est nilpotent. Réciproquement la propriété d'avoir des commutateurs nilpotents est stable par quotient et comme il existe des commutateurs non nilpotents dans $\mathcal{L}(E)$, d'après le théorème de Burnside, $\mathcal{A}$ est réductible; la triangularisation découle alors du principe général.

Remarque: on peut ainsi voir la triangularisabilité comme une généralisation de la commutativité; on relâche la condition d'être nul pour un commutateur, en demandant qu'il soit nilpotent.

Théorème 88. ((McCoy)) *La paire $\{A, B\}$ est triangularisable si et seulement si $p(A, B)(AB - BA)$ est nilpotent pour tout polynôme commutatif p en A et B .*

Preuve : Le sens direct découle de la remarque 82. Pour la réciproque d'après le principe général il suffit de montrer que l'algèbre $\mathcal{A}$ engendrée par

A, B est réductible dès que $\dim E > 1$. Si $AB = BA$ c'est clair, sinon soit $x \in E$ tel que $(AB - BA)x \neq 0$ et $C \in \mathcal{L}(E)$ vérifiant $C(AB - BA)x = x$. Si $\mathcal{A}$ était irréductible alors d'après le théorème de Burnside elle serait égale à $\mathcal{L}(E)$ et donc $C \in \mathcal{A}$ et donc de la forme $p(A, B)$. La contradiction découle alors du fait que $C(AB - BA)$ n'est pas nilpotent.

Corollaire 89. *Une sous-algèbre unitaire $\mathcal{A}$ de $\mathcal{L}(E)$ est triangularisable si et seulement si $\mathcal{A}/\text{Rad}\mathcal{A}$ est commutatif, où*

$$\text{Rad}\mathcal{A} = \{A \in \mathcal{A} : \sigma(AB) \subset \{0\} \ \forall B \in \mathcal{A}\}$$

est le radical de $\mathcal{A}$, i.e. l'intersection de tous les idéaux à droite (ou à gauche) maximaux de $\mathcal{A}$.

Preuve : Si $\mathcal{A}$ est triangularisable et si $B, C \in \mathcal{A}$ alors pour tout $A \in \mathcal{A}$ d'après la remarque 82, on a $\sigma((BC - CB)A) = \{0\}$ et donc $BC - CB \in \text{Rad}\mathcal{A}$ i.e. $\mathcal{A}/\text{Rad}\mathcal{A}$ est commutatif.

Réciproquement si $\mathcal{A}/\text{Rad}\mathcal{A}$ est commutatif alors $BC - CB \in \text{Rad}\mathcal{A}$ pour tout $B, C \in \mathcal{A}$ de sorte que $\mathcal{A}$ est triangularisable d'après le corollaire 86.

Nous allons nous intéresser à des sous-espaces vectoriels de $\mathcal{L}(E)$ stable par certaines multiplications non associatives comme par exemple les *algèbres de Lie* stables sous le crochet de Lie $[A, B] = AB - BA$, ou les algèbre de Jordan stable sous $(A, B) \mapsto AB + BA$. Le résultat le plus célèbre est le théorème de Engel ci-dessous.

Théorème 90. *Soit $\mathcal{N}$ un sous-ensemble du cône nilpotent de $\mathcal{L}(E)$ vérifiant la propriété suivante : pour tout $A, B \in \mathcal{N}$ il existe un polynôme non commutatif p en A et B tel que $AB + p(A, B)A \in \mathcal{N}$. Alors $\mathcal{N}$ est triangularisable.*

Preuve : On raisonne par récurrence sur la dimension n de E ; le cas $n = 1$ étant trivial supposons donc le résultat acquis jusqu'au rang n et traitons celui de $n + 1$. Soit $\mathcal{F}$ l'ensemble des sous-espaces de E qui sont des intersections de noyaux d'éléments de $\mathcal{N}$, i.e. de la forme $V_{\mathfrak{S}} = \bigcap_{A \in \mathfrak{S}} \text{Ker } A$ où $\mathfrak{S} \subset \mathcal{N}$, et soit $K \in \mathcal{F}$ de dimension minimale non nulle. Notons alors $\mathcal{N}_0 = \{A \in \mathcal{N} : Ax = 0 \ \forall x \in K\}$; l'ensemble $\overline{\mathcal{N}}_0 \subset \mathcal{L}(E/K)$ vérifie les hypothèses de l'énoncé de sorte que d'après l'hypothèse de récurrence $\overline{\mathcal{N}}_0$ est triangularisable et donc $\mathcal{N}_0$ aussi car ses éléments sont nuls sur K .

Il suffit alors de montrer que $\mathcal{N}_0 = \mathcal{N}$. Dans le cas contraire soit $B \in \mathcal{N}$ et $x \in K$ tels que $Bx \neq 0$. Si K était un sous-espace stable de l'endomorphisme nilpotent B , il existerait $x_0 \in K$ tel que $Bx_0 = 0$ et $\text{Ker } B \cap K$ serait un élément non nul de $\mathcal{F}$ de dimension strictement plus petite que celle de K ce qui n'est pas par hypothèse. Soit alors $x_1 \in K$ et $A_1 \in \mathcal{N}_0$ tel que $A_1 B x_1 \neq 0$ et soit p_1 un polynôme non commutatif en A_1 et B tel que $B_1 = A_1 B + p_1(A_1, B)A_1 \in \mathcal{N}$. Comme $B_1 x_1 \neq 0$ comme précédemment K

n'est pas un sous-espace stable de B_1 ; soit alors $A_2 \in \mathcal{N}_0$ et $x_2 \in K$ tels que $A_2 B_1 x_2 \neq 0$. Soit p_2 tel que $B_2 = A_2 B_1 + p_2(A_2, B_1)A_2 \in \mathcal{N}$. En continuant le processus, on construit

$$\{A_1, A_2, \dots, A_{n+1}\} \subset \mathcal{N}_0, \quad \{B_1, B_2, \dots, B_n\} \subset \mathcal{N}$$

et des vecteurs $\{x_1, \dots, x_{n+1}\}$ tels que

$$A_{n+1}A_n \cdots A_2 A_1 B x_{n+1} = A_{n+1}B_n x_{n+1} \neq 0.$$

Comme $\mathcal{N}_0$ est triangularisable, tout produit de $n+1$ de ses éléments est nul de sorte que $A_{n+1} \cdots A_1 = 0$ ce qui contredit $A_{n+1} \cdots A_1 B x_{n+1} \neq 0$. On en déduit donc que $\mathcal{N}_0 = \mathcal{N}$ qui est trigonalisable.

Corollaire 91. *Un ensemble $\mathcal{N}$ d'éléments nilpotents de $\mathcal{L}(E)$ qui vérifie une des propriétés suivantes est trigonalisable :*

- **théorème de Jacobson** : pour tout $A, B \in \mathcal{N}$, il existe un scalaire c tel que $AB - cBA \in \mathcal{N}$;
- **théorème de Engel** : pour tout $A, B \in \mathcal{N}$, $[A, B] = AB - BA \in \mathcal{N}$;
- pour tout $A, B \in \mathcal{N}$, $AB + BA \in \mathcal{N}$.

Corollaire 92. *Soit $\mathcal{E} \subset \mathcal{L}(E)$ stable par le crochet de Lie. Alors $\mathcal{E}$ est triangularisable si et seulement si tous ses commutateurs sont nilpotents.*

Preuve : Le sens direct découle de la remarque 82. Réciproquement comme la propriété est clairement stable par quotient d'après le principe général il suffit de montrer que $\mathcal{E}$ admet un sous-espace stable. Soit $\mathcal{N}$ l'ensemble des commutateurs de $\mathcal{E}$; si $\mathcal{N} = \{0\}$ alors $\mathcal{E}$ est commutatif admet donc un sous-espace stable. Sinon d'après le corollaire précédent $\mathcal{N}$ est trigonalisable de sorte que $K = \bigcap_{N \in \mathcal{N}} \text{Ker } N$ est un sous-espace non nul stable par tous les éléments de $\mathcal{E}$. En effet pour $x \in K$ et $B \in \mathcal{E}$, pour tout $A \in \mathcal{N}$, on a $Ax = 0$ et $(AB - BA)x = 0$ de sorte que $ABx = 0$ soit $Bx \in K$, d'où le résultat.

Théorème 93. (Levitzki) *Tout semi-groupe S d'éléments nilpotents de $\mathcal{L}(E)$ est triangularisable.*

Preuve : La nilpotence étant une propriété stable par quotient, il suffit d'après le principe général de montrer la réductibilité de S dès que $\dim E > 1$. La trace est une forme linéaire qui s'annule sur S et donc sur l'algèbre engendrée par S qui est simplement l'espace vectoriel engendré par S . Le résultat découle alors du théorème de Burnside et du fait qu'il existe des éléments de $\mathcal{L}(E)$ de trace non nulle.

Remarque: on aurait aussi pu déduire ce résultat du théorème de Jacobson pour $c = 0$.

Théorème 94. (Kolchin) *Si tout élément du semi-groupe S est unipotent alors S est triangularisable.*

Preuve : L'unipotence étant une propriété stable par quotient il suffit d'après le principe général de montrer la réductibilité de S dès que $\dim E > 1$. Si tous les commutateurs sont nuls alors S est abélien et donc réductible. Sinon soit $C = AB - BA \neq 0$ et soit $\mathcal{A}$ l'algèbre engendrée par S . L'idéal bilatère engendré par C est alors contenu dans le noyau de la forme linéaire trace : en effet XCY s'écrit comme une combinaison linéaire de produit DCE avec $D, E \in S$ de sorte que $\text{tr}(DCE) = \text{tr}(DABE) - \text{tr}(DBAE) = n - n = 0$. Comme $\mathcal{L}(E)$ n'admet pas d'idéal propre non nul, on en déduit que $\mathcal{A}$ n'est pas égal à $\mathcal{L}(E)$ et est donc réductible d'après le théorème de Burnside.

Si $\mathcal{E} \subset \mathcal{L}(E)$ est triangularisable alors la fonction trace est permutable sur $\mathcal{E}$, i.e. pour tout $A_1, \dots, A_m \in \mathcal{E}$ et pour toute permutation $\sigma \in \mathfrak{S}_m$, on a

$$\text{tr}(A_1 A_2 \cdots A_m) = \text{tr}(A_{\sigma(1)} A_{\sigma(2)} \cdots A_{\sigma(m)}).$$

La réciproque est vraie et découle simplement du théorème 87.

Proposition 95. *Soit $\mathcal{E} \subset \mathcal{L}(E)$, alors $\mathcal{E}$ est triangularisable si et seulement si la trace est permutable sur $\mathcal{E}$.*

Preuve : Pour tout $A, B, C \in \mathcal{E}$, on a $\text{tr}((AB - BA)C) = 0$. En particulier on en déduit que pour tout $k > 0$, $\text{tr}(AB - BA)^k = 0$ et donc $AB - BA$ est nilpotent d'où le résultat d'après 87.

Corollaire 96. *Si S est un semi-groupe qui vérifie une des propriétés suivantes, est triangularisable :*

- (i) **Kaplansky** : la trace est constante sur S ;
- (ii) la trace est multiplicative sur S .

En outre toutes dans la première situation les termes diagonaux dans une telle triangularisation ne dépendent que de S et sont soit égaux à 0 ou 1 ; dans la deuxième il existe j qui ne dépend que de S tels que tous les termes diagonaux dans une triangularisation, hors celui en (j, j) , sont nuls.

Preuve : (i) La triangularisabilité découle de la proposition précédente. Soit alors $A \in S$ de sorte que les $\text{tr} A^k$ sont constants pour tout $k \geq 1$ égaux à c . Notons $\lambda_1, \dots, \lambda_m$, les valeurs propres non nulles de A ; il suffit alors de montrer que pour tout $\text{tr} A = m$. En effet d'après les relations de Newton, l'ensemble $\{\lambda_1, \dots, \lambda_m\}$ est déterminé uniquement par les $S_k = \lambda_1^k + \dots + \lambda_m^k$ pour $k = 1, \dots, m$; si tous les S_k sont égaux à 1 alors $\lambda_1 = \dots = \lambda_m = 1$ convient trivialement. Notons σ_i les fonctions symétriques usuelles des λ_j de sorte que l'on a

$$S_{m+1} - S_m \sigma_1 + S_{m-1} \sigma_2 + \cdots + (-1)^m S_1 = 0 \sigma_n$$

$$S_m - \sigma_1 S_{m-1} + \cdots + (-1)^m m \sigma_m = 0$$

de sorte que comme $\sigma_m \neq 0$, on obtient $c = m$.

Ainsi les valeurs propres de A sont 0 ou 1 ; si les termes diagonaux de $A, B \in S$ n'étaient pas égaux alors on aurait $\text{tr}(ST) < \text{tr}S$ ce qui n'est pas.

(ii) La triangularisabilité découle de la proposition précédente. Soit $A \in S$ de valeurs propres $\lambda_1, \dots, \lambda_n$. Par hypothèse on a $\sum_i \lambda_i^k = (\sum_i \lambda_i)^k$. Si $\sum_i \lambda_i = 0$ alors tous les λ_i sont nuls d'où le résultat. Si $\sum_i \lambda_i = b \neq 0$ alors pour tout $k \geq 1$, $\sum_i (a_i/b)^k = 1$ de sorte que d'après la preuve de (i), il existe un unique j tel que $\lambda_j \neq 0$. La multiplicativité de la trace implique que ce j est le même pour tous les $A \in S$.

Corollaire 97. *Soit G un sous-groupe de $GL(E)$; les propriétés suivantes sont alors équivalentes :*

- (i) G est triangularisable ;
- (ii) pour tout $g \in G$, la trace est constante sur $gD(G)$;
- (iii) la trace est constante sur $D(G)$;
- (iv) pour tout $A \in D(G)$, $\sigma(A) = \{1\}$, i.e. A est unipotent.

Preuve : (i) $\Rightarrow$ (ii) : d'après la proposition précédente, la trace est permutable sur G de sorte que pour tout $g \in G$ et $h \in D(G)$, on a $\text{tr}(gh) = \text{tr}(g\text{Id}) = \text{tr}g$.

(ii) $\Rightarrow$ (iii) : immédiat

(iii) $\Leftrightarrow$ (iv) : découle du théorème de Kaplansky prouvé ci-dessus en utilisant que les éléments de G sont inversibles.

(iv) $\Rightarrow$ (i) : comme la propriété (iv) est clairement stable par quotient, il suffit donc de montrer que G est réductible. D'après le théorème de Kolchin, $D(G)$ est triangularisable : soit $\{0\} = F_0 \subset F_1 \subset \dots \subset F_{n-1} \subset F_n = E$ le drapeau complet correspondant. Supposons $D(G) \neq \{\text{Id}\}$ car sinon G est commutatif et le résultat est clair. Soit F le sous-espace vectoriel engendré par $\bigcup_{h \in D(G)} (h - \text{Id})(E)$. Comme pour tout $h \in D(G)$, $(h - \text{Id})(E) \subset F_{n-1}$, F est un sous-espace strict de E qui de plus est invariant sous G : en effet soit $g \in G$ alors pour tout $h \in D(G)$, on a $g(h - \text{Id}) = (ghg^{-1} - \text{Id})g$ et comme $ghg^{-1} \in D(G)$, on a $g((h - \text{Id})(E)) \subset (ghg^{-1} - \text{Id})(E) \subset F$, d'où le résultat.

Corollaire 98. *Soit G un sous-groupe de $GL(E)$ tel que pour tout $A, B, C \in G$, $\sigma(ABC) = \sigma(BAC)$ alors G est triangularisable.*

Remarque : comme $\sigma(AB) = \sigma(BA)$, la propriété de l'énoncé revient à dire que le spectre est permutable, i.e. pour tout $\{A_1, \dots, A_m\} \subset G$, et pour toute permutation $\sigma \in \mathfrak{S}_m$, on a $\sigma(A_1 A_2 \dots A_m) = \sigma(A_{\sigma(1)} \dots A_{\sigma(m)})$. Le lecteur notera bien que cette propriété plus faible que celle de permutabilité de la trace car ici on demande juste que les ensembles de valeurs propres sont égaux, sans tenir compte des multiplicités.

Preuve : Le résultat découle directement de l'implication (iv) $\Rightarrow$ (i) dans le corollaire précédent. En effet pour tout $h \in D(G)$, on a $\sigma(h) = \sigma(\text{Id}) = \{1\}$ de sorte que h est unipotent.

3.5 Noyaux emboîtés

Soient E un $\mathbb{K}$ -espace vectoriel de dimension finie n et $u \in \mathcal{L}(E)$. Pour tout $\lambda \in \mathbb{K}$ et $r \geq 1$, on note

$$K_r(\lambda) := \text{Ker}(u - \lambda \text{Id})^r \quad \text{et} \quad I_r(\lambda) := \text{Im}(u - \lambda \text{Id})^r,$$

et on note $dK_r(\lambda) := \dim_{\mathbb{K}} K_r(\lambda)$ et $dI_r(\lambda) := \dim_{\mathbb{K}} I_r(\lambda)$. On pose aussi $dK_0(\lambda) = 0$ et $dI_0(\lambda) = n$.

Proposition 99. *La suite $dK_r(\lambda)$ (resp. $dI_r(\lambda)$) est tout d'abord strictement croissante (resp. décroissante) puis stationnaire à partir d'un indice r_0 (resp. le même indice r_0). Par ailleurs la suite*

$$\delta_r(\lambda) := dK_r(\lambda) - dK_{r-1}(\lambda)$$

pour $r \geq 1$ est décroissante jusqu'au rang r_0 puis stationnaire égale à 0.

Preuve : Quitte à considérer $u - \lambda \text{Id}$, on suppose $\lambda = 0$ et on note simplement K_r pour $K_r(0)$. Soit alors r tel que $K_r = K_{r+1} \subset K_{r+2}$. Pour $x \in K_{r+2}$ on a $u(x) \in K_{r+1} = K_r$ et donc $u^{r+1}(x) = 0$, i.e. $x \in K_{r+1}$ et donc $K_{r+2} = K_{r+1}$, ce qui montre la première partie de l'énoncé puisqu'avec le théorème du rang $dK_r(\lambda) + dI_r(\lambda) = n$.

En ce qui concerne δ_r , considérons l'endomorphisme $K_r \rightarrow K_{r-1}/K_{r-2}$ qui envoie x sur l'image de $u(x) \in K_{r-1}/K_{r-2}$. Son noyau est clairement K_{r-1} de sorte qu'on a une injection

$$K_r/K_{r-1} \hookrightarrow K_{r-1}/K_{r-2}$$

et donc $\delta_r \geq \delta_{r-1}$, d'où le résultat.

Proposition 100. *Soit $u \in \mathcal{L}(E)$ dont le polynôme caractéristique est scindé. On note*

$$E = \bigoplus_{\lambda \in \text{Spec}(u)} E(\lambda)$$

la décomposition de E en sous-espace caractéristiques. Pour chaque $\lambda \in \text{Spec}(u)$, il existe une base de $E(\lambda)$ dans laquelle la matrice de u est diagonale par blocs du type

$$J_k(\lambda) = \begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \lambda & 1 \\ 0 & \cdots & \cdots & 0 & \lambda \end{pmatrix}$$

où la taille des blocs sont données par les longueurs des lignes du tableau de Young, cf. la figure ??, associé à u et λ dont les colonnes sont de taille $\delta_r(\lambda)$.

Une façon de construire ce tableau de Young est la suivante : on prend un vecteur e_1 de $K_r - K_{r-1}$ et on note pour $i = 1, \dots, r-1$, $e_{i+1} = u^i(e_1)$. Si $\dim K_r/K_{r-1} > 1$, on choisit un vecteur $e_{r+1} \in K_r$ tel que les images de e_1, e_{r+1} dans K_r/K_{r-1} soient libres et on pose pour $i = 1, \dots, r-1$, $e_{r+1+i} = u^i(e_{r+1})$. On continue le procédé jusqu'à obtenir une base $e_1, e_{r+1}, \dots, e_{kr+1}$ de K_r/K_{r-1} . On choisit alors un vecteur $e_{(k+1)r+1}$ de K_{r-1} tel que les images de $u(e_1), \dots, u(e_{kr+1}), e_{(k+1)r+1}$ forment une famille libre de K_{r-1}/K_{r-2} et on pose pour tout $i = 1, r-2$, $e_{(k+1)r+1+i} = u(e_{(k+1)r+1})$. On continue ce procédé jusqu'à épuiser tous les K_i . Parallèlement on remplit le tableau de Young comme dans la figure ?? dans laquelle l'image de e_1 est une base de $\text{Ker } u^6 / \text{Ker } u^5$, les images de $u(e_1), e_7, e_{12}$ forment une base de $\text{Ker } u^5 / \text{Ker } u^4$, les images de $u^4(e_1), u^3(e_7), u^3(e_{12}), e_{17}$ forment une base de $\text{Ker } u^2 / \text{Ker } u$ et

$$u^5(e_1), u^4(e_7), u^4(e_{12}), u(e_{17}), e_{19}$$

forment une base de $\text{Ker } u$.

Définition 101. La maison de u est l'ensemble des tableaux de Young de u pour $\lambda \in \text{Spec}(u)$.

Remarque: sur un corps algébriquement clos, la classe de conjugaison d'un endomorphisme correspond à sa maison au sens de la définition précédente. Une application intéressante de ce résultat est le théorème de Brauer suivant.

Théorème 102. (de Brauer) Soit K un corps quelconque et pour $\sigma \in \mathfrak{S}_n$, on note $M(\sigma)$ la matrice de permutation associée, i.e. définie par $M(\sigma)(e_i) = e_{\sigma(i)}$ pour tout $i = 1, \dots, n$. Alors σ et σ' sont conjugués dans $\mathfrak{S}_n$ si et seulement si $M(\sigma)$ et $M(\sigma')$ sont semblables dans $GL_n(K)$.

Preuve : Le sens direct est évident puisque si $\sigma' = \tau\sigma\tau^{-1}$ alors $M(\sigma') = PM(\sigma)P^{-1}$ avec $P = M(\tau)$.

Réciproquement supposons que $M(\sigma)$ et $M(\sigma')$ sont semblables. Notons $V^\sigma = \text{Ker}(M(\sigma) - \text{Id})$ l'espace des invariants sous $M(\sigma)$, i.e. l'ensemble des $v = \sum_{i=1}^n \lambda_i e_i$ tels que λ_i ne dépend que de l'orbite de i sous l'action de σ . Ainsi $\dim V^\sigma$ est égal au nombre d'orbites sous σ . Soit alors $\sigma = c_1 \cdots c_r$ la décomposition en cycles à supports disjoints de σ et notons, pour $k = 1, \dots, n$, $n_k(\sigma)$ le nombre de cycles c_i de longueurs k . On a alors

$$\dim V^\sigma = \sum_{k=1}^n n_k(\sigma) = \sum_{k=1}^n n_k(\sigma') = \dim V^{\sigma'}.$$

De même comme $M(\sigma^r) = M(\sigma)^n$, on a aussi $\dim V^{\sigma^r} = \dim V^{(\sigma')^r}$ et comme, si c est un cycle de longueur k alors c^r s'écrit comme le produit de $k \wedge r$ - cycles à supports disjoints tous de même longueur $\frac{k}{k \wedge r}$, on en déduit que pour tout r , on a

$$\sum_{k=1}^n (k \wedge r) n_k(\sigma) = \sum_{k=1}^n (k \wedge r) n_k(\sigma'),$$

ce qui s'écrit matriciellement $SX = SX'$ où $S := (i \wedge j)_{1 \leq i, j \leq n}$ est la matrice des pgcd et X (resp. X') est la matrice colonne des $n_k(\sigma)$ (resp. $n_k(\sigma')$). Soit alors $A = (a_{i,j})_{1 \leq i, j \leq n}$ la matrice définie par $a_{i,j} = 1$ si j divise i et 0 sinon. La relation $\sum_{d|m} \varphi(d) = m$ s'écrit alors matriciellement sous la forme

$$\text{Adiag}(\varphi(1), \dots, \varphi(n))^t A = S$$

de sorte que la matrice S est inversible et donc $X = X'$, i.e. σ et σ' ont « la même » décomposition en cycles à supports disjoints et sont donc conjugués.

3.6 Endomorphismes cycliques

Soit E un K -espace vectoriel de dimension finie.

Définition 103. Un endomorphisme $f \in \mathcal{L}(E)$ est dit cyclique si et seulement s'il existe $v \in E$ tel que $E = \{P(f)(v) : P \in K[X]\}$.

Remarque: si n est la dimension de E alors $E = \{P(f)(v) : P \in K_{n-1}[X]\}$ et comme $(1, X, \dots, X^{n-1})$ est une base de $K_{n-1}[X]$, on en déduit que $(v, f(v), \dots, f^{n-1}(v))$ est une base de V . Dans cette base la matrice de f est de la forme

$$\begin{pmatrix} 0 & \cdots & \cdots & 0 & a_0 \\ 1 & 0 & & \vdots & a_1 \\ 0 & \ddots & \ddots & & \vdots \\ \cdots & & & 1 & 0 & a_{n-2} \\ 0 & \cdots & \cdots & 1 & a_{n-1} \end{pmatrix}.$$

On vérifie aisément que le polynôme caractéristique de cette matrice est $P(X) = X^n - a_{n-1}X^{n-1} - \dots - a_0$ et on dit que la matrice précédente est la matrice compagnon de $P(X) = X^n - a_{n-1}X^{n-1} - \dots - a_0$.

Lemme 104. Le polynôme minimal d'un endomorphisme cyclique est égal à son polynôme caractéristique.

Preuve : Comme $v, f(v), \dots, f^{n-1}(v)$ est une famille libre, tout polynôme Q de degré $\leq n-1$ vérifie alors $Q(f)(v) \neq 0$, de sorte que le polynôme minimal est de degré $\geq n$. Comme par ailleurs il divise le polynôme caractéristique, lequel est de degré n , on en déduit qu'il lui est égal.

Proposition 105. Soit $g \in \mathcal{L}(E)$ tel que $gf = fg$; il existe alors $P \in K[X]$ tel que $g = P(f)$.

Remarque: autrement dit le commutant d'un endomorphisme cyclique est $\{P(f) : P \in \mathbb{K}[X]/(\pi_f)\}$.

Preuve : On écrit $g(v) = \alpha_0 v + \dots + \alpha_{n-1} f^{n-1}(v)$ et on pose $Q(X) = \alpha_0 + \dots + \alpha_{n-1} X^{n-1}$. Pour vérifier l'égalité $g = Q(f)$ il suffit de la vérifier sur la base $(v, f(v), \dots, f^{n-1}(v))$, soit

$$g(f^i(v)) = f^i(g(v)) = f^i(Q(f)(v)) = Q(f)(f^i(v)),$$

d'où le résultat.

3.7 Invariants de similitude

Soit V un K -espace vectoriel de dimension finie et $f \in \mathcal{L}(E)$. L'idée nouvelle est alors de considérer V muni de l'endomorphisme f , comme un $K[X]$ -module comme suit, et d'utiliser le théorème de structure des modules de type fini sur un anneau principal, cf. le §??.

Définition 106. *On munit V d'une structure de $K[X]$ -module en posant $X.v := f(v)$ et par linéarité pour tout $P \in K[X]$, on a $P.v = P(f)(v)$. On notera V_f l'espace V muni de cette structure de $K[X]$ -module.*

Proposition 107. *Deux endomorphismes f et g sont semblables si et seulement si les deux structures de $K[X]$ -module induites sur V sont isomorphes, i.e. $V_f \simeq V_g$.*

Preuve : Soit $h \in \mathcal{L}(E)$ tel que $g = hf h^{-1}$ alors pour $v \in V_f$, on a $h(X.v) = hf(v) = g(h(v)) = X.h(v)$, i.e. h induit un isomorphisme de $K[X]$ -modules : $V_f \simeq V_g$.

Réciproquement si $h : V_f \simeq V_g$ alors $h(X.v) = X.h(v)$, i.e. $h(f(v)) = g(h(v))$ et donc $hf = gh$ soit $g = hf h^{-1}$ et donc f et g sont semblables.

De la théorie générale des modules sur un anneau principal, cf. le théorème 57, on en déduit le théorème suivant qu'il faut comprendre comme une décomposition de l'espace en somme directe de sous-espaces cycliques.

Théorème 108. *Il existe une unique suite de polynômes non constants $P_1(X) | \dots | P_r(X)$ tels que*

$$V_f \simeq K[X]/(P_1) \times \dots \times K[X]/(P_r). \quad (3)$$

En particulier on a $P_r = \pi_f$ et $P_1 \dots P_r = \chi_f$.

Exemple : si $V_f \simeq K[X]/(X^n)$ alors la matrice de f dans la base d'image $(1, X, \dots, X^{n-1})$ est la matrice de Jordan $J_n(0)$. Ainsi si P_r est totalement décomposé, par application du lemme chinois, on passe de la décomposition (3) à celle de Jordan.

Proposition 109. (Décomposition de Dunford) *Soit $f \in \mathcal{L}(E)$ tel que son polynôme caractéristique est séparable. Alors f s'écrit de manière unique sous la forme $d + n$ où*

- n est nilpotent et d est semi-simple,
- d et n commutent.

En outre d et n sont des polynômes en f .

Remarque: d'après la définition ??, un polynôme est séparable si tous ses facteurs irréductibles le sont. Par ailleurs un polynôme irréductible est séparable si et seulement si ses racines sont simples dans son corps de décomposition qui est alors une extension galoisienne du corps de départ. Rappelons enfin qu'en caractéristique nulle tout polynôme est séparable.

Remarque: comme le polynôme caractéristique et le polynôme minimal de f ont les mêmes facteurs irréductibles, ils sont soit tous les deux séparables ou pas.

Preuve : Commençons par l'existence : en utilisant (3) et le lemme chinois, il suffit de traiter le cas où f est cyclique avec $\chi_f = \pi^r$ pour π un polynôme irréductible de $K[X]$. Notons $L = \text{Dec}_K(\pi)$ un corps de décomposition de π sur K de sorte que l'extension L/K est galoisienne. On écrit

$$L[X]/(\pi^r) \simeq L[X]/(X - \lambda_1)^r \times \cdots \times L[X]/(X - \lambda_n)^r,$$

où les λ_i sont les racines de π dans L , i.e. $\pi(X) = \prod_{i=1}^n (X - \lambda_i)$. Sur chaque composante $L[X]/(X - \lambda_i)^r$, on pose $d_i = \lambda_i \text{Id}$ et $n_i = f - \lambda_i$ qui est nilpotent d'ordre r . On note alors, d'après le lemme chinois, P le polynôme unitaire de degré $< nr$ tel que $P(X) \equiv \lambda_i \pmod{(X - \lambda_i)^r}$ pour tout $i = 1, \dots, n$. Soit alors $\sigma \in \text{Gal}(L/K)$ qui permute les λ_i : en particulier on remarque que $\sigma(P)$ vérifie les mêmes congruences de sorte que $\sigma(P) = P$ et donc finalement $P(X) \in K[X]$ et donc la décomposition $f = d + n$ est bien définie sur K où d , et donc aussi n , est un polynôme en f .

Considérons enfin le problème de l'unicité : soit $d' + n' = d + n$ avec $d'n' = n'd'$ et où d, n est défini comme ci-avant, i.e. d et n sont des polynômes en f . Ainsi d' commute avec d et donc d, d' , et donc aussi $d - d'$, sont simultanément diagonalisables dans une extension finie. De même n' commute avec n et donc $n - n'$ est nilpotent et semi-simple, puisque égal à $d' - d$ ce qui impose que $d' - d = 0 = n - n'$, i.e. $d = d'$ et $n = n'$, d'où le résultat.

Contre exemple : considérons $K = \mathbb{F}_p((T))$ et le K -espace vectoriel $E = K[X]/(X^p - T)$ muni de l'endomorphisme f défini par la multiplication par X . Notons que $X^p - T$ étant irréductible E est un corps de sorte que pour tout $Q \in K[X]$, l'endomorphisme $Q(f)$ égal donc à la multiplication par $Q(X)$ est soit nulle soit un isomorphisme. Ainsi si l'énoncé précédent était valable, l'endomorphisme nilpotent $n = Q(f)$ est nécessairement nul et donc f serait semi-simple, i.e. diagonalisable dans $E' := K'[X]/(X^p - T)$ où $K' = \mathbb{F}_p((T^{\frac{1}{p}}))$. Or dans K' on a $X^p - T = (X - T^{\frac{1}{p}})^p$ et donc f vu comme endomorphisme de E' admet une unique valeur propre : s'il était diagonalisable ce serait donc une homothétie, ce qui n'est visiblement pas le cas.

Un défaut de la preuve précédente est qu'elle nécessite la connaissance des racines du polynôme minimal π_f de f et n'est donc pas constructive, i.e. on ne peut pas programmer un ordinateur pour calculer la décomposition de Dunford en se basant sur la preuve précédente. On propose à présent, en adaptant la méthode de Newton classique, une construction algorithmique de d en remarquant qu'il doit annuler le polynôme $P(X)$ défini, en caractéristique nulle, par

$$P(X) := \frac{\pi_f(X)}{\pi_f(X) \wedge \pi'_f(X)}.$$

On introduit ainsi la suite $(f_n)_{n \in \mathbb{N}}$ définie par récurrence

$$f_0 = f, \quad f_{n+1} = f_n - P(f_n)P'(f_n)^{-1}.$$

- Vérifions tout d'abord que cette suite est bien définie, i.e. que $P'(f_n)$ est une matrice inversible. Pour ce faire on raisonne par récurrence et on ajoute à l'hypothèse de récurrence la propriété suivante :

$$P(f_n) = P(f)^{2^n} Q_n(f), \quad Q_n \in K[X],$$

de sorte que $P(f_n)$ est nilpotent puisque par $P(f)$ l'est : en effet pour r plus grand que la plus grande multiplicité d'une racine de π_f , le polynôme P^r est divisible par π_f et donc $P(f)^r$ est nul.

- Au rang $n = 0$, en posant $Q_0 = 1$ on a bien $P(f) = P(f)^{2^0} Q_0(f)$. On écrit ensuite une relation de Bezout entre P et P' qui par hypothèse sont premier entre eux ce qui donne $U(f)P(f) + V(f)P'(f) = \text{Id}$. Comme $P(f)$ est nilpotent, on en déduit que $\text{Id} - U(f)P(f)$ est inversible et donc aussi $P'(f)$.
- Supposons donc le résultat acquis jusqu'au rang n , de sorte que f_{n+1} est bien défini : on notera au passage que puisque $P'(f_n)$ est un polynôme en f_n , on a $P(f_n)P'(f_n)^{-1} = P'(f_n)^{-1}P(f_n)$. On calcule alors $P(f_{n+1})$ en utilisant la formule de Taylor :

$$\begin{aligned} P(f_{n+1}) &= P(f_n) + (f_{n+1} - f_n)P'(f_n) + (f_{n+1} - f_n)^2 Q(f_n) \\ &= (f_{n+1} - f_n)Q(f_n) \\ &= P(f_n)^2 (P'(f_n)^{-2} Q(f_n)) \end{aligned}$$

qui est donc bien de la forme $P(f)^{2^{n+1}} Q_{n+1}(f)$, d'après l'hypothèse de récurrence et en utilisant que f_n est un polynôme en f .

- Ainsi donc pour n assez grand, $P(f_n)$ est nul et la suite f_n devient stationnaire égale à d qui par construction est un polynôme en f et vérifie $P(d) = 0$. Ainsi d est semi-simple. Par ailleurs on a

$$d - f = (f_n - f_{n-1}) + (f_{n-1} - f_{n-2}) + \cdots + (f_1 - f_0)$$

où chacun $f_{i+1} - f_i = -P(f_i)P'(f_i)^{-1}$ est un polynôme en f et nilpotent. Ainsi ces endomorphismes nilpotents commutent entre eux deux à deux et leur somme est donc nilpotente, d'où le résultat.

3.8 Sous-espaces stables

Un sous-espace $W \subset V$ est stable par f si et seulement si W est un sous- $K[X]$ -module de V_f .

Lemme 110. *Soit f un endomorphisme cyclique. Les sous-espaces stables par f sont les $\mathfrak{S}P(f) = \text{Ker } \frac{\chi_f}{P}(f)$ où P décrit les diviseurs de χ_f .*

Remarque: en particulier un endomorphisme cyclique n'admet qu'un nombre fini de sous-espaces stables. On renvoie à l'exercice ?? pour la réciproque.

Preuve : Tout sous-espace stable de $V_f \simeq K[X]/(\chi_f)$ sont ses sous-modules et donc aux idéaux $(P(X))$ pour P un diviseur de χ_f , d'où le résultat.

Définition 111. *Un endomorphisme est dit indécomposable si on ne peut pas décomposer l'espace en une somme directe de deux sous-espaces stables stricts. Il est dit semi-simple si tout sous-espace stable admet un supplémentaire stable.*

Proposition 112. *Un endomorphisme f est indécomposable si et seulement s'il est cyclique et son polynôme caractéristique est la puissance d'un irréductible.*

Preuve : D'après la décomposition (3) en sous-espaces cycliques, l'endomorphisme doit être nécessairement cyclique. En outre d'après le lemme chinois le polynôme caractéristique doit être la puissance d'un irréductible.

Réciproquement supposons que f est cyclique avec χ_f la puissance d'un irréductible. Si on avait $V = V_0 \oplus V_1$ alors pour $W_0 \simeq K[X]/(P_0)$ (resp. $W_1 \simeq K[X]/(P_1)$) un sous-espace cyclique de V_0 , on devrait avoir, d'après le théorème chinois $P_0 \wedge P_1 = 1$ et $P_0 P_1$ qui diviser χ_f ce qui n'est pas.

Proposition 113. *Un endomorphisme f est semi-simple si et seulement s'il est sans facteur carré.*

Preuve : Supposons que f est semi-simple et, par l'absurde que $\pi_f = P^2 Q$. Le sous espace $W = \text{Ker } P(u)$ est stable et admet donc un supplémentaire stable W' . Montrons alors que $PQ(f)$ est nul sur W et W' et est donc divisible par π_f ce qui sera contradictoire. Clairement $QP(f)$ s'annule sur W et pour $w' \in W'$, on a $P(f) \circ (PQ(f))(w') = 0$ et donc $PQ(f)(w') \in W$: mais comme W' est stable par f , on a aussi $PQ(f)(w') \in W'$ et donc comme $W \cap W' = \{0\}$, $PQ(f)(w') = 0$.

Supposons à présent que π_f est irréductible de sorte que V est un E -espace vectoriel où $E = K[X]/(\pi_f)$ est un corps. Un sous-espace stable de V est alors un E -espace vectoriel qui admet donc un E -supplémentaire qui est un K -espace vectoriel stable par f et donc f est bien semi-simple.

Soit enfin f tel que $\pi_f = \prod_i P_i$ est sans facteur carré. La décomposition (3) s'écrit aussi sous la forme $V = \bigoplus_i V_i$ où $V_i = \text{Ker } P_i(f)$ est un E_i -espace vectoriel où $E_i = K[X]/(P_i)$ est une extension finie de K . Soit alors W un sous-espace stable de V et soit $W_i = W \cap V_i = \text{Ker } P_i(f)|_W$ de sorte que d'après le lemme des noyaux

$$W = \bigoplus_i W_i.$$

Comme $f|_{V_i}$ est semi-simple, on note W'_i un supplémentaire stable de W_i dans V_i de sorte que $W' := \bigoplus_i W'_i$ est un supplémentaire stable de W .

Remarque: sur un corps algébriquement clos, semi-simple est équivalent à la notion plus classique d'endomorphisme diagonalisable au sens de la définition suivante.

Définition 114. *Un endomorphisme est dit diagonalisable s'il existe une base de vecteurs propres, i.e. s'il existe une base dans laquelle sa matrice est diagonale.*

Remarque: sur $\mathbb{R}$ un endomorphisme semi-simple est semblable à une matrice diagonale par blocs dont les blocs sont soit de dimension 1 soit de dimension 2 de la forme $\begin{pmatrix} a & -b \\ b & a \end{pmatrix}$.

Théorème 115. *L'endomorphisme f est diagonalisable si et seulement si son polynôme minimal est scindé à racines simples.*

Remarque: par exemple si le polynôme caractéristique est scindé à racines simples alors π_f aussi ; plus généralement si P est un polynôme annulateur de f scindé à racines simples alors μ_f le sera aussi, puis que μ_f divise tout polynôme annulateur.

Nous avons vu que dans le cas d'un endomorphisme cyclique f , le commutant de f est $\{Q(f) : Q \in K[X]\}$. Considérons le cas $V_f \simeq K[X]/(P) \oplus K[X]/(PQ)$ et l'endomorphisme $g = A_1(f_1) \oplus A_2(f_2)$ où f_1 (resp. f_2) est la restriction de f au premier facteur $K[X]/(P)$ (resp. $K[X]/(PQ)$). On choisit alors A_2 tel que $A_2 \not\equiv A_1 \pmod{P}$. S'il existe $B \in K[X]$ tel que $g = B(f)$ alors $B \equiv A_1 \pmod{P}$ et $B \equiv A_2 \pmod{PQ}$ ce qui n'est pas puisque $A_2 \not\equiv A_1 \pmod{P}$. On a ainsi construit un endomorphisme g commutant avec f mais qui n'est pas un polynôme en f . En particulier on en déduit que si f est tel que son commutant est $\{Q(f) : Q \in K[X]\}$ alors f est un endomorphisme cyclique.

Proposition 116. *Avec les notations précédentes, on suppose g commute avec f et que tout sous-espace stable F par f est aussi stable par g . Alors g est de la forme $Q(f)$ pour $Q \in K[X]$.*

Preuve : Considérons la décomposition de

$$V_f = K[X]/(P_1) \oplus K[X]/(P_2) \oplus \cdots \oplus K[X]/(P_r)$$

en sous-espace cyclique. D'après le cas cyclique, comme la restriction g_i de g à chacun des facteurs $V_i := K[X]/(P_i)$ commute à celle f_i de f , on en déduit qu'il existe des polynômes $Q_1, \dots, Q_r$ tels que $g_i = Q_i(f_i)$ et il s'agit de montrer que $Q_r \equiv Q_i \pmod{P_i}$ pour tout $i = 1, \dots, r$. Notons v_i un vecteur engendrant l'espace cyclique V_i et soit $w_i = v_i + \frac{P_r}{P_i}(f)(v_r)$ de sorte que l'espace cyclique engendré par w_i est isomorphe à $K[X]/(P_i)$. Comme

par hypothèse g stabilise cet espace cyclique, sa restriction est de la forme $Q(f)$, i.e.

$$\begin{aligned} g(w_i) &= Q(f)(w_i) = Q(f)(v_i) + Q(f)\left(\frac{P_r}{P_i}(f)(v_r)\right) \\ &= Q_i(f)(v_i) + Q_r(f)\left(\frac{P_r}{P_i}(f)(v_r)\right), \end{aligned}$$

et donc, dans $V_i \oplus V_r$,

$$Q \equiv Q_i \pmod{P_i} \text{ et } Q \equiv Q_r \pmod{P_i},$$

d'où le résultat.

3.9 Exercices

Exercice 1. Montrer que si un ouvert de $\mathbb{M}_n(\mathbb{C})$ contient les matrices diagonales et est stable par similitude, alors il est égal à $\mathbb{M}_n(\mathbb{C})$ tout entier.

Preuve : Soit F le fermé complémentaire ; s'il était non vide il contiendrait une matrice $M = S + N$, sa décomposition de Dunford, et contiendrait aussi sa partie semi-simple S , laquelle est dans l'adhérence de la classe de similitude de M , d'où la contradiction.

Exercice 2. Montrer que sur un corps algébriquement clos, deux matrices A et B sont semblables si et seulement si, pour tout $\lambda \in K$ et pour tout $k \geq 0$, on a $\text{rg}(A - \lambda I)^k = \text{rg}(B - \lambda I)^k$.

Preuve : À partir de la forme de Jordan, on rappelle que pour $k \geq 1$, $d_k(A) := \dim \text{Ker}(A - \lambda I)^k - \dim \text{Ker}(A - \lambda I)^{k-1}$ est égal au nombre de bloc Jordan pour la valeur propre λ qui sont de taille plus grande que k . D'après le théorème du rang, pour tout $k \geq 1$, on a $d_k(A) = d_k(B)$ de sorte que A et B ont les mêmes réduites de Jordan et sont donc semblables.

Exercice 3. Montrer que les valeurs propres de $A = (a_{i,j})_{1 \leq i,j \leq n}$ sont dans la réunion des disques fermés centrés en $a_{i,i}$ et de rayon $\sum_{j \neq i} |a_{i,j}|$ (ce sont les disques de Gershgorin).

Preuve : Le résultat découle directement du lemme d'Hadamard appliqué à $A - \lambda \text{Id}$. Rappelons que ce lemme dit que si pour tout $1 \leq i \leq n$, on a $|a_{i,i}| > \sum_{j \neq i} |a_{i,j}|$ alors A est inversible. En effet soit X de coordonnées $(x_i)_{1 \leq i \leq n}$ dans le noyau de A et soit i_0 tel que $|x_{i_0}|$ soit maximal parmi les $|x_i|$. De l'égalité $a_{i_0,i_0}x_{i_0} = -\sum_{j \neq i_0} a_{i_0,j}x_j$ on en déduit la majoration $|a_{i_0,i_0}x_{i_0}| \leq |x_{i_0}| \sum_{j \neq i_0} |a_{i_0,j}|$ et donc $x_{i_0} = 0$ soit $X = 0$.

Exercice 4. Soit A une matrice vérifiant $(A - I)^2(A - 2I) = 0$. Calculer $\exp(A)$ sous la forme d'un polynôme en A .

Preuve : Le lemme des noyaux permet de décomposer l'espace $E = \text{Ker}(A - I)^2 \oplus \text{Ker}(A - 2I)$. On considère le projecteur q (resp. p) sur $\text{Ker}(A - I)^2$ (resp. $\text{Ker}(A - 2I)$) parallèlement à $\text{Ker}(A - 2I)$ (resp. $\text{Ker}(A - I)^2$). De l'égalité $p + q = \text{Id}$, on obtient $\exp(A) = \exp(A)p + \exp(A)q$. Or $\exp(A)p = e^2 \exp(A - 2I) = e^2 \sum_{n \geq 0} \frac{(A - 2I)^n}{n!} \circ p = e^2 p$ car $(A - 2I) \circ p = 0$. De même on a $\exp(A)q = eAq$. De l'identité de Bezout $1 = (X - 1)^2 - X(X - 2)$, on en déduit que $p = -A(A - 2I)$ et $q = (A - I)^2$ et donc

$$\exp(A) = -e^2 A(A - 2I) + eA(A - I)^2$$

Exercice 5. Montrer que l'ensemble des matrices diagonalisables de $M_n(\mathbb{C})$ est connexe et dense. Quel est son intérieur ? Ce dernier est-il encore connexe ?

Preuve : *Connexité :* on a une application surjective $GL_n(\mathbb{C}) \times (\mathbb{C}^\times)^n$ sur l'ensemble des matrices diagonalisables : on envoie $(P, (a_1, \dots, a_n))$ sur $P \text{diag}(a_1, \dots, a_n) P^{-1}$. L'ensemble $GL_n(\mathbb{C}) \times (\mathbb{C}^\times)^n$ étant connexe, il en est de même de l'ensemble des matrices diagonalisables.

On rappelle que $GL_n(\mathbb{C})$ est connexe : soient P_1, P_2 deux matrices inversibles. On considère le polynôme $\det(P_1 z + (1 - z)P_2)$. Le complémentaire de l'ensemble (fini) des zéros de ce polynôme est connexe ; on considère alors un chemin qui relie 0 à 1 dans ce complémentaire, ce qui fournit un chemin de P_1 à P_2 dans $GL_n(\mathbb{C})$.

Densité : soit A une matrice complexe que l'on trigonalise $PAP^{-1} = T$. Soit alors $\epsilon_1, \dots, \epsilon_n$ petits tels que les $t_{i,i} + \epsilon_i$ sont tous distincts. La matrice $T + \text{diag}(\epsilon_1, \dots, \epsilon_n)$ est alors diagonalisable car elle a n valeurs propres distinctes.

Intérieur : étant donné une matrice A diagonalisable avec une valeur propre multiple ; $P^{-1}AP = \text{diag}(a_1, a_2, \dots, a_n)$ avec $a_1 = a_2$, alors $A + PE_{1,2}P^{-1}$ n'est plus diagonalisable.

Réciproquement si A est diagonalisable à valeurs propres distinctes, vu que le polynôme caractéristique dépend continûment de A , et que les racines d'un polynôme dépendent continûment de ses coefficients, on en déduit que si A' est proche de A , il aura aussi n valeurs propres distinctes et sera donc diagonalisable.

Par ailleurs l'ensemble des matrices à valeurs propres distinctes est encore connexe. En effet c'est le complémentaire des zéros du polynôme en n^2 variable définit comme le discriminant du polynôme caractéristique.

Exercice 6. Caractériser les matrices qui sont des carrés : traiter le cas de $\mathbb{C}$ puis de $\mathbb{R}$.

Preuve : Remarquons déjà que le problème se ramène à traiter le cas nilpotent et le cas inversible. En effet M peut s'écrire sous la forme $P \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix} P^{-1}$ avec A nilpotent et B inversible. Si on a $X^2 = M$ alors M et X commutent

de sorte que $X = P \begin{pmatrix} X_1 & X_3 \\ X_4 & X_2 \end{pmatrix} P^{-1}$ avec $AX_3 = X_3B$ et $AX_4 = X_4B$; on en déduit alors que $\chi_A(A)X_3 = X_3\chi_A(B) = 0$. Or χ_A et χ_B sont premiers entre eux et donc $\chi_A(B)$ est inversible et donc $X_3 = 0$. De la même façon on a $X_4 = 0$.

Le cas nilpotent est traité dans les questions sur la leçon *Endomorphismes nilpotents*. Traitons alors le cas inversible.

Cas complexe : on se ramène comme précédemment au cas $A = \lambda I + N$ avec N nilpotent. On écrit $A = \lambda(I + \frac{N}{\lambda})$ qui a pour racine carrée $\sqrt{\lambda}(I + \frac{N}{\lambda})^{1/2}$ où $(I + \frac{N}{\lambda})^{1/2}$ est défini par la série qui est finie car N est nilpotente.

Cas réel : $A = X^2$ avec A et X réelle. Le cas des valeurs propres strictement positives se traite comme ci-dessus. En ce qui concerne les valeurs propres négatives elles ne peuvent provenir que des valeurs propres imaginaires pures de X ; cette dernière étant réelle les blocs de Jordan associés à $\lambda \in \mathbb{C}$ sont les mêmes que ceux associés à $\bar{\lambda}$. Par ailleurs pour N nilpotente de noyau de dimension 1, on a $(\lambda \text{Id} + N)^2 = \lambda^2 \text{Id} + N'$ avec $N' = 2\lambda N + N^2$ nilpotente de noyau de dimension 1 (écrire N sous la forme de Jordan) de sorte que N' est semblable à N . Ainsi $J_n(\lambda)^2$ est semblable à $J_n(\lambda^2)$ et on remarque donc que les blocs de Jordan de A relativement aux valeurs propres négatives sont, pour chaque dimension, en nombre pair.

Il ne reste alors plus qu'à traiter les matrices réelles A sans valeur propre réelle : A est alors semblable à une somme directe de matrices de la forme $J_n(\lambda) \oplus J_n(\bar{\lambda})$. On écrit alors $J_n(\lambda) = X^2$. La matrice diagonale par blocs $\text{diag}(X, \bar{X})$ est semblable à une matrice réelle : ainsi A est semblable au carré d'une matrice elle-même semblable à une matrice réelle de sorte que A est semblable au carré d'une matrice réelle ; il est alors classique que l'on peut prendre la matrice de passage réelle et donc A est un carré.

Exercice 7. *Montrer que toute matrice est de façon effective semblable à une matrice de Hessenberg, i.e. telle que tous les termes au dessous de sa sous-diagonale sont nuls.*

Preuve : La méthode est celle du pivot de Gauss : on opère, à gauche, sur les lignes sans toucher à la première de façon à obtenir une première colonne dont tous les termes sont nuls sauf éventuellement les deux premiers. On applique alors la même transformation à droite : comme à gauche on n'avait pas modifié la première ligne, à droite on ne modifie pas la première colonne. On raisonne ensuite par récurrence.

Exercice 8. *Montrer que l'application $A \mapsto A \exp(A)$ est surjective sur $\mathbb{C}$.*

Preuve :

- Le plus difficile est le cas $n = 1$: il s'avère qu'en dehors de $z \neq 0$ qui n'a qu'un seul antécédent, tous les autres complexes en ont chacun une infinité par $f(z) = z \exp z$.

- Pour une matrice nilpotente : d'après Jordan, on se ramène à un seul bloc de Jordan plein J_n . On raisonne par analyse et synthèse. On remarque alors que si $J_n = A \exp A$ alors A est nilpotent avec $\dim \text{Ker } A = 1$ ce qui impose que A est semblable à $J_n : A = P J_n P^{-1}$. En synthèse on remarque simplement que, d'après Jordan, $J_n + J_n^2 + \frac{J_n^3}{2} + \dots + \frac{J_n^{n-1}}{(n-2)!}$ et J_n sont semblables.
- Il ne reste plus qu'à traiter le cas inversible où l'on se ramène à $M = \lambda I_n + J_n$ et où on écrit $\lambda = \mu \exp(\mu)$ avec $\mu \neq 0$ et même $\mu \neq -1$. On a alors $f(\mu I_n + N) = f(\mu) I_n + f'(\mu) N + N^2 p(N)$ où $p(N)$ est un polynôme en N . Comme $f'(\mu) \neq 0$, on peut alors procéder comme dans le cas nilpotent.

Exercice 9. Soit V un $\mathbf{C}$ espace vectoriel de dimension finie $n \geq 1$ et $u \in \text{End}_{\mathbf{C}}(V)$. On munit alors V de sa structure de $\mathbf{C}[X]$ -module associée à u .

- On suppose que V ne possède aucun sous-module non trivial. Montrer que $n = 1$. On remplace maintenant $\mathbf{C}$ par $\mathbf{R}$. L'énoncé est-il encore vrai ?
- On note $P_u = P_1.P_2^2 \dots P_l^l$ le polynôme caractéristique de u , où les P_i sont deux à deux premiers entre eux, sans facteurs carrés et unitaires. Vérifier qu'une telle écriture est possible et est unique.
- Avec les notations du b), on suppose de plus que V est somme directe de sous- $\mathbf{C}[X]$ -modules de dimension 1 (en tant que $\mathbf{C}$ -espaces vectoriels). Calculer les invariants de similitude de u .
- Sous l'hypothèse du c), on se donne de plus en élément $v \in \text{End}_{\mathbf{C}}(V)$ tel que $v \circ u = u \circ v$. Montrer qu'il existe une base de V où u et v sont simultanément diagonalisables.

Preuve : On remarque tout d'abord que u possède une unique valeur propre car dans le cas contraire, pour λ_1 et λ_2 des valeurs propres distinctes, $W = \text{Ker}(u - \lambda_1 \text{Id})$ et $W' = \text{Ker}(u - \lambda_2 \text{Id})$ auraient une intersection réduite au vecteur nul. Soit alors λ l'unique valeur propre de u (sur $\mathbf{C}$, un endomorphisme possède toujours au moins une valeur propre). On remarque alors que $\text{Ker}(u - \lambda \text{Id})$ est de dimension 1, car sinon pour x_1 et x_2 des vecteurs propres non colinéaires, $W = \mathbb{C}x_1$ et $W' = \mathbb{C}x_2$ auraient une intersection réduite vecteur nul. On en déduit donc que u admet un unique facteur invariant égal à son polynôme minimal et à son polynôme caractéristique, soit $(X - \lambda)^n$.

Exercice 10. Soit E un espace vectoriel sur $\mathbb{R}$ de dimension n , u un endomorphisme de polynôme minimal P . On suppose que $P = P_1 P_2$, P_1 et P_2 étant des polynômes unitaires non constants premiers entre eux. On note E_u l'espace E muni de la structure de $\mathbb{R}[X]$ -module définie par l'endomorphisme u .

(a) Montrer que pour $i = 1, 2$,

$$E_i = \{x \in E \mid P_i(u)(x) = 0\}$$

sont des sous-modules de E_u .

(b) Montrer que $E_u = E_1 \oplus E_2$.

(c) Montrer que P_1 est le polynôme minimal de $u|_{E_1}$.

Preuve : (a) Sur $\mathbb{C}$ tout endomorphisme possède une valeur propre et donc un vecteur propre v de sorte que $\mathbb{C}v$ est un sous-espace stable non réduit au vecteur nul de sorte que par hypothèse il est égal à l'espace tout entier qui est donc de dimension 1.

Sur $\mathbb{R}$, l'énoncé est faux, il suffit de considérer dans $\mathbb{R}^2$, une matrice de rotation d'angle $0 < \theta < \pi$.

(b) On décompose P_u , qui par convention est unitaire, en produits de facteurs irréductibles $Q_1^{\alpha_1} \cdots Q_r^{\alpha_r}$ et on remarque que P_i se définit comme le produit des Q_j tels que $\alpha_j = i$.

(c) En tant que $\mathbb{C}[X]$ -module, V est de la forme $(\mathbb{C}[X]/(X - \lambda_1))^{\alpha_1} \oplus \cdots \oplus (\mathbb{C}[X]/(X - \lambda_r))^{\alpha_r}$, où les λ_i sont les valeurs propres de u et α_i leur multiplicité dans le polynôme caractéristique. Avec les notations de (b), on a $P_i = \prod_{j/\alpha_j=i} (X - \lambda_j)$. Les facteurs invariants sont de la forme $\mu_1 | \mu_2 | \cdots | \mu_l$ où chacun des μ_j est de la forme $\prod_{i \in I_j} (X - \lambda_i)$ où I_j est un certain sous-ensemble de $\{1, \dots, r\}$ tel que $I_j \subset I_{j+1}$. Ainsi les éléments de I_1 sont répétés l fois, ceux de I_2 le sont $(l-1)$ fois et de manière générale ceux de I_j le sont $(l+1-j)$ fois. On en déduit donc que $l = \max_i \{\alpha_i\}$ puis que I_j est l'ensemble des i tels que $\alpha_i \geq l+1-j$ de sorte que les facteurs invariants sont $P_l, P_l P_{l-1}, P_l P_{l-1} P_{l-2}, \dots, P_l \cdots P_1$.

Exercice 11. Pour $n > 1$, on note $J_n \in \mathbb{M}_n(\mathbb{C})$ la matrice nilpotente dont tous les coefficients sont nuls, sauf ceux de la première sur-diagonale $j_{i,i+1}$ pour $1 \leq i < n$ qui sont égaux à 1. On considère les matrices suivantes, écrites par blocs :

(a) $A_1 = \text{diag}(aI_3, bI_2, cI_1)$;

(b) $A_2 = \text{diag}(I_3, I_2 + J_2, I_2 + J_2, I_3 + J_3, I_3 + J_3, 2I_2, 2I_3 + J_3, 3I_2, 3I_2 + J_2)$;

(c) $A_3 = \begin{pmatrix} 0 & 0 & \cdots & 0 & a_1 \\ 0 & 1 & \ddots & \vdots & a_2 \\ \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & \cdots & 0 & 1 & a_{n-1} \\ 0 & \cdots & 0 & 0 & a_n \end{pmatrix}$

Déterminer dans chaque cas :

(i) les invariants de similitudes ;

(ii) les polynômes minimaux et caractéristiques ;

(iii) la suite des dimensions des noyaux $\text{Ker}(A_i - \alpha)^k$ où α est une valeur propre.

Preuve : On note $V = \mathbb{C}^n$ l'espace vectoriel en question, que l'on munit de la structure de $A = \mathbb{C}[X]$ -module définie par la matrice à étudier ; on notera $a_r(X) | \cdots | a_1(X)$, ses invariants de similitude. Le polynôme minimal est alors $a_1(X)$ et le polynôme caractéristique est le produit des invariants de similitude. Pour toute valeur propre $\alpha \in \mathbb{C}$, on notera $K_\alpha^i = \text{Ker}(u - \alpha \text{Id})^i$ où u est l'endomorphisme associé à la matrice en question dans la base canonique ; on note aussi r_α l'indice i tel que $K_\alpha^{i-1} \neq K_\alpha^i = K_\alpha^j$ pour tout $j \geq i$; $K_\alpha^{r_\alpha}$ est appelé le sous-espace caractéristique associé à α . L'entier r_α est la multiplicité de α dans $a_1(X)$ tandis que sa dimension est la multiplicité de α dans le produit des a_i . On note $\delta_\alpha^i = \dim K_\alpha^i - \dim K_\alpha^{i-1}$; partant de la forme de Jordan il est aisé de voir que δ_α^i est égal au nombre de a_k divisible par $(X - \alpha)^i$. On remarque ainsi que le nombre r d'invariants de similitude est égal au maximum des dimensions des sous-espaces propres.

(a) Le A -module V est clairement isomorphe à $(A/(X-a))^3 \times (A/(X-b))^2 \times A/(X-c)$; on calcule alors les invariants de similitude via le théorème chinois comme dans la feuille précédente ce qui donne : $(X-a)$, $(X-a)(X-b)$ et $(X-a)(X-b)(X-c)$. La matrice étant diagonalisable, les sous-espaces propres sont les sous-espaces caractéristiques, i.e. tous les δ_α^i sont nuls.

(b) De même on a

$$V \simeq (A/(X-1))^3 \times (A/(X-1)^2)^2 \times (A/(X-1)^3)^2 \times (A/(X-2))^2 \times A/(X-2)^3 \times (A/(X-3))^2 \times A/(X-3)^2$$

les invariants de similitude donnés comme d'habitude par application du théorème chinois sont alors

$$(X-1)^3(X-2)^3(X-3)^2, \quad (X-1)^3(X-2)(X-3), \quad (X-1)^2(X-2)(X-3), \\ (X-1)^2, \quad (X-1), \quad (X-1), \quad (X-1).$$

Avec les notations introduites ci-dessus, on a $\delta_1^1 = 7$ (resp. $\delta_2^1 = 3$, resp. $\delta_3^1 = 3$), puis $\delta_1^2 = 4$ (resp. $\delta_2^2 = 1$, resp. $\delta_3^2 = 1$), et $\delta_1^3 = 2$ (resp. $\delta_2^3 = 1$, resp. $\delta_3^3 = 0$) tous les δ_i^k étant nuls pour $k > 3$. On obtient alors $\dim K_1^1 = 7$ (resp. $\dim K_2^1 = 3$, resp. $\dim K_3^1 = 3$), $\dim K_1^2 = 11$ (resp. $\dim K_2^2 = 4$, resp. $\dim K_3^2 = 4$) et $\dim K_1^3 = 13$ (resp. $\dim K_2^3 = 5$, resp. $\dim K_3^3 = 4$) avec $r_1 = 3$ (resp. $r_2 = 3$, resp. $r_3 = 2$).

(c) Le sous-espace propre associé à la valeur propre 0 (resp. 1) est de dimension supérieure ou égale à 1 (resp. $n-2$) ; dans $\mathbb{C}$, la dernière valeur propre est déterminée via la trace de la matrice dont on sait qu'elle est égale à la somme des valeurs propres comptées avec multiplicité (en effet toute matrice complexe est trigonalisable) ; ainsi on a $1 \cdot 0 + (n-2) \cdot 1 + x = n-2 + a_n$ de sorte que la dernière valeur propre est a_n . Si $a_n \neq 0, 1$ alors la somme des dimensions des sous-espaces propres associés aux valeurs propres 0, 1, a_n est n de sorte que la matrice est diagonalisable et donc

$$V \simeq A/(X) \times A/(X-a_n) \times (A/(X-1))^{n-2}$$

et les invariants de similitude sont

$$a_1(X) = (X-1)X(X-a_n), \quad a_2(X) = X-1, \quad \dots \quad a_{n-2}(X) = X-1.$$

Si $a_n = a_1 = 0$, on est dans la même situation, car le noyau de la matrice est alors de dimension 2 car son rang est de manière évidente $n-2$; les invariants de similitude sont alors

$$a_1(X) = X(X-1), \quad a_2(X) = X(X-1), \quad a_3(X) = \dots = a_{n-2}(X) = X-1.$$

Dans le cas où $a_n = 0$ et a_1 non nul, on a alors $r_0 = 2$ avec $\dim K_0^1 = 1$ de sorte que

$$V \simeq A/(X^2) \times (A/(X-1))^{n-2}$$

soit

$$a_1(X) = X^2(X-1), \quad a_2(X) = \dots = a_{n-2}(X) = (X-1).$$

Exercice 12. Soit u un endomorphisme de $\mathbb{C}^n$, dont les valeurs propres sont 0 et 1; on note K_0^i (resp. K_1^i) le noyau de u^i (resp. $(u - \text{Id})^i$) et soit d_0^i (resp. d_1^i) sa dimension. On suppose que la suite (d_0^i) (resp. (d_1^i)) est égale à $(4, 7, 9, 10, 10, \dots)$ (resp. $(3, 4, 5, 5, \dots)$). Déterminer alors les invariants de similitudes de u .

Preuve : D'après les rappels donnés dans l'exercice précédent, le nombre d'invariants de similitude est égal à la dimension maximale des sous-espaces propres soit donc ici 4 invariants de similitude a_1, a_2, a_3, a_4 . Le polynôme minimal s'écrit sous la forme $X^{\alpha_1}(X-1)^{\beta_1}$ avec $\alpha_1 = r_0$ et $\beta_1 = r_1$ où l'on rappelle que r_0 (resp. r_1) est l'indice i tel que $K_0^{i-1} \neq K_0^i = K_0^{i+k}$ (resp. $K_1^{i-1} \neq K_1^i = K_1^{i+k}$) pour tout $k \geq 0$, soit donc ici $a_1(X) = X^4(X-1)^3$. De même on écrit les $a_i(X)$ sous la forme $a_i(X) = X^{\alpha_i}(X-1)^{\beta_i}$ pour $2 \leq i \leq 4$ avec $\alpha_i \geq \alpha_{i+1}$ (resp. $\beta_i \geq \beta_{i+1}$) et $\sum_{i=1}^4 \alpha_i = 10$ (resp. $\sum_{i=1}^4 \beta_i = 5$).

On introduit comme dans l'exercice précédent $\delta_0^i = \dim K_0^i - \dim K_0^{i-1}$ (resp. $\delta_1^i = \dim K_1^i - \dim K_1^{i-1}$); on rappelle que δ_0^i (resp. δ_1^i) est le nombre d'invariants de similitude divisibles par X^i (resp. $(X-1)^i$) (pour le voir il suffit de raisonner sur la forme de Jordan). En ce qui concerne la valeur propre 0 : on a $\delta_0^4 = 1$ de sorte que $\alpha_2 \leq 3$, en outre $\delta_0^3 = 2$ impose $\alpha_2 \geq 3$ soit $\alpha_2 = 3$ et $\alpha_3 \leq 2$. Enfin $\delta_0^2 = 3$ donne $\alpha_3 = 2$ et $\alpha_4 = 1$.

En ce qui concerne la valeur propre 1, on trouve de la même façon, $\beta_2 = \beta_3 = \beta_4 = 1$ ce qui fournit finalement

$$a_1(X) = X^4(X-1)^3, \quad a_2(X) = X^3(X-1), \quad a_3(X) = X^2(X-1), \quad a_4(X) = X(X-1).$$

Exercice 13. Ecrire sous la forme de Jordan et donner la suite des dimensions des noyaux $\text{Ker}(u - \lambda \text{Id})^i$ des endomorphismes u dont les invariants de similitudes sont :

- (a) $P_1(X) = X$;
- (b) $P_1(X) = X(X - 1)$;
- (c) $P_1(X) = X$ et $P_2(X) = X^2$;
- (d) $P_1(X) = X$ et $P_2(X) = X(X - 1)$;
- (e) $P_1(X) = X^2(X - 1)$, $P_2(X) = X^2(X - 1)(X - 2)$, $P_3(X) = X^3(X - 1)^2(X - 2)$ et $P_4(X) = X^4(X - 1)^3(X - 2)^4$;

Preuve : On reprend les notations des exercices précédents. On rappelle que la dimension n de l'espace vectoriel en question est la somme des degrés des invariants de similitude.

- (a) Ici $n = 1$ et l'endomorphisme en question est l'identité.
- (b) On a $n = 2$ et un espace cyclique avec deux valeurs propres distinctes ; u est donc diagonalisable et sa matrice dans une base de diagonalisation est la matrice diagonale $\text{diag}(0, 1)$.
- (c) $n = 3$ et 0 est la seule valeur propre avec $\delta_0^1 = 2$ et $\delta_0^2 = 1$ soit $\dim K_0^1 = 2$ et $\dim K_0^2 = 1$ et la matrice de Jordan associée est $\text{diag}(0, J_2)$.
- (d) $n = 3$ et 0, 1 sont les valeurs propres de u avec $\delta_0^1 = 2$ (resp. $\delta_1^1 = 1$) et $\delta_1^i = \delta_0^i = 0$ pour $i > 1$. On obtient alors $\dim K_0^1 = 2$ et $\dim K_1^1 = 1$, l'endomorphisme est donc diagonalisable.
- (e) $n = 24$, les valeurs propres étant 0, 1, 2 ; la suite δ_0^i (resp. δ_1^i , resp. δ_2^i) est $(4, 4, 2, 1, 0, \dots)$ (resp. $(4, 2, 1, 0, \dots)$, resp. $(3, 1, 1, 1, 0, \dots)$) de sorte que la suite des dimensions des K_0^i (resp. K_1^i , resp. K_2^i) est $(4, 8, 10, 11, \dots)$ (resp. $(4, 6, 7, \dots)$, resp. $(3, 4, 5, 6, \dots)$). La forme de Jordan est la matrice diagonale par blocs

$$\text{diag}(J_2, J_2, J_3, J_4, I_1, I_1, I_2, I_3, 2I_2, 2I_4 + J_4).$$

Exercice 14. On se propose dans ce problème de donner un algorithme pour calculer la décomposition de Dunford sans calculer les valeurs propres (ce qui algorithmiquement ne peut en général se faire que de manière approchée).

Soit $n \geq 1$ et $A \in \mathcal{M}_n(K)$ une matrice carrée à coefficients dans le corps $K \subset \mathbb{C}$. On note χ_A le polynôme caractéristique de A . Dans $\mathbb{C}$, $\chi_A(X)$ se décompose sous la forme $\prod_i (X - \lambda_i)^{n_i}$ avec $\sum_i n_i = n$. On introduit alors le polynôme $P(X) = \prod_i (X - \lambda_i)$.

(a) Montrer que $P(X) = \lambda \frac{\chi_A(X)}{\chi_A(X) \wedge \chi'_A(X)}$, où $\chi_A(X) \wedge \chi'_A(X)$ désigne le pgcd de χ_A avec son polynôme dérivé et $\lambda \in K$. En déduire alors que $P(X)$ est un polynôme à coefficients dans K .

(b) Soient U et N des matrices de $\mathcal{M}_n(K)$ respectivement inversible et nilpotente, qui commutent entre elles. Montrer que $U - N$ est inversible. Montrer alors que $P'(A)$ est une matrice inversible de $\mathcal{M}_n(K)$ dont l'inverse commute avec A .

(c) On considère alors la suite suivante : $A_0 := A$ et $A_{n+1} = A_n - P'(A_n) \cdot (P'(A_n))^{-1}$. On veut montrer par récurrence sur n que la suite est bien définie, i.e. que $P'(A_n)$ est une matrice inversible.

- (i) Montrer que pour tout polynôme $Q \in K[X]$, il existe $\tilde{Q} \in K[X, Y]$ tel que $Q(X + Y) = Q(X) + YQ'(X) + Y^2\tilde{Q}(X, Y)$.
- (ii) En supposant la suite A_n définie jusqu'au rang n , montrer que $P(A_n)$ s'écrit sous la forme $P(A)^{2^n} \cdot B_n$ où B_n est une matrice de $\mathcal{M}_n(K)$ qui est un polynôme en A .
- (iii) En utilisant une formule de Taylor pour le polynôme P' , écrire $P'(A_{n+1})$ comme la somme d'une matrice inversible $P'(A_n)$ et d'une matrice nilpotente qui commutent entre elles.
- (d) Montrer que pour tout polynôme $Q \in K[X]$, il existe $\tilde{Q} \in K[X, Y]$ tel que $Q(X + Y) = Q(X) + YQ'(X) + Y^2\tilde{Q}(X, Y)$. Montrer alors par récurrence sur n que $P(A_n)$ s'écrit sous la forme $P(A)^{2^n} \cdot B_n$ où B_n est une matrice de $\mathcal{M}_n(K)$.
- (e) En déduire que la suite A_n est stationnaire de limite D avec D diagonalisable sur $\mathbb{C}$ et $N := A - D$ nilpotente vérifiant $DN = ND$.

Preuve : (a) On remarque que la multiplicité de λ_i dans P' est égale à $n_i - 1$ de sorte que λ_i est une racine à l'ordre 1 de $\frac{\mu_A(X)}{\mu_A(X) \wedge \mu_{A'}(X)}$ et qu'en outre ce sont ces seules racines d'où le résultat. On notera en particulier que la connaissance des λ_i n'est pas nécessaire pour calculer P qui peut se calculer via l'algorithme d'Euclide.

(b) - L'idée est d'utiliser la relation formelle $(1-x)(1+x+x^2+\dots+x^k) = 1 - x^{k+1}$ avec $x = U^{-1}N$ et k tel que $N^{k+1} = 0$ soit $(1 - U^{-1}N)(1 + U^{-1}N + \dots + (U^{-1}N)^k) = I_n$ car $(U^{-1}N)^{k+1} = U^{-k-1}N^{k+1}$ car U et N commutent entre eux; soit en multipliant à gauche par U et à droite par U^{-1} , $(U - N)(U^{-1} + U^{-2}N + \dots + U^{-k-1}N^k) = I_n$.

- Les valeurs propres de A ne sont pas des racines de P' et $P \wedge P' = 1$. On considère alors une relation de Bezout $RP' + SP = 1$ pour P et P' qui en l'appliquant à A , donne $R(A)P'(A) = 1 - N$ avec $N = S(A)P(A)$. Or d'après le théorème de Cayley-Hamilton, on a $P^r(A) = 0$ pour $r \geq \max_i(n_i)$ de sorte que N est nilpotent et donc par application de ce qui précède $P'(A)$ est inversible dont l'inverse commute avec A en tant que polynôme en A .

(c) Il s'agit de la méthode de Newton appliqué aux matrices, le but étant de construire une racine de P , i.e. de trouver la partie diagonalisable de A dans sa décomposition de Dunford. Remarquons que pour $n = 0$, $P'(A_0)$ est inversible d'après la question précédente.

- (i) Il suffit par exemple de le vérifier sur les monômes X^m , soit

$$(X + Y)^m = X^m + mYX^{m-1} + Y^2 \sum_{k=2}^m \binom{k}{m} Y^{k-2} X^{m-k}.$$

(ii) Il est clair d'après (a) que pour tout $0 \leq k \leq n$, A_k est un polynôme en A . On raisonne par récurrence sur n . Pour $n = 0$, on a $P(A_0) = P(A)$. Supposons donc le résultat acquis jusqu'au rang k . D'après (i), on écrit $P(A_{k+1}) = P(A_k + Y) = P(A_k) + YP'(A_k) + Y^2\tilde{Q}(A_k, Y)$ avec Y tel que

$P(A_k) + YP'(A_k) = 0$. D'après (a) $Y = P(A_k)Q(A_k)$ et donc $P(A_{k+1})$ est de la forme $P(A)^{2^{k+1}}B_{k+1}$ pour une matrice B_{k+1} qui en tant que polynôme en A_k commute avec A .

(iii) La formule de Taylor donne $P'(A_{n+1}) - P'(A_n) = (A_{n+1} - A_n)Q(A_n)$ où $Q \in K[X]$. Or $A_{n+1} - A_n$ est de la forme $P(A_n)\tilde{Q}(A_n)$ et est donc nilpotent et commute avec A_n qui est un polynôme en A . On en déduit alors que $P'(A_{n+1})$ est inversible d'après (a).

(e) On rappelle que $P^r(A) = 0$ pour $r = \max_i \{n_i\}$ de sorte que la sous-suite $(A_k)_{k \geq n}$ est constante dès que $2^n \geq r$. La limite D est un polynôme en A tel que $P(D) = 0$ de sorte que D est diagonalisable car elle possède un polynôme annulateur scindé à racines simples (dans $\mathbb{C}$). Par ailleurs, pour n tel que $2^n \geq r$, on a $A - D = A_0 - A_n = \sum_{i=0}^{n-1} (A_i - A_{i+1})$ avec $A_i - A_{i+1}$ nilpotent et qui est un polynôme en A . Ainsi les $A_i - A_{i+1}$ commutent en eux de sorte que leur somme est nilpotente d'où le résultat.

Exercice 15. Soit u un endomorphisme de $V \simeq K^n$ dont on note χ_u et π_u respectivement les polynômes caractéristique et minimal :

- (1) χ_u est irréductible si et seulement si V n'a pas de sous-espace stable par u ;
- (2) u est cyclique avec π_u une puissance d'un polynôme irréductible si et seulement si V est indécomposable sous u ;
- (3) proposer un algorithme pour tester si u est semi-simple.

Preuve : (1) Si V a un sous-espace stable W par u , en complétant une base de W en une base de V , la matrice de u y est diagonale par bloc et son polynôme caractéristique est divisible par celui de $u|_W$. Réciproquement si χ_u est de la forme PQ avec P et Q premier entre eux le lemme des noyaux décompose l'espace en une somme directe de $\text{Ker } P(u)$ et de $\text{Ker } Q(u)$. Si $\chi_u = P^r$ avec P irréductible, on a alors $E = \text{Ker } P$ i.e. $\pi_u = P$. Si on prend x quelconque non nul, l'espace vectoriel engendré par $x, u(x), u^2(x), \dots$ est donc au plus de dimension $\deg P$ (en fait on a égalité), et par hypothèse est donc égal à l'espace tout entier, i.e. $r = 1$.

(2) Dans le sens direct, en utilisant la structure de $K[X]$ -module sur V induite par u , on a $V \simeq K[X]/(\pi_u)$. Si V était décomposable il serait en tant que $K[X]$ -module isomorphe à un produit direct $K[X]/P_1 \times K[X]/P_2$, ce qui impose $P_1 = P^r$ et $P_2 = P^s$ avec P irréductible et $\pi_u = P^{r+s}$. Or le polynôme minimal de ce produit direct est visiblement $P^{\max(r,s)}$ soit donc $\min(r, s) = 0$.

Réciproquement si V est indécomposable alors u est cyclique. Par ailleurs si son polynôme minimal n'était pas une puissance d'un polynôme irréductible, alors le lemme des noyaux contredirait l'indécomposabilité de V .

(3) u est semi-simple si et seulement si π_u est sans multiplicité, i.e. est premier avec π'_u . On peut tester si u est semi-simple de manière algorithmique : on calcule le polynôme caractéristique χ_u et on teste si $\frac{\chi_u}{\chi_u \wedge \chi'_u}$ annule u .

Exercice 16. *Quels sont les endomorphismes complexes u qui ne possèdent qu'un nombre fini de sous-espaces stables ?*

Preuve : On se place évidemment sur un corps infini. Il faut déjà que les sous-espaces propres soient de dimension 1 sinon, il y aurait une infinité de droite dans un de ces sous-espaces propres, qui sont bien évidemment stables.

Les sous-espaces propres étant de dimension 1, on en déduit que l'endomorphisme est cyclique (c'est vrai sur chaque sous-espace caractéristique, on applique ensuite le théorème chinois). L'espace muni de la structure de $K[X]$ -module induite par u , est alors isomorphe à $K[X]/(P(X))$ et les sous-espaces stables sont en bijection avec les diviseurs de P .

Exercice 17. *Quels sont les endomorphismes u tels que tout sous-espace stable est de la forme $\text{Ker } P(u)$ ou $\text{Im } P(u)$ pour P un polynôme.*

Preuve : On peut déjà remarquer que les $\text{Ker } P(u)$ et $\text{Im } P(u)$ décrivent un ensemble fini de sous-espaces : en effet si P est premier avec le polynôme minimal de u alors $\text{Ker } P(u) = (0)$ et $\text{Im } P(u) = E$. Comme dans la question précédente, il faut que les sous-espaces propres soient de dimension 1, et alors l'espace étant cyclique, les sous-espaces stables seront les $\text{Ker } Q(X) = \text{Im } \frac{P(X)}{Q(X)}$.

Exercice 18. *Étant donné un drapeau et le parabolique associé montrer que les sous-espaces stables par tous les éléments de sous-groupe parabolique sont ceux du drapeau.*

Preuve : Soit E un sous-espace stable et soit i tel que $V_i \subset E \not\subset V_{i+1}$. Supposons que $E \neq V_i$ et soit $x \in (V_{i+1} \setminus V_i) \cap E$ et soit $y \in V_{i+1} \setminus (V_i \cup E)$. Il existe alors $g \in P_W$ tel que $g|_{V_i} = \text{Id}$ et $g(x) = y$. Or comme E est stable, on a $y \in E$ d'où la contradiction.

Exercice 19. *Donner la "liste" de tous les sous-espaces stables d'une isométrie vectorielle de $\mathbb{R}^3$.*

Preuve : L'identité, les retournements et les réflexions sont diagonalisables ; on connaît alors la "liste" de leurs sous-espaces stables. En particulier on ne peut pas distinguer les retournements des réflexions. En ce qui concerne une rotation générique, elle est semi-simple et ne possède qu'une seule droite propre et donc aussi un seul plan propre (comme elle est semi-simple un plan propre possède un supplémentaire stable qui est donc l'unique droite propre et qui est en plus orthogonal au plan stable).

Exercice 20. (1) *Soit E un K -espace vectoriel et F un sous-espace. Montrer que l'ensemble des supplémentaires de F dans E est un espace affine de direction $\text{Hom}_K(E/F, F)$.*

- (2) Soit u un endomorphisme de E et soit F un sous-espace stable par u . Montrer que l'ensemble des supplémentaires de F stables par u est, quand il est non vide, un espace affine de direction le sous-espace vectoriel de $\text{Hom}_K(E/F, F)$ des s tels que $s \circ \bar{u} - u|_F \circ s = 0$.
- (3) Soit F un sous-espace stable sous M ; on suppose qu'il existe un supplémentaire G stable $E = F \oplus G$. Donner une CNS pour que G soit l'unique supplémentaire stable de F .

Preuve : (1) On identifie les supplémentaires de F dans E avec les sections s de la surjection canonique $\pi : E \rightarrow E/F$. L'espace affine des supplémentaires de F dans E est alors l'ensemble des solutions de l'équation linéaire avec second membre $\pi \circ s = \text{Id}_{E/F}$. La direction de cet espace affine est donc l'ensemble des s tel que $\pi \circ s = 0$ soit donc l'ensemble des $s' : E/F \rightarrow F$.

(2) Un tel supplémentaire sera alors stable si et seulement si $s \circ \bar{u} - u \circ s = 0$. La direction est alors l'intersection des deux espaces vectoriels : $\pi \circ s = 0$ et $s \circ \bar{u} - u \circ s = 0$, soit le sous-espace de $\text{Hom}_K(E/F, F)$ des s tels que $s \circ \bar{u} - u|_F \circ s = 0$.

(3) D'après la question précédente, la direction de l'espace affine des supplémentaires stables de F est celle de l'espace vectoriel des matrices rectangulaires $X \in \mathbb{M}_{p, n-p}(K)$ telles que $AX - XB = 0$ avec $p = \dim F$. On va montrer qu'une condition nécessaire et suffisante est que les polynômes caractéristiques de A et B sont premiers entre eux.

Soit μ_A et μ_B les polynômes minimaux respectifs de A et B et soit Q un facteur irréductible de leur pgcd. D'après la théorie de la réduction, on peut décomposer F (resp. G) en sous-espace stable par A (resp. B) sous la forme $F_A \oplus F'_A$ (resp. $G_B \oplus G'_B$), tels que F_A (resp. G_B) est cycliques relativement à A (resp. B) de polynôme caractéristique Q . Soit alors $X : G \rightarrow F$ défini comme suit : il induit un isomorphisme de G_B sur F_A et est nul sur G'_B . On a alors $AX = XB$.

Supposons que μ_A et μ_B sont premiers entre eux ce qui est équivalent au fait que leur polynôme caractéristique le sont. Par linéarité on a $\mu_A(A)X = X\mu_A(B) = 0$ d'après Cayley-Hamilton. Or comme μ_A est premier avec μ_B , une relation de Bézout $P_A\mu_A + P_B\mu_B = 1$ donne que $\mu_A(B)$ est inversible d'inverse $P_A(B)$ de sorte que X est nulle.

Exercice 21. Donner les sous-groupes commutatifs d'exposant r de $GL_n(\mathbb{C})$. En déduire que $GL_n(\mathbb{C}) \simeq GL_m(\mathbb{C})$ si et seulement si $n = m$.

Preuve : Soit G un sous-groupe commutatif de $GL_n(\mathbb{C})$ d'exposant r . Pour tout $g \in G$, on a donc $g^r = \text{Id}$ i.e. $X^r - 1$ est un polynôme annulateur scindé à racines simples de g de sorte que g est diagonalisable dans une base $(e_1, \dots, e_n)$. Pour tout $g' \in G$, g et g' commutent et sont diagonalisables, on en déduit donc qu'ils sont simultanément diagonalisables. Ainsi la matrice de tout $g \in G$ dans la base $(e_1, \dots, e_n)$, est diagonale de la forme

$\text{diag}(\xi_1, \dots, \xi_n)$ où ξ_i est une racine r -ième de l'unité. On en déduit donc que $G \simeq (\mathbb{Z}/r\mathbb{Z})^n$.

Si les deux groupes $GL_n(\mathbb{C})$ et $GL_m(\mathbb{C})$ sont isomorphes, leurs sous-groupes commutatifs d'exposant r se correspondent soit $(\mathbb{Z}/r\mathbb{Z})^n \simeq (\mathbb{Z}/r\mathbb{Z})^m$ et donc $n = m$ par cardinalité.

Exercice 22. *Montrer que A de $GL_n(\mathbb{C})$ est diagonalisable si et seulement si il existe k tel que A^k l'est.*

Preuve : Le sens direct est évident. Dans l'autre sens, on raisonne dans chacun des espaces caractéristiques de sorte que l'on se ramène à une unique valeur propre λ . On écrit A sous la forme $\lambda(I_n + N)$ avec N nilpotent. On a alors $(I_n + N)^k = I_n + kN + \dots$. Or on remarque que $kN + \dots$ est nilpotent et semblable à N (utiliser Jordan), d'où le résultat.

Exercice 23. *Proposer un test effectif pour savoir si une matrice complexe est diagonalisable à valeurs propres distinctes. Traiter ensuite le cas des matrices réelles.*

Preuve : Une matrice de $M_n(K)$ est diagonalisable à valeurs propres distinctes si et seulement si son polynôme minimal est de degré n et est scindé à racines simples. Sur $\mathbb{C}$, il suffit alors de tester si le polynôme caractéristique χ est à racines simples. Pour cela il suffit de vérifier qu'un pgcd de χ et χ' est égal à 1.

Sur $\mathbb{R}$, il faut en plus tester si χ est scindé. Pour cela on dispose des suites de Sturm qui nous donnent le nombre de racines réelles de χ .

Exercice 24. *Soient A et B deux matrices simultanément diagonalisables. Existe-t-il un polynôme P tel que $B = P(A)$? Montrer qu'il existe une matrice C ainsi que des polynômes P_A, P_B tels que $A = P_A(C)$ et $B = P_B(C)$.*

Preuve : Si on prend $A = I_n$ alors $P(A) = P(1)I_n$ de sorte que si B est diagonalisable sans être scalaire, il ne peut exister un tel P .

Le problème dans la question précédente venait des racines multiples. Soit P la matrice de passage de la base canonique à la base de diagonalisation commune de A et B : $PAP^{-1} = \text{diag}(a_1, \dots, a_n)$ et $PBP^{-1} = \text{diag}(b_1, \dots, b_n)$. Prenons alors C tel que $PCP^{-1} = \text{diag}(c_1, \dots, c_n)$ avec les c_i distincts deux à deux, par exemple $c_i = i$. On choisit alors P_A (resp. P_B) tel que pour tout i , $P_A(c_i) = a_i$ (resp. $P_B(c_i) = b_i$) : c'est possible en utilisant par exemple les polynômes d'interpolation de Lagrange. On a donc bien $A = P_A(C)$ et $B = P_B(C)$.

Exercice 25. *Soit a un endomorphisme non diagonalisable, trouver les polynômes P tel que $P(a)$ soit diagonalisable (traiter le cas de $\mathbb{C}$ puis de $\mathbb{R}$).*

Preuve : Sur $\mathbb{C}$: on note $\mu_a(X) = \prod_{\lambda} (X - \lambda)^{r_{\lambda}}$ le polynôme minimal de a . Pour toute valeur propre λ , on se place sur le sous-espace caractéristique associé $E_{(\lambda)}$. Si $P(a)$ est diagonalisable alors sa restriction à $E_{(\lambda)}$ est $P(\lambda)\text{Id}$. Or on a $P(a) - P(\lambda)\text{Id} = (a - \lambda\text{Id})Q_{\lambda}(a)$; pour que ce dernier soit nul il faut que $Q_{\lambda}(a)(E_{(\lambda)}) \subset E_{\lambda}$ où E_{λ} est le sous-espace propre. Ainsi il faut que $(X - \lambda)^{r_{\lambda}-1}$ divise $Q_{\lambda}(X)$ ce qui est équivalent, pour $r_{\lambda} > 1$ à

$$P'(\lambda) = P''(\lambda) = \dots = P^{(r_{\lambda}-1)}(\lambda) = 0$$

Réciproquement si cette dernière condition est vérifiée pour toute valeur propre λ alors $P(a)$ est diagonalisable.

Sur $\mathbb{R}$: on se place dans $\mathbb{C}$ de sorte qu'il faut que la condition précédente soit vérifiée pour toute valeur propre λ . En outre il faut s'assurer que les valeurs propres $P(\lambda)$ sont réelles. Réciproquement si ces deux conditions sont vérifiées alors la matrice de a dans la base canonique est semblable sur $\mathbb{C}$ à une matrice diagonale réelle. Il est bien connu alors qu'elle y est semblable sur $\mathbb{R}$.

Exercice 26. Donner les invariants de similitude d'un endomorphisme diagonalisable.

Preuve : On écrit le polynôme caractéristique de a , $\chi(X) = \prod_{\lambda} (X - \lambda)^{r_{\lambda}}$. Celui-ci étant supposé diagonalisable, son polynôme minimal est $\mu(X) = \prod_{\lambda} (X - \lambda)$. On rappelle que les invariants de similitude sont des polynômes

$$\mu_1 | \mu_2 | \dots | \mu_r$$

avec $\mu = \mu_r$ et $\chi = \prod_i \mu_i$. On en déduit ainsi que tous les invariants de similitudes sont sans multiplicité, que $r = \max_{\lambda} (r_{\lambda})$ et que

$$\mu_i(X) = \prod_{\lambda / r_{\lambda} \geq r-i+1} (X - \lambda)$$

Exercice 27. La matrice $A = \begin{pmatrix} 1 & 1 & \dots & 1 & 1 \\ 1 & 0 & \dots & 0 & 1 \\ \vdots & \vdots & & \vdots & \vdots \\ 1 & 0 & \dots & 0 & 1 \\ 1 & 1 & \dots & 1 & 1 \end{pmatrix}$ est-elle diagonalisable ? Donner ses valeurs propres.

Preuve : La matrice A est symétrique réelle donc diagonalisable. Par ailleurs son rang est clairement égal à 2 de sorte que 0 est valeur propre à l'ordre $n - 2$; reste alors à trouver deux autres valeurs propres λ_1 et λ_2 . La trace nous donne $\lambda_1 + \lambda_2 = 2$ tandis que la trace de A^2 donne $\lambda_1^2 + \lambda_2^2 = 2n + n - 2$ ce qui donne λ_1 et λ_2 .

Exercice 28. Les matrices suivantes sont-elles des carrés dans $\mathbb{M}_2(\mathbb{R})$:

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

Preuve : Soit $A = \begin{pmatrix} -1 & 0 \\ 0 & -4 \end{pmatrix}$. Supposons qu'il existe B réelle telle que $A = B^2$. Le polynôme $(X^2 + 1)(X^2 + 4)$ est alors un polynôme annulateur de B de sorte que son polynôme minimal, qui est de degré 1 ou 2, doit être $X^2 + 1$ ou $X^2 + 4$ (qui sont irréductibles sur $\mathbb{R}$). Mais alors on aurait $A + \text{Id} = 0$ ou $A + 4\text{Id} = 0$ ce qui n'est pas.

Soit $A = \begin{pmatrix} -1 & 1 \\ 0 & -1 \end{pmatrix}$. Supposons qu'il existe B tel que $A = B^2$. Le polynôme $(X^2 + 1)^2$ est alors un polynôme annulateur de B de sorte que son polynôme minimal de B est $X^2 + 1$ et donc $A + \text{Id} = 0$ ce qui n'est pas.

Exercice 29. Montrer que si un ouvert de $\mathbb{M}_n(\mathbb{C})$ contient les matrices diagonales et est stable par similitude, alors il est égal à $\mathbb{M}_n(\mathbb{C})$ tout entier.

Preuve : Soit F le fermé complémentaire ; s'il était non vide il contiendrait une matrice $M = S + N$, sa décomposition de Dunford, et contiendrait aussi sa partie semi-simple S , laquelle est dans l'adhérence de la classe de similitude de M , d'où la contradiction.

Exercice 30. (1) Sur $\mathbb{R}$ ou $\mathbb{C}$, montrer que

- (i) l'intérieur de l'ensemble $\mathcal{D}$ des matrices diagonalisables est l'ensemble $\mathcal{D}^1$ des matrices avec n valeurs propres distinctes ;
 - (ii) l'adhérence de $\mathcal{D}$ est l'ensemble $\mathcal{T}$ des matrices trigonalisables.
- (2) Sur $\mathbb{C}$, montrer la connexité de $\mathcal{D}$ et $\mathcal{D}^1$.

Preuve : (1) (i) Soit A une matrice possédant n valeurs propres distinctes de sorte que son polynôme caractéristique est scindé à racines simples. D'après la continuité du polynôme caractéristique et de celle des racines d'un polynôme, on en déduit que pour tout A' proche de A , le polynôme caractéristique de A' possède, sur $\mathbb{C}$, n racines distinctes. Si A et A' sont réelles, alors leurs racines le sont aussi : pour A c'est vrai par hypothèse, pour A' , ses racines sont complexes conjuguées et proches de celles de χ_A , on conclut en remarquant que ces dernières sont simples.

Ainsi $\mathcal{D}^1$ est ouvert, en outre toute matrice diagonalisable est limite de matrices de $\mathcal{D}^1$: en effet pour tout $a_1, \dots, a_n$ et pour tout $\epsilon > 0$, il existe $\epsilon_1, \dots, \epsilon_n$ tels que $0 < \epsilon_i < \epsilon$ et les $a_i + \epsilon_i$ sont distincts deux à deux. Ainsi pour $A = P \text{diag}(a_1, \dots, a_n) P^{-1}$, la boule de centre A et de rayon ϵ contient $A' = P \text{diag}(a_1 + \epsilon_1, \dots, a_n + \epsilon_n) \in \mathcal{D}^1$.

On en déduit alors que l'adhérence de $\mathcal{D}^1$ contient $\mathcal{D}$ et donc que $\mathcal{D}$ est l'intérieur de $\mathcal{D}$.

(ii) Soit $(A_k)_{k \in \mathbb{N}}$ une suite d'éléments de $\mathcal{D}$ qui converge vers A . D'après la continuité du polynôme caractéristique, χ_k converge vers χ . Les χ_k étant

scindé, on en déduit comme ci-dessus que χ est scindé : ses racines complexes sont limites des racines de χ_k , si celles-ci sont toutes réelles, leurs limites aussi. On rappelle alors que A est trigonalisable.

Réciproquement si A est trigonalisable $PAP^{-1} = T$, soit alors $\epsilon_1, \dots, \epsilon_n$ petits tels que les $t_{i,i} + \epsilon_i$ sont tous distincts. La matrice $T + \text{diag}(\epsilon_1, \dots, \epsilon_n)$ est alors diagonalisable car elle a n valeurs propres distinctes. Ainsi A est dans l'adhérence de $\mathcal{D}$.

(2) Pour $\mathcal{D}$: on a une application surjective $GL_n(\mathbb{C}) \times (\mathbb{C}^\times)^n$ sur l'ensemble des matrices diagonalisables : on envoie $(P, (a_1, \dots, a_n))$ sur $P \text{diag}(a_1, \dots, a_n) P^{-1}$. L'ensemble $GL_n(\mathbb{C}) \times (\mathbb{C}^\times)^n$ étant connexe, il en est de même de l'ensemble des matrices diagonalisables.

On rappelle que $GL_n(\mathbb{C})$ est connexe : soient P_1, P_2 deux matrices inversibles. On considère le polynôme $\det(P_1 z + (1 - z)P_2)$. Le complémentaire de l'ensemble (fini) des zéros de ce polynôme est connexe ; on considère alors un chemin qui relie 0 à 1 dans ce complémentaire, ce qui fournit un chemin de P_1 à P_2 dans $GL_n(\mathbb{C})$.

Pour $\mathcal{D}^1$: on remarque que $\mathcal{D}^1$ est le complémentaire des zéros du polynôme en n^2 variable défini comme le discriminant du polynôme caractéristique.

Exercice 31. Décrire l'adhérence de l'orbite d'un bloc de Jordan

$$J_n = \begin{pmatrix} 0 & 1 & 0 & 0 \\ \vdots & \ddots & \ddots & 0 \\ \vdots & & 0 & 1 \\ 0 & \dots & \dots & 0 \end{pmatrix}$$

Preuve : On va montrer que cette adhérence est l'ensemble des matrices nilpotentes. Rappelons que d'après la décomposition de Jordan, tout matrice nilpotente est semblable à une matrice diagonale par blocs, avec sur la diagonale des blocs de Jordan de taille distinctes $J(n_1, \dots, n_r) := \text{diag}(J_{n_1}, \dots, J_{n_r})$ avec $\sum_i n_i = n$. On remarque alors que J_n est semblable à

$$\begin{pmatrix} J_{n_1} & \epsilon E_{n_1, n_2} & 0 & \dots & 0 \\ 0 & J_{n_2} & \epsilon E_{n_2, n_3} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & 0 & J_{n_{r-1}} & \epsilon E_{n_{r-1}, n_r} \\ 0 & \dots & \dots & 0 & J_{n_r} \end{pmatrix}$$

où $E_{i,j}$ est la matrice de taille $i \times j$ dont tous les coefficients sont nuls sauf celui du coin en bas à gauche qui vaut 1, et où $\epsilon > 0$. En faisant tendre ϵ vers zéro on en déduit que $J(n_1, \dots, n_r)$ est dans l'adhérence de l'orbite de J_n .

Exercice 32. (a) Montrer que si $\text{Ker } A^2 = \text{Ker } A$, alors il existe un pseudo-inverse X , i.e. tel que $AX = XA$, $AXA = A$ et $XAX = X$.

(b) A quelle condition sur les invariants de similitude de A a-t-on $\dim \text{Ker } A^2 = 2 \dim \text{Ker } A$?

Preuve : (a) On décompose l'espace en sous-espace caractéristique pour A . Sur les espaces caractéristiques associés aux valeurs propres non nulles, on pose $X = A^{-1}$. Sur l'espace caractéristique associé à la valeur propre 0, l'hypothèse implique que A y est nulle, on prend donc X quelconque.

(b) Cela correspond à dire que les deux première colonnes du tableau de Young associé à A sont de même longueur ce qui est équivalent à demander qu'il n'y ait aucun bloc de Jordan de taille 1.

Exercice 33. On considère la suite des dimensions des noyaux emboîtés. Décrire l'ensemble des suites obtenues.

Preuve : Notons pour $k \geq 0$, $a_k := \dim \text{Ker } a^k$. Il s'agit d'une suite croissante majorée par la dimension de l'espace n . On a $a_0 = 0$ et si $a_1 = 0$ alors pour tout k , $a_k = 0$. Plus généralement soit r le premier indice tel que $a_r = a_{r+1}$. Soit alors $x \in \text{Ker } a^{r+2}$ de sorte que $a(x) \in \text{Ker } a^{r+1} = \text{Ker } a^r$ et donc $a^{r+1}(x) = 0$ soit $x \in \text{Ker } a^{r+1}$ et donc $a_{r+1} = a_r$ et par récurrence $a_r = a_{r+k}$ pour tout $k \geq 0$.

On introduit la suite $d_k := a_k - a_{k-1}$ pour tout $k \geq 1$. Remarquons que cette suite est décroissante : en effet a induit un endomorphisme injectif de $\text{Ker } a^k / \text{Ker } a^{k-1}$ dans $\text{Ker } a^{k-1} / \text{Ker } a^{k-2}$.

Réciproquement soit $(a_k)_{k \geq 0}$ une suite croissante majorée par n telle que la suite des différences d_k est décroissante. On considère alors la matrice nilpotente A sous forme de Jordan dont le nombre de blocs de Jordan de taille r est égal à $d_r - d_{r+1}$. On vérifie alors aisément que $a_k = \dim \text{Ker } A^k$.

Par ailleurs, une manière graphique de représenter les classes de similitude de matrices nilpotentes, est d'introduire le diagramme de Young dont les colonnes sont les d_i pour $i = 1, \dots, r$. Les blocs de Jordan se lisent alors sur les lignes.

Exercice 34. Montrer qu'un sous- $\mathbb{R}$ -espace vectoriel du cône nilpotent est de dimension inférieur à $\frac{n(n-1)}{2}$ et que ce maximum est atteint.

Preuve : On considère la forme quadratique q définie sur l'espace des matrices qui à X associe $\text{tr} X^2$. De manière évidente le cône isotrope est constitué de vecteurs isotropes de sorte que l'espace vectoriel en question sera totalement isotrope.

Par ailleurs, si $X \neq 0$ est symétrique (resp. antisymétrique) alors $q(X) > 0$ (resp. $q(X) < 0$) de sorte que la signature de q est $(\frac{n(n+1)}{2}, \frac{n(n-1)}{2})$. Ainsi un sous-espace totalement isotrope est de dimension inférieure ou égale à $\frac{n(n-1)}{2}$.

L'égalité est clairement atteinte pour les matrices strictement triangulaires supérieures.

Exercice 35. Soit u un endomorphisme nilpotent ; pour tout $i \geq 0$, on note K_0^i le noyau de u^i soit d_0^i sa dimension. On suppose que la suite (d_0^i) est égale à $(4, 7, 9, 10, 10, \dots)$. Déterminer alors les invariants de similitudes de u .

Preuve : On note $V = K^n$ l'espace vectoriel en question, que l'on munit de la structure de $A = K[X]$ -module définie par la matrice à étudier ; on notera $a_r(X) | \dots | a_1(X)$, ses invariants de similitude. Le polynôme minimal est alors $a_1(X)$ et le polynôme caractéristique est le produit des invariants de similitude. On note r l'indice i tel que $K_0^{i-1} \neq K_0^i = K_0^j$ pour tout $j \geq i$. L'entier r est la multiplicité de 0 dans $a_1(X)$ tandis que sa dimension est la multiplicité de 0 dans le produit des a_i . On note $\delta_0^i = \dim K_0^i - \dim K_0^{i-1}$; partant de la forme de Jordan il est aisé de voir que δ_0^i est égal au nombre de a_k divisible par X^i . On remarque ainsi que le nombre r d'invariants de similitude est égal au maximum des dimensions des sous-espaces propres.

Ainsi le nombre d'invariants de similitude est égal à la dimension du noyau soit donc 4 invariants de similitude a_1, a_2, a_3, a_4 . Le polynôme minimal s'écrit sous la forme X^{α_1} avec $\alpha_1 = r_0$ où r_0 est l'indice i tel que $K_0^{i-1} \neq K_0^i = K_0^{i+k}$ pour tout $k \geq 0$, soit donc ici $a_1(X) = X^4$. De même on écrit les $a_i(X)$ sous la forme $a_i(X) = X^{\alpha_i}$ pour $2 \leq i \leq 4$ avec $\alpha_i \geq \alpha_{i+1}$ et $\sum_{i=1}^4 \alpha_i = 10$.

On introduit comme ci-avant $\delta_0^i = \dim K_0^i - \dim K_0^{i-1}$; δ_0^i est le nombre d'invariants de similitude divisibles par X^i . De $\delta_0^4 = 1$ on déduit $\alpha_2 \leq 3$; en outre $\delta_0^3 = 2$ impose $\alpha_2 \geq 3$ soit $\alpha_2 = 3$ et $\alpha_3 \leq 2$. Enfin $\delta_0^2 = 3$ donne $\alpha_3 = 2$ et $\alpha_4 = 1$.

Exercice 36. A quelles conditions sur les invariants de similitude de A nilpotent, l'équation $X^2 = A$ a-t-elle des solutions ?

Preuve : On raisonne par analyse et synthèse. Soit donc X nilpotent que l'on écrit sous forme de Jordan : x_k est le nombre de blocs de Jordan de taille k . La décomposition de Jordan de J_k^2 comporte deux blocs $J_{\lfloor \frac{k}{2} \rfloor}$ et $J_{\lceil \frac{k}{2} \rceil}$.

On en déduit alors que le tableau de Young associé à X^2 vérifie l'une des conditions équivalentes suivantes :

- il ne contient pas deux colonnes consécutives de même longueur impaire ;
- si on groupe les lignes deux par deux en partant du haut (en partant de la convention que l'on a une dernière ligne de longueur nulle dans le cas où le noyau est de dimension impaire), alors les lignes d'une même paire diffèrent d'au plus une case.

La synthèse est alors évidente.

Exercice 37. Donner, en fonction des invariants de similitude, la dimension du commutant d'un endomorphisme nilpotent.

Preuve : Il s'agit de déterminer le nombre de degré de liberté dans le choix d'un opérateur M qui commute avec A . On raisonne dans une base de Jordanisation de A . On rappelle que l'on a

$$\text{Ker } A \subsetneq \text{Ker } A^2 \subsetneq \cdots \subsetneq \text{Ker } A^r = \text{Ker } A^{r+1}$$

On considère une base $e_n, \dots, e_{n-d_r+1}$ de $\text{Ker } A^r - \text{Ker } A^{r-1}$ de cardinal la longueur d_r de la dernière colonne du tableau de Young associé à A . L'image de cette base est totalement libre ce qui donne $d_r n$ degré de liberté; en contrepartie l'image des u^k de ces vecteurs sont fixés. Soit alors r_1 maximal tel que $d_{r_1} \neq d_r$; on obtient alors $d_{r_1} \dim \text{Ker } A^{r_1}$ nouveaux degrés de liberté. On procède ainsi de suite jusqu'à épuiser tout l'espace.

On vérifie alors aisément qu'on obtient un nombre de degré de liberté égal à la somme des carrés des longueurs des colonnes du tableau de Young.

Exercice 38. (a) Montrer que le sous-espace vectoriel engendré par le cône nilpotent est l'hyperplan des matrices de trace nulle.

(b) Montrer que le sous-espace vectoriel engendré par les matrices nilpotentes de rang 1 est le même hyperplan.

(c) En déduire que le sous-espace vectoriel engendré par les matrices d'une classe de similitude quelconque de matrices nilpotentes est l'hyperplan des matrices de trace nulle.

Preuve : Dans les trois cas, l'inclusion est immédiate. On va montrer directement (b). Comme d'habitude cela repose sur un petit calcul en dimension 2, à savoir : $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ est semblable à $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ en considérant la nouvelle base $e_1 + e_2$ et $e_1 - e_2$.

Soit alors A une matrice de trace nulle; en ajoutant une combinaison linéaire de matrice nilpotente de rang 1, on se ramène à A diagonale $\text{diag}(a_1, \dots, a_n)$ avec $\sum_i a_i = 0$ que l'on écrit sous la forme

$$\text{diag}(a_1, -a_1, 0, \dots, 0) + \text{diag}(0, a_2 + a_1, a_3, \dots, a_n).$$

D'après le calcul précédent la première matrice est semblable à une combinaison linéaire de matrice nilpotentes de rang 1; la deuxième aussi par hypothèse de récurrence.

(c) L'orbite d'une classe de similitude quelconque contient dans son adhérence la classe de similitude des matrices nilpotentes de rang 1. On conclut alors d'après (b).

Exercice 39. Montrer que tout hyperplan H de $\mathbb{M}(n, \mathbb{C})$ contient au moins $n^2 - n - 1$ matrices nilpotentes linéairement indépendantes.

Preuve : On se ramène au cas où H a pour équation $\text{tr}(TX) = 0$ avec T triangulaire et on considère les intersections de H avec les sous-espaces des matrices nilpotentes triangulaires supérieures ou inférieures.

Exercice 40. Soit $M = S + N$ la décomposition de Dunford de M en semi-simple plus nilpotent. Montrer que S est dans l'adhérence de la classe de similitude de M .

Preuve : Cela découle simplement du fait que 0 est dans l'adhérence de la classe de similitude de N .

4 Algèbre bilinéaire

4.1 Formes sesquilineaires : généralités

Rappelons qu'étant donné un automorphisme σ du corps $\mathbb{K}$, par exemple la conjugaison complexe de $\mathbb{C}$, une *application semi-linéaire* est une application θ telle que pour tout $x, y \in E$ et $\lambda \in \mathbb{K}$ on a

$$\theta(x + \lambda y) = \theta(x) + \lambda^\sigma \theta(y)$$

où par convention on note λ^σ pour $\sigma(\lambda)$.

Définition 117. On appelle forme σ -sesquilineaire toute application $\phi : E \times E \rightarrow \mathbb{K}$ vérifiant les conditions suivantes :

- pour tout $x \in E$, l'application $\phi_x : y \in E \mapsto \phi(x, y)$ est linéaire ;
- pour tout $y \in E$ l'application $\phi_y : x \in E \mapsto \phi(x, y)$ est σ -linéaire.

Remarque: les notations ϕ_x et ϕ_y ne sont pas exemplaires, on veillera à ne pas se mélanger.

Proposition 118. Soit E un espace vectoriel muni d'une base $(e_i)_{1 \leq i \leq n}$. On note $A_\phi = (\phi(e_i, e_j)) \in \mathbb{M}_n(\mathbb{K})$ la matrice associée à la forme sesquilineaire ϕ , relativement à la base $(e_i)_i$. Pour tous vecteurs $x, y \in E$ de vecteur colonne coordonnées X et Y , on a alors

$$\phi(x, y) = {}^t X^\sigma A_\phi Y.$$

Remarque: si $P_{(e_i) \leftarrow (e'_i)}$ est la matrice de changement de base de $(e_i)_i$ à $(e'_i)_i$ alors la matrice A'_ϕ relativement à cette nouvelle base est telle que

$$A'_\phi = {}^t P_{(e_i) \leftarrow (e'_i)}^\sigma A_\phi P_{(e_i) \leftarrow (e'_i)}.$$

En particulier le déterminant de A_ϕ , qu'on appelle le discriminant de ϕ , n'est défini que comme élément de $\mathbb{K}/N(\mathbb{K})$ où $N(\mathbb{K}) = \{\lambda \lambda^\sigma, \lambda \in \mathbb{K}\}$.

Définition 119. Pour M une partie de E , on note

$$M^\perp = \{y \in E, \phi(M, y) = 0\}, \quad {}^\perp M = \{x \in E, \phi(x, M) = 0\}.$$

On dit que $M^\perp$ (resp. ${}^\perp M$) est l'orthogonal à droite (resp. à gauche) de M .

Remarque: $M^\perp$ et ${}^\perp M$ sont clairement des sous-espaces de E .

Lemme 120. *Si M est un sous-espace de E alors*

$$\dim M + \dim M^\perp = \dim E + \dim(M \cap {}^\perp E).$$

Preuve : Soit $f : E \longrightarrow E^*$ l'application semi-linéaire qui à x associe $y \mapsto \phi(x, y)$. Par définition $M^\perp$ est l'orthogonal au sens usuel, cf. la proposition 26, de $f(M)$ de sorte que

$$\dim M^\perp = \dim E - \dim f(M) = \dim E - (\dim M - \dim(M \cap {}^\perp E))$$

puisque ${}^\perp E = \text{Ker } f$.

Remarque: si on applique la formule à $M = E$, on obtient $\dim E^\perp = \dim {}^\perp E$.

Définition 121. *On dit que ϕ est non dégénérée si $E^\perp = \{0\}$ (resp. ${}^\perp E = \{0\}$). Le rang de A_ϕ est appelé le rang de ϕ , il est égal à la codimension de $E^\perp$ et ${}^\perp E$.*

Remarque: lorsque ϕ est non dégénérée, l'application $x \mapsto \phi_x$ induit un isomorphisme semi-linéaire *canonique* de E vers son dual E^* .

Lemme 122. *On a ${}^\perp(M^\perp) = M + {}^\perp E$.*

Remarque: En particulier si ϕ est non dégénérée, on retrouve la propriété habituelle de bidualité ${}^\perp(M^\perp) = M$.

Preuve : On a clairement $M + {}^\perp E \subset {}^\perp(M^\perp)$ de sorte qu'il suffit de montrer l'égalité des dimensions. En raisonnant comme dans le lemme précédent, on a

$$\dim N + \dim {}^\perp N = \dim E + \dim(N \cap E^\perp).$$

On applique la formule à $N = M^\perp$ de sorte qu'en remarquant que $E^\perp \subset M^\perp$, on obtient

$$\dim {}^\perp(M^\perp) = \dim E - \dim M^\perp + \dim E^\perp = \dim M + \dim {}^\perp E - \dim(M \cap {}^\perp E)$$

car $\dim E^\perp = \dim {}^\perp E$ d'après ce qui précède, et on reconnaît la formule de Grassman qui donne la dimension de $M + {}^\perp E$.

Définition 123. *Une forme σ -sesquilinéaire est dite réflexive si pour tout $x, y \in E$, $\phi(x, y) = 0$ équivaut à $\phi(y, x) = 0$. Elle est dite hermitienne (resp. antihermitienne) si $\phi(x, y) = \epsilon \left(\phi(y, x) \right)^\sigma$ avec $\epsilon = 1$ (resp. $\epsilon = -1$).*

Remarque: pour une forme hermitienne ou antihermitienne, σ est nécessairement une involution ; dans le cas antihermitien en caractéristique différente de 2, on a même $\sigma = \text{Id}$ et on dit simplement que ϕ est *anti-symétrique*.

Proposition 124. *On suppose que ϕ est une forme hermitienne ou anti-hermitienne non dégénérée. Pour tout $u \in \mathcal{L}(E)$, il existe un unique endomorphisme $u^* \in \mathcal{L}(E)$, appelé adjoint de u , tel que pour tous $x, y \in E$:*

$$\phi(u(x), y) = \phi(x, u^*(y)).$$

En outre on a aussi $\phi(x, u(y)) = \phi(u^(x), y)$.*

Remarque: la donnée d'une forme σ -sesquilinéaire non dégénérée, induit un isomorphisme semi-linéaire canonique entre E et E^* de sorte que l'adjoint habituel d'un endomorphisme vu dans $\mathcal{L}(E^*)$ se voit comme un endomorphisme de E . L'égalité $\phi(u(x), y) = \phi(x, u^*(y))$ de la proposition est alors une simple traduction de l'isomorphisme entre $\mathcal{L}(E^*)$ et $\mathcal{L}(E)$ induit par ϕ .

Preuve : Considérons pour y fixé, la forme linéaire $x \mapsto \phi(y, u(x))$; comme ϕ est non dégénérée, il existe un vecteur dépendant de y que l'on note $u^*(y)$ telle que pour tout x on ait $\phi(u^*(y), x) = \phi(y, u(x))$ et donc aussi, en utilisant la nature hermitienne de ϕ , $\phi(u(x), y) = \phi(x, u^*(y))$. Enfin on vérifie aisément que $y \mapsto u^*(y)$ est linéaire.

On suppose à présent que ϕ est une forme hermitienne ou antihermitienne, auquel cas la caractéristique est en outre supposée différente de 2.

Définitions 125. — *Un vecteur x de E est dit isotrope si $\phi(x, x) = 0$.*
— *Un sous-espace F de E est dit isotrope si $F \cap F^\perp \neq \{0\}$.*
— *Un sous-espace F de E est dit totalement isotrope et on écrit SETI, si $F \subset F^\perp$.*
— *Un sési est dit maximal et on écrit SETIM, si pour tout SETI G contenant F alors $G = F$.*

Remarque: comme on est en dimension finie, tout SETI est contenu dans un SETIM.

Remarque: si F est non isotrope alors $E = F \oplus F^\perp$.

Proposition 126. *Tous les SETIM ont la même dimension appelée indice de ϕ .*

Preuve : Soient U et V deux SETIM et introduisons M et N qui sont respectivement des supplémentaires de $U \cap V$ dans U et V . On suppose par l'absurde $\dim U > \dim V$ de sorte que $r := \dim M > s := \dim N$. Pour $f_1, \dots, f_s$ une base de N , considérons l'application $M \rightarrow \mathbb{K}^s$ définie par

$$m \mapsto (\phi(f_1, m), \dots, \phi(f_s, m)).$$

Son noyau est $M \cap N^\perp$ de sorte que d'après le théorème du rang

$$\dim(M \cap N^\perp) = \dim M - \dim \Im \geq r - s > 0.$$

Considérons alors un vecteur non nul $x \in M \cap N^\perp \subset U$. Montrons que $x \in V^\perp$: soit donc $v \in V$ que l'on décompose $v = u + n$. On a alors

$\phi(x, v) = \phi(x, u) + \phi(x, n)$ avec $\phi(x, u) = 0$ car U est un SETI et $\phi(x, n) = 0$ car $x \in N^\perp$. Considérons alors $W = V + \mathbb{K}x$: pour tout $w = v + \lambda x$ on a

$$\phi(w, w) = \phi(v, v) + \lambda^2 \phi(x, x) + \bar{\lambda} \phi(v, x) + \lambda \phi(x, v),$$

où $\phi(v, v) = 0$ (resp. $\phi(x, x) = 0$) car V (resp. U) est un SETIM, et $\phi(v, x) = \phi(x, v) = 0$ car $x \in V^\perp$. Ainsi donc W est un SETI contenant strictement V alors que ce dernier était supposé maximal, d'où la contradiction.

Proposition 127. *Si ϕ est non dégénérée, il existe alors une décomposition dite de Witt de l'espace $E = F \oplus F' \oplus G$ avec F, F' des sétim et G un sous-espace non isotrope telle que la matrice de ϕ dans une base adaptée soit de la forme*

$$\begin{pmatrix} 0 & I_r & 0 \\ \epsilon I_r & 0 & 0 \\ 0 & 0 & B \end{pmatrix}.$$

Remarque: on donnera une preuve plus loin de ce résultat dans le cas où $\sigma = \text{Id}$, i.e. pour les formes quadratiques.

4.2 Endomorphismes remarquables

On suppose à présent E muni d'une formes hermitienne non dégénérée en caractéristique différente de 2.

Définition 128. *Un automorphisme u de E est dit unitaire si pour tous $x, y \in E$:*

$$\phi(u(x), u(y)) = \phi(x, y).$$

L'ensemble des automorphismes unitaires est un sous-groupe de $GL(E)$ noté U_ϕ et appelé le groupe unitaire de ϕ . Le noyau du morphisme déterminant d'image $\mathbb{H} = \{\lambda \in \mathbb{K} : \lambda \lambda^\sigma = 1\}$, est noté SU_ϕ et s'appelle le groupe spécial unitaire de ϕ .

Remarque: u est unitaire si et seulement si $u^{-1} = u^*$. Matriciellement la matrice U de u est unitaire si et seulement si ${}^t U^\sigma A_\phi U = A_\phi$. Dans le cas où $A_\phi = I_n$, on retrouve la condition usuelle ${}^t U^\sigma U = I_n$.

Définition 129. *Une similitude de rapport $\lambda \in \mathbb{K}^\times$ est un automorphisme tel que pour tous $x, y \in E$:*

$$\phi(u(x), u(y)) = \lambda \phi(x, y).$$

Le groupe des similitudes se note GU_ϕ .

Remarque: une définition classique d'une similitude consiste à demander :

$$\phi(x, y) = 0 \Rightarrow \phi(u(x), u(y)) = 0.$$

Définition 130. Un endomorphisme u est dit *auto-adjoint* s'il vérifie $u = u^*$.

Remarque: dans le cas réel (resp. complexe) on dit aussi *symétrique* (resp. *hermitien*).

Définition 131. Un endomorphisme u d'un espace hermitien est dit *normal* si $u \circ u^* = u^* \circ u$.

Théorème 132. Un endomorphisme normal est *semi-simple*.

Preuve : Nous allons montrer plus précisément que si F est stable par u normal alors $F^\perp$ est aussi u -stable. Considérons une base orthonormée de F que l'on complète en une base orthonormée de $F^\perp$. La matrice de u dans cette base est de la forme $\begin{pmatrix} A & B \\ 0 & C \end{pmatrix}$ et on calcule

$$\begin{aligned} MM^* &= \begin{pmatrix} A & B \\ 0 & C \end{pmatrix} \begin{pmatrix} A^* & 0 \\ B^* & C^* \end{pmatrix} = \begin{pmatrix} AA^* + BB^* & BC^* \\ CB^* & CC^* \end{pmatrix} \\ M^*M &= \begin{pmatrix} A^* & 0 \\ B^* & C^* \end{pmatrix} \begin{pmatrix} A & B \\ 0 & C \end{pmatrix} = \begin{pmatrix} A^*A & A^*B \\ B^*A & B^*B + C^*C \end{pmatrix} \end{aligned}$$

ce qui fournit les égalités

$$\begin{cases} AA^* + BB^* = A^*A \\ BC^* = A^*B \\ CC^* = B^*B + C^*C. \end{cases}$$

Comme $\text{tr}(AA^*) = \text{tr}(A^*A)$ on en déduit que $\text{tr}(BB^*) = 0$ et donc $B = 0$ d'où le résultat

Remarque: pour f normal, on a $\|f(x)\| = \|f^*(x)\|$ pour tout x et donc $\text{Ker } f = \text{Ker } f^*$. Plus généralement si λ est une valeur propre de f , en utilisant que $f - \lambda \text{Id}$ est aussi normal, on en déduit que le sous-espace propre $\text{Ker}(f - \lambda \text{Id})$ est égal au sous-espace propre $\text{Ker}(f^* - \bar{\lambda} \text{Id})$. Leur orthogonal commun est donc stabilisé à la fois par f et f^* . Ainsi dans le cas où le corps est algébriquement clos, on peut utiliser ces arguments pour prouver que f et f^* sont simultanément diagonalisables.

4.3 Formes quadratiques

On reprend les définitions du paragraphe précédent dans le cas où $\sigma = \text{Id}$, on parle alors de forme bilinéaire symétrique ϕ et $q(x) := \phi(x, x)$ est une *forme quadratique*. En caractéristique différente de deux, la formule de polarisation

$$\phi(x, y) = \frac{1}{4}(q(x+y) - q(x-y))$$

permet d'identifier les formes bilinéaires symétriques et les formes quadratiques. On garde le vocabulaire du paragraphe précédent avec les notions

d'isotropie, de forme non dégénérée permettant d'identifier canoniquement l'espace E avec son dual E^* . La traduction matricielle de $\phi(x, y)$ est tXAY et le changement de base s'exprime $A' = {}^tPAP$.

Définition 133. *Étant donné deux formes bilinéaires symétriques (V, ϕ) et (V', ϕ') , la somme orthogonale $(V, \phi) \perp (V', \phi')$ est la forme bilinéaire symétrique $\phi \perp \phi'$ définie sur $V \oplus V'$ par la formule*

$$(\phi \perp \phi')(x + x', y + y') = \phi(x, y) + \phi'(x', y'), \quad \forall x, y \in V, \forall x', y' \in V'.$$

Remarque: si V est muni d'une forme bilinéaire symétrique ϕ et de deux sous-espaces U, W tels que $V = U \oplus W$ et $\phi(u, w) = 0$ pour tout $u \in U$ et $w \in W$, alors

$$(V, \phi) \simeq (U, \phi|_U) \perp (W, \phi|_W).$$

En particulier on a $(V, \phi) \simeq (V^\perp, 0) \perp (V/V^\perp, \phi')$ avec ϕ' non dégénérée.

Lemme 134. *Soit (V, ϕ) une forme bilinéaire symétrique et W un sous-espace de V tel que $b|_W$ est non dégénérée. On a alors*

$$(V, \phi) = (W, \phi|_W) \perp (W^\perp, b|_{W^\perp}).$$

Preuve : Commençons par montrer que $V = W \oplus W^\perp$. Comme ϕ_W est non dégénérée, on a $W \cap W^\perp = \{0\}$. Soit alors $v \in V$ et $f_v \in W^*$ défini par

$$f_v : w \in W \mapsto \phi(w, v).$$

Comme $\phi|_W$ est non dégénérée l'application $w \in W \mapsto f_w \in W^*$ est un isomorphisme et il existe donc $w \in W$ tel que $f_v = f_w$ et donc $v - w \in W^\perp$ d'où l'affirmation. L'isomorphisme de l'énoncé découle alors de la remarque précédente.

Notation 3. *Pour $a_1, \dots, a_n \in K^\times$, on note $\langle a_1, \dots, a_n \rangle$ la forme bilinéaire symétrique sur K^n définie par*

$$(x, y) \in K^n \mapsto \sum_{i=1}^n a_i x_i y_i \in K.$$

Remarque: la matrice de $\langle a_1, \dots, a_n \rangle$ dans la base canonique est la matrice diagonale $\text{diag}(a_1, \dots, a_n)$. Comme les a_i sont non nuls, cette matrice est inversible et $\langle a_1, \dots, a_n \rangle$ est non dégénérée.

Théorème 135. *Toute forme bilinéaire symétrique est équivalente à une forme $\langle a_1, \dots, a_r \rangle$ pour des $a_i \in K^\times$ et où r est le rang de la forme.*

Preuve : On raisonne par récurrence sur la dimension de l'espace ; le cas d'une droite étant évident. La forme bilinéaire ϕ étant non nulle, de la formule de polarisation, on en déduit l'existence d'un vecteur v tel que $q(v) \neq 0$ de sorte que la restriction de ϕ à $W = K.v$ est non dégénérée. D'après le lemme précédent, on a $(V, \phi) = (W, \phi|_W) \perp (W^\perp, \phi|_{W^\perp})$ et on conclut par récurrence.

Remarque: l'énoncé signifie qu'il existe une base $(e_1, \dots, e_n)$ telle que $q(e_i) = a_i$ où pour $i > r$ on a posé $a_i = 0$. Notons $l_1, \dots, l_n$ la base duale associée, on a alors

$$q(x) = a_1 l_1(x)^2 + \dots + a_r l_r(x)^2.$$

L'algorithme de décomposition de Gauss permet de trouver directement les formes linéaires indépendantes $l_1, \dots, l_r$. Notons en outre que changer e_i en λe_i modifie a_i par $\lambda^2 a_i$, de sorte que les a_i ne sont à priori bien définis que dans $K^\times / K^{\times,2}$. Pour autant ce ne sont pas nécessairement des invariants au sens où $\langle a_1, \dots, a_n \rangle$ peut être isomorphe à $\langle a'_1, \dots, a'_n \rangle$ même, modulo permutations, les $a_i a'_i \notin K^{\times,2}$. Pour

- $K = \mathbb{C}$, comme tout élément non nul est un carré, $\phi \simeq \langle 1, \dots, 1 \rangle$ où le nombre de 1 est égal au rang de ϕ , qui est bien un invariant.
- Pour $K = \mathbb{R}$, on a $\phi \simeq \langle 1, \dots, 1, -1, \dots, -1 \rangle$. On note r (resp. s) le nombre 1 (resp. de -1); le couple (r, s) s'appelle la signature de ϕ . Pour montrer que ce sont des invariants, il suffit par exemple de noter que r (resp. s) est la dimension maximale d'un sous-espace sur lequel la restriction de ϕ est définie positive (resp. négative).
- Sur un corps fini $K = \mathbb{F}_q$, en dimension n on dispose de deux classes de formes bilinéaires symétriques non dégénérées à savoir $\langle 1, \dots, 1 \rangle$ et $\langle 1, \dots, 1, \alpha \rangle$ où α n'est pas un carré de $\mathbb{F}_q^\times$. Pour le démontrer on raisonne par récurrence sur n , le cas $n = 1$ étant évident. Soit alors $n \geq 2$; pour tout $\alpha, \beta \in \mathbb{F}_q^\times$, les ensembles $\{\alpha x^2 : x \in \mathbb{F}_q\}$ et $\{1 - \beta y^2 : y \in \mathbb{F}_q\}$ sont de cardinal $\frac{q+1}{2}$ et ne peuvent pas être disjoints. Ainsi l'équation $\alpha x^2 + \beta y^2 = 1$ admet une solution, i.e. il existe v tel que $q(v) = 1$. On applique alors l'hypothèse de récurrence à $(Kv)^\perp$.

Définition 136. Étant donné un vecteur v tel que $q(v) \neq 0$, la réflexion τ_v par rapport à $(Kv)^\perp$ est définie par

$$\tau_v(x) = x - 2 \frac{\phi(x, v)}{q(v)} v.$$

Pour $x = x_v + x' \in (Kv) \oplus (Kv)^\perp$, on a $\tau_v(x) = -x_v + x'$ et τ_v est une isométrie relativement à q , i.e. $\phi(x, y) = \phi(\tau_v(x), \tau_v(y))$.

Lemme 137. Soit ϕ une forme bilinéaire symétrique sur un K -espace vectoriel V et soient $x, y \in V$ tels que $q(x) = q(y) \neq 0$. Il existe alors une isométrie τ de V pour ϕ telle que $\tau(x) = y$.

Preuve : Soit $u = \frac{x+y}{2}$ et $v = \frac{x-y}{2}$ avec donc $x = u + v$ et $y = u - v$. Comme $q(x) = q(y)$, on calcule $\phi(u, v) = 0$ et $0 \neq q(x) = q(u) + q(v)$ de sorte que, quitte à échanger le rôle de x et y , on peut supposer que $q(v) \neq 0$. La réflexion τ_v de la définition précédente vérifie alors $\tau_v(x) = y$ d'où le résultat.

Théorème 138. (de Witt) Soient (V_1, ϕ_1) , (V_2, ϕ_2) et (V, ϕ) trois espaces munis d'une forme bilinéaire symétrique avec ϕ non dégénérée. Alors

$$(V_1, \phi_1) \perp (V, \phi) \simeq (V_2, \phi_2) \perp (V, \phi) \Leftrightarrow (V_1, \phi_1) \simeq (V_2, \phi_2).$$

Preuve : Comme ϕ est non dégénérée, elle est diagonalisable et il suffit donc de traiter le cas où $(V, \phi) = (Kv, \langle a \rangle)$. Pour $i = 1, 2$, on note $v_i = (0, v) \in V_i \perp V$ et $v'_2 = f(v_1)$ où $f : (V_1, \phi_1) \perp (Kv, \langle a \rangle) \simeq (V_2, \phi_2) \perp (Kv, \langle a \rangle)$. On applique alors le lemme précédent à v_2 et v'_2 avec $\tau(v'_2) = v_2$ de sorte que $\tau \circ f$ envoie v_1 sur v_2 et induit donc un isomorphisme entre leurs orthogonaux i.e. entre (V_1, ϕ_1) et (V_2, ϕ_2) d'où le résultat.

Proposition 139. Soit V un K -espace vectoriel de dimension 2 muni d'une forme bilinéaire symétrique ϕ non dégénérée. Les propriétés suivantes sont équivalentes :

- (i) ϕ est isotrope ;
- (ii) $\det \phi = -1 \in K^\times / K^{\times, 2}$;
- (iii) $\phi \simeq \langle 1, -1 \rangle$;
- (iv) il existe une base dans laquelle la matrice de ϕ est $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$.

Preuve : Nous allons suivre la chaine d'implications (iii) $\Rightarrow$ (ii) $\Rightarrow$ (i) $\Rightarrow$ (iv) $\Rightarrow$ (iii). Les seules qui ne sont pas évidentes sont (ii) $\Rightarrow$ (i) $\Rightarrow$ (iv).

- Montrons donc (ii) $\Rightarrow$ (i) : on a $\phi \langle \alpha, \beta \rangle$ avec $\alpha, \beta \in K^\times$ et $\alpha\beta = -1 \in K^\times / K^{\times, 2}$. On a donc $\phi \simeq \langle \alpha, -\alpha \rangle$ qui est bien isotrope puisque $q(e_1 + e_2) = 0$.

- Montrons (i) $\Rightarrow$ (iv) : soit alors $x \neq 0$ tel que $q(x) = 0$. Comme ϕ est non dégénérée, il existe y tel que $\phi(x, y) \neq 0$ et $\phi(x, x + \lambda y) = \lambda \phi(x, y)$. On peut donc choisir y tel que $\phi(x, y) = 1$ avec donc nécessairement y non colinéaire à x . On note $\alpha = q(y)$ et on pose $z = y - \frac{\alpha}{2}x$ de sorte que $q(z) = 0$ et $\phi(x, z) = 1$. Dans la base (x, z) , la matrice de ϕ est celle de (iv).

Définition 140. Une forme qui satisfait l'une des propriétés équivalentes de la proposition précédente est appelée plan hyperbolique : on le note H . Une base telle que la matrice de ϕ soit comme dans (iv) est dite hyperbolique. On appelle enfin forme hyperbolique une somme orthogonale de plans hyperboliques.

Lemme 141. Soit ϕ une forme bilinéaire symétrique non dégénérée sur un K -espace vectoriel V . Les assertions suivantes sont équivalentes :

- (i) la forme ϕ est hyperbolique ;
- (ii) il existe un sous-espace W de V avec $2 \dim W = \dim V$ avec $\phi|_W = 0$;
- (iii) il existe un sous-espace W de V tel que $W^\perp W$.

Preuve : - (i) $\Rightarrow$ (ii) : notons $(e_1, \dots, e_{2n})$ la base de V dans laquelle $\phi \simeq \langle 1, -1, \dots, 1, -1 \rangle$. Le sous-espace vectoriel W engendré par $e_{2i-1} + e_{2i}$ pour $i = 1, \dots, n$ vérifie alors (ii).

- L'implication (ii) $\Rightarrow$ (iii) vérifions (iii) $\Rightarrow$ (i). Soit $(e_1, \dots, e_n)$ une base de W . Pour tout $i = 2, \dots, n$, le sous-espace $(Ke_1)^\perp$ est de dimension $2n-1$ et contient W : d'après le théorème du rang il existe alors y orthogonal à $e_2, \dots, e_n$ et n'appartenant pas à W . D'après la proposition précédente, le plan H engendré par e_1 et y est alors hyperbolique et $W \cap H^\perp$ vérifie (iii) relativement à $H^\perp$. On conclut alors par récurrence sur la dimension.

Remarque: en particulier $(V, \phi) \perp (V, -\phi)$ est une forme hyperbolique puisque le sous-espace $W = \{(v, v) : v \in V\}$ vérifie le point (ii) du lemme précédent.

Lemme 142. *Soit ϕ une forme hyperbolique et ψ une forme non dégénérée. Alors $\phi \otimes \psi$ est hyperbolique.*

Preuve : Par hypothèse $\phi = \langle 1, -1 \rangle \perp \dots \perp \langle 1, -1 \rangle$ de sorte que

$$\phi \otimes \psi \simeq (\psi \perp -\psi) \perp \dots \perp (\psi \perp -\psi)$$

et le résultat découle de la remarque précédente.

Proposition 143. (Décomposition de Witt) *Soit ϕ une forme non dégénérée isotrope. Il existe alors une forme non dégénérée anisotrope (V_a, ϕ_a) et une forme hyperbolique (V_h, ϕ_h) telles que*

$$(V, \phi) \simeq (V_h, \phi_h) \perp (V_a, \phi_a).$$

Une telle décomposition est en outre unique à isomorphisme près.

Remarque: autrement dit pour étudier une forme bilinéaire symétrique non dégénérée, on se ramène aux formes anisotropes.

Preuve : Si ϕ est anisotrope il n'y a rien à faire. Sinon soit $x \neq 0$ tel que $q(x) = 0$ puis, cf. la preuve de la proposition précédente, y tel que $\phi(x, y) = 1$. Comme (x, y) est nécessairement libre, ils engendrent un plan hyperbolique W et la restriction ϕ' de ϕ à $W^\perp$ est non dégénérée ce qui permet d'itérer la construction jusqu'à obtenir une forme anisotrope.

Le fait que la décomposition soit unique à isomorphisme près découle du théorème de Witt 138 et du fait qu'une forme hyperbolique est de la forme $H \perp \dots \perp H$.

Définition 144. *Le nombre de facteurs H dans*

$$(V_h, \phi_h) \simeq H \perp \dots \perp H$$

est appelé l'indice de Witt de ϕ .

Notation 4. Soit $\phi \simeq \langle a_1, \dots, a_n \rangle$ et $\phi' \simeq \langle a'_1, \dots, a'_m \rangle$ des formes bilinéaires symétriques non dégénérées. On note $\phi \otimes \phi'$ la forme bilinéaire symétrique non dégénérée isomorphe à

$$\langle a_1 a'_1, \dots, a_1 a'_m, a_2 a'_1, \dots, a_n a'_m \rangle.$$

Remarque: la notation ci-dessus est un artifice pour ne pas avoir à définir canoniquement le produit tensoriel de deux espaces vectoriels.

Définition 145. On note $\mathcal{M}_K$ l'ensemble des classes d'isomorphismes de formes bilinéaires symétriques non dégénérées sur K et on le munit de la relation d'équivalence

$$\phi \sim \phi' \Leftrightarrow \phi_a \simeq \phi'_a.$$

On note $W(K)$ l'ensemble quotient.

Remarque: on vérifie aisément que $W(K)$ muni de la somme directe orthogonale $\perp$ et du produit tensoriel $\otimes$ est un anneau commutatif.

Exemples :

- dans $\mathbb{C}$, -1 est un carré et donc $\langle 1, \dots, 1 \rangle \simeq \langle 1, -1, 1, \dots \rangle$ qui est une forme hyperbolique si et seulement si la dimension de l'espace est pair. Ainsi $W(\mathbb{Z}) \simeq \mathbb{Z}/2\mathbb{Z}$ où la flèche est donnée par la dimension de l'espace modulo 2.
- Le même raisonnement dans le cas $K = \mathbb{R}$, montre que $W(\mathbb{R}) \simeq \mathbb{Z}$ où la flèche est donnée par $r - s$ où (r, s) est la signature.
- Pour $K = \mathbb{F}_p$ où $p \equiv 1 \pmod{4}$ comme -1 est un carré, on se ramène comme précédemment aux formes suivantes

$$\langle 1 \rangle, \langle \alpha \rangle, \langle 1, \alpha \rangle, H$$

où $\alpha \in \mathbb{F}_p^\times - \mathbb{F}_p^{\times, 2}$. Ces quatre formes sont distinctes dans $W(\mathbb{F}_p)$ et on vérifie aisément qu'elles sont toutes d'ordre 2

$$\langle 1 \rangle \perp \langle 1 \rangle = H \sim 0, \quad \langle \alpha \rangle \perp \langle \alpha \rangle \simeq 0, \quad \langle 1, \alpha \rangle \perp \langle 1, \alpha \rangle = \langle 1, 1 \rangle \perp \langle \alpha, \alpha \rangle \sim 0.$$

Ainsi $W(\mathbb{F}_p) \simeq (\mathbb{Z}/2\mathbb{Z})^2$.

- Terminons par le $K = \mathbb{F}_p$ avec $p \equiv 3 \pmod{4}$, on obtient alors les formes

$$\langle 1 \rangle, \langle -1 \rangle, \langle 1, 1 \rangle, H = \langle 1 - 1 \rangle$$

qui sont distinctes dans $W(\mathbb{F}_p)$. Or $\langle 1 \rangle \perp \langle 1 \rangle = \langle 1, 1 \rangle \not\sim H$ et donc $W(\mathbb{F}_p) \simeq \mathbb{Z}/4\mathbb{Z}$.

4.4 Le cas réel

Dans ce cas on a nécessairement $\sigma = \text{Id}$ et on revient donc sur les formes quadratiques du paragraphe précédent avec les notions de positivité.

Définition 146. Une forme bilinéaire symétrique est dite :

- positive (resp. négative) si pour tout $x \in E$, on a $\phi(x, x) \geq 0$ (resp. $\phi(x, x) \leq 0$) ;
- définie positive (resp. définie négative) si elle est positive (resp. négative) et que $\phi(x, x) = 0$ si et seulement si x est le vecteur nul.

Théorème 147. (Loi d'inertie de Sylvester)

Soit ϕ une forme bilinéaire symétrique.

- (i) Il existe alors une décomposition

$$E = E^\perp \oplus E^+ \oplus E^-$$

telle que la restriction de ϕ à E^+ (resp. E^-) est définie positive (resp. définie négative). Une telle décomposition n'est pas unique mais les dimensions r de E^+ et s de E^- sont les mêmes pour toute telle décomposition et sont respectivement égale au maximum des dimensions des sous-espaces F de E tels que la restriction de ϕ y soit définie positive (resp. négative). On dit que le couple (r, s) est la signature de ϕ .

- (ii) Le rang de ϕ est égal à $s + r$ et son indice est égal à $(n - \text{rg}\phi) + \min\{s, r\}$.

Preuve : (i) L'existence de la décomposition a été vue plus haut dans un cadre général, il reste alors à vérifier la caractérisation de r et s . Soit donc un espace F sur lequel q est défini positif et supposons par l'absurde que sa dimension est $> r$. D'après le théorème du rang, il intersecte alors non trivialement $E^\perp \oplus E^-$ espace sur lequel est $q(x) \leq 0$ d'où la contradiction. Ainsi r est bien égal à la dimension maximale d'un espace sur lequel q est définie positive.

(ii) Le rang de ϕ est clairement égal à $r + s$. Rappelons que son indice est la dimension d'une SETIM. Notons $(e_1^+, \dots, e_r^+)$ (resp. $(e_1^-, \dots, e_s^-)$) une base de E^+ (resp. E^-) ainsi que $(e_1^0, \dots, e_t^0)$ une base de $E^\perp$. Considérons alors

$$F = \text{Vect}(e_1^+ + e_1^-, \dots, e_{\min(r,s)}^+ + e_{\min(r,s)}^-, e_1^0, \dots, e_t^0).$$

On vérifie aisément que F est totalement isotrope de dimension $(n - \text{rg}\phi) + \min(r, s)$.

Supposons alors par l'absurde qu'il existe un SETI F de dimension $> (n - \text{rg}\phi) + \min(r, s)$. D'après le théorème du rang il intersecte alors non trivialement E^+ si $r > s$ et E^- sinon alors que pour tout x non nul de E^+ (resp. E^-) on a $q(x) > 0$ (resp. $q(x) < 0$), d'où la contradiction.

Application : tout sous- $\mathbb{R}$ -espace vectoriel du cône nilpotent, i.e. de l'ensemble des matrices nilpotentes de $M_n(\mathbb{R})$, est de dimension inférieur à $\frac{n(n-1)}{2}$. En effet on considère la forme quadratique q définie sur l'espace des matrices qui à X associe $\text{tr}X^2$. De manière évidente le cône isotrope est

constitué de vecteurs isotropes de sorte que l'espace vectoriel en question sera totalement isotrope. Par ailleurs, si $X \neq 0$ est symétrique (resp. anti-symétrique) alors $q(X) > 0$ (resp. $q(X) < 0$) de sorte que la signature de q est $(\frac{n(n+1)}{2}, \frac{n(n-1)}{2})$. Ainsi un sous-espace totalement isotrope est de dimension inférieure ou égale à $\frac{n(n-1)}{2}$. L'égalité est clairement atteinte pour les matrices strictement triangulaires supérieures.

Définition 148. Soit F un sous-espace non isotrope de E ; l'unique involution unitaire u telle que $F = \text{Im}(u + \text{Id})$ s'appelle la symétrie orthogonale par rapport au sous-espace non isotrope F . Si F est un hyperplan, on dit que u est une réflexion; si F est de codimension 2 on dit que u est un retournement ou un renversement.

Remarque: si v est un vecteur normé orthogonal à un hyperplan H non isotrope, alors la réflexion par rapport à H est l'application $x \mapsto x - 2\phi(x, v)v$.

Théorème 149. (de Cartan-Dieudonné)

Soit q une forme quadratique non dégénérée sur E .

- (i) Tout élément u de $O(q)$ peut s'écrire comme un produit de $p := \dim \mathfrak{S}(u - \text{Id})$ réflexions; en outre toute écriture de u comme composé de réflexions exige au moins p réflexions différentes.
- (ii) En dimension ≥ 3 , tout élément de $SO(q)$ peut s'écrire comme un produit de q retournements avec $q \leq \dim E$.

Remarque: En dimension 2, on a un isomorphisme de groupes $\mathbb{R}/2\pi\mathbb{Z} \rightarrow SO(2)$ qui à θ associe la matrice

$$R(\theta) := \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}.$$

On remarque alors qu'il n'est pas possible d'exhiber une sous-famille génératrice particulière.

Preuve : (i) Notons tout d'abord que si r_H désigne la réflexion orthogonale relativement à l'hyperplan H , alors $u = r_{H_1} \circ \dots \circ r_{H_r}$ vérifie $H_1 \cap \dots \cap H_r \subset \text{Ker}(u - \text{Id})$ de sorte que $\dim \mathfrak{S}(u - \text{Id}) = n - \dim \text{Ker}(u - \text{Id}) \leq r$ et donc toute écriture de u comme un composé de réflexions nécessite au moins $\dim \mathfrak{S}(u - \text{Id})$ facteurs.

Montrons alors par récurrence sur $p = \dim \mathfrak{S}(u - \text{Id})$, qu'il existe une écriture $u = r_1 \circ \dots \circ r_q$ avec $q \leq p$. Pour $p = 0$ c'est clair puisqu'alors $u = \text{Id}$. Supposons le résultat acquis jusqu'à $p-1$. Considérons alors $x \in \text{Ker}(u - \text{Id})^\perp$ tel que $u(x) \neq x$ et soit r_0 la réflexion orthogonale relativement à l'hyperplan H orthogonal à $u(x) - x$. Comme $u \in O(q)$, on a $u(x) \in \text{Ker}(u - \text{Id})^\perp$ et donc $\text{Ker}(u - \text{Id}) \subset H$ et donc $\text{Ker}(u - \text{Id}) + \text{Vect}(x) \subset \text{Ker}(r_0 \circ u - \text{Id})$. Ainsi $\dim \mathfrak{S}(r_0 \circ u - \text{Id}) \leq p-1$ et d'après l'hypothèse de récurrence, il existe une écriture $r_0 \circ u = r_1 \circ \dots \circ r_q$ avec $q \leq p-1$ et le résultat s'en déduit en composant à gauche par r_0 .

(ii) Il suffit de noter que pour r une réflexion orthogonale relativement à un hyperplan H , alors $-r$ est un retournement. Ainsi pour $u = (r_1 \circ r_2) \circ \cdots \circ (r_{2k-1} \circ r_{2k})$, où on notera que le nombre de facteurs est nécessairement pair pour $u \in SO(q)$, en écrivant $r_{2i-1} \circ r_{2i} = (-r_{2i-1}) \circ (-r_{2i})$, on fait bien apparaître u comme un composé de retournements.

En dimension 3 toute matrice de $SO(3)$ est semblable à une matrice de la forme

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

On a aussi un isomorphisme de groupes entre $SO(3)$ et le groupe des quaternions de norme 1 quotienté par $\{\pm 1\}$. Cette description est particulièrement utile quand il s'agit de composer des rotations de l'espace.

Proposition 150. *Un endomorphisme u est normal si et seulement si dans une base orthonormée il admet une matrice diagonale par blocs de taille 1 ou 2 de la forme $\begin{pmatrix} a & b \\ -b & a \end{pmatrix}$.*

Preuve : Comme sur $\mathbb{R}$ un polynôme irréductible est de degré au plus 2, il résulte de (3) et en utilisant que les sous-modules de $\mathbb{R}[X]/(P)$ correspondent aux diviseurs de P , que u admet soit une droite soit un plan stable. Si on a une droite stable et comme d'après 132 un endomorphisme normal est semi-simple, on conclut par récurrence sur la dimension. Si on a un plan stable et comme la restriction de u à un sous-espace stable est encore normal, pour $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, l'égalité $MM^* = M^*M$ donne

$$\begin{cases} b^2 = c^2 \\ ac + bd = ab + dc \end{cases}$$

et donc

- soit $c = b$ auquel cas la matrice est symétrique avec un polynôme caractéristique scindé, i.e. u admet une droite stable ;
- soit $c = -d$ et $a = d$ et on retrouve la matrice de l'énoncé.

On conclut encore par récurrence sur la dimension en utilisant la semi-simplicité de u .

Réciproquement, il suffit de vérifier que la matrice $\begin{pmatrix} a & b \\ -b & a \end{pmatrix}$ est normale, ce qui a déjà été fait ci-avant.

Corollaire 151. *Un endomorphisme u de E est*

- *symétrique si et seulement si il est diagonalisable en base orthonormée.*
- *anti-symétrique si et seulement si, dans une base orthonormée, sa matrice est diagonale par blocs nuls ou de la forme $\begin{pmatrix} 0 & b \\ -b & 0 \end{pmatrix}$.*

- *orthogonal si et seulement si, dans une base orthonormée, sa matrice est diagonale par blocs où les blocs sont soit ± 1 soit des matrices de rotation $\begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}$.*

Remarque : on peut utiliser ce théorème de réduction pour montrer, par exemple, que le groupe spécial orthogonal est connexe par arcs, en reliant toute matrice orthogonale positive O à la matrice identité. Pour ce faire, on transforme toute matrice de rotation de O et d'angle θ , en une matrice de rotation d'angle $t\theta$ pour $t \in [0, 1]$, et en regroupant les blocs de -1 deux par deux pour les identifier à une matrice de rotation d'angle π , que l'on transforme en une matrice d'angle $t\pi$.

Proposition 152. *Tout sous-groupe compact G de $GL_n(\mathbb{R})$ est contenu dans un conjugué du groupe orthogonal.*

Preuve : Il s'agit donc de montrer que G stabilise une forme quadratique définie positive. En effet si A est la matrice d'une telle forme, elle est diagonalisable en base orthonormée ${}^tPAP = D$ avec ${}^tP = P^{-1}$ et $D = \text{diag}(\lambda_1, \dots, \lambda_n)$ où les λ_i sont strictement positifs. On peut alors écrire $D = D_1^2$ où D_1 est une matrice diagonale à valeurs propres strictement positives de sorte que $A = B^2$ avec $B = {}^tPD_1P$. Ainsi pour tout $M \in G$, on a ${}^tBB = {}^tMB^2M$ soit

$${}^t(BMB^{-1})(BMB^{-1}) = 1$$

i.e. BMB^{-1} est orthogonale.

Considérons alors l'action de G sur l'ensemble $\mathfrak{S}_n^{++}$ des formes quadratiques définies positives selon la formule ${}^tBB \mapsto {}^tM{}^tBBM = {}^t(BM)BM$. L'idée est d'utiliser un théorème de point fixe ce qui nécessite de se restreindre à un convexe compact. Pour avoir un compact il suffit de considérer $\{{}^tBB : B \in G\}$ et pour la convexité de prendre l'enveloppe convexe de ce compact qui reste donc compact d'après le théorème de Carathéodory. On a ainsi l'action d'un groupe compact sur un convexe compact et le lemme 154 ci-après assure l'existence d'un point fixe. Il ne reste plus qu'à vérifier alors que le point fixe correspond à une forme définie positive, ce qui découle du lemme suivant.

Lemme 153. *L'ensemble $\mathfrak{S}_n^{++}$ des formes quadratiques définies positives est convexe.*

Preuve : L'idée est de faire de la congruence simultanée. Soient donc A et B deux matrices symétriques définies positives : A définit alors un produit scalaire ϕ_A . La matrice $U = A^{-1}B$ définit un endomorphisme symétrique relativement à ϕ_A puisque $AU = {}^tUA$ et donc il existe une base orthonormée pour ϕ_A diagonalisant U , i.e. ${}^tPAP = I_n$ et $P^{-1}UP = D$ et donc

$(P^{-1}A^{-1}P^{-1})^tPBP = D$ soit ${}^tPBP = D$. Ainsi pour $0 \leq t \leq 1$, on a

$$tA + (1-t)B = {}^tP^{-1}(tI_n + (1-t)D)P^{-1}$$

qui est bien définie positive.

On conclut à présent avec le lemme de point fixe nécessaire pour finir de prouver la proposition précédente.

Lemme 154. *Soit $G \subset GL(E)$ un groupe compact agissant sur un compact convexe K d'un espace vectoriel euclidien E . Il existe alors un point de K fixé par tous les éléments de G .*

Preuve : On introduit l'application

$$x \in E \mapsto N_G(x) = \sup_{g \in G} \|g(x)\|,$$

où le sup est en fait un maximum, i.e. est atteint, car G est compact. On vérifie aisément que N_G est une norme qui est de plus strictement convexe puisque

$$N_G(x+y) = \|g_0(x) + g_0(y)\| \leq \|g_0(x)\| + \|g_0(y)\| \leq N_G(x) + N_G(y)$$

et pour avoir égalité il faut en particulier que la première inégalité soit une égalité et donc qu'il existe $\lambda > 0$ tel que $g_0(x) = \lambda g_0(y)$ et comme g_0 est linéaire et inversible, $x = \lambda y$, d'où l'affirmation.

Ainsi il existe un unique $x_0 \in K$ minimisant N_G et comme trivialement pour tout x on a $N_G(g(x)) = N_G(x)$, d'après l'unicité du minimum on a $g(x_0) = x_0$, d'où le résultat.

4.5 Le cas hermitien

Dans ce paragraphe pour $\mathbb{K} = \mathbb{C}$, pour ne pas répéter le cadre des formes quadratiques, on considère le cas où σ est la conjugaison complexe. Pour $A \in GL_n(\mathbb{C})$, on note A^* pour ${}^t\bar{A}$. Notons en particulier que toute forme hermitienne ϕ vérifie $\phi(x, x) \in \mathbb{R}$. On dit alors qu'elle est *positive* (resp. *négative*) si $\phi(x, x) \geq 0$ (resp. ≤ 0) pour tout $x \in E$ et on dit qu'elle est en outre *définie* si $\phi(x, x) = 0 \Rightarrow x = 0$.

Remarque: si E est muni d'une forme hermitienne définie positive on dit que E est un *espace hermitien*.

Proposition 155. *Si ϕ est positive alors*

$$\phi(x, y)\overline{\phi(x, y)} \leq \phi(x, x)\phi(y, y).$$

Théorème 156. *Comme dans le cas réel,*

- il existe une décomposition $E = E^\perp \oplus E^+ \oplus E^-$ telle que la restriction de ϕ à E^+ (resp. E^-) est définie positive (resp. négative). En outre la dimension s de E^+ et t de E^- sont indépendantes de cette décomposition et le couple (s, t) s'appelle la signature de ϕ .
- Il existe une base $(e_i)_{1 \leq i \leq n}$ telle que

$$\phi\left(\sum_{i=1}^n \lambda_i e_i, \sum_{j=1}^n \mu_j e_j\right) = \sum_{i=1}^s \lambda_i \overline{\mu_i} - \sum_{i=s+1}^{s+t} \lambda_i \overline{\mu_i}.$$

- Le rang de ϕ est $s + t$ et son indice $n - (s + t) + \min\{s, t\}$.

Terminons ces rappels algébriques par un exemple fondamental de matrice hermitienne.

Définition 157. Soit $(x_1, \dots, x_m)$ une famille de vecteurs d'un espace hermitien E . La matrice de Gram définie par

$$\text{Gram}(x_1, \dots, x_m) = \left(\phi(x_i, x_j) \right)_{1 \leq i, j \leq m}$$

est une matrice hermitienne dont on note $G(x_1, \dots, x_m)$ le déterminant.

Proposition 158. Pour tout $(x_1, \dots, x_m)$, le déterminant de Gram $G(x_1, \dots, x_m)$ appartient aux réels positifs; il est non nul si et seulement si la famille $(x_1, \dots, x_m)$ est libre.

Corollaire 159. Soient $x \in E$ et F un sous-espace de E ; si $(x_1, \dots, x_m)$ est une base de F alors la distance $d(x, F)$ de x à F est donnée par

$$d(x, F)^2 = \frac{G(x, x_1, \dots, x_m)}{G(x_1, \dots, x_m)}.$$

Proposition 160. Un endomorphisme est normal si et seulement s'il est diagonalisable en base orthonormée.

Preuve : Si u est normal, il est alors semi-simple d'après 132 : pour x un vecteur propre de u , on a $E = \langle x \rangle \oplus \langle x \rangle^\perp$ et on conclut par récurrence sur la dimension.

La réciproque est évidente et découle du fait qu'une matrice diagonale commute avec son adjointe diagonale.

Corollaire 161. Un endomorphisme d'un espace hermitien est

- hermitienne (resp. anti-hermitienne) si et seulement si, dans une base orthonormée, sa matrice est diagonale réelle (resp. imaginaire pure).
- unitaire si et seulement si, dans une base orthonormée, sa matrice est diagonale avec des coefficients de module 1.

4.6 Classes de congruences

Définition 162. Deux matrices carrées A et B de $\mathbb{M}_n(\mathbb{C})$ sont dites congruentes, s'il existe une matrice inversible P telle que $B = P^*AP$.

La relation de congruence est clairement une relation d'équivalence. Cette relation n'est réellement utilisée que pour des matrices hermitiennes et correspond à l'effet d'un changement de base sur la matrice associée à un produit hermitien, cf. la proposition 118. Dans ce cadre, nous verrons que les classes sont paramétrées par la signature, cf. le théorème 147 dans le cas réel et 156 dans le cas hermitien. En revanche contrairement au cas des classes de similitudes, on a un énoncé de congruence simultanée suivant.

Proposition 163. Soit deux matrices hermitiennes A et B avec B définie positives. Il existe alors une matrice inversible P vérifiant

$$P^*AP = D \quad \text{et} \quad P^*BP = I_n,$$

où D est une matrice diagonale.

Remarque: on notera bien que la matrice P n'est pas unitaire de sorte que A et B ne sont pas simultanément diagonalisable !

Preuve : On note Q une matrice dont les vecteurs colonnes forment une base orthonormée pour le produit hermitien défini par B , i.e. $(X, Y) \mapsto X^*BY$. On a ainsi $Q^*BQ = I_n$. Comme Q^*AQ est visiblement hermitienne, on la diagonalise en base orthonormée pour le produit hermitien canonique, i.e. il existe une matrice unitaire R telle que $R^*(Q^*AQ)R = D$. On a alors $R^*(Q^*BQ)R = R^*R = I_n$, de sorte que la matrice $P = QR$ convient.

4.7 Classes de similitudes unitaires

On s'intéresse ici aux classes de similitudes unitaires. Notons tout d'abord que si A et B sont dans la même classe de similitude unitaire alors

$$\sum_{i,j=1}^n |b_{ij}|^2 = \sum_{i,j=1}^n |a_{i,j}|^2$$

En effet cela découle de l'égalité $\sum_{i,j=1}^n |a_{i,j}|^2 = \text{tr}(A^*A)$. Plus généralement étant donnés un mot $M(s, t) = s^{m_1}t^{n_1}s^{m_2}t^{n_2}\dots s^{m_k}t^{n_k}$ en deux variables s, t avec $m_1, n_1, \dots, m_k, n_k \geq 0$: le degré de $M(s, t)$ est par définition égal à $m_1 + n_1 + \dots + m_k + n_k$. Pour $A \in \mathbb{M}_n(\mathbb{C})$, on pose

$$M(A, A^*) = A^{m_1}(A^*)^{n_1}\dots A^{m_k}(A^*)^{n_k}$$

On remarque alors que pour tout $M(s, t)$, $\text{tr}M(A, A^*)$ est constant sur la classe de similitude unitaire.

Théorème 164. (Specht) Deux matrices A, B sont unitairement semblables si et seulement si $\text{tr}M(A, A^*) = \text{tr}M(B, B^*)$ pour tout mot $M(s, t)$.

L'inconvénient évident du théorème de Specht est qu'il faut vérifier une infinité de conditions. Le théorème de Percy suivant permet de se ramener à un nombre fini.

Théorème 165. (Percy [?]) Deux matrices A, B sont unitairement semblables si et seulement si $\text{tr}M(A, A^*) = \text{tr}M(B, B^*)$ pour tout mot $M(s, t)$ de degré au plus $2n^2$.

Remarque: un rapide calcul montre qu'il y a au plus 4^{n^2} mots distincts de degré au plus $2n^2$. Pour $n = 2$, il est facile de montrer qu'il suffit en fait de considérer les mots s, s^2 et ts . Pour $n = 3$, on peut montrer qu'il suffit de considérer les 9 mots, $s, s^2, ts, ts^2, t^2s^2, tsts, ts^2ts, ts^2t^2s$ au lieu de tous les mots de degré au plus 18.

4.8 Valeurs propres de matrices hermitiennes

Pour une matrice complexe générale, les valeurs propres sont les racines du polynôme caractéristique. Si la matrice est hermitienne, les valeurs propres sont en outre les solutions de plusieurs problèmes d'optimisation : ce paragraphe est consacré à la présentation de quelques uns de ceux-ci, le lecteur intéressé pourra consulter [?] §4.2. Dans la suite A désigne une matrice complexe hermitienne dont les valeurs propres réelles sont classées par ordre croissant :

$$\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n.$$

Théorème 166. (Courant-Fisher) Pour tout $1 \leq k \leq n$, on a

$$\lambda_k = \min_{F \in \mathcal{E}_k} \max_{0 \neq x \in F} \frac{x^* Ax}{x^* x} = \max_{F \in \mathcal{E}_{n-k+1}} \min_{0 \neq x \in F} \frac{x^* Ax}{x^* x}$$

où $\mathcal{E}_k$ désigne l'ensemble des sous-espaces de dimension k de $\mathbb{C}^n$.

Remarque: le cas de λ_1 et λ_n est connu sous le nom du théorème de Rayleigh-Ritz.

Preuve : Notons pour tout $1 \leq k \leq n$, x_k un vecteur propre unitaire pour la valeur propre λ_k . Soit alors $F \in \mathcal{E}_k$ de sorte que

$$F \cap \text{vect}(x_k, x_{k+1}, \dots, x_n)$$

est de dimension supérieure à 1. Pour $x = \sum_{i=k}^n \alpha_i x_i \neq 0$ un vecteur unitaire de cette intersection, on a

$$x^* Ax = \sum_{i=k}^n \lambda_i |\alpha_i|^2 \geq \lambda_k \sum_{i=k}^n |\alpha_i|^2 = \lambda_k.$$

On en déduit alors l'inégalité

$$\lambda_k \geq \min_{F \in \mathcal{E}_k} \max_{0 \neq x \in F} \frac{x^* A x}{x^* x}.$$

Par ailleurs pour $F = \text{vect}(x_1, \dots, x_k)$, on a

$$\lambda_k = \max_{0 \neq x \in F} \frac{x^* A x}{x^* x},$$

d'où l'égalité. L'autre cas se traite de manière strictement identique.

Citons quelques applications du théorème précédent, cf. [?] §4.3.

Théorème 167. (Weyl) Soient A, B des matrices hermitiennes de valeurs propres $\lambda_i(A), \lambda_i(B)$ classées par ordre croissant. En classant de même les valeurs propres de $A + B$, on obtient :

- (a) $\lambda_k(A) + \lambda_1(B) \leq \lambda_k(A+B) \leq \lambda_k(A) + \lambda_n(B)$, pour tout $k = 1, \dots, n$;
- (b) si B est de rang au plus r :
 - $\lambda_k(A+B) \leq \lambda_{k+r}(A) \leq \lambda_{k+2r}(A+B)$ pour $k = 1, \dots, n-2r$;
 - $\lambda_k(A) \leq \lambda_{k+r}(A+B) \leq \lambda_{k+2r}(A)$, pour $k = 1, \dots, n-2r$;
- (c) $\lambda_{j+k-n}(A+B) \leq \lambda_j(A) + \lambda_k(B)$, pour tout $1 \leq j, k \leq n$ tels que $j+k \geq n+1$;
- (d) $\lambda_j(A) + \lambda_k(B) \leq \lambda_{j+k-1}(A+B)$, pour tout $1 \leq j, k \leq n$ tels que $j+k \leq n+1$.

Preuve : (a) On a $\frac{x^*(A+B)x}{x^*x} = \frac{x^*Ax}{x^*x} + \frac{x^*Bx}{x^*x}$; le résultat découle alors simplement de l'encadrement $\lambda_1(B) \leq \frac{x^*Bx}{x^*x} \leq \lambda_n(B)$ et du théorème 166.

(b) On écrit B sous la forme $\alpha_1 u_1 u_1^* + \dots + \alpha_r u_r u_r^*$ où les vecteurs u_i de $\mathbb{C}^n$ ne sont pas nécessairement indépendants. On a alors

$$\begin{aligned} \lambda_{k+2r}(A+B) &= \min_{F \in \mathcal{E}_{k+2r}} \max_{0 \neq x \in F} \frac{x^*(A+B)x}{x^*x} \\ &\geq \min_{F \in \mathcal{E}_{k+2r}} \max_{0 \neq x \in F \cap \text{vect}(u_1, \dots, u_r)^\perp} \frac{x^*(A+B)x}{x^*x}. \end{aligned}$$

En notant $\mathcal{E}'_{k+r}$ l'ensemble des sous-espaces de $\text{vect}(u_1, \dots, u_r)^\perp$ de dimension $k+r$, le dernier terme ci-dessus est égal à

$$\min_{F \in \mathcal{E}'_{k+r}} \max_{0 \neq x \in F} \frac{x^* A x}{x^* x} \geq \min_{F \in \mathcal{E}_{k+r}} \max_{0 \neq x \in F} \frac{x^* A x}{x^* x} = \lambda_{k+r}(A).$$

Selon le même schéma, on a

$$\begin{aligned} \lambda_k(A+B) &= \max_{F \in \mathcal{E}_{n-k+1}} \min_{0 \neq x \in F} \frac{x^*(A+B)x}{x^*x} \\ &\leq \max_{F \in \mathcal{E}_{n-k+1}} \min_{0 \neq x \in F \cap \text{vect}(u_1, \dots, u_r)^\perp} \frac{x^*(A+B)x}{x^*x} \\ &= \max_{F \in \mathcal{E}'_{n-k+1-r}} \min_{0 \neq x \in F} \frac{x^* A x}{x^* x} \\ &\leq \max_{F \in \mathcal{E}_{n-k+1-r}} \min_{0 \neq x \in F} \frac{x^* A x}{x^* x} \\ &= \lambda_{k+r}(A) \end{aligned}$$

Ces deux familles d'inégalités fournissent alors celles de l'énoncé.

(c) On diagonalise $A = UD_1U^*$ et $B = VD_2V^*$ et on note u_i (resp. v_i) les vecteurs colonnes de la matrice unitaire U (resp. V). Pour un couple (j, k) vérifiant les conditions de l'énoncé, on note pour β assez grand tel que pour tout $j+1 \leq i \leq n$ et pour tout $k+1 \leq i' \leq n$, $\lambda_i(A) - \beta < \lambda_j(A)$ et $\lambda_{i'}(B) - \beta < \lambda_k(B)$:¹

$$A_j = \beta(u_n u_n^* + \cdots + u_{j+1} u_{j+1}^*) \quad B_k = \beta(v_n^* + \cdots + v_{k+1} v_{k+1}^*).$$

En remarquant que $(A - A_j)u_i = (\lambda_i - \beta)u_i$ pour $i = j+1, \dots, n$, et est égal à $\lambda_i u_i$ pour $i = 1, \dots, j$, on note que $\lambda_n(A - A_j) = \lambda_j(A)$. De même on a $\lambda_n(B - B_k) = \lambda_k(B)$.

Par ailleurs comme A_j (resp. B_k) est de rang $n-j$ (resp. $n-k$), $A_j + B_k$ est de rang au plus $2n-j-k$ et donc d'après (b)

$$\begin{aligned} \lambda(A - A_j + B - B_k) &= \lambda_n(A + B - (A_j + B_k)) \\ &\geq \lambda_{n-(2n-j-k)}(A + B) \\ &= \lambda_{j+k-n}(A + B) \end{aligned}$$

D'après (a) pour $k = n$, on a aussi

$$\lambda_n(A - A_j + B - B_k) \leq \lambda_n(A - A_j) + \lambda_n(B - B_k)$$

de sorte que

$$\begin{aligned} \lambda_k(A) + \lambda_k(B) &= \lambda(A - A_j) + \lambda_n(B - B_k) \geq \lambda_n(A - A_j + B - B_k) \\ &= \lambda_n((A + B) - (A_j + B_k)) \geq \lambda_{j+k-n}(A + B) \end{aligned}$$

(d) Le résultat découle directement de (c) en considérant $-A$ et $-B$ et en notant que $\lambda_i(-A) = -\lambda_{n-i+1}(A)$.

Remarque: dans (b) pour le cas $r = 1$, les valeurs propres sont entrelacées, i.e.

$$\lambda_k(A + B) \leq \lambda_k(A) \leq \lambda_{k+2}(A + B) \quad \lambda_k(A) \leq \lambda_{k+1}(A + B) \leq \lambda_{k+2}(A)$$

Le même phénomène se produit dans la situation suivante.

Théorème 168. Soit $A \in \mathbb{M}_{n+1}(\mathbb{C})$ hermitienne de valeurs propres $\lambda_1 \leq \cdots \leq \lambda_{n+1}$. Alors les valeurs propres $\lambda'_1 \leq \cdots \leq \lambda'_n$ d'une matrice extraite principale² de A vérifie les inégalités suivantes :

$$\lambda_1 \leq \lambda'_1 \leq \lambda_2 \leq \lambda'_2 \leq \cdots \leq \lambda'_{n-1} \leq \lambda_n \leq \lambda'_n \leq \lambda_{n+1}$$

1. On fera attention que dans [?] théorème 4.3.6 (c), l'auteur oublie que $\lambda_j(A)$ peut être strictement négatif, de sorte que $\lambda_n(A - B)$ serait nul et non égal à $\lambda_{n-r}(A)$.

2. i.e. pour un indice $1 \leq i \leq n+1$, on enlève à A sa i -ème colonne et sa i -ème ligne

Preuve : Par souci de simplicité supposons que $i = n + 1$, on a alors

$$\begin{aligned}\lambda_{k+1} &= \min_{F \in \mathcal{E}_{k+1}} \max_{0 \neq x \in F} \frac{x^* A x}{x^* x} \\ &\geq \min_{F \in \mathcal{E}_{k+1}} \max_{0 \neq x \in F \cap \text{vect}(e_{n+1})^\perp} \frac{x^* A x}{x^* x} \\ &= \min_{F' \in \mathcal{E}'_k} \max_{0 \neq x' \in F'} \frac{(x')^* A' x'}{(x')^* x'} \\ &= \lambda'_k\end{aligned}$$

où on a écrit pour $x \in \mathbb{C}^{n+1}$, $x' \in \mathbb{C}^n$ est le vecteur obtenu en supprimant la dernière coordonnée : on a adopté des notations similaires pour F' et $\mathcal{E}'_k$ désigne les sous-espaces de dimension k de $\mathbb{C}^n$. La majoration $\lambda_k \leq \lambda'_k$ se prouve de la même manière en considérant $\max_{F \in \mathcal{E}_{n-k+1}} \min_{0 \neq x \in F}$.

Remarque: évidemment la démonstration s'adapte simplement au cas où l'on considère une matrice extraite principale A_r de A , où l'on a supprimé r -lignes et les r -colonnes correspondantes. Le résultat est alors :

$$\lambda_k(A) \leq \lambda_k(A_r) \leq \lambda_{k+n-r}(A).$$

Le théorème précédent admet aussi une réciproque :

Théorème 169. Soient pour un entier $n \geq 1$, des nombres réels tels que :

$$\lambda_1 \leq \lambda'_1 \leq \lambda_2 \leq \lambda'_2 \leq \dots \leq \lambda'_{n-1} \leq \lambda_n \leq \lambda'_n \leq \lambda_{n+1}$$

Soit $A' = \text{diag}(\lambda'_1, \dots, \lambda'_n)$, il existe alors un réel a et un vecteur $y \in \mathbb{R}^n$ tels que $\{\lambda_1, \dots, \lambda_{n+1}\}$ est l'ensemble des valeurs propres de la matrice symétrique : $A = \begin{pmatrix} A' & y \\ {}^t y & a \end{pmatrix}$.

Preuve : Le calcul de la trace donne $a = \sum_{i=1}^{n+1} \lambda_i - \sum_{i=1}^n \lambda'_i$. Pour tout t distincts des λ_i , on a l'égalité suivante :

$$\begin{aligned}\begin{pmatrix} I & 0 \\ {}^t((tI - A')^{-1}y) & 1 \end{pmatrix} \begin{pmatrix} tI - A' & -y \\ -{}^t y & t - a \end{pmatrix} \begin{pmatrix} I & (tI - A')^{-1}y \\ 0 & 1 \end{pmatrix} \\ = \begin{pmatrix} tI - A' & 0 \\ 0 & (t - a) - {}^t y(tI - A')^{-1}y \end{pmatrix}\end{aligned}$$

En prenant les déterminants, on obtient alors l'égalité suivante entre les polynômes caractéristiques :

$$\chi_A(t) = \chi_{A'}(t) \left[(t - a) - \sum_{i=1}^n y_i^2 \frac{1}{t - \lambda_i} \right]$$

avec donc $\chi_{A'}(t) = \prod_{i=1}^n (t - \lambda_i)$. Il s'agit alors de montrer l'existence de n réels y_i tels que $\chi_A(\lambda_k) = 0$ pour tout $k = 1, \dots, n + 1$. On considère alors la division euclidienne $\xi_A(t) := \prod_{i=1}^{n+1} (t - \lambda_i) = Q(t)\chi_{A'}(t) + R(t)$ avec donc $Q(t) = t - a$ et pour tout $1 \leq k \leq n$, $R(\lambda'_k) = \xi_A(\lambda'_k)$, ce qui détermine

uniquement le polynôme R de degré inférieur ou à n en utilisant par exemple les polynômes interpolateurs de Lagrange. Supposons pour simplifier que tous les λ'_i sont distincts, on a alors

$$R(t) = \sum_{i=1}^n \xi_A(\lambda'_i) \frac{\chi_{A'}(t)}{\chi'_{A'}(t)(t - \lambda'_i)}$$

de sorte que

$$\frac{\xi_A(t)}{\chi_{A'}(t)} = (t - a) - \sum_{i=1}^n \frac{-\xi_A(\lambda'_i)}{\chi'_{A'}(t)} \frac{1}{t - \lambda'_i}$$

Il suffit alors de montrer que pour tout $i = 1, \dots, n$, $\xi_A(\lambda'_i) \chi'_{A'}(\lambda'_i) \leq 0$ de sorte qu'en posant $y_i^2 = -\frac{\xi_A(\lambda'_i)}{\chi'_{A'}(\lambda'_i)}$, on aura $\chi_A(t) = \xi_A(t)$.

Il s'agit alors d'utiliser l'hypothèse d'entrelacement des valeurs propres : on remarque ainsi que

$$\xi_A(\lambda'_i) = (-1)^{n-i+1} \prod_{j=1}^{n+1} (\lambda'_i - \lambda_j) \quad \chi'_{A'}(\lambda'_i) = (-1)^{n-i} \prod_{\substack{j=1 \\ j \neq i}}^n (\lambda'_i - \lambda'_j)$$

sont effectivement de signes opposés.

Remarque: dans le cas où certains des λ'_i sont égaux, par exemple $\lambda'_1 = \lambda'_2 = \dots = \lambda'_k < \lambda'_{k+1} \leq \dots$, on remarque alors que $(t - \lambda'_1)^{k-1}$ divise $\xi_A(t)$ et on reprend le raisonnement précédent en divisant $\xi_A(t)$ et $\chi_{A'}(t)$ par $(t - \lambda'_1)^{k-1}$.

Corollaire 170. Soit $A \in \mathbb{M}_n(\mathbb{C})$ hermitienne ; pour $1 \leq r \leq n$, U désigne une matrice de $\mathbb{M}_{n,r}(\mathbb{C})$ telle que ses vecteurs colonnes forment une famille orthonormale, i.e. $U^*U = I \in \mathbb{M}_r(\mathbb{C})$. On a alors les propriétés suivantes :

- (i) pour tout $k = 1, 2, \dots, r$, $\lambda_k(A) \leq \lambda_k(U^*AU) \leq \lambda_{k+n-r}(A)$;
- (ii) $\lambda_1(A) + \dots + \lambda_r(A) = \min_{U^*U=I \in \mathbb{M}_r(\mathbb{C})} \text{tr}(U^*AU)$ et

$$\lambda_{n-r+1}(A) + \dots + \lambda_n(A) = \max_{U^*U=I \in \mathbb{M}_r(\mathbb{C})} \text{tr}(U^*AU).$$

Remarque: (i) est connu comme le théorème de séparation de Poincaré et est utilisé en mécanique quantique où on a accès aux calculs de $u_i^* A u_j$ pour une famille orthonormée $(u_i)_{1 \leq i \leq r}$.

Preuve : (i) si $r < n$, on complète les vecteurs colonnes de U en une base orthonormée ; la matrice U' est alors unitaire et $(U')^* A U'$ a les mêmes valeurs propres que A et $U^* A U$ en est une matrice extraite principale, le résultat découle alors de 168, ou plutôt de la remarque qui suit.

(ii) les majorations découlent directement de (i), les égalités sont alors obtenues si les colonnes de U correspondent aux vecteurs propres des r plus petites valeurs propres.

On a le même genre de résultat pour les valeurs singulières.

Théorème 171. Soit $A \in \mathbb{M}_{m,n}(\mathbb{C})$ et $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_q$ ses valeurs singulières pour $q = \min\{m, n\}$. Pour $1 \leq k \leq q$, on a

$$\sigma_k = \min_{F \in \mathcal{E}_k} \max_{0 \neq x \in F} \frac{\|Ax\|_2}{\|x\|_2} = \min_{F \in \mathcal{E}_{n-k+1}} \max_{0 \neq x \in F} \frac{\|Ax\|_2}{\|x\|_2}$$

Remarque: suivant le même schéma que précédemment, on peut obtenir des résultats similaires sur l'enchevêtrement des valeurs singulières. Par exemple pour $A \in \mathbb{M}_{m,n}(\mathbb{C})$ avec $m \geq n$, pour $1 \leq i \leq n$, $A^{(i)}$ désigne la matrice extraite de A en enlevant sa i -ème colonne et ligne. On note σ_i (resp. σ'_i) les valeurs singulières de A (resp. A'), de sorte que l'on a

$$\sigma_1 \geq \sigma'_1 \geq \sigma_2 \geq \sigma'_2 \geq \dots \geq \sigma'_{n-1} \geq \sigma_n \geq 0$$

On en déduit alors cf. [?] 3.1.3, que si A_r désigne une sous-matrice de A obtenue en lui ôtant r lignes et/ou colonnes alors $\sigma_k(A) \geq \sigma_k(A_r) \geq \sigma_{k+r}(A)$.

Corollaire 172. Soit $A \in \mathbb{M}_n(\mathbb{C})$ de valeurs singulières $\sigma_1 \geq \dots \geq \sigma_n$ et soit $H(A) = \frac{1}{2}(A + A^*)$ sa partie hermitienne de valeurs propres $\lambda_1 \geq \dots \geq \lambda_n$. Pour tout $k = 1, \dots, n$, on a $\lambda_k \leq \sigma_k$.

Preuve : Pour $x \in \mathbb{C}^n$ unitaire, on a $x^* H(A) x = \operatorname{Re} (x^* A x) \leq |x^* A x| \leq \|x\|_2 \cdot \|Ax\|_2 = \|Ax\|_2$. Ainsi on a

$$\lambda_k = \min_{F \in \mathcal{E}_k} \max_{\substack{0 \neq x \in F \\ \|x\|_2=1}} x^* H(A) x \leq \min_{F \in \mathcal{E}_k} \max_{\substack{0 \neq x \in F \\ \|x\|_2=1}} \|Ax\|_2 = \sigma_k$$

Corollaire 173. (cf. [?] 3.3.2) Soit $A \in \mathbb{M}_n(\mathbb{C})$ de valeurs singulières $\sigma_1 \geq \dots \geq \sigma_n$ et de valeurs propres $\{\lambda_1, \dots, \lambda_n\}$ ordonnées de sorte que $|\lambda_1| \geq \dots \geq |\lambda_n|$. On a alors

$$|\lambda_1 \cdots \lambda_k| \leq \sigma_1 \cdots \sigma_k \quad \forall k = 1, \dots, n$$

avec égalité pour $k = n$.

Preuve : Soit U unitaire telle que $U^* A U = T$ est triangulaire supérieure et où la diagonale de T est $(\lambda_1, \dots, \lambda_n)$. Soit $U_k \in \mathbb{M}_{n,k}(\mathbb{C})$ la matrice extraite de U constituée par ses k premières colonnes de sorte que $U^* A U = \begin{pmatrix} U_k^* A U_k & * \\ * & * \end{pmatrix}$. La matrice $T_k = U_k^* A U_k$ est donc triangulaire supérieure de diagonale égale à $\lambda_1, \dots, \lambda_k$ de sorte que $|\det T_k| = |\lambda_1 \cdots \lambda_k|$ est égal au produit $\sigma_1(T_k) \cdots \sigma_k(T_k)$ des valeurs singulières de T_k . Le résultat découle alors de la remarque qui suit le théorème 171, i.e. $\sigma_1(T_k) \cdots \sigma_k(T_k) \leq \sigma_1 \cdots \sigma_k$.

4.9 Exercices

Exercice 1. Montrer que deux matrices réelles unitairement semblables sont orthogonalement semblables.

Preuve : On a $A = UBU^{-1}$ et ${}^tA = U{}^tBU^{-1}$. On rappelle que deux matrices réelles qui sont semblables sur $\mathbb{C}$ sont semblables sur $\mathbb{C}$. Il existe donc $P \in GL_n(\mathbb{R})$ telle que $A = PBP^{-1}$ et ${}^tA = P{}^tBP^{-1}$. Soit $P = OS$ la décomposition polaire de P et $A = OSB^{-1}B^{-1}o^{-1}$. Le résultat découle alors du fait que B et S commutent : en effet on a $PBP^{-1} = A = {}^t({}^tA) = {}^tP^{-1}B{}^tP$. Ainsi B commute avec tPP donc aussi avec S qui est un polynôme en tPP .

Remarque : Comme application, on pourra en déduire la réduction des isométries, à partir de la diagonalisation des endomorphismes unitaires.

Exercice 2. Montrer que les seules matrices de Bourdaud (i.e. les valeurs propres se lisent sur la diagonale) réelles symétriques sont les matrices diagonales.

Preuve : Soit λ_1 la plus grande des valeurs propres, on a $\lambda_1 = \sum_{\|x\|=1} (A(x), x)$ et là où le sup est atteint, on est en présence d'un vecteur propre associé à λ_1 : en effet la différentielle s'annule en x_0 sur l'espace tangent à la sphère en x_0 (extrema liés), soit $(A(x_0), y) = 0$ pour tout y tel que $(x_0, y) = 0$ de sorte que $A(x_0)$ est colinéaire à x_0 .

On remarque ensuite qu'il existe k tel que $\lambda_1 = a_{k,k} = (A(e_k), e_k)$ et on raisonne par récurrence sur l'orthogonal à e_k .

On aurait aussi pu utiliser la forme quadratique $A \mapsto \text{tr}({}^tAA) = \sum_{i,j} \lambda_{i,j}^2$ qui est clairement invariante sous l'action par conjugaison du groupe orthogonal de sorte que $\text{tr}({}^tAA) = \sum_i \lambda_{i,i}^2$, d'où le résultat.

Exercice 3. Soit G un sous-groupe de $O(n)$ alors G fini si et seulement si G est d'exposant fini si et seulement si l'ensemble des traces des éléments de G est fini.

Preuve : Si G est fini il est clairement d'exposant fini. Si G est d'exposant fini alors tout élément de G est semblable à une matrice diagonale par blocs avec sur la diagonale I_r , $-I_s$ et des matrices de taille 2, de rotations dont les angles sont alors de la forme $\frac{2k\pi}{n}$, ce qui donne donc un ensemble fini de traces possibles.

Si l'ensemble des traces est fini, considérons $\text{vect}G$ le sous-espace vectoriel de $\mathcal{L}_n(\mathbb{R})$ engendré par les éléments de G dont on fixe une base $g_1, \dots, g_r$. On considère aussi le produit scalaire $(A, B) := \text{tr}({}^tAB)$. On note $a_i(g)$ la composante de g sur g_i , soit $g = \sum_i a_i(g)g_i$. On compose à droite avec g_j^{-1} et on prend la trace. Comme $g_i^{-1} = {}^tg_j$, on a $\text{tr}(g_i g_j^{-1}) = (g_i, g_j)$ et donc

$$\text{tr}(g g_j^{-1}) = \sum_i a_i(g) (g_i, g_j)$$

et comme les g_i sont linéairement indépendants, la matrice M dont les éléments sont les (g_i, g_j) est inversible donc

$$a_i(g) = \sum_j (M^{-1})_{ij} \operatorname{tr}(g g_j^{-1})$$

de sorte que l'on obtient un nombre fini de $a_i(g)$ et donc de g .

Exercice 4. Une rotation de $SO(3)$ sera notée par $r = (k, \theta)$ où k est le vecteur unitaire de l'axe de la rotation et θ son angle.

Soient alors $r = (OA, 2\alpha)$ et $s = (OB, 2\beta)$ deux rotations telles que $\frac{\alpha}{\pi}$ et $\frac{\beta}{\pi}$ soient irrationnels. Montrer que si l'on excepté une infinité dénombrable de valeurs pour la mesure c de l'angle entre les axes OA et OB , le groupe engendré par r et s est dense dans $SO(3)$.

Preuve : Si P_3 est le plan OAB et $P_2 = (OA, -\alpha)(P_3)$ alors r s'écrit comme le produit des réflexions par rapport aux plans P_2 et P_3 . De même s est le produit de P_1 et P_2 où $P_1 = (OB, \beta)(P_3)$.

Afin d'approcher une rotation $(k, 2\theta)$, on approche son axe puis son angle. Pour approcher $\mathbb{R}k$, on approche les plans qu'il détermine avec OA , et OB . D'après le théorème de Jacobi-Kronecker, ils sont respectivement approchés par $P'_2 = (OA, -p\alpha)(P_3)$ et $P'_1 = (OB, q\beta)(P_3)$ si p et q sont des entiers adéquats. Ainsi $\mathbb{R}k' = P'_1 \cap P'_2$ approche $\mathbb{R}k$.

Puisque $r^p = (OA, 2p\alpha) = (P_3)(P'_2)$ et $s^q = (OB, 2q\beta) = (P'_1)(P_3)$, on a $s^q r^p = (P'_1)(P'_2)$ dont la mesure $2\gamma'$ de l'angle est donnée par la formule fondamentale de la trigonométrie sphérique

$$\cos \gamma' = \sin(p\alpha) \sin(q\beta) \cos c - \cos(p\alpha) \cos(q\beta)$$

On cherche $\frac{\gamma'}{\pi}$ irrationnel ; la formule précédente montre que si p et q décrivent les entiers et si $\frac{\gamma'}{\pi}$ décrit les rationnels, $\cos c$ ne prend qu'une infinité dénombrable de valeurs. On choisit alors c pour que $\cos c$ n'appartienne pas à cet ensemble de valeurs. Il en résulte alors que $\frac{\gamma'}{\pi}$ est irrationnel pour tout p, q . Le théorème de Jacobi-Kronecker montre alors que l'on peut choisir n pour que $2n\gamma'$ approche 2θ de sorte que $(s^q r^p)^n$ approche $(k, 2\theta)$.

Exercice 5. (Tauvel p.417) Montrer que u est diagonalisable de spectre réel si et seulement si u est le produit de deux endomorphismes hermitiens, l'un au moins d'entre eux étant défini positif

Preuve : Supposons que u est diagonalisable de spectre réel : soit $(e_1, \dots, e_n)$ une base orthonormée et soit $(x_1, \dots, x_n)$ une base formée de vecteurs propres de u de valeurs propres réelles $\lambda_1, \dots, \lambda_n$. On définit f et g par $f(x_i) = e_i$ et $g(e_i) = \lambda_i e_i$ de sorte que $u = f^{-1} \circ g \circ f$. Soit alors $f = qr$ la décomposition polaire de f avec q unitaire et r hermitienne définie positive.

On obtient ainsi en posant $l = r \circ l \circ r^{-1} = q^* \circ g \circ q$ qui est hermitienne, alors $u = r^{-1} \circ l \circ r = (r^{-1} \circ l \circ r^{-1}) \circ r^2$ avec donc $r^{-1} \circ l \circ r^{-1}$ et r^2 hermitiennes.

Réciproquement si $u = v \circ w$ avec w (resp. v) hermitienne (resp. hermitienne définie positive). Notons $l = v^{1/2}$ qui est hermitienne définie positive, on a $u = l \circ (l \circ w \circ l) \circ l^{-1}$. Comme $l \circ w \circ l$ est hermitienne, elle est diagonalisable à spectre réel et il en est donc de même de u qui lui est semblable.

Exercice 6. (Tauvel p.418) Pour $n \geq 2$ et A hermitienne non nulle de taille n , montrer que A est définie positive ou négative si et seulement si pour toute matrice B hermitienne, AB est diagonalisable.

Preuve : Le sens direct découle de l'exercice précédent. Supposons donc que pour toute matrice hermitienne B , AB est diagonalisable. Soit P unitaire telle que $A = PDP^*$ avec D diagonale réelle. Comme $AB = P(DP^*BP)P^{-1}$, on voit que AB est diagonalisable si et seulement si DP^*BP l'est. Par ailleurs comme P^*BP est hermitienne, on peut supposer $A = D$.

On se ramène aisément en dimension 2 avec $A = \text{diag}(x^2, -y^2)$ avec x, y réels. Soient alors $B = \begin{pmatrix} y^2 & xy \\ xy & x^2 \end{pmatrix}$ et $B' = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$. Si $y \neq 0$, on a $AB \neq 0$ et $(AB)^2 = 0$. Si $y = 0$ alors $AB' \neq 0$ et $(AB')^2 = 0$. Ainsi AB (resp. AB') est nilpotente non nulle et donc non diagonalisable.

Exercice 7. (M-T p.219 ou M p.92) Soient $\lambda_1, \dots, \lambda_n$ les valeurs propres d'une matrice complexe $A = (a_{i,j})$. Montrer que A est normale si et seulement si $\sum_{i,j} |a_{i,j}|^2 = \sum_i |\lambda_i|^2$.

Preuve : Rappelons que $\text{tr}(AA^*) = \sum_{i,j} |a_{i,j}|^2$ est $U(n)$ -invariante de sorte que si A est normale elle est alors unitairement semblable à la matrice diagonale des λ_i d'où le sens direct. Réciproquement toute matrice complexe est unitairement semblable à une matrice triangulaire T de diagonale formée des λ_i . L'égalité implique alors que les termes qui ne sont pas sur la diagonale sont nulle, i.e. que T est diagonale.

Exercice 8. Montrer que A est normale si et seulement si $\text{tr}(AA^*)^2 = \text{tr}A^2A^{*2}$.

Preuve : D'après l'exercice précédent, la matrice hermitienne $H = AA^* - A^*A$ est nulle si et seulement si $\text{tr}H^2 = 0$. Ainsi A est normale si et seulement si la trace de $(AA^*)^2 - A(A^*)^2A - A^*A^2A^* + (A^*A)^2$ est nulle. Or rappelons que $\text{tr}AB = \text{tr}BA$ de sorte que $\text{tr}(A^*A)^2 = \text{tr}(AA^*)^2$ et $\text{tr}A^*A^2A^* = \text{tr}A(A^*)^2A = \text{tr}A^2A^{*2}$, d'où le résultat.

Exercice 9. Montrer que deux matrices sont unitairement équivalentes si et seulement si elles ont les mêmes valeurs singulières

Preuve : Le sens direct découle de $PAP^* = B$, $PA^*P^* = B^*$ soit $PAA^*P^* = BB^*$. Réciproquement, la décomposition polaire donne $A = H_A U_A$, $B =$

$H_B U_B$ avec H_A, H_B hermitiennes définies positives, et U_A, U_B unitaires. On a donc $AA^* = H_A H_A^*$ et $BB^* = H_B H_B^*$. Les matrices H_A, H_B sont diagonalisables de sorte que si AA^* et BB^* ont même valeurs propres alors les valeurs propres de H_A sont égales à celles de H_B au signe près, soit $H_A = P_A D P_A^*$ et $H_B = P_B D D' P_B^*$ avec $D' = \text{diag}(\epsilon_1, \dots, \epsilon_n)$ où les ϵ_i sont égaux à ± 1 de sorte que D' est unitaire. On a donc $D = P_B^* U_B^* B P_B D'$ et $A = U_A P_A P_B^* U_B^* B P_B D' P_A^*$ d'où le résultat.

Exercice 10. Donner le centre Z de $O(q)$ (resp. Z^+ de $O^+(q)$) et montrer que $O(q)$ est un produit semi-direct de $O^+(q)$ par $\mathbb{Z}/2\mathbb{Z}$; à quelle condition ce produit semi-direct peut-il être pris direct ?

Preuve : Il est clair que $\{Id, -Id\} \subset Z$; réciproquement soit $z \in Z$ et τ_D une réflexion de droite D . On a $z \tau_D z^{-1} = \tau_D = \tau_{z(D)}$ de sorte que z laisse stable toutes les droites de l'espace; c'est donc une homothétie (résultat classique) et donc $z = \pm Id$.

En ce qui concerne Z^+ remarquons que $-Id$ appartient à $O^+(q)$ si et seulement si n est pair. Pour $n \geq 3$ soit τ_P un renversement de plan P ; on a $z \tau_P z^{-1} = \tau_P = \tau_{z(P)}$ de sorte que z laisse stable tous les plans de l'espace. Toute droite étant l'intersection de deux plans, on en déduit de même que z laisse stable toutes les droites de l'espace, soit $Z^+ = \{Id\}$ pour n impair et sinon $Z^+ = Z$ pour n pair. Pour $n = 2$, il est bien connu que O^+ est commutatif.

Il est clair que la suite exacte $1 \rightarrow O^+(q) \rightarrow O(q) \rightarrow \mathbb{Z}/2\mathbb{Z} \rightarrow 0$ est scindée, un relèvement étant donné par exemple par une réflexion quelconque. Pour obtenir un produit direct, il faut trouver un élément d'ordre 2 qui n'est pas dans O^+ et qui commute à tous les éléments de O^+ ; la seule possibilité est alors $-Id$ en dimension impaire.

□

Exercice 11. Soit $u \in O(q)$ et $F_u = \{x \in E / u(x) = x\}$ et on note $p_u = n - \dim F_u$. Montrer par récurrence sur p_u , que u est le produit d'au plus p_u réflexions. Montrer ensuite que u est le produit d'au moins p_u réflexions.

Preuve : On raisonne par récurrence sur p_u , le cas $p_u = 0$ correspondant à $u = Id$. Supposons donc $p_u > 0$ et soit $x \in F_u^\perp$ non nul et soit $y = u(x) \neq x$ car $x \notin F_u$; on a $y \in F_u^\perp$ car F_u étant stable par u , $F_u^\perp$ l'est aussi. De plus comme x et y ont même norme, on en déduit que $(x - y, x + y) = 0$ (triangle isocèle). On considère alors la réflexion τ définie par $x - y$ de sorte que $\tau(x - y) = y - x$ et $\tau(x + y) = x + y$ soit donc $\tau(y) = x$ avec $\tau|_{F_u} = Id$. Ainsi on a $F_u \subset F_{\tau \circ u}$ ce dernier contenant x de sorte que $p_{\tau \circ u} < p_u$ et on conclut par récurrence.

En outre si u est le produit de r réflexions alors F_u est clairement de dimension supérieure ou égale à $n - r$ (l'intersection de r hyperplans) soit donc $p_u \leq r$.

□

Exercice 12. Montrer que pour $n \geq 3$, tout élément de $O^+(q)$ est produit d'au plus n renversements.

Preuve : Le cas $n = 3$ est évident en remarquant que si τ est une réflexion, alors $-\tau$ est un renversement de sorte que le produit de deux réflexions (et donc tout produit d'un nombre pair) est un produit de deux renversements $\tau_1 \circ \tau_2 = (-\tau_1) \circ (-\tau_2)$.

Pour $n \geq 3$, soient τ_1 et τ_2 des réflexions par rapport aux hyperplans H_1 et H_2 et $u = \tau_1 \circ \tau_2$. Soit alors $V \subset H_1 \cap H_2$ un sous-espace de dimension $n-3$: $u|_V = Id$ et $V^\perp$ est stable sous u . D'après le cas $n = 3$, on a $u|_{V^\perp} = \sigma_1 \circ \sigma_2$ où σ_1, σ_2 sont des renversements de $V^\perp$. On obtient le résultat en prolongeant les σ_i par l'identité sur V . □

Exercice 13. Soient u_1 et u_2 deux symétries orthogonales de même nature (i.e. tels que $\dim \text{Ker}(u_1 - Id) = \dim \text{Ker}(u_2 - Id)$). Montrer que u_1 et u_2 sont conjuguées par $O^+(q)$. En déduire alors que $D(O(q)) = D(O^+(q)) = O^+(q)$.

Preuve : On décompose l'espace $E = E_1 \oplus E_1^\perp = E_2 \oplus E_2^\perp$ où $E_i = \text{Ker}(u_i - Id)$. On choisit alors des bases orthonormées (e_i^1) et (e_i^2) de E adaptées à ces décompositions. Soit alors u tel que $u(e_i^1) = e_i^2$; u est une isométrie et quitte à changer e_1 en $-e_1$, on peut supposer que u est positive. On vérifie alors immédiatement que $u \circ u_1 \circ u^{-1} = u_2$.

L'inclusion $D(O(q)) \subset O^+(q)$ est évidente ; réciproquement soient τ_1 et τ_2 deux réflexions et soit u tel que $u \circ \tau_1 \circ u^{-1} = \tau_2$ de sorte que $\tau_1 \circ \tau_2 = [\tau_1, u]$. Comme tout élément de $O^+(q)$ est le produit d'un nombre pair de réflexions, on obtient bien l'inclusion réciproque.

De même pour montrer que $O^+(q) \subset D(O^+(q))$ pour $n \geq 3$, il suffit de montrer que tout renversement est un commutateur. Soit V un sous-espace de dimension 3 et (e_1, e_2, e_3) une base orthonormée. Soient $\sigma_1, \sigma_2, \sigma_3$ les renversements définis par $(\sigma_i)|_{V^\perp} = Id$ et $\sigma_i(e_i) = e_i$ et donc $\sigma_i(e_j) = -e_j$ pour $i \neq j$. On a alors $\sigma_3 = \sigma_1 \circ \sigma_2$. En outre il existe $u \in O^+(q)$ tel que $\sigma_2 = u \circ \sigma_1 \circ u^{-1}$ et donc $\sigma_3 = [\sigma_1, u]$. □

Exercice 14. Montrer que pour tout $u \in O(q)$, il existe une décomposition orthogonale

$$E = \text{Ker}(u - Id) \oplus \text{Ker}(u + Id) \oplus P_1 \oplus \cdots \oplus P_r$$

où les P_i sont des plans stables par u , tels que la restriction de u y soit une rotation.

Preuve : On procède par récurrence sur la dimension, les cas $n = 1$ et $n = 2$ étant bien connus. Si u admet une valeur propre réelle (forcément ± 1), c'est terminé (en particulier si n est impair). Sinon soit $\lambda \in \mathbb{C}$ une

valeur propre du complexifié de $u_{\mathbb{C}}$, de sorte que $\bar{\lambda}$ est aussi valeur propre. Soit alors $x \in E \otimes_{\mathbb{R}} \mathbb{C}$ un vecteur propre du complexifié relativement à λ et soit $\bar{x}$ son conjugué qui est alors propre pour $\bar{\lambda}$ relativement à $u_{\mathbb{C}}$. Le plan complexe $P = \mathbb{C}x + \mathbb{C}\bar{x}$ est alors invariant par $u_{\mathbb{C}}$. On remarque alors que les vecteurs $\frac{x+\bar{x}}{2}$ et $\frac{x-\bar{x}}{2i}$ sont réels et forment une base de P de sorte que le plan réel qu'ils engendrent est stable sous u . □

Exercice 15. - On veut prouver la simplicité de $O^+(3, \mathbb{R})$. Soit donc N un sous-groupe distingué non réduit à l'identité ; expliquer pourquoi il suffit de montrer que N contient un renversement.

- Soit alors $u \in N$, une rotation d'axe D et soit P le plan orthogonal à D à l'origine de sorte que la restriction de u à P est une rotation d'angle θ que l'on suppose $0 < \theta < \pi$. Soient alors x et $y = u(x)$ des points de la sphère unité de E ; on note d la distance entre x et y . Montrer que pour tout $0 \leq d' \leq d$, il existe x_1, x_2 des points de la sphère unité à distance d' l'un de l'autre et tels que $x_2 = u(x_1)$.

- Dédurre de ce qui précède qu'étant donnés y_1, y_2 des points de la sphère unité distant de d' avec $0 \leq d' \leq d$, il existe $u' \in N$ tels que $u'(y_1) = y_2$. En considérant la rotation d'axe z et d'angle π/m pour m assez grand, construire un retournement de N et conclure.

Preuve : - Comme les renversements engendrent $O^+(3, \mathbb{R})$ et sont conjugués sous $O^+(3, \mathbb{R})$, il suffit de montrer que N en contient un.

- Un calcul classique donne $d^2 = 2(1 - \cos \theta)$. Soit a un des points de $D \cap S^2$; le résultat découle de l'observation que u envoie le méridien contenant a et x , sur celui contenant a et y et que lorsque x_1 varie de x à a , la distance $\|x_1 - u(x_1)\|$ varie continûment de d à 0. De façon précise, on considère $x + \lambda a$ de norme au carré égale à $1 + \lambda^2$ de sorte que $x_1 = \frac{x + \lambda a}{\sqrt{1 + \lambda^2}} \in S^2$. On a alors $\|u(x_1) - x_1\| = \frac{d}{\sqrt{1 + \lambda^2}}$ de sorte qu'il suffit de prendre $\lambda = \frac{\sqrt{d^2 - m^2}}{m}$.

- Soit x_3 (resp. y_3) un vecteur de norme 1 orthogonal au plan engendré par x_1 et x_2 (resp. y_1 et y_2) et soit u tel que $s(x_i) = y_i$ pour $i = 1, 2, 3$. Il est clair que s conserve le produit scalaire et donc $u \in O(3, \mathbb{R})$; quitte à changer y_3 en $-y_3$, on peut supposer que s est positive. On pose $u' := s \circ u \circ s^{-1} \in N$ et $u'(y_1) = y_2$. Soit alors r_n la rotation d'angle π/n et d'axe a . Comme $\mathbb{R}$ est archimédien, le rapport π/n tend vers 0 quand n tend vers $+\infty$ et donc pour n assez grand $\|x - r_n(x)\| \leq d$. On pose alors $x_0 = x$ et $x_{i+1} = r_n(x_i)$ avec donc $x_n = -x$. Comme on a $\|x_{i+1} - x_i\| \leq d$ il existe alors $u_i \in N$ tel que $u(x_i) = x_{i+1}$ de sorte que $v = u_n \circ \dots \circ u_1 \in N$ et $v(x) = -x$ et v est donc un renversement, d'où le résultat. □

Exercice 16. On note H le corps des quaternions et soit G ceux de norme 1 : $G = \{a + bi + cj + dk / a^2 + b^2 + c^2 + d^2 = 1\}$. On considère alors l'action

de G sur H par automorphismes intérieurs. En restreignant cette action à l'ensemble P des quaternions purs, montrer que l'on obtient alors un isomorphisme $G/\{\pm 1\} \simeq O(3, \mathbb{R})^+$. La suite exacte associée est-elle scindée ?

Preuve : On a $P \simeq \mathbb{R}^3$ et on vérifié aisément que l'action de conjugaison de G est $\mathbb{R}$ -linéaire et conserve la norme de sorte qu'elle définit un morphisme de groupes $G \longrightarrow O(3, \mathbb{R})$. On note en outre que $G \simeq S^3$ est connexe et que le morphisme précédent est continue de sorte que l'image de $G \rightarrow O(3, \mathbb{R}) \rightarrow \{\pm 1\}$ est connexe et donc égale à $\{1\}$. On obtient donc bien un morphisme de groupe $\phi : G \longrightarrow O^+(3, \mathbb{R})$. Montrons la surjectivité : soit $p \in P \cap G$, on a $\phi_p(p) = p$ ce qui prouve que ϕ_p fixe p (et est non triviale), c'est donc une rotation d'axe p . En outre on a $p^2 = -1$ soit ϕ_p d'ordre 2 ; c'est donc un renversement. On obtient donc tous les renversements, or ceux-ci engendrent $O^+(3, \mathbb{R})$, d'où la surjectivité. Pour le noyau, on a $\phi_g(p) = p$ pour tout $p \in P$ si et seulement si g commute à tous les éléments de P et donc à tous les éléments de H , soit donc $g \in \mathbb{R} \cap G = \{\pm 1\}$.

Si la suite exacte

$$1 \rightarrow \{\pm 1\} \longrightarrow G \xrightarrow{\phi} O^+(3, \mathbb{R}) \rightarrow 1$$

était scindée, on aurait un sous-groupe H de G tel que $\phi|_H$ soit un isomorphisme de H sur $O^+(3, \mathbb{R})$. Mais alors pour $g \in G$, on aurait g ou $-g$ qui appartiendrait à H . En prenant $o \in P \cap G$, on a $p^2 = (-p)^2 = -1$ soit donc $-1 \in H$, contradiction. □

Exercice 17. On considère l'action de $G \times G$ sur H définie par $(q_1, q_2).q := q_1 q \bar{q}_2$. Montrer que l'on définit ainsi un isomorphisme $G \times G / \{(1, 1), (-1, -1)\} \simeq O(4, \mathbb{R})^+$ et en déduire que $PO(4, \mathbb{R})^+ \simeq O(3, \mathbb{R})^+ \times O(3, \mathbb{R})^+$.

Preuve : L'application ϕ_{q_1, q_2} est clairement $\mathbb{R}$ -linéaire et conserve la norme. Par continuité, on conclut comme précédemment que son image est contenue dans les isométries positives soit donc

$$\phi : G \times G \longrightarrow O^+(4, \mathbb{R})$$

Soit $(q_1, q_2) \in \text{Ker } \phi$, i.e. $q_1 q \bar{q}_2 = q$ pour tout $q \in H$. Pour $q = 1$, on trouve $q_1 = q_2$ de sorte qu'ensuite q_1 est central et donc $\text{Ker } \phi = \{(1, 1), (-1, -1)\}$.

Pour la surjectivité, soit $u \in O^+(4, \mathbb{R})$, si on a $u(1) = 1$, comme $P = 1^\perp$, on a $u(P) = P$ avec $u|_P \in O^+(3, \mathbb{R})$ et d'après ce qui il existe $q \in G$ tel que $\phi_{q, q} = u$. Si on a $u(1) = g$, on a alors $\phi_{\bar{g}, 1} \circ u(1) = 1$ et on conclut grâce au cas précédent. Finalement on obtient donc

$$G \times G / \{(1, 1), (-1, -1)\} \simeq O(4, \mathbb{R})^+$$

En passant au groupe projectif, on cherche les couples (q_1, q_2) tels que $\phi_{q_1, q_2} = -Id$, i.e. $q_1 q \bar{q}_2 = -q$ pour tout $q \in H$. En faisant $q = 1$, on obtient $q_1 = -q_2$, puis on voit que q_1 est central soit alors

$$G \times G/V \simeq PO(4, \mathbb{R})^+$$

où $V = \{(1, 1), (1, -1), (-1, 1), (-1, -1)\}$. En outre la projection canonique $G \rightarrow G/\{\pm 1\}$ induit un isomorphisme

$$(G \times G)/V \simeq G/\{\pm 1\} \times G/\{\pm 1\}$$

et donc d'après ce qui précède

$$PO(4, \mathbb{R})^+ \simeq O(3, \mathbb{R})^+ \times O(3, \mathbb{R})^+.$$

□

Exercice 18. (Tauvel p.417) Montrer que u est diagonalisable de spectre réel si et seulement si u est le produit de deux endomorphismes hermitiens, l'un au moins d'entre eux étant défini positif

Preuve : Supposons que u est diagonalisable de spectre réel : soit $(e_1, \dots, e_n)$ une base orthonormée et soit $(x_1, \dots, x_n)$ une base formée de vecteurs propres de u de valeurs propres réelles $\lambda_1, \dots, \lambda_n$. On définit f et g par $f(x_i) = e_i$ et $g(e_i) = \lambda_i e_i$ de sorte que $u = f^{-1} \circ g \circ f$. Soit alors $f = qr$ la décomposition polaire de f avec q unitaire et r hermitienne définie positive. On obtient ainsi en posant $l = r \circ l \circ r^{-1} = q^* \circ g \circ q$ qui est hermitienne, alors $u = r^{-1} \circ l \circ r = (r^{-1} \circ l \circ r^{-1}) \circ r^2$ avec donc $r^{-1} \circ l \circ r^{-1}$ et r^2 hermitiennes.

Réciproquement si $u = v \circ w$ avec w (resp. v) hermitienne (resp. hermitienne définie positive). Notons $l = v^{1/2}$ qui est hermitienne définie positive, on a $u = l \circ (l \circ w \circ l) \circ l^{-1}$. Comme $l \circ w \circ l$ est hermitienne, elle est diagonalisable à spectre réel et il en est donc de même de u qui lui est semblable.

Exercice 19. (Tauvel p.418) Pour $n \geq 2$ et A hermitienne non nulle de taille n , montrer que A est définie positive ou négative si et seulement si pour toute matrice B hermitienne, AB est diagonalisable.

Preuve : Le sens direct découle de l'exercice précédent. Supposons donc que pour toute matrice hermitienne B , AB est diagonalisable. Soit P unitaire telle que $A = PDP^*$ avec D diagonale réelle. Comme $AB = P(DP^*BP)P^{-1}$, on voit que AB est diagonalisable si et seulement si DP^*BP l'est. Par ailleurs comme P^*BP est hermitienne, on peut supposer $A = D$.

On se ramène aisément en dimension 2 avec $A = \text{diag}(x^2, -y^2)$ avec x, y réels. Soient alors $B = \begin{pmatrix} y^2 & xy \\ xy & x^2 \end{pmatrix}$ et $B' = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$. Si $y \neq 0$, on a $AB \neq 0$ et $(AB)^2 = 0$. Si $y = 0$ alors $AB' \neq 0$ et $(AB')^2 = 0$. Ainsi AB (resp. AB') est nilpotente non nulle et donc non diagonalisable.

Exercice 20. A est normale si et seulement si A^* est un polynôme en A

Preuve : La réciproque étant évidente, montrons le sens direct. Soit $(e_1, \dots, e_n)$ une base orthonormée de vecteurs propres pour $A : Ae_i = \lambda_i e_i$. De même on a $A^*e_i = \overline{\lambda_i}e_i$. On construit alors un polynôme interpolateur de Lagrange P tel que $P(\lambda_i) = \overline{\lambda_i}$ de sorte que $A^* = P(A)$.

Exercice 21. *Si A est normale, montrer que l'orbite de A sous l'action du groupe unitaire est homéomorphe à $U(n)/U(k_1) \times \dots \times U(k_r)$.*

Preuve : Il suffit de remarquer que le centralisateur de A est isomorphe à $U(k_1) \times \dots \times U(k_r)$ où les k_i sont les dimensions des sous-espaces propres.

Exercice 22. *Montrer que deux matrices sont unitairement équivalentes si et seulement si elles ont les mêmes valeurs singulières*

Preuve : Le sens direct découle de $PAP^* = B$, $PA^*P^* = B^*$ soit $PAA^*P^* = BB^*$. Réciproquement, la décomposition polaire donne $A = H_A U_A$, $B = H_B U_B$ avec H_A, H_B hermitiennes définies positives, et U_A, U_B unitaires. On a donc $AA^* = H_A H_A^*$ et $BB^* = H_B H_B^*$. Les matrices H_A, H_B sont diagonalisables de sorte que si AA^* et BB^* ont même valeurs propres alors les valeurs propres de H_A sont égales à celles de H_B au signe près, soit $H_A = P_A D P_A^*$ et $H_B = P_B D D' P_B^*$ avec $D' = \text{diag}(\epsilon_1, \dots, \epsilon_n)$ où les ϵ_i sont égaux à ± 1 de sorte que D' est unitaire. On a donc $D = P_B^* U_B^* B P_B D'$ et $A = U_A P_A P_B^* U_B^* B P_B D' P_A^*$ d'où le résultat.

Exercice 23. *(M-T p.187) Les matrices de Householder sont exactement les matrices de $U(n)$ qui sont hermitiennes de signature $(n-1, 1)$.*

Preuve : Rappelons qu'une matrice de Householder associée au vecteur colonne $v \in \mathbb{C}^n - \{0\}$ est $H(v) = I - 2 \frac{vv^*}{v^*v}$. La matrice vv^* est hermitienne de rang 1 ; ses valeurs propres sont 0 à l'ordre $n-1$ et $\text{tr}(vv^*) = v^*v$. Ainsi les valeurs propres de $H(v)$ qui est clairement hermitienne et unitaire sont 1 à l'ordre $n-1$ et -1 à l'ordre 1.

Réciproquement si H est une matrice hermitienne unitaire, ses valeurs propres sont réelles de module 1 ; vu l'hypothèse sur la signature et le fait que H est diagonalisable dans une base orthonormée $H = U D U^*$ avec $D = \text{diag}(-1, 1, \dots, 1)$, c'est à dire $H = U H(e_1) U^{-1} = H(U(e_1))$.

5 Liste des projets possibles

1. Dimension infinie avec notamment espaces de Hilbert, Banach et Fréchet
2. Utilisations du conditionnement des matrices
3. Algorithme de Fadæev pour le calcul du polynôme minimal : cas des autres invariants de similitude.
4. Théorème de Perron-Frobenius, chaînes de Markov et le pagerank de google
5. Exponentielle de matrices
6. $PSL_2(\mathbb{Z})$: générateurs et relations ; groupes libres et paradoxe de Banach Tarski
7. Théorie des caractères
8. Sous-groupes compacts de $GL_n(\mathbb{R})$
9. Simplicité de SL_n et SO_3 .
10. Endoscopie
11. Matrices bistochastiques
12. Matrices de Hadamard