Chapitre 1

Dualité, orthogonalité, espaces vectoriels quotients

Ce premier chapitre comporte deux volets bien distincts. Le premier volet est surtout consacré à des rappels concernant les notions de dual et de bidual pour un espace vectoriel, de base duale et antéduale, d'application transposée, de crochet de dualité et d'orthogonalité (au sens de la dualité, à ne pas confondre avec l'orthogonalité au sens d'un produit scalaire donné). Ces notions apparaîtrons tout au long du semestre dans les différents chapitres du cours et dans les exercices des fiches de TD. Le second volet, plus original, est consacré à la notion d'espace vectoriel quotient. Il s'agit là d'un outil conceptuel important qui permet, entre autres choses, de produire des isomorphismes canoniques entre certains espaces vectoriels.

Dans tout ce chapitre, $\mathbb{K}$ désigne un corps quelconque, par exemple $\mathbb{Q}, \mathbb{R}, \mathbb{C}$, et E désigne un $\mathbb{K}$-espace vectoriel de dimension quelconque (sauf mention explicite du contraire).

1.1 Dualité et orthogonalité

1.1.A Généralités

Définition 1.1. On note

$$E^* := \mathcal{L}(E, \mathbb{K})$$

le *dual* de E. Les éléments de E^* sont appelés *formes linéaires* sur E. On appelle *hyperplan* de E le noyau d'une forme linéaire non-nulle sur E.

Remarque 1.2.

(i) Si $\dim(E) < \infty$, alors $\dim(E^*) = \dim(E)\dim(\mathbb{K}) = \dim(E)$.

(ii) Deux formes linéaires non-nulles sur E ont le même noyau si et seulement si elles sont proportionnelles (cf. exercice 6 de la fiche 1 du TD).

Exemple 1.3.

- Si $E = \mathbb{K}^3$, alors $E \to \mathbb{K},\ (x, y, z) \mapsto x + y + z$ est une forme linéaire non-nulle sur E. L'hyperplan associé est le sous-espace vectoriel de E défini par l'équation $x + y + z = 0$.
- Si E est un $\mathbb{R}$-espace vectoriel et $f \in \mathcal{C}^1(E, \mathbb{R})$, alors on a $df_x \in E^*$ pour tout $x \in E$.

Une fois que l'on dispose de la notion d'espace dual, il est naturel de se demander si une application linéaire entre espaces vectoriels induit une application linéaires entre les espaces duaux. La réponse est oui mais à condition d'échanger la source et le but après passage au dual :

Définition 1.4. À toute application linéaire $\varphi: E \to F$ entre espaces vectoriels, on associe sa *transposée* définie par

$$^t\varphi: F^* \to E^*,\ \epsilon \mapsto \epsilon \circ \varphi.$$

Proposition 1.5 (Propriétés de la transposée)**.** Soient $\varphi: E \to F$ et $\psi: F \to G$ deux applications linéaires entre espaces vectoriels. On a alors les propriétés suivantes :
(i) $^t(\psi \circ \varphi) = {}^t\varphi \circ {}^t\psi$
(ii) $^t\mathrm{Id}_E = \mathrm{Id}_{E^*}$
(iii) L'application $\mathcal{L}(E,F) \to \mathcal{L}(F^*,E^*),\ \varphi \mapsto {}^t\varphi$ est linéaire.
(iv) Si φ est bijective, alors $({}^t\varphi)^{-1} = {}^t(\varphi^{-1})$.
(v) Si φ est surjective, alors $^t\varphi$ est injective.
(vi) Si φ est injective, alors $^t\varphi$ est surjective.

Démonstration. (i) : Pour tout $\epsilon \in G^*$, on a

$$^t(\psi \circ \varphi)(\epsilon) = \epsilon \circ \psi \circ \varphi = {}^t\varphi(\epsilon \circ \psi) = ({}^t\varphi \circ {}^t\psi)(\epsilon),$$

et donc $^t(\psi \circ \varphi)(\epsilon) = ({}^t\varphi \circ {}^t\psi)(\epsilon)$. Puisque cette égalité est vraie pour tout $\epsilon \in G^*$, on en déduit que $^t(\psi \circ \varphi) = {}^t\varphi \circ {}^t\psi$. (ii) : Pour tout $\epsilon \in E^*$, on a $({}^t\mathrm{Id}_E)(\epsilon) = \epsilon \circ \mathrm{Id}_E = \epsilon = \mathrm{Id}_{E^*}(\epsilon)$, d'où $^t\mathrm{Id}_E = \mathrm{Id}_{E^*}$.
(iii) : Soient $\varphi, \varphi' \in \mathcal{L}(E,F)$ et $\lambda \in \mathbb{K}$, alors pour tout $\epsilon \in E^*$, on a

$$^t(\varphi + \varphi')(\epsilon) = \epsilon \circ (\varphi + \varphi') = \epsilon \circ \varphi + \epsilon \circ \varphi' = ({}^t\varphi + {}^t\varphi')(\epsilon)\ ;\ \text{et}$$
$$^t(\lambda\varphi)(\epsilon) = \epsilon \circ (\lambda\varphi) = \lambda(\epsilon \circ \varphi) = \lambda{}^t\varphi(\epsilon).$$

Et donc l'application $\mathcal{L}(E,F) \to \mathcal{L}(F^*,E^*),\ \varphi \mapsto {}^t\varphi$ est linéaire.
(iv) : Pour tout $\epsilon \in E^*$, on a

$$({}^t\varphi \circ {}^t(\varphi^{-1}))(\epsilon) = {}^t\varphi(\epsilon \circ \varphi^{-1}) = \epsilon \circ \varphi^{-1} \circ \varphi = \epsilon.$$

Donc $^t\varphi \circ {}^t(\varphi^{-1}) = \mathrm{Id}_{E^*}$, d'où $^t(\varphi^{-1}) = ({}^t\varphi)^{-1}$.
(v) : On suppose que φ est surjective, et on considère $^t\varphi: F^* \to E^*$. Soit $\epsilon \in \mathrm{Ker}({}^t\varphi)$. Alors $^t\varphi(\epsilon) = \epsilon \circ \varphi = 0$, c'est-à-dire

$$\forall x \in E,\ (\epsilon \circ \varphi)(x) = \epsilon(\varphi(x)) = 0.$$

Mais alors la surjectivité de φ implique $\epsilon = 0$, et donc $^t\varphi$ est injective.
(vi) : En exercice ! (Un peu plus difficile que le point (v).) □

Définition 1.6. Le *bidual* de E est

$$E^{**} = (E^*)^* = \mathcal{L}(E^*, \mathbb{K}).$$

Théorème 1.7. L'application

$$\iota: E \to E^{**} = \mathcal{L}(E^*, \mathbb{K}),\ x \mapsto (\epsilon \mapsto \epsilon(x))$$

est une application linéaire injective *canonique* (c'est-à-dire dont la définition ne dépend pas du choix d'une base de E) de E dans E^{**}. De plus, ι est un isomorphisme d'espaces vectoriels si et seulement $\dim(E) < \infty$.

Démonstration. • Linéarité de ι : Pour tous $x, x' \in E$, pour tout $\lambda \in \mathbb{K}$, on a

$$\begin{aligned}\iota(x + \lambda x') &= (\epsilon \mapsto \epsilon(x + \lambda x')) \\ &= (\epsilon \mapsto \epsilon(x) + \lambda \epsilon(x')) \\ &= (\epsilon \mapsto \epsilon(x)) + \lambda(\epsilon \mapsto \epsilon(x')) \\ &= \iota(x) + \lambda \iota(x').\end{aligned}$$

Il s'ensuit que l'application ι est linéaire.

- Injectivité de ι : Soit $x \in E \setminus \{0\}$. On veut montrer que $\iota(x) \neq 0$. Pour ce faire, on procède par l'absurde et on suppose donc que $\iota(x) = 0$, c'est-à-dire, que pour tout $\epsilon \in E^*$, $\epsilon(x) = 0$.

  On considère la droite vectoriel $\langle x \rangle$ dans E, et on fixe un sous-espace vectoriel supplémentaire F de $\langle x \rangle$ dans E (un tel supplémentaire existe toujours, mais en dimension infinie montrer son existence nécessite l'usage du lemme de Zorn qui est équivalent à l'axiome du choix). On a donc $E = \langle x \rangle \oplus F$. Ainsi, tout élément $v \in E$ s'écrit $v = \lambda x + f$ de manière unique, avec $\lambda \in \mathbb{K}$ et $f \in F$. On définit alors une forme linéaire non-nulle $\epsilon_0 \in E^*$ en posant $\epsilon_0(v) = \lambda$. Cette forme linéaire vérifie $\epsilon_0(x) = 1 \neq 0$, ce qui contredit le fait que pour tout $\epsilon \in E^*$, $\epsilon(x) = 0$. Il s'ensuit que ι est injective.
- Si $\dim(E) < \infty$, alors $\dim(E^{**}) = \dim(E^*) = \dim(E)$, et donc comme ι est injective on en déduit que ι est un isomorphisme d'espaces vectoriels.
- Si $\dim(E) = \infty$, alors on peut montrer que ι n'est pas surjective, mais la démonstration de cette affirmation est plus difficile que ce qui précède. (L'étudiant.e motivé.e qui souhaite lire une démonstration de ce résultat peut consulter les pages 25–28 du livre "Finite-Dimensional Vector Spaces (2nde édition, 1974)" de Paul Richard HALMOS.)

□

Remarque 1.8. On peut montrer (mais c'est hors du cadre de ce cours) que la seule application canonique de E dans E^* est l'application nulle. En particulier, E n'est pas canoniquement isomorphe à E^* (sauf dans le cas trivial $E = \{0\}$).

1.1.B Cas de la dimension finie

Dans toute cette section on suppose que $\dim(E) = n < \infty$. L'avantage de travailler dans ce cadre est qu'on dispose d'un outil supplémentaire : les matrices.

Définition-Proposition 1.9. Soit $\mathcal{B} = \{e_1, e_2, \dots, e_n\}$ une base de E. Pour tout $i \in \{1, \dots, n\}$, on définit la forme linéaire $e_i^* \in E^*$ par

$$e_i^*(e_j) = \delta_{ij} := \begin{cases} 1 & \text{si } i = j \ ; \\ 0 & \text{sinon.} \end{cases}$$

Alors la famille $\mathcal{B}^* = \{e_1^*, \dots, e_n^*\}$ est une base de E^* appelée *base duale* de $\mathcal{B}$. En particulier, on retrouve que $\dim(E^*) = \dim(E)$.

Démonstration. Vérifions que la famille $\mathcal{B}^*$ est libre :
Soient $\lambda_1, \dots, \lambda_n \in \mathbb{K}$ tels que $\sum_{i=1}^n \lambda_i e_i^* = 0$. Alors, en évaluant cette égalité en e_j avec $j \in \{1, \dots, n\}$, on obtient $\lambda_j = 0$. Donc $\mathcal{B}^*$ est libre.

Vérifions que la famille $\mathcal{B}^*$ est génératrice :
On fixe $\epsilon \in E^*$. Pour tout $x \in E$, on a la décomposition $x = \sum_{i=1}^n \lambda_i e_i$. Donc

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

Ainsi $\epsilon = \sum_{i=1}^n \epsilon(e_i)e_i^*$, et donc $\epsilon \in \mathrm{Vect}(\mathcal{B}^*)$, autrement dit la famille $\mathcal{B}^*$ est génératrice.

On a donc montré que la famille $\mathcal{B}^*$ est à la fois libre et génératrice, c'est donc une base de l'espace dual E^*. □

Remarque 1.10. Soient $\mathcal{B} = \{e_1, \ldots, e_n\}$, une base de E, et $\mathcal{B}^*$ la base duale de $\mathcal{B}$. Alors on peut vérifier que l'on a les égalités suivantes :

$$\forall x \in E,\ x = \sum_{i=1}^n e_i^*(x)e_i \quad \text{et} \quad \forall \epsilon \in E^*,\ \epsilon = \sum_{i=1}^n \epsilon(e_i)e_i^*.$$

Remarque 1.11. La notation $\{e_1^*, \ldots, e_n^*\}$ peut laisser penser que chaque e_i^* ne dépend que de e_i, ce qui est faux : chaque e_i^* dépend de tous les éléments de $\mathcal{B} = \{e_1, \ldots, e_n\}$. Considérons par exemple le cas suivant : on suppose que $\dim(E) = 2$, soit $\mathcal{B} = \{e_1, e_2\}$ une base de E, et soit $\mathcal{B}' = \{f_1, f_2\}$, avec $f_1 = e_1$ et $f_2 = e_1 + e_2$, une autre base de E. Alors $f_1 = e_1$ mais $f_1^* = e_1^* - e_2^* \neq e_1^*$.

Définition-Proposition 1.12. Soit $\mathbb{C} = \{\epsilon_1, \ldots, \epsilon_n\}$ une base de E^*. Il existe une unique base $\mathbb{C}^\circ = \{e_1, \ldots, e_n\}$ de E telle que $e_i^* = \epsilon_i$ pour tout $i \in \{1, \ldots, n\}$. Cette base s'appelle base *antéduale* de $\{\epsilon_1, \ldots, \epsilon_n\}$.

Démonstration. On note $\{\epsilon_1^*, \ldots, \epsilon_n^*\}$ la base de E^{**} obtenue en prenant la base duale de $\{\epsilon_1, \ldots, \epsilon_n\}$.
Existence d'une base antéduale : D'après le théorème 1.7, pour tout $i \in \{1, \ldots, n\}$, il existe $e_i \in E$ tel que $\iota(e_i) = \epsilon_i^*$. Déjà, comme $\iota: E \to E^{**}$ est un isomorphisme d'espace vectoriels, on en déduit que $\{e_1, \ldots, e_n\}$ est une base de E puisque c'est l'image de la base $\{\epsilon_1^*, \ldots, \epsilon_n^*\}$ par l'isomorphisme $\iota^{-1}: E^{**} \to E$. Ensuite, pour tout $j \in \{1, \ldots, n\}$, on a

$$\epsilon_j(e_i) = \iota(e_i)(\epsilon_j) = \epsilon_i^*(\epsilon_j) = \delta_{ij}.$$

On a donc bien $e_j^* = \epsilon_j$, pour tout $j \in \{1, \ldots, n\}$, et donc $\{e_1, \ldots, e_n\}$ est une base antéduale de $\{\epsilon_1, \ldots, \epsilon_n\}$.
Unicité de la base antéduale : Si $\{e_1, \ldots, e_n\}$ est une base antéduale de $\{\epsilon_1, \ldots, \epsilon_n\}$ (c'est-à-dire une base de E telle que $e_i^* = \epsilon_i$ pour tout $i \in \{1, \ldots, n\}$), alors

$$\forall i, j \in \{1, \ldots, n\},\ \iota(e_i)(\epsilon_j) = \epsilon_j(e_i) = e_j^*(e_i) = \delta_{ij}.$$

Donc $\iota(e_i) = \epsilon_i^*$. Mais on a vu que ι est un isomorphisme (théorème 1.7), et donc chaque $e_i = \iota^{-1}(\epsilon_i^*)$ est unique. □

Le résultat qui suit est une conséquence immédiate des définitions-propositions 1.9 et 1.12.

Corollaire 1.13. Il y a une bijection naturelle entre l'ensemble des bases de E et l'ensemble des bases de E^* ; cette bijection est donnée par

$$\begin{array}{ccc} \{\text{bases de } E\} & \to & \{\text{bases de } E^*\} \\ \mathcal{B} & \mapsto & \mathcal{B}^* \\ \mathcal{C}^\circ & \leftarrow\!\shortmid & \mathcal{C} \end{array}$$

Nous allons maintenant nous intéresser au lien entre application transposée (au sens de la définition 1.4) et matrice transposée (au sens habituel).

Proposition 1.14. Soient E et F deux $\mathbb{K}$-espaces vectoriels de dimension n et p respectivement. On fixe des bases $\mathcal{B}$ = et $\mathcal{C}$ de E et F respectivement. Si $\varphi: E \to F$ est une application linéaire, alors

$$\mathrm{Mat}_{\mathcal{C}^*,\mathcal{B}^*}({}^t\varphi) = {}^t\mathrm{Mat}_{\mathcal{B},\mathcal{C}}(\varphi).$$

Démonstration. On pose

$$\begin{aligned}
&\mathcal{B} = \{e_1, \ldots, e_n\};\\
&\mathcal{C} = \{f_1, \ldots, f_p\};\\
&\mathrm{Mat}_{\mathcal{B},\mathcal{C}}(\varphi) = (\alpha_{ij})_{ij} \in \mathcal{M}_{p,n}(\mathbb{K}); \text{ et}\\
&\mathrm{Mat}_{\mathcal{C}^*,\mathcal{B}^*}({}^t\varphi) = (\beta_{hk})_{hk} \in \mathcal{M}_{n,p}(\mathbb{K}).
\end{aligned}$$

Alors, pour tous $h \in \{1, \ldots, n\}$, $k \in \{1, \ldots, p\}$, on a

$$\beta_{hk} = \left(\sum_{j=1}^{n} \beta_{jk} e_j^*\right)(e_h) = \left({}^t\varphi(f_k^*)\right)(e_h) = (f_k^* \circ \varphi)(e_h) = f_k^*(\varphi(e_h)) = f_k^*\left(\sum_{i=1}^{p} \alpha_{ih} f_i\right) = \alpha_{kh}.$$

On a donc $\beta_{hk} = \alpha_{kh}$, pour tous $h \in \{1, \ldots, n\}$, $k \in \{1, \ldots, p\}$, ce qui revient bien à dire que la matrice $\mathrm{Mat}_{\mathcal{B},\mathcal{C}}(\varphi)$ est la transposée de la matrice $\mathrm{Mat}_{\mathcal{C}^*,\mathcal{B}^*}({}^t\varphi)$. □

Corollaire 1.15. Soit $\varphi: E \to F$ une application linéaire entre espaces vectoriels de dimension finie. Alors $\mathrm{rg}(\varphi) = \mathrm{rg}({}^t\varphi)$.

Démonstration. En effet, si l'on fixe des bases $\mathcal{B}$ et $\mathcal{C}$ de E et F respectivement, on a

$$\mathrm{rg}(\varphi) = \mathrm{rg}(\mathrm{Mat}_{\mathcal{B},\mathcal{C}}(\varphi)) = \mathrm{rg}({}^t\mathrm{Mat}_{\mathcal{B},\mathcal{C}}(\varphi)) = \mathrm{rg}(\mathrm{Mat}_{\mathcal{C}^*,\mathcal{B}^*}({}^t\varphi)) = \mathrm{rg}({}^t\varphi).$$

□

On va maintenant rappeler comment calculer une base duale ou antéduale via le calcul matriciel. Soient E un $\mathbb{K}$-espace vectoriel de dimension finie n, $\mathcal{B} = \{e_1, \ldots, e_n\}$ une base de E, et $\mathcal{B}^* = \{e_1^*, \ldots, e_n^*\}$ la base duale de $\mathcal{B}$. On considère une seconde base $\mathcal{B}' = \{f_1, \ldots, f_n\}$ de E.

Question 1.16. Comment calculer les coordonnées des f_i^* relativement à la base $\mathcal{B}^*$ connaissant les coordonnées des f_i relativement à la base $\mathcal{B}$?

Soient $i, j \in \{1, \ldots, n\}$. On note $f_i = \sum_{k=1}^{n} \lambda_{ki} e_k$ et $f_j^* = \sum_{k=1}^{n} \mu_{kj} e_k^*$. On pose

$$F_i := \mathrm{Mat}_{\mathcal{B}}(f_i) = \begin{bmatrix} \lambda_{1i} \\ \vdots \\ \lambda_{ni} \end{bmatrix} \quad \text{et} \quad T_j := \mathrm{Mat}_{\mathcal{B}^*}(f_j^*) = \begin{bmatrix} \mu_{1j} \\ \vdots \\ \mu_{nj} \end{bmatrix},$$

qui sont des vecteurs colonne de taille n. Puis on pose $F = \begin{bmatrix} F_1 & \cdots & F_n \end{bmatrix} = (\lambda_{ij})_{ij}$ et $T = \begin{bmatrix} T_1 & \cdots & T_n \end{bmatrix} = (\mu_{ij})_{ij}$, qui sont deux matrices carrées de taille n. Par hypothèse, on connaît les coefficients de la matrice F, et on cherche à calculer les coefficients de la matrice T (puisque les coordonnées des f_i^* relativement à la base $\mathcal{B}^*$ sont les coefficients de la i-ème colonne de la matrice T).

L'observation clef est la suivante : Pour tous $i, j \in \{1, \dots, n\}$, on a

$$ {}^tT_jF_i = {}^t\mathrm{Mat}_{\mathcal{B}^*}(f_j^*)\mathrm{Mat}_{\mathcal{B}}(f_i) = f_j^*(f_i) = \delta_{ij}. $$

Il s'ensuit que ${}^tTF = I_n$, et donc $T = {}^tF^{-1}$. Ainsi, les coordonnées des f_i^* relativement à la base $\mathcal{B}^*$ sont les coefficients de la i-ème colonne de la matrice ${}^tF^{-1}$.

En conclusion, calculer une base duale revient à inverser une matrice (puis à prendre la matrice transposée, mais cette seconde opération est aisée). Et si jamais on souhaite calculer une base antéduale, le principe est le même sauf que dans ce cas on connaît la matrice T et on cherche à calculer la matrice F qui est alors égale à ${}^tT^{-1}$.

Ce qui précède peut aussi se reformuler ainsi :

Proposition 1.17. Soient $\mathcal{B}$ et $\mathcal{B}'$ deux bases quelconques de l'espace vectoriel E.
Si $\mathrm{Pass}(\mathcal{B} \to \mathcal{B}') = P$, alors $\mathrm{Pass}(\mathcal{B}^* \to \mathcal{B}'^*) = {}^tP^{-1}$.

1.1.C Orthogonalité (au sens de la dualité)

Définition 1.18. Soit E un $\mathbb{K}$-espace vectoriel.
- On a une forme bilinéaire canonique

$$ \langle - | - \rangle \colon E^* \times E \to \mathbb{K}, \ (\epsilon, x) \mapsto \epsilon(x) $$

  appelée *crochet de dualité*.
- Deux éléments $\epsilon \in E^*$ et $x \in E$ sont dits *orthogonaux* si $\langle \epsilon | x \rangle = 0$.

Remarque 1.19. On veillera bien à ne pas confondre le crochet de dualité de la définition 1.18 avec un quelconque produit scalaire. En effet, ici on ne suppose pose que E est un $\mathbb{R}$-espace vectoriel euclidien ! (Mais bien sûr, les notations ne sont pas choisies au hasard, il y a bien un lien entre ces deux notions.)

Définition 1.20 (Orthogonal d'un sous-ensemble)**.** • Si F est un sous-ensemble de E, alors

$$ F^\perp := \{\epsilon \in E^* \mid \forall x \in F, \ \langle \epsilon | x \rangle = 0\} $$

  est un sous-espace vectoriel de E^*.
- Si G est un sous-ensemble de E^*, alors

$$ G^\circ := \{x \in E \mid \forall \epsilon \in G, \ \langle \epsilon | x \rangle = 0\} = \bigcap_{\epsilon \in G} \mathrm{Ker}(\epsilon) $$

  est un sous-espace vectoriel de E.

Proposition 1.21. On désigne par F, F_1, F_2 des sous-ensembles de E et par G, G_1, G_2 des sous-ensembles de E^*.
(i) Si $F_1 \subset F_2$, alors $F_2^\perp \subset F_1^\perp$.
(ii) Si $G_1 \subset G_2$, alors $G_2^\circ \subset G_1^\circ$.
(iii) Si $F \subset E$, alors $F^\perp = \mathrm{Vect}(F)^\perp$.
(iv) Si $G \subset E^*$, alors $G^\circ = \mathrm{Vect}(G)^\circ$.

Démonstration. Cf. exercice 8 de la fiche 1 de TD. □

Proposition 1.22. On suppose que $\dim(E) = n < \infty$.
(i) Si F est un sous-espace vectoriel de E, alors

$$\dim(F) + \dim(F^{\perp}) = \dim(E) \quad \text{et} \quad (F^{\perp})^{\circ} = F.$$

(ii) Si G est un sous-espace vectoriel de E^*, alors

$$\dim(G) + \dim(G^{\circ}) = \dim(E) \quad \text{et} \quad (G^{\circ})^{\perp} = G.$$

Démonstration de (i). On note $r = \dim(F)$ et soit $\{e_1, \dots, e_r\}$ une base de F que l'on complète en une base $\{e_1, \dots, e_n\}$ de E. On a $F = \mathrm{Vect}(\{e_1, \dots, e_r\})$, et donc d'après la proposition 1.21, on a $F^{\perp} = \{e_1, \dots, e_r\}^{\perp}$.

On note $\epsilon = \sum_{i=1}^{n} \lambda_i e_i^* \in E^*$. Alors

$$\epsilon \in F^{\perp} \Leftrightarrow \epsilon \in \{e_1, \dots, e_r\}^{\perp} \Leftrightarrow \forall j \in \{1, \dots, r\},\ \lambda_j = \epsilon(e_j) = \langle \epsilon | e_j \rangle = 0.$$

Donc $\epsilon \in F^{\perp}$ si et seulement si $\epsilon \in \mathrm{Vect}(\{e_{r+1}^*, \dots, e_n^*\})$, et donc $\dim(F^{\perp}) = n - r$. De plus, $(F^{\perp})^{\circ} = (\{e_1, \dots, e_r\}^{\perp})^{\circ} = (\{e_{r+1}^*, \dots, e_n^*\})^{\circ}$, d'où

$$x = \sum_{i=1}^{n} \alpha_i e_i \in (F^{\perp})^{\circ} \Leftrightarrow \forall j \in \{r+1, \dots, n\},\ \alpha_j = \langle e_j^* | x \rangle = 0,$$

ce qui implique

$$(F^{\perp})^{\circ} = \mathrm{Vect}(\{e_1, \dots, e_r\}) = F.$$

Démonstration de (ii). La démonstration est similaire à celle du point (i) et est laissée en exercice. $\square$

Remarque 1.23. On peut montrer que l'égalité $(F^{\perp})^{\circ} = F$ reste vraie lorsque $\dim(E) = \infty$. En revanche, seule l'inclusion $G \subset (G^{\circ})^{\perp}$ reste vraie lorsque $\dim(E) = \infty$. (On laisse le soin à l'étudiant.e motivé.e de justifier ces deux assertions !)

Corollaire 1.24. (Équations d'un sous-espace vectoriel en dimension finie.)
Soit E un $\mathbb{K}$-espace vectoriel de dimension finie n.

- Soient $\epsilon_1, \dots, \epsilon_p \in E^*$ telles que $\mathrm{rg}(\{\epsilon_1, \dots, \epsilon_p\}) = r$. Alors le sous-espace vectoriel $F = \{\epsilon_1, \dots, \epsilon_p\}^{\circ}$ est de dimension $n - r$.
- Réciproquement, si F est un sous-espace vectoriel de E de dimension $n - q$, il existe q formes linéaires $\epsilon_1, \dots, \epsilon_q \in E^*$ telles que $\mathrm{rg}(\epsilon_1, \dots, \epsilon_q) = q$ et $F = \{\epsilon_1, \dots, \epsilon_q\}^{\circ} = \cap_{i=1}^{q} \mathrm{Ker}(\epsilon_i)$.

Démonstration. Il s'agit essentiellement d'une reformulation de la proposition 1.22. On laisse donc la démonstration de ce corollaire en exercice. $\square$

Soit F un sous-espace vectoriel de E et $\varphi \in \mathcal{L}(E)$. On dit que F est *stable* par φ si $\varphi(F) \subset F$. Le résultat qui suit jouera un rôle clé dans le prochain chapitre.

Proposition 1.25. On suppose que $\dim(E) < \infty$. Soient F un sous-espace vectoriel de E et $\varphi \in \mathcal{L}(E)$. Alors F est stable par φ si et seulement si $F^{\perp}$ est stable par ${}^t\varphi$.

Démonstration. ($\Rightarrow$) : On suppose que $\varphi(F) \subset F$. Alors, pour tout $\epsilon \in F^\perp$, on a

$$0 = \langle \epsilon | \varphi(F) \rangle = \epsilon(\varphi(F)) = ({}^t\varphi(\epsilon))(F) = \langle {}^t\varphi(\epsilon) | F \rangle.$$

Il s'ensuit que ${}^t\varphi(\epsilon) \in F^\perp$, c'est-à-dire que $F^\perp$ est stable par ${}^t\varphi$.

($\Leftarrow$) : D'après le corollaire 1.24 (ici on utilise l'hypothèse $\dim(E) < \infty$), il existe r formes linéaires $\epsilon_1, \ldots, \epsilon_r \in E^*$ telles que $F = \{\epsilon_1, \ldots, \epsilon_r\}^\circ = \cap_{i=1}^r \mathrm{Ker}(\epsilon_i)$. Donc, pour tout $i \in \{1, \ldots, r\}$, on a $F \subset \mathrm{Ker}(\epsilon_i)$, c'est-à-dire $\epsilon_i \in F^\perp$. Comme $F^\perp$ est stable par ${}^t\varphi$, on a ${}^t\varphi(\epsilon_i)(F) = 0$, autrement dit, $\epsilon_i(\varphi(F)) = 0$, pour tout $i \in \{1, \ldots, r\}$, et donc $\varphi(F) \subset \cap_{i=1}^r \mathrm{Ker}(\epsilon_i) = F$. □

Remarque 1.26.

- Nous verrons au cours des prochaines séances que la proposition 1.25 est très utile dans certains raisonnements par récurrence en dimension finie pour montrer l'existence d'un sous-espace vectoriel stable non-trivial. Par exemple, en utilisant la proposition 1.25 on peut facilement montrer que n'importe quelle matrice complexe est trigonalisable.
- La proposition 1.25 reste vraie lorsque $\dim(E) = \infty$ mais la démonstration est plus compliquée et nécessite un recours à l'axiome du choix.

1.2 Espaces vectoriels quotients

1.2.A Généralités

Définition-Proposition 1.27. Soient E un $\mathbb{K}$-espace vectoriel et F un sous-espace vectoriel de E

- La relation binaire $\mathcal{R}$ sur E définie par $x\mathcal{R}y$ si $x - y \in F$ est une relation d'équivalence sur E.
- L'espace quotient $E/\mathcal{R}$, noté E/F, est un $\mathbb{K}$-espace vectoriel muni des lois

$$\overline{x} + \overline{y} := \overline{x+y} \quad \text{et} \quad \lambda\overline{x} := \overline{\lambda x}, \tag{1.1}$$

où $\overline{x}$ désigne l'image de $x \in E$ dans l'espace quotient $E/\mathcal{R}$.

Démonstration. On vérifie aisément que la relation binaire $\mathcal{R}$ sur E est une relation d'équivalence (c'est-à-dire que $\mathcal{R}$ est réflexive, symétrique et transitive). Ainsi l'espace quotient $E/F := E/\mathcal{R}$ est bien défini comme ensemble ; il reste à vérifier que E/F est naturellement muni d'une structure d'espace vectoriel, c'est à dire que l'addition et la multiplication externe définies par (1.1) sont bien définies et que ces deux lois vérifient toutes les propriétés qui définissent un espace vectoriel.

Pour montrer que l'addition et la multiplication externe données par (1.1) sont bien définies, on choisit $x_1, x_2, y_1, y_2 \in E$ tels que $\overline{x_1} = \overline{x_2}$ et $\overline{y_1} = \overline{y_2}$. On a alors

$$(x_1 + y_1) - (x_2 + y_2) = (x_1 - x_2) + (y_1 - y_2) \in F, \quad \text{et donc } \overline{x_1 + y_1} = \overline{x_2 + y_2}.$$

De même

$$\lambda x_2 - \lambda x_1 = \lambda(x_2 - x_1) \in F, \quad \text{et donc } \overline{\lambda x_1} = \overline{\lambda x_2}.$$

Ainsi l'addition et la multiplication externe définies par (1.1) sont bien définies.

On laisse alors le soin au lecteur/à la lectrice de ces notes de vérifier que ces deux lois vérifient toutes les propriétés qui définissent un espace vectoriel (ce n'est pas difficile mais un peu rébarbatif). □

Définition-Proposition 1.28. L'application

$$q\colon E \to E/F,\ x \mapsto \overline{x}$$

est une application linéaire surjective, appelé *morphisme de passage au quotient* ; elle vérifie $\mathrm{Ker}(q) = F$.

Démonstration. Le fait que l'application q soit linéaire découle de (1.1). En effet, on a

$$\forall x_1, x_2 \in E, \forall \lambda \in \mathbb{K},\ q(x_1 + \lambda x_2) = \overline{x_1 + \lambda x_2} = \overline{x_1} + \overline{\lambda x_2} = \overline{x_1} + \lambda \overline{x_2} = q(x_1) + \lambda q(x_2).$$

De plus, q est surjective par définition du quotient E/F. Enfin

$$\mathrm{Ker}(q) = \{x \in E \mid \overline{x} = \overline{0}\} = \{x \in E \mid x - 0 \in F\} = F.$$

□

Proposition 1.29. Soient E un espace vectoriel et F un sous-espace vectoriel de E. Alors

$$\dim(F) + \dim(E/F) = \dim(E).$$

Démonstration. Si $\dim(E) < \infty$, alors le résultat découle du théorème du rang appliqué au morphisme de passage au quotient $q\colon E \to E/F$.

Si $\dim(E) = \infty$, il faut vérifier que $\dim(F) = \infty$ ou $\dim(E/F) = \infty$. On procède par l'absurde en supposant que $\dim(F)$ et $\dim(E/F)$ sont finies. Soient $\mathcal{B}_1 = \{e_1, \dots, e_r\}$ une base de F et $\mathcal{B}_2 = \{t_1, \dots, t_s\}$ une base de E/F. Pour chaque $t_i \in E/F$, on choisit $b_i \in E$ tel que $q(b_i) = t_i$. On note alors $\mathcal{L}_2 := \{b_1, \dots, b_s\}$ qui est une famille d'éléments de E.

Montrons que $\mathcal{B}_1 \cup \mathcal{L}_2$ est une famille génératrice de E. En effet, si $x \in E$, alors

$$q(x) = \overline{x} = \sum_{i=1}^{s} \lambda_i t_i = q\left(\sum_{i=1}^{s} \lambda_i b_i\right).$$

Donc $x - \sum_i \lambda_i b_i \in \mathrm{Ker}(q) = F$. On peut donc décomposer $x - \sum_i \lambda_i b_i$ dans la base $\mathcal{B}_1$, ce qui donne

$$x - \sum_{i=1}^{s} \lambda_i b_i = \sum_{j=1}^{r} \mu_j e_j, \text{ et donc } x = \sum_{i=1}^{s} \lambda_i b_i + \sum_{j=1}^{r} \mu_j e_j.$$

Ainsi, $\mathcal{B}_1 \cup \mathcal{L}_2$ est une famille génératrice de E, d'où $\dim(E) < \infty$, ce qui contredit notre hypothèse $\dim(E) = \infty$. On a donc bien $\dim(F) = \infty$ ou $\dim(E/F) = \infty$. □

Définition 1.30. On appelle *codimension* de F dans E l'élément de $\mathbb{N} \cup \{\infty\}$ défini par

$$\mathrm{codim}_E(F) := \dim(E/F).$$

Proposition 1.31. Un sous-espace vectoriel F de E est de codimension finie dans E si et seulement si F admet un supplémentaire S dans E de dimension finie. On a alors $\mathrm{codim}_E(F) = \dim(S)$.

Démonstration. ($\Rightarrow$) : On suppose que F est de codimension finie dans E, c'est-à-dire que $\dim(E/F) < \infty$. On note $s := \dim(E/F)$. Soit $\{t_1, \dots, t_s\}$ une base de E/F, et soit $\mathcal{L} := \{b_1, \dots, b_s\}$ une famille d'éléments de E tels que $q(b_i) = t_i$ pour tout $i \in \{1, \dots, s\}$. On pose $S := \mathrm{Vect}(\{b_1, \dots, b_s\})$. Alors $\dim(S) < \infty$, et il reste à voir que $E = F \oplus S$.

- Si $v \in F \cap S$, alors $v = \sum_{i=1}^{s} \lambda_i b_i$ et $q(v) = \overline{0}$ car $\mathrm{Ker}(q) = F$. Il s'ensuit que

$$\overline{0} = q(v) = q\left(\sum_{i=1}^{s} \lambda_i b_i\right) = \sum_{i=1}^{s} \lambda_i q(b_i) = \sum_{i=1}^{s} \lambda_i t_i.$$

  Or la famille $\{t_1, \dots, t_s\}$ est libre, donc tous les λ_i sont nuls et $v = 0$. Autrement dit, $F \cap S = \{0\}$.
- Soit $v \in E$. Alors $q(v) = \sum_{i=1}^{s} \lambda_i t_i = \sum_{i=1}^{s} \lambda_i q(b_i) = q(\sum_{i=1}^{s} \lambda_i b_i)$. Il s'ensuit que

$$v - \sum_{i=1}^{s} \lambda_i b_i \in \mathrm{Ker}(q) = F, \quad \text{et donc } v = \sum_{i=1}^{s} \lambda_i b_i + f,$$

  pour un certain $f \in F$. Autrement dit, $E = S + F$.

On a donc bien $E = F \oplus S$. En outre, on vérifie aisément que la famille $\mathcal{L} \coloneqq \{b_1, \dots, b_s\}$ est libre, et donc que c'est une base de S. Il en découle que $\mathrm{codim}_E(F) = \dim(E/F) = s = \dim(S)$.

($\Leftarrow$) : On suppose que $E = F \oplus S$ avec S est un sous-espace vectoriel de dimension finie. On note $\mathcal{B} \coloneqq \{b_1, \dots, b_s\}$ une base de S. Nous allons montrer que $\overline{\mathcal{B}} \coloneqq \{\overline{b_1}, \dots, \overline{b_s}\}$ est une base de E/F.

- *$\overline{\mathcal{B}}$ est libre :* Soient $\lambda_1, \dots, \lambda_s \in \mathbb{K}$ tels que $\sum_{i=1}^{s} \lambda_i \overline{b_i} = \overline{0}$. Alors $\sum_{i=1}^{s} \lambda_i b_i \in F \cap S = \{0\}$, et donc tous les λ_i sont nuls puisque $\mathcal{B}$ est une base de S.
- *$\overline{\mathcal{B}}$ est génératrice :* On fixe $\overline{x} \in E/F$, et soit $x \in E$ tel que $q(x) = \overline{x}$. On a $x = \sum_{i=1}^{s} \lambda_i b_i + f$, pour un certain $f \in F$, et donc

$$\overline{x} = q(x) = q\left(\sum_{i=1}^{s} \lambda_i b_i + f\right) = \sum_{i=1}^{s} \lambda_i q(b_i) + q(f) = \sum_{i=1}^{s} \lambda_i q(b_i) = \sum_{i=1}^{s} \lambda_i \overline{b_i}.$$

On a donc montré que la famille $\overline{\mathcal{B}}$ est la fois libre et génératrice, autrement dit c'est une base de E/F. En particulier, $\dim(S) = s = \dim(E/F) = \mathrm{codim}_E(F)$. □

Remarque 1.32. Avec les notations de la proposition 1.31, on peut vérifier que la restriction du morphisme de passage au quotient $q_{|S} \colon S \to E/F$ est un isomorphisme d'espaces vectoriels. Et donc, moralement, travailler avec l'espace quotient E/F revient à travailler avec un supplémentaire de F dans E mais sans faire le choix explicite de ce supplémentaire.

Théorème 1.33. Soient E un $\mathbb{K}$-espace vectoriel et F un sous-espace vectoriel de E. Les assertions suivantes sont équivalentes :

(i) Le sous-espace vectoriel F est un hyperplan de E (cf. définition 1.1).
(ii) On a $\dim(E/F) = 1$.
(iii) Il existe une droite vectorielle D de E telle que $E = F \oplus D$.

Démonstration. (ii) $\Leftrightarrow$ (iii) **:** Il s'agit de la proposition 1.31 avec $S = D$.

(ii) $\Rightarrow$ (i) **:** Étant donné que $\dim(E/F) = 1$, il existe un isomorphisme d'espaces vectoriels $\psi \colon E/F \simeq \mathbb{K}$, et alors $\psi \circ q$ est une forme linéaire surjective (donc non-nulle) sur E telle que $\mathrm{Ker}(\psi \circ q) = \mathrm{Ker}(q) = F$. Ainsi, F est le noyau d'une forme linéaire non-nulle sur E, donc un hyperplan de E.

(i) $\Rightarrow$ (iii) **:** Soit $\epsilon \in E^*$ une forme linéaire non-nulle sur E telle que $F = \mathrm{Ker}(\epsilon)$. Puisque $\epsilon \neq 0_{E^*}$, il existe $v_0 \in E$ tel que $\epsilon(v_0) = 1$. On pose $D = \langle v_0 \rangle$ la droite vectorielle engendrée par v_0 dans E. Nous allons voir que $E = F \oplus D$.

- Si $v \in D \cap F$, alors $v = \lambda v_0$ et $q(v) = \overline{0}$. Il s'ensuit que $\lambda = \lambda q(v_0) = q(\lambda v_0) = q(v) = \overline{0}$. Donc $D \cap F = \{0\}$.

- Soit $v \in E$. Alors $\epsilon(v) = \mu = \mu\epsilon(v_0) = \epsilon(\mu v_0)$. Ainsi $v - \mu v_0 \in \mathrm{Ker}(\epsilon) = F$, et donc $v = \mu v_0 + f$, pour un certain $f \in F$, c'est-à-dire $E = D + F$.

On a donc $E = D \oplus F$, ce qui termine la démonstration de la proposition. □

Remarque 1.34. Dans la littérature, on trouve parfois qu'un hyperplan F de E est défini comme un sous-espace vectoriel F de E tel que $\dim(E/F) = 1$. D'après le théorème 1.33, cette définition est équivalente à la définition 1.1 ; on pourra donc utiliser indifféremment n'importe laquelle de ces deux définitions lorsqu'on parle d'hyperplan.

1.2.B Propriété universelle du morphisme de passage au quotient

Théorème 1.35. Soit $\varphi: E \to E'$ une application linéaire entre deux espaces vectoriels E et E'. Soit F un sous-espace vectoriel de E tel que $\varphi(F) = \{0\}$. Alors il existe une unique application linéaire

$$\overline{\varphi}: E/F \to E' \text{ telle que } \varphi = \overline{\varphi} \circ q. \tag{1.2}$$

Autrement dit, on a le diagramme commutatif suivant :

$$\begin{tikzcd} E \arrow[rr, "\varphi"] \arrow[dr, "q"'] & & E' \\ & E/F \arrow[ur, "\overline{\varphi}"'] & \end{tikzcd}$$

Démonstration. Si l'application $\overline{\varphi}$ existe, alors elle est uniquement déterminée par l'égalité (1.2) puisque $q: E \to E/F$ est surjective, et on doit alors avoir

$$\forall \overline{x} \in E/F,\ \overline{\varphi}(\overline{x}) := \varphi(x).$$

Vérifions que l'application $\overline{\varphi}$ est bien définie. Soient $x_1, x_2 \in E$ tels que $\overline{x_1} = \overline{x_2}$. Alors $\varphi(x_1) - \varphi(x_2) = \varphi(x_1 - x_2) \in \varphi(F) = \{0\}$, et donc la condition $\overline{x_1} = \overline{x_2}$ implique $\varphi(x_1) = \varphi(x_2)$, ce qui signifie que l'élément $\overline{\varphi}(\overline{x}) := \varphi(x)$ ne dépend pas du choix de x dans sa classe d'équivalence $\overline{x}$.

Enfin, une fois que l'on a vérifié l'existence de $\overline{\varphi}$, on vérifie aisément que cette application est linéaire en utilisant (1.1), (1.2) et la linéarité de φ. En effet, on a

$$\begin{aligned} \forall \overline{x_1}, \overline{x_2} \in E/F, \forall \lambda \in \mathbb{K},\ \overline{\varphi}(\overline{x_1} + \lambda\overline{x_2}) &= \overline{\varphi}(\overline{x_1 + \lambda x_2}) = \overline{\varphi}(q(x_1 + \lambda x_2)) = \varphi(x_1 + \lambda x_2) \\ &= \varphi(x_1) + \lambda\varphi(x_2) = \overline{\varphi}(q(x_1)) + \lambda\overline{\varphi}(q(x_2)) = \overline{\varphi}(\overline{x_1}) + \lambda\overline{\varphi}(\overline{x_2}). \end{aligned}$$

□

L'intérêt du théorème 1.35 est qu'il permet de construire des isomorphismes canoniques entre certains $\mathbb{K}$-espaces vectoriels (cf. les deux prochaines propositions ainsi que les exercices de la fiche 1 de TD).

Proposition 1.36. Soit $\varphi: E \to E'$ une application linéaire entre espaces vectoriels. Alors φ induit l'isomorphisme d'espaces vectoriels

$$E/\mathrm{Ker}(\varphi) \simeq \mathrm{Im}(\varphi).$$

Démonstration. On note $\varphi' : E \to \operatorname{Im}(\varphi),\ x \mapsto \varphi(x)$ qui est une application linéaire surjective telle que $\operatorname{Ker}(\varphi) = \operatorname{Ker}(\varphi')$. (En pratique on a juste remplacé l'espace vectoriel E' par $\operatorname{Im}(\varphi)$ pour se ramener au cas d'une application linéaire surjective.)

D'après le théorème 1.35 appliqué à φ' avec $F := \operatorname{Ker}(\varphi') = \operatorname{Ker}(\varphi)$, on a le diagramme commutatif suivant :

$$\begin{tikzcd}
E \arrow[rr, "\varphi'"] \arrow[dr, "q"'] & & \operatorname{Im}(\varphi) \\
& E/\operatorname{Ker}(\varphi) \arrow[ur, "\overline{\varphi'}"'] &
\end{tikzcd}$$

Étant donné que $\varphi' = \overline{\varphi'} \circ q$ est aussi surjective, on en déduit que $\overline{\varphi'}$ est surjective. Par ailleurs

$$\operatorname{Ker}(\overline{\varphi'}) = \{\overline{x} \mid \overline{\varphi'}(\overline{x}) = 0\} = \{\overline{x} \mid \varphi'(x) = 0\} = \{\overline{x} \mid x \in \operatorname{Ker}(\varphi') = \operatorname{Ker}(\varphi)\} = \{\overline{0}\},$$

et donc $\overline{\varphi'}$ est injective. Il s'ensuit que $\overline{\varphi'}$ est un isomorphisme d'espaces vectoriels. □

Remarque 1.37. Si $\dim(E) < \infty$, on retrouve le théorème du rang en passant aux dimensions dans

$$E/\operatorname{Ker}(\varphi) \simeq \operatorname{Im}(\varphi)$$

et en appliquant la proposition 1.29 au membre de gauche.

On termine ce premier chapitre avec le résultat suivant qui fait le lien entre les sections 1.1 et 1.2.

Proposition 1.38. Soit F un sous-espace vectoriel de E. Alors il y a un isomorphisme canonique d'espaces vectoriels

$$(E/F)^* \simeq F^\perp.$$

Démonstration. Cf. l'exercice 15 de la fiche 1 de TD. □

Quelques références bibliographiques

Voici des références générales (liste non-exhaustive bien sûr) concernant la dualité et les espaces vectoriels quotients :

- Xavier Gourdon. *Algèbre*, Chapitre 3.
- Joseph Grifone. *Algèbre linéaire*, Chapitre 3 et Appendice A.3.
- Edmond Ramis, Claude Deschamps, Jacques Odoux.
  Le cours de mathématiques Tome 1- Algèbre, Chapitre 9.

Chapitre 2

Réduction des endomorphismes

Dans ce deuxième chapitre on étudie la réduction des endomorphismes. Étant donnés un espace vectoriel E de dimension finie $n \geq 1$ et $f \in \mathcal{L}(E)$, le but est de trouver une base $\mathcal{B}$ de E telle que la matrice $\mathrm{Mat}_{\mathcal{B}}(f)$ soit la plus "simple" possible (par exemple avec un maximum de zéros en dehors de la diagonale). De manière équivalente, étant donnée une matrice $M \in \mathcal{M}_n(\mathbb{K})$, on cherche $P \in \mathrm{GL}_n(\mathbb{K})$ telle que la matrice $P^{-1}MP$ soit la plus "simple" possible.

Les deux réductions que nous allons étudier dans ce chapitre sont la réduction de Frobenius (valide sur n'importe quel corps, commode pour déterminer si deux endomorphismes sont semblables mais pas pour les applications numériques) et la réduction de Jordan (valide uniquement lorsque le polynôme caractéristique est scindé sur le corps $\mathbb{K}$, très commode pour les applications numériques telles que le calcul de puissances ou de l'exponentielle).

Les sections 2.1 et 2.2 sont essentiellement des rappels de résultats d'algèbre linéaires vus les années précédentes, la section 2.3 est consacrée à la réduction de Frobenius et la section 2.4 est consacrée à la réduction de Jordan.

Dans tout ce chapitre, $\mathbb{K}$ désigne un corps quelconque, par exemple $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$... et E désigne un $\mathbb{K}$-espace vectoriel de dimension finie $n \geq 1$.

2.1 Endomorphismes nilpotents, diagonalisables, trigonalisables

2.1.A Polynôme caractéristique

Définition-Proposition 2.1.

- Si $M \in \mathcal{M}_n(\mathbb{K})$, on définit le *polynôme caractéristique* de M par

$$\chi_M(X) := \det(M - XI_n)$$

qui est un polynôme de degré n dont le coefficient dominant est $(-1)^n$.

- De même, si $f \in \mathcal{L}(E)$, on définit le *polynôme caractéristique* de f par

$$\chi_f(X) := \det(\mathrm{Mat}_{\mathcal{B}}(f) - XI_n)$$

qui est un polynôme de degré n dont le coefficient dominant est $(-1)^n$, et qui ne dépend pas du choix de la base $\mathcal{B}$ de E.

Démonstration. La seule chose à vérifier est que le polynôme χ_f défini ci-dessus est indépendant du choix de la base $\mathcal{B}$ de E. On fixe donc deux bases $\mathcal{B}_1$ et $\mathcal{B}_2$ de E, et on note

$$M_1 := \mathrm{Mat}_{\mathcal{B}_1}(f), \quad M_2 := \mathrm{Mat}_{\mathcal{B}_2}(f) \quad \text{et} \quad P := \mathrm{Pass}(\mathcal{B}_1 \to \mathcal{B}_2).$$

Alors $M_2 = P^{-1} M_1 P$ et

$$\begin{aligned}
\chi_{M_2}(X) &= \det(M_2 - XI_n) \\
&= \det(P^{-1}M_1P - XI_n) \\
&= \det(P^{-1}(M_1 - XI_n)P) \\
&= \det(P^{-1})\det(M_1 - XI_n)\det(P) \\
&= \det(M_1 - XI_n) \\
&= \chi_{M_1}(X),
\end{aligned}$$

ce qui montre que le polynôme χ_f ne dépend pas du choix de la base $\mathcal{B}$ de E. □

On rappelle qu'une *valeur propre* de $f \in \mathcal{L}(E)$ est un élément $\lambda \in \mathbb{K}$ pour lequel il existe un *vecteur propre* $v \in E$, c'est-à-dire un élément $v \in E \setminus \{0\}$ tel que $f(v) = \lambda v$. En pratique, on utilise souvent le résultat suivant pour déterminer les valeurs propres d'un endomorphisme donné.

Proposition 2.2. Soient $f \in \mathcal{L}(E)$ et $\lambda \in \mathbb{K}$. Alors λ est valeur propre de f si et seulement si $\chi_f(\lambda) = 0$.

Démonstration. Le résultat découle du fait que l'on a les équivalences suivantes :

$$\begin{aligned}
\lambda \text{ est valeur propre de } f &\Leftrightarrow \exists v \in E \setminus \{0\} \text{ tel que } f(v) = \lambda v \\
&\Leftrightarrow \mathrm{Ker}(f - \lambda \mathrm{Id}_E) \neq \{0\} \\
&\Leftrightarrow \text{la matrice } \mathrm{Mat}_{\mathcal{B}}(f - \lambda \mathrm{Id}_E) \text{ n'est pas inversible} \\
&\Leftrightarrow \chi_f(\lambda) := \det(\mathrm{Mat}_{\mathcal{B}}(f) - \lambda I_n) = \det(\mathrm{Mat}_{\mathcal{B}}(f - \lambda \mathrm{Id}_E)) = 0,
\end{aligned}$$

où $\mathcal{B}$ est n'importe quelle base de E. □

Proposition 2.3. Soit $f \in \mathcal{L}(E)$ et $F \subseteq E$ un sous-espace vectoriel non-nul de E *stable* par f, c'est-à-dire tel que $f(F) \subseteq F$. Soit $g := f_{|F} \in \mathcal{L}(F)$. Alors $\chi_g \mid \chi_f$.

Démonstration. Soit $\mathcal{B}_0$ une base de F que l'on complète en une base $\mathcal{B}$ de E. On note $r < n$ la dimension du sous-espace vectoriel F. Alors la matrice de f dans la base $\mathcal{B}$ est diagonale par blocs, avec en haut à gauche le bloc $\mathrm{Mat}_{\mathcal{B}_0}(f_{|F}) = \mathrm{Mat}_{\mathcal{B}_0}(g)$, d'où $\mathrm{Mat}_{\mathcal{B}}(f) = \begin{bmatrix} \mathrm{Mat}_{\mathcal{B}_0}(g) & M_1 \\ 0 & M_2 \end{bmatrix}$ avec $M_1 \in \mathcal{M}_{r,n-r}(\mathbb{K})$ et $M_2 \in \mathcal{M}_{n-r,n-r}(\mathbb{K})$. On en déduit

$$\chi_f(X) = \det(\mathrm{Mat}_{\mathcal{B}}(f) - XI_n) = \det(\mathrm{Mat}_{\mathcal{B}_0}(g) - XI_r)\det(M_2 - XI_{n-r}) = \chi_g(X)\chi_{M_2}(X).$$

En particulier, $\chi_g \mid \chi_f$. □

Corollaire 2.4. Soit λ une valeur propre de $f \in \mathcal{L}(E)$. Alors la dimension de l'espace propre E_λ est inférieure ou égale à la multiplicité de λ comme racine de χ_f.

Démonstration. Il suffit d'appliquer le résultat précédent avec $F = E_\lambda$. En effet, on a alors $g = f_{|E_\lambda} = \lambda \mathrm{Id}_{E_\lambda}$, et donc $\chi_g(X) = (\lambda - X)^{\dim(E_\lambda)}$. D'après la proposition 2.3, on a donc que $(\lambda - X)^{\dim(E_\lambda)}$ divise χ_f, autrement dit, que la multiplicité de λ comme racine de χ_f est au moins égale à la dimension de l'espace propre E_λ. □

2.1.B Endomorphismes nilpotents

Définition 2.5. Un endomorphisme $f \in \mathcal{L}(E)$ est dit *nilpotent* s'il existe $r \geq 1$ tel que $f^r = f \circ \cdots \circ f = 0$. Le plus petit entier $r \geq 1$ tel que $f^r = 0$ est appelé *l'indice de nilpotence* de f.

Exemple 2.6. On fixe $n \geq 1$ et on considère $E = \mathbb{K}[X]_{\leq n-1}$. Alors les endomorphismes de E définis par

$$f_1(P)(X) := P'(X) \quad \text{et} \quad f_2(P)(X) := P(X+1) - P(X)$$

sont nilpotents d'indice n.

Le résultat qui suit nous permet de caractériser les endomorphismes nilpotents au moyen de leur polynôme caractéristique. Cette observation jouera un rôle crucial pour déduire la réduction de Jordan à partir de la réduction de Frobenius dans la section 2.4.

Proposition 2.7. Soit $f \in \mathcal{L}(E)$. Alors f est nilpotent si et seulement si $\chi_f(X) = (-1)^n X^n$.

Démonstration. ($\Rightarrow$) On va démontrer la proposition par récurrence sur $n = \dim(E) \geq 1$. Soit $f \in \mathcal{L}(E)$ tel que $f^r = 0$ pour un certain $r \geq 1$. Si $n = 1$, alors $f^r = 0$ implique $f = 0$, et alors $\chi_f(X) = -X$. On suppose maintenant le résultat vrai au rang $n-1 \geq 1$ et on va le prouver au rang n. On a $0 = \det(f^r) = \det(f)^r$, et donc $\det(f) = 0$, autrement dit, $\mathrm{Ker}(f) \neq \{0\}$. Soit $e_1 \in \mathrm{Ker}(f) \setminus \{0\}$. Alors $f(e_1) = 0$. On complète $\{e_1\}$ en une base $\mathcal{B} = \{e_1, \dots, e_n\}$ de E. Alors

$$\mathrm{Mat}_{\mathcal{B}}(f) = \begin{bmatrix} 0 & * & \cdots & * \\ 0 & & & \\ \vdots & & M & \\ 0 & & & \end{bmatrix},$$

avec $M \in \mathcal{M}_{n-1}(\mathbb{K})$, et donc

$$0 = \mathrm{Mat}_{\mathcal{B}}(f^r) = \mathrm{Mat}_{\mathcal{B}}(f)^r = \begin{bmatrix} 0 & * & \cdots & * \\ 0 & & & \\ \vdots & & M^r & \\ 0 & & & \end{bmatrix},$$

d'où $M^r = 0$. D'après l'hypothèse de récurrence, appliquée à l'endomorphisme représenté par la matrice M dans la base $\{e_2, \dots, e_n\}$, on a $\chi_M(X) = (-1)^{n-1} X^{n-1}$. Il s'ensuit que

$$\chi_f(X) = \det(\mathrm{Mat}_{\mathcal{B}}(f) - X I_n) = \det\left(\begin{bmatrix} -X & * & \cdots & * \\ 0 & & & \\ \vdots & & M - X I_{n-1} & \\ 0 & & & \end{bmatrix}\right) = -X \chi_M(X) = (-1)^n X^n.$$

($\Leftarrow$) L'autre implication est une conséquence du théorème de Cayley-Hamilton qui affirme que $\chi_f(f) = 0$. En effet, si $\chi_f(X) = (-1)^n X^n$ et $\chi_f(f) = 0$, alors $f^n = (-1)^n \chi(f) = 0$, et donc l'endomorphisme f est nilpotent. (En particulier on voit que l'indice de nilpotence d'un endomorphisme nilpotent de E est toujours borné par la dimension de E.) $\square$

2.1.C Endomorphismes diagonalisables

Définition 2.8. Soit $f \in \mathcal{L}(E)$. On dit que f est *diagonalisable* s'il existe une base de vecteurs propres de f, autrement dit, s'il existe une base $\mathcal{B}$ de E telle que $\mathrm{Mat}_{\mathcal{B}}(f)$ soit diagonale.

Exemples 2.9.

(i) La matrice $M_1 := \begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix}$ est diagonalisable mais pas $M_2 := \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$.

(ii) Soient $d \geq 1$ et $E = \mathcal{M}_d(\mathbb{K})$. L'application

$$\mathcal{M}_d(\mathbb{K}) \to \mathcal{M}_d(\mathbb{K}),\ M \mapsto {}^tM$$

est diagonalisable, et on a $E = E_{-1} \oplus E_1$.

Théorème 2.10. Soit $f \in \mathcal{L}(E)$. Les assertions suivantes sont équivalentes :

(i) f est diagonalisable

(ii) χ_f est scindé (sur $\mathbb{K}$) et, pour toute racine λ_i d'ordre α_i de χ_f, on a $\dim(E_{\lambda_i}) = \alpha_i$

(iii) les valeurs propres $\lambda_1, \ldots, \lambda_p$ de f satisfont $E = E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_p}$

Démonstration. (i) $\Rightarrow$ (ii) : Si f est diagonalisable, alors il existe une base $\mathcal{B}$ de E telle que $\mathrm{Mat}_{\mathcal{B}}(f) = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$. Alors $\chi_f(X) = \prod_{i=1}^n (\lambda_i - X)$, ce qui implique le point (ii).
(ii) $\Rightarrow$ (iii) : On écrit $\chi_f(X) = (-1)^n \prod_{i=1}^p (X - \lambda_i)^{\alpha_i}$, où les λ_i sont deux à deux distincts et $\alpha_i \geq 1$ est l'ordre de λ_i comme racine de χ_f. Par hypothèse, on a $\alpha_i = \dim(E_{\lambda_i})$. On pose alors $F := \bigoplus_{i=1}^p E_{\lambda_i}$, qui est un sous-espace vectoriel de E dont la dimension vaut

$$\dim(F) = \sum_{i=1}^p \dim(E_{\lambda_i}) = \sum_{i=1}^p \alpha_i = \deg(\chi_f) = n = \dim(E).$$

Il s'ensuit que $E = \bigoplus_{i=1}^p E_{\lambda_i}$.
(iii) $\Rightarrow$ (i) : Pour tout $i \in \{1, \ldots, p\}$, on fixe une base $\mathcal{B}_i$ de E_{λ_i}. Alors l'hypothèse $E = \bigoplus_{i=1}^p E_{\lambda_i}$ implique que $\mathcal{B} = \mathcal{B}_1 \cup \cdots \cup \mathcal{B}_p$ est une base de E formée de vecteurs propres de f, et donc f est diagonalisable. □

Le résultat qui suit, qui est une conséquence immédiate du théorème 2.10, donne une condition suffisante généralement facile à vérifier pour qu'un endomorphisme soit diagonalisable. On prendra cependant bien garde au fait que cette condition n'est pas une équivalence !

Corollaire 2.11. Soit $f \in \mathcal{L}(E)$. Si χ_f est scindé (sur $\mathbb{K}$) est à racines simples, alors f est diagonalisable.

Démonstration. On écrit $\chi_f(X) = (-1)^n \prod_{i=1}^n (X - \lambda_i)$, où les λ_i sont deux à deux distincts. D'après le corollaire 2.4, on a $\dim(E_{\lambda_i}) = 1$ pour tout $i \in \{1, \ldots, n\}$. Le résultat découle alors de l'équivalence entre les points (ii) et (i) dans le théorème 2.10. □

Enfin, le résultat qui suit sera prouvé dans la section 2.2 (remarque 2.27), mais nous allons en avoir besoin pour démontrer le théorème 2.19 (diagonalisation simultanée), d'où le choix de le faire figurer dans cette section.

Proposition 2.12. Soient $f \in \mathcal{L}(E)$ un endomorphisme diagonalisable et F un sous-espace vectoriel de E stable par f. Alors l'endomorphisme $f_{|F} \in \mathcal{L}(F)$ est diagonalisable.

2.1.D Endomorphismes trigonalisables

Définition 2.13. Soit $f \in \mathcal{L}(E)$. On dit que f est *trigonalisable* s'il existe une base $\mathcal{B}$ de E telle que $\mathrm{Mat}_{\mathcal{B}}(f)$ soit triangulaire supérieure.

Exemple 2.14. La matrice $M_1 := \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$ n'est pas trigonalisable dans $\mathcal{M}_2(\mathbb{R})$ mais $M_2 := \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$ l'est, elle est même diagonalisable dans $\mathcal{M}_2(\mathbb{R})$. En revanche, M_1 et M_2 sont toutes les deux trigonalisables dans $\mathcal{M}_2(\mathbb{C})$.

Remarque 2.15. Un endomorphisme diagonalisable est trigonalisable. Mais la réciproque est fausse, par exemple la matrice $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ est trigonalisable (et même triangulaire supérieure) mais n'est pas diagonalisable.

Théorème 2.16. Soit $f \in \mathcal{L}(E)$. Alors f est trigonalisable (sur $\mathbb{K}$) si et seulement si χ_f est scindé (sur $\mathbb{K}$).

Démonstration. ($\Rightarrow$) : Soit $\mathcal{B}$ une base de E telle que $M := \mathrm{Mat}_{\mathcal{B}}(f) = \begin{bmatrix} \lambda_1 & * & * & * \\ 0 & \lambda_2 & * & * \\ 0 & 0 & \ddots & * \\ 0 & 0 & 0 & \lambda_n \end{bmatrix}$. Alors $\chi_f(X) = \chi_M(X) = \prod_{i=1}^n (\lambda_i - X)$ est scindé (sur $\mathbb{K}$).

($\Leftarrow$) : On va procéder par récurrence sur $n = \dim(E) \geq 1$. Si $n = 1$, alors le résultat est immédiat. On suppose maintenant le résultat vrai au rang $n-1 \geq 1$ et on va le prouver au rang n. On commence par remarquer que l'on a l'égalité de polynômes caractéristiques $\chi_{{}^t f} = \chi_f$, où l'on rappelle que l'on note ${}^t f \colon E^* \to E^*$ l'application *transposée* de f (c.f. définition 1.4). En effet, pour toute base $\mathcal{B}$ de E, on a

$$\begin{aligned}
\chi_{{}^t f}(X) &= \det(\mathrm{Mat}_{\mathcal{B}^*}({}^t f) - X I_n) \\
&= \det({}^t\mathrm{Mat}_{\mathcal{B}}(f) - X I_n) \quad \text{d'après la proposition 1.14} \\
&= \det({}^t\mathrm{Mat}_{\mathcal{B}}(f - X\mathrm{Id}_E)) \\
&= \det(\mathrm{Mat}_{\mathcal{B}}(f - X\mathrm{Id}_E)) \quad \text{car on a toujours } \det({}^t M) = \det(M) \\
&= \det(\mathrm{Mat}_{\mathcal{B}}(f) - X I_n) = \chi_f(X),
\end{aligned}$$

où $\mathcal{B}^*$ est la base duale de $\mathcal{B}$. En particulier, $\chi_{{}^t f}$ est scindé puisque χ_f est scindé par hypothèse. Donc l'endomorphisme ${}^t f \in \mathcal{L}(E^*)$ admet un vecteur propre $\epsilon_0 \in E^* \smallsetminus \{0\}$. Étant donnée que la droite vectorielle $\langle \epsilon_0 \rangle$ dans E^* est stable par ${}^t f$, on déduit de la proposition 1.25 que l'hyperplan $H := \langle \epsilon_0 \rangle^\circ$ de E est stable par f. Ensuite, d'après la proposition 2.3, le polynôme caractéristique $\chi_{f_{|H}}$ divise χ_f, et donc $\chi_{f_{|H}}$ est scindé. D'après l'hypothèse de récurrence, il existe donc une base $\mathcal{B}_0 = \{e_1, \dots, e_{n-1}\}$ de H telle que $\mathrm{Mat}_{\mathcal{B}_0}(f_{|H})$ est triangulaire supérieure. On complète cette base $\mathcal{B}_0$ de H en une base $\mathcal{B}$ de E, on a alors $\mathrm{Mat}_{\mathcal{B}}(f) = \begin{bmatrix} & & & * \\ & \mathrm{Mat}_{\mathcal{B}_0}(f_{|H}) & & * \\ & & & * \\ 0 & \cdots & 0 & * \end{bmatrix}$ qui est triangulaire supérieure. Ainsi f est trigonalisable, ce qui conclut la démonstration. $\square$

Remarques 2.17.

(i) Le théorème 2.16 implique que tout endomorphisme est trigonalisable sur un corps algébriquement clos (par exemple lorsque $\mathbb{K} = \mathbb{C}$) car alors χ_f est toujours scindé.

(ii) Si $f \in \mathcal{L}(E)$ et $F \subseteq E$ est un sous-espace vectoriel de E stable par f, alors le théorème 2.16 implique que $g := f_{|F} \in \mathcal{L}(F)$ est trigonalisable puisque $\chi_g \mid \chi_f$ d'après la proposition 2.3.

2.1.E Réductions simultanées

Lemme 2.18. Soient $f, g \in \mathcal{L}(E)$ tels que $f \circ g = g \circ f$. Alors :

(i) Tout espace propre de f est stable par g. En particulier $\mathrm{Ker}(f) = E_0$ est stable par g.

(ii) Le sous-espace vectoriel $\mathrm{Im}(f)$ est stable par g.

Démonstration. (i) : Soit E_λ un espace propre de f associé à la valeur propre λ.

$$\forall x \in E_\lambda,\ \ f(g(x)) = g(f(x)) = g(\lambda x) = \lambda g(x),$$

et donc $g(x) \in E_\lambda$; autrement dit, E_λ est stable par g.

(ii) : Soit $y \in \mathrm{Im}(f)$. Alors il existe $x \in E$ tel que $f(x) = y$, puis

$$g(y) = g(f(x)) = f(g(x)) \in \mathrm{Im}(f),$$

et donc $\mathrm{Im}(f)$ est stable par g. □

Théorème 2.19 (Diagonalisation simultanée)**.**
Soient $f, g \in \mathcal{L}(E)$. Si f et g commutent et sont diagonalisables, alors il existe une base commune de diagonalisation de f et g. (On dit alors que f et g sont *codiagonalisables.*)

Démonstration. Soient $\lambda_1, \ldots, \lambda_p$ les valeurs propres de f et $E_{\lambda_1}, \ldots, E_{\lambda_p}$ les espaces propres associés. D'après le lemme 2.18, chaque espace propre E_{λ_i} est stable par g. Puis, d'après la proposition 2.12, chaque endomorphisme $g_{|E_{\lambda_i}} \in \mathcal{L}(E_{\lambda_i})$ est diagonalisable. Ainsi, pour chaque $i \in \{1, \ldots, p\}$, il existe une base $\mathcal{B}_i$ de E_{λ_i} formée de vecteurs propres de g, et aussi de f puisque $f_{|E_{\lambda_i}} = \lambda_i \mathrm{Id}_{E_{\lambda_i}}$. Étant donnée que $E = \bigoplus_{i=1}^{p} E_{\lambda_i}$, on en déduit que $\mathcal{B} = \mathcal{B}_1 \cup \ldots \mathcal{B}_p$ est une base de E qui diagonalise simultanément f et g. □

Exemples 2.20.

(i) Les matrices $\begin{bmatrix} 1 & -1 \\ 0 & 2 \end{bmatrix}$ et $\begin{bmatrix} 3 & -1 \\ 0 & 4 \end{bmatrix}$ sont codiagonalisables. On laisse le soin au lecteur ou à la lectrice de trouver une base de diagonalisation commune pour ces deux matrices.

(ii) Soient $A = \begin{bmatrix} 3 & -1 & -1 \\ 1 & 1 & -1 \\ 1 & -1 & 1 \end{bmatrix}$ et $B = \begin{bmatrix} -1 & 2 & 2 \\ 0 & 3 & 0 \\ -4 & 2 & 5 \end{bmatrix}$ deux matrices de $\mathcal{M}_3(\mathbb{K})$. Nous allons montrer que A et B sont codiagonalisables et détailler le calcul d'une base de diagonalisation commune pour A et B (en suivant la stratégie de la démonstration du théorème 2.19).

On commence par calculer le polynôme caractéristique de A, on trouve

$$\chi_A(X) = -X^3 + 5X^2 - 8X + 4 = (2 - X)^2(1 - X).$$

La matrice A admet donc pour valeurs propres 1 et 2. On vérifie que les espaces propres correspondants sont

$$E_1 = \mathrm{Vect}(\begin{bmatrix}1\\1\\1\end{bmatrix}) \quad \text{et} \quad E_2 = \mathrm{Vect}(\begin{bmatrix}0\\1\\-1\end{bmatrix}, \begin{bmatrix}1\\0\\1\end{bmatrix}).$$

On a $\dim(E_1) + \dim(E_2) = 1 + 2 = 3$, donc A est diagonalisable.

De même, le calcul du polynôme caractéristique de B donne

$$\chi_B(X) = -X^3 + 7X^2 - 15X + 9 = (3 - X)^2(1 - X).$$

Et on vérifie que $\mu_B(X) = (3 - X)(1 - X)$, ce qui revient à dire que B est diagonalisable.

Ensuite, on vérifie que

$$AB = BA = \begin{bmatrix}1 & 1 & 1\\3 & 3 & -3\\-5 & 1 & 7\end{bmatrix}.$$

Donc, d'après le théorème 2.19, les matrices A et B sont codiagonalisables. De plus, en notant $u \in \mathcal{L}(\mathbb{K}^3)$ l'endomorphisme dont la matrice dans la base canonique de $\mathbb{K}^3$ est B, on a que les $\mathbb{K}$-sous-espaces vectoriels E_1 et E_2 sont stables par u (Lemme 2.18). Comme $\dim(E_1) = 1$, cela implique que E_1 est aussi un espace propre pour u, et donc pour B. On calcule

$$B\begin{bmatrix}1\\1\\1\end{bmatrix} = 3\begin{bmatrix}1\\1\\1\end{bmatrix}, \text{ et donc } \begin{bmatrix}1\\1\\1\end{bmatrix} \text{ est un vecteur propre de } B \text{ associé à la valeur propre 3.}$$

Par ailleurs, comme $u_{|E_2}$ est diagonalisable, le $\mathbb{K}$-sous-espace vectoriel E_2 se décompose comme $E_2 = L_1 \oplus L_3$, avec L_1 et L_3 deux droites B-stables sur lesquelles B agit par multiplication par 1 et 3 respectivement. On note $v_1 = \begin{bmatrix}0\\1\\-1\end{bmatrix}$ et $v_2 = \begin{bmatrix}1\\0\\1\end{bmatrix}$. Alors $E_2 = \mathrm{Vect}(v_1, v_2)$, et on cherche $\alpha_1, \alpha_2, \beta_1, \beta_2 \in \mathbb{K}$ tels que

$$\begin{cases} B(\alpha_1 v_1 + \alpha_2 v_2) = \alpha_1 v_1 + \alpha_2 v_2; \text{ et} \\ B(\beta_1 v_1 + \beta_2 v_2) = 3(\beta_1 v_1 + \beta_2 v_2). \end{cases}$$

La résolution de ces deux systèmes linéaires à deux inconnues donne

$$\alpha_1 = 0, \ \alpha_2 = 1, \ \beta_1 = 1, \ \text{et } \beta_2 = 0.$$

Autrement dit, on a

$$L_1 = \mathrm{Vect}(\begin{bmatrix}1\\0\\1\end{bmatrix}) \quad \text{et} \quad L_3 = \mathrm{Vect}(\begin{bmatrix}0\\1\\-1\end{bmatrix})$$

On en déduit qu'une base de diagonalisation commune pour A et B est donnée par

$$\mathcal{B} = \{\begin{bmatrix}1\\1\\1\end{bmatrix}, \begin{bmatrix}0\\1\\-1\end{bmatrix}, \begin{bmatrix}1\\0\\1\end{bmatrix}\}.$$

Et donc, si l'on note $P = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & -1 & 1 \end{bmatrix}$ la matrice de passage de la base canonique de $\mathbb{K}^3$ vers la base $\mathcal{B}$, on obtient

$$P^{-1}AP = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix} \quad \text{et} \quad P^{-1}BP = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Théorème 2.21 (Trigonalisation simultanée)**.**
Soient $f, g \in \mathcal{L}(E)$. Si f et g commutent et sont trigonalisables, alors il existe une base commune de trigonalisation de f et g. (On dit alors que f et g sont *cotrigonalisables*.)

Démonstration. On commence par une observation préliminaire : avec les hypothèses du théorème, f et g ont un vecteur propre commun. En effet, f est trigonalisable, donc f admet une valeur propre $\lambda \in \mathbb{K}$, et d'après le lemme 2.18 et la remarque 2.17 (ii), l'endomorphisme $g_{|E_\lambda} \in \mathcal{L}(E_\lambda)$ est trigonalisable, donc admet également un vecteur propre $v \in E_\lambda$ (associé à une certaine valeur propre λ'). Finalement, $v \in E$ est un vecteur propre commun à f et g.

Nous allons maintenant montrer le théorème par récurrence sur $n = \dim(E) \geq 1$. Si $n = 1$, alors il n'y a rien à prouver puisque tout endomorphisme est trigonalisable. On suppose le résultat vrai au rang $n-1 \geq 1$ et on va le montrer au rang n. Pour ce faire, on va utiliser à nouveau la dualité, comme dans la démonstration du théorème 2.16. Puisque $f \circ g = g \circ f$, le point (i) de la proposition 1.5 implique que ${}^t g \circ {}^t f = {}^t f \circ {}^t g$. Donc, d'après l'observation préliminaire appliquée à ${}^t f$ et ${}^t g$, il existe un vecteur propre $\epsilon_0 \in E^* \setminus \{0\}$ commun à ${}^t f$ et ${}^t g$. Ainsi, la droite vectorielle $\langle \epsilon_0 \rangle$ de E^* est stable par ${}^t f$ et ${}^t g$, et donc la proposition 1.25 implique que l'hyperplan $H := \langle \epsilon_0 \rangle^\circ$ de E est stable par f et g. D'après l'hypothèse de récurrence, les endomorphismes $f_{|H}$ et $g_{|H}$ de H sont trigonalisables dans une même base $\mathcal{B}_0$ de H. On complète $\mathcal{B}_0$ en une base $\mathcal{B}$ de E, alors les matrices

$$\mathrm{Mat}_{\mathcal{B}}(f) = \begin{bmatrix} & & & * \\ & \mathrm{Mat}_{\mathcal{B}_0}(f_{|H}) & & * \\ & & & * \\ 0 & \cdots & 0 & * \end{bmatrix} \quad \text{et} \quad \mathrm{Mat}_{\mathcal{B}}(g) = \begin{bmatrix} & & & * \\ & \mathrm{Mat}_{\mathcal{B}_0}(g_{|H}) & & * \\ & & & * \\ 0 & \cdots & 0 & * \end{bmatrix}$$

sont triangulaires supérieures, d'où le résultat. □

Remarque 2.22. Les théorèmes 2.19 et 2.21 se généralisent au cas de $r(\geq 1)$ endomorphismes de E qui sont diagonalisables et commutent deux à deux (c.f. le livre "Algèbre" de Gourdon, § 4.1, Exercice 4).

2.2 Polynômes d'endomorphismes

2.2.A Généralités

Soit $P \in \mathbb{K}[X]$ de degré $p \geq 0$. On note $P(X) = a_0 + a_1 X + \cdots + a_p X^p$, avec $a_p \neq 0$. Pour toute matrice $M \in \mathcal{M}_n(\mathbb{K})$, on définit

$$P(M) = a_0 I_n + a_1 M + \cdots + a_p M^p \in \mathcal{M}_n(\mathbb{K}),$$

et pour tout endomorphisme $f \in \mathcal{L}(E)$, on définit

$$P(f) = a_0 \mathrm{Id}_E + a_1 f + \cdots + a_p f^p \in \mathcal{L}(E).$$

On vérifie alors que, pour toute base $\mathcal{B}$ de E, on a $\mathrm{Mat}_{\mathcal{B}}(P(f)) = P(\mathrm{Mat}_{\mathcal{B}}(f))$. De plus, pour tous $P, Q \in \mathbb{K}[X]$ et pour tout $f \in \mathcal{L}(E)$, on a

$$(PQ)(f) = P(f) \circ Q(f) = Q(f) \circ P(f) = (QP)(f).$$

Lemme 2.23. Soient $P \in \mathbb{K}[X]$ et $f \in \mathcal{L}(E)$. Si $\lambda \in \mathbb{K}$ est une valeur propre de f, alors $P(\lambda)$ est une valeur propre de $P(f)$.

Démonstration. Par définition d'une valeur propre de f, il existe $v \in E \setminus \{0\}$ tel que $f(v) = \lambda v$. On montre alors par une récurrence facile que $f^r(v) = \lambda^r v$, pour tout $r \geq 0$. Il s'ensuit que

$$P(f)(v) = \sum_{i=0}^{p} a_i f^i(v) = \sum_{i=0}^{p} a_i \lambda^i v = \left(\sum_{i=0}^{p} a_i \lambda^i \right) v = P(\lambda) v,$$

et donc $P(\lambda)$ est une valeur propre de l'endomorphisme $P(f)$. $\square$

Soient $P \in \mathbb{K}[X]$ et $M \in \mathcal{M}_n(\mathbb{K})$. Si $M = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ est diagonale, alors on vérifie que $P(M) = \mathrm{diag}(P(\lambda_1), \ldots, P(\lambda_n))$ est diagonale. De même, si $M = \begin{bmatrix} \lambda_1 & * & * & * \\ 0 & \lambda_2 & * & * \\ 0 & 0 & \ddots & * \\ 0 & 0 & 0 & \lambda_n \end{bmatrix}$ est triangulaire supérieure, alors on vérifie que $P(M) = \begin{bmatrix} P(\lambda_1) & * & * & * \\ 0 & P(\lambda_2) & * & * \\ 0 & 0 & \ddots & * \\ 0 & 0 & 0 & P(\lambda_n) \end{bmatrix}$ est triangulaire supérieure.

Plus généralement, si $M \in \mathcal{M}_n(\mathbb{K})$ est diagonalisable resp. trigonalisable, alors P(M) est diagonalisable resp. trigonalisable. En revanche, il n'est généralement pas vrai que si M est nilpotente alors $P(M)$ est nilpotente (prendre par exemple pour P un polynôme constant non-nul), sauf si $P(0) = 0$.

2.2.B Polynôme minimal

On commence par rappeler un théorème classique déjà vu en cours l'an dernier :

Théorème 2.24 (Cayley-Hamilton)**.** Soient $f \in \mathcal{L}(E)$ et χ_f le polynôme caractéristique de f. Alors $\chi_f(f) = 0$.

Démonstration. C.f. le livre "Algèbre" de Gourdon (§ 4.2., théorème 3) ou votre cours de l'an dernier. $\square$

Définition-Proposition 2.25. Soit $f \in \mathcal{L}(E)$. Il existe un unique polynôme unitaire de degré minimal dans $\mathbb{K}[X]$, que l'on notera μ_f, tel que pour tout polynôme $P \in \mathbb{K}[X]$ vérifiant $P(f) = 0$, on a $\mu_f \mid P$. Le polynôme μ_f est appelé le *polynôme minimal de f*.

Démonstration. Le point-clé pour montrer l'existence de μ_f est le fait que $\mathbb{K}[X]$ est un *anneau euclidien*, c'est-à-dire que $\mathbb{K}[X]$ est muni d'une *division euclidienne* :

$$\forall (A,B) \in \mathbb{K}[X]^2,\ \exists!(Q,R) \in \mathbb{K}[X]^2 \text{ tel que } A = BQ + R \text{ avec } \deg(R) < \deg(B).$$

On considère l'ensemble

$$I := \{P \in \mathbb{K}[X], P(f) = 0\}.$$

L'ensemble I n'est pas réduit à $\{0\}$ car il contient χ_f d'après le théorème de Cayley-Hamilton. Soit T un polynôme non-nul de degré minimal contenu dans I. Un tel polynôme existe puisque l'ensemble $\{\deg(P),\ P \in I \setminus \{0\}\} \subseteq \mathbb{N}^*$ est minoré. Alors, pour tout polynôme $P \in I$, il existe un unique couple $(Q,R) \in \mathbb{K}[X]^2$ tel que $P = TQ + R$ avec $\deg(R) < \deg(T)$. En outre, $P, T \in I$, donc $R(f) = P(f) - (TQ)(f) = P(f) - Q(f) \circ T(f) = 0$, et donc $R \in I$. Mais $\deg(R) < \deg(T)$ et T est un polynôme non-nul de degré minimal contenu dans I, donc nécessairement $R = 0$. Autrement dit, T divise tout polynôme $P \in \mathbb{K}[X]$ vérifiant $P(f) = 0$.

Ensuite, quitte à multiplier T par une constante non-nulle, on peut supposer que T est unitaire. Enfin, si T_1, T_2 sont deux polynômes unitaires de degré minimal contenus dans I, alors par ce qui précède on a $T_1 \mid T_2$ et $T_2 \mid T_1$, et donc $T_1 = T_2$ puisque deux polynômes qui se divisent l'un l'autre sont égaux à une constante près et que T_1, T_2 sont unitaires. On a donc bien montré l'existence d'un unique polynôme unitaire de degré minimal dans $\mathbb{K}[X]$, que l'on notera dorénavant μ_f, tel que pour tout polynôme $P \in \mathbb{K}[X]$ vérifiant $P(f) = 0$, on a $\mu_f \mid P$. □

Proposition 2.26. Soit $f \in \mathcal{L}(E)$. Un scalaire $\lambda \in \mathbb{K}$ est valeur propre de f si et seulement si λ est racine de μ_f.

Démonstration. ($\Rightarrow$) : Soit $v \in E \setminus \{0\}$ un vecteur propre de f associé à la valeur propre λ. D'après le lemme 2.23, on a $\mu_f(f)(v) = \mu_f(\lambda)v$. Or $\mu_f(f) = 0$, par définition de μ_f, et donc $\mu_f(\lambda)(v) = 0$, d'où $\mu_f(\lambda) = 0$ puisque $v \neq 0$.

($\Leftarrow$) : Soit λ une racine de μ_f. Alors on peut factoriser μ_f par $(X - \lambda)$, autrement dit, il existe $P \in \mathbb{K}[X]$ tel que $\mu_f(X) = (X - \lambda)P(X)$. On remarque que $P(f) \neq 0$ étant donné que $\deg(P) < \deg(\mu_f)$ et que μ_f divise tous les polynômes annulateurs de f. Or

$$(f - \lambda \mathrm{Id}_E) \circ P(f) = \mu_f(f) = 0,$$

et donc l'endomorphisme $(f - \lambda \mathrm{Id}_E) \in \mathcal{L}(E)$ n'est pas inversible, sinon on pourrait multiplier à gauche par son inverse ce qui nous donnerait $P(f) = 0$. Il s'ensuit que $\mathrm{Ker}(f - \lambda \mathrm{Id}_E) \neq \{0\}$, autrement dit, λ est valeur propre de f. □

Remarque 2.27. Soit $f \in \mathcal{L}(E)$. Il découle des propositions 2.2 et 2.26 que les polynômes χ_f et μ_f ont les mêmes racines (mais pas nécessairement avec les mêmes multiplicités). En outre, le théorème de Cayley-Hamilton implique que $\mu_f \mid \chi_f$, et donc que la multiplicité de λ comme racine de μ_f est inférieure ou égale à la multiplicité de λ comme racine de χ_f.

2.2.C Lemme des noyaux et applications

Le prochain résultat est fondamental en algèbre linéaire. Il intervient en particulier dans la preuve de la décomposition de Dunford et dans la preuve de la réduction de Jordan.

Théorème 2.28 (Lemme des noyaux)**.** Soient $f \in \mathcal{L}(E)$ et $P = P_1 \cdots P_r \in \mathbb{K}[X]$, où les polynômes $P_i \in \mathbb{K}[X]$ sont premiers entre eux deux à deux. Alors

$$\operatorname{Ker}(P(f)) = \operatorname{Ker}(P_1(f)) \oplus \cdots \oplus \operatorname{Ker}(P_r(f)).$$

Démonstration. La preuve se fait par récurrence sur $r \geq 1$ en utilisant une identité de Bézout dans $\mathbb{K}[X]$ (d'abord pour traiter le cas $r = 2$, puis pour faire l'hérédité). On renvoie l'étudiant.e à son cours de l'an dernier ou au livre "Algèbre" de Gourdon (§ 4.2.1, théorème 1) pour la preuve. □

Le résultat qui suit donne un critère pour savoir si un endomorphisme dont on connaît le polynôme minimal est diagonalisable (à comparer avec le théorème 2.10).

Théorème 2.29. Soit $f \in \mathcal{L}(E)$. Les assertions suivantes sont équivalentes :
(i) f est diagonalisable ;
(ii) il existe $P \in \mathbb{K}[X]$ scindé à racines simples tel que $P(f) = 0$; et
(iii) μ_f est scindé à racines simples.

Démonstration. L'équivalence entre les points (ii) et (iii) est claire par définition du polynôme minimal μ_f (voir la définition-proposition 2.25).

(i) $\Rightarrow$ (ii) : Soient $\lambda_1, \ldots, \lambda_p \in \mathbb{K}$ les valeurs propres de f et $E_{\lambda_1}, \ldots, E_{\lambda_p}$ les espaces propres de f correspondants. On pose $P(X) = \prod_{i=1}^{p}(X - \lambda_i) \in \mathbb{K}[X]$ qui est un polynôme scindé à racines simples. D'après le lemme des noyaux, on a

$$\operatorname{Ker}(P(f)) = \bigoplus_{i=1}^{p} \operatorname{Ker}(f - \lambda_i \mathrm{Id}_E) = \bigoplus_{i=1}^{p} E_{\lambda_i} = E,$$

où la dernière égalité découle du fait que f est diagonalisable par hypothèse. On a donc $\operatorname{Ker}(P(f)) = E$, c'est-à-dire $P(f) = 0$, et donc il existe $P \in \mathbb{K}[X]$ scindé à racines simples tel que $P(f) = 0$.

(ii) $\Rightarrow$ (i) : Soit $P \in \mathbb{K}[X]$ scindé à racines simples tel que $P(f) = 0$. On écrit $P(X) = \alpha \prod_{i=1}^{p}(X - \lambda_i)$ avec les $\lambda_i \in \mathbb{K}$ deux à deux distincts et $\alpha \in \mathbb{K}^*$. D'après le lemme des noyaux, on a

$$E = \operatorname{Ker}(P(f)) = \bigoplus_{i=1}^{p} \operatorname{Ker}(f - \lambda_i \mathrm{Id}_E),$$

où la première égalité découle du fait que $P(f)$ est l'endomorphisme nul. Ensuite, on note $S = \{i \in \{1, \ldots, p\},\ \dim(\operatorname{Ker}(f - \lambda_i \mathrm{Id}_E)) \geq 1\}$. Pour tout $i \in S$, λ_i est valeur propre de f et $\operatorname{Ker}(f - \lambda_i \mathrm{Id}_E)$ est l'espace propre associé. On déduit donc de l'égalité ci-dessus que $E = \bigoplus_{i \in S} E_{\lambda_i}$, c'est-à-dire que f est diagonalisable. □

Remarque 2.30. On est maintenant en mesure de prouver la proposition 2.12. En effet, si f est diagonalisable, alors μ_f est scindé à racines simples (implication (i) $\Rightarrow$ (iii) du théorème 2.29). Or $\mu_f(f_{|F}) = \mu_f(f)_{|F} = 0$, donc il existe $P \in \mathbb{K}[X]$ scindé à racines simples tel que $P(f_{|F}) = 0$ (prendre $P = \mu_f$), et donc $f_{|F}$ est diagonalisable (implication (ii) $\Rightarrow$ (i) du théorème 2.29).

Exemples 2.31.
(i) Les *projecteurs* sont définis comme les endomorphismes $p \in \mathcal{L}(E)$ tels que $p \circ p = p$. En particulier, ils sont annulés par $P(X) = X^2 - X = X(X - 1)$ qui est scindé à racines simples. Il s'ensuit que les projecteurs sont toujours diagonalisables et à valeurs propres dans $\{0, 1\}$.

(ii) Les *symétries* sont définis comme les endomorphismes $s \in \mathcal{L}(E)$ tels que $s \circ s = \mathrm{Id}_E$. En particulier, elles sont annulés par $P(X) = X^2 - 1 = (X-1)(X+1)$ qui est scindé à racines simples. Il s'ensuit que les symétries sont toujours diagonalisables et à valeurs propres dans $\{-1, 1\}$. Voici trois exemples de symétries :

$$\begin{aligned}
&\mathcal{M}_n(\mathbb{K}) \to \mathcal{M}_n(\mathbb{K}), \quad M \mapsto {}^tM \\
&\mathbb{K}[X]_{\geq n} \to \mathbb{K}[X]_{\geq n}, \quad P(X) \mapsto P(1-X) \\
&\mathbb{K}[X]_{\geq n} \to \mathbb{K}[X]_{\geq n}, \quad P(X) \mapsto X^n P(\frac{1}{X})
\end{aligned}$$

Enfin, le résultat qui suit donne un critère pour savoir si un endomorphisme dont on connaît le polynôme minimal est trigonalisable (à comparer avec le théorème 2.16).

Théorème 2.32. Soit $f \in \mathcal{L}(E)$. Les assertions suivantes sont équivalentes :
(i) f est trigonalisable ;
(ii) il existe $P \in \mathbb{K}[X]$ scindé tel que $P(f) = 0$; et
(iii) μ_f est scindé.

Démonstration. La preuve est similaire à celle du théorème 2.29 et est laissée en exercice. □

2.2.D Sous-espaces caractéristiques et décomposition de Dunford

Soit $f \in \mathcal{L}(E)$ dont le polynôme caractéristique est scindé. On écrit

$$\chi_f(X) = (-1)^n \prod_{i=1}^{p} (X - \lambda_i)^{\alpha_i},$$

où les $\lambda_i \in \mathbb{K}$ sont deux à deux distincts et chaque $\alpha_i \geq 1$. On définit *l'espace caractéristique* de f associé à la valeur propre λ_i par

$$N_i := \mathrm{Ker}\left((f - \lambda_i \mathrm{Id}_E)^{\alpha_i}\right)$$

qui est un sous-espace vectoriel de E contenant l'espace propre E_{λ_i}.

Proposition 2.33. Les espaces caractéristiques de $f \in \mathcal{L}(E)$ vérifient les propriétés suivantes :
- Pour tout $i \in \{1, \ldots, p\}$, l'espace caractéristique N_i est stable par f.
- On a la décomposition en somme directe $E = \bigoplus_{i=1}^{p} N_i$.
- Pour tout $i \in \{1, \ldots, p\}$, on a $\dim(N_i) = \alpha_i$.

Démonstration. La preuve est essentiellement une conséquence du théorème de Cayley-Hamilton et du lemme des noyaux appliqué à χ_f. On renvoie l'étudiant.e à son cours de l'an dernier ou au livre "Algèbre" de Gourdon (§ 4.4.1, définition 1) pour les détails de la démonstration. □

Une autre conséquence du lemme des noyaux et de la décomposition en espaces caractéristiques est la fameuse décomposition de Dunford, prouvée en cours l'an dernier, dont on rappelle ci-dessous l'énoncé afin d'étudier ultérieurement le lien de celle-ci avec la réduction de Jordan dans la section 2.4.

Théorème 2.34 (Décomposition de Dunford)**.** Soit $f \in \mathcal{L}(E)$ **tel que** χ_f **est scindé**. Il existe un unique couple $(d, n) \in \mathcal{L}(E)^2$, avec d diagonalisable et n nilpotent, tels que

(i) $f = d + n$; et
(ii) $d \circ n = n \circ d$.
De plus, d et n sont des polynômes en f, c'est-à-dire qu'il existe $P_1, P_2 \in \mathbb{K}[X]$ tels que $d = P_1(f)$ et $n = P_2(f)$.

Démonstration. C.f. le livre "Algèbre" de Gourdon (§ 4.4.2, théorème 2 et théorème 3) ou votre cours de l'an dernier. □

Exemple 2.35. La décomposition de Dunford de

$$\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \text{ est } \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix},$$

mais la décomposition de Dunford de

$$\begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix} \text{ est } \begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

Je tiens à insister sur le fait qu'il s'agit là d'une *décomposition*, et non pas d'une *réduction* au sens où cet énoncé ne donne pas l'existence d'une certaine base $\mathcal{B}$ de E telle que la matrice $\mathrm{Mat}_{\mathcal{B}}(f)$ ait une forme "sympathique". Néanmoins, si l'on travaille avec des matrices, la décomposition de Dunford permet tout de même d'écrire une matrice donnée $M \in \mathcal{M}_n(\mathbb{K})$, dont le polynôme caractéristique est scindé, comme une somme de deux matrices $M = D + N$ avec D une matrice diagonalisable (mais pas nécessairement diagonale) et N une matrice nilpotente telles que $ND = DN$ (ce qui est commode lorsque l'on souhaite par exemple appliquer la formule du binôme de Newton afin de calculer $M^r = (D+N)^r$ avec $r \in \mathbb{N}^*$). On renvoie au cours de l'an dernier ou bien au livre "Algèbre" de Gourdon (§ 4.4.2) pour le calcul pratique de la décomposition de Dunford ainsi que son application au calcul d'exponentielle.

2.3 Réduction de Frobenius

Dans cette section on prouve l'existence de la réduction de Frobenius (qui doit son nom au mathématicien allemand Ferdinand Georg Frobenius, 1849-1917) pour n'importe quel endomorphisme $f \in \mathcal{L}(E)$, sans aucune hypothèse sur le polynôme caractéristique de f.

2.3.A Endomorphismes cycliques

Dans toute cette section on fixe $f \in \mathcal{L}(E)$. On commence par introduire quelques notations dont on aura besoin par la suite.

Définition 2.36. On note :
- $I = \{P \in \mathbb{K}[X] \mid P(f) = 0\} \subseteq \mathbb{K}[X]$;
- $\mathcal{L} := \{P(f),\ P \in \mathbb{K}[X]\} \subseteq \mathcal{L}(E)$;
- $\forall x \in E,\ I_x := \{P \in \mathbb{K}[X] \mid P(f)(x) = 0\} \subseteq \mathbb{K}[X]$; et
- $\forall x \in E,\ E_x := \{P(f)(x),\ P \in \mathbb{K}[X]\} = \mathrm{Vect}(f^i(x),\ i \in \mathbb{N}) \subseteq E$.

De plus, pour tout $x \in E$, on montre (comme dans la démonstration de la définition-proposition 2.25) qu'il existe un unique polynôme unitaire de degré minimal dans $\mathbb{K}[X]$, que l'on notera $\mu_{f,x}$, tel que pour tout polynôme $P \in \mathbb{K}[X]$ vérifiant $P(f)(x) = 0$, on a $\mu_{f,x} \mid P$. Le polynôme $\mu_{f,x}$ est appelé le *polynôme minimal ponctuel de f en x*.

Soit $x \in E$. Remarquons que, puisque $\mu_f(f) = 0$, on a $\mu_f(f)(x) = 0$, et donc $\mu_{f,x} \mid \mu_f$. Nous verrons dans la preuve de la proposition 2.38 que, pour un élément $x \in E$ générique, on a en fait l'égalité.

La proposition qui suit, dont la preuve fait appel à la notion d'espace vectoriel quotient (introduite dans la section 1.2), nous sera utile pour démontrer le théorème 2.44 qui est le résultat clé pour prouver l'existence et l'unicité d'une forme normale de Frobenius dans la section 2.3.C.

Proposition 2.37. (Dans toute la suite on garde les notations de la définition 2.36.)

(i) On note $k := \deg(\mu_f)$. L'ensemble $\mathcal{L}$ est un sous-espace vectoriel de $\mathcal{L}(E)$, de dimension k, dont une base est donnée par $\{\mathrm{Id}_E, f, f^2, \ldots, f^{k-1}\}$.

(ii) On note $l := \deg(\mu_{f,x})$. L'ensemble E_x est un sous-espace vectoriel de E, de dimension l, dont une base est donnée par $\{x, f(x), f^2(x), \ldots, f^{l-1}(x)\}$.

Démonstration. (i) : On considère l'application linéaire

$$\psi\colon \mathbb{K}[X] \to \mathcal{L}(E),\ P(X) \mapsto P(f).$$

On a $\mathrm{Im}(\psi) = \mathcal{L}$, par définition de $\mathcal{L}$, et $\mathrm{Ker}(\psi) = \{P \in \mathbb{K}[X] \mid P(f) = 0\} = I$. D'après la proposition 1.36, ψ induit un isomorphisme d'espaces vectoriels

$$\overline{\psi}\colon \mathbb{K}[X]/I \simeq \mathcal{L},\ \overline{P(X)} \mapsto P(f), \quad \text{où l'on note } \overline{P(X)} \text{ l'image de } P(X) \text{ dans } \mathbb{K}[X]/I.$$

Étant donnée qu'une base de l'espace vectoriel quotient $\mathbb{K}[X]/I$ est donnée par $\{\overline{1}, \overline{X}, \ldots, \overline{X}^{k-1}\}$, avec $\overline{X}$ l'image de X dans le quotient $\mathbb{K}[X]/I$, on en déduit que $\dim(\mathcal{L}) = \dim(\mathbb{K}[X]/I) = k$. Enfin, puisque l'image d'une base de $\mathbb{K}[X]/I$ par l'isomorphisme $\overline{\psi}$ est une base de $\mathcal{L}$, on en déduit que $\{\mathrm{Id}_E, f, f^2, \ldots, f^{k-1}\}$ est une base de $\mathcal{L}$.

(ii) : La preuve de (ii) est analogue à celle de (i) en considérant l'application linéaire

$$\psi_x\colon \mathbb{K}[X] \to E,\ P \mapsto P(f)(x)$$

à la place de ψ. On laisse au lecteur ou à la lectrice le soin de rédiger les détails de la démonstration. □

Proposition 2.38. Il existe $x \in E$ tel que $\mu_{f,x} = \mu_f$.

Démonstration. Soient $Q_1, \ldots, Q_r \in \mathbb{K}[X]$ les polynômes unitaires non-constants qui divisent μ_f et tels que $Q_i \neq \mu_f$. Alors, pour tout $i \in \{1, \ldots, r\}$, $\mathrm{Ker}(Q_i(f))$ est un sous-espace vectoriel *strict* de E (c'est-à-dire distinct de E) sinon on aurait $E = \mathrm{Ker}(Q_i(f))$, autrement dit, $Q_i(f) = 0$, ce qui est impossible puisque Q_i est un diviseur strict de μ_f. On en déduit que $\bigcup_{i=1}^r \mathrm{Ker}(Q_i(f)) \subsetneq E$.

Par ailleurs, on a déjà remarqué que, pour tout $x \in E$, le polynôme unitaire $\mu_{f,x}$ divise μ_f, et donc on a l'alternative suivante : $\mu_{f,x} = \mu_f$ ou bien $\mu_{f,x} = Q_i$ pour un certain $i \in \{1, \ldots, r\}$. Mais $\mu_{f,x} = Q_i$ implique $Q_i(f)(x) = \mu_{f,x}(f)(x) = 0$, c'est-à-dire $x \in \mathrm{Ker}(Q_i(f))$. On en déduit, par la contraposée, que si $x \notin \bigcup_{i=1}^r \mathrm{Ker}(Q_i(f))$, alors $\mu_{f,x} = \mu_f$. Ainsi, il suffit de choisir $x \in E \setminus (\bigcup_{i=1}^r \mathrm{Ker}(Q_i(f)))$ pour avoir $\mu_{f,x} = \mu_f$. □

Remarque 2.39. Notons que dans la preuve de la proposition 2.38, on utilise implicitement le *lemme d'évitement* qui affirme qu'un espace vectoriel n'est pas égal à une réunion finie de sous-espaces vectoriels stricts dès lors que le corps de base $\mathbb{K}$ est infini. (N'hésitez pas à venir me voir si vous voulez que je vous en donne la démonstration ; celle-ci étant assez élémentaire.)

Définition 2.40. On dit que $f \in \mathcal{L}(E)$ est *cyclique* s'il existe $x \in E$ tel que $E_x = E$. Dans ce cas, on a $E = \mathrm{Vect}(f^i(x),\ i \in \mathbb{N}) = \mathrm{Vect}(f^i(x),\ 0 \le i \le n)$, où la dernière égalité découle de la proposition 2.37 (ii).

Étant donné que, pour tout $x \in E$, on a

- $\mu_{f,x} \mid \mu_f \mid \chi_f$;
- $\dim(E) = n = \deg(\chi_f)$; et
- $\dim(E_x) = \deg(\mu_{f,x})$;

on en déduit qu'un endomorphisme $f \in \mathcal{L}(E)$ est cyclique si et seulement s'il existe $x \in E$ tel que

$$\mu_{f,x} = \mu_f = (-1)^n \chi_f.$$

Définition 2.41. Soit $P \in \mathbb{K}[X]$ un polynôme unitaire de degré $p \ge 1$, que l'on note $P(X) = X^p + a_{p-1}X^{p-1} + \cdots + a_0$. La *matrice compagnon* du polynôme P est la matrice

$$\mathcal{C}(P) := \begin{bmatrix} 0 & \cdots & 0 & -a_0 \\ 1 & \ddots & \vdots & \vdots \\ & \ddots & 0 & -a_{p-2} \\ & (0) & 1 & -a_{p-1} \end{bmatrix} \in \mathcal{M}_p(\mathbb{K}).$$

Proposition 2.42. Soient $P \in \mathbb{K}[X]$ un polynôme unitaire, de degré $p \ge 1$, et $\mathcal{C}(P)$ sa matrice compagnon. Alors $\chi_{\mathcal{C}(P)} = (-1)^p P$.

Démonstration. La preuve se fait par récurrence sur $p = \deg(P) \ge 1$. Si $p = 1$, alors $P(X) = X + a_0$ et $\mathcal{C}(P) = \begin{bmatrix} -a_0 \end{bmatrix}$. On a donc bien $\chi_{\mathcal{C}(P)}(X) = (-X - a_0) = (-1)^1 P(X)$.

On suppose maintenant que la propriété est vraie au rang $p-1 \ge 1$ et on va la montrer au rang p. En développant $\chi_{\mathcal{C}(P)}(X) = \det(\mathcal{C}(P) - XI_p)$ par rapport à la première ligne, il vient

$$\begin{aligned}
\chi_{\mathcal{C}(P)}(X) &= \det\left(\begin{bmatrix} -X & \cdots & (0) & -a_0 \\ 1 & \ddots & \vdots & \vdots \\ & \ddots & -X & -a_{p-2} \\ & (0) & 1 & -X - a_{p-1} \end{bmatrix}\right) \\
&= -X \det\left(\begin{bmatrix} -X & \cdots & (0) & -a_1 \\ 1 & \ddots & \vdots & \vdots \\ & \ddots & -X & -a_{p-2} \\ & (0) & 1 & -X - a_{p-1} \end{bmatrix}\right) + (-1)^p a_0 \det\left(\begin{bmatrix} 1 & -X & & (0) \\ 0 & 1 & \ddots & \\ \vdots & \ddots & \ddots & -X \\ 0 & \cdots & 0 & 1 \end{bmatrix}\right) \\
&= -X \times \chi_{\mathcal{C}(Q)} + (-1)^p a_0.
\end{aligned}$$

avec $Q(X) = X^{p-1} + a_{p-1}X^{p-2} + \cdots + a_1$, qui est unitaire de degré $p-1$. Et donc, en appliquant l'hypothèse de récurrence au polynôme Q, on obtient

$$\chi_{\mathcal{C}(P)}(X) = -X(-1)^{p-1}Q(X) + (-1)^p a_0 = (-1)^p(X(Q(X) + a_0) = (-1)^p P(X),$$

ce qui termine la preuve de la proposition. $\square$

Le résultat qui suit donne une caractérisation des endomorphismes cycliques.

Proposition 2.43. Un endomorphisme $f \in \mathcal{L}(E)$ est cyclique si et seulement s'il existe une base $\mathcal{B}$ de E telle que $\mathrm{Mat}_{\mathcal{B}}(f) = \mathcal{C}(\mu_f)$.

Démonstration. On note $\mu_f(X) = X^p + a_{p-1}X^{p-1} + \cdots + a_0$ le polynôme minimal de f.
$(\Rightarrow)$: Si $f \in \mathcal{L}(E)$ est cyclique, alors $p = n$, il existe $x \in E$ tel que $E_x = E$, et $\mathcal{B} := \{x, f(x), \ldots, f^{n-1}(x)\}$ est une base de E_x d'après la proposition 2.37 (ii). Dans cette base, on a bien $\mathrm{Mat}_{\mathcal{B}}(f) = \mathcal{C}(\mu_f)$ puisque $f(f^i(x)) = f^{i+1}(x)$, pour tout $i \in \{0, \ldots, n-2\}$, et $f(f^{n-1}(x)) = f^n(x) = -a_{p-1}f^{p-1}(x) - \cdots - a_1 f(x) - a_0 x$ car $\mu_f(f)(x) = 0$.
$(\Leftarrow)$: On suppose maintenant qu'il existe une base $\mathcal{B} := \{e_1, \ldots, e_n\}$ de E telle que $\mathrm{Mat}_{\mathcal{B}}(f)$ soit une matrice compagnon. Alors $f(e_i) = e_{i+1}$, pour tout $i \in \{1, \ldots, n-1\}$, et donc si l'on note $x := e_1$, on voit que $\mathcal{B} = \{x, f(x), \ldots, f^{n-1}(x)\}$, c'est-à-dire que f est cyclique. $\square$

2.3.B Invariants de similitude

(Les principales références que j'ai utilisées pour écrire cette section et la section suivante sont l'annexe B du livre "Algèbre" de Gourdon et le chapitre 6.5 du livre Objectif Agrégation" de Beck-Malick-Peyré.)

Le résultat qui suit assure l'existence d'une décomposition de n'importe quel endomorphisme $f \in \mathcal{L}(E)$ en une somme d'endomorphismes cycliques. Cette décomposition n'est pas unique, mais nous allons voir que l'on peut lui associer une unique suite de polynômes unitaires non-constants qui vérifient certaines relations de divisibilité.

Théorème 2.44. Soit $f \in \mathcal{L}(E)$. Il existe des sous-espaces vectoriels $F_1, \ldots, F_r$ de E, tous stables par f, tels que :
(i) $E = F_1 \oplus \cdots \oplus F_r$;
(ii) pour tout $i \in \{1, \ldots, r\}$, la restriction $f_i := f_{|F_i} \in \mathcal{L}(F_i)$ est cyclique ; et
(iii) en notant $P_i := \mu_{f_i}$, on a $P_{i+1} \mid P_i$, pour tout $i \in \{1, \ldots, r-1\}$.
De plus, l'entier $r \geq 1$ et la suite des polynômes $P_1, \ldots, P_r$ ne dépend que de f, et non du choix des F_i (qui eux ne sont pas uniques en général), et on a $P_1 = \mu_f$.

Démonstration. On va procéder en deux temps et montrer d'abord l'existence (par un argument de récurrence) de sous-espaces vectoriels $F_1, \ldots, F_r$ de E, tous stables par f et vérifiant les points (i)-(ii)-(iii) ci dessus, puis l'unicité de l'entier $r \geq 1$ et de la suite des polynômes $P_1, \ldots, P_r$.
Existence : Soit $k := \deg(\mu_f)$ et soit $x_0 \in E$ tels que $\mu_{f,x_0} = \mu_f$ (un tel x_0 existe toujours d'après la proposition 2.38). Le sous-espace vectoriel $F_1 := E_{x_0}$ est de dimension k (proposition 2.37 (i)) et il est stable par f par définition de E_{x_0} (définition 2.36). Comme $\deg(\mu_{f,x_0}) = \deg(\mu_f) = k$, une base de $F_1 = E_{x_0}$ est donnée par

$$\mathcal{B}_1 := \{e_i,\ 1 \leq i \leq k\}, \quad \text{avec } e_i := f^{i-1}(x_0).$$

On complète alors $\mathcal{B}_1$ en une base $\mathcal{B} := \{e_1, \ldots, e_n\}$ de E, et on désigne par $\mathcal{B}^* := \{e_1^*, \ldots, e_n^*\}$ la base duale associée. On définit alors

$$\Gamma = \{{}^t f^i(e_k^*) \mid i \in \mathbb{N}\} = \{e_k^* \circ f^i,\ i \in \mathbb{N}\} \subseteq E^*$$

et son orthogonal (au sens de la dualité)

$$G := \Gamma^\circ = \{x \in E \mid \forall i \in \mathbb{N},\ e_k^* \circ f^i(x) = 0\}$$

qui est un sous-espace vectoriel de E dont on vérifie aisément qu'il est stable par f. Nous allons montrer que $E = F_1 \oplus G$, et pour cela, nous prouvons successivement $F_1 \cap G = \{0\}$ et $\dim(F_1) + \dim(G) = n$ ($\Leftrightarrow \dim(G) = n - k$).

- Soit $y \in F_1 \cap G$. Si $y \neq 0$, alors on peut écrire $y = \sum_{i=1}^{p} a_i e_i$ avec $p \leq k$ et $a_p \neq 0$ (car $y \in F_1$). En composant par ${}^t f^{k-p}(e_k^*) = e_k^* \circ f^{k-p}$, et comme $y \in G = \Gamma^\circ$, on obtient

$$0 = e_k^* \circ f^{k-p}(y) = e_k^* \circ f^{k-p}\left(\sum_{i=1}^{p} a_i e_i\right) = e_k^*\left(\sum_{i=1}^{p} a_i e_{k-p+i}\right) = \sum_{i=1}^{p} a_i e_k^*\left(e_{k-p+i}\right) = a_p \neq 0,$$

ce qui est absurde. Il s'ensuit que $F_1 \cap G = \{0\}$.
- Étant donné que $G = \Gamma^\circ = \mathrm{Vect}(\Gamma)^\circ$, on a $\dim(G) = n - k$ si et seulement si $\dim(\mathrm{Vect}(\Gamma)) = k$ (proposition 1.22). On considère alors l'application linéaire

$$\varphi\colon \mathcal{L} = \{P(f) \mid P \in \mathbb{K}[X]\} \to \mathrm{Vect}(\Gamma),\ \ P(f) \mapsto e_k^* \circ P(f).$$

Par définition de $\mathrm{Vect}(\Gamma)$, l'application φ est surjective. De plus, φ est également injective. En effet, si $e_k^* \circ P(f) = 0$ avec $P(f) \neq 0$, on peut écrire $P(f) = \sum_{i=0}^{p} a_i f^i \in \mathcal{L}$ avec $p \leq k-1$ et $a_p \neq 0$, et alors

$$0 = e_k^* \circ P(f)(f^{k-p-1}(x_0)) = e_k^*\left(\sum_{i=0}^{p} a_i f^{i+k-p-1}(x_0)\right) = e_k^*\left(\sum_{i=0}^{p} a_i e_{i+k-p}\right) = \sum_{i=0}^{p} a_i e_k^*\left(e_{i+k-p}\right) = a_p \neq 0$$

ce qui est absurde. Il s'ensuit que φ est un isomorphisme, et donc $\dim(\mathrm{Vect}(\Gamma)) = \dim(\mathcal{L}) = \deg(\mu_{f,x_0}) = k$ (proposition 2.37).

On a donc trouvé un sous-espace vectoriel $G \subseteq E$, stable par f, tel que $E = F_1 \oplus G$ et $\dim(G) < \dim(E)$ (puisque $\dim(F_1) = \deg(\mu_f) \geq 1$). On note $P_1 := \mu_{f_{|F_1}}$ et $P_2 := \mu_{f_{|G}}$. Alors

$$\mu_f = \mu_{f,x_0} \mid \mu_{f_{|F_1}} \mid \mu_f, \quad \text{et donc} \quad P_1 := \mu_{f_{|F_1}} = \mu_f \ \text{ et } \ P_2 := \mu_{f_{|G}} \mid \mu_f = P_1.$$

On peut maintenant appliquer à nouveau ce qui précède en remplaçant le couple (E, f) par le couple $(G, f_{|G})$ et on obtient par récurrence l'existence de sous-espaces vectoriels $F_1, \ldots, F_r$ de E, tous stables par f et vérifiant les conditions (i)-(ii)-(iii) de l'énoncé.

Unicité : On suppose l'existence de deux suites de sous-espaces vectoriels $F_1, \ldots, F_r$ et $G_1, \ldots, G_s$ de E, tous stables par f, et vérifiant les conditions (i)-(ii)-(iii) de l'énoncé. On note

$$P_i := \mu_{f_{|F_i}} \quad \text{et} \quad Q_j := \mu_{f_{|G_j}}.$$

Déjà la condition (iii) implique que $P_1 = \mu_f = Q_1$. Supposons que la liste $(P_1, \ldots, P_r)$ est différente de la liste $(Q_1, \ldots, Q_s)$ et notons $j \geq 2$ le premier indice tel que $P_j \neq Q_j$; un tel indice existe toujours, même si $r \neq s$, car

$$\sum_{i=1}^{r} \deg(P_i) = \sum_{i=1}^{r} \dim(F_i) = n = \sum_{i=1}^{s} \dim(G_i) = \sum_{i=1}^{s} \deg(Q_j).$$

L'égalité $E = F_1 \oplus \cdots \oplus F_r$ avec les F_i stables par f, et la propriété $P_j(f)(F_k) = 0$ pour $k \geq j$, entraînent

$$P_j(f)(E) = P_j(f)(F_1) \oplus \cdots \oplus P_j(f)(F_{j-1}), \tag{2.1}$$

et par ailleurs l'égalité $E = G_1 \oplus \cdots \oplus G_s$ avec les G_i stables par f entraîne

$$P_j(f)(E) = P_j(f)(G_1) \oplus \cdots \oplus P_j(f)(G_{j-1}) \oplus P_j(f)(G_j) \oplus \cdots \oplus P_j(f)(G_s). \tag{2.2}$$

Or, on peut vérifier que l'on a $\dim(P_j(f)(F_i)) = \dim(P_j(f)(G_i))$ pour $1 \le i \le j-1$. En passant aux dimensions dans (2.1) et (2.2), on obtient

$$0 = \dim(P_j(f)(G_j)) = \cdots = \dim(P_j(f)(G_s)),$$

ce qui prouve que $Q_j \mid P_j$. Par symétrie du rôle des P_i et des Q_i, on a aussi $P_j \mid Q_j$, et donc $P_j = Q_j$ puisque ce sont des polynômes unitaires. Mais alors on aboutit à une absurdité puisque j est le premier indice tel que $P_j \neq Q_j$. On en déduit que nécessairement $r = s$ et $P_i = Q_i$ pour tout $i \in \{1, \dots, r\}$. Ainsi, on a l'unicité de l'entier $r \ge 1$ et de la suite des polynômes $P_1, \dots, P_r$. □

Définition 2.45. Étant donné $f \in \mathcal{L}(E)$, on appelle la suite de polynômes $P_1, \dots, P_r$, donnée par le théorème 2.44, les *invariants de similitude de* f (la terminologie deviendra claire dans la section 2.3.C). En particulier, le premier invariant de similitude de f est $P_1 = \mu_f$.

2.3.C Réduction de Frobenius

Théorème 2.46 (Réduction de Frobenius)**.** Soient $f \in \mathcal{L}(E)$ et $(P_1, \dots, P_r)$ la suite des invariants de similitude de f. Alors :

- Il existe une base $\mathcal{B}$ de E telle que

$$\mathrm{Mat}_{\mathcal{B}}(f) = \begin{bmatrix} \mathcal{C}(P_1) & & & \\ & \mathcal{C}(P_2) & & (0) \\ & & \ddots & \\ & (0) & & \mathcal{C}(P_r) \end{bmatrix},$$

  où chaque $\mathcal{C}(P_i)$ est la matrice compagnon associée au polynôme P_i.
- De plus, on a l'unicité de cette réduction au sens suivant : Si $(Q_1, \dots, Q_s)$ est une suite de polynômes unitaires non-constants, avec $Q_s \mid \cdots \mid Q_2 \mid Q_1$, et qu'il existe une base $\mathcal{B}'$ telle que

$$\mathrm{Mat}_{\mathcal{B}'}(f) = \begin{bmatrix} \mathcal{C}(Q_1) & & & \\ & \mathcal{C}(Q_2) & & (0) \\ & & \ddots & \\ & (0) & & \mathcal{C}(Q_s) \end{bmatrix},$$

  alors $r = s$ et $Q_i = P_i$ pour tout $i \in \{1, \dots, r\}$.
- Enfin, on a $P_1 = \mu_f$ et $P_1 \dots P_r = (-1)^n \chi_f$.

Démonstration. Avec les notation du théorème 2.44, il suffit pour tout $i \in \{1, \dots, r\}$ de considérer une base $\mathcal{B}_i$ de F_i dans laquelle $\mathrm{Mat}_{\mathcal{B}_i}(f_{|F_i}) = \mathcal{C}(P_i)$, ce qui est possible d'après la proposition 2.43, puis de prendre pour $\mathcal{B}$ la base définie par $\mathcal{B} := \mathcal{B}_1 \cup \cdots \cup \mathcal{B}_r$. De plus, la partie unicité du théorème se déduit de la partie unicité du théorème 2.44 et la justification détaillée de cette partie est laissée en exercice.

Il reste à vérifier la dernière assertion de l'énoncé. On a déjà vu que $P_1 = \mu_f$. Puis, comme $\chi_{\mathcal{C}(P_i)} = (-1)^{\deg(P_i)} P_i$, un calcul par blocs donne

$$\chi_f = \chi_{\mathrm{Mat}_{\mathcal{B}}(f)} = \prod_{i=1}^{r} \chi_{\mathcal{C}(P_i)} = (-1)^n \prod_{i=1}^{r} P_i,$$

et donc $P_1 \dots P_r = (-1)^n \chi_f$. □

Une matrice de la forme $\begin{bmatrix} \mathcal{C}(P_1) & & & \\ & \mathcal{C}(P_2) & & (0) \\ & & \ddots & \\ & (0) & & \mathcal{C}(P_r) \end{bmatrix}$, où les P_i sont des polynômes unitaires tels que $P_r \mid \cdots \mid P_1$, est appelée une *réduite de Frobenius*. Faire une réduction de Frobenius pour un endomorphisme $f \in \mathcal{L}(E)$ revient donc à trouver une base $\mathcal{B}$ de E telle que $\mathrm{Mat}_{\mathcal{B}}(f)$ soit une réduite de Frobenius.

Définition 2.47.
- On dit que $M, N \in \mathcal{M}_n(\mathbb{K})$ sont *semblables* s'il existe $P \in \mathrm{GL}_n(\mathbb{K})$ telle que $M = PNP^{-1}$.
- On dit que $f, g \in \mathcal{L}(E)$ sont *semblables* s'il existe $h \in \mathcal{GL}(E)$ tel que $f = h \circ g \circ h^{-1}$.

Et bien sûr, f et g sont semblables si et seulement $\mathrm{Mat}_{\mathcal{B}}(f)$ et $\mathrm{Mat}_{\mathcal{B}}(g)$ sont semblables (pour n'importe quelle base $\mathcal{B}$ de E).

On vérifie aisément que la relation binaire "f est semblable à g" est une relation d'équivalence sur $\mathcal{L}(E)$ dont les classes d'équivalence sont appelées *classes de similitude*. Un problème naturel est alors de caractériser les classes de similitude de $\mathcal{L}(E)$, c'est-à-dire d'obtenir un critère explicite pour déterminer lorsque deux endomorphismes donnés $f, g \in \mathcal{L}(E)$ sont oui ou non dans la même classe de similitude. Une réponse complète à ce problème est apportée par les invariants de similitude (d'où leur nom !).

Corollaire 2.48. Deux endomorphismes $f, g \in \mathcal{L}(E)$ sont semblables si et seulement si f et g ont les même invariants de similitude.

Démonstration. ($\Leftarrow$) : Si $f, g \in \mathcal{L}(E)$ ont les même invariants de similitude, alors d'après le théorème 2.46, il existe des bases $\mathcal{B}_1$ et $\mathcal{B}_2$ de E telles que

$$\mathrm{Mat}_{\mathcal{B}_1}(f) = \begin{bmatrix} \mathcal{C}(P_1) & & & \\ & \mathcal{C}(P_2) & & (0) \\ & & \ddots & \\ & (0) & & \mathcal{C}(P_r) \end{bmatrix} = \mathrm{Mat}_{\mathcal{B}_2}(g).$$

On pose $P \coloneqq \mathrm{Pass}(\mathcal{B}_1 \to \mathcal{B}_2)$, alors

$$\mathrm{Mat}_{\mathcal{B}_1}(f) = \mathrm{Mat}_{\mathcal{B}_2}(g) = P^{-1}\mathrm{Mat}_{\mathcal{B}_1}(g)P,$$

et donc $\mathrm{Mat}_{\mathcal{B}_1}(f)$ et $\mathrm{Mat}_{\mathcal{B}_1}(g)$ sont semblables, ce qui signifie que f et g sont semblables.
($\Rightarrow$) : Si $f, g \in \mathcal{L}(E)$ sont semblables alors il suffit de montrer que $\mu_f = \mu_g$ puis de reprendre la preuve du théorème 2.44 pour voir que f et g ont les mêmes invariants de similitude. Afin de se simplifier la vie, on va donc se contenter de montrer que $\mu_f = \mu_g$, sachant que $f = h \circ g \circ h^{-1}$. On a

$$0 = h \circ \mu_g(g) \circ h^{-1} = \mu_g(h \circ g \circ h^{-1}) = \mu_g(f), \quad \text{et donc } \mu_f \mid \mu_g.$$

De même, puisque $g = h^{-1} \circ f \circ h$, on a $\mu_g \mid \mu_f$. Ces deux polynômes étant unitaires, on en déduit l'égalité $\mu_f = \mu_g$. $\square$

Application 2.49. En dimension 2 et 3, deux endomorphismes sont semblables si et seulement s'ils ont le même polynôme minimal et le même polynôme caractéristique.

Démonstration. Si $f, g \in \mathcal{L}(E)$ sont semblables, alors on a déjà vu qu'ils ont le même polynôme minimal et le même polynôme caractéristique.

On suppose maintenant que $f, g \in \mathcal{L}(E)$ ont le même polynôme minimal et le même polynôme caractéristique. D'après le corollaire 2.48, il nous faut montrer que f et g ont alors les mêmes invariants de similitude. Étant donné que le produit des invariants de similitude est égal au polynôme caractéristique (au signe près), on voit que le nombre d'invariants de similitude est égal au plus à $n = \dim(E)$. On va distinguer les deux cas $n = 2$ et $n = 3$:

- **Cas** $n = 2$**.** Ici f et g ont au plus deux invariants de similitude. Si $\deg(\mu_f) = 2$, alors f et g ont un seul invariant de similitude $P_1 = \mu_f = \mu_g$. Si $\deg(\mu_f) = 1$, alors $\mu_f(X) = X - \alpha = \mu_g(X)$ avec $\alpha \in \mathbb{K}$, et donc $f = g = \alpha \mathrm{Id}_E$; en particulier f et g sont semblables (et même égaux !) avec pour invariants de similitude $(\mu_f, \chi_f/\mu_f)$.
- **Cas** $n = 3$**.** Ici f et g ont au plus trois invariants de similitude. Si $\deg(\mu_f) = 3$, alors f et g ont un seul invariant de similitude $P_1 = \mu_f = \mu_g$. Si $\deg(\mu_f) = 2$, alors les f et g ont nécessairement deux invariants de similitude qui sont $P_1 = \mu_f = \mu_g$ et $P_2 = \chi_f/\mu_f = \chi_g/\mu_g$. Enfin, si $\deg(\mu_f) = 1$, alors f et g admettent exactement trois invariants de similitude (P_1, P_2, P_3) donnés par $P_1 = P_2 = P_3 = \mu_f = \mu_g$ puisque $P_3 \mid P_2 \mid P_1$. (Alternativement, on peut argumenter comme dans le cas $n = 2$ pour montrer que $f = g = \alpha \mathrm{Id}_E$ avec $\mu_f(X) = X - \alpha = \mu_g(X)$.

□

L'exemple qui suit montre néanmoins que la suite des invariants de similitude n'est pas déterminée uniquement par le polynôme minimal et le polynôme annulateur lorsque $n \geq 4$.

Exemple 2.50. Soient $E = \mathbb{K}^4$ muni de la base canonique $\mathcal{B}$. On considère les deux endomorphismes de E dont les matrices dans la base $\mathcal{B}$ sont données par

$$U := \begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad \text{et} \quad V := \begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}.$$

Alors $\chi_U(X) = \chi_V(X) = X^4$ et $\mu_U(X) = \mu_V(X) = X^2$, mais U et V ne sont pas semblables car elles ne sont pas de même rang. En fait on voit directement que la suite des invariants de similitude de U est (X^2, X, X) tandis que la suite des invariants de similitude de V est (X^2, X^2).

Le résultat qui suit est une autre application importante des invariants de similitude.

Application 2.51. Soient $M, N \in \mathcal{M}_n(\mathbb{R})$. Alors M et N sont semblables dans $\mathcal{M}_n(\mathbb{R})$ si et seulement si M et N sont semblables dans $\mathcal{M}_n(\mathbb{C})$.

Démonstration. Si M et N sont semblables dans $\mathcal{M}_n(\mathbb{R})$, alors il existe $P \in \mathrm{GL}_n(\mathbb{R})$ telle que $PMP^{-1} = N$, et donc M et N sont aussi semblables dans $\mathcal{M}_n(\mathbb{C})$.

Réciproquement, on suppose que M et N sont semblables dans $\mathcal{M}_n(\mathbb{C})$, c'est-à-dire qu'il existe $P \in \mathrm{GL}_n(\mathbb{C})$ telle que $PMP^{-1} = N$. Il s'agit de montrer que l'on peut choisir $P \in \mathrm{GL}_n(\mathbb{R})$, ce qui n'a rien d'évident si on essaye de le faire à la main. En revanche, on peut vérifier que les invariants de similitude de $M \in \mathcal{M}_n(\mathbb{R})$ coïncident avec les invariants de similitude de $M \in \mathcal{M}_n(\mathbb{C})$ (conséquence de l'unicité des invariants de similitude). Il en découle que si M et N ont les mêmes invariants de similitude vues comme matrices dans $\mathcal{M}_n(\mathbb{C})$, alors elles ont aussi les mêmes invariants de similitude vues comme matrices dans $\mathcal{M}_n(\mathbb{R})$, et donc elles sont semblables dans $\mathcal{M}_n(\mathbb{R})$. □

2.4 Réduction de Jordan

Dans cette section on prouve l'existence de la réduction de Jordan (qui doit son nom au mathématicien français Camille Jordan, 1838-1922) pour n'importe quel endomorphisme $f \in \mathcal{L}(E)$ **dont le polynôme caractéristique est scindé**. Cette réduction est tellement employée, en particulier en analyse pour la résolution d'équations différentielles ou pour déterminer le terme général de certaines suites récurrentes, qu'on la nomme parfois "jordanisation" des endomorphismes.

2.4.A Cas d'un endomorphisme nilpotent

On considère dans un premier temps la réduction de Jordan pour un endomorphisme nilpotent $f \in \mathcal{L}(E)$. D'après la proposition 2.7, on a alors $\chi_f(X) = (-1)^n X^n$ qui est scindé.

Théorème 2.52 (Réduction de Jordan pour les endomorphismes nilpotents)**.**
Soit $f \in \mathcal{L}(E)$ un endomorphisme nilpotent. Alors il existe une base $\mathcal{B}$ de E telle que

$$\mathrm{Mat}_{\mathcal{B}}(f) = \begin{bmatrix} 0 & \epsilon_1 & & (0) \\ \vdots & \ddots & \ddots & \\ \vdots & & \ddots & \epsilon_{n-1} \\ 0 & \cdots & \cdots & 0 \end{bmatrix}, \quad \text{avec } \epsilon_i \in \{0,1\} \text{ pour tout } i \in \{1,\dots,n-1\}.$$

Démonstration. Soit $(P_1,\dots,P_r)$ la suite des invariants de similitude de f. D'après le théorème 2.46, on a $P_1 \cdots P_r = (-1)^n \chi_f(X) = X^n$, où la dernière égalité découle du fait que f est nilpotent. Il s'ensuit que, pour tout $i \in \{1,\dots,r\}$, on a $P_i(X) = X^{n_i}$ avec $n_i \in \mathbb{N}^*$ (et $n_1 \geq \cdots \geq n_r$ puisque $P_r \mid \cdots \mid P_1$). Toujours d'après le théorème 2.46, il existe une base $\mathcal{B}' = \{e_1,\dots,e_n\}$ de E telle que

$$\mathrm{Mat}_{\mathcal{B}'}(f) = \begin{bmatrix} \mathcal{C}(P_1) & & & \\ & \mathcal{C}(P_2) & & (0) \\ & & \ddots & \\ & (0) & & \mathcal{C}(P_r) \end{bmatrix} = \begin{bmatrix} \mathcal{C}(X^{n_1}) & & & \\ & \mathcal{C}(X^{n_2}) & & (0) \\ & & \ddots & \\ & (0) & & \mathcal{C}(X^{n_r}) \end{bmatrix},$$

avec $\mathcal{C}(X^{n_i}) = \begin{bmatrix} 0 & \cdots & 0 & 0 \\ 1 & \ddots & \vdots & \vdots \\ & \ddots & 0 & 0 \\ (0) & & 1 & 0 \end{bmatrix} \in \mathcal{M}_{n_i}(\mathbb{K})$. Il suffit alors de prendre pour $\mathcal{B}$ la base $\{e_n,\dots,e_1\}$ pour avoir

$$\mathrm{Mat}_{\mathcal{B}}(f) = \begin{bmatrix} 0 & \epsilon_1 & & (0) \\ \vdots & \ddots & \ddots & \\ \vdots & & \ddots & \epsilon_{n-1} \\ 0 & \cdots & \cdots & 0 \end{bmatrix}, \quad \text{avec } \epsilon_i \in \{0,1\} \text{ pour tout } i \in \{1,\dots,n-1\}.$$

□

Définition 2.53. Pour tout $m \in \mathbb{N}^*$, la matrice

$$J_m := \begin{bmatrix} 0 & 1 & & (0) \\ \vdots & \ddots & \ddots & \\ \vdots & & \ddots & 1 \\ 0 & \cdots & \cdots & 0 \end{bmatrix} \in \mathcal{M}_m(\mathbb{K})$$

est appelé *bloc de Jordan* nilpotent de taille m.

Quitte à permuter les vecteurs de la base $\mathcal{B}$ dans le théorème 2.52, on peut toujours supposer que

$$\operatorname{Mat}_{\mathcal{B}}(f) = \begin{bmatrix} J_{n_1} & & & \\ & J_{n_2} & & (0) \\ & & \ddots & \\ & (0) & & J_{n_r} \end{bmatrix}, \text{ avec } n_1 \geq n_2 \geq \cdots \geq n_r \geq 1.$$

En particulier, on voit qu'à un endomorphisme nilpotent $f \in \mathcal{L}(E)$, on peut associer une suite d'entiers $(n_1, \ldots, n_r)$ telle que $n_1 \geq \cdots \geq n_r \geq 1$ avec $\sum_i n_i = n$. Une telle suite d'entiers est appelée une *partition* de n.

Corollaire 2.54. L'application qui à un endomorphisme nilpotent associe la partition $(n_1, \ldots, n_r)$ de n comme ci-dessus induit une bijection entre les classes de similitudes d'endomorphismes nilpotents de E et les partitions de n.

Démonstration. D'après ce qui précède, la donnée de la partition $(n_1, \ldots, n_r)$ de n est équivalente à la donnée de la suite des invariants de similitude $(X^{n_1}, \ldots, X^{n_r})$ de f. On en déduit que si f et g sont deux endomorphismes nilpotents qui sont semblables, alors ils ont les mêmes invariants de similitude, et donc donnent la même partition de n. En outre, étant donnée une partition $(n_1, \ldots, n_r)$ de n, on peut toujours lui associer la classe de similitude de l'endomorphisme nilpotent $f \in \mathcal{L}(E)$ dont la matrice dans une base $\mathcal{B}$ fixée de E est

$$\operatorname{Mat}_{\mathcal{B}}(f) = \begin{bmatrix} J_{n_1} & & & \\ & J_{n_2} & & (0) \\ & & \ddots & \\ & (0) & & J_{n_r} \end{bmatrix}.$$

On a donc bien une correspondance bijective entre les classes de similitudes d'endomorphismes nilpotents de E et les partitions de n. □

Exemple 2.55.

- Si $n = 1$, alors il y a une seule partition (1) qui correspond à la matrice nulle.
- Si $n = 2$, alors il y a deux partitions (2) et $(1,1)$ qui correspondent aux matrices J_2 et $\begin{bmatrix} 0 \end{bmatrix}$ respectivement.
- Si $n = 3$, alors il y a trois partitions (3), $(2,1)$ et $(1,1,1)$ qui correspondent aux matrices J_3, $\begin{bmatrix} J_2 & \\ & 0 \end{bmatrix}$ et $\begin{bmatrix} 0 \end{bmatrix}$ respectivement.
- Si $n = 4$, alors il y a cinq partitions (4), $(3,1)$, $(2,2)$, $(2,1,1)$ et $(1,1,1,1)$ qui correspondent aux matrices J_4, $\begin{bmatrix} J_3 & \\ & 0 \end{bmatrix}$, $\begin{bmatrix} J_2 & \\ & J_2 \end{bmatrix}$, $\begin{bmatrix} J_2 & & \\ & 0 & \\ & & 0 \end{bmatrix}$ et $\begin{bmatrix} 0 \end{bmatrix}$ respectivement.

2.4.B Cas d'un endomorphisme quelconque

On va maintenant considérer la réduction de Jordan pour un endomorphisme quelconque $f \in \mathcal{L}(E)$ dont le polynôme caractéristique χ_f est scindé.

Théorème 2.56 (Réduction de Jordan).
Soit $f \in \mathcal{L}(E)$ un endomorphisme dont le polynôme caractéristique est scindé et s'écrit

$$\chi_f(X) = (-1)^n \prod_{i=1}^{r} (X - \lambda_i)^{\alpha_i},$$

avec $\alpha_i \in \mathbb{N}^*$ et les λ_i deux à deux distincts. Alors il existe une base $\mathcal{B}$ de E telle que

$$\mathrm{Mat}_{\mathcal{B}}(f) = \begin{bmatrix} K_1 & & & \\ & K_2 & & (0) \\ & & \ddots & \\ & (0) & & K_r \end{bmatrix},$$

avec, pour tout $j \in \{1, \dots, r\}$,

$$K_j = \begin{bmatrix} \lambda_j & \epsilon_{j,1} & & (0) \\ \vdots & \ddots & \ddots & \\ \vdots & & \ddots & \epsilon_{j,\alpha_j - 1} \\ 0 & \cdots & \cdots & \lambda_j \end{bmatrix} \in \mathcal{M}_{\alpha_j}(\mathbb{K}), \quad \text{avec chaque } \epsilon_{j,l} \in \{0, 1\}.$$

Démonstration. Nous allons voir que ce théorème est essentiellement une conséquence du lemme des noyaux (théorème 2.28) et de la réduction de Jordan pour les endomorphismes nilpotents (théorème 2.52).

Pour tout $i \in \{1, \dots, r\}$, on rappelle que l'on note

$$N_i := \mathrm{Ker}\left((f - \lambda_i \mathrm{Id}_E)^{\alpha_i}\right)$$

l'espace caractéristique de f associé à la valeur propre λ_i (voir section 2.2.D). On rappelle aussi que $E = \bigoplus_{i=1}^{r} N_i$ et que les N_i sont stables par f (proposition 2.33). Pour tout $i \in \{1, \dots, r\}$, on note $\beta_i = \dim(N_i)$, $f_i := f_{|N_i}$ et $g_i := f_i - \lambda_i \mathrm{Id}_{N_i}$. L'endomorphisme $g_i \in \mathcal{L}(N_i)$ est nilpotent puisque $g_i^{\alpha_i} = 0$. D'après la réduction de Jordan pour les endomorphismes nilpotents (théorème 2.52), il existe une base $\mathcal{B}_i$ de N_i telle que

$$\mathrm{Mat}_{\mathcal{B}_i}(g_i) = \begin{bmatrix} 0 & \epsilon_1 & & (0) \\ \vdots & \ddots & \ddots & \\ \vdots & & \ddots & \epsilon_{\beta_i - 1} \\ 0 & \cdots & \cdots & 0 \end{bmatrix} \in \mathcal{M}_{\beta_i}(\mathbb{K}), \quad \text{avec } \epsilon_l \in \{0, 1\} \text{ pour tout } l \in \{1, \dots, \beta_i - 1\}.$$

Or,

$$\mathrm{Mat}_{\mathcal{B}_i}(g_i) = \mathrm{Mat}_{\mathcal{B}_i}(f_i - \lambda_i \mathrm{Id}_{N_i}) = \mathrm{Mat}_{\mathcal{B}_i}(f_i) - \lambda_i \mathrm{Id},$$

et donc

$$\mathrm{Mat}_{\mathcal{B}_i}(f_i) = \lambda_i \mathrm{Id} + \mathrm{Mat}_{\mathcal{B}_i}(g_i) = \begin{bmatrix} \lambda_i & \epsilon_1 & & (0) \\ \vdots & \ddots & \ddots & \\ \vdots & & \ddots & \epsilon_{\beta_i - 1} \\ 0 & \cdots & \cdots & \lambda_i \end{bmatrix} \in \mathcal{M}_{\beta_i}(\mathbb{K}), \quad \text{avec } \epsilon_l \in \{0, 1\} \text{ pour tout } l \in \{1, \dots, \beta_i - 1\}.$$

Ainsi, en posant $\mathcal{B} = \mathcal{B}_1 \cup \cdots \cup \mathcal{B}_r$, on obtient une base $\mathcal{B}$ de E telle que $\mathrm{Mat}_{\mathcal{B}}(f)$ soit de la forme souhaitée. □

Définition 2.57. Pour tout $\lambda \in \mathbb{K}$ et pour tout $m \in \mathbb{N}^*$, la matrice

$$J_{\lambda,m} := \begin{bmatrix} \lambda & 1 & & (0) \\ \vdots & \ddots & \ddots & \\ \vdots & & \ddots & 1 \\ 0 & \cdots & \cdots & \lambda \end{bmatrix} \in \mathcal{M}_m(\mathbb{K})$$

est appelé *bloc de Jordan* de taille m associé à la valeur propre λ.

Une matrice diagonale par blocs avec uniquement des blocs de Jordan sur la diagonale est appelée une *réduite de Jordan*. Faire une réduction de Jordan pour un endomorphisme $f \in \mathcal{L}(E)$ revient donc à trouver une base $\mathcal{B}$ de E telle que $\mathrm{Mat}_{\mathcal{B}}(f)$ soit une réduite de Jordan.

Remarques 2.58.
(1) Si $f \in \mathcal{L}(E)$ est diagonalisable, alors sa réduite de Jordan est une matrice diagonale.
(2) Le théorème 2.56 implique en particulier le théorème 2.16, mais l'existence d'une réduite de Jordan est en fait un résultat beaucoup plus fin que l'existence d'une trigonalisation.
(3) La réduction de Jordan implique également la décomposition de Dunford (c'est en quelque sorte une version effective et optimale de la décomposition de Dunford). En effet, si $\mathrm{Mat}_{\mathcal{B}}(f) = J_{\lambda,m}$ avec $\lambda \in \mathbb{K}$ et $m \in \mathbb{N}^*$, alors en définissant $d, n \in \mathcal{L}(E)$ comme les endomorphismes satisfaisant $\mathrm{Mat}_{\mathcal{B}}(d) = \lambda I_m$ et $\mathrm{Mat}_{\mathcal{B}}(n) = J_m$, on retrouve que d est diagonalisable, n est nilpotent et $d \circ n = n \circ d$. Et il en va de même pour n'importe quelle réduite de Jordan en faisant une décomposition de Dunford bloc à bloc.
(4) Lorsque $\mathbb{K} = \mathbb{R}$, le polynôme χ_f n'est pas toujours scindé. Néanmoins il existe une version alternative de la réduction de Jordan dans le cas réel (c.f. par exemple le théorème 5 du polycopié de Grégory Vial : `https://w3.ens-rennes.fr/math/people/gregory.vial/files/cplts/jordan.pdf`).

En pratique, étant donné un endomorphisme $f \in \mathcal{L}(E)$ avec E de petite dimension (disons $n \leq 6$) la donnée de χ_f, μ_f et des dimensions des espaces propres permet de trouver la réduite de Jordan de f sans avoir à calculer explicitement la base $\mathcal{B}$. En effet, on peut vérifier (en exercice !) que :
- Le nombre de blocs de Jordan de associés à la valeur propre λ est égal à la dimension de l'espace propre E_λ.
- La taille maximal d'un bloc de Jordan associé à la valeur propre λ est égal à la multiplicité de λ comme racine de μ_f.

Pour finir, nous allons voir deux méthodes pour calculer une *base de Jordan* de $f \in \mathcal{L}(E)$ (c'est-à-dire une base de E telle que $\mathrm{Mat}_{\mathcal{B}}(f)$ soit une réduite de Jordan). Du point de vue matriciel, se donner $f \in \mathcal{L}(E)$ revient à se donner une matrice $M \in \mathcal{M}_n(\mathbb{K})$ (dès lors que l'on a fixé une base quelconque de E), et se donner une base de Jordan de f revient alors à se donner $P \in \mathrm{GL}_n(\mathbb{K})$ telle que $J := P^{-1}MP$ soit une réduite de Jordan.

- **Méthode 1 :** Pour chaque valeur propre λ_i de $f \in \mathcal{L}(E)$, on part de $w_1 \in E_{\lambda_i} = \mathrm{Ker}(f - \lambda_i \mathrm{Id}_E)$ et on calcule les $w_j \in E$ de proche en proche tels que

$$(f - \lambda_i \mathrm{Id}_E)(w_{j+1}) = w_j \text{ avec } 1 \leq j \leq \alpha_i = \dim(N_i).$$

Alors $\{w_1, w_2, \ldots\}$ une base de Jordan correspondant à un bloc J_{m,λ_i}.

– **Exemple 1 :** On considère la matrice $A := \begin{bmatrix} 1 & 1 & 0 \\ -1 & 0 & -1 \\ 0 & -1 & 1 \end{bmatrix}$. On peut vérifier que $\chi_A(X) = -X(1-X)^2$ et $\dim(E_1) = 1$. Donc la réduite de Jordan de A est la matrice $J_A = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$.
On fixe $v_1 \in E_0$ et $v_2 \in E_1$. On cherche alors $v_3 \in \mathbb{K}^3$ tel que $Av_3 = v_2 + v_3$, c'est-à-dire, tel que v_3 est solution du système linéaire $(A - I_3)v_3 = v_2$. Par exemple on peut vérifier que

$$\begin{bmatrix} v_1, & v_2, & v_3 \end{bmatrix} = \begin{bmatrix} 1 & 1 & -1 \\ -1 & 0 & 1 \\ -1 & -1 & 0 \end{bmatrix}$$

convient. On pose alors $P := \begin{bmatrix} v_1, & v_2, & v_3 \end{bmatrix}$ et on vérifie que l'on a bien $J_A = P^{-1}AP$.

– **Exemple 2 :** On considère la matrice $B := \begin{bmatrix} 4 & 1 & -1 \\ -2 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix}$. On peut vérifier que $\chi_B(X) = (2-X)^3$ et $\dim(E_2) = 1$. Donc la réduite de Jordan de A est la matrice $J_B = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{bmatrix}$. On fixe $v_1 \in E_2$. On cherche alors $v_2 \in \mathbb{K}^3$ tel que $Bv_2 = 2v_2 + v_1$, c'est-à-dire, tel que v_2 est solution du système linéaire $(B - 2I_3)v_2 = v_1$. Enfin, on cherche $v_3 \in \mathbb{K}^3$ tel que tel que v_3 est solution du système linéaire $(B - 2I_3)v_3 = v_2$. Par exemple on peut vérifier que

$$\begin{bmatrix} v_1, & v_2, & v_3 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ -1 & -1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$$

convient. On pose alors $P := \begin{bmatrix} v_1, & v_2, & v_3 \end{bmatrix}$ et on vérifie que l'on a bien $J_A = P^{-1}AP$.

- **<u>Méthode 2</u> :** Pour chaque valeur propre λ_i de $f \in \mathcal{L}(E)$, on part de $w_1 \in N_i = \mathrm{Ker}((f - \lambda_i \mathrm{Id}_E)^{\beta_i}) \setminus \mathrm{Ker}((f - \lambda_i \mathrm{Id}_E)^{\beta_i - 1})$, avec β_i l'ordre de λ_i comme racine de μ_f, et on calcule les $w_j \in E$ de proche en proche en posant

$$w_{j+1} := (f - \lambda_i \mathrm{Id}_E)(w_j) \text{ avec } 1 \leq j \leq \beta_i.$$

Alors $\{\ldots, w_2, w_1\}$ une base de Jordan correspondant à un bloc J_{m,λ_i}.

– **Exemple 1 :** Avec la matrice A précédente, on fixe $v_1 \in E_0$, on choisit

$$v_3 \in \mathrm{Ker}((A - I_3)^2) \setminus \mathrm{Ker}(A - I_3)$$

et on pose $v_2 = (A - I_3)v_3$. Par exemple on peut vérifier que

$$\begin{bmatrix} v_1, & v_2, & v_3 \end{bmatrix} = \begin{bmatrix} 1 & 1 & -1 \\ -1 & 0 & 1 \\ -1 & -1 & 0 \end{bmatrix}$$

convient. On pose alors $P := \begin{bmatrix} v_1, & v_2, & v_3 \end{bmatrix}$ et on vérifie que l'on a bien $J_A = P^{-1}AP$.

– **Exemple 2 :** Avec la matrice B précédente, on choisit

$$v_3 \in \operatorname{Ker}\left((B-2I_3)^3\right) \smallsetminus \operatorname{Ker}\left((B-I_3)^2\right)$$

et on pose $v_2 = (B-I_3)v_3$ puis $v_1 = (B-I_3)v_2$. Par exemple on peut vérifier que

$$\begin{bmatrix} v_1, & v_2, & v_3 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ -1 & -1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$$

convient. On pose alors $P := \begin{bmatrix} v_1, & v_2, & v_3 \end{bmatrix}$ et on vérifie que l'on a bien $J_B = P^{-1}BP$.

Quelques références bibliographiques

- Xavier Gourdon. *Algèbre*, Chapitre 4 et appendice B.
- Vincent Beck, Jérôme Malick et Gabriel Peyré. *Objectif Agrégation*, Chapitres 4.2 et 6.5.
- Joseph Grifone. *Algèbre linéaire*, Chapitre 6.
- Edmond Ramis, Claude Deschamps, Jacques Odoux.
  Le cours de mathématiques Tome 1- Algèbre, Chapitre 12.

Chapitre 3

Formes quadratiques

Dans tout ce chapitre $\mathbb{K}$ désigne un corps de caractéristique $\neq 2$, par exemple $\mathbb{Q}, \mathbb{R}, \mathbb{C}$.

Motivation et but

L'archétype de forme quadratique est la forme $x^2 + y^2 + z^2$ sur $\mathbb{R}^3$, qui définit une structure euclidienne et dont la racine carrée permet de calculer la norme d'un vecteur. Un autre exemple très classique en physique théorique est la forme $x^2 + y^2 + z^2 - t^2$ sur $\mathbb{R}^4$, qui permet de définir l'espace de Minkowski utilisé en relativité restreinte.

Les formes quadratiques interviennent dans de nombreux domaines des mathématiques :
- en géométrie affine avec la classification des coniques et des quadriques ;
- en géométrie différentielle avec les première et seconde formes fondamentales ;
- en analyse avec les développements limités dont le terme d'ordre 2 est une forme quadratique qui correspond à la différentielle seconde de la fonction numérique considérée ;
- en optimisation où la recherche d'extrema locaux passe souvent par l'étude de la matrice Hessienne de la fonction numérique considérée ;
- en statistique et en théorie des probabilités avec la variance qui est une forme quadratique qui mesure de la dispersion des valeurs d'un échantillon ou d'une distribution de probabilité ;
- en théorie des groupes puisque tout groupe fini peut-être réalisé comme sous-groupe du groupe des isométries d'une certaine forme quadratique ;
- en théorie des nombres où l'on s'intéresse par exemple aux formes quadratiques entières (c'est-à-dire à coefficients entiers) pour la résolution d'équations diophantiennes comme le théorème des deux carrés de Fermat ; etc.

Dans ce chapitre nous exposons les bases de la théorie des formes quadratiques avec en ligne de mire la réduction de Gauss, qui permet d'écrire toute forme quadratique comme une combinaison linéaire de carrés de formes linéaires linéairement indépendantes (cf. section 3.4), la classification des formes quadratiques lorsque $\mathbb{K} = \mathbb{C}$ ou $\mathbb{R}$ (cf. section 3.5), en termes de certains invariants numériques (rang, signature, discriminant), ainsi que le théorème d'orthogonalisation simultanée dans le cas euclidien (cf. section 3.6).

3.1 Formes bilinéaires et formes quadratiques

Soit E un $\mathbb{K}$-espace vectoriel (pas nécessairement de dimension finie pour le moment).

Définitions 3.1.

- Une *forme bilinéaire* sur E est une application
$$\phi: E \times E \to \mathbb{K},\ (x,y) \mapsto \phi(x,y)$$
telle que, pour tous $x_0, y_0 \in E$, les applications $\phi(x_0, -): E \to \mathbb{K}$ et $\phi(-, y_0): E \to \mathbb{K}$ sont des formes linéaires. On note $\mathcal{B}il(E,\mathbb{K})$ l'ensemble des formes bilinéaires sur E ; cet ensemble est naturellement muni d'une structure de $\mathbb{K}$-espace vectoriel induite par celle de $\mathbb{K}$.
- Une forme bilinéaire $\phi: E\times E \to \mathbb{K}$ est dite *symétrique* si, pour tous $x, y \in E$, on a $\phi(x,y) = \phi(y,x)$. De même, ϕ est dite *antisymétrique* si, pour tous $x, y \in E$, on a $\phi(x,y) = -\phi(y,x)$.
- Soit $\phi: E \times E \to \mathbb{K}$ une forme bilinéaire symétrique. On appelle *noyau* de ϕ le $\mathbb{K}$-sous-espace vectoriel de E défini par
$$\mathrm{Ker}(\phi) := \bigcap_{x\in E} \mathrm{Ker}(\phi(x,-)) = \bigcap_{y\in E} \mathrm{Ker}(\phi(-,y)).$$
(Le fait que ces deux $\mathbb{K}$-sous-espaces vectoriels coïncident est une conséquence du fait que ϕ est symétrique.) Une forme bilinéaire symétrique ϕ est dite *non-dégénérée* si $\mathrm{Ker}(\phi) = \{0\}$, et elle est dite *dégénérée* si $\mathrm{Ker}(\phi) \neq \{0\}$.

Exemples 3.2.

(i) Si $E = \mathbb{K}^2$, alors
$$\phi_0: E \times E \to \mathbb{K},\ (u,v) \mapsto \det(u,v)$$
est une forme bilinéaire antisymétrique sur E.

(ii) Si $E = \mathbb{K}^n$, alors
$$\phi_1: E \times E \to \mathbb{K},\ (x,y) \mapsto \sum_{i=1}^{n} x_i y_i$$
est une forme bilinéaire symétrique non-dégénérée sur E. Ici on note $(x_1, \dots, x_n)$ et $(y_1, \dots, y_n)$ les coordonnées des vecteurs x et y relativement à la base canonique de E.

(iii) Si $E = \mathcal{C}^1(\mathbb{R},\mathbb{R})$, alors
$$\phi_2: E \times E \to \mathbb{K},\ (f,g) \mapsto \int_0^1 f(t)g(t)dt$$
est une forme bilinéaire symétrique non-dégénérée sur E.

Remarque 3.3. Il y a un isomorphisme entre $\mathcal{B}il(E,\mathbb{K})$ et $\mathcal{L}(E,E^*)$ donné par

$$\begin{array}{ccc} \mathcal{B}il(E,\mathbb{K}) & \simeq & \mathcal{L}(E,E^*) \\ \phi & \mapsto & (x \mapsto \phi(x,-)) \\ (x,y) \mapsto f(x)(y) & \leftarrow\!\shortmid & f \end{array}$$

On en déduit que $\dim(\mathcal{B}il(E,\mathbb{K})) = \dim(\mathcal{L}(E,E^*)) = n^2$ lorsque $\dim(E) = n < \infty$. De plus, si ϕ est une forme bilinéaire symétrique correspondante à l'application linéaire f via l'isomorphisme ci-dessus, alors on vérifie que $\mathrm{Ker}(\phi) = \mathrm{Ker}(f)$ et $\mathrm{rg}(\phi) = \mathrm{rg}(f)$ (le rang d'une forme bilinéaire sera défini dans la section 3.2, cf. définition 3.14).

Définition 3.4. Une *forme quadratique* sur E est une application
$$q: E \to \mathbb{K},\ x \mapsto \phi(x,x)$$
où $\phi: E \times E \to \mathbb{K}$ est une forme bilinéaire symétrique appelée la *forme polaire* de q.

La forme polaire ϕ d'une forme quadratique q est uniquement déterminée par q. En effet, on a les *formules de polarisation* suivantes :

$$\forall x, y \in E,\ \phi(x,y) = \frac{1}{2}[q(x+y) - q(x) - q(y)] = \frac{1}{4}[q(x+y) - q(x-y)].$$

Remarque 3.5. En pratique, une application $q: E \to \mathbb{K}$ est une forme quadratique sur E si
(i) $\forall x \in E$, $\forall \lambda \in \mathbb{K}$, $q(\lambda x) = \lambda^2 q(x)$; et
(ii) $\forall x, y \in E$, l'application $(x,y) \mapsto \frac{1}{2}(q(x+y) - q(x) - q(y))$ est bilinéaire.
Précisons que ces deux conditions ne sont pas équivalentes. Par exemple la fonction $f: \mathbb{K} \to \mathbb{K}$, $t \mapsto t^2 + t$ vérifie (ii) mais pas (i), tandis que la fonction $g: \mathbb{R}^2 \to \mathbb{R}$, $(x_1, x_2) \mapsto (|x_1| + |x_2|)^2$ vérifie ((i)) mais pas (ii).

Exemples 3.6.

(i) Si $E = \mathbb{K}^n$, alors

$$q_1: E \to \mathbb{K},\ x \mapsto \sum_{i=1}^n x_i^2$$

est une forme quadratique sur E (lorsque $\mathbb{K} = \mathbb{R}$ on retrouve la norme euclidienne).

(ii) Si $E = \mathcal{M}_n(\mathbb{K})$, alors

$$\begin{aligned} &q_2: E \to \mathbb{K},\ A \mapsto \operatorname{tr}(A)^2 \ ; \\ &q_3: E \to \mathbb{K},\ A \mapsto \operatorname{tr}(A^2) \ ; \text{et} \\ &q_4: E \to \mathbb{K},\ A \mapsto \operatorname{tr}({}^tAA) \end{aligned}$$

sont des formes quadratiques sur E.

(iii) Si $E = \mathbb{K}[X]$, alors

$$\begin{aligned} &q_5: E \to \mathbb{K},\ P \mapsto \int_0^1 P(t)P'(t)dt \ ; \text{et} \\ &q_6: E \to \mathbb{K},\ P \mapsto \int_0^1 (P(t)^2)dt - \left(\int_0^1 P(t)dt\right)^2 \end{aligned}$$

sont des formes quadratiques sur E.

Le lemme qui suit justifie pourquoi on identifie généralement les formes quadratiques sur un espace vectoriel de dimension finie n aux polynômes homogènes de degré 2 en n variables.

Lemme 3.7. On suppose que $dim(E) = n < \infty$. Une application $q: E \to \mathbb{K}$ est une forme quadratique si et seulement si, étant donnée une base $\mathcal{B} = \{e_1, \dots, e_n\}$ de E (ou, d'une façon équivalente, pour toute base $\mathcal{B} = \{e_1, \dots, e_n\}$ de E), pour tout $x = (x_1, \dots, x_n) \in E$, l'élément $q(x)$ est un polynôme homogène de degré 2 en les composantes x_i de x dans la base $\mathcal{B}$.

Démonstration. On écrit $x = \sum_{i=1}^n x_i e_i$ et $y = \sum_{i=1}^n y_i e_i$.

($\Rightarrow$) : Si q est une forme quadratique sur E, on note $\phi: E \times E \to \mathbb{K}$ sa forme polaire. Alors

$$\phi(x,y) = \phi\left(\sum_{i=1}^n x_i e_i, \sum_{j=1}^n y_j e_j\right) = \sum_{i,j=1}^n x_i y_j \phi(e_i, e_j) = \sum_{i,j=1}^n \lambda_{ij} x_i y_j,$$

avec $\lambda_{ij} = \phi(e_i, e_j) \in \mathbb{K}$. Et donc

$$\forall x = (x_1, \dots, x_n) \in E,\ q(x) = \phi(x,x) = \sum_{i,j=1}^{n} \lambda_{ij} x_i x_j$$

est un polynôme homogène de degré 2 en les composantes x_i de x dans la base $\mathcal{B}$.

($\Leftarrow$) : Soit l'application $q\colon E \to \mathbb{K}$ telle que $q(x) = \sum_{1 \le i \le j \le n} \mu_{ij} x_i x_j$, avec $\mu_{ij} \in \mathbb{K}$. Alors, en posant $\lambda_{ii} = \mu_{ii}$ et $\lambda_{ij} = \lambda_{ji} = \frac{1}{2}\mu_{ij}$ pour tous $1 \le i < j \le n$, on obtient

$$q(x) = \phi(x,x), \quad \text{où } \phi(x,y) = \sum_{i,j=1}^{n} \lambda_{ij} x_i y_j$$

est une forme bilinéaire symétrique sur E. Et donc q est une forme quadratique sur E. $\square$

⚠ **On suppose dorénavant que E est un $\mathbb{K}$-espace vectoriel de dimension finie $n \in \mathbb{N}$.**

3.2 Expression matricielle

On fixe une base $\mathcal{B} = \{e_1, \dots, e_n\}$ de E. Soient $x, y \in E$ et soit $\phi\colon E \times E \to \mathbb{K}$ une forme bilinéaire sur E. On écrit $x = \sum_{i=1}^{n} x_i e_i$ et $y = \sum_{i=1}^{n} y_i e_i$. Alors

$$\phi(x,y) = \phi\left(\sum_{i=1}^{n} x_i e_i, \sum_{j=1}^{n} y_j e_j\right) = \sum_{i,j=1}^{n} x_i y_j \phi(e_i, e_j) = {}^t X \operatorname{Mat}_{\mathcal{B}}(\phi)\, Y$$

avec $X \coloneqq \operatorname{Mat}_{\mathcal{B}}(x) = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$, $Y \coloneqq \operatorname{Mat}_{\mathcal{B}}(y) = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$ et $\operatorname{Mat}_{\mathcal{B}}(\phi) \in \mathcal{M}_n(\mathbb{K})$ la matrice carrée dont le coefficient (i,j) vaut $\phi(e_i, e_j)$.

En particulier, la forme bilinéaire ϕ est (anti)symétrique si et seulement si la matrice $\operatorname{Mat}_{\mathcal{B}}(\phi)$ est (anti)symétrique.

Exemples 3.8.

(i) Soient $E = \mathbb{K}^3$, muni de la base canonique, notée $\mathcal{B}$, et $\phi\colon E \times E \to \mathbb{K},\ (x,y) \mapsto x_1 y_1 + x_2 y_1 + x_3 y_2$. Alors la matrice $\operatorname{Mat}_{\mathcal{B}}(\phi) = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$ n'est ni symétrique ni antisymétrique puisque la forme bilinéaire ϕ n'est ni symétrique ni antisymétrique.

(ii) Soient $E = \mathbb{K}^2$, muni de la base $\mathcal{B} = \{(1,0),(1,1)\}$, et $\phi_0\colon E \times E \to \mathbb{K},\ (u,v) \mapsto \det(u,v)$. Alors la matrice $\operatorname{Mat}_{\mathcal{B}}(\phi_0) = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$ est antisymétrique puisque la forme bilinéaire ϕ_0 est antisymétrique.

(iii) Soient $E = \mathcal{M}_2(K)$, muni de la base canonique $\mathcal{B} = \{E_{11}, E_{12}, E_{21}, E_{22}\}$, et $\phi_3\colon E \times E \to \mathbb{K},\ (A,B) \mapsto \operatorname{tr}(AB)$. Alors la matrice $\operatorname{Mat}_{\mathcal{B}}(\phi_3) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$ est symétrique puisque la forme bilinéaire ϕ_3 est symétrique.

Remarque 3.9. L'application $\phi \in \mathcal{B}il(E,\mathbb{K}) \mapsto \mathrm{Mat}_{\mathcal{B}}(\phi) \in \mathcal{M}_n(\mathbb{K})$ est un isomorphisme. Celui-ci se restreint en un isomorphisme entre les formes bilinéaires (anti)symétriques sur E est les matrices (anti)symétriques dans $\mathcal{M}_n(\mathbb{K})$.

On rappelle que $\mathcal{M}_n(K) = \mathcal{S}_n(\mathbb{K}) \oplus \mathcal{A}_n(K)$, où $\mathcal{S}_n(\mathbb{K})$ est le $\mathbb{K}$-sous-espace vectoriel des matrices symétriques (de dimension $\frac{1}{2}n(n+1)$) et $\mathcal{A}_n(\mathbb{K})$ est le $\mathbb{K}$-sous-espace vectoriel des matrices antisymétriques (de dimension $\frac{1}{2}n(n-1)$). En effet, il est clair que $\mathcal{S}_n(\mathbb{K}) \cap \mathcal{A}_n(\mathbb{K}) = \{0\}$, et toute matrice $M \in \mathcal{M}_n(\mathbb{K})$ vérifie $M = \frac{1}{2}(M + {}^tM) + \frac{1}{2}(M - {}^tM)$, qui est la somme d'une matrice symétrique et d'une matrice antisymétrique.

En particulier, on en déduit que

- l'espace vectoriel des formes bilinéaire symétriques sur E est de dimension $\frac{1}{2}n(n+1)$;
- l'espace vectoriel des formes bilinéaire antisymétriques sur E est de dimension $\frac{1}{2}n(n-1)$; et
- toute forme bilinéaire sur E peut s'écrire de manière unique comme la somme d'une forme bilinéaire symétrique et d'une forme bilinéaire antisymétrique.

Lemme 3.10. Soit ϕ une forme bilinéaire symétrique sur E, et soit $\mathcal{B}$ une base de E. Alors $\mathrm{Ker}(\mathrm{Mat}_{\mathcal{B}}(\phi)) = \mathrm{Ker}(\phi)$ ne dépend pas du choix de la base $\mathcal{B}$. En particulier, ϕ est non-dégénérée si et seulement $\det(\mathrm{Mat}_{\mathcal{B}}(\phi)) \neq 0$, quelle que soit la base $\mathcal{B}$ considérée.

Démonstration. Comme précédemment, si $x, y \in E$, on note $X := \mathrm{Mat}_{\mathcal{B}}(x)$, $Y := \mathrm{Mat}_{\mathcal{B}}(y)$. Alors

$$\begin{aligned}
\mathrm{Ker}(\phi) &= \{y \in E \mid \forall x \in E,\ \phi(x,y) = 0\} \\
&= \{Y \in \mathcal{M}_{n,1}(\mathbb{K}) \mid \forall X \in \mathcal{M}_{n,1}(\mathbb{K}),\ {}^tX\mathrm{Mat}_{\mathcal{B}}(\phi)Y = 0\} \\
&= \{Y \in \mathcal{M}_{n,1}(\mathbb{K}) \mid \mathrm{Mat}_{\mathcal{B}}(\phi)Y = 0\} \\
&= \mathrm{Ker}(\mathrm{Mat}_{\mathcal{B}}(\phi)).
\end{aligned}$$

□

Lemme 3.11 (Formule de changement de base)**.** Soient $\mathcal{B}$ et $\mathcal{B}'$ deux bases de E, et soit $P = \mathrm{Pass}(\mathcal{B} \to \mathcal{B}')$. Si ϕ est une forme bilinéaire sur E, alors

$$\mathrm{Mat}_{\mathcal{B}'}(\phi) = {}^tP\,\mathrm{Mat}_{\mathcal{B}}(\phi)\,P.$$

Démonstration. On note $X := \mathrm{Mat}_{\mathcal{B}}(x)$, $Y := \mathrm{Mat}_{\mathcal{B}}(y)$, $X' := \mathrm{Mat}_{\mathcal{B}'}(x)$ et $Y' := \mathrm{Mat}_{\mathcal{B}'}(y)$. Alors $Y = PY'$ et $X = PX'$, et donc

$${}^tX\,\mathrm{Mat}_{\mathcal{B}}(\phi)\,Y = {}^t(PX')\,\mathrm{Mat}_{\mathcal{B}}(\phi)\,PY' = {}^tX'({}^tP\,\mathrm{Mat}_{\mathcal{B}}(\phi)\,P)Y' = {}^tX'\,\mathrm{Mat}_{\mathcal{B}'}(\phi)\,Y',$$

d'où $\mathrm{Mat}_{\mathcal{B}'}(\phi) = {}^tP\,\mathrm{Mat}_{\mathcal{B}}(\phi)\,P$. □

Warning 3.12. Il faut veiller à bien distinguer la formule de changement de base pour les formes bilinéaires et la formule de changement de base pour les applications linéaires.

Exemple 3.13. Soient $E = \mathbb{K}[X]_{\leq 2}$, $\mathcal{B} = \{1, X, X^2\}$, $\mathcal{B}' = \{X - X^2, X^2, 1 - X\}$ et

$$\phi_2 : E \to \mathbb{K},\ (P,Q) \mapsto \int_0^1 (P(t)Q'(t) + P'(t)Q(t))dt,$$

alors

$$\mathrm{Mat}_{\mathcal{B}'}(\phi_2) = {}^t\mathrm{Pass}(\mathcal{B} \to \mathcal{B}')\mathrm{Mat}_{\mathcal{B}}(\phi_2)\mathrm{Pass}(\mathcal{B} \to \mathcal{B}') = \begin{bmatrix} 0 & 1 & -1 \\ 0 & 0 & 1 \\ 1 & -1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & -1 \\ -1 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}.$$

Définition 3.14. Soit ϕ une forme bilinéaire sur E. On appelle *rang* de ϕ l'entier

$$\operatorname{rg}(\phi) = \operatorname{rg}(\operatorname{Mat}_{\mathcal{B}}(\phi)),$$

où $\mathcal{B}$ est n'importe quelle base de E. Le fait que cet entier ne dépend pas du choix de $\mathcal{B}$ est une conséquence de la formule de changement de base ci-dessus.

Remarque 3.15. Étant donnée qu'il y a une correspondance bijective entre les formes quadratiques et les formes bilinéaire symétriques sur E (cf. section 3.1), on fera parfois l'abus d'écrire

- $\operatorname{Ker}(q)$ au lieu de $\operatorname{Ker}(\phi)$;
- $\operatorname{rg}(q)$ au lieu de $\operatorname{rg}(\phi)$; et
- $\operatorname{Mat}_{\mathcal{B}}(q)$ au lieu de $\operatorname{Mat}_{\mathcal{B}}(\phi)$, lorsque $\mathcal{B}$ est une base de E.

Soit ϕ une forme bilinéaire symétrique sur E, et soit $\mathcal{B}$ une base de E. Étant donné que $\operatorname{Ker}(\phi) = \operatorname{Ker}(\operatorname{Mat}_{\mathcal{B}}(\phi))$ et $\operatorname{rg}(\phi) = \operatorname{rg}(\operatorname{Mat}_{\mathcal{B}}(\phi))$, le théorème du rang pour les matrices nous donne un *théorème du rang pour les formes bilinéaire symétriques* :

$$\dim(\operatorname{Ker}(\phi)) + \operatorname{rg}(\phi) = \dim(E).$$

En particulier, la forme bilinéaire symétrique ϕ est non-dégénérée si et seulement si $\operatorname{rg}(\phi) = \dim(E)$.

Exemples 3.16.

(i) Si $E = \mathbb{K}^n$, $\mathcal{B}$ est la base canonique de E, et

$$\phi_1 \colon E \times E \to \mathbb{K},\ (x, y) \mapsto \sum_{i=1}^{n} x_i y_i,$$

alors $\operatorname{Mat}_{\mathcal{B}}(\phi_1) = I_n$. En particulier, $\operatorname{Ker}(\phi_1) = \{0\}$ et $\operatorname{rg}(\phi_1) = n$.

(ii) Si $E = \mathbb{K}[X]_{\leq 2}$, $\mathcal{B} = \{1, X, X^2\}$, et

$$\phi_2 \colon E \to \mathbb{K},\ (P, Q) \mapsto \int_0^1 (P(t)Q'(t) + P'(t)Q(t))dt,$$

alors $\operatorname{Mat}_{\mathcal{B}}(\phi_2) = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$. On vérifie que $\operatorname{Ker}(\phi_2) = \operatorname{Vect}\left(\begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}\right)$ et $\operatorname{rg}(\phi_2) = 2$.

Définition 3.17. Deux matrices $M_1, M_2 \in \mathcal{M}_n(\mathbb{K})$ sont dites *congruentes* s'il existe $P \in \mathrm{GL}_n(\mathbb{K})$ telle que $M_2 = {}^tPM_1P$. C'est une relation d'équivalence (à vérifier !) sur $\mathcal{M}_n(\mathbb{K})$ dont les classes d'équivalence sont appelées *classes de congruence* sur $\mathcal{M}_n(\mathbb{K})$.

De plus, la relation de congruence se restreint au $\mathbb{K}$-sous-espace vectoriel des matrices (anti)symétriques dans $\mathcal{M}_n(\mathbb{K})$. Autrement dit, si M_1 est (anti)symétrique et M_2 est congruente à M_1, alors M_2 est également (anti)symétrique.

Warning 3.18. Il faut veiller à bien distinguer les trois relations d'équivalence suivantes sur les matrices : matrices semblables (ou conjuguées), matrices congruentes, matrices équivalentes.

Le corollaire qui suit est une conséquence immédiate de la formule de changement de base (lemme 3.11) et de la définition de matrices congruentes (définition 3.17).

Corollaire 3.19. Soient $\mathcal{B}$ et $\mathcal{B}'$ deux bases de E et soit ϕ une forme bilinéaire symétrique sur E. Alors les matrices $\operatorname{Mat}_{\mathcal{B}}(\phi)$ et $\operatorname{Mat}_{\mathcal{B}'}(\phi)$ sont congruentes.

Nous verrons dans la section 3.5 que le problème de la classification des classes d'équivalence des formes quadratiques sur E (au sens de la définition 3.30) se ramène au problème de la classification des classes de congruence des matrices symétriques dans $\mathcal{M}_n(\mathbb{K})$.

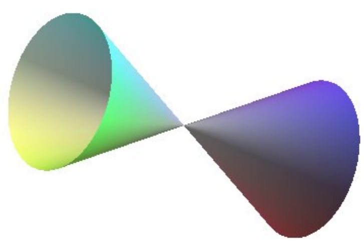
3.3 Orthogonalité et isotropie

Définitions 3.20. Soit q une forme quadratique sur E de forme polaire ϕ.

- Le *ϕ-orthogonal* (ou *q-orthogonal*) d'un $\mathbb{K}$-sous-espace vectoriel F de E est le $\mathbb{K}$-sous-espace vectoriel
$$F^{\perp} = \{y \in E \mid \forall x \in F, \phi(x,y) = 0\}.$$
- Un vecteur $x \in E$ est dit *isotrope* pour q si $q(x) = \phi(x,x) = 0$, autrement dit si $x \in Vect(x)^{\perp}$. L'ensemble des vecteurs isotropes pour q est un cône de E, appelé *cône isotrope de q*, qui contient $\mathrm{Ker}(q)$. La forme quadratique q est dite *définie* si son cône isotrope est trivial.
- Une base $\mathcal{B}$ de E est dite *ϕ-orthogonale* (ou *q-orthogonale*) si pour tous $i,j \in \{1,\dots,n\}$, on a $\phi(e_i,e_j) = 0$ lorsque $i \neq j$.

Remarque 3.21. Une base $\mathcal{B}$ de E est ϕ-orthogonale si et seulement si la matrice $\mathrm{Mat}_{\mathcal{B}}(\phi)$ est diagonale.

Exemple 3.22. On considère la forme quadratique non-dégénérée
$$q\colon \mathbb{R}^3 \to \mathbb{R},\ (x_1,x_2,x_3) \mapsto x_1^2 - x_2^2 + x_3^2.$$
Alors $x = (x_1,x_2,x_3) \in \mathbb{R}^3$ est un vecteur isotrope de q si et seulement si $x_1^2 - x_2^2 + x_3^2 = 0$. L'ensemble des vecteurs isotropes de q est donc un cône de révolution d'axe dirigé par $\begin{bmatrix}0\\1\\0\end{bmatrix}$ et d'angle $\frac{\pi}{4}$.



Par ailleurs, si $F = \mathrm{Vect}\left(\begin{bmatrix}1\\1\\0\end{bmatrix}\right)$, alors on vérifie que $F^{\perp} = \mathrm{Vect}\left(\begin{bmatrix}1\\1\\0\end{bmatrix},\begin{bmatrix}0\\0\\1\end{bmatrix}\right)$ qui est un $\mathbb{R}$-sous espace vectoriel de $\mathbb{R}^3$ qui contient F. En particulier, $F \cap F^{\perp} = F \neq \{0\}$.

Théorème 3.23. Toute forme quadratique $q\colon E \to \mathbb{K}$, de forme polaire ϕ, admet une base ϕ-orthogonale.

Démonstration. Si q est la forme quadratique nulle sur E, alors n'importe quelle base de E convient. Sinon on fixe $x \in E$ tel que $q(x) \neq 0$, et soit $F = \mathrm{Vect}(x)^{\perp}$ qui est un hyperplan de E. En effet, $F = \mathrm{Ker}(\phi(x,-))$ est le noyau d'une forme linéaire non-nulle puisque $\phi(x,x) \neq 0$. On en déduit une décomposition de E en somme directe ϕ-orthogonale $E = \mathrm{Vect}(x) \oplus^{\perp} F$. On va donc pouvoir démontrer le théorème par récurrence sur $n = \dim(E)$:

- Si $n = 1$, alors n'importe quelle base de E convient.
- Si $n \geq 2$ et que le théorème est vrai jusqu'en dimension $n-1$, alors la forme quadratique $q_{|F}\colon F \to \mathbb{K}$ admet une base $\phi_{|F\times F}$-orthogonale $\{f_1,\dots,f_{n-1}\}$. Il s'ensuit que $\{x, f_1,\dots,f_{n-1}\}$ est une base ϕ-orthogonale de $E = \mathrm{Vect}(x) \oplus^{\perp} F$. $\square$

Nous verrons dans la prochaine section comment, à l'aide de la réduction de Gauss, construire une base ϕ-orthogonale de E pour une forme quadratique donnée q sur E.

L'existence d'une base ϕ-orthogonale se traduit matriciellement par le résultat suivant :

Corollaire 3.24 (Version matricielle)**.** Toute matrice symétrique est congruente à une matrice diagonale.

Démonstration. On pose $E = \mathbb{K}^n$ muni de sa base canonique, notée $\mathcal{B}$, et soit $M \in \mathcal{M}_n(\mathbb{K})$ une matrice symétrique. On note ϕ la forme bilinéaire symétrique sur E telle que $\mathrm{Mat}_{\mathcal{B}}(\phi) = M$. D'après le théorème 3.23, l'espace vectoriel E admet une base $\mathcal{B}'$ telle que $D := \mathrm{Mat}_{\mathcal{B}'}(\phi)$ est diagonale. Donc, en notant $P = \mathrm{Pass}(\mathcal{B} \to \mathcal{B}')$, on obtient

$$^tPMP = {}^tP\mathrm{Mat}_{\mathcal{B}}(\phi)P = \mathrm{Mat}_{\mathcal{B}'}(\phi) = D,$$

autrement dit, la matrice symétrique M est congruente à la matrice diagonale D. □

Remarque 3.25. Lorsque $\mathbb{K} = \mathbb{R}$, le corollaire 3.24 peut être vu comme une conséquence du *théorème spectral* (cf. théorème 3.46) qui affirme que, pour toute matrice $M \in \mathcal{S}_n(\mathbb{R})$, on peut trouver $P \in \mathrm{O}_n(\mathbb{R})$ telle que $P^{-1}MP = {}^tPMP$ est diagonale. L'intérêt du corollaire 3.24 réside dans le fait qu'il est valide quelque soit le corps $\mathbb{K}$, tandis que le théorème spectral n'est valide que lorsque $\mathbb{K} = \mathbb{R}$.

Pour l'étudiant.e qui souhaite pousser plus loin l'étude des formes quadratiques, de leurs cônes isotropes, et de leurs groupes d'isométries lorsque $\mathbb{K}$ est un corps quelconque, je recommande la lecture des chapitres V–VIII du livre "Cours d'algèbre" de Daniel Perrin.

3.4 Réduction de Gauss

Soit $q: E \to \mathbb{K}$ une forme quadratique sur E de forme polaire ϕ. Pour déterminer une base ϕ-orthogonale de E, la méthode de Gram-Schmidt ne s'applique pas en raison de la présence possible de vecteurs isotropes au cours de l'une des étapes de l'algorithme comme l'illustre l'exemple suivant.

Exemple 3.26. Soit la forme quadratique

$$q: \mathbb{K}^2 \to \mathbb{K},\ x = (x_1, x_2) \mapsto x_1^2 - x_2^2,$$

de forme polaire $\phi: (x, y) \mapsto x_1y_1 - x_2y_2$. On note $\{e_1, e_2\}$ la base canonique de $\mathbb{K}^2$. Soient $v_1 = e_1 + e_2$ et $v_2 = e_1 - e_2$, alors $\{v_1, v_2\}$ est une base de $\mathbb{K}^2$ formée de vecteurs isotropes pour q. Essayons d'appliquer la méthode de Gram-Schmidt pour orthogonaliser la base $\{v_1, v_2\}$: On cherche alors $\alpha \in \mathbb{K}$ tel que $\phi(v_1, w_2) = 0$ avec $w_2 = v_2 + \alpha v_1$. Or, pour tout $\alpha \in \mathbb{K}$, on a

$$\phi(v_1, w_2) = \phi(v_1, v_2) + \alpha q(v_1) = \phi(v_1, v_2) = 2 \neq 0.$$

On a donc recours à une autre méthode, la *réduction en une combinaison linéaire de carrés de formes linéaires linéairement indépendantes*, due au mathématicien Carl Friedrich Gauss et que nous allons maintenant détailler.

On se donne une forme quadratique q sur E, et on fixe une base $\mathcal{B}$ de E. Alors, d'après le lemme 3.7, pour tout $x \in E$, on peut écrire $q(x)$ comme un polynôme homogène de degré 2 en les coefficients $(x_1, \dots, x_n)$ de x dans la base $\mathcal{B}$. On a donc

$$\forall x \in E,\ q(x) = \sum_{i=1}^{n} a_{i,i} x_i^2 + \sum_{1 \le i < j \le n} a_{i,j} x_i x_j,$$

avec $a_{i,j} \in \mathbb{K}$ pour tous $1 \le i \le j \le n$.

En procédant par récurrence sur n, nous allons écrire q comme combinaison linéaire de carrés de formes linéaires linéairement indépendantes. Lorsque $n = 1$, alors $q(x) = a_{1,1} x^2$ est un multiple (éventuellement nul) du carré de la forme linéaire $x \mapsto x$, et on a donc le résultat souhaité.

On suppose maintenant que $n \ge 2$. On distingue deux cas :

(i) **Il existe au moins un indice i tel que $a_{i,i} \neq 0$.** Quitte à permuter les indices, on peut supposer que $a_{1,1} \neq 0$. On peut alors réécrire $q(x)$ comme

$$q(x) = a_{1,1} x_1^2 + x_1 \epsilon_0(x_2, \dots, x_n) + q_0(x_2, \dots, x_n),$$

où $\epsilon_0 \colon \mathbb{K}^{n-1} \to \mathbb{K}$ est une forme linéaire en $(x_2, \dots, x_n)$ et $q_0 \colon \mathbb{K}^{n-1} \to \mathbb{K}$ est une forme quadratique en $(x_2, \dots, x_n)$. Il s'ensuit que

$$q(x) = a_{1,1}\left(x_1 + \frac{1}{2a_{1,1}} \epsilon_0(x_2, \dots, x_n)\right)^2 + \left(q_0(x_2, \dots, x_n) - \frac{1}{4a_{1,1}} \epsilon_0(x_2, \dots, x_n)^2\right).$$

En posant

$$\epsilon_1 \colon \mathbb{K}^n \to \mathbb{K},\ (x_1, \dots, x_n) \mapsto x_1 + \frac{1}{2a_{1,1}} \epsilon_0(x_2, \dots, x_n),$$

qui est une forme linéaire qui dépend de x_1, et

$$q_1 \colon \mathbb{K}^{n-1} \to \mathbb{K},\ (x_2, \dots, x_n) \mapsto q_0(x_2, \dots, x_n) - \frac{1}{4a_{1,1}} \epsilon_0(x_2, \dots, x_n)^2,$$

qui est une forme quadratique qui ne dépend pas de x_1, on obtient

$$q(x_1, \dots, x_n) = a_{1,1} \epsilon_1(x_1, \dots, x_n)^2 + q_1(x_2, \dots, x_n).$$

On réitère alors la méthode de Gauss en remplaçant q par q_1 qui est une forme quadratique sur $\mathbb{K}^{n-1}$, et après un nombre fini d'étapes on obtient finalement la réduction souhaitée.

(ii) **Pour tout $i \in \{1, \dots, n\}$, on a $a_{i,i} = 0$.** Si q est nulle alors il n'y a rien à faire. Sinon, il existe au moins un $a_{i,j} \neq 0$ avec $1 \le i < j \le n$. Quitte à permuter les indices, on peut supposer que $a_{1,2} \neq 0$. On peut alors écrire $q(x)$ sous la forme

$$q(x) = a_{1,2} x_1 x_2 + x_1 \epsilon_2(x_3, \dots, x_n) + x_2 \epsilon_3(x_3, \dots, x_n) + q_2(x_3, \dots, x_n),$$

où ϵ_2 et ϵ_3 sont des formes linéaires en $(x_3, \dots, x_n)$ et q_2 est une forme quadratique en $(x_3, \dots, x_n)$. On vérifie alors que l'on a les égalités suivantes :

$$\begin{aligned} q(x) &= a_{1,2}\left(x_1 + \frac{1}{a_{1,2}} \epsilon_3\right)\left(x_2 + \frac{1}{a_{1,2}} \epsilon_2\right) + \left(q_2 - \frac{1}{a_{1,2}} \epsilon_2 \epsilon_3\right) \\ &= \frac{a_{1,2}}{4}\left(\left(x_1 + x_2 + \frac{1}{a_{1,2}}(\epsilon_2 + \epsilon_3)\right)^2 - \left(x_1 - x_2 + \frac{1}{a_{1,2}}(\epsilon_3 - \epsilon_2)\right)^2\right) + \left(q_2 - \frac{1}{a_{1,2}} \epsilon_2 \epsilon_3\right) \end{aligned}$$

En posant
$$\epsilon_4 : \mathbb{K}^n \to \mathbb{K},\ (x_1, \dots, x_n) \mapsto x_1 + x_2 + \frac{1}{a_{1,2}}(\epsilon_2 + \epsilon_3)$$
et
$$\epsilon_5 : \mathbb{K}^n \to \mathbb{K},\ (x_1, \dots, x_n) \mapsto x_1 - x_2 + \frac{1}{a_{1,2}}(\epsilon_3 - \epsilon_2),$$
qui sont des formes linéaires linéairement indépendantes (qui dépendent de x_1), et
$$q_3 : \mathbb{K}^{n-2} \to \mathbb{K},\ (x_3, \dots, x_n) \mapsto q_2 - \frac{1}{a_{1,2}}\epsilon_2\epsilon_3,$$
qui est une forme quadratique qui ne dépend pas de x_1, x_2, on obtient
$$q(x_1, \dots, x_n) = \frac{a_{1,2}}{4}(\epsilon_4(x_1, \dots, x_n)^2 - \epsilon_5(x_1, \dots, x_n)^2) + q_3(x_3, \dots, x_n).$$
On réitère alors la méthode de Gauss en remplaçant q par q_3 qui est une forme quadratique sur $\mathbb{K}^{n-2}$, et après un nombre fini d'étapes on obtient finalement la réduction souhaitée.

Théorème 3.27. Soit q une forme quadratique sur E, de forme polaire ϕ. On note $r := \mathrm{rg}(\phi)$. Alors il existe r formes linéaires linéairement indépendantes $\ell_1, \dots, \ell_r \in E^*$, et r scalaires non-nuls $\lambda_1, \dots, \lambda_r \in \mathbb{K}^*$, tels que
$$\forall x \in E,\ q(x) = \sum_{j=1}^{r} \lambda_j (\ell_j(x))^2.$$
De plus, si l'on complète la famille libre $\ell_1, \dots, \ell_r$ en une base $\mathcal{C}$ de E^*, et que l'on note $\mathcal{B} = \{b_1, \dots, b_n\}$ la base antéduale de $\mathcal{C}$, alors la matrice $\mathrm{Mat}_{\mathcal{B}}(\phi)$ est diagonale, autrement dit $\mathcal{B}$ est une base ϕ-orthogonale. Enfin, on a
$$\mathrm{Ker}(\phi) = \mathrm{Vect}(b_{r+1}, \dots, b_n) = \mathrm{Vect}(\ell_1, \dots, \ell_r)^\circ.$$

Démonstration. Si l'on applique la réduction de Gauss à la forme quadratique q, on obtient l'existence de formes linéaires linéairement indépendantes $\ell_1, \dots, \ell_t \in E^*$, et de scalaires non-nuls $\lambda_1, \dots, \lambda_t \in \mathbb{K}^*$, avec $t \in \{1, \dots, n\}$, tels que
$$\forall x \in E,\ q(x) = \sum_{j=1}^{t} \lambda_j (\ell_j(x))^2.$$
On vérifie alors que la forme polaire ϕ de q est donnée par
$$\forall x, y \in E,\ \phi(x, y) = \sum_{j=1}^{t} \lambda_j \ell_j(x) \ell_j(y).$$
Si l'on complète la famille libre $\ell_1, \dots, \ell_t$ en une base $\mathcal{C}$ de E^*, et que l'on note $\mathcal{B} = \{b_1, \dots, b_n\}$ la base antéduale de $\mathcal{C}$, on obtient alors que
$$\forall r, s \in \{1, \dots, n\},\ \phi(b_r, b_s) = \sum_{j=1}^{t} \lambda_j \ell_j(b_r) \ell_j(b_s) = \begin{cases} \lambda_s & \text{si } r = s \in \{1, \dots, t\}\ ;\ \text{et} \\ 0 & \text{sinon.} \end{cases}$$

Donc $\mathrm{Mat}_{\mathcal{B}}(\phi) = \mathrm{diag}(\lambda_1, \dots, \lambda_t, 0, \dots, 0)$. Or le rang de cette dernière matrice est t puisque tous les λ_j avec $j \in \{1, \dots, t\}$ sont non-nuls. On a donc

$$r = \mathrm{rg}(\phi) = \mathrm{rg}(\mathrm{Mat}_{\mathcal{B}}(\phi)) = \mathrm{rg}(\mathrm{diag}(\lambda_1, \dots, \lambda_t, 0, \dots, 0)) = t.$$

Enfin, d'après le lemme 3.10, on a

$$\mathrm{Ker}(\phi) = \mathrm{Ker}(\mathrm{Mat}_{\mathcal{B}}(\phi)) = \mathrm{Ker}(\mathrm{diag}(\lambda_1, \dots, \lambda_r, 0, \dots, 0)) = \mathrm{Vect}(b_{r+1}, \dots, b_n) = \mathrm{Vect}(\ell_1, \dots, \ell_r)^\circ,$$

où la dernière égalité découle du fait que $\mathcal{B}$ est la base antéduale de $\mathbb{C}$. $\square$

Remarque 3.28. Concernant l'énoncé du théorème 3.27 :

- Si $\mathbb{K} = \mathbb{C}$, alors on peut toujours se ramener au cas où tous les λ_i sont égaux à 1.
- Si $\mathbb{K} = \mathbb{R}$, alors on peut toujours se ramener au cas où tous les λ_i sont égaux à ± 1.
- Si $\mathbb{K} = \mathbb{Q}$, alors on peut toujours se ramener au cas où toues les λ_i sont des entiers relatifs sans facteurs carrés.

Exemples 3.29.

(i) On considère la forme quadratique

$$q\colon \mathbb{K}^3 \to \mathbb{K}, (x_1, x_2, x_3) \mapsto x_1^2 + 2x_2^2 + 5x_3^2 + 2x_1x_2 - 4x_2x_3.$$

On est dans le cas (i), et donc la méthode de Gauss donne

$$\begin{aligned} q(x) &= (x_1 + x_2)^2 + (2x_2^2 + 5x_3^2 - 4x_2x_3 - x_2^2) \\ &= (x_1 + x_2)^2 + (x_2 - 2x_3)^2 + x_3^2 \\ &= (\ell_1(x))^2 + (\ell_2(x))^2 + (\ell_3(x))^2. \end{aligned}$$

En particulier, on a $\mathrm{rg}(q) = 3$. La famille $\mathcal{C} = \{\ell_1, \ell_2, \ell_3\}$ est une base de $(\mathbb{K}^3)^*$, dont la base antéduale est $\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -2 \\ 2 \\ 1 \end{bmatrix} \right\}$, où les vecteurs de $\mathcal{B}$ sont exprimés dans la base canonique de $\mathbb{K}^3$, et on vérifie que $\mathrm{Mat}_{\mathcal{B}}(q) = I_3$.

(ii) On considère la forme quadratique

$$q\colon \mathbb{K}^4 \to \mathbb{K}, (x_1, x_2, x_3, x_4) \mapsto 5x_1x_2 + 6x_1x_3 + 3x_2x_3.$$

On est dans le cas (ii), et donc la méthode de Gauss donne

$$\begin{aligned} q(x) &= \frac{5}{4}((x_1 + x_2 + \frac{9}{5}x_3)^2 - (x_1 - x_2 - \frac{3}{5}x_3)^2) - \frac{18}{5}x_3^2 \\ &= \frac{5}{4}(\ell_1(x))^2 - \frac{5}{4}(\ell_2(x))^2 - \frac{18}{5}(\ell_3(x))^2. \end{aligned}$$

En particulier, on a $\mathrm{rg}(q) = 3$. On complète la famille libre $\{\ell_1, \ell_2, \ell_3\}$ en une base $\mathcal{C} = \{\ell_1, \ell_2, \ell_3, \ell_4\}$ de $(\mathbb{K}^4)^*$ en posant par exemple $\ell_4(x) = x_4$. Alors la base antéduale de $\mathcal{C}$ est $\mathcal{B} = \left\{ \frac{1}{2}\begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \frac{1}{2}\begin{bmatrix} 1 \\ -1 \\ 0 \\ 0 \end{bmatrix}, \frac{1}{5}\begin{bmatrix} -3 \\ -6 \\ 5 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \right\}$ où les vecteurs de $\mathcal{B}$ sont exprimés dans la base canonique de $\mathbb{K}^4$. Enfin, on vérifie que $\mathrm{Mat}_{\mathcal{B}}(q) = \mathrm{diag}(\frac{5}{4}, \frac{-5}{4}, \frac{-18}{5}, 0)$, et donc $\mathrm{Ker}(q) = \mathrm{Vect}\left(\begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \right)$.

3.5 Classification des formes quadratiques

Dans cette section, on s'intéresse à la classification des formes quadratiques sur un $\mathbb{K}$-espace vectoriel E de dimension finie $n \in \mathbb{N}$.

3.5.A Généralités

Définition 3.30. Deux formes quadratiques q_1 et q_2 sur E sont dites *équivalentes* s'il existe $u \in \mathcal{GL}(E)$ tel que, pour tout $x \in E$, on ait $q_2(x) = q_1(u(x))$.

On fixe une base $\mathcal{B}$ de E, et on note ϕ_1 et ϕ_2 les formes polaires de q_1 et q_2 respectivement. Alors q_1 et q_2 sont équivalentes si et seulement si $\mathrm{Mat}_{\mathcal{B}}(\phi_1)$ et $\mathrm{Mat}_{\mathcal{B}}(\phi_2)$ sont congruentes. En effet, si l'on pose $X = \mathrm{Mat}_{\mathcal{B}}(x)$ et $U = \mathrm{Mat}_{\mathcal{B}}(u)$, alors

$$\begin{aligned}\forall x \in E,\ q_2(x) = q_1(u(x)) &\Leftrightarrow \forall X \in \mathcal{M}_{n,1}(\mathbb{K}),\ {}^t X \,\mathrm{Mat}_{\mathcal{B}}(\phi_2)\, X = {}^t(UX)\,\mathrm{Mat}_{\mathcal{B}}(\phi_1)\, UX \\ &\Leftrightarrow \mathrm{Mat}_{\mathcal{B}}(\phi_2) = {}^t U \,\mathrm{Mat}_{\mathcal{B}}(\phi_1)\, U.\end{aligned}$$

Reformulation équivalente : Les deux formes quadratiques q_1 et q_2 sont équivalentes s'il existe des bases $\mathcal{B}_1$ et $\mathcal{B}_2$ de E telles que $\mathrm{Mat}_{\mathcal{B}_1}(\phi_1) = \mathrm{Mat}_{\mathcal{B}_2}(\phi_2)$.

En effet, avec les notation précédentes, on a

$$\begin{aligned}\forall x \in E,\ q_2(x) = q_1(u(x)) &\Leftrightarrow \mathrm{Mat}_{\mathcal{B}}(\phi_2) = {}^t U \,\mathrm{Mat}_{\mathcal{B}}(\phi_1)\, U \\ &\Leftrightarrow \mathrm{Mat}_{\mathcal{B}_2}(\phi_2) = \mathrm{Mat}_{\mathcal{B}_1}(\phi_1)\end{aligned}$$

en posant $\mathcal{B}_2 = \mathcal{B}$ et $\mathcal{B}_1$ la base de E telle que $\mathrm{Pass}(\mathcal{B} \to \mathcal{B}_1) = U$.

En résumé, il y a donc une correspondance bijective entre les classes d'équivalence des formes quadratiques sur E d'une part, et les classes de congruence des matrices symétriques dans $\mathcal{M}_n(\mathbb{K})$ d'autre part. Cette correspondance est donnée par l'application $q \mapsto \mathrm{Mat}_{\mathcal{B}}(\phi)$, où ϕ est la forme polaire de q et $\mathcal{B}$ est n'importe quelle base de E préalablement fixée.

Dans ce qui suit, nous allons nous intéresser à la classification des formes quadratiques en termes d'invariants numériques (par exemple le rang d'une forme quadratique est un tel invariant). Autrement dit, nous allons chercher à déterminer des invariants numériques tels que deux formes quadratiques sont équivalentes si et seulement si elles ont les mêmes invariants numériques.

3.5.B Cas où $\mathbb{K} = \mathbb{C}$

Théorème 3.31. Si $\mathbb{K}$ est algébriquement clos (par exemple si $\mathbb{K} = \mathbb{C}$), alors deux formes quadratiques q_1 et q_2 sur E sont équivalentes si et seulement si $\mathrm{rg}(q_1) = \mathrm{rg}(q_2)$.

Démonstration. On fixe une base $\mathcal{B}$ de E. On note ϕ_1 et ϕ_2 les formes polaires de q_1 et q_2 respectivement.

($\Rightarrow$) : Si q_1 et q_2 sont équivalentes, alors il existe $U \in \mathrm{GL}_n(\mathbb{K})$ telle que $\mathrm{Mat}_{\mathcal{B}}(\phi_2) = {}^tU\mathrm{Mat}_{\mathcal{B}}(\phi_1)U$, et on a les égalités suivantes :

$$\mathrm{rg}(q_2) = \mathrm{rg}(\mathrm{Mat}_{\mathcal{B}}(\phi_2)) = \mathrm{rg}({}^t U \,\mathrm{Mat}_{\mathcal{B}}(\phi_1) U) = \mathrm{rg}(\mathrm{Mat}_{\mathcal{B}}(\phi_1)) = \mathrm{rg}(q_1).$$

($\Leftarrow$) : Soit q une forme quadratique sur E, de forme polaire ϕ, et on note $r := \mathrm{rg}(q)$. Nous allons d'abord montrer que $\mathrm{Mat}_{\mathcal{B}}(\phi)$ est congrue à la matrice $J_r := \begin{bmatrix} I_r & 0_{r,n-r} \\ 0_{n-r,r} & 0_{n-r,n-r} \end{bmatrix}$.

Soit $\mathcal{B}'$ une base ϕ-orthogonale de E (qui existe toujours d'après le théorème 3.23), et soit $P = \mathrm{Pass}(\mathcal{B} \to \mathcal{B}')$. Alors, quitte à permuter les vecteurs de la base $\mathcal{B}'$, on a

$$\mathrm{Mat}_{\mathcal{B}}(\phi) = {}^tP\mathrm{Mat}_{\mathcal{B}'}(\phi)P = {}^tP\mathrm{diag}(\mu_1, \dots, \mu_r, 0, \dots, 0)P, \quad \text{avec } \mu_1, \dots, \mu_r \in \mathbb{K}^*.$$

Comme $\mathbb{K}$ est algébriquement clos, pour tout $i \in \{1, \dots, r\}$, l'équation $X^2 - \mu_i$ admet une solution (ce qui ne serait pas vrai pour $\mathbb{K} = \mathbb{R}$). Soit $\gamma_i \in \mathbb{K}$ tel que $\gamma_i^2 = \mu_i$ pour tout $i \in \{1, \dots, r\}$. Alors, en notant $R = \mathrm{diag}(\gamma_1, \dots, \gamma_r, 1, \dots, 1)$, on voit que

$$\mathrm{Mat}_{\mathcal{B}}(\phi) = {}^tP\mathrm{diag}(\mu_1, \dots, \mu_r, 0, \dots, 0)P = {}^tP{}^tRJ_rRP = {}^t(RP)J_rRP.$$

Autrement dit, les matrices $\mathrm{Mat}_{\mathcal{B}}(\phi)$ et J_r sont congruentes.

Soient maintenant q_1 et q_2 deux formes quadratiques sur E de même rang r. Alors, d'après ce qui précède, les matrices $\mathrm{Mat}_{\mathcal{B}}(\phi_1)$ et $\mathrm{Mat}_{\mathcal{B}}(\phi_2)$ sont toutes les deux congruentes à J_r, donc congruentes entres elles (puisque la relation de congruence est transitive), et donc q_1 et q_2 sont équivalentes. □

Corollaire 3.32. Si $\mathbb{K}$ est algébriquement clos (par exemple si $\mathbb{K} = \mathbb{C}$), alors il y a $n+1$ classes d'équivalence de formes quadratiques sur E, et donc $n+1$ classes de congruence dans $\mathcal{S}_n(\mathbb{C})$.

Démonstration. D'après le théorème 3.31, l'application qui à une forme quadratique q sur E associe son rang $\mathrm{rg}(q) \in \{0, \dots, n\}$ induit une bijection entre l'ensemble des classes d'équivalence de formes quadratiques sur E et l'ensemble $\{0, \dots, n\}$.

La dernière assertion du corollaire est une conséquence de la correspondance bijective entre classes d'équivalence de formes quadratiques sur E et classes de congruences dans $\mathcal{S}_n(\mathbb{C})$. □

Exemples 3.33.

(i) Les formes quadratiques

$$\begin{aligned} q_0&: \mathbb{C}^2 \to \mathbb{C},\ (x_1, x_2) \mapsto 0 \\ q_1&: \mathbb{C}^2 \to \mathbb{C},\ (x_1, x_2) \mapsto x_1^2 \\ q_2&: \mathbb{C}^2 \to \mathbb{C},\ (x_1, x_2) \mapsto x_1^2 + x_2^2 \end{aligned}$$

sont deux à deux non équivalentes, et toute forme quadratique sur $\mathbb{C}^2$ est équivalente à l'une d'elles puisque $\mathrm{rg}(q_i) = i$ avec $i \in \{0, 1, 2\}$.

(ii) Les formes quadratiques

$$\begin{aligned} q_a&: \mathbb{C}^2 \to \mathbb{C},\ (x_1, x_2) \mapsto x_1^2 + x_2^2 \\ q_b&: \mathbb{C}^2 \to \mathbb{C},\ (x_1, x_2) \mapsto x_1^2 - x_2^2 \\ q_c&: \mathbb{C}^2 \to \mathbb{C},\ (x_1, x_2) \mapsto -x_1^2 - x_2^2 \end{aligned}$$

sont deux à deux équivalentes car elles ont toutes les trois le même rang.

3.5.C Cas où $\mathbb{K} = \mathbb{R}$

Théorème 3.34 (Loi d'inertie de Sylvester)**.** On suppose que $\mathbb{K} = \mathbb{R}$. Soit q une forme quadratique sur E, de forme polaire ϕ. Il existe une base ϕ-orthogonale $\mathcal{B}_0 = \{b_1, \dots, b_n\}$ de E telle que

$$\forall x \in E,\ q(x) = x_1^2 + \cdots + x_p^2 - x_{p+1}^2 - \cdots - x_r^2,$$

où $(x_1, \dots, x_n)$ sont les composantes de x dans la base $\mathcal{B}_0$, $r = \mathrm{rg}(\phi)$, et $p \in \{1, \dots, r\}$ est un entier qui ne dépend que de q (et pas du choix de la base $\mathcal{B}_0$). Le couple $(p, r-p)$ est appelé la *signature* de la forme quadratique q. Deux formes quadratiques sur E sont équivalentes si et seulement si elles ont la même signature.

Démonstration. Soit $\mathcal{B}$ une base ϕ-orthogonale de E, alors quitte à permuter les éléments de $\mathcal{B}$, on a

$$\mathrm{Mat}_{\mathcal{B}}(\phi) = \mathrm{diag}(\mu_1, \dots, \mu_r, 0, \dots, 0),\ \text{avec } \mu_1, \dots, \mu_r \in \mathbb{K}^*.$$

Pour tout $i \in \{1, \dots, r\}$, on fixe $\alpha_i \in \mathbb{R}^*$ tel que $\alpha_i^2 = |\mu_i|$. Alors

$$\mathrm{diag}(\mu_1, \dots, \mu_r, 0, \dots, 0) = {}^t\mathrm{diag}(\alpha_1, \dots, \alpha_r, 1, \dots, 1)\mathrm{diag}(\delta_1, \dots, \delta_r, 0, \dots, 0)\mathrm{diag}(\alpha_1, \dots, \alpha_r, 1, \dots, 1),$$

avec $\delta_i \in \{\pm 1\}$ pour tout $i \in \{1, \dots, r\}$. Soit $\mathcal{B}_0$ la base de E telle que

$$\mathrm{Pass}(\mathcal{B} \to \mathcal{B}_0) = \mathrm{diag}(\alpha_1, \dots, \alpha_r, 1, \dots, 1).$$

Alors

$$\mathrm{Mat}_{\mathcal{B}_0}(\phi) = \mathrm{diag}(\delta_1, \dots, \delta_r, 0, \dots, 0).$$

Et donc, quitte à permuter les éléments de $\mathcal{B}_0$, on obtient une base ϕ-orthogonale de E telle que

$$\forall x \in E,\ q(x) = x_1^2 + \cdots + x_p^2 - x_{p+1}^2 - \cdots - x_r^2,$$

où $(x_1, \dots, x_n)$ sont les composantes de x dans la base $\mathcal{B}_0$ et $r = \mathrm{rg}(\phi)$.

Montrons maintenant que l'entier p ne dépend pas du choix de la base $\mathcal{B}_0$. On fixe deux bases $\mathcal{B}_1 = \{f_1, \dots, f_n\}$ et $\mathcal{B}_2 = \{g_1, \dots, g_n\}$ de E telles que

$$\forall x \in E,\ q(x) = y_1^2 + \cdots + y_{p_1}^2 - y_{p_1+1}^2 - \cdots - y_r^2,\ \text{avec } p_1 \in \{1, \dots, r\},$$

où $(y_1, \dots, y_n)$ sont les composantes de x dans la base $\mathcal{B}_1$, et

$$\forall x \in E,\ q(x) = z_1^2 + \cdots + z_{p_2}^2 - z_{p_2+1}^2 - \cdots - z_r^2,\ \text{avec } p_2 \in \{1, \dots, r\},$$

où $(z_1, \dots, z_n)$ sont les composantes de x dans la base $\mathcal{B}_2$. Alors

$$\begin{aligned}
&\forall x \in F_1 = Vect(f_1, \dots, f_{p_1}) \setminus \{0\},\ q(x) > 0\ ;\\
&\forall x \in G_1 = Vect(f_{p_1+1}, \dots, f_n),\ q(x) \leq 0\ ;\\
&\forall x \in F_2 = Vect(g_1, \dots, g_{p_2}) \setminus \{0\},\ q(x) > 0\ ;\ \text{et}\\
&\forall x \in G_2 = Vect(g_{p_2+1}, \dots, g_n),\ q(x) \leq 0.
\end{aligned}$$

Il s'ensuit que

$$F_1 \cap G_2 = F_2 \cap G_1 = \{0\}.$$

Et donc
$$p_1 + (n - p_2) = \dim(F_1) + \dim(G_2) \le \dim(E) = n \iff p_1 \le p_2$$
et
$$p_2 + (n - p_1) = \dim(F_2) + \dim(G_1) \le \dim(E) = n \iff p_2 \le p_1,$$
ce qui implique $p_1 = p_2$. Il s'ensuit que l'entier p ne dépend pas du choix de la base $\mathcal{B}_0$.

Il reste à démontrer que deux formes quadratiques sur E sont équivalentes si et seulement si elles ont la même signature. Soient p, s deux entiers tels que $0 \le p + s \le n$. On définit la matrice diagonale par blocs
$$K_{p,s} = \begin{bmatrix} I_p & 0_{p,s} & 0_{p,n-p-s} \\ 0_{s,p} & -I_s & 0_{s,n-p-s} \\ 0_{n-p-s,p} & 0_{n-p-s,s} & 0_{n-p-s,n-p-s} \end{bmatrix} \in \mathcal{M}_n(\mathbb{K}).$$
D'après ce qui précède, pour toute forme quadratique q sur E de signature (p,s), il existe une base $\mathcal{B}_0$ telle que $\mathrm{Mat}_{\mathcal{B}_0}(\phi) = K_{p,s}$. Donc, si deux formes quadratiques q_1 et q_2 sur E ont la même signature (p,s), alors il existe deux bases $\mathcal{B}_1$ et $\mathcal{B}_2$ de E telles que
$$\mathrm{Mat}_{\mathcal{B}_1}(\phi_1) = K_{p,s} = \mathrm{Mat}_{\mathcal{B}_2}(\phi_2),$$
et donc q_1 et q_2 sont équivalentes.

Réciproquement, soient q_1 et q_2 deux formes quadratiques sur E qui sont équivalentes et de signatures respectives (p_1, s_1) et (p_2, s_2). Alors il existe une base $\mathcal{B}_1$ de E et une matrice $P \in \mathrm{GL}_n(\mathbb{K})$ telles que
$$K_{p_1,s_1} = \mathrm{Mat}_{\mathcal{B}_1}(\phi_1) = {}^tP\mathrm{Mat}_{\mathcal{B}_1}(\phi_2)P = \mathrm{Mat}_{\mathcal{B}_2}(\phi_2),$$
où $\mathcal{B}_2$ est la base de E telle que $P = \mathrm{Pass}(\mathcal{B}_1 \to \mathcal{B}_2)$. Et donc (p_1, s_1) est aussi la signature de la forme quadratique q_2, autrement dit $(p_1, s_1) = (p_2, s_2)$. □

Corollaire 3.35. Soit q une forme quadratique sur un $\mathbb{R}$-espace vectoriel E de dimension n.
(i) q est non-dégénérée si et seulement si $\mathrm{sign}(q) = (p, n-p)$ pour un certain $p \in \{1, \dots, n\}$;
(ii) q est définie positive si et seulement si $\mathrm{sign}(q) = (n, 0)$; et
(iii) q est définie négative si et seulement si $\mathrm{sign}(q) = (0, n)$.

Démonstration. En exercice ! □

Corollaire 3.36. Si $\mathbb{K} = \mathbb{R}$, alors il y a $\frac{1}{2}(n+1)(n+2)$ classes d'équivalence de formes quadratiques sur E, et donc $\frac{1}{2}(n+1)(n+2)$ classes de congruence dans $\mathcal{S}_n(\mathbb{R})$.

Démonstration. D'après le théorème 3.34, l'application qui à une forme quadratique q sur E associe sa signature (p,s) induit une bijection entre l'ensemble des classes d'équivalence de formes quadratiques sur E et l'ensemble $\{(p,s),\ 0 \le p+s \le n\}$. Or le cardinal de ce dernier ensemble est
$$\sum_{p+s=0}^{n} 1 = \sum_{r=0}^{n} (r+1) = \sum_{r'=1}^{n+1} r' = \frac{1}{2}(n+1)(n+2).$$

La dernière assertion du corollaire est une conséquence de la correspondance bijective entre classes d'équivalence de formes quadratiques sur E et classes de congruences dans $\mathcal{S}_n(\mathbb{R})$.

□

Exemples 3.37.

(i) Les formes quadratiques

$$\begin{aligned} q_a&: \mathbb{R}^2 \to \mathbb{R},\ (x_1, x_2) \mapsto x_1^2 + x_2^2 \\ q_b&: \mathbb{R}^2 \to \mathbb{R},\ (x_1, x_2) \mapsto x_1^2 - x_2^2 \\ q_c&: \mathbb{R}^2 \to \mathbb{R},\ (x_1, x_2) \mapsto -x_1^2 - x_2^2 \end{aligned}$$

sont deux à deux non-équivalentes car elles ont pour signatures $(2,0)$, $(1,1)$ et $(0,2)$ respectivement. Mais toute forme quadratique non-dégénérée sur $\mathbb{R}^2$ est équivalente à l'une d'elles.

(ii) On considère la forme quadratique

$$q: \mathbb{R}^3 \to \mathbb{R}, (x_1, x_2, x_3) \mapsto x_1^2 + 2x_2 + 15x_3^2 - 4x_1x_2 + 6x_1x_3 - 8x_2x_3.$$

Alors la réduction de Gauss donne

$$q(x) = (x_1 - 2x_2 + 3x_3)^2 - 2(x_2 - x_3)^2 + 8x_3^2,$$

et donc $\operatorname{sign}(q) = (2,1)$. En particulier, q est non-dégénérée.

(iii) On considère la forme quadratique sur $E = \mathbb{R}[X]_{\leq n-1}$ définie par

$$q: E \to \mathbb{R},\ P \mapsto \int_0^1 P^2 - \left(\int_0^1 P\right)^2.$$

Alors on peut vérifier que $\operatorname{sign}(q) = (n-1, 0)$. En particulier, q est dégénérée.

3.5.D Discriminant d'une forme quadratique

Définition 3.38 (Quotient $\mathbb{K}^*/(\mathbb{K}^*)^2$)**.**
Soient $s, t \in \mathbb{K}^*$. On note $s \sim t$ s'il existe $\alpha \in \mathbb{K}^*$ tel que $s = t\alpha^2$. On vérifie alors que $\sim$ est une relation d'équivalence sur $\mathbb{K}^*$, et on note $\mathbb{K}^*/(\mathbb{K}^*)^2$ le groupe quotient correspondant.

Exemples 3.39.

- Si $\mathbb{K} = \mathbb{C}$, alors tout nombre complexe $s \in \mathbb{C}^*$ vérifie $s = \alpha^2$ pour un certain $\alpha \in \mathbb{C}^*$, et donc $s \sim 1$ pour tout $s \in \mathbb{C}^*$. Il s'ensuit que $\mathbb{C}^*/(\mathbb{C}^*)^2 = \{[1]\}$.
- Si $\mathbb{K} = \mathbb{R}$, alors tout nombre réel s vérifie

$$\begin{cases} s = \alpha^2, & \text{avec } \alpha = \sqrt{s}, \quad \text{si } s > 0 \ ; \text{ et} \\ s = -\alpha^2 & \text{avec } \alpha = \sqrt{-s}, \quad \text{si } s < 0. \end{cases}$$

De plus -1 n'est pas un carré dans $\mathbb{R}^*$, et donc $1 \not\sim -1$. Il s'ensuit que $\mathbb{R}^*/(\mathbb{R}^*)^2 = \{[-1], [1]\}$.
- Si $\mathbb{K} = \mathbb{Q}$, alors l'ensemble $\mathbb{Q}^*/(\mathbb{Q}^*)^2$ est infini. En effet, si $p, q \in \mathbb{N}$ sont deux nombres premiers distincts, alors on vérifie que $p \not\sim q$, et donc l'ensemble des nombres premiers, qui est infini, s'injecte dans l'ensemble $\mathbb{Q}^*/(\mathbb{Q}^*)^2$.

Théorème 3.40. Si q est une forme quadratique non-dégénérée sur E, de forme polaire ϕ, alors on peut lui associer la classe d'équivalence $\delta(q) := [\det(\operatorname{Mat}_{\mathcal{B}}(\phi))] \in \mathbb{K}^*/(\mathbb{K}^*)^2$, où $\mathcal{B}$ est n'importe quelle base de E. L'élément $\delta(q)$ ne dépend pas du choix de la base $\mathcal{B}$. De plus, si q_1 et q_2 sont deux formes quadratiques non-dégénérées équivalentes, alors $\delta(q_1) = \delta(q_2)$.

Démonstration. Si $\mathcal{B}_1$ et $\mathcal{B}_2$ sont deux bases de E, alors d'après la formule de changement de base on a

$$\mathrm{Mat}_{\mathcal{B}_2}(\phi) = {}^tP\mathrm{Mat}_{\mathcal{B}_1}(\phi)P, \quad \text{avec } P = \mathrm{Pass}(\mathcal{B}_1 \to \mathcal{B}_2).$$

En passant au déterminant des deux côté de l'égalité, on obtient

$$\det(\mathrm{Mat}_{\mathcal{B}_2}(\phi)) = \det(P)^2 \det(\mathrm{Mat}_{\mathcal{B}_1}(\phi)),$$

d'où

$$[\det(\mathrm{Mat}_{\mathcal{B}_2}(\phi))] = [\det(\mathrm{Mat}_{\mathcal{B}_2}(\phi))] \text{ dans } \mathbb{K}^*/(\mathbb{K}^*)^2.$$

Par conséquent, l'élément $\delta(q)$ ne dépend pas du choix de la base $\mathcal{B}$ de E.

Maintenant, soient q_1 et q_2 deux formes quadratiques non-dégénérées équivalentes, de formes polaires ϕ_1 et ϕ_2 respectivement. Alors, pour toute base $\mathcal{B}$ de E, les matrices $\mathrm{Mat}_{\mathcal{B}}(\phi_1)$ et $\mathrm{Mat}_{\mathcal{B}}(\phi_2)$ sont congruentes, et donc

$$\delta(q_1) = [\det(\mathrm{Mat}_{\mathcal{B}}(\phi_1))] = [\det(\mathrm{Mat}_{\mathcal{B}}(\phi_2))] = \delta(q_2) \quad \text{dans } \mathbb{K}^*/(\mathbb{K}^*)^2.$$

Ainsi, si q_1 et q_2 sont deux formes quadratiques non-dégénérées équivalentes, alors on a bien $\delta(q_1) = \delta(q_2)$. □

Définition 3.41. Soit q une forme quadratique non-dégénérée sur E. L'élément $\delta(q) \in \mathbb{K}^*/(\mathbb{K}^*)^2$, défini dans le théorème 3.40, est appelé le *discriminant* de q.

Le discriminant d'une forme quadratique non-dégénérée est donc un nouvel invariant numérique constant sur chaque classe d'équivalence de formes quadratiques non-dégénérées sur E. Cet invariant présente l'avantage d'être défini quel que soit le corps $\mathbb{K}$, mais il ne suffit pas à classifier les formes quadratiques non-dégénérées sur E. Par exemple, si $\mathbb{K} = \mathbb{R}$, alors $\delta(q)$ peut prendre deux valeurs ($[1]$ et $[-1]$), alors que la loi d'inertie de Sylvester (théorème 3.34) implique qu'il existe $n+1$ classes d'équivalence de formes quadratiques non-dégénérées sur E.

Exemples 3.42.

- Les formes quadratiques non-dégénérées

$$\begin{aligned} q_a&: \mathbb{R}^2 \to \mathbb{R},\ (x_1, x_2) \mapsto x_1^2 + x_2^2 \\ q_b&: \mathbb{R}^2 \to \mathbb{R},\ (x_1, x_2) \mapsto x_1^2 - x_2^2 \\ q_c&: \mathbb{R}^2 \to \mathbb{R},\ (x_1, x_2) \mapsto -x_1^2 - x_2^2 \end{aligned}$$

  ont pour discriminants respectifs $[1]$, $[-1]$ et $[1]$ dans $\mathbb{R}^*/(\mathbb{R}^*)^2 = \{[-1], [1]\}$.
  Le théorème 3.40 implique donc que q_b n'est équivalente ni à q_a ni à q_c, en revanche elle ne permet pas de conclure quant à savoir si q_a et q_c sont équivalentes (il est alors nécessaire de comparer les signatures de q_a et q_c pour conclure).
- On considère les formes quadratiques non-dégénérées

$$\begin{aligned} q_d&: \mathbb{K}^2 \to \mathbb{K},\ (x_1, x_2) \mapsto x_1^2 + x_2^2\ ; \text{ et} \\ q_e&: \mathbb{K}^2 \to \mathbb{K},\ (x_1, x_2) \mapsto x_1^2 + 2x_2^2. \end{aligned}$$

  Si $\mathbb{K} = \mathbb{C}$ ou $\mathbb{K} = \mathbb{R}$, alors q_d et q_e sont équivalentes. En revanche, si $\mathbb{K} = \mathbb{Q}$, alors

$$\delta(q_d) = [1] \neq [2] = \delta(q_e) \text{ dans } \mathbb{Q}^*/(\mathbb{Q}^*)^2 \text{ puisque } \sqrt{2} \notin \mathbb{Q}.$$

  Et donc q_d et q_e ne sont pas équivalentes.

- On considère les formes quadratiques non-dégénérées

$$q_f: \mathbb{Q}^2 \to \mathbb{Q},\ (x_1, x_2) \mapsto 3x_1^2 + 2x_2^2\ ;\ \text{et}$$
$$q_g: \mathbb{Q}^2 \to \mathbb{Q},\ (x_1, x_2) \mapsto 6x_1^2 + x_2^2.$$

  Alors $\delta(q_f) = [6] = \delta(q_g)$ dans $\mathbb{Q}^*/(\mathbb{Q}^*)^2$, mais on peut vérifier (par exemple matriciellement) que q_f et q_g ne sont pas équivalentes sur $\mathbb{Q}^2$.

  Cet exemple illustre à nouveau le fait que le discriminant n'est pas suffisant pour classifier les formes quadratiques non-dégénérées ; il reste néanmoins un outil utile pour montrer que deux formes quadratiques non-dégénérées ne sont pas équivalentes dès lors qu'elles n'ont pas le même discriminant.

Pour l'étudiant.e qui souhaite en savoir plus sur la classification des formes quadratiques lorsque $\mathbb{K} = \mathbb{Q}$ ou lorsque $\mathbb{K}$ est un corps fini, je conseille de consulter le livre "Cours d'arithmétique" de Jean-Pierre Serre.

3.6 Formes quadratiques sur un espace euclidien

⚠ On suppose dorénavant que $\mathbb{K} = \mathbb{R}$ et que E est un espace euclidien de dimension n, c'est-à-dire un $\mathbb{R}$-espace vectoriel muni d'un produit scalaire que l'on notera généralement par $\langle -, - \rangle$.

Remarquons que

$$E \times E \to \mathbb{R},\ (x, y) \mapsto \langle x, y \rangle$$

est une forme bilinéaire symétrique définie positive dont la forme quadratique associée est le carré de la norme euclidienne

$$E \to \mathbb{R},\ x \mapsto \|x\|_2 = \sqrt{\langle x, x \rangle}.$$

Définition 3.43. Soient $f, g \in \mathcal{L}(E)$. On dit que g est *l'adjoint* de f (par rapport au produit scalaire $\langle -, - \rangle$) si

$$\forall x, y \in E,\ \langle f(x), y \rangle = \langle x, g(y) \rangle\,.$$

On note alors $g = {}^t f$. Lorsque $f = {}^t f$, on dit que f est *auto-adjoint*.

Lemme 3.44. Pour tout $f \in \mathcal{L}(E)$, l'endomorphisme adjoint ${}^t f$ existe et il est unique. De plus, si $\mathcal{B}$ est une base orthonormée de $(E, \langle -, - \rangle)$, alors $\mathrm{Mat}_{\mathcal{B}}({}^t f) = {}^t\mathrm{Mat}_{\mathcal{B}}(f)$.

Démonstration. On fixe une base orthonormée $\mathcal{B}$ de $(E, \langle -, - \rangle)$. Alors $\mathrm{Mat}_{\mathcal{B}}(\langle -, - \rangle) = I_n$, et donc en notant $M = \mathrm{Mat}_{\mathcal{B}}(f)$, $X = \mathrm{Mat}_B(x)$ et $Y = \mathrm{Mat}_{\mathcal{B}}(y)$, on a

$$\langle f(x), y \rangle = {}^t(MX) I_n Y = {}^t X {}^t M Y = {}^t X I_n ({}^t M Y) = \left\langle x, {}^t f y \right\rangle,$$

avec ${}^t f \in \mathcal{L}(E)$ l'unique endomorphisme tel que $\mathrm{Mat}_{\mathcal{B}}({}^t f) = {}^t M = {}^t\mathrm{Mat}_{\mathcal{B}}(f)$. □

Remarque 3.45. Soit $\mathcal{B}$ une base orthonormée de $(E, \langle -, - \rangle)$. Alors $f \in \mathcal{L}(E)$ est auto-adjoint si et seulement si la matrice $\mathrm{Mat}_{\mathcal{B}}(f)$ est symétrique.

Le théorème qui suit est classique et sa démonstration a déjà été étudiée l'an dernier. On ne la rappellera donc pas. Mais l'étudiant.e qui souhaite se rafraîchir la mémoire peut consulter la section 5.2.4 du livre "Algèbre" de Xavier Gourdon.

Théorème 3.46 (Théorème spectral)**.** Si $f \in \mathcal{L}(E)$ est auto-adjoint, alors f est diagonalisable dans une base orthonormée de $(E, \langle -, - \rangle)$, et de plus les valeurs propres de f sont réelles.

Corollaire 3.47 (Version matricielle du théorème spectral)**.** Si $M \in \mathcal{S}_n(\mathbb{R})$, alors il existe $P \in \mathrm{O}_n(\mathbb{R})$ telle que

$$P^{-1}MP = {}^tPMP = D,$$

où D est une matrice diagonale réelle.

Le résultat le plus important de cette section 3.6, qui est essentiellement une conséquence du théorème spectral, est le suivant :

Théorème 3.48 (Pseudo-réduction simultanée)**.** Soit q une forme quadratique sur l'espace euclidien $(E, \langle -, - \rangle)$, de forme polaire ϕ. Alors il existe une base $\mathcal{B}$ de E qui est à la fois orthonormée pour $\langle -, - \rangle$ et orthogonale pour ϕ.

Démonstration. Soit $\mathcal{B}_0$ une base orthonormée de $(E, \langle -, - \rangle)$, et soit $M := \mathrm{Mat}_{\mathcal{B}_0}(\phi)$ qui est une matrice symétrique réelle et correspond donc à un endomorphisme auto-adjoint $f \in \mathcal{L}(E)$ défini par $\mathrm{Mat}_{\mathcal{B}_0}(f) = M$. D'après le théorème spectral, il existe une base orthonormée $\mathcal{B}$ telle que $D := \mathrm{Mat}_{\mathcal{B}}(f)$ est diagonale. Alors, en notant $P := \mathrm{Pass}(\mathcal{B}_0 \to \mathcal{B})$, qui est une matrice orthogonale puisque $\mathcal{B}_0$ et $\mathcal{B}$ sont des bases orthonormées, on obtient que

$$\mathrm{Mat}_{\mathcal{B}}(\phi) = {}^tP\mathrm{Mat}_{\mathcal{B}_0}(\phi)P = {}^tPMP = P^{-1}MP = P^{-1}\mathrm{Mat}_{\mathcal{B}_0}(f)P = \mathrm{Mat}_{\mathcal{B}}(f) = D.$$

Et donc $\mathcal{B}$ est une base de E qui est à la fois orthonormée pour $\langle -, - \rangle$ et orthogonale pour ϕ. □

Remarque 3.49. La différence entre le théorème 3.48 et le théorème 3.23 réside dans le fait qu'ici la base $\mathcal{B}$ est non seulement ϕ-orthogonale mais aussi orthonormée pour le produit scalaire $\langle -, - \rangle$ sur l'espace euclidien E.

Corollaire 3.50 (Version matricielle de la pseudo-réduction simultanée)**.**
Soient $Q, M \in \mathcal{S}_n(\mathbb{R})$. On suppose de plus que la matrice Q est définie positive. Alors il existe une matrice $P \in \mathrm{GL}_n(\mathbb{R})$ telle que

$${}^tPQP = I_n \quad \text{et} \quad {}^tPMP = D,$$

où D est une matrice diagonale réelle.

Démonstration. On fixe une base quelconque de $E = \mathbb{R}^n$, que l'on note $\mathcal{B}_0$, et soient ϕ_1 et ϕ_2 les deux formes bilinéaires symétriques sur E telles que $Q = \mathrm{Mat}_{\mathcal{B}_0}(\phi_1)$ et $M = \mathrm{Mat}_{\mathcal{B}_0}(\phi_2)$.

Étant donné que la matrice Q est définie positive, la forme bilinéaire symétrique ϕ_1 est un produit scalaire sur E, et donc (E, ϕ_1) est un espace euclidien. Alors, d'après le théorème 3.48, il existe une base $\mathcal{B}$ de E qui est à la fois orthonormée pour ϕ_1 et orthogonale pour ϕ_2. En notant $P = \mathrm{Pass}(\mathcal{B}_0 \to \mathcal{B}) \in \mathrm{GL}_n(\mathbb{R})$, on a donc bien

$${}^tPQP = {}^tP\mathrm{Mat}_{\mathcal{B}_0}(\phi_1)P = \mathrm{Mat}_{\mathcal{B}}(\phi_1) = I_n$$

et

$${}^tPMP = {}^tP\mathrm{Mat}_{\mathcal{B}_0}(\phi_2)P = \mathrm{Mat}_{\mathcal{B}}(\phi_2) = D,$$

où D est une matrice diagonale réelle. □

Nous allons maintenant indiquer une méthode pour déterminer une base $\mathcal{B}$ de E qui soit à la fois orthonormée pour $\langle -,- \rangle$ et orthogonale pour ϕ. Nous aurons besoin du lemme suivant :

Lemme 3.51. Soit q une forme quadratique, de forme polaire ϕ, sur l'espace euclidien $(E, \langle -,- \rangle)$, et soit $\mathcal{B}_0$ une base orthonormée de $(E, \langle -,- \rangle)$. Alors les espaces propres de la matrice $\mathrm{Mat}_{\mathcal{B}_0}(\phi)$ sont deux à deux orthogonaux pour $\langle -,- \rangle$.

Démonstration. Si la matrice $M := \mathrm{Mat}_{\mathcal{B}_0}(\phi)$ admet une unique valeur propre, alors il n'y à rien à prouver. Sinon, on fixe λ_1 et λ_2 deux valeurs propres distinctes de M. Soient v_1 et v_2 des vecteurs propres (non-nuls) associés aux valeurs propres λ_1 et λ_2 respectivement. Alors

$$\begin{aligned}
\lambda_1 \langle v_1, v_2 \rangle &= \langle \lambda_1 v_1, v_2 \rangle \\
&= \langle M v_1, v_2 \rangle \\
&= \left\langle v_1, {}^t M v_2 \right\rangle \quad \text{car } \mathcal{B}_0 \text{ est orthonormée pour } \langle -,- \rangle \\
&= \langle v_1, M v_2 \rangle \quad \text{car } M \text{ est symétrique} \\
&= \langle v_1, \lambda_2 v_2 \rangle \\
&= \lambda_2 \langle v_1, v_2 \rangle
\end{aligned}$$

Et donc, on obtient

$$(\lambda_1 - \lambda_2) \langle v_1, v_2 \rangle = 0,$$

ce qui implique $\langle v_1, v_2 \rangle = 0$ puisque $\lambda_1 \neq \lambda_2$. Autrement dit, les espaces propres de M sont deux à deux orthogonaux pour $\langle -,- \rangle$. □

Soit q une forme quadratique, de forme polaire ϕ, sur l'espace euclidien $(E, \langle -,- \rangle)$. En pratique, pour déterminer une base $\mathcal{B}$ de E qui soit à la fois orthonormée pour $\langle -,- \rangle$ et orthogonale pour ϕ, on procède de la façon suivante :

(i) On fixe une base orthonormée $\mathcal{B}_0$ de $(E, \langle -,- \rangle)$ et on calcule les espaces propres de la matrice $M := \mathrm{Mat}_{\mathcal{B}_0}(\phi) \in \mathcal{S}_n(\mathbb{R})$.

(ii) Pour chaque espace propre E_λ de M, on détermine une base orthonormée $\mathcal{B}_\lambda$ (par exemple à l'aide de la méthode de Gram-Schmidt).

(iii) La réunion des bases $\mathcal{B}_\lambda$ nous donne une base orthonormée $\mathcal{B}$ de $(E, \langle -,- \rangle)$ qui est également ϕ-orthogonale. En effet,

$$\mathrm{Mat}_{\mathcal{B}}(\phi) = {}^t\mathrm{Pass}(\mathcal{B}_0 \to \mathcal{B}) M \mathrm{Pass}(\mathcal{B}_0 \to \mathcal{B}) = \mathrm{Pass}(\mathcal{B}_0 \to \mathcal{B})^{-1} M \mathrm{Pass}(\mathcal{B}_0 \to \mathcal{B}) = D,$$

où D est une matrice diagonale réelle et $\mathrm{Pass}(\mathcal{B}_0 \to \mathcal{B}) \in \mathrm{O}_n(\mathbb{R})$.

3.7 Étude des coniques et des quadriques

Le théorème de pseudo-réduction simultanée a de nombreuses applications. L'une d'elles est l'étude des coniques dans le plan euclidien et des quadriques dans l'espace euclidien.

Soit $(E, \langle -,- \rangle)$ un espace euclidien de dimension $n \in \mathbb{N}$, soit q une forme quadratique sur E, et soit $r \in \mathbb{R}_+^*$ une constante positive fixée.

Définition 3.52. L'ensemble

$$\Gamma_{q,r} := \{x \in E \mid q(x) = r\}$$

est appelé une *conique* dans E lorsque $\dim(E) = 2$ resp. une *quadrique* dans E lorsque $\dim(E) \geq 3$.

D'après le théorème 3.48, il existe une base $\mathcal{B} = \{b_1, \dots, b_n\}$ de E qui est à la fois orthonormée pour $\langle -, - \rangle$ et q-orthogonale. Ainsi

$$\forall x \in E,\ q(x) = \sum_{i=1}^{n} a_i x_i^2$$

avec $x = \sum_{i=1}^{n} x_i e_i$ et $a_i \in \mathbb{R}$ pour tout $i \in \{1, \dots, n\}$. De plus, quitte à remplacer chaque vecteur b_i de la base $\mathcal{B}$ par un multiple bien choisi, on peut même supposer que $a_i \in \{-1, 0, 1\}$ pour tout $i \in \{1, \dots, n\}$. La base obtenue n'est alors plus une base orthonormée pour $\langle -, - \rangle$, mais elle reste orthogonale simultanément pour $\langle -, - \rangle$ et q, ce qui simplifie souvent les calculs lorsqu'on étudie ces objets du point de vue de la géométrie affine.

En outre, si $u \in \mathcal{L}(E)$, alors on peut vérifier que $u(\Gamma_{q,r})$ a le même type géométrique et la même topologie de $\Gamma_{q,r}$ (seuls les aspects métriques de $\Gamma_{q,r}$ changent). Ainsi, le type géométrique et la topologie de $\Gamma_{q,r}$ ne dépendent que de la signature de la forme quadratique q sur E, et non pas du choix de la base q-orthogonale considérée.
Lorsque $\dim(E) = 2$, le tableau suivant récapitule les différentes possibilités pour la conique $\Gamma_{q,r}$:

signature de q	type géométrique de $\Gamma_{q,r}$	topologie de $\Gamma_{q,r}$
(2,0)	ellipse	compact et connexe
(1,1)	hyperbole	pas compact, deux composantes connexes
(1,0)	réunion de deux droites	pas compact, deux composantes connexes
(0,2), (0,1) ou (0,0)	$\varnothing$	

Lorsque $\dim(E) = 3$, le tableau suivant récapitule les différentes possibilités pour la quadrique $\Gamma_{q,r}$:

signature de q	type géométrique de $\Gamma_{q,r}$	topologie de $\Gamma_{q,r}$
(3,0)	ellipsoïde	compact et connexe
(2,1)	hyperboloïde à une nappe	pas compact, connexe
(1,2)	hyperboloïde à deux nappes	pas compact, deux composantes connexes
(2,0)	cylindre elliptique	pas compact, connexe
(1,1)	cylindre hyperbolique	pas compact, deux composantes connexes
(1,0)	réunion de deux plans	pas compact, deux composantes connexes
(0,3), (0,2), (0,1) ou (0,0)	$\varnothing$	

Quelques références bibliographiques

- Les références les plus élémentaires concernant la théorie des formes bilinéaires symétriques et des formes quadratiques :
  - Rémi Goblot. *Algèbre linéaire.*
  - Xavier Gourdon. *Algèbre*, Chapitre 5.
  - Joseph Grifone. *Algèbre linéaire*, Chapitres 7 et 9.
- Une référence pour aller plus loin dans la théorie des formes quadratiques et l'étude des groupes orthogonaux :
  - Daniel Perrin. *Cours d'algèbre*, Chapitres VI et VIII.
- Une référence pour l'étude des groupes orthogonaux réels généralisés $\mathrm{O}(s,t)$ associés à une forme quadratique de signature (s,t) :
  - Denis Serre. *Les matrices : théorie et pratique*, Chapitre 7.
- Une référence pour la classification des formes quadratiques lorsque $\mathbb{K} = \mathbb{Q}$ ou $\mathbb{K} = \mathbb{F}_q$:
  - Jean-Pierre Serre. *Cours d'arithmétique*, Chapitre IV.1.

Chapitre 4

Formes hermitiennes et espaces hermitiens

Motivation et but

Un *produit scalaire euclidien* est une forme bilinéaire symétrique et définie positive. Par exemple, le produit scalaire euclidien usuel sur $\mathbb{R}^3$ est défini par

$$\phi(x,y) = x_1y_1 + x_2y_2 + x_3y_3,$$

et alors l'application $x \mapsto \sqrt{\phi(x,x)}$ définit une norme sur $\mathbb{R}^3$ qu'on appelle la norme euclidienne usuelle. Malheureusement, la même formule ne donne pas une norme sur $\mathbb{C}^3$ puisque $x_1^2+x_2^2+x_3^2 \notin \mathbb{R}_+$ en général. Pour pallier ce problème, il faut substituer à la notion de produit scalaire euclidien celle de produit scalaire hermitien. Un *produit scalaire hermitien* est une forme sesquilinéaire à symétrie hermitienne et définie positive. Par exemple, le produit scalaire hermitien usuel sur $\mathbb{C}^3$ est défini par

$$\varphi(x,y) = \overline{x_1}y_1 + \overline{x_2}y_2 + \overline{x_3}y_3,$$

et alors l'application $x \mapsto \sqrt{\varphi(x,x)}$ définit une norme sur $\mathbb{C}^3$ qu'on appelle la norme hermitienne usuelle.

Un espace hermitien est un $\mathbb{C}$-espace vectoriel de dimension finie muni d'un produit scalaire hermitien. Cette structure, qui doit son nom au mathématicien français Charles Hermite (1822-1901), est donc l'analogue complexe d'un espace euclidien ; d'ailleurs, de nombreuses propriétés sont communes aux deux structures : existence d'une base orthonormée, algorithme de Gram-Schmidt, réduction des endomorphismes auto-adjoints, etc. En outre, le fait que le corps $\mathbb{C}$ soit algébriquement clos permet d'obtenir des résultats de réduction souvent "plus forts" dans le cadre hermitien que dans le cadre euclidien. Par exemple, on peut toujours diagonaliser les endomorphismes unitaires d'un espace hermitien tandis que l'on peut seulement diagonaliser par blocs les isométries d'un espace euclidien.

Dans ce quatrième chapitre, nous exposons les bases de la théorie des formes hermitiennes et des espaces hermitiens, avec en ligne de mire l'étude des endomorphismes unitaires (cf. section 4.5) et la décomposition polaire des matrices complexes inversibles (cf. section 4.6).

4.1 Formes sesquilinéaires et formes hermitiennes

Soit E un $\mathbb{C}$-espace vectoriel (pas nécessairement de dimension finie pour le moment).

Définitions 4.1.

- Une *forme sesquilinéaire* sur E est une application

$$\varphi: E \times E \to \mathbb{C}, \ (x, y) \mapsto \varphi(x, y)$$

telle que, pour tous $x_0, y_0 \in E$, l'application $\varphi(x_0, -): E \to \mathbb{C}$ est linéaire et l'application $\varphi(-, y_0): E \to \mathbb{C}$ est *antilinéaire* ; cette dernière condition signifie que

$$\forall x_1, x_2 \in E, \ \forall \lambda \in \mathbb{C}, \quad \varphi(\lambda x_1 + x_2, y_0) = \overline{\lambda}\varphi(x_1, y_0) + \varphi(x_2, y_0).$$

On note $\mathcal{S}es(E, \mathbb{C})$ l'ensemble des formes sesquilinéaires sur E ; cet ensemble est naturellement muni d'une structure de $\mathbb{C}$-espace vectoriel. En effet, on a

$$\forall \varphi_1, \varphi_2 \in \mathcal{S}es(E, \mathbb{C}), \ \forall \mu \in \mathbb{C}, \ \mu\varphi_1 + \varphi_2 \in \mathcal{S}es(E, \mathbb{C}).$$

- Une forme sesquilinéaire $\varphi: E \times E \to \mathbb{C}$ est dite *à symétrie hermitienne* si

$$\forall x, y \in E, \ \varphi(y, x) = \overline{\varphi(x, y)}.$$

Alors, pour tout $x \in E$, on a $\varphi(x, x) = \overline{\varphi(x, x)} \in \mathbb{R}$. Les formes sesquilinéaires à symétrie hermitienne forment un $\mathbb{R}$-sous-espace vectoriel de $\mathcal{S}es(E, \mathbb{C})$, mais pas un $\mathbb{C}$-sous-espace vectoriel puisque ce sous-ensemble n'est pas stable par multiplication par un élément de $\mathbb{C} \setminus \mathbb{R}$.

⚠ **Remarque 4.2.** Dans ces notes, nous suivons la convention qui semble être la plus répandue en définissant une forme sesquilinéaire comme étant linéaire à droite et anti-linéaire à gauche. Mais dans certains ouvrages, une forme sesquilinéaire est définie comme étant linéaire à gauche et anti-linéaire à droite.

Exemples 4.3.

(i) Si $E = \mathbb{C}^n$, alors

$$\varphi_1: E \times E \to \mathbb{C}, \ (x, y) \mapsto \sum_{j=1}^{n} \overline{x_j} y_j$$

est une forme sesquilinéaire à symétrie hermitienne sur E. Ici on note $(x_1, \ldots, x_n)$ et $(y_1, \ldots, y_n)$ les coordonnées des vecteurs x et y relativement à la base canonique de E.

(ii) Si $E = \mathbb{C}[X]$, alors

$$\varphi_2: E \times E \to \mathbb{C}, \ (P, Q) \mapsto \overline{P(0)}Q(1) + (i - 3)\,\overline{P(1)}Q(0)$$

est une forme sesquilinéaire sur E, mais pas à symétrie hermitienne.

(iii) Si $E = \mathcal{C}^1(\mathbb{R}, \mathbb{C})$, alors

$$\varphi_3: E \times E \to \mathbb{C}, \ (f, g) \mapsto \int_0^1 \overline{f(t)} g(t) dt$$

est une forme sesquilinéaire à symétrie hermitienne sur E.

Définition 4.4. Une *forme hermitienne* sur E est une application

$$\Phi: E \to \mathbb{R},\ x \mapsto \varphi(x,x),$$

où $\varphi: E \times E \to \mathbb{C}$ est une forme sesquilinéaire à symétrie hermitienne appelée la *forme polaire* de Φ.

La forme polaire φ d'une forme hermitienne Φ est uniquement déterminée par Φ. En effet, on a la *formule de polarisation* suivante (à vérifier en exercice) :

$$\forall x, y \in E,\ \varphi(x,y) = \frac{1}{4}[\Phi(x+y) - \Phi(x-y) + i\Phi(x-iy) - i\Phi(x+iy)].$$

Remarque 4.5. En pratique, une application $\Phi: E \to \mathbb{R}$ est une forme hermitienne sur E si
(i) $\forall x \in E$, $\forall \lambda \in \mathbb{C}$, $\Phi(\lambda x) = |\lambda|^2 \Phi(x)$; et
(ii) $\forall x, y \in E$, l'application $(x,y) \mapsto \frac{1}{4}[\Phi(x+y) - \Phi(x-y) + i\Phi(x-iy) - i\Phi(x+iy)]$ est sesquilinéaire à symétrie hermitienne.

Ces deux conditions ne sont pas équivalentes. Par exemple la fonction $f: \mathbb{C} \to \mathbb{R},\ t \mapsto |t|^2 + \alpha$, avec $\alpha \in \mathbb{R}$, vérifie (ii) mais pas (i), tandis que la fonction $g: \mathbb{C}^2 \to \mathbb{R},\ (x_1, x_2) \mapsto |x_1 x_2|$ vérifie (i) mais pas (ii).

Exemples 4.6.
(i) Si $E = \mathbb{C}^n$, alors

$$\Phi_1: E \to \mathbb{R},\ x \mapsto \sum_{j=1}^{n} |x_j|^2$$

est une forme hermitienne sur E.
(ii) Si $E = \mathcal{M}_n(\mathbb{C})$, alors

$$\begin{aligned} &\Phi_2: E \to \mathbb{R},\ A \mapsto |\mathrm{tr}(A)|^2 \ ; \\ &\Phi_3: E \to \mathbb{R},\ A \mapsto \mathrm{tr}(\overline{A}A) \ ; \text{et} \\ &\Phi_4: E \to \mathbb{R},\ A \mapsto \mathrm{tr}({}^t\overline{A}A) \end{aligned}$$

sont des formes hermitiennes sur E.
(iii) Si $E = \mathbb{C}[X]$, alors

$$\begin{aligned} &\Phi_5: E \to \mathbb{R},\ P \mapsto \overline{P(0)}P'(1) + \overline{P'(1)}P(0) \ ; \text{et} \\ &\Phi_6: E \to \mathbb{R},\ P \mapsto \int_0^1 |P(t)|^2 dt - \left|\left(\int_0^1 P(t)dt\right)\right|^2 \end{aligned}$$

sont des formes hermitiennes sur E.

⚠ **On suppose dorénavant que E est un $\mathbb{C}$-espace vectoriel de dimension finie $n \in \mathbb{N}$.**

4.2 Expression matricielle et matrices hermitiennes

On fixe une base $\mathcal{B} = \{e_1, \dots, e_n\}$ de E. Soient $x, y \in E$ et soit $\varphi: E \times E \to \mathbb{C}$ une forme sesquilinéaire sur E. On écrit $x = \sum_{j=1}^n x_j e_j$ et $y = \sum_{k=1}^n y_k e_k$. Alors

$$\varphi(x,y) = \varphi\left(\sum_{j=1}^n x_j e_j, \sum_{k=1}^n y_k e_k\right) = \sum_{j,k=1}^n \overline{x_j} y_k \varphi(e_j, e_k) = {}^t\overline{X}\,\mathrm{Mat}_{\mathcal{B}}(\varphi)\,Y$$

avec $X := \mathrm{Mat}_{\mathcal{B}}(x) = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$, $Y := \mathrm{Mat}_{\mathcal{B}}(y) = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$ et $\mathrm{Mat}_{\mathcal{B}}(\varphi) \in \mathcal{M}_n(\mathbb{C})$ la matrice carrée dont le coefficient (j,k) vaut $\varphi(e_j, e_k)$.

Définition 4.7. Une matrice $M \in \mathcal{M}_n(\mathbb{C})$ est dite *hermitienne* si elle vérifie ${}^t\overline{M} = M$. (En particulier, si $M \in \mathcal{M}_n(\mathbb{R})$, alors M est hermitienne si et seulement M est symétrique.) L'ensemble des matrices hermitiennes dans $\mathcal{M}_n(\mathbb{C})$ est un $\mathbb{R}$-sous-espace vectoriel que l'on note $\mathcal{H}_n(\mathbb{C})$.

On vérifie aisément que la forme sesquilinéaire φ est à symétrie hermitienne si et seulement si la matrice $\mathrm{Mat}_{\mathcal{B}}(\varphi)$ est hermitienne.

Exemples 4.8.

(i) Soit $E = \mathbb{C}^3$, muni de la base canonique, notée $\mathcal{B}$, et on considère la forme sesquilinéaire

$$\varphi: E \times E \to \mathbb{C},\ (x,y) \mapsto \overline{x_1} y_1 + (5+i)\,\overline{x_2} y_1 - 3i\,\overline{x_3} y_2.$$

Alors la matrice $\mathrm{Mat}_{\mathcal{B}}(\varphi) = \begin{bmatrix} 1 & 0 & 0 \\ 5+i & 0 & 0 \\ 0 & -3i & 0 \end{bmatrix}$ n'est pas hermitienne.

(ii) Soient $E = \mathcal{M}_2(\mathbb{C})$, muni de la base canonique $\mathcal{B} = \{E_{11}, E_{12}, E_{21}, E_{22}\}$, et

$$\varphi_3: E \times E \to \mathbb{C},\ (A,B) \mapsto \mathrm{tr}(\overline{A}B).$$

Alors la matrice $\mathrm{Mat}_{\mathcal{B}}(\varphi_3) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$ est réelle et symétrique, donc hermitienne.

Remarque 4.9. L'application $\varphi \in \mathcal{S}es(E, \mathbb{C}) \mapsto \mathrm{Mat}_{\mathcal{B}}(\varphi) \in \mathcal{M}_n(\mathbb{C})$ est un isomorphisme de $\mathbb{C}$-espaces vectoriels. Celui-ci se restreint en un isomorphisme de $\mathbb{R}$-espaces vectoriels entre les formes sesquilinéaires à symétrie hermitienne sur E est les matrices hermitiennes dans $\mathcal{M}_n(\mathbb{C})$.

De plus, en notant $M = M_1 + iM_2$ avec $M_1, M_2 \in \mathcal{M}_n(\mathbb{R})$, on vérifie que

$$\begin{aligned} M \in \mathcal{H}_n(\mathbb{C}) &\Leftrightarrow {}^t(M_1 + iM_2) = \overline{M_1 + iM_2} \\ &\Leftrightarrow {}^tM_1 + i\,{}^tM_2 = M_1 - iM_2 \\ &\Leftrightarrow {}^tM_1 = M_1 \text{ et } {}^tM_2 = -M_2 \\ &\Leftrightarrow M_1 \in \mathcal{S}_n(\mathbb{R}) \text{ et } M_2 \in \mathcal{A}_n(\mathbb{R}), \end{aligned}$$

d'où

$$\dim_{\mathbb{R}}(\mathcal{H}_n(\mathbb{C})) = \dim_{\mathbb{R}}(\mathcal{S}_n(\mathbb{R})) + \dim_{\mathbb{R}}(\mathcal{A}_n(\mathbb{R})) = \frac{1}{2}n(n+1) + \frac{1}{2}n(n-1) = n^2,$$

et donc le $\mathbb{R}$-sous-espace vectoriel de $\mathcal{S}es(E,\mathbb{C})$ constitué des formes sesquilinéaires à symétrie hermitienne est de dimension réelle n^2. (Et le $\mathbb{C}$-espace vectoriel $\mathcal{S}es(E,\mathbb{C}) \simeq \mathcal{M}_n(\mathbb{C})$ est de dimension complexe n^2, donc de dimension réelle $2n^2$.)

Lemme 4.10 (Formule de changement de base)**.** Soient $\mathcal{B}$ et $\mathcal{B}'$ deux bases de E, et soit $P = \mathrm{Pass}(\mathcal{B} \to \mathcal{B}')$. Si φ est une forme sesquilinéaire sur E, alors

$$\mathrm{Mat}_{\mathcal{B}'}(\varphi) = {}^t\overline{P}\,\mathrm{Mat}_{\mathcal{B}}(\varphi)\,P.$$

Démonstration. La démonstration de ce lemme est analogue à celle du lemme 3.11. □

Exemple 4.11. Soient $E = \mathbb{C}^3$, $\mathcal{B}$ la base canonique de E, et $\mathcal{B}' = \left\{\begin{bmatrix}1\\1\\2+i\end{bmatrix}, \begin{bmatrix}1\\0\\1-i\end{bmatrix}, \begin{bmatrix}0\\-1\\0\end{bmatrix}\right\}$ une autre base de E. On considère la forme sesquilinéaire à symétrie hermitienne

$$\varphi: E \times E \to \mathbb{C},\ (x,y) \mapsto \overline{x_1}y_1 - 2\overline{x_2}y_2 + (7+i)\,\overline{x_1}y_2 + (7-i)\,\overline{x_2}y_1 + i\sqrt{5}\,\overline{x_3}y_2 - i\sqrt{5}\,\overline{x_2}y_3.$$

Alors $\mathrm{Mat}_{\mathcal{B}}(\varphi) = \begin{bmatrix}1 & 7+i & 0\\ 7-i & -2 & -i\sqrt{5}\\ 0 & i\sqrt{5} & 0\end{bmatrix}$, et la formule de changement de base donne

$$\begin{aligned}\mathrm{Mat}_{\mathcal{B}'}(\varphi) &= \begin{bmatrix}1 & 1 & 2-i\\ 1 & 0 & 1+i\\ 0 & -1 & 0\end{bmatrix}\begin{bmatrix}1 & 7+i & 0\\ 7-i & -2 & -i\sqrt{5}\\ 0 & i\sqrt{5} & 0\end{bmatrix}\begin{bmatrix}1 & 1 & 0\\ 1 & 0 & -1\\ 2+i & 1-i & 0\end{bmatrix}\\ &= \begin{bmatrix}2\sqrt{5}+13 & 8-\sqrt{5}-i(1+\sqrt{5}) & -5-\sqrt{5}-i(1+2\sqrt{5})\\ 8-\sqrt{5}+i(1+\sqrt{5}) & 1 & -7+\sqrt{5}-i(1+\sqrt{5})\\ -5-\sqrt{5}+i(1+2\sqrt{5}) & -7+\sqrt{5}+i(1+\sqrt{5}) & -2\end{bmatrix}.\end{aligned}$$

4.3 Analogie entre la théorie des formes quadratiques réelles et la théorie des formes hermitiennes

Pour la plupart des énoncés portant sur les formes quadratiques réelles, il existe un énoncé analogue portant sur les formes hermitiennes. Dans cette section on évoque quelques uns de ces énoncés sans entre dans les détails. (On renvoie le lecteur ou la lectrice de ces notes au chapitre 5 du livre "Algèbre" de Gourdon pour un traitement commun de ces deux théories.)

- On peut définir la notion de noyau et de rang d'une forme hermitienne $\Phi: E \to \mathbb{R}$, de forme polaire φ, de la même façon que pour les formes quadratiques, et ces notions coïncident bien sûr avec le noyau et le rang de la matrice $\mathrm{Mat}_{\mathcal{B}}(\varphi)$, avec $\mathcal{B}$ une base quelconque de E. En particulier, il existe un théorème du rang pour les formes hermitiennes.
- On peut parler d'orthogonalité et d'isotropie pour les formes hermitiennes, et toute forme hermitienne $\Phi: E \to \mathbb{C}$, de forme polaire φ, admet une base φ-orthogonale (la démonstration de ce résultat est analogue à la démonstration du théorème 3.23). En particulier, cet énoncé se traduit matriciellement par le résultat suivant :

$$\forall M \in \mathcal{H}_n(\mathbb{C}), \exists P \in \mathrm{GL}_n(\mathbb{C}),\ {}^t\overline{P}MP \text{ est une matrice diagonale réelle.}$$

- Il existe une méthode de Gauss pour réduire n'importe quelle forme hermitienne en une combinaison linéaire de carrés de modules de formes linéaires sur E linéairement indépendantes, et le théorème 3.27 admet un analogue pour les formes hermitiennes. Par exemple :

$$\Phi_1(x,y) := \overline{x}y + x\overline{y} = \frac{1}{2}|x+y|^2 - \frac{1}{2}|x-y|^2 \text{ ; et}$$
$$\Phi_2(x,y,z) := \overline{x}x + \overline{y}y - 2i\overline{x}y + 2i\overline{y}x + 2\overline{z}y + 2\overline{y}z = |x-2iy|^2 - 3|y - \frac{2}{3}z|^2 + \frac{4}{3}|z|^2.$$

- Il existe une loi d'inertie de Sylvester pour les formes hermitiennes qui permet d'associer à toute forme hermitienne un couple d'entiers, qu'on appelle la signature de la forme hermitienne, et deux formes hermitiennes sur E sont équivalentes si et seulement si elles ont la même signature. Par exemple la signature de Φ_1 est $(1,1)$, et la signature de Φ_2 est $(2,1)$.

4.4 Produit scalaire hermitien

Soit Φ une forme hermitienne sur un $\mathbb{C}$-espace vectoriel E de dimension n, et soit φ la forme polaire de Φ.

Définition 4.12. On dit que φ est un *produit scalaire (hermitien)* sur E si Φ est définie positive, c'est-à-dire
$$\forall x \in E,\ \Phi(x) \geq 0, \quad \text{et} \quad \Phi(x) = 0 \Rightarrow x = 0.$$
On dit alors que (E,φ) est un *espace hermitien.*

Exemples 4.13.

(i) Si $E = \mathbb{C}^n$ et $\gamma_1, \ldots, \gamma_n \in \mathbb{R}_+^*$ (par exemple $\gamma_1 = \cdots = \gamma_n = 1$), alors
$$\varphi: E \times E \to \mathbb{C},\ (x,y) \mapsto \sum_{j=1}^{n} \gamma_j \overline{x_j} y_j$$
est un produit scalaire (hermitien) sur E.

(ii) Soit $E = \mathbb{C}[X]_{\leq n-1}$. On se donne n points distincts $z_1, \ldots, z_n \in \mathbb{C}$. Alors
$$\varphi_1: E \times E \to \mathbb{C},\ (P,Q) \mapsto \sum_{j=1}^{n} \overline{P(z_j)} Q(z_j) \ \text{ et } \varphi_2: E \times E \to \mathbb{C},\ \left(\sum_{j=0}^{n-1} a_j X^j, \sum_{j=0}^{n-1} b_j X^j\right) \mapsto \sum_{j=0}^{n-1} \overline{a_j} b_j$$
sont des produits scalaires (hermitiens) sur E.

(iii) Si $E = \mathcal{M}_n(\mathbb{C})$, alors
$$\varphi: E \times E \to \mathbb{C},\ (A,B) \mapsto \mathrm{tr}({}^t\overline{A}B)$$
est un produit scalaire sur E dont la norme associée est appelée *norme de Frobenius.*

Remarques 4.14.

- Comme dans le cadre euclidien, si (E,φ) est un espace hermitien, alors on peut orthonormaliser n'importe quelle base de E à l'aide d'un analogue de l'algorithme Gram-schmidt.
- Un espace hermitien de dimension complexe n est aussi un espace euclidien de dimension réelle $2n$. En effet, si φ est un produit scalaire (hermitien) sur E, alors $(x,y) \mapsto \mathcal{R}e(\varphi(x,y))$ est un produit scalaire (euclidien) sur E.

Proposition 4.15 (Inégalité de Cauchy-Schwarz)**.** Si $\Phi: E \to \mathbb{R}$ est une forme hermitienne dont la forme polaire $\varphi: E \times E \to \mathbb{C}$ est un produit scalaire (hermitien), alors

$$\forall x, y \in E,\ |\varphi(x,y)|^2 \leq \Phi(x)\Phi(y).$$

De plus, on a égalité si et seulement si x et y sont colinéaires.

Démonstration. Si $x = 0$ ou $y = 0$, alors l'inégalité est évidente puisque les deux membres valent alors 0. On peut donc supposer que x et y sont non-nuls.

Si $\varphi(x,y) = \rho e^{i\theta}$ avec $\rho \in \mathbb{R}_+$ et $\theta \in \mathbb{R}$, on peut toujours remplacer x par $xe^{-i\theta}$ pour se ramener au cas où $\varphi(x,y) \in \mathbb{R}$. On a alors

$$\forall \lambda \in \mathbb{R},\ \Phi(\lambda x + y) = \lambda^2 \Phi(x) + 2\lambda\varphi(x,y) + \Phi(y) \geq 0 \ \ \text{(car la forme hermitienne } \Phi \text{ est positive)}.$$

On a donc un trinôme du second degré en λ avec un discriminant négatif, c'est-à-dire

$$\Delta := 4|\varphi(x,y)|^2 - 4\Phi(x)\Phi(y) \leq 0,\ \text{d'où l'inégalité annoncée.}$$

Ensuite, on a égalité si et seulement $\Delta = 0$, ce qui signifie qu'il existe $\lambda_0 \in \mathbb{R}$ tel que $\Phi(\lambda_0 x + y) = 0$. Comme la forme hermitienne Φ est définie, on en déduit $\lambda_0 x + y = 0$, c'est-à-dire que les vecteurs x et y sont colinéaires. □

Corollaire 4.16 (Inégalité de Minkowski)**.** Si la forme hermitienne $\Phi: E \to \mathbb{R}$ est définie positive, alors

$$\forall x, y \in E,\ \sqrt{\Phi(x+y)} \leq \sqrt{\Phi(x)} + \sqrt{\Phi(y)}.$$

De plus, il y a égalité si et seulement x et y sont positivement liés.

Démonstration. On a

$$\begin{aligned}
\Phi(x+y) &= \Phi(x) + 2\mathcal{R}e(\varphi(x,y)) + \Phi(y) \\
&\leq \Phi(x) + 2|\varphi(x,y)| + \Phi(y) \\
&\leq \Phi(x) + 2\sqrt{\Phi(x)\Phi(y)} + \Phi(y) \ \ \text{(d'après l'inégalité de Cauchy-Schwarz)} \\
&= \left(\sqrt{\Phi(x)} + \sqrt{\Phi(y)}\right)^2.
\end{aligned}$$

Et le cas d'égalité découle du cas d'égalité dans l'inégalité de Cauchy-Schwarz. □

Soit Φ une forme hermitienne et définie positive sur E. On déduit de l'inégalité de Minkowski que l'application

$$E \to \mathbb{R}_+,\ x \mapsto \sqrt{\Phi(x)}$$

est une norme qui munit E d'une structure d'espace de Banach.

4.5 Endomorphismes unitaires et matrices unitaires

Dorénavant, on note $\langle -, - \rangle$ un produit scalaire (hermitien) sur le $\mathbb{C}$-espace vectoriel E de dimension n, et on note $\| - \|$ la norme correspondante, c'est-à-dire $\|x\| = \sqrt{\langle x, x\rangle}$ pour tout $x \in E$.

Définition 4.17. Un endomorphisme $u \in \mathcal{L}(E)$ tel que, pour tout $x \in E$, on a $\|u(x)\| = \|x\|$ est appelé un *endomorphisme unitaire* de E. L'ensemble des endomorphismes unitaires de E est un sous-groupe de $\mathcal{GL}(E)$, noté $\mathcal{U}(E)$, que l'on appelle *groupe unitaire*.

Remarques 4.18.

- Si $u \in \mathcal{L}(E)$ est un endomorphisme unitaire, alors $u \in \mathcal{GL}(E)$ puisque

$$\forall x \in \mathrm{Ker}(u),\ 0 = \|u(x)\| = \|x\|,\ \text{ et donc } x = 0.$$

- Il découle de la formule de polarisation que $u \in \mathcal{L}(E)$ est unitaire si et seulement si

$$\forall x, y \in E,\ \langle u(x), u(y) \rangle = \langle x, y \rangle\,.$$

- La caractérisation ci-dessus implique que $u \in \mathcal{L}(E)$ est unitaire si et seulement si l'image d'une base orthonormée de E est envoyé sur une base orthonormée de E.

Définitions 4.19. L'ensemble

$$\mathrm{U}_n(\mathbb{C}) := \{M \in \mathcal{M}_n(\mathbb{C}) \mid {}^t\overline{M}M = I_n\}$$

est un sous-groupe de $\mathrm{GL}_n(\mathbb{C})$ appelé *groupe unitaire*. De plus, l'ensemble

$$\mathrm{SU}_n(\mathbb{C}) := \{M \in \mathcal{M}_n(\mathbb{C}) \mid {}^t\overline{M}M = I_n \text{ et } \det(M) = 1\} = \mathrm{U}_n(\mathbb{C}) \cap \mathrm{SL}_n(\mathbb{C})$$

est un sous-groupe distingué de $\mathrm{U}_n(\mathbb{C})$ appelé *groupe spécial unitaire*.

Remarque 4.20. Les groupes matriciels $\mathrm{U}_n(\mathbb{C})$ et $\mathrm{SU}_n(\mathbb{C})$ sont les analogues "hermitiens" des groupes matriciels $\mathrm{O}_n(\mathbb{R})$ et $\mathrm{SO}_n(\mathbb{R})$. D'ailleurs, on vérifie aisément que

$$\mathrm{U}_n(\mathbb{C}) \cap \mathrm{GL}_n(\mathbb{R}) = \mathrm{O}_n(\mathbb{R}) \quad \text{et} \quad \mathrm{SU}_n(\mathbb{C}) \cap \mathrm{GL}_n(\mathbb{R}) = \mathrm{SO}_n(\mathbb{R}).$$

Soit $M \in \mathrm{U}_n(\mathbb{C})$. Alors, il découle de la définition de $\mathrm{U}_n(\mathbb{C})$ que $M^{-1} = {}^t\overline{M}$, ce qui rend le calcul de l'inverse aisé pour les matrices unitaires. Par ailleurs, si l'on calcule le déterminant des deux côtés de l'égalité ${}^t\overline{M}M = I_n$, on obtient

$$|\det(M)|^2 = \overline{\det(M)}\det(M) = \det(\overline{M})\det(M) = \det({}^t\overline{M})\det(M) = \det({}^t\overline{M}M) = \det(I_n) = 1,$$

et donc $|\det(M)| = 1$. Enfin, on peut identifier $\mathrm{SU}_n(\mathbb{C})$ avec le noyau de l'homomorphisme

$$\det\colon \mathrm{U}_n(\mathbb{C}) \to \mathbb{S}^1 := \{z \in \mathbb{C},\ |z| = 1\},\ M \mapsto \det(M).$$

Le résultat qui suit nous donne une quatrième caractérisation des endomorphismes unitaires et fait le lien entre endomorphismes unitaires et matrices unitaires.

Proposition 4.21. Soit $\mathcal{B}$ une base orthonormée de l'espace hermitien $(E, \langle -, - \rangle)$, et soit $u \in \mathcal{L}(E)$. Alors $u \in \mathcal{U}(E)$ si et seulement si $\mathrm{Mat}_{\mathcal{B}}(u) \in \mathrm{U}_n(\mathbb{C})$. En particulier, si $u \in \mathcal{U}(E)$, alors $|\det(u)| = 1$.

Démonstration. On note $M := \mathrm{Mat}_{\mathcal{B}}(u)$. On a vu que l'endomorphisme u est unitaire si et seulement si

$$\forall x, y \in E,\ \langle u(x), u(y)\rangle = \langle x, y\rangle\,,$$

ce qui donne matriciellement (puisque la base $\mathcal{B}$ est orthonormée) :

$$\forall X, Y \in \mathcal{M}_{n,1}(\mathbb{C}),\ {}^t(\overline{MX})MY = {}^t\overline{X}{}^t\overline{M}MY = {}^t\overline{X}Y.$$

Et cette dernière égalité est équivalente à ${}^t\overline{M}M = I_n$. (En effet, si $A = (a_{j,k})_{1\leq j,k\leq n} \in \mathcal{M}_n(\mathbb{C})$, alors ${}^t\overline{e_j}Ae_k = {}^te_jAe_k = a_{j,k}$, où l'on note $\{e_1, \ldots, e_n\}$ la base canonique de $\mathbb{C}^n$.)

Enfin, la dernière assertion résulte des égalités $\det(u) = \det(M)$ et $|\det(M)| = 1$. $\square$

Dans le cadre euclidien, on ne peut pas toujours diagonaliser les isométries, mais seulement les diagonaliser par blocs en général. En revanche, on peut toujours diagonaliser les endomorphismes unitaires dans le cadre hermitien.

Théorème 4.22 (Réduction des endomorphismes unitaires)**.**
Soit $(E, \langle -,-\rangle)$ un espace hermitien et $u \in \mathcal{U}(E)$. Alors il existe une base orthonormée de E qui diagonalise u, et toutes les valeurs propres de u sont de module 1.

Démonstration. Déjà, si $\lambda \in \mathbb{C}$ est une valeur propre de $u \in \mathcal{U}(E)$, alors pour tout vecteur propre $x \in E_\lambda \setminus \{0\}$, on a

$$\|x\| = \|u(x)\| = \|\lambda x\| = |\lambda|\,\|x\|,\ \text{et donc } |\lambda| = 1.$$

Pour montrer l'existence d'une base orthonormée de $(E, \langle -,-\rangle)$ qui diagonalise u, on procède par récurrence sur la dimension de E. Si $\dim(E) = 1$, alors il n'y a rien à faire. On suppose maintenant que le résultat est vrai pour tout endomorphisme unitaire $u' \in \mathcal{U}(E')$ avec $\dim(E') \leq n-1$, et on va démontrer le résultat pour $u \in \mathcal{U}(E)$.

Étant donné que $\mathbb{C}$ est algébriquement clos, u admet au moins une valeur propre $\lambda_0 \in \mathbb{C}$. Soit $x_0 \in E_{\lambda_0}$ tel que $\|x_0\| = 1$. Alors $F = \mathrm{Vect}(x_0)$ est une droite de E stable par u, et on peut vérifier que $F^\perp = \{y \in E \mid \langle x_0, y\rangle = 0\}$ est un supplémentaire de F dans E stable par u. En effet, $F^\perp$est un hyperplan de E, puisque c'est le noyau de la forme linéaire (non-nulle) $y \mapsto \langle x_0, y\rangle$, et $x_0 \notin F^\perp$ puisque $\langle x_0, x_0\rangle = \|x_0\|^2 = 1$. Donc $E = F \oplus F^\perp$. En outre, si $y \in F^\perp$, alors $u(y) \in F^\perp$ car

$$\langle x_0, u(y)\rangle = \left\langle \frac{1}{\lambda_0}\lambda_0 x_0, u(y)\right\rangle = \frac{1}{\overline{\lambda_0}}\langle \lambda_0 x_0, u(y)\rangle = \frac{1}{\overline{\lambda_0}}\langle u(x_0), u(y)\rangle = \frac{1}{\overline{\lambda_0}}\langle x_0, y\rangle = 0.$$

L'endomorphisme $u_{|F^\perp}$ est unitaire, et donc d'après l'hypothèse de récurrence, il existe une base orthonormée $\mathcal{B}_0$ de $F^\perp$ qui diagonalise $u_{|F^\perp}$. En posant $\mathcal{B} = \{x_0\} \cup \mathcal{B}_0$, on obtient une base de E qui diagonalise u, ce qui achève la démonstration du théorème. $\square$

Corollaire 4.23 (Version matricielle)**.** Soit $M \in \mathrm{U}_n(\mathbb{C})$. Il existe $P \in \mathrm{U}_n(\mathbb{C})$ telle que

$$P^{-1}MP = {}^t\overline{P}MP = D,$$

où D est une matrice diagonale dont les coefficients sont des nombres complexes de module 1.

Démonstration. On note $\mathcal{B}_0$ la base canonique de $\mathbb{C}^n$, qui est orthonormée pour le produit scalaire hermitien usuel $\langle x, y\rangle = \sum_{j=1}^n \overline{x_j}y_j$. Soit $u \in \mathcal{U}(\mathbb{C}^n)$ l'endomorphisme unitaire défini pas l'égalité $M = \mathrm{Mat}_{\mathcal{B}_0}(u)$. D'après le théorème 4.22, il existe une base orthonormée $\mathcal{B}$ de $(\mathbb{C}^n, \langle -,-\rangle)$ telle que

$\mathrm{Mat}_{\mathcal{B}}(u)$ est diagonale. La matrice de passage $P = \mathrm{Pass}(\mathcal{B}_0 \to \mathcal{B})$ est unitaire, puisque $\mathcal{B}_0$ et $\mathcal{B}$ sont deux bases orthonormées de $(\mathbb{C}^n, \langle -, - \rangle)$, et on a

$$ {}^t\overline{P}MP = P^{-1}MP = \mathrm{Pass}(\mathcal{B}_0 \to \mathcal{B})^{-1}\mathrm{Mat}_{\mathcal{B}_0}(u)\mathrm{Pass}(\mathcal{B}_0 \to \mathcal{B}) = \mathrm{Mat}_{\mathcal{B}}(u) =: D $$

qui est une matrice diagonale. Enfin, toujours d'après le théorème 4.22, les valeurs propres de u, qui sont précisément les coefficients de la matrice diagonale D, sont des nombres complexes de module 1, ce qui termine la démonstration du corollaire. □

Dans le cadre hermitien, il existe une large classe d'endomorphismes "naturels" (les *endomorphismes normaux*) dont on peut montrer qu'ils sont diagonalisables dans une base orthonormée. En particulier, les endomorphismes unitaires et auto-adjoints sont normaux.

L'étudiant.e qui souhaite en savoir plus sur la réduction des endomorphismes normaux peut consulter par exemple le chapitre 5.3 du livre "Algèbre" de Gourdon.

4.6 Décomposition polaire

Le but de cette dernière section est de démontrer la décomposition polaire pour les matrices complexes inversibles. Il s'agit d'un outil très utile dans l'étude des propriétés topologiques des groupes matriciels.

On rappelle que sur un $\mathbb{K}$-espace vectoriel de dimension finie (avec $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$), toutes les normes sont équivalentes et définissent donc la même topologie. Ainsi, sur l'espace des matrices $\mathcal{M}_n(\mathbb{C})$, on peut par exemple choisir de travailler avec la norme

$$ \| - \| \colon M \mapsto \sqrt{\mathrm{tr}({}^t\overline{M}M)}. $$

Et si X est n'importe quelle partie de $\mathcal{M}_n(\mathbb{C})$, par exemple $\mathrm{GL}_n(\mathbb{C})$, $\mathcal{H}_n(\mathbb{C})$ ou $\mathrm{U}_n(\mathbb{C})$, alors on peut munir X de la topologie induite par celle de $\mathcal{M}_n(\mathbb{C})$.

Le résultat qui suit nous sera utile dans la démonstration de la décomposition polaire.

Lemme 4.24. Le groupe unitaire $\mathrm{U}_n(\mathbb{C})$ est une partie compacte de $\mathcal{M}_n(\mathbb{C})$.

Démonstration. Étant donné que $\mathcal{M}_n(\mathbb{C})$ est un espace vectoriel normé, donc un espace métrique, il s'agit de montrer que $\mathrm{U}_n(\mathbb{C})$ est une partie fermée et bornée de $\mathcal{M}_n(\mathbb{C})$.

On considère l'application

$$ f \colon \mathcal{M}_n(\mathbb{C}) \to \mathcal{M}_n(\mathbb{C}),\ M \mapsto {}^t\overline{M}M. $$

Cette application est continue ; en effet, les applications

$$ f_1 \colon M_1 \mapsto (\overline{M_1}, M_1) \quad \text{et} \quad f_2 \colon (M_2, M_3) \mapsto {}^tM_2M_3 $$

sont continues, et $f = f_2 \circ f_1$. En outre, $\mathrm{U}_n(\mathbb{C}) = f^{-1}(\{I_n\})$ est l'image inverse du singleton $\{I_n\}$ qui est une partie fermée de $\mathcal{M}_n(\mathbb{C})$, et donc $\mathrm{U}_n(\mathbb{C})$ est aussi une partie fermée de $\mathcal{M}_n(\mathbb{C})$.

Ensuite, pour tout $M \in \mathrm{U}_n(\mathbb{C})$, on a $\|M\|^2 = \mathrm{tr}({}^t\overline{M}M) = \mathrm{tr}(I_n) = n$, et donc $\|M\| = \sqrt{n}$. En particulier, $\mathrm{U}_n(\mathbb{C})$ est contenue dans la boule fermée de centre $\{0\}$ et de rayon $\sqrt{n}$, c'est donc une partie bornée de $\mathcal{M}_n(\mathbb{C})$.

En conclusion, $\mathrm{U}_n(\mathbb{C})$ est une partie fermée et bornée de $\mathcal{M}_n(\mathbb{C})$, donc une partie compacte de $\mathcal{M}_n(\mathbb{C})$. □

Définitions 4.25. On note

$$\mathcal{H}_n^+(\mathbb{C}) = \{M \in \mathcal{H}_n(\mathbb{C}) \mid M \text{ est positive}\}; \text{ et}$$
$$\mathcal{H}_n^{++}(\mathbb{C}) = \{M \in \mathcal{H}_n(\mathbb{C}) \mid M \text{ est définie positive}\}.$$

(Une matrice $M \in \mathcal{H}_n(\mathbb{C})$ est dite *positive* si, pour tout $X \in \mathcal{M}_{n,1}(\mathbb{C})$, on a ${}^t\overline{X}MX \geq 0$, et elle est dite *définie positive* si, pour tout $X \in \mathcal{M}_{n,1}(\mathbb{C}) \smallsetminus \{0\}$, on a ${}^t\overline{X}MX > 0$.)

On rappelle (cf. remarque 4.9) que $\mathcal{H}_n(\mathbb{C})$ est un $\mathbb{R}$-sous-espace vectoriel de dimension n^2 dans $\mathcal{M}_n(\mathbb{C})$. On a bien sûr les inclusions

$$\mathcal{H}_n^{++}(\mathbb{C}) \subset \mathcal{H}_n^+(\mathbb{C}) \subset \mathcal{H}_n(\mathbb{C}),$$

et on peut vérifier que $\mathcal{H}_n^+(\mathbb{C})$ est un cône fermé dans $\mathcal{H}_n(\mathbb{C})$, et que $\mathcal{H}_n^{++}(\mathbb{C})$ est un cône ouvert dans $\mathcal{H}_n(\mathbb{C})$ qui coïncide avec l'intérieur de $\mathcal{H}_n^+(\mathbb{C})$.

Exemples 4.26.

(i) Si $n = 1$, alors $\mathcal{H}_1(\mathbb{C}) = \mathcal{M}_1(\mathbb{R}) \simeq \mathbb{R}$, $\mathcal{H}_1^+(\mathbb{C}) \simeq \mathbb{R}_+$, et $\mathcal{H}_1^{++}(\mathbb{C}) \simeq \mathbb{R}_+^*$.
(ii) Si $n = 2$, alors on vérifie que

$$\mathcal{H}_2(\mathbb{C}) = \left\{ \begin{bmatrix} a & b \\ \overline{b} & c \end{bmatrix};\ a, c \in \mathbb{R} \right\} \simeq \mathbb{R}^4;$$
$$\mathcal{H}_2^+(\mathbb{C}) = \left\{ \begin{bmatrix} a & b \\ \overline{b} & c \end{bmatrix};\ a, c \in \mathbb{R},\ a + c \geq 0, \text{ et } ac \geq |b|^2 \right\}; \text{ et}$$
$$\mathcal{H}_2^{++}(\mathbb{C}) = \left\{ \begin{bmatrix} a & b \\ \overline{b} & c \end{bmatrix};\ a, c \in \mathbb{R},\ a + c > 0, \text{ et } ac > |b|^2 \right\}.$$

On a mentionné dans la section 4.3 que tout matrice hermitienne est diagonalisable et à valeurs propres réelles. En fait, on peut même montrer un *théorème spectral* pour les matrices hermitiennes (dont la démonstration est analogue à celle du théorème spectral pour les matrices symétriques réelles) :

$$\forall M \in \mathcal{H}_n(\mathbb{C}),\ \exists P \in U_n(\mathbb{C}),\ {}^t\overline{P}MP = P^{-1}MP = D$$

avec D une matrice diagonale réelle. Il s'ensuit que :

- si $M \in \mathcal{H}_n^+(\mathbb{C})$, alors M est diagonalisable et ses valeurs propres sont dans $\mathbb{R}_+$; et
- si $M \in \mathcal{H}_n^{++}(\mathbb{C})$, alors M est diagonalisable et ses valeurs propres sont dans $\mathbb{R}_+^*$.

Enfin, on rappelle qu'une application $f\colon X \to Y$ entre deux espaces topologiques est appelé un *homéomorphisme* si f est continue, bijective, et d'inverse continu. On est maintenant prêts à énoncer puis à démontrer le principal résultat de cette section :

Théorème 4.27 (Décomposition polaire)**.**
L'application

$$\pi\colon \mathrm{U}_n(\mathbb{C}) \times \mathcal{H}_n^{++}(\mathbb{C}) \to \mathrm{GL}_n(\mathbb{C}),\ (O, S) \mapsto OS$$

est un homéomorphisme (ici on considère $\mathrm{U}_n(\mathbb{C}) \times \mathcal{H}_n^{++}(\mathbb{C})$ muni de la topologie produit). En particulier, toute matrice complexe inversible se décompose de façon unique en produit d'une matrice unitaire et d'une matrice hermitienne définie positive.

Démonstration. Pour alléger les notations, dans toute la suite on notera $M^* := {}^t\overline{M}$ lorsque $M \in \mathcal{M}_n(\mathbb{C})$.

• **Surjectivité de** π **:** Soit $M \in \mathrm{GL}_n(\mathbb{C})$. On cherche $(O,S) \in \mathrm{U}_n(\mathbb{C}) \times \mathcal{H}_n^{++}(\mathbb{C})$ tel que $M = OS$. Alors $M^* = S^*O^*$, et donc on doit avoir

$$M^*M = S^*O^*OS = S^*S = S^2, \text{ puisque } O^* = O^{-1} \text{ et } S = S^*.$$

De plus, on vérifie que $M^*M \in \mathcal{H}_n^{++}(\mathbb{C})$. Donc, d'après le théorème spectral pour les matrices hermitiennes, il existe $P \in \mathrm{U}_n(\mathbb{C})$ et une matrice diagonale $D = \mathrm{diag}(\lambda_1, \dots, \lambda_n)$, avec les $\lambda_i \in \mathbb{R}_+^*$, telles que $M^*M = PDP^{-1} = PDP^*$. On pose alors $S := PD'P^{-1} = PD'P^*$, avec $D' = \mathrm{diag}(\sqrt{\lambda_1}, \dots, \sqrt{\lambda_n})$ (de sorte à avoir $S^2 = M^*M$) ; c'est un élément de $\mathcal{H}_n^{++}(\mathbb{C})$ puisque $S^* = (PD'P^*)^* = PD'^*P^* = PD'P^* = S$ et les valeurs propres de S sont dans $\mathbb{R}_+^*$.

Ensuite, on pose $O := MS^{-1}$. Alors

$$O^*O = (MS^{-1})^*MS^{-1} = (S^{-1})^*M^*MS^{-1} = (S^*)^{-1}M^*MS^{-1} = S^{-1}S^2S^{-1} = I_n.$$

Autrement dit, $O \in U_n(\mathbb{C})$ et on a bien $M = OS$ avec $S \in \mathcal{H}_n^{++}(\mathbb{C})$. On a donc prouvé que l'application π est surjective, c'est-à-dire que toute toute matrice complexe inversible se décompose en produit d'une matrice unitaire et d'une matrice hermitienne définie positive. On va maintenant montrer que cette décomposition est unique, c'est-à-dire, que π est injective.

• **Injectivité de** π **:** On fixe $M \in \mathrm{GL}_n(\mathbb{C})$. Si $M = OS$ avec $(O,S) \in \mathrm{U}_n(\mathbb{C}) \times \mathcal{H}_n^{++}(\mathbb{C})$, alors $O = MS^{-1}$ et donc O est déterminée de manière unique par S. Autrement dit, pour montrer l'injectivité de π il suffit de montrer l'unicité de S. On a vu précédemment que l'égalité $M = OS$ implique $M^*M = S^2$, et donc M détermine S^2. Il suffit donc de montrer que si $T \in \mathcal{H}_n^{++}(\mathbb{C})$, alors il existe une unique matrice $R \in \mathcal{H}_n^{++}(\mathbb{C})$ telle que $R^2 = T$. Ceci impliquera que S^2 détermine S de manière unique, et donc que M détermine S de manière unique (et O également comme $O = MS^{-1}$).

On fixe $T \in \mathcal{H}_n^{++}(\mathbb{C})$, et soient $R_1, R_2 \in \mathcal{H}_n^{++}(\mathbb{C})$ telles que $R_1^2 = R_2^2 = T$. Alors R_1 commute avec $R_1^2 = R_2^2$. Or, il existe $f \in \mathbb{C}[X]$ tel que $R_2 = f(R_2^2)$. En effet, d'après le théorème spectral pour les matrices hermitiennes, il existe $P \in \mathrm{U}_n(\mathbb{C})$ et une matrice diagonale $D = \mathrm{diag}(\lambda_1, \dots, \lambda_n)$, avec les $\lambda_i \in \mathbb{R}_+^*$, telles que $R_2 = PDP^{-1} = PDP^*$. Il suffit donc de choisir un polynôme $f \in \mathbb{C}[X]$ tel que $f(\lambda_i^2) = \lambda_i$ pour $i = 1, \dots, n$ (un tel polynôme existe toujours par interpolation lagrangienne), et on obtient

$$f(R_2^2) = f(PD^2P^{-1}) = Pf(D^2)P^{-1} = P\mathrm{diag}(f(\lambda_1^2), \dots, f(\lambda_n^2))P^{-1} = P\mathrm{diag}(\lambda_1, \dots, \lambda_n)P^{-1} = PDP^{-1} = R_2.$$

Donc, R_1 commute avec $f(R_2^2) = R_2$. On peut donc diagonaliser simultanément R_1 et R_2 : Il existe $Q \in \mathrm{GL}_n(\mathbb{C})$ et deux matrices diagonales à coefficients dans $\mathbb{R}_+^*$, notées D_1 et D_2, telles que $R_1 = QD_1Q^{-1}$ et $R_2 = QD_2Q^{-1}$. Puis

$$R_1^2 = R_2^2 \Leftrightarrow D_1^2 = D_2^2 \Leftrightarrow D_1 = D_2 \text{ (car } x \mapsto \sqrt{x} \text{ est une bijection de } \mathbb{R}_+^* \text{ dans } \mathbb{R}_+^*) \Leftrightarrow R_1 = R_2.$$

Et donc, il existe une unique matrice $R \in \mathcal{H}_n^{++}(\mathbb{C})$ telle que $R^2 = T$.

À ce stade, on a montré que l'application π est bijective, et elle est continue puisque les coefficients de la matrice OS sont des polynômes en les coefficients des matrices O et S. Pour terminer la démonstration, il reste donc à montrer que l'application réciproque de π, que l'on va noter κ, est également continue.

• **Continuité de** $\kappa = \pi^{-1}$ **:** Étant donné qu'il semble ardu d'expliciter l'application κ, on va montrer sa continuité via la caractérisation séquentielle de la continuité. On considère une suite

$(M_p)_{p\in\mathbb{N}}$ d'éléments dans $\mathrm{GL}_n(\mathbb{C})$ qui converge vers une matrice $M \in \mathrm{GL}_n(\mathbb{C})$. Pour chaque $p \in \mathbb{N}$, on a montré l'existence d'un couple $(O_p, S_p) \in \mathrm{U}_n(\mathbb{C}) \times \mathcal{H}_n^{++}(\mathbb{C})$ tel que $M_p = O_p S_p$, et on a aussi l'existence d'un couple $(O, S) \in \mathrm{U}_n(\mathbb{C}) \times \mathcal{H}_n^{++}(\mathbb{C})$ tel que $M = OS$. Il s'agit de vérifier que la suite $(O_p)_{p\in\mathbb{N}}$ converge vers O et que la suite $(S_p)_{p\in\mathbb{N}}$ converge vers S pour avoir la continuité de κ.

D'après le lemme 4.24, $\mathrm{U}_n(\mathbb{C})$ est une partie compacte de $\mathcal{M}_n(\mathbb{C})$. Donc, d'après le théorème de Bolzano-Weierstrass, la suite $(O_p)_{p\in\mathbb{N}}$ admet une sous-suite convergente $(O_{\psi(p)})_{p\in\mathbb{N}}$ avec $\psi: \mathbb{N} \to \mathbb{N}$ une fonction strictement croissante. On note $O' \in \mathrm{U}_n(\mathbb{C})$ la limite de cette sous-suite. Alors la sous-suite $(S_{\psi(p)})_{p\in\mathbb{N}}$ converge vers une matrice $S' := O'^{-1} M \in \mathcal{H}_n^{+}(\mathbb{C}) \cap \mathrm{GL}_n(\mathbb{C}) = \mathcal{H}_n^{++}(\mathbb{C})$, et on a

$$OS = M = \lim_{p\to+\infty} M_{\psi(p)} = \lim_{p\to+\infty} O_{\psi(p)} S_{\psi(p)} = O'S'.$$

Et donc, par unicité de la décomposition polaire, on a forcément $O = O'$ et $S = S'$. Donc la suite $(O_p)_{p\in\mathbb{N}}$ admet une unique valeur d'adhérence, à savoir O, et donc elle converge vers celle-ci. Il s'ensuit que la suite $(S_p)_{p\in\mathbb{N}}$ est également convergente de limite $O^{-1}M = S$. On a donc montré la continuité de κ, ce qui termine la démonstration du théorème. □

Remarques 4.28.

(i) Si $n = 1$, alors le théorème 4.27 redonne la décomposition polaire classique $z = \rho e^{i\theta}$, ce qui justifie le nom de ce théorème.

(ii) On peut aisément étendre l'existence d'une décomposition polaire aux matrices complexes non-inversibles, mais alors on perd l'unicité de la décomposition. Plus précisément, toute matrice complexe se décompose en produit d'une matrice unitaire et d'une matrice hermitienne positive, mais pas nécessairement de façon unique.

(iii) Le fait qu'une application bijective entre deux espaces topologiques soit continue n'implique pas que son inverse soit continu est général. Par exemple l'application

$$h: [0; 2\pi[\to \mathbb{S}^1 = \{z \in \mathbb{C},\ |z| = 1\},\ \theta \mapsto e^{i\theta}$$

est une bijection continue, mais ce n'est pas un homéomorphisme puisque $\mathbb{S}^1$ est une partie compacte de $\mathbb{C}$ tandis que $[0; 2\pi[$ n'est pas une partie compacte de $\mathbb{R}$.

Quelques références bibliographiques

- Des références générales concernant la théorie des formes sesquilinéaires et des formes hermitiennes :
  - Rémi Goblot. *Algèbre linéaire.*
  - Xavier Gourdon. *Algèbre*, Chapitre 5.
  - Joseph Grifone. *Algèbre linéaire*, Chapitres 8 et 10.
- Une référence pour l'étude des groupes unitaires généralisés $\mathrm{U}(s,t)$ associés à une forme hermitienne de signature (s,t) :
  - Denis Serre. *Les matrices : théorie et pratique*, Chapitre 7.
- Une référence pour aller plus loin dans l'étude des propriétés topologiques des groupes de matrices :
  - Rached Mneimné et Frédéric Testard. *Introduction à la théorie des groupes de Lie classiques*, Chapitre 1.
- Une référence pour aller plus loin dans la théorie des formes sesquilinéaires et des formes hermitiennes, lorsque l'on remplace $\mathbb{C}$ par un corps quelconque :
  - Daniel Perrin. *Cours d'algèbre*, Chapitre V.