UNIVERSITÉ DE POITIERS
Mathématiques
Agrégation 2008/2009
Paul Broussous

Préparation à l'Agrégation de Mathématiques

Algèbre linéaire
Réduction des endomorphismes

Avant Propos

Nous supposerons connues les notions de base d'algèbre linéaire (programme de L1 et L2). En particulier nous demandons aux étudiants de réviser les notions de *base*, *dimension*, *rang* et *déterminant*. Le *théorème de la base incomplète* sera d'utilisation systématique.

Faute de temps, nous ne développerons que certains points du programme : ceux qui semblent poser le plus de difficultés aux agrégatifs. En particulier nous traiterons les points suivants : *espaces quotients*, *dualité*, *réduction des endomorphismes*. Nous illustrerons le cours par des exercices et des extraits de problèmes d'écrit d'Agrégation récents.

La référence principale à la base de ces notes est *Algèbre*, Patrice Tauvel, 2nde édition, Dunod. Nous conseillons aussi le *Cours de Mathématiques*, tome 1, *Algèbre*, d'Arnaudiès et Fraysse, chez Dunod.

La page web de Pascal Boyer (http ://www.institut.math.jussieu.fr/ boyer/) est très bien faite et propose des pistes intéressantes d'exercices et de développements à l'oral.

Nous utiliserons les notations suivantes.

– k sera le plus souvent le corps de base.
– k^n est l'espace vectoriels des vecteurs lignes ou n-uplets.
– E, F, G, U, V, W, ..., désigneront des espaces vectoriels.
– u, v, w, ..., désigneront des vecteurs, et parfois des applications linéaires.
– 0_E désigne le vecteur nul de E et id_E l'identité de E.
– f, g, h, φ, ψ désignerons des applications linéaires.
– $\mathfrak{L}_k(E,F) = \mathfrak{L}(E,F)$ est l'espace vectoriel des applications linéaire de E dans F.
– $\mathfrak{L}(E) = \mathrm{End}_k(E) = \mathrm{End}(E)$ est l'algèbre des endomorphismes de E.
– Ker et Im désigne le noyau et l'image d'une application linéaire.
– dim désigne la dimension d'un espace vectoriel.
– rg désigne le rang d'une application linéaire, d'une matrice, ou d'une famille de vecteurs.
– $\mathrm{Mat}(\mathcal{B}, \mathcal{B}', f)$ désigne la matrice de l'application linéaire f dans des bases $\mathcal{B}$ et $\mathcal{B}'$.
– Det désigne le déterminant d'une matrice ou d'un endomorphisme.
– $\mathrm{M}(n,k)$, $\mathrm{M}(n,m,k)$ sont les algèbres des matrices carrées $n \times n$ ou rectangulaires $n \times m$.

Première partie

1. Sommes, produits, espaces quotients, supplémentaires

Soit I un ensemble fini ou infini d'indices et $(E_i)_{i\in I}$ une famille d'espaces vectoriels. On munit le produit $\prod_{i\in I} E_i$, c'est-à-dire l'ensemble des suites $(v_i)_{i\in I}$, où pour chaque i, $v_i \in E_i$, des lois suivantes :

– addition $(v_i)_{i\in I} + (w_i)_{i\in I} = (v_i + w_i)_{i\in I}$,
– multiplication par un scalaire $\lambda(v_i)_{i\in I} = (\lambda v_i)_{i\in I}$, $\lambda \in k$.

1.1. Lemme-Définition. i) *Muni de ces deux lois $\prod_{i\in I} E_i$ est un k-espace vectoriel appelé l'espace vectoriel produit de la famille $(E_i)_{i\in I}$.*
ii) *Les projections canoniques $p_{i_o} : \prod_{i\in I} E_i \longrightarrow E_{i_o}$, $(v_i)_{i\in I} \mapsto v_{i_o}$, $i_o \in I$ sont linéaires.*

Démonstration. Laissée en exercice.

Rappelons que si E et F sont des espaces vectoriels, alors $\mathcal{L}(E, F)$ est naturellement un k-espace vectoriel.

1.2. Proposition. *Soit E un espace vectoriel et $(F_i)_{i\in I}$ une famille d'espaces vectoriels. Il existe un isomorphisme naturel entre espaces vectoriels :*

$$\varphi \ : \ \prod_{i\in I} \mathcal{L}(E, F_i) \longrightarrow \mathcal{L}(E, \prod_{i\in I} F_i) \ .$$

Démonstration. Un élément du produit $\prod_{i\in I} \mathcal{L}(E, F_i)$ est une suite $f = (f_i)_{i\in I}$, où pour chaque $i \in I$, $f_i : E \longrightarrow F_i$ est une application linéaire. Définissons $\varphi(f)$ comme associant à $v \in E$, la suite $(f_i(v))_{i\in I}$.

Notons $p_{i_o} : \prod_{i\in I} F_i \longrightarrow F_{i_o}$, $i_o \in I$, les projections canoniques. Définissons une application linéaire

$$\psi \ : \ \mathcal{L}(E, \prod_{i\in I} F_i) \longrightarrow \prod_{i\in I} \mathcal{L}(E, F_i)$$

par $\psi(g) = (p_i \circ g)_{i\in I}$. Si $f \in \prod_{i\in I} \mathcal{L}(E, F_i)$, $\psi \circ \varphi(f)$ est la famille $(p_i \circ \varphi(f))_{i\in I}$, avec pour $v \in E$, $p_i \circ \varphi(f)(v) = p_i(f_j(v))_{j\in I} = f_i(v)$, c'est-à-dire $p_i \circ \varphi(f) = f_i$. Donc $\psi \circ \varphi$ est l'application identité de l'espace $\prod_{i\in I} \mathcal{L}(E, F_i)$. Nous laissons le soin au lecteur de montrer en exercice que $\varphi \circ \psi$ est l'application identité de $\mathcal{L}(E, \prod_{i\in I} F_i)$.

Si $(E_i)_{i\in I}$ est une famille d'espaces vectoriels, on note

$$\coprod_{i\in} E_i \ \text{ ou encore } \bigoplus_{i\in I} E_i \ ,$$

le sous-ensemble de $\prod_{i\in I} E_i$ formé des suites $(v_i)_{i\in I}$ telles que v_i est nul sauf pour un nombre fini d'indices i.

1.3. Lemme-Définition. i) *Avec les notations précédentes, $\coprod_{i\in I} E_i$ est un sous-espace vectoriel de $\prod_{i\in I} E_i$ qu'on appelle* somme directe (externe) *de la famille $(E_i)_{i\in I}$.*
ii) *Pour chaque $i_o \in I$, la restriction de p_{i_o} s'appelle la* projection canonique *de $\coprod_{i\in I} E_i$ sur E_{i_o}.*

iii) *Pour chaque $i_o \in I$, l'application $j_{i_o} : E_{i_o} \longrightarrow \coprod_{i\in I} E_i$, donné par $[j_{i_o}(v)]_i = v$, si $i = i_o$, $[j_{i_o}(v)]_i = 0$, si $i \neq i_o$, est linéaire et s'appelle* l'injection canonique *de E_{i_o} dans $\coprod_{i\in I} E_i$.*

Démonstration. Laissée en exercice.

1.4. *Remarques.* i) Supposons I fini. Alors le produit et la somme directe d'une famille d'espaces vectoriels coïncident.
ii) Supposons que les E_i soient tous égaux à un même espace vectoriel E. Alors $\prod_{i\in I} E_i$ s'identifie à l'espace vectoriel E^I des applications de I dans E. Dans cette identification, l'espace $\coprod_{i\in I} E_i$ correspond au sous-espace $E^{(I)}$ de E^I formé des applications à support fini ; le support d'une application $f : I \longrightarrow E$ étant $\mathrm{Supp}(f) = \{i \in I \ ; \ f(i) \neq 0_E\}$.

1.5. Proposition-Définition. *Soit I un ensemble. Pour $i_o \in I$, soit $f_{i_o} \in k^{(I)}$ l'application définie par $f_{i_o}(i) = \delta_{ii_o}$. Alors la famille $(f_i)_{i\in I}$ est une base de $k^{(I)}$,* appelée base canonique.

Démonstation. Laissée en exercice.

1.6. Proposition. *Soient $(E_i)_{i\in I}$ une famille d'espaces vectoriels et F un espace vectoriel. Il existe un isomorphisme naturel d'espaces vectoriels :*

$$\varphi \ : \mathfrak{L}(\coprod_{i\in I} E_i, F) \longrightarrow \prod_{i\in I} \mathfrak{L}(E_i, F) \ .$$

Démonstration. En effet φ associe à une application linéaire $f : \coprod_{i\in I} E_i \longrightarrow F$ la famille d'applications linéaires $(f \circ j_i)_{i\in I}$, où les j_i sont les injections canoniques. On vérifie facilement que φ a pour inverse l'application ψ donnée par

$$\psi((f_i)_{i\in I}) \ : \ (v_i)_{i\in I} \in \coprod_{i\in I} E_i \mapsto \sum_{i\in I} f_i(v_i) \in F \ .$$

Soit E un espace vectoriel et $(E_i)_{i\in I}$ une famille de sous-espaces vectoriels de E. Rappelons que la somme $\sum_{i\in I} E_i$ est le sous-espace vectoriel de E engendré par la réunion des sous-ensembles E_i, $i \in I$. C'est aussi l'ensemble des sommes $\sum_{i\in I} v_i$, où $(v_i)_{i\in I}$ décrit l'ensembles des familles de vecteurs nulles sauf pour un nombre fini d'indices. On a donc une application surjective :

$$p \ : \coprod_{i\in I} E_i \longrightarrow \sum_{i\in I} E_i \subset E \ , \ , (v_i)_{i\in I} \mapsto \sum_{i\in I} v_i \ .$$

On dit que la somme $\displaystyle\sum_{i\in I} E_i$ est *directe* si p est injective, c'est-à-dire si p est un isomorphisme. Si c'est le cas, on écrit

$$\bigoplus_{i\in I} E_i \text{ pour } \sum_{i\in I} E_i$$

Exercice **1.** Soit $(E_i)_{i\in I}$ une famille de sous-espaces vectoriels d'un espace vectoriel E. Posons

$$F = \sum_{i\in I} E_i \ , \ G = \coprod_{i\in I} E_i \ .$$

Montrer que les conditions suivantes sont équivalentes :

(i) La somme $\sum_{i\in I} E_i$ est directe.
(ii) Pour tout $x \in F$, il existe un unique suite $(x_i) \in G$ telle que $x = \sum_{i\in I} x_i$.
(iii) Si $(x_i) \in G$ vérifie $\sum_i x_i = 0$, alors $x_i = 0$ quelque soit $i \in I$.
(iv) Pour tout $i \in I$, on a

$$E_i \cap \big(\sum_{j\in I\setminus\{i\}} E_j\big) = \{0\} \ .$$

En outre, si I est l'ensemble fini $\{1, 2, \ldots, n\}$, les conditions précédentes sont équivalentes à

$$E_i \cap (E_1 + \cdots + E_{i-1}) = \{0\}$$

pour $i = 2, \ldots, n$.

Exercice **2.** Soit E le $\mathbb{R}$-espace vectoriel des fonctions $f : \mathbb{R} \longrightarrow \mathbb{R}$ continues, nulles sur $\mathbb{Z}$ et nulles en dehors d'un intervalle borné (dépendant de f). Pour chaque $n \in \mathbb{Z}$, soit E_n le sous-espace vectoriel de E formé des applications continues nulles en dehors de $[n, n + 1]$. Montrer que $E = \bigoplus_{n\in\mathbb{Z}} E_n$.

Deux sous-espaces vectoriels U et V d'un espace vectoriel E sont dits *supplémentaires* si $U \oplus V = E$. Tout vecteur de x de E s'écrit alors de façon unique $u + v$, avec $u \in U$ et $v \in V$. On dit que V est un *supplémentaire* de U L'application p qui à $x \in E$ associe u s'appelle la projection sur U parallèlement à V. Cette application est un *projecteur* (ou *idempotent*) de l'anneau $\mathrm{End}(E)$ au sens où $p \circ p = p$. Réciproquement, si p est un projecteur de E, on vérifie aisément que $\mathrm{Ker}\ p$ et $\mathrm{Im}\ p$ sont supplémentaires.

Exercice **3.** Soit $(p_i)_{i\in 1,\ldots n}$ une famille de projecteurs d'un espace vectoriel E. On dit que cette famille est *orthogonale* si pour tout couple d'indices (i, j) vérifiant $i \neq j$, on a $p_i \circ p_j = 0$. Posons $E_i = \mathrm{Im}\ p_i$, $i = 1, \ldots, n$. Montrer que les assertions suivantes sont équivalentes :

(i) E est la somme directe des E_i, $i = 1, \ldots, n$.
(ii) La famille de projecteurs $(p_i)_{i=1,\ldots,n}$ est orthogonale et on a $\mathrm{id}_E = p_1 + \cdots + p_n$.

Soient E un espace vectoriel et F un sous-espace de E. Considérons dans un premier temps E et F comme groupes abéliens (pour l'addition des vecteurs). Rappelons que le groupe quotient $(E/F, +)$ est l'ensemble quotient $E/\mathcal{R}$ pour la relation d'équivalence :

$$x \ \mathcal{R} \ y \ \ \text{ssi} \ \ x - y \in F \ ,$$

muni de l'addition

$$\bar{x} + \bar{y} = \overline{x + y} \ .$$

Ici $\bar{u}$ désigne la classe d'équivalence pour la relation $\mathcal{R}$ d'un élément u de E, et s'appelle la *classe de u modulo F*. Si $u \in E$, la classe de x est l'ensemble $\{u + f \ ; \ f \in F\}$ et se note aussi $u + F$. Noter que le groupe $(E/F, +)$ est abélien.

1.7. Proposition-Définition. i) *L'application*

$$m \ : \ k \times E/F \longrightarrow E/F \ , \ (\lambda, \overline{x}) \mapsto \overline{\lambda x}$$

est bien définie et munit le groupe abélien $(E/F, +)$ d'une structure de k-espace vectoriel appelé espace vectoriel quotient *de E par F.*
ii) *L'application $p : E \longrightarrow E/F$, $x \mapsto \overline{x}$ est linéaire et surjective, et s'appelle la* projection canonique *de E sur F*

Démonstration. Il s'agit dans un premier temps de montrer que $m(\lambda, x + F) = \overline{\lambda x}$ ne dépend pas du représentant x choisi dans la classe $\overline{x}$. Soit $y \in E$ un autre représentant, de sorte que $y = x + f$, avec $f \in F$. On a $\lambda y = \lambda x + \lambda f$, avec $\lambda f \in F$ puisque F est un sous-espace de E. Donc $\overline{\lambda x} = \overline{\lambda y}$, ce qu'il fallait démontrer.

Dans un second temps, il s'agit de démontrer que E muni de l'addition $+$ et de la loi externe m vérifie les axiomes d'un espace vectoriel. Montrons par exemple que $m(\lambda, (\overline{x} + \overline{y}) = m(\lambda, \overline{x}) + m(\lambda, \overline{y})$, quel que soit $\lambda \in k$ et quels ques soient $\overline{x}, \overline{y} \in E/F$. Le membre de gauche est par définition $\overline{\lambda(x + y)}$, tandis que le second est $\overline{(\lambda x) + (\lambda y)}$, et ils coïncident. (Noter que l'on s'est implicitement servi du fait que les formules donnant l'addition et la multiplication externe ne dépendent pas des représentants choisis pour faire le calcul). La vérification des autres axiomes est similaire. Notons simplement que l'élément neutre de E/F est $0_{E/F} = \overline{0}$.

L'application p est surjective car c'est le projection canonique d'un ensemble sur son quotient par une relation d'équivalence. Elle est linéaire par définition même de m.

Dorénavant nous noterons $\lambda \bar{x}$ pour $m(\lambda, \bar{x})$.

1.8. Théorème. (Propriété universelle du quotient). *Soient E, G deux espaces vectoriels et F un sous-espace de E. Soit $u : E \longrightarrow G$ une application linéaire telle que* Ker $u \supset F$. *Alors il existe une unique application $\bar{u} : E/F \longrightarrow G$ telle que $\bar{u} \circ p = u$, où $p : E \longrightarrow E/F$ est la projection canonique. Cette application est de plus linéaire.*

Démonstration. Supposons que $\bar{u}$ existe et soit $\mathbf{x} \in E/F$. Alors $\bar{u}(\mathbf{x}) = a$ pour tout $a \in E$ tel que $p(a) = \mathbf{x}$, c'est-à-dire pour tout $a \in \mathbf{x}$. Ceci prouve que $\bar{u}$ est uniquement déterminée. Réciproquement, la relation $\bar{u}(\mathbf{x}) = u(a)$ si $a \in \mathbf{x}$ définit une application $\bar{u} \ : E/F \longrightarrow G$ sans ambiguité. En effet si $a, b \in \mathbf{x}$, on a $a - b \in F \subset$ Ker u, de sorte que $u(a) = u(b)$. On vérifie sans peine que $\bar{u}$ ainsi définie est linéaire.

1.9. Théorème. (Factorisation canonique d'une application linéaire.) *Soient $u : E \longrightarrow F$ une application linéaire. Notons $p : E \longrightarrow E/$*Ker u *et $i :$* Im $u \longrightarrow F$ *la projection et l'injection canoniques. Il existe une unique application $\bar{u}$: $E/$*Ker $u \longrightarrow$ Im u *telle que $i \circ \bar{u} \circ p = u$, et cette application est un isomorphisme de k-espaces vectoriels.*

Démonstration. L'unicité modulo l'existence se montre comme dans le théorème prédédent. Ce même théorème fournit l'existence d'une application $\bar{u} : E/$Ker $u \longrightarrow F$ telle que $\bar{u} \circ p = u$ et il s'agit de montrer que $\bar{u}$ est injective, d'image Im u. Si $y \in$ Im u, alors y s'écrit $u(x)$, $x \in E$, donc s'écrit $\bar{u}(\bar{x})$ par définition de $\bar{u}$. Ainsi $\bar{u}$ est surjective. Soit

$\mathbf{x} \in \text{Ker } \bar{u}$ avec $\mathbf{x} = \bar{x}$, $x \in E$. Par définition, on a $u(x) = 0$, donc $x \in \text{Ker } u$, donc $\mathbf{x} = 0$, par définition du quotient $E/\text{Ker } u$.

1.10. Théorème. *Soient E un espace vectoriel, F un sous-espace de E et $p : E \longrightarrow E/F$ l'application canonique. Alors l'application $\pi : V \mapsto p(V)$ est une bijection de l'ensemble $\mathcal{V}_F(E)$ des sous-espaces vectoriels de E contenant F sur l'ensemble $\mathcal{V}(E/F)$ des sous-espaces vectoriels de E/F. La bijection réciproque est $\sigma : W \mapsto p^{-1}(W)$.*

Démonstration. Laissée au lecteur en exercice.

Exercice **4.** Soit E un espace vectoriel et $G \subset F \subset E$ des sous-espaces. Montrer que F/G est naturellement un sous-espace vectoriel de E/G. Montrer qu'il existe un isomorphisme naturel entre les quotients E/F et $(E/G)/(F/G)$ (*premier théorème d'isomorphisme* de Noether).

Exercice **5.** Soit E un espace vectoriel et F et G deux sous-espaces de E. Montrer qu'il existe un isomorphisme naturel entre les quotients $(F+G)/G$ et $F/F \cap G$ (*second théorème d'isomorphisme* de Noether).

Exercice **6.** Soit $(E_i)_{i\in I}$ une famille d'espace vectoriel et, pour chaque $i \in I$, soit F_i un sous-espace de E_i. Montrer qu'il existe des isomorphismes naturels :

$$\prod_{i\in I}(E_i/F_i) \simeq (\prod_{i\in I} E_i)/(\prod_{i\in I} F_i)$$

$$\coprod_{i\in I}(E_i/F_i) \simeq (\coprod_{i\in I} E_i)/(\coprod_{i\in I} F_i)$$

1.11. Théorème-Définition. i) *Si F est un sous-espace d'un espace vectoriel E, tout supplémentaire de F dans E est isomorphe à E/F. En particulier, si E/F est de dimension finie, tous les supplémentaires de F ont la même dimension, nombre qui s'appelle la* codimension *de F dans E, et se note* $\text{Codim}_E(F)$.
ii) *Soit E un espace vectoriel et F un sous-espace de E. . Les assertions suivantes sont équivalentes :*

a) *Il existe un forme linéaire non-nulle $f \in \mathcal{L}(E,k)$ telle que $F = \text{Ker } f$.*
b) *Le sous-espace F est de codimension 1 dans E.*
c) *Le sous-espace F admet une droite vectorielle comme supplémentaire.*

Si ces conditions sont vérifiées, on dit que F est un hyperplan de E.

Démonstration. i) Soit G un supplémentaire de F et $f : G \longrightarrow E/F$ la restriction de la projection canonique $p : E \longrightarrow E/F$ à G. On a $\text{Ker } f = \text{Ker} p \cap G = F \cap G = \{0\}$, donc f est injective. Soit $\mathbf{x} \in E/F$. Un représentant de $\mathbf{x}$ s'écrit $f + g$, $f \in F$, $g \in G$, d'où l'on tire $\mathbf{x} = \bar{f + g} = \bar{g} = f(g)$. Donc f est surjective.
ii) L'équivalence entre b) et c) découle du point i). Supposons que $F = \text{Ker } f$ avec $f \in \mathcal{L}(E,k)$ non nulle. Alors f est surjective et induit un isomorphisme $E/F \simeq k$, ce qui prouve que E/F est de dimension 1. Réciproquement, si E/F est de dimension 1, alors il existe un isomorphisme $\varphi : E/F \longrightarrow k$. Alors la forme linéaire non nulle $f \circ p$, où $p : E \longrightarrow E/F$ est la projection canonique, a pour noyau f.

2. Dualité

Le *dual algébrique*, ou plus simplement *dual*, d'un k-espace vectoriel E est l'espace $\mathcal{L}(E,k)$ des formes linéaires de E. Il se note E^*. A partir d'un vecteur $x \in E$ et d'une forme linéaire $\varphi \in E^*$, on peut construire le nombre $\varphi(x)$. Celui-ci se note aussi $\langle x, \varphi \rangle$. Cette notation se comprend mieux si l'on réalise que l'application :

$$E \times E^* \longrightarrow k \ , \ (x,\varphi) \mapsto \langle x, \varphi \rangle \ ,$$

est *bilinéaire* (c'est-à-dire linéaire en x et en φ). On l'appelle l'appelle la *forme bilinéaire canonique* sur $E \times E^*$, ou encore *crochet de dualité*.

Soient E et F deux espaces vectoriels et $u : E \longrightarrow F$ une application linéaire. On appelle *transposée* de u l'application ${}^t u : F^* \longrightarrow E^*$, donnée par ${}^t u(\varphi) = \varphi \circ u$. On vérifie aisément que ${}^t u$ est linéaire.

2.1. Proposition. *Soient E, F, G trois espaces vectoriels.* i) *L'application*

$$\mathcal{L}(E,F) \longrightarrow \mathcal{L}(F^*,E^*) \ , \ u \mapsto {}^t u$$

est linéaire.
ii) *On a* ${}^t\mathrm{id}_E = \mathrm{id}_{E^*}$.
iii) *Si $u \in \mathcal{L}(E,F)$ et $v \in \mathcal{L}(F,G)$, on a*

$${}^t(v \circ u) = {}^t u \circ {}^t v \ .$$

iv) *Si $u \in \mathcal{L}(E,F)$ est un isomorphisme, alors ${}^t u \in \mathcal{L}(F^*,E^*)$ est un isomorphisme d'inverse*

$$({}^t u)^{-1} = {}^t(u^{-1}) \ .$$

Démonstration. Routine laissée au lecteur.

Noter que si $u \in \mathcal{L}(E,F)$, alors ${}^t u$ est l'unique application $v : F^* \longrightarrow E^*$ telle que

$$\langle x, v(\varphi) \rangle = \langle u(x), \varphi \rangle \ , \ (x,\varphi) \in E \times F^* \ .$$

Noter ainsi la similarité avec l'adjoint d'un endomorphisme relativement à une forme bilinéaire non dégénérée. Exploitant cette similarité, nous introduisons le vocabulaire suivant.

Soit E un espace vectoriel.

Un élément x de E est dit *orthogonal* à un élément φ de E^* si $\langle x, \varphi \rangle = 0$.

Soit $X \subset E$. L'*orthogonal* de X, noté $X^\perp$, est l'ensemble des $\varphi \in E^*$ vérifiant $\langle x, \varphi \rangle = 0$, pour tout $x \in X$. C'est un sous-espace vectoriel de E^*.

Soit $X^* \subset E^*$. L'*orthogonal* $(X^*)^\top$ de X^* est l'ensemble des $x \in E$, vérifiant $\langle x, \varphi \rangle = 0$, pour tout $\varphi \in X^*$. C'est le sous-espace vectoriel de E donné par $\cap_{\varphi \in X^*} \mathrm{Ker}\ \varphi$.

2.2. Proposition. *Soient X, Y, $(X_i)_{i \in I}$ des parties de E, et X^*, Y^*, $(x_i^*)_{i \in I}$ des parties de E^*.*

i) *on a les relations suivantes :*

$$X \subset Y \implies Y^\perp \subset X^\perp \ , \ X^* \subset Y^* \implies (Y^*)^\top \subset (X^*)^\top \ ,$$

$$X^\perp = \mathrm{Vect}(X)^\perp \ , \ (X^*)^\top = \mathrm{Vect}(X^*)^\top \ , \ X \subset (X^\perp)^\top \ , \ X^* \subset ((X^*)^\top)^\perp \ ,$$

$$(\bigcup_{i\in I} X_i))^\perp = \bigcap_{i\in I} X_i^\perp \ , \ (\bigcup X_i^*)^\top = \bigcap_{i\in I} (X_i^*)^\top \ .$$

ii) *Les conditions suivantes sont équivalentes :*

a) $\langle x, x^* \rangle = 0$ *pour tout* $(x, x^*) \in X \times X^*$.
b) $X^* \subset X^\perp$
c) $X \subset (X^*)^\top$.

Démonstration. Facile. Laissée au lecteur en exercice.

Le procédé suivant est une astuce centrale en mathématiques. Notons E^{**} le dual du dual de E. Cet espace est appelé le *bidual* de E. On notera encore par un crochet l'application bilinéaire canonique $E^* \times E^{**} \longrightarrow k$. Soit $x \in E$. L'évaluation d'un forme linéaire en x, c'est-à-dire l'application

$$E^* \longrightarrow k \ , \ x^* \mapsto x^*(x) = \langle x, x^* \rangle$$

est linéaire et définit donc un élément $\hat{x}$ du bidual E^*. L'application

$$c_E \ : \ E \longrightarrow E^{**} \ , \ x \mapsto \hat{x}$$

est linéaire et s'appelle l'*application canonique de E dans son bidual.*

2.3. Théorème. *L'application c_E est injective. En outre, si E est de dimension finie, elle est bijective.*

Démonstration. Soit $(e_i)_{i\in I}$ une base de E (on admet l'existence de bases en dimension infinie). Tout $x \in E$ s'écrit de façon unique

$$x = \sum_{i\in I} x_i e_i$$

pour une famille de scalaires $(x_i)_{i\in I} \in k^{(I)}$. Pour chaque $i \in I$, l'application $f_i : E \longrightarrow k$, $x \mapsto x_i$, est un forme linéaire. Soit x un élément de Ker c_E. Avec les notations précédentes, on a $x_i = f_i(x) = \hat{x}(f_i) = 0$, quelque soit i. Donc $x = 0$.

Supposons à présent I fini (*i.e.* E de dimension finie Card(I)). Si $f \in E^*$ et $x \in E$, on a

$$f(x) = \sum_{i\in I} x_i f(e_i) = \sum_{i\in I} f(e_i) f_i(x) \ .$$

Donc la famille $(f_i)_{i\in I}$ est génératrice de E^*. Si $(\lambda_i)_{i\in I}$ est une famille de scalaires, on a

$$\sum_{i\in I} \lambda_i f_i = 0 \quad \implies \quad \lambda_j = \sum_{i\in I} \lambda_i f_i(e_j) = 0 \ , \ j \in I \ .$$

Ainsi $(f_i)_{i\in I}$ est une base de E^*. Ceci prouve que $\dim(E) = \dim(E^*)$. De même $\dim(E^*) = \dim(E^{**})$ et donc $\dim(E) = \dim(E^{**})$. L'application c_E étant injective, elle est aussi surjective.

Si E est de dimension finie, la base $(f_i)_{i\in I}$ s'appelle la *base duale* de la base $(e_i)_{i\in I}$ et se note $(e_i^*)_{i\in I}$.

Exercice **7.** (i) Soit I un ensemble infini et $E = k^{(I)}$. Montrer que E^* est naturellement isomorphe à k^I de sorte que le crochet de dualité entre E et E^* est donné par :

$$\langle (x_i)_{i\in I}, (\lambda_i)_{i\in I}\rangle = \sum_{i\in I} x_i\lambda_i \ , \ (x_i)_{i\in I} \in k^{(I)} \ , \ (\lambda_i)_{i\in I} \in k^I \ .$$

(ii) Identifiant E^* et k^I, montrer qu'il existe une forme linéaire non nulle $\varphi \in E^{**}$ qui contienne $k^{(I)}$ dans son noyau. Montrer que cette forme linéaire ne peut être dans l'image de c_E.

Exercice **8.** Soit E un espace vectoriel. Montrer que $(E^*)^\top = \{0\}$.

2.4. Théorème. *Soit E un espace vectoriel et x^*, y^* des formes linéaires sur E. On a* $\mathrm{Ker}\ x^* = \mathrm{Ker}\ y^*$ *si et seulement s'il existe $\lambda \in k$, $\lambda \neq 0$, tel que $y^* = \lambda x^*$.*

Démonstration. Supposons (seul cas non trivial) que les formes linéaires ne sont pas nulles. Si $\mathrm{Ker}\ x^* = \mathrm{Ker}\ y^*$, alors ces deux noyaux sont un hyperplan H de E. Soit D un supplémentaire de cet hyperplan et a un vecteur qui engendre D. Alors $x^*(a)$ et $y^*(a)$ ne sont pas nuls. Si $\lambda = y^*(a)/x^*(a)$, on vérifie facilement que $y^* = \lambda x^*$. La récriproque est évidente.

2.5. Proposition. *Soit E un espace vectoriel et soit F un sous-espace de E différent de E. Soit $\mathcal{H}$ l'ensemble des hyperplans de E contenant F· Alors F est l'intersection des éléments de $\mathcal{H}$. Autrement dit :*

$$F = \bigcap_{x^*\in E^* \ , \ \mathrm{Ker}\ x^* \supset F} \mathrm{Ker}\ x^* \ .$$

Démonstration. Notons d'abord que $\mathcal{H}$ n'est pas vide. En effet l'espace quotient E/F n'est pas nul et possède donc une forme linéaire φ non nulle. Cette forme peut se voir comme une forme linéaire de E dont le noyau contient F. (Variante : construire une telle forme à partir d'une base de E complétée d'une base de F).

Une inclusion est évidente. Soit $x \notin F$ et soit G un supplémentaire de $F \oplus kx$. Alors $F \oplus G$ est un hyperplan et il existe un forme linéaire φ de noyau $F \oplus G$. En particulier $x \notin \mathrm{Ker}\ \varphi$, ce qui prouve l'inclusion opposée.

2.6. Lemme. *Soient F un sous-espace de E et j l'injection canonique de F dans E. Alors ${}^t j$ est surjective. Si F est distinct de E, ${}^t j$ n'est pas injective.*

Démonstration. Il faut montrer que si $y^* \in F^*$, il existe $x^* \in E^*$ tel que $y^* = {}^t x^* = x^* \circ j$. Autrement dit, il faut montrer que y^* se prolonge en une forme linéaire de E. Pour cela on peut par exemple prolonger y^* par zéro sur un supplémentaire de F dans E. Si $F \neq E$, un tel supplémentaire est non nul, de sorte qu'il y a différentes façons de prolonger y^* : ${}^t j$ n'est pas injective.

2.7. Théorème. *Soient E, F des espaces vectoriels et $u \in \mathcal{L}(E, F)$.*

(i) *L'application* $\Psi : \mathcal{L}(E,F) \longrightarrow \mathcal{L}(F^*,E^*)$, *donnée par* $v \mapsto {}^t v$ *est injective. Si* E *et* F *sont de dimension finie, elle est bijective.*
(ii) *On a* $\text{Im } {}^t u = (\text{Ker } u)^\perp$, *et* ${}^t u$ *est surjectif si et seulement si* u *est injectif.*
(iii) *On a* $\ker {}^t u = (\text{Im } u)^\perp$, *et* ${}^t u$ *est injectif si et seulement si* u *est surjective.*

Démonstration. (i) Soit $v \in \mathcal{L}(E,F)$ tel que ${}^t v = 0$. Alors pour tout $y^* \in F^*$ et tout $x \in E$, on a $\langle y^*, v(x)\rangle = \langle {}^t v(y^*), x\rangle = 0$. Donc $\text{Im } v \subset (F^*)^\top = 0$ (exercice 8). Donc $v = 0$.

(ii) La factorisation canonique de $u : E \longrightarrow F$, s'écrit $u = j \circ \bar{u} \circ p$, où $\bar{u} : E/(\text{Ker } u) \longrightarrow \text{Im } u$ est un isomorphisme et $p : E \longrightarrow E/(\text{Ker } u)$ et $j : \text{Im } u \longrightarrow F$ sont les applications canoniques. On obtient ${}^t u = {}^t p \circ {}^t \bar{u} \circ {}^t j$. Or ${}^t \bar{u}$ est bijective et ${}^t j$ est surjective (2.6). Donc $\text{Im } {}^t u = \text{Im } {}^t p$. Mais un élément y^* de E^* est dans l'image de ${}^t p$ si et seulement s'il se factorise à travers $E/(\text{Ker } u)$, c'est-à-dire s'il est trivial sur $\text{Ker } u$. D'où le résultat.

(iii) C'est évident puisque, pour $x^* \in F^*$, on a $x^* \in \text{Ker } {}^t u$ si et seulement si $x^* \circ u = 0$, c'est-à-dire si x^* s'annule sur l'image de u.

2.8. Théorème. *Soit* E *un espace vectoriel de dimension finie. L'application associant à une base de* E *sa duale est une bijection de l'ensemble des bases de* E *sur l'ensemble des bases de* E^*.

Démonstration. La surjectivité de cette aplication se montre de la façon suivante. Soit $\boldsymbol{e}^* = (e_i^*)_{i=1,\dots,n}$ une base de E^*. Il faut montrer qu'elle est la duale d'une base de E. Soit $\boldsymbol{e}^{**} = (e_i^{**})_{i=1,\dots,n}$ la duale de $\boldsymbol{e}^*$ dans E^{**}. On peut rapatrier cette base dans E via c^{-1}, où $c : E \longrightarrow E^{**}$ est la bijection canonique. La base $\boldsymbol{e} = (e_i)_{i=1,\dots,n}$, $e_i = c^{-1}(e_i^{**})$, $i = 1,\dots,n$ a pour duale $\boldsymbol{e}^*$. En effet pour tous $i,j \in \{1,\dots,n\}$, on a

$$\delta_{ij} = \langle e_i^{**}, e_j^*\rangle = \langle e_j^*, c^{-1}(e_i^{**})\rangle \ ,$$

par définition de c. d'où le résultat.

Prouvons l'injectivité de l'application. Si $\boldsymbol{e} = (e_i)_{i=1,\dots,n}$ et $\boldsymbol{f} = (f_i)_{i=1,\dots,n}$ sont des bases ayant toutes deux $\boldsymbol{e}^*$ comme base duale, on a

$$\langle e_j^*, e_i - f_i\rangle = 0 \ , \ i, \ j \in \{1,\dots n\} \ .$$

Donc pour tout i, $e_i - f_i \in (E^*)^\top = \{0\}$ et $\boldsymbol{e} = \boldsymbol{f}$.

Exercice **9.** Supposons le corps k de caractéristique 0, et soit E l'espace vectoriel $k_n[X]$, pour un entier $n \geqslant 0$· Pour $k = 0,\dots,n$, on note φ_k le forme linéaire sur E donnée par

$$P \mapsto P^{(k)}(0) \ .$$

Montrer que $(\varphi_0,\dots,\varphi_n)$ est une base de E^* et déterminer la base de E dont elle est la duale.

Exercice **9.** Soit E l'espace vectoriel $\mathcal{F}(k,k)$ des applications de k dans lui-même. On considère des éléments f_1, f_2, ..., f_n de E linéairement indépendants. En raisonnant par dualité, montrer qu'il existe a_1, a_2, ..., a_n dans k tels que la matrice $[f_i(a_j)]_{i,j\in\{1,\dots,n\}}$ soit inversible.

Exercice **10.** On suppose $k = \mathbb{Q}$ et l'on donne sur $\mathbb{Q}^3$ les trois formes linéaires

$$\varphi_1(x,y,z) = x + 2y - 3z \ , \ \varphi_2(x,y,z) = 5x - 2y \ , \ \varphi_3(x,y,z) = 2x - y - z \ .$$

Montrer que $(\varphi_1, \varphi_2, \varphi_3)$ est une base de $(\mathbb{Q}^3)^*$ et chercher la base de $\mathbb{Q}^3$ dont elle est la duale.

Exercice **11.** Supposons E de dimension finie non nulle et posons $\mathcal{L} = \mathrm{End}(E)$. Pour chaque $v \in \mathcal{L}$ on définit $T_v \in \mathcal{L}^*$ par $T_v(u) = \mathrm{Tr}(uv)$ (Tr désigne l'application trace).

(a) Montrer que l'application $v \mapsto T_v$ est un isomorphisme de $\mathcal{L}$ sur $\mathcal{L}^*$.

(b) Montrer que l'orthogonale de l'espace engendré par $\{\alpha \circ \beta - \beta \circ \alpha \ ; \ \alpha, \ \beta \in \mathcal{L}\}$ est la droite kT_{id_E}.

(c) Pour α, $\beta \in \mathcal{L}$, on pose $[\alpha, \beta] = \alpha\beta - \beta\alpha$ et

$$\mathcal{C}_\alpha = \{u \in \mathcal{L} \ ; \ \alpha \circ u = u \circ \alpha\} \ .$$

Montrer que pour α, $\beta \in \mathcal{L}$ et $\gamma \in \mathcal{C}_\alpha$, on a $\mathrm{Tr}([\alpha, \beta]\gamma) = 0$.

(d) Soient α, $u \in \mathcal{L}$ tels que T_u s'annule sur $\mathcal{C}_\alpha$. Montrer qu'il existe $v \in \mathcal{L}$ tel que $u = [\alpha, v]$.

Exercice **12.** Appliquer l'algorithme de Gauss à la forme quadratique sur $\mathbb{Q}^3$ donnée par $q(x,y,z) = 2x^2 + y^2 + 4xy + xz + yz$. En utilisant des résultats de dualité, déterminer une base de $\mathbb{Q}^3$ orthogonale pour q.

2.9. Théorème. *Soient E un espace vectoriel de dimension finie n , F un sous-espace de E et Φ un sous-espace de E^*.*

(a) *On a :*

$$\dim F + \dim F^\perp = \dim \Phi + \dim \Phi^\top = n \ .$$

(b) *Soient de plus G un sous-espace de E et Ψ un sous-espace de E^*. Alors :*

$$F = (F^\perp)^\top \ , \ (F+G)^\perp = F^\perp \cap G^\perp \ , \ (F \cap G)^\perp = F^\perp + G^\perp \ ,$$

$$\Phi = (\Phi^\top)^\perp \ , \ (\Phi + \Psi)^\top = \Phi^\top \cap \Psi^\top \ , \ (\Phi \cap \Psi)^\top = \Phi^\top + \Psi^\top \ ,$$

Démonstration. (a) Puisque $(E/F)^* \simeq F^\perp$, il vient

$$\dim F^\perp = \dim (E/F)^* = \dim E/F = \dim E - \dim F \ .$$

D'où la première égalité. Pour traiter la seconde, soit $\boldsymbol{e}^* = (e_1^*, \ldots, e_n^*)$ une base de E^* complétée d'une base $(e_1^*, \ldots, e_p^*)$ de Φ. Si $(e_1, e_2, \ldots, e_n)$ est la base de E dont $\boldsymbol{e}^*$ est la duale, il est immédiat que $\Phi^\top = \mathrm{Vect}(e_{p+1}, \ldots, e_n)$. D'où $n = \dim \Phi + \dim \Phi^\top$.
(b) On a facilement $F \subset (F^\perp)^\top$ et $\Phi \subset (\Phi^\top)^\perp$. Mais les égalités de (a) entraînent que ces inclusions sont des inclusions d'espaces de même dimension.

Les égalités $(F+G)^\perp = F^\perp \cap G^\perp$ et $(\Phi + \Psi)^\top = \Phi^\top \cap \Psi^\top$ sont faciles. Il en est de même de l'inclusion $F^\perp + G^\perp \subset F^\perp \cap G^\perp$. Or, en notant $d = \dim (F^\perp + G^\perp)$, on a :

$$d = \dim F^\perp + \dim G^\perp - \dim(F^\perp \cap G^\perp) = \dim F^\perp + \dim G^\perp - \dim(F+G)^\perp$$

$$= n - \dim\, F - \dim\, G - \dim\, (F+G) = n - \dim\, (F \cap G) = \dim(F \cap G)^{\perp}\ .$$

D'où l'égalité dans l'inclusion. La dernière égalité se prouve de façon similaire.

2.10. Théorème. *Soient E et F des espaces vectoriels de dimension finie.*

(a) *L'application $G \mapsto G^{\perp}$ est une bijection décroissante (pour la relation d'inclusion) entre l'ensemble des sous-espaces de E et l'ensemble des sous-espaces de E^*. Sa bijection récipoque est donnée par $\Phi \mapsto \Phi^{\top}$.*

(b) *Pour $u \in \mathcal{L}(E,F)$, on a* $\mathrm{rg}({}^t u) = \mathrm{rg}(u)$.

Démonstration. Le point (a) découle du théorème précédent. Le point (b) découle du même théorème en se servant de l'égalité $\mathrm{Im}\ {}^t u = (\mathrm{Ker}\ u)^{\perp}$.

Soit E un espace vectoriel de dimension finie n. Les derniers résultats que nous avons vus permettent d'introduire la notion d'équation *cartésienne* pour un sous-espace F de dimension p. On a $\dim F^{\perp} = n - p$.

Soit $(x_1^*, \ldots, x_n^*)$ un système de générateurs de $F^{\perp}$. Alors $m \geqslant n - p$, et F est l'orthogonal de $\mathrm{Vect}(x_1^*, \ldots, x_n^*)$. Donc, si $x \in E$, on a :

$$x \in F \quad \text{ssi} \quad \langle x_i^*, x\rangle = 0 \text{ pour} 1 \leqslant i \leqslant m\ .$$

On dit que les relations $\langle x_i^*, x\rangle = 0$ constituent un *système d'équations* du sous-espace F. Noter qu'on peut choisir les x_i^* pour avoir $m = n - p$.

Réciproquement, soit $\boldsymbol{x}^* = (x_1^*, \ldots, x_m^*)$ un système de rang r d'éléments non nuls de E^*. Le système d'équations $\langle x_i^*, x\rangle = 0$, $i = 1$, ..., m, caractérise le sous-espace vectoriel $F = [\mathrm{Vect}(\boldsymbol{x}^*)]^{\top}$ et l'on a $\dim F = n - r$.

On peut extraire de $\boldsymbol{x}^*$ une base $(y_1^*, \ldots, y_r^*)$ de $[\mathrm{Vect}(\boldsymbol{x}^*)]^{\top}$. Si l'on pose $H_i = \mathrm{Ker}\ y_i^*$, on a alors $F = H_1 \cap H_2 \cap \cdots \cap H_r$.

Ainsi, un sous-espace F de dimension p est l'intersection de $n-p$ hyperplans. D'autre part, si P_1, ..., P_s sont des hyperplans d'intersection F, on a $s \geqslant n - p$.

Exercice **13.** (i) Soit u un endomorphisme d'un espace vectoriel E laissant stable toute droite de E. Montrer que u est une homothétie.
(ii) Montrer qu'on a la même conclusion si u laisse stable tous les hyperplans de E. On donnera deux démonstrations, dont l'une utilisera la transposée ${}^t u$.

Seconde partie

3. Polynômes d'endomorphismes

Notons $k[X]$ l'algèbre des polynômes à une indéterminée X. Elle est munie d'un application *degré* : $k[X] \longrightarrow \mathbb{N} \cup \{-\infty\}$, définie de la façon suivante. Si $P = \sum_{n \geqslant 0} a_n X^n$, alors $\deg(P) = -\infty$, si $P = 0$, et $\deg(P) = \max\{n \in \mathbb{N} \ ; \ a_n \neq 0\}$, si $P \neq 0$. Rappelons que l'anneau $k[X]$ est intègre et muni d'une division euclidienne :

Si A et B sont deux polynômes et si B est non nul, il existe un unique couple (Q, R) de polynômes tels que

$$A = BQ + R \text{ et } \deg(R) < \deg(B) \ .$$

On dit que Q est le quotient *et R le* reste *dans la* division euclidienne *de A par B.*

L'existence d'une division euclidienne dans $k[X]$ entraîne que cet anneau est *principal* : tout idéal $\mathcal{I}$ de $k[X]$ est monogène, c'est-à-dire de la forme $(P_\circ) := \{PP_\circ \ ; \ P \in k[X]\}$ pour un polynôme $P_\circ$. Le polynôme $P_\circ$ est uniquement déterminé à une unité de $k[X]$ près, c'est-à-dire à un scalaire non nul près.

Rappelons aussi que l'anneau $k[X]$ est *factoriel* : tout polynôme non nul P s'écrit de façon unique (à l'ordre près)

$$P = u \prod_{i=1}^{s} P_i^{\nu_i} \ ,$$

où u est un scalaire non-nul, les ν_i des entiers non nuls, et les P_i des polynômes *irréductibles unitaires* deux à deux distincts. En particulier on a les notions de pgcd d'une famille de polynômes, de *ppcm*, de famille de polynômes *premiers entre eux* (ou *étrangers*) *dans leur ensemble*. Si $(P_i)_{i=1,\ldots,s}$ est une famille de polynômes de pgcd P, on a l'égalité d'idéaux :

$$(P_1) + \cdots + (P_s) = (P)$$

En particulier si les P_i sont premiers entre eux dans leur ensemble, on a

$$(P_1) + \cdots + (P_s) = (1) = k[X]$$

et il existe des polynômes Q_1, ..., Q_s tels que

$$Q_1 P_1 + \cdots + Q_s P_s = 1 \ .$$

(la réciproque étant bien sûr vraie : c'est le lemme de Bezout).

Dans toute la suite E désigne un espace vectoriel de dimension finie sur $k\cdot$ Si u est un endomorphisme de E et $P = \sum_{k=0}^{d} a_k X^k$ un polynôme, on pose

$$P(u) = a_0 \mathrm{id}_E + a_1 u + a_2 u^2 + \cdots + a_d u^d \ .$$

Si $P(u) = 0$, on dit que P *annule* u, ou encore que P est un *polynôme annulateur* de u.

3.1. Lemme. (i) *Avec les notations précédentes, l'application φ_u, de $k[X]$ dans $\mathcal{L}(E)$, qui à P associe $P(u)$, est un morphisme de k-algèbres (elle est compatible avec l'addition, la multiplication interne et la multiplication par un scalaire).*

(ii) *Le noyau de φ_u n'est pas réduit à $\{0\}$: il existe un polynôme non nul $P_\circ$ tel que $P_\circ(u) = 0$.*

Démonstration. Le point (i) est facile et laissé au lecteur. Pour (ii), notons que puisque l'algèbre $\mathcal{L}(E)$ est de dimension finie, la famille $\{u^k \ ; \ k \in \mathbb{N}\}$ est liée. D'où l'existence de $P_\circ$.

3.2. Lemme-Définition. (i) *Le noyau de φ_u est un idéal de $k[X]$. Puisque cet anneau est principal et que cet idéal est non nul ,* Ker φ_u *est de la forme (μ_u) pour un polynôme unitaire (non nul) μ_u bien déterminé. On l'appelle le* polynôme minimal *de u.*
(ii) *Le polynôme minimal de u est aussi le polynôme unitaire (non nul) de degré minimal parmi les polynômes annulateurs de u.*

Démonstration. Ces assertions très classiques sont laissées au lecteur.

Voici quelques exemples de polynômes minimaux (vérifications laissées en exercices).

1. Le polynôme minimal de l'endomorphisme nul est X.
2. Le polynôme minimal d'un endomorphisme n nilpotent (vérifiant $n^k = 0_E$, pour $k \in \mathbb{N}$ assez grand) est de la forme X^l, pour un certain entier $l \geqslant 1$.
3. Le polynôme minimal d'une homothétie $\lambda \mathrm{id}_E$, $\lambda \in k$, est $X - \lambda$.
4. Si $E = k^2$, x, $y \in k$, avec $y \neq 0$, et u est l'endomorphisme de matrice

$$\begin{pmatrix} x & y \\ 0 & x \end{pmatrix}$$

dans la base canonique, alors $\mu_u(X) = (X - x)^2$.
5. Si $E = k^2$, si -1 n'est pas un carré dans k et si u est l'endomorphisme de matrice

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

alors $\mu_u(x) = X^2 + 1$.

On définit de même le polynôme minimal d'une matrice carrée $A \in \mathrm{M}(n, k)$, $n \geqslant 1$. On remarque que si P est un polynôme, $A \in \mathrm{M}(n, k)$ et $U \in \mathrm{GL}(n, k)$, alors

$$P(UAU^{-1}) = UP(A)U^{-1} \ .$$

Il s'ensuit que deux matrices semblables ont le même polynôme minimal. De plus si $\dim E = n$ et si A est la matrice d'un endomorphisme u de E dans une base $\mathcal{B}$, alors les polynômes minimaux de u et A coïncident.

Exercice **14.** On suppose que la caractéristique de k n'est pas 2. Soit

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

une matrice dans $\mathrm{M}(2, k)$. Soit

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

le polynôme caractéristique de A.
(i) Vérifier à la main, sans utiliser de théorème, que $\chi_A(A) = 0_2$.
(ii) Discuter, selon les valeurs de a, b, c, d et du discriminant $(a-d)^2+4bc$ de χ_A, les formes possibles du polynôme minimal de A. (Plus précisément on distinguera les cas : $(a-d)^2+4bc$ est nul, est un carré dans k, n'est pas un carré dans k.)
(iii) Que peut-on dire en caractéristique 2 ?

Soit $n \geqslant 1$ un entier. Rappelons que l'application déterminant $\mathrm{Det} : \mathrm{M}(n, A) \longrightarrow A$ est définie pour tout anneau commutatif A et jouit de propriétés similaires à celle du déterminant sur un corps. En particulier, si $A \in \mathrm{M}(n,k)$ et si X est une indéterminée, la matrice $XI_n - A$ est à coefficients dans $k[X]$. Son déterminant – un polynôme – est appelé le *polynôme caractéristique* de A et est noté χ_A. Si $n = 0$, on posera $\chi_A(X) = 1$ par convention. On voit facilement que quelle que soit la matrice carrée A de taille n, χ_A est un polynôme unitaire de degré n.

Il est remarquable que tout polynôme unitaire de $k[X]$ est le polynôme caractéristique d'au moins une matrice. Si P est donné par $P(X) = X^m + a_{m-1}X^{m-1} + \cdots + a_1X + a_0$, une telle matrice est la *matrice compagnon* (ou *partenaire*) de P, donnée par

$$\kappa(P) = \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 & -a_0 \\ 1 & 0 & 0 & \cdots & 0 & -a_1 \\ 0 & 1 & 0 & \cdots & 0 & -a_2 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & -a_{m-2} \\ 0 & 0 & 0 & \cdots & 1 & -a_{m-1} \end{pmatrix} .$$

Exercice **15.** Prouver que le polynôme caractéristique de $\kappa(P)$ est $P(X)$!

3.3. Proposition. (i) *Soient $P \in \mathrm{GL}(n,k)$ et $A \in \mathrm{M}(n,k)$. Alors A et PAP^{-1} ont même polynôme caractéristique. Il en est de même de A et ${}^tA$.*
(ii) *Soient n, $p \in \mathrm{M}(n,k)$, $A \in \mathrm{M}(n,p,k)$ et $B \in \mathrm{M}(p,n,k)$. Alors on a l'égalité $X^p\chi_{AB}(X) = X^n\chi_{BA}(X)$. En particulier, si A et B sont de même taille, AB et BA ont même polynôme caractéristique.*

Démonstration. Le point (i) découle de propriétés standard du déterminant. Pour prouver (ii), on utilise l'astuce suivante. On a l'égalité de polynômes :

$$\mathrm{Det}\begin{pmatrix} XI_n - AB & A \\ 0 & XI_p \end{pmatrix} = X^p\chi_{AB}(X) \ , \ \mathrm{Det}\begin{pmatrix} XI_n & A \\ 0 & XI_p - BA \end{pmatrix} = X^n\chi_{BA}(X) \ .$$

D'un autre côté, un calcul par blocs montre que

$$\begin{pmatrix} XI_n - AB & A \\ 0 & XI_p \end{pmatrix}\begin{pmatrix} I_n & 0 \\ B & I_p \end{pmatrix} = \begin{pmatrix} XI_n & A \\ XB & XI_p \end{pmatrix}$$

et

$$\begin{pmatrix} I_n & 0 \\ B & I_p \end{pmatrix}\begin{pmatrix} XI_n & A \\ 0 & XI_p - BA \end{pmatrix} = \begin{pmatrix} XI_n & A \\ XB & XI_p \end{pmatrix}$$

On conclut en prenant le déterminant et en remarquant que

$$\mathrm{Det}\begin{pmatrix} I_n & 0 \\ B & I_p \end{pmatrix} = 1 \ .$$

Remarque. Il est faux en général que $\mu_{AB} = \mu_{BA}$ pour deux matrice carrées de même taille A et B. Le vérifier par exemple avec

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \text{ et } B = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \ .$$

3.4. Proposition. *Soit $A \in \mathrm{M}(n,k)$ et $P \in k[X]$. Les assertions suivantes sont équivalentes :*

(i) $P(A) = 0$,
(ii) *Il existe une matrice B de $\mathrm{M}(n, k[X])$ telle que $(XI_n - A)B = P(X)I_n$.*

Démonstration. Le cas $P = 0$ étant évident, on suppose $P \neq 0$ et l'on pose $d = \deg(P)$.
(i) $\implies$ (ii) Supposons $P(A) = 0$ et écrivons $P(X) = a_0 + a_1X + \cdots + a_dX^d$. On écrit astucieusement :

$$P(X)I_n = P(X)I_n - P(A) = \sum_{i=1}^{p} a_i(X^iI_n - A^i)$$

$$= (XI_n - A)\sum_{i=1}^{p} a_i(X^{i-1}I_n + X^{i-2}A + \cdots + XA^{i-2} + A^{i-1}) \ .$$

D'où l'implication.

(ii) $\implies$ (i) Ecrivons $B = \sum_{i \geqslant 0} X^iB_i$, avec $B_i \in \mathrm{M}(n,k)$. Pour des raisons de degrés, on a $B_i = 0$ si $i \geqslant d$. L'égalité

$$(XI_n - A)(B_0 + XB_1 + \cdots + X^{d-1}B_{d-1}) = a_0I_n + a_1XI_n + \cdots + a_dX^dI_n$$

dans $\mathrm{M}(n, k[X])$ signifie que

$$\begin{array}{rcll} B_{d-1} & = & a_dI_n & (1) \\ B_{d-2} - AB_{d-1} & = & a_{d-1}I_n & (2) \\ \cdots & \cdots & \cdots & \\ B_0 - AB_1 & = & a_1I_n & (\mathrm{d}) \\ -AB_0 & = & a_0I_n & (\mathrm{d}{+}1) \end{array}$$

En multipliant la relation (k) à gauche par A^{d+1-k} et en sommant, on obtient bien $P(A) = 0$.

Remarque. Pour montrer (ii) $\implies$ (i), on peut être tenté de substituer A à X dans la relation $(XI_n - A)B = P(X)I_n$. Cependant substituer A à X dans la matrice B à coefficients dans $k[X]$ n'a pas de sens dans ce contexte.

3.5. Théorème. (Dit de Cayley-Hamilton.) *Soient $n \in \mathbb{N}^*$ et $A \in \mathrm{M}(n,k)$.*

(i) *Le polynôme caractéristique de A annule A.*
(ii) *Le polynôme minimal de A divise son polynôme caractéristique.*

Démonstration. Notons que (i) et (ii) sont équivalents par définition même du polynôme minimal. Rappelons que si M est une matrice carrée à coefficients dans un anneau commutatif et si $\mathrm{com}(M)$ est la matrice complémentaire (ou comatrice) de M (c'est-à-dire la transposée de la matrice des cofacteurs de M), alors

$$M\mathrm{com}(M) = \mathrm{com}(M)M = \mathrm{Det}(M)I_n \ .$$

En appliquant ceci à $M = XI_n - A$, on obtient

$$(XI_n - A)\mathrm{com}(XI_n - A) = \chi_A(X)I_n \ .$$

Il suffit à présent d'appliquer la proposition 3.4 avec $B = \mathrm{com}(XI_n - A)$.

Revenons à la situation d'un endomorphisme u d'une espace vectoriel E de dimension finie non nulle n. Puisque le polynôme caractéristique d'une matrice est invariant par conjugaison, il est licite de définir le polynôme caractéritique χ_u de u comme étant celui de sa matrice dans une base arbitrairement choisie. Le théorème de Cayley-Hamilton est vrai pour les endomorphismes : μ_u divise χ_u.

Supposons que F soit un sous-espace de E stable par u. Alors u induit un endomorphisme $\bar{u}$ de E/F défini par

$$\bar{u}(\bar{x}) = \overline{u(x)} \ , \ x \in E \ ,$$

où, rappelons le, pour $y \in E$, on note $\bar{y}$ la classe d'équivalence $y+F$ de y. On dit que $\bar{u}$ est l'endomorphisme *déduit de u par passage au quotient*. Noter l'intérêt de ce procédé : bien que F n'ait pas forcément de supplémentaire stable par u, u induit un endomorphisme de E/F, espace isomorphe à un supplémentaire quelconque de F.

3.6. Proposition. (i) *Soit F un sous-espace de E stable par u. Notons $v = u_{|F}$ et $w \in \mathcal{L}(E/F)$ l'endomorphisme déduit de u par passage au quotient. Alors :*

$$\chi_u = \chi_v \chi_w \ .$$

(ii) *Soient E_1, ..., E_p des sous-espaces de E stables par u dont E est la somme directe. Pour $i = 1$, ..., p, notons $u_i = u_{|E_i}$. Alors :*

$$\chi_u = \chi_{u_1} \cdots \chi_{u_p} \ .$$

Démonstration. (i) Soit $\boldsymbol{e} = (e_1, \ldots, e_n)$ une base de E complétée d'une base $\boldsymbol{f} = (e_1, \ldots, e_p)$ de F. Alors $\boldsymbol{g} = (\bar{e}_{p+1}, \ldots, \bar{e}_n)$ est une base de E/F (exercice !). Ainsi si $A = \mathrm{Mat}(v, \boldsymbol{f})$ et $B = \mathrm{Mat}(w, \boldsymbol{g})$, il est clair que

$$\mathrm{Mat}(u, \boldsymbol{e}) = \begin{pmatrix} A & C \\ 0 & B \end{pmatrix}$$

pour une certaine matrice $C \in \mathrm{M}(p, n-p, k)$. D'où l'assertion par la formule donnant le déterminant d'une matrice triangulaire supérieure par blocs.
(ii) La preuve est similaire en travaillant dans une base de E obtenue en réunissant des bases des E_i.

Rermarque. Cette proposition n'est plus vraie en général si l'on remplace polynôme caractéristique par polynôme minimal. Contre-exemple : prendre $u = \mathrm{id}_E$ et $1 \leqslant \dim F < \dim E$; on obtient $\mu_u(X) = X - 1 \neq \mu_v(X)\mu_w(X) = (X-1)^2$.

4. Réduction des endomorphismes : les idées fondamentales

Etant donné un endomorphisme u d'un espace vectoriel E (toujours supposé de dimension $n > 0$) , il est souvent utile de se ramener à une des situations suivantes :

– travailler avec une base dans laquelle la matrice de u ait une forme simple ou remarquable,
– ou décomposer l'espace E en une somme directe $E_1 \oplus \cdots \oplus E_r$ de sous-espaces stables par u telle que les restrictions $u_{|E_i}$ aient des propriétés remarquables.

Le pendant matriciel de ceci est la recherche, pour une matrice donnée, d'une matrice à forme simple qui lui soit semblable (travailler avec l'endomorphisme ayant cette matrice dans une base donnée).
Il y a diverses motivations à ce programme, dont :

– permettre des calculs explicites (fonctions de matrices telles que l'exponentielle, commutant de u dans $\mathfrak{L}(E)$, ...),
– pouvoir démontrer des propriétés de u.

Les outils de bases pour trouver des sous-espaces stables par u, ou pour décomposer E en une somme directe de sous-espaces stables, sont, outre l'emploi systématique des propriété de l'anneau $k[X]$, l'usage des deux résultats suivants.

4.1. Lemme *Soient v un endomorphisme de E qui commute avec u. Alors pour tout polynôme $P \in k[X]$, les sous-espaces* Ker $P(v)$ *et* Im $P(v)$ *sont stables par u.*

Démonstration. Cela vient du fait que $P(v)$ et u commutent. Soit $x \in \mathrm{Ker}\ P(v)$, alors $P(v)(u(x)) = u(P(v)(x)) = 0$, donc $u(x) \in \mathrm{Ker}\ P(v)$. De même si $y \in \mathrm{Im}\ P(v)$, on a $y = P(v)(x)$ pour un $x \in E$, d'où $u(y) = u(P(v)(x)) = P(v)(u(x)) \in \mathrm{Im}\ P(v)$.

En particulier, en faisant $v = u$, on a que pour tout polynôme P, Ker $P(u)$ et Im $P(u)$ sont des sous-espaces de E stables par u.

4.2. Théorème (**Théorème de décomposition des noyaux**). *Soient P_1, ..., P_r des polynômes premiers entre deux à deux et posons $P = P_1 \cdots P_r$. Alors si $P(u) = 0$, on a*

$$E = \mathrm{Ker}\ P_1(u) \oplus \cdots \oplus \mathrm{Ker}\ P_n(u)\ ,$$

et les projecteurs associés à cette décomposition de E sont des polynômes en u.

Commençons par établir le lemme suivant.

4.3. Lemme. *Soient un entier $r \geqslant 2$ et P_1, ..., P_r des polynômes de $k[X]$ premiers entre eux deux à deux. Pour $i = 1$, ..., r, posons $Q_i = P_1 \cdots P_{i-1} P_{i+1} \cdots P_r$. Alors les polynômes Q_1, ..., Q_r sont premiers entre eux dans leur ensemble.*

Démonstration. Supposons qu'il existe un polynôme irréductible P divisant tous les Q_i. Alors P divise l'un des P_i disons P_{i_o}. Puisque P divise Q_{i_o}, il doit diviser l'un des P_j pour $j \neq i_o$, ce qui contredit l'hypothèse que les polynômes P_i sont premiers entre eux deux à deux.

Démonstration du théorème 4.2. Posons $Q_i = P_1 \cdots P_{i-1} P_{i+1} \cdots P_n$, $1 \leqslant i \leqslant n$. D'après le lemme précédent et le lemme de Bezout, il existe des polynômes R_1, ..., R_n, tels que $R_1 Q_1 + \cdots R_n + Q_n = 1$.

Posons $T_i = R_i Q_i$, $1 \leqslant i \leqslant n$. On a pour tout $x \in E$, $x = T_1(u)(x) + \cdots + T_n(u)(x)$. Par suite, en posant $E_i = \operatorname{Im} T_i(u)$, on a $E = E_1 + \cdots + E_n$.

Or, si $i \neq j$, $Q_i Q_j$ est un multiple de P, et donc $T_i(u) \circ T_j(u) = 0$. On en déduit que les $T_i(u)$ sont des projecteurs orthogonaux et que E est la somme directe des E_i. On aura fini en prouvant que $E_i = \operatorname{Ker} P_i(u)$.

Puisque $P_i T_i = R_i P$ on a $P_i(u) \circ T_i(u) = 0$. Donc $E_i = \operatorname{Im} T_i(u) \subset \operatorname{Ker} P_i(u)$. Mais si $i \neq j$, P_i divise Q_j, de sorte que si $x \in \operatorname{Ker} P_i$, on a

$$x = T_i(u)(x) + \sum_{j \neq i} R_j(u) \circ Q_j(u)(x) = T_i(u)(x) \ .$$

Donc $\ker P_i(u) \subset E_i$.

Noter que le théorème 4.2 donne une décomposition de E en somme directe de sous-espaces stables par u.

Soient $u \in \mathcal{L}(E)$ et $x \in E$. Posons :

$$I_u^x = \{P \in k[X] \ ; \ P(u)(x) = 0\} \ , \ E_u^x = \{P(u)(x) \ ; \ P \in k[X]\} \ .$$

Il est facile de voir que I_u^x est un idéal non nul de $k[X]$. Il contient l'idéal $(\mu_u(X))$. On désigne par μ_u^x l'unique polynôme unitaire tel que $I_u^x = (\mu_u^x(X))$. Observons que E_u^x est le sous-espace de E engendré par les $u^k(x)$, $k \in \mathbb{N}$. Il est clairement stable par u.

4.4. Lemme. *Soient $u \in \mathcal{L}(E)$ et x, $y \in E$.*
(i) *On a $\deg(\mu_u^x) = 0$ si et seulement si $x = 0$.*
(ii) *On a $\dim E_u^x = \deg(\mu_u^x)$.*
(iii) *Si μ_u^x et μ_u^y sont premiers entre eux, alors $\mu_u^{x+y} = \mu_u^x \mu_u^y$.*

Démonstration. (i) On a clairement $\mu_u^0 = 1$ et, si $\mu_u^x = 1$, alors $1.x = \mathrm{id}_E(x) = x = 0$.
(ii) On peut supposer ici que $x \neq 0$. Soit $n = \deg(\mu_u^x) > 0$. Puisque chaque $P \in k[X]$ s'écrit $Q\mu_u^x + R$ avec $\deg(R) < n$, on voit que l'espace vectoriel E_u^x est engendré par x, $u(x)$, ..., $u^{n-1}(x)$. Soient λ_0, ..., λ_{n-1} des scalaires tels que

$$\lambda_0 x + \lambda_1 u(x) + \cdots + \lambda_{n-1} u^{n-1}(x) = 0 \ .$$

Alors si $Q = \lambda_0 + \lambda_1 X + \cdots + \lambda_n X^{n-1}$, on a $Q(u)(x) = 0$ et donc μ_u^x divise Q. On en conclut que $Q = 0$ puisque $\deg(Q) < n$. Ainsi $(x, u(x), ..., u^{n-1})$ est une base de E_u^x et l'égalité du lemme en découle.
(iii) Si $P = \mu_u^x \mu_u^y$ est le ppcm unitaire de μ_u^x et μ_u^y, on a $P(u)(x+y) = 0$. Donc μ_u^{x+y} divise $\mu_u^x \mu_u^y$. De même, des égalités $x = x + y - y$ et $\mu_u^y = \mu_u^{-y}$, on tire que μ_u^x divise $\mu_u^{x+y} \mu_u^y$, donc que μ_u^x divise μ_u^{x+y}, puisque μ_u^x et μ_u^y sont premiers entre eux. De même μ_u^y divise μ_u^{x+y}. On en déduit que P divise μ_u^{x+y}, d'où l'égalité du lemme.

4.5. Théorème. *Soit $u \in \mathcal{L}(E)$. Il existe un vecteur x de E tel que $\mu_u = \mu_u^x$.*

Démonstration. Ecrivons $\mu_u = P_1^{n_1} \cdots P_r^{n_r}$, où les polynômes P_i sont premiers entre eux et les $n_i > 0$. Pour $i = 1, ..., r$, posons $E_i = \text{Ker } P_i(u)$. Par le théorème de décomposition des noyaux, on a la décomposition de E en sous-espaces stables par u :

$$E = \sum_{i=1,...,r} E_i \ .$$

Si $x_i \in E_i$, on a $P_i^{n_i}(u)(x_i) = 0$, de sorte que $\mu_u^{x_i}$ divise $P_i^{n_i}$.

Supposons qu'il existe un indice k tel que $\mu_u^{x_k} \neq P_k^{n_k}$, pour tout $x_k \in E_k$. Alors $P_k^{n_k - 1}(u)(x_k) = 0$, pour tout $x_k \in E_k$. Posons

$$Q = P_1^{n_1} \cdots P_{k-1}^{n_{k-1}} P_k^{n_k - 1} P_{k+1}^{n_{k+1}} \cdots P_r^{n_r} \ .$$

Alors $Q(u)(x) = 0$, pour tout $x \in E$ (décomposer x en $x_1 + \cdots + x_r$, $x_i \in E_i$). Comme $\deg(Q) < \deg(\mu_u)$, c'est absurde.

On a donc prouvé que, pour $1 \leqslant i \leqslant r$, il existe $x_i \in E_i$ tel que $\mu_u^{x_i} = P_i^{n_i}$. En posant $x = x_1 + \cdots + x_r$, une généralisation par récurrence du lemme précédent montre que $\mu_u^x = \mu_u$.

Le résultat suivant réduit l'étude d'un endomorphisme au cas où le polynôme minimal est puissance d'un polynôme irréductible.

4.6. Théorème. *Soit $u \in \mathcal{L}(E)$ de polynôme minimal $\mu_u = P_1^{n_1} \cdots P_r^{n_r}$, où les P_i sont unitaires irréductibles deux à deux distincts et les n_i strictement positifs. Posons $E_i = \text{Ker } P_i^{n_i}(u)$. Alors :*
(i) *Pour $i = 1, ..., r$, E_i est stable par u et E est la somme directe des E_i. Posons alors $u_i = u_{|E_i}$.*
(ii) *Pour $i = 1, ..., r$, le polynôme minimal de u_i est $P_i^{n_i}$.*

Démonstration. Le premier point résulte du théorème de décomposition des noyaux. Le second est une conséquence de la preuve du théorème précédent.

Terminons cette section par la notion d'endomorphisme *semisimple*, notion plus souple que celle d'endomorphisme *diagonalisable*, qui sera étudiée plus loin dans ces notes.

4.7. Définition. (i) *On dit qu'un polynôme $P \in k[X]$ est* sans facteur multiple *s'il s'écrit $P = P_1 \cdots P_r$, où les P_i sont irréductibles deux à deux non associés (c'est-à-dire que deux d'entre eux ne diffèrent pas d'un facteur constant).*

(ii) *On dit que* $u \in \mathcal{L}(E)$ *est* semisimple *si tout sous-espace de* E *stable par* u *possède un supplémentaire dans* E *stable par* u.

Soit $u \in \mathcal{L}(E)$ de polynôme minimal $P_1^{n_1} \cdots P_r^{n_r}$, où les P_i sont unitaires irréductibles deux à deux distincts et les $n_i > 0$.

4.8. Lemme. *Soit* F *un sous-espace de* E *stable par* u. *Alors :*

$$F = (F \cap E_1) \oplus \cdots \oplus (F \cap E_r) \ .$$

Démonstration. Les projecteurs p_1, ..., p_r associés à la décomposition $E = E_1 \oplus \cdots E_r$ sont des polynômes en u. Ils laissent donc F stable. Ainsi si $x \in F$, on a

$$x = p_1(x) + \cdots p_r(x) \text{ avec } p_i(x) \in F \cap E_i \ .$$

Ceci prouve l'inclusion $F \subset (F \cap E_1) \oplus \cdots \oplus (F \cap E_r)$, l'autre inclusion étant évidente.

4.9. Lemme. *Pour* $i = 1, ..., r$, *notons* $u_i = u_{|E_i}$. *Les conditions suivantes sont équivalentes :*

(i) u *est semisimple ;*
(ii) *les* v_i, $i = 1, ..., r$, *sont semisimples.*

Démonstration. (i) $\Longrightarrow$ (ii) Fixons un $i \in \{1, ..., r\}$. Soit F_i un sous-espace de E_i stable par u_i. Cela signifie qu'il est stable par u. Par hypothèse, il existe un supplémentaire G de F_i dans E stable par u. Alors $G_i = G \cap E_i$ est un supplémentaire de F_i dans E_i et il est stable par u comme intersection de deux sous-espaces de E stables par u.

(ii) $\Longrightarrow$ (i) Soit F un sous-espace de E stable par u. Par le lemme précédent, on a $F = F_1 \oplus \cdots \oplus F_r$, où $F_i = F \cap E_i$, $i = 1, ..., r$. Chaque F_i est stable par u_i et par hypothèse, il existe pour chaque i un supplémentaire G_i de F_i dans E_i stable par u_i (donc par u). Il est alors immédiat que $G = G_1 \oplus \cdots \oplus G_r$ est un supplémentaire de F stable par u.

4.10. Théorème. *Si* $u \in \mathcal{L}(E)$, *les conditions suivantes sont équivalentes :*

(i) u *est semisimple ;*
(ii) *le polynôme minimal de* u *est sans facteur multiple.*

Démonstration. D'après le lemme 4.9, on peut se ramener au cas où $\mu_u = P^m$, P irréductible unitaire et $m > 0$. Il s'agit alors de montrer que u est semisimple si et seulement si $m = 1$.

Supposons $m = 1$. Alors pour tout $x \in E\backslash\{0\}$, on a $1 \neq \mu_u^x | P$, et donc $\mu_u^x = P$.

Remarquons que pour tout $x \in E\backslash\{0\}$ et pour tout $y \in E_u^x\backslash\{0\}$, on a $E_u^y = E_u^x$. En effet E_u^x étant stable par u, on a $E_u^y \subset E_u^x$. D'un autre côté, on a l'égalité de dimensions : $\dim(E_u^x) = \dim(E_u^y) = \deg(P)$.

Soit alors F un sous-espace de E stable par u. Si $F = E$, $\{0\}$ est un supplémentaire de F u-stable. Sinon fixons $x \in E\backslash F$. Alors $E_u^x \cap F = \{0\}$. En effet, supposons par l'absurde qu'il existe $y \in E_u^x \cap F\backslash\{0\}$. Alors $E_u^y = E_u^x$ et, puisque F est u-stable, $E_u^y \subset F$, ce qui est absurde. On peut donc construire par récurrence des vecteurs x_1, ..., x_p de E tels que

$$E = F \oplus E_u^{x_1} \oplus \cdots \oplus E_u^{x_p} \ .$$

Le sous-espace $G = E_u^{x_1} \oplus \cdots \oplus E_u^{x_p}$ est alors un supplémentaire de F stable par u.

Supposons au contraire que $n > 1$ et posons $K = \text{Ker } P(u)$. Par définition de μ_u on a $K \neq E$. De plus comme $P^m(u) = P(u) \circ \cdots \circ P(u) = 0$ (m facteurs), on a que $P(u)$ est non inversible, c'est-à-dire $K \neq \{0\}$. Supposons par l'absurde que K possède un supplémentaire u-stable G et soit $x \in G$, $x \neq 0$. Il existe un entier $q \leqslant m$ tel que $P^q(u)(x) = 0$. De plus, on a $q \geqslant 2$, car $x \notin K$. En choisissant q minimal, on a que $y := P^{q-1}(u)(x)$ doit être un élément non nul de G. Or $P(y) = 0$ et donc $y \in K$, ce qui est une contradiction.

Exercice **16.** Montrer qu'un endomorphisme nilpotent est semisimple si et seulement si il est nul.

Exercice **17.** Supposons $k = \mathbb{R}$, $E = k^2$ et $u \in \mathcal{L}(E)$ de matrice $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ dans la base canonique. Donner une condition nécessaire et suffisante portant sur a, b, c et d pour que u soit semisimple.

Exercice **18.** Supposons que $k = \mathbb{R}$ et soit $u \in \mathcal{L}(E)$. Montrer que si le polynôme caractéristique de u est scindé à racines simples sur $\mathbb{C}$, alors u est semisimple. Donner une CNS portant sur $\mu_u(X) \in \mathbb{C}[X]$ pour que u soit semisimple.

Exercice **19.** Supposons $k = \mathbb{R}$ et E euclidien.

(a) Montrer que tout endomorphisme orthogonal de E est semisimple. (On pourra, au choix, utiliser le théorème de réduction des endomorphismes orthogonaux, ou bien rechercher un supplémentaire stable d'un sous-espace stable en utilisant l'orthogonalité.)
(b) De façon similaire, traiter le cas d'un endomorphisme autoadjoint.
(c) Un endomorphisme normal est-il semisimple ?

4.11. Théorème. *Soit E un espace vectoriel de dimension finie $n > 0$. Soit $u \in \mathcal{L}(E)$.*
(i) *Le polynôme χ_u divise le polynôme $(\mu_u)^n$ dans $k[X]$.*
(ii) *Les polynômes χ_u est μ_u ont mêmes facteurs irréductibles dans $k|X]$, donc mêmes racines dans k.*
(iii) *χ_u est scindé sur k si et seulement si μ_u l'est.*

Démonstration. Les points (i) et (iii) découlent clairement de (ii), que nous démontrons. Par le théorème de décomposition des noyaux et puisque le polynôme caractéristique est multiplicatif dans les somme directes stables, il suffit de traiter le cas où $\mu_u = P^q$, où P est unitaire irréductible et $q \in \mathbb{N}^*$. Le cas $n = 1$ étant clair, nous raisonnons par récurrence sur n.

Supposons d'abord qu'il existe un sous-espace F de E stable par u tel que $1 \leqslant \dim(F) \leqslant \dim(E)$. Soit $v = u_{|F}$ et $w \in \mathcal{L}(E/F)$ l'endomorphisme obtenu par passage au quotient. Les polynômes μ_v et μ_w sont des puissances de P (car v et w sont annulés par μ_u). On a alors le résultat par l'hypothèse de récurrence grâce à (3.6.i).

Si l'hypothèse précédente n'est pas vérifiée, on a $E_u^x = E$, pour tout $x \in E\backslash\{0\}$. On a alors $\dim(E) = \deg(\mu_u)$, et puisque μ divise χ_u, il vient $\chi_u = \mu_u$

Corollaire 4.12. *Une matrice $A \in M(n, k)$ est nilpotente si et seulement si $A^n = 0$.*

Démonstration. En effet A est nilpotent si et seulement si le seul facteur irréductible de μ_A est X.

5. Diagonalisation et trigonalisation

Commençons par fixer des notations et de la terminologie. On fixe un espace vectoriel non nul E, de dimension n, et un endomorphisme $u \in \mathcal{L}(E)$. On appelle *spectre* de u, et on note $\mathrm{spec}(u)$, l'ensemble des racines (dans k) du polynôme caractéristique χ_u. Si $\lambda \in \mathrm{spec}(u)$ est une racine de χ_u, on note m_λ l'ordre de λ comme zéro de $\chi_u(X)$. Le polynôme caractéristique de u peut donc s'écrire

$$\chi_u(X) = \prod_{\lambda \in \mathrm{spec}(u)} (X-\lambda)^{m_\lambda} Q(X)$$

où $Q(X) \in k[X]$ est un polynôme sans racine dans k.

Rappelons que les polynômes minimal et caractéristique ont mêmes facteurs irréductibles, de sorte que $\mathrm{spec}(u)$ est aussi l'ensemble des racines de μ_u dans k.

Une *valeur propre* de u est un salaire λ tel que $u - \lambda \mathrm{id}_E$ soit non injectif. Si λ est valeur propre de u, le noyau $\mathrm{Ker}\,(u-\lambda\mathrm{id}_E)$ s'appelle le *sous-espace propre* associé à λ et se note $E_\lambda(u)$. Ses éléments non nuls sont appelés les vecteurs propres relatifs à λ. Un *vecteur propre* de u est un vecteur propre relatif à une valeur propre de u.

5.1. Théorème. *Soit $\lambda \in k$. Les conditions suivantes sont équivalentes :*

(i) *λ est valeur propre de u ;*
(ii) *$\lambda \in \mathrm{spec}(u)$;*
(iii) *λ est racine de μ_u.*

Démonstration. L'équivalence entre (ii) et (iii) a déjà été vue. L'équivalence entre (i) et (ii) découle du fait que l'endomorphisme $u - \lambda\mathrm{id}_E$ est non-injectif si et seulement si son déterminant est nul.

Remarque. On prend aussi comme définition première du spectre de u l'ensemble des valeurs propres. Ce problème de priorité est sans conséquence grâce aux équivalences du théorème.

On utilisera un vocabulaire similaire en remplaçant u par une matrice carrée A de $\mathrm{M}(n,k)$: on associe à A l'endomorphisme u_A de k^n qui admet A comme matrice dans la base canonique ; le spectre, les valeurs propres, les vecteurs propres de A sont ceux de u_A.

5.2. Corollaire. (i) *On a* $\mathrm{card}(\mathrm{spec}(u)) \leqslant \dim(E)$. *De plus si k est algébriquement clos,* $\mathrm{spec}(u)$ *est non vide.*
(ii) *Pour tout λ racine de χ_u, on a* $\dim E_\lambda(u) \leqslant m_\lambda$.

Démonstration. (i) C'est clair puisque χ_u est de degré $\dim(E)$.

(ii) Le sous-espace $F = E_\lambda(u)$ est stable par u de sorte que u induit un endomorphisme $\bar{u}$ sur le quotient E/F. Le polynôme caractéristique de u est alors donné par

$$\chi_u(X) = \chi_{u|F}\chi_{\bar{u}} \ .$$

Par construction, la restriction $u_{|F}$ est une homothétie de rapport λ, de sorte que $\chi_{u|F} = (X-\lambda)^{\dim(E_\lambda(u))}$. L'inégalité en découle.

Soit $\lambda \in \mathrm{spec}(u)$. Par définition le *sous-espace caractéristique* attaché à λ est $\mathrm{Ker}\,(u - \lambda\mathrm{id}_E)^{m_\lambda}$. On le note $F_\lambda(u)$. Notons que c'est un sous-espace de E stable par u.

5.3. Proposition. (i) *Les sous-espaces propres E_λ, $\lambda \in \mathrm{spec}(u)$, sont en somme directe.*
(ii) *Les sous-espaces caractéristiques F_λ, $\lambda \in \mathrm{spec}(u)$, sont en somme directe.*
(iii) *Si k est algébriquement clos, on a la décomposition*

$$E = \bigoplus_{\lambda \in \mathrm{spec}(u)} F_\lambda(u) \ .$$

Démonstration. Les trois points découlent du théorème de décomposition des noyaux. Prendre la famille de polynômes $P_\lambda(X) = X - \lambda$, $\lambda \in \mathrm{spec}(u)$, pour le point (i), et la famille $P_\lambda(X) = (X-\lambda)^{m_\lambda}$, $\lambda \in \mathrm{spec}(u)$, pour les points (ii) et (iii).

On dit que l'endomorphisme u est *diagonalisable* s'il existe une base de E dans laquelle sa matrice est diagonale. De même une matrice A est dite diagonalisable, si l'endomorphisme u_A associé est diagonalisable. Ceci est équivalent à l'existence d'un matrice inversible P telle que $P^{-1}AP$ est diagonale. Diagonaliser A signifie exhiber une telle matrice P. De même diagonaliser un endomorphisme diagonalisable u signifie produire une base où sa matrice est diagonale.

Le résultat suivant est immédiat et sa démonstration est laissée en exercice.

5.4. Proposition. *Un endomorphisme u est diagonalisable si et seulement si E est somme directe des sous-espaces propres $E_\lambda(u)$, $\lambda \in \mathrm{spec}(u)$.*

Une homothétie $u = t\mathrm{id}_E$ est l'archétype d'un endomorphisme diagonalisable. On a ici $\mathrm{spec}(u) = \{t\}$ et $E = E_t(u)$.

Un endomorphisme nilpotent non nul n'est jamais diagonalisable. En effet supposons $u \in \mathcal{L}(E)$ nilpotent et diagonalisable. Alors par définition d'un endomorphisme nilpotent, le polynôme minimal μ_u est de la forme x^r, $r \in \mathbb{N}$ et donc $\mathrm{spec}(u) = \{0\}$. Puisque u est diagonalisable, on a $E = E_0(u)$. Mais ceci signifie que u est l'homothétie de rapport 0. Donc $u = 0$.

Exercice **20.** (Patrice Tauvel) Exhiber une matrice de $\mathrm{M}(2,k)$ qui ne soit jamais diagonalisable quelle que soit la caractéristique de k.

Exercice **21.** Montrer qu'une matrice de $\mathrm{M}(2,\mathbb{C})$ n'est pas diagonalisable si et seulement si elle est semblable à une matrice de la forme $\begin{pmatrix} a & b \\ 0 & a \end{pmatrix}$, $b \neq 0$.

Exercice **21+1/2.** Soient A et B 2 matrices carrées à coefficients réels. Supposons qu'elles soient semblables vues comme matrices complexes. Montrer qu'elles sont semblables comme matrices réelles. *Indication* : soit $P = U + iV$ une matrice complexe conjuguant A en B (U et V sont deux matrices réelles). Chercher une matrice réelle conjuguant A en B sou la forme $P' = U + xV$, $x \in \mathbb{R}$.

5.5. Théorème. *Soit u un endomorphisme de E. Les conditions suivantes sont équivalentes :*

(i) *u est annulé par un polynôme P non constant scindé à racines simples ;*

(ii) *le polynôme minimal de u est scindé à racines simples ;*

(iii) *l'endomorphisme u est diagonalisable ;*

(iv) *tout sous-espace de E (stable ou non par u) possède un supplémentaire stable par u.*

Si toutes ces conditions équivalentes sont réalisées, et si F est un sous-espaces stable par u, alors F est la somme directe de ses intersections avec les sous-espaces propres de u.

Démonstration. (i) $\implies$ (ii) Découle du fait que μ_u divise P.
(ii) $\implies$ (iii) Découle du théorème de décomposition des noyaux appliqué à la famille des facteurs irréductibles (de degré 1) de μ_u.
(iii) $\implies$ (iv) Soit F un sous-espace de E et $\mathcal{B} = (e_1, \ldots, e_n)$ une base de E formée de vecteurs propres de u. Complétons une base de F en une base de E en ajoutant les vecteurs e_i, $i \in I$, pour un certain sous-ensemble I de $\{1, ..., n\}$ (théorème de la base incomplète). Alors $\bigoplus_{i \in I} ke_i$ est un supplémentaire de F stable par u.
(iv) $\implies$ (iii) Puisque tout hyperplan de E admet un supplémentaire stable, u fixe une droite, donc admet un vecteur propre e_1 (prendre un générateur d'une telle droite). Supposons, par récurrence, avoir construit p vecteurs e_1, ..., e_p linéairement indépendants et propres pour u, où $1 \leqslant p \leqslant n-1$. Soit H un hyperplan contenant $\mathrm{Vect}(e_1, ..., e_p)$ et ke_{p+1} un supplémentaire de H u-stable. Alors e_{p+1} est un vecteur propre linéairement indépendants des e_i, $i < p+1$. On obtient donc ainsi une base de vecteurs propres.
(iii) $\implies$ (i). Si $\mathcal{B} = (e_i)_{i=1,\ldots,n}$ est une base de vecteurs propres de u et si $u(e_i) = \lambda_i e_i$, $i = 1$, ..., n, la matrice de u dans $\mathcal{B}$ est $A = \mathrm{diag}(\lambda_1, ..., \lambda_n)$. Un calcul immédiat montre que A est annulé par le polynôme $P(X)$ produit des $X - \lambda$, où λ parcourt l'*ensemble* $\{\lambda_1, \ldots, \lambda_n\}$ (sans répétition). Donc u est annulé par ce même polynôme P, qui est scindé à racines simples.

D'après le point (iv) du théorème, tout endomorphisme diagonalisable est semisimple. Les deux notions coïncident si k est algébriquement clos (en rédiger soigneusement la démonstration en exercice). Toutefois si k n'est pas algébriquement clos, un endomorphisme semisimple n'est pas forcément diagonalisable. Par exemple, prenons $k = \mathbb{R}$, $E = \mathbb{R}^2$, et pour u l'endomorphisme de matrice $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ dans la base canonique. Un calcul immédiat donne $\chi_u(X) = \mu_u(X) = X^2 + 1$. Ainsi u est semisimple, puisque $\mu(X)$ est irréductible, mais non diagonalisable car $\mathrm{spec}(u) = \emptyset$.

5.6 Corollaire. (i) (*Critère simple de diagonalisabilité*) *Si $u \in \mathcal{L}(E)$ possède n valeurs propres distinctes, il est automatiquement diagonalisable. La réciproque est fausse dès que $n > 1$.*

(ii) *Soient $u \in \mathcal{L}(E)$ et F un sous-espace de E stable par u. Notons $\bar{u}$ l'endomorphisme de E/F obtenu par passage au quotient. Alors si u est diagonalisable, $u_{|F}$ et $\bar{u}$ le sont aussi.*
(iii) *Soit I un ensemble fini ou infini et $(u_i)_{i\in I}$ une famille d'endomorphismes de E diagonalisables et commutant deux à deux. Alors les u_i possèdent une base commune de diagonalisation.*

Démonstration. (i) Découle du théorème précédent par le fait que χ_u est scindé à racines simples (on a d'ailleurs $\mu_u = \chi_u$). Si $n > 1$, l'endomorphisme nul est diagonalisable, pourtant son spectre a un seul élément.

(ii) Il est facile de vérifier que μ_u, qui est scindé à racines simples, annule $u_{|F}$ et $\bar{u}$. Le résultat en découle par le théorème précédent.

(iii) Démontrons le résultat par récurrence sur n. Si $n = 1$, le résultat est immédiat car tout endomorphisme est alors diagonalisable. Supposons $n > 1$. Si tous les endomorphismes u_i sont des homothéties le résultat est évident. Supposons donc que u_{i_o} ne soit par une homothétie pour un certain $i_o \in I$. Les sous-espaces $E_\lambda(u_{i_o})$, $\lambda \in \mathrm{spec}(u_{i_o})$ sont stables par les u_i, $i \in I$, et sont de dimensions strictement inférieures à n. Soit $\lambda \in \mathrm{spec}(u_{i_o})$. En appliquant l'hypothèse de récurrence à la famille obtenue en restreignant les u_i à $E_\lambda(u_{i_o})$, on obtient une base $\mathcal{B}_\lambda$ de $E_\lambda(u_{i_o})$ formées de vecteurs propres pour les u_i, $i \in I$. Une base $\mathcal{B}$ de E obtenue en réunissant les bases $\mathcal{B}_\lambda$, $\lambda \in \mathrm{spec}(u_{i_o})$, est alors une base de diagonalisation commune.

On dit qu'un endomorphisme u de E est *trigonalisable* s'il existe une base de E dans laquelle sa matrice est triangulaire supérieure. Une matrice A de $\mathrm{M}(n,k)$ est dite *trigonalisable* si l'endomorphisme correspondant de k^n est trigonalisable, c'est-à-dire si A est semblable à une matrice triangulaire supérieure.

On montrera en exercice qu'un endomorphisme est trigonalisable si et seulement si il existe une base dans laquelle sa matrice est triangulaire inférieure.

5.7. Théorème. *Soit u un endomorphisme de E. Les conditions suivantes sont équivalentes :*
(i) *L'endomorphisme u est trigonalisable.*
(ii) *Le polynôme χ_u est scindé sur k.*
(iii) *Le polynôme μ_u est scindé sur k.*

Démonstration. L'équivalence entre (ii) et (iii) vient du fait que χ_u et μ_u ont les mêmes facteurs irréductibles. Si u est trigonalisable, de matrice triangulaire supérieure A dans une certaine base, alors un calcul direct de $\mathrm{Det}(XI_n - A)$ prouve que $\chi_A = \chi_u$ est scindé.

Montrons par récurrence sur n que (ii) entraîne (i). Si $n = 1$ l'implication est triviale et nous supposons $n \geqslant 2$. Soit donc $u \in \mathcal{L}(E)$ à polynôme caractéristique scindé. Par hypothèse u possède au moins une valeur propre λ. Posons $v = u - \lambda \mathrm{id}_E$. Le sous-espace $\mathrm{Im}\ v$ est strict et est donc contenu dans un hyperplan H. On a alors $v(H) \subset \mathrm{Im}\ v \subset H$: H est stable par v. On en déduit que H est aussi stable par $u = v + \lambda \mathrm{id}_E$. Posons $w = u_{|H}$. Alors puisque χ_w divise χ_u, le polynôme χ_w est scindé. Par récurrence, il existe une base $\boldsymbol{h} = (e_1, ..., e_{n-1})$ de H telle que $\mathrm{Mat}(\boldsymbol{h}, w)$ est triangulaire supérieure. Si maintenant e_n est un vecteur quelconque de $E \backslash H$ et $\boldsymbol{e} = (e_1, ..., e_n)$, la matrice $\mathrm{Mat}(\boldsymbol{e}, u)$ est triangulaire supérieure.

Exercice **22.** Soit $u \in \mathcal{L}(E)$ et ${}^t u$ l'endomorphisme transposé de E^*. Montrer que u et ${}^t u$ ont même polynôme caractéristique. On suppose que u possède une valeur propre λ. En utilisant la dualité, montrer que u stabilise un hyperplan de E.

5.8. Théorème. (i) *Si k est algébriquement clos, tout endomorphisme de E est trigonalisable.*
(ii) *Soient F un sous-espace de E stable par $u \in \mathcal{L}(E)$, et $w \in \mathcal{L}(E/F)$ l'endomorphisme obtenu par passage au quotient. Alors u est trigonalisable si et seulement si $u_{|F}$ et w le sont.*

Démonstration. Le point (i) et clair et (ii) découle du fait que $\chi_u = \chi_{u|F}\chi_w$.

Supposons u trigonalisable, de valeurs propres λ_1, ..., λ_n, comptées avec leurs multiplicités. Soit A la matrice de u dans une base donnée. En déterminant directement le coefficient de X^{n-1} et le terme constant dans le déterminant $\mathrm{Det}(XI_n - A)$, on trouve :

$$\chi_u = X^n - \mathrm{Tr}(A)X^{n-1} + \cdots + (-1)^n \mathrm{Det}(A) \ .$$

En effectuant le même calcul dans une base de trigonalisation, il vient :

$$\mathrm{Tr}(u) = \lambda_1 + \cdots + \lambda_n \text{ et } \mathrm{Det}(u) = \lambda_1 \cdots \lambda_n \ .$$

Plus généralement, on observe que si u est trigonalisable dans une base $\mathcal{B}$ et $P \in k[X]$, alors la matrice de $P(u)$ dans $\mathcal{B}$ est triangulaire supérieure et ses coefficients diagonaux sont $P(\lambda_1)$, ..., $P(\lambda_n)$ (à l'ordre près). On en déduit le résultat suivant :

5.9. Proposition. *Soient $u \in \mathcal{L}(E)$ un endomorphisme trigonalisable et P un polynôme. Avec les notations précédentes, on a*

$$\chi_{P(u)} = (X - P(\lambda_1)) \cdots (X - P(\lambda_n)) \ .$$

En particulier $\mathrm{Tr}(P(u)) = P(\lambda_1) + \cdots + P(\lambda_n)$ *et* $\mathrm{Det}(P(u)) = P(\lambda_1) \cdots P(\lambda_n)$.

5.10. Théorème. *Soient $u \in \mathcal{L}(E)$ un endomorphisme trigonalisable. Alors :*
(i) *E est somme directe des sous-espaces caractéristiques $F_\lambda(u)$, où λ décrit* $\mathrm{spec}(u)$.
(ii) *Pour chaque valeur propre λ de u, la dimension de $F_\lambda(u)$ est la multiplicité m_λ de λ dans χ_u.*

Démonstration. (i) On a $\chi_u(X) = \prod_{\lambda \in \mathrm{spec}(u)} (X - \lambda)^{m_\lambda}$ et $F_\lambda(u) = \mathrm{Ker}\ (u - \lambda \mathrm{id}_E)^{m_\lambda}$, $\lambda \in \mathrm{spec}(u)$. L'assertion découle donc du théorème de décomposition des noyaux.
(ii) Pour chaque valeur propre λ, notons $u_\lambda = u_{|F_\lambda(u)}$. Par définition même de $F_\lambda(u)$, on a que u_λ a un spectre réduit à $\{\lambda\}$. Donc $\chi_{u_\lambda} = (X - \lambda)^{\dim(F_\lambda(u))}$. D'un autre côté, on a

$$\chi_u(X) = \prod_{\lambda \in \mathrm{spec}(u)} \chi_{u_\lambda}(X) \ .$$

On en déduit $\chi_{u_\lambda} = (X - \lambda)^{m_\lambda}$ et l'égalité désirée.

5.11. Théorème. (i) *Soient $u \in \mathcal{L}(E)$ un endomorphisme trigonalisable et F un sous-espace de E stable par u. Alors F est somme directe de ses intersections avec les sous-espaces caractéristiques de u.*
(ii) *Soient I un ensemble fini ou infini et $(u_i)_{i\in I}$ une famille d'endomorphismes trigonalisables de E commutant deux à deux. Alors les u_i $i \in I$, possèdent une base de triangulation commune.*

Démonstration. La démonstration du point (i) est similaire à celle de 5.6 (iii). Pour le point (ii) raisonnons par récurrence sur $n = \dim(E)$. Pour $n = 1$, le résultat est trivial. Il l'est aussi si tous les endomorphismes sont des homothéties. Supposons donc $n > 1$ et qu'il existe un endomorphisme u de la famille qui ne soit pas une homothétie. Soient F un sous-espace propre de u et $x \mapsto \bar{x}$ la projection canonique $E \longrightarrow E/F$. Le sous-espace F est strict et stable par tous les u_i. Notons w_i l'endomorphisme de E/F déduit de u_i par passage au quotient. Notons $p = \dim(F)$. Par hypothèse de récurrence :

– il existe une base $(e_1, ..., e_p)$ de F dans laquelle u_i a une matrice triangulaire supérieure pour tout i ;

– il existe une base $(\varepsilon_{p+1}, ..., \varepsilon_n)$ de E/F dans laquelle w_i a une matrice triangulaire supérieure pour tout i.

Pour $i = p + 1$, ..., n, soit e_i un vecteur de F tel que $\bar{e}_i = \varepsilon_i$. Nous laissons le soin au lecteur de montrer que $(e_1, ..., e_n)$ est une base de E dans laquelle u_i a une matrice triangulaire supérieure pour tout i.

5.12. Théorème. (**Décomposition de Dunford**) *Soit u un endomorphisme trigonalisable de E. Il existe un unique couple (d, ν) d'endomorphismes de E vérifiant les conditions suivantes :*
(i) $u = d + \nu$;
(ii) *d et ν commutent* ;
(iii) *d est diagonalisable et ν est nilpotent.*
De plus, on a $\chi_u = \chi_d$ et il existe des polynômes P et Q de $k[X]$ sans terme constant tels que $d = P(u)$ et $\nu = Q(u)$.

Démonstration. Notons $r = \mathrm{Card}(\mathrm{spec}(u))$, λ_1, ..., λ_r les valeurs propres distinctes de u et F_1, ..., F_r les sous-espaces caractéristiques. Rappelons que E et la somme directe des F_i. On peut donc définir un endomorphisme diagonalisable d de E en imposant $d_{|F_i} = \lambda_i \mathrm{id}_{F_i}$, $i = 1, ..., r$. Soit $\nu = u - d$. Par définition même des sous-espaces caractéristiques, pour chaque i, la restriction de ν à F_i est nilpotente d'indice divisant m_{λ_i}. Il s'ensuit que ν est nilpotent d'indice divisant le ppcm des m_{λ_i}.

Soit p_i le projecteur sur F_i relativement à la décomposition $E = F_1 \oplus \cdots \oplus F_r$. D'après le théorème de décomposition des noyaux, p_i est un polynôme en u. Or on a l'écriture $d = \lambda_1 p_1 + \cdots + \lambda_r p_r$, de sorte que d (et donc $v = u - \nu$) est un polynôme en u. Il s'ensuit que d et ν commutent.

Soit donc $P \in k[X]$ tel que $d = P(u)$, de sorte que $\nu = Q(v)$, où $Q(X) = X - P(X)$. Puisque u est trigonalisable, le spectre de ν est $\{Q(\lambda_1), ..., Q(\lambda_r)\}$. Il s'ensuit que Q s'annule sur $\mathrm{spec}(u)$. Si $0 \in \mathrm{spec}(u)$, on a alors $Q(0) = 0$ et Q (donc P) est sans terme constant. Si par contre u est inversible, le théorème de Cayley-Hamilton fournit un polynôme R sans terme constant tel que $R(u) = \mathrm{id}_E$. On peut alors remplacer P par PR (et $Q(X)$ par $X - P(X)R(X)$) pour être sûr que P (et donc Q) est sans terme constant.

Supposons qu'il existe un deuxième couple d'endomorphismes (D, N) tel que D soit diagonalisable, N nilpotent, $DN = ND$ et $u = D+N$. Comme d et ν sont des polynômes en u, ils commutent avec D et N. Il s'ensuit qu'on peut simultanément diagonaliser d et D et simultanément trigonaliser ν et N. En particulier $d - D$ est diagonalisable et $N - \nu$ est nilpotent (car ayant une matrice triangulaire supérieure stricte dans une base). Or $d - D = N - \nu$ et le seul endomorphisme à la fois diagonalisable et nilpotent est l'endomorphisme nul. Ainsi $d = D$ et $\nu = N$, ce qui prouve l'unicité dans le théorème.

Remarque. Il découle de la démonstration du théorème précédent que les endomorphismes d et n ont même spectre.

6. Endomorphismes nilpotents

A l'aide des théorèmes de Cayley-Hamilton et de décomposition des noyaux, nous avons vu que, pour chaque endomorphisme u de E, on a une décomposition de l'espace en somme directe de sous-espaces stables par u telle que la restriction de u à chaque sous-espace soit la somme d'une homothétie et d'un endomorphisme nilpotent. Les homothéties prenant une forme diagonale dans n'importe quelle base, nous sommes ramenés à trouver de jolies bases pour les endomorphismes nilpotents. C'est l'objet de cette section. Comme corollaire, nous obtiendrons la réduction de Jordan d'un endomorphisme ainsi que des propriétés des classes de conjugaison d'endomorphimes nilpotents.

Commençons par quelques compléments découlant du théorème de Cayley Hamilton. Celui-ci dit en particulier qu'un endomorphisme u est nilpotent si et seulement si $\chi_u(X) = X^n$, c'est-à-dire si χ_u est scindé et $\mathrm{spec}(u) = \{0\}$. Nous obtenons donc :

6.1. Théorème. *Soient E un espace vectoriel de dimension finie n et u un endomorphisme de E. Les conditions suivantes sont équivalentes :*

(i) *u est nilpotent.*
(ii) *Il existe une base dans laquelle la matrice de u est triangulaire supérieure stricte.*
(ii)' *Il existe une base dans laquelle la matrice de u est triangulaire inférieure stricte.*

Démonstration. L'équivalence entre (ii) et (ii)' vient de ce que si un endomorphisme a une matrice triangulaire supérieure dans la base $(e_1, ..., e_n)$, il a une matrice triangulaire inférieure dans la base $(e_n, e_{n-1}, ..., e_1)$ (*le vérifier*!).

6.2. Corollaire. *Si $u \in \mathcal{L}(E)$ est nilpotent, alors $\mathrm{Tr}(u^p) = 0$ pour tout $p \in \mathbb{N}^*$.*

Démonstration. Cela vient du fait que si une matrice A est triangulaire supérieure stricte, A^p l'est aussi pour tout $p > 0$.

6.3. Proposition. *Supposons le corps k de caractéristique nulle. Pour $u \in \mathcal{L}(E)$, les conditions suivantes sont équivalentes :*

(i) *u est nilpotent.*
(ii) *On a $\mathrm{Tr}(u^p) = 0$ pour tout $p \in \mathbb{N}^*$.*

Démonstration. Nous devons démontrer que (ii) entraîne (i). Raisonnons par récurrence sur n, le cas $n = 1$ étant clair. Ecrivons

$$\chi_u(X) = X^n + a_{n-1}X^{n-1} + \cdots + a_0 \ .$$

En prenant la trace dans l'égalité $\chi_u(u) = 0$, on obtient $a_0\mathrm{Tr}(\mathrm{id}_E) = na_0 = 0$. Puisque la caractéristique de k n'est pas nulle, on en déduit que $a_0 = (-1)^n\mathrm{Det}(u) = 0$. Ainsi u est non injectif. Soit alors $(e_1, ..., e_n)$ une base de E telle que e_1 soit dans le noyau de u. La matrice A de u dans cette base est de la forme

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

où $B \in \mathrm{M}(n-1, k)$ et C est une matrice ligne. Un calcul par blocs montre que $\mathrm{Tr}(B^p) = 0$ pour tout $p \in \mathbb{N}^*$. Par hypothèse de récurrence B est nilpotente et donc A l'est aussi.

L'étude fine des endomorphismes nilpotents est basée sur la suite des noyaux itérés. Fixons pour le moment un endomorphisme quelconque u de E. Pour $k \in \mathbb{N}$, on pose $N_k = \ker\, u^k$, $d_k = \dim(N_k)$, et pour $k \geqslant 0$, $n_k = d_k - d_{k-1}$.

6.4. Lemme. (i) *La suite $(N_k)_{k\in\mathbb{N}}$ est croissante pour la relation d'inclusion. La suite $(d_k)_{k\in\mathbb{N}}$ est donc croissante.*
(ii) *La suite $(n_k)_{k>0}$ est décroissante.*

Démonstration. (i) Si $x \in N_k$, alors $u^{k+1}(x) = u(u^k(x)) = 0$ et donc $x \in N_{k+1}$. D'où le résultat.
(ii) Soit $k \geqslant 0$. Si $x \in N_{k+1}$, alors $u^k(u(x)) = 0$, de sorte que $u(x) \in N_k$. Il s'ensuit que u induit une application bien définie et linéaire $\bar{u} : N_{k+2}/N_{k+1} \longrightarrow N_{k+1}/N_k$ donnée par $\bar{u}(x + N_{k+1}) = u(x) + N_k$. Cette application est injective. En effet soit $\bar{x} = x + N_{k+1} \in N_{k+2}/N_{k+1}$. Si $\bar{u}(\bar{x}) = 0$, on a $u(x) \in N_k$, d'où $x \in N_{k+1}$ et donc $\bar{x} = 0$. Ainsi $\dim(N_{k+2}/N_{k+1}) \leqslant \dim(N_{k+1}/N_k)$, autrement dit $d_{k+2} - d_{k+1} \leqslant d_{k+1} - d_k$, comme voulu.

6.5. Corollaire. *Avec les notations du lemme précédent, la suite $(d_k)_{k\in\mathbb{N}}$ est strictement croissante jusqu'à un certain rang à partir duquel elle devient stationnaire.*

Démonstration. La suite $(d_k)_{k\in\mathbb{N}}$ est bornée par $n = \dim(E)$. Elle est donc stationnaire à partir d'un certain rang. Il nous faut prouver que si k_o est le plus petit entier tel que $d_{k_o} = d_{k_o+1}$, alors $(d_k)_{k\geqslant k_o}$ est constante. Or nous avons vu que la suite $(d_{k+1} - d_k)_{k\geqslant 0}$ est décroissante et positive. Donc si elle est nulle pour $k = k_o$, elle l'est pour $k \geqslant k_o$.

Dorénavant, nous supposerons que l'*endomorphisme u est nilpotent*. On note $r > 0$ son indice de nilpotence : on a $u^r = 0$ et $u^{r-1} \neq 0$.

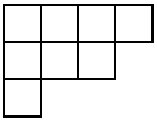
6.6. Corollaire. *Soit $u \in \mathcal{L}(E)$ un endomorphisme nilpotent.*
(i) *On a $d_0 = 0$ et $d_k = \dim(E)$ pour k grand. L'entier k_o à partir duquel la suite (d_k) stationne est l'indice de nilpotence de u.*
(ii) *Le r-uplet $(n_1, n_2, ..., n_r)$ est une* partition *de n au sens où :*

$$n_1 \geqslant n_2 \geqslant \cdots \geqslant n_r \text{ et } n_1 + n_2 + \cdots n_r = n \ .$$

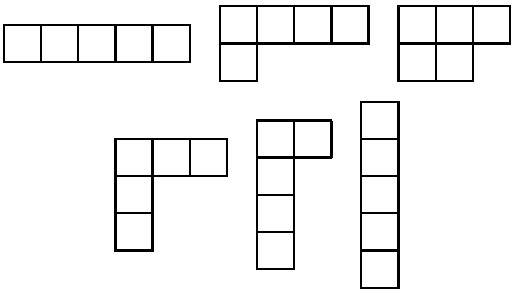
Démonstration. (i) On a évidemment $N_0 = \ker u^0 = \ker \mathrm{id}_E = \{0\}$ et donc $d_0 = 0$. Puisque $u^k = 0$, pour k grand, $d_k = n = \dim(E)$, pour k grand. La suite (d_k) stationne donc avec la valeur n. Puisque $\dim(N_k) = n$ si et seulement si $N_k = E$, on a que $k = k_o$ si et seulement si $N_{k_o} = E$ et $N_{k_o-1} \neq E$, c'est-à-dire si $k_o = r$.
(ii) On a

$$n_1 + \cdots + n_r = (d_1 - d_0) + (d_2 - d_1) + \cdots (d_r - d_{r-1}) = d_r - d_0 = n \ .$$

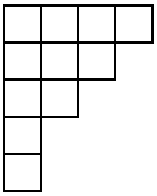
A chaque partition (décroissante par définition) $(n_1, n_2, ..., n_r)$ d'un entier n, on associe un tableau de Young. Il s'agit d'un tableau formé de cases que l'on dispose en colonnes : la première colonne comporte n_1 cases, la deuxième n_2 cases, etc. Voici par exemple le tableau de Young de la partition $(3, 2, 2, 1)$ de l'entier 8 :



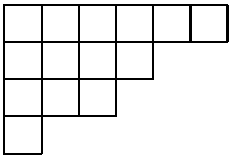
Voici les tableaux associés aux six partitions de l'entier 5 :



Noter, qu'étant donné un tableau de Young, la lecture de la suite des longueurs des lignes (en lisant de haut en bas) donne une autre partition de n. C'est aussi la partition associée au tableau *transposé* obtenu en mettant les lignes en colonnes. Par exemple à partir de la partition $(6, 4, 3, 1)$ de l'entier 14, on a le tableau de Young



dont le tableau transposé est



correspondant à la partition de 14 donnée par $(4, 3, 3, 2, 1, 1)$. On dit que la partition associée au tableau transposé est la *partition duale* de la partition initiale.

Nous verrons plus loin que la partition (ou de façon équivalent le tableau de Young) associé à l'endomorphisme nilpotent u détermine sa classe de conjugaison dans $\mathrm{GL}(E)$. En effet nous ferons apparaître par un autre point de vue le tableau de Young transposé.

Si $p \in \mathbb{N}^*$, on note $J_p = (a_{ij}) \in \mathrm{M}(p,k)$ la matrice définie par $a_{i,i+1} = 1$ pour $i = 1, ..., p-1$ et $a_{ij} = 0$ sinon. Ainsi $J_1 = 0$ et si $p \geqslant 2$, on a

$$J_p = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & 0 & \cdots & 0 & 0 \end{pmatrix} .$$

Un calcul facile montre que J_p est nilpotente d'indice p. En effet, si $k = 1, ..., p-1$, J_p^k se déduit de p en remontant la diagonale de 1 de $k-1$ lignes (*faites ce calcul !*). On dit que J_p est la matrice nilpotente de Jordan d'ordre p.

6.7. Théorème. *Soient E un espace vectoriel de dimension finie $n > 0$ et u un endomorphisme nilpotent d'indice p de E. Les conditions suivantes sont équivalentes :*
(i) $p = n$.
(ii) *Il existe une base de E dans laquelle la matrice de u est J_n.*
(iii) *Il existe $x \in E$ tel que $(x, u(x), ..., u^{n-1}(x))$ soit une base de E.*

Démonstration. (i) $\Longrightarrow$ (iii) Par hypothèse, il existe $x \in E$ tel que $u^{n-1}(x) \neq 0$ et $u^n(x) = 0$. Montrons que la famille $(x, u(x), ..., u^{n-1}(x))$ est libre. Soient a_0, a_1, ...,a_{n-1} des scalaires tels que

$$a_0 x + a_1 u(x) + \cdots + a_{n-1} u^{n-1}(x) = 0 .$$

En appliquant u^{n-1} à cette égalité, on obtient $a_0 u^{n-1}(x) = 0$, donc $a_0 = 0$. De proche en proche, on obtient que tous les scalaires sont nuls, comme annoncé.
(iii) $\Longrightarrow$ (ii) Par hypothèse $(u^{n-1}(x), ..., u(x), x)$ est aussi une base de E et un calcul immédiat montre que la matrice de u dans cette base est J_n.
(ii) $\Longrightarrow$ (i) Car J_n est nilpotent d'ordre n.

6.8. Théorème. *Soit E un espace vectoriel de dimension finie $n > 0$ et u un endomorphisme nilpotent d'ordre p de E. Il existe une partition $(m_1, m_2, ..., m_s)$ de n et une base de E dans laquelle la matrice de u est la matrice diagonale par blocs*

$$\mathrm{Diag}(J_{m_1}, J_{m_2}, ..., J_{m_s}) ,$$

et telle que l'entier m_1 est égal à l'indice de nilpotence p.

Démonstration. Le résultat est clair si $n = 1$. Supposons donc que $n \geqslant 2$ et raisonnons par récurrence sur n. Soit $v = {}^t u \in \mathcal{L}(E^*)$ l'endomorphisme transposé.

Puisque pour tout entier $k \geqslant 0$, $v^k = {}^t(u^k)$, v est nilpotent d'indice p. Il existe donc une forme linéaire $\varphi \in E^*$ telle que $v^{p-1}(\varphi) = \varphi \circ u^{p-1} \neq 0$. On a ainsi l'existence d'un $x \in E$ tel que $\varphi \circ u^{p-1}(x) \neq 0$. Soient $F = kx + ku(x) + \cdots + ku^{p-1}(x) \subset E$ et

$\Phi = k\varphi + kv(\varphi) + \cdots + kv^{p-1}(\varphi) \subset E^*$. Comme dans la preuve du théorème 6.7, on a que $\mathcal{B}_1 = (u^{p-1}(x), ..., u(x), x)$ et $(\varphi, ..., v^{p-1}(\varphi))$ sont des bases de F et Φ respectivement. Il est de plus clair que F et Φ sont stables par u et v respectivement. .

Soit G l'orthogonal de Φ dans E :

$$G = \{x \in E \ ; \ \psi(x) = 0 \text{ pour tout } \psi \in \Phi\} \ .$$

Il est stable par u. En effet si $x \in G$ et $\psi \in \Phi$, on a $\psi(u(x)) = v(\psi)(x) = 0$, car $v(\psi) \in G$. De plus la dimension de G est $n - \dim(\Phi) = n - p$. Montrons que $F \cap G = \{0\}$, et donc que $F \oplus G = \{0\}$. SI $y \in F \cap G$, il s'écrit $y = a_0 x + a_1 u(x) + \cdots + a_{p-1} u^{p-1}(x)$. Supposons y non nul et soit k le plus petit entier tel que $a_k \neq 0$.Puisque $y \in G$, on a $u^{p-1-k}(\varphi)(y) = 0$, c'est-à-dire $\varphi \circ u^{p-1-k}(y) = a_k \varphi(u^{p-1}(x)) = 0$, une contradiction.

On a donc une décomposition de E en une somme directe $F \oplus G$ de sous-espace stables par u et une base $\mathcal{B}_1$ de F telle que la matrice de $u_{|F}$ dans $\mathcal{B}_1$ soit J_p. L'endomorphisme $u_{|G}$ est nilpotent d'ordre inférieur ou égal à p. Donc par hypothèse de récurrence, il existe une base de G dans laquelle la matrice de $u_{|G}$ soit de la forme

$$\mathrm{Diag}(J_{m_2}, ..., J_{m_s}) \ ,$$

où $m_2 \leqslant p$ et $m_2 \geqslant m_3 \geqslant \cdots \geqslant m_s$. Le résultat en découle.

6.9. Lemme. *Soient $(m_1, m_2, ..., m_s)$ une partition de n et A la matrice diagonale par blocs de $\mathrm{M}(n, k)$ donnée par*

$$A = \mathrm{Diag}(J_{m_1}, J_{m_2}, ..., J_{m_s}) \ .$$

Soit u_A l'endomorphisme nilpotent de k^n associé à A. Posons $d_k = \dim(\ker u_A^k)$, $k \geqslant 0$ et $n_k = d_k - d_{k-1}$, $k \geqslant 1$.

(i) *L'indice de nilpotence de u_A est m_1 et, pour $k = 1, ..., m_1$, on a la formule :*

$$d_k = \min(k, m_1) + \min(k, m_2) + \cdots + \min(k, m_s) \ .$$

(ii) *La partition $(n_1, n_2, ..., n_{m_1})$ est associée au tableau transposé de celui de $(m_1, m_2, ..., m_s)$.*

Démonstration. (i) Une bonne figure vaut mieux qu'un long discours. On se convainc de cette formule en faisant un dessin de la matrice A et en notant que le noyau de l'endomorphisme naturellement associé à J_p^k a pour dimension k, si $k \leqslant p$, et dimension p si $k \geqslant p$; il a donc dimension $\min(k, p)$.
(ii) En faisant une figure, on note que d_k est la somme des longueurs des k premières lignes du tableau de Young associé à $(m_1, m_2, ..., m_s)$. Donc la longueur de la $k$ème ligne de ce tableau de Young est $d_k - d_{k-1}$.

Comme corollaire, on a le résultat suivant.

6.10. Théorème. (i) *Soit u un endomorphisme nilpotent de E. La partition de $n = \dim(E)$ associée à u par le théorème 6.8 ne dépend que de u et pas des divers choix dans la démonstration du théorème.*

(ii) *Deux endomorphismes nilpotents sont conjugués dans* $\mathrm{GL}(E)$ *si et seulement si les partitions associées sont égales. En particulier les classes de conjugaison d'endomorphismes nilpotents sont en nombre fini et en bijection avec les partitions de* n.
(iii) *Les classes de similitude de matrices nilpotentes de* $\mathrm{M}(n,k)$ *sont en nombre fini et en bijection avec les partitions de* n.

Pour $p \in \mathbb{N}^*$ est $\lambda \in k$, notons $J_p(\lambda)$ la matrice triangulaire supérieure $\lambda I_p + J_p$. On dit que $J_p(\lambda)$ est la matrice de Jordan d'ordre p associée à λ. Le polynôme caractéristique de cette matrice est $(X-\lambda)^p$.

Réciproquement, soit $u \in \mathcal{L}(E)$ un endomorphisme de polynôme caractéristique de la forme $(X-\lambda)^n$. D'après le théorème de Cayley-Hamilton, l'endomorphisme $v = u - \lambda \mathrm{id}_E$ est nilpotent et il existe donc une partition $(n_1, ..., n_r)$ de n et une base de E dans laquelle la matrice de v est $\mathrm{Diag}(J_{n_1}, ..., J_{n_r})$. Dans cette base la matrice de u est

$$\mathrm{Diag}(J_{n_1}(\lambda), ..., J_{n_r}(\lambda)) \ .$$

Soit à présent un endomorphisme diagonalisable u de E. Comme conséquence du fait que E est somme directe des sous-espaces caractéristiques de E, on obtient le résultat suivant.

6.11. Théorème. (**Réduction de Jordan**) *Soient* E *un espace vectoriel de dimension finie non nulle et* u *un endomorphisme trigonalisable de* E. *Il existe une base de* E, *des entiers* p_1, ..., p_s, *des scalaires* μ_1, μ_2, ..., μ_s, *tels que la matrice de* u *dans cette base soit*

$$\mathrm{Diag}(J_{p_1}(\mu_1), ..., J_{p_s}(\mu_s)) \ .$$

L'ensemble des μ_i *est égal à l'ensemble des valeurs propres de* u.

Exercice **23.** Soit $(n_1, ..., n_r)$ une partition de n, de partition duale $(m_1, ..., m_s)$. Montrer que pour $1 \leqslant i \leqslant s$, m_i est le cardinal de l'ensemble des entiers $j \in \{1, ..., r\}$ qui vérifient $n_j \geqslant i$. *Faire un dessin !*

Exercice **24.** Soit $u \in \mathcal{L}(E)$ un endomorphisme nilpotent de partition $(n_1, n_2, ..., n_r)$. Donner une formule donnant le nombre de blocs de Jordan de taille $d \times d$ en fonction de la partition duale, puis en fonction des n_i.

Exercice **25.** Soit $u \in \mathcal{L}(E)$ un endomorphisme nilpotent de partition $(n_1, n_2, ..., n_r)$. Soit $(m_1, ..., m_s)$ la partition duale. Le commutant $C(u)$ de u dans $\mathcal{L}(E)$ est par définition

$$C(u) = \{v \in \mathcal{L}(E) \ ; \ v \circ u = u \circ v\} \ .$$

Montrer que $C(u)$ est une sous-algèbre de $\mathcal{L}(u)$ et que sa dimension en tant qu'espace vectoriel est

$$n_1^2 + n_2^2 + \cdots + n_s^2 \ .$$

Exercice **26.** Soit u et v deux endomorphismes trigonalisables de E. Montrer l'équivalence entre les deux assertions suivantes :
(i) u et v sont conjugués sous $\mathrm{GL}(E)$.
(ii) quel que soit le scalaire $\lambda \in k$ et quel que soit l'entier $k > 0$, on a

$$\dim(\ker \ (u - \lambda \mathrm{id}_E)^k) = \dim(\ker \ (v - \lambda \mathrm{id}_E)^k) \ .$$

Bibliographie

Arnaudiès, J.M. et Fraysse, H., *Cours de Mathématiques T.1, Algèbre*, Dunod Université, 1987.

Mneimné, Rached, *Réduction des endomorphismes*, Calvage et Mounet, 2006.

Paugam, Annette, *Questions délicates en algèbre et géométrie*, Dunod, collection Sciences Sup, 2007

Tauvel, Patrice, *Algèbre*, Dunod, collection Sciences Sup, 2° édition, 2005.