Sorbonne Université

Licence de Mathématiques

Année 2019–2020

LU2MA123 - Algèbre linéaire et bilinéaire IIb

RÉDUCTION DES ENDOMORPHISMES

par Alexandru OANCEA

Remerciements :

à Vincent Humilière, Patrick Polo, Pierre-Antoine Guihéneuf, Marco Maculan pour m'avoir prêté leurs polycopiés et archives ;

à Adrien Deloro pour sa relecture tout en finesse.

Bibliographie :

X. Gourdon, *Les maths en tête : Algèbre*, Ellipses, 2009.

L. Koelblen, P. Polo, V. Humilière, *Algèbre et géométrie/Algèbre linéaire 2, espaces affines*, polycopiés 2009–2016, Université Pierre et Marie Curie.

https://webusers.imj-prg.fr/~patrick.polo/LM270/polyLM270-2013.pdf

https://webusers.imj-prg.fr/~vincent.humiliere/2M270-2016/Poly2M270-2016.pdf

M. Audin, O. Debarre, *Algèbre linéaire 2*, polycopié 1998/1999, Université Louis Pasteur, Strasbourg.

https://www.math.ens.fr/~debarre/DEUG2.pdf

Page web du cours :

https://webusers.imj-prg.fr/~alexandru.oancea/2020-L2-LU2MA123/LU2MA123-2020.html

Alexandru OANCEA

Sorbonne Université

Institut de Mathématiques de Jussieu – Paris Rive Gauche (UMR 7586 du CNRS)

URL : http://webusers.imj-prg.fr/~alexandru.oancea

E-mail : alexandru.oancea@imj-prg.fr

TABLE DES MATIÈRES

CHAPITRE 0

RAPPELS : VALEURS PROPRES, VECTEURS PROPRES, DIAGONALISABILITÉ

Ce chapitre 0 est constitué de **rappels** sur la diagonalisation des endomorphismes, telle que vue dans le cours LU2MA221 "Algèbre linéaire et bilinéaire I" en L2 S3. La référence est la section §4.2 du polycopié de Yves Coudène, qui est disponible à l'adresse

`https://www.lpsm.paris/pageperso/coudene/2MA221-algebre-lineaire-bilineaire-I.html`

Nous allons revisiter ces notions dans le cours. Nous supposons connue la méthode du pivot pour la résolution des systèmes linéaires $AX = Y$.

0.1. Somme directe de sous-espaces

***Définition 0.1.1* (Sous-espaces en somme directe).** — Soient V un k-espace vectoriel, $E_1, \ldots, E_n$ des sous-espaces de V. (Ni V ni les E_i ne sont supposés de dimension finie.)

(1) La *somme des sous-espaces* $E_1, \ldots, E_n$, notée $E_1 + \cdots + E_n$ ou $\sum_{i=1}^n E_i$, est le sous-espace de V engendré par $E_1 \cup \cdots \cup E_n$. L'on montre que

$$(*) \qquad E_1 + \cdots + E_n = \{x_1 + \cdots + x_n \,:\, \forall i,\, x_i \in E_i\}.$$

(2) On dit que les E_i *sont en somme directe* si pour tous $x_1 \in E_1$, ..., $x_n \in E_n$, l'égalité $x_1 + \cdots + x_n = 0$ entraîne $x_1 = 0 = \cdots = x_n$. Ceci équivaut à dire que tout élément x de $E_1 + \cdots + E_n$ s'écrit *de façon unique* $x = x_1 + \cdots + x_n$ avec $x_i \in E_i$. Dans ce cas, $E_1 + \cdots + E_n$ est noté $E_1 \oplus \cdots \oplus E_n$ ou $\bigoplus_{i=1}^n E_i$.

Si les sous-espaces E_i, $i = 1, \ldots, n$ sont de dimension finie, alors ils sont en somme directe si et seulement si

$$(\star) \qquad \dim(E_1 + \cdots + E_n) = \dim(E_1) + \cdots + \dim(E_n).$$

***Terminologie 0.1.2*.** — Si $E_1, \ldots, E_n$ sont en somme directe et si de plus $E_1 \oplus \cdots \oplus E_n$ égale V, alors on dit que V est la somme directe des E_i.

***Remarques 0.1.3*.** — (1) Il résulte de la définition que $E_1, \ldots, E_n$ sont en somme directe si et seulement si, pour tout $i = 1, \ldots, n$, on a : $E_i \cap \sum_{j \neq i} E_j = 0$.

(2) En particulier, si $n = 2$, alors E_1 et E_2 sont en somme directe si et seulement si $E_1 \cap E_2 = (0)$.

(0) version du 21/1/2020

(3) **Attention !** Si des sous-espaces sont en somme directe, leur somme n'est pas nécessairement égale à l'espace tout entier : par exemple si E_1, E_2 sont deux droites distinctes dans $\mathbb{R}^3$, leur somme est directe, et c'est un plan de $\mathbb{R}^3$, et non $\mathbb{R}^3$ tout entier !

(4) **Attention !** Si $n \geq 3$, la condition $E_i \cap E_j = \{0\}$ pour $i \neq j$ *n'entraîne pas* que la somme des E_i soit directe : par exemple si E_1, E_2, E_3 sont trois droites distinctes dans $\mathbb{R}^2$, elles vérifient $E_i \cap E_j = \{0\}$ pour $i \neq j$, mais leur somme n'est pas directe (car $E_1 + E_2$ égale $\mathbb{R}^2$ donc contient E_3).

Définition 0.1.4 **(Sous-espaces supplémentaires)**. — Soient V un espace vectoriel, E, F deux sous-espaces de V. On dit que E et F sont des sous-espaces *supplémentaires* si $V = E \oplus F$, c.-à-d., si $E \cap F = (0)$ et $E + F = V$.

Si V est de dimension finie, ceci équivaut à dire que $E \cap F = (0)$ et $\dim(E) + \dim(F) = \dim(V)$.

Proposition 0.1.5. — *Soit V un k-espace vectoriel de dimension finie n. Tout sous-espace E de V admet un supplémentaire.* □

Remarque 0.1.6. — Soient V un espace vectoriel et E, F deux sous-espaces de dimension finie. Alors

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

0.2. Espaces propres et critères de diagonalisabilité

Définition 0.2.1 **(valeurs propres, vecteurs propres)**. — Soit V un k-espace vectoriel et u un endomorphisme de V. Un scalaire $\lambda \in k$ est une *valeur propre de u* s'il existe $x \in V \setminus \{0\}$ tel que $u(x) = \lambda x$, ou encore si

$$\ker(u - \lambda \mathrm{Id}) \neq 0.$$

Un élément $x \in \ker(u - \lambda \mathrm{Id})$ s'appelle *vecteur propre pour la valeur propre λ*. Le sous-espace

$$V_\lambda = \ker(u - \lambda \mathrm{Id}) \subseteq V$$

s'appelle *sous-espace propre pour la valeur propre λ*.

Théorème 0.2.2. — *Soient V un k-espace vectoriel (pas nécessairement de dimension finie), u un endomorphisme de V, et $\lambda_1, \ldots, \lambda_r$ des valeurs propres de u deux à deux distinctes. Pour $i = 1, \ldots, r$, on note*

$$E_i = V_{\lambda_i} = \{v \in V \mid u(v) = \lambda_i v\}$$

le sous-espace propre associé. Alors les V_{λ_i} sont en somme directe.

Démonstration. — Montrons par récurrence sur $r \geq 1$ l'assertion suivante : si l'on a une égalité $x_1 + \cdots + x_r = 0$ avec $x_i \in V_{\lambda_i}$, alors $x_1 = \cdots = x_r = 0$. Il n'y a rien à démontrer pour $r = 1$, donc on peut supposer $r \geq 2$ et l'assertion établie pour $r - 1$. L'on a $u(x_1 + \cdots + x_r) - \lambda_r(x_1 + \cdots + x_r) = 0$, ou encore

$$(\lambda_1 - \lambda_r)x_1 + \cdots + (\lambda_{r-1} - \lambda_r)x_{r-1} = 0.$$

L'hypothèse de récurrence assure que, pour tout $i = 1, \ldots, r-1$, l'on a $(\lambda_i - \lambda_r)x_i = 0$, ou encore $x_i = 0$ puisque $\lambda_i - \lambda_r \neq 0$. Or $x_1 + \cdots + x_r = 0$, ce qui entraîne $x_r = 0$ et achève la démonstration. □

La somme $\bigoplus_{i=1}^r V_{\lambda_i}$ n'est pas nécessairement égale à V ; si V est de dimension finie, c'est le cas si et seulement si il existe une base formée de vecteurs propres de u.

***Définition 0.2.3* (Endomorphismes diagonalisables).** — Soient V un k-espace vectoriel de dimension finie et $u \in \mathrm{End}_k(V)$. Les conditions suivantes sont équivalentes :

(i) il existe une base de V dans laquelle la matrice de u est diagonale.
(ii) V admet une base formée de vecteurs propres de u ;
(iii) les vecteurs propres de u engendrent V ;
(iv) la somme des espaces propres de u égale V ;
(v) V est la somme directe des espaces propres de u.

Si ces conditions sont vérifiées, on dit que u est *diagonalisable*.

***Proposition 0.2.4* (Valeurs propres distinctes).** — *Soit $u \in \mathrm{End}_k(V)$ ($\dim_k(V) = n$). Si u possède n valeurs propres distinctes, alors u est diagonalisable.*

Démonstration. — L'endomorphisme u possède n espaces propres distincts $V_1, \ldots, V_n$, qui sont en somme directe d'après le théorème précédent. Alors le sous-espace $E = V_1 \oplus \cdots \oplus V_n$ de V est de dimension

$$n \geq \dim E = \sum_{i=1}^{n} \dim V_i \geq n = \dim V.$$

Ainsi $\dim E = \dim V$ et par conséquent $V = E = V_1 \oplus \cdots \oplus V_n$. De plus, chaque V_i est nécessairement de dimension 1. □

Cette proposition fournit une condition *suffisante* de diagonalisabilité. Bien entendu, cette condition n'est **pas nécessaire** : par exemple la matrice identité I_n (≥ 2) est diagonale et a toutes ses valeurs propres égales à 1.

0.3. Polynôme caractéristique d'un endomorphisme

Soit V un k-espace vectoriel de dimension finie n. L'outil fondamental pour trouver les valeurs propres d'un endomorphisme $u \in \mathrm{End}_k(V)$ est son *polynôme caractéristique*

$$P_u(X) = \det(u - XI_n) \in k[X].$$

Celui-ci est défini de la manière suivante : l'on choisit une base quelconque de V, l'on représente u par une matrice $A \in M_n(k)$ et l'on pose

$$P_u(X) = \det(A - XI_n) \in k[X].$$

La définition ne dépend pas du choix de la base.

***Proposition 0.3.1*.** — *Un scalaire $\lambda \in k$ est valeur propre de u si et seulement si il est racine du polynôme caractéristique.*

Démonstration. — L'équation $u(x) = \lambda x$, $x \in V$ équivaut à $(u - \lambda\mathrm{Id})x = 0$. L'existence d'une solution non-nulle équivaut à la non-injectivité de l'application linéaire $u - \lambda\mathrm{Id} : V \to V$. Puisque V est de dimension finie, ceci équivaut à l'annulation du déterminant $\det(u - \lambda I_n) = P_u(\lambda)$. □

***Corollaire 0.3.2*.** — *Soit $u \in \mathrm{End}_k(V)$ ($\dim_k(V) = n$). Si le polynôme caractéristique $P_u(X)$ possède n racines distinctes, alors u est diagonalisable.* □

CHAPITRE 1

POLYNÔMES D'ENDOMORPHISMES

La terme de "réduction des endomorphismes" fait référence au problème suivant.

Problème (réduction des endomorphismes). Étant donné un k-espace vectoriel V (de dimension finie n) et un endomorphisme $u \in \mathrm{End}_k(V)$, trouver une base de V dans laquelle la matrice de u prend la forme "la plus simple possible".

Les applications les plus simples sont les homothéties $\lambda\mathrm{Id}$, $\lambda \in k$. C'est une des raisons pour lesquelles l'on a introduit dans la section précédente les notions de *valeur propre* et *sous-espace propre* de u. Dans ce chapitre, nous allons raffiner cette étude. On y démontre des critères de diagonalisation, un théorème de trigonalisation, un théorème de réduction à une forme diagonale par blocs via l'étude des sous-espaces caractéristiques, ainsi que deux théorèmes importants : le théorème de Cayley-Hamilton et le lemme des noyaux.

Les appendices contiennent des discussions de la notion de somme directe d'espaces vectoriels, du théorème de Bézout et du fait que le corps $\mathbb{C}$ est algébriquement clos.

Sauf mention contraire tous les espaces vectoriels considérés dans ce chapitre sont de dimension finie.

1.1. Critère de diagonalisabilité

Dans la section précédente nous avons vu une condition suffisante de diagonalisabilité. Celle-ci peut être complétée en une condition nécessaire et suffisante (CNS).

Définition et proposition 1.1.1 **(Multiplicités algébrique et géométrique d'une valeur propre)**

Soient V un $\mathbb{C}$-espace vectoriel de dimension n, $u \in \mathrm{End}_{\mathbb{C}}(V)$, $\lambda_1, \ldots, \lambda_r$ les racines (deux à deux distinctes) du polynôme caractéristique $P_u(X)$ dans $\mathbb{C}$. D'une part, $P_u(X)$ se factorise

$$P_u(X) = (-1)^n (X - \lambda_1)^{m_1} \cdots (X - \lambda_r)^{m_r}$$

où m_i est la multiplicité de λ_i comme racine de $P_u(X)$. D'autre part, d'après la Proposition (0.3.1), $\lambda_1, \ldots, \lambda_r$ sont les valeurs propres de u.

On appelle multiplicité algébrique *(resp.* géométrique*) de la valeur propre λ_i sa multiplicité m_i comme racine de $P_u(X)$ (resp. la dimension n_i de l'espace propre V_{λ_i}).*

(1) *On a* $\dim V_{\lambda_i} \leq m_i$ *pour tout i.*

(2) *u est diagonalisable* $\iff \dim V_{\lambda_i} = m_i$ *pour tout i.*

(0) version du 9/02/2020

Démonstration. — (1) Pour tout i, soit $\mathscr{C}^i$ une base de V_{λ_i}. Comme les espaces propres sont en somme directe, la famille $\mathscr{C} = \mathscr{C}^1 \cup \cdots \cup \mathscr{C}^r$ est une famille libre, donc on peut la compléter en une base $\mathscr{B}$ de V. Alors $A = \mathrm{Mat}_{\mathscr{B}}(u)$ est de la forme suivante :

$$A = \left(\begin{array}{c|c|c|c|c} \lambda_1 I_{n_1} & 0 & \cdots & 0 & * \\ \hline 0 & \lambda_2 I_{n_2} & \ddots & \vdots & * \\ \hline 0 & \ddots & \ddots & 0 & * \\ \hline \vdots & \ddots & 0 & \lambda_r I_{n_r} & * \\ \hline 0 & \cdots & 0 & 0 & B \end{array}\right)$$

où B est une matrice carrée de taille $p = n-(n_1+\cdots+n_r)$. En particulier, A est triangulaire par blocs et l'on en déduit l'égalité

$$P_u(X) = \det(A - XI_n) = P_B(X)\prod_{i=1}^{r}(\lambda_i - X)^{n_i}.$$

Donc $\prod_{i=1}^r(\lambda_i - X)^{n_i}$ divise $P_u(X)$, d'où $n_i \leq m_i$ pour tout i, ce qui prouve (1).

(2) Si $n_i = m_i$ pour tout i, alors le sous-espace $E = \bigoplus_{i=1}^r V_{\lambda_i}$ est de dimension $\sum_{i=1}^r m_i = n$, donc égale V, donc u est diagonalisable. Réciproquement, si u est diagonalisable, il existe une base $\mathscr{B}$ de V telle que

$$A = \mathrm{Mat}_{\mathscr{B}}(u) = \left(\begin{array}{c|c|c|c} \lambda_1 I_{n_1} & 0 & \cdots & 0 \\ \hline 0 & \lambda_2 I_{n_2} & \ddots & \vdots \\ \hline \vdots & \ddots & \ddots & \vdots \\ \hline 0 & \cdots & 0 & \lambda_r I_{n_r} \end{array}\right)$$

alors $P_u(X) = \det(A-XI_n) = \prod_{i=1}^r(\lambda_i - X)^{n_i} = (-1)^n\prod_{i=1}^r(X-\lambda_i)^{n_i}$, d'où $n_i = m_i$ pour tout i. □

Le critère (1.1.1) est utile en pratique : étant connues les racines du polynôme caractéristique avec leurs multiplicités, l'on calcule les sous-espaces propres correspondants. Pour déterminer si l'endomorphisme u est diagonalisable l'on compare la multiplicité algébrique à la multiplicité géométrique.

Les espaces propres sont en somme directe, mais leur somme n'est E que si u est diagonalisable. Par exemple, la seule valeur propre de la matrice carrée

$$U_s = \begin{pmatrix} 0 & 1 & 0 & \ldots & \ldots & 0 \\ 0 & 0 & 1 & 0 & \ldots & 0 \\ \vdots & & \ddots & \ddots & \ddots & \vdots \\ \vdots & & & \ddots & \ddots & 0 \\ 0 & \ldots & \ldots & \ldots & 0 & 1 \\ 0 & \ldots & \ldots & \ldots & \ldots & 0 \end{pmatrix}$$

est 0 et l'espace propre associé est de dimension 1.

1.2. Sous-espaces stables

La notion clé dans le problème de la réduction des endomorphismes est celle de sous-espace stable.

***Définition 1.2.1* (Sous-espace stable).** — Soit $u \in \mathrm{End}_k(V)$. On dit qu'un sous-espace E de V est *stable par* u si

$$u(E) \subseteq E.$$

Dans ce cas, la restriction de u à E induit un endomorphisme de E, noté $u|_E$ ou u_E.

Exemple 1.2.2. — Les sous-espaces propres d'un endomorphisme u sont stables par u. La restriction de u à chaque sous-espace propre est une homothétie.

Remarque 1.2.3. — Supposons que E est un sous-espace stable par u, choisissons une base de E et complétons-là en une base de V. La matrice de u dans cette base est triangulaire supérieure par blocs

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

où A représente la matrice de $u|_E$ dans la base choisie.

Supposons que l'on a trouvé une décomposition $V = E_1 \oplus \cdots \oplus E_r$ en somme directe de sous-espaces stables par u, de sorte que $u(E_i) \subseteq E_i$ pour tout i. Définissons une base de V en concaténant des bases de $E_1, \ldots, E_r$. La matrice de u dans cette base est diagonale par blocs

$$\begin{pmatrix} A_1 & 0 & \cdots & 0 \\ 0 & A_2 & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & \cdots & 0 & A_r \end{pmatrix},$$

où A_i représente la matrice de $u|_{E_i}$ dans la base choisie.

Lorsque l'endomorphisme u est diagonalisable la situation peut être reformulée en termes de sous-espaces stables de la manière suivante : l'on a une décomposition $V = V_{\lambda_1} \oplus \cdots \oplus V_{\lambda_r}$ en somme directe de sous-espace propres, chaque V_{λ_i} est stable par u et sa dimension est égale à la multiplicité algébrique de λ_i. La restriction $u|_{V_{\lambda_i}}$ est une homothétie de rapport λ_i et la matrice de u dans une base constituée de vecteurs propres est diagonale.

Pour mieux saisir la notion de sous-espace stable, nous donnons maintenant un résultat et un exemple qui en font usage.

***Proposition 1.2.4* (Restriction d'un endomorphisme diagonalisable)**

Soient V un k-espace vectoriel de dimension finie, $u \in \mathrm{End}_k(V)$ un endomorphisme diagonalisable, E un sous-espace de V stable par u. Alors E admet une base formée de vecteurs propres de u, i.e. la restriction u_E de u à E est diagonalisable.

Démonstration. — La preuve est une variation sur celle du Théorème 0.2.2. D'après 0.2.3, il suffit de montrer que E est engendré par des vecteurs propres de u. Comme u est diagonalisable, tout $x \in E$ s'écrit dans V comme une somme de vecteurs propres :

$$(\dagger) \qquad x = x_1 + \cdots + x_r, \qquad \text{avec} \quad x_i \in V_{\mu_i} \quad \text{et} \quad \mu_i \neq \mu_j \text{ si } i \neq j.$$

Montrons par récurrence sur r que pour tout $x \in E$ et toute écriture (†) comme ci-dessus, chaque x_i appartient à E (ce qui prouvera le théorème). C'est OK pour $r = 1$, donc on peut supposer $r \geq 2$ et le résultat démontré pour $r - 1$. Appliquant $u - \mu_r \mathrm{Id}_V$ à (†) on obtient

$$x' = (u - \mu_r \mathrm{Id}_V)(x) = \sum_{i=1}^{r-1} (\mu_i - \mu_r) x_i$$

et $x' \in E$ puisque E est stable par u. Donc par hypothèse de récurrence, chacun des vecteurs propres $(\mu_i - \mu_r)x_i$, $i = 1, \dots, r-1$ appartient à E, donc x_i appartient aussi à E puisque $\mu_i - \mu_r \neq 0$. En reportant ceci dans (†) on obtient $x_r \in E$, ce qui prouve le théorème. □

Rappels 1.2.5. — Soit k un corps. Si $n \cdot 1_k = 1_k + \cdots + 1_k$ (n termes) est $\neq 0$ pour tout entier $n > 0$, on dit que k *est de caractéristique* 0 ; c'est le cas par exemple pour $\mathbb{Q}, \mathbb{R}, \mathbb{C}$. Sinon, le plus petit entier $p > 0$ tel que $p \cdot 1_k = 0$ est nécessairement un nombre premier (car si $p = rs$ avec $r, s \geq 1$, l'égalité $0 = (r \cdot 1_k)(s \cdot 1_k)$ entraîne que $r \cdot 1_k = 0$ ou $s \cdot 1_k = 0$, disons $r \cdot 1_k = 0$, mais alors la minimalité de p entraîne que $r = p$) ; dans ce cas on dit que k *est de caractéristique* p. D'autre part, si V est un k-espace vectoriel et $p \in \mathrm{End}_k(V)$, rappelons qu'on dit que p est un **projecteur** si $p^2 = p \circ p$ est égal à p.

Exemple 1.2.6 **(Symétries)**. — Soient k un corps de caractéristique $\neq 2$ (par exemple, $k = \mathbb{Q}, \mathbb{R}$ ou $\mathbb{C}$), V un k-espace vectoriel de dimension n, et $s \in \mathrm{End}_k(V)$ tel que $s^2 = \mathrm{Id}_V$. Alors s est diagonalisable ; plus précisément, soient

$$p_+ = \frac{\mathrm{Id}_V + s}{2}, \qquad p_- = \frac{\mathrm{Id}_V - s}{2}, \qquad V_\pm = \mathrm{Im}(p_\pm).$$

Alors p_+ et p_- sont des projecteurs et l'on a :

$$V = V_+ \oplus V_- \qquad \text{et} \qquad \forall x \in V_\pm, \quad s(x) = \pm x.$$

Donc, si $s \neq \pm \mathrm{Id}_V$, alors V_+ et V_- sont non-nuls et $V_\pm$ est l'espace propre associé à la valeur propre ± 1 ; dans ce cas, s est la symétrie par rapport à V_+ parallèlement à V_-.

Pour montrer ces affirmations notons les identités suivantes : $p_+^2 = p_+$, $p_-^2 = p_-$ et $p_- = \mathrm{Id}_V - p_+$ d'où $p_+ p_- = 0 = p_- p_+$. Ainsi p_+ et p_- sont des projecteurs et l'on a $V = V_+ \oplus V_-$. De plus, si $x \in V_\pm$, on voit aussitôt que $s(x) = \pm x$, ce qui achève la preuve.

Remarque 1.2.6.1. — Attention, si k est de caractéristique 2, c.-à-d., si $2 = 0$ dans k (par exemple, si k est le corps à deux éléments $\mathbb{F}_2 = \mathbb{Z}/2\mathbb{Z}$), la matrice $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \in M_2(k)$ vérifie $A^2 = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} = I_2$ mais A n'est pas diagonalisable : en effet sa seule valeur propre est 1, donc si A était diagonalisable on aurait $A = I_2$, ce qui n'est pas le cas.

1.3. Trigonalisation

Définition 1.3.1 **(Endomorphismes trigonalisables)**. — Soit $u \in \mathrm{End}_k(V)$. On dit que u est *trigonalisable* s'il existe une base $\mathscr{B}$ de V dans laquelle la matrice $A = \mathrm{Mat}_{\mathscr{B}}(u)$ est triangulaire, disons supérieure. [(1)]

Dans ce cas, soient $\lambda_1, \dots, \lambda_n$ les coefficients diagonaux ($n = \dim_k V$), et soit X une indéterminée. Alors

$$P_u(X) = \det(A - XI_n) = \prod_{i=1}^{n} (\lambda_i - X)$$

donc $\lambda_1, \dots, \lambda_n$ sont les n racines (comptées avec multiplicités) du polynôme caractéristique $P_u(X)$. On voit donc qu'une condition *nécessaire* pour que u soit trigonalisable est que $P_u(X)$ ait toutes ses racines dans k. Ceci conduit à la définition suivante :

(1) Si la matrice de u dans une base $(v_1, \dots, v_n)$ est triangulaire supérieure, alors la matrice dans la base $(v_n, \dots, v_1)$ est triangulaire inférieure, et vice-versa, donc on pourrait dans la définition remplacer le mot « supérieure » par « inférieure ».

***Définition 1.3.2* (Polynômes scindés, corps algébriquement clos)**

Soient k un corps et $P \in k[X]$ un polynôme de degré $n \geq 1$.

(1) On dit que P est *scindé* dans $k[X]$ s'il admet n racines $\lambda_1, \ldots, \lambda_n$ dans k (comptées avec multiplicités), c.-à-d., si P se factorise dans $k[X]$ en produit de facteurs de degré 1 :

$$P = a(X - \lambda_1) \cdots (X - \lambda_n)$$

(où a est le coefficient dominant de P).

(2) On dit que k est *algébriquement clos* si tout polynôme $P \in k[X]$ de degré ≥ 1 est scindé. Par exemple, on sait que $\mathbb{C}$ est algébriquement clos (une démonstration est donnée dans un appendice à ce chapitre).

***Théorème 1.3.3* (Trigonalisation).** — *Un endomorphisme $u \in \mathrm{End}_k(V)$ est trigonalisable si et seulement si son polynôme caractéristique est scindé dans k.*

En particulier, lorsque $k = \mathbb{C}$, tout endomorphisme est trigonalisable.

Démonstration. — L'implication directe a été démontrée plus haut, il s'agit de prouver l'implication inverse.

On procède par récurrence sur la dimension n de V. Lorsque $n = 1$ il n'y a rien à démontrer, on suppose donc $n \geq 2$ et l'affirmation démontrée pour $n - 1$. Il existe par hypothèse une racine $\lambda \in k$ du polynôme caractéristique, c'est-à-dire une valeur propre. Soit e un vecteur propre associé, soit $E = \mathrm{Vect}(e)$ et soit E' un supplémentaire de E dans V. Comme discuté dans la Remarque 1.2.3, la matrice de u dans une base de V constituée de e et d'une base $\mathscr{C}$ de E' est "triangulaire supérieure par blocs" du type

$$\begin{pmatrix} \lambda & B \\ 0 & C \end{pmatrix}.$$

Par ailleurs, l'on a $P_u(X) = (\lambda - X)P_C(X)$ et, puisque $P_u(X)$ est scindé sur k, il en est de même pour $P_C(X)$. Soit $p : V \to E'$ est la projection sur E' parallèlement à E. Puisque $\dim E' = n - 1$ nous pouvons appliquer l'hypothèse de récurrence à l'endomorphisme $p \circ u|_{E'} \in \mathrm{End}_k(E')$, dont la matrice dans la base $\mathscr{C}$ est C : dans une base convenable $\mathscr{C}'$ de E' sa matrice est triangulaire supérieure. Si l'on ajoute cette base à e l'on obtient une base de V dans laquelle la matrice de u est triangulaire supérieure. □

***Corollaire 1.3.4* (Déterminant (resp. trace) = produit (resp. somme) des valeurs propres)**

Soient $A \in M_n(\mathbb{C})$ et $\lambda_1, \ldots, \lambda_n$ les n racines dans $\mathbb{C}$ (comptées avec multiplicité) du polynôme caractéristique $P_A(X)$. Alors

$$\det(A) = \lambda_1 \cdots \lambda_n, \qquad \mathrm{Tr}(A) = \lambda_1 + \cdots + \lambda_n.$$

Démonstration. — D'après le théorème 1.3.3, il existe $P \in \mathrm{GL}_n(\mathbb{C})$ telle que $A' = P^{-1}AP$ soit triangulaire ; notons $\lambda_1, \ldots, \lambda_n$ ses coefficients diagonaux, alors

$$P_{A'}(X) = \prod_{i=1}^{n}(\lambda_i - X), \qquad \det(A') = \lambda_1 \cdots \lambda_n, \qquad \mathrm{Tr}(A') = \lambda_1 + \cdots + \lambda_n.$$

D'autre part A' et A ont même polynôme caractéristique, même déterminant et même trace. Alors $\lambda_1, \ldots, \lambda_n$ sont les racines de $P_A(X) = P_{A'}(X)$ et l'on a $\det(A) = \det(A') = \lambda_1 \cdots \lambda_n$ et $\mathrm{Tr}(A) = \mathrm{Tr}(A') = \lambda_1 + \cdots + \lambda_n$. □

Remarque 1.3.5. — Par définition, un endomorphisme $u \in \mathrm{End}_k(V)$ est trigonalisable s'il existe une base $\mathscr{B} = (v_1, \ldots, v_n)$ de V telle que la matrice de u dans la base $\mathscr{B}$ soit triangulaire supérieure. Ceci revient à dire que, pour tout $j = 1, \ldots, n$, le vecteur $u(v_j)$ est combinaison linéaire des vecteurs v_i avec $i \leq j$, ou encore qu'il existe des scalaires $a_{ij} \in k$ tels que

$$u(v_j) = \sum_{i=1}^{j} a_{ij} v_i.$$

La démonstration du Théorème 1.3.3 est effective. Elle permet en pratique de trigonaliser un endomorphisme grâce à l'algorithme suivant.

Soit $u \in \mathrm{End}_k(V)$ avec $\dim V = n$.

Test de trigonalisabilité. On calcule d'abord le polynôme caractéristique $P_u(X) \in k[X]$.

– Si $P_u(X)$ n'est pas scindé sur k alors l'endomorphisme u n'est pas trigonalisable et l'algorithme ne fonctionnera pas.
– On suppose donc dorénavant que P_u est scindé sur k. Il s'écrit

$$P_u(X) = (-1)^n (X - \lambda_1)^{m_1} \cdots (X - \lambda_r)^{m_r}$$

avec $\lambda_i \in k$, $m_i \in \mathbb{N}^*$ pour $i = 1, \ldots, r$ et $m_1 + \cdots + m_r = n$. L'endomorphisme u est alors trigonalisable et l'algorithme fonctionnera.

L'algorithme reçoit en entrée un sous-espace vectoriel $E \subset V$ et un endomorphisme trigonalisable $f \in \mathrm{End}_k(E)$ dont on connaît les valeurs propres. Il produit à la sortie une base $\mathscr{B}$ de l'espace vectoriel E dans laquelle la matrice de f est triangulaire supérieure.

Pour déterminer une base $\mathscr{B}$ dans laquelle u est trigonalisable, on appelle l'algorithme avec $E = V$ et $f = u$. Les valeurs propres de u sont $\lambda_1, \ldots, \lambda_r$.

Algorithme 1 **(Algorithme de trigonalisation – version 1.0)**

Étape 1. (Calcul d'espaces propres). *On calcule les espaces propres de f. Si la somme de leur dimension vaut $\dim E$ alors f est diagonalisable dans une base de vecteurs propres ; on calcule une telle base et l'algorithme la renvoie comme résultat. Sinon, on passe au 2.*

Étape 2. (Calcul d'un supplémentaire). *On calcule un supplémentaire dans E pour la somme directe des espaces propres de f. On le note E'.*

Étape 3. (Trigonalisation de dimension inférieure). *Soit $p : E \to E'$ la projection linéaire sur E' parallèlement à la somme des espaces propres de f. Nous définissons*

$$f' = p \circ f|_{E'} \in \mathrm{End}_k(E').$$

On appelle l'algorithme de façon récursive avec en entrée le couple constitué de E' et $f' \in \mathrm{End}_k(E')$. On reçoit une base $\mathscr{B}'$ de E'. L'algorithme renvoie une base $\mathscr{B} = \mathscr{C} \cup \mathscr{B}'$, où $\mathscr{C}$ est une base de la somme des espaces propres de f.

Exemple 1.3.6. — Voici comment fonctionne l'algorithme en pratique sur un exemple. Considérons la matrice

$$A = \begin{pmatrix} -2 & 3 & 3 \\ -2 & 1 & 2 \\ -3 & 3 & 4 \end{pmatrix} \in M_3(\mathbb{R}).$$

Son polynôme caractéristique est $P_A(X) = -(X-1)^3$. Il est scindé sur $\mathbb{R}$ et la matrice est donc trigonalisable. Elle possède une unique valeur propre $\lambda = 1$. On note e_1, e_2, e_3 les vecteurs de la base canonique de $\mathbb{R}^3$.

On lance l'algorithme avec $E = \mathbb{R}^3$ et $f = A$.

On calcule une base de l'espace propre $E_1 = \ker(A - I_3)$ par la méthode du pivot de Gauss et l'on trouve $E_1 = \mathrm{Vect}(e_1 + e_3)$. La dimension de E_1 est strictement plus petite que 3 et par conséquent la matrice A n'est pas diagonalisable. On choisit un supplémentaire de E_1, par exemple $E' = \mathrm{Vect}(e_2, e_3)$. Dans la base $(e_1 + e_3, e_2, e_3)$ l'endomorphisme f est exprimé par la matrice

Étape 1

Étape 2

$$\widetilde{A} = P^{-1}AP = \begin{pmatrix} 1 & 3 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}, \quad \text{avec} \quad P = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}.$$

En effet, nous avons $A \cdot (e_1 + e_3) = e_1 + e_3$, $Ae_2 = 3(e_1 + e_3) + e_2$, $Ae_3 = 3(e_1 + e_3) + 2e_2 + e_3$. Nous avons eu de la chance et la matrice A' est déjà triangulaire supérieure ! [(2)]

Fin

Voyons comment le choix du supplémentaire influe sur le déroulement de l'algorithme. Choisissons comme supplémentaire $E' = \mathrm{Vect}(e_2 + e_3, e_3)$. Dans la base $(e_1 + e_3, e_2 + e_3, e_3)$ l'endomorphisme f est exprimé par la matrice

Étape 2

$$\widetilde{A} = P^{-1}AP = \begin{pmatrix} 1 & 6 & 3 \\ 0 & 3 & 2 \\ 0 & -2 & -1 \end{pmatrix}, \quad \text{avec} \quad P = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix}.$$

En effet, nous avons $A \cdot (e_1 + e_3) = e_1 + e_3$, $A \cdot (e_2 + e_3) = 6(e_1 + e_3) + 3(e_2 + e_3) - 2e_3$, $Ae_3 = 3(e_1 + e_3) + 2(e_2 + e_3) - e_3$. Puisque celle-ci n'est pas triangulaire supérieure, il faut itérer l'algorithme.

Appel récursif

Considérons l'endomorphisme f' de $E' = \mathrm{Vect}(e_2 + e_3, e_3)$ donné dans cette base par le carré inférieur droit de la matrice $\widetilde{A}$, à savoir

$$A' = \begin{pmatrix} 3 & 2 \\ -2 & -1 \end{pmatrix}.$$

(Celui-ci est précisément la composition de la projection sur E' parallèlement à E_1 avec $f|_{E'}$.) Pour calculer $E'_1 = \ker(f' - \mathrm{Id}_{E'})$, nous identifions E' à $\mathbb{R}^2$ en faisant correspondre à la base $(e_2 + e_3, e_3)$ la base canonique (e'_1, e'_2) de $\mathbb{R}^2$. On trouve $\ker(A' - I_2) = \mathrm{Vect}(e'_1 - e'_2)$, ou encore $E'_1 = \ker(f' - \mathrm{Id}_{E'}) = \mathrm{Vect}\big((e_2 + e_3) - e_3\big) = \mathrm{Vect}(e_2)$. On choisit comme supplémentaire de E'_1 dans E' le sous-espace $E'' = \mathrm{Vect}(e_3)$. Dans la base $(e_1 + e_3, e_2, e_3)$ la matrice de l'endomorphisme f est triangulaire supérieure

Étape 1'

Étape 2'

$$\widetilde{A} = P^{-1}AP = \begin{pmatrix} 1 & 3 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}, \quad \text{avec} \quad P = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix},$$

puisque $A \cdot (e_1 + e_3) = e_1 + e_3$, $Ae_2 = 3(e_1 + e_3) + e_2$, $Ae_3 = 3(e_1 + e_3) + 2e_2 + e_3$ tel que déjà calculé plus haut.

Fin

[(2)] En réalité l'algorithme continue à la manière du cas examiné plus bas, mais le résultat final reste inchangé. On aurait pu inclure dans l'algorithme un test de trigonalité pour l'arrêter dès qu'une base de trigonalisation est trouvée, mais nous avons choisi de ne pas le faire pour plus de lisibilité.

1.4. Polynômes d'endomorphismes

Le but de cette section et de la suivante est d'appliquer notre compréhension de l'arithmétique des polynômes à l'algèbre linéaire, via le concept de *polynôme d'endomorphisme*. En particulier le théorème purement arithmétique de Bézout joue un rôle essentiel dans la réduction des endomorphismes.

***Définition 1.4.1* (Polynômes d'endomorphismes).** — Soient V un k-espace vectoriel et $u \in \mathrm{End}_k(V)$. Étant donné un polynôme

$$Q = a_0 + a_1 X + \cdots + a_d X^d$$

à coefficients dans k l'on pose

$$Q(u) = a_0 \,\mathrm{Id}_V + a_1 u + \cdots + a_d u^d \in \mathrm{End}_k(V),$$

où u^i désigne $u \circ \cdots \circ u$ (i fois) et $u^0 = \mathrm{Id}_V$. Un élément de $\mathrm{End}_k(V)$ est un *polynôme en* u s'il est de la forme $Q(u)$ pour un certain $Q \in k[X]$.

L'application

$$\phi : k[X] \to \mathrm{End}_k(V), \qquad Q \mapsto Q(u)$$

est un *homomorphisme de k-algèbres*. Par définition, ceci signifie que :

- ϕ est k-linéaire.
- ϕ est unitaire : $\phi(1) = \mathrm{Id}_V$.
- ϕ préserve la multiplication : $\phi(QR) = \phi(Q)\phi(R)$ pour tous $Q, R \in k[X]$.

Cette dernière relation entraîne notamment

$$Q(u) \circ R(u) = (QR)(u) = (RQ)(u) = R(u) \circ Q(u).$$

En particulier, « les polynômes en u commutent entre eux » et « la composition de deux polynômes en u est encore un polynôme en u ».

***Définition 1.4.2* (Polynômes de matrices).** — Étant donnés $A \in M_n(k)$ et un polynôme $Q = a_0 + a_1 X + \cdots + a_d X^d$ à coefficients dans k, l'on note $A^0 = I_n$ et

$$Q(A) = a_0 I_n + a_1 A + \cdots + a_d A^d \in M_n(k).$$

Une matrice est un *polynôme en A* si elle est de la forme $Q(A)$ pour un certain $Q \in k[X]$.

On obtient comme avant un morphisme de k-algèbres

$$\phi : k[X] \to M_n(k), \qquad Q \mapsto Q(A),$$

« les polynômes en A commutent entre eux », i.e. pour tous $Q, R \in k[X]$ on a

$$Q(A)R(A) = (QR)(A) = (RQ)(A) = R(A)Q(A),$$

et « la multiplication de deux polynômes en A est encore un poynôme en A ».

***Exercice 1.4.3*.** — Soit V un espace vectoriel de dimension n et $u \in \mathrm{End}_k(V)$. Si A est la matrice de u dans une base $\mathscr{B}$, alors la matrice de $Q(u)$ dans la base $\mathscr{B}$ est $Q(A)$. (Après avoir résolu cet exercice une première fois, rédiger une approche plus abstraite employant le concept d'isomorphisme de k-algèbres.)

***Lemme 1.4.4*.** — *Soient $u \in \mathrm{End}_k(V)$ et $x \in V$ un vecteur propre pour une valeur propre $\lambda \in k$. Alors $Q(u)(x) = Q(\lambda)\, x$ pour tout polynôme $Q \in k[X]$.*

En particulier, si $Q(u) = 0$, alors λ est une racine de Q.

Démonstration. — Nous avons

$$u^2(x) = u(u(x)) = u(\lambda x) = \lambda u(x) = \lambda^2 x$$

et l'on montre ainsi, par récurrence sur i, que $u^i(x) = \lambda^i x$ pour tout $i \in \mathbb{N}^*$ (et aussi pour $i = 0$, avec la convention $u^0 = \mathrm{Id}_V$ et $\lambda^0 = 1$). La conclusion du lemme en découle. □

Nous démontrons maintenant le Lemme des noyaux. C'est l'un des résultats les plus utiles pour la réduction des endomorphismes. Il traduit en algèbre linéaire le théorème de Bézout, démontré en appendice à la fin de ce chapitre.

***Théorème 1.4.5* (Théorème de Bézout).** — *Soient $P_1, \dots, P_r \in k[X]$ des polynômes premiers entre eux. Il existe alors des polynômes $S_1, \dots, S_r \in k[X]$ tels que*

$$P_1 S_1 + \cdots + P_r S_r = 1.$$

Démonstration. — Voir l'appendice B. □

***Lemme 1.4.6* (Lemme des noyaux).** — *Soient $P_1, \dots, P_r \in k[X]$ des polynômes premiers entre eux deux à deux. On pose $P = P_1 \cdots P_r$. Soient V un k-espace vectoriel et $u \in \mathrm{End}_k(V)$.*

1) On a $\ker P(u) = \ker P_1(u) \oplus \cdots \oplus \ker P_r(u)$.

2) Les projections $\ker P(u) \to \ker P_i(u) \subset \ker P(u)$ relatives à cette décomposition en somme directe sont des polynômes en u.

Dans la démonstration qui suit, on notera le fait que le cœur de l'argument se situe au cas $r = 2$. D'ailleurs, le passage de $r = 1$ à $r = 2$ illustre déjà l'argument de récurrence.

Démonstration. — On procède par récurrence sur $r \geq 2$ (lorsque $r = 1$ il n'y a rien à démontrer).

Soit $r = 2$. Le théorème de Bézout entraîne l'existence de polynômes S_1 et S_2 tels que $1 = S_1 P_1 + S_2 P_2$. Si $x \in \ker P_1(u) \cap \ker P_2(u)$, on a $P_1(u)(x) = P_2(u)(x) = 0$, donc

$$x = S_1(u)P_1(u)(x) + S_2(u)P_2(u)(x) = 0.$$

D'autre part $\ker P_i(u)$ est contenu dans $\ker P(u)$ pour $i = 1, 2$. Si $x \in \ker P(u)$ l'on a

$$x = \underbrace{P_1(u)(S_1(u)(x))}_{\in \ker P_2(u)} + \underbrace{P_2(u)(S_2(u)(x))}_{\in \ker P_1(u)},$$

donc $V = \ker P_1(u) + \ker P_2(u)$. Ceci montre 1). De plus, la projection sur $\ker P_1(u)$ est, sur $\ker P(u)$, égale à $(S_2 P_2)(u)$, et celle sur $\ker P_2(u)$ à $(S_1 P_1)(u)$. Ceci prouve 2).

Supposons le résultat vrai au rang r et montrons-le au rang $r+1$. On a $P = Q_1 Q_2$ avec $Q_1 = P_1 \dots P_r$ et $Q_2 = P_{r+1}$. Les polynômes Q_1 et Q_2 sont premiers entre eux puisque P_{r+1} est premier avec chacun des polynômes $P_1, \dots, P_r$. D'après le cas $r = 2$ nous avons $\ker P(u) = \ker Q_1(u) \oplus \ker Q_2(u)$ et les projections sur $\ker Q_1(u)$ et $\ker Q_2(u)$ sont des polynômes en u. Par hypothèse de récurrence on a $\ker Q_1(u) = \ker P_1(u) \oplus \dots \ker P_r(u)$ et les projections sur les facteurs $\ker P_i(u)$, $i = 1, \dots, r$ sont des polynômes en u. On déduit

$$\ker P(u) = \big(\ker P_1(u) \oplus \cdots \oplus \ker P_r(u)\big) \oplus \ker P_{r+1}(u) = \ker P_1(u) \oplus \cdots \oplus \ker P_{r+1}(u).$$

Les projections sont des polynômes en u comme composées de polynômes en u. □

Nous donnons maintenant un critère de diagonalisabilité formulé en termes de polynômes d'endomorphismes. Celui-ci redémontre en particulier le Corollaire 0.3.2.

***Proposition 1.4.7* (Polynômes sans racines multiples).** — *Soit V un k-espace vectoriel de dimension finie. Un endomorphisme $u \in \mathrm{End}_k(V)$ est diagonalisable si et seulement s'il existe $P \in k[X]$ scindé sur k ayant toutes ses racines simples tel que $P(u) = 0$.*

Démonstration. — *Condition nécessaire.* Supposons u diagonalisable. Notons $\lambda_1, \ldots, \lambda_r$ les valeurs propres (distinctes) de u et $V_{\lambda_1}, \ldots, V_{\lambda_r}$ les sous-espaces propres correspondants, de sorte que $V = V_{\lambda_1} \oplus \cdots \oplus V_{\lambda_r}$. Posons

$$P = (X - \lambda_1) \ldots (X - \lambda_r) \in k[X].$$

Le polynôme P est scindé dans k et il a toutes ses racines simples. Nous affirmons que $P(u) = 0$ puisque $P(u)$ s'annule sur une base constituée de vecteurs propres de u. En effet, la diagonalisabilité de u implique l'existence d'une telle base et, si x_ℓ est un vecteur propre correspondant à la valeur propre λ_ℓ, alors

$$P(u)(x_\ell) = \Big(\prod_{i=1}^{r}(u - \lambda_i \,\mathrm{Id})\Big)(x_\ell) = \Big(\prod_{i\neq\ell}(u - \lambda_i \,\mathrm{Id})\Big)(u - \lambda_\ell \,\mathrm{Id})(x_\ell) = 0.$$

Dans la deuxième égalité nous utilisons le fait que les polynômes en u commutent.

Condition suffisante. Soit $P \in k[X]$ scindé sur k avec racines simples tel que $P(u) = 0$. Écrivons $P = c(X-\lambda_1) \ldots (X-\lambda_r)$ avec les $\lambda_i \in k$ distincts et $c \neq 0$. Puisque les polynômes $X - \lambda_i$ sont premiers entre eux deux à deux le Lemme des noyaux 1.4.6 s'applique et nous obtenons

$$(\dagger) \qquad V = \ker P(u) = \ker(u - \lambda_1 \,\mathrm{Id}) \oplus \cdots \oplus \ker(u - \lambda_r \,\mathrm{Id}).$$

Chaque $\ker(u - \lambda_i \,\mathrm{Id})$ qui est non-nul est un sous-espace propre de u. En ne retenant dans $(\dagger)$ que les termes non-nuls on déduit que V est somme directe de sous-espaces propres de u. Ceci équivaut à la diagonalisabilité de u. □

***Corollaire 1.4.8* (Automorphismes d'ordre fini de $\mathbb{C}^n$).** — *Soient V un $\mathbb{C}$-espace vectoriel de dimension n et $u \in \mathrm{End}_{\mathbb{C}}(V)$ tel que $u^d = \mathrm{Id}_V$ pour un entier $d \geq 1$ (dans ce cas, on dit que u* est un automorphisme d'ordre fini*). Alors u est diagonalisable (et ses valeurs propres sont des racines d-èmes de l'unité).* □

1.5. Polynôme minimal

Définition 1.5.1. — Soit V un k-espace vectoriel de dimension finie et $u \in \mathrm{End}_k(V)$. *L'idéal des polynômes qui annulent u* est le sous-espace vectoriel

$$I_u = \{Q \in k[X] \mid Q(u) = 0\}.$$

La terminologie est justifiée par le fait que I_u est bien un idéal de $k[X]$ au sens suivant : si $Q \in I_u$ alors $PQ \in I_u$ pour tout $P \in k[X]$. On renvoie à l'Appendice B pour une discussion plus détaillée.

Il est clair que

$$I_u \subsetneq k[X]$$

puisque $1 \notin I$. Par ailleurs

$$I_u \neq \{0\}.$$

En effet, nous pouvons écrire de façon alternative $I_u = \ker \phi$, avec $\phi : k[X] \to \mathrm{End}_k(V)$, $Q \mapsto Q(u)$ le morphisme de k-algèbres considéré après la Définition 1.4.1. Puisque $k[X]$ est un k-espace vectoriel de dimension infinie et $\mathrm{End}_k(V)$ est un k-espace vectoriel de dimension finie (égale à n^2), il s'ensuit que ϕ ne peut pas être injectif.

Définition et proposition 1.5.2. — *Il existe un unique polynôme unitaire μ_u qui engendre l'idéal I_u, au sens où*

$$I_u = (\mu_u) = \{P\mu_u \mid P \in k[X]\}.$$

Ce polynôme est appelé polynôme minimal de u.

Démonstration. — Ceci est un cas particulier du Théorème B.3 dans l'Appendice B. □

Remarque 1.5.3. — Soient V un k-espace vectoriel de dimension n et $u \in \mathrm{End}_k(V)$. Puisque $\dim \mathrm{End}_k(V) = n^2$, les n^2+1 vecteurs de $\mathrm{End}_k(V)$ donnés par les endomorphismes $\mathrm{Id}_V, u, \ldots, u^{n^2}$ forment nécessairement une famille liée sur k. Autrement dit, il existe un polynôme non nul $Q \in k[X]$ de degré $\leq n^2$ qui annule u. Ceci fournit en particulier une borne sur le degré du polynôme minimal, à savoir $\deg \mu_u \leq n^2$. Nous verrons plus bas que le théorème de Cayley-Hamilton améliore cette borne de façon inattendue à $\deg \mu_u \leq n$.

Proposition 1.5.4. — *Soit V un k-espace vectoriel de dimension finie. Un endomorphisme $u \in \mathrm{End}_k(V)$ est diagonalisable si et seulement si son polynôme minimal $\mu_u \in k[X]$ est scindé sur k avec racines simples.*

Démonstration. — $\Leftarrow$ Si μ_u est scindé sur k avec racines simples, alors u est diagonalisable d'après la proposition 1.4.7 (qui utilise le théorème de Bézout).

$\Rightarrow$ Si u est diagonalisable nous avons vu qu'il existe un polynôme scindé sur k avec racines simples tel que $P(u) = 0$. Donc $P \in I_u$ et μ_u divise P_u. Ceci entraîne que μ_u est scindé sur k avec racines simples. □

Remarque 1.5.5. — Il est en général délicat de calculer le polynôme minimal d'un endomorphisme. Néanmoins, son existence peut souvent être très utile.

1.6. Théorème de Cayley-Hamilton

Dans cette section nous définissons la notion d'espace caractéristique et nous en déduisons une autre sorte de réduction d'un endomorphisme u, à savoir sous forme de matrice diagonale par blocs dont la taille est égale à la multiplicité des valeurs propres. Ceci s'appuie sur un théorème important dit « de Cayley-Hamilton », [(3)] et sur le Lemme des noyaux 1.4.6.

***Théorème 1.6.1* (Théorème de Cayley-Hamilton)**. — *Soit V un $\mathbb{C}$-espace vectoriel de dimension n, et soient $u \in \mathrm{End}_{\mathbb{C}}(V)$ et $P_u(X)$ son polynôme caractéristique. Alors l'endomorphisme $P_u(u)$ est nul, c.-à-d. : « u* est annulé par son polynôme caractéristique *».*

Démonstration. — On a $P_u(X) = (-1)^n \prod_{i=1}^n (X - \lambda_i)$, où $\lambda_1, \ldots, \lambda_n$ sont les n racines (comptées avec multiplicité) de $P_u(X)$ dans $\mathbb{C}$. D'après le théorème 1.3.3, il existe une base $\mathscr{B} = (f_1, \ldots, f_n)$ de V dans laquelle la matrice de u est triangulaire supérieure, avec $\lambda_1, \ldots, \lambda_n$ sur la diagonale. Ceci équivaut à dire que, pour tout $i = 1, \ldots, n$, le sous-espace F_i de V engendré par $f_1, \ldots, f_i$ est stable par u et, plus précisément, que l'on a, pour $i = 1, \ldots, n$:

$$(u - \lambda_i \mathrm{Id}_V)(F_i) \subseteq F_{i-1},$$

[(3)] Le théorème a été observé et démontré pour certaines matrices 4×4 par William Rowan Hamilton (1805–1865) en 1853 et pour des matrices 3×3 par Arthur Cayley (1821–1895) en 1858. Le théorème a été démontré en toute généralité par Ferdinand Georg Frobenius (1849–1917) en 1878. Cf. `https://en.wikipedia.org/wiki/Cayley-Hamilton_theorem`.

avec la convention $F_0 = \{0\}$. Comme $F_n = V$, on déduit des inclusions ci-dessus que $(u - \lambda_n \operatorname{Id}_V)(V) \subseteq F_{n-1}$, puis que $(u - \lambda_{n-1} \operatorname{Id}_V)(u - \lambda_n \operatorname{Id}_V)(V) \subseteq F_{n-2}$, etc., d'où finalement :

$$(u - \lambda_1 \operatorname{Id}_V) \cdots (u - \lambda_n \operatorname{Id}_V)(V) \subseteq F_0 = \{0\}.$$

Ceci montre que $(-1)^n P_u(u) = 0$, ou encore $P_u(u) = 0$. □

***Corollaire 1.6.2* (Cayley-Hamilton pour $k \subseteq \mathbb{C}$).** — *Soient k un sous-corps de $\mathbb{C}$ (par exemple $k = \mathbb{R}$) et V un k-espace vectoriel de dimension n. Soient $u \in \operatorname{End}_k(V)$ et $P_u(X)$ son polynôme caractéristique. Alors $P_u(u) = 0$.*

Démonstration. — Soient $\mathscr{B}$ une base de V et $A = \operatorname{Mat}_{\mathscr{B}}(u) \in M_n(k)$. D'une part, $P_u(X) = P_A(X)$; notons P ce polynôme. D'autre part, comme $k \subseteq \mathbb{C}$, on peut considérer A comme élément de $M_n(\mathbb{C})$, donc d'après le théorème de Cayley-Hamilton on a $P(A) = 0$. Or $P(A)$ est la matrice dans la base $\mathscr{B}$ de $P(u)$, d'où $P(u) = 0$. □

***Remarque 1.6.3*.** — Tout corps k peut être réalisé comme sous-corps d'un corps $\overline{k}$ « algébriquement clos », i.e., tel que tout polynôme de à coefficients dans $\overline{k}$ soit scindé dans $\overline{k}$ (on dit aussi que $\overline{k}$ est une « clôture algébrique de k »). Les preuves du Théorème 1.6.1 et du Corollaire 1.6.2 peuvent être directement adaptées pour montrer que le théorème de Cayley-Hamilton reste valable pour des k-espaces vectoriels de dimension finie avec k un corps quelconque.

***Remarque 1.6.4*.** — Il existe des dizaines (!) de démonstrations différentes du théorème de Cayley-Hamilton, ce qui indique son statut central en algèbre linéaire. C'est un phénomène incontournable, digne d'une étude approfondie.

Dans la suite nous allons nous restreindre pour plus de commodité à des $\mathbb{C}$-espaces vectoriels, mais la remarque précédente montre que la plupart des résultats restent valables pour des coefficients dans un corps k quelconque.

***Corollaire 1.6.5*.** — *Soit $u \in \operatorname{End}_k(V)$. Le polynôme minimal μ_u divise le polynôme caractéristique P_u. Leurs racines sur k sont les mêmes, à savoir les valeurs propres de u.*

Démonstration. — La première affirmation est conséquence de la définition de μ_u et du fait que P_u annule u (théorème de Cayley-Hamilton).

Si λ est valeur propre de u alors $\mu_u(\lambda) = 0$ par le Lemme 1.4.4. Réciproquement, si λ est racine de μ_u alors $P_u(\lambda) = 0$ puisque μ_u divise P_u, donc λ est valeur propre de u. □

***Remarque 1.6.6*.** — Le polynôme minimal joue un rôle essentiel dans la théorie des extensions finies de corps, avec des applications notamment en cryptographie. On parle par exemple de *polynôme minimal d'un nombre algébrique sur* $\mathbb{Q}$: c'est un cas particulier de la notion que nous étudions ici.[(4)]

[(4)] Voir A. Chambert-Loir, *Algèbre corporelle*, Éditions de l'École Polytechnique, 2005. `http://www.cmls.polytechnique.fr/perso/chambert/teach/algebre.pdf` (consultée le 10 février 2020), ou B. Martin, *Codage, cryptologie et applications*, Presses Polytechniques et Universitaires Romandes (PPUR), 2004, chapitre 4 (« Codes cycliques »).

1.7. Espaces caractéristiques

***Définition 1.7.1* (Espaces caractéristiques).** — Soient V un $\mathbb{C}$-espace vectoriel de dimension finie, $u \in \mathrm{End}_{\mathbb{C}}(V)$ et $\lambda \in \mathbb{C}$ une valeur propre de u dont la multiplicité algébrique (i.e., sa multiplicité comme racine de $P_u(X)$) vaut m. On pose

$$V_{(\lambda)} = \mathrm{Ker}(u - \lambda \,\mathrm{Id}_V)^m$$

et on l'appelle *l'espace caractéristique associé à* λ.

Remarque 1.7.2. — 1. Tout espace caractéristique $V_{(\lambda)}$ est stable par u. En effet, $V_{(\lambda)}$ est stable par $u - \lambda \,\mathrm{Id}_V$ donc aussi par $u = (u - \lambda \,\mathrm{Id}_V) + \lambda \,\mathrm{Id}_V$.

2. L'espace caractéristique $V_{(\lambda)}$ contient l'espace propre $V_\lambda = \mathrm{Ker}(u - \lambda \,\mathrm{Id}_V)$. L'inclusion est stricte en général.

Exemple 1.7.3. — Si u est l'endomorphisme de $\mathbb{C}^2$ de matrice $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ (i.e. , $u(e_1) = 0$ et $u(e_2) = e_1$), alors $P_u(X) = X^2$ a 0 comme unique racine. Par un calcul direct (ou encore d'après tle théorème de Cayley-Hamilton) on a $u^2 = 0$, donc $V_{(0)} = \mathbb{C}^2$ tandis que $V_0 = \mathrm{Ker}(u) = \mathbb{C}e_1$.

***Théorème 1.7.4* (Décomposition en espaces caractéristiques)**

Soient V un $\mathbb{C}$-espace vectoriel de dimension n et $u \in \mathrm{End}_{\mathbb{C}}(V)$. Écrivons $P_u(X) = (-1)^n \prod_{i=1}^r (X - \lambda_i)^{m_i}$, où $\lambda_1, \ldots, \lambda_r$ sont les valeurs propres, deux à deux distinctes, de u et m_i est la multiplicité algébrique de λ_i. Alors :

(1) *V est la somme directe des espaces caractéristiques $V_{(\lambda_1)}$, …, $V_{(\lambda_r)}$ et les projections $V \to V_{(\lambda_i)}$ relatives à cette décomposition sont des polynômes en u.*

(2) *On a $\dim V_{(\lambda_i)} = m_i$ pour tout i.*

Démonstration. — Pour tout i, posons

$$P_i = (X - \lambda_i)^{m_i},$$

de sorte que

$$V_{(\lambda_i)} = \ker P_i(u).$$

Les polynômes P_i sont clairement premiers deux à deux et le Lemme des noyaux 1.4.6 s'applique. On en déduit que

$$\ker(P_1 \ldots P_r)(u) = V_{(\lambda_1)} \oplus \cdots \oplus V_{(\lambda_r)}.$$

Or $P_1 \ldots P_r = (-1)^n P_u$ et donc, par le théorème de Cayley-Hamilton, on a

$$\ker(P_1 \ldots P_r)(u) = \ker P_u(u) = V.$$

Par ailleurs, il découle aussi du lemme des noyaux que les projections $V \to V_{(\lambda_i)}$ sont des polynômes en u. Ceci démontre (1).

Pour tout $i = 1, \ldots, r$ posons $d_i = \dim V_{(\lambda_i)}$, choisissons une base $\mathscr{B}_i$ de $V_{(\lambda_i)}$ et notons $u_i = u|_{V_{(\lambda_i)}}$. Alors $\mathscr{B} = \mathscr{B}_1 \cup \cdots \cup \mathscr{B}_r$ est une base de V et, notant $A_i = \mathrm{Mat}_{\mathscr{B}_i}(u_i)$, on a :

$$(\dagger) \qquad \mathrm{Mat}_{\mathscr{B}}(u) = \left(\begin{array}{c|c|c|c} A_1 & 0 & \cdots & 0 \\ \hline 0 & A_2 & \ddots & \vdots \\ \hline \vdots & \ddots & \ddots & 0 \\ \hline 0 & \cdots & 0 & A_r \end{array}\right), \qquad \text{d'où} \qquad P_u(X) = \prod_{i=1}^r P_{u_i}(X).$$

Montrons que λ_i est la seule valeur propre de u_i. À cet effet, soit μ une valeur propre quelconque de u_i et $x \in V_{(\lambda_i)}$ un vecteur propre associé. Alors $(u_i - \lambda_i \operatorname{Id}_V)(x) = (\mu - \lambda_i)x$, d'où $(u_i - \lambda_i \operatorname{Id}_V)^p(x) = (\mu - \lambda_i)^p x$ pour tout $p \in \mathbb{N}^*$. Or $(u_i - \lambda_i \operatorname{Id}_V)^{m_i}(x) = 0$ par définition de $V_{(\lambda_i)}$, ce qui entraîne $(\mu - \lambda_i)^{m_i} x = 0$ et donc $\mu = \lambda_i$ puisque $x \neq 0$.

On en déduit que $P_{u_i}(X) = (-1)^{d_i}(X - \lambda_i)^{d_i}$ et, d'après (†), on obtient $P_u(X) = (-1)^n \prod_{i=1}^r (X - \lambda_i)^{d_i}$. Ceci entraîne $d_i = m_i$ pour tout i et démontre (2). □

Remarque 1.7.5. — Le Théorème de décomposition en espaces caractéristiques 1.7.4 est effectif, puisque les sous-espaces caractéristiques peuvent être calculés explicitement pourvu que les valeurs propres soient connues. [(5)] Il fournit une décomposition de V en une somme directe de sous-espaces stables de dimensions égales à la multiplicité algébrique des valeurs propres. En termes de matrices, il fournit une réduction de toute matrice $A \in M_n(\mathbb{C})$ à une matrice diagonale par blocs comme dans (†) dont la taille est égale à la multiplicité algébrique de chaque valeur propre. Cette décomposition est revisitée dans la section suivante en lien avec la trigonalisation, et dans le chapitre suivant en lien avec la décomposition de Dunford.

1.8. Suite des noyaux et algorithme de trigonalisation – version 2.0

Nous nous appuyons sur la décomposition en espaces caractéristiques pour donner un algorithme de trigonalisation plus efficace que celui de la section 1.3.

Définition et proposition 1.8.1 **(Suite des noyaux et indice)**

Soit $u \in \operatorname{End}_k(V)$ avec $\dim V$ finie. Il existe un unique entier $s \in \mathbb{N}$ tel que

$$\{0\} = \ker u^0 \subsetneq \ker u \subsetneq \ker u^2 \subsetneq \cdots \subsetneq \ker u^s = \ker u^{s+1} = \ker u^{s+2} = \ldots$$

L'entier s s'appelle indice de u. *C'est le plus petit entier naturel p tel que* $\ker u^p = \ker u^{p+1}$.

Démonstration. — On a $\ker u^p \subset \ker u^{p+1}$ pour $p \in \mathbb{N}$. La suite croissante $(\dim \ker u^p)_{p\in\mathbb{N}}$ prend ses valeurs dans l'ensemble fini $\{0, \ldots, \dim V\}$, elle est donc stationnaire à partir d'un certain rang. Soit $s = \min\{p \in \mathbb{N} \mid \dim \ker u^p = \dim \ker u^{p+1}\} \in \mathbb{N}$. Par définition l'on a $\ker u^p \subsetneq \ker u^{p+1}$ pour tout $p < s$ et $\ker u^s = \ker u^{s+1}$. Montrons l'égalité $\ker u^p = \ker u^{p+1}$ pour tout $p > s$. L'inclusion $\ker u^p \subset \ker u^{p+1}$ a déjà été discutée. Pour prouver l'inclusion inverse $\ker u^{p+1} \subset \ker u^p$, on considère un vecteur $x \in \ker u^{p+1}$ et on montre que $u^p(x) = 0$. Or l'égalité

$$u^{p+1}(x) = u^{s+1}(u^{p-s}(x)) = 0$$

entraîne bien

$$u^s(u^{p-s}(x)) = u^p(x) = 0$$

puisque $\ker u^s = \ker u^{s+1}$. □

Corollaire 1.8.2. — *(i) Soient V un $\mathbb{C}$-espace vectoriel de dimension finie, $u \in \operatorname{End}_{\mathbb{C}}(V)$ et $\lambda \in \mathbb{C}$ une valeur propre de u de multiplicité algébrique m. La suite des noyaux de $(u -$*

(5) Étant donné un polynôme général de degré ≥ 5, il n'existe pas de formule « universelle » pour trouver ses racines qui soit analogue de la bien-connue « $x_{1,2} = (-b \pm \sqrt{\Delta})/2a$ avec $\Delta = b^2 - 4ac$ » (Le lecteur intéressé par le sujet devra suivre un cours de théorie de Galois, ou bien consulter un livre tel A. Chambert-Loir, *loc. cit.* De telles formules existent par ailleurs en degrés 3 et 4.) Dans cette généralité, la discussion faite ici reste purement théorique. En pratique, on utilise des méthodes d'algèbre linéaire pour approximer les racines des polynômes, et en particulier pour calculer des valeurs propres de matrices.

$\lambda \operatorname{Id}_V)^p$, $p \in \mathbb{N}$, *est strictement croissante avant d'être stationnaire ; elle a comme premier terme non-nul l'espace propre* V_λ *et comme terme stationnaire l'espace caractéristique* $V_{(\lambda)}$ *:*

$$\{0\} \subsetneq V_\lambda = \ker(u - \lambda \operatorname{Id}_V) \subsetneq \cdots \subsetneq V_{(\lambda)} = \ker(u - \lambda \operatorname{Id}_V)^s = \cdots = \ker(u - \lambda \operatorname{Id}_V)^m.$$

(ii) Soit $\mathscr{B}$ *une base de* $V_{(\lambda)} = \ker(u - \lambda \operatorname{Id}_V)^m$ *obtenue en considérant d'abord une base de* $V_\lambda = \ker(u - \lambda \operatorname{Id}_V)$*, que l'on complète ensuite en une base de* $\ker(u - \lambda \operatorname{Id}_V)^2$*, que l'on complète ensuite en une base de* $\ker(u - \lambda \operatorname{Id}_V)^3$ *etc. La matrice de* $u|_{V_{(\lambda)}}$ *dans la base* $\mathscr{B}$ *est triangulaire supérieure.*

Démonstration. — (i) La conclusion est équivalente à montrer que l'indice s de $u - \lambda \operatorname{Id}_V$ est au plus égal à m. Soient $\lambda_1, \ldots, \lambda_r$ les valeurs propres de u distinctes deux à deux, avec $\lambda_1 = \lambda$. Considérons la décomposition en espaces caractéristiques $V = V_{(\lambda_1 = \lambda)} \oplus V_{(\lambda_2)} \oplus \cdots \oplus V_{(\lambda_r)}$. Chaque espace caractéristique est stable par u et donc par $u - \lambda \operatorname{Id}_V$. Puisque $u - \lambda \operatorname{Id}_V$ est inversible sur chaque $V_{(\lambda_i)}$ avec $i \geq 2$, on déduit $\ker(u - \lambda \operatorname{Id}_V)^s \subset V_{(\lambda)} = \ker(u - \lambda \operatorname{Id}_V)^m$, donc $s \leq m$ par définition de l'indice s.

(ii) Notons $E_i = \ker(u - \lambda \operatorname{Id}_V)^i$ pour $i = 0, \ldots, s$, de sorte que $\{0\} = E_0 \subsetneq E_1 \subsetneq \cdots \subsetneq E_s$. La conclusion équivaut à montrer que $u(E_i) \subseteq E_i$ pour tout i. Or

$$(u - \lambda \operatorname{Id}_V) E_i \subseteq E_{i-1} \quad \text{pour tout} \quad i \geq 1.$$

En effet, si $x \in E_i$ alors $(u - \lambda \operatorname{Id}_V)^{i-1}(u - \lambda \operatorname{Id}_V)(x) = (u - \lambda \operatorname{Id}_V)^i(x) = 0$. On en déduit que $u(E_i) \subseteq \lambda E_i + E_{i-1} \subseteq E_i$. □

Ce résultat fournit l'agorithme de trigonalisation suivant. L'algorithme reçoit en entrée un espace vectoriel V et un endomorphisme $u \in \operatorname{End}_{\mathbb{C}}(V)$ dont on connaît les valeurs propres $\lambda_1, \ldots, \lambda_r$ et leurs multiplicités respectives $m_1, \ldots, m_r$. Il fournit en sortie une base $\mathscr{B}$ de V dans laquelle la matrice de u est triangulaire supérieure.

***Algorithme 2* (Algorithme de trigonalisation – version 2.0)**

Pour chaque $i = 1, \ldots, r$ *:*

On calcule une base $\mathscr{B}_i$ *de* $V_{(\lambda_i)} = \ker(u - \lambda_i \operatorname{Id}_V)^{m_i}$ *en calculant d'abord une base de* $V_{\lambda_i} = \ker u - \lambda_i \operatorname{Id}_V$*, en la complétant en une base de* $\ker(u - \lambda_i \operatorname{Id}_V)^2$*, en complétant celle-ci en une base de* $\ker(u - \lambda_i \operatorname{Id}_V)^3$ *etc., jusqu'à arriver à une base de* $\ker(u - \lambda_i \operatorname{Id}_V)^{m_i}$*.*

La base $\mathscr{B}$ *recherchée est* $\mathscr{B} = \mathscr{B}_1 \sqcup \cdots \sqcup \mathscr{B}_r$*.*

Exemple 1.8.3. — Voici comment fonctionne l'algorithme sur la matrice considérée dans l'exemple 1.3.6, à savoir

$$A = \begin{pmatrix} -2 & 3 & 3 \\ -2 & 1 & 2 \\ -3 & 3 & 4 \end{pmatrix} \in M_3(\mathbb{R}).$$

Le polynôme caractéristique est $P_A(X) = -(X-1)^3$, il est scindé sur $\mathbb{R}$ et la matrice est trigonalisable. Elle possède une unique valeur propre $\lambda = 1$ de multiplicité algébrique 3. On a

$$A - I_3 = \begin{pmatrix} -3 & 3 & 3 \\ -2 & 0 & 2 \\ -3 & 3 & 3 \end{pmatrix}.$$

Nous calculons une base de l'espace propre $V_1 = \ker(A - I_3)$ par l'algorithme du pivot de Gauss sur colonnes, en rajoutant la première colonne à la deuxième et à la troisième. La matrice devient échelonnée et nous trouvons $V_1 = \operatorname{Vect}(e_1 + e_3)$.

On a

$$(A - I_3)^2 = \begin{pmatrix} -6 & 0 & 6 \\ 0 & 0 & 0 \\ -6 & 0 & 6 \end{pmatrix}$$

et nous calculons une base de $\ker(A-I_3)^2$ par l'algorithme du pivot de Gauss sur colonnes en rajoutant la première colonne à la troisième. La matrice devient échelonnée et nous trouvons $\ker(A - I_3)^2 = \mathrm{Vect}(e_1 + e_3, e_2)$.

Enfin, nous savons que $V_{(1)} = \mathbb{R}^3 = \ker(A - I_3)^3$, ou encore $(A - I_3)^3 = 0$. (Ceci peut bien sûr être vérifié directement.) Nous complétons la base $(e_1 + e_3, e_2)$ de $\ker(A - I_3)^2$ en une base de $\mathbb{R}^3$ en lui adjoignant par exemple le vecteur e_3 (mais nous pourrions aussi lui adjoindre e_1, ou e_2+e_3). En conclusion, nous trouvons que la matrice A devient triangulaire supérieure dans la base $(e_1 + e_3, e_2, e_3)$ de $\mathbb{R}^3$.

Vérifions la réponse : la matrice

$$\widetilde{A} = P^{-1}AP = \begin{pmatrix} 1 & 3 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}, \quad \text{avec} \quad P = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix},$$

est bien triangulaire supérieure.

A. Appendice (†) : somme directe externe d'espaces vectoriels

***Définition A.1* (Sommes directes).** — Soient $V_1, \ldots, V_n$ des k-espaces vectoriels. L'ensemble produit

$$V_1 \times \cdots \times V_n = \{(v_1, \ldots, v_n) \mid v_i \in V_i\}$$

est muni d'une structure d'espace vectoriel définie « composante par composante », c.-à-d., $t \cdot (v_1, \ldots, v_n) = (t \cdot v_1, \ldots, t \cdot v_n)$ et $(v_1, \ldots, v_n) + (v'_1, \ldots, v'_n) = (v_1 + v'_1, \ldots, v_n + v'_n)$.

On l'appelle la *somme directe* (externe) des V_i et on le note

$$V_1 \oplus \cdots \oplus V_n \qquad \text{ou} \qquad \bigoplus_{i=1}^{n} V_i.$$

De même, un n-uplet $(v_1, \ldots, v_n)$ (avec $v_i \in V_i$) est aussi noté $v_1 + \cdots + v_n$ ou $\sum_{i=1}^n v_i$, c.-à-d., on identifie l'élément $v_i \in V_i$ au n-uplet $(0, \ldots, 0, v_i, 0, \ldots, 0)$ (où v_i est à la i-ème place).

Une base de $V_1 \oplus \cdots \oplus V_n$ est obtenue en adjoignant des bases de $V_1, \ldots, V_n$, de sorte que

$$\dim \bigoplus_{i=1}^{n} V_i = \sum_{i=1}^{n} \dim V_i.$$

Remarque A.2. — Supposons maintenant que $E_1, \ldots, E_n$ soient des sous-espaces d'un k-espace vectoriel V. D'une part, on note $E_1 + \cdots + E_n$ le sous-espace de V engendré par $E_1 \cup \cdots \cup E_n$, défini comme étant l'ensemble de toutes les sommes

$$x_1 + \cdots + x_n, \qquad \text{avec } x_i \in E_i.$$

D'autre part, on peut former, la somme directe externe $S = E_1 \oplus \cdots \oplus E_n$ des E_i ; ce n'est pas un sous-espace de V, mais on a une application linéaire naturelle

$$\sigma : E_1 \oplus \cdots \oplus E_n \to V, \qquad (x_1, \ldots, x_n) \mapsto x_1 + \cdots + x_n$$

dont l'image est le sous-espace $E_1 + \cdots + E_n$ de V, et le noyau est le sous-espace de S formé des n-uplets $(x_1, \ldots, x_n)$ tels que $x_1 + \cdots + x_n = 0$.

On voit donc que $\operatorname{Ker}\sigma = \{0\}$ si et seulement si les sous-espaces $E_1, \dots, E_n$ sont en somme directe dans V, et dans ce cas σ est un isomorphisme de la somme directe externe S sur le sous-espace de V noté $E_1 \oplus \cdots \oplus E_n$ en 0.1.1. Ceci justifie l'usage de la notation $E_1 \oplus \cdots \oplus E_n$ dans les deux cas. Pour des espaces vectoriels arbitraires $E_1, \dots, E_n$, la somme directe « externe » $E_1 \oplus \cdots \oplus E_n$ sera appelée simplement « somme directe ».

B. Appendice (†) : division euclidienne dans $\mathbb{C}[X]$ et théorème de Bézout

Théorème B.1 **(Division euclidienne dans $k[X]$).** — *Soit k un corps et soit $P \in k[X]$ un polynôme de degré $d \geq 1$. Pour tout $F \in k[X]$, il existe un unique couple (Q, R) d'éléments de $k[X]$ tel que*

$$F = PQ + R, \qquad \text{et} \quad R = 0 \text{ ou bien } \deg(R) < \deg(P).$$

On appelle Q (resp. R) le quotient (resp. le reste) de la division euclidienne de F par P.

Démonstration. — Montrons l'existence de Q, R en procédant par récurrence sur $\deg(F)$. Si $F = 0$ ou si $\deg(F) < d = \deg(P)$, on prend $Q = 0$ et $R = F$. Soit $n \geq d$ et supposons l'existence établie pour les degrés $< n$. Soit F de degré n. Notons a le coefficient dominant de F et c celui de P. Alors $ac^{-1}X^{n-d}P$ est de degré n et de coefficient dominant a, donc $F - ac^{-1}X^{n-d}P$ est de degré $< n$. Par hypothèse de récurrence, il existe $Q, R \in k[X]$ tels que

$$F - ac^{-1}X^{n-d}P = PQ + R, \qquad \text{et} \quad R = 0 \text{ ou bien } \deg(R) < \deg(P).$$

Alors $F = P(Q + ac^{-1}X^{n-d}) + R$, ce qui prouve le résultat d'existence.

Montrons l'unicité : si Q_1, R_1 vérifient les mêmes conditions, les égalités $PQ + R = F = PQ_1 + R_1$ donnent

$$P(Q - Q_1) = R_1 - R.$$

Si $Q - Q_1$ était $\neq 0$ alors $P(Q - Q_1)$ serait de degré $d + \deg(Q - Q_1) \geq d$. Or, $R_1 - R$ est nul ou de degré $< d$. Donc nécessairement $Q - Q_1 = 0$, d'où $R_1 - R = 0$, d'où $Q = Q_1$ et $R = R_1$. Ceci prouve l'unicité. □

Définitions B.2. — 1) Soit I un sous-ensemble de $k[X]$. On dit que I est un *idéal* de $k[X]$ si c'est un sous-espace vectoriel et si, pour tout $P \in I$ et $S \in k[X]$, on a $SP \in I$.

2) L'intersection de deux idéaux est un idéal. Étant donnés des polynômes $P_1, \dots, P_r \in k[X]$, l'idéal

$$(P_1, \dots, P_r) = \bigcap_{P_1, \dots, P_r \in I \text{ idéal}} I$$

est appelé *l'idéal engendré par* $P_1, \dots, P_r$. C'est l'idéal le plus petit, pour l'ordre partiel donné par l'inclusion, qui contient $P_1, \dots, P_r$. L'on démontre l'égalité

$$(P_1, \dots, P_r) = \{S_1P_1 + \cdots + S_rP_r \mid S_1, \dots, S_r \in k[X]\}.$$

3) On dit qu'un idéal I de $k[X]$ est *principal* s'il peut être engendré par un seul élément, c.-à-d., s'il existe $P \in I$ tel que $I = \{SP \mid S \in k[X]\} = (P)$.

Théorème B.3. — *Soit k un corps. Tout idéal I de $k[X]$ est principal. Plus précisément, si I est un idéal non nul de $k[X]$, il existe un unique polynôme unitaire $P \in I$ tel que $I = (P)$.*

Démonstration. — Si I est l'idéal nul $\{0\}$, il est engendré par le polynôme nul 0. Donc on peut supposer $I \neq \{0\}$. Dans ce cas, l'ensemble $\{\deg(Q) \mid Q \in I - \{0\}\}$ est un sous-ensemble non-vide de $\mathbb{N}$, donc admet un plus petit élément d. Soit $P \in I - \{0\}$ tel que $\deg(P) = d$; quitte à remplacer P par $a^{-1}P$, où a est le coefficient dominant de P, on peut supposer P unitaire.

Soit F un élément arbitraire de I, d'après le théorème B.1, on peut écrire $F = PQ + R$, avec $R = 0$ ou bien $\deg(R) < \deg(P) = d$. Comme I est un idéal, alors $PQ \in I$ et donc $R = F - PQ$ appartient à I. Si on avait $R \neq 0$, ce serait un élément non nul de I de degré $< d$, contredisant la minimalité de d. Donc $R = 0$ et donc $F = PQ$. Il en résulte que $I = \{PQ \mid Q \in k[X]\} = (P)$, i.e. I est principal, engendré par le polynôme unitaire P. De plus, P est unique. En effet, si P_1 est un second polynôme unitaire tel que $I = (P_1)$, alors il existe $Q, Q_1 \in k[X]$ tels que $P_1 = PQ$ et $P = P_1Q_1$. Il en résulte que Q et Q_1 sont de degré zéro, donc des éléments de k, et comme P et P_1 sont unitaires, l'égalité $P_1 = PQ$ entraîne $Q = 1$ d'où $P_1 = P$. □

***Théorème B.4* (Théorème de Bézout sur $\mathbb{C}$).** — *Soient $P_1, \ldots, P_r \in \mathbb{C}[X]$ des polynômes non nuls, sans racine commune. Alors il existe $S_1, \ldots, S_r \in \mathbb{C}[X]$ tels que $S_1P_1 + \cdots + S_rP_r = 1$.*

Démonstration. — Soit $I = (P_1, \ldots, P_r) \subset \mathbb{C}[X]$ l'idéal engendré par $P_1, \ldots, P_r$ et $D \in \mathbb{C}[X]$ son unique générateur unitaire. Puisque les polynômes $P_1, \ldots, P_r$ n'ont pas de racine commune et $\mathbb{C}$ est algébriquement clos, il en découle que $\deg D = 0$, ou encore $D = 1$. En effet, dans le cas contraire on aurait $\deg D > 0$, donc D aurait au moins une racine dans $\mathbb{C}$, qui serait aussi racine commune pour $P_1, \ldots, P_r$ puisque D divise chacun des P_i.

Or l'idéal $(P_1, \ldots, P_r)$ est l'ensemble des sommes $S_1P_1 + \cdots + S_rP_r$, avec $S_1, \ldots S_r \in \mathbb{C}[X]$. En particulier $D = 1$ peut être exprimé comme une telle somme, ce qui finit la preuve. □

Le théorème de Bézout se généralise au cas des polynômes à coefficients dans un corps quelconque (en particulier $k = \mathbb{R}$) de la manière suivante.

***Définition B.5*.** — Les polynômes $P_1, \ldots, P_r \in k[X]$ sont dits *premiers entre eux* si tout polynôme D qui les divise simultanément est nécessairement de degré 0, ou encore s'il n'existe pas de polynôme de degré ≥ 1 qui les divise simultanément.

La démonstration du Théorème B.4 s'adapte directement pour démontrer le

***Théorème B.6* (Théorème de Bézout).** — *Soient $P_1, \ldots, P_r \in k[X]$ des polynômes non nuls premiers entre eux. Il existe $S_1, \ldots, S_r \in k[X]$ tels que $S_1P_1 + \cdots + S_rP_r = 1$.* □

***Remarque B.7*.** — Soit k un corps algébriquement clos (par exemple $k = \mathbb{C}$). Les polynômes $P_1, \ldots, P_r \in k[X]$ sont premiers entre eux si et seulement s'ils n'ont pas de racine commune. De façon équivalente, il existe $D \in k[X]$ avec $\deg D \geq 1$ qui divise simultanément les polynômes $P_1, \ldots, P_r$ si et seulement si les P_i ont une racine commune dans k. L'implication directe découle du fait qu'un tel polynôme D a nécessairement une racine dans k, et celle-ci est aussi une racine commune des P_i. L'implication inverse découle du fait que, si $\alpha \in k$ est racine commune des P_i, alors $X - \alpha$ les divise simultanément.

C. Appendice (†) : $\mathbb{C}$ est algébriquement clos

***Théorème C.1*.** — *$\mathbb{C}$ est algébriquement clos, c.-à-d., tout polynôme $P \in \mathbb{C}[X]$ non constant admet une racine dans $\mathbb{C}$.*

Démonstration. — Soit $P \in \mathbb{C}[X]$ un polynôme de degré $n \geq 1$. Sans perte de généralité, on peut supposer P *unitaire*, i.e. de coefficient dominant égal à 1. Écrivons

$$P = a_n X^n + a_{n-1} X^{n-1} + \cdots + a_0$$

avec $a_n \neq 0$. Raisonnons par l'absurde et supposons que P ne s'annule pas sur $\mathbb{C}$. Alors, en particulier, $a_0 \neq 0$. Notons $|\cdot|$ la norme usuelle sur $\mathbb{C}$, c.-à-d., $|z| = \sqrt{z\overline{z}}$.

Montrons d'abord que la fonction continue $f : \mathbb{C} \to \mathbb{R}_+$, $z \mapsto |P(z)|$ atteint son minimum $r_0 > 0$ sur $\mathbb{C}$. Rappelons que $\lim_{|z|\to+\infty} |P(z)| = +\infty$, de sorte qu'il existe $R > 0$ tel que

$$|z| \geq R \Longrightarrow |P(z)| \geq |a_0|. \tag{1}$$

Comme le disque fermé D de centre 0 et rayon R est compact, la fonction f y atteint son minimum r_0, et $r_0 > 0$ puisqu'on a supposé que P ne s'annule pas. Puisque pour tout $z \notin D$ on a $f(z) = |P(z)| \geq |a_0| = |P(0)| \geq r_0$, alors r_0 est le minimum de f sur $\mathbb{C}$.

Soit $z_0 \in D$ tel que $f(z_0) = r_0$. En remplaçant P par le polynôme de même degré $Q(X) = P(z_0)^{-1} P(X + z_0)$, on se ramène au cas où $z_0 = 0$ et où $P(0) = 1$ est le minimum de f sur $\mathbb{C}$.

Notons k l'ordre d'annulation en 0 de $P - 1$. On peut alors écrire

$$P(X) = 1 + a_k X^k + \cdots + a_n X^n$$

avec $1 \leq k < n$ et $a_k, a_n \neq 0$. (Le cas où $k = n$ est exclu puisque le polynôme $1 + a_n X^n$ possède des racines complexes.) Écrivons $a_k = re^{i\theta}$ avec $r > 0$ et $\theta \in [0, 2\pi[$. Posons $z_\varepsilon = \varepsilon e^{i(\pi-\theta)/k}$ pour tout $\varepsilon \in \mathbb{R}_+^*$. Comme $z_\varepsilon^k = \varepsilon^k e^{i(\pi-\theta)} = -\varepsilon^k e^{-i\theta}$, alors

$$P(z_\varepsilon) = 1 - r\varepsilon^k + \varepsilon^k h(\varepsilon), \qquad \text{où} \qquad h(\varepsilon) = \sum_{j=1}^{n-k} a_{k+j} z_\varepsilon^j.$$

Comme $\lim_{\varepsilon\to 0} \varepsilon^k = 0$ et $\lim_{\varepsilon\to 0} h(\varepsilon) = 0$, il existe $\varepsilon_0 \in]0, 1[$ tel que

$$\forall \varepsilon \leq \varepsilon_0, \qquad r\varepsilon^k < 1 \quad \text{et} \quad |h(\varepsilon)| \leq \frac{r}{2}.$$

On a alors

$$|P(z_{\varepsilon_0})| = |1 - r\varepsilon_0^k + \varepsilon_0^k h(\varepsilon_0)| \leq |1 - r\varepsilon_0^k| + \frac{r}{2}\varepsilon_0^k = 1 - r\varepsilon_0^k + \frac{r}{2}\varepsilon_0^k = 1 - \frac{r}{2}\varepsilon_0^k < 1.$$

Ceci contredit l'hypothèse que $1 = P(0)$ était le minimum de $f = |P|$ sur $\mathbb{C}$. Cette contradiction montre que l'hypothèse que P ne s'annule pas sur $\mathbb{C}$ est impossible. Ceci achève la démonstration du théorème C.1. □

Corollaire C.2. — *Tout polynôme $P \in \mathbb{C}[X]$ de degré $n \geq 1$ se factorise en produit de facteurs de degré 1, i.e.*

$$P = a(X - \lambda_1) \ldots (X - \lambda_n),$$

où $a \in \mathbb{C}^*$ est le coefficient dominant de P.

Démonstration. — Nous procédons par récurrence sur $n \geq 1$. L'affirmation est évidente pour $n = 1$, nous pouvons donc supposer $n \geq 2$ et le résultat établi pour $n - 1$. Soit P de degré n et de coefficient dominant a. D'après le Théorème C.1 le polynôme P admet au moins une racine λ_1 dans $\mathbb{C}$. Faisant la division euclidienne de P par $X - \lambda_1$, on peut écrire

$$P = (X - \lambda_1)P_1 + R, \qquad \text{avec} \qquad R = 0 \text{ ou bien } \deg(R) < 1.$$

Donc $R = 0$ ou bien R est une constante $c \neq 0$. Or, évaluant l'égalité ci-dessus en $X = \lambda_1$, on trouve $R(\lambda_1) = P(\lambda_1) = 0$, donc nécessairement $R = 0$. Ainsi $P = (X - \lambda_1)P_1$, avec P_1 non nul, de degré $n-1$ et de coefficient dominant a. Par hypothèse de récurrence, P_1 se factorise en $P_1 = a(X - \lambda_2) \ldots (X - \lambda_n)$, et donc $P = (X - \lambda_1)P_1$ égale $a(X - \lambda_1) \ldots (X - \lambda_n)$. □