Chapitre I

Réduction des endomorphismes

1. Valeurs propres, vecteurs propres

1.1. Définition. Soit E un espace vectoriel sur le corps $\mathbf{K} = \mathbf{R}$ ou $\mathbf{C}$, et u un endomorphisme de E, c'est-à-dire une application linéaire de E dans lui-même.

1. Un scalaire $\lambda \in \mathbf{K}$ est dit *valeur propre* de u s'il existe un élément non nul x de E tel que $u(x) = \lambda x$. On dit alors que x est *vecteur propre* de u associé à la valeur propre λ.

2. Soit λ une valeur propre de u. L'ensemble des $x \in E$ tels que $u(x) = \lambda x$ est un sous-espace vectoriel de E, appelé *sous-espace propre* de u associé à la valeur propre λ.

1.2. Proposition. *Soit E un espace vectoriel sur le corps $\mathbf{K} = \mathbf{R}$ ou $\mathbf{C}$, u un endomorphisme de E, $\lambda_1, \dots, \lambda_p$ des valeurs propres deux à deux distinctes de u. Les sous-espaces propres $F_1, \dots, F_p$ qui leur sont associés sont en somme directe, ce qui signifie que*

$$x_1 + \cdots + x_p = 0\,, \qquad \textit{avec} \quad x_1 \in F_1\,, \dots x_p \in F_p\,,$$

implique

$$x_1 = 0\ \dots, x_p = 0\,.$$

Démonstration. La propriété est vraie pour $p = 1$. Supposons-la vraie pour $p \leq q - 1$, et montrons qu'elle est vraie pour $p = q$. On peut écrire

$$x_1 = -x_2 - \cdots - x_q\,, \tag{$*$}$$

d'où, en appliquant u aux deux membres,

$$u(x_1) = \lambda_1 x_1 = -\lambda_2 x_2 - \cdots - \lambda_q x_q\,,$$

et par suite, compte tenu de $(*)$,

$$(\lambda_2 - \lambda_1)x_2 + \cdots + (\lambda_q - \lambda_1)x_q = 0\,.$$

Mais d'après l'hypothèse de récurrence, ceci implique

$$(\lambda_2 - \lambda_1)x_2 = 0\,, \dots, (\lambda_q - \lambda_1)x_q = 0\,,$$

ou, puisque les valeurs propres sont supposées deux à deux distinctes,

$$x_2 = 0\,, \dots, x_q = 0\,.$$

Mais alors, compte tenu de $(*)$, on a aussi

$$x_1 = 0\,. \qquad \square$$

1.3. Définition. Soit E un espace vectoriel de dimension finie sur le corps $\mathbf{K} = \mathbf{R}$ ou $\mathbf{C}$, et u un endomorphisme de E. On appelle *polynôme caractéristique de* u le polynôme, à une indéterminée X,

$$P_u(X) = \det(u - X\,\mathrm{id}_E)\,,$$

où id_E désigne l'application identique de E dans lui-même.

1.4. Théorème. *Avec les hypothèses et notations de 1.3, un scalaire $\lambda \in \mathbf{K}$ est valeur propre de u si et seulement si $P_u(\lambda) = 0$.*

En effet, puisque E est de dimension finie, $u - \lambda\,\mathrm{id}_E$ n'est pas un isomorphisme si et seulement si son noyau est non nul, c'est-à-dire si et seulement si λ est valeur propre de u.

1.5. Remarques.

1. Le degré du polynôme caractéristique $P_u(X)$ est égal à la dimension de l'espace vectoriel E.

2. Si $\mathbf{K} = \mathbf{C}$, on sait, d'après le théorème de d'Alembert, que $P_u(X)$ possède au moins une racine, donc que u possède au moins une valeur propre. Par contre, un endomorphisme d'un espace vectoriel réel peut n'avoir aucune valeur propre. Il en est ainsi par exemple de l'endomorphisme de $\mathbf{R}^2$:

$$(x, y) \mapsto (X, Y)\,, \qquad \text{avec} \quad X = y\,, \quad Y = -x\,.$$

1.6. Proposition. *Les hypothèses et notations sont celles de 1.3 et, de plus, on suppose $\mathbf{K} = \mathbf{C}$. On note n la dimension de E et $\lambda_1, \dots, \lambda_n$ les racines de $P_u(X)$, rangées dans un ordre quelconque, chacune d'elles apparaissant un nombre de fois égal à sa multiplicité. Il existe une base $(e_1, \dots, e_n)$ de E telle que la matrice M de u dans cette base ait les propriétés suivantes :*

(i) *la matrice M est triangulaire inférieure, ce qui exprime que pour tout i, $1 \le i \le n$, les composantes de $u(e_i)$ sur les vecteurs de base e_j, avec $j < i$, sont nulles;*

(ii) *les éléments de la diagonale de M sont $\lambda_1, \dots, \lambda_n$, rangés dans cet ordre, ce qui exprime que pour tout i, $1 \le i \le n$, la composante de $u(e_i)$ sur le vecteur de base e_i est λ_i.*

Démonstration. La propriété est vraie si la dimension de E est 1. Procédons par récurrence, en la supposant vraie pour tout espace vectoriel de dimension $\le n - 1$.

L'application linéaire $u - \lambda_1 \operatorname{id}_E$ n'étant pas un isomorphisme, son image est contenue dans un hyperplan de E, c'est-à-dire dans un sous-espace vectoriel F de E de dimension $n-1$. Montrons que u applique F dans lui-même. Soit $x \in F$. On peut peut écrire

$$u(x) = (u - \lambda_1 \operatorname{id}_E)(x) + \lambda_1 x\,.$$

Le premier terme du membre de droite, $(u - \lambda_1 \operatorname{id}_E)(x)$, appartient à l'image de $u - \lambda_1 \operatorname{id}_E$, donc à F. Le second terme, $\lambda_1 x$, appartient à F par hypothèse. Donc $u(x) \in F$, et on a bien $u(F) \subset F$.

Soit V la restriction de u à F, e_1 un élément de E n'appartenant pas à F et $(e'_2, \dots, e'_n)$ une base de F. Alors $(e_1, e'_2, \dots, e'_n)$ est une base de E. On peut écrire

$$u(e_1) = \lambda_1 e_1 + (u - \lambda_1 \operatorname{id}_E)(e_1)\,.$$

Comme $(u - \lambda_1 \operatorname{id}_E)(e_1)$ est élément de F, ceci prouve que le coefficient de la matrice de u, dans la base $(e_1, e'2, \dots, e'_n)$, situé dans l'angle supérieur gauche, est λ_1. Comme u applique les autres vecteurs de base $e'_2, \dots, e'_n$, dans F, cette matrice est de la forme

$$\begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ N' & & N & \end{pmatrix},$$

où N' est une matrice-colonne à $n-1$ éléments, et N une matrice $(n-1) \times (n-1)$, qui n'est autre que la matrice de la restriction v de u à F, dans la base $(e'_2, \dots, e'_n)$. Par suite, le polynôme caractéristique de u est

$$P_u(X) = (\lambda_1 - X) P_v(X)\,,$$

où $P_v(X)$ est le polynôme caractéristique de v.

D'après l'hypothèse de récurrence, il existe une base $(e_2 \dots, e_n)$ de F dans laquelle v a une matrice triangulaire inférieure, dont la diagonale est $(\lambda_2, \dots, \lambda_n$. On voit alors que $(e_1, e_2, \dots, e_n)$ est une base de E dans laquelle la matrice de u a la forme désirée. $\square$

2. Le théorème de Cayley-Hamilton

2.1. Notations. Soit E un espace vectoriel sur le corps $\mathbf{K} = \mathbf{R}$ ou $\mathbf{C}$, et u un endomorphisme de E. Pour tout polynôme

$$Q(X) = a_0 + a_1 X + a_2 X^2 + \cdots + a_m X^m\,,$$

à coefficients $a_0, a_1, a_2, \dots, a_m$ éléments de $\mathbf{K}$, on note $Q(u)$ l'endomorphisme de E

$$Q(u) = a_0 \operatorname{id}_E + a_1 u + a_2 u^2 + \cdots + a_m u^m\,,$$

avec la convention habituelle

$$u^2 = u \circ u\,, \quad \text{et, pour tout entier } k > 2\,, \quad u^k = u \circ u^{k-1}\,.$$

2.2. Théorème de Cayley-Hamilton. *Soit E un espace vectoriel de dimension finie sur le corps $\mathbf{K} = \mathbf{R}$ ou $\mathbf{C}$, u un endomorphisme de E, P_u le polynôme caractéristique de u. On a*

$$P_u(u) = 0\,.$$

Démonstration. On considère d'abord le cas où $\mathbf{K} = \mathbf{C}$. Soit n la dimension de E, $\lambda_1, \ldots, \lambda_n$ les valeurs propres de u rangées dans un ordre quelconque, chacune d'elles figurant un nombre de fois égal à sa multiplicité. D'après la proposition 1.6, il existe une base $(e_1, \ldots, e_n)$ de E dans laquelle la matrice de u a des coefficients λ_{ij}, $1 \leq i \leq n$, $1 \leq j \leq n$, qui vérifient

$$\lambda_{ij} = \begin{cases} 0 & \text{si } i < j, \\ \lambda_i & \text{si } i = j, \\ \lambda_{ij} & \text{quelconque si } i > j. \end{cases}$$

Le polynôme caractéristique de u a pour expression

$$P_u(X) = (\lambda_n - X)(\lambda_{n-1} - X) \ldots (\lambda_1 - X)\,.$$

D'après la forme de la matrice de u, on a

$$(\lambda_n \operatorname{id}_E - u)(e_n) = 0\,.$$

On va montrer que pour tout i vérifiant $1 \leq i \leq n-1$, l'endomorphisme

$$(\lambda_n \operatorname{id}_E - u) \circ (\lambda_{n-1} \operatorname{id}_E - u) \circ \cdots \circ (\lambda_i \operatorname{id}_E - u)$$

annule $e_i, e_{i+1}, \ldots, e_n$, c'est-à-dire applique chacun de ces vecteurs sur le vecteur nul.

On vient de voir que $(\lambda_n \operatorname{id}_E - u)$ annule e_n. On a

$$(\lambda_n \operatorname{id}_E - u) \circ (\lambda_{n-1} \operatorname{id}_E - u)(e_n) = (\lambda_{n-1} \operatorname{id}_E - u) \circ (\lambda_n \operatorname{id}_E - u)(e_n) = 0\,,$$

et de même

$$(\lambda_n \operatorname{id}_E - u) \circ (\lambda_{n-1} \operatorname{id}_E - u)(e_{n-1}) = (\lambda_n \operatorname{id}_E - u)(-\lambda_{n,n-1} e_n) = 0\,,$$

car $u(e_{n-1}) = \lambda_{n-1} e_{n-1} + \lambda_{n,n-1} e_n$. La propriété est donc démontrée pour $i = n-1$. On procède par récurrence descendante sur i, en supposant la propriété vraie pour $j+1 \leq i \leq n-1$. L'endomorphisme

$$(\lambda_n \operatorname{id}_E - u) \circ \ldots \circ (\lambda_{j+1} \operatorname{id}_E - u) \circ (\lambda_j \operatorname{id}_E - u) \tag{$*$}$$

annule $e_{j+1}, \dots, e_n$, d'après l'hypothèse de récurrence, car les facteurs qui le composent commutent deux à deux. D'autre part, d'après la forme de la matrice de u, $(\lambda_j \,\mathrm{id}_E - u)(e_j)$ est combinaison linéaire de $e_{j+1}, \dots, e_n$. L'hypothèse de récurrence montre alors que l'endomorphisme $(*)$ annule e_j.

Il suffit alors de faire $i = 1$ pour prouver que $P_u(u)$ annule tous les vecteurs de la base $(e_1, \dots, e_n)$, donc que $P_u(u) = 0$.

Considérons maintenant le cas où $\mathbf{K} = \mathbf{R}$. Soit $E_{\mathbf{C}}$ le complexifié de E, c'est-à-dire le produit $E \times E$, muni de la structure d'espace vectoriel complexe pour laquelle le produit par un scalaire complexe est défini par

$$(a + ib)(x, y) = (ax - by, ay + bx)\,, \quad \text{avec} \quad a \in \mathbf{R}\,,\ b \in \mathbf{R}\,,\ i = \sqrt{-1}\,,\ (x, y) \in E \times E\,.$$

On identifie E au sous-ensemble $E \times \{0\}$ de $E_{\mathbf{C}}$, ce qui permet d'écrire l'élément (x, y) de $E_{\mathbf{C}}$ sous la forme

$$(x, y) = x + iy\,,$$

et on prolonge u en un endomorphisme $u_{\mathbf{C}}$ de $E_{\mathbf{C}}$, en posant

$$u_{\mathbf{C}}(x, y) = \bigl(u(x), u(y)\bigr)\,.$$

Toute base de E, pour sa structure d'espace vectoriel réel, est aussi une base de $E_{\mathbf{C}}$ pour sa structure d'espace vectoriel complexe. Par suite, le polynôme caractéristique $P_u(X)$ de u est aussi le polynôme caractéristique de $u_{\mathbf{C}}$. Mais u étant la restriction de $u_{\mathbf{C}}$ à E, $P_u(u)$ est la restriction de $P_u(u_{\mathbf{C}})$ à E, donc est nul. □

2.3. Remarques.

1. La démonstration du théorème de Cayley-Hamilton donnée ici, qui repose sur la proposition 1.6 permettant de triangulariser la matrice d'un endomorphisme, est valable pour $\mathbf{K} = \mathbf{R}$ ou $\mathbf{C}$. Le lecteur intéressé trouvera une démonstration directe, n'utilisant pas la triangularisation, par exemple dans le livre de J. Lelong-Ferrand et J.-M. Arnaudiès (cours de mathématiques, tome 1, algèbre, Dunod, Paris 1977). Le théorème de Cayley-Hamilton est applicable à tout endomorphisme d'un espace vectoriel de dimension finie sur un corps quelconque, et même à tout endomorphisme d'un module de dimension finie sur un anneau unifère quelconque.

2. Dans les hypothèses du théorème 2.2, l'existence d'un polynôme Q tel que $Q(u) = 0$ pouvait être affirmée *a priori*. En effet, l'espace des endomorphismes de E est de dimension n^2, et par suite les $n^2 + 1$ endomorphismes $\mathrm{id}_E, u, u^2, \dots, u^{n^2}$ ne forment pas une famille libre. Il existe donc nécessairement un polynôme Q, à coefficients dans $\mathbf{K}$, de degré $\leq n^2$, tel que $Q(u) = 0$. Le théorème de Cayley-Hamilton donne l'expression explicite d'un tel polynôme, et montre qu'on peut lui imposer d'être de degré n.

2.4. Définition. Les hypothèses et notations étant celles du théorème de Cayley-Hamilton (2.2), on appelle *polynôme minimal* de u, et on note $Q_u(X)$, le polynôme normalisé de degré minimal qui annule u, c'est-à-dire qui vérifie $Q_u(u) = 0$.

Cette définition est justifiée par les considérations suivantes. L'ensemble des polynômes $Q(X)$, à coefficients dans $\mathbf{K}$, tels que $Q(u) = 0$, est un idéal de l'anneau $\mathbf{K}[X]$ des polynômes à coefficients dans $\mathbf{K}$, car si $Q(X)$ est un polynôme tel que $Q(u) = 0$, et $R(X)$ un polynôme quelconque,

$$(QR)(u) = Q(u) \circ R(u) = R(u) \circ Q(u) = 0\,.$$

Or on sait que tout idéal de $\mathbf{K}[X]$ est principal, c'est-à-dire est engendré par un élément unique de $\mathbf{K}[X]$. On peut donc affirmer qu'il existe un polynôme, unique si on lui impose d'être normalisé, qui est le polynôme $Q(X)$ de plus bas degré tel que $Q(u) = 0$. C'est le polynôme minimal $Q_u(X)$ de u. Tout autre polynôme $R(X) \in \mathbf{K}[X]$ vérifie $R(u) = 0$ si et seulement s'il est multiple de $Q_u(X)$. En particulier, le polynôme caractéristique $P_u(X)$ est multiple du polynôme minimal.

3. Sous-espaces caractéristiques

3.1. Notations. Dans tout ce paragraphe, E est un espace vectoriel de dimension finie n sur le corps $\mathbf{K} = \mathbf{R}$ ou $\mathbf{C}$, u un endomorphisme de E, $P_u(X)$ son polynôme caractéristique, $Q_u(X)$ son polynôme minimal. On dira qu'un polynôme $P(X)$, à coefficients dans $\mathbf{K}$, *annule* u, si $P(u) = 0$. En particulier, le polynôme caractéristique et le polynôme minimal de u annulent u.

3.2. Théorème. *Soient $S_1(X), S_2(X), \ldots, S_p(X)$ des polynômes à coefficients dans $\mathbf{K}$, deux à deux premiers entre eux. Le noyau N de $S_1(u) \circ S_2(u) \circ \cdots \circ S_p(u)$ est somme directe des noyaux N_1 de $S_1(u)$, N_2 de $S_2(u)$, …, N_p de $S_p(u)$.*

Démonstration. Le théorème est trivial pour $p = 1$. Montrons-le pour $p = 2$. D'après le théorème de Bezout, il existe deux polynômes $T_1(X)$ et $T_2(X)$ tels que

$$T_1(X)S_1(X) + T_2(X)S_2(X) = 1\,,$$

d'où

$$T_1(u) \circ S_1(u) + T_2(u) \circ S_2(u) = \mathrm{id}_E\ .$$

On a donc, pour tout $x \in E$,

$$T_1(u) \circ S_1(u)(x) + T_2(u) \circ S_2(u)(x) = x\,. \qquad (*)$$

Supposons $x \in N$. On a, puisque l'anneau $\mathbf{K}[u]$ des endomorphismes s'exprimant comme des fonctions polynômes de l'endomorphisme u est commutatif,

$$S_2(u) \circ T_1(u) \circ S_1(u)(x) = T_1(u) \circ S_1(u) \circ S_2(u)(x) = 0\,,$$

ce qui prouve que $T_1(u) \circ S_1(u)(x) \in N_2$.

De même, $T_2(u) \circ S_2(u)(x) \in N_1$.

L'égalité $(*)$ montre alors que tout élément de N est somme d'un élément de N_1 et d'un élément de N_2, donc que $N \subset N_1 + N_2$. Comme d'autre part $N_1 \subset N$ et $N_2 \subset N$, on a $N = N_1 + N_2$. Enfin si $x \in N_1 \cap N_2$, l'égalité (1) prouve que $x = 0$. Donc N est somme directe de N_1 et de N_2.

Supposons maintenant $p > 2$, et procédons par récurrence, en supposant la propriété établie lorsqu'on remplace p par $p-1$. On peut écrire

$$S(X) = S_1(X)S_2(X)\dots S_p(X) = R(X)S_p(X)\,, \quad \text{avec} \quad R(X) = S_1(X)S_2(X)\dots S_{p-1}(X)\,.$$

Les polynômes $R(X)$ et $S_p(X)$ étant premiers entre eux, le raisonnement fait pour $p = 2$ montre que le noyau N de $R(u) \circ S_p(u)$ est somme directe du noyau de $R(u)$ et du noyau N_p de $S_p(u)$. L'hypothèse de récurrence montre que le noyau de $R(u)$ est somme directe des noyaux N_1 de $S_1(u)$, N_2 de $S_2(u)$, …, N_{p-1} de $S_{p-1}(u)$. Donc N est somme directe de N_1, N_2, …, N_p. □

3.3. Corollaire. *Soient $S_1(X), S_2(X), \dots, S_p(X)$ des polynômes à coefficients dans $\mathbf{K}$, deux à deux premiers entre eux. On suppose que leur produit $S(X) = S_1(X)S_2(X)\dots S_p(X)$ annule u. Alors l'espace vectoriel E est somme directe des noyaux N_1 de $S_1(u)$, N_2 de $S_2(u)$, …, N_p de $S_p(u)$.*

Démonstration. Puisque $S(u) = 0$, son noyau est E. □

3.4. Théorème. *Soit E un espace vectoriel de dimension finie n sur $\mathbf{C}$, u un endomorphisme de E, $\lambda_1, \dots, \lambda_p$ ses valeurs propres (deux à deux distinctes), et $n_1, \dots, n_p$ leurs multiplicités, en tant que racines du polynôme caractéristique de u. On note E_i le noyau de $(\lambda_i \operatorname{id}_E - u)^{n_i}$, $(1 \leq i \leq p)$. Alors chacun des sous-espaces E_i est invariant par u, et a pour dimension n_i. De plus, l'espace E est somme directe des sous-espaces E_i $(1 \leq i \leq p)$.*

On dit que E_i est le *sous-espace caractéristique* de u associé à la valeur propre λ_i.

Démonstration. Puisque u commute avec $(\lambda_i \operatorname{id}_E - u)^{n_i}$, si $x \in E_i$, on a

$$(\lambda_i \operatorname{id}_E - u)^{n_i} \circ u(x) = u \circ (\lambda_i \operatorname{id}_E - u)^{n_i}(x) = 0\,,$$

donc $u(x) \in E_i$.

Le polynôme caractéristique de u est

$$P_u(X) = \prod_{i=1}^{p} (\lambda_i - X)^{n_i}\,.$$

Les polynômes $S_i(X) = (\lambda_i - X)^{n_i}$ étant deux à deux premiers entre eux, puisque les λ_i sont deux à deux distincts, le corollaire 3.3 montre que E est somme directe des sous-espaces E_i.

Dans une base de E formée par la réunion de bases de chacun des sosus-espaces E_i, la matrice de u est diagonale par blocs, chaque bloc étant la matrice de la restriction u_i de u

à E_i. Le polynôme caractéristique de u est donc le produit des polynômes caractéristiques des u_i, $1 \leq i \leq p$. La seule valeur propre de u_i étant λ_i, le polynôme caractéristique de u_i est nécessairement $(\lambda_i - X)^{n_i}$, et par suite la dimension de E_i est n_i. □

3.5. Remarques. Dans les hypothèses du théorème 3.4, soit $Q_u(x)$ le polynôme minimal de u. Puisqu'il divise $P_u(X)$ et qu'il est normalisé, il est nécessairement de la forme

$$Q_u(X) = \prod_{i=1}^{p} (X - \lambda_i)^{m_i}, \quad \text{avec} \quad 0 \leq m_i \leq n_i .$$

D'après 3.3, E est somme directe des noyaux des endomorphismes $(u - \lambda_i)^{m_i}$. Mais puisque $m_i \leq n_i$, le noyau de $(u - \lambda_i)^{m_i}$ est contenu dans le noyau E_i de $(\lambda_i - u)^{n_i}$. Comme E est aussi somme directe des E_i, le noyau de $(u - \lambda_i)^{m_i}$ est nécessairement E_i. Par conséquent, on a $m_i \geq 1$.

D'autre part, le polynôme minimal de u_i est nécessairement $(X - \lambda_i)^{m_i}$. En effet, il divise $(X - \lambda_i)^{m_i}$, donc est de la forme $(X - \lambda_i)^{k_i}$, avec $k_i \leq m_i$. Mais alors $\prod_{i=1}^{p} (X - \lambda_i)^{k_i}$ annule u, donc est un multiple du polynôme minimal de u, ce qui prouve que $k_i = m_i$ pour tout i, $1 \leq i \leq p$.