Réduction des endomorphismes

Table des matières

La plupart des notions définies dans ce cours pour des *endomorphismes* a un analogue pour les *matrices* ; cet analogue sera souvent sous-entendu.

1 Sous-espaces stables et polynômes d'endomorphismes

Soit E un espace vectoriel de dimension finie sur un corps k, et $f : E \to E$ un endomorphisme linéaire de E. *Réduire* f, c'est chercher une base de E dans laquelle la matrice de f est la plus simple possible. Une telle base possède un intérêt calculatoire évident et également une signification géométrique ou physique importante. La situation idéale est celle où l'espace E est somme directe de droites stables sous f ; en général, on essaie de se rapprocher le plus possible de cette situation en décomposant E en somme directe de sous-espaces stables les plus petits possibles. Introduisons maintenant formellement ces diverses notions.

1.1 Définitions. Soient E un k-espace vectoriel de dimension finie et f un endomorphisme de E.

(1) Un sous-espace vectoriel $F \subset E$ est *stable sous* f ou f*-stable* si on a $f(F) \subset F$.

(2) Un *vecteur propre* est un vecteur non nul $x \in E$ qui engendre une droite f-stable.

(3) Une *valeur propre* est un scalaire $\lambda \in k$ tel que $f - \lambda\,\mathrm{Id}$ n'est pas injective, c'est-à-dire tel qu'il existe $x \in E \setminus \{0\}$ avec $f(x) = \lambda x$.

(4) Le *spectre de* f noté $\mathrm{Sp}(f)$ est l'ensemble des valeurs propres de f.

(5) Si $\lambda \in \mathrm{Sp}(f)$, on appelle *sous-espace propre de* f *associé à la valeur propre* λ le sous-espace $\ker(f - \lambda\,\mathrm{Id})$, souvent noté E_λ.

Si $\lambda \in k$ et $x \in E \setminus \{0\}$, c'est donc la même chose de dire que λ est une valeur propre de vecteur propre x, ou que x est un vecteur propre associé à la valeur propre λ. Observons que l'endomorphisme $f - \lambda\,\mathrm{Id}$ est un exemple de polynôme en f puisqu'il est égal à $P(f)$ avec $P(X) = X - \lambda$. Plus généralement, les noyaux de polynômes en f donneront des exemples de sous-espaces stables :

1.2 Définition. Soit $P = a_0 + a_1 X + \cdots + a_d X^d$ un polynôme à coefficients dans k. On note $P(f)$ l'endomorphisme de E égal à $a_0\,\mathrm{Id} + a_1 f + \cdots + f^d$, où f^i désigne l'endomorphisme f itéré i fois, c'est-à-dire $f \circ f \circ \cdots \circ f$. On appelle *polynôme en* f tout endomorphisme de la forme $P(f)$.

1.3 Proposition. *Soit f un endomorphisme de E et P, Q deux polynômes à coefficients dans k.*

(1) $P(f) + Q(f) = (P + Q)(f)$.

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

(3) $\ker P(f)$ *et* $\operatorname{im} P(f)$ *sont des sous-espaces f-stables de E.*

Preuve : Notons $P = a_0 + a_1 X + \cdots + X^d$ et $Q = b_0 + b_1 X + \cdots + b_e X^e$. On a :

$$P(f)+Q(f) = a_0 \operatorname{Id} + a_1 f + \cdots + f^d + b_0 \operatorname{Id} + b_1 f + \cdots + b_e d^e = (a_0+b_0) \operatorname{Id} + (a_1+b_1) f + \cdots = (P+Q)(f)$$

ce qui démontre le point (1). Maintenant, rappelons-nous que d'après les règles de multiplication des polynômes, le coefficient de X^l dans le produit PQ est égal à $\sum_{i+j=l} a_i b_j$. Par conséquent,

$$\begin{aligned}
P(f) \circ Q(f) &= \left(\sum_{i=0}^{d} a_i f^i\right) \circ \left(\sum_{j=0}^{e} b_j f^j\right) = \sum_{i=0}^{d} a_i \left(f^i \circ \left(\sum_{j=0}^{e} b_j f^j\right)\right) \text{ par définition de } \sum a_i f^i, \\
&= \sum_{i=0}^{d} a_i \sum_{j=0}^{e} b_j (f^i \circ f^j) \text{ car } f^i \text{ est linéaire,} \\
&= \sum_{l=0}^{\max(d,e)} \left(\sum_{i+j=l} a_i b_j\right) f^k = (PQ)(f).
\end{aligned}$$

Comme $(PQ)(f) = (QP)(f)$, on obtient ainsi le point (2). Pour (3), on note tout d'abord que $\ker(P(f))$ et $\operatorname{im}(P(f))$ sont des sous-espaces vectoriels, comme tout noyau et image d'un endomorphisme. Le fait que ces sous-espaces soient f-stables est une conséquence du point (2). En effet, si $P(f)(x) = 0$, alors $P(f)(f(x)) = f(P(f)(x)) = f(0) = 0$ donc $\ker P(f)$ est stable, et si $x = P(f)(y)$, alors $f(x) = f(P(f)(y)) = P(f)(f(y))$ donc $\operatorname{im} P(f)$ est stable. □

Les propriétés (1) et (2) de la proposition 1.3 disent que l'application linéaire $\mathrm{ev}_f : k[X] \to L(E)$ qui envoie P sur $P(f)$ est un morphisme d'anneaux unitaires. Nous allons voir maintenant qu'à cause de ceci, l'arithmétique de l'anneau $k[X]$ (et plus précisément le fait que ce soit un anneau *principal*) a des conséquences fortes dans $L(E)$.

Dans la suite, on notera le plus souvent $P(f)Q(f)$ au lieu de $P(f) \circ Q(f)$.

1.4 Lemme. *Soient deux polynômes P et Q tels que P divise Q. Alors* $\ker P(f) \subset \ker Q(f)$ *et* $\operatorname{im} Q(f) \subset \operatorname{im} P(f)$.

Preuve : Ceci est laissé en exercice. Attention au sens de l'inclusion dans chaque cas ! □

1.5 Proposition (Lemme des noyaux). *Soient E un k-espace vectoriel de dimension finie, f un endomorphisme de E, et $P_1, \ldots, P_s$ des polynômes premiers entre eux. Soit P le produit des P_i. Alors, on a* $\ker(P(f)) = \ker(P_1(f)) \oplus \cdots \oplus \ker(P_s(f))$.

Preuve : Pour simplifier, nous donnons la démonstration lorsque $s = 2$. On doit donc montrer que $\ker P(f) = \ker P_1(f) \oplus \ker P_2(f)$. Notons d'abord que $\ker P_i(f) \subset \ker P(f)$ pour $i = 1, 2$ d'après le lemme. Comme P_1 et P_2 sont premiers entre eux par hypothèse, d'après le théorème de Bézout (voir le cours sur les anneaux) il existe deux polynômes U_1 et U_2 tels que $U_1 P_1 + U_2 P_2 = 1$. En évaluant cette égalité sur l'endomorphisme f, on trouve

$$U_1(f) P_1(f) + U_2(f) P_2(f) = \operatorname{Id}. \tag{1}$$

Si $x \in \ker P(f)$, posons $x_1 = (U_2(f)P_2(f))(x)$ et $x_2 = (U_1(f)P_1(f))(x)$. Comme $P_1(f)P_2(f) = P(f)$, on a $x_1 \in \ker P_1(f)$ et $x_2 \in \ker P_2(f)$. D'après (1), on a $x = x_1 + x_2$ ce qui montre que $\ker P_1(f)$ et $\ker P_2(f)$ engendrent $\ker P(f)$. Par ailleurs, si $x \in \ker P_1(f) \cap \ker P_2(f)$ on a $x_1 = x_2 = 0$ d'après leur définition, donc $x = 0$. Ainsi $\ker P_1(f)$ et $\ker P_2(f)$ sont en somme directe. □

1.6 Corollaire. *Soit $f \in L(E)$ et $\lambda_1, \dots, \lambda_s$ des valeurs propres distinctes. Alors,*

(1) *les sous-espaces propres E_{λ_i} sont en somme directe. En particulier $s \leqslant n = \dim(E)$.*

(2) *Si $s = n$, les espaces propres E_{λ_i} sont des droites et E en est la somme directe.*

Preuve : (1) C'est le cas particulier du lemme des noyaux dans lequel $P_i = X - \lambda_i$. Le fait que $s \leqslant n$ provient du fait que chaque sous-espace propre est de dimension au moins 1, donc leur somme directe est de dimension au moins s.
(2) On a $n \leqslant \dim(E_{\lambda_1}) + \cdots + \dim(E_{\lambda_n}) \leqslant n$ donc chaque dimension est égale à 1 et la somme directe des E_{λ_i} est E. □

2 Polynôme minimal et polynôme caractéristique

Nous avons vu qu'il existe une application linéaire $\mathrm{ev}_f : k[X] \to L(E)$ telle que $\mathrm{ev}_f(P) = P(f)$. Comme $k[X]$ est de dimension infinie alors que $L(E)$ est de dimension finie, ce morphisme n'est pas injectif, donc $\ker(\mathrm{ev}_f) \neq \{0\}$.

2.1 Définition. On apppelle *polynôme annulateur de f* un polynôme P tel que $P(f) = 0$, c'est-à-dire un élément de $\ker(\mathrm{ev}_f)$.

Cette partie est consacrée à deux polynômes annulateurs importants. Pour définir le premier, il faut noter que ev_f est aussi un morphisme d'anneaux. Ainsi son noyau $\ker(\mathrm{ev}_f)$ est un *idéal* de $k[X]$, et comme $k[X]$ est principal cet idéal peut être engendré par un seul élément (voir le cours sur les anneaux pour plus de détails).

2.2 Définition. Soit $f \in L(E)$ un endomorphisme. On appelle *polynôme minimal de f* et on note μ_f le générateur unitaire de l'idéal des polynômes annulateurs de f. C'est aussi le polynôme annulateur de f unitaire et de plus petit degré.

2.3 Exemple. Un endomorphisme est nilpotent si et seulement si son polynôme minimal est de la forme X^d, pour un certain entier d appelé l'*indice de nilpotence*. Un endomorphisme distinct de 0 et de l'identité est un projecteur si et seulement si son polynôme minimal est $X^2 - X$.

Il y a un autre polynôme naturellement associé à f, et crucial à cause du lemme 2.6 ci-dessous :

2.4 Définition. Le *polynôme caractéristique de f* est le polynôme χ_f défini par

$$\chi_f(X) = \det(X \operatorname{Id} - f).$$

Dans cette définition, on doit voir X comme un *scalaire*. Ceci signifie que si on choisit une base de E et que l'on note A la matrice de f dans cette base, alors $\chi_f(X)$ est le déterminant de la matrice $X \operatorname{Id} - A$ qui est une matrice à coefficients dans le corps de fractions rationnelles $k' = k(X)$.

2.5 Remarque. Soit $n = \dim(E)$. Le polynôme caractéristique est de degré n. Plus précisément, son terme de plus haut degré est X^n et il est en particulier unitaire. Certains auteurs définissent le polynôme caractéristique par $\chi_f(X) = \det(f - X\,\mathrm{Id})$, mais alors son coefficient dominant est $(-1)^n$ ce qui est un peu moins agréable.

2.6 Lemme. *Les racines de χ_f sont les valeurs propres de f.*

Preuve : Un scalaire $\lambda \in k$ est racine de χ_f si et seulement si $\det(\lambda\,\mathrm{Id} - f) = 0$, c'est-à-dire si et seulement si $\lambda\,\mathrm{Id} - f$ n'est pas injective. Ceci est la définit d'une valeur propre. □

2.7 Lemme. *Soient E un k-espace vectoriel de dimension finie et $f \in L(E)$. On considère un sous-espace vectoriel f-stable F et on note $f|_F$ la restriction de f à ce sous-espace. Alors,*

(1) *$\mu_{f|_F}$ divise μ_f,*

(2) *$\chi_{f|_F}$ divise χ_f.*

Preuve : (1) Notons que pour tout polynôme P, on a $P(f)|_F = P(f|_F)$. En particulier $\mu_f(f|_F) = \mu_f(f)|_F = 0$. Or par définition de $\mu_{f|_F}$, tout polynôme qui annule $f|_F$ est un multiple de $\mu_{f|_F}$. On obtient donc que $\mu_{f|_F}$ divise μ_f.
(2) Maintenant, soit $\{e_1, \dots, e_m\}$ une base de F, G un supplémentaire de F, $\{e_{m+1}, \dots, e_n\}$ une base de G, de sorte que la réunion $\{e_1, \dots, e_n\}$ est une base de E. Dans cette base, la matrice de f est triangulaire par blocs, de la forme $A = \left(\begin{smallmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{smallmatrix}\right)$ où A_{11} est la matrice de $f|_F$. Le polynôme caractéristique de f est le déterminant de la matrice

$$X\,\mathrm{Id} - A = \begin{pmatrix} X\,\mathrm{Id} - A_{11} & -A_{12} \\ 0 & X\,\mathrm{Id} - A_{22} \end{pmatrix}.$$

Comme le déterminant d'une matrice triangulaire par blocs est produit des déterminants des blocs diagonaux, on trouve $\chi_f(X) = \det(X\,\mathrm{Id} - A_{11}) \det(X\,\mathrm{Id} - A_{22}) = \chi_{f|_F}(X) \det(X\,\mathrm{Id} - A_{22})$ ce qui démontre que $\chi_{f|_F}$ divise χ_f. □

2.8 Théorème (Cayley-Hamilton). *Pour tout endomorphisme $f \in L(E)$ de polynôme caractéristique noté χ, on a $\chi(f) = 0$. En d'autres termes χ est un polynôme annulateur pour f, ou encore, μ divise χ.*

Preuve : Montrer que $\chi(f) = 0$ revient à montrer que $(\chi(f))(x) = 0$ pour tout vecteur x. Fixons x et notons m le plus petit entier tel qu'il existe une combinaison linéaire

$$f^m(x) + a_{m-1} f^{m-1}(x) + \cdots + a_1 f(x) + a_0 x = 0. \tag{2}$$

Par ce choix de m, la famille $\mathcal{B} = \{x, f(x), \dots, f^{m-1}(x)\}$ est libre de sorte que le sous-espace $F = \mathrm{Vect}(\mathcal{B})$ est f-stable, de dimension m, de base $\mathcal{B}$. Dans cette base, la matrice de $f|_F$ est

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

Notons $\chi' = \chi_{f|_F}$. On sait calculer le polynôme caractéristique de la matrice ci-dessus (c'est un exercice classique), c'est $\chi'(X) = X^m + a_{m-1} X^{m-1} + \cdots + a_1 X + a_0$. L'équation (2) nous dit précisément que

$\chi'(f)(x) = 0$. Or d'après le lemme 2.7, il existe un polynôme P tel que $\chi = P\chi'$, donc $\chi(f) = P(f)\chi'(f)$ et

$$(\chi(f))(x) = P(f)(\chi'(f)(x)) = P(f)(0) = 0,$$

comme on voulait démontrer. □

2.9 Corollaire. *On a* $\deg(\mu_f) \leqslant n$.

Preuve : μ_f divise χ_f d'après le théorème Cayley-Hamilton, donc $\deg(\mu_f) \leqslant \deg(\chi_f) = n$. □

3 Endomorphismes trigonalisables et diagonalisables

3.1 Définition. Soient E un k-espace vectoriel de dimension finie et $f \in L(E)$. On dit que f est *trigonalisable* (resp. *diagonalisable*), s'il existe une base de E dans laquelle la matrice de f est triangulaire supérieure (resp. diagonale).

Soit $\mathcal{B}$ une base de E et A la matrice de f dans cette base. Compte tenu du fait que la matrice de f dans une base $\mathcal{B}'$ est $P^{-1}AP$, où P est la matrice de passage de $\mathcal{B}$ à $\mathcal{B}'$, dire que f (ou A) est trigonalisable (resp. diagonalisable) signifie qu'il existe une matrice inversible P telle que $P^{-1}AP$ est triangulaire supérieure (resp. diagonale).

3.2 Théorème. *Soit f un endomorphisme d'un k-espace vectoriel de dimension finie. Alors les conditions suivantes sont équivalentes :*

(1) *f est trigonalisable,*

(2) *le polynôme caractéristique χ de f est scindé.*

On rappelle qu'un polynôme est *scindé* s'il est produit de polynômes de degré 1.

Preuve : (1) $\Rightarrow$ (2). Le polynôme caractéristique de f est le polynôme caractéristique de la matrice de f dans une base quelconque. Si f est trigonalisable, on peut choisir une base dans laquelle la matrice de f est triangulaire supérieure. On a alors

$$\chi(X) = \det\left(X\,\mathrm{Id} - \begin{pmatrix} a_{11} & * & * \\ & \ddots & * \\ 0 & & a_{nn} \end{pmatrix} \right) = \det \begin{pmatrix} X - a_{11} & * & * \\ & \ddots & * \\ 0 & & X - a_{nn} \end{pmatrix},$$

c'est-à-dire $\chi(X) = (X - a_{11}) \dots (X - a_{nn})$ qui est scindé.

(2) $\Rightarrow$ (1). Montrons par récurrence sur n que si χ est scindé, f est trigonalisable. Pour $n = 1$, il n'y a rien à démontrer. Supposons la propriété vraie au rang $n-1$ et notons $\lambda_1, \dots, \lambda_n$ les racines non nécessairement distinctes de χ. Soit e_1 un vecteur propre de f relatif à la valeur propre λ_1 et complétons-le en une base $\{e_1, e_2', \dots, e_n'\}$ de E. La matrice de f dans cette base est de la forme

$$A = \begin{pmatrix} \lambda_1 & B \\ 0 & C \end{pmatrix}.$$

Un calcul direct montre que $\chi_A(X) = (X - \lambda_1)\chi_C(X)$ et donc χ_C est scindé, de degré $n-1$. D'après l'hypothèse de récurrence, il existe une matrice $Q \in \mathrm{GL}_{n-1}(k)$ telle que $Q^{-1}CQ$ est une matrice triangulaire supérieure T. Introduisons la matrice diagonale par blocs $P = \left(\begin{smallmatrix} 1 & 0 \\ 0 & Q \end{smallmatrix}\right)$. On voit que

$$P^{-1}AP = \begin{pmatrix} \lambda_1 & BQ \\ 0 & Q^{-1}CQ \end{pmatrix} = \begin{pmatrix} \lambda_1 & BQ \\ 0 & T \end{pmatrix},$$

qui est bien une matrice triangulaire supérieure. $\square$

3.3 Corollaire. *Si k est un corps algébriquement clos, par exemple le corps $\mathbb{C}$ des nombres complexes, alors tout endomorphisme est trigonalisable.*

Preuve : Rappelons qu'un corps k est dit *algébriquement clos* lorsque tout polynôme à coefficients dans k est scindé. Ce corollaire est donc une application directe du théorème précédent. Dans le cas de $\mathbb{C}$, il faut invoquer le théorème fondamental de l'Algèbre (ou théorème de d'Alembert-Gauss) qui dit que $\mathbb{C}$ est algébriquement clos. $\square$

3.4 Théorème. *Soit f un endomorphisme d'un k-espace vectoriel de dimension finie. Alors les conditions suivantes sont équivalentes :*

(1) *f est diagonalisable,*

(2) *le polynôme minimal μ de f est scindé à racines simples,*

(3) *il existe un polynôme scindé à racines simples P qui annule f.*

Preuve : (1) $\Rightarrow$ (2). Si f est diagonalisable, il existe une base $\{e_1,\dots,e_n\}$ de E dans laquelle la matrice de f est diagonale, avec pour coefficients diagonaux les valeurs propres $\lambda_1,\dots,\lambda_s$ apparaissant avec des multiplicités α_i. D'après le lemme des noyaux 1.5, l'espace E est somme directe des sous-espaces propres E_{λ_i}, et il est alors clair que le polynôme $P=(X-\lambda_1)\dots(X-\lambda_s)$ annule f. Le polynôme minimal μ est un diviseur de P, donc il est scindé à racines simples. En fait on peut montrer que $\mu=P$: il suffit de noter que si Q est un diviseur strict de P, l'une des valeurs propres λ_i n'est pas racine de Q et on voit que la restriction de $Q(f)$ à E_{λ_i} n'est pas nulle.

(2) $\Rightarrow$ (3) est évident.

(3) $\Rightarrow$ (1). Soit P un polynôme scindé à racines simples tel que $P(f)=0$. Écrivons $P=(X-\lambda_1)\dots(X-\lambda_s)$. D'après le lemme des noyaux, on a $E=\ker P(f)=E_{\lambda_1}\oplus\dots\oplus E_{\lambda_s}$. Pour chaque i, notons $\mathcal{B}_i$ une base de E_{λ_i} puis $\mathcal{B}=\mathcal{B}_1\cup\dots\cup\mathcal{B}_s$. On obtient ainsi une base de E dans laquelle la matrice de f est diagonale. $\square$

3.5 Corollaire. *Soit f un endomorphisme d'un k-espace vectoriel de dimension finie E, et soit $F\subset E$ un sous-espace f-stable. Si f est diagonalisable, la restriction $f|_F$ est diagonalisable.*

Preuve : D'après le théorème, le polynôme minimal de f est scindé à racines simples. Par ailleurs μ annule f, donc il annule $f|_F$. Par une nouvelle application du théorème, précisément l'implication (3) $\Rightarrow$ (1), la restriction $f|_F$ est diagonalisable. $\square$

4 Sous-espaces caractéristiques et calcul du polynôme minimal

Dans cette section, à l'aide de la notion de sous-espace caractéristique, nous expliquons comment utiliser tous les concepts développés précédemment pour étudier un endomorphisme f donné. Nous en tirons ensuite quelques conséquences concernant le calcul pratique des objets les plus importants attachés à f, notamment son polynôme minimal. Nous ne considérerons que le cas où le polynôme caractéristique est *scindé*, c'est-à-dire qu'il est produit de polynômes de degré 1. Par exemple, ceci est toujours le cas lorsque $k=\mathbb{C}$.

4.1 Définition. Soit f un endomorphisme d'un k-espace vectoriel de dimension finie, $\lambda \in k$ une valeur propre de f et α la multiplicité de f dans le polynôme caractéristique χ_f. On appelle *sous-espace caractéristique de f associé à la valeur propre λ* le sous-espace $F_\lambda = \ker((f - \lambda\,\mathrm{Id})^\alpha)$.

Si l'on désigne par E_λ le sous-espace propre associé à la valeur propre λ, alors $E_\lambda \subset F_\lambda$. Ceci vient simplement du fait que $\ker(f - \lambda\,\mathrm{Id}) \subset \ker((f - \lambda\,\mathrm{Id})^\alpha)$.

4.2 Théorème. *Soient k un corps, E un k-espace vectoriel de dimension finie, f un endomorphisme de E. On suppose que f a un polynôme caractéristique χ scindé : $\chi(X) = (X - \lambda_1)^{\alpha_1} \dots (X - \lambda_r)^{\alpha_r}$ où $\lambda_1, \dots, \lambda_r$ sont les valeurs propres distinctes de f et $\alpha_i \geqslant 1$ est la multiplicité de λ_i. Enfin, on note μ le polynôme minimal de f.*

(1) *Soit F_i le sous-espace caractéristique associé à λ_i. On a $E = F_1 \oplus \cdots \oplus F_r$.*

(2) *Le sous-espace F_i est f-stable. Soit $f_i = f|_{F_i} : F_i \to F_i$ la restriction de f à F_i. L'endomorphisme $f_i - \lambda_i\,\mathrm{Id}$ est nilpotent et la seule valeur propre de f_i est λ_i.*

(3) *Soit χ_i le polynôme caractéristique de f_i. Alors $\chi = \chi_1 \dots \chi_r$.*

(4) *Soit μ_i le polynôme minimal de f_i. Alors $\mu = \mu_1 \dots \mu_r$.*

(5) *On a $\chi_i(X) = (X - \lambda_i)^{\alpha_i}$. En particulier $\dim(F_i) = \deg(\chi_i) = \alpha_i$.*

(6) *On a $\mu_i(X) = (X - \lambda_i)^{\beta_i}$, où $\beta_i \leqslant \alpha_i$ est l'indice de nilpotence de $f_i - \lambda_i\,\mathrm{Id}$.*

(7) *Soit $\mathcal{B}_i$ une base de F_i qui trigonalise f_i (une telle base existe d'après le théorème 3.2). Soit $\mathcal{B} = \mathcal{B} \cup \cdots \cup \mathcal{B}_r$ qui est une base de E, et soit M la matrice de f dans la base $\mathcal{B}$. Alors M est triangulaire, diagonale par blocs (triangulaires) de tailles α_i, et f est diagonalisable si et seulement si $\beta_i = 1$ pour tout i, si et seulement si M est diagonale.*

Preuve : (1) D'après le théorème de Cayley-Hamilton (théorème 2.8), on a $\chi(f) = 0$. Il s'ensuit que $\ker \chi(f) = E$. Maintenant, notons que les polynômes $P_i(X) = (X - \lambda_i)^{\alpha_i}$ sont premiers entre eux deux à deux, et appliquons le lemme des noyaux à $\chi = P_1 \dots P_r$. On trouve $E = \ker \chi(f) = \ker P_1(f) \oplus \cdots \oplus \ker P_r(f) = F_1 \oplus \cdots \oplus F_r$.

(2) L'endomorphisme $(f - \lambda_i\,\mathrm{Id})^{\alpha_i} = P_i(f)$ est un polynôme en f. Donc F_i, qui n'est autre que son noyau, est f-stable par la proposition 1.3. Par ailleurs, on a $(f_i - \lambda_i\,\mathrm{Id})^{\alpha_i} = (f - \lambda_i\,\mathrm{Id})^{\alpha_i}|_{F_i}$ qui est nul d'après la définition de F_i. Ainsi $f_i - \lambda_i\,\mathrm{Id}$ est nilpotent d'indice $\leqslant \alpha_i$. Donc sa seule valeur propre est 0. Si λ est une valeur propre de f_i et $x \neq 0$ un vecteur propre associé, on a $(f_i - \lambda_i\,\mathrm{Id})(x) = f_i(x) - \lambda_i x = \lambda x - \lambda_i x = (\lambda - \lambda_i)x$. Ainsi x est un vecteur propre (non nul) de $f_i - \lambda_i\,\mathrm{Id}$ pour la valeur propre $\lambda - \lambda_i$. Puisque la seule valeur propre de $f_i - \lambda_i\,\mathrm{Id}$ est 0, il s'ensuit que $\lambda = \lambda_i$ donc la seule valeur propre de f_i est λ_i.

(3) Choisissons pour chaque i une base $\mathcal{B}_i$ de F_i. D'après le point (1), la réunion $\mathcal{B} = \mathcal{B}_1 \cup \cdots \cup \mathcal{B}_r$ est une base de E. La matrice de f dans $\mathcal{B}$ est diagonale par blocs, et le i-ième bloc n'est autre que la matrice de f_i dans la base $\mathcal{B}_i$. Comme le déterminant d'une matrice diagonale par blocs est le produit des déterminants des blocs diagonaux, on voit que $\chi = \chi_1 \dots \chi_r$.

(4) Notons $\pi = \mu_1 \dots \mu_r$. Pour montrer que $\mu = \pi$, il suffit de montrer que μ divise π puis que π divise μ, car alors $\mu = u\pi$ pour un scalaire $u \in k$, et $u = 1$ puisque μ et π sont tous deux unitaires. Commençons par montrer que μ divise π. Pour chaque i, par définition du polynôme minimal μ_i on a $\mu_i(f_i) = 0$. Comme μ_i divise π, on en déduit que $\pi(f_i) = 0$. Comme $\pi(f_i) = \pi(f)|_{F_i}$, ceci signifie que $\pi(f)$ est un endomorphisme qui est nul sur chaque sous-espace F_i. Comme les F_i engendrent E d'après le point (1), on en déduit que $\pi(f) = 0$ c'est-à-dire que π est un polynôme annulateur de f. Or les polynômes annulateurs sont tous multiples du polynôme minimal ; donc μ divise π. Montrons maintenant que π divise μ. On sait par le théorème de Cayley-Hamilton que μ_i divise χ_i. Étant donné que les χ_i sont premiers entre eux deux à deux, ceci montre que les μ_i sont premiers entre eux deux à

deux. On sait aussi par le lemme 2.7 que μ_i divise μ. Alors μ qui est divisible par chacun des μ_i premiers entre eux deux à deux doit être divisible par leur produit, c'est-à-dire π. Ceci termine la preuve du fait que $\mu = \mu_1 \dots \mu_r$.

(5) Rappelons que $\chi(X) = P_1 \dots P_r$ avec $P_i(X) = (X - \lambda_i)^{\alpha_i}$. Comme χ_i divise χ d'après le lemme 2.7 et que de plus la seule valeur propre de f_i est λ_i, on déduit que χ_i divise P_i. En d'autres termes, il existe $\gamma_i \leqslant \alpha_i$ tel que $\chi_i(X) = (X - \lambda_i)^{\gamma_i}$. Mais comme $\chi = \chi_1 \dots \chi_r$, la seule possibilité est que $\gamma_i = \alpha_i$. Comme le degré du polynôme caractéristique est égal à la dimension de l'espace vectoriel ambiant, on obtient $\dim(F_i) = \deg(\chi_i) = \alpha_i$.

(6) L'indice de nilpotence β_i est par définition le plus petit entier m tel que $(f_i - \lambda_i \,\mathrm{Id})^m = 0$. Ceci signifie premièrement que $(X-\lambda_i)^{\beta_i}$ est un polynôme annulateur de f_i, et deuxièmement que $(X-\lambda_i)^m$ n'est pas un polynôme annulateur si $m < \beta_i$. Le premier fait implique que le polynôme minimal de f_i est de la forme $\mu_i(X) = (X - \lambda_i)^{\delta_i}$ avec $\delta_i \leqslant \beta_i$, et le deuxième fait implique que $\delta_i = \beta_i$.

(7) Le fait que la réunion des $\mathcal{B}_i$ est une base de E provient de la décomposition $E = F_1 \oplus \dots \oplus F_r$. Le fait que la matrice de f dans la base $\mathcal{B}$, c'est-à-dire M, soit diagonale par blocs provient du fait que F_i est stable par f. Chaque bloc est bien sûr triangulaire par choix de $\mathcal{B}_i$. Enfin, si M est diagonale alors bien sûr f est diagonalisable, et réciproquement, si f est diagonalisable, alors pour tout i le sous-espace caractéristique F_i est égal au sous-espace propre E_i, donc $f_i : F_i \to F_i$ est l'homothétie de rapport λ_i et sa matrice, qui est le i-ième bloc de M, est $\lambda_i \,\mathrm{Id}$. Il s'ensuit que M est diagonale. □

4.3 Remarques. (1) D'un point de vue pratique, dans le cas où l'on sait trouver toutes les racines de χ, ce résultat explique comment calculer le polynôme minimal, et décider si oui ou non f est diagonalisable. Plus précisément, la méthode est de calculer une base $\mathcal{B}_i$ de chaque sous-espace caractéristique F_i. Si f est une homothétie sur chaque F_i (de rapport λ_i), alors f est diagonalisable et son polynôme minimal est $\mu(X) = (X - \lambda_1) \dots (X - \lambda_r)$. De manière générale, pour chaque i on dispose de la matrice M_i de f_i dans la base $\mathcal{B}_i$; on calcule les puissances successives de $M_i - \lambda_i \,\mathrm{Id}$, et la première puissance nulle donne l'indice de nilpotence β_i. On sait alors que $\mu(X) = (X - \lambda_1)^{\beta_1} \dots (X - \lambda_r)^{\beta_r}$.

(2) Supposons que f est donné, après choix d'une base de E, par une matrice M. Si l'on a de bonnes raisons de penser *a priori* que f (ou M) est diagonalisable, alors la méthode la plus rapide est de poser et résoudre directement le système $MX = \lambda X$ pour un λ arbitraire. En appliquant par exemple la méthode du pivot de Gauss pour essayer de trouver des solutions X, on voit au cours du procédé que certaines valeurs de λ posent problème i.e. donnent un système avec des lignes linéairement dépendantes. En fait, la valeur $\chi(\lambda)$ (où χ est le polynôme caractéristique) doit apparaître naturellement, ce qui s'explique par le fait que les manipulations lignes-colonnes opérées pour résoudre le système reviennent à calculer un déterminant. Pour chaque λ racine de χ, on continue le calcul en donnant une base de l'espace propre associé. On obtient ainsi directement les valeurs propres et, en même temps, une matrice de passage dans une base de diagonalisation.

(3) En appliquant le lemme des noyaux pour $\mu = \mu_1 \dots \mu_r$, on voit que $E = \ker \mu_1(f) \oplus \dots \oplus \ker \mu_r(f)$. Étant donné que μ_i divise χ_i, on a par ailleurs $\ker \mu_i(f) \subset F_i$ et on en déduit que cette inclusion est une égalité. Donc $F_i = \ker((f - \lambda_i)^{\beta_i})$ et on montre que $F_i = \ker((f - \lambda_i)^m$ pour tout $m \geqslant \beta_i$, par exemple pour $m = \dim(E)$.