Chapitre 4

Reduction Des Endomorphismes

Ce chapitre est probablement le plus important du programme d'algèbre de cycle préparatoire. On y met en application tous les résultats obtenus jusqu'ici, pour donner une manière efficace de réduction des endomorphismes en dimension finie : intuitivement, il s'agit de trouver « la meilleure base » dans laquelle se placer pour étudier un endomorphisme donné. Comme d'habitude, la lettre E désigne un $\mathbb{K}$-espace vectoriel.

4.1 Diagonalisation et trigonalisation

4.1.1 Éléments propres d'un endomorphisme

Les plus simples des endomorphismes de E sont les homothéties, c'est-à-dire les applications $\lambda\,\mathrm{id}_E$ avec $\lambda \in \mathbb{K}$. Bien entendu, un endomorphisme donné de E ne sera pas, en général, une homothétie ; en revanche, on peut espérer trouver des sous-espaces de E sur lesquels f se comporte comme une homothétie.

Dans toute la suite du cours, si $f \in \mathscr{L}(E)$ et $\lambda \in \mathbb{K}$, le sous-espace $\mathrm{Ker}\,(f - \lambda\,\mathrm{id}_E)$ sera noté $E_\lambda(f)$, ou plus simplement E_λ si f est clairement fixé. Il s'agit évidemment d'un sous-espace de E (c'est un noyau d'endomorphisme de E) et par définition,

$$E_\lambda = \mathrm{Ker}\,(f - \lambda\mathrm{id}_E) = \{x \in E \mid f x = \lambda x\}$$

Il s'agit donc du sous-espace de E, le plus grand possible, sur lequel f induit l'homothétie $\lambda\,\mathrm{id}_E$. Ces observations motivent l'introduction des notions d'éléments propres d'un endomorphisme de E.

Définition 4.1.1 (Valeurs propres, vecteurs propres)
Soient $f \in \mathscr{L}(E)$ et $\lambda \in \mathbb{K}$.

1. On dit que λ est *valeur propre de* f si, et seulement si, $E_\lambda(f) \neq \{0\}$.
2. L'ensemble des valeurs propres de f est appelé *le spectre de* f, et on le note $\mathrm{Sp}\,f$.
3. Si λ est valeur propre de f, le sous-espace $E_\lambda(f)$ est appelé *sous-espace propre de* E *pour* f *relatif à* λ.
4. Si λ est valeur propre de f, les vecteurs non-nuls de $E_\lambda(f)$ sont appelés *vecteurs propres de* f *relatifs à la valeur propre* λ.

Il convient de faire quelques remarques à la suite de ces définitions :

- Un vecteur propre n'est pas nul ; une valeur propre peut parfaitement être nulle. Plus précisément, 0 est valeur propre de f si, et seulement si, $E_0(f) \neq \{0\}$, ce qui équivaut à dire que f n'est pas injective.
- Insistons : un vecteur propre n'est pas nul. Plus précisément, un vecteur $x \in E$ est vecteur propre de f si, et seulement si, $x \neq 0$ et il existe $\lambda \in \mathbb{K}$ tel que $fx = \lambda x$. Auquel cas, le sous-espace propre $E_\lambda(f)$ est :
$$E_\lambda(f) = \{0\} \cup \{x \in E \mid x \text{ vecteur propre de } f \text{ relatif à } \lambda\}$$
- Si $\lambda \in \mathbb{K}$, le sous-espace $E_\lambda(f)$ est stable par f (que λ soit valeur propre de f ou pas). Il est nul si et seulement si λ n'est pas valeur propre de f. C'est un sous-espace propre si, et seulement si, il n'est pas nul, ce qui équivaut à dire que λ est valeur propre de f.
- Si E est de dimension finie n non nulle et si $\lambda \in \mathbb{K}$ est valeur propre de f, notons $p \in [\![1\,;n]\!]$ la dimension de $E_\lambda(f)$. On en prend une base $\mathscr{B}_\lambda = (e_1, \ldots, e_p)$, et on la complète (si nécessaire) en une base $\mathscr{B} = (e_1, \ldots, e_n)$ de E. Alors
$$\mathrm{Mat}_{\mathscr{B}}\, f = \begin{bmatrix} \lambda I_p & B \\ 0_{n-p,p} & C \end{bmatrix} \qquad \text{avec} \qquad B \in M_{p,n-p}(\mathbb{K}) \quad C \in M_{n-p}(\mathbb{K})$$

Ensuite, une proposition simple, mais fondamentale :

Proposition 4.1.2
Soient $f \in \mathscr{L}(E)$ et Λ un sous-ensemble de $\mathbb{K}$, non vide, fini. Les sous-espaces $(E_\lambda(f))_{\lambda\in\Lambda}$ sont en somme directe.

Preuve : On procède par récurrence en définissant, pour tout entier $n \geqslant 2$, l'assertion $\mathscr{P}(n)$: « Soient $f \in \mathscr{L}(E)$ et $\Lambda \subset \mathbb{K}$, fini, de cardinal n. Les sous-espaces $(E_\lambda(f))_{\lambda\in\Lambda}$ sont en somme directe. »

- $\mathscr{P}(2)$ est vraie : si λ_1 et λ_2 sont deux scalaire distincts, on se donne $x \in E_{\lambda_1}(f) \cap E_{\lambda_2}(f)$. Alors
$$fx = \lambda_1 x = \lambda_2 x \qquad \text{et} \qquad \underbrace{(\lambda_1 - \lambda_2)}_{\neq 0} x = 0$$
ce qui prouve que $x = 0$.
- Soit $n \geqslant 2$ un entier pour lequel $\mathscr{P}(n)$ est vraie. Soit $\Lambda \subset \mathbb{K}$, fini, de cardinal $n+1$. Soit $(x_\lambda)_{\lambda\in\Lambda} \in \prod_{\lambda\in\Lambda} E_\lambda(f)$, tel que
$$\sum_{\lambda\in\Lambda} x_\lambda = 0$$
On fixe un $\mu \in \Lambda$ et on note $\Lambda' = \Lambda \setminus \{\mu\}$, qui est de cardinal n. On a donc
$$x_\mu + \sum_{\lambda\in\Lambda'} x_\lambda = 0 \tag{1}$$
On applique f :
$$0 = \underbrace{f(x_\mu)}_{=\mu x_\mu} + \sum_{\lambda\in\Lambda'} \underbrace{f(x_\lambda)}_{=\lambda x_\lambda} = \mu x_\mu + \sum_{\lambda\in\Lambda'} \lambda x_\lambda \tag{2}$$
Puis on soustrait (**2**) à (**1**) multipliée par μ, de manière à se débarasser des termes en x_μ :
$$\sum_{\lambda\in\Lambda'} (\mu - \lambda) x_\lambda = 0$$

D'après l'hypothèse de récurrence, les $(\mathrm{E}_\lambda)_{\lambda\in\Lambda'}$ sont en somme directe donc

$$\forall\lambda\in\Lambda' \qquad \underbrace{(\mu-\lambda)}_{\neq 0}\, x_\lambda = 0$$

et
$$\forall\lambda\in\Lambda' \qquad x_\lambda = 0$$

Enfin, en reprenant (**1**), il reste $x_\mu = 0$ également. Ceci prouve $\mathscr{P}(n+1)$.

- Conclusion : Par récurrence, $\mathscr{P}(n)$ est vraie pour tout entier $n \geqslant 2$. □

Corollaire 4.1.3
Soit $n\in\mathbb{N}^\star$. On suppose E de dimension finie n. Soit $f\in\mathscr{L}(\mathrm{E})$. Alors $\mathrm{Sp}\, f$ est fini et $|\mathrm{Sp}\, f|\leqslant n$.

Preuve : Trivial : les sous-espaces propres sont de dimension au moins 1 ; et les sous-espaces propres relatifs à des valeurs propres distinctes deux-à-deux sont en somme directe. Donc il y a au plus n valeurs propres distinctes. □

Cette proposition, couplée avec la remarque faite plus haut sur la matrice de f dans une base adaptée à un sous-espace propre, montrent que plus on trouve de sous-espaces propres de f, plus on en simplifie la matrice dans une « bonne base ». Et intuitivement, si on peut décomposer E en somme directe de sous-espaces propres, alors la matrice de f dans une base adaptée sera d'une forme extrêmement agréable : diagonale.

Définition 4.1.4 (Endomorphisme diagonalisable)
Soit $f\in\mathscr{L}(\mathrm{E})$. On dit que f *est diagonalisable* si, et seulement si, $\mathrm{Sp}\, f$ n'est pas vide et E est somme directe des sous-espaces propres de f.

Proposition 4.1.5
On suppose E de dimension finie non nulle. Soit $f\in\mathscr{L}(\mathrm{E})$. Il est diagonalisable si, et seulement si, il existe une base de E dans laquelle la matrice de f est diagonale.

Preuve : Notons $n = \dim \mathrm{E} \neq 0$. Supposons f diagonalisable ; le spectre de f est fini, non vide. Notons $p = |\mathrm{Sp}\, f| \in [\![1\,;n]\!]$ et $\lambda_1,\ldots,\lambda_p$ les p valeurs propres de f, distinctes deux-à-deux. On se donne des bases $\mathscr{B}_1,\ldots,\mathscr{B}_p$ de $\mathrm{E}_{\lambda_1},\ldots,\mathrm{E}_{\lambda_p}$ respectivement ; comme $\mathrm{E} = \bigoplus_{k=1}^{p} \mathrm{E}_{\lambda_k}$, on obtient en les réunissant une base $\mathscr{B} = (\mathscr{B}_1,\ldots,\mathscr{B}_p)$ de E. Et

$$\mathrm{Mat}_{\mathscr{B}}\, f = \begin{bmatrix} \lambda_1 \mathrm{I}_{n_1} & & \\ & \ddots & \\ & & \lambda_p \mathrm{I}_{n_p} \end{bmatrix} \qquad \text{avec} \qquad \forall i\in[\![1\,;p]\!] \quad n_i = \dim \mathrm{E}_{\lambda_i}$$

Réciproquement, supposons qu'il existe une base $\mathscr{B} = (e_1,\ldots,e_n)$ dans laquelle la matrice A de f est diagonale. Notons $m_1,\ldots,m_n\in\mathbb{K}$ les coefficients diagonaux de cette matrice ainsi que $\mathrm{M} = \{m_1,\ldots,m_n\}$. Si $\lambda\in\mathbb{K}$,

$$\mathrm{Mat}_{\mathscr{B}}\,(f-\lambda\,\mathrm{id}_{\mathrm{E}}) = \mathrm{A} - \lambda\,\mathrm{I}_n = \begin{bmatrix} m_1-\lambda & & \\ & \ddots & m_n-\lambda \end{bmatrix}$$

- Si $\lambda\in\mathrm{M}$, alors $f-\lambda\mathrm{id}_{\mathrm{E}}$ a un noyau non trivial (une des colonnes au moins de cette matrice est nulle) : λ est valeur propre de f. En outre, si on note

$$\mathrm{I}_\lambda = \{k \in [\![1\,;n]\!] \mid m_k = \lambda\}$$

et
$$\mathrm{F}_\lambda = \mathrm{Vect}\,(e_k)_{k\in \mathrm{I}_\lambda}$$

alors $\mathrm{F}_\lambda \subset \mathrm{E}_\lambda(f)$ puisque pour chaque $k \in \mathrm{I}_\lambda$, la k-ème colonne de $\mathrm{A}-\lambda \mathrm{I}_n$ est nulle.

Mais d'autre part, le rang de cette matrice se lit aisément parce qu'elle est diagonale : c'est le nombre de coefficients diagonaux non nuls. C'est-à-dire $n - |\mathrm{I}_\lambda| = n - \dim \mathrm{F}_\lambda$. D'après le théorème du rang, $\dim \mathrm{E}_\lambda(f) = \dim \mathrm{F}_\lambda$: F_λ est le sous-espace propre relatif à λ.

- Si $\lambda \notin \mathrm{M}$, les colonnes de $\mathrm{A}-\lambda \mathrm{I}_n$ sont libres ; donc $f - \lambda\,\mathrm{id}_\mathrm{E}$ est de rang n, surjective, donc injective : λ n'est pas valeur propre de f.

On a donc prouvé que $\mathrm{M} = \mathrm{Sp}\, f$ et que pour tout $\lambda \in \mathrm{M}$, le sous-espace propre relatif à λ est F_λ, qui est de dimension $|\mathrm{I}_\lambda|$.

De plus, les $(\mathrm{I}_\lambda)_{\lambda\in\mathrm{M}}$ sont disjoints deux-à-deux et leur réunion est $[\![1\,;n]\!]$ (trivial d'après leur définition) ce qui prouve que

$$n = \sum_{\lambda\in\mathrm{M}} |\mathrm{I}_\lambda| = \sum_{\lambda\in\mathrm{Sp}\, f} \dim \mathrm{E}_\lambda(f)$$

D'où
$$\mathrm{E} = \bigoplus_{\lambda\in\mathrm{Sp}\, f} \mathrm{E}_\lambda(f)$$

et f est diagonalisable. □

Évidemment, on aimerait maintenant savoir deux choses :

1. Comment trouver les valeurs propres d'un endomorphisme ?
2. Y a-t-il des conditions permettant d'assurer qu'un endomorphisme est diagonalisable ?

4.1.2 Le polynôme caractéristique

Dans toute cette section, E est un $\mathbb{K}$-espace vectoriel de dimension finie n non nulle et f est un endomorphisme de E. D'après le théorème du rang et le cours sur le déterminant,

$$\begin{aligned}\forall \lambda \in \mathbb{K} \qquad \lambda \in \mathrm{Sp}\, f &\iff \mathrm{E}_\lambda(f) \neq \{0\}\\ &\iff (f - \lambda\,\mathrm{id}_\mathrm{E}) \text{ non injective}\\ &\iff (f - \lambda\,\mathrm{id}_\mathrm{E}) \text{ non bijective}\\ &\iff \det(f - \lambda\,\mathrm{id}_\mathrm{E}) = 0\end{aligned}$$

Ceci nous donne envie de considérer l'application $\lambda \longmapsto \det(f - \lambda\,\mathrm{id}_\mathrm{E})$ pour cherche où elle s'annule. Gardons ces réflexions de côté et commençons par étudier de quel type de fonction il s'agit.

Définition 4.1.6 (Polynôme caractéristique d'une matrice)
Soit $\mathrm{A} \in \mathrm{M}_n(\mathbb{K})$. On appelle *polynôme caractéristique de* A la fraction $\chi_\mathrm{A} = \det(\mathrm{A} - \mathrm{X}\mathrm{I}_n)$.

> Une remarque s'impose sur la nature de χ_A, et pourquoi on parle ici de « fraction ».
> On a développé notre théorie de l'algèbre linéaire autour d'un corps et d'un espace vectoriel sur ce corps. On ne sait calculer des déterminants que pour des matrices à coefficients dans un corps. La matrices A est dans $\mathrm{M}_n(\mathbb{K})$, tandis que $\mathrm{X}\mathrm{I}_n$ est dans $\mathrm{M}_n(\mathbb{K}(\mathrm{X}))$. Donc elles sont toutes les deux dans $\mathrm{M}_n(\mathbb{K}(\mathrm{X}))$, et c'est dans le corps $\mathbb{K}(\mathrm{X})$ qu'on calcule notre déter-

minant. *A priori*, sans réflexion supplémentaire, χ_{A} est donc un élément de $\mathbb{K}(\mathrm{X})$.

Évidemment, on va prouver qu'il s'agit en fait d'un polynôme, d'où le nom de *polynôme caractéristique de* A.

Proposition 4.1.7
Soit $\mathrm{A} \in \mathrm{M}_n(\mathbb{K})$. *Le polynôme caractéristique de* A *est un polynôme à coefficients dans* $\mathbb{K}$, *de degré* n. *De plus,*

- *son coefficient de degré n est* $(-1)^n$ *;*
- *son coefficient de degré* $n-1$ *est* $(-1)^{n-1}$ *Tr*A *;*
- *son coefficient de degré 0 est det*A.

Preuve : Notons $\mathrm{A} = (a_{i,j})_{1\leqslant i,j\leqslant n}$ et posons

$$\forall i,j \in [\![1\,;n]\!] \qquad b_{i,j} = \begin{cases} a_{i,i} - \mathrm{X} & \text{si } i = j \\ a_{i,j} & \text{si } i \neq j \end{cases} \qquad \mathrm{B} = (b_{i,j})_{1\leqslant i,j\leqslant n} \in \mathrm{M}_n(\mathbb{K}(\mathrm{X}))$$

On observe que $\quad \mathrm{B} = \mathrm{A} - \mathrm{XI}_n \quad$ et $\quad \forall i,j \in [\![1\,;n]\!] \quad \deg(b_{i,j}) = \begin{cases} 1 & \text{si } i = j \\ 0 & \text{si } i \neq j \end{cases}$

D'après notre formule préférée du cours sur les déterminants et en notant $\mathscr{B} = (e_1,\ldots,e_n)$ la base canonique de $(\mathbb{K}(\mathrm{X}))^n$,

$$\chi_{\mathrm{A}} = \sum_{i_1=1}^{n} \cdots \sum_{i_n=1}^{n} \Big(\prod_{k=1}^{n} b_{i_k,k}\Big) \det_{\mathscr{B}}(e_{i_1},\ldots,e_{i_n})$$

On peut observer que chaque terme dans cette somme horrible est un polynôme de degré n au plus. Donc $\chi_{\mathrm{A}} \in \mathbb{K}_n[\mathrm{X}]$. On s'intéresse ensuite aux termes qu'on peut calculer simplement. Si $(i_1,\ldots,i_n) \in [\![1\,;n]\!]^n$, on a

$$\deg \prod_{k=1}^{n} b_{i_k,k} = \sum_{k=1}^{n} \deg b_{i_k,k} = |\{k \in [\![1\,;n]\!] \mid i_k = k\}| \qquad (\star)$$

- Calcul du terme de degré n : Il vient immédiatement

  $$\Big(\deg \prod_{k=1}^{n} b_{i_k,k} = n\Big) \iff \Big(\big|\{k \in [\![1\,;n]\!] \mid i_k = k\}\big| = n\Big) \iff \Big(\forall k \in [\![1\,;n]\!] \quad i_k = k\Big)$$

  Donc le terme de degré n provient uniquement de

  $$\Big(\prod_{k=1}^{n} b_{k,k}\Big) \det_{\mathscr{B}}(e_1,\ldots,e_n) = \prod_{k=1}^{n} (a_{k,k} - \mathrm{X})$$

  Ce dernier produit un $(-\mathrm{X})^n$, ainsi qu'une contribution de degré $n-1$ qui est

  $$(-\mathrm{X})^{n-1} \sum_{k=1}^{n} a_{k,k} = (-\mathrm{X})^{n-1}\,\mathrm{Tr}\,\mathrm{A}$$

- Calcul du terme de degré $n-1$: On commence par rappeler que, parce que le déterminant est alterné, le déterminant $\det_{\mathscr{B}}(e_{i_1},\ldots,e_{i_n})$ est nul s'il existe $j \neq k$ dans $[\![1\,;n]\!]$ tels que $i_j = i_k$. On peut donc supposer que l'application $j \longmapsto i_j$ est injective de $[\![1\,;n]\!]$ dans lui-même ; et donc bijective.

Supposons que cette application n'est pas l'identité. Il existe alors $j \in [\![1\,;\,n]\!]$ tel que $i_j \neq j$; et il existe aussi $k \in [\![1\,;\,n]\!]$ tel que $i_k = j$. Bien sûr, il est impossible que $k = j$, puisque $i_k \neq i_j$. Tout cela pour dire que

$$\left|\{m \in [\![1\,;\,n]\!] \mid i_m = m\}\right| \leqslant n-2$$

car cet ensemble ne contient ni k, ni j. D'après $(\star)$,

$$\deg \prod_{k=1}^{n} b_{i_k,k} \leqslant n-2 < n-1$$

Dans le développement de χ_A, le seul terme de degré $n-1$ est donc celui déjà calculé plus haut : $(-X)^{n-1}\operatorname{Tr}A$.

- Calcul du terme de degré 0 : Là, c'est facile, parce que χ_A est un polynôme donc son coefficient de degré 0 est $\chi_A(0)$:

$$\chi_A(0) = \sum_{i_1=1}^{n} \cdots \sum_{i_n=1}^{n} \left(\prod_{k=1}^{n} \underbrace{b_{i_k,k}(0)}_{=a_{i_k,k}}\right) \det_{\mathscr{B}}(e_{i_1},\dots,e_{i_n}) = \det A \qquad \square$$

La proposition suivante est triviale, conséquence des propriétés connues du déterminant.

Proposition 4.1.8
Soit $A \in M_n(\mathbb{K})$.

- A *et* ${}^tA$ *ont le même polynôme caractéristique.*
- *Pour toute matrice* $P \in GL_n(\mathbb{K})$, A *et* PAP^{-1} *ont le même polynôme caractéristique.*

Maintenant, si $\mathscr{B}$ et $\mathscr{C}$ sont deux bases de E, et si on note M et N les matrices de f dans les bases $\mathscr{B}$ et $\mathscr{C}$ respectivement, on a $M = PNP^{-1}$, avec $P = P_{\mathscr{B}}^{\mathscr{C}}$. La proposition précédente permet alors de définir le polynôme caractéristique de f :

Définition 4.1.9 (Polynôme caractéristique d'un endomorphisme)
On appelle *polynôme caractéristique de* f le polynôme caractéristique de $\operatorname{Mat}_{\mathscr{B}} f$ lorsque $\mathscr{B}$ est une base de E. Sa valeur ne dépend pas du choix de $\mathscr{B}$, et on le note χ_f.

- χ_f est un polynôme de degré n ;
- son coefficient de degré n est $(-1)^n$;
- son coefficient de degré $n-1$ est $(-1)^{n-1}\operatorname{Tr} f$;
- son coefficient de degré 0 est $\det f$.

La relation fondamentale entre valeurs propres et polynôme caractéristique a été expliquée au début de cette section :

Théorème 4.1.10
Soit $\lambda \in \mathbb{K}$. *C'est une valeur propre de* f *si, et seulement si,* λ *est racine de* χ_f.

On voit tout de suite qu'il est possible que f n'ait aucune valeur propre : il suffit pour cela que χ_f n'ait pas de racine dans $\mathbb{K}$. Cela peut arriver : par exemple l'endomorphisme de $\mathbb{R}^2$ canoniquement associé à $\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ a pour polynôme caractéristique X^2+1.

4.1.3 Multiplicités d'une valeur propre

Ici, f est toujours un endomorphisme fixé de E, espace de dimension finie n. On peut associer aux valeurs propres de f deux notions de multiplicité : l'une algébrique (sa multiplicité dans la polynôme caractéristique), l'autre géométrique (la dimension du sous-espace propre associé). Il y a une relation entre ces deux entiers, qu'on peut utiliser pour dire si f est diagonalisable.

Définition 4.1.11 (Multiplicités d'une valeur propre)
Soit $\lambda \in \mathbb{K}$ une valeur propre de f.

- On appelle *multiplicité algébrique de* λ son ordre de multiplicité comme racine de χ_f. Généralement, elle sera notée m_λ dans la suite de ce cours.
- On appelle *multiplicité géométrique de* λ la dimension du sous-espace propre $E_\lambda(f)$. Elle sera notée n_λ lorsque rien n'est précisé.

On remarquera que, par définition :

- si $\lambda \in \mathrm{Sp}\, f$, alors m_λ et n_λ sont deux entiers, tous deux non nuls.
- si χ_f est scindé dans $\mathbb{K}[X]$, alors $\sum\limits_{\lambda \in \mathrm{Sp}\, f} m_\lambda = n$ et

$$\chi_f = (-1)^n \prod_{\lambda \in \mathrm{Sp}\, f} (X-\lambda)^{m_\lambda} = \prod_{\lambda \in \mathrm{Sp}\, f} (\lambda - X)^{m_\lambda}$$

De plus, on a très facilement (relations coefficients-racines) :

Proposition 4.1.12
On suppose que χ_f est scindé.

- *Le produit des valeurs propres de f (comptées avec multiplicités algébriques) est $\det f$.*
- *La somme des valeurs propres de f (comptées avec multiplicités algébriques) est $\mathrm{Tr}\, f$.*

Autrement dit,

$$\sum_{\lambda \in Sp f} m_\lambda \lambda = Tr f \qquad \prod_{\lambda \in Sp f} \lambda^{m_\lambda} = det f$$

Tout aussi facilement :

Théorème 4.1.13
On suppose que χ_f est scindé à racines simples. Alors f est diagonalisable.

Preuve : Si χ_f est scindé à racines simples, f a n valeurs propres distinctes. Les sous-espaces propres associés sont de dimension au moins 1, en somme directe. Donc leur somme (directe) est de dimension au moins n. Mais E est de dimension n : leur somme est précisément égale à E, et f est diagonalisable. □

Proposition 4.1.14
La multiplicité géométrique est toujours inférieure à la multiplicité algébrique d'une valeur propre. Autrement dit,

$$\forall \lambda \in Sp f \qquad 1 \leqslant n_\lambda \leqslant m_\lambda$$

Preuve : Soit λ valeur propre de f. On se donne $(e_1, \ldots, e_{n_\lambda})$, base de $E_\lambda(f)$, qu'on complète en une base de E notée $\mathscr{B} = (e_1, \ldots, e_n)$. Alors

$$\mathrm{Mat}_{\mathscr{B}} f = \begin{bmatrix} \lambda I_{n_\lambda} & B \\ 0 & D \end{bmatrix}$$

avec les matrices B et D qui ont la bonne taille. De ce fait,

$$\chi_f = (\lambda - X)^{n_\lambda} \det(D - XI_{n-n_\lambda}) = (\lambda - X)^{n_\lambda} \chi_D$$

et λ est de multiplicité algébrique au moins n_λ ; c'est-à-dire que $n_\lambda \leqslant m_\lambda$. □

Enfin, un critère de diagonalisabilité :

Théorème 4.1.15
f est diagonalisable si, et seulement si, χ_f est scindé dans $\mathbb{K}[X]$ et, pour chaque valeur propre, les multiplicités algébrique et géométrique sont égales.

Preuve : Supposons f diagonalisable ; dans une base de E, sa matrice est diagonale et on voit immédiatement que χ_f est scindé. Ensuite, E est somme directe des sous-espaces propres, et χ_f est de dimension n donc

$$n = \dim E = \sum_{\lambda \in \mathrm{Sp}\, f} m_\lambda \leqslant \sum_{\lambda \in \mathrm{Sp}\, f} m_\lambda \leqslant n$$

Ces inégalités sont alors toutes des égalités et il s'ensuit que

$$\forall \lambda \in \mathrm{Sp}\, f \qquad n_\lambda = m_\lambda$$

Réciproquement, si cette condition est vérifiée et χ_f est scindé, on a

$$n = \sum_{\lambda \in \mathrm{Sp}\, f} m_\lambda = \sum_{\lambda \in \mathrm{Sp}\, f} n_\lambda = \sum_{\lambda \in \mathrm{Sp}\, f} \dim E_\lambda(f) = \dim\Big(\bigoplus_{\lambda \in \mathrm{Sp}\, f} E_\lambda(f)\Big)$$

d'où
$$E = \bigoplus_{\lambda \in \mathrm{Sp}\, f} E_\lambda(f)$$
□

Ainsi, le plan de réduction d'un endomorphisme f est relativement simple :

1. Calculer le polynôme caractéristique et le factoriser. Si χ_f n'est pas scindé, f n'est pas diagonalisable.
2. Si χ_f est scindé, pour chaque racine λ, calculer le sous-espace propre $E_\lambda(f)$, c'est-à-dire rechercher $\mathrm{Ker}(f - \lambda\, \mathrm{id}_E)$. Trouver ensuite une base pour cet espace.
3. Si les multiplicités algébriques sont égales aux multiplicités géométriques, f est diagonalisable. En réunissant les bases des sous-espaces propres, on obtient une base $\mathscr{B}$ de E dans laquelle f a une matrice diagonale.
   Si f est donné par une matrice $n \times n$, il sera peut-être utile de calculer la matrice de passage de la base canonique de $\mathbb{K}^n$ vers $\mathscr{B}$, ainsi que son inverse.

Exemple 4.1.16
Supposons qu'on veuille réduire dans $M_3(\mathbb{R})$ la matrice

$$A = \begin{bmatrix} 0 & 2 & -1 \\ 3 & -2 & 0 \\ -2 & 2 & 1 \end{bmatrix}$$

Avant de commencer les calculs, on observe quand même A pendant une minute pour voir si certaines choses évidentes (noyau, valeurs propres triviales) apparaissent. Mais ce n'est pas le cas, donc on calcule le polynôme caractéristique. Notre but est de le factoriser, donc on essaie de le calculer intelligemment : pas de développement direct, mais des opérations sur les lignes/colonnes. Il n'y a pas de garantie qu'on puisse faire apparaître des racines évidentes ; mais c'est mieux que de développer comme un idiot.

$$\chi_A = \begin{vmatrix} -X & 2 & -1 \\ 3 & -2-X & 0 \\ -2 & 2 & 1-X \end{vmatrix} \underset{\substack{C_1 \leftarrow C_1 - XC_3 \\ C_2 \leftarrow C_2 + 2C_3}}{=} \begin{vmatrix} 0 & 0 & -1 \\ 3 & -2-X & 0 \\ X^2-X-2 & 4-2X & 1-X \end{vmatrix} = -\begin{vmatrix} 3 & -2-X \\ X^2-X-2 & 4-2X \end{vmatrix}$$

$$= (2-X)\begin{vmatrix} 3 & -2-X \\ X+1 & -2 \end{vmatrix} = (2-X)\big(-6+(X+2)(X+1)\big) = (2-X)(X^2+3X-4) = (2-X)(1-X)(-4-X)$$

χ_A est scindé à racines simples dans $\mathbb{R}$; A est diagonalisable. On calcule ses sous-espaces propres :

- Recherche de $E_2 = \operatorname{Ker}(A-2I_3)$. On effectue des opérations sur les lignes de $A-2I_3$:

$$A-2I_3 = \begin{bmatrix} -2 & 2 & -1 \\ 3 & -4 & 0 \\ -2 & 2 & -1 \end{bmatrix} \xrightarrow[L_1 \leftarrow L_1+L_2]{L_3 \leftarrow L_3 - L_1} \begin{bmatrix} 1 & -2 & -1 \\ 3 & -4 & 0 \\ 0 & 0 & 0 \end{bmatrix} \xrightarrow{L_2 \leftarrow L_2 - 3L_1} \begin{bmatrix} 1 & -2 & -1 \\ 0 & 2 & 3 \\ 0 & 0 & 0 \end{bmatrix} \xrightarrow{L_1 \leftarrow L_1 + L_2} \begin{bmatrix} 1 & 0 & 2 \\ 0 & 2 & 3 \\ 0 & 0 & 0 \end{bmatrix}$$

  et on voit que $$\operatorname{Ker}(A-2I_3) = \operatorname{Vect} u_1 \qquad \text{avec} \qquad u_1 = \begin{bmatrix} 4 \\ 3 \\ -2 \end{bmatrix}$$

- Recherche de $E_1 = \operatorname{Ker}(A-I_3)$. Même travail qu'avant : à l'aide d'opérations sur les lignes, on peut transformer $A-I_3$ en

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

  d'où $$\operatorname{Ker}(A-I_3) = \operatorname{Vect} u_2 \qquad \text{avec} \qquad u_2 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

- Recherche de $E_{-4} = \operatorname{Ker}(A+4I_3)$. Même chose : à l'aide d'opérations sur les lignes, on transforme $A+4I_3$ en

$$\begin{bmatrix} 1 & 0 & -1 \\ 0 & 2 & 3 \\ 0 & 0 & 0 \end{bmatrix}$$

  et $$\operatorname{Ker}(A+4I_3) = \operatorname{Vect} u_3 \qquad \text{avec} \qquad u_3 = \begin{bmatrix} 2 \\ -3 \\ 2 \end{bmatrix}$$

On n'oublie surtout pas de vérifier rapidement que

$$Au_1 = 2u_1 \qquad Au_2 = u_2 \qquad Au_3 = -4u_3$$

Évidemment, ces formules doivent être vraies ; mais cette vérification permet de s'assurer qu'on n'a pas fait d'erreur de calcul. Si on note P la matrice de passage de la base canonique de $\mathbb{R}^3$ vers (u_1, u_2, u_3), on a

$$A = P\begin{bmatrix} 2 & & \\ & 1 & \\ & & -4 \end{bmatrix}P^{-1} \qquad \text{avec} \qquad P = \begin{bmatrix} 4 & 1 & 2 \\ 3 & 1 & -3 \\ -2 & 1 & 2 \end{bmatrix}$$

À partir de là, tout dépend de ce qu'on veut faire. Diagonaliser A permet en particulier de calculer facilement ses puissances puisque

$$\forall n \in \mathbb{N} \qquad \mathrm{A}^n = \mathrm{P}\begin{bmatrix} 2^n & & \\ & 1 & \\ & & (-4)^n \end{bmatrix}\mathrm{P}^{-1}$$

On aura évidemment besoin de calculer P^{-1}. Ici, ce n'est pas un problème puisqu'on a des matrices de petite taille.

Exemple 4.1.17
Prenons encore un exemple simple : réduire dans $\mathrm{M}_2(\mathbb{R})$ la matrice

$$\mathrm{A} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$$

On a $$\chi_{\mathrm{A}} = (-\mathrm{X})^2 + (\mathrm{Tr}\,\mathrm{A})\,(-\mathrm{X}) + \det \mathrm{A} = \mathrm{X}^2 + 1$$

qui n'est pas scindé dans $\mathbb{R}[\mathrm{X}]$. Donc $\mathrm{Sp}_{\mathbb{R}}(\mathrm{A}) = \varnothing$ et A n'est pas diagonalisable dans $\mathrm{M}_2(\mathbb{R})$

Mais A est aussi dans $\mathrm{M}_2(\mathbb{C})$ et on peut s'intéresser à sa diagonalisabilité sur le corps des complexes. Dans $\mathbb{C}[\mathrm{X}]$, χ_{A} est scindé :

$$\chi_{\mathrm{A}} = (\mathrm{X}-\mathrm{i})(\mathrm{X}+\mathrm{i}) \qquad \mathrm{Sp}_{\mathbb{C}}(\mathrm{A}) = \{\pm \mathrm{i}\}$$

À nouveau, les racines sont simples donc on sait que A est diagonalisable sur $\mathbb{C}$. Le calcul des sous-espaces propres est le même que dans l'exemple précédent :

$$\mathrm{A} = \mathrm{P}\begin{bmatrix} \mathrm{i} & \\ & -\mathrm{i} \end{bmatrix}\mathrm{P}^{-1} \qquad \text{avec} \qquad \mathrm{P} = \begin{bmatrix} 1 & 1 \\ -\mathrm{i} & \mathrm{i} \end{bmatrix}$$

4.1.4 Endomorphismes trigonalisables

Dans ce qui suit, E est un espace de dimension finie n non nulle, et f est un endomorphisme de E. On a vu que f est diagonalisable si, et seulement si, son polynôme caractéristique est scindé et ses valeurs propres ont les mêmes multiplicités algébrique et géométrique.

Ainsi, même si χ_f est scindé, il reste possible que f ne soit pas diagonalisable, parce qu'un (au moins) des sous-espaces propres n'est pas assez grand. Essentiellement, cela veut dire qu'on demande quelque chose de trop fort (une base où f a une matrice diagonale) ; on peut essayer d'être moins gourmand et se contenter d'une base où f a une matrice triangulaire.

Définition 4.1.18 (Endomorphismes trigonalisables)
On dit que f est trigonalisable si, et seulement si, il existe une base de E dans laquelle la matrice de f est triangulaire supérieure.

Cette notion est beaucoup moins astreignante que celle de diagonalisabilité, puisqu'on a :

Théorème 4.1.19 (Caractérisation des endomorphismes trigonalisables)
f est trigonalisable si, et seulement si, χ_f est scindé.

Preuve : Si f est trigonalisable, il existe une base $\mathscr{B}$ de E telle que $\mathrm{Mat}_{\mathscr{B}} f$ est triangulaire supérieure. En notant $\mathrm{M} = (m_{i,j})_{1 \leqslant i,j \leqslant n}$ cette matrice, le polynôme caractéristique de f se calcule facilement :

$$\chi_f = \chi_M = \prod_{k=1}^{n} (m_{k,k} - X)$$

et il est scindé.

On va prouver la réciproque par récurrence sur la dimension de E. Plus précisément, si n est un entier non nul, on note $\mathscr{P}(n)$ l'assertion : « Si E est un $\mathbb{K}$-espace vectoriel de dimension n et $f \in \mathscr{L}(E)$ a son polynôme caractéristique scindé, alors f est diagonalisable. »

$\mathscr{P}(1)$ est trivialement vraie, et on se donne un entier n non nul pour lequel $\mathscr{P}(n)$ est vraie. Soient E un $\mathbb{K}$-espace vectoriel de dimension $n+1$ et $f \in \mathscr{L}(E)$, avec un polynôme caractéristique scindé. On peut alors prendre λ_1 racine de χ_f ; et λ_1 est aussi valeur propre de f, de sorte qu'on peut trouver e_1 non nul tel que $f e_1 = \lambda_1 e_1$. On complète alors en une base $\mathscr{B} = (e_1, \ldots, e_{n+1})$ de E et on a

$$\mathrm{Mat}_{\mathscr{B}} f = \begin{bmatrix} \lambda_1 & L \\ 0_{n,1} & M \end{bmatrix} \qquad \text{avec} \qquad L \in M_{1,n}(\mathbb{K}) \quad M \in M_n(\mathbb{K})$$

Du coup,

$$\chi_f = (\lambda_1 - X)\,\chi_M$$

et on voit que $\chi_M | \chi_f$. Donc χ_M est scindé. D'après l'hypothèse de récurrence, M (vu comme endomorphisme de $\mathbb{K}^n$) est trigonalisable : il existe une matrice $P \in GL_n(\mathbb{K})$ et une matrice T triangulaire supérieure telles que $M = PTP^{-1}$. Un simple calcul prouve alors que

$$\begin{bmatrix} 1 & 0_{1,n} \\ 0_{n,1} & P \end{bmatrix} \begin{bmatrix} \lambda_1 & LP \\ 0_{n,1} & T \end{bmatrix} \begin{bmatrix} 1 & 0_{1,n} \\ 0_{n,1} & P^{-1} \end{bmatrix} = \mathrm{Mat}_{\mathscr{B}} f$$

De plus,

$$\begin{bmatrix} 1 & 0_{1,n} \\ 0_{n,1} & P^{-1} \end{bmatrix} = \begin{bmatrix} 1 & 0_{1,n} \\ 0_{n,1} & P \end{bmatrix}^{-1}$$

Il existe donc une base $\mathscr{C}$ de E (la matrice de passage de $\mathscr{B}$ vers $\mathscr{C}$ est $\begin{bmatrix} 1 & 0_{1,n} \\ 0_{n,1} & P \end{bmatrix}$) dans laquelle f a une matrice triangulaire supérieure. Ceci prouve $\mathscr{P}(n+1)$.

Par récurrence, $\mathscr{P}(n)$ est vraie pour tout $n \in \mathbb{N}^\star$, ce qui achève la preuve du théorème. □

Une conséquence immédiate est la suivante, à l'aide du théorème de d'Alembert :

Corollaire 4.1.20
Tout endomorphisme d'un $\mathbb{C}$-espace vectoriel de dimension finie non nulle est trigonalisable.

Une méthode pour trigonaliser une application linéaire dont le polynôme caractéristique est scindé est donnée par la preuve du **théorème 1.19**. Une étude attentionnée montrera que cette méthode n'est probablement pas optimale : en effet, le choix de la valeur propre λ_1 et de la base $\mathscr{B}$ sont assez arbitraires et on peut penser qu'il y a des choix plus intéressants que d'autres. Voyons ça sur un exemple.

Exemple 4.1.21
Considérons l'endomorphisme de $\mathbb{R}^3$ défini par la matrice

$$A = \begin{bmatrix} 2 & 1 & -2 \\ 3 & 4 & -10 \\ 2 & 2 & -5 \end{bmatrix}$$

On calcule

$$\chi_A = -(1-X)^2(1+X) \qquad \mathrm{Sp}_{\mathbb{R}} A = \{-1\,;\,1\}$$

puis $\quad \operatorname{Ker}(\mathrm{A}-\mathrm{I}_3) = \operatorname{Vect} u_1 \quad \operatorname{Ker}(\mathrm{A}+\mathrm{I}_3) = \operatorname{Vect} u_2 \quad$ avec $\quad u_1 = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} \quad u_2 = \begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix}$

χ_A est scindé dans $\mathbb{R}[\mathrm{X}]$, mais 1 est valeur propre double, tandis que le sous-espace associé est de dimension 1. Donc A n'est pas diagonalisable. Dans un premier temps, on applique mot-pour-mot la preuve du **théorème 1.19**. On pose

$$v_1 = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} \qquad v_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \qquad v_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \qquad \mathrm{P} = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

et on vérifie facilement que $\mathscr{B} = (v_1, v_2, v_3)$ est une base de $\mathbb{R}^3$, par exemple parce que P est inversible, d'inverse

$$\mathrm{P}^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Par suite, $\quad \operatorname{Mat}_{\mathscr{B}} \mathrm{A} = \mathrm{P}^{-1}\mathrm{AP} = \begin{bmatrix} 1 & 1 & -2 \\ 0 & 5 & -12 \\ 0 & 2 & -5 \end{bmatrix}$

On trigonalise alors la matrice $\begin{bmatrix} 5 & -12 \\ 2 & -5 \end{bmatrix}$. On trouve (comme prévu, voir la preuve du **théorème 1.19**) que son polynôme caractéristique est $\mathrm{X}^2 - 1$ et on n'a aucun problème pour la diagonaliser :

$$\begin{bmatrix} 5 & -12 \\ 2 & -5 \end{bmatrix} = \begin{bmatrix} 3 & 2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 1 & -2 \\ -1 & 3 \end{bmatrix}$$

D'où $\quad \operatorname{Mat}_{\mathscr{B}} \mathrm{A} = \begin{bmatrix} 1 & & \\ & 3 & 2 \\ & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & \\ & 1 & \\ & & -1 \end{bmatrix} \begin{bmatrix} 1 & & \\ & 1 & -2 \\ & -1 & 3 \end{bmatrix}$

et $\quad \mathrm{A} = \mathrm{P}(\operatorname{Mat}_{\mathscr{B}} \mathrm{A})\mathrm{P}^{-1} = \begin{bmatrix} 1 & & \\ -1 & 3 & 2 \\ & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & \\ & 1 & \\ & & -1 \end{bmatrix} \begin{bmatrix} 1 & & \\ 1 & 1 & -2 \\ -1 & -1 & -3 \end{bmatrix}$

On a trigonalisé A : en posant

$$m_1 = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} \qquad m_2 = \begin{bmatrix} 0 \\ 3 \\ 1 \end{bmatrix} \qquad m_3 = \begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix}$$

l'endomorphisme de $\mathbb{R}^3$ a une matrice triangulaire supérieure dans la base (m_1, m_2, m_3) et on a

$$\mathrm{A}m_1 = m_1 \qquad \mathrm{A}m_2 = m_1 + m_2 \qquad \mathrm{A}m_3 = -m_3$$

Exemple 4.1.22
On reprend l'exemple précédent, à partir du moment où on s'est aperçu que A n'est pas diagonalisable. Au lieu d'appliquer la preuve du **théorème 1.19**, on pose

$$v_1 = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} \qquad v_2 = \begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix} \qquad v_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

C'est-à-dire qu'on a choisi pour v_1 et v_2 les vecteurs propres de A déjà calculés (relatifs aux valeurs propres 1 et -1) ; et on a pris un vecteur v_3 qui permet d'avoir une base $\mathscr{B} = (v_1, v_2, v_3)$ de $\mathbb{R}^3$. On calcule la matrice de passage de la base canonique vers $\mathscr{B}$ et son inverse :

$$P = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 2 & 0 \\ 0 & 1 & 1 \end{bmatrix} \qquad P^{-1} = \frac{1}{2}\begin{bmatrix} 2 & & \\ 1 & 1 & \\ -1 & -1 & 2 \end{bmatrix}$$

ce qui permet d'obtenir $$[Av_3]_{\mathscr{B}} = P^{-1}Av_3 = P^{-1}\begin{bmatrix} -2 \\ -10 \\ -5 \end{bmatrix} = \begin{bmatrix} -2 \\ -6 \\ 1 \end{bmatrix}$$

d'où $$A = P\begin{bmatrix} 1 & & -2 \\ & -1 & -6 \\ & & 1 \end{bmatrix}P^{-1}$$

À nouveau, on a triagonalisé A. Mais cette fois, parce qu'on a choisi notre base $\mathscr{B}$ de manière plus intelligente, la procédure était plus rapide.

Cet exemple montre que deux méthodes de trigonalisation différentes aboutiront à des formes triangulaires nettement différentes. Au contraire, si l'endomorphisme est diagonalisable, toutes les formes diagonales seront « presque » la même : la seule liberté qu'on a est l'ordre dans lequel on présente les éléments diagonaux (les valeurs propres).

À titre indicatif, on fait remarquer qu'il existe une méthode de trigonalisation « optimale », qui fournit une matrice triangulaire supérieure avec « le plus grand nombre de zéros possible. » Il s'agit de la forme de Jordan, dont l'étude n'est malheureusement pas au programme. Elle sera néanmoins présentée, pour référence, à la fin du chapitre. Dans les deux exemples ci-dessus, le premier a fourni la forme de Jordan de A, mais c'est totalement par hasard : la preuve du **théorème 1.19** ne fait pas attention au nombre de zéros dans la forme triangulaire supérieure obtenue.

4.2 Polynômes d'endomorphismes

La partie précédente présentait surtout l'aspect calculatoire de la théorie de la réduction des endomorphismes : comment diagonaliser et trigonaliser un endomorphisme, en pratique. La partie théorique restait extrêmement simple.

Dans cette section, on pousse plus loin l'étude théorique, en faisant agir l'anneau $\mathbb{K}[X]$ sur un endomorphisme de E.

4.2.1 Sous-espaces stables

Rappelons d'abord la notion de sous-espaces stable et quelques propriétés. Dans ce qui suit, E est un $\mathbb{K}$-espace vectoriel et f est un endomorphisme de E.

Définition 4.2.1 (Sous-espace stable)
Soit F un sous-espace de E. On dit que F *est stable par f* si, et seulement si, $f(F) \subset F$, c'est-à-dire

$$\forall x \in F \qquad fx \in F$$

Dans ce cas, on peut définir un endomorphisme $\tilde{f}$ de F en posant

$$\begin{array}{rrcl} \tilde{f}: & \mathrm{F} & \longrightarrow & \mathrm{F} \\ & x & \longmapsto & fx \end{array}$$

$\tilde{f}$ est appelé *l'endomorphisme induit par f sur* F.

On fera bien attention au fait que, si F est stable par f, l'endomorphisme induit par f sur F **n'est pas** la restriction de f à F. Précisément, c'est la restriction de f à F comme ensemble de départ **et** d'arrivée.

L'endomorphisme induit par f sur F est dans $\mathscr{L}(\mathrm{F})$. La restriction de f à F est dans $\mathscr{L}(\mathrm{F}, \mathrm{E})$. Si E est de dimension finie, on peut remarquer que le premier sera représenté par des matrices carrées ; le second par des matrices rectangulaires. Cela devrait suffire à convaincre qu'il ne s'agit pas du même objet mathématique.

L'intérêt des sous-espaces stables est évident lorsqu'on adopte un point de vue matriciel : il permet d'avoir des matrices avec beaucoup de zéros.

Proposition 4.2.2 (Sous-espaces stables : interprétation matricielle)
On suppose E *de dimension finie $n \geqslant 2$. Soit* F *un sous-espace de* E *de dimension $p \in [[1\,;\, n-1]]$, avec une base $(e_1, \ldots, e_p)$ complétée en $\mathscr{B} = (e_1, \ldots, e_n)$ base de* E. *Alors* F *est stable par f si, et seulement si, $Mat_{\mathscr{B}}\, f$ a la forme*

$$\begin{bmatrix} \mathrm{A} & \mathrm{B} \\ 0_{n-p,p} & \mathrm{D} \end{bmatrix} \qquad \textit{avec} \quad \mathrm{A} \in \mathrm{M}_p(\mathbb{K}) \quad \mathrm{B} \in \mathrm{M}_{p,n-p}(\mathbb{K}) \quad \mathrm{D} \in \mathrm{M}_{n-p}(\mathbb{K})$$

Dans ce cas, A *est la matrice dans $(e_1, \ldots, e_p)$ de l'endomorphisme $\tilde{f}$ de* F *induit par f. De plus, $\chi_{\tilde{f}}$ divise χ_f.*

Tout ceci est une conséquence triviale des définitions. On peut même généraliser cette observation au cas de plusieurs sous-espaces stables. Et en fait, la situation « idéale » est lorsqu'on peut décomposer f en somme directe de sous-espaces stables : les endomorphismes induits par f sur chacun de ces sous-espaces sont, intuitivement, plus simples que f. Précisément :

Proposition 4.2.3
On suppose E *de dimension finie. Soient $p \in \mathbb{N}^\star$, des sous-espaces $\mathrm{E}_1, \ldots, \mathrm{E}_p$, non nuls, en somme directe et tels que $\bigoplus_{k=1}^{p} \mathrm{E}_k = \mathrm{E}$. On note $q_1, \ldots, q_p$ les projections associées à cette décomposition et*

$$\forall k \in [[1\,;\, p]] \qquad n_k = dim\, \mathrm{E}_k$$

On prend des bases $\mathscr{B}_1, \ldots, \mathscr{B}_p$ de $\mathrm{E}_1, \ldots, \mathrm{E}_p$, qu'on réunit en une base $\mathscr{B} = (\mathscr{B}_1, \ldots, \mathscr{B}_p)$ de E. *Alors $\mathrm{E}_1, \ldots, \mathrm{E}_p$ sont tous stables par f si, et seulement si,*

$$Mat_{\mathscr{B}}\, f = \begin{bmatrix} \mathrm{A}_1 & & \\ & \ddots & \\ & & \mathrm{A}_p \end{bmatrix} \qquad \textit{avec} \qquad \forall k \in [[1\,;\, p]] \quad \mathrm{A}_k \in \mathrm{M}_{n_k}(\mathbb{K})$$

Dans ce cas, en notant $f_1, \ldots, f_p$ les endomorphismes induits par f sur $\mathrm{E}_1, \ldots, \mathrm{E}_p$, on a

$$\forall k \in [[1\,;\, p]] \qquad \mathrm{A}_k = Mat_{\mathscr{B}_k}\, f_k$$

En particulier,

$$\chi_f = \prod_{k=1}^{p} \chi_{f_k}$$

et $$\forall x \in \mathrm{E} \qquad f x = \sum_{k=1}^{p} f_k(q_k x)$$

Des endomorphismes qui commutent ont le bon goût de stabiliser leurs sous-espaces caractéristiques mutuels :

Proposition 4.2.4
Soient f et g deux endomorphismes de E, tels que $fg = gf$. Alors pour tout $\lambda \in \mathbb{K}$, $Ker(f - \lambda\, id_{\mathrm{E}})$ et $Im(f - \lambda\, id_{\mathrm{E}})$ sont stables par g.

Preuve : C'est trivial.

- Si $x \in \operatorname{Ker} f$, alors $fgx = gfx = 0$ donc $gx \in \operatorname{Ker} f$.
- Si $y \in \operatorname{Im} f$, il existe $x \in \mathrm{E}$ tel que $y = fx$. Du coup, $gy = gfx = fgx$ donc $gy \in \operatorname{Im} f$.

Ensuite, si $\lambda \in \mathbb{K}$, $f - \lambda\,\mathrm{id}$ et g commutent, donc on applique ce qu'on vient de prouver à ces deux endomorphismes de E. □

4.2.2 Polynômes d'un endomorphisme

Si $f \in \mathscr{L}(\mathrm{E})$ et $\mathrm{P} \in \mathbb{K}[\mathrm{X}]$, on définit un nouvel endomorphisme de E, noté $\mathrm{P}(f)$:

$$\mathrm{P}(f) = \sum_{k=0}^{\infty} \mathrm{P}_k f^k$$

Ceci définit une application $$\Phi : \begin{array}{rcl} \mathbb{K}[\mathrm{X}] & \longrightarrow & \mathscr{L}(\mathrm{E}) \\ f & \longmapsto & \mathrm{P}(f) \end{array}$$

Proposition 4.2.5
Φ est un morphisme d'algèbres :

$$\forall \mathrm{P},\mathrm{Q} \in \mathbb{K}[\mathrm{X}] \quad \forall \lambda \in \mathbb{K} \qquad (\lambda\mathrm{P} + \mathrm{Q})(f) = \lambda\,\mathrm{P}(f) + \mathrm{Q}(f)$$

$$\forall \mathrm{P},\mathrm{Q} \in \mathbb{K}[\mathrm{X}] \qquad (\mathrm{PQ})(f) = (\mathrm{P}(f)) \circ (\mathrm{Q}(f)) = (\mathrm{Q}(f)) \circ (\mathrm{P}(f))$$

Preuve : C'est une simple question de vérifier les axiomes. La linéarité est absolument triviale. La compatibilité avec le produit semble (bizarrement) poser des problèmes aux élèves, donc prouvons-la. Soient P et Q dans $\mathbb{K}[\mathrm{X}]$. On a

$$\mathrm{P}(f) = \sum_{k=0}^{\infty} \mathrm{P}_k f^k \qquad \mathrm{Q}(f) = \sum_{k=0}^{\infty} \mathrm{Q}_k f^k$$

Comme $(\mathscr{L}(\mathrm{E}), +, \circ)$ est un anneau,

$$(\mathrm{Q}(f)) \circ (\mathrm{P}(f)) = \sum_{k=0}^{\infty}\sum_{\ell=0}^{\infty} \mathrm{Q}_k \mathrm{P}_\ell f^{k+\ell}$$

et comme $\mathbb{K}[\mathrm{X}]$ est un anneau,

$$\mathrm{QP} = \sum_{k=0}^{\infty}\sum_{\ell=0}^{\infty} \mathrm{Q}_k \mathrm{P}_\ell \mathrm{X}^{k+\ell}$$

D'après la linéarité de Φ,

$$(\mathrm{QP})(f) = \sum_{k=0}^{\infty}\sum_{\ell=0}^{\infty} \underbrace{(\mathrm{Q}_k\mathrm{P}_\ell\mathrm{X}^{k+\ell})(f)}_{=\mathrm{Q}_k\mathrm{P}_\ell f^{k+\ell}} = (\mathrm{Q}(f))\circ(\mathrm{P}(f)) \qquad \square$$

On remarquera en particulier que f et $\mathrm{P}(f)$ commutent, pour tout $\mathrm{P}\in\mathbb{K}[\mathrm{X}]$.

Définition 4.2.6 (Polynôme annulateur)
Soit $\mathrm{P}\in\mathbb{K}[\mathrm{X}]$. On dit que P *annule* f, ou encore que P *est un polynôme annulateur de* f si, et seulement si, $\mathrm{P}(f)=0$.

Le reste du chapitre est consacré à tout ce qu'on peut tirer des polynômes annulateurs de f. Mais évidemment, on se demande d'abord si f a des polynômes annulateurs autres que 0. La réponse est positive en dimension finie.

Proposition 4.2.7
On suppose E *de dimension finie non nulle. Alors f a des polynômes annulateurs non nuls.*

Preuve : Notons n la dimension de E. Comme $\mathscr{L}(\mathrm{E})$ est de dimension n^2, la famille $(1,\dots,f^{n^2})$ est liée : on peut trouver $a_0,\dots,a_{n^2}$, non tous nuls, dans $\mathbb{K}$ tels que

$$\sum_{k=0}^{n^2} a_k f^k = 0$$

Donc

$$\underbrace{\Big(\sum_{k=0}^{n^2} a_k\mathrm{X}^k\Big)}_{\neq 0}(f) = 0 \qquad \square$$

En fait, il se trouve qu'on peut être beaucoup plus précis : la preuve précédente donne l'existence d'un polynôme annulateur de degré n^2 au plus. On peut faire beaucoup mieux ; et comme on le verra plus tard, il est utile de connaître un polynôme annulateur de degré le plus petit possible. On montrera plus tard que χ_f, le polynôme caractéristique de f, est aussi un polynôme annulateur. Un des miracles de l'algèbre.

Mais cela devra attendre d'avoir un peu plus développé notre théorie ; on en est aux choses triviales pour l'instant.

Proposition 4.2.8
Soit P *un polynôme annulateur de f. Si $\lambda\in\mathbb{K}$ est valeur propre de f, alors il est aussi racine de* P.

Preuve : Soit x un vecteur propre associé à λ ; par définition, $x\neq 0$ et $fx=\lambda x$. On a $\mathrm{P}(f)=0$ donc

$$0 = (\mathrm{P}(f))(x) = \sum_{k=0}^{\infty} \mathrm{P}_k \underbrace{f^k x}_{=\lambda^k x} = \mathrm{P}(\lambda)\,x$$

Il s'ensuit que $\mathrm{P}(\lambda)=0$. $\square$

Ainsi, théoriquement, si on a trouvé (par chance ?) un polynôme annulateur de f, les valeurs propres de f sont parmi les racines de P. Si (par chance ?) P est facile à factoriser, on a sensiblement réduit la taille de l'ensemble où l'on doit chercher les valeurs propres de f.

Lorsque le polynôme caractéristique de f est scindé, il est facile de calculer le polynôme caractéristique de $\mathrm{P}(f)$, en trigonalisant.

Proposition 4.2.9
On suppose χ_f *scindé, de sorte que* $\chi_f = \prod_{\lambda \in Sp f} (\lambda - \mathrm{X})^{m_\lambda}$. *Si* $\mathrm{P} \in \mathbb{K}[\mathrm{X}]$, *alors*

$$\chi_{\mathrm{P}(f)} = \prod_{\lambda \in Sp f} (\mathrm{P}(\lambda) - \mathrm{X})^{m_\lambda}$$

Preuve : Il suffit de trigonaliser f : puisque χ_f est scindé, il existe une base $\mathscr{B}$ de E telle que

$$\mathrm{Mat}_{\mathscr{B}} f = \begin{bmatrix} \lambda_1 & \star & \cdots & \star \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \star \\ 0 & \cdots & 0 & \lambda_n \end{bmatrix}$$

et une récurrence immédiate prouve que

$$\forall k \in \mathbb{N} \qquad \mathrm{Mat}_{\mathscr{B}} f^k = (\mathrm{Mat}_{\mathscr{B}} f)^k = \begin{bmatrix} \lambda_1^k & \star & \cdots & \star \\ 0 & \lambda_2^k & \ddots & \vdots \\ \vdots & \ddots & \ddots & \star \\ 0 & \cdots & 0 & \lambda_n^k \end{bmatrix}$$

d'où par linéarité $\qquad \forall \mathrm{P} \in \mathbb{K}[\mathrm{X}] \qquad \mathrm{Mat}_{\mathscr{B}} \mathrm{P}(f) = \begin{bmatrix} \mathrm{P}(\lambda_1) & \star & \cdots & \star \\ 0 & \mathrm{P}(\lambda_2) & \ddots & \vdots \\ \vdots & \ddots & \ddots & \star \\ 0 & \cdots & 0 & \mathrm{P}(\lambda_n) \end{bmatrix}$

et $\qquad \forall \mathrm{P} \in \mathbb{K}[\mathrm{X}] \qquad \chi_{\mathrm{P}(f)} = \prod_{k=1}^{n} (\mathrm{P}(\lambda_k) - \mathrm{X}) = \prod_{\lambda \in \mathrm{Sp} f} \big(\mathrm{P}(\lambda) - \mathrm{X}\big)^{m_\lambda}$ $\square$

On fera bien attention à interpéter correctement ce théorème : il **ne dit pas** que si λ est valeur propre de f, alors $\mathrm{P}(\lambda)$ est valeur propre de $\mathrm{P}(f)$ avec la même multiplicité algébrique. En effet, il est parfaitement possible que pour deux valeurs propres distinctes λ et μ, on ait $\mathrm{P}(\lambda) = \mathrm{P}(\mu)$. Par exemple, si $\chi_f = (1 - \mathrm{X})(-1 - \mathrm{X})$, alors $\chi_{f^2} = (1 - \mathrm{X})^2$; du coup, f a deux valeurs propres simples, tandis que f^2 a une valeur propre double.

4.2.3 Le lemme des noyaux

Théorème 4.2.10 (Lemme des noyaux)
Soient $m \in \mathbb{N}^\star$ *et* $\mathrm{P}_1, \ldots, \mathrm{P}_m$ *dans* $\mathbb{K}[\mathrm{X}]$, *premiers entre eux deux-à-deux. On note* $\mathrm{P} = \prod_{k=1}^{m} \mathrm{P}_k$. *Alors*

$$Ker \mathrm{P}(f) = \bigoplus_{k=1}^{m} Ker \mathrm{P}_k(f)$$

De plus, si on note $p_1, \ldots, p_m$ *les projections associées à la somme directe, il existe* $\mathrm{Q}_1, \ldots, \mathrm{Q}_m \in \mathbb{K}[\mathrm{X}]$ *tels que*

$$\forall k \in [\![1\,; m]\!] \qquad p_k = \mathrm{Q}_k(f)$$

Preuve : Lorsque $m = 1$, il n'y a rien à prouver. On suppose $m \geqslant 2$. Il sera utile pour la suite de poser

$$\forall k \in [\![1\,;\,m]\!] \qquad \mathrm{V}_k = \prod_{\substack{j=1\\ j\neq k}}^{m} \mathrm{P}_j$$

et on remarquera que $\qquad \forall k \in [\![1\,;\,m]\!] \qquad \mathrm{P}_k \mathrm{V}_k = \mathrm{P}$

Un des lemmes du cours sur les fractions rationnelles assure qu'on peut trouver des polynômes $\mathrm{U}_1, \ldots, \mathrm{U}_m$ tels que

$$\frac{1}{\mathrm{P}} = \sum_{k=1}^{m} \frac{\mathrm{U}_k}{\mathrm{P}_k} \qquad \text{et} \qquad \forall k \in [\![1\,;\,m]\!] \quad \deg \mathrm{U}_k < \deg \mathrm{P}_k$$

En regroupant la fraction de droite, on se retrouve avec

$$\frac{1}{\mathrm{P}} = \frac{\sum\limits_{k=1}^{m} \mathrm{U}_k \mathrm{V}_k}{\mathrm{P}}$$

d'où

$$1 = \sum_{k=1}^{m} \mathrm{U}_k \mathrm{V}_k$$

et

$$\mathrm{id}_{\mathrm{E}} = \sum_{k=1}^{m} \mathrm{U}_k(f)\,\mathrm{V}_k(f)$$

Ceci étant fait, on peut prouver le théorème des noyaux.

- Soit $k \in [\![1\,;\,m]\!]$; on montre d'abord que $\operatorname{Ker} \mathrm{P}_k(f) \subset \operatorname{Ker} \mathrm{P}(f)$. C'est simple : en effet, si x est dans $\operatorname{Ker} \mathrm{P}_k(f)$,

$$\mathrm{P}(f)(x) = (\mathrm{V}_k \mathrm{P}_k)(f)(x) = (\mathrm{V}_k(f))\big(\mathrm{P}_k(f)(x)\big) = 0$$

- On montre ensuite que les $(\operatorname{Ker} \mathrm{P}_k(f))_{k \in [\![1\,;\,m]\!]}$ sont en somme directe. Soient $x_1, \ldots, x_m$ dans $\operatorname{Ker} \mathrm{P}_1(f), \ldots, \operatorname{Ker} \mathrm{P}_m(f)$ respectivement, tels que

$$\sum_{k=1}^{m} x_k = 0$$

On fixe k dans $[\![1\,;\,m]\!]$. On a

$$\forall j \in [\![1\,;\,m]\!] \setminus \{k\} \qquad \mathrm{P}_k | \mathrm{V}_j$$

donc $\qquad \forall j \in [\![1\,;\,m]\!] \setminus \{k\} \qquad \mathrm{V}_j(f)(x_k) = 0$

et $\qquad \forall j \in [\![1\,;\,m]\!] \setminus \{k\} \qquad \mathrm{U}_j(f)\mathrm{V}_j(f)(x_k) = 0$

d'où $\qquad x_k = \mathrm{id}_{\mathrm{E}}(x_k) = \sum\limits_{j=1}^{m} \mathrm{U}_j(f)\mathrm{V}_j(f)(x_k) = \mathrm{U}_k(f)\mathrm{V}_k(f)(x_k)$

Mais on a aussi

$$\forall j \in [\![1\,;\,m]\!] \setminus \{k\} \qquad \mathrm{U}_k(f)\mathrm{V}_k(f)(x_j) = 0$$

parce que $\mathrm{P}_j | \mathrm{V}_k$ lorsque $j \in [\![1\,;\,m]\!] \setminus \{k\}$. Par suite,

$$x_k = \mathrm{U}_k(f)\mathrm{V}_k(f)(x_k) = \mathrm{U}_k(f)\mathrm{V}_k(f)\Big(\sum_{j=1}^{m} x_j\Big) = \mathrm{U}_k(f)\mathrm{V}_k(f)(0) = 0$$

Comme k était quelconque dans $[\![1\,;m]\!]$, on a prouvé que les noyaux de $P_1(f),\dots,P_m(f)$ sont en somme directe.

- On montre que ces noyaux engendrent $\operatorname{Ker} P(f)$. On se donne un $x \in \operatorname{Ker} P$; on sait que

$$x = \mathrm{id}_E(x) = \sum_{k=1}^{m} U_k(f)V_k(f)(x)$$

Posons $\qquad \forall k \in [\![1\,;m]\!] \qquad x_k = U_k(f)V_k(f)(x)$

On a $\qquad \forall k \in [\![1\,;m]\!] \qquad P_k(f)(x_k) = P_k(f)U_k(f)V_k(f)(x) = U_k(f)\underbrace{P_k(f)V_k(f)}_{=P(f)}(x) = 0$

donc $\qquad \forall k \in [\![1\,;m]\!] \qquad x_k \in \operatorname{Ker} P_k(f) \qquad \text{et} \qquad x = \sum_{k=1}^{m} x_k$

Ceci prouve bien le résultat annoncé.

- Enfin, maintenant qu'on sait que $\operatorname{Ker} P(f) = \bigoplus_{k=1}^{m} \operatorname{Ker} P_k(f)$, on peut définir les projections $p_1,\dots,p_m$ associées à cette décomposition. Par définition et d'après ce qu'on a vu au point précédent,

$$\forall x \in \operatorname{Ker} P(f) \quad \forall k \in [\![1\,;m]\!] \qquad p_k(x) = (U_kV_k)(f)(x)$$

Une conséquence immédiate de ce lemme est le très joli théorème suivant, qui caractérise les endomorphismes diagonalisables :

Théorème 4.2.11
f est diagonalisable si, et seulement si, il admet un polynôme annulateur scindé à racines simples.

Preuve : Supposons f diagonalisable : on sait que $E = \bigoplus_{\lambda \in \operatorname{Sp} f} E_\lambda(f)$. Posons

$$P = \prod_{\lambda \in \operatorname{Sp} f} (X - \lambda) \quad \text{de sorte que} \quad P(f) = \prod_{\lambda \in \operatorname{Sp} f} (f - \lambda \mathrm{id}_E)$$

Si μ est une valeur propre de f et x est dans $E_\mu(f)$, alors

$$P(f)(x) = \prod_{\substack{\lambda \in \operatorname{Sp} f \\ \lambda \neq \mu}} (f - \lambda \mathrm{id}_E)\, \underbrace{(f - \mu \mathrm{id}_E)(x)}_{=0}$$

Donc $P(f)$ est nul sur tous les sous-espaces propres de f. Par suite, $P(f)$ est nul sur E.

Réciproquement, si f admet un polynôme annulateur scindé à racines simples. On le note P ; on note $m \in \mathbb{N}^\star$ le nombre de racines de P, et celles-ci sont notées $\lambda_1,\dots,\lambda_m$. Et on suppose que P est unitaire, pour simplifier : son coefficient dominant ne change absolument pas le fait que P annule f. Avec ces notations,

$$P = \prod_{k=1}^{m} (X - \lambda_k)$$

Les polynômes $X - \lambda_1,\dots,X - \lambda_m$ sont premiers entre eux deux-à-deux et d'après le **lemme des noyaux**,

$$E = \operatorname{Ker} P(f) = \bigoplus_{k=1}^{m} \operatorname{Ker}(f - \lambda_k \mathrm{id}_E) = \bigoplus_{k=1}^{m} E_{\lambda_k}(f)$$

Ceci prouve que f est diagonalisable : on rappelle en effet que les valeurs propres de f sont parmi les racines de P (car P annule f) ; et que si $\lambda \in \mathbb{K}$ n'est pas valeur propre de f, alors $E_\lambda(f) = \{0\}$. □

Ce théorème a deux corollaires importants :

Corollaire 4.2.12
On suppose f diagonalisable. Si F *est un sous-espace de* E*, stable par f, alors l'endomorphisme induit par f sur* F *est diagonalisable.*

Preuve : f est diagonalisable donc il existe $P \in \mathbb{K}[X]$, scindé, à racines simples, qui annule f. Notons $\tilde{f}$ l'endomorphisme induit par f sur F. On a

$$\forall x \in F \quad \forall k \in \mathbb{N} \qquad \tilde{f}^k(x) = f^k(x)$$

donc
$$\forall x \in F \qquad P(\tilde{f})(x) = P(f)(x) = 0$$

Donc P annule $\tilde{f}$. D'après le **théorème 2.11**, $\tilde{f}$ est diagonalisable. □

> On observera la difficulté qu'il peut y avoir à prouver directement ce théorème, sans utiliser la théorie des polynômes annulateurs. En effet, on pourrait avoir envie de commencer la preuve ainsi : f étant diagonalisable, on a la décomposition $E = \bigoplus_{\lambda \in \mathrm{Sp} f} E_\lambda(f)$. On aimerait alors dire que $F = \bigoplus_{\lambda \in \mathrm{Sp} f} (E_\lambda(f) \cap F)$. Cependant, cette relation est difficile à prouver directement, bien qu'elle soit vraie.
>
> En effet, en général, si $E = E_1 \oplus E_2$ est une décomposition de E, et si G est un sous-espace de E, le sous-espace $(E_1 \cap F) \oplus (E_2 \cap F)$ n'est pas égal à F. Prendre par exemple dans $\mathbb{R}^2$:
>
> $$E_1 = \mathrm{Vect}\begin{bmatrix}1\\0\end{bmatrix} \qquad E_2 = \mathrm{Vect}\begin{bmatrix}0\\1\end{bmatrix} \qquad G = \mathrm{Vect}\begin{bmatrix}1\\1\end{bmatrix}$$
>
> Dans ce cas, $E_1 \cap G$ et $E_2 \cap G$ sont $\{0\}$.

Corollaire 4.2.13 (Diagonalisation d'endomorphismes diagonalisables qui commutent)
On suppose E *de dimension finie non nulle. Soit $(f_i)_{i \in I}$ une famille d'endomorphismes de* E*, diagonalisables, et qui commutent deux-à-deux. Alors il existe une base de* E *dans laquelle les matrices de toutes les $(f_i)_{i \in I}$ sont diagonales.*

> Avant de prouver ce théorème, on remarquera bien qu'on ne fait pas d'hypothèse sur I : il peut être fini ou infini. En revanche, on suppose qu'on sait déjà que les $(f_i)_{i \in I}$ sont tous diagonalisables.

Preuve : On procède par récurrence sur la dimension de l'espace. On définit, pour tout $n \in \mathbb{N}^\star$, l'assertion $\mathscr{P}(n)$: « Soit E de dimension finie n. Si $(f_i)_{i \in I}$ est une famille d'endomorphismes de E, diagonalisables, commutant deux-à-deux, alors on peut les diagonaliser dans une même base. »

- $\mathscr{P}(1)$ est vraie, puisqu'en dimension 1, tout endomorphisme est diagonalisable : il n'y a rien à prouver.

- Soit n un entier non nul tel que $\mathscr{P}(n)$ est vraie. On se donne un espace vectoriel E de dimension $n+1$ et une famille $(f_i)_{i\in I}$ d'endomorphismes de E, diagonalisables, qui commutent deux-à-deux.
  Si tous les $(f_i)_{i\in I}$ sont des homothéties, il n'y a rien à prouver. Donc on suppose qu'il existe $a \in I$ tel que f_a n'est pas une homothétie. Puisque f_a est diagonalisable, on peut choisir une de ses valeurs propres μ. De sorte que
  $$E = E_\mu(f_a) \oplus G \qquad \text{avec} \qquad G = \bigoplus_{\substack{\lambda\in \mathrm{Sp} f\\ \lambda\neq\mu}} E_\lambda(f)$$
  Le point important, c'est que les $(E_\lambda(f))_{\lambda\in\mathrm{Sp} f}$ sont stables par tous les $(f_i)_{i\in I}$ car ils commutent tous avec f_a et id_E. En particulier, $E_\mu(f)$ et G sont stables par les $(f_i)_{i\in I}$. Donc pour chaque $i \in I$, f_i induit un endomorphisme f_i^μ de $E_\mu(f)$, et un endomorphisme f_i^G de G. Évidemment, les $(f_i^\mu)_{i\in I}$ commutent deux-à-deux, parce que
  $$\forall x \in E_\mu(f) \quad \forall i, j \in I \qquad f_i^\mu f_j^\mu(x) = f_i(f_j(x)) = f_j(f_i(x)) = f_j^\mu f_i^\mu(x)$$
  Pour les $(f_i^G)_{i\in I}$, c'est la même chose. Enfin, les $(f_i^\mu)_{i\in I}$ et les $(f_i^G)_{i\in I}$ sont tous diagonalisables d'après le **corollaire 2.12**.
  Le fait que f_a ne soit pas une homothétie implique que $\dim E_\mu(f_a) < n+1$; et le fait que μ est valeur propre de f_a assure que $\dim E_\mu(f_a) \neq 0$. Autrement dit,
  $$\dim E_\mu(f_a) \in [\![1\,;\,n]\!] \qquad \dim G = n+1-\dim E_\mu(f_a) \in [\![1\,;\,n]\!]$$
  Ceci permet d'utiliser $\mathscr{P}(n)$: on peut trouver une base $\mathscr{B}_\mu$ de $E_\mu(f_a)$ qui diagonalise les $(f_i^\mu)_{i\in I}$; et une base $\mathscr{B}_G$ de G qui diagonalise les $(f_i^G)_{i\in I}$. En les réunissant, on obtient une base $\mathscr{B} = (\mathscr{B}_\mu, \mathscr{B}_G)$ qui diagonalise les $(f_i)_{i\in I}$. On a prouvé $\mathscr{P}(n+1)$.
- Par récurrence, $\mathscr{P}(n)$ est vraie pour tout $n \in \mathbb{N}^\star$, ce qui prouve le théorème. □

4.2.4 Le polynôme minimal

L'utilité d'un polynôme annulateur de f est évidente, à la lumière des résultats précédents. Mais ces théorèmes sont surtout des outils de travail théoriques. En pratique, on peut avoir envie d'un polynôme annulateur, si possible de degré minimal. Car cela permet de calculer plus facilement, par exemple, les puissances d'un endomorphisme.

En effet, supposons que $P \in \mathbb{K}[X]$ est un polynôme annulateur de f ; on note p son degré. Pour chaque $n \in \mathbb{N}$, on note Q_n et R_n le quotient et le reste de la division euclidienne de X^n par P :
$$X^n = PQ_n + R_n \qquad \text{avec} \qquad \deg R_n \leqslant p-1$$
Alors
$$f^n = P(f)Q_n(f) + R_n(f) = R_n(f)$$
Donc si on est capable de calculer R_n, on aura une expression de f^n en fonction des $(f^k)_{0\leqslant k\leqslant p-1}$. Et plus p est petit, plus ce calcul sera simple. C'est une des motivations pour introduire le polynôme minimal de f.

Mais avant cela, illustrons ce qui précède sur un exemple.

Exemple 4.2.14
Prenons la matrice A de l'**exemple 1.21**. On peut vérifier que $P = (X-1)^2(X+1)$ annule A. Si n est un entier non nul, notons R_n le reste de la division euclidienne de X^n par P. On a $\deg R_n \leqslant 2$ donc il existe des scalaires a, b et c tels que

$$\mathrm{R}_n = a\mathrm{X}^2 + b\mathrm{X} + c$$

Comme $$\mathrm{X}^n = (\mathrm{X}-1)^2(\mathrm{X}+1)\mathrm{Q}_n + \mathrm{R}_n$$

on a tout de suite $$\begin{cases} 1 = \mathrm{R}_n(1) = a+b+c \\ (-1)^n = \mathrm{R}_n(-1) = a-b+c \end{cases}$$

Pour obtenir une troisième relation, on utilise le fait que 1 est racine double de $(\mathrm{X}-1)^2(\mathrm{X}+1)\mathrm{Q}_n$, donc racine de la dérivée. Alors

$$n\,1^{n-1} = \mathrm{R}'_n(1) = 2a+b$$

Ainsi, $$\begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & 1 \\ 2 & 1 & 0 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 1 \\ (-1)^n \\ n \end{bmatrix}$$

et $$a = \frac{n}{2} + \frac{(-1)^n - 1}{4} \qquad b = \frac{1-(-1)^n}{2} \qquad c = \frac{(-1)^n+3}{4} - \frac{n}{2}$$

Finalement, $$\forall n \in \mathbb{N}^\star \qquad \mathrm{A}^n = \frac{2n+(-1)^n-1}{4}\mathrm{A}^2 + \frac{1-(-1)^n}{2}\mathrm{A} + \frac{(-1)^n+3-2n}{4}\mathrm{I}_n$$

Formule qu'il est difficile de deviner si on veut la prouver par récurrence.

La théorie des idéaux de $\mathbb{K}[\mathrm{X}]$ va maintenant revenir en force. On rappelle que, pendant le cours de révisions sur les polynômes, on a introduit la notion d'idéal d'un anneau :

Définition 4.2.15 (Idéaux d'un anneau)
Soient A un anneau commutatif et $\mathscr{I} \subset \mathrm{A}$. On dit que $\mathscr{I}$ *est un idéal de* A si, et seulement si, c'est un sous-groupe de $(\mathrm{A},+)$ et

$$\forall a \in \mathrm{A} \quad \forall x \in \mathscr{I} \qquad ax \in \mathscr{I}$$

et on a identifié les idéaux de $\mathbb{K}[\mathrm{X}]$:

Théorème 4.2.16
Soit $\mathscr{I}$ un idéal non nul de $\mathbb{K}[\mathrm{X}]$. Il existe un unique $\mathrm{P} \in \mathscr{I}$, unitaire, tel que

$$\mathscr{I} = \mathrm{P}\mathbb{K}[\mathrm{X}] = \{\mathrm{PQ} \mid \mathrm{Q} \in \mathbb{K}[\mathrm{X}]\}$$

Autrement dit, P *est caractérisé par le fait que*

$$\begin{cases} \mathrm{P} \textit{ est unitaire} \\ \forall \mathrm{Q} \in \mathrm{A} \qquad \mathrm{Q} \in \mathscr{I} \iff \mathrm{P}|\mathrm{Q} \end{cases}$$

Il se trouve que les polynômes annulateurs d'un endomorphisme forment un idéal de $\mathbb{K}[\mathrm{X}]$:

Proposition 4.2.17
On note $\mathrm{Ann}\, f$ l'ensemble des polynômes annulateurs de f. C'est un idéal de $\mathbb{K}[\mathrm{X}]$; si E *est de dimension finie, $\mathrm{Ann}\, f$ n'est pas réduit à* $\{0\}$.

Preuve : C'est juste une question de vérifier les axiomes qui définissent un idéal. Si E est de dimension finie, la **proposition 2.7** dit exactement que $\mathrm{Ann}\, f$ n'est pas $\{0\}$. □

Définition 4.2.18 (Polynôme minimal)
On suppose E de dimension finie, non nulle. Le générateur principal unitaire de $\operatorname{Ann} f$ est appelé *le polynôme minimal de* f et on le note π_f. Il est caractérisé par les propriétés :

$$\begin{cases} \pi_f \text{ est unitaire} \\ \forall \mathrm{P} \in \mathbb{K}[\mathrm{X}] \qquad \mathrm{P}(f) = 0 \iff \pi_f | \mathrm{P} \end{cases}$$

En dimension finie, puisque π_f divise tous les polynômes annulateurs de f, c'est le polynôme annulateur de f *unitaire* de degré minimal, d'où son nom.

Trouver π_f n'est généralement pas facile. Mais si on a par chance un polynôme P, annulateur de f, il « suffit » d'essayer tous les diviseurs de P. Nécessairement, l'un d'eux doit être π_f.

On sait aussi que les valeurs propres de f sont racines de π_f (**proposition 2.8**) ; donc si on a réussi à calculer χ_f et à le factoriser, on n'a qu'un nombre fini de possibilités pour π_f. En fait, on a même mieux : on peut montrer que χ_f est un polynôme annulateur de f ; donc π_f divise χ_f.

Théorème 4.2.19 (Théorème de Cayley-Hamilton)
On suppose E *de dimension finie non nulle. Si* $f \in \mathscr{L}(\mathrm{E})$, *alors* $\pi_f | \chi_f$ *et* $\chi_f(f) = 0$.

Preuve : Notons $n \in \mathbb{N}^\star$ la dimension de E. Soit $f \in \mathscr{L}(\mathrm{E})$. On fixe $x \in \mathrm{E}$, non nul, et on va prouver que $\chi_f(f)(x) = 0$. Cela se fait en plusieurs étapes. On pose

$$\mathrm{U} = \operatorname{Vect}(f^k(x))_{k\in\mathbb{N}} = \{\mathrm{P}(f)(x) \mid \mathrm{P} \in \mathbb{K}[\mathrm{X}]\} \qquad \mathscr{I} = \{\mathrm{P} \in \mathbb{K}[\mathrm{X}] \mid \mathrm{P}(f)(x) = 0\}$$

- U est un sous-espace de E, par définition. De plus,

$$\forall \mathrm{P} \in \mathbb{K}[\mathrm{X}] \qquad f(\mathrm{P}(f)(x)) = (f\mathrm{P}(f))(x) = (\mathrm{XP})(f)(x) \in \mathrm{U}$$

  donc U est stable par f. Enfin, il contient x donc sa dimension, notée m, n'est pas nulle.
- $\mathscr{I}$ est un idéal : en effet,

$$\forall \mathrm{P}, \mathrm{Q} \in \mathscr{I} \qquad (\mathrm{P}+\mathrm{Q})(f)(x) = \underbrace{\mathrm{P}(f)(x)}_{=0} + \underbrace{\mathrm{Q}(f)(x)}_{=0} = 0$$

  et $$\forall \mathrm{P} \in \mathscr{I} \quad \forall \mathrm{Q} \in \mathbb{K}[\mathrm{X}] \qquad (\mathrm{QP})(f)(x) = \mathrm{Q}(f)\big(\mathrm{P}(f)(x)\big) = \mathrm{Q}(f)(0) = 0$$

  De plus, il contient π_f parce que $\pi_f(f) = 0$ et $\pi_f(f)(x) = 0$. Par conséquent, il existe un (unique) polynôme unitaire $\mathrm{P}_{f,x}$ dans $\mathscr{I}$ tel que $\mathscr{I} = \mathrm{P}_{f,x}\mathbb{K}[\mathrm{X}]$. Son degré est noté p. Et p n'est pas nul, car $f^0(x) \neq 0$.
- On montre ensuite que $m = p$. Soit $y \in \mathrm{U}$: par définition, il existe $\mathrm{P} \in \mathbb{K}[\mathrm{X}]$ tel que $y = \mathrm{P}(f)(x)$. On effectue la division euclidienne de P par $\mathrm{P}_{f,x}$ pour obtenir Q et R dans $\mathbb{K}[\mathrm{X}]$, tels que

$$\mathrm{P} = \mathrm{P}_{f,x}\mathrm{Q} + \mathrm{R} \qquad \text{et} \qquad \deg \mathrm{R} \leqslant p-1$$

  Alors $$\mathrm{P}(f) = \mathrm{Q}(f)\,\mathrm{P}_{f,x}(f) + \mathrm{R}(f)$$

  et $$y = \mathrm{P}(f)(x) = \mathrm{Q}(f)\big(\underbrace{\mathrm{P}_{f,x}(f)(x)}_{=0}\big) + \mathrm{R}(f)(x) = \mathrm{R}(f)(x)$$

  Comme R est de degré plus petit que $p-1$, y est combinaison linéaire des $(f^k(x))_{0\leqslant k\leqslant p-1}$. Cette famille engendre donc U et il s'ensuit que $m \leqslant p$.

On montre en plus que cette famille est libre : soient $\lambda_0,\dots,\lambda_{p-1}$ sont dans $\mathbb{K}$ tels que

$$\sum_{k=0}^{p-1}\lambda_k f^k(x)=0$$

alors
$$\Big(\sum_{k=0}^{p-1}\lambda_k X^k\Big)(f)(x)=0$$

et il vient
$$\sum_{k=0}^{p-1}\lambda_k X^k\in\mathscr{I}\qquad\text{ou encore}\qquad P_{f,x}\,\Big|\sum_{k=0}^{p-1}\lambda_k X^k$$

Mais $P_{f,x}$ est de degré p, donc $\sum_{k=0}^{p-1}\lambda_k X^k=0$. Les scalaires $\lambda_0,\dots,\lambda_{p-1}$ sont nuls.

Donc $\mathscr{B}_x=(f^k(x))_{0\leqslant k\leqslant p-1}$ est une base de U, ce qui prouve que U est de dimension p.

- On note $\tilde f$ l'endomorphisme de U induit par f et on calcule sa matrice, notée C_f, dans $\mathscr{B}_x$. Pour cela, il sera utile d'introduire les coefficients de $P_{f,x}$:

$$P_{f,x}=X^p+\sum_{k=0}^{p-1}P_k X^k$$

Par définition,
$$P_{f,x}(f)(x)=0=f^p(x)+\sum_{k=0}^{p-1}P_k f^k(x)$$

Il vient
$$C_f=\mathrm{Mat}_{\mathscr{B}_x}\tilde f=\begin{bmatrix}0&\cdots&0&-P_0\\ 1&\ddots&\vdots&\vdots\\ &\ddots&0&-P_{p-2}\\ &&1&-P_{p-1}\end{bmatrix}$$

Son polynôme caractéristique se calcule facilement par récurrence et on trouve que c'est tout simplement $(-1)^pP_{f,x}$:

$$\chi_{C_f}=\chi_{\tilde f}=(-1)^pP_{f,x}$$

- On conclut : On a déjà vu (**proposition 2.2**) que $\chi_{\tilde f}$ divise χ_f. Mais $\chi_{\tilde f}=(-1)^pP_{f,x}$ donc χ_f est dans l'idéal $\mathscr{I}$. Et il s'ensuit que $\chi_f(f)(x)=0$.

On a prouvé
$$\forall x\in E\qquad \chi_f(f)(x)=0$$

donc χ_f annule f. Et par suite, π_f divise χ_f. □

Corollaire 4.2.20
On suppose E *de dimension finie, non nulle. Alors* $\deg\pi_f\leqslant\dim E$. *De plus, si* χ_f *est scindé, alors* π_f *est scindé et il existe des entiers* $(q_\lambda)_{\lambda\in Sp f}$ *tels que*

$$\pi_f=\prod_{\lambda\in Sp f}(X-\lambda)^{q_\lambda}\qquad\textit{et}\qquad\forall\lambda\in Sp f\quad 1\leqslant q_\lambda\leqslant m_\lambda$$

Preuve : On sait que π_f divise χ_f donc $\deg \pi_f \leqslant \deg \chi_f = \dim \mathrm{E}$.

Supposons χ_f scindé. Si P est un facteur irréductible de π_f, alors il divise χ_f. Mais les facteurs irréductibles de χ_f sont tous degré 1 donc P est de degré 1. Par suite, π_f est scindé. Comme il divise χ_f, toute racine de π_f est racine de χ_f. Donc il existe $(q_\lambda)_{\lambda \in \mathrm{Sp} f}$ entiers tels que

$$\pi_f = \prod_{\lambda \in \mathrm{Sp} f} (\mathrm{X} - \lambda)^{q_\lambda} \qquad \text{et} \qquad \forall \lambda \in \mathrm{Sp} f \quad 0 \leqslant q_\lambda \leqslant m_\lambda$$

Enfin, si $\lambda \in \mathrm{Sp} f$, la proposition **2.8** assure que λ est aussi racine de π_f, puisque celui-ci annule f. Donc $q_\lambda \geqslant 1$. □

Autrement dit, si χ_f est scindé, π_f aussi ; toutes les valeurs propres de f apparaissent au moins une fois dans le polynôme minimal ; et les seules racines de celui-ci sont les valeurs propres de f.

On termine le paragraphe avec une dernière technicalité : relier les polynômes minimaux de f et $f - \lambda \, \mathrm{id}_\mathrm{E}$.

Proposition 4.2.21
On suppose E *de dimension finie, non nulle. Soit* $\lambda \in \mathbb{K}$. *Alors*

$$\pi_{f - \lambda id_\mathrm{E}} = \pi_f \circ (\mathrm{X} + \lambda)$$

Preuve : C'est un jeu très amusant avec la définition des polynômes minimaux. Par définition, $\pi_{f-\lambda \mathrm{id}_\mathrm{E}}$ annule $f - \lambda \, \mathrm{id}_\mathrm{E}$. Mais

$$\pi_{f-\lambda \mathrm{id}_\mathrm{E}}(f - \lambda \, \mathrm{id}_\mathrm{E}) = \left[\pi_{f-\lambda \mathrm{id}_\mathrm{E}} \circ (\mathrm{X} - \lambda)\right](f)$$

d'où $\quad \pi_f | \pi_{f-\lambda \mathrm{id}_\mathrm{E}} \circ (\mathrm{X} - \lambda) \quad$ ce qui équivaut à dire $\quad \pi_f \circ (\mathrm{X} + \lambda) | \pi_{f-\lambda \mathrm{id}_\mathrm{E}}$

L'autre direction se prouve de la même manière : π_f annule f. Mais on peut écrire

$$\pi_f(f) = \pi_f(f - \lambda \, \mathrm{id}_\mathrm{E} + \lambda \, \mathrm{id}_\mathrm{E}) = \left[\pi_f \circ (\mathrm{X} + \lambda)\right](f - \lambda \, \mathrm{id}_\mathrm{E})$$

Par conséquent, $\pi_{f-\lambda \mathrm{id}_\mathrm{E}}$ divise $\pi_f(\mathrm{X} + \lambda)$.

Comme ces deux polynômes sont unitaires et se divisent mutuellement, ils sont égaux. □

4.3 Approfondissements

Notre programme d'études s'arrête là concernant la réduction. Néanmoins, la théorie semble bien incomplète : on n'a donné que la définition du polynôme minimal, sans vraiment en donner d'application, ou de méthode déterministe pour le trouver.

Cette section contient des approfondissements de la théorie, pour qu'elle soit satisfaisante. On va pousser la réduction des endomorphismes le plus loin possible pour arriver à la fameuse réduction de Jordan.

4.3.1 Endomorphismes nilpotents

Le but de ce paragraphe est l'étude poussée des endomorphismes nilpotents, fondamentale pour comprendre la réduction de Jordan. On souhaite d'abord prouver une réciproque au **corollaire 2.20**. Dans toute cette partie, E est de dimension finie non nulle, notée n, et f est un endomorphisme de E. On commence par

Lemme 4.3.1
Si f est nilpotent, il existe une base de E *dans laquelle la matrice de f est triangulaire supérieure stricte.*

Preuve : Pour tout entier $n \in \mathbb{N}^\star$, on note $\mathscr{P}(n)$ l'assertion : « Soit F un $\mathbb{K}$-espace vectoriel de dimension n non nulle. Si $u \in \mathscr{L}(\mathrm{F})$ est nilpotent, il existe une base de F dans laquelle la matrice de f est triangulaire supérieure stricte. »

- $\mathscr{P}(1)$ est vraie : en effet, si F est de dimension 1 et $u \in \mathscr{L}(\mathrm{F})$ est nilpotent, il n'est pas inversible. Dans n'importe quelle base, sa matrice est 1×1 non inversible : elle est nulle (et en particulier triangulaire supérieure stricte).
- Soit $n \in \mathbb{N}^\star$ tel que $\mathscr{P}(n)$ est vraie. On se donne F de dimension $n+1$ et $u \in \mathscr{L}(\mathrm{F})$ nilpotent. Comme $u^{n+1} = 0$, u n'est pas bijectif, donc non injectif. Il existe $e_1 \in \mathrm{F}$, non nul, tel que $u(e_1) = 0$. On forme alors une base $\mathscr{B} = (e_1, \dots, e_n)$ de F à l'aide du **théorème de la base incomplète** et on a

$$\mathrm{Mat}_{\mathscr{B}} f = \begin{bmatrix} 0 & \mathrm{L} \\ 0_{n,1} & \mathrm{M} \end{bmatrix} \qquad \text{avec} \qquad \mathrm{L} \in \mathrm{M}_{1,n}(\mathbb{K}) \quad \mathrm{M} \in \mathrm{M}_n(\mathbb{K})$$

  On vérifie par récurrence que

$$\mathrm{Mat}_{\mathscr{B}} f^{n+1} = (\mathrm{Mat}_{\mathscr{B}} f)^{n+1} = \begin{bmatrix} 0 & \mathrm{LM}^n \\ 0 & \mathrm{M}^{n+1} \end{bmatrix}$$

  Mais cette matrice est nulle parce que $f^{n+1} = 0$. Donc $\mathrm{M}^{n+1} = 0$. Mais M est un endomorphisme de $\mathbb{K}^n$, qui est de dimension n ; d'après $\mathscr{P}(n)$, il existe une matrice $\mathrm{P} \in \mathrm{GL}_n(\mathbb{K})$ telle que $\mathrm{P}^{-1}\mathrm{MP}$ est triangulaire supérieure stricte. Posons

$$\mathrm{Q} = \begin{bmatrix} 1 & 0_{1,n} \\ 0_{n,1} & \mathrm{P} \end{bmatrix}$$

  Il est facile de vérifier que Q est inversible et que Q^{-1} a la même forme, avec P^{-1} dans le bloc en bas à droite. D'où

$$\mathrm{Q}^{-1}\,\mathrm{Mat}_{\mathscr{B}} f\,\mathrm{Q} = \begin{bmatrix} 1 & 0_{1,n} \\ 0_{n,1} & \mathrm{P}^{-1} \end{bmatrix} \begin{bmatrix} 0 & \mathrm{L} \\ 0_{n,1} & \mathrm{M} \end{bmatrix} \begin{bmatrix} 1 & 0_{1,n} \\ 0_{n,1} & \mathrm{P} \end{bmatrix} = \begin{bmatrix} 0 & \mathrm{LP} \\ 0_{n,1} & \mathrm{P}^{-1}\mathrm{MP} \end{bmatrix}$$

  Cette matrice est triangulaire supérieure stricte. D'après les relations de passage et de changement de base, dans la base dont les vecteurs ont pour coordonnées les colonnes de Q dans $\mathscr{B}$, cette matrice est la matrice de f. Donc $\mathscr{P}(n+1)$ est vraie.
- Par récurrence, $\mathscr{P}(n)$ est vraie pour tout $n \in \mathbb{N}^\star$, ce qui prouve le théorème. □

Corollaire 4.3.2
Les assertions suivantes sont équivalente :

1. *f est nilpotent ;*
2. *Il existe une base de* E *dans laquelle f a une matrice triangulaire supérieure stricte ;*
3. *$\chi_f = (-\mathrm{X})^n$;*

Preuve : Le **lemme 3.1** montre que (1) $\Longrightarrow$ (2). Et (3) $\Longrightarrow$ (1) découle du théorème de Cayley-Hamilton.

Enfin, si on peut trouver une base dans laquelle f a une matrice triangulaire supérieure stricte, on se place dans cette base pour calculer le polynôme caractéristique et on trouve bien $(-X)^n$. □

Prouvons maintenant notre réciproque au **corollaire 2.20** :

Théorème 4.3.3
Les assertions suivantes sont équivalentes :

1. χ_f *est scindé ;*
2. f *a un polynôme annulateur scindé ;*
3. π_f *est scindé.*

Dans ce cas, si $p \in \mathbb{N}^\star$ *est le nombre de racines de* π_f *et si ces racines sont notées* $\lambda_1, \dots, \lambda_p$ *avec multiplicités* $q_1, \dots, q_p$*, alors*

$$\pi_f = \prod_{k=1}^{p} (X - \lambda_k)^{q_k} \qquad \chi_f = \prod_{k=1}^{p} (\lambda_k - X)^{n_k} \qquad \textit{avec} \qquad \forall k \in [\![1\,;\,p]\!] \quad n_k = \dim \mathit{Ker}(f - \lambda_k\, \mathit{id}_E)^{q_k}$$

Preuve : Le théorème de Cayley-Hamilton prouve que (1) $\Longrightarrow$ (2). La preuve de (2) $\Longrightarrow$ (3) se passe exactement de la même manière que dans le **corollaire 2.20** : si P est un polynôme scindé qui annule f, alors π_f le divise. Et les facteurs irréductibles de π_f sont tous de degré 1.

Enfin, supposons que (3) est vraie. Notons p le nombre de racines de π_f ; on les note $\lambda_1, \dots, \lambda_p$, et leurs multiplicités $q_1, \dots, q_p$. Alors

$$\pi_f = \prod_{k=1}^{p} (X - \lambda_k)^{q_k}$$

Comme $\pi_f(f) = 0$, le **lemme des noyaux** donne

$$E = \bigoplus_{k=1}^{p} \mathrm{Ker}\,(f - \lambda_k)^{q_k}$$

Posons $\qquad \forall k \in [\![1\,;\,p]\!] \qquad E_k = \mathrm{Ker}\,(f - \lambda_k\, \mathrm{id}_E)^{q_k} \qquad n_k = \dim E_k$

Les sous-espaces $(E_k)_{1 \leqslant k \leqslant p}$ sont tous stables par f, qui induit des endomorphismes $(f_k)_{1 \leqslant k \leqslant p}$ sur chacun d'eux.

Si $k \in [\![1\,;\,p]\!]$, $f_k - \lambda_k\, \mathrm{id}_{E_k}$ est nilpotent par définition de E_k. D'après le **corollaire 3.2**, son polynôme caractéristique est $(-X)^{n_k}$, d'où $\chi_{f_k} = (\lambda - X)^{n_k}$. Enfin, la **proposition 2.3** assure que

$$\chi_f = \prod_{k=1}^{m} (\lambda_k - X)^{n_k}$$ □

4.3.2 Les noyaux emboîtés

Dans la preuve du **théorème 3.3**, on a pu voir apparaître (en gardant les mêmes notations) les sous-espaces $\mathrm{Ker}\,(f - \lambda_k)^{q_k}$. Ils semblent particulièrement intéressants puisqu'ils sont reliés en même temps

- au polynôme minimal : l'exposant q_k est l'ordre de multiplicité de λ_k dans π_f ;
- au polynôme caractéristique : la dimension de $\mathrm{Ker}\,(f - \lambda_k)^{q_k}$ est précisément l'ordre de multiplicité de λ_k dans χ_f.

Et effectivement, ils sont au centre de la théorie de la réduction. Cependant, on remarquera bien que les résultats obtenus dans ce théorème supposaient que π_f (ou χ_f) est scindé dans $\mathbb{K}[X]$. Nous allons voir si l'on peut se passer de cette hypothèse, et prouver aussi que q_k a une interprétation géométrique.

Dans cette partie, E est toujours de dimension finie n non nulle, f est un endomorphisme de E.

Définition 4.3.4 (Sous-espaces caractéristiques)
Soit λ une valeur propre de f. On note $q_\lambda \in [\![1;n]\!]$ sa multiplicité dans π_f. Alors $\mathrm{Ker}(f-\lambda)^{q_\lambda}$ est appelé *sous-espace caractéristique de f relatif à* λ. Il sera noté $K_\lambda(f)$ dans tout le cours.

La démonstration du **théorème 3.3** peut alors être généralisée sans trop de difficultés :

Proposition 4.3.5
Soit λ une valeur propre de f. La dimension de $K_\lambda(f)$ est la multiplicité algébrique de λ.

Preuve : Puisque λ a pour multiplicité q_λ dans π_f, on peut trouver $Q \in \mathbb{K}[X]$ tel que

$$\pi_f = (X-\lambda)^{q_\lambda} Q \qquad \text{et} \qquad Q(\lambda) \neq 0$$

De ce fait, Q et $(X-\lambda)^q$ sont premiers entre eux. D'après le **lemme des noyaux**,

$$E = \mathrm{Ker}\,\pi_f(f) = K_\lambda(f) \oplus \mathrm{Ker}\,Q(f)$$

Posons $$p = \dim K_\lambda(f)$$

$K_\lambda(f)$ et $\mathrm{Ker}\,Q(f)$ sont stables par f. On note f_1 et f_2 les endomorphismes induits sur ces deux espaces respectivement. On a alors

$$\chi_f = \chi_{f_1}\chi_{f_2}$$

Parce que $f_1 - \lambda\mathrm{id}_{K_\lambda(f)}$ est nilpotent sur $K_\lambda(f)$, le **corollaire 3.2** donne son polynôme caractéristique qui est $(-X)^p$. Donc

$$\chi_{f_1} = \chi_{f_1-\lambda\mathrm{id}_{K_\lambda(f)}}(X-\lambda) = (\lambda - X)^p$$

D'autre part, comme Q annule f_2 sur $\mathrm{Ker}\,Q(f)$, le polynôme minimal π_{f_2} divise Q. Mais $Q(\lambda)\neq 0$ donc $\pi_{f_2}(\lambda)\neq 0$. Par suite, λ ne peut pas être racine de χ_{f_2}. Autrement dit,

$$\chi_f = (\lambda - X)^p\chi_{f_2} \qquad \text{avec} \qquad \chi_{f_2}(\lambda)\neq 0$$

ce qui signifie exactement que p est la multiplicité de λ dans χ_f. □

On poursuit l'étude des sous-espaces caractéristiques, avec le **théorème des noyaux emboîtés**, déjà étudié en exercices plusieurs fois.

Lemme 4.3.6 (Théorème des noyaux emboîtés)
On définit

$$\forall p \in \mathbb{N} \qquad N_p = Ker f^p \qquad I_p = Im f^p \qquad d_p = dim N_{p+1} - dim N_p$$

- *Les suites $(N_p)_{p\in\mathbb{N}}$ et $(I_p)_{p\in\mathbb{N}}$ sont respectivement croissante et décroissante.*
- *Les ensembles* $$\{p\in\mathbb{N} \mid N_p = N_{p+1}\} \qquad \text{et} \qquad \{p\in\mathbb{N}\mid I_p = I_{p+1}\}$$

  ne sont pas vides. Leurs plus petits éléments sont égaux et on note leur valeur commune p_0. Alors $p_0 \leqslant n$,

$$\forall p \geqslant p_0 \qquad \mathrm{N}_p = \mathrm{N}_{p_0} \quad \mathrm{I}_p = \mathrm{I}_{p_0}$$

et les espaces N_{p_0} *et* I_{p_0} *sont supplémentaires dans* E.

- *La suite* $(d_p)_{p\in\mathbb{N}}$ *est décroissante jusqu'au rang* p_0*, après lequel elle est nulle.*
- *Si* $p \in \mathbb{N}$ *et* H *est un sous-espace de* N_{p+2}*, en somme directe avec* N_{p+1}*, alors* $f(\mathrm{H})$ *est un sous-espace de* N_{p+1}*, en somme directe avec* N_p*. De plus,* f *est injective sur* H.
- *Si* $p \in \mathbb{N}$ *et* $x \in \mathrm{N}_{p+1} \setminus \mathrm{N}_p$*, la famille* $(x, \dots, f^p(x))$ *est libre.*
- p_0 *est l'ordre de multiplicité de* 0 *dans* π_f.

Preuve : On procède à peu près dans l'ordre.

- Soit $p \in \mathbb{N}$. Si $x \in \mathrm{N}_p$, alors

  $$f^{p+1}(x) = f(f^p(x)) = f(0) = 0$$

  ce qui prouve $\mathrm{N}_p \subset \mathrm{N}_{p+1}$.

  Si $y \in \mathrm{I}_{p+1}$, il existe $x \in \mathrm{I}_p$ tel que $y = f^{p+1}(x) = f^p(f(x))$. Donc $\mathrm{I}_{p+1} \subset \mathrm{I}_p$.

- Si l'ensemble $\{p \in \mathbb{N} \mid \mathrm{N}_p = \mathrm{N}_{p+1}\}$ est inclus dans $[\![n+1\,;+\infty]\!]$, la suite $(\mathrm{N}_p)_{0\leqslant p\leqslant n}$ est strictement croissante ; la suite des dimensions aussi et l'on a

  $$\forall p \in [\![0\,; n]\!] \qquad 1 + \dim \mathrm{N}_p \leqslant \dim \mathrm{N}_{p+1}$$

  Par récurrence, $$\dim \mathrm{N}_n \geqslant n+1$$

  ce qui est absurde. Il existe donc $p_0 = \mathrm{Min}\{p \in \mathbb{N} \mid \mathrm{N}_p = \mathrm{N}_{p+1}\}$ et $p_0 \leqslant n$. On observe que d'après le **théorème du rang**,

  $$\forall p \in \mathbb{N} \qquad \mathrm{N}_p = \mathrm{N}_{p+1} \iff \mathrm{I}_p = \mathrm{I}_{p+1}$$

  ce qui prouve que $$p_0 = \mathrm{Min}\{p \in \mathbb{N} \mid \mathrm{I}_p = \mathrm{I}_{p+1}\}$$

  Si on définit, pour $p \in \mathbb{N}$, l'assertion $\mathscr{P}(p)$: « $\mathrm{K}_p = \mathrm{K}_{p_0}$ », alors $\mathscr{P}(0)$ est vraie par définition de p_0. Et si $p \in \mathbb{N}$ est tel que $\mathscr{P}(p)$ est vraie, on se donne $x \in \mathrm{N}_{p+1}$. Alors

  $$0 = f^{p+1}(x) = f^p(f(x) \qquad \text{d'où} \qquad f(x) \in \mathrm{N}_p = \mathrm{N}_{p_0}$$

  Il s'ensuit $$0 = f^{p_0}(f(x)) = f^{p_0+1}(x) \qquad \text{donc} \qquad x \in \mathrm{N}_{p_0+1} = \mathrm{N}_{p_0}$$

  Ceci prouve $$\mathrm{N}_{p+1} \subset \mathrm{N}_{p_0}$$

  L'autre inclusion a déjà été prouvée. Donc $\mathscr{P}(p+1)$ est vraie. Ce qui prouve par récurrence

  $$\forall p \in \mathbb{N} \qquad p \geqslant p_0 \implies \mathrm{N}_p = \mathrm{N}_{p_0}$$

D'après le **théorème du rang,** on sait déjà que

$$\dim \mathrm{N}_{p_0} + \dim \mathrm{I}_{p_0} = \dim \mathrm{E}$$

De plus, si $x \in \mathrm{N}_{p_0} \cap \mathrm{I}_{p_0}$, alors il existe $y \in \mathrm{E}$ tel que $x = f^{p_0}(y)$. En outre,

$$0 = f^{p_0}(x) = f^{2p_0}(y) \qquad \text{donc} \qquad y \in \mathrm{N}_{2p_0} = \mathrm{N}_{p_0}$$

ce qui assure que $x = f^{p_0}(y) = 0$. Donc N_{p_0} et I_{p_0} sont en somme directe. Avec l'aide de la relation entre les dimensions, on sait qu'ils sont supplémentaires dans E.

- On montre en même temps les deux points suivants. Soient $p \in \mathbb{N}$ et H un sous-espace de N_{p+2}, en somme directe avec N_{p+1}. On a

$$\forall x \in \mathrm{H} \qquad f^{p+1}(f(x)) = f^{p+2}(x) = 0$$

ce qui prouve que $f(\mathrm{H}) \subset \mathrm{N}_{p+1}$.

Montrons l'injectivité de f sur H : soit $x \in \mathrm{H}$ tel que $f(x) = 0$. Alors $f^{p+1}(x) = 0$ et il vient que $x \in \mathrm{H} \cap \mathrm{N}_{p+1} = \{0\}$.

Ensuite, prouvons que $f(\mathrm{H})$ est en somme directe avec N_p : si $y \in f(\mathrm{H}) \cap \mathrm{N}_p$, il existe $x \in \mathrm{H}$ tel que $y = f(x)$. En outre, $f^{p+1}(x) = f^{p+2}(y) = 0$ donc x est aussi dans N_{p+1}. Or, $\mathrm{H} \cap \mathrm{N}_{p+1} = \{0\}$ donc $x = 0$.

En particulier, si on prend pour H un supplémentaire de N_{p+1} dans N_{p+2}, alors

$$\dim \mathrm{H} = \dim \mathrm{N}_{p+2} - \dim \mathrm{N}_{p+1} = d_{p+1}$$

D'autre part, $f(\mathrm{H})$ est en somme directe avec N_p et inclus dans N_{p+1} donc

$$\dim f(\mathrm{H}) + \dim \mathrm{N}_p = \dim\big(f(\mathrm{H}) \oplus \mathrm{N}_p\big) \leqslant \dim \mathrm{N}_{p+1}$$

et

$$\dim f(\mathrm{H}) \leqslant d_p$$

Or, f est injective sur H donc $f(\mathrm{H})$ a la même dimension de H, c'est-à-dire d_{p+1}. On a bien prouvé $d_{p+1} \leqslant d_p$.

Par définition de p_0, la suite $(d_p)_{p\in\mathbb{N}}$ décroît strictement jusqu'au rang p_0, à partir duquel elle est constante nulle.

- Soient $p \in \mathbb{N}$ et $x \in \mathrm{K}_{p+1} \setminus \mathrm{K}_p$. Par définition, $f^p(x) \neq 0$ et $f^{p+1}(x) = 0$. Soient $\lambda_0, \dots, \lambda_p$ dans $\mathbb{K}$ tels que

$$\sum_{k=0}^{p} \lambda_k f^k(x) = 0$$

On applique f^p pour avoir

$$0 = \sum_{k=0}^{p} \lambda_k \underbrace{f^{k+p}(x)}_{=0 \text{ si } k\geqslant 1} = \lambda_0 f^p(x)$$

On en déduit $\lambda_0 = 0$. Une récurrence immédiate prouve qu'en appliquant successivement $f^p, \dots, f^0$, on obtiendra (dans l'ordre) que $\lambda_0, \dots, \lambda_p$ sont nuls. La famille $(x, \dots, f^p(x))$ est libre.

- Enfin, notons q la multiplicité de 0 dans π_f. Les sous-espaces N_{p_0} et I_{p_0} sont stables par f car f et f^{p_0} commutent. On note f_N et f_I les endomorphismes induits par f sur ces deux sous-espaces.

  f_N est nilpotent sur $N_{p_0} = \operatorname{Ker} f^{p_0}$ et $f_N^{p_0} = 0$. Donc le polynôme minimal π_{f_N} divise X^{p_0}. Mais par définition de

  $$p_0 = \operatorname{Min}\{p \in \mathbb{N} \mid N_p = N_{p+1}\}$$

  on a $N_{p_0-1} \subsetneq N_{p_0}$. Cela signifie que $f_N^{p_0-1} \neq 0$ donc π_{f_N} ne divise pas X^{p_0-1}. Par suite,

  $$\pi_{f_N} = X^{p_0}$$

  D'autre part, $\pi_f(f) = 0$ sur E, donc $\pi_f(f) = 0$ sur N_{p_0}. Ce qui veut dire que $\pi_f(f_N) = 0$. Il s'ensuit que $\pi_{f_N} = X^{p_0}$ divise π_f. D'où $p_0 \leqslant q$.

  Pour l'autre inégalité : $X^{p_0}\chi_{f_I}$ annule f sur N_{p_0} (puisque $f^{p_0} = 0$ sur N_{p_0}) et sur I_{p_0} (parce que $\chi_{f_I}(f_I) = 0$ d'après **Cayley-Hamilton**). Comme $N_{p_0} \oplus I_{p_0} = E$, $X^{p_0}\chi_{f_I}$ annule f sur E : le polynôme minimal π_f divise $X^{p_0}\chi_{f_I}$. Comme 0 n'est pas racine de χ_{f_I}, on a $q \leqslant p_0$. □

Ce lemme, technique mais assez simple est au cœur de la réduction des endomorphismes nilpotents, comme on le verra dans la prochaine section. Il nous donne, en outre, l'interprétation géométrique qui nous manquait pour l'ordre de multiplicité d'une valeur propre dans le polynôme minimal.

Corollaire 4.3.7
Soit λ une valeur propre de f. On note $q_\lambda \in [\![1\,;n]\!]$ sa multiplicité dans π_f. Alors

$$q_\lambda = Min\{q \in \mathbb{N} \mid Ker(f-\lambda)^q = Ker(f-\lambda)^{q+1}\}$$

Preuve : Il suffit essentiellement d'appliquer ce qui précède à l'endomorphisme $u = f - \lambda \mathrm{id}_E$. Notre **théorème des noyaux emboîtés** assure que l'ordre de multiplicité de 0 comme racine de π_u est précisément

$$\operatorname{Min}\{q \in \mathbb{N} \mid \operatorname{Ker} u^q = \operatorname{Ker} u^{q+1}\} = \operatorname{Min}\{q \in \mathbb{N} \mid \operatorname{Ker}(f-\lambda)^q = \operatorname{Ker}(f-\lambda)^{q+1}\}$$

Mais d'après la **proposition 2.21**, $\pi_f = \pi_{u+\lambda \mathrm{id}_E} = \pi_u \circ (X-\lambda)$; donc la multiplicité de 0 dans π_u est aussi la multiplicité de λ dans π_f. □

À l'aide de tout ce qu'on a fait, on a enfin une méthode géométrique pour trouver le polynôme minimal de f, en supposant que χ_f est scindé et qu'on a réussi à le factoriser. En effet,

- Les racines de π_f sont exactement les valeurs propres de f (**corollaire 2.20**).
- On prend une valeur propre λ et on calcule les $(\operatorname{Ker}(f-\lambda)^p)_{p\in\mathbb{N}}$, jusqu'au premier moment où la dimension atteint la multiplicité algébrique m_λ (**proposition 3.5**).
- L'entier p_0 obtenu à ce moment est précisément q_λ, multiplicité de λ dans π_f (**corollaire 3.7**).

4.3.3 La décomposition de Jordan

La décomposition de Jordan est la forme la plus poussée de réduction d'un endomorphisme : il s'agit de la forme matricielle qui a le plus de zéros possibles. Commençons par définir ce qu'est un bloc de Jordan :

Définition 4.3.8 (Blocs de Jordan)
Soient $k \geqslant 2$ un entier et $\alpha \in \mathbb{K}$. On appelle *bloc de Jordan d'ordre k relatif à* α la matrice de type $k \times k$ suivante :

$$\mathrm{J}_k(\alpha) = \begin{bmatrix} \alpha & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & \alpha \end{bmatrix}$$

On appelle *bloc de Jordan d'ordre* 1 *relatif à* α la matrice

$$\mathrm{J}_1(\alpha) = [\ \alpha\ ] \in \mathrm{M}_1(\mathbb{K})$$

Le **théorème de Jordan** va nous assurer qu'un endomorphisme dont le polynôme caractéristique est scindé peut être exprimé, dans une base bien choisie, par une matrice comportant uniquement des blocs de Jordan sur la diagonale (tous les autres blocs étant nuls).

Commençons par analyser les blocs de Jordan eux-même, avec quelques observations simples. On fixe $k \in \mathbb{N}^\star$ et on note $(e_1, \ldots, e_k)$ la base canonique de $\mathbb{K}^k$.

1. Si $\alpha \in \mathbb{K}$, alors $\mathrm{J}_k(\alpha) - \alpha \mathrm{I}_k = \mathrm{J}_k(0)$. Donc on s'intéresse aux propriétés de $\mathrm{J}_k(0)$, qu'on pourra alors transférer à $\mathrm{J}_k(\alpha)$.
2. Le bloc $\mathrm{J}_1(0)$ est l'endomorphisme nul de $\mathbb{K}$. On a

$$\forall n \in \mathbb{N}^\star \qquad \mathrm{J}_1(0)^n = 0 \quad \text{et} \quad \mathrm{Ker}\, \mathrm{J}_1(0) = \mathbb{K}$$

   Cet endomorphisme étant extrêmement simple à comprendre, on suppose que $k \geqslant 2$.
3. Le comportement de $\mathrm{J}_k(0)$ sur la base $(e_1, \ldots, e_k)$ est facile à lire : chaque vecteur de base est envoyé sur le précédent, sauf e_1 qui part sur 0.

$$\forall p \in [\![1\,;\, k]\!] \qquad \mathrm{J}_k e_p = \begin{cases} 0 & \text{si } p = 1 \\ e_{p-1} & \text{sinon} \end{cases}$$

4. Il est alors évident que si $m \in [\![1\,;\, k-1]\!]$ et $p \in [\![1\,;\, k]\!]$, la matrice $\mathrm{J}_k(0)^m$ envoie $e_1, \ldots, e_m$ sur 0, et décale $e_{m+1}, \ldots, e_k$ vers $e_1, \ldots, e_{k-m}$:

$$\mathrm{J}_k(0)^m = \begin{bmatrix} 0 & \cdots & 0 & 1 & & \\ & \ddots & & \ddots & \ddots & \\ & & 0 & \cdots & 0 & 1 \\ & & & 0 & \cdots & 0 \\ & & & & \ddots & \vdots \\ & & & & & 0 \end{bmatrix}$$

   Du coup, $\qquad \mathrm{Im}\, \mathrm{J}_k(0)^m = \mathrm{Vect}\,(e_1, \ldots, e_{k-m}) \qquad \mathrm{Ker}\, \mathrm{J}_k(0)^m = \mathrm{Vect}\,(e_1, \ldots, e_m)$

La première étape (et la plus importante) pour prouver notre théorème consiste à réduire les endomorphismes nilpotents.

Lemme 4.3.9 (Forme de Jordan pour les nilpotents)
Soient E *de dimension finie non nulle,* f *un endomorphisme nilpotent de* E. *Alors il existe une base* $\mathscr{B}$ *de* E *telle que* $\mathrm{Mat}_{\mathscr{B}}\, f$ *est diagonale par blocs, avec des blocs de Jordan relatifs à 0 sur la diagonale.*

Preuve : On commence par remarquer que le théorème est facile dans deux cas :

- Si $f = 0$. En effet, en choisissant n'importe quelle base de E, la matrice de f est nulle : elle est donc diagonale par blocs, avec n blocs diagonaux $\mathrm{J}_1(0)$.
- Si $f^n = 0$ et $f^{n-1} \neq 0$. Comme vu dans le **lemme des noyaux emboîtés**, si $x \in \mathrm{E}$ est tel que $f^{n-1}x \neq 0$, alors $\mathscr{B} = (f^{n-1}x, \dots, x)$ est libre et c'est une base de E. La matrice de f dans $\mathscr{B}$ est le bloc de Jordan $\mathrm{J}_n(0)$.

La suite est une preuve par récurrence assez technique, sur la dimension de l'espace. On définit donc, pour tout $n \in \mathbb{N}$, l'assertion $\mathscr{P}(n)$: « Soient E de dimension n, f un endomorphisme nilpotent de E. Il existe une base de E dans laquelle la matrice de f est diagonale par blocs, avec des blocs de Jordan relatifs à 0 sur la diagonale. »

$\mathscr{P}(1)$ ne pose aucun problème : en dimension 1, les nilpotents sont nuls et ce cas a déjà été traité ci-dessus.

C'est maintenant que les choses se compliquent. On se donne $n \in \mathbb{N}^\star$ tel que $\mathscr{P}(1), \dots, \mathscr{P}(n)$ sont vraies. On prend E de dimension $n+1$ et $f \in \mathscr{L}(\mathrm{E})$, nilpotent. On note p le degré de π_f. Si $p = 1$, c'est que $f = 0$ et le problème est déjà réglé. Si $p = n+1$, c'est bon également. On suppose donc que f n'est pas nul, et que $p \in [\![1\,;n]\!]$. On rappelle que

$$p = \mathrm{Min}\,\{k \in \mathbb{N} \mid f^k = 0\} = \mathrm{Min}\,\{k \in \mathbb{N} \mid \mathrm{Ker}\, f^k = \mathrm{Ker}\, f^{k+1}\}$$

- Première étape : construire un premier bloc de Jordan.
  Par définition de p, $f^{p-1} \neq 0$ et $f^p = 0$. On prend un x dans $\mathrm{Ker}\, f^p \setminus \mathrm{Ker}\, f^{p-1}$. D'après le **lemme des noyaux emboîtés**, la famille $(x, \dots, f^{p-1})$ est libre. On pose

  $$\forall k \in [\![0\,;p-1]\!] \qquad e_k = f^{p-1-k}x \qquad \mathscr{B} = (e_0, \dots, e_{p-1}) \qquad \mathrm{U} = \mathrm{Vect}\,\mathscr{B}$$

  On a $$\forall k \in [\![0\,;p-1]\!] \qquad f(e_k) = f^{p-k}x = \begin{cases} 0 & \text{si } k = 0 \\ e_{k-1} & \text{si } k \neq 0 \end{cases}$$

  U est stable par f, qui induit un endomorphisme f_U sur cet espace. En outre, les relations ci-dessus montrent que $\mathrm{Mat}_\mathscr{B}\, f_\mathrm{U}$ est le bloc de Jordan $\mathrm{J}_p(0)$.
- Deuxième étape : trouver un supplémentaire de U stable par f.
  On complète $\mathscr{B}$ en une base $\mathscr{C} = (e_0, \dots, e_n)$. Notons que c'est parfaitement possible, parce que U est de dimension $p < n+1$. On construit alors la base duale $\mathscr{C}^\star = (e_0^\star, \dots, e_n^\star)$ et on considère l'ensemble

  $$\mathrm{F} = \{v \in \mathrm{E} \mid \forall m \in \mathbb{N} \quad e_0^\star(f^m(v)) = 0\}$$

  Voyons ses propriétés.

  1. F est un espace vectoriel et on étudie sa dimension. F est un pré-orthogonal :

     $$\mathrm{F} = \{e_0^\star \circ f^m \mid m \in \mathbb{N}\}^\circ$$

     Donc c'est un sous-espace de E. Mais en fait, $f^p = 0$ par définition de p donc

     $$\mathrm{F} = \{e_0^\star \circ f^m \mid m \in [\![0\,;p-1]\!]\}^\circ = \big(\mathrm{Vect}\,(e_0^\star \circ f^m)_{0 \leqslant m \leqslant p-1}\big)^\circ$$

     Or, $$\dim \mathrm{Vect}\,(e_0^\star \circ f^m)_{0 \leqslant m \leqslant p-1} \leqslant p$$

     donc $$\dim \mathrm{F} = n+1 - \dim \mathrm{Vect}\,(e_0^\star \circ f^m)_{0 \leqslant m \leqslant p-1} \geqslant n+1-p > 0$$

2. F est stable par f : en effet, si $v \in \mathrm{F}$, par définition
$$\forall m \in \mathbb{N} \qquad e_0^\star(f^m(v)) = 0$$
et
$$\forall m \in \mathbb{N} \qquad e_0^\star(f^m(f(v)) = e_0^\star(f^{m+1}(v)) = 0$$
ce qui prouve que $f(v) \in \mathrm{F}$.
3. F est en somme directe avec U : Si $v \in \mathrm{F} \cap \mathrm{U}$, il est combinaison linéaire de $e_0, \ldots, e_{p-1}$ et on peut trouver $\lambda_0, \ldots, \lambda_{p-1} \in \mathbb{K}$ tels que
$$v = \sum_{k=0}^{p-1} \lambda_k e_k$$
Alors
$$\forall m \in [\![1\,;p-1]\!] \qquad f^m(v) = \sum_{k=0}^{p-1} \lambda_k \underbrace{f^{m+p-1-k} x}_{=0 \text{ si } m+p-1-k \geqslant p} = \sum_{k=m}^{p} \lambda_k e_{k-m}$$
Mais par définition de la base duale, $e_0^\star$ s'annule sur $e_1, \ldots, e_{p-1}$ et vaut 1 sur e_0. Donc
$$\forall m \in [\![1\,;p-1]\!] \qquad 0 \underbrace{=}_{\text{car } v \in \mathrm{F}} e_0^\star(f^m(v)) = \sum_{k=m-1}^{p-1} \lambda_k \underbrace{e_0^\star(e_{k-m})}_{=0 \text{ si } k-m>0} = \lambda_m$$
Ainsi, $\lambda_0, \ldots, \lambda_{p-1}$ sont nuls : F et U sont en somme directe.

Mais alors
$$\dim(\mathrm{U} \oplus \mathrm{F}) = \dim \mathrm{U} + \dim \mathrm{F} \geqslant p + n + 1 - p = n + 1$$
d'où
$$\mathrm{U} \oplus \mathrm{F} = \mathrm{E}$$

- On réduit f sur F : Puisque F est stable par f, on a un endomorphisme induit f_{F}. Bien sûr, f_{F} est nilpotent puisque f l'est. De plus,
$$\dim \mathrm{U} = p \in [\![1\,;n]\!] \qquad \text{donc} \qquad \dim \mathrm{F} \in [\![1\,;n]\!]$$
Ceci nous permet d'appliquer l'hypothèse de récurrence : on peut trouver une base $\mathscr{D}$ de F telle que $\mathrm{Mat}_{\mathscr{D}} f_{\mathrm{F}}$ est diagonale par blocs, avec des blocs de Jordan relatifs à 0 sur la diagonale.
Alors $(\mathscr{B}, \mathscr{D})$ est une base de E dans laquelle la matrice de f est sous forme diagonale par blocs, avec des blocs de Jordan sur la diagonale. Ceci achève de prouver $\mathscr{P}(n+1)$. □

Ceci étant fait, on s'intéresse maintenant au nombre de blocs de Jordan de chaque taille, après avoir réduit un endomorphisme nilpotent. On se place dans un espace vectoriel E de dimension $n \neq 0$, et on prend un endomorphisme nilpotent f. On note, comme dans le **lemme des noyaux emboîtés** :
$$\forall p \in \mathbb{N} \qquad \mathrm{N}_p = \mathrm{Ker}\, f^p \qquad d_p = \dim \mathrm{Ker}\, f^{p+1} - \mathrm{Ker}\, f^p$$
et
$$q = \mathrm{Min}\,\{p \in \mathbb{N} \mid \mathrm{N}_p = \mathrm{N}_{p+1}\} = \mathrm{Min}\,\{p \in \mathbb{N} \mid \mathrm{N}_p = \mathrm{E}\}$$
Alors
$$\chi_f = (-\mathrm{X})^n \qquad \pi_f = \mathrm{X}^q$$

et on rappelle que $$\mathrm{N}_0 = \{0\} \qquad \mathrm{N}_q = \mathrm{E}$$

D'après le **lemme 3.9**, on peut trouver une base $\mathscr{B}$ de E dans laquelle la matrice de f, notée M, est diagonale par blocs, avec des blocs de Jordan relatifs à 0 sur la diagonale. Pour chaque $k \in [\![1\,;n]\!]$, on note a_k le nombre de blocs $\mathrm{J}_k(0)$ dans cette matrice.

1. La matrice de f^q dans $\mathscr{B}$ est la matrice nulle. Soit $k \in [\![1\,;n]\!]$ tel que $a_k \neq 0$. Cette matrice comporte donc au moins un bloc $\mathrm{J}_k(0)^q$ sur la diagonale. Donc $\mathrm{J}_k(0)^q = 0$; mais le polynôme minimal de J_k est X^k, comme l'a montré l'étude préliminaire des blocs de Jordan. D'où $k \leqslant q$.
$$\forall k \in [\![1\,;n]\!] \qquad a_k \geqslant 1 \implies k \leqslant q$$
2. La matrice de f est de taille $n \times n$ donc
$$\sum_{k=1}^{q} k a_k = n = \dim \mathrm{N}_q \tag{$\star$}$$
3. Fixons $p \in [\![1\,;q-1]\!]$. La matrice de f^p dans $\mathscr{B}$ est diagonale par blocs ; si $k \in [\![1\,;q]\!]$, elle comporte a_k blocs $\mathrm{J}_k(0)^p$ sur la diagonale.
   - Si $k \leqslant p$, alors $\mathrm{J}_k(0)^p = 0$ et cela fournit k vecteurs de $\mathscr{B}$ qui sont dans $\operatorname{Ker} f^p$.
   - Si $p+1 \leqslant k \leqslant q$, on a vu que $\mathrm{J}_k(0)^p$ a un noyau de dimension p. On obtient donc p vecteurs de $\mathscr{B}$ qui sont dans $\operatorname{Ker} f^p$. De plus, $\mathrm{J}_k(0)^p$ est de rang $k-p$.

   Tous les vecteurs de $\operatorname{Ker} f^p$ obtenus en faisant varier k dans $[\![1\,;q]\!]$ forment une sous-famille de $\mathscr{B}$, donc sont libres.

   D'où les relations $$\dim \operatorname{Ker} f^p \geqslant \sum_{k=1}^{p} k a_k + p \sum_{k=p+1}^{q} a_k \qquad \operatorname{rg} f^p \geqslant \sum_{k=p+1}^{q} (k-p) a_k$$

   Si une de ces inégalités était stricte, on obtiendrait $n > n$, en utilisant $(\star)$, et le théorème du rang. Ainsi,
$$\forall p \in [\![1\,;q-1]\!] \qquad \dim \mathrm{N}_p = \sum_{k=1}^{p} k a_k + p \sum_{k=p+1}^{q} a_k$$
4. On obtient alors $$\forall p \in [\![0\,;q-1]\!] \qquad d_p = \dim \mathrm{N}_{p+1} - \dim \mathrm{N}_p = \sum_{k=p+1}^{q} a_k$$

   C'est un simple calcul, qui utilise le fait que $\mathrm{N}_0 = \{0\}$, et la relation $(\star)$ pour obtenir d_{q-1}. Il permet d'en déduire
$$\forall p \in [\![1\,;q]\!] \qquad d_{p-1} - d_p = a_p$$

   À nouveau, il faut faire un petit cas particulier pour avoir $d_{q-1} - d_q = a_q$: on utilise le fait que
$$d_q = \dim \mathrm{N}_{q+1} - \dim \mathrm{N}_q = n - n = 0$$

   tandis que $$d_{q-1} = \sum_{k=q}^{q} a_k = a_q$$

Ces relations sont suffisamment intéressantes pour en faire un théorème :

Lemme 4.3.10 (Forme de Jordan pour les nilpotents – Version 2)
Soient E *de dimension finie et* $f \in \mathscr{L}(\mathrm{E})$*, nilpotent. Il existe une base de* E *dans laquelle la matrice de* f *est diagonale par blocs, avec des blocs de Jordan relatifs à 0 sur la diagonale. De plus, pour chaque* $k \in \mathbb{N}^\star$*, le nombre de blocs* $\mathrm{J}_k(0)$ *est*

$$dim\,Ker\,f^{k-1} - 2dim\,Ker\,f^k + dim\,Ker\,f^{k+1}$$

Voyons sur un exemple comment obtenir en pratique la forme de Jordan.

Exemple 4.3.11
On suppose E de dimension 28. Soit $f \in \mathscr{L}(\mathrm{E})$, nilpotent. Après avoir fait plein de calculs de noyaux, on a trouvé que f est nilpotent d'ordre 4 et on a identifié les dimensions de tous les sous-espaces $(\mathrm{Ker}\, f^p)_{p\in\mathbb{N}}$.

p	$\dim \mathrm{Ker}\, f^p$	d_p
0	0	11
1	11	8
2	19	6
3	25	3
4	28	0

k	Nombre de blocs $\mathrm{J}_k(0)$
1	$d_0 - d_1 = 3$
2	$d_1 - d_2 = 2$
3	$d_2 - d_3 = 3$
4	$d_3 - d_4 = 3$

On doit pouvoir trouver une base de E dans laquelle la matrice de f est

$$\begin{bmatrix}
\begin{matrix}0&1&0&0\\0&0&1&0\\0&0&0&1\\0&0&0&0\end{matrix} \\
& \begin{matrix}0&1&0&0\\0&0&1&0\\0&0&0&1\\0&0&0&0\end{matrix} \\
& & \begin{matrix}0&1&0&0\\0&0&1&0\\0&0&0&1\\0&0&0&0\end{matrix} \\
& & & \begin{matrix}0&1&0\\0&0&1\\0&0&0\end{matrix} \\
& & & & \begin{matrix}0&1&0\\0&0&1\\0&0&0\end{matrix} \\
& & & & & \begin{matrix}0&1&0\\0&0&1\\0&0&0\end{matrix} \\
& & & & & & \begin{matrix}0&1\\0&0\end{matrix} \\
& & & & & & & \begin{matrix}0&1\\0&0\end{matrix} \\
& & & & & & & & 0 \\
& & & & & & & & & 0 \\
& & & & & & & & & & 0
\end{bmatrix}$$

Pour trouver une telle base, il suffit de suivre les instructions données par la matrice. Le schéma suivant permet aussi de suivre la procédure :

$$\begin{array}{llllr} \mathrm{Ker}\, f^4 & & & & 3 \text{ boîtes} \\ & \mathrm{Ker}\, f^3 & & & 6 \text{ boîtes} \\ & & \mathrm{Ker}\, f^2 & & 8 \text{ boîtes} \\ & & & \mathrm{Ker}\, f & 11 \text{ boîtes} \end{array}$$

Les « boîtes » représentent les vecteurs de base qu'il faut trouver pour constituer une base de E. Il y a 11 boîtes pour $\mathrm{Ker}\, f$, 8 de plus pour $\mathrm{Ker}\, f^2$ (soit un total de 19, qui est la dimension de $\mathrm{Ker}\, f^2$), 6 de plus pour compléter $\mathrm{Ker}\, f^3$ (qui est de dimension 25), puis les 3 qui manquent pour avoir $\mathrm{Ker}\, f^4$, qui est égal à E. La procédure qui suit décrit comment les remplir.

- Commencer par prendre une famille libre (e_4, e_8, e_{12}) dans E, qui ne sont pas dans $\mathrm{Ker}\, f^3$. Construire ensuite

$$\begin{array}{ccc} e_3 = f(e_4) & e_2 = f(e_3) & e_1 = f(e_2) \\ e_7 = f(e_8) & e_6 = f(e_7) & e_5 = f(e_6) \\ e_{11} = f(e_{12} & e_{10} = f(e_{11}) & e_9 = f(e_{10}) \end{array}$$

  La famille $(e_1, \dots, e_{12})$ est libre : cela se vérifie simplement par la définition et en utilisant le fait que $\mathrm{Vect}(e_4, e_8, e_{12})$ est en somme directe avec $\mathrm{Ker}\, f^3$. Nos boîtes commencent donc à se remplir :

$$\begin{array}{l|lll} f & e_4 & e_8 & e_{12} \\ f & e_3 & e_7 & e_{11} \\ f & e_2 & e_6 & e_{10} \\ & e_1 & e_5 & e_9 \end{array}$$

- On a déjà 3 vecteurs libres (e_3, e_7, e_{11}) dans $\mathrm{Ker}\, f^3 \setminus \mathrm{Ker}\, f^2$. On cherche alors trois vecteurs (e_{15}, e_{18}, e_{21}) qui les complètent pour former un supplémentaire de $\mathrm{Ker}\, f^2$ dans $\mathrm{Ker}\, f^3$. Puis on définit

$$\begin{array}{cc} e_{14} = f(e_{15}) & e_{13} = f(e_{14}) \\ e_{17} = f(e_{18}) & e_{16} = f(e_{17}) \\ e_{20} = f(e_{21}) & e_{19} = f(e_{20}) \end{array}$$

  et on vérifie que $(e_1, \dots, e_{21})$ est libre. Nos « dimensions » sont remplies de la manière suivante :

$$\begin{array}{l|llllll} f & e_4 & e_8 & e_{12} \\ f & e_3 & e_7 & e_{11} & e_{15} & e_{18} & e_{21} \\ f & e_2 & e_6 & e_{10} & e_{14} & e_{17} & e_{20} \\ & e_1 & e_5 & e_9 & e_{13} & e_{16} & e_{19} \end{array}$$

- On choisit deux vecteurs e_{23}, e_{25} dans $\mathrm{Ker}\, f^2$ qui complètent $(e_2, e_6, e_{10}, e_{14}, e_{17}, e_{20})$ pour former un supplémentaire de $\mathrm{Ker}\, f$ dans $\mathrm{Ker}\, f^2$. À partir d'eux, on définit

$$\begin{array}{c} e_{22} = f(e_{23}) \\ e_{24} = f(e_{25}) \end{array}$$

  et la famille $(e_1, \dots, e_{25})$ est libre :

$$\begin{array}{l|llllllll} f & e_4 & e_8 & e_{12} \\ f & e_3 & e_7 & e_{11} & e_{15} & e_{18} & e_{21} \\ f & e_2 & e_6 & e_{10} & e_{14} & e_{17} & e_{20} & e_{23} & e_{25} \\ & e_1 & e_5 & e_9 & e_{13} & e_{16} & e_{19} & e_{22} & e_{24} \end{array}$$

- Enfin, on trouve trois vecteurs (e_{26}, e_{27}, e_{28}) dans $\mathrm{Ker}\, f$ qui complètent $(e_1, e_5, e_9, e_{13}, e_{16}, e_{19}, e_{22}, e_{24})$ en un supplémentaire de $\{0\}$ dans $\mathrm{Ker}\, f$. Et on a terminé de replir l'espace E :

$$\begin{array}{l|lllllllllll}
 & e_4 & e_8 & e_{12} \\
f \downarrow & e_3 & e_7 & e_{11} & e_{15} & e_{18} & e_{21} \\
f \downarrow & e_2 & e_6 & e_{10} & e_{14} & e_{17} & e_{20} & e_{23} & e_{25} \\
f \downarrow & e_1 & e_5 & e_9 & e_{13} & e_{16} & e_{19} & e_{22} & e_{24} & e_{26} & e_{27} & e_{28}
\end{array}$$

Par construction, $(e_1, \ldots, e_{28})$ est une base de E dans laquelle la matrice de f a la forme de Jordan trouvée précédemment.

En résumé, les vecteurs qu'on a cherchés sont $e_4, e_8, e_{12}, e_{15}, e_{18}, e_{21}, e_{23}, e_{25}, e_{26}, e_{27}$ et e_{28} ; tous les autres sont trouvés à partir d'eux, par applications successives de f.

Corollaire 4.3.12 (Décomposition de Jordan)
On suppose E *de dimension finie non nulle. Soit $f \in \mathscr{L}(\mathrm{E})$ tel que π_f (ou χ_f) est scindé. Il existe une base de* E *dans laquelle la matrice de f est diagonale par blocs, avec des blocs de Jordan relatifs aux valeurs propres de f sur la diagonale.*

Preuve : Il suffit d'appliquer ce qui précède aux sous-espaces caractéristiques. On rappelle qu'on a noté

$$\pi_f = \prod_{\lambda \in \mathrm{Sp}\, f} (\mathrm{X} - \lambda)^{q_\lambda}$$

Ce polynôme annule f donc

$$\mathrm{E} = \bigoplus_{\lambda \in \mathrm{Sp}\, f} \mathrm{K}_\lambda(f)$$

Si $\lambda \in \mathrm{Sp}\, f$, f induit un endomorphisme f_λ de $\mathrm{K}_\lambda(f)$. Et $f_\lambda - \lambda\, \mathrm{id}_{\mathrm{K}_\lambda(f)}$ est nilpotent. Donc on peut trouver une base $\mathscr{B}_\lambda$ de $\mathrm{K}_\lambda(f)$ dans laquelle la matrice de $f_\lambda - \lambda\, \mathrm{id}_{\mathrm{K}_\lambda(f)}$ est diagonale par blocs, avec des blocs de Jordan relatifs à 0 sur la diagonale. Du coup, la matrice de f_λ dans $\mathscr{B}_\lambda$ est de la même forme, avec des blocs de Jordan relatifs à λ.

En réunissant les bases $(\mathscr{B}_\lambda)_{\lambda \in \mathrm{Sp}\, f}$, on obtient une base de E dans laquelle la matrice de f a la forme voulue. $\square$