Réduction des endomorphismes

Thomas Dedieu

Rentrée 2022
compilé le 27 juin 2023

Avertissement. Malgré mes efforts, ce texte n'est pas encore complètement rédigé. En particulier il comporte des trous et des passages en style sténodactylographique. Néanmoins il a vocation à ne pas contenir d'erreur. Il en contient malgré tout certainement, et si vous en trouvez il m'intéresse que vous me les communiquiez.

Table des matières

1 – Introduction

1.1 – Composantes selon une somme directe

1.1 Définition. Soit $E_1, \dots, E_r$ des sous-espaces vectoriels de E. On dit que $E_1, \dots, E_r$ sont en *somme directe*, ou que la somme $\sum_{i=1}^r E_i$ est directe, si la condition suivante est vérifiée :

$$\forall (x_1, \dots, x_r) \in E_1 \times \cdots \times E_r : \quad x_1 + \cdots + x_r = 0 \iff x_1 = \cdots = x_r = 0. \tag{1.1.1}$$

On laisse au lecteur le soin de vérifier que la condition (1.1.1) est équivalente à la condition suivante :

$$\forall (x_1, \dots, x_r), (x'_1, \dots, x'_r) \in E_1 \times \cdots \times E_r : \quad x_1 + \cdots + x_r = x'_1 + \cdots + x'_r \iff x_1 = x'_1, \dots, x_r = x'_r. \tag{1.1.2}$$

Ainsi, si la somme $\sum_{i=1}^r E_i$ est directe, alors pour tout vecteur $x \in \sum_{i=1}^r E_i$, il existe un unique r-uplet $(x_1, \dots, x_r) \in E_1 \times \cdots \times E_r$ tel que $x = x_1 + \cdots + x_r$. Dans ces conditions, pour tout $i_0 = 1, \dots, r$, on dit que le vecteur $x_{i_0} \in E_{i_0}$ est la *composante de x selon E_{i_0} relativement à la somme directe* $\bigoplus_{i=1}^r E_i$. Souvent, on omettra par abus de langage de dire "relativement à la somme directe $\bigoplus_{i=1}^r E_i$" ; il est important de garder à l'esprit qu'il s'agit d'un abus de langage.

1.2 Proposition. *Soit $E_1, \dots, E_r$ des sous-espaces vectoriels de E tels que $E = \bigoplus_{i=1}^r E_i$. Soit $i_0 \in [\![1, r]\!]$. L'application qui à tout vecteur $x \in E$ associe sa composante selon E_{i_0} relativement à la somme directe $\bigoplus_{i=1}^r E_i$ est une application linéaire de E dans E_{i_0}.*

Dans l'énoncé ci-dessus, on sera attentif au fait que l'hypothèse "$E = \bigoplus_{i=1}^r E_i$" affirme deux choses : (i) $E = \sum_{i=1}^r E_i$, autrement dit tout vecteur $x \in E$ peut s'écrire comme somme de vecteurs $x_1 \in E_1, \dots, x_r \in E_r$, et (ii) la somme $\sum_{i=1}^r E_i$ est directe. La condition "$E = \bigoplus_{i=1}^r E_i$" est équivalente à la condition suivante :

$$\forall x \in E : \quad \exists ! (x_1, \dots, x_r) \in E_1 \times \cdots \times E_r \quad \text{tel que} \quad x = x_1 + \cdots + x_r.$$

Une nouvelle fois, le lecteur auquel est destiné ce livre devrait être en mesure d'établir lui-même la Proposition 1.2, et nous lui conseillons vivement de le faire.

1.3 Proposition. *Soit E, F deux espaces vectoriels, munis respectivement des décompositions en sommes directes $E = \bigoplus_{j=1}^r E_j$ et $F = \bigoplus_{i=1}^s F_i$.*

1.3.1. *Soit $f \in \mathcal{L}(E, F)$. Pour tout $(i, j) \in [\![1, s]\!] \times [\![1, r]\!]$, l'application f_{ij} qui à $x \in E_j$ associe la composante de $f(x)$ selon F_i relativement à la somme directe $F_1 \oplus \cdots \oplus F_s$ est une application linéaire $f_{ij} \in \mathcal{L}(E_j, F_i)$.*

1.3.2. *Pour toute famille d'applications linéaires $(\tilde{f}_{ij}) \in \prod_{(i,j) \in [\![1,s]\!] \times [\![1,r]\!]} \mathcal{L}(E_j, F_i)$, il existe une unique application linéaire $f \in \mathcal{L}(E, F)$ telle que $f_{ij} = \tilde{f}_{ij}$, où les f_{ij} sont les applications associées à f comme en 1.3.1.*

On appellera *composantes de f selon les décompositions $E = \bigoplus_{j=1}^r E_j$ et $F = \bigoplus_{i=1}^s F_i$* les applications linéaires $f_{ij} \in \mathcal{L}(E_j, F_i)$ comme en 1.3.1 ci-dessus.

Preuve. Pour 1.3.1, il suffit d'observer que

$$f_{ij} = \tilde{p}_i \circ f|_{E_j} ,$$

où $\tilde{p}_i$ est l'application i-ème composante relativement à la décomposition $F = F_1 \oplus \cdots \oplus F_s$ comme en 1.2 : la restriction $f|_{E_j}$ est une application linéaire $E_i \to F$, et l'application i-ème composante $\tilde{p}_i$ est linéaire $F \to F_i$, donc la composition $\tilde{p}_i \circ f|_{E_j}$ est bien une application linéaire $E_j \to F_i$.

Pour 1.3.2, commençons par la partie "unicité". Soit $f, g \in \mathcal{L}(E,F)$ telles que $f_{ij} = g_{ij}$ pour tout $(i,j) \in [\![1,s]\!] \times [\![1,r]\!]$, où f_{ij} et g_{ij} sont définies à partir de f et g respectivement comme en 1.3.1. Il s'agit de démontrer que $f = g$. Soit $x \in E$, et écrivons le $x = x_1 + \cdots + x_r$, avec $(x_1, \ldots, x_r) \in E_1 \times \cdots \times E_r$. On a

$$\begin{aligned} f(x) = f\Big(\sum_{j=1}^{r} x_j\Big) &= \sum_{j=1}^{r} f(x_j) \\ &= \sum_{j=1}^{r}\sum_{i=1}^{s} f_{ij}(x_j); \end{aligned} \tag{1.3.1}$$

l'égalité de droite sur la première ligne provient de la linéarité de f, et celle de la seconde ligne du fait que pour tout j, $f_{1j}(x_j), \ldots, f_{sj}(x_j)$ sont les composantes de $f(x_j)$ selon la somme directe $F_1 \oplus \cdots \oplus F_s$. Ainsi, puisque $f_{ij} = g_{ij}$ pour tout (i,j), on a

$$\sum_{j=1}^{r}\sum_{i=1}^{s} f_{ij}(x_j) = \sum_{j=1}^{r}\sum_{i=1}^{s} g_{ij}(x_j),$$

et donc $f(x) = g(x)$, ce qui conclut la preuve du fait que $f = g$.

Enfin, démontrons la partie existence de 1.3.2. La formule (1.3.1) ci-dessus nous dit comment définir f, cette définition étant sans ambiguïté puisque $x_1, \ldots, x_r$ sont uniquement déterminés par x par (1.1.2). Une autre façon de formuler les choses est de dire que l'application linéaire

$$f = \sum_{j=1}^{r}\sum_{i=1}^{s} \iota_i \circ \tilde{f}_{ij} \circ \tilde{q}_j$$

convient, où $\tilde{q}_j \in \mathcal{L}(E, E_j)$ est l'application j-ème composante relativement à la décomposition $E = E_1 \oplus \cdots \oplus E_r$, et $\iota_i \in \mathcal{L}(F_i, F)$ est l'injection canonique $F_i \hookrightarrow F$. Noter que pour tout (i,j), $\iota_i \circ \tilde{f}_{ij} \circ \tilde{q}_j \in \mathcal{L}(E,F)$, donc la somme de toutes ces applications linéaires est bien définie. □

1.4 Exemple : projecteurs. Considérons le cas particulier de la Proposition 1.3 où $E = F$ est muni d'une seule décomposition $E = \bigoplus_{i=1}^{r} E_i$. Pour tout $i_0 \in [\![1,r]\!]$, il existe un unique endomorphisme $p_{i_0} \in \mathcal{L}(E)$ tel que

$$\forall x \in E_{i_0} :\ p_{i_0}(x) = x, \quad \text{et} \quad \forall x \in \bigoplus_{i \neq i_0} E_i :\ p_{i_0}(x) = 0. \tag{1.4.1}$$

On peut le voir en appliquant 1.3.2 à la collection $(\tilde{f}_{ij})_{(i,j)\in[\![1,r]\!]^2}$ définie par

$$\tilde{f}_{ij} = \begin{cases} \mathrm{id}_{E_{i_0}} & \text{si } (i,j) = (i_0, i_0) \\ 0 & \text{sinon.} \end{cases}$$

L'endomorphisme $p_{i_0} \in \mathcal{L}(E)$ s'appelle le *projecteur sur E_{i_0} relativement à la décomposition* $E = \bigoplus_{i=1}^{r} E_i$. On notera la subtile différence entre $p_{i_0} \in \mathcal{L}(E)$ et l'application i_0-ème composante $\tilde{p}_{i_0} \in \mathcal{L}(E, E_{i_0})$: pour $x \in E$ décomposé en $x = x_1 + \cdots + x_r$ selon $E = \bigoplus_{i=1}^{r} E_i$, on a $p_{i_0}(x) = \tilde{p}_{i_0}(x) = x_{i_0}$, mais dans le premier cas x_{i_0} est considéré comme un vecteur de E, tandis que dans le second cas il est considéré comme un vecteur de E_{i_0} ; autrement dit on a la relation $p_i = \iota_i \circ \tilde{p}_i$, où ι_i désigne l'inclusion $E_i \hookrightarrow E$. Il suit des conditions (1.4.1) que le projecteur p_i dépend seulement de E_{i_0} et de son supplémentaire $\bigoplus_{i \neq i_0} E_i$, et non pas de toute la décomposition $E = \bigoplus_{i=1}^{r} E_i$.

Soit F et G deux sous-espaces vectoriels supplémentaires dans E. Le *projecteur sur F dans la direction de G* est l'endomorphisme $p \in \mathcal{L}(E)$ défini par les conditions :

$$\forall x \in F :\ p(x) = x, \quad \text{et} \quad \forall x \in G :\ p(x) = 0.$$

On prendra bien garde au fait que le projecteur p dépend à la fois de F et de G, comme nous l'illustrons dans la figure ci-dessous [1] (on note p et q les projections sur F et G relativement à $F \oplus G$, et p' et q' les projections sur F et G' relativement à $F \oplus G'$).

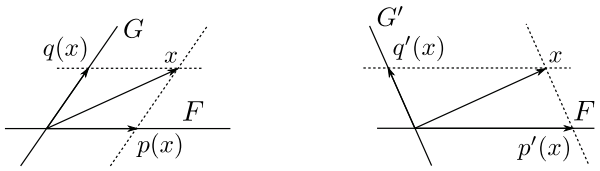


FIGURE 1 – Projections sur F dans les directions de G et G' respectivement

Un endomorphisme $f \in \mathcal{L}(E)$ est un *projecteur* s'il existe deux sous-espaces F et G supplémentaires dans E tels que f est le projecteur sur F dans la direction de G.

1.5 Proposition. *Un endomorphisme $p \in \mathcal{L}(E)$ est un projecteur si et seulement si $p \circ p = p$.*

Nous proposons la preuve de cette proposition à titre d'exercice. Il faut démontrer que p est le projecteur sur $\mathrm{im}(p)$ dans la direction de $\ker(p)$ (en particulier, si $p \circ p = p$, ces deux sous-espaces sont supplémentaires).

Pour poursuivre l'exercice, voici un autre énoncé à démontrer.

1.6 Lemme. *Considérons une décomposition $E = \bigoplus_{i=1}^{r} E_i$, et $p_1, \ldots, p_r \in \mathcal{L}(E)$ les projecteurs associés. Alors on a*

$$p_1 + \cdots + p_r = \mathrm{id}_E.$$

1.2 – Matrices par blocs

1.7 Sommes directes. On dit que $E_1 \oplus \cdots \oplus E_r = E$ si pour tout $x \in E$ il existe un unique $(x_1, \ldots, x_r) \in E_1 \times \cdots \times E_r$ tel que $x = x_1 + \cdots + x_r$.

1.8 Décomposition d'une application linéaire selon une somme directe. On suppose $E = E_1 \oplus \cdots \oplus E_r$ et $F = F_1 \oplus \cdots \oplus F_s$. Soit $f \in \mathcal{L}(E, F)$. Il existe une unique famille d'applications linéaires $f_{ij} \in \mathcal{L}(E_j, F_i)$ telle que

$$\forall (x_1, \ldots, x_r) \in E_1 \times \cdots \times E_r,\ f(x_1 + \cdots + x_r) = \sum\nolimits_{ij} f_{ij}(x_j). \tag{1.8.1}$$

Réciproquement, pour toute famille d'applications linéaires $f_{ij} \in \mathcal{L}(E_j, F_i)$, il existe une unique $f \in \mathcal{L}(E, F)$ telle que (1.8.1) soit vérifiée.

1.8.1 Exemple. Pour $E = E_1 + E_2$, définition de l'endomorphisme de E projecteur sur E_1 dans la direction de E_2, ainsi que de la projection de E sur E_1 dans la direction de E_2.

1.8.2 Remarque. On a $f_{ij} = \pi_{F,F_i} \circ f \circ \iota_{E_j,E} \in \mathcal{L}(E_j, F_i)$.

1.9 Bases compatibles à une somme directe. Si $\mathcal{B}_1, \ldots, \mathcal{B}_r$ sont des bases de $E_1, \ldots, E_r$ respectivement et $E = E_1 \oplus \cdots \oplus E_r$, alors $\mathcal{B} = (\mathcal{B}_1, \ldots, \mathcal{B}_r)$ est une base de E.

Dans les conditions du 1.8, si $\mathcal{B} = (\mathcal{B}_1, \ldots, \mathcal{B}_r)$ et $\mathcal{D} = (\mathcal{D}_1, \ldots, \mathcal{D}_s)$ sont des bases de E et F respectivement, compatibles aux sommes directes, alors $\mathrm{Mat}_{\mathcal{B},\mathcal{D}}(f)$ se décompose par blocs $\mathrm{Mat}_{\mathcal{B}_j,\mathcal{D}_i}(f_{ij})$.

1. admirez au passage l'illusion d'optique : nous vous jurons que si vous superposez les deux figures, vous verrez le vecteur x est bien le même à gauche et à droite.

1.10 Proposition. *Soit $E = \bigoplus_{j=1}^{r} E_j$, $F = \bigoplus_{i=1}^{s} F_i$, et $f \in \mathcal{L}(E,F)$. Dans des bases $\mathcal{B} = (\mathcal{B}_1, \ldots, \mathcal{B}_r)$ et $\mathcal{D} = (\mathcal{D}_1, \ldots, \mathcal{D}_s)$ compatibles aux décompositions de E et F respectivement, la matrice de f se décompose par blocs*

$$\mathrm{Mat}_{\mathcal{B},\mathcal{D}}(f) = \left(\begin{array}{c|c|c} M_{11} & \cdots & M_{1r} \\ \hline \vdots & & \vdots \\ \hline M_{s1} & \cdots & M_{sr} \end{array}\right)$$

où pour tout i, j, M_{ij} est la matrice dans les bases $\mathcal{B}_j$ et $\mathcal{D}_i$ de la composante $f_{ij} \in \mathcal{L}(E_j, F_i)$ de f.

Preuve. On note $\mathcal{B}_j = (e_1^j, \ldots, e_{n_j}^j)$ pour tout $j = 1, \ldots, r$ (ainsi $n_1, \ldots, n_r$ sont les dimensions respectives de $E_1, \ldots, E_r$), et $\mathcal{D}_i = (\varepsilon_1^i, \ldots, \varepsilon_{m_i}^i)$ pour tout $i = 1, \ldots, s$ (ainsi $m_1, \ldots, m_s$ sont les dimensions respectives de $F_1, \ldots, F_s$).

On écrit $\mathrm{Mat}_{\mathcal{B},\mathcal{D}}(f) = (a_k^l)_{1 \leqslant k \leqslant m,\, 1 \leqslant l \leqslant n}$, avec $m = m_1 + \cdots + m_s = \dim(F)$ et $n = n_1 + \cdots + n_s = \dim(E)$, et on découpe cette matrice en blocs de tailles $m_i \times n_j$ comme indiqué ci-dessous.

$$\begin{array}{c} \\ m_1 \updownarrow \\ \\ m_s \updownarrow \end{array} \begin{array}{c} \overset{\longleftrightarrow}{n_1} \qquad\qquad \overset{\longleftrightarrow}{n_r} \\ \left(\begin{array}{c|c|c} M_{11} & \cdots & M_{1r} \\ \hline \vdots & & \vdots \\ \hline M_{s1} & \cdots & M_{sr} \end{array}\right) \end{array}$$

Fixons $i_0 \in [\![1, s]\!]$ et $j_0 \in [\![1, r]\!]$. Le bloc $M_{i_0 j_0}$, de taille $m_{i_0} \times n_{j_0}$, est

$$M_{i_0 j_0} = \begin{pmatrix} a_{m_1+\cdots+m_{i_0-1}+1}^{n_1+\cdots+n_{j_0-1}+1} & \cdots & a_{m_1+\cdots+m_{i_0-1}+1}^{n_1+\cdots+n_{j_0-1}+n_{j_0}} \\ \vdots & & \vdots \\ a_{m_1+\cdots+m_{i_0-1}+m_{i_0}}^{n_1+\cdots+n_{j_0-1}+1} & \cdots & a_{m_1+\cdots+m_{i_0-1}+m_{i_0}}^{n_1+\cdots+n_{j_0-1}+n_{j_0}} \end{pmatrix}.$$

On veut montrer que c'est la matrice de la composante $f_{i_0 j_0}$ dans les bases $\mathcal{B}_{j_0}$ et $\mathcal{D}_{i_0}$. Il faut donc montrer que pour tout $q \in [\![1, n_j]\!]$,

$$f_{i_0 j_0}(e_q^{j_0}) = \sum_{p=1}^{m_{i_0}} a_{m_1+\cdots+m_{i_0-1}+p}^{n_1+\cdots+n_{j_0-1}+q} \cdot \varepsilon_p^{i_0}.$$

Puisque (a_k^l) est la matrice de f dans les bases $\mathcal{B}$ et $\mathcal{D}$, on a (en regardant la colonne correspondant à $e_q^{j_0}$) :

$$f(e_q^{j_0}) = \sum_{i=1}^{s} \sum_{p=1}^{m_i} a_{m_1+\cdots+m_{i-1}+p}^{n_1+\cdots+n_{j_0-1}+q} \cdot \varepsilon_p^i.$$

Pour tout $i = 1, \ldots, s$, le vecteur $\sum_{p=1}^{m_i} (a_{m_1+\cdots+m_{i-1}+p}^{n_1+\cdots+n_{j_0-1}+q} \cdot \varepsilon_p^i)$ appartient à F_i car c'est une combinaison linéaire de $\varepsilon_1^i, \ldots, \varepsilon_{m_i}^i$. On en déduit que les composantes du vecteur $f(e_q^{j_0})$ selon $F_1, \ldots, F_s$ respectivement sont

$$\sum_{p=1}^{m_1} a_p^{n_1+\cdots+n_{j_0-1}+q} \cdot \varepsilon_p^1, \quad \ldots \quad , \quad \sum_{p=1}^{m_s} a_{m_1+\cdots+m_{s-1}+p}^{n_1+\cdots+n_{j_0-1}+q} \cdot \varepsilon_p^s.$$

En particulier, la composante selon F_{i_0} est

$$f_{i_0 j_0}(e_q^{j_0}) = \sum_{p=1}^{m_{i_0}} a_{m_1+\cdots+m_{i_0-1}+p}^{n_1+\cdots+n_{j_0-1}+q} \cdot \varepsilon_p^{i_0}$$

comme il fallait démontrer. □

1.11 Composition/produit par blocs. On considère trois espaces vectoriels décomposés en sommes directes $E = E_1 \oplus \cdots \oplus E_r$, $F = F_1 \oplus \cdots \oplus F_s$, $G = G_1 \oplus \cdots \oplus G_t$, et deux applications linéaires $f \in \mathcal{L}(E,F)$ et $g \in \mathcal{L}(F,G)$ décomposées respectivement en f_{ij} et g_{ij} comme en (1.8.1). On a pour tout $j = 1, \ldots, r$ et $i = 1, \ldots, t$

$$(g \circ f)_{ij} = \sum\nolimits_{1 \leqslant k \leqslant s} g_{ik} \circ f_{kj}.$$

Le calcul à faire pour démontrer cette formule est exactement le même que celui qui établit la formule

$$\mathrm{Mat}(g \circ f) = \mathrm{Mat}(g) \times \mathrm{Mat}(f).$$

C'est ce qui explique les formules pour un produit de matrices par blocs, qui se résument en disant "qu'on peut prétendre que les blocs sont des scalaires". Attention juste au fait que les produits de blocs ne sont pas commutatifs.

1.3 – Qu'est-ce que la réduction des endomorphismes ?

1.12 But. On veut réduire l'étude d'un endomorphisme $f \in \mathcal{L}(E)$ à l'étude d'endomorphismes f_i définis sur des espaces vectoriels de dimension plus petite que E.

De la sorte, on peut espérer se ramener par récurrence sur la dimension à décrire tout endomorphisme en termes d'endomorphismes atomiques, c'est-à-dire d'endomorphismes qu'on ne peut plus casser en endomorphismes plus petits. On va voir cependant que la situation n'est pas tout-à-fait aussi simple (sans mauvais jeu de mot).

1.13 Applications. En se ramenant à des endomorphismes atomiques, on espère pouvoir se limiter à une courte liste de briques élémentaires permettant de reconstruire n'importe quel endomorphisme. On pourra alors analyser un endomorphisme arbitraire, par exemple en le décrivant par une matrice constituée de blocs bien compris. Une telle matrice est ce qu'on appelle une *forme normale*.

Un aspect de ce problème est la recherche d'une classification des endomorphismes "à équivalence près". La relation d'équivalence qui nous intéresse est la *similitude* : deux endomorphismes $f \in \mathcal{L}(E)$ et $f' \in \mathcal{L}(E')$ sont *semblables* s'il existe un isomorphisme $\varphi : E \simeq E'$ tel que $f' = \varphi \circ f \circ \varphi^{-1}$. Les classes d'équivalence pour la relation de similitude s'appellent les classes de similitude (c'est bien trouvé, non ?).

La recherche de formes normales consiste à donner un représentant emblématique de chaque classe de similitude. Ainsi deux endomorphismes seront semblables si et seulement si ils ont la même forme normale. Un exemple fameux est la forme normale de Jordan pour les endomorphismes trigonalisables.

On s'attachera aussi à la recherche d'*invariants de similitude*, c'est-à-dire d'objets associés à tout endomorphisme qui ne dépendent que de sa classe similitude. Des exemples bien connus sont les polynômes caractéristique et minimal. On souhaite connaître suffisamment d'invariants pour pouvoir distinguer les classes de similitude. Les polynômes caractéristique et minimal sont notoirement insuffisants pour cela, mais nous réaliserons complètement ce souhait d'avoir suffisamment d'invariants dans la section 3.6 (voir aussi le théorème 2.52 pour le cas particulier des endomorphismes trigonalisables).

1.14. Pour commencer la théorie de la réduction des endomorphismes, il faut en premier lieu formuler en termes utilisables (de manière équivalente, en termes précis) la condition « l'endomorphisme $f \in \mathcal{L}(E)$ est entièrement encodé par la donnée d'endomorphismes sur des sous-espaces stricts de E. » [2] Après un moment de réflexion, on arrive à la condition suivante : *il existe une décomposition* $E = \bigoplus_{i=1}^{r} E_i$ *en somme de sous-espaces stricts et des endomorphismes* $f_i \in \mathcal{L}(E_i)$ *pour tout* $i = 1, \ldots, r$, *tels que pour tout* $(x_1, \ldots, x_r) \in E_1 \times \cdots \times E_r$,

$$f(x_1 + \cdots + x_r) = f_1(x_1) + \cdots + f_r(x_r).$$

2. Un sous-espace vectoriel $F \subseteq E$ est dit *strict* si $\{0\} \subsetneq F \subsetneq E$.

Nous aurons besoin de la définition suivante pour reformuler cette condition.

1.15 Définition. Soit $f \in \mathcal{L}(E)$. Un sous-espace vectoriel F de E est *stable par* f si pour tout $x \in F$, on a $f(x) \in F$.

En ces termes, la condition du 1.14 s'exprime de la façon suivante : *il existe une décomposition* $E = \bigoplus_{i=1}^{r} E_i$ *en somme de sous-espaces stricts stables par* f.

? Penser à renvoyer au paragraphe où l'on voit que la donnée de f_F et de $\bar{f}_F$ ne suffit pas à déterminer f.

1.16. Considérons une décomposition $E = \bigoplus_{i=1}^{r} E_i$, et notons f_{ij}, $1 \leqslant i, j \leqslant r$, les composantes de f selon cette décomposition. Soit $i_0 \in [\![1, r]\!]$. Le sous-espace E_{i_0} est stable par f si et seulement si pour tout $i \neq i_0$, la composante $f_{ii_0} \in \mathcal{L}(E_{i_0}, E_i)$ est nulle (c'est tautologique).

Ainsi lorsque tous les E_i sont stables par f, les seules composantes non-nulles de f sont les f_{ii}, $1 \leqslant i \leqslant r$, et dans une base compatible à la décomposition $E = \bigoplus_{i=1}^{r} E_i$ la matrice de f est diagonale par blocs.

On dit que $f_{ii} \in \mathcal{L}(E_i)$ est l'endomorphisme induit par f sur E_i. Plus généralement, on pose la définition suivante.

1.17 Définition. Soit $f \in \mathcal{L}(E)$. Si F est un sous-espace de E stable par f, l'*endomorphisme induit* par f sur F est l'endomorphisme $f_F \in \mathcal{L}(F)$ défini par la condition

$$\forall x \in F : \quad f_F(x) = f(x).$$

Avant d'aller plus loin, voyons un lemme élémentaire qui explique comment la réduction des endomorphismes peut être appliquée à la classification à similitude près.

1.18 Lemme. *Soit $f, g \in \mathcal{L}(E)$. On suppose qu'il existe deux décompositions $E = \bigoplus_{i=1}^{r} F_i$ et $E = \bigoplus_{i=1}^{r} G_i$ en sommes de sous-espaces stables par f et g respectivement, telles que pour tout $i = 1, \ldots, r$, il existe un isomorphisme $\varphi_i : F_i \simeq G_i$ tel que $g_{G_i} = \varphi_i f_{F_i} \varphi_i^{-1}$. Alors les endomorphismes f et g sont semblables.*

Preuve. Il s'agit de construire à partir des φ_i un isomorphisme φ de E "diagonal par blocs" qui va conjuguer f et g. Précisément, on définit φ selon ses composantes $\varphi_{ij} \in \mathcal{L}(F_i, G_j)$ relativement aux décompositions $E = \bigoplus_{i=1}^{r} F_i$ et $E = \bigoplus_{i=1}^{r} G_i$, grâce à la Proposition 1.3 : on considère l'unique $\varphi \in \mathcal{L}(E)$ tel que pour tout $i, j \in [\![1, r]\!]$,

$$\varphi_{ij} = \begin{cases} \varphi_i & \text{si } i = j\,; \\ 0 & \text{sinon.} \end{cases}$$

On vérifie alors sans mal que φ est un isomorphisme puisque $\varphi_1, \ldots, \varphi_r$ sont des isomorphismes, et $g = \varphi f \varphi^{-1}$. □

La condition pour qu'un endomorphisme soit atomique découle naturellement de la condition donnée en 1.14. La terminologie classique veut qu'on appelle ces endomorphismes simples plutôt qu'atomiques.

1.19 Définition. Soit $f \in \mathcal{L}(E)$. L'endomorphisme f est *simple* si aucun sous-espace strict de E n'est stable par f.

1.20 Exemples. Si E est de dimension 1, tout endomorphisme de E est simple.

Si $E = \mathbf{R}^2$, l'endomorphisme r_θ défini dans la base canonique par la matrice

$$R_\theta = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

est simple si $\theta \not\equiv 0 \bmod \pi$. Géométriquement c'est clair, puisque cet endomorphisme est une rotation d'angle θ ; cependant dans le contexte où nous nous trouvons il n'y a pas de produit scalaire, et donc pas d'angles. Voici une justification plus algébrique (qui anticipe un peu sur les notions que nous verrons plus tard) : puisque E est de dimension 2, tout sous-espace stable strict est de dimension 1, donc contenu dans un sous-espace propre. Or le polynôme caractéristique de R_θ est

$$X^2 - (2\cos\theta)X + 1 = (X - e^{i\theta})(X + e^{-i\theta}),$$

donc R_θ n'a aucune valeur propre réelle. Ainsi r_θ ne peut pas avoir de sous-espace stable strict.

Si $E = \mathbf{C}^2$, l'endomorphisme $\tilde{r}_\theta$ défini dans la base canonique par la matrice R_θ n'est pas simple : les deux sous-espaces propres relatifs aux valeurs propres $e^{i\theta}$ et $e^{-i\theta}$ sont tous les deux des droites de $\mathbf{C}^2$ stables par $\tilde{r}_\theta$.

Comme annoncé plus haut, il est illusoire d'espérer pouvoir réduire en toute généralité un endomorphisme en somme d'endomorphismes simples. L'exemple suivant illustre bien les problèmes qui peuvent se poser.

1.21 Exemple. Soit $E = \mathbf{k}^2$ et f l'endomorphisme donné dans la base canonique (e_1, e_2) par la matrice

$$N = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

Le seul sous-espace strict de E stable par f est la droite $\mathrm{Vect}(e_1)$. En effet, par le même argument qu'en 1.20 ci-dessus, tout sous-espace stable par f est contenu dans un sous-espace propre de f. Or son polynôme caractéristique est X^2, donc 0 est son unique valeur propre, et son seul sous-espace propre est $\ker(f) = \mathrm{Vect}(e_1)$.

Il est donc impossible de décomposer E en somme directe de sous-espaces stricts stables par f.

On donne un nom aux endomorphismes qu'il est possible de décomposer en sommes d'endomorphismes simples.

1.22 Définition. Un endomorphisme $f \in \mathcal{L}(E)$ est *semi-simple* s'il est somme directe d'endomorphismes simples, c'est-à-dire s'il existe une décomposition $E = \bigoplus_{i=1}^{r} E_i$ en somme de sous-espaces stables par f telle que pour tout $i = 1, \ldots, r$, l'endomorphisme induit $f_{E_i} \in \mathcal{L}(E_i)$ est simple.

L'endomorphisme de l'Exemple 1.21 n'est ni simple, ni semi-simple. L'endomorphisme r_θ de 1.20 est semi-simple. On verra plus loin que tout endomorphisme diagonalisable est semi-simple. Dans une large mesure, les endomorphismes semi-simples sont ceux qu'il est possible de diagonaliser quitte à étendre les scalaires. Dans la section 3.3 nous donnons des conditions nécessaires et suffisantes pour la simplicité et la semi-simplicité portant sur les polynômes minimaux et caractéristiques.

Dans la littérature on trouve souvent une autre définition de la semi-simplicité, qui est la suivante. Nous allons voir que les deux versions sont équivalentes, même si ce n'est pas si évident à première vue.

1.23 Définition. Un endomorphisme $f \in \mathcal{L}(E)$ est *b-semi-simple* si pour tout sous-espace F stable par f, il existe un supplémentaire F' de F dans E qui est lui aussi stable par f.

1.24 Théorème. *Les conditions de semi-simplicité et de b-semi-simplicité sont équivalentes.*

Le reste de cette section est consacré à la preuve de ce théorème, et à divers compléments. Nous donnons ici une preuve élémentaire, mais il est aussi possible d'utiliser des outils plus élaborés de réduction des endomorphismes, voir la section 3.3.

Preuve du théorème 1.24. L'implication « semi-simple ⇒ b-semi-simple » est le contenu de la proposition 1.27, et sa réciproque « b-semi-simple ⇒ semi-simple » est le contenu de la proposition 1.25. □

1.25 Proposition. *Soit $f \in \mathcal{L}(E)$ un endomorphisme b-semi-simple. Pour tout sous-espace F stable par E, l'endomorphisme induit $f_F \in \mathcal{L}(F)$ est lui aussi b-semi-simple.*

Preuve. Soit $F_0 \subseteq F$ un sous-espace stable par f_F, ou de manière équivalente stable par f. Par semi-simplicité de f, il existe donc $\tilde{F}_0$ supplémentaire de F_0 dans E stable par f. Nous allons montrer que le sous-espace $F_0' = \tilde{F}_0 \cap F$ est un supplémentaire de F_0 dans F, stable par f_F, ce qui conclura la preuve.

La stabilité de F_0' par f est claire, et puisque $F_0' \subseteq F$, elle est équivalente à la stabilité par f_F. On a $F_0 \cap F_0' \subseteq F_0 \cap \tilde{F}_0 = \{0\}$, donc F_0' et F_0 sont en somme directe. Il reste à montrer que $F_0 \oplus F_0' = F$. Puisque F_0 et F_0' sont tous les deux contenus dans F, on a l'inclusion $F_0 \oplus F_0' \subseteq F$. Pour montrer l'inclusion inverse, considérons $x \in F$ et décomposons le selon $F_0 \oplus \tilde{F}_0 = E$: on écrit $x = x_0 + \tilde{x}_0$, avec $x_0 \in F_0$ et $\tilde{x}_0 \in \tilde{F}_0$. Alors $\tilde{x}_0 = x - x_0 \in F$, donc $\tilde{x}_0 \in \tilde{F}_0 \cap F = F_0'$. Ceci prouve que $x \in F_0 + F_0'$, et la preuve est terminée. □

1.26 Corollaire. *Soit $f \in \mathcal{L}(E)$ b-semi-simple. Alors f est semi-simple.*

Preuve du corollaire. On raisonne par récurrence sur la dimension n de E. Si $n \leqslant 1$, f est lui-même simple et il n'y a rien à démontrer. Supposons donc $n > 1$ et le résultat démontré pour $\dim(E) < n$. Si f est simple, le résultat est trivial. Sinon, il existe F sous-espace strict de E stable par f. Puisque f est semi-simple, il existe F' supplémentaire de F dans E lui aussi stable par f. Puisque F est strict, on a $\dim(F) < n$ et $\dim(F') < n$. D'après la proposition 1.25, les endomorphismes induits f_F et $f_{F'}$ sont tous les deux semi-simples. On peut donc appliquer l'hypothèse de récurrence, qui nous dit qu'il existe deux décompositions $F = \bigoplus_i F_i$ et $F' = \bigoplus_i F_i'$ en sommes de sous-espaces stables par f_F et $f_{F'}$ respectivement, et donc stables par f, tels que les $(f_F)_{F_i} = f_{F_i}$ et $(f_{F'})_{F_i'} = f_{F_i'}$ sont simples. On a donc une décomposition $E = \bigoplus_i F_i \oplus \bigoplus_i F_i'$ comme on voulait. □

1.27 Proposition. *Soit $f \in \mathcal{L}(E)$. Si f est semi-simple, alors il est aussi b-semi-simple.*

Preuve. Supposons qu'il existe une décomposition $E = \bigoplus_{i=1}^r E_i$ telle que pour tout $i \in [\![1, r]\!]$, E_i est stable par f et l'endomorphisme induit f_{E_i} est simple, et montrons qu'alors f est semi-simple. Soit F un sous-espace stable par f. Il s'agit de trouver un supplémentaire à F lui aussi stable par f. Considérons

$$\mathcal{I} = \big\{ I \subseteq [\![1, r]\!] \text{ t.q. } F \cap (\textstyle\bigoplus_{i \in I} E_i) = \{0\} \big\}.$$

L'ensemble $\mathcal{I}$ est fini, donc il possède des éléments maximaux pour l'inclusion. Soit I_0 un élément maximal de $\mathcal{I}$. Le sous-espace $E_{I_0} = \bigoplus_{i \in I_0} E_i$ est stable par f, et nous allons montrer que c'est un supplémentaire de F dans E, ce qui concluera la preuve.

Les sous-espaces F et E_{I_0} sont en somme directe puisque $I_0 \in \mathcal{I}$, donc il suffit de prouver que $F \oplus E_{I_0} = E$. Puisque $E = \bigoplus_{i=1}^r E_i$, il suffit de prouver que E_{i_1} est contenu dans $F \oplus E_{I_0}$ pour tout $i_1 \in [\![1, r]\!]$. Si $i_1 \in I_0$, c'est clair. Sinon, $I_0 \cup \{i_1\}$ contient strictement I_0, donc par maximalité de I_0, l'ensemble $I_0 \cup \{i_1\}$ n'appartient pas à $\mathcal{I}$, et donc $F \cap (E_{I_0} \oplus E_{i_1}) \neq \{0\}$. Ainsi il existe $y \in F$ non nul, $x_0 \in E_{I_0}$ et $x_1 \in E_{i_1}$ tels que $y = x_0 + x_1$. Puisque $F \cap E_{I_0} = \{0\}$, x_1 est non nul, et puisque $x_1 = y - x_0$, $x_1 \in E_{i_1} \cap (F \oplus E_{I_0})$. Ainsi, $E_{i_1} \cap (F \oplus E_{I_0}) \neq \{0\}$. Mais $E_{i_1} \cap (F \oplus E_{I_0})$ est un sous-espace de E_{i_1} stable par f, donc c'est un sous-espace stable par $f_{E_{i_1}}$. Puisque $f_{E_{i_1}}$ est simple, on en déduit que $E_{i_1} \cap (F \oplus E_{I_0}) = E_{i_1}$, et donc E_{i_1} est tout entier contenu dans $F \oplus E_{I_0}$ comme on voulait démontrer. □

Pour conclure cette section, nous allons donner une preuve différente de l'implication « semi-simple ⇒ b-semi-simple » qui sera l'occasion de voir quelques lemmes utiles et instructifs.

1.28 Lemme. *On considère un endomorphisme $f \in \mathcal{L}(E)$.*
1.28.1. *Soit $E = F \oplus F'$ une décomposition telle que F et F' sont tous les deux stables par f. Alors les deux projecteurs $p, p' \in \mathcal{L}(E)$ sur F et F' respectivement relativement à cette décomposition commutent à f.*
1.28.2. *Soit $p \in \mathcal{L}(E)$ un projecteur commutant à f. Pour tout sous-espace F stable par f, le sous-espace $p(F)$ est stable par f.*

Preuve. Commençons par prouver 1.28.1. Par symétrie il suffit de démontrer le résultat pour p. Considérons un vecteur $x \in E$, et décomposons le en $x = x_F + x'_F$ avec $x_F \in F$, $x'_F \in F'$. On a

$$f(x) = f(x_F) + f(x'_F)$$

par linéarité, et $f(x_F) \in F$, $f(x'_F) \in F'$ par stabilité de F et F' par f. Donc

$$p(f(x)) = f(x_F) = f(p(x))$$

comme il fallait démontrer.

Venons en maintenant à 1.28.2. Soit F stable par f. Soit $x \in p(F)$, et montrons que $f(x) \in p(F)$. Il existe $y \in F$ tel que $x = p(y)$, et alors

$$f(x) = f(p(y)) = p(f(y))$$

appartient à $p(F)$ puisque $f(y) \in F$ par stabilité de F. □

1.29 Lemme. *Soit E un espace vectoriel muni d'une décomposition $E = F \oplus F'$, et considérons $p \in \mathcal{L}(E)$ le projecteur sur F dans la direction de F'. Pour tout sous-espace $L \subseteq E$, on a*

$$p(L) = (L + F') \cap F.$$

Preuve. Soit $x \in p(L)$. Il existe $y \in L$ tel que $x = p(y)$, et donc $y = x + x'$ pour un certain $x' \in F'$. Alors $x = y - x' \in L + F'$. D'autre part $x \in F$ puisque p est un projecteur sur F. Ceci prouve l'inclusion $p(L) \subseteq (L + F') \cap F$.

Pour montrer l'inclusion inverse, considérons $x \in (L + F') \cap F$. Il existe $y \in L$ et $x' \in F'$ tels que $x = y + x'$. Alors

$$x = p(x) = p(y + x') = p(y),$$

donc $x \in p(L)$ comme il fallait démontrer. □

Preuve alternative de la proposition 1.27. Soit $f \in L(E)$. Montrons par récurrence sur $r \in \mathbf{N}^*$ que s'il existe une décomposition $E = \bigoplus_{i=1}^{r} E_i$ telle que pour tout $i \in [\![1, r]\!]$, E_i est stable par f et l'endomorphisme induit f_{E_i} est simple, alors f est b-semi-simple. Si $r = 1$ le résultat est vrai, puisque alors f est simple et donc b-semi-simple. Supposons donc $r \geqslant 2$, et le résultat démontré pour un nombre $r' < r$ de facteurs. Posons $E_0 = \bigoplus_{i<r} E_i$; l'hypothèse de récurrence assure que l'endomorphisme $f_{E_0} \in \mathcal{L}(E_0)$ induit par f est b-semi-simple.

Soit F un sous-espace de E stable par f. L'intersection $F_0 = F \cap E_0$ est stable par f_{E_0}, donc il existe F'_0 supplémentaire de F_0 dans E_0 stable par f_{E_0}, et donc par f.

Montrons pour commencer que

$$(\star) \qquad F'_0 \oplus F = F + E_0.$$

Puisque $F'_0 \subseteq E_0$, on a $F'_0 \cap F \subseteq F'_0 \cap F_0 = \{0\}$, donc la somme est bien directe. L'inclusion $F'_0 \oplus F \subseteq F + E_0$ est conséquence des inclusions $F'_0 \subseteq E_0$ et $F \subseteq F$. Montrons l'inclusion inverse $F + E_0 \subseteq F'_0 \oplus F$. Puisque $F \subseteq F'_0 \oplus F$, il suffit de vérifier que $E_0 \subseteq F'_0 \oplus F$, qui est bien vraie puisque $E_0 = F'_0 \oplus F_0$ et $F_0 = F \cap E_0 \subseteq F$. Ceci conclut la preuve de $(\star)$.

Le sous-espace $F + E_0$ est stable par f, donc $(F + E_0) \cap E_r$ aussi, et puisque f_{E_r} est simple, on a ou bien $(F + E_0) \cap E_r = \{0\}$ ou bien $(F + E_0) \cap E_r = E_r$. Notons que $(F + E_0) \cap E_r$ est la projection de F sur E_r dans la direction de E_0, d'après le lemme 1.29.

a) Si $(F + E_0) \cap E_r = E_r$, alors $F + E_0 = E$. En effet, $F + E_0$ contient alors à la fois E_0 et E_r, qui sont supplémentaires dans E. Dans ce cas, F'_0 est un supplémentaire de F dans E stable par f, et la preuve est terminée.

b) Si $(F + E_0) \cap E_r = \{0\}$, alors $F + E_0 = E_0$. En effet, F est alors contenu dans E_0 (puisque sa projection sur E_r est nulle). Dans ce cas, $F'_0 \oplus F = E_0$, et $F'_0 \oplus E_r$ est un supplémentaire de F dans E stable par f, ce qui conclut la preuve. □

1.30 Exercice. Soit $f \in \mathcal{L}(E)$.
1) Montrer que si $F \subseteq E$ est un sous-espace stable de dimension 1, alors il existe $\lambda_F \in \mathbf{k}$ tel que

$$\forall x \in F : \ f(x) = \lambda_F.x.$$

En déduire que F est contenu dans un sous-espace propre de f.
2) Montrer que si f laisse toutes les droites de E stables, alors f est une homothétie.
(*Indication :* étant donné deux droites F et F', pour montrer que $\lambda_F = \lambda_{F'}$, considérer une droite $\mathrm{Vect}(x + x')$ avec $x \in F$ et $x' \in F'$.)

1.4 – Quotients

Il sera commode d'utiliser la notion de quotient pour raisonner par récurrence dans le cadre de la réduction des endomorphismes. Ici nous donnons quelques rappels autour de cette notion.

1.31 Quotient par un sous-espace vectoriel. Soit E un $\mathbf{k}$-espace vectoriel et F un sous-espace vectoriel de E. $(E, +)$ est un groupe abélien, et $(F, +)$ est un sous-groupe (distingué, forcément) de $(E, +)$. On a donc un groupe quotient $(E/F, +)$. L'opération de multiplication par les scalaires définie par

$$\forall \bar{x} \in E, \forall \lambda \in \mathbf{k} : \ \lambda.\bar{x} = \overline{\lambda.x}$$

munit E/F d'une structure de $\mathbf{k}$-espace vectoriel canoniquement induite par celle de E. (On laisse toutes les vérifications élémentaires mais nécessaires au lecteur).

L'application $x \in E \mapsto \overline{x} \in E/F$ est une application linéaire surjective, dont le noyau est F. On l'appelle la *projection canonique*.

1.32 Lemme. *Soit F un sous-espace vectoriel de E, et considérons une base $\mathcal{B} = (e_1, \ldots, e_p, \ldots, e_n)$ telle que $(e_1, \ldots, e_p)$ est une base de F. Alors $(\bar{e}_{p+1}, \ldots, \bar{e}_n)$ est une base de E/F.*

Une base de E satisfaisant à l'hypothèse du lemme ci-dessus sera dite *compatible à* F. Étant donné une base compatible à F, les bases $(e_1, \ldots, e_p)$ de F et $(\bar{e}_{p+1}, \ldots, \bar{e}_n)$ de E/F comme ci-dessus seront dites *induites* par $\mathcal{B}$.

Preuve. Montrons que la famille $(\bar{e}_{p+1}, \ldots, \bar{e}_n)$ est libre. Soit $\lambda_{p+1}, \ldots, \lambda_n \in \mathbf{k}$ tels que $\lambda_{p+1}.\bar{e}_{p+1} + \cdots + \lambda_n.\bar{e}_n = 0$. Alors

$$\overline{\lambda_{p+1}.e_{p+1} + \cdots + \lambda_n.e_n} = \lambda_{p+1}.\overline{e_{p+1}} + \cdots + \lambda_n.\overline{e_n} = 0,$$

donc $\lambda_{p+1}.e_{p+1} + \cdots + \lambda_n.e_n \in F$. Puisque $(e_1, \ldots, e_p)$ est une base de F, il existe donc des scalaires $\mu_1, \ldots, \mu_p \in \mathbf{k}$ tels que

$$\begin{aligned} \lambda_{p+1}.e_{p+1} + \cdots + \lambda_n.e_n = \mu_1.e_1 + \cdots + \mu_p.e_p \\ \iff -\mu_1.e_1 - \cdots - \mu_p.e_p + \lambda_{p+1}.e_{p+1} + \cdots + \lambda_n.e_n = 0. \end{aligned}$$

Par liberté de la famille $(e_1, \ldots, e_p, \ldots, e_n)$, l'égalité de droite implique $\mu_1 = \cdots = \mu_p = \lambda_{p+1} = \cdots = \lambda_n = 0$. On a ainsi bien montré que $(\bar{e}_{p+1}, \ldots, \bar{e}_n)$ est libre.

Montrons que la famille $(\bar{e}_{p+1}, \ldots, \bar{e}_n)$ engendre E/F. Soit $\xi \in E/F$. On choisit un représentant $x \in E$ de ξ, que l'on écrit dans la base $\mathcal{B}$, $x = a_1.e_1 + \cdots + a_n.e_n$. Passant aux classes modulo F :

$$\xi = \overline{x} = a_1.\bar{e}_1 + \cdots + a_p.\bar{e}_p + a_{p+1}.\bar{e}_{p+1} + \cdots + a_n.\bar{e}_n = a_{p+1}.\bar{e}_{p+1} + \cdots + a_n.\bar{e}_n,$$

où la dernière égalité vient du fait que $\bar{e}_1 = \cdots = \bar{e}_p = 0$ puisque $e_1, \ldots, e_p \in F$. Ainsi tout $\xi \in E/F$ est combinaison linéaire de $\bar{e}_{p+1}, \ldots, \bar{e}_n$, et on a bien démontré que la famille $(\bar{e}_{p+1}, \ldots, \bar{e}_n)$ engendre E/F. □

1.33 Corollaire. *Le quotient E/F est de dimension finie, et*

$$\dim(E/F) = \dim E - \dim F.$$

Le cas $p = 2$ du résultat suivant est en quelque sort le lemme 1.32 "dans l'autre sens". Il sera commode d'avoir la version pour p arbitraire toute prête. Une suite de sous-espaces vectoriels $F_0, F_1, \ldots, F_p$ comme dans l'énoncé ci-dessous s'appelle une *filtration* de E.

1.34 Lemme. *Considérons une suite de sous-espaces vectoriels*

$$\{0\} = F_0 \subseteq F_1 \subseteq \cdots \subseteq F_p = E \ ;$$

on appelle une telle suite On considère des vecteurs

$$\begin{gathered} \varepsilon_{1,1}, \ldots, \varepsilon_{1,r_1} \in F_1 \\ \vdots \\ \varepsilon_{p,1}, \ldots, \varepsilon_{p,r_p} \in F_p \end{gathered}$$

tels que pour tout $i = 1, \ldots, p$ les classes $\bar{\varepsilon}_{1,1}, \ldots, \bar{\varepsilon}_{i,r_i}$ modulo F_{i-1} constituent une base de F_i/F_{i-1}. Alors la famille (ε_{ij}) est une base de E.

Preuve. Par récurrence sur p, il suffit de savoir le faire pour $p = 2$. Soit F un sous-espace vectoriel de E, $(a_1, \ldots, a_s)$ base de F, et $b_1, \ldots, b_r \in E$ tels que $(\bar{b}_1, \ldots, \bar{b}_r)$ soit une base du quotient E/F. Montrons que $(a_1, \ldots, a_s, b_1, \ldots, b_r)$ est une base de E.

Soit $\lambda_1, \ldots, \lambda_s, \mu_1, \ldots, \mu_r$ tels que

$$\lambda_1.a_1 + \cdots + \lambda_p.a_s + \mu_1.b_1 + \cdots + \mu_r.b_r = 0. \tag{1.34.1}$$

Après projection dans E/F, reste

$$\mu_1.\bar{b}_1 + \cdots + \mu_r.\bar{b}_r = 0,$$

qui implique $\mu_1 = \cdots = \mu_r = 0$. (1.34.1) devient alors une combinaison linéaire nulle de $e_1, \ldots, e_p$, qui à son tour donne $\lambda_1 = \cdots = \lambda_s = 0$. Ceci prouve la liberté.

D'autre part, soit $x \in E$. Sa projection $\bar{x}$ dans le quotient E/F s'écrit comme une combinaison linéaire

$$\bar{x} = \mu_1.\bar{b}_1 + \cdots + \mu_r.\bar{b}_r.$$

Ainsi $x - (\mu_1.b_1 + \cdots + \mu_r.b_r)$ est dans F, puisque sa classe modulo F est nulle, en conséquence de quoi il s'écrit comme combinaison linéaire de $a_1, \ldots, a_s$. □

Quotients et supplémentaires

À présent, nous voulons expliquer les relations entre le quotient E/F et les supplémentaires de F dans E.

1.35 Lemme. *On considère F sous-espace vectoriel de E, et F' un supplémentaire de F dans E.*
(a) *La projection canonique $\pi : E \to E/F$ induit par restriction un isomorphisme $F' \xrightarrow{\cong} E/F$.*
(b) *L'application composante selon F' dans la décomposition $E = F \oplus F'$, $\tilde{p} : E \to F'$, induit par passage au quotient un isomorphisme $E/F \xrightarrow{\cong} F'$.*

Preuve. (a) La projection canonique $\pi : E \to E/F$ est surjective, de noyau F. Le noyau de la restriction $\pi|_{F'}$ est donc $F' \cap F = \{0\}$, et ainsi $\pi|_{F'}$ est injective. Montrons qu'elle est également surjective. Soit $\xi \in E/F$. On choisit $x \in E$ représentant ξ, et on l'écrit $x = x_F + x'_F$, avec $x_F \in F$ et $x'_F \in F'$. Alors

$$\xi = \overline{x} = \pi(x) = \pi(x_F + x'_F) = \pi(x'_F),$$

la dernière égalité provenant de la linéarité de π et du fait que $\pi(x_F) = 0$. Ainsi $\pi|_{F'}$ est surjective et inective, et nous avons démontré que c'est un isomorphisme $F' \xrightarrow{\cong} E/F$. □

1.36 Quotient et supplémentaire. Choisissons un supplémentaire F' de F dans E. On dispose de la projection $\tilde{p} : E \to F'$ dans la direction de F, qui est surjective et de noyau F. D'après la propriété universelle du quotient, il existe un isomorphisme canonique entre F' et E/F. Cette réalisation du quotient dépend d'un choix et n'est donc pas canonique.

Le choix d'un supplémentaire de F dans E équivaut au choix d'une injection linéaire $j : E/F \hookrightarrow E$ telle que $\mathrm{im}(j) \cap F = \{0\}$.

Il est important de noter qu'il n'existe pas de supplémentaire canonique à F (sauf à adjoindre des structures supplémentaires, par exemple une structure euclidienne). Cependant le bloc B de (1.38.1) a un sens canonique : il ne dépend que du F, et pas du choix du supplémentaire F'. Pour le voir, nous allons considérer le quotient E/F.

Propriété universelle du quotient

Ce qui suit n'est pas strictement nécessaire pour la lecture des sections sur la réduction des endomorphismes. C'est cependant important du point de vue théorique, et nous l'utiliserons à l'occasion. Il n'est pas nécessaire ici de supposer E de dimension finie.

1.37 Proposition. *Soit E un espace vectoriel, et F un sous-espace vectoriel de E. Le quotient E/F jouit des deux propriétés suivantes :*
(a) il existe une application linéaire $\pi : E \to E/F$ surjective et de noyau F ;
(b) pour tout espace vectoriel G, l'application linéaire (c'en est une)

$$u \in \mathcal{L}(E/F, G) \mapsto u \circ \pi \in \mathcal{L}(E, G)$$

induit un isomorphisme (fonctoriel...)

$$\Phi : \mathcal{L}(E/F, G) \cong \ker\big(\mathrm{restr} : \mathcal{L}(E, G) \to \mathcal{L}(F, G)\big).$$

S'il existe E' jouissant lui aussi de l'une ou l'autre de ces deux propriétés, alors il existe un unique isomorphisme $\varphi : E/F \cong E'$ tel que $\pi' = \varphi \circ \pi$.

1.37.1 Preuves. On laisse au lecteur la preuve de la propriété (a) pour E/F et sa projection canonique. On va montrer que les propriétés (a) et (b) sont équivalentes, puis qu'un $\mathbf{k}$-ev satisfaisant à ces propriétés s'identifie canoniquement à E/F.

Soit E' vérifiant (a) (i.e. on suppose E' $\mathbf{k}$-ev muni de $\pi' : E \to E'$ linéaire, surjective et de noyau F), et montrons que (b) vaut pour E' (cela montrera au passage que (b) vaut bien pour le quotient). On commence par constater que $u \mapsto u \circ \pi'$ donne bien une application linéaire

$$\Phi_{\pi'} : \mathcal{L}(E', G) \to \ker\big(\mathrm{restr} : \mathcal{L}(E, G) \to \mathcal{L}(F, G)\big)$$

pour tout $\mathbf{k}$-ev G. Reste à voir que c'est effectivement un isomorphisme. Si $u \circ \pi' = 0$, alors $u = 0$ car π' est surjective, d'où l'injectivité de $\Phi_{\pi'}$.

Pour la surjectivité, on construit à la main un antécédent pour toute application linéaire $v : E \to G$ s'annulant sur F. Pour tout $y \in E'$ il existe $x \in E$ tel que $y = \pi'(x)$, et on pose $u(y) := v(x)$; ça ne dépend pas du choix de x car $\ker(\pi') = F$ et $v|_F = 0$. Ceci définit une application linéaire $u : E' \to G$ qui visiblement factorise comme il faut (i.e. $v = u \circ \pi'$). □

Réciproquement, soit E' vérifiant (b), et montrons directement que (a) vaut pour E' (à nouveau, cela montrera au passage que le quotient catégoriel est bien le quotient défini par la relation de congruence). Précisément, on suppose qu'il existe une application linéaire $\pi' : E \to E'$ tel que $\Phi_{\pi'}$ soit un isomorphisme pour tout G, et il s'agit de montrer que π' est surjective et de noyau F.

On regarde le $\Phi_{\pi'}$ pour $G = E'$: puisque $\pi' = \Phi_{\pi'}(\mathrm{id}_{E'})$, π' est dans le noyau de la restriction à F, donc $F \subseteq \ker(\pi')$. D'autre part, on regarde le $\Phi_{\pi'}$ pour $G = E/F$: puisque π est dans

le noyau de la restriction à F, il existe un $u : E' \to E/F$ tel que $\pi = u \circ \pi'$. Ceci implique $\ker(\pi') \subseteq \ker(\pi) = F$, donc on conclut que $\ker(\pi') = F$.

D'autre part, si π' n'était pas surjective on pourrait contredire l'injectivité de $\Phi_{\pi'}$ pour n'importe quel $G \neq \{0\}$ de la manière suivante. Soit $x \in E'$ non atteint par π'. On considère G un $\mathbf{k}$-ev non nul et $u' : E' \to G$. Alors pour g n'importe quel vecteur non nul de G et ℓ une forme linéaire de noyau un hyperplan transverse à x, on a

$$(u' + \ell.g) \circ \pi' = u' \circ \pi',$$

ce qui contredit l'injectivité de $\Phi_{\pi'}$. □

Soit E' vérifiant (a). Alors $\pi' \in \mathcal{L}(E, E')$ est dans le noyau de la restriction à F, donc il existe un unique $\varphi : E/F \to E'$ tel que $\pi' = \varphi \circ \pi$. Reste à voir que ce φ est un isomorphisme. Puisque π est surjective, $\ker \varphi \neq \{0\} \Rightarrow \ker \pi' \supsetneq F$ et donc nécessairement φ est injective. D'autre part φ est surjective car π' l'est. □

1.5 – Sous-espaces stables et matrices triangulaires par blocs

On a vu en 1.16 que l'existence d'une décomposition de l'espace ambient E en somme directe de sous-espaces stables par un endomorphisme $f \in \mathcal{L}(E)$ se traduit matriciellement par l'existence d'une base de E dans laquelle la matrice de f est diagonale par blocs. Nous aurons besoin de l'énoncé un peu plus robuste ci-dessous, qui se démontre de la même façon (nous laissons donc la preuve comme exercice au lecteur).

1.38 Proposition. *Soit $f \in \mathcal{L}(E)$ et F un sous-espace de E stable par f. Dans une base $\mathcal{B}$ de E compatible à F, la matrice $\mathrm{Mat}_{\mathcal{B}}(f)$ est triangulaire supérieure par blocs*

$$\begin{pmatrix} A & C \\ 0 & B \end{pmatrix} \tag{1.38.1}$$

avec $A \in \mathcal{M}_p(\mathbf{k})$ et $B \in \mathcal{M}_{n-p}(\mathbf{k})$ carréees (en notant $n = \dim(E)$ et $p = \dim(F)$).

Réciproquement, si dans une base $\mathcal{B} = (e_1, \ldots, e_n)$ la matrice de f est triangulaire supérieure par blocs comme ci-dessus, alors le sous-espace $\mathrm{Vect}(e_1, \ldots, e_p)$ est stable par f.

Nous allons voir que les classes de conjugaison des blocs diagonaux A et B comme en (1.38.1) sont canoniquement attachées à la donnée de f et du sous-espace stable F (Corollaire 1.44). En revanche le bloc C dépend fortement du choix de la base $\mathcal{B}$, et n'encode en général aucune information intrinsèque à la paire (f, F), comme le démontre l'exemple 1.39 ci-dessous.

1.39 Exemple. Soit f l'endomorphisme de $\mathbf{k}^2$ donné dans la base canonique (e_1, e_2) par la matrice

$$\begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix}.$$

On considère le sous-espace stable $F = \mathrm{Vect}(e_1)$; la base canonique est compatible à F. La base $(e_1, e_1 + e_2)$ est elle aussi compatible à f, et dans cette base la matrice de f est

$$\begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}.$$

L'information géométrique contenue dans les blocs A et B s'exprime en fonction des objets suivants.

1.40 Endomorphismes induits. Soit $f \in \mathcal{L}(E)$ et F un sous-espace de E stable par f. La condition

$$\forall x \in F : f_F(x) = f(x).$$

définit un endomorphisme $f_F \in \mathcal{L}(F)$, dit *endomorphisme induit* par f sur F (déjà introduit en 1.17 ci-dessus).

D'autre part, la condition

$$\forall \bar{x} \in E/F : \bar{f}_F(\bar{x}) = \overline{f(x)}$$

définit un endomorphisme $\bar{f}_F \in \mathcal{L}(E/F)$, dit *endomorphisme sur le quotient induit* par f.

Il convient de vérifier que $\bar{f}_F$ est bien défini par la condition ci-dessus, ce qui revient à vérifier que $\overline{f(x)}$ ne dépend pas du choix du représentant de la classe $\overline{x}$. Soit $x' \in E$ un autre représentant de $\bar{x}$. On a $x' - x \in F$, et donc puisque F est stable et f est linéaire, $f(x') - f(x) \in F$. Ainsi $\overline{f(x')} = \overline{f(x)}$ comme il fallait démontrer.

1.40.1 Mise en garde. Nous recommandons d'être attentif au fait que la donnée de f_F et $\bar{f}_F$ est insuffisante pour reconstruire f, voir l'exemple 1.39 ci-dessus.

1.41 Proposition. *Soit $f \in \mathcal{L}(E)$, F stable par f, et $\mathcal{B}$ une base de E compatible à F. Alors les blocs A et B de la matrice* $\mathrm{Mat}_{\mathcal{B}}(f)$ *écrite comme en* (1.38.1) *représentent f_F et $\bar{f}_F$ respectivement, dans les bases de F et de E/F induites par $\mathcal{B}$.*

Avant d'attaquer la preuve, précisons ce que nous appelons les bases de F et de E/F induites par $\mathcal{B}$. Puisque $\mathcal{B}$ est compatible à F, elle est la concaténation d'une base $\mathcal{B}_F$ de F et d'une famille libre $\mathcal{B}'$; d'après le lemme 1.32, les classes modulo F des vecteurs de $\mathcal{B}'$ forment une base $\bar{\mathcal{B}}_F$ de E/F. Les bases induites dont il est question dans l'énoncé ci-dessus sont $\mathcal{B}_F$ et $\bar{\mathcal{B}}_F$.

Preuve. Le fait que $A = \mathrm{Mat}_{\mathcal{B}_F}(f_F)$ est contenu dans la proposition 1.10. Pour montrer que $B = \mathrm{Mat}_{\bar{\mathcal{B}}_F}(\bar{f}_F)$, notons $\mathcal{B}_F = (e_1, \ldots, e_p)$ et $\mathcal{B}' = (e_{p+1}, \ldots, e_n)$. Soit $j \in [\![1, n-p]\!]$. Notons en outre $C = (c_{ij})_{\substack{1 \leqslant i \leqslant p \\ 1 \leqslant j \leqslant n-p}}$ et $B = (b_{ij})_{1 \leqslant i,j \leqslant n-p}$ les blocs de (1.38.1). On a

$$\begin{aligned} \bar{f}_F(\bar{e}_{p+j}) &= \overline{f(e_{p+j})} \\ (*) \qquad &= \underbrace{\overline{\sum_{i=1}^{p} c_{ij}.e_i}}_{=\bar{0}} + \overline{\sum_{i=p+1}^{n} b_{i-p,j}.e_i} \\ &= \sum_{i=1}^{n-p} b_{ij}.\bar{e}_{p+i}, \end{aligned}$$

ce qui prouve bien que $B = \mathrm{Mat}_{\bar{\mathcal{B}}_F}(\bar{f}_F)$. (* La première somme est nulle car c'est une combinaison linéaire de $\overline{e_1}, \ldots, \overline{e_p}$ qui sont tous nuls, puisque $e_1, \ldots, e_p \in F$). □

1.42 Remarque. Il est instructif de suivre les choix faits lors de l'écriture de la matrice (1.38.1). Au commencement on a F stable par f. Le choix d'une base de F détermine le bloc A, qui est la matrice de f_F dans cette base. Ensuite le choix d'une base $\bar{\mathcal{B}} = (\varepsilon_1, \ldots, \varepsilon_{n-p})$ de E/F détermine le bloc B, qui est la matrice de $\bar{f}_F$ dans cette base. Jusqu'ici on n'a pas eu à choisir de supplémentaire à F.

Enfin, il faut choisir des représentants $e'_1, \ldots, e'_{n-p} \in E$ de $\varepsilon_1, \ldots, \varepsilon_{n-p} \in E/F$ pour déterminer le bloc C. Ce choix détermine un supplémentaire $F' = \mathrm{Vect}(e'_1, \ldots, e'_{n-p})$ de F dans E, qui incarne le quotient E/F dans E, voir 1.36. On prendra garde au fait que ce supplémentaire F' n'a aucune raison d'être stable par f (d'ailleurs on a déjà vu qu'il est tout-à-fait possible qu'un sous-espace stable ne possède aucun supplémentaire stable, voir l'exemple 1.21).

1.43 Exercice. Soit $f, g \in \mathcal{L}(E)$, et F un sous-espace stable par f et par g. Montrer que F est stable par $f \circ g$, et

$$(f \circ g)_F = f_F \circ g_F \quad \text{et} \quad (\overline{f \circ g})_F = \bar{f}_F \circ \bar{g}_F.$$

En déduire que dans un produit MN de matrices triangulaires supérieures par blocs (de tailles compatibles), les blocs diagonaux de MN sont les produits des blocs diagonaux de M et N respectivement.

1.44 Corollaire. *Soit $f \in \mathcal{L}(E)$ et F stable par f. Considérons les matrices*

$$\begin{pmatrix} A & C \\ 0 & B \end{pmatrix} \quad \textit{et} \quad \begin{pmatrix} A' & C' \\ 0 & B' \end{pmatrix}$$

de f dans deux bases $\mathcal{B}$ et $\mathcal{B}'$ compatibles à F. Alors les matrices A et A' d'une part, et B et B' d'autre part, sont conjuguées.

Preuve. D'après la Proposition 1.41, les matrices A et A' (resp. B et B') représentent le même endomorphisme f_F (resp. $\bar{f}_F$) dans les bases de F (resp. E/F) induites par $\mathcal{B}$ et $\mathcal{B}'$ respectivement. Ceci implique le résultat. □

1.45 Proposition. *Soit $f \in \mathcal{L}(E)$.*
(i) Soit F un sous-espace stable par f. Alors $\chi_f = \chi_{f_F} \cdot \chi_{\bar{f}_F}$.
(ii) On suppose qu'il existe une décomposition $E = \bigoplus_{i=1}^r F_i$ où tous les F_i sont stables par f. Alors on a l'égalité

$$\chi_f = \prod_{i=1}^r \chi_{f_{F_i}}.$$

Preuve. L'assertion (i) est une conséquence directe de la proposition 1.41 ci-dessus et du calcul des déterminants triangulaires par blocs **??**.

L'assertion (ii) s'obtient de la même façon en observant que la matrice de f dans une base compatible à la décomposition $E = \bigoplus_{i=1}^r F_i$ est diagonale par blocs en application de 1.16 et 1.10. □

2 – Analyse

2.1 – Polynômes d'endomorphismes

2.1 Définition. On considère le morphisme de $\mathbf{k}$-algèbres $\theta_f : \mathbf{k}[X] \to \mathcal{L}_{\mathbf{k}}(E)$, dit *morphisme d'évaluation* en f, défini par la condition $\theta_f(X) = f$.

Puisque X engendre $\mathbf{k}[X]$ comme $\mathbf{k}$-algèbre, cette condition détermine de manière unique le morphisme θ_f. Pour $P = a_n X^n + \cdots + a_1 X + a_0$, avec $a_i \in \mathbf{k}$ pour tout $i = 0, \ldots, n$, on a

$$\begin{aligned} \theta_f(P) &= a_n \theta_f(X)^n + \cdots + a_1 \theta_f(X) + a_0 \theta_f(1) \\ &= a_n f^n + \cdots + a_1 f + a_0 \mathrm{id}_E, \end{aligned}$$

où pour tout $i \in \mathbf{N}$, $f^i = f \circ \cdots \circ f$ (i facteurs).

On notera $P(f) = \theta_f(P) \in \mathcal{L}(E)$ pour tout $P \in \mathbf{k}[X]$. On désignera par $\mathbf{k}[f]$ l'image du morphisme d'évaluation θ_f ; un *polynôme en f* est un éléments de $\mathbf{k}[f]$. On notera que $\mathbf{k}[f]$ est une sous-algèbre de $\mathcal{L}(E)$. Elle est de dimension finie puisque $\mathcal{L}(E)$ est de dimension finie égale à $\dim(E)^2$, et commutative puisque $\mathbf{k}[X]$ est une algèbre commutative. Ainsi, deux polynômes en f commutent toujours.

L'utilisation de polynômes en f sera fondamentale pour notre approche de la réduction des endomorphismes. Pour commencer, remarquons que la plupart des endomorphismes habituellement associés à f sont en fait des polynômes en f.

2.2 Exemples. L'inverse de f (lorsqu'il existe), l'exponentielle de f, sont des polynômes en f. Pour l'inverse voir le corollaire 2.8 et la remarque 2.8.1 ci-dessous. Pour l'exponentielle, cela vient du fait que $\exp(f)$ est une limite de polynômes en f, et que $\mathbf{k}[f]$ est fermé dans $\mathcal{L}(E)$.

On verra plus tard des exemples de projecteurs intrinsèquement liés à f (à savoir les projecteurs spectraux, voir 2.44) qui sont des polynômes en f.

Un premier intérêt de considérer des polynômes d'endomorphismes pour la réduction est qu'ils permettent de produire plein de sous-espaces stables.

2.3 Lemme. *Soit $f, g \in \mathcal{L}(E)$ deux endomorphismes qui commutent, et $P \in \mathbf{k}[X]$. Alors le sous-espace vectoriel* $\ker(P(f))$ *est stable par g.*

En particulier, en prenant $g = f$, on obtient que $\ker(P(f))$ est stable par f. La preuve de ce lemme est élémentaire et laissée au lecteur.

2.3.1 Mise en garde. Même si f et g commutent, il est grossièrement faux que tout sous-espace stable par f est stable par g. Par exemple, $g = \mathrm{id}$ commute avec tout f, tout sous-espace de E est stable par id, mais bien sûr il existe en général des sous-espaces de E qui ne sont pas stables par f.

2.4 Valeurs propres, espaces propres. Soit $\lambda \in \mathbf{k}$. On note $E_\lambda(f) = \ker(f - \lambda.\mathrm{id}_E)$ (ou simplement E_λ s'il n'y a pas de risque de confusion), qu'on appelle *espace propre* de f pour la valeur λ. On dit que λ est *valeur propre* de f si $E_\lambda(f) \supsetneq \{0\}$. L'ensemble des valeurs propres de f est appelé le *spectre* de f, noté $\mathrm{Sp}(f)$. Un scalaire $\lambda \in \mathbf{k}$ est valeur propre de f si et seulement si λ est racine du polynôme caractéristique $\chi_f = \det(X\mathrm{id}_E - f)$.

On a $E_\lambda(f) = \ker\big((X - \lambda)(f)\big)$, donc les espaces propres de f sont stables par tout endomorphisme commutant avec f.

2.5 Polynômes annulateurs, polynôme minimal. Un *polynôme annulateur* pour f est un polynôme $P \in \mathbf{k}[X]$ tel que $P(f) = 0$. Le *polynôme minimal* de f est l'unique générateur unitaire du noyau du morphisme d'évaluation θ_f.[3] On le note μ_f, ou μ si on ne craint pas de confusion.

Le polynôme minimal de f est caractérisé par les deux conditions suivantes :

(i) $\mu(f) = 0 \in \mathcal{L}(E)$;

(ii) pour tout $P \in \mathbf{k}[X]$, si $P(f) = 0$ alors μ divise P.

Ainsi μ est le polynôme unitaire annulant f dont le degré est le plus petit possible.

Par définition, on a $\ker(\theta_f) = (\mu_f)$, et donc un isomorphisme canonique $\mathbf{k}[f] \cong \mathbf{k}[X]/(\mu_f)$. En particulier, la dimension de la $\mathbf{k}$-algèbre $\mathbf{k}[f]$ est $\deg(\mu_f)$.

2.6 Proposition. *Le polynôme minimal est un invariant de similitude. Il est stable par extension des scalaires.*

2.6.1 Exercice. Montrer que

$$\deg(\mu_f) = \min\big\{p : (\mathrm{id}, f, \dots, f^p) \text{ est liée}\big\}$$

(où l'on considère $(\mathrm{id}, f, \dots, f^p)$ comme une famille de vecteurs de $\mathcal{L}(E)$).

2.6.2 Application. Invariance de μ par extension de corps, *via* l'invariance du rang. (La bonne preuve est que μ est un invariant de similitude, et que ceux-ci se calculent par un pivot de Gauss).

Soit $\mathbf{k} \to \mathbf{k}'$ une extension de corps. Certainement $\mu_{u^{\mathbf{k}'}}(u^{\mathbf{k}}) = 0$, donc $\mu_{u^{\mathbf{k}}} | \mu_{u^{\mathbf{k}'}}$. Ces deux polynômes ont le même degré par 2.6.1 et l'invariance du rang par extension de corps (voir **??** du Prologue), donc diffèrent d'une constante multiplicative. □

3. on rappelle à cet égard que l'anneau $\mathbf{k}[X]$ est principal, donc il existe $P \in \mathbf{k}[X]$ tel que $\ker(\theta_f) = (P)$. Puisque $\mathbf{k}[X]$ est de dimension infinie et $\mathcal{L}(E)$ est de dimension finie, θ_f ne peut pas être injectif, et donc $\ker(\theta_f) \supsetneq (0)$. Ainsi $P \neq 0$, et puisque les générateurs de l'idéal $\ker(\theta_f)$ sont les aP avec $a \in \mathbf{k}^*$, il existe un unique générateur de $\ker(\theta_f)$ qui est unitaire.

2.7 Proposition. *Soit $f \in \mathcal{L}(E)$, $P \in \mathbf{k}[X]$. Les deux propositions suivantes sont équivalentes :*
(i) $P(f)$ est inversible ;
(ii) $P(f)$ est inversible et $P(f)^{-1} \in \mathbf{k}[f]$;
(iii) P et μ_f sont premiers entre eux.

Preuve. L'implication (ii)⇒(i) est triviale. Montrons que (i) ⇒ (iii) : supposons que $P(f)$ est inversible. Soit R un facteur irréductible de P. Alors par minimalité de μ, R ne divise pas μ : en effet s'il existait Q tel que $\mu = QR$, alors puisque $R(f)$ est inversible [4] on aurait nécessairement $Q(f) = 0$. Ceci implique que P et μ sont premiers entre eux.

Montrons que (iii) ⇒ (ii) : supposons P et μ premiers entre eux. Alors il existe $U, V \in \mathbf{k}[X]$ tels que $PU + \mu V = 1$. En évaluant en f, on obtient

$$P(f)Q(f) + \underbrace{\mu(f)}_{=0} V(f) = \mathrm{id} \quad \Longleftrightarrow \quad P(f)Q(f) = \mathrm{id},$$

ainsi $P(f)$ est inversible d'inverse $Q(f)$. □

2.8 Corollaire. *Soit $f \in \mathcal{L}(E)$. L'endomorphisme f est inversible si et seulement si $\mu_f(0) \neq 0$. Dans ce cas $f^{-1} \in \mathbf{k}[f]$.*

Preuve. On applique la proposition avec $P = X$; les polynômes X et μ sont premiers entre eux si et seulement si X ne divise pas μ, ce qui équivaut à la condition $\mu(0) \neq 0$. □

2.8.1 Remarque. Dans la situation du corollaire, il est facile en pratique de trouver un polynôme donnant l'inverse. Si $\mu = a_p X^p + \cdots + a_1 X + a_0$ avec $a_0 \neq 0$, alors

$$a_p f^p + \cdots + a_1 f + a_0.\mathrm{id} = f(a_p f^{p-1} + \cdots + a_1.\mathrm{id}) + a_0.\mathrm{id} = 0$$

donc $f^{-1} = -a_0^{-1}(a_p f^{p-1} + \cdots + a_1.\mathrm{id})$.

2.9 Proposition. *Soit $f \in \mathcal{L}(E)$ et P un polynôme annulateur de f (par exemple $P = \mu_f$ ou χ_f). Pour tout $A \in \mathbf{k}[X]$, on a*

$$A(f) = R(f)$$

où R est le reste de la division euclidienne de A par P.

La preuve est élémentaire et laissée au lecteur.

2.9.1 Application : calcul de puissance. La proposition ci-dessus fournit un moyen efficace de calculer f^k pour de grands entiers k. Une autre méthode consiste à utiliser la décomposition de Dunford de f, voir **??**.

2.10 Lemme. *Les polynômes χ et μ ont les mêmes racines.*

On verra plus loin (Prop. 3.7) qu'en fait, χ et μ ont toujours les mêmes facteurs irréductibles, même s'ils ne sont pas scindés.

Preuve. Soit $\lambda \in \mathbf{k}$. On a la suite d'équivalences :

$$\begin{aligned} \chi(\lambda) = 0 &\iff f - \lambda.\mathrm{id} \text{ non inversible} \\ &\iff X - \lambda \text{ et } \mu \text{ non premiers entre eux} \qquad \text{(par la proposition 2.7)} \\ &\iff \mu(\lambda) = 0. \end{aligned}$$

□

4. il existe S tel que $P = RS$, $P(f) = S(f) \circ R(f)$ est inversible, donc nécessairement $R(f)$ est inversible.

2.10.1 Remarque. On peut aussi démontrer que toute valeur propre de f est racine de μ_f par le calcul suivant. Soit $P \in \mathbf{k}[X]$, et $e \in E_\lambda(f)$. On a

$$P(f)(e) = P(\lambda).e.$$

Si λ est valeur propre, on peut supposer $e \neq 0$; le calcul précédent indique alors que λ est racine de tout polynôme annulateur de f.

Avant d'aborder le prochain résultat, il est bon de faire l'observation suivante.

2.11 Lemme. *Soit $f \in \mathcal{L}(E)$ et F un sous-espace stable par f. Pour tout polynôme $P \in \mathbf{k}[X]$, le sous-espace F est stable par $P(f)$, et on a*

$$P(f)_F = P(f_F) \in \mathcal{L}(F) \quad \textit{et} \quad \overline{P(f)}_F = P(\bar{f}_F) \in \mathcal{L}(E/F).$$

Ce lemme est essentiellement tautologique et nous l'utiliserons sans même y penser. Il faut toutefois s'assurer qu'on sait le démontrer. Sans surprise la preuve n'est pas très instructive, mais elle est particulièrement fastidieuse.

Preuve. Pour commencer, on démontre par récurrence sur $k \in \mathbf{N}$ que F est stable par f^k, $(f^k)_F = (f_F)^k$, et $\overline{(f^k)}_F = (\bar{f}_F)^k$. Pour $k = 0$, F est bien stable par $f^0 = \mathrm{id}_E$, et on a bien $(\mathrm{id}_E)_F = \mathrm{id}_F$, et $\overline{(\mathrm{id}_E)}_F = \mathrm{id}_{E/F}$. Soit $k \in \mathbf{N}^*$, et supposons le résultat démontré pour tout $k' < k$. Montrons que F est stable par f^k : soit $x \in F$; on a $f^{k-1}(x) \in F$ par hypothèse de récurrence, donc $f^k(x) = f(f^{k-1}(x)) \in F$ puisque F est stable par f. Montrons que $(f^k)_F = (f_F)^k$: soit $x \in F$; on a $(f^k)_F(x) = f^k(x)$ par définition, donc $(f^k)_F(x) = f(f^{k-1}(x)) = f((f_F)^{k-1}(x))$ par hypothèse de récurrence, et ainsi par définition de f_F et puisque $(f_F)^{k-1}(x) \in F$, $(f^k)_F(x) = f_F((f_F)^{k-1}(x)) = (f_F)^k(x)$ comme il fallait démontrer. La preuve du fait que $\overline{(f^k)}_F = (\bar{f}_F)^k$ est similaire, et nous ne l'infligerons à personne.

À présent considérons $P \in \mathbf{k}[X]$ que nous écrirons $P = a_k X^k + \cdots + a_0$. L'endomorphisme $P(f)$ est une combinaison linéaire de $f^k, \ldots, f, \mathrm{id}_E$, qui laissent tous F stable d'après ce qui précède, donc F est stable par $P(f)$. Montrons que $P(f)_F = P(f_F)$: soit $x \in F$; on a

$$\begin{aligned} P(f)_F(x) = P(f)(x) &= a_k f^k(x) + \cdots + a_1 f(x) + a_0 x \\ &= a_k f_F^k(x) + \cdots + a_1 f_F(x) + a_0 x = P(f_F)(x). \end{aligned}$$

La preuve de l'identité $\overline{P(f)}_F = P(\bar{f}_F)$ est similaire. □

Ce résultat semble moins abscons écrit matriciellement ; en vertu de la proposition 1.38, ceci prend la formule suivante : pour une matrice M triangulaire supérieure par blocs comme ci-dessous, pour tout $P \in \mathbf{k}[X]$, la matrice $P(M)$ est comme indiqué ci-dessous :

$$M = \begin{pmatrix} A & C \\ 0 & B \end{pmatrix} \; ; \qquad P(M) = \begin{pmatrix} P(A) & * \\ 0 & P(B) \end{pmatrix}. \tag{2.11.1}$$

Cependant pour garder les choses à l'endroit, il faut bien se souvenir que les règles de calcul pour les matrices triangulaires par blocs proviennent des règles de composition pour les endomorphismes laissant un même sous-espace stable, voir exercice 1.43, et donc *in fine* c'est bien la formule (2.11.1) qui provient du lemme 2.11 et pas l'inverse.

2.12 Proposition. *Soit $f \in \mathcal{L}(E)$ et F un sous-espace stable par f. Alors μ_{f_F} et $\mu_{\bar{f}_F}$ divisent tous les deux μ_f ; de manière équivalente,* $\mathrm{ppcm}(\mu_{f_F}, \mu_{\bar{f}_F})$ *divise* μ_f.

Si de plus F possède un supplémentaire stable, alors on a égalité $\mu_f = \mathrm{ppcm}(\mu_{f_F}, \mu_{\bar{f}_F})$.

Preuve. Il résulte des définitions des endomorphismes induits $f_F \in \mathcal{L}(F)$ et $\bar{f}_F \in \mathcal{L}(E/F)$ que tout polynôme annulateur pour f est *a fortiori* annulateur pour f_F et $\bar{f}_F$. Autrement dit on a les deux inclusions d'idéaux

$$(\mu_f) \subseteq (\mu_{f_F}) \quad \text{et} \quad (\mu_f) \subseteq (\mu_{\bar{f}_F}),$$

respectivement équivalentes aux deux relations de divisibilité $\mu_{f_F} | \mu_f$ et $\mu_{\bar{a}f_F} | \mu_f$. □

2.2 – Endomorphismes trigonalisables

2.13 Définition. Un endomorphisme $f \in \mathcal{L}(E)$ est *trigonalisable* s'il existe une base $\mathcal{B}$ de E telle que la matrice $\mathrm{Mat}_{\mathcal{B}}(f)$ est triangulaire supérieure.

2.13.1 Remarque.
f trigonalisable $\Leftrightarrow$ $\exists$ une base $(e_1, \ldots, e_n)$ t.q. chaque $\mathrm{Vect}(e_1, \ldots, e_p)$ est stable par f
$\Leftrightarrow$ $\exists$ une filtration $\{0\} = F_0 \subsetneq F_1 \subsetneq \ldots \subsetneq F_n = E$ t.q. chaque F_p est stable par f.

2.13.2 Exercice. Les deux notions de trigonalisabilité supérieure et inférieure sont elles équivalentes ?

2.14 Théorème. *Les trois propositions suivantes sont équivalentes :*
(i) *f trigonalisable ;*
(ii) *χ_f scindé ;*
(iii) *μ_f scindé.*

L'énoncé contient en particulier l'équivalence "χ scindé $\Leftrightarrow$ μ scindé" qui n'a rien d'évident.

L'implication (ii) $\Rightarrow$(iii) est un corollaire direct du théorème de Cayley–Hamilton. Ici nous allons démontrer directement (i) $\Leftrightarrow$ (ii) et (i) $\Leftrightarrow$ (iii), sans utiliser Cayley–Hamilton. Ceci nous permettra de donner une nouvelle preuve de ce théorème, voir 2.15.

Preuve. L'implication (i) $\Rightarrow$ (ii) est une conséquence directe du calcul du déterminant d'une matrice triangulaire supérieure.

Montrons (ii) $\Rightarrow$ (i) par récurrence sur la dimension. Si $\dim(E) \leqslant 1$ le résultat est trivial. Supposons $\dim(E) > 1$ et le résultat démontré pour les endomorphismes d'un espace vectoriel de dimension strictement inférieure. On suppose χ_f scindé. Puisque $\dim(E) > 0$ on a $\chi_f \neq 1$, donc χ_f possède au moins une racine $\alpha \in \mathbf{k}$. Cet α est une valeur propre de f, considérons l'espace propre F_α correspondant. C'est un sous-espace stable par f, et on a un endomorphisme $f_{\bar{F}_\alpha}$ induit sur le quotient E/F_α. Le polynôme caractéristique de $f_{\bar{F}_\alpha}$ divise celui de f d'après 1.45, donc il est scindé lui aussi ; puisque α est valeur propre, on a $F_a \neq \{0\}$ et donc $\dim(E/F_\alpha) < \dim(E)$. Par hypothèse de récurrence, on a donc que $f_{\bar{F}_\alpha}$ est trigonalisable.

Soit $(\bar{e}_{p+1}, \ldots, \bar{e}_n)$ une base de E/F_α trigonalisante pour $f_{\bar{F}_\alpha}$, et $(e_1, \ldots, e_p)$ n'importe quelle base pour F_α ($p = \dim(F_\alpha)$, $n = \dim(E)$). Alors $(e_1, \ldots, e_n)$ est une base de E d'après le lemme 1.34 du Prologue. La matrice de f dans cette base est triangulaire supérieure par blocs

$$\begin{pmatrix} \alpha.\mathrm{Id}_p & * \\ 0 & T \end{pmatrix}$$

où T est la matrice de $f_{\bar{F}_\alpha}$ dans la base $(\bar{e}_{p+1}, \ldots, \bar{e}_n)$, et est donc triangulaire supérieure. On conclut donc que la base $(e_1, \ldots, e_n)$ est trigonalisante pour f.

Montrons (i) $\Rightarrow$ (iii). On suppose f trigonalisable ; soit $(e_1, \ldots, e_n)$ une base de E dans laquelle la matrice de f est

$$\begin{pmatrix} a_1 & * & \ldots & * \\ & \ddots & \ddots & \vdots \\ & & \ddots & * \\ & & & a_n \end{pmatrix},$$

et notons $F_0 = \{0\}$, et $F_p = \mathrm{Vect}(e_1, \ldots, e_p)$ pour tout $p = 1, \ldots, n$. Pour tout p, on a $(f - a_p.\mathrm{id})(F_p) \subseteq F_{p-1}$, donc

$$(f - a_1.\mathrm{id}) \cdots (f - a_n.\mathrm{id})(F_n = E) \subseteq F_0 = \{0\}.$$

Ainsi le polynôme $\prod_p (X - a_p)$ est annulateur pour f, et comme il est scindé μ_f est nécessairement scindé lui aussi.

Montrons (iii) $\Rightarrow$ (i) par récurrence sur la dimension. On suppose μ_f scindé ; soit $\alpha \in \mathbf{k}$ une de ses racines. On commence par montrer que α est valeur propre de f. Pour cela on écrit $\mu_f = (X-\alpha)^k \tilde{\mu}$ avec $\tilde{\mu}(\alpha) \neq 0$. Si α n'était pas valeur propre de f, alors $f - \alpha.\mathrm{id}$ serait inversible et donc $\mu_f(f) = 0 \Leftrightarrow \tilde{\mu}(f) = 0$, ce qui contredirait la minimalité de μ_f.

Ensuite la preuve ressemble beaucoup à celle de (ii) $\Rightarrow$ (i). On considère F_α l'espace propre associé à α. C'est un sous-espace stable par f, et on peut considérer l'endomorphisme $f_{\bar{F}_\alpha}$ induit sur le quotient E/F_α. Le polynôme μ_f annule $f_{\bar{F}_\alpha}$, donc le polynôme minimal de $f_{\bar{F}_\alpha}$ est scindé. Puisque α est valeur propre, on a $F_a \neq \{0\}$ et donc $\dim(E/F_\alpha) < \dim(E)$. Par hypothèse de récurrence, on a donc que $f_{\bar{F}_\alpha}$ est trigonalisable. On conclut alors exactement comme dans la preuve de (ii) $\Rightarrow$ (i). □

2.15 Preuve de Cayley–Hamilton en partant du cas trigonalisable. Dans la preuve de (i) $\Rightarrow$ (iii) du théorème 2.14 ci-dessus, on a en fait démontré que si f est trigonalisable alors χ_f annule f.

En général, il existe une extension $\mathbf{k}'$ de $\mathbf{k}$ telle que χ_f soit scindé sur $\mathbf{k}'$. Alors f est trigonalisable sur $\mathbf{k}'$, et on en déduit que χ_f (qui ne dépend pas du corps de base) est annulateur pour f.

On a ainsi démontré le théorème de Cayley–Hamilton pour un endomorphisme d'un espace vectoriel de dimension finie, ou de manière équivalente pour les matrices à coefficients dans un corps. Ceci permet par l'argument habituel de démontrer l'énoncé pour des coefficients dans un anneau (commutatif) arbitraire : on considère une matrice A dont les coefficients sont des indéterminées (a_{ij}). En raisonnant dans le corps $\mathbf{Q}(a_{ij})$ on peut appliquer l'énoncé démontré plus haut, et obtenir ainsi l'identité $\chi_A(A) = 0$, qui vit en fait dans $\mathcal{M}_n(\mathbf{Z}[a_{ij}])$. Par spécialisation, ceci prouve $\chi_A(A) = 0$ pour A à coefficients dans n'importe quel anneau commutatif. □

2.16 Proposition. Si f et g sont trigonalisables et $fg = gf$, alors ils sont simultanément trigonalisables.

2.16.1 Remarque. La réciproque est fausse. Il suffit de se baisser pour ramasser un exemple. Pour ceux qui ont mal au dos, voir 2.48.1.

2.17. Si χ_f est scindé, on a la décomposition de Dunford qui nous dit que f se réduit à une partie diagonalisable et une partie nilpotente ; en particulier f est diagonalisable si et seulement si la partie nilpotente est triviale.

Ceci devrait suffire à motiver l'étude des diagonalisables à la Section 2.3 et des nilpotents à la Section 2.5. Cette étude nous permettra de donner un critère de similitude pour deux endomorphismes à polynôme caractéristique scindé (i.e. un critère de similitude pour les endomorphismes trigonalisables), sans utiliser de technique sophistiquée — croit-on.

2.3 – Lemme des noyaux

2.18 Proposition (Lemme des noyaux)**.** *Soit $P_1, \ldots, P_r \in \mathbf{k}[X]$ des polynômes deux à deux premiers entre eux. Alors on a une décomposition en somme directe*

$$\ker\big(P_1 \cdots P_r(f)\big) = \bigoplus\nolimits_{i=1}^{r} \ker\big(P_i(f)\big).$$

Les projecteurs associés à cette décomposition sont des polynômes en f que l'on peut calculer explicitement.

Il y a un léger abus de langage dans l'énoncé. Voici la version tout-à-fait rigoureuse. Notons $K = \ker\big(P_1 \cdots P_r(f)\big)$; c'est un sous-espace stable par f. Pour tout $i = 1, \ldots, r$, le projecteur $p_i \in \mathcal{L}(K)$ sur $\ker\big(P_i(f)\big)$ dans la direction de $\bigoplus_{i' \neq i} \ker\big(P_{i'}(f)\big)$ est un polynôme en $f_K \in \mathcal{L}(K)$.

2.19 Conditions d'application du lemme des noyaux. La condition "$P_1, \dots, P_r$ deux à deux premiers entre eux" signifie "pour tout $i \neq j$, $\operatorname{pgcd}(P_i, P_j) = 1$", et est plus forte que la condition "$P_1, \dots, P_r$ premiers entre eux (dans leur ensemble)", qui signifie "$\operatorname{pgcd}(P_1, \dots, P_r) = 1$".

2.19.1 Exercice. Donner un exemple de trois polynômes qui sont premiers entre eux dans leur ensemble mais pas deux à deux.
(*Indication* : on peut prendre Q_1, Q_2, Q_3 construits comme en 2.20 ci-dessous).

2.19.2 Situation classique. Le cas archétypique d'application du lemme des noyaux est quand $P_i = H_i^{\alpha_i}$ avec les H_i irréductibles deux à deux distincts. On laisse au lecteur le soin de vérifier que dans ce cas les P_i sont effectivement deux à deux premiers entre eux.

2.20. Si $P_1, \dots, P_r \in \mathbf{k}[X]$ sont deux à deux premiers entre eux, alors les $Q_i := \prod_{j \neq i} P_j$ sont premiers entre eux dans leur ensemble.[5] Puisque $\operatorname{pgcd}(Q_1, \dots, Q_r)$ est le générateur unitaire de l'idéal $(Q_1, \dots, Q_r)$, on a dans ce cas $(Q_1, \dots, Q_r) = (1)$. Ceci implique l'existence de $U_1, \dots, U_r \in \mathbf{k}[X]$ tels que

$$1 = U_1 Q_1 + \cdots + U_r Q_r.$$

2.20.1 Calcul pratique des U_i. Le principe est de raisonner par récurrence sur r et d'utiliser l'algorithme d'Euclide. Pour $r = 2$, il suffit d'appliquer directement l'algorithme d'Euclide à P_1 et P_2. Pour $r \geqslant 3$, supposons par récurrence avoir trouvé $V_1, \dots, V_{r-1}$ tels que

$$1 = V_1 R_1 + \cdots + V_{r-1} R_{r-1},$$

où $R_i = \prod_{\substack{j \neq i \\ j < r}} P_j$ pour tout $i = 1, \dots, r-1$. En multipliant par P_r, on obtient

$$P_r = V_1 Q_1 + \cdots + V_{r-1} Q_{r-1}. \tag{2.2.i}$$

Puisque P_r est premier avec chacun des P_i, $i = 1, \dots, r-1$, il est premier avec le produit $P_1 \cdots P_{r-1} = Q_r$. En appliquant l'algorithme d'Euclide, on trouve S et T tels que

$$SQ_r + TP_r = 1.$$

En remplaçant P_r par son expression dans (2.2.i), on obtient

$$SQ_r + TV_1 Q_1 + \cdots + TV_{r-1} Q_{r-1} = 1,$$

et donc $U_1 = TV_1, \dots, U_{r-1} = TV_{r-1}$ et $U_r = S$ conviennent.

2.21 Preuve du lemme des noyaux. On note $Q_i = \prod_{j \neq i} P_j$ pour tout $i = 1, \dots, r$, et on considère des polynômes $U_1, \dots, U_r \in \mathbf{k}[X]$ tels que $1 = U_1 Q_1 + \cdots + U_r Q_r$, dont l'existence est garantie par le fait que les P_i sont deux à deux premiers entre eux (exactement comme au paragraphe 2.20 ci-dessus). Évaluée en f, cette identité donne

$$\mathrm{id} = U_1(f) Q_1(f) + \cdots + U_r(f) Q_r(f). \tag{2.21.1}$$

Notons $K = \ker\big(P_1 \cdots P_r(f)\big)$, et $K_i = \ker\big(P_i(f)\big)$ pour tout $i = 1, \dots, r$. Le sous-espace K est stable par f, et l'énoncé se réduit en fait à une affirmation à propos de l'endomorphisme induit $f_K \in \mathcal{L}(K)$. On a $K_i \subseteq K$ pour tout i, et donc $\sum_i K_i \subseteq K$. Le résultat est donné par les propriétés suivantes.

1) Pour tout $i = 1, \dots, r$, $U_i Q_i(f)_K \in \mathcal{L}(K)$ est le projecteur sur K_i dans la direction de $\sum_{j \neq i} K_j$.

5. soit par l'absurde H facteur irréductible commun à tous les Q_i. H divise Q_1, donc il doit diviser un P_i, $i > 1$, disons P_2. H divise aussi Q_2, donc nécessairement aussi un P_j avec $j \neq 2$. H serait alors diviseur irréuctible commun à P_2 et P_j, une contradiction.

a) Pour tout i, $\mathrm{im}\big(U_iQ_i(f)_K\big) \subseteq K_i$. En effet, si $x \in \mathrm{im}\big(U_iQ_i(f)_K\big)$, alors il existe $y \in K$ tel que $x = U_iQ_i(f)(y)$, et donc $P_i(f)(x) = U_i(f)\big(P_iQ_i(f)(y)\big) = 0$ puisque $y \in K$.

b) Pour tout i, $\sum_{j\neq i} K_j \subseteq \ker\big(U_iQ_i(f)\big)$. En effet, pour tout $j \neq i$, si $x \in K_j$ alors $Q_i(f)(x) = 0$ puisque $P_j|Q_i$, et donc $K_j \subseteq \ker\big(U_iQ_i(f)\big)$.

a') Pour tout i, $K_i \subseteq \mathrm{im}\big(U_iQ_i(f)_K\big)$. En effet, si $x \in K_i$, alors en évaluant (2.21.1) en x on obtient

$$x = \sum\nolimits_j U_jQ_j(f)(x) = U_iQ_i(f)(x) \tag{2.21.2}$$

puisque $Q_j(f)(x) = 0$ comme on l'a déjà vu. Ainsi $x \in \mathrm{im}\big(U_iQ_i(f)_K\big)$, puisque $x \in K_i \subseteq K$.

b') Pour tout i, $\ker\big(U_iQ_i(f)_K\big) \subseteq \sum_{j\neq i} K_j$. En effet, si $x \in \ker\big(U_iQ_i(f)_K\big)$, alors en évaluant (2.21.1) en x on obtient $x = \sum_{j\neq i} U_jQ_j(f)(x)$, et donc $x \in \sum_{j\neq i} K_j$ puisque $\mathrm{im}\big(U_jQ_j(f)_K\big) \subseteq K_j$ comme on l'a vu en a).

Tout ceci implique 1) : a) et a') d'une part, et b) et b') d'autre part, disent que

$$\mathrm{im}\big(U_iQ_i(f)_K\big) = K_i \quad \text{et} \quad \ker\big(U_iQ_i(f)_K\big) = \sum\nolimits_{j\neq i} K_j,$$

et (2.21.2) dit que $U_iQ_i(f)_K$ agit comme l'identité sur K_i, donc $U_iQ_i(f)_K$ est bien le projecteur de K sur K_i dans la direction de $\sum_{j\neq i} K_j$.

2) $K = \bigoplus_i K_i$.

a) $K = \sum_i K_i$. En effet, pour tout $x \in K$ on a $x = \sum_i U_iQ_i(f)(x)$ par (2.21.1), et $U_iQ_i(f)(x) \in K_i$ d'après 1).

b) La somme $\sum_i K_i$ est directe. En effet, soit $(x_1, \ldots, x_r) \in K_1 \times \cdots \times K_r$ tel que $x_1 + \cdots + x_r = 0$. Alors pour tout i, l'image de ce vecteur par $U_iQ_i(f)$ est nulle, donc d'après 1),

$$U_iQ_i(f)(x_1 + \cdots + x_r) = U_iQ_i(f)(x_i) = x_i = 0.$$

Ceci prouve bien que les K_i sont en somme directe.

À présent 1) et 2) sont bien démontrées, et ceci conlut la preuve de la proposition. □

2.22 Corollaire. *Si P et Q sont premiers entre eux, alors $P(f)_{\ker Q(f)}$ est injectif.*

Preuve. Le noyau de $P(f)_{\ker(Q(f))}$ est $\ker(P(f)) \cap \ker(Q(f))$. D'après le lemme des noyaux, et puisque P et Q sont premiers entre eux, cette intersection est réduite à $\{0\}$. □

2.23 Exercice. 1) Soit $p \in \mathcal{L}(E)$ tel que $p^2 = p$. Montrer que p est un projecteur, c'est-à-dire qu'il existe F et G sous-espaces supplémentaires de E tels que pour tout $x_F \in F$ et $x_G \in G$: $p(x_F + x_G) = x_F$ (p est alors le projecteur sur F dans la direction de G).

2) Soit $s \in \mathcal{L}(E)$ tel que $s^2 = \mathrm{id}$. Montrer que s est une symétrie, c'est-à-dire qu'il existe F et G sous-espaces supplémentaires de E tels que pour tout $x_F \in F$ et $x_G \in G$: $s(x_F + x_G) = x_F - x_G$ (s est alors la symétrie par rapport à F dans la direction de G).

2.4 – Endomorphismes diagonalisables

2.24 Définition. Un endomorphisme $f \in \mathcal{L}(E)$ est *diagonalisable* si E se décompose en somme directe de sous-espaces propres de f.

Une condition équivalente est qu'il existe une base de E constituée de vecteurs propres pour f ; la matrice de f dans une telle base est diagonale (et pas seulement diagonale par blocs).

2.25 Notation. Pour tout scalaire α, on note $E_\alpha = \ker(f - \alpha.\mathrm{id})$ l'*espace propre* de f relativement à α. Attention, on dit que α est *valeur propre* de f seulement si (et si) $E_\alpha \neq \{0\}$. On notera éventuellement $E_\alpha(f) = E_\alpha$ en cas de risque de confusion.

2.26 Théorème. *Soit $f \in \mathcal{L}(E)$. Les trois propositions suivantes sont équivalentes :*
(i) *l'endomorphisme f est diagonalisable ;*
(ii) *il existe un polynôme annulateur de f scindé à racines simples ;*
(iii) *le polynôme minimal μ_f est scindé à racines simples.*

Preuve. Montrons "(i) $\Rightarrow$ (ii)". Si f est diagonalisable, il existe une base de E dans laquelle la matrice de f est diagonale. Notons $\alpha_1, \ldots, \alpha_r$ les scalaires apparaissant sur la diagonale de cette matrice ; on autorise les répétitions, et suppose ainsi les α_i deux à deux distincts. Un calcul direct montre que $(X - \alpha_1) \cdots (X - \alpha_r)$ annule f. Puisque les α_i sont deux à deux distincts, ce polynôme est scindé à racines simples, et donc (ii) est vérifiée.

L'implication "(ii) $\Rightarrow$ (iii)" est purement algébrique. Soit P polynôme annulateur de f scindé à racines simples. Alors μ_f divise P, donc μ_f est lui-même scindé à racines simples, ce qui prouve (iii).

Nous allons maintenant prouver l'implication "(iii) $\Rightarrow$ (i)", ce qui concluera la preuve du théorème. Notons $\mu_f = \prod_{i=1}^r (X - \alpha_i)$, où les α_i sont des scalaires deux à deux distincts. Ainsi les polynômes $X - \alpha_1, \ldots, X - \alpha_r$ sont deux à deux premiers entre eux. On peut donc leur appliquer le lemme des noyaux, ce qui donne

$$\ker(\mu_f(f)) = \bigoplus\nolimits_{i=1}^{r} E_{\alpha_i}(f).$$

Puisque μ_f annule f, on a $\ker(\mu_f(f)) = E$, et donc l'égalité ci-dessus est une décomposition de E en somme de sous-espaces propres pour f, ce qui prouve (i). □

2.27 Proposition. *Si χ_f est scindé à racines simples, alors f est diagonalisable.*

2.27.1 Mise en garde. La réciproque de la proposition 2.27 est notoirement fausse. La diagonalisabilité de f implique que χ_f est scindé (c'est un cas particulier du théorème 2.14 !), mais pas nécessairement à racines simples.

La proposition 2.27 peut s'obtenir comme un corollaire du théorème précédent, en utilisant le théorème de Cayley–Hamilton, mais on a fait le choix de ne voir le théorème de Cayley–Hamilton que comme une conséquence de tout le reste, et donc de ne jamais l'utiliser dans les preuves.

Preuve. Si χ_f est scindé à racines simples, alors il existe $\alpha_1, \ldots, \alpha_n \in \mathbf{k}$ ($n = \dim(E)$) deux à deux distincts tels que $\chi_f = \prod_{i=1}^n (X - \alpha_i)$; en effet, le polynôme caractéristique est de degré n. Pour tout $i = 1, \ldots, n$ l'endomorphisme $f - \alpha_i.\mathrm{id}$ n'est pas inversible donc son noyau E_{α_i} est non nul. D'après le lemme des noyaux,

$$\ker(\chi_f(f)) = \bigoplus\nolimits_{i=1}^{n} E_{\alpha_i},$$

donc $\ker(\chi_f(f))$ a dimension au moins n. Pour conclure, on a nécessairement $\ker(\chi_f(f)) = E$ et chaque E_{α_i} est de dimension 1. □

2.28 Proposition. *Un endomorphisme diagonalisable est semi-simple.*

Preuve. Soit $f \in \mathcal{L}(E)$ diagonalisable, et $(e_1, \ldots, e_n)$ une base de E constituée de vecteurs propres. Posons $F_i = \mathrm{Vect}(e_i)$ pour tout $i = 1, \ldots, n$. Alors chaque F_i est stable par f, et $E = \bigoplus_{i=1}^n F_i$. Pour tout i, on a $\dim(F_i) = 1$, donc l'endomorphisme induit f_{F_i} est simple. □

En vertu du théorème 1.24, les endomorphismes diagonalisables sont donc également b-semi-simples (voir définition 1.23). Il est bon toutefois de savoir démontrer cette propriété directement ; c'est la proposition 2.30 ci-dessous.

2.29 Proposition. *Soit $f \in \mathcal{L}(E)$ un endomorphisme diagonalisable. Tout sous-espace stable par f se décompose selon les sous-espaces propres de f ; autrement dit, si F est stable par f, alors*

$$F = \bigoplus\nolimits_{\alpha\in \mathrm{Sp}(f)} \big(F \cap E_\alpha(f)\big).$$

2.29.1 Attention. Pour une décomposition en somme directe $E = \bigoplus E_i$ arbitraire et F un sous-espace de E, si les $F \cap E_i$ sont bien en somme directe, il est grossièrement faux que leur somme égale F : en général l'inclusion $\bigoplus_i (F \cap E_i) \subseteq F \cap (\bigoplus_i E_i)$ peut être stricte.. C'est le cas dans la situation (très simple) représentée ci-dessous.

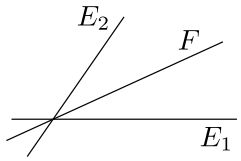


FIGURE 2 – Une situation dans laquelle $\bigoplus_i (F \cap E_i) \subsetneq F \cap (\bigoplus_i E_i)$

Preuve du lemme. Puisque F est stable par f, on peut considérer l'endomorphisme induit $f_F \in \mathcal{L}(F)$. On a $F \cap E_\alpha(f) = \ker(f_F - \alpha.\mathrm{id}_F)$ pour tout $\alpha \in \mathbf{k}$. Le polynôme minimal μ_f est scindé à racines simples, et annule f donc aussi f_F. On en déduit que f_F est diagonalisable, et donc

$$F = \bigoplus\nolimits_{\alpha\in\mathbf{k}} \ker(f_F - \alpha.\mathrm{id}_F) = \bigoplus\nolimits_{\alpha\in \mathrm{Sp}(f)} \big(F \cap E_\alpha(f)\big).$$

□

2.30 Proposition. *Si $f \in \mathcal{L}(E)$ est diagonalisable, alors tout sous-espace stable par f possède un supplémentaire dans E stable par f lui aussi.*

Preuve. Soit F stable par f. On a

$$F = \bigoplus\nolimits_{\alpha\in \mathrm{Sp}(f)} \big(F \cap E_\alpha(f)\big).$$

d'après la proposition 2.29. Pour chaque $\alpha \in \mathrm{Sp}(f)$, choisissons F'_α supplémentaire de $F \cap E_\alpha(f)$ dans $E_\alpha(f)$. Alors le sous-espace

$$F' = \bigoplus\nolimits_{\alpha\in \mathrm{Sp}(f)} F'_\alpha$$

est un supplémentaire de F dans E, et il est stable par f. □

2.31 Lemme. *Soit $f \in \mathcal{L}(E)$ un endomorphisme diagonalisable. Pour tout $\alpha \in \mathbf{k}$, la dimension de l'espace propre E_α est égale à la multiplicité de α comme racine du polynôme caractéristique χ_f.*

Preuve. Puisque f est diagonalisable, on a $E = \bigoplus_{\alpha\in \mathrm{Sp}(f)} E_\alpha$. Puisque les sous-espaces propres sont stables par f, on a donc (corollaire **??**)

$$\chi_f = \prod\nolimits_{\alpha\in \mathrm{Sp}(f)} \chi_{f_{E_\alpha}}.$$

Or pour tout scalaire α, $f_{E_\alpha} = \mathrm{id}_{E_\alpha}$, donc $\chi_{f_{E_\alpha}} = (X-\alpha)^{\dim(E_\alpha)}$, donc

$$\chi_f = \prod\nolimits_{\alpha\in \mathrm{Sp}(f)} (X-\alpha)^{\dim(E_\alpha)},$$

et le résultat est démontré. □

2.32 Proposition. *Soit $f, g \in \mathcal{L}(E)$ deux endomorphismes diagonalisables. Alors les trois propositions suivantes sont équivalentes :*
(i) *f et g sont semblables ;*
(ii) $\chi_f = \chi_g$ *;*
(iii) *pour tout $\lambda \in \mathbf{k}$, $\dim(E_\lambda(f)) = \dim(E_\lambda(g))$.*

Preuve. L'implication "(i) $\Rightarrow$ (ii)" a déjà été vue ref. L'implication "(ii) $\Rightarrow$ (iii)" est une conséquence directe du lemme 2.31.

Nous allons montrer "(iii) $\Rightarrow$ (i)", ce qui achèvera la preuve de la proposition. La condition (i) nous dit d'ores et déjà que f et g ont le même spectre. De plus, pour toute valeur propre λ de f et g, puisque $\dim(E_\lambda(f)) = \dim(E_\lambda(g))$, il existe un isomorphisme $\varphi_\lambda : E_\lambda(f) \simeq E_\lambda(g)$. Puisque $f_{E_\lambda(f)} = \mathrm{id}_{E_\lambda(f)}$ et $g_{E_\lambda(g)} = \mathrm{id}_{E_\lambda(g)}$, on a

$$f_{E_\lambda(f)} = \varphi_\lambda \circ g_{E_\lambda(g)} \circ \varphi_\lambda^{-1}.$$

On en déduit que f et g sont semblables en appliquant le lemme 1.18. □

On dit que deux endomorphismes f et g sont *simultanément diagonalisables* s'il exsite une base de E constituée de vecteurs qui sont à la fois propres pour f et pour g.

2.33 Théorème. *Soit $f, g \in \mathcal{L}(E)$ deux endomorphismes diagonalisables. Les deux conditions suivantes sont équivalentes :*
(i) *f et g sont simultanément diagonalisables ;*
(ii) *f et g commutent.*

Preuve. L'implication "(ii) $\Rightarrow$ (i)" est une conséquence directe du fait que le produit de deux matrices diagonalisables est commutatif : si $D_1 = \mathrm{diag}(\alpha_1, \dots, \alpha_n)$ et $D_2 = \mathrm{diag}(\beta_1, \dots, \beta_n)$, alors

$$D_1 D_2 = \mathrm{diag}(\alpha_1\beta_1, \dots, \alpha_n\beta_n) = D_2 D_1.$$

Montrons la réciproque. On considère la décomposition de E selon les sous-espaces propres de f,

$$E = \bigoplus\nolimits_{\lambda \in \mathrm{Sp}(f)} E_\lambda(f).$$

Puisque g commute avec f, il résulte du lemme 2.3 que pour tout $\lambda \in \mathrm{Sp}(f)$, $E_\lambda(f)$ est stable par g. L'endomorphisme induit $g_{E_\lambda(f)} \in \mathcal{L}(E_\lambda(f))$ est diagonalisable, car il est annulé par μ_g qui est scindé à racines simples. Il existe donc une base $\mathcal{B}_\lambda$ de $E_\lambda(f)$ qui est constituée de vecteurs propres pour $g_{E_\lambda(f)}$, et donc aussi pour g. Ces vecteurs sont des vecteurs de $E_\lambda(f)$, donc ils sont automatiquement propres pour f. La concaténation des familles $\mathcal{B}_\lambda$, $\lambda \in \mathrm{Sp}(f)$, est une base de E, qui est constituée de vecteurs propres à la fois pour f et pour g. Ceci prouve que f et g sont simultanément diagonalisables. □

Il est aussi possible de prouver le théorème 2.33 directement par le calcul. Ceci prend la forme suivante.

2.34 Version matricielle de la simultanée diagonalisabilité. Soit $n \in \mathbf{N}^*$, et $p_1, \dots, p_r \in \mathbf{N}^*$ tels que $n = p_1 + \cdots + p_r$. Considérons les deux matrices carrées de taille n

$$A = \begin{pmatrix} \alpha_1 \mathrm{Id}_{p_1} & & \\ & \ddots & \\ & & \alpha_r \mathrm{Id}_{p_r} \end{pmatrix} \quad \text{et} \quad B = \begin{pmatrix} B_{11} & \cdots & B_{1r} \\ \vdots & & \vdots \\ B_{r1} & \cdots & B_{rr} \end{pmatrix} :$$

A est diagonale, et B est une matrice par blocs où chaque B_{ij} est de taille $p_i \times p_j$. On suppose $\alpha_1, \dots, \alpha_r$ deux à deux disjoints. Alors A et B commutent si et seulement si B est diagonale par blocs,

$$B = \begin{pmatrix} B_{11} & & 0 \\ & \ddots & \\ 0 & & B_{rr} \end{pmatrix}.$$

En effet, il résulte des règles de calcul pour les matrices par blocs que

$$AB = \begin{pmatrix} \alpha_1 B_{11} & \cdots & \alpha_1 B_{1r} \\ \vdots & & \vdots \\ \alpha_r B_{r1} & \cdots & \alpha_r B_{rr} \end{pmatrix} \quad : \quad \text{et} \quad BA = \begin{pmatrix} \alpha_1 B_{11} & \cdots & \alpha_r B_{1r} \\ \vdots & & \vdots \\ \alpha_1 B_{r1} & \cdots & \alpha_r B_{rr} \end{pmatrix}.$$

Puisque les α_i sont deux à deux distincts, on a donc $AB = BA$ si et seulement si $B_{ij} = 0$ si $i \neq j$, comme il fallait démontrer.

Ceci permet de suivre matriciellement la preuve du critère de diagonalisation simultanée. Soit $f, g \in \mathcal{L}(E)$ deux endomorphismes diagonalisables qui commutent. En considérant $\mathcal{B}$ une base de E diagonalisante pour f, on a d'après ce qui précède

$$A = \begin{pmatrix} \alpha_1 \mathrm{Id}_{p_1} & & \\ & \ddots & \\ & & \alpha_r \mathrm{Id}_{p_r} \end{pmatrix} \quad \text{et} \quad B = \begin{pmatrix} B_1 & & 0 \\ & \ddots & \\ 0 & & B_r \end{pmatrix},$$

où $\{\alpha_1, \ldots, \alpha_r\} = \mathrm{Sp}(f)$. Puisque g est diagonalisable, chaque bloc B_i est diagonalisable, donc il existe des matrices inversibles $P_1, \ldots, P_r$ de tailles $p_1, \ldots, p_r$ telles que pour tout i, $P_i^{-1} B_i P_i = \Delta_i$ est une matrice diagonale. Alors $P = \mathrm{diag}(P_1, \ldots, P_r)$ est une matrice (diagonale par blocs) inversible, et

$$P^{-1}AP = \begin{pmatrix} \alpha_1 \mathrm{Id}_{p_1} & & \\ & \ddots & \\ & & \alpha_r \mathrm{Id}_{p_r} \end{pmatrix} \quad \text{et} \quad P^{-1}BP = \begin{pmatrix} \Delta_1 & & 0 \\ & \ddots & \\ 0 & & \Delta_r \end{pmatrix}$$

sont toutes les deux diagonales (pas seulement diagonales par blocs), ce qui prouve que f et g sont simultanément diagonalisables.

2.35 Exercice. Soit $f_1, \ldots, f_N \in \mathcal{L}(E)$ des endomorphismes diagonalisables commutant deux à deux. Montrer qu'ils sont simultanément diagonalisables dans leur ensemble.

Cette propriété permet de démontrer que les représentations irréductibles complexes d'un groupe abélien fini sont toutes de dimension 1.

2.5 – Endomorphismes nilpotents

Dans cette section, étant donné un endomorphisme $f \in \mathcal{L}(E)$, nous noterons $K_i = \ker(f^i)$ pour tout $i \in \mathbf{N}$ (ou $K_i(f)$ s'il y a un risque d'ambiguïté), et $k_i = \dim(K_i)$. La proposition suivante ne nécessite pas que f soit nilpotent.

2.36 Proposition. *Soit $f \in \mathcal{L}(E)$.*
(i) *La suite de sous-espaces $(K_i)_{i \geqslant 0}$ est croissante.*
(ii) *La suite $(\dim K_{i+1} - \dim K_i)_{i \geqslant 0}$ est décroissante.*

En particulier, il existe $i_0 \in \mathbf{N}$ tel que :
(i) pour tout $i \leqslant i_0$, $K_{i-1} \subsetneq K_i$;
(ii) pour tout $i \geqslant i_0$, $K_i = K_{i_0}$.
Autrement dit, la suite des noyaux K_i est "d'abord strictement croissante, puis constante".

Preuve. (i) La suite est croissante.

Supposons $K_i = K_{i+1}$. Montrons que $K_{i+1} = K_{i+2}$. Puisque la suite est croissante, il suffit de montrer que $K_{i+2} \subseteq K_{i+1}$. Soit $x \in K_{i+2}$. Alors $f(x) \in K_{i+1} = K_i$, donc $x \in K_{i+1}$.

(ii) On a $f(K_{i+1}) \subseteq K_i$, donc en composant f avec la projection $K_i \to K_i/K_{i-1}$, on obtient une application linéaire ${}^\sharp\bar{f}_{i+1} : K_{i+1} \to K_i/K_{i-1}$. Son noyau est K_i :

$${}^\sharp\bar{f}_{i+1}(x) = 0 \Leftrightarrow f(x) \in K_{i-1} \Leftrightarrow f^i(x) = 0.$$

Ainsi, ${}^\sharp\bar{f}_{i+1}$ induit une application linéaire injective

$$\bar{f}_{i+1} : K_{i+1}/K_i \to K_i/K_{i-1}, \tag{2.36.1}$$

et donc $k_{i+1} - k_i \leqslant k_i - k_{i-1}$. □

2.36.1 Remarque. On a des résultats analogues avec la suite des images des itérés de f. On peut les prouver directement comme ci-dessus, ou bien les déduire sans effort en appliquant l'énoncé donné plus haut à l'endomorphisme transposé $f^\mathsf{T} \in \mathcal{L}(E^\star)$.

2.37 Corollaire. Si f nilpotent, alors son indice de nilpotence est $\leqslant \dim E$.

2.38 Théorème. *Soit $f, g \in \mathcal{L}(E)$ deux endomorphismes nilpotents. Les deux propositions suivantes sont équivalentes :*
(i) f et g sont semblables ;
(ii) pour tout i, $\dim\big(\ker(f^i)\big) = \dim\big(\ker(g^i)\big)$.

2.38.1 Remarque. Pour que l'énoncé ci-dessus contienne le théorème de Jordan, il manque l'unicité de la forme normale à permutation des blocs près. Ceci est très bien expliqué dans les notes [**?**]. Une bonne façon de démontrer l'unicité est de montrer l'injectivité de l'application qui à une forme réduite de Jordan associe la suite des dimensions des noyaux de ses itérés, voir 2.39.

Avant de se lancer dans la preuve, il est bon de se souvenir qu'elle s'écrit de droite à gauche.

Preuve. $(i) \Rightarrow (ii)$ est immédiat. La stratégie pour prouver la réciproque est d'écrire tout endomorphisme nilpotent sous une forme normale ne dépendant que des dimensions des noyaux de ses itérés ; ainsi si f et g satisfont à la propriété (ii), ils ont une forme normale commune et sont donc semblables.

Soit donc f nilpotent d'indice de nilpotence p, et notons pour tout i, $k_i = \dim K_i = \dim\big(\ker(f^i)\big)$ et $s_i = k_i - k_{i-1}$. On commence par choisir une base $(\bar{e}_1, \dots, \bar{e}_{s_p})$ de $K_p/K_{p-1} = E/K_{p-1}$. Puisque l'application linéaire

$$\bar{f}_p : \bar{x} \in K_p/K_{p-1} \mapsto \overline{f(x)} \in K_{p-1}/K_{p-2}$$

de (2.36.1) est injective, $\big(\overline{f(e_1)}, \dots, \overline{f(e_{s_p})}\big)$ est une famille libre de vecteurs de K_{p-1}/K_{p-2}. On la complète en une base

$$\big(\overline{f(e_1)}, \dots, \overline{f(e_{s_p})}, \bar{e}_{s_p+1}, \dots, \bar{e}_{s_{p-1}}\big),$$

remarquant au passage que la notation est consistante puisque $s_p \leqslant s_{p-1}$.

On obtient ainsi par récurrence des vecteurs

$$e_1, \dots, e_{s_p}, e_{s_p+1}, \dots\dots, e_{s_2}, e_{s_2+1}, \dots, e_{s_1} \in E$$

(à nouveau : $s_p \leqslant \dots \leqslant s_2 \leqslant s_1$) tels que pour tout $i = 1, \dots, p$, (i) les vecteurs $e_{s_{i+1}+1}, \dots, e_{s_i}$ sont dans K_i, et (ii) les classes des vecteurs

$$f^{p-i}(e_1), \dots, f^{p-i}(e_{s_p}), \dots, f(e_{s_{i+2}+1}), \dots, f(e_{s_{i+1}}), e_{s_{i+1}+1}, \dots, e_{s_i}$$

constituent une base de K_i/K_{i-1}.

Le Lemme 1.34 du Prologue nous assure alors que la famille

$$\begin{array}{cccc} e_1, \dots, e_{s_p}, & & & \\ f(e_1), \dots, f(e_{s_p}), & & e_{s_p+1}, \dots, e_{s_{p-1}}, & \\ \vdots & & \vdots & \ddots \\ f^{p-1}(e_1), \dots, f^{p-1}(e_{s_p}), & f^{p-2}(e_{s_p+1}), \dots, f^{p-2}(e_{s_{p-1}}), & \dots\dots, & e_{s_{1+1}+1}, \dots, e_{s_1} \end{array}$$

est une base de E. On obtient une forme normale de Jordan dans la base obtenue en renumérotant la base ci-dessus en lisant colonne par colonne de gauche à droite et de haut en bas. [6] Le nombre de blocs de Jordan de taille i est $s_i - s_{i+1}$ (notons que $s_i = 0$ pour $i > p$), entièrement déterminé par la suite $(k_i)_{i\geqslant 0}$. □

2.39 Unicité de la forme normale de Jordan. (Examen 2019 ; inspiré par [**?**]).
1) Soit $n \geqslant 1$. On considère la matrice $A_n = (a_{ij})_{1\leqslant i,j\leqslant n} \in \mathcal{M}_n(\mathbf{Q})$ définie par

$$\forall i, j \in [\![1, n]\!] \quad a_{ij} = \begin{cases} j & \text{si } j \leqslant i \\ i & \text{si } j > i. \end{cases}$$

a) Montrer que A_3 est inversible et calculer son inverse.
b) Montrer que la matrice

$$P_n = \begin{pmatrix} 2 & -1 & & & \\ -1 & 1 & & & \\ -1 & 0 & 1 & & \\ \vdots & \vdots & & \ddots & \\ -1 & 0 & & & 1 \end{pmatrix} \in \mathcal{M}_n(\mathbf{Q})$$

est inversible, puis calculer $P_n \times A_n$.
c) Montrer que A_n est inversible pour tout $n \geqslant 1$.
2) On considère $F_1, \ldots, F_r$ sous-espaces vectoriels de E tels que $E = \bigoplus_{1\leqslant i\leqslant r} F_i$.
a) Soit $g \in \mathcal{L}(E)$ tel que pour tout $i = 1, \ldots, r$, F_i est stable par g. Montrer que

$$\ker(g) = \bigoplus_{1\leqslant i\leqslant r} \big(\ker(g) \cap F_i\big).$$

b) Soit $f \in \mathcal{L}(E)$ tel que pour tout $i = 1, \ldots, r$, F_i est stable par f. Montrer que pour tout $P \in \mathbf{k}[X]$,

$$\ker\big(P(f)\big) = \bigoplus_{1\leqslant i\leqslant r} \big(\ker(P(f)) \cap F_i\big).$$

3) On considère l'endomorphisme f de $\mathbf{k}^r$ défini par la multiplication à gauche par la matrice

$$J_r = \begin{pmatrix} 0 & & & \\ 1 & 0 & & \\ & \ddots & \ddots & \\ & & 1 & 0 \end{pmatrix}.$$

Calculer $\dim\big(\ker(f^i)\big)$ pour tout $i \in \mathbf{N}$.
4) Soit $f \in \mathcal{L}(E)$ et $b_1, \ldots, b_n$ ($n = \dim(E)$) des entiers. On suppose qu'il existe une décomposition

$$E = \bigoplus_{1\leqslant r\leqslant n} \bigoplus_{1\leqslant a\leqslant b_r} F_{r,a}$$

telle que chaque $F_{r,a}$ est de dimension r et stable par f, et $f_{F_{r,a}}$ est semblable à l'endomorphisme de $\mathbf{k}^r$ de la question 3.
a) En utilisant les questions 2b et 3, calculer $k_i := \dim\big(\ker(f^i)\big)$ pour tout $i \in \mathbf{N}$. En déduire que

$$\begin{pmatrix} k_1 \\ \vdots \\ k_n \end{pmatrix} = A_n \times \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$$

6. ceci fournit des blocs de Jordan sous forme de matrice compagnon, avec les '1' sous la diagonale ; pour les avoir au dessus, il faut lire les colonnes de bas en haut.

où A_n est la matrice de la question 1.
b) Conclure que s'il existe une autre décomposition

$$E = \bigoplus_{1\leqslant r\leqslant n} \bigoplus_{1\leqslant a\leqslant b'_r} F'_{r,a}$$

telle que chaque $F'_{r,a}$ est de dimension r et stable par f, et $f_{F'_{r,a}}$ est semblable à l'endomorphisme de $\mathbf{k}^r$ de la question 3, alors $b_i = b'_i$ pour tout $i = 1, \ldots, n$. □

2.40 Curiosité. Quelles valeurs la suite des k_i peut-elle prendre ? Les seules contraintes sont celles imposées par les inégalités du 2.36, qui se réduisent à

$$0 \leqslant k_n - k_{n-1} \leqslant \ldots \leqslant k_2 - k_1 \leqslant k_1 - k_0 = k_1.$$

En effet, ces inégalités donnent une condition nécessaire et suffisante pour que

$$A^{-1} \times \begin{pmatrix} k_1 \\ \vdots \\ k_n \end{pmatrix} = \begin{pmatrix} 2 & -1 & & & \\ -1 & 2 & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & -1 & 2 & 1 \\ & & & -1 & 1 \end{pmatrix} \times \begin{pmatrix} k_1 \\ \vdots \\ k_n \end{pmatrix}$$

ait toutes ses entrées $\geqslant 0$.

On conclut par le traditionnel énoncé de simultanée propriété-ité.

2.41 Proposition. *Si n et n' sont nilpotents et commutent, alors $n + n'$ est nilpotent.*

2.42 Proposition. *Un endomorphisme nilpotent est semi-simple si et seulement si il est nul.*

Preuve. Soit $f \in L(E)$ nilpotent et semi-simple. Alors f est aussi b-semi-simple d'après le théorème 1.24. Considérons le noyau de f ; il est stable par f, et doit posséder un supplémentaire F dans E, lui aussi stable par f. Alors l'endomorphisme induit $f_F \in \mathcal{L}(F)$ est injectif, puisque $\ker(f_F) = \ker(f) \cap F = \{0\}$. D'autre part, puisque pour tout entier $N \geqslant 0$, $(f_F)^N = (f^N)_F$, f_F est lui aussi nilpotent. Ceci impose $F = \{0\}$, et donc $\ker(f) = E$, autrement dit $f = 0$. □

2.6 – Endomorphismes à polynôme caractéristique scindé

En vertu du Théorème 2.14, cette section pourrait s'appeler *Endomorphismes trigonalisables, II.* Dans toute cette section, on suppose que χ_f est scindé, ou de manière équivalente μ_f est scindé.

2.43 Notation. Écrivons

$$\mu = \prod\nolimits_{\lambda\in \mathrm{Sp}} (X-\lambda)^{a_\lambda} \quad \text{et} \quad \chi = \prod\nolimits_{\lambda\in \mathrm{Sp}} (X-\lambda)^{b_\lambda}.$$

2.44 Définition. Les *sous-espaces caractéristiques* de f sont les $C_\lambda = \ker\big((f - \lambda.\mathrm{id})^{a_\lambda}\big)$ pour $\lambda \in \mathrm{Sp}(f)$. On a la décomposition

$$E = \bigoplus\nolimits_{\lambda\in \mathrm{Sp}} C_\lambda \tag{2.44.1}$$

en somme de sous-espaces stables par f (et par tout g commutant avec f). Les *projecteurs spectraux* $p_\lambda \in \mathcal{L}(E)$, qui s'écrivent

$$p_{\lambda_i}(x_{\lambda_1}, \ldots, x_{\lambda_r}) = (0, \ldots, x_{\lambda_i}, \ldots, 0)$$

dans la décomposition (3.8.1), sont des polynômes en f que l'on sait calculer explicitement, voir le lemme des noyaux 2.18.

2.45 Proposition. *Soit $\lambda \in \mathrm{Sp}(f)$.*
i) L'entier a_λ est l'indice de stagnation de la suite des $K_i(\lambda) = \ker\big((f-\lambda.\mathrm{id})^i\big)$, et l'indice de nilpotence de $f_{C_\lambda} - \lambda.\mathrm{id}_{C_\lambda} \in \mathcal{L}(C_\lambda)$. En particulier, $f_{C_\lambda} - \lambda.\mathrm{id}_{C_\lambda} \in \mathcal{L}(C_\lambda)$ est nilpotent.
ii) L'entier b_λ est la dimension du sous-espace caractéristique C_λ.

On verra plus loin une version un peu plus générale de ce résultat (Proposition 3.9). Les preuves sont strictement identiques, et en fait la formulation pour 3.9 me semble plus lisible car moins polluée par les notations.

Preuve. i) D'après le lemme des noyaux, $f - \lambda.\mathrm{id}$ est inversible sur $\bigoplus_{\lambda' \neq \lambda} C_{\lambda'}$, donc $K_i(\lambda) = C_\lambda \cap \ker\big((f-\lambda.\mathrm{id})^i\big)$ par le Lemme **??**, et ce noyau s'identifie canoniquement à $\ker\big((f_{C_\lambda} - \lambda.\mathrm{id}_{C_\lambda})^i\big)$.

Par définition de C_λ, on a $K_{a_\lambda}(\lambda) = C_\lambda$. Pour $i \geqslant a_\lambda$, on a $K_{a_\lambda}(\lambda) \subseteq K_i(\lambda) \subseteq C_\lambda$ et donc $K_i(\lambda) = C_\lambda$. Il reste à démontrer que pour $i < a_\lambda$, on a $K_i(\lambda) \subsetneq C_\lambda$, ce qui est garanti par la minimalité de μ.

Précisons ce dernier point. Notons $Q = \prod_{\lambda' \neq \lambda} (X - \lambda')^{a_{\lambda'}}$. Le polynôme Q annule l'endomorphisme de $\bigoplus_{\lambda' \neq \lambda} C_{\lambda'}$ induit par f, puisque $(f - \lambda.\mathrm{id})^{a_\lambda} \circ Q(f) = 0$ et $f - \lambda.\mathrm{id}$ est inversible sur $\bigoplus_{\lambda' \neq \lambda} C_{\lambda'}$. Ainsi si on a $K_i = C_\lambda$, alors le polynôme $(X-\lambda)^i Q$ annule f, donc il est divisible par μ, et nécessairement $i \geqslant a_\lambda$.

ii) La décomposition $\bigoplus C_\lambda$ est une somme de sous-espaces stables, donc $\chi_f = \prod_\lambda \chi_{f_{C_\lambda}}$. Or pour chaque valeur propre λ, $\chi_{f_{C_\lambda}} = (X-\lambda)^{\dim C_\lambda}$, puisque $(f - \lambda.\mathrm{id})_{C_\lambda}$ est nilpotent. On a donc nécessairement $\dim(C_\lambda) = b_\lambda$. □

2.46 Application au théorème de Cayley–Hamilton. La Proposition 2.45 ci-dessus offre en corollaire le théorème de Cayley–Hamilton pour les endomorphismes trigonalisables. Pour démontrer le théorème en général, on se ramène au cas trigonalisable par un argument d'extension des scalaires comme en 2.15.

Soit donc $f \in \mathcal{L}(E)$ trigonalisable, et conservons les notations introduites ci-dessus. Pour chaque $\lambda \in \mathrm{Sp}(f)$, a_λ est l'indice de nilpotence de $f_{C_\lambda} - \lambda.\mathrm{id} \in \mathcal{L}(C_\lambda)$, donc $a_\lambda \leqslant \dim(C_\lambda)$. Puisque d'autre part $b_\lambda = \dim(C_\lambda)$, on a donc $a_\lambda \leqslant b_\lambda$ pour tout $\lambda \in \mathrm{Sp}(f)$, et ainsi $\mu_f | \chi_f$. □

2.47 Théorème (Décomposition de Jordan–Chevalley, aussi dite de Dunford)**.** *Soit $f \in \mathcal{L}(E)$ trigonalisable. Il existe une unique paire (d, n) d'endomorphismes de E satisfaisant aux quatre conditions suivantes :*
(i) *$f = d + n$;*
(ii) *d est diagonalisable ;*
(iii) *n est nilpotent ;*
(iv) *$dn = nd$.*
De plus, d et n sont tous les deux des polynômes en f, qu'on sait calculer explicitement.

La condition (iv) de commutativité est aussi importante que le reste ; elle assure que les endomorphismes d et n sont "compatibles". Dans l'exemple 2.48, nous allons voir qu'on perd l'unicité si on retire la condition de commutativité, et que la décomposition vérifiant cette condition est "la bonne". Dans la remarque 2.49, nous allons voir que la condition de commutativité assure qu'on peut trouver une base de E dans laquelle d et n sont tous les deux sous forme normale.

Nous donnerons la preuve du Théorème 2.47 après cet exemple et cette remarque.

2.48 Exemple. Considérons la matrice

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix}.$$

Elle est diagonalisable puisque son polynôme caractéristique est $(X-1)(X-2)$ qui est scindé à racines simples, donc sa décomposition est

$$A = A + 0$$

où A est diagonalisable et 0 est nilpotente.

2.48.1 Remarque. Cet exemple donne de façon amusante un exemple de deux matrices triangulaires supérieures qui ne commutent pas. En effet, posons

$$T_1 = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} \quad \text{et} \quad T_2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

On a (i) T_1 diagonalisable, (ii) T_2 nilpotente, et (iii) $A = T_1 + T_2$. Par unicité de la décomposition, $T_1 + T_2$ n'est pas la décomposition de A (puisque c'est $A + 0$). On a donc nécessairement $T_1 \times T_2 \neq T_2 \times T_1$. Bien sûr, on peut calculer explicitement les deux produits T_2T_1 et T_1T_2 et voir de ses propres yeux qu'ils sont différents.

2.49 Remarque. Il existe une base dans laquelle d a une matrice diagonale et n est sous forme réduite de Jordan. Une telle base est en particulier trigonalisante pour $f = d + n$.

En effet, considérons la décomposition de E selon les sous-espaces propres $E_\lambda = E_\lambda(d)$ de d : $E = \bigoplus_{\lambda \in \mathrm{Sp}(d)} E_\lambda$. Puisque d et n commutent, ces sous-espaces sont stables par n (et donc aussi par $f = d + n$). Pour chaque valeur propre λ de d, considérons une base $\mathcal{B}_\lambda$ de E_λ dans laquelle la matrice de l'endomorphisme de E_λ est sous forme réduite de Jordan. Alors $\mathcal{B} = (\mathcal{B}_\lambda)_{\lambda \in \mathrm{Sp}(d)}$ est une base de E comme on voulait. En contemplant les écritures de f, d, n dans cette base, on constate que $\mathrm{Sp}(d) = \mathrm{Sp}(f)$, et les sous-espaces propres de d sont les sous-espaces caractéristiques de f, autrement dit $E_\lambda(d) = C_\lambda(f)$.

Écrivons tout ceci un peu plus explicitement. Appelons $\lambda_1, \ldots, \lambda_r$ les valeurs propres de d, et pour tout $i = 1, \ldots, r$, notons E_i l'espace propre $E_{\lambda_i}(d)$, $\mathcal{B}_i$ la base $\mathcal{B}_{\lambda_i}$, f_i, d_i et n_i les endomorphismes de E_i induits respectivement par f, d et n. Par construction, on a

$$\mathrm{Mat}_{\mathcal{B}_i}(d_i) = \lambda_i.\mathrm{Id}_{b_i} \quad \text{et} \quad \mathrm{Mat}_{\mathcal{B}_i}(n_i) = \mathrm{diag}(J_{c_1^i}, \ldots, J_{c_{s_i}^i}),$$

où la notation J_c désigne un bloc de Jordan de taille c. Ainsi

$$\mathrm{Mat}_{\mathcal{B}_i}(f_i) = \mathrm{diag}\big(J_{c_1^i}(\lambda_i), \ldots, J_{c_{s_i}^i}(\lambda_i)\big),$$

où la notation J_c désigne $J_c + \lambda.\mathrm{Id}_c$. Puisque les E_i sont stables par f, on a

$$\begin{aligned} \mathrm{Mat}_{\mathcal{B}}(f) &= \mathrm{diag}\big(\mathrm{Mat}_{\mathcal{B}_1}(f_1), \ldots, \mathrm{Mat}_{\mathcal{B}_r}(f_r)\big) \\ &= \mathrm{diag}\big(J_{c_1^1}(\lambda_1), \ldots, J_{c_{s_1}^1}(\lambda_1),\ \ldots,\ J_{c_1^r}(\lambda_r), \ldots, J_{c_{s_r}^r}(\lambda_r)\big). \end{aligned}$$

La multiplicité de λ_i comme racine de χ_f est $b_i = \sum_{j=1}^{s_i} c_j^i$, et comme racine de μ_f est $a_i = \max(c_1^i, \ldots, c_{s_i}^i)$. D'après le Théorème 2.38, les tailles $c_1^i, \ldots, c_{s_i}^i$ des blocs de Jordan de n_i sont déterminées par les dimensions des noyaux des $n_i^k = (f - \lambda_i.\mathrm{id})_{E_i}$, $k \in \mathbf{N}$.

Les considérations de la remarque précédente contiennent tous les ingrédients nécessaires à la construction d'une décomposition de Dunford. La preuve ci-dessous en fait la synthèse. Notons aussi que la remarque précédente contient les arguments pour démontrer le Théorème 2.52 plus loin.

Preuve du Théorème 2.47. On note $p_\lambda \in \mathcal{L}(E)$ les projecteurs spectraux ; ce sont des polynômes en f. On a $\mathrm{id}_E = \sum_{\lambda \in \mathrm{Sp}(f)} p_\lambda$, et donc

$$f = f \circ \sum\nolimits_{\lambda \in \mathrm{Sp}(f)} p_\lambda = \underbrace{\sum\nolimits_\lambda \lambda.p_\lambda}_{=:d} + \underbrace{\sum\nolimits_\lambda (f - \lambda.\mathrm{id}) \circ p_\lambda}_{=:n}.$$

Manifestement, d et n satisfont aux propriétés (i)–(iv) du Théorème 2.47, et l'existence d'une décomposition est bien démontrée. En outre, les d et n que nous avons construits sont bien des polynômes en f.

Montrons l'unicité de la décomposition. Soit (d', n') une autre paire satisfaisant aux conditions (i)–(iv). *A priori* rien ne dit que d' et n' sont des polynômes en f. Cependant, d' commute à lui-même et à n', donc aussi à $f = d' + n'$. Puisque $d \in \mathbf{k}[f]$, d' commute aussi à d. De la même façon, n' commute à n. On en déduit que $d' - d$ est diagonalisable, et $n - n'$ nilpotent. Enfin, puisque $d + n = d' + n'$, on a $d' - d = n - n'$. Cet endomorphisme est à la fois nilpotent et diagonalisable, donc il est nul (puisqu'il est nilpotent, 0 est son unique valeur propre, et puisqu'il est diagonalisable il est donc nécessairement nul ; on peut aussi appliquer la Proposition 2.42). Ainsi $d = d'$ et $n = n'$, ce qui prouve l'unicité. □

2.50 Exemple. Considérons la matrice

$$A = \begin{pmatrix} -3 & -2 & -2 \\ -2 & 0 & -1 \\ 10 & 5 & 6 \end{pmatrix}.$$

Son polynôme caractéristique est $\chi = (X-1)^3$, donc A possède un seul sous-espace caractéristique $C_1 = \mathbf{k}^3$, et la décomposition de Jordan–Chevalley de A est

$$A = \mathrm{Id}_3 + (A - \mathrm{Id}_3),$$

où Id_3 est diagonalisable et $A - \mathrm{Id}_3$ est nilpotente (on peut vérifier par le calcul que $(A - \mathrm{Id})^2 = 0$; voir aussi l'exemple 3.41).

2.51 Proposition. *Soit $f \in \mathcal{L}(E)$ trigonalisable. Alors f est semi-simple si et seulement si il est diagonalisable.*

Le théorème suivant permet de décider en pratique si deux endomorphismes trigonalisables sont semblables, voir 2.53.

2.52 Théorème (Caractérisation des classes de similitude)**.** *Soit f et g deux endomorphismes trigonalisables. Les deux conditions suivantes sont équivalentes :*

(i) *f et g sont semblables ;*

(ii) *pour tout $\lambda \in \mathbf{k}$ et tout $s \in \mathbf{N}$,*

$$\dim\big(\ker(f - \lambda.\mathrm{id})^s\big) = \dim\big(\ker(g - \lambda.\mathrm{id})^s\big).$$

La preuve de ce théorème est la synthèse d'arguments apparaissant plus haut dans la Remarque 2.49.

Preuve du Théorème 2.52. L'implication "(i) ⇒ (ii)" est claire ; nous allons démontrer la réciproque. Grâce au lemme 1.18, il suffit de savoir le faire sous-espace caractéristique par sous-espace caractéristique, et sur chaque sous-espace caractéristique nous allons appliquer le critère 2.38.

Soit $\lambda \in \mathbf{k}$. Les endomorphismes $(f - \lambda\mathrm{id})_{C_\lambda(f)} \in \mathcal{L}(C_\lambda(f))$ et $(g - \lambda\mathrm{id})_{C_\lambda(g)} \in \mathcal{L}(C_\lambda(g))$ sont nilpotents ; puisque

$$\ker(f - \lambda\mathrm{id})^i = C_\lambda(f) \cap \ker(f - \lambda\mathrm{id})^i \cong \ker\big((f - \lambda\mathrm{id})_{C_\lambda(f)}\big)^i,$$

et de même $\ker(g - \lambda\mathrm{id})^i \cong \ker(g - \lambda\mathrm{id})^i_{C_\lambda(g)}$, la condition (ii) nous dit que la condition (ii) du théorème 2.38 est vérifiée pour $(f - \lambda\mathrm{id})_{C_\lambda(f)} \in \mathcal{L}(C_\lambda(f))$ et $(g - \lambda\mathrm{id})_{C_\lambda(g)} \in \mathcal{L}(C_\lambda(g))$. On en déduit qu'il existe un isomorphisme $\varphi_\lambda : C_\lambda(f) \simeq C_\lambda(g)$ tel que

$$(f - \lambda\mathrm{id})_{C_\lambda(f)} = \varphi_\lambda \circ (g - \lambda\mathrm{id})_{C_\lambda(g)} \circ \varphi_\lambda^{-1},$$

ce qui implique que $f_{C_\lambda(f)} = \varphi_\lambda \circ g_{C_\lambda(g)} \circ \varphi_\lambda^{-1}$. On peut alors conclure grâce au lemme 1.18 que f et g sont semblables. □

2.53. Le critère de similitude du Théorème 2.52 permet de décider de manière algorithmique si deux endomorphismes trigonalisables sont semblables. Faisons-le directement avec des matrices.

On considère $A, B \in \mathcal{M}_n(\mathbf{k})$ et on se demande si elles sont semblables. On commence par calculer les polynômes caractéristiques χ_A et χ_B. S'ils sont distincts on peut s'arrêter tout de suite, A et B ne sont pas semblables. Si $\chi_A = \chi_B$, on décompose ce polynôme en produit de facteurs premiers. S'il n'est pas scindé, pour l'instant on ne sait pas répondre à la question et il faut utiliser des techniques plus élaborées, voir section 3.6.

Supposons donc que $\chi_A = \chi_B$ est scindé, et l'ensemble de ses racines est $\{\lambda_1, \ldots, \lambda_r\}$. Si $\lambda \in \mathbf{k} \setminus \{\lambda_1, \ldots, \lambda_r\}$, $A - \lambda\mathrm{Id}$ et $B - \lambda\mathrm{Id}$ sont inversibles, donc

$$\dim\big(\ker(A - \lambda.\mathrm{Id})^s\big) = \dim\big(\ker(B - \lambda.\mathrm{Id})^s\big) = 0$$

pour tout $s \in \mathbf{N}$. Il faut donc nous concentrer sur $\lambda_1, \ldots, \lambda_r$. Pour chaque $i = 1, \ldots, r$, on calcule successivement les

$$k_s(\lambda_i, A) = \dim\big(\ker(A - \lambda_i.\mathrm{Id})^s\big)$$

jusqu'à trouver un $s_0 \in \mathbf{N}$ tel que $k_{s_0}(\lambda_i, A) = k_{s_0+1}(\lambda_i, A)$. On sait alors par la Proposition 2.36 que $k_s(\lambda_i, A) = k_{s_0}(\lambda_i, A)$ pour tout $s \geqslant s_0$. On fait la même chose pour les

$$k_s(\lambda_i, B) = \dim\big(\ker(B - \lambda_i.\mathrm{Id})^s\big).$$

Il n'y a plus qu'à comparer les valeurs obtenues pour les différents $k_s(\lambda_i, A)$ et $k_s(\lambda_i, B)$ pour décider si A et B sont semblables. On retiendra qu'il suffit de calculer un nombre fini de dimensions de noyaux pour pouvoir conclure.

2.54 Exercice. Soit $\lambda, \mu \in \mathbf{k}$ distincts. Donner un représentant de toutes les classes de similitude de matrices de $\mathcal{M}_5(\mathbf{k})$ dont le polynôme minimal est $(X-\lambda)^2(X-\mu)$. Bien justifier que les matrices que vous donnerez sont deux à deux non semblables.

Voir aussi Exercice 3.47 pour des questions du même goût.

3 – Synthèse

3.1 – Sous-espaces cycliques

3.1 Définition. Soit $f \in \mathcal{L}(E)$ et $x \in E$. On appelle *sous-espace cyclique* associé à x, noté F_x, le plus petit sous-espace vectoriel de E stable par f et contenant x.

On appelle *polynôme minimal local* de f en x, noté μ_x, le générateur unitaire de l'idéal $\{P \in \mathbf{k}[X] \mid P(f)(x) = 0\}$.

Pour commencer, notons qu'il existe effectivement un élément minimal pour l'inclusion dans l'ensemble des sous-espaces stables par f contenant x, puisque "être stable" et "contenir x" sont deux propriétés stables par intersection. Ainsi, on a

$$F_x = \bigcap_{\substack{F \text{ stable par } f \\ \text{et } x \in F}} F.$$

3.2 Proposition. *Soit $f \in \mathcal{L}(E)$ et $x \in E$. Le polynôme minimal $\mu_{f_{F_x}}$ de l'endomorphisme $f_{F_x} \in \mathcal{L}(F_x)$ induit par f sur le sous-espace cyclique de x est le polynôme μ_x, polynôme minimal de f en x.*

Preuve. On a $\mu_{f_{F_x}}(f_{F_x}) = 0$, donc $\mu_{f_{F_x}}(f_{F_x})(x) = \mu_{f_{F_x}}(f)(x) = 0$, et ainsi $\mu_x | \mu_{f_{F_x}}$.

Réciproquement, $\ker \mu_x(f)$ est un sous-espace stable par f qui contient x, donc $F_x \subseteq \ker \mu_x(f)$. Autrement dit, $\mu_x(f_{F_x}) = 0$, et $\mu_{f_{F_x}} | \mu_x$.

On a finalement bien $\mu_{f_{F_x}} = \mu_x$ puisque ces deux polynômes sont unitaires. □

3.3 Proposition. *Soit $x \in E$. On note $p = \deg(\mu_x)$ et*

$$\mu_x = X^p + a_{p-1}X^{p-1} + \cdots + a_0.$$

Alors la famille $(x, f(x), \ldots, f^{p-1}(x))$ est une base de F_x.

Preuve. On a nécessairement $x \in F_x$. Puisque F_x est stable par f, on a donc aussi $f(x) \in F_x$, puis $f(f(x)) = f^2(x) \in F_x$, et ainsi par récurrence $f^k(x) \in F_x$ pour tout $k \in \mathbf{N}$. On a donc $\mathrm{Vect}\big(f^k(x), k \in \mathbf{N}\big) \subseteq F_x$. D'autre part $\mathrm{Vect}\big(f^k(x), k \in \mathbf{N}\big)$ est manifestement un sous-espace vectoriel de E, stable par f et contenant x, donc par minimalité de F_x on a aussi l'inclusion inverse $F_x \subseteq \mathrm{Vect}\big(f^k(x), k \in \mathbf{N}\big)$, et finalement $F_x = \mathrm{Vect}\big(f^k(x), k \in \mathbf{N}\big)$.

Montrons par récurrence sur $k \in \mathbf{N}$ que $f^k(x)$ est combinaison linéaire de $x, f(x), \ldots, f^{p-1}(x)$. Si $k \leqslant p-1$ c'est trivial. Si $k \geqslant p$, on suppose $f^{k-1}(x)$ combinaison linéaire de $x, f(x), \ldots, f^{p-1}(x)$ par hypothèse de récurrence. Alors $f^k(x) = f(f^{k-1}(x))$ est combinaison linéaire de $f(x), f^2(x), \ldots, f^p(x)$. Puisque

$$\mu_x(f)(x) = f^p(x) + a_{p-1}f^{p-1}(x) + \cdots + a_0\mathrm{id}(x) = 0,$$

$f^p(x)$ est combinaison linéaire de $x, f(x), \ldots, f^{p-1}(x)$, et finalement $f^k(x)$ est bien combinaison linéaire de $x, f(x), \ldots, f^{p-1}(x)$. Ceci démontre que $(x, f(x), \ldots, f^{p-1}(x))$ engendre F_x.

Montrons maintenant que cette famille est libre. Soit $\alpha_0, \ldots, \alpha_{p-1} \in \mathbf{k}$ tels que

$$\alpha_0 x + \cdots + \alpha_{p-1}f^{p-1}(x) = 0.$$

Alors le polynôme $\alpha_0 + \cdots + \alpha_{p-1}X^{p-1}$ appartient à l'idéal $\{P \in \mathbf{k}[X] \mid P(f)(x) = 0\}$, donc il est divisible par μ_x. Puisque $\deg(\mu_x) = p$, ceci impose que $\alpha_0 = \cdots = \alpha_{p-1} = 0$. □

3.4. Écrivons la matrice C_x de $f_{F_x} \in \mathcal{L}(F_x)$ dans la base $(x, f(x), \ldots, f^{p-1}(x))$ (on conserve les notations introduites ci-dessus). Pour $k < p-1$ on a $f(f^k(x)) = f^{k+1}(x)$ (pour $k \geqslant p-1$ aussi, d'ailleurs !), tandis que pour $k = p$ on a la relation observée ci-dessus

$$f^p(x) = -a_{p-1}f^{p-1}(x) - \cdots - a_0\mathrm{id}(x).$$

On a donc

$$C_x = \begin{pmatrix} 0 & & & -a_0 \\ 1 & \ddots & & \vdots \\ & \ddots & 0 & -a_{p-2} \\ & & 1 & -a_{p-1} \end{pmatrix}.$$

En particulier, on a $\mu_{C_x} = \mu_{f_{F_x}}$, et donc d'après la Proposition 3.2,

$$\mu_{C_x} = \mu_x = X^p + a_{p-1}X^{p-1} + \cdots + a_0. \tag{3.4.1}$$

3.5 Définition. Soit $P = X^p + a_{p-1}X^{p-1} + \cdots + a_1X + a_0 \in \mathbf{k}[X]$ un polynôme unitaire de degré p. On appelle *matrice compagnon* de P la matrice

$$\mathrm{Comp}(P) = \begin{pmatrix} 0 & & & -a_0 \\ 1 & \ddots & & \vdots \\ & \ddots & 0 & -a_{p-2} \\ & & 1 & -a_{p-1} \end{pmatrix} \in \mathcal{M}_p(\mathbf{k}).$$

3.6 Proposition. *On a $\chi_{\mathrm{Comp}(P)} = \mu_{\mathrm{Comp}(P)} = P$.*

Preuve. L'affirmation sur le polynôme minimal a été démontrée ci-dessus, voir (3.4.1). Pour démontrer l'autre identité, nous allons calculer explicitement le polynôme caractéristique de la matrice $\mathrm{Comp}(P)$. On développe $\det(X.\mathrm{Id}_p - \mathrm{Comp}(P))$ par rapport à la dernière colonne :

$$\begin{aligned}\chi_{C_x} &= \sum_{k=1}^{p-1}(-1)^{k+p}a_{k-1}\cdot\begin{vmatrix} X & & & & & & \\ -1 & X & & & & & \\ & \ddots & \ddots & & & & \\ & & -1 & X & 0 & & \\ & & & -1 & X & & \\ & & & & \ddots & \ddots & \\ & & & & & -1 & X \\ & & & & & & -1 \end{vmatrix} + (X+a_{p-1})\cdot\begin{vmatrix} X & & & \\ -1 & X & & \\ & \ddots & \ddots & \\ & & -1 & X \end{vmatrix} \\ &= \sum_{k=1}^{p-1}(-1)^{k+p}a_{k-1}\cdot(-1)^{p-k}X^{k-1} + (X+a_{p-1})\cdot X^{p-1} \\ &= \sum_{k=0}^{p-2}a_kX^k + a_{p-1}X^{p-1} + X^p\end{aligned}$$

comme il fallait. □

Le développement par rapport à la dernière colonne effectué ci-dessus peut éventuellement nécessiter plusieurs tentatives avant d'être réussi. Il est un peu moins acrobatique de faire une récurrence sur p en développant par rapport à la première ligne. Nous recommandons plutôt cette seconde méthode en conditions de stress.

3.2 – Sous-espaces caractéristiques

Avant de définir les sous-espaces caractéristiques, nous allons démontrer deux énoncés reliant les polynômes caractéristiques et scindés.

Le premier est le fameux théorème de Cayley–Hamilton. La preuve donnée ici est directe, en ce qu'elle ne passe pas par un argument d'extension des scalaires pour se ramener au cas trigonalisable. Cette preuve due à Frobenius est la première qui ait été donnée du théorème de Cayley–Hamilton en toute généralité : Cayley ne l'avait démontré que pour des matrices de taille 2 ou 3, et Hamilton pour certaines matrices de taille 4 en utilisant "ses" quaternions.

Le second résultat a lui aussi été démontré précédemment avec l'hypothèse supplémentaire que f est trigonalisable, c'est le Lemme 2.10.

3.7 Proposition. *Soit $f \in \mathcal{L}(E)$.*
(i) *Le polynôme minimal μ_f divise le polynôme caractéristique χ_f.*
(ii) *Les polynômes μ_f et χ_f ont les mêmes facteurs irréductibles.*

Preuve. Pour montrer (i) il suffit de démontrer que χ_f annule f, ou de manière équivalente que $\chi_f(f)(x) = 0$ pour tout $x \in E$. Soit $x \in E$. On considère le sous-espace cyclique F_x associé à x. Puisqu'il est stable par f, $\chi_{f_{F_x}}$ divise χ_f (Proposition 1.45). D'autre part, il résulte des résultats de la section 3.1 que $\chi_{f_{F_x}} = \mu_x$, le polynôme minimal de f en x. On a donc $\chi_{f_{F_x}}(f)(x) = 0$, et par suite $\chi_f(f)(x) = 0$ comme on voulait.

Pour démontrer (ii), maintenant qu'on sait que μ_f divise χ_f, il suffit de démontrer que tout facteur irréductible de χ_f est nécessairement facteur irréductible de μ_f. Faisons-le par récurrence sur la dimension de E. Si $\dim(E) = 1$ le résultat est vrai car μ_f et χ_f sont tous les deux de degré 1. Si $\dim(E) \geqslant 2$, considérons $x \in E$ non-nul et F_x le sous-espace cyclique associé. On a $\chi_f = \chi_{f_{F_x}}\chi_{\bar{f}_{F_x}}$, et $\mu_{f_{F_x}}$ et $\mu_{\bar{f}_{F_x}}$ divisent tous les deux μ_f (Propositions 1.45 et 2.12). Soit P polynôme irréductible divisant χ_f. Puisqu'il est irréductible il divise $\chi_{f_{F_x}}$ ou $\chi_{\bar{f}_{F_x}}$. Si P divise

$\chi_{f_{F_x}}$, alors puisque $\chi_{f_{F_x}} = \mu_{f_{F_x}}$ (voir section 3.1), P divise $\mu_{f_{F_x}}$, et donc μ_f comme il fallait démontrer. Si P divise $\chi_{\bar{f}_{F_x}}$, il divise $\mu_{\bar{f}_{F_x}}$ par hypothèse de récurrence (on a choisi $x \neq 0$, donc $F_x \supsetneq \{0\}$, et ainsi $\dim(E/F_x) < \dim(E)$), et donc aussi μ_f. □

3.8 Sous-espaces caractéristiques. On appelle $\mathrm{Irr} \subseteq \mathbf{k}[x]$ l'ensemble des facteurs irréductibles de χ et/ou μ. On note

$$\mu = \prod\nolimits_{P\in\mathrm{Irr}} P^{a_P} \quad \text{et} \quad \chi = \prod\nolimits_{P\in\mathrm{Irr}} P^{b_P}.$$

Les *sous-espaces caractéristiques* de f sont les $C_P = \ker\big(P^{a_P}(f)\big)$ pour $P \in \mathrm{Irr}$. D'après le lemme des noyaux, on a la décomposition

$$E = \bigoplus\nolimits_{P\in\mathrm{Irr}} C_P \tag{3.8.1}$$

en somme de sous-espaces stables par f, et par tout endomorphisme commutant à f. Les *projecteurs spectraux* sont les projecteurs relativement à cette décomposition, et ce sont des polynômes en f que l'on sait calculer explicitement.

3.9 Proposition. *Soit P facteur irréductible de χ et/ou μ. On utilise les notations introduites ci-dessus.*
(i) *L'entier a_P est caractérisé par le fait que P^{a_P} est le polynôme minimal de l'endomorphisme induit f_{C_P} sur le sous-espace caractéristique C_P.*
(ii) *L'entier b_P est* $\dim(C_P)/\deg(P)$.

Preuve. On considère pour chaque facteur irréductible P de μ l'endomorphisme $f_{C_P} \in \mathcal{L}(C_P)$ induit sur le sous-espace caractéristique associé à P. Par définition de C_P, le polynôme P^{a_P} annule f_{C_P}. Puisque P est irréductible, on en déduit qu'il existe un entier naturel $a'_P \leqslant a_P$ tel que $\mu_{f_{C_P}} = P^{a'_P}$.

Puisque la décomposition en sous-espaces caractéristiques est une décomposition en sous-espaces stables, on a par la proposition 2.12

$$\begin{aligned}\mu_f &= \mathrm{ppcm}(\mu_{f_{C_P}}, P \in \mathrm{Irr}) \\ &= \prod\nolimits_{P\in\mathrm{Irr}} P^{a'_P}.\end{aligned}$$

Par unicité de la décomposition en produit de facteurs irréductibles (c'est la factorialité de $\mathbf{k}[X]$), on a donc $a'_P = a_P$ pour tout $P \in \mathrm{Irr}$, ce qui prouve (i).

Pour chaque $P \in \mathrm{Irr}$ on a $\mu_{f_{C_P}} = P^{a'_P}$, donc il existe $b'_P \in \mathbf{N}$ tel que $\chi_{f_{C_P}} = P^{b'_P}$ par la proposition 3.7, partie (ii). En regardant le degré, on voit que $b'_P = \dim(C_P)/\deg(C_P)$. Enfin, à nouveau grâce à la décomposition en sous-espaces caractéristiques, on a $\chi_f = \prod_{P\in\mathrm{Irr}} \chi_{f_{C_P}} = \prod_{P\in\mathrm{Irr}} P^{b'_P}$, donc à nouveau par unicité de la décomposition en produit de facteurs irréductibles, on a $b'_P = b_P$, ce qui prouve (ii). □

3.3 – Simplicité et semi-simplicité

Nous revenons ici sur les notions de simplicité et semi-simplicité (et b-semi-simplicité) introduites et étudiées en section 1.3. Nous allons notamment donner des caractérisations de la simplicité et semi-simplicité en termes des polynômes caractéristique et minimal.

Pour mémoire, nous avons convenu en section 1.3 qu'un endomorphisme est semi-simple s'il est somme directe d'endomorphismes simples, et b-semi-simple s'il possède la propriété que tout sous-espace stable possède un supplémentaire stable. Nous avons montré que la semi-simplicité et la b-semi-simplicité sont deux propriétés équivalentes, nous allons le redémontrer ici.

3.10 Proposition. *Soit $f \in \mathcal{L}(E)$. Les propositions suivantes sont équivalentes :*
(i) *f simple ;*
(ii) *le polynôme caractérisique χ_f irréductible.*

Preuve. Supposons f simple, et montre que ceci impose à χ_f d'être irréductible. Soit $S, T \in \mathbf{k}[X]$ tels que $\chi_f = ST$. On a $S(f)T(f) = 0$, donc $S(f)$ ou $T(f)$ a un noyau non-nul. Disons que c'est S, et soit $x \in \ker(S(f))$ non nul. Par simplicité de f, F_x le plus petit sous-espace stable contenant x est E tout entier, donc $\deg(\mu_x) = n$ (on utilise les notations de la section 3.1). Ceci implique $\deg S \geqslant n$, et donc $T \in \mathbf{k}$. Ceci prouve que χ est irréductible.

Si au contraire f possède un sous-espace stable non-trivial F, alors χ_{f_F} est diviseur strict de χ_f, puisque $\deg(\chi_{f_F}) < \deg(\chi_f)$, donc χ n'est pas irréductible. □

3.11. En particulier, la proposition 3.10 nous dit que si $\mathbf{k}$ est algébriquement clos, alors les seuls endomorphismes simples sont les homothéties d'une droite. (En effet, si $\mathbf{k}$ est algébriquement clos les seuls polynômes irréductibles sont les polynômes de degré 1). Dans ce cas, les endomorphismes semi-simples, qui par définition sont sommes directes d'endomorphismes simples, sont nécessairement diagonalisables.

Remarquons au passage que Le caractère simple dépend du choix du corps de base. Pour l'illustrer, considérons l'endomorphisme associé à la matrice

$$R_\theta = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

($\theta \in \mathbf{R}$ non congru à 0 modulo π). L'endomorphisme de $\mathbf{R}^2$ associé à cette matrice est simple (voir exemple 1.20, mais l'endomorphisme de $\mathbf{C}^2$ associé à la même matrice n'est pas simple (il a deux sous-espaces propres de dimension 1).

3.12 Proposition. *Soit $f \in \mathcal{L}(E)$. Les propositions suivantes sont équivalentes :*
(i) *f est b-semi-simple ;*
(ii) *aucun carré non constant ne divise le polynôme minimal μ_f.*

Preuve. Montrons "(i) ⇒ (ii)". Soit T facteur irréductible de μ_f. Il s'agit de montrer que T^2 ne divise pas μ_f. Pour cela, considérons $C_T^0 := \ker(T(f))$.[7] C'est un sous-espace stable par f. Si f est b-semi-simple, C_T^0 possède un supplémentaire stable F. L'endomorphisme $T(f)_F \in \mathcal{L}(F)$ est injectif puisque

$$\ker\big(T(f)_F\big) = \ker\big(T(f)\big) \cap F = \{0\} \ ;$$

il est donc inversible pour raisons de dimension. On en déduit que T et μ_{f_F} sont premiers entre eux, et donc T ne divise pas μ_{f_F}. Puisque $T \cdot \mu_{f_F}$ annule f, μ_f divise $T \cdot \mu_{f_F}$, et donc T^2 ne divise pas μ_f, comme il fallait démontrer.

Montrons que réciproquement, "(ii) ⇒ (i)". Écrivons $\mu_f = T_1 \cdots T_r$, où les T_i sont irréductibles deux à deux non proportionnels, et considérons la décomposition en sous-espaces caractéristiques $E = C_1 \oplus \cdots \oplus C_r$ et les projecteurs spectraux $p_1, \ldots, p_r$ (pour tout $i = 1, \ldots, r$, on note $C_i = \ker(T_i(f))$). Soit F un sous-espace stable par f. Puisque les p_i sont des polynômes en f d'après le lemme des noyaux, chacun laisse F stable. On en déduit la décomposition

$$F = (F \cap C_1) \oplus \cdots \oplus (F \cap C_r). \tag{3.12.1}$$

On va montrer que chaque $F \cap C_i$ possède un supplémentaire G_i dans C_i stable par f, ou de manière équivalente par f_{C_i}. Alors $G := \bigoplus G_i$ sera un supplémentaire de F dans E stable par f, et on aura gagné.

Pour montrer l'existence d'un supplémentaire stable à $F \cap C_i$ dans C_i, on utilise un argument un peu baroque. Soit $\mathbf{k}_i := \mathbf{k}[X]/(T_i)$; puisque T_i est irréductible, $\mathbf{k}_i$ est un corps. L'opération de composition externe

$$(\bar{A}, x) \in \mathbf{k}_i \times C_i \mapsto \bar{A}.x := A(f)(x)$$

7. attention, en général ce n'est pas le sous-espace caractéristique associé à T ; justement ça l'est si et seulement si T^2 ne divise pas μ.

munit le $\mathbf{k}$-espace vectoriel C_i d'une structure de $\mathbf{k}_i$-espace vectoriel. On observe que les $\mathbf{k}_i$-sous-espaces vectoriels de C_i correspondent ensemblistement aux $\mathbf{k}$-sous-espaces vectoriels de C_i qui sont stables par f. Ainsi $F \cap C_i$ est un $\mathbf{k}_i$-sev de C_i. On lui choisit un $\mathbf{k}_i$-supplémentaire : c'est un $\mathbf{k}$-sev de C_i supplémentaire à $F \cap C_i$ qui est stable par f, et nous avons donc trouvé notre G_i. □

3.13 Remarque. Il est toujours vrai, sans condition sur f, que tout sous-espace stable F se décompose selon les sous-espaces caractéristiques comme en (3.12.1).

On peut le voir en appliquant le même argument que dans la preuve ci-dessus, ou bien de la manière suivante. On considère $\mu_f = \prod P_i^{a_i}$ la décomposition du polynôme minimal en produit de facteurs irréductibles. Le polynôme μ_f est annulateur aussi pour l'endomorphisme induit $f_F \in \mathcal{L}(F)$, et donc on obtient par le lemme des noyaux

$$F = \bigoplus_i \ker\big(P_i^{a_i}(f_F)\big),$$

qui est exactement la décomposition cherchée puisque $\ker(P_i^{a_i}(f_F)) = F \cap \ker(P_i^{a_i}(f))$.

En revanche, il est grossièrement faux que pour une décomposition arbitraire $E = \bigoplus H_i$ on a $F = \bigoplus(F \cap H_i)$. Nous l'avons déjà observé en 2.29.1.

3.14 Corollaire. *Soit $f \in \mathcal{L}(E)$ b-semi-simple. Alors pour tout sous-espace F stable par f, l'endomorphisme induit $f_F \in \mathcal{L}(F)$ est b-semi-simple.*

Preuve. Puisque f est semi-simple, μ_f est sans facteur carré. Comme μ_{f_F} divise μ_f, il est lui aussi sans facteur carré, et donc f_F est lui-même b-semi-simple. □

3.15 Proposition. *Soit $f \in \mathcal{L}(E)$. Les propositions suivantes sont équivalentes :*
(i) *f est semi-simple ;*
(ii) *f est b-semi-simple.*

Preuve. Supposons pour commencer que f est semi-simple. Alors par définition il existe une décomposition $E = \bigoplus_{i=1}^r F_i$ en sous-espaces stables telle que pour tout $i = 1, \dots, r$, l'endomorphisme induit $f_{F_i} \in \mathcal{L}(F_i)$ est simple. D'après la proposition 3.10, le polynôme caractéristique $\chi_{f_{F_i}}$ est un polynôme irréductible P_i. On a donc $\mu_{f_{F_i}} = P_i$ d'après le théorème de Cayley–Hamilton. Pour conclure,

$$\mu_f = \operatorname{ppcm}(\mu_{f_{F_1}}, \dots, \mu_{f_{F_r}}) = \operatorname{ppcm}(P_1, \dots, P_r)$$

est sans facteur carré puisque $P_1, \dots, P_r$ sont irréductibles. D'après la proposition 3.12, ceci prouve que f est b-semi-simple comme il fallait démontrer.

Réciproquement, nous allons démontrer par récurrence sur $n = \dim(E) \geqslant 0$ que tout endomorphisme de E b-semi-simple est semi-simple. Si $n = 0$, le résultat est trivial. Supposons donc $n \geqslant 1$ et le résultat démontré pour tout $n' < n$. Soit $f \in \mathcal{L}(E)$ b-semi-simple. Si f est simple, il n'y a rien à démontrer. Sinon il existe F sous-espace stable par f tel que

$$0 < \dim(F) < \dim(E). \tag{3.15.1}$$

Puisque f est b-semi-simple, il existe un supplémentaire F' de F stable par f. Les inégalités (3.15.1) entraînent des inégalités identiques pour $\dim(F')$, et on peut donc appliquer l'hypothèse de récurrence à $f_F \in \mathcal{L}(F)$ et $f_{F'} \in \mathcal{L}(F')$ qui sont tous les deux b-semi-simples d'après le corollaire 3.14. On en déduit qu'il existe des décompositions $F = \bigoplus_{i=1}^r F_i$ et $F' = \bigoplus_{i=1}^s F_i'$ en sous-espaces stables par f_F et $f_{F'}$ repectivement, telles que les f_{F_i} et $f_{F_i'}$ sont simples. Finalement,

$$E = \big(\bigoplus\nolimits_{i=1}^r F_i\big) \oplus \big(\bigoplus\nolimits_{i=1}^s F_i'\big)$$

est une décomposition de E en somme directe de sous-espaces stables par f tels que les endomorphismes induits sur chaque terme de la somme sont tous simples. Ceci prouve bien que f est semi-simple. □

3.16. Les endomorphismes qui ne sont archétypiquement pas semi-simples sont les nilpotents (non nuls). Pour ceux-là, $X^2|\mu_f$ (sauf si $f=0$), et le noyau est un sous-espace stable sans supplémentaire stable (voir la preuve de la proposition 2.42.

A contrario l'archétype de l'endomorphisme semi-simple est l'endomorphisme diagonalisable. D'ailleurs, la preuve donnée ci-dessus du fait que si μ est sans facteur carré alors f est semi-simple est parallèle à celle du fait que les diagonalisables sont semi-simples ; la différence est que dans le cas diagonalisable, on a $\mathbf{k}_i \cong \mathbf{k}$ et donc l'argument baroque est invisible (mais bien là tapis dans l'ombre).

Nous avons évoqué à plusieurs reprises le fait que les endomorphismes semi-simples sont essentiellement ceux qu'il est possible de diagonaliser quitte à étendre les scalaires. Pour que ce soit tout-à-fait vrai, il faut faire une hypothèse sur le corps de base $\mathbf{k}$.

3.17 Rappel de théorie des corps. Un corps $\mathbf{k}$ est *parfait* si tout polynôme irréductible de $\mathbf{k}[X]$ est scindé à racines simples dans une extension de décomposition.

Pour illustrer la notion, le mieux est sans doute de donner un exemple de corps qui n'est pas parfait (le lecteur doute peut-être du fait qu'une telle horreur puisse exister). Soit $\mathbf{k}$ un corps de caractéristique $p>0$ tel qu'il existe $\alpha \in \mathbf{k}$ qui n'est pas une puissance p-ième : on écrit alors $\alpha \notin \mathbf{k}^p$.[8] Nous allons voir que le polynôme $X^p-\alpha \in \mathbf{k}[X]$ est irréductible, mais que c'est une puissance p-ième dans toute extension de décomposition. Considérons une extension $\mathbf{k}'$ de $\mathbf{k}$ dans laquelle α possède une racine p-ième $\beta \in \mathbf{k}'$. Alors

$$X^p-\alpha = X^p-\beta^p = (X-\beta)^p$$

dans $\mathbf{k}'[X]$.

Reste à voir que $X^p-\alpha$ est irréductible dans $\mathbf{k}[X]$. Dans $\mathbf{k}'[X]$, $X^p-\alpha$ a un unique facteur irréductible, $X-\beta$. Soit $P \in \mathbf{k}[X]$ facteur irréductible de $X^p-\alpha$. Il existe un entier $q \in [\![1,p]\!]$ tel que, dans $\mathbf{k}'[X]$,

$$P = (X-\beta)^q = X^q - q\beta X^{q-1} + \cdots .$$

Puisque $\beta \notin \mathbf{k}$ et $P \in \mathbf{k}[X]$, on a nécessairement $q\beta = 0$, et donc $q=p$. Ainsi l'unique facteur irréductible de $X^p-\alpha$ dans $\mathbf{k}[X]$ est $X^p-\alpha$ lui-même, comme il fallait démontrer.

Un exemple explicite réalisant la situation ci-dessus est le suivant. On prend $\mathbf{k} = \mathbf{F}_p(T)$, le corps des fractions rationnelles à coefficients dans $\mathbf{F}_p = \mathbf{Z}/p\mathbf{Z}$, p premier, et $\alpha = T$.

On retiendra (nous ne le démontrerons pas ici) que les corps parfaits sont exactement les corps de caractéristique 0, et les corps de caractéristique $p>0$ tels que $\mathbf{k}^p = \mathbf{k}$. Cette dernière catégorie inclut tous les corps finis.

3.18 Proposition. *Soit E un $\mathbf{k}$-espace vectoriel de dimension finie, et $f \in \mathcal{L}(E)$.*
(i) Si $\mathbf{k}$ est algébriquement clos, alors f semi-simple $\Leftrightarrow$ f diagonalisable.
(ii) Si $\mathbf{k}$ est parfait, alors $M \in \mathcal{M}_n(\mathbf{k})$ est semi-simple si et seulement si elle est diagonalisable dans une extension finie de $\mathbf{k}$.

Preuve. D'après les propositions 3.12 et 3.15, f est semi-simple si et seulement si son polynôme minimal μ_f est sans facteur constant. Si $\mathbf{k}$ est algébriquement clos, les irréductibles de $\mathbf{k}[X]$ sont les polynômes de degré 1, et cette condition équivaut à ce que μ_f soit scindé à racines simples, et donc à ce que f soit diagonalisable. Ceci prouve (i).

Montrons (ii). S'il existe $\mathbf{k}'$ extension de $\mathbf{k}$ telle que le polynôme minimal de M est scindé à racines simples sur $\mathbf{k}'$, alors μ_M est sans facteur carré dans $\mathbf{k}'[X]$ et donc *a fortiori* dans $\mathbf{k}[X]$. On en déduit que M est semi-simple.

Réciproquement, si $M \in \mathcal{M}_n(\mathbf{k})$ est semi-simple, alors il existe $P_1,\dots,P_r \in \mathbf{k}[X]$ irréductibles deux à deux non proportionnels tels que $\mu_M = P_1\cdots P_r$. Considérons une extension $\mathbf{k}'$ de $\mathbf{k}$ dans laquelle $P_1,\dots,P_r$ sont tous décomposés. Puisque $\mathbf{k}$ est un corps parfait, $P_1,\dots,P_r$ sont tous

8. attention à ne pas confondre avec le produit cartésien p-fois de $\mathbf{k}$, que nous notons de la même façon...

scindés à racines simples sur $\mathbf{k}'$. Comme de plus ils sont deux à deux premiers entre eux, le produit $P_1 \cdots P_r$ est lui-même scindé à racines simples sur $\mathbf{k}'$. On en déduit que M est diagonalisable dans $\mathcal{M}_n(\mathbf{k}')$, comme il fallait démontrer. □

3.19 Corollaire. *Si M est une matrice définie sur un corps parfait, alors le caractère semi-simple ou non est invariant par extension des scalaires.*

Preuve. D'après la proposition précédente, la matrice M est semi-simple si et seulement si elle est diagonalisable sur une clôture algébrique de $\mathbf{k}$. Cette condition ne change pas quand on passe à une extension $\mathbf{k}'$ de $\mathbf{k}$. □

3.20 Exemple. Si en revanche $\mathbf{k}$ n'est pas parfait, l'énoncé du corollaire ci-desssus est faux en général. Considérons par exemple la matrice compagnon

$$\mathrm{Comp}(X^p - T) \in \mathcal{M}_p(\mathbf{F}_p(T))$$

pour p premier. Son polynôme caractéristique est $X^p - T$ qui est irréductible (voir 3.17 ci-dessus), donc cette matrice est simple, et *a fortiori* semi-simple. En revanche, son polynôme minimal est $(X - S)^p$ dans l'extension $\mathbf{k}' = \mathbf{k}[S]/(S^p - T)$ de $\mathbf{k}$, et sur $\mathbf{k}'$ cette matrice n'est pas semi-simple, et encore moins diagonalisable.

3.4 – Structure de l'algèbre $\mathbf{k}[f]$

3.21 L'algèbre $\mathbf{k}[f]$. On a déjà vu en 2.5 qu'elle est isomorphe à $\mathbf{k}[X]/(\mu_f)$, donc de dimension finie $\deg \mu_f$ sur $\mathbf{k}$.

Parmi ses « propriétés globales », outre sa dimension il faut citer son commutant dans $\mathcal{L}(E)$, i.e. l'ensemble des endomorphismes de E qui commutent avec f. Ceci s'identifie naturellement à l'espace des endomorphismes de E comme $\mathbf{k}[X]$-module :

$$\mathrm{Com}_{\mathcal{L}(E)}(f) \cong \mathrm{End}_{\mathbf{k}[X]}\big(M_{E,f}\big)$$

(source : [**?**, Vol. 2, II.2]).

3.22. La première chose à dire c'est que, notant $\mu_f = \prod T_i^{a_i}$ la décomposition en produit de facteurs irréductible, on a par le lemme chinois

$$\mathbf{k}[f] \cong \bigoplus_i \mathbf{k}[X]/(T_i^{a_i}), \tag{3.22.1}$$

qui présente un parallèle évident avec la décomposition en sous-espaces caractéristiques.

Sans expliciter beaucoup plus cette décomposition, on peut déjà prouver les deux énoncés suivants.

3.23 Semi-simplicité et nilpotents. *L'endomorphisme f est semi-simple ssi l'algèbre $\mathbf{k}[f]$ n'a pas d'élément nilpotent non-nul.*

Preuve. Le point clef est que semi-simple $\Leftrightarrow$ aucun carré non constant ne divise μ. Ainsi, vu la décomposition (3.22.1), il s'agit de se convaincre que $\mathbf{k}[X]/\big(T^a\big)$, T irréductible, possède des nilpotents non-triviaux ssi $a > 1$.

Si $a > 1$, $\bar{T}$ est un nilpotent non-nul. Si $a = 1$, soit $\bar{F}$ un nilpotent : il existe $s \geqslant 1$ tel que $\bar{F}^s = 0$, i.e. $T|F^s$. Comme T irréductible, par le lemme de Gauss ceci implique $T|F$, i.e. $\bar{F} = 0$. □

3.24 Idempotents et projecteurs sur les sous-espaces caractéristiques. *Les idempotents de l'algèbre* $\mathbf{k}[f]$ *sont les (sommes de) projecteurs sur les sous-espaces caractéristiques.*

Autrement dit, pour tout $g \in \mathbf{k}[f]$ *tel que* $g^2 = g$*, il existe* $I \subseteq [\![1, r]\!]$ *tel que* $g = \sum_{i\in I} p_i$*, où* $p_1, \dots, p_r$ *sont les projecteurs sur les sous-espaces caractéristiques de* f*.*

Preuve. Soit $g \in \mathbf{k}[f]$ idempotent. Puisque $g^2 = g$, g est un projecteur, et puisque g est un polynôme en f, il laisse stable chacun des sous-espaces caractéristiques de f. Nous allons démontrer que pour chaque T facteur irréductible de μ_f, l'endomorphisme induit g_{C_T} est 0 ou id_{C_T}, ce qui démontrera le résultat annoncé.

Ainsi, il suffit de démontrer que pour f de polynôme minimal $\mu_f = T^a$, T irréductible, tout $g \in \mathbf{k}[f]$ idempotent est 0 ou id. Il existe $A \in \mathbf{k}[X]$ tel que $g = A(f)$. Si A est premier à T, alors il est premier à $T^a = \mu_f$, donc $g = A(f)$ est un isomorphisme. Puisque c'est un projecteur, on a nécessairement $g = \mathrm{id}$. Si *a contrario* T divise A, alors μ_f divise A^a et donc $g^a = 0$. Puisque g est idempotent, $g^a = g$, et donc $g = 0$ dans ce cas. □

Il est bon néanmoins d'explorer plus en profondeur les relations entre la décomposition de $\mathbf{k}[f]$ donnée par l'isomorphisme chinois et la décomposition de E donnée par le lemme des noyaux.

3.25 Isomorphisme chinois. Soit $P_1, \dots, P_r$ des polynômes deux à deux premiers entre eux. L'application

$$\varphi : (A \bmod P_1 \cdots P_r) \in \frac{\mathbf{k}[X]}{(P_1 \cdots P_r)} \longmapsto (A \bmod P_i)_{1\leqslant i\leqslant r} \in \prod\nolimits_{1\leqslant i\leqslant r} \frac{\mathbf{k}[X]}{(P_i)}$$

est un isomorphisme de $\mathbf{k}$-algèbres.

Posons $Q_i = \prod_{j\neq i} P_j$ pour tout $i = 1, \dots, r$. Les polynômes Q_i sont premiers entre eux dans leur ensemble comme on l'a vu lors de la preuve du lemme des noyaux. Ils sont donc liés par une relation de Bezout : il existe $U_1, \dots, U_r \in \mathbf{k}[X]$ tels que

$$U_1Q_1 + \cdots + U_rQ_r = 1.$$

L'application

$$\psi : (A_i \bmod P_i)_{1\leqslant i\leqslant r} \longmapsto U_1Q_1A_1 + \cdots + U_rQ_rA_r \bmod P_1 \cdots P_r$$

est le morphisme de $\mathbf{k}$-algèbres réciproque de φ.

Preuve. On laisse au lecteur le soin de vérifier que l'application φ est bien définie, et est un morphisme de $\mathbf{k}$-algèbres. Montrons que ce morphisme est injectif : Soit $\bar{A} \in \ker(\varphi)$. Pour tout $i = 1, \dots, r$, $A \bmod P_i$ est nul, autrement dit A est divisible par P_i. Puisque les P_i sont deux à deux premiers entre eux, A est donc divisible par $P_1, \dots, P_r$, autrement dit $\bar{A} = 0$. Ceci suffit pour conclure que φ est un isomorphisme, puisque les deux algèbres $\mathbf{k}[X]/(P_1 \cdots P_r)$ et $\prod_{1\leqslant i\leqslant r}\big(\mathbf{k}[X]/(P_i)\big)$ sont de dimensions égales

$$\deg(P_1 \cdots P_r) = \deg(P_1) + \cdots + \deg(P_r).$$

Le fait que ψ soit le morphisme réciproque de φ prouvera explicitement la surjectivité de φ.

à propos de l'application ψ, commençons par montrer qu'elle est bien définie : soit A_i et B_i deux polynômes congrus modulo P_i. Alors il existe $K_i \in \mathbf{k}[X]$ tel que $B_i - A_i = K_iP_i$, et donc

$$U_iQ_iB_i - U_iQ_iA_i = K_iP_iQ_i = K_iP_1 \cdots P_r,$$

autrement dit $U_iQ_iA_i$ et $U_iQ_iB_i$ sont congrus modulo $P_1 \cdots P_r$ comme on voulait démontrer. Ensuite montrons que ψ est un morphisme de $\mathbf{k}$-algèbres : toutes les vérifications sont élémentaires

(donc laissées au lecteur), sauf peut-être celle de la multiplicativité, qui fonctionne comme suit. Soit $A_1, B_1, \dots, A_r, B_r \in \mathbf{k}[X]$. On a [9]

$$\begin{aligned}\psi\big((A_1,\dots,A_r)\cdot(B_1,\dots,B_r)\big) &= \psi(A_1B_1,\dots,A_rB_r)\\ &= U_1Q_1A_1B_1+\cdots+U_rQ_rA_rB_r\end{aligned}$$

et

$$\psi(A_1,\dots,A_r)\cdot\psi(B_1,\dots,B_r)=\sum_{1\leqslant i,j\leqslant r} U_iQ_iA_i\cdot U_jQ_jB_j.$$

Si $i\neq j$, Q_iQ_j est nul modulo $P_1\cdots P_r$, donc en fait

$$\psi(A_1,\dots,A_r)\cdot\psi(B_1,\dots,B_r)=\sum_{1\leqslant i\leqslant r}(U_iQ_i)^2A_iB_i.$$

Enfin, pour tout $i=1,\dots,r$, en multipliant la relation de Bezout par U_iQ_i il vient

$$U_iQ_i=\sum_{1\leqslant j\leqslant n}U_iQ_iU_jQ_j=(U_iQ_i)^2,$$

où la dernière égalité est donnée à nouveau par le fait que Q_iQ_j est divisible par $P_1\cdots P_r$ si $i\neq j$. Finalement on a donc bien

$$\begin{aligned}\psi(A_1,\dots,A_r)\cdot\psi(B_1,\dots,B_r)&=\sum_{1\leqslant i\leqslant r}(U_iQ_i)^2A_iB_i=\sum_{1\leqslant i\leqslant r}U_iQ_iA_iB_i\\ &=\psi\big((A_1,\dots,A_r)\cdot(B_1,\dots,B_r)\big).\end{aligned}$$

Il reste à voir que les morphismes φ et ψ sont réciproques l'un de l'autre. Pour $A\in\mathbf{k}[X]$ on a

$$\psi\circ\varphi(A)=(U_1Q_1+\cdots+U_rQ_r)A=A,$$

et pour $A_1,\dots,A_r\in\mathbf{k}[X]$,

$$\begin{aligned}\varphi\circ\psi(A_1,\dots,A_r)&=(U_1Q_1A_1+\cdots U_rQ_rA_r \bmod P_i)_{1\leqslant i\leqslant r}\\ &=(U_iQ_iA_i \bmod P_i)_{1\leqslant i\leqslant r}\\ &=(A_i \bmod P_i)_{1\leqslant i\leqslant r},\end{aligned}$$

où la dernière égalité est donnée par le fait qu'en réduisant l'identité de Bezout modulo P_i, il vient $\overline{U_iQ_i}=\overline{1}$. □

3.26 Projecteurs spectraux. On constate que les projecteurs spectraux sont les éléments de $\mathbf{k}[f]$ qui dans la décomposition (3.22.1) s'écrivent $(0,\dots,1,\dots,0)$. Pour cela, on relit l'expression des projecteurs spectraux comme polynômes en f fournie par la preuve du lemme des noyaux 2.18, puis celle du morphisme réciproque ψ ci-dessus.

On en déduit une autre preuve de la caractéristation des idempotents de $\mathbf{k}[f]$.

Seconde preuve de 3.24. Il s'agit de montrer que pour $T\in\mathbf{k}[X]$ irréductible,les idempotents de $\mathbf{k}[X]/(T^a)$ sont 0 et 1. $\bar{A}$ est idempotent ssi

$$\begin{aligned}A^2\equiv A \bmod T^a &\Leftrightarrow T^a|A(A-1)\\ &\Leftrightarrow T^a|A \text{ ou } T^a|(A-1)\end{aligned}$$

car T irréductible et A et $A-1$ premiers entre eux. □

9. pour alléger les notations, on confond polynômes et classes de congruences, c'est dans notre intérêt à tous

3.5 – Endomorphismes cycliques

3.27 Définition. On dit qu'un endomorphisme $f \in \mathcal{L}(E)$ est *cyclique* s'il existe un vecteur $x \in E$ tel que le plus petit sous-espace stable par f contenant x est E tout entier.

On rappelle (voir section 3.1) que le plus petit sous-espace stable par f contenant x s'appelle le sous-espace cyclique associé à x, et nous le notons F_x.

D'après les résultats de la section 3.1, f est cyclique si et seulement si il existe une base de E dans laquelle f est donné par une matrice compagnon ; cette matrice compagnon est nécessairement associée au polynôme caractéristique χ_f.

Nous allons donner une condition nécessaire et suffisante pour qu'un endomorphisme soit cyclique, mais avant cela examinons à la main le cas des endomorphismes nilpotents et diagonalisables.

3.28 Exemple. Soit $f \in \mathcal{L}(E)$ nilpotent. Pour tout vecteur $x \in E$, soit i le plus petit entier tel que

$$x \in K_i(f) = \ker(f^i).$$

Alors $\mu_x = X^i$, et le sous-espace cyclique F_x a dimension i. On en déduit que f est cyclique si et seulement si son indice de nilpotence est égal à la dimension de E.

Le fait que $\mu_x = X^i$ résulte des arguments donnés dans la preuve du Théorème 2.38 : par minimalité de i, $\overline{x} \in K_i/K_{i-1}$ est non-nulle, donc $\overline{f_i(\overline{x})} = \overline{f(x)} \in K_{i-1}/K_{i-2}$ est non-nulle si $i > 1$. On montre ainsi par récurrence que $\overline{x}, \overline{f(x)}, \ldots, \overline{f^{i-1}(x)}$ sont tous non-nuls dans $K_i/K_{i-1}, K_{i-1}/K_{i-2}, \ldots, K_1/K_0$ respectivement. On en déduit par le lemme 1.34 que la famille $x, f(x), \ldots, f^{i-1}(x)$ est libre, et donc $\deg \mu_x \geqslant i$. Puisque $f^i(x) = 0$, on en déduit $\mu_x = X^i$ comme on voulait.

3.29 Exemple. Soit $f \in \mathcal{L}(E)$ diagonalisable. On note $\lambda_1, \ldots, \lambda_r$ ses valeurs propres deux à deux distinctes, et $E_1, \ldots, E_r$ les sous-espaces propres correspondants. Pour tout vecteur $x \in E$, on a la décomposition

$$x = x_1 + \cdots + x_r$$

selon la décomposition $E = \bigoplus_{i=1}^r E_i$. Alors $F_x = \mathrm{Vect}(x_1, \ldots, x_r)$ a dimension au plus r (égale au nombre de x_i non-nuls). On en déduit que f est cyclique si et seulement si tous ses sous-espaces propres sont de dimension 1, autrement dit si son polynôme caractéristique est scindé à racines simples.

Pour montrer que $F_x = \mathrm{Vect}(x_1, \ldots, x_r)$, nous proposons deux méthodes. Dans les deux cas, on commence par remarquer que $\mathrm{Vect}(x_1, \ldots, x_r)$ est un sous-espace stable contenant x, donc on a l'inclusion $F_x \subseteq \mathrm{Vect}(x_1, \ldots, x_r)$. La première méthode pour montrer l'inclusion inverse est de se souvenir que d'après la Proposition 2.29, tout sous-espace F stable par f est de la forme

$$F = F_1 \oplus \cdots \oplus F_r,$$

où chaque F_i est un sous-espace arbitraire de l'espace propre E_i. On en déduit que tout sous-espace stable par f contenant $\mathrm{Vect}(x_1, \ldots, x_r)$ est de la forme $\mathrm{Vect}(x_{i_1}, \ldots, x_{i_s})$, avec $1 \leqslant i_1 < \cdots < i_s \leqslant r$. Si un tel sous-espace est contenu strictement dans $\mathrm{Vect}(x_1, \ldots, x_r)$, c'est qu'il existe $i_0 \in [\![1, r]\!] \setminus \{i_1, \ldots, i_s\}$ tel que $x_{i_0} \neq 0$, et dans ce cas $\mathrm{Vect}(x_{i_1}, \ldots, x_{i_s}) \subseteq \bigoplus_{i \neq i_0} E_i$ ne peut pas contenir x, puisque $x \notin \bigoplus_{i \neq i_0} E_i$ car $x_{i_0} \neq 0$. On en déduit que le plus petit sous-espace stable par f contenant x est $\mathrm{Vect}(x_1, \ldots, x_r)$ comme il fallait démontrer.

L'autre façon de faire consiste à démontrer par un calcul explicite que

$$\mathrm{Vect}\big(x, f(x), \ldots, f^{r-1}(x)\big) = \mathrm{Vect}(x_1, \ldots, x_r),$$

ce qui implique $\mathrm{Vect}(x_1, \ldots, x_r) \subseteq F_x$ et permet donc de conclure. Le calcul est le suivant :

$$\begin{aligned} x &= x_1 + \cdots + x_r \\ f(x) &= \lambda_1 x_1 + \cdots + \lambda_r x_r \\ &\vdots \\ f^{r-1}(x) &= \lambda_1^{r-1} x_1 + \cdots + \lambda_r^{r-1} x_r. \end{aligned}$$

Il implique aussitôt l'inclusion $\mathrm{Vect}\big(x, f(x), \ldots, f^{r-1}(x)\big) \supseteq \mathrm{Vect}(x_1, \ldots, x_r)$, et l'inclusion inverse vient du fait qu'on peut inverser le système ci-dessus, la matrice de Vandermonde $V(\lambda_1, \ldots, \lambda_r)$ étant inversible dans la mesure où les valeurs propres $\lambda_1, \ldots, \lambda_r$ sont deux à deux disjointes.

3.30 Proposition. *Pour tout $f \in \mathcal{L}(E)$, il existe un vecteur $x \in E$ tel que $\mu_x = \mu_f$.*

Preuve. On écrit la décomposition en produit de facteurs premiers $\mu_f = P_1^{a_1} \cdots P_r^{a_r}$ du polynôme minimal, et on regarde la décomposition en sous-espaces caractéristiques

$$E = C_1 \oplus \cdots \oplus C_r.$$

Pour chaque $i = 1, \ldots, r$, il existe un $x_i \in C_i$ tel que $\mu_{x_i} = P_i^{a_i}$. En effet, pour tout $x \in C_i$, $\mu_x | P_i^{a_i}$ car $P_i^{a_i}$ annule l'endomorphisme $f_{C_i} \in \mathcal{L}(C_i)$ induit par f, donc il existe $a(x)$ tel que $\mu_x = P_i^{a(x)}$. Si pour tout $x \in C_i$ on a $a(x) < a_i$, alors $P_i^{a_i - 1}(f_{C_i}) = 0$, et ceci contredit le fait que $P_i^{a_i}$ est le polynôme minimal de f_{C_i} (voir Proposition 3.9).

Ensuite $x = x_1 + \cdots + x_r$ convient car pour tout $Q \in \mathbf{k}[X]$:

$$Q(f)(x) = Q(f)(x_1) + \cdots + Q(f)(x_r) = 0 \Leftrightarrow Q(f)(x_1) = \cdots = Q(f)(x_r) = 0,$$

puisque $Q(f)(x_i) \in C_i$ pour tout i, et les sous-espaces caractéristiques sont en somme directe. On a donc

$$\begin{aligned} Q(f)(x) = 0 &\iff P_1^{a_1}, \ldots, P_r^{a_r} \text{ divisent } Q \\ &\iff P_1^{a_1} \cdots P_r^{a_r} | Q, \end{aligned}$$

autrement dit $\mu_x = P_1^{a_1} \cdots P_r^{a_r} = \mu_f$ comme on voulait. □

3.31 Proposition. *Soit $f \in \mathcal{L}(E)$. Les deux propositions suivantes sont équivalentes :*
(i) *f cyclique ;*
(ii) *$\chi_f = \mu_f$.*

Preuve. (i) ⇒ (ii). Cette implication est une traduction du fait qu'une matrice compagnon $\mathrm{Comp}(P)$ a polynômes caractéristique et minimal tous les deux égaux à P, voir le lemme 3.6.
(ii) ⇒ (i). D'après la proposition 3.30, il existe x tel que $\mu_x = \mu_f$; on a donc $\mu_x = \chi_f$ puisque $\chi_f = \mu_f$. Du coup, $\dim F_x = \deg \mu_x = \deg \chi_f = \dim E$, et donc $E = F_x$. □

On retrouve bien avec ce critère à quelle condition un endomorphisme nilpotent ou diagonalisable est cyclique, voir les exemples 3.28 et 3.29.

3.32. Le fait qu'un endomorphisme à polynôme caractéristique scindé à racines simples (en particulier, diagonalisable) soit cyclique nous dit que les deux matrices ci-dessous sont semblables si et seulement si $\lambda_1, \ldots, \lambda_n$ sont deux à deux distincts, ce qui ne semble pas évident *a priori*.

$$\begin{pmatrix} \lambda_1 & & & \\ & \ddots & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix} \quad \text{et} \quad \begin{pmatrix} 0 & & & (-1)^{n-1}\sigma_n \\ 1 & \ddots & & \\ & \ddots & 0 & -\sigma_2 \\ & & 1 & \sigma_1 \end{pmatrix},$$

où $\sigma_1, \ldots, \sigma_n$ sont les fonctions symétriques élémentaires en $\lambda_1, \ldots, \lambda_n$. Plus généralement, on obtient le résultat suivant en appliquant les Propositions 3.30 et/ou 3.31.

3.33. Soit $Q_1, \dots, Q_r \in \mathbf{k}[X]$. Les deux matrices

$$\mathrm{diag}(\mathrm{Comp}(Q_1), \dots, \mathrm{Comp}(Q_r)) \quad \text{et} \quad \mathrm{Comp}(Q_1 \cdots Q_r)$$

(la première est une matrice diagonale par blocs, dont les blocs diagonaux sont les matrices compagnons $\mathrm{Comp}(Q_1), \dots, \mathrm{Comp}(Q_r)$) sont semblables si et seulement si $Q_1, \dots, Q_r$ sont deux à deux premiers entre eux. Un cas typique où cette condition est vérifiée est si $Q_i = P_i^{a_i}$ pour tout $i = 1, \dots, r$, où $P_1, \dots, P_r$ sont des polynômes irréductibles deux à deux non proportionnels, et $a_1, \dots, a_r \in \mathbf{N}$.

Pour prouver notre affirmation, commençons par observer qu'une matrice est semblable à la matrice compagnon $\mathrm{Comp}(Q_1 \cdots Q_r)$ si et seulement si c'est la matrice d'un endomorphisme cyclique de polynôme minimal $Q_1 \cdots Q_r$. Or la matrice de gauche est la matrice d'un endomorphisme f de $\mathbf{k}^N$, $N = \sum_{i=1}^r \deg(Q_i)$, pour lequel il existe une décomposition en sous-espaces stables $k^N = \bigoplus_{i=1}^r E_i$ telle que pour tout i, l'endomorphisme induit f_{E_i} a polynômes minimal et caractéristique tous les deux égaux à Q_i. On a donc

$$\chi_f = Q_1 \cdots Q_r \quad \text{et} \quad \mu_f = \mathrm{ppcm}(Q_1, \dots, Q_r),$$

et les deux matrices $\mathrm{diag}(\mathrm{Comp}(Q_1), \dots, \mathrm{Comp}(Q_r))$ et $\mathrm{Comp}(Q_1 \cdots Q_r)$ sont semblables si et seulement si

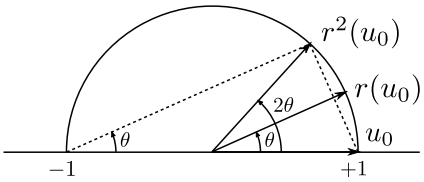
$$\mathrm{ppcm}(Q_1, \dots, Q_r) = Q_1 \cdots Q_r,$$

ce qui équivaut à ce que les polynômes $Q_1, \dots, Q_r$ soient deux à deux premiers entre eux.

3.34 Exemple. Illustrons 3.32 par un exemple permettant de faire un joli dessin. Pour tout $\theta \in \mathbf{R}$ non congru à 0 modulo π, les deux matrices

$$R_\theta = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \quad \text{et} \quad \begin{pmatrix} 0 & -1 \\ 1 & 2\cos\theta \end{pmatrix}$$

sont semblables. En effet, elles ont toutes les deux $\chi = \mu = X^2 - (2\cos\theta)X + 1$; une autre façon de le dire est que la matrice de gauche vue à coefficients dans $\mathbf{C}$ est diagonalisable de polynôme caractéristique $(X - e^{i\theta})(X - e^{-i\theta}) = -(2\cos\theta)X + 1$, donc redevable de 3.32. Géométriquement ces matrices sont les matrices de la rotation d'angle θ dans deux bases différentes, comme le montre le dessin ci-dessous (en toute rigueur, pour parler de rotation d'angle θ il faut au préalable munir le plan $\mathbf{R}^2$ d'une structure euclidienne).



Le triangle en pointillés est rectangle puisqu'il s'appuie sur un diamètre du cercle. L'angle de gauche de ce triangle est θ à cause de la relation entre angle au centre et angle inscrit, et l'hypothénuse a longueur 2. On a donc bien

$$r^2(u_0) = -u_0 + (2\cos\theta).r(u_0).$$

Si en revanche θ est congru à 0 modulo π, la matrice R_θ n'est pas cyclique (c'est une homothétie), et n'est donc pas semblable à une matrice compagnon.

Dans la situation "orthogonale" à celle de 3.32 (c'est-à-dire celle où le polynôme caractéristique au lieu d'être à racines simples est une puissance de $X - \lambda$), on a le fait suivant.

3.35. Pour tout $\lambda \in \mathbf{k}$, les matrices

$$\begin{pmatrix} \lambda & & & \\ 1 & \ddots & & \\ & \ddots & \lambda & \\ & & 1 & \lambda \end{pmatrix}, \quad \text{et} \quad \mathrm{Comp}\big((X-\lambda)^n\big) = \begin{pmatrix} 0 & & & (-1)^{n-1}\binom{n}{n}\lambda^n \\ 1 & \ddots & & \\ & \ddots & 0 & -\binom{n}{2}\lambda^2 \\ & & 1 & \binom{n}{1}\lambda \end{pmatrix}$$

sont semblables. En effet, la matrice de gauche a polynôme minimal $(X-\lambda)^n$ où n est la taille de la matrice, et pour raisons de degré c'est aussi son polynôme caractéristique. On en déduit que c'est une matrice cyclique, et ainsi qu'elle est semblable à $\mathrm{Comp}\big((X-\lambda)^n\big)$.

3.6 – Décomposition de Frobenius et invariants de similitude

Cette section est consacrée à la preuve et à l'étude des conséquences du résultat suivant.

3.36 Théorème (Décomposition de Frobenius)**.** *Pour tout $f \in \mathcal{L}(E)$, il existe une unique suite de polynômes unitaires non constants $P_1, \dots, P_r \in \mathbf{k}[X]$ satisfaisant aux deux conditions :*

(i) *il existe une décomposition $E = \bigoplus_{1 \leqslant i \leqslant r} F_i$ en sous-espaces stables par f, telle que pour tout $i = 1, \dots, r$ l'endomorphisme induit $f_{F_i} \in \mathcal{L}(F_i)$ est cyclique de polynôme caractéristique et minimal P_i ; et*

(ii) *$P_1 | P_2 | \cdots | P_r$.*

Nous allons démontrer à part l'existence et l'unicité de la décomposition de Frobenius, mais avant de procéder tâchons d'illustrer l'importance de ce résultat. Les points à retenir sont (i) que la suite de polynômes P_i caractérise la classe de similitude de f, et (ii) que ces polynômes peuvent se calculer algorithmiquement, en appliquant le pivot de Gauss à la matrice $X.\mathrm{Id} - A$ à coefficients dans l'anneau principal $\mathbf{k}[X]$, après avoir choisi une base pour écrire la matrice A de f.

Une décomposition $E = \bigoplus_{1 \leqslant i \leqslant r} F_i$ comme dans l'énoncé du théorème est appelée une *décomposition de Frobenius* de f.

3.37 Définition. Soit $f \in \mathcal{L}(E)$. Les polynômes $P_1, \dots, P_r$ associés à f comme dans le Théorème 3.36 sont appelés les *invariants de similitude* de f.

Cette terminologie est justifiée par l'énoncé suivant.

3.38 Corollaire. *Deux endomorphismes $f, g \in \mathcal{L}(E)$ sont semblables si et seulement si ils ont les mêmes invariants de similitude.*

Preuve. Notons $(P_1, \dots, P_r)$ et $(Q_1, \dots, Q_s)$ les invariants de similitude de f et g respectivement. Si f et g sont semblables, alors toute décomposition de Frobenius pour f est aussi une décomposition de Frobenius pour g, et donc nécessairement $(P_1, \dots, P_r) = (Q_1, \dots, Q_s)$.

Réciproquement, s'il existe deux décompositions $E = \bigoplus_{1 \leqslant i \leqslant r} F_i$ et $E = \bigoplus_{1 \leqslant i \leqslant r} G_i$ en sous-espaces F_i stables par f et G_i stables par g telles que f_{F_i} et g_{G_i} sont cycliques de polynôme caractéristique P_i, alors il existe pour tout $i = 1, \dots, r$ un isomorphisme $\varphi_i : G_i \cong F_i$ tel que $g_{G_i} = \varphi_i^{-1} \circ f_{F_i} \circ \varphi_i$, donc f et g sont semblables par le Lemme 1.18. □

La partie 'unicité' du Théorème 3.36 est une conséquence directe de la Proposition suivante, que nous démontrerons plus loin. Pour la notion de facteurs invariants d'une matrice à coefficients dans un anneau principal, nous renvoyons à la section A.

3.39 Proposition. *Soit $f \in \mathcal{L}(E)$, et A la matrice de f dans une base $\mathcal{B}$ de E. Les invariants de similitude de f sont les facteurs invariants de la matrice $X.\mathrm{Id} - A$ à coefficients dans l'anneau principal $\mathbf{k}[X]$ (auxquels il faut retirer les '1' initiaux).*

Autrement dit, les facteurs invariants de la matrice $X.\mathrm{Id} - A$ sont $(1, \dots, 1, P_1, \dots, P_r)$ où $(P_1, \dots, P_r)$ sont les invariants de similitude de f.

3.40. Une conséquence particulièrement intéressante est qu'on peut déterminer les invariants de similitude en opérant le pivot de Gauss (des matrices à coefficients dans $\mathbf{k}[X]$) à la matrice $X.\mathrm{Id}_n - A$, voir section A.

3.41 Exemple. Calculons les invariants de similitude de l'endomorphisme associé à la matrice

$$A = \begin{pmatrix} -3 & -2 & -2 \\ -2 & 0 & -1 \\ 10 & 5 & 6 \end{pmatrix}.$$

On calcule $X.\mathrm{Id} - A$, puis on échange les deux premières lignes, avant d'effectuer une permutation circulaire sur les colonnes pour mettre un 1 en haut à gauche :

$$X.\mathrm{Id} - A = \begin{pmatrix} X+3 & 2 & 2 \\ 2 & X & 1 \\ -10 & -5 & X-6 \end{pmatrix} \xrightarrow[\text{ii) } C_3 \to C_1 \to C_2 \to C_3]{\text{i) } L_1 \leftrightarrow L_2} \begin{pmatrix} 1 & 2 & X \\ 2 & X+3 & 2 \\ X-6 & -10 & -5 \end{pmatrix} ;$$

ensuite on remplace L_2 par $L_2 - 2L_1$ et L_3 par $L_3 - (X-6)L_1$, on obtient

$$\begin{pmatrix} 1 & 2 & X \\ 0 & X-1 & -2(X-1) \\ 0 & -2(X-1) & -(X-1)(X-5) \end{pmatrix} ;$$

enfin on remplace L_3 par $L_3 + 2L_2$, on obtient :

$$\begin{pmatrix} 1 & 2 & X \\ 0 & X-1 & -2(X-1) \\ 0 & 0 & -(X-1)^2 \end{pmatrix}.$$

On voit déjà apparaître les facteurs invariants de la matrice, il n'y a plus qu'à faire du nettoyage en opérant sur les colonnes pour arriver à une forme diagonale : on fait tout d'abord $C_2 \leftarrow C_2 - 2C_1$ et $C_3 \leftarrow C_3 - XC_1$, ce qui donne la première matrice ci-dessous, puis $C_3 \leftarrow -C_3 - 2C_2$, ce qui donne la seconde matrice ci-dessous.

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & X-1 & -2(X-1) \\ 0 & 0 & -(X-1)^2 \end{pmatrix} ; \qquad \begin{pmatrix} 1 & 0 & 0 \\ 0 & X-1 & 0 \\ 0 & 0 & (X-1)^2 \end{pmatrix}.$$

On en conclut que les invariants de similitude de la matrice A sont $X-1$ et $(X-1)^2$, donc A est semblable à la matrice

$$\begin{pmatrix} \mathrm{Comp}(X-1)) & \\ & \mathrm{Comp}\big((X-1)^2\big) \end{pmatrix} = \begin{pmatrix} 1 & & \\ & 0 & -1 \\ & 1 & 2 \end{pmatrix},$$

qui d'après 3.35 (voir aussi plus généralement 3.45 plus loin), est elle-même semblable à la matrice ci-dessous,

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

On laisse au lecteur le soin de vérifier qu'effectivement $\ker\big(A - \mathrm{Id}\big)$ et $\ker\big(A - \mathrm{Id}\big)^2\big)$ ont dimension 2 et 3 respectivement.

3.42 Remarque. Les invariants de similitude contiennent l'information des polynômes minimal et caractéristique. Dans les notations du Théorème 3.36, on a

$$\mu_f = P_r \quad \text{et} \quad \chi_f = P_1 \cdots P_r$$

En effet, toujours dans les notations du Théorème 3.36, on a pour chaque $i = 1, \dots, r$, $P_i = \chi_{f_{F_i}} = \mu_{f_{F_i}}$, et donc puisque tous les F_i sont stables,

$$\begin{aligned} \chi_f &= \chi_{f_{F_1}} \cdots \chi_{f_{F_r}} \qquad \text{et} \qquad \mu_f = \operatorname{ppcm}(\mu_{f_{F_1}}, \dots, \mu_{f_{F_1}}) \\ &= P_1 \cdots P_r \qquad\qquad\qquad\quad = \operatorname{ppcm}(P_1, \dots, P_r) = P_r. \end{aligned}$$

En particulier, on connaît désormais un algorithme pour calculer le polynôme minimal (en fait on en connaissait déjà un : calculer les puissances successives de A jusqu'à trouver le plus petit entier k tel que $\mathrm{Id}, A, \dots, A^k$ soient liés comme vecteurs de $\mathcal{M}_n(\mathbf{k})$, puis chercher une relation de dépendance linéaire entre ces matrices ; l'algorithme du pivot de Gauss est nettement plus efficace).

3.43 Corollaire. *Les invariants de similitude sont invariants par extension des scalaires.*

En particulier, puisque le polynôme minimal est l'un des invariants de similitude, il est invariant par extension des scalaires. Nous renvoyons à l'exemple 3.45.1 plus loin pour une illustration instructive de ce résultat.

Preuve. Les invariants de similitude se calculent par l'algorithme de Gauss, qui ne fait pas sortir du corps de départ. Précisément, si on part d'une matrice $A \in \mathcal{M}_n(\mathbf{k})$, avec donc $X.\mathrm{Id} - A$ à coefficients dans $\mathbf{k}[X]$, étendre les scalaires revient à considérer A à coefficients dans $\mathbf{k}'$, une extension de corps de $\mathbf{k}$, et ainsi $X.\mathrm{Id} - A$ à coefficients dans $\mathbf{k}'[X]$. Mais si on opère le pivot de Gauss sur $X.\mathrm{Id} - A \in \mathcal{M}_n(\mathbf{k}'[X])$, on fait les mêmes opérations que si l'on opérait le pivot de Gauss sur $X.\mathrm{Id} - A \in \mathcal{M}_n(\mathbf{k}[X])$, et donc le résultat final est le même ! □

Une autre façon de prouver le corollaire est de dire directement que les facteurs invariants sont eux-mêmes invariants par extension des scalaires. Dans notre contexte, cela s'écrit de la manière suivante. Si $X.\mathrm{Id} - A$ est équivalente à $\operatorname{diag}(1, \dots, 1, P_1, \dots, P_r)$ dans $\mathcal{M}_n(\mathbf{k}[X])$, alors *a fortiori* ces deux matrices sont équivalentes dans $\mathcal{M}_n(\mathbf{k}'[X])$, donc les facteurs invariants de $X.\mathrm{Id} - A$ vue dans $\mathcal{M}_n(\mathbf{k}'[X])$ sont bien $1, \dots, 1, P_1, \dots, P_r$.

3.44 Remarque. Dans le cas des endomorphismes trigonalisables, nous avions déjà des invariants suffisants pour distinguer les classes de similitude (Théorème 2.52) et calculables algorithmiquement, ainsi qu'une forme normale représentant chacune de ces classes (Remarque 2.49). Ce que l'on gagne dans ce cas, c'est donc une forme un peu plus compacte pour organiser les invariants, et un algorithme lui aussi plus compact pour les calculer.

Justement, voyons comment déterminer les invariants de similitude d'un endomorphisme trigonalisable à partir de sa forme normale de Jordan.

3.45 Exemple. On considère la matrice diagonale par blocs

$$A = \operatorname{diag}\big(J_{a_{1,1}}(\lambda_1), \dots, J_{a_{1,k_1}}(\lambda_1),\ \dots\ , J_{a_{r,1}}(\lambda_r), \dots, J_{a_{1,k_r}}(\lambda_r)\big),$$

avec r valeurs propres $\lambda_1, \dots, \lambda_r$ deux à deux distinctes, et pour tout $i = 1, \dots, r$, k_i blocs de Jordan associés à la valeur propre λ_i, de tailles respectives $a_{i,1} \leqslant \dots \leqslant a_{i,k_i}$.

Quitte à rajouter formellement des blocs de taille 0, on peut supposer qu'il y a autant de blocs pour chaque valeur propre, autrement dit $k_1 = \dots = k_r$; nous noterons k ce nombre. En permutant les blocs, on obtient la matrice ci-dessous, semblable à A :

$$A' = \operatorname{diag}\big(J_{a_{1,1}}(\lambda_1), \dots, J_{a_{r,1}}(\lambda_r),\ \dots\ , J_{a_{1,k}}(\lambda_1), \dots, J_{a_{r,k}}(\lambda_r)\big),$$

D'après 3.35, le bloc de Jordan $J_a(\lambda)$ est semblable à la matrice compagnon de $(X - \lambda)^a$; pour tout $s = 1, \dots, k$, les deux matrices ci-dessous sont donc semblables,

$$\operatorname{diag}\big(J_{a_{1,s}}(\lambda_1), \dots, J_{a_{r,s}}(\lambda_r)\big) \quad \text{et} \quad A_s = \operatorname{diag}\big(\operatorname{Comp}\big((X - \lambda_1)^{a_{1,s}}\big), \dots, \operatorname{Comp}\big((X - \lambda_r)^{a_{r,s}}\big)\big).$$

Puisque $\lambda_1, \ldots, \lambda_r$ sont deux à deux distinctes, les polynômes $(X-\lambda_1)^{a_{1,s}}, \ldots, (X-\lambda_r)^{a_{r,s}}$ sont deux à deux premiers entre eux, donc d'après 3.33 la matrice A_s est cyclique associée au polynôme $\prod_{i=1}^r (X-\lambda_i)^{a_{i,s}}$ pour tout $s = 1, \ldots, k$.

En conclusion, la matrice A est semblable à la matrice diagonale par blocs dont les blocs diagonaux sont les matrices compagnons associées aux polynômes $\prod_{i=1}^r (X-\lambda_i)^{a_{i,s}}$ pour $s = 1, \ldots, k$. Puisque pour tout $i = 1, \ldots, r$, on a choisi $a_{i,1} \leqslant \ldots \leqslant a_{i,k_i}$, ces polynômes vérifient les relations de divisibilité requises de sorte que

$$\prod_{i=1}^r (X-\lambda_i)^{a_{i,1}}, \ \ldots \ , \ \prod_{i=1}^r (X-\lambda_i)^{a_{i,k}}$$

sont les invariants de similitude de la matrice A.

3.45.1. Le cas particulier ci-dessous de 3.45 est intéressant du point de vue de l'invariance par extension des scalaires. On considère la matrice

$$R = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \in \mathcal{M}_2(\mathbf{R}).$$

Dans $\mathcal{M}_2(\mathbf{C})$, cette matrice est diagonalisable, semblable à la matrice diagonale $\mathrm{diag}(e^{i\theta}, e^{-i\theta})$. D'après ce qui précède, elle a un seul invariant de similitude,

$$(X - e^{i\theta})(X - e^{-i\theta})$$

qui est bien un polynôme à coefficients réels.

3.46 Exercice. Montrer qu'un endomorphisme trigonalisable est cyclique si et seulement si tous ses sous-espaces propres sont des droites.

3.47 Exercice. 1) Soit $\alpha_0 \in \mathbf{C}$. Donner un représentant de chacune des classes de similitude dans $\mathcal{M}_4(\mathbf{C})$ constituées de matrices ayant α comme unique valeur propre.

2) Pour chaque classe de similitude comme dans la question précédente, représentée par une matrice A, calculer :
a) les polynômes caractéristique et minimal χ_A et μ_A ;
b) pour tout $\alpha \in \mathbf{C}$ et $i \in \mathbf{N}$, la dimension du noyau de $(\alpha \mathrm{Id}_4 - A)^i$;
c) les invariants de similitude de A.

3) On suppose à présent $\alpha_0 \in \mathbf{R}$. Déterminer deux matrices $A', A'' \in \mathcal{M}_4(\mathbf{R})$ qui ne sont pas semblables, mais qui sont telles que pour $A = A', A''$:

$$\text{(i)} \ \dim\big(\ker(\alpha_0 \mathrm{Id}_4 - A)^i\big) = \begin{cases} i & \text{si } i \leqslant 1 \\ 2 & \text{sinon} \end{cases} \ ; \qquad \text{et (ii)} \ \dim\big(\ker(\alpha \mathrm{Id}_4 - A)^i\big) = 0$$

pour tout $\alpha \in \mathbf{R} - \{\alpha_0\}$ et tout $i \in \mathbf{N}$.
(*Indication :* nous remercions nos aimables lecteurs de chercher A' et A'' non-trigonalisables, pour éviter de trouver un contre-exemple à notre théorème 2.52).

Existence d'une décomposition de Frobenius

Le point difficile est le lemme 3.49. Une fois ce point acquis, la preuve est une récurrence sans grand mystère. La preuve du lemme 3.49 fait intervenir la dualité très élégamment.

3.48 Preuve de la partie existence du théorème 3.36. On raisonne par récurrence sur la dimension de E. Si $\dim(E) \leqslant 1$, le résultat est trivial. Considérons donc E de dimension au moins 2, et supposons l'existence démontrée en dimension plus petite.

D'après la proposition 3.30 il existe $x \in E$ tel que $\mu_x = \mu_f$. Soit F_x le sous-espace cyclique associé à x. D'après le lemme 3.49 ci-dessous, il existe un supplémentaire G_x de F_x dans E qui est stable par f. On a $\dim(G_x) < \dim(E)$, donc par hypothèse de récurrence appliquée à l'endomorphisme induit $f_{G_x} \in \mathcal{L}(G_x)$ il existe des polynômes $P_1 | \cdots | P_r$ et une décomposition $G_x = \bigoplus_{i=1}^r G_i$ telle que chaque f_{G_i} soit cyclique de polynôme P_i. On a alors $E = F_x \oplus \bigoplus_{i=1}^r G_i$, et f_{F_x} est cyclique de polynôme $\mu_x = \mu_f$. On a $P_r = \mu_{f_{G_x}}$, voir remarque 3.42, donc $P_r | \mu_f$ (c'est la proposition 2.12). Ainsi $(P_1, \ldots, P_r, \mu_f)$ est une suite de polynômes comme on voulait. □

3.49 Lemme. *Soit $x \in E$ tel que $\mu_x = \mu_f$. Alors le sous-espace cyclique associé à x possède un supplémentaire dans E stable par f.*

3.49.1 Attention. En général il est faux que le sous-espace cyclique associé à x quelconque possède un supplémentaire stable.

Par exemple, si f est nilpotent cyclique, c'est-à-dire nilpotent d'indice $n = \dim(E)$, pour $x \in \ker(f^i) \setminus \ker(f^{i-1})$, $i < n$, le sous-espace cyclique F_x n'a pas de supplémentaire stable. En effet, on a $0 < \dim(F_x) < n$, donc si on avait une décomposition $E = F_x \oplus G$ avec G stable lui aussi, l'indice de nilpotence serait inférieur à $\max(\dim(F_x), \dim(G))$, donc strictement inférieur à n. Dans ce cas, $\mu_x = X^i$ et $\mu_f = X^n$.

Preuve. Notons p le degré de $\mu_f = \mu_x$. Nous allons définir explicitement un supplémentaire stable à F_x par p équations linéaires indépendantes. La famille $(x, f(x), \ldots, f^{p-1}(x))$ est une base de F_x. En particulier c'est une famille libre, et nous pouvons donc licitement considérer une forme linéaire $\ell \in E^\star$ telle que $\ell(f^i(x)) = 0$ si $i < p-1$ et $\ell(f^{p-1}(x)) = 1$.

Montrons que les formes linéaires $\ell \circ f^i$, $i = 0, \ldots, p-1$ sont linéairement indépendantes. Il suffit de vérifier que leurs restrictions à F_x le sont. Or la matrice de $(\ell, \ldots, \ell \circ f^{p-1}) \in \mathcal{L}(F_x, \mathbf{k}^p)$ dans la base $(x, f(x), \ldots, f^{p-1}(x))$ est de la forme

$$\begin{pmatrix} 0 & \ldots & 0 & 1 \\ \vdots & \iddots & 1 & * \\ 0 & \iddots & \iddots & \vdots \\ 1 & * & \ldots & * \end{pmatrix}, \tag{3.49.1}$$

donc elle est inversible, et ainsi les restrictions des $\ell \circ f^i$ $(i = 0, \ldots, p-1)$ à F_x sont linéairement indépendantes comme il fallait.

On définit $G = (\ell, \ldots, \ell \circ f^{p-1})^\perp$. Puisque $\ell, \ldots, \ell \circ f^{p-1}$ sont linéairement indépendantes, c'est un sous-espace de dimension $n-p$ de E. Montrons que c'est un supplémentaire de F_x. Il suffit de montrer que $F_x \cap G = \{0\}$. Or un vecteur $z \in F_x$ est dans $(\ell, \ldots, \ell \circ f^{p-1})^\perp$ si et seulement si ses coordonnées dans la base $(x, f(x), \ldots, f^{p-1}(x))$ sont dans le noyau de la matrice (3.49.1). Celle-ci étant inversible, on a bien $F_x \cap G = \{0\}$.

Il reste à montrer que G est stable par f. Soit $z \in G$. On a

$$\ell(z) = \ell(f(z)) = \cdots = \ell(f^{p-1}(z)) = 0.$$

Puisque $\deg(\mu_f) = p$, $f^p(z)$ est une combinaison linéaire de $z, f(z), \ldots, f^{p-1}(z))$. Ceci implique que $\ell(f^p(z))$ est une combinaison linéaire de $\ell(z), \ell(f(z)), \ldots, \ell(f^{p-1}(z)))$, et donc que $\ell(f^p(z)) = 0$. Finalement on a donc

$$\begin{aligned} \ell(f(z)) = \cdots = \ell(f^{p-1}(z)) = \ell(f^p(z)) = 0 \\ \iff \quad \ell(f(z)) = \cdots = \ell(f^{p-2}(f(z))) = \ell(f^{p-1}(f(z))) = 0 \end{aligned}$$

et ainsi $f(z) \in G$ comme il fallait démontrer. □

3.50 Remarque. On peut comprendre la preuve ci-dessus un peu plus conceptuellement. En fait, le $\ell \in E^\star$ qu'on considère est tel que $\mu_\ell = \mu_{f^\intercal}$: le polynôme minimal local de $f^\intercal \in \mathcal{L}(E^\star)$ égale le polynôme minimal de $f^\intercal$. En effet f et $f^\intercal$ ont même polynôme minimal, [10] de degré p, et on montre dans la preuve que $\ell, \ell \circ f = f^\intercal(\ell), \dots, \ell \circ f^{p-1} = (f^\intercal)^{p-1}(\ell)$ sont linéairement indépendantes, donc $\deg \mu_\ell \geqslant p$, et finalement $\mu_\ell = \mu_{f^\intercal}$.

Alors

$$\mathrm{Vect}(\ell, \dots, \ell \circ f^{p-1}) = \mathrm{Vect}\big(\ell, f^\intercal(\ell), \dots, ({}^\intercal f)^{p-1}(\ell)\big) = F_\ell$$

est le sous-espace cyclique de $f^\intercal$ associé à ℓ. Il est donc automatiquement stable par f, et de dimension p.

Enfin ℓ a été choisi de sorte qu'en plus $F_x \subseteq E$ et $F_\ell \subseteq E^\star$ sont naturellement en dualité, au sens où par le morphisme de restriction $\varphi \in E^\star \mapsto \varphi|_F \in F_x^\star$, F_ℓ est isomorphe à $F_x^\star$. Si on préfère (mais j'en doute), F_ℓ est un supplémentaire de $F_x^\perp$, donc il est isomorphe au quotient $E^\star/F_x^\perp$, qui lui est canoniquement isomorphe à $F_x^\star$ (voir ref).

3.50.1 Attention. La famille des restrictions de $\ell, \dots, \ell \circ f^{p-1}$ à F_x n'est en général pas la base duale de $(f^{p-1}(x), \dots, f(x), x)$ (et encore moins celle de $(x, f(x), \dots, f^{p-1}(x))$). à titre d'exercice, on pourra écrire explicitement la matrice (3.49.1) en fonction des coefficients du polynôme minimal $\mu_x = \mu_f$.

Unicité de la décomposition de Frobenius

Comme nous l'avons annoncé plus haut, la partie unicité du théorème 3.36 est une conséquence directe de l'identification des invariants de similitude de f aux facteurs invariants non triviaux de la matrice $X.\mathrm{Id} - A \in \mathcal{M}_n(\mathbf{k}[X])$, où A est la matrice de f dans une base arbitraire (c'est la proposition 3.39). Nous allons donc ici démontrer cette proposition. Avant de le faire, nous attirons l'attention du lecteur sur une subtilité de l'énoncé du théorème 3.36.

3.51 Mise en garde. Si les invariants de similitude de f sont bien uniques, en général la décomposition en somme de sous-espaces cycliques elle ne l'est pas.

Par exemple, si f est une homothétie n'importe quelle décomposition de E en somme directe de droites est une décomposition de Frobenius de f.

Un autre exemple : si f est nilpotent avec k blocs de Jordan tous de la même taille a, de manière équivalente si f est nilpotent d'indice a tel que $\dim(\ker f^i) - \dim(\ker f^{i-1}) = k$ pour tout $i = 1, \dots, a$, alors pour toute famille de vecteurs $x_1, \dots, x_k$ telle que $(\bar{x}_1, \dots, \bar{x}_k)$ est une base de $E/\ker(f^{a-1})$, les sous-espaces cycliques $F_{x_1}, \dots, F_{x_k}$ fournissent une décomposition de Frobenius de f.

Nous prouvons la proposition 3.39 par une mise en oeuvre explicite du pivot de Gauss sur $X.\mathrm{Id} - A$ pour une matrice A sous forme décomposée de Frobenius.

3.52 Preuve de la proposition 3.39. Soit $P_1, \dots, P_r$ des polynômes comme dans le théorème 3.36. Alors il existe une base de E dans laquelle la matrice de f est diagonale par blocs

$$A = \mathrm{diag}(\mathrm{Comp}(P_1), \dots, \mathrm{Comp}(P_r)) \in \mathcal{M}_n(\mathbf{k}).$$

D'après le lemme 3.53 ci-dessous, pour chaque $i = 1, \dots, r$ la matrice $X.\mathrm{Id}_{n_i} - \mathrm{Comp}(P_i)$ est équivalente dans $\mathcal{M}_{n_i}(\mathbf{k}[X])$, $n_i = \deg(P_i)$, à la matrice diagonale $\mathrm{diag}(1, \dots, 1, P_i)$. On en déduit que la matrice $X.\mathrm{Id}_n - A$ est équivalente dans $\mathcal{M}_n(\mathbf{k}[X])$ à la matrice diagonale

$$\mathrm{diag}(1, \dots, \dots, 1, P_1, \dots, P_r) \in \mathcal{M}_n(\mathbf{k}[X]).$$

□

10. on peut s'en convaincre directement à titre d'exercice ; sinon les matrices de f et $f^\intercal$ dans des bases duales l'une de l'autre sont transposées l'une de l'autre, donc elles ont même polynôme minimal.

3.53 Lemme. *Soit $P \in \mathbf{k}[X]$ unitaire non constant. Les facteurs invariants de la matrice $X.\mathrm{Id} - \mathrm{Comp}(P)$ sont $(1, \dots, 1, P)$.*

Preuve. On écrit $P = X^n + a_{n-1}X + \cdots + a_0$. On effectue les opérations élémentaires suivantes sur les lignes de la matrice

$$X.\mathrm{Id} - \mathrm{Comp}(P) = \begin{pmatrix} X & & & & a_0 \\ -1 & \ddots & & & a_1 \\ & \ddots & \ddots & & \vdots \\ & & -1 & X & a_{n-2} \\ & & & -1 & X + a_{n-1} \end{pmatrix},$$

qu'on appelle $L_1, \dots, L_n$: on effectue la permutation circulaire $L_n \to L_{n-1} \to \cdots \to L_1 \to L_n$, on obtient

$$\begin{pmatrix} -1 & X & & & a_1 \\ & \ddots & \ddots & & \vdots \\ & & -1 & X & a_{n-2} \\ & & & -1 & X + a_{n-1} \\ X & & & & a_0 \end{pmatrix},$$

puis on remplace L_n par

$$L_n + XL_1 + X^2L_2 + \cdots + X^{n-1}L_{n-1},$$

ce qui donne

$$\begin{pmatrix} -1 & X & & & a_1 \\ & \ddots & \ddots & & \vdots \\ & & -1 & X & a_{n-2} \\ & & & -1 & X + a_{n-1} \\ & & & & P(X) \end{pmatrix}.$$

Il ne reste alors plus qu'à faire du nettoyage automatique en opérant sur les colonnes pour arriver à la matrice diagonale $\mathrm{diag}(1, \dots, 1, P(X))$.

Précisément, on remplace C_2 par $C_2 + XC_1$ et C_n par $C_n + a_1C_1$ pour mettre des 0 sur la première ligne, puis C_3 par $C_3 + XC_2$ et C_n par $C_n + a_2C_2$ pour mettre des 0 sur la seconde ligne, et ainsi de suite jusqu'à remplacer C_n par $C_n + (X + a_{n-1})C_{n-1}$ pour mettre un 0 sur l'avant-dernière ligne. On a alors obtenu la matrice $\mathrm{diag}(-1, \dots, -1, P(X))$ qui convient à notre bonheur, qu'on peut transformer en $\mathrm{diag}(1, \dots, 1, P(X))$ en multipliant les $n-1$ premières lignes par -1. □

3.7 – Interprétation en termes de $\mathbf{k}[X]$-modules

3.54 Modules. Définition d'un module. Ce qui change par rapport aux espaces vectoriels : il n'est pas toujours possible de résoudre les systèmes d'équations linéaires, en particulier ciao la théorie de la dimension. Dans la catégorie des espaces vectoriels, (i) E de dimension n possède des sev de toutes les dimensions $\leqslant n$, (ii) F sev possède toujours un supplémentaire, et (iii) toute suite exacte est scindée. Dans la catégorie des modules ces trois énoncés sont faux (on verra qu'ils le sont aussi dans la catégorie des groupes).

3.55 $\mathbf{k}[X]$-module associé à un endomorphisme. Le $\mathbf{k}[X]$-module $M_{E,f}$. Ceci revient à considérer que la donnée de $f \in \mathcal{L}(E)$ induit une représentation de l'algèbre $\mathbf{k}[X]$. La recherche de décompositions de ces représentations est exactement la théorie de la réduction des endomorphismes.

3.56 Proposition. *Les $\mathbf{k}[X]$-modules isomorphes à $M_{E,f}$ sont exactement les $M_{E',f'}$ pour lesquels qu'il existe $\varphi : E \cong E'$ tel que $f = \varphi^{-1} \circ f' \circ \varphi$..*

Preuve. Soit $\varphi : M_{E,f} \cong M$ isomorphisme de $\mathbf{k}[X]$-modules. Je définis $E' := M$ muni de la structure de $\mathbf{k}$-ev sous-jacente, et $f' := m_X \in \mathcal{L}(E')$. La relation $f = \varphi^{-1} \circ f' \circ \varphi$ est offerte par le fait que φ est un isomorphisme de $\mathbf{k}[X]$-modules, donc commute à la multiplication par $X \in \mathbf{k}[X]$. □

3.57 Sous-modules. Les sous modules de $M_{E,f}$ sont les M_{F,f_F}, F stable par f. Ceci dicte les définitions :
(i) f simple s'il n'a pas de sev stable non trivial ;
(ii) f semi-simple si tout stable possède un supplémentaire stable.

Invariants de similitude : on les a définis plus haut, et on sait qu'on peut les trouver de la manière suivante : on prend une base, on fabrique $X \cdot \mathrm{Id}_n - M$, et on la met sous forme réduite.

3.58 Théorème. *Deux matrices sont semblables ssi elles ont les mêmes invariants de similitude.*

3.58.1. *Le $\mathbf{k}[X]$-module $M_{E,f}$ est isomorphe au quotient $E[X]/\mathrm{im}(X \cdot \mathrm{id} - f)$* via

$$\pi : \sum e_i X^i \mapsto \sum f^i(e_i).$$

Preuve. on vérifie que $\pi\big((X\mathrm{id} - f)(\sum e_i X^i)\big) = 0$, et réciproquement si $\sum e_i X^i \in \ker \pi$ alors

$$\begin{aligned}\sum e_i X^i &= \sum e_i X^i - 0 \\ &= \sum e_i X^i - \sum f^i(e_i) \\ &= \sum (X^i \mathrm{id} - f^i)(e_i),\end{aligned}$$

et chaque $X^i\mathrm{id} - f^i = (X\mathrm{id})^i - f^i$ se factorise par $X\mathrm{id} - f$. □

3.58.2 Conclusion.

$$\frac{E[X]}{\mathrm{im}(X \cdot \mathrm{id} - f)} \cong \frac{\mathbf{k}[X]}{(P_1)} \oplus \cdots \oplus \frac{\mathbf{k}[X]}{(P_n)}.$$

□

3.59 Remarque. On a démontré que les invariants de similitude peuvent être définis à partir du $\mathbf{k}[X]$-module $M_{E,f}$, en utilisant le fait qu'il est de type fini.

3.60 Décomposition de Frobenius. Chaque $E_i = \mathbf{k}[X]/(P_i)$ est un sev stable par f et $f_i := f_{E_i}$ est cyclique avec $\chi_{f_i} = \mu_{f_i} = P_i$. La décomposition

$$M_{E,f} = E_1 \oplus \cdots \oplus E_n$$

est une somme directe de sous-espaces cycliques pour f.

On en déduit $\chi_f = \chi_{f_1} \cdots \chi_{f_n} = P_1 \cdots P_n$, et $\mu_f = P_n$. Ceci dévoile le Théorème de Cayley–Hamilton (je ne dirais pas que ça le trivialise...), y compris sa version améliorée "χ et μ ont les mêmes facteurs irréductibles".

3.60.1 Exercice. f cyclique $\Leftrightarrow$ tous les P_i sauf le dernier sont égaux à 1.
[En particulier, on fait remarquer qu'en général il y a un certain nombre d'invariants égaux à 1, et que ceux-ci donnent des facteurs triviaux dans la décomposition cyclique.]

3.61 Décomposition de Dunford–Jordan d'un bloc cyclique.

$$\frac{\mathbf{k}[X]}{\big((X-\lambda_1)^{a_1}\cdots(X-\lambda_r)^{a_r}\big)} \;\cong\; \frac{\mathbf{k}[X]}{(X-\lambda_1)^{a_1}} \oplus \cdots \oplus \frac{\mathbf{k}[X]}{(X-\lambda_r)^{a_r}}$$

par le lemme chinois ; pour chaque morceau de la décomposition de droite, la multiplication par X s'écrit

$$\begin{pmatrix} \lambda & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & \lambda \end{pmatrix} \tag{3.61.1}$$

dans la base $\big((X-\lambda)^{a-1},\dots,X-\lambda,1\big)$, puisque

$$X\cdot(X-\lambda)^i = (X-\lambda)^{i+1} + \lambda(X-\lambda)^i.$$

(3.61.1) est un bloc de Dunford–Jordan de taille a.

3.62 Exercice. Soit $P \in \mathbf{k}[X]$. On considère le $\mathbf{k}$-espace vectoriel $E_P = \mathbf{k}[X]/(P)$ (il se trouve que c'est aussi une $\mathbf{k}[X]$-algèbre), et l'endomorphisme $m_P \in \mathcal{L}(E_P)$ de multiplication par X :

$$\forall H \in \mathbf{k}[X], \quad m_P(\bar{H}) = \overline{X\cdot H}.$$

1) On note $n = \deg(P)$. Montrer que E_P est de dimension n, muni de la base canonique $(\bar{1},\bar{X},\dots,\bar{X}^{n-1})$.

2) Démontrer que $\mu_{m_P} = P$.

3) Écrire la matrice de m_P dans la base canonique. En conclure que le polynôme minimal de la matrice compagnon associée à P est P lui-même.

A – Facteurs invariants

A.1 Théorème (Théorème des facteurs invariants)**.** *Soit A un anneau principal, $M \in \mathcal{M}_{n,m}(A)$ ($n \leqslant m$ disons). La matrice M est équivalente dans $\mathcal{M}_{n,m}(A)$ à une matrice*

$$\begin{pmatrix} a_1 & & & 0 & \cdots & 0 \\ & \ddots & & \vdots & & \vdots \\ & & a_m & 0 & \cdots & 0 \end{pmatrix} \tag{A.1.1}$$

où $a_1|\cdots|a_m$. Les a_i sont uniquement déterminés à multiplication par un inversible de A près.