Algèbre linéaire - 2
année universitaire 2018-2019
Université de Lorraine

Pierre-Emmanuel Chaput

Ce polycopié est très largement inspiré du polycopié utilisé par mon prédecesseur J.F. Grosjean, que je remercie pour son travail de rédaction.

Table des matières

Chapitre 1

Déterminants

1.1 Permutations

1.1.1 Permutations : définitions

Définition 1.1 *Une* **permutation** *de l'ensemble* $F_n = \{1, ..., n\}$ *est une bijection de* F_n *sur* F_n.

Exemples de permutations: Soit $F_4 = \{1, 2, 3, 4\}$ et soit $\sigma \in \mathfrak{S}_4$ définie par : $\sigma(1) = 4$, $\sigma(2) = 3$, $\sigma(3) = 1$ et $\sigma(4) = 2$. On écrira alors :

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 3 & 1 & 2 \end{pmatrix}$$

On notera $\mathfrak{S}_n$ l'ensemble des permutations.

Remarque 1.2 *La donnée d'un élément* σ *de* $\mathfrak{S}_n$ *est définie par les données successives de* $\sigma(1) \in F_n$, $\sigma(2) \in F_n \setminus \{\sigma(1)\}, \cdots, \sigma(n) \in F_n \setminus \{\sigma(1), \cdots, \sigma(n-1)\}$.

On en déduit que card $(\mathfrak{S}_n) = n!$.

Exercice de cours 1.1 *Combien y a-t-il d'éléments dans* $\mathfrak{S}_3$ *? Donner la liste des éléments de* $\mathfrak{S}_3$.

Remarque 1.3 $\mathfrak{S}_n$ *muni de la loi de composition* $\circ$ *est un groupe. En effet* $\circ$ *est une loi interne. En effet si* $(\sigma, \sigma') \in \mathfrak{S}_n^2$*, alors* $\sigma \circ \sigma' \in \mathfrak{S}_n$ *car la composée de deux bijections est une bijection.*

La loi $\circ$ *est associative. En effet si* $(\sigma, \sigma', \sigma'') \in \mathfrak{S}_n^3$*, alors on a évidemment* $(\sigma \circ \sigma') \circ \sigma'' = \sigma \circ (\sigma' \circ \sigma'')$.

L'application identité notée id ($id(k) = k$ *pour tout* $k \in F_n$*) est une permutation de* F_n *et est l'élément neutre pour* $\circ$*. Pour tout* $\sigma \in \mathfrak{S}_n$ *on a* $\sigma \circ Id = Id \circ \sigma = \sigma$.

Tout élément σ *de* $\mathfrak{S}_n$ *admet un inverse qui n'est rien d'autre que la bijection réciproque* σ^{-1}.

Définition 1.4 $\mathfrak{S}_n$ *muni de la loi* $\circ$ *s'appelle le* **groupe des permutations** *ou* **groupe symétrique**.

Définition 1.5 *On suppose que $n \geqslant 2$. Pour tout $(i,j) \in F_n^2$ tel que $i < j$, on appelle* **transposition** *échangeant i et j et on note τ_{ij} la permutation de F_n définie par $\tau_{ij}(i) = j$, $\tau_{ij}(j) = i$ et $\tau_{ij}(k) = k$, pour tout $k \in F_n \setminus \{i,j\}$.*

On écrira aussi $\tau_{ij} = (i,j)$.

Exemple 1.6 $\tau = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 2 & 3 & 1 \end{pmatrix}$ *est une transposition et on a $\tau = (1,4)$.*

Nous admettons maintenant le théorème suivant qui nous sera utile par la suite.

Théorème 1.7 *Tout élément de $\mathfrak{S}_n$ est produit de transpositions.*

1.1.2 Signature d'une permutation

Définition 1.8 *Soient $\sigma \in \mathfrak{S}_n$ et $i,j \in \{1, \cdots, n\}$. On dit que le couple (i,j)* **présente une inversion pour** *σ (ou est une* **inversion de** *σ) si : $i < j$ et $\sigma(i) > \sigma(j)$. On note $\ell(\sigma)$ le nombre d'inversions de σ.*

Exemple 1.9 *La permutation Id ne possédant aucune inversion, on a $\ell(Id) = 0$.*

Exemple 1.10 *Considérons $\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 3 & 1 & 2 \end{pmatrix}$. Les couples $(1,2)$, $(1,3)$, $(1,4)$, $(2,3), (2,4)$, sont des inversions de σ. En revanche $(3,4)$ n'est pas une inversion. Donc $\ell(\sigma) = 5$.*

Soit $\tau = (i, i+1)$ une transposition élémentaire et soit $\sigma \in \mathfrak{S}_n$. Etudions $I(\sigma \circ \tau)$. Pour $k < l$, on a $(k,l) \in I(\sigma \circ \tau)$ ssi $\sigma(\tau(k)) > \sigma(\tau(l))$. Il y a deux cas de figure : si $(k,l) \neq (i, i+1)$ alors $\tau(k) < \tau(l)$. Alors $(k,l) \in I(\sigma \circ \tau)$ ssi $(\tau(k), \tau(l)) \in I(\sigma)$ ssi $(k,l) \in \tau(I(\sigma))$. Si $(k,l) = (i, i+1)$, alors $(k,l) \in I(\sigma \circ \tau)$ ssi $(i, i+1) \notin I(\sigma)$. On a donc montré le résultat suivant :

Exemple 1.11 *Soit $\tau = (i, i+1)$ une transposition élémentaire et soit $\sigma \in \mathfrak{S}_n$.*
- *Si $(i, i+1) \in I(\sigma)$ alors $I(\sigma \circ \tau) = \tau(I(\sigma)) \setminus \{(i+1, i)\}$. Dans ce cas $\ell(\sigma \circ \tau) = \ell(\sigma) - 1$.*
- *Si $(i, i+1) \notin I(\sigma)$ alors $I(\sigma \circ \tau) = \tau(I(\sigma)) \cup \{(i, i+1)\}$. Dans ce cas $\ell(\sigma \circ \tau) = \ell(\sigma) + 1$.*

Exercice de cours 1.2 *Montrer que pour tout $\sigma \in \mathfrak{S}_4$, on a $\ell(\sigma) \leq 6$. Existe-t-il une permutation $\sigma \in \mathfrak{S}_4$ telle que $\ell(\sigma) = 6$? Les trouver toutes.*

Exercice de cours 1.3 *Soit σ la permutation de l'exemple 1.10. Vérifier le Théorème 1.7 pour σ en écrivant σ comme produit de cinq transpositions de la forme $(i, i+1)$ avec $i \in \{1,2,3\}$.*

On pourra utiliser le fait que pour toute permutation σ', si $\sigma'(i) > \sigma'(i+1)$, alors $\ell(\sigma' \circ (i, i+1)) = \ell(\sigma') - 1$, et trouver ainsi des entiers i_1, i_2, i_3, i_4, i_5 tels que $\ell(\sigma \circ (i_1, i_1+1)) = \ell(\sigma) - 1$, $\ell(\sigma \circ (i_1, i_1+1) \circ (i_2, i_2+1)) = \ell(\sigma \circ (i_1, i_1+1)) - 1$, et ainsi de suite. On observera alors que $\sigma \circ (i_1, i_1+1) \circ (i_2, i_2+1) \circ (i_3, i_3+1) \circ (i_4, i_4+1) \circ (i_5, i_5+1)$ est de longueur 0, donc est l'élément neutre.

Définition 1.12 *On appelle* **signature** *de $\sigma \in \mathfrak{S}_n$, le nombre $\varepsilon(\sigma) = (-1)^{\ell(\sigma)}$.*

Remarque 1.13 $\varepsilon(\sigma) \in \{-1, 1\}$.

Définition 1.14 *Soit $\sigma \in \mathfrak{S}_n$. On dit que σ est* **paire** *si $\varepsilon(\sigma) = 1$ et σ est* **impaire** *si $\varepsilon(\sigma) = -1$.*

Théorème 1.15 *Soient $\sigma, \sigma' \in \mathfrak{S}_n$. Alors :*

1. $\varepsilon(\sigma) = \displaystyle\prod_{1 \leqslant i < j \leqslant n} \frac{\sigma(j) - \sigma(i)}{j - i}$.
2. $\varepsilon(\sigma \circ \sigma') = \varepsilon(\sigma)\varepsilon(\sigma')$.
3. $\varepsilon(\sigma^{-1}) = \varepsilon(\sigma)$.

Preuve:

1. Remarquons d'abord que :

$$\begin{array}{rccc} \Phi & : & F_n^2 \setminus \{(i,i)\ ;i \in F_n\} & \longrightarrow & F_n^2 \setminus \{(i,i)\ ;i \in F_n\} \\ & & (i,j) & \longmapsto & (\sigma(i), \sigma(j)) \end{array}$$

est une bijection. Donc :

$$\prod_{1 \leqslant i \neq j \leqslant n} \frac{\sigma(j) - \sigma(i)}{j - i} = 1$$

Or

$$\begin{aligned} \prod_{1 \leqslant i \neq j \leqslant n} \frac{\sigma(j) - \sigma(i)}{j - i} &= \left(\prod_{1 \leqslant i < j \leqslant n} \frac{\sigma(j) - \sigma(i)}{j - i} \right) \left(\prod_{1 \leqslant j < i \leqslant n} \frac{\sigma(j) - \sigma(i)}{j - i} \right) \\ &= \left(\prod_{1 \leqslant i < j \leqslant n} \frac{\sigma(j) - \sigma(i)}{j - i} \right)^2 = 1 \end{aligned}$$

Donc $\displaystyle\prod_{1 \leqslant i < j \leqslant n} \frac{\sigma(j) - \sigma(i)}{j - i} = \pm 1$. Puisque pour tout $i < j$, $\dfrac{\sigma(j) - \sigma(i)}{j - i}$ est positif s'il y a inversion en (i,j) et négatif sinon, le signe de $\displaystyle\prod_{1 \leqslant i < j \leqslant n} \frac{\sigma(j) - \sigma(i)}{j - i}$ est $(-1)^{\ell(\sigma)}$. On en déduit que $\displaystyle\prod_{1 \leqslant i < j \leqslant n} \frac{\sigma(j) - \sigma(i)}{j - i} = (-1)^{\ell(\sigma)} = \varepsilon(\sigma)$.

2. Montrons la deuxième assertion.

$$\begin{aligned}
\varepsilon(\sigma\circ\sigma') &= \prod_{1\leqslant i<j\leqslant n} \frac{\sigma\circ\sigma'(j)-\sigma\circ\sigma'(i)}{j-i} \\
&= \prod_{1\leqslant i<j\leqslant n} \frac{\sigma\circ\sigma'(j)-\sigma\circ\sigma'(i)}{\sigma'(j)-\sigma'(i)}\frac{\sigma'(j)-\sigma'(i)}{j-i} \\
&= \left(\prod_{1\leqslant i<j\leqslant n} \frac{\sigma(\sigma'(j))-\sigma(\sigma'(i))}{\sigma'(j)-\sigma'(i)}\right)\varepsilon(\sigma') \\
&= \left(\prod_{\substack{1\leqslant i,j\leqslant n\\ \sigma'(i)<\sigma'(j)}} \frac{\sigma(\sigma'(j))-\sigma(\sigma'(i))}{\sigma'(j)-\sigma'(i)}\right)\varepsilon(\sigma')
\end{aligned}$$

Maintenant, l'application

$$\begin{array}{ccc}
\{(i,j)/\sigma'(i)<\sigma'(j)\} & \longrightarrow & \{(k,l)/k<l\} \\
(i,j) & \longmapsto & (\sigma'(i),\sigma'(j))
\end{array}$$

est une bijection. Donc

$$\prod_{\substack{1\leqslant i,j\leqslant n\\ \sigma'(i)<\sigma'(j)}} \frac{\sigma(\sigma'(j))-\sigma(\sigma'(i))}{\sigma'(j)-\sigma'(i)} = \prod_{1\leqslant k<l\leqslant n}\frac{\sigma(l)-\sigma(k)}{l-k} = \varepsilon(\sigma)$$

Ceci achève la preuve de l'assertion 2.

3. De 2. on déduit immédiatement que pour tout $\sigma\in\mathfrak{S}_n$, $\varepsilon(id)=\varepsilon(\sigma\circ\sigma^{-1})=\varepsilon(\sigma)\varepsilon(\sigma^{-1})=1$. Donc $\varepsilon(\sigma)$ et $\varepsilon(\sigma^{-1})$ ont même signe et $\varepsilon(\sigma)=\varepsilon(\sigma^{-1})$.

□

Proposition 1.16 *Soit τ une transposition de $\mathfrak{S}_n$. Alors τ est impaire.*

PREUVE: Supposons que $\tau=(i,j)$ avec $i<j$. Nous allons compter le nombre d'inversions en tous les couples (k,l), $k<l$. Si $k\neq i,j$ et $l\neq i,j$, on n'a pas d'inversion en (k,l). Donc regardons les cas où l'un des deux entiers est i ou j.

Si $k<i$, il n'y a pas d'inversion en (k,i) et (k,j). De même si $k>j$, il n'y a pas d'inversion en (i,k) et (j,k). Par contre si $j-i\geqslant 2$ et $i<k<j$, il y a inversion en (i,k) et (k,j), ce qui fait $2(j-i-1)$ inversions pour tous les couples de cette forme. Enfin comme il y a inversion en (i,j), on en déduit que $\ell(\tau)=1+2(j-i-1)$ si $j-i\geqslant 2$ et $\ell(\tau)=1$ si $j-i=1$. On obtient le résultat désiré puisque $\varepsilon(\tau)=(-1)^{\ell(\tau)}$. □

Exercice de cours 1.4 *Montrer que la transposition $(2,4)$, dans $\mathfrak{S}_5$, est de longueur 3, donc que cette transposition est bien impaire.*

1.2 Applications multilinéaires

Dans toute la suite, $\mathbb{K}$ désigne le corps des réels ou des complexes.

1.2.1 Applications multinéaires : définition

Définition 1.17 *Soient E et F deux espaces vectoriels sur $\mathbb{K}$. On dit que l'application :*

$$\begin{array}{cccc} f & : & E^p = E \times \cdots \times E & \longrightarrow & F \\ & & (x_1, \cdots, x_p) & \longmapsto & f(x_1, \cdots, x_p) \end{array}$$

est une **application multilinéaire** *ou* p**-linéaire** *si pour tout i compris entre 1 et p et pour tous vecteurs x_1,...,x_{i-1}, x_{i+1},...,x_p fixés, l'application :*

$$x \longmapsto f(x_1, \cdots, x_{i-1}, x, x_{i+1}, \cdots, x_p)$$

est une application linéaire de E dans F. Si de plus $F = \mathbb{K}$, on dit que f est une **forme p-linéaire**.

Autrement dit on a pour tous vecteurs x et y de E :

$$\begin{array}{l} f(x_1, \cdots, x_{i-1}, \lambda x + \mu y, x_{i+1}, \cdots, x_p) = \\ \lambda f(x_1, \cdots, x_{i-1}, x, x_{i+1}, \cdots, x_p) + \mu f(x_1, \cdots, x_{i-1}, y, x_{i+1}, \cdots, x_p) \end{array}$$

Exercice de cours 1.5 *Soit $E = \mathbb{R}^2$ et $F = \mathbb{R}^3$, et soit $p = 2$. Montrer que l'application*

$$\begin{array}{cccc} f & : & E^2 = \mathbb{R}^2 \times \mathbb{R}^2 & \longrightarrow & \mathbb{R}^3 \\ & & ((x_1, y_1), (x_2, y_2)) & \longmapsto & (0, x_1 y_2 - y_1 x_2, 0) \end{array} \tag{1.1}$$

est bilinéaire, mais pas celle définie de façon similaire par le vecteur $(0, x_1 y_1 - x_2 y_2, 0)$.

Remarque 1.18 *Notons $\mathcal{L}^p(E, F)$ l'ensemble des applications p-linéaires de E dans F. Étant donné $\lambda \in \mathbb{K}$, $f, g \in \mathcal{L}^p(E, F)$, on peut définir les deux opérations suivantes :*

1. $(f + g)(x_1, \cdots, x_p) = f(x_1, \cdots, x_p) + g(x_1, \cdots, x_p)$.
2. $(\lambda f)(x_1, \cdots, x_p) = \lambda f(x_1, \cdots, x_p)$.

Alors il est facile de vérifier que $\mathcal{L}^p(E, F)$ muni de ces deux opérations est un espace vectoriel sur $\mathbb{K}$.

1.2.2 Applications multilinéaires alternées

Définition 1.19 *Soient E et F deux espaces vectoriels sur $\mathbb{K}$ et $v : E^p = E \times \cdots \times E \longrightarrow F$ une application multilinéaire. On dit que v est* **alternée** *si pour tout couple $(i, j) \in F_p^2$ tel que $i \neq j$ et pour tout $(x_1, \cdots x_p) \in E^p$:*

$$x_i = x_j \quad \textit{implique} \quad v(x_1, \cdots, x_p) = 0$$

Exercice de cours 1.6 *Montrer que l'application f définie par (1.1) est alternée.*

Proposition 1.20 *Si $v : E^p = E \times \cdots \times E \longrightarrow F$ est une application multilinéaire alternée, alors pour tout couple $(i,j) \in \{1,\ldots,p\}^2$ tel que $i < j$, on a :*

$$v(x_1,\ldots,x_i,\ldots,x_j,\ldots,x_p) = -v(x_1,\ldots,x_j,\ldots,x_i,\ldots,x_p)$$

Preuve: Comme v est multilinéaire et alternée on a :

$$\begin{aligned}
0 =& v(x_1,\ldots,x_i+x_j,\ldots,x_i+x_j,\ldots,x_p)\\
&= v(x_1,\ldots,x_i,\ldots,x_i,\ldots,x_p) + v(x_1,\ldots,x_i,\ldots,x_j,\ldots,x_p)\\
&+ v(x_1,\ldots,x_j,\ldots,x_i,\ldots,x_p) + v(x_1,\ldots,x_j,\ldots,x_j,\ldots,x_p)\\
&= v(x_1,\ldots,x_i,\ldots,x_j,\ldots,x_p) + v(x_1,\ldots,x_j,\ldots,x_i,\ldots,x_p)
\end{aligned}$$

d'où le résultat. □

Exercice de cours 1.7 *Montrer que le produit vectoriel sur $\mathbb{R}^3$ est une application bilinéaire alternée. On pourra utiliser la formule qui définit ce produit vectoriel, $(x_1,y_1,z_1) \wedge (x_2,y_2,z_2) = (y_1z_2 - y_2z_1, z_1x_2 - z_2x_1, x_1y_2 - x_2y_1)$, ou la règle des trois doigts de Maxwell.*

Remarque 1.21 *L'ensemble $\Lambda^p(E,F)$ des applications multilinéaires alternées de E dans F est un sous-espace vectoriel de $\mathcal{L}^p(E,F)$ pour les opérations définies dans la remarque 1.18.*

Remarque 1.22 *De la proposition 1.20, on déduit immédiatement que si τ est une transposition de $\mathfrak{S}_p$, $v(x_1,\cdots,x_p) = \varepsilon(\tau)v(x_{\tau(1)},\cdots,x_{\tau(p)})$.*

Cette remarque nous conduit à la proposition suivante :

Proposition 1.23 *Soit $v : E^p = E \times \cdots \times E \longrightarrow F$ une application multilinéaire alternée et soit $\sigma \in \mathfrak{S}_p$. Alors pour tout vecteur $(x_1,\cdots,x_p) \in E^p$, on a :*

$$v(x_1,\cdots,x_p) = \varepsilon(\sigma)v(x_{\sigma(1)},\cdots,x_{\sigma(p)})$$

Preuve: La preuve résulte d'une application directe du théorème 1.7. En effet toute permutation s'écrivant comme un produit de transpositions, il existe des transpositions $\tau_1,...,\tau_k$ telles que $\sigma = \tau_1 \circ \cdots \circ \tau_k$. De la remarque ci-dessus, il découle que :

$$\begin{aligned}
v(x_1,\cdots,x_p) &= \varepsilon(\tau_1)v(x_{\tau_1(1)},\cdots,x_{\tau_1(p)})\\
&= \varepsilon(\tau_1)\varepsilon(\tau_2)v(x_{\tau_1\circ\tau_2(1)},\cdots,x_{\tau_1\circ\tau_2(p)})\\
&= \cdots\cdots\cdots\cdots\cdots\\
&= \varepsilon(\tau_1)\varepsilon(\tau_2)\cdots\varepsilon(\tau_k)v(x_{\tau_1\circ\tau_2\circ\cdots\circ\tau_k(1)},\cdots,x_{\tau_1\circ\tau_2\circ\cdots\circ\tau_k(p)})\\
&= \varepsilon(\tau_1)\varepsilon(\tau_2)\cdots\varepsilon(\tau_k)v(x_{\sigma(1)},\cdots,x_{\sigma(p)})
\end{aligned}$$

La deuxième égalité découle de la relation

$$v(y_1, \cdots, y_p) = \varepsilon(\tau_2) v(y_{\tau_2(1)}, \cdots, y_{\tau_2(p)})$$

appliquée à $y_1, \cdots, y_p$ définis par $y_j = x_{\tau_1(j)}$.
et utilisant la propriété 2. du théorème 1.15, on déduit que $\varepsilon(\tau_1)\varepsilon(\tau_2)\cdots\varepsilon(\tau_k) = \varepsilon(\tau_1 \circ \tau_2 \circ \cdots \circ \tau_k) = \varepsilon(\sigma)$, d'où le résultat recherché. $\square$

Proposition 1.24 *Soit E un espace vectoriel de dimension n sur $\mathbb{K}$.*

1. *Soit $\mathcal{B} = (e_1, \ldots, e_n)$ une base de E. Pour tous vecteurs x_1,...,x_n de E, on note $\begin{pmatrix} x_{11} \\ \vdots \\ x_{n1} \end{pmatrix}$,..., $\begin{pmatrix} x_{1n} \\ \vdots \\ x_{nn} \end{pmatrix}$ leurs coordonnées respectives dans la base $\mathcal{B}$. Définissons ω par :*

$$\begin{array}{cccc} \omega \ : & E^n & \longrightarrow & \mathbb{K} \\ & (x_1, \ldots x_n) & \longmapsto & \omega(x_1, \ldots, x_n) = \displaystyle\sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) \prod_{i=1}^{n} x_{\sigma(i),i}. \end{array}$$

   Alors ω est l'unique forme n-linéaire alternée $E^n \to \mathbb{K}$ vérifiant $\omega(e_1, \ldots, e_n) = 1$.

2. *$\forall v \in \Lambda^n(E, \mathbb{K})$ on a $v = v(e_1, \ldots, e_n)\omega$.*
3. *$\dim(\Lambda^n(E, \mathbb{K})) = 1$.*

Exercice de cours 1.8 *Supposer que $n = 2$. Calculer $\omega\left(\begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 3 \\ 4 \end{pmatrix}\right)$.*

Exercice de cours 1.9 *Supposer que $n = 3$. Montrer que $\omega\left(\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \begin{pmatrix} 2 \\ 3 \\ 4 \end{pmatrix}, \begin{pmatrix} 3 \\ 5 \\ 7 \end{pmatrix}\right) = 0$.*

Preuve de la proposition : Montrons que ω est une n-forme linéaire alternée. Soit $x'_i \in E$ de coordonnées $\begin{pmatrix} x'_{1i} \\ \vdots \\ x'_{ni} \end{pmatrix}$ dans $(e_1, \cdots, e_n)$ et soient λ, $\lambda' \in \mathbb{K}$. Alors :

$$\begin{aligned} \omega(x_1, \cdots, \lambda x_i + \lambda' x'_i, \cdots, x_n) &= \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) x_{\sigma(1)1} \cdots (\lambda x_{\sigma(i)i} + \lambda' x'_{\sigma(i)i}) \cdots x_{\sigma(n)n} \\ &= \lambda \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) x_{\sigma(1)1} \cdots x_{\sigma(i)i} \cdots x_{\sigma(n)n} \\ &+ \lambda' \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) x_{\sigma(1)1} \cdots x'_{\sigma(i)i} \cdots x_{\sigma(n)n} \\ &= \lambda \omega(x_1, \cdots, x_i, \cdots, x_n) + \lambda' \omega(x_1, \cdots, x'_i, \cdots, x_n) \end{aligned}$$

Donc ω est multilinéaire. Montrons maintenant que ω est alternée. Soit $(i,j) \in \{1,\ldots,p\}^2$ tel que $i < j$. Supposons que $x_i = x_j$. Considérons $\tau = (i,j)$ la transposition qui échange i et j. Alors :

$$\begin{aligned}
\omega(x_1,\cdots,x_i,\cdots,x_j,\cdots,x_n) &= \omega(x_{\tau(1)},\cdots,x_{\tau(n)}) \\
&= \sum_{\sigma\in\mathfrak{S}_n} \varepsilon(\sigma) x_{\sigma(1)\tau(1)}\cdots x_{\sigma(n)\tau(n)} \\
&= \sum_{\sigma\in\mathfrak{S}_n} \varepsilon(\sigma) x_{\sigma(\tau(1))\tau(\tau(1))}\cdots x_{\sigma(\tau(n))\tau(\tau(n))}.
\end{aligned}$$

La dernière égalité provient de l'égalité

$$\prod_{j=1}^n x_{\sigma(j),\tau(j)} = \prod_{k=1}^n x_{\sigma(k),\tau(k)} = \prod_{j=1}^n x_{\sigma(\tau(j)),\tau(\tau(j))}.$$

Or $\tau\circ\tau = id$, donc :

$$\begin{aligned}
\omega(x_1,\cdots,x_i,\cdots,x_j,\cdots,x_n) &= \sum_{\sigma\in\mathfrak{S}_n} \varepsilon(\sigma) x_{\sigma\circ\tau(1)1}\cdots x_{\sigma\circ\tau(n)n} \\
&= -\sum_{\sigma\in\mathfrak{S}_n} \varepsilon(\sigma\circ\tau) x_{\sigma\circ\tau(1)1}\cdots x_{\sigma\circ\tau(n)n}
\end{aligned}$$

car $\varepsilon(\sigma\circ\tau) = \varepsilon(\sigma)\varepsilon(\tau) = -\varepsilon(\sigma)$ et $\varepsilon(\tau) = -1$. Comme $\sigma \longmapsto \sigma\circ\tau$ est une bijection de $\mathfrak{S}_n$ sur lui-même, on déduit que :

$$\begin{aligned}
\omega(x_1,\cdots,x_i,\cdots,x_j,\cdots,x_n) &= -\sum_{\sigma\in\mathfrak{S}_n} \varepsilon(\sigma) x_{\sigma(1)1}\cdots x_{\sigma(n)n} \\
&= -\omega(x_1,\cdots,x_i,\cdots,x_j,\cdots,x_n)
\end{aligned}$$

Donc $2\omega(x_1,\cdots,x_i,\cdots,x_j,\cdots,x_n) = 0$ et comme $\mathbb{K}$ est $\mathbb{R}$ ou $\mathbb{C}$, on déduit que

$$\omega(x_1,\cdots,x_i,\cdots,x_j,\cdots,x_n) = 0.$$

Donc ω est alternée.

D'autre part si les vecteurs x_1,...,x_n sont les vecteurs e_1,...,e_n de la base canonique alors $x_{ij} = 0$ si $i \neq j$ et $x_{ii} = 1$. Donc $x_{\sigma(1)1}\cdots x_{\sigma(n)n} \neq 0$ si et seulement si $\sigma(i) = i$ pour tout i compris entre 1 et n c'est-à-dire si et seulement si $\sigma = id$. Donc :

$$\omega(e_1,\cdots,e_n) = x_{11}\cdots x_{nn} = 1$$

Soit maintenant v une forme n-linéaire alternée. Alors :

$$\begin{aligned} v(x_1, \cdots, x_n) &= v\left(\sum_{i_1=1}^{n} x_{i_1 1} e_{i_1}, \sum_{i_2=1}^{n} x_{i_2 2} e_{i_2}, \cdots, \sum_{i_n=1}^{n} x_{i_n n} e_{i_n}\right) \\ &= \sum_{i_1=1}^{n} \sum_{i_2=1}^{n} \cdots \sum_{i_n=1}^{n} x_{i_1 1} x_{i_2 2} \cdots x_{i_n n} v(e_{i_1}, \cdots, e_{i_n}) \end{aligned}$$

Comme v est multilinéaire alternée, dès que pour $k \neq l$ on a $i_k = i_l$, alors $v(e_{i_1}, \cdots, e_{i_n}) = 0$. Donc on ne va considérer que les termes où tous les i_j sont différents. Cela revient à dire que l'application $j \longmapsto i_j$ qui va de $\{1, \cdots, n\}$ dans lui-même est une bijection donc une permutation σ de $\mathfrak{S}_n$. Donc en réécrivant $i_j = \sigma(j)$, on a :

$$\begin{aligned} v(x_1, \cdots, x_n) &= \sum_{\sigma \in \mathfrak{S}_n} x_{\sigma(1)1} \cdots x_{\sigma(n)n} v(e_{\sigma(1)}, \cdots, e_{\sigma(n)}) \\ &= \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) x_{\sigma(1)1} \cdots x_{\sigma(n)n} v(e_1, \cdots, e_n) \end{aligned}$$

où dans cette dernière égalité on a appliqué la proposition 1.23. Donc $v = v(e_1, \ldots, e_n)\omega$.

En particulier, si v est une forme n-linéaire alternée telle que $v(e_1, \ldots, e_n) = 1$, on obtient $v = \omega$, ce qui montre l'unicité de ω affirmée au premier point.

Enfin comme $v = v(e_1, \ldots, e_n)\omega$, il est facile de voir que ω est l'unique n-forme alternée telle que $\omega(e_1, \ldots, e_n) = 1$.

□

1.3 Définition du déterminant et premières propriétés

1.3.1 Déterminant d'une famille de n-vecteurs

Notation 1.25 *D'après la proposition précédente si E est un $\mathbb{K}$-espace vectoriel de dimension n muni d'une base $\mathcal{B} = (e_1, \cdots, e_n)$, alors il existe une unique n-forme linéaire alternée ω telle que $\omega(e_1, \cdots, e_n) = 1$. Notons $det_{\mathcal{B}} = \omega$.*

Définition 1.26 *Soit E un espace vectoriel de dimension n muni d'une base $\mathcal{B} = (e_1, \ldots, e_n)$. Alors pour toute famille de n vecteurs $(u_1, \ldots, u_n)$ le scalaire $det_{\mathcal{B}}(u_1, \ldots, u_n)$ s'appelle le* **déterminant de la famille** *$(u_1, \ldots, u_n)$ dans la base $\mathcal{B}$.*

Remarque 1.27 *De la proposition 1.24 on déduit immédiatement que :*

$$\forall v \in \Lambda^n(E, \mathbb{K}), \forall (u_1, \cdots, u_n) \in E^n, v(u_1, \cdots, u_n) = v(e_1, \cdots, e_n) det_{\mathcal{B}}(u_1, \cdots, u_n)$$

Preuve: En effet, la fonction v est n-linéaire et alternée, donc elle est égale à $\lambda \cdot \det_{\mathcal{B}}$ pour un certain scalaire λ. En calculant sa valeur sur $(e_1, \ldots, e_n)$, on voit que $\lambda = v(e_1, \ldots, e_n)$. □

Proposition 1.28 *Soit $(u_1, \cdots, u_n) \in E^n$. Alors la famille $(u_1, \cdots, u_n)$ est liée si et seulement si $det_{\mathcal{B}}(u_1, \cdots, u_n) = 0$.*

Preuve: Supposons que $(u_1, \cdots, u_n)$ est liée. Alors il existe des scalaires $\lambda_1, \cdots, \lambda_n$ non tous nuls tels que :

$$\lambda_1 u_1 + \cdots + \lambda_n u_n = 0$$

Comme $\lambda_1, \cdots, \lambda_n$ non tous nuls, il existe $i \in \{1, \cdots, n\}$ tel que $\lambda_i \neq 0$. Donc :

$$u_i = -\sum_{\substack{1\leqslant j\leqslant n\\ j\neq i}} \frac{\lambda_j}{\lambda_i} u_j = \sum_{\substack{1\leqslant j\leqslant n\\ j\neq i}} \alpha_j u_j$$

et

$$\det_{\mathcal{B}}(u_1, \cdots, u_n) = \sum_{\substack{1\leqslant j\leqslant n\\ j\neq i}} \alpha_j \det_{\mathcal{B}}(u_1, \cdots, u_{i-1}, u_j, u_{i-1}, \cdots, u_n) = 0$$

Réciproquement si $\det_{\mathcal{B}}(u_1, \cdots, u_n) = 0$. Supposons par l'absurde que $(u_1, \cdots, u_n)$ est libre. Comme E est de dimension n, $\mathcal{B}' = (u_1, \cdots, u_n)$ est une base de E. Donc, d'après la Remarque 1.27 appliquée à $v = \det_{\mathcal{B}'}$:

$$\det_{\mathcal{B}'}(u_1, \cdots, u_n) = \det_{\mathcal{B}'}(e_1, \cdots, e_n)\det_{\mathcal{B}}(u_1, \cdots, u_n) = 0$$

Ceci contredit le fait que $\det_{\mathcal{B}'}(u_1, \cdots, u_n) = 1$ donc $(u_1, \cdots, u_n)$ est liée. □

Exercice de cours 1.10 *Montrer que les calculs de déterminants effectués dans les deux exercices suivant la Proposition 1.24 sont en accord avec la Proposition 1.28.*

1.3.2 Déterminant d'un endomorphisme

Proposition et définition 1.29 *Soit E un espace vectoriel sur $\mathbb{K}$ de dimension finie n, f un endomorphisme de E et $v : E^n \longrightarrow \mathbb{K}$ une forme n-linéaire alternée non nulle. Soit $f^\star(v)$ l'application de E^n dans $\mathbb{K}$ définie pour tous $x_1, \cdots, x_n$ par :*

$$f^\star(v)(x_1, \cdots, x_n) = v(f(x_1), \cdots, f(x_n))$$

Alors $f^\star(v)$ est une forme n-linéaire alternée et $f^\star(v) = \lambda v$. De plus λ ne dépend pas du choix de v et s'appelle le **déterminant** *de f. On note $\lambda = det(f)$.*

Preuve: La preuve du fait que $f^\star(v)$ est multilinéaire alternée est laissée au lecteur. La proposition 1.24 nous disant que l'espace des formes multilinéaires alternées est un espace vectoriel de dimension 1, il s'ensuit que v étant non nul est une base de cet espace et donc $f^\star(v) = \lambda v$. Maintenant si v' est une autre forme n-linéaire alternée non nulle, alors $v' = \mu v$ et on en déduit facilement que $f^\star(v') = \lambda v'$ ce qui prouve l'unicité de λ. □

Exercice de cours 1.11 *Soit $\omega : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ la forme bilinéaire alternée définie par la formule $\omega((x_1, y_1), (x_2, y_2)) = x_1 y_2 - x_2 y_1$ et soit $f_1, f_2, f_3 : \mathbb{R}^2 \to \mathbb{R}^2$ définies par $f_1(x, y) = (2y, 3x)$, $f_2(x, y) = (x + 2y, y)$ et $f_3(x, y) = (2x + y, 4x + 2y)$. Calculer $f_1^\star\omega, f_2^\star\omega$ et $f_3^\star\omega$.*

Proposition 1.30 *Soit E un espace vectoriel de dimension n muni d'une base $\mathcal{B} = (e_1, \dots, e_n)$ et soit $f \in \mathcal{L}(E)$. Alors :*

$$det(f) = det_{\mathcal{B}}(f(e_1), \dots, f(e_n))$$

Preuve: D'après la définition de $\det(f)$ on a

$$\det_{\mathcal{B}}(f(e_1), \dots, f(e_n)) = f^\star(\det_{\mathcal{B}})(e_1, \dots, e_n) = \det(f)\det_{\mathcal{B}}(e_1, \dots, e_n) = \det(f).$$

□

Proposition 1.31 *Soit E un espace vectoriel de dimension finie n sur $\mathbb{K}$. Alors pour tout endomorphisme f de E et tout $\lambda \in \mathbb{K}$ on a :*

$$det(\lambda f) = \lambda^n det(f)$$

Preuve: Soit $\mathcal{B} = (e_1, \cdots, e_n)$ une base de E alors :

$$\begin{aligned}\det(\lambda f) &= \det_{\mathcal{B}}((\lambda f)(e_1), \cdots, (\lambda f)(e_n)) = \det_{\mathcal{B}}(\lambda f(e_1), \cdots, \lambda f(e_n)) \\ &= \lambda^n \det_{\mathcal{B}}(f(e_1), \cdots, f(e_n))\end{aligned}$$

Cette dernière égalité provient du fait que $\det_{\mathcal{B}}$ est une forme n-linéaire. On en déduit donc le résultat recherché. □

Proposition 1.32 *Soit E un espace vectoriel de dimension finie n sur $\mathbb{K}$. Soient f et g deux endomorphismes de E. Alors :*

$$det(f \circ g) = det(f) det(g)$$

Preuve: Soit v une forme n-linéaire alternée non nulle. Alors :

$$\begin{aligned}((f \circ g)^\star v)(x_1, \cdots, x_n) &= v((f \circ g)(x_1), \cdots, (f \circ g)(x_n) \\ &= v(f(g(x_1)), \cdots, f(g(x_n))) \\ &= \det(f) v(g(x_1), \cdots, g(x_n)) \\ &= \det(f)\det(g) v(x_1, \cdots, x_n)\end{aligned}$$

Ceci permet de conclure que $(f \circ g)^\star(v) = \det(f)\det(g)v$. Mais d'autre part, on a aussi $(f \circ g)^\star(v) = \det(f \circ g)v$, d'où le résultat. □

Exercice de cours 1.12 *Pour les fonctions f_1, f_2, f_3 de l'exercice précédent, calculer $f_1 \circ f_2$, et $det(f_1 \circ f_2)$ de deux manières différentes.*

Proposition 1.33 *Soit E un espace vectoriel de dimension finie n.*

1. *Alors $det(Id_E) = 1$.*
2. *Un endomorphisme f de E est un automorphisme de E si et seulement si $det(f) \neq 0$. Si f est un automorphisme alors : $det(f^{-1}) = \dfrac{1}{det(f)}$.*

PREUVE: Étant donnée v une forme n-linéaire alternée non nulle on a de manière immédiate $\mathrm{Id}_E^\star v = v$, donc $\det(\mathrm{Id}_E) = 1$ ce qui montre le point 1.

Supposons maintenant que f est un automorphisme. Alors d'après la proposition 1.32 on a :

$$\det(f \circ f^{-1}) = \det(f)\det(f^{-1}) = \det(\mathrm{Id}_E) = 1$$

Ce qui montre que $\det(f) \neq 0$ ainsi que la relation demandée.

Réciproquement supposons que $\det(f) \neq 0$. Soit $\mathcal{B} = (e_1, \ldots, e_n)$ une base de E. Alors

$$\det(f) = \det_{\mathcal{B}}(f(e_1), \ldots, f(e_n))$$

Comme $\det(f) \neq 0$, d'après la proposition 1.28 $(f(e_1), \ldots, f(e_n))$ est une famille libre de E, donc une base de E. Il s'ensuit que f est surjective et comme les espaces de départ et d'arrivée sont de même dimension, f est un automorphisme. □

Proposition 1.34 *Soit E un espace vectoriel de dimension n sur $\mathbb{K}$ muni d'une base $\mathcal{B} = (e_1, \ldots, e_n)$. Soient $f \in \mathcal{L}(E)$ et $A = (a_{ij})_{1\leqslant i,j\leqslant n}$ la matrice de f dans $\mathcal{B}$. Alors*

$$det(f) = \sum_{\sigma\in\mathfrak{S}_n} \varepsilon(\sigma) \prod_{i=1}^{n} a_{\sigma(i)i}$$

PREUVE: D'après la proposition 1.30 $\det(f) = \det_{\mathcal{B}}(f(e_1), \ldots, f(e_n))$. D'autre part la j-ième colonne de A n'étant rien d'autre que les coordonnées dans la base $\mathcal{B}$ de $f(e_j)$, on déduit du point 1. de la proposition 1.24 la relation souhaitée. □

1.3.3 Déterminant d'une matrice

Soit $A \in \mathcal{M}_{n,n}(\mathbb{K})$. On note $A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}$.

Définition 1.35 *On note $f_A : \mathbb{K}^n \to \mathbb{K}^n$ l'application linéaire définie par : $f_A(X) = AX$. Le déterminant de A est alors défini par la formule*

$$det(A) = det(f_A). \tag{1.2}$$

Remarque 1.36

— *On a évidemment $det(I_n) = 1$. En effet :*

$$det(I_n) = det(Id_{\mathbb{K}^n}) = 1$$

— *D'autre part, la Proposition 1.34 donne $det(A) = \displaystyle\sum_{\sigma\in\mathfrak{S}_n} \varepsilon(\sigma) \prod_{i=1}^{n} a_{\sigma(i)i}$.*

Comme corollaire de la Définition 1.35 et des propositions 1.31, 1.32 et 1.33 on a les théorèmes suivants :

Théorème 1.37 *Soit* $A \in \mathcal{M}_n(\mathbb{K})$ *et* $\lambda \in \mathbb{K}$. *Alors :*

$$det(\lambda A) = \lambda^n det(A)$$

Théorème 1.38 *Soient* $A, B \in \mathcal{M}_n(\mathbb{K})$. *Alors* $det(AB) = det(A)det(B)$.

Théorème 1.39 *Soit* $A \in \mathcal{M}_n(\mathbb{K})$. *Alors* A *est inversible si et seulement si* $det(A) \neq 0$. *Si* A *est inversible alors :* $det(A^{-1}) = \dfrac{1}{det(A)}$.

Nous verrons dans la partie 4.2 que l'on peut retrouver ce résultat de manière différente.

Proposition 1.40 *Soit* E *un espace vectoriel de dimension* n *sur un corps* $\mathbb{K}$ *muni d'une base* $\mathcal{B}$ *et soit* $(u_1, \cdots, u_n)$ *une famille de* n *vecteurs. Soit* P *la matrice de* $(u_1, \cdots, u_n)$ *dans* $\mathcal{B}$. *Alors*

$$det_{\mathcal{B}}(u_1, \cdots, u_n) = det(P)$$

Preuve: La matrice P est en effet la matrice, dans la base $\mathcal{B}$, de l'application linéaire f qui envoie les vecteurs de la base $\mathcal{B}$ sur les vecteurs $u_1, \cdots, u_n$. Or, par définition, le déterminant d'une matrice est le déterminant de l'application linéaire qui lui est associée. Donc $\det(P) = \det(f)$. Le résultat découle donc de la Proposition 1.30, qui donne $\det(f) = \det_{\mathcal{B}}(f(e_1), \cdots, f(e_n))$. □

1.3.4 Déterminant des matrices carrées 2×2

Soit $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$, alors $\det(A) = \sum_{\sigma \in \mathfrak{S}_2} \varepsilon(\sigma) a_{\sigma(1)1} a_{\sigma(2)2}$. Il n'y a que deux permutations dans $\mathfrak{S}_2$, les permutations $id = \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix}$ et $\tau = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$. On a alors $\varepsilon(id) = 1$ et $\varepsilon(\tau) = -1$.

Donc :

Proposition 1.41 *Soit* $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$, *alors*

$$det(A) = a_{11}a_{22} - a_{21}a_{12}.$$

1.3.5 Matrices transposées

Définition 1.42 *Soit* $A = (a_{ij})_{1 \leqslant i,j \leqslant n}$. *La transposée* ${}^tA$ *de* A *est la matrice dont les coefficients sont les* $\tilde{a}_{i,j}$, *avec* $\tilde{a}_{ij} = a_{ji}$.

Exercice de cours 1.13 *Soit* $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathcal{M}_{2,2}(\mathbb{K})$. *Montrer que* $det({}^tA) = det(A)$.

Proposition 1.43 *Soit* $A \in \mathcal{M}_n(\mathbb{K})$, *alors* $det({}^tA) = det(A)$.

Preuve: Soit $A = (a_{ij})_{1\leqslant i,j\leqslant n}$, on a ${}^tA = (\tilde{a}_{ij})_{1\leqslant i,j\leqslant n}$ avec $\tilde{a}_{ij} = a_{ji}$. Donc :

$$\begin{aligned}\det({}^tA) &= \sum_{\sigma\in\mathfrak{S}_n} \varepsilon(\sigma)\tilde{a}_{\sigma(1)1}\cdots\tilde{a}_{\sigma(n)n}\\ &= \sum_{\sigma\in\mathfrak{S}_n} \varepsilon(\sigma)a_{1\sigma(1)}\cdots a_{n\sigma(n)}\end{aligned}$$

On peut permuter les facteurs $a_{1\sigma(1)}$,..., $a_{n\sigma(n)}$. On a alors :

$$a_{1\sigma(1)}\cdots a_{n\sigma(n)} = a_{\sigma'(1)\sigma(\sigma'(1))}\cdots a_{\sigma'(n)\sigma(\sigma'(n))}$$

pour tout $\sigma' \in \mathfrak{S}_n$. C'est en particulier vrai si on prend $\sigma' = \sigma^{-1}$. On a donc :

$$\det({}^tA) = \sum_{\sigma\in\mathfrak{S}_n} \varepsilon(\sigma)a_{\sigma^{-1}(1)1}\cdots a_{\sigma^{-1}(n)n}$$

mais $\varepsilon(\sigma) = \varepsilon(\sigma^{-1})$, donc

$$\det({}^tA) = \sum_{\sigma\in\mathfrak{S}_n} \varepsilon(\sigma^{-1})a_{\sigma^{-1}(1)1}\cdots a_{\sigma^{-1}(n)n}$$

Or l'application $\sigma \longmapsto \sigma^{-1}$ est une bijection de $\mathfrak{S}_n$ dans lui-même. Donc finalement,

$$\det({}^tA) = \sum_{\sigma\in\mathfrak{S}_n} \varepsilon(\sigma)a_{\sigma(1)1}\cdots a_{\sigma(n)n} = \det(A)$$

□

Remarque 1.44 *Le fait que* $\det({}^tA) = \det(A)$ *a pour conséquence que toutes les propriétés du déterminant démontrées sur les colonnes restent vraies sur les lignes et vice versa.*

1.3.6 Opérations sur les lignes et les colonnes

Soit $A = \begin{pmatrix} a_{11} & \cdots & a_{1n}\\ \vdots & & \vdots\\ a_{n1} & \cdots & a_{nn}\end{pmatrix} \in \mathcal{M}_n(\mathbb{K})$.

On notera par la suite :

$$C_1(A) = \begin{pmatrix} a_{11}\\ \vdots\\ a_{n1}\end{pmatrix}, \cdots, C_n(A) = \begin{pmatrix} a_{1n}\\ \vdots\\ a_{nn}\end{pmatrix}$$

les colonnes de A et :

$$L_1(A), \cdots, L_n(A)$$

les lignes de A.

Alors $\det(A) = \det_{\mathcal{C}}(C_1(A), \cdots, C_n(A)) = \det({}^tA) = \det_{\mathcal{C}}({}^tL_1(A), \cdots, {}^tL_n(A))$. Les deux propositions qui suivent sont des conséquences immédiates du fait que $\det_{\mathcal{C}}$ est multilinéaire alternée.

Proposition 1.45 *Si une colonne (ou une ligne) est combinaison linéaire d'autres colonnes (ou lignes) alors* $det(A) = 0$. *En particulier, si deux colonnes (ou deux lignes) sont égales, alors* $det(A) = 0$.

Exercice de cours 1.14 *Soit* $A \in \mathcal{M}_{n,n}(\mathbb{K})$ *une matrice telle que* $L_1(A) = L_2(A)$. *Donner les détails de la preuve de l'égalité* $det(A) = 0$.

Proposition 1.46 (Opérations sur les colonnes)

1. *Si à la colonne* $C_j(A)$ *(ou à la ligne* $L_i(A)$*) on ajoute une combinaison linéaire des autres colonnes* $\sum_{k\neq j} \lambda_k C_k(A)$ *(ou des autres lignes* $\sum_{k\neq i} \lambda_k L_k(A)$*),* **le déterminant reste inchangé**.
2. *Si on permute deux colonnes (ou deux lignes) de* A, **le déterminant change de signe**.
3. *Si on multiplie une colonne (ou une ligne) de* A *par* μ, *alors* **le déterminant de** A **est multiplié par** μ.

Exercice de cours 1.15 *Montrer le cas particulier suivant de la Proposition ci-dessus :*

$$det_{\mathcal{C}}(C_1(A) + 2C_2(A), C_2(A), \cdots, C_n(A)) = det_{\mathcal{C}}(C_1(A), C_2(A), \cdots, C_n(A)).$$

Exercice de cours 1.16 *Calculer le déterminant de la matrice* $\begin{pmatrix} 1 & 2 \\ 3 & 6 \end{pmatrix}$. *Vérifier le résultat obtenu à l'aide de la Proposition 1.45.*

1.4 Développement d'un déterminant suivant une ligne ou une colonne

1.4.1 Cofacteurs et comatrices

Soit $A \in \mathcal{M}_n(\mathbb{K})$. On note $A(i,j)$ la matrice de $\mathcal{M}_{n-1}(\mathbb{K})$ obtenue en supprimant la i-ième ligne et la j-ième colonne de A.

Exemple 1.47 *Soit* $A = \begin{pmatrix} 3 & 3 & 7 \\ 0 & 2 & 2 \\ -3 & -1 & -4 \end{pmatrix}$. *Alors :*

$$A(1,2) = \begin{pmatrix} 0 & 2 \\ -3 & -4 \end{pmatrix}$$

Définitions 1.48 *Soit* $A \in \mathcal{M}_n(\mathbb{K})$.

- *On appelle* **cofacteur** *de* $A = (a_{ij})_{1\leqslant i,j\leqslant n}$ *d'indice* (i,j) *le coefficient*

$$\tilde{a}_{ij} = (-1)^{i+j} det(A(i,j))$$

- *La matrice* $\tilde{A} = (\tilde{a}_{ij})_{1\leqslant i,j\leqslant n}$ *de* $\mathcal{M}_n(\mathbb{K})$ *s'appelle la* **matrice des cofacteurs** *ou* **comatrice**.

Exercice de cours 1.17 *Soit A la matrice* $\begin{pmatrix} 3 & 3 & 7 \\ 0 & 2 & 2 \\ -3 & -1 & -4 \end{pmatrix}$. *Montrer que* $\tilde{A} = \begin{pmatrix} -6 & -6 & 6 \\ 5 & 9 & -6 \\ -8 & -6 & 6 \end{pmatrix}$.

Avant d'énoncer le théorème principal, nous avons besoin des deux lemmes suivants :

Lemme 1.49 *Soit $B \in \mathcal{M}_{n-1}(\mathbb{K})$ et* $A = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & B & \\ 0 & & & \end{pmatrix}$. *Alors $det(A) = det(B)$.*

Preuve: Considérons les applications :

$$\begin{array}{rcccc} \Phi & : & \mathcal{M}_{n-1,1}(\mathbb{K}) & \longrightarrow & \mathcal{M}_{n,1}(\mathbb{K}) \\ & & X = \begin{pmatrix} x_1 \\ \vdots \\ x_{n-1} \end{pmatrix} & \longmapsto & \Phi(X) = \begin{pmatrix} 0 \\ x_1 \\ \vdots \\ x_{n-1} \end{pmatrix} \end{array}$$

et

$$\begin{array}{rcccc} V & : & \mathcal{M}_{n-1,1}(\mathbb{K})^{n-1} & \longrightarrow & \mathbb{K} \\ & & (X_1, \cdots, X_{n-1}) & \longmapsto & \det_{\mathcal{C}}(E_1, \Phi(X_1), \cdots, \Phi(X_{n-1})) \end{array}$$

Alors Φ est une application linéaire et de manière immédiate, V est une forme $n-1$-linéaire alternée. Soit $\mathcal{C}' = (E'_1, \cdots, E'_{n-1})$ la base canonique de $\mathcal{M}_{n-1,1}(\mathbb{K})$. Alors

$$V(E'_1, \cdots, E'_{n-1}) = \det_{\mathcal{C}}(E_1, \Phi(E'_1), \cdots, \Phi(E'_{n-1})) = \det_{\mathcal{C}}(E_1, \cdots, E_n) = 1$$

D'après la proposition 1.24, il existe une unique $n-1$-forme alternée sur $\mathcal{M}_{n-1,1}(\mathbb{K})$ vérifiant $V(E'_1, \ldots, E'_{n-1}) = 1$. Donc $V = \det_{\mathcal{C}'}$ et on a :

$$\begin{aligned} \det(B) &= \det_{\mathcal{C}'}(C_1(B), \cdots, C_{n-1}(B)) \\ &= \det_{\mathcal{C}}[E_1, \Phi(C_1(A)), \cdots, \Phi(C_{n-1}(A))] \\ &= \begin{vmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & B & \\ 0 & & & \end{vmatrix} \end{aligned}$$

ce qui achève la preuve. □

Lemme 1.50 *Soit* $C = \begin{pmatrix} a_{11} & \cdots & a_{1j-1} & 0 & a_{1j+1} & \cdots & a_{1n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{i-11} & \cdots & a_{i-1j-1} & 0 & a_{i-1j+1} & \cdots & a_{i-1n} \\ a_{i1} & \cdots & a_{ij-1} & 1 & a_{ij+1} & \cdots & a_{in} \\ a_{i+11} & \cdots & a_{i+1j-1} & 0 & a_{i+1j+1} & \cdots & a_{i+1n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & \cdots & a_{nj-1} & 0 & a_{nj+1} & \cdots & a_{nn} \end{pmatrix}$.

Alors $det(C) = (-1)^{i+j} det(A(i,j))$.

PREUVE: Après avoir échangé la ligne i avec la ligne $i-1$, puis la ligne $i-1$ avec la ligne $i-2$, et ainsi de suite jusqu'à la première ligne, et après avoir fait de même avec les colonnes, on obtient d'après les règles de la Proposition 1.46 :

$$\det(C) = (-1)^{i+j}\begin{vmatrix} 1 & a_{i1} & \cdots & a_{ij-1} & a_{ij+1} & \cdots & a_{in} \\ 0 & a_{11} & \cdots & a_{1j-1} & a_{1j+1} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 0 & a_{i-11} & \cdots & a_{i-1j-1} & a_{i-1j+1} & \cdots & a_{i-1n} \\ 0 & a_{i+11} & \cdots & a_{i+1j-1} & a_{i+1j+1} & \cdots & a_{i+1n} \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 0 & a_{n1} & \cdots & a_{nj-1} & a_{nj+1} & \cdots & a_{nn} \end{vmatrix}$$

$$= (-1)^{i+j}\begin{vmatrix} 1 & a_{i1} \quad \cdots \quad a_{ij-1} \quad a_{ij+1} \quad \cdots \quad a_{in} \\ 0 & \\ \vdots & \\ \vdots & A(i,j) \\ \vdots & \\ 0 & \end{vmatrix}$$

$$= (-1)^{i+j}\begin{vmatrix} 1 & 0 \quad \cdots \quad \cdots \quad \cdots \quad \cdots \quad 0 \\ 0 & \\ \vdots & \\ \vdots & A(i,j) \\ \vdots & \\ 0 & \end{vmatrix} = (-1)^{i+j}\det(A(i,j))$$

d'après le lemme précédent. □

Nous sommes maintenant en mesure d'énoncer un théorème fondamental sur le calcul des déterminants qui sert constamment.

Théorème 1.51 *Soit $A = (a_{ij}) \in \mathcal{M}_n(\mathbb{K})$ et k un entier tel que $1 \leqslant k \leqslant n$. Alors :*

1. *$det(A) = \sum_{i=1}^{n}(-1)^{i+k}a_{ik}\,det(A(i,k)) = \sum_{i=1}^{n} a_{ik}\tilde{a}_{ik}$. On dit qu'on* **développe suivant la colonne** *k.*
2. *$det(A) = \sum_{j=1}^{n}(-1)^{k+j}a_{kj}\,det(A(k,j)) = \sum_{j=1}^{n} a_{kj}\tilde{a}_{kj}$. On dit qu'on* **développe suivant la ligne** *k.*

Exemple 1.52 *Calculons $det(A) = \begin{vmatrix} 3 & 3 & 7 \\ 0 & 2 & 2 \\ -3 & -1 & -4 \end{vmatrix}$. Nous allons développer ce déterminant suivant la deuxième colonne. Pour cela affectons chaque coefficients a_{ij} de la matrice, du signe $(-1)^{i+j}$. On a alors :*

$$\begin{vmatrix} 3 & 3 & 7 \\ 0 & 2 & 2 \\ -3 & -1 & -4 \end{vmatrix} = \begin{vmatrix} 3^+ & \mathbf{3}^- & 7^+ \\ 0^- & \mathbf{2}^+ & 2^- \\ -3^+ & -\mathbf{1}^- & -4^+ \end{vmatrix}$$
$$= -3\begin{vmatrix} 0 & 2 \\ -3 & -4 \end{vmatrix} + 2\begin{vmatrix} 3 & 7 \\ -3 & -4 \end{vmatrix} - (-1)\begin{vmatrix} 3 & 7 \\ 0 & 2 \end{vmatrix}$$

où $\begin{vmatrix} 0 & 2 \\ -3 & -4 \end{vmatrix}$ *a été obtenue en supprimant la première ligne et la deuxième colonne,* $\begin{vmatrix} 3 & 7 \\ -3 & -4 \end{vmatrix}$ *en supprimant la deuxième ligne et la deuxième colonne et* $\begin{vmatrix} 3 & 7 \\ 0 & 2 \end{vmatrix}$ *la troisième ligne et la deuxième colonne.*

D'où $det(A) = -3 \cdot 6 + 2 \cdot 9 + 1 \cdot 6 = 6.$

Preuve du théorème 1.51: Montrons le 1. du théorème. La k-ième colonne de A peut sécrire :

$$\begin{pmatrix} a_{1k} \\ \vdots \\ a_{nk} \end{pmatrix} = a_{1k}\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \cdots + a_{nk}\begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$$

comme le déterminant est multilinéaire, on a immédiatement,

$$\det(A) = \sum_{i=1}^{n} a_{ik} \begin{vmatrix} a_{11} & \cdots & a_{1k-1} & 0 & a_{1k+1} & \cdots & a_{1n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{i-11} & \cdots & a_{i-1k-1} & 0 & a_{i-1k+1} & \cdots & a_{i-1n} \\ a_{i1} & \cdots & a_{ik-1} & 1 & a_{ik+1} & \cdots & a_{in} \\ a_{i+11} & \cdots & a_{i+1k-1} & 0 & a_{i+1k+1} & \cdots & a_{i+1n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & \cdots & a_{nk-1} & 0 & a_{nk+1} & \cdots & a_{nn} \end{vmatrix}$$

Et d'après le lemme 1.50, on a :

$$\det(A) = \sum_{i=1}^{n} (-1)^{i+k} a_{ik} \det(A(i,k)).$$

D'autre part, comme $\det({}^tA) = \det(A)$, on obtient immédiatement l'assertion 2.

□

Proposition 1.53 (Règle de Sarrus) *On a :*

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}a_{22}a_{33} + a_{21}a_{32}a_{13} + a_{31}a_{12}a_{23}$$
$$- (a_{13}a_{22}a_{31} + a_{23}a_{32}a_{11} + a_{33}a_{12}a_{21})$$

La preuve s'obtient en développant par rapport à n'importe quelle ligne ou colonne.

Remarque 1.54 *On obtient facilement le résultat en écrivant :*

$$\begin{array}{l}\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \\ \begin{matrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{matrix}\end{array} = a_{11}a_{22}a_{33} + a_{21}a_{32}a_{13} + a_{31}a_{12}a_{23}$$
$$- (a_{13}a_{22}a_{31} + a_{23}a_{32}a_{11} + a_{33}a_{12}a_{21})$$

Attention ! **La règle de Sarrus fonctionne uniquement pour les matrices** $(3, 3)$.

1.4.2 Matrices triangulaires

Définition 1.55 *Une matrice $A = (a_{i,j})$ est dite triangulaire si elle est triangulaire inférieure (ie $i < j \Rightarrow a_{i,j} = 0$) ou triangulaire supérieure (ie $i > j \Rightarrow a_{i,j} = 0$).*

Proposition 1.56 *Soit $A = (a_{ij})_{1 \leqslant i,j \leqslant n}$ une matrice triangulaire. Alors $det(A) = a_{11} \cdots a_{nn}$.*

Preuve: On démontre le résultat par récurrence sur n pour les matrices triangulaires supérieures. La relation est vraie pour $n = 1$. Supposons maintenant qu'elle est vraie pour les matrices de $\mathcal{M}_n(\mathbb{K})$ pour un entier $n \geqslant 1$ et montrons qu'elle reste vraie pour les matrices de $\mathcal{M}_{n+1}(\mathbb{K})$. Soit A une matrice triangulaire supérieure de $\mathcal{M}_{n+1}(\mathbb{K})$. Alors

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & \cdots & a_{1n} \\ 0 & a_{22} & & & \vdots \\ \vdots & 0 & \ddots & & \vdots \\ \vdots & \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & a_{n+1\ n+1} \end{pmatrix}$$

En développant suivant la dernière ligne on a :

$$\det(A) = a_{n+1\ n+1} \begin{vmatrix} a_{11} & a_{12} & \cdots & \cdots & a_{1n} \\ 0 & a_{22} & & & \vdots \\ \vdots & 0 & \ddots & & \vdots \\ \vdots & \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & a_{nn} \end{vmatrix}$$

Par hypothèse de récurrence le résultat étant vrai pour les matrices de $\mathcal{M}_n(\mathbb{K})$ on obtient le résultat voulu.

D'autre part, comme pour toute matrice carrée A, $\det(A) = \det({}^tA)$, la proposition reste vraie pour les matrices triangulaires inférieures. $\square$

Exercice de cours 1.18 *Calculer le déterminant de la matrice* $\begin{pmatrix} 1 & 2 & 3 \\ 0 & 2 & 3 \\ 0 & 0 & 3 \end{pmatrix}$.

1.4.3 Matrices inverses et déterminant

Théorème 1.57 *Pour tout* $A \in \mathcal{M}_n(\mathbb{K})$*, on a :*

$$A\,{}^t\tilde{A} = {}^t\tilde{A}A = det(A)I_n$$

PREUVE: Posons $A = (a_{ij})$ et ${}^t\tilde{A} = (b_{jk})$. Alors $b_{jk} = \tilde{a}_{kj}$
Soit $C = A\,{}^t\tilde{A}$. Montrons que $C = \det(A)I_n$. Pour cela posons $C = (c_{ik})$.

Soit $D(i,k)$ la matrice $\begin{pmatrix} a_{11} & \cdots\cdots\cdots & a_{1n} \\ \vdots & & \vdots \\ a_{i1} & \cdots\cdots\cdots & a_{in} \\ \vdots & & \vdots \\ a_{k-11} & \cdots\cdots\cdots & a_{k-1n} \\ a_{i1} & \cdots\cdots\cdots & a_{in} \\ a_{k+11} & \cdots\cdots\cdots & a_{k+1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots\cdots\cdots & a_{nn} \end{pmatrix}$

D'après le 1. du théorème 1.51, on a

$$\begin{aligned} \det(D(i,k)) &= \sum_{j=1}^{n} (-1)^{k+j} a_{ij} \det(A(k,j)) = \sum_{j=1}^{n} a_{ij}\tilde{a}_{kj} \\ &= \sum_{j=1}^{n} a_{ij} b_{jk} = c_{i,k} \end{aligned}$$

Si $i = k$, alors $D(i,k) = A$, donc $c_{ii} = \det(A)$. Si $i \neq k$, les lignes k et i de $D(i,k)$ sont les mêmes. Donc $c_{i,k} = 0$. On en déduit donc que $C = \det(A)I_n$.

□

Exercice de cours 1.19 *Soit* A *la matrice* $\begin{pmatrix} 3 & 3 & 7 \\ 0 & 2 & 2 \\ -3 & -1 & -4 \end{pmatrix}$. *Montrer par le calcul que*

$$A\,{}^t\tilde{A} = {}^t\tilde{A}A = det(A)I_3.$$

On en déduit alors immédiatement le corollaire suivant :

Corollaire 1.58 *Soit* $A \in \mathcal{M}_n(\mathbb{K})$*. Alors* A *est inversible si et seulement si* $det(A) \neq 0$ *(résultat déjà énoncé dans le théorème 1.39). Si* A *est inversible,* $A^{-1} = \dfrac{{}^t\tilde{A}}{det(A)}$.

Exercice de cours 1.20 *Soit A la matrice* $\begin{pmatrix} -3 & 2 & -1 \\ 3 & 1 & 2 \\ 2 & 0 & 1 \end{pmatrix}$. *Calculer* $det(A)$. *Monter que A est inversible, et que son inverse a des coefficients entiers. Calculer A^{-1} par deux méthodes différentes.*

1.5 Formules de Cramer

On rappelle que si $A \in \mathcal{M}_n(\mathbb{K})$ et $(S) : AX = Y$ est le système de n équations à n inconnues associé, alors les propriétés suivantes sont équivalentes :

— (S) est un système de Cramer.
— (S) admet une solution et une seule.
— $\mathrm{rg}(A) = n$.
— A est inversible.
— $\det(A) \neq 0$.

Si (S) est de Cramer, la solution est donnée par $X = A^{-1}Y = \dfrac{{}^t\tilde{A}Y}{\det(A)}$. Si on note $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ et $Y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$, on a alors la proposition suivante :

Proposition 1.59 *(***Formules de Cramer***) Si (S) est de Cramer, alors :*

$$x_i = \frac{\begin{vmatrix} a_{11} & \cdots & a_{1i-1} & y_1 & a_{1i+1} & \cdots & a_{1n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & \cdots & a_{ni-1} & y_n & a_{ni+1} & \cdots & a_{nn} \end{vmatrix}}{det(A)}$$

Preuve: On développe le déterminant de la proposition par rapport à la i-ième colonne :

$$\begin{aligned} &\frac{1}{\det(A)} \begin{vmatrix} a_{11} & \cdots & a_{1i-1} & y_1 & a_{1i+1} & \cdots & a_{1n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & \cdots & a_{ni-1} & y_n & a_{ni+1} & \cdots & a_{nn} \end{vmatrix} \\ =&\frac{1}{\det(A)} \sum_{j=1}^{n} y_j (-1)^{i+j} \det(A(j,i)) \end{aligned}$$

Posons $\tilde{A} = (\tilde{a}_{ij})$. On a $X = A^{-1}Y = \dfrac{{}^t\tilde{A}Y}{\det(A)}$, donc :

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \frac{1}{\det(A)} \begin{pmatrix} \tilde{a}_{11} & \cdots & \tilde{a}_{n1} \\ \vdots & & \vdots \\ \tilde{a}_{1n} & \cdots & \tilde{a}_{nn} \end{pmatrix} \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

On obtient donc

$$\frac{1}{\det(A)}\sum_{j=1}^{n} y_j(-1)^{i+j}\det(A(j,i)) = \frac{1}{\det(A)}\sum_{j=1}^{n} y_j\tilde{a}_{ji} = x_i.$$

□

1.6 Rang d'une matrice

Définition 1.60 (Matrice extraite) *Soit $A \in \mathcal{M}_{p,q}(\mathbb{K})$, $A = (a_{ij})$. Soit $r \leqslant p$ et $s \leqslant q$. On dit que $A' \in \mathcal{M}_{r,s}(\mathbb{K})$ est une* **matrice extraite** *de A si A' est de la forme $A' = (a_{i_k j_l})_{\substack{1\leqslant k\leqslant r\\ 1\leqslant l\leqslant s}}$ où $1 \leqslant i_1 < \cdots < i_r \leqslant p$ et $1 \leqslant j_1 < \cdots < j_s \leqslant q$.*

Exemple 1.61 *Soit* $A = \begin{pmatrix} \mathbf{4} & -1 & \mathbf{3} & \mathbf{6} & 7 \\ 2 & 1 & 5 & -1 & 0 \\ \mathbf{3} & 6 & \mathbf{8} & \mathbf{2} & 4 \\ 1 & 1 & 2 & -1 & 1 \end{pmatrix}$

Alors la matrice $A' = \begin{pmatrix} 4 & 3 & 6 \\ 3 & 8 & 2 \end{pmatrix}$ *est une matrice extraite de A.*

Rappel 1.62 *Soit $A \in \mathcal{M}_{p,q}(\mathbb{K})$. Le rang de A noté $rg(A)$ est le rang de la famille des vecteurs colonnes de A. On a $rg(A) = rg({}^tA)$. Le rang de A est donc aussi le rang de la famille des vecteurs lignes de A.*

Théorème 1.63 *Soit $A \in \mathcal{M}_{p,q}(\mathbb{K})$. Alors A est de rang r si et seulement si :*

1. *Il existe une matrice carrée A' extraite de A de format (r,r) telle que $det(A') \neq 0$.*
2. *Pour tout $s > r$, toute matrice carrée A'' de format (s,s) extraite de A vérifie $det(A'') = 0$.*

PREUVE:

1. Montrons tout d'abord que si A est une matrice de $\mathcal{M}_{p,q}(\mathbb{K})$ de rang r, alors il existe une matrice carrée A' extraite de A de format (r,r) telle que $\det(A') \neq 0$:
   Comme $\mathrm{rg}(A) = \mathrm{rg}(C_1(A),\ldots,C_q(A)) = r$, on peut extraire de $(C_1(A),\ldots,C_q(A))$ une famille libre $(C_{j_1}(A),\ldots,C_{j_r}(A))$ de r vecteurs $(j_1 < \cdots < j_r)$.
   Notons A' la matrice de $\mathcal{M}_{p,r}(\mathbb{K})$ dont la ℓ-ième colonne est $C_{j_\ell}(A)$. Alors :
   $$r = \mathrm{rg}(A') = \mathrm{rg}({}^tA') = \mathrm{rg}(L_1(A'),\cdots,L_p(A'))$$
   Donc on peut extraire de la famille $(L_1(A'),\cdots,L_p(A'))$ une famille libre de r vecteurs $(L_{i_1}(A'),\cdots,L_{i_r}(A'))$ $(i_1 < \cdots < i_r)$. Soit A'' la matrice de $\mathcal{M}_r(\mathbb{K})$ dont la k-ième ligne est $L_{i_k}(A')$. En fait $A'' = (a_{i_k j_\ell})_{1\leqslant k,\ell\leqslant r}$ et $\mathrm{rg}(A'') = r$. Donc $\det(A'') \neq 0$.
2. Montrons maintenant que s'il existe une matrice carrée $A' \in \mathcal{M}_s(\mathbb{K})$ extraite de A telle que $\det(A') \neq 0$ alors $\mathrm{rg}(A) \geqslant s$:
   Tout d'abord $A' = (a_{i_k j_\ell})_{1\leqslant k,\ell\leqslant s}$ avec $i_1 < \cdots < i_s$ et $j_1 < \cdots < j_s$.
   Puisque $\det(A') \neq 0$ $\mathrm{rg}(A') = s$ et $(C_1(A'),\cdots,C_s(A'))$ est une famille libre. Mais cela implique que la famille $(C_{j_1}(A),\cdots,C_{j_s}(A))$ est libre. Donc $\mathrm{rg}(A) \geqslant s$.

Des points 1. et 2. on déduit de manière immédiate le théorème.

□

Exemple 1.64 *Soit* $A = \begin{pmatrix} 2 & 3 & 5 & 4 \\ -1 & 0 & -1 & 1 \\ 1 & 2 & 3 & 3 \end{pmatrix}$. *Alors :*

$$\begin{vmatrix} 2 & 3 & 5 \\ -1 & 0 & -1 \\ 1 & 2 & 3 \end{vmatrix} = \begin{vmatrix} 3 & 5 & 4 \\ 0 & -1 & 1 \\ 2 & 3 & 3 \end{vmatrix} = \begin{vmatrix} 2 & 3 & 4 \\ -1 & 0 & 1 \\ 1 & 2 & 3 \end{vmatrix} = \begin{vmatrix} 2 & 5 & 4 \\ -1 & -1 & 1 \\ 1 & 3 & 3 \end{vmatrix} = 0$$

Donc tous les déterminants de matrices $(3,3)$ *extraites de* A *sont nuls, donc* $rg(A) < 3$.

Pour montrer que A *est de rang* 2, *il suffit de trouver une matrice* $(2,2)$ *extraite de déterminant non nul. Or*

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

donc $rg(A) = 2$.

Chapitre 2

Réduction des endomorphismes

Dans toute la suite, $\mathbb{K}$ désigne le corps $\mathbb{R}$ ou $\mathbb{C}$ et E est un espace vectoriel sur $\mathbb{K}$ de dimension finie n D'autre part $\mathcal{L}(E)$ désigne le $\mathbb{K}$-espace vectoriel des endomorphismes de E.

2.1 Sous-espaces propres et diagonalisation

Étant donné un espace vectoriel E sur un corps $\mathbb{K}$ de dimension finie, les endomorphismes les plus simples sont les homothéties $f = \lambda \mathrm{Id}_E$. La matrice de ces endomorphismes est la même dans toutes les bases de E. C'est une matrice diagonale avec des λ sur la diagonale et des 0 partout ailleurs. Autrement dit il n'y a pas de direction privilégiée.

Étant donné maintenant un endomorphisme u quelconque, on va essayer dans ce chapitre de le "dévisser" en "cassant" l'espace E en une somme directe de sous-espaces stables (précisément les sous-espaces propres) sur lesquels u se comportera comme une homothétie. Si cela est possible en prenant une base dans chacun de ces sous-espaces et en les réunissant, on obtiendra une base de E dans laquelle la matrice de u sera diagonale.

2.1.1 Valeurs propres, vecteurs propres

Définition 2.1 *Soient $u \in \mathcal{L}(E)$ et $\lambda \in \mathbb{K}$. S'il existe un vecteur* **non nul** *x de E tel que $u(x) = \lambda x$, alors on dit que :*

1. *λ est une* **valeur propre** *de u.*
2. *x est un* **vecteur propre** *de u associé à λ.*

On appelle **spectre** *de u et on note* $\mathrm{Sp}(u)$ *l'ensemble des valeurs propres de u.*

Remarque 2.2
1. *Il résulte de la définition qu'un vecteur propre n'est jamais nul.*
2. *λ est une valeur propre de u si et seulement si $Ker(u - \lambda Id_E) \neq \{0\}$.*

En effet λ est une valeur propre de u si et seulement si il existe un vecteur non nul x de E tel que $u(x) = \lambda x$. Or $0 = u(x) - \lambda x = (u - \lambda \mathrm{Id}_E)(x)$. Il s'ensuit que λ est une valeur propre de u si et seulement si $\mathrm{Ker}(u - \lambda \mathrm{Id}_E) \neq \{0\}$.

Définition 2.3 *Soient $A \in \mathcal{M}_n(\mathbb{K})$ et $\lambda \in \mathbb{K}$. S'il existe un vecteur colonne* **non nul** *X de $\mathcal{M}_{n,1}(\mathbb{K})$ tel que $AX = \lambda X$, alors on dit que :*

1. *λ est une* **valeur propre** *de A.*
2. *X est un* **vecteur propre** *de A associé à λ.*

On appelle **spectre** *de A et on note $\mathrm{Sp}(A)$ l'ensemble des valeurs propres de A.*

Exercice de cours 2.1 *Soit $A = \begin{pmatrix} 3 & 2 \\ -4 & -3 \end{pmatrix}$. Déterminer le spectre de A. On pourra résoudre le système*

$$\begin{cases} 3x + 2y &= \lambda x \\ -4x - 3y &= \lambda y\,, \end{cases}$$

dépendant du paramètre λ.

Remarque 2.4 *Soient $\mathcal{B}$ une base de E, $u \in \mathcal{L}(E)$ et $A = Mat_{\mathcal{B}}(u)$. Alors $\mathrm{Sp}(u) = \mathrm{Sp}(A)$. Plus précisément :*

1. *λ est une valeur propre de u si et seulement si λ est une valeur propre de A.*
2. *x est un vecteur propre de u si et seulement si $X = Mat_{\mathcal{B}}(x)$ est un vecteur propre de A.*

Définition 2.5 *Soit $A = (a_{ij})_{1 \leqslant i,j \leqslant n}$ une matrice carrée de $\mathcal{M}_n(\mathbb{K})$. On appelle trace de A le scalaire :*

$$tr(A) := \sum_{i=1}^{n} a_{ii}$$

Proposition 2.6 *Soient $(A, B) \in \mathcal{M}_n(\mathbb{K})^2$. Alors $tr(AB) = tr(BA)$.*

Exercice de cours 2.2 *Donner une preuve de ce résultat. En cas de difficulté, commencer par le cas où $n = 2$.*

Proposition et définition 2.7 *Soient $u \in \mathcal{L}(E)$ et $\mathcal{B}$ une base de E. Alors le scalaire $tr(Mat_{\mathcal{B}}(u))$ ne dépend pas de la base choisie et s'appelle la* **trace** *de u. On le note $tr(u)$.*

Preuve: Soit $\mathcal{B}'$ une autre base, et soit P la matrice de passage de $\mathcal{B}$ à $\mathcal{B}'$. Soient $A = \mathrm{Mat}_{\mathcal{B}}(u)$ et $A' = \mathrm{Mat}_{\mathcal{B}'}(u)$. On a $A' = P^{-1}AP$, donc $\mathrm{tr}(A') = \mathrm{tr}(P^{-1}AP) = \mathrm{tr}(APP^{-1}) = \mathrm{tr}(A)$. $\square$

Exercice de cours 2.3 *Soit A la matrice de l'exercice 2.1, $A = \begin{pmatrix} 3 & 2 \\ -4 & -3 \end{pmatrix}$ et soit x un réel. Montrer que $det(A - xI_2) = x^2 - 1$.*

Proposition et définition 2.8 *Soit $u \in \mathcal{L}(E)$. La fonction*

$$\begin{array}{rccc} \chi_u & : & \mathbb{K} & \longrightarrow & \mathbb{K} \\ & & x & \longmapsto & det(u - xId_E) \end{array}$$

est une fonction polynomiale de degré n. De plus :

$$\chi_u(X) = (-1)^n X^n + (-1)^{n-1} tr(u) X^{n-1} + \cdots + det(u)$$

χ_u s'appelle le **polynôme caractéristique** *de u.*

Proposition et définition 2.9 *Soit $A \in \mathcal{M}_n(\mathbb{K})$. La fonction*

$$\chi_A \ : \ \begin{array}{rcl} \mathbb{K} & \longrightarrow & \mathbb{K} \\ x & \longmapsto & det(A - xI_n) \end{array}$$

est une fonction polynomiale de degré n. De plus :

$$\chi_A(X) = (-1)^n X^n + (-1)^{n-1} tr(A) X^{n-1} + \cdots + det(A)$$

χ_A s'appelle le **polynôme caractéristique** *de A.*

Remarques 2.10

— *Soient $\mathcal{B}$ une base de E, $u \in \mathcal{L}(E)$ et $A = Mat_{\mathcal{B}}(u)$. Alors $\chi_A = \chi_u$.*

— *Dans le cas $n = 2$, observer que χ_u est déterminé par $tr(u)$ et $det(u)$.*

Exercice de cours 2.4 *Vérifier que ces énoncés sont compatibles avec le calcul que vous avez fait dans l'exercice 2.3.*

Preuve des énoncés 2.8 et 2.9 : Soit $A = (a_{ij})_{1 \leqslant i,j \leqslant n}$. Alors :

$$\chi_A(x) = \det(A - xI_n) = \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) \prod_{i=1}^{n} (a_{\sigma(i)i} - x\delta_{\sigma(i)i})$$

Chacun des produits $\displaystyle\prod_{i=1}^{n} (a_{\sigma(i)i} - x\delta_{\sigma(i)i})$ est une fonction polynomiale en x de degré inférieur ou égal à n. Il en est de même de leur somme et χ_A est un polynôme de degré au plus n. Or :

$$\chi_A(x) = \sum_{\sigma \in \mathfrak{S}_n \setminus \{\mathrm{Id}\}} \varepsilon(\sigma) \prod_{i=1}^{n} (a_{\sigma(i)i} - x\delta_{\sigma(i)i}) + \prod_{i=1}^{n} (a_{ii} - x)$$

Remarquons que si $\sigma \in \mathfrak{S}_n \setminus \{\mathrm{Id}\}$, alors il existe $j \in \{1, \cdots, n\}$ tel que $\sigma(j) \neq j$. En posant $k = \sigma^{-1}(j)$, on a $k \neq j$ et $\sigma(k) \neq k$. Donc $\delta_{\sigma(j)j} = \delta_{\sigma(k)k} = 0$. Ainsi :

$$\prod_{i=1}^{n} (a_{\sigma(i)i} - x\delta_{\sigma(i)i}) = a_{\sigma(j)j} a_{\sigma(k)k} \prod_{\substack{i=1 \\ i \neq j,k}}^{n} (a_{\sigma(i)i} - x\delta_{\sigma(i)i})$$

et donc pour chaque $\sigma \in \mathfrak{S}_n \setminus \{\mathrm{Id}\}$, $\displaystyle\prod_{i=1}^{n} (a_{\sigma(i)i} - x\delta_{\sigma(i)i})$ est une fonction polynomiale de degré au plus $n-2$ et il en est de même de $\displaystyle\sum_{\sigma \in \mathfrak{S}_n \setminus \{\mathrm{Id}\}} \varepsilon(\sigma) \prod_{i=1}^{n} (a_{\sigma(i)i} - x\delta_{\sigma(i)i})$. Or comme $\displaystyle\prod_{i=1}^{n} (a_{ii} - x)$ est de degré n il s'ensuit que χ_A est de degré n et que les coefficients de x^n et x^{n-1} dans χ_u sont les coefficients de x^n et x^{n-1} dans $\displaystyle\prod_{i=1}^{n} (a_{ii} - x)$. Donc les coefficients de x^n et x^{n-1} dans χ_A sont respectivement $(-1)^n$ et $(-1)^{n-1} \displaystyle\sum_{i=1}^{n} a_{ii} = (-1)^{n-1} \mathrm{tr}(A)$.

Le terme constant de χ_A est $\chi_A(0) = \det(A)$. Ceci prouve l'énoncé 2.9. La preuve de 2.8 est une conséquence immédiate de la remarque 2.10 et de la proposition et définition 2.7. $\square$

En raison des remarques 2.4 et 2.10, on se contentera dans la majorité des cas d'énoncer les résultats dans le cadre des endomorphismes.

Proposition 2.11 *Soit $u \in \mathcal{L}(E)$. Alors $\lambda \in \mathrm{Sp}(u)$ si et seulement si λ est racine de χ_u.*

PREUVE: On a vu d'après la remarque 2.2 que λ est valeur propre de u si et seulement si $\mathrm{Ker}(u - \lambda \mathrm{Id}_E) \neq \{0\}$. Or les 4 assertions ci-dessous sont équivalentes :

1. $\mathrm{Ker}(u - \lambda \mathrm{Id}_E) \neq \{0\}$
2. $u - \lambda \mathrm{Id}_E$ n'est pas injective.
3. $u - \lambda \mathrm{Id}_E$ n'est pas un isomorphisme.
4. $\det(u - \lambda \mathrm{Id}_E) = 0$ (voir proposition 1.33).

Donc λ est valeur propre de u si et seulement si $\chi_u(\lambda) = 0$. $\square$

Exercice de cours 2.5 *Résoudre à nouveau l'exercice 2.1 en utilisant le polynôme caractéristique de A.*

Corollaire 2.12 *Tout endomorphisme u de E a au plus n valeurs propres.*

Définition 2.13 *Soient $u \in \mathcal{L}(E)$ et λ une valeur propre de u. On dira que λ est une valeur propre d'ordre m si λ est une racine d'ordre m de χ_u. L'entier m s'appelle la* **multiplicité** *de λ et sera noté $m(\lambda)$.*

Exercice de cours 2.6 *Déterminer le spectre et la multiplicité des valeurs propres des matrices suivantes : $B = \begin{pmatrix} 3 & 1 & 2 \\ 1 & 2 & 1 \\ -2 & -1 & -1 \end{pmatrix}$ et $C = \begin{pmatrix} 0 & -1 & 0 \\ 2 & 3 & 0 \\ -3 & -3 & 1 \end{pmatrix}$.*

2.1.2 Sous-espaces propres

Définition 2.14 *Soient $u \in \mathcal{L}(E)$ et λ une valeur propre de u. Le sous-espace vectoriel $E_\lambda(u) := Ker(u - \lambda Id_E)$ s'appelle le* **sous-espace propre** *associé à la valeur propre λ.*

Remarque 2.15

1. *Si λ est une valeur propre de u, on a toujours $E_\lambda(u) \neq \{0\}$.*
2. *$E_\lambda(u) \setminus \{0\}$ est l'ensemble des vecteurs propres de u.*
3. *Les sous-espaces propres de u sont stables par u (pour toute valeur propre λ, $u(E_\lambda(u)) \subset E_\lambda(u)$).*

Définition 2.16 *Soit $A \in \mathcal{M}_n(\mathbb{K})$ et λ une valeur propre de A. Le sous-espace vectoriel $M_\lambda(A) := Ker(A - \lambda I_n)$ de $\mathcal{M}_{1,n}(\mathbb{K})$ s'appelle le* **sous-espace propre** *associé à la valeur propre λ.*

Exercice de cours 2.7 *Calculer les sous-espaces propres des matrices A, B et C des exercices 2.1 et 2.6.*

Exercice de cours 2.8 *Montrer que $u(E_\lambda(u)) \subset E_\lambda(u)$ pour toute valeur propre λ de u.*

Proposition 2.17 *Soient $u \in \mathcal{L}(E)$ et λ une valeur propre de u. Alors :*

$$1 \leqslant \dim(E_\lambda(u)) \leqslant m(\lambda)$$

PREUVE: Posons $p = \dim(E_\lambda(u))$. Soit $(e_1, \cdots, e_p)$ une base de $E_\lambda(u)$ que l'on complète par $e_{p+1}, \cdots, e_n$ pour obtenir une base $\mathcal{B}$ de E. Dans ce cas on a :

$$\mathrm{Mat}_{\mathcal{B}}(u) = \begin{pmatrix} \lambda & 0 & \cdots & 0 & a_{1\ p+1} & \cdots & a_{1n} \\ 0 & \ddots & \ddots & \vdots & \vdots & & \vdots \\ \vdots & \ddots & \ddots & 0 & \vdots & & \vdots \\ 0 & \cdots & 0 & \lambda & a_{p\ p+1} & \cdots & a_{pn} \\ 0 & \cdots & \cdots & 0 & a_{p+1\ p+1} & \cdots & a_{p+1\ n} \\ \vdots & & & \vdots & \vdots & & \vdots \\ 0 & \cdots & \cdots & 0 & a_{n\ p+1} & \cdots & a_{nn} \end{pmatrix}$$

Le polynôme caractéristique s'écrit donc :

$$\chi_u(x) = (\lambda - x)^p \chi_M(x)$$

où $M = (a_{ij})_{p+1 \leqslant i,j \leqslant n}$. Donc $(\lambda - x)^p$ divise χ_u et $p \leqslant m(\lambda)$. □

Théorème 2.18 *Soit $u \in \mathcal{L}(E)$. Alors les sous-espaces propres de u sont en somme directe.*

PREUVE: Soient $\lambda_1, \cdots, \lambda_r$ les valeurs propres de u. Le but est de montrer que si $(v_1, \cdots, v_r) \in E_{\lambda_1}(u) \times \cdots \times E_{\lambda_r}(u)$ et $v_1 + \cdots + v_r = 0$ alors $v_i = 0$ pour tout $i \in \{1, \cdots, r\}$.

On va procéder par récurrence en montrant que $E_{\lambda_1}(u), \cdots, E_{\lambda_j}(u)$ sont en somme directe pour tout $j \in \{1, \cdots, r\}$.

Pour $j = 1$, la propriété est vraie.

On suppose maintenant la propriété vraie au rang $j \in \{1, \cdots, r-1\}$. Montrons qu'elle est vraie au rang $j + 1$.

Soit $(v_1, \cdots, v_{j+1}) \in E_{\lambda_1}(u) \times \cdots \times E_{\lambda_{j+1}}(u)$ tel que $\displaystyle\sum_{i=1}^{j+1} v_i = 0$. Alors :

$$u\left(\sum_{i=1}^{j+1} v_i\right) = \sum_{i=1}^{j+1} u(v_i) = \sum_{i=1}^{j+1} \lambda_i v_i = 0$$

Donc :

$$0 = \sum_{i=1}^{j+1} \lambda_i v_i - \lambda_{j+1} \sum_{i=1}^{j+1} v_i = \sum_{i=1}^{j+1} (\lambda_i - \lambda_{j+1}) v_i = \sum_{i=1}^{j} (\lambda_i - \lambda_{j+1}) v_i$$

Or pour tout $i \in \{1, \cdots, j\}$, $(\lambda_i - \lambda_{j+1})v_i \in E_{\lambda_i}(u)$. Comme par hypothèse de récurrence, $E_{\lambda_1}(u), \cdots, E_{\lambda_j}(u)$ sont en somme directe, il s'ensuit que pour tout $i \in \{1, \cdots, j\}$, $(\lambda_i - \lambda_{j+1})v_i = 0$. De plus $\lambda_i - \lambda_{j+1} \neq 0$ pour tout $i \in \{1, \cdots, j\}$. Donc :

$$v_1 = \cdots = v_j = 0$$

et $\displaystyle\sum_{i=1}^{j+1} v_i = v_{j+1} = 0$. Donc $E_{\lambda_1}(u), \cdots, E_{\lambda_{j+1}}(u)$ sont en somme directe. Par récurrence sur j la propriété est vraie pour tout $j \in \{1, \cdots, r\}$. □

Exercice de cours 2.9 *Vérifier que les espaces propres obtenus pour A, B et C dans l'exercice 2.7 sont en somme directe.*

2.1.3 Endomorphismes et matrices diagonalisables

Théorème 2.19 *Soit $u \in \mathcal{L}(E)$. Alors les propositions suivantes sont équivalentes :*

1. *Il existe une base $\mathcal{B}$ de E telle que $Mat_{\mathcal{B}}(u)$ soit diagonale.*
2. *Il existe une base $\mathcal{B}$ de E constituée de vecteurs propres de u.*
3. *E est la somme directe des sous-espaces propres de u.*
4. *χ_u est scindé sur $\mathbb{K}$ (i.e. χ_u a toute ses racines dans $\mathbb{K}$) et pour toute valeur propre de u, $\dim(E_\lambda(u)) = m(\lambda)$.*

Preuve: $1 \Rightarrow 2$. Il existe une base $\mathcal{B} = (e_1, \cdots, e_n)$ de E telle que :

$$\mathrm{Mat}_{\mathcal{B}}(u) = \begin{pmatrix} \alpha_1 & & 0 \\ & \ddots & \\ 0 & & \alpha_n \end{pmatrix}$$

Par définition de la matrice d'un endomorphisme dans une base, on a pour tout $i \in \{1, \cdots, n\}$, $u(e_i) = \alpha_i e_i$ ce qui prouve 2.

$2 \Rightarrow 3$. Soit $\mathcal{B} = (e_1, \cdots, e_n)$ une base formée de vecteurs propres de u. D'après le théorème 2.18 les sous-espaces propres sont en somme directe et $\mathcal{B} = (e_1, \cdots, e_n)$ est une famille libre de $\displaystyle\bigoplus_{i=1}^{r} E_{\lambda_i}(u)$. Donc :

$$\dim\left(\bigoplus_{i=1}^{r} E_{\lambda_i}(u)\right) \geqslant n = \dim(E)$$

et $E = \displaystyle\bigoplus_{i=1}^{r} E_{\lambda_i}(u)$.

$3 \Rightarrow 1$ et $3 \Rightarrow 4$. Soient $\lambda_1, \cdots, \lambda_r$ les valeurs propres de u. Posons $n_i = \dim(E_{\lambda_i}(u))$. Soient $\mathcal{B}_1, \cdots, \mathcal{B}_r$ des bases respectivement de $E_{\lambda_1}(u), \cdots, E_{\lambda_r}(u)$. Comme par hypothèse $E =$

$\bigoplus_{i=1}^{r} E_{\lambda_i}(u)$, $\mathcal{B} = \bigcup_{i=1}^{r} \mathcal{B}_i$ est une base de E constituée de vecteurs propres. Donc :

$$\operatorname{Mat}_{\mathcal{B}}(u) = \begin{pmatrix} \lambda_1 & 0 & \cdots & \cdots & \cdots & \cdots & 0 \\ 0 & \ddots & \ddots & & & & \vdots \\ \vdots & \ddots & \lambda_1 & \ddots & & & \vdots \\ \vdots & & \ddots & \ddots & \ddots & & \vdots \\ \vdots & & & \ddots & \lambda_r & \ddots & \vdots \\ \vdots & & & & \ddots & \ddots & 0 \\ 0 & \cdots & \cdots & \cdots & \cdots & 0 & \lambda_r \end{pmatrix}$$

ce qui prouve 1.

D'autre part $\chi_u(x) = (\lambda_1 - x)^{n_1} \cdots (\lambda_r - x)^{n_r}$. Il s'ensuit que le polynôme est scindé et que la multiplicité des valeurs propres vérifie $m(\lambda_i) = n_i = \dim(E_{\lambda_i}(u))$ ce qui prouve 4.

$4 \Rightarrow 3$ On suppose que χ_u est scindé sur $\mathbb{K}$ et que pour toute valeur propre de u, $\dim(E_\lambda(u)) = m(\lambda)$. Donc :

$$\chi_u(x) = \prod_{i=1}^{r} (\lambda_i - x)^{m(\lambda_i)}$$

et $\dim(E_{\lambda_i}(u)) = m(\lambda_i)$. On a donc :

$$\dim\left(\bigoplus_{i=1}^{r} E_{\lambda_i}(u)\right) = \sum_{i=1}^{r} \dim(E_{\lambda_i}) = \sum_{i=1}^{r} m(\lambda_i) = n$$

Par conséquent $\bigoplus_{i=1}^{r} E_{\lambda_i}(u) = E$. □

Exercice de cours 2.10 *Pour les matrices A, B, C ci-dessus et pour chaque valeur propre λ, déterminer* $\dim E_\lambda$ *et $m(\lambda)$.*

Définition 2.20 *Soit $u \in \mathcal{L}(E)$. On dit que u est* **diagonalisable** *si l'une des 4 conditions équivalentes du théorème 2.19 est vérifiée.*

Exercice de cours 2.11 *Les endomorphismes de $\mathbb{K}^n$ associés aux matrices A, B et C ci-dessus sont-ils diagonalisables ?*

Définition 2.21 *Soit $A \in \mathcal{M}_n(\mathbb{K})$. On dit que A est* **diagonalisable** *si A est semblable à une matrice diagonale. Autrement dit A est diagonalisable s'il existe une matrice $P \in GL_n(\mathbb{K})$ et une matrice diagonale $D \in \mathcal{M}_n(\mathbb{K})$ telles que : $D = P^{-1}AP$.*

Remarque 2.22 *Lorsqu'on voudra insister sur le fait que la matrice P est à coefficients dans $\mathbb{K}$, on précisera que A est diagonalisable* sur $\mathbb{K}$. *Ainsi, par exemple, une matrice à coefficients dans $\mathbb{R}$ pourra être diagonalisable sur $\mathbb{C}$ mais non sur $\mathbb{R}$: ceci signifie qu'il existe $P \in GL_n(\mathbb{C})$ telle que $P^{-1}AP$ est une matrice diagonale à coefficients complexes, mais qu'il n'existe pas de matrice $P \in GL_n(\mathbb{R})$ telle que $P^{-1}AP$ est une matrice diagonale à coefficients réels.*

Exercice de cours 2.12 *Donner un exemple d'une matrice carrée d'ordre 2 à coefficients réels dont les valeurs propres sont deux nombres complexes non réels (on pourra d'abord choisir un polynôme χ ayant deux racines complexes, puis une matrice dont le polynôme caractéristique est χ). Montrer que cette matrice est diagonalisable sur $\mathbb{C}$ mais pas sur $\mathbb{R}$.*

Exercice de cours 2.13 (Exercice facultatif) *Soit A une matrice à coefficients réels et λ une valeur propre réelle de A. Montrer que les espaces propres réel et complexe de A pour la valeur propre λ ont la même dimension.*

Soit $A \in \mathcal{M}_n(\mathbb{R})$ de polynôme caractéristique scindé sur $\mathbb{R}$. Montrer que A est diagonalisable sur $\mathbb{R}$ si et seulement si elle est diagonalisable sur $\mathbb{C}$.

Remarque 2.23

1. *Soient $\mathcal{B}$ une base de E, $u \in \mathcal{L}(E)$ et $A = Mat_{\mathcal{B}}(u)$. Alors u est diagonalisable si et seulement si A est diagonalisable.*
2. *Si A est diagonalisable alors il existe $P \in GL_n(\mathbb{K})$ telle que $D = P^{-1}AP$ est diagonale. Dans ce cas $(C_1(P), \cdots, C_n(P))$ est une base de $\mathcal{M}_{n,1}(\mathbb{K})$ constituée de vecteurs propres de A.*

Corollaire 2.24 *Soit $u \in \mathcal{L}(E)$. Alors si χ_u est scindé sur $\mathbb{K}$ et si toutes les racines de χ_u sont simples alors u est diagonalisable.*

Exercice de cours 2.14 *Donner un exemple d'endomorphisme u diagonalisable mais tel que les racines de χ_u ne sont pas simples.*

2.1.4 Calcul des puissances d'une matrice

Soit $A \in \mathcal{M}_k(\mathbb{K})$. Si A est diagonalisable alors il existe une matrice diagonale $D \in \mathcal{M}_k(\mathbb{K})$ et $P \in GL_k(\mathbb{K})$ telles que $D = P^{-1}AP$. Or on sait calculer D^n. En effet si $D = \begin{pmatrix} \alpha_1 & & 0 \\ & \ddots & \\ 0 & & \alpha_k \end{pmatrix}$, alors $D^n = \begin{pmatrix} \alpha_1^n & & 0 \\ & \ddots & \\ 0 & & \alpha_k^n \end{pmatrix}$. D'autre part on peut montrer par récurrence que :

$$D^n = (P^{-1}AP)^n = P^{-1}A^nP$$

d'où :

$$A^n = PD^nP^{-1}$$

Exercice de cours 2.15 *Calculer les puissances de la matrice A de l'exercice 2.1.*

2.2 Trigonalisation et théorème de Cayley-Hamilton

2.2.1 Endomorphismes trigonalisables

Définition 2.25 *Soit $u \in \mathcal{L}(E)$. On dit que u est* **trigonalisable** *s'il existe une base $\mathcal{B}$ de E telle que $Mat_{\mathcal{B}}(u)$ est triangulaire supérieure.*

Définition 2.26 *Soit $A \in \mathcal{M}_n(\mathbb{K})$. On dit que A est* **trigonalisable** *s'il existe $P \in GL_n(\mathbb{K})$ telle que $P^{-1}AP$ soit triangulaire supérieure.*

Remarque 2.27 *Soient $\mathcal{B}$ une base de E, $u \in \mathcal{L}(E)$ et $A = Mat_{\mathcal{B}}(u)$. Alors u est trigonalisable si et seulement si A est trigonalisable.*

Théorème 2.28 *Soit $u \in \mathcal{L}(E)$. Alors u est trigonalisable si et seulement si χ_u est scindé dans $\mathbb{K}[X]$ (i.e. χ_u se décompose en un produit de polynômes du premier degré).*

Preuve:

— ($\Rightarrow$) Si u est trigonalisable alors il existe une base $\mathcal{B}$ telle que :

$$\mathrm{Mat}_{\mathcal{B}}(u) = \begin{pmatrix} a_{11} & \cdots & \cdots & a_{1n} \\ 0 & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & 0 & a_{nn} \end{pmatrix}$$

Notant $T = \mathrm{Mat}_{\mathcal{B}}(u)$, on en déduit immédiatement que $\chi_u(x) = \det(T - xI_n) = \prod_{i=1}^{n}(a_{ii} - x)$ ce qui signifie que χ_u est scindé.

— ($\Leftarrow$) On suppose que χ_u est scindé et on veut montrer que u est trigonalisable.
Nous allons raisonner par récurrence sur $n = \dim(E)$. Si $n = 1$ l'affirmation est immédiate. En supposant que pour $n \geqslant 2$ l'affirmation est vraie quand $\dim(E) = n-1$, montrons qu'elle est encore vraie quand $\dim(E) = n$.
Donc soit E un $\mathbb{K}$-espace vectoriel de dimension n et soit $u \in \mathcal{L}(E)$ tel que χ_u soit scindé. Cela entraîne que $\mathrm{Sp}(u) \neq \emptyset$. Soit $\lambda \in \mathrm{Sp}(u)$ et soit e_1 un vecteur propre associé. Complétons e_1 par des vecteurs $e_2, \cdots, e_n$ pour que $\mathcal{B} = (e_1, \cdots, e_n)$ soit une base de E. Alors :

$$\mathrm{Mat}_{\mathcal{B}}(u) = M = \begin{pmatrix} \lambda & m_{12} & \cdots & \cdots & \cdots & m_{1n} \\ 0 & & & & & \\ \vdots & & & & & \\ \vdots & & & M_1 & & \\ \vdots & & & & & \\ 0 & & & & & \end{pmatrix}$$

Donc $\chi_u(x) = (\lambda - x)\det(M_1 - xI_{n-1})$. D'autre part on a $E = \mathbb{K}e_1 \oplus F$ où $F = \mathrm{Vect}(e_2, \cdots, e_n)$. Soit g l'endomorphisme de E défini par :

$$\mathrm{Mat}_{\mathcal{B}}(g) = \begin{pmatrix} 0 & 0 & \cdots & \cdots & \cdots & 0 \\ 0 & & & & & \\ \vdots & & & & & \\ \vdots & & & M_1 & & \\ \vdots & & & & & \\ 0 & & & & & \end{pmatrix}$$

Remarquons que pour tout $i \in \{2, \cdots, n\}$:

$$u(e_i) = m_{1i}e_1 + g(e_i)$$

Ainsi pour tout $v \in F$ il existe $a \in \mathbb{K}$ tel que :

$$u(v) = ae_1 + g(v)$$

D'autre part on a évidemment $g(F) \subset F$. Donc $g_{|F} \in \mathcal{L}(F)$ et $\mathrm{Mat}_{\{e_2,\cdots,e_n\}}(g_{|F}) = M_1$. Or $\chi_u(x) = (\lambda - x)\det(M_1 - xI_{n-1}) = (\lambda - x)\chi_{M_1}(x)$. Comme χ_u est scindé, il s'ensuit que χ_{M_1} est scindé et on peut appliquer l'hypothèse de récurrence. Il existe une base $(f_2, \cdots, f_n)$ de F telle que :

$$\mathrm{Mat}_{\{f_2,\cdots,f_n\}}(g_{|F}) = \begin{pmatrix} \alpha_{22} & \alpha_{23} & \cdots & \alpha_{2n} \\ 0 & \alpha_{33} & \cdots & \alpha_{3n} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & \alpha_{nn} \end{pmatrix}$$

Comme $E = \mathbb{K}e_1 \oplus F$, $\mathcal{B}' = \{e_1, f_2, \cdots, f_n\}$ est une base de E. Pour tout $j \in \{2, \cdots, n\}$ il existe $a_{1j} \in \mathbb{K}$ tel que :

$$u(f_j) = a_{1j}e_1 + g(f_j) = a_{1j}e_1 + \sum_{i=2}^{j} \alpha_{ij} f_i$$

Il s'ensuit que $\mathrm{Mat}_{\mathcal{B}'}(u)$ est triangulaire supérieure.

□

On donne une autre preuve de l'implication χ_u scindé $\Rightarrow u$ trigonalisable.

Définition 2.29 (Espace vectoriel quotient, ensemble) *Soit E un $\mathbb{K}$-espace vectoriel et $F \subset E$ un sous-espace vectoriel. Soit $u \in E$.*

— *La classe de u modulo F, notée $\overline{u}$, est le sous-ensemble $u + F$ des vecteurs de E de la forme $u + w$ avec $w \in F$: $\overline{u} = u + F = \{u + w, w \in F\}$.*

— *L'ensemble quotient E/F est l'ensemble des $\overline{u}$, lorsque u parcourt E.*

On aura donc compris que les éléments de E/F sont par définition des sous-ensembles de E. Le fait que des éléments soient des ensembles est une des raisons qui font que la notion d'espace vectoriel quotient est difficile à appréhender !

Exercice de cours 2.16 *Pour $u, v \in E$, montrer en utilisant le premier point, que $\overline{u} = \overline{v}$ si et seulement si $u - v \in F$.*

Exemple 2.30 *Soit $E = \mathbb{R}^2$ et $F = \{(x, y) : y = x\}$. On a $\overline{(1,0)} = \overline{(0,-1)} = \{(1+t, t), t \in \mathbb{R}\}$ mais $\overline{(1,0)} \neq \overline{(0,1)}$. Graphiquement, les éléments de E/F sont les parallèles à la droite F d'équation $y = x$.*

Définition 2.31 (Espace vectoriel quotient, structure d'espace vectoriel) *Soit E un $\mathbb{K}$-espace vectoriel et $F \subset E$ un sous-espace vectoriel. Il existe une unique structure d'espace vectoriel sur l'ensemble quotient E/F telle que $\pi : E \to E/F$ soit une application linéaire.*

Preuve: En effet, on a nécessairement

$$\lambda \cdot \overline{v} = \overline{\lambda v} \text{ et } \overline{u+v} = \overline{u} + \overline{v} \tag{2.1}$$

si π est linéaire. Réciproquement, la formule (2.1) définit une structure d'espace vectoriel sur E/F. En effet, on vérifie qu'elle ne dépend pas du choix de u et de v, et qu'elle définit donc de manière univoque $\lambda \cdot \alpha$ et $\alpha + \beta$ pour $\lambda \in \mathbb{K}$ et $\alpha, \beta \in E/F$. Pour cela, il faut montrer que si $\overline{u}' = \alpha$ et $\overline{v}' = \beta$, alors $\overline{\lambda u'} = \overline{\lambda u}$ et $\overline{u' + v'} = \overline{u+v}$. Or, on a, par l'exercice précédent, l'existence de $w, t \in F$ tels que $u' = u + w$ et $v' = v + t$. Ceci donne $\overline{\lambda u'} = \overline{\lambda(u+w)} = \overline{\lambda u + \lambda w} = \overline{\lambda u}$ car $\lambda w \in F$. De même, $\overline{u' + v'} = \overline{u + v + w + t} = \overline{u+v}$ car $w + t \in F$.

On vérifie alors de manière assez automatique que tous les axiomes des espaces vectoriels sont vérifiés avec ces deux lois. □

Lemme 2.32 (Espace vectoriel quotient, base) *Soit E un $\mathbb{K}$-espace vectoriel et $F \subset E$ un sous-espace vectoriel de dimension d. Soit $\mathcal{B} = (e_1, \cdots, e_n)$ une base de E avec $(e_1, \cdots, e_d)$ une base de F. Alors $\overline{\mathcal{B}} = (\overline{e_{d+1}}, \cdots, \overline{e_n})$ est une base de E/F. En particulier,* $\dim(E/F) = \dim(E) - \dim(F)$.

Inversement, soit $(e_1, \cdots, e_n)$ une famille de vecteurs de E ayant les propriétés suivantes :

- *$(e_1, \cdots, e_d)$ est une base de F.*
- *$(\overline{e_{d+1}}, \cdots, \overline{e_n})$ est une base de E/F.*

Alors $(e_1, \cdots, e_n)$ est une base de E.

Preuve: Soit $\alpha \in E/F$ et soit $u \in E$ tel que $\overline{u} = \alpha$. Alors il existe des scalaires λ_i tels que $u = \sum_{i \geq 1} \lambda_i e_i$. Comme $\overline{e_i} = \overline{0}$ si $i \leq d$, on en déduit que $\overline{u} = \sum_{i \geq d+1} \lambda_i \overline{e_i}$. Ainsi $\alpha = \sum_{i \geq d+1} \lambda_i \overline{e_i}$. Donc $(\overline{e_{d+1}}, \cdots, \overline{e_n})$ est une famille génératrice de E/F.

Par ailleurs, si on a une relation $\sum_{i \geq d+1} \lambda_i \overline{e_i} = 0$, alors on en déduit que $\sum_{i \geq d+1} \lambda_i e_i \in F$ par définition de E/F, donc il existe $\lambda_1, \cdots, \lambda_d$ tels que $\sum_{i \geq d+1} \lambda_i e_i = -\sum_{j \leq d} \lambda_j e_j$. Alors $\sum_{i \geq 1} \lambda_i e_i = 0$, et donc tous les λ_i sont nuls, puisque $\mathcal{B}$ est une famille libre.

Pour la partie réciproque, il suffit de montrer que $(e_1, \cdots, e_n)$ est une famille libre. Or, si $\sum \lambda_i e_i$ est une combinaison linéaire nulle, en appliquant π on en déduit que $\sum_{i > d} \overline{e_i} = 0$. Par le deuxième point, $\lambda_i = 0$ si $i > d$. Mais on a alors $\sum_{i \leq d} \lambda_i e_i = 0$, ce qui donne $\lambda_i = 0$ pour $i \leq d$ par le premier point. □

Proposition 2.33 (Passage au quotient) *Soit $f : E \to G$ une application linéaire et $F \subset E$ un sous-espace vectoriel. On peut former un diagramme commutatif*
$$\begin{tikzcd} E \arrow[r, "f"] \arrow[d, "\pi"] & G \\ E/F \arrow[ur, dashed, "\overline{f}"'] & \end{tikzcd}$$
si et seulement si $f(F) = \{0\}$, soit $F \subset \ker(f)$. Dans ce cas, il existe une unique application $\overline{f} : E/F \to G$ telle que ce diagramme commute.

La proposition "on peut former un diagramme commutatif ..." signifie en des termes plus habituels "il existe $\overline{f} : E/F \to G$ telle que $f = \overline{f} \circ \pi$".

Preuve: Supposons qu'on peut définir $\overline{f}$. Soit alors $v \in F$, on a d'après l'exercice 2.16 $\overline{v} = \overline{0}$. On en déduit que $f(v) = \overline{f}(\pi(v)) = \overline{f}(\overline{0}) = 0$.

Réciproquement, si $f(F) = \{0\}$, alors la formule $\overline{f}(\alpha) = f(v)$ si $\overline{v} = \alpha$ définit bien une application $\overline{f} : E/F \to G$ (ie, $\overline{f}(\alpha)$ ne dépend pas du choix de v tel que $\overline{v} = \alpha$), et c'est la seule qui fait commuter le diagramme. $\square$

On revient à la preuve du théorème. Soit donc $u : E \to E$ telle que χ_u est scindé. Soit $\lambda \in \mathbb{K}$ tel que $\chi_u(\lambda) = 0$. Alors λ est une valeur propre de u, soit donc $v \in E$ tel que $u(v) = \lambda v$. On définit $F = \mathbb{K} \cdot v$ et on forme le diagramme
$$\begin{tikzcd} E \arrow[r, "u"] \arrow[dr, "u'"'] & E \arrow[d, "\pi"] \\ & E/F \end{tikzcd}.$$
Comme $u(v) = \lambda v$, on a $u(F) \subset F$ et donc $u'(F) = 0$. D'après le Proposition 2.33, on obtient un diagramme
$$\begin{tikzcd} E \arrow[r, "u'"] \arrow[d, "\pi"] & E/F \\ E/F \arrow[ur, dashed, "\overline{u}"'] & \end{tikzcd}$$
En mettant ensemble les deux diagrammes précédent, on obtient
$$\begin{tikzcd} E \arrow[r, "u"] \arrow[d, "\pi"'] \arrow[dr, "u'"'] & E \arrow[d, "\pi"] \\ E/F \arrow[r, "\overline{u}"'] & E/F \end{tikzcd}$$
On a alors le lemme suivant, qui relie la matrice de $\overline{u}$ à celle de u :

Lemme 2.34 *Soit $\mathcal{B} = (e_1, \cdots, e_n)$ une base de E avec $e_1 = v$. Soit $\overline{\mathcal{B}} = (\overline{e_2}, \cdots, \overline{e_n})$ une base de E/F, par le Lemme 2.32. Alors, $Mat_{\overline{\mathcal{B}}}(\overline{u})$ est la matrice extraite de $Mat_{\mathcal{B}}(u)$ obtenue en supprimant la première ligne et la première colonne.*

Preuve: Par le diagramme commutatif définissant $\overline{u}$, si $u(e_j) = \sum_i a_{i,j} e_i$, alors $\overline{u}(\overline{e_j}) = \sum_{i \geq 2} a_{i,j} \overline{e_i}$. Ceci montre que la matrice de $\overline{u}$ est la matrice extraite annoncée. $\square$

On termine alors la preuve du théorème comme suit : d'après le Lemme 2.34, la matrice de u dans la base $\mathcal{B}$ s'écrit

$$\mathrm{Mat}_{\mathcal{B}}(u) = \begin{pmatrix} \lambda & a_{12} & \cdots & \cdots & \cdots & a_{1n} \\ 0 & & & & & \\ \vdots & & & & & \\ \vdots & & & \mathrm{Mat}_{\overline{\mathcal{B}}}(\overline{u}) & & \\ \vdots & & & & & \\ 0 & & & & & \end{pmatrix} \tag{2.2}$$

En développant le déterminant de $\mathrm{Mat}_{\mathcal{B}}(u) - xI_n$ par rapport à la première colonne, on en déduit que $\chi_u(x) = (\lambda - x)\chi_{\overline{u}}(x)$. Comme χ_u est scindé, si on note μ_i ses racines comptées avec multiplicités, de sorte que $\chi_u(x) = \prod_i(\mu_i - x)$, on a $\prod_i(\mu_i - \lambda) = 0$. On en déduit l'existence d'un entier i tel que $\mu_i = \lambda$. Alors, $\chi_{\overline{u}}(x) = \prod_{j\neq i}(\mu_j - x)$. En particulier, $\chi_{\overline{u}}$ est scindé.

On a $\dim(E/F) = \dim(E) - 1$; en argumentant par récurrence sur $\dim(E)$, on peut donc supposer qu'il existe une base $\overline{\mathcal{B}'} = (\overline{e'_2}, \cdots, \overline{e'_n})$ de E/F telle que $\mathrm{Mat}_{\overline{\mathcal{B}'}}(\overline{u})$ soit triangulaire supérieure. Si on pose $\mathcal{B}' = (e_1, e'_2, \cdots, e'_n)$, le Lemme 2.32 montre que $\mathcal{B}'$ est une base de E. De plus, d'après la relation (2.2) appliquée avec $\mathcal{B}'$, la matrice de u dans la base $\mathcal{B}'$ est triangulaire supérieure.

Exercice de cours 2.17 *Soit A une matrice carrée d'ordre 2 à coefficients réels et soit χ_A son polynôme caractéristique. Soit Δ le discriminant de χ_A.*

1. *Si $\Delta < 0$, montrer que A est diagonalisable sur $\mathbb{C}$ mais pas sur $\mathbb{R}$.*
2. *Si $\Delta = 0$, montrer que A est trigonalisable sur $\mathbb{R}$. Donner un exemple où A n'est pas diagonalisable sur $\mathbb{R}$.*
3. *Si $\Delta > 0$, montrer que A est diagonalisable sur $\mathbb{R}$.*

Corollaire 2.35 *Soit E un espace vectoriel sur $\mathbb{C}$ de dimension finie. Alors tout endomorphisme de E est trigonalisable.*

PREUVE: C'est une conséquence du fait que $\mathbb{C}$ est algébriquement clos (i.e. tout polynôme à coefficients dans $\mathbb{C}$ est scindé). $\square$

Remarque 2.36 *Si un endomorphisme u est trigonalisable alors sa matrice dans une base appropriée est triangulaire supérieure et sur la diagonale principale figurent les valeurs propres de u, chacune d'elle y figurant autant de fois que son ordre de multiplicité dans χ_u.*

Corollaire 2.37 *Si χ_u est scindé, c'est-à-dire si $\chi_u(x) = \prod_{i=1}^{n}(\lambda_i - x)$ alors $tr(u) = \sum_{i=1}^{n}\lambda_i$ et $det(u) = \prod_{i=1}^{n}\lambda_i$.*

2.2.2 Polynômes d'endomorphismes

Soit u un endomorphisme de E. On rappelle que pour tout $i \in \mathbb{N}^\star$, u^i est défini par récurrence de la façon suivante :

1. $u^0 = \mathrm{Id}_E$.
2. Pour tout entier $i \geqslant 1$, $u^i = u \circ u^{i-1}$.

On a alors les propriétés suivantes, pour tout $(p, q) \in \mathbb{N}$ on a :

1. $u^p \circ u^q = u^{p+q} = u^q \circ u^p$.
2. $(u^p)^q = u^{pq} = (u^q)^p$

Définition 2.38 *Soit u un endomorphisme de E. On appelle polynôme de l'endomorphisme u tout endomorphisme de la forme :*

$$P(u) = \sum_{i=0}^{k} a_i u^i$$

où $P(X) = \sum_{i=0}^{k} a_i X^i$ est un polynôme de $\mathbb{K}[X]$.

Remarque 2.39 *De même si A est une matrice de $\mathcal{M}_n(\mathbb{K})$ et $P(X) = \sum_{i=0}^{k} a_i X^i$ un polynôme de $\mathbb{K}[X]$, on pose :*

$$P(A) = \sum_{i=0}^{k} a_i A^i$$

Exercice de cours 2.18 *Calculer $P(A), P(B)$ et $P(C)$ pour $P(X) = X^2 + X + 1$, lorsque A, B, C sont les matrices des exercices 2.1 et 2.6.*

Proposition 2.40 *Soient u et v deux endomorphismes de E, P et Q deux polynômes de $\mathbb{K}[X]$ et $(\lambda, \mu) \in \mathbb{K}^2$. Alors :*

1. $(\lambda P + \mu Q)(u) = \lambda P(u) + \mu Q(u)$.
2. $(PQ)(u) = P(u) \circ Q(u) = Q(u) \circ P(u) = (QP)(u)$.
3. *Si $\mathcal{B}$ est une base de E, alors $Mat_{\mathcal{B}}(P(u)) = P(Mat_{\mathcal{B}}(u))$.*

Remarque 2.41 *Cette proposition ne dit rien d'autre que pour tout $u \in \mathcal{L}(E)$ l'application :*

$$\begin{array}{rccc} \Phi_u & : & \mathbb{K}[X] & \longrightarrow & \mathcal{L}(E) \\ & & P(X) & \longmapsto & P(u) \end{array}$$

est un morphisme de $\mathbb{K}$-algèbres. Plus précisément cela signifie que :

— *Φ_u est une application linéaire du $\mathbb{K}$-espace vectoriel $\mathbb{K}[X]$ dans le $\mathbb{K}$-espace vectoriel $\mathcal{L}(E)$.*
— $\forall (P, Q) \in \mathbb{K}[X]^2$, $\Phi_u(PQ) = \Phi_u(P) \circ \Phi_u(Q)$.
— $\Phi_u(1) = u^0 = Id_E$.

D'autre part, le point 2. de la proposition dit que si $P(X) = (X - a_1)^{\alpha_1} \cdots (X - a_r)^{\alpha_r}$ alors :

$$P(u) = (u - a_1 Id_E)^{\alpha_1} \circ \cdots \circ (u - a_r Id_E)^{\alpha_r}$$

2.2.3 Théorème de Cayley-Hamilton

Théorème 2.42 (Cayley-Hamilton) *Soit u un endomorphisme de E, alors :*

$$\chi_u(u) = 0$$

Commençons par traiter un exemple : soit M la matrice ayant tous les coefficients nuls sauf sur la sur-diagonale, où ils sont égaux à 1 : $M = (a_{i,j})$ avec $a_{i,i+1} = 1$ et $a_{i,j} = 0$ pour $j \neq i+1$. On a alors $\chi_M(x) = (-x)^n$. Montrons que $M^n = 0$. Pour cela, soit $V_i \subset \mathbb{K}^n$ le sous-espace vectoriel engendré par $e_1, \cdots, e_i$, avec la convention $V_0 = \{0\}$. On a $M \cdot e_i = e_{i-1}$ donc $M \cdot V_i \subset V_{i-1}$. On en déduit que

$$M^n \cdot V_n \subset M^{n-1} \cdot V_{n-1} \subset M^{n-2} \cdot V_{n-2} \subset \cdots ,$$

de telle sorte que finalement $M^n \cdot V_n = \{0\}$ et donc $M^n = 0$.

Preuve: Comme $\mathcal{M}_n(\mathbb{R}) \subset \mathcal{M}_n(\mathbb{C})$, il suffit de montrer le théorème dans le cas où $\mathbb{K} = \mathbb{C}$. Ainsi u est trigonalisable. Donc soit $\mathcal{B} = \{e_1, \cdots, e_n\}$ une base dans laquelle $T = \mathrm{Mat}_{\mathcal{B}}(u)$ est triangulaire supérieure. Notons $\lambda_1, \cdots, \lambda_n$, les valeurs propres de u telles que :

$$T = \begin{pmatrix} \lambda_1 & & \\ O & \ddots & \\ & & \lambda_n \end{pmatrix}$$

Posons $V_0 = \{0\}$ et pour tout $k \in \{1, \cdots, n\}$ soit $V_k := \mathrm{Vect}(e_1, \cdots, e_k)$. Alors on a le lemme suivant :

Lemme 2.43 *Pour tout $k \in \{1, \cdots, n\}$, $(u - \lambda_k Id_E\ )(V_k) \subset V_{k-1}$.*

Preuve: Soit $k \in \{1, \cdots, n\}$ et $j \in \{1, \cdots, k\}$. Comme T est triangulaire supérieure il existe des scalaires $\alpha_{1j}, \cdots, \alpha_{jj}$ tels que :

$$(u - \lambda_k \mathrm{Id}_E\ )(e_j) = \sum_{i=1}^{j} \alpha_{ij} e_i - \lambda_k e_j$$

Si $j < k$, alors $(u - \lambda_k \mathrm{Id}_E\ )(e_j) \in V_{k-1}$. Si $j = k$, comme $\alpha_{kk} = \lambda_k$, on a aussi $(u - \lambda_k \mathrm{Id}_E\ )(e_j) \in V_{k-1}$. □

Poursuivons maintenant la preuve du théorème. Soit maintenant $x \in E = V_n$. Alors d'après le lemme précédent :

$$(u - \lambda_n \mathrm{Id}_E\ )(x) \in V_{n-1}\ ,\ (u - \lambda_{n-1} \mathrm{Id}_E\ ) \circ (u - \lambda_n \mathrm{Id}_E\ )(x) \in V_{n-2}\ , \cdots ,$$
$$(u - \lambda_2 \mathrm{Id}_E\ ) \circ \cdots \circ (u - \lambda_n \mathrm{Id}_E\ )(x) \in V_1$$

Comme V_1 n'est rien d'autre que le sous-espace propre associé à λ_1, on a :

$$(u - \lambda_1 \mathrm{Id}_E\ ) \circ \cdots \circ (u - \lambda_n \mathrm{Id}_E\ )(x) = 0$$

et comme $\chi_u(u)(x) = (-1)^n (u - \lambda_1 \mathrm{Id}_E\ ) \circ \cdots \circ (u - \lambda_n \mathrm{Id}_E\ )(x)$, il s'ensuit que $\chi_u(u) = 0$. □

Exemple 2.44 *On considère deux exemples pour illustrer ce théorème et la théorie du polynôme minimal. Soit* $U = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}$ *et* $V = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$*. Come ces matrices sont triangulaires supérieures, on calcul facilement leurs polynômes caractéristiques :* $\chi_U(X) = \chi_V(X) = (X-2)^2$*. On vérifie le théorème de Cayley-Hamilton :* $U - 2I_2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ *et* $V - 2I_2 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$*. On a donc bien* $(U-2I_2)^2 = 0$ *et* $(V-2I_2)^2 = 0$*.*

Exercice de cours 2.19 *Vérifier le théorème de Cayley-Hamilton pour les matrices* A, B, C *ci-dessus.*

Exercice de cours 2.20 *Soit* M *la matrice*

$$M = \begin{pmatrix} 2 & 1 & 1 & 2 & 0 \\ 4 & 4 & 3 & 5 & -2 \\ 0 & 0 & 1 & 1 & 0 \\ -2 & -1 & -1 & -1 & 1 \\ 2 & 2 & 2 & 4 & 1 \end{pmatrix}$$

Montrer que $\chi_M(X) = (X-1)^3(X-2)^2$*. Vérifier que* $\chi_M(M) = 0$ *(pour cela garder la forme factorisée de* χ_M *!).*

2.3 Polynôme minimal

2.3.1 Théorème de décomposition des noyaux

Théorème 2.45 *Soient* $u \in \mathcal{L}(E)$*,* $P_1, \cdots, P_k$ *des polynômes de* $\mathbb{K}[X]$ *premiers entre eux deux à deux et* $P = \prod_{i=1}^{k} P_i$*. Alors :*

$$Ker(P(u)) = \bigoplus_{i=1}^{k} Ker(P_i(u))$$

Preuve: Si $k = 1$, le résultat est évident.

Si $k = 2$, alors $P = P_1 P_2$ et P_1 et P_2 sont premiers entre eux. D'après le théorème de Bezout, il existe deux polynômes U_1 et U_2 tels que :

$$1 = U_1 P_1 + U_2 P_2$$

Il s'ensuit que :

$$\mathrm{Id}_E = U_1(u) \circ P_1(u) + U_2(u) \circ P_2(u)$$

Soit $x \in \mathrm{Ker}(P_1(u)) \cap \mathrm{Ker}(P_2(u))$. Alors la relation précédente entraîne que :

$$x = U_1(u)(P_1(u)(x)) + U_2(u)(P_2(u)(x)) = 0$$

Donc $\mathrm{Ker}(P_1(u))$ et $\mathrm{Ker}(P_2(u))$ sont en somme directe.

Montrons maintenant que $\mathrm{Ker}(P(u)) \subset \mathrm{Ker}(P_1(u)) + \mathrm{Ker}(P_2(u))$:

Soit $x \in \mathrm{Ker}(P(u))$, alors :

$$x = U_1(u)(P_1(u)(x)) + U_2(u)(P_2(u)(x))$$

Or :

$$P_2(u)(U_1(u)(P_1(u)(x))) = U_1(u)(P(u)(x)) = 0$$

Donc $U_1(u)(P_1(u)(x)) \in \mathrm{Ker}(P_2(u))$. De même $U_2(u)(P_2(u)(x)) \in \mathrm{Ker}(P_1(u))$. Finalement $x \in \mathrm{Ker}(P_1(u)) + \mathrm{Ker}(P_2(u))$.

Montrons que $\mathrm{Ker}(P_1(u)) + \mathrm{Ker}(P_2(u)) \subset \mathrm{Ker}(P(u))$:

Soit $x \in \mathrm{Ker}(P_1(u))$, alors $P(u)(x) = P_2(u)(P_1(u)(x)) = 0$ donc $x \in \mathrm{Ker}(P(u))$ et $\mathrm{Ker}(P_1(u)) \subset \mathrm{Ker}(P(u))$. De même $\mathrm{Ker}(P_2(u)) \subset \mathrm{Ker}(P(u))$. Donc $\mathrm{Ker}(P_1(u)) + \mathrm{Ker}(P_2(u)) \subset \mathrm{Ker}(P(u))$. On a donc montré que :

$$\mathrm{Ker}(P(u)) = \mathrm{Ker}(P_1(u)) \oplus \mathrm{Ker}(P_2(u))$$

Soit $k \geqslant 3$ et supposons le théorème démontré au rang $k-1$. Montrons que le résultat est vrai au rang k. Soient $P_1, \cdots, P_k$ des polynômes de $\mathbb{K}[X]$ premiers entre eux deux à deux et $P = \prod_{i=1}^{k} P_i$. Alors les polynômes P_1 et $P_2 \cdots P_k$ sont premiers entre eux et d'après ce qui a été montré précédemment on a :

$$\mathrm{Ker}(P(u)) = \mathrm{Ker}(P_1(u)) \oplus \mathrm{Ker}(P_2 \cdots P_k)(u) = \bigoplus_{i=1}^{k} \mathrm{Ker}(P_i(u))$$

□

Exercice de cours 2.21 *Pour la matrice M ci-dessus, déterminer les sous-espaces vectoriels Ker$((M - I_5)^3)$ et Ker$((M - 2I_5)^2)$. Vérifier qu'ils sont supplémentaires dans $\mathbb{R}^5$. Retrouver ce résultat en utilisant le théorème de décomposition des noyaux.*

2.3.2 Polynômes annulateurs, polynôme minimal

Commençons par quelques rappels.

Définition 2.46 *Soit $(A, +, \times)$ un anneau. Alors on dit qu'un sous-ensemble $\mathcal{I}$ de A est un* **idéal** *si :*

1. *$(\mathcal{I}, +)$ est un sous-groupe de $(A, +)$.*
2. *$\forall a \in A \ , \ \forall x \in \mathcal{I} \ , \ ax \in \mathcal{I}$.*

Un idéal est dit **principal** *si $\mathcal{I}$ est engendré par un seul élément. Autrement dit $\mathcal{I}$ est principal s'il existe $\alpha \in \mathcal{I}$ tel que :*

$$\mathcal{I} = \{y\alpha \ ; \ y \in A\}$$

Un **anneau principal** *est un anneau dont tous les idéaux sont principaux.*

L'anneau des polynômes $\mathbb{K}[X]$ est un anneau principal. C'est une conséquence de l'algorithme d'Euclide. En effet soit $\mathcal{I}$ un idéal de $\mathbb{K}[X]$. Alors $\mathcal{I}$ est non vide. Si $\mathcal{I} = \{0\}$, le résultat est évident.

Supposons $\mathcal{I} \neq \{0\}$. Soit P_0 un polynôme non nul de $\mathcal{I}$ de degré minimal. Soit $P \in \mathcal{I}$. En effectuant la division euclidienne de P par P_0, il existe des polynômes Q et R tels que :

$$P(X) = Q(X)P_0(X) + R(X)$$

où $\deg(R) < \deg(P_0)$. Comme $(\mathcal{I}, +)$ est un idéal de $\mathbb{K}[X]$, $P(X) - Q(X)P_0(X) \in \mathcal{I}$. Donc $R(X) \in \mathcal{I}$. Mais comme P_0 un polynôme non nul de $\mathcal{I}$ de degré minimal, il s'ensuit que $R(X) = 0$ et :

$$\mathcal{I} = \{Q(X)P_0(X)\ ;\ Q(X) \in \mathbb{K}[X]\}$$

Donc $\mathbb{K}[X]$ est un anneau principal et tout idéal $\mathcal{I}$ de $\mathbb{K}[X]$ est engendré par tout polynôme non nul de $\mathcal{I}$ de degré minimal. De plus il n'est pas difficile de voir qu'il existe un unique polynôme **unitaire** (i.e. dont le terme de plus haut degré vaut 1) de $\mathcal{I}$ de degré minimal qui engendre $\mathcal{I}$. Résumons tout ceci dans la proposition qui suit :

Proposition 2.47 *L'anneau $\mathbb{K}[X]$ des polynômes est principal et tout idéal $\mathcal{I}$ tel que $\mathcal{I} \neq \{0\}$ est engendré par un unique polynôme unitaire de $\mathcal{I}$ de degré minimal.*

Proposition et définition 2.48 *Soit $u \in \mathcal{L}(E)$. Considérons le sous-ensemble de $\mathbb{K}[X]$ suivant :*

$$\mathcal{I}(u) = \{P \in \mathbb{K}[X]\ ;\ P(u) = 0\}$$

Alors $\mathcal{I}(u)$ est un idéal de $\mathbb{K}[X]$ appelé **idéal annulateur** *de u.*

Tout polynôme P de $\mathcal{I}(u)$ est appelé **polynôme annulateur** *de u.*

Preuve: Le polynôme nul appartient à $\mathcal{I}(u)$, donc $\mathcal{I}(u) \neq \emptyset$. Soient $(P, Q) \in (\mathcal{I}(u))^2$. Alors :

$$(P - Q)(u) = P(u) - Q(u) = 0$$

Donc $\mathcal{I}(u)$ est un sous-groupe de $(\mathbb{K}[X], +)$. D'autre part pour tout $A \in \mathbb{K}[X]$:

$$(AP)(u) = A(u) \circ P(u) = 0$$

Donc $AP \in \mathcal{I}(u)$ et $\mathcal{I}(u)$ est un idéal de $\mathbb{K}[X]$. □

Exemple 2.49 *Soit $W = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$. Alors $W^n = 0$ si $n \geq 2$. Donc, pour $P(X) = a_0 + a_1X + a_2X^2 + \cdots$, on a $P(W) = a_0I_2 + a_1W$. L'annulateur de W est donc l'ensemble des polynômes $P(X) = a_0 + a_1X + a_2X^2 + \cdots$ avec $a = 0 = a_1 = 0$. On vérifiera que c'est un idéal de $\mathbb{K}[X]$.*

On va maintenant étudier de façon similaire l'idéal annulateur des matrices U et V de l'Exemple 2.44. Pour cela, on établit un lemme préliminaire :

Lemme 2.50 *Soit $P \in \mathbb{K}[X]$ de degré n et $c \in \mathbb{K}$. Il existe d'uniques coefficients $a_0, \cdots, a_n \in \mathbb{K}$ tels que*

$$P(X) = a_n(X-c)^n + a_{n-1}(X-c)^{n-1} + \cdots + a_1(X-c) + a_0 \tag{2.3}$$

PREUVE: En effet, l'équation 2.3 est équivalente à l'équation

$$P(X+c) = a_nX^n + a_{n-1}X^{n-1} + \cdots + a_1X + a_0.$$

Les coefficients a_i sont donc les coefficients du polynôme $P(X+c)$. □

Exemple 2.51 *Un polynôme P est dans l'idéal annulateur de U, resp. V, si et seulement si 2 est une racine double, resp. simple de P.*

PREUVE: Soit $P \in \mathbb{K}[X]$; écrivons le à l'aide du Lemme 2.50 sous la forme

$$P(X) = a_n(X-2)^n + a_{n-1}(X-2)^{n-1} + \cdots + a_2(X-2)^2 + a_1(X-2) + a_0.$$

On a alors

$$P(U) = a_n(U-2I_2)^n + a_{n-1}(U-2I_2)^{n-1} + \cdots + a_2(U-2I_2)^2 + a_1(U-2I_2) + a_0I_2.$$

Comme $U - 2I_2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ est de carré nul, on a donc $P(U) = a_1(U-2I_2) + a_0I_2$. Ainsi, $P(U)$ est nul si et seulement si $a_0 = a_1 = 0$, ce qui équivaut à dire que 2 est une racine double de P.

Pour V, le raisonnement est similaire, à ceci près que $V - 2I_2 = 0$. On a donc $P(V) = a_0I_2$. Ainsi, $P(V)$ est nul si et seulement si $a_0 = 0$, ce qui équivaut à dire que 2 est une racine de P. □

Proposition et définition 2.52 *Soit $u \in \mathcal{L}(E)$. Alors $\mathcal{I}(u) \neq \{0\}$ et il existe un unique polynôme unitaire μ_u de $\mathcal{I}(u)$ de degré minimal qui engendre $\mathcal{I}(u)$.*

Autrement dit μ_u est l'unique polynôme annulateur de u, unitaire de degré minimal qui divise tous les autres polynômes annulateurs de u.

μ_u s'appelle le **polynôme minimal** *de u.*

PREUVE: D'après le théorème de Cayley-Hamilton, $\chi_u(u) = 0$. Comme $\deg \chi_u = n$, χ_u est un polynôme non nul et $\mathcal{I}(u) \neq \{0\}$. D'après la proposition 2.47 il existe un unique polynôme unitaire μ_u de $\mathcal{I}(u)$ de degré minimal qui engendre $\mathcal{I}(u)$. Cela revient à dire que :

$$\mathcal{I}(u) = \{P(X)\mu_u(X) \,;\, P(X) \in \mathbb{K}[X]\}$$

□

Exercice de cours 2.22 *Soit $N \in \mathcal{M}_2(\mathbb{R})$ une matrice telle que $N^2 = I_2$. Montrer qu'on a trois possibilités pour μ_N : soit $\mu_N(X) = X-1$, soit $\mu_N(X) = X+1$, soit $\mu_N(X) = X^2-1$.*

Remarque 2.53 *Le degré m du polynôme minimal vérifie toujours $m \geqslant 1$.*

Exemple 2.54 *L'exemple 2.49 montre que l'idéal annulateur de W est l'ensemble des multiples de X^2. Ainsi, $\mu_W(X) = X^2$. De façon similaire, l'exemple 2.51 montre que l'idéal annulateur de U resp. V est l'ensemble des multiples de $(X-2)^2$ resp. $X-2$. On a donc $\mu_U(X) = (X-2)^2$ et $\mu_V(X) = X-2$.*

Exercice de cours 2.23 *Soit A la matrice de l'exercice 2.1 et soit μ_A son polynôme minimal.*

1. *Montrer que μ_A divise $(X+1)(X-1)$.*
   Il y a donc trois cas de figure possible : μ_A est un des trois polynômes $(X+1), (X-1)$ et $(X+1)(X-1)$.
2. *Les polynômes $X-1$ et $X+1$ annulent-ils A ?*
3. *En déduire que $\mu_A(X) = (X+1)(X-1)$.*

On remarque que l'unique racine de χ_W et μ_W est 0, et que l'unique racine de χ_U, χ_V, μ_U et μ_V est 2.

Proposition 2.55 *Le polynôme minimal et le polynôme caractéristique ont exactement les mêmes racines.*

Preuve: Soit λ une racine de χ_u. Alors λ est une valeur propre de u. Soit x un vecteur propre de u associé à λ. Il est facile de voir que pour tout entier $i \geqslant 0$: $u^i(x) = \lambda^i x$. Donc si $\mu_u(X) = \sum_{i=0}^{k} a_i X^i$, alors :

$$0 = \mu_u(u)(x) = \sum_{i=0}^{k} a_i u^i(x) = \left(\sum_{i=0}^{k} a_i \lambda^i \right) x = \mu_u(\lambda) x$$

Comme x est non nul car x est un vecteur propre, $\mu_u(\lambda) = 0$ et λ est une racine de μ_u.

Réciproquement, comme μ_u divise χ_u, les racines de μ_u sont des racines de χ_u. $\square$

Exercice de cours 2.24 *Pour les matrices B et C de l'exercice 2.6, montrer qu'on a*

$$\mu_B(X) = (X-2)(X-1)^2 \ ; \ \mu_C(X) = (X-2)(X-1).$$

Pour cela, on pourra utiliser la relation $\mu_B | \chi_B$, et la proposition précédante.

Exercice de cours 2.25 *Soit M la matrice de l'exercice 2.20. Montrer que*

$$\mu_M(X) = \chi_M(X) = (X-1)^3 (X-2)^2.$$

Exercice de cours 2.26 *Soit $N \in \mathcal{M}_2(\mathbb{R})$ une matrice telle que $N^2 = I_2$. Montrer que les racines de χ_N appartiennent à $\{-1, 1\}$.*

Observons que $\mu_U(X) = (X-2)^2, \mu_V(X) = X-2$ et $\mu_W(X) = X^2$. Par ailleurs, V est diagonale, U et W sont triangulaires supérieures. Les matrices U et W ne sont pas diagonalisables car leur unique espace propre est de dimension 1 (en effet $U - 2I_2$ et W sont de rang 1). Ceci illustre le critère de diagonalisabilité suivant :

Théorème 2.56 *Soit $u \in \mathcal{L}(E)$. Alors u est diagonalisable si et seulement si son polynôme minimal est scindé et si toutes ses racines sont simples.*

PREUVE:

— ($\Rightarrow$) Supposons que u est diagonalisable. Soient $\lambda_1, \cdots, \lambda_r$ les valeurs propres de u (comptées sans multiplicité - autrement dit $\lambda_1, \cdots, \lambda_r$ sont distincts). Considérons le polynôme :

$$P(X) = (X - \lambda_1) \cdots (X - \lambda_r)$$

Comme u est diagonalisable, $E = \bigoplus_{i=1}^{r} E_{\lambda_i}(u)$. Donc soit $x \in E$, alors il existe $(x_1, \cdots, x_r) \in E_{\lambda_1}(u) \times \cdots \times E_{\lambda_r}(u)$ tel que :

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

Or :

$$P(u)(x_i) = \left(\prod_{\substack{1 \leqslant j \leqslant r \\ j \neq i}} (u - \lambda_j \mathrm{Id}_E) \right) ((u - \lambda_i \mathrm{Id}_E)(x_i)) = 0$$

Donc $P(u)(x) = 0$ pour tout $x \in E$ et $P(u) = 0$. Donc P est un polynôme annulateur de u et μ_u divise P. Comme les racines de μ_u sont exactement les valeurs propres $\lambda_1, \cdots, \lambda_r$, il s'ensuit que $\mu_u = P$ et μ_u est scindé et n'a que des racines simples.

— ($\Leftarrow$) Supposons que μ_u est scindé et n'a que des racines simples. Alors :

$$\mu_u(X) = (X - \lambda_1) \cdots (X - \lambda_r)$$

et $\lambda_1, \cdots, \lambda_r$ sont les valeurs propres de u. D'après le théorème 2.45 :

$$E = \mathrm{Ker}(\mu_u(u)) = \bigoplus_{i=1}^{r} \mathrm{Ker}(u - \lambda_i \mathrm{Id}_E) = \bigoplus_{i=1}^{r} E_{\lambda_i}(u)$$

Donc E est la somme directe des sous-espaces propres de u ce qui signifie que u est diagonalisable.

□

Corollaire 2.57 *Soit $u \in \mathcal{L}(E)$. Alors u est diagonalisable si et seulement si il existe un polynôme annulateur de u qui est scindé et qui n'admet que des racines simples.*

PREUVE:

— ($\Rightarrow$) On suppose que u est diagonalisable. Alors d'après le théorème précédent μ_u est scindé et n'a que des racines simples. De plus c'est un polynôme annulateur.

— ($\Leftarrow$) Réciproquement supposons qu'il existe un polynôme annulateur P de u qui est scindé et qui n'admet que des racines simples. Comme μ_u divise P, il s'ensuit que μ_u est scindé et n'a que des racines simples ce qui implique d'après le théorème précédent que u est diagonalisable.

□

Exercice de cours 2.27 *Soit $N \in \mathcal{M}_2(\mathbb{R})$ une matrice telle que $N^2 = I_2$. Montrer que N est diagonalisable. Quelles peuvent être les valeurs propres de N ? Montrer que la trace de N vaut $-2, 0$ ou 2.*

2.4 Projecteurs

Soient F et G deux sous-espaces vectoriels supplémentaires dans E : $E = F \oplus G$. Pour tout x de E il existe un unique couple $(p(x), q(x)) \in F \times G$ tel que $x = p(x) + q(x)$.

Proposition 2.58 *Les applications p et q sont des endomorphismes de E. p s'appelle le* **projecteur** *sur F* **parallèlement** *à G et q s'appelle le* **projecteur** *sur G* **parallèlement** *à F.*

Remarque 2.59 *On a $p + q = Id_E$.*

Proposition 2.60

1. *Soient F et G deux sous-espaces vectoriels de E supplémentaires dans E et p le projecteur sur F parallèlement à G. Alors on a : $p \circ p = p$, $Im(p) = F$ et $Ker(p) = G$.*
   De plus si q est le projecteur sur G parallèlement à F, alors $p \circ q = q \circ p = 0$.
2. *Réciproquement, soit $p \in \mathcal{L}(E)$ tel que $p \circ p = p$. Alors :*

$$E = Im(p) \oplus Ker(p)$$

et p est le projecteur sur $Im(p)$ parallèlement à $Ker(p)$.

Preuve: Prouvons d'abord le point 1). Soit $x \in E$. Alors $p(x) = p(x) + 0$ et $(p(x), 0) \in F \times G$. Donc par unicité de la décomposition $p(p(x)) = p(x)$ et $p \circ p = p$.

Déterminons $\mathrm{Im}(p)$. Comme pour tout $x \in E$, $p(x) \in F$, il est évident que $\mathrm{Im}(p) \subset F$. D'autre part pour tout x dans F, on a $x = x + 0$ et $(x, 0) \in F \times G$. Donc $x = p(x) \in \mathrm{Im}(p)$. Ainsi $\mathrm{Im}(p) = F$.

Déterminons $\mathrm{Ker}(p)$. Pour tout x de G, $x = 0 + x$ et $(0, x) \in F \times G$, donc $p(x) = 0$, d'où $G \subset \mathrm{Ker}(p)$. Réciproquement soit $x \in G$, alors $x = p(x) + q(x) = 0 + x$ et par unicité de la décomposition $p(x) = 0$. Donc $\mathrm{Ker}(p) \subset G$. Ainsi $\mathrm{Ker}(p) = G$.

Montrons maintenant le point 2). Soit $x \in \mathrm{Ker}(p) \cap \mathrm{Im}(p)$. Alors $p(x) = 0$ et il existe $y \in E$ tel que $p(y) = x$. D'où :

$$0 = p(x) = p(p(y)) = (p \circ p)(y) = p(y) = x$$

Ceci montre que $\mathrm{Im}(p) \cap \mathrm{Ker}(p) = \{0\}$ et donc $\mathrm{Im}(p)$ et $\mathrm{Ker}(p)$ sont en somme directe. Soit $x \in E$. Alors $x = p(x) + (x - p(x))$. Or $p(x - p(x)) = p(x) - (p \circ p)(x) = 0$. Donc $x - p(x) \in \mathrm{Ker}(p)$. Ceci montre que $E = \mathrm{Im}(p) + \mathrm{Ker}(p)$. Nous avons donc montré que $\mathrm{Im}(p)$ et $\mathrm{Ker}(p)$ sont supplémentaires dans E.

Puisque pour tout $x \in E$, $x = p(x) + (x - p(x))$ avec $p(x) \in \mathrm{Im}(p)$ et $x - p(x) \in \mathrm{Ker}(p)$, p est le projecteur sur $\mathrm{Im}(p)$ parallèlement à $\mathrm{Ker}(p)$. □

Exercice de cours 2.28 *Soit, dans $\mathbb{R}^2$, les sous-espaces vectoriels F et G engendrés respectivement par $(1, 1)$ et $(1, -1)$. Ecrire les matrices P resp. Q des projections sur F (resp. G) parallèlement à G (resp. F). Vérifier que $P + Q = I_2, P^2 = P, Q^2 = Q$.*

Remarque 2.61 *L'application linéaire $q = Id_E - p$ est le projecteur sur $Ker(p)$ parallèlement à $Im(p)$, $Ker(q) = Im(p)$ et $Im(q) = Ker(p)$.*

Proposition et définition 2.62 *Soient $F_1, \ldots, F_k$ des sous-espaces vectoriels de E tels que $E = \bigoplus_{i=1}^{k} F_i$. Alors pour tout $x \in E$, il existe un unique k-uplet $(p_1(x), \ldots, p_k(x)) \in F_1 \times \ldots \times F_k$ tel que : $x = \sum_{i=1}^{k} p_i(x)$.*

Les applications p_i sont des endomorphismes et $(p_i)_{1 \leqslant i \leqslant k}$ s'appelle **famille de projecteurs** *de E associée à la décomposition $E = \bigoplus_{i=1}^{k} F_i$.*

Proposition 2.63

1. *Soit $(p_i)_{1 \leqslant i \leqslant k}$ une famille de projecteurs de E associée à la décomposition $E = \bigoplus_{i=1}^{k} F_i$.*

   Alors : pour tout $(i,j) \in \{1, \ldots, k\}^2$:
   - *(i) $Id_E = \sum_{i=1}^{k} p_i$.*
   - *(ii) $p_i \circ p_i = p_i$.*
   - *(iii) $p_i \circ p_j = 0$ si $i \neq j$.*

   De plus $Im(p_i) = F_i$ et $Ker(p_i) = \bigoplus_{\substack{j=1 \\ j \neq i}}^{k} F_j$ pour tout $i \in \{1, \ldots, k\}$.
2. *Réciproquement soit $(p_i)_{1 \leqslant i \leqslant k}$ une famille d'endomorphismes de E vérifiant les propriétés (i), (ii) et (iii). Alors $E = \bigoplus_{i=1}^{k} Im(p_i)$ et $(p_i)_{1 \leqslant i \leqslant k}$ est la famille de projecteurs associés à cette décomposition.*

Remarque 2.64 *Soit $E = \bigoplus_{i=1}^{r} F_i$ et $(p_i)_{1 \leqslant i \leqslant r}$ la famille de projecteurs associée à $(F_i)_{1 \leqslant i \leqslant r}$. Pour v un endomorphisme tel que pour tout $i \in \{1, \cdots, r\}$, F_i est stable par v, on a $v \circ p_i = p_i \circ v$.*

Preuve: Pour montrer l'égalité des applications linéaires $v \circ p_i$ et $p_i \circ v$, il suffit de montrer que pour $x \in F_j$ avec $1 \leq j \leq r$, on a $v \circ p_i(x) = p_i \circ v(x)$. Or, si $j \neq i$, on a $p_i(V_j) = \{0\}$, donc $v \circ p_i(x) = p_i \circ v(x) = 0$. Si $j = i$, on a $v \circ p_i(x) = p_i \circ v(x) = v(x)$. $\square$

Proposition 2.65 *Soit $u \in \mathcal{L}(E)$. Alors u est diagonalisable si et seulement si il existe des scalaires $\lambda_1, \cdots, \lambda_r$ et une famille de sous-espaces vectoriels $(F_i)_{1 \leqslant i \leqslant r}$ tels que :*

$$E = \bigoplus_{i=1}^{r} F_i \quad \text{et} \quad u = \sum_{i=1}^{r} \lambda_i p_i$$

où $(p_i)_{1 \leqslant i \leqslant r}$ est la famille de projecteurs associée à $(F_i)_{1 \leqslant i \leqslant r}$. Dans ce cas $\mathrm{Sp}(u) = \{\lambda_1, \cdots, \lambda_r\}$ et $F_i = E_{\lambda_i}(u)$ pour tout $i \in \{1, \cdots, r\}$.

Preuve: Supposons que u est diagonalisable. Par le théorème 2.19, E est somme directe de ses sous-espaces propres :

$$E = \bigoplus_i E_{\lambda_i}(u) .$$

Si on note $F_i = E_{\lambda_i}(u)$ et (p_i) la famille de projecteurs associée à cette décomposisition, il reste à montrer que $u = \sum_i \lambda_i p_i$. Pour montrer que pour tout x de E, $u(x) = \sum_i \lambda_i p_i(x)$, on peut, par linéarité, supposer qu'il existe i tel que $x \in F_i$. Alors, on a $u(x) = \lambda_i x$ et $\sum_j \lambda_j p_j(x) = \lambda_i x$. Donc l'égalité est établie.

Supposons maintenant qu'il existe (F_i) et (λ_i) tels que

$$\begin{cases} E = \bigoplus_i F_i \\ p = \sum_i \lambda_i p_i \end{cases}$$

Alors, on vérifie comme ci-dessus que pour $x \in F_i$, on a $\sum_j \lambda_j p_j(x) = \lambda_i x$, de sorte que $F_i \subset E_{\lambda_i}(u)$. On raisonne sur les dimensions pour en déduire que $F_i = E_{\lambda_i}(u)$: l'inclusion déjà établie donne $\dim(F_i) \leq \dim(E_{\lambda_i}(u))$. D'où $\dim E = \sum \dim(F_i) \leq \sum \dim(E_{\lambda_i}(u)) \leq \dim E$, de sorte que toutes ces inégalités sont des égalités, et donc $F_i = E_{\lambda_i}(u)$. On en déduit $E = \bigoplus E_{\lambda_i}(u)$, et donc que u est diagonalisable. □

Exercice de cours 2.29 *Soit $F \subset \mathbb{R}^3$ le sous-espace vectoriel engendré par $(1, -2, 3)$, et soit G le sous-espace vectoriel engendré par $(1, -1, 0)$ et $(0, 0, 1)$. Calculer la matrice P de la projection sur F parallèlement à G et la matrice Q de la projection sur G parallèlement à F. Calculer $2P + Q$ et comparer le résultat à la matrice C de l'exercice 2.6. Pouvez vous expliquer ce que vous constatez ?*

Remarque 2.66

— *Si $E = \bigoplus_{i=1}^{r} F_i$ et $u = \sum_{i=1}^{r} \lambda_i p_i$ où $(p_i)_{1\leqslant i\leqslant r}$ est la famille de projecteurs associée à $(F_i)_{1\leqslant i\leqslant r}$, alors pour tout entier $k \geqslant 0$:*

$$u^k = \sum_{i=1}^{r} \lambda_i^k p_i$$

— *Plus généralement si $(u_1, \cdots, u_r) \in \mathcal{L}(E)^r$ et si pour tout $i \in \{1, \cdots, r\}$ les sous-espaces F_i sont stables par u_i, alors l'endomorphisme $u = \sum_{i=1}^{r} u_i \circ p_i$ vérifie pour tout entier $k \geqslant 0$:*

$$u^k = \sum_{i=1}^{r} u_i^k \circ p_i$$

2.5 Sous-espaces caractéristiques et décomposition de Dunford-Jordan

2.5.1 Indice d'un endomorphisme et endomorphismes nilpotents

Proposition 2.67 *Soit $u \in \mathcal{L}(E)$. Alors :*

1. *Pour tout entier* $k \geqslant 0$, $Ker(u^k) \subset Ker(u^{k+1})$.
2. *La suite* $(Ker(u^k))_{k\in\mathbb{N}}$ *est stationnaire. Autrement dit :*
$$\exists j \geqslant 0 \ ; \ \forall k \geqslant j \ , \ Ker(u^j) = Ker(u^k)$$
3. *Si* j *est un entier tel que* $Ker(u^j) = Ker(u^{j+1})$ *alors :*
$$\forall k \geqslant j \ , \ Ker(u^j) = Ker(u^k)$$
4. *Soit* i_0 *le plus petit entier tel que* $Ker(u^{i_0}) = Ker(u^{i_0+1})$. *Alors* $i_0 \leqslant \dim(E)$.

Preuve:

1. Soit $k \in \mathbb{N}$ et soit $x \in \mathrm{Ker}(u^k)$, alors :
$$u^{k+1}(x) = u(u^k(x)) = u(0) = 0$$
Donc $x \in \mathrm{Ker}(u^{k+1})$ et $\mathrm{Ker}(u^k) \subset \mathrm{Ker}(u^{k+1})$.
2. Considérons la suite $(\dim(\mathrm{Ker}(u^k)))_{k\in\mathbb{N}}$. Comme $\mathrm{Ker}(u^k) \subset \mathrm{Ker}(u^{k+1}) \subset E$, $\dim(\mathrm{Ker}(u^k)) \leqslant \dim(\mathrm{Ker}(u^{k+1})) \leqslant \dim(E) = n$ donc $(\dim(\mathrm{Ker}(u^k)))_{k\in\mathbb{N}}$ est une suite croissante d'entiers majorée. Donc à partir d'un certain rang, la suite $(\dim(\mathrm{Ker}(u^k)))_{k\in\mathbb{N}}$ est stationnaire. Donc :
$$\exists j \geqslant 0 \ ; \ \forall k \geqslant j \ , \ \dim(\mathrm{Ker}(u^j)) = \dim(\mathrm{Ker}(u^k))$$
Comme $\mathrm{Ker}(u^j) \subset \mathrm{Ker}(u^k)$ pour tout $k \geqslant j$, il s'ensuit que $\mathrm{Ker}(u^j) = \mathrm{Ker}(u^k)$ pour tout $k \geqslant j$.
3. Supposons que $\mathrm{Ker}(u^j) = \mathrm{Ker}(u^{j+1})$. Montrons que pour tout $k \geqslant j$, $\mathrm{Ker}(u^k) = \mathrm{Ker}(u^{k+1})$. Soit $x \in \mathrm{Ker}(u^{k+1})$, alors :
$$u^{k+1}(x) = u^{j+1}(u^{k-j}(x)) = 0$$
Donc $u^{k-j}(x) \in \mathrm{Ker}(u^{j+1})$, mais comme $\mathrm{Ker}(u^j) = \mathrm{Ker}(u^{j+1})$, $u^{k-j}(x) \in \mathrm{Ker}(u^j)$ et :
$$u^j(u^{k-j}(x)) = 0 = u^k(x)$$
Donc $x \in \mathrm{Ker}(u^k)$ et $\mathrm{Ker}(u^{k+1}) \subset \mathrm{Ker}(u^k)$. Mais d'après le 1. on a aussi $\mathrm{Ker}(u^k) \subset \mathrm{Ker}(u^{k+1})$. Finalement $\mathrm{Ker}(u^k) = \mathrm{Ker}(u^{k+1})$.
4. Soit i_0 le plus petit entier tel que $\mathrm{Ker}(u^{i_0}) = \mathrm{Ker}(u^{i_0+1})$. Considérons l'application :
$$\begin{array}{rcccc} F & : & \{0, \cdots, i_0\} & \longrightarrow & \{0, \cdots, n\} \\ & & k & \longmapsto & \dim(\mathrm{Ker}(u^k)) \end{array}$$
D'après le point précédent, on a :
$$\dim(\mathrm{Ker}(u^0)) < \cdots < \dim(\mathrm{Ker}(u^{i_0}))$$
Donc F est injective et il s'ensuit que $i_0 \leqslant n$.

□

Définition 2.68 *Soit $u \in \mathcal{L}(E)$. On appelle* **indice** *de u le plus petit entier i_0 tel que $Ker(u^{i_0}) = Ker(u^{i_0+1})$.*

Exercice de cours 2.30 *Déterminer l'indice de $B - I_3$ et celui de $C - I_3$.*

Remarque 2.69

— *Si $i_0 > 0$, alors d'après la définition de i_0 et la proposition on a pour tout $p \geqslant 1$:*

$$\{0\} = Ker(u^0) \subsetneq \cdots \subsetneq Ker(u^{i_0}) = Ker(u^{i_0+1}) = \cdots = Ker(u^{i_0+p})$$

— *Soit $p \in \mathbb{N}^\star$. Alors $Ker(u^p) = \{0\}$ si et seulement si $Ker(u) = \{0\}$. En effet $Ker(u) \subset Ker(u^p)$ donc l'égalité $Ker(u^p) = \{0\}$ implique $Ker(u) = \{0\}$. Réciproquement, si $Ker(u) = \{0\}$, alors u est injective, et il en est de même de u^p pour tout p.*

Le deuxième point résulte aussi de la Proposition 2.67 dans le cas particulier $i_0 = 0$.

Définition 2.70 *Soit $\nu \in \mathcal{L}(E)$. On dit que ν est un endomorphisme* **nilpotent** *s'il existe un entier k tel que $\nu^k = 0$.*

Proposition 2.71 *Soit ν un endomorphisme nilpotent de E. Soit n_0 l'indice de ν. Alors n_0 est le plus petit entier tel que $\nu^{n_0} = 0$ et on a $n_0 \leqslant \dim(E)$.*

Preuve: Soit m le plus petit entier tel que $\nu^m = 0$. Alors $\nu^{m-1} \neq 0$ et :

$$\mathrm{Ker}(\nu^{m-1}) \subsetneq \mathrm{Ker}(\nu^m) = E$$

Il s'ensuit que $\mathrm{Ker}(\nu^m) = \mathrm{Ker}(\nu^{m+1})$ et m est l'indice de ν. □

Exercice de cours 2.31 *Donner un exemple d'endomorphisme nilpotent non nul.*

Proposition 2.72 *Soit ν un endomorphisme nilpotent de E. Alors $\mathrm{Sp}(\nu) = \{0\}$. De plus ν est diagonalisable si et seulement si $\nu = 0$.*

Preuve: Soit n_0 l'indice de ν. Alors $\nu^{n_0} = 0$ et $\nu^{n_0-1} \neq 0$. Soit $P(X) = X^{n_0}$. Alors P est un polynôme annulateur et comme le polynôme minimal de ν divise P, il s'ensuit que $\mu_\nu(X) = X^k$ avec $k \leqslant n_0$.

D'après le théorème 2.56, ν est diagonalisable si et seulement si μ_ν est scindé (ce qui est le cas) et μ_ν n'a que des racines simples. Or μ_ν n'a que des racines simples si et seulement si $\mu_\nu(X) = X$. Mais dans ce cas $\mu_\nu(\nu) = \nu = 0$. □

Proposition 2.73 *Soit d un endomorphisme diagonalisable de E. Alors d est nilpotent si et seulement si $d = 0$.*

Preuve: Dans une base adaptée, la matrice D de d est diagonale : écrivons $D = Diag(\lambda_1, \ldots, \lambda_n)$. Alors $D^k = Diag(\lambda_1^k, \cdots, \lambda_n^k)$. Donc, il existe k tel que $D^k = 0$ si et seulement si $\forall i, \lambda_i = 0$, soit $D = 0$. □

Exercice de cours 2.32 *Soit ν un endomorphisme nilpotent de E et soit n_0 son indice de nilpotence.*

1. *Montrer que ν annule un polynôme scindé.*
2. *Quelles sont les valeurs propres de ν ?*
3. *Montrer qu'il existe une base de E dans laquelle la matrice de ν est strictement triangulaire supérieure (c'est-à-dire triangulaire supérieure avec des 0 sur la diagonale).*
4. *Montrer que $\nu^{\dim(E)} = 0$.*
5. *Retrouver le résultat de la Proposition 2.71 selon lequel $n_0 \leq \dim(E)$.*

2.5.2 Sous-espaces caractéristiques

Dans ce paragraphe u est un endomorphisme de E tel que χ_u soit scindé sur $\mathbb{K}$. Soient $\lambda_1, \cdots, \lambda_r$ les valeurs propres (distinctes) de u. Alors :

$$\chi_u(X) = (-1)^n \prod_{i=1}^{r} (X - \lambda_i)^{\alpha_i} \quad , \quad \mu_u(X) = \prod_{i=1}^{r} (X - \lambda_i)^{\beta_i}$$

On note d'autre part :

$$C_{\lambda_i}(u) := \mathrm{Ker}(u - \lambda_i \mathrm{Id}_E)^{\alpha_i} \quad \text{et} \quad M_{\lambda_i}(u) := \mathrm{Ker}(u - \lambda_i \mathrm{Id}_E)^{\beta_i}$$

On a bien entendu $1 \leqslant \beta_i \leqslant \alpha_i$.

Définition 2.74 *Soit $i \in \{1, \cdots, r\}$. Le sous-espace vectoriel $C_{\lambda_i}(u)$ s'appelle le* **sous-espace caractéristique** *associé à la valeur propre λ_i.*

Exercice de cours 2.33 *Déterminer $C_1(A), M_1(A), C_{-1}(A), M_{-1}(A)$. Déterminer aussi $C_1(B)$ et $M_1(B)$.*

Remarque 2.75

— *D'après le paragraphe précédent, on a :*

$$E_{\lambda_i}(u) \subset M_{\lambda_i}(u) \subset C_{\lambda_i}(u)$$

— *$M_{\lambda_i}(u)$ et $C_{\lambda_i}(u)$ sont stables par u (i.e. $u(M_{\lambda_i}(u)) \subset M_{\lambda_i}(u)$ et $u(C_{\lambda_i}(u)) \subset C_{\lambda_i}(u)$).*

Proposition 2.76 *Soit $i \in \{1, \cdots, r\}$. L'indice de $u - \lambda_i Id_E$ est β_i. En particulier on a :*

$$M_{\lambda_i}(u) = C_{\lambda_i}(u)$$

Preuve: Les polynômes des ensembles $\{(X - \lambda_i)^{\alpha_i} \ ; \ 1 \leqslant i \leqslant r\}$ et $\{(X - \lambda_i)^{\beta_i} \ ; \ 1 \leqslant i \leqslant r\}$ sont premiers deux à deux. Comme χ_u et μ_u sont des polynômes annulateurs on a $\mathrm{Ker}(\chi_u(u)) = \mathrm{Ker}(\mu_u(u)) = E$ et d'après le théorème 2.45 on a :

$$E = \bigoplus_{i=1}^{r} M_{\lambda_i}(u) = \bigoplus_{i=1}^{r} C_{\lambda_i}(u)$$

Donc :

$$\dim(E) = \sum_{i=1}^{r} \dim(M_{\lambda_i}(u)) = \sum_{i=1}^{r} \dim(C_{\lambda_i}(u))$$

Donc

$$\sum_{i=1}^{r} (\dim(C_{\lambda_i}(u) - \dim(M_{\lambda_i}(u)) = 0$$

Comme $\dim(C_{\lambda_i}(u)) - \dim(M_{\lambda_i}(u)) \geqslant 0$, On a $\dim(M_{\lambda_i}(u)) = \dim(C_{\lambda_i}(u))$ et $M_{\lambda_i}(u) = C_{\lambda_i}(u)$. On en déduit donc que l'indice i_0 de $u - \lambda_i \mathrm{Id}_E$ vérifie $i_0 \leqslant \beta_i$.

Supposons que $i_0 < \beta_i$. Cela signifie alors que $\mathrm{Ker}(u - \lambda_i \mathrm{Id}_E)^{\beta_i - 1} = \mathrm{Ker}(u - \lambda_i \mathrm{Id}_E)^{\beta_i}$. Considérons alors le polynôme :

$$Q(X) = \left(\prod_{\substack{1 \leqslant j \leqslant r \\ j \neq i}} (X - \lambda_j)^{\beta_j} \right) (X - \lambda_i)^{\beta_i - 1}$$

Soit $x \in E$. Comme $E = \bigoplus_{j=1}^{r} M_{\lambda_j}(u)$, il existe $(x_1, \cdots, x_r) \in M_{\lambda_1}(u) \times \cdots \times M_{\lambda_r}(u)$ tel que : $x = x_1 + \cdots + x_r$. Pour tout $j \in \{1, \cdots, r\}$ tel que $j \neq i$, on a $(u - \lambda_j \mathrm{Id}_E)^{\beta_j}(x_j) = 0$. De plus comme $\mathrm{Ker}(u - \lambda_i \mathrm{Id}_E)^{\beta_i - 1} = \mathrm{Ker}(u - \lambda_i \mathrm{Id}_E)^{\beta_i} = M_{\lambda_i}(u)$ on a $(u - \lambda_i \mathrm{Id}_E)^{\beta_i - 1}(x_i) = 0$. Donc pour tout $j \in \{1, \cdots, r\}$ $Q(u)(x_j) = 0$ et :

$$Q(u)(x) = 0$$

Ceci pour tout $x \in E$. Donc Q est un polynôme annulateur de u. Mais $\deg(Q) < \deg(\mu_u)$ ce qui contredit la définition de μ_u. Donc $i_0 = \beta_i$. □

Proposition 2.77 *Pour tout $i \in \{1, \cdots, r\}$ on a :*

$$\dim(C_{\lambda_i}(u)) = m(\lambda_i)$$

Preuve: On montre d'abord que $\dim(C_{\lambda_i}(u)) \leq m(\lambda_i)$ par une preuve analogue à celle de la Proposition 2.17 : posons $n_i = \dim(C_{\lambda_i}(u))$. Comme $C_{\lambda_i}(u)$ est stable par u, la restriction $u_i = u_{|C_{\lambda_i}(u)}$ est un endomorphisme de $C_{\lambda_i}(u)$. De plus on a $\mathrm{Ker}(u_i - \lambda_i \mathrm{Id}_{C_{\lambda_i}(u)})^{\alpha_i} = C_{\lambda_i}(u)$. Donc le polynôme $Q(X) = (X - \lambda_i)^{\alpha_i}$ est un polynôme annulateur de u_i et ceci montre que λ_i est l'unique valeur propre de u_i. Soit $\mathcal{B}_i$ une base de $C_{\lambda_i}(u)$ dans laquelle la matrice de u_i est triangulaire supérieure. On complète $\mathcal{B}_i$ par une famille de vecteurs $\mathcal{L}$ de façon à ce que $\mathcal{B} = \mathcal{B}_i \cup \mathcal{L}$ soit une base de E. Alors :

$$\mathrm{Mat}_{\mathcal{B}}(u) = \begin{pmatrix} \lambda_i & a_{12} & \cdots & a_{1n_i} & \\ 0 & \ddots & & \vdots & \\ \vdots & \ddots & \ddots & a_{n_i-1,n_i} & A \\ 0 & \cdots & 0 & \lambda_i & \\ & & O & & B \end{pmatrix}$$

On en déduit alors que :

$$\chi_u(X) = (\lambda_i - X)^{n_i}\chi_B(X)$$

et donc $(\lambda_i - X)^{n_i}$ divise $\chi_u(X)$.

Cela implique que $n_i \leqslant \alpha_i$. On a donc :

$$\dim(E) = \sum_{i=1}^{r} \dim(C_{\lambda_i}(u)) = \sum_{i=1}^{r} n_i \leqslant \sum_{i=1}^{r} \alpha_i = \dim(E)$$

On déduit que $n_i = \alpha_i$. □

Exercice de cours 2.34 *Montrer que ce résultat est compatible avec les calculs d'espaces caractéristiques que vous avez faits dans l'exercice 2.33.*

2.5.3 Décomposition de Dunford-Jordan

Théorème 2.78 (Décomposisition de Dunford-Jordan) *Soit $u \in \mathcal{L}(E)$ tel que χ_u soit scindé dans $\mathbb{K}[X]$. Alors il existe un unique couple (d, ν) d'endomorphismes de E tel que :*

1. *$u = d + \nu$.*
2. *$d \circ \nu = \nu \circ d$.*
3. *d est diagonalisable et ν est nilpotent.*

De plus si $\mu_u(X) = \prod_{i=1}^{r}(X - \lambda_i)^{\beta_i}$, alors $d = \sum_{i=1}^{r} \lambda_i p_i$ où $(p_i)_{1\leqslant i\leqslant r}$ est la famille de projecteurs associés à la décomposition $E = \bigoplus_{i=1}^{r} C_{\lambda_i}(u)$, et l'indice de ν est $N = \max_{1\leqslant i\leqslant r}(\beta_i)$.

Preuve: On a $E = \bigoplus_{i=1}^{r} C_{\lambda_i}(u)$. Soit $(p_i)_{1\leqslant i\leqslant r}$ la famille de projecteurs associés à cette décomposition. Posons $d = \sum_{i=1}^{r} \lambda_i p_i$. Alors d est un endomorphisme diagonalisable (voir Proposition 2.65). Posons maintenant $\nu = u - d$. Alors :

$$\nu = u - \sum_{i=1}^{r} \lambda_i p_i = u \circ \sum_{i=1}^{r} p_i - \sum_{i=1}^{r} \lambda_i p_i = \sum_{i=1}^{r} (u - \lambda_i \mathrm{Id}_E) \circ p_i$$

Montrons que $\nu \circ d = d \circ \nu$.

$$\begin{aligned}
\nu \circ d &= \left(\sum_{i=1}^{r} (u - \lambda_i \mathrm{Id}_E) \circ p_i\right) \circ \sum_{j=1}^{r} \lambda_j p_j \\
&= \sum_{1\leqslant i,j\leqslant r} \lambda_j (u - \lambda_i \mathrm{Id}_E) \circ p_i \circ p_j \\
&= \sum_{1\leqslant i,j\leqslant r} \lambda_j (u - \lambda_i \mathrm{Id}_E) \circ p_j \circ p_i
\end{aligned}$$

Comme pour tout couple $(i,j) \in \{1,\cdots,r\}^2$, les sous-espaces $C_{\lambda_j}(u)$ sont stables par $(u-\lambda_i \mathrm{Id}_E)$ alors $(u-\lambda_i \mathrm{Id}_E) \circ p_j = p_j \circ (u-\lambda_i \mathrm{Id}_E)$, par la Remarque 2.64. Donc :

$$\begin{aligned} \nu \circ d &= \sum_{1 \leqslant i,j \leqslant r} \lambda_j p_j \circ (u-\lambda_i \mathrm{Id}_E) \circ p_i \\ &= d \circ \nu \end{aligned}$$

Montrons maintenant que ν est nilpotent d'indice $N = \max\limits_{1 \leqslant i \leqslant r}(\beta_i)$. Comme pour tout $i \in \{1,\cdots,r\}$, les sous-espaces $C_{\lambda_i}(u)$ sont stables par $(u-\lambda_i \mathrm{Id}_E)$, on a (voir la remarque 2.66) :

$$\nu^N = \sum_{i=1}^{r} (u-\lambda_i \mathrm{Id}_E)^N \circ p_i$$

Or $N \geqslant \beta_i$ pour tout $i \in \{1,\cdots,r\}$ et β_i est l'indice de $u-\lambda_i \mathrm{Id}_E$ d'après la proposition 2.76. Donc $\mathrm{Ker}(u-\lambda_i \mathrm{Id}_E)^N = \mathrm{Ker}(u-\lambda_i \mathrm{Id}_E)^{\beta_i} = C_{\lambda_i}(u)$.

Soit $x \in E$. Alors pour tout $i \in \{1,\cdots,r\}$, $p_i(x) \in \mathrm{Ker}(u-\lambda_i \mathrm{Id}_E)^N$. Donc pour tout $x \in E$, $\nu^N(x) = 0$ ce qui signifie que $\nu^N = 0$ et ν est nilpotent. Soit $j \in \{1,\cdots,r\}$ tel que $\beta_j = N$. Alors dans ce cas $\mathrm{Ker}(u-\lambda_j \mathrm{Id}_E)^{N-1} \subsetneq \mathrm{Ker}(u-\lambda_j \mathrm{Id}_E)^N = C_{\lambda_j}(u)$. Donc soit $x \in C_{\lambda_j}(u) \setminus \mathrm{Ker}(u-\lambda_j \mathrm{Id}_E)^{N-1}$. Alors $p_i(x) = 0$ pour tout $i \neq j$ et :

$$\nu^{N-1}(x) = \sum_{i=1}^{r} (u-\lambda_i \mathrm{Id}_E)^{N-1}(p_i(x)) = (u-\lambda_j \mathrm{Id}_E)^{N-1}(x) \neq 0$$

Donc ν est nilpotent d'indice N.

Il nous reste à montrer l'unicité du couple (d,ν). Soit (d',ν') vérifiant les hypothèses 1, 2 et 3 du théorème. Notons N' l'indice de ν'. Soit $\lambda' \in \mathrm{Sp}(d')$ et $x \in E_{\lambda'}(d') \setminus \{0\}$. Alors comme u et d' commutent :

$$\begin{aligned} 0 = \nu'^{N'}(x) = (u-d')^{N'}(x) &= \sum_{k=0}^{N'} \binom{N'}{k} (-1)^{N'-k} (u^k \circ d'^{N'-k})(x) \\ &= \sum_{k=0}^{N'} \binom{N'}{k} (-1)^{N'-k} (u^k \circ (\lambda' \mathrm{Id}_E)^{N'-k})(x) \\ &= (u-\lambda' \mathrm{Id}_E)^{N'}(x) \end{aligned}$$

Donc $\mathrm{Ker}(u-\lambda' \mathrm{Id}_E)^{N'} \neq \{0\}$. Il s'ensuit d'après la remarque 2.69 que $\mathrm{Ker}(u-\lambda' \mathrm{Id}_E) \neq \{0\}$ et $\lambda' \in \mathrm{Sp}(u)$. Donc $\mathrm{Sp}(d') \subset \mathrm{Sp}(u)$. De plus on déduit de ce qui précède que :

$$E_{\lambda'}(d') \subset C_{\lambda'}(u)$$

Notons $\mathrm{Sp}(d') = \{\lambda'_1,\cdots,\lambda'_s\}$. Comme d' est diagonalisable :

$$E = \bigoplus_{i=1}^{s} E_{\lambda'_i}(d') \subset \bigoplus_{i=1}^{s} C_{\lambda'_i}(u)$$

Donc $E = \bigoplus_{i=1}^{s} C_{\lambda'_i}(u)$. On en déduit que $\{\lambda'_1, \cdots, \lambda'_s\} = \{\lambda_1, \cdots, \lambda_r\}$ et $\mathrm{Sp}(d') = \mathrm{Sp}(u)$. Donc pour tout $i \in \{1, \cdots, r\}$, $E_{\lambda_i}(d') = C_{\lambda_i}(u)$ et pour tout $x \in E$ on a :

$$d'(x) = \sum_{i=1}^{r} d'(p_i(x)) = \sum_{i=1}^{r} \lambda_i p_i(x) = d(x)$$

Donc $d = d'$ et $\nu = \nu'$. $\square$

Remarque 2.79 *Si u est diagonalisable, sa décomposition de Dunford-Jordan est donc $u = u+0$ (c'est-à-dire $d = u$ et $\nu = 0$).*

Exercice de cours 2.35

1. *Déterminer la décomposition de Dunford-Jordan des matrices A et C.*
2. *Déterminer la décomposition de Dunford-Jordan de la matrice $\begin{pmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{pmatrix}$. Attention c'est un piège.*

Remarque 2.80

1. $\mathrm{Sp}(d) = \mathrm{Sp}(u)$.
2. $\forall \lambda \in \mathrm{Sp}(u)$, $E_\lambda(d) = C_\lambda(u)$.
3. $\forall \lambda \in \mathrm{Sp}(u)$, $C_\lambda(u)$ *est stable par ν et par d.*

2.6 Applications et exemples : suites récurrentes

2.6.1 Calcul des puissances d'un endomorphisme

Soit $u \in \mathcal{L}(E)$ tel que χ_u soit scindé dans $\mathbb{K}[X]$. Soient $\lambda_1, \cdots, \lambda_r$ les valeurs propres (distinctes) de u. Alors :

$$\chi_u(X) = (-1)^n \prod_{i=1}^{r} (X - \lambda_i)^{\alpha_i} \quad , \quad \mu_u(X) = \prod_{i=1}^{r} (X - \lambda_i)^{\beta_i}$$

On rappelle d'autre part que :

$$\mathrm{Ker}(u - \lambda_i \mathrm{Id}_E)^{\beta_i - 1} \subsetneq \mathrm{Ker}(u - \lambda_i \mathrm{Id}_E)^{\beta_i} = \mathrm{Ker}(u - \lambda_i \mathrm{Id}_E)^{\alpha_i} = C_{\lambda_i}(u)$$

et que $1 \leqslant \beta_i \leqslant \alpha_i = \dim(C_{\lambda_i}(u))$.

D'après le théorème 2.78 il existe un unique couple (d, ν) d'endomorphismes de E vérifiant les propriétés 1,2,3 du théorème. Notons m l'indice de ν. Soit $k \in \mathbb{N}$ alors comme d et ν commutent on peut appliquer la formule du binôme de Newton ce qui donne :

$$u^k = \sum_{i=0}^{k} \binom{k}{i} d^{k-i} \nu^i$$

Si $k \geqslant m-1$, alors :

$$u^k = \sum_{i=0}^{m-1} \binom{k}{i} d^{k-i} \nu^i$$

Maintenant soient des bases $\mathcal{B}_1, \cdots, \mathcal{B}_r$ respectivement de chacun des sous-espaces caractéristiques $C_{\lambda_1}(u), \cdots, C_{\lambda_r}(u)$. Considérons la base de E définie par $\mathcal{B} = \mathcal{B}_1 \cup \cdots \cup \mathcal{B}_r$. Dans ce cas comme $E_{\lambda_i}(u) = C_{\lambda_i}(u)$ pour tout $i \in \{1, \cdots, r\}$ on a :

$$\mathrm{Mat}_{\mathcal{B}}(d) = \begin{pmatrix} \lambda_1 & & & & & & \\ & \ddots & & & & & \\ & & \lambda_1 & & & & \\ & O & & \ddots & & O & \\ & & & & \lambda_r & & \\ & & & & & \ddots & \\ & & & & & & \lambda_r \end{pmatrix}$$

et $\mathrm{Mat}_{\mathcal{B}}(\nu) = \mathrm{Mat}_{\mathcal{B}}(u-d)$. De plus comme $C_{\lambda_i}(u)$ est stable par ν, en posant $A = \mathrm{Mat}_{\mathcal{B}}(u)$, $D = \mathrm{Mat}_{\mathcal{B}}(d)$ et $N = \mathrm{Mat}_{\mathcal{B}}(\nu)$ alors A est une matrice diagonale par blocs et pour $k \geqslant m-1$ on a :

$$A^k = \sum_{i=0}^{m-1} \binom{k}{i} D^{k-i} N^i$$

On sait calculer les puissances de D. Il reste à calculer $N^2, \cdots, N^{m-1}$.

2.6.2 Une méthode de trigonalisation

On peut cependant faire les choses plus finement. Pour cela pour tout $i \in \{1, \cdots, r\}$ comme :

$$\mathrm{Ker}(u - \lambda_i \mathrm{Id}_E) \subsetneq \cdots \subsetneq \mathrm{Ker}(u - \lambda_i \mathrm{Id}_E)^{\beta_i - 1} \subsetneq \mathrm{Ker}(u - \lambda_i \mathrm{Id}_E)^{\beta_i}$$

On construit des bases $\mathcal{B}_{i,1}, \cdots, \mathcal{B}_{i,\beta_i}$ respectivement de $\mathrm{Ker}(u - \lambda_i \mathrm{Id}_E), \cdots, \mathrm{Ker}(u - \lambda_i \mathrm{Id}_E)^{\beta_i}$ de telle manière que :

$$\mathcal{B}_{i,1} \subset \cdots \subset \mathcal{B}_{i,\beta_i}$$

Autrement dit chaque base $\mathcal{B}_{i,j-1}$ de $\mathrm{Ker}(u - \lambda_i \mathrm{Id}_E)^{j-1}$ est complétée par une famille libre $\mathcal{L}_j$ de $\mathrm{Ker}(u - \lambda_i \mathrm{Id}_E)^j$ pour obtenir une base $\mathcal{B}_{i,j}$ de $\mathrm{Ker}(u - \lambda_i \mathrm{Id}_E)^j$. Remarquons alors que si $v \in \mathcal{L}_j$ alors $(u - \lambda_i \mathrm{Id}_E)^{j-1}((u - \lambda_i \mathrm{Id}_E)(v)) = 0$. Donc $u(v) - \lambda_i v \in \mathrm{Ker}(u - \lambda_i \mathrm{Id}_E)^{j-1} = \mathrm{Vect}(\mathcal{B}_{i,j-1})$. Il s'ensuit qu'on a :

$$\mathrm{Mat}_{\mathcal{B}}(u) = \begin{pmatrix} \lambda_1 & * & \cdots & \cdots & \cdots & \cdots & * \\ & \ddots & \ddots & & & & \vdots \\ & & \lambda_1 & \ddots & & & \vdots \\ & O & & \ddots & \ddots & & \vdots \\ & & & & \lambda_r & \ddots & \vdots \\ & & & & & \ddots & * \\ & & & & & & \lambda_r \end{pmatrix}$$

et dans cette base N est une matrice triangulaire supérieure stricte.

2.6.3 Suites récurrentes

Soit (u_n) une suite à valeurs complexes.

Définition 2.81 *Une relation de récurrence d'ordre k est une relation vérifiée par la suite (u_n) du type :*

$$u_{n+k} = a_{k-1}u_{n+k-1} + a_{k-2}u_{n+k-2} + \cdots + a_0u_n \tag{2.4}$$

pour tout $n \geqslant 0$.

Cette relation de récurrence est équivalente à la relation de récurrence matricielle :

$$X_{n+1} = AX_n$$

pour tout $n \geqslant 0$ où $X_n = \begin{pmatrix} u_{n+k-1} \\ \vdots \\ u_n \end{pmatrix} \in \mathcal{M}_{k,1}(\mathbb{K})$ et $A = \begin{pmatrix} a_{k-1} & a_{k-2} & \cdots & \cdots & a_0 \\ 1 & 0 & \cdots & \cdots & 0 \\ 0 & 1 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 1 & 0 \end{pmatrix} \in \mathcal{M}_k(\mathbb{K})$. Un calcul facile permet de voir que pour tout $n \geqslant 0$, $X_n = A^nX_0$. Calculons le polynôme caractéristique χ_A. Alors :

$$\begin{aligned} \chi_A(\lambda) &= \begin{vmatrix} a_{k-1}-\lambda & a_{k-2} & \cdots & \cdots & a_0 \\ 1 & -\lambda & \cdots & \cdots & 0 \\ 0 & 1 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 1 & -\lambda \end{vmatrix} \\ &= \begin{vmatrix} -\lambda & 0 & \cdots & \cdots & 0 \\ 0 & -\lambda & \cdots & \cdots & 0 \\ 0 & 1 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 1 & -\lambda \end{vmatrix} + D_{k-1} = (-\lambda)^k + D_{k-1} \end{aligned}$$

où :

$$D_i = \begin{vmatrix} a_i & a_{i-1} & \cdots & \cdots & a_0 \\ 1 & -\lambda & \cdots & \cdots & 0 \\ 0 & 1 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 1 & -\lambda \end{vmatrix}$$

En développant suivant la première colonne on remarque que $D_i = (-1)^i a_i \lambda^i - D_{i-1}$ et par récurrence on montre que :

$$D_i = (-1)^i \sum_{j=0}^{i} a_j \lambda^j$$

et :

$$\chi_A(\lambda) = (-1)^k \left(\lambda^k - a_{k-1}\lambda^{k-1} - \cdots - a_1\lambda - a_0\right)$$

Si on travaille dans $\mathbb{C}$, χ_A est scindé.

Supposons d'abord que toutes les racines de χ_A soient distinctes. Notons $\lambda_1, \cdots, \lambda_k$ ces racines. Alors pour tout $i \in \{1, \cdots, k\}$, $V_i = \begin{pmatrix} \lambda_i^{k-1} \\ \lambda_i^{k-2} \\ \vdots \\ \lambda_i \\ 1 \end{pmatrix}$ est un vecteur propre associé à λ_i car $a_{k-1}\lambda_i^k + \cdots + a_1\lambda_i + a_0 = \lambda_i^k$.

D'après la remarque 2.23, on a $D = P^{-1}AP$ où P est la matrice de $\mathcal{M}_k(\mathbb{K})$ dont les vecteurs colonne $C_i(P)$ vérifient $C_i(P) = V_i$, et :

$$D = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_k \end{pmatrix}$$

Donc notant $P^{-1}X_0 = \begin{pmatrix} c_1 \\ \vdots \\ c_k \end{pmatrix}$

$$X_n = A^n X_0 = PD^nP^{-1}X_0 = P \begin{pmatrix} \lambda_1^n & & 0 \\ & \ddots & \\ 0 & & \lambda_k^n \end{pmatrix} \begin{pmatrix} c_1 \\ \vdots \\ c_k \end{pmatrix} = P \begin{pmatrix} c_1\lambda_1^n \\ \vdots \\ c_k\lambda_k^n \end{pmatrix}$$

Donc $X_n = c_1\lambda_1^n V_1 + \cdots + c_k\lambda_k^n V_k$ d'où :

Proposition 2.82 *Supposons que $\chi_A(X) = (-1)^k \left(X^k - a_{k-1}X^{k-1} - \cdots - a_1X - a_0\right)$ a k racines distinctes $\lambda_i, \ldots, \lambda_k$. Alors, les suites qui vérifient la relation de récurrence (2.4) sont les suites définies par*

$$u_n = c_1\lambda_1^n + \cdots + c_k\lambda_k^n \ ,$$

où $c_1, \cdots, c_k$ sont des constantes quelconques.

Remarque 2.83 *Les coefficients inconnus $c_1, \cdots, c_k$ sont déterminés par les k premières valeurs $u_0, \cdots, u_{k-1}$ de la suite.*

Le cas où les racines ne sont pas distinctes est plus complexe. Traitons ce cas au travers d'un exemple. Considérons la suite récurrente suivante vérifiant pour tout $n \geqslant 0$:

$$u_{n+3} = u_{n+2} + 8u_{n+1} - 12u_n$$

La matrice A définie comme ci-dessus vaut :

$$A = \begin{pmatrix} 1 & 8 & -12 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

Et d'après ce qui a été fait plus haut :

$$\chi_A(\lambda) = -(\lambda^3 - \lambda^2 - 8\lambda + 12) = -(\lambda - 2)^2(\lambda + 3)$$

Donc les valeurs propres de A sont 2 et -3 de multiplicités respectives 2 et 1.

D'autre part $\dim(E_2(A)) = 1$ et $E_2(A) = \mathrm{Vect}(U)$ où $U = \begin{pmatrix} 4 \\ 2 \\ 1 \end{pmatrix}$. De plus le système :

$$(A - 2I_3)^2 \begin{pmatrix} x \\ y \\ z \end{pmatrix} = 0$$

est équivalent à $x - 4y + 4z = 0$ d'où l'on déduit qu'en posant $V = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$, $\{U, V\}$ est une base de l'espace $C_2(A)$. De plus $AV = 2V - U$.

Comme $E_{-3}(A) = \mathrm{Vect}(W)$ où $W = \begin{pmatrix} -9 \\ 3 \\ -1 \end{pmatrix}$, (U, V, W) est une base de $\mathbb{R}^3$. Donc si on pose $P = \begin{pmatrix} 4 & 0 & -9 \\ 2 & 1 & 3 \\ 1 & 1 & -1 \end{pmatrix}$ on a :

$$P^{-1}AP = \begin{pmatrix} 2 & -1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -3 \end{pmatrix} = D + N$$

où $D = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -3 \end{pmatrix}$ et $N = \begin{pmatrix} 0 & -1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$. Comme N est nilpotente d'indice 2, on déduit que pour $n \geqslant 1$:

$$\begin{pmatrix} u_{n+2} \\ u_{n+1} \\ u_n \end{pmatrix} = A^n \begin{pmatrix} u_2 \\ u_1 \\ u_0 \end{pmatrix} = PD^nP^{-1} \begin{pmatrix} u_2 \\ u_1 \\ u_0 \end{pmatrix} = P(D^n + nD^{n-1}N)C$$

où $C = P^{-1}\begin{pmatrix}u_2\\u_1\\u_0\end{pmatrix} = \begin{pmatrix}c_1\\c_2\\c_3\end{pmatrix}$. Donc :

$$\begin{pmatrix}u_{n+2}\\u_{n+1}\\u_n\end{pmatrix} = \begin{pmatrix}4 & 0 & -9\\2 & 1 & 3\\1 & 1 & -1\end{pmatrix}\begin{pmatrix}2^n & -n2^{n-1} & 0\\0 & 2^n & 0\\0 & 0 & (-3)^n\end{pmatrix}\begin{pmatrix}c_1\\c_2\\c_3\end{pmatrix}$$

et

$$u_n = c_1 2^n + c_2(-n2^{n-1} + 2^n) - c_3(-3)^n$$

Chapitre 3

Systèmes différentiels linéaires à coefficients constants

Dans toute la suite le corps $\mathbb{K}$ considéré est $\mathbb{R}$ ou $\mathbb{C}$.

3.1 Exponentielle d'une matrice ou d'un endomorphisme

3.1.1 Espaces vectoriels normés

La notion d'espace vectoriel normé fait l'objet d'un cours spécifique au S4. Nous nous contentons ici d'énoncer les résultats de cette théorie qui nous seront utiles pour résoudre les systèmes différentiels.

Définition 3.1 *Soit E un $\mathbb{K}$-espace vectoriel. On appelle* **norme** *sur E toute application $N : E \longrightarrow \mathbb{R}^+$ vérifiant :*

1. $N(x) = 0 \Leftrightarrow x = 0$.
2. $\forall(\lambda, x) \in \mathbb{K} \times E \quad ; \quad N(\lambda x) = |\lambda| N(x)$.
3. $\forall(x, y) \in E \times E \quad ; \quad N(x+y) \leqslant N(x) + N(y)$.

Un espace vectoriel muni d'une norme est appelé **espace vectoriel normé**.

Dans la pratique on note pour tout $x \in E$, $N(x) = \|x\|$ et $(E, \| \quad \|)$ désigne l'espace vectoriel normé associé.

Exercice de cours 3.1 *Soit $E = \mathbb{R}^2$. On note, pour $(x, y) \in E$,*

— $\|(x,y)\|_1 = |x| + |y|$;
— $\|(x,y)\|_2 = \sqrt{x^2 + y^2}$;
— $\|(x,y)\|_\infty = \sup(|x|, |y|)$

Montrer que $\| \ \|_1$, $\| \ \|_2$ et $\| \ \|_\infty$ sont des normes. Quelles sont les "cercles unités" pour ces normes, définis par l'équation $\|(x,y)\|_i = 1$?

Voici maintenant quelques résultats importants pour la suite :

Définition 3.2 *On dit qu'une suite* $(x_n)_{n\in\mathbb{N}}$ *de* E *tend vers* $x \in E$ *si* $\|x_n - x\| \longrightarrow 0$ *quand* $n \longrightarrow \infty$.

Définition 3.3 *Deux normes* $\| \ \|_1$ *et* $\| \ \|_2$ *sont* **équivalentes** *s'il existe* $\alpha, \beta > 0$ *tels que pour tout* $x \in E$ *:*

$$\alpha\|x\|_1 \leqslant \|x\|_2 \leqslant \beta\|x\|_1$$

Les suites convergentes pour $\| \ \|_1$ *ou* $\| \ \|_2$ *sont alors les mêmes et ont la même limite.*

Théorème 3.4 *Si* E *est de* **dimension finie** *alors toutes les normes sont équivalentes.*

Proposition 3.5 *Si* E *est de dimension finie muni d'une base* $(e_1, \cdots, e_n)$ *alors toute suite* $(x_k)_{k\in\mathbb{N}}$ *de* E *telle que* $x_k = \sum_{i=1}^{n} x_{k,i}e_i$ *pour tout* $k \in \mathbb{N}$ *tend vers* $x = \sum_{i=1}^{n} x_i e_i$ *si et seulement si pour tout* $i \in \{1, \cdots, n\}$ *:*

$$x_{k,i} \longrightarrow x_i \quad \textit{quand} \quad k \longrightarrow \infty$$

Preuve: Il suffit de prendre la norme sup, définie par $\| \sum_{i=1}^{n} x_i e_i\|_\infty = \sup_i |x_i|$, pour obtenir l'équivalence indiquée. □

Proposition 3.6 *Si* $(E, \| \ \|)$ *est un* $\mathbb{K}$*-espace vectoriel normé de dimension finie alors* $(E, \| \ \|)$ *est* **complet**. *Autrement dit toute suite de* **Cauchy** *dans* E *est convergente. On rappelle qu'une suite* $(x_n)_{n\in\mathbb{N}}$ *est de Cauchy si :*

$$\forall \varepsilon > 0 \quad , \quad \exists N \quad ; \quad \forall p, q \geqslant N \quad , \quad \|x_p - x_q\| < \varepsilon$$

Preuve: Si on utilise la norme sup $\| \ \|_\infty$, cela résulte de la Proposition 3.5. □

Définition 3.7 *Si* $(E, \| \ \|_E)$ *et* $(F, \| \ \|_F)$ *sont deux espaces vectoriels normés, alors* $f : E \longrightarrow F$ *est continue en* $x \in E$ *si pour toute suite* $(x_n)_{n\in\mathbb{N}}$ *de* E *telle que* $\lim_{n\longrightarrow\infty} x_n = x$ *pour la norme* $\| \ \|_E$ *alors* $\lim_{n\longrightarrow\infty} f(x_n) = f(x)$ *pour la norme* $\| \ \|_F$. *Là encore si* E *et* F *sont de dimension finie l'ensemble des fonctions continues est le même quel que soit le choix des normes sur* E *et* F.

À partir de maintenant $(E, \| \ \|)$ est un espace vectoriel normé de dimension finie n.

Donnons un exemple important de norme dans un espace vectoriel de dimension finie : soit $(e_1, \cdots, e_n)$ est une base de E. Pour tout $x = \sum_{i=1}^{n} x_i e_i$, posons :

$$\|x\|_1 = \sum_{i=1}^{n} |x_i|$$

Alors $\| \ \|_1$ est une norme sur E.

Définition 3.8 *Soit* $\sum_{k\geqslant 0} u_k$ *une série de* E*. On dit* $\sum_{k\geqslant 0} u_k$ **converge absolument** *si la série* $\sum_{k\geqslant 0} \|u_k\|$ *converge.*

Proposition 3.9 *Soit* $\sum_{k\geqslant 0} u_k$ *une série de* E *qui converge absolument. Alors* $\sum_{k\geqslant 0} u_k$ *est convergente.*

Preuve: Posons $S_N = \sum_{k=0}^{N} u_k$. Alors pour tout $p > q$:

$$\|S_p - S_q\| = \left\| \sum_{k=q+1}^{p} u_k \right\| \leqslant \sum_{k=q+1}^{p} \|u_k\|$$

Comme $\sum_{k\geqslant 0} \|u_k\|$ est une série convergente la suite $\left(\sum_{k=0}^{N} \|u_k\| \right)_{N\in\mathbb{N}}$ est de Cauchy. Alors, pour tout $\varepsilon > 0$:

$$\exists N \quad ; \quad \forall p,q \geqslant N \quad , \quad \sum_{k=q+1}^{p} \|u_k\| < \varepsilon$$

Donc :

$$\forall p,q \geqslant N \quad , \quad \|S_p - S_q\| < \varepsilon$$

Donc $(S_N)_{N\in\mathbb{N}}$ est de Cauchy dans E qui est complet car de dimension finie. Donc $(S_N)_{N\in\mathbb{N}}$ est une suite convergente ce qui signifie que $\sum_{k\geqslant 0} u_k$ converge. □

3.1.2 Norme sur l'espace des endomorphismes et sur les matrices

Dans toute cette partie $(E, \| \ \|)$ est une espace vectoriel normé de dimension finie n.

Proposition 3.10 *Soit* $u \in \mathcal{L}(E)$*. Alors il existe une constante* M *telle que pour tout* $x \in E$ *:*

$$\|u(x)\| \leqslant M\|x\|$$

De plus u *est continue sur* $(E, \| \ \|)$*.*

Preuve: Soit $(e_1, \cdots, e_n)$ une base de E et $x = \sum_{i=1}^{n} x_i e_i$ alors :

$$\|u(x)\| = \left\| \sum_{i=1}^{n} x_i u(e_i) \right\| \leqslant \sum_{i=1}^{n} |x_i| \|u(e_i)\| \leqslant \max_{1\leqslant i\leqslant n} (\|u(e_i)\|) \sum_{i=1}^{n} |x_i| = \max_{1\leqslant i\leqslant n} (\|u(e_i)\|) \|x\|_1$$

Or toutes les normes étant équivalentes il existe un réel $\beta > 0$ tel que pour tout $x \in E$, $\|x\|_1 \leqslant \beta\|x\|$. Donc posant $M = \beta \max\limits_{1 \leqslant i \leqslant n} (\|u(e_i)\|)$, on a prouvé que pour tout $x \in E$:

$$\|u(x)\| \leqslant M\|x\|$$

Maintenant si $(x_k)_{k\in\mathbb{N}}$ est une suite de E qui converge vers x, alors :

$$\|u(x_k) - u(x)\| = \|u(x_k - x)\| \leqslant M\|x_k - x\|$$

Comme $\lim\limits_{k \longrightarrow \infty} \|x_k - x\| = 0$, on conclut que $\lim\limits_{k \longrightarrow \infty} u(x_k) = u(x)$ et u est continue. □

Exemple 3.11 *Soit $f_A : \mathbb{R}^n \to \mathbb{R}^n$ l'application linéaire associée à une matrice $A = (a_{i,j})$, et soit la norme $\| \ \|_1$ sur $\mathbb{R}^n$. On a $\|AX\|_1 \leq \left(\sum_{i,j} |a_{i,j}|\right) \|X\|_1$.*

Preuve: En effet,

$$\begin{array}{rcl} \sum_i \sum_j |a_{i,j}x_j| & \leq & \sum_{i,j} |a_{i,j}x_j| \\ & = & \sum_{i,j} |a_{i,j}|\,|x_j| \\ & \leq & \left(\sum_{i,j} |a_{i,j}|\right)\ (\sum_k |x_k|) \end{array}$$

□

De la proposition on déduit que $\sup\limits_{\substack{x\in E \\ \|x\|=1}} \|u(x)\|$ est bien défini.

Proposition 3.12 *Pour tout $u \in \mathcal{L}(E)$ définissons :*

$$|||u||| := \sup_{\substack{x\in E \\ \|x\|=1}} \|u(x)\|$$

Alors $||| \ \ |||$ est une norme sur E qui satisfait :

1. *$\forall x \in E$, $\|u(x)\| \leqslant |||u|||\|x\|$.*
2. *$\forall (u, v) \in (\mathcal{L}(E))^2$, $|||u \circ v||| \leqslant |||u||||||v|||$.*

$||| \ \ |||$ s'appelle **norme d'algèbre**.

Preuve: En effet, on déduit de la définition de $|||u|||$ que $\|u(x)\| \leq |||u|||\ \|x\|$ pour tout $x \in E$. Soit maintenant $x \in E$, on a

$$\|u \circ v(x)\| \leq |||u|||\ \|v(x)\| \leq |||u|||\ |||v|||\ \|x\| \ ,$$

en appliquant deux fois cette propriété. □

Proposition 3.13 *On considère l'espace vectoriel des matrices colonnes $E = \mathcal{M}_{n,1}(\mathbb{K})$ muni d'une norme $\| \ \ \|$. Soit $A \in \mathcal{M}_n(\mathbb{K})$ et définissons :*

$$|||A||| := \sup_{\substack{X\in E \\ \|X\|=1}} \|AX\|$$

Alors $||| \ \ |||$ est une norme sur E qui satisfait :

1. $\forall X \in E \ , \ \|AX\| \leqslant |||A||| \|X\|$.

2. $\forall (A,B) \in (\mathcal{M}_n(\mathbb{K}))^2 \ , \ |||AB||| \leqslant |||A||| |||B|||$.

$|||\ \ |||$ *s'appelle* **norme matricielle**.

PREUVE: Pour tout $A \in \mathcal{M}_n(\mathbb{K})$, considérons l'application :

$$\begin{array}{cccc} f_A & : & E & \longrightarrow & E \\ & & X & \longmapsto & AX \end{array}$$

Alors f_A est un endomorphisme de E et $|||A||| = |||f_A|||$. L'assertion 1. est évidente. Pour l'assertion 2. on a :

$$|||AB||| = |||f_{AB}||| = |||f_A \circ f_B||| \leqslant |||f_A||| |||f_B||| = |||A||| |||B|||$$

$\square$

3.1.3 Dérivabilité et intégration des fonctions à variable réelle et à valeurs dans un espace vectoriel

Dans toute cette partie $(E, \|\ \ \|)$ est une espace vectoriel normé de dimension finie n.

Définition 3.14 *Soit I un intervalle non vide de $\mathbb{R}$. On dit que la fonction $f : I \longrightarrow E$ est* **dérivable** *en $t_0 \in I$, s'il existe un vecteur de E noté $f'(t_0)$ tel que :*

$$\left\| \frac{1}{h}(f(t_0+h) - f(t_0)) - f'(t_0) \right\| \longrightarrow 0$$

quand $h \longrightarrow 0$.

Si f est dérivable sur I, l'application :

$$\begin{array}{cccc} f' & : & I & \longrightarrow & E \\ & & t & \longmapsto & f'(t) \end{array}$$

s'appelle la dérivée de f sur I.

Proposition 3.15 *Soit I un intervalle non vide de $\mathbb{R}$. On suppose que E est muni d'une base $\mathcal{B}$. Soit $f : I \longrightarrow E$ une fonction. Notons $(f_1, \cdots, f_n)$ les fonctions composantes de f dans la base $\mathcal{B}$ (i.e. $\forall t \in I \ , \ f(t) = \sum_{i=1}^{n} f_i(t) e_i$). Alors f est dérivable en $t_0 \in I$ si et seulement si les fonctions f_i sont dérivables en $t_0 \in I$ pour tout $i \in \{1, \cdots, n\}$, et dans ce cas $f'(t_0) = \sum_i f_i'(t_0) e_i$.*

PREUVE: On peut supposer que $\|x\| = \|x\|_1$ où $\|x\|_1$ est défini via la base $\mathcal{B}$. Soit $d = \sum a_i e_i$: on a

$$\left\| \frac{1}{h}(f(t_0+h) - f(t_0)) - \sum a_i e_i \right\| \longrightarrow 0$$

si et seulement si

$$\sum_i \left| \frac{1}{h}(f_i(t_0+h) - f_i(t_0)) - a_i \right| \longrightarrow 0\,,$$

ce qui équivaut finalement à

$$\forall i, \left| \frac{1}{h}(f_i(t_0+h) - f_i(t_0)) - a_i \right| \longrightarrow 0\,,$$

soit le fait que chaque fonction f_i est dérivable en t_0 de dérivée $f_i'(t_0) = a_i$. □

Proposition 3.16 *Soit I un intervalle non vide de $\mathbb{R}$ et soit $f : I \longrightarrow E$ une fonction.*

1. *Soient $\mathcal{B}$ et $\mathcal{C}$ deux bases de E. Alors les fonctions composantes de f dans $\mathcal{B}$ sont intégrables au sens de Riemann sur I si et seulement si les fonctions composantes de f dans $\mathcal{C}$ sont intégrables au sens de Riemann sur I. Si les fonctions composantes de f dans une base sont intégrables, on dira alors que f est intégrable.*
2. *Soit $\mathcal{B} = (u_1, \cdots, u_n)$ une base de E et $(f_1, \cdots, f_n)$ les composantes de f dans $\mathcal{B}$. Alors la quantité :*

$$\sum_{i=1}^{n} \int_I (f_i(t)dt)\, u_i$$

ne dépend pas de la base choisie et est notée $\int_I f(t)dt$.

Preuve: On note $\mathcal{B} = (u_1, \cdots, u_n)$ et $\mathcal{C} = (v_1, \cdots, v_n)$. On note aussi P la matrice de passage de $\mathcal{B}$ à $\mathcal{C}$, de sorte que par définition $\forall j, v_j = \sum_i P_{i,j} u_i$. Soient (f_i) les coordonnées de f dans $\mathcal{B}$ et (g_j) les coordonnées de f dans $\mathcal{C}$: on va exprimer une relation entre les fonctions (f_i) et (g_j). Pour cela, on calcule :

$$\sum_i f_i(t) u_i = f(t) = \sum_j g_j(t) v_j = \sum_j g_j(t) \left(\sum_i P_{i,j} u_i \right) = \sum_i \left(\sum_j P_{i,j} g_j(t) \right) u_i\,.$$

On en déduit que

$$\forall i, f_i(t) = \sum_j P_{i,j} g_j(t)\,.$$

Ainsi, si les fonctions g_j sont intégrables, les fonctions f_i le sont aussi. Par symétrie, on obtient l'équivalence.

De plus,

$$\begin{array}{rcl} \sum_i \left(\int_I f_i(t)dt \right) u_i & = & \sum_i \left(\int_I \sum_j P_{i,j} g_j(t)dt \right) u_i \\ & = & \sum_j \left(\int_I g_j(t)dt \right) \sum_i P_{i,j} u_i \\ & = & \sum_j \left(\int_I g_j(t)dt \right) v_j\,. \end{array}$$

Ceci montre que la quantité $\sum_{i=1}^n \int_I (f_i(t)dt)\, u_i$ ne dépend pas de la base (u_i) choisie. □

Désormais l'espace $E = \mathcal{M}_{n,1}(\mathbb{K})$ est muni d'une norme et $\mathcal{M}_n(\mathbb{K})$ est muni de la norme matricielle associée.

Proposition 3.17 *Soient :*

$$A : \begin{array}{ccc} I & \longrightarrow & \mathcal{M}_{m,n}(\mathbb{K}) \\ t & \longmapsto & A(t) \end{array}, \quad B : \begin{array}{ccc} I & \longrightarrow & \mathcal{M}_{n,p}(\mathbb{K}) \\ t & \longmapsto & B(t) \end{array}$$

deux fonctions dérivables sur I. Alors la fonction

$$AB : \begin{array}{ccc} I & \longrightarrow & \mathcal{M}_n(\mathbb{K}) \\ t & \longmapsto & A(t)B(t) \end{array}$$

est dérivable sur I et :

$$\forall t \in I, (AB)'(t) = A'(t)B(t) + A(t)B'(t).$$

Preuve: Soit $t, h \in \mathbb{R}$ tels que $[t-h, t+h] \subset I$. Comme A et B sont dérivables en t, on a

$$\begin{aligned} A(t+h) &= A(t) + hA'(t) + o(h) \\ B(t+h) &= B(t) + hB'(t) + o(h) \end{aligned}$$

On en déduit que $A(t+h)B(t+h) = A(t)B(t) + h(A(t)B'(t) + A'(t)B(t)) + o(h)$. Ainsi, AB est dérivable en t et $(AB)'(t) = A(t)B'(t) + A'(t)B(t)$. □

3.1.4 Exponentielle d'une matrice

Proposition et définition 3.18 *Pour tout $A \in \mathcal{M}_n(\mathbb{K})$, la série $\sum_{k \geqslant 0} \frac{1}{k!} A^k$ est convergente. Sa somme notée e^A ou $\exp(A)$ s'appelle l'***exponentielle** *de A. De plus :*

$$|||e^A||| \leqslant e^{|||A|||}$$

Preuve: $\left|\left|\left|\frac{1}{k!}A^k\right|\right|\right| \leqslant \frac{1}{k!}|||A^k||| \leqslant \frac{1}{k!}|||A|||^k$. Or la série $\sum_{k \geqslant 0} \frac{1}{k!}|||A|||^k$ est convergente donc $\sum_{k \geqslant 0} \frac{1}{k!} A^k$ est convergente. Posons $s_k = \sum_{i=0}^{k} \frac{1}{i!} A^i$. Comme :

$$\left|\left|\left|\sum_{i=0}^{k} \frac{1}{i!} A^i\right|\right|\right| \leqslant \sum_{i=0}^{k} \left|\left|\left|\frac{1}{i!} A^i\right|\right|\right| \leqslant \sum_{i=0}^{k} \frac{1}{i!}|||A|||^i \leqslant e^{|||A|||},$$

on a $|||s_k||| \leq e^{|||A|||}$ pour tout k. Soit $\epsilon > 0$ et soit k_0 tel que $|||e^A - s_k||| \leq \epsilon$ pour tout $k \geq k_0$: on en déduit que $|||e^A||| \leq |||s_k + (e^A - s_k)||| \leq |||s_k||| + |||e^A - s_k||| \leq e^A + \epsilon$. Comme cette inégalité est vraie pour tout ϵ, on a $|||e^A||| \leqslant e^{|||A|||}$. □

Proposition 3.19 *1. Soit $A \in \mathcal{M}_n(\mathbb{K})$ et $P \in GL_n(\mathbb{K})$. Alors :*

$$e^{P^{-1}AP} = P^{-1} e^A P$$

2. *Soit $(A,B) \in \mathcal{M}_n(\mathbb{K})^2$. Si $AB = BA$ alors :*
$$e^A e^B = e^{A+B} = e^B e^A$$
et
$$Be^A = e^A B$$
3. *Soit $A \in \mathcal{M}_n(\mathbb{K})$, alors e^A est toujours inversible et :*
$$(e^A)^{-1} = e^{-A}$$

Preuve:

1. Soit $A \in \mathcal{M}_n(\mathbb{K})$ et $P \in GL_n(\mathbb{K})$. Alors :
$$P^{-1}\left(\sum_{i=0}^{k} \frac{1}{i!}A^i\right)P = \sum_{i=0}^{k} \frac{1}{i!}P^{-1}A^iP = \sum_{i=0}^{k} \frac{1}{i!}(P^{-1}AP)^i$$
Or l'application $X \longmapsto P^{-1}XP$ est un endomorphisme de $\mathcal{M}_n(\mathbb{K})$ qui est de dimension finie, donc elle est continue et :
$$P^{-1}e^AP = P^{-1}\left(\lim_{k\longrightarrow\infty}\sum_{i=0}^{k} \frac{1}{i!}A^i\right)P = \lim_{k\longrightarrow\infty} P^{-1}\left(\sum_{i=0}^{k} \frac{1}{i!}A^i\right)P\,.$$
Cette limite vaut donc
$$\lim_{k\longrightarrow\infty}\sum_{i=0}^{k} \frac{1}{i!}(P^{-1}AP)^i = e^{P^{-1}AP}\,.$$
2. Pour montrer le point 2. nous avons besoin du lemme suivant :

Lemme 3.20 *Soient $A = \sum_{k\geqslant 0} A_k$ et $B = \sum_{k\geqslant 0} B_k$ deux séries absolument convergentes de $\mathcal{M}_n(\mathbb{K})$. Alors la série $C = \sum_{k\geqslant 0}\left(\sum_{l=0}^{k} A_kB_{k-l}\right)$ est absolument convergente et $C = AB$.*

Notons, par convergence uniforme de A et B,
$$a := \sum_{i\geq 0} |||A_i||| \quad \text{et} \quad b := \sum_{i\geq 0} |||B_j|||\,.$$
La convergence uniforme est conséquence de la remarque suivante :
$$\begin{aligned}\textstyle\sum_{l=0}^{n} |||C_l||| \leq \sum_{l=0}^{n}\left(\sum_{k=0}^{l} |||A_kB_{l-k}|||\right) &\leq \textstyle\sum_{l=0}^{n}\sum_{k=0}^{l} |||A_k|||\ |||B_{l-k}||| \\ &\leq \textstyle\sum_{i=0}^{n}\sum_{j=0}^{n} |||A_i|||\ |||B_j||| \\ &\leq \textstyle\left(\sum_{i=0}^{n} |||A_i|||\right)\left(\sum_{j=0}^{n} |||B_j|||\right) \\ &\leq \textstyle\left(\sum_{i=0}^{\infty} |||A_i|||\right)\left(\sum_{j=0}^{\infty} |||B_j|||\right) \\ &\leq ab\,.\end{aligned}$$

Ces inégalités montrent d'ailleurs que la série de terme général $c_l := \sum_{k=0}^{l} |||A_k|||\ |||B_{l-k}|||$ est convergente. Soit donc N tel que $\forall n \geq N$,

$$\left\{\begin{array}{rcl} \sum_{i>n} |||A_i||| & < & \epsilon \\ \sum_{j>n} |||B_j||| & < & \epsilon \\ \sum_{l>n} c_l & < & \epsilon \end{array}\right.$$

La dernière inégalité implique que $|||C - \sum_{l=0}^{n} C_l||| \leq \sum_{l>n} c_l \leq \epsilon$.
Pour montrer que $C = AB$, on fixe $\epsilon > 0$ et on montre que $|||C - AB||| \leq (a + b + 2)\epsilon$.
On a :

$$\begin{array}{rcl} |||C - \sum_{i,j\leq n} A_iB_j||| & \leq & |||C - \sum_{l=0}^{n} C_l||| + |||\sum_{l=0}^{n} C_l - \sum_{i,j\leq n} A_iB_j||| \\ & \leq & \epsilon + \sum_{i,j\leq n; i+j>n} |||A_iB_j||| \\ & \leq & \epsilon + \sum_{i,j\leq n; i+j>n} |||A_i|||\ |||B_j||| \\ & \leq & \epsilon + \sum_{i,j: n<i+j\leq 2n} |||A_i|||\ |||B_j||| \\ & = & \epsilon + \sum_{l=n+1}^{2n} c_l \\ & \leq & 2\epsilon \end{array}$$

Par ailleurs, on a aussi :

$$\begin{array}{rcl} |||AB - \sum_{i,j\leq n} A_iB_j||| & = & |||\sum_{i\geq 0} A_iB - \sum_{i,j\leq n} A_iB_j||| \\ & = & |||\sum_{i\leq n} A_i(B - \sum_{j\leq n} B_j) + \sum_{i>n} A_iB||| \\ & \leq & \sum_{i\leq n} |||A_i|||\ |||B - \sum_{j\leq n} B_j||| + |||(\sum_{i>n} A_i)B||| \\ & \leq & \sum_{i\leq n} |||A_i|||\epsilon + |||\sum_{i>n} A_i|||\ |||B||| \\ & \leq & a\epsilon + \epsilon b \\ & \leq & (a+b)\epsilon \end{array}$$

En mettant ensemble les deux dernières inégalités obtenues, on déduit :

$$|||C - AB||| \leq |||C - \sum_{i,j\leq n} A_iB_j||| + |||\sum_{i,j\leq n} A_iB_j - AB||| \leq 2\epsilon + (a+b)\epsilon\,,$$

comme souhaité.

Montrons maintenant le point 2. Comme A et B commutent on peut appliquer la formule du binôme de Newton :

$$\begin{aligned} \frac{1}{k!}(A+B)^k &= \frac{1}{k!}\sum_{i=0}^{k}\binom{k}{i} A^iB^{k-i} \\ &= \frac{1}{k!}\sum_{i=0}^{k}\frac{k!}{i!(k-i)!}A^iB^{k-i} \\ &= \sum_{i=0}^{k}\left(\frac{1}{i!}A^i\right)\left(\frac{1}{(k-i)!}B^{k-i}\right) \end{aligned}$$

Comme $\sum_{k\geqslant 0}\frac{1}{k!}A^k$ et $\sum_{k\geqslant 0}\frac{1}{k!}B^k$ sont absolument convergentes, d'après le lemme la série $\sum_{k\geqslant 0}\frac{1}{k!}(A+B)^k$ est absolument convergente et :

$$\sum_{k=0}^{\infty}\frac{1}{k!}(A+B)^k=\sum_{k=0}^{\infty}\sum_{i=0}^{k}\left(\frac{1}{i!}A^i\right)\left(\frac{1}{(k-i)!}B^{k-i}\right)=\left(\sum_{k=0}^{\infty}\frac{1}{k!}A^k\right)\left(\sum_{k=0}^{\infty}\frac{1}{k!}B^k\right)$$

d'où le résultat.

$\square$

Proposition 3.21 *Soit $A\in\mathcal{M}_n(\mathbb{K})$. Alors l'application $\mathbb{R}\longrightarrow\mathcal{M}_n(\mathbb{K})$, $t\longmapsto e^{tA}$ est de classe C^∞ et $(e^{tA})'=Ae^{tA}$.*

Preuve: Montrons d'abord que $t\longmapsto e^{tA}$ est dérivable :

$$\begin{aligned}\left|\left|\left|\frac{1}{h}(e^{(t+h)A}-e^{tA})-Ae^{tA}\right|\right|\right| &= \left|\left|\left|e^{tA}\left(\frac{1}{h}(e^{hA}-I)-A\right)\right|\right|\right| \\ &= \left|\left|\left|e^{tA}\left(\sum_{k=2}^{\infty}\frac{h^{k-1}}{k!}A^k\right)\right|\right|\right| \\ &= |h|\left|\left|\left|e^{tA}\left(\sum_{k=2}^{\infty}\frac{h^{k-2}}{k!}A^k\right)\right|\right|\right|\end{aligned}$$

Supposons $|h|\leqslant 1$. Donc :

$$\begin{aligned}\left|\left|\left|\frac{1}{h}(e^{(t+h)A}-e^{tA})-Ae^{tA}\right|\right|\right| &\leqslant |h||||e^{tA}|||\sum_{k=2}^{\infty}\frac{1}{k!}|||A|||^k \\ &\leqslant |h||||e^{tA}|||e^{|||A|||}\longrightarrow 0\end{aligned}$$

quand $h\longrightarrow 0$. Ceci montre que $t\mapsto e^{tA}$ est dérivable de dérivée $t\mapsto Ae^{tA}$. $\square$

Proposition 3.22 *Soit $D=\begin{pmatrix}\lambda_1 & 0 & \cdots & 0\\ 0 & \ddots & \ddots & \vdots\\ \vdots & \ddots & \ddots & 0\\ 0 & \cdots & 0 & \lambda_n\end{pmatrix}$ une matrice diagonale. Alors :*

$$e^D=\begin{pmatrix}e^{\lambda_1} & 0 & \cdots & 0\\ 0 & \ddots & \ddots & \vdots\\ \vdots & \ddots & \ddots & 0\\ 0 & \cdots & 0 & e^{\lambda_n}\end{pmatrix}$$

En particulier pour tout $\lambda\in\mathbb{K}$, on a $e^{\lambda I_n}=e^{\lambda}I_n$.

PREUVE: $e^D = \lim_{N \longrightarrow +\infty} \left(\sum_{k=0}^{N} \frac{1}{k!} D^k \right)$. Or :

$$\sum_{k=0}^{N} \frac{1}{k!} D^k = \sum_{k=0}^{N} \frac{1}{k!} \begin{pmatrix} \lambda_1^k & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n^k \end{pmatrix} = \begin{pmatrix} \sum_{k=0}^{N} \frac{1}{k!} \lambda_1^k & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \sum_{k=0}^{N} \frac{1}{k!} \lambda_n^k \end{pmatrix}$$

Notons $\sum_{k=0}^{N} \frac{1}{k!} D^k = (x_{ij}^N)_{1 \leqslant i,j \leqslant n}$ et $e^D = (x_{ij})_{1 \leqslant i,j \leqslant n}$. Pour tout $i, j \in \{1, \cdots, n\}$, notons E_{ij} la matrice dont tous les coefficients sont nuls sauf celui de la ligne i et de la colonne j qui est égal à 1.

Les coefficients de $\sum_{k=0}^{N} \frac{1}{k!} D^k$ et e^D ne sont rien d'autre que les coordonnées respectivement de $\sum_{k=0}^{N} \frac{1}{k!} D^k$ et de e^D dans la base $\{E_{ij}\}_{1 \leqslant i,j \leqslant n}$. Comme $\lim_{N \longrightarrow +\infty} \left(\sum_{k=0}^{N} \frac{1}{k!} D^k \right) = e^D$, $x_{ij}^N \longrightarrow x_{ij}$ ce qui prouve que :

$$e^D = \begin{pmatrix} \sum_{k=0}^{+\infty} \frac{1}{k!} \lambda_1^k & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \sum_{k=0}^{+\infty} \frac{1}{k!} \lambda_n^k \end{pmatrix} = \begin{pmatrix} e^{\lambda_1} & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & e^{\lambda_n} \end{pmatrix}$$

□

3.2 Calcul pratique de l'exponentielle d'une matrice

De manière générale si $A \in \mathcal{M}_n(\mathbb{K})$ et si son polynôme caractéristique est scindé, alors d'après la décomposition de Dunford, il existe une matrice D diagonalisable et une matrice N nilpotente telles que :

1. $A = D + N$.
2. $DN = ND$.

Donc comme D et N commutent :

$$e^A = e^{D+N} = e^D e^N$$

Comme D est diagonalisable il existe $P \in GL_n(\mathbb{K})$ telle que :

$$P^{-1}DP = \begin{pmatrix} \mu_1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \mu_n \end{pmatrix}$$

D'autre part N est nilpotente donc il existe un entier $p \leqslant n$ tel que $N^p = 0$. Donc :

$$e^A = P \begin{pmatrix} e^{\mu_1} & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & e^{\mu_n} \end{pmatrix} P^{-1} \left(I_n + N + \frac{1}{2!}N^2 + \cdots + \frac{1}{(p-1)!}N^{p-1} \right)$$

En fait on n'a pas besoin de calculer D et N. En effet soient $\lambda_1, \cdots, \lambda_r$ les valeurs propres de A et pour tout $i \in \{1, \cdots, r\}$, $C_{\lambda_i} = \mathrm{Ker}(A - \lambda_i I_n)^{\alpha_i}$ les sous-espaces caractéristiques de A. Pour tout $i \in \{1, \cdots, r\}$ on construit une base $\mathcal{B}_i$ de chacun des sous-espaces C_{λ_i}. Comme $\mathcal{M}_{n,1}(\mathbb{K}) = \bigoplus_{i=1}^{r} C_{\lambda_i}$, $\mathcal{B} = \bigcup_{i=1}^{r} \mathcal{B}_i$ est une base de $\mathcal{M}_{n,1}(\mathbb{K})$. Soit P la matrice de $\mathcal{B}$. Comme les sous-espaces caractéristiques sont stables on a :

$$A' = P^{-1}AP = \begin{pmatrix} A_1 & & O \\ & \ddots & \\ O & & A_r \end{pmatrix}$$

où les matrices $A_i \in \mathcal{M}_{\alpha_i}(\mathbb{K})$. Dans cette base D est diagonale, c'est-à-dire que

$$P^{-1}DP = \begin{pmatrix} \lambda_1 I_{\alpha_1} & & O \\ & \ddots & \\ O & & \lambda_r I_{\alpha_r} \end{pmatrix}$$

Posons $N' = P^{-1}NP$ alors :

$$N' = P^{-1}NP = P^{-1}(A-D)P = A' - \begin{pmatrix} \lambda_1 I_{\alpha_1} & & O \\ & \ddots & \\ O & & \lambda_r I_{\alpha_r} \end{pmatrix}$$

et :

$$e^A = P \begin{pmatrix} e^{\lambda_1} I_{\alpha_1} & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & e^{\lambda_r} I_{\alpha_r} \end{pmatrix} \left(I_n + N' + \frac{1}{2!}N'^2 + \cdots + \frac{1}{(p-1)!}N'^{p-1} \right) P^{-1}$$

On peut encore raffiner les choses. De manière similaire à la preuve de la proposition 3.22 on peut montrer que :

$$e^A = e^{PA'P^{-1}} = Pe^{A'}P^{-1} = P \begin{pmatrix} e^{A_1} & & O \\ & \ddots & \\ O & & e^{A_r} \end{pmatrix} P^{-1}$$

Calculons e^{A_i} pour tout $i \in \{1, \cdots, r\}$. Alors :

$$e^{A_i} = e^{(A_i - \lambda_i I_{\alpha_i}) + \lambda_i I_{\alpha_i}} = e^{\lambda_i} e^{(A_i - \lambda_i I_{\alpha_i})}$$

Or :

$$N' = A' - \begin{pmatrix} \lambda_1 I_{\alpha_1} & & O \\ & \ddots & \\ O & & \lambda_r I_{\alpha_r} \end{pmatrix} = \begin{pmatrix} A_1 - \lambda_1 I_{\alpha_1} & & O \\ & \ddots & \\ O & & A_r - \lambda_r I_{\alpha_r} \end{pmatrix}$$

On a donc pour tout $i \in \{1, \cdots, r\}$, $(A_i - \lambda_i I_{\alpha_i})^p = 0$ et même $(A_i - \lambda_i I_{\alpha_i})^{\beta_i} = 0$ où β_i est l'indice de $A_i - \lambda_i I_{\alpha_i}$:

$$e^{A_i} = e^{\lambda_i} \left(I_{\alpha_i} + N_i' + \frac{1}{2!} N_i'^2 + \cdots + \frac{1}{(\beta_i - 1)!} N_i'^{\beta_i - 1} \right)$$

où $N_i' = A_i - \lambda_i I_{\alpha_i}$.

Remarque 3.23 *On a vu dans la proposition 3.12 que l'on pouvait munir $\mathcal{L}(E)$ de la norme d'algèbre $|||\ \ |||$. De la même manière que pour les matrices on peut définir l'exponentielle d'un endomorphisme et tous les énoncés précédents sont transposables dans le cadre des endomorphismes.*

3.3 Systèmes différentiels linéaires à coefficients constants

$\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$ et I désigne un intervalle non vide de $\mathbb{R}$. À partir de maintenant on identifie $\mathcal{M}_{n,1}(\mathbb{K})$ et $\mathbb{K}^n$.

Définition 3.24 *Soit $n \in \mathbb{N}^\star$. Pour tout $i \in \{1, \cdots, n\}$, soit $b_i : I \longrightarrow \mathbb{K}$ une fonction continue et pour tout $(i,j) \in \{1, \cdots, n\}^2$ soit $a_{ij} \in \mathbb{K}$.*

1. *Le système :*

$$(\mathcal{SL}) \qquad \begin{cases} y_1' & = & a_{11}y_1 + & \cdots & + a_{1n}y_n + b_1 \\ \vdots & & \vdots & & \vdots \\ y_n' & = & a_{n1}y_1 + & \cdots & + a_{nn}y_n + b_n \end{cases}$$

s'appelle ***système différentiel linéaire d'ordre*** 1 ***à coefficients constants****.*

2. *Le système :*

$$(\mathcal{SH}) \qquad \begin{cases} y_1' & = & a_{11}y_1 + & \cdots & + a_{1n}y_n \\ \vdots & & \vdots & & \vdots \\ y_n' & = & a_{n1}y_1 + & \cdots & + a_{nn}y_n \end{cases}$$

s'appelle ***système différentiel linéaire homogène associé à*** $(\mathcal{SL})$*.*

Résoudre $(\mathcal{SL})$ ou $(\mathcal{SH})$ revient à trouver un intervalle non vide $J \subset I$ et des fonctions $y_1, \cdots, y_n$ dérivables sur J et vérifiant le système considéré. Si on note $A = (a_{ij})_{1 \leqslant i,j \leqslant n}$, $Y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$ et $B = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$ alors :

$$(\mathcal{SL}) \Longleftrightarrow (\mathcal{L}) \qquad Y' = AY + B$$

et

$$(\mathcal{SH}) \Longleftrightarrow (\mathcal{H}) \qquad Y' = AY$$

Résoudre $(\mathcal{L})$ ou $(\mathcal{H})$ revient à trouver un intervalle non vide $J \subset I$ et une fonction $Y : J \longrightarrow \mathbb{K}^n$ dérivable sur J et vérifiant respectivement $(\mathcal{L})$ ou $(\mathcal{H})$.

3.3.1 Système différentiel linéaire homogène

Notons $\mathcal{C}(I, \mathbb{K}^n)$ l'espace vectoriel des fonctions continues de I dans $\mathbb{K}^n$.

Théorème 3.25 *L'ensemble $S_{\mathcal{H}}$ des solutions sur $\mathbb{R}$ de $\mathcal{H} : Y' = AY$ est un sous-espace vectoriel de $\mathcal{C}(\mathbb{R}, \mathbb{K}^n)$ de dimension n formé de fonctions de classe C^∞.*

Plus précisément :

$$S_{\mathcal{H}} := \{t \longmapsto e^{tA}C \quad \text{où} \quad C \in \mathbb{K}^n\}$$

De plus pour tout $(t_0, Y_0) \in \mathbb{R} \times \mathbb{K}^n$ il existe une unique solution Y de $\mathcal{H}$ telle que $Y(t_0) = Y_0$.

Preuve: Il est immédiat que la fonction nulle appartient à $S_{\mathcal{H}}$ et que si $(Y_1, Y_2) \in S_{\mathcal{H}}^2$ alors pour tout $(\lambda_1, \lambda_2) \in \mathbb{K}^2$:

$$\lambda_1 Y_1 + \lambda_2 Y_2 \in S_{\mathcal{H}}$$

Donc $S_{\mathcal{H}}$ est un sous-espace vectoriel de $\mathcal{C}(\mathbb{R}, \mathbb{K}^n)$.

Maintenant soit $Y \in S_{\mathcal{H}}$ et soit $Z(t) = e^{-tA}Y(t)$. Alors :

$$Z'(t) = -Ae^{-tA}Y(t) + e^{-tA}Y'(t) = -Ae^{-tA}Y(t) + e^{-tA}AY(t)$$

Or A et e^{-tA} commutent donc $Z'(t) = 0$ pour tout $t \in \mathbb{R}$. Donc il existe $C \in \mathbb{K}^n$ tel que $Z(t) = C$ pour tout $t \in \mathbb{R}$ et $Y(t) = e^{tA}C$.

Réciproquement on vérifie facilement que pour tout $C \in \mathbb{K}^n$, la fonction $t \mapsto Y(t) = e^{tA}C$ est solution de $S_{\mathcal{H}}$.

D'autre part soit $(t_0, Y_0) \in \mathbb{R} \times \mathbb{K}^n$. Alors :

$$Y(t_0) = Y_0 \Longleftrightarrow e^{t_0 A}C = Y_0 \Longleftrightarrow C = e^{-t_0 A}Y_0$$

□

Théorème 3.26 *Soit $A \in \mathcal{M}_n(\mathbb{K})$ telle que χ_A est scindé. Notons $\mu_1, \cdots, \mu_n$ les valeurs propres de A comptées avec multiplicité et soit $(U_1, \cdots, U_n)$ une base de $\mathbb{K}^n$ constituée de vecteurs appartenant aux sous-espaces caractéristiques. Plus précisément on suppose que :*

$$\forall i \in \{1, \cdots, n\} \ , \ U_i \in C_{\mu_i}(A)$$

Alors :

$$S_{\mathcal{H}} = \left\{ t \longmapsto \sum_{i=1}^{n} c_i e^{\mu_i t} \left(\sum_{k=0}^{\beta_i - 1} \frac{t^k}{k!} (A - \mu_i I_n)^k U_i \right) \ \text{où} \quad \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} \in \mathbb{K}^n \right\}$$

où β_i est la multiplicité de μ_i dans μ_A.

En particulier si A est diagonalisable alors :

$$S_{\mathcal{H}} = \left\{ t \longmapsto \sum_{i=1}^{n} c_i e^{\mu_i t} U_i \ \text{où} \quad \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} \in \mathbb{K}^n \right\}$$

Et dans ce cas $(U_1, \cdots, U_n)$ est une base de $\mathbb{K}^n$ constituée de vecteurs propres de A.

Preuve: Le polynôme caractéristique χ_A étant scindé on a :

$$\chi_A(\lambda) = (-1)^n \prod_{i=1}^{r} (\lambda_i - \lambda)^{\alpha_i}$$

et comme $\mathbb{K}^n = \bigoplus_{i=1}^{r} C_{\lambda_i}(A)$, on choisit une base $\mathcal{B} = (U_1, \cdots, U_n)$ de E telle que $\mathcal{B} = \bigcup_{i=1}^{r} \mathcal{B}_i$ où pour tout $i \in \{1, \cdots, r\}$, $\mathcal{B}_i$ est une base de $C_{\lambda_i}(A)$.

Y est une solution de $\mathcal{H} : Y' = AY$ sur $\mathbb{R}$ si et seulement si $Y(t) = e^{tA}C$ où $C \in \mathbb{K}^n$. Notons $(c_1, \cdots, c_n)$ les coordonnées de C dans la base $\mathcal{B}$. Alors $C = \sum_{i=1}^{n} c_i U_i$. Pour tout $i \in \{1, \cdots, n\}$, $U_i \in C_{\mu_i}(A)$ où $\mu_1, \cdots, \mu_n$ sont les valeurs propres de A comptées avec multiplicité. Donc :

$$\begin{aligned} Y(t) &= \sum_{i=1}^{n} c_i e^{tA} U_i = \sum_{i=1}^{n} c_i e^{\mu_i t I_n + t(A - \mu_i I_n)} U_i \\ &= \sum_{i=1}^{n} c_i e^{\mu_i t I_n} e^{t(A - \mu_i I_n)} U_i = \sum_{i=1}^{n} c_i e^{\mu_i t} e^{t(A - \mu_i I_n)} U_i \end{aligned}$$

D'autre part on a vu que si γ_i désigne la multiplicité de μ_i dans le polynôme minimal de A, alors :

$$\operatorname{Ker}(A - \mu_i I_n)^{\gamma_i - 1} \subsetneq \operatorname{Ker}(A - \mu_i I_n)^{\gamma_i} = C_{\mu_i}(A)$$

De plus comme $U_i \in C_{\mu_i}(A)$, pour tout $k \geqslant \gamma_i$, $(A - \mu_i I_n)^k U_i = 0$ et il s'ensuit que :

$$Y(t) = \sum_{i=1}^{n} c_i e^{\mu_i t} \left(\sum_{k=0}^{\gamma_i - 1} \frac{t^k}{k!} (A - \mu_i I_n)^k U_i \right)$$

□

Remarque 3.27 *1. Les formules du théorème donnent un calcul implicite de e^{tA}. En effet, en vertu de la formule de changement de coordonnées, et comme $Y(0) = C = \sum c_i U_i$, on a $\begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} = P^{-1}C$ où P est la matrice de passage de la base canonique de $\mathbb{K}^n$ à $\mathcal{B}$. Autrement dit les colonnes de P sont $U_1, \cdots, U_n$ et*

$$Y(t) = e^{tA} P \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix}$$

ce qui permet de déduire que les colonnes de $e^{tA}P$ sont respectivement :

$$e^{\mu_1 t}\left(\sum_{k=0}^{\gamma_1 - 1} \frac{t^k}{k!}(A - \mu_1 I_n)^k U_1\right), \cdots, e^{\mu_n t}\left(\sum_{k=0}^{\gamma_n - 1} \frac{t^k}{k!}(A - \mu_n I_n)^k U_n\right)$$

2. On peut remarquer d'autre part que si N est la matrice nilpotente de la décomposition de Jordan de A, alors en utilisant la forme de N apparaissant dans la preuve du théorème 2.78 du chapitre 2 on a pour tout $i \in \{1, \cdots, n\}$:

$$e^{tN} U_i = \left(\sum_{k=0}^{\gamma_i - 1} \frac{t^k}{k!}(A - \mu_i I_n)^k U_i\right)$$

3.3.2 Système différentiel linéaire non homogène

Soient I un intervalle de $\mathbb{R}$, $B : I \longrightarrow \mathbb{K}^n$ une fonction continue et $A \in \mathcal{M}_n(\mathbb{K})$.

Théorème 3.28 *On suppose que l'ensemble $S_{\mathcal{L}}$ des solutions de $\mathcal{L} : Y' = AY + B$ est non vide. Soit $\varphi_0 \in S_{\mathcal{L}}$. Alors :*

$$S_{\mathcal{L}} := \{I \longrightarrow \mathbb{K}^n, t \longmapsto e^{tA}C + \varphi_0(t) \quad \text{où} \quad C \in \mathbb{K}^n\}$$

Preuve: Soit Y une solution de $\mathcal{L}$ sur I. Alors :

$$(Y - \varphi_0)'(t) = Y'(t) - \varphi_0'(t) = (AY(t) + B(t)) - (A\varphi_0(t) + B(t)) = A(Y(t) - \varphi_0(t))$$

Donc $Y - \varphi_0$ est solution de l'équation homogène $\mathcal{H} : Y' = AY$. Donc $Y(t) = e^{tA}C + \varphi_0(t)$ où $C \in \mathbb{K}^n$.

Réciproquement, on vérifie facilement que toutes les fonctions de la forme $Y(t) = e^{tA}C + \varphi_0(t)$ pour tout $t \in I$ sont solutions.

□

Théorème 3.29 ***(Méthode de la variation de la constante)*** *L'ensemble $S_{\mathcal{L}}$ est non vide. Plus précisément pour tout $t_0 \in I$:*

$$\begin{array}{rcccl} \varphi_0 & : & I & \longrightarrow & \mathbb{K}^n \\ & & t & \longmapsto & \varphi_0(t) := \displaystyle\int_{t_0}^{t} e^{(t-x)A} B(x)dx \end{array}$$

est une solution de $\mathcal{L}$.

PREUVE: Soit pour tout $t \in I$, la fonction $\varphi_0(t) := \int_{t_0}^{t} e^{(t-x)A}B(x)dx$. Montrons que φ_0 est solution de $\mathcal{L}$.

Remarquons d'abord que puisque tA et $-xA$ commutent alors $e^{(t-x)A} = e^{tA}e^{-xA}$ et $\varphi_0(t) = e^{tA}\int_{t_0}^{t} e^{-xA}B(x)dx$. De plus :

$$\varphi_0'(t) = Ae^{tA}\int_{t_0}^{t} e^{-xA}B(x)dx + e^{tA}e^{-tA}B(t) = A\varphi_0(t) + B(t)$$

D'où le résultat. □

Remarque 3.30 — *Les solutions de $\mathcal{H} : Y' = AY$ s'écrivent $Y(t) = e^{tA}C$ avec $C \in \mathbb{K}^n$. Si l'on cherche une fonction dérivable $C : I \longrightarrow \mathbb{K}^n$ telle que $\varphi_0(t) = e^{tA}C(t)$ soit solution de $\mathcal{L}$ on a :*

$$\begin{aligned} 0 &= \varphi_0'(t) - (A\varphi_0(t) + B(t)) = Ae^{tA}C(t) + e^{tA}C'(t) - (Ae^{tA}C(t) + B(t)) \\ &= e^{tA}C'(t) + B(t). \end{aligned}$$

Donc $C'(t) = e^{-tA}B(t)$ et on retrouve l'expression donnée dans le théorème. Cette façon de trouver une solution de $\mathcal{L}$ explique l'expression de "Méthode de la variation de la constante".

— *Pour résoudre $\mathcal{L}$ il suffit tout simplement de trouver une solution particulière φ_0 de $\mathcal{L}$ et toutes les autres solutions s'écrivent comme la somme de φ_0 et d'une solution de $\mathcal{H}$.*

3.4 Équations différentielles linéaires à coefficients constants

$\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$ et I désigne un intervalle non vide de $\mathbb{R}$.

Définition 3.31 *Soient $n \geqslant 1$, $(a_0, \cdots, a_{n-1}) \in \mathbb{K}^n$ et $b : I \longrightarrow \mathbb{K}$ continue.*

1. *L'équation :*
$$(\mathcal{L}) \qquad y^{(n)} + a_{n-1}y^{(n-1)} + \cdots + a_1y' + a_0y = b$$
*s'appelle **équation différentielle linéaire à coefficients constants**.*
2. *L'équation :*
$$(\mathcal{H}) \qquad y^{(n)} + a_{n-1}y^{(n-1)} + \cdots + a_1y' + a_0y = 0$$
*s'appelle **équation différentielle linéaire homogène associée à** $\mathcal{L}$.*

Résoudre $(\mathcal{L})$ ou $(\mathcal{H})$ revient à trouver un intervalle non vide $J \subset I$ et une fonction y dérivable sur J vérifiant $(\mathcal{L})$ ou $(\mathcal{H})$.

Définition 3.32 *Soient $n \geqslant 1$, $(a_0, \cdots, a_{n-1}) \in \mathbb{K}^n$ et $b : I \longrightarrow \mathbb{K}$ continue. On appelle **équation caractéristique** de l'équation différentielle :*

$$(\mathcal{L}) \qquad y^{(n)} + a_{n-1}y^{(n-1)} + \cdots + a_1y' + a_0y = b$$

l'équation :

$$x^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0 = 0$$

Théorème 3.33 *Soient $n \geqslant 1$, $(a_0, \cdots, a_{n-1}) \in \mathbb{K}^n$. Considérons l'équation homogène :*

$$(\mathcal{H}) \qquad y^{(n)} + a_{n-1}y^{(n-1)} + \cdots + a_1 y' + a_0 y = 0$$

Supposons que les racines de l'équation caractéristique soient $\lambda_1, \cdots, \lambda_r$ de multiplicité respectives $\alpha_1, \cdots, \alpha_r$. Alors l'ensemble des solutions $S_{\mathcal{H}}$ définies sur $\mathbb{R}$ de $(\mathcal{H})$ est un sous-espace vectoriel de $\mathcal{C}(\mathbb{R}, \mathbb{C})$. Plus précisément :

$$S_{\mathcal{H}} := \{y \in \mathcal{C}^\infty(\mathbb{R}, \mathbb{C}) \ ; \ \exists (P_1, \cdots, P_r) \in \mathbb{C}_{\alpha_1 - 1}[X] \times \cdots \times \mathbb{C}_{\alpha_r - 1}[X] \ ; \ y(t) = \sum_{i=1}^r e^{\lambda_i t} P_i(t)\}$$

Preuve: Posons $Y = \begin{pmatrix} y^{(n-1)} \\ \vdots \\ y' \\ y \end{pmatrix}$. Alors :

$$Y' = \begin{pmatrix} -a_{n-1} & -a_{n-2} & \cdots & \cdots & -a_0 \\ 1 & 0 & \cdots & \cdots & 0 \\ 0 & 1 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 1 & 0 \end{pmatrix} Y$$

Notons A la matrice du système différentiel obtenu. Les solutions de ce système sont de la forme :

$$Y(t) = e^{tA} C$$

où $C \in \mathbb{C}^n$. Calculons le polynôme caractéristique de A. Alors :

$$\chi_A(\lambda) = (-1)^n \left(\lambda^n + a_{n-1}\lambda^{n-1} + \cdots + a_1\lambda + a_0\right)$$

et les racines de χ_A sont exactement les racines de l'équation caractéristique. Donc

$$\chi_A(X) = (-1)^n \prod_{i=1}^r (X - \lambda_i)^{\alpha_i}$$

Comme $\mathbb{C}^n = \bigoplus_{i=1}^r C_{\lambda_i}(A)$, il existe $(U_1, \cdots, U_r) \in \prod_{i=1}^r C_{\lambda_i}(A)$ tel que $C = \sum_{i=1}^r U_i$. Donc par un raisonnement similaire à la preuve du théorème 3.26 :

$$Y(t) = \sum_{i=1}^r e^{tA} U_i = \sum_{i=1}^r e^{\lambda_i t I_n + t(A - \lambda_i I_n)} U_i = \sum_{i=1}^r e^{\lambda_i t} e^{t(A - \lambda_i I_n)} U_i = \sum_{i=1}^r e^{\lambda_i t} \sum_{k=0}^{\alpha_i - 1} \frac{t^k}{k!}(A - \lambda_i I_n)^k U_i$$

Il s'ensuit que la dernière coordonnée de $Y(t)$ qui n'est rien d'autre que $y(t)$ vérifie :

$$y(t) = \sum_{i=1}^r e^{\lambda_i t} \sum_{k=0}^{\alpha_i - 1} \alpha_{ki} t^k = \sum_{i=1}^r e^{\lambda_i t} P_i(t)$$

où P_i est un polynôme de degré au plus $\alpha_i - 1$.

Réciproquement, l'espace des solutions $\mathcal{S}_H$ est donc inclus dans l'espace $\mathcal{S}'$ des éléments de la forme $y(t) = \sum_{i=1}^{r} e^{\lambda_i t} P_i(t)$ avec chaque P_i de degré au plus $\alpha_i - 1$. Comme $\mathcal{S}'$ a comme famille génératrice les éléments $e^{\lambda_i t} t^m$ avec $0 \leq m \leq \alpha_i - 1$, qui sont en nombre $\sum \alpha_i = n$, on a $\dim \mathcal{S}' \leq n$.

Par ailleurs le $\mathbb{K}$-espace vectoriel des $y(t)$ vérifiant $(\mathcal{H})$ et le $\mathbb{K}$-espace vectoriel des $Y(t)$ vérifiant $Y' = AY$ sont isomorphes. En effet, à y on peut associer le vecteur $Y = (y^{(n-1)}, \ldots, y', y)$ et à $Y = (y_1, \ldots, y_n)$ on peut associer y_n : ce sont des isomorphismes inverses l'un de l'autre. On en déduit que $\dim \mathcal{S}_H = n$. Comme $\mathcal{S}_H \subset \mathcal{S}'$, on obtient $\mathcal{S} = \mathcal{S}'$.

Deuxième preuve de la réciproque, plus explicite :

Réciproquement il faut vérifier que si $y(t) = \displaystyle\sum_{i=1}^{r} e^{\lambda_i t} P_i(t)$ où P_i est un polynôme de degré au plus $\alpha_i - 1$ alors y est solution de $(\mathcal{H})$. Considérons l'endomorphisme suivant :

$$\begin{array}{rcccc} p & : & E & \longrightarrow & E \\ & & y & \longmapsto & p(y) = y' \end{array}$$

où $E = \mathcal{C}^\infty(\mathbb{R}, \mathbb{C})$. Alors si $y \in S_{\mathcal{H}}$, on a :

$$y^{(n)} + a_{n-1} y^{(n-1)} + \cdots + a_1 y' + a_0 y = (p^n + a_{n-1} p^{(n-1)} + \cdots + a_1 p' + a_0 p)(y) = \prod_{i=0}^{r} (p - \lambda_i \mathrm{Id}_E)^{\alpha_i}(y) = 0$$

On va alors montrer le lemme suivant :

Lemme 3.34 *Soit $\lambda \in \mathbb{C}$ et $\alpha \geqslant 1$ un entier. Alors :*

$$(p - \lambda Id_E\ )^\alpha (t^{\alpha-1} e^{\lambda t}) = 0$$

Preuve: On a :

$$\begin{aligned} (p - \lambda \mathrm{Id}_E\ )(t^{\alpha-1} e^{\lambda t}) &= p(t^{\alpha-1} e^{\lambda t}) - \lambda t^{\alpha-1} e^{\lambda t} \\ &= (\alpha - 1) t^{\alpha-2} e^{\lambda t} + \lambda t^{\alpha-1} e^{\lambda t} - \lambda t^{\alpha-1} e^{\lambda t} \\ &= (\alpha - 1) t^{\alpha-2} e^{\lambda t} \end{aligned}$$

Soit $j \in \{1, \ldots, \alpha - 1\}$. Pour $j = 1$, on vient de montrer que :

$$(p - \lambda \mathrm{Id}_E\ )^j (t^{\alpha-1} e^{\lambda t}) = (\alpha - 1) t^{\alpha-1-j} e^{\lambda t}$$

Pour $j \geqslant 2$, supposons que :

$$(p - \lambda \mathrm{Id}_E\ )^j (t^{\alpha-1} e^{\lambda t}) = (\alpha - 1)(\alpha - 2) \ldots (\alpha - j) t^{\alpha-1-j} e^{\lambda t}$$

et montrons que cette relation reste vraie au rang $j + 1$.

$$\begin{aligned} (p - \lambda \mathrm{Id}_E\ )^{j+1} (t^{\alpha-1} e^{\lambda t}) &= (p - \lambda \mathrm{Id}_E\ )((\alpha - 1)(\alpha - 2) \ldots (\alpha - j) t^{\alpha-1-j} e^{\lambda t}) \\ &= (\alpha - 1)(\alpha - 2) \ldots (\alpha - j)(p(t^{\alpha-1-j} e^{\lambda t}) - \lambda t^{\alpha-1-j} e^{\lambda t}) \\ &= (\alpha - 1)(\alpha - 2) \ldots (\alpha - j)(\alpha - 1 - j) t^{\alpha-1-(j+1)} e^{\lambda t} \end{aligned}$$

La relation est donc vraie au rang $j+1$ et on conclut qu'elle est vraie pour tout $j \in \{1, \ldots, \alpha\}$. En prenant $j = \alpha$, on obtient la relation désirée. $\square$

De ce lemme on déduit immédiatement que pour tout $P_i \in \mathbb{C}_{\alpha_i - 1}[X]$, la fonction $t \longmapsto e^{\lambda_i t} P_i(t) \in S_{\mathcal{H}}$.

$\square$