UNIVERSITÉ PARIS 7
DENIS DIDEROT

MI3

Algèbre et analyse fondamentales I

CHAPITRE IV

RÉDUCTION DES ENDOMORPHISMES

année 2008-2009

Auteur : Thierry Joly

Département de Formation de 1er Cycle de Sciences Exactes

RÉDUCTION DES ENDOMORPHISMES

Plan du chapitre :

RÉDUCTION DES ENDOMORPHISMES

N.B. Dans tout ce chapitre, la lettre $\mathbf{K}$ désigne l'un des ensembles $\mathbb{R}$ ou $\mathbb{C}$.

1 Sommes directes de sous-espaces vectoriels (rappels)

Définition On appelle *somme* de sous-espaces $E_1, \dots, E_n$ d'un $\mathbf{K}$-espace vectoriel E l'ensemble noté $E_1 + \cdots + E_n$ des vecteurs de E de la forme $x_1 + \cdots + x_n$, où $x_1 \in E_1, \dots, x_n \in E_n$:

$$E_1 + \cdots + E_n = \big\{x_1 + \cdots + x_n \ ; \ x_1 \in E_1, \dots, x_n \in E_n\big\}.$$

Proposition 1 *La somme $E_1 + \cdots + E_n$ de sous-espaces quelconques $E_1, \dots, E_n$ d'un $\mathbf{K}$-espace vectoriel E est un sous-espace de E.*

Démonstration Pour tous $x, y \in E_1 + \cdots + E_n$ et tout $k \in \mathbf{K}$, il existe par définition des vecteurs $x_1 \in E_1, \dots, x_n \in E_n$, $y_1 \in E_1, \dots, y_n \in E_n$ tels que : $x = x_1 + \cdots + x_n$ et $y = y_1 + \cdots + y_n$, donc $x + y = (x_1 + y_1) + \cdots + (x_n + y_n) \in E_1 + \cdots + E_n$ et $kx = kx_1 + \cdots + kx_n \in E_1 + \cdots + E_n$.

□

Définition On dit que la somme $E_1 + \cdots + E_n$ de sous-espaces $E_1, \dots, E_n$ d'un $\mathbf{K}$-espace vectoriel E est *directe* lorsque pour tous $x_1 \in E_1, \dots, x_n \in E_n$:

$$x_1 + \cdots + x_n = 0 \quad \Longrightarrow \quad x_1 = \ldots = x_n = 0.$$

Si tel est le cas, la somme $E_1 + \cdots + E_n$ est notée : $E_1 \oplus \cdots \oplus E_n$.

Proposition 2 *Soit $E_1 \oplus \cdots \oplus E_n$ une somme directe de sous-espaces d'un $\mathbf{K}$-espace vectoriel E. Alors tout vecteur $x \in E_1 \oplus \cdots \oplus E_n$ se décompose* de façon unique *en une somme :*

$$x = x_1 + \cdots + x_n, \qquad x_1 \in E_1, \dots, x_n \in E_n.$$

Démonstration Si $x = x_1 + \cdots + x_n = x'_1 + \cdots + x'_n$ avec $x_i, x'_i \in E_i$ pour tout i, alors $(x_1 - x'_1) + \cdots + (x_n - x'_n) = 0$. Il s'ensuit : $x_1 - x'_1 = \ldots = x_n - x'_n = 0$, soit encore : $x_1 = x'_1, \dots, x_n = x'_n$.

□

Proposition 3 *La somme $E_1 + E_2$ de* deux *sous-espaces vectoriels E_1, E_2 d'un $\mathbf{K}$-espace vectoriel E est directe ssi $E_1 \cap E_2 = \{0\}$.*

Démonstration Si la somme $E_1 + E_2$ est directe, alors pour tout $x \in E_1 \cap E_2$, on a : $x \in E_1$, $-x \in E_2$ et la relation $x + (-x) = 0$ entraîne donc $x = -x = 0$, d'où : $E_1 \cap E_2 = \{0\}$. Réciproquement, si $E_1 \cap E_2 = \{0\}$, alors pour tous $x_1 \in E_1$, $x_2 \in E_2$ tels que $x_1 + x_2 = 0$, on a $x_1 = -x_2 \in E_1 \cap E_2$, donc $x_1 = -x_2 = 0$ et la somme $E_1 + E_2$ est directe.

$\square$

Remarques • Dans le cas particulier où chaque sous-espace E_i d'une somme $E_1 + \ldots + E_n$ est engendré par un unique vecteur non nul v_i, alors la somme $E_1 + \ldots + E_n$ est directe ssi les vecteurs $v_1, \ldots, v_n$ sont linéairement indépendants.
• La notion de somme directe de sous-espaces peut donc être vue comme une généralisation de la notion d'indépendance linéaire de vecteurs et la proposition 2 est à rapprocher de l'unicité des coefficients k_i d'une combinaison linéaire $\sum_{i=1}^{n} k_i v_i$ de vecteurs linéairement indépendants $v_1, \ldots, v_n$.
• Tout naturellement, les notions de somme directe et de systèmes linéairement indépendants présentent aussi les mêmes écueils. Par exemple, de même qu'il est tout à fait faux de dire que des vecteurs $v_1, \ldots, v_n$ sont linéairement indépendants ssi ils deux à deux non colinéaires (erreur fréquente), il faut se garder de généraliser abusivement la proposition 3 en prétendant qu'une somme $E_1 + \ldots + E_n$ est directe ssi $E_i \cap E_j = \{0\}$ pour chaque paire de sous-espaces $E_i \neq E_j$.

Théorème 4 *Soit $E_1 \oplus \cdots \oplus E_n$ une somme directe de sous-espaces d'un $\mathbf{K}$-espace vectoriel E. Si $(u_{11}, \ldots, u_{1p_1})$ est une base quelconque de E_1, $(u_{21}, \ldots, u_{2p_2})$ une base quelconque de $E_2, \ldots$ $\ldots$ et $(u_{n1}, \ldots, u_{np_n})$ une base quelconque de E_n, alors la suite de vecteurs obtenue en accolant toutes ces bases :*

$$(u_{11}, \ldots, u_{1p_1}, u_{21}, \ldots, u_{2p_2}, \ldots\ldots, u_{n1}, \ldots, u_{np_n})$$

est une base de $E_1 \oplus \cdots \oplus E_n$.

Démonstration Il s'agit d'établir que tout vecteur x de $E_1 \oplus \cdots \oplus E_n$ s'écrit de façon unique sous la forme :

$$x = k_{11}u_{11} + \cdots + k_{1p_1}u_{1p_1} + k_{21}u_{21} + \cdots + k_{2p_2}u_{2p_2} + \cdots\cdots + k_{n1}u_{n1} + \cdots + k_{np_n}u_{np_n}. \qquad (*)$$

Tout vecteur $x \in E_1 \oplus \cdots \oplus E_n$ s'écrit sous la forme $x = x_1 + \cdots + x_n$, où $x_1 \in E_1, \ldots, x_n \in E_n$. Comme $(u_{i1}, \ldots, u_{ip_i})$ est une base de E_i, chacun des vecteurs x_i s'écrit à son tour sous la forme $x_i = k_{i1}u_{i1} + \cdots + k_{ip_i}u_{ip_i}$. En remplaçant ces expressions dans la somme $x = x_1 + \cdots + x_n$, on obtient la relation $(*)$.

Montrons à présent que les scalaires k_{ij} de $(*)$ sont uniques. Supposons que l'on a aussi :

$$x = k'_{11}u_{11} + \cdots + k'_{1p_1}u_{1p_1} + k'_{21}u_{21} + \cdots + k'_{2p_2}u_{2p_2} + \cdots\cdots + k'_{n1}u_{n1} + \cdots + k'_{np_n}u_{np_n}.$$

Alors pour tout i, $x'_i = k'_{i1}u_{i1} + \cdots + k'_{ip_i}u_{ip_i}$ est un vecteur de E_i et l'on a : $x = x'_1 + \cdots + x'_n$. La proposition 2 entraîne donc $x_1 = x'_1, x_2 = x'_2, \ldots, x_n = x'_n$, d'où pour chaque i :

$$x_i = k_{i1}u_{i1} + \cdots + k_{ip_i}u_{ip_i} = k'_{i1}u_{i1} + \cdots + k'_{ip_i}u_{ip_i}.$$

Comme les coordonnées du vecteur x_i dans la base $(u_{i1}, \ldots, u_{ip_i})$ de E_i sont uniques, il s'ensuit $k_{ij} = k'_{ij}$ pour tous i, j.

$\square$

Corollaire 5 $\dim(E_1 \oplus \cdots \oplus E_n) = \dim E_1 + \cdots + \dim E_n$.

2 Diagonalisation

2.1 Matrices diagonales – endomorphismes diagonalisables

Définition Si $k_1, \dots, k_n$ sont des scalaires, on note $\mathrm{Diag}(k_1, \dots, k_n)$ la matrice carrée $n \times n$:

$$\mathrm{Diag}(k_1, \dots, k_n) = \begin{pmatrix} k_1 & 0 & 0 & \cdots & 0 \\ 0 & k_2 & 0 & & 0 \\ 0 & 0 & k_3 & & 0 \\ \vdots & & & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & k_n \end{pmatrix}$$

Les matrices de la forme $\mathrm{Diag}(k_1, \dots, k_n)$ sont appelées *matrices diagonales.*

Définition On dit qu'un endomorphisme f de d'un $\mathbf{K}$-espace vectoriel E est *diagonalisable* s'il existe une base de E dans laquelle la matrice représentant f est diagonale.

Diagonaliser f signifie : rechercher une telle base. Si la matrice de f dans la base $(u_1, \dots, u_n)$ est $\mathrm{Diag}(k_1, \dots, k_n)$, on a pour tout i : $f(u_i) = k_i u_i$, autrement dit u_i est un vecteur propre associé à la valeur propre k_i. Diagonaliser f revient donc à rechercher une base de E *uniquement constituée de vecteurs propres.*

Exemple Soit f l'endomorphisme de $\mathbb{R}^2$ dont la matrice dans la base canonique (e_1, e_2) de $\mathbb{R}^2$ est :

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

Bien que A ne soit pas diagonale, f est diagonalisable. En effet, les vecteurs $u_1 = e_1 + e_2$ et $u_2 = e_1 + 2e_2$ ne sont à l'évidence pas colinéaires, donc le système (u_1, u_2) est libre et forme une base de $\mathbb{R}^2$. De plus, la matrice A nous donne : $f(e_1) = -2e_2$ et $f(e_2) = e_1 + 3e_2$, donc :

- $f(u_1) = f(e_1) + f(e_2) = -2e_2 + (e_1 + 3e_2) = e_1 + e_2 = u_1$,
- $f(u_2) = f(e_1) + 2f(e_2) = -2e_2 + 2(e_1 + 3e_2) = 2e_1 + 4e_2 = 2u_2$.

Ainsi, la matrice de f dans la base (u_1, u_2) est :

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

Comme les coordonnées dans la base canonique (e_1, e_2) des vecteurs u_1, u_2 sont respectivement $(1, 1)$ et $(1, 2)$, la matrice de passage de la base canonique (e_1, e_2) à cette nouvelle base (u_1, u_2) s'écrit :

$$P = \begin{array}{cc} \begin{array}{cc} \overset{u_1}{\downarrow} & \overset{u_2}{\downarrow} \end{array} & \\ \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} & \begin{array}{l} \leftarrow e_1 \\ \leftarrow e_2 \end{array} \end{array}$$

Rappelons que si X (respectivement X') est le vecteur colonne des coordonnées dans la base (e_1, e_2) (respectivement dans la base (u_1, u_2)) d'un même vecteur de $\mathbb{R}^2$, alors $X = PX'$, $X' = P^{-1}X$, et que ceci entraîne les relations :

$$D = P^{-1}AP, \qquad A = PDP^{-1}.$$

Remarque Par abus de langage, on dit aussi que l'on a "diagonalisé" la matrice A : cela signifie simplement que l'on a trouvé une matrice *inversible* P (la matrice de passage) telle que $D = P^{-1}AP$ soit diagonale.

2.2 Applications de la diagonalisation

Indiquons dès à présent quelques problèmes où la diagonalisation des matrices s'avère précieuse :

- Calcul des puissances d'une matrice. Une vertu des matrices diagonales est qu'elles sont particulièrement faciles à multiplier entre elles ; en effet, on vérifie sans peine la relation :

$$\mathrm{Diag}\big(k_1,\ldots,k_p\big).\mathrm{Diag}\big(k'_1,\ldots,k'_p\big) = \mathrm{Diag}\big(k_1k'_1,\ldots,k_pk'_p\big).$$

Cette dernière entraîne facilement par récurrence sur $n \in \mathbb{N}$:

$$\big(\mathrm{Diag}(k_1,\ldots,k_p)\big)^n = \mathrm{Diag}\big(k_1^n,\ldots,k_p^n\big).$$

Ainsi, alors que l'on ne voit pas bien comment calculer directement A^n pour la matrice A de l'exemple précédent, on peut immédiatement écrire :

$$D^n = \begin{pmatrix} 1 & 0 \\ 0 & 2^n \end{pmatrix}.$$

Or la relation $A = PDP^{-1}$ entraîne :

$$A^n = \underbrace{(PDP^{-1})\cdots(PDP^{-1})}_{n\ \text{facteurs}} = PD^nP^{-1},$$

de sorte que l'on obtient A^n en inversant P puis en calculant le produit PD^nP^{-1} :

$$P^{-1} = \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix}, \qquad A^n = PD^nP^{-1} = \begin{pmatrix} 2-2^n & 2^n-1 \\ 2-2^{n+1} & 2^{n+1}-1 \end{pmatrix}.$$

- Calcul du terme général d'une suite récurrente linéaire. Il est bien connu qu'une suite géométrique $(u_n)_{n\in\mathbb{N}}$ de raison a, i.e. telle que $u_{n+1} = a\,u_n$, a pour terme général : $u_n = a^n u_0$. Le calcul matriciel permet d'exprimer de même le terme général d'une suite définie à partir de ses k premiers termes par une relation de la forme :

$$u_{n+k} = a_1\,u_{n+k-1} + a_2\,u_{n+k-2} + \cdots + a_k\,u_n.$$

Soit, par exemple, la suite $(u_n)_{n\in\mathbb{N}}$ définie par :

$$\left[\begin{array}{l} u_0 = 4, \\ u_1 = 7, \\ u_{n+2} = 3u_{n+1} - 2u_n. \end{array}\right.$$

Quitte à être redondant, la relation de récurrence de cette définition peut aussi s'exprimer par le système :

$$\left[\begin{array}{l} u_{n+1} = \phantom{-2u_n +{}} u_{n+1} \\ u_{n+2} = -2u_n + 3u_{n+1} \end{array}\right. \qquad \text{soit encore} \qquad \begin{pmatrix} u_{n+1} \\ u_{n+2} \end{pmatrix} = A \begin{pmatrix} u_n \\ u_{n+1} \end{pmatrix},$$

où A est toujours la même matrice que précédemment. En posant $U_n = \begin{pmatrix} u_n \\ u_{n+1} \end{pmatrix}$ pour tout $n \in \mathbb{N}$, on a donc :

$$U_0 = \begin{pmatrix} 4 \\ 7 \end{pmatrix}, \qquad U_{n+1} = A\,U_n,$$

d'où, par récurrence sur $n \in \mathbb{N}$:

$$U_n = A^n\,U_0.$$

À l'aide du calcul de A^n plus haut, on obtient $u_n = (2-2^n).4 + (2^n-1).7$, soit encore :

$$u_n = 3.2^n + 1.$$

2.3 Sous-espaces propres d'un endomorphisme

Comme on a déjà remarqué plus haut, diagonaliser un endomorphisme f d'un $\mathbf{K}$-espace vectoriel E consiste à former une base de E à l'aide de vecteurs propres de f. Puisque l'on sait déjà déterminer les valeurs propres de f (il s'agit des racines de son polynôme caractéristique), il nous reste à étudier pour chaque valeur propre λ l'ensemble E_λ des vecteurs propres associés à λ. Cet ensemble E_λ est en fait le noyau de l'application linéaire $f - \lambda\,\mathrm{Id}_E$:

$$f(v) = \lambda v \quad \Longleftrightarrow \quad (f - \lambda\,\mathrm{Id}_E)(v) = f(v) - \lambda v = 0 \quad \Longleftrightarrow \quad v \in \mathrm{Ker}(f - \lambda\,\mathrm{Id}_E).$$

Définition Soit f un endomorphisme d'un $\mathbf{K}$-espace vectoriel E. Pour toute valeur propre λ de f, le sous-espace vectoriel $E_\lambda = \mathrm{Ker}(f - \lambda\,\mathrm{Id}_E) = \{v \in E\,;\, f(v) = \lambda v\}$ est appelé *sous-espace propre* de f associé à la valeur propre λ.

Rappelons que la *multiplicité d'une racine* α d'un polynôme $P(x)$ est le plus grand entier m tel que $(x-\alpha)^m$ divise $P(x)$, *i.e.* tel que $P(x)$ puisse s'écrire sous la forme : $P(x) = (x-\alpha)^m Q(x)$, où $Q(x)$ est un polynôme.

Théorème 6 *Soit f un endomorphisme d'un $\mathbf{K}$-espace vectoriel E de dimension finie, P_f son polynôme caractéristique et $\lambda_1, \ldots, \lambda_p$ les racines de P_f, que l'on suppose deux à deux distinctes et de multiplicités respectives $m_1, \ldots, m_p$. Alors :*

- *La somme des sous-espaces propres $E_{\lambda_i} = \mathrm{Ker}\,(f - \lambda_i \mathrm{Id}_E)$ de f est* directe *:*

$$E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_p} \subseteq E.$$

- *La dimension de chaque sous-espace propre E_{λ_i} vérifie :*

$$\dim E_{\lambda_i} \leqslant m_i.$$

Démonstration Établissons par récurrence sur $n \in \{1, \ldots, p\}$ que la somme $E_{\lambda_1} + \cdots + E_{\lambda_n}$ est directe. Lorsque $n = 1$, cette somme est trivialement directe, puisqu'elle ne comporte qu'un seul terme. Supposons le résultat établi au rang $n-1$ et établissons-le au rang n. Soit donc $v_1 \in E_{\lambda_1}$, $v_2 \in E_{\lambda_2}, \ldots, v_n \in E_{\lambda_n}$ tels que :

$$v_1 + \cdots + v_{n-1} + v_n = 0.$$

En appliquant f à cette somme, on obtient en vertu de la linéarité de f et des relations $f(v_i) = \lambda_i v_i$:

$$\lambda_1 v_1 + \cdots + \lambda_{n-1} v_{n-1} + \lambda_n v_n = 0,$$

et en multipliant cette même somme par λ_n :

$$\lambda_n v_1 + \cdots + \lambda_n v_{n-1} + \lambda_n v_n = 0.$$

Retranchons ces deux dernières égalités :

$$(\lambda_1 - \lambda_n) v_1 + \cdots + (\lambda_{n-1} - \lambda_n) v_{n-1} = 0.$$

Comme $(\lambda_i - \lambda_n) v_i \in E_{\lambda_i}$ pour tout $i \in \{1, \ldots, n-1\}$, l'hypothèse de récurrence entraîne alors : $(\lambda_i - \lambda_n) v_i = 0$ $(i = 1, \ldots, n-1)$, or $\lambda_i - \lambda_n \neq 0$ (car $\lambda_1, \ldots, \lambda_{n-1}$ sont deux à deux distincts), donc $v_i = 0$ pour tout $i \in \{1, \ldots, n-1\}$. Il s'ensuit évidemment $v_n = 0$; ainsi tous les vecteurs v_i sont nuls, ce qui établit que la somme des sous-espaces E_{λ_i} est directe.

Fixons maintenant un sous-espace propre E_{λ_i} et montrons que l'on a : $d = \dim E_{\lambda_i} \leqslant m_i$. Pour ce faire, considérons une base quelconque $(u_1, \ldots, u_k)$ de E_{λ_i}, que l'on complète en une base $(u_1, \ldots, u_n)$ de E. Comme $f(u_i) = \lambda_i u_i$ pour tout $i \leqslant d$, la matrice A de f dans la base $(u_1, \ldots, u_n)$ est de la forme :

$$A = \left(\begin{array}{c|c} \underbrace{\begin{matrix} \lambda_i & & 0 \\ & \ddots & \\ & & \lambda_i \end{matrix}}_{d} & B \\ \hline 0 & C \end{array}\right)$$

$P_f(x)$ est donc le déterminant de la matrice :

$$A - xI_n = \left(\begin{array}{c|c} \underbrace{\begin{matrix} \lambda_i - x & & 0 \\ & \ddots & \\ & & \lambda_i - x \end{matrix}}_{d} & B \\ \hline 0 & C - xI_{n\text{-}d} \end{array}\right)$$

En itérant d développements selon la première colonne de ce déterminant, on obtient donc :

$$P_f(x) = (\lambda_i - x)^d \det(C - xI_{n-d}).$$

De plus, $\det(C - xI_{n-d})$ est bien un polynôme, puisque qu'il s'agit du polynôme caractéristique de l'endomorphisme de $\mathbf{K}^{n-d}$ représenté par la matrice C. Ainsi, la multiplicité de la racine λ_i de P_f est au moins égale à d, autrement dit : $\dim E_{\lambda_i} = d \leqslant m_i$.

$\square$

2.4 Critères de diagonalisation

Définition On dit qu'un polynôme $P(x)$ est *scindé dans* $\mathbf{K}$ s'il est décomposable en un produit de facteurs du premier degré à coefficients dans $\mathbf{K}$, *i.e.* s'il peut s'écrire sous la forme :

$$P(x) \;=\; a \prod_{i=1}^{n} (x - \alpha_i), \qquad a, \alpha_1, \ldots, \alpha_n \in \mathbf{K}.$$

Remarque Si le polynôme caractéristique $P_f(x)$ d'un endomorphisme f est scindé dans $\mathbf{K}$, alors on peut l'écrire sous la forme : $P_f(x) = a \prod_{i=1}^{p} (\lambda_i - x)^{m_i}$, où $\lambda_1, \ldots, \lambda_p \in \mathbf{K}$ sont ses racines deux à deux distinctes dans $\mathbf{K}$. De plus, a est alors le coefficient de plus haut degré du polynôme $P_f(-x) = a \prod_{i=1}^{p} (x + \lambda_i)^{m_i}$ et vaut donc 1 en vertu de la proposition 12 du premier chapitre, d'où la forme suivante de $P_f(x)$:

$$P_f(x) \;=\; \prod_{i=1}^{n} (\lambda_i - x)^{m_i}.$$

Théorème 7 *Soit f un endomorphisme d'un* $\mathbf{K}$*-espace vectoriel E de dimension finie, P_f son polynôme caractéristique, $\lambda_1, \ldots, \lambda_p$ une liste sans répétition de toutes ses valeurs propres et $E_{\lambda_i} = \operatorname{Ker}(f - \lambda_i \mathrm{Id}_E)$ $(1 \leqslant i \leqslant p)$ ses sous-espaces propres associés. Les énoncés suivants sont alors équivalents :*

1. *f est diagonalisable*
2. *P_f est scindé, mettons : $P_f(x) = \prod_{i=1}^{p}(\lambda_i - x)^{m_i}$, et la multiplicité de* chaque *racine λ_i de P_f est égale à la dimension du sous-espace propre associé à λ_i :* $\dim E_{\lambda_i} = m_i$, $1 \leqslant i \leqslant p$.
3. $\dim E = \dim E_{\lambda_1} + \cdots + \dim E_{\lambda_p}$
4. *E* est *la somme (directe) des sous-espaces propres de f : $E = E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_p}$.*

Démonstration *1 ⇒ 2.* Par hypothèse, E possède une base $(u_1, \ldots, u_n)$ constituée de vecteurs propres de f. Quitte à réordonner les vecteurs de cette base, on peut supposer que la matrice de f dans la base $(u_1, \ldots, u_n)$ est, pour des scalaires $\lambda_1, \ldots, \lambda_p$ deux à deux distincts, de la forme : $D = \mathrm{Diag}(\underbrace{\lambda_1, \ldots, \lambda_1}_{m_1}, \ldots, \underbrace{\lambda_p, \ldots, \lambda_p}_{m_p})$. En itérant des développements selon la première colonne, on obtient :

$$P_f(x) = \det \mathrm{Diag}(\underbrace{\lambda_1 - x, \ldots, \lambda_1 - x}_{m_1}, \ldots, \underbrace{\lambda_p - x, \ldots, \lambda_p - x}_{m_p}) = \prod_{i=1}^{p}(\lambda_i - x)^{m_i}.$$

Ainsi, P_f est scindé et m_i est bien la multiplicité de la racine λ_i de P_f pour tout $i \in \{1, \ldots, p\}$. De plus, en vertu de la forme de la matrice D, la base $(u_1, \ldots, u_n)$ contient clairement m_i vecteurs (linéairement indépendants) de E_{λ_i}, d'où : $m_i \leqslant \dim E_{\lambda_i}$. On en déduit à l'aide du théorème 7 : $\dim E_{\lambda_i} = m_i$ $(1 \leqslant i \leqslant p)$.
2 ⇒ 3. On a par hypothèse : $\dim E_{\lambda_1} + \cdots + \dim E_{\lambda_p} = m_1 + \cdots + m_p = \deg P_f = \dim E$.
3 ⇒ 4. Selon le corollaire 5, on a alors : $\dim(E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_p}) = \dim E_{\lambda_1} + \cdots + \dim E_{\lambda_p} = \dim E$, autrement dit $E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_p}$ est un sous-espace vectoriel de E de même dimension que E, d'où l'égalité : $E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_p} = E$.
4 ⇒ 1. Pour tout $i \in \{1, \ldots, p\}$, soit $(u_{i1}, \ldots, u_{in_i})$ une base quelconque de E_{λ_i}. Par définition, les vecteurs u_{ij} sont des vecteurs propres de f. De plus, $(u_{11}, \ldots, u_{1n_1}, \ldots\ldots, u_{p1}, \ldots, u_{pn_p})$ constitue une base de $E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_p}$ d'après le théorème 4. Ainsi, l'hypothèse $E = E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_p}$ entraîne que ces vecteurs propres de f forment une base de E et f est bien diagonalisable.

□

Remarque Lorsque P_f est scindé et ne possède que des racines simples (*i.e.* de multiplicité 1), alors f est nécessairement diagonalisable en vertu du critère *2* ci-dessus.

2.5 Méthode de diagonalisation – Exemples

Afin de diagonaliser un endomorphisme f, on peut procéder comme suit :

1. Calcul et scindage de P_f : $P_f(x) = \prod_{i=1}^{p}(\lambda_i - x)^{m_i}$. Si P_f n'est pas scindé, alors f n'est pas diagonalisable.
2. Pour chaque racine λ_i de P_f, détermination d'une base $(u_{i1}, \ldots, u_{in_i})$ du sous-espace propre : $E_{\lambda_i} = \operatorname{Ker}(f - \lambda_i \mathrm{Id}_E)$.
   - Si l'une de ces bases vérifie : $n_i = \dim E_{\lambda_i} < m_i$, alors f n'est pas diagonalisable.

- Sinon, on a $n_i = \dim E_{\lambda_i} = m_i$ pour tout i et l'on obtient une base de E en les juxtaposant. La matrice de passage à cette nouvelle base et la matrice diagonale représentant f dans cette dernière s'en déduisent immédiatement :

$$P = \left(\begin{array}{c|c|c|c|c|c|c|c|c|c} u_{11} & \cdots & u_{1n_1} & u_{21} & \cdots & u_{2n_2} & \cdots & u_{p1} & \cdots & u_{pn_p} \end{array} \right)$$

$$D = \begin{pmatrix} \underbrace{\lambda_1 \cdots \lambda_1}_{n_1} & & & 0 \\ & \underbrace{\lambda_2 \cdots \lambda_2}_{n_2} & & \\ & & \ddots & \\ 0 & & & \underbrace{\lambda_p \cdots \lambda_p}_{n_p} \end{pmatrix}$$

Exemple 1. Soit f l'endomorphisme de $\mathbb{R}^2$ représenté dans la base canonique de $\mathbb{R}^2$ par la matrice :

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

Le polynôme $P_f(x) = \begin{vmatrix} -x & 1 \\ -1 & -x \end{vmatrix} = x^2 + 1$ n'est pas scindé dans $\mathbb{R}$, donc f n'est pas diagonalisable. Considérons maintenant l'endomorphisme g de $\mathbb{C}^2$ représenté par la matrice A : on a cette fois $P_g(x) = x^2 + 1 = (x - i)(x + i)$ et g a deux valeurs propres simples : i et $-i$. Selon le théorème 6, les deux sous-espaces propres correspondants E_i, E_{-i} sont donc de dimension 1 et g est à coup sûr diagonalisable, puisque : $\dim E_i + \dim E_{-i} = 2 = \dim \mathbb{C}^2$. Déterminons une base de E_i. Les vecteurs de $E_i = \mathrm{Ker}(g - i\,\mathrm{Id}_{\mathbb{C}^2})$ sont les vecteurs $(z_1, z_2) \in \mathbb{C}^2$ tels que :

$$(A - iI_2)\begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = \begin{pmatrix} -i & 1 \\ -1 & -i \end{pmatrix}\begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \qquad \text{i.e.} \qquad \left[\begin{array}{l} -iz_1 + z_2 = 0 \\ -z_1 - iz_2 = 0 \end{array} \right.$$

autrement dit tels que $z_2 = iz_1$, puisque les deux équations du système équivalent à cette dernière. On peut donc choisir comme base de E_i le vecteur $u_1 = (1, i)$. On trouve de même que E_{-i} est l'ensemble des vecteurs $(z_1, z_2) \in \mathbb{C}^2$ tels que $z_2 = -iz_1$, et l'on peut donc choisir comme base de E_{-i} le vecteur $u_2 = (1, -i)$. La matrice de passage à la base (u_1, u_2) et la matrice de g dans cette dernière sont respectivement :

$$P = \begin{pmatrix} 1 & 1 \\ i & -i \end{pmatrix}, \qquad D = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}.$$

En conclusion, la matrice A est diagonalisable dans $\mathbb{C}$, mais pas dans $\mathbb{R}$.

Exemple 2. Pour tout $a \in \mathbb{R}$, soit f_a l'endomorphisme de $\mathbb{R}^3$ dont la matrice dans la base canonique de $\mathbb{R}^3$ est :

$$M_a = \begin{pmatrix} 4 & 0 & -2 \\ 2 & 5 & 4 \\ 4 & 2 & a \end{pmatrix}.$$

Déterminons pour quelles valeurs de a l'endomorphisme f_a est diagonalisable et diagonalisons f_a pour ces valeurs. Pour ce faire, on commence par calculer le polynôme caractéristique de f_a :

$$\begin{aligned} P_{f_a}(x) = \det(M_a - xI_3) &= \begin{vmatrix} 4-x & 0 & -2 \\ 2 & 5-x & 4 \\ 4 & 2 & a-x \end{vmatrix} = (4-x)\begin{vmatrix} 5-x & 4 \\ 2 & a-x \end{vmatrix} - 2\begin{vmatrix} 2 & 5-x \\ 4 & 2 \end{vmatrix} \\ &= (4-x)\big((5-x)(a-x) - 8\big) - 2\big(4 - 4(5-x)\big) = (4-x)\big((5-x)(a-x) - 8\big) + 8(4-x) \\ &= (4-x)(5-x)(a-x). \end{aligned}$$

- Si $a = 4$, P_{f_a} a comme racines la racine double 4 et la racine simple 5. En vertu du théorème 6, on en déduit que les deux sous-espaces propres E_4, E_5 de f_4 vérifient : $\dim E_4 = 1$ ou 2, $\dim E_5 = 1$. Pour savoir si f_4 est diagonalisable, il faut donc déterminer la dimension de E_4. Clairement, la matrice :

  $$M_4 - 4I_3 = \begin{pmatrix} 0 & 0 & -2 \\ 2 & 1 & 4 \\ 4 & 2 & 0 \end{pmatrix}$$

  a pour rang 2 (ses deux premières colonnes sont proportionnelles entre elles, mais pas à la troisième). On a donc : $\dim E_4 = \dim \operatorname{Ker}(f_4 - 4\,\mathrm{Id}_{\mathbb{R}^3}) = 3 - \operatorname{rang}(f_4 - 4\,\mathrm{Id}_{\mathbb{R}^3}) = 1$. Ainsi, $\dim E_4 + \dim E_5 = 1 + 1 \neq \dim \mathbb{R}^3$, donc f_4 n'est pas diagonalisable.

- Si $a = 5$, P_{f_a} a comme racines la racine simple 4 et la racine double 5. Les deux sous-espaces propres E_4, E_5 de f_5 vérifient donc : $\dim E_4 = 1$ et $\dim E_5 = 1$ ou 2. La matrice :

  $$M_5 - 5I_3 = \begin{pmatrix} -1 & 0 & -2 \\ 2 & 0 & 4 \\ 4 & 2 & 0 \end{pmatrix}$$

  a pour rang 2 (ses deux premières lignes sont proportionnelles, mais pas à la troisième). Le rang de l'application $f_5 - 5\,\mathrm{Id}_{\mathbb{R}^3}$ est donc 2, d'où : $\dim E_5 = 3 - \operatorname{rang}(f_5 - 5\,\mathrm{Id}_{\mathbb{R}^3}) = 1$. Ainsi, on a : $\dim E_4 + \dim E_5 = 2 \neq \dim \mathbb{R}^3$, donc f_5 n'est pas diagonalisable.

- Si $a \neq 4$ et $a \neq 5$, alors P_{f_a} a trois racines simples distinctes : $4, 5$ et a. Les trois sous-espaces propres correspondants de f_a vérifient donc : $\dim E_4 = \dim E_5 = \dim E_a = 1$. On a alors $\dim E_4 + \dim E_5 + \dim E_a = \dim \mathbb{R}^3$ et le théorème 7 entraîne que f_a est diagonalisable.

  Diagonalisons f_a dans ce cas. Toujours selon le théorème 7, on a alors $\mathbb{R}^3 = E_4 \oplus E_5 \oplus E_a$, de sorte qu'il suffit de trouver des vecteurs non nuls u_1, u_2, u_3 dans E_4, E_5, E_a respectivement pour constituer une base de diagonalisation pour f (en effet, chacun de ces vecteurs constituera automatiquement une base du sous-espace correspondant et (u_1, u_2, u_3) sera donc bien une base de E, en vertu du théorème 4). E_4 est l'ensemble des vecteurs $(x, y, z) \in \mathbb{R}^3$ tels que :

  $$(M_a - 4I_3)\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 & 0 & -2 \\ 2 & 1 & 4 \\ 4 & 2 & a-4 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

  On peut donc prendre $u_1 = (1, -2, 0)$.

E_5 est l'ensemble des vecteurs $(x, y, z) \in \mathbb{R}^3$ tels que :

$$(M_a - 5I_3)\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -1 & 0 & -2 \\ 2 & 0 & 4 \\ 4 & 2 & a-5 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},$$

soit encore :

$$\left[\begin{array}{l} x + 2z = 0 \\ 4x + 2y + (a-5)z = 0. \end{array}\right.$$

On peut donc prendre $x = 2$, $z = -1$ et $2y = a - 5 - 4x = a - 13$, i.e. $u_2 = \left(2, \dfrac{a-13}{2}, -1\right)$.

E_a est l'ensemble des vecteurs $(x, y, z) \in \mathbb{R}^3$ tels que :

$$(M_a - aI_3)\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 4-a & 0 & -2 \\ 2 & 5-a & 4 \\ 4 & 2 & 0 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

soit encore :

$$\left[\begin{array}{l} (4-a)x - 2z = 0 \\ 2x + (5-a)y + 4z = 0 \\ 4x + 2y = 0. \end{array}\right.$$

En prenant $x = -1$, $y = 2$ pour vérifier la troisième équation, on vérifie facilement que $z = \dfrac{a-4}{2}$ satisfait les deux premières, de sorte que l'on peut choisir $u_3 = \left(-1, 2, \dfrac{a-4}{2}\right)$.
Finalement, la matrice de passage à la base (u_1, u_2, u_3) et la matrice de f_a dans cette dernière sont respectivement :

$$P = \begin{pmatrix} 1 & 2 & -1 \\ -2 & \frac{a-13}{2} & 2 \\ 0 & -1 & \frac{a-4}{2} \end{pmatrix}, \qquad D = \begin{pmatrix} 4 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & a \end{pmatrix}.$$

Remarque On pourra vérifier que les vecteurs u_1, u_2, u_3 trouvés plus haut sont des vecteurs propres de f_a, même lorsque $a = 4$ ou $= 5$. Cependant, la diagonalisation ci-dessus cesse d'être valide dans ces deux cas, car (u_1, u_2, u_3) cesse d'être une base : on a $u_1 = -u_3$ si $a = 4$ et $u_2 = -2u_3$ si $a = 5$.

3 Trigonalisation

3.1 Matrices triangulaires – endomorphismes trigonalisables

Définition On dit qu'une matrice carrée $A = (a_{ij})_{\substack{1 \leqslant i \leqslant n \\ 1 \leqslant j \leqslant n}}$ est *triangulaire supérieure* si l'on a $a_{ij} = 0$ pour tous (i, j) tels que $i > j$:

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ 0 & a_{22} & a_{23} & & a_{2n} \\ 0 & 0 & a_{33} & & a_{3n} \\ \vdots & & & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & a_{nn} \end{pmatrix}.$$

Définition Un endomorphisme f d'un espace vectoriel E est dit *trigonalisable* s'il existe une base de E dans laquelle f est représenté par une matrice triangulaire supérieure.

Trigonaliser f signifie : rechercher une telle base. Si f a dans la base $(u_1, \ldots, u_n)$ une matrice triangulaire supérieure, mettons la matrice A comme plus haut, alors pour tout j : $f(u_j) = \sum_{i=1}^{j} a_{ij} u_i$.

Trigonaliser $f : E \to E$ revient donc à chercher une base $(u_1, u_2, \ldots, u_n)$ de E telle que pour tout $j \in \{1, \ldots, n\}$, $f(u_j)$ appartient au sous-espace engendré par les vecteurs $u_1, u_2, \ldots, u_j$:

$$f(u_j) \in \mathrm{Vect}(u_1, u_2, \ldots, u_j).$$

(En particulier, u_1 est nécessairement un vecteur propre de f.)

3.2 Critère de trigonalisation

Théorème 8 *Soit E un $\mathbf{K}$-espace vectoriel de dimension finie, $f : E \to E$ un endomorphisme et $P_f(x)$ son polynôme caractéristique. Alors :*

$$f \text{ trigonalisable} \iff P_f \text{ scindé.}$$

En particulier, lorsque $\mathbf{K} = \mathbb{C}$, f est toujours trigonalisable.

Démonstration Si f est représenté par une matrice triangulaire supérieure $T = (a_{ij})_{\substack{1 \leqslant i \leqslant n \\ 1 \leqslant j \leqslant n}}$, on a en itérant des développements de déterminants selon leur première colonne :

$$P_f(x) = \det(T - xI_n) = \begin{vmatrix} a_{11}-x & a_{12} & a_{13} & \cdots & a_{1n} \\ 0 & a_{22}-x & a_{23} & & a_{2n} \\ 0 & 0 & a_{33}-x & & a_{3n} \\ \vdots & & & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & a_{nn}-x \end{vmatrix} = (a_{11}-x) \begin{vmatrix} a_{22}-x & a_{23} & & a_{2n} \\ 0 & a_{33}-x & & a_{3n} \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn}-x \end{vmatrix}$$

$$= (a_{11} - x)(a_{22} - x) \begin{vmatrix} a_{33}-x & a_{34} & & a_{3n} \\ 0 & a_{44}-x & & a_{4n} \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn}-x \end{vmatrix} = \cdots\cdots = \prod_{i=1}^{n} (a_{ii} - x).$$

Ainsi, le polynôme caractéristique de f est scindé.

Réciproquement, si P_f est scindé, alors la remarque suivant la définition de polynôme scindé (début de la section 2.4) entraîne que P_f est de la forme : $P_f(x) = \prod_{i=1}^{n}(\lambda_i - x)$, où les scalaires λ_i ne sont pas nécessairement distincts. Nous allons montrer par récurrence sur n que si :

$$P_f(x) = \prod_{i=1}^{n} (\lambda_i - x),$$

alors il existe une base dans laquelle la matrice de f est de la forme :

$$\begin{pmatrix} \lambda_1 & a_{12} & a_{13} & \cdots & a_{1n} \\ 0 & \lambda_2 & a_{23} & & a_{2n} \\ 0 & 0 & \lambda_3 & & a_{3n} \\ \vdots & & & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_n \end{pmatrix}.$$

Si $n = 1$, alors la matrice de f dans toute base est la matrice $1{\times}1$: $\begin{pmatrix} \lambda_1 \end{pmatrix}$ et il n'y a rien à prouver. Supposons donc ce fait établi au rang $n - 1$ et montrons-le au rang n. Comme λ_1 est valeur propre de f, il existe un vecteur non nul $u_1 \in E$ tel que $f(u_1) = \lambda_1 u_1$. On peut alors trouver des vecteurs $u_2, \ldots, u_n$ tels que $\mathcal{B} = (u_1, u_2, \ldots, u_n)$ soit une base de E. Soit $F = \mathrm{Vect}(u_2, \ldots, u_p)$ le sous-espace engendré par $u_2, \ldots, u_n$; on a donc : $E = \mathrm{Vect}(u_1) \oplus F$. Soit $p : E \to F$ la projecteur sur F parallèlement à u_1, autrement dit l'application linéaire qui à tout vecteur de coordonnées $(x_1, x_2, \dots, x_n)$ dans la base $\mathcal{B}$ de E associe le vecteur de coordonnées $(x_2, \dots, x_n)$ dans la base $(u_2, \dots, u_n)$ de F. Soit enfin $g : F \to F$ l'endomorphisme défini par $g(v) = p(f(v))$ pour tout $v \in F$ et C sa matrice dans la base $(u_2, \dots, u_n)$ de F. La matrice A de f dans la base $\mathcal{B}$ est donc nécessairement de la forme :

$$A = \left(\begin{array}{c|c} \lambda_1 & b_2 \cdots b_n \\ \hline 0 & \\ \vdots & C \\ 0 & \end{array}\right).$$

À l'aide d'un développement selon la première colonne, il s'ensuit :

$$P_f(x) = \det(A - xI_n) = \left|\begin{array}{c|c} \lambda_1 - x & b_2 \cdots b_n \\ \hline 0 & \\ \vdots & C - xI_{n-1} \\ 0 & \end{array}\right| = (\lambda_1 - x)\det(C - xI_{n-1}).$$

Ainsi, le polynôme caractéristique de g est :

$$P_g(x) = \det(C - xI_{n-1}) = \prod_{i=2}^{n} (\lambda_i - x).$$

De par l'hypothèse de récurrence, il existe donc une base $(v_2, \dots, v_n)$ de F dans laquelle la matrice de g est de la forme :

$$T = \begin{pmatrix} \lambda_2 & a_{22} & \cdots & a_{2n} \\ 0 & \lambda_3 & & a_{3n} \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix};$$

Comme $E = \mathrm{Vect}(u_1) \oplus F$, le théorème 4 entraîne que $\mathcal{B}' = (u_1, v_2, \dots, v_n)$ est une base de E. De plus, le projecteur p n'est autre que l'application qui à tout vecteur de coordonnées $(x_1, x_2, \dots, x_n)$ dans la base $\mathcal{B}'$ de E associe le vecteur de coordonnées $(x_2, \dots, x_n)$ dans la base $(v_2, \dots, v_n)$ de F. Or selon la forme de la matrice T, les coordonnées dans cette dernière base du vecteur $g(v_j) = p(f(v_j))$ sont $(a_{2j}, \dots, a_{j-1\,j}, \lambda_j, 0, \dots, 0)$ pour tout $j \in \{2, \dots, n\}$, donc les coordonnées de $f(v_j)$ $(j = 2, \dots, n)$ dans $\mathcal{B}'$ sont de la forme : $(b'_j, a_{2j}, \dots, a_{j-1\,j}, \lambda_j, 0, \dots, 0)$. Ainsi, la matrice de f dans $\mathcal{B}'$ est :

$$\left(\begin{array}{c|c} \lambda_1 & b'_2 \cdots b'_n \\ \hline 0 & \\ \vdots & T \\ 0 & \end{array}\right),$$

i.e. une matrice de la forme désirée.

□

Remarque Il ressort de cette démonstration que *les coefficients diagonaux d'une matrice triangulaire représentant un endomorphisme f sont toujours les valeurs propres de f* (chacune étant répétée autant de fois que sa multiplicité dans le polynôme P_f). Par ailleurs, on y a montré que si $P_f(x) = \prod_{i=1}^{n} (\lambda_i - x)$ (les scalaires λ_i n'étant pas nécessairement distincts), alors f possède dans une certaine base une matrice de la forme :

$$\begin{pmatrix} \lambda_1 & a_{12} & a_{13} & \cdots & a_{1n} \\ 0 & \lambda_2 & a_{23} & & a_{2n} \\ 0 & 0 & \lambda_3 & & a_{3n} \\ \vdots & & & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_n \end{pmatrix}.$$

Puisque l'ordre des scalaires λ_i dans la suite $\lambda_1, \ldots, \lambda_n$ n'a absolument aucune influence sur l'expression $\prod_{i=1}^n (\lambda_i - x)$, il s'ensuit que dans la forme d'une matrice triangulaire supérieure représentant f, *on peut choisir arbitrairement l'ordre des valeurs propres sur la diagonale* (à condition, bien sûr, de respecter leur multiplicité). Il est particulièrement utile de garder à l'esprit ces deux faits lorsque l'on cherche à trigonaliser un endomorphisme (cf section suivante).

Concluons cette section par d'autres faits utiles concernant les valeurs propres d'un endormorphisme f. Rappelons que la *trace* $\operatorname{tr} f$ de f est la somme des coefficients diagonaux de n'importe quelle matrice représentant f (cf Chapitre III, p.24).

Proposition 9 *Soit f un endomorphisme d'un espace de dimension n dont le polynôme caractéristique P_f est scindé, autrement dit tel que P_f possède n racines $\lambda_1, \ldots, \lambda_n$ (non nécessairement distinctes). On a alors :*

$$\operatorname{tr} f = \sum_{i=1}^n \lambda_i, \qquad \det f = \prod_{i=1}^n \lambda_i.$$

Démonstration En effet, f est alors représenté dans une certaine base par une matrice triangulaire supérieure T dont les coefficients diagonaux sont précisément les n racines $\lambda_1, \ldots, \lambda_n$ de $P_f(x) = \prod_{i=1}^n (\lambda_i - x)$ (comptées autant de fois que leur multiplicité). On en déduit immédiatement :

$$\operatorname{tr} f = \sum_{i=1}^n \lambda_i \qquad \text{et :} \qquad \det f = \det(T - 0.I_n) = P_f(0) = \prod_{i=1}^n \lambda_i.$$

□

Remarque Cette dernière proposition peut être mise à profit pour la détermination de valeurs propres d'un endomorphisme :

- Si un endomorphisme f est représenté par une matrice $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, alors ses deux valeurs propres λ_1, λ_2 ont pour somme S et pour produit P :

  $$S = \operatorname{tr} f = a + d, \qquad P = \det f = ad - bc.$$

  λ_1, λ_2 sont alors les racines du polynôme $x^2 - Sx + P$ (qui est de fait le polynôme caractéristique de f).

- Aux dimensions supérieures à 2, la trace et le déterminant ne suffisent plus à déterminer les valeurs propres d'un endomorphisme f. Toutefois, la trace de f (qui est toujours rapidement calculée à partir d'une matrice représentant f) permet de vérifier la somme des valeurs propres trouvées et fournit donc un moyen simple de détecter d'éventuelles erreurs de calculs, à l'image de la célèbre "preuve par 9".

3.3 Méthode de trigonalisation – Exemple

Afin de trigonaliser un endomorphisme $f : E \to E$, on peut commencer par calculer et factoriser son polynôme caractéristique P_f.

- Si P_f n'est pas scindé dans $\mathbf{K}$, alors f n'est pas trigonalisable.
- Sinon, P_f est de la forme : $P_f(x) = \prod_{i=1}^n (\lambda_i - x)$ (où $\lambda_1, \ldots, \lambda_n$ ne sont pas nécessairement distincts) et il s'agit de trouver une base $(u_1, \ldots, u_n)$ dans laquelle la matrice de f est triangulaire supérieure.

On a alors tout intérêt à placer dans cette base $(u_1, \ldots, u_n)$ le plus grand nombre possible de vecteurs propres de f, en déterminant une base $(u_{i1}, \ldots, u_{ip_i})$ de chaque sous-espace propre E_i $(1 \leqslant i \leqslant s)$ de f. En effet, une fois connues les valeurs propres de f, l'obtention de telles bases est relativement rapide ; de plus, nous n'avons pas à nous soucier de ce que la réunion de ces bases est bien un début de base de E possible, autrement dit un système libre, puisque c'est automatiquement le cas par le théorème 4, du fait que la somme $E_1 + \ldots + E_s$ est directe (selon le théorème 6). Nous pouvons donc choisir :

$$(u_1, \ldots, u_p) = (u_{11}, \ldots, u_{1p_1}, u_{21}, \ldots, u_{2p_2}, \ldots\ldots, u_{s1}, \ldots, u_{sp_s}), \quad p = \dim E_1 + \cdots + \dim E_s.$$

Quitte à réordonner la suite de scalaires $\lambda_1, \ldots, \lambda_n$, nous pouvons supposer que la valeur propre associée à chaque vecteur propre u_i $(1 \leqslant i \leqslant p)$ est λ_i. Ce choix des p premiers vecteurs de base impose que la matrice triangulaire supérieure représentant f sera de la forme :

$$T = \begin{pmatrix} \lambda_1 & 0 & 0 & \cdots & 0 & a_{1\,p+1} & a_{1\,p+2} & \cdots\cdots & a_{1n} \\ & \lambda_2 & 0 & & 0 & a_{2\,p+1} & a_{2\,p+2} & & a_{2n} \\ & & \lambda_3 & & \vdots & \vdots & \vdots & & \vdots \\ & & & \ddots & 0 & \vdots & \vdots & & \vdots \\ & & & & \lambda_p & a_{p\,p+1} & \vdots & & \vdots \\ & & & & & \lambda_{p+1} & a_{p+1\,p+2} & & \vdots \\ & & & & & & \lambda_{p+2} & & \vdots \\ & & 0 & & & & & \ddots & a_{n-1\,n} \\ & & & & & & & & \lambda_n \end{pmatrix}.$$

En revanche, la remarque faite à la suite de la démonstration du théorème 8 nous permet de ranger dans n'importe quel ordre les valeurs propres restantes $\lambda_{p+1}, \lambda_{p+2}, \ldots, \lambda_n$ (en tenant compte de leur multiplicité). Si $p = n$, nous avons déjà fini en obtenant la plus belle trigonalisation possible : une diagonalisation. Sinon, il reste à choisir l'un après l'autre les vecteurs de base $u_{p+1}, \ldots, u_n$. On peut s'y prendre de la façon suivante : mettons que l'on a déjà déterminé $u_1, \ldots, u_j$ $(p \leqslant j < n)$. Afin de choisir u_{j+1} de sorte que $(u_1, \ldots, u_j, u_{j+1})$ soit encore un système libre, on commence par compléter arbitrairement le système $(u_1, \ldots, u_j)$ en une base $(u_1, \ldots, u_j, v_{j+1}, \ldots, v_n)$ de E, puis on cherche u_{j+1} sous la forme d'une combinaison linéaire des vecteurs $v_{j+1}, \ldots, v_n$:

$$u_{j+1} = \sum_{i=j+1}^{n} x_i v_i \qquad (x_{j+1}, \ldots, x_n \in \mathbf{K}).$$

La forme de matrice T impose alors :

$$f(u_{j+1}) \;=\; \sum_{i=1}^{j} a_{ij} u_i \;+\; \lambda_{j+1} u_{j+1}.$$

En explicitant cette dernière relation, on obtient n équations linéaires dont les inconnues sont les n scalaires $a_{1j}, a_{2j}, \ldots, a_{jj}, x_{j+1}, \ldots, x_n$. En effet, la linéarité de f permet de la réécrire sous la forme :

$$\sum_{i=j+1}^{n} x_i f(v_i) \;=\; \sum_{i=1}^{j} a_{ij} u_i \;+\; \lambda_{j+1} \sum_{i=j+1}^{n} x_i v_i,$$

soit encore :

$$\sum_{i=1}^{j} a_{ij} u_i \;+\; \sum_{i=j+1}^{n} x_i \big(\lambda_{j+1} v_i - f(v_i)\big) \;=\; 0,$$

ce qui constitue bien un système de n équations linéaires, puisqu'il s'agit d'une relation vectorielle dans un espace de dimension n. Toute solution non nulle de ce système fournit d'un même coup un vecteur u_{j+1} possible et les coefficients correspondants de la $j^{\text{ème}}$ colonne de T : $a_{1j}, a_{2j}, \ldots, a_{jj}$.

Exemple Soit f l'endomorphisme de $\mathbb{R}^4$ dont la matrice dans la base canonique de $\mathbb{R}^4$ est :

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

Commençons par vérifier si f est trigonalisable en factorisant son polynôme caractéristique :

$$P_f(x) = \det(A - xI_4) = \begin{vmatrix} 1-x & 0 & 0 & 0 \\ 4 & -2-x & -3 & -3 \\ 0 & 4 & 4-x & 3 \\ -2 & -1 & 0 & 1-x \end{vmatrix} = (1-x)\begin{vmatrix} -2-x & -3 & -3 \\ 4 & 4-x & 3 \\ -1 & 0 & 1-x \end{vmatrix}$$

$$= (1-x)\left(-\begin{vmatrix} -3 & -3 \\ 4-x & 3 \end{vmatrix} + (1-x)\begin{vmatrix} -2-x & -3 \\ 4 & 4-x \end{vmatrix}\right) = (1-x)\left(3x-3 + (1-x)\begin{vmatrix} -2-x & -3 \\ 4 & 4-x \end{vmatrix}\right)$$

$$= (1-x)^2\left(-3 + \begin{vmatrix} -2-x & -3 \\ 4 & 4-x \end{vmatrix}\right) = (1-x)^2\left(1-2x-x^2\right) = (1-x)^4.$$

Puisque P_f est scindé, f est bien trigonalisable. Déterminons une base de son unique sous-espace propre E_1. E_1 est l'ensemble des vecteurs $(x, y, z, t) \in \mathbb{R}^4$ tels que :

$$(A - I_4)\begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 4 & -3 & -3 & -3 \\ 0 & 4 & 3 & 3 \\ -2 & -1 & 0 & 0 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad \text{soit encore : } \left[\begin{array}{l} 4x - 3y - 3z - 3t = 0 \\ 4y + 3z + 3t = 0 \\ -2x - y = 0, \end{array}\right.$$

ou encore, en rajoutant à la première équation les deux autres :

$$\left[\begin{array}{l} 2x = 0 \\ 4y + 3z + 3t = 0 \\ -2x - y = 0, \end{array}\right. \qquad \text{c'est-à-dire : } \qquad \left[\begin{array}{l} x = y = 0 \\ z + t = 0. \end{array}\right.$$

Ainsi, on a $E_1 = \mathrm{Ker}(f - \mathrm{Id}_{\mathbb{R}^4}) = \{(0, 0, z, -z)\ ;\ z \in \mathbb{R}\}$, autrement dit, E_1 est la droite vectorielle $\mathrm{Vect}(u_1)$ engendrée par le vecteur $u_1 = (0, 0, 1, -1)$. Comme $P_f(x) = (1-x)^4$, toute matrice triangulaire supérieure représentant f sera nécessairement de la forme :

$$T = \begin{pmatrix} 1 & a & b & c \\ 0 & 1 & b' & c' \\ 0 & 0 & 1 & c'' \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Afin de trouver un vecteur de base u_2 convenable, complétons notre unique vecteur u_1 en une base de $\mathbb{R}^4$ par les vecteurs $e_1 = (1, 0, 0, 0)$ et $e_2 = (0, 1, 0, 0)$ et $e_3 = (0, 0, 1, 0)$ de sa base canonique $\mathcal{B}$. (u_1, e_1, e_2, e_3) est bien une base de $\mathbb{R}^4$ car :

$$\det_{\mathcal{B}}(e_1, e_2, e_3, u_1) = \det_{\mathcal{B}}(e_1, e_2, e_3, e_3) - \det_{\mathcal{B}}(e_1, e_2, e_3, e_4) = 0 - 1 \neq 0.$$

Cherchons donc u_2 sous la forme d'une combinaison linéaire de ces vecteurs additionnels e_1, e_2, e_3 : $u_2 = xe_1 + ye_2 + ze_3 = (x, y, z, 0)$. La seconde colonne de la matrice T impose : $f(u_2) = au_1 + u_2$. Cette relation s'écrit sur la base canonique $\mathcal{B}$:

$$A\begin{pmatrix} x \\ y \\ z \\ 0 \end{pmatrix} = a\begin{pmatrix} 0 \\ 0 \\ 1 \\ -1 \end{pmatrix} + \begin{pmatrix} x \\ y \\ z \\ 0 \end{pmatrix}, \qquad \text{soit encore : } \qquad \left[\begin{array}{rcr} x & = & x \\ 4x - 2y - 3z & = & y \\ 4y + 4z & = & a + z \\ -2x - y & = & -a\,. \end{array}\right.$$

On vérifie facilement que ce système a pour solutions les quadruplets (a, x, y, z) tels que : $x = 0$, $y = a$ et $z = -y$. On peut donc choisir $a = 1$ et $u_2 = (0, 1, -1, 0)$.

Afin de trouver un vecteur u_3 convenable, complétons maintenant le système libre (u_1, u_2) en une base de $\mathbb{R}^4$ par les vecteurs e_1, e_2 de sa base canonique $\mathcal{B}$. (u_1, u_2, e_1, e_2) est bien une base de $\mathbb{R}^4$ car :

$$\det_{\mathcal{B}}(e_1, e_2, u_2, u_1) = \det_{\mathcal{B}}(e_1, e_2, e_2, u_1) - \det_{\mathcal{B}}(e_1, e_2, e_3, u_1) = 0 - (-1) \neq 0.$$

Cherchons donc u_3 sous la forme d'une combinaison linéaire de ces vecteurs additionnels e_1, e_2 : $u_2 = xe_1 + ye_2 = (x, y, 0, 0)$. La troisième colonne de la matrice T impose : $f(u_3) = bu_1 + b'u_2 + u_3$. Cette relation s'écrit sur la base canonique $\mathcal{B}$:

$$A\begin{pmatrix} x \\ y \\ 0 \\ 0 \end{pmatrix} = b\begin{pmatrix} 0 \\ 0 \\ 1 \\ -1 \end{pmatrix} + b'\begin{pmatrix} 0 \\ 1 \\ -1 \\ 0 \end{pmatrix} + \begin{pmatrix} x \\ y \\ 0 \\ 0 \end{pmatrix}, \quad \text{soit encore :} \quad \left[\begin{array}{rcl} x & = & x \\ 4x - 2y & = & b' + y \\ 4y & = & b - b' \\ -2x - y & = & -b\,. \end{array}\right.$$

On vérifie facilement que ce système a pour solutions les quadruplets (b, b', x, y) tels que : $x = 0$, $y = b$ et $b' = -3b$. On peut donc choisir $b = 1$, $b' = -3$ et $u_3 = (0, 1, 0, 0)$.

Nous pouvons achever cette trigonalisation en complétant le système libre (u_1, u_2, u_3) par n'importe quel vecteur u_4 tel que (u_1, u_2, u_3, u_4) soit une base de $\mathbb{R}^4$. Choisissons par exemple $u_4 = e_1$. (u_1, u_2, u_3, u_4) est alors bien une base de $\mathbb{R}^4$ car il s'agit — à l'ordre près des vecteurs — de la base (u_1, u_2, e_1, e_2) considérée plus haut. La dernière colonne de la matrice T impose la relation : $f(u_4) = cu_1 + c'u_2 + c''u_3 + u_4$, et celle-ci s'écrit sur la base canonique $\mathcal{B}$:

$$A\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} = c\begin{pmatrix} 0 \\ 0 \\ 1 \\ -1 \end{pmatrix} + c'\begin{pmatrix} 0 \\ 1 \\ -1 \\ 0 \end{pmatrix} + c''\begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad \text{soit encore :} \quad \left[\begin{array}{rcl} 1 & = & 1 \\ 4 & = & c' + c'' \\ 0 & = & c - c' \\ -2 & = & -c\,, \end{array}\right.$$

d'où : $c = c' = c'' = 2$.

Ainsi, la matrice de passage de la base canonique à la base (u_1, u_2, u_3, u_4) et la matrice de f dans cette dernière sont respectivement :

$$P = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 1 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix}, \qquad T = \begin{pmatrix} 1 & 1 & 1 & 2 \\ 0 & 1 & -3 & 2 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

3.4 Application aux systèmes différentiels linéaires

Définition • On appelle *système différentiel linéaire à coefficients constants avec second membre* un système de la forme :

$$\left[\begin{array}{rcl} x_1'(t) & = & a_{11}x_1(t) + \cdots + a_{1n}x_n(t) + b_1(t) \\ & \vdots & \\ x_n'(t) & = & a_{n1}x_1(t) + \cdots + a_{nn}x_n(t) + b_n(t) \end{array}\right. \qquad (S)$$

où $b_1, \ldots, b_n : I \to \mathbf{K}$ sont des fonctions *continues* sur un intervalle I de $\mathbb{R}$ et à valeurs dans $\mathbf{K}$.

Une solution sur I de (S) consiste en n fonctions $x_1, \dots, x_n$ dérivables sur I et à valeurs dans $\mathbf{K}$ vérifiant le système pour tout $t \in I$. En posant :

$$A = \begin{pmatrix} a_{11} \dots a_{1n} \\ \vdots \qquad \vdots \\ a_{n1} \dots a_{nn} \end{pmatrix} \in M_n(\mathbf{K}), \qquad \begin{array}{rcl} B : I & \to & \mathbf{K}^n \\ t & \mapsto & B(t) = \end{array} \begin{pmatrix} b_1(t) \\ \vdots \\ b_n(t) \end{pmatrix} \quad \text{et} \quad \begin{array}{rcl} X : I & \to & \mathbf{K}^n \\ t & \mapsto & X(t) = \end{array} \begin{pmatrix} x_1(t) \\ \vdots \\ x_n(t) \end{pmatrix},$$

le système (S) s'écrit plus simplement :

$$X'(t) = \underset{\substack{\uparrow \\ \text{matrice “à coefficients constants”}}}{A}\, X(t) + \underset{\substack{\uparrow \\ \text{“second membre” } (X'(t) - A\,X(t) = \underline{B(t)})}}{B(t)}.$$

• On appelle *condition initiale* du système (S) la donnée d'une “date” $t_0 \in I$ et d'une “position” $X_0 \in \mathbf{K}^n$. Une solution sur I de (S) vérifiant $X(t_0) = X_0$ est alors appelée *solution de (S) sur I pour la condition initiale $X(t_0) = X_0$*.

• On appelle *système différentiel linéaire homogène* ou encore *système différentiel linéaire sans second membre* associé à (S) le système : $X'(t) = A\,X(t)$.

Commençons par remarquer que les solutions de tels systèmes se décrivent en termes d'espaces vectoriels et affines.

Proposition 10 *Soit $X'(t) = A\,X(t) + B(t)$ un système différentiel linéaire avec second membre, $X_P : I \to \mathbf{K}^n$ une solution particulière de ce système et E l'ensemble des solutions sur un intervalle $I \subset \mathbb{R}$ de son système linéaire homogène associé : $X'(t) = A\,X(t)$. Alors :*

- *E est un $\mathbf{K}$-espace vectoriel pour la multiplication par un scalaire et l'addition usuelles des fonctions de I dans $\mathbf{K}^n$.*
- *L'ensemble des solutions du système avec second membre est :*

$$F = \{X_P + X \ ; \ X \in E\}.$$

Autrement dit, l'ensemble des solutions sur I de $X'(t) = A\,X(t) + B(t)$ est le sous-espace affine F parallèle à E et passant par le point X_P.

Démonstration Si $k, l \in \mathbf{K}$ et $X, Y \in E$, alors pour tout $t \in I$:

$$(kX + lY)'(t) = k\,X'(t) + l\,Y'(t) = k\,A\,X(t) + l\,A\,Y(t) = A\,(k\,X(t) + l\,Y(t)) = A\,(kX + lY)(t),$$

d'où $kX + lY \in E$, ce qui établit que E est bien un $\mathbf{K}$-espace vectoriel. De plus, pour toute fonction $X : I \to \mathbf{K}^n$ et tout $t \in I$, on a :

$$\begin{aligned} X'(t) = A\,X(t) \quad &\iff \quad X'(t) + X_P'(t) = A\,X(t) + A\,X_P(t) + B(t) \\ &\iff \quad (X_P + X)'(t) = A\,(X_P + X)(t) + B(t), \end{aligned}$$

i.e. X est solution sur I de $X'(t) = A\,X(t)$ ssi $X_P + X$ est solution sur I de $X'(t) = A\,X(t) + B(t)$.

□

La seconde assertion de cette proposition ne fait qu'exprimer en termes géométriques la règle :

Solution générale de l'équation avec 2^{nd} membre	$=$	solution particulière de l'équation avec 2^{nd} membre	$+$	solution générale de l'équation homogène associée.

Rappels sur le cas $n = 1$ (cf MI2).

L'établissement de toutes les solutions d'une équation linéaire homogène du 1^er^ ordre est remarquablement simple :

Proposition 11 *Les solutions sur un intervalle I de $\mathbb{R}$ de l'équation linéaire homogène du 1^er^ ordre $x'(t) = a\,x(t)$ ($a \in \mathbf{K}$) sont les fonctions de la forme : $x(t) = k\,e^{at}$ ($k \in \mathbf{K}$).*

Démonstration Les fonctions $x(t) = k\,e^{at}$ sont clairement solutions de l'équation $x'(t) = a\,x(t)$. Réciproquement, si $x : I \to \mathbf{K}$ est solution sur I de l'équation $x'(t) = a\,x(t)$, alors x est dérivable sur I ainsi que la fonction u définie sur I par $u(t) = x(t)\,e^{-at}$ et l'on obtient pour tout $t \in I$: $x(t) = u(t)\,e^{at}$, $x'(t) = u'(t)\,e^{at} + a\,u(t)\,e^{at}$. En remplaçant ces expressions dans $x'(t) = a\,x(t)$, il vient : $u'(t) = 0$, de sorte que u est une fonction constante sur I, mettons $u(t) = k$, d'où pour tout $t \in I$: $x(t) = k\,e^{at}$.

$\square$

Méthode de variation des constantes. Cette méthode permet de trouver les solutions d'une équation linéaire du 1^er^ ordre avec second membre $x'(t) = a\,x(t) + b(t)$ ($a \in \mathbf{K}$, $b : I \to \mathbf{K}$) à partir de la solution générale de l'équation linéaire homogène associée $x'(t) = a\,x(t)$, en reprenant l'idée de la démonstration ci-dessus : On fait "varier la constante" k de la solution générale de l'équation homogène $x(t) = k\,e^{at}$; autrement dit, on cherche les solutions de $x'(t) = a\,x(t) + b(t)$ sous la forme $x(t) = k(t)\,e^{at}$. Il vient alors : $k'(t)\,e^{at} + a\,k(t)\,e^{at} = a\,k(t)\,e^{at} + b(t)$, d'où : $k'(t) = e^{-at}b(t)$. Comme la solutions de cette dernière équation sont les fonctions : $k(t) = \int e^{-at}b(t)\,dt + C$, la solution génerale de l'équation avec second membre est : $x(t) = e^{at}\int e^{-at}b(t)\,dt + C\,e^{at}$.

Remarque Le premier terme $e^{at}\int e^{-at}b(t)\,dt$ de cette dernière expression est une solution particulière de l'équation avec second membre, tandis que son second terme $C\,e^{at}$ est la solution générale de l'équation homogène associée. On retrouve donc bien la règle :

$$\begin{array}{c} \text{Solution générale de} \\ \text{l'équation avec } 2^{\text{nd}} \text{ membre} \end{array} = \begin{array}{c} \text{solution particulière de} \\ \text{l'équation avec } 2^{\text{nd}} \text{ membre} \end{array} + \begin{array}{c} \text{solution générale de l'équation} \\ \text{homogène associée.} \end{array}$$

Théorème 12 *(existence globale et unicité des solutions des systèmes linéaires avec second membre) Soit I un intervalle de $\mathbb{R}$, $t_0 \in I$, $A \in M_n(\mathbf{K})$ et $B : I \to \mathbf{K}^n$ une fonction continue sur I ($\mathbf{K} = \mathbb{R}$ ou $\mathbb{C}$). Alors pour tout $X_0 \in \mathbf{K}$, il existe une solution et une seule sur I du système $X'(t) = A\,X(t) + B(t)$ pour la condition initiale $X(t_0) = X_0$.*

Démonstration Commençons par le cas le plus simple à traiter : $\mathbf{K} = \mathbb{C}$.
On établit alors le théorème par récurrence sur n. Lorsque $n = 1$, cela résulte des rappels ci-dessus : la seule solution sur I d'une équation $z'(t) = a\,z(t) + b(t)$ pour la condition initiale $z(t_0) = z_0$ est : $z(t) = e^{at}\int e^{-at}b(t)\,dt + k_0e^{at} = e^{at}F(t) + k_0e^{at}$, où $k_0 = z_0e^{-at_0} - F(t_0)$. Supposons le théorème établi au rang $n-1$ et considérons un système $Z'(t) = A\,Z(t) + B(t)$, où $A \in M_n(\mathbb{C})$, $B : I \to \mathbb{C}^n$. Soit λ une valeur propre quelconque de A (il en existe puisque $\mathbf{K} = \mathbb{C}$) et v_n un vecteur propre associé que l'on complète en une base $(v_1, \ldots, v_n)$. L'endomorphisme représenté sur la base canonique de $\mathbb{C}^n$ par la matrice A a alors sur la base $(v_1, \ldots, v_n)$ une matrice C de la forme :

$$C = \left(\begin{array}{c|c} & 0 \\ E & \vdots \\ & 0 \\ \hline F & \lambda \end{array}\right)$$

où $E \in M_{n-1}(\mathbb{C})$ et en notant P la matrice de passage de la base canonique à la base $(v_1, \ldots, v_n)$, on a : $A = PCP^{-1}$. Soit $D : I \to \mathbb{C}^n$ la fonction définie sur I par $D(t) = P^{-1}B(t)$ et pour toute fonction $Z : I \to \mathbb{C}^n$, notons U la fonction définie sur I par $U(t) = P^{-1}Z(t)$. Comme P^{-1} est une matrice à coefficients constants, on vérifie alors facilement (par linéarité de la dérivation) : $U'(t) = P^{-1}Z'(t)$. On a alors pour tout $t \in I$:

$$Z'(t) = A\,Z(t)+B(t) \quad \Leftrightarrow \quad P^{-1}Z'(t) = P^{-1}(PCP^{-1})Z(t)+P^{-1}B(t) \quad \Leftrightarrow \quad U'(t) = C\,U(t)+D(t),$$

de sorte que Z est solution sur I du système $Z'(t) = A\,Z(t) + B(t)$ pour la condition initiale $Z'(t_0) = Z_0$ si et seulement si U est solution sur I du système $U'(t) = C\,U(t) + D(t)$ pour la condition initiale $U'(t_0) = P^{-1}Z_0$. Ainsi, il suffit d'établir l'existence et l'unicité d'une solution sur I de $U'(t) = C\,U(t) + D(t)$ pour une condition initiale donnée $U'(t_0) = U_0$.
En posant :

$$D(t) = \begin{pmatrix} d_1(t) \\ \vdots \\ \vdots \\ d_n(t) \end{pmatrix}, \quad G(t) = \begin{pmatrix} d_1(t) \\ \vdots \\ d_{n-1}(t) \end{pmatrix}, \quad U(t) = \begin{pmatrix} u_1(t) \\ \vdots \\ \vdots \\ u_n(t) \end{pmatrix}, \quad V(t) = \begin{pmatrix} u_1(t) \\ \vdots \\ u_{n-1}(t) \end{pmatrix}, \quad U_0 = \begin{pmatrix} V_0 \\ \hline k \end{pmatrix},$$

on a :

$$\left[\begin{array}{l} \forall\, t\in I \quad U'(t) = C\,U(t) + D(t) \\ \\ U(t_0) = U_0 \end{array}\right. \quad \Leftrightarrow \quad \left[\begin{array}{ll} \forall\, t\in I \quad V'(t) = E\,Z(t) + G(t) & (1) \\ \forall\, t\in I \quad u_n'(t) = \lambda\, u_n(t) + F\,V(t) + d_n(t) & (2) \\ V(t_0) = V_0 & (3) \\ u_n(t_0) = k & (4) \end{array}\right.$$

L'hypothèse de récurrence entraîne qu'il existe une unique fonction $V : I \to \mathbb{C}^{n-1}$ satisfaisant (1) et (3) et, pour cette fonction V, le cas $n = 1$ entraîne qu'il existe une fonction $u_n : I \to \mathbb{C}$ et une seule satisfaisant (2) et (4). Cela établit l'existence et l'unicité de la solution sur I de $U'(t) = C\,U(t) + D(t)$ pour la condition initiale $U'(t_0) = U_0$.

Le cas où $\mathbf{K} = \mathbb{R}$ se déduit facilement du précédent. En effet, l'unicité d'une solution à valeurs réelles résulte immédiatement de l'unicité d'une solution à valeurs complexes. De plus, tout système $X'(t) = A\,X(t) + B(t)$ $(A \in M_n(\mathbb{R}),\ B : I \to \mathbb{R}^n)$ possède une unique solution $Z : I \to \mathbb{C}^n$ satisfaisant une condition initiale $Z(t_0) = X_0 \in \mathbb{R}^n$ donnée. En écrivant $Z(t)$ sous la forme $Z(t) = X(t) + i\,Y(t)$ $(X(t) \in \mathbb{R}^n,\ Y(t) \in \mathbb{R}^n)$, on obtient :

$$X'(t) + i\,Y'(t) = A(X(t) + i\,Y(t)) + B(t) = (A\,X(t) + B(t)) + i\,A\,Y(t),$$

d'où pour tout $t \in I$: $Y'(t) = A\,Y(t)$. Ainsi, Y est l'unique solution de $Y'(t) = A\,Y(t)$ pour la condition initiale $Y(t_0) = \mathrm{Im}(X_0) = 0$. Comme $Y = 0$ est une solution évidente de $Y'(t) = A\,Y(t)$ pour cette condition initiale, on en déduit pour tout $t \in I$: $Y(t) = 0$; autrement dit, l'unique solution du système $Z'(t) = A\,Z(t) + B(t)$ pour la condition initiale $Z(t_0) = X_0 \in \mathbb{R}^n$ est $Z = X$.

□

Méthode pratique de résolution des systèmes différentiels linéaires à coefficients constants.

Cette méthode suit à peu de choses près le canevas de la démonstration ci-dessus. Considérons un système $X'(t) = A\,X(t) + B(t)$ $(A \in M_n(\mathbf{K}),\ B : I \to \mathbf{K}^n)$. Pour le résoudre, on commence par trigonaliser dans le pire des cas — sinon diagonaliser — la matrice A ; mettons $A = PTP^{-1}$, où :

$$T = \begin{pmatrix} c_{1,1} & c_{1,2} & \cdots & \cdots & c_{1,n} \\ & c_{2,2} & & & c_{2,n} \\ & & \ddots & & \vdots \\ & & & c_{n-1,n-1} & c_{n-1,n} \\ \text{\Large 0} & & & & c_{n,n} \end{pmatrix}.$$

Puis on effectue le même changement de variables que dans la preuve du théorème 12. En posant :

$$U(t) = \begin{pmatrix} u_1(t) \\ \vdots \\ u_n(t) \end{pmatrix} = P^{-1}X(t), \qquad D(t) = \begin{pmatrix} d_1(t) \\ \vdots \\ d_n(t) \end{pmatrix} = P^{-1}B(t),$$

le système $X'(t) = A\,X(t) + B(t)$ équivaut au système $U'(t) = T\,U(t) + D(t)$, lequel s'écrit :

$$\left[\begin{array}{rcl} u_1'(t) & = & c_{1,1}u_1(t) + \big(c_{1,2}u_2(t) + c_{1,3}u_3(t) + \cdots + c_{1,n}u_n(t) + d_1(t)\big) \\ u_2'(t) & = & c_{2,2}u_1(t) + \big(c_{2,3}u_2(t) + \cdots + c_{2,n}u_n(t) + d_2(t)\big) \\ & \vdots & \\ u_{n-1}'(t) & = & c_{n-1,n-1}u_{n-1}(t) + \big(c_{n-1,n}u_n(t) + d_{n-1}(t)\big) \\ u_n'(t) & = & c_{n,n}u_n(t) + d_n(t) \end{array}\right.$$

La résolution de ce système se ramène à celle de n équations linéaires du 1[er] ordre avec second membre, quand on les resoud de bas en haut en remplaçant dans chacune les solutions des équations inférieures.

Exercice. À l'aide de la trigonalisation effectuée dans la section précédente, résoudre le système :

$$\left[\begin{array}{l} x_1'(t) = x_1(t) \\ x_2'(t) = 4x_1(t) - 2x_2(t) - 3x_3(t) - 3x_4(t) \\ x_3'(t) = 4x_2(t) + 4x_3(t) + 3x_4(t) \\ x_4'(t) = -2x_1(t) - x_2(t) + x_4(t) \end{array}\right.$$

Enfin remarquons que le théorème d'existence et d'unicité ci-dessus permet de préciser la première assertion de la proposition 10 :

Proposition 13 *Soit $X'(t) = A\,X(t)$ un système différentiel linéaire* homogène *à coefficients constants et E l'ensemble de ses solutions sur un intervalle $I \neq \varnothing$ de $\mathbb{R}$ quelconque. Alors :*

- *E est un $\mathbf{K}$-espace vectoriel* de dimension n *pour la multiplication par un scalaire et l'addition usuelles des fonctions,*
- *pour tout $t_0 \in I$, l'application $\varphi_{t_0} : E \to \mathbb{R}^n$ définie par $\varphi_{t_0}(X) = X(t_0)$ est un isomorphisme d'espace vectoriels.*

Démonstration Soit $t_0 \in I$. Pour tout $k, l \in \mathbf{K}$:

$$\varphi_{t_0}(kX + lY) = (kX + lY)(t_0) = k\,X(t_0) + l\,Y(t_0) = k\,\varphi_{t_0}(X) + l\,\varphi_{t_0}(Y),$$

donc φ_{t_0} est une application linéaire de E dans $\mathbf{K}^n$. Le théorème 12 exprime très exactement la bijectivité de φ_{t_0}. Ainsi, E et $\mathbf{K}^n$ ont même dimension, à savoir n.

$\square$