北航中法工程师学院
Centrale Pékin

Algèbre 3

École Centrale de Pékin

Cours de mathématiques du cycle préparatoire

9 décembre 2021

Table des matières

Avant-propos

Vous trouverez au fil de ce cours différents symboles :

- Le symbole "", situé dans la marge, signifie que le point correspondant est un point délicat (il s'agit d'un *virage dangereux*). Plus ce symbole est gros, plus le point en question est subtil.
- Le symbole "◄", situé dans la marge, signifie qu'un point technique dans le fil du cours n'a pas été totalement explicité. Il s'agit d'un choix volontaire afin d'obliger l'étudiant à écrire entièrement le détail du point manquant. Si ce point n'est pas détaillé, c'est qu'il ne *doit pas* présenter de difficulté.
- Le symbole "□" est un marqueur signifiant la fin d'une démonstration.
- Ce cours peut comporter des fautes de frappe, des coquilles, voire des erreurs d'argumentation. Ainsi, il faut toujours être vigilant lorsque vous suivez et que vous travaillez ce cours. Vérifier que les exemples sont justes et que les preuves n'ont pas de fautes est un exercice très utile (et indispensable) en mathématiques pour comprendre les notions et comprendre leurs utilisations.

Vous trouverez aussi des notations mathématiques :

E	un ensemble ou un espace vectoriel
F	un sous-espace vectoriel de l'espace vectoriel E
u	un endomorphisme de l'espace vectoriel E
$\sigma(E)$	l'ensemble des bijections d'un ensemble E dans lui-même
$\circ$	la composition des applications, donc $(\sigma(E), \circ)$ est un groupe.
$\mathcal{S}_n$	l'ensemble des permutations de $[\![1, n]\!]$ (le groupe symétrique d'ordre n)
(i, j)	une transposition
$(a_1, a_2, \ldots, a_p)$	un p-cycle
$I(\sigma)$	le nombre d'inversions de σ
$\varepsilon(\sigma)$	la signature de σ
$\mathcal{A}_n$	groupe alterné d'ordre n
$\mathcal{L}_p(E, F)$	l'espace vectoriel des applications p-linéaires de E dans un $\mathbb{K}$-espace vectoriel F
$\Lambda^p(E)$	l'espace vectoriel des formes p-linéaires de E dans un $\mathbb{K}$-espace vectoriel F
$\det_{\mathcal{B}}(u_1, \ldots, u_n)$	le déterminant de la famille de vecteurs $(u_1, \ldots, u_n)$ par rapport à la base $\mathcal{B}$
$\det(f)$	le déterminant de l'endomorphisme f
$\det(A)$	le déterminant d'une matrice carrée A
$\mathrm{com}\,(A)$	la comatrice
$\mathrm{Tr}\,(u)$	la trace d'un endomorphisme u
$\mathrm{Tr}\,(M)$	la trace d'une matrice carrée M
$E_\lambda(u)$	le sous-espace vectoriel propre à une valeur propre λ de l'endomorphisme u
$F_\lambda(u)$	le sous-espace vectoriel caractéristique à une valeur propre λ de l'endomorphisme u
$\mathrm{Spec}\,(u)$	le spectre d'un endomorphisme u
$\chi_A(X)$	le polynôme caractéristique d'une matrice carrée A
$\chi_u(X)$	le polynôme caractéristique d'un endomorphisme u
$C_{P(X)}$	la matrice compagnon d'un polynôme P
$m(\lambda)$	l'ordre de multiplicité d'une valeur propre λ
$r(\lambda)$	l'indice d'une valeur propre λ
e_u	le morphisme d'évaluation des polynômes en un endomorphisme u
e_A	le morphisme d'évaluation des polynômes en une matrice carré A
$\mathbb{K}[u]$	la sous-algèbre engendrée par un endomorphisme u
$\mathcal{I}_u$	l'idéal annulateur de u
M_u	le polynôme minimal d'un endomorphisme u
M_A	le polynôme minimal d'une matrice carrée A
$< x \| y >$	le produit scalaire de deux vecteurs(éléments) x et y
$N(x), \|\|x\|\|$	la norme de x
$d(x, y)$	la distance entre deux vecteurs(éléments) x et y
$x \perp y$	deux vecteurs(éléments) x et y sont orthogonaux
$A^\perp$	l'orthogonal ou le supplémentaire orthogonal de A

$E = F \overset{\perp}{\oplus} F^{\perp}$	F et $F^{\perp}$ sont supplémentaires de E
ϕ_a	la forme linéaire $\phi_a(x) =< a\|x >$
P_F	la projection orthogonale sur F parallèlement à $F^{\perp}$

Chapitre 1 Algèbre bilinéaire

Table des matières du chapitre

1.1 Définitions

1.1.1 Formes bilinéaires, formes bilinéaires symétriques

Dans ce cours, $\mathbb{K}$ désigne un corps (comme $\mathbb{R}$, $\mathbb{Q}$, $\mathbb{C}$, $\mathbb{Z}/p\mathbb{Z}$). Pour certains objets (ex : produit scalaire) nous nous restreindrons à $\mathbb{Q}, \mathbb{R}, \mathbb{C}$, voire seulement $\mathbb{R}$.

Définition 1
Soit E un $\mathbb{K}$-espace vectoriel E. Soit $f : E \times E \longrightarrow \mathbb{K}$.
On dit que f est une **forme bilinéaire** \双线性函数\, si pour tous $x, y \in E$, les fonctions

$$\begin{array}{rccc} f(.,y) : & E & \longrightarrow & \mathbb{K} \\ & x' & \longmapsto & f(x',y) \end{array} \quad \text{et} \quad \begin{array}{rccc} f(x,.) : & E & \longrightarrow & \mathbb{K} \\ & y' & \longmapsto & f(x,y') \end{array}$$

sont des applications linéaires.

Exemples 2

1. *Pour $E = \mathbb{R}^2$, la fonction $f : (u,v) \in \mathbb{R}^2 \times \mathbb{R}^2 \mapsto xx' + yy' \in \mathbb{R}$ est une forme bilinéaire.*
2. *Pour $E = \mathbb{C}$, la fonction $f : (z,z') \in \mathbb{C} \times \mathbb{C} \mapsto zz' \in \mathbb{C}$ est une forme bilinéaire.*
3. *Pour $E = \mathbb{R}[X]$, la fonction $f : (P,Q) \in E^2 \mapsto P(0)Q(0) \in \mathbb{R}$ est une forme bilinéaire.*
4. *Pour $E = \mathcal{C}^0([a,b],\mathbb{R})$, avec $a < b$, la fonction $S : (f,g) \in E^2 \mapsto \int_a^b f(t)g(t)dt \in \mathbb{R}$ est une forme bilinéaire.*

Remarque 3 — *Une forme bilinéaire $f : E \times E \to \mathbb{K}$ est très différente d'une forme linéaire $g : E \times E \to \mathbb{K}$. Pour $(x,y), (x',y') \in E \times E$, on a $g(x+x', y+y') = g(x,y) + g(x',y')$ par linéarité, tandis que*

$$f(x+x', y+y') = f(x,y) + f(x,y') + f(x',y) + f(x',y'), \text{ par bilinéarité .}$$

De même, pour $\lambda \in \mathbb{K}$ on a $g(\lambda x, \lambda y) = \lambda g(x,y)$ par linéarité, tandis que

$$f(\lambda x, \lambda y) = \lambda f(x, \lambda y) = \lambda^2 f(x,y), \text{ par bilinéarité .}$$

Remarque 4 — *Pour $f, g : E \times E \to \mathbb{K}$ des formes bilinéaires et $\lambda \in \mathbb{K}$, la fonction $f + \lambda g$ est encore une forme bilinéaire.*

En effet, pour tous $x, y \in E$, les fonctions $f(x,.) + \lambda g(x,.)$ et $f(.,y) + \lambda g(.,y)$ sont des applications linéaires (toute combinaison linéaire d'applications linéaires est une application linéaire).

Donc, l'ensemble des formes bilinéaires sur E est un $\mathbb{K}$-espace vectoriel.

1.1.2 Matrice d'une forme bilinéaire

Proposition-Définition 5
Soit E un $\mathbb{K}$-espace vectoriel de dimension finie n.
Soient $\mathcal{B} = (e_1, \cdots, e_n)$ une base de E et $f : E^2 \to \mathbb{K}$ une forme bilinéaire.

Soit $A = (a_{ij})_{i,j} \in \mathscr{M}_n(\mathbb{K})$ avec $a_{i,j} = f(e_i, e_j)$. On dit que A est la matrice de la forme bilinéaire f dans la base $\mathcal{B}$, notée $Mat_{\mathcal{B}}(f)$.

Pour $x, y \in E$ avec $x = x_1e_1 + \cdots + x_ne_n$ et $y = y_1e_1 + \cdots + y_ne_n$, on a :

$$f(x,y) = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_iy_i = \underbrace{\begin{pmatrix} x_1 & x_2 & \cdots & x_n \end{pmatrix}}_{{}^tX} \underbrace{\begin{pmatrix} a_{1,1} & a_{1,2} & \ldots & a_{1,p} \\ a_{2,1} & a_{2,2} & \ldots & a_{2,p} \\ \vdots & \vdots & & \vdots \\ a_{n,1} & a_{n,2} & \ldots & a_{n,p} \end{pmatrix}}_{A} \underbrace{\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}}_{Y} = {}^tX \cdot A \cdot Y.$$

Preuve —

$$\begin{aligned} f(x,y) &= f(\sum_{i=1}^{n} x_ie_i, \sum_{j=1}^{n} y_je_j) = \sum_{i=1}^{n} x_if(e_i, \sum_{j=1}^{n} y_je_j) = \sum_{i=1}^{n}\sum_{j=1}^{n} x_iy_jf(e_i,e_j) \quad \text{par linéarité de } f(.,y) \text{ et } f(e_i,.). \\ &= \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_iy_j = \sum_{i=1}^{n} x_i \left(\sum_{j=1}^{n} a_{ij}y_j\right) = {}^tX \cdot (A \cdot Y)\,. \end{aligned}$$

□

Définition 6

Soient E un $\mathbb{K}$-espace vectoriel (abrégé e.v.) et $f : E \times E \to \mathbb{K}$ une forme bilinéaire. On dit que f est une forme bilinéaire **symétrique** si l'on a $f(x,y) = f(y,x)$ pour tous $x, y \in E$.

Exemple 7 —

- *Pour* $E = \mathbb{K}[X]$ *et* $\varphi : (P,Q) \in E^2 \mapsto P(1)Q(1) \in \mathbb{K}$, φ *est une forme bilinéaire symétrique.*
- *Pour* $E = \mathbb{R}^2$, $\varphi : (x,y) \in E^2 \mapsto x_1y_1 + y_2y_2 \in \mathbb{R}$ *est une forme bilinéaire symétrique.*
- *Pour* $E = \mathbb{R}^2$, $\varphi : (x,y) \in E^2 \mapsto x_1y_2 - x_2y_1 \in \mathbb{R}$ *et* $\psi : (x,y) \in E^2 \mapsto x_1y_2 \in \mathbb{R}$ *sont des formes bilinéaires qui ne sont pas symétriques* ($\varphi((1,0),(0,1)) = 1 \neq -1 = \varphi((0,1),(1,0))$ *).*

Proposition 8

Soient E un $\mathbb{K}$-e.v. de dimension finie n, et $\mathcal{B} = (e_1, \ldots, e_n)$ une base de E. Soit $f : E \times E \to \mathbb{K}$ une forme bilinéaire.

Alors f est une forme bilinéaire symétrique si et seulement si $Mat_{\mathcal{B}}(f)$ est une matrice symétrique. La forme bilinéaire φ est symétrique si, et seulement si, la matrice A est symétrique.

Preuve —

- Les coefficients de la matrice $A = Mat_{\mathcal{B}}(f)$ sont les $a_{i,j} = f(e_i, e_j)$, $1 \leq i, j \leq n$.
  Si f est symétrique, alors on a $f(e_i, e_j) = f(e_j, e_i)$, d'où $a_{ij} = a_{ji}$, donc A est une matrice symétrique ($A = {}^tA$).
- Réciproquement, supposons A est symétrique. On a $f(x,y) = {}^tX \cdot A \cdot Y$. La matrice ${}^tX \cdot A \cdot Y$ est une matrice 1×1. Elle est donc égale à sa transposée. D'où :
  $$f(x,y) = {}^tX \cdot A \cdot Y = {}^t({}^tX \cdot A \cdot Y) = {}^tY \cdot {}^tA \cdot {}^t({}^tX) = {}^tY \cdot A \cdot X = f(y,x).$$
  Donc f est une forme bilinéaire symétrique.

□

Il ne faut pas confondre la matrice d'une forme bilinéaire $\varphi : E \times E \to \mathbb{K}$ et la matrice d'une application linéaire $f : E \to E$. Pour $\mathcal{B} = (e_1, e_2, \cdots, e_n)$ de E, une même matrice $A \in \mathcal{M}_n(\mathbb{K})$ peut représenter un endomorphisme $f(X \mapsto AX)$ et une forme bilinéaire φ $((X,Y) \mapsto {}^tXAY)$.

Si l'on change de base de E, ces matrices ne changent pas de la même manière.

Proposition 9

Soit E un $\mathbb{K}$-e.v. de dimension finie n. Soient $\mathcal{B}$, $\mathcal{B}'$ deux bases de E. Soient $f : E \to E$ une application linéaire et $\varphi : E^2 \to \mathbb{K}$ une forme bilinéaire. Soit P la matrice de passage de la base $\mathcal{B}$ vers $\mathcal{B}'$. Alors on a :

- $Mat_{\mathcal{B}'}(f) = P^{-1}Mat_{\mathcal{B}}(f)P$;
- $Mat_{\mathcal{B}'}(\varphi) = {}^tPMat_{\mathcal{B}}(\varphi)P$.

Preuve —

endomorphisme f	forme bilinéaire φ
$y = f(x) \Leftrightarrow Y = Mat_{\mathcal{B}}(f)X$	$\varphi(x,y) = {}^tXMat_{\mathcal{B}}(\varphi)Y$
$\Leftrightarrow Y' = BX'$	$= {}^tX'CY'$
$X = PX'$	
$Y = PY'$	
D'où $B = P^{-1}Mat_{\mathcal{B}}(f)P$.	D'où $C = {}^tPMat_{\mathcal{B}}(\varphi)P$.

□

1.1.3 Produit scalaire

DÉFINITION 10
Soit E un $\mathbb{R}$-espace vectoriel. Soit $S : E^2 \to \mathbb{R}$ une forme bilinéaire.
On dit que S est un **produit scalaire** \内积/数量积\ si celle-ci vérifie :

- $\forall x, y \in E$, $S(x,y) = S(y,x)$ (S est symétrique) ;
- $\forall x \in E$, $S(x,x) \geq 0$, avec $S(x,x) = 0 \iff x = 0$ (S est **définie positive**).

Un produit scalaire (ou forme bilinéaire symétrique définie positive) est noté $\langle x|y\rangle$, ou $\langle x,y\rangle$ ou $x \cdot y$.

REMARQUE 11 — *En géométrie, on utilise souvent la notation $\overrightarrow{u} \cdot \overrightarrow{v}$ pour désigner le produit scalaire des vecteurs $\overrightarrow{u}$ et $\overrightarrow{v}$.*

La notion de produit scalairé nécessite que le corps $\mathbb{K}$ possède des éléments "positifs". Nous n'étudierons le produit scalaire que pour $\mathbb{K} = \mathbb{R}$.

THÉORÈME 12 (Inégalité de Cauchy-Schwarz)
Soit E un $\mathbb{R}$-e.v. muni d'un produit scalaire $< .|. >$.

1. Pour x et y dans E, on a : $\langle x|y\rangle^2 \leqslant \langle x|x\rangle\langle y|y\rangle$.
2. Cette inégalité est une égalité si, et seulement si, x et y sont colinéaires ($x = \lambda y$ ou $y = \lambda x$, pour un $\lambda \in \mathbb{R}$).

Preuve

1.
   - Si $y = 0$, l'inégalité est évidente (c'est une égalité).
   - Sinon, posons $P(\lambda) =< x + \lambda\, y|x + \lambda\, y >= \lambda^2 \ < y|y > +2\lambda \ < x|y > + < x|x >$.
     Alors P est une fonction polynomiale de degré 2 (puisque $\forall y \neq 0, \langle y|y\rangle > 0$) telle que $P(\lambda) \geqslant 0$, $\forall \lambda \in \mathbb{R}$. Son discriminant :
     $$\Delta = 4 < x|y >^2 -4 < x|x >< y|y >$$
     est donc négatif ou nul, ce qui donne l'inégalité annoncée.
2.
   - Si x et y sont proportionnels, il existe un scalaire λ tel que $y = \lambda\, x$ ou $x = \lambda\, y$.
     Supposons par exemple $y = \lambda\, x$. Alors on a : $< x|y >^2=< x|\lambda\, x >^2= \lambda^2 \ < x|x >^2=< x|x >< y|y >$.
   - Réciproquement, supposons que $< x|y >^2=< x|x >< y|y >$.
     — Si $y = 0$, alors x et y sont proportionnels.
     — Sinon, le polynôme P est de degré 2 avec un discriminant nul. Il existe donc $\lambda \in \mathbb{R}$ tel que $P(\lambda) = 0$.
     Cela donne $< x + \lambda\, y|x + \lambda\, y >= 0$. Par définition du produit scalaire, on en déduit que $x + \lambda\, y = 0$, et donc que x est proportionnel à y.

□

EXEMPLES 13

1. *Pour $E = \mathbb{R}^n$, $n \geq 1$, la forme bilinéaire symétrique $S : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ définie par :*
   $$S\big((x_1, x_2, \ldots, x_n), (y_1, y_2, \ldots, y_n)\big) = \sum_{i=1}^{n} x_i\, y_i$$
   est un produit scalaire.
   On l'appelle le ***produit scalaire canonique*** \标准内积\ *sur $\mathbb{R}^n$.*
2. *Pour $E = \mathcal{C}^0([a,b], \mathbb{R})$ et $S : (f,g) \in E^2 \mapsto \int_0^1 f(t)g(t)dt \in \mathbb{R}$. S est une forme bilinéaire symétrique ainsi qu'un produit scalaire.*

REMARQUE 14 — *Pour $S : E \times E \to \mathbb{R}$ une forme bilinéaire symétrique telle que $S(x,x) \geq 0$, S vérifie l'inégalité de Cauchy-Schwarz.*
Le cas d'égalité est par contre faux si S n'est pas un produit scalaire (si $S(x,x) = 0$ n'implique pas $x = 0$).

REMARQUE 15 — *Pour vérifier qu'une forme bilinéaire symétrique S est un produit scalaire, il faut regarder si $S(x,x)$ est toujours positif ou nul, et si $S(x,x) = 0 \Leftrightarrow x = 0$.*

REMARQUE 16 — *Soient $n \geq 1$ et $A \in \mathscr{M}_n(\mathbb{R})$. Alors la forme bilinéaire $(X,Y) \mapsto {}^tXAY$ est un produit scalaire si et seulement si :*

- *A est une matrice symétrique ;*
- *${}^tXAX \geq 0$ pour tout vecteur colonne $X \in \mathbb{R}^n$;*
- *${}^tXAX = 0$ si et seulement si $X = 0$.*

Exemple 17 — *Soit $E = \mathbb{R}^n$. Soit $A \in \mathscr{M}_n(\mathbb{R})$ inversible. On pose $B = {}^tAA$.*
Alors la fonction $S : (X, Y) \in \mathbb{R}^n \times \mathbb{R}^n \mapsto {}^tXBY \in \mathbb{R}$ est un produit scalaire.
Cette fonction est bien une forme bilinéaire, comme vu précédemment.
La matrice B est symétrique, donc cette forme bilinéaire est symétrique.

Soit $X \in \mathbb{R}^n$. En posant $Z = AX$, on a $S(X, X) = {}^tX{}^tAAY = {}^tZZ$. Pour $\langle .|. \rangle$ le produit scalaire canonique sur $\mathbb{R}^n$, on a donc :

$$S(X, X) = \langle Z|Z\rangle = \langle AX|AX\rangle \geq 0.$$

La forme bilin. sym. S est donc positive.

Enfin, si l'on a $S(X, X) = 0$, alors on a $\langle AX|AX\rangle = 0$, donc $AX = 0$. Comme la matrice A est inversible, on obtient $X = 0$. Donc S est une forme bilin. sym. définie positive, c'est-à-dire un produit scalaire.

1.1.4 Norme euclidienne

Définition 18 (Rappel)
Soit E un $\mathbb{R}$-e.v. Soit $N : E \to \mathbb{R}_+$.
On dit que N est une **norme** \范数\ sur E si cette fonction vérifie les axiomes suivants :

- $N(x) = 0 \iff x = 0$ (axiome de séparation \分离性\) ;
- $N(\lambda \cdot x) = |\lambda|\, N(x)$, $\forall (\lambda, x) \in \mathbb{R} \times E$ (homogénéité \正齐次性\) ;
- $N(x + y) \leqslant N(x) + N(y)$, $\forall (x, y) \in E^2$ (inégalité triangulaire \三角不等式\).

Proposition-Définition 19
Soit E un $\mathbb{R}$-e.v. muni d'un produit scalaire $< .|. >$. Alors, la fonction :

$$\begin{array}{rccl} ||\cdot|| : E & \longrightarrow & \mathbb{R}_+ \\ x & \longmapsto & ||x|| = \sqrt{< x|x >} \end{array}$$

est une norme sur E.

Cette norme est appelée **norme euclidienne** \欧几里德范数\ associée au produit scalaire $< .|. >$.
La distance associée $d : (x, y) \mapsto \|x - y\|$ est appelée **distance euclidienne** \欧几里德距离\.
En utilisant cette norme, l'inégalité de Cauchy-Schwarz s'écrit :

$$\big|< x|y >\big| \leqslant ||x||\ ||y||\,.$$

Preuve —

- Cette fonction est bien définie sur E, puisque $\langle x|x\rangle \geqslant 0$, $\forall x \in E$, et elle est à valeurs dans $\mathbb{R}_+$.
- Soient $x \in E$ et $\lambda \in \mathbb{R}$. Par bilinéarité du produit scalaire, on a :
$$< \lambda x|\lambda x >= \lambda\ < x|\lambda x >= \lambda^2\ < x|x >,$$
et donc $||\lambda x|| = |\lambda|\ ||x||$.
- Pour $x \in E$, on a : $||x|| = 0 \iff < x|x >= 0 \iff x = 0$.
- Pour $(x, y) \in E^2$, on a :
$$\begin{aligned} ||x + y||^2 &=< x + y|x + y >=< x|x > + < y|y > +2 < x|y >\leqslant ||x||^2 + ||y||^2 + 2\,|< x|y >| \\ &\leqslant ||x||^2 + ||y||^2 + 2\,||x||\ ||y|| \qquad \text{d'après l'inégalité de Cauchy-Schwarz} \\ &\leq (||x|| + ||y||)^2, \end{aligned}$$
donc $\|\cdot\|$ est bien une norme sur E.

□

Exemples 20

1. *Pour $n \geq 1$, soit $\phi : (x, y) \in \mathbb{R}^n \times \mathbb{R}^n \mapsto \sum_{i=1}^n x_i\, y_i \in \mathbb{R}$ le produit scalaire canonique de $\mathbb{R}^n$. Alors la norme euclidienne associée est :*
$$||(x_1, x_2, \ldots, x_n)|| = \sqrt{\sum_{i=1}^n x_i^2}.$$
L'inégalité de Cauchy-Schwarz pour ce produit scalaire s'écrit : ◄
$$\left|\sum_{i=1}^n x_i y_i\right| \leqslant \left(\sum_{i=1}^n x_i^2\right)^{1/2} \left(\sum_{i=1}^n y_i^2\right)^{1/2}.$$

2. *Dans un espace vectoriel de dimension n muni d'une base $\mathcal{B}$, on peut définir un produit scalaire en posant :*
$$\langle x|y\rangle = \sum_{i=1}^{n} x_i\, y_i,$$
où x_1, x_2,..., x_n (resp. y_1, y_2,..., y_n) sont les composantes dans la base $\mathcal{B}$ du vecteur x (resp. y).
3. *Pour $E = \mathbb{C}^0\big([a,b]\big)$, la fonction $(f,g) \mapsto \int_a^b f(x)\, g(x)\, dx$ est un produit scalaire sur cet e.v.. L'inégalité de Cauchy-Schwarz correspondante est :* ◄
$$\left|\int_a^b f(x)\, g(x)\, dx\right| \leqslant \left(\int_a^b f^2(x)\, dx\right)^{1/2} \left(\int_a^b g^2(x)\, dx\right)^{1/2}.$$
4. *Soit $E = C^0(\mathbb{R}/2\pi\mathbb{Z})$ l'espace vectoriel des fonctions continues et 2π-périodiques sur $\mathbb{R}$. La fonction $(f,g) \mapsto \frac{1}{\pi}\int_0^{2\pi} f(x)\, g(x)\, dx$ est un produit scalaire sur cet e.v..*
5. *Pour $E = \mathbb{R}^3$, la forme bilinéaire symétrique S définie par :*
$$S\big((x,y,z),(x',y',z')\big) = x\,x' + y\,y' + z\,z' + \frac{1}{2}(x\,y' + x'\,y + x\,z' + x'\,z + y\,z' + y'\,z)$$
est un produit scalaire. En effet, on a : ◄
$$\begin{aligned} S\big((x,y,z),(x,y,z)\big) &= x^2 + y^2 + z^2 + x\,y + x\,z + y\,z \\ &= \frac{1}{2}\left((x+y)^2 + (y+z)^2 + (z+x)^2\right). \end{aligned}$$
Donc $S\big((x,y,z),(x,y,z)\big)$ est positif et ne peut être nul que si $x = y = z = 0$.

PROPOSITION 21
Soit E un $\mathbb{R}$-e.v. muni d'un produit scalaire. Soient $x, y \in E$.
Alors, on a $\|x+y\| = \|x\| + \|y\|$ si et seulement si $x = \lambda y$ ou $y = \lambda x$ pour un $\lambda \in \mathbb{R}_+$ (ssi x et y sont *positivement liés*).

Preuve — D'après la preuve de la Proposition-Définition 19, x et y vérifient le cas d'égalité dans l'inégalité triangulaire si et seulement si $2\langle x|y\rangle = 2\|x\|\|y\|$.
Ainsi, x et y vérifient le cas d'égalité dans l'inégalité de Cauchy-Schwarz. On a donc $x = \lambda y$ ou $y = \lambda x$ pour un $\lambda \in \mathbb{R}$.
Comme $\langle x|y\rangle = \|x\|\|y\| \geq 0$, on doit avoir $\lambda \geq 0$. □

REMARQUE 22 — *Pour E un $\mathbb{R}$-e.v. et $N : E \to \mathbb{R}_+$ une fonction. Si l'on peut trouver un produit scalaire $\langle .|.\rangle$ tel que $N(x)^2 = \langle x|x\rangle$, alors on aura montré que N est une norme.*

1.1.5 Espaces vectoriels euclidiens

DÉFINITION 23
Soit E un $\mathbb{R}$-e.v. muni d'un produit scalaire. On dit alors que E est un **espace vectoriel euclidien** \欧几里德空间\.

Si $(E, <.|.>)$ est un espace euclidien, il est donc naturellement muni de la norme euclidienne associée à son produit scalaire. \欧几里德空间或准希尔伯特空间可以利用内积定义一个范数\

EXEMPLE 24 — *Tous les produits scalaires considérés précédemment munissent leur espace vectoriel associé d'une structure d'espace euclidien.*

REMARQUE 25 — *La norme euclidienne $\|\cdot\|$ d'un espace vectoriel euclidien E est définie à partir de son produit scalaire $\langle .|.\rangle$.*

Nous allons montrer que la réciproque est vraie : Si l'on connaît toutes les valeurs de la norme euclidienne $\|\cdot\|$, alors on peut retrouver les valeurs du produit scalaire $\langle .|.\rangle$. Pour cela, nous aurons besoin des égalités suivantes.

PROPOSITION 26
Soit E un espace vectoriel euclidien. Soit $(x,y) \in E^2$. On a :

- Identités de polarisation \极化恒等式\ :
  1. $||x+y||^2 = ||x||^2 + ||y||^2 + 2 < x|y >$;
  2. $||x-y||^2 = ||x||^2 + ||y||^2 - 2 < x|y >$;
  3. $||x+y||^2 - ||x-y||^2 = 4 < x|y >$.
- Identité du parallélogramme \平行四边形法则\ :
  $||x+y||^2 + ||x-y||^2 = 2\left(||x||^2 + ||y||^2\right)$ [1].

Preuve —

- On a : $||x+y||^2 =< x+y|x+y >=< x|x > + < x|y > + < y|x > + < y|y >=< x|x > +2 < x|y > + < y|y >$.
- Appliquer l'égalité précédente à x et $-y$ pour développer $||x+y||^2$.
- Les deux dernières égalités se déduisent des précédentes par somme et différence.

□

Remarque 27 — *Pour $\|.\|$ une norme issue d'un produit scalaire, les identités de polarisation nous disent alors que*

$$\langle x|y\rangle = \frac{||x+y||^2 - ||x-y||^2}{4}.$$

On peut donc bien déterminer les valeurs de $\langle .|.\rangle$ en fonction des valeurs de $\|.\|$.
L'identité du parallélogramme est une identité que vérifient toutes les normes euclidiennes. Mais certaines normes ne sont pas euclidiennes. Une façon qui permet de le montrer est de trouver x et y qui ne vérifient pas l'identité du parallélogramme.

Exemples 28

1. *Sur $\mathbb{R}^2$, on définit $N((x,y)) = \sqrt{x^2+2xy+3y^2}$. Pour montrer que définit une norme euclidienne sur $\mathbb{R}^2$, on commence par vérifier que :*
$$x^2+2xy+3y^2 = (x+y)^2 + 2y^2$$
est positif et ne peut être nul que si $(x,y) = (0,0)$.
Il faut alors exhiber le produit scalaire dont N provient. D'après la dernière identité de polarisation, il doit être égal à :
$$\begin{aligned} S\big((x,y),(x',y')\big) &= \frac{N((x+x',y+y'))^2 - N((x-x',y-y'))^2}{4} \\ &= x\,x' + x\,y' + y\,x' + 3y\,y'. \end{aligned}$$
On remarque que la fonction S est bien une forme bilinéaire symétrique S, et que l'on a :
$$\forall (x,y) \in \mathbb{R}^2\,,\ S\big((x,y),(x,y)\big) = N((x,y))^2,$$
Ainsi, S est bien un produit scalaire, donc N est bien une norme euclidienne (et donc une norme).
2. *Sur $\mathbb{R}^2$, on définit la norme « infinie »* \无穷范数\ *par :*
$$||(x,y)||_\infty = \max\big(|x|,|y|\big)$$
Cette fonction est bien une norme (voir Analyse 4) mais ce n'est pas une norme euclidienne.
En effet, pour $u = (2,1)$ et $v = (1,2)$, on a :
$$||u+v|| = 3 \qquad ||u-v|| = 1 \qquad ||u|| = ||v|| = 2$$
et donc : $||u+v||^2 + ||u-v||^2 = 10 \neq 8 = 2\left(||u||^2 + ||v||^2\right)$.
3. *Pour $E = \mathbb{R}^n$, on définit la norme « ℓ^1 » par :*
$$\|x\|_{\ell^1} = \sum_{i=1}^{n} |x_i|.$$
Cette fonction est bien une norme (voir Analyse 4), mais ce n'est pas une norme euclidienne. (Le montrer)

1. La somme des carrés des quatre côtés est égale à la somme des carrés des deux diagonales.

Le produit scalaire permet par exemple de définir la notion d'angle entre deux vecteurs :

Corollaire 29 (Lien entre produit scalaire et angles en géométrie)
Soit E un e.v. euclidien. Soient $x, y \in E$ non-nuls. Alors :

$$\exists ! \theta \in [0, \pi], \text{ tel que } \langle x|y \rangle = \|x\| \cdot \|y\| \cdot \cos\theta.$$

Preuve — D'après l'inégalité de Cauchy-Schwarz, on a $|\langle x|y\rangle| \leqslant \|x\| \cdot \|y\|$. Si les vecteurs x et y ne sont pas nuls, on en déduit que $\dfrac{|\langle x|y\rangle|}{\|x\| \cdot \|y\|} \leqslant 1$. D'où $-1 \leqslant \dfrac{\langle x|y\rangle}{\|x\| \cdot \|y\|} \leqslant +1$, donc $\exists ! \theta \in [0, \pi]$, tel que $\dfrac{\langle x|y\rangle}{\|x\| \cdot \|y\|} = \cos\theta$. □

Chapitre 2 Notion de déterminant

Ce chapitre est constitué de trois étapes :

1. Nous commencerons par rappeler les propriétés fondamentales du groupe symétrique $\mathcal{S}_n$, notamment les propriétés de la *signature* d'une permutation.
2. Avec le groupe symétrique, nous définissons les *formes multi-linéaires alternées*.
3. Avec les formes multi-linéaires alternées, nous définissons le *déterminant*. Nous étudions alors les nombreuses propriétés de cette fonction.

Table des matières du chapitre

2.1 Rappels sur le groupe symétrique $\mathcal{S}_n$

Définition 1
Soit $n \in \mathbb{N}^*$. L'ensemble $\mathrm{Bij}([\![1, n]\!])$ des fonctions bijectives de $\{1, \dots, n\}$ dans $1, \dots, n\}$ est un groupe pour la composition de fonctions $\circ$. On le note $\mathcal{S}_n$.
On appelle le groupe $(\mathcal{S}_n, \circ)$ le **groupe symétrique d'ordre** n \ n阶对称群\.
Un élément de $\mathcal{S}_n$ est appelé une **permutation** \置换\.

Décrire une permutation $\sigma : \{1, 2, \dots, n\} \longrightarrow \{1, 2, \dots, n\}$ revient à donner les images de chaque i par σ, pour $1 \leq i \leq n$. On peut ainsi noter σ par :

$$\begin{bmatrix} 1 & 2 & \cdots & n \\ \sigma(1) & \sigma(2) & \cdots & \sigma(n) \end{bmatrix}.$$

Par exemple, pour :

$$\begin{bmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 3 & 7 & 5 & 4 & 6 & 1 & 2 \end{bmatrix} \sigma$$

on a $\sigma(1) = 3$, $\sigma(2) = 7$, $\sigma(3) = 5$, $\sigma(4) = 4$, $\sigma(5) = 6$, $\sigma(6) = 1$, $\sigma(7) = 2$.

Définition 2 (Transposition)
Soit $n \geq 2$. Soient $i, j \in [\![1, n]\!]$, avec $i \neq j$. On définit la fonction $\tau : [\![1, n]\!] \to [\![1, n]\!]$ par $\tau(i) = j$, $\tau(j) = i$, et $\tau(k) = k$ pour tout $k \neq i, j$.
La fonction τ est alors une permutation de $\mathcal{S}_n$. On la note $\tau = (i, j)$.
On appelle cette permutation une **transposition** \对换\.

La transposition (i, j) intervertit i et j et ne change pas les autres élément de $[\![1, n]\!]$.

Définition 3
Soit $n \geq 2$. Soit $2 \leq p \leq n$ un entier. Soient $a_1, \ldots, a_p \in [\![1, n]\!]$ des éléments distincts. On définit la fonction $\sigma : [\![1, n]\!] \to [\![1, n]\!]$ par :

$$\begin{aligned}\sigma(x) &= x, \, \forall x \notin \{a_1, a_2, \ldots, a_p\} \, ; \\ \sigma(a_i) &= a_{i+1}, \, \forall i \in [\![1, p-1]\!] \, ; \\ \sigma(a_p) &= a_1.\end{aligned}$$

La fonction σ est alors une permutation de $\mathcal{S}_n$. On la note $\sigma = (a_1, a_2, \ldots, a_p)$.
On appelle cette permutation un **p-cycle** ou **cycle d'ordre** p \\p-轮换\\.

Théorème 4
Soit $n \geq 2$.
Alors toute permutation de $\mathcal{S}_n$ peut s'écrire comme un produit de transpositions.

Exemple 5 — *Pour $a_1, \ldots, a_p \in [\![1, n]\!]$ distincts, le p-cycle $\gamma = (a_1, a_2, \ldots, a_p)$ se décompose en :*

$$\gamma = (a_1, a_2)(a_2, a_3) \ldots (a_{p-1}, a_p).$$

Remarque 6 — *Attention : Comme l'opération entre permutations est une composée de fonctions, pour calculer l'image de $k \in [\![1, n]\!]$ par une tel produit de transpositions, il faut commencer par la permutation de droite. ($\sigma\tau(k) = \sigma(\tau(k))$)*

Définition 7
Soient $n \geq 2$ et $\sigma \in \mathcal{S}_n$. On définit le **support** de σ comme l'ensemble des entiers k tels que $\sigma(k) \neq k$.

Remarque 8 —

1. *Le support d'un p-cycle $(a_1, \ldots, a_p)$ est l'ensemble $\{a_1, \ldots, a_p\}$.*
2. *Soient σ, τ deux permutations avec des supports disjoints. Alors σ et τ commutent.* ◄

Théorème 9 (Décomposition en produit de cycles à support disjoint)
Soit $n \geq 2$. Soit $\sigma \in \mathcal{S}_n$.
Alors il existe $\sigma_1, \sigma_2, \ldots, \sigma_p$ des cycles dont les supports sont disjoints, tels que :

$$\sigma = \sigma_1 \circ \ldots \circ \sigma_p.$$

On dit que toute permutation de $\mathcal{S}_n$ se décompose en produit de cycles à supports disjoints.
De plus cette décomposition est unique à l'ordre près. \\任何一个非单位元置换都可以表示成一群两两不相交的轮换的乘积，并且除了轮换的排列次序外表示是唯一的。\\

Exemple 10 — *La décomposition de $\sigma = \left[\begin{smallmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 5 & 2 & 1 & 8 & 3 & 7 & 6 & 4 \end{smallmatrix}\right]$ en produit de cycles à supports disjoints est $\sigma = (1\ 5\ 3) \circ (4\ 8) \circ (6\ 7)$.*
Pour l'unicité il faut comprendre : unique à l'écriture de chaque cycle près (exemple : $(1\ 5\ 3) = (5\ 3\ 1)$) et à l'ordre près (exemple : $(4\ 8) \circ (1\ 5\ 3) = (1\ 5\ 3) \circ (4\ 8)$).

Définition 11 (Signature d'une permutation)
Soient $n \geq 2$ et $\sigma \in \mathcal{S}_n$. On définit la **signature** \\符号\\ de σ comme le nombre

$$\varepsilon(\sigma) = \Pi_{1 \leq i < j \neq 1} \frac{\sigma(j) - \sigma(i)}{j - i}.$$

Proposition 12
La signature d'une transposition est égale à -1.

Théorème 13
Soit $n \geq 2$. Soient $\sigma, \tau \in \mathcal{S}_n$.Alors, on a :

$$\varepsilon(\sigma\,\tau) = \varepsilon(\sigma)\,\varepsilon(\tau).$$

La fonction $\varepsilon : \mathcal{S}_n \to \{-1, 1\}$ est un morphisme de groupes de $\mathcal{S}_n$ vers $\big(\{-1, 1\}, \times\big)$.

Définition 14
Soient $n \geq 2$ et $\sigma \in \mathcal{S}_n$.
On dit que la permutation σ est **paire** \\偶排列\\ si sa signature vaut 1.

On dit que la permutation σ est **impaire** \奇排列\ si sa signature vaut -1.

EXEMPLE 15 — *Soit $p \geq 2$. La décomposition $(a_1, a_2, \ldots, a_p) = (a_1, a_2)(a_2, a_3)\ldots(a_{p-1}, a_p)$ prouve qu'un p-cycle a pour signature $(-1)^{p-1}$.*

COROLLAIRE 16 (Calcul de $\varepsilon(\sigma)$)
Soit $n \geq 2$. Soit $\sigma \in \mathcal{S}_n$. Soit $\sigma = \sigma_1 \sigma_2 \ldots \sigma_p$ la décomposition de σ en produit de cycles à support disjoint. Soient $\alpha_1, \ldots, \alpha_p$ les longueurs de ces cycles.
Alors, on a $\varepsilon(\sigma) = (-1)^{\alpha_1+\ldots+\alpha_p-p}$.

DÉFINITION 17 (Groupe alterné)
Soit $n \geq 1$. On pose $\mathcal{A}_n$ l'ensemble des permutations de $\mathcal{S}_n$ de signature 1 (permutations paires).
On appelle $\mathcal{A}_n$ le **groupe alterné d'ordre** n \\$n$元交错群\.

REMARQUE 18 — *$\mathcal{A}_n$ est un sous-groupe de $\mathcal{S}_n$, car cet ensemble est le noyau de la signature (qui est un morphisme de groupes).*

PROPOSITION 19
Soit $n \geq 2$. Soit $\tau \in \mathcal{S}_n$ une permutation impaire. Alors, l'ensemble des permutations impaires est égal à $\mathcal{A}_n \tau = \{\sigma\tau, \sigma \in \mathcal{A}_n\}$.

2.2 FORMES p-LINÉAIRES

Dans le reste du chapitre, $\mathbb{K}$ désignera un corps (ex : $\mathbb{R}$,$\mathbb{Q}$,$\mathbb{C}$).

2.2.1 Définition

DÉFINITION 20
Soit E un $\mathbb{K}$-espace vectoriel. Soit $p \geq 1$ un entier. Soit $f : E^p \to \mathbb{K}$ une fonction. On dit que la fonction f est une **forme p-linéaire** si, pour toute famille $(a_1, a_2, \ldots, a_p) \in E^p$ et pour tout $i \in [\![1, p]\!]$, la i-ème fonction :

$$\begin{array}{rccl} f_i : & E & \longrightarrow & \mathbb{K} \\ & x & \longmapsto & f(a_1, a_2, \ldots, a_{i-1}, x, a_{i+1}, \ldots, a_p) \end{array}$$

est une application linéaire.

EXEMPLES 21

1. *Les formes 1-linéaires sont les formes linéaires.*
2. *Les formes bilinéaires sont les formes 2-linéaires.*
3. *Dans un espace euclidien, le produit scalaire est une forme bilinéaire.*
4. *La fonction :*

$$\begin{array}{rccl} \phi : & \mathbb{C}^0\big([0,1], \mathbb{R}\big)^2 & \longrightarrow & \mathbb{R} \\ & (u, v) & \longmapsto & \int_0^1 u(t)\, v(t)\, dt \end{array}$$

   est ainsi une forme bilinéaire sur $E = \mathbb{C}^0\big([0,1], \mathbb{R}\big)$.

5. *La fonction :*

$$\begin{array}{rccl} \psi : & \mathbb{C}^p & \longrightarrow & \mathbb{C} \\ & (z_1, z_2, \ldots, z_p) & \longmapsto & z_1\, z_2 \ldots z_p \end{array}$$

   est une forme p-linéaire sur $\mathbb{C}$.

Autrement dit, une forme p-linéaire est une fonction $f : (u_1, \ldots, u_p) \in E^p \to \mathbb{K}$ qui est *linéaire par rapport à chaque variable* $u_1, \ldots, u_p$ (linéaire en u_1, linéaire en u_2, …, linéaire en u_p).

2.2.2 Expression d'une forme p-linéaire en dimension finie

Dans la suite de ce chapitre, nous ne nous intéresserons qu'aux formes p-linéaires sur des $\mathbb{K}$-e.v. de dimension finie.

REMARQUE 22 (Cas des formes bilinéaires ($p = 2$)) — *Soit E un $\mathbb{K}$-e.v. de dimension n. Soit $\mathcal{B} = (e_1, \dots, e_n)$ une base de E. Soit f une forme bilinéaire sur E. Soient $u, v \in E$ tels que :*

$$u = \sum_{i=1}^{n} a_i\, e_i \quad \text{et} \quad v = \sum_{i=1}^{n} b_i\, e_i.$$

Alors, on a :

$$f(u,v) = f\left(\sum_{k=1}^{n} a_k\, e_k, \sum_{l=1}^{n} b_l\, e_l\right) = \sum_{k=1}^{n} a_k f\left(e_k, \sum_{l=1}^{n} b_l\, e_l\right) = \sum_{k=1}^{n} a_k \sum_{l=1}^{n} b_l\, f(e_k, e_l) = \sum_{k,l=1}^{n} a_k\, b_l\, f(e_k, e_l).$$

Réciproquement, si $(y_{i,j})_{(i,j)\in[\![1,n]\!]^2}$ est une famille d'éléments de $\mathbb{K}$, la fonction : ◀

$$\left(\sum_{i=1}^{n} a_i\, e_i, \sum_{i=1}^{n} b_i\, e_i\right) \longmapsto \sum_{(k,l)\in[\![1,n]\!]^2} a_k\, b_l\, y_{k,l}$$

est une forme bilinéaire sur E.

REMARQUE 23 — *Une forme bilinéaire f peut s'exprimer autrement, comme nous l'avons vu au chapitre précédent. On pose $A = (a_{ij})_{i,j} \in \mathscr{M}_n(\mathbb{K})$ la matrice telle que $a_{i,j} = f(e_i, e_j)$. Pour U, V les vecteurs colonnes associés à u, v :*

$$U = {}^t\!\begin{pmatrix} a_1 & a_2 & \cdots & a_n \end{pmatrix} \quad \text{et} \quad V = {}^t\!\begin{pmatrix} b_1 & b_2 & \cdots & b_n \end{pmatrix},$$

on a alors :

$$f(u,v) = \sum_{k,l=1}^{n} a_k\, b_l\, f(e_k, e_l) = {}^t\!UAV.$$

La matrice A est la matrice de la forme bilinéaire f dans la base $\mathcal{B}$.

§ 1. Applications trilinéaires ($p = 3$) —

REMARQUE 24 (Cas des formes trilinéaires ($p = 3$)) — *Soit E un $\mathbb{K}$-e.v. de dimension n. Soit $\mathcal{B} = (e_1, \dots, e_n)$ une base de E. Soit f une forme 3-linéaire (ou trilinéaire) sur E. Soient $u_1, u_2, u_3 \in E$ tels que :*

$$u_j = \sum_{i=1}^{n} a_{i,j}\, e_i, \ \forall j \in [\![1,3]\!].$$

Alors, on a :

$$\begin{aligned}
f(u_1,u_2,u_3) &= f\left(\sum_{k=1}^{n} a_{k,1}\, e_k, \sum_{l=1}^{n} a_{l,2}\, e_l, \sum_{m=1}^{n} a_{m,3}\, e_m\right) = \sum_{k=1}^{n} a_{k,1}\, f\left(e_k, \sum_{l=1}^{n} a_{l,2}\, e_l, \sum_{m=1}^{n} a_{m,3}\, e_m\right) \\
&= \sum_{k=1}^{n} a_{k,1} \left(\sum_{l=1}^{n} a_{l,2}\, f\Big(e_k, e_l, \sum_{m=1}^{n} a_{m,3}\, e_m\Big)\right) = \sum_{k=1}^{n} a_{k,1} \sum_{l=1}^{n} a_{l,2} \Big(\sum_{m=1}^{n} a_{m,3}\, f(e_k, e_l, e_m)\Big) \\
&= \sum_{k=1}^{n}\sum_{l=1}^{n}\sum_{m=1}^{n} a_{k,1}\, a_{l,2}\, a_{m,3}\, f(e_k, e_l, e_m).
\end{aligned}$$

Réciproquement, si $(y_{i,j,k})_{(i,j,k)\in[\![1,n]\!]^3}$ est une famille d'éléments de $\mathbb{K}$, la fonction : ◀

$$\left(\sum_{i=1}^{n} a_{i,1}\, e_i, \sum_{i=1}^{n} a_{i,2}\, e_i, \sum_{i=1}^{n} a_{i,3}\, e_i\right) \longmapsto \sum_{(k,l,m)\in[\![1,n]\!]^3} a_{k,1}\, a_{l,2}\, a_{m,3}\, y_{k,l,m}$$

est une application trilinéaire sur E.

Le cas général

Proposition 25
Soient E un $\mathbb{K}$-e.v. de dimension n et $\mathcal{B}=(e_1,\ldots,e_n)$ une base de E. Soit $p\geq 1$. Soit f une forme p-linéaire (ou trilinéaire) sur E.
Soient $u_1,\ldots,u_p\in E$ tels que :

$$u_j=\sum_{i=1}^{n} a_{i,j}\,e_i,\ \forall j\in [\![1,p]\!].$$

Alors, on a :

$$f(u_1,u_2,\ldots,u_p)=\sum_{1\leqslant i_1\leqslant n,1\leqslant i_2\leqslant n,\ldots,1\leqslant i_p\leqslant n} a_{i_1,1}\,a_{i_2,2}\ldots a_{i_p,p}\,f(e_{i_1},e_{i_2},\ldots,e_{i_p}).$$

Preuve — Vérifions cette formule par récurrence sur $p\geq 1$.

- Soit f une forme 1-linéaire sur E, c'est-à-dire une forme linéaire sur E.
  Si $u_1=\sum\limits_{i=1}^{n} a_{i,1}\,e_i$, on a $f(u_1)=\sum\limits_{i=1}^{n} a_{i,1}\,f(e_i)$, ce qui démontre la relation pour $p=1$.
- Supposons que la relation est vérifiée au rang $p-1$, avec $p\geqslant 2$. Pour f une forme p-linéaire f sur E et u_1,u_2,…, u_p tels que :

  $$\forall j\in [\![1,p]\!]\,,\ u_j=\sum_{i=1}^{n} a_{i,j}\,e_i$$

  la linéarité de l'application $x\mapsto f(u_1,u_2,\ldots,u_{p-1},x)$ donne :

  $$f(u_1,u_2,\ldots,u_p)=\sum_{i_p=1}^{n} a_{i_p,p}\,f(u_1,u_2,\ldots,u_{p-1},e_{i_p}).\qquad (*)$$

  Pour tout entier $i_p\in [\![1,n]\!]$, l'application :

  $$(x_1,x_2,\ldots,x_{p-1})\mapsto f(x_1,x_2,\ldots,x_{p-1},e_{i_p})$$

  est une forme $(p-1)$-linéaire. L'hypothèse de récurrence donne alors :

  $$f(u_1,u_2,\ldots,u_{p-1},e_{i_p})=\sum_{1\leqslant i_1\leqslant n,1\leqslant i_2\leqslant n,\ldots,1\leqslant i_{p-1}\leqslant n} a_{i_1,1}\,a_{i_2,2}\ldots a_{i_{p-1},p-1}\,f(e_{i_1},e_{i_2},\ldots,e_{i_p}).$$

  En remplaçant les termes dans l'égalité $(*)$, on obtient :

  $$f(u_1,u_2,\ldots,u_p)=\sum_{i_p=1}^{n} a_{i_p,p}\left(\sum_{1\leqslant i_1\leqslant n,1\leqslant i_2\leqslant n,\ldots,1\leqslant i_{p-1}\leqslant n} a_{i_1,1}\,a_{i_2,2}\ldots a_{i_{p-1},p-1}\,f(e_{i_1},e_{i_2},\ldots,e_{i_p})\right),$$

  ce qui prouve la relation au rang p.

□

Réciproquement, si $(y_\sigma)_{\sigma\in[\![1,n]\!]^p}$ est une famille d'éléments de $\mathbb{K}$, la fonction :

$$\left(\sum_{i=1}^{n} a_{i,1}\,e_i,\sum_{i=1}^{n} a_{i,2}\,e_i,\ldots,\sum_{i=1}^{n} a_{i,p}\,e_i\right)\longmapsto \sum_{\sigma\in[\![1,n]\!]^p} a_{\sigma_1,1}\,a_{\sigma_2,2}\cdots a_{\sigma_p,p}\,y_\sigma$$

est une forme p-linéaire sur E.

L'espace vectoriel des formes p-linéaires

L'ensemble des formes p-linéaires de E dans un $\mathbb{K}$-espace vectoriel F est non vide et stable par combinaison linéaire. C'est donc un sous-espace vectoriel de l'espace vectoriel $\mathcal{F}(E^p,\mathbb{K})$. On note cet espace $\mathcal{L}_p(E,\mathbb{K})$ ou $\Lambda^p(E)$.
Pour calculer la dimension de $\Lambda^p(E)$, on utilise les isomorphismes.

Construisons un **isomorphisme** : Soit $(e_1,\ldots,e_n)$ une base de E,. Pour $1\leq i\leq n$, on définit $e_i^*:E\to\mathbb{K}$ l'application linéaire telle que $e_i^*(e_j)=0$ si $j\neq i$ et $e_i^*(e_i)=1$.
Pour $x=x_1e_1+\ldots+x_ne_n$ on a donc $e_i^*(x)=x_i$.

Soit $(i_1,\ldots,i_p)$ une liste de p éléments de $[\![1,n]\!]$. On définit la forme p-linéaire sur E :

$$\begin{array}{rcl} u_{(i_1,\ldots,i_p)}\,:\,E^p & \longrightarrow & \mathbb{K}\\ (x_1,\ldots,x_p) & \longmapsto & e_{i_1}^*(x_1)\cdots e_{i_p}^*(x_p).\end{array}$$

D'après la partie précédente, cette famille de formes p-linéaires est une famille génératrice de $\Lambda^p(E)$.

On remarque de plus que pour $j_1, \ldots, j_p \in \{1, \ldots, n\}$, on a $u_{(i_1,\ldots,i_p)}(e_{j_1}, \ldots, e_{j_p}) = e^*_{i_1}(e_{j_1}) \cdots e^*_{i_p}(e_{j_p}) = 1$ si $j_1 = i_1$,$j_2 = i_2$,...,$j_p = i_p$ et 0 sinon. On peut alors montrer que la famille des formes p-linéaires $u_{(i_1,\ldots,i_p)}$ est libre. Cette famille est donc une base de $\Lambda^p(E)$, ce qui donne : ◀

$$\boxed{\dim(\mathcal{L}_p(E, \mathbb{K})) = \dim(E)^p.}$$

2.2.3 Formes p-linéaires alternées

Définition 26
Soient E un $\mathbb{K}$-e.v. et $p \geq 2$. Soit $f : E^p \to \mathbb{K}$ une forme p-linéaire sur E. On dit que f est une forme p-linéaire **alternée** \交替多重线性映射\ si, pour tout $(u_1, u_2, \ldots, u_p) \in E^p$ et pour tout $i \neq j$, on a :

$$u_i = u_j \Longrightarrow f(u_1, u_2, \ldots, u_p) = 0.$$

Définition 27
Soient E un $\mathbb{K}$-e.v. et $p \geq 2$. Soit $f : E^p \to \mathbb{K}$ une forme p-linéaire sur E. On dit que f est une forme p-linéaire **antisymétrique** \反对称\ si, pour tout $(u_1, u_2, \ldots, u_p) \in E^p$ et pour tout $i < j$, on a :

$$f(u_1, \ldots, \underset{\substack{\uparrow \\ i\text{-ième}}}{u_i}, \ldots \underset{\substack{\uparrow \\ j\text{-ième}}}{u_j}, \ldots, u_p) = -f(u_1, \ldots, \underset{\substack{\uparrow \\ i\text{-ième}}}{u_j}, \ldots \underset{\substack{\uparrow \\ j\text{-ième}}}{u_i}, \ldots, u_p).$$

Proposition 28
Soient E un $\mathbb{K}$-e.v., $p \geq 2$, et f une forme p-linéaire sur E.
Si f est alternée, alors elle est antisymétrique.
Si est $\mathbb{K}$ un corps tel que $1 + 1 \neq 0$ et si f est antisymétrique, alors f est alternée.

Preuve —

- "$\Rightarrow$" Soient $u_1, \ldots, u_p \in E$. Soient i, j tels que $i < j$. Alors, la fonction :

$$\begin{array}{rcl} g : E^2 & \longrightarrow & F \\ (x, y) & \longmapsto & f(u_1, \ldots, \underset{\substack{\uparrow \\ i\text{-ième}}}{x}, \ldots \underset{\substack{\uparrow \\ j\text{-ième}}}{y}, \ldots, u_p) \end{array}$$

  est une forme bilinéaire sur E, avec :

$$g(x, x) = 0, \forall x \in E.$$

  Soient $(x, y) \in E^2$. Comme g est bilinéaire, on a :

$$\begin{aligned} 0 = g(x + y, x + y) &= g(x, x + y) + g(y, x + y) = g(x, x) + g(x, y) + g(y, x) + g(y, y) \\ &= 0 + g(x, y) + g(y, x) + 0. \end{aligned}$$

  On a donc :

$$g(x, y) = -g(y, x), \forall (x, y) \in E^2.$$

- "$\Leftarrow$" Soit f une forme bilinéaire antisymétrique. Soient i, j avec $i < j$. On a alors par antisymétrie :

$$f(u_1, \ldots, \underset{\substack{\uparrow \\ u_i}}{x}, \ldots \underset{\substack{\uparrow \\ u_j}}{x}, \ldots, u_p) = -f(u_1, \ldots, \underset{\substack{\uparrow \\ u_j}}{x}, \ldots \underset{\substack{\uparrow \\ u_j}}{x}, \ldots, u_p),$$

  cela donne : $2.f(u_1, \ldots, x, \ldots, x, \ldots, u_p) = 0$.
  Comme $1 + 1 = 2 \neq 0$ dans $\mathbb{K}$, on obtient alors $f(u_1, \ldots, x, \ldots, x, \ldots, u_p) = 0$, ce qui conclut.

□

Exemples 29

1. *La forme nulle $f : (u_1, \ldots, u_p) \in E^p \mapsto 0 \in \mathbb{K}$ est une forme p-linéaire alternée.*
2. *La fonction*

$$\begin{array}{rcl} f : (\mathbb{K}^2)^2 & \longrightarrow & \mathbb{K} \\ ((x_1, x_2), (y_1, y_2)) & \longmapsto & x_1 y_2 - x_2 y_1 \end{array}$$

   est une forme bilinéaire alternée. Par contre, $(x, y) \mapsto x_1 y_2 + x_2 y_1$ n'est pas alternée. ◀
3. *Pour E un $\mathbb{R}$-e.v. euclidien, le produit scalaire sur E est une forme bilinéaire qui n'est pas alternée.*
4. *Si $1 + 1 \neq 0$ dans $\mathbb{K}$, la seule forme bilinéaire symétrique qui est aussi alternée est la forme nulle : $f : (x, y) \in E^2 \mapsto 0$.* ◀

5. *Si $1+1=0$ dans $\mathbb{K}$ (par exemple dans $\mathbb{Z}/2\mathbb{Z}$), alors on a $1=-1$, donc les formes p-linéaires antisymétriques sont exactement les formes p-linéaires symétriques.*

Remarque 30 —

- *Lorsque $\mathbb{K}$ est un corps dans lequel $1+1=0$, comme par exemple $\mathbb{Z}/2\mathbb{Z}$, une forme p-linéaire antisymétrique n'est pas forcément alternée.*
  La preuve de l'implication repose sur une division par 2. Or, dans un tel corps $\mathbb{K}$, le nombre 2 ne possède pas d'inverse (et diviser par 2 n'a alors pas de sens).
  Cela veut dire que les notions de forme alternée et de forme antisymétrique ne sont pas toujours égales, même si dans les corps usuels $\mathbb{Q}, \mathbb{R}, \mathbb{C}$ (et dans la majorité des $\mathbb{Z}/p\mathbb{Z}$) il y a équivalence.
- *Dans le reste de ce chapitre, nous allons nous intéresser à des formes p-linéaires alternées. Donc, peu importe le corps $\mathbb{K}$, ces formes seront aussi antisymétriques.*

Proposition 31
Soient E un $\mathbb{K}$-e.v., $p \geq 2$, et f une forme p-linéaire alternée sur E. Soit $(u_1, u_2, \ldots, u_p)$ une famille de vecteurs de E qui est liée. Alors, on a $f(u_1, u_2, \ldots, u_p) = 0$.

Preuve — Comme la famille $(u_1, u_2, \ldots, u_p)$ est liée, il existe i tel que u_i soit combinaison linéaire $\sum_{k\neq i} \lambda_k u_k$ des autres vecteurs. On a donc :

$$f(u_1, u_2, \ldots, u_p) = f(u_1, \ldots, u_{i-1}, \sum_{k\neq i} \lambda_k u_k, u_{i+1}, \ldots, u_p) = \sum_{k\neq i} \lambda_k f(u_1, \ldots, u_{i-1}, u_k, u_{i+1}, \ldots u_p)$$

qui est nul parce que chacune des familles $(u_1, \ldots, u_{i-1}, u_k, u_{i+1}, \ldots, u_p)$, pour $k \neq i$, contient deux vecteurs égaux et que f est alternée. □

Corollaire 32
Soient une forme p-linéaire alternée f sur E et $(u_1, u_2, \ldots, u_p) \in E^p$. Alors le nombre $f(u_1, u_2, \ldots, u_p) \in \mathbb{K}$ est inchangé si l'on ajoute à l'un des u_j une combinaison linaire des vecteurs u_i, $i \neq j$.

Preuve — Si x est une combinaison linaire de la famille $(u_1, \ldots, u_{j-1}, u_{j+1}, \ldots, u_p)$ alors, d'après la proposition précédente :

$$f(u_1, \ldots, u_{j-1}, x, u_{j+1}, \ldots, u_p) = 0,$$

et, par la p-linéarité de f, on a :

$$\begin{aligned} f(u_1, \ldots, u_{j-1}, u_j + x, u_{j+1}, \ldots, u_p) &= f(u_1, \ldots, u_{j-1}, u_j, u_{j+1}, \ldots, u_p) + f(u_1, \ldots, u_{j-1}, x, u_{j+1}, \ldots, u_p) \\ &= f(u_1, \ldots, u_{j-1}, u_j, u_{j+1}, \ldots, u_p). \end{aligned}$$

□

Formes p-linéaires alternées et permutation

Proposition 33
Soient E un $\mathbb{K}$-e.v. et $f : E^n \to \mathbb{K}$ une forme n-linéaire alternée sur E. Soit $\sigma \in \mathcal{S}_n$ une permutation. Alors, on a :

$$f(u_{\sigma(1)}, u_{\sigma(2)}, \ldots, u_{\sigma(n)}) = \varepsilon(\sigma)\, f(u_1, u_2, \ldots, u_n),\ \forall (u_1, \ldots, u_n) \in E^n.$$

Preuve — Soit $\tau \in \mathcal{S}_n$ une transposition. Soit $(u_1, \ldots, u_n) \in E^n$. Comme f est une forme n-linéaire alternée, la proposition 28 donne :

$$f(u_{\tau(1)}, u_{\tau(2)}, \ldots, u_{\tau(n)}) = -f(u_1, u_2, \ldots, u_n).$$

On définit la fonction g_τ par :

$$\begin{array}{rcl} g_\tau : E^n & \longrightarrow & E^n \\ (u_1, u_2, \ldots, u_n) & \longmapsto & (u_{\tau(1)}, u_{\tau(2)}, \ldots, u_{\tau(n)}). \end{array}$$

On a alors :

$$f(u_{\tau(1)}, u_{\tau(2)}, \ldots, u_{\tau(n)}) = \varepsilon(\tau)\, f(u_1, u_2, \ldots, u_n) = f \circ g_\tau(u_1, u_2, \ldots, u_n).$$

Donc $f \circ g_\tau = \varepsilon(\tau)\, f$.

Soit $\sigma \in \mathcal{S}_n$. Alors σ s'écrit comme un produit de transpositions $\sigma = \tau_1 \tau_2 \cdots \tau_{p-1} \tau_p$. Comme la signature est un morphisme de groupes, on a donc :

$$\begin{array}{rcl} f(u_{\sigma(1)}, u_{\sigma(2)}, \ldots, u_{\sigma(n)}) & = & f \circ g_\sigma(u_1, u_2, \ldots, u_n) \\ & = & f \circ g_{\tau_1} \circ g_{\tau_2} \circ \cdots \circ g_{\tau_{p-1}} \circ g_{\tau_p}(u_1, u_2, \ldots, u_n) \\ & = & \varepsilon(\tau_p)\, f \circ g_{\tau_1} \circ g_{\tau_2} \circ \cdots \circ g_{\tau_{p-1}}(u_1, u_2, \ldots, u_n) \\ & & \cdots \\ & = & \varepsilon(\tau_1)\, \varepsilon(\tau_2) \cdots \varepsilon(\tau_{p-1}) \varepsilon(\tau_p)\, f(u_1, u_2, \ldots, u_n) \\ & = & \varepsilon(\sigma)\, f(u_1, u_2, \ldots, u_n). \end{array}$$

□

Remarque 34 — *Avec cela, on déduit l'écriture d'une forme n-linéaire alternée dans une base. Soit $\mathcal{B} = (e_1, e_2, \ldots, e_n)$ une base de E. Soit ϕ une forme n-linéaire alternée sur E. Soient u_1, u_2,..., $u_n \in E$ avec : $u_j = \sum_{i=1}^n a_{i,j}\, e_i$, $\forall j \in [\![1, n]\!]$. Alors, la proposition 25 nous donne :*

$$\begin{aligned}
\phi(u_1, u_2, \ldots, u_n) &= \sum_{1 \leqslant i_1 \leqslant n, 1 \leqslant i_2 \leqslant n, \ldots, 1 \leqslant i_n \leqslant n} a_{i_1,1}\, a_{i_2,2} \ldots a_{i_n,n}\, \phi(e_{i_1}, e_{i_2}, \ldots, e_{i_n}) \\
& \textit{considérons } (i_1, i_2, \ldots, i_n) = (\sigma(1), \sigma(2), \ldots, \sigma(n)), \sigma \in [\![1, n]\!]^{[\![1,n]\!]}. \\
&= \sum_{\sigma \in \mathcal{F}([\![1,n]\!],[\![1,n]\!])} a_{\sigma(1),1}\, a_{\sigma(2),2} \ldots a_{\sigma(n),n}\, \phi(e_{\sigma(1)}, e_{\sigma(2)}, \ldots, e_{\sigma(n)}) \\
&= \sum_{\sigma \in Bij([\![1,n]\!],[\![1,n]\!])} a_{\sigma(1),1}\, a_{\sigma(2),2} \ldots a_{\sigma(n),n}\, \phi(e_{\sigma(1)}, e_{\sigma(2)}, \ldots, e_{\sigma(n)}).
\end{aligned}$$

On obtient la dernière égalité en remarquant que lorsque σ n'est pas bijective (pas une permutation), la famille $(e_{\sigma(1)}, e_{\sigma(2)}, \ldots, e_{\sigma(n)})$ possède au moins deux vecteurs égaux. Comme ϕ est une forme n-linéaire alternée, on a alors $\phi(e_{\sigma(1)}, \ldots, e_{\sigma(n)}) = 0$.
Enfin, la proposition 33 nous donne :

$$\begin{aligned}
\phi(u_1, u_2, \ldots, u_n) &= \sum_{\sigma \in \mathcal{S}_n} \varepsilon(\sigma)\, a_{\sigma(1),1}\, a_{\sigma(2),2} \ldots a_{\sigma(n),n}\, \phi(e_1, e_2, \ldots, e_n) \\
&= \left(\sum_{\sigma \in \mathcal{S}_n} \varepsilon(\sigma)\, a_{\sigma(1),1}\, a_{\sigma(2),2} \ldots a_{\sigma(n),n} \right) \phi(e_1, e_2, \ldots, e_n). \qquad (*)
\end{aligned}$$

2.2.4 Formes n-linéaires alternées en dimension n

Théorème 35
Soient E un $\mathbb{K}$-e.v. de dimension n et $\mathcal{B} = (e_1, e_2, \ldots, e_n)$ une base de E. Alors :

1. Il existe une unique forme n-linéaire alternée ϕ_0 sur E telle que :
$$\phi_0(e_1, e_2, \ldots, e_n) = 1.$$
2. Toute forme n-linéaire alternée sur E est proportionnelle à ϕ_0.

Corollaire 36
Soit E un $\mathbb{K}$-e.v. de dimension n. Alors l'ensemble des forme n-linéaires alternées sur E est un e.v. de dimension 1.

Preuve —
Cas de la dimension 2. Soient E un $\mathbb{K}$-espace vectoriel de dimension 2, $\mathcal{B} = (e_1, e_2)$ une base, et ϕ une forme bilinéaire alternée sur E. Soient $u, v \in E$ avec :
$$u = a_1\, e_1 + a_2\, e_2 \quad \text{et} \quad v = b_1\, e_1 + b_2\, e_2,$$
la bilinéarité de ϕ permet d'écrire :
$$\begin{aligned}
\varphi(u, v) &= \varphi(a_1\, e_1 + a_2\, e_2, b_1\, e_1 + b_2\, e_2) \\
&= a_1\, b_1\, \varphi(e_1, e_1) + a_1\, b_2\, \varphi(e_1, e_2) + a_2\, b_1\, \varphi(e_2, e_1) + a_2\, b_2\, \varphi(e_2, e_2).
\end{aligned}$$
Comme φ est alternée, on a :
$$\varphi(e_1, e_1) = \varphi(e_2, e_2) = 0 \quad \text{et} \quad \varphi(e_2, e_1) = -\varphi(e_1, e_2),$$
ce qui donne :
$$\varphi(u, v) = (a_1\, b_2 - a_2\, b_1)\, \varphi(e_1, e_2).$$
On voit :

- que φ est proportionnelle à la forme bilinéaire φ_0 :
$$\begin{array}{rccc} \varphi_0 : & E^2 & \longrightarrow & \mathbb{K} \\ & (u, v) & \longmapsto & a_1\, b_2 - a_2\, b_1 \end{array}$$
- que si l'on impose $\varphi(e_1, e_2) = 1$, la forme φ est alors égale à φ_0,
- que φ_0 est une forme bilinéaire alternée non nulle puisque $\varphi_0(e_1, e_2) = 1$. ◀

Le cas général. Soit $\mathcal{B} = (e_1, e_2, \ldots, e_n)$ une base de E. Posons : $u_j = \displaystyle\sum_{i=1}^n a_{i,j}\, e_i$, $\forall j \in [\![1, n]\!]$ et notons $\phi_0 : E^n \to \mathbb{K}$ l'application définie par :
$$\phi_0(u_1, u_2, \ldots, u_n) = \sum_{\sigma \in \mathcal{S}_n} \varepsilon(\sigma) a_{\sigma(1),1}\, a_{\sigma(2),2} \ldots a_{\sigma(n),n}$$

- Pour ϕ une forme n-linéaire alternée sur E, on a obtenu précédemment que :
$$\begin{aligned}\phi(u_1,u_2,\ldots,u_n) &= \sum_{\sigma\in\mathcal{S}_n} \varepsilon(\sigma)\, a_{\sigma(1),1}\, a_{\sigma(2),2}\cdots a_{\sigma(n),n}\, \phi(e_1,e_2,\ldots,e_n)\\ &= \phi(e_1,e_2,\ldots,e_n) \sum_{\sigma\in\mathcal{S}_n} \varepsilon(\sigma)\, a_{\sigma(1),1}\, a_{\sigma(2),2}\cdots a_{\sigma(n),n}\\ &= \phi(e_1,e_2,\ldots,e_n)\, \phi_0(u_1,u_2,\ldots,u_n),\end{aligned}$$
donc,
$$\phi = \phi(e_1,e_2,\ldots,e_n)\,\phi_0,$$
c'est-à-dire ϕ est proportionnelle à ϕ_0. Si l'on a $\phi(e_1,e_2,\ldots e_n)=1$, alors on a $\phi=\phi_0$, ce qui prouve l'unicité.
Il reste à montrer que ϕ_0 est une forme n-linéaire alternée vérifiant $\phi_0(e_1,e_2,\ldots,e_n)=1$.
- ϕ_0 est bien une forme n-linéaire, d'après la caractérisation des formes n-linéaires obtenue précédemment. ◀
- Soient $i\neq j$ tels que l'on ait $u_i=u_j$. Montrons que l'on a alors $\phi_0(u_1,u_2,\ldots,u_n)=0$.
Notons τ la transposition (i,j). On sait alors que l'ensemble des permutations impaires est $\mathcal{A}_n\,\tau = \{\sigma\tau, \sigma\in\mathcal{A}_n\}$. On peut donc écrire :
$$\begin{aligned}\phi_0(u_1,u_2,\ldots,u_n) &= \sum_{\sigma\in\mathcal{S}_n}\varepsilon(\sigma)\prod_{k=1}^n a_{\sigma(k),k}\\ &= \sum_{\sigma\in\mathcal{A}_n}\varepsilon(\sigma)\prod_{k=1}^n a_{\sigma(k),k} + \sum_{\sigma\in\mathcal{A}_n}\varepsilon(\sigma\tau)\prod_{k=1}^n a_{\sigma(\tau(k)),k}\\ &= \sum_{\sigma\in\mathcal{A}_n}\prod_{k=1}^n a_{\sigma(k),k} - \sum_{\sigma\in\mathcal{A}_n}\prod_{k=1}^n a_{\sigma(\tau(k)),k} = \sum_{\sigma\in\mathcal{A}_n}\Big(\prod_{k=1}^n a_{\sigma(k),k} - \prod_{k=1}^n a_{\sigma(\tau(k)),k}\Big).\end{aligned}$$
Or, on a $u_i=u_j$, donc $a_{i,l}=a_{j,l}$ pour tout $1\le l\le n$. Cela donne :
$$a_{\sigma(\tau(i)),i} = a_{\sigma(j),i} = a_{\sigma(j),j} \quad\text{et}\quad a_{\sigma(\tau(j)),j} = a_{\sigma(i),j} = a_{\sigma(i),i}.$$
Pour $k\notin\{i,j\}$, on a $a_{\sigma(\tau(k)),k} = a_{\sigma(k),k}$. Ainsi ;
$$\prod_{k=1}^n a_{\sigma(\tau(k)),k} = \prod_{k=1}^n a_{\sigma(k),k}$$
et par suite $\phi_0(u_1,u_2,\ldots,u_n)=0$. ϕ_0 est donc bien alternée.
- Enfin, vérifions que $\phi_0(e_1,e_2,\ldots,e_n)=1$. Par définition de ϕ_0, on a :
$$\phi_0(e_1,e_2,\ldots,e_n) = \sum_{\sigma\in\mathcal{S}_n}\varepsilon(\sigma)\,\delta_{\sigma(1),1}\,\delta_{\sigma(2),2}\cdots\delta_{\sigma(n),n}$$
où $(\delta_{i,j})_{1\leqslant i,j\leqslant n}$ est la matrice des composantes des vecteurs e_1, e_2,.., e_n dans la base $(e_1,e_2,\ldots,e_n)$. Cette matrice est exactement I_n.
Donc, dans la somme précédente, on a :
  - Si σ n'est pas l'identité, il existe $i\in[\![1,n]\!]$ tel que $\sigma(i)\neq i$ et donc $\delta_{\sigma(i),i}=0$.
  Donc le produit $\varepsilon(\sigma)\,\delta_{\sigma(1),1}\,\delta_{\sigma(2),2}\cdots\delta_{\sigma(n),n}$ est nul.
  - Si $\sigma=\mathrm{Id}_{[\![1,n]\!]}$, on a $\varepsilon(\sigma)\,\delta_{\sigma(1),1}\,\delta_{\sigma(2),2}\cdots\delta_{\sigma(n),n}=1$.

Donc $\phi_0(e_1,e_2,\ldots,e_n)=1$.

□

2.3 Déterminant

2.3.1 Diverses notions de déterminants

Déterminant d'une famille de vecteurs dans une base

Définition 37
Soient E un $\mathbb{K}$-e.v. de dimension n, $\mathcal{B}$ une base de E, et ϕ_0 l'unique forme n-linéaire alternée telle que $\phi_0(\mathcal{B})=1$. Soient $u_1,\ldots,u_n\in E$.
On dit que le nombre $\phi_0(u_1,u_2,\ldots,u_n)$ s'appelle **déterminant de la famille** $(u_1,u_2,\ldots,u_n)$ **dans la base** $\mathcal{B}$ \向量组$(u_1,u_2,\ldots,u_n)$在基$\mathcal{B}$下的行列式\. On le note $\det_{\mathcal{B}}(u_1,u_2,\ldots,u_n)$.

On note $\det_{\mathcal{B}}$ la fonction $\det_{\mathcal{B}} : (u_1,u_2,\ldots,u_n)\in E^n \mapsto \det_{\mathcal{B}}(u_1,u_2,\ldots,u_n)\in\mathbb{K}$.

Remarque 38 — *Comme l'ensemble des formes n-linéaires alternées sur E est un e.v. de dimension 1 et que $\det_{\mathcal{B}}$ est un élément non nul de cet e.v., $(\det_{\mathcal{B}})$ est donc une base de cet e.v.. Toute forme n-linéaire alternée sur E est donc proportionnelle à $\det_{\mathcal{B}}$.*
Soit $\mathcal{B}'$ est une autre base de E. Alors il existe un scalaire λ tel que $\det_{\mathcal{B}'}=\lambda\det_{\mathcal{B}}$, c'est-à-dire :
$$\det_{\mathcal{B}'}(u_1,u_2,\ldots,u_n) = \lambda\det_{\mathcal{B}}(u_1,u_2,\ldots,u_n),\ \forall(u_1,u_2,\ldots,u_n)\in E^n.$$
Pour $\mathcal{B}=(e_1,\ldots,e_n)$, on obtient $\lambda=\det_{\mathcal{B}'}(e_1,\ldots,e_n)=\det_{\mathcal{B}'}(\mathcal{B})$. Pour $\mathcal{B}'=(f_1,\ldots,f_n)$, on a alors :
$$1 = \det_{\mathcal{B}'}(\mathcal{B})\ \det_{\mathcal{B}}(f_1,\ldots,f_n) = \det_{\mathcal{B}'}(\mathcal{B})\ \det_{\mathcal{B}}(\mathcal{B}').$$

Déterminant d'un endomorphisme

Proposition-Définition 39
Soient E un $\mathbb{K}$-e.v. de dimension n et $f : E \to E$ un endomorphisme. Alors il existe un unique $\lambda \in \mathbb{K}$ tel que pour toute base $\mathcal{B}$ de E et pour tout $(u_1, u_2, \ldots, u_n) \in E^n$, on ait :

$$\det_{\mathcal{B}}\big(f(u_1), f(u_2), \ldots, f(u_n)\big) = \lambda \det_{\mathcal{B}}(u_1, u_2, \ldots, u_n).$$

Ce nombre est appelé **déterminant** de f \自同态f的行列式\. On le note $\det(f)$ ou $\det f$.
Pour $\mathcal{B} = (e_1, \ldots, e_n)$ une base de E, on a :

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

Preuve
Unicité. Soit $\mathcal{B} = (e_1, \ldots, e_n)$ une base de E. Si λ convient, alors pour $(u_1, u_2, \ldots, u_n) \in E^n$, on a :

$$\det_{\mathcal{B}}\big(f(u_1), f(u_2), \ldots, f(u_n)\big) = \lambda \det_{\mathcal{B}}(u_1, u_2, \ldots, u_n).$$

En prenant $(u_1, u_2, \ldots, u_n) == (e_1, \ldots, e_n) = \mathcal{B}$, on obtient :

$$\lambda = \det_{\mathcal{B}}\big(f(\mathcal{B})\big),$$

ce qui prouve l'unicité.
Existence.

- Soit $\mathcal{B}_0$ une base de E. L'application :

$$(u_1, u_2, \ldots, u_n) \longmapsto \det_{\mathcal{B}_0}\big(f(u_1), f(u_2), \ldots, f(u_n)\big)$$

de E^n dans $\mathbb{K}$ est une forme n-linéaire alternée. Elle est donc proportionnelle à $\det_{\mathcal{B}_0}$ d'après les résultats précédents. ◀
Il existe donc $\lambda \in \mathbb{K}$ tel que pour tout $(u_1, u_2, \ldots, u_n) \in E^n$, on ait :

$$\det_{\mathcal{B}_0}\big(f(u_1), f(u_2), \ldots, f(u_n)\big) = \lambda \det_{\mathcal{B}_0}(u_1, u_2, \ldots, u_n). \qquad (*)$$

- Soit $\mathcal{B}$ une autre base de E. La forme n-linéaire $\det_{\mathcal{B}}$ est proportionnelle à $\det_{\mathcal{B}_0}$, donc il existe $\alpha \in \mathbb{K}$ tel que $\det_{\mathcal{B}} = \alpha \det_{\mathcal{B}_0}$.
Ainsi, pour tout $(u_1, u_2, \ldots, u_n) \in E^n$, on a :

$$\begin{aligned} \det_{\mathcal{B}}\big(f(u_1), f(u_2), \ldots, f(u_n)\big) &= \alpha \det_{\mathcal{B}_0}\big(f(u_1), f(u_2), \ldots, f(u_n)\big) \\ &= \alpha\big(\lambda \det_{\mathcal{B}_0}(u_1, u_2, \ldots, u_n)\big) \\ &= \lambda \det_{\mathcal{B}}(u_1, u_2, \ldots, u_n), \end{aligned}$$

ce qui prouve le résultat.

□

\注意区别:
$\det_{\mathcal{B}}$是一个n维空间中的n重交替线性函数且由$\mathcal{B}$决定;
$\det f$是一个固定的值，与基的选择无关。\

Remarque 40 — *Pour $B = (e_1, \ldots, e_n)$ une base de E, et $u_1, \ldots, u_n \in E$ avec $u_j = \sum_{i=1}^{n} a_{i,j} e_i$, on a :*

$$det_B(u_1, \ldots, u_n) = \sum_{\sigma \in \mathcal{S}_n} \varepsilon(\sigma) a_{\sigma(1),1}\, a_{\sigma(2),2} \ldots a_{\sigma(n),n}.$$

Cette expression est la formule générique du déterminant d'une famille de n vecteurs dans une base.
Pour $f : E \to E$ une application linéaire, $det(f)$ est égal au déterminant de la famille $f(e_1), \ldots, f(e_n)$ dans la base B.
Dans le cas où $E = \mathbb{K}^n$ et où B est la base canonique de $\mathbb{K}^n$, on notera parfois det à la place de det_B.

Exemples 41

1. *Pour $\mathcal{B}$ une base de E, on a* $\det \mathrm{Id}_E = \det_{\mathcal{B}}(\mathrm{Id}_E(\mathcal{B})) = \det_{\mathcal{B}}(\mathcal{B}) = 1$.
2. *Soient F et G deux sous-espaces supplémentaires de E. Soit s la symétrie par rapport à F parallèlement à G. Soien $(e_1, e_2, \ldots, e_p)$ et $(e_{p+1}, e_{p+2}, \ldots, e_n)$ des bases de F et de F. Alors $\mathcal{B} = (e_1, e_2, \ldots, e_n)$ est une base de E , et on a :*

$$\begin{aligned} \det s &= \det_{\mathcal{B}}\big(s(e_1), \ldots, s(e_n)\big) = \det_{\mathcal{B}}(e_1, \ldots, e_p, -e_{p+1}, \ldots, -e_n) \\ &= (-1)^{n-p} = (-1)^{\dim G}. \end{aligned}$$

3. *Soient $\sigma \in \mathcal{S}_n$ et $\mathcal{B} = (e_1, e_2, \ldots, e_n)$ une base de E. On définit l'application linéaire $f : E \to E$ par :*

$$f(e_i) = e_{\sigma(i)}, \forall i \in [\![1, n]\!].$$

Alors le déterminant de f vaut :

$$\det f = \det_{\mathcal{B}}(e_{\sigma(1)}, e_{\sigma(2)}, \ldots, e_{\sigma(n)}) = \varepsilon(\sigma) \det_{\mathcal{B}}(\mathcal{B}) = \varepsilon(\sigma).$$

Déterminant d'une matrice carrée

DÉFINITION 42
Soit $n \geq 1$. Soit $A \in \mathscr{M}_n(\mathbb{K})$. Soient $C_1, \ldots, C_n$ les colonnes de A, vues comme vecteurs colonnes de $\mathbb{K}^n$. Soit B la base canonique de $\mathbb{K}^n$.
On définit le **déterminant de la matrice** A, noté $\det(A)$ ou $\det A$, par $det(A) = det_B(C_1, \ldots, C_n)$.

Le déterminant d'une matrice carrée A est donc le déterminant de ses vecteurs colonnes, par rapport à la base canonique de $\mathbb{K}^n$.

REMARQUE 43 — *Pour $A = (a_{i,j})_{\substack{1\leqslant i\leqslant n\\ 1\leqslant j\leqslant n}}$, le déterminant de A se note aussi :*

$$\det A = \begin{vmatrix} a_{1,1} & \cdots & a_{1,j} & \cdots & a_{1,n} \\ \vdots & & \vdots & & \vdots \\ a_{i,1} & \cdots & a_{i,j} & \cdots & a_{i,n} \\ \vdots & & \vdots & & \vdots \\ a_{n,1} & \cdots & a_{n,j} & \cdots & a_{n,n}. \end{vmatrix}.$$

On a :

$$det(A) = \sum_{\sigma \in \mathcal{S}_n} \varepsilon(\sigma) a_{\sigma(1),1}\, a_{\sigma(2),2} \ldots a_{\sigma(n),n}.$$

Ceci est l'expression générique du déterminant d'une matrice carrée.

EXEMPLES 44

1. *Pour $n = 2$, on a :*

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc.$$

2. *Pour $n = 3$, on a :*

$$\begin{vmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{vmatrix} = \begin{array}{l} a_{1,1}\, a_{2,2}\, a_{3,3} + a_{1,2}\, a_{2,3}\, a_{3,1} + a_{1,3}\, a_{2,1}\, a_{3,2} \\ -a_{1,1}\, a_{2,3}\, a_{3,2} - a_{1,2}\, a_{2,1}\, a_{3,3} - a_{1,3}\, a_{2,2}\, a_{3,1}. \end{array}$$

*Cette formule peut se retrouver à l'aide de la **méthode** dite **de Sarrus** : On recopie les deux premières lignes de la matrice sous la troisième, et on effectue les "produits" en diagonale, chacun étant affecté du signe + ou − selon le fait que la diagonale et montante ou descendante[1].*

Maintenant que le déterminant d'une matrice carrée a été défini, nous allons pouvoir énoncer des propriétés vérifiées par cette fonction. Ces propriétés sont nombreuses, ce qui montre l'importance de cette fonction.

PROPOSITION 45
Soient $\mathbb{K}$ un corps et $n \geq 1$. Alors la fonction $\det : A \in \mathscr{M}_n(\mathbb{K}) \mapsto \det(A) \in \mathbb{K}$ est une fonction polynômiale en les coefficients de la matrice d'entrée A.

Preuve — Cela découle de l'expression de $\det(A)$. □

PROPOSITION 46 (Déterminant d'une matrice et déterminant d'une famille de vecteurs)
Soit $\mathcal{B} = (e_1, \ldots, e_n)$ une base de E. Soit $(u_1, u_2, \ldots, u_n)$ une famille de n vecteurs de E. Soit A la matrice des coefficients de $(u_1, \ldots, u_n)$ dans la base $\mathcal{B}$. Alors, on a : $\det_{\mathcal{B}}(u_1, u_2, \ldots, u_n) = \det(A)$.

1. ATTENTION — cette méthode n'est pas généralisable à des matrices d'ordre strictement supérieur à 3.

Preuve — Cela découle de l'expression de $\det(A)$. $\square$

PROPOSITION 47 (Déterminant d'une matrice et déterminant d'une application linéaire)
Soient $f : E \to E$ une application linéaire et $\mathcal{B} = (e_1, \dots, e_n)$ une base de E. Soit $A = \mathrm{Mat}_{\mathcal{B}}(f)$ la matrice de f dans la base $\mathcal{B}$. Alors on a : $\det f = \det A$.

Preuve — Ces deux déterminants sont égaux au déterminant de la famille $\big(f(e_1), f(e_2), \dots, f(e_n)\big)$ dans la base $\mathcal{B}$. $\square$

EXEMPLE 48 — *Si D est une matrice diagonale d'éléments diagonaux $\lambda_1, \lambda_2, \dots, \lambda_n$, on a :*

$$\det D = \prod_{i=1}^{n} \lambda_i .$$

La matrice D est la matrice de la famille $(\lambda_1\, e_1, \lambda_2\, e_2, \dots, \lambda_n\, e_n)$ dans la base $\mathcal{B} = (e_1, e_2, \dots, e_n)$. Par n-linéarité du déterminant, on a donc :

$$\det D = \det_{\mathcal{B}}(\lambda_1\, e_1, \lambda_2\, e_2, \dots, \lambda_n\, e_n) = \left(\prod_{i=1}^{n} \lambda_i\right) \det_{\mathcal{B}}(\mathcal{B}) = \prod_{i=1}^{n} \lambda_i .$$

PROPOSITION 49 (Déterminant et géométrie du plan)
Soit $n = 2$. Soit $\mathcal{B} = (\vec{\imath}, \vec{\jmath})$ une base orthonormée du plan $\mathbb{R}^2$. Alors, $|\det_{\mathcal{B}}(\vec{u}, \vec{v})|$ est égal à la surface du parallélogramme de côtés $\vec{u} = \begin{pmatrix} a \\ b \end{pmatrix}$ et $\vec{v} = \begin{pmatrix} c \\ d \end{pmatrix}$ non nuls.

Preuve — La surface de ce parallélogramme est $S = \|\vec{u}\|\,\|\vec{v}\|\,|sin(\vec{u}, \vec{v})|$. Or

$$|sin(\vec{u}, \vec{v})| = \sqrt{1 - \cos^2(\vec{u}, \vec{v})} \qquad \text{et} \qquad \cos(\vec{u}, \vec{v}) = \frac{\vec{u} \cdot \vec{v}}{\|\vec{u}\|\,\|\vec{v}\|}.$$

D'où $S = \sqrt{\|\vec{u}\|^2\,\|\vec{v}\|^2 - (\vec{u} \cdot \vec{v})^2} = \sqrt{(a^2 + b^2)(c^2 + d^2) - (ac + bd)^2} = \sqrt{(ad - bc)^2}$. Donc $S = |ad - bc| = |\det_{\mathcal{B}}(\vec{u}, \vec{v})|$. $\square$

REMARQUE 50 (Déterminant et géométrie de l'espace) — *Soit $n = 3$. Soit $\mathcal{B} = (\vec{\imath}, \vec{\jmath}, \vec{k})$ une base orthonormée directe de l'espace $\mathbb{R}^3$. Alors,*

$$\det_{\mathcal{B}}(\vec{u}, \vec{v}, \vec{w}) = \vec{u} \cdot (\vec{v} \wedge \vec{w}) \qquad \text{et} \qquad V = |\det_{\mathcal{B}}(\vec{u}, \vec{v}, \vec{w})|$$

est le volume du parallépipède[2] *de côtés $\vec{u}$, $\vec{v}$ et $\vec{w}$.*

2.3.2 Propriétés du déterminant

THÉORÈME 51
Soient E un $\mathbb{K}$-e.v. de dimension n et $\mathcal{B}$ une base de E. Soit $(u_1, u_2, \dots, u_n)$ est une famille de n vecteurs de E. Alors, les propriétés suivantes sont équivalentes :

i) $(u_i)_{1 \leqslant i \leqslant n}$ est une base de E,

ii) $\det_{\mathcal{B}}(u_1, u_2, \dots, u_n) \neq 0$.

Preuve —

i) $\Longrightarrow$ *ii)* Si $\mathcal{B}' = (u_i)_{1 \leqslant i \leqslant n}$ est une base de E, on a vu précédemment qu'il existe $\lambda \in \mathbb{K}$ tel que $\det_{\mathcal{B}'} = \lambda \det_{\mathcal{B}}$. Comme on a :

$$1 = \det_{\mathcal{B}'}(u_1, u_2, \dots, u_n) = \lambda \det_{\mathcal{B}}(u_1, u_2, \dots, u_n),$$

on en déduit que $\det_{\mathcal{B}}(u_1, u_2, \dots, u_n) \neq 0$.

ii) $\Longrightarrow$ *i)* En utilisant la proposition 31, on obtient la contraposée, à savoir : Si $(u_i)_{1 \leqslant i \leqslant n}$ n'est pas une base de E, alors la famille $(u_i)_{1 \leqslant i \leqslant n}$ est liée, et $\det_{\mathcal{B}}(u_1, u_2, \dots, u_n) = 0$.

$\square$

PROPOSITION 52
Soient E un $\mathbb{K}$-e.v. de dimension n, f, g deux endomorphismes de E, et $\lambda \in \mathbb{K}$. Alors, on a :

- $\det(\lambda\, f) = \lambda^n \det(f)$;
- $\det(f \circ g) = \det(f)\,\det(g)$.

2. \平行六面体\

Preuve — Soit $\mathcal{B} = (e_i)_{1 \leqslant i \leqslant n}$ une base de E. Par n-linéarité du déterminant, on a :

$$\det(\lambda f) = \det_{\mathcal{B}}\big(\lambda f(e_1), \lambda f(e_2), \dots, \lambda f(e_n)\big) = \lambda^n \det_{\mathcal{B}}\big(f(e_1), f(e_2), \dots, f(e_n)\big) = \lambda^n \det(f).$$

D'après les propriétés du déterminant d'une application linéaire, on a :

$$\begin{aligned}\det(f \circ g) &= \det_{\mathcal{B}}\Big(f \circ g(e_1), f \circ g(e_2), \dots, f \circ g(e_n)\Big) = \det_{\mathcal{B}}\Big(f(g(e_1)), f\big(g(e_2)\big), \dots, f\big(g(e_n)\big)\Big) \\ &= \det f \det_{\mathcal{B}}\big(g(e_1), g(e_2), \dots, g(e_n)\big) = \det f \, \det g.\end{aligned}$$

□

PROPOSITION 53
Soit $n \geq 1$. Soient $A, B \in \mathscr{M}_n(\mathbb{K})$. Soit $\lambda \in \mathbb{K}$. Alors on a :

- $\det(\lambda\, A) = \lambda^n \det(A)$;
- $\det(A\, B) = \det(A) \det(B)$.

Preuve —

- Soit f l'endomorphisme de $\mathbb{K}^n$ canoniquement associé à A. La matrice de $\lambda\, f$ par rapport à la base canonique de $\mathbb{K}^n$ est $\lambda\, A$ et l'on a :
$$\det(\lambda\, A) = \det(\lambda\, f) = \lambda^n \det f = \lambda^n \det A.$$
- Soit g l'endomorphisme de $\mathbb{K}^n$ canoniquement associé à B. La matrice de $f \circ g$ par rapport à la base canonique de $\mathbb{K}^n$ est $A\, B$ et l'on a :
$$\det(A\, B) = \det(f \circ g) = \det(f) \det(g) = \det(A) \det(B).$$

□

REMARQUE 54 — *Pour* $(A, B) \in \mathscr{M}_n(\mathbb{K})^2$, *on a donc* $\det(A\, B) = \det(A) \det(B) = \det(B\, A)$.

PROPOSITION 55
Soit $n \geq 1$. Soit E un $\mathbb{K}$-e.v. de dimension n.

1. Soit $f : E \to E$ une application linéaire. Alors, f est bijective si, et seulement si, $\det f \neq 0$.
Dans ce cas, on obtient :
$$\det\left(f^{-1}\right) = (\det f)^{-1}.$$
2. Soit $A \in \mathscr{M}_n(\mathbb{K})$. Alors, A est inversible si, et seulement si, $\det A \neq 0$.
Dans ce cas, on obtient :
$$\det\left(A^{-1}\right) = (\det A)^{-1}.$$

Preuve

1. Soit $\mathcal{B} = (e_i)_{1 \leqslant i \leqslant n}$ une base de E. On a :
$$\det f = \det_{\mathcal{B}}\big(f(e_1), f(e_2), \dots, f(e_n)\big).$$
D'après le théorème 51, le scalaire $\det f$ est non nul si, et seulement si, $\big(f(e_1), f(e_2), \dots, f(e_n)\big)$ est une base de E, c'est-à-dire si, et seulement si, f est bijective sur E (d'après le théorème du rang).
La proposition 52 permet alors d'écrire :
$$\det f \, \det\left(f^{-1}\right) = \det(f \circ f^{-1}) = \det \mathrm{Id}_E = 1,$$
ce qui donne $\det\left(f^{-1}\right) = (\det f)^{-1}$.
2. Soit f l'endomorphisme de $\mathbb{K}^n$ canoniquement associé à A. On sait que $\det A = \det f$ et que A est inversible si, et seulement si, f est bijectif. Grâce au résultat précédent, on en déduit que A est inversible si, et seulement si, $\det A \neq 0$.
La proposition 53 permet alors d'écrire :
$$\det A \, \det\left(A^{-1}\right) = \det(A\, A^{-1}) = \det I_n = 1,$$
ce qui donne $\det\left(A^{-1}\right) = (\det A)^{-1}$.

□

PROPOSITION 56
Soit $n \geq 2$. Soit $A \in \mathscr{M}_n(\mathbb{K})$. Alors, on a :
$$\det A = \det\left({}^tA\right).$$

Notons $L_1, \dots, L_n$ les lignes de A, vues comme vecteurs lignes de $\mathbb{K}^n$. Alors, on a $\det(A) = \det({}^tL_1, \dots, {}^tL_1)$. La fonction $A \mapsto \det(A)$ est une forme n-linéaire alternée en les lignes de la matrice A.

Preuve — On a :

$$\begin{aligned}\det A &= \sum_{\sigma\in\mathcal{S}_n} \varepsilon(\sigma)\, a_{\sigma(1),1}\, a_{\sigma(2),2}\cdots a_{\sigma(n),n} = \sum_{\sigma\in\mathcal{S}_n} \varepsilon(\sigma)\, a_{1,\sigma^{-1}(1)}\, a_{2,\sigma^{-1}(2)}\cdots a_{n,\sigma^{-1}(n)}\\ &= \sum_{\sigma\in\mathcal{S}_n} \varepsilon(\sigma^{-1})\, a_{1,\sigma^{-1}(1)}\, a_{2,\sigma^{-1}(2)}\cdots a_{n,\sigma^{-1}(n)} = \sum_{\tau\in\mathcal{S}_n} \varepsilon(\tau)\, a_{1,\tau(1)}\, a_{2,\tau(2)}\cdots a_{n,\tau(n)}\\ &= \det({}^tA).\end{aligned}$$

Pour $L_1,\dots,L_n$ les lignes de A, vues comme vecteurs lignes de $\mathbb{K}^n$, les colonnes de la matrice ${}^tA$ sont exactement les ${}^tL_1,\dots,{}^tL_n$. On obtient donc que :

$$\det(A) = \det({}^tA) = \det({}^tL_1,\dots,{}^tL_1).$$

Comme $\det({}^tA)$ est une forme n-linéaire alternée en les colonnes de ${}^tA$, $\det(A)$ est donc une forme n-linéaire alternée en les lignes de A. □

Remarque 57 — *Pour $A\in\mathscr{M}_n(\mathbb{K})$, la quantité $\det(A)$ reste la même que l'on considère A comme une matrice composée de n vecteurs colonnes $C_1,\dots,C_n$ ou une matrice de n vecteurs lignes $L_1,\dots,L_n$.*
On peut ainsi calculer le déterminant d'une matrice A en effectuant des opérations sur lignes ou sur les colonnes de A (que des opérations sur les lignes, ou que des opérations sur les colonnes, ou un mélange d'opérations sur les lignes et d'opérations sur les colonnes).

2.3.3 Opérations sur les lignes ou les colonnes d'un déterminant

Comme on l'a précédemment vu , le déterminant d'une famille de vecteurs par rapport à une base, ou du déterminant d'un endomorphisme, est le déterminant d'une certaine matrice carrée. Nous nous intéresserons donc aux méthodes permettant de calculer le déterminant d'une matrice A.

Le déterminant d'une matrice étant une forme n-linéaire alternée des colonnes ou des lignes de cette matrice, les propriétés des formes n-linéaires alternées permettent d'énoncer les règles suivantes.

Proposition 58 (Déterminant et opérations sur les lignes/colonnes)
Soit $n\geq 2$. Soit $A\in\mathscr{M}_n(\mathbb{K})$.

- Si A a deux colonnes (resp. deux lignes) identiques, alors $\det(A)=0$.
- L'échange de deux colonnes de A (resp. deux lignes) multiplie son déterminant par -1.
- Si une colonne (resp. une ligne) de A est combinaison linaire des **autres** colonnes (resp. des **autres** lignes), alors $\det(A)=0$.
- Si une colonne (resp. une ligne) de A est formée de 0, alors $\det(A)=0$.
- La valeur de $\det(A)$ est inchangée si l'on ajoute à une colonne (resp. à une ligne) de A une combinaison linaire des **autres** colonnes (resp. des **autres** lignes).
- Si l'on multiplie une colonne de A (resp. une ligne) par λ, alors son déterminant est multiplié par λ. Donc, si l'on multiplie la matrice A par λ, son déterminant est multiplié par λ^n.

Remarque 59 — *Ces règles de transformation d'un déterminant permettent :*

- *soit de prouver qu'il est nul,*
- *soit d'introduire dans une colonne (resp. une ligne) un maximum de 0 afin de pouvoir utiliser les résultats qui vont suivre*[3]*.*

Cette proposition décrit le comportement du déterminant lorsque l'on applique à une matrice A des opérations élémentaires sur ses lignes ou sur ses colonnes ($L_i\leftarrow L_i+\lambda L_j$, $L_i\leftarrow\lambda L_i$, $L_i\leftrightarrow L_j$).
*On peut ainsi utiliser la **méthode du Pivot** (ou Pivot de Gauss) pour calculer $\det(A)$.*
La méthode du Pivot permet de se ramener à une matrice échelonnée, et nous allons voir des façons de calculer le déterminant d'une matrice échelonnée.

Exemples 60

1. *On a :*

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

puisque la matrice a deux colonnes proportionnelles.

3. \尽可能最大程度地在矩阵的行或列中变换出0，以便于使用后面介绍的方法。\

2. *On a :*

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

puisque la deuxième ligne est la demi-somme des deux autres. ($L_3 = \frac{L_1+L_2}{2}$)

2.3.4 Développement d'un déterminant selon une colonne ou une ligne

Proposition 61
Soit $n \geq 2$. Soit $A \in \mathscr{M}_n(\mathbb{K})$ de la forme :

$$A = \begin{pmatrix} & & & 0 \\ & A' & & \vdots \\ & & & 0 \\ * & \dots & * & a_{n,n} \end{pmatrix}.$$

Alors, on a $\det A = a_{n,n} \det A'$.

Preuve — On a par définition :

$$\det A = \sum_{\sigma \in \mathcal{S}_n} \varepsilon(\sigma)\, a_{\sigma(1),1}\, a_{\sigma(2),2} \dots a_{\sigma(n),n}.$$

Soit $\sigma \in \mathcal{S}_n$. Si $\sigma(n) \neq n$, alors $a_{\sigma(n),n} = 0$. On a donc :

$$\det A = \sum_{\sigma \in \mathcal{S}_n, \sigma(n)=n} \varepsilon(\sigma)\, a_{\sigma(1),1}\, a_{\sigma(2),2} \dots a_{\sigma(n-1),n-1}\, a_{n,n}.$$

Les permutations σ de $\mathcal{S}_n$ telles que $\sigma(n) = n$ sont exactement de la forme $\sigma = s$, pour $s \in \mathcal{S}_{n-1}$. Les propriétés de la signature sur $\mathcal{S}_n$ et $\mathcal{S}_{n-1}$ nous disent qu'on a alors $\varepsilon(s) = \varepsilon(\sigma)$.
Ainsi, on a :

$$\det A = a_{n,n} \sum_{s \in \mathcal{S}_{n-1}} \varepsilon(s)\, a_{s(1),1}\, a_{s(2),2} \dots a_{s(n-1),n-1} = a_{n,n} \det A'.$$

□

Remarque 62 — *De façon similaire, si* $A = \begin{pmatrix} a_{1,1} & * & \dots & * \\ 0 & & & \\ \vdots & & A' & \\ 0 & & & \end{pmatrix}$, *alors* $\det A = a_{1,1} \det A'$.

Corollaire 63
Soit $A = (a_{i,j})_{i,j} \in \mathscr{M}_n(\mathbb{K})$. Si A est une matrice triangulaire, alors on a : $\det A = \prod_{i=1}^{n} a_{i,i}$. [4]

Preuve

- Comme $\det(A) = \det({}^tA)$, il suffit de démontrer le résultat pour les matrices triangulaires inférieures.
- Si A est triangulaire inférieure, la proposition précédente nous donne $\det A = a_{n,n} \det A'$ où A' est la matrice A privée de sa dernière ligne et de sa dernière colonne. Une récurrence facile nous donne alors le résultat.

□

Exemple 64 — *Soient* $a, b, c \in \mathbb{K}$. *On veut calculer le déterminant :*

$$\Delta = \begin{vmatrix} 1 & a & a^2 \\ 1 & b & b^2 \\ 1 & c & c^2 \end{vmatrix}.$$

4. \计算方阵行列式的重要性质之一。\

On a successivement :

$$\begin{aligned}\Delta &= \begin{vmatrix}1 & a & a^2\\ 0 & b-a & b^2-a^2\\ 0 & c-a & c^2-a^2\end{vmatrix} && L_2 \leftarrow L_2 - L_1,\ L_3 \leftarrow L_3 - L_1\\ &= (b-a)(c-a)\begin{vmatrix}1 & a & a^2\\ 0 & 1 & b+a\\ 0 & 1 & c+a\end{vmatrix} && \textit{mise en facteur dans } L_2 \textit{ et } L_3\\ &= (b-a)(c-a)\begin{vmatrix}1 & a & a^2\\ 0 & 1 & b+a\\ 0 & 0 & c-b\end{vmatrix} && L_3 \leftarrow L_3 - L_2\\ &= (b-a)(c-a)(c-b) && \textit{d'après le corollaire 63.}\end{aligned}$$

Définition 65
Soit $n \geq 2$. Soit $A = (a_{i,j}) \in \mathscr{M}_n(\mathbb{K})$. Soient $i, j \in [\![1, n]\!]$.
On note $\Delta_{i,j}$ le déterminant de la matrice extraite de A obtenue en supprimant la i-ième ligne et la j-ième colonne de A.
Ce déterminant est appelé **mineur** \余子式\ de A.[5]
Le nombre $(-1)^{i+j}\Delta_{i,j}$ est appelé **cofacteur** \代数余子式\ de A.

Théorème 66 (Développement du déterminant selon une colonne \行列式按一列展开\)
Soit $n \geq 2$. Soit $A = (a_{i,j})_{i,j} \in \mathscr{M}_n(\mathbb{K})$. On a :

$$\det A = \sum_{i=1}^{n} a_{i,j}\,(-1)^{i+j}\Delta_{i,j},\ \forall j \in [\![1, n]\!].^{6}\ \text{(développement selon la j-ème colonne de A)}$$

Preuve — Soit $\mathcal{B}$ la base canonique de $\mathbb{K}^n$. Soient $C_1, C_2, \ldots, C_n$ les colonnes de A. Soit $j \in [\![1, n]\!]$. On a $C_j = \sum_{i=1}^{n} a_{i,j}\, e_i$, ce qui donne :

$$\begin{aligned}\det A &= \det_{\mathcal{B}}\left(C_1, C_2, \ldots, C_{j-1}, \sum_{i=1}^{n} a_{i,j}\, e_i, C_{j+1}, \ldots, C_n\right)\\ &= \sum_{i=1}^{n} a_{i,j} \det_{\mathcal{B}}(C_1, C_2, \ldots, C_{j-1}, e_i, C_{j+1}, \ldots, C_n).\end{aligned}$$

Notons :

$$\begin{aligned}D_{i,j} &= \det_{\mathcal{B}}(C_1, C_2, \ldots, C_{j-1}, e_i, C_{j+1}, \ldots, C_n)\\ &= \begin{vmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,j-1} & 0 & a_{1,j+1} & \cdots & a_{1,n}\\ a_{2,1} & a_{2,2} & \cdots & a_{2,j-1} & 0 & a_{2,j+1} & \cdots & a_{2,n}\\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots\\ a_{i-1,1} & a_{i-1,2} & \cdots & a_{i-1,j-1} & 0 & a_{i-1,j+1} & \cdots & a_{i-1,n}\\ a_{i,1} & a_{i,2} & \cdots & a_{i,j-1} & 1 & a_{i,j+1} & \cdots & a_{i,n}\\ a_{i+1,1} & a_{i+1,2} & \cdots & a_{i+1,j-1} & 0 & a_{i+1,j+1} & \cdots & a_{i+1,n}\\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots\\ a_{n,1} & a_{n,2} & \cdots & a_{n,j-1} & 0 & a_{n,j+1} & \cdots & a_{n,n}\end{vmatrix}\end{aligned}$$

On peut opérer sur $D_{i,j}$ une suite de $n-j$ échanges de colonnes pour amener la j-ième colonne en dernière position, puis une suite de $n-i$ échanges de lignes pour amener la i-ième ligne en dernière position. Le déterminant est alors multiplié

5. \把$a_{i,j}$所在的第i行与第j列划去后，所留下来的$n-1$ 阶子矩阵的行列式叫做的$a_{i,j}$ 的余子式。\
6. \计算方阵行列式的重要性质之二。\

par $(-1)^{n-j}(-1)^{n-i} = (-1)^{i+j}$. Cela donne :

$$\begin{aligned}
D_{i,j} &= (-1)^{n-j}\begin{vmatrix}
a_{1,1} & a_{1,2} & \cdots & a_{1,j-1} & a_{1,j+1} & \cdots & a_{1,n} & 0\\
a_{2,1} & a_{2,2} & \cdots & a_{2,j-1} & a_{2,j+1} & \cdots & a_{2,n} & 0\\
\vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots\\
a_{i-1,1} & a_{i-1,2} & \cdots & a_{i-1,j-1} & a_{i-1,j+1} & \cdots & a_{i-1,n} & 0\\
a_{i,1} & a_{i,2} & \cdots & a_{i,j-1} & a_{i,j+1} & \cdots & a_{i,n} & 1\\
a_{i+1,1} & a_{i+1,2} & \cdots & a_{i+1,j-1} & a_{i+1,j+1} & \cdots & a_{i+1,n} & 0\\
\vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots\\
a_{n,1} & a_{n,2} & \cdots & a_{n,j-1} & a_{n,j+1} & \cdots & a_{n,n} & 0
\end{vmatrix}\\
&= (-1)^{n-j}(-1)^{n-i}\begin{vmatrix}
a_{1,1} & a_{1,2} & \cdots & a_{1,j-1} & a_{1,j+1} & \cdots & a_{1,n} & 0\\
a_{2,1} & a_{2,2} & \cdots & a_{2,j-1} & a_{2,j+1} & \cdots & a_{2,n} & 0\\
\vdots & \vdots & & \vdots & \vdots & & \vdots & \vdots\\
a_{i-1,1} & a_{i-1,2} & \cdots & a_{i-1,j-1} & a_{i-1,j+1} & \cdots & a_{i-1,n} & 0\\
a_{i+1,1} & a_{i+1,2} & \cdots & a_{i+1,j-1} & a_{i+1,j+1} & \cdots & a_{i+1,n} & 0\\
\vdots & \vdots & & \vdots & \vdots & & \vdots & \vdots\\
a_{n,1} & a_{n,2} & \cdots & a_{n,j-1} & a_{n,j+1} & \cdots & a_{n,n} & 0\\
a_{i,1} & a_{i,2} & \cdots & a_{i,j-1} & a_{i,j+1} & \cdots & a_{i,n} & 1
\end{vmatrix} = (-1)^{i+j}\Delta_{i,j},
\end{aligned}$$

d'après la proposition précédente et d'après la définition du mineur $\Delta_{i,j}$. On a donc :

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

□ En appliquant ce résultat à la transposée de A, on obtient :

Théorème 67 (Développement du déterminant selon une ligne \行列式按一行展开\)
Soit $n \geq 2$. Soit $A = (a_{i,j})_{i,j} \in \mathscr{M}_n(\mathbb{K})$. On a :

$$\det A = \sum_{j=1}^{n} a_{i,j}(-1)^{i+j}\Delta_{i,j},\ \forall i \in [\![1,n]\!].\ \text{(développement selon la i-ème ligne de A)}$$

Exemples 68

1. *Pour un déterminant 3×3, on a donc : (développement selon la 1-ère ligne)*

$$\begin{vmatrix} a_{1,1} & a_{1,2} & a_{1,3}\\ a_{2,1} & a_{2,2} & a_{2,3}\\ a_{3,1} & a_{3,2} & a_{3,3}\end{vmatrix} = a_{1,1}\begin{vmatrix} a_{2,2} & a_{2,3}\\ a_{3,2} & a_{3,3}\end{vmatrix} - a_{1,2}\begin{vmatrix} a_{2,1} & a_{2,3}\\ a_{3,1} & a_{3,3}\end{vmatrix} + a_{1,3}\begin{vmatrix} a_{2,1} & a_{2,2}\\ a_{3,1} & a_{3,2}\end{vmatrix}.$$

   Mais aussi : (développement selon la 2-ème colonne)

$$\begin{vmatrix} a_{1,1} & a_{1,2} & a_{1,3}\\ a_{2,1} & a_{2,2} & a_{2,3}\\ a_{3,1} & a_{3,2} & a_{3,3}\end{vmatrix} = -a_{1,2}\begin{vmatrix} a_{2,1} & a_{2,3}\\ a_{3,1} & a_{3,3}\end{vmatrix} + a_{2,2}\begin{vmatrix} a_{1,1} & a_{1,3}\\ a_{3,1} & a_{3,3}\end{vmatrix} - a_{3,2}\begin{vmatrix} a_{1,1} & a_{1,3}\\ a_{2,1} & a_{2,3}\end{vmatrix}.$$

2. *Le déterminant de Vandermonde \范德蒙德行列式\.*
   Soient $x_0, x_1, \ldots, x_n \in \mathbb{K}$. On définit $V(x_0, x_1, \ldots, x_n)$ le déterminant de Vandermonde par :

$$V(x_0,x_1,\ldots,x_n) = \begin{vmatrix} 1 & 1 & \cdots & 1\\ x_0 & x_1 & \cdots & x_n\\ x_0^2 & x_1^2 & \cdots & x_n^2\\ \vdots & \vdots & & \vdots\\ x_0^n & x_1^n & \cdots & x_n^n\end{vmatrix}$$

   On a alors :

$$V(x_0,x_1,\ldots,x_n) = \prod_{0\leqslant j<i\leqslant n}(x_i - x_j).$$

   Cela montre que :

$$V(x_0,x_1,\ldots,x_n) = 0 \iff \exists i,j \in [\![0,n]\!] \text{ tels que } i \neq j \text{ et } x_i = x_j.$$

   Montrons cela par récurrence sur n. On définit la propriété H_n : Soient $x_0, x_1, \ldots, x_n \in \mathbb{K}$. On a :

$$V(x_0,x_1,\ldots,x_n) = \prod_{0\leqslant j<i\leqslant n}(x_i - x_j).$$

- *H_1 est vraie, car pour $x_0, x_1 \in \mathbb{K}$, on a :*

$$V(x_0, x_1) = \begin{vmatrix} 1 & 1 \\ x_0 & x_1 \end{vmatrix} = x_1 - x_0.$$

- *Soit $n \geq 2$. Supposons que H_{n-1} est vraie. Soient $x_0, x_1, \ldots, x_n \in \mathbb{K}$.*
  - *Si les scalaires $x_0, x_1, \ldots, x_n$ ne sont pas distincts deux à deux, le déterminant $V(x_0, x_1, \ldots, x_n)$ a deux colonnes identiques et est donc nul. Dans ce cas $V(x_0, x_1, \ldots, x_n) = \prod_{0 \leqslant j < i \leqslant n} (x_i - x_j) = 0$.*
  - *Si les scalaires $x_0, x_1, \ldots, x_n$ sont distincts deux à deux, on développe ce déterminant par rapport à la dernière colonne. Cela montre que la fonction $f : x \mapsto V(x_0, x_1, \ldots, x_{n-1}, x)$ est une fonction polynomiale en x, de degré inférieur ou égal à n, et dont le terme de degré n est :*

$$V(x_0, x_1, \ldots, x_{n-1})\, x^n.$$

D'après H_{n-1}, on a :

$$V(x_0, x_1, \ldots, x_{n-1}) = \prod_{0 \leqslant j < i \leqslant n-1} (x_i - x_j)$$

Cette quantité est donc non nulle car $x_0, x_1, \ldots, x_{n-1}$ sont deux à deux distincts. Ainsi, f est une fonction polynomiale de degré n.
Or, f admet comme racines $x_0, x_1, \ldots, x_{n-1}$ car pour tout $i \in [\![0, n-1]\!]$, la quantité $f(x_i)$ est un déterminant admettant deux colonnes identiques. On connaît donc toutes les racines de f. Cela donne :

$$V(x_0, x_1, \ldots, x_{n-1}, x) = V(x_0, x_1, \ldots, x_{n-1}) \prod_{j=0}^{n-1} (x - x_j), \ \forall x \in \mathbb{K}.$$

Ceci implique :

$$V(x_0, x_1, \ldots, x_n) = \prod_{0 \leqslant j < i \leqslant n-1} (x_i - x_j) \prod_{j=0}^{n-1} (x_n - x_j) = \prod_{0 \leqslant j < i \leqslant n} (x_i - x_j),$$

ce qui montre que H_n est vraie, ce qui termine la récurrence.

2.3.5 Déterminant d'une matrice par blocs

Proposition 69
Soit $n \geq 2$. Soit $A \in \mathscr{M}_n(\mathbb{K})$ possédant un bloc de zéros en bas à gauche. On l'écrit sous la forme :

$$A = \begin{pmatrix} C & D \\ 0 & E \end{pmatrix}$$

où C est une matrice $p \times p$, D une matrice $p \times (n-p)$, 0 la matrice nulle $(n-p) \times p$, et E une matrice $(n-p) \times (n-p)$. Alors on a :

$$\det A = (\det C) \cdot (\det E).$$

Preuve — On a :

$$\det A = \sum_{\sigma \in \mathcal{S}_n} \epsilon(\sigma) A_{\sigma(1)1} A_{\sigma(2)2} \cdots A_{\sigma(n)n}.$$

Or, si $j \leq p$ et $i \geq p+1$, on a $A_{ij} = 0$. On peut donc enlever de la somme toutes les permutations σ pour lesquelles il existe $j \in [\![1, p]\!]$ avec $\sigma(j) \notin [\![1, p]\!]$. Donc il reste seulement les permutations σ telles que $\sigma([\![1, p]\!]) \subset [\![1, p]\!]$. Comme σ est une bijection sur $[\![1, n]\!]$, on a forcément

$$\sigma([\![1, p]\!]) = [\![1, p]\!] \qquad \text{et} \qquad \sigma([\![p+1, n]\!]) = [\![p+1, n]\!].$$

Cela signifie que les σ qui restent dans la somme se décomposent en composée (qui commute) d'une bijection de $[\![1, p]\!]$ et d'une bijection de $[\![p+1, n]\!]$. Donc :

$$\det A = \sum_{\alpha \in \mathrm{Bij}([\![1,p]\!]), \beta \in \mathrm{Bij}([\![p+1,n]\!])} \epsilon(\alpha \circ \beta) A_{\alpha(1)1} \cdots A_{\alpha(p)p} \cdot A_{\beta(p+1),p+1} \cdots A_{\beta(n),n}.$$

Puisque $\epsilon(\alpha \circ \beta) = \epsilon(\alpha)\epsilon(\beta)$, on obtient ensuite :

$$\det A = \left(\sum_{\alpha \in \mathrm{Bij}([\![1,p]\!])} \epsilon(\alpha) A_{\alpha(1)1} \cdots A_{\alpha(p)p} \right) \cdot \left(\sum_{\beta \in \mathrm{Bij}([\![p+1,n]\!])} \epsilon(\beta) A_{\beta(p+1),p+1} \cdots A_{\beta(n),n} \right).$$

Le premier facteur est égal à :

$$\sum_{\alpha \in \mathrm{Bij}(p)} \epsilon(\alpha) C_{\alpha(1)1} \cdots C_{\alpha(p)p} = \det C.$$

Pour le second facteur, si on associe à $\beta \in \mathrm{Bij}([\![p+1,n]\!])$ la permutation β' de $[\![1,n-p]\!]$ définie par $\beta'(k)=\beta(k+p)-p$, on a $\epsilon(\beta)=\epsilon(\beta')$, et le second facteur s'écrit alors :

$$\sum_{\beta\in\mathcal{S}_(n-p)} \epsilon(\beta)E_{\beta'(1),1}\cdots E_{\beta'(n-p),n-p} = \det E.$$

D'où le résultat. □

Par itération de ce théorème, on peut calculer facilement les déterminants de matrices "triangulaires par blocs".

Exemples 70

1. $$\det\begin{bmatrix} 1 & -1 & 4 & 3 & 7 & -2 \\ 1 & 2 & -2 & -3 & 5 & -6 \\ 0 & 0 & 4 & 2 & -3 & 7 \\ 0 & 0 & -1 & -3 & 5 & 2 \\ 0 & 0 & 0 & 0 & -2 & 3 \\ 0 & 0 & 0 & 0 & 3 & -2 \end{bmatrix} = \det\begin{bmatrix} 1 & -1 \\ 1 & 2 \end{bmatrix}\cdot\det\begin{bmatrix} 4 & 2 \\ -1 & -3 \end{bmatrix}\cdot\det\begin{bmatrix} -2 & 3 \\ 3 & -2 \end{bmatrix}.$$
2. *Comme le déterminant est invariant par transposition, on peut aussi calculer des déterminants "triangulaires inférieurs par blocs" :*
$$\det\begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 2 & 0 & 0 \\ 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \end{bmatrix} = \det\begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}\cdot\det\begin{bmatrix} 3 & 4 \\ 7 & 8 \end{bmatrix}.$$

Le déterminant d'une matrice "diagonale par blocs" est facile à calculer : c'est le produit des déterminants de tous les blocs.

Corollaire 71

Soit $n\geq 2$. Soit $1\leq k\geq n$ et $n_1,\ldots,n_k\in\mathbb{N}^*$ tels que $n_1+\ldots+n_k=n$. Soit $A\in\mathscr{M}_n(\mathbb{K})$

1. Si A une matrice diagonale par blocs :
$$A=\begin{pmatrix} A_1 & 0 & \cdots & 0 \\ 0 & A_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \cdots & A_k \end{pmatrix},\text{ avec } A_i\in\mathscr{M}_{n_i}(\mathbb{K}),$$
alors on a $\det A=\prod_{i=1}^k \det A_i$.
2. Soit E un $\mathbb{K}$-e.v. de dimension n.Soient $E_1,\ldots,E_k$ des sous-e.v. de E en somme directe : $E=\bigoplus_{i=1}^k E_i$. Soit $u: E\to E$ un endomorphisme, tel que chaque sous-e.v. E_i est stable par u. On note $u_i : E_i\to E_i$ l'endomorphisme induit par u sur le sous-espace E_i.
Alors, on a $\det u=\prod_{i=1}^k \det u_i$.

Preuve — La seconde partie est seulement la traduction de la première partie au niveau des endomorphismes.
La première partie se prouve par récurrence sur k, en se ramenant à la formule $\det\begin{pmatrix} A_1 & 0 \\ 0 & A_2\end{pmatrix} = (\det A_1)\cdot(\det A_2)$. □ ◄

2.3.6 Comatrice

Définition 72

Soit $n\geq 2$. Soit $A=(a_{i,j})_{i,j}\in\mathscr{M}_n(\mathbb{K})$. On définit la matrice $\mathrm{com}\,(A)=(b_{i,j})_{i,j}$, avec $b_{i,j}=(-1)^{i+j}\Delta_{i,j}$.
Cette matrice $\mathrm{com}\,(A)$ est appelée **comatrice** \伴随矩阵\.
C'est la matrice des cofacteurs de A ($b_{i,j}$ est le cofacteur (i,j) de A).

Proposition 73

Soient $n\geq 2$ et $A=(a_{i,j})_{i,j}\in\mathscr{M}_n(\mathbb{K})$. Alors, on a :

$$A\,{}^t\mathrm{com}\,(A) = {}^t\mathrm{com}\,(A)\;A = (\det A)\,I_n.$$

Preuve

- Posons $A\,{}^t\mathrm{com}\,(A) = (c_{i,j})_{1\leqslant i\leqslant n, 1\leqslant j\leqslant n}$.
Soient $i,k\in[\![1,n]\!]$. On a par définition :
$$c_{i,k}=\sum_{j=1}^n a_{i,j}\,b_{k,j}=\sum_{j=1}^n(-1)^{j+k}\Delta_{k,j}\,a_{i,j}.$$

— Si $k = i$, la dernière somme ci-dessus représente le développement du déterminant de A suivant la i-ième ligne. On a donc :

$$c_{i,i} = \sum_{j=1}^{n} (-1)^{i+j} \Delta_{i,j}\, a_{i,j} = \det A.$$

— Si $k \neq i$, soit $A' = (a'_{i,j})_{i,j}$ la matrice obtenue à partir de A en remplaçant sa k-ème ligne L_k par sa i-ème ligne L_i ($L_k \to L_i$). Comme A' admet deux lignes identiques, son déterminant est nul. En développant $\det(A')$ suivant sa k-ième ligne, on a :

$$0 = \det A' = \sum_{j=1}^{n} (-1)^{j+k} \Delta'_{k,j}\, a'_{k,j}.$$

Par construction, on a $a'_{k,j} = a_{i,j}$ et, puisque les lignes de A et de A' autres que les k-ièmes sont identiques, on a $\Delta'_{k,j} = \Delta_{k,j}$. La relation précédente devient donc :

$$0 = \det A' = \sum_{j=1}^{n} a_{i,j} (-1)^{j+k} \Delta_{k,j} = c_{i,k}.$$

On obtient ainsi que $A\, {}^t\mathrm{com}\,(A) = (\det A)\, I_n$.

- On démontre de manière analogue, en utilisant des développements par rapport aux colonnes, la relation ${}^t\mathrm{com}\,(A)\; A = (\det A)\, I_n$.

□

Corollaire 74
Soient $n \geq 2$ et $A = (a_{i,j})_{i,j} \in \mathscr{M}_n(\mathbb{K})$. Si A est inversible, on a alors :

$$A^{-1} = \frac{1}{\det A}\, {}^t\mathrm{com}\,(A)\,.$$

Exemple 75 — *Si* $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, *on a* $\mathrm{com}\,(A) = \begin{pmatrix} d & -c \\ -b & a \end{pmatrix}$.
Si A est inversible, on a :

$$A^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a. \end{pmatrix}$$

Remarques 76

- *À l'exception de ce cas des matrices 2×2 on utilise rarement la formule précédente pour inverser une matrice. En effet,* $\mathrm{com}\,(A)$ *est une matrice dont chacun des coefficients est un déterminant de taille* $(n-1) \times (n-1)$.
- *Par contre, la matrice* $\mathrm{com}\,(A)$ *est une matrice dont tous les coefficients sont des polynômes en les coefficients de A. Ainsi, A^{-1} est une matrice dont tous les coefficients sont un quotient de deux polynômes en les coefficients de A.*
Cette formule peut être très utile lorsque la matrice A dépend d'un paramètre, pour étudier les propriétés de continuité, de dérivabilité… des coefficients de A^{-1}.
Cette formule peut aussi être très utile en algèbre pour caractériser les coefficients de A^{-1} par rapport à ceux de A. Comme les coefficients de A^{-1} sont obtenus à partir de polynômes à coefficients entiers en les coefficients de A, cela montre que l'expression de A^{-1} ne dépend pas du corps $\mathbb{K}$ choisi. Pour une matrice A à coefficients rationnels, A est dans $\mathscr{M}_n(\mathbb{Q})$,$\mathscr{M}_n(\mathbb{R})$,$\mathscr{M}_n(\mathbb{C})$. Peu importe le corps choisi, la valeur de $\det(A)$ *reste la même, et si A est inversible alors A^{-1} sera à coefficients rationnels.*

2.3.7 Formules de Cramer

Proposition 77 (\克拉默法则\)
Soit $n \geq 2$. Soient $A \in \mathscr{M}_n(\mathbb{K})$ inversible et $Y \in \mathscr{M}_{1,n}(\mathbb{K})$ un vecteur colonne. Soit $(\mathscr{S})$ un système linéaire d'écriture matricielle $A\,X = Y$.
Alors, le système $(\mathscr{S})$ possède une unique solution (on dit qe c'est un systme **de Cramer**), et on a :

$$x_i = \frac{\det A_i}{\det A},\ \forall i \in [\![1, n]\!],$$

où A_i est la matrice obtenue à partir de A en remplaçant sa i-ième colonne C_i par le vecteur colonne Y.

Preuve — Si $(\mathscr{S})$ est un système de Cramer, l'unique solution $(x_1, x_2, \ldots, x_n)$ de $(\mathscr{S})$ vérifie la relation vectorielle :

$$\sum_{j=1}^{n} x_j\, C_j = B$$

ce qui implique :

$$\forall i \in [\![1,n]\!]\,,\ \det A_i = \det(C_1,\dots,C_{i-1},B,C_{i+1},\dots,C_n) = \sum_{j=1}^{n} x_j\ \det(C_1,\dots,C_{i-1},C_j,C_{i+1},\dots,C_n).$$

Comme $\det(C_1,\dots,C_{i-1},C_j,C_{i+1},\dots,C_n)$ est nul dès que $j \neq i$, on obtient :

$$\forall i \in [\![1,n]\!]\,,\ \det A_i = x_i \det A$$

ce qu'il fallait démontrer. □

Remarques 78

- *Dans la pratique, ces formules ne sont pas utilisées car elles nécessitent de calculer $n+1$ déterminants de taille $n \times n$.*
  On ne les utilise que dans le cas $n = 2$ et dans les cas où l'on sait calculer facilement les déterminants des matrices A_i et A.
- *Par contre, lorsque le système dépend d'un paramètre, ces formules permettent d'étudier la continuité et la dérivabilité... des solutions en fonction de ce paramètre. On peut faire pareil avec la relation $AX = Y \Leftrightarrow X = A^{-1}Y$, et $A^{-1} = (\det(A))^{-1}\,\mathrm{com}\,(A)$.*

Exemple 79 — *Pour $n = 2$, le système :*

$$\begin{cases} a\,x + b\,y = e \\ c\,x + d\,y = f \end{cases}$$

est un système de Cramer si, et seulement si, on a :

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} \neq 0$$

et la solution en est alors :

$$x = \frac{\begin{vmatrix} e & b \\ f & d \end{vmatrix}}{\begin{vmatrix} a & b \\ c & d \end{vmatrix}} \quad \text{et} \quad y = \frac{\begin{vmatrix} a & e \\ c & f \end{vmatrix}}{\begin{vmatrix} a & b \\ c & d \end{vmatrix}}.$$

2.4 Rappels sur la trace

Trace d'une matrice

Définition 80

Soit $M = (m_{i,j})_{(i,j)\in[\![1,n]\!]^2} \in \mathscr{M}_n(\mathbb{K})$. On appelle *trace* de la matrice M le nombre :

$$\mathrm{Tr}\,(M) = \sum_{i=1}^{n} m_{i,i}.$$

Proposition 81

Soient $(A, B) \in \mathscr{M}_n(\mathbb{K})^2$. On a :

$$\mathrm{Tr}\,(AB) = \mathrm{Tr}\,(BA)\,.$$

Preuve — Posons $A = (a_{i,j})$ et $B = (b_{i,j})$. Alors :

$$\mathrm{Tr}\,(AB) = \sum_{i=1}^{n}\sum_{k=1}^{n} a_{i,k}b_{k,i} = \sum_{k=1}^{n}\sum_{i=1}^{n} b_{k,i}a_{i,k} = \mathrm{Tr}\,(BA)\,.$$

□

On remarque en particulier le fait que, pour toute matrice $M \in \mathscr{M}_n(\mathbb{K})$ et pour toute matrice inversible $P \in \mathrm{GL}_n(\mathbb{K})$, on a :

$$\mathrm{Tr}\left(P^{-1}MP\right) = \mathrm{Tr}\left(MPP^{-1}\right) = \mathrm{Tr}\,(M)\,.$$

Corollaire 82

Soit $n \geq 1$. Soit $\phi \in \mathcal{L}(\mathscr{M}_n(\mathbb{K}), \mathbb{K})$ une forme linéaire. Alors, il existe une unique matrice $A \in \mathscr{M}_n(\mathbb{K})$ telle que :

$$\phi : M \longmapsto \mathrm{Tr}\,(AM)\,.$$

De plus, toute forme linéaire $\phi \in \mathcal{L}(\mathscr{M}_n(\mathbb{K}), \mathbb{K})$ vérifiant

$$\phi(MN) = \phi(NM),\ \forall (M, N) \in \mathscr{M}_n(\mathbb{K})^2,$$

est de la forme :

$$\phi : M \longmapsto \alpha \operatorname{Tr}(M)$$

pour un unique $\alpha \in \mathbb{K}$.

Preuve — La famille $(E_{i,j})_{(i,j)\in[\![1,n]\!]^2}$ est la base canonique de $\mathscr{M}_n(\mathbb{K})$.

- Soit $\phi \in \mathcal{L}(\mathscr{M}_n(\mathbb{K}), \mathbb{K})$. Soit $M = (m_{i,j}) \in \mathscr{M}_n(\mathbb{K})$. On a :

$$\phi(M) = \sum_{(i,j)\in[\![1,n]\!]^2} m_{i,j}\phi(E_{i,j}).$$

  La matrice $A = (a_{i,j})$ avec $a_{i,j} = \phi(E_{j,i})$ est donc l'unique matrice vérifiant $\phi(M) = \operatorname{Tr}(AM)$ pour tout M. Réciproquement, la fonction $M \mapsto \operatorname{Tr}(AM) \in \mathbb{K}$ est bien une forme linéaire.
- Soit $\phi \in \mathcal{L}(\mathscr{M}_n(\mathbb{K}), \mathbb{K})$ vérifiant $\phi(MN) = \phi(NM)$ pour tout $(M, N) \in \mathscr{M}_n(\mathbb{K})^2$.
  Pour tous $i, j, k, l \in \{1, \dots, n\}$, on a :

$$E_{i,j}E_{k,l} = \delta_{j,k}E_{i,l}$$

  Ainsi pour tout (i, j) avec $i = j$, on obtient :

$$\phi(E_{i,i}) = \phi(E_{i,j}E_{j,i}) = \phi(E_{j,i}E_{i,j}) = \phi(E_{j,j}) = \phi(E_{1,1}).$$

  Pour tout (i, j) avec $i \neq j$, on a :

$$\phi(E_{i,j}) = \phi(E_{i,i}E_{i,j}) = \phi(E_{i,j}E_{i,i}) = \phi(0) = 0.$$

  On obtient donc :

$$\phi(M) = \sum_{i=1}^{n} m_{i,i}\phi(E_{1,1}) = \phi(E_{1,1})\sum_{i=1}^{n} m_{i,i} = \phi(E_{1,1})\operatorname{Tr}(M).$$

□

Trace d'un endomorphisme

Proposition-Définition 83
Soit E un $\mathbb{K}$-e.v. de dimension n. Soit $B = (e_1, \dots, e_n)$ une base de E. Soit $f : E \to E$ un endomorphisme. On définit la **trace** \迹\ de f, notée $\operatorname{Tr}(f)$, comme :

$$\operatorname{Tr}(f) = \operatorname{Tr}(\mathrm{Mat}_B(f)).$$

Cette définition ne dépend pas de la base B choisie.

Preuve — Soit B' une autre base de E, et soit $P \in \mathscr{M}_n(\mathbb{K})$ la matrice de passage de B vers B'. Alors, on a : $\mathrm{Mat}_{B'}(f) = P^{-1}\mathrm{Mat}_B(f)P$. Cela donne :

$$\operatorname{Tr}(\mathrm{Mat}_B(f)) = \operatorname{Tr}\left(P^{-1}\mathrm{Mat}_B(f)P\right) = \operatorname{Tr}\left(\mathrm{Mat}_B(f)PP^{-1}\right) = \operatorname{Tr}(\mathrm{Mat}_B(f)),$$

ce qui prouve que cette trace ne dépend pas de la base choisie. Donc, $\operatorname{Tr}(f)$ est bien défini. □

Exemple 84 — *Soit E un $\mathbb{K}$-e.v. de dimension n. Soit p un projecteur sur E. Alors :*

$$\boxed{\mathrm{rg}(p) = \operatorname{Tr}(p).}$$

En effet, en prenant une base $(e_1, \dots, e_n)$ adaptée à la décomposition en somme directe

$$E = \operatorname{Im}(u) \oplus \operatorname{Ker}(u),$$

alors la matrice de p dans cette base est

$$\begin{pmatrix} I_r & O_{r,n-r} \\ O_{n-r,r} & O_{n-r} \end{pmatrix}.$$

D'où le résultat.

Proposition 85
Soient $\mathbb{K}$ un corps et $n \geq 2$.

Alors le déterminant det(.) et la trace Tr (.) sont des **invariants de similitude** sur $\mathscr{M}_n(\mathbb{K})$:
Pour tout $A \in \mathscr{M}_n(\mathbb{K})$, pour tout $M \in \mathscr{M}_n(\mathbb{K})$ inversible, on a :

$$\det(MAM^{-1}) = \det(A)$$
$$\mathrm{Tr}\left(MAM^{-1}\right) = \mathrm{Tr}\,(A)\,.$$

Autrement dit, pour A et B deux matrices qui sont semblables ($B = MAM^{-1}$ pour un M inversible), alors A et B ont même trace et même déterminant.

Chapitre 3 Diagonalisation

Les notions de cette section concernent les endomorphismes. **Elles concernent aussi les matrices en considérant les endomorphismes canoniquement associés $X \mapsto AX$.** Nous allons prinipalement apprendre à chercher des sous-espaces stables à un endomorphisme afin de mieux le comprendre.

Nous verrons que l'existence de sous-espaces vectoriels stables est très importante dans l'étude géométrique et algébrique d'un endomorphisme. On s'intéresse particulièrement aux sous-espaces vectoriels stables qui sont différents de $\{0\}$ et minimaux, ou aux noyaux des endomorphismes qui s'écrivent comme un "polynôme" en l'endomorphisme u.

Dans ce chapitre, l'ensemble $\mathbb{K}$ désigne un corps. On utilisera souvent $\mathbb{K} = \mathbb{R}, \mathbb{C}$ ou $\mathbb{Q}$. Tous les objets que nous verrons seront définis sur des $\mathbb{K}$-espaces vectoriels E. Par contre, certains résultats ne seront vrais que pour des e.v. E de dimension finie.

Table des matières du chapitre

3.1 Valeurs propres, vecteurs propres & sous-espaces vectoriels propres

Définition 1
Soit E un $\mathbb{K}$-e.v. Soit $u : E \to E$ un endomorphisme. Soient $\lambda \in \mathbb{K}$ et $x \in E$, $x \neq 0$.

1. S'il existe $y \in E$, $y \neq 0$ tel que $u(y) = \lambda y$, on dit alors que λ est une **valeur propre** \特征值\ de u.
2. S'il existe $\gamma \in \mathbb{K}$ tel que $u(x) = \gamma x$, on dit que x est un **vecteur propre** \特征向量\ de u.
3. On définit $\mathrm{Spec}_{\mathbb{K}}(u)$ l'ensemble des valeurs propres de u sur $\mathbb{K}$.
   Cet ensemble est appelé le **spectre** de u sur $\mathbb{K}$.
4. Pour $\lambda \in \mathrm{Spec}_{\mathbb{K}}(u)$, on définit $E_\lambda(u) = \mathrm{Ker}\,(u - \lambda Id_E)$ l'ensemble des vecteurs propres de u associés à la valeur propre λ (auquel on rajoute 0).
   Cet ensemble est appelé **le sous-espace vectoriel propre**\特征子空间\ de u pour la valeur propre λ.

Soient $n \in \mathbb{N}^*$, $A \in \mathscr{M}_n(\mathbb{K})$ une matrice carrée, et $X \in \mathscr{M}_{n,1}(\mathbb{K})$ un vecteur colonne.

1. S'il existe un vecteur colonne $Y \in \mathscr{M}_{n,1}(\mathbb{K})$, $Y \neq 0$ tel que $AY = \lambda Y$, on dit alors que λ est une **valeur propre** \特征值\ de A.
2. S'il existe $\gamma \in \mathbb{K}$ tel que $AX = \gamma X$, on dit que X est un **vecteur propre** \特征向量\ de A.
3. On définit $\mathrm{Spec}_{\mathbb{K}}(A)$ l'ensemble des valeurs propres de A sur $\mathbb{K}$.
   Cet ensemble est appelé le **spectre** de A sur $\mathbb{K}$.

4. Pour $\lambda \in \mathrm{Spec}_{\mathbb{K}}(A)$, on définit $E_\lambda(A) = \mathrm{Ker}\,(A - \lambda I_n)$ l'ensemble des vecteurs propres de A associés à la valeur propre λ (auquel on rajoute 0).
Cet ensemble est appelé **le sous-espace vectoriel propre**\特征子空间\ de A pour la valeur propre λ.

Remarque 2 — *Si E est de dimension n, en prenant B une base de E et en représentant le vecteur $x \in E$ par une vecteur colonne $X \in \mathscr{M}_{n,1}(\mathbb{K})$, on a*

$$u(x) = \lambda x \iff Mat_B(u)X = \lambda X.$$

Ainsi, x est un vecteur propre de u si, et seulement si, X est un vecteur propre de $A = Mat_B(u)$, pour la même valeur propre λ.

Si x est un vecteur propre de u, alors x est associé à une unique valeur propre λ telle que $u(x) = \lambda x$.
Si λ est une valeur propre de u, l'ensemble des vecteurs propres associés est l'ensemble des vecteurs non nuls de $\mathrm{Ker}\,(u - \lambda\,\mathrm{Id}_E)$. ◄

Exemples 3
Soit E un $\mathbb{K}$-e.v. et $u : E \to E$ un endomorphisme.

1. *$E_0(u) = \mathrm{Ker}\,(u)$ est le noyau de u.*
2. *$E_1(u) = \mathrm{Ker}\,(u - \mathrm{Id}_E) = \{x \in E, u(x) = x\}$ est le sous-espace des vecteurs de E invariants par u.*
3. *Pour $u = \lambda Id_E$, on a $Spec_{\mathbb{K}}(u) = \{\lambda\}$. Le sous-espace propre associé à λ est $E_\lambda(u) = E$.*
4. *Pour $E = C^\infty(\mathbb{R}, \mathbb{R})$ et $D : f \mapsto f'$ l'endomorphisme de dérivation, on étudie les valeurs propres de D en résolvant les équations différentielles :*

$$f' = \lambda f, \text{ pour un } \lambda \in \mathbb{R}.$$

La théorie des équations différentielles linéaires du premier ordre montre que tout $\lambda \in \mathbb{R}$ est une valeur propre de D et que le sous-espace vectoriel propre associé à λ est $E_\lambda(D) = Vect(x \mapsto e^{\lambda x})$. Cet endomorphisme a donc une infinité de valeurs propres, et chaque sous-espace propre est de dimension 1.

5. *Soient $E = \mathbb{R}^2$, $t \in \mathbb{R}$, et $A = \begin{pmatrix} \cos(t) & -\sin(t) \\ \sin(t) & \cos(t) \end{pmatrix}$. Alors :*

$$Spec_{\mathbb{R}}(A) = \left| \begin{array}{ll} \{1\} & si\ t = 2k\pi,\ k \in \mathbb{Z}, \\ \{-1\} & si\ t = \pi + 2k\pi,\ k \in \mathbb{Z}, \\ \emptyset & sinon. \end{array} \right.$$

En effet, on a $\det(A - \lambda I_2) = (\cos(t) - \lambda)^2 + \sin(t)^2$. Or, on a $\mathrm{Ker}\,(A - \lambda I_2) \neq \{0\}$ si et seulement si $A - \lambda I_2$ est non-inversible, ssi $\det(A - \lambda I_2) = 0$. C'est-à-dire si et seulement si $\sin(t) = 0$ et $\lambda = \cos(t)$. Lorsque $t = 2k\pi$ on a $A = I_2$, donc $E_1(A) = \mathbb{R}^2$. Lorsque $t = \pi + 2k\pi$ on a $A = -I_2$, donc $E_{-1}(A) = \mathbb{R}^2$. Par contre, pour $\mathbb{K} = \mathbb{C}$ et $E' = \mathbb{C}^2$, le spectre de la matrice A vaut :

$$Spec_{\mathbb{C}}(A) = \{e^{it}, e^{-it}\}.$$

*En effet, sur $\mathbb{C}$ le polynôme $(\cos(t) - X)^2 + \sin(t)^2$ admet pour racines e^{it} et e^{-it}.
Le calcul montre que $(1, -i)$ est un vecteur propre de A pour e^{it}, et que $(1, i)$ est un vecteur propre de A pour e^{-it}.*

Remarques 4

1. *Pour le cas de la dimension 0, l'unique endomorphisme de l'espace vectoriel réduit à $\{0\}$ ne possède pas de valeur propre puisque tout élément de $\mathcal{L}(\{0\})$ est bijectif.*
2. *Un endomorphisme d'un espace vectoriel non réduit à $\{0\}$ n'admet pas nécessairement de valeur propre ; c'est ce que nous avons vu en exemple pour les rotation d'angle $\theta \notin \pi\mathbb{Z}$ sur $\mathbb{R}^2$.*
3. *Pour A une matrice à coefficients réels, on peut avoir $Spec_{\mathbb{R}}(A) \neq Spec_{\mathbb{C}}(A)$. Les valeurs propres de A dépendent du corps $\mathbb{K}$ avec lequel on considère tous nos objets.*
4. *Si au contraire il n'y a pas d'ambiguité sur le corps $\mathbb{K}$, on notera parfois $\mathrm{Spec}\,(u)$ au lieu de $Spec_{\mathbb{K}}(u)$.*

Exemple 5 — *Voici les valeurs propres de quelques endomorphismes :*

1. *Pour une homothétie $h = \lambda\,\mathrm{Id}_E$, on a $\mathrm{Spec}\,(h) = \{\lambda\}$ et $E_\lambda(h) = E$;*

2. *Pour un projecteur p non-trivial ($p \neq 0$, $p \neq Id_E$), on a* $\mathrm{Spec}\,(p) = \{0,1\}$, $E_0(p) = \mathrm{Ker}\,(p)$ *et* $E_1(p) = \mathrm{Im}\,(p)$ *;*
3. *Pour une symétrie s non-triviale ($s \neq Id_E$,$s \neq -Id_E$), on a* $\mathrm{Spec}\,(s) = \{-1,1\}$, $E_{-1}(s) = \mathrm{Ker}\,(s + \mathrm{Id}_E)$ *et* $E_1(s) = \mathrm{Ker}\,(s - \mathrm{Id}_E)$.

Proposition 6
Soit E un $\mathbb{K}$-e.v. de dimension finie. Soient $u : E \to E$ un endomorphisme et $\lambda \in \mathbb{K}$. Alors, on a :

$$\begin{aligned}\lambda \in \mathrm{Spec}\,(u) &\iff \mathrm{Ker}\,(u - \lambda\,\mathrm{Id}_E) \neq \{0_E\} \iff u - \lambda\,\mathrm{Id}_E \text{ non injective}\\ &\iff u - \lambda\,\mathrm{Id}_E \text{ non bijective} \iff \det(u - \lambda\,\mathrm{Id}_E) = 0\end{aligned}$$

Preuve — En dimension finie, un endomorphisme est injectif ssi il est surjectif ssi il est bijectif. □

Remarque 7 — *Pour E un $\mathbb{K}$-e.v. de dimensioon finie, on a donc que u est bijectif si et seulement si* $0 \notin \mathrm{Sp}(u) \iff \det(u) = 0$.
De même, $A \in \mathscr{M}_n(\mathbb{K})$ est inversible si et seulement si $0 \notin \mathrm{Sp}(A) \iff \det A \neq 0$.

Remarque 8 — *Pour A une matrice à coefficients réels, on peut considérer que A appartient à $\mathscr{M}_n(\mathbb{R})$ ou à $\mathscr{M}_n(\mathbb{C})$. Or, pour $\lambda \in \mathbb{R}$ $\det(A - \lambda I_n)$ est un nombre qui ne dépend pas du fait que l'on ait pris $\mathbb{K} = \mathbb{R}$ ou $\mathbb{K} = \mathbb{C}$ (voir chapitre Déterminant).*
C'est-à-dire, $A - \lambda I_n$ est inversible dans $\mathscr{M}_n(\mathbb{R})$ ssi $A - \lambda I_n$ est inversible dans $\mathscr{M}_n(\mathbb{C})$. On a donc :

$$\mathrm{Spec}_{\mathbb{R}}\,(A) = \mathrm{Spec}_{\mathbb{C}}\,(A) \cap \mathbb{R}.$$

De même, pour $B \in \mathscr{M}_n(\mathbb{Q})$, on a :

$$Spec_{\mathbb{Q}}(B) = \mathrm{Spec}_{\mathbb{R}}\,(B) \cap \mathbb{Q} = \mathrm{Spec}_{\mathbb{C}}\,(B) \cap \mathbb{Q}.$$

Exemple 9 —
- *Pour $A = \left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$ on a vu que* $\mathrm{Spec}_{\mathbb{R}}\,(A) = \emptyset$ *mais que* $\mathrm{Spec}_{\mathbb{C}}\,(()\,A) = \{-i, i\}$.
- *Pour $A = \left(\begin{smallmatrix} 0 & 2 \\ 1 & 0 \end{smallmatrix}\right)$ on a* $\det(A - \lambda I_2) = \lambda^2 - 2$. *Ainsi, on a $Spec_{\mathbb{Q}}(A) = \emptyset$ et* $\mathrm{Spec}_{\mathbb{R}}\,(()\,A) = \mathrm{Spec}_{\mathbb{C}}\,(()\,A) = \{-\sqrt{2}, \sqrt{2}\}$.

Proposition 10
Soient E un $\mathbb{K}$-e.v., $u : E \to E$ un endomorphisme, $\phi : E \to E$ un isomorphisme.
Alors, on a $\mathrm{Spec}\,(u) = \mathrm{Spec}\,(\phi \circ u \circ \phi^{-1})$.
De plus, pour tout $\lambda \in \mathrm{Spec}\,(u)$, on a $\dim(E_\lambda(u)) = \dim(E_\lambda(\phi u \phi^{-1}))$.
Soient $n \geq 1$, $A, P \in \mathscr{M}_n(\mathbb{K})$ avec P inversible.
Alors, on a $\mathrm{Spec}\,(A) = \mathrm{Spec}\,(PAP^{-1})$.
De plus, pour tout $\lambda \in \mathrm{Spec}\,(A)$, on a $\dim(E_\lambda(A)) = \dim(E_\lambda(PAP^{-1}))$.
Le spectre d'un endomorphisme et la dimension des sous-espaces propres sont des invariants de similitude.

Preuve — Soit $\phi : E \to E$ un isomorphisme. Soient $x \in E$ non-nul et $\lambda \in \mathbb{K}$. On a :

$$\begin{aligned} u(x) = \lambda x &\iff u \circ \phi^{-1} \circ \phi(x) = \lambda x \\ &\iff (\phi \circ u \circ \phi^{-1})(\phi(x)) = \phi(\lambda x) = \lambda \phi(x). \end{aligned}$$

Comme ϕ est un isomorphisme, on a $\phi(x) \neq 0$ si et seulement si $x \neq 0$.
Donc, x est un vecteur propre pour u associé à λ si et seulement si $\phi(x)$ est un vecteur propre pour $\phi u \phi^{-1}$ associé à λ.

Ces endomorphismes ont donc les mêmes valeurs propres, et l'on a :

$$E_\lambda\left(\phi \circ u \circ \phi^{-1}\right) = \phi\left(E_\lambda(u)\right).$$

Leurs sous-espaces propres associés à λ sont donc isomorphes, ce qui implique qu'ils ont la même dimension. □

Remarque 11 — *Soit E un $\mathbb{K}$-e.v. de dimension n. Soit $u : E \to E$ un endomorphisme. La proposition précédente nous dit que pour n'importe quelle base B de E, la matrice $Mat_B(u)$ possède le même spectre que u, et que leurs sous-espaces propres associés à une valeur propre donnée sont de même dimension.*
On peut ainsi étudier le spectre et les espaces propres de u, ou ceux de $Mat_B(u)$. Ce résultat est très utile s'il existe une base B telle que $Mat_B(u)$ a une expression qui permet de calculer plus facilement $\det(Mat_B(u) - \lambda I_n)$.

3.2 Polynôme caractéristique

On rappelle que pour $\mathbb{K}$ un corps, l'ensemble $\mathbb{K}(X)$ est le corps des fractions rationnelles à coefficients dans $\mathbb{K}$. Cet ensemble est un corps, qui contient $\mathbb{K}[X]$.

3.2.1 Définition du polynôme caractéristique

Proposition-Définition 12
Soient $n \geq 1$ et $A \in \mathscr{M}_n(\mathbb{K})$.
Alors, le déterminant de la matrice $XI_n - A \in \mathscr{M}_n(\mathbb{K}(X))$, $\det(XI_n - A)$, est un polynôme. On le note $\chi_A(X)$.
Le polynôme $\chi_A(X)$ est appelé **polynôme caractéristique** \特征多项式\ **de** A.
$\chi_A(X)$ est un polynôme unitaire, de degré n, avec :

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

Preuve — Soit $A = (\alpha_{i,j})$. La matrice $(XI_n - A)$ est une matrice à coefficients dans le corps des fractions rationnelles $\mathbb{K}(X)$. Son déterminant, donné par :

$$\chi_A(X) = \sum_{\sigma \in \mathcal{S}_n} \varepsilon(\sigma)\left(X\delta_{\sigma(1),1} - \alpha_{\sigma(1),1}\right) \ldots \left(X\delta_{\sigma(n),n} - \alpha_{\sigma(n),n}\right),$$

est une somme de produits de polynômes. C'est donc un polynôme.
De plus, ce polynôme est de degré inférieur ou égal à n.

Soit $\sigma \in \mathcal{S}_n$. Si σ n'est pas l'identité, alors il existe au moins deux éléments distincts $i, j \in [\![1, n]\!]$ tels que $\sigma(i) \neq i$ et $\sigma(j) \neq j$. On a donc $\delta_{\sigma(i),i} = 0 = \delta_{\sigma(j),j}$, et le terme :

$$\varepsilon(\sigma)\left(X\delta_{\sigma(1),1} - \alpha_{\sigma(1),1}\right) \ldots \left(X\delta_{\sigma(n),n} - \alpha_{\sigma(n),n}\right)$$

est donc un polynôme de degré inférieur ou égal à $n-2$.

Les termes de degré n et $n-1$ de $\chi_A(X)$ sont donc ceux du produit :

$$(X - \alpha_{1,1}) \ldots (X - \alpha_{n,n}),$$

c'est-à-dire X^n et $-(\alpha_{1,1} + \cdots + \alpha_{n,n})X^{n-1} = -\operatorname{Tr}(A)X^{n-1}$.

Enfin, le terme constant de $\chi_A(X)$ vaut $\chi_A(0) = \det(-A) = (-1)^n \det(A)$. □

Proposition 13
Soient $n \geq 1$ et $A, P \in \mathscr{M}_n(\mathbb{K})$ avec P inversible.
Alors, on a $\chi_{PAP^{-1}}(X) = \chi_A(X)$.
Le polynôme caractéristique est un invariant de similitude.

Preuve — On a :
$$\det\left(XI_n - PAP^{-1}\right) = \det\left(XPI_nP^{-1} - PAP^{-1}\right) = \det\left(P\left(XI_n - A\right)P^{-1}\right) = \det(P)\det(XI_n - A)\det\left(P^{-1}\right) = \det(XI_n - A).$$
□

Proposition-Définition 14
Soit E un $\mathbb{K}$-e.v. de dimension n. Soit $u :\in \mathcal{L}(E)$ un endomorphisme.
Alors il existe un unique polynôme, noté $\chi_u(X)$, tel que pour toute base B de E on ait $\chi_u(X) = \chi_{\operatorname{Mat}_B(u)}(X)$.
Ce polynôme est appelé le **polynôme caractéristique de** u.
Le polynôme χ_u est unitaire, de degré n, avec :

$$\chi_u(X) = X^n - (\operatorname{Tr}(u))\,X^{n-1} + \cdots + (-1)^n \det(u).$$

Pour tout $\lambda \in \mathbb{K}$, on a :

$$\chi_u(\lambda) = \det(\lambda \operatorname{Id}_E - u).$$

Preuve — Pour B et B' deux bases de E, les matrices $\operatorname{Mat}_B(u)$ et $\operatorname{Mat}_{B'}(u)$ sont semblables. D'après la proposition précédente on a donc $\chi_{\operatorname{Mat}_B(u)}(X) = \chi_{\operatorname{Mat}_{B'}(u))}(X)$. Les propriétés de $\operatorname{Mat}_B(u)$ permettent alors de conclure. □

Exemples 15
1. *On peut éventuellement poser la convention suivante : le polynôme caractéristique de l'endomorphisme d'un espace vectoriel réduit à* $\{0\}$ *ainsi que celui de la matrice vide, est le polynôme* 1.
2. *Le polynôme caractéristique de* $(\alpha) \in \mathscr{M}_1(\mathbb{K})$ *est* $X - \alpha$.
3. *Le polynôme caractéristique de* $A = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \in \mathscr{M}_2(\mathbb{K})$ *est :*
$$\chi_A(X) = X^2 - (\alpha + \delta)X + (\alpha\delta - \beta\gamma) = X^2 - \operatorname{Tr}(A)\,X + \det(A).$$
4. *Le polynôme caractéristique de la matrice :*
$$A = \begin{pmatrix} \alpha & \beta & \gamma \\ \alpha' & \beta' & \gamma' \\ \alpha'' & \beta'' & \gamma'' \end{pmatrix}$$
de $\mathscr{M}_3(\mathbb{K})$ *est :* $X^3 - \operatorname{Tr}(A)\,X^2 + \left((\alpha\beta' - \alpha'\beta) + (\alpha\gamma'' - \alpha''\gamma) + (\beta'\gamma'' - \beta''\gamma')\right)X - \det(A)$.

5. *Pour $A \in \mathscr{M}_n(\mathbb{C})$, on a $\chi_A = \overline{\chi_{\overline{A}}}$. complexes.*
6. *Soit E un $\mathbb{K}$-e.v. de dimension n. Soit $u : E \to E$ un endomorphisme de rang 1. Dans une base B adaptée au sous-espace vectoriel $\mathrm{Im}\,(u)$, la matrice de u est de la forme :*

$$Mat_B(u) = \begin{pmatrix} \alpha_{1,1} & \dots & \dots & \alpha_{1,n} \\ 0 & \dots & \dots & 0 \\ \vdots & & \vdots & \\ 0 & \dots & \dots & 0 \end{pmatrix}.$$

Son polynôme caractéristique est donc $X^{n-1}(X - \alpha_{1,1})$. Comme $\alpha_{1,1} = \mathrm{Tr}\,(Mat_B(u)) = \mathrm{Tr}\,(u)$, on obtient :

$$\chi_u(X) = X^{n-1}\,(X - \mathrm{Tr}\,(u))\,.$$

Remarque 16 — *Soit $A \in \mathscr{M}_n(\mathbb{K})$. Alors on a $\det(XI_n - {}^tA) = \det({}^tXI_n - A) = \det(XI_n - A)$. Donc, $\chi_{{}^tA}(X) = \chi_A(X)$.*

Exemples 17 (Matrices triangulaires)

1. *Soit A une matrice triangulaire supérieure, de diagonale $(\alpha_1, \dots, \alpha_n)$. On a alors :*

$$\chi_A(X) = \begin{vmatrix} X-\alpha_1 & \dots & \dots & * \\ 0 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & X-\alpha_n \end{vmatrix} = \prod_{k=1}^{n} (X - \alpha_k)\,.$$

2. *Pour M une matrice triangulaire supérieure par blocs de la forme :*

$$M = \begin{pmatrix} A & B \\ 0 & D \end{pmatrix},$$

on a :

$$\chi_M(X) = \begin{vmatrix} XI_p - A & -C \\ 0 & XI_q - D \end{vmatrix} = \chi_A(X)\chi_D(X).$$

Ces résultats sont aussi vrais pour des matrices triangulaires inférieures.

Proposition 18
Soit $n \geq 1$. Soient $n_1, \dots, n_r \in \{1, \dots, n\}$ tels que $n_1 + \dots, n_r = n$. Soit $M \in \mathscr{M}_n(\mathbb{K})$ une matrice triangulaire supérieure par blocs :

$$M = \begin{pmatrix} A_1 & * & \dots & * \\ 0 & A_2 & * & \vdots \\ \vdots & \ddots & \ddots & * \\ 0 & \dots & 0 & A_r \end{pmatrix}, \text{ pour } A_i \in \mathscr{M}_{n_i}(\mathbb{K}).$$

Alors, on a $\chi_M(X) = \chi_{A_1}(X) \dots \chi_{A_r}(X) = \Pi_{i=1}^{r} \chi_{A_i}(X)$.
Le résultat est aussi vrai si M est une matrice triangulaire inférieure par blocs.

Preuve — On démontre le résultat par récurrence sur $r \geq 2$, en utilisant le fait que la matrice M est aussi de la forme $M = \begin{pmatrix} A_1 & B \\ 0 & A' \end{pmatrix}$, où A' est une matrice triangulaire supérieure par blocs avec $r-1$ blocs diagonaux, et que $XI_n - M$ est aussi une matrice triangulaire supérieure par blocs, avec r blocs diagonaux.
Si M est une matrice triangulaire inférieure par blocs, alors ${}^tM$ est une matrice triangulaire supérieure par blocs. La propriété $\chi_{{}^tM}(X) = \chi_M(X)$ permet alors de conclure. $\square$

3.2.2 Polynôme caractéristique et valeurs propres

Proposition 19
Soit E un $\mathbb{K}$-e.v. de dimension finie. Soit $u : E \to E$ un endomorphisme.
Alors, le spectre de u est égal à l'ensemble des racines du polynôme caractéristique $\chi_u(X)$.
Soit $n \geq 1$. Soit $A \in \mathscr{M}_n(\mathbb{K})$.
Alors, on a $\mathrm{Spec}\,(A) = \{\lambda \in \mathbb{K},\, \chi_A(\lambda) = 0\}$.

Preuve — On a montré qu'en dimension finie, on a $\lambda \in \operatorname{Spec}(u)$ si et seulement si $\det(u-\lambda Id_E) = 0$. Or, on a $\chi_u(\lambda) = \det(\lambda Id_E - u) = (-1)^n \det(u - \lambda Id_E)$.
La preuve est identique pour une matrice A. □

Remarque 20 — *Ainsi, un seul calcul de déterminant peut suffire pour déterminer toutes les valeurs propres d'une matrice A (ou d'un endomorphisme u), du moment que l'on arrive à factoriser le polynôme caractéristique $\chi_A(X)$.*

Corollaire 21
Soit E un $\mathbb{K}$-e.v. de dimension finie. Soient $u : E \to E$ un endomorphisme et $A \in \mathscr{M}_n(\mathbb{K})$.
Alors u (ou A) possède au plus n valeurs propres distinctes.

- Si $\mathbb{K} = \mathbb{C}$, alors u (ou A) a au moins une valeur propre.
- Si $\mathbb{K} = \mathbb{R}$ et si n est impair, alors u (ou A) a au moins une valeur propre.

Preuve — Les valeurs propres de u sont les racines de $\chi_u(X)$. Comme ce polynôme est de degré n, il possède au plus n racines distinctes. □

Le corollaire précédent montre d'obtenir le polynôme caractéristique sous forme factorisée est très important en pratique. On factorise en général $\chi_A(X)$ en calculant calculant le déterminant $\det(XI_n - A)$ par opérations élémentaires afin de faire apparaître des facteurs communs dans les lignes ou les colonnes.

Exemple 22 — *Soit*

$$A = \begin{pmatrix} 2 & 5 & -6 \\ 4 & 6 & -9 \\ 3 & 6 & -8 \end{pmatrix}.$$

Le polynôme caractéristique de A est :

$$\chi_A(X) = \begin{vmatrix} X-2 & -5 & 6 \\ -4 & X-6 & 9 \\ -3 & -6 & X+8 \end{vmatrix}.$$

On a : La somme des coefficients des lignes du déterminant ci-dessus étant $X-1$, l'opération $C_1 \leftarrow C_1 + C_2 + C_3$ donne :

$$\begin{vmatrix} X-2 & -5 & 6 \\ -4 & X-6 & 9 \\ -3 & -6 & X+8 \end{vmatrix} \underset{C_1 \leftarrow C_1+C_2+C_3}{=} (X-1) \begin{vmatrix} 1 & -5 & 6 \\ 1 & X-6 & 9 \\ 1 & -6 & X+8 \end{vmatrix} \underset{\substack{L2 \leftarrow L_2 - L_1 \\ L_3 \leftarrow L_3 - L_1}}{=} (X-1) \begin{vmatrix} 1 & -5 & 6 \\ 0 & X-1 & 3 \\ 0 & -1 & X+2 \end{vmatrix}.$$

On obtient donc :

$$\chi_A(X) = (X-1)((X-1)(X+2)+3) = (X-1)(X^2+X+1).$$

Le spectre de A est donc $\left\{1, j, j^2\right\}$ dans $\mathbb{C}$ et seulement $\{1\}$ dans $\mathbb{R}$.

Définition 23
Soit E un $\mathbb{K}$-e.v. de dimension finie. Soit $u : E \to E$ un endomorphisme. Soit $\lambda \in \mathbb{K}$ une valeur propre de u.
On définit $m_u(\lambda)$ la multiplicité du facteur $(X - \lambda)$ pour le polynôme caractéristique $\chi_u(X)$.
L'entier $m(\lambda)$ est appelé **multiplicité de la valeur propre** λ, pour u.
Pour $n \geq 1$, $A \in \mathscr{M}_n(\mathbb{K})$, et $\lambda \in \operatorname{Spec}(A)$, on définit de même $m_A(\lambda)$ la multiplicité du facteur $(X - \lambda)$ pour le polynôme caractéristique $\chi_A(X)$.

Exemple 24 —

- *Soit $A \in \mathscr{M}_5(\mathbb{K})$ une matrice triangulaire supérieure dont les coefficients diagonaux sont* $(1, 1, 2, 2, 3)$.
  On a alors $\chi_A(X) = (X-1)(X-1)(X-2)(X-2)(X-3)$. On a donc $\operatorname{Spec}(A) = \{1, 2, 3\}$, avec $m_A(1) = 2$,$m_A(2) = 2$,$m_A(3) = 1$.

- *Pour $C = I_2$, on a $\chi_C(X) = (X-1)^2$, donc $\operatorname{Spec}(C) = \{1\}$, avec $m_C(1) = 2$. On a aussi $E_1(B) = \operatorname{Ker}(B - I_2) = \mathbb{K}^2$, donc $\dim(E_1(C)) = 2$.*

- *Soit $B = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$. On a $\chi_B(X) = (X-1)^2$, donc $\operatorname{Spec}(B) = \{1\}$, avec $m_B(1) = 2$. Par contre, on a $E_1(B) = \operatorname{Ker}(B - I_2) =$ Vect$((1,0))$, donc $\dim(E_1(B)) = 1$.*
  On a ici $\dim(E_1(B)) \neq m(1)$. On peut de plus remarquer que B n'est pas semblable à I_2 car $\dim(E_1(B)) \neq \dim(E_1(I_2))$.

Proposition 25
Soient E un $\mathbb{K}$-e.v., $u : E \to E$ un endomorphisme, $\phi : E \to E$ un isomorphisme. Soit $\lambda \in \mathrm{Spec}\,(u)$.
Alors, on a $m_u(\lambda) = m_{\phi^{-1}u\phi}(\lambda)$. Soient $n \geq 1$ et $A, P \in \mathscr{M}_n(\mathbb{K})$ avec P inversible. Soit $\lambda \in \mathrm{Spec}\,(A)$.
Alors, on a $m_A(\lambda) = m_{P^{-1}AP}(\lambda)$.
La multiplicité des valeurs propres est un invariant de similitude.

Preuve — On a vu que u et $\phi^{-1}u\phi$ ont le même spectre et le même polynôme caractéristique, ce qui conclut. Même chose pour A et $P^{-1}AP$. □

Remarque 26 — *Soient $A, B \in \mathscr{M}_n(\mathbb{K})$ deux matrices carrées (respectivement $u, v : E \to E$ deux endomorphismes).*
Si A et B (resp. u et v) n'ont pas le même spectre ou pas le même polynôme caractéristique ou pas les mêmes multiplicités de valeurs propres ou pas les mêmes dimensions d'espaces propres, alors A et B (resp. u et v) ne sont pas semblables.
En effet, tous ces objets sont des invariants de similitudes.
Par contre, si deux matrices A et B qui ont le même spectre/poly. caractéristique/multiplicités des valeurs propres/dimensions des sous-espaces propres, on ne sait pas si A et B sont semblables ou non.

Exemple 27 (Exemple avec des matrices nilpotentes) — *Soient* $A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$ *et* $B = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$.
Les matrices A et B sont triangulaires supérieures par blocs, avec deux blocs diagonaux de taille 2×2. Le calcul donne alors :

$$\chi_A(X) = X^2 \times X^2 = X^4 \text{ et } \chi_B(X) = X^2 \times X^2 = X^4.$$

On a donc $\chi_A(X) = \chi_B(X)$. De plus, on a $\mathrm{Spec}\,(A) = \mathrm{Spec}\,(B) = \{0\}$, et $m_A(0) = 4 = m_B(0)$.
Le calcul donne : $E_0(A) = Vect(e_1, e_3)$ et $E_0(B) = Vect(e_1, e_2)$. On a donc $\dim(E_0(A)) = 2 = \dim(E_0(B))$.
Par contre, on a $A^2 = 0$ et $B^2 = E_{1,4} \neq 0$. Les matrices A et B ne sont donc pas semblables, car on aurait sinon $B^2 = (P^{-1}A)^2 = P^{-1}A^2P = 0$, ce qui n'est pas le cas.

Cet exemple montre qu'il existe des matrices qui ont le même spectre/poly. caractéristique/multiplicités des valeurs propres/dimensions des sous-espaces propres, mais qui ne sont pas semblables.

Remarque 28 — *Pour u un endomorphisme et $\lambda \in \mathrm{Spec}\,(u)$, attention à ne pas confondre $m_u(\lambda)$ la multiplicité de $(X - \lambda)$ dans $\chi_u(X)$ et $\dim(E_\lambda(u))$ la dimension du sous-espace propre associé à λ.*
Ces deux quantités ne sont en général pas égales.

3.3 Sous-espaces vectoriels stables par un endomorphisme

3.3.1 Définition

Définition 29
Soient E un $\mathbb{K}$-e.v. et $u : E \to E$ un endomorphisme. Soit F un sous-espace vectoriel de E.
Si l'on a $u(F) \subset F$, on dit alors que le sous-e.v. F est **stable par** u, ou u-stable. \一般称F为u的不变子空间，简称u-子空间。\

Exemples 30

1. *Pour tout $u : E \to E$ endomorphisme, les sous-espaces vectoriels $\{0\}$ et E sont stables par u.*
2. *Il existe des endomorphismes u dont les seuls sous-e.v. stables sont $\{0\}$ et E. Par exemple, $R_{\frac{\pi}{2}} : \mathbb{R}^2 \to \mathbb{R}^2$ la rotation d'angle $\frac{\pi}{2}$ dans $\mathbb{R}^2$ ne possède aucun sous-espace stable non-trivial (différent de $\{0\}$ et $\mathbb{R}^2$).*
3. *Pour $u : x \mapsto \lambda x$ une homothétie de rapport λ, on vérifie que tout sous-ev F de E est stable par u.*
   On peut montrer que les seuls endomorphismes v de E tels que tout sous-e.v. de E est stable par v sont les homothéties. ◀
4. *Soit $u : E \to E$ un endomorphisme. Pour $F \subset \mathrm{Ker}\,(u)$, alors F est stable par u (car $u(F) = \{0\}$). Pour $G \supset Im(u)$, alors G est stable par u (car $u(G) \subset Im(u)$).*
5. *Soit $u : E \to E$ un endomorphisme. Pour F, G deux sous-ev stables par u, alors $F \cap G$ et $F + G$ sont des sous-ev stables par u.* ◀

3.3.2 Endomorphisme induit par stabilité

Définition

DÉFINITION 31
Soient E un $\mathbb{K}$-e.v., $u : E \to E$ un endomorphisme, et F un sous-e.v. de E.
Si F est stable par u, on peut alors définir la fonction $u_F : x \in F \mapsto u(x) \in F$.
Cete fonction est un endomorphisme, appelé **l'endomorphisme induit par u sur F**.

REMARQUE 32 — *Le fait que $u(F) \subset F$ permet de remplacer l'espace d'arrivée E par son sous-espace vectoriel F.*
Attention à ne pas confondre la restriction $u_{|F} : F \to E$ de u à F, qui est une application linéaire de F vers E que l'on peut définir pour tout u, et l'endomorphisme $u_F : F \to F$ induit par u sur F, qui est un endomorphisme sur F que l'on ne peut définir que lorsque F est stable par u.
On notera que l'image de u_F est $Im(u_F) = u(F)$ et que le noyau de u_F est $\operatorname{Ker}(u_F) = \operatorname{Ker}(u) \cap F$. ◄

PROPOSITION 33
Soient E un $\mathbb{K}$-e.v., $u : E \to E$ un endomorphisme, et F un sous-e.v. de E. Si F est stable par u, alors les valeurs propres de l'endomorphisme u_F induit par u sur F sont les valeurs propres de u telles que $E_\lambda(u) \cap F \neq \{0\}$.
On a alors :

$$E_\lambda(u_F) = E_\lambda(u) \cap F.$$

Preuve — Soit F est stable par u. Alors pour tout $\lambda \in \mathbb{K}$, F est stable par $u - \lambda Id_E$. On a $(u - \lambda Id_E)_F = u_F - \lambda Id_F$.
Avec la remarque précédente, on obtient : $\operatorname{Ker}(u_F - \lambda Id_F) = \operatorname{Ker}(u - \lambda Id_F) \cap F$. Ce qui permet de conclure. □

À l'aide des vecteurs propres d'un endomorphisme u, il est facile de construire certains sous-espaces stables par u.

PROPOSITION 34 (Sous-espaces stables engendrés par des vecteurs propres)
Soient E un $\mathbb{K}$-e.v. et $u : E \to E$ un endomorphisme. Soient $v_1, \dots, v_k$ des vecteurs propres de u (pas forcément de même valeur propre). Alors $F = \operatorname{Vect}(v_1, \dots, v_k)$ est stable par u.

Preuve — Soit $x \in F$. Alors on peut écrire $x = \sum_{j=1}^k \alpha_j v_j$. Notons λ_j la valeur propre associée à v_j. On a alors :

$$u(x) = u\left(\sum_{j=1}^k \alpha_j v_j\right) = \sum_{j=1}^k \alpha_j u(v_j) = \sum_{j=1}^k (\alpha_j \lambda_j) v_j \in F.$$

□

Endomorphisme induit et polynôme caractéristique Si un sous-espace vectoriel F est stable par u, cela peut se voir sur certaines représentations matricielles de u.

PROPOSITION 35 (Sous-espaces stables et représentation matricielle)
Soient E un $\mathbb{K}$-e.v. de dimension n et $u : E \to E$ un endomorphisme. Soit F un sous-espace de E. Soit $\mathcal{B} = (e_1, \dots, e_n)$ une base de E telle que $F = \operatorname{Vect}(e_1, \dots, e_p)$ pour un certain $p \in [\![1, n]\!]$.
Alors, le sous-e.v. F est stable par u si et seulement si la matrice $\operatorname{Mat}_{\mathcal{B}}(u)$ est de la forme :

$$\operatorname{Mat}_{\mathcal{B}}(u) = \begin{pmatrix} a_{1,1} & \cdots & * & * & \cdots & * \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{p,1} & \cdots & a_{p,p} & a_{p,p+1} & \cdots & * \\ 0 & \cdots & 0 & a_{p+1,p+1} & \cdots & * \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & a_{n,p+1} & \cdots & a_{n,n} \end{pmatrix}$$

PreuvePosons $\operatorname{Mat}_{\mathcal{B}}(u) = (a_{i,j})_{i,j}$.

- Si F est stable par u, alors on a
$$u(e_j) \in F, \forall j \in [\![1, p]\!]$$
Donc :
$$\sum_{i=1}^n a_{i,j} e_i \in \operatorname{Vect}(e_1, \dots, e_p), \forall j \in [\![1, p]\!]$$
Ceci implique que :
$$a_{i,j} = 0, \forall j \in [\![1, p]\!], \forall i \in [\![p+1, n]\!],$$
ce qui est la forme voulue (triangulaire supérieure par blocs, avec un bloc $0_{n-p,p}$ en bas à gauche).

- Réciproquement, supposons que
$$a_{i,j} = 0, \forall j \in [\![1,p]\!], \forall i \in [\![p+1,n]\!].$$
On a alors :
$$\sum_{i=1}^{n} A_{i,j} e_i \in \text{Vect}\,(e_1,\ldots,e_p)\,, \forall j \in [\![1,p]\!].$$
Cela veut dire que $u(e_1),\ldots,u(e_p) \in F = \text{Vect}(e_1,\ldots,e_p)$. Donc F est stable par u.

$\square$

Proposition 36 (Endomorphisme induit et polynôme caractéristique)
Soient E un $\mathbb{K}$-e.v. de dimension n et $u : E \to E$ un endomorphisme. Soit F un sous-espace de E stable par u. Alors, $\chi_{u_F}(X)$ divise $\chi_u(X)$, le polynôme caractéristique de u.

Preuve — Soit $\mathcal{B} = (e_1,\ldots,e_n)$ une base de E telle que $F = \text{Vect}(e_1,\ldots,e_p)$. La matrice de u dans $\mathcal{B}$ est aors de la forme :
$$\text{Mat}_{\mathcal{B}}(u) = \begin{pmatrix} A & B \\ 0_{n-p,p} & D \end{pmatrix}.$$
On a alors $A = \text{Mat}_{(e_1,\ldots,e_p)}(u_F)$, par définition de l'endomorphisme induit u_F. Le polynôme caractéristique de u vérifie donc $\chi_u(X) = \chi_{\text{Mat}_{\mathcal{B}}(u)}(X) = \chi_A(X)\chi_D(X)$. Ce polynôme est donc divisible par $\chi_A(X)$, qui est le polynôme caractéristique de u_F, ce qui conclut. $\square$

Ainsi, trouver des sous-espaces F stables par un endomorphisme u permet de trouver des diviseurs du polynôme caractéristque de u $\chi_u(X)$. Cela aide à factoriser $\chi_u(X)$.

Remarque 37 — *Soient E un $\mathbb{K}$-e.v. de dimension n, et $u : E \to E$ un endomorphisme. Si $\chi_u(X)$ ne possède aucun diviseur de degré p, alors aucun sous-ev F de E de dimension p n'est stable par u.*
En effet, pour F sous-ev stable par u, $\chi_{u_F}(X)$ est de degré $\dim(F)$ et est un diviseur de χ_u.
De plus, si $\chi_u(X)$ est irréductible dans $\mathbb{K}[X]$, alors u ne possède aucun sous-ev stable non trivial (différent de $\{0\}$ et E).
Par exemple, pour $A = \begin{pmatrix} 0 & 2 \\ 1 & 0 \end{pmatrix}$, on a $\chi_A(X) = X^2-2$. Ce polynôme est irréductible dans $\mathbb{Q}[X]$, donc l'application $X \mapsto AX$ ne possède aucun sous-ev stable non trivial dans $\mathbb{Q}^2$. Dans $\mathbb{R}$ on a par contre $\chi_A(X) = (X-\sqrt{2})(X+\sqrt{2})$ donc A possède des valeurs propres, donc A possède des sous-ev stables non triviaux.

Proposition 38
Soient E un $\mathbb{K}$-e.v. de dimension n, et $u : E \to E$ un endomorphisme. Soit $\lambda \in \text{Spec}\,(u)$. Alors, on a :
$$1 \leq \dim(E_\lambda(u)) \leq m_u(\lambda).$$

Preuve — Soit $\lambda \in \text{Spec}\,(u)$. Le sous-espace propre $F = E_\lambda(u)$ étant-non réduit à $\{0\}$, il est de dimension au moins 1. Ce sous-espace vectoriel est aussi stable par u.
On remarque que l'endomorphisme induit u_F vérifie : $u_F = \lambda Id_F$.

Ainsi, le polynôme caractéristique de u_F est $\chi_{u_F}(X) = (X-\lambda)^{\dim(E_\lambda(u))}$.
Comme ce polynôme divise $\chi_u(X)$, on obtient $\dim(E_\lambda(u)) \leqslant m(\lambda)$. $\square$

Exemples 39
1. *Si $\text{rg}(u) = r$ on a $\dim E_0(u) = \text{Ker}\,(u) = n-r$, donc $\chi_u(X)$ est divisible par X^{n-r}.* ◀
2. *La matrice :*
$$A = \begin{pmatrix} 0 & \ldots & 0 & \alpha_1 \\ \vdots & & \vdots & \vdots \\ 0 & \ldots & 0 & \alpha_{n-1} \\ \alpha_1 & \ldots & \alpha_{n-1} & 0 \end{pmatrix}$$
est de rang égal à 2 (ou 0). Donc, son polynôme charactéristique est de la forme :
$$\chi_A(X) = X^{n-2}\left(X^2+uX+v\right) = X^n + uX^{n-1} + vX^{n-2}.$$
On a de plus $u = -\,\text{Tr}\,(A) = 0$. En développant le déterminant :
$$\det(XiI_n - A) = \begin{vmatrix} X & \ldots & 0 & -\alpha_1 \\ \vdots & & \vdots & \vdots \\ 0 & \ldots & X & -\alpha_{n-1} \\ -\alpha_1 & \ldots & -\alpha_{n-1} & X \end{vmatrix}$$

les termes non nuls de degré $n-2$ sont obtenus pour les transpositions $\tau = (i, n)$ avec $i < n$. On obtient donc :

$$v = -\alpha_1^2 - \cdots - \alpha_{n-1}^2.$$

Ainsi, on a $\chi_A(X) = X^{n-2}(X^2 - (\alpha_1^2 + \ldots + \alpha_n^2))$.

§ 1. *Sous-espaces propres et commutativité*

PROPOSITION 40
Soient E un $\mathbb{K}$-e.v. et $u, v : E \to E$ des endomorphismes.
Si u et v commutent ($u \circ v = v \circ u$), alors tout sous-espace propre de v, $E_\lambda(v)$, est stable par u.

Preuve — Soit $\lambda \in \operatorname{Spec}(v)$. Soit $x \in E_\lambda(v)$. On a $v(x) = \lambda x$.
Cela donne :

$$v(u(x)) = u(v(x)) = u(\lambda x) = \lambda u(x),$$

donc $u(x) \in E_\lambda(v)$, ce qui conclut. □

REMARQUE 41 — *Attention, ce résultat est faux pour un sous-espace stable F de v quelconque. Par exemple, pour $E = \mathbb{K}^2$, $u(x, y) = (y, x)$, $v = Id_E$, le sous-ev $F = Vect((1, 0))$ est stable par v, mais il n'est pas du tout stable par u.*
Il faut bien remarquer dans la preuve de la proposition précédente que $F = E_\lambda(v)$ et que $v(x) = \lambda x$ sont des informations nécessaires.

3.3.3 Sous-espaces cycliques

PROPOSITION-DÉFINITION 42
Soient E un $\mathbb{K}$-e.v., $u : E \to E$ un endomorphisme, et $x \in E$.
On définit $S_u(x) = \operatorname{Vect}(u^k(x),\ k \in \mathbb{N})$.
Ce sous-espace vectoriel est appelé **le sous-espace cyclique** engendré par x \由x生成的循环子空间\.
C'est le plus petit sous-espace vectoriel qui contient x et qui est stable par u.

Preuve —
"⊃" Soit $a \in S_u(x)$, On a $a_0, \ldots, a_n \in \mathbb{K}$ tels $a = a_0 x + a_1 u(x) + \ldots + a_n u^n(x)$. Ainsi, on a $u(a) = a_0 u(x) + a_1 u^2(x) + \ldots + a_n u^{n+1}(x)$. Donc $u(a) \in S_u(x)$. Ce sous-ev est bien stable par u, et il contient x.
"⊂" Soit F un sous-ev stable par u et contenant x. Alors F contient $x, u(x), u^2(x), \ldots$, donc F contient $\operatorname{Vect}(x, u(x), u^2(x), \ldots) = S_u(x)$.

□

REMARQUE 43 — *En fait, pour F un sous-ev de E contenant x et stable par u, on a :* ◄

$$\operatorname{Vect}(x, u(x), \cdots, u^p(x)) \subset S_u(x) \subset F \subset E,\ \forall p \in \mathbb{N}$$

PROPOSITION 44
Soient E un $\mathbb{K}$-e.v. de dimension n, $u : E \to E$ un endomorphisme. Soit $x \in E$ non-nul.
Soit p le plus grand entier tel que la famille $\mathcal{B} = (x, u(x), \cdots, u^p(x))$ est libre.
Alors, on a $S_u(x) = \operatorname{Vect}(\mathcal{B})$. Donc $\dim(S_u(x)) = p + 1$.
De plus, pour $a_0, \ldots, a_p \in \mathbb{K}$ tels que $u^{p+1}(x) = -a_0 x - a_1 u(x) - \ldots - a_p u^p(x)$, on a :

$$\operatorname{Mat}_{\mathcal{B}}(u_{S_u(x)}) = \begin{pmatrix} 0 & \cdots & \cdots & \cdots & 0 & -a_0 \\ 1 & \ddots & & & \vdots & -a_1 \\ 0 & \ddots & \ddots & & \vdots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots & \vdots \\ 0 & \cdots & 0 & 1 & 0 & -a_{p-1} \\ 0 & \cdots & 0 & 0 & 1 & -a_p \end{pmatrix}.$$

Preuve — Comme $x \neq 0$, la famille (x) est libre. Comme $\dim(E) = n$, la famille $(x, u(x), \cdots, u^n(x))$ n'est pas libre. Donc il existe un plus grand entier p tel que la famille $(x, u(x), \cdots, u^p(x))$ soit libre. On a de plus $p < n$.
Posons $\mathcal{B} = (x, u(x), \cdots, u^p(x))$ et $F = \operatorname{Vect}(\mathcal{B})$.
Par définition de l''entier p, la famille $\big(x, u(x), \cdots, u^p(x), u^{p+1}(x)\big)$ est liée. Il existe donc $p+2$ coefficients $\alpha_0, \alpha_1, \ldots, \alpha_{p+1} \in \mathbb{K}$, non tous nuls, tels que

$$\alpha_0 x + \alpha_1 u(x) + \cdots + \alpha_p u^p(x) + \alpha_{p+1} u^{p+1}(x) = 0_E.$$

On a de plus $\alpha_{p+1} \neq 0$ car sinon la famille $\mathcal{B}$ serait liée. Quitte à diviser tous les coefficients par α_{p+1}, on peut supposer que $\alpha_{p+1} = 1$. On a donc des coefficients $a_0, a_1, \cdots, a_p \in \mathbb{K}$ de $\mathbb{K}$ tels que $a_0 x + a_1 u(x) + \cdots a_p u^p(x) + u^{p+1}(x) = 0_E$. On en déduit que $u^{p+1}(x) \in F$.

Cela permet de démontrer par récurrence sur $k \geq 0$ que $u^k(x) \in F$. En effet on a $u^0(x) = x \in F$, et si $u^k(x) \in F$ alors $u^k(x)$ est une combinaison linéaire de $x, u(x), \ldots, u^p(x)$. Donc $u^{k+1}(x)$ est une combinaison linéaire de $u(x), \ldots, u^{p+1}(x)$, qui sont des éléments de F.

Ainsi, on a donc $S_u(x) = \mathrm{Vect}(u^k(x),\, k \in \mathbb{N}) \subset F$, et la remarque précédente nous donne l'inclusion réciproque.
Comme $F = S_u(x)$, F est un sous-espace stable par u. La relation $u^{p+1}(x) = -a_0 x a_1 u(x) - \cdots - a_p u^p(x)$ nous donne alors l'expression de la matrice $\mathrm{Mat}_{\mathcal{B}}(u_F)$. □

Ce résultat nous permet d'obtenir d'une autre façon des sous-espaces vectoriels F stables par un endomorphisme u. Comme $\chi_{u_F}(X)$ divise $\chi_u(X)$, cela nous permet d'obtenir à nouveau des diviseurs de $\chi_u(X)$. Le résultat suivant nous donne l'expression du polynôme caractéristique de l'endomorphisme induit de la proposition précédente.

Proposition-Définition 45
Soit $n \geq 1$. Soit $P(X) = X^n + \alpha_{n-1} X^{n-1} + \cdots \alpha_1 X + \alpha_0 \in \mathbb{K}[X]$ un polynôme unitaire.
On appelle **matrice compagnon de** P \多项式P的伴随矩阵\ [1], la matrice :

$$C_{P(X)} = \begin{pmatrix} 0 & 0 & \ldots & \ldots & 0 & -\alpha_0 \\ 1 & 0 & \ldots & \ldots & 0 & -\alpha_1 \\ 0 & 1 & \ddots & & & \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & \ldots & 1 & 0 & -\alpha_{n-2} \\ 0 & 0 & \ldots & 0 & 1 & -\alpha_{n-1} \end{pmatrix} \in \mathscr{M}_n(\mathbb{K}).$$

On a : $\chi_{C_{P(X)}}(X) = P(X)$.
Le polynôme caractéristique de la matrice compagnon de P est le polynôme $P(X)$.

Preuve — Notons $L_1, \ldots, L_{n+1}$ les lignes de la matrice $C_{P(X)}$. Son polynôme caractéristique est :

$$\chi_{C_{P(X)}} = \begin{vmatrix} X & 0 & \ldots & \ldots & 0 & \alpha_0 \\ -1 & X & \ldots & \ldots & 0 & \alpha_1 \\ 0 & -1 & \ddots & & & \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & \ldots & -1 & X & \alpha_{n-2} \\ 0 & 0 & \ldots & 0 & -1 & X + \alpha_{n-1} \end{vmatrix}$$

Avec l'opération $L_1 \leftarrow L_1 + XL_2 + \cdots + X^{n-1} L_n$, ce déterminant devient :

$$\chi_{C_{P(X)}} = \begin{vmatrix} 0 & 0 & \ldots & \ldots & 0 & P(X) \\ -1 & X & \ldots & \ldots & 0 & \alpha_1 \\ 0 & -1 & \ddots & & & \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & \ldots & -1 & X & \alpha_{n-2} \\ 0 & 0 & \ldots & 0 & -1 & X + \alpha_{n-1} \end{vmatrix}.$$

Le développement selon la première ligne donne : $\chi_{C_{P(X)}} = (-1)^{n+1} P(X)(-1)^{n-1} = P(X)$. □

Exemple 46 — *Pour $E = \mathbb{R}^3$, on considère l'endomorphisme $u : (x, y, z) \mapsto (y, z, x)$.*
Trouvons trois vecteurs x_1, x_2, x_3 de $\mathbb{R}^3$ tels que $\dim S_u(x_1) = 1 \quad \dim S_u(x_2) = 2 \quad \dim S_u(x_3) = 3$:

- *Le vecteur doit être x_1 un vecteur propre non nul. C'est le cas par exemple de $(1, 1, 1)$, associé à la valeur propre 1. On a donc $(X - 1) \mid \chi_u(X)$.*
- *L'espace stable engendré par x_2 doit être dimension 2. Cela signifie qu'il faut trouver un sous-espace vectoriel de dimension 2 stable par u, et choisir dans ce sous-e.v. un vecteur qui n'est pas un vecteur propre. On peut remarquer que l'hyperplan $x + y + z = 0$ est stable par u. Prenons $x_2 = (1, -1, 0)$. Alors $u(x_2) = (-1, 0, 1)$ et $u^2(x_2) = (0, 1, -1) = -x_2 - u(x_2)$. Donc le sous-e.v. $S_u(x_2)$ est bien de dimension 2. On a donc $X^2 + X + 1 \mid \chi_u(X)$. Comme $(X - 1)$ et $X^2 + X + 1$ sont premiers entre eux, on a donc $(X - 1)(X^2 + X + 1) \mid \chi_u(X)$, d'où $\chi_u(X) = (X - 1)(X^2 + X + 1)$ en comparant les degrés.*
  En déterminant des sous-espaces stables par u de la forme $S_u(x)$, on a ainsi factorisé $\chi_u(X)$.
- *En prenant $x_3 = (1, 0, 0)$, on remarque que $(x_3, u(x_3), u^2(x_3))$ engendre $\mathbb{R}^3$. Donc $S_u(x_3) = \mathbb{R}^3$.*

1. \注意不要和comatrice混淆\

3.3.4 Polynômes caractéristiques scindés

Rappelons que l'on dit qu'un polynôme $P(X) \in \mathbb{K}[X]$ est **scindé** sur $\mathbb{K}$ s'il peut s'écrire comme produit de facteurs du premier degré de $\mathbb{K}[X]$ et qu'il est scindé à racines simples sur $\mathbb{K}$ si de plus ses racines sont de multiplicité 1.

Exemples 47

1. *Soit E un $\mathbb{C}$-e.v. de dimension finie. Pour tout endomorphisme $u : E \to E$, le polynôme caractéristique $\chi_u(X)$ est scindé, puisque tout polynôme de $\mathbb{C}[X]$ est scindé.*
2. *Pour $E = \mathbb{R}^2$ et $R(\theta) = \begin{pmatrix} cos\theta & -sin\theta \\ sin\theta & cos\theta \end{pmatrix}$, avec $\theta \in \mathbb{R}$, son polynôme caractéristique est $\chi_{R(\theta)}(X) = X^2 - 2Xcos\theta + 1$. Si $\theta \not\equiv 0 \bmod \pi$, ce polynôme est irréductible dans $\mathbb{R}[X]$. Il n'est donc pas scindé. Par contre, $\chi_{R(\theta)}(X)$ est scindé à racines simples dans $\mathbb{C}[X]$.*

Proposition 48
Soit E un $\mathbb{K}$-e.v. de dimension finie. Soit $u : E \to E$ un endomorphisme. Si χ_u est scindé, alors pour tout sous-ev F stable par u, le polynôme χ_{u_F} est lui aussi scindé.

Preuve — Le polynôme χ_{u_F} divise le polynôme scindé χ_u. □

Proposition 49
Soit E un $\mathbb{K}$-e.v. de dimension n. Soit $u : E \to E$ un endomorphisme.
Si χ_u est scindé, alors pour $(\mu_1, \dots, \mu_n)$ les racines de χ_u comptées avec multiplicité, on a :

$$\mathrm{Tr}\,(u) = \mu_1 + \dots + \mu_n \quad \text{et} \quad \det(u) = \mu_1 \dots \mu_n.$$

Preuve — On a $\chi_u(X) = X^n - \mathrm{Tr}\,(u)\,X^{n-1} + \ldots + (-1)^n \det(u) = (X - \mu_1)\ldots(X - \mu_n) = X^n - (\mu_1 + \ldots + \mu_n)X^{n-1} + \ldots + (-1)^n \mu_1 \ldots \mu_n$. □

Remarque 50 — *Pour $A = (\alpha_{i,j})_n$ une matrice triangulaire supérieure, son polynôme caractéristique est scindé et ses racines (comptées avec multiplicité) sont $(\alpha_{1,1}, \ldots, \alpha_{n,n})$.*

Proposition 51
Soient u et v deux endomorphismes de E qui commutent i.e. $u \circ v = v \circ u$. Alors l'image et le noyau de u sont stables par v, et l'image et le noyau de v sont stables par u.

Preuve — Il suffit par symétrie de montrer que l'image et le noyau de u sont stables par v. Montrons que l'image de u est stable par v : soit $x \in \mathrm{Im}\,(u)$. Alors il existe $y \in E$ avec $x = u(y)$. Dès lors

$$v(x) = v(u(y)) = u(v(y)) \in \mathrm{Im}\,(u)\,.$$

La stabilité de $\mathrm{Im}\,(u)$ par v est prouvée.

Montrons que le noyau de u est stable par v : soit $x \in \mathrm{Ker}\,(u)$. Alors

$$u(v(x)) = v(u(x)) = v(0) = 0.$$

Donc $v(x) \in \mathrm{Ker}\,(u)$, et la stabilité de $\mathrm{Ker}\,(u)$ par v est prouvée. □

En particulier, si u et v commutent, alors tout sous-espace propre de u est stable par v. En effet, comme u, v commutent, on a aussi, pour tout $\lambda \in K$, la relation

$$(u - \lambda\,\mathrm{Id}_E) \circ v = u \circ v - \lambda v = v \circ u - \lambda v = v \circ (u - \lambda\,\mathrm{Id}_E).$$

Donc le noyau de $u - \lambda\,\mathrm{Id}_E$ (à savoir l'espace propre de u correspondant à la valeur propre λ) est stable par v, grâce au théorème que nous venons de prouver.

Corollaire 52
Les sous-espaces vectoriels propres d'un endomorphisme u de E sont stables par tout endomorphisme v commutant avec u.

3.4 Sommes de sous-espaces propres

3.4.1 Rappels sur les sommes finies de sous-espaces vectoriels

Définition 53
Soit E un $\mathbb{K}$-e.v. Soient $E_1, \ldots, E_p$ des sous-espaces vectoriels de E.

On appelle **somme des sous-espaces vectoriels** $(E_i)_{1\leq i\leq p}$ le sous-espace vectoriel :

$$E_1 + \ldots + E_p = \sum_{i=1}^{p} E_i = \left\{ \sum_{i=1}^{p} x_i,\ \text{avec } (x_i)_{1\leq i\leq p} \in \prod_{1\leq i\leq p} E_i \right\}.$$

Preuve — L'ensemble $\sum_{1\leq i\neq p} E_i$ est un sous-espace vectoriel, comme image de l'application linéaire

$$(x_i)_{1\leq i\leq p} \in \prod_{1\leq i\leq p} E_i \mapsto \sum_{1\leq i\leq p} x_i \in E$$

. □

Lorsque les E_i sont de dimension finie, l'application linéaire

$$(x_i)_{1\leq i\leq p} \in \prod_{1\leq i\leq p} E_i \mapsto \sum_{1\leq i\leq p} x_i \in E$$

donne, avec le théorème du rang :

$$\dim\left(\sum_{i\in I} Ei\right) \leqslant \sum_{i\in I} \dim\left(E_i\right).$$

§ *1. Somme directe d'une famille finie de sous-espaces vectoriels*

Définition 54
Soit E un $\mathbb{K}$-e.v. Soient $E_1, \ldots, E_p$ des sous-espaces vectoriels de E.
On dit que la famille $(E_1, \ldots, E_p)$ est en **somme directe** si l'on a :

$$\forall (x_1, \ldots, x_p) \in E_1 \times \ldots \times E_p, \quad \sum_{i=1}^{p} x_i = 0 \implies x_i = 0\ \forall 1 \leq i \leq p.$$

C'est-à-dire, si la seule somme de vecteurs de $E_1, \ldots, E_p$ qui est nulle est la somme de vecteurs nuls. Quand la famille $(E_1, \ldots, E_p)$ est en somme directe, on note alors sa somme $E_1 \oplus \ldots \oplus E_p = \bigoplus_{1\leq i\leq p} E_i$.

Exemples 55

1. *Une famille $(e_i)_{i\in I}$ de vecteurs non nuls est libre si, et seulement si, la famille $(\mathbb{K}e_i)_{i\in I}$ est en somme directe.*
2. *Un couple (F, G) de sous-espaces vectoriels est en somme directe si et seulement si $F \cap G = \{0\}$.*
3. *Soient E un ev. et $(E_1, \ldots, E_p)$ des sous-ev de E en somme directe. Toute sous-famille $(E_{i_1}, \ldots, E_{i_r})$, $r \leq p$ est encore en somme directe.*
4. *Soient E un ev. et $(E_1, \ldots, E_p)$ des sous-ev de E en somme directe. Soient $F_1, \ldots, F_p$ des sous-ev de $E_1, \ldots, E_p$. Alors les sous-ev $(F_1, \ldots, F_p)$ sont encore en somme directe.*

Remarque 56 — *Attention, une famille $(E_i)_{i\in I}$ avec au moins trois éléments qui vérifient $E_i \cap E_j = \{0\}$ pour tout $i \neq j$ n'est pas nécessairement en somme directe.*
Par exemple, la famille de sous-ev de $\mathbb{K}^2$

$$\mathbb{K}(1,0), \mathbb{K}(0,1), \mathbb{K}(1,1),$$

n'est pas en somme directe même si les intersections deux à deux de ces sous-espaces vectoriels sont réduites à $\{0\}$.

Proposition 57
Soit E un $\mathbb{K}$-ev. Soient $E_1, \ldots, E_p$ des sous-ev de E de dimension finie.
Alors, la famille $(E_1, \ldots, E_p)$ est en somme directe si, et seulement si, on a :

$$\dim\left(\sum_{i\in I} E_i\right) = \sum_{i\in I} \dim\left(E_i\right).$$

Preuve — En dimension finie, l'application linéaire surjective $(x_i) \in \prod_{i=1}^{p} E_i \mapsto \sum_{1=1}^{p} x_i \in \sum_{i=1}^{p} E_i$ est injective si et seulement si elle est bijective, si et seulement si les espaces vectoriels de départ et d'arrivée ont même dimension. □

Définition 58
Soit E un $\mathbb{K}$-e.v. Soient $E_1, \ldots, E_p$ des sous-e.v. de E. Si la famille $(E_1, \ldots, E_p)$ est en somme directe avec $E = \bigoplus_{1 \leq i \leq p} E_i$, on dit alors que $(E_1, \ldots, E_p)$ est une **décomposition en somme directe** \直和分解\ de E.

Remarque 59 — *Attention, si F est un sous-espace vectoriel de E, on a :*

$$\bigoplus_{1 \leq i \leq p} (F \cap E_i) \subset F,$$

mais pas nécessairement :

$$\bigoplus_{1 \leq i \leq p} (F \cap E_i) = F.$$

Par exemple, on a $\mathbb{K}^2 = \mathbb{K}(1,0) \oplus \mathbb{K}(0,1)$ mais le sous-espace vectoriel $F = \mathbb{K}(1,1)$ n'est pas égal à la somme directe :

$$(F \cap \mathbb{K}(1,0)) \oplus (F \cap \mathbb{K}(0,1))$$

qui est réduite à $\{0\}$.

§ *2. Projections associées à une décomposition en somme directe*

Proposition-Définition 60
Soit E un $\mathbb{K}$-e.v. avec une décomposition en somme directe $E = \bigoplus_{1 \leq i \leq p} E_i,$.

Alors tout élément x de E s'écrit de façon sous la forme

$$x = \sum_{i=1}^{p} x_i,\ x_i \in E_i.$$

Le vecteur x_i s'appelle **la i-ème composante de** x dans la décomposition $E = \bigoplus_{i \in I} E_i$.
La fonction $p_i : x \in E \mapsto x_i \in E$ est la projection sur E_i parallèlement à $\bigoplus_{j \neq i} E_j$.
On l'appelle **la i-ème projection associée à la décomposition** $E = \bigoplus_{i \in I} E_i$.
Ces projections vérifient les relations :

1. pour tout $i \in I$, $p_i^2 = p_i$; ◀
2. $p_i \circ p_j = 0$ pour tous $i \neq j$; ◀
3. $\sum_{i \in I} p_i = \mathrm{Id}_E$. ◀

Proposition 61
Soit E un $\mathbb{K}$-e.v. avec une décomposition en somme directe $E = \bigoplus_{1 \leq i \leq p} E_i,$. Soient $\mathcal{B}_1, \ldots, \mathcal{B}_p$ des bases de $E_1, \ldots, E_p$.
Alors, $\mathcal{B} = \mathcal{B}_1 \cup \ldots \cup \mathcal{B}_p$ est une base de E.

Preuve — Il est évident que la réunion $\mathcal{B}$ des $\mathcal{B}_i$ engendre la somme des E_i. Pour montrer qu'elle est libre, considérons une combinaison linéaire $\sum_{e \in \mathcal{B}} \alpha_e e$ nulle. On a alors :

$$\sum_{i \in I} \left(\sum_{e \in \mathcal{B}_i} \alpha_e e \right) = \sum_{e \in \mathcal{B}} \alpha_e e = 0.$$

Puisque les vecteurs $\sum_{e \in \mathcal{B}_i} \alpha_e e$ appartiennent aux E_i, la propriété de somme directe entraîne $\sum_{e \in \mathcal{B}_i} \alpha_e e = 0$. On en déduit que, pour tout e de $\mathcal{B}_i$, le scalaire α_e est nul. Cela valant pour tout i, la famille $(\alpha_e)_{e \in \mathcal{B}}$ est nulle. □

3.4.2 Sous-espaces stables et matrices triangulaires

Proposition 62
Soient E un $\mathbb{K}$-e.v. de dimension n et $E_1, \ldots, E_r$ des sous-ev de E. On suppose que E est la somme directe des sous-ev E_i :

$$E = \bigoplus_{i=1}^{r} E_i.$$

Soit $n_i = \dim(E_i)$. Soient $\mathcal{B}_1, \ldots, \mathcal{B}_r$ des bases de $E_1, \ldots, E_r$, et $\mathcal{B} = \bigcup_{i=1}^{r} \mathcal{B}_i$, qui est une base de E. Soit $u : E \to E$ un endomorphisme.

Alors, les sous-ev $E_1, \dots, E_r$ sont stables par u si et seulement si la matrice de u par rapport à $\mathcal{B}$ est une matrice diagonale par blocs :

$$\mathrm{Mat}_{\mathcal{B}}(u) = \begin{pmatrix} A_1 & 0 & \cdots & 0 \\ 0 & A_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \cdots & A_r \end{pmatrix},$$

où $A_i \in \mathscr{M}_{n_i}(\mathbb{K})$, $\forall 1 \leq i \leq r$, et où les zéros de la matrice représentent des blocs de zéros.

Remarque 63 — *Dans un tel cas, on a alors*

$$\chi_u(X) = \chi_{A_1}(X) \cdots \chi_{A_r}(X).$$

Corollaire 64 (Caractérisation des matrices triangulaires)
Soient E un $\mathbb{K}$-e.v. de dimension n et $\mathcal{B} = (e_1, \dots, e_n)$ une base de E. Soit $u : E \to E$ un endomorphisme de E. Alors, la matrice $\mathrm{Mat}_{\mathcal{B}}(u)$ est triangulaire supérieure si et seulement si les n sous-espaces vectoriels :

$$\mathrm{Vect}\,(e_1)\,, \mathrm{Vect}\,(e_1, e_2)\,, \dots, \mathrm{Vect}\,(e_1, \dots, e_k)\,, \dots, \mathrm{Vect}\,(e_1, \dots, e_n)\,,$$

sont tous stables par u.

Preuve — Il suffit d'observer qu'une matrice triangulaire possède en bas à gauche un bloc de zéros de taille $(n-p) \times p$ pour tout $p \in [\![1, n]\!]$, et d'appliquer le théorème qui décrit les matrices des endomorphismes pour lesquels $\mathrm{Vect}(e_1, \dots, e_p)$ est stable. □

3.4.3 Sous-espaces propres et sommes directes

Théorème 65
Soient E un $\mathbb{K}$-e.v. et $u : E \to E$ un endomorphisme. Soient $\lambda_1, \dots, \lambda_p \in \mathrm{Spec}\,(u)$ des valeurs propres de u distinctes. Soient $x_1, \dots, x_p \in E$ des vecteurs propres associés à $\lambda_1, \dots, \lambda_p$.
Alors, la famille $(x_1, \dots, x_p)$ est une famille libre. \属于不同特征值的特征向量是线性无关的\

Preuve — Nous allons démontrer cela par récurrence sur $p \geq 1$.

- Si $p = 1$, la famille (x_1) est libre car un vecteur propre n'est jamais nul par définition.
- Supposons le résultat vrai pour toute famille de taille $p-1$. Soient $\lambda_1, \dots, \lambda_p$ p valeurs propres distinctes de u et $x_1, \dots, x_p$ p vecteurs propres associés. Il faut montrer l'implication

$$\sum_{k=1}^{p} a_k x_k = 0 \implies a_1 = a_2 = \cdots = a_p = 0$$

Supposons que $\sum_{k=1}^{p} a_k x_k = 0$. En appliquant la fonction u, on obtient :

$$0 = u(0) = \sum_{k=1}^{p} a_k u(x_k) = \sum_{k=1}^{p} a_k \lambda_k x_k$$

On a donc :

$$0 = \lambda_p \cdot (\sum_{k=1}^{n} a_k x_k) - \sum_{k=1}^{p} a_k \lambda_k x_k = \sum_{k=1}^{p-1} a_k (\lambda_p - \lambda_k) x_k$$

D'après l'hypothèse de récurrence, la famille $(x_1, \cdots, x_{p-1})$ est libre. On a donc :

$$a_k(\lambda_k - \lambda_p) = 0,\ \forall k \in [\![1, p-1]\!].$$

Comme on a $\lambda_k - \lambda_p \neq 0$, on obtient $a_k = 0$ pour tout $1 \leq k \leq p-1$. Il reste alors dans la somme $0 + a_p x_p = 0$, ce qui donne $a_p = 0$. Cela conclut la récurrence.

□

Autrement dit, toute famille finie de vecteurs propres associés à des valeurs propres deux à deux distinctes est libre. On notera que cette propriété est valable en dimension infinie.

Théorème 66
Soient E un $\mathbb{K}$-e.v. et $u : E \to E$ un endomorphisme. Soient $\lambda_1, \dots, \lambda_p \in \mathrm{Spec}\,(u)$ des valeurs propres de u distinctes.
Alors, la famille des sous-espaces propes $(E_{\lambda_1}(u), \dots, E_{\lambda_p}(u))$ est en somme directe.

Preuve — Soit $(x_1, \dots, x_p) \in E_{\lambda_1}(u) \times \dots \times E_{\lambda_p}(u)$ une famille de vecteurs telle que

$$\sum_{i=1}^{p} x_i = 0.$$

Supposons par l'absurde que les vecteurs $x_1, \ldots, x_p$ ne sont pas tous nuls. Quitte à réordonner les vecteurs, on peut supposer que les vecteurs $x_1, \ldots, x_r$ sont non-nuls, pour un $1 \leq r \leq p$, et que les vecteurs $x_{r+1}, \ldots, x_p$ sont nuls.
Les vecteurs $x_1, \ldots, x_r$ sont des vecteurs propres non-nuls, associés à des valeurs propres distinctes, pour lesquels on aurait $x_1 + \ldots + x_r = 0$.

D'après le théorème précédent, la famille $(x_1, \ldots, x_r)$ est libre, ce qui contredit la relation $x_1 + \ldots + x_r = 0$.

Donc tous les vecteurs $x_1, \ldots, x_p$ sont nuls. Ainsi, les espaces propres $E_{\lambda_1}(u), \ldots, E_{\lambda_p}(u)$ sont en somme directe. □

EXEMPLE 67 — *Soient $E = \mathcal{C}^\infty(\mathbb{R}, \mathbb{C})$ et $D : f \in E \mapsto f' \in E$ l'application linéaire de dérivation. Pour tout $\lambda \in \mathbb{C}$, la fonction $x \in \mathbb{R} \longmapsto e^{\lambda x}$ est un vecteur propre de D pour la valeur propre λ. La famille $\left(x \in \mathbb{R} \longmapsto e^{\lambda x} \in \mathbb{C}\right)_{\lambda \in \mathbb{C}}$ est donc libre, d'après les résultats précédents (toute combinaison linéaire d'un nombre fini de ces fonctions qui est nulle est forcément la combinaison linéaire nulle). Comme cette famille est infinie, on en déduit que l'espace vectoriel $\mathcal{C}^\infty(\mathbb{R}, \mathbb{C})$ est de dimension infinie (ce résultat a déjà été démontré de plusieurs manières dans les cours précédents).*

3.5 DIAGONALISABILITÉ

3.5.1 Définition et caractérisations élémentaires

DÉFINITION 68
Soit E un $\mathbb{K}$-e.v. de dimension n. Soit $u : E \to E$ un endomorphisme.
On dit que l'endomorphisme u est **diagonalisable** \可对角化矩阵\ s'il existe une base B de E telle que $\mathrm{Mat}_B(E)$ est une matrice diagonale.
Une base B dans laquelle la matrice $\mathrm{Mat}_B(E)$ est diagonale est appelée une **base de diagonalisation** de u.
Soient $n \geq 1$ et $A \in \mathscr{M}_n(\mathbb{K})$.
On dit que la matrice A est **diagonalisable** \可对角化矩阵\ s'il existe une matrice inversible P telle que $P^{-1}AP$ est une matrice diagonale.

REMARQUES 69
- *Pour $\mathcal{B} = (e_1, \ldots, e_n)$ une base de diagonalisation de u, les vecteurs $e_1, \ldots, e_n$ sont alors des vecteurs propres de u.*
- *Une matrice $A \in \mathscr{M}_n(\mathbb{K})$, est diagonalisable si et seulement si l'application linéaire $X \in \mathbb{K}^n \mapsto AX$ est diagonalisable.*
  En effet, un changement de base vers la base B correspond à la matrice PAP^{-1}, où P est la matrice de passage de la base canonique vers la base B

THÉORÈME 70
Soient E un $\mathbb{K}$-e.v. de dimension n et $u : E \to E$ un endomorphisme.
On a les équivalences suivantes :

i) u est diagonalisable ;

ii) Il existe une base $\mathcal{B}$ de E constituée de vecteurs propres de u ;

iii) E est la somme directe de sous-espaces propres de u :

$$E = \bigoplus_{\lambda \in \mathrm{Spec}(u)} E_\lambda(u) \,;$$

\ u可以对角化 $\Longleftrightarrow$ 所有特征子空间的直和是全空间\

iv) Pour $\mathrm{Spec}\,(u) = \{\lambda_1, \ldots, \lambda_r\}$, on a :

$$\sum_{k=1}^{r} \dim\left(E_{\lambda_k}(u)\right) = n \,;$$

v) Le polynôme caractéristique χ_u est scindé, et pour $\mathrm{Spec}\,(u) = \{\lambda_1, \ldots, \lambda_r\}$, on a :

$$\forall k \in [\![1, r]\!], \dim\left(E_{\lambda_k}(u)\right) = m(\lambda_k).$$

\ u可以对角化 $\Longleftrightarrow$ 每个特征值的几何重数等于代数重数\

Preuve — On pose $\mathrm{Spec}\,(u) = \{\lambda_1, \ldots, \lambda_r\}$. Pour $1 \leq i \leq r$, on pose $F_i = E_{\lambda_i}(u)$.
i) $\Longrightarrow$ *ii)* Soit $\mathcal{B} = (e_1, \ldots, e_n)$ une base de diagonalisation de u. Comme $\mathrm{Mat}_{\mathcal{B}}(u)$ est diagonale, les vecteurs $e_1, \ldots, e_n$ sont des vecteurs propres de u.

ii) $\Longrightarrow$ *iii)* On a montré que les sous-espaces propres F_i étaient en somme directe. S'il existe une base $B = (e_1, \ldots, e_n)$ de E constituée de vecteurs propres, alors $\bigoplus_{i=1}^k F_i$ contient une base de E. On a donc $\dim(F_1 \oplus \ldots \oplus F_r) \geq n = \dim(E)$. Donc ce sous-ev est de dimension n, ce qui donne :

$$E = \bigoplus_{i=1}^{r} F_i.$$

iii) $\Longrightarrow$ *iv)* On a $\dim(E) = n = \dim(F_1 \oplus \ldots \oplus F_r) = \sum_{k=1}^n \dim(F_k)$.

iv) $\Longrightarrow$ *i)* Soient $\mathcal{B}_1, \ldots, \mathcal{B}_r$ des bases de $F_1, \ldots, F_r$. Comme ces sous-espaces propres sont en somme directe, les résultats de la section précédente nous disent que $\mathcal{B} = \mathcal{B}_1 \cup \ldots \cup \mathcal{B}_r$ est une base de $F_i \oplus \ldots \oplus F_r$. Cette base contient $\sum_{k=1}^n \dim(F_k) = n = \dim(E)$ éléments. C'est donc une base de E. Cette base étant une base formée de vecteurs propres de u, la matrice $\mathrm{Mat}_{\mathcal{B}}(u)$ est alors une matrice diagonale.

iv) $\Longleftrightarrow$ *v)* On a montré précédemment que l'on a $0 \leq \dim\left(E_{\lambda_k}(u)\right) \leqslant m_u(\lambda_k)$ pour tout $1 \leq k \leq n$. On sait de plus que $m_u(\lambda_1) + \ldots + m_u(\lambda_r) \leq \deg(\chi_u(X)) = n$. Cela donne :

$$0 \leq \sum_{k=1}^{r} \dim\left(E_{\lambda_k}(u)\right) \leqslant \sum_{k=1}^{r} m(\lambda_k) \leqslant n.$$

Ainsi, l'égalité :

$$\sum_{k=1}^{r} \dim\left(E_{\lambda_k}(u)\right) = n$$

est équivalente à :

$$\sum_{k=1}^{r} m(\lambda_k) = n \quad \text{et} \quad \dim\left(E_{\lambda_k}(u)\right) = m(\lambda_k),\ \forall 1 \leq k \leq r.$$

□

Corollaire 71

Soient E un $\mathbb{K}$-e.v. de dimension n et $u : E \to E$ un endomorphisme.
Si le polynôme caractéristique χ_u est scindé à racines simples, alors l'endomorphisme u possède n valeurs propres distinctes et est diagonalisable.
Pour $\mathrm{Spec}\,(u) = \{\lambda_1, \ldots, \lambda_n\}$ et pour $e_1, \ldots, e_n$ des vecteurs propres de u associés aux valeurs propres $\lambda_1, \ldots, \lambda_n$, la famille $\mathcal{B} = (e_1, \ldots, e_n)$ est une base de E, et $\mathrm{Mat}_{\mathcal{B}}(u) = \mathrm{Diag}(\lambda_1, \ldots, \lambda_n)$.

Exemple 72 — *Soit $k \in \mathbb{C}$. On définit la matrice de $\mathscr{M}_4(\mathbb{C})$:*

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

Cette matrice est de rang 2, et de trace k. Un calcul de déterminant nous donne $\chi_A(X) = X^2(X^2 - kX - 3)$.
Le sous espace-propre associé à 0, qui est $\mathrm{Ker}\,(A)$*, s'obtient en résolvant le système :*

$$\begin{cases} y = 0 \\ x + ky + z + t = 0 \end{cases}$$

Il est de dimension 2, engendré par les vecteurs :

$$\begin{pmatrix} 1 \\ 0 \\ -1 \\ 0 \end{pmatrix} \quad et \quad \begin{pmatrix} 1 \\ 0 \\ 0 \\ -1 \end{pmatrix}.$$

On retrouve d'ailleurs que A est de rang 2.

Ainsi :

- *Si k vérifie $k^2 + 12 \neq 0$, le polnôme $X^2 - kX - 3$ a deux racines λ_1 et λ_2 distinctes non nulles. Les sous-espaces vectoriels propres associés étant de dimension non-nulle, la matrice A est donc diagonalisable.*
- *Si k est égal à $\pm 2i\sqrt{3}$, le polynôme $X^2 - kX - 3$ a une racine $\lambda = \frac{k}{2}$ de multiplicité 2. On détermine le sous-espace propre associé en résolvant :*

$$\begin{cases} -\lambda x & + & y & & & & & = & 0 \\ x & + & (k-\lambda)y & + & z & + & t & = & 0 \\ & & y & - & \lambda z & & & = & 0 \\ & & y & - & & & \lambda t & = & 0 \end{cases}.$$

Une résolution avec la méthode du Pivot montre que ce sous-espace propre est de dimension 1*, engendré par*

$$\begin{pmatrix}1\\ \lambda\\ 1\\ 1\end{pmatrix}.$$

La dimension de la somme des sous-espaces propres de A *vaut* 3*, donc la matrice* A *n'est pas diagonalisable dans ce cas.*

Remarque 73 — *Soient* E *un e.v. de dimension et* n $u : E \to E$ *un endomorphisme. Si* $\chi_u(X) = (X-\lambda)^n$*,* u *n'a qu'une seule valeur propre* λ*.*
Alors u *est diagonalisable si et seulement si* $Ker(u - \lambda Id_E) = E$*, si et seulement si* $u = \lambda \operatorname{Id}_E$ *.*
Par exemple, la matrice $A = \begin{pmatrix}2 & a & b\\ 0 & 2 & c\\ 0 & 0 & 2\end{pmatrix}$ *est diagonalisable si et seulement si* $a = b = c = 0$*.*

Exemple 74 — *Soit* E *un* $\mathbb{K}$*-e.v. Soit* $u : E \to E$ *un endomorphisme*
S'il existe un entier $r \in \mathbb{N}$ *tel que* $u^r = 0$*, on dit que l'endomorphisme* u ***nilpotent***[2]*.*
*On appelle **indice de nilpotence** de* u *l'entier* $\operatorname{nil}(u) = \min\{k \in \mathbb{N}, u^k = 0\}$*. Cet entier existe en tant que minimum d'une partie non vide de* $\mathbb{N}$*.*
Soit E *de dimension* n*. Supposons que pour* u *nilpotent son polynôme caractéristique* χ_u *est scindé. Alors, on a* $\chi_u(X) = X^n$*.*
En effet, pour $\lambda \in \operatorname{Spec}(u)$ *et* x *un vecteur propre associé à* λ*, on a*

$$0 = u^{\operatorname{nil}(u)}(x) = \lambda^{\operatorname{nil}(u)} x,$$

ce qui implique que $\lambda = 0$*. Comme* χ_u *est scindé, cela implique que* $\chi_u(X) = X^n$*.*
Ainsi, un endomorphisme nilpotent dont le polynôme caractéristique est scindé est diagonalisable si et seulement s'il est nul.
Par exemple, $A = \begin{pmatrix}a & 1\\ a^2 & -a\end{pmatrix}$ *est une matrice nilpotente qui n'est pas diagonalisable.*
Nous verrons dans le chapitre suivant que tout endomorphisme nilpotent u *a son polynôme caractéristique scindé. (et que donc* $\chi_u(X) = X^n$*)*

Méthode 75 (**Etude de la diagonalisabilité**) — *Pour déterminer si un endomorphisme* u *est diagonalisable, on procède en général de la façon suivante :*

1. *On détermine* χ_u *le polynôme caractéristique de* u *;*
2. *On factorise ce polynôme pour obtenir le spectre de* u *;*
   Cela passe par un calcul de $\det(XI_n - Mat_B(u))$ *via la méthode du Pivot pour trouver des facteurs de* χ_u*, par une recherche de vecteurs propres évidents pour trouver des facteurs* $(X-\lambda)^k$ *de* χ_u*, ou par une recherche de sous-espaces stables par* u *pour trouver des diviseurs de* χ_u*.*
   Si le polynôme caractéristique χ_u *n'est pas scindé, alors* u *n'est pas diagonalisable.*
3. *On détermine les sous-espaces propres* $E_\lambda(u)$ *de* u*. (une base* $\mathcal{B}_\lambda$ *et leur dimension)*
4. *Conclure : Si la somme des dimensions des sous-espaces propres* $E_\lambda(u)$ *de* u *ne vaut pas* n*, alors* u *n'est pas diagonalisable.*
   Sinon, u *est diagonalisable et* $\mathcal{B} = \bigcup_{\lambda \in \operatorname{Spec}(u)} B_\lambda$ *est une base de* E *qui diagonalise* u*.*

Exemples 76

1. *On veut étudier la matrice :*

$$A = \begin{pmatrix}0 & 3 & 2\\ -2 & 5 & 2\\ 2 & -3 & 0\end{pmatrix} \in \mathscr{M}_3(\mathbb{K}).$$

Le polynôme caractéristique de A *vaut : (factorisation via la méthode du Pivot ou via des racines évidentes)*

$$\chi_A(X) = X^3 - 5X^2 + 8X - 4 = (X-2)^2(X-1).$$

Les racines de χ_A *comptées avec multiplicité sont* $1, 2, 2$*. Le sous-espace propre* $E_1(A) = Ker(A - I_3)$ *est alors de dimension* 1*. Trouvons-en une base. On résout l'équation* $AX = X$*, soit :*

$$\begin{cases} -x + 3y + 2z &= 0\\ -2x + 4y + 2z &= 0\\ 2x - 3y - z &= 0 \end{cases}.$$

2. \a(m).幂零的\

Ce système linéaire est équivalent à $x = y = -z$. Ainsi on a $E_1(A) = \mathbb{K}v_1$, pour

$$v_1 = \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}.$$

Le sous-espace propre $E_2(A) = Ker(A2I_3)$ est de dimension 1 ou 2. Trouvons-en une base. On résout l'équation $AX - 2X = 0$, soit :

$$\begin{cases} -2x + 3y + 2z &= 0 \\ -2x + 3y + 2z &= 0 \\ 2x - 3y - 2z &= 0 \end{cases}.$$

Ce système linéaire est équivalent à $2x - 3y - 2z = 0$. Ainsi on a $E_2(A) = \mathbb{K}v_2 \oplus \mathbb{K}v_3$ où

$$v_2 = \begin{pmatrix} 3 \\ 2 \\ 0 \end{pmatrix} \quad et \quad v_3 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}.$$

La matrice A est donc diagonalisable car la somme des dimensions de ses sous-espaces propres vaut 3. Une base de diagonalisation de A est $\mathcal{C} = (v_1, v_2, v_3)$. La matrice de passage de la base canonique à $\mathcal{C}$ est :

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

La matrice de l'endomorphisme $X \mapsto AX$ dans la base $\mathcal{C}$ est $P^{-1}AP = \mathrm{Diag}\,(1, 2, 2) = D$. On obtient alors $A = PDP^{-1}$, soit après calcul de P^{-1} :

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

2. *On se place sur $E = \mathbb{R}_n[X]$. On définit la fonction :*

$$\begin{array}{rcl} u \, : \mathbb{R}_n[X] & \longrightarrow & \mathbb{R}_n[X] \\ P(X) & \longmapsto & (1 - X^2)P'(X) + nXP(X) \end{array}$$

. Cette fonction est bien définie car :

$$u(X^k) = (n-k)X^{k+1} + kX^{k-1} \in \mathbb{R}_n[X], \quad \forall k \in [\![0, n]\!],$$

(pour $k = 0$, l'expression kX^{k-1} est nulle, et pour $k = n$ on a $u(X^n) = nX^{n-1}$). La fonction u est alors un endomorphisme sur $\mathbb{R}_n[X]$. La matrice de u dans la base canonique $(1, X, \dots, X^n)$ est :

$$\begin{pmatrix} 0 & 1 & & & (0) \\ n & 0 & 2 & & \\ & n-1 & \ddots & \ddots & \\ & & \ddots & \ddots & n \\ (0) & & & 1 & 0 \end{pmatrix}.$$

Son polynôme caractéristique ne parait pas facile à calculer.[3] *On recherche directement les valeurs propres de u. Un polynôme P non nul de E est un vecteur propre de u s'il existe $\lambda \in \mathbb{R}$ tel que l'on ait :*

$$(1 - X^2)P'(X) + nXP(X) = \lambda P(X).$$

Dans $\mathbb{R}(X)$, cela donne :

$$\frac{P'(X)}{P(X)} = \frac{nX - \lambda}{X^2 - 1} = \frac{n - \lambda}{2(X-1)} + \frac{n + \lambda}{2(X+1)}.$$

Si la factorisation de $P(X)$ dans $\mathbb{C}[X]$ est $a_m \prod_{k=1}^{r} (X - \alpha_k)^{n_k}$, on a alors :

$$\frac{P'(X)}{P(X)} = \sum_{k=1}^{r} \frac{n_k}{X - \alpha_k}.$$

3. On peut pourtant le calculer avec une récurrence double par un développement selon certaines lignes/colonnes. Mais cela sera pour un autre exemple ou en exercice.

Ainsi, le polynôme $P(X)$ doit être de degré $\frac{n-\lambda}{2} + \frac{n+\lambda}{2} = n$, l'ensemble de ses racines sur $\mathbb{C}$ est contenu dans $\{1, -1\}$, et $\frac{n-\lambda}{2}, \frac{n+\lambda}{2}$ doivent être des entiers compris entre 0 et n.
Pour $0 \leq k \leq n$, on pose $P_k(X) = (X-1)^k(X+1)^{n-k}$. Un calcul montre que l'on a alors :

$$u(P_k(X)) = (n-2k)P_k(X).$$

Comme les nombres réels $(n-2k)_{k\in[\![0,n]\!]}$ sont deux à deux distincts, l'endomorphisme u possède $n+1 = \dim(\mathbb{R}_n[X])$ valeurs propres distinctes. Ainsi, la famille $(P_k(X))_{k\in[\![0,n]\!]}$ de vecteurs propres de u est une base de $\mathbb{R}_n[X]$. Cet endomorphisme est donc diagonalisable, avec pour valeurs propres $(n-2k)_{k\in[\![0,n]\!]}$.

3.5.2 Réduction des endomorphismes diagonalisables

Remarque 77 — *Soit E un $\mathbb{K}$-e.v. de dimension n. Soit $u : E \to E$ un endomorphisme. Supposons que u est diagonalisable. Soit $\mathrm{Spec}\,(u) = \{\lambda_1, \dots, \lambda_r\}$. On a montré dans cette section que l'on a alors :*

- $\sum_{k=1}^r \dim(E_{\lambda_k}(u)) = n$;
- $E = E_{\lambda_1}(u) \oplus \ldots \oplus E_{\lambda_r}(u)$;
- $\dim(E_{\lambda_k}(u)) = m_u(\lambda_k)$, $\forall 1 \leq k \leq r$;
- $\chi_u(X) = \Pi_{i=1}^r (X-\lambda_k)^{\dim(E_{\lambda_k}(u))}$.
- *Pour $\mathcal{B}_1, \dots, \mathcal{B}_r$ des bases de $E_{\lambda_1}(u), \dots, E_{\lambda_r}(u)$, la famille $\mathcal{B} = \mathcal{B}_1 \cup \ldots \cup \mathcal{B}_r$ est une base de E.*
- *La matrice $Mat_{\mathcal{B}}(u)$ est une matrice diagonale.*

Pour tout $1 \leq k \leq r$, on a $u_{E_{\lambda_k}(u)} = \lambda_k Id_{E_{\lambda_1}(u)}$.
Ainsi, en notant $D_k = \lambda_k I_{\dim(E_{\lambda_k}(u))}$, la matrice $Mat_{\mathcal{B}}(u)$ est diagonale par blocs, avec

$$Mat_{\mathcal{B}}(u) = Diag(D_1, \dots, D_r).$$

Pour $1 \leq k \leq r$, soit $p_k : E \to E$ la projection sur $E_{\lambda_k}(u)$ parallèlement à $\bigoplus_{j\neq k} E_{\lambda_j}(u)$. On a alors :

$$u = \lambda_1 p_i + \ldots + \lambda_r p_r,$$

car ces deux endomorphismes sont égaux sur la base $\mathcal{B}$.

Remarque 78 (Puissances d'un endomorphisme diagonalisable) — *Soit E un $\mathbb{K}$-e.v. de dimension n. Soit $u : E \to E$ un endomorphisme. Supposons que u est diagonalisable, et reprenons les notations de la remarque précédente.*
Soit $m \geq 0$. On a alors $u^m_{E_{\lambda_k}(u)} = \lambda_k^m Id_{E_{\lambda_1}(u)}$.
Ainsi, on a :

$$Mat_{\mathcal{B}}(u^m) = Diag(D_1^m, \dots, D_r^m).$$

En utilisant les projections p_k, cela donne :

$$u^m = \lambda_1^m p_1 + \ldots + \lambda_r^m p_r,$$

car ces deux endomorphismes sont égaux sur la base $\mathcal{B}$.
Ainsi, l'endomorphisme u^m est totalement décrit à l'aide du spectre et de la base de diagonalisation de u.
En particulier, u^m est diagonalisable, et les racines de son polynôme caractéristique χ_{u^m} sont $\lambda_1^m, \dots, \lambda_r^m$ (certaines racines pouvant être égales). C'est-à-dire :

$$\chi_{u^m}(X) = \Pi_{i=1}^k (X-\lambda_k^m)^{\dim(E_{\lambda_k}(u))}.$$

D'un point de vue matriciel, si $A \in \mathscr{M}_n(\mathbb{K})$ est diagonalisable, on a P une matrice inversible et $D = Diag(a_1, \dots, a_n)$ une matrice diagonale telles que $A = PDP^{-1}$.
Pour tout $m \geq 0$, on a alors :

$$A^m = (PDP^{-1})^m = PD^mP^{-1} = PDiag(a_1^m, \dots, a_n^m)P^{-1}.$$

On peut alors calculer facilement la matrice A^m à l'aide de deux produits matriciels.

Exemple 79 — *Reprenons le premier exemple de la page 48. Nous avons vu que la matrice*

$$A = \begin{pmatrix} 0 & 3 & 2 \\ -2 & 5 & 2 \\ 2 & -3 & 0 \end{pmatrix}$$

est diagonalisable. Elle s'écrit $A = P \cdot \mathrm{Diag}\,(1,2,2) \cdot P^{-1}$ *avec :*

$$P = \begin{pmatrix} 1 & 3 & 1 \\ 1 & 2 & 0 \\ -1 & 0 & 1 \end{pmatrix} \qquad \textit{et} \quad P^{-1} = \begin{pmatrix} 2 & -3 & -2 \\ -1 & 2 & 1 \\ 2 & -3 & -1 \end{pmatrix}.$$

Pour $k \in \mathbb{N}$*, obtient alors les puissances de* A *par la relation* $A^k = P \cdot \mathrm{Diag}\left(1^k, 2^k, 2^k\right) \cdot P^{-1}$*. Cela donne :*

$$A^k = 1^k P \cdot \mathrm{Diag}\,(1,0,0) \cdot P^{-1} + 2^k P \cdot \mathrm{Diag}\,(0,1,1) \cdot P^{-1} = \begin{pmatrix} 2 & -3 & -2 \\ 2 & -3 & -2 \\ -2 & 3 & 2 \end{pmatrix} + 2^k \begin{pmatrix} -1 & 3 & 2 \\ -2 & 4 & 2 \\ 2 & -3 & -1 \end{pmatrix}.$$

On remarque que les matrices devant 1^k *et* 2^k *dans l'écriture ci-dessus correspondent aux matrices dans la base canonique des projections* p_1 *et* p_2 *associées à la somme directe de sous-espaces propres* $\mathbb{K}^3 = E_1(A) \oplus E_2(A)$*.*

Remarque 80 — *Dans l'exemple précédent, pour* M_1 *et* M_2 *les matrices associées aux projections* p_1 *et* p_2*, on a :*

$$M_1 = \begin{pmatrix} 2 & -3 & -2 \\ 2 & -3 & -2 \\ -2 & 3 & 2 \end{pmatrix} = (2I_3 - A), \quad \textit{et } M_2 = \begin{pmatrix} -1 & 3 & 2 \\ -2 & 4 & 2 \\ 2 & -3 & -1 \end{pmatrix} = (A - I_3).$$

Les polynômes $L_1(X) = (2 - X)$ *et* $L_2(X) = (X - 1)$ *sont les polynômes d'interpolation associés à l'ensemble* $\{1, 2\}$ *(*$L_1(1) = 1$*,* $L_1(2) = 0$*,* $L_2(2) = 2$*,* $L_2(1) = 0$*).*
Nous verrons dans le chapitre suivant que lorsque A *est une matrice diagonalisable, les projections* p_k *associées aux sous-espaces propres de* A *peuvent être calculées avec des polynômes d'interpolation de Lagrange.*
Ce calcul ne nécessite pas de déterminer une base de vecteurs propres (donc la matrice de passage P*), ni de faire un inverse de matrice (calculer* P^{-1}*).*

Chapitre 4 Polynômes d'endomorphisme

Table des matières du chapitre

Dans ce chapitre, l'ensemble $\mathbb{K}$ désigne un corps. On utilisera souvent $\mathbb{K} = \mathbb{R}, \mathbb{C}$ ou $\mathbb{Q}$. Tous les objets que nous verrons seront définis sur des $\mathbb{K}$-espaces vectoriels E. Par contre, certains résultats ne seront vrais que pour des e.v. E de dimension finie.

4.1 Morphisme d'évaluation

Définition 1

Soient E un $\mathbb{K}$-e.v., $u \in \mathcal{L}(E)$ et $P = \sum_{k=0}^{n} a_k X^k \in \mathbb{K}[X]$.

- On appelle **évaluation** du polynôme P en u l'endomorphisme de E suivant :

$$P(u) = \sum_{k=0}^{n} a_k u^k = a_0 \operatorname{Id}_E + a_1 u + \cdots + a_n u^n,$$

où $u^k = \underbrace{u \circ u \circ \cdots \circ u}_{k \text{ fois}}$ est la composée k-ème de u, et $u^0 = \operatorname{Id}_E$.
La fonction $P \mapsto P(u)$ est appelée **morphisme d'évaluation** en u.

- Soit $A \in \mathscr{M}_n(\mathbb{K})$. On définit **l'évaluation** du polynôme P en A par la matrice :

$$P(A) = \sum_{k=0}^{n} a_k A^k = a_0 I_n + a_1 A + \cdots + a_n A^n.$$

La fonction $P \mapsto P(A)$ est appelée **morphisme d'évaluation** en A.

Remarque 2 — *Soit E un $\mathbb{K}$-e.v. de dimension n. Soient $u \in \mathcal{L}(E)$, et $\mathcal{B}$ une base de E. On a déjà vu que $\operatorname{Mat}_{\mathcal{B}}(u^k) = \operatorname{Mat}_{\mathcal{B}}(u)^k$, pour tout $k \geq 0$.*
Ainsi, pour tout polynôme $P \in \mathbb{K}[X]$, on a :

$$\operatorname{Mat}_{\mathcal{B}}(P(u)) = P\left(\operatorname{Mat}_{\mathcal{B}}(u)\right).$$

◀

Exemples 3

1. *Soit E un $\mathbb{K}$-e.v. et $u : E \to E$ un endomorphisme. Alors p est une projection si, et seulement si, $P(p) = 0$ où $P = X^2 - X$.*
2. *Soit $\lambda \in \mathbb{K}$. Pour $u = \lambda \cdot id_E$ et $P \in \mathbb{K}[X]$, on trouve alors $P(u) = P(\lambda) \cdot id_E$.*
3. *Pour $A = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$ et $P(X) = (X-1)(X-2) = X^2 - 3X + 2$, on a :*

$$P(A) = A^2 - 3A + 2 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} = 0.$$

4. *Pour $E = \mathcal{C}^\infty(\mathbb{R}, \mathbb{C})$, on pose $u : f \in E \mapsto f' \in E$ l'endomorphisme de dérivation. Soit $P = \sum_{k=0}^n a_k X^k \in \mathbb{C}[X]$. Alors, l'endomorphisme $P(u)$ est l'opérateur différentiel sur $\mathcal{C}^\infty(\mathbb{R}, \mathbb{C})$ suivant :*

$$\begin{array}{rcl} P(u) \,:\, \mathcal{C}^\infty(\mathbb{R}, \mathbb{C}) & \longrightarrow & \mathcal{C}^\infty(\mathbb{R}, \mathbb{C}) \\ f & \longmapsto & a_0 f + a_1 f' + \cdots + a_n f^{(n)}. \end{array}$$

5. *L'évaluation de $P(X) = 3 - 2X + X^2$ en $A = \begin{pmatrix} 1 & -1 \\ 1 & 2 \end{pmatrix}$ est :*

$$P(A) = 3\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - 2\begin{pmatrix} 1 & -1 \\ 1 & 2 \end{pmatrix} + \begin{pmatrix} 1 & -1 \\ 1 & 2 \end{pmatrix}^2 = \begin{pmatrix} 1 & -1 \\ 1 & 2 \end{pmatrix} = A.$$

Proposition 4

Soient E un $\mathbb{K}$-e.v. et $u : E \to E$ un endomorphisme. Soient $P, Q \in \mathbb{K}[X]$, $\lambda \in \mathbb{K}$. Alors on a :

1. $(\lambda P)(u) = \lambda P(u)$;
2. $(P + Q)(u) = P(u) + Q(u)$;
   Le morphisme d'évaluation en u, $P \in \mathbb{K}[X] \mapsto P(u) \in \mathcal{L}(E)$, est une application linéaire.
3. $(PQ)(u) = P(u) \circ Q(u)$;
   Le morphisme d'évaluation en u, $P \in \mathbb{K}[X] \mapsto P(u) \in \mathcal{L}(E)$, est un morphisme d'anneaux.

Soient $n \geq 1$ et $A \in \mathscr{M}_n(\mathbb{K})$.
Alors le morphisme d'évaluation en A, $P \in \mathbb{K}[X] \mapsto P(A) \in \mathscr{M}_n(\mathbb{K})$, est une application linéaire et un morphisme d'anneaux.

Preuve — En prenant $m = \max(\deg(P), \deg(Q))$, écrivons P et Q de la forme :

$$P = \sum_{k=0}^m a_k X^k \quad \text{et} \quad Q = \sum_{k=0}^m b_k X^k.$$

- Soit $\lambda \in \mathbb{K}$. On a :
$$(\lambda P)(u) = \lambda a_0 Id_E + \lambda a_1 u + \ldots + \lambda a_n u^n = \lambda P(u).$$
- On a :
$$\begin{array}{rcl} (P+Q)(u) & = & \left(\sum_{k=0}^m (a_k + b_k) X^k\right)(u) = \sum_{k=0}^m (a_k + b_k) u^k \\ & = & \sum_{k=0}^m a_k u^k + \sum_{k=0}^m b_k u^k = P(u) + Q(u). \end{array}$$
  Cela montre aussi que $P \mapsto P(u)$ est une application linéaire.
- Vérifions d'abord la dernière relation pour des monômes. Soient $P_1(X) = X^p$ et $Q_1(X) = X^q$, $p, q \in \mathbb{N}$. On a :
$$(P_i Q_1)(u) = (X^{p+q})(u) = u^{p+q} = u^p \circ u^q = P_1(u) \circ P_2(u).$$
  Démontrons maintenant le cas général. On a :
$$\begin{array}{rcl} P(u) \circ Q(u) & = & (\sum_{k=0}^n a_k u^k) \circ (\sum_{l=0}^n b_l u^l) = \sum_{k=0}^n a_k (u^k \circ (\sum_{l=0}^n b_l u^l)) \\ & = & \sum_{k=0}^n a_k (\sum_{l=0}^n b_l u^k \circ u^l)), \text{ par linéarité de } u^k \\ & = & \sum_{k=0}^n \sum_{l=0}^n a_k b_l u^{k+l}) = \sum_{k=0}^n \sum_{l=0}^n a_k b_l X^{k+l}(u) \\ & = & (\sum_{k=0}^n \sum_{l=0}^n a_k b_l X^{k+l})(u), \text{ d'après les points précédents} \\ & = & (PQ)(u). \end{array}$$
  On vérifie de plus que $1(u) = u^0 = Id_E$, pour en déduire que $P \mapsto P(u)$ est un morphisme d'anneaux.

La preuve est identique pour une matrice $A \in \mathscr{M}_n(\mathbb{K})$ (ou bien on peut utiliser le lien entre A et l'endomorphisme $X \mapsto AX$). □

Remarque 5 — *Pour $u : E \to E$ un endomorphisme, l'image du morphisme d'évaluation $P \mapsto P(u)$ est donc un sous-anneau de $\mathcal{L}(E)$ contenant u, et même une sous-algèbre de la $\mathbb{K}$-algèbre $(\mathcal{L}(E), +, \circ, .)$.*

Proposition 6
Soient E un $\mathbb{K}$-e.v. et $u : E \to E$ un endomorphisme.
On note $\mathbb{K}[u]$ l'ensemble $\{P(u),\, P \in \mathbb{K}[X]\}$ des polynômes en u. Alors :

- $\mathbb{K}[u]$ est une sous-algèbre de $\mathcal{L}(E)$, qui est commutative ;
- $\mathbb{K}[u]$ est la plus petite sous-algèbre de $\mathcal{L}(E)$ qui contient u.

Soient $n \geq 1$ et $A \in \mathscr{M}_n(\mathbb{K})$.
On note $\mathbb{K}[A]$ l'ensemble $\{P(A),\, P \in \mathbb{K}[X]\}$ des polynômes en A.
Alors, $\mathbb{K}[A]$ est une sous-algèbre commutative de $\mathscr{M}_n(\mathbb{K})$. C'est la plus petite sous-algèbre de $\mathscr{M}_n(\mathbb{K})$ qui contient A.

Preuve — $\mathbb{K}[u]$ est l'image de $\mathbb{K}[X]$ par le morphisme d'évaluation $P \mapsto P(u) \in \mathcal{L}(E)$. Comme cette fonction est un morphisme d'anneaux et une application linéaire, c'est un morphisme d'algèbre. Son image est alors une sous-algèbre de $(\mathcal{L}(E), +, \circ, .)$.

Pour $v_1, v_2 \in \mathbb{K}[u]$, on a $P, Q \in \mathbb{K}[X]$ tels que $v_1 = P(u)$,$v_2 = Q(u)$. Cela donne :

$$v_1 \circ v_2 = P(u) \circ Q(u) = (PQ)(u) = (QP)(u) = Q(u) \circ P(u) = v_2 \circ v_1,$$

donc v_1 et v_2 commutent.

Enfin, soit S une sous-algèbre de $\mathcal{L}(E)$ contenant u. Alors S contient Id_E et les u^k, pour tout $k \leq 1$. Donc S contient toutes les combinaisons linéaires de $Id_E, u, u^2, \ldots, u^k, \ldots$. Donc S contient tous les polynômes en u. Donc S contient $\mathbb{K}[u]$.
La preuve est identique pour une matrice $A \in \mathscr{M}_n(\mathbb{K})$. □

Remarque 7 — *Nous avons vu à plusieurs reprises que les anneaux $\mathcal{L}(E)$ et $\mathscr{M}_n(\mathbb{K})$ ne sont pas commutatifs, ce qui empêche d'effectuer certaines opérations et rend faux certains résultats.*
Le fait que les sous-anneaux $\mathbb{K}[u]$ et $\mathbb{K}[A]$ soient commutatifs est très important.

Remarque 8 — *Soient E un $\mathbb{K}$-ev avec $\dim E \geq 2$ et $u : E \to E$ un endomorphisme. Alors le morphisme d'évaluation n'est pas surjectif car l'anneau $\mathcal{L}(E)$ n'est pas commutatif alors que $\mathbb{K}[u]$ oui. (pareil pour $\mathscr{M}_n(\mathbb{K})$ avec $n \geq 2$)*

Remarque 9 — *Soit E un $\mathbb{K}$-ev de dimension finie. Soit $u : E \to E$ un endomorphisme. Alors $\mathbb{K}[u]$ est un sous-ev de $\mathcal{L}(E)$, qui est un $\mathbb{K}$-ev de dimension finie.*

Donc, $\mathbb{K}[u]$ est un ev de dimension finie. Or, $\mathbb{K}[X]$ est un $\mathbb{K}$-ev de dimension infinie. Le morphisme d'évaluation $P \mapsto P(u)$ n'est donc pas injectif.

Cela veut dire que pour tout $v \in \mathbb{K}[u]$, il existe plusieurs polynômes P, Q tels que $P(u) = Q(u)$. Autrement dit, l'écriture $v = P(u)$ n'est pas unique.

Exemple 10 — *Soit $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \in \mathscr{M}_2(R)$. La matrice $B = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ appartient à $\mathbb{R}[A]$ car on a :*

$$B = P(A) \text{ où } P = (X-1).$$

Mais on a aussi :

$$B = Q(A) \text{ où } Q = (X-1)^2 + (X-1).$$

Proposition 11
Soient E un $\mathbb{K}$-e.v., et u un endomorphisme sur E. Soit ϕ un endomorphisme inversible sur E. Alors, on a :

$$P(\phi u \phi^{-1}) = \phi P(u) \phi^{-1},\ \forall P \in \mathbb{K}[X].$$

Soient $n \geq 1$ et $A, P \in \mathscr{M}_n(\mathbb{K})$ avec P inversible. Alors, on a :

$$P(PAP^{-1}) = PP(A)P^{-1},\ \forall P \in \mathbb{K}[X].$$

Si deux endomorphismes u, v (ou matrices A, B) sont semblables, alors $P(u)$ et $P(v)$ (ou $P(A)$ et $P(B)$) sont semblables.

Preuve — Soit $k \geq 1$. On a

$$(\phi u \phi^{-1})^k = (\phi u \phi^{-1})(\phi u \phi^{-1}) \ldots (\phi u \phi^{-1}) = \phi u^k \phi^{-1}.$$

Ainsi, pour $P(X) = a_0 + \ldots + a_n X^n$, on a :

$$P(\phi u \phi^{-1}) = a_0 Id_E + a_1 \phi u \phi^{-1} + \ldots + a_n (\phi u \phi^{-1})^n = \phi(a_0 Id_E + a_1 u + \ldots + a_n u^n)\phi^{-1} = \phi P(u) \phi^{-1}.$$

La preuve est identique pour les matrices. □

Proposition 12
Soient $n \geq 1$ et $A \in \mathscr{M}_n(\mathbb{K})$. Alors :

$$P({}^tA) = {}^tP(A), \forall P \in \mathbb{K}[X].$$

Preuve — On remarque que pour tout $k \geq 1$, on a ${}^tA^k = {}^tA.A.\ldots, A = {}^tA.{}^{tA} \ldots {}^tA = {}^tA^k$.

Ainsi, pour $P(X) = a_0 + \ldots + a_n X^n$, on a :

$${}^tP(A) = {}^ta_0 I_n + a_1 A + \ldots + a_n A^n = a_0 I_n + a_1 {}^tA + \ldots + a_n {}^tA^n = P({}^tA).$$

□

Proposition 13 (Polynômes de matrices triangulaires)
Soit $n \geq 1$. Soit $A \in \mathscr{M}_n(\mathbb{K})$ une matrice triangulaire supérieure par blocs :

$$A = \begin{pmatrix} A_1 & * & \ldots & & * \\ 0 & A_2 & * & & \vdots \\ 0 & \ddots & \ddots & & \vdots \\ 0 & \ldots & 0 & A_{r-1} & * \\ 0 & \ldots & & 0 & A_r \end{pmatrix}, \; A_i \in \mathscr{M}_{n_i}(\mathbb{K}), \; n_1 + \ldots + n_r = n.$$

Alors, pour tout $P \in \mathbb{K}[X]$, $P(A)$ est une matrice triangulaire supérieure par blocs :

$$P(A) = \begin{pmatrix} P(A_1) & * & \ldots & & * \\ 0 & P(A_2) & * & & \vdots \\ 0 & \ddots & \ddots & & \vdots \\ 0 & \ldots & 0 & P(A_{r-1}) & * \\ 0 & \ldots & & 0 & P(A_r) \end{pmatrix}$$

Preuve — La définition du produit matriciel fait que, pour tout $k \geq 1$, on a (démontré dans des cours précédents) :

$$A^k = \begin{pmatrix} A_1^k & * & \ldots & & * \\ 0 & A_2^k & * & & \vdots \\ 0 & \ddots & \ddots & & \vdots \\ 0 & \ldots & 0 & A_{r-1}^k & * \\ 0 & \ldots & & 0 & A_r^k \end{pmatrix}.$$

On a de plus $A^0 = I_n = \mathrm{Diag}(I_{n_1}, \ldots, I_{n_r}) = \mathrm{Diag}(A_1^0, \ldots A_r^0)$.
On obtient ainsi le résultat par linéarité. □

Corollaire 14
Soient E un $\mathbb{K}$-e.v. de dimension n et $u : E \to E$ un endomorphisme.
Si u est diagonalisable, alors $P(u)$ est diagonalisable pour tout $P \in \mathbb{K}[X]$.

Preuve — Si u est diagonalisable, il existe une base $\mathcal{B}$ de E telle que $\mathrm{Mat}_{\mathcal{B}}(u)$ soit diagonale.

Soit $P \in \mathbb{K}[X]$. D'après la proposition précédente, la matrice $\mathrm{Mat}_{\mathcal{B}}(P(u)) = P(\mathrm{Mat}_{\mathcal{B}}(u))$ est elle aussi diagonale, donc $P(u)$ est diagonalisable. □

Les deux propositions suivantes vont se révéler par la suite particulièrement utile.

Proposition 15
Soient E un $\mathbb{K}$-e.v. et $u \in \mathcal{L}(E)$. Soit F un sous-espace stable par u.
Alors, pour tout $P \in \mathbb{K}[X]$, le sous-espace F est stable par $P(u)$. On a de plus :

$$P(u)_F = P(u_F).$$

Preuve — Soit $x \in F$ et soit $P = \sum_{i=0}^n a_i X^i$. Alors : $P(u)(x) = \sum_{i=0}^n a_i u^i(x)$.

Comme F est stable par u, les vecteurs $u^i(x)$ appartiennent à F, donc $P(u)(x)$ est une combinaison linéaire d'éléments de F. Ainsi, $P(u)(x) \in F$ et F est stable par $P(u)$.

Soit $x \in F$. Par définition de $P(u)_F$, on a :

$$P(u)_F(x) = P(u)(x) = \sum_{i=0}^{n} a_i u^i(x) = \sum_{i=0}^{n} a_i u_F^i(x) = P(u_F)(x).$$

Ce qui conclut. □

Proposition 16
Soient E un $\mathbb{K}$-e.v., $u \in \mathcal{L}(E)$, et $P \in \mathbb{K}[X]$.
Alors, les sous-espaces vectoriels $\operatorname{Im}(P(u))$ et $\operatorname{Ker}(P(u))$ sont stables par u.

Preuve — Les sous-ev $\operatorname{Im}(P(u))$ et $\operatorname{Ker}(P(u))$ sont stables par $P(u)$. Comme u et $P(u)$ commutent, ces sous-ev sont stables par u. □

4.2 Idéal annulateur et Polynôme minimal

4.2.1 Polynômes annulateurs

Définition 17
Soient E un $\mathbb{K}$-e.v., $u \in \mathcal{L}(E)$, et $P \in \mathbb{K}[X]$.
On dit que P est un **polynôme annulateur**\零化多项式\ **de** u si $P(u) = 0$.
Soient $n \geq 1$ et $A \in \mathscr{M}_n(\mathbb{K})$.
On dit que P est un **polynôme annulateur**\零化多项式\ **de** A si $P(A) = 0$.

Remarque 18 — *Pour $e_u : P \in \mathbb{K}[X] \mapsto P(u) \in \mathcal{L}(X)$, l'ensemble des polynômes annulateurs de u est exactement $Ker(e_u)$.*
Comme e_u est une application linéaire et un morphisme d'anneaux, ensemble est ainsi un sous-ev de $\mathbb{K}[X]$ et un idéal de $\mathbb{K}[X]$.

On rappelle que les idéaux de $\mathbb{K}[X]$ sont de la forme $M.\mathbb{K}[X]$, pour un $M \in \mathbb{K}[X]$ (ces idéaux sont principaux). De plus :

- *Si $M = 0$ on a $M\mathbb{K}[X] = \{0\}$;*
- *Si $M \neq 0$, alors il existe un unique polynôme unitaire N tel que $N\mathbb{K}[X] = M\mathbb{K}[X]$.*

Nous verrons qu'en dimension infinie, un endomorphisme peut avoir ou ne pas avoir de polynôme annulateur non nul. En dimension finie, tout endomorphisme possède un polynôme annulateur non nul.

Exemples 19

1. *Pour tout endomorphisme u, le polynôme nul $P = 0$ est un polynôme annulateur de u.*
2. *Le polynôme constant $P = 1$ n'est le polynôme annulateur d'aucun endomorphisme u.*
3. *Pour p une projection, p est annulée par le polynôme $X^2 - X$ (car $p^2 = p$).*
4. *Pour s une symétrie, s est annulée par le polynôme $X^2 - 1$ (car $s^2 = Id_E$).*
5. *Pour $u = \lambda Id_E$, $P(X) = (X - \lambda)$ est un polynôme annulateur de u.*
6. *Soit $A \in \mathscr{M}_n(\mathbb{K})$ une matrice diagonalisable. On a $P \in \mathscr{M}_n(\mathbb{K})$ inversible et $\lambda_1, \ldots, \lambda_n \in \mathbb{K}$ tels que $A = PDiag(\lambda_1, \ldots, \lambda_n)P^{-1}$.*
   Posons $Q(X) = (X - \lambda_1) \ldots (X - \lambda_n)$. On a alors :
   $$\begin{aligned} Q(A) &= Q(PDiag(\lambda_1, \ldots, \lambda_n)P^{-1}) = PQ(Diag(\lambda_1, \ldots, \lambda_n))P^{-1} \\ &= PDiag(Q(\lambda_1), \ldots, Q(\lambda_n))P^{-1} = PDiag(0, \ldots, 0)P^{-1} \\ &= 0, \end{aligned}$$
   donc Q est un polynôme annulateur de A.
   D'après les résultats du chapitre précédent sur les matrices diagonalsables, on remarque que $Q(X) = \pi_{i=1}^{n}(X - \lambda_i) = \chi_A(X)$. Donc χ_A est un polynôme annulateur de A. Nous reviendrons sur ce résultat dans ce chapitre (Théorème de Cayley-Hamilton).
7. *Dans $E = \mathcal{C}^\infty(\mathbb{R}, \mathbb{C})$, l'endomorphisme de dérivation $D : f \mapsto f'$ ne possède pas de polynôme annulateur non nul. Soit $P \in \mathbb{C}[X]$. Pour tout $\lambda \in \mathbb{C}$, posons $f_\lambda : x \mapsto e^{\lambda x}$.*
   Alors on a $D(f_\lambda) = \lambda f_\lambda$, donc $P(D)(f_\lambda) = P(\lambda) f_\lambda$.
   Si l'on a $P(D) = 0$, alors on a $P(\lambda) = 0$ pour tout $\lambda \in \mathbb{C}$, ce qui implique que $P(X) = 0$.

4.2.2 Polynôme minimal, cas de la dimension finie

PROPOSITION 20
Soit E un $\mathbb{K}$-e.v. de dimension de n. Soit $u \in \mathcal{L}(E)$.
Alors il existe $P \in \mathbb{K}[X]$ non-nul tel que $P(u) = 0$.
Soit $A \in \mathscr{M}_n(\mathbb{K})$.
Alors la matrice A possède un polynôme annulateur non-nul.

Preuve — L'espace vectoriel $\mathcal{L}(E)$ est de dimension n^2, donc la famille

$$\left(u^0, u, u^2, \ldots, u^{n^2}\right)$$

est liée. Il existe donc une relation de dépendance linéaire non triviale :

$$\lambda_0 \operatorname{Id}_E + \lambda_1 u + \ldots + \lambda_{n^2} u^{n^2} = 0_{\mathcal{L}(E)}, \text{ où } \lambda_0, \ldots, \lambda_{n^2} \in \mathbb{K} \text{ ne sont pas tous nuls .}$$

Le polynôme

$$P = \sum_{i=0}^{n^2} \lambda_i X^i$$

est ainsi un polynôme annulateur non nul de u.

La preuve est identique pour la matrice A. □

DÉFINITION 21
Soit E un $\mathbb{K}$-e.v. Soit $u : E \to E$ un endomorphisme qui possède un polynôme annulateur non-nul. Pour $e_u : P \in \mathbb{K}[X] \mapsto P(u) \in \mathcal{L}(E)$, le noyau de ce morphisme n'est pas réduit à $\{0\}$.
Il existe donc un unique polynôme $M \in \mathbb{K}[X]$, unitaire, tel que $\operatorname{Ker}(e_u) = M\mathbb{K}[X]$.
Ce polynôme est appelé **polynôme minimal** de u \最小多项式\. On le note μ_u, ou M_u.

REMARQUE 22 — *Si E est de dimension finie, alors tout endomorphisme u sur E possède un polynôme minimal. Pour u un endomorphisme possédant un polynôme minimal, on a :*

$$P(u) = 0 \text{ si et seulement si } P \mid \mu_u.$$

Le polynôme minimal de u, μ_u, est le polynôme unitaire annulant de u de plus petit degré.
*Pour $A \in \mathscr{M}_n(\mathbb{K})$, on définit de même le **polynôme minimal** de A que l'on note μ_A.*

La question qui se pose alors est : comment calculer le polynôme annulateur d'un endomorphisme ou d'une matrice en dimension finie ?
Ce problème est très souvent difficile. Il existe cependant une méthode générale qui fonctionne, mais elle est trop longue en général pour être utilisée d'un point de vue algorithmique :

1. On commence par calculer un polynôme annulateur de u ou de A. Nous verrons dans le chapitre suivant comment trouver un tel polynôme dans le cas de la dimension finie (théorème de Cayley-Hamilton).
2. Pour P un polynôme annulateur de u (ou A), on factorise P dans $\mathbb{K}[X]$.
3. Parmi tous les diviseurs de P, on cherche ceux qui annulent u (ou A) et qui sont de plus petit degré possible.

PROPOSITION 23
Soit E un $\mathbb{K}$-e.v. de dimension n. Soient $u, v \in \mathcal{L}(E)$ avec v inversible. Soit $\mathcal{B}$ une base de E. On a :

1. $$\mu_{\operatorname{Mat}_{\mathcal{B}}(u)}(X) = \mu_u(X) ;$$
2. $$\mu_{vuv^{-1}}(X) = \mu_u(X) ;$$
3. Soient $A, P \in \mathscr{M}_n(\mathbb{K})$, avec P inversible. Alors :
$$\mu_{PAP^{-1}}(X) = \mu_A(X).$$
Le polynôme minimal est un invariant de similitude.
4. On a de plus :
$$\mu_{{}^tA}(X) = \mu_A(X).$$

Preuve — Soit $P \in \mathbb{K}[X]$. Les résultats précédents donnent :

1. $P(u) = 0$ si et seulement si $0 = \mathrm{Mat}_{\mathcal{B}}(P(u)) = P(\mathrm{Mat}_{\mathcal{B}}(u))$.
2. $P(vuv^{-1}) = vP(u)v^{-1}$. Donc $P(PAP^{-1}) = 0$ si et seulement si $P(u) = 0$.
3. $P(PAP^{-1}) = P.P(A).P^{-1}$. Donc $P(PAP^{-1}) = 0$ si et seulement si $P(A) = 0$.
4. $P(A) = 0$ si et seulement si $0 = {}^tP(A) = P({}^tA)$.

Comme le polynôme minimal d'un endomorphisme f est l'unique polynôme annulateur de f qui est non-nul, unitaire, et de degré minimal, cela conclut la preuve. □

Exemples 24

1. *Pour $u = 0$ l'endomorphisme nul, on a $\mu_0(X) = X$.*
2. *On a $\mu_{Id_E}(X) = X - 1$.*
3. *Le polynôme minimal d'un endomorphisme est un polynôme de degré supérieur ou égal à 1.*
4. *Pour u un endomorphisme, on a $u = \lambda Id_E$ si et seulement si $\mu_u(X) = X - \lambda$.* ◀
5. *Soit p une projection. Si $p = 0$ on a $\mu_p(X) = X$. Si $p = Id_E$ on a $\mu_p(X) = X - 1$. Sinon, on a $\mu_p(X) = X^2 - X$.*
   En effet, $X^2 - X$ est un polynôme annulateur de p, donc son polynôme minimal μ_p est un diviseur unitaire de $X^2 - X$, de degré au moins 1. Ce polynôme vaut donc $X^2 - X$ ou X ou $X - 1$. Comme on a supposé $p \neq 0$ et $p \neq Id_E$, on en déduit que $\mu_p(X) = X^2 - X$.
6. *De la même façon, si s est une symétrie différente de id et de $-\mathrm{id}$, alors $\mu_s(X) = X^2 - 1$.*
7. *Pour u un endomorphisme nilpotent ($\exists k \geq 1$ tel que $u^k = 0$), on a $\mu_u(X) = X^r$, où r est l'indice de nilpotence de u. Réciproquement, si $\mu_u(X) = X^r$, alors u est nilpotent d'indice r.*

Nous terminons cette section par une propriété très utile que nous avons déjà vue pour le polynôme caractéristique χ_u.

Proposition 25
Soient E un $\mathbb{K}$-e.v. et u un endomorphisme sur E. Soit F un sous-ev de E stable par u.
Si u admet un polynôme minimal, alors l'endomorphisme induit u_F possède un polynôme minimal, et l'on a :

$$\mu_{u_F} \text{ divise } \mu_u.$$

Preuve — Soit $P \in \mathbb{K}[X]$. On a vu que $P(u_F) = P(u)_F$. Ainsi, μ_u est un polynôme annulateur de u_F. Donc u_F possède un polynôme minimal, et celui-ci divise μ_u. □

Exemple 26 — *Soit A la matrice diagonale par blocs suivante :*

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 3 \\ 0 & 0 & 2 \end{pmatrix} \in \mathscr{M}_3(\mathbb{R}).$$

On associe A à l'endomorphisme $u : X \in \mathbb{R}^3 \mapsto AX \in \mathbb{R}^3$. Comme A est diagonale par blocs, on sait alors que les sous-ev $F_1 = \mathbb{R}e_1$ et $F_2 = \mathrm{Vect}(e_2, e_3)$ sont stables par u.

- *On a $u_{F_1} = Id_{F_1}$, donc $\mu_{u_{F_1}} = X - 1$.*
- *Sur F_2, on a :*
  $$\mathrm{Mat}_{(e_2,e_3)}(u_{F_2}) = \begin{pmatrix} 2 & 3 \\ 0 & 2 \end{pmatrix}.$$
  Comme
  $$\left(\begin{pmatrix} 2 & 3 \\ 0 & 2 \end{pmatrix} - 2I_2\right)^2 = 0,$$
  on en déduit que $(X-2)^2$ est un polynôme annulateur de u_{F_2}. Comme $(X-2)$ n'est pas un polynôme annulateur de u_{F_2}, on a alors $\mu_{u_{F_2}} = (X-2)^2$.

Ainsi le polynôme minimal de A est un multiple de $(X-1)$ et de $(X-2)^2$. Comme $(X-1)(X-2)^2$ est un polynôme annulateur de A, on en déduit que $\mu_A = (X-1)(X-2)^2$.

4.2.3 Calculs de polynômes d'endomorphismes ou de matrices

Proposition 27
Soient E un $\mathbb{K}$-e.v. et u un endomorphisme sur E.
Si u possède un polynôme annulateur P tel que $P(0) \neq 0$, alors u est inversible.
Pour $P(X) = a_0 + \ldots + a_n X^n$, on a $u^{-1} = \sum_{k=1}^{n} \frac{-a_k}{a_0} u^{k-1}$.

Preuve — On a :

$$0 = P(u) = a_0 Id_E + a_1 u + \ldots + a_n u^n = a_0 Id_E + \sum_{k=1}^{n} a_k u^k = a_0 Id_E + u(\sum_{k=1}^{n} a_k u^{k-1}).$$

Ainsi, on obtient :

$$Id_E = u(\frac{-a_k}{a_0} u^{k-1}) = (\frac{-a_k}{a_0} u^{k-1})u,$$

donc u est inversible, d'inverse $u^{-1} = \frac{-a_k}{a_0} u^{k-1}$. □

PROPOSITION 28
Soient E un $\mathbb{K}$-e.v. et u un endomorphisme sur E qui a un polynôme annulateur non-nul P. Soit $M \in \mathbb{K}[X]$. Soit $R \in \mathbb{K}[X]$ le reste de la division euclidienne de M par P. On a alors :

$$M(u) = R(u).$$

Ainsi, tout polynôme en u est égal à un polynôme de degré au plus $\deg(\mu_u) - 1$ en u, que l'on peut déterminer à l'aide d'une division euclidienne.

Preuve — Soit $M = PQ + R$, $\deg(R) < \deg(P)$ la division euclidienne de M par P. On a alors :

$$M(u) = (PQ)(u) + R(u) = Q(u) \circ P(u) + R(u) = 0 + R(u),$$

ce qui conclut. □

La connaissance d'un polynôme annulateur non nul d'un endomorphisme u, ou de son polynôme minimal, est ainsi très utile. Nous verrons par la suite en quoi ces polynômes donnent beaucoup d'informations sur u.

EXEMPLE 29 — *Soit $n \geqslant 2$. On considère la matrice :*

$$J = \begin{pmatrix} 0 & 1 & \ldots & 1 \\ 1 & 0 & (1) & \vdots \\ \vdots & (1) & \ddots & 1 \\ 1 & \ldots & 1 & 0 \end{pmatrix}.$$

- *La relation $(I_n + J)^2 = n(I_n + J)$ montre que l'on a :*

$$J^2 = (n-1)I_n + (n-2)J.$$

  Ainsi le polynôme $P(X) = X^2 - (n-2)X - (n-1) = (X+1)(X-(n-1))$ est un polynôme annulateur de J. Vu que J n'est pas une matrice diagonale, elle ne peut être annulée par un polynôme de degré 1. On a donc $\mu_J(X) = X^2 - (n-2)X - (n-1) = (X+1)(X-(n-1))$.
  Comme $\mu_J(0) \neq 0$, J est inversible, d'inverse :

$$J^{-1} = \frac{1}{n-1}\left(J - (n-2)I_n\right).$$

- *On calcule alors J^k, pour $k \in \mathbb{N}$, de la manière suivante :*
  — *On effectue la division euclidienne de X^k par μ_J :*

$$X^k = \mu_J Q + \alpha_k X + \beta_k.$$

  On obtient α_k et β_k en évaluant X^k en -1 et $n-1$. Cela donne :

$$\alpha_k = \frac{1}{n}\left((n-1)^k - (-1)^k\right) \quad \text{et} \quad \beta_k = (-1)^k + \frac{1}{n}\left((n-1)^k - (-1)^k\right).$$

  — *On obtient ainsi :*

$$J^k = \alpha_k J + \beta_k I_n = (n-1)^k \frac{1}{n}(I_n + J) + (-1)^k \left(\frac{n-1}{n} I_n - \frac{1}{n} J\right) \quad \forall k \in \mathbb{N}^*.$$

Nous terminons notre étude sur le polynôme minimal par un exemple que l'on retrouve fréquemment, et qui se révèle très utile : celui de la *matrice compagnon*[1].

1. Voir la proposition-définition 45.

Proposition 30
Soit $n \geq 1$. Soit $P(X) = a_0 + a_1X + \ldots + X^n$ un polynôme unitaire de degré n. Soit $C_P \in \mathscr{M}_{n+1}(\mathbb{K})$ la matrice compagnon de P.
Alors, on a $\mu_{C_P}(X) = P(X)$.

Preuve — Notons $(e_0, \ldots, e_{n-1})$ la base canonique de $\mathbb{K}^n$. La forme de la matrice C_P donne alors :

$$C_P e_i = e_{i+1},\ \forall i = 0, \ldots, n-2 \quad \text{et} \quad C_p e_{n-1} = -\sum_{k=0}^{n-1} \alpha_k e_k.$$

On a donc :

$$e_1 = C_P e_0,\ e_2 = C_P^2 e_0,\ \ldots,\ e_{n-1} = C_P^{n-1} e_0.$$

et

$$C_P^n e_0 = -\alpha_0 e_0 - \ldots - \alpha_{n-1} C_P^{n-1} e_0.$$

On obtient alors que :

$$P(C_p)e_0 = 0.$$

Ainsi, pour tout $k = 1, \ldots, n-1$, on a :

$$P(C_P)e_k = P(C_P)C_P^k e_0 = C_P^k P(C_P)e_0 = 0.$$

Cela montre que le polynôme P est un polynôme annulateur de C_P.

Soit $R = \sum_{k=0}^r \beta_k X^k$ un polynôme, avec $r < \deg(P) = n$,et $R(C_P) = 0$. On a alors :

$$R(C_P)e_0 = 0 = \sum_{k=0}^r \beta_k C_P^k e_0 = \sum_{k=0}^r \beta_k e_k.$$

Il s'agit d'une relation de dépendance linéaire portant les vecteurs $e_1, \ldots, e_r$. Ces vecteurs forment une famille libre, donc les coefficients β_k sont nuls, ce qui donne $R = 0$. Le polynôme P est donc le polynôme unitaire annulateur de C_P de plus petit degré, donc $P = \mu_{C_P}$. □

Remarque 31 — *Soient E un $\mathbb{K}$-e.v. de dimension n et $u \in \mathcal{L}(E)$ un endomorphisme. En prenant $x \in E$ non-nul, on peut déterminer le sous-ev $S_u(x) = Vect(x^k,\ k \geq 0)$ en trouvant le plus petit entier k tel que la famille $(x, u(x), \ldots, u^k(x))$ soit liée.*
On a vu dans le chapitre précédent qu'il existe alors $a_0, \ldots, a_{k-1} \in \mathbb{K}$ tels que $u^k(x) = -a_0x - a_1u(x) - \ldots - a_{k-1}u^{k-1}(x)$. De plus, en notant $P(X) = a_0 + a_1X + \ldots + a_{k-1}X^{k-1} + X^k$, on a $Mat_{(x,u(x),\ldots,u^{k-1}(x))}(u) = C_P$ et $\chi_{u_{S_u(x)}} = C_P$.

La proposition précédente nous dit alors que $P = \mu_{u_{S_u(x)}}$. Le polynôme P est donc un diviseur du polynôme minimal de u, μ_u.

4.3 Polynômes d'endomorphisme & éléments propres

Les notions développées dans cette section concernent les endomorphismes. Elles s'appliquent aux matrices de $\mathscr{M}_n(\mathbb{K})$ en considérant les endomorphismes canoniquement associés sur $\mathbb{K}^n$.

4.3.1 Valeurs propres

Proposition 32
Soient E un $\mathbb{K}$-e.v., $u \in \mathcal{L}(E)$, $\lambda \in \operatorname{Spec}(u)$, et $P \in \mathbb{K}[X]$.
Pour tout $x \in E_\lambda(u)$, on a $P(u)(x) = P(\lambda)x$.
Donc, $P(\lambda) \in \operatorname{Spec}(P)(u)$.

Preuve — Comme $x \in E_\lambda(u)$, on a $u^k(x) = \lambda^k x$ pour tout k, d'où :

$$P(u)(x) = \left(\sum_{k=0}^p \alpha^k u^k\right)(x) = \sum_{k=0}^p \alpha^k u^k(x) = \left(\sum_{k=0}^p \alpha^k \lambda^k\right) x = P(\lambda)x.$$

En prenant $x \neq 0$, on obtient donc $P(u)(x) = P(\lambda)x$ avec $x \neq 0$, ce qui conclut. □

Corollaire 33
Soient E un $\mathbb{K}$-e.v. et $u \in \mathcal{L}(E)$. Soit P un polynôme annulateur de u.
Alors, les valeurs propres de u sont des racines de P.

Preuve — Les relations $P(u) = 0$ et $u(x) = \lambda x$ avec $x \neq 0$ entraînent $P(\lambda)x = 0$ et donc $P(\lambda) = 0$. □

EXEMPLE 34 — *Soit $E = F \oplus G$ un espace vectoriel avec une décomposition en somme directe. On suppose que F et G ne sont pas résuits à $\{0\}$. Soit $\alpha \in \mathbb{K}$. On définit l'application linéaire :*

$$\begin{array}{rccl} a : E & \longrightarrow & E & \\ x & \longmapsto & x_F + \alpha x_G, & \text{où} \quad x = x_F + x_G, \quad x_F \in F, \quad x_G \in G. \end{array}$$

Le polynôme $(X-1)(X-\alpha)$ est un polynôme annulateur de a, donc le spectre de a est contenu dans $\{1, \alpha\}$. ◄
Comme les noyaux $\mathrm{Ker}\left((a - \mathrm{Id}_E)\right) = F$ et $\mathrm{Ker}\left((a - \alpha \mathrm{Id}_E)\right) = G$ sont non nuls, on a $\mathrm{Spec}\,(a) = \{1, \alpha\}$, avec $E_1(a) = F$ et $E_\alpha(a) = G$.
Pour $\alpha = 0$ on retrouve la projection sur F parallèlement à G, et pour $\alpha = -1$ on retrouve la symétrie par rapport à F parallèlement à G.

Lorsque u possède un polynôme annulateur (par exemple en dimension finie) on a le résultat suivant qui donne une caractérisation du spectre de u :

THÉORÈME 35
Soient E un $\mathbb{K}$-e.v. et $u \in \mathcal{L}(E)$ possédant un polynôme annulateur non-nul.
Alors, les valeurs propres de u sont exactement les racines de $\mu_u(X)$ dans $\mathbb{K}$.

Preuve —

- Comme $\mu_u(u) = 0$, les valeurs propres de u sont des racines de μ_u.
- Soit α une racine de $M\mu_u$. Ona donc $M_u(X) = (X-\alpha)N(X)$. Par minimalité de μ_u, l'endomorphisme $N(u)$ est non nul. Donc $\mathrm{Im}\,(N(u)) \neq 0$. Soit $x = N(u)(z) \in \mathrm{Im}\,(N(u))$ avec $x \neq 0$. On a alors :

$$(u - \alpha\,\mathrm{Id}_E)(x) = (u - \alpha\,\mathrm{Id}_E) \circ N(u)(z) = mu_u(u)(z) = 0,$$

donc α est une valeur propre de u.

□

4.3.2 Théorème de Cayley-Hamilton

Pour le reste de ce chapitre, les espaces E seront de dimension finie.

THÉORÈME 36 (Théorème de Cayley-Hamilton)
Soient E un $\mathbb{K}$-e.v. de dimension n et $u : E \to E$ un endomorphisme.
Alors, le polynôme caractéristique de u annule u :

$$\chi_u(u) = 0.$$

Preuve —

Méthode No 1 : Point de vue des endomorphismes. Soit $x \in E$ non-nul. Soit et p le plus petit entier tel que $u^p(x)$ appartienne au sous-espace cyclique $F = \mathrm{Vect}\{x, u(x), \ldots, u^{p-1}(x)\}$. Il existe donc une famille $\alpha_0, \ldots, \alpha_{p-1} \in \mathbb{K}$ tels que

$$u^p(x) = \sum_{k=0}^{p-1} -\alpha_k u^k(x).$$

Le sous-ev F est stable par u et la matrice de l'endomorphisme u_F dans la base $(x, u(x), \ldots, u^{p-1}(x))$ est :

$$A = \begin{pmatrix} 0 & 0 & \cdots & 0 & 0 & -\alpha_0 \\ 1 & 0 & \cdots & 0 & 0 & -\alpha_1 \\ 0 & 1 & \ddots & \vdots & \vdots & \vdots \\ \vdots & & \ddots & 0 & 0 & -\alpha_{p-3} \\ 0 & 0 & \cdots & 1 & 0 & -\alpha_{p-2} \\ 0 & 0 & \cdots & 0 & 1 & -\alpha_{p-1} \end{pmatrix}.$$

Cette matrice est la matrice compagnon du polynôme $P(X) = X^p + \sum_{k=0}^{p-1} \alpha_k X^k$.
On a vu d'une part que $\chi_{u_F}(X) = \chi_A(X) = P(X)$, et d'autre part que $\mu_{u_F} = \mu_A(X) = P(X)$. Comme on a $\mu_{u_F}(u_F) = 0$, on obtient donc que $\chi_{u_F}(u_F) = 0$.
Comme F est stable ar u, χ_{u_F} divise χ_u. Il existe $Q \in \mathbb{K}[X]$ tel que $\chi_u = Q\chi_{u_F}$. On en déduit :

$$\chi_u(u)(x) = Q(u) \circ \chi_{u_F}(u)(x) = Q(u)\left(\chi_{u_F}(u_F)(x)\right) = Q(u)(0) = 0.$$

Donc $\chi_u(u)(x) = 0$ pour tout $x \in E$. On en conclut que $\chi_u(u) = 0$.

Méthode No 2 : Point de vue matriciel. [2] Notons A la matrice de u dans une base quelconque d'un espace vectoriel de dimension n. Il suffit de démontrer que $\chi_A(A) = 0$, puisque $\chi_u = \chi_A$ et $\chi_u(u) = 0 \Leftrightarrow \chi_A(A) = 0$.

2. Cette preuve est valable dans $\mathbb{C}$ ou $\mathbb{R}$.

Notons $\operatorname{com}(\lambda I_n - A) = B(\lambda)$. On a alors $B(\lambda)(\lambda I_n - A) = \det(\lambda I_n - A)I_n = \chi_A(\lambda)I_n$. Comme $B(\lambda)$ est une comatrice, ses coefficients sont des polynômes en les coefficients de $\lambda I_n - A$ de degré au plus $n-1$. Cette matrice est donc de la forme : $B(\lambda) = \lambda^{n-1}B_{n-1} + \lambda^{n-2}B_{n-2} + \ldots + \lambda B_1 + B_0$, où les matrices $B_0, \ldots, B_{n-1}$ sont des matrices dont les coefficients sont des polynômes en les coefficients de A. Cela donne donc :

$$\begin{aligned} B(\lambda)(\lambda I_n - A) &= (\lambda^{n-1}B_{n-1} + \lambda^{n-2}B_{n-2} + \ldots + \lambda B_1 + B_0)(\lambda I_n - A) \\ &= \lambda^n B_{n-1} + \lambda^{n-1}(B_{n-2} - B_{n-1}A) + \lambda^{n-2}(B_{n-3} - B_{n-2}A) + \ldots \\ &\quad +\lambda^2(B_1 - B_2A) + \lambda(B_0 - B_1A) - B_0A. \end{aligned}$$

$\chi_A(X)$ est un polynôme unitaire de degré n. Notons $\chi_A(\lambda) = \lambda^n + a_{n-1}\lambda^{n-1} + \ldots + a_1\lambda + a_0$. On a donc $\chi_A(\lambda)I_n = \lambda^n I_n + a_{n-1}\lambda^{n-1}I_n + \ldots + a_1\lambda + a_0 I_n$. Comme $B(\lambda)(\lambda I_n - A) = \chi_A(\lambda)I_n$, on obtient :

$$\begin{aligned} B_{n-1} &= I_n \\ B_{n-2} - B_{n-1}A &= a_{n-1}I_n \\ B_{n-3} - B_{n-2}A &= a_{n-2}I_n \\ &\ldots \\ B_0 - B_1A &= a_1 I_n \\ -B_0A &= a_0 I_n. \end{aligned}$$

Ainsi, $\chi_A(A)$ vaut :

$$\begin{aligned} \chi_A(A) &= A^n + a_{n-1}A^{n-1} + \ldots + a_1A + a_0I_n \\ &= B_{n-1}A^n + (B_{n-2} - B_{n-1}A)A^{n-1} + (B_{n-3} - B_{n-2}A)A^{n-2} + \ldots + (B_0 - B_1A)A - B_0A \\ &= B_{n-1}A^n + B_{n-2}A^{n-1} - B_{n-1}A^n + B_{n-3}A^{n-2} - B_{n-2}A^{n-1} + \ldots + B_0A - B_1A^2 - B_0A = 0. \end{aligned}$$

□

Corollaire 37
Soient E un e.v. de dimension n et $u \in \mathcal{L}(E)$.
Alors $\mu_u(X)$ divise $\chi_u(X)$.
Le polynôme minimal de u divise son polynôme caractéristique.

Corollaire 38
Soit u est un endomorphisme d'un $\mathbb{R}$-e.v. de dimension finie.
Alors il existe une droite ou un plan vectoriel qui est stable par u.

Preuve — Si u possède une valeur propre λ, il existe $x \in E$ non nul tel que $u(x) = \lambda x$. $\operatorname{Vect} x$ est alors une droite stable par u.

Si u ne possède pas de valeur propre réelle, alors son polynôme caractéristique se factorise sur $\mathbb{R}$ sous la forme d'un produit de irréductibles de degré 2 :

$$\chi_u(X) = \prod_{k=1}^{t}(X^2 + p_kX + q_k).$$

Le théorème de Cayley-Hamilton montre que l'on a :

$$\prod_{k=1}^{t}(u^2 + p_ku + q_k \operatorname{Id}_E) = \chi_u(u) = 0.$$

Il existe donc un k tel que $u^2 + p_ku + q_k \operatorname{Id}_E$ ne soit pas inversible, donc non-injectif. Soit $x \in \operatorname{Ker}\left(u^2 + p_ku + q_k \operatorname{Id}_E\right)$ non-nul. On a donc

$$u^2(x) = -q_kx - p_ku(x),$$

c'est-à-dire que $\operatorname{Vect}\{x, u(x)\}$ est stable par u et de dimension 1 ou 2. (ce sous-ev est de dimension 2 car la famille $(x, u(x))$ est libre)

□

4.3.3 Indice d'un endomorphisme & Endomorphisme nilpotent

Dans ce paragraphe, nous introduisons un outil général puis nous l'appliquons au cas des endomorphismes nilpotents. Ces deux points serons utiles pour la suite de notre étude.

Proposition-Définition 39
Soient E un $\mathbb{K}$-e.v. de dimension finie, et $f \in \mathcal{L}(E)$.
Alors, il existe un unique entier naturel $r \in \mathbb{N}$ tel que :

$$\{0\} = \operatorname{Ker}\left(f^0 = \operatorname{Id}_E\right) \subsetneq \operatorname{Ker}(f) \subsetneq \cdots \subsetneq \operatorname{Ker}(f^r) = \operatorname{Ker}\left(f^{r+1}\right) = \cdots = \operatorname{Ker}(f^q) = \cdots$$

Cet entier est appelé **l'indice** de l'endomorphisme f.
On a aussi :

$$r(f) = \min\left\{k \in \mathbb{N}, \operatorname{Ker}\left(f^k\right) = \operatorname{Ker}\left(f^{k+1}\right)\right\} = \min\left\{k \in \mathbb{N}, \operatorname{Ker}\left(f^k\right) = \operatorname{Ker}\left(f^l\right)\ \forall l \geqslant k\right\}.$$

Démonstration. La suite des noyaux des f^k est croissante :

$$\operatorname{Ker}\left(f^k\right) \subset \operatorname{Ker}\left(f^{k+1}\right),\ \forall k \in \mathbb{N}$$

Ainsi, la suite $(\dim \operatorname{Ker}\left(f^k\right))_{k\geq 0}$ est une suite croissante d'entiers, majorée par $\dim(E)$. Il existe donc un rang $r \geq 0$ tel que $\dim \operatorname{Ker}(f^r) = \dim \operatorname{Ker}\left(f^{r+1}\right)$, c'est-à-dire $\operatorname{Ker}(f^r) = \operatorname{Ker}\left(f^{r+1}\right)$.
Montrons par récurrence sur $k \geq 1$ que l'on a : $\operatorname{Ker}\left(f^{r^k}\right) = \operatorname{Ker}(f^r)$, $\forall k \geq 1$.

- On a $\operatorname{Ker}(f^r) = \operatorname{Ker}\left(f^{r+1}\right)$, donc cela est vrai pour $k = 1$.
- Soit $k \geq 2$. Supposons cela vrai pour $k-1$. Soit $x \in \operatorname{Ker}\left(f^{r+k}\right)$. On a donc

$$0 = f^{r+k}(x) = f^{r+k-1}(f(x)).$$

  Ainsi, on a $f(x) \in \operatorname{Ker}\left(f^{r+k-1}\right) = \operatorname{Ker}(f^r)$, donc $0 = f^r(f(x)) = f^{r+1}(x)$.
  On a donc $x \in \operatorname{Ker}\left(f^{r+1}\right) = \operatorname{Ker}(f^r)$, ce qui donne $\operatorname{Ker}\left(f^{r+k}\right) \subset \operatorname{Ker}(f^r)$. L'inclusion réciproque étant toujours vraie, on en déduit que $\operatorname{Ker}(f^r) = \operatorname{Ker}\left(f^{r+k}\right)$.

Ainsi, en prenant r le plus petit entier positif tel que $\operatorname{Ker}(f^r) = \operatorname{Ker}\left(f^{r+1}\right)$, on a :

$$\{0\} = \operatorname{Ker}\left(f^0\right) = \mathrm{Id}_E \subsetneq \operatorname{Ker}(f) \subsetneq \cdots \subsetneq \operatorname{Ker}(f^r),\ \text{ et }\ \operatorname{Ker}(f^q) = \operatorname{Ker}(f^r),\forall q \geq r.$$

Cela montre aussi que cet entier r est unique. □

EXEMPLES 40

1. *Soit f un automorphisme de E. Alors son indice vaut 0 car $\operatorname{Ker}(f) = \operatorname{Ker}(Id_E) = \{0\}$.*
2. *Soit f une pojection de E différente de Id_E. Alors son indice vaut 1 car $\operatorname{Ker}(f) \neq \{0\}$ et car $f^2 = f$.*
3. *L'endomorphisme $f : P \in \mathbb{K}_n[X] \mapsto P' \in \mathbb{K}_n[X]$ est d'indice $n+1$.*

PROPOSITION 41
Soient E un $\mathbb{K}$-e.v. de dimension n et $f \in \mathcal{L}(E)$ un endomorphisme nilpotent.
Alors, on a : $\chi_f(X) = X^n$.

Preuve — Soit $q \in \mathbb{N}^*$ tel que $f^q = 0$. Nous donnons deux méthodes de démonstration de ce résultat.

Méthode No 1 : Passage aux complexes. Cette preuve est valable si $\mathbb{K} = \mathbb{C}$ ou si $\mathbb{K}$ est un sous-corps de $\mathbb{C}$. Soit B une base de E et $A = \mathrm{Mat}_B(f)$.
Comme on a $A^q = 0$, pour toute valeur propre $\lambda \in \mathbb{K}$ de A, on a $\lambda^q = 0$, donc A n'admet qu'une seule valeur propre, 0. Comme $\mathbb{K}$ est inclus dans $\mathbb{C}$, on peut voir A comme une matrice de $\mathscr{M}_n(\mathbb{C})$. Donc $\chi_A(X) \in \mathbb{C}[X]$. Ce polynôme est donc scindé sur $\mathbb{C}$, et ses racines sont des valeurs propres de A. On obtient donc :

$$\chi_f(X) = \chi_A(X) = X^n.$$

Méthode No 2 : Récurrence sur la dimension de l'espace. Cette preuve est vraie peu importe le corps $\mathbb{K}$.

- Pour $n = 1$: Tous les endomorphismes sont de la forme $f = \lambda Id_E$. La seule homothétie qui est nilpotente est l'endomorphisme nul. Dans ce cas, on a $\chi_f(X) = X$.
- Soit $n \geq 2$. Supposons le résultat vrai au rang $n-1$ et montrons-le au rang n.
  On a $f^q = 0$ donc f n'est pas inversible. Ainsi, on a $\operatorname{Ker}(f) \neq \{0\}$ et on peut considérer un vecteur non nul $e_1 \in \operatorname{Ker}(f)$. Complétons la famille (e_1) en une base $\mathcal{B}$ de E. Alors la matrice de f dans cette base est de la forme :

$$\mathrm{Mat}_{\mathcal{B}}(f) = \begin{pmatrix} 0 & \# & \cdots & \# \\ 0 & & M & \end{pmatrix}.$$

  Cette matrice est triangulaire supérieure par blocs. On a donc :

$$0 = \mathrm{Mat}_{\mathcal{B}}(f^q) = \begin{pmatrix} 0^q & \# & \cdots & \# \\ 0 & & M^q & \end{pmatrix}.$$

  On en déduit que $M^q = 0$, et donc que M est une matrice nilpotente. Comme $M \in \mathscr{M}_{n-1}(\mathbb{K})$, l'hypothèse de récurrence donne $\chi_M(X) = X^{n-1}$.
  On a alors $\chi_f(X) = \chi_{\mathrm{Mat}_{\mathcal{B}}(f)}(X) = (X-0)\chi_M(X) = X^n$.

□

REMARQUE 42 — *Cette proposition permet de caractériser les endomorphismes nilpotents : f est nilpotentesi et seulement si son polynôme caractéristique vaut $X^{\dim(E)}$.*

En guise de synthèse, on obtient la proposition suivante qui s'avérer riche pour la suite de notre étude.

PROPOSITION 43
Soient E un $\mathbb{K}$-e.v. de dimension finie, et $f \in \mathcal{L}(E)$ un endomorphisme nilpotent.
Alors, les nombres entiers suivants sont égaux :

- L'indice de nilpotence de f ;
- L'indice de f en tant qu'endomorphisme sur un ev de dimension finie ;
- Le degré du polynôme minimal de f.

Preuve — Notons r l'indice de nilpotence de f, r' son indice et r'' le degré de son polynôme minimal. On a alors :
- Pour tout $k \geqslant r$, $\text{Ker}\left(f^k\right) = E = \text{Ker}\left(f^{k+1}\right)$ et donc $r' \leqslant r$.
- Pour tout $k \geqslant r'$, $\text{Ker}\left(f^k\right) = \text{Ker}\left(f^{r'}\right)$. Or $f^r = 0$. Donc $\text{Ker}\left(f^r\right) = E$ et donc $\text{Ker}\left(f^{r'}\right) = E$ ou encore $f^{r'} = 0$. Cela donne $r \leqslant r'$.
- Par définition du polynôme minimal, on a $f^{r''} = 0$, donc $r \leqslant r''$.
- Comme $f^r = 0$, le polynôme X^r est annulateur de f et $\mu_f | X^r$, ce qui donne $r'' \leqslant r$.

On obtient ainsi l'égalité entre r, r' et r''. □

4.3.4 Lemme des noyaux

Théorème 44 (Lemme des noyaux)
Soient E un $\mathbb{K}$-e.v. et $u : E \to E$ un endomorphisme. Soient $P_1, ..., P_r \mathbb{K}[X]$ des polynômes premiers entre eux deux à deux, et soit $P = P_1 \dots P_r$. Alors, on a la décomposition en somme directe :

$$\text{Ker}\left(P(u)\right) = \bigoplus_{k=1}^{r} \text{Ker}\left(P_k(u)\right).$$

De plus, la projection p_k de $\text{Ker}\left(P(u)\right)$ sur $\text{Ker}\left(P_k(u)\right)$ parallèlement à $\bigoplus_{i \neq k} \text{Ker}\left(P_i(u)\right)$ est de la forme $U_k(u)_{\text{Ker}(P(u))}$ pour un polynôme $U_k \in \mathbb{K}[X]$.

Preuve — Notons $F = \text{Ker}\left(P(u)\right)$. Nous allons démontrer le résultat par récurrence sur $r \geq 2$.
- Supposons tout d'abord $r = 2$. Comme P_1 et P_2 sont premiers entre eux, le théorème de Bézout donne un couple $(A_1, A_2) \in \mathbb{K}[X]^2$ tel que :
$$A_1 P_1 + A_2 P_2 = 1.$$
On note $p_1 = (A_1 P_1)(u)_F$ et $p_2 = (A_2 P_2)(u)_F$. On a alors :
$$(A_1 P_1)(u)_F + (A_2 P_2)(u)_F = p_1 + p_2 = \text{Id}_F. \qquad (B)$$
Soit $x \in \text{Ker}\left(P_1(u)\right)$. On a $P(u)(x) = \left(P_2(u) \circ P_1(u)\right)(x) = 0$, donc $x \in F$, et $\text{Ker}\left(P_1(u)\right) \subset F$. On obtient de même que $\text{Ker}\left(P_2(u)\right) \subset F$. Ainsi, on a :
$$\left(\text{Ker}\left(P_1(u)\right) + \text{Ker}\left(P_2(u)\right)\right) \subset F.$$
Soit $x \in E$. La relation (B) nous permet d'écrire :
$$x = x_1 + x_2 \quad \text{avec} \quad x_1 = (A_1 P_1)(u)(x) = p_1(x) \quad \text{et} \quad x_2 = (A_2 P_2)(u)(x) = p_2(x). \qquad (P)$$
On a alors : $P_2(u)(x_1) = (P_2 A_1 P_1)(u)(x) = (A_1)(u) \circ (P)(u)(x) = 0$, donc $x_1 \in \text{Ker}\left(P_2(u)\right)$. De même, on a $P_1(u)(x_2) = 0$, donc $x_2 \in \text{Ker}\left(P_1(u)\right)$. Donc, on a $\text{Ker}\left(P_1(u)\right) + \text{Ker}\left(P_2(u)\right) = F$.
- Soit maintenant $x \in \text{Ker}\left(P_1(u)\right) \cap \text{Ker}\left(P_2(u)\right)$. La relation (B) donne :
$$x = A_2(u) \circ P_2(u)(x) + A_1(u) \circ P_1(u)(x) = 0 + 0 = 0.$$
On obtient finalement la somme directe : $\text{Ker}\left(P_1(u)\right) \oplus \text{Ker}\left(P_2(u)\right) = F$.
- On remarque que l'on a de plus :
  - $\forall x \in F$, $(p_1 \circ p_2)(x) = ((A_1 P_1)(u) \circ (A_2 P_2)(u))(x) = (A_1 A_2)(u) \circ P(u)(x) = 0$, donc $p_1 \circ p_2 = 0$;
  - $p_1 = p_1 \circ \text{Id}_F = p_1 \circ (p_1 + p_2) = p_1 \circ p_1$ et de même $p_2 = p_2 \circ p_2$.

  Ainsi, les applications linéaires p_1 et p_2 sont des projections sur F.
- Il reste à montrer que $\text{Im}\left(p_1\right) = \text{Ker}\left(P_2(u)\right)$.
  - Soit $y \in \text{Im}\left(p_1\right)$. On a $\exists x \in F$ tel que $y = p_1(x)$, ce qui donne $P_2(u)(y) = (P_2(u) \circ p_1)(x) = (A_1)(u) \circ P(u)(x) = 0$. Donc $y \in \text{Ker}\left(P_2(u)\right)$, d'où $\text{Im}\left(p_1\right) \subset \text{Ker}\left(P_2(u)\right)$.
  - Réciproquement, soit $x \in \text{Ker}\left(P_2(u)\right)$. On a $x = p_1(x) + p_2(x)$, et $p_2(x) = (A_2 P_2)(u)(x) = A_2(u) \circ P_2(u)(x) = 0$. Donc, $x = p_1(x) \in \text{Im}\left(p_1\right)$.

  On montre de la même façon que $\text{Im}\left(p_2\right) = \text{Ker}\left(P_1(u)\right)$. Ainsi, les projections $p_1 = (A_1 P_1)(u)_F$ et $p_2 = (A_2 P_2)(u)_F$ sont les projections associées à la somme directe que l'on a obtenue.

Soit maintenant $r \geq 2$. Supposons le résultat vrai pour $r - 1$.
- Comme $P_1, \dots, P_r$ sont premiers entre eux deux à deux, P_1 et $R_1 = P_2 \cdots P_r$ sont premiers entre eux. On a ainsi :
$$F = \text{Ker}\left(P(u)\right) = \text{Ker}\left(P_1(u)\right) \oplus \text{Ker}\left(R_1(u)\right) \qquad (R)$$
d'après la preuve du cas $r = 2$.
- L'hypothèse de récurrence nous fournit alors la décomposition :
$$\text{Ker}\left(R_1(u)\right) = \bigoplus_{k=2}^{r} \text{Ker}\left(P_k(u)\right).$$
Cela montre la somme directe désirée.

- De plus, la projection $p_1 : F \to F$ sur $\mathrm{Ker}\,(P_1(u))$ parallèlement à $\bigoplus_{k=2}^{r} \mathrm{Ker}\,(P_k(u))$ est la projection associée à la somme directe (R). D'après la preuve du cas $r = 2$, il existe donc un polynôme $U_1 \in \mathbb{K}[X]$ tel que $p_1 = U_1(u)_F$. Par permutation, c'est aussi le cas de des projections $p_2, \ldots, p_r$.

□

Le lemme des noyaux est un résultat très important en algèbre linéaire.

Corollaire 45
Soient E un $\mathbb{K}$-e.v. et $u \in \mathcal{L}(E)$. Soit P un polynôme annulateur non-nul de u. Soient $P_1 \ldots, P_r$ des polynômes premiers entre eux deux à deux tels que $P = P_1 \cdots P_r$.
Alors, on a :

$$E = \bigoplus_{k=1}^{r} \mathrm{Ker}\,(P_k(u))\,.$$

Exemple 46 — *Pour* $A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix} \in \mathscr{M}_3(\mathbb{R})$, *on a vu que* $(X-1)(X-2)^2$ *est un polynôme annulateur de* A. *Ainsi, on a* $\mathbb{R}^3 = \mathrm{Ker}\,(A - I_3) \oplus \mathrm{Ker}\,\big((A-2I_3)^2\big)$.

Remarque 47 — *Pour* u *un endomorphisme et* $\lambda_1, \ldots, \lambda_r$ *des valeurs propres de* u *distinctes, les polynômes* $X - \lambda_1, \ldots, X - \lambda_r$ *sont premiers entre eux deux à deux.*
On retrouve avec le lemme des noyaux le fait que $\mathrm{Ker}\,(u - \lambda_1 Id_E)\,, \ldots, \mathrm{Ker}\,(u - \lambda_r Id_E)$ *sont en somme directe.*

Remarque 48 — *Soient* u *un endomorphisme et* $P_1, \ldots, P_r$ *des polynômes premiers entre eux deux à deux. Notons* $Q_k = \prod_{i \neq k} P_i$. *Alors les polynômes* $Q_1, \ldots, Q_r$ *sont premiers entre eux dans leur ensemble. D'après le théorème de Bezout généralisé, il existe donc* $A_1, \ldots, A_r \in \mathbb{K}[X]$ *tels que* $Q_1 A_1 + \ldots + Q_r A_r = 1$, *de sorte que*

$$\mathrm{Id}_E = Q_1(u) \circ A_1(u) + \ldots + Q_r(u) \circ A_r(u).$$

En reprenant les éléments de la preuve du lemme des noyaux, pour $F = \mathrm{Ker}\,((P_1 \ldots P_r)(u))$ *la projection* p_k *dans* F *sur* $\mathrm{Ker}\,(P_k(u))$ *parallèlement à* $\bigoplus_{i \neq k}^{r} \mathrm{Ker}\,(P_i(u))$ *est alors égale à :* $p_k = Q_k A_k(u)_F$.

4.3.5 Synthèse

Faisons un bilan de certains résultats obtenus dans ce chapitre et dans le chapitre précédent.

Soit E un $\mathbb{K}$-e.v. E de dimension n. Soit $u \in \mathcal{L}(E)$ un endomorphisme.

§ *1. Polynôme caractéristique* — Le premier élément d'information que nous avons pour u est son *polynôme caractéristique* $\chi_u(X)$. On a :

$$\chi_u(X) = \prod_{k=1}^{r} (X - \lambda_k)^{m_u(\lambda_k)} \pi_{l=1}^{s} P_i^{\alpha_i},$$

avec

$$\mathrm{Spec}\,(u) = \{\lambda_1, \ldots, \lambda_r\}\,, \text{ et } P_1 \ldots, P_s \text{ irréductibles et de degré } > 1.$$

Pour pour toute valeur propre λ_k, l'entier $m_u(\lambda_k)$ est appelé la *multiplicité* de λ_k pour u.
Comme χ_u est unitaire et de degré n, on a $m_u(\lambda_1) + \ldots + m_u(\lambda_r) \leq n$.

§ *2. Polynôme minimal* — D'après le théorème de Caley-Hamilton, le polynôme minimal de u, μ_u, divise son polynôme caractéristique. Ainsi, on a $\deg(\mu_u) \leq \deg(\chi_u) \leq n$.
De plus, les valeurs propres de u sont les racines de μ_u dans $\mathbb{K}$.
On obtient donc :

$$\mu_u(X) = \prod_{k=1}^{r} (X - \lambda_k)^{r_u(\lambda_k)} \pi_{l=1}^{s} P_i^{\beta_i},$$

avec $1 \leqslant r_u(\lambda_k) \leqslant m_u(\lambda_k)$ pour tout $1 \leq k \leq r$ et $\beta_i \leq \alpha_i$ pour tout $1 \leq i \leq s$.
L'entier $r_u(\lambda_k)$ est appelé **l'indice de la valeur propre** λ_k. (c'est la multiplicité de $X - \lambda_k$ dans $\mu_u(X)$) λ_k pour u.

§ *3. Sous-espace propres* — Nous avons vu que les sous-espaces propres de u,

$$E_{\lambda_k}(u) = \mathrm{Ker}\,(u - \lambda_k\,\mathrm{Id}_E)\,,\ 1 \leq k \leq r,$$

sont en somme directe. Leur dimension vérifie :

$$1 \leqslant \dim\,(E_{\lambda_k}(u)) \leqslant m_u(\lambda_k),\ \forall 1 \leq k \leq r.$$

§ *4. Sous-espaces caractéristiques* — Pour toute valeur propre λ_k de u, nous pouvons écrire, comme dans le paragraphe 4.3.3, la suite croissante des noyaux de $(u - \lambda_k \operatorname{Id}_E)^k$:

$$\{0\} \subsetneq \operatorname{Ker}(u - \lambda_k \operatorname{Id}_E) = E_{\lambda_k}(u) \subsetneq \cdots \subsetneq \operatorname{Ker}\left((u - \lambda_k \operatorname{Id}_E)^{r'_u(\lambda_k)}\right)$$
$$= \operatorname{Ker}\left((u - \lambda_k \operatorname{Id}_E)^{r'_u(\lambda_k)+1}\right) = \cdots = \operatorname{Ker}\left((u - \lambda_k \operatorname{Id}_E)^{q}\right) = \cdots$$

où $r'_u(\lambda_k)$ est donc l'indice de l'endomorphisme $u - \lambda_k \operatorname{Id}_E$.
On définit alors le **sous-espace caractéristique** \根子空间\ de u comme :

$$F_{\lambda_k}(u) = \operatorname{Ker}\left((u - \lambda_k \operatorname{Id}_E)^{r'_u(\lambda_k)}\right).$$

On obtient alors la proposition suivante :

Proposition 49
Soient E un $\mathbb{K}$-e.v. de dimension n, et $u : E \to E$ un endomorphisme avec $\operatorname{Spec}(u) = \{\lambda_1, \ldots, \lambda_r\}$.
Avec les notations précédentes, on a alors :

$$r'_u(\lambda_k) = r_u(\lambda_k),\ \forall 1 \leq k \leq r.$$

L'indice $r'_u(\lambda_k)$ de l'endomorphisme $u - \lambda_k \operatorname{Id}_E$ est égal à l'indice $r_u(\lambda_k)$ de la valeur propre λ_k (la multiplicité de λ_k dans $\mu_u(X)$).

Preuve — Soit $1 \leq k \leq r$. Le polynôme minimal de u s'écrit alors :

$$\mu_u(X) = (X - \lambda_k)^{r_u(\lambda_k)} Q(X)$$

où $Q(X)$ est un polynôme premier avec $(X - \lambda_k)$.

Soit $q \in \mathbb{N}$. Supposons $q \leq r_u(\lambda_k)$. Comme μ_u divise $(X - \lambda_k)^q Q(X)$, le lemme des noyaux donne :

$$E = \operatorname{Ker}(P_q(u)) = \operatorname{Ker}((u - \lambda_k \operatorname{Id}_E)^q) \oplus \operatorname{Ker}(Q(u)).$$

On obtient donc que :

$$\operatorname{Ker}((u - \lambda_k \operatorname{Id}_E)^q) = \operatorname{Ker}\left((u - \lambda_k \operatorname{Id}_E)^{r_u(\lambda_k)}\right),\ \forall q \leq r_u(\lambda_k).$$

Supposons maintenant que $q \leq r_u(\lambda_k)$. Alors μ_u ne divise pas $(X - \lambda_k)^q Q(X)$, donc $(u - \lambda_k Id_E)^q Q(u) \neq 0$. Le lemme des noyaux donne donc :

$$E \neq \operatorname{Ker}(P_q(u)) = \operatorname{Ker}((u - \lambda_k \operatorname{Id}_E)^q) \oplus \operatorname{Ker}(Q(u)).$$

On en déduit que

$$\operatorname{Ker}((u - \lambda_k \operatorname{Id}_E)^q) \neq \operatorname{Ker}\left((u - \lambda_k \operatorname{Id}_E)^{r_u(\lambda_k)}\right).$$

Donc, les propriétés de l'indice de $u - \lambda_k Id_E$, $r'_u(\lambda_k)$, nous disent que cet indice vaut exactement $r_u(\lambda_k)$. Ce qui conclut. □

Remarque 50 — *On déduit donc de cette propriété la façon dont les noyaux itérés de $(u - \lambda_k \operatorname{Id}_E)$ grossissent :*

$$\{0\} \subsetneq \operatorname{Ker}(u - \lambda_k \operatorname{Id}_E) = E_{\lambda_k}(u) \subsetneq \cdots \subsetneq \operatorname{Ker}\left((u - \lambda_k \operatorname{Id}_E)^{r_u(\lambda_k)}\right) = \cdots = \operatorname{Ker}\left((u - \lambda_k \operatorname{Id}_E)^{m_u(\lambda_k)}\right) = \cdots$$

Proposition 51
Soient E un $\mathbb{K}$-e.v. de dimension n, et $u : E \to E$ un endomorphisme avec $\operatorname{Spec}(u) = \{\lambda_1, \ldots, \lambda_r\}$.
Avec les notations précédentes, on a alors :

$$\dim\left(\operatorname{Ker}\left((u - \lambda_k \operatorname{Id}_E)^{r_u(\lambda_k)}\right)\right) = m_u(\lambda_k),\ \forall 1 \leq k \leq r$$

Preuve — On a $\chi_u(X) = \prod_{k=1}^r (X - \lambda_k)^{m_u(\lambda_k)} Q(X)$, avec Q un polynôme qui n'a pas de racines.
On pose $F_k = \operatorname{Ker}\left((u - \lambda_k \operatorname{Id}_E)^{m_u(\lambda_k)}\right)$ le sous-espace caractéristique associé à la valeur propre λ_k.
On pose $F = \operatorname{Ker}(Q(u))$.
Le théorème de Cayley-Hamilton et le lemme des noyaux donnent la somme directe :

$$E = \operatorname{Ker}(\chi_u(u)) = \oplus_{k=1}^r F_k \oplus F.$$

Comme $F_k = \operatorname{Ker}\left((u - \lambda_k Id_E)^{m_u(\lambda_k)}\right)$, pour l'endomorphisme induit u_{F_k}, on a ainsi :

$$(u_{F_k} - \lambda_k Id_{F_k})^{m_u(\lambda_k)} = (u - \lambda_k Id_E)^{m_u(\lambda_k)}_{F_k} = 0.$$

Donc $(X - \lambda_k)^{m_u(\lambda_k)}$ est un polynôme annulateur de u_{F_k}.
Cela veut dire que $(u_{F_k} - \lambda_k Id_{F_k})$ est un endomorphisme nilpotent.
Le polynôme caractéristique d'un endomorphisme nilpotent est de la forme X^m.
On a donc $\chi_{u_{F_k} - \lambda_k Id_{F_k}} = X^{\dim(F_k)}$.
Ainsi :

$$\chi_{u_{F_k}} = \det(X Id_{F_k} - u_{F_k}) = \det((X - \lambda_k) Id_{F_k} - (u_{F_k} - \lambda_k Id_{F_k}) = (X - \lambda_k)^{\dim(F_k)}.$$

Sur $F = \text{Ker}\,(Q(u))$, on a de même $Q(u_F) = Q(u)_F = 0$.
Cela veut dire que $Q(X)$ est un polynôme annulateur de u_F. Le polynôme minimal de u_F est donc un diviseur de $Q(X)$.
Comme Q n'a pas de racines, μ_{u_F} n'a pas de racines non plus.
D'après un résultat précédent, on a donc $\text{Spec}\,(u_F) = \emptyset$.
Le polynôme caractéristique de u_F n'a pas de racines non plus, donc $\chi_{u_F}(X) = R(X)$, avec R polynôme sans racines.
Vu que

$$E = \text{Ker}\,(\chi_u(u)) = \oplus_{k=1}^r F_k \oplus F,$$

on a la factorisation suivante : $\chi_u(X) = \Pi_{k=1}^r \chi_{u_{F_k}} . \chi_{u_F}$.
En conclusion, nous avons deux factorisations du polynôme caractéristique de u :

$$\prod_{k=1}^r (X - \lambda_k)^{m_u(\lambda_k)} \times Q(X) = \chi_u(X) = \prod_{k=1}^r (X - \lambda_k)^{\dim\left(F_{\lambda_k}(u)\right)} \times R(X).$$

Cela donne donc :

$$\chi_{u_F}(X) = R(X) = Q(X), \text{ et } m_u(\lambda_k) = \dim\left(F_{\lambda_k}(u)\right), \forall 1 \leq k \leq r.$$

□

4.4 Diagnalisation, trigonalisation

Les notions développées dans cette section concernent les endomorphismes. Elles s'appliquent aux matrices carrées de $\mathscr{M}_n(\mathbb{K})$ en considérant les endomorphismes associés sur $\mathbb{K}^n$.

4.4.1 Diagonalisation

Soient E un ev de dimension finie et $u \in \mathcal{L}(E)$ un endomorphisme. On a vu que u est diagonalisable si et seulement si $\chi_u(X)$ est scindé, et si la dimension du sous-espace propre $E_{\lambda_k}(u)$ vaut $m(\lambda_k)$ pour tout $k = 1, \dots, r$. Cela donne le résultat :

Théorème 52
Soient E un ev de dimension finie et $u \in \mathcal{L}(E)$ un endomorphisme.
Alors, u est diagonalisable si et seulement si son polynôme minimal est scindé à racines simples.

Preuve — On pose $\text{Spec}\,(u) = \{\lambda_1, \dots, \lambda_r\}$.

- Si u est diagonalisable, on a $E = \bigoplus_{i=1}^r \text{Ker}\,(u - \lambda_i Id_E)$.
  Pour $P(X) = (X - \lambda_1) \dots (X - \lambda_r)$, le lemme des noyaux donne :
  $$E = \text{Ker}\,(P(u)).$$
  C'est-à-dire, $P(u) = 0$. Donc le polynôme minimal de u, χ_u, divise P. Comme P est un polynôme scindé à racines simples, μ_u est alors un polynôme scindé à racines simples.
  Comme $\lambda_1, \dots, \lambda_r$ sont les racines de μ_u, on a même $\mu_u(X) = P(X)$.
- Réciproquement, on suppose que μ_u est scindé à racines simples. Comme les racines de μ_u sont les valeurs propres de u, on a donc $\mu_u(X) = (X - \lambda_1) \dots (X - \lambda_r)$. Le lemme des noyaux nous donne alors :
  $$E = \text{Ker}\,(\mu_u(u)) = \bigoplus_{k=1}^s \text{Ker}\,((u - \alpha_k \,\text{Id}_E)).$$
  Donc E est la somme directe des sous-espaces propres de E. Cela veut dire que u est diagonalisable.

□

Corollaire 53
Soient E un ev de dimension finie et $u \in \mathcal{L}(E)$ un endomorphisme.
Alors, u est diagonalisable si et seulement s'il est annulé par un polynôme scindé à racines simples.

Nous avons obtenu une nouvelle condition qui caractérise la diagonalisabilité.

Corollaire 54
Soient E un ev de dimension finie et $u \in \mathcal{L}(E)$ un endomorphisme.
Soit F un sous-ev de E stable par u. Si u est diagonalisable, alors u_F est diagonalisable.

Preuve — On sait que u_F est annulé par μ_u, ce qui conclut. □

Proposition 55 (Rappel)
Soient E un ev de dimension finie et $u \in \mathcal{L}(E)$ un endomorphisme diagonalisable avec $\text{Spec}\,(u) = \{\lambda_1, \dots, \lambda_r\}$.

Soit $p_k : E \to E$ la projection sur $E_{\lambda_k}(u)$ parallèlement à $\bigoplus_{i\neq k} E_{\lambda_i}(u)$. Alors, on a :

$$u = \lambda_1 p_1 + \ldots + \lambda_r p_r$$
$$u^m = \lambda_1^m p_1 + \ldots + \lambda_r^m p_r, \forall m \geq 0$$

Preuve — Voir Chapitre Diagonalisation. □

PROPOSITION 56
Soient E un ev de dimension finie et $u \in \mathcal{L}(E)$ un endomorphisme diagonalisable avec $\mathrm{Spec}\,(u) = \{\lambda_1, \ldots, \lambda_r\}$. On a $u = \lambda_1 p_1 + \ldots + \lambda_r p_r$, où $p_k : E \to E$ est la projection sur $E_{\lambda_k}(u)$ parallèlement à $\bigoplus_{i\neq k} E_{\lambda_i}(u)$.
Soient $L_1, \ldots, L_r \in \mathbb{K}[X]$ les polynômes d'interpolation de Lagrange associés à l'ensemble $\{\lambda_1, \ldots, \lambda_r\}$. Alors, on a :

$$p_k = L_k(u) = \Pi_{i\neq k} \frac{u - \lambda_i}{\lambda_k - \lambda_i}, \forall 1 \leq k \leq r.$$

Preuve — Les propriétés des polynômes d'interpolation de Lagrange nous disent que $L_k(\lambda_i) = 0$ si $i \neq k$ et $L_k(\lambda_k) = 1$.
Or, on a :

$$L_k(u_{E_{\lambda_i}(u)}) = L_k(\lambda_i Id_{E_{\lambda_i}(u)}) = L_k(\lambda_i) Id_{E_{\lambda_i}(u)}.$$

Donc $L_k(u_{E_{\lambda_i}(u)}) = 0$ si $i \neq k$ et $L_k(u_{E_{\lambda_k}(u)}) = Id_{E_{\lambda_i}(u)}$.
L'endomorphisme $L_k(u)$ est donc égal à p_k sur $E_{\lambda_i}(u)$, pour tout $1 \leq i \leq r$. Comme ces sous-espaces sont en somme directe dans E, on en déduit que $L_k(u) = p_k$, ce qui conclut. □

COROLLAIRE 57
Soit $u \in \mathcal{L}(E)$ un endomorphisme diagonalisable, avec $\mathrm{Spec}\,(u) = \{\lambda_1, \ldots, \lambda_r\}$, dont on connaît $A = \mathrm{Mat}_{\mathcal{B}}(u)$. On peut alors calculer u^m de deux façons :

- *Méthode 1 :* On détermine une base $\mathcal{B}'$ de vecteurs propres de u.
  On détermine P la matrice de passage de $\mathcal{B}$ vers $\mathcal{B}'$. On calcule P^{-1} l'inverse de P.
  Alors, on a $\mathrm{Mat}_{\mathcal{B}}(u) = A = PDP^{-1}$, où D est une matrice diagonale. Cela donne :
  $$\mathrm{Mat}_{\mathcal{B}}(u^m) = \mathrm{Mat}_{\mathcal{B}}(u)^m = A^m = PD^mP^{-1}, \forall m \geq 0.$$
  On détermine donc u^m en calculant D^m (facile), puis en effectuant le produit PD^mP^{-1}.
- *Méthode 2 :* On détermine les polynômes interpolateurs de Lagrange $L_1, \ldots, L_r$ associés à $\{\lambda_1, \ldots, \lambda_r\}$ en développant l'expression $L_k(X) = \Pi_{i\neq k} \frac{X-\lambda_i}{\lambda_k-\lambda_i}$.
  Alors, on a
  $$u^m = \lambda_1^m L_1(u) + \ldots + \lambda_r^m L_r(u), \forall m \geq 0.$$

EXEMPLE 58 — *On pose* $A = \begin{pmatrix} -1 & 0 & 1 \\ 2 & 0 & 2 \\ 2 & 0 & 0 \end{pmatrix}$. *Montrer que A est diagonalisable, et A^m pour tout $m \geq 0$.*
On commence par caluler le polynôme caractéristique de A. Un calcul de déterminant donne : $\chi_A(X) = X(X-1)(X+2)$.
Le polynôme caractéristique de A est scindé à racines simples, donc A est diagonalisable. De plus, on obtient que $\mathrm{Spec}\,(A) = \{-2, 0, 1\}$.
Soient L_1, L_2, L_3 les polynômes d'interpolation de Lagrange associés à $\{-2, 0, 1\}$. On a :

$$L_1(X) = \frac{(X-0)(X-1)}{(-2-0)(-2-1)} = \frac{X^2 - X}{6}$$
$$L_2(X) = \frac{(X+2)(X-1)}{(0+2)(0-1)} = \frac{X^2 + X - 2}{-2}$$
$$L_3(X) = \frac{(X+2)(X-0)}{(1+2)(1-0)} = \frac{X^2 + 2X}{3}$$

Pour tout $m \geq 0$, on a donc :

$$A^m = (-2)^m L_1(A) + 0^m L_2(A) + 1^m L_3(A).$$

Pour $m \geq 1$, cela donne $A^m = (-2)^m \frac{1}{6}(A^2 - A) + 0 + \frac{1}{3}(A^2 + 2A)$.
On calcule :

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

On obtient donc :

$$A^m = \frac{(-2)^m}{6}\begin{pmatrix} 4 & 0 & -2 \\ 0 & 0 & 0 \\ -4 & 0 & 2 \end{pmatrix} + \frac{1}{3}\begin{pmatrix} 1 & 0 & 1 \\ 6 & 0 & 6 \\ 2 & 0 & 2 \end{pmatrix}.$$

Définition 59
Soient E un ev de dimension finie et $(u_i)_{i\in I}$ une famille d'endomorphismes de E.
On dit que les endomorphismes $(u_i)_{i\in I}$ sont **simultanément diagonalisables** s'il existe une base $\mathcal{B}$ de E dans laquelle toutes les matrices $\mathrm{Mat}_{\mathcal{B}}(u_i)$ sont diagonales.
Une telle base B est appelée une **base de diagonalisation simultanée**.

Proposition 60
Soient E un ev de dimension finie et $(u_i)_{i\in I}$ une famille d'endomorphismes de E.
Les endomorphismes $(u_i)_{i\in I}$ sont simultanément diagonalisables si et seulement si, u_i est diagonalisable et si u_i commute avec u_j, pour tous $i, j \in I$.

Preuve —

"⇒" Pour $\mathcal{B}$ une base de diagonalisation simultanée, $\mathrm{Mat}_{\mathcal{B}}(u_i)$ est diagonale. Or, deux matrices diagonales commutent.

"⇐" Montrons la réciproque par récurrence forte sur la dimension n de E.
Si $n \geq 1$, tous les endomorphismes sont de la forme $u_i = \lambda_i Id_E$, et le résultat est vrai.
Soit $n \geq 1$. Supposons le résultat vrai sur tout espace de dimension $\leq n$.
Soit E de dimension $n+1$. On distingue deux cas.

- Si tous les u_i sont des homothéties, $u_i = \lambda_i Id_E$, n'importe quelle base convient.
- Sinon, il existe $i_0 \in I$ tel que u_{i_0} n'est pas une homothétie. Comme u_{i_0} est diagonalisable, il possède donc au moins deux valeurs propres. On peut alors écrire :
$$E = F \oplus G$$
où F est un sous-espace vectoriel propre de u_{i_0} et G la somme directe des autres sous-espaces vectoriels propres de u_{i_0}. Ces sous-espaces vectoriels sont de la forme $\mathrm{Ker}\,(P(u_{i_0}))$ pour des polynômes P.
Ils sont donc stables pour chaque endomorphisme u_i car u_i commute avec $P(u_{i_0})$. On note $(u'_i)_i$ et $(u''_i)_i$ les familles d'endomorphismes induits sur F et G. Ce sont des familles d'endomorphismes diagonalisables qui commutent deux à deux.
Comme les dimensions de F et G sont strictement inférieures à celle de E, par hypothèse de récurrence il existe $\mathcal{B}'$ et $\mathcal{B}''$ des bases de F et G de diagonalisation simultanée. En posant $\mathcal{B} = \mathcal{B}' \cup \mathcal{B}''$, on obtient une base de diagonalisation simultanée des (u_i).

□

4.4.2 Trigonalisation

Définition 61
Soient E un ev de dimension finie et $u \in \mathcal{L}(E)$ un endomorphisme.
On dit que u est **trigonalisable** s'il existe une base $\mathcal{B}$ de E dans laquelle la matrice $\mathrm{Mat}_{\mathcal{B}}(u)$ est triangulaire supérieure.
La base $\mathcal{B}$ est appelée une **base de trigonalisation** de u. [3].
Soit $A \in \mathscr{M}_n(\mathbb{K})$. On dit que A est **trigonalisable** s'il existe une matrice inversible P telle que $P^{-1}AP$ soit triangulaire supérieure.

Remarque 62 — *Si la matrice de u dans la base $(e_1, \ldots, e_n)$ est triangulaire supérieure, alors la matrice de u dans la base inversée $(e_n, \ldots, e_1)$ est une matrice triangulaire inférieure.*

Théorème 63
Soient E un ev de dimension finie et $u \in \mathcal{L}(E)$ un endomorphisme.
Alors, u est trigonalisable si et seulement s'il existe $P \in \mathbb{K}[X]$ scindé tel que $P(u) = 0$.

Preuve —

"⇒" Si u est trigonalisable, on a une base $\mathcal{B}$ de E telle que $A = \mathrm{Mat}_{\mathcal{B}}(u)$ soit triangulaire supérieure. On a vu que le polynôme caractéristique d'une matrice triangulaire supérieure est : $\chi_A(X) = (X - a_{1,1}) \ldots (X - a_{n,n})$. C'est un polynôme scindé.
Le théorème de Cayley-Hamilton nous dit alors que : $\chi_A(u) = \chi_u(u) = 0$.

"⇐" Nous démontrons la réciproque par récurrence sur n, la dimension de E.
Si $n = 1$ on a $u = \lambda Id_E$ et le résultat est vrai.
Soit $n > 1$, supposons que tout endomorphisme d'un espace vectoriel de dimension $n-1$ qui annule un polynôme scindé est trigonalisable.
Si u est annulé par un polynôme scindé $P(X) = \prod_{k=1}^{m}(X - \beta_k)$, le lemme des noyaux nous dit qu'au moins une des racines de P est une valeur propre de u. Quitte à réordonner, prenons β_1 comme valeur propre de u. Il existe donc $e_1 \in E$ non-nul tel

3. Nous avons vu que c'est le cas si tous les sous-espaces vectoriels $\mathrm{Vect}\,(e_1, \ldots, e_i)$ sont stables par u. Voir le corollaire 64.

que $u(e_1) = \beta_1 e_1$.
On complète la famille (e_1) en une base $\mathcal{B} = (e_1, \dots, e_n)$ de E. On a alors :

$$B = \mathrm{Mat}_{\mathcal{B}}(u) = \begin{pmatrix} \beta_1 & L \\ 0 & B' \end{pmatrix}.$$

Cette matrice est triangulaire supérieure par blocs. On a donc :

$$B^k = \begin{pmatrix} \beta_1^k & * \\ 0 & B'^k \end{pmatrix} \Rightarrow 0 = P(B) = \begin{pmatrix} P(\beta_1) & * \\ 0 & P(B') \end{pmatrix}.$$

On en déduit que $P(B') = 0$, c'est-à-dire que la matrice $B' \in \mathscr{M}_{n-1}(\mathbb{K})$ est annulée par un polynôme scindé. Par hypothèse de récurrence, B' est alors semblable à une matrice triangulaire supérieure. Ainsi, B est trigonalisable et donc u l'est aussi.

□

Corollaire 64
Soient E un $\mathbb{K}$-ev de dimension finie et $u \in \mathcal{L}(E)$ un endomorphisme.
On a les équivalences :

1. u est trigonalisable ;
2. χ_u est scindé ;
3. μ_u est scindé ;
4. Spec (u) est non-vide, et pour Spec $(u) = \{\lambda_1, \dots, \lambda_r\}$, on a

$$m_u(\lambda_1) + \ldots + m_u(\lambda_r) = \dim(E).$$

Preuve — L'équivalence entre les trois premiers points est immédiate.
Le polynôme caractéristique de u, χ_u, est scindé si et seulement s'il est de la forme :

$$\chi_u(X) = \Pi_{i=1}^{r}(X - \lambda_i)^{m_u(\lambda_i)}.$$

Or, les racines de χ_u sont exactement les valeurs propres de u, et on a $\sum_{i=1}^{r} m_u(\lambda_i) = \deg(\chi_u) = n$. □

Remarque 65 — *Tout endomorphisme sur un $\mathbb{C}$-espace vectoriel de dimension finie est donc trigonalisable. De même, toute matrice complexe A est trigonalisable et s'écrit :*

$$A = PTP^{-1}$$

avec P inversible et T triangulaire supérieure.
Ce n'est pas le cas sur $\mathbb{R}$ où un endomorphisme (une matrice) est trigonalisable si, et seulement si, sn polynôme caractéristique est scindé. Par exemple une matrice de rotation rotation $R(\theta)$ ($\theta \not\equiv 0 \mathrm{mod}\ \pi$) de $\mathbb{R}^2$ n'est pas trigonalisable.

Remarque 66 — *Si u est un endomorphisme trigonalisable, alors* Spec (u) *est non-vide.*
Pour tout F sous-ev stable par u, si u est trigonalisable alors u_F est trigonalisable, donc u_F a un spectre non-vide.

Exemple 67 — *On vérifie avec la méthode du Pivot que le polynôme caractéristique de la matrice*

$$A = \begin{pmatrix} 0 & 3 & 3 \\ -1 & 8 & 6 \\ 2 & -14 & -10 \end{pmatrix} \in \mathscr{M}_3(\mathbb{R})$$

vaut $X(X+1)^2$ et que les vecteurs :

$$X_1 = \begin{pmatrix} 2 \\ 1 \\ -1 \end{pmatrix} \quad \text{et} \quad X_2 = \begin{pmatrix} 3 \\ 3 \\ -4 \end{pmatrix}$$

sont des vecteurs propres associés aux valeurs 0 et -1. Prenons pour X_3 le premier vecteur de la base canonique. Alors (X_1, X_2, X_3) est une base de trigonalisation de A, et la matrice de $u_A : X \mapsto AX$ dans cette base est :

$$P^{-1}AP = \begin{pmatrix} 0 & 0 & 2 \\ 0 & -1 & -1 \\ 0 & 0 & -1 \end{pmatrix}, \quad \text{avec } P = \begin{pmatrix} 2 & 3 & 1 \\ 1 & 3 & 0 \\ -1 & -4 & 0 \end{pmatrix}.$$

On cherche un vecteur X_4 tel que $AX_4 = X_2 - X_4$. Après résolution, on trouve $X_4 = {}^t(0, -1, 2)$. La famille (X_1, X_2, X_4) est alors une autre base de trigonalisation, et la matrice de u_A dans cette base est :

$$J = Q^{-1}AQ = \begin{pmatrix} 0 & 0 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & -1 \end{pmatrix}, \quad \text{avec } Q = \begin{pmatrix} 2 & 3 & 0 \\ 1 & 3 & -1 \\ -1 & -4 & 2 \end{pmatrix}.$$

Cette fois la matrice J est diagonale par blocs, avec des blocs diagonaux qui sont triangulaires supérieurs.

Proposition 68 (Trigonalisation simultanée)
Soient E un $\mathbb{K}$-ev de dimension finie et $u_i \in \mathcal{L}(E)$, $i \in I$, des endomorphisme.
Si les endomorphismes u_i sont trigonalisables et commutent deux à deux, alors il existe une base $\mathcal{B}$ de E dans laquelle les matrices de tous les u_i sont triangulaires supérieures.

Preuve — Nous démontrons cela par récurrence sur n la dimension de E.
Si $n = 1$, on a $u_i = \lambda_i Id_E$, et cela est vrai.
Soit $n \geq 2$. Supposons le résultat vrai au rang $n-1$.
Soit $i \in I$. Comme u_i est trigonalisable, son spectre est non-vide. Soit $\lambda \in \mathrm{Spec}\,(u_i)$. Le sous-espace $\mathrm{Ker}\,(u_i - \lambda Id_E)$ est stable par tous les u_j, $j \in I$. Donc, pour $x \in \mathrm{Ker}\,(u_i - \lambda Id_E)$ non-nul, on a $u_j(x) = \lambda x$, pour tout $j \in J$. On poursuit la preuve de la même façon que celle de 60. □

Proposition 69
Soient E un $\mathbb{K}$-ev de dimension finie et $u \in \mathcal{L}(E)$ un endomorphisme.
Alors, u est nilpotent si, et seulement si, il existe une base $\mathcal{B}$ de E dans laquelle la matrice $\mathrm{Mat}_{\mathcal{B}}(u)$ est triangulaire supérieure stricte. (de diagonale nulle)

Preuve —

“⇒” Si u est nilpotent on a $\chi_u(X) = X^n$, donc u trigonalisable. Pour $\mathcal{B}$ une base de trigonalisation de u, $A = \mathrm{Mat}_{\mathcal{B}}(u)$ est triangulaire supérieure. Or, les coefficients diagonaux de A sont les racines de χ_u, comptées avec multiplicité. Donc, la matrice A est triangulaire supérieure stricte.

“⇐” S'il existe une base $\mathcal{B}$ dans laquelle la matrice $A = \mathrm{Mat}_{\mathcal{B}}(u)$ est triangulaire supérieure stricte, on obtient $\chi_u(X) = \chi_A(X) = X^n$, donc u^n d'après le théorème de Cayley-Hamilton, donc u est nilpotent.

□

4.4.3 Décomposition dite de Dunford

Nous allons étudier une nouvelle forme de réduction plus poussée que la trigonalisation.

Proposition 70 (Décomposition de Jordan-Chevalley (dite de Dunford))
Soient E un $\mathbb{K}$-ev de dimension finie et $u \in \mathcal{L}(E)$ un endomorphisme.
Si χ_u est scindé, alors il existe un unique couple d'endomorphismes $(d, n) \in (\mathcal{L}(E))^2$ avec d diagonalisable et n nilpotent tels que

$$u = d + n \quad \text{et} \quad d \circ n = n \circ d.$$

De plus, ces endomorphismes sont des polynômes en u.

Preuve — On écrit $\chi_u(X) = \prod_{k=1}^{r}(X - \lambda_k)^{m_k}$ et $F_k = \mathrm{Ker}\,((u - \lambda_k\,\mathrm{Id}_E)^{m_k})$ le sous-espace caractéristique de λ_k.

Existence. D'après le lemme des noyaux on a $E = \bigoplus_{k=1}^{r} F_k$. D'après la remarque 48, la projection p_k sur F_k parallèlement à $\bigoplus_{i \neq k}^{r} F_i$ est un polynôme en u. On le note $p_k = P_k(u)$. On pose alors :

$$\boldsymbol{d = \sum_{k=1}^{r} \lambda_k p_k} \qquad \text{et} \qquad \boldsymbol{n = u - d = \sum_{k=1}^{r} (u - \lambda_k\,\mathrm{Id}_E) \circ p_k.}$$

On rappelle que $p_i \circ p_j = \delta_{i,j} p_i$, pour tous $1 \leq i, j \leq r$. Ainsi, d laisse stable chaque F_k, et l'on a $d_{F_k} = \lambda_k Id_{F_k}$. Donc d est diagonalisable.
De même, on montre par récurrence sur $q \geq 0$ que

$$n^q = \sum_{k=1}^{r} (u - \lambda_k\,\mathrm{Id}_E)^q \circ p_k,\ \forall\forall q \in \mathbb{N}^*.$$

Pour $q = \max\{m_k, k = 1, 2, \ldots, r\}$, on a $(u - \lambda_k\,\mathrm{Id}_E)^q_{F_k} = 0$ car $F_k = \mathrm{Ker}\,((u - \lambda_k\,\mathrm{Id}_E)^{m_k})$.
Cela donne $(u - \lambda_k\,\mathrm{Id}_E)^q \circ p_k = 0$. Ainsi, on a $n^q = 0$, donc n est nilpotent.
Enfin, d et n sont des polynômes en u, donc ils commutent.

Unicité. Soit (d', n') un autre couple vérifiant les conditions. Comme d' et n' commutent, ils commuent avec $u = d' + n'$, donc ils commuent aussi avec d et n qui sont des polynômes en u. Ainsi, d et d' sont simultanément diagonalisables dans une même base, ce qui implique que $d - d'$ est diagonalisable.
D'autre part, n et n' commutent et sont trigonalisables, donc ils sont trigonalisables dans une même base $\mathcal{B}$. Alors, $\mathrm{Mat}_{\mathcal{B}}(n' - n) = \mathrm{Mat}_{\mathcal{B}}(n') - \mathrm{Mat}_{\mathcal{B}}(n)$ est une matrice triangulaire supérieure de diagonale nulle. On en déduit que $\chi_{n'-n}(X) = X^{\dim(E)}$ et donc que $n' - n$ est nilpotent.
Comme on a $d - d' = n' - n$, $d - d'$ est donc nilpotent. Le seul endomorphisme diagonalisable et nilpotent étant l'endomorphisme nul, on obtien $d - d' = n' - n = 0$, ce qui donne l'unicité.

□

Remarques 71

1. *Ainsi, tout endomorphisme trigonalisable u possède une décomposition de Dunford.*
2. *En fait, comme u laisse ses sous-espaces caractéristiques stables, on a défini d et n sur chaque sous-espace caractéristique F_k par :*

$$d_{F_k} = \lambda_k Id_{F_k}, \text{ et } n_{F_k} = u_{F_k} - d_{F_k} = u_{F_k} - \lambda_k Id_{F_k}.$$

   On remarque alors que d_{F_k} et n_{F_k} sont des polynôme en u_{F_k}, et que ces endomorphismes commutent. Le lemme des noyaux nous dit que les espaces F_k sont en somme directe, avec $E = \bigoplus_{k=1}^{r} F_k$, ce qui permet de faire remonter le comportement sur chaque F_k à un comportement sur E.
3. *Soit $\mathcal{B}$ une base adaptée à la somme directe des sous-espaces caractéristiques $F_1, \dots, F_r$, on a alors :*

$$\mathrm{Mat}_{\mathcal{B}}(u) = \begin{array}{c} \\ m_1 \\ m_2 \\ \vdots \\ m_r \end{array} \begin{array}{c} \begin{array}{cccc} \overset{m_1}{\longleftrightarrow} & \overset{m_2}{\longleftrightarrow} & \cdots & \overset{m_r}{\longleftrightarrow} \end{array} \\ \begin{pmatrix} A_1 & & & (0) \\ & A_2 & & \\ & & \ddots & \\ (0) & & & A_r \end{pmatrix} \end{array}.$$

   Cette matrice est diagonale par blocs. Chaque bloc A_k, de taille $m_k \times m_k$, est la matrice de la restrictition u_k de u au sous-espace caractéristique F_k.
   Comme $\chi_{u_k} = (X - \lambda_k)^{m_k}$, pour $N_k = A_k - \lambda_k I_{m_k}$ on en déduit que $\chi_{N_k} = X^{m_k}$, et donc que N_k est nilpotente.
   On a donc $A_k = \lambda_k I_{m_k} + N_k$ avec $N_k = A_k - \lambda_k I_{m_k}$ nilpotente.
   Cela donne alors :

$$\mathrm{Mat}_{\mathcal{B}}(d) = \begin{pmatrix} \lambda_1 I_{m_1} & & & \\ & \lambda_2 I_{m_2} & & \\ & & \ddots & \\ & & & \lambda_s I_{m_r} \end{pmatrix} \quad \text{et} \quad \mathrm{Mat}_{\mathcal{B}}(n) = \begin{pmatrix} N_1 & & & \\ & N_2 & & \\ & & \ddots & \\ & & & N_r \end{pmatrix}.$$

   Et ces deux matrices commutent : $DN = ND$.
4. *Comme $\chi_{u_k}(X) = (X - \lambda)^{m_k}$, on peut aller un peu plus loin en prenant une base $\mathcal{B}_k$ de F_k qui trigonalise u_k. Dans cette base, la matrice de d_k sera encore $\lambda_k I_{m_k}$, et la matrice de n_k sera une matrice triangulaire supérieure stricte.*
   Donc, dès que l'endomorphisme u est trigonalisable, on peut trouver une base $\mathcal{B}$ de E dans laquelle la matrice $\mathrm{Mat}_{\mathcal{B}}(u)$ est diagonale par blocs, avec des blocs triangulaires supérieurs.
5. *L'écriture $u = d + n$ donnée par la décomposition de Dunford s'utilise pour calculer u^p :*

$$u^p = (d+n)^p = \sum_{k=0}^{p} \binom{p}{k} d^k \circ n^{p-k}.$$

   Dans l'expression ci-dessus, on peut retirer les termes de la somme pour lesquels $p - k$ est plus grand que l'indice de nilpotence de n.

4.4.4 Réduction de Jordan

Nous allons donner une réduction encore plus poussée que la précédente.

Définition 72
Soient $n \geq 1$ et $\lambda \in \mathbb{K}$. On appelle **bloc de Jordan** \若尔当块\ de taille n pour λ la matrice de la forme

$$\mathcal{J}_n(\lambda) = \begin{pmatrix} \lambda & 1 & 0 & & (0) \\ & \lambda & 1 & \ddots & \\ & & \ddots & \ddots & 0 \\ & (0) & & \lambda & 1 \\ & & & & \lambda \end{pmatrix} \in \mathscr{M}_n(\mathbb{K}).$$

Remarque 73 —

1. *$\mathcal{J}_n(\lambda)$ est nilpotente si, et seulement si, $\lambda = 0$.*
2. *On a $\mathrm{Spec}\,(\mathcal{J}_n(\lambda)) = \{\lambda\}$ et $\chi_{\mathcal{J}_n(\lambda)}(X) = \mu_{\mathcal{J}_n(\lambda)} = (X-\lambda)^n$.*
3. *Pour $1 \leq m \leqslant n$, on a $\mathcal{J}_n(0)^m = \begin{pmatrix} O_{n-m,m} & I_{n-m} \\ O_m & O_{m,n-m} \end{pmatrix}$.*

Proposition 74

Soient E un $\mathbb{K}$-ev de dimension finie et $u \in \mathcal{L}(E)$ un endomorphisme. Soient $\lambda \in \mathrm{Spec}\,(u)$ et $r(\lambda)$ l'indice de $(u - \lambda Id_E)$.

Pour $m \leq r(\lambda)$, soit $\alpha \in \mathrm{Ker}\,((u-\lambda\,\mathrm{Id}_E)^m) \setminus \mathrm{Ker}\,\big((u-\lambda\,\mathrm{Id}_E)^{m-1}\big)$.

On pose $\alpha_i = (u-\lambda\,\mathrm{Id}_E)^{m-i}(\alpha)$, $\forall i \in [\![1,m]\!]$. Alors, on a :

1. La famille $\mathcal{B} = (\alpha_1,\dots,\alpha_m) = ((u-\lambda\,\mathrm{Id}_E)^{m-1}(\alpha),\dots,(u-\lambda\,\mathrm{Id}_E)(\alpha),\alpha)$ est libre.
2. Le sous-espace $F = \mathrm{Vect}(\mathcal{B})$ est stable par u, et on a $\mathrm{Mat}_{\mathcal{B}}(u_F) = \mathcal{J}_m(\lambda)$.

Preuve —

1. D'après la définition, on on a $\alpha = \alpha_m$ et $\alpha_i \neq 0_E$, $\forall i \in [\![1,m]\!]$.
   Soient $\lambda_1,\dots,\lambda_m \in \mathbb{K}$ tels que $\sum_{i=1}^m \lambda_i\alpha_i = 0_E$. Supposons que les λ_i ne sont pas tous nuls. On pose $m' = \max\{i, \lambda_i \neq 0\} \leqslant m$.
   Cela donne :
   $$\sum_{i=1}^{m'} \lambda_i\alpha_i = 0_E, \text{ avec } \lambda_{m'} \neq 0.$$
   On a donc $0_E = (u-\lambda\,\mathrm{Id}_E)^{m'-1}(\sum_{i=1}^{m'}\lambda_i\alpha_i) = \lambda_{m'}\alpha_1$, d'où $\lambda_{m'} = 0$, ce qui est une contradiction.
   On en déduit que la famille $\mathcal{B} = (\alpha_1,\dots,\alpha_m)$ est libre.
2. La famille $\mathcal{B}$ est donc une base de F.
   Or, on a $(u-\lambda\,\mathrm{Id}_E)(\alpha_i) = \alpha_{i-1}$, $\forall i \in [\![2,m]\!]$, donc $u(\alpha_i) = \lambda\alpha_i + \alpha_{i-1} \in F$.
   Enfin, on a $(u-\lambda\,\mathrm{Id}_E)(\alpha_1) = (u-\lambda\,\mathrm{Id}_E)^m(\alpha) = 0$, donc $u(\alpha_1) = \lambda\alpha_1 \in F$. Ce qui donne le résultat.

□

Remarque 75 — *Dans la proposition précédente, on a en fait $F = S_{u-\lambda\,\mathrm{Id}_E}(\alpha)$, le sous-espace cyclique engendré par α pour $u-\lambda\,\mathrm{Id}_E$. Cet espace est de dimension m, et une base est engendrée par α :*

$$(u-\lambda\,\mathrm{Id}_E)^m\alpha = 0_E \leftarrow (u-\lambda\,\mathrm{Id}_E)^{m-1}\alpha = \alpha_1 \leftarrow (u-\lambda\,\mathrm{Id}_E)^{m-2}\alpha = \alpha_2 \leftarrow \cdots \leftarrow (u-\lambda\,\mathrm{Id}_E)\alpha = \alpha_{m-1} \leftarrow \alpha = \alpha_m.$$

Afin de montrer que toute matrice trigonalisable possède une décomposition en blocs de Jordan, nous allons montrer cela dans le cas nilpotent. Nous aurons besoin du lemme suivant, issu du cours sur les applications linéaires.

Lemme 76

Soient E, F des $\mathbb{K}$-ev de dimension finie et $u : E \to F$ une application linéaire. Soit $\mathcal{B}_0 = (x_1, x_2, \dots, x_r)$ une base de $\mathrm{Ker}\,(u)$ et $\mathcal{B}_1 = (y_1, y_2, \dots, y_s)$ une famille de vecteurs de E. On note $\mathcal{B} = \mathcal{B}_0 \cup \mathcal{B}_1$. Alors, on a :

1. $\mathcal{B}$ est libre si, et seulement si, $u(\mathcal{B}_1)$ est libre. ◀
2. $\mathcal{B}$ est une base de E si, et seulement si, $u(\mathcal{B}_1)$ une base de $\mathrm{Im}\,(u)$. ◀

Proposition 77

Soient E un $\mathbb{K}$-ev de dimension finie et $u \in \mathcal{L}(E)$ un endomorphisme nilpotent.

Alors, il existe des sous-espaces cycliques de u, $S_u(x_1),\dots,S_u(x_r)$, tels que :

$$E = \bigoplus_{i=1}^{r} S_u(x_i).$$

Preuve — Nous démontrons le résultat par récurrence sur $n \geq 1$ la dimension de E.
Si $\dim\,(E) = 1$, on a $u = 0$ et le résultat est vrai.
Soit $n \geq 2$. Supposons le résultat vrai au rang $n-1$, et démontrons-le pour E de dimension n.

Comme $u(E)$ est stable par u, on considère u' l'endomorphisme induit par u sur $u(E)$. Comme u est nilpotent, u n'est pas inversible, donc $\dim\,(u(E)) < n$. De plus, u' est encore nilpotent. Par hypothèse de récurrence, $u(E)$ s'écrit comme une somme directe de sous-espaces cycliques pour u'.
Ainsi, il existe une base de $u(E)$ de la forme :

$$B_{U(E)} = \left\{ \begin{array}{cccc} y_1 & y_2 & \dots & y_t \\ u(y_1) & u(y_2) & \dots & u(y_t) \\ \vdots & \vdots & & \vdots \\ u^{k_1-1}(y_1) & u^{k_2-1}(y_2) & \dots & u^{k_t-1}(y_t), \\ (u^{k_1}(y_1) = 0_E) & (u^{k_2}(y_2) = 0_E) & \dots & (u^{k_t}(y_t) = 0_E) \end{array} \right\}.$$

Les vecteurs de la i-ième colonne forment une base du sous-espace cyclique engendrée par y_i, $S_{u'}(y_i)$, de dimension $k_i > 0$.

Comme $y_1, \dots, y_t \in u(E)$, soient $x_1, \dots, x_t$ tels que $y_i = u(x_i), i \in [\![1, t]\!]$. On obtient les vecteurs ci-dessous :

$$\begin{array}{c|ccccc}
u^{-1}(\mathcal{B}_{U(E)}) & x_1 & x_2 & \dots & x_t & \\
 & u(x_1) & u(x_2) & \dots & u(x_t) & \\
 & \vdots & \vdots & & \vdots & \\
 & u^{k_1-1}(x_1) & u^{k_2-1}(x_2) & \dots & u^{k_t-1}(x_t) & \\
\hline
\text{base de } \operatorname{Ker}(u) \ : & u^{k_1}(x_1) & u^{k_2}(x_2) & \dots & u^{k_t}(x_t), & x_{t+1}, \dots, x_s. \\
 & (u^{k_1+1}(x_1) = 0_E) & (u^{k_2+1}(x_2) = 0_E) & \dots & (u^{k_t+1}(x_t) = 0_E) & \text{(complétés)}
\end{array}$$

La famille $(u^{k_1}(x_1), u^{k_2}(x_2), \dots, u^{k_t}(x_t))$ est libre, car sous-famille d'une base de $u(E)$, et est contenue dans $\operatorname{Ker}(u)$. On peut donc la compléter en une base de $\operatorname{Ker}(u)$, en rajoutant des vecteurs que l'on notera $x_{t+1}, \dots, x_s$ à la dernière ligne du tableau ci-dessus. Les autres vecteurs dans ce tableau sont les images réciproques d'une base de $u(E)$. D'apres le lemme 76, ces vecteurs du tableau ci-dessus forment une base de E. Les vecteurs dans la i-ième colonne sont une base du sous-espace cyclique engendré par x_i, $S_u(x_i)$. Cela termine la récurrence. □

Remarque 78 — *Dans la preuve précédente, en ordonnant les vecteurs de la base $\mathcal{B}$ selon leurs sous-espaces cycliques respectifs $\mathcal{B} = (u^{k_1}(x_1), u^{k_1-1}(x_1), \dots, u(x_1), x_1, u^{k_2}(x_2), u^{k_2-1}(x_2), \dots, u(x_2), x_2, \dots, u^{k_t}(x_t), x_t, x_{t+1}, \dots, x_s)$, on obtient :*

$$\operatorname{Mat}_{\mathcal{B}}(u) \begin{pmatrix} \mathcal{J}_{k_1}(0) & & & \\ & \mathcal{J}_{k_2}(0) & & \\ & & \ddots & \\ & & & \mathcal{J}_{k_s}(0) \end{pmatrix}.$$

Théorème 79 (Décomposition de Jordan)
Soient E un $\mathbb{K}$-ev de dimension finie et $u \in \mathcal{L}(E)$ un endomorphisme.
Si χ_u est scindé, alors il existe une base $\mathcal{B}$ de E dans laquelle la matrice de u est diagonale par blocs avec des blocs de Jordan.
On dit que u a **une forme de Jordan** :

$$\operatorname{Mat}_{\mathcal{B}}(u) = \begin{pmatrix} \mathcal{J}_{k_1}(\lambda_1) & & & \\ & \mathcal{J}_{k_2}(\lambda_2) & & \\ & & \ddots & \\ & & & \mathcal{J}_{k_s}(\lambda_s) \end{pmatrix}.$$

Les λ_i sont les valeurs propres de u (qui peuvent apparaître plusieurs fois). [4]

Preuve — Supposons que $\chi_u(X) = \prod_{i=1}^{r} (X - \lambda_i)^{m_i}$ avec λ_i, $i \in [\![1, r]\!]$ deux à deux différents. On pose $F_i = \operatorname{Ker}((u - \lambda_i \operatorname{Id}_E)^{m_i})$ le sous-espace caractéristique associé à λ_i.
On pose $v_i = (u - \lambda_i \operatorname{Id}_E)_{F_i}$. Comme $\chi_{u_{F_i}}(X) = (X - \lambda_i)^{m_i}$, on a $\chi_{v_i}(X) = X^{m_i}$. Donc, v_i est nilpotent.
D'après la proposition 77, il existe une base $\mathcal{B}_i$ de F_i dans laquelle la matrice de v_i a une forme de Jordan. Comme on a $u_{F_i} = v_i + \lambda_i \operatorname{Id}_E$, on obtient :

$$\operatorname{Mat}_{\mathcal{B}_i}(u_{F_i}) = \begin{pmatrix} \mathcal{J}_{k_1}(\lambda_i) & & & \\ & \mathcal{J}_{k_2}(\lambda_i) & & \\ & & \ddots & \\ & & & \mathcal{J}_{k_t}(\lambda_i) \end{pmatrix}.$$

L'espace E est la somme directe des sous-ev F_i, $i \in [\![1, r]\!]$. En posant $\mathcal{B} = \mathcal{B}_1 \cup \ldots \cup \mathcal{B}_r$, on obtient donc une base de E dans laquelle la matrice de u est de la forme voulue. □

Remarque 80 — *Soit $A \in \mathscr{M}_n(\mathbb{K})$.*

- *Si A est diagonalisable, on a $A = PDP^{-1}$, et $A^m = PD^m P^{-1}$.*
  Les puissances de A se calculent facilement.
- *Si A est trigonalisable, on a $A = PTP^{-1}$, et $A^m = PT^m P^{-1}$.*
  Calculer les puissances de A revient à calculer les puissances d'une matrice triangulaire supérieure quelconque. Ce n'est pas si facile.
- *Si A est trigonalisable, la décomposition dite de Dunford donne $A = D + N$, et $A^m = (D + N)^m$.*
  Les puissances de A s'expriment facilement en fonction de D et de N.
  Pour les calculer, il faut cependant calculer les puissances de la matrice nilpotente N. Cela prend un peu de temps.
  Dans la décomposition $A = PTP^{-1}$, cela revient à séparer T en $T = D' + N'$, avec N' une matrice triangulaire supérieure stricte.
  Cette forme de matrice nilpotente est un peu plus facile à gérer, mais il faut quand même calculer $N'^2, N'^3, \dots, N'^{n-1}$ pour calculer les puissances de T.

4. Il y a en général plus d'un bloc de Jordan pour chaque valeur propre, plusieurs λ_i peuvent avoir la même valeur.

- *Si A est trigonalisable, la décomposition de Jordan donne $A = PJP^{-1}$ et $A^m = PJ^mP^{-1}$.*
  Comme les puissances d'un bloc de Jordan se calculent facilement, les puissances de la matrice de Jordan J se calculent facilement. On peut donc calculer facilement les puissances de A.

La réduction de Jordan est une réduction de Dunford améliorée (ainsi qu'une trigonalisation améliorée), où la partie nilpotente/triangulaire stricte a une expression simplifiée, ce qui permet de calculer des puissances ou de la manipuler plus facilement.
En retour, trouver une base dans laquelle on a une réduction de Jordan demande plus de temps que pour trouver une base de trigonalisation.

Exemples 81

1. *Dans l'exemple 67, la matrice J est la forme de Jordan de la matrice A.*
2. *On pose la matrice :*

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

   Le calcul du polynôme caractéristique donne $\chi_A(X) = (X-1)(X-2)(X+1)^2$, donc $\mathrm{Spec}\,(A)\,\{-1, 1, 2\}$.
   Les trois sous-espaces caractéristiques sont :

   - *la droite* $\mathrm{Ker}\,(A - I_4)$*, engendrée par* $X_1 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ -1 \end{pmatrix}$.
   - *la droite* $\mathrm{Ker}\,(A - 2I_4)$*, engendrée par* $X_2 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ -2 \end{pmatrix}$.
   - *le plan* $\mathrm{Ker}\,\big((A + I_4)^2\big)$.
     On détermine d'abord le sous-espace propre qu'il contient : $\mathrm{Ker}\,(A + I_4)$ *est la droite engendrée par* $X_3 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$.
     On cherche un antécédent de X_3 par $A + I_4$. On obtient : $X_4 = \begin{pmatrix} 0 \\ -1 \\ 1 \\ -1 \end{pmatrix}$.

   Donc, pour $P = \begin{pmatrix} X_1 & X_2 & X_3 & X_4 \end{pmatrix} \in \mathrm{GL}_4(\mathbb{R})$*, on a :*

$$P^{-1}AP = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & -1 \end{pmatrix}.$$

Exemple 82 — *On pose la matrice :*

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

Un calcul de déterminant donne : $\chi_A(X) = (X-1)(X-2)^5$.
Le sous-espace caractéristique $\mathrm{Ker}\,(A - I_6)$ *est engendré par la première colonne de A, que l'on note X_1.*
Notons $B = A - 2I$, on a :

$$B = \begin{pmatrix} -1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 2 & 3 \\ 0 & 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \quad B^2 = \begin{pmatrix} 1 & -1 & -1 & 0 & 3 & 3 \\ 0 & 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \quad B^3 = B^4 = \begin{pmatrix} -1 & 1 & 1 & 0 & -2 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Soit u l'endomorphisme canoniquement associé à A et $v = (u - 2\,\mathrm{Id}_E)_{F_2(u)}$.
On a alors $\mu_u(X) = \mu_A(X) = (X-1)(X-2)^3$ et $\dim\,(F_2(u)) = \dim\big(\mathrm{Ker}\,\big(B^3\big)\big) = 5$. *De plus, v est un endomorphisme nilpotent sur $F_2(u)$ qui est d'indice* 3.
On cherche des sous-espaces cycliques dont la somme directe donne $F_2(u)$.

- *On résout le système $BY = 0$ pour trouver une base de* $\mathrm{Ker}\,(v)$ *:*

$$\mathcal{B}_1 = \mathcal{S}_1 = \{Y_1, Y_2\} = \left\{ \begin{pmatrix} 1\\1\\0\\0\\0\\0 \end{pmatrix}, \begin{pmatrix} 1\\0\\1\\0\\0\\0 \end{pmatrix} \right\}.$$

- *On résout le système $B^2Y = 0$ pour compléter la famille $\mathcal{B}_1$ en une base de* $\mathrm{Ker}\left(v^2\right)$ *:*

$$\mathcal{B}_2 = \mathcal{S}_1 \cup \mathcal{S}_2,\ \ \mathcal{S}_2 = \{Y_3, Y_4\} = \left\{ \begin{pmatrix} 0\\0\\0\\1\\0\\0 \end{pmatrix}, \begin{pmatrix} 0\\0\\3\\0\\2\\-1 \end{pmatrix} \right\}.$$

- *On résout le système $B^2Y = 0$ pour compléter la famille $\mathcal{B}_1$ en une base de* $\mathrm{Ker}\left(v^2\right)$ *:*

$$\mathcal{B}_3 = \mathcal{B}_2 \cup \mathcal{S}_3,\ \ \mathcal{S}_3 = \{Y_5\} = \left\{ \begin{pmatrix} 0\\0\\0\\0\\1\\-2 \end{pmatrix} \right\}.$$

- *On regarde la famille*

$$\mathcal{T} = \{B^2Y_5, BY_3, BY_4, Y_1, Y_2\} = \left\{ \begin{pmatrix} -3\\-3\\0\\0\\0\\0 \end{pmatrix}, \begin{pmatrix} 1\\1\\0\\0\\0\\0 \end{pmatrix}, \begin{pmatrix} 4\\1\\3\\0\\0\\0 \end{pmatrix}, \begin{pmatrix} 1\\1\\0\\0\\0\\0 \end{pmatrix}, \begin{pmatrix} 1\\0\\1\\0\\0\\0 \end{pmatrix} \right\},$$

  qui est une famille de vecteurs de $\mathrm{Ker}\,(B)$*, et on cherche à en extraire une famille libre de rang maximal. Cette famille est de rang 2, et une sous-famille libre de rang maximal est :* $\mathcal{T}^* = \{B^2Y_5, BY_4\}$*.*
  Alors, $F_2(u)$ est la somme directe des sous-espaces cycliques engendrés par Y_5 et Y_4. Une base de $F_2(u)$ est donc :

$$\mathcal{B}_3' = \begin{pmatrix} 0 & \leftarrow u^2(Y_5) = B^2Y_5 & \leftarrow u(Y_5) = BY_5 & \leftarrow Y_5 \\ 0 & \leftarrow u(Y_4) = BY_4 & \leftarrow Y_4 & \end{pmatrix}.$$

Pour l'endomorphisme v, cela donne :

$$\mathrm{Mat}_{\mathcal{B}_3'}(v) = \begin{pmatrix} 0&1&0&0&0\\0&0&1&0&0\\0&0&0&0&0\\0&0&0&0&1\\0&0&0&0&0 \end{pmatrix} = \begin{pmatrix} \mathcal{J}_3(0) & \\ & \mathcal{J}_2(0) \end{pmatrix},$$

Donc, pour la matrice $P = \begin{pmatrix} X_1 & B^2Y_5 & BY_5 & Y_5 & BY_4 & Y_4 \end{pmatrix} \in \mathrm{GL}_6(\mathbb{R})$*, on a :*

$$P^{-1}AP = \begin{pmatrix} 1&0&0&0&0&0\\0&2&1&0&0&0\\0&0&2&1&0&0\\0&0&0&2&0&0\\0&0&0&0&2&1\\0&0&0&0&0&2 \end{pmatrix} = \begin{pmatrix} 1 & & \\ & \mathcal{J}_3(2) & \\ & & \mathcal{J}_2(2) \end{pmatrix}.$$

4.5 Applications

4.5.1 Applications aux équations différentielles

Soit l'équation différentielle linéaire à coefficients constants d'ordre p :

$$f^{(p)} + \alpha_{p-1} f^{(p-1)} + \cdots + \alpha_1 f' + \alpha_0 f = 0 \qquad (ED)$$

avec $(\alpha_0, \ldots, \alpha_{p-1}) \in \mathbb{C}^p$, d'inconnue $f \in \mathcal{C}^\infty(\mathbb{R}, \mathbb{C})$.

On appelle **polynôme caractéristique** de (ED) le polynôme :

$$\chi(X) = X^p + \sum_{k=0}^{p-1} \alpha_k X^k.$$

On note sa factorisation dans $\mathbb{C}[X]$:

$$\chi(X) = \prod_{k=1}^{r} (X - \lambda_k)^{p_k},$$

où les λ_k sont deux à deux distincts et les p_k sont non-nuls.
Soit D l'endomorphisme de dérivation sur $\mathcal{C}^\infty(\mathbb{R},\mathbb{C})$. Alors, l'ensemble $\mathcal{S}$ des solutions de (ED) est le noyau de l'endomorphisme $\chi(D)$.
Le lemme des noyaux nous fournit alors la décomposition en somme directe :

$$\mathcal{S} = \bigoplus_{k=1}^{r} \operatorname{Ker}\left((D-\lambda_k \operatorname{Id}_E)^{p_k}\right).$$

Résoudre l'équation (ED) se ramène donc à résoudre r équations différentielles linéaires plus simples.
Soient $1 \leq k \neq r$ et $f \in \mathcal{C}^\infty(\mathbb{R},\mathbb{C})$. On démontre par récurrence sur $p_k \geq 1$ la relation suivante :

$$(D-\lambda_k \operatorname{Id}_E)^{p_k}\left(f(x)e^{\lambda_k x}\right) = e^{\lambda_k x} D^{p_k}(f)(x).$$

On en déduit alors que :

$$\operatorname{Ker}\left((D-\lambda_k \operatorname{Id}_E)^{p_k}\right) = \left\{x \in \mathbb{R} \longmapsto P(x)e^{\lambda_k x},\ P \in \mathbb{C}_{p_k-1}[X]\right\}.$$

Ainsi, tout élément f de $\mathcal{S}$ s'écrit de façon unique comme :

$$x \in \mathbb{R} \longmapsto f(x) = \sum_{k=1}^{r} P_k(x)e^{\lambda_k x},\ P_1,\ldots,P_r \in \mathbb{C}[X],\ \deg(P_k) \leq p_k - 1.$$

L'espace vectoriel des solutions $\mathcal{S}$ est donc de dimension $p_1 + \ldots + p_k = p$.
On retrouve ce dernier point comme conséquence du théorème de Cauchy sur les équations linéaires à coefficients constants d'ordre 1 ou 2.

Exemple 83 — *Résoudre sur $\mathbb{R}$ l'équation différentielle :*

$$y^{(5)} - 13y^{(4)} + 67y^{(3)} - 171y" + 216y' = 108y.\ (E)$$

L'équation (E) est une équa. diff. linéaire à coefficients constants, d'ordre 5.
Pour $D : y \mapsto y'$ l'endomorphisme de dérivation sur $C^\infty(\mathbb{R},\mathbb{C})$, les solutions de (E) sont exactement les fonctions dans $Ker(P(D))$, avec $P(X) = X^5 - 13X^4 + 67X^3 - 171X^2 + 216X - 108$. (les fonctions y telles que $P(D)(y) = 0$)
On remarque que 2 est une racine de P. En calculant P', on remarque que $P'(2) = 0$, donc 2 est une racine double de P.
On obtient alors en factorisant : $P(X) = (X^2 - 4X + 4)(X^3 - 9X^2 + 27X - 27)$.
On reconnaît le polynôme $(X-3)^3$. On a donc $P(X) = (X-2)^2(X-3)^3$.
D'après le Lemme des noyaux, on a donc :

$$Ker(P(D)) = Ker((D-2Id)^2) \oplus Ker((D-3Id)^3).$$

D'après les résultats de la section, on a ainsi :

$$Ker(P(D)) = Vect(x \mapsto \exp(2x), x \mapsto x\exp(2x), x \mapsto \exp(3x), x \mapsto x\exp(3x), x \mapsto x^2\exp(3x)).$$

Une fonction y est solution de (E) si et seulement s'il existe $a_1,\ldots,a_5 \in \mathbb{C}$ tels que

$$y(x) = (a_1x + a_2)\exp(2x) + (a_3x^2 + a_4x + a_5)\exp(3x),\ \forall x \in \mathbb{R}.$$

4.5.2 Applications aux suites récurrentes linéaires à coefficients constants

On prend $\mathbb{K} = \mathbb{Q}, \mathbb{R}$, ou $\mathbb{C}$. (ou un sous-corps de $\mathbb{C}$)
On considère l'*équation récurrente linéaire* à coefficients constants d'ordre p :

$$\forall n \in \mathbb{N}, u_{n+p} + \alpha_{p-1}u_{n+p-1} + \cdots + \alpha_1 u_{n+1} + \alpha_0 u_n = 0 \qquad (ER)$$

avec $\alpha_0,\ldots,\alpha_{p-1} \in \mathbb{K}$. L'inconnue est une suite $u = (u_n)_n \in \mathbb{K}^{\mathbb{N}}$.

On appelle **polynôme caractéristique** de (ER) le polynôme :

$$\chi(X) = X^p + \sum_{k=0}^{p-1} \alpha_k X^k.$$

On suppose que le polynôme $\chi(X)$ est scindé sur $\mathbb{K}$, de factorisation :

$$\chi(X) = \prod_{k=1}^{r} (X - \lambda_k)^{p_k},$$

avec les racines λ_k deux à deux distinctes et $p_k > 0$.
On note T l'endomorphisme de translation sur $\mathcal{L}(\mathbb{K}^{\mathbb{N}})$, défini par :

$$\begin{array}{rcl} T \,:\, \mathbb{K}^{\mathbb{N}} & \longrightarrow & \mathbb{K}^{\mathbb{N}} \\ (u_n)_{n\in\mathbb{N}} & \longmapsto & (u_{n+1})_{n\in\mathbb{N}}. \end{array}$$

L'ensemble $\mathcal{S}$ des solutions de l'équation (ER) est le noyau de l'endomorphisme $\chi(T)$ sur $\mathbb{K}^{\mathbb{N}}$.
Comme une suite dans $\mathcal{S}$ est déterminée par ses p premiers termes, l'application

$$\begin{array}{rcl} u \,:\, \mathcal{S} & \longrightarrow & \mathbb{K}^p \\ (u_n)_{n\in\mathbb{N}} & \longmapsto & (u_0, \ldots, u_{p-1}) \end{array}$$

est un isomorphisme, donc $\dim(\mathcal{S}) = p$.
De plus, le lemme des noyaux nous fournit la décomposition en somme directe :

$$\mathcal{S} = \bigoplus_{k=1}^{r} \operatorname{Ker}\left((T - \lambda_k \operatorname{Id}_{\mathbb{K}^{\mathbb{N}}})^{p_k}\right).$$

Il faut alors déterminer, pour tout k, le noyau de l'endomorphisme $(T - \lambda_k \operatorname{Id}_{\mathbb{K}^{\mathbb{N}}})^{p_k}$.
Ce sous-espace est de dimension p_k.

- Si $\lambda_k = 0$, il s'agit du noyau de T^{p_k}. C'est évidemment l'espace vectoriel des suites (u_n) telle que $u_n = 0$ pour tout $n \geqslant p_k$. On retrouve qu'il est de dimension p_k.
- Supposons $\lambda_k \neq 0$. Soit $U \in \mathbb{K}[X]$ un polynôme. Un calcul montre que

  $$(T - \lambda_k \operatorname{Id}_E)((\lambda_k^n U(n))_{n\in\mathbb{N}}) = \left(\lambda_k^{n+1} V(n)\right)_{n\in\mathbb{N}},$$

  avec $V(n)$ le polynôme $U(n+1) - U(n)$.
  On montre alors par récurrence sur p_k que $\operatorname{Ker}\left((T - \lambda_k \operatorname{Id}_E)^{p_k}\right)$ contient l'ensemble :

  $$\left\{(\lambda_k^n U(n))_{n\in\mathbb{N}},\, U \in \mathbb{K}[X],\, \deg(U) \leq p_k - 1\right\}.$$

  Cet ensemble est lui aussi un sous-espace vectoriel de dimension p_k, puisqu'il est l'image de $\mathbb{K}_{p_k-1}[X]$ par l'application linéaire injective

  $$\begin{array}{rcl} \Phi \,:\, \mathbb{K}_{p_k-1}[X] & \longrightarrow & \operatorname{Ker}\left((T - \lambda_k \operatorname{Id}_E)^{p_k}\right) \\ U & \longmapsto & (\lambda_k^n U(n))_{n\in\mathbb{N}}. \end{array}$$

  Notons qu'on utilise ici le fait que $\mathbb{K}$ contienne $\mathbb{N}$ (c'est faux si $\mathbb{K} = \mathbb{Z}/p\mathbb{Z}$ par ex.)
  Par égalité des dimensions, on obtient alors :

  $$\operatorname{Ker}\left((T - \lambda_k \operatorname{Id}_E)^{p_k}\right) = \left\{(\lambda_k^n U(n))_{n\in\mathbb{N}},\, U \in \mathbb{K}_{p_k-1}[X]\right\}.$$

En supposant que $\alpha_0 \neq 0$, on trouve finalement que toute solution u de (ER) est de la forme :

$$\forall n \in \mathbb{N}, u_n = \sum_{k=1}^{r} \lambda_k^n U_k(n)$$

pour une unique famille $U_1, \ldots, U_r \in \mathbb{K}[X]$ avec $\deg(U_k) \leq p_k - 1$.
On retrouve aussi que l'espace $\mathcal{S}$ est de dimension p.

Exemple 84 — *Trouver toutes les suites $(u_n)_n$ à coefficients complexes vérifiant*

$$u_{n+4} = 4u_{n+3} - 3u_{n+2} - 4u_{n+1} + 4u_n,\ \forall n \geq 0\,(E)$$

Pour $T : (u_n)_n \mapsto (u_{n+1})_n$ l'application linéaire de décalage à gauche sur $(\mathbb{C})^{\mathbb{N}}$, les suites $u = (u_n)_n$ vérifiant (E) sont exactement les suites telles que $P(T)(u) = 0$, pour P le polynôme $P(X) = X^4 - 4X^3 + 3X^2 + 4X - 4$. On trouve que $1, -1, 2$ sont des racines de P. On obtient $P(X) = (X-1)(X+1)(X-2)^2$.
Ainsi, on a $Ker(P(T)) = Vect((1^n)_n, ((-1)^n)_n, (2^n)_n, (n2^n)_n)$.
Donc, une suite u vérifie (E) si et seulement s'il existe $a, b, c, d \in \mathbb{C}$ tels que

$$u_n = a + b(-1)^n + (cn + d)2^n,\ \forall n \geq 0.$$

Chapitre 5 Orthogonalité

5.1 Bases orthonormées

Nous travaillerons jusqu'à la fin de ce chapitre sur un $\mathbb{R}$-e.v. euclidien E désigne un espace vectoriel euclidien, muni de sa norme euclidienne $||\cdot||$ et de la distance d associée.

5.1.1 Familles orthonormées

Définition 1
Soit E un $\mathbb{R}$-e.v. euclidien. Soit $x \in E$.
On dit que x est **unitaire** \单位向量\, ou **normé** , si $\|x\| = 1$.

Exemples 2

1. *Dans* $\mathbb{R}^2$, $x = (\frac{1}{\sqrt{2}}, \frac{-1}{\sqrt{2}})$ *est unitaire.*
2. *Dans* $\mathbb{R}^n$ *muni du produit scalaire canonique, les vecteurs de la base canonique sont unitaires.*
3. *Pour* $E = C^0(\mathbb{R}/2\pi Z)$ *l'e.v. des fonctions continues et* 2π*-périodiques sur* $\mathbb{R}$*, muni du produit scalaire :*
$$(f, g) \longmapsto \frac{1}{\pi}\int_0^{2\pi} f(x)\, g(x)\, dx,$$
les fonctions sin *et* cos *sont unitaires.*

Définition 3
Soit E un $\mathbb{R}$-e.v. euclidien. Soient $x, y \in E$.
On dit que x et y sont **orthogonaux** \正交的\ si l'on a $< x|y >= 0$. On note alors $x \perp y$.

Remarque 4 — *Comme* $\langle .|. \rangle$ *est symétrique, on a* $< x|y >= 0$ *si et seulement si* $< y|x >= 0$*. Donc la relation précédente est symétrique.*

Exemples 5

1. *Dans* $\mathbb{R}^2$*, on a* $x(\frac{1}{\sqrt{2}}, \frac{-1}{\sqrt{2}}) \perp (\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})$.
2. *Pour* $\mathbb{R}^n$ *muni du produit scalaire canonique, les vecteurs* $e_1, \dots, e_n$ *de la base canonique sont orthogonaux deux à deux.*
3. *Pour* $E = C^0(\mathbb{R}/2\pi\mathbb{Z})$ *l'e.v. des fonctions continues et* 2π*-périodiques sur* $\mathbb{R}$*, muni du produit scalaire :*
$$(f, g) \longmapsto \frac{1}{\pi}\int_0^{2\pi} f(x)\, g(x)\, dx,$$
les fonctions sin *et* cos *sont orthogonales.*

Définition 6
Soit E un $\mathbb{R}$-e.v. euclidien. Soit A une partie de E.
On définit **l'orthogonal de** A, noté $A^\perp$, par :
$$A^\perp = \{x \in E \text{ tels que } x \perp y,\, \forall y \in A\}\,.$$

Proposition 7
Soient E un $\mathbb{R}$-e.v. euclidien, et A une partie de E.
Alors $A^\perp$ est un sous-espace vectoriel de E.

Preuve — Soit A une partie de E.

- $A^\perp$ contient le vecteur nul puisque celui-ci est orthogonal à tout élément de A.
- Soient $x, y \in A^\perp$ et $\lambda \in \mathbb{R}$. Pour tout $y \in A$, on a :
$$\langle t|x + \lambda y\rangle = \langle t|x\rangle + \lambda\langle t|y\rangle = 0.$$
Ainsi, $x + \lambda y \in A^\perp$, donc $A^\perp$ est un sous-espace vectoriel de E.

□

Exemples 8

1. *On a* $\{0\}^\perp = E$.
2. *On a* $E^\perp = \{0\}$. *En effet :*
   - $E^\perp$ *est un sous-e.v. de* E, *donc il contient* 0.
   - *Pour* $x \in E^\perp$, *alors on a* $x \perp x$, *donc* $||x||^2 =< x|x >= 0$. *Cela donne* $x = 0$, *donc* $E^\perp = \{0\}$.
3. *Soient* A, B *des sous-parties de* E. *Si* $A \subset B$, *alors* $B^\perp \subset A^\perp$. ◀
4. *Soit* $a \in E$ *non-nul. Alors le sous-e.v.* $\{a\}^\perp$ *est exactement le noyau de la forme linéaire non nulle* $\phi_a : x \mapsto < a|x >$. *Ainsi,* $\{a\}^\perp$ *est un hyperplan de* E. ◀

Définition 9
Soit E un $\mathbb{R}$-e.v. euclidien. Soient F, G deux sous-e.v. de E.
On dit que F et G sont **orthogonaux** \正交子空间\ si l'on a $x \perp y$, $\forall (x, y) \in F \times G$.

Exemples 10

1. *Pour* F *un sous-e.v. de* E, *les sous-e.v.* F *et* $F^\perp$ *sont orthogonaux, puisque par définition les éléments de* $F^\perp$ *sont orthogonaux à tous les éléments de* F.
2. *Les sous-espace vectoriels* F *et* G *sont orthogonaux si, et seulement si,* $F \subset G^\perp$. ◀

Théorème 11 (Théorème de Pythagore \勾股定理\)
Soient E un $\mathbb{R}$-e.v. euclidien et $x, y \in E$.
Alors x et y sont orthogonaux si et seulement si $||x+y||^2 = ||x||^2 + ||y||^2$.

Preuve — Conséquence de la Proposition 26 et de la définition de l'orthogonalité. □

Définition 12
Soit E un $\mathbb{R}$-e.v. euclidien. Soit $(e_i)_{i\in I}$ une famille de vecteurs de E.
On dit que la famille $(e_i)_{i\in I}$ est une **famille orthogonale** \正交向量组\ si tous ses vecteurs sont deux à deux orthogonaux.
On dit que la famille $(e_i)_{i\in I}$ est une **famille orthonormée** \标准正交向量组\ (ou **orthonormale**) si tous ses vecteurs sont unitaires et deux à deux orthogonaux.

Exemples 13

1. *Dans* $\mathbb{R}^2$ *muni du produit scalaire canonique, la famille* $((\frac{1}{\sqrt{2}}, \frac{-1}{\sqrt{2}}), (\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}))$ *est orthonormée.*
2. *Dans* $\mathbb{R}^n$ *muni de son produit scalaire canonique, la base canonique* $(e_1, \ldots, e_n)$ *est une famille orthonormée.*
3. *Pour* $E = C^0(\mathbb{R}/2\pi Z)$ *l'e.v. des fonctions continues et* 2π*-périodiques sur* $\mathbb{R}$, *muni du produit scalaire :* ◀

$$(f, g) \longmapsto \frac{1}{\pi}\int_0^{2\pi} f(x)\,g(x)\,dx,$$

$$\text{la famille } \{x \mapsto \frac{1}{\sqrt{2}}, \quad x \mapsto \cos(x), \quad x \mapsto \cos(2x), \quad x \mapsto \cos(nx), \forall n \in \mathbb{N}^*,$$
$$x \longmapsto \sin(x), \quad x \longmapsto \sin(2x), \quad x \mapsto \sin(nx), \forall n \in \mathbb{N}^*\} \text{ est orthonormée.}$$

Proposition 14
Soit E un $\mathbb{R}$-e.v. euclidien. Soit $(x_1, x_2, \ldots, x_n)$ une famille de vecteurs de E.
Si cette famille est orthogonale, alors on a :

$$\left|\left|\sum_{i=1}^n x_i\right|\right|^2 = \sum_{i=1}^n ||x_i||^2.$$

Preuve — Par bilinéarité du produit scalaire, on a :

$$< \sum_{i=1}^n x_i | \sum_{i=1}^n x_i >= \sum_{1\leqslant i,j\leqslant n} < x_i|x_j >= \sum_{i=1}^n < x_i|x_i >$$

puisque $< x_i|x_j >= 0$ si $i \neq j$. □

Proposition 15
Soit E un $\mathbb{R}$-e.v. euclidien. Soit $(e_1, e_2, \ldots, e_n)$ une famille de vecteurs de E.
Si cette famille est orthonormée, alors on a :

1. Pour $x = \sum_{i=1}^{n} \lambda_i e_i \in \mathrm{Vect}(e_1, \ldots, e_n)$, on a $\lambda_i =< e_i|x >$, $\forall 1 \leq i \leq n$.
2. La famille $(e_1, e_2, \ldots, e_n)$ est libre.

Preuve

1. Par linéarité à droite du produit scalaire, on a :

$$< e_i|x >= \sum_{j=1}^{n} \lambda_j \ < e_i|e_j >= \lambda_i.$$

2. Soit $(\lambda_i)_{1\leqslant i\leqslant n} \in \mathbb{R}^n$ tel que $\sum_{k=1}^{n} \lambda_k e_k = 0$. Alors on :

$$0 = \quad < e_i|0 >=< e_i| \sum_{i=k}^{n} \lambda_k e_k >= \sum_{i=k}^{n} \lambda_k < e_i|e_k >= \lambda_i, \ \forall i \in [\![1, n]\!].$$

Donc cette famille est bien libre.

□

5.1.2 Bases orthonormées

Définition 16
Soit E un $\mathbb{R}$-e.v. euclidien de dimension finie n. Soit $\mathcal{B} = (e_1, e_2, \ldots, e_n)$ une base de E.
On dit que la base $\mathcal{B}$ est une **base orthonormée** \标准正交基\ de E (ou b.o.n.) si la famille $(e_1, e_2, \ldots, e_n)$ est orthonormée.

Exemple 17 — *Dans $\mathbb{R}^2$ muni du produit scalaire canonique, la famille $((\frac{1}{\sqrt{2}}, \frac{-1}{\sqrt{2}}), (\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}))$ est une base orthonormée.*

Dans $\mathbb{R}^n$ muni du produit scalaire canonique, la base canonique $(e_1, \ldots, e_n)$ est une base orthonormée.

Proposition 18
Soit E un $\mathbb{R}$-e.v. euclidien de dimension finie n.
Alors E possède au moins une base orthonormée.

Preuve — Raisonnons par récurrence sur n la dimension de E.

- Si $\dim(E) = 1$, pour $x \neq 0$, la famille $(\frac{x}{\|x\|})$ une base orthonormée de E.
- Soit $n \geqslant 1$. Supposons que tout e.v. euclidien de dimension $n-1$ possède une base orthonormée. Soit E un e.v. euclidien de dimension n.
  Soit $a \in E$ non-nul. Quitte à prendre $\frac{a}{\|a\|}$, on peut supposer que a est unitaire.
  On a vu que l'ensemble $\{a\}^\perp$ des vecteurs orthogonaux à a est un hyperplan de E. Comme $dim(\{a\}) = n-1$, l'hypothèse de récurrence nous dit que $\{a\}^\perp$ possède une base orthonormée $(e_1, e_2, \ldots, e_{n-1})$. Comme a est unitaire et orthogonal à tous les e_i, la famille $(e_1, e_2, \ldots, e_{n-1}, a)$ est donc une famille orthonormée. D'après la proposition précédente, cette famille est libre. Comme elle contient n vecteurs, c'est donc une base orthonormée de E. Ce qui conclut.

□

Remarque 19 — *Soit E un $\mathbb{R}$-e.v. euclidien de dimension finie. Tout sous-espace vectoriel F de E est euclidien pour le produit scalaire induit, et donc admet au moins une base orthonormée.*

Proposition 20
Soient E un $\mathbb{R}$-e.v. euclidien de dimension n, et $\mathcal{B} = (e_1, e_2, \ldots, e_n)$ une base orthonormée de E. Soient $x, y \in E$.
On a :

1. $x = \sum_{i=1}^{n} < x|e_i > e_i$.
2. Pour $x = \sum_{i=1}^{n} x_i e_i$ et $y = \sum_{i=1}^{n} y_i e_i$, on a

$$< x|y >= \sum_{i=1}^{n} x_i y_i \quad \text{et} \quad ||x||^2 = \sum_{i=1}^{n} x_i^2.$$

Preuve — L'égalité $x = \sum_{i=1}^{n} < x|e_i > e_i$ est une conséquence de la Proposition 15.

Pour $x = \sum_{i=1}^n x_i\, e_i$ et $y = \sum_{i=1}^n y_i\, e_i$, la bilinéarité du produit scalaire donne :

$$\langle x|y\rangle = \langle \sum_{i=1}^n x_i\, e_i| \sum_{j=1}^n y_j\, e_j\rangle = \sum_{i=1}^n \sum_{j=1}^n x_i y_j \langle e_i|e_j\rangle = \sum_{i=1}^n x_i\, y_i.$$

On en déduit alors que $||x||^2 = \langle x|x\rangle = \sum_{i=1}^n x_i^2$ □

Corollaire 21
Soit E un $\mathbb{R}$-e.v. euclidien de dimension n, et $\mathcal{B} = (e_1, e_2, \ldots, e_n)$ une base orthonormée de E. Alors, on a :

- La fonction $f : x = x_1 e_1 + \ldots + x_n e_n \in E \mapsto (x_1, \ldots, x_n) \in \mathbb{R}^n$ isomorphisme d'espaces vectoriels de E sur $\mathbb{R}^n$.
- Si l'on munit $\mathbb{R}^n$ de sa structure euclidienne canonique, cet isomorphisme f conserve la norme et le produit scalaire.
- Pour $x, y \in E$ de coordonnées $(x_1, x_2, \ldots, x_n)$ et $(y_1, y_2, \ldots, y_n)$, on a :

$$d(x, y) = \|x - y\| = \left(\sum_{i=1}^n (x_i - y_i)^2\right)^{1/2}.$$

Ainsi, un e.v. euclidien de dimension n s'étudie et se manipule de la même façon que $(\mathbb{R}^n, \langle .|.\rangle)$.

5.1.3 Procédé d'orthonormalisation de Schmidt

Théorème 22
Soit E un $\mathbb{R}$-e.v. euclidien de dimension n. Soit $\mathcal{B} = (e_1, e_2, \ldots, e_n)$ une base de E.
Alors il existe une base orthonormée $(f_1, f_2, \ldots, f_n)$ de E telle que :

$$\mathrm{Vect}(e_1, e_2, \ldots, e_p) = \mathrm{Vect}(f_1, f_2, \ldots, f_p),\ \forall p \in [\![1, n]\!].$$

On peut construire cette base à l'aide d'un algorithme.

Preuve — Nous allons démontrer cela par récurrence sur p. Cette démonstration donne en fait un algorithme efficace qui permet de construire cette base. On l'appelle le **procédé d'orthonormalisation de Schmidt** \施密特正交化\.

- Le vecteur f_1 doit être un vecteur unitaire et colinéaire à e_1. On prend alors $f_1 = \frac{e_1}{||e_1||}$.
- Soit $1 \leq p \leq n$. Supposons avoir une famille $(f_1,f_2,\ldots,f_p)$ orthonormée telle que :

$$\mathrm{Vect}(e_1, e_2, \ldots, e_k) = \mathrm{Vect}(f_1, f_2, \ldots, f_k),\ \forall k \in [\![1, p]\!].$$

Comme on a $\mathrm{Vect}(e_1, e_2, \ldots, e_p) = \mathrm{Vect}(f_1, f_2, \ldots, f_p)$, tout vecteur de $\mathrm{Vect}(e_1, e_2, \ldots, e_{p+1})$ peut s'écrire comme combinaison linéaire de $f_1, f_2, \ldots, f_p, e_{p+1}$.
On cherche donc g_{p+1} orthogonal à $f_1,f_2,\ldots,f_p$, de la forme :

$$g_{p+1} = e_{p+1} - \sum_{i=1}^p \lambda_i\, f_i.$$

L'orthogonalité avec $f_1, \ldots, f_p$ donne :

$$0 =< f_i|g_{p+1} >=< f_i|e_{p+1} > -\lambda_i,\ \forall 1 \leq i \leqslant p$$

En prenant $\lambda_i =< f_i|e_{p+1} >$, le vecteur g_{p+1} est donc orthogonal à $f_1,\ldots,f_p$. Le vecteur g_{p+1} est aussi non nul puisque :

$$e_{p+1} \notin \mathrm{Vect}(e_1, e_2, \ldots, e_p) = \mathrm{Vect}(f_1, f_2, \ldots, f_p).$$

On pose alors $f_{p+1} = \frac{g_{p+1}}{||g_{p+1}||}$.
La famille $(f_1, f_2, \ldots, f_{p+1})$ est alors une famille orthonormée, donc libre, dans $\mathrm{Vect}(e_1, e_2, \ldots, e_{p+1})$. Comme elle possède $p + 1$ vecteurs, c'est donc une base de ce sous-e.v. :

$$\mathrm{Vect}(f_1, f_2, \ldots, f_{p+1}) = \mathrm{Vect}(e_1, e_2, \ldots, e_{p+1}).$$

□

Exemple 23 — *On prend $E = \mathbb{R}^3$, muni du produit scalaire (voir Exemple 5) :*

$$< (x, y, z)|(x', y', z') >= x\,x' + y\,y' + z\,z' + \frac{1}{2}(x\,y' + x'\,y + x\,z' + x'\,z + y\,z' + y'\,z)$$

dont la norme associée est :

$$||(x, y, z)|| = \sqrt{x^2 + y^2 + z^2 + x\,y + x\,z + y\,z}.$$

A partir de la base canonique (e_1, e_2, e_3), construisons une base orthonormée (f_1, f_2, f_3) de $\mathbb{R}^3$ avec le procédé de Schmidt.

- *Le vecteur* $e_1 = (1,0,0)$ *est unitaire, donc on peut prendre* $f_1 = e_1$.
- *Cherchons* g_2 *orthogonal à* f_1 *de la forme :* $g_2 = e_2 - \lambda f_1$.
  On a $< f_1|g_2 >=< f_1|e_2 > -\lambda$*, donc il suffit de prendre* $\lambda =< f_1|e_2 >= \dfrac{1}{2}$. *Cela donne :*

$$g_2 = \left(-\frac{1}{2}, 1, 0\right) \quad \textit{et} \quad f_2 = \frac{1}{\sqrt{3}}(-1,2,0).$$

- *Cherchons* g_3 *orthogonal à* f_1 *et* f_2 *de la forme :* $g_3 = e_3 - \lambda f_1 - \mu f_2$. *Il suffit de prendre :*

$$\lambda =< f_1|e_3 >= \frac{1}{2} \quad \textit{et} \quad \mu =< f_2|e_3 >= \frac{1}{2\sqrt{3}}.$$

  Cela donne :

$$g_3 = \left(-\frac{1}{3}, -\frac{1}{3}, 1\right) \quad \textit{et} \quad f_3 = \frac{1}{\sqrt{6}}(-1,-1,3).$$

Remarque 24 — *Soit* E *un* $\mathbb{R}$*-e.v. euclidien de dimension* n*,* $\mathcal{B} = (e_1, e_2, \cdots, e_n)$ *une base de* E*,et* $\mathcal{C} = (f_1, f_2, \cdots, f_n)$ *une base orthonormée obtenue avec le procédé de Schmidt.*
Soit P *La matrice de passage de* $\mathcal{B}$ *vers* $\mathcal{C}$*. Alors la matrice* P *est triangulaire supérieure et ses éléments diagonaux sont non-nuls (*f_i *est une combinaison linéaire de* $e_1, \ldots, e_i$*).*

5.1.4 Supplémentaire orthogonal

Proposition-Définition 25
Soit E un $\mathbb{R}$-e.v. euclidien de dimension n. Soit F un sous-e.v. de E. Alors, on a :

1. F et $F^\perp$ sont supplémentaires dans E ;
2. $\dim\left(F^\perp\right) + \dim(F) = \dim(E)$;
3. $(F^\perp)^\perp = F$.

Le sous-ev $F^\perp$ est appelé le **supplémentaire orthogonal** \正交补\ de F. On écrit alors $E = F \overset{\perp}{\oplus} F^\perp$, pour signifier que la somme directe entre F et $F^\perp$ est orthogonale.

Preuve —

1.
   - On a $F \cap F^\perp = \{0\}$ puisqu'un vecteur x dans F et $F^\perp$ est orthogonal à lui-même, ce qui implique que $x = 0$.
   - D'après le théorème de Scmhidt, il existe $\mathcal{B} = (e_1, e_2, \ldots, e_p)$ une base orthonormée de F. Soit $x \in E$. On pose $y = \sum_{i=1}^{p} < e_i|x > e_i$. On a donc :
   $$< e_i|y >=< e_i|x >, \forall i \in [\![1,p]\!].$$
   En posant $z = x - y$, le vecteur z est ainsi orthogonal à $e_1, \ldots, e_p$, donc à $\mathrm{Vect}(e_1, \ldots, e_p) = F$. Ainsi, on a $x = y + z$, avec $y \in F$ et $z \in F^\perp$. Donc F et $F^\perp$ sont supplémentaires dans E (le vecteur y est appelé projeté orthogonal de x sur F).
2. Conséquence du fait que F et $F^\perp$ sont supplémentaires.
3. Comme tout élément de F est orthogonal à tout élément de $F^\perp$ on a $F \subset (F^\perp)^\perp$. Comme on a :
$$\dim\left(\left(F^\perp\right)^\perp\right) = n - \dim\left(F^\perp\right) = \dim(F),$$
on en déduit que $F = (F^\perp)^\perp$.

□

Proposition 26
Soit E un $\mathbb{R}$-e.v. euclidien de dimension n.
Alors, toute famille orthonormée de E peut être complétée en une base orthonormée de E.

Preuve — Soit $(e_1, e_2, \ldots, e_p)$ une famille orthonormée de E. On pose $F = \mathrm{Vect}(e_1, \ldots, e_p)$. Alors $F^\perp$ est le supplémentaire orthogonal de F dans E. D'après le procédé de Schmidt, $F^\perp$ possède une base orthonormée. Notons-la $(e_{p+1}, \ldots, e_n)$. On obtient alors que $(e_1, e_2, \ldots, e_n)$ est une famille orthonormée de E à $n = \dim(E)$ vecteurs, donc c'est une b.o.n. de E. □

Remarques 27

- *Comme nous l'avons vu dans la démonstration précédente, si* $E = F \overset{\perp}{\oplus} G$*, la réunion d'une base orthonormée de* F *et d'une base orthonormée de* G *est une base orthonormée de* E.

- *Réciproquement, si $\mathcal{B} = (e_1, e_2, \ldots, e_n)$ est une base orthonormée de E et si $p \leqslant n$, alors les sous-espaces vectoriels :*
$$F = \text{Vect}\{e_1, \ldots, e_p\} \quad \text{et} \quad G = \text{Vect}\{e_{p+1}, \ldots, e_n\}$$
sont deux sous-espaces vectoriels orthogonaux et supplémentaires dans E.

5.1.5 Équations d'un hyperplan

Remarque 28 — *Soient E un $\mathbb{R}$-e.v. euclidien et $a \in E$. On rappelle que l'on note ϕ_a la forme linéaire définie sur E par $\phi_a(x) = < a|x >$, et que l'on a $\{a\}^\perp = \text{Ker}\,(\phi_a)$.*

Définition 29
Soit E un $\mathbb{R}$-e.v. euclidien. Soient H un hyperplan de E et $u \in E$. On dit que le vecteur u est un **vecteur normal** \法线\ l'hyperplan H si u est non-nul et si u est orthogonal à H.

Proposition 30
Soit E un $\mathbb{R}$-e.v. euclidien de dimension n. Soient $\mathcal{B} = (e_1, \ldots, e_n)$ une b.o.n de E, H un hyperplan de E et $a \in E$ non-nul, avec $a = a_1e_1 + \ldots + a_ne_n$.
Alors, les propriétés suivantes sont équivalentes :
(i) $H = \text{Ker}\,(\phi_a)$;
(ii) $H = \{x = x_1e_1 + \ldots x_ne_n,$ tels que $\sum_{i=1}^n a_i\,x_i = 0.\}$
(iii) a est un vecteur normal à H.

Preuve
i) $\Longleftrightarrow$ *ii)* Puisque $\mathcal{B}$ est orthonormée, on a $\phi_a(x) = 0$ si et seulement si $\sum_{i=1}^n a_i\,x_i = 0$.
iii) $\Longrightarrow$ *i)* Si a est un vecteur normal à H, alors le noyau de ϕ_a est un hyperplan contenant H (de dim $n-1$). Il est donc égal à H.
i) $\Longrightarrow$ *iii)* Si $H = \text{Ker}\,(\phi_a)$, alors tous les éléments de H sont orthogonaux au vecteur a. Donc a est un vecteur normal à H.

□

Proposition 31
Soit E un $\mathbb{R}$-e.v. euclidien de dimension n. Soit $f : E \to \mathbb{R}$ une forme linéaire sur E.
Alors il existe un unique $a \in E$ tel que $f = \phi_a$.

Preuve
Unicité. Si $a, b \in E$ sont tels que $\phi_a = \phi_b$, alors on a :
$$\forall x \in E\,,\ < a|x > = < b|x > \ \text{ c'est-à-dire } \forall x \in E\,,\ < a - b|x > = 0.$$
Le vecteur $a - b$ est donc orthogonal à tous les éléments de E, et par conséquent il est nul.
Existence.
- Si $f = 0$, alors le vecteur nul convient.
- Sinon, le noyau de f est un hyperplan H. Comme $\dim(H) = n - 1$, l'orthogonal $H^\perp$ de H est de dimension 1. Pour $a \in H^\perp$ non-nul, on a alors $H = \text{Ker}\,(\phi_a)$. Ces deux formes linéaires non nulles ont le même noyau. Elles sont donc proportionnelles : il existe $\lambda \in \mathbb{R}$ tel que $f = \lambda\,\phi_a = \phi_{\lambda\,a}$.

□

Remarque 32 — *On peut aussi démontrer cette proposition en utilisant le théorème du rang. La fonction $f : a \in E \mapsto \phi_a \in \mathcal{L}(E, \mathbb{R})$ est une application linéaire de E dans $\mathcal{L}(E, \mathbb{R})$, l'espace vectoriel des formes linéaires sur E.*
- *Son noyau est réduit à 0, car si $\phi_a = 0$, alors tous les vecteurs de E sont orthogonaux au vecteur a et donc $a = 0$. Donc f est injective.*
- *Comme $\dim(E) = n = \dim(\mathcal{L}(E, \mathbb{R}))$, f est donc isomorphisme. Ainsi, toute forme linéaire s'écrit de façon unique sous la forme ϕ_a.*

Ce résultat est connu sous le nom de ***théorème de représentation de Riesz*** ***Riesz*** 表示定理\\.

5.2 Projections orthogonales

5.2.1 Projections vectorielles

Définition 33
Soit E un $\mathbb{R}$-e.v. euclidien de dimension finie. Soit F un sous-e.v. de E. Soit P la projection sur F parallèlement

à $F^\perp$.
La projection P est appelée **projection orthogonale** \正交投影\ sur F. On la note $P = p_F$.

PROPOSITION 34
Soient E un $\mathbb{R}$-e.v. euclidien de dimension finie, et F un sous-e.v. de E. Soit $\mathcal{B} = (e_1, e_2, \ldots, e_p)$ une base orthonormée de F. Alors, pour tout $x \in E$, on a :

$$p_F(x) = \sum_{i=1}^{p} < x|e_i > e_i.$$

Preuve — On a $x = p_F(x) + (x - p_F(x))$. Comme $x - p_F(x) \in F^\perp$, on a $\langle x - p_F(x)|e_i\rangle = 0$, c'est-à-dire $< x|e_i > = < p_F(x)|e_i >$. Comme $p_F(x) \in F$, on obtient le résultat. □

REMARQUE 35 — *Ainsi, il suffit d'avoir une base orthonormée du sous-espace F, par exemple avec l'algorithme de Gram-Schmidt, pour pouvoir calculer très facilement la projection orthogonale sur F.*

EXEMPLES 36

1. *Soit E un $\mathbb{R}$-e.v. euclidien de dimension n. Soit $a \in E$ unitaire. La proposition précédente nous donne alors l'expression de π, la projection orthogonale sur la droite vectorielle engendrée par a : $\pi(x) = < x|a > a$. Si le vecteur a n'est pas unitaire et est non-nul, alors $a' = \frac{a}{\|a\|}$ est unitaire. La projection sur la droite vectorielle engendrée par a vaut alors :*

$$\pi(x) = \langle x|a'\rangle a' = \frac{\langle x|a\rangle}{\|a\|}\frac{a}{\|a\|} = \frac{\langle x|a\rangle}{\langle a|a\rangle}a.$$

Soit $\mathcal{B}$ une base orthonormée de E. Si $(x_1, x_2, \ldots, x_n)$ (resp. $(a_1, a_2, \ldots, a_n)$) sont les composantes de x (resp. de a) dans la base $\mathcal{B}$, on a donc :

$$\pi(x) = \frac{\sum_{i=1}^{n} a_i\, x_i}{\sum_{i=1}^{n} a_i^2}\, a.$$

2. *Dans le procédé d'orthonormalisation de Schmidt, on construit le vecteur g_{p+1} comme :*

$$g_{p+1} = e_{p+1} - \sum_{i=1}^{p} \lambda_i\, f_i \text{ avec } \lambda_i = < f_i|e_{p+1} > .$$

Ainsi, pour $F = \mathrm{Vect}(e_1, e_2, \ldots, e_p)$, on a $g_{p+1} = e_{p+1} - p_F(e_{p+1})$. Le théorème de Pythagore nous donne alors :

$$||g_{p+1}||^2 = ||e_{p+1}||^2 - ||p_F(e_{p+1})||^2 = ||e_{p+1}||^2 - \sum_{i=1}^{p} \lambda_i^2 = \|e_{p+1}\|^2 - \sum_{i=1}^{p}\langle e_{p+1}|f_i\rangle^2.$$

Cela permet de calculer plus rapidement $\|g_{p+1}\|$.

5.2.2 Distance à un sous-espace

DÉFINITION 37
Soit E un $\mathbb{R}$-e.v. normé. Soient A une partie non-vide de E, et $x \in E$. On définit la **distance** de x à A comme :

$$d(x, A) = \inf_{y\in A} d(x, y) = \inf_{y\in A} \|x - y\|.$$

L'existence de cette quantité $d(x, A)$ vient du fait que $\{d(x, y),\ y \in A\}$ est une partie non vide de $\mathbb{R}_+$.

PROPOSITION 38
Soit E un $\mathbb{R}$-e.v. euclidien de dimension finie. Soient F un sous-ev de E et $x \in E$.
Alors, la distance de x à F vérifie $d(x, F) = \|x - p_F(x)\|$, où $p_F(x)$ est le projeté orthogonal de x sur F.
De plus, $p_F(x)$ et l'unique vecteur $y \in F$ tel que $d(x, F) = \|x - y\|$.

Preuve — Soit $y \in F$. D'après le théorème de Pythagore, on a :

$$d(x,y)^2 = \|x-y\|^2 = \|(x-p_F(x)) + (p_F(x)-y)\|^2 = \|x-p_F(x)\|^2 + \|p_F(x)-y\|^2 \geq \|x-p_F(x)\|^2 = d(x,p_F(x))^2.$$

On a donc $d(x,F) \geqslant d(y,p_F(x))$. Et comme $p_F(y) \in \mathbb{F}$, on obtient $d(x,F) = d(x,p_F(x))$.

De plus, on a $d(x,y) = d(x,F) = d(x,p_F(x))$ si et seulement si $\|p_F(x)-y\|^2 = 0$, ce qui est équivalent à $y = p_F(x)$. □

Remarque 39 —
• *Pour chercher à minimiser $\|x-y\|$ avec $y \in F$, on peut chercher à minimiser $\|x-y\|^2$, comme on le fait dans la preuve de la proposition.*
Comme $\|.\|$ est une norme euclidienne, on a $\|x-y\|^2 = \langle x-y, x-y\rangle$, et on peut utiliser les propriétés du produit scalaire dans les calculs.
• *Si $(e_1,\ldots,e_r)$ est une base orthonormée de F, on a $p_F(x) = \sum_{k=1}^r \langle x,e_k\rangle e_k$.*
Ainsi, on a

$$d(x,F)^2 = \|x-p_F(x)\|^2 = \|x - \sum_{k=1}^r \langle x,e_k\rangle e_k\|^2.$$

Si on complète la famille orthonormée $(e_1,\ldots,e_r)$ en une base orthonormée $(e_1,\ldots,e_n)$ de E, on a alors $x = \sum_{i=1}^n \langle x,e_i\rangle e_i$.
Cela donne donc

$$d(x,F)^2 = \|x - \sum_{k=1}^r \langle x,e_k\rangle e_k\|^2 = \|\sum_{i=r+1}^n \langle x,e_i\rangle e_i\|^2 = \sum_{i=r+1}^n \langle x,e_i\rangle^2.$$

Avec une base orthonormée de F complétée en une base orthonormée de E, on obtient une expression très simple de la distance d'un vecteur x à F.

Proposition 40
Soit E un $\mathbb{R}$-e.v. euclidien de dimension finie. Soient H un hyperplan de E et a un vecteur normal à H. Soit $x \in E$. On a alors :

$$d(x,H) = \frac{|\langle a|x\rangle|}{||a||}$$

Preuve — La distance d'un vecteur x à H est la norme du vecteur $x - p_H(x)$. Or, $x - p_H(x) \in H^\perp = \mathrm{Vect}(a)$. On a donc :

$$x - p_H(x) = \lambda \frac{a}{||a||}$$

pour un certain $\lambda \in \mathbb{R}$. On trouve :

$$\lambda = < (x-p_H(x))| \frac{a}{||a||} > = < x|\frac{a}{||a||} > + \underbrace{< p_H(x)|\frac{a}{||a||} >}_{=0}.$$

D'où la formule cherchée. □