# Chapitre 0

# Actions de Groupes

## Sommaire

# 1 Groupes, sous-groupes, morphismes, groupes engendrés

Voici quelques exemples de groupes : $\mathbb{Z}$, $S_n$, $\mathbb{Z}/n\mathbb{Z}$, $\mathrm{GL}_n(K)$, $O_n(\mathbb{R})\ldots$

# 2 Actions de Groupes

Une action d'un groupe $G$ sur un ensemble $X$ est la donnée d'un morphisme de groupes $\sigma : G \longrightarrow \mathrm{Bij}(X)$. On note souvent $g.x := \sigma(g)(x)$. Les propriétés de bases sont

$$g.(hx) = (gh).x \quad e.x = x,$$

avec des notations évidentes.
Lorsque $X$ est muni d'une structure additionelle, on considère souvent des actions qui respectent cette structure : $\sigma : G \longrightarrow \mathrm{Aut}(X)$. Exemples : actions linéaires, actions par automorphismes de groupes, actions par homéomorphismes...
Orbite de $x \in X$ :

$$G.x = \{g.x \mid g \in G\} \subset X$$

et stabilisateur de $X$ :

$$G_x = \{g \in G \,|\, g.x = x\}$$

est un sous-groupe. Ces deux notions sont reliées par une bijection : $G/G_x \longrightarrow G.x$. En particulier, si $G$ est fini on a $\sharp G.x = \frac{\sharp G}{\sharp G_x}$. Ici $G/H = \{gH \,|\, g \in H\}$ est une collection de parties de $G$ (avec $H = G_x$).
Les points d'une même orbite doivent être pensés comme équivalants, jouant le même rôle. Par exemple, leurs stabilisateurs sont conjugués :

$$G_{gx} = gG_xg^{-1}.$$

**Exemple 1.**
1. 3 Actions de $G$ sur lui-même ($g$ est un élément de $G$ et $h$ est pensé comme un point c'est-à-dire un élément de $X = G$) :

$$g \cdot h := gh \quad g \cdot h = hg^{-1} \quad g \cdot h = ghg^{-1}.$$

2. Action de $S_n$ sur l'ensemble $\{1, \ldots, n\}$. Il en découle des actions de $S_n$ sur les les parties à $k$ éléments de $\{1, \ldots, n\}$.
3. Action de $\mathrm{GL}_n(\mathbb{R})$ sur $\mathbb{R}^n$ ou de $\mathrm{GL}(V)$ sur $V$. Il en découle des actions de $\mathrm{GL}(V)$ sur l'ensemble des sev de $V$ de dimension donné $k$.
4. Trois actions de $\mathrm{GL}_n(K)$ sur $\mathcal{M}_n(K)$ :

$$g \cdot M = gM \quad g \cdot M = gMg^{-1} \quad g \cdot M = gM^tg.$$

**Proposition .1: Equation aux classes**

Les orbites forment une partition de l'ensemble $X$. En particulier, si $X$ est fini, et $x_1 \in \mathcal{O}_1, \ldots, x_s \in \mathcal{O}_s$ sont les orbites, on a

$$\sharp X = \sum_{i=1}^{s} \frac{\sharp G}{\sharp G_{x_i}}.$$

Une variante de cette *équation aux classes* est la **formule de Burnside** :

$$s = \sharp(X/G) = \frac{1}{\sharp G} \sum_{g \in G} \sharp\{x \in X \,:\, gx = x\}.$$

Pour démontrer cette formule on exprime le cardinal de

$$\mathcal{X} = \{(g, x) \in G \times X \,:\, gx = x\}$$

en projetant sur $X$ et sur $G$. Puis on égale les résultats.

De nombreux problèmes de classification s'expriment au moyen d'action de groupes. Pensez aux endomorphismes à changement de base près et à l'action de $\mathrm{GL}(E)$ par conjugaison. Il s'agit alors de comprendre les orbites. Un moyen de faire cela est de choisir un représentant pour chaque orbite. D'où la définition suivante.

**Définition .2**

Soit $G$ un groupe agissant sur un ensemble $X$. Un ensemble $(x_i)_{i\in I}$ d'éléments de $X$ est un *système complet de représentants* si

$$\forall \text{ orbite } \mathcal{O} \subset X \qquad \exists! i \in I \quad x_i \in \mathcal{O}.$$

Dans ce cours, on va construire des ensembles complets de représentants pour de nombreuses actions issues de l'algèbre linéaire.

# 3 Groupe quotient

Soit $H$ un sous-groupe d'un groupe $G$. On a l'application

$$\pi : G \longrightarrow G/H,\ g \longmapsto gH.$$

Est-il possible de munir $G/H$ d'une structure de groupe de telle sorte que $\pi$ soit un morphisme. Nous n'avons pas le choix, il faut poser

$$gH \cdot g'H = (gg')H.$$

Mais cela n'a pas toujours de sens car dépend du représentant.
On dit que $H$ est *distingué* si pour tout $g \in G$, on a $gHg^{-1} \subset H$.

**Proposition .3: Groupe quotient**

On peut munir $G/H$ d'une structure de groupe de telle sorte que $\pi$ soit un morphisme si et seulement si $H$ est distingué.

**Exemple 2.**
1. $\mathbb{Z}/n\mathbb{Z}$.
2. $\mathrm{PGL}_n(\mathbb{R})$ le quotient de $\mathrm{GL}_n(\mathbb{R})$ par son centre.

# 4 Reconstruire $G$ à partir de $G$ et $G/H$

Disons le d'amblé, cela n'est pas toujours possible et nécessite une donnée supplémentaire. Supposons (pour simplifier) que $G$ est fini et choisissons un élément $g_i$ dans chaque classe de $G/H$ :

$$G/H = \{g_1H, \ldots, g_sH\}.$$

Alors, l'application

$$\rho : G/H \times H \longrightarrow G,\ (g_i, h) \longrightarrow g_ih$$

est une bijection. Si bien, que l'on peut penser aux éléments de $G$ comme à des couples. Cela dépend du chooix des $g_i$.

**Produit semi-direct interne.** Supposons que $G$ contient un sous-groupe $K$ pour lequel la restriction de $\pi$ est un isomorphisme. Alors $K \cap H$ est trivial. Ceci est une hypothèse forte, mais la discuter nous emmènerait en cohomology des groupes (bien trop loin donc).
On vérifie alors que l'application $K \times H \longrightarrow G, (k, h) \longmapsto kh$ est une bijection. Pour comprendre le produit de deux paires, il faut faire :

$$(kh).(k'h') = (kk')((k'^{-1}hk)h').$$

Les données utiles pour ce calcul sont

1. la multiplication dans $K$ ;
2. la multiplication dans $H$ ;

3. la conjugaison d'un élément de $H$ par un élément de $K$.

Ce constat nous menne à la notion de produit semi-direct externe.

**Produit semi-direct interne.** Soit $K$ et $H$ deux groupes et $\rho : K \longrightarrow \mathrm{Aut}(H)$ un morphisme de groupes (c'est-à-dire une action de $K$ sur $H$ par automorphismes de groupes).
On définit alors une loi de groupe sur le produit ensembliste $K \times H$ par la formule :

$$(k, h).(k', h') = (kk', \rho(k'^{-1})(h)h'). \tag{4.1}$$

Le produit $K \times H$ muni de cette loi est noté $H \rtimes_\rho K$.

**Exercice 1.** *1. Montrer que $H$ est $K$ sont isomorphes à des sous-groupes de $H \rtimes_\rho K$. On identifie maintenant $H$ et $K$ à des sous-groupes.*

*2. Montrer que $H \rtimes_\rho K = KH = HK$.*

*3. Ecrire la loi sur le produit $H \times K$ induit par cette bijection.*

# Chapitre 1

# Algèbre Linéaire

## Sommaire

# 1 Matrices

Soit $K$ un corps. Soit $p$ et $q$ deux entiers naturels non nuls. Soit $E_{ij}$ la matrice élémentaire de la base canonique de $\mathcal{M}_{pq}(K)$. La formule suivante est utile en pratique

$$E_{ij}E_{kl} = \delta_j^k E_{il}.$$

C'est une version de la fameuse relation de Chasles.

Nous fixons ici une notation pour la matrice associée à une application linéaire entre espaces vectoriels munis de bases et illustrons sa pertinence sur les formules de changement de bases.
Soit $E$ et $F$ deux espaces vectoriels de dimensions finies et $f : E \longrightarrow F$ une application linéaire. Soit $\mathcal{B}_E$ et $\mathcal{B}_F$ des bases de $E$ et $F$ respectivement. On note

$$\mathrm{Mat}_{\mathcal{B}_F\,\mathcal{B}_E}(f), \tag{1.1}$$

la matrice de $f$. Le coefficient $(i, j)$ (ligne $i$ et colonne $j$) est la $i^{eme}$ coordonnées de l'image par $f$ du $j^{eme}$ vecteur de $\mathcal{B}_E$.
On remarquera que la base de l'espace d'arrivée survient en premier dans la notation. Cela est en cohérence avec la notation $M_{ij}$, ou encore avec le fait que

$$\mathrm{Mat}_{\mathcal{B}_F\,\mathcal{B}_E}(f) \in \mathcal{M}_{\sharp\mathcal{B}_F\,\sharp\mathcal{B}_E}(K).$$

Pour $v \in E$, on note $\mathrm{Mat}_{\mathcal{B}_E}(v)$ le vecteur colonne constitué des coordonnées de $v$ dans la base $\mathcal{B}_E$. Les seules formules à connaitre sont

$$\mathrm{Mat}_{\mathcal{B}_G\,\mathcal{B}_E}(g \circ f) = \mathrm{Mat}_{\mathcal{B}_G\,\mathcal{B}_F}(g)\mathrm{Mat}_{\mathcal{B}_F\,\mathcal{B}_E}(f) \qquad \mathrm{Mat}_{\mathcal{B}_F}(f(v)) = \mathrm{Mat}_{\mathcal{B}_F\,\mathcal{B}_E}(f)\mathrm{Mat}_{\mathcal{B}_E}(v), \tag{1.2}$$

où $v \in E$, $f : E \longrightarrow F$ et $g : F \longrightarrow G$.
Pour retenir la position des bases dans les formules (1.2), on pourra remarquer l'analogie avec la formule de Chasles : $\overrightarrow{GE} = \overrightarrow{GF} + \overrightarrow{FE}$.

**Changement de bases.** Si $\mathcal{B}'_E$ est une seconde base de $E$, on considère les matrices de passage $\mathrm{Mat}_{\mathcal{B}'_E\,\mathcal{B}_E}(\mathrm{Id}_E)$ et $\mathrm{Mat}_{\mathcal{B}_E\,\mathcal{B}'_E}(\mathrm{Id}_E)$. En appliquant les formules (1.2) à $\mathrm{Id}_E \circ \mathrm{Id}_E = \mathrm{Id}_E$ et à $f \circ \mathrm{Id}_E = f$ on trouve

$$\mathrm{Mat}_{\mathcal{B}_E\,\mathcal{B}'_E}(\mathrm{Id}_E)\mathrm{Mat}_{\mathcal{B}'_E\,\mathcal{B}_E}(\mathrm{Id}_E) = I_n, \qquad \mathrm{Mat}_{\mathcal{B}_F\,\mathcal{B}_E}(f)\mathrm{Mat}_{\mathcal{B}_E\,\mathcal{B}'_E}(\mathrm{Id}_E) = \mathrm{Mat}_{\mathcal{B}_F\,\mathcal{B}'_E}(f),$$

où $n$ est la dimension de $E$. De manière analoge, on obtient les formules de changement de base $PMP^{-1}$ et $PMQ^{-1}$.
On a aussi

$$\mathrm{Mat}_{\mathcal{B}_E\,\mathcal{B}'_E}(\mathrm{Id}_E) = \begin{matrix} \mathcal{B}'_E \\ \begin{pmatrix} \\ \end{pmatrix} \end{matrix} \; \mathcal{B}_E.$$

**Opérations élémentaires.**
On définit pour $\lambda \in K$ et $i, j$ des indices :

$$T_{ij}(\lambda) = I_n + \lambda E_{ij} \qquad D_i(\lambda) = I_n + (\lambda - 1)E_{ii} \qquad P_{ij} = I_n - E_{ii} - E_{jj} + E_{ij} + E_{ji}.$$

On a :
- $L_i \to L_i + \lambda L_j$ s'obtient par multiplication à gauche par $T_{ij}(\lambda)$ (pour tout $\lambda \in K$).
- $L_i \leftrightarrow L_j$ s'obtient par multiplication à gauche par $P_{ij}$.
- $L_i \to \lambda L_i$ s'obtient par multiplication à gauche par $D_i(\lambda)$ (pour tout $\lambda$ **non nul** dans $K$).

On a les mêmes opérations sur les colonnes en multipliant à droite.

# 2 Dualité

## 2.1 Formes linéaires et bases duales

Soit $E$ un espace vectoriel de dimension finie. Une *forme linéaire sur $E$* est une application linéaire de $E$ dans $K$. On note $E^*$ l'espace vectoriel constitué des formes linéaires. Observer que (1) est une base de $K$. Si $\mathcal{B}$ est une base de $E$ et $\varphi \in E^*$, $\mathrm{Mat}_{(1)\,\mathcal{B}}(\varphi)$ est une matrice **ligne** dont les entrées sont les valeurs de $\varphi$ aux éléments de $\mathcal{B}$.

**Exemple 3.** Toute forme linéaire sur $\mathbb{R}^2$ est donnée par une formule

$$\begin{array}{ccc} \mathbb{R}^2 & \longrightarrow & \mathbb{R} \\ \begin{pmatrix} x \\ y \end{pmatrix} & \longmapsto & ax + by \end{array}$$

pour des nombres réels fixés $a$ et $b$.

Un exemple de forme linéaire sur $\mathbb{R}^4$ est

$$\begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} \longmapsto 2x + 4y - z + t.$$

**Les formes linéaires coordonnées.** Explicitons $\mathcal{B} = (e_1, \ldots, e_n)$. L'application $e_i^*$ qui à un vecteur associe sa $i^{eme}$ coordonnée dans la base $\mathcal{B}$ est une fome linéaire. On a

$$\mathrm{Mat}_{(1)\,\mathcal{B}}(e_i^*) = E_{1i}.$$

De plus, $\mathrm{Mat}_{(1)\,\mathcal{B}}(\varphi) = (\lambda_1, \ldots, \lambda_n)$ si et seulement si

$$\varphi = \lambda_1 e_1^* + \cdots + \lambda_n e_n^*.$$

On en déduit que $(e_1^*, \ldots, e_n^*)$ est une base, notée $\mathcal{B}^*$ de $E^*$. La base $\mathcal{B}^*$ est appelée *base duale de $\mathcal{B}$*. En particulier $\dim(E^*) = n$. Remarquons aussi que $\mathrm{Mat}_{\mathcal{B}^*}(\varphi) = {}^t\mathrm{Mat}_{(1)\,\mathcal{B}}(\varphi)$. Observons aussi que la coordonnée $\lambda_i$ de $\varphi$ dans la base duale est

$$\lambda_i = \varphi(e_i) \quad \text{car} \quad e_j^*(e_i) = \delta_i^j. \tag{2.1}$$

Réciproquement, étant donnée une base $\mathcal{C}$ de $E^*$, on vérifie qu'il existe une unique base $\mathcal{B}$ de $E$ telle que $\mathcal{B}^* = \mathcal{C}$. La base $\mathcal{B}$ est appelée *base anteduale de $\mathcal{C}$*.

*Remarque.* Contrairement à ce que peut laisser croire la notation, $e_i^*$ dépend de la base $\mathcal{B}$ entière et pas seulement de $e_i$. Pour se convaincre de cela, regardons l'exemple suivant $\mathcal{B} = (e_1, e_2)$ est la base canonique de $\mathbb{R}^2$ et $\mathcal{C} = (\epsilon_1, \epsilon_2)$ est donnée par $\epsilon_1 = e_1$ et $\epsilon_2 = e_1 + e_2$. Soit $v = xe_1 + ye_2$ un vecteur. On a $v = (x-y)e_1 + y(e_1 + e_2) = (x-y)\epsilon_1 + y\epsilon_2$. Donc $\epsilon_1^*(v) = x - y$ ou encore $\epsilon_1^* = e_1^* - e_2^*$, alors que $\epsilon_1 = e_1$.

## 2.2 Hyperplans

Le noyau $\mathrm{Ker}\varphi$ d'une forme linéaire linéaire non nulle est un sous-espace vectoriel de $E$ de dimension $n-1$ (par application directe du théorème de rang). Réciproquement pour tout sous-espace $F$ de dimension $n-1$ de $E$, il existe une forme linéaire linéaire non nulle telle que $\mathrm{Ker}\varphi = F$. Un tel sous-espace vectoriel de $F$ est appelé *hyperplan*.
Pour montrer la réciproque, on peut par exemple construire une base de $E$ qui commence par une de $F$ et considérer la base duale.

## 2.3 Bidualité

L'espace vectoriel $E^*$ a lui-même un espace vectoriel dual $E^{**}$ appelé bidual de $E$.

**Théorème I.4. Isomorphisme avec le bidual**

L'application linéaire

$$\begin{array}{rccc} \iota : & E & \longrightarrow & E^{**} \\ & v & \longmapsto & \iota(v), \end{array}$$

où

$$\begin{array}{rccc} \iota(v) : & E^* & \longrightarrow & k \\ & \varphi & \longmapsto & \varphi(v) \end{array}$$

est bien définie et est un isomorphisme.

La démonstration est laissée en exercice. On pourra aussi montrer que si $\mathcal{C}$ est une base de $E^*$, on a

$$\left(\iota^{-1}\mathcal{C}^*\right)^* = \mathcal{C}.$$

Soit $\mathcal{B}$ une base de $E$. Alors

$$\mathrm{Mat}_{\mathcal{B}^{**}\mathcal{B}}(\iota) = I_n.$$

## 2.4 Orthogonalité

Pour $F$ un sous-espace de $E$, on appelle

$$F^\perp = \{\varphi \in E^* \,:\, \forall v \in F \quad \varphi(v) = 0\},$$

l'orthogonal de $F$. On vérifie que

1. $F^\perp$ est un sous-espace vectoriel de $E^*$ ;
2. si $\mathcal{B} = (e_1, \ldots, e_k, e_{k+1}, \ldots, e_n)$ est une base de $E$ telle que $(e_1, \ldots, e_k)$ est une base de $F$ alors $(e^*_{k+1}, \ldots, e^*_n)$ est une base de $F^\perp$ ;
3. $\dim F + \dim F^\perp = \dim E$.

De manière similaire, pour $G$ un sous-espace de $E^*$, on appelle

$$G^\circ = \{v \in E \,:\, \forall \varphi \in G \quad \varphi(v) = 0\},$$

l'ante-orthogonal de $G$. On vérifie que

1. $G^\circ$ est un sous-espace vectoriel de $E$ ;
2. si $\mathcal{C} = (e^*_1, \ldots, e^*_k, e^*_{k+1}, \ldots, e^*_n)$ est une base de $E^*$ telle que $(e^*_1, \ldots, e^*_k)$ est une base de $G$ alors $(e_{k+1}, \ldots, e_n)$ est une base de $G^\circ$ ;
3. $\dim G + \dim G^\circ = \dim E$.

Quelques propriétés de ces constructions :

**Proposition I.5: Propriétés de l'orthogonal et de l'anteorthogonal**

1. Si $F$ est un sous-espace vectoriel de $E$, on a $(F^\perp)^\circ = F$.
2. Si $G$ est un sous-espace vectoriel de $E^*$, on a $(G^\circ)^\perp = G$. De plus, $\iota(G^\circ) = G^\perp$.
3. Soit $F_1$ et $F_2$ deux sous-espaces vectoriels de $E$. Alors

   $$(F_1 + F_2)^\perp = F_1^\perp \cap F_2^\perp,$$

   et

   $$(F_1 \cap F_2)^\perp = F_1^\perp + F_2^\perp.$$

4. Soit $G_1$ et $G_2$ deux sous-espaces vectoriels de $E^*$. Alors

   $$(G_1 + G_2)^\circ = G_1^\circ \cap G_2^\circ,$$

et

$$(G_1 \cap G_2)^\circ = G_1^\circ + G_2^\circ.$$

**Preuve**

Les deux premières assertions découlent directement de la description des orthogonaux avec des bases.
En utilisant la deuxième assertion, la troisième est une conséquence de la dernière.
Montrons la dernière. Comme $G_1 \subset G_1+G_2$, on a $(G_1+G_2)^\circ \subset G_1^\circ$ puis que $(G_1+G_2)^\circ \subset G_1^\circ \cap G_2^\circ$.
Réciproquement soit $x \in G_1^\circ \cap G_2^\circ$. Soit $\psi_1 \in G_1$ et $\psi_2 \in G_2$. Alors

$$(\psi_1 + \psi_2)(x) = \psi_1(x) + \psi_2(x) = 0 + 0 = 0$$

et $x$ appartient à $(G_1 + G_2)^\circ$. Finalement $(G_1 + G_2)^\circ = G_1^\circ \cap G_2^\circ$.

Comme $G_1 \cap G_2 \subset G_1$, on a $(G_1 \cap G_2)^\circ \supset G_1^\circ$. Comme $(G_1 \cap G_2)^\circ$ est un espace vectoriel, on en déduit par symétrie que $(G_1 \cap G_2)^\circ \supset G_1^\circ + G_2^\circ$.
Par ailleurs, on a

$$\begin{aligned}\dim(G_1^\circ + G_2^\circ) &= \dim(G_1^\circ) + \dim(G_2^\circ) - \dim(G_1^\circ \cap G_2^\circ) = \dim(G_1^\circ) + \dim(G_2^\circ) - \dim((G_1 + G_2)^\circ) \\ &= n - \dim(G_1) - \dim(G_2) + \dim(G_1) + \dim(G_2) - \dim(G_1 \cap G_2) \\ &= \dim((G_1 \cap G_2)^\circ)\end{aligned}$$

Dans la première ligne, on a utilisé la formule de Grassmann puis l'égalité $(G_1 + G_2)^\circ = G_1^\circ \cap G_2^\circ$.
A la deuxième on utilise encore la formule de Grassmann.

## 2.5 Transposition

### 2.5.1 Définition et Matrices

Soit $f : E \longrightarrow F$ une application linéaire. On appelle *transposé de* $f$ l'application linéaire suivante :

$$\begin{array}{rccc} {}^t f : & F^* & \longrightarrow & E^* \\ & \varphi & \longmapsto & \varphi \circ f. \end{array}$$

Le lien avec la transposition des matrices est le suivant :

**Proposition I.6: Matrice de la transposé**

Soit $\mathcal{B}_E$ et $\mathcal{B}_F$ des bases de $E$ et $F$ respectivement. Alors

$$\mathrm{Mat}_{\mathcal{B}_E^* \mathcal{B}_F^*}({}^t f) = {}^t\mathrm{Mat}_{\mathcal{B}_F \mathcal{B}_E}(f).$$

**Preuve**

Notons $(e_1, \ldots, e_p)$ la base de $E$ et $(\epsilon_1, \ldots, \epsilon_q)$ celle de $F$. L'entrée $m_{ij}$ à la ligne $i$ et la colonne $j$ de $\mathrm{Mat}_{\mathcal{B}_E^* \mathcal{B}_F^*}({}^t f)$ est la $i^{eme}$ coordonnées de ${}^t f(\epsilon_j^*)$. D'après (2.1),

$$m_{ij} = \epsilon_j^* \circ f(e_i)$$

est la coordonnée en $\epsilon_j$ de $f(e_i)$, c'est-à-dire l'entrée $(j, i)$ de $\mathrm{Mat}_{\mathcal{B}_F \mathcal{B}_E}(f)$.

Des formules faciles à démontrer

$${}^t(f + \lambda g) = {}^t f + \lambda {}^t g \qquad {}^t(f \circ g) = {}^t g \circ {}^t f.$$

De plus, modulo les identifications $\iota_E$ et $\iota_F$, on a ${}^t({}^t f) = f$.

### 2.5.2 Noyaux et Images

**Proposition I.7: Noyau et Image de la transposé**

Soit $f : E \longrightarrow F$ et ${}^tf : F^* \longrightarrow E^*$. On a :

1. $\mathrm{Ker}{}^tf = (\mathrm{Im}f)^\perp$ ;
2. $\mathrm{Im}{}^tf = (\mathrm{Ker}f)^\perp$.

**Preuve**

Soit $\varphi \in F^*$. On a $\varphi \in \mathrm{Ker}{}^tf$ si et seulement si $\varphi \circ f = 0$ si et seulement si $\varphi \circ f(v) = 0$ pour tout $v \in E$ si et seulement si $\varphi(v) = 0$ pour tout $v \in \mathrm{Im}(f)$ si et seulement si $\varphi \in (\mathrm{Im}f)^\perp$.

Soit $\varphi \in F^*$ et donc ${}^tf(\varphi) \in \mathrm{Im}({}^tf)$. Soit $v \in \mathrm{Ker}f$. On a

$$ {}^tf(\varphi)(v) = \varphi(f(v)) = \varphi(0) = 0. $$

Donc ${}^tf(\varphi) \in (\mathrm{Ker}f)^\perp$ et $\mathrm{Im}{}^tf \subset (\mathrm{Ker}f)^\perp$.
Par ailleurs, on a :

$$ \begin{aligned} \dim(\mathrm{Im}{}^tf) &= \dim(F^*) - \dim(\mathrm{Ker}{}^tf) = \dim(F^*) - \dim((\mathrm{Im}f)^\perp) \\ &= \dim(F^*) - \dim(F) + \dim(\mathrm{Im}f) = \dim(E) - \dim(\mathrm{Ker}f) \\ &= \dim((\mathrm{Ker}f)^\perp). \end{aligned} $$

Avec l'inclusion déjà montrée cela permet de conclure que $\mathrm{Im}{}^tf = (\mathrm{Ker}f)^\perp$.

# 3 Réduction des endomorphismes

## 3.1 Polynôme minimal

Soit $A \in \mathcal{M}_n(K)$. Si $P = a_0 + a_1X + \cdots + a_dX^d$ est un polynôme, on pose

$$ P(A) = a_0I_n + a_1A + \cdots + a_dA^d. $$

On obtient ainsi un morphisme d'algèbres

$$ \varphi : K[X] \longrightarrow \mathcal{M}_n(K),\ A \longmapsto P(A). $$

L'ensemble $I$ des polynômes $P \in K[X]$ tels que $P(A) = 0$ est un idéal. Comme $K[X]$ est euclidien, cet idéal est engendré par un unique polynôme unitaire $\mu_A$, appelé polynôme minimal de $A$. On obtient un isomorphisme

$$ K[X]/(\mu_A) \simeq K[A] \subset \mathcal{M}_n(K), $$

où $K[A]$ est l'image de $\varphi$. Comme $\dim_K(\mathcal{M}_n(K)) \leqslant n^2$, le degré de $\mu_A$ est au plus $n^2$. Le théorème suivant donne mieux :

**Théorème I.8. Cayley-Hamilton**

Soit $\chi_A = \det(A - XI_n)$ le polynôme caractéristique de $A$. On a

$$ \chi_A(A) = 0. $$

Donc $\mu_A$ divise $\chi_A$.

**Preuve**

Hamilton-Cayley est vrai pour la matrice générique : c'est la matrice carrée $G = (T_{i,j})_{i,j=1,\dots,n}$ que l'on considère comme matrice à coefficients dans l'anneau de polynômes $R = \mathbb{Z}[Ti, j]$ que l'on peut plonger dans son corps de fractions $\mathbb{Q}(Ti, j)$. Si $M \in \mathcal{M}_n(K)$, il existe un unique morphisme $\theta : R \longrightarrow K$ qui envoie $G$ sur $M$ (on spécialise $G$ en $M$ par $\theta$). Le morphisme $\theta$ envoie toute identité algébrique vérifiée par les coefficients de $G$ sur l'identité correspondante pour $M$. En particulier, il suffit de vérifier Hamilton-Cayley pour $G$ pour l'obtenir pour n'importe quelle matrice à coefficients dans $K$ (et même dans un anneau commutatif).
Or $G$ vérifie Hamilton-Cayley parce qu'elle est diagonalisable sur le corps de décomposition de son polynôme caractéristique. Il suffit de vérifier que les valeurs propres de $G$ sont distinctes, c'est-à-dire que le discriminant $\Delta \in R$ de son polynôme caractéristique est non nul. Pour le montrer on peut spécialiser $G$ en une matrice diagonale $M$ à éléments diagonaux distincts sur $\mathbb{C}$, par exemple. Alors $\Delta$ s'envoie sur le discriminant du polynôme caractéristique de $M$ qui est non nul.

## 3.2 Lemme des noyaux

**Lemme I.9**

Soit $A$ un anneau euclidien (factoriel suffit). Soit $a_1, \dots, a_k$ non nuls dans $A$ et deux à deux premiers entre eux.
Posons $b = a_1 \times \cdots \times a_k$ et $c_j = b/a_j$.
Alors $c_1, \dots, c_k$ sont globalement premier entre eux.

**Preuve**

Soit $d$ un diviseur irréductible commun à tous les $c_j$.
Comme $d$ divise $c_1 = a_2 \dots a_k$ il divise un des $a_j$. Disons $d$ divise $a_{j_0}$. Comme $a_{j_0}$ est premier avec $a_1$, $d$ ne divise pas $a_1$.
Comme $d$ divise $c_2 = a_1 \times (a_3 \dots a_k)$, $d$ divise $(a_3 \dots a_k)$. Donc $d$ divise un des $a_j$ pour $j \geqslant 3$. Donc $d$ ne divise pas $a_2$.
Par récurrence, on obtient ainsi que $d$ ne divise aucun des $a_1, \dots, a_{k-1}$. Ce qui contredit $d$ divise $c_k$.

**Théorème I.10. Lemme des noyaux ou TDN**

Soit $E$ un $K$-espace vectoriel et $u$ un endomorphisme de $E$. Soit $P_1, \dots, P_k$ des polynômes deux à deux premiers entre eux. Alors

$$\mathrm{Ker}(P_1 \dots P_k)(u) = \mathrm{Ker}P_1(u) \oplus \cdots \oplus \mathrm{Ker}P_k(u).$$

Si, de plus, $(P_1 \dots P_k)(u) = 0$ alors, pour tout $j$, les projections $\pi_j$ sur $\mathrm{Ker}P_j(u)$ parallèlement à la somme des autres $\mathrm{Ker}P_i(u)$ est un polynôme en $u$.

**Preuve**

Posons $P = P_1 \dots P_k$ et $Q_j = P/P_j$. L'hypothèse implique que les $Q_j$ sont globalement premiers entre eux. Le théorème de Bezout donne l'existence de polynômes $V_i$ tels que $\sum V_i Q_i = 1$.
Soit $x \in E$. En évaluant cette relation en $u$, puis en $x$ on obtient

$$\sum V_i(u) \circ Q_i(u)(x) = x. \tag{3.1}$$

Si de plus, $x \in KerP(u)$ alors $V_i(u) \circ Q_i(u)(x) \in \mathrm{Ker}P_i(u)(x)$. Donc $KerP(u)$ est bien la somme des $\mathrm{Ker}P_i(u)(x)$.
Pour $j$ fixé, soit

$$x \in \mathrm{Ker}P_j(u) \cap (\sum_{i \neq j} \mathrm{Ker}P_i(u)).$$

Pour montrer que la somme est directe, nous devons montrer que $x = 0$.
Pour tout $i \neq j$, comme $P_j$ divise $Q_i$ et $x \in \mathrm{Ker}P_j(u)$, on a $Q_i(u)(x) = 0$. De même, $x \in \sum_{i \neq j} \mathrm{Ker}P_i(u)$ implique que $Q_j(u)(x) = 0$. Mais alors, la relation (3.1) montre que $x = 0$.

Reste à montrer que les $\pi_j$ sont des polynômes en $u$ avec l'hypothèse supplementaire. On va montrer que $\pi_i = (V_iQ_i)(u)$. On a déjà vu que $(V_iQ_i)(u)(x) \in \mathrm{Ker}P_i(u)$ pour tout $x$. Mais alors (3.1) et l'unicité de la décomposition en somme directe suffisent à conclure.

Précision sur CH :

**Théorème I.11**

Les polynômes $\mu_u$ et $\chi_u$ ont les mêmes facteurs irréductibles.

**Preuve**

Ecrivons $\chi_u = P_1^{\alpha_1} \dots P_s^{\alpha_s}$ sa décomposition en produit d'irréductibles. Le TDN et CH donnent

$$E = \mathrm{Ker}P_1^{\alpha_1}(u) \oplus \cdots \oplus \mathrm{Ker}P_s^{\alpha_s}(u).$$

On pose $E_i = \mathrm{Ker}P_i^{\alpha_i}(u)$ et $u_i$ la restriction de $u$ à $E_i$. Un point clé est de montrer que $E_i$ est non nul pour chaque $i$. On choisit une base et travaille matriciellement. Soit $\lambda \in K'$ une racine de $P_i$. Ici $K'$ est par exemple le corps de rupture de $P_i$. Il existe un vecteur propre $X$ de valeur propre $\lambda$. La relation $MX = \lambda X$ implique que $X \in E_i$ qui est donc non nul.
Par Cayley-Hamilton, on a $\mu_u = P_1^{\beta_1} \dots P_s^{\beta_s}$ avec $0 \leqslant \beta_i \leqslant \alpha_i$. Il s'agit de montrer que $\beta_i \neq 0$. Encore le TDN donne

$$E = \mathrm{Ker}P_1^{\beta_1}(u) \oplus \cdots \oplus \mathrm{Ker}P_s^{\beta_s}(u).$$

Or $\mathrm{Ker}P_1^{\beta_i}(u) \subset \mathrm{Ker}P_1^{\alpha_i}(u)$, pour tout $i$. En comparant les deux décompositions il vient $\mathrm{Ker}P_i^{\beta_i}(u) = \mathrm{Ker}P_i^{\alpha_i}(u) \neq \{0\}$. Donc $\beta_i \neq 0$.

## 3.3 Vecteur totalisateur

Si $x \in E$, on note $I_x$ l'idéal de $k[X]$ constitué des polynômes $P$ tels que $P(u)(x) = 0$. On note $\mu_{u,x}$ le générateur unitaire de cet idéal. On voit facilement que $\mu_{u,x}$ divise $\mu_u$. On a de plus le :

**Lemme I.12**

Il existe $x$ dans $E$ tel que $\mu_{u,x}$=$\mu_u$.

**Preuve**

On écrit $\mu_u = P_1^{\alpha_1} \cdots P_s^{\alpha_s}$, avec $P_i$ des polynômes irréductibles 2 à 2 distincts et $\alpha_i$ des entiers strictement positifs. D'après le **théorème de décomposition des noyaux**, on a :

$$E = \mathrm{Ker}(\mu_u(u)) = \mathrm{Ker}(P_1^{\alpha_1}(u)) \oplus \cdots \oplus \mathrm{Ker}(P_s^{\alpha_s}(u)).$$

On vérifie que pour tout $i$, $\mathrm{Ker}(P_i^{\alpha_i}(u))$ est stable par $u$ et $\mu_{u_{|\mathrm{Ker}(P_i^{\alpha_i}(u))}} = P_i^{\alpha_i}$. En effet, sinon $\mu_u$ serait plus petit ! Soit $x_i \in \mathrm{Ker}(P_i^{\alpha_i}(u)) - \mathrm{Ker}(P_i^{\alpha_i - 1}(u))$. Posons $x = x_1 + \cdots + x_s$.

On vérifie alors, que $P(u)(x) = 0$ si et seulement si pour tout $i$, $P(u)(x_i) = 0$, c'est-à-dire si et seulement si pour tout $i$, $P_i^{\alpha_i} = \mu_{u_{|\mathrm{Ker}(P_i^{\alpha_i}(u))}}$ divise $P$ ; si et seulement si $\mu_u$ divise $P$. On en déduit que $\mu_u$ divise $\mu_{u,x}$. CQFD.

# 4 Théorème de Jordan

## 4.1

Soit $E$ un $k$-espace vectoriel de dimension finie. Le théorème de Jordan est le suivant :

**Théorème I.13. Jordan**

Soit $u$ un endomorphisme de $E$ dont le polynôme caractéristique est **scindé**.
Alors, il existe une base de $E$ dans laquelle la matrice de $u$ est diagonale par blocs carrés du type

$$J_{\lambda,l} = \left.\begin{pmatrix} \lambda & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & \lambda \end{pmatrix}\right\} l,$$

où $\lambda \in k$.

De plus, l'ensemble (avec multiplicités) des couples $(\lambda, l)$ tels que $J_{\lambda,l}$ apparaît ainsi ne dépend que de $u$ et pas de la base.

Une matrice diagonale par blocs du type $J_{\lambda,l}$ est appelé *matrice de Jordan.*

## 4.2 Existence

Nous donnons ici les grandes lignes de la démonstration de l'existence :
Étape 1 : Réduction au cas nilpotent
Grâce au **théorème de décomposition des noyaux** et à celui de **Cayley-Hamilton** on montre que $E$ est la somme directe des sous-espaces caractéristiques de $u$. Cette décomposition est stable par $u$, il suffit donc de montrer l'existence dans le cas où $u$ a un seul sous-espace caractéristique c'est-à-dire $u = \lambda Id + n$, avec $n$ nilpotente. Il suffit alors, de traiter le cas où $u$ est nilpotente.

Étape 2 : Le cas nilpotent
La démonstration se fait par récurrence sur la dimension de $E$.
Supposons que $u^s = 0$ et $u^{s-1} \neq 0$. Soit $x$ dans $E$ tel que $u^{s-1}(x) \neq 0$. Alors, la famille $\mathcal{F} = (u^{s-1}(x), \cdots, u(x), x)$ est libre, et le sous-espace $F$ engendré par $\mathcal{F}$ est stable par $u$. De plus, la matrice de la restriction de $u$ à $F$ dans la base $\mathcal{F}$ est $J_{0,s}$.

Pour pouvoir appliquer l'hypothèse de récurrence, il suffit maintenant de trouver un supplémentaire à $F$ stable par $u$.
Comme $u^{s-1} \neq 0$, il existe $\varphi \in E^*$ tel que

$$\varphi(u^{s-1}(x)) \neq 0. \tag{4.1}$$

Soit $G'$ le sous-espace de $E^*$ engendré par les ${}^tu^i(\varphi)$. D'après (4.1), ${}^tu^{s-1}(\varphi) \neq 0$. De plus, $({}^tu)^s = {}^t(u^s) = 0$. On en déduit $({}^tu^{s-1}(\varphi), \cdots, \varphi)$ est une base de $G'$ et que $G'$ est stable par ${}^tu$.
Mais alors, l'orthogonal de $G'$ est stable par $u$ et a la dimension d'un supplémentaire de $F$. Il nous reste donc à montrer que cet orthogonal est en somme directe avec $F$.
Soit $y = \sum_{i=0}^{s-1} a_i u^i(x) \in F$. Supposons que $y$ est dans l'orthogonal de $G'$ et montrons que $y$ est nul. Si ce n'est pas le cas notons $i$ l'indice minimal tel que $a_i \neq 0$. On a : $0 = {}^tu^{s-1-i}(\varphi)(y) = a_i\varphi(u^{s-1}(x))$. Mais alors, la condition (4.1) montre que $a_i = 0$. Contradiction.

## 4.3 Unicité

La démonstration de l'unicité que nous esquissons ci-dessous est aussi un moyen de calculer une matrice de Jordan à laquelle est semblable une matrice donnée.

Supposons que $u$ a une matrice de Jordan $M$ dans une certaine base de $E$. On va expliquer que l'on peut décrire les blocs de $M$ en ne parlant que de $u$ : cela montrera bien l'unicité !
On remarque tout d'abord que l'ensemble des $\lambda$ qui apparaissent est le spectre de $u$. Fixons $\lambda$ dans le spectre de $u$ et considérons la suite : $d_0 = 0$, $d_i = \dim(\mathrm{Ker}(u - \lambda\mathrm{Id})^i)$. On vérifie que si $u$ est un $J_{\lambda,l}$ (dans une base) cette suite augmente de 1 en 1 jusqu'à atteindre la valeur $l$ puis stationne. On en déduit alors que $d_i - d_{i-1}$ est le nombre de blocs $J_{\lambda,l}$ de $M$ vérifiant $l \geqslant i$. Posons $\delta_i = d_i - d_{i-1}$. Notons $l_1 \geqslant l_2 \geqslant \cdots \geqslant l_r$, la suite ordonnée des tailles des blocs $J_{\lambda,l}$ de $M$. On remarque alors que $l_i$ est le nombre de $\delta_j$ supérieurs à $i$. Nous avons bien exprimé les $l$ tels que $J_{\lambda,l}$ apparaît dans $M$ en fonction des $\dim(\mathrm{Ker}(u - \lambda\mathrm{Id})^i)$.

## 4.4 Calcul de la matrice de Jordan d'un endomorphisme

Une question naturelle est maintenant : Étant donnée une matrice $A$ explicite, comment calculer une matrice de Jordan $J$ à laquelle elle est semblable et comment calculer une matrice de passage $P$ ?

Ces deux questions qui se ressemblent sont en fait très différentes. Pour le comprendre, on peut regarder le cas d'une matrice $A$ diagonalisable à valeurs propres distinctes. Dès que l'on a scindé le polynôme caractéristique de $A$, on sait à quelle matrice diagonale $A$ est semblable. Trouver une matrice de passage revient à trouver une base constituée de vecteurs propres et revient donc à résoudre autant de systèmes linéaires qu'il y a de valeurs propres.

Dans les deux cas, il faut commencer par calculer puis scinder le polynôme caractéristique de $A$. Pour chaque valeur propre $\lambda$, on calcule alors la suite des sous-espaces $\mathrm{Ker}(A - \lambda I_n)^i$.

Pour répondre à la première question, on a presque terminé puisque l'on peut alors facilement calculer les $\delta_i$ de la section précédente, puis les $l_i$ associés. Autrement dit, on refait de manière explicite la démonstration de l'unicité pour calculer $J$.

En revanche, pour trouver une matrice de passage, on refait explicitement la démonstration de l'existence.

**Exemple 4.** Calculer la réduite de Jordan et une base de Jordan pour

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

On vérifie que

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

$A^3 = 0$ et que le rang de $A$ vaut 3. Donc, la suite des dimensions des noyaux des puissances de $A$ vaut $0 < 2 < 4 < 5$. Les différences de telles dimensions successives valent $2 \geqslant 2 \geqslant 1$. On en déduit que $A$ est semblable à la matrice :

$$B := \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Cherchons maintenant une base dans laquelle l'endomorphisme associé à $A$ a $B$ pour matrice. On commence par choisir un vecteur qui n'est pas dans le noyau de $A^2$. **On prend un vecteur de la base canonique pour simplifier les calculs !** Posons

$$V_1 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \qquad V_2 = A(V_1) = \begin{pmatrix} 2 \\ 0 \\ 0 \\ 0 \\ -1 \end{pmatrix} \qquad V_3 = A^2(V_1) = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

On cherche maintenant un sous-espace stable par $A$ et supplémentaire à $\mathrm{Vect}(V_1, V_2, V_3)$. Soit $\varphi = e_1^*$ une forme linéaire simple (base canonique) telle que $\varphi(V_3) \neq 0$. On calcule les itérés de $\varphi$ par ${}^tA$ :

$$\varphi = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \qquad {}^tA(\varphi) = \begin{pmatrix} 0 \\ -1 \\ 2 \\ -2 \\ -1 \end{pmatrix} \qquad {}^t(A^2)(\varphi) = \begin{pmatrix} 0 \\ -1 \\ 1 \\ -1 \\ 0 \end{pmatrix}$$

Soit $G'$ le sous-espace engendré par ces trois formes linéaires. On cherche maintenant $V_4$ orthogonal à ces trois formes linéaires tel que $A^1(V_4) \neq 0$ (car on cherche maintenant un bloc de Jordan $2 \times 2$). L'orthogonal de $G'$ est

$$\left\{ \begin{pmatrix} 0 \\ x \\ y \\ y - x \\ x \end{pmatrix} \ : \text{ tels que } x,\, y \in \mathbb{R} \right\}.$$

Autrement dit,

$$\left( \begin{pmatrix} 0 \\ 1 \\ 0 \\ -1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \\ 0 \end{pmatrix} \right)$$

est une base de l'orthogonal de $G'$. On vérifie qu'en prenant pour $V_4$ le premier vecteur de la base ci-dessus, on a bien $A(V_4) \neq 0$. On calcule :

$$A(V_4) = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \\ 0 \end{pmatrix} =: V_5.$$

La matrice de $A$ dans la base $(V_3, V_2, V_1, V_5, V_4)$ est $B$. Autrement dit, la matrice

$$P := \begin{bmatrix} 1 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & -1 \\ 0 & -1 & 0 & 0 & 1 \end{bmatrix}$$

est une matrice de passage de $A$ à $B$, c'est-à-dire $P^{-1}AP = B$.

**Exercice 2.** *Calculer la réduite de Jordan puis une base de Jordan des matrices :*

$$M = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ -2 & -2 & 0 & 1 \\ -2 & 0 & -1 & -2 \end{bmatrix} \quad \text{et} \quad N := \begin{bmatrix} 6 & 1 & -2 & 1 \\ 8 & 5 & -2 & 3 \\ 11 & 0 & -4 & 1 \\ -22 & -11 & 6 & -7 \end{bmatrix}.$$

*Afin de limiter vos calculs, on vous indique que* $N^2 = 0$.

### 4.5 Applications

Les propriétés suivantes se montrent grâce au théorème de Jordan et sont difficiles sans celui-ci :

1. Toute matrice de $\mathcal{M}_n(\mathbb{C})$ est semblable à sa transposée. Le résultat est vrai sur $\mathbb{R}$ car deux matrices réelles semblables sur $\mathbb{C}$ sont semblables sur $\mathbb{R}$.

2. Toute matrice de $M_n(\mathbb{C})$ est semblable à une matrice symétrique. Sur $\mathbb{R}$, les matrices semblables à des matrices symétriques sont les matrices diagonalisables.

3. Soit $A \in \mathcal{M}_n(\mathbb{C})$. Posons $\mathrm{Com}(A) := \{B \in \mathcal{M}_n(\mathbb{C}) \ : \ AB = BA\}$. Posons $\mathrm{Com}(\mathrm{Com}(A)) := \{C \in \mathcal{M}_n(\mathbb{C}) \ : \ BC = CB \ \ \forall B \in \mathrm{Com}(A)\}$. Alors, $\mathrm{Com}(\mathrm{Com}(A))$ est l'ensemble des polynômes en $A$.

*Remarque.* Il va de soit que cette liste n'est pas exhaustive et ne demande qu'à être complétée.

# 5 Décomposition de Dunford

## 5.1 Enoncé : version 1

**Théorème I.14. Décomposition de Dunford : cas $\mu_u$ scindé**

Soit $u$ un endomorphisme dont le polynôme caractéristique est scindé.
Alors, il existe un unique couple $(d, n)$ d'endomorphismes tels que

1. $d$ est diagonalisable ;
2. $n$ est nilpotent ;
3. $d \circ n = n \circ d$ ;
4. $u = d + n$.

De plus, $d$ et $n$ sont des polynômes en $u$.

**Preuve**

**Avec le théorème Chinois.** L'existence est une conséquence du TDN. On écrit $\chi_u = \prod_{i=1}^{s}(X - \lambda_i)^{\alpha_i}$ avec $\alpha_i \in \mathbb{N}^*$ et $\lambda_i \in K$ 2 à 2 distincts. On pose $E_i = \mathrm{Ker}(X-\lambda_i)^{\alpha_i}(u)$. Par Cayley-Hamilton, $\chi_u(u) = 0$. Or les $(X - \lambda_i)^{\alpha_i}$ sont 2 à 2 premiers entre eux. Donc le TDN donne

$$E = \bigoplus_{i=1}^{s} E_i. \tag{5.1}$$

Comme $u$ commute à $(X - \lambda_i)^{\alpha_i}(u)$, $E_i$ est stable par $u$. Autrement dit, dans une base de $E$ obtenue en concaténant des bases des $E_i$, la matrice de $u$ est diagonale par blocs. Notons $u_i$ l'endomorphisme de $E_i$ induit par $u$. Autrement dit, le $i$-ème bloc diagonal.
Comme $E_i = \mathrm{Ker}(X - \lambda_i)^{\alpha_i}(u)$, $u_i - \lambda_i \mathrm{Id}_{E_i}$ est nilpotent d'ordre au plus $\alpha_i$. Soit maintenant $d_0$ l'endomorphisme de $E$ qui multiplie les vecteurs de $E_i$ par $\lambda_i$. Par construction $d_0$ est diagonalisable, et $n_0 := u - d_n$ est nilpotent. Le produit par bloc montre, en outre, que $n_0$ et $d_0$ commutent.
Nous venons de montrer l'existence sans le caractère polynomial.

**Un point important** à comprendre est que cette existence ne suffit pas à montrer l'unicité. En effet, dans la preuve de l'unicité ci-après, nous utilisons fortement l'existence de $d$ et $n$ dans $K[u]$.

On cherche donc à trouver un polynôme $P \in K[X]$ tel que

$$d_0 = P(u).$$

Comme $d_0$ et $P(u)$ stabilisent chaque $E_i$, il suffit de montrer que

$$\forall i = 1, \ldots, s \qquad \lambda_i \mathrm{Id}_{E_i} = P(u_i)$$

c'est-à-dire que

$$\forall i = 1, \ldots, s \qquad (P - \lambda_i)(u_i) = 0. \tag{5.2}$$

Le polynôme minimal de $u_i$ divise $(X - \lambda_i)^{\alpha_i}$. Il est donc égal à $(X - \lambda_i)^{\beta_i}$ pour $0 < \beta_i \leqslant \alpha_i$. Mais alors (5.2) est équivalent à

$$\forall i = 1, \ldots, s \qquad (X - \lambda_i)^{\beta_i} \text{ divise } P - \lambda_i, \tag{5.3}$$

ou encore à

$$\forall i = 1, \ldots, s \qquad P \equiv \lambda_i \quad [(X - \lambda_i)^{\beta_i}]. \tag{5.4}$$

On reconnait en (5.4) un système de congruence. Comme les $(X - \lambda_i)^{\beta_i}$ sont deux à deux premiers entre eux, le théorème Chinois assure l'existe de solution $P_0$. Ainsi $d_0 = P_0(u)$ est un polynôme en $u$. Mais alors, $n_0 = (X - P_0)(u)$ est aussi un polynôme en $u$.

Montrons l'unicité. Soit $(n, d)$ une autre solution. On a $u = n + d = n_0 + d_0$ donc $n - n_0 = d_0 - d$. Comme $d$ commute à $u$ et $d_0$ est un polynôme en $u$, $d$ et $d_0$ commutent. Mais alors, ils sont simultanément diagonalisables. En particulier $d_0 - d$ est diagonalisable.
De même, $n$ et $n_0$ commutent. Donc

$$(n - n_0)^{2N} = \sum_{k=0}^{2N} (-1)^k \binom{2N}{k} n^k n_0 2N - k.$$

Pour tout $0 \leqslant k \leqslant 2N$, $k \geqslant N$ ou $2N - K \geqslant N$. Donc $n^k n_0 2N - k = 0$. Donc $n - n_0$ est nilpotente. On vient de montrer que l'endomorphisme $n - n_0 = d_0 - d$ est à la fois diagonalisable et nilpotent. Il est donc nul.

## 5.2 Avec Newton

On se pose la question suivante :

Peut-on calculer $d$ et $n$ sans connaître les valeurs propres de $u$ ?

La démonstration ci-dessus suggère une réponse négative. Mais, il s'avère que la réponse est oui ! En fait, une adaptation de la méthode de Newton pour l'approximation des racines conduit à une démonstration effective (i.e. transformable, en un algorithme qui a le bon gout d'être très rapide). Cette preuve est dûe à Claude Chevalley (1909 – 1984).

**Rappel : Algorithme de Newton.** L'algorithme de Newton est un algorithme itératif sensé converger vers une solution d'une équation du type $f(x) = 0$ avec $f$ dérivable et de dérivé jamais nulle. On définit une suite par récurence par la fomule :

$$x_{n+1} = x_n - f(x_n)/f'(x_n).$$

Dans les bons cas $x_n$ converge ; la limite est alors une (la) solution de $f(x) = 0$.

**Preuve de l'existence forte**

On va exhiber une « équation » satisfaite par $d$ et vérifier que l'algorithme de Newton converge pour cette équation.
Le point de départ est donc la constation que les racines du polynôme minimal de $d$ sont simples et sont les valeurs propres de $u$. Soit $\chi_u = \prod_{i=1}^{s} (X - \lambda_i)^{\alpha_i}$ le polynôme carctéristique de $u$ et posons

$$Q = \prod_{i=1}^{s} (X - \lambda_i). \tag{5.5}$$

On aura

$$Q(d) = 0.$$

Considérons donc la suite d'endomorphismes suivante :

$$\begin{aligned} u_0 &= u, \\ u_{n+1} &= u_n - Q(u_n) * Q'(u_n)^{-1}. \end{aligned}$$

Remarquons qu'il faut justifier que $Q'(u_n)$ est bien inversible. On montre en fait par une même récurence sur $n$ les trois propriétés suivantes :

1. $u_n \in k[u]$ ;
2. $Q'(u_n)$ est inversible pour tout $n$ ; et
3. $Q(u_n) \in Q(u)^{2^n} k[u]$, en particulier $Q(u_n)$ est nilpotent.

Pour $n = 0$, $Q'$ est premier avec le polynôme minimal de $u$ donc $Q'(u)$ est inversible. Supposons l'hypothèse satisfaite pour $u_n$ : ceci implique que $u_{n+1}$ est bien défini. Comme l'inverse d'un endomorphisme inversible $v$ appartient à $k[v]$, $u_{n+1} \in k[u]$. D'après le lemme I.15 appliqué dans l'anneau $k[u]$, il suffit de montrer que $Q'(u_{n+1}) - Q'(u_n)$ est nilpotent pour avoir l'inversibilité de $Q'(u_{n+1})$. Or le lemme I.16 implique que $Q'(u_{n+1}) - Q'(u_n) \in (u_{n+1} - u_n)k[u]$. Or, la formule de récurence montre que $(u_{n+1} - u_n)k[u] \subset Q(u_n)k[u]$. Par récurence, $Q(u_n)$ est nilpotent ; et l'inversibilité de $Q'(u_{n+1})$ suit.
Il reste à montrer que $Q(u_{n+1}) \in Q(u)^{2^{n+1}} k[u]$. Pour cela écrivons

$$Q(u_{n+1}) = Q(u_n + y),$$

avec $y = -\frac{Q(u_n)}{Q'(u_n)}$. Le lemme I.17 montre alors qu'il existe $v \in k[u]$ tel que

$$Q(u_{n+1}) = Q(u_n) + yQ'(u_n) + y^2 v.$$

Or $Q(u_n) + yQ'(u_n) = 0$ et $y^2 \in Q(u_n)^2 k[u]$. Ainsi

$$Q(u_{n+1}) \in Q(u_n)^2 k[u] \subset Q(u)^{2^{n+1}} k[u].$$

La dernière assertion de la récurrence est démontrée.

La dernière assertion montre que $Q(u_n) = 0$ pour $n$ asser grand. Remarquons que $n \geqslant \ln_2(\dim E)$ suffit. Ainsi, à partir de cette valeur $u_n$ stationne.

Notons $d$ l'endomorphisme sur lequel $u_n$ stationne. On a $Q(d) = 0$ ; donc $d$ est diagonalisable.

Il nous reste à voir que $u - d$ est nilpotente. Or, d'après ce qui précède, $u - d = (u_0 - u_1) + (u_1 - u_2) + \cdots + (u_{n-1} - u_n)$ (si $d = u_n$) appartient à $Q(u).k[u]$. On en déduit qu'il est nilpotent.

Voici les énoncés des lemmes que nous avons utilisés :

**Lemme I.15**

Soit $A$ un anneau commutatif unitaire. Soit $a$ un élément inversible de $A$ et $n$ un élément nilpotent. Alors, $a + n$ est inversible.

**Preuve**

La démonstration s'inspire de la formule :

$$\frac{1}{1-x} = 1 + x + x^2 + x^3 + \dots$$

Ecrivons $a + n = a(1 - y)$ avec $y = -a^{-1}n$. Remarquons que $y$ est nilpotent et fixons un entier

naturel $k$ tel que $y^k = 0$. On vérifie alors que

$$(1-y) * (1 + y + y^2 + \cdots + y^{k-1}) = 1 - y^k = 1.$$

Alors $1 - y$ et $a + n$ sont inversibles.

**Lemme I.16**

Pour tout polynôme $R \in K[X]$ il existe un polynôme en deux variables $S$ tel que :

$$R(Y) - R(X) = (Y - X)S(X, Y).$$

Ici $X$ et $Y$ sont deux indéterminées.

**Preuve**

L'ensemble des polynômes $R$ qui vérifient le lemme est un sous-espace vectoriel de l'espace des polynômes. Or la formule

$$Y^d - X^d = (Y - X)(Y^{d-1} + Y^{d-2}X + \cdots + X^{d-1}),$$

montre que $R = X^d$ appartient à cet ev. Le lemme en découle.

**Lemme I.17**

Pour tout polynôme $R \in K[X]$ il existe un polynôme en deux variables $S$ tel que :

$$R(X + Y) = R(X) + YR'(X) + Y^2S(X, Y).$$

**Preuve**

Même preuve que le lemme précédent avec la formule

$$(X + Y)^d = X^d + YdX^{d-1} + Y^2(\cdots).$$

**Corollaire I.18**

Soit $K \subset L$ deux corps. Soit $M \in M_n(K)$. On suppose que $\chi_M$ est scindé sur $L$. Soit $M = D + N$ sa décomposition de Dunford sur $L$.
Alors $D$ et $N$ sont dans $\mathcal{M}_n(K)$.

**Preuve**

Considérons le polynôme caractéristique $\chi_M$ de $M$ et sa factorisation

$$\chi_M = \prod_{i=1}^{s} (X - \lambda_i)^{\alpha_i}.$$

En vertu du lemme **??**, le polynôme $Q = \prod_{i=1}^{s}(X - \lambda_i)$ est à coefficients dans $K$.

Mais alors, la preuve avec l'algorithme de Newton montre que $D \in \mathcal{M}_n(K)$. Donc $N = M - D \in \mathcal{M}_n(K)$.

**Lemme I.19**

Soit $K \subset L$ deux corps. On suppose que $K$ est de caractéristique nulle ou que est fini.
Soit

$$P = \prod_{i=1}^{s} (X - \lambda_i)^{\alpha_i}$$

un polynôme scindé sur $L$. On suppose les $\lambda_i \in L$ 2 à 2 distincts et $\alpha_i \in \mathbb{N}^*$.
Alors le polynôme $Q = \prod_{i=1}^{s} (X - \lambda_i)$ est à coefficients dans $K$.

**Preuve**

En caractéristique zéro, cela vient de la formule

$$Q = \frac{\chi_u}{\chi_u \wedge \chi_u'}.$$

En effet, le pgcd divise $P$ et est donc de la forme $\prod_{i=1}^{s} (X - \lambda_i)^{\beta_i}$ avec $0 \leqslant \beta_i \leqslant \alpha_i$. Par ailleurs,

$$P' = \sum_{i=1}^{s} \alpha_i \frac{\chi_u}{X - \lambda_i} \tag{5.6}$$

et la plus grande puissance de $(X - \lambda_i)$ qui divise $P'$ est $(X - \lambda_i)^{\alpha_i - 1}$.

Supposons maintenant que $K$ est fini de caractéristique $p > 0$. Quitte à remplacer $L$ par le corps engendré par $K$ et les $\lambda_i$, on peut supposer que $L$ est aussi fini.
Considérons le Frobenus $\mathrm{Fr} : L \longrightarrow L, x \longmapsto x^p$. C'est un morphisme de corps donc injectif. Par finitude Fr est en fait bijectif. De même sa restriction $\mathrm{Fr} : K \longrightarrow K$ est bijectif.
On fait un récurrence sur le maximum des $\alpha_i$. Si celui-ci vaut 1, il n'y a rien à montrer.
Considérons d'abord le cas où $P' = 0$. En écrivant $P = \sum a_i X^i$, on voit immédiatement que l'hypothèse $P' = 0$ signifie que $P$ est un polynôme en $X^p$ : $P(X) = U(X^p)$. Ecrivons $R = \sum_i a_i X^i$, avec $a_i \in K$. Comme Fr est sujectif, il existe $b_i \in K$ tels que $b_i^p = a_i$. Posons $V = \sum_i b_i X^i$. Comme Fr est un morphisme de corps, on a $U^p = R(X^p) = P(X)$. En particulier, $U$ et $P$ ont les même racines. L'hypoth !se de récurrence appliquée à $U$ montre que $Q \in K[X]$.

Supposons à présent que $P' \neq 0$. La formule (**??**) est encore valable et montre qu'au moins un des $\alpha_i$ (disons $\alpha_1$ quitte à renuméroter) est non nul.
Le calcul fait en caractéristique nulle s'adapte sans peine et montre que $P \wedge P' = \prod_{i=1}^{s} (X - \lambda_i)^{\beta_i}$ avec

$$\beta_i = \begin{cases} \alpha_i - 1 & \text{si } p \text{ ne divise pas } \alpha_i \\ \alpha_i & \text{sinon.} \end{cases}$$

Mais alors

$$S_1 := \frac{\chi_u}{\chi_u \wedge \chi_u'} = \prod_{p \nmid \alpha_i} (X - \lambda_i) \in K[X].$$

Si $n$ est le degré de $P$, on a de plus

$$T_1 := S_1^n \wedge \chi_u = \prod_{p \nmid \alpha_i} (X - \lambda_i)^{\alpha_i} \in K[X]$$

et

$$T_2 = \frac{P}{T_1} = \prod_{p \mid \alpha_i} (X - \lambda_i)^{\alpha_i} \in K[X].$$

Appliquant le premier cas à $S_2$ (ou l'hypothèse de récurrence) on obtient que

$$S_2 = \prod_{p \mid \alpha_i} (X - \lambda_i) \in K[X].$$

Cela conclut puisque $Q = S_1 S_2$.

*Remarques.*

1. Dans le démonstration précédente on a seulement utilisé que Fr : $K \longrightarrow K$ est surjectif. Dans ce cas, on dit que le corps $K$ est parfait. C'est l'hypothèse minimale pour avoir Dunford.

2. L'endomorphisme de Frobenius n'est pas toujours un automorphisme. Il est tooujours injectif mais par forcément surjectif. Prenons par exemple $K = \mathbb{F}_p(t)$. Alors l'image du Frobenius est $\mathbb{F}_p(t^p)$.

   Le lemme est faux avec ce corps. Par exemple, sur $K = \mathbb{F}_2(t)$, le polynôme $P = X^2 + t = (X + \sqrt{t})^2$ et $X + \sqrt{t} \notin K[X]$.

   Ici $\sqrt{t}$ est un symbole formel représentant la classe de $X$ dans le quotient $K[X]/P$ et vérifie donc $(\sqrt{t})^2 = t$.

## 5.3 Implémentation

La preuve de Chevalley est parfaitement constructive pour peu que nous sachions calculer le polynôme $Q$ défini par la formule (5.5). La démonstration du lemme I.19 étant constructive on peut calculer $Q$ grâce à la division euclidienne et l'algorithme d'euclide.

Il s'agit ici d'implémenter l'algorithme. Voici les étapes de l'algo :

1. Calculer $P_u$, puis $Q$ et $Q'$.
2. Calculer $u_n$ pour $2^n \leqslant \dim(E)$.
3. Conclure.

Voici une implémentation en SageMath.

```
def dunford(A):
    p=A.charpoly(x);
    q=p//(p.gcd(derivative(p)))
    qder=derivative(p)
    An=A;
    B=q(An)
    while B!=0 :
        An=An-B*qder(An)^(-1)
        B=q(An)
    return An,A-An
```

Voici un exemple $A$

$$\begin{pmatrix}
-239 & -219 & -201 & -182 & -164 & -149 & -135 & -120 & -105 & -90 & -75 & -60 & -45 & -30 & -15 \\
22 & 21 & 21 & 17 & 11 & 10 & 10 & 8 & 7 & 6 & 5 & 4 & 3 & 2 & 1 \\
612 & 560 & 518 & 478 & 440 & 400 & 360 & 321 & 280 & 240 & 200 & 160 & 120 & 80 & 40 \\
-416 & -392 & -379 & -356 & -330 & -300 & -270 & -240 & -209 & -180 & -150 & -120 & -90 & -60 & -30 \\
183 & 194 & 201 & 190 & 176 & 150 & 125 & 102 & 82 & 67 & 55 & 44 & 33 & 22 & 11 \\
-326 & -328 & -324 & -311 & -297 & -275 & -234 & -197 & -162 & -132 & -109 & -88 & -66 & -44 & -22 \\
17 & 17 & 17 & 17 & 17 & 17 & 16 & 16 & 12 & 6 & 5 & 5 & 3 & 2 & 1 \\
412 & 412 & 412 & 412 & 412 & 412 & 360 & 318 & 278 & 240 & 200 & 160 & 121 & 80 & 40 \\
-266 & -266 & -266 & -266 & -266 & -266 & -242 & -229 & -206 & -180 & -150 & -120 & -90 & -59 & -30 \\
128 & 128 & 128 & 128 & 128 & 128 & 139 & 146 & 135 & 121 & 95 & 70 & 47 & 27 & 12 \\
-217 & -217 & -217 & -217 & -217 & -217 & -219 & -215 & -202 & -188 & -166 & -125 & -88 & -53 & -23 \\
12 & 12 & 12 & 12 & 12 & 12 & 12 & 12 & 12 & 12 & 12 & 11 & 11 & 7 & 1 \\
212 & 212 & 212 & 212 & 212 & 212 & 212 & 212 & 212 & 212 & 212 & 160 & 118 & 78 & 40 \\
-116 & -116 & -116 & -116 & -116 & -116 & -116 & -116 & -116 & -116 & -116 & -92 & -79 & -56 & -30 \\
33 & 33 & 33 & 33 & 33 & 33 & 33 & 33 & 33 & 33 & 33 & 44 & 51 & 40 & 26
\end{pmatrix}.$$

Alors $D = A_2$ est

$$\begin{pmatrix}
-240 & -220 & -202 & -183 & -165 & -150 & -135 & -120 & -105 & -90 & -75 & -60 & -45 & -30 & -15 \\
22 & 21 & 21 & 17 & 11 & 10 & 9 & 8 & 7 & 6 & 5 & 4 & 3 & 2 & 1 \\
612 & 560 & 518 & 478 & 440 & 400 & 360 & 320 & 280 & 240 & 200 & 160 & 120 & 80 & 40 \\
-416 & -392 & -379 & -356 & -330 & -300 & -270 & -240 & -210 & -180 & -150 & -120 & -90 & -60 & -30 \\
183 & 194 & 201 & 190 & 176 & 150 & 125 & 102 & 82 & 66 & 55 & 44 & 33 & 22 & 11 \\
-326 & -328 & -324 & -311 & -297 & -275 & -234 & -197 & -162 & -132 & -110 & -88 & -66 & -44 & -22 \\
17 & 17 & 17 & 17 & 17 & 17 & 16 & 16 & 12 & 6 & 5 & 4 & 3 & 2 & 1 \\
412 & 412 & 412 & 412 & 412 & 412 & 360 & 318 & 278 & 240 & 200 & 160 & 120 & 80 & 40 \\
-266 & -266 & -266 & -266 & -266 & -266 & -242 & -229 & -206 & -180 & -150 & -120 & -90 & -60 & -30 \\
128 & 128 & 128 & 128 & 128 & 128 & 139 & 146 & 135 & 121 & 95 & 70 & 47 & 27 & 11 \\
-216 & -216 & -216 & -216 & -216 & -216 & -218 & -214 & -201 & -187 & -165 & -124 & -87 & -52 & -22 \\
12 & 12 & 12 & 12 & 12 & 12 & 12 & 12 & 12 & 12 & 12 & 11 & 11 & 7 & 1 \\
212 & 212 & 212 & 212 & 212 & 212 & 212 & 212 & 212 & 212 & 212 & 160 & 118 & 78 & 40 \\
-116 & -116 & -116 & -116 & -116 & -116 & -116 & -116 & -116 & -116 & -116 & -92 & -79 & -56 & -30 \\
33 & 33 & 33 & 33 & 33 & 33 & 33 & 33 & 33 & 33 & 33 & 44 & 51 & 40 & 26
\end{pmatrix}$$

$N$ est

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

## 5.4 Sans l'hypothèse scindé

### 5.4.1 Endomorphismes semi-simples

On veut capturer l'hypothèse diagonalisable sur une extension en restant dans le corps de base. La première idée est, que sur $\mathbb{C}$, un endomorphisme est diagonalisable si et seulement si son polynôme minimal est à racines simples, c'est-à-dire sans facteur carré.

**Proposition I.20: Endomorphismes Semi-Simples**

Soit $k$ un corps, $E$ un $k$-ev dimension finie et $u$ un endomorphisme de $E$. Alors, se valent

1. le polynôme minimal $\mu_u$ de $u$ est sans facteur carrré.
2. Tout sous-espace vectoriel $u$-stable de $E$ admet un supplémentaire stable.

**Preuve**

Commençons par 2 implique 1. Supposons par l'absurde que $\mu_u = P^2Q$, pour deux polynômes $P$ et $Q$. Notons $F$, le noyau de $P(u)$ et $G$ un supplémentaire stable.
On remarque que $P(u)\Big(PQ(u)(G)\Big) = \{0\}$, donc $PQ(u)(G) \subset F$. Or $PQ(u)(G) \subset G$. Donc $PQ(u)(G) = \{0\}$.
Comme, de plus, $QP(u)(F) = \{0\}$ ; on en déduit que $PQ$ annule $u$. Contradiction.

Supposons 1 et soit $F$ un sous-espace stable. Ecrivons $\mu_u = Q_1 \dots Q_s$. Le TDN donne $\oplus_i \mathrm{Ker} Q_i(u)$. Le TDN à la restriction de $u$ à $F$ donne $F = \oplus_i(\mathrm{Ker} Q_i(u)) \cap F$.
On peut donc supposer $\mu_u$ irréductible. Soit $L = k[X]/(\mu_u)$ qui est un corps. On remarque que $E$ est un $L$-espace vectoriel $P.v = P(u)v$. Alors un sous-$k$-espace de $V$ est stable si et seulement s'il est un $L$-sous-espace. L'existence de supplémentaires pour les $L$-espaces vectoriels permet de conclure.

### 5.4.2 Dunford version 2

**Théorème I.21. Décomposition de Dunford : cas général**

Soit $k$ un corps parfait (par exemple fini ou de caractéristique nulle). Soit $u$ un endomorphisme sur un $k$-espace vectoriel de dimension fini.
Alors, il existe un unique couple $(d, n)$ d'endomorphismes tels que

1. $d$ est semi-simple ;
2. $n$ est nilpotent ;
3. $d \circ n = n \circ d$ ;
4. $u = d + n$.

De plus, $d$ et $n$ sont des polynômes en $u$.

Le résultat découle directement du théorème I.14, du corollaire I.18 et de la proposition I.20.

# 6 Endomorphismes cycliques

Soit $P = a_0 + a_1 X + \cdots + a_{n-1} X^{n-1} + X^n$ un polynôme unitaire de degré $n$. On appelle *matrice compagnon de $P$* la matrice $n \times n$ :

$$C_P = \begin{pmatrix} 0 & 0 & \cdots & 0 & -c_0 \\ 1 & 0 & \cdots & 0 & -c_1 \\ & & \ddots & \vdots & \\ 0 & & & & \\ 0 & 0 & \cdots & & -c_{n-1} \end{pmatrix}.$$

**Théorème I.22. Endomorphisme cyclique**

Soit $u$ un endormorphisme de $E$, de dimension $n$. Se valent :

1. il existe $x \in E$ tel que $(x, u(x), \ldots, u^{n-1}(x))$ est une base ;
2. $\deg(\mu_u) = n$ ;
3. $\mu_u = \pm\chi_u$ ;
4. $\mathrm{Com}(u) = K[u]$ ;
5. $\dim(\mathrm{Com}(u)) = n$.

On dit alors que $u$ est cyclique.

**Preuve**

Voir TD.

# Chapitre 2

# Algèbre Bilinéaire

## Sommaire

# 1 Formes bilinéaires et quadratiques et leurs matrices

Soit $E$ un $K$-espace vectoriel de dimension finie. Une forme bilinéaire symétrique $\varphi$ sur $E$ est une application $\varphi : E \times E \longrightarrow K$ telle que pour tout $x, y, z \in E$ et $\lambda \in K$, on a

1. $\varphi(x + \lambda y, z) = \varphi(x, z) + \lambda\varphi(y, z)$ ;
2. $\varphi(x, y) = \varphi(y, x)$.

A une forme bilinéaire on associe

$$\begin{array}{rccc} \tilde{\varphi} : & E & \longrightarrow & E^* \\ & y & \longrightarrow & \left( \begin{array}{ccc} E & \longrightarrow & K \\ x & \longmapsto & \varphi(x, y) \end{array} \right). \end{array}$$

On vérifie aisément que $\tilde{\varphi}$ est bien définie et linéaire en utilisant le premier axiome (et son analogue en deuxième position).
La transposé de $\tilde{\varphi}$ est une application linéaire de $E^{**}$ dans $E^*$. On vérifie que le second axiome signifie que

$$^t\tilde{\varphi} = \tilde{\varphi} \circ \iota. \tag{1.1}$$

Autrement dit, $\varphi$ est symétrique si $\tilde{\varphi}$ est égale à sa transposé. La nature est bien faite !
On définit alors

$$\mathrm{rang}(\varphi) := \mathrm{rang}(\tilde{\varphi})$$

et

$$\mathrm{Ker}(\varphi) := \mathrm{Ker}(\tilde{\varphi}) = \{y \in E \,|\, \forall x \in E \quad \varphi(x, y) = 0\}.$$

Le théorème du rang appliqué à $\tilde{\varphi}$ donne

$$\dim \mathrm{Ker}(\varphi) + \mathrm{rang}(\varphi) = \dim(E).$$

On dit que la forme bilinéaire est *non dégénérée* si son rang vaut $\dim(E)$, autrement dit si $\tilde{\varphi}$ est un isomorphisme d'espaces vectoriels.
Soit maintenant une base $\mathcal{B}$ de $E$. On pose

$$\mathrm{Mat}_{\mathcal{B}}(\varphi) := \mathrm{Mat}_{\mathcal{B}^*\mathcal{B}}(\tilde{\varphi}).$$

Si $\mathcal{B} = (e_1, \ldots, e_n)$ on remarque que le coefficient $(i, j)$ de la matrice est $\varphi(e_i, e_j)$.

**Changement de bases.** Soit $\mathcal{C}$ une seconde base $E$. En écrivant $\tilde{\varphi} = \mathrm{Id}_{E^*} \circ \tilde{\varphi} \circ \mathrm{Id}_E$, on obtient

$$\mathrm{Mat}_{\mathcal{C}}(\varphi) = \mathrm{Mat}_{\mathcal{C}^*\mathcal{C}}(\tilde{\varphi}) = \mathrm{Mat}_{\mathcal{C}^*\mathcal{B}^*}(\mathrm{Id}_E)\mathrm{Mat}_{\mathcal{B}^*\mathcal{B}}(\tilde{\varphi})\mathrm{Mat}_{\mathcal{B}\mathcal{C}}(\mathrm{Id}_E).$$

Or $^t\mathrm{Id}_{E^*} = \mathrm{Id}_E$, donc en posant $P = \mathrm{Mat}_{\mathcal{B}\mathcal{C}}(\mathrm{Id}_E)$, $M = \mathrm{Mat}_{\mathcal{B}}(\varphi)$ et $N = \mathrm{Mat}_{\mathcal{C}}(\varphi)$, on obtient

$$N = {}^tPMP.$$

Le groupe $\mathrm{GL}_n(K)$ agit sur l'espace $S_n(K)$ des matrices carrées de taille $n$ par

$$P.S := PS\,{}^tP.$$

Deux matrices d'une même orbite sont dîtes *congruentes*. On vient donc de voir que deux matrices sont congruentes si et seulement elles représentent la même forme bilinéaire dans deux bases.

**Orthogonalité.** On dit que deux vecteurs $x$ et $y$ de $E$ sont orthogonaux si $\varphi(x, y) = 0$. On dit que $x$ est isotrope si $\varphi(x, x) = 0$.
On suppose ici que $\varphi$ est **non dégénérée**. Comme $\tilde{\varphi}$ est un isomorphisme, on identifie $E$ et $E^*$. Alors, on peut transporter les sev orthogonaux au sens de la dualité dans $E$. Si $F$ est un sous-espace vectoriel de $E$, on définit

$$F^{\perp_\varphi} := \{x \in E \,:\, \varphi(x, y) = 0 \quad \forall y \in F\}.$$

Les propriétés (dimension, bi-orthogonal, somme, intersection) de la section 2.4 restent trivialement vraies.
Une nouveauté : on peut comparer $F$ et son orthogonal.

**Exercice 3.** *Soit $E$ un espace vectoriel réel de dimension $n \geqslant 2$ et $1/leqd \leqslant n-1$.*

1. *Soit $F$ et $G$ deux sous-espaces vectoriels de $E$ de dimension $d$ et $n-d$ respectivement. Montrer que* $\dim(F \cap G) \leqslant \min(d, n-d)$.
2. *Soit $0 \leqslant k \leqslant \min(d, n-d)$. Montrer qu'il existe $F$ et $G$ comme ci-dessus tels que* $\dim(F \cap G) = k$.
3. *Montrer qu'il existe une forme bilinéaire non dégénéré $\varphi$ sur $E$ et un sous-espace vectoriel $F$ de $E$ de dimension $d$ tel que* $\dim(F \cap F^{\perp_\varphi}) = k$.

# 2 Forme quadratique et Réduction de Gauss

## 2.1 Formes quadratiques

On suppose ici que la caractéristique du corps n'est pas 2. Une *forme quadratique* sur $E$ est une fonction $Q$ de $E$ dans $K$ qui s'exprime comme un polynôme homogène de degré 2 (dans une/toute base).
Les formules suivantes, appelées formules de polarisation donnent une correspondance entre formes quadratiques et formes bilinéaires symétriques.

$$\begin{gathered} Q(x) = \varphi(x,x) \\ \varphi(x,y) = \tfrac{1}{2}(Q(x+y) - Q(x) - Q(y)) \\ \varphi(x,y) = \tfrac{1}{4}(Q(x+y) - Q(x-y)) \end{gathered}$$

## 2.2 Algorithme de Gauss

Soit $E$ un $K$-espace vectoriel de dimension finie $n$, et soit $\mathcal{B} = (e_1, \ldots, e_n)$ une base de $V$. Soit $q : E \to K$ une forme quadratique, et soit $M = (a_{ij})_{1 \leqslant i,j \leqslant n}$ sa matrice représentative dans la base $\mathcal{B}$. Si $x = \sum_{i=1}^n x_i e_i$, on a donc

$$q(x) = \sum_{i=1}^{n} a_{ii} x_i^2 + 2 \sum_{1 \leqslant i < j \leqslant n} a_{ij} x_i x_j = P(x_1, \ldots, x_n).$$

On procède par récurrence sur le nombre de variables. A chaque étape, il y a deux cas.

**Avec des carrés.** S'il existe un indice $k$ tel que $a_{kk} \neq 0$, on regroupe tous les termes faisant intervenir la variable $x_k$, et on complète le carré. On écrit

$$P(x_1, \ldots, x_n) = a_{kk} x_k^2 + 2 f_k x_k + P_0,$$

où $f_k$ est une forme linéaire en les variables $x_i, i \neq k$, et $P_0$ est une forme quadratique en les variables $x_i, i \neq k$.
On a alors

$$\begin{aligned} P(x_1, \ldots, x_n) &= a_{kk}(x_k^2 + \tfrac{2}{a_{kk}} f_k x_k) + P_0 \\ &= a_{kk}((x_k + \tfrac{f_k}{a_{kk}})^2 - \tfrac{f_k^2}{a_{kk}^2}) + P_0 \end{aligned} .$$

On peut donc écrire

$$P(x_1, \ldots, x_n) = a_{kk}(x_k + \frac{f_k}{a_{kk}})^2 + P_1,$$

où $P_1$ est une forme quadratique en les variables $x_i, i \neq k$.

**Sans carré.** Si $a_{kk} = 0$ pour tout $k$, mais qu'il existe $k$ et $\ell$ tels que $k < \ell$ et $a_{k\ell} \neq 0$. C'est le cas délicat.
On écrit

$$P(x_1, \ldots, x_n) = 2a_{k\ell} x_k x_\ell + 2 f_k x_k + 2 f_\ell x_\ell + P_0,$$

où $f_k$ et $f_\ell$ sont des formes linéaires en les variables $x_i, (i \neq k, \ell)$, et $P_0$ est une forme quadratique en les variables $x_i, (i \neq k, \ell)$.
On a ainsi

$$P(x_1, \ldots, x_n) = 2a_{k\ell}(x_k + \frac{1}{a_{k\ell}} f_\ell)(x_\ell + \frac{1}{a_{k\ell}} f_k) - \frac{2}{a_{k\ell}} f_k f_\ell + P_0.$$

On a donc

$$P(x_1, \ldots, x_n) = 2a_{k\ell} AB + P_1,$$

avec $A = x_k + \frac{1}{a_{k\ell}} f_\ell, B = x_\ell + \frac{1}{a_{k\ell}} f_k$, et $P_1$ est une forme quadratique en les variables $x_i, i \neq k, \ell$. On a alors

$$P(x_1, \ldots, x_n) = \frac{a_{k\ell}}{2}((A+B)^2 - (A-B)^2) + P_1.$$

Si $P_1 = 0$, on arrête. Sinon, on recommence le procédé avec $P_1$.

On peut montrer que l'on obtient alors une écriture de la forme

$$q(x) = \alpha_1(L_1(x))^2 + \ldots + \alpha_r(L_r(x))^2,$$

où :

1. chaque $\alpha_i \in \mathbb{R}^*$
2. chaque $L_i$ est une forme linéaire sur $V$
3. la famille de formes $(L_1, \ldots, L_r)$ est libre.

Si $q$ n'est pas de rang $n$ ($r \neq n$), on complète par des formes linéaires $L_{r+1}, L_{r+2}, \ldots, L_n$ (on les choisit par exemple parmi les formes coordonnées $x_1, ..., x_n$) pour que la famille $(L_1, \ldots, L_n)$ soit libre et on écrit

$$q(x) = \alpha_1(L_1(x))^2 + \ldots + \alpha_r(L_r(x))^2 + 0(L_{r+1})^2 + \ldots + 0(L_n(x)))^2, \tag{2.1}$$

et on pose $\alpha_{r+1} = \cdots = \alpha_n = 0$.

**Calcul de la base $q$-orthogonale**
On considère la base anté-duale $\mathcal{C} = (f_1, \ldots, f_n)$ de la base $(L_1, \ldots, L_n)$. Alors, l'expression (2.1) devient

$$q = \sum_i \alpha_i f_i^*$$

si bien que $\mathrm{Mat}_{\mathcal{C}}(q) = \mathrm{diag}(\alpha_1, \ldots, \alpha_n)$.
L'algorithme démontre donc le résultat suivant.

**Théorème II.23. Réduction de Gauss**

1. Soit $Q$ une forme quadratique et $\varphi$ la forme bilinéaire associée. Alors il existe une base $\mathcal{C}$ telle que $\mathrm{Mat}_{\mathcal{C}}(\varphi)$ est diagonale.
2. Soit $S$ une matrice symétrique. Il existe une matrice inversible $P$ telle que ${}^tPMP$ est diagonale.

**Calcul explicite.** Comment calculer $\mathcal{C}$ ?
L'algorithme fournit les formes linéaires $L_i$. On note $\mathcal{C}^* = (L_1, \ldots, L_n)$. On collecte cette information dans la matrice $P = \mathrm{Mat}_{\mathcal{B}^*\mathcal{C}^*}(\mathrm{Id}_{E^*})$ : ses colonnes sont les coordonnées de $L_j$.
Alors ${}^tP^{-1} = \mathrm{Mat}_{\mathcal{B}\mathcal{C}}(\mathrm{Id}_E)$ : ses colonnes sont les coordonnées de $f_j$ dans la base $\mathcal{B}$.

**Exemple 1 :**
On considère la forme quadratique $q$ définie sur $\mathbb{R}^2$ par

$$q(x, y) = x^2 + 4xy$$

On élimine la variable $x$ en formant un carré contenant tous les termes dépendant de $x$ (forme canonique d'un polynôme du second degré en $x$ dépendant de $y$ vu comme paramètre)

$$q(x, y) = (x + 2y)^2 - 4y^2 = x'^2 - 4y'^2, \quad x' = x + 2y, y' = y$$

Pour trouver la base $q$-orthogonale, il suffit de chercher son premier vecteur de base $x' = 1, y' = 0$ donc $y = 0$ puis $x = 1$, puis son deuxième vecteur de base $x' = 0, y' = 1$ donc $y = 1$ puis $x = -2y = -2$. La matrice de passage de la base canonique à la base $q$-orthogonale est donc

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

on peut vérifier $M := [[1,2],[2,0]]; P := [[1,-2],[0,1]]; tran(P) * M * P$.

**Exemple 2**
On considère la forme quadratique $q$ définie sur $\mathbb{R}^3$ par

$$q(x,y,z) = x^2 + 2xy + 4xz + 2yz$$

On élimine la variable $x$

$$q(x,y,z) = (x+y+2z)^2 - (y+2z)^2 + 2yz = (x+y+2z)^2 - y^2 - 4z^2 - 2yz$$

Puis on élimine $y$ dans ce qui reste

$$q(x,y,z) = (x+y+2z)^2 - (y+z)^2 - 3z^2 = x'^2 - y'^2 - 3z'^2$$

Pour trouver la base $q$-orthogonale correspondante, on résoud le système

$$\left\{\begin{array}{rcl} x+y+2z & = & x' \\ y+z & = & y' \\ z & = & z' \end{array}\right.$$

pour $(x',y',z') = (1,0,0)$ (premier vecteur de la base $q$-orthogonale) puis $(x',y',z') = (0,1,0)$ (deuxième vecteur de la base $q$-orthogonale) et $(x',y',z') = (0,0,1)$ (troisième vecteur de la base $q$-orthogonale).

**Exemple 3 :**

Soit $q : \mathbb{R}^4 \to \mathbb{R}$ l'application qui a $\mathbf{u} = \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix}$ associe

$$q(\mathbf{u}) = x^2 + 2xy + 2xz + 2xt + y^2 + 6yz - 2yt + z^2 + 10zt + t^2.$$

L'application $q$ est bien une forme quadratique car c'est un polynôme de degré 2 homogène.
Appliquons l'algorithme de Gauss à $q$ pour trouver une base $q$-orthogonale. On a

$$\begin{array}{rcl} q(\mathbf{u}) & = & x^2 + 2(y+z+t)x + y^2 + 6yz - 2yt + z^2 + 10zt + t^2 \\ & = & (x+y+z+t)^2 - (y+z+t)^2 + y^2 + 6yz - 2yt + z^2 + 10zt + t^2 \\ & = & (x+y+z+t)^2 + 4yz - 4yt + 8zt. \end{array}$$

On a maintenant

$$\begin{array}{rcl} 4yz - 4yt + 8zt & = & 4(yz + (-t)y + (2t)z) \\ & = & 4((y+2t)(z-t) + 2t^2) \\ & = & 4(y+2t)(z-t) + 8t^2 \\ & = & (y+z+t)^2 - (y-z+3t)^2 + 8t^2 \end{array} .$$

Finalement, on obtient

$$q(\mathbf{u}) = (x+y+z+t)^2 + (y+z+t)^2 - (y-z+3t)^2 + 8t^2.$$

On a donc $\mathrm{rg}(q) = 4$. On a

$$\left\{\begin{array}{rcl} L_1(u) & = & x+y+z+t \\ L_2(u) & = & y+z+t \\ L_3(u) & = & y-z+3t \\ L_4(u) & = & t \end{array}\right.$$

Calcul de $e_1'$ : on a $L_1(e_1') = 1, L_2(e_1') = L_3(e_1') = L_4(e_1') = 0$. Si $(x,y,z,t)$ sont les coordonnées de $e_1'$

$$\left\{\begin{array}{rcccccl} x & +y & +z & +t & = & 1 \\ & y & +z & +t & = & 0 \\ & y & -z & +3t & = & 0 \\ & & & t & = & 0 \end{array}\right.$$

donc $x = 1$, $y = z = t = 0$. La matrice du système est donnée par

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

La matrice du système est presque triangulaire supérieure, il y a donc assez peu de manipulation à faire pour résoudre le système. Avec un logiciel ou à la main, on calcule $M^{-1}$ et on lit $e'_1$ dans la 1ère colonne de $M$, $e'_2$ dans la deuxième colonne, etc.

$$e'_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix},\ e'_2 = \begin{pmatrix} -1 \\ 1/2 \\ 1/2 \\ 0 \end{pmatrix},\ e'_3 = \begin{pmatrix} 0 \\ 1/2 \\ -1/2 \\ 0 \end{pmatrix},\ e'_4 = \begin{pmatrix} 0 \\ -2 \\ 1 \\ 1 \end{pmatrix}$$

Ces vecteurs $(e'_1, e'_2, e'_3, e'_4)$ forment donc une base $q$-orthogonale. On vérifie en appliquant la formule de changement de base de la base $(e'_1, e'_2, e'_3, e'_4)$ où $q$ est diagonale (de coefficients 1, 1, -1 et 8) vers la base canonique.

## 2.3 Sur $\mathbb{C}$

Soit $n \in \mathbb{N}^*$. Pour $0 \leqslant r \leqslant n$, soit $I_r$ la matrice carré diagonale de taille $n$, contenant $r$ « 1 » sur les colonnes $1, \ldots, r$.

**Théorème II.24. Formes quadratiques complexes : classification**

Soit $S \in \mathcal{M}_n(\mathbb{C})$ symétrique. Il existe un unique $0 \leqslant r \leqslant n$ pour lequel il existe $P \in \mathrm{GL}_n(\mathbb{C})$ tel que

$$S = P I_r \, {}^t P.$$

**Preuve**

Soit $q$ la forme quadratique sur $\mathbb{C}^n$ associée à $S$ : $q(X) = {}^t X S X$. D'après l'algorithme de Gauss on peut écrire $q$ comme dans la formule (2.1). Pour chaque $\alpha_i \neq 0$, on choisit $\beta_i$ tel que $\beta_i^2 = \alpha_i$ et on pose $L'_i = \beta_i L_i$. Pour $i$ tel que $\alpha_i = 0$ on pose $L'_i = L_i$. Alors

$$q = \sum_{i=1}^{r} (L'_i)^r.$$

Le raisonement après la formule (2.1) montre l'existence de $r$ et $P$.
L'unicité vient de

$$r = \mathrm{rg}(I_r) = \mathrm{rg}(P I_r \, {}^t P) = \mathrm{rg}(S).$$

Une reformulation :

**Corollaire II.25**

L'ensemble $\{I_r \ : \ 0 \leqslant r \leqslant n\}$ est un système complet de représentants de l'action de $\mathrm{GL}_n(\mathbb{C})$ sur $S_n(\mathbb{C})$ données par

$$P.S = P S \, {}^t P.$$

## 2.4 Sur $\mathbb{R}$

Soit $E$ un $\mathbb{R}$-espace vectoriel de dimension $n$ et $q$ une forme quadratique sur $E$. On dit que $q$ est *positive* si

$$\forall v \in E \qquad q(v) \geqslant 0.$$

On dit que $q$ est *définie positive* si

$$\forall v \in E - \{0\} \qquad q(v) > 0.$$

**Théorème II.26. Théorème d'inertie de Sylvester**

Soit $S \in \mathcal{M}_n(\mathbb{R})$ symétrique. Il existe un unique couple d'entiers naturels $(r_-, r_+)$ tels que $r_- + r_+ \leqslant n$ et pour lequel il existe $P \in \mathrm{GL}_n(\mathbb{R})$ tel que

$$S = P \begin{pmatrix} I_{r_+} & & \\ & -I_{r_-} & \\ & & 0 \end{pmatrix} {}^tP.$$

**Preuve**

Soit $q$ la forme quadratique sur $\mathbb{R}^n$ associée à $S$ : $q(X) = {}^tXSX$. D'après l'algorithme de Gauss on peut écrire $q$ comme dans la formule (2.1). Pour chaque $\alpha_i \neq 0$, on choisit $\beta_i \in \mathbb{R}$ tel que $\pm\beta_i^2 = \alpha_i$ et on pose $L_i' = \beta_i L_i$. Pour $i$ tel que $\alpha_i = 0$ on pose $L_i' = L_i$. Alors

$$q = \sum_{i=1}^{r} \pm (L_i')^r.$$

On renumérote les $L_i$ pour avoir les $+$ puis les $-$. Le raisonement après la formule (2.1) montre l'existence de $r_\pm$ et $P$.

Comme sur $\mathbb{C}$, on a

$$r_+ + r_- = \mathrm{rg}(S).$$

Pour l'unicité, il suffit donc de montrer que

$$r_+ = \max\{\dim W \ : \ W \text{ est un sev de } \mathbb{R}^n \text{ t.q. } q_{|W} \text{ est définie positive}\}. \tag{2.2}$$

Soit $\mathcal{C} = (f_1, \ldots, f_n)$ une base de $\mathbb{R}^n$ telle que $\mathrm{Mat}_{\mathcal{C}}(q) = \begin{pmatrix} I_{r_+} & & \\ & -I_{r_-} & \\ & & 0 \end{pmatrix}$. Posons $W_0 = \mathrm{Vect}(f_1, \ldots, f_{r_+})$ et $G_0 = \mathrm{Vect}(f_{r_+ + 1}, \ldots, f_n)$. On remarque que

1. $\forall v \in W_0 - \{0\} \qquad q(v) > 0$ ; et
2. $\forall v \in G_0 \qquad q(v) \leqslant 0$.

La première propriété montre que le max de (2.2) vaut au moins $r_+$.
Pour montrer l'autre inégalité, fixons $W$ tel que $\dim(W) > r_+$ et montrons que $q_{|W}$ n'est pas définie positive. Puisque $\dim(W) + \dim(G_0) > n$, la formule Grassmann implique que $W \cap G_0$ n'est pas réduit à zéro. Soit $v$ non nul dans $W \cap G_0$. D'après la deuxième propriété, on $q(v) \leqslant 0$. Donc $q_{|W}$ n'est pas définie positive.

Le couple $(r_+, r_-)$ est appelé la *signature* de $q$.

## 2.5 Sur un corps fini $\mathbb{F}_q$

**Sur un corps quelconque.** On a déjà vu que le rang est un invariant. En revanche, le déterminant n'en est pas un puisque

$$\det(PS\,{}^tP) = (\det P)^2 \det(S)$$

peut être différent de $\det(S)$. Sur les réels, cette formule montre néanmoins que le signe se $\det(S)$ ne change pas. En général, la classe de $\det(S)$ dans $\mathbb{K}/(\mathbb{K}^*)2$ est invariante, on l'appelle le *discriminant* de la forme quadratique.

On a besoin d'un rafinement : *le discrimnant réduit*. Soit $E$ un $\mathbb{K}$-espace vectoriel de dimension finie et $Q$ une forme quadratique sur $E$. Soit $K$ le noyau de la forme bilinéaire $\varphi$ associée à $Q$. Alors, l'application

$$\begin{array}{rccc} \bar{\varphi} : & E/K \times E/K & \longrightarrow & \mathbb{K} \\ & (v,w) & \longmapsto & \varphi(v,w) \end{array}$$

est bien définie, bilinéaire et non dégénérée. Le discriminant de $\bar{\varphi}$ est donc un élément de $\mathbb{K}^*/(\mathbb{K}^*)^2$ bien défini et appelé *le discrimnant réduit* de $\varphi$ ou $Q$. On le note $\bar{\delta}(Q)$.

Pour calculer $\bar{\delta}(Q)$ on peut procéder comme suit. Soit $(e_1,\dots,e_s)$ une base de $\mathrm{Ker}(Q)$ que l'on complète en une base $\mathcal{B} = (e_1,\dots,e_n)$ de $E$. Alors

$$\mathrm{Mat}_{\mathcal{B}}(Q) = \begin{pmatrix} 0 & 0 \\ 0 & A \end{pmatrix}$$

pour une matrice carrée et inversible $A$. Alors, la classe de $\det(A)$ dans $\mathbb{K}/(\mathbb{K}^*)2$ est $\bar{\delta}(Q)$.
**Sur un corps fini $\mathbb{F}_q$.** Soit $p$ un premier différent de 2 et $q = p^\alpha$ avec $\alpha \in \mathbb{N}^*$. L'application carré

$$\mathbb{F}_p^* \longrightarrow \mathbb{F}_p^*, x \longmapsto x^2$$

est un morphisme de groupe dont le noyau est $\{\pm 1\}$. Donc son image $(\mathbb{F}_p^*)^2$ est d'indice 2. Fixons $\tau_0 \in \mathbb{F}_q^* - (\mathbb{F}_p^*)^2$. On a

$$\mathbb{F}_q^*/(\mathbb{F}_p^*)^2 = \{\bar{1}, \bar{\tau}_0\}.$$

**Théorème II.27. Formes quadratiques sur un corps fini : classification**

L'ensemble suivant est un système complet de représentants pour l'action de $\mathrm{GL}_n(\mathbb{F}_q)$ sur $S_n(\mathbb{F}_q)$ :

$$\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} \qquad \text{avec } 0 \leqslant r \leqslant n$$

$$\begin{pmatrix} I_{r-1} & 0 & 0 \\ 0 & \tau_0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \qquad \text{avec } 1 \leqslant r \leqslant n.$$

**Preuve**

Le rang et le discriminant réduit montrent que ces matrices sont 2 à 2 non congruentes.

Soit $S$ une matrice symétrique. Il reste à montrer que $S$ est congruente à une des matrices du théorème. L'algorithme de Gauss nous permet de supposer que $S$ est diagonale. Quitte à utiliser une matrice $P$ de permutation, on peut supposer que les coefficients non nuls de $S$ sont en haut à gauche.

**Lemme II.28**

Soit $a, b \in \mathbb{F}_q^*$. Alors il existe $x, y \in \mathbb{F}_q$ tels que $ax^2 + by^2 = 1$.

**Preuve du lemme**

On réécrit l'équation sous la forme

$$ax^2 = 1 - by^2.$$

Avec le 0, on a vu que $\mathbb{F}_q$ contient $\frac{q+1}{2}$ carrés. Mais alors,

$$\sharp\{ax^2 \ : \ x \in \mathbb{F}_q\} = \sharp\{1 - by^2 \ : \ x \in \mathbb{F}_q\} = \frac{q+1}{2}.$$

Ces deux ensembles sont des parties de $\mathbb{F}_q$ qui est de cardinal $q$. Ils ne peuvent être disjoint. Le lemme en découle.

Notons $Q$ la forme quadratique sur $\mathbb{F}_q^n$ associée à $S$. Soit $(e_1, \dots, e_n)$ la base canonique de $\mathbb{F}_q^n$. Si $\mathrm{rg}(S) \leqslant 1$, il n'y a rien à montrer. Sinon considérons $F = \overrightarrow{(}e_1, e_2)$. Soit $a$ et $b$ les deux premiers coefficents diagonaux de $S$. On a $Q(ae_1 + be_2) = ax^2 + by^2$. Le lemme donne donc $v \in F$ tel que $Q(v) = 1$. Soit $w$ un générateur de l'orthogonal de $v$ dans dans $F$ pour $Q_{|F}$.
La matrice de $Q$ dans la base $(v, w, e_3, \dots, e_n)$ est diagonale et commence par 1.
Maintenant, une récurence immédiate montre qu'il existe une base dans laquelle la matrice de $Q$ est diagonale et commence par $\mathrm{rg}(Q) - 1$ « 1 ».
On peut alors multiplier le $r^{ieme}$ vecteur de base par un scalaire pour obtenir 1 ou $\tau_0$ comme $r^{ieme}$ coefficent diagonal.

## 2.6 Sur $\mathbb{Q}$

Dans tous les cas précédent ($\mathbb{K} = \mathbb{C}$, $\mathbb{R}$ et $\mathbb{F}_q$) il n'y a qu'un nombre fini de classes de congruence dans $S_n(\mathbb{K})$. Nous allons voir que ce n'est pas le cas lorsque $\mathbb{K} = \mathbb{Q}$.

**Proposition II.29**

Dans $S_n(\mathbb{Q})$ les matrices

$$\begin{pmatrix} 1 & & & \\ & \ddots & & \\ & & 1 & \\ & & & p \end{pmatrix}$$

avec $p$ premier sont 2 à 2 non congruentes.
En particulier, $\mathrm{GL}_n(QQ)$ a une infinité d'orbites dans $S_n(\mathbb{Q})$.

**Preuve**

Le discriminant vaut $p$. Soit $p \neq q$ deux nombre prmiers. Il suffit de montrer que la classe de $p$ et $q$ dans $\mathbb{Q}^*/(\mathbb{Q}^*)^2$ sont distinctes.
Supposons par l'absurde qu'il existe $a, b \in \mathbb{Z} - \{0\}$ tels que

$$p = q(\frac{a}{b})^2$$

c'est-à-dire $b^2p = qa^2$. La valuation $p$ adique de $b^2p$ est impaire et celle de $qa^2$ paire. Contradiction.

# 3 Espaces Euclidiens

Dans cette section le corps est $\mathbb{R}$.

## 3.1 Définition

Soit $E$ un $\mathbb{R}$-espace vectoriel de dimension finie $n$.

**Définition II.30**

Une forme bilinéaire symétrique $B$ sur $E$ est appelée un *produit scalaire* si elle est définie positive.
On notera souvent $B(v,w)$ par $\langle v,w\rangle$.
L'espace vectoriel $E$ muni d'un praduit scalaire est appelé un *espace euclidien.*

Soit $M$ la matrice dans une base donnée de la forme bilnéaire symétrique $B$. Alors $B$ est un produit scalaire si et seulement si

$$\forall X \in \mathcal{M}_{n,1}(\mathbb{R}) \quad X \neq 0 \;\Rightarrow\; {}^tXMX > 0.$$

C'est aussi équivalent à ce que $q_B = \varphi_1^2 + \cdots + \varphi_n^2$ pour une base $(\varphi_1, \ldots, \varphi_n)$ de $E^*$. C'est aussi équivalent à ce que $q_B$ soit de signatue $(n,0)$.

**Définition II.31: Base Orthonormée**

Soit $(E, \langle\ ,\ \rangle)$ un espace euclidien. Une base $\mathcal{B} = (e_1, \ldots, e_n)$ de $E$ est dite orthonormée si

$$\forall i,j \qquad \langle e_i, e_j\rangle = \delta_i^j,$$

où $\delta$ est le symbole de Kronecker.

La base $\mathcal{B}$ est orthonormée si et seulement si $\mathrm{Mat}_{\mathcal{B}}(\langle\ ,\ \rangle) = I_n$. En particulier, le théorème de réduction montre qu'il existe toujours des bases orthonormées. Nous verrons plus tard l'algorithme de Gram-Schmidt qui est une alternative à Gauss pour en calculer.

**Exemple 5.**

1. Sur $\mathbb{R}^n$, la formule suivante définit un produit scalaire (appelé produit scalaire canonique) :
$$\left\langle \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \right\rangle = \sum_i x_i y_i$$
Vérifier que la base canonique est orthonormée.
2. Sur $\mathcal{M}_n(\mathbb{R})$, la formule suivante définit un produit scalaire :
$$\langle A, B\rangle = \mathrm{tr}({}^tAB).$$
Montrer que la base canonique est orthonormée.
3. Soit $E = \mathbb{R}_n[X]$. La formule suivante définit un produit scalaire :
$$\langle P, Q\rangle = \int_0^1 PQ(t)dt.$$

## 3.2 Cauchy-Schwarz et Minkowski

**Théorème II.32. Inégalité de Cauchy-Schwarz**

Soit E un espace vectoriel réel de dimension finie muni d'un produit scalaire $\langle\ ,\ \rangle$. Soit $u$ et $v$ dans $E$. Alors

$$|\langle u,v\rangle| \leqslant \sqrt{\langle u,u\rangle}\sqrt{\langle v,v\rangle}.$$

De plus, on a égalité si et seulement si la famille $(u,v)$ est liée.

**Preuve**

Considérons l'application

$$\begin{array}{rccc} \varphi : & \mathbb{R} & \longrightarrow & \mathbb{R}^+ \\ & t & \longmapsto & \langle u+tv, u+tv\rangle. \end{array}$$

L'application est bien à valeur dans $\mathbb{R}^+$ car $\langle\ ,\ \rangle$ est un produit scalaire. De plus comme $\varphi(t) = \langle v,v\rangle t^2 + 2\langle u,v\rangle t + \langle u,u\rangle$, $\varphi$ est un polynôme de degré au plus deux. Son discriminant est négatif ou nul :

$$\Delta = \langle u,v\rangle^2 - \langle v,v\rangle\langle u,u\rangle \leqslant 0.$$

Comme $\sqrt{\ }$ est croissante, l'égalité cherchée en découle.

Quand a t-on égalité ?
Si $v = 0$, on a égalité et la famille $(u,v)$ est liée. Supposons à présent $v \neq 0$. Alors, $\varphi$ est de degré 2. On a égalité dans l'inégalité du théorème si et seulement si

$$\begin{array}{lll} \Delta = 0 & \iff \varphi(-\frac{b}{2a}) = 0 & \text{où } a, b \text{ et } c \text{ sont les coefficients de } \varphi \\ & \iff (u - \frac{\langle u,v\rangle}{\langle v,v\rangle}v, u - \frac{\langle u,v\rangle}{\langle v,v\rangle}v) = 0 & \\ & \iff u - \frac{\langle u,v\rangle}{\langle v,v\rangle}v = 0 & \end{array}$$

Ceci implique bien que $(u,v)$ est liée. Réciproquement si la famille est liée alors $u = \lambda v$ (car $v$ est non nul) et on vérifie sans peine la relation.

On pose

$$\|u\| = \sqrt{(u,u)}. \tag{3.1}$$

Il s'agit d'une **norme** en vertu du

**Corollaire II.33: Inégalité de Minkowski**

Soit E un espace vectoriel réel de dimension finie muni d'un produit scalaire $\langle\ ,\ \rangle$. Soit $u$ et $v$ dans $E$. Alors

$$\|u+v\| \leqslant \|u\| + \|v\|.$$

De plus, on a égalité si et seulement si la famille $(u,v)$ est positivement liée c'est-à-dire il existe $\lambda$ et $\mu$ dans $\mathbb{R}^+$ non tous les deux nuls tels que

$$\lambda u + \mu v = 0.$$

**Preuve**

Regardons

$$\begin{array}{ll} (\|u\| + \|v\|)^2 - \|u+v\|^2 & = \|u\|^2 + \|v\|^2 + 2\|u\|\cdot\|v\| - \langle u+v, u+v\rangle \\ & = 2(\|u\|\cdot\|v\| - \langle u,v\rangle) \geqslant 0. \end{array}$$

L'inégalité à montrer en découle.
Si $u = \lambda v$ avec $\lambda \in \mathbb{R}^+$, un calcul direct montre qu'on a bien égalité.
Réciproquement supposons que l'on a égalité. D'après la suite d'inégalité ci-dessus, on a égalité dans l'inégalité de Cauchy-Schwarz. Quitte à échanger $u$ et $v$, on peut supposer qu'il existe $\lambda \in \mathbb{R}$ tel que $v = \lambda u$. Alors

$$\|u+v\| = |1+\lambda|\|u\|$$

et

$$\|u\| + \|v\| = (1+|\lambda|)\|u\|$$

. Donc $\lambda \geqslant 0$ ou $u = 0$. La conclusion est satisfaite dans les deux cas.

L'exemple de $\mathbb{R}^n$ dans la section précédente révèle que nos plan et espace ambiants sont euclidiens. En particulier la géométrie usuelle est euclidienne.
On dit que deux vecteurs $u$ et $v$ sont *orthogonaux* si $\langle u, v\rangle = 0$.

**Théorème II.34. Pythagore**

Les vecteurs $u$ et $v$ sont orthogonaux si et seulement si

$$\|u + v\|^2 = \|u\|^2 + \|v\|^2.$$

La preuve est directe. Illustrer cet énoncé par un dessin dans le plan qui explique son nom.

## 3.3 Projection orthogonale

Commençons par un résultat sur l'orthogonal d'un sous-espace.

**Proposition II.35: L'Orthogonal est Supplémentaire**

Soit E un espace vectoriel réel de dimension finie muni d'un produit scalaire $\langle\ ,\ \rangle$ et $F$ un sous-espace vectoriel de $E$.
Alors

$$F \oplus F^{\perp_{\langle\ ,\ \rangle}} = E.$$

**Preuve**

Comme $\langle\ ,\ \rangle$ est une forme bilinéaire non dégénérée

$$\dim(F^{\perp_{\langle\ ,\ \rangle}}) = \dim(E) - \dim(F).$$

Il suffit alors de montrer que $F \cap F^{\perp_{\langle\ ,\ \rangle}} = \{0\}$. Soit $v$ un vecteur de cette intersection. Alors $0 = \langle v, v\rangle = \|v\|^2$. Donc $v = 0$.

La *projection orthogonale sur* $F$ est l'application linéaire

$$\begin{array}{rccc} p_F : & E = F \oplus F^{\perp_{\langle\ ,\ \rangle}} & \longrightarrow & F \\ & x + y & \longmapsto & x \end{array}$$

On vérifie sans peine que $p_F$ est linéaire et, $\mathrm{Ker}(p_F) = F^{\perp_{\langle\ ,\ \rangle}}$ et $\mathrm{Im}(p_F) = F$.

Le théorème suivant donne une interprétation de $p_F(v)$. Il affirme que ce point est le point de $F$ le plus proche de $v$. Cela est très utilisé en optimisation et en statistique. On pose

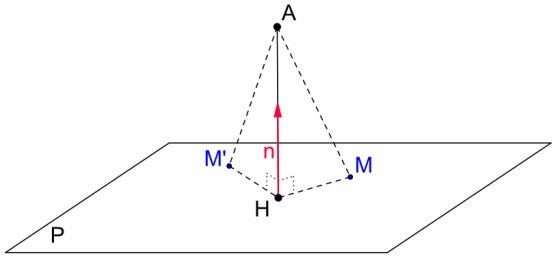
$$d(v, F) = \inf_{x \in F} \|v - x\|,$$

la distance de $v$ à $F$.

**Théorème II.36. Projection orthogonale**

Soit E un espace vectoriel réel de dimension finie muni d'un produit scalaire $\langle\ ,\ \rangle$. Soit $F$ un sous-espace vectoriel de $E$ et $v$ un vecteur de $E$. Le point $p_F(v)$ est l'unique point de $F$ tel que

$$\|v - p_F(v)\| = d(v, F). \tag{3.2}$$




**Preuve**

Il s'agit de montrer que pour tout $w \in F$, on a

$$\|v - p_F(v)\|^2 \leqslant \|v - w\|^2.$$

Écrivons $v - w = (v - p_F(v)) + (p_F(v) - w)$. Comme $v - p_F(v)$ est dans l'orthogonal de $F$, il est orthogonal à $p_F(v) - w$. Mais alors, le théorème de Pythagore implique que

$$\|v - w\|^2 = \|v - p_F(v)\|^2 + \|p_F(v) - w\|^2.$$

L'inégalité recherchée en découle.

## 3.4 Algorithme de Gram-Schmidt

**Théorème II.37. Gram-Schmidt**

Soit $\mathcal{C} = (\varepsilon_1, \dots, \varepsilon_n)$ une base de $E$. Il existe une unique base orthonormée $(e_1, \dots, e_n)$ telle que

1. Pour tout $k \in \{1, \dots, n\}$, on a $\text{Vect}(\varepsilon_1, \dots, \varepsilon_k) = \text{Vect}(e_1, \dots, e_k)$ ;
2. Pour tout $k \in \{1, \dots, n\}$, on a $\langle e_k, \varepsilon_k \rangle > 0$.

**Preuve**

La preuve se fait par récurrence et est essentiellement un algorithme. Il convient d'être capable de calculer la base orthonormée sur de petits exemples.
L'hypothèse de récurrence est : il existe une unique famille $(e_1, \dots, e_k)$ tels que

1. $\forall k' \leqslant k \qquad \text{Vect}(\varepsilon_1, \dots, \varepsilon_{k'}) = \text{Vect}(e_1, \dots, e_{k'})$ ;
2. $\forall k' \leqslant k \qquad \langle e_{k'}, \varepsilon_{k'} \rangle > 0$ ;
3. $\forall i, j \leqslant k \qquad \langle e_i, e_j \rangle = \delta_i^j$.

L'initialisation est laissée en exercice.
Hérédité. Supposons l'hypothèse de récurrence vrai au rang $k-1$. Prenons un vecteur $v$ de $\text{Vect}(\varepsilon_1, \dots, \varepsilon_k)$. Comme $\text{Vect}(\varepsilon_1, \dots, \varepsilon_{k-1}) = \text{Vect}(e_1, \dots, e_{k-1})$, on a une expression

$$v = \beta\varepsilon_k + a_{k-1}e_{k-1} + \cdots + a_1 e_1.$$

A quelles conditions $v = e_k$ convient ? On veut que pour tout $i < k$,

$$0 = \langle e_k, e_i \rangle = \beta\langle \varepsilon_k, e_i \rangle + a_i. \tag{3.3}$$

L'astuce consiste à considérer d'abord le cas où $\beta = 1$. Alors on pose alors

$$\tilde{a}_i = -\langle \varepsilon_k, e_i \rangle \qquad \text{et} \qquad \tilde{e}_k = \varepsilon_k + \tilde{a}_{k-1} e_{k-1} + \cdots + \tilde{a}_1 e_1.$$

On vérifie que $\tilde{e}_k$ satisfait toutes les conditions sauf

1. $\langle e_k, e_k \rangle = 1$ ;
2. $\langle e_k, \varepsilon_k \rangle > 0$.

Alors $e_k = \pm \frac{\tilde{e}_k}{\|\tilde{e}_k\|}$ convient.
Une relecture attentive de cette preuve montre que nous n'avions aucun choix et donc l'unicité.

## 3.5 Isométries et Groupe Orthogonal

Soit $E$ un espace euclidien.

**Définition II.38: Isométrie**

Une *isométrie* $u$ *de* $E$ est un endomorphisme de $E$ tel que

$$\forall x \in E \qquad \|u(x)\| = \|x\|.$$

Avec les formules de polarisation on montre facilement que

$$\forall x, y \in E \qquad \langle u(x), u(y) \rangle = \langle x, y \rangle. \tag{3.4}$$

On vérifie aussi que toute isométrie est inversible (son noyau étant réduit à $\{0\}$). De plus, l'ensemble des isométries est stable par composition et par inverse. Ainsi, l'ensemble $O(E)$ des isométries de $E$ est un **sous-groupe** de $\mathrm{GL}(E)$.

**Proposition II.39: Isométries et bases orthonormées**

Soit $\mathcal{B}$ une base orthonormée de $E$. Alors $u$ est une isométrie si et seulement si $u(\mathcal{B})$ est une base orthonormée.

**Preuve**

On peut supposer que $u$ est inversible. On pense alors à $u$ comme a un changement de base de $\mathcal{B}$ à $u(\mathcal{B})$. D'après (3.4), $u$ est une isométrie si et seulement si

$$\mathrm{Mat}_{\mathcal{B}}(\langle\ ,\ \rangle) = \mathrm{Mat}_{u(\mathcal{B})}(\langle\ ,\ \rangle).$$

Or $\mathrm{Mat}_{u(\mathcal{B})}(\langle\ ,\ \rangle)$ est la matrice dont les coefficients vaut $\langle u(e_i), u(e_j) \rangle$.

La preuve de la proposition montre aussi que

$$u \in O(E) \quad \Longleftrightarrow \quad {}^t\mathrm{Mat}_{\mathcal{B}}(u)\mathrm{Mat}_{\mathcal{B}}(u) = I_n.$$

On définit donc

$$O_n(\mathbb{R}) = \{A \in \mathcal{M}_n(\mathbb{R}) \ : \ {}^tAA = I_n\}.$$

**Exercice 4.** *1. Montrer que*

$$O_2(\mathbb{R}) = \{\begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix} \ : \ t \in \mathbb{R}\} \cup \{\begin{pmatrix} \cos t & \sin t \\ \sin t & -\cos t \end{pmatrix} \ : \ t \in \mathbb{R}\}.$$

*2. En déduire que $O_2(\mathbb{R})$ est constitué des rotations et des symétries orthogonales.*

*3. Montrer que plus généralement, toute symétrie orthogonale est une isométrie.*

**Théorème II.40. Groupe orthogonal**

$O_n(\mathbb{R})$ est un sous-groupe compact de $\mathrm{GL}_n(\mathbb{R})$.

**Preuve**

Nous avons déjà vu que c'est un sous-groupe. Il est fermé comme préimage de $\{I_n\}$ par l'application $A \mapsto {}^tAA$.
Il est borné car chaque colonne $C$ des éléments de $O_n(\mathbb{R})$ vérifient $\|C\| = 1$.

## 3.6 Adjoint et endomorphismes symétrique

Soit $(E, \langle\ ,\ \rangle)$ un espace euclidien.

**Lemme II.41**

Soit $f$ un endomorphisme de $E$. Alors, il existe un unique endomorphisme $f^*$ de $E$ tel que

$$\forall v, w \in E \qquad \langle v, f(w)\rangle = \langle f^*(v), w\rangle.$$

**Preuve**

Le produit scalaire permet d'identifier $E$ à $E^*$ :

$$\begin{array}{rccccc} \iota : & E & \longrightarrow & E^* & & \\ & v & \longmapsto & E & \longrightarrow & \mathbb{R} \\ & & & w & \longrightarrow & \langle v, w\rangle. \end{array}$$

On vérifie que l'identité du lemme signifie que

$${}^tf \circ \iota = \iota \circ f^*.$$

En effet, pour tout $v, w \in E$, on a

$$({}^tf \circ \iota(v))(w) = \langle v, f(w)\rangle$$

et

$$(\iota \circ f^*(v))(w) = \langle f^*(v), w\rangle.$$

Le lemme en découle.

**Preuve matricielle.** Soit $\mathcal{B}$ une base orthomormé de $E$ et $A$ la matrice de $f$ dans la base $\mathcal{B}$. Soit $v, w \in E$ et $X, Y \in \mathcal{M}_{n1}(\mathbb{R})$ leurs vecteurs coordonnées. Alors

$$\langle v, f(w)\rangle = {}^tXAY.$$

Si $B$ est la matrice de $f^*$ alors

$$\langle f^*(v), w\rangle = {}^t(BX)Y = {}^tX{}^tBY.$$

Donc $B = {}^tA$ convient. L'unicité provient de l'unicité de la matrice d'une forme bilinéaire qui elle-même est une coonséquence de

$$B_{ij} = ({}^tB)_{ji} = {}^tE_j\,{}^tBE_i,$$

où $E_i$ et $E_j$ sont les matrices de la base canonique.

La seconde preuve du lemme (et en fait aussi la première) montre que, **dans une base orthonormée** $\mathcal{B}$, on a

$$\mathrm{Mat}_{\mathcal{B}}(f^*) = {}^t\mathrm{Mat}_{\mathcal{B}}(f).$$

**Définition II.42**

On dit que $f$ est *symétrique* si $f^* = f$ c'est-à-dire si la matrice de $f$ dans une base orthonormée est symétrique.

**Théorème II.43. Réducrion des endomorphismes symétriques**

1. Soit $(E, \langle\ ,\ \rangle)$ un espace euclidien et $f$ un endomorphisme symétrique. Alors, il existe une base orthormée $\mathcal{B}$ telle que la matrice $\mathrm{Mat}_{\mathcal{B}}(f)$ est diagonale.
2. Soit $(E, \langle\ ,\ \rangle)$ un espace euclidien et $f$ un endomorphisme symétrique. Alors, il existe une base orthormée $\mathcal{B}$ constituée de vecteurs propres de $f$.
3. Soit $S$ une matrice symétrique de taille $n$. Alors, il existe $O \in O_n(\mathbb{R})$ telle que $OS\,{}^tO$ est diagonale.

Pour $n \in \mathbb{N}^*$, posons

$$\mathcal{D}_n := \{\mathrm{diag}(\lambda_1, \ldots, \lambda_n)\ :\ \lambda_1 \geqslant \cdots \geqslant \lambda_n \in \mathbb{R}\}.$$

**Corollaire II.44**

L'ensemble $\mathcal{D}_n$ est un système complet de représentants pour l'action de $O_n(\mathbb{R})$ sur $S_n(\mathbb{R})$.

Le corollaire est une conséquence directe du théorème. Les 3 versions sont trivialement équivalentes. Montrons la version matricielle.

**Preuve**

Commençons par montrer que $S$ a une valeur propre. Pour cela, posons $S^{n-1} := \{X \in \mathbb{R}^n\ :\ \|X\| = 1\}$ et considèrons la fonction

$$\begin{array}{rccc} \varphi\ : & S^{n-1} & \longrightarrow & \mathbb{R} \\ & X & \longmapsto & \langle SX, X\rangle. \end{array}$$

Comme $S^{n-1}$ est compact et $\varphi$ est continue, le maximum de $\varphi$ est atteint disons en $X_0$.
On étend $\varphi$ en $\tilde{\varphi}$ par la même formule. On a

$$\tilde{\varphi}(X_0 + tv) = \langle S(X_0 + tv), X + tv\rangle = \tilde{\varphi}(X_0) + 2t\langle SX_0, v\rangle + t^2\langle Sv, v\rangle$$

et la différentielle de $\tilde{\varphi}$ en $X_0$ est

$$T_{X_0}\tilde{\varphi} : v \longmapsto 2\langle SX_0, v\rangle.$$

De même, la différentielle de $X \mapsto \langle X, X\rangle$ en $X_0$ est $T_{X_0}N : v \longmapsto 2\langle X_0, v\rangle$. Le théorème des extrema liés assure alors que $T_{X_0}\tilde{\varphi}$ s'annule sur le noyau de $T_{X_0}N$. En remarquant, qu'il s'agit de

deux formes linéaires, on en déduit qu'elles sont proportionnelles :

$$\exists \lambda \in \mathbb{R} \qquad T_{X_0}\tilde{\varphi} = \lambda T_{X_0} N.$$

Comme $\langle\ ,\ \rangle$ est non-dégénré, il vient $SX_0 = \lambda X_0$.
Remarquons qu'en évaluant en $X_0$, il vient $\varphi(X_0) = \lambda$.

Soit $F$ un sous-espace vectoriel de $\mathbb{R}^n$ stable par $S$. Montrons que $F^\perp$ est stable par $S$.
Une première façons d'obtenir cela est d'invoquer une base orthonormés $\mathcal{C}$ qui commence par une base de $F$ (et se poursuit donc par une base de $F^\perp$). Comme la matrice de (l'endomorphisme associé à) $S$ dans la base $\mathcal{C}$ est symétrique, $F^\perp$ est stable.
On peut aussi raisonner directement. Soit $Y \in F^\perp$. Il s'agit de montrer que $SY \in F^\perp$ c'est-à-dire

$$\langle SY, X\rangle = 0 \qquad \forall X \in F.$$

Cela découle de $\langle SY, X\rangle = \langle Y, SX\rangle$ et $SX \in F$.

Nous sommes maintenant en position de finir la preuve par récurrence sur la taille de matrice $S$.
Si $S$ est de taille 1, il n'y a rien à montrer. Supposons le résultat connu pour toute matrice de taille inférieure à $n-1$.
D'après la première étape, il existe une valeur propre $\lambda$. Considérons le sous-espace propre associé $E_\lambda$ et $k$ sa dimension. D'après la seconde étape, il existe $O \in O_n(\mathbb{R})$ telle que

$$OS\,{}^tO = \begin{pmatrix} \lambda I_k & 0 \\ 0 & T \end{pmatrix}$$

On applique alors l'hypothèse de récurence à $T$ pour conclure.

# Chapitre 3

# Formes hermitiennes

## Sommaire

# 1 Introduction

Dans ce chapitre on s'intéresse aux espaces vectoriels complexes. Soit $E$ un $\mathbb{C}$-espace vectoriel de dimension finie. On remarque que $E$ est à fortiori un espace vectoriel réel. On va dans ce chapitre étudier certaines normes euclidiennes adaptée à la structure d'espace vectoriel complexe.
Commençons par regarder le cas le plus simple : $E = \mathbb{C}$ munni du module. Alors $E$ est un espace vectoriel réel de dimension deux. Soit $z \in \mathbb{C}$. Lorsque l'on pense à $\mathbb{C}$ comme à un $\mathbb{R}$-espace vectoriel, on écrit $z = x + iy$ avec $x, y \in \mathbb{R}$. Alors

$$|z|^2 = x^2 + y^2$$

est bien une norme euclidienne. Elle vérifie $|iz| = |z|$. De plus,

$$|z|^2 = z\bar{z}.$$

Cette dernière écriture insite à considérer la conjugaison complexe et les formes antilinéaires.

# 2 Espaces vectoriels complexes et leur espace antidual

## 2.1 Espaces vectoriels réels et complexes

Comme nous l'avons déjà dit un espace vectoriel complexe est un espace vectoriel réel munni d'une structure supplémentaire. Soit $E$ un espace vectoriel réel de dimension finie. Pour le munir d'une structure de $\mathbb{C}$-espace vectoriel il faut déterminer $zv$ pour tout $v \in E$ et $z \in \mathbb{C}$. Si l'on écrit $z = x + iy$, on doit avoir

$$zv = (xv) + y(iv).$$

Donc il suffit de connaitre $iv$.

**Définition III.45: Espace vectoriel complexe**

Un espace vectoriel complexe $E$ est un espace vectoriel réel munni d'un endomorphisme $I$ tel que $I^2 = -\mathrm{Id}_E$.

Etant donné un espace vectoriel complexe $E$, pour $v \in E$ et $z = x + iy \in \mathbb{C}$ on pose

$$zv = xv + yI(v).$$

On vérifie que cette formule vérifie les axiomes usuels de $\mathbb{C}$-espace vectoriel.

## 2.2 Antidual

**Définition III.46**

Une forme *antilinéaire* $\varphi : E \longrightarrow \mathbb{C}$ est application telle que

1. $\forall v, w \in E \qquad \varphi(v + w) = \varphi(v) + \varphi(w)$ ;
2. $\forall v \in E \quad \forall \lambda \in \mathbb{C} \qquad \varphi(\lambda v) = \bar{\lambda}\varphi(v)$.

Notons $\bar{E}^*$ l'ensemble des formes antilinéaires.

Remarquons que $\bar{E}^*$ est un $\mathbb{C}$-espace vectoriel (sev de toutes les fonctions de $E$ dans $\mathbb{C}$). De plus, l'application $\bar{E}^* \longrightarrow E^*; \varphi \longmapsto \bar{\varphi}$ est bijective (mais pas $\mathbb{C}$-linéaire !). On en déduit que si $\mathcal{B} = (e_1, \dots, e_n)$ est une base de $E$ alors $\bar{\mathcal{B}}^* := (\bar{e}_1^*, \dots, \bar{e}_n^*)$ est une base de $\bar{E}^*$. Ici $\bar{e}_k^*$ est défini par

$$\begin{array}{rccc} \bar{e}_k^* : & E & \longrightarrow & \mathbb{C} \\ & \sum_k \lambda_k e_k & \longmapsto & \bar{\lambda}_k. \end{array}$$

On a

$$\bar{e}_k^*(e_l) = \delta_k^l \qquad \bar{e}_k^*(ze_l) = \bar{z}\delta_k^l.$$

Si $E = \mathbb{C}^n$ et $\mathcal{B} = (e_1, \dots, e_n)$ est la base canonique, $\bar{e}_k^*$ est souvent noté $\bar{z}_k$.

**Proposition III.47**

Soit $\mathcal{B} = (e_1, \ldots, e_n)$ et $\mathcal{C} = (f_1, \ldots, f_n)$ deux bases de $E$. Soit $\varphi \in \bar{E}^*$.

1. Les coordonnées de $\varphi$ dans $\bar{\mathcal{B}}^*$ sont les $\varphi(e_i)$. Ainsi
$$\varphi = \sum_{k=1}^{n} \varphi(e_k)\bar{e}_k^*.$$
2. Si $C = \mathrm{Mat}_{\bar{\mathcal{B}}^*}(\varphi)$ et $X = \mathrm{Mat}_{\mathcal{B}}(x)$, on a
$$\varphi(x) = {}^tC\bar{X}.$$
3. On a la relation
$$\mathrm{Mat}_{\bar{\mathcal{B}}^*\,\bar{\mathcal{C}}^*}(Id_{\bar{E}^*}) = {}^t\overline{\mathrm{Mat}_{\mathcal{C}\,\mathcal{B}}(Id_{\bar{E}})} = \overline{\mathrm{Mat}_{\mathcal{B}^*\,\mathcal{C}^*}(Id_{E^*})}.$$

**Preuve**

Pour la première assertion, on calcule (pour $z_k \in \mathbb{C}$) d'un côté
$$\varphi(z_1e_1 + \cdots + z_ne_n) = \bar{z}_1\varphi(e_1) + \cdots + \bar{z}_n\varphi(e_n)$$
et de l'autre
$$\begin{array}{ll}(\sum_{k=1}^n \varphi(e_k)\bar{e}_k^*)(z_1e_1 + \cdots + z_ne_n) & = \sum_k \varphi(e_k)(\bar{e}_k^*)(z_1e_1 + \cdots + z_ne_n \\ & = \sum_{k,l} \varphi(e_k)\bar{z}_l\ \bar{e}_k^*(e_l) \\ & = \sum_k \varphi(e_k)\bar{z}_l.\end{array}$$

Si on écrit $\varphi = \sum_k \lambda_k\bar{e}_k$ et $x = \sum_k z_ke_k$, les deux côtés de l'égalité de la seconde assertion donnent
$$\sum_k \lambda_k\bar{z}_k.$$

Posons $X = \mathrm{Mat}_{\bar{\mathcal{B}}^*}(\varphi)$ et $Y = \mathrm{Mat}_{\bar{\mathcal{C}}^*}(\varphi)$. La première assertion montre que $\bar{X} = \mathrm{Mat}_{\mathcal{B}^*}(\bar{\varphi})$ et $\bar{Y} = \mathrm{Mat}_{\mathcal{C}^*}(\bar{\varphi})$. On a donc
$$\bar{X} = \mathrm{Mat}_{\mathcal{B}^*\,\mathcal{C}^*}(Id_{E^*})\bar{Y},$$
et
$$X = \overline{\mathrm{Mat}_{\mathcal{B}^*\,\mathcal{C}^*}(Id_{E^*})}Y.$$
D'où la dernière assertion.

## 2.3 Formes sesquilinéaires

Soit $E$ un espace vectoriel complexe de dimension $n$. Une *forme sesquilinéaire sur $E$* est une application linéaire
$$\begin{array}{rccc}\tilde{\phi} : & E & \longrightarrow & \bar{E}^* \\ & x & \longmapsto & \left(\begin{array}{ccc} E & \longrightarrow & \mathbb{C} \\ y & \longmapsto & \tilde{\phi}(x)(y)\end{array}\right).\end{array}$$

On écrit, plus souvent $\tilde{\phi}(x)(y) = \phi(x, y)$. La forme est dite *hermitienne* si $\phi(y, x) = \overline{\phi}(x, y)$.
Soit $\mathcal{B} = (e_1, \ldots, e_n)$ une base de $E$ et $\bar{\mathcal{B}} := (\bar{e}_1^*, \ldots, \bar{e}_n^*)$ la base de $\bar{E}^*$ associée. Par définition, la matrice de $\phi$ est la transposée $A$ de la matrice de $\tilde{\phi}$.
Comme dans le cas réel, on définit le rang et le noyau de $\phi$ comme étant les noyaux et rang de $\tilde{\phi}$. On dit qu'elle est non dégénérée si son rang est la dimension de $E$.

Les propriétés fondamentales de cette matrice son résumé dans l'énoncé suivant.

**Proposition III.48**

Soit $\phi$ une forme sesquilinéaire sur $E$ munni de la base $\mathcal{B} = (e_1, \dots, e_n)$. Soit $A$ sa matrice. On a :

1. Le coefficient $(k, l)$ de $A$ vaut $\phi(e_k, e_l)$.
2. Soit $x$ et $y$ dans $E$ et $X$ et $Y$ leurs vecteurs coordonnées. On a :

$$\phi(x, y) = {}^tXA\bar{Y}.$$

**Preuve**

La ligne $k$ de $A$ est par définition ${}^t\mathrm{Mat}_{\bar{\mathcal{B}}^*}(\tilde{\phi}(e_k))$. D'après la Proposition III.47, le $l^{\text{eme}}$ coefficient de ${}^t\mathrm{Mat}_{\bar{\mathcal{B}}^*}(\tilde{\phi}(e_k))$ est $\tilde{\phi}(e_k)(e_l) = \phi(e_k, e_l)$. Ceci prouve la première assertion.
On a, par définition, $\mathrm{Mat}_{\bar{\mathcal{B}}^*}(\tilde{\phi}(x)) = {}^tAX$. D'après la Proposition III.47, il vient alors

$$\tilde{\phi}(x)(y) = {}^t({}^tAX)\bar{Y} = {}^tXA\bar{Y}.$$

# 3 Espaces Hermitiens

**Définition III.49**

Une forme sesquilinéaire et hermitienne est dite *définie positive* si pour tout $x \in E$ non nul on a

$$\phi(x, x) > 0.$$

On montre que dans ce cas, la forme est non dégénérée et que $\|x\| = \sqrt{\phi(x, x)}$ définit une norme sur $E$ qui satisfait $\|zx\| = |z|\|x\|$. On dit qu'une telle norme *dérive d'un produit hermitien.*
Il y a un seul produit hermitien sur $E$ à changement de base près :

**Théorème III.50**

Il existe des bases $\mathcal{B} = (e_1, \dots, e_n)$ de $E$ telles que

$$(e_i, e_j) = \delta_i^j$$

c'est-à-dire la matrice de $(\cdot, \cdot)$ est $I_n$. Une telle base est dite *unitaire.*

**Preuve**

Les deux preuves du cas réels, j'ai nommé algorithme de Gauss et Gram-Schmidt s'adaptent au cas complexe.

Le lien avec les normes dérivant d'un produit scalaire est le suivant.

**Théorème III.51**

Soit $E$ un espace vectoriel complexe et $I$ la multiplication par $i \in \mathbb{C}$. Soit $\|\cdot\|$ une norme sur $E$ pensé comme espace vectoriel réel. Alors, se valent

1. $\|\cdot\|$ dérive d'un produit scalaire et $I$ est une isométrie.
2. $\|\cdot\|$ dérive d'un produit hermitien.

**Preuve**

# 4 Réduction des endomorphismes normaux

## 4.1 Adjoint d'un endomorphisme

Soit $(E,(\cdot,\cdot))$ un espace hermitien et $f \in \mathcal{L}(E)$. Alors, il existe un unique endomorphisme $f^* \in \mathcal{L}(E)$ tel que

$$\forall x,y \in E \qquad (x,f(y)) = (f^*(x),y). \tag{4.1}$$

Cette équation s'écrit

$$\forall x,y \in E \qquad \tilde{\phi}(x)\circ f(y) = \tilde{\phi}(f^*(x))(y)$$

ou encore

$$\forall x \in E \qquad \tilde{\phi}(x)\circ f = (\tilde{\phi}\circ f^*)(x)$$

c'est-à-dire

$$\forall x \in E \qquad (\tilde{\phi})^{-1}\left(\tilde{\phi}(x)\circ f\right) = f^*(x).$$

D'où l'existence et l'unicité de $f^*$ comme application de $E$ dans $E$. La linéarité de $f^*$ est facile à vérifier.

**Proposition III.52**

Soit $\mathcal{B}$ une base de $E$. Soit $A$ la matrice de $\phi$, $M$ la matrice de $f$ et $M^*$ celle de $f^*$ dans la base $\mathcal{B}$. Alors

$$M^* = {}^tA^{-1}{}^t\bar{M}{}^tA.$$

En particulier, si la base est unitaire, on a $M^* = {}^t\bar{M}$.

**Preuve**

Soit $X$ et $Y$ deux vecteurs colonnes. Notons aussi D'après la proposition V.74, la condition (4.1) s'écrit

$${}^tXA\overline{MY} = {}^t(M^*X)A\bar{Y},$$

ou encore

$${}^tX(A\bar{M})\bar{Y} = {}^tX({}^tM^*A)\bar{Y}.$$

Vu que cela est vrai pour tout $X$ et $Y$, on obtient $A\bar{M} = {}^tM^*A$ ou encore $M^* = {}^tA^{-1}{}^t\bar{M}{}^tA$.
Si la base est unitaire alors $A = I_n$.

Un lemme utile pour la suite.

**Lemme III.53**

Si $F$ est stable par $f$ alors $F^\perp$ est stable par $f^*$.

**Preuve**

Soit $x \in F^\perp$. Il s'agit de montrer que $f^*(x)$ est orthogonal à tout $y \in F$. Pour un tel $y$, on a

$$(f^*(x), y) = (x, f(y)) = 0,$$

car $f(x) \in F$.

## 4.2 Réduction des endomorphismes normaux

On dit que $f$ est *normal* si

$$f \circ f^* = f^* \circ f.$$

**Théorème III.54**

Tout endomorphisme normal est diagonalisable en base unitaire.

**Preuve**

On procède par récurrence sur la dimension de $E$. Si elle vaut 1, il n'y a rien à démontrer.
Comme $\mathbb{C}$ est algébrique clos, il existe une valeur propre $\lambda$. Notons $E_\lambda$ le sous-espace propre associé. Comme $f$ et $f^*$ commutent, $E_\lambda$ est stable par $f^*$. En vertu du lemme III.53, $E_\lambda^\perp$ est stable par $(f^*)^* = f$. On recommence pour la restriction de $f$ à $E_\lambda^\perp$.
On obtient alors que $E$ est la somme orthogonale des sous-espaces propres de $E$.

# Chapitre 4

# Matrices à coefficients dans un anneau euclidien

**Sommaire**

# 1 Rappels sur les anneaux euclidiens

Def et exemples
Gauss, Bezout, ppcm, pgcd, Euclide, factoriel.
Pas de preuve.

# 2 Théorème de Smith

Soit $A$ un anneau euclidien de Sthasme $N$.

## 2.1 Le groupe $\mathrm{GL}_n(A)$

Soit $M \in \mathcal{M}_n(A)$ une matrice carré à coefficents dans $A$.
On note $\mathrm{Com}(A)$ la matrice de $\mathcal{M}_n(A)$ dont le coefficient est $(-1)^{i+j} \det_{n-1}(A^{ij})$, où $A^{ij}$ est le mineur obtenu en supprimant la ligne $i$ et la colonne $j$.

**Proposition IV.55: Formule de Cramer**

On a

$$ {}^t\mathrm{Com}(A).A = A.{}^t\mathrm{Com}(A) = \det(A) I_n. $$

Ceci est conséquence simple de la $n$-linéarité du déterminant, c'est-à-dire des règmes de développement par rapport aux lignes et aux colonnes.

On définit le groupe $\mathrm{GL}_n(A)$ comme l'ensemble des éléments inversibles de l'anneau $\mathcal{M}_n(A)$ :

$$ \mathrm{GL}_n(A) = \{M \in \mathcal{M}_n(A) \ : \ \exists N \in \mathcal{M}_n(A) \qquad MN = NM = I_n\}. $$

**Corollaire IV.56**

Une matrice $M$ appartient à $\mathrm{GL}_n(A)$ si et seulement si son déterminant est inversible dans $A$.

## 2.2 Réduction de Smith

**Théorème IV.57. Forme normale de Smith**

Rappelons que $A$ est euclidien. Soit $M$ une matrice de taille $p \times q$ à coefficients dans $A$. Alors, il existe $(P, Q)$ dans $\mathrm{GL}_p(A) \times \mathrm{GL}_p(A)$, $d_1, \cdots d_s$ dans $A$ tels que :

1. $d_1 | d_2 | \cdots | d_s$,
2. et

$$ PMQ = \begin{pmatrix} d_1 & & & & & \\ & \ddots & & & & \\ & & d_s & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{pmatrix} $$

De plus, $s$ et les $d_i$ ne dépendent que de (l'image de) $M$ à association près.

Commençons par regarder le cas où $p = 2$ et $q = 1$, donc $M = \begin{pmatrix} a \\ b \end{pmatrix}$. Soit $P = \begin{pmatrix} u & v \\ u' & v' \end{pmatrix}$ dans $\mathrm{GL}_2(A)$. Alors $PM$ a la forme voulue si et seulement si (à association près)

$$ \{uv' - u'v = 1 ua + vb = du'a + v'b = 0 $$

Soit $p$ le pgcd de $a$ et $b$. On écrit $a = pa'$ et $b = pb'$. Alors

$$\{uv' - u'v = 1ua' + vb' = d/pu'a' + v'b' = 0$$

La dernière équation implique $u' = b'$ et $v' = -a'$ sachant que $\text{pgcd}(u', v') = \text{pgcd}(a', b') = 1$. Mais alors les deux première équations donnent $d = p$.

Fort de ce lien entre facteurs invariants et pgcd, on peut réinterpréter l'algorithme d'Euclide :

*Par des opérations élémentaires sur les lignes la matrice $M$ peut être transformé en la matrice* $\begin{pmatrix} \text{pgcd}(a,b) \\ 0 \end{pmatrix}$.

**Preuve**

**Existence.** Algorithme
Commençons par mentionner une conséquence de l'existence.

**Corollaire IV.58**

Le groupe $\text{GL}(A)$ est engendré par les matrices de transverction, permutatin et dilatation.

**Unicité.** Il s'agit de calculer les $d_i$ en des termes visiblement invariants :

**Lemme IV.59**

1. Le pgcd des mineurs de taille $r$ est invariant.
2. Pour la matrice diagonale de l'énnoncé Le pgcd des mineurs de taille $r$ vaut $d_1 \dots d_r$.

**Exemple 6.** Calculer $P$, $Q$ et les invariants de la matrice :

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 6 & 6 \end{pmatrix}$$

On part avec

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 6 & 6 \end{pmatrix} \quad \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

# 3 Systèmes diophantiens

Soit $MX = B$ un système linéaire à coefficients dans $A$.

# 4 Theorème de la base adaptée

**Théorème IV.60**

Pour tout sous-$A$-module $\Lambda$ de $A^n$, il existe une base $(e_1, \dots, e_n)$ de $A^n$ et des $d_i$ se divisant tels que

$$\Lambda = Ad_1e_1 \oplus \dots \oplus Ad_se_s.$$

**Preuve**

On commence par montrer par récurrence sur $n$ qu'un sous-module de $A^n$ est de type fini en écrivant $A^n = A^{n-1} \times A$. Si $n = 1$, un sous-module est un idéal, il est donc engendré par un élément.
Soit $\pi : A^n \longrightarrow A$ la projection sur le dernier facteur. Soit $N$ un sous-module de $A^n$.
On observe d'abord que $\pi(N)$ est un idéal de $A$, donc $\pi(N) = (a)$ avec $a \in A$. Si $a = 0$, $N$ est inclus dans $A^{n-1}$ et on conclut par récurrence.
Supposons $a \neq 0$. Soit $x$ tel que $\pi(x) = a$. On vérifier que $N = (\mathrm{Ker}\pi \cap N) + Ax$. Alors l'hypothèse de récurrence permet de conclure.

Remarquons que la preuve ci-dessus montre même plus : $N$ est engendré par au plus $n$ éléments. Soit $v_1, \ldots, v_k$ des générateurs de $N$. Ecrivant chaque $v_i$ comme un élément de $A^n$, on obtient une matrice $M \in \mathcal{M}_{nk}(A)$. On lui applique le théorème de réduction de Smith : $M = PDQ$.

Un des points importants dans la preuve ci-dessus est qu'un sous-module d'un $A$-module de type fini est de type fini. Voici un exemple qui montre que cela n'est pas trivial.

**Exemple 7.** Soit $k$ un corps et soit $A := k[X_i : i \in \mathbb{N}]$ l'anneau de polynômes en une infinité de variables (la réunion croissante des $k[X_1, \ldots, X_n]$) C'est un $A$-module de type fini, engendré par 1. Cependant, le sous-A-module $I := (X_i : i \in \mathbb{N})$ (c'est bien un sous-$A$-module car c'est un idéal) n'est pas de type fini. En effet, si $(f_1, \ldots, f_k) = I$ alors soit $n$ tel que $f_i \in k[X_1, \ldots, X_n]$ pour tout $i$. On a donc $X_{n+1} = \sum_i g_i f_i$ pour certains $g_i \in A$, et on trouve une contradiction en substituant 0 à $X_1, \ldots, X_n$ (les éléments de $I$ n'ont pas de coefficient constant).

# 5 Groupes abéliens de type fini

Enoncé + Exemple.

**Exemple 8.** Soit $G = \mathbb{Z}/12\mathbb{Z} \times \mathbb{Z}/15\mathbb{Z}$. Trouver les invariants de $G$.

# 6 Réduction de Frobenius

On se donne un endomorphisme $u$ d'un $k$-espace vectoriel de dimension finie $E$. Alors, $E$ est muni d'une structure de $k[X]$-module par la formule :

$$P.x = P(u)(x) \quad \forall x \in E \ \forall P \in k[X].$$

On fixe une base $\mathcal{B} = (e_1, \cdots, e_n)$ de $E$. Considérons alors l'application

$$\begin{array}{rccc} \varphi : & k[X]^n & \longrightarrow & E \\ & (P_1, \cdots, P_n) & \longmapsto & \sum_i P_i.e_i. \end{array}$$

D'après le théorème **??**, il existe une base $(E_1, \cdots, E_n)$ de $k[X]^n$ et des polynômes non nuls $P_1, \cdots, P_s$ tels que $(P_1.E_1, \cdots, P_s.E_s)$ est une base de $\mathrm{Ker}\varphi$. De plus, $k[X]^n/\mathrm{Ker}\varphi$ est isomorphe à $E$, qui est l'image de $\varphi$. En particulier, ce $k$-espace vectoriel est de dimension finie. On en déduit que $s = n$.
Plus précisément,

$$k[X]^n/\mathrm{Ker}\varphi \simeq k[X]/(P_1) \times \cdots \times k[X]/(P_n).$$

Notons $d_i$ le degré de $P_i$. Posons

$$\mathcal{B}_1 = (E_1, XE_1, \cdots, X^{d_1-1}E_1, E_2, XE_2, \cdots, X^{d_n-1}E_n).$$

Alors, $\varphi(\mathcal{B}_1)$ est une base du $k$-espace vectoriel $E$. On vérifie alors aisément que la matrice de $u$ dans cette base est la juxtaposition diagonale des matrices compagnons des polynômes $P_1, \cdots, P_n$. On retrouve donc le résultat d'existence du théorème suivant :

**Théorème IV.61. Frobenius**

Soit $u \in \mathcal{L}(E)$. Il existe une décomposition (non unique) de $E$ en somme directe de sous-espaces cycliques : $E = F_1 \oplus \cdots \oplus F_s$ telle que

$$\mu_{u_{|F_s}} | \mu_{u_{|F_{s-1}}} | \cdots | \mu_{u_{|F_1}}.$$

De plus, les polynômes $\mu_{u_{|F_i}}$ ne dépendent pas de la décomposition : on les appelle les facteurs invariants de $u$.

*Remarque.* Dans la discussion ci-dessus, il se peut que certains polynômes soient égaux à 1. Il n'interviennent alors nulle part. On pourra en guise d'exercice caractériser les endomorphismes pour lesquels aucun de ces polynômes n'est égal à 1.

Soit $M$ la matrice de $u$ dans la base de départ $\mathcal{B}$. Considérons la matrice $\tilde{M} = M - XI_n$ à coefficients dans $k[X]$. Cette matrice induit un morphisme de $k[X]$-module de $k[X]^n$ dans lui-même. On a alors le

**Théorème IV.62**

L'image de $\tilde{M}$ est égale au noyau de $\varphi$.

**Preuve**

Notons $\mathcal{I}$ l'image de $\tilde{M}$ et $\mathcal{N}$ le noyau de $\varphi$. Pour montrer que $\mathcal{I}$ est inclus $\mathcal{N}$, il suffit de montrer que chaque vecteur colonne de $\tilde{M}$ appartient à $\mathcal{N}$. Par exemple, le premier vecteur colonne de $\tilde{M}$ est

$$\begin{pmatrix} e_1^*(Me_1) - X \\ e_2^*(Me_1) \\ \vdots \\ e_n^*(Me_1) \end{pmatrix}.$$

Son image par $\varphi$ est donc : $(e_1^*(Me_1)e_1 - Me_1 + \sum_{j \geqslant 2} e_j^*(Me_1)e_j = Me_1 - Me_1 = 0$.

Ensuite, on raisonne par un argument de dimension, en considérant $k[X]^n/\mathcal{I}$ et $k[X]^n/\mathcal{N}$ comme des $k$-espaces vectoriels (et non plus comme des $k[X]$-modules). Nous savons déjà que $k[X]^n/\mathcal{N}$ est un $k$-espace vectoriel de dimension $n$. Le lecteur vérifiera qu'il suffit donc de montrer la même propriété pour $k[X]^n/\mathcal{I}$.
D'après le théorème **??**, il existe $P$ et $Q$ dans $\mathrm{GL}_n(k[X])$ et $Q_1, \cdots, Q_n$ tels que

$$P\tilde{M}Q^{-1} = \begin{pmatrix} Q_1 & & \\ & \ddots & \\ & & Q_n \end{pmatrix}.$$

Mais alors, $k[X]^n/\mathcal{I}$ est isomorphe comme $k$-espace vectoriel à $k[X]/(Q_1) \times \cdots \times k[X]/(Q_n)$. En particulier, sa dimension est la somme des degrés des $Q_i$, c'est-à-dire le degré de $\prod_{i=1}^n Q_i = \det(P\tilde{M}Q^{-1}) = \det(P).\det(Q)^{-1}\det(\tilde{M})$. Or, $\det(P)$ et $\det(Q)$ sont des éléments inversibles de $k[X]$, c'est-à-dire des éléments de $k$. Finalement, la dimension du $k$-espace vectoriel $k[X]^n/\mathcal{I}$ est égal au degré de $\det(\tilde{M})$ c'est-à-dire à $n$.

Le théorème IV.62 est un moyen très pratique de calculer les facteurs invariants d'une matrice donnée. En effet, on considère la matrice $M - XI_n$, on lui applique l'algorithme qui permet de montrer le théorème **??**. On calcule ainsi des polynômes unitaires $P_1, \cdots, P_n$ se divisant les uns les autres tels qu'il existe $P$

et $Q$ dans $\mathrm{GL}_n(k[X])$ tels que

$$P\tilde{M}Q^{-1} = \begin{pmatrix} P_1 & & \\ & \ddots & \\ & & P_n \end{pmatrix}.$$

Les polynômes $P_i$ non égaux à un ainsi obtenus sont les facteurs invariants de $M$.

**Exercice 5.** *Déterminer les facteurs invariants de la matrice :*

$$A = \begin{bmatrix} -17 & 8 & 12 & -14 \\ -46 & 22 & 35 & -41 \\ 2 & -1 & -4 & 4 \\ -4 & 2 & 2 & -3 \end{bmatrix}$$

# Chapitre 5

# Topologie en algèbre linéaire

## Sommaire

# 1 Polynômes

## 1.1 Résultant et Discriminant

Soit $K$ un corps quelconque. Etant donnés deux polynômes $P$ et $Q$, on cherche un moyen calculatoire de décider sur $P$ et $Q$ sont premiers entre eux. La clé est l'observation suivante : $P$ et $Q$ ont un facteur commun non constant si et seulement si

$$\exists (A,B) \in (K[X])^2 \text{ non nuls tels que } \begin{cases} \deg(A) < \deg(Q) \\ \deg(B) < \deg(P) \\ AP + BQ = 0. \end{cases}$$

Posons $p = \deg(P)$ et $q = \deg(Q)$ et considérons

$$\begin{array}{rccc} \varphi : & K_{q-1}[X] \times K_{p-1}[X] & \longrightarrow & K_{p+q-1}[X] \\ & (A,B) & \longmapsto & AP + BQ. \end{array}$$

Il s'agit de décider di $\varphi$ est injective ou pas. Or, les espaces d'arrivée et de départ sont de même dimension. On peut donc utiliser le déterminant.
**Attention.** Le déterminant de $\varphi$ n'existe pas, c'est-à-dire il dépend de choix de bases.
Munissons $K_{p+q-1}[X]$ de sa base canonique $\mathcal{C}$ et $K_{q-1}[X] \times K_{p-1}[X]$ de la base $\mathcal{B} := ((1,0), \ldots, (X^{q-1}, 0), (0,1), \ldots, (0, X^{p}$
Posons

$$M = \mathrm{Mat}_{\mathcal{CB}}(\varphi).$$

On définit alors

$$\mathrm{Res}(P,Q) := \det M.$$

On a

**Théorème V.63. Résultant**

Avec les notations ci-dessus,

$$P \wedge Q = 1 \quad \Longleftrightarrow \quad \mathrm{Res}(P,Q) \neq 0.$$

**Exercice 6.** *Décrire la matrice $M$ explicitement, en termes des coefficients de $P$ et $Q$.*

On définit aussi le disciment de $P$ par

$$\delta(P) := \mathrm{Res}(P, P').$$

On obtient le

**Théorème V.64. Résultant**

Avec les notations ci-dessus,

$$P \text{ est sans facteur carré} \quad \Longleftrightarrow \quad \delta(P) \neq 0.$$

## 1.2 Densité et polynômes à plusieurs variables

Ici $K = \mathbb{R}$ ou $\mathbb{C}$.

**Théorème V.65**

Soit $P \in K[x_1 \ldots, x_n]$ un polynôme non nul à $n$ variables.
Alors l'ensemble

$$U_P := \{(x_1, \ldots, x_n) \in K^n \ : \ P(x_1, \ldots, x_n) \neq 0\}$$

est un ouvert dense de $K^n$.

**Preuve**

Notons $Z_P$ le complémentaire de $U_P$. Il s'agit de montrer que $U_P$ rencontre toutes les boules ouvertes de $K^n$. Autrement dit qu'aucune boule de $K^n$ n'est incluse dans $Z_P$. Pour être plus précis on va montrer que si $Z_P$ contient une boule ouverte alors $P$ est nul.
Comme toutes les normes sont équivalentes, nous avons le choix. On choisit la norme $\max_i |x_i|$. Alors les boules sont des produits d'intervalles de $\mathbb{R}$ dans le cas réel, et de boules de $\mathbb{C}$ dans le cas complexe.
Il suffit donc de montrer que si $I_1, \dots, I_n$ sont des parties infinies de $K$ et que $P$ s'annule sur $I_1 \times \cdots \times I_n$ alors $P = 0$.
Nous allons montrer cela par récurrence sur $n$. Le cas $n = 1$ est bien connu : le seul polynôme qui a une infinité de racines est le polynôme nul. Supposons l'énoncé connu sur $K^{n-1}$. On écrit $P$ sous la forme

$$P = P_0 + P_1 x^n + \cdots P_d x_n^d$$

où $d \in \mathbb{N}$ et les $P_i \in K[x_1 \dots, x_{n-1}]$.
Fixons $(t_1, \dots, t_{n-1}) \in I_1 \times \cdots \times I_{n-1}$. Alors le polynôme en une variable $x_n$ suivant

$$P(t_1, \dots, t_{n-1}, x_n)$$

s'annule sur $I_n$. Il a donc une infinité de racines et il est nul. Donc ses coefficients $P_0(t_1, \dots, t_{n-1}), \dots, P_d(t_1, \dots, t_{n-1})$ sont tous nuls.
On vient de montrer que les polynômes $P_i$ sont nuls sur $I_1 \times \cdots \times I_{n-1}$. Donc par hypothèse de récurrence ils sont nuls. Mais alors $P = 0$.

## 2 Densité

Les résultats de la section précédente s'appliquent à l'étude des matrices.

**Théorème V.66**

1. Dans $\mathcal{M}_n(\mathbb{C})$, l'ensemble $\mathcal{D}_n(\mathbb{C})$ des matrices diagonalisables contient un ouvert dense.
2. L'ensemble des endomorphismes cycliques est dense dans $\mathcal{M}_n(K)$ pour $K = \mathbb{R}$ ou $\mathbb{C}$.
3. La classe de conjugaison de $M \in \mathcal{M}_n(\mathbb{C})$ est fermée si et seulement si $M$ est diagonalisable.
4. La classe de conjugaison de $M \in \mathcal{M}_n(\mathbb{C})$ contient 0 dans son adhérence si et seulement si $M$ est nilpotente.

**Preuve**

**Exercice 7.** *1. Le théorème décrit l'adhérence de $\mathcal{D}_n(\mathbb{C})$. Déterminer son intérieur.*
*2. Montrer que l'énoncé 1 est faux dans $\mathcal{M}_2(\mathbb{R})$.*

Densité de diagonalisable, cyclique.
Preuve de chiAB=chiBA et Cayley-Hamilton sur les complexes. Présentation de l'argument pour passer à tout corps via $\mathbb{Z}$.

## 3 Orbites

fermées, contenant 0 dans l'adhérence.

# 4 Exponentielle matricielle

Ici $K = \mathbb{R}$ ou $\mathbb{C}$.

## 4.1 Définition et calculs

Pour $A \in \mathcal{M}_n(K)$, on pose

$$e^A = \exp(A) = \sum_{n=0}^{\infty} \frac{A^n}{n!}.$$

Il faut d'abord montrer que la série converge. Pour cela choisissons une norme d'algèbre $\|\cdot\|$ sur $\mathcal{M}_n(K)$. Cela signifie que $\|MN\| \leqslant \|M\| \, \| \leqslant \|N\|$. Une norme d'opérateur ou la norme de Frobenus font l'affaire. Alors on a

$$\|\frac{A^n}{n!}\| \leqslant \frac{\|A\|^n}{n!}$$

et la série converge normalement. Par complétude de $\mathcal{M}_n(K)$, la série $\sum_{n=0}^{\infty} \frac{A^n}{n!}$ converge.
Les premières propriétés de l'exponentielle sont résumées dans le théorème suivant.

**Théorème V.67**

Soit $A, B \in \mathcal{M}_n(K)$ et $P \in \mathrm{GL}_n(K)$. On a :

1. $e^0 = I_n$.
2. $e^{-A}e^A = I_n$.
3. Si $AB = BA$ alors $e^A e^B = e^{A+B}$.
4. $Pe^A P^{-1} = e^{PAP^{-1}}$.
5. ${}^t e^A = e^{{}^t A}$.
6. $\det(e^A) = e^{\mathrm{tr}(A)}$.

**Preuve**

Les propriétés 4 et 5 sont vrai sur les sommes partielles $\sum_{n=0}^{\mathsf{N}} \frac{A^n}{n!}$. Elles s'obtiennent donc par passage à la limite quand $N \to +\infty$.
Pour la dernière, on peut penser à $A \in \mathcal{M}_n(\mathbb{C})$ (même si $K = \mathbb{R}$. Alors, il existe $Q \in \mathrm{GL}_n(\mathbb{C})$ et $T$ triangulaire supérieure telles que $A = QTQ^{-1}$. Notons $T_{ii}$ pour $i = 1, \ldots, n$ les éléments diagonaux de $T$. On a

$$\det(e^A) = \det(Q^{-1}e^A Q) = \det(e^{Q^{-1}AQ}) = \det(T) = \prod_{i=1}^{n} e^{T_{ii}} = e^{\sum_i^n T_{ii}} = e^{\mathrm{tr}(T)} = e^{\mathrm{tr}(A)}.$$

La propriété 1 est évidente. La deuxième est une conséquence de la troisième.
Supposons $AB = BA$. On a

$$\frac{(A+B)^n}{n!} = \frac{1}{n!}\sum_{k=0}^{n} \binom{n}{k} A^k B^{n-k} = \sum_{k=0}^{n} \binom{n}{k} \frac{A^k}{k!}\frac{B^{n-k}}{(n-k)!}.$$

Par ailleurs, par absolue convergence, $e^A e^B$ est le produit de Cauchy des deux séries. La propriété 3 en découle.

## 4.2 Surjectivité

**Théorème V.68**

L'application

$$\exp : \mathcal{M}_n(\mathbb{C}) \longrightarrow \mathrm{GL}_n(\mathbb{C})$$

est surjective.

**Preuve**

On rappelle que $\exp : \mathbb{C} \longrightarrow \mathbb{C}^*$ est surjective.
Soit $A \in \mathrm{GL}_n(\mathbb{C})$ et $A = D + N$ sa décomposition de Dunford. On pose $N' = D^{-1}N$ qui est nilpotente car $D$ et $N$ commuttent. On a $A = D(I_n + N')$.
La démonstration consiste à trouver deux matrices $U$ et $V$ telles que

1. $e^U = I_n + N'$ ;
2. $e^V = D$ ;
3. $UV = VU$.

On aura alors $e^{U+V} = e^U e^V = A$.
Pour $U$, on utilise la série de $\ln(1+n)$ :

$$U := \sum_{k=1}^{n-1} (-1)^{k-1} \frac{N'^k}{k}. \tag{4.1}$$

On diagonalise maintenant : on écrit $\Delta = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ et $D = Q\Delta Q^{-1}$. Pour chaque $i$, on choisit $\mu_i \in \mathbb{C}$ tel que $e^{\mu_i} = \lambda_i$ et on pose $\delta = \mathrm{diag}(\mu_1, \ldots, \mu_n)$. Alors $e^{Q\gamma Q^{-1}} = Qe^{\delta}Q^{-1} = Q\Delta Q^{-1} = D$. On pose donc $V = e^{Q\gamma Q^{-1}}$.
Dans la construction précédente, on peut supposer que

$$\lambda_i = \lambda_j \iff \mu_i = \mu_j.$$

Alors, $\mathrm{Com}(\delta) = Com(\Delta)$. Donc $\mathrm{Com}(D) = \mathrm{Com}(V)$.
On sait que $D$ et $N$ commuttent. Donc $V$ et $N$ commuttent. Mais d'après (4.1) $U$ est un polynôme en $N$. Donc $UV = VU$.

## 4.3 Régularité

**Théorème V.69**

1. L'application $\exp : \mathcal{M}_n(\mathbb{C}) \longrightarrow \mathrm{GL}_n(\mathbb{C})$ est différentiable.
2. La différentielle de exp en 0 est l'identité.
3. Pour $A \in \mathcal{M}_n(\mathbb{C})$, l'application $\mathbb{R} \longrightarrow \mathrm{GL}_n(\mathbb{C})$, $t \longmapsto e^{tA}$ est $C^\infty$ et sa dérivée est $Ae^{tA}$.

**Preuve**

**Théorème V.70**

L'application

$$\begin{array}{ccc} S_n(\mathbb{R}) & \longrightarrow & S_n^{++}(\mathbb{R}) \\ S & \longmapsto & e^S \end{array}$$

est un homéomorphisme.

**Preuve**

*Bien définie.* Comme exp et transposé commuttent, l'image est constituée de matrices symétriques. Soit $S \in S_n(\mathbb{R})$. Alors $S = O\Delta O^{-1}$ avec $\Delta$ diagonale et $O$ orthogonale. On en déduit que $e^S = Oe^{\Delta}O^{-1}$ a des valeurs propres strictement positives. Donc $e^S \in S_n^{++}$.

*Surjectivité.* Soit $T \in S_n^{++}$. Alors $T = O\Delta O^{-1}$ avec $\Delta$ diagonale et $O$ orthogonale. Alors $\Delta_{ii} > 0$. Posons $\delta = \mathrm{diag}(\ln(\Delta_{11}), \ldots, \ln(\Delta_{nn}))$. On a $e^{\delta} = \Delta$ et $e^{O\delta O^{-1}} = T$. De plus, comme $O$ est orthogonale $S := O\delta O^{-1}$ est symétrique.
Remarquons que $\mathrm{Com}(S) = \mathrm{Com}(T)$.

*Injectivité.* Soit $T \in S_n^{++}$ et $S$ l'antécédent construit ci-dessus. Soit $S'$ un autre antécédent. On a : $ST = TS$, $S'T = TS'$ et $\mathrm{Com}(T) = \mathrm{Com}(S)$. Donc $SS' = S'S$. Mais alors, $S, S'$ et $T$ sont simultanément diagonalisables. Or pour des matrices réelles et diagonale $e^s = e^{s'}$ implique $s = s'$. On en déduit que $S = S'$.

*Continuité.* Restriction de exp qui est continue sur $\mathcal{M}_n(\mathbb{C})$.

*Continuité de la réciproque.* On utilise le critère séquenciel. Soit $(A_k)_{k\in\mathbb{N}}$ une suite d'éléments de $S_n^{++}$ qui tend vers $A \in S_n^{++}$. Soit $B_k \in S_n$ telle que $e^{B_k} = A_k$ et $B \in S_n$ telle que $e^B = A$. Montrons que la suite $B_k$ tend vers $B$.
Par continuité de exp, la seule valeur d'adhérence de la suite $(B_k)_{k\in\mathbb{N}}$ est $B$. Il suffit donc de montrer que la suite est bornée.
Pour cela choisissons la norme triple associée à la norme euclidienne canonique sur $\mathbb{R}^n$. Pour une matrice symétrique, celle-ci vérifie

$$\forall M \in S_n(\mathbb{R}) \qquad \rho(M) = |||M|||_2 \tag{4.2}$$

Comme $(A_k)_{k\in\mathbb{N}}$ converge, elle est bornée. Par continuité de l'inverse, $({A_k}^{-1})_{k\in\mathbb{N}}$ converge aussi (vers $A^{-1}$) et est bornée. On en déduit avec (4.2) qu'il existe $\epsilon > 0$ et $M \in \mathbb{R}^+$ telle que

$$\forall k \in \mathbb{N} \qquad \mathrm{Sp}(A_k) \subset [\epsilon; M].$$

Mais alors,

$$\forall k \in \mathbb{N} \qquad \mathrm{Sp}(B_k) \subset [\ln(\epsilon); \ln(M)].$$

A nouveau avec (4.2), on en déduit que la suite $(B_k)_{k\in\mathbb{N}}$ est bornée.

## 4.4 Calcul de $e^A$ et $A^n$

Dans cette section, nous allons voir un algorithme pour calculer :

1. $A^n$ avec $n$ une variable ;
2. $e^{tA}$ ;
3. la décomposition de Dunford de $A$.

Pour cela nous supposons que nous avons scindé un polynôme annulateur de $A$.
Pour $\lambda \in k^*$ et $j \in \mathbb{N}$, on définit la suite

$$\theta_\lambda^j(n) = n^j \lambda^n.$$

On définit aussi

$$\theta_0^j(n) = \delta_n^j,$$

où $\delta$ est le symbole de Kronecker. Ainsi

$$\begin{aligned}\theta_0^0 &= 1\,0\,0\,0\,0\,0\ldots\\ \theta_0^1 &= 0\,1\,0\,0\,0\,0\ldots\\ \theta_0^2 &= 0\,0\,1\,0\,0\,0\ldots\end{aligned}$$

**Lemme V.71**

La famille $(\theta_\lambda^j)_{j\in\mathbb{N},\,\lambda\in k}$ de suites de $k^{\mathbb{N}}$ est libre.

*Démonstration.* La famille des $\theta_0^j(n)$ étant libre et nulle pour $n$ assez grand, il suffit de montrer que, si

$$\alpha_1\theta_{\lambda_1}^{j_1}(n)+\cdots+\alpha_s\theta_{\lambda_s}^{j_s}(n)=0$$

pour $n$ assez grand, alors chaque $\alpha_i$ est nul. Ici les $\alpha_i$ sont des scalaires quelconques et les paires $(\lambda_i,j_i)$ sont deux à deux distinctes.
Soit $j_{max}$ la valeur maximale des $j_i$ qui apparaissent. On montre cette propriété par récurence sur

$$(j_{max},\sharp\{i\,:\,j_i=j_{max}\}).$$

Si $j_{max}=0$, tous les $j_i$ sont nuls. L'énoncé se montre alors par récurence sur $s$ en considérant $E(n+1)-\lambda_1E(n)$, où $E(n)=\alpha_1\theta_{\lambda_1}^0(n)+\cdots+\alpha_s\theta_{\lambda_s}^0(n)$.
Quitte à renuméroter supposons que $j_{max}=j_1$. Considérons à présent $E(n+1)-\lambda_1E(n)$. On remarque que $\theta_\lambda^j(n+1)-\lambda\theta_\lambda^j(n)$ est une combinaison linéaire des $\theta_\lambda^{j'}$ pour $j'<j$. De même, $\theta_\lambda^j(n+1)-\lambda'\theta_\lambda^j(n)$ est une combinaison linéaire des $\theta_\lambda^{j'}$ pour $j'\leqslant j$ si $\lambda\neq\lambda'$.
Ainsi, $E(n+1)-\lambda_1E(n)$ a la même forme que $E(n)$ avec

1. un des coefficients égal à $\alpha_1j_1$ ;
2. soit un $j_{max}$ strictement inférieur, soit $\sharp\{i\,:\,j_i=j_{max}\}$ strictement inférieur.

L'hypothèse de récurrence implique donc que $\alpha_1=0$. Contradiction. □

**Théorème V.72**

Soit $k$ un corps de caractéristique nulle et $M$ une matrice carré à coefficients dans $k$. Soit $\mu=\prod_{i=1}^s(X-\lambda_i)^{\alpha_i}$ avec les $\lambda_i$ 2 à 2 distincts et $\alpha_i\in\mathbb{N}^*$. On suppose que $\mu(M)=0$ mais pas nécessairement que $\mu$ est le polynôme minimal de $M$.
Considérons l'ensemble $\mathcal{E}$ des paires $(i,j)\in\mathbb{N}^2$ telles que $1\leqslant i\leqslant s$ et $0\leqslant j<\alpha_i$. Alors, il existe une unique famille de matrices $(A_{(i,j)})_{(i,j)\in\mathcal{E}}$ indexée par $\mathcal{E}$ telle que

$$M^n=\sum_{(i,j)\in\mathcal{E}}A_{(i,j)}\theta_{\lambda_i}^j(n).$$

*Remarque.* Si 0 n'est pas une racine de $\mu$, l'énoncé dit que chaque coefficient de $M^n$ est de la forme

$$\sum_{i=1}^s P_i(n)\lambda_i^n,$$

où $P_i$ est un polynôme de degré au plus $\alpha_i-1$.
Si $0=\lambda_1$ est une racine de $\mu$ d'ardre $\alpha_0$, l'énoncé dit que chaque coefficient de $M^n$, **pour** $n\geqslant\alpha_0$, est de la forme

$$\sum_{i=1}^s P_i(n)\lambda_i^n,$$

où $P_i$ est un polynôme de degré au plus $\alpha_i-1$.

*Démonstration.* L'unicité est une conséquence directe de la liberté des $\theta_\lambda^j$.
Il existe une matrice inversible $P$ telle que $B=PAP^{-1}$ est diagonale par blocs avec des blocs de la forme $\lambda_iI_{k_i}+N_i$, avec $N_i$ nilpotente d'ordre au plus $\alpha_i$. Un calcul par blocs montre facilement que $B^n$ est de la forme cherhcée. On remarque alors que chaque coefficient de $P^{-1}BP$ est une combinaison linéaire des coefficients de $B$. Le théorème suit. □

*Remarque.* Pour la démonstration de l'existence, on peut aussi théoriser la technique bien connue de calcul des puissances d'une matrice par division euclidienne de $X^n$ par un polynôme annulateur.
Soit $R_n$ le reste de la division euclidienne de $X^n$ par $\mu_M$. On écrit

$$X^n = Q_n\mu_M + R_n,$$

avec $Q_n \in k[X]$. Soit $\varphi : k_{d-1}[X] \longrightarrow k^d$, qui envoie $R$ sur $(R^{(j)}(\lambda_i))$ pour $i$ et $j$ comme ci-dessus. Alors $\varphi$ est une application linéaire entre deux espaces vectoriels de dimension $d$. De plus, si $P$ est dans le noyau de $\varphi$ alors $\lambda_i$ est une racine d'ordre au moins $\alpha_i$ de $P$. Comme $\sum_i \alpha_i = d > \deg(P)$, cela impose que $P$ est nul. Ainsi $\mathrm{Ker}\varphi = \{0\}$. Finalement $\varphi$ est un isomorphisme.
On remarque alors que, pour tout $i < \alpha_i$, on a

$$\varphi(R_n) = \varphi(X^n) = n(n-1)\dots(n-j+1)\lambda_i^{n-j} \in \mathrm{Vect}\{\theta^k_{\lambda_i} \,:\, 0 \leqslant k < \alpha_i\}.$$

On en déduit que les coefficients de $R_n$ qui s'obtiennent en calculant $\varphi^{-1}(\varphi(R_n))$ sont dans l'espace vectoriel voulu.

## 4.5 Application au calcul de la décomposition de Dunford

Remarquons tout d'abord que la formule (**??**) montre une version faible du thèorème dans laquelle il faut prendre $j \leqslant i_N = \max \alpha_1, \dots, \alpha_s$ :

*Pour tout $n \geqslant i_N$, chaque coefficient de de la suite $M^n$ est une combinaisons linéaires des suites $\theta^j_{\lambda_i}$ pour $1 \leqslant i \leqslant s$ et $0 \leqslant j < i_N$.*

En effet, pour tout $n \geqslant i_N$, on a

$$M^n = \sum_{k=0}^{i_N-1} \binom{n}{k} P\Delta^{n-k}P^{-1}N^k.$$

Or chaque coefficient de $\Delta^{n-k}$ est nul ou multiple d'un $\theta^0_{\lambda_i}(n)$. Alors, chaque coefficient de $P\Delta^{n-k}P^{-1}$ est une combinaison linéaire des $\theta^0_{\lambda_i}(n)$. De plus, $\binom{n}{k}$ est égal au polynôme $n(n-1)\dots(n-k+1)$ à une constante multiplicative près. Ainsi, les coefficients de $\binom{n}{k} P\Delta^{n-k}P^{-1}N^k$ sont des combinaisons linéaires des suites $\theta^j_{\lambda_i}$ pour $1 \leqslant i \leqslant s$ et $0 \leqslant j < i_N$ :

$$\forall n \geqslant i_N \qquad M^n = \sum_{\substack{1 \leqslant i \leqslant s \\ 0 \leqslant j < i_N}} A_{(i,j)}\theta^j_{\lambda_i}(n).$$

Une lecture attentive de la preuve ci-dessus monte qu'en ne gardant dans (**??**) que les termes correspondant à $k = 0$, on obtient

$$D^n = P\Delta^n P^{-1} = \sum_{1 \leqslant i \leqslant s} A_{(i,0)}\theta^0_{\lambda_i}(n).$$

Donc

$$D = \sum_{1 \leqslant i \leqslant s} A_{(i,0)}\lambda_i.$$

On en déduit l'algorithme suivant du calcul de la décomposition de Dunford :

1. `On calcule le polynôme caractéristique` $\mu_M$ `de` $M$ `de manière scindé.`
2. `On calcule les` $I, A, \dots, A^{taille-1}$.
3. `On écrit` $A^n$ `comme dans le théorème V.72 avec des matrices indéterminées` $A_{(i,j)}$.
4. `On calcule ces indéterminées en résolvant des systèmes (on utilise les premières valeurs de` $A^n$ `calculées).`

5. On écrit les coefficients de $A^n$ (pour $n$ assez grand) sous la forme $\sum_i P_i \lambda_i^n$. Alors

$$D = \sum_i P_i(0)\lambda_i^n.$$

La remarque finale est que par liberté des suites $\theta_i^j$, on peut identifier la formule du théorème à (**??**). On obtient

$$D = \sum_i A_i^1 \lambda_i.$$

Ainsi le calcul de $M^n$ permet de retrouver la décomposition de Dunford et la boucle est bouclée.

**Exemple.** Considérons

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

On écrit

$$M^n = \begin{pmatrix} 1 & a_n & b_n \\ 0 & 2^n & c_n \\ 0 & 0 & 2^n \end{pmatrix}.$$

D'après le théorème la suite $a_n = a^1 + a^2 2^n + a^3 n 2^n$. On calcule alors $a_n$ en résolvant le système $a_0 = 0$, $a_1 = 1$ et $a_2 = 2$. De même, on peut calculer $b_n$ et $c_n$. On trouve :

$$M^n = \begin{pmatrix} 1 & -1+2^n & 1-2^n+\frac{1}{2}n2^n \\ 0 & 2^n & \frac{1}{2}n2^n \\ 0 & 0 & 2^n \end{pmatrix}.$$

On en déduit que

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

puis que

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

# 5 Décomposition polaire

Pour $M \in \mathcal{M}_n(\mathbb{R})$, on pose

$$\mathrm{Com}(M) := \{N \in \mathcal{M}_n(\mathbb{R}) \ : \ MN = NM\}.$$

**Théorème V.73. Décomposition polaire de $\mathrm{GL}_n(\mathbb{R})$**

Soit $n \in \mathbb{N}^*$. L'application suivante est un homéomorphisme

$$\begin{array}{ccc} S_n^{++} \times O_n(\mathbb{R}) & \longrightarrow & \mathrm{GL}_n(\mathbb{R}) \\ (S, O) & \longmapsto & SO \end{array}$$

est un homéomorphisme.

**Preuve**

Analyse. Soit $A \in \mathrm{GL}_n(\mathbb{R})$. On cherche à écrire $A = SO$. Supposons cela fait. La première astuce

consiste à éliminer $O$ de l'équation :

$$A\,{}^tA = SO\,{}^tO\,{}^tS = S^2.$$

On va commencer par résoudre cette équation d'inconnue $S$.
Synthèse. Pour tout $X \in \mathbb{R} - \{0\}$, on a

$${}^tX\,A\,{}^tAX = \|{}^tAX\|^2 > 0,$$

car $A$ est inversible. Donc $A\,{}^tA$ est symétrique définie positive. Par le théorème spectral, il existe $P \in O_n(\mathbb{R})$ et $\lambda_1, \dots, \lambda_n \in \mathbb{R}^{*,+}$ tels que

$$PA\,{}^tA\,{}^tP = \mathrm{diag}(\lambda_1, \dots, \lambda_n).$$

On pose alors

$$S = {}^tP\mathrm{diag}(\sqrt{\lambda_1}, \dots, \sqrt{\lambda_n})P.$$

Notons $\Delta$ la dernière matrice diagonale. Cette matrice est bien définie positive. On pose alors $O = S^{-1}A$. Alors

$$O\,{}^tO = S^{-1}A\,{}^tAS^{-1} = {}^tP\Delta^{-1}P\,{}^tP\Delta^2P\,{}^tP\Delta^{-1}P = I_n.$$

Donc $O$ est orthogonale et $A = SO$. Ceci achève la preuve de la surjectivité.

Soit $S$ la matrice ci-dessus. Montrons que

$$\mathrm{Com}(S) = \mathrm{Com}(S^2). \tag{5.1}$$

Ces deux ensembles sont des sous-espaces vectoriels de $\mathcal{M}_n(\mathbb{R})$. De plus, $\mathrm{Com}(S) \subset \mathrm{Com}(S^2)$. Enfin, ces deux espaces vectoriels ont la même dimension. En effet, la dimension de $\mathrm{Com}(S^2)$ ne dépend que de l'ensemble des couples $(i, j)$ tels que $\lambda_i = \lambda_j$.

Montrons l'injectivité. Supposons que $(S', O')$ vérifie aussi $A = S'O'$. On a alors $S^2 = S'^2$. En particulier, $S'$ commute à $S^2$. D'après (5.1), $S'$ commute à $S$. Donc $S$ et $S'$ sont simultanément diagonalisables : il existe $Q \in \mathrm{GL}_n(\mathbb{R})$ (et même $O_n(\mathbb{R})$) telle que

$$S = Q\mathrm{diag}(\mu_1, \dots, \mu_n)Q^{-1} \qquad S' = Q\mathrm{diag}(\mu'_1, \dots, \mu'_n)Q^{-1}.$$

Mais alors, $S^2 = S'^2$ donne $\mu_i^2 = (\mu'_i)^2$ pour tout $i$. Comme $S$ et $S'$ sont définie positives, les $\mu_i > 0$ et $\mu'_i > 0$. D'où l'on déduit que $\mu_i = \mu'_i$, puis $S = S'$. Alors $O = S^{-1}A = S'^{-1}A = O'$. D'où l'unicité.

Montrons la continuité de la réciproque par le critère séquenciel. Soit donc $(A_n)_{n\in\mathbb{N}}$ une suite de matrices de $\mathrm{GL}_n(\mathbb{R})$ qui converge vers une matrice $A$. Soit $(S_n, O_n)$ et $(S, O)$ les décompositions polaires de $A_n$ et $A$. Il s'agit de montrer que $(S_n, O_n)$ tend vers $(S, O)$.
Soit $O'$ une valeur d'adhérence de la suite $O_n$ dans $O_n(\mathbb{R})$, disons limite de la suite extraite $O_{\phi(n)}$. Alors $S_{\phi(n)} = A\,{}^tO_{\phi(n)}$ converge. Notons $S'$ sa limite. Comme l'ensemble des matrices symétriques positive est fermé, $S'$ est symétrique positive. Par ailleurs, $S' = {}^tO'A$ est inversible. Donc $S'$ est définie positive. Mais alors, $(S', O')$ est une décomposition polaire de $A$. Par unicité, $(S', O') = (S, O)$.
Nous venons de montrer que la suite $O_n$ du compact $O_n(\mathbb{R})$ a une seule valeur d'adhérence qui est $O$. Donc $O_n$ tend vers $O$. Mais alors, le raisonnement précédent montre que $S_n$ tend vers $S$.

**Corollaire V.74**

L'ensemble $S_n^{++}$ est un système complet de représentants pour l'action

$$\begin{array}{ccc} O_n(\mathbb{R}) \times \mathrm{GL}_n(\mathbb{R}) & \longrightarrow & \mathrm{GL}_n(\mathbb{R}) \\ (O, M) & \longmapsto & MO^{-1}. \end{array}$$

Soit

$$A = \begin{pmatrix} 1 & 2 & 0 & -1 \\ 2 & 6 & 0 & -1 \\ -1 & 4 & -1 & 7 \\ -2 & -4 & -3 & 8 \end{pmatrix}$$

1. Calculer la décomposition $LU$ de la matrice $A$.
2. En déduire, les solutions du système

$$Ax = \begin{pmatrix} -3 \\ -10 \\ -10 \\ 3 \end{pmatrix},$$

d'inconnu $X \in \mathbb{R}^4$.

Pour info on a $A = LU$ avec

$$L = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ -1 & 3 & 1 & 0 \\ -2 & 0 & 3 & 1 \end{pmatrix} \quad \text{et} \quad U = \begin{pmatrix} 1 & 2 & 0 & -1 \\ 0 & 2 & 0 & 1 \\ 0 & 0 & -1 & 3 \\ 0 & 0 & 0 & -3 \end{pmatrix},$$

et

$$X = \begin{pmatrix} 1 \\ -2 \\ 1 \\ 0 \end{pmatrix}.$$