Dimension d'un espace vectoriel. Rang. Exemples et applications

prérequis : les notions de base sur les espaces vectoriels, matrices équivalentes, semblables.
Remarques d'ordre général : il est clair qu'il faut commencer par un premier paragraphe sur la théorie de la dimension : il faudra veiller à le rédiger avec soin afin de respecter un ordre logique parmi les énoncés. Parmi les passages obligés, il faut bien entendu donner des exemples et des applications sur les applications linéaires. Il parait difficile de ne pas évoquer la signature, la multiplicativité des degrés dans les extensions de corps, la géométrie affine. Il peut être intéressant de se lancer sur les codes correcteurs d'erreurs.

Table des matières

1. Théorie de la dimension

1.1. Familles libres, génératrices et bases. —

Définition 1.1. — Une famille $\{(e_i)_{i\in I}\}$ de vecteurs d'un espace vectoriel E est dit *libre* si pour tout famille $(\lambda_i)_{i\in I} \in \mathbb{K}^{(I)}$ à support fini

$$\sum_{i\in I} \lambda_i e_i = 0 \Rightarrow \forall i \in I,\ \lambda_i = 0.$$

Elle est dite *génératrice* si $< \{e_i :\ i \in I\} >= E$, i.e. si tout vecteur de E peut s'écrire comme une combinaison linéaire à support fini des e_i.

Remarque : la famille $(X^i)_{i\in\mathbb{N}} \in \mathbb{K}[X]$ est libre et génératrice. En revanche elle n'est pas génératrice dans $K[[X]]$.

Remarque : la famille $(e_i)_{i\in I}$ est dite *liée* si elle n'est pas libre, i.e. s'il existe une famille $(\lambda_i)_{i\in I} \in \mathbb{K}^{(I)}$ non nulle telle que $\sum_{i\in I} \lambda_i e_i = 0$.

Définition 1.2. — Une famille $(e_i)_{i\in I}$ de vecteurs de E est *une base* si elle est libre et génératrice.

1.2. Théorème de la base incomplète. —

Théorème 1.3. — *Soient $\{f_1, \cdots, f_p\}$ une famille libre de vecteurs et $\{g_1, \cdots, g_q\}$ une famille génératrice de E. Il existe alors un entier $n \geq p$ et une base $\{e_1, \cdots, e_n\}$ de E telle que $e_i = f_i$ pour $1 \leq i \leq p$ et $e_j \in \{g_1, \cdots, g_q\}$ pour $p+1 \leq j \leq n$.*

Remarque : ainsi tout espace contenant une famille génératrice finie admet une base.

Corollaire 1.4. — *Soit E un espace vectoriel muni d'une base de cardinal n. Alors toute famille de cardinal strictement supérieur à n est liée.*

Remarque : on en déduit alors que le cardinal de toute base de E est toujours le même ; on l'appelle *la dimension* de E.
Application : en géométrie affine, une droite (resp. un plan...) est un espace vectoriel de dimension 1 (resp. 2, ...) ayant perdu son origine.

1.3. Sous espace vectoriel, espace quotient, dual. —

Proposition 1.5. — *Tout sous-espace vectoriel F de E est de dimension inférieure ou égale à celle de E avec égalité si et seulement si $F = E$.*

Définition 1.6. — On appelle hyperplan d'un espace vectoriel E de dimension finie, tout sous-espace de dimension $n-1$.

Remarque : en dimension infinie, un hyperplan est un sous-espace tel que E/F est de dimension 1. La dimension de l'espace quotient E/F s'appelle *la codimension* de F dans E.
Remarque : la dimension de $E \times F$ est le produit des dimensions de E et F.
Remarque : toute famille libre est de cardinal $\leq n$ avec égalité si et seulement si c'est une base.

Proposition 1.7. — *Soient E un espace vectoriel de dimension finie et F un sous-espace de E. Alors F et E/F sont de dimension finie et* $\dim_{\mathbb{K}} E = \dim_{\mathbb{K}} F + \dim_{\mathbb{K}} E/F$.

Application : soit $u : E \longrightarrow F$ une application linéaire avec F (resp. E) de dimension finie. Alors $\operatorname{Im} u$ (resp. $\operatorname{Ker} u$) est de dimension finie appelée le rang de u. Si E et F sont de dimension finie alors $\operatorname{rg}(u) + \dim_{\mathbb{K}} \operatorname{Ker}(u) = \dim_{\mathbb{K}} E$.

Corollaire 1.8. — *Les classes d'équivalences des $u : E \longrightarrow F$ sont paramétrées par le rang.*

Aspect algorithmique : le rang d'une famille donnée se calcule matriciellement en pivotant à droite sur la matrice dont les vecteurs colonnes sont ceux de la famille écrits dans une base fixée.

Proposition 1.9. — *Soient F et G deux sous-espaces vectoriels d'un espace de dimension finie E. Alors* $\dim_{\mathbb{K}}(F+G) = \dim_{\mathbb{K}} F + \dim_{\mathbb{K}} G - \dim_{\mathbb{K}}(F \cap G)$.

Application : deux droites distinctes du plan projectif s'intersectent toujours en un point.
Remarque : le produit tensoriel de deux espaces de dimension finie est de dimension finie égale au produit des dimensions.

Corollaire 1.10. — *Soient $k \subset K \subset E$ une extension de corps. Les degrés sont reliés par la formule $[E : k] = [E : K].[K : k]$.*

Lemme 1.11. — *Un espace vectoriel V sur un corps infini k n'est pas réunion finie de sous-espaces stricts $V_1, \cdots, V_t$.*

Preuve : C'est clair pour $t = 1$ et on suppose, par hypothèse de récurrence, que $t \geq 2$ et le résultat établi pour $t - 1$. Ainsi il existe $u, v \in V$ tels que $u \notin V_t$ et $v \notin V_1 \cup \cdots \cup V_{t-1}$. Si on avait $V = V_1 \cup \cdots \cup V_t$ alors $v \in V_t$ et comme l'ensemble des $x_\lambda := u + \lambda v$ pour $\lambda \in k$ est infini, il existe $\lambda \neq \mu$ tels que x_λ et x_μ appartiennent au même V_j. On ne peut avoir $j = t$ car sinon on aurait $u \in V_t$; donc $j < t$ et V_j contient $x_\lambda - x_\mu = (\lambda - \mu)v$ et donc v aussi ce qui n'est pas.
Application au théorème de l'élément primitif : il s'agit de montrer que si K/k est une extension séparable de degré fini, alors K admet un élément primitif ζ sur k, i.e. $K = k[\zeta]$.

1) Si k est un corps fini alors $K^\times$ est cyclique et $K = k[\zeta]$ pour tout générateur ζ de $K^\times$.

2) Supposons à présent que k est infini et notons $n = [K : k]$. Pour Ω une clôture algébrique de K, K/k étant séparable, il existe des k-isomorphismes $K \to \Omega$ deux à deux distincts $\tau_1, \cdots, \tau_n$ de sorte que pour tous $i \neq j$, $\mathrm{Ker}(\tau_i - \tau_j)$ est un sous-espace strict de K. D'après le lemme il existe $x \in K$ n'appartenant à aucun des $\mathrm{Ker}(\tau_i - \tau_j)$ de sorte que les $\tau_i(x)$ sont deux à deux distincts. Comme ce sont des racines dans Ω de $\mu_{x,k}$ de sorte que

$$n \leq [k[x] : k] \leq [K : k] = n$$

et donc $K = k[x]$.

Pour E de dimension finie n, son espace dual $E^* = \hom_{\mathbb{K}}(E, \mathbb{K})$ est aussi de dimension n. Pour $(e_i)_{i=1,\cdots,n}$ une base de E, on peut définir sa base duale $(e_i^*)_{i=1,\cdots,n}$ définie par $e_j^*(e_i) = \delta_{i,j}$.

Définition 1.12. — Soit F un sous-espace vectoriel de E alors $F^\perp = \{\psi \in E^* : \ \forall f \in F, \ \psi(f) = 0\}$.

Proposition 1.13. — *Si F est de dimension r alors $F^\perp$ est de dimension $n - r$.*

Aspect algorithmique : on écrit la matrice formée des vecteurs colonnes, dans une base fixée de E, d'une famille génératrice de F auxquels on rajoute le vecteur colonne ${}^t(x_1, \cdots, x_n)$. On pivote alors à droite jusqu'a faire apparaitre les équations linéaires.

1.4. Un mot sur la dimension infinie. —

Proposition 1.14. — *Soit E un $\mathbb{K}$-espace vectoriel et V, W_1, W_2 des sous-espaces tels que $V \cap W_1 = \{0\}$ et $V + W_2 = E$. Il existe alors un supplémentaire W de V contenu dans W_2 et contenant W_1.*

Preuve : Considérons l'ensemble $\mathcal{E}$ des sous-espaces de E contenant W_1 et contenus dans W_2 ; $\mathcal{E}$ n'est pas vide car $W_1 \in \mathcal{E}$. En outre $\mathcal{E}$ est partiellement ordonné par la relation d'inclusion et est inductif. Rappelons que cela signifie que toute chaine totalement ordonnée admet un majorant : ici pour une telle chaine, un majorant est simplement donné par la réunion qui est clairement un sous-espace.

D'après le lemme de Zorn, $\mathcal{E}$ admet un élément maximal, notons le W. Par définition on a donc $W \cap V = \{0\}$ et $W_1 \subset W \subset W_2$. Il reste alors à prouver que $V + W = E$; tout élément $x \in E$ s'écrit $x = v + w_2$ avec $v \in V$ et $w_2 \in W_2$. Si $w_2 \in W$ alors c'est gagné, sinon on considère le sous-espace engendré X par W et w_2. Par maximalité de W, X $\mathcal{E}$ de sorte qu'il existe $0 \neq y \in X \cap V$; ainsi $y = w + \lambda w_2 \in V$ et donc $y \in W \cap V$ ce qui n'est pas.
Remarque : le lecteur notera bien l'utilisation essentielle du lemme de Zorn qui rappelons le est équivalent à l'axiome du choix. Ainsi notre preuve n'est pas du tout constructive.

Corollaire 1.15. — *Tout sous-espace V de E admet un supplémentaire.*

Corollaire 1.16. — *Tout espace vectoriel non nul admet une base.*

Preuve : Considérons l'ensemble $\mathcal{A}$ des familles libres de E ; c'est clairement un ensemble non vide, partiellement ordonné par l'inclusion et inductif. D'après le lemme de Zorn, il possède un élément maximal qui est donc une famille libre maximal c'est donc nécessairement une famille génératrice et donc une base.
Remarque : le lecteur pourra s'exercer sur $\mathbb{K}^{\mathbb{N}}$ en vérifiant que toute base est nécessairement non dénombrable.

Corollaire 1.17. — ***(Théorème de la base incomplète)***
Soit $(e_i)_{i\in I}$ une partie génératrice de E. Soit $J \subset I$ tel que $(e_i)_{i\in J}$ est libre, il existe alors $J \subset K \subset I$ tel que $(e_i)_{i\in K}$ soit une base.

Preuve : On considère l'ensemble $\mathcal{A}$ des familles libres $(e_i)_{i\in A}$ pour $A \subset I$. C'est un ensemble non vide partiellement ordonné par l'inclusion et clairement inductif. D'après le lemme de Zorn, $\mathcal{A}$ possède une élément maximal K ; comme précédement $(e_i)_{i\in K}$ est libre et génératrice par maximalité de K.
Remarque : citons enfin le cas des espaces de Hilbert, i.e. des espaces hermitiens, au sens du paragraphe sur l'algèbre bilinéaire, qui sont complets, i.e. toutes les suites de Cauchy sont convergentes.

Définition 1.18. — On dit que $(e_i)_{i\in I}$ est *une base de Hilbert* d'un espace de Hilbert H si et seulement si :
- c'est une base orthonormée, i.e. $< e_i, e_j >= \delta_{i,j}$;
- la famille est complète au sens que pour tout $x \in H$ il existe $(\lambda_i)_{i\in I}$ telle que $\sum_{i\in I} \lambda_i e_i = x$, i.e. la série correspondante dans H est convergente de limite x.

Remarque : le lecteur vérifiera aisément qu'une base au sens de Hilbert n'est pas une base au sens classique, cf. par exemple les espaces L^2.

2. Premiers exemples

2.1. Noyaux emboités et invariants de similitude. — Soient E un $\mathbb{K}$-espace vectoriel de dimension finie n et $u \in \mathcal{L}(E)$. Pour tout $\lambda \in \mathbb{K}$ et $r \geq 1$, on note

$$K_r(\lambda) := \mathrm{Ker}(u - \lambda \mathrm{Id})^r \quad \text{et} \quad I_r(\lambda) := \mathrm{Im}(u - \lambda \mathrm{Id})^r,$$

et on note $dK_r(\lambda) := \dim_{\mathbb{K}} K_r(\lambda)$ et $dI_r(\lambda) := \dim_{\mathbb{K}} I_r(\lambda)$. On pose aussi $dK_0(\lambda) = 0$ et $dI_0(\lambda) = n$.

Proposition 2.1. — *La suite $dK_r(\lambda)$ (resp. $dI_r(\lambda)$) est tout d'abord strictement croissante (resp. décroissante) puis stationnaire à partir d'un indice r_0 (resp. le même indice r_0). Par ailleurs la suite*

$$\delta_r(\lambda) := dK_r(\lambda) - dK_{r-1}(\lambda)$$

pour $r \geq 1$ est décroissante jusqu'au rang r_0 puis stationnaire égale à 0.

Preuve : La croissance de $dK_r(\lambda)$ (resp. la décroissance de $dI_r(\lambda)$) est claire puisque $K_{r-1}(\lambda) \subset K_r(\lambda)$ (resp. $I_r(\lambda) \subset I_{r-1}(\lambda)$). Par ailleurs si $K_r(\lambda) = K_{r+1}(\lambda)$ alors pour $x \in K_{r+2}(\lambda)$ on a $u(x) \in K_{r+1}(\lambda) = K_r(\lambda)$ et donc $u^{r+1}(x) = 0$ i.e. $x \in K_{r+1}(\lambda)$. Le théorème du rang donne enfin que l'indice où $dI_r(\lambda)$ stationne est le même que celui de $dK_r(\lambda)$. On remarque ensuite que u induit une injection

$$K_{r+1}(\lambda)/K_r(\lambda) \hookrightarrow K_r(\lambda)/K_{r-1}(\lambda)$$

ce qui implique la décroissance de la suite $\delta_r(\lambda)$ jusque 0.

La suite $\delta_r(\lambda)$ définit une partition du sous-espace caractéristique $E_\lambda(u)$ que l'on peut représenter sous la forme d'un tableau de Young. Ainsi pour $\lambda = 0$ et $d_r := dK_r(0 =$, le tableau de Young associé à u est tel que ses colonnes sont de taille $d_i - d_{i-1}$; ses lignes définissent alors une partition $(n_1 \geq n_2 \geq \cdots \geq n_s)$ de $n = n_1 + \cdots + n_s$ qui correspond à la forme de Jordan de u.

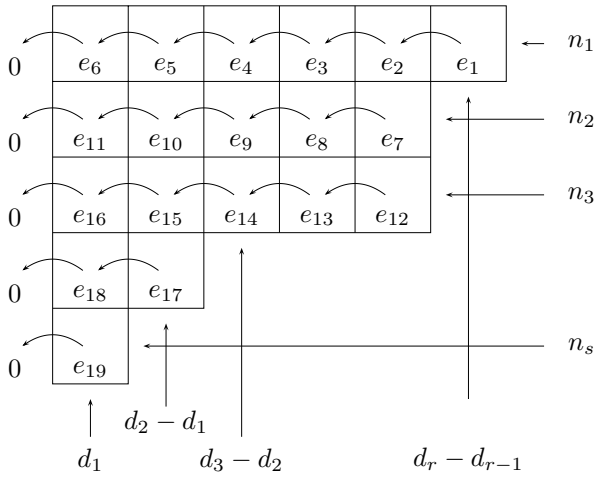

FIGURE 1. Tableau de Young associé à un endomorphisme

Une façon de construire ce tableau de Young est la suivante : on prend un vecteur e_1 de $K_r - K_{r-1}$ et on note pour $i = 1, \cdots, r-1$, $e_{i+1} = u^i(e_1)$. Si $\dim K_r/K_{r-1} > 1$ on choisit un vecteur $e_{r+1} \in K_r$ tel que les images de e_1, e_{r+1} dans N_r/N_{r-1} soient libres et on pose pour $i = 1, \cdots, r-1$, $e_{r+1+i} = u^i(e_{r+1})$. On continue le procédé jusqu'a obtenir une base $e_1, e_{r+1}, \cdots, e_{kr+1}$ de K_r/K_{r-1}. On choisit alors un vecteur $e_{(k+1)r+1}$ de K_{r-1} tel que les

images de $u(e_1), \cdots, u_(e_{kr+1}), e_{(k+1)r+1}$ forment une famille libre de K_{r-1}/K_{r-2} et on pose pour tout $i = 1, r-2$, $e_{(k+1)r+1+i} = u(e_{(k+1)r+1})$. On continue ce procédé jusqu'à épuiser tous les K_i. Parallèlement on remplit le tableau de Young comme dans la figure 1 dans laquelle l'image de e_1 est une base de $\operatorname{Ker} u^6/\operatorname{Ker} u^5$, les images de $u(e_1), e_7, e_{12}$ forment une base de $\operatorname{Ker} u^5/\operatorname{Ker} u^4$, les images de $u^4(e_1), u^3(e_7), u^3(e_{12}), e_{17}$ forment une base de $\operatorname{Ker} u^2/\operatorname{Ker} u$ et $u^5(e_1), u^4(e_7), u^4(e_{12}), u(e_{17}), e_{19}$ forment une base de $\operatorname{Ker} u$.
Remarque : dans la base construite précédemment la matrice de u est diagonale par blocs, les blocs étant des matrices de Jordan de taille $n_1, \cdots, n_s$. La structure de $K[X]$-module sur E induite par u donne un isomorphisme $E \simeq K[X]/(X^{n_1}) \times \cdots \times K[X]/(X^{n_s})$.
Application :

- *Deux matrices de permutations* sont semblables si et seulement si les permutations associées sont conjuguées (le faire avec la dimension des E^{σ^m}.
- *Racines carrées* : la réduite de Jordan de J_n est $\operatorname{diag}(J_{\lceil n/2 \rceil}, J_{\lfloor n/2 \rfloor})$, i.e. si n est pair (resp. impair) alors le tableau de Young de J_n^2 admet deux lignes de même taille $n/2$ (resp. de taille $(n+1)/2$ et $(n-1)/2$). Ainsi l'équation matricielle $X^2 = A$ a des solutions si et seulement si le tableau de Young de A vérifie une des conditions équivalentes suivantes :
  - en regroupant les lignes deux par deux en partant du haut (avec la convention que la dernière ligne est nulle si $\dim N_1$ est impaire), les lignes d'une même paire différent d'au plus une case ;
  - il n'y a pas deux colonnes consécutives de même longueur impaire.
- *Dimension du commutant* : il s'agit de déterminer le nombre de degré de liberté dans le choix d'un opérateur M qui commute avec A. On raisonne dans une base de Jordanisation de A. On rappelle que l'on a

$$\operatorname{Ker} A \subsetneq \operatorname{Ker} A^2 \subsetneq \cdots \subsetneq \operatorname{Ker} A^r = \operatorname{Ker} A^{r+1}.$$

  On considère une base $e_n, \cdots, e_{n-d_r+1}$ de $\operatorname{Ker} A^r - \operatorname{Ker} A^{r-1}$ de cardinal la longueur d_r de la dernière colonne du tableau de Young associé à A. L'image de cette base est totalement libre ce qui donne $d_r n$ degré de liberté ; en contrepartie l'image des u^k de ces vecteurs sont fixés. Soit alors r_1 maximal tel que $d_{r_1} \neq d_r$; on obtient alors $d_{r_1} \dim \operatorname{Ker} A^{r_1}$ nouveau degrés de liberté. On procède ainsi de suite jusqu'à épuiser tout l'espace. On vérifie alors aisément qu'on obtient un nombre de degré de liberté égal à la somme des carrés des longueurs des colonnes du tableau de Young.
- *Adhérence des orbites* : l'adhérence de l'orbite d'un bloc de Jordan de taille maximale est l'ensemble des nilpotents. Plus généralement l'ordre de Chevalley sur les orbites nilpotentes, est défini par $\mathcal{O}_1 \leq \mathcal{O}_2$ si et seulement si $\mathcal{O}_1$ est dans l'adhérence de $\mathcal{O}_2$: celui-ci correspond à l'ordre habituel sur les tableaux de Young, i.e. $(n_1 \geq n_2 \geq \cdots) \geq (m_1 \geq m_2 \geq \cdots)$ si et seulement si $n_1 \geq m_1$, $n_1 + n_2 \geq m_1 + m_2$ ...

2.2. Formes quadratiques. — Rappelons qu'étant donné un automorphisme σ du corps $\mathbb{K}$, par exemple la conjugaison complexe de $\mathbb{C}$, *une application semi-linéaire* est une application θ telle que pour tout $x, y \in E$ et $\lambda \in \mathbb{K}$ on a

$$\theta(x + \lambda y) = \theta(x) + \lambda^\sigma \theta(y)$$

où par convention on note λ^σ pour $\sigma(\lambda)$.

Définition 2.2. — On appelle *forme σ-sesquilinéaire* toute application $\phi : E \times E \to \mathbb{K}$ vérifiant les conditions suivantes :

— pour tout $x \in E$, l'application $\phi_x : y \in E \mapsto \phi(x,y)$ est linéaire ;
— pour tout $y \in E$ l'application $\phi_y : x \in E \mapsto \phi(x,y)$ est σ-linéaire.

Remarque : les notations ϕ_x et ϕ_y ne sont pas exemplaires, on veillera à ne pas se mélanger.

Définition 2.3. — Pour M une partie de E, on note

$$M^\perp = \{y \in E,\ \phi(M,y) = 0\}, \qquad {}^\perp M = \{x \in E,\ \phi(x,M) = 0\}.$$

On dit que $M^\perp$ (resp. ${}^\perp M$) est l'orthogonal à droite (resp. à gauche) de M.

Remarque : $M^\perp$ et ${}^\perp M$ sont clairement des sous-espaces de E. En outre $E^\perp = \operatorname{Ker}\phi_y$ et ${}^\perp E = \operatorname{Ker}\phi_x$.

Définition 2.4. — On dit que ϕ est *non dégénérée* si $E^\perp = \{0\}$ (resp. ${}^\perp E = \{0\}$. Le rang de A_ϕ est appelé *le rang* de ϕ, il est égal à la codimension de $E^\perp$ et ${}^\perp E$.

Remarque : pour M un sous-espace de E on a

$$\dim M + \dim M^\perp = \dim E + \dim(M \cap {}^\perp E).$$

On a aussi que ${}^\perp(M^\perp) = M + {}^\perp E$ et donc pour ϕ non dégénérée on retrouve la propriété habituelle ${}^\perp(M^\perp) = M$.

Définition 2.5. — Une forme σ-sesquilinéaire est dite *réflexive* si pour tout $x, y \in E$, $\phi(x,y) = 0$ équivaut à $\phi(y,x) = 0$.. Elle est dite hermitienne (resp. antihermitienne) si $\phi(x,y) = \epsilon\Big(\phi(y,x)\Big)^\sigma$ avec $\epsilon = 1$ (resp. $\epsilon = -1$).

Remarque : pour une forme hermitienne ou antihermitienne, σ est nécessairement une involution ; dans le cas antihermitien en caractéristique différente de 2, on a même $\sigma = \mathrm{Id}$ et on dit simplement que ϕ est *anti-symétrique*.

On suppose à présent que ϕ est une forme hermitienne ou antihermitienne, auquel cas la caractéristique est en outre supposée différente de 2.

Définitions 2.6. — — *Un vecteur x de E est dit* isotrope *si $\phi(x,x) = 0$.*
— *Un sous-espace F de E est dit* isotrope *si $F \cap F^\perp \neq \{0\}$.*
— *Un sous-espace F de E est dit* totalement isotrope *et on écrit* séti*, si $F \subset F^\perp$.*
— *Un séti est dit maximal et on écrit* sétim*, si pour tout séti G contenant F alors $G = F$.*

Remarque : comme on est en dimension finie, tout séti est contenu dans un sétim.
Remarque : si F est non isotrope alors $E = F \oplus F^\perp$; dans le cas où ϕ est non dégénéré c'est même une équivalence.

Proposition 2.7. — *Si ϕ est non dégénéré il existe alors une décomposition dite de Witt de l'espace $E = F \oplus F' \oplus G$ avec F, F' des sétim et G un sous- espace non isotrope telle que la matrice de ϕ dans une base adaptée soit de la forme*

$$\begin{pmatrix} 0 & I_r & 0 \\ \epsilon I_r & 0 & 0 \\ 0 & 0 & B \end{pmatrix}.$$

Remarque : sous-entendu dans l'énoncé précédent est que toute les sétim ont la même dimension appelée *l'indice* de ϕ.

Considérons à présent le cas où $\mathbb{K} = \mathbb{R}$. Dans ce cas $\sigma = \mathrm{Id}$ et on parle alors de forme bilinéaire symétrique et antisymétrique.

Définition 2.8. — Une forme bilinéaire symétrique est dite :
— positive (resp. négative) si pour tout $x \in E$, on a $\phi(x,x) \geq 0$ (resp. $\phi(x,x) \leq 0$) ;
— définie positive (resp. définie négative) si elle est positive (resp. négative) et que $\phi(x,x) = 0$ si et seulement si x est le vecteur nul.

Théorème 2.9. — ***(Loi d'inertie de Sylvester)***
Soit ϕ une forme bilinéaire symétrique.
— *Il existe alors une décomposition*

$$E = E^{\perp} \oplus E^{+} \oplus E^{-}$$

telle que la restriction de ϕ à E^+ (resp. E^-) est définie positive (resp. définie négative). Une telle décomposition n'est pas unique mais les dimensions s de E^+ et t de E^- sont les mêmes pour toute telle décomposition et sont respectivement égale au maximum des dimensions des sous-espaces F de E tels que la restriction de ϕ y soit définie positive (resp. négative). On dit que le couple (s,t) est la signature *de ϕ.*
— *Il existe une base $(e_i)_{1\leq i\leq n}$ de E telle que*

$$\phi(\sum_{i=1}^{n} \lambda_i e_i,\ \sum_{j=1}^{n} \mu_j e_j) = \sum_{i=1}^{s} \lambda_i\lambda_i\mu_i - \sum_{i=s+1}^{s+t} \lambda_i\mu_i.$$

— *Le rang de ϕ est égal à $s+t$ et son indice est égal à $(n - \mathrm{rg}\phi) + \min\{s,t\}$.*

Considérons à présent le cas $\mathbb{K} = \mathbb{C}$ et σ égal à la conjugaison complexe. Pour $A \in GL_n(\mathbb{C})$, on note A^* pour ${}^t\overline{A}$. Notons en particulier que tout forme hermitienne ϕ vérifie $\phi(x,x) \in \mathbb{R}$. On dit alors qu'elle est *positive* (resp. négative) si $\phi(x,x) \geq 0$ (resp. ≤ 0) pour tout $x \in E$ et on dit qu'elle est en outre *définie* si $\phi(x,x) = 0 \Rightarrow x = 0$.
Remarque : si E est muni d'une forme hermitienne définie positive on dit que E est *un espace hermitien.*

Théorème 2.10. — *Comme dans le cas réel,*
— *il existe une décomposition $E = E^{\perp} \oplus E^{+} \oplus E^{-}$ telle que la restriction de ϕ à E^+ (resp. E^-) est définie positive (resp. négative). En outre la dimension s de E^+ et t de E^- sont indépendantes de cette décomposition et le couple (s,t) s'appelle* la signature *de ϕ.*
— *Il existe une base $(e_i)_{1\leq i\leq n}$ telle que*

$$\phi\Big(\sum_{i=1}^{n} \lambda_i e_i, \sum_{j=1}^{n} \mu_j e_j\Big) = \sum_{i=1}^{s} \lambda_i\overline{\mu_i} - \sum_{i=s+1}^{s+t} \lambda_i\overline{\mu_i}.$$

— *Le rang de ϕ est $s+t$ et son indice $n - (s+t) + \min\{s,t\}$.*

2.3. Quelques sous-espaces de matrices. — Rappelons que $\mathcal{L}(E) \simeq E \otimes E^*$ et est donc de dimension n^2. On en déduit en particulier l'existence du polynôme minimal d'un endomorphisme qui est donc de degré $\leq n^2$.

L'ensemble $\mathcal{N}$ des matrices nilpotentes est un cône, i.e. si N est nilpotente alors tN aussi pour tout $t \in K$. Evidemment $\mathcal{N}$ n'est pas un espace vectoriel : par exemple $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ et $B = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$ sont nilpotentes alors que $A + B = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ ne l'est pas.

Proposition 2.11. — *Le sous-espace vectoriel engendré par l'un des ensembles suivant*

(a) $\mathcal{N}$ *;*
(b) les matrices nilpotentes de rang 1 *;*
(c) les matrices d'une classe de similitude quelconque de matrices nilpotentes, est l'hyperplan des matrices de trace nulle.

Preuve : Dans les trois cas, l'inclusion est immédiate. On va montrer directement (b). Comme d'habitude cela repose sur un petit calcul en dimension 2, à savoir : $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ est semblable à $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ en considérant la nouvelle base $e_1 + e_2$ et $e_1 - e_2$. Soit alors A une matrice de trace nulle ; en ajoutant une combinaison linéaire de matrice nilpotente de rang 1, on se ramène à A diagonale $\mathrm{diag}(a_1, \cdots, a_b)$ avec $\sum_i a_i = 0$ que l'on écrit sous la forme

$$\mathrm{diag}(a_1, -a_1, 0, \cdots, 0) + \mathrm{diag}(0, a_2 + a_1, a_3, \cdots, a_n).$$

D'après le calcul précédent la première matrice est semblable à une combinaison linéaire de matrice nilpotentes de rang 1 ; la deuxième aussi par hypothèse de récurrence.

(c) L'orbite d'une classe de similitude quelconque contient dans son adhérence la classe de similitude des matrices nilpotentes de rang 1. On conclut alors d'après (b).

Proposition 2.12*. — Tout hyperplan* H *de* $\mathbb{M}(n, \mathbb{C})$ *contient au moins* $n^2 - n - 1$ *matrices nilpotentes linéairement indépendantes.*

Preuve : On se ramène au cas où H a pour équation $tr(TX) = 0$ avec T triangulaire et on considère les intersections de H avec les sous-espaces des matrices nilpotentes triangulaires supérieures ou inférieures.

Proposition 2.13*. — Tout sous-*$\mathbb{R}$*-espace vectoriel de* $\mathcal{N}$ *est de dimension inférieur à* $\frac{n(n-1)}{2}$.

Preuve : On considère la forme quadratique q définie sur l'espace des matrices qui à X associe $\mathrm{tr}X^2$. De manière évidente le cône isotrope est constitué de vecteurs isotropes de sorte que l'espace vectoriel en question sera totalement isotrope. Par ailleurs, si $X \neq 0$ est symétrique (resp. antisymétrique) alors $q(X) > 0$ (resp. $q(X) < 0$) de sorte que la signature de q est $(\frac{n(n+1)}{2}, \frac{n(n-1)}{2})$. Ainsi un sous-espace totalement isotrope est de dimension inférieure ou égale à $\frac{n(n-1)}{2}$. L'égalité est clairement atteinte pour les matrices strictement triangulaires supérieures.
Remarque : évidemment ce maximum est atteint pour les matrices strictement triangulaires supérieures.

Théorème 2.14*. — (dit de Flanders)*
Un sous-espace vectoriel de $\mathbb{M}_n(\mathbb{R})$ *formé de matrices de rang* $\leq r$ *est de dimension* $\leq nr$.

Proposition 2.15*. — Un sous-espace affine de* $\mathbb{M}_n(\mathbb{R})$ *de dimension* $n^2 - \binom{k+1}{2} + 1$ *contient une matrice de rang au plus* $k - 1$.

Remarque : un sous-espace V de $\mathbb{M}_n(\mathbb{R})$ de dimension $> n(n-1)/2$ contient une matrice possédant une valeur propre non nulle. En effet, en utilisant la formule de Grassman, $V \cap \mathrm{Sym}(\mathbb{R}^n)$ en un sous-espace de dimension strictement positive. Soit alors M une matrice non nulle de cette intersection. Etant diagonalisable et non nulle, elle possède une valeur propre non nulle.

2.4. Trigonalisation simultanée. — On rappelle que si u est un endomorphisme d'un $\mathbb{C}$-espace vectoriel E et si $F \subset E$ est un sous-espace stable par u alors u induit un endomorphisme $\bar{u}$ de E/F. Étant donné un sous-ensemble $\mathcal{E}$ de $\mathcal{L}(E)$ et $G \subset F \subset E$ des sous-espaces stables par tous les éléments u de $\mathcal{E}$, l'ensemble des quotients de $\mathcal{E}$ pour $\{G \subset F\}$ est par définition $\{\bar{u} \in \mathcal{L}(F/G) :\ u \in \mathcal{E}\}$. *Une propriété P sera dite stable par quotients* si pour tout ensemble $\mathcal{E} \subset \mathcal{L}(E)$ constitués d'éléments satisfaisant P alors l'ensemble quotient de $\mathcal{E}$ pour $\{G \subset F\}$ est aussi constitué d'éléments de $\mathcal{L}(F/G)$ satisfaisant P.
Principe général : soit $\mathcal{P}$ est un ensemble de propriétés stables par quotients et vérifiant la propriété suivante : pour tout $\mathcal{E} \subset \mathcal{L}(E)$ constitué d'éléments vérifiant $\mathcal{P}$ avec $\dim E > 1$, $\mathcal{E}$ est réductible i.e. il existe un sous-espace vectoriel non trivial F de E stable par tous les éléments de $\mathcal{E}$. Alors $\mathcal{E}$ est triangularisable.
Exemple : tout sous-ensemble commutatif de $\mathcal{L}(E)$ est triangularisable. En effet la commutativité est clairement une propriété stable par quotient. La propriété de réductibilité découla alors du fait que tout sous-espace propre de A est stable par toute matrice B commutant avec A.
Remarque : si $\mathcal{A}$ est une sous-algèbre de $\mathcal{L}(E)$ l'ensemble $\mathcal{A}.x := \{Ax :\ A \in \mathcal{A}\}$, où $x \in E$, est un sous-espace stable par $\mathcal{A}$. Si $\mathcal{A}.x = E$, on dit que x est un vecteur cyclique pour $\mathcal{A}$. La détermination des sous-algèbres de $\mathcal{L}(E)$ qui possèdent des sous-espaces invariants non triviaux est réglée par le théorème suivant qui s'occupe de la partie réductibilité du principe général énoncé plus haut dans le cas des sous-algèbres de $\mathcal{L}(E)$.

Théorème 2.16. — *(**Burnside** cf.* [**6**] *1.2.2) Toute sous-algèbre propre de $\mathcal{L}(E)$ est réductible.*

Preuve : Soit $\mathcal{A}$ une sous-algèbre irréductible de $\mathcal{L}(E)$; comme tout endomorphisme est une somme d'endomorphisme de rang 1, nous allons montrer que tout endomorphisme de rang 1 appartient à $\mathcal{A}$.

Montrons tout d'abord que $\mathcal{A}$ contient un élément de rang 1. Soit $u_0 \in \mathcal{A}$ non nul de rang minimal ; si ce rang est strictement plus grand que 1, alors il existe des vecteurs x_1 et x_2 tels que $(u_0(x_1), u_0(x_2))$ est linéairement indépendant. Comme $\{u \circ u_0(x_1) :\ u \in \mathcal{A}\} = E$, il existe $u_1 \in \mathcal{A}$ tel que $u_1 \circ u_0(x_1) = x_2$ et donc $(u_0 \circ u_1 \circ u_0(x_1), u_0(x_1))$ est libre. Soit alors λ tel que la restriction de $u_1 \circ u_0 - \lambda\mathrm{Id}$ à $u_0(E)$ n'est pas inversible ; $(u_0 \circ u_1 - \lambda\mathrm{Id})u_0$ est non nul car l'image de x_1 est non nulle, et $(u_0 \circ u_1 - \lambda\mathrm{Id}) \circ u_0$ est de rang strictement plus petit que celui de u_0, d'où la contradiction et donc u_0 est de rang 1.

Pour y_0 dans l'image de u_0, on considère la forme linéaire ψ_0 définie par $u_0(x) = \psi_0(x)y_0$. Soit alors $u \in \mathcal{L}(E)$ défini par $u(x) = \psi(x)y$ où $y \in E$ et $\psi \in E^*$. Montrons alors que u appartient à $\mathcal{A}$. Pour $v \in \mathcal{A}$, on a $u_0 \circ v \in \mathcal{A}$ et $u_0 \circ v(x) = \psi_0(v(x))y_0$. Soit alors $F' \subset E^*$, l'ensemble des formes linéaires ψ telles que $x \mapsto \psi(x)y_0$ appartienne à $\mathcal{A}$: F' est clairement un sous-espace de E^*. Si ce sous-espace était strict, il existerait $x_0 \neq 0$ tel que $\psi(x_0) = 0$ pour tout $\psi \in F'$ (un espace vectoriel de dimension finie est réflexif). La contradiction découle alors du fait que $\psi_0(v(x_0)) = 0$ pour tout $v \in \mathcal{A}$ implique que x_0 est nul car $\{v(x_0) :\ v \in \mathcal{A}\} = E$. Soit donc $v_1 \in \mathcal{A}$ tel que $\psi = \psi_0 \circ v_1$.

De même comme $y_0 \neq 0$, alors $\{v(x_0) :\ v \in \mathcal{A}\} = E$ et donc pour tout $y \in E$ soit $v_2 \in \mathcal{A}$ tel que $v_2(y_0) = y$ et donc $u = v_2 \circ u_0 \circ v_1$.

Corollaire 2.17. — *(cf.* [**6**] *1.2.3) Les seuls idéaux bilatères de $\mathcal{L}(E)$ sont $\{0\}$ et $\mathcal{L}(E)$.*

Preuve : Soit $\mathcal{I}$ un idéal bilatère de $\mathcal{L}(E)$ non réduit à 0. Il suffit alors de montrer que $\mathcal{I}$ est irréductible. Si $u \neq 0$ appartient à $\mathcal{I}$, pour tout $0 \neq x \in E$, il existe $v \in \mathcal{L}(E)$ tel que $u \circ v(x) \neq 0$. Soit $y \in E$ et $w \in \mathcal{L}(E)$ tel que $w \circ u \circ v(x) = y$. On a $w \circ u \circ v \in \mathcal{I}$ de sorte que tout vecteur $x \neq 0$ est cyclique pour $\mathcal{I}$ et donc $\mathcal{I}$ est irréductible.
Remarque : on renvoie à **??** pour une preuve directe.

Corollaire 2.18. — *(cf.* [**6**] *1.2.4) Soit E est $\mathbb{C}$-espace vectoriel de dimension finie alors tout automorphisme d'algèbre ϕ de $\mathcal{L}(E)$ est intérieur, i.e. il existe $P \in GL(E)$ tel que pour tout $A \in \mathcal{L}(E)$, $\phi(A) = PAP^{-1}$.*

Preuve : Soit $A_0 \in \mathcal{L}(E)$ un idempotent de rang 1, $\phi(A_0)$ est alors un idempotent, montrons qu'il est aussi de rang 1. L'ensemble $\{A_0BA_0 : \ B \in \mathcal{L}(E)\}$ est un sous-espace vectoriel de $\mathcal{L}(E)$ de dimension 1 : on peut l'identifier avec $\mathcal{L}(\operatorname{Im} A_0)$. Son image par ϕ, $\{\phi(A_0)C\phi(A_0) : \ C \in \mathcal{L}(E)\}$ est donc aussi un sous-espace de $\mathcal{L}(E)$ de dimension 1 identifié à $\mathcal{L}(\operatorname{Im}\phi(A_0))$ de sorte que $\phi(A_0)$ est de rang 1. Comme tous les idempotents de rang 1 sont semblables à $\operatorname{diag}(1, 0, \cdots, 0)$, quitte à composer ϕ par $A \mapsto PAP^{-1}$, on peut supposer que $\phi(A_0) = A_0$.

Notons x_0 un vecteur directeur de $\operatorname{Im} A_0$ et soit $P \in \mathcal{L}(E)$ défini par $P(Bx_0) = \phi(B)x_0$: si $B_1x_0 = B_2x_0$ alors comme $A_0x_0 = x_0$, on a $(B_1 - B_2)A_0 = 0$ et donc $(\phi(B_1) - \phi(B_2))A_0 = 0$ de sorte que $\phi(B_1)x_0 = \phi(B_2)x_0$ et P est bien définie et évidemment linéaire. Supposons que $\phi(B)x_0 = 0$ de sorte que $\phi(B)\phi(A_0) = \phi(BA_0) = 0$ et donc $BA_0 = 0$ soit $Bx_0 = 0$ ce qui prouve l'injectivité de P et comme on est en dimension finie $P \in GL(E)$.

Soit alors $A \in \mathcal{L}(E)$, on a $P(AB)x_0 = \phi(AB)x_0 = \phi(A)\phi(B)x_0 = \phi(A)PBx_0$ et donc $PAy = \phi(A)Py$ pour tout $y = Bx_0$. Quand B décrit $\mathcal{L}(E)$, y décrit E et donc $PA = \phi(A)P$ pour tout $A \in \mathcal{L}(E)$ d'où le résultat.

Corollaire 2.19. — *(cf.* [**6**] *1.3.1) Toute algèbre d'endomorphismes nilpotents est triangularisable.*

Preuve : La propriété d'être nilpotent est stable par quotient comme en outre il existe des éléments de $\mathcal{L}(E)$ qui ne sont pas nilpotents, toute algèbre constituée d'endomorphismes nilpotents est, d'après le théorème de Burnside, réductible. La triangularisation découle alors du principe général énoncé plus haut.

3. Applications

3.1. Polynômes de Lagrange. — Soit $\mathbb{K}$ un corps et pour $a \in \mathbb{K}$ soit f_a la forme linéaire définie par $Q \in \mathbb{K}_{n-1}[X] \mapsto Q(i) \in \mathbb{K}$.

Proposition 3.1. — *Pour $i = 1, \cdots, n$ soient $a_i \in \mathbb{K}$ distincts deux à deux. La famille $(f_{a_1}, \cdots, f_{a_n})$ est alors une base de l'espace $\left(\mathbb{K}_{n-1}[X]\right)^*$ de base duale*

$$L_i(X) = \prod_{\substack{1 \leq j \leq n \\ j \neq i}} \frac{X - a_j}{a_i - a_j}.$$

Preuve : Notons que $f_{a_i}(L_j) = \delta_{i,j}$. Notons alors que la famille des $(f_{a_i})_{1 \leq i \leq n}$ est libre : en effet étant donnée une relation $\sum_i \lambda_i f_i = 0$, en la testant sur L_i, on obtient $\lambda_i = 0$. Comme cette famille est de cardinal la dimension de $\mathbb{K}_{n-1}[x]$ on en déduit que c'est une base.

Remarque : les polynômes L_i sont dits de Lagrange. La décomposition de $P \in \mathbb{K}_{n-1}[X]$ dans la base $(L_i)_{1\leq i\leq n}$ est aisée et donnée par

$$P(X) = \sum_{i=0}^{n-1} P(a_i)L_i(X).$$

Application au partage de secret : soit p un nombre premier "grand" ; tous les entiers considérés dans la suite seront supposés inférieur à p. Soit s_0 un entier. On choisit alors $n-1$ entiers $s_1, \cdots, s_{n-1}$ « au hasard » et inférieur à p. Soit P le polynôme $\sum_{i=0}^{n-1} s_i X^i$. En appliquant ce qui précède aux $a_i = i$ pour $i = 1, \cdots, n$, on en déduit que la connaissance des $P(i)$ pour $1 \leq i \leq n$, permet de retrouver $s_0 = P(0)$. En revanche si on suppose connu les $P(i)$ pour $1 \leq i \neq i_0 \leq n$ alors on ne connaît s_0 qu'à un multiple de $\frac{n!}{i_0}$ près. En effet soit $Q \in \mathbb{Q}_{n-1}[X]$ tel que $Q(i) = P(i)$ pour tout $0 \leq i \neq i_0 \leq n-1$. Ainsi $P - Q$ appartient à $\bigcap_{\substack{0\leq i\leq n-1 \\ i\neq i_0}} \operatorname{Ker} f_i$ qui est de dimension 1 engendré par $\prod_{0\leq i\neq i_0\leq n-1}(X-i)$ de sorte qu'il existe $\lambda \in \mathbb{Q}$ tel que $Q(X) = P(X) + \lambda \prod_{i\neq i_0}(X-i)$; or Q est à coefficients dans $\mathbb{Z}$ de sorte que $\lambda \in \mathbb{Z}$. Ainsi pour le coefficient constant de Q on obtient $s_0 + (-1)^{n-1}\lambda\frac{n!}{i_0}$ où $\lambda \in \mathbb{Z}$ est non déterminé ; on connaît alors s_0 à un multiple de $\frac{n!}{i_0}$ près.
Remarque : si $p < \frac{n!}{i_0}$, alors s_0 est connu.

On suppose désormais connue la congruence modulo p des $P(i)$ pour $i \neq i_0$ alors comme l'ensemble des restes de la division euclidienne par p de $\lambda\frac{n!}{i_0}$ lorsque λ décrit $\mathbb{Z}$, est égal à $\{0, 1, \cdots, p-1\}$, on ne sait strictement rien sur s_0 modulo p.

Le code pour déclencher une frappe nucléaire est un nombre inférieur à p que seul le président connaît. Au cas où celui-ci serait dans l'impossibilité d'agir, il est prévu que son état major constitué de n membres puissent déclencher la frappe sans que toutefois $n-1$ parmi eux y parviennent.

Supposons que le code soit s_0. On tire au sort les s_i, et on transmet $P(i)$ modulo p à la personne numérotée i. D'après ce qui précède, les n personnes réunies peuvent reconstituer s_0 alors que d'après (3), $n-1$ quelconques ne le peuvent pas.

De la même façon soit $P(X) = \sum_{i=0}^{k-1} s_i X^i$ et on transmet $P(i)$ à la personne i pour $1 \leq i \leq n$. Comme précédemment, k personnes quelconques peuvent reconstituer P et donc s_0 alors que $k-1$ quelconques ne le peuvent pas
Remarque : Si une personne malintentionnée i_0 transmet une mauvaise valeur distincte de $P(i_0)$ alors que toutes les autres transmettent leur $P(i)$, la personne i_0 sera la seule à connaître le code s_0. Bien sur s'il y a deux qui trichent, personne ne sait rien.

3.2. Algorithme de Berlekamp. — Soit $PA = \prod_{i=1}^{r} A_i \in \mathbb{F}_q[X]$ sans facteur carré où les A_i sont irréductibles. On note $R = \mathbb{F}_q[X]/(A)$ qui est une $\mathbb{F}_q$-algèbre de dimension $\deg A = n$.

Définition 3.2. — Soit f le morphisme d'élévation à la puissance q dans R, i.e. $f(Q) = Q^q$, et on défini l'algèbre de Berlekamp

$$R^f := \operatorname{Ker}(f - \mathrm{Id}) = \{Q \in R : \ R^q = R\}.$$

Proposition 3.3. — *La dimension* $\dim_{\mathbb{F}_q}(R^f) = r$*, le nombre de facteurs irréductibles de* A*.*

Preuve : D'après le théorème chinois on a $R \simeq \prod_{i=1}^{r} \mathbb{K}_i$ où $\mathbb{K}_i := \mathbb{F}_q[X]/(A_i)$ est une extension de degré $n_i := \deg A_i$ de $\mathbb{F}_q$. On note alors $f_i : \mathbb{K}_i \longrightarrow \mathbb{K}_i$ le morphisme d'élévation à la puissance q i.e. le morphisme de Frobenius qui engendre $\mathbb{K}_i/\mathbb{F}_q$. En particulier on a $\mathbb{K}_i^{f_i} = \mathbb{F}_q$ et donc $R^f \simeq \prod_{i=1}^{r} \mathbb{K}_i^{f_i} \simeq \mathbb{F}_q^r$.

Notation 3.4. — *Soit* $M = (m_{i,j})_{1 \leq i,j \leq n}$ *la matrice de* $f - \mathrm{Id}$ *dans la base canonique* $(X^i)_{0 \leq i \leq n-1}$ *de* R.

Proposition 3.5. — *Si* $r > 1$, *il existe* $G \in R^f$ *non constant et une décomposition non triviale*

$$A = \prod_{a \in \mathbb{F}_q} \mathrm{pgcd}(A, G - a).$$

Preuve : Comme $\dim_{\mathbb{F}_q} R^f = r > 1$, il existe nécessairement un élément G non constant dans R^f. En substituant G à la formule $X^q - X = \prod_{a \in \mathbb{F}_q}(X - a)$, on obtient

$$G(X)^q - G(X) = \prod_{a \in \mathbb{F}_q} (G(X) - a).$$

Comme les $G(X) - a$ sont premiers entre eux deux à deux, P et $G^q - G$ sont sans facteurs multiples et donc

$$\mathrm{pgcd}(A, G^q - G) = \prod_{a \in \mathbb{F}_q} \mathrm{pgcd}(A, G - a).$$

Pour $G \in R^f$, on a $G^q - G = 0$ et donc $\mathrm{pgcd}(P, G^q - G) = P$. La décomposition est en outre non triviale car sinon on aurait $P | G - a$ et donc $G = a$ dans R ce qui n'est pas.

L'algorithme de Berlekamp procède ainsi :

— On calcule la matrice M de Berlekamp.
— On calcule la dimension r du noyau de M ; si $r = 1$ alors P est irréductible.
— On cherche un vecteur G du noyau de M qui n'est pas un polynôme constant.
— Avec l'algorithme d'Euclide, on calcule les pgcd de A et $G - a$ pour tout $a \in \mathbb{F}_q$.
— Avec chacun des facteurs trouvés dans l'étape précédente, on recommence le processus.

3.3. Points de Gauss. —

Définition 3.6. — Soit $\mathcal{L}$ une forme linéaire sur $\mathbb{C}[X]$; une suite $(P_n)_{n \in \mathbb{N}}$ de polynômes à coefficients complexes est dite orthogonale relativement à $\mathcal{L}$ si :

— pour tout $n \in \mathbb{N}$, $\deg P_n = n$;
— pour tout $n \neq m$, $\mathcal{L}(P_n P_m) = 0$;
— pour tout $n \in \mathbb{N}$, $\mathcal{L}(P_n^2) \neq 0$.

Si en outre $\mathcal{L}(P_n^2) = 1$ pour tout $n \in \mathbb{N}$, la suite est dite orthonormale relativement à $\mathcal{L}$.

Remarque : si $(P_n)_{n \in \mathbb{N}}$ est une suite orthogonale relativement à $\mathcal{L}$, en abrégé on notera $SPO(\mathcal{L})$, alors $\mathcal{L}(P.P_n) = 0$ pour tout $P \in \mathbb{C}_{n-1}[X]$ ou autrement dit P_n dirige la droite orthogonale à l'hyperplan $\mathbb{C}_{n-1}[X]$ de $\mathbb{C}_n[X]$. La famille des P_n étant étagée, elle forme une base de $\mathbb{C}[X]$ et tout polynôme $P \in \mathbb{C}[X]$ s'écrit

$$P(X) = \sum_{k=0}^{\infty} \frac{\mathcal{L}(P.P_k)}{\mathcal{L}(P_k^2)}.P_k.$$

Une façon d'obtenir l'unicité d'une $SPO(\mathcal{L})$ est d'imposer que les P_n soient unitaires. En ce qui concerne l'existence elle est assurée par le théorème suivant.
Remarque : dans le cas réel, $\mathcal{L}(P)$ est souvent défini comme l'intégrale de P contre une fonction de poids à valeurs positives, i.e.

$$\mathcal{L}(P) = \int_a^b P(x)w(x)dx.$$

Théorème 3.7. — *Notons pour tout $n \in \mathbb{N}$, $\mu_n = \mathcal{L}(X^n)$ et soit*

$$\Delta_n = \det(\mu_{i+j})_{0\leq i,j\leq n} = \begin{vmatrix} \mu_0 & \mu_1 & \cdots & \mu_n \\ \mu_1 & \mu_2 & \cdots & \mu_{n+1} \\ \vdots & \vdots & & \vdots \\ \mu_n & \mu_{n+1} & \cdots & \mu_{2n} \end{vmatrix}$$

Une condition nécessaire et suffisante pour l'existence d'une $SPO(\mathcal{L})$ est

$$\Delta_n \neq 0 \qquad \forall n \in \mathbb{N}.$$

Preuve : Supposons que la suite des $P_n(X) = \sum_{k=0}^n c_{n,k}X^k$ est une $SPO(\mathcal{L})$ de sorte que pour tout n, m on a les relations $\mathcal{L}(X^m P_n(X)) = \delta_{m,n}K_n$ avec $K_n \neq 0$, ce qui matriciellement s'écrit

$$\begin{pmatrix} \mu_0 & \mu_1 & \cdots & \mu_n \\ \mu_1 & \mu_2 & \cdots & \mu_{n+1} \\ \vdots & \vdots & & \vdots \\ \mu_n & \mu_{n+1} & \cdots & \mu_{2n} \end{pmatrix} \begin{pmatrix} c_{n,0} \\ c_{n,1} \\ \vdots \\ c_{n,n} \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ K_n \end{pmatrix}.$$

Comme une SPO est uniquement déterminée par les K_n, l'équation matricielle ci-dessus a une unique solution de sorte que $\Delta_n \neq 0$.

Réciproquement si $\Delta_n \neq 0$ pour tout $n \in \mathbb{N}$, pour tout suite arbitraire $K_n \neq 0$, on construit les P_n dont les coefficients sont l'unique solution du système ci-dessus. En particulier pour tout $n \geq 1$, on a

$$c_{n,n} = K_n \frac{\Delta_{n-1}}{\Delta_n} \neq 0$$

de sorte que P_n est de degré n et orthogonal à $\mathbb{C}_{n-1}[X]$ et donc $(P_n)_{n\in\mathbb{N}}$ est bien une $SPO(\mathcal{L})$.

Définition 3.8. — La forme linéaire $\mathcal{L}$ est dite définie positive si $\mathcal{L}(P(X)) > 0$ pour tout polynôme à coefficients réels non nul prenant des valeurs positives sur $\mathbb{R}$.

Remarque : si $\mathcal{L}$ est définie positive alors $\mathcal{L}(X^{2k}) > 0$ et en utilisant

$$0 < \mathcal{L}\Big((X+1)^{2n}\Big) = \sum_{k=0}^{2n} \binom{2n}{k} \mu_{2n-k}$$

on obtient, en raisonnant par récurrence, que les μ_{2k+1} sont réels. Le procédé d'orthonormalisation de Gram-Schmidt permet alors de construire simplement une $SPO(\mathcal{L})$ de polynômes à coefficients réels. La caractérisation des $\mathcal{L}$ définies positives est donnée par la proposition suivante.

Proposition 3.9. — *La forme linéaire $\mathcal{L}$ est définie positive si et seulement si pour tout $n \in \mathbb{N}$, $\mu_n \in \mathbb{R}$ et $\Delta_n > 0$.*

Preuve : Supposons les μ_n réels et $\Delta_n > 0$. Soit d'après le théorème précédent $(P_n)_{n\in\mathbb{N}}$ la $SPO(\mathcal{L})$ unitaire avec donc $\mathcal{L}(P_n^2) = \Delta_n/\Delta_{n-1} > 0$ et $P_n \in \mathbb{R}_n[X]$. Un polynôme de degré m à coefficients réels $P(X)$ s'écrit alors sous la forme $\sum_{k=0}^m a_k P_k$ avec $a_k \in \mathbb{R}$ et

$$\mathcal{L}(P^2) = \sum_{j,k=0}^{m} a_j a_k \mathcal{L}(P_j P_k) = \sum_{k=0}^{m} a_k^2 \mathcal{L}(P^2) > 0.$$

Or si A est un polynôme à coefficients réels prenant des valeurs positives sur $\mathbb{R}$, alors d'après la proposition **??**, il existe $P, Q \in \mathbb{R}[X]$ tels que $A = P^2 + Q^2$ d'où le résultat.

Réciproquement si $\mathcal{L}$ est définie positive alors on a vu que les μ_n sont tous réels ; soit alors $(P_n)_{n\in\mathbb{N}}$ la $SPO(\mathcal{L})$ unitaire, de sorte que les $P_n \in \mathbb{R}[X]$ et donc

$$0 < \mathcal{L}(P_n^2) = \Delta_n/\Delta_{n-1}$$

Or comme $\Delta_{-1} = 1$, on en déduit par récurrence que tous les Δ_n sont strictement positifs.

Définition 3.10. — La forme linéaire $\mathcal{L}$ est dite définie positive sur un sous-ensemble $E \subset \mathbb{R}$, si $\mathcal{L}(P) > 0$ pour tout polynôme non identiquement nul sur E n'y prenant que des valeurs positives ; E est alors dit un support de $\mathcal{L}$.

Remarque : si E est un support infini de $\mathcal{L}$ alors tout sous-ensemble dense de E est aussi un support de $\mathcal{L}$. Ainsi en général il n'y a pas de plus petit support pour $\mathcal{L}$. Dans le cas où $\mathcal{L}$ admet un support borné, on peut définir son plus petit support fermé, cf. [**1**].

Dans la suite $\mathcal{L}$ désigne une forme linéaire sur $\mathbb{C}[X]$ définie positive et $(P_n)_{n\in\mathbb{N}}$ est la $SOP(\mathcal{L})$.

Théorème 3.11. — *Soit I un intervalle de $\mathbb{R}$ qui est un support de $\mathcal{L}$; alors les zéros de P_n sont réels, simples et appartiennent à l'intérieur de I.*

Preuve : Comme $\mathcal{L}(P_n) = 0$ P_n ne garde pas un signe constant sur I et par continuité il s'y annule au moins une fois avec une multiplicité impaire. Notons $x_1, \cdots, x_k$ les zéros de P_n dans I de multiplicité impaire et soit $P(X) = \prod_{i=1}^k (X - x_k)$. Ainsi $P(X)P_n(X)$ garde un signe constant sur I et donc $\mathcal{L}(PP_n) \neq 0$ de sorte que P ne peut pas appartenir à $\mathbb{C}_{n-1}[X]$ et donc $k = n$ d'où le résultat.

Théorème 3.12. — ***(Points de Gauss)*** *Soit $\mathcal{L}$ définie positive ; il existe alors des réels positifs $a_{n,1}, \cdots, a_{n,n}$ de somme égal à μ_0 tels que pour tout polynôme $\pi(x)$ de degré au plus $2n-1$, on ait*

$$\mathcal{L}(\pi(x)) = \sum_{k=1}^{n} a_{n,k} \pi(x_{n,k}).$$

Preuve : Soit $L_n(x) = \sum_{k=1}^n \pi(x_{n,k} l_k(x)$ le polynôme interpolateur de Lagrange de π pour les points $x_{n,1}, \cdots, x_{n,n}$, où $l_k(x) = \frac{P_n(x)}{(x-x_{n,k})P_n'(x_{n,k})}$. Ainsi $Q(x) = \pi(x) - L_n(x)$ est un polynôme de degré au plus $2n-1$ qui s'annule aux points $x_{n,i}$ pour $i = 1, \cdots, n$ et donc $Q(x) = R(x)P_n(x)$ avec $\deg R \leq n-1$. On a alors

$$\mathcal{L}(\pi(x)) = \mathcal{L}(L_n(x)) + \mathcal{L}(RP_n) = \mathcal{L}(L_n) = \sum_{k=1}^{n} \pi_n(x_{n,k}) \mathcal{L}(l_k)$$

ce qui donne le résultat en posant $a_{n,k} = \mathcal{L}(l_k)$. Pour $\pi = l_m^2$, on obtient

$$0 < \mathcal{L}(l_m^2) = \sum_{k=1}^{n} a_{n,k} l_m^2(x_{n,k}) = a_{n,m}.$$

Enfin pour $\pi(x) = 1$, on obtient $a_{n,1} + \cdots + a_{n,n} = \mu_0$.
Remarque : dans le cas où $\mathcal{L}(P) = \int_a^b P(x)w(x)dx$, la formule précédente permet de calculer les valeurs de cette intégrale sur $\mathbb{R}_{2n-1}[X]$ en fonction des valeurs prises par P en les points $x_{n,k}$. La méthode des rectangles et des trapèzes sont des cas particuliers de cette technique.
Remarque : on notera bien que si on avait pris des points d'interpolation au hasard, une formule comme dans le théorème précédent n'aurait été valide que sur $\mathbb{R}_{n-1}[X]$.

3.4. Codes correcteurs. — La problématique des codes correcteurs est la suivante : A veut transmettre une information à B via un canal bruité (les ondes dans l'air ambiant, un flux d'électrons dans un cable...) de sorte que B le reçoit avec éventuellement des erreurs que l'on supposera pas trop nombreuses (sinon il faut changer de mode de transmission). Il s'agit alors pour B de détecter ces erreurs et si possible, les corriger. L'idée est alors pour A de rajouter de la redondance à son message ; citons l'exemple un peu bête suivant.

On fixe un alphabet fini F de cardinal q (rapidement F sera un corps fini) de sorte que tous les messages à transmettre constituent un sous-ensemble de F^k. La phase d'encodage consiste ensuite à choisir $n > k$ puis à associer injectivement à chaque information $I \in F^k$ un message $M \in F^n$; le sous-ensemble obtenu de F^n s'appelle *le code C de longueur n*. Le rapport k/n qui mesure la redondance s'appelle *le taux d'information* du code. On dit que le message $(m_1, \cdots, m_n)$ est affecté de r erreurs si r de ses coordonnées ne sont pas correctes.

Définition 3.13. — Soient $(x_1, \cdots, x_n)$ et $(y_1, \cdots, y_n)$ deux éléments de F^n ; la distance de Hamming entre x et y notée $d_H(x, y)$ est le nombre d'indices $1 \leq i \leq n$ tels que $x_i \neq y_i$.

Remarque : $d_h : F^n \times F^n \to \mathbb{N}$ mérite bien le nom de distance comme le lecteur le vérifiera facilement.

Lors de la phase de décodage, on supposera toujours que le nombre d'erreurs possibles sur un mot est limité de sorte que si le message reçu R appartient au code alors le nombre d'erreurs est nul et sinon le message initial M est un mot du code C qui minimise la distance de Hamming. On peut alors formaliser le processus de décodage comme une application $D : F^n \to F^n$ dont l'image appartient à C et qui est l'identité sur C. Pour que tout cela fonctionne correctement, il y a un certain nombre de contraintes que nous allons essayer d'exposer.

Définition 3.14. — Soit C un code sur F, on appelle distance minimum de C l'entier

$$d = \min\{d_H(x, y) : \ x \neq y \in C\}.$$

S'il existe un mot $m' \in C$ tel que $d_H(m', R) < r$, alors clairement $m' \neq m$ et donc il ne faut pas décoder par m' même s'il s'agit du mot de C le plus proche de R ; en résumé il faut supposer que le nombre d'erreur r est tel que $r \leq \lfloor d/2 \rfloor$: en effet si on avait $d_H(m', R) < s$ alors d'après l'inégalité triangulaire on aurait $d_H(m, m') < 2r \leq d$ ce qui contredit la définition de la distance minimum d de C. Pour d pair et $r = d/2$, il n'est pas non plus exclu qu'il y ait deux mots distincts de C à distance r de R ce qui ne permet pas de décoder correctement.

Définition 3.15. — La capacité de correction de C, notée souvent t, est l'entier

$$t = \lfloor \frac{d-1}{2} \rfloor.$$

On dit alors que C est un code t-correcteur.

Remarque : ainsi pour tout $m \neq m'$ dans C, les boules fermées $B(m,t)$ et $B(m,t')$ sont disjointes et pour tout $x \in F^n$, la boule $B(x,t)$ contient au plus un mot de C. Signalons la situation idéale, mais rare, suivante où tout mot de F^n peut se décoder.

Définition 3.16. — Un code C est dit parfait si F^n est la réunion disjointe des boules fermées $B(m,t)$ où m décrit C.

Remarque : en utilisant que le cardinal de toute boule fermée de rayon r est de cardinal $\sum_{i=0}^{r} \binom{i}{n} (q-1)^i$, le code C est parfait si et seulement si on a

$$|C|. \sum_{i=0}^{t} \binom{i}{n} (q-1)^i = q^n.$$

Un code est bon si $|C|$ et d sont grands ; évidemment ces exigences sont contradictoires.

Considérons ici le cas des *codes linéaires*. On prend pour F le corps $\mathbb{F}_q$; le code C est dit linéaire si C est un sous-espace vectoriel de $\mathbb{F}_q^n$ de dimension k. *Le poids* $\omega(x)$ d'un élément $x \in \mathbb{F}_q^n$ est le nombre de ses composantes non nulles, soit aussi $d_H(x,0)$. Ainsi on a $d = \min_{0 \neq x \in C} \omega(x)$.

Proposition 3.17. — ***(Borne du singleton)*** *On a l'égalité*

$$d \leq n - k + 1.$$

Preuve : Notons E le sous-espace vectoriel de $\mathbb{F}_q^n$ formé des éléments dont les $k-1$ dernières composantes sont nulles de sorte qu'en notant $(e_i)_{1 \leq i \leq n}$ la base canonique, E est engendré par $e_1, \cdots, e_{n-k+1}$. Comme $\dim E + \dim C > n$, $E \cap C$ n'est pas réduit à 0 et il existe donc x tel que $\omega(x) \leq n - k + 1$ d'où le résultat.
Remarque : la distance relative d/n de C et son taux d'information k/n ne peuvent pas être simultanément proche de 1 vu que leur somme est plus petite que $1 + 1/n$. On dit que C est *un code MDS*, en anglais Maximum Distance Separable, si on a $d = n - k + 1$.

Définition 3.18. — Une matrice génératrice G d'un code C est une matrice dont les lignes forment une base. Une matrice vérificatrice H d'un code C est une matrice telle que $x \in C \Leftrightarrow Hx = 0$.

Proposition 3.19. — *La matrice H est vérificatrice si et seulement si elle est de rang $n-k$ et $G^t H = 0$.*

Preuve : Le résultat découle directement du fait qu'une matrice est vérificatrice pour $C = \{(u_1, \cdots, u_k)G : \ (u_1, \cdots, u_k) \in \mathbb{F}_q^k\}$ si et seulement si ses lignes forment une base des formes linéaires de $C^\perp$ s'annulant sur C.

Remarque : rappelons comment on calcule une base des formes linéaires s'annulant sur C. On considère la matrice $\tilde{G}$ construite à partir de G en rajoutant une dernière ligne $(x_1 \cdots x_n)$. En opérant sur les colonnes de $\tilde{G}$, on se ramène alors à une matrice étagée de la forme

$$\begin{pmatrix} * & 0 & \cdots & 0 & \cdots & \cdots & \cdots & \cdots & 0 \\ \vdots & * & \cdots & * & 0 & \cdots & \cdots & \cdots & 0 \\ \vdots & & & & \ddots & 0 & & & \vdots \\ \vdots & & & & & * & 0 & \cdots & 0 \\ f_1 & f_2 & \cdots & \cdots & \cdots & f_k & f_{k+1} & \cdots & f_n \end{pmatrix}$$

et $f_{k+1}, \cdots, f_n$ forment une base de $C^\perp$. Une autre façon de procéder est d'utiliser les codes systématiques.

Définition 3.20. — Un code C est dit systématique s'il existe une matrice B ayant k lignes et $n-k$ colonnes telles que $(I_k|B)$ soit une matrice génératrice de C ; une matrice de cette forme est dite normalisée.

Remarque : si elle existe, la matrice normalisée est nécessairement unique. L'avantage de celle-ci est que le message se lit directement sur les k-premières composantes. En opérant sur les lignes, C de matrice génératrice G est systématique si et seulement si la matrice extraite des k premières colonnes et lignes, est inversible. En s'autorisant aussi à permuter les coordonnées, on se ramène toujours à un code systématique. Le lemme suivant fournit alors un algorithme pour construire H à partir de G.

Lemme 3.21. — *Si $G = (I_k|B)$ est la matrice génératrice normalisée de C alors $H = (-{}^tB|I_{n-k})$ est une matrice de contrôle.*

Correction des erreurs : supposons que le code est 1-correcteur et notons m' le message reçu différant du message envoyé x en au plus une coordonnée alors l'erreur à corrigée est $\epsilon = x' - x$ avec ϵ égal au vecteur e_i de la base canonique tel que $He_i = Hx'$. Plus généralement pour décoder un message, on commence par calculer tous les Hx pour les x tels que $\omega(x) \leq t$ de sorte que lorsque l'on reçoit un message m', la correction à apporter est ϵ tel que $\omega(\epsilon) \leq t$ et $H(\epsilon) = H(m')$.

Proposition 3.22. — *Soit H une matrice de contrôle de C ; la distance d de C est égal au nombre minimum de colonnes de H qui en tant que vecteurs de $\mathbb{F}_q^{n-k}$, sont linéairement dépendantes.*

Preuve : Le résultat découle des observations évidentes suivantes : s'il existe dans c un mot $(x_1, \cdots, x_n)$ de poids r alors de la relation $Hx = 0$, on en déduit qu'il existe r colonnes de H linéairement dépendantes ; la réciproque est identique.

Code de Hamming de longueur 7 : prenons l'ensemble des mots de sept chiffres binaires, $q = 2$, $n = 7$ et $\mathcal{C}$ le code ayant pour base

$$e_0 = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad e_1 = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad e_2 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \quad e_3 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 1 \\ 0 \\ 1 \end{pmatrix},$$

Pour transmettre un message $m = (m_0, m_1, m_2, m_3)$, on transmet $x = m_0e_0 + m_1e_1 + m_2e_2 + m_3e_3$. Les paramètres de ce code sont $(7, 4, 3)$. Une matrice vérificatrice est

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

Remarque : la matrice H a été obtenue en opérant sur les lignes comme annoncés précédemment ; le lecteur pourra aussi vérifier qu'elle est de rang 3 et que $G^tH = 0$. En outre toutes les colonnes de H sont distinctes de sorte qu'on obtient toutes les vecteurs non nuls de $\mathbb{F}_3$. On en déduit alors que deux colonnes sont obligatoirement libres et qu'étant données deux colonnes quelconques de H, leur somme est une colonne de H. De la proposition précédente on en déduit que $d = 3$. Ainsi le code de Hamming de longueur 7 est 1-correcteur parfait mais il n'est pas MDS. Etant donné un message reçu x' on dira que le message initial était $x = x' + e_i$ où i est l'indice de la colonne de H égale à Hx'. En ce qui concerne l'information de départ $y = A^{-1} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}$ où $A = \begin{pmatrix} 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}$.

3.5. En analyse. — On peut parler des équations différentielles linéaires à coefficients constants et des suites récurrentes linéaires. On peut aussi se lancer sur le théorème dit du rang constant suivant.

Théorème 3.23. — *Soit U un ouvert de $\mathbb{R}^n$ et $a \in U$. Si $f : U \longrightarrow \mathbb{R}^p$ est une application de classe C^1 dont la différentielle est de rang constant égal à r alors il existe*

— *un C^1 difféomorphisme φ d'un ouvert V de $\mathbb{R}^n$ contenant 0 sur U tel que $\varphi(0) = a$;*

— *un C^1 difféomorphisme ψ de $f(U)$ sur un ouvert de $\mathbb{R}^p$*

tels que pour tout $x \in V$

$$\psi \circ f \circ \varphi(x_1, \cdots, x_n) = (x_1, \cdots, x_r, 0, \cdots, 0).$$

4. Développements

h

— existence de la dimension [**2**] [**?**]

— dimension du corps de rupture, mutiplicativité des degrés, application à la constructibilité [**5**]

— théorème de la base de Burnside [**3**]

— idéaux de $\mathcal{L}(E)$

— théorème de Rouché-Fontené ou détermination du rang à l'aide des bordants [**?**]
— algorithme de détermination du rang à l'aide des opérations élémentaires
— classification des classes de similitudes des endomorphismes nilpotents [**4**]
— deux matrices de permutations sont conjuguées si et seulement si les permutations sont conjuguées (une preuve qui utilise la notion de dimension)
— théorème de Hahn Banach en dimension fini [**7**]
— théorème du rang constant [**7**]

5. Questions

6. Solutions

Références

[1] T.S. Chihara. *An introduction to orthogonal polynomials*. Mathematics and its applications, 1978.

[2] J. Fresnel. *Algèbre des matrices*. Hermann, 1997.

[3] I. Hall. *Theory of groups*.

[4] R. Mneiné. *Réduction des endomorphismes*. Calvage et Mounet, 2006.

[5] D. Perrin. *Cours d'algèbre*. Ellipses, 1998.

[6] H. Radjavi and P. Rosenthal. *Simultaneous Triangularization*. Springer, 2000.

[7] F. Rouvière. *Petit guide de calcul différentiel à l'usage de la licence et de l'agrégation*. Cassini, 2003.