Algèbres $\mathbb{K}[u]$, réduction de Jordan, exponentielle matricielle et calcul fonctionnel

Hugo Eulry

Résumé

On illustre dans ces notes de cours les 3 points fondamentaux que sont la réduction de Jordan, la décomposition de Dunford et l'exponentielle matricielle. En étudiant d'abord l'algèbre $\mathbb{K}[u]$, on donne la démarche classique pour traiter des ces thèmes, puis on mime le formalisme de l'exponentielle matricielle pour définir le *calcul fonctionnel* holomorphe et obtenir les mêmes résultats de façon purement analytique.

Table des matières

1 – Algèbres $\mathbb{K}[u]$ et réduction

1.1 – Notions de bases

Définition 1. *Étant donné un corps $\mathbb{K}$, une $\mathbb{K}$-algèbre A est un $\mathbb{K}$-espace vectoriel que l'on munit d'une structure d'anneau $(A, +, \times)$ dont la multiplication est bilinéaire.*
L'algèbre A sera dite associative/unitaire si l'anneau A l'est.
Pour deux algèbres A, B, une application $f : A \to B$ est un morphisme d'algèbre si f est à la fois une application linéaire et un morphisme d'anneaux entre A et B. On parlera d'isomorphisme de $\mathbb{K}$-algèbre si f est de plus bijective.

Dans tout ce qui suit, on supposera implicitement que nos algèbres sont associatives et unitaires, ce qui est le cas dans les exemples qu'on considérera. La notion d'algèbre sur un anneau commutatif peut être définie de façon similaire mais on se contente ici du cas des corps pour avoir accès aux techniques d'algèbre linéaire usuelles et à la notion de dimension.

On a une façon générique de créer des morphismes d'algèbre via le morphisme d'évaluation. Étant donné un élément $x \in A$, on définit son morphisme d'évaluation par :

$$ev_x : \left| \begin{array}{l} \mathbb{K}[X] \to A \\ \quad P \mapsto P(x) \end{array} \right.$$

Son noyau est un idéal de $\mathbb{K}[X]$: l'idéal des polynômes annulateurs de x. Comme $\mathbb{K}[X]$ est principal, soit $Ker(ev_x)$ est trivial, soit il est engendré par un unique polynôme unitaire noté Π_x : le polynôme minimal de x.

Définition 2. *Soit $\mathbb{K}$ un corps, A une $\mathbb{K}$-algèbre et $x \in A$. Si ev_x est injectif, on dit que x est un élément transcendant. Sinon, x est dit algébrique et $Ker(ev_x)$ est un idéal de $\mathbb{K}[X]$ engendré par un unique générateur unitaire : Π_x.*

Notons que $ev_x(\mathbb{K}[X])$ est une sous algèbre de A, notée $\mathbb{K}[x]$, appelée algèbre engendrée par x. C'est la plus petite sous-algèbre de A qui contient x. Dès lors on a :

$$\langle x\rangle := \mathbb{K}[x] = \bigcap_{x\in\tilde{A} \text{ sous-algèbre de } A} \tilde{A}$$

Notons également que ev_x induit un isomorphisme :

$$\mathbb{K}[X]/(\Pi_x) \xrightarrow{\sim} \mathbb{K}[x]$$

De ce fait, si x est transcendant sur A, alors on a $\mathbb{K}[X] \simeq \mathbb{K}[x]$.

On s'intéressera dans la suite uniquement aux algèbres qui sont de dimension finie (en tant qu'espace vectoriel).

Exemples 3. — *$\mathbb{C}$ est une $\mathbb{R}$ algèbre de dimension 2*
- *$\mathcal{M}_n(\mathbb{K})$, ou de façon équivalente $\mathcal{L}(E)$ où E est un espace vectoriel de dimension finie n, est une algèbre de dimension n^2.*
- *Si $P \in \mathbb{K}[X]$ est non nul, $\mathbb{K}[X]/(P)$ est une algèbre de dimension finie.*
- *Étant donnée une chaîne (appelée drapeau) de sous-espaces vectoriels de E : $0 = F_0 \subset F_1 \subset \cdots \subset F_r = E$, l'espace des endomorphismes qui préservent chacun des sous-espaces F_i est une sous-algèbre de $\mathcal{L}(E)$. De façon équivalente, si $a = (a_1 \cdots, a_r)$ avec $a_i = dim(F_i/F_{i-1})$, $T_a(\mathbb{K})$ l'espace des matrices triangulaires par blocs avec blocs diagonaux de taille respective a_i est une sous-algèbre de $\mathcal{M}_n(\mathbb{K})$.*
- *Étant donnée une décomposition en somme directe $E = F_1 \oplus \cdots \oplus F_r$, l'espace des endomorphismes qui préservent cette décomposition est une sous-algèbre de $\mathcal{L}(E)$. De même, si $a = (a_1, \cdots, a_r)$ où $a_i = dim(F_i)$, l'espace des matrices diagonales par blocs avec des blocs de taille respective a_i est une sous-algèbre de $\mathcal{M}_n(\mathbb{K})$.*

1.2 – Algèbre des polynômes en un endomorphisme

Comme vu ci-dessus, pour un élément $u \in \mathcal{L}(E)$, la structure de l'algèbre $\mathbb{K}[u]$ est entièrement déterminée par la donnée du polynôme minimal de u. Notons qu'on a même la décomposition suivante :

$$\mathbb{K}[u] \simeq \mathbb{K}[X]/(P_1^{\alpha_1}) \times \cdots \times \mathbb{K}[X]/(P_r^{\alpha_r})$$

si $\Pi_u = P_1^{\alpha_1} \cdots P_r^{\alpha_r}$ est la décomposition en irréductibles de Π_u.

Définition 4. *Soit F un s.e.v. de E tel que $u(F) \subset F$ (F est stable par u), la restriction de u à F définit bien un endomorphisme de F, appelé endomorphisme induit et noté $u_{|F}$.*

Exemple 5. *Si $P \in \mathbb{K}[X]$, $Ker(P(u))$ et $Im(P(u))$ sont des sous-espaces stables par u.*

Il est naturel de déduire des propriétés de $u_{|F}$ à partir de celles de u, notamment la suivante :

Proposition 6. *Si F est stable par u, alors $\Pi_{u_{|F}}$ divise Π_u. Plus généralement, si $E = F_1 \oplus \cdots \oplus F_r$ est une décomposition en sous-espaces stables par u, alors Π_u est le PPCM des $\Pi_{u_{|F_i}}$.*

Une telle décomposition est utile en pratique car elle permet de ramener l'étude de u à celle de ses endomorphismes induits, ou de façon équivalente, ramener l'étude d'une matrice à celle de plusieurs matrices de plus petite taille. C'est le principe de la

réduction : trouver des sous-espaces qui forment à la fois une décomposition en somme directe, et qui sont stables par l'endomorphisme étudié.
Le résultat suivant donne une "recette" pour obtenir des décompositions intéressantes.

Théorème 7 (Lemme des noyaux). *Soient E un $\mathbb{K}$-e.v., $u \in \mathcal{L}(E)$ et $P_1, \cdots, P_r \in \mathbb{K}[X]$ des polynômes premiers entre eux deux à deux. Alors, en notant $P = \prod_{i=1}^r P_i$, on a*

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

En particulier, si $P(u) = 0$, on a une décomposition de l'espace tout entier.

Dans le cas de Π_u et avec les notations précédentes, on a

$$E = \bigoplus_{i=1}^{r} Ker(P_i^{\alpha_i}(u)) = \bigoplus_{i=1}^{r} F_i$$

et de plus $\Pi_{u_{|F_i}} = P_i^{\alpha_i}$. En particulier, on retrouve le résultat suivant :

Proposition 8. *Soient E un $\mathbb{K}$-e.v. et $u \in \mathcal{L}(E)$, si u admet un polynômes annulateur scindé à racines simples alors u est diagonalisable. En outre, Π_u est scindé à racines simples sur $\mathbb{K}$ si et seulement si u est diagonalisable.*

Bien que le deuxième critère fournisse une équivalence, il est plus difficile à utiliser en pratique car nécessite de calculer le polynôme minimal. On lui préfère en général le premier, qui n'est qu'une condition nécessaire.

Exemple 9. *Si G est un groupe abélien fini et $\rho : G \to GL_n(\mathbb{C})$ une représentation linéaire de G, tout élément de la forme $\rho(g)$ est diagonalisable. En effet, si $g \in G$ on a $\rho(g)^{|G|} = \rho\left(g^{|G|}\right) = I_n$ et comme $X^{|G|} - 1$ est scindé à racines simples sur $\mathbb{C}$, on en déduit que $\rho(g)$ est diagonalisable.*

Dans le cas particulier des corps finis, on peut préciser la proposition précédente.

Proposition 10. *$M \in \mathcal{M}_n(\mathbb{F}_q)$ est diagonalisable sur $\mathbb{F}_q$ si et seulement si $X^q - X$ est un polynôme annulateur de M.*

Démonstration : La condition est clairement suffisante car $X^q - X$ est scindé à racines simples sur $\mathbb{F}_q$. Réciproquement, si M est diagonalisable sur $\mathbb{F}_q$ alors il existe $P \in GL_n(\mathbb{F}_q)$ et $\lambda_1, \cdots, \lambda_n \in \mathbb{F}_q$ tels que $M = PDP^{-1}$, où $D = Diag(\lambda_i\ ,\ i \in [\![1, n]\!])$. Comme $x^q = x$ pour tout élément $x \in \mathbb{F}_q$ (le Frobenius est précisément l'identité sur un corps fini) donc

$$M^q = PD^qP^{-1} = PDP^{-1} = M$$

ce qui permet de conclure.

□

Le résultat suivant est un élément fondamental dans la démonstration du théorème de réduction de Frobenius. Il fait le lien entre polynôme minimal et polynôme minimal en un vecteur donné.

Proposition 11. *Soit $u \in \mathcal{L}(E)$ où E est de dimension finie. Étant donné un vecteur $x \in E$, on note Π_u^x le générateur unitaire de l'idéal*

$$\{P \in \mathbb{K}[X]\ ,\ P(u)(x) = 0\}$$

Alors il existe $x \in E$ tel que $\Pi_u^x = \Pi_u$.

Démonstration : Dans le cas où $\Pi_u = P^\alpha$ avec P irréductible, supposons que pour tout $x \in E$, $\Pi_u^x \neq P^\alpha$. $\Pi_u = P^\alpha$ étant le polynôme minimal, on a $\Pi_u^x | P^\alpha$ et comme P est irréductible, $\Pi_u^x | P^{\alpha-1}$. Mais alors $P^{\alpha-1}(u)(x) = 0$ et donc finalement $P^{\alpha-1}(u) = 0$, ce qui est absurde par minimalité de Π_u.
Pour le cas général, considérons $\Pi_u = P_1^{\alpha_1} \cdots P_r^{\alpha_r}$ la décomposition de Π_u en facteurs irréductibles. Par lemme des noyaux, on a la décomposition en somme directe :

$$E = \bigoplus_{i=1}^{r} Ker(P_i^{\alpha_i}(u))$$

En notant $F_i := Ker(P_i^{\alpha_i}(u))$ et $u_i := u_{|F_i}$, ce qui précède assure l'existence de vecteurs $x_i \in F_i$ pour tout $i \in [\![1, r]\!]$ tels que $\Pi_{u_i}^{x_i} = \Pi_{u_i}$. Soit alors $x := x_1 + \cdots + x_r$, on montre que x convient. Comme $\Pi_u^x | \Pi_u$, il suffit de montrer que $\Pi_u | \Pi_u^x$, les deux polynômes étant unitaires, c'est à dire, que $\Pi_u^x(u) = 0$. On a

$$0 = \Pi_u^x(u)(x) = \Pi_u^x(x_1) + \cdots + \Pi_u^x(x_r)$$

Comme chaque F_i est stable par u (et donc par Π_u^x), la décomposition précédente est encore une somme directe, ce qui assure que $\Pi_u^x(u)(x_i) = 0$ pour tout $i \in [\![1, r]\!]$. Alors, $\Pi_{u_i}^{x_i} | \Pi_u^x$ d'une part, et $\Pi_{u_i} = \Pi_{u_i}^{x_i}$ d'autre part en utilisant le cas précédent. On a alors $\Pi_u^x(u) = 0$ sur chaque F_i, donc sur E tout entier, ce qui permet de conclure.

□

2 – Réduction de Jordan

2.1 – Diagonalisation, trigonalisation

Pour étudier un endomorphisme u en dimension finie, il est naturel de chercher des sous-espaces sur lesquels l'endomorphisme est plus simple, le cas idéal étant d'avoir une base de diagonalisation. La notion de vecteur propre/valeur propre apparaît alors naturellement si on cherche des droites stables pour un endomorphisme, on aura alors une décomposition de l'espace la plus simple possible pour u.

Définition 12. *Soit $u \in \mathcal{L}(E)$, $x \in E \setminus \{0\}$ est appelé vecteur propre associé à la valeur propre $\lambda \in \mathbb{K}$ si $x \in Ker(u - \lambda \cdot id_E)$, i.e. $u(x) = \lambda x$.*

La notion de valeur propre correspond donc à un défaut d'injectivité de $u - \lambda \cdot id_E$, donc à un défaut d'inversibilité si E est de dimension finie (dans le cas de la dimension infinie il faut bien distinguer valeur propre, injectivité, et valeur spectrale, inversibilité).
On a alors une caractérisation "simple" des valeurs propres en dimension finie :

$$\lambda \in \mathbb{K} \text{ est valeur propre de } u \Leftrightarrow det(u - \lambda \cdot id_E) = 0$$

ou son équivalent matriciel $det(A - \lambda \cdot I_n) = 0$. Notons que même si la version matricielle requiert à-priori le choix d'une base pour écrire l'endomorphisme u, l'équation précédente est invariante par conjugaison $A \mapsto PAP^{-1}$. On obtient donc un objet algébrique qui caractérise exactement les valeurs propres.

Définition 13. *Soit $u \in \mathcal{L}(E)$, on appelle polynôme caractéristique de u le polynôme définit par :*

$$\chi_u(X) := det(u - X \cdot id_E) \in \mathbb{K}[X]$$

On ramène donc le problème de recherche de valeurs propres à celui de résolution d'une équation polynomiale. En particulier, le choix du corps de base est crucial ici, si le corps est algébriquement clos il existera toujours des valeurs propres.

Exemple 14. $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ *n'admet pas de valeur propre sur $\mathbb{R}$ mais admet $\pm i$ sur $\mathbb{C}$.*

Remarque : *Bien que le fait de considérer une extension du corps puisse assurer l'existence de valeurs propres, χ_u et Π_u sont quant à eux invariants par extension de corps.*

Le fait d'avoir toutes ses valeurs propres dans le corps de base est équivalent à ce que χ_u soi scindé sur $\mathbb{K}$. Cela fournit une condition nécessaire et suffisante de trigonalisation.

Proposition 15. *$u \in \mathcal{L}(E)$ est trigonalisable sur $\mathbb{K}$ si et seulement si χ_u est scindé sur $\mathbb{K}$. Si de plus χ_u est à racines simples alors u est diagonalisable.*

Remarque : *Si u est trigonalisable, on a effectivement que toutes ses valeurs propre sont dans $\mathbb{K}$. La réciproque se montre par récurrence. Notons que le dernier critère est une condition suffisante mais évidemment pas nécessaire en regardant le cas de $u = id_E$.*

Ce résultat assure 2 points importants : si $\mathbb{K}$ est algébriquement clos, alors u est trigonalisable, et il en est de même si u est nilpotent (peu importe le corps, celui-ci contient toujours 0, seule valeur propre des endomorphismes nilpotents).

Pour obtenir une condition nécessaire de diagonalisablilité, on doit introduire la notion d'espace propre.

Définition 16. *Soit $u \in \mathcal{L}(E)$ et λ une valeur propre de u. On appelle multiplicité géométrique de λ la dimension de $E_u(\lambda)$ et multiplicité algébrique, notée m_λ, sa multiplicité en tant que racine de χ_u. On appelle sous-espace propre de u associé à λ le sous-espace :*

$$E_\lambda(u) := Ker(u - \lambda \cdot id_E)$$

et sous-espace caractéristique

$$E_\lambda^{m_\lambda}(u) := Ker\left((u - \lambda \cdot id_E)^{m_\lambda}\right)$$

Notons que par définition de χ_u, on a

$$\sum_{\lambda \in Sp(u)} m_u(\lambda) = n$$

alors qu'en général on a seulement

$$\sum_{\lambda \in Sp(u)} \dim(E_\lambda(u)) \leq n$$

Cependant, on peut montrer que les sous-espaces propres sont en somme directe et qu'alors :

$$\bigoplus_{\lambda \in Sp(u)} E_\lambda(u) \subset E$$

Le défaut d'égalité dans les deux dernières identité est exactement ce qu'il manque à u pour être diagonalisable.

Proposition 17. *$u \in \mathcal{L}(E)$ est diagonalisable si et seulement si $m_u(\lambda) = dim(E_\lambda(u))$ pour tout $\lambda \in Sp(u)$, ou de façon équivalente, $\sum_\lambda dim(E_\lambda(u)) = n$.*

Démonstration : La condition est clairement nécessaire au regard de l'inclusion précédente. Réciproquement, supposons u diagonalisable, notons $\lambda_1, \cdots, \lambda_r$ ses valeurs propres distinctes. Alors il existe une base de E dans laquelle la matrice de u est diagonale par blocs avec des blocs de la forme $\lambda_i \cdot I_{m_i}$. Dans ce cas le polynôme caractéristique est donc $\chi_u(X) = \prod_{i=1}^r (X - \lambda_i)^{m_i}$, d'où le résultat.

□

Dans le cas d'un endomorphisme diagonalisable, il est direct de voir que $\chi_u(u) = 0$. Ce résultat reste vrai si u est quelconque.

Théorème 18 (Cayley-Hamilton). *Pour tout endomorphisme $u \in \mathcal{L}(E)$, on a*

$$\chi_u(u) = 0$$

Démonstration : Dans le cas où $\mathbb{K} = \mathbb{C}$, on peut résonner par densité des matrices diagonalisables et continuité du déterminant. Dans un cadre plus général, soit $A \in \mathcal{M}_n(\mathbb{K})$, on a la relation $MCom(M)^T = det(M)I_n$ valable pour toute matrice carrée à coefficients dans un anneau. En particulier pour $M = A - XI_n$ on obtient

$$(A - XI_n)Com(A - XI_n)^T = \chi_A(X)I_n$$

$Com(A - XI_n)^T$ est un polynôme à coefficients matriciels, de degré au plus $n - 1$, qu'on peut écrire sous la forme $Com(A - XI_n)^T = \sum_{j=0}^{n-1} B_j X^j$. On écrit de même $\chi_A(X) = \sum_{j=0}^{n} a_j X^j$.
En développant le produit dans le membre de gauche on obtient la relation

$$(A - XI_n)Com(A - XI_n)^T = -X^n B_{n-1} + \sum_{j=1}^{n-1} X^j(AB_j - B_{j-1}) + AB_0 = \sum_{j=0}^{n} a_j X^j I_n$$

En identifiant les coefficients de chaque côté, on obtient :

$$a_n I_n = -B_{n-1}, \quad a_j I_n = AB_j - B_{j-1}, \quad a_0 I_n = AB_0$$

On a alors :

$$\begin{aligned}
\chi_A(A) &= \sum_{j=0}^{n} a_j A^j = \sum_{j=0}^{n} A^j a_j I_n && (1)\\
&= -A^n B_{n-1} + \sum_{j=1}^{n-1} A^j(AB_j - B_{j-1}) + AB_0 && (2)\\
&= \sum_{j=0}^{n-1} A^{j+1}B^j - \sum_{j=1}^{n} A^j B_{j-1} = 0 && (3)
\end{aligned}$$

□

On a donc un exemple "canonique" de polynôme annulateur d'un endomorphisme. Dans ce cas, en utilisant le lemme des noyaux, on a une décomposition de l'espace en sous-espaces stables par u

$$E = \bigoplus_{i=1}^{r} Ker\left((u - \lambda_i \cdot id_E)^{m_i}\right)$$

où sur chaque sous-espace $Ker(u - \lambda_i \cdot id_E)^{m_i}$, la restriction de $u - \lambda_i \cdot id_E$ est un endomorphisme nilpotent.

2.2 – Réduction de Jordan

On a donc ramené le problème de réduire un endomorphisme quelconque au cas des endomorphismes nilpotents. C'est exactement l'objet de la réduction de Jordan. Commençons par définir les matrice qui joueront un rôle crucial dans la suite .

Définition 19. *Soient $r \in \mathbb{N}^*$ et $\lambda \in \mathbb{K}$, on définit le bloc de Jordan de taille r :*

$$J_{r,\lambda} := \begin{pmatrix} \lambda & 1 & & (0) \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ (0) & & & \lambda \end{pmatrix} \in \mathcal{M}_r(\mathbb{K})$$

Le but de cette section est de démontrer le théorème de réduction suivant :

Théorème 20 (Jordan). *Soit $u \in \mathcal{L}(E)$, toutes les valeurs propres de u sont dans $\mathbb{K}$ si et seulement si il existe une base de E dans laquelle la matrice de u est diagonale par blocs avec des blocs de Jordan du type $J_{r,\lambda}$ sur la diagonale, $\lambda \in Sp(u)$.*

Notons qu'en utilisant la décomposition en somme directe ci-dessus, il suffit en fait de traiter le cas des endomorphismes nilpotents. On commence par un lemme préliminaire.

Lemme 21. *Soit $u \in \mathcal{L}(E)$ nilpotent d'indice $r \geq 1$, alors :*

$$\{0\} = Ker(u^0) \subsetneq Ker(u) \subsetneq \cdots \subsetneq Ker(u^{r-1}) \subsetneq Ker(u^r) = E$$

Démonstration : La chaîne d'inclusion est directe à obtenir et se termine par E car u est nilpotent d'indice r. Supposons par l'absurde qu'on ait $Ker(u^{i-1}) = Ker(u^i)$ pour un certain $i \in [\![1, r]\!]$ et considérons $x \in Ker(u^{i+1})$. Alors $u(x) \in Ker(u^i) = Ker(u^{i-1})$ et on a

$$u^i(x) = u^{i-1}(u(x)) = 0$$

et donc $x \in Ker(u^i)$. On a donc par récurrence

$$Ker(u^{i-1}) = Ker(u^i) = Ker(u^{i+1}) = \cdots = Ker(u^r) = E$$

soit encore $u^{i-1} = 0$ avec $i - 1 < r$, ce qui est absurde par minimalité de r.

□

On peut maintenant démontrer le résultat suivant.

Proposition 22 (Jordan pour les nilpotents). *Soit $u \in \mathcal{L}(E)$ nilpotent, alors il existe une base de E dans laquelle la matrice de u est diagonale par blocs avec des blocs de la forme $J_{r_i,0}$ où $r_1 + \cdots + r_p = dim(E)$.*

Démonstration : On commence par construire par récurrence une suite de supplémentaires aux noyaux itérés de u, c'est à dire des sous-espaces $F_0, F_1, \cdots, F_r, F_{r+1}$ de E vérifiant

$$Ker(u^i) = Ker(u^{i-1}) \oplus F_i \quad \text{et} \quad u(F_i) \subset F_{i-1} \quad \text{pour } i \in [\![1, r]\!]$$

où r désigne l'indice de nilpotence de u. En posant $F_{r+1} = \{0\}$ les propriétés sont respectées, peu importe le choix de F_r. Supposons à présent la construction faite au rang $i \in [\![1, r+1]\!]$ et construisons F_{i-1}.
Si $i = 1$, comme $Ker(u^0) = \{0\}$ on a $F_1 = Ker(u)$ et poser $F_0 = \{0\}$ satisfait les propriétés attendues.
Si $i \geq 2$, on peut tout d'abord noter que $u(F_i)$ et $Ker(u^{i-2})$ sont en somme directe. En effet, si $y \in F_i$ et $x = u(y)$ vérifie $u^{i-2}(x) = 0$ alors $y \in F_i \cap Ker(u^{i-1})$ et donc $y = 0$ par construction de F_i, d'où $x = 0$. On définit alors F_{i-1} comme étant un supplémentaire de $Ker(u^{i-2})$ dans $Ker(u^{i-1})$, contenant $u(F_i)$. F_{i-1} ainsi construit vérifie les hypothèse qu'on en attend et achève la construction.
Par construction des F_i on peut noter d'une part que la restriction de u à F_i est injective pour $i \geq 2$ et que d'autre part

$$E = F_1 \oplus \cdots \oplus F_r$$

On va alors construire une base de E en construisant une base $(e_{i,1}, \cdots, e_{i,d_i})$ de F_i pour $i \in [\![1, r]\!]$.
Fixons pour commencer la construction une base $(e_{r,1}, \cdots, e_{r,d_r})$ de F_r. Supposons la construction faite au rang $i \in [\![2, r]\!]$, i.e. on a une base $(e_{i,1}, \cdots, e_{i,d_i})$ de F_i. Comme $u_{|F_i}$ est injective, l'image de la base par u est une famille libre de F_{i-1}, que l'on peut compléter en une base $(e_{i-1,1}, \cdots, e_{i-1,d_{i-1}})$ de F_{i-1}, notons qu'on a alors $e_{i-1,k} = u(e_{i,k})$ pour tout $k \in [\![1, d_i]\!]$.
On a donc construit une base de E :

$$(e_{1,1}, \cdots, e_{1,d_1}, \cdots, e_{r,1}, \cdots, e_{r,d_r})$$

reste à ordonner les vecteurs de sorte qu'on obtienne bien une base de Jordan.
On a un tableau de la forme suivante :

$$
\begin{array}{llll}
F_r & e_{r,1}, \cdots, e_{r,d_r} & & \\
F_{r-1} & u(e_{r,1}), \cdots, u(e_{r,d_r}) & e_{r-1,d_r+1}, \cdots, e_{r-1,d_{r-1}} & \\
F_{r-2} & u^2(e_{r,1}), \cdots, u^2(e_{r,d_r}) & u(e_{r-1,d_r+1}), \cdots, u(e_{r-1,d_{r-1}}) & e_{r-2,d_{r-1}+1}, \cdots, e_{r-2,d_{r-2}} \\
\vdots & & & \\
F_1 & u^{r-1}(e_{r,1}), \cdots, u^{r-1}(e_{r,d_r}) & u^{r-2}(e_{r-1,d_r+1}), \cdots & \cdots
\end{array}
$$

En lisant le tableau colonne par colonne, de haut en bas, puis de gauche à droite, on ordonne les vecteurs de base en une famille $\mathcal{B} = \{e_1, \cdots, e_n\}$ de sorte que $u(e_j) = e_{j+1}$ si e_j n'est pas sur la dernière ligne, et 0 sinon, $\mathcal{B}$ est donc bien une base de Jordan pour u.

□

Le théorème de réduction de Jordan découle directement de la décomposition

$$E = \bigoplus_{i=1}^{r} Ker\left((u - \lambda_i \cdot id_E)^{m_i}\right)$$

en appliquant le résultat pour les nilpotents sur chaque sous-espace caractéristique de u.

3 – Décomposition de Dunford/Jordan-Chevalley

On a vu qu'à défaut de toujours pouvoir proposer une base diagonalisation, on peut au moins appliquer la réduction de Jordan dans le cas d'un corps algébriquement clos et obtenir une base de trigonalisation. En prenant l'exemple du calcul des puissances d'une matrice donnée, il est certes plus simple de le faire si la matrice en question T est triangulaire mais l'exercice reste pénible. L'astuce réside ici dans le fait que si $T \in \mathcal{M}_n(\mathbb{K})$ est triangulaire, alors $T = D + N$ où D est sa partie diagonale et N est triangulaire supérieure stricte donc nilpotente. Comme D et N commutent il est alors facile de calculer les puissances de T via la formule de Newton :

$$T^m = \sum_{k=0}^{n-1} \binom{m}{k} D^{m-k} N^k$$

Le but de la décomposition de Dunford (ou Jordan-Chevalley) est de montrer qu'on peut toujours se ramener à un tel calcul après changement de base approprié.

Définition 23 (Projecteurs spectraux). *Soient $u \in \mathcal{L}(E)$ tel que χ_u soit scindé sur $\mathbb{K}$ et $\lambda \in \mathbb{K}$ une valeur propre de u. On appelle projecteur spectral associé à λ l'unique endomorphisme $p_\lambda \in \mathcal{L}(E)$ vérifiant :*

$$
\begin{aligned}
&p_\lambda(x) = x \text{ pour } x \in Ker\left((u - \lambda \cdot id_E)^{m_\lambda}\right) \\
&p_\lambda(x) = 0 \text{ pour } x \in Ker\left((u - \mu \cdot id_E)^{m_\mu}\right) \text{ si } \mu \neq \lambda
\end{aligned}
$$

En d'autres termes, p_λ est la projection sur $Ker\left((u - \lambda \cdot id_E)^{m_\lambda}\right)$ parallèlement à $\bigoplus_{\mu \neq \lambda} Ker\left((u - \lambda_\mu \cdot id_E)^{m_\mu}\right)$.

Ces projecteurs apparaissent dans la démonstration de la décomposition de Dunford sous la forme du lemme suivant :

Lemme 24. *Soit $u \in \mathcal{L}(E)$ tel que χ_u soit scindé :*

$$\chi_u = \prod_{\lambda \in Sp(u)} (X - \lambda)^{m_\lambda}$$

On note pour $\lambda \in Sp(u)$

$$Q_\lambda := \prod_{\mu \neq \lambda} (X-\mu)^{m_\mu} = \frac{\chi_u}{(X-\lambda)^{m_\lambda}}$$

Alors il existe des polynômes $U_\lambda \in \mathbb{K}[X]$ *tels que*

$$\frac{1}{\chi_u} = \sum_{\lambda \in Sp(u)} \frac{U_\lambda}{(X-\lambda)^{m_\lambda}}$$

et de plus les projecteurs spectraux s'écrivent

$$p_\lambda = (Q_\lambda U_\lambda)(u)$$

Démonstration : En écrivant la décomposition en éléments simples de $\frac{1}{\chi_u}$ et en réduisant au même dénominateur les termes avec le même pôle on obtient l'existence des U_λ. Si maintenant on remplace la définition de Q_λ on obtient :

$$\frac{1}{\chi_u} = \sum_{\lambda \in Sp(u)} \frac{U_\lambda}{(x-\lambda)^{m_\lambda}} = \sum_{\lambda \in Sp(u)} \frac{Q_\lambda U_\lambda}{\chi_u}$$

ce qui assure que

$$\sum_{\lambda \in Sp(u)} Q_\lambda U_\lambda = 1$$

Si $\mu \neq \lambda$, alors $(X-\mu)^\mu$ divise Q_λ et alors $(Q_\lambda U_\lambda)(u)(x) = 0$ pour $x \in Ker\left((u - \mu \cdot id_E)^{m_\mu}\right)$. En utilisant l'identité si dessus évaluée en u puis en x, tous les termes de la somme s'annulent sauf celui correspondant à λ. On a alors

$$x = \sum_{\mu \in Sp(u)} (Q_\mu U_\mu)(u)(x) = (Q_\lambda U_\lambda)(u)(x)$$

ce qui caractérise bien p_λ.

□

On peut maintenant passer à la preuve de la décomposition annoncée plus haut :

Théorème 25 (Décomposition de Dunford). *Soit* $u \in \mathcal{L}(E)$ *tel que* χ_u *soit scindé, alors il existe deux endomorphismes* $d, \nu \in \mathcal{L}(E)$ *qui commutent tels que* d *soit diagonalisable,* ν *nilpotent et* $u = d + \nu$.
Une telle décomposition est unique, en particulier d *et* ν *sont des polynômes en* u *et*

$$d = \sum_{\lambda \in Sp(u)} \lambda p_\lambda$$

Démonstration :
• **Existence :** Soit $d = \sum_{\lambda \in Sp(u)} \lambda p_\lambda$ comme suggéré et $\nu = u - d$, on montre que cette décomposition convient. D'après le lemme précédent, les p_λ sont des polynômes en u, c'est donc aussi le cas pour d et ν, en particulier d et n commutent.
De plus, $E_\lambda^{m_\lambda}(u) = Ker((u - \lambda id_E)^{m_\lambda})$ est stable par u donc par d, on alors par définition de p_λ :

$$d_{|E_\lambda^{m_\lambda}(u)} = \lambda id_{E_\lambda^{m_\lambda}(u)}$$

ce qui prouve que d est diagonalisable en utilisant le lemme des noyaux pour avoir la somme directe

$$E = \bigoplus_{i=1}^{r} Ker\left((u - \lambda_i \cdot id_E)^{m_i}\right)$$

D'autre part, $E_\lambda^{m_\lambda}(u)$ est également stable par ν donc en considérant la restriction on a

$$\nu_{|E_\lambda^{m_\lambda}(u)} = (u - \lambda id_E)_{|E_\lambda^{m_\lambda}(u)}$$

$\nu_{|E_\lambda^{m_\lambda}(u)}$ est donc nilpotent d'indice $\dim(E_\lambda^{m_\lambda}(u)) \leq n$, en particulier on a

$$(\nu^n)_{|E_\lambda^{m_\lambda}(u)} = \left(\nu_{|E_\lambda^{m_\lambda}(u)}\right)^n = 0$$

Toujours de la même décomposition en somme directe, on déduit que ν est nilpotent.
• **Unicité :** supposons qu'on ait une deuxième décomposition $u = d' + \nu'$ comme annoncé. On a alors $d - d' = \nu' - \nu$ et comme d' commute avec ν', donc avec u, d' commute avec tout polynôme en u, en particulier avec d. Le même raisonnement assure que ν' et ν commutent et alors, $d - d'$ est diagonalisable et $\nu' - \nu$ est nilpotent et de l'égalité $d - d' = \nu' - \nu$ on a $d = d'$ et $\nu = \nu'$.

□

Cette démonstration fournit une recette explicite pour fabriquer la décomposition de Dunford à condition de pouvoir calculer explicitement les valeurs propres et les projecteurs spectraux. Ce calcul peut être assez compliqué en général du fait de devoir factoriser le polynôme χ_u, on propose une autre méthode démonstration plus algorithmique qui permet d'éviter ce calcul (en tout cas si $\mathbb{K}$ est de caractéristique nulle).

Théorème 26 (Méthode de Newton pour la décomposition de Dunford)**.** *Soit $u \in \mathcal{L}(E)$ tel que χ_u soit scindé sur $\mathbb{K}$ et posons $P = \prod_{\lambda \in Sp(u)}(X - \lambda)$. On définit la suite $(u_n) \subset \mathcal{L}(E)$ par :*

$$u_0 = u \quad et \quad \forall n \in \mathbb{N}\,,\ u_{n+1} = u_n - P(u_n) \circ (P'(u_n))^{-1}$$

Alors (u_n) est bien définie et est stationnaire. Si n est un rang où u_n est stationnaire, $d := u_n$ et $\nu = u - u_n$ est la décomposition de Dunford de u.

Démonstration : Par définition de P, on a $P | \chi_u$ et si $r = \max(m_\lambda(u))$ alors on a aussi $\chi_u | P^r$. En particulier, $P(u)^r = P^r(u) = 0$ donc $P(u)$ est nilpotent.
D'autre part, P est à racines simples donc P et P' sont sans racines communes et P' et χ_u sont premiers entre eux. Il existe alors $U, V \in \mathbb{K}[X]$ tels qu'on ait une relation de Bézout :

$$UP' + V\chi_u = 1$$

et par Cayley-Hamilton, $U(u) \circ P'(u) = (UP')(u) = id_E$. $P'(u)$ est donc inversible et son inverse est un polynôme en u.
On montre maintenant que la définition de u_n est bien valide à tout rang n. On pose l'hypothèse de récurrence

$$(H_n)\ :\ u_n \in \mathbb{K}[u]\,,\ P'(u_n) \in \mathbb{K}[u]^*\,,\ P(u_n) \in P(u)^{2^n}\mathbb{K}[u]$$

Supposons (H_n) vérifiée jusqu'à un certain rang n. Comme $P'(u_n)$ est inversible dans $\mathbb{K}[u]$ et que $P(u_n)$ est un polynôme en u_n, donc en u d'après H_n, on a bien $u_{n+1} \in \mathbb{K}[u]$.
On a également l'identité

$$P'(u_{n+1}) - P'(u_n) = (u_{n+1} - u_n) \circ Q(u)$$

pour un certain $Q \in \mathbb{K}[X]$. En effet, les termes en id_E de $P'(u_{n+1})$ et de $P'(u_n)$ s'annulent et il n'y a alors que des termes de la forme

$$u_{n+1}^k - u_n^k = (u_{n+1} - u_n)\left(u_{n+1}^{k-1} + u_{n+1}^{k-2}u_n + \cdots + u_{n+1}u_n^{k-2} + u_n^{k-1}\right)$$

où le terme le plus à droite est juste un polynôme en u. En reprenant la définition de u_{n+1}, on a $u_{n+1} - u_n = P(u_n) \circ (-P'(u_n)^{-1})$ et comme $P'(u_n)^{-1} \in \mathbb{K}[u]$ et

$P(u_n) \in P(u)^{2^n}\mathbb{K}[u]$ où $P(u)$ est nilpotent, il en est de même pour $u_{n+1} - u_n$, et donc aussi pour $P'(u_{n+1}) - P'(u_n)$. Ainsi,

$$P'(u_{n+1}) = (P'(u_{n+1}) - P'(u_n)) + P'(u_n)$$

est la somme d'un nilpotent et d'un inversible de $\mathbb{K}[u]$, donc est lui-même inversible dans $\mathbb{K}[u]$.
Enfin, il reste à voir que $P(u_{n+1}) \in P(u)^{2^{n+1}}\mathbb{K}[u]$. Pour ça il suffit de remarquer que si $Q \in \mathbb{K}[X]$, on a (il suffit de le montrer sur les monômes)

$$Q(X+Y) = Q(X) + Q'(X)Y + Y^2\mathbb{K}[X,Y]$$

et comme u_n et $-P(u_n) \circ P'(u_n)^{-1}$ commutent, on peut les substituer respectivement à X et Y. On obtient alors en prenant $Q = P$:

$$P(u_{n+1}) = P\left(u_n - P(u_n) \circ P'(u_n)^{-1}\right) \in P(u_n)^2\mathbb{K}[u]$$

d'où le résultat comme $P(u_n) \in P(u)^{2^n}\mathbb{K}[u]$.
Ceci étant montré, il suffit de choisir un indice m assez grand de sorte que $2^m \geq r$. On a alors $P(u)^{2^m} = 0$ et $P(u_m) = 0$, la suite est bien stationnaire.
Posons alors $d = u_m$ et $\nu = u - d$, par construction ce sont des polynômes en u, en particulier ils commutent. De plus $P(d) = 0$ et P est scindé à racines simples, donc d est diagonalisable.
Enfin, notons qu'on a à tout rang j

$$u_{j+1} - u_j \in P(u)\mathbb{K}[u]$$

et alors en sommant sur j, $\nu = u_0 - u_m \in P(u)\mathbb{K}[u]$ ce qui assure que ν est nilpotent car $P(u)$ l'est.

□

Cette méthode de démonstration alternative a l'avantage de ne pas nécessiter le calcul des valeurs propres de u si $\mathbb{K}$ est de caractéristique nulle. En effet, dans ce cas le fait que λ soit racine de multiplicité m de χ_u implique que λ est racine de multiplicité $m-1$ de χ_u'. P s'obtient alors de façon explicite par

$$P = \frac{\chi_u}{\chi_u \wedge \chi_u'}$$

où le PGCD $\chi_u \wedge \chi_u'$ peut se calculer explicitement via l'algorithme d'Euclide. On a donc une méthode exacte et purement algorithmique pour déterminer la décomposition de Dunford.

4 – Exponentielle dans une algèbre normée

Dans toute cette partie, les résultats mettront en jeu (plus ou moins explicitement) des séries convergentes, on considère donc pour la suite que $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$ afin d'assurer une bonne topologie. Dès lors qu'on travaille dans une algèbre de dimension finie sur $\mathbb{K}$, on peut se munir d'une norme quelconque (toutes les normes étant équivalentes) et en particulier d'une norme d'algèbre.

Définition 27. *Étant donnée une $\mathbb{K}$-algèbre de dimension finie A, une norme $\|\cdot\|$ sur A est dite d'algèbre (ou sous-multiplicative) si*

$$\forall x, y \in A\,,\ \|x \times y\| \leq \|x\| \cdot \|y\|$$

Par exemple, sur les espaces de matrices/applications linéaires, toute norme subordonnée est une norme d'algèbre. En particulier, pour une norme $\|\cdot\|$ quelconque sur A, la norme suivante est une norme d'algèbre sur A :

$$\forall x \in A\,,\ \|x\|_A := \sup_{\|y\|=1} \|x \times y\|$$

A étant de dimension finie, on peut s'intéresser à la convergence de séries dans A via leur convergence normale (en dimension infinie, on aurait besoin d'avoir également de la complétude, c'est la notion d'algèbre de Banach).

Définition 28. *Pour une $\mathbb{K}$-algèbre de dimension finie A et $x \in A$, on définit l'exponentielle par la somme de la série normalement convergente suivante*

$$\exp(x) := \sum_{n \geq 0} \frac{1}{n!} x^n$$

On peut définir de façon analogue n'importe quelle somme de série entière sur A, pourvu que la série converge normalement. L'exponentielle joue néanmoins un rôle important dans de nombreux domaines, notamment pour la résolution d'équations différentielles ou l'étude d'algèbres de Lie.
Notons que comme l'espace des polynômes en un élément est un sous-espace vectoriel de A, il est fermé et donc $\exp(x)$ est un polynôme en x.
On déduit facilement de la définition les propriétés élémentaires suivantes.

Proposition 29. *Soient $x, y \in A$, si x et y commutent alors x et $\exp(y)$ commutent, de plus*

$$\exp(x+y) = \exp(x)\exp(y)$$

Pour tout élément $x \in A$, $\exp(x)$ est inversible dans A et

$$\exp(x)^{-1} = \exp(-x)$$

Si $M \in A = \mathcal{M}_n(\mathbb{K})$ et $P \in GL_n(\mathbb{K})$ alors

$$P\exp(M)P^{-1} = \exp(PMP^{-1})$$

La difficulté première dans les applications de l'exponentielle matricielle concerne son calcul pratique. La formule donnant $\exp(M)$ rendant les choses assez compliquées pour une matrice quelconque M, il existe quelques cas particuliers où le calcul est relativement simple, en voici une liste non-exhaustive.

— Si $M \in \mathcal{M}_n(\mathbb{K})$ est diagonalisable :

$$\text{si } M = P \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} P^{-1} \quad \text{alors} \quad \exp(M) = P \begin{pmatrix} e^{\lambda_1} & & \\ & \ddots & \\ & & e^{\lambda_n} \end{pmatrix} P^{-1}$$

— Si M est nilpotente, le calcul est simplement une somme finie.

— Si on a accès à la décomposition de Dunford $M = D + N$, alors comme D et N commutent on peut combiner les deux cas précédents :

$$\exp(M) = \exp(D)\exp(N) = \exp(D)\left(I_n + N + \frac{1}{2!}N^2 + \cdots + \frac{1}{(r-1)!}N^{r-1}\right)$$

Dans les autres cas (au moins sur $\mathbb{C}$) on peut utiliser la décomposition de Jordan de M pour calculer $\exp(M)$, moyennant le calcul de l'exponentielle d'un bloc de Jordan.

Proposition 30. *Soient $t, \lambda \in \mathbb{K}$ et $r \geq 1$, on a*

$$\exp(tJ_{r,\lambda}) = e^{\lambda t}\exp(tJ_{r,0}) \begin{pmatrix} 1 & t & \frac{1}{2!}t^2 & \cdots & \frac{1}{(r-1)!}t^{r-1} \\ & \ddots & \ddots & \ddots & \vdots \\ & & \ddots & \ddots & \frac{1}{2!}t^2 \\ & & & \ddots & t \\ & & & & 1 \end{pmatrix}$$

Ayant en tête l'application aux équations différentielles, il est judicieux d'avoir le résultat pour $t \in \mathbb{K}$, d'autant que la substitution $t = 1$ ne pose pas tellement de problème. Dans l'espoir de définir un potentiel "logarithme", il peut être utile de décrire précisément l'image de $\mathcal{M}_n(\mathbb{K})$ par l'exponentielle suivant que $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$. On commence par le résultat important suivant.

Théorème 31. *Soit $M \in \mathcal{M}_n(\mathbb{C})$, alors*

$$\exp(\mathbb{C}[M]) = \mathbb{C}[M]^*$$

Démonstration : Notons tout d'abord que

$$\mathbb{C}[M]^* = \mathbb{C}[M] \cap GL_n(\mathbb{C})$$

En effet, la première inclusion est immédiate et pour la deuxième, étant donnée une matrice $B \in \mathbb{C}[M] \cap GL_n(\mathbb{C})$, on a $\det(B) \neq 0$ et le théorème de Cayley-Hamilton fournit alors un inverse à B sous la forme d'un polynôme en B, donc un inverse dans $\mathbb{C}[M]$ car B est elle-même un polynôme en M.
On montre ensuite que $\mathbb{C}[M]^*$ est un connexe de $\mathbb{C}[M]$. En effet, soient $M_0, M_1 \in \mathbb{C}[M]^*$, le chemin naïf

$$M_t := tM_1 + (1-t)M_0, \ , \ t \in [0,1]$$

reste dans $\mathbb{C}[M]$ mais pas nécessairement dans $GL_n(\mathbb{C})$. Cependant, la fonction $t \in \mathbb{C} \mapsto \det(M_t)$ est polynomiale donc ne s'annule qu'un nombre fini de fois sur $\mathbb{C}$. $\mathbb{C}$ privé de ces points étant connexe par arcs, il existe un chemin continu $\gamma : [0,1] \mapsto \mathbb{C}$ qui relie 0 à 1 en évitant ces points. Ainsi, l'application continue $t \in [0,1] \mapsto M_{\gamma(t)}$ définit un chemin continu de M_0 à M_1 qui reste dans $\mathbb{C}[M]^*$, ce qui assure que $\mathbb{C}[M]^*$ est connexe par arcs donc connexe.
Notons qu'on a déjà l'inclusion $\exp(\mathbb{C}[M]) \subset \mathbb{C}[M]^*$, il suffit donc de montrer que $\exp(\mathbb{C}[M])$ est ouvert et fermé dans $\mathbb{C}[M]^*$, comme $\exp(\mathbb{C}[M])$ est non vide, on aura le résultat par connexité.
Montrons tout d'abord que c'est un ouvert. En calculant la différentielle en 0 on a

$$d_0 \exp = id_{\mathbb{C}[M]}$$

qui est inversible. On peut donc appliquer le théorème d'inversion locale dans $\mathbb{C}[M]$, ce qui fournit un voisinage ouvert $\mathcal{U}$ de 0 dans $\mathbb{C}[M]$ et $\mathcal{V}$ de I_n dans $\mathbb{C}[M]^*$ de sorte que exp envoie $\mathcal{U}$ sur $\mathcal{V}$. Pour $B \in \mathbb{C}[M]$, on définit un voisinage ouvert de $\exp(B)$ dans $\exp(\mathbb{C}[M])$ par

$$\mathcal{V}_B := \{V \exp(B)\, , \ V \in \mathcal{V}\}$$

En effet, si $V \in \mathcal{V}$, il existe $U \in \mathcal{U}$ tel que $\exp(U) = V$, comme toutes les matrices en jeu sont des polynômes en M elles commutent et on a

$$V \exp(M) = \exp(U)\exp(B) = \exp(U + B) \in \exp(\mathbb{C}[M])$$

ce qui prouve que $\exp(\mathbb{C}[M])$ est ouvert dans $\mathbb{C}[M]^*$.
Pour montrer que c'est un fermé, notons $\mathcal{X} := \mathbb{C}[M]^* \setminus \exp(\mathbb{C}[M])$ son complémentaire dans $\mathbb{C}[M]^*$ et montrons que c'est un ouvert. Il suffit de montrer qu'en fait

$$\mathcal{X} := \bigcup_{B \in \mathcal{X}} B \exp(\mathbb{C}[M])$$

on conclut alors en utilisant le fait que $B \exp(\mathbb{C}[M])$ est ouvert dès que B est inversible. La première inclusion est immédiate car $I_n \in \exp(\mathbb{C}[M])$. Réciproquement, soient $B \in \mathcal{X}$ et $P \in \mathbb{C}[X]$, on définit la matrice $N := B \exp(P(M))$. Alors $N \notin \exp(\mathbb{C}[M]))$ car sinon on aurait $B = N \exp(-P(M)) \in \exp(\mathbb{C}[M])$. On a donc bien l'inclusion inverse, ce qui achève la preuve.

□

Remarque : *On utilise le théorème d'inversion locale bien entre les espaces* $\mathbb{C}[A]$ *et* $\exp(\mathbb{C}[A])$ *pour obtenir des voisinages dans les bons espaces, et pas seulement dans* $\mathcal{M}_n(\mathbb{C})$*. La démonstration repose de façon cruciale sur le fait que le corps de base est* $\mathbb{C}$ *pour utiliser la connexité, en particulier, ces arguments tombent en défaut dans le cas de* $\mathbb{R}$*.*

Corollaire 32 (Surjectivité de l'exponentielle). *L'application* $\exp : \mathcal{M}_n(\mathbb{C}) \to GL_n(\mathbb{C})$ *est surjective.*
L'image de $\exp : \mathcal{M}_n(\mathbb{R}) \to GL_n(\mathbb{R})$ *est*

$$\left\{M^2\,,\ M \in GL_n(\mathbb{R})\right\}$$

Démonstration : Soit $M \in GL_n(\mathbb{C})$, on a $M \in \mathbb{C}[M]^*$ et donc par le théorème précédent, il existe $P \in \mathbb{C}[X]$ tel que $M = \exp(P(M))$, ce qui prouve la surjectivité dans la cas complexe.
Pour le cas réel, l'inclusion directe découle de l'identité

$$\left(\exp(\frac{1}{2}M)\right)^2 = \exp(M)$$

Réciproquement, comme $M \in GL_n(\mathbb{R})$, il existe $P \in \mathbb{C}[X]$ tel que $M = \exp(P(M))$. Notons qu'on a aussi $M = \overline{M} = \exp(\overline{P}(M))$ car M est à coefficients réels. On conclut avec l'identité

$$M^2 = \exp\left((P + \overline{P})(M)\right)$$

□

5 – Calcul fonctionnel

Dans cette section on digresse et on s'intéresse à une généralisation de la section précédente, toujours an travaillant avec $\mathbb{C}$ comme corps de base donc. La question centrale est de définir $f(A)$ pour un ensemble de fonctions suffisamment large. Dans le cas où f est polynomiale, il n'y a aucune ambiguïté de définition car on n'effectue que des opérations d'algèbre. Si A est diagonalisable, le calcul est facile en le reportant sur chacune des directions propres. On propose ici une vision plus générale basée sur la formule de Cauchy, on retrouvera au passage les résultats classiques énoncés plus tôt (Cayley-Hamilton, Jordan, Dunford...). L'avantage principal de cette méthode est, d'une part, qu'elle ne restreint pas l'ensemble des matrices et peu l'ensemble des fonctions auxquelles elle s'applique, et d'autre part qu'elle s'adapte bien à la dimension infinie en remplaçant les matrices par des opérateurs compacts.
A titre d'exemple, si $A \in \mathcal{M}_n(\mathbb{C})$ est diagonalisable

$$A = P \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} P^{-1}$$

et si f est définie sur un voisinage de $\sigma(A) := Sp(A)$ on définit $f(A)$ comme suit

$$f(A) = P \begin{pmatrix} f(\lambda_1) & & \\ & \ddots & \\ & & f(\lambda_n) \end{pmatrix} P^{-1}$$

On voit alors que toute l'information importante pour définir $f(A)$ est contenue dans le spectre. On se fixe une norme subordonnée $|||\cdot|||$ sur $\mathcal{M}_n(\mathbb{C})$ et on commence par quelques résultats d'ordre généraux.

Proposition 33 (Résolvante). *Soit $A \in \mathcal{M}_n(\mathbb{C})$, l'application*

$$R_A : z \mapsto (zI_n - A)^{-1}$$

est méromorphe et ses pôles coïncident avec les valeurs propres de A, d'ordre inférieur ou égal à leur multiplicité algébrique.

Démonstration : La formule de la commatrice permet d'écrire les coefficients de $(zI_n - A)^{-1}$ comme des fractions rationnelles en z dont les potentiels pôles sont exactement les valeurs propres de A. Reste à montrer que ce sont effectivement des pôles : si $z \mapsto (zI_n - A)^{-1}$ est borné au voisinage d'un certain $\lambda \in \mathbb{C}$, alors elle admet une valeur d'adhérence R_{λ_0} en λ_0 comme suite bornée en dimension finie. Mais alors, en passant à la limite on obtient

$$(\lambda_0 I_n - A)R_{\lambda_0} = I_n$$

et donc $\lambda_0 \notin \sigma(A)$.

□

Le fait que R_A définisse une fonction méromorphe nous permet en fait d'en considérer l'intégrale sur n'importe quel contour qui ne rencontre pas $\sigma(A)$, c'est de cette manière qu'on va considérer $f(A)$, dans un premier temps pour des fonctions entières, puis d'une façon plus générale.

Proposition 34 (Formule de Cauchy matricielle). *Soit $A \in \mathcal{M}_n(\mathbb{C})$ et f une fonction holomorphe sur $\mathbb{C}$. On note*

$$f(A) := \sum_{k\geq 0} \frac{f^{(k)}(0)}{k!} A^k$$

Alors, pour tout lacet simple γ orienté positivement et contenant $\sigma(A)$ dans son intérieur on a la formule de Cauchy :

$$f(A) = \frac{1}{2i\pi} \int_\gamma f(z)(zI_n - A)^{-1} dz$$

Démonstration : Comme R_A est holomorphe en dehors du spectre, l'intégrale ne dépend du contour qu'au travers de la propriété "lacet simple entourant positivement $\sigma(A)$". On peut donc l'élargir pour qu'il contienne 0 et que $|z| > \|\|A\|$ pour tout $z \in \gamma$, cette deuxième condition permet d'écrire

$$R_A(z) = (zI_n - A)^{-1} = \sum_{k\geq 0} z^{-(k+1)} A^k$$

et en poursuivant le calcul :

$$\begin{aligned}
\frac{1}{2i\pi} \int_\gamma f(z)(zI_n - A)^{-1} dz &= \frac{1}{2i\pi} \sum_{k\geq 0} \int_\gamma f(z) z^{-(k+1)} A^k dz \\
&= \sum_{m,k\geq 0} \frac{1}{2i\pi} \int_\gamma \frac{f^{(m)}(0)}{m!} z^{m-(k+1)} A^k dz \\
&= \sum_{m,k\geq 0} \frac{f^{(m)}(0)}{m!} \left(\frac{1}{2i\pi} \int_\gamma z^{m-(k+1)} dz \right) A^k \\
&= \sum_{m\geq 0} \frac{f^{(m)}}{m!} A^m = f(A)
\end{aligned}$$

□

La définition par une intégrale de $f(A)$ se démarque de celle par les séries entières dans la liberté de choix du contour γ. En effet, on peut localiser aussi proche de $\sigma(A)$ que l'on veut et donc considérer $f(A)$ pour des fonctions que ne sont holomorphes que sur un voisinage de $\sigma(A)$.

Théorème 35 (Calcul fonctionnel). *Soit $A \in \mathcal{M}_n(\mathbb{C})$ et Ω un voisinage ouvert de $\sigma(A)$, il existe une unique application $\Phi_A : \mathcal{H}(\Omega) \to \mathcal{M}_n(\mathbb{C})$ vérifiant :*
— *Φ_A est un morphisme d'algèbre*
— *Φ_A est continue pour la topologie de la convergence uniforme sur tout compact*
— *$\Phi_A(id_\Omega) = A$ et $\Phi_A(1) = I_n$*
De plus on a les propriétés suivantes :
— *si γ est une réunion de cercles disjoints, simples et positivement orientés, centrés en des valeurs propres de A alors pour toute fonction $f \in \mathcal{H}(\Omega)$:*

$$\Phi_A(f) = \frac{1}{2i\pi}\int_\gamma f(z)(zI_n - A)^{-1}dz$$

— *si $f \in \mathcal{H}(\Omega)$, $\sigma(\Phi_A(f)) = f(\sigma(A))$*
— *$\Phi_A(f)$ ne dépend pas du choix du voisinage Ω tel que $f \in \mathcal{H}(\Omega)$*

Démonstration :

• **Existence :** il suffit de vérifier que la formule de Cauchy proposée convient. Concernant la continuité, si $f \in \mathcal{H}(\Omega)$, alors pour toute norme $\|\cdot\|$ sur $\mathcal{M}_n(\mathbb{C})$:

$$\|\Phi_A(f)\| \leq \frac{1}{2\pi}\int_\gamma |f(z)|\|(zI_n - A)^{-1}\||dz| \leq \left(\frac{|\gamma|}{2\pi}\sup_{z\in\gamma}\|(zI_n - A)^{-1}\|\right)\sup_{z\in\gamma}|f(z)|$$

ce qui assure la continuité pour la topologie de la convergence uniforme sur tout compact. On a déjà remarqué que l'intégrale ne dépendait pas du contour, on ne change donc pas $\Phi_A(f)$ en faisant un choix de voisinage ouvert tel que $f \in \mathcal{H}(\Omega)$. En particulier pour le calcul de $\Phi_A(z \mapsto z)$ et $\Phi_A(1)$ on peut considérer γ aussi large que nécessaire (comme dans la preuve précédente) et on obtient alors :

$$\Phi_A(z \mapsto z) = \frac{1}{2i\pi}\int_\gamma z(zI_n - A)^{-1}dz = \sum_{k\geq 0}\left(\frac{1}{2i\pi}\int_\gamma z^{-k}dz\right)A^k = A$$

et le même calcul donne aussi $\Phi_A(1) = I_n$. Il reste alors à vérifier la propriété de morphisme d'algèbre. Considérons $f, g \in \mathcal{H}(\Omega)$, donné chacun par une intégrale sur un contour γ, η respectivement. Encore une fois, on peut déformer les lacets sans changer la valeur de l'intégrale, on peut donc supposer librement que η est contenu dans l'intérieur de γ et écrire :

$$\Phi_A(f)\Phi_A(g) = \left(\frac{1}{2i\pi}\right)^2\int_{z\in\gamma,\zeta\in\eta} f(z)g(\zeta)(zI_n - A)^{-1}(\zeta I_n - A)^{-1}dzd\zeta$$

Pour calculer le produit $(zI_n - A)^{-1}(\zeta I_n - A)^{-1}$ on utilise l'identité de résolvante :

$$(\zeta - z)R_A(z)R_A(\zeta) = R_A(z) - R_A(\zeta)$$

En effet, en remarquant que

$$(zI_n - A)^{-1} = (zI_n - A)^{-1}(\zeta I_n - A)(\zeta I_n - A)^{-1}$$

et l'analogue pour ζ

$$(\zeta I_n - A)^{-1} = (zI_n - A)^{-1}(zI_n - A)(\zeta I_n - A)^{-1}$$

on obtient l'identité souhaitée en factorisant. On peut alors effectuer l'intégration par rapport à ζ :

$$\Phi_A(f)\Phi_A(g) = \frac{1}{2i\pi}\int_\gamma f(z)\left(\frac{1}{2i\pi}\int_\eta \frac{g(\zeta)}{\zeta - z}\left((zI_n - A)^{-1} - (\zeta I_n - A)^{-1}\right)d\zeta\right)dz$$

Par choix de γ et η, pour $z \in \gamma$ fixé, la fonction $\zeta \mapsto \frac{g(\zeta)}{\zeta - z}$ n'a pas de pôle à l'intérieur du contour η, d'où par formule de Cauchy :

$$\frac{1}{2i\pi}\int_\gamma f(z)\left(\frac{1}{2i\pi}\int_\eta \frac{g(\zeta)}{\zeta - z}d\zeta\right)(zI_n - A)^{-1}dz = 0$$

et par le même raisonnement pour $z \mapsto \frac{f(z)}{\zeta - z}$ on a

$$\frac{1}{2i\pi}\int_\gamma f(z)\left(\frac{1}{2i\pi}\int_\eta \frac{g(\zeta)}{\zeta - z}(\zeta I_n - A)^{-1}d\zeta\right)dz = -\frac{1}{2i\pi}\int_\eta f(\zeta)g(\zeta)(\zeta I_n - A)^{-1}d\zeta$$

ce qui prouve bien la propriété de morphisme d'algèbre.
• **Unicité :** ayant la continuité, notons que la propriété de morphisme d'algèbre et le fait de spécifier $\Phi_A(id_\Omega)$ et $\Phi_A(1)$ déterminent entièrement Φ_A sur l'espace des fonctions polynomiales, ce qui donne l'unicité par densité pour la topologie de la convergence uniforme sur tout compact.
Enfin, concernant la propriété de spectre, supposons que $\lambda \in \sigma(A)$, alors $f(\lambda)I_n - f(A)$ n'est pas inversible. En effet en utilisant la propriété de morphisme on a

$$f(\lambda)I_n - f(A) = \Phi_A\left(z \mapsto \frac{f(\lambda) - f(z)}{\lambda - z}\right)(\lambda I_n - A)$$

où la fonction $z \mapsto \frac{f(\lambda)-f(z)}{\lambda - z}$ est holomorphe dans un voisinage de $\sigma(A)$. Donc $f(\lambda)I_n - f(A)$ n'est pas inversible car $\lambda I_n - A$ ne l'est pas, et alors $f(\lambda) \in \sigma(f(A))$.
Réciproquement, si $\mu \notin f(\sigma(A))$, $\mu I_n - f(A)$ est inversible d'inverse $\Phi_A(z \mapsto (\mu - f(z))^{-1})$, où la fonction $z \mapsto (\mu - f(z))^{-1}$ est holomorphe au voisinage de $\sigma(A)$.

□

Le théorème de calcul fonctionnel est fondamental en théorie spectrale et permet de revisiter les points algébriques les plus importants vus dans ces notes, mais avec des outils d'analyse complexe. En suivant l'ordre logique qu'on a présenté ici, commençons par des résultats de réduction.

Lemme 36. *Soit $A \in \mathcal{M}_n(\mathbb{C})$ telle que $\sigma(A) = \{0\}$, alors A est nilpotente.*

Démonstration : En utilisant le théorème de Cayley-Hamilton, le résultat est direct, mais on peut s'en passer.
On a montré que la fonction $R_A : z \mapsto (zI_n - A)^{-1}$ a pour seul pôle 0, et est méromorphe sur $\mathbb{C} \setminus \{0\}$. En particulier son développement en série de Laurent n'admet qu'un nombre fini de termes d'indice négatif. Pour z assez grand, on a explicitement

$$R_A(z) = \sum_{k \geq 0} z^{-(k+1)}A^k$$

et alors A^k doit être nul à partir d'un certain rang.

□

Muni de ce lemme, on peut montrer le théorème de décomposition spectrale qui généralise à la fois la réduction de Jordan et la décomposition de Dunford.

Théorème 37 (Décomposition spectrale). *Soit $A \in \mathcal{M}_n(\mathbb{C})$, notons $\lambda_1, \cdots, \lambda_r$ les valeurs propres distinctes de A. Alors il existe des matrices $\Pi_1, \cdots, \Pi_r, N_1, \cdots, N_r$ définies comme des fonctions de A telles que :*
— *les Π_i sont les matrices des projecteurs spectraux respectivement sur $E_{\lambda_i}^{m_i}$, parallèlement à $\sum_{j \neq i} E_{\lambda_j}^{m_j}$*
— *les N_i sont nilpotentes, d'indice n_i inférieur ou égal à $rg(\Pi_i)$*
—

$$A = \sum_{i=1}^r (\lambda_i I_n + N_i)\Pi_i$$

De plus, si f est holomorphe au voisinage de $\sigma(A)$, alors

$$\Phi_A(f) =: f(A) = \sum_{i=1}^{r}\sum_{j=0}^{n_i-1}\frac{f^{(j)}(\lambda_i)}{j!}N_i^j\Pi_i$$

On remarque que toutes les matrices en jeu commutent (comme fonctions de A), on a alors exactement l'énoncé de la décomposition de Dunford (en séparant suivant que $j=0$ ou non) et de la réduction de Jordan (quitte à choisir une base adaptée aux Π_i, N_i). On pourrait dans ce cas se contenter des démonstrations vues plus haut, mais elles ne donnent les Π_i, N_i que de manière peu explicite comme des polynômes en A.

Démonstration : On cherche les Π_i comme des fonctions de A qui ne prennent en compte que la direction de $E_{\lambda_i}^{m_i}$ pour obtenir un projecteur, en d'autres termes on cherche des fonctions (g_i) qui localisent chacune autour de λ_i.
Précisément, on se donne une famille de fonctions (g_i) définies au voisinage de $\sigma(A)$ vérifiant que d'une part $g_i \equiv 1$ au voisinage de λ_i et $g_i \equiv 0$ au voisinage de $\sigma(A)\setminus\{\lambda_i\}$ (localiser), et d'autre part $g_ig_j = 0$ si $i \neq j$ et $\sum_i g_i \equiv 1$ (séparer l'espace). En particulier de telles fonctions sont holomorphes au voisinage de $\sigma(A)$ (quitte à prendre un voisinage encore plus petit).
On pose alors $\Pi_i := \Phi_A(g_i) = g_i(A)$ et $N_i := h_i(A)$ où $h_i : z \mapsto (z-\lambda_i)g_i(z)$. En particulier h_i est identiquement nulle sur un voisinage de $\sigma(A)\setminus\{\lambda_i\}$ et coïncide avec $z-\lambda_i$ sur un voisinage de λ_i. Dans ces conditions on a

$$id = \sum_{i=1}^{r}(\lambda_i + h_i)g_i$$

de laquelle découle la décomposition annoncée.
On peut aussi remarquer que Π_i est un projecteur en utilisant la propriété de morphisme et le fait que $g_i^2 = g_i$ sur un voisinage de $\sigma(A)$.
Pour les N_i, il vient du théorème que $\sigma(N_i) = h_i(\sigma(A)) = \{0\}$ et $N_i = (A-\lambda_iI_n)\Pi_i$ a nécessairement un rang inférieur ou égal à celui de Π_i.
La nilpotence de N_i implique que $\mathrm{Im}(\Pi_i)$ est inclus dans $E_{\lambda_i}^{m_i} = \mathrm{Ker}\left((\lambda_iI_n - A)^{m_i}\right)$ par définition. Réciproquement, si f_i est définie par $f_i \equiv 0$ au voisinage de λ_i et coïncide avec $z \mapsto (z-\lambda_i)^{-1}$ au voisinage de $\sigma(A)\setminus\{\lambda_i\}$, on a $f_i(A)(\lambda_iI_n - A) = \Phi_A\left(z \mapsto f_i(z)(\lambda_i - z)\right)$. Dans ces conditions, en utilisant le fait que les Π_i sont supplémentaires, on a

$$f_i(A)^{m_i}(\lambda_iI_n - A)^{m_i} = \sum_{j\neq i}\Pi_j = I_n - \Pi_i$$

ce qui prouve l'inclusion $E_{\lambda_i}^{m_i} \subset \mathrm{Im}(\Pi_i)$.
Enfin, si f est holomorphe sur un voisinage de $\sigma(A)$, il suffit d'injecter dans l'intégrale définissant $f(A)$ l'expression suivante de la résolvante :

$$\begin{aligned}
(zI_n - A)^{-1} &= \left(\sum_{i=1}^{r}\left(zI_n - (\lambda_iI_n + N_i)\right)\Pi_i\right)^{-1} \\
&= \sum_{i=1}^{r}\left(zI_n - (\lambda_iI_n + N_i)\right)^{-1}\Pi_i \\
&= \sum_{i=1}^{r}\sum_{j=0}^{n_i-1}(z-\lambda_i)^{-(j+1)}N_i^j\Pi_i
\end{aligned}$$

on obtient alors par formule de Cauchy

$$f(A) = \sum_{i=1}^{r}\sum_{j=0}^{n_i-1}\left(\frac{1}{2i\pi}\int_\gamma f(z)(z-\lambda_i)^{-(j+1)}dz\right)N_i^k\Pi_i = \sum_{i=1}^{r}\sum_{j=0}^{n_i-1}\frac{f^{(j)}(\lambda_i)}{j!}N_i^j\Pi_i$$

$\square$

On peut remarquer que, de la même façon que le théorème de calcul fonctionnel permet de passer des polynômes aux fonctions holomorphes, le résultat précédent permet d'étendre la définition de $f(A)$ en ne demandant qu'une quantité restreinte de régularité à la fonction f.

Par lecture directe du théorème précédent, on retrouve des résultats bien connus. Par exemple si A est à valeurs propres distinctes, alors (avec les notations précédentes) $r = n$ et dans ce cas chaque N_i est nilpotent d'indice 1, c'est-à-dire nul, A s'écrit donc comme somme de projecteurs supplémentaires, donc diagonalisable. De même, si A est annulé par un polynôme scindé à racines simples, alors $\mathrm{rg}(\Pi_i) = 1$ pour tout i et donc A est encore diagonalisable.
On peut également retrouver le théorème de Cayley-Hamilton en notant que le polynôme minimal de A est en fait $\prod_{i=1}^{r}(X-\lambda_i)^{n_i}$.
Enfin, on peut également remarquer que toute fonction de A (définie comme ci-dessus) peut en fait être réalisée comme un polynôme en A de degré au plus $\sum_i n_i - 1$. En effet, il faut et il suffit de déterminer par interpolation un polynôme P vérifiant :

$$\forall i \in [\![1, r]\!]\,,\ \forall j \in [\![0, n_i - 1]\!]\,,\ P^{(j)}(\lambda_i) = f^{(j)}(\lambda_i)$$

ce qui impose $\sum_i n_i$ conditions et donc nécessite un polynôme de degré au plus $\sum_i n_i - 1$.

Pour conclure cette section, on donne une nouvelle démonstration plus concise de la surjectivité de l'exponentielle.

Théorème 38 (Surjectivité de l'exponentielle). *Soit $A \in GL_n(\mathbb{C})$, il existe $B \in \mathcal{M}_n(\mathbb{C})$ telle que $\exp(B) = A$.*

Démonstration : Soit log une détermination du logarithme qui soit holomorphe au voisinage de $\sigma(A)$ et posons $B = \log(A)$. Soit γ l'union de cercles (positivement orientés) centrés en chacune des valeurs propres de A, suffisamment petits pour qu'ils soient contenus dans le domaine de définition de log. On définit η de la même façon avec des cercles contenus dans ceux définissant γ. Avec ces précautions, on peut remarquer que $\log(\gamma)$ entoure le spectre de B de façon cohérente avec tout ce qui précède. Dans ce cas on peut écrire :

$$\begin{aligned}
\exp(B) &= \frac{1}{2i\pi}\int_{\log(\gamma)} e^z (zI_n - B)^{-1} dz \\
&= \left(\frac{1}{2i\pi}\right)^2 \int_{\log(\gamma)}\int_{\eta} \frac{e^z}{z - \log(\zeta)} (\zeta I_n - A)^{-1} d\zeta dz \\
&= \left(\frac{1}{2i\pi}\right)^2 \int_{\gamma}\int_{\eta} \frac{1}{\log(s) - \log(\zeta)} (\zeta I_n - A)^{-1} d\zeta\, ds \\
&= \frac{1}{2i\pi}\int_{\eta} \left(\frac{1}{2i\pi}\int_{\gamma} \frac{s-\zeta}{\log(s) - \log(\zeta)} \frac{1}{s-\zeta} ds\right) (\zeta I_n - A)^{-1} d\zeta
\end{aligned}$$

Pour calculer l'intégrale sur γ, il suffit de remarquer que $s \mapsto \frac{s-\zeta}{\log(s)-\log(\zeta)}$ est holomorphe au voisinage de ζ et admet pour limite $\frac{1}{\log'(\zeta)} = \zeta$. On a donc, encore par formule de Cauchy,

$$\exp(B) = \frac{1}{2i\pi}\int_{\eta} \zeta(\zeta I_n - A)^{-1} d\zeta = A$$

$\square$

Références

[1] G. Berhuy, *Algèbre : le grand combat.* Calvage & Mounet (2020).

[2] D. Borthwick, *Spectral Theory.* Springer (2020).

[3] X. Gourdon, *Les maths en tête : Algèbre.* Ellipses (1994).

[4] P. Lévy-Bruhl *Introduction à la théorie spectrale.* Dunod (2003).

[5] R. Mansuy, R. Mneimé *Algèbre linéaire : Réduction des endomorphismes.* Vuibert (2012).

[6] A. Mouzard, *Algèbres de dimension finie.* Notes de cours.

[7] L.M. Rodrigues, *Complément sur les valeurs propres.* Notes de cours.

[8] M. Romagny, *Algèbres de dimension finie, décomposition de Jordan-Chevalley, exponentielle.* Notes de cours.