DM n° 17 : Algèbre linéaire, matrices

Correction du problème 1 – Réduction de Jordan

Partie I – Réduction du problème

1. On a $u \circ (u - \lambda_i \mathrm{id})^{\alpha_i} = (u - \lambda_i)^{\alpha_i} \circ u$, donc, pour tout $x \in \mathrm{Ker}((u - \lambda_i \mathrm{id})^{\alpha_i})$,
$$(u - \lambda_i \mathrm{id})^{\alpha_i}(u(x)) = u\,(u - \lambda_i \mathrm{id})^{\alpha_i}(x)) = u(0) = 0.$$
Ainsi, $u(x) \in \mathrm{Ker}(u - \lambda_i \mathrm{id})^{\alpha_i}$. On en déduit que $\boxed{\mathrm{Ker}((u - \lambda_i \mathrm{id})^{\alpha_i}) \text{ est stable par } u}$.
Plus généralement, le même raisonnement prouverait que pour tous polynômes P et Q, $\mathrm{Ker}(P(u))$ est stable par $Q(u)$. En particulier, $\boxed{\mathrm{Ker}((u - \lambda_i \mathrm{id})^{\alpha_i}) \text{ est aussi stable par } u - \lambda_i \mathrm{id}.}$

2. Soit A la matrice de u dans la base $\mathcal{B}$, et $A = (A_{i,j})_{1 \leqslant i,j \leqslant k}$ le découpage de A par blocs correspondant à la décomposition de E en somme des E_i. En notant p_i la projection sur E_i parallèlement à $\bigoplus\limits_{j \neq i} E_j$, le bloc $A_{i,j}$ représente alors
$$A_{i,j} = \mathrm{Mat}_{\mathcal{B}_j, \mathcal{B}_i} p_i \circ u_{|E_j}.$$
Or, si $i \neq j$, par stabilité des E_j par u, $\mathrm{Im}(u_{|E_j}) \subset E_j$, donc $p_i \circ u_{|E_j} = 0$. Ainsi, tous les blocs $A_{i,j}$ situés ailleurs que sur la diagonale sont nuls. Les blocs diagonaux, quant à eux, représentent $p_i \circ u_{|E_i}$ (c'est-à-dire u_i) relativement à la base $\mathcal{B}_i$; on a la représentation par blocs :
$$\boxed{\mathrm{Mat}_{\mathcal{B}}(u) = \begin{pmatrix} \boxed{\mathrm{Mat}_{\mathcal{B}_1}(u_1)} & 0 & \cdots & 0 \\ 0 & \boxed{\mathrm{Mat}_{\mathcal{B}_2}(u_2)} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \boxed{\mathrm{Mat}_{\mathcal{B}_k}(u_k)} \end{pmatrix}}$$

3. Puisque $E_i = \mathrm{Ker}((u - \lambda_i \mathrm{id})^{\alpha_i})$, pour tout $x \in E_i$, $v_i^{\alpha_i}(x) = 0$. Ainsi, l'endomophisme $v_i^{\alpha_i}$ de E_i est nul. On en déduit que $\boxed{v_i \text{ est nilpotent.}}$

4. Si tout endomorphisme nilpotent d'un $\mathbb{C}$-ev de dimension finie admet une décomposition de Jordan, alors les v_i admettent une décomposition de Jordan. Ainsi, il existe une base $\mathcal{B}_i$ relativement à laquelle $\mathrm{Mat}_{\mathcal{B}_i}(v_i)$ est diagonale par blocs, chaque bloc diagonal étant une matrice de Jordan de diagonale nulle $\begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & 1 \\ 0 & \cdots & 0 & 0 \end{pmatrix}$.
Ainsi, dans cette base, $\mathrm{Mat}_{\mathcal{B}_i}(u_i)$ est également diagonale par blocs, chaque bloc étant une matrice de Jordan de diagonale λ_i : $\begin{pmatrix} \lambda_i & 1 & 0 & 0 \\ 0 & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & 1 \\ 0 & \cdots & 0 & \lambda_i \end{pmatrix}$.

La question 2 nous assure alors que dans la base obtenue par concaténation des $\mathcal{B}_i$, la matrice de u est diagonale par blocs, chaque bloc étant une matrice de Jordan ; il peut y avoir plusieurs blocs de même diagonale λ_i (et de taille éventuellement différentes), correspondant aux différents blocs de la décomposition de u_i.

Ainsi, si le résultat est acquis pour les endomorphismes nilpotents,

$\boxed{\text{tout endomorphisme d'un } \mathbb{C}\text{-ev de dimension finie admet une réduction de Jordan}}$.

Partie II – Réduction de Jordan d'un endomorphisme nilpotent

1. Argument classique. Soit $\lambda_0, \dots, \lambda_{p-1}$ tels que

$$\lambda_0 x + \cdots + \lambda_{p-1} u^{p-1}(x) = 0.$$

Supposons les λ_i non tous nuls. Soit i le plus petit indice tel que $\lambda_i \neq 0$. En composant par u^{p-1-i}, il vient alors :

$$\lambda_i u^{p-1} + \lambda_{i+1} u^p + \cdots + \lambda_{p-1} u^{2p-2-i}(x) = 0$$

Comme $u^p = 0$ et $u^{p-1}(x) \neq 0$, il vient

$$\lambda_i u^{p-1}(x) = 0 \qquad \text{puis:} \qquad \lambda_i = 0,$$

d'où une contradiction. Ainsi, les λ_i sont tous nuls, ce qui signifie que :

$\boxed{\text{la famille } (u^{p-1}(x), u^{p-2}(x), \dots, u(x), x) \text{ est libre.}}$

On note F le sous-espace engendré par cette famille.

2. Il suffit de montrer que l'image par u des éléments de la base $(u^{p-1}(x), \dots, u(x), x)$ de F est dans F. Or, pour tout $i \in [\![0, p-1]\!]$, $u(u^i(x)) = u^{i+1}(x) \in F$, puisqu'il s'agit d'un autre élément de la base si $i \neq p-1$, et de $0 \in F$ si $i = p-1$. Ainsi, $\boxed{F \text{ est stable par } u.}$

3. Soit $k \in [\![1, p]\!]$.
   - La somme est directe car $u^{p-k}(x) \notin \mathrm{Ker}(u^{k-1})$, puisque $u^{p-1}(x) \neq 0$.
   - Clairement et classiquement, $\mathrm{Ker}(u^{k-1}) \subset \mathrm{Ker}(u^k)$, et $u^k(u^{p-k}(x)) = u^p(x) = 0$, donc $\mathrm{Vect}(u^{p-k}(x)) \subset \mathrm{Ker}(u^k)$.

   Ainsi : $\boxed{\mathrm{Ker}(u^{k-1}) \oplus \mathrm{Vect}(u^{p-k}(x)) \subset \mathrm{Ker}(u^k)}$

4. Soit S_p un supplémentaire de $\mathrm{Ker}(u^{p-1}) \oplus \mathrm{Vect}(x)$ dans $\mathrm{Ker}(u^p)$. On a alors :

$$\mathrm{Ker}(u^{p-2}) \oplus \mathrm{Vect}(u(x)) \oplus u(S_p) \subset \mathrm{Ker}(u^{p-1}).$$

En effet :
   - La première somme est directe d'après la question précédente.
   - Soit $y \in (\mathrm{Ker}(u^{p-2}) \oplus \mathrm{Vect}(u(x))) \cap u(S_p)$. Il existe donc $x_1 \in \mathrm{Ker}(u^{p-2})$, $\lambda \in \mathbb{C}$, et $x_2 \in S_p$ tels que

$$y = x_1 + \lambda u(x) \qquad \text{et} \qquad y = u(x_2), \qquad \text{donc:} \qquad u(x_2) = x_1 + \lambda u(x).$$

   En appliquant u^{p-2}, il vient donc :

$$u^{p-1}(x_2) = \lambda u^{p-1}(x).$$

   Or, u^{p-1} est injective sur $S_p \oplus \mathrm{Vect}(x)$, supplémentaire de $\mathrm{Ker}(u^{p-1})$ dans $\mathrm{Ker}(u^p) = E$, et x_2 et x sont dans $S_p \oplus \mathrm{Vect}(x)$. Ainsi, $x_2 = \lambda x$. Comme $S_p \cap \mathrm{Vect}(x) = \{0\}$, on en déduit que $x_2 = 0$, puis $y = 0$.
   Ainsi, la somme $\mathrm{Ker}(u^{p-2}) \oplus \mathrm{Vect}(u(x)) \oplus u(S_p)$ est directe.
   - On a $\mathrm{Ker}(u^{p-2}) \oplus \mathrm{Vect}(u(x)) \subset \mathrm{Ker}(u^{p-1})$ d'après la question précédente, et $u(S_p) \subset \mathrm{Ker}(u^{p-1})$ (car $u^p = 0$). Donc

$$\mathrm{Ker}(u^{p-2}) \oplus \mathrm{Vect}(u(x)) \oplus u(S_p) \subset \mathrm{Ker}(u^{p-1}).$$

   Soit T_{p-1} un supplémentaire de $\mathrm{Ker}(u^{p-2}) \oplus \mathrm{Vect}(u(x)) \oplus u(S_p)$ dans $\mathrm{Ker}(u^{p-1})$. On a donc

$$\mathrm{Ker}(u^{p-2}) \oplus \mathrm{Vect}(u(x)) \oplus u(S_p) \oplus T_{p-1} = \mathrm{Ker}(u^{p-1}),$$

   donc, $\boxed{S_{p-1} = T_{p-1} \oplus u(S_p) \text{ est un supplémentaire de } \mathrm{Ker}(u^{p-2}) \oplus \mathrm{Vect}(u(x)) \text{ dans } \mathrm{Ker}(u^{p-1}) \text{ contenant } u(S_p).}$

5. Le raisonnement est le même. Supposons $S_p, \dots, S_{k+1}$ construits.
   - On a $\mathrm{Ker}(u^{k-1}) \oplus \mathrm{Vect}(u^{p-k}) \subset \mathrm{Ker}(u^k)$ d'après la question 3.
   - Soit $y \in (\mathrm{Ker}(u^{k-1}) \oplus \mathrm{Vect}(u^{p-k})) \cap u(S_{k+1})$. Il existe $x_1 \in \mathrm{Ker}(u^{k-1})$, $\lambda \in \mathbb{C}$ et $x_2 \in S_{k+1}$ tels que

     $$y = u(x_2) = x_1 + \lambda u^{p-k}(x).$$

     En appliquant u^{k-1} à cette égalité, il vient

     $$u^k(x_2) = u^k(\lambda u^{p-k-1}(x)).$$

     Or, u^k est injective sur $S_{k+1} \oplus \mathrm{Vect}(u^{p-k-1})$ (car cet espace est en somme directe avec $\mathrm{Ker}(u^{k+1})$, d'après la construction de S_{k+1}). Ainsi, $x_2 = \lambda u^{p-k-1}(x)$. Comme $S_{k+1} \cap \mathrm{Vect}(u^{p-k-1}(x)) = \{0\}$, il vient $x_2 = 0$, puis $y = 0$.
     Ainsi, la somme $\mathrm{Ker}(u^{k-1}) \oplus \mathrm{Vect}(u^{p-k}) \oplus u(S_{k+1})$ est directe.
   - Comme $S_{k+1} \subset \mathrm{Ker}(u^{k+1})$, on a $u(S_{k+1}) \subset \mathrm{Ker}(u^k)$. Les autres inclusions ayant déjà été montrées, il vient donc :

     $$\mathrm{Ker}(u^{k-1}) \oplus \mathrm{Vect}(u^{p-k}) \oplus u(S_{k+1}) \subset \mathrm{Ker}(u^k).$$

     Soit T_k un supplémentaire de $\mathrm{Ker}(u^{k-1}) \oplus \mathrm{Vect}(u^{p-k}) \oplus u(S_{k+1})$ dans $\mathrm{Ker}(u^k)$ et $S_k = T_k \oplus u(S_{k+1})$.
     Alors $\boxed{S_k \text{ est un supplémentaire de } \mathrm{Ker}(u^{k-1}) \oplus \mathrm{Vect}(u^{p-k}) \text{ dans } \mathrm{Ker}(u^k)\text{, contenant } u(S_{k+1}).}$

6. Soit $T = S_1 + \dots + S_p$.
   - Par définition de S_1, on a $\mathrm{Vect}(u^{p-1}(x)) \oplus S_1 = \mathrm{Ker}(u)$.
   - On a ensuite $\mathrm{Ker}(u) \oplus \mathrm{Vect}(u^{p-2}(x)) \oplus S_2 = \mathrm{Ker}(u^2)$, donc d'aprè le point précéent :

     $$\mathrm{Vect}(u^{p-1}(x), u^{p-2}(x)) \oplus S_1 \oplus S_2 = \mathrm{Ker}(u^2).$$

   - Plus généralement, pour tout $k \in [1, p]$,

     $$\mathrm{Vect}(u^{p-1}(x), \dots, u^{p-k}(x)) \oplus S_1 \oplus \dots \oplus S_k = \mathrm{Ker}(u^k).$$

     En effet, si le résultat est aquis pour $k - 1 \geqslant 1$, il vient :

     $$\begin{aligned} &\mathrm{Vect}(u^{p-1}(x), \dots, u^{p-k}(x)) \oplus S_1 \oplus \dots \oplus S_k \\ &\qquad = \mathrm{Vect}(u^{p-1}(x), \dots, u^{p-k+1}(x)) \oplus S_1 \oplus \dots \oplus S_{k-1} \oplus \mathrm{Vect}(u^{p-k}(x)) \oplus S_k \\ &\qquad = \mathrm{Ker}(u^{k-1}) \oplus \mathrm{Vect}(u^{p-k}(x)) \oplus S_k = \mathrm{Ker}(u^k). \end{aligned}$$

   - En particulier, pour $k = p$, on trouve que $\boxed{T \text{ est un supplémentaire de } F \text{ dans } E = \mathrm{Ker}(u^p).}$
   - Par ailleurs,

     $$u(T) \subset u(S_1) + \dots + u(S_p) \subset \{0\} + S_1 + \dots + S_{p-1} \subset T.$$

     Ainsi, $\boxed{T \text{ est stable par } u.}$

7. On montre alors le résultat par récurrence forte sur la dimension n de l'espace E.
   Lorsque $\dim E = 0$, il n'y a rien à montrer.
   Soit $n \in \mathbb{N}^*$. Supposons que tout endomorphisme nilpotent d'un $\mathbb{C}$-ev de dimension $k < n$ admet une décomposition de Jordan. Soit u un endomorphisme nilpotent d'un $\mathbb{C}$-ev E de dimension n, d'indice de nilpotentce p. Soit $x \notin \mathrm{Ker}(u^{p-1})$, et F et T construits comme ci-dessus. Alors F et T sont des sous-espaces stables par u. Soient u_1 et u_2 les endomorphismes de F et T induits par u.
   - Comme $\dim(F) > 0$, on a $\dim(T) < n$, et on peut appliquer l'hypothèse de récurrence à T : il existe une base $\mathcal{B}_2$ de T relativement à laquelle $\mathrm{Mat}_{\mathcal{B}_2}(u_2)$ est diagonale par bloc, chaque bloc étant une matrice de Jordan de diagonale nulle.
   - On ne peut pas appliquer l'hypothèse de récurrence à F (car rien n'assure que $\dim(F) < n$: T peut être nul). En revanche, il est facile de trouver une base relativement à laquelle la matrice de u_1 est une matrice de Jordan : $\mathcal{B}_1 = (u^{k-p}(x), \dots, u(x), x)$.

- Soit $\mathcal{B}$ la juxtaposition des bases $\mathcal{B}_1$ et $\mathcal{B}_2$. Alors la matrice de u relativement à la base $\mathcal{B}$ est diagonale par blocs :
$$\mathrm{Mat}_{\mathcal{B}}(u) = \left(\begin{array}{c|c} \mathrm{Mat}_{\mathcal{B}_1}(u_1) & 0 \\ \hline 0 & \mathrm{Mat}_{\mathcal{B}_2}(u_2) \end{array}\right).$$
D'après les deux points précédents, cette matrice est constituée de blocs diagonaux égaux à des matrices de Jordan, ce qui achève la preuve.

Ainsi, $\boxed{\text{toute matrice nilpotente admet une réduction de Jordan.}}$

Correction du problème 2 – Exponentielles et logarithmes de matrices

Partie I – Premier exemple

1. (a) $A^0 = I_4$, $A^1 = A$, et :
$$\boxed{A^2 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}} \qquad \boxed{A^3 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}} \qquad \boxed{A^4 = 0}, \qquad \text{puis, } \boxed{\text{pour tout } n \geqslant 4,\ A^n = 0}.$$

(b) Ainsi, la somme $\sum \frac{A^n}{n!}$ est une somme finie, donc forcément convergente. Par conséquent, $\exp(A)$ existe, et :
$$\exp(A) = \sum_{n=0}^{3} \frac{A^n}{n!} = I_4 + A + \frac{1}{2}A^2 + \frac{1}{6}A^3 \qquad \text{soit:} \qquad \boxed{\exp(A) = \begin{pmatrix} 1 & 1 & \frac{1}{2} & \frac{1}{6} \\ 0 & 1 & 1 & \frac{1}{2} \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}}$$

(c) On peut bien sûr inverser la matrice obtenue par la méthode du pivot. On peut aussi tenter l'analogie avec l'exponentielle réelle ou complexe, l'inverse de $\exp(x)$ étant $\exp(-x)$. Or le calcul de $\exp(-A)$ et totalement similaire à celui de $\exp(A)$, hormis un signe qui alterne. On trouve :
$$\boxed{\exp(-A) = \begin{pmatrix} 1 & -1 & \frac{1}{2} & -\frac{1}{6} \\ 0 & 1 & -1 & \frac{1}{2} \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 1 \end{pmatrix}}$$
On obtient sans difficulté $\exp(A)\exp(-A) = I_4$, donc $\boxed{\text{l'inverse de } \exp(A) \text{ est bien } \exp(-A)}$.

Ce n'est pas étonnant puisque une telle propriété provient pour les réels d'une étude complètement formelle des séries. La convergence des séries suffit à justifier une telle propriété. Ici, la convergence ne pose pas de problème, puisque les sommes sont finies.

(d) Soit $C = \exp(A) - I_4$. Alors :
$$\boxed{C = \begin{pmatrix} 0 & 1 & \frac{1}{2} & \frac{1}{6} \\ 0 & 0 & 1 & \frac{1}{2} \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}}, \qquad \boxed{C^2 = \begin{pmatrix} 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}}, \qquad \boxed{C^3 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}}, \qquad \boxed{\forall n \geqslant 4,\ C^n = 0}.$$

(e) La série $\sum \frac{(-1)^{n+1}C^n}{n}$ est finie, donc convergente, ce qui équivaut à l'existence de $\ln(\exp(A))$. Alors :
$$\ln(\exp(A)) = \sum_{n=1}^{3} \frac{(-1)^{n+1}C^n}{n} = \begin{pmatrix} 0 & 1 & \frac{1}{2} & \frac{1}{6} \\ 0 & 0 & 1 & \frac{1}{2} \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} - \frac{1}{2}\begin{pmatrix} 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} + \frac{1}{3}\begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$
et on obtient sans peine et sans surprise $\boxed{\ln(\exp(A)) = A}$.

(f) On a $I_4 - A = \begin{pmatrix} 1 & -1 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 1 \end{pmatrix}$. Cette matrice est inversible, car triangulaire à coefficients diagonaux non nuls. La moitié dupivot est déjà faite : il reste à faire le pivot remontant, en effectuant les mêmes opérations sur la matrice identité :

$$\left(\begin{array}{cccc|cccc} 1 & -1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \end{array}\right) \underset{L_3 \leftarrow L_3 + L_4}{\longrightarrow} \left(\begin{array}{cccc|cccc} 1 & -1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \end{array}\right)$$

$$\underset{L_2 \leftarrow L_2 + L_3}{\longrightarrow} \left(\begin{array}{cccc|cccc} 1 & -1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \end{array}\right)$$

$$\underset{L_1 \leftarrow L_1 + L_2}{\longrightarrow} \left(\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \end{array}\right)$$

Ainsi, $\boxed{(I_4 - A)^{-1} = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}}$

(g) Comme pour les sommes précédentes, la somme $\sum A^n$ est finie donc convergente, et on trouve sans difficulté

$$\boxed{\sum_{n=0}^{+\infty} A^n = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix} = (I_n - A)^{-1}.}$$

Faites l'analogie avec la formule de sommation des séries géométriques.

2. • Notons J la matrice obtenue lorsque $a = 1$. On a alors $A = aJ$. Par ailleurs, la description du produit matricielle par colonnes permet d'obtenir la description suivante du produit d'une matrice décrite par ses colonnes par J :

$$\left(\begin{array}{c|c|c} C_1 & \cdots & C_p \end{array}\right) J = \left(\begin{array}{c|c|c|c} 0 & C_1 & \cdots & C_{p-1} \end{array}\right)$$

Ainsi, $J^n = J^n I_p$ est la matrice obtenue en décalant les colonnes de la matrice I_p de n crans vers la droite (les n dernières disparaissent), les premières colonnes étant remplacées par des colonnes nulles. Ainsi

$$\forall n \in [1, p-1], \quad J^n = \begin{pmatrix} 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ \vdots & \ddots & & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & & \ddots & \ddots & 0 \\ \vdots & & & \ddots & & \ddots & 1 \\ \vdots & & & & \ddots & & 0 \\ \vdots & & & & & \ddots & \vdots \\ 0 & \cdots & \cdots & \cdots & \cdots & \cdots & 0 \end{pmatrix}$$

Ainsi, pour $n \in [0, p-1]$, $A^n = a^n J^n$ est la matrice contituée d'une diagonale de a^n partant de la position $(1, n+1)$, et de 0 partout ailleurs. Pour tout $n \geqslant p$, on a $A^n = 0$. Ce dernier point nous assure la convergence de toutes les sommes que nous manipulerons, et en particulier, l'existence des exponentielles et logarithmes.

- On obtient donc :

$$\boxed{\exp(A)=\begin{pmatrix}1 & a & \cdots & \frac{a^n}{n!} & \cdots & \frac{a^p}{p!}\\ 0 & \ddots & \ddots & & \ddots & \vdots\\ \vdots & \ddots & \ddots & \ddots & & \frac{a^n}{n!}\\ \vdots & & \ddots & \ddots & \ddots & \vdots\\ \vdots & & & \ddots & \ddots & a\\ \vdots & \cdots & \cdots & \cdots & 0 & 1\end{pmatrix}}\ \text{et de même}\ \exp(-A)=\begin{pmatrix}1 & -a & \cdots & \frac{(-a)^n}{n!} & \cdots & \frac{(-a)^p}{p!}\\ 0 & \ddots & \ddots & & \ddots & \vdots\\ \vdots & \ddots & \ddots & \ddots & & \frac{(-a)^n}{n!}\\ \vdots & & \ddots & \ddots & \ddots & \vdots\\ \vdots & & & \ddots & \ddots & -a\\ \vdots & \cdots & \cdots & \cdots & 0 & 1\end{pmatrix}$$

- La vérification du fait que l'inverse de $\exp(A)$ est $\exp(-A)$ est plus commode sur la définition par une somme, toutes les manipulations étant justifiées par le fait que seul un nombre fini de termes est non nul :

$$\exp(A)\exp(-A)=\sum_{k=0}^{+\infty}\frac{A^k}{k!}\cdot\sum_{k=0}^{+\infty}\frac{(-A)^k}{k!}=\sum_{\ell=0}^{+\infty}c_\ell A^\ell,$$

  où pour tout $\ell\in\mathbb{N}$,

$$c_\ell=\sum_{\substack{i+j=\ell\\ i\in[0,n-1]\\ j\in[1,n-1]}}\frac{1}{i!}\cdot\frac{(-1)^j}{j!}=\frac{1}{\ell!}\sum_{\substack{i+j=\ell\\ i\in[0,n-1]\\ j\in[1,n-1]}}\binom{\ell}{j}\cdot(-1)^j=\frac{1}{\ell!}(1-1)^\ell,$$

  d'après la formule du binôme. On en déduit que $c_0=1$ et pour tout $\ell\in\mathbb{N}^*$, $c_\ell=0$.
  Ainsi, $\boxed{\exp(A)\exp(-A)=I_p.}$ La matrice $\exp(-A)$ est donc bien inversible, d'inverse $\exp(-A)$.
- Le terme $\exp(A)$ s'écrivant $AP(A)$, où P est un polynômes, $(\exp(A)^p=A^pP(A)^p=0$, Ainsi, le logarithme est bien défini. Vérifier que $\ln(\exp(A))$ n'est pas chose facile, si on s'y prend directement. On peut en revanche remarquer que, étant donné un entier $d\geqslant 1$ le développement de la somme $\displaystyle\sum_{k=1}^{d}\frac{(-1)^{k+1}}{k}\left(\sum_{i=1}^{d}\frac{A^i}{i!}\right)^k$ peut se faire formellement sur le polynôme $\displaystyle\sum_{k=1}^{d}\frac{(-1)^{k+1}}{k}\left(\sum_{i=1}^{d}\frac{X^i}{i!}\right)^k$. La fonction polynomiale associée à la troncature au degré d de ce polynôme fournit alors le développement limité de $\ln(\exp(x))$ à l'ordre d au voisinage de 0, à savoir $x+o(x^d)$. Par unicité du développement limité, et par identification des polynômes formels et des fonctions polynomiales sur $\mathbb{R}$, on obtient donc :

$$\forall d\geqslant 1, T_d\left(\sum_{k=1}^{d}\frac{(-1)^{k+1}}{k}\left(\sum_{i=1}^{d}\frac{X^i}{i!}\right)^k\right)=X.$$

  En prenant $d=2p$, et en évaluant en A, les puissances de A d'ordre supérieur à p étant nulles, on obtient :

$$\sum_{k=1}^{p-1}\frac{(-1)^{k+1}}{k}\left(\sum_{i=1}^{p-1}\frac{A^i}{i!}\right)^k=A\qquad\text{soit:}\qquad\boxed{\ln(\exp(A))=A}.$$

- Même si un pivot n'est pas commode à réalisé sur une matrice de taille générique $n\times n$, la situation est suffisamment simple ici sur la matrice I_p-A pour s'en sortir : on se rend compte assez facilement qu'il va falloir d'abord retrancher aL_p à L_{p-1}, puis aL_{p-1} à L_{p-2} et ainsi de suite. On se convainc assez facilement qu'on obtient :

$$(I_p-A)^{-1}=\begin{pmatrix}1 & a & a^2 & \cdots & a^p\\ 0 & \ddots & \ddots & \ddots & \vdots\\ \vdots & \ddots & \ddots & \ddots & a^2\\ \vdots & & \ddots & \ddots & a\\ 0 & \cdots & \cdots & 0 & 1\end{pmatrix}$$

- Il est ici beaucoup plus commode de raisonner sur les sommes pour trouver ce résultat (en le devinant par analogie au cas réel) :

$$(I_p - A)\cdot\sum_{k=0}^{+\infty} A^k = (I_p - A)\cdot\sum_{k=0}^{p-1} A^k = \sum_{k=0}^{p-1} A^k - A^{k-1}.$$

Un petit télescopage donne alors :

$$(I_p - A)\cdot\sum_{k=0}^{+\infty} A^k = I_p - A^p = I_p.$$

Ainsi, l'inverse de $I_p - A$ est bien :

$$\boxed{(I_p - A)^{-1} = \sum_{k=0}^{+\infty} A^k = \sum_{k=0}^{+\infty} a^k J^k = \begin{pmatrix} 1 & a & a^2 & \cdots & a^p \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & a^2 \\ \vdots & & \ddots & \ddots & a \\ 0 & \cdots & \cdots & 0 & 1 \end{pmatrix}}$$

Partie II – Deuxième exemple

On pose $A = \begin{pmatrix} -\frac{5}{3} & -\frac{2}{3} \\ -\frac{1}{3} & -\frac{4}{3} \end{pmatrix}$.

1. (a)
   - On recherche des vecteurs b_1 et b_2 tels qu'il existe λ_1 et λ_2 tels que $f(b_1) = \lambda_1 b_1$ et $f(b_2) = \lambda_2 b_2$, donc tels que $b_1 \in \mathrm{Ker}(f - \lambda_1 \mathrm{id})$ et $b_2 \in \mathrm{Ker}(f - \lambda_2 \mathrm{id})$. Réciproquement, si b_1 et b_2 vérifient une telle condition, et ne sont pas liés (en particulier non nuls), ils forment une base telle que requise. Il s'agit donc de trouver deux telles valeurs de λ_1 et λ_2 telles que les noyaux $\mathrm{Ker}(f - \lambda_1 \mathrm{id})$ et $\mathrm{Ker}(f - \lambda_2 \mathrm{id})$ ne soient pas réduits à 0, donc tels que les endomorphismes $f - \lambda_1 \mathrm{id}$ et $f - \lambda_2 \mathrm{id}$ ne soient pas injectifs, donc ne soient pas des automorphismes, donc tels que leur matrice associée dans la base canonique ne soit pas inversible.
   - On utilise la caractérisation par les déterminants :
     $A - \lambda I_2$ est non inversible si et seulement si $\left|\begin{pmatrix} -\frac{5}{3} - \lambda & -\frac{2}{3} \\ -\frac{1}{3} & -\frac{4}{3} - \lambda \end{pmatrix}\right| = 0$, soit :

     $$\left(-\frac{5}{3} - \lambda\right)\left(-\frac{4}{3} - \lambda\right) - \frac{2}{9} = 0 \qquad \text{soit:} \qquad \lambda^2 + 3\lambda + 2.$$

     Seules 2 valeurs de λ conviennent : $\lambda = -1$ et $\lambda = -2$ (ce sont les valeurs propres de A).
   - On choisit donc b_1 non nul dans $\mathrm{Ker}(A + I_2) = \mathrm{Ker}\begin{pmatrix} -\frac{2}{3} & -\frac{2}{3} \\ -\frac{1}{3} & -\frac{1}{3} \end{pmatrix}$. Les colonnes de cette matrice vérifiant $C_1 - C_2 = 0$, un vecteur de son noyau est par exemple $\boxed{b_1 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}}$.
   - On choisit b_2 non nul dans $\mathrm{Ker}(A + 2I_2) = \mathrm{Ker}\begin{pmatrix} \frac{1}{3} & -\frac{2}{3} \\ -\frac{1}{3} & \frac{2}{3} \end{pmatrix}$. Les colonnes de cette matrice vérifiant $2C_1 + C_2 = 0$, un vecteur de son noyau est par exemple $\boxed{b_2 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}}$.
   - Par construction, $Ab_1 = b_1$ et $Ab_2 = 2b_2$, et (b_1, b_2) forme une base (car non colinéaires dans un espace de dimension 2).

   (b) On a alors $\boxed{P = \begin{pmatrix} 1 & 2 \\ -1 & 1 \end{pmatrix}}$.

Son déterminant vaut $\det(P) = 3 \neq 0$, donc P est inversible, et

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

On vérifie sans trop de peine que $\boxed{P^{-1}AP = \begin{pmatrix} -1 & 0 \\ 0 & -2 \end{pmatrix} = D}$.

On vient d'effectuer la diagonalisation de la matrice A. Remarquez que la matrice diagonale obtenue est constituée des valeurs propres de A.

Remarquez aussi que la matrice D n'est pas unique (on aurait pu échanger l'ordre des deux valeurs propres, mais c'est tout), en encore moins P (non seulement on peut échanger l'ordre des valeurs propres, mais pour chaque valeur propre, le choix des b_i associés n'est pas unique). Cependant, la relation $D = P^{-1}AP$ ou $A = PDP^{-1}$ est toujours la même.

(c) • Pour tout $n \in \mathbb{N}$, on a $D^n = \begin{pmatrix} (-1)^n & 0 \\ 0 & (-2)^n \end{pmatrix}$, donc

$$\forall n \in \mathbb{N}, \quad \sum_{k=0}^{n} \frac{D^k}{k!} = \begin{pmatrix} \sum\limits_{k=0}^{n} \frac{(-1)^k}{k!} & 0 \\ 0 & \sum\limits_{k=0}^{n} \frac{(-2)^k}{k!} \end{pmatrix}.$$

On a convergence de chacun de ces coefficients, par conséquent, $\exp(D)$ est bien défini, et

$$\boxed{\exp(D) = \begin{pmatrix} e^{-1} & 0 \\ 0 & e^{-2} \end{pmatrix}}.$$

• On a, pour tout $n \in \mathbb{N}$, $A^n = PDP^{-1}PDP^{-1} \cdots PDP^{-1} = PD^nP^{-1}$. Ainsi, pour tout $n \in \mathbb{N}$,

$$\sum_{k=0}^{n} \frac{A^k}{k!} = P\left(\sum_{k=0}^{n} D^n\right)P^{-1}.$$

Or, multiplier par une matrice revient à faire une combinaison linéaire (finie) des coefficients. Comme on a convergence coefficient par coefficient de la somme $\sum D^n$, il en est donc de même, d'après les règles sur les limites, de $P \cdot \sum D^n$, puis de $P\left(\sum D^n\right)P^{-1}$. Par conséquent, la somme définissant $\exp(A)$ converge, et les limites de réels étant compatibles avec les sommes et les produits (ce sont les seules opérations intervenant dans la description du produit matriciel), il vient :

$$\boxed{\exp(A) = P\exp(D)P^{-1}}.$$

(d) Un calcul, qu'il faudrait préciser un peu, amène alors $\boxed{\exp(A) = \frac{1}{3}\begin{pmatrix} 2e^{-2} + e^{-1} & 2e^{-2} - 2e^{-1} \\ e^{-2} - e^{-1} & e^{-2} + 2e^{-2} \end{pmatrix}}$.

2. • On obtient de la même manière $\exp(-A) = P\exp(-D)P^{-1} = P\begin{pmatrix} e^2 & 0 \\ 0 & e \end{pmatrix}P^{-1} = \frac{1}{3}\begin{pmatrix} 2e^2 + e & 2e^2 - 2e \\ e^2 - e & e^2 + 2e^2 \end{pmatrix}$.

• Un calcul (sur les résultats obtenus, afin d'éviter les discussions sur les convergences), un peu fastidieux mais sans difficulté, amène alors, de façon très surprenante : $\boxed{\exp(A)\exp(-A) = I_2.}$

Devinez alors qui est l'inverse de $\exp(A)$?

3. On note $B = \exp(A)$. Notre but est de déterminer $\ln(B) = \ln(\exp(A))$.

(a) On a $B = \exp(A) = P\exp(D)P^{-1}$, donc $B - I_2 = P(\exp D - I_2)P^{-1}$. Ainsi, il suffit de poser

$$\boxed{D' = \exp(D) - I_2 = \begin{pmatrix} e^{-2} - 1 & 0 \\ 0 & e^{-1} - 1 \end{pmatrix}}.$$

(b) On a donc, pour tout $n \in \mathbb{N}$, $\displaystyle\sum_{k=1}^{n} \frac{(-1)^{k+1}}{k} D'^k = \begin{pmatrix} \sum_{k=1}^{n} \frac{(-1)^{k+1}}{k}(\mathrm{e}^{-2}-1)^k & 0 \\ 0 & \sum_{k=1}^{n} \frac{(-1)^{k+1}}{k}(\mathrm{e}^{-1}-1)^k \end{pmatrix}$

Or, puisque $|\mathrm{e}^{-2}-1| < 1$ et $|\mathrm{e}^{-1}-1| < 1$, les séries $\displaystyle\sum_{n=1}^{+\infty} \frac{(-1)^{n+1}}{n}(\mathrm{e}^{-2}-1)^n$ et $\displaystyle\sum_{n=1}^{+\infty} \frac{(-1)^{n+1}}{n}(\mathrm{e}^{-1}-1)^n$ convergent, respectivement vers $\ln(\mathrm{e}^{-2}) = -2$ et $\ln(\mathrm{e}^{-1}) = -1$. Par conséquent, $\ln(D'+I_2)$ existe, et

$$\boxed{\ln(D'+I_2) = \begin{pmatrix} -2 & 0 \\ 0 & -1 \end{pmatrix} = D}.$$

On a utilisé ici le développement en série entière du logarithme, démontré à l'aide des formules de Taylor :

$$\forall x \in]-1,1[, \quad \ln(1+x) = \sum_{n=1}^{+\infty} \frac{(-1)^{n+1}}{n} x^n.$$

(c) Comme plus haut, multiplier par une matrice ne change pas la convergence, donc la série définissant $\ln(B)$, obtenue en multipliant la série définissant $\ln(D'+I_2)$ par P à gauche et par P^{-1} à droite converge, d'où l'existence de $\ln(B)$, et comme précédemment,

$$\ln(\exp(A)) = \ln(B) = P\ln(D'+I_2)P^{-1} = PDP^{-1} \qquad \text{donc:} \qquad \ln(\exp(A) = A.$$

Cela n'a plus de quoi nous étonner.

Partie III – Exponentielle générale

On se fixe $n \in \mathbb{N}^*$. On définit, pour tout $A = (a_{i,j}) \in \mathcal{M}_n(\mathbb{C})$:

$$\|A\| = \max_{1 \leqslant i,j \leqslant n} |a_{i,j}|.$$

1. (a)
- De façon évidente, $\|0_n\| = 0$. Réciproquement, si $A^n = 0$, alors $\displaystyle\max_{1 \leqslant i,j \leqslant n} |a_{i,j}| = 0$. Tous ces termes étant positifs, on en déduit que pour tout (i,j), $|a_{i,j}| = 0$, donc $A = 0$.
- Soit $\lambda \in \mathbb{C}$, et $A \in \mathcal{M}_n(\mathbb{C})$.

$$\|\lambda A\| = \max_{1 \leqslant i,j \leqslant n} |\lambda a_{i,j}| = \max_{1 \leqslant i,j \leqslant n} |\lambda| \cdot |a_{i,j}| = |\lambda| \max_{1 \leqslant i,j \leqslant n} \cdot |a_{i,j}| = |\lambda| \|A\|.$$

- Soit $A = (a_{i,j})$ et $B = (b_{i,j})$:

$$\|A+B\| = \max_{1 \leqslant i,j \leqslant n} |a_{i,j} + b_{i,j}| \leqslant \max_{1 \leqslant i,j \leqslant n} (|a_{i,j}| + |b_{i,j}|).$$

Or, pour tout (i_0, j_0),

$$|a_{i_0,j_0}| \leqslant \max_{1 \leqslant i,j \leqslant n} |a_{i,j}| \qquad \text{et} \qquad |b_{i_0,j_0}| \leqslant \max_{1 \leqslant i,j \leqslant n} |b_{i,j}| \qquad \text{donc:} \qquad |a_{i_0,j_0}| + |b_{i_0,j_0}| \leqslant \|A\| + \|B\|.$$

Cette inégalité étant vraie pour tout $(i_0, j_0)| \in [1,n]^2$, il vient :

$$\max_{1 \leqslant i,j \leqslant n} (|a_{i,j}| + |b_{i,j}|) \leqslant \|A\| + \|B\| \qquad \text{puis:} \qquad \|A+B\| \leqslant \|A\| + \|B\|.$$

On en déduit que $\boxed{A \mapsto \|A\| \text{ est bien une norme sur } \mathcal{M}_n(\mathbb{R})}$.

(b) Soit (A_k) une suite de matrice.
- Supposons que (A_k) converge (coefficient par coefficient) vers A. Notons $A = (a_{i,j})$ et pour tout k, $A_k = (a_{i,j,k})$. Soit $\varepsilon > 0$. Pour tout (i,j), il existe $k_{i,j}$ tel que pour tout $k \geqslant k_{i,j}$,

$$|a_{i,j,k} - a_{i,j}| \leqslant \varepsilon.$$

Soit $k_0 = \displaystyle\max_{1 \leqslant i,j \leqslant n} k_{i,j}$. On a alors, pour tout $k \geqslant k_0$,

$$\forall (i,j) \in [1,n]^, \quad |a_{i,j,k} - a_{i,j}| \leqslant \varepsilon \qquad \text{donc:} \qquad \max_{1 \leqslant i,j \leqslant n} |a_{i,j,k} - a_{i,j}| \leqslant \varepsilon.$$

Ainsi, pour tout $k \geqslant k_0$, $\|A_k - A\| \leqslant \varepsilon$. Cela prouve bien la convergence de $\|A_k - A\|$ vers 0

- Réciproquement, supposons que $\|A_k - A\|$ converge vers 0. Alors, pour tout $(i,j) \in [1,n]$, et pour tout $k \in \mathbb{N}$,

$$0 \leqslant |a_{i,j,k} - a_{i,j}| \leqslant \|A_k - A\| \to 0.$$

Ainsi, le théorème d'encadrement nous assure que $a_{i,j,k} \to a_{i,j}$. On a bien convergence coefficient par coefficient vers A.

Conclusion : $\boxed{A_k \text{ converge vers une matrice } A \text{ si et seulement si } \|A_k - A\| \to 0.}$

(c) On reprend le même raisonnement que dans la question précedente, mais en se ramenant au critère de Cauchy de la convergence de chaque coefficient.

2. Soit $A, B \in \mathcal{M}_n(\mathbb{R})$. On pose $AB = (c_{i,k})$. On a :

$$\forall (i,k) \in [1,n]^2, \quad |c_{i,k}| = \left|\sum_{j=1}^{n} a_{i,j} b_{j,k}\right| \leqslant \sum_{j=1}^{n} |a_{i,j}| \cdot |b_{j,k}| \leqslant \sum_{j=1}^{n} \|A\| \cdot \|B\| = n\|A\| \cdot \|B\|.$$

Cette inégalité étant vraie pour tout $(i,k) \in [1,n]^2$, on en déduit :

$$\boxed{\|AB\| \leqslant n\|A\| \cdot \|B\|}$$

.

3. Une récurrence immédiate à partir de l'expression de la question précédente amène, pour tout $k \in \mathbb{N}^*$, $\boxed{\|A^k\| \leqslant n^{k-1}\|A\|^k}$. On a alors, pour tout (k,p) tel que $k < p$

$$\left\|\sum_{\ell=k+1}^{p} \frac{A^\ell}{\ell!}\right\| \leqslant \sum_{\ell=k+1}^{p} \frac{n^{\ell-1}\|A\|^\ell}{\ell!} = \frac{1}{n} \sum_{\ell=k+1}^{p} \frac{(n\|A\|)^\ell}{\ell!}.$$

Soit $\varepsilon > 0$. Le critère de Cauchy appliqué à la convergence de la série exponentielle réelle de paramètre $n\|A\|$ nous donne l'existence de N tel que pour tout (k,p) tels que $N \leqslant k < p$,

$$\frac{1}{n} \sum_{\ell=k+1}^{p} \frac{(n\|A\|)^\ell}{\ell!} \leqslant \varepsilon,$$

et par suite :

$$\left\|\sum_{\ell=k+1}^{p} \frac{A^\ell}{\ell!}\right\| \leqslant \varepsilon.$$

La question 1(c) appliquée à la somme partielle de la série exponentielle de A amène alors la convergence de cette série. Ainsi, $\boxed{\exp(A) \text{ existe}}$.

4. On a, pour tout $N \in \mathbb{N}$, du fait de la commutation $AB = BA$ qui permet d'utiliser la formule du binôme :

$$\begin{aligned}
\left\|\left(\sum_{k=0}^{N} \frac{A^k}{k!}\right)\left(\sum_{\ell=0}^{N} \frac{B^\ell}{\ell!}\right) - \sum_{i=0}^{N} \frac{(A+B)^i}{i!}\right\| &= \left\|\sum_{k=0}^{N}\sum_{\ell=0}^{N} \frac{A^k B^\ell}{k!\ell!} - \sum_{i=0}^{N}\sum_{k=0}^{i} \binom{i}{k} \frac{A^k B^{i-k}}{i!}\right\| \\
&= \left\|\sum_{k=0}^{N}\sum_{\ell=0}^{N} \frac{A^k B^\ell}{k!\ell!} - \sum_{i=0}^{N}\sum_{k=0}^{i} \frac{A^k B^{i-k}}{k!(i-k)!}\right\| \\
&= \left\|\sum_{k=0}^{N}\sum_{\ell=0}^{N} \frac{A^k B^\ell}{k!\ell!} - \sum_{k=0}^{N}\sum_{i=k}^{N} \frac{A^k B^{i-k}}{k!(i-k)!}\right\| \\
&= \left\|\sum_{k=0}^{N}\sum_{\ell=0}^{N} \frac{A^k B^\ell}{k!\ell!} - \sum_{k=0}^{N}\sum_{\ell=0}^{N-k} \frac{A^k B^\ell}{k!\ell!}\right\| \\
&= \left\|\sum_{k=0}^{N}\sum_{\ell=N-k+1}^{N} \frac{A^k B^\ell}{k!\ell!}\right\| \\
&\leqslant \sum_{k=0}^{N}\sum_{\ell=N-k+1}^{N} \frac{n\|A^k\|\|B^\ell\|}{k!\ell!}
\end{aligned}$$

En reprenant alors tout le calcul en sens inverse, il vient :

$$\left\| \left(\sum_{k=0}^{n} \frac{A^k}{k!}\right)\left(\sum_{\ell=0}^{n} \frac{B^\ell}{\ell!}\right) - \sum_{i=0}^{n} \frac{(A+B)^i}{i!} \right\| \leqslant n\left(\left(\sum_{k=0}^{N} \frac{\|A\|^k}{k!}\right)\left(\sum_{\ell=0}^{N} \frac{\|B\|^\ell}{\ell!}\right) - \sum_{i=0}^{N} \frac{(\|A\|+\|B\|)^i}{i!}\right)$$
$$\underset{N\to+\infty}{\longrightarrow} n(\mathrm{e}^{\|A\|}\mathrm{e}^{\|B\|} - \mathrm{e}^{\|A\|+\|B\|}) = 0,$$

d'après les règles concernant l'exponentielle complexe.

Ainsi, d'après le théorème d'encadrement, $\left\| \left(\sum_{k=0}^{n} \frac{A^k}{k!}\right)\left(\sum_{\ell=0}^{n} \frac{B^\ell}{\ell!}\right) - \sum_{i=0}^{n} \frac{(A+B)^i}{i!} \right\| \longrightarrow 0$, et les règles sur les limites matricielles, déjà utilisées plus haut, permettent de conclure que

$$\boxed{\exp(A)\exp(B) = \exp(A+B)}.$$

Remarques, à lire pour votre édification personnelle

La majoration ne doit pas être faite trop brutalement, il faut d'abord compenser tous les termes qui se compensent, sinon on ne parvient pas à exploiter la propriété similaire de l'exponentielle complexe.

Remarquez aussi qu'on s'est ramené à l'exponentielle complexe, pour laquelle le résultat est connu pour des raisons analytiques, vu notre définition de l'exponentielle (soit par l'équation différentielle $y' = y$, soit comme réciproque de ln, soit encore et de façon plus directe comme solution continue de $f(x+y) = f(x)f(y)$, avec les conditions initiales requises). Dans ce cas, on fait le lien entre la fonction exponentielle et sa série par les formules de Taylor.

Cependant, on pourrait très bien définir l'exponentielle réelle (et complexe) par la somme de la série exponentielle, et établir ses propriétés à partir de là. Dans ce cas, la preuve qu'on a faite ne fait que nous ramener au problème similaire pour l'exponentielle réelle. On tourne en rond. On peut s'en sortir par une idée proche à celle développée ci-dessus, lorsque x et y sont positifs, en remarquant que :

$$\sum_{k=0}^{\lfloor \frac{N}{2} \rfloor} \sum_{\ell=0}^{\lfloor \frac{N}{2} \rfloor} \frac{x^k}{k!}\frac{y^\ell}{\ell!} \leqslant \sum_{i=0}^{N} \frac{(x+y)^i}{i!} \leqslant \sum_{k=0}^{N}\sum_{\ell=0}^{N} \frac{x^k}{k!}\frac{y^\ell}{\ell!}.$$

En effet, après développement du terme du milieu, on se retrouve à sommer des termes identitiques et positifs, sur des ensembles d'indices inclus les uns dans les autres (un carré de coté $\lfloor \frac{N}{2} \rfloor$, inclus dans un triangle $k+\ell \leqslant N$, $k, \ell \geqslant 0$, inclus dans un carré de côté N : faites un dessin !). Or, les deux sommes de droite et de gauche tendent vers $\exp(A)\exp(B)$. Cela assure la convergence de celle du milieu (mais dans la situation présente on le savait déjà), et surtout l'égalité des limites, qui nous fournit ce qu'on voulait. En utilisant cela avec $x = \|A\|$ et $y = \|B\|$, cela complète la preuve.

Le procédé exposé ci-dessus est valable pour montrer la convergence de $\sum c_n$ vers $(\sum a_n)(\sum b_n)$, lorsque toutes les séries sont à termes positifs, et (c_n) est défini par le produit de convolution

$$\forall n \in \mathbb{N}, \quad c_n = \sum_{k=0}^{n} a_k b_{n-k}.$$

C'est ce qu'on appelle le produit de Cauchy des séries.

À partir de la propriété pour les séries à termes positifs, on récupère le cas des séries absolument convergentes (c'est-à-dire telles que $\sum |a_n|$ converge, ce qui implique la convergence de $\sum a_n$), par le même argument que celui qui nous a permis de nous ramener de l'exponentielle matricielle à l'exponentielle réelle. On obtient donc la validité du produit de Cauchy pour des séries absolument convergentes.

5. D'après ce qui précède, il faut bien sûr chercher un contre-exemple tel que A et B ne commutent pas. On essaye avec $A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ et $B = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$. On trouve :

$$\forall k \geqslant 1, A^n = A,$$

et donc (attention au fait que l'égalité précédente n'est pas valable pour $k = 0$)

$$\exp(A) = A \sum_{k=0}^{+\infty} \frac{1}{k} - A + I_2 = \begin{pmatrix} e & 0 \\ 0 & 1 \end{pmatrix}.$$

Remarquez au passage (fait déjà entrevu plus haut), qu'appliquer l'exponentielle à une matrice diagonale revient à appliquer l'exponentielle à ses coefficients diagonaux.
Pour B, on a directement $B^2 = 0$, donc

$$\exp(B) = I_2 + B = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

Enfin, soit $C = A + B = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$. On calcule sans peine :

$$\forall k \geqslant 1, \quad C^k = C,$$

et le même argument que pour A donne :

$$\exp(C) = eC - C + I_2 = \begin{pmatrix} e & e-1 \\ 0 & 1 \end{pmatrix}.$$

Or,

$$\exp(A)\exp(B) = \begin{pmatrix} e & e \\ 0 & 1 \end{pmatrix}.$$

Ainsi, $\boxed{exp(A + B) \neq \exp(A)\exp(B)}$.

6. Nous posons l'hypothèse $AB = BA$, qui nous permet d'utiliser l'exponentielle. Il semble évident que $\cos(A) + \mathrm{i}\sin(A) = \exp(\mathrm{i}\,A)$: considérez les sommes partielles de rang N, cela donne en formant l'expression $\cos(A) + \mathrm{i}\sin(A)$ la somme partielle de rang $2N$ de $\exp(\mathrm{i}\,A)$. Ainsi, en combinant les relations obtenues pour A et $-A$, on obtient l'analogue matriciel des formules d'Euler :

$$\cos(A) = \frac{1}{2}(\exp(\mathrm{i}\,A) + \exp(-\mathrm{i}\,A)) \qquad \text{et} \qquad \cos(A) = \frac{1}{2\,\mathrm{i}}(\exp(\mathrm{i}\,A) - \exp(-\mathrm{i}\,A))$$

On peut aussi constater que si A et B sont à coefficients réels, comme stipulé par l'énoncé, et en notant, pour $C \in \mathcal{M}_n(\mathbb{C})$, $\mathrm{Re}(C)$ la matrice des parties réelles des coefficients de C, et de même pour $\mathrm{Im}(C)$, on a la relation :

$$\cos(A) = \mathrm{Re}(\exp(\mathrm{i}\,A)) \qquad \text{et} \qquad \sin(A) = \mathrm{Im}(\exp(\mathrm{i}\,A)).$$

On a alors, en utilisant la commutativité de A et B (donc de $\mathrm{i}\,A$ et $\mathrm{i}\,B$) :

$$\cos(A + B) = \mathrm{Re}(\exp(\mathrm{i}\,A + \mathrm{i}\,B)) = \mathrm{Re}(\exp(\mathrm{i}\,A)\exp(\mathrm{i}\,B)) = \mathrm{Re}\left((\cos(A) + \mathrm{i}\sin(A))(\cos(A) - \mathrm{i}\sin(A))\right)$$

(les définitions par les séries donnent sur cos et sin les propriétés de parité que l'on connait bien sur $\mathbb{R}$). Ainsi, en développant et en extrayant la partie réelle, il vient :

$$\boxed{\cos(A + B) = \cos(A)\cos(B) - \sin(A)\sin(B)}.$$

De même, $\sin(A + B)$ est la partie imaginaire de l'expression ci-dessus :

$$\boxed{\sin(A + B) = \sin(A)\cos(B) + \cos(A)\sin(B)}.$$

On retrouve sans surprise les formules bien connues sur $\mathbb{R}$ (mais attention à l'hypothèse $AB = BA$ nécessaire pour les obtenir !)
On pourrait étendre ces formules au cas ou A et B sont à ceofficients complexes, mais dans ce cas, on ne peut plus raisonner avec les parties réelles et imaginaires : il faut tout faire à partir des formules d'Euler. On peut aussi définir de la sorte $\mathrm{ch}(A)$ et $\mathrm{sh}(A)$. On se convaincra sans peine de la validité de toutes les formules de trigonométrie circulaire ou hyperbolique portant sur ces fonctions, sous l'hypothèse de commutation.