D.M. 13 : Cauchy-Lipschitz et l'exponentielle matricielle, solutions

Partie I :

1) Par déf.

$$|||A||| = \sup_{\|X\|=1} \|AX\ \| = \sup_{\|X\|\leq 1} \|AX\| = \sup_{X\in F\setminus\{0\}} \frac{\|AX\|}{\|X\|}.$$

La dernière définition équivaut encore, par linéarité à dire que $|||A|||$ est le plus petit rapport de Lipschitz de $X \mapsto AX$.

2) Par récurrence sur n :

— Φ_0 est bien définie et de classe C^1 (fonction constante).

— supposons que pour un $n \geq 0$ Φ_n est bien définie et de classe C^1. Alors Φ_n est en particulier continue, ainsi par hypothèse que A et B. Il en résulte que la fonction $s \longmapsto A(s).\Phi_n(s) + B(s)$ est aussi continue sur I

Ainsi Φ_{n+1} est donc bien définie et, par théorème sur fondamental sur le lien intégrale/primitive pour les fonctions continues Φ_{n+1} est de classe C^1.

La récurrence est établie.

3. a) Les fonctions A et $\Phi_1 - \Phi_0$ sont continues, donc bornées sur le segment $[\alpha, \beta]$, d'o l'existence de μ et M.

b) La linéarité de l'intégrale et de $A(s)$ donnent

$$\Phi_{n+1}(t) - \Phi_n(t) = \int_{t_0}^{t} A(s)\left(\Phi_n(s) - \Phi_{n-1}(s)\right) \mathrm{d}s.$$

Ensuite, d'après l'inégalité triangulaire sur les intégrales et la déf. de la norme subordonnée rappelée au 1) :

$$\|\Phi_{n+1}(t) - \Phi_n(t)\| \leqslant \int_{\min(t_0,t)}^{\max(t_0,t)} \|A(s)\left(\Phi_n(s) - \Phi_{n-1}(s)\right)\| \mathrm{d}s \leqslant \int_{\min(t_0,t)}^{\max(t_0,t)} |||A(s)|||.\, \|\Phi_n(s) - \Phi_{n-1}(s)\| \mathrm{d}s.$$

Enfin, $\|\Phi_{n+1}(t) - \Phi_n(t)\| \leqslant \mu \int_{\min(t_0,t)}^{\max(t_0,t)} \|\Phi_n(s) - \Phi_{n-1}(s)\| \mathrm{d}s$ puisque $[\min(t_0,t), \max(t_0,t)] \subset [\alpha, \beta]$.

c) Preuve facile par récurrence sur n, en utilisant l'inégalité du b).

N.B. : l'hypothèse de récurrence s'applique tout $s \in [\min(t_0,t), \max(t_0,t)]$ car ce segment est inclus dans $[\alpha, \beta]$.

4. a) Avec les notations du 3. on a :

$$\forall n \in \mathbb{N}, \forall t \in [\alpha, \beta], \|\Phi_{n+1}(t) - \Phi_n(t)\| \leqslant \frac{M\mu^n(\beta-\alpha)^n}{n!}$$

majorant qui est le terme général d'une série (exponentielle) convergente indépendante de t. Ainsi la série $\sum(\Phi_{n+1} - \Phi_n)$ converge normalement, donc uniformément, sur tout segment de I qui contient t_0. Comme tout segment de I est inclus dans un segment de I contenant t_0, $\sum(\Phi_{n+1} - \Phi_n)$ converge uniformément sur tout segment de I.

Par lien suites-séries (sommes télescopiques), la suite (Φ_n) converge uniformément sur tout segment de I.

4.b) Pour $t \in I$ posons $\Psi(t) = A(t).\Phi(t) + b(t)$. Par linéarité de la multiplication par $A(t)$,

$$\Psi_n(t) - \Psi(t) = A(t)\left(\Phi_n(t) - \Phi(t)\right).$$

On fixe un segment $[\alpha, \beta]$ de I et on définit μ comme au 3.a). Ensuite, pour $n \in \mathbb{N}$, on pose $\varepsilon_n = \sup_{t\in[\alpha,\beta]} \|\Phi_n(t) - \Phi(t)\|$. Selon a), la suite (ε_n) converge vers 0 .

Pour tout $t \in [\alpha, \beta]$,

$$\|\Psi_n(t) - \Psi(t)\| = \|A(t)\left(\Phi_n(t) - \Phi(t)\right)\| \leqslant |||A(t)|||\, \|\Phi_n(t) - \Phi(t)\| \leqslant \mu\varepsilon_n$$

Le majorant tend vers 0 et ne dépend pas de t. La suite (Ψ_n) converge donc vers Ψ uniformément sur tout segment de I.
c) Fixons $t \in I$. Par définition de Φ_n et Ψ_n, pour tout $n \in \mathbb{N}$,

$$\Phi_{n+1}(t) = X_0 + \int_{t_0}^{t} \Psi_n(s)\mathrm{d}s.$$

Par passage aux limites, a) et b) donnent

$$\Phi(t) = X_0 + \int_{t_0}^{t} \Psi(s)\mathrm{d}s = X_0 + \int_{t_0}^{t} (A(s)\Phi(s) + b(s))\mathrm{d}s.$$

Cela signifie exactement que Φ est solution du problème de Cauchy $((E), (t_0, X_0))$ (ici écrit sous forme intégrale), cf 1) b).
5. a) Δ est de classe C^1 car X_1 et X_2 le sont et $\Delta(t_0) = 0$; par conséquent :

$$\begin{aligned}\Delta(t) = \int_{t_0}^{t} \Delta'(s)\mathrm{d}s &= \int_{t_0}^{t} \left(X_1'(s) - X_2'(s)\right)\mathrm{d}s \\ &= \int_{t_0}^{t} A(s).(\Delta(s))\mathrm{d}s \text{ puisque } X_1 \text{ et } X_2 \text{ sont solutions de } (E).\end{aligned}$$

b) On reproduit pratiquement la preuve de 3. b) et c).

— on montre d'abord comme en 3.b) :

$$\forall t \in [\alpha, \beta], \|\Delta(t)\| \leqslant \mu \int_{\min(t_0,t)}^{\max(t_0,t)} \|\Delta(s)\|\mathrm{d}s$$

— on prouve ensuite comme en 3.c) l'inégalité demandée par récurrence sur n en majorant dans l'intégrale $\|\Delta(s)\|$ l'aide de l'hypothèse de récurrence.

c) Le membre de droite de l'ingalité du b) tend vers 0 quand n tend vers l'infini. Le passage la limite donne donc $\Delta(t) = 0$, c'est–dire $X_1(t) = X_2(t)$, pour tout $t \in [\alpha, \beta]$. Cela étant valable pour tout segment $[\alpha, \beta]$ contenant t_0, on conclut que $X_1 = X_2$.
On a ainsi démontré l'existence d'une solution à un problème de Cauchy au 4. et son unicité au 5.

II La surjectivité de l'exp. matricielle

6) a) Il s'agit d'une question de cours sur la décomposition sur les s.e.v. caractéristiques.
Soit $E = \mathbb{C}^n$ et $a \in \mathscr{L}(E)$ l'endomorphisme canoniquement associé à A. Par Cayley-Hamilton, on sait que $\chi_A(a) = 0$.
Par théorème de décomposition des noyaux, $E = \oplus_{i=1}^{k} \ker(a - \lambda_i \,\mathrm{id})^{r_i}$.
Soit $\mathcal{B} = (\mathbb{B}_1, \dots, \mathcal{B}_k)$ une base de E adaptée à cette décomposition en somme directe, où $\mathcal{B}_i$ est une base de $E_i \coloneqq \ker(a - \lambda_i \,\mathrm{id})^{r_i}$.
Comme chacun des ces sous-espaces E_i est stable par a, la matrice A' de a dans la base $\mathcal{B}$ est de la forme : $A' = \mathrm{diag}(M_1, \dots, M_k)$ où $M_k = \mathrm{Mat}_{\mathcal{B}_i}(a_{E_i})$.
Comme $(a_{E_i} - \lambda_i)^{r_i} = 0$, on peut écrire $a_{E_i} = \lambda_i \,\mathrm{id}_{E_i} + \nu_i$ avec ν_i nilpotente.
Matriciellement $M_i = \lambda_i I + N_i$, ce qui donne la conclusion demandée : $\mathbb{A}' = operatornameDiag\,(\lambda_1 I_{r_1} + N_1, \dots, \lambda_k$
b) (i) Notons $f_k(t) = M(t)^k$ et démontrons le résultat demandé par récurrence sur k. C'est évident pour $k = 1$. Supposons le résultat acquis au rang k. $f_{k+1}(t) = f_1(t)f_k(t)$ donc par l'hypothèse de récurrence f_{k+1} est de classe C^1 et pour tout $t \in \mathbb{R}$:

$$\begin{aligned}f_{k+1}'(t) = f_1'(t)f_k(t) + f_1(t)f_k'(t) &= M'(t)M(t)^k + kM(t)M'(t)M(t)^{k-1} \\ &= M'(t)M(t)^k + kM'(t)M(t)^k = (k+1)M'(t)M(t)^k.\end{aligned}$$

La récurrence est établie.
(ii) b) Notons $f(t) = \exp M(t) = \sum_{k=0}^{+\infty} u_k(t)$, où $u_k(t) = \frac{M(t)^k}{k!}$.

Selon (i), les u_k sont de classe C^1, et $u_0' = 0$ et $u_k'(t) = \frac{M'(t)M(t)^{k-1}}{(k-1)!}$ pour $k \geqslant 1$.

Soit $\|\cdot\|$ une norme d'algèbre sur $\mathscr{M}_n(\mathbb{C})$.

On en déduit que pour tout $(k,t) \in \mathbb{N}^* \times \mathbb{R}$, $\|u_k'(t)\| \leqslant \frac{\|M'(t)\|\|M(t)\|^{k-1}}{(k-1)!}$.

Soit maintenant $R \in \mathbb{R}_+^*$, les fonctions M et M' sont continues, donc bornées sur $[-R,R]$, d'où l'existence de deux constantes K et K' telles que :

$$\forall t \in [-R,R], \|u_k'(t)\| \leqslant K' \frac{(K)^{k-1}}{(k-1)!},$$

majoration par le terme général d'une série convergente indépendante de t.

La série $\sum u_k'$ converge donc normalement sur tout segment de $\mathbb{R}$.

Par théorème, il en résulte que f est de classe C^1 et que pour tout $t \in \mathbb{R}$:

$$f'(t) = \sum_{k=0}^{+\infty} u_k'(t) = \sum_{k=1}^{+\infty} \frac{M'(t)M(t)^{k-1}}{(k-1)!} = M'(t)\exp M(t)$$

7. a) Grâce à l'identité géométrique $(1-X)(1+X+\cdots+X^{n-1}) = 1-X^n$, on calcule ici :

$$(I_n+tN)\,L'(t) = (I_n+tN)\sum_{k=1}^{n-1}(-1)^{k-1}t^{k-1}N^k = (I_n-(-tN))\left(\sum_{k=0}^{n-2}(-tN)^k\right)N$$

$$= \left(I_n-(-tN)^{n-1}\right)N = N \text{ car } N^n = 0.$$

b) $L(t)$ et $L'(t)$ commutent car ce sont des polynômes en N. On peut donc appliquer le 6 b), ce qui donne, avec la formule sur la dérivée d'un produit :

$$\varphi'(t) = N\exp(-L(t)) - (I_n+tN)\,L'(t)\exp(-L(t))$$

et donc avec le a), on obtient ;

$$\forall\, t \in \mathbb{R}, \varphi'(t) = 0.$$

c) On déduit du b) que φ est constante et en particulier que $\varphi(1) = \varphi(0)$. Cela donne

$$(I_n+N)\exp(-L(1)) = I_n,$$

puis

$$\exp L(1) = I_n + N$$

d) En posant $\lambda = re^{i\theta}$ avec $(r,\theta) \in \mathbb{R}_+^* \times \mathbb{R}$, les complexes μ tels que $e^\mu = \lambda$ sont les $\ln r + i(\theta+2k\pi)$ avec $k \in \mathbb{Z}$.

e) N/λ est nilpotente comme N donc, d'après c), il existe $M_0 \in \mathscr{M}_n(\mathbb{C})$ telle que

$$\exp M_0 = I_n + N/\lambda.$$

Posons $M = \mu I_n + M_0$, où $\mu \in \mathbb{C}$ vérifie $e^\mu = \lambda$. Alors

$$\exp M = e^\mu \exp M_0 = \lambda\,(I_n + N/\lambda) = \lambda I_n + N$$

8) Soit $M \in GL_n(\mathbb{C})$. Par 6) a), il existe $P \in GL_n(\mathbb{C})$ telle que :

$$M = P\,\mathrm{diag}(M_1,\ldots,M_k)P^{-1}$$

où $M_i = \lambda_i I + N_i$ et N_i nilpotente, pour tout $i = 1,\ldots,k$ et tous les λ_i sont non nuls (valeurs propres de M).

Par 7) e), pour chaque i il existe une matrice R_i telle que $\exp(R_i) = M_i$.

Soit alors $R = P\,\mathrm{diag}(R_1,\ldots,R_k)P^{-1}$. On a bien $\exp(R) = M$ et la surjectivité annoncée.

9) Soit $A \in \mathrm{GL}_n(\mathbb{C})$. Selon 8., il existe $M \in \mathscr{M}_n(\mathbb{C})$ telle que $\exp M = A$. On en déduit que $A = (\exp(M/k))^k$, avec $\exp(M/k) \in \mathrm{GL}_n(\mathbb{C})$, d'où le résultat.

10. a) On a vu la question 7., dont on conserve les notations, que $A = \lambda I_n + N = \exp M$, avec :

$$M = \mu I_n + M_0 = \mu I_n + \sum_{k=1}^{n-1} \frac{(-1)^{k-1}}{k} \left(\frac{N}{\lambda}\right)^k = \mu I_n + \sum_{k=1}^{n-1} \frac{(-1)^{k-1}}{k} \left(\frac{A}{\lambda} - I_n\right)^k.$$

En effet, $M_0 = \widetilde{L}(1)$, où $\widetilde{L}$ est définie à partir de N/λ comme L l'est à partir de N. On constate sur son expression que $M \in \mathbb{C}[A]$.

b) Remarquons d'abord que comme N_i est nilpotente, $(X - \lambda_i)^{r_i}$ annule A'_i, donc μ_i divise $(X - \lambda_i)^{r_i}$, donc les μ_i sont deux à deux premiers entre eux.

La question posée revient en fait au théorème des restes chinois pour le système $\begin{cases} P \equiv P_1 \ [\mu_1] \\ \vdots \\ P \equiv P_k \ [\mu_k] \end{cases}$.

On va redémontrer ce théorème dans ce cadre polynomial ici, en utilisant (ce n'est pas vraiment nécessaire en fait) la structure vectorielle. Comparer à la preuve du cours d'arithmétique.

Posons $\mu = \prod_{i=1}^k \mu_i$, et notons $d_i = \deg \mu_i, d = \deg \mu = \sum_{i=1}^k d_i$.

Pour $i \in [\![1, k]\!]$ et $P \in \mathbb{C}[X]$ on note $R_i(P)$ le reste de la division euclidienne de P par μ_i. Considérons alors l'application, manifestement linéaire,

$$R : \left\{ \begin{array}{ccc} \mathbb{C}_{d-1}[X] & \longrightarrow & \mathbb{C}_{d_1-1}[X] \times \ldots \times \mathbb{C}_{d_k-1}[X] \\ P & \longmapsto & (R_1(P), \ldots, R_k(P)). \end{array} \right.$$

Si $P \in \operatorname{Ker} R$, alors P est divisible par tous les μ_i, donc aussi par leur produit μ puisque les μ_i sont deux à deux premiers entre eux ; comme $\deg P < \deg \mu$, cela implique que $P = 0$. L'application R est donc injective. De plus, l'espace de départ et l'espace d'arrivée de R sont tous les deux de dimension d, donc R est bijective. En particulier, il existe un (unique) $P \in \mathbb{C}_{d-1}[X]$ tel que $R(P) = (R_1(P_1), \ldots, R_k(P_k))$. Ainsi, pour tout $i \in [\![1, k]\!], P$ et P_i ont le même reste dans la division par μ_i, ce qui signifie que μ_i divise $P - P_i$.

c) μ_i divise $P - P_i$, donc $(P - P_i)(A'_i) = 0$, donc $P(A'_i) = P_i(A'_i)$, donc $\exp P(A'_i) = A'_i$. Il vient alors, par blocs, $P(A') = \operatorname{Diag}(P(A'_1), \ldots, P(A'_k))$ puis :

$$\exp P(A') = \operatorname{Diag}(\exp P(A'_1), \ldots, \exp P(A'_k)) = \operatorname{Diag}(A'_1, \ldots, A'_k) = A'.$$

Ensuite, comme au 8., en changeant les notations $A = UA'U^{-1}$ avec $U \in GL_n(\mathbb{C})$ d'où $P(A) = UP(A')U^{-1}$ et finalement :

$$\exp P(A) = U(\exp P(A'))U^{-1} = UA'U^{-1} = A.$$

11. a) Si $A = \exp M$, alors $A = \exp(M/2)^2$ et $\exp(M/2) \in \mathrm{GL}_n(\mathbb{R})$.

b) C'est le résultat du 10. appliqué à R vu comme un élément de $\mathrm{GL}_n(\mathbb{C})$.

c) La matrice R étant réelle, on a aussi :

$$R = \overline{\exp P(R)} = \exp \overline{P(R)} = \exp \bar{P}(R).$$

Comme $P(R)$ et $\bar{P}(R)$ commutent, on sait que

$$A = R^2 = (\exp P(R))(\exp \bar{P}(R)) = \exp(P + \bar{P})(R)$$

et $(P + \bar{P})(R) \in \mathscr{M}_n(\mathbb{R})$ puisque $P + \bar{P} \in \mathbb{R}[X]$ et $R \in \mathscr{M}_n(\mathbb{R})$.

d) Pour l'inclusion (facile) : Si $A = \exp M$ avec $M \in \mathscr{M}_n(\mathbb{R})$, alors $A = \exp(M/2)^2$ donc $\det A = \det(\exp(M/2))^2 > 0$, donc $A \in \mathrm{GL}_n^+(\mathbb{R})$.

Montrons que l'inclusion est stricte en considérant la matrice $D = \operatorname{Diag}(1, 2, \ldots, n-2, -1, -2)$. On voit que $\det D = 2(n-2)! > 0$, donc $D \in \mathrm{GL}_n^+(\mathbb{R})$.

Par l'absurde, supposons qu'il existe $M \in \mathscr{M}_n(\mathbb{R})$ tel que $D = \exp M$.

Alors on sait que D commute avec M. Or pour D diagonale à termes diagonaux deux à deux distincts, on sait que les matrices qui commutent à D sont aussi diagonales.

Donc M est diagonale, donc $e^{m_{n,n}} = -2$, ce qui est absurde. On a ainsi établi l'inclusion stricte demande.

Partie III : formule $\exp(A+B)$, extrait de Mines-Ponts MP 2022

12) a) Si $AB = BA$, alors par une récurrence immédiate, on a $AB^k = B^kA$ pour tout k entier naturel puis, par combinaisons linéaires, $A\cdot Q(B) = Q(B)\cdot A$ pour tout polynôme Q de $\mathbb{K}[X]$. En particulier, pour tout k entier naturel,

$$A\cdot\left(\sum_{j=0}^{k}\frac{B^j}{j!}\right) = \left(\sum_{j=0}^{k}\frac{B^j}{j!}\right)\cdot A$$

En faisant tendre k vers l'infini, par continuité du produit matriciel, il vient $A\cdot e^B = e^B\cdot A$.

b) Par dérivation d'un produit (justifié car le produit matriciel est bilinéaire), on obtient

$$g'(t) = (A+B)\cdot e^{t(A+B)}e^{-tB} - e^{t(A+B)}\cdot B\cdot e^{-tB}.$$

Mais, si A et B commutent, alors B et $t(A+B)$ commutent pour tout t, donc B et $e^{t(A+B)}$ commutent d'après le a). Finalement,

$$g'(t) = (A+B)\cdot e^{t(A+B)}e^{-tB} - B\cdot e^{t(A+B)}e^{-tB} = A\cdot g(t).$$

D'autre part, $f_A'(t) = A\cdot f_A(t)$. Comme $f_A(0) = g(0) = I_n$, les fonctions f_A et g sont solutions du même problème de Cauchy $(\mathbf{P})$: $\begin{cases} M'(t) = AM(t) \\ M(0) = I_n \end{cases}$.

c) Notons que l'équation différentielle $M'(t) = AM(t)$, où la fonction inconnue est $M : \mathbb{R}\to\mathcal{M}_n(\mathbb{K})$ de classe $\mathcal{C}^1$, peut s'écrire sous la forme $M'(t) = \alpha(M(t))$, où α est l'endomorphisme de $\mathcal{M}_n(\mathbb{K})$ défini par $\alpha(M) = AM$. Il s'agit donc bien d'un problème de Cauchy dans les termes du programme MP, et le théorème de Cauchy linéaire démontré dans la partie I, donne l'unicité de la solution d'un tel problème de Cauchy, on déduit que $g = f_A$. En particulier, dans le cas où $A = 0_n$, on obtient $e^{tB}e^{-tB} = I_n$, ce qui montre que la matrice e^{tB} est inversible, d'inverse e^{-tB}. On en déduit immédiatement la relation $(*)$ de l'énoncé.

d) Par produit de Cauchy de série absolument convergente, cf. chapitre T1.

13) Supposons la relation $(*)$ satisfaite. Les deux membres sont évidemment des fonctions de classe $\mathcal{C}^\infty$ de la variable t. Une première dérivation donne par exemple

$$\forall t\in\mathbb{R}\quad (A+B)e^{t(A+B)} = e^{tA}\cdot A\cdot e^{tB} + e^{tA}\cdot B\cdot e^{tB} = e^{tA}(A+B)e^{tB}$$

Une deuxième dérivation fournit
$\forall t\in\mathbb{R}\quad (A+B)^2e^{t(A+B)} = e^{tA}\cdot A(A+B)e^{tB} + e^{tA}(A+B)B\cdot e^{tB} = e^{tA}\left(A^2+2AB+B^2\right)e^{tB}.$

En évaluant pour $t=0$, on obtient $(A+B)^2 = A^2+2AB+B^2$, ce qui se simplifie en $BA = AB$.

14) Pour une norme d'algèbre, on montre immédiatement que pour toute matrice A, on a $\|\exp(A)\| \le \exp(\|A\|)$ et donc :

$$\|X_k\| \le \left\|\exp\left(\frac{A}{k}\right)\right\|\left\|\exp\left(\frac{B}{k}\right)\right\| \le \exp\left(\frac{\|A\|}{k}\right)\exp\left(\frac{\|B\|}{k}\right) = \exp\left(\frac{\|A\|+\|B\|}{k}\right),$$

et, par inégalité triangulaire et croissance de la fonction exponentielle de $\mathbb{R}$ dans $\mathbb{R}$:

$$\|Y_k\| \le \exp\left(\frac{\|A+B\|}{k}\right) \le \exp\left(\frac{\|A\|+\|B\|}{k}\right)$$

15) La fonction h est de classe $\mathcal{C}^2$ sur $\mathbb{R}$ par théorèmes classiques d'opérations, et on calcule

$$\forall t\in\mathbb{R}\quad h'(t) = e^{tA}(A+B)e^{tB} - (A+B)e^{t(A+B)}.$$

On observe notamment que $h(0) = 0_n$ et $h'(0) = 0_n$. La formule de Taylor-Young à l'ordre 2 pour les fonctions vectorielles donne alors le développement

$$h(t) = 0_n + t0_n + \frac{t^2}{2}h''(0) + o\left(t^2\right)$$

dont nous retiendrons seulement que

$$h(t) = O\left(t^2\right) \qquad \text{lorsque } t \to 0.$$

Il en résulte que

$$X_k - Y_k = h\left(\frac{1}{k}\right) = O\left(\frac{1}{k^2}\right) \qquad \text{lorsque } k \to +\infty.$$

16) On reconnaît une somme télescopique :

$$\sum_{i=0}^{k-1} X_k^i \left(X_k - Y_k\right) Y_k^{k-i-1} = \sum_{i=0}^{k-1} \left(X_k^{i+1} Y_k^{k-(i+1)} - X_k^i Y_k^{k-i}\right) = X_k^k - Y_k^k$$

17) La question 15) nous apprend qu'il existe un réel strictement positif C tel que, pour k entier naturel assez grand, on ait $\|X_k - Y_k\| \leq \frac{C}{k^2}$. À l'aide des propriétés de la norme $\|\cdot\|$, on majore :

$$\left\|X_k^k - Y_k^k\right\| \leq \sum_{i=0}^{k-1} \left\|X_k^i \left(X_k - Y_k\right) Y_k^{k-i-1}\right\| \leq \sum_{i=0}^{k-1} \|X_k\|^i \|X_k - Y_k\| \|Y_k\|^{k-i-1} = \|X_k - Y_k\| \sum_{i=0}^{k-1} \|X_k\|^i \|Y_k\|^{k-i-1}$$

Par ailleurs, grâce à 14) ,

$\|X_k\|^i \|Y_k\|^{k-i-1} \leq \left[\exp\left(\frac{\|A\|+\|B\|}{k}\right)\right]^i \left[\exp\left(\frac{\|A\|+\|B\|}{k}\right)\right]^{k-i-1} = \exp\left((k-1)\frac{\|A\|+\|B\|}{k}\right)$

donc $\|X_k\|^i \|Y_k\|^{k-i-1} \leq \exp(\|A\| + \|B\|)$. Comme la somme comporte k termes, on aboutit donc à

$$\left\|X_k^k - Y_k^k\right\| \leq k \cdot \exp(\|A\| + \|B\|) \cdot \frac{C}{k^2}$$

il en résulte que $\lim_{k\to+\infty} \left(X_k^k - Y_k^k\right) = 0$. Comme $\left(Y_k^k\right)$ est une suite constante de valeur e^{A+B}, on conclut que $\lim_{k\to+\infty} X_k^k = e^{A+B}$, ce qu'il fallait démontrer.