Image de l'exponentielle réelle

October 20, 2006

Un nombre réel est une exponentielle si, et seulement si, c'est un carré non nul. Nous allons voir que cela se généralise au cas des matrices, c'est l'objet du théorème suivant.

Théorème 0.1 (Image de l'exponentielle de matrice réelle)

$$\exp(M_p(\mathbb{R})) = \left\{R^2, \quad R \in GL_p(\mathbb{R})\right\}.$$

Autrement dit, une matrice réelle inversible est une exponentielle si, et seulement si, c'est un carré.
Ce résultat est assez remarquable: pour savoir si une matrice inversible réelle est une exponentielle, il suffit de tester si c'est une matrice carrée, ce qui est généralement plus simple.

Ce texte comprend trois preuves différentes.

1 Preuve n°1

La preuve que nous allons donner est tirée d'un article de Michel Coste [www1], enseignant à la prepa-agreg de Rennes.
Nous aurons auparavant besoin de démontrer un lemme qui raffine la surjectivité de exp : $M_p(\mathbb{C}) \to Gl_p(\mathbb{C})$.

Lemme 1.1 (Raffinement de la surjectivité de l'exponentielle complexe) *Pour toute matrice M de $GL_p(\mathbb{C})$, il existe un polynôme $P \in \mathbb{C}[X]$ tel que $M = \exp(P(M))$.*

Preuve:

- Soit $M \in GL_p(\mathbb{C})$. On écrit sa décomposition de Dunford (χ_M est scindé sur $\mathbb{C}$):
$M = D + N$ avec D diagonalisable sur $\mathbb{C}$, N nilpotente, et D et N qui commutent. On utilisera en plus que D et N sont des polynômes en M, c'est un ingrédient essentiel.
Puisque D est inversible, $M = D(I_p + D^{-1}N)$. L'idée est de trouver deux polynômes U et V tels que $D = \exp(U(M))$ et $I_p + D^{-1}N = \exp(U(M))$.
Puisque $U(M)$ et $V(M)$ commutent, on aura alors $M = \exp(U(M) + V(M))$ ce qui démontre le lemme avec $P = U + V$.

- On a $D = P\text{diag}(\lambda_1, ..., \lambda_p)P^{-1}$ où les λ_i sont les valeurs propres toutes non nulles puisque M est inversible.
Pour tout $\lambda \in \text{Spec}(M)$, il existe un nombre complexe μ_λ tel que $\exp \mu_\lambda = \lambda$ puisque $\exp : \mathbb{C} \to \mathbb{C}^*$ est surjective.
Il existe alors un polynôme interpolateur de Lagrange L tel que pour tout $\mu \in \text{Spec}(M)$, $L(\lambda) = \mu_\lambda$

puisque les λ sont disctincts 2à 2. Alors

$$\begin{aligned} D &= P\text{diag}(\lambda_1, ..., \lambda_p)P^{-1} \\ &= P\text{diag}(\exp\mu_1, ..., \exp\mu_p)P^{-1} \\ &= \exp\left(P\text{diag}(\mu_1, ..., \mu_p)P^{-1}\right) \\ &= \exp\left(P\text{diag}(L(\lambda_1), ..., L(\lambda_p))P^{-1}\right) \\ &= \exp\left((L(P\text{diag}(\lambda_1, ..., \lambda_p)P^{-1}\right) \\ &= \exp(L(D)) \end{aligned}$$

Comme D est un polynôme en M, $L(D)$ est aussi un polynôme en M, ce qui donne l'existence du polynôme U.

- Montrons que $I_p + D^{-1}N$ est unipotente. La matrice D^{-1} est un polynôme en D. En effet, d'après le théorème de Cayley Hamilton, $\chi_D(D) = 0$, ce qui donne

$$(-1)^p D^p + a_{p-1}D^{p-1} + ... + a_1 D + a_0 I_p = 0$$

avec $a_0 = \det D \neq 0$. En composant par D^{-1} puis en divisant par a_0, on a donc

$$D^{-1} = {a_0}^{-1}\left((-1)^p D^{p-1} + a_{p-1}D^{p-2} + ... + a_1 I_p\right),$$

ce qui donne le résultat.
N commute avec D donc avec D^{-1} puisque c'est un polynôme en D.
On a alors $(D^{-1}N)^p = (D^{-1})^p N^p = 0$ ce qui prouve que $D^{-1}N$ est nilpotente et donc que $I_p + D^{-1}N$ est unipotente. D'après la proposition **??**, elle est l'exponentielle d'un logarithme, plus précisément,

$$I_p + D^{-1}N = \exp\left(\sum_{k=1}^{p} \frac{(-1)^{k-1}}{k}(D^{-1}N)^k\right).$$

Ce logarithme est une somme finie, c'est un polynôme en $D^{-1}N$, donc un polynôme en M car D^{-1} et N le sont aussi. $I_p + D^{-1}N$ est donc un polynôme en M, ce qui donne l'existence du polynôme V et achève la preuve du lemme. □

Preuve du théorème:

- Si $M = \exp(T)$, avec T réelle, M est inversible et $M = (\exp(T/2))^2$ donc est un carré.

- Passons à la réciproque. On suppose que $M = R^2$ avec $R \in GL_p(\mathbb{R})$.
On applique le lemme à R, il existe un polynôme complexe P tel que $R = \exp(P(R))$. Puisque R est réelle, en passant au conjugué, on a aussi $R = \exp(\overline{P}(R))$.
$P(R)$ et $\overline{P}(R)$ commutent, donc $R^2 = \exp(P(R) + \overline{P}(R))$, ce qui montre que $M = R^2$ est l'exponentielle de la matrice réelle $P(R) + \overline{P}(R)$. □

Remarque: Contrairement à la dimension 1, exp **n'est pas injective** sur $M_p(\mathbb{R})$ pour $p \neq 1$.
Par exemple pour $p = 2$, trouvons une matrice réelle qui a pour valeurs propres $2i\pi$ et $-2i\pi$. Le polynôme caractéristique vaut alors $X^2 + 4\pi^2$, il n'y a qu'à prendre sa matrice compagnon $A = \begin{pmatrix} 0 & -4\pi^2 \\ 1 & 0 \end{pmatrix}$. A est diagonalisable sur $\mathbb{C}$, semblable à $\text{diag}(2i\pi, -2i\pi)$ donc son exponentielle est semblable à $\text{diag}(\exp(2i\pi), \exp(-2i\pi)) = I_2$. On a donc

$$\exp(A) = \exp(0) = I_2,$$

ce qui montre qu'il n'y a pas injectivité.

2 Preuve n°2: raffinement de la surjectivité de l'exponentielle complexe

Cet exercice propose une jolie preuve du lemme 1.1.

1. En dimension 1
   (a) Montrer que $H = \exp(\mathbb{C})$ est un sous-groupe ouvert de $\mathbb{C}^*$.
   (b) Montrer que H est aussi fermé dans $\mathbb{C}^*$, conclure que $\exp : \mathbb{C} \to \mathbb{C}^*$ est surjective.
2. En dimension quelconque
   Soit $A \in M_n(\mathbb{C})$. On veut montrer qu'il existe $P \in \mathbb{C}[X]$ tel que $A = \exp(P(A))$.
   (a) $\exp(M_n(\mathbb{C}))$ est-il un sous-groupe de $GL_n(\mathbb{C})$?
   (b) On note $\mathbb{C}[A]^* = \mathbb{C}[A] \cap GL_n(\mathbb{C})$. Justifier que $\exp : \mathbb{C}[A] \to \mathbb{C}[A]^*$ est bien définie et que $\mathbb{C}[A]^*$ est un ouvert de $\mathbb{C}[A]$.
   (c) Montrer que exp est de classe C^1 sur $M_n(\mathbb{C})$.
   (d) Montrer que $H = \exp(\mathbb{C}[A])$ est un sous-groupe ouvert de $\mathbb{C}[A]^*$ puis conclure.

Commentaires. Nous reprenons ici la jolie preuve de Pommelet [Po] de la surjectivité de $\exp : \mathbb{C} \to \mathbb{C}^*$. Si on essaye de la généraliser à $GL_n(\mathbb{C})$, ce qui bloque c'est le défaut de commutativité, pour pallier à ce problème, l'idée est de restreindre la source à un ensemble de matrices qui commutent, on choisit alors naturellement $\mathbb{C}[A]$.
Cet exercice met en jeu topologie, calcul différentiel et "sans le dire" un petit résultat sur les groupes topologiques (groupe muni d'une topologie séparée pour laquelle la multiplication et l'inverse sont continues).
Plus précisément, on verra que si H est un sous-groupe de G, H est ouvert dans G si, et seulement si, l'élément neutre est intérieur à H . De plus, si H est ouvert, H est aussi fermé. Donc si G est connexe...

Corrigé.

1. (a) Pour tout $a, b \in \mathbb{C}$, $\exp(a+b) = \exp(a)\exp(b)$. La fonction $\exp : \mathbb{C} \to \mathbb{C}^*$ est donc un morphisme de groupes, par suite $\exp(\mathbb{C})$ est un sous-groupe de $\mathbb{C}^*$.
   Montrons que 1 est intérieur à H. exp est holomorphe de dérivée elle même, elle donc de classe C^1 si on la regarde comme une fonction de deux variables réelles, sa différentielle en 0 est la multiplication par $\exp(0) = 1$, c'est donc l'identité sur $\mathbb{R}^2$. D'après le théorème d'inversion locale, exp est donc localement inversible, il existe un voisinage ouvert V de 1 dont tous les éléments sont des exponentielles, donc $1 \in V \subset H$, ce qui montre montre que 1 est intérieur à H.
   Si $a \in H$, aV est un voisinage de a, ouvert car image de V par l'homéomorphisme $h \mapsto ah$ (sa réciproque est $h \mapsto a^{-1}h$), et inclus dans H car H est stable par multiplication. H est donc ouvert.

   (b) On partitionne G en ses classes modulo H ($x \sim y \Leftrightarrow xy^{-1} \in H$). Chacune de ses classes bH est ouverte comme image de l'ouvert H par $h \mapsto bh$. Le complémentaire de H dans $\mathbb{C}^*$ est donc la réunion des ouverts bH avec $b \notin H$, c'est donc une partie ouverte de $\mathbb{C}^*$, ce qui donne H fermé dans $\mathbb{C}^*$.
   Concluons: H est à la fois fermé et ouvert dans $\mathbb{C}^*$ qui est connexe, comme H n'est pas vide, $H = \mathbb{C}^*$.

2. (a) Dès que $n \geq 2$, comme deux matrices ne commutent pas forcément, on n'a plus $\exp(A+B) = \exp(A)\exp(B)$,
   $\exp : M_n(\mathbb{C}) \to GL_n(\mathbb{C})$ n'est donc plus un morphisme! On ne peut donc pas généraliser ainsi la preuve. Il nous faut de la commutativité. On pense alors aux algèbres de polynômes...

(b) Si $M = P(A)$ où P est un polynôme,

$$\exp(M) = \lim_{n\to+\infty} \sum_{k=0}^{n} \frac{(P(A))^n}{n!}.$$

Comme $\mathbb{C}[A]$ est une $\mathbb{C}$-algèbre de dimension finie, (voir chapitre sur les espaces stables) c'est donc une partie fermée de $M_n(\mathbb{C})$. Par suite $\exp(M)$ est donc un polynôme en A, puisque limite d'une suite de polynômes en A. On sait en plus que $\exp(M)$ est inversible.
$\mathbb{C}[A]^* = \{M \in \mathbb{C}[A], \det M \neq 0\}$, donc est une partie ouverte de $\mathbb{C}[A]$ comme image réciproque de l'ouvert $\mathbb{C}^*$ par l'application continue det (car polynomiale).

(c) Une méthode simple, consiste à montrer que la série des différentielles de $A \mapsto \sum_{n\geq 0} \frac{A^n}{n!}$ converge uniformément sur les compacts de $M_n(\mathbb{C})$; on pourra alors légitimement dériver terme à terme.
On choisit sur $M_n(\mathbb{C})$ une norme $||\ ||$ sous-multiplicative, c'est à dire telle que $||AB|| \leq ||A||\,||B||$.
Pour tout entier $k \geq 1$, l'application $X \mapsto X^k$ est de classe C^1 sur $M_n(\mathbb{C})$ car polynomiale en les coefficients de X. Pour calculer sa différentielle en X, on isole les termes linéaires en H obtenus en développant $(X+H)^k$. Les autres termes (en nombre fini) ont au moins deux fois le "facteur" H, ce sont donc des $O(||H||^2)$.

$$(X+H)^k = (X+H)...(X+H) = X^k + HX^{k-1} + XHX^{k-2} + ... + X^{k-1}H + O(||H||^2).$$

On en déduit que $f_k : X \mapsto \frac{X^k}{k!}$ est de clase C^1 sur $M_n(\mathbb{C})$ et sa différentielle en X est, pour $k \geq 1$, l'application linéaire de $M_n(\mathbb{C})$ dans $M_n(\mathbb{C})$ définie par

$$\mathrm{d}f_k(X).H = \frac{1}{k!}(HX^{k-1} + XHX^{k-2} + ... + X^{k-1}H).$$

On note $\mathcal{L}$ l'espace vectoriel des applications linéaires de $M_n(\mathbb{C})$ dans $M_n(\mathbb{C})$. On le munit de la norme $\mathcal{N}$ subordonnée à $||\ ||$, ce qui lui confère une structure d'espace complet. Il s'agit de montrer que $\sum_k \mathrm{d}f_k$ converge uniformément sur les compacts de $M_n(\mathbb{C})$.
$||\mathrm{d}f_k(X).H|| \leq \frac{1}{k!}k||X||^{k-1}||H||$, donc $\mathcal{N}(\mathrm{d}f_k(X)) \leq \frac{||X||^{k-1}}{(k-1)!}$.
La série $\sum_{n\geq 0} \mathrm{d}f_k(X)$ est donc uniformément convergente sur les boules $\{X \in M_n(\mathbb{C}),\ ||X|| \leq R\}$ de $M_n(\mathbb{C})$, la série $\sum_{k\geq 0} f_k$ converge simplement vers exp, on en déduit que exp est de classe C^1 sur ces boules, donc sur $M_n(\mathbb{C})$.
Remarques:

- Vous pouvez regarder Rouvière [Ro] exercice 38, pour une preuve du théorème de dérivation terme à terme.
- On peut montrer que exp est de classe C^∞, on pourra consulter Lafontaine [Laf] p 36 pour une preuve différente à l'aide d'analyse complexe.
- Attention, on ne peut justifier que exp est de classe C^1 par l'argument $\exp A$ est un polynôme en A. En effet, les coefficients du polynôme dépendent de A... Avec ce même raisonnement, on comettrait l'erreur de dire que l'application qui à une matrice associe son polynôme minimal est continue... Ceci est faux, la suite $A_n = \begin{pmatrix} 0 & \frac{1}{n} \\ 0 & 0 \end{pmatrix}$ tend vers la matrice nulle. Pour tout n, le polynôme minimal de A_n est X^2, il ne peut donc tendre vers X le polynôme minimal de la matrice nulle.

(d) A partir de là, c'est facile, il n'y a plus qu'à imiter la preuve de la question 1.
Comme tous les polynômes en A, commutent entre eux, $\exp : \mathbb{C}[A] \to \mathbb{C}[A]^*$ est un morphisme de groupes, et $H = \exp(\mathbb{C}[A])$ est un sous-groupe de $\mathbb{C}[A]^*$.
exp est de classe C^1, sa différentielle en 0 est l'identité donc est inversible;

$$\exp(H) = I_n + H + O(||H||^2).$$

On en déduit l'existence d'un voisinage ouvert $\mathcal{V}$ tel que $I_n \in \mathcal{V} \subset H$. Pour tout $M \in H$, $M\mathcal{V}$ est un voisinage ouvert de M dans H, car $B \mapsto MB$ est un homéomorphisme sur

$GL_n(\mathbb{C})$ (c'est une application linéaire), d'inverse $B \mapsto M^{-1}B$).
H est donc un sous-groupe ouvert de $\mathbb{C}[A]^*$, il est donc aussi fermé (même preuve qu'en 1.b).
Si M et N sont dans $\mathbb{C}[A]^*$, la fonction polynomiale de $\mathbb{C}$ dans $\mathbb{C}$

$$z \mapsto \det\left((1-z)M + zN\right)$$

n'admet qu'un nombre fini de zéros et ne s'annule ni en 0, ni en 1, donc il existe un chemin continue γ joignant 0 à 1 qui évite ces zéros. Alors

$$t \mapsto \left(1-\gamma(t)\right)M + \gamma(t)N$$

joint M et N et est à valeurs dans $\mathbb{C}[A]^*$, ce qui prouve que $\mathbb{C}[A]^*$ est connexe par arcs. Finalement, H non vide est ouvert et fermé dans le connexe $\mathbb{C}[A]^*$, donc $H = \mathbb{C}[A]^*$, ce qui achève la preuve.

Remarque: $GL_n(\mathbb{C})$ est un groupe topologique, sa topologie provient de l'espace vectoriel normé $M_n(\mathbb{C})$, et les applications $(A,B) \mapsto AB$ et $A \mapsto A^{-1}$ sont continues (comme $A^{-1} = \frac{{}^t\mathrm{com}A}{\det A}$, l'application inverse est une fraction rationnelle en les coefficients de A) .

3 Preuve n°3: selon quelques indications de Mneimé dans réduction des endomorphismes

Lemme 3.1 *Si A est une matrice réelle qui possède un polynôme caractéristique sans racines réelles, alors A est l'exponentielle d'une matrice réelle.*

Preuve:
χ_A est un produit de puissances de polynômes irréductibles sur $\mathbb{R}$ de degré 2, grâce au théorème de décomposition des noyaux, il suffit de prouver le résulat pour $A \in M_{2n}(\mathbb{R})$ avec $\chi_A = \left(X^2 - 2\Re(\lambda)X + |\lambda|^2\right)^n$ et λ non réel. On a $\chi_A = (X-\lambda)^n(X-\overline{\lambda})^n$. On note respectivement $\mathcal{C}$ et $\overline{\mathcal{C}}$ les sous-espaces caractéristiques de A associés à λ et $\overline{\lambda}$. Ils sont de dimension n.
Si $(X_1,\ldots,X_n)$ est une base de $\mathcal{C}$, alors $(\overline{X}_1,\ldots,\overline{X}_n)$ est une base de $\overline{\mathcal{C}}$. En effet, on vérifie facilement que $X \in \ker(A-\lambda I_{2n}) \Leftrightarrow \overline{X} \in \ker\left(A-\overline{\lambda}I_{2n}\right)$ et que $(\overline{X}_1,\ldots,\overline{X}_n)$ est une famille libre.
Comme $\mathcal{C}$ est stable par A, il existe des nombres complexes v_{ij} tels que pour tout $j \in \{1,\ldots,n\}$, $AX_j = \sum_{i=1}^n v_{ij}X_i$, ce qui donne $A\overline{X_j} = \sum_{i=1}^n \overline{v_{ij}}\overline{X_i}$.
On peut donc en déduire que A est semblable sur $\mathbb{C}$ à $\mathrm{diag}(V,\overline{V})$ où $V = (v_{ij}) \in GL_n(\mathbb{C})$ (sinon 0 serait valeur propre de A).
Comme $\exp : M_n(\mathbb{C}) \to GL_n(\mathbb{C})$ est surjective, il existe $U \in M_n(\mathbb{C})$ tel que $V = \exp U$.
A est donc semblable à $\exp\left(\mathrm{diag}(U,\overline{U})\right)$. Or la matrice $\mathrm{diag}(U,\overline{U})$ est semblable à une matrice réelle (pourquoi ?) (en dimension 2, on montre que $\mathrm{diag}(\lambda,\overline{\lambda})$ est semblable à la matrice réelle associée à la similitude de rapport λ, pour cela on montre qu'elles ont les mêmes polynômes caractéristiques et minimaux, donc sont semblables. En dimension supérieure à 4, ce n'est plus vrai, et il faut certainement utiliser des invariants de similitude...)
Finallement, A est semblable sur $\mathbb{C}$ à l'exponentielle d'une matrice réelle, comme A est aussi réelle, elles sont en fait semblables sur $\mathbb{R}$, ce qui permet de conclure. □.

Preuve du théorème principal: soit $M = R^2$ avec $R \in GL_p(\mathbb{R})$.
Cas 1: on suppose que les valeurs propres réelles de R sont strictement positives. Soit a l'une d'elles, on note $\mathcal{C} = \ker(R-aI_p)^\alpha$, l'espace caractéristique associé. Dans une base de $\mathcal{C}$, la matrice de $R|_{\mathcal{C}}$ s'écrit $aI_\alpha + N$ avec N nilpotente.
En utilisant une base associée à la décomposition en facteurs irréductibles de χ_R, on obtient que R est semblable sur $\mathbb{R}$ à une matrice diagonale par blocs où les blocs sont de la forme:

- $aI_\alpha + N$ avec N nilpotente et $a > 0$
- S où S est une matrice réelle sans valeur propre réelle.

Or $aI_\alpha + N = a(I_\alpha + \frac{N}{a})$ est donc une exponentielle car $a = e^{\ln a}$ et $I_\alpha + \frac{N}{a}$ est une exponentielle puisqu'elle est unipotente.
Ensuite d'après le lemme, S est aussi une exponentielle.
On en déduit que R est semblable sur $\mathbb{R}$ à une diagonale d'exponentielles de matrices, ie

$$R^2 = P\text{diag}\left(\exp(S_1), \ldots, \exp(S_k)\right) P^{-1} = P \exp\left(\text{diag}(S_1, \ldots, S_k)\right) P^{-1} = \exp\left(P\text{diag}(S_1, \ldots, S_k)P^{-1}\right),$$

R est donc une exponentielle, R^2 aussi, ce qui permet de conclure.
Cas 2: R possède des valeurs propres réelles négatives. On note alors R' la matrice obtenue en multipliant par -1 les **composantes caractéristiques** de R, c'est à dire que si $aI_\alpha + N$ est un bloc de R avec $a < 0$, on remplace ce bloc par son opposé $-aI_\alpha - N$. On a alors par construction $R'^2 = R^2 = M$, et donc cette fois-ci R' est à valeurs propres réelles positives, on est donc ramené au cas 1, ce qui achève la preuve. □