La décomposition de Jordan-Chevalley

Je propose deux niveaux de lecture pour ce qui suit : un niveau plutôt abstrait pour des algèbres sur un corps relativement quelconque (la note est écrite dans cet esprit), et un niveau plus concret pour l'algèbre des matrices sur $\mathbb{R}$ ou $\mathbb{C}$. Les deux niveaux sont également profitables pour l'agrégation.

1 Énoncé du théorème

Le théorème qui affirme que toute matrice $M \in \mathsf{M}_n(\mathbb{C})$ est somme d'une matrice diagonalisable et d'une matrice nilpotente qui commutent est souvent appelé *décomposition de Dunford* dans l'enseignement français. Les raisons de cette dénomination sont étranges ; si Nelson Dunford a démontré un théorème de ce genre, c'était certainement postérieur aux travaux pionniers de Jordan (vers 1870) et à ceux de Chevalley (vers 1950) qui ont donné sa forme moderne au théorème. La dénomination anglo-saxonne de *décomposition de Jordan-Chevalley* est bien plus fidèle à l'histoire, et le très intéressant article [CEZ] l'explique clairement. Nous utiliserons donc cette dernière terminologie. Nous allons présenter une version très générale du théorème, dont la preuve, due à Chevalley, est basée sur la méthode de Newton d'approximation des solutions d'une équation. Elle présente l'avantage de ne pas nécessiter le calcul des valeurs propres de la matrice ; elle est entièrement effective. Pour énoncer le théorème, nous rappelons quelques définitions.

1.1 Définitions. (1) Une algèbre (associative, unitaire) sur un corps k est un k-espace vectoriel A muni d'une structure d'anneau (associatif, unitaire) dont la multiplication est bilinéaire.

(2) Un élément $a \in A$ est *algébrique* si le morphisme d'évaluation $\mathsf{ev}_a : k[X] \to A$, $P \mapsto P(a)$ n'est pas injectif. On appelle alors *polynôme minimal de a* noté μ_a le générateur unitaire de $\ker(\mathsf{ev}_a)$.

(3) On dit qu'un élément algébrique a est *nilpotent* si $\mu_a = X^r$ pour un certain $r \geqslant 1$, *unipotent* si $a - 1$ est nilpotent, et *semi-simple* si μ_a est sans facteur carré.

1.2 Exercice. La donnée d'une structure de k-algèbre (associative, unitaire) sur A est équivalente à la donnée d'une structure d'anneau (associatif, unitaire) et d'un morphisme d'anneaux $k \to A$ dont l'image est incluse dans le centre de A.

Rappelons qu'un corps k est dit *parfait* s'il est de caractéristique 0, ou s'il est de caractéristique $p > 0$ et son endomorphisme de Frobenius $\mathrm{Fr} : k \to k$, $x \mapsto x^p$ est surjectif (donc bijectif). Les corps finis ou algébriquement clos sont parfaits. Si k_0 est de caractéristique p, le corps de fractions rationnelles $k = k_0(X)$ n'est pas parfait, car X n'est pas une puissance p-ième (c'est un élément de degré 1 alors que les puissances p-ièmes sont de degré multiple de p).

1.3 Théorème. *Soit k un corps parfait et A une k-algèbre de dimension finie. Alors pour tout $a \in A$ il existe un unique couple (s, n) dans A tel que $a = s + n$, s est semi-simple, n est nilpotent, et $sn = ns$. De plus s et n sont des polynômes en a, et μ_s divise μ_a.*

1.4 Remarques. (1) Supposons que $k = \mathbb{C}$ et $A = \mathsf{M}_n(\mathbb{C})$. Un polynôme est sans facteur carré si et seulement s'il est à racines simples, donc les notions « semi-simple » et « diagonalisable » coïncident. On voit que le théorème ci-dessus redonne la décomposition $D + N$ habituelle.

(2) Supposons que $k = \mathbb{R}$. Soit $M \in \mathsf{M}_n(\mathbb{R})$. Si on croit à la décomposition $D + N$ sur $\mathbb{C}$, on peut écrire $M = S + N$ avec $S, N \in \mathsf{M}_n(\mathbb{C})$. En prenant les conjuguées complexes, on a alors $\overline{S} + \overline{N} = \overline{M} = M = S + N$. Par unicité des parties diagonalisable et nilpotente, on voit que $\overline{S} = S$ et $\overline{N} = N$, i.e. $S, N \in \mathsf{M}_n(\mathbb{R})$. La matrice S est diagonalisable sur $\mathbb{C}$ donc semi-simple sur $\mathbb{R}$. Ce cas particulier facile se trouve dans [BMP], application 4.32 et il est utile par exemple pour décrire l'image de l'exponentielle matricielle réelle, cf [BMP], exercice 4.17.

On a une version multiplicative de la décomposition de Jordan-Chevalley.

1.5 Théorème. *Soit k un corps parfait et A une k-algèbre de dimension finie. Pour tout $a \in A$ inversible, il existe un unique couple (s, u) dans A tel que $a = su$, s est semi-simple inversible, u est unipotent, et $su = us$. De plus s et n sont des polynômes en a, et μ_s divise μ_a.*

Ce résultat se déduit directement de la décomposition $a = s + n$. On voit que $s = a - n$ est somme d'un inversible et d'un nilpotent qui commutent, il est donc inversible. Posons $u = 1 + s^{-1}n$. Alors $a = su$ et cette décomposition possède les propriétés annoncées.

2 Représentations linéaires et semi-simplicité

2.1 L'idée des représentations. Les ensembles de bijections d'ensembles ou de transformations bijectives de certaines structures comme les espaces vectoriels sont des *groupes*. Ceci a pour conséquence qu'on peut « représenter » un groupe donné G comme groupe de bijections au moyen de morphismes $\rho : G \to \mathfrak{S}_X$ ou $\rho : G \to \mathsf{GL}(E)$. Ceci permet de mieux l'étudier, utilisant par exemple toute notre connaissance de l'algèbre linéaire. On parle de *représentations ensemblistes* ou *linéaires* de G. Dans le cas linéaire, l'espace vectoriel E est appelé l'*espace de la représentation.*

De la même manière, les ensembles d'endomorphismes de k-espaces vectoriels sont des *k-algèbres* (associatives unitaires) et on peut « représenter » une k-algèbre donnée A comme ensemble d'endomorphismes linéaires au moyen de morphismes $\rho : A \to \mathsf{End}(E)$. On parle de *représentations linéaires* de A. Pour que la théorie soit souple et complète, on autorise tous les morphismes ρ, mais les représentations d'un groupe ou d'une algèbre qui la reflètent le mieux sont celles pour lesquelles ρ est injectif, qu'on appelle représentations *fidèles.*

2.2 Exemples. Voici les deux familles les plus importantes de représentations fidèles.

(1) Représentations naturelles. Si $G \subset \mathfrak{S}_n$, il a une représentation dite *naturelle* dans l'ensemble $\{1, \ldots, n\}$ correspondant au morphisme d'inclusion donné $\rho : G \hookrightarrow \mathfrak{S}_n$. Si $G \subset \mathsf{GL}(E)$, il a une représentation naturelle dans l'espace vectoriel E. Si $A \subset \mathsf{End}(E)$, elle a une représentation naturelle dans l'espace vectoriel E.

(2) Représentations régulières (gauches).
(a) Tout groupe G peut se représenter comme groupe de bijections sur lui-même par les multiplications à gauche $\rho_g : G \to G$, $x \mapsto gx$. La représentation correspondante est la *représentation régulière* $G \hookrightarrow \mathfrak{S}_G$, c'est celle qui apparaît dans le théorème de Cauchy en théorie des groupes finis.

(b) Tout groupe G peut se représenter comme groupe d'automorphismes linéaires sur l'espace vectoriel $E = k^G = \oplus_{g\in G} k e_g$ via les morphismes $\rho_g : E \to E$, $e_h \mapsto e_{gh}$. Il s'agit de la représentation régulière $G \hookrightarrow \mathsf{GL}(k^G)$ que nous reverrons dans la théorie des représentations linéaires de groupes finis.
(c) Toute k-algèbre A se représente comme algèbre d'endomorphismes par les multiplications à gauche $\rho_a : A \to A$, $x \mapsto ax$. C'est la représentation régulière $A \hookrightarrow \mathsf{End}(A)$ d'une algèbre. On voit en particulier que toute algèbre de dimension finie est une sous-algèbre d'une algèbre de matrices, puisque $\mathsf{End}(A) \simeq \mathsf{M}_n(k)$ si $\mathsf{dim}(A) = n$.

2.3 Remarque. La représentation régulière est utile par exemple pour construire l'algèbre des quaternions. En effet, si on se souvient des relations $i^2 = j^2 = k^2 = -1$, $ij = -ji = k$, $jk = -kj = i$, $ki = -ik = j$, on connaît l'action de i et j par multiplication à gauche sur l'algèbre $\mathbb{H}$. En d'autres termes, les images de i et j par $\mathbb{H} \to \mathsf{End}(\mathbb{H}) \simeq \mathsf{M}_4(\mathbb{R})$ sont des matrices explicites. On peut donc définir $\mathbb{H}$ comme la sous-algèbre de $\mathsf{M}_4(\mathbb{R})$ engendrée par ces matrices.

2.4 Endomorphismes semi-simples. Nous nous contentons ici de rappeler l'énoncé du théorème central sur les endomorphismes semi-simples pour les replacer dans le contexte de l'algèbre linéaire. Ce théorème n'aura pas d'importance pour la suite puisque nous utiliserons exclusivement la caractérisation de la définition 1.1. Nous ne donnons pas de preuve, et nous renvoyons à [Tau] § 10.14, [BMP] th. 4.29, [Gou], problèmes du chapitre 4. Il n'y aura pas de complément de cours spécifique sur cette notion.

2.5 Théorème. *Soient k un corps parfait, A une k-algèbre de dimension finie, $\rho : A \to \mathsf{End}(E)$ une représentation linéaire fidèle de dimension finie, et $a \in A$. Les conditions suivantes sont équivalentes :*

(1) *Tout sous-espace $\rho(a)$-stable de E possède un supplémentaire stable.*

(2) *L'endomorphisme $\rho(a)$ est diagonalisable après passage à une clôture algébrique de k.*

(3) *Le polynôme minimal de a est produit de polynômes irréductibles distincts.*

D'un certain point de vue, le rôle de A est secondaire. En effet, comme ρ est injectif, on peut identifier a et $\rho(a)$ qui ont d'ailleurs mêmes polynômes minimaux. Alors le théorème donne un résultat sur un élément d'une algèbre de matrices $\mathsf{End}(E)$. D'un autre point de vue, le point (3) montre que la propriété de semi-simplicité de a ne dépend pas de la représentation linéaire fidèle E choisie.

3 Que se passe-t-il sur les corps non parfaits ?

Pour nous, les propriétés importantes des corps parfaits seront les suivantes.

3.1 Lemme. *Soit k un corps de caractéristique $p > 0$.*

(1) *Si $t \in k$ n'est pas une puissance p-ième, le polynôme $X^p - t \in k[X]$ est irréductible.*

(2) *Les conditions suivantes sont équivalentes :*

(a) *k est parfait.*

(b) *tout polynôme $P \in k[X]$ de dérivée nulle est une puissance p-ième.*

(c) *tout polynôme $P \in k[X]$ irréductible a une dérivée non nulle.*

(d) *tout polynôme $P \in k[X]$ irréductible a ses racines* simples *dans une clôture algébrique $\overline{k}/k$.*

Preuve : (1) Soit $P = X^p - t$ et notons α une racine de P dans une clôture algébrique $\overline{k}$. Soit $Q \in k[X]$ un facteur de P, unitaire de degré $i > 0$. Puisque $P = X^p - t = X^p - \alpha^p = (X-\alpha)^p$, on peut écrire $Q = (X-\alpha)^i = X^p - i\alpha X^{i-1} + \dots$. Comme Q est à coefficients dans k et $\alpha \notin k$, on doit avoir $i = 0$ dans k, c'est-à-dire $i = p$.

(2) (a) $\Rightarrow$ (b). Soit P tel que $P' = 0$. En écrivant $P = \sum a_i X^i$, on voit que les a_i avec i premier à p sont nuls. Il reste les a_i avec $i = jp$. Si on écrit $a_{jp} = b_j^p$, on obtient $P = (\sum b_j X^j)^p$.

(b) $\Rightarrow$ (c). Un polynôme de dérivée nulle est une puissance p-ième donc n'est pas irréductible.

(c) $\Rightarrow$ (d). Soit P irréductible. D'après (c) le polynôme P' est non nul, donc premier avec P. Une racine multiple de P dans $\overline{k}$ serait aussi racine de P', donc ne peut exister.

(d) $\Rightarrow$ (a). Si $t \in k$ n'est pas une puissance p-ième, d'après (1) le polynôme $P = X^p - t$ est irréductible et possède une racine p-uple dans $\overline{k}$. □

3.2 Un contre-exemple. Sur un corps k non parfait, les problèmes apparaissent pour les endomorphismes dont le polynôme minimal contient des facteurs irréductibles à dérivée nulle. L'exemple typique d'un tel polynôme est $P = X^p - t$ avec un terme constant t qui n'est pas une puissance p-ième dans k. Si a est un élément d'une algèbre A avec polynôme minimal $\mu_a = P$, il est semi-simple (voir 2.5) donc admet une décomposition $s + n$ avec $s = a$ et $n = 0$. Pour obtenir un exemple où la décomposition de Jordan-Chevalley n'existe pas, nous devons supposer que $P^2 \mid \mu_a$, par exemple $\mu_a = P^2$. On peut construire un tel a à l'aide d'une matrice compagnon, ce qui revient à prendre $A = k[X]/(P^2)$ et a égal à la classe de X agissant par multiplication sur A. Plaçons-nous dans cette situation. Supposons que l'on puisse écrire $a = s + n$ avec s semi-simple, n nilpotent, et $sn = ns$. On note pour la suite que les éléments nilpotents de l'algèbre A sont les multiples de $P(a)$, et ils sont tous de carré nul puisque $P(a)^2 = 0$. Comme $n^2 = 0$, on a :

$$0 = \mu_a(a) = \mu_a(s+n) = \mu_a(s) + n\mu_a'(s).$$

Il en découle que $\mu_a(s) = -n\mu_a'(s)$ est nilpotent, donc $\mu_a^2(s) = P^4(s) = 0$. On en déduit que $\mu_s | P^4$. Comme s est semi-simple, on trouve $\mu_s = P$ donc $P(s) = s^p - t = 0$. Alors $a^p = s^p + n^p = s^p = t$ donc P annule a, ce qui est impossible.

4 Preuve de la décomposition de Jordan-Chevalley

La seule référence disponible pour ce qui suit est [RB], problème 3.1, corrigé en fin d'ouvrage. Ces auteurs se placent dans $\mathsf{M}_n(\mathbb{C})$, ce qui n'enlève rien à l'intérêt du développement. Nous adoptons autant que possible des notations proches des leurs.

4.1 Unicité. Supposons que $a = s + n = s' + n'$. On peut plonger A dans $\mathsf{End}(A)$ par la représentation régulière. En choisissant une base de A, on peut identifier $\mathsf{End}(A)$ avec une algèbre de matrices $\mathsf{M}_n(k)$ puis la plonger dans $\mathsf{M}_n(\overline{k})$. Alors s, s' sont diagonalisables. La preuve classique de l'unicité s'applique : on montre que s et s' commutent donc $s - s'$ est diagonalisable, puis que n et n' commutent donc $n' - n$ est nilpotente, et finalement $s - s' = n' - n$ est nulle.

4.2 Idée de la preuve. Par condition nécessaire, si s est semi-simple son polynôme minimal μ_s est sans facteur carré. Si de plus μ_s divise μ_a, l'élément s est annulé par le polynôme P produit des facteurs irréductibles distincts de μ_a, qu'on appelle le *radical* de μ_a. L'idée de la preuve est de construire s comme une racine de l'équation $P(s) = 0$ à l'aide de la méthode d'approximation de Newton, en partant de $a_0 = a$. La situation est particulièrement favorable pour deux raisons :

— le point de départ de l'algorithme $a = s + n$ est très proche de la racine cherchée s, puisque les nilpotents sont les infinitésimaux de l'algèbre ;

— comme k est parfait, on a $P' \neq 0$ premier avec P donc $P'(a) \neq 0$.

4.3 La preuve.

4.3.1. Cadre. Tous les calculs vont se passer dans $A_0 = k[a] \simeq k[X]/(\mu_a)$, la sous-algèbre engendrée par a. Quitte à remplacer A par A_0, on suppose donc que $A = A_0$. En particulier A est commutative et engendrée par a. On note $\mu = \mu_a$ et $p = \mathrm{car}(k) \geqslant 0$.

4.3.2. Fonction à laquelle on applique l'algorithme. Soit P le produit des facteurs irréductibles distincts de μ, aussi appelé *radical* de μ. D'un point de vue théorique, on peut écrire $P = P_1 \dots P_s$ si $\mu = P_1^{r_1} \dots P_s^{r_s}$ est la décomposition en produit d'irréductibles de μ. Le calcul est cependant effectif :

— si $p = 0$, ou plus généralement si les multiplicités r_i sont premières à p, on a l'expression explicite $P = \mu/(\mu, \mu')$ où le pgcd se calcule par l'algorithme d'Euclide.

— en général, le calcul est effectif grâce à l'algorithme décrit dans 4.4 ci-dessous.

4.3.3. Valeur intiale de l'algorithme : estimation de $P(a)$ et $P'(a)$. Par construction, il existe un entier r tel que μ divise P^r. On peut prendre $r = \max(r_1, \dots, r_s)$ si on a des informations sur les multiplicités, et en tout cas $r = \deg(\mu)$ convient. En particulier $P(a)^r = 0$ i.e. $\epsilon := P(a)$ est nilpotent d'indice $\leqslant r$. Dans la suite, nous noterons $x = \mathrm{O}(\epsilon^n)$ pour dire qu'un élément x appartient à l'idéal engendré par ϵ^n. Comme k est parfait, les polynômes P_i sont tous à racines simples dans $\overline{k}$. En particulier μ et P' sont premiers entre eux. En prenant une relation de Bézout, on voit que $P'(a)$ est inversible dans A.

4.3.4. La méthode de Newton. On considère ensuite la suite définie par récurrence :

$$\begin{cases} a_0 = a \\ a_{n+1} = a_n - \frac{P(a_n)}{P'(a_n)}. \end{cases}$$

Montrons par récurrence que (i) a_n est bien défini (ii) $P(a_n) = \mathrm{O}(\epsilon^{2^n})$ et (iii) $a_n - a = \mathrm{O}(\epsilon)$. Pour $n = 0$, ces trois propriétés sont évidentes. Pour $n \geqslant 1$, du fait que $a_n = a + \mathrm{O}(\epsilon)$ on déduit que $P'(a_n) = P'(a) + \mathrm{O}(\epsilon)$. Cet élément est somme d'un inversible et d'un nilpotent, il est donc inversible. Il s'ensuit que a_{n+1} est bien défini ce qui établit (i). Maintenant notons $P(X + Y) = P(X) + YP'(X) + Y^2Q(X, Y)$ où Q est un certain polynôme. On calcule alors

$$P(a_{n+1}) = P\left(a_n - \frac{P(a_n)}{P'(a_n)}\right) = P(a_n) - \frac{P(a_n)}{P'(a_n)}P'(a_n) + \left(\frac{-P(a_n)}{P'(a_n)}\right)^2 Q\left(a_n, \frac{-P(a_n)}{P'(a_n)}\right) = \mathrm{O}(\epsilon^{2^{n+1}})$$

ce qui établit (ii). Enfin $a_{n+1} - a = a_n - \frac{P(a_n)}{P'(a_n)} - a = \mathrm{O}(\epsilon)$ ce qui établit (iii). Ceci conclut la preuve des trois propriétés.

4.3.5. Stationnarité. Lorsque $n = \lceil \log_2(r) \rceil$, on a $\epsilon^{2^n} = 0$ donc $P(a_n) = 0$ et la suite stationne à $s = a_\infty = a_n$. Cet élément est une racine de P donc semi-simple. De plus, $a - s = a - a_n = \mathrm{O}(\epsilon)$ est nilpotent. On a obtenu la décomposition $a = s + n$ recherchée.

4.4 Calcul du radical d'un polynôme sur un corps parfait. Dans la mise en place de la méthode de Newton, nous avons affirmé que le calcul du radical d'un polynôme à coefficients dans un corps parfait est effectif. Nous présentons un algorithme (tiré de [BCGLLSS], chap. 32, § 3) qui réalise ce calcul dans un corps parfait k où le calcul de la racine p-ième dans k est effectif. Par exemple, si k est un corps fini de cardinal $q = p^d$, on a $\mathrm{Fr}^d = \mathsf{Id}_k$ donc la racine p-ième est l'itéré $(d-1)$-uple du morphisme de Frobenius. Pour un polynôme à coefficients dans k de dérivée nulle, le calcul de la racine p-ième est alors effectif également puisqu'il se fait coefficient par coefficient.

Lemme. *L'algorithme* rad *suivant calcule le radical d'un polynôme non nul $P \in k[X]$:*

Si $\deg(P) = 0$ *faire* $R := 1$ *et sortir*
Si $P' = 0$ *faire* $R := \mathrm{rad}(P^{1/p})$
Si $P' \neq 0$ *faire* $U := P \wedge P'$
$V := P/U$
$n := \deg(P)$
$W := U \wedge V^{n-1}$
$R := V * \mathrm{rad}(U/W)$
Afficher R.

Preuve : Écrivons la décomposition de P en facteurs irréductibles $P = \prod_{i\in I} Q_i^{\alpha_i} = EF$, où E regroupe les facteurs de multiplicités α_i premières à p, et F les facteurs de multiplicités α_i multiples de p. Appliquons l'algorithme à P. Les éventualités $\deg(P) = 0$ ou $P' = 0$ sont faciles. Supposons donc que $P' \neq 0$. Lorsqu'on dérive l'expression $P = EF$, le facteur F a une dérivée nulle et se comporte donc comme une constante. On obtient $P' = E'F$ et $P \wedge P' = (E \wedge E')F$. Le calcul de $E \wedge E'$ est facile : si Q^α est l'un des facteurs qui apparaissent dans E, on a $E = Q^\alpha G$ d'où $E' = Q^{\alpha-1}(\alpha Q'G + QG')$. On voit que $Q^{\alpha-1} \mid E'$ mais $Q^\alpha \nmid E'$ car $\alpha \not\equiv 0 \mod p$. Finalement :

$$U = P \wedge P' = \prod_{p\nmid\alpha_i} Q_i^{\alpha_i-1} \cdot \prod_{p\mid\alpha_i} Q_i^{\alpha_i}.$$

Ensuite $V = P/U = \prod_{p\nmid\alpha_i} Q_i$ donc $V^{n-1} = \prod_{p\nmid\alpha_i} Q_i^{n-1}$. Comme $\alpha_i \leqslant n$ pour tout i, on obtient $W = U \wedge V^{n-1} = \prod_{p\nmid\alpha_i} Q_i^{\alpha_i-1}$ et $U/W = \prod_{p\mid\alpha_i} Q_i^{\alpha_i}$. Ici n'apparaissent plus que les multiplicités multiples de p, et $\mathrm{rad}(P) = \mathrm{rad}(V)\,\mathrm{rad}(U/W)$. Comme les degrés de V et U/W sont strictement inférieurs au degré de P, l'algorithme se termine. □

L'exponentielle dans les algèbres de dimension finie

5 L'exponentielle dans une algèbre de dimension finie

Soit k l'un des corps $\mathbb{R}$ ou $\mathbb{C}$ (l'hypothèse utile est d'avoir un corps normé complet, par exemple le corps des nombres p-adiques $\mathbb{Q}_p$ conviendrait). Soit A une k-algèbre (associative et unitaire) de dimension finie. Pour tout choix d'une norme de k-espace vectoriel $\|\cdot\|$ sur A, la formule

$$N(a) = \sup_{\|x\|=1} \|ax\|$$

définit une norme d'algèbre, i.e. une norme d'espace vectoriel sous-multiplicative. Il s'ensuit que pour tout $x \in A$ et tout entier N, on a l'inégalité :

$$\left\| \sum_{n=0}^{N} \frac{1}{n!} x^n \right\| \leqslant \sum_{n=0}^{N} \frac{1}{n!} \|x\|^n \leqslant e^{\|x\|},$$

donc la série $\sum_{n\geqslant 0} \frac{1}{n!} x^n$ est normalement convergente.

5.1 Définition. Pour $x \in A$, on appelle *exponentielle de x dans A* et on note $\exp_A(x)$ la somme de la série $\sum_{n\geqslant 0} \frac{1}{n!} x^n$.

5.2 Remarque. Comme nous sommes en dimension finie, le sous-espace vectoriel $k[x]$ des polynômes en x est fermé. Il s'ensuit que $\exp_A(x)$, qui est la limite de la suite $\sum_{n=0}^{N} \frac{1}{n!} x^n$ d'éléments de $k[x]$, est un élement de $k[x]$. Autrement dit $\exp_A(x)$ est un polynôme en x.

5.3 Proposition. *Soit $f : A \to B$ un morphisme de k-algèbres. Alors, pour tout $x \in A$, on a $f(\exp_A(x)) = \exp_B(f(x))$.*

Preuve : Comme f est un morphisme, on a $f(\sum_{n=0}^{N} \frac{1}{n!} x^n) = \sum_{n=0}^{N} \frac{1}{n!} f(x)^n$. Or f est une application linéaire entre deux espaces vectoriels de dimension finie, elle est donc continue. En passant à la limite sur N, on obtient l'égalité désirée. □

5.4 Remarque. En utilisant la représentation régulière (gauche) $A \hookrightarrow \mathrm{End}_k(A) \simeq \mathrm{M}_n(k)$ où $n = \dim(A)$, d'après la proposition on aura :

$$\exp_A = \exp_{\mathrm{End}_k(A)|A}.$$

Ceci montre que si on le souhaite, on peut toujours se placer dans une algèbre de matrices pour calculer une exponentielle. En fait, souvent c'est la démarche contraire qui sera payante : on sera dans une algèbre de matrices et on calculera l'exponentielle dans une algèbre plus petite. C'est exactement la même chose que ce qui se passe avec les groupes finis, qui peuvent tous être plongés dans un groupe symétrique via le théorème de Cauchy, ce que l'on fait rarement en pratique.

5.5 Exemple. (1) Soit A une $\mathbb{R}$-algèbre de dimension finie et $I \in A$ tel que $I^2 = -1$. (Ici $1 = 1_A$ est le neutre multiplicatif de l'algèbre). Alors le polynôme minimal de I sur $\mathbb{R}$ est X^2+1, et le morphisme de k-algèbres $f_0 : \mathbb{R}[X] \to A$ défini par $f_0(X) = I$ se factorise en un morphisme $f : \mathbb{R}[X]/(X^2+1) \simeq \mathbb{C} \to A$. Dit d'une autre manière, la sous-$\mathbb{R}$-algèbre de A engendrée par I est $\mathbb{R} \oplus \mathbb{R}I$, isomorphe à $\mathbb{C}$. En on déduit que pour tout $t \in \mathbb{R}$:

$$\exp_A(tI) = \exp_A(f(ti)) = f(\exp_{\mathbb{C}}(ti)) = f(\cos(t) + i\sin(t)) = \cos(t) + I\sin(t)$$

où l'on a noté $\cos(t)$ au lieu de $\cos(t)1$.

(2) On peut prendre $A = \mathsf{M}_2(\mathbb{R})$ et $I = \left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$ (ou son opposée). Le résultat s'écrit :

$$\exp\begin{pmatrix} 0 & -t \\ t & 0 \end{pmatrix} = \begin{pmatrix} \cos(t) & -\sin(t) \\ \sin(t) & \cos(t) \end{pmatrix}.$$

C'est la fameuse matrice de rotation R_t, qui est donc une exponentielle. Ce calcul est parfois fait « à la main » dans les livres, en calculant les puissances de tI, en regroupant les puissances paires et en reconnaissant la série du cosinus, etc. Bien sûr c'est plus lourd, mais surtout cela revient à refaire le calcul de $\exp(it)$ dans $\mathbb{C}$!

(3) On peut prendre $A = \mathbb{H}$ et $I = r \in \mathbb{H}$ un quaternion imaginaire pur de norme 1. En effet, si r est imaginaire pur, on a $\overline{r} = -r$ et s'il est de norme 1, on a $r\overline{r} = 1$. On en déduit que $r^2 = -1$. Pour un quaternion quelconque $q \in \mathbb{H}$, notons $a \in \mathbb{R}$ sa partie réelle et t la norme de $q - a$. On obtient $q = a + tr$ avec r imaginaire pur de norme 1, donc $\exp(q) = \exp(a + rt) = e^a(\cos(t) + r\sin(t))$.

(4) On tire de ce calcul un exemple de deux matrices qui ne commutent pas et telles que $\exp(A+B) \neq \exp(A)\exp(B)$. Prenons $A = \left(\begin{smallmatrix} 0 & -t \\ 0 & 0 \end{smallmatrix}\right)$ et $B = \left(\begin{smallmatrix} 0 & 0 \\ t & 0 \end{smallmatrix}\right)$ qui sont deux matrices de carré nul. On a :

$$\exp(A+B) = \exp(tI) = \cos(t) + I\sin(t) \;\neq\; \exp(A)\exp(B) = (1+A)(1+B) = 1 + A + B + AB.$$

Ceci étant dit, dès que A et B sont deux matrices de carré nul qui ne commutent pas, on voit que $\exp(A)\exp(B) \neq \exp(B)\exp(A)$ qui ne peut donc être égal à $\exp(A+B)$.

5.6 Théorème. *Soit $k = \mathbb{R}$ ou $\mathbb{C}$ et A une k-algèbre de dimension finie. Soit $a = s + n$ la décomposition de Jordan-Chevalley additive de a. Alors $\exp(a) = \exp(s)\exp(n)$ est la décomposition de Jordan-Chevalley multiplicative de $\exp(a)$.*

Preuve : En effet s est diagonalisable sur $\mathbb{C}$, donc $\exp(s)$ également, ce qui montre que $\exp(s)$ est semi-simple. De plus $\exp(n) = 1 + n + \ldots$ est unipotent. Enfin s et n sont des polynômes en a, donc aussi $\exp(s)$ et $\exp(n)$, ce qui montre qu'ils commutent. Le résultat en découle par unicité de la décomposition de Jordan-Chevalley. □

6 L'exponentielle des endomorphismes nilpotents

Soit A une algèbre de dimension finie sur un corps k. Si $x \in A$ est nilpotent, on peut définir son exponentielle $\exp_A(x) = \sum_{n\geqslant 0} \frac{1}{n!}x^n$ sans hypothèse sur k puisque cette série est un polynôme. De même, si $u \in A$ est unipotent, on peut définir son logarithme par la formule $\log_A(u) = \sum_{n\geqslant 1} \frac{(-1)^{n+1}}{n}(u-1)^n$, puisque $u - 1$ est nilpotent. Ces deux applications sont en fait inverses l'une de l'autre, comme le montre le résultat suivant. Une référence est [MT], 3.3.3, p. 60, mais nous n'en suivons pas vraiment la preuve.

6.1 Théorème. *Soit* $A = \mathsf{M}_n(k)$ *l'algèbre des matrices sur un corps de caractéristique* 0 *ou* $p > n$. *Alors on a des bijections réciproques*

$$\{\text{éléments nilpotents de } A\} \underset{\log_A}{\overset{\exp_A}{\rightleftarrows}} \{\text{éléments unipotents de } A\}$$

Si k *égale* $\mathbb{R}$ *ou* $\mathbb{C}$, *ces bijections sont des homéomorphismes.*

6.2 Remarques. (1) En fait l'énoncé précédent est valable dans toute algèbre de dimension finie, en remplaçant n par le maximum des indices de nilpotence des éléments nilpotents de A, qui en tout état de cause sont bornés par la dimension de A.

(2) Si k égale $\mathbb{R}$ ou $\mathbb{C}$, les bijections $\exp_A$ et $\log_A$ sont données par des expressions polynomiales, donc on pourrait avoir le désir d'annoncer une régularité bien meilleure que celle d'un simple homéomorphisme. Malheureusement, ce n'est pas possible car les ensembles source et but n'ont pas de structure différentiable. Par exemple, on peut vérifier en exercice que l'ensemble des matrices nilpotentes de taille $(2,2)$ s'identifie au sous-ensemble de k^3 d'équation $a^2 + bc = 0$. Il s'agit d'une variété algébrique *non différentiable* qui présente une singularité en $(0,0,0)$.

Preuve : Si x est nilpotent, $\exp(x) = 1 + x + \dots$ est unipotent. Réciproquement si u est unipotent, $\log(u) = (u-1) - \frac{1}{2}(u-1)^2 + \dots$ est nilpotent. Il reste à montrer que $\log$ et $\exp$ sont inverses l'une de l'autre. Partons de l'égalité de séries formelles $\log(\exp(X)) = X$, valable dans $\mathbb{Q}[[X]]$. En regardant cela modulo X^n, on obtient l'égalité analogue dans l'anneau quotient $\mathbb{Q}[[X]]/(X^n) \simeq \mathbb{Q}[X]/(X^n)$. On constate alors que les seuls dénominateurs qui interviennent ont des facteurs premiers $\leqslant n-1$. Donc, cette dernière égalité est valable dans le sous-anneau $\mathbb{Z}[\frac{1}{(n-1)!}][X]/(X^n)$. Si k est de caractéristique 0, on peut spécialiser X en un élément nilpotent de A dans l'identité $\log(\exp(X)) = X$. Si k est de caractéristique $p > n$, comme p ne divise pas $(n-1)!$ on peut réduire modulo p pour obtenir l'identité dans $\mathbb{F}_p[X]/(X^n)$ puis la plonger dans $k[X]/(X^n)$. Comme précédemment, on spécialise ensuite X en un élément nilpotent de A. On procède de la même manière pour démontrer que $\exp(\log(U)) = U$ vue comme égalité de séries formelles dans $\mathbb{Q}[[X]]$ où $X = U - 1$. □

7 L'exponentielle des groupes de Lie

Les groupes de Lie sont un sujet d'étude très riche, au croisement de plusieurs disciplines (algèbre, analyse, topologie, physique, théorie des groupes, théorie des nombres...). Pour l'agrégation, les références sont [MT] et [Fa]. On renvoie aussi aux compléments de cours de Géométrie Différentielle faits par Jürgen Angst début octobre.

7.1 Définition. On appelle *groupe de Lie linéaire* un sous-groupe fermé de $\mathsf{GL}_n(\mathbb{R})$.

On pourrait aussi considérer les sous-groupes fermés de $\mathsf{GL}_n(\mathbb{C})$, mais comme ce dernier est un sous-groupe fermé de $\mathsf{GL}_{2n}(\mathbb{R})$ cela n'ajouterait pas de généralité. Un point essentiel de la théorie est le théorème de Cartan-von Neumann ([MT] p. 68 et théorème 3.4.3 p. 66) qui affirme que tout sous-groupe fermé non discret de $\mathsf{GL}_n(\mathbb{R})$ est une sous-variété. Dans notre contexte, ce résultat est important également parce que sa preuve utilise l'exponentielle que nous allons définir ci-dessous.

Cependant, on n'insiste pas ici sur ces points ; à l'agrégation, l'intérêt des groupes de Lie résulte surtout dans l'étude d'exemples explicites comme $\mathsf{GL}_n(\mathbb{R})$, $\mathsf{SL}_n(\mathbb{R})$, $\mathrm{O}_n(\mathbb{R})$, $\mathsf{SO}_n(\mathbb{R})$, $\mathsf{Sp}_n(\mathbb{R})$, $\mathrm{U}_n(\mathbb{C})$, ou éventuellement le groupe des quaternions de norme 1.

7.2 Définition. Soit $G \subset \mathsf{GL}_n(\mathbb{R})$ un groupe de Lie linéaire. On appelle *algèbre de Lie de* G et on note $\mathsf{Lie}(G)$ l'ensemble $\mathfrak{g} = \{X \in \mathsf{M}_n(\mathbb{R}) \text{ t.q. } \exp(tX) \in G \text{ pour tout } t \in \mathbb{R}\}$.

Si on sait qu'un groupe de Lie linéaire est une sous-variété de $\mathsf{GL}_n(\mathbb{R})$, on peut aussi définir son algèbre de Lie comme l'espace tangent en l'élément identité. Cette définition montre en particulier que $\mathsf{Lie}(G)$ ne dépend que d'un voisinage de l'identité dans G ; par exemple $\mathsf{Lie}\,\mathrm{O}_n(\mathbb{R}) = \mathsf{Lie}\,\mathsf{SO}_n(\mathbb{R})$.

7.3 Exemples. (1) Si $G = \mathsf{GL}_n(\mathbb{R})$ alors $\mathfrak{g} = \mathfrak{gl}_n(\mathbb{R}) := \mathsf{M}_n(\mathbb{R})$ (évident).
(2) Si $G = \mathsf{SL}_n(\mathbb{R})$ alors $\mathfrak{g} = \mathfrak{sl}_n(\mathbb{R})$ est l'ensemble des matrices de trace nulle. En effet, grâce à la formule $\det(\exp(M)) = \exp(\mathrm{tr}(M))$, l'identité $\det(\exp(tX)) = 1$ pour tout t donne, à l'ordre 1, la condition $\mathrm{tr}(X) = 0$.
(3) Si $G = \mathsf{SO}_n(\mathbb{R})$ alors $\mathfrak{g} = \mathfrak{so}_n(\mathbb{R})$ est l'ensemble des matrices antisymétriques. En effet, l'identité $\exp(tX)\exp(tX^*) = \mathrm{Id}$ implique que $\exp(tX)$ et $\exp(tX^*)$ commutent donc $\exp(tX + tX^*) = \mathrm{Id}$ d'où, à l'ordre 1, la condition $X + X^* = 0$.

Sur ces exemples, on constate que $\mathsf{Lie}(G)$ est un $\mathbb{R}$-espace vectoriel stable par le crochet de Lie $[X, Y] := XY - YX$. C'est vrai en général et c'est un calcul assez facile pour lequel on renvoie à [MT], 3.4.1 ou [Fa], III.2.1 et II.2.4. Nous arrivons à cinq résultats importants concernant nos exemples.

7.4 Théorème. *L'application* $\exp : \mathfrak{gl}_n(\mathbb{C}) \to \mathsf{GL}_n(\mathbb{C})$ *est surjective.*

Ce résultat est fondamental et c'est le plus classique ; on le trouve dans [Gou], problèmes du chap. 4, ou [BMP], exercice 4.17. Nous le démontrons ci-dessous.

7.5 Théorème. *L'application* $\exp : \mathfrak{so}_n(\mathbb{R}) \to \mathsf{SO}_n(\mathbb{R})$ *est surjective.*

Ce résultat n'est pas vraiment énoncé dans [MT] même s'il est dans l'air ([MT], 2.6.4 montre que $\mathsf{SO}_n(\mathbb{R})$ est connexe en utilisant la réduction des endomorphismes orthogonaux comme nous le ferons ci-dessous, et [MT] 3.4.3.2 calcule l'algèbre de Lie $\mathfrak{so}_n(\mathbb{R})$). Nous le démontrons ci-dessous.

7.6 Corollaire. *Les groupes de Lie* $\mathsf{GL}_n(\mathbb{C})$ *et* $\mathsf{SO}_n(\mathbb{R})$ *sont connexes par arcs.*

Ceci se déduit immédiatement de 7.4 et 7.5 puisque les espaces vectoriels $\mathfrak{gl}_n(\mathbb{C})$ et $\mathfrak{so}_n(\mathbb{R})$ sont connexes par arcs et $\exp$ est une application continue.

7.7 Théorème. *L'application* $\exp : \mathfrak{gl}_n(\mathbb{R}) \to \mathsf{GL}_n(\mathbb{R})$ *n'est pas surjective.*

Ce résultat est clair, car $\det(\exp(M)) = \exp(\mathrm{tr}(M)) > 0$ donne l'inclusion $\exp(\mathfrak{gl}_n(\mathbb{R})) \subset \mathsf{GL}_n^+(\mathbb{R})$. On peut préciser en montrant que les matrices qui sont dans l'image de l'exponentielle sont celles qui sont le carré d'une autre matrice, voir [BMP], exercice 4.17. Nous n'en dirons pas plus sur ce résultat.

7.8 Théorème. *L'application* $\exp : \mathfrak{sl}_n(\mathbb{C}) \to \mathsf{SL}_n(\mathbb{C})$ *n'est pas surjective.*

Ce résultat peut être démontré avec les outils dont nous disposons (notamment la décomposition de Jordan-Chevalley) mais nous n'en dirons poas plus ici. Nous renvoyons à la note [Ro].

Il nous reste à donner deux démonstrations.

7.9 Preuve de 7.4. Soit $M \in \mathsf{GL}_n(\mathbb{C})$ et $M = DU$ sa décomposition de Jordan-Chevalley multiplicative. D'après le théorème 6.1, il existe $N \in \mathsf{M}_n(\mathbb{C})$ telle que $\exp(N) = U$. De plus, l'expression explicite $N = \log(U)$ montre que N est un polynôme en U donc un polynôme en M. Notons $D = P^{-1}\,\mathsf{diag}(\lambda_1, \dots, \lambda_n)P$ la partie diagonalisable. Comme D est inversible, les λ_i sont non nuls et la surjectivité de l'exponentielle complexe montre qu'il existe μ_i tel que $\exp(\mu_i) = \lambda_i$. Choisissons un polynôme d'interpolation F tel que $F(\lambda_i) = \mu_i$ pour tout i. On a alors $D' := F(D) = P^{-1}\,\mathsf{diag}(\mu_1, \dots, \mu_n)P$ est un polynôme en D, donc un polynôme en M, qui vérifie $\exp(D') = D$. Finalement D' et N sont des polynômes en M donc elles commutent entre elles, et $\exp(D' + N) = \exp(D')\exp(N) = DU = M$.

7.10 Preuve de 7.5. Soit $M \in \mathsf{SO}_n(\mathbb{R})$. Notons

$$R_\theta = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}$$

la matrice de rotation. On sait que c'est l'exponentielle de θI où

$$I = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \in \mathfrak{so}_2(\mathbb{R}).$$

D'après le théorème de réduction des matrices orthogonales, il existe $P \in \mathrm{O}_n(\mathbb{R})$ telle que

$$PMP^{-1} = \mathsf{diag}(\mathsf{Id}_r, R_{\theta_1}, \dots, R_{\theta_s}).$$

C'est donc l'exponentielle de la matrice diagonale par blocs antisymétrique :

$$Q = \mathsf{diag}(0_r, \theta_1 I, \dots, \theta_s I).$$

Comme l'exponentielle respecte la conjugaison, M est donc l'exponentielle de la matrice $P^{-1}QP$. Comme P est orthogonale, il est immédiat de vérifier que cette matrice est encore antisymétrique.

Références

[BCGLLSS] A. Bostan, F. Chyzak, M. Giusti, R. Lebreton, G. Lecerf, B. Salvy, É. Schost, *Algorithmes Efficaces en Calcul Formel*, polycopié de cours disponible à l'adresse `perso.ens-lyon.fr/bruno.salvy/mpri/poly.pdf`

[BMP] V. Beck, J. Malick, G. Peyré, *Objectif Agrégation*, H & K, 2005.

[CEZ] D. Couty, J. Esterle, R. Zarouf, *Décomposition effective de Jordan-Chevalley*, Gazette des Mathématiciens no 129 (2011), 29–49, disponible à l'adresse `http://smf4.emath.fr/Publications/Gazette/2011/129/smf_gazette_129_29-49.pdf`

[CG2] Ph. Caldero, J. Germoni, *Histoires hédonistes de groupes et de géométries*, tome II, Calvage & Mounet, 2015.

[Fa] J. Faraut, *Analyse sur les groupes de Lie*, Calvage & Mounet, 2006.

[Gou] X. Gourdon, *Algèbre*, Ellipses, 2009.

[Me] J.-Y. Mérindol, *Nombres et algèbre*, EDP Sciences, 2006.

[MT] R. Mneimné, F. Testard, *Introduction à la théorie des groupes de Lie classiques*, Hermann, 1986, réédité en 2005.

[RB] J.-J. Risler, P. Boyer, *Algèbre pour la Licence 3*, Dunod, 2006.

[Ro] M. Romagny, *L'exponentielle de* $\mathrm{SL}_n(\mathbb{C})$ *n'est pas surjective*, note disponible à l'adresse `https://perso.univ-rennes1.fr/matthieu.romagny/agreg/dvt/exp_non_surjective.pdf`.

[Sz] A. Szpirglas, *Algèbre L3*, Pearson, 2009.

[Tau] P. Tauvel, *Algèbre*, Dunod, 2005.