exercices de mathématiques

oraux x-ens

algèbre 2

Serge Francinou
Hervé Gianella
Serge Nicolas

CASSINI

EXERCICES DE MATHÉMATIQUES
ORAUX X-ENS

Enseignement des mathématiques

1. J.-Y. Ouvrard, *Probabilités I*
2. J. Hubbard, B. West, *Équations différentielles et systèmes dynamiques*
3. M. Cottrell, V. Genon-Catalot, Ch. Duhamel, Th. Meyre, *Exercices de probabilités*
4. F. Rouvière, *Petit guide de calcul différentiel à l'usage de la licence et de l'agrégation*
5. J.-Y. Ouvrard, *Probabilités II*
6. G. Zémor, *Cours de cryptographie*
7. A. Szpirglas, *Exercices d'algèbre*
8. B. Perrin-Riou, *Algèbre, arithmétique et Maple*

10-14. S. Francinou, H. Gianella, S. Nicolas, *Exercices de mathématiques — Oraux X-ENS*

15. H. Krivine, *Exercices de mathématiques pour physiciens*
16. J. Jacod, Ph. Protter, *Les bases de la théorie des probabilités*
17. M. Willem, *Analyse fonctionnelle élémentaire*
18. É. Amar, É. Matheron, *Analyse complexe*
19. B. Randé, *Problèmes corrigés. Concours 2002 et 2003 (MP)*
20. D. Perrin, *Mathématiques d'école*
21. B. Randé, *Problèmes corrigés. Concours 2004 (MP)*
22. P. Bourgade, *Olympiades internationales de mathématiques 1976-2005*

SERGE FRANCINOU
HERVÉ GIANELLA
SERGE NICOLAS

Exercices de mathématiques des oraux de l'École polytechnique et des Écoles normales supérieures

Algèbre. Tome II

CASSINI

SERGE FRANCINOU, ancien élève de l'École normale supérieure et agrégé de mathématiques, est actuellement professeur en classe préparatoire au lycée Henri IV.

HERVÉ GIANELLA, ancien élève de l'École normale supérieure et agrégé de mathématiques, est actuellement professeur en classe préparatoire au lycée Saint-Louis.

SERGE NICOLAS, ancien élève de l'École normale supérieure et agrégé de mathématiques, est actuellement professeur en classe préparatoire au lycée Henri IV.

ISBN 2-8425-091-5

Table des matières

Introduction

Cet ouvrage est le second tome d'algèbre d'un recueil d'exercices de mathématiques destiné à la préparation des oraux des concours d'entrée aux Écoles normales supérieures et à l'École polytechnique. Il comportera six tomes, trois d'algèbre et trois d'analyse.

La vocation première des Écoles normales est de former des chercheurs ou des enseignants-chercheurs. Le concours d'entrée vise donc à détecter les qualités scientifiques du candidat, son aptitude à la recherche. À l'oral, on jugera avant tout la capacité de prendre des initiatives, d'utiliser une indication, de mener à bien une démarche. On ne sera pas surpris que les exercices posés aient un contenu mathématique riche, qu'ils soient très éloignés du simple exercice technique, d'application du cours, qu'ils soient souvent difficiles. Ils visent la plupart du temps à la démonstration d'un résultat mathématique significatif. Ils pourraient apparaître excessivement difficiles si on perdait de vue le déroulement concret de l'épreuve. L'oral des ENS est un long dialogue (l'épreuve dure environ cinquante minutes, comme d'ailleurs à l'École polytechnique) entre le candidat et l'examinateur, qui tout au long de l'épreuve fournit des indications, quand c'est nécessaire, pour relancer la réflexion du candidat et tester ses réactions. Il est d'ailleurs impossible de rendre pleinement compte dans un recueil d'exercices du caractère oral de l'épreuve.

L'École polytechnique, quant à elle, est plus généraliste. Les exercices posés au concours sont de facture plus classique et, en règle générale, l'examinateur intervient moins. C'est au candidat de montrer sa maîtrise du programme dans la résolution d'un exercice dont la difficulté est cependant très variable. Certains sont proches des exercices d'ENS. Les énoncés circulent d'ailleurs d'un concours à l'autre, ou peuvent même être repris d'exercices d'Olympiades.

Les énoncés qui figurent dans ce recueil ont été donnés entre 1993 et 2004. Ils sont extraits pour l'essentiel des listes publiées chaque année par la RMS (*Revue des mathématiques de l'enseignement supérieur* aux éditions Vuibert jusqu'en 2003 et désormais *Revue de la filière Mathématiques* aux éditions e.net) dont nous remercions les auteurs pour l'aide précieuse qu'ils apportent ainsi aux élèves et aux professeurs des classes préparatoires. Il s'agit de versions communiquées par les étudiants, reflétant la compréhension que ceux-ci ont eue de l'exercice et le déroulement

conjoncturel de leur oral, comme le montrent les variations d'une année à l'autre pour un même exercice. Nous n'avons pas hésité à les modifier, pour rectifier des erreurs, compléter un énoncé quand manifestement l'exercice s'est arrêté avant que le résultat que l'examinateur avait en vue ne soit atteint, ou ajouter des indications.

Nous avons choisi de laisser quelques énoncés « bruts », ceux pour lesquels nous estimons qu'une démarche naturelle (qui peut être longue et ardue) permet de conduire à la solution. Pour d'autres exercices, nous avons pris la liberté de rajouter des questions intermédiaires, qui auraient pu être celles posées par l'examinateur. Quitte à perdre en concision, nous avons tenu à rédiger les solutions les plus pédagogiques possible, essayant d'exposer clairement les idées et démarches des raisonnements sans pour autant escamoter les détails ou calculs qui peuvent paraître évidents. On évite autant que possible l'introduction d'une astuce ou d'un objet *ad hoc* permettant d'atteindre rapidement la solution. S'il n'y a pas moyen d'expliquer l'origine de cette astuce, c'est que l'exercice est peu intéressant et que l'étudiant en tirera peu de profit.

À l'intérieur de chaque chapitre, les exercices ont été regroupés thématiquement, et à l'intérieur de chaque thème, souvent par ordre de difficulté croissante. Ainsi regroupés, ils apparaîtront plus accessibles, car plongés dans leur contexte mathématique, éclairés par d'autres exercices voisins. Les introductions historiques qui ouvrent chaque chapitre, outre leur intérêt propre, visent au même but. Enfin, nous avons agrémenté les énoncés de quelques remarques préliminaires. Sans faire de rappels de cours systématiques, nous avons énoncé, voire redémontré certains résultats : lemmes classiques, intervenant dans la résolution d'un grand nombre d'exercices, ou résultats au contraire à la lisière du programme, mais utiles, pour lesquels des éclaircissements étaient nécessaires. On trouvera aussi des remarques de synthèse ou des généralisations qui, nous l'espérons, pourront amener le candidat curieux à approfondir ses connaissances. Les quelques indications bibliographiques ont le même objectif.

Le lecteur ne tirera profit de ce livre d'exercices que s'il cherche des solutions personnelles avant d'en étudier les corrigés. Une bonne connaissance du cours est indispensable. En effet, les théorèmes du programme fournissent bon nombre de schémas de démonstration. Rappelons aussi quelques démarches générales qui peuvent faciliter l'appréhension des exercices difficiles :

▷ ne pas hésiter à faire les calculs ou à étudier le problème en dimension 2 ou 3 : par exemple pour l'étude d'un déterminant ou la réduction d'une matrice.

▷ le renforcement des hypothèses peut aboutir à un problème plus simple : cas où le corps est algébriquement clos, cas où les matrices sont diagonalisables ou inversibles (avec extension possible du résultat par un argument de densité),

▷ considérer des sous-espaces stables permet souvent d'envisager un raisonnement par récurrence sur la dimension.

Au-delà des étudiants en classe préparatoire, ces ouvrages intéresseront aussi les candidats au CAPES et à l'Agrégation, qui y trouveront matière à réviser les principales notions du programme, ainsi que des exemples pour nourrir un développement pour leur oral.

Voyons maintenant plus précisément le contenu de ce second tome d'algèbre. Le premier chapitre est consacré au déterminant et peut être abordé dès la première année en classe préparatoire. Le second chapitre sur la réduction des endomorphismes constitue le cœur du programme d'algèbre de seconde année et il est le plus riche des trois. Le dernier chapitre, aux exercices plus difficiles, est dédié à l'étude du groupe linéaire.

Comme dans les autres tomes, les exercices sont classés par thème. La difficulté est toutefois plutôt croissante : les chapitres commencent par des questions techniques ou des savoir-faire indispensables (calculs de déterminants, recherche de valeurs propres, générateurs du groupe linéaire,...) et se terminent souvent par des exercices plus théoriques.

Le troisième et dernier tome d'algèbre portera sur les espaces euclidiens, les espaces hermitiens, les formes quadratiques et la géométrie.

Nous remercions André et Catherine Bellaïche et nos élèves Nicolas Curien, Julien Élie et Baptiste Teyssier pour leur relecture approfondie de l'ouvrage et leurs nombreuses suggestions, tant sur le fond que sur la forme.

Enfin, si vous souhaitez nous contacter pour nous faire part de vos remarques, vous pouvez envoyer un courriel à l'adresse *fgn.cassini@free.fr*.

Chapitre 1
Formes multilinéaires et déterminants

L'utilisation des matrices et des déterminants trouve son origine dans l'étude systématique des systèmes linéaires menée à partir du XVIII^{e} *siècle. Alors que Leibniz et Mac Laurin avaient déjà introduit les notations à indices et résolu les systèmes à deux ou trois inconnues, Cramer, en 1754, comprend que les solutions d'un système linéaire s'expriment comme quotient de deux expressions polynomiales multilinéaires des coefficients du système. Ces expressions représentent des déterminants mais ces derniers, étudiés notamment par Vandermonde et Laplace ne sont définis alors que par récurrence sur la taille (autrement dit par le développement par rapport à une rangée). On doit également à Laplace l'interprétation du déterminant en termes de volume. Par la suite, au début du* XIX^{e} *siècle, Gauss, dans ses recherches sur les formes quadratiques, représente les changements de base dans* $\mathbb{R}^3$ *à l'aide de tableaux de nombres (les matrices) et introduit le produit de deux de ces tableaux pour obtenir la composée de deux changements de bases. Cela devait suggérer en 1812 à Cauchy la règle générale du produit de deux déterminants et il lui revient d'imposer la terminologie moderne.*

Les premiers exercices portent sur les formes multilinéaires alternées. Si E *est un* K*-espace vectoriel de dimension* n*, l'espace des formes* n*-linéaires alternées sur* E *est une droite vectorielle. Si* B *est une base de* E*, l'application déterminant dans la base* B *est un élément non nul de cette droite. L'exercice suivant contient ce résultat et demande plus généralement la dimension de l'espace des formes* n*-linéaires alternées sur un espace de dimension* d *quelconque.*

1.1. Dimension de l'espace des formes n-linéaires alternées

Soit E un K-espace vectoriel de dimension finie $d \geqslant 1$. Pour $n \geqslant 1$, on note $\mathcal{A}_n(\mathrm{E})$ l'espace des formes n-linéaires alternées sur E. Calculer $\dim \mathcal{A}_n(\mathrm{E})$.

(ENS Ulm)

▷ **Solution.**

Soit $\mathcal{B} = (e_1, \ldots, e_d)$ une base de E. À tout $(x_1, \ldots, x_n) \in \mathrm{E}^n$, on associe la matrice $\mathrm{A} = (a_{ij})_{\substack{1 \leqslant i \leqslant d \\ 1 \leqslant j \leqslant n}} \in \mathcal{M}_{d,n}(\mathrm{K})$ des coordonnées des vecteurs x_j dans la base B :

$$\forall j \in [\![1, n]\!], \quad x_j = a_{1j}e_1 + a_{2j}e_2 + \cdots + a_{dj}e_d.$$

Soit $f \in \mathcal{A}_n$. Développons $f(x_1, \ldots, x_n)$ en utilisant la multilinéarité : $f\left(\sum_{i_1=1}^{d} a_{i_1 1}e_{i_1}, \ldots, \sum_{i_n=1}^{d} a_{i_n n}e_{i_n}\right)$ est égal à

$$\sum_{(i_1,\ldots,i_n)\in[\![1,d]\!]^n} a_{i_1 1}a_{i_2 2}\cdots a_{i_n n} f(e_{i_1}, e_{i_2}, \ldots, e_{i_n}).$$

Dans la somme obtenue on peut se limiter aux indices $(i_1, \ldots, i_n)$ deux à deux distincts puisque f est alternée. En particulier, si $n > d$, f est nulle et $\dim \mathcal{A}_n(\mathrm{E}) = 0$. On supposera donc dans la suite $n \leqslant d$.

Comme f est alternée elle est aussi antisymétrique. Pour toute permutation $\sigma \in \mathcal{S}_n$ on a donc

$$f(e_{i_{\sigma(1)}}, \ldots, e_{i_{\sigma(n)}}) = \varepsilon(\sigma) f(e_{i_1}, e_{i_2}, \ldots, e_{i_n}).$$

Il en résulte que $f(x_1, \ldots, x_n)$ vaut

$$\sum_{1 \leqslant i_1 < i_2 < \cdots < i_n \leqslant d} \left(\sum_{\sigma \in \mathcal{S}_n} \varepsilon(\sigma) a_{i_{\sigma(1)}1} \cdots a_{i_{\sigma(n)}n} \right) f(e_{i_1}, e_{i_2}, \ldots, e_{i_n}).$$

Notons $\mathcal{P}_n([\![1, d]\!])$ l'ensemble des parties à n éléments de $[\![1, d]\!]$. Pour toute partie $\mathrm{I} = \{i_1 < i_2 < \cdots < i_n\} \in \mathcal{P}_n([\![1, d]\!])$, considérons la forme n-linéaire $g_{\mathrm{I}} : \mathrm{E}^n \to \mathrm{K}$ définie pour tout $(x_1, \ldots, x_n) \in \mathrm{E}^n$ par

$$g_{\mathrm{I}}(x_1, \ldots, x_n) = \sum_{\sigma \in \mathcal{S}_n} \varepsilon(\sigma) a_{i_{\sigma(1)}1} \cdots a_{i_{\sigma(n)}n}.$$

On vient de prouver que toute forme linéaire alternée f est combinaison linéaire des formes n-linéaires g_{I}. Par ailleurs, ces formes n-linéaires g_{I} sont bien alternées pour toute partie I : cela résulte du cours. En effet, $\mathrm{A} = (a_{ij})_{\substack{1 \leqslant i \leqslant d \\ 1 \leqslant j \leqslant n}}$ désignant toujours la matrice de la famille $(x_1, \ldots, x_n) \in \mathrm{E}^n$ dans la base $\mathcal{B}$, g_{I} est le déterminant de la sous-matrice de A obtenue en ne gardant que les lignes d'indice $i_1, i_2, \ldots, i_n$. Il est alors clair que si deux des x_i sont égaux, cette sous-matrice a un déterminant nul. Donc $(g_{\mathrm{I}})_{\mathrm{I} \in \mathcal{P}_n([\![1,d]\!])}$ est une partie génératrice de $\mathcal{A}_n(\mathrm{E})$.

Il ne reste plus qu'à vérifier que cette famille est libre. Supposons donc que

$$\sum_{I\in\mathcal{P}_n([\![1,d]\!])} \lambda_I g_I = 0.$$

Si on applique cette égalité à la famille $(e_{i_1}, \ldots, e_{i_n})$ avec $i_1 < i_2 < \cdots < i_n$, la seule partie I telle que $g_I(e_{i_1}, \ldots, e_{i_n})$ soit non nul est $I = \{i_1, i_2, \ldots, i_n\}$ pour laquelle on trouve 1. Il en résulte que tous les λ_I sont nuls et la famille est bien libre. Ainsi, $(g_I)_{I\in\mathcal{P}_n([\![1,d]\!])}$ est une base de $\mathcal{A}_n(\mathrm{E})$ et

$$\boxed{\dim \mathcal{A}_n(\mathrm{E}) = \mathrm{Card}\,\mathcal{P}_n([\![1,d]\!]) = \mathrm{C}_d^n}. \quad \lhd$$

Pour $n = 1$ on retrouve la dimension du dual de E *qui est celle de* E *et pour $n = d$ on obtient* 1.

Dans l'exercice suivant on montre que pour qu'une forme n-linéaire soit alternée, il suffit qu'elle s'annule sur tout n-uplet de vecteurs contenant deux vecteurs consécutifs égaux. L'objectif de l'exercice est de montrer qu'une famille de n vecteurs qui annule toute forme n-linéaire alternée est forcément liée (dans un espace E *de dimension quelconque).*

1.2. Formes n-linéaires alternées

Soit E un K-espace vectoriel, $\mathcal{A}_n(\mathrm{E})$ l'espace des formes n-linéaires alternées sur E.

1. Montrer que si φ est une forme n-linéaire vérifiant $\varphi(x_1, \ldots, x_n) = 0$ dès qu'il existe $i \in [\![1, n-1]\!]$ tel que $x_i = x_{i+1}$, alors φ est alternée.

2. Soit $f \in \mathcal{A}_n(\mathrm{E})$ et μ une forme linéaire sur E. On pose

$$g(x_1, \ldots, x_{n+1}) = \sum_{i=1}^{n+1} (-1)^i \mu(x_i) f(x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_{n+1}).$$

Montrer que $g \in \mathcal{A}_{n+1}(\mathrm{E})$.

3. Soit $(x_1, \ldots, x_n) \in \mathrm{E}^n$ tel que $g(x_1, \ldots, x_n) = 0$ pour tout $g \in \mathcal{A}_n(\mathrm{E})$. Montrer que pour toute forme $f \in \mathcal{A}_{n-1}(\mathrm{E})$ on a

$$\sum_{i=1}^{n} (-1)^i x_i f(x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n) = 0.$$

Montrer que la famille $(x_1, \ldots, x_n)$ est liée.

(ENS Lyon)

▷ **Solution.**

1. Soit φ une forme n-linéaire vérifiant $\varphi(x_1, \ldots, x_n) = 0$ dès qu'il existe $i \in [\![1, n-1]\!]$ tel que $x_i = x_{i+1}$. Soit $(x_1, \ldots, x_n) \in \mathrm{E}^n$ tel qu'il existe $i < j$ avec $x_i = x_j$. Montrons que $\varphi(x_1, \ldots, x_n) = 0$. C'est le cas si $j = i+1$. Sinon, développons

$$\varphi(x_1, \ldots, x_{i-1}, x_i - x_{i+1}, x_{i+1} - x_i, x_{i+2}, \ldots, x_n)$$

qui, par hypothèse, est nul :

$$\begin{aligned} 0 \;=\; & \varphi(x_1, \ldots, x_i, x_{i+1}, \ldots, x_n) - \varphi(x_1, \ldots, x_i, x_i, \ldots, x_n) \\ & + \varphi(x_1, \ldots, x_{i+1}, x_i, \ldots, x_n) - \varphi(x_1, \ldots, x_{i+1}, x_{i+1}, \ldots, x_n), \end{aligned}$$

ce qui donne $\varphi(x_1, \ldots, x_i, x_{i+1}, \ldots, x_n) = -\varphi(x_1, \ldots, x_{i+1}, x_i, \ldots, x_n)$. En réitérant le procédé, $\varphi(x_1, \ldots, x_i, x_{i+1}, \ldots, x_n)$ est égal à

$$(-1)^{j-i-1}\varphi(x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_{j-1}, x_i, x_j, \ldots, x_n) = 0.$$

2. La $(n+1)$-linéarité de g est claire. Utilisons la question 1 pour montrer que g est alternée. Soit $(x_1, \ldots, x_n) \in \mathrm{E}^n$ avec $x_i = x_{i+1}$ pour un certain $i \in [\![1, n-1]\!]$. Dans l'écriture de $g(x_1, \ldots, x_n)$, il y a $n-2$ termes nuls et il reste

$$\begin{aligned} g(x_1, \ldots, x_n) \;&=\; (-1)^i \mu(x_i) f(x_1, \ldots, x_{i-1}, x_{i+1}, x_{i+2}, \ldots, x_n) \\ &\quad + (-1)^{i+1} \mu(x_{i+1}) f(x_1, \ldots, x_{i-1}, x_i, x_{i+2}, \ldots, x_n) \\ &=\; 0. \end{aligned}$$

Donc g est alternée.

3. Soit $f \in \mathcal{A}_{n-1}(\mathrm{E})$. L'image du vecteur

$$\sum_{i=1}^{n} (-1)^i x_i f(x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)$$

par toute forme linéaire $\mu \in \mathrm{E}^*$ est nulle d'après la question 2. Il en résulte que ce vecteur est nul.

Pour montrer que la famille $(x_1, \ldots, x_n)$ est liée, on raisonne par récurrence sur n, le cas $n = 1$ étant immédiat. Supposons le résultat acquis au rang $n-1$ et considérons $(x_1, \ldots, x_n) \in \mathrm{E}^n$ annulant toute forme n-linéaire alternée. S'il existe une forme $(n-1)$-linéaire f telle que $f(x_1, \ldots, x_{n-1}) \neq 0$ c'est terminé car la relation ci-dessus est une relation de liaison. Sinon, l'hypothèse de récurrence s'applique à $(x_1, \ldots, x_{n-1})$ et le résultat en découle car si $(x_1, \ldots, x_{n-1})$ est liée, $(x_1, \ldots, x_n)$ l'est aussi. ◁

La forme $(n+1)$-linéaire alternée g définie dans la question 2 est le produit extérieur de la forme linéaire μ et de la forme n-linéaire alternée

f. Cette notion de produit extérieur est plus générale et permet d'associer à une forme n-linéaire alternée f et une forme m-linéaire g, une forme $(n+m)$-linéaire alternée. Ce calcul extérieur se révèle très important dans l'étude des formes différentielles (voir par exemple CARTAN HENRI. *Cours de calcul différentiel, Hermann).*

Dans les exercices qui suivent, on se préoccupe essentiellement de déterminants de matrices. La multilinéarité et le caractère alterné montrent que le déterminant d'une matrice ne change pas si on ajoute à une colonne (ou une ligne) une combinaison linéaire des autres colonnes (ou lignes). Cela est pratiquement utilisé dans tous les calculs.

1.3. Borne supérieure du déterminant sur une boule unité

On note S l'ensemble des matrices $A = (a_{ij}) \in \mathcal{M}_n(\mathbb{R})$ telle que $|a_{ij}| \leqslant 1$ pour tout couple (i,j) et on considère $\alpha = \sup_{A \in S} \det A$.
Montrer que α est fini, puis que α est un entier divisible par 2^{n-1}.
(**École polytechnique**)

▷ **Solution.**

L'ensemble S est une partie compacte de $\mathcal{M}_n(\mathbb{R})$. Le déterminant étant continu, il est borné sur S et atteint sa borne supérieure. On peut donc affirmer que α est fini et qu'il est atteint. Soit $A = (a_{ij})_{1 \leqslant i,j \leqslant n}$ dans S une matrice telle que $\det A = \alpha$. On va montrer qu'il est possible de choisir A avec tous ses coefficients dans $\{\pm 1\}$. Le déterminant est une fonction affine de chaque coefficient. Pour $(i,j) \in [\![1,n]\!]^2$, il existe donc des fonctions polynomiales U_{ij} et V_{ij} des coefficients m_{kl}, mais indépendantes de m_{ij}, telles que pour toute matrice $M = (m_{kl})_{1 \leqslant k,l \leqslant n}$ on ait $\det M = U_{ij}(M) m_{ij} + V_{ij}(M)$.

Si $U_{ij}(A) = 0$, rien de nous empêche de remplacer le coefficient a_{ij} par 1 : cela ne change pas le déterminant. Si en revanche, $U_{ij}(A)$ est non nul, alors a_{ij} est nécessairement égal à ± 1 car, sur le segment $[-1,1]$, une fonction affine non constante $ax+b$ atteint nécessairement son maximum en 1 ou en -1. Comme cela vaut pour tout couple (i,j), on a montré qu'il existe une matrice $A \in S$ telle que $\det A = \alpha$ et dont tous les coefficients valent ± 1. Le déterminant d'une matrice à coefficients dans $\mathbb{Z}$ est un entier donc α est entier. Il ne reste plus qu'à prouver qu'il est divisible par 2^{n-1}. Pour cela on utilise la multilinéarité. Ajoutons la première colonne de A à chacune des autres. On peut alors mettre 2 en facteur dans chaque colonne d'indice 2 à n. La multilinéarité du déterminant permet de conclure. ◁

Les prochains exercices sont consacrés à des calculs de déterminants.

1.4. Calcul d'un déterminant (1)

Calculer

$$\det \begin{pmatrix} a_1+b_1 & b_1 & \dots & b_1 \\ b_2 & a_2+b_2 & \dots & b_2 \\ \vdots & \vdots & \ddots & \vdots \\ b_n & b_n & \dots & a_n+b_n \end{pmatrix}.$$

(École polytechnique)

▷ **Solution.**

Notons $\mathrm{C} = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}$, $(e_1, e_2, \dots, e_n)$ la base canonique de K^n et

$$\mathrm{M} = \begin{pmatrix} a_1+b_1 & b_1 & \dots & b_1 \\ b_2 & a_2+b_2 & \dots & b_2 \\ \vdots & & \ddots & \vdots \\ b_n & b_n & \dots & a_n+b_n \end{pmatrix}.$$

La transposée de M admet pour i-ème colonne $b_i\mathrm{C} + a_i e_i$. Ainsi,

$$\det \mathrm{M} = \det({}^t\mathrm{M}) = \det(a_1 e_1 + b_1\mathrm{C}, a_2 e_2 + b_2\mathrm{C}, \dots, a_n e_n + b_n\mathrm{C}).$$

On développe cette expression en utilisant la multilinéarité du déterminant. Chaque terme où apparaissent $b_i\mathrm{C}$ et $b_j\mathrm{C}$ avec $i \neq j$ est nul, car le déterminant est une forme n-linéaire alternée. Il reste

$$\begin{aligned} \det \mathrm{M} &= \det(a_1 e_1, \dots, a_n e_n) \\ &\quad + \sum_{k=1}^{n} \det(a_1 e_1, \dots, a_{k-1} e_{k-1}, b_k\mathrm{C}, a_{k+1} e_{k+1}, \dots, a_n e_n) \\ &= a_1 \dots a_n + \sum_{k=1}^{n} b_k \Big(\prod_{i \neq k} a_i \Big) \det(e_1, \dots e_{k-1}, \mathrm{C}, e_{k+1}, \dots, e_n). \end{aligned}$$

Si à C on enlève $e_1 + \dots + e_{k-1} + e_{k+1} + \dots + e_n$ $(1 \leqslant k \leqslant n)$, il reste e_k. Comme det est une forme multilinéaire alternée,

$$\det(e_1, \dots e_{k-1}, \mathrm{C}, e_k, \dots, e_n) = \det(e_1, \dots, e_{k-1}, e_k, e_{k+1}, \dots, e_n) = 1$$

et finalement

$$\det \mathrm{M} = \prod_{i=1}^{n} a_i + \sum_{k=1}^{n} b_k \prod_{i \neq k} a_i \,. \triangleleft$$

L'exemple suivant est typique des calculs de déterminants : on effectue des opérations sur la matrice de manière à annuler certains coefficients, puis on développe selon une ligne ou une colonne pour se ramener à des déterminants de taille inférieure.

1.5. Calcul d'un déterminant (2)

Calculer le déterminant de la matrice $(\mathrm{C}_{i+j-2}^{j-1})_{1 \leqslant i,j \leqslant n+1}$.
(**École polytechnique**)

▷ **Solution.**
Il s'agit de calculer

$$\mathrm{D} = \begin{vmatrix} \mathrm{C}_0^0 & \mathrm{C}_1^1 & \dots & \mathrm{C}_{n-1}^{n-1} & \mathrm{C}_n^n \\ \mathrm{C}_1^0 & \mathrm{C}_2^1 & \dots & \mathrm{C}_n^{n-1} & \mathrm{C}_{n+1}^n \\ \mathrm{C}_2^0 & \mathrm{C}_3^1 & \dots & \mathrm{C}_{n+1}^{n-1} & \mathrm{C}_{n+2}^n \\ \vdots & \vdots & & \vdots & \vdots \\ \mathrm{C}_n^0 & \mathrm{C}_{n+1}^1 & \dots & \mathrm{C}_{2n-1}^{n-1} & \mathrm{C}_{2n}^n \end{vmatrix}.$$

On va effectuer sur les lignes de la matrice les opérations élémentaires suivantes : $\mathrm{L}_{n+1} \leftarrow \mathrm{L}_{n+1} - \mathrm{L}_n$, $\mathrm{L}_n \leftarrow \mathrm{L}_n - \mathrm{L}_{n-1}$, $\dots$, $\mathrm{L}_2 \leftarrow \mathrm{L}_2 - \mathrm{L}_1$. Compte-tenu de la relation de Pascal, on a $\mathrm{C}_{i+j-2}^{j-1} - \mathrm{C}_{i+j-3}^{j-1} = \mathrm{C}_{i+j-3}^{j-2}$ pour tout $1 \leqslant i \leqslant n+1$ et $2 \leqslant j \leqslant n$, et il reste donc

$$\mathrm{D} = \begin{vmatrix} 1 & 1 & 1 & \dots & 1 \\ 0 & \mathrm{C}_1^0 & \mathrm{C}_2^1 & \dots & \mathrm{C}_n^{n-1} \\ 0 & \mathrm{C}_2^0 & \mathrm{C}_3^1 & \dots & \mathrm{C}_{n+1}^{n-1} \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & \mathrm{C}_n^0 & \mathrm{C}_n^1 & \dots & \mathrm{C}_{2n-1}^{n-1} \end{vmatrix} = \begin{vmatrix} \mathrm{C}_1^0 & \mathrm{C}_2^1 & \dots & \mathrm{C}_n^{n-1} \\ \mathrm{C}_2^0 & \mathrm{C}_3^1 & \dots & \mathrm{C}_{n+1}^{n-1} \\ \vdots & \vdots & & \vdots \\ \mathrm{C}_n^0 & \mathrm{C}_n^1 & \dots & \mathrm{C}_{2n-1}^{n-1} \end{vmatrix}.$$

Sur ce dernier déterminant, on réitère la même suite d'opérations : $\mathrm{L}_n \leftarrow \mathrm{L}_n - \mathrm{L}_{n-1}$, $\dots$, $\mathrm{L}_2 \leftarrow \mathrm{L}_2 - \mathrm{L}_1$:

$$\mathrm{D} = \begin{vmatrix} 1 & \mathrm{C}_2^1 & \dots & \mathrm{C}_n^{n-1} \\ 0 & \mathrm{C}_2^0 & \dots & \mathrm{C}_{n-2}^{n-1} \\ \vdots & \vdots & & \vdots \\ 0 & \mathrm{C}_{n-1}^0 & \dots & \mathrm{C}_{2n-2}^{n-2} \end{vmatrix} = \begin{vmatrix} \mathrm{C}_2^0 & \dots & \mathrm{C}_{n-2}^{n-1} \\ \vdots & & \vdots \\ \mathrm{C}_{n-1}^0 & \dots & \mathrm{C}_{2n-2}^{n-2} \end{vmatrix}$$

Ainsi, de proche en proche, au bout de n séries d'opérations, il reste un déterminant de taille $(1,1)$:

$$\mathrm{D} = \left|\mathrm{C}_{2n-n}^{n-n}\right| = \left|\mathrm{C}_n^0\right| = 1. \lhd$$

Précisons que la notation avec deux grandes barres verticales pour signifier un déterminant est due à Cayley (1841).

1.6. Calcul d'un déterminant (3)

Calculer le déterminant

$$\begin{vmatrix} a_1 & a_0 & 0 & \dots & 0 & 0 \\ a_2 & a_1 & a_0 & & 0 & 0 \\ \vdots & & \ddots & \ddots & & \vdots \\ \vdots & & & \ddots & \ddots & 0 \\ a_{n-1} & a_{n-2} & \dots & \dots & a_1 & a_0 \\ a_n & a_{n-1} & a_{n-2} & \dots & a_2 & a_1 \end{vmatrix}$$

avec $a_0 = 1$, $a_1 = \dfrac{1}{2}$, $\dots$, $a_n = \dfrac{1 \cdot 3 \cdot 5 \cdots (2n-1)}{2 \cdot 4 \cdot 6 \cdots (2n)}$.

(École polytechnique)

▷ **Solution.**

Appelons D_n le déterminant que nous devons calculer. On va essayer de trouver une relation de récurrence entre les D_n et pour cela on va développer D_n par rapport à la dernière ligne. Le cofacteur de a_k est

$$(-1)^{k+1} \begin{vmatrix} a_1 & a_0 & 0 & \dots & 0 & 0 & \dots & \dots & 0 \\ a_2 & a_1 & a_0 & & \vdots & \vdots & & & \vdots \\ \vdots & & & \ddots & \vdots & \vdots & & & \vdots \\ a_{n-k-1} & & & & a_0 & 0 & \dots & \dots & 0 \\ a_{n-k} & & & & a_1 & 0 & 0 & \dots & 0 \\ a_{n-k+1} & & & & a_2 & a_0 & 0 & \dots & 0 \\ \vdots & & & & \vdots & a_1 & a_0 & & 0 \\ \vdots & & & & \vdots & \vdots & & \ddots & \vdots \\ a_{n-1} & & & & a_k & a_{k-2} & & & a_0 \end{vmatrix},$$

ce qui donne, d'après les règles de calcul des déterminants par blocs, $(-1)^{k+1}\mathrm{D}_{n-k}a_0^{k-1} = (-1)^{k+1}\mathrm{D}_{n-k}$. On en déduit que

$$D_n = (-1)^{n+1}a_nD_0 + (-1)^n a_{n-1}D_1 + \cdots + (-1)^{k+1}a_kD_{n-k} + \cdots + a_1D_{n-1},$$

avec la convention $D_0 = 1$. Autrement dit. si on pose

$$u_n = \sum_{k=0}^{n}(-1)^k a_k D_{n-k},$$

on a $u_0 = 1$ et $u_n = 0$ pour $n \geqslant 1$.

On reconnaît dans la suite (u_n) le produit de Cauchy de la suite $((-1)^n a_n)$ et de la suite (D_n). On considère alors les séries entières $f(x) = \sum_{n=0}^{+\infty} D_n x^n$ et $g(x) = \sum_{n=0}^{+\infty}(-1)^n a_n x^n$. La seconde a un rayon de convergence égal à 1 et on reconnaît le développement en série entière de $(1+x)^{-1/2}$. En effet, rappelons que pour tout α réel et tout $x \in]-1, 1[$, on a

$$(1+x)^\alpha = 1 + \sum_{n=1}^{+\infty} \frac{\alpha(\alpha-1)\cdots(\alpha-n+1)}{n!}x^n.$$

En prenant $\alpha = -\frac{1}{2}$, on obtient

$$\frac{\alpha(\alpha-1)\cdots(\alpha-n+1)}{n!} = \frac{(-1/2).(-3/2).\cdots.(1/2-n)}{n!} = (-1)^n a_n.$$

Notons que a_n vaut aussi $\frac{(2n)!}{(2^n n!)^2} = \frac{C_{2n}^n}{4^n}$. Or, $\frac{1}{g(x)} = \sqrt{1+x}$ est aussi développable en série entière avec un rayon égal à 1 et, comme $\frac{d}{dx}(\sqrt{1+x}) = \frac{1}{2\sqrt{x+1}} = \frac{1}{2}g(x)$, on a

$$\sqrt{1+x} = \frac{1}{g(x)} = 1 + \sum_{n=1}^{+\infty} \frac{(-1)^{n-1}a_{n-1}}{2n}x^n.$$

Si l'on écrit pour $|x| < 1$, $\sqrt{1+x} = \sum_{n=0}^{+\infty} D'_n x^n$, la règle du produit de Cauchy donne

$$\sum_{k=0}^{n}(-1)^k a_k D'_{n-k} = \begin{cases} 1 \text{ si } n = 0 \\ 0 \text{ si } n \geqslant 1 \end{cases}.$$

Comme $D_0 = D'_0 = 1$, par une récurrence immédiate, on en déduit $D'_n = D_n$ et on obtient donc pour tout $n \geqslant 1$,

$$\boxed{D_n = (-1)^{n-1}\frac{a_{n-1}}{2n}}.\triangleleft$$

L'expression polynomiale du déterminant sert rarement dans les calculs pratiques. C'est toutefois le cas dans l'exercice suivant.

1.7. Calcul d'un déterminant (4)

Soit $A = (a_{ij}) \in \mathcal{M}_n(\mathbb{C})$ telle que :
(*i*) pour $i \neq j$, $a_{ij} \neq 0 \Longrightarrow a_{ji} = 0$;
(*ii*) ($a_{ij} \neq 0$ et $a_{jk} \neq 0$) $\Longrightarrow a_{ik} \neq 0$ dès que i, j, k sont deux à deux distincts.
Calculer $\det A$.

(ENS Ulm)

▷ **Solution.**
On va utiliser la définition du déterminant

$$\det A = \sum_{\sigma \in \mathcal{S}_n} \varepsilon(\sigma) \prod_{i=1}^{n} a_{i\sigma(i)}.$$

Soit σ une permutation distincte de l'identité. Elle admet une décomposition en produit de cycles à supports disjoints. Soit $c = (i_1, i_2, \ldots, i_k)$, $k \geqslant 2$, l'un des cycles en question. Dans le produit $\prod_{i=1}^{n} a_{i\sigma(i)}$ intervient le facteur $a_{i_1 i_2} a_{i_2 i_3} \cdots a_{i_{k-1} i_k} a_{i_k i_1}$. Mais les propriétés (*i*) et (*ii*) impliquent la nullité d'au moins un terme de ce produit (s'ils sont tous non nuls, la transitivité conduit à $a_{i_1 i_k} \neq 0$ et alors $a_{i_k i_1} = 0$ par (*ii*) : contradiction). La seule contribution au déterminant est donc donnée par l'identité et

$$\boxed{\det A = \prod_{i=1}^{n} a_{ii}}. \quad \triangleleft$$

Les exercices ci-après concernent des déterminants très classiques, à commencer par celui de Vandermonde.

1.8. Déterminant de Vandermonde

1. Calculer le déterminant

$$V(x_1, \ldots, x_n) = \begin{vmatrix} 1 & x_1 & x_1^2 & \ldots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \ldots & x_2^{n-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_n & x_n^2 & \ldots & x_n^{n-1} \end{vmatrix}.$$

2. Montrer que, pour tout $(m_1, \ldots, m_n) \in \mathbb{Z}^n$, l'entier $\prod_{k=1}^{n-1} k!$ divise $\prod_{1 \leqslant i < j \leqslant n} (m_j - m_i)$.

(École polytechnique)

▷ **Solution.**

1. Si deux des x_i sont égaux, $\mathrm{V}(x_1, \ldots, x_n) = 0$ (deux lignes sont identiques). Supposons les x_i deux à deux distincts et posons

$$\mathrm{P}(\mathrm{X}) = \mathrm{V}(x_1, x_2, \ldots, x_{n-1}, \mathrm{X}).$$

C'est un polynôme de degré au plus $n-1$. Le développement suivant la dernière ligne nous persuade que le coefficient en X^{n-1} est $\mathrm{V}(x_1, \ldots, x_{n-1})$. De plus, si on substitue x_i à X ($1 \leqslant i \leqslant n-1$), deux lignes sont identiques et $\mathrm{P}(x_i) = 0$. Les x_i étant deux à deux distincts, P est divisible par $\prod_{i=1}^{n-1} (\mathrm{X} - x_i)$, polynôme unitaire de degré $n-1$. Il en résulte que

$$\mathrm{P}(\mathrm{X}) = \mathrm{V}(x_1, \ldots, x_{n-1}) \prod_{i=1}^{n-1} (\mathrm{X} - x_i),$$

et en particulier

$$\mathrm{P}(x_n) = \mathrm{V}(x_1, \ldots, x_{n-1}) \prod_{i=1}^{n-1} (x_n - x_i).$$

De proche en proche, on en tire la célèbre formule de Vandermonde

$$\boxed{\mathrm{V}(x_1, \ldots, x_n) = \prod_{1 \leqslant i < j \leqslant n} (x_j - x_i)}.$$

2. On reconnaît dans le second entier $\mathrm{V}(m_1, \ldots, m_n)$. Or, par des opérations sur les colonnes, il est clair que $\mathrm{V}(m_1, \ldots, m_n)$ vaut aussi

$$\begin{vmatrix} 1 & m_1 & m_1(m_1-1) & \ldots & m_1(m_1-1)\cdots(m_1-n+1) \\ 1 & m_2 & m_2(m_2-1) & \ldots & m_2(m_2-1)\cdots(m_2-n+1) \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & m_n & m_n(m_n-1) & \ldots & m_n(m_n-1)\cdots(m_n-n+1) \end{vmatrix}.$$

Comme un produit de k entiers consécutifs est toujours divisible par $k!$ (cf. par exemple exercice 5.20 du tome 1 pour une preuve de ce résultat classique), les entiers de la $(k+1)$-ième colonne sont tous divisibles

par $k!$. Par multilinéarité on en déduit que $\mathrm{V}(m_1, \ldots, m_n)$ est divisible par $\prod_{k=1}^{n-1} k!$. $\lhd$

Le lecteur trouvera une seconde solution de la question 2 (moins élégante) dans l'exercice 5.49 du tome 1.

Les cinq exercices qui suivent supposent connu le déterminant de Vandermonde. On imagine que lors de l'oral le candidat a certainement été amené à rapidement retrouver ce résultat classique.

1.9. Matrice de Vandermonde incomplète

Pour $0 \leqslant k \leqslant n$, calculer le déterminant :

$$\mathrm{D}_k = \begin{vmatrix} 1 & x_1 & \ldots & x_1^{k-1} & x_1^{k+1} & \ldots & x_1^n \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 1 & x_n & \ldots & x_n^{k-1} & x_n^{k+1} & \ldots & x_n^n \end{vmatrix}.$$

(École polytechnique)

$\rhd$ **Solution.**

L'idée est simplement de se ramener à un déterminant de Vandermonde en rajoutant la colonne manquante, une ligne supplémentaire et d'interpréter le déterminant proposé comme un mineur de la matrice obtenue. On considère donc le polynôme

$$\mathrm{P}(\mathrm{X}) = \begin{vmatrix} 1 & x_1 & x_1^2 & \ldots & x_1^{n-1} & x_1^n \\ 1 & x_2 & x_2^2 & \ldots & x_2^{n-1} & x_2^n \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 1 & x_n & x_n^2 & \ldots & x_n^{n-1} & x_n^n \\ 1 & \mathrm{X} & \mathrm{X}^2 & \ldots & \mathrm{X}^{n-1} & \mathrm{X}^n \end{vmatrix}.$$

Le déterminant de Vandermonde incomplet D_k s'interprète comme le mineur relatif à X^k. On a, d'après l'exercice précédent,

$$\begin{aligned} \mathrm{P}(\mathrm{X}) = \mathrm{V}(x_1, \ldots, x_n, \mathrm{X}) &= \mathrm{V}(x_1, \ldots, x_n) \prod_{k=1}^{n} (\mathrm{X} - x_k) \\ &= \mathrm{V}(x_1, \ldots, x_n)(\mathrm{X}^n - \sigma_1 \mathrm{X}^{n-1} + \cdots + (-1)^n \sigma_n) \end{aligned}$$

où les σ_k sont les fonctions symétriques élémentaires des x_k. Le développement par rapport à la dernière ligne nous assure que

$$(-1)^{n+1+k+1}\mathrm{D}_k = (-1)^{n-k}\sigma_{n-k}\mathrm{V}(x_1, \ldots, x_n)$$

et finalement

$$\boxed{\mathrm{D}_k = \sigma_{n-k}\mathrm{V}(x_1, \ldots, x_n)}. \lhd$$

Dans l'exercice suivant on considère des déterminants qui ressemblent à celui de Vandermonde, mais les exposants sont des réels strictement positifs quelconques.

1.10. Vandermonde généralisé

Soient $\alpha_1 < \cdots < \alpha_n$ des réels strictement positifs et $a_1, \ldots, a_n$ des réels non tous nuls.

1. Montrer que la fonction $f(x) = a_1x^{\alpha_1} + \cdots + a_nx^{\alpha_n}$ admet au plus $n-1$ zéros distincts dans $\mathbb{R}_+^*$.

2. Soient $t_1, \ldots, t_n$ des réels tels que $0 < t_1 < \cdots < t_n$. Montrer que le déterminant de la matrice $(t_i^{\alpha_j})_{1\leqslant i,j\leqslant n}$ est strictement positif.

(**École polytechnique**)

▷ **Solution.**

1. On procède par récurrence sur n, le résultat étant clair pour $n = 1$, puisque une fonction puissance $x \mapsto x^\alpha$ ne s'annule pas sur $\mathbb{R}_+^*$. Supposons le résultat vrai au rang $n-1$. Soient $\alpha_1 < \cdots < \alpha_n$ des réels strictement positifs et $f(x) = a_1x^{\alpha_1} + \cdots + a_nx^{\alpha_n}$. On peut supposer qu'aucun a_i n'est nul sinon l'hypothèse de récurrence s'applique directement. Supposons par l'absurde que f admette n zéros distincts $0 < x_1 < x_2 < \cdots < x_n$. En divisant par x^{α_1}, on constate que la fonction g définie sur $\mathbb{R}_+^*$ par

$$g(x) = a_1 + a_2x^{\alpha_2-\alpha_1} + \cdots + a_nx^{\alpha_n-\alpha_1}$$

s'annule aussi en les x_k. Le théorème de Rolle appliqué sur chaque intervalle $[x_k, x_{k+1}]$ montre que g' admet au moins $n-1$ zéros deux à deux distincts. Il en est de même de la fonction $x \longmapsto x^{\alpha_1}g'(x) = \sum_{k=2}^{n} a_k(\alpha_k - \alpha_1)x^{\alpha_k}$. Cela contredit l'hypothèse de récurrence.

2. On va procéder par récurrence sur n. Pour $n = 1$ le déterminant vaut $t_1^{\alpha_1}$ et il est strictement positif. Supposons le résultat vrai au rang $n-1$ et soit $t_1, \ldots, t_n$ des réels tels que $0 < t_1 < \cdots < t_n$. On considère la fonction

par $k!$. Par multilinéarité on en déduit que $\mathrm{V}(m_1, \ldots, m_n)$ est divisible par $\prod_{k=1}^{n-1} k!$. $\lhd$

Le lecteur trouvera une seconde solution de la question 2 (moins élégante) dans l'exercice 5.49 du tome 1.

Les cinq exercices qui suivent supposent connu le déterminant de Vandermonde. On imagine que lors de l'oral le candidat a certainement été amené à rapidement retrouver ce résultat classique.

1.9. Matrice de Vandermonde incomplète

Pour $0 \leqslant k \leqslant n$, calculer le déterminant :

$$\mathrm{D}_k = \begin{vmatrix} 1 & x_1 & \ldots & x_1^{k-1} & x_1^{k+1} & \ldots & x_1^n \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 1 & x_n & \ldots & x_n^{k-1} & x_n^{k+1} & \ldots & x_n^n \end{vmatrix}.$$

(École polytechnique)

$\rhd$ **Solution.**

L'idée est simplement de se ramener à un déterminant de Vandermonde en rajoutant la colonne manquante, une ligne supplémentaire et d'interpréter le déterminant proposé comme un mineur de la matrice obtenue. On considère donc le polynôme

$$\mathrm{P}(\mathrm{X}) = \begin{vmatrix} 1 & x_1 & x_1^2 & \ldots & x_1^{n-1} & x_1^n \\ 1 & x_2 & x_2^2 & \ldots & x_2^{n-1} & x_2^n \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 1 & x_n & x_n^2 & \ldots & x_n^{n-1} & x_n^n \\ 1 & \mathrm{X} & \mathrm{X}^2 & \ldots & \mathrm{X}^{n-1} & \mathrm{X}^n \end{vmatrix}.$$

Le déterminant de Vandermonde incomplet D_k s'interprète comme le mineur relatif à X^k. On a, d'après l'exercice précédent,

$$\begin{aligned} \mathrm{P}(\mathrm{X}) &= \mathrm{V}(x_1, \ldots, x_n, \mathrm{X}) = \mathrm{V}(x_1, \ldots, x_n) \prod_{k=1}^{n} (\mathrm{X} - x_k) \\ &= \mathrm{V}(x_1, \ldots, x_n)(\mathrm{X}^n - \sigma_1 \mathrm{X}^{n-1} + \cdots + (-1)^n \sigma_n) \end{aligned}$$

où les σ_k sont les fonctions symétriques élémentaires des x_k. Le développement par rapport à la dernière ligne nous assure que

$$(-1)^{n+1+k+1}\mathrm{D}_k = (-1)^{n-k}\sigma_{n-k}\mathrm{V}(x_1, \ldots, x_n)$$

et finalement

$$\boxed{\mathrm{D}_k = \sigma_{n-k}\mathrm{V}(x_1, \ldots, x_n)}. \lhd$$

Dans l'exercice suivant on considère des déterminants qui ressemblent à celui de Vandermonde, mais les exposants sont des réels strictement positifs quelconques.

1.10. Vandermonde généralisé

Soient $\alpha_1 < \cdots < \alpha_n$ des réels strictement positifs et $a_1, \ldots, a_n$ des réels non tous nuls.

1. Montrer que la fonction $f(x) = a_1x^{\alpha_1} + \cdots + a_nx^{\alpha_n}$ admet au plus $n-1$ zéros distincts dans $\mathbb{R}_+^*$.

2. Soient $t_1, \ldots, t_n$ des réels tels que $0 < t_1 < \cdots < t_n$. Montrer que le déterminant de la matrice $(t_i^{\alpha_j})_{1\leqslant i,j\leqslant n}$ est strictement positif.

(**École polytechnique**)

▷ **Solution.**

1. On procède par récurrence sur n, le résultat étant clair pour $n = 1$, puisque une fonction puissance $x \mapsto x^\alpha$ ne s'annule pas sur $\mathbb{R}_+^*$. Supposons le résultat vrai au rang $n-1$. Soient $\alpha_1 < \cdots < \alpha_n$ des réels strictement positifs et $f(x) = a_1x^{\alpha_1} + \cdots + a_nx^{\alpha_n}$. On peut supposer qu'aucun a_i n'est nul sinon l'hypothèse de récurrence s'applique directement. Supposons par l'absurde que f admette n zéros distincts $0 < x_1 < x_2 < \cdots < x_n$. En divisant par x^{α_1}, on constate que la fonction g définie sur $\mathbb{R}_+^*$ par

$$g(x) = a_1 + a_2x^{\alpha_2-\alpha_1} + \cdots + a_nx^{\alpha_n-\alpha_1}$$

s'annule aussi en les x_k. Le théorème de Rolle appliqué sur chaque intervalle $[x_k, x_{k+1}]$ montre que g' admet au moins $n-1$ zéros deux à deux distincts. Il en est de même de la fonction $x \longmapsto x^{\alpha_1}g'(x) = \sum_{k=2}^{n} a_k(\alpha_k - \alpha_1)x^{\alpha_k}$. Cela contredit l'hypothèse de récurrence.

2. On va procéder par récurrence sur n. Pour $n = 1$ le déterminant vaut $t_1^{\alpha_1}$ et il est strictement positif. Supposons le résultat vrai au rang $n-1$ et soit $t_1, \ldots, t_n$ des réels tels que $0 < t_1 < \cdots < t_n$. On considère la fonction

$$f(t) = \begin{vmatrix} t_1^{\alpha_1} & t_1^{\alpha_2} & \dots & t_1^{\alpha_n} \\ t_2^{\alpha_1} & t_2^{\alpha_2} & \dots & t_2^{\alpha_n} \\ \vdots & \vdots & & \vdots \\ t^{\alpha_1} & t^{\alpha_2} & \dots & t^{\alpha_n} \end{vmatrix}.$$

Le problème consiste à montrer que $f(t_n) > 0$. Or, si on développe selon la dernière ligne, on constate que f est de la forme

$$t \longmapsto a_1 t^{\alpha_1} + \cdots + a_n t^{\alpha_n}.$$

L'hypothèse de récurrence permet de dire que le coefficient a_n est strictement positif. La question 1 montre que f admet au plus $n-1$ zéros distincts. Or, il est clair que $f(t_1) = f(t_2) = \cdots = f(t_{n-1}) = 0$ (car pour ces valeurs deux des lignes de la matrice sont égales). Il en résulte que f ne s'annule pas sur l'intervalle $]t_{n-1}, +\infty[$, et comme elle est continue, elle y garde un signe constant. Il suffit d'observer que $f(t)$ tend vers $+\infty$ quand t tend vers $+\infty$ (car $a_n > 0$) pour conclure que f est strictement positive sur $]t_{n-1}, +\infty[$, ce qui termine la récurrence. ◁

1.11. Une autre généralisation du Vandermonde

Soit $i_1, i_2, \dots, i_k$ dans $\mathbb{N}^*$ et $n = i_1 + \cdots + i_k$. Pour $x \in \mathbb{R}$, on note $\mathrm{C}(x) = (1, x, x^2, \dots, x^{n-1})$. On se donne $x_1, \dots, x_k$ dans $\mathbb{R}$ et on considère la matrice M dont les lignes successives sont $\mathrm{C}(x_1), \frac{\mathrm{C}'(x_1)}{1!}, \dots, \frac{\mathrm{C}^{(i_1-1)}(x_1)}{(i_1-1)!}, \dots, \mathrm{C}(x_k), \frac{\mathrm{C}'(x_k)}{1!}, \dots, \frac{\mathrm{C}^{(i_k-1)}(x_k)}{(i_k-1)!}$. Calculer $\det \mathrm{M}$.

(École polytechnique)

▷ **Solution.**

On considère un polynôme P réel unitaire de degré $n-1$ qu'on écrit $\mathrm{P} = a_0 + a_1\mathrm{X} + \cdots + a_{n-2}\mathrm{X}^{n-2} + \mathrm{X}^{n-1}$. On ajoute à la dernière colonne de M la première multipliée par a_0, la seconde multipliée par a_1, jusqu'à l'avant dernière multipliée par a_{n-2}. Cela ne change pas le déterminant et fait apparaître dans la dernière colonne les coefficients

$$\mathrm{P}(x_1), \frac{\mathrm{P}'(x_1)}{1!}, \dots, \frac{\mathrm{P}^{(i_1-1)}(x_1)}{(i_1-1)!}, \dots, \mathrm{P}(x_k), \frac{\mathrm{P}'(x_x)}{1!}, \dots, \frac{\mathrm{P}^{(i_k-1)}(x_k)}{(i_k-1)!}.$$

L'idée est de choisir le polynôme P de manière à avoir beaucoup de 0 dans cette dernière colonne. En fait, on peut s'arranger pour que $\mathrm{P}(x_1) = \mathrm{P}'(x_1) = \cdots = \mathrm{P}^{(i_1-1)}(x_1) = 0$: il suffit que x_1 soit racine d'ordre i_1 de P. De même pour les autres x_i. Toutefois le degré de P doit être égal

à $n-1$ et P ne peut donc avoir que $n-1$ racines avec multiplicité. On prend alors

$$\mathrm{P}(\mathrm{X}) = (\mathrm{X}-x_1)^{i_1} \ldots (\mathrm{X}-x_{k-1})^{i_{k-1}}(\mathrm{X}-x_k)^{i_k-1}.$$

Avec ce polynôme, le seul coefficient non nul dans la dernière colonne est $\dfrac{\mathrm{P}^{(i_k-1)}(x_k)}{(i_k-1)!}$. Or, il est aisé de voir que

$$\mathrm{P}^{(i_k-1)}(x_k) = (i_k-1)!(x_k-x_1)^{i_1} \ldots (x_k-x_{k-1})^{i_{k-1}}.$$

En développant selon la dernière colonne, on voit que

$$\det \mathrm{M} = \frac{\mathrm{P}^{(i_k-1)}(x_k)}{(i_k-1)!} \times \det \mathrm{N} = \prod_{1\leqslant j\leqslant k-1} (x_k-x_j)^{i_j} \times \det \mathrm{N},$$

où N est la matrice de taille $n-1$ fabriquée tout comme M mais avec x_k à l'ordre i_k-1. Une récurrence aisée sur n, montre alors que

$$\boxed{\det \mathrm{M} = \prod_{j<k} (x_k-x_l)^{i_j i_k}}. \quad \triangleleft$$

Les deux exercices suivants utilisent des déterminants de Vandermonde.

1.12. Application du déterminant de Vandermonde (1)

Soit $\mathrm{A} = (a_{ij}) \in \mathcal{M}_n(\mathrm{K})$ avec, pour $1 \leqslant i,j \leqslant n$, $a_{ij} = 1 + x_i^j$ où $x_1, ..., x_n$ sont des scalaires. Calculer le déterminant de A.

(École polytechnique)

▷ **Solution.**

En soustrayant à chaque colonne d'indice $j \in [\![2, n]\!]$ la colonne précédente, on obtient

$$\det \mathrm{A} = \begin{vmatrix} 1+x_1 & x_1(x_1-1) & \ldots & x_1^{n-1}(x_1-1) \\ 1+x_2 & x_2(x_2-1) & \ldots & x_2^{n-1}(x_2-1) \\ \vdots & & & \vdots \\ \vdots & & & \vdots \\ 1+x_n & x_n(x_n-1) & \ldots & x_n^{n-1}(x_n-1) \end{vmatrix}.$$

En écrivant $1+x_i = 2x_i - (x_i-1)$ et en utilisant la linéarité du déterminant, on obtient

$$\det A = \begin{vmatrix} 2x_1 & x_1(x_1-1) & \dots & x_1^{n-1}(x_1-1) \\ 2x_2 & x_2(x_2-1) & \dots & x_2^{n-1}(x_2-1) \\ \vdots & & & \vdots \\ \vdots & & & \vdots \\ 2x_n & x_n(x_n-1) & \dots & x_n^{n-1}(x_n-1) \end{vmatrix}$$

$$- \begin{vmatrix} x_1-1 & x_1(x_1-1) & \dots & x_1^{n-1}(x_1-1) \\ x_2-1 & x_1(x_2-1) & \dots & x_2^{n-1}(x_2-1) \\ \vdots & & & \vdots \\ \vdots & & & \vdots \\ x_n-1 & x_n(x_n-1) & \dots & x_n^{n-1}(x_n-1) \end{vmatrix}.$$

Dans le premier déterminant, après avoir factorisé par 2, on additionne successivement chaque colonne à la suivante et on obtient

$$2 \begin{vmatrix} x_1 & x_1^2 & \dots & x_1^n \\ x_2 & x_2^2 & \dots & x_2^n \\ \vdots & & & \vdots \\ \vdots & & & \vdots \\ x_n & x_n^2 & \dots & x_n^n \end{vmatrix}$$

qui est égal à $2x_1 \cdots x_n$ fois le déterminant de Vandermonde $V(x_1, \dots, x_n)$.

Dans le second déterminant, après avoir mis en facteur $x_1 - 1$, $x_2 - 1$, $\dots$, $x_n - 1$ dans les lignes 1 à n, on obtient de nouveau le déterminant de Vandermonde. On a donc

$$\det A = 2x_1 \dots x_n V(x_1, \dots, x_n) - (x_1 - 1) \dots (x_n - 1) V(x_1, \dots, x_n)$$

$$= \prod_{i<j} (x_j - x_i) \left[2x_1 x_2 \dots x_n - (x_1 - 1)(x_2 - 1) \dots (x_n - 1) \right]. \lhd$$

1.13. Application du déterminant de Vandermonde (2)

Soit $m < n$ deux entiers naturels, $P_1, \dots, P_n$ des polynômes de $\mathbb{R}_m[X]$ et $(a_1, \dots a_n)$ des réels.

1. Calculer

$$\begin{vmatrix} P_1(a_1) & \dots & P_1(a_n) \\ P_2(a_1) & \dots & P_2(a_n) \\ \vdots & & \vdots \\ P_n(a_1) & \dots & P_n(a_n) \end{vmatrix}.$$

À quelle condition le déterminant est-il non nul ?

2. En déduire pour $m < n$, la valeur de

$$\begin{vmatrix} 1^m & 2^m & \dots & n^m \\ 2^m & 3^m & \dots & (n+1)^m \\ \vdots & \vdots & & \vdots \\ n^m & (n+1)^m & \dots & (2n-1)^m \end{vmatrix}.$$

(École polytechnique)

▷ **Solution.**

1. Si la famille $(P_1, \dots, P_n)$ est liée, le déterminant est nul. C'est le cas en particulier si $n > m+1 = \dim \mathbb{R}_m[X]$. Supposons donc $n = m+1$.

• La première méthode est inspirée du calcul du déterminant de Vandermonde : si deux des a_i sont égaux, le déterminant est nul. Supposons donc que les a_i sont deux à deux distincts.

Les polynômes $P_1, \dots, P_n$ étant fixés, notons $D(a_1, \dots, a_n)$ le déterminant cherché et posons $P(X) = D(a_1, \dots, a_{n-1}, X)$. Le polynôme P est de degré inférieur ou égal à $m = n-1$ et possède $n-1$ racines distinctes $a_1, \dots, a_{n-1}$. Il s'écrit $D_1(a_1, \dots, a_{n-1}) \prod_{i=1}^{n-1} (X - a_i)$, où D_1 est une fonction polynomiale. En particulier, on obtient

$$D(a_1, \dots, a_n) = D_1(a_1, \dots, a_{n-1}) \prod_{i=1}^{n-1} (a_n - a_i).$$

En réitérant le raisonnement avec a_{n-1} au lieu de a_n, on montre que $D_1(a_1, \dots, a_{n-1}) = D_2(a_1, \dots, a_{n-2}) \prod_{i=1}^{n-2} (a_{n-1} - a_i)$ et finalement, on obtient que

$$D(a_1, \dots, a_n) = KV(a_1, \dots, a_n),$$

où K ne dépend pas des a_i mais seulement des polynômes P_j.

Posons désormais $K = S(P_1, \dots, P_n)$. Il est clair que S est une forme n-linéaire alternée. En notant det le déterminant dans la base canonique de $\mathbb{R}_{n-1}[X]$, on obtient qu'il existe $k \in \mathbb{R}$ tel que, pour tous polynômes $P_1, \dots, P_n$,

$$D(a_1, \dots, a_n) = kV(a_1, \dots, a_n) \det(P_1, \dots, P_n).$$

Si $(P_1, \dots, P_n)$ est la base canonique de $\mathbb{R}_{n-1}[X]$, on a alors $D(a_1, \dots, a_n) = V(a_1, \dots, a_n)$, ce qui montre que $k = 1$ et

$$\mathrm{D}(a_1, \ldots, a_n) = \mathrm{V}(a_1, \ldots, a_n) \det(\mathrm{P}_1, \ldots, \mathrm{P}_n).$$

Cette formule reste vraie si les a_i ne sont pas distincts.

Cela montre que le déterminant est nul si, et seulement si, $n > m+1$ ou $n = m+1$ et ou bien deux des a_i sont égaux ou bien la famille $(\mathrm{P}_1, \ldots, \mathrm{P}_n)$ est liée.

• La deuxième méthode consiste à écrire la matrice des $\mathrm{P}_i(a_j)$ comme un produit : on se place comme précédemment dans le cas $m = n-1$. Si on écrit pour $1 \leqslant i \leqslant n$, $\mathrm{P}_i = c_{i,0} + c_{i,1}\mathrm{X} + \cdots + c_{i,n-1}\mathrm{X}^{n-1}$, on a

$$\mathrm{P}_i(a_j) = c_{i,0}a_j^0 + c_{i,1}a_j^1 + \cdots + c_{i,n-1}a_j^{n-1},$$

si bien que la matrice des $\mathrm{P}_i(a_j)$ apparaît comme le produit de la matrice $(c_{i,j})_{\substack{1\leqslant i\leqslant n\\ 0\leqslant j\leqslant n-1}}$ (qui n'est rien d'autre que la transposée de matrice de $(\mathrm{P}_1, \ldots, \mathrm{P}_n)$ dans la base $(1, \mathrm{X}, \ldots, \mathrm{X}^{n-1})$) par la matrice de Vandermonde $(a_j^{i-1})_{1\leqslant i,j\leqslant n}$. On en déduit que le déterminant cherché est $\det(\mathrm{P}_1, \ldots, \mathrm{P}_n)\mathrm{V}(a_1, \ldots, a_n)$.

2. En posant $\mathrm{P}_1 = \mathrm{X}^m$, $\mathrm{P}_2 = (\mathrm{X}+1)^m$, $\ldots$, $\mathrm{P}_n = (\mathrm{X}+n-1)^m$ et $a_i = i$ pour $1 \leqslant i \leqslant n$, le déterminant à calculer, que l'on note D_n, est

$$\begin{vmatrix} \mathrm{P}_1(a_1) & \ldots & \mathrm{P}_1(a_n) \\ \mathrm{P}_2(a_1) & \ldots & \mathrm{P}_2(a_n) \\ \vdots & & \vdots \\ \mathrm{P}_n(a_1) & \ldots & \mathrm{P}_n(a_n) \end{vmatrix}.$$

On applique le résultat de la question 1. Si $n > m+1$, $\mathrm{D}_n = 0$. On suppose $n = m+1$. On calcule

$$\mathrm{V}(a_1, \ldots, a_n) = \mathrm{V}(1, \ldots, n) = \prod_{j=2}^{n}\prod_{i=1}^{j-1}(j-i) = \prod_{j=2}^{n}(j-1)! = \prod_{j=1}^{n-1} j!$$

et $\det(\mathrm{P}_1, \ldots, \mathrm{P}_n)$. Comme $\mathrm{P}_j = \sum_{i=0}^{n-1} \mathrm{C}_{n-1}^i (j-1)^{n-1-i}\mathrm{X}^i$, on obtient, en mettant C_{n-1}^i en facteur dans la $(i+1)$-ième ligne,

$$\det(\mathrm{P}_1, \ldots, \mathrm{P}_n) = \mathrm{C}_{n-1}^0 \mathrm{C}_{n-1}^1 \ldots \mathrm{C}_{n-1}^{n-1} \begin{vmatrix} 0 & 1 & 2^{n-1} & \ldots & (n-1)^{n-1} \\ 0 & 1 & 2^{n-2} & \ldots & (n-1)^{n-2} \\ \vdots & & & & \vdots \\ 0 & 1 & 2 & \ldots & (n-1) \\ 1 & 1 & 1 & \ldots & 1 \end{vmatrix}$$

$$= \frac{(n-1)!^n}{(0!1!\ldots(n-1)!)^2}(-1)^{\frac{n(n-1)}{2}}\mathrm{V}(0, 1, \ldots, n-1),$$

le déterminant est nul dès que deux α_i sont égaux (resp. deux β_j) puisque deux lignes (resp. deux colonnes) sont alors identiques. Nous supposerons maintenant les α_i (resp. les β_j) deux à deux distincts.

Posons

$$F(X) = \begin{vmatrix} \frac{1}{\alpha_1+\beta_1} & \frac{1}{\alpha_1+\beta_2} & \cdots & \frac{1}{\alpha_1+\beta_n} \\ \vdots & \vdots & & \vdots \\ \frac{1}{\alpha_{n-1}+\beta_1} & \frac{1}{\alpha_{n-1}+\beta_2} & \cdots & \frac{1}{\alpha_{n-1}+\beta_n} \\ \frac{1}{X+\beta_1} & \frac{1}{X+\beta_2} & \cdots & \frac{1}{X+\beta_n} \end{vmatrix} \in \mathbb{C}(X).$$

En imaginant le développement de ce déterminant selon la dernière ligne, on convient que F est la somme de n fractions nulles ou de degré -1 et donc $\deg F \leqslant -1$. Toujours avec ce développement, on peut réduire au même dénominateur pour mettre F sous la forme

$$F = \frac{P(X)}{(X+\beta_1)(X+\beta_2)\dots(X+\beta_n)}.$$

On a donc $\deg P \leqslant n-1$. Les α_i $(1 \leqslant i \leqslant n-1)$ ne sont pas pôles de F et $F(\alpha_i) = 0$ (en substituant α_i à X, le déterminant présente deux lignes identiques). Donc $P(\alpha_i) = 0$. Les α_i étant deux à deux distincts et $\deg P \leqslant n-1$, il existe $\lambda \in \mathbb{R}$ avec

$$P = \lambda(X-\alpha_1)(X-\alpha_2)\dots(X-\alpha_{n-1}).$$

Il s'agit de calculer λ. Multiplions F par $X+\beta_n$:

$$(X+\beta_n)F = \begin{vmatrix} \frac{1}{\alpha_1+\beta_1} & \cdots & \cdots & \frac{1}{\alpha_1+\beta_n} \\ \vdots & \vdots & & \vdots \\ \frac{1}{\alpha_{n-1}+\beta_1} & \cdots & \cdots & \frac{1}{\alpha_{n-1}+\beta_n} \\ \frac{X+\beta_n}{X+\beta_1} & \cdots & \frac{X+\beta_n}{X+\beta_{n-1}} & 1 \end{vmatrix}$$
$$= \lambda \frac{(X-\alpha_1)\dots(X-\alpha_{n-1})}{(X+\beta_1)\dots(X+\beta_{n-1})}$$

et évaluons en $-\beta_n$:

$$\begin{vmatrix} \frac{1}{\alpha_1+\beta_1} & \cdots & \frac{1}{\alpha_1+\beta_{n-1}} & \times \\ \vdots & & \vdots & \vdots \\ \frac{1}{\alpha_{n-1}+\beta_1} & \cdots & \frac{1}{\alpha_{n-1}+\beta_{n-1}} & \times \\ 0 & \dots & 0 & 1 \end{vmatrix} = \lambda \frac{(-1)^{n-1}(\beta_n+\alpha_1)\dots(\beta_n+\alpha_{n-1})}{(-1)^{n-1}(\beta_n-\beta_1)\dots(\beta_n-\beta_{n-1})}.$$

En développant ce déterminant par rapport à la dernière ligne, il vient

$$1 \times \begin{vmatrix} \frac{1}{\alpha_1+\beta_1} & \cdots & \frac{1}{\alpha_1+\beta_{n-1}} \\ \vdots & & \vdots \\ \frac{1}{\alpha_{n-1}+\beta_1} & \cdots & \frac{1}{\alpha_{n-1}+\beta_{n-1}} \end{vmatrix} = \lambda \frac{(\beta_n+\alpha_1)\dots(\beta_n+\alpha_{n-1})}{(\beta_n-\beta_1)\dots(\beta_n-\beta_{n-1})},$$

d'où

$$\lambda = \begin{vmatrix} \frac{1}{\alpha_1+\beta_1} & \cdots & \frac{1}{\alpha_1+\beta_{n-1}} \\ \vdots & & \vdots \\ \frac{1}{\alpha_{n-1}+\beta_1} & \cdots & \frac{1}{\alpha_{n-1}+\beta_{n-1}} \end{vmatrix} \frac{(\beta_n-\beta_1)\dots(\beta_n-\beta_{n-1})}{(\beta_n+\alpha_1)\dots(\beta_n+\alpha_{n-1})}.$$

Le déterminant cherché, $\mathrm{F}(\alpha_n)$ est donc égal à

$$\begin{vmatrix} \frac{1}{\alpha_1+\beta_1} & \cdots & \frac{1}{\alpha_1+\beta_{n-1}} \\ \vdots & & \vdots \\ \frac{1}{\alpha_{n-1}+\beta_1} & \cdots & \frac{1}{\alpha_{n-1}+\beta_{n-1}} \end{vmatrix} \frac{\prod_{i<n}(\beta_n-\beta_i)\prod_{i<n}(\alpha_n-\alpha_i)}{\prod_{i<n}(\beta_n+\alpha_i)\prod_{i\leqslant n}(\alpha_i+\beta_n)}.$$

On établit par récurrence la formule du déterminant de Cauchy. ◁

1.15. Déterminant de Hürwitz

Calculer le déterminant

$$\begin{vmatrix} r_1 & a & \dots & a \\ b & r_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & a \\ b & \dots & b & r_n \end{vmatrix}$$

dans $\mathbb{R}$, puis dans un corps commutatif quelconque.

(ENS Lyon)

▷ **Solution.**

Notre démarche va consister à rajouter une indéterminée aux coefficients de la matrice. Posons

$$\mathrm{P(X)} = \begin{vmatrix} r_1 + \mathrm{X} & a + \mathrm{X} & \dots & a + \mathrm{X} \\ b + \mathrm{X} & r_2 + \mathrm{X} & \ddots & \vdots \\ \vdots & \ddots & \ddots & a + \mathrm{X} \\ b + \mathrm{X} & \dots & b + \mathrm{X} & r_n + \mathrm{X} \end{vmatrix}.$$

Il s'agit d'un polynôme de degré inférieur ou égal à 1. Pour s'en convaincre, il suffit de retrancher la première colonne à toutes les autres avant de développer par rapport à cette colonne. Il suffit donc de connaître les valeurs de P en deux points distincts.

Supposons dans un premier temps $a \neq b$. On a on a $\mathrm{P}(-a) = \mathrm{Q}(a)$ et $\mathrm{P}(-b) = \mathrm{Q}(b)$, en posant $\mathrm{Q} = (r_1 - \mathrm{X}) \dots (r_n - \mathrm{X})$. La valeur qui nous intéresse est $\mathrm{P}(0)$. Si on écrit $\mathrm{P(X)} = \alpha\mathrm{X} + \beta$, on obtient $\mathrm{P}(-a) = -\alpha a + \beta$ et $\mathrm{P}(-b) = -\alpha b + \beta$ de sorte que

$$\mathrm{P}(0) = \beta = \frac{b\mathrm{P}(-a) - a\mathrm{P}(-b)}{b - a} = \boxed{\frac{b\mathrm{Q}(a) - a\mathrm{Q}(b)}{b - a}}$$

Il reste à traiter le cas où $a = b$. Dans le cas où le corps de base est $\mathbb{R}$ il suffit de faire tendre b vers a dans l'expression précédente. On obtient

$$\lim_{b \to a} \frac{b\mathrm{Q}(a) - a\mathrm{Q}(b)}{b - a} = \lim_{b \to a} \mathrm{Q}(a) - a\frac{\mathrm{Q}(b) - \mathrm{Q}(a)}{b - a} = \boxed{\mathrm{Q}(a) - a\mathrm{Q}'(a)}.$$

En fait ce résultat reste encore valable avec un corps quelconque : on retrouve un cas particulier du déterminant de l'exercice 1.4 (avec les notations de cet exercice il suffit en effet de prendre les b_k tous égaux à a et $a_k = r_k - a$ pour tout k).◁

Le calcul de ce cas particulier a aussi été détaillé dans l'exercice 1.11 du tome 1 où la matrice de Hürwitz intervenait dans un problème combinatoire.

L'exercice suivant reprend avec une autre argumentation le résultat de l'exercice 6.28 du premier tome d'Algèbre.

1.16. Liberté d'une famille de fonctions

Soit X un ensemble, K un corps commutatif et n fonctions $f_1, f_2, \dots, f_n$ de X dans K. Montrer que les applications $f_1, f_2, \dots, f_n$ sont libres dans $\mathcal{F}(\mathrm{X}, \mathrm{K})$ si, et seulement si, il existe $x_1, x_2, \dots, x_n \in \mathrm{X}$ tel que le déterminant de $(f_i(x_j))_{1 \leqslant i,j \leqslant n}$ est non nul.

(École Polytechnique)

▷ **Solution.**

Si on suppose la famille $(f_1, \ldots, f_n)$ liée, il existe des scalaires $\lambda_1, \lambda_2, \ldots, \lambda_n$ non tous nuls tels que $\lambda_1 f_1 + \lambda_2 f_2 + \cdots + \lambda_n f_n = 0$. Si on évalue en x_j, il vient $\lambda_1 f_1(x_j) + \lambda_2 f_2(x_j) + \cdots + \lambda_n f_n(x_j) = 0$, ce qui donne une relation de liaison sur les lignes de la matrice $(f_i(x_j))_{1 \leqslant i,j \leqslant n}$. Son déterminant est donc nul.

Montrons maintenant par récurrence sur $n \geqslant 1$ que si $(f_1, f_2, \ldots, f_n)$ est libre, il existe $x_1, x_2, \ldots, x_n \in \mathrm{X}$ tels que le déterminant de $(f_i(x_j))_{1 \leqslant i,j \leqslant n}$ est non nul. Si $n = 1$ et (f_1) est libre, $f_1 \neq 0$ et il existe donc x_1 tel que $f_1(x_1) \neq 0$. Supposons $n \geqslant 2$ et les f_i libres. La sous-famille $(f_1, \ldots, f_{n-1})$ est donc libre et par hypothèse de récurrence, il existe $(x_1, \ldots, x_{n-1}) \in \mathrm{X}^{n-1}$ tel que $\det(f_i(x_j))_{1 \leqslant i,j \leqslant n-1}$ est non nul. Considérons, pour $x \in \mathrm{X}$, le déterminant

$$\Phi(x) = \begin{vmatrix} f_1(x_1) & \cdots & f_1(x_{n-1}) & f_1(x) \\ \vdots & & & \vdots \\ f_{n-1}(x_1) & \cdots & f_{n-1}(x_{n-1}) & f_{n-1}(x) \\ f_n(x_1) & \cdots & f_n(x_{n-1}) & f_n(x) \end{vmatrix}.$$

En développant $\Phi(x)$ par rapport à la dernière colonne, on obtient

$$\Phi(x) = \mathrm{D}_1 f_1(x) + \cdots + \mathrm{D}_{n-1} f_{n-1}(x) + \mathrm{D}_n f_n(x),$$

avec D_i cofacteur de $f_i(x)$. La fonction Φ s'écrit

$$\Phi = \mathrm{D}_1 f_1 + \cdots + \mathrm{D}_{n-1} f_{n-1} + \mathrm{D}_n f_n.$$

Comme $\mathrm{D}_n = \det(f_i(x_j))_{1 \leqslant i,j \leqslant n-1} \neq 0$ et les f_i sont libres, Φ ne peut être identiquement nulle (car sinon on aurait une relation de liaison). Il existe donc x_n tel que $\Phi(x_n) \neq 0$. Le n-uplet $(x_1, \ldots, x_{n-1}, x_n)$ répond à la question. ◁

Deux matrices semblables ont le même déterminant. La réduction est donc un outil puissant qui peut intervenir dans le calcul des déterminants.

1.17. Matrice circulante

Soit p premier et $(a_0, a_1, \ldots, a_{p-1}) \in \mathbb{Z}^p$. Montrer que

$$\begin{vmatrix} a_0 & a_1 & a_2 & \cdots & a_{p-1} \\ a_{p-1} & a_0 & a_1 & \cdots & a_{p-2} \\ a_{p-2} & a_{p-1} & a_0 & \cdots & a_{p-3} \\ \vdots & & & \ddots & \vdots \\ a_1 & a_2 & a_3 & \cdots & a_0 \end{vmatrix} \equiv a_0 + a_1 + \cdots + a_{p-1} \quad [p].$$

(École polytechnique)

▷ **Solution.**

La disposition des a_k dans la matrice proposée invite à introduire la matrice de permutation J correspondant au p-cycle $(p, p-1, \ldots, 1)$:

$$\mathrm{J} = \begin{pmatrix} 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 1 & & \vdots \\ \vdots & & \ddots & \ddots & \vdots \\ \vdots & & & 0 & 1 \\ 1 & 0 & \ldots & 0 & 0 \end{pmatrix}.$$

On a $\mathrm{J}^p = \mathrm{I}$ et $\mathrm{A} = a_0\mathrm{I} + a_1\mathrm{J} + a_2\mathrm{J}^2 + \cdots + a_{p-1}\mathrm{J}^{p-1}$. Notons $\bar{\mathrm{A}}$ la matrice de $\mathcal{M}_p(\mathbb{Z}/p\mathbb{Z})$ obtenue à partir de A en remplaçant les a_i par leurs classes modulo p. Le déterminant de $\bar{\mathrm{A}}$ est alors égal à la classe modulo p du déterminant de A. On est donc ramené à prouver que $\det \bar{\mathrm{A}} = \sum_{i=0}^{p-1} \bar{a}_i$.

On a $\bar{\mathrm{A}} = \sum_{k=0}^{p-1} \bar{a}_k \bar{\mathrm{J}}^k$. On va trigonaliser la matrice $\bar{\mathrm{J}}$. Son polynôme caractéristique est

$$\chi_{\bar{\mathrm{J}}}(\mathrm{X}) = (-1)^p(\mathrm{X}^p - 1) = (-1)^p(\mathrm{X} - 1)^p \in \mathbb{Z}/p\mathbb{Z}[\mathrm{X}]$$

car il est bien connu que p divise C_p^k pour $1 \leqslant k \leqslant p-1$ (cf. exercice 4.24 du tome 1). Comme $\chi_{\bar{\mathrm{J}}}$ est scindé, $\bar{\mathrm{J}}$ est trigonalisable, semblable à une matrice triangulaire supérieure dont les termes diagonaux sont tous égaux à $\bar{1}$. Il en résulte que $\bar{\mathrm{A}}$ est semblable à une matrice triangulaire supérieure ayant pour coefficients diagonaux $\sum_{k=0}^{p-1} \bar{a}_k$. Ainsi,

$$\det \bar{\mathrm{A}} = \left(\sum_{k=0}^{p-1} \bar{a}_k \right)^p = \sum_{k=0}^{p-1} \bar{a}_k{}^p = \sum_{k=0}^{p-1} \bar{a}_k$$

la dernière égalité résultant du petit théorème de Fermat. D'où le résultat demandé. ◁

On s'intéressera de manière plus générale à la réduction des matrices circulantes dans le second chapitre (exercices 2.20 et 2.21).

Le thème des matrices stochastiques est un sujet riche qui sera largement développé dans le chapitre réduction (exercices 2.6 et suivants).

L'énoncé qui vient propose d'étudier la borne supérieure du déterminant sur l'ensemble des matrices stochastiques.

1.18. Majoration du déterminant d'une matrice stochastique

Soit $M = (m_{ij})_{1\leqslant i,j\leqslant n} \in \mathcal{M}_n(\mathbb{R})$ telle que $m_{ij} \geqslant 0$ pour tout couple (i,j) et $\sum_{j=1}^{n} m_{ij} = 1$ pour tout $i \in [\![1,n]\!]$.

1. Montrer que $|\det M| \leqslant 1$.

2. Étudier le cas d'égalité.

(École Polytechnique)

▷ **Solution.**

1. Plusieurs approches sont possibles pour cette question. On peut par exemple montrer que toutes les valeurs propres complexes d'une matrice stochastique sont de module $\leqslant 1$ comme cela est fait dans l'exercice 2.6 du chapitre diagonalisation. Cela implique le résultat puisque le déterminant est le produit des valeurs propres. On va ici montrer le résultat par récurrence sur n, le cas $n = 1$ étant évident. Soit $M = (m_{ij})_{1\leqslant i,j\leqslant n} \in \mathcal{M}_n(\mathbb{R})$ vérifiant les hypothèses. Pour se ramener à des matrices de taille $n-1$ on développe le déterminant de M selon la première ligne. On a

$$\det M = \sum_{j=1}^{n} m_{1j}(-1)^{1+j} \det M_{1j}$$

où M_{1j} est extraite de M en ôtant la première ligne et la j-ième colonne. Les coefficients de ces matrices M_{1j} sont bien positifs mais la somme des coefficients de chaque ligne ne vaut plus nécessairement 1 : ces sommes sont toutes $\leqslant 1$. Ce petit problème nous empêche d'appliquer l'hypothèse de récurrence à ces matrices. En fait, on s'en sort très facilement en démontrant un résultat plus fort ! On prend comme propriété

(H_n) : *toute matrice* $M \in \mathcal{M}_n(\mathbb{R})$ *dont les coefficients sont positifs ou nuls et telle que la somme des coefficients de chaque ligne est inférieure ou égale à* 1 *vérifie* $|\det M| \leqslant 1$.

L'assertion (H_1) est toujours claire. Si on suppose (H_{n-1}) vraie, et si $M = (m_{ij})_{1\leqslant i,j\leqslant n} \in \mathcal{M}_n(\mathbb{R})$ vérifie les hypothèses de (H_n), on a par le développement ci-dessus,

$$|\det M| \leqslant \sum_{j=1}^{n} m_{1j} |\det M_{1j}| \leqslant \sum_{j=1}^{n} m_{1j} \leqslant 1$$

car toutes les matrices extraites M_{1j} vérifient les hypothèses de (H_{n-1}). Le résultat est donc prouvé.

2. Étudions le cas d'égalité.

Pour $n = 1$ la seule matrice qui réalise l'égalité est la matrice identité.

On observe dans la majoration réalisée ci-dessus que, s'il y a égalité, tous les cofacteurs des coefficients m_{1j} non nuls doivent valoir 1 en valeur absolue et que la somme des coefficients de la première ligne doit valoir 1. En développant selon une autre ligne, on a plus généralement le cofacteur de chaque coefficient non nul qui vaut 1 en valeur absolue, et la somme des coefficients de chaque ligne doit faire 1, c'est-à-dire que la matrice est stochastique.

Regardons alors ce qui se passe pour une matrice $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ de taille $(2,2)$. Si $a > 0$ on a donc $|d| = 1$ et donc $d = 1$ puisque les coefficients sont positifs. On a alors $c = 0$ car $c+d = 1$. De $|\det M| = 1$, on tire $a = 1$ puis $b = 0$: la matrice M est la matrice identité. Lorsque $a = 0$, on voit de même que $M = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$. On a alors l'intuition du résultat suivant : les seules matrices stochastiques M telles que $|\det M| = 1$ sont les matrices de permutation (qui conviennent clairement).

Supposons cela vrai au rang $n-1$ et soit $M = (m_{ij})_{1 \leqslant i,j \leqslant n} \in \mathcal{M}_n(\mathbb{R})$ une matrice qui réalisé l'égalité au rang n. L'un au moins des coefficients de la première ligne est non nul. Quitte à permuter les colonnes de M on peut supposer sans perte de généralité que $m_{11} > 0$. La matrice extraite $M' = (m_{ij})_{2 \leqslant i,j \leqslant n}$ est de taille $n-1$ et doit vérifier $|\det M'| = 1$ d'après ce qui est dit plus haut. Par hypothèse de récurrence il s'agit d'une matrice de permutation. Comme la somme des coefficients de chaque ligne de M vaut 1 on a nécessairement $m_{21} = m_{31} = \cdots = m_{n1} = 0$. Ainsi, si on développe selon la première colonne, $\det M = m_{11} \det M'$. Cela impose donc $m_{11} = 1$ et par suite $m_{12} = \cdots = m_{1n} = 0$. La matrice M est une matrice de permutation. Cela termine la récurrence et prouve le résultat annoncé. ◁

La comatrice est le thème des exercices suivants. Le seul résultat vraiment important à retenir est la relation

$$^t\operatorname{Com} A\, A = A\, ^t\operatorname{Com} A = (\det A) I_n$$

valable pour toute matrice $A \in \mathcal{M}_n(K)$. *On en déduit notamment le rang de la comatrice en fonction du rang de* A. *Si* A *est inversible i.e. de rang* n, *on a* $\det A \neq 0$ *et la comatrice de* A *est inversible. Si* $\operatorname{rg} A < n-1$, *la comatrice de* A *est nulle car tous les mineurs de taille* $n-1$ *extraits de* A *sont nuls. Dans le cas où* $\operatorname{rg} A = n-1$, *la comatrice n'est pas nulle.*

Mais la relation ${}^t\mathrm{Com}\,\mathrm{A}\,\mathrm{A} = 0$ *montre que l'image de* A *est incluse dans le noyau de* ${}^t\mathrm{Com}\,\mathrm{A}$. *On a donc dans ce cas* $\mathrm{rg}\,{}^t\mathrm{Com}\,\mathrm{A} = \mathrm{rg}\,\mathrm{Com}\,\mathrm{A} = 1$. *On se rendra compte à travers les exercices suivants qu'on est souvent amené à discuter selon ces trois cas.*

1.19. Calcul d'une trace

Soit A, B dans $\mathcal{M}_n(\mathbb{R})$ et X un vecteur non nul de $\mathbb{R}^n$. On suppose que $\mathrm{AX} = 0$ et qu'il existe $\mathrm{Y} \in \mathbb{R}^n$ tel que $\mathrm{AY} = \mathrm{BX}$. On note A_j la matrice obtenue en remplaçant la j-ième colonne de A par celle de B. Montrer que $\sum_{j=1}^{n} \det \mathrm{A}_j = 0$.

(ENS Ulm)

▷ **Solution.**

Interprétons d'abord le déterminant de A_j. Notons Δ_{ij} les cofacteurs de la matrice A et (b_{ij}) les coefficients de B. Si on développe le déterminant de A_j par rapport à la j-ième colonne on obtient

$$\det \mathrm{A}_j = \sum_{i=1}^{n} b_{ij}\Delta_{ij}.$$

Cette somme s'interprète comme le coefficient d'indice (j, j) de la matrice produit $\widetilde{\mathrm{A}}\mathrm{B}$, où $\widetilde{\mathrm{A}}$ désigne la transposée de la comatrice de A. La somme que l'on demande d'estimer n'est donc autre que $\mathrm{Tr}(\widetilde{\mathrm{A}}\mathrm{B})$.

La matrice A n'est pas inversible par hypothèse donc $\mathrm{rg}(\mathrm{A}) \leqslant n-1$. Si $\mathrm{rg}(\mathrm{A}) < n-1$, le rang étant la taille maximale des sous-matrices de A inversibles, tous les mineurs d'ordre $n-1$ sont nuls et la comatrice est nulle. Le résultat s'en déduit. Supposons donc que $\mathrm{rg}\,\mathrm{A} = n-1$. Son noyau est alors la droite engendrée par X. On sait que $\mathrm{A}\widetilde{\mathrm{A}} = (\det \mathrm{A})\mathrm{I} = 0$. En particulier $\mathrm{Im}\,\widetilde{\mathrm{A}} \subset \mathrm{Ker}\,\mathrm{A}$ de sorte que $\widetilde{\mathrm{A}}$ est de rang au plus 1 (en fait de rang 1 exactement car A admet au moins un mineur de taille $n-1$ non nul).

Posons $\mathrm{M} = \widetilde{\mathrm{A}}\mathrm{B}$. On a $\mathrm{rg}(\mathrm{M}) \leqslant 1$ et $\mathrm{Im}(\mathrm{M}) \subset \mathrm{Ker}\,\mathrm{A} = \mathbb{R}.\mathrm{X}$ vu ce qui précède. Si $\mathrm{M} = 0$, on a fini. Sinon, on prend une base $(e_1, \ldots, e_{n-1})$ du noyau de M que l'on complète par e_n en une base de $\mathbb{R}^n$. La matrice de l'endomorphisme canoniquement associé à M dans la base $(e_1, \ldots, e_n)$ est de la forme

$$\begin{pmatrix} 0 & 0 & \dots & 0 & \times \\ 0 & 0 & & 0 & \times \\ \vdots & \vdots & & \vdots & \times \\ 0 & 0 & & 0 & \times \\ 0 & 0 & \dots & 0 & d \end{pmatrix}.$$

On en déduit que $d = \operatorname{Tr} M = 0$ si, et seulement si, $\operatorname{Im} M \subset \operatorname{Ker} M$. Or justement, BX est dans l'image de A donc dans le noyau de $\widetilde{A}$ et $MX = 0$. D'où le résultat. $\lhd$

Dans les deux exercices suivants et dans bien d'autres de cet ouvrage, on utilise la densité des matrices inversibles dans $\mathcal{M}_n(K)$ lorsque $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$: pour $A \in \mathcal{M}_n(\mathbb{K})$, la fonction polynôme $\lambda \longmapsto \det(A - \lambda I_n)$ n'est pas identiquement nulle donc a des zéros isolés. Pour λ assez petit non nul, $A - \lambda I_n$ est inversible. Pour $k \in \mathbb{N}^*$ assez grand, la matrice $A_k = A - \frac{1}{k} I_n$ est inversible et la suite (A_k) converge vers A.

La densité du groupe linéaire permet d'étendre par continuité à toute matrice de $\mathcal{M}_n(\mathbb{K})$ des propriétés valables pour les matrices inversibles.

1.20. Spectre de la comatrice

Soit $A \in \mathcal{M}_n(\mathbb{R})$ non inversible. On note $\widetilde{A}$ la transposée de la comatrice de A. Quelles sont les valeurs propres de $\widetilde{A}$?

(**École polytechnique**)

$\rhd$ **Solution.**

On distingue trois cas selon le rang de A.

• Si A est inversible, la relation $\widetilde{A}A = A\widetilde{A} = (\det A) I_n$ montre que $\widetilde{A} = (\det A) A^{-1}$. Le spectre de la $\widetilde{A}$ est donc

$$\left\{ \frac{\det A}{\lambda}, \lambda \in \operatorname{Sp}(A) \right\}.$$

• Si $\operatorname{rg} A < n - 1$, tous les mineurs d'ordre $n - 1$ de A sont nuls et la comatrice de A est nulle. Le spectre de $\widetilde{A}$ est donc réduit à $\{0\}$.

• Il reste le cas $\operatorname{rg} A = n - 1$: il existe alors un mineur d'ordre $n - 1$ non nul, puisque le rang est la taille maximale des sous-matrices carrées inversibles et donc $\operatorname{rg} \widetilde{A} \geqslant 1$. D'autre part, $A\widetilde{A} = (\det A) I_n = 0$. Donc $\operatorname{Im} \widetilde{A} \subset \operatorname{Ker} A$ qui, d'après le théorème du rang, est une droite de $\mathbb{R}^n$. Il s'ensuit que $\operatorname{Im} \widetilde{A} = \operatorname{Ker} A$ et donc $\dim \operatorname{Ker} \widetilde{A} = n - 1$ (toujours en vertu du théorème du rang). Donc 0 est valeur propre de $\widetilde{A}$ avec une

multiplicité au moins égale à $n-1$. Nous savons que la somme des valeurs propres comptées avec multiplicité est égale à la trace. On en déduit donc que le spectre de la $\widetilde{A}$ est $\{0, d\}$ avec $d = \operatorname{Tr}\widetilde{A}$. Intéressons-nous à $\operatorname{Tr}\widetilde{A}$.

Revenons au cas A inversible : notons $\lambda_1, \ldots, \lambda_n$ les racines du polynôme caractéristique qui sont donc non nulles. Les scalaires $\dfrac{\det A}{\lambda_i}$ sont les valeurs propres distinctes ou confondues de $\widetilde{A}$. Donc sa trace est

$$\operatorname{Tr}\widetilde{A} = \det A \sum_{i=1}^{n} \frac{1}{\lambda_i} = \sum_{i=1}^{n} \prod_{j \neq i} \lambda_j = \sigma_{n-1}(A)$$

où $\sigma_{n-1}(A)$ désigne l'expression symétrique élémentaire des produits de $n-1$ facteurs parmi les racines distinctes ou confondues de χ_A. En posant $\chi_A = X^n + a_{n-1}X^{n-1} + \cdots + a_1X + a_0$, on obtient $\sigma_{n-1}(A) = (-1)^{n-1}a_1$. Puisque $\mathrm{GL}_n(\mathbb{R})$ est dense dans $\mathcal{M}_n(\mathbb{R})$, que $A \longmapsto a_1$ est continue, car a_1 est une fonction polynomiale des coefficients de A, et que l'application $A \longmapsto \operatorname{Tr}\widetilde{A}$ est également continue, la relation $\operatorname{Tr}\widetilde{A} = (-1)^{n-1}a_1$ reste valable pour une matrice non inversible.

Dans le cas où A est de rang $n-1$, le spectre de la comatrice est $\{0, (-1)^{n-1}a_1\}$ où a_1 est le coefficient de X dans χ_A. Si 0 est une racine simple de χ_A, il s'agit simplement du produit des $n-1$ autres racines. Sinon $\widetilde{A}$ ne possède qu'une valeur propre 0. ◁

1.21. La transposée de la comatrice de A est un polynôme en A

Soit $A \in \mathcal{M}_n(K)$. On note $\widetilde{A}$ la transposée de la comatrice de A. Montrer que $\widetilde{A}$ est un polynôme en A. On traitera successivement les cas $K = \mathbb{R}$, K infini, K quelconque.

(ENS Lyon)

▷ **Solution.**

• Comme nous y invite l'énoncé, prenons d'abord $K = \mathbb{R}$. D'après le théorème de Cayley-Hamilton, si $\chi_A = X^n + a_{n-1}X^{n-1} + \cdots + a_1X + a_0$, on a

$$\chi_A(A) = A^n + a_{n-1}A^{n-1} + \cdots + a_1A + a_0I_n = 0.$$

Supposons A inversible. Alors $\widetilde{A} = \det(A)A^{-1}$. Comme

$$A(A^{n-1} + a_{n-1}A^{n-2} + \cdots + a_2A + a_1I_n) = -a_0I_n = -\det(A)I_n,$$

il vient

$$\widetilde{A} = \det(A)A^{-1} = -(A^{n-1} + a_{n-1}A^{n-2} + \cdots + a_2A + a_1I_n) \in \mathbb{R}[A].$$

Le résultat est donc vrai pour une matrice inversible.

Nous allons étendre la propriété sur $\mathcal{M}_n(\mathbb{R})$ par densité du groupe linéaire $\mathrm{GL}_n(\mathbb{R})$. Soit $(\mathrm{A}_p)_{p\in\mathbb{N}}$ une suite de matrices inversibles qui converge vers A. Pour $\mathrm{M}\in\mathcal{M}_n(\mathbb{R})$, écrivons

$$\chi_{\mathrm{M}}(\mathrm{X}) = \mathrm{X}^n + a_{n-1}(\mathrm{M})\mathrm{X}^{n-1} + \cdots + a_1(\mathrm{M})\mathrm{X} + a_0(\mathrm{M}).$$

Les a_i apparaissent comme des fonctions polynomiales des coefficients de M. En particulier, ce sont des fonctions continues et par conséquent $\lim_{p\to+\infty} a_i(\mathrm{A}_p) = a_i(\mathrm{A})$. D'autre part, $\mathrm{M}\longmapsto\widetilde{\mathrm{M}}$ est continue car polynomiale. On a, pour tout p,

$$\widetilde{\mathrm{A}_p} = -(\mathrm{A}_p^{n-1} + a_{n-1}(\mathrm{A}_p)\mathrm{A}_p^{n-2} + \cdots + a_2(\mathrm{A}_p)\mathrm{A}_p + a_1(\mathrm{A}_p)\mathrm{I}_n)$$

ce qui donne en faisant tendre p vers l'infini :

$$\widetilde{\mathrm{A}} = -(\mathrm{A}^{n-1} + a_{n-1}(\mathrm{A})\mathrm{A}^{n-2} + \cdots + a_2(\mathrm{A})\mathrm{A}_p + a_1(\mathrm{A})\mathrm{I}_n) \in \mathbb{R}[\mathrm{A}].$$

Le résultat est donc prouvé pour $\mathrm{K} = \mathbb{R}$.

• Supposons maintenant K infini. Comme précédemment, pour M dans $\mathcal{M}_n(\mathbb{R})$, écrivons

$$\chi_{\mathrm{M}}(\mathrm{X}) = \mathrm{X}^n + a_{n-1}(\mathrm{M})\mathrm{X}^{n-1} + \cdots + a_1(\mathrm{M})\mathrm{X} + a_0(\mathrm{M})$$

avec les a_i fonctions polynomiales des coefficients de M. Soit $\lambda\in\mathrm{K}$. Si λ n'est pas valeur propre de A, $\mathrm{A}-\lambda\mathrm{I}_n$ est inversible et le même calcul que dans le cas réel assure :

$$\widetilde{(\mathrm{A}-\lambda\mathrm{I}_n)} = -((\mathrm{A}-\lambda\mathrm{I})^{n-1} + a_{n-1}(\mathrm{A}-\lambda\mathrm{I})(\mathrm{A}-\lambda\mathrm{I})^{n-2} + \cdots + a_1(\mathrm{A}-\lambda\mathrm{I})\mathrm{I}).$$

Il s'agit là d'une identité polynomiale en λ, vraie pour λ en dehors du spectre de A *i.e.* pour une infinité de valeurs. Le polynôme

$$\widetilde{(\mathrm{A}-X\mathrm{I}_n)} + (\mathrm{A}-X\mathrm{I})^{n-1} + a_{n-1}(\mathrm{A}-X\mathrm{I})(\mathrm{A}-X\mathrm{I})^{n-2} + \cdots + a_1(\mathrm{A}-X\mathrm{I})\mathrm{I}$$

est donc nul. En $\mathrm{X}=0$, cela donne

$$\widetilde{\mathrm{A}} = -(\mathrm{A}^{n-1} + a_{n-1}(\mathrm{A})\mathrm{A}^{n-2} + \cdots + a_2(\mathrm{A})\mathrm{A}_p + a_1(\mathrm{A})\mathrm{I}_n) \in \mathrm{K}[\mathrm{A}].$$

Le résultat est donc vrai pour K infini.

• Reste à traiter le cas où le corps K est fini. Nous avons démontré que lorsque K est infini, $\widetilde{\mathrm{A}}$ est combinaison linéaire des A^i avec des coefficients qui sont au signe près les coefficients du polynôme caractéristique de A. Si le corps K est fini, on le plonge dans un surcorps infini, par exemple

$L = K(Y)$. Alors $\widetilde{A}$ est combinaison linéaire des A^i à coefficients dans L. Plus précisément, ces coefficients sont au signe près les coefficients du polynôme caractéristique de A, donc des éléments de K. La transposée de la comatrice est bien un polynôme en A à coefficients dans K. ◁

En particulier, toute matrice qui commute avec A *commute avec* $\widetilde{A}$. *Cette question a été posée indépendamment lors d'un autre oral à l'École normale supérieure.*

1.22. Comatrice d'un produit

Soit R un anneau de caractéristique différente de 2. Pour A dans $\mathcal{M}_n(R)$ on note Com A la comatrice de A.

Montrer que $\mathrm{Com}(AB) = \mathrm{Com}\,A\,\mathrm{Com}\,B$. Avant d'aborder le cas général, on traitera successivement les cas suivants :

1. R est le corps des réels ou des complexes ;
2. R est un corps quelconque ;
3. R est un anneau intègre.

(ENS Lyon)

▷ **Solution.**

1. On sait que, si A est inversible et si R est un corps, on a l'égalité $\mathrm{Com}\,A = (\det A)\,{}^tA^{-1}$. Ainsi dans le cas où A et B sont toutes deux inversibles,

$$\begin{aligned}\mathrm{Com}(AB) &= \det(AB)\,{}^t(AB)^{-1} = (\det A \det B)({}^tB\,{}^tA)^{-1}\\ &= \det A({}^tA)^{-1}\det B({}^tB)^{-1} = \mathrm{Com}\,A\,\mathrm{Com}\,B.\end{aligned}$$

Dans le cas $R = \mathbb{R}$ (resp. $R = \mathbb{C}$), comme $\mathrm{GL}_n(\mathbb{R})$ (resp. $\mathrm{GL}_n(\mathbb{C})$) est dense dans $\mathcal{M}_n(\mathbb{R})$ (resp. $\mathcal{M}_n(\mathbb{C})$) et que l'application qui à A associe Com A est continue, le résultat demeure pour toutes matrices A et B de $\mathcal{M}_n(R)$, par passage à la limite.

2. Si R est un corps quelconque on ne peut plus invoquer un argument topologique pour étendre la formule aux matrices non inversibles. On plonge alors $\mathcal{M}_n(R)$ dans $\mathcal{M}_n(R(X))$, où R(X) est le corps des fractions rationnelles à une indéterminée sur R. Les matrices $A - XI$ et $B - XI$ sont inversibles dans $\mathcal{M}_n(R(X))$ (car $\det(A - XI) = \chi_A(X) \neq 0$). D'après la question précédente,

$$\mathrm{Com}((A - XI)(B - XI)) = \mathrm{Com}(A - XI)\ \mathrm{Com}(B - XI)$$

Cette égalité matricielle équivaut à l'égalité de n^2 polynômes formels en X. Si on évalue ces polynômes en 0 on obtient l'égalité de tous les coefficients des matrices Com(AB) et Com A Com B.

3. Si R est un anneau intègre, on le plonge dans son corps des fractions pour se ramener à la question précédente. ◁

Si R *est un anneau commutatif quelconque, le résultat est encore vrai. Soit* $A = (a_{ij})$ *et* $B = (b_{ij})$ *deux matrices de* $\mathcal{M}_n(R)$. *Soient* $(X_{ij})_{1\leqslant i,j\leqslant n}$ *et* $(Y_{ij})_{1\leqslant i,j\leqslant n}$ $2n^2$ *indéterminées. Les matrices* $M = (X_{ij})_{1\leqslant i,j\leqslant n}$ *et* $N = (Y_{ij})_{1\leqslant i,j\leqslant n}$ *sont à coefficients dans l'anneau intègre* R′ *des polynômes en* X_{ij} *et* Y_{ij} *à coefficients dans* $\mathbb{Z}$. *D'après ce qui précède on a* $\mathrm{Com}(MN) - \mathrm{Com}\,M\,\mathrm{Com}\,N = 0$. *Par la propriété universelle de* R′ *il existe un unique morphisme d'anneau* ψ *de* R′ *dans* R *tel que* $\psi(X_{ij}) = a_{ij}$ *et* $\psi(Y_{ij}) = b_{ij}$ *pour tout* (i,j). *L'identité* $\mathrm{Com}(MN) - \mathrm{Com}\,M\,\mathrm{Com}\,N = 0$ *équivaut à la nullité de* n^2 *polynômes de* R′. *Les images de ces polynômes par* ψ *sont également nulles, ce qui conduit à la nullité de tous les coefficients de la matrice* $\mathrm{Com}(AB) - \mathrm{Com}\,A\,\mathrm{Com}\,B$.

On pourra utiliser le résultat qui vient d'être démontré dans l'exercice suivant.

1.23. Équation matricielle faisant intervenir la comatrice

Soit $n \geqslant 2$ un entier. Déterminer les matrices $A \in \mathcal{M}_n(\mathbb{C})$ telles que $A + \widetilde{A}$ soit scalaire, où $\widetilde{A}$ est la transposée de la comatrice de A.
(**ENS Ulm, École polytechnique**)

▷ **Solution.**

Notons pour tout $\lambda \in \mathbb{C}$, E_λ l'ensemble des matrices $A \in \mathcal{M}_n(\mathbb{C})$ telles que $A + \widetilde{A} = \lambda I_n$. Commençons par une observation fondamentale : si A est dans E_λ, toute la classe de similitude de A est incluse dans E_λ. En effet, si $B = P^{-1}AP$ avec $P \in \mathrm{GL}_n(\mathbb{C})$, on a, d'après l'exercice 1.22,

$$\begin{aligned}\widetilde{B} &= {}^t\mathrm{Com}(P^{-1}AP) = {}^t(\mathrm{Com}\,P^{-1}\ \mathrm{Com}\,A\ \mathrm{Com}\,P)\\ &= \widetilde{P}\widetilde{A}\widetilde{P^{-1}} = (\det P)P^{-1}\widetilde{A}(\det P^{-1})P\\ &= P^{-1}\widetilde{A}P.\end{aligned}$$

Il en résulte donc que $B + \widetilde{B} = P^{-1}(A + \widetilde{A})P = \lambda I_n$, ce qui prouve le résultat annoncé.

On se donne dans la suite une matrice $A \in \mathcal{M}_n(\mathbb{C})$ et un scalaire $\lambda \in \mathbb{C}$ tels que $A + \widetilde{A} = \lambda I_n$. Observons qu'en multipliant cette égalité par A, il vient $A^2 - \lambda A + \det(A)I = 0$. On va mener l'analyse en discutant selon le rang de A. On fera la synthèse directement dans chaque cas.

• Supposons $\mathrm{rg}(A) < n-1$. Alors tous les mineurs de taille $n-1$ de A sont nuls de sorte que $\widetilde{A} = 0$. La matrice A est alors scalaire et donc nulle puisque son rang n'est pas égal à n. Réciproquement, la matrice nulle est bien dans E_0.

• Supposons que $\mathrm{rg}(A) = n-1$ et que le scalaire λ n'est pas nul. On a alors $A^2 = \lambda A$ de sorte que la matrice A est annulée par le polynôme scindé à racines simples $X(X-\lambda)$. Elle donc diagonalisable. Comme son rang est égal à $n-1$, elle est semblable à la matrice diagonale $D = \mathrm{Diag}(\lambda, \lambda, \ldots, \lambda, 0)$. Un calcul immédiat montre que $\widetilde{D} = \mathrm{Diag}(0, 0, \ldots, 0, \lambda^{n-1})$. La matrice D est donc dans E_λ si et seulement si $\lambda = \lambda^{n-1}$, c'est-à-dire si et seulement si λ est une racine $(n-2)$-ième de l'unité (on a supposé ici $\lambda \neq 0$).

Réciproquement, pour toute racine $(n-2)$-ième de l'unité λ la classe de similitude de la matrice $\mathrm{Diag}(\lambda, \lambda, \ldots, \lambda, 0)$ est incluse dans E_λ. On obtient donc $n-2$ classes de similitudes de ce type.

• Supposons toujours que $\mathrm{rg}(A) = n-1$ mais avec $\lambda = 0$. On a alors $A^2 = 0$ c'est-à-dire $\mathrm{Im}\, A \subset \mathrm{Ker}\, A$. Cela impose $2\,\mathrm{rg}(A) \leqslant n$ c'est-à-dire $2(n-1) \leqslant n$, soit encore $n \leqslant 2$. Donc E_0 ne contient pas de matrice de rang $n-1$ lorsque $n > 2$. Pour $n = 2$, on a, en posant $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, $\widetilde{A} = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$ de sorte que $A + \widetilde{A} = \mathrm{Tr}(A) I_n$. Ainsi, pour $n = 2$ le problème est entièrement résolu : toutes les matrices sont solutions, A étant dans $E_{\mathrm{Tr}\, A}$. Dans la suite de la discussion on supposera donc $n \geqslant 3$.

• Il reste pour finir le cas où A est inversible. Notons Δ le discriminant du polynôme $P(X) = X^2 - \lambda X + \det A$.

∗ Si $\Delta \neq 0$, P est scindé à racines simples et la matrice A est diagonalisable. Notons $\lambda_1 \neq \lambda_2$ les deux racines de P et $p \in [\![1, n-1]\!]$ la dimension du sous-espace propre associé à λ_1. La matrice A est semblable à $D = \mathrm{Diag}(\lambda_1, \ldots, \lambda_1, \lambda_2, \ldots, \lambda_2)$, où λ_1 apparaît p fois et λ_2 apparaît $n-p$ fois. On a les relations $\lambda_1 + \lambda_2 = \lambda$ et $\lambda_1 \lambda_2 = \det A$ (relation coefficients-racines pour l'équation du second degré). Par ailleurs, on a aussi $\det A = \det D = \lambda_1^p \lambda_2^{n-p}$. Il en résulte que λ_1 et λ_2 doivent vérifier la relation $\lambda_1^{1-p} = \lambda_2^{n-p-1}$.

Étudions la réciproque. Soit λ_1 un nombre complexe non nul quelconque et p un entier naturel compris entre 1 et $n-1$. Soit λ_2 un nombre complexe distinct de λ_1 vérifiant $\lambda_2^{n-p-1} = \lambda_1^{1-p}$. Si $p \neq n-1$, il y a $n-p-1$ possibilités pour λ_2 sauf lorsque λ_1 est une racine $(n-2)$-ième de l'unité où il y en a une de moins. Lorsque $p = n-1$, λ_1 doit nécessairement être une racine $(n-1)$-ième de 1 et λ_2 est un complexe non nul quelconque, différent de λ_1. Posons $D = \mathrm{Diag}(\lambda_1, \ldots, \lambda_1, \lambda_2, \ldots, \lambda_2)$

où λ_1 apparaît p fois et λ_2 apparaît $n-p$ fois. On a, grâce à la relation vérifiée par λ_1 et λ_2,

$$\widetilde{D} = (\det D)D^{-1} = \lambda_1\lambda_2 D^{-1} = \mathrm{Diag}(\underbrace{\lambda_2, \ldots, \lambda_2}_{p \text{ fois}}, \underbrace{\lambda_1, \ldots, \lambda_1}_{n-p \text{ fois}})$$

Et donc $D + \widetilde{D} = (\lambda_1 + \lambda_2)I_n = \lambda I_n$ avec $\lambda = \lambda_1 + \lambda_2$. Toute la classe de similitude de D est incluse dans E_λ.

$*$ Il reste à étudier le cas où $\Delta = 0$, c'est-à-dire le cas où $\lambda^2 = 4 \det A$. Le polynôme du second degré P admet alors $\mu = \frac{\lambda}{2}$ comme racine double et $(A - \mu I_n)^2 = 0$. Mais alors μ est l'unique valeur propre de A et on a donc $\det A = \mu^n$. Comme $\mu^2 = \det A$, μ doit être une racine $(n-2)$-ième de l'unité ($\mu \neq 0$ car $\det A \neq 0$). Réciproquement, soit μ une racine $(n-2)$-ième de l'unité et B une matrice quelconque vérifiant $B^2 = 0$. La matrice $A = \mu I_n + B$ est inversible, d'inverse $A^{-1} = \frac{1}{\mu} I_n - \frac{1}{\mu^2} B$. On a $\det A = \mu^n = \mu^2$. Donc, $\widetilde{A} = (\det A)A^{-1} = \mu I_n - B$ et $\widetilde{A} + A = 2\mu I_n$ est bien scalaire. $\lhd$

1.24. Expression de la transposée de la comatrice de $XI_n - A$

Soit K un corps commutatif et $A \in \mathcal{M}_n(K)$. On note ${}^t\mathrm{Com}(XI_n - A)$ la matrice transposée de la matrice des cofacteurs de $XI_n - A$. Montrer qu'il existe une base $(U_0, U_1, \ldots, U_{n-1})$ de $K_{n-1}[X]$ telle que ${}^t\mathrm{Com}(XI_n - A) = \sum_{i=0}^{n-1} U_i(X)A^i$.

(ENS Ulm)

$\triangleright$ **Solution.**

Chaque cofacteur de la matrice $XI_n - A$ est un polynôme en X de degré au plus $n-1$, ce qui montre l'existence de matrices $B_0, \ldots, B_{n-1}$ telles que

$${}^t\mathrm{Com}(XI_n - A) = B_0X^{n-1} + B_1X^{n-2} + \cdots + B_{n-1}.$$

Cette décomposition est unique car, si on écrit $B_k = (b_{i,j}^k)$, pour tout $(i,j) \in [\![1,n]\!]^2$, la décomposition dans la base canonique de $K_{n-1}[X]$ du coefficient d'indice (i,j) de ${}^t\mathrm{Com}(XI_n - A)$ est $b_{i,j}^0X^{n-1} + \cdots + b_{i,j}^{n-1}$.

Nous allons montrer que les matrices B_k sont des polynômes en A en utilisant la relation fondamentale vérifiée par la comatrice. Écrivons $\chi_A = a_nX^n + a_{n-1}X^{n-1} + \cdots + a_1X + a_0$ le polynôme caractéristique

de A, avec $a_n = 1$. La relation

$$^t\mathrm{Com}(\mathrm{XI}_n - \mathrm{A})(\mathrm{XI}_n - \mathrm{A}) = \det(\mathrm{XI}_n - \mathrm{A})\mathrm{I}_n$$

s'écrit

$$\mathrm{B}_0\mathrm{X}^n + (\mathrm{B}_1 - \mathrm{B}_0\mathrm{A})\mathrm{X}^{n-1} + \cdots + (\mathrm{B}_{n-1} - \mathrm{B}_{n-2}\mathrm{A})\mathrm{X} - \mathrm{B}_{n-1}\mathrm{A}$$
$$= \chi_\mathrm{A}(\mathrm{X})\mathrm{I}_n = a_n\mathrm{X}^n\mathrm{I}_n + a_{n-1}\mathrm{X}^{n-1}\mathrm{I}_n + \cdots + a_0\mathrm{I}_n,$$

ce qui par unicité de la décomposition donne

$$\mathrm{B}_0 = a_n\mathrm{I}_n \quad \text{et} \quad \text{pour } 1 \leqslant k \leqslant n-1,\ \mathrm{B}_k - \mathrm{B}_{k-1}\mathrm{A} = a_{n-k}\mathrm{I}_n.$$

On en déduit $\mathrm{B}_0 = a_n\mathrm{I}_n$, $\mathrm{B}_1 = a_n\mathrm{A} + a_{n-1}\mathrm{I}_n$ et, plus généralement, par une récurrence immédiate, pour $0 \leqslant k \leqslant n-1$,

$$\mathrm{B}_k = a_n\mathrm{A}^k + a_{n-1}\mathrm{A}^{k-1} + \cdots + a_{n-k+1}\mathrm{A} + a_{n-k}\mathrm{I}_n = \sum_{i=0}^{k} a_{n-k+i}\mathrm{A}^i.$$

On obtient donc

$$^t\mathrm{Com}(\mathrm{XI}_n - \mathrm{A}) = \sum_{k=0}^{n-1} \mathrm{B}_{n-1-k}\mathrm{X}^k = \sum_{k=0}^{n-1}\sum_{i=0}^{n-k-1} a_{k+i+1}\mathrm{A}^i\mathrm{X}^k$$
$$= \sum_{i=0}^{n-1}\left(\sum_{k=0}^{n-i-1} a_{k+i+1}\mathrm{X}^k\right)\mathrm{A}^i.$$

C'est la décomposition cherchée avec $\mathrm{U}_i(\mathrm{X}) = \sum_{k=0}^{n-i-1} a_{k+i+1}\mathrm{X}^k$. Chaque polynôme U_i est de degré $n-i-1$ car le coefficient de X^{n-i-1} est $a_n = 1$. Ces n polynômes étant échelonnés en degré, ils forment une famille libre donc une base de $\mathrm{K}_{n-1}[\mathrm{X}]$. ◁

Dans l'exercice suivant, on détermine la différentielle du déterminant. La relation fondamentale vérifiée par la comatrice est l'outil essentiel.

1.25. Différentielle du déterminant

1. Prouver que l'application $\varphi : \mathrm{GL}_n(\mathbb{R}) \to \mathcal{M}_n(\mathbb{R})$ qui à X associe $(\det \mathrm{X})\mathrm{X}^{-1}$ admet un et un seul prolongement continu $\bar{\varphi}$ à $\mathcal{M}_n(\mathbb{R})$.

2. Soient A, B dans $\mathcal{M}_n(\mathbb{R})$. Prouver que

$$\left(\frac{d}{dt}(\det(\mathrm{A} + t\mathrm{B}))\right)_0 = \mathrm{Tr}(\bar{\varphi}(\mathrm{A})\mathrm{B}).$$

(École polytechnique)

▷ **Solution.**

1. Pour tout $X \in \mathcal{M}_n(\mathbb{R})$, notons $\widetilde{X}$ la transposée de la comatrice de X. On a la relation $\widetilde{X}X = X\widetilde{X} = \det X I_n$. En particulier, si X est inversible, alors $\widetilde{X} = (\det X)X^{-1} = \varphi(X)$. L'application $\overline{\varphi} : X \longmapsto \widetilde{X}$ est donc un prolongement de φ à $\mathcal{M}_n(\mathbb{R})$. Elle est continue, car les coefficients de $\widetilde{X}$ sont des fonctions polynomiales donc continues des coefficients de X.

L'application $\overline{\varphi}$ est le seul prolongement continu de φ à $\mathcal{M}_n(\mathbb{R})$, car $\mathrm{GL}_n(\mathbb{R})$ est dense dans $\mathcal{M}_n(\mathbb{R})$.

2. Considérons la fonction $f : t \in \mathbb{R} \longmapsto \det(A + tB) \in \mathbb{R}$. Il s'agit de déterminer $f'(0)$.

Notons $C_1, C_2, \ldots, C_n$ (respectivement $C'_1, C'_2, \ldots, C'_n$) les vecteurs colonnes de A (respectivement de B). On a, pour tout réel t,

$$f(t) = \det(C_1 + tC'_1, \ldots, C_n + tC'_n).$$

Par multilinéarité du déterminant, f est une fonction polynôme. Son nombre dérivée en 0 est égal au coefficient du terme de degré 1. Par multilinéarité, on trouve

$$f'(0) = \sum_{j=1}^{n} \det(C_1, \ldots, C_{j-1}, C'_j, C_{j+1}, \ldots, C_n).$$

Calculons $\det(C_1, \ldots, C_{j-1}, C'_j, C_{j+1}, \ldots, C_n)$ en le développant selon la $j^{\text{ème}}$ colonne. On obtient $\sum_{i=1}^{n} b_{ij}\Delta_{ij}$, où Δ_{ij} est le cofacteur du coefficient a_{ij} de la matrice A. On a donc $f'(0) = \sum_{1 \leqslant i,j \leqslant n} \Delta_{ij} b_{ij}$. Comme Δ_{ij} est le terme d'indice (j,i) de $\overline{\varphi}(A)$, $\sum_{i=1}^{n} \Delta_{ij} b_{ij}$ est le terme d'indice jj de $\overline{\varphi}(A)B$ et donc $f'(0) = \mathrm{Tr}(\overline{\varphi}(A)B)$. On conclut :

$$\boxed{\left(\frac{d}{dt}(\det(A + tB))\right)_0 = \mathrm{Tr}(\overline{\varphi}(A)B)}. \quad \triangleleft$$

On vient donc de prouver que la différentielle de l'application déterminant en un point A *est la forme linéaire* $M \mapsto \mathrm{Tr}(\overline{\varphi}(A)M)$.

Une propriété fondamentale du déterminant est que $\det(AB) = \det A \det B$ *si* A, B *sont deux matrices carrées de même taille. Toutefois, le produit* AB *peut être une matrice carrée sans que* A *et* B *soient carrées : il suffit que* A *soit de taille* (n,m) *et* B *de taille* (m,n). *Que dire du déterminant de* AB *dans ce cas ? La formule de Cauchy-Binet répond à cette question.*

1.26. Formule de Cauchy-Binet

Pour une matrice A et deux p-uplets d'indices $i_1 < i_2 < \cdots < i_p$ et $k_1 < k_2 < \cdots < k_p$ on note $\mathrm{A}\begin{pmatrix} i_1 i_2 \ldots i_p \\ k_1 k_2 \ldots k_p \end{pmatrix}$ le mineur de A obtenu en prenant les lignes $i_1, \ldots, i_p$ et les colonnes $k_1, \ldots, k_p$. Soit $\mathrm{A} \in \mathcal{M}_{n,m}(\mathrm{K})$. $\mathrm{B} \in \mathcal{M}_{m,n}(\mathrm{K})$ et $\mathrm{C} = \mathrm{AB}$. Montrer que si $n > m$ alors $\det \mathrm{C} = 0$. Si $n \leqslant m$ prouver que

$$\det \mathrm{C} = \sum_{1 \leqslant k_1 < k_2 < \cdots < k_n \leqslant m} \mathrm{A}\begin{pmatrix} 1\,2 \ \ldots n \\ k_1 k_2 \ldots k_n \end{pmatrix} \mathrm{B}\begin{pmatrix} k_1 k_2 \ldots k_n \\ 1\,2 \ \ldots n \end{pmatrix}$$

(ENS Ulm)

▷ **Solution.**

La matrice C est carrée de taille (n, n) : parler de son déterminant a bien un sens. Si $n > m$, $\mathrm{rg}(\mathrm{C}) \leqslant \mathrm{rg}(\mathrm{A}) \leqslant m < n$ donc $\det \mathrm{C} = 0$. Dans la suite on suppose $n \leqslant m$ et on note $\mathrm{A}_1, \ldots \mathrm{A}_m$ les colonnes de la matrice A. On a

$$\begin{aligned}
\det \mathrm{C} &= \det\left(\sum_{k_1=1}^{m} b_{k_1 1} \mathrm{A}_{k_1}, \ \ldots, \ \sum_{k_n=1}^{m} b_{k_n n} \mathrm{A}_{k_n} \right) \\
&= \sum_{(k_1,\ldots,k_n) \in [\![1.m]\!]^n} \mathrm{A}\begin{pmatrix} 1\,2 \ \ldots n \\ k_1 k_2 \ldots k_n \end{pmatrix} b_{k_1 1} b_{k_2 2} \ldots b_{k_n n} \\
&= \sum_{k_1 < \cdots < k_n} \sum_{\sigma \in \mathcal{S}_n} \underbrace{\mathrm{A}\begin{pmatrix} 1\,2 \ \ldots n \\ k_{\sigma(1)} k_{\sigma(2)} \ldots k_{\sigma(n)} \end{pmatrix}}_{= \varepsilon(\sigma) \mathrm{A}\left(\begin{smallmatrix} 1\,2 \ \ldots n \\ k_1 k_2 \ldots k_n \end{smallmatrix} \right)} b_{k_{\sigma(1)} 1} b_{k_{\sigma(2)} 2} \ldots b_{k_{\sigma(n)} n} \\
&= \sum_{k_1 < \cdots < k_n} \mathrm{A}\begin{pmatrix} 1\,2 \ \ldots n \\ k_1 k_2 \ldots k_n \end{pmatrix} \sum_{\sigma \in \mathcal{S}_n} \varepsilon(\sigma) b_{k_{\sigma(1)} 1} b_{k_{\sigma(2)} 2} \ldots b_{k_{\sigma(n)} n} \\
&= \sum_{1 \leqslant k_1 < k_2 < \cdots < k_n \leqslant m} \mathrm{A}\begin{pmatrix} 1\,2 \ \ldots n \\ k_1 k_2 \ldots k_n \end{pmatrix} \mathrm{B}\begin{pmatrix} k_1 k_2 \ldots k_n \\ 1\,2 \ \ldots n \end{pmatrix}.
\end{aligned}$$

D'où le résultat. ◁

On a regroupé ci-après les calculs de déterminants par blocs. Le seul résultat à connaître concerne les matrices triangulaires par blocs : le déterminant est alors égal au produit des déterminants des blocs diagonaux. Dans la plupart des cas on essaie de se ramener à cette situation.

1.27. Déterminant d'une matrice par blocs (1)

Soient A, B, C et D quatre matrices de $\mathcal{M}_n(\mathbb{C})$ telles que D et C commutent. Calculer $\det \begin{pmatrix} A & B \\ C & D \end{pmatrix}$.

(**École polytechnique**)

▷ **Solution.**

L'idée est d'écrire M comme produit de deux matrices triangulaires par blocs (l'une supérieure et l'autre inférieure) dont on pourra calculer facilement le déterminant. On peut essayer par exemple de trouver des blocs U, V, W de taille n tels que

$$M = \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} U & V \\ 0 & D \end{pmatrix} \begin{pmatrix} I & 0 \\ W & I \end{pmatrix}.$$

Pour cela, il suffit que $U + VW = A$, $V = B$ et $DW = C$. Supposons d'abord que D est inversible. Il suffit de prendre $W = D^{-1}C$, $V = B$ et $U = A - BD^{-1}C$ pour que la décomposition soit valide et dans ces conditions

$$\det M = \det(A - BD^{-1}C) \det D = \det(AD - BD^{-1}CD) = \det(AD - BC)$$

car D et C commutent.

Pour étendre le résultat au cas où D n'est pas inversible nous proposons trois méthodes :

• Première méthode : par passage à la limite. On se donne une suite (D_k) formée de matrices inversibles commutant avec C et qui converge vers D. On peut prendre par exemple $D_k = D - \frac{1}{k}I$: la matrice D_k est inversible si $\frac{1}{k}$ est inférieur à la plus petite valeur propre réelle > 0 de D (s'il en existe). Posons $M_k = \begin{pmatrix} A & B \\ C & D_k \end{pmatrix}$. La suite $(\det M_k)$ converge vers $\det M$ par continuité du déterminant. Or d'après ce qui précède,

$$\det M_k = \det(AD_k - BC) \xrightarrow[k\to\infty]{} \det(AD - BC).$$

Par unicité de la limite, on a bien $\det M = \det(AD - BC)$.

• Deuxième méthode : cette méthode n'est pas très éloignée de la précédente, mais elle n'utilise pas de passage à la limite. Pour tout $\lambda \in \mathbb{C}$, $D - \lambda I$ commute avec C et si λ est en dehors du spectre de D, cette matrice est inversible. Dans ce cas, on a

$$\det\begin{pmatrix} A & B \\ C & D-\lambda I \end{pmatrix} = \det(A(D-\lambda I) - BC)$$

Puisque le spectre de D est fini, le polynôme

$$P = \det\begin{pmatrix} A & B \\ C & D-XI \end{pmatrix} - \det(A(D-XI) - BC)$$

s'annule en une infinité de valeurs, il est donc nul. L'identité $P(0) = 0$ donne l'identité recherchée. Cette méthode est encore valable lorsqu'on remplace $\mathbb{C}$ par n'importe quel corps commutatif infini.

• Troisième méthode : par plongement de $\mathcal{M}_{2n}(\mathbb{C})$ dans $\mathcal{M}_{2n}(\mathbb{C}(X))$. On pose $M(X) = \begin{pmatrix} A - XI_n & B \\ C & D - XI_n \end{pmatrix}$: c'est une matrice de taille $2n$ à coefficients dans le corps $\mathbb{C}(X)$. La matrice $D - XI_n$ commute encore avec C. D'autre part, $\det(D - XI_n) = (-1)^n \chi_{D(X)} \neq 0$ puisque $\chi_{D(X)}$ est de degré n unitaire. En particulier, $D - XI_n$ est inversible, c'est-à-dire $D - XI_n \in GL_n(\mathbb{C}(X))$. Par conséquent

$$\det M(X) = \det((A - XI_n)(D - XI_n) - BC).$$

En faisant $X = 0$, on trouve $\det M = \det M(0) = \det(AD - BC)$. L'avantage de cette solution est qu'elle reste valable lorsqu'on remplace $\mathbb{C}$ par n'importe quel corps commutatif. ◁

Voici une formule qui généralise le résultat de l'exercice précédent.

1.28. Déterminant d'une matrice par blocs (2) : formule de Williamson

Soit $(A_{ij})_{1 \leqslant i,j \leqslant m}$ une famille de matrices de $\mathcal{M}_n(\mathbb{C})$, qui commutent deux à deux. On considère les matrices

$$A = \begin{pmatrix} A_{11} & \dots & A_{1m} \\ \vdots & & \vdots \\ A_{m1} & \dots & A_{mm} \end{pmatrix} \in \mathcal{M}_{nm}(\mathbb{C}) \quad \text{et} \quad B = \sum_{\sigma \in \mathcal{S}_m} \varepsilon(\sigma) \prod_{i=1}^{m} A_{\sigma(i)i}.$$

Montrer que $\det A = \det B$.

(École polytechnique)

▷ **Solution.**

On va procéder par récurrence sur m le résultat étant évident pour $m = 1$ puisqu'on a alors $A = B = A_{11}$. Supposons donc le résultat vrai au rang $m-1$ avec $m \geqslant 2$. Comme d'habitude avec des déterminants par blocs on va essayer de se ramener à une matrice triangulaire par blocs pour utiliser ensuite l'hypothèse de récurrence.

Supposons tout d'abord que A_{mm} est inversible. En s'inspirant des étapes élémentaires dans l'algorithme du pivot de Gauss, il est facile d'annuler tous les blocs de la dernière colonne de A excepté le bloc diagonal. Il suffit pour cela de multiplier A à gauche par la matrice

$$M = \begin{pmatrix} I_n & 0 & \dots & 0 & -A_{1m}A_{mm}^{-1} \\ 0 & I_n & 0 & \dots & -A_{2m}A_{mm}^{-1} \\ \vdots & & \ddots & & \vdots \\ 0 & & & \ddots & -A_{m-1m}A_{mm}^{-1} \\ 0 & \dots & & 0 & I_n \end{pmatrix}.$$

Notons C_{ij} les blocs qui interviennent dans la matrice $C = MA$: on a $C_{im} = 0$ pour tout $i < m$, $C_{mj} = A_{mj}$ pour tout $j \in [\![1, m]\!]$ et enfin $C_{ij} = A_{ij} - A_{im}A_{mm}^{-1}A_{mj}$ pour $(i,j) \in [\![1, m-1]\!]^2$. Comme M est de déterminant 1, on a

$$\det A = \det C = \det A_{mm} \det(C_{ij})_{1 \leqslant i,j \leqslant m-1}.$$

Or les matrices C_{ij} commutent deux à deux et il est donc possible d'appliquer l'hypothèse de récurrence à la matrice C. On a donc

$$\det A = \det A_{mm} \det \left(\sum_{\sigma \in \mathcal{S}_{m-1}} \varepsilon(\sigma) \prod_{i=1}^{m-1} C_{\sigma(i)i} \right).$$

Notons R le sous-anneau de $\mathcal{M}_n(\mathbb{C})$ engendré par les matrices A_{ij}. C'est un anneau commutatif (pas forcément intègre). La matrice B représente en fait le déterminant de la matrice A regardée comme une matrice de taille m à coefficients dans l'anneau R. Si on regarde le calcul précédent comme étant effectué dans $\mathcal{M}_m(R)$, on constate que

$$A_{mm} \sum_{\sigma \in \mathcal{S}_{m-1}} \varepsilon(\sigma) \prod_{i=1}^{m-1} C_{\sigma(i)i} = \sum_{\sigma \in \mathcal{S}_m} \varepsilon(\sigma) \prod_{i=1}^{m} A_{\sigma(i)i} = B.$$

On a donc bien $\det A = \det B$.

Pour le cas général on peut conclure en utilisant le densité du groupe linéaire puisque les deux termes de l'égalité dépendent continûment de la matrice A_{mm}. ◁

1.29. Déterminant d'une matrice par blocs (3)

Soit $M = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$ une matrice carrée complexe inversible avec A et D elles aussi carrées. On écrit $M^{-1} = \begin{pmatrix} A' & B' \\ C' & D' \end{pmatrix}$ avec le même découpage. Trouver une relation entre $\det M$, $\det D$ et $\det A'$. **(ENS Ulm)**

▷ **Solution.**

On notera m la taille de A et n celle de D. Pour avoir une idée du résultat regardons le cas particulier où $C = 0$. La matrice M est alors triangulaire par blocs et il vient $\det M = \det A \det D$. Nécessairement A est inversible et d'inverse A'. On a donc $\det D = \det M \det A'$. Nous allons tâcher de montrer ce résultat pour C quelconque.

Comme $M^{-1}M = I_{m+n}$, on peut écrire

$$\begin{pmatrix} I_m & 0 \\ 0 & I_n \end{pmatrix} = \begin{pmatrix} A' & B' \\ C' & D' \end{pmatrix} \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \left(\begin{array}{c|c} A'A + B'C & A'B + B'D \\ \hline C'A + D'C & C'B + D'D \end{array}\right).$$

On a donc, en particulier, $A'B + B'D = 0$ et $C'B + D'D = I_n$.

Compte-tenu de ces relations, calculons

$$\begin{aligned} M^{-1} \times \begin{pmatrix} I_m & B \\ 0 & D \end{pmatrix} &= \begin{pmatrix} A' & B' \\ C' & D' \end{pmatrix} \begin{pmatrix} I_m & B \\ 0 & D \end{pmatrix} \\ &= \left(\begin{array}{c|c} A' & A'B + B'D \\ \hline C' & C'B + D'D \end{array}\right) \\ &= \begin{pmatrix} A' & 0 \\ C' & I_n \end{pmatrix}. \end{aligned}$$

En passant aux déterminants on a

$$\det\left(M^{-1} \times \begin{pmatrix} I_m & B \\ 0 & D \end{pmatrix}\right) = \det M^{-1} \det D = \det A' \det I_n,$$

d'où finalement $\boxed{\det M \det A' = \det D}$. ◁

1.30. Déterminant d'une matrice par blocs (4)

Soit $n \geqslant 3$, $(a,b,c,d) \in \mathbb{C}^4$ et $\mathrm{N} \in \mathcal{M}_{n-2}(\mathbb{C})$. On considère la matrice $\mathrm{M} = \begin{pmatrix} a & * & b \\ * & \mathrm{N} & * \\ c & * & d \end{pmatrix} \in \mathcal{M}_n(\mathbb{C})$, ainsi que $\mathrm{A} = \begin{pmatrix} a & * \\ * & \mathrm{N} \end{pmatrix}$, $\mathrm{B} = \begin{pmatrix} * & b \\ \mathrm{N} & * \end{pmatrix}$, $\mathrm{C} = \begin{pmatrix} * & \mathrm{N} \\ c & * \end{pmatrix}$ et $\mathrm{D} = \begin{pmatrix} \mathrm{N} & * \\ * & d \end{pmatrix}$.

1. Montrer que $\det \mathrm{A} = \det \mathrm{B} = 0$ implique $\det \mathrm{M} \det \mathrm{N} = 0$.

2. Montrer que $\det \mathrm{M} \det \mathrm{N} = \det \mathrm{A} \det \mathrm{D} - \det \mathrm{B} \det \mathrm{C}$.

(ENS Lyon)

▷ **Solution.**

1. Supposons que $\det \mathrm{A} = \det \mathrm{B} = 0$ et que N est inversible. Dans ce cas la matrice A n'est pas inversible, donc de rang $\leqslant n-2$, mais contient une sous-matrice de taille $n-2$ inversible à savoir N. On a donc $\mathrm{rg}(\mathrm{A}) = n-2$ et si on note $\mathrm{A}_1, \ldots, \mathrm{A}_{n-1}$ les colonnes de A, alors les colonnes $\mathrm{A}_2, \ldots, \mathrm{A}_{n-1}$ engendrent l'image de A (on confond la matrice A et l'endomorphisme de $\mathbb{C}^{n-1}$ qui lui est canoniquement associé). En particulier A_1 peut s'écrire sous la forme $\mathrm{A}_1 = \lambda_2 \mathrm{A}_2 + \cdots + \lambda_{n-1}\mathrm{A}_{n-1}$. On calcule alors le déterminant de M en effectuant sur les colonnes la manipulation $\mathrm{C}_1 \leftarrow \mathrm{C}_1 - \lambda_2 \mathrm{C}_2 - \cdots - \lambda_{n-1}\mathrm{C}_{n-1}$ puis en développant selon la première colonne. Il vient :

$$\det \mathrm{M} = \begin{vmatrix} 0 & * & b \\ 0 & \mathrm{N} & * \\ c' & * & d \end{vmatrix} = (-1)^{n+1} c' \det \mathrm{B} = 0.$$

2. Supposons dans un premier temps que N est inversible. Le vecteur colonne de $\mathbb{C}^{n-2}$ qui apparaît dans la première colonne de M entre les coefficients a et c peut s'écrire comme combinaison linéaire des colonnes de N. En effectuant la manipulation correspondante sur les colonnes de M, on a

$$\det \mathrm{M} = \begin{vmatrix} a' & * & b \\ 0 & \mathrm{N} & * \\ c' & * & d \end{vmatrix}.$$

De la même manière, on peut faire apparaître des zéros entre b et d :

$$\det \mathrm{M} = \begin{vmatrix} a' & * & b' \\ 0 & \mathrm{N} & 0 \\ c' & * & d' \end{vmatrix}.$$

En manipulant sur les lignes on peut également remplacer les $*$ par des zéros :

$$\det \mathrm{M} = \begin{vmatrix} a' & 0 & b' \\ 0 & \mathrm{N} & 0 \\ c' & 0 & d' \end{vmatrix}$$

(les coefficients a', b', c', d' ne sont pas affectés par ces dernières opérations). On développe maintenant par rapport à la première colonne :

$$\begin{aligned} \det \mathrm{M} &= a' \begin{vmatrix} \mathrm{N} & 0 \\ 0 & d' \end{vmatrix} + (-1)^{n+1} c' \begin{vmatrix} 0 & b' \\ \mathrm{N} & 0 \end{vmatrix} \\ &= a'd' \det \mathrm{N} + (-1)^{n+1}(-1)^n b'c' \det \mathrm{N}. \end{aligned}$$

Mais on a $a' \det \mathrm{N} = \det \mathrm{A}$, $b' \det \mathrm{N} = \det \mathrm{B}$, $c' \det \mathrm{N} = \det \mathrm{C}$ et $d' \det \mathrm{N} = \det \mathrm{D}$, car les manipulations précédentes n'ont pas modifié les déterminants des matrices extraites A, B, C, D. En multipliant l'égalité ci-dessus par $\det \mathrm{N}$ on obtient donc

$$\boxed{\det \mathrm{M} \det \mathrm{N} = \det \mathrm{A} \det \mathrm{D} - \det \mathrm{B} \det \mathrm{C}}.$$

Par densité de $\mathrm{GL}_{n-2}(\mathbb{C})$ dans $\mathcal{M}_{n-2}(\mathbb{C})$, le résultat demeure vrai pour N quelconque. $\lhd$

Dans l'exercice suivant on calcule le déterminant d'une application linéaire sur $\mathcal{L}(\mathrm{E})$ qui fait intervenir des matrices par blocs.

1.31. Déterminant d'un automorphisme intérieur

Soit E un K-espace vectoriel de dimension finie et $g \in \mathrm{GL}(\mathrm{E})$. On considère $\varphi : u \longmapsto gug^{-1}$. Déterminer $\det \varphi$ et $\mathrm{Tr}\, \varphi$.

(**École polytechnique**)

$\rhd$ **Solution.**

On va commencer par le déterminant. Le problème devient plus simple si on le généralise. En effet, multiplier à gauche et à droite sont deux opérations indépendantes : si on note φ_1 (resp. φ_2) l'application $u \longmapsto gu$ (resp. $u \longmapsto ug^{-1}$), on a $\varphi = \varphi_1 \circ \varphi_2 = \varphi_2 \circ \varphi_1$ et par conséquent $\det \varphi = \det \varphi_1 \det \varphi_2$.

On regarde alors le problème matriciellement. Si A est une matrice de $\mathcal{M}_n(\mathrm{K})$, on notera $\mathrm{G_A} : \mathrm{M} \longmapsto \mathrm{AM}$ (resp. $\mathrm{D_A} : \mathrm{M} \longmapsto \mathrm{MA}$) l'application consistant à multiplier à gauche (resp. à droite) par A. Considérons la base canonique $(\mathrm{E}_{ij})_{1 \leqslant i,j \leqslant n}$ de $\mathcal{M}_n(\mathrm{K})$ ordonnée de la manière suivante :

$$(E_{11}, E_{21}, \ldots, E_{n1}, E_{12}, E_{22}, \ldots, E_{n2}, \ldots, E_{1n}, E_{2n}, \ldots, E_{nn}).$$

Si $A = (a_{ij})_{1 \leqslant i,j \leqslant n}$, on a :

$$G_A(E_{ij}) = AE_{ij} = a_{1i}E_{1j} + \cdots + a_{ni}E_{nj}.$$

Il en résulte que la matrice de G_A est une matrice diagonale par blocs formée de n matrices A. On a donc $\det G_A = (\det A)^n$.

Pour la multiplication à droite, il est inutile de faire une étude supplémentaire. En effet, notons $T : M \longmapsto {}^tM$ la transposition. C'est un isomorphisme involutif de $\mathcal{M}_n(K)$. Pour toutes matrices A et M on a $T(AM) = T(M)T(A)$, c'est-à-dire que $T \circ G_A = D_{{}^tA} \circ T$. En passant au déterminant on a donc $\det G_A = \det D_{{}^tA}$ (car $\det T \neq 0$). Une matrice et sa transposée ayant même déterminant, on en déduit que $\det D_A = (\det A)^n$.

Dans le cas qui nous intéresse, on choisit une base de E et on on note A la matrice de g dans une base. Le déterminant de φ est celui de $G_A \circ D_{A^{-1}}$. On a donc $\det \varphi = \det G_A \det D_{A^{-1}} = (\det A)^n (\det A^{-1})^n = 1$.

Passons à la trace. Ici la décomposition de φ en $G_A \circ D_{A^{-1}}$ ne sert à rien (la trace n'est pas multiplicative !). Il n'y a guère d'autre choix que de regarder directement la matrice de φ dans la base $(E_{ij})_{1 \leqslant i,j \leqslant n}$. On pose $A^{-1} = (b_{ij})_{1 \leqslant i,j \leqslant n}$. On a

$$\varphi(E_{ij}) = AE_{ij}A^{-1} = \sum_{1 \leqslant r,s \leqslant n} a_{ri} b_{js} E_{rs}.$$

La coordonnée de $\varphi(E_{ij})$ selon E_{ij} vaut donc $a_{ii}b_{jj}$. Il en résulte que

$$\operatorname{Tr} \varphi = \sum_{1 \leqslant i,j \leqslant n} a_{ii}b_{jj} = \operatorname{Tr} A \operatorname{Tr} A^{-1} = \boxed{\operatorname{Tr} g \operatorname{Tr} g^{-1}}. \lhd$$

Les exercices suivants concernent des matrices à coefficients entiers. Le premier se rattache à la théorie des ensembles.

1.32. Déterminant de la matrice d'incidence des parties non vides d'un ensemble

Soit $I = \{1, 2, ..., n\}$, $(A_i)_{1 \leqslant i \leqslant 2^n - 1}$ les parties non vides de I numérotées dans un ordre quelconque et $B \in \mathcal{M}_{2^n-1}(\mathbb{R})$ la matrice définie par $b_{ij} = 1$ si $A_i \cap A_j \neq \emptyset$ et $b_{ij} = 0$ sinon.

Montrer que le déterminant de B ne dépend pas de la numérotation choisie et le calculer.

▷ **Solution.**

Changer la numérotation des parties non vides de I revient à considérer une permutation σ de $\{1, \ldots, 2^n - 1\}$ et à noter $(A_{\sigma(i)})_{1 \leqslant i \leqslant 2^n - 1}$ les parties non vides de I. On considère alors la matrice $B' \in \mathcal{M}_{2^n - 1}(\mathbb{R})$ définie par $b'_{ij} = 1$ si $A_{\sigma(i)} \cap A_{\sigma(j)} \neq \emptyset$ et $b'_{ij} = 0$ sinon. On a donc $b'_{ij} = b_{\sigma(i),\sigma(j)}$. La matrice de permutation $P_\sigma = (p_{ij}) \in \mathcal{M}_{2^n - 1}(\mathbb{R})$ définie par $p_{ij} = 1$ si $\sigma(i) = j$ et 0 sinon, vérifie $(P_\sigma)^{-1} = P_{\sigma^{-1}} = {}^tP_\sigma$. On en déduit que

$$b'_{ij} = b_{\sigma(i),\sigma(j)} = \sum_{1 \leqslant k,l \leqslant 2^n - 1} p_{ik} b_{kl} p_{jl},$$

ce qui montre que $B' = P_\sigma B (P_\sigma)^{-1}$. Les matrices B et B′ sont semblables, donc ont même déterminant.

On note désormais B_n une des matrices précédentes et D_n son déterminant. On le calcule en raisonnant par récurrence. On a clairement $D_1 = 1$. Pour un entier $n \geqslant 2$, on numérote les parties non vides de I de la façon suivante : d'abord les parties ne contenant pas n, puis les parties contenant n, la dernière étant $\{n\}$. On suppose de plus que, pour $1 \leqslant i \leqslant 2^{n-1} - 1$, on a $A_{i+2^{n-1}-1} = A_i \cup \{n\}$. On obtient

$$B_n = \left(\begin{array}{ccc|c} & & & 0 \\ & B_{n-1} & & \vdots \\ & & & 0 \\ \hline & & & 1 \\ & B_{n-1} & & \vdots \\ & & & 1 \\ \hline 0 & \ldots & 0 \;|\; 1 & \ldots \quad 1 & 1 \end{array}\right),$$

où $J \in M_{2^{n-1}-1}(\mathbb{R})$ est la matrice dont tous les coefficients sont égaux à 1.

En soustrayant la dernière colonne aux $2^{n-1} - 1$ précédentes, il vient

$$D_n = \det \left(\begin{array}{ccc|c} & & & 0 \\ & B_{n-1} & & \vdots \\ & & & 0 \\ \hline & & & 1 \\ & B_{n-1} & & \vdots \\ & & & 1 \\ \hline 0 & \ldots & 0 \;|\; 0 & \ldots \quad 0 & 1 \end{array}\right)$$

$$= \det \left(\begin{array}{c|c} \mathrm{B}_{n-1} & \mathrm{B}_{n-1} \\ \hline \mathrm{B}_{n-1} & 0 \end{array} \right) - \det \left(\begin{array}{c|c} \mathrm{B}_{n-1} & \mathrm{B}_{n-1} \\ \hline 0 & -\mathrm{B}_{n-1} \end{array} \right)$$
$$= \det \mathrm{B}_{n-1} \det(-\mathrm{B}_{n-1}) = -(\det \mathrm{B}_{n-1})^2 = -\mathrm{D}_{n-1}^2,$$

puisque la taille de B_{n-1} est impaire. Compte tenu de la valeur de D_1, on obtient, pour tout $n \geqslant 2$, $\mathrm{D}_n = -1$. ◁

L'exercice qui suit utilise un déterminant pour résoudre un problème de dénombrement.

1.33. Dérangements pairs et impairs

Un élément de $\mathcal{S}_n$ est appelé dérangement si c'est une permutation sans point fixe. Y a-t-il plus de dérangements impairs ou de dérangements pairs dans $\mathcal{S}_n$?

(École polytechnique)

▷ **Solution.**

En notant $\mathcal{D}_n$ l'ensemble des dérangements d'ordre n et Δ_n la différence entre le nombre de dérangements pairs et le nombre de dérangements impairs, on obtient

$$\Delta_n = \sum_{\sigma \in \mathcal{D}_n} \varepsilon(\sigma).$$

Considérons la matrice $\mathrm{A} = (a_{ij}) \in \mathcal{M}_n(\mathbb{R})$ dont tous les coefficients sont égaux à 1 sauf ceux de la diagonale qui sont nuls. On écrit $\det \mathrm{A} = \sum_{\sigma \in \mathcal{S}_n} \varepsilon(\sigma) a_{1\sigma(1)} \dots a_{n\sigma(n)}$. Le produit $a_{1\sigma(1)} \dots a_{n\sigma(n)}$ est nul s'il existe $i \in [\![1, n]\!]$ tel que $\sigma(i) = i$, c'est-à-dire si σ n'est pas un dérangement. Sinon, $a_{1\sigma(1)} \dots a_{n\sigma(n)} = 1$. On obtient donc

$$\det \mathrm{A} = \sum_{\sigma \in \mathcal{D}_n} \varepsilon(\sigma) = \Delta_n.$$

Reste à calculer le déterminant. C'est un cas particulier du calcul effectué dans l'exercice 1.4. Si on note $(e_1, \dots, e_n)$ la base canonique de $\mathbb{R}^n$ et u le vecteur $(1, \dots, 1)$, on obtient

$$\Delta_n = \det(u - e_1, \dots, u - e_n) = \det(-e_1, \dots, -e_n) + \sum_{i=1}^{n} \det(-e_1, \dots, -e_{i-1}, u, -e_{i+1}, \dots, -e_n)$$
$$= (-1)^n + n(-1)^{n-1} = (-1)^{n-1}(n-1).$$

Il y a donc plus de dérangements pairs si $n-1$ est pair, *i.e.* si n est impair. $\triangleleft$

L'énoncé qui suit présente un calcul particulièrement astucieux.

1.34. Déterminant de Smith (1875)

Soit $A = (a_{ij}) \in \mathcal{M}_n(\mathbb{R})$ où $a_{ij} = \operatorname{pgcd}(i,j)$. Calculer $\det A$. On utilisera la relation $n = \sum_{d|n} \varphi(d)$ où φ est l'indicatrice d'Euler.

(ENS Ulm)

$\triangleright$ **Solution.**

Notons que le lecteur trouvera une preuve de l'identité vérifiée par l'indicatrice d'Euler dans l'exercice 4.17 du tome 1. L'idée du calcul de $\det A$ est vraiment très astucieuse : il s'agit de décomposer A en un produit de matrices plus simples, en utilisant justement la relation $n = \sum_{d|n} \varphi(d)$. On a effectivement

$$a_{ij} = \operatorname{pgcd}(i,j) = \sum_{d|\operatorname{pgcd}(i,j)} \varphi(d) = \sum_{d|i \ d|j} \varphi(d) = \sum_{d=1}^{n} \varphi(d) b_{di} b_{dj},$$

avec $b_{kl} = 1$ si k divise l et $b_{kl} = 0$ sinon. La matrice $B = (b_{ij})$ de $\mathcal{M}_n(\mathbb{R})$ est la matrice d'incidence de la relation de divisibilité. C'est une matrice triangulaire supérieure unipotente (*i.e.* avec des 1 sur la diagonale) et la relation ci-dessus équivaut à l'identité matricielle

$$A = {}^t B \operatorname{Diag}(\varphi(1), \ldots, \varphi(n)) B.$$

En passant au déterminant, on obtient donc le joli résultat

$$\boxed{\det A = \prod_{k=1}^{n} \varphi(k)}. \ \triangleleft$$

L'introduction de la matrice B *permet aussi de calculer le déterminant de la matrice dont le terme (i,j) est le nombre de diviseurs communs à i et j.*

Dans l'énoncé suivant on étudie à quelle condition un vecteur de $\mathbb{Z}^n$ peut être complété pour former une base du réseau $\mathbb{Z}^n$. Le lecteur trouvera une autre approche de ce problème (qui n'utilise pas les déterminants) dans l'exercice 7.18 du tome 1.

1.35. Première colonne d'une matrice inversible de $\mathcal{M}_n(\mathbb{Z})$

1. À quelle condition une matrice de $\mathcal{M}_n(\mathbb{Z})$ est-elle inversible dans $\mathcal{M}_n(\mathbb{Z})$?

2. À quelle condition un vecteur de $\mathbb{Z}^n$ est-il la première colonne d'une matrice inversible de $\mathcal{M}_n(\mathbb{Z})$?

(ENS Cachan)

▷ **Solution.**

1. Soit $\mathrm{M} \in \mathcal{M}_n(\mathbb{Z})$.

• Supposons M inversible dans $\mathcal{M}_n(\mathbb{Z})$, *i.e.* $\mathrm{M} \in \mathrm{GL}_n(\mathbb{Z})$. Alors M^{-1} est à coefficients entiers. On a $\det \mathrm{M} \det \mathrm{M}^{-1} = \det \mathrm{I}_n = 1$. Or $\det \mathrm{M}$ et $\det \mathrm{M}^{-1}$ sont des entiers. Donc $\det \mathrm{M}$ est inversible dans $\mathbb{Z}$. Par conséquent, $\det \mathrm{M} = \pm 1$.

• Réciproquement, supposons que $\det \mathrm{M} = \varepsilon = \pm 1$. Notons N la transposée de la comatrice de M. C'est une matrice à coefficients entiers. On a la célèbre relation : $\mathrm{MN} = \mathrm{NM} = \varepsilon \mathrm{I}_n$ et $\mathrm{M}(\varepsilon \mathrm{N}) = (\varepsilon \mathrm{N})\mathrm{M} = \mathrm{I}_n$. Comme $\varepsilon \mathrm{N} \in \mathcal{M}_n(\mathbb{Z})$, M est inversible dans $\mathcal{M}_n(\mathbb{Z})$. On conclut :

$$\boxed{\mathrm{M} \in \mathrm{GL}_n(\mathbb{Z}) \iff \det \mathrm{M} = \pm 1}$$

Plus généralement, si A *est un anneau commutatif et* $\mathrm{M} \in \mathcal{M}_n(\mathrm{A})$, M *est un élément inversible de* $\mathcal{M}_n(\mathrm{A})$ *si et seulement si* $\det \mathrm{M}$ *est un inversible de* A.

2. Soit $\mathrm{X} = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} \in \mathbb{Z}^n$. Montrons qu'il existe une matrice inversible de $\mathcal{M}_n(\mathbb{Z})$ dont la première colonne est X si et seulement si le pgcd des a_i est égal à 1.

• La condition est nécessaire : soit A une matrice inversible de $\mathcal{M}_n(\mathbb{Z})$ dont la première colonne est X ; on calcule le déterminant de A en développant selon la première colonne : $\det \mathrm{A} = \sum_{i=1}^{n} a_i \mathrm{A}_i$, où A_i est le cofacteur de a_i. Par hypothèse, les A_i sont entiers et $\det \mathrm{A} = \pm 1$ d'après la première question, donc les a_i sont premiers entre eux d'après le théorème de Bezout.

• Montrons que la condition est suffisante en raisonnant par récurrence sur n. C'est évident pour $n = 1$: $a_1 = \pm 1$. Supposons la pro-

priété vérifiée au rang $n-1$ et considérons $X = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} \in \mathbb{Z}^n$ tel que $\operatorname{pgcd}(a_1, a_2, \ldots, a_n) = 1$. Posons $d = \operatorname{pgcd}(a_1, a_2, \ldots, a_{n-1})$ et, pour $1 \leqslant k \leqslant n-1$, $a'_k = \frac{a_k}{d}$. D'après l'hypothèse de récurrence, il existe une matrice inversible B' de $\mathcal{M}_{n-1}(\mathbb{Z})$ dont la première colonne est $\begin{pmatrix} a'_1 \\ a'_2 \\ \vdots \\ a'_{n-1} \end{pmatrix}$. En multipliant sa première colonne par d, on obtient une matrice B de première colonne $\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_{n-1} \end{pmatrix}$, de taille $n-1$ et de déterminant $\pm d$.

On a par hypothèse $\operatorname{pgcd}(d, a_n) = 1$ et donc $\operatorname{pgcd}(\det B, a_n) = 1$. Il existe $(u, v) \in \mathbb{Z}^2$ tel que $u \det B + v a_n = 1$. Cherchons une matrice A de taille n répondant au problème, de la forme

$$A = \left(\begin{array}{c|c} B & \begin{matrix} x_1 \\ \vdots \\ x_{n-1} \end{matrix} \\ \hline a_n\ 0\ \ldots\ 0 & u \end{array}\right) \quad \text{où } (x_1, x_2, \ldots, x_{n-1}) \in \mathbb{Z}^{n-1} \text{ est à choisir.}$$

Développons le déterminant de A selon la dernière ligne. Pour cela, posons :

$$B = \left(\begin{array}{c|c} \begin{matrix} a_1 \\ \vdots \\ a_{n-1} \end{matrix} & C \end{array}\right) \quad \text{et} \quad D = \left(\begin{array}{c|c} C & \begin{matrix} x_1 \\ \vdots \\ x_{n-1} \end{matrix} \end{array}\right).$$

Dans ces conditions $\det A = u \det B + (-1)^{n+1} a_n \det D$. Pour avoir $\det A = 1$, il suffit d'avoir $\det D = (-1)^{n+1} v$. Si $(x_1, \ldots, x_{n-1}) = (a'_1, \ldots, a'_{n-1})$, alors la matrice D est obtenue à partir de B' en faisant une permutation circulaire des colonnes. Alors $\det D = \pm \det B' = \pm 1$.

En prenant $(x_1, \ldots, x_{n-1}) = \varepsilon v (a'_1, \ldots, a'_{n-1})$, avec $\varepsilon = \pm 1$, on peut choisir ε pour que $\det D = (-1)^{n+1} v$ et donc $\det A = 1$. $\triangleleft$

1.36. Opération de $\mathbf{GL}_n(\mathbb{Z})$ sur le réseau $\mathbb{Z}^n$

Pour toute matrice $X \in \mathcal{M}_{n,1}(\mathbb{Z}) = \mathbb{Z}^n$ on note $\Lambda(X)$ le pgcd des coefficients de X. Pour $A \in \mathcal{M}_n(\mathbb{Z})$ montrer qu'il y a équivalence entre :

(i) $\det A = \pm 1$;

(ii) $\forall X \in \mathbb{Z}^n.\ \ \Lambda(AX) = \Lambda(X)$.

(ENS Ulm)

▷ **Solution.**

• $(i) \Longrightarrow (ii)$. Par homogénéité du pgcd, il suffit de montrer la relation (ii) lorsque $\Lambda(X) = 1$. Dans ce cas, le théorème de Bezout donne l'existence de $U \in \mathbb{Z}^n$ tel que ${}^tUX = 1$. On sait que A est inversible et que A^{-1} est dans $\mathcal{M}_n(\mathbb{Z})$ car $\det A = \pm 1$ (voir l'exercice précédent). Posons $V = {}^tA^{-1}U \in \mathbb{Z}^n$. On a alors ${}^tVAX = 1$, ce qui prouve, par la réciproque du théorème de Bezout, que $\Lambda(AX) = 1$.

• $(ii) \Longrightarrow (i)$. Montrons que la matrice A est inversible. En effet, si ce n'est pas le cas on peut trouver $X \in \mathbb{Z}^n$ non nul tel que $AX = 0$ ce qui contredit (ii). Donc $\det A \in \mathbb{Z}^*$.

Supposons par l'absurde $|\det A| \neq 1$ et soit p premier divisant $\det A$. On note $\widetilde{A} \in \mathcal{M}_n(\mathbb{Z}/p\mathbb{Z})$ la matrice obtenue à partir de A par réduction modulo p. Comme $\det \widetilde{A} = 0$, cette matrice n'est pas inversible. Il existe donc $Y \in \mathcal{M}_{n,1}(\mathbb{Z}/p\mathbb{Z}) = (\mathbb{Z}/p\mathbb{Z})^n$ non nul tel que $\widetilde{A}Y = 0$. On peut alors écrire $Y = \begin{pmatrix} \overline{x_1} \\ \overline{x_2} \\ \vdots \\ \overline{x_n} \end{pmatrix}$ avec les x_i entiers. Si $X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{Z}^n$, le nombre premier p divise $\Lambda(AX)$ mais sans diviser $\Lambda(X)$. Contradiction. ◁

L'implication $(i) \Longrightarrow (ii)$ montre que si on fait opérer le groupe $\mathrm{GL}_n(\mathbb{Z})$ sur l'ensemble $\mathbb{Z}^n$ de manière naturelle, deux points X et Y qui sont dans la même orbite vérifient $\Lambda(Y) = \Lambda(X)$. Il est naturel de se demander si cela caractérise les orbites. C'est bien le cas. En effet, si $\Lambda(X) = 1$, X est dans l'orbite du vecteur $(1.0, \ldots, 0)$ car il existe une matrice M de $\mathrm{GL}_n(\mathbb{Z})$ dont la première colonne est X (cf. exercice 1.35). Par suite, pour tout $d \in \mathbb{N}$, les éléments X de $\mathbb{Z}^n$ vérifiant $\lambda(X) = d$ sont dans l'orbite de $(d, 0, \ldots 0)$ (il suffit de multiplier par d).

L'énoncé suivant utilise aussi la réduction modulo un nombre premier.

1.37. Un problème de poids

1. Soit $A = (a_{ij}) \in \mathcal{M}_n(\mathbb{R})$ avec $a_{ii} = 0$ pour tout i et $a_{ij} \in \{\pm 1\}$ pour $i \neq j$. Si n est pair, montrer que A est inversible.

2. On dispose de $2n+1$ cailloux, $n \geqslant 1$. On suppose que chaque sous-ensemble de $2n$ cailloux peut se partager en deux paquets de n cailloux de même masse totale. Montrer que tous les cailloux ont la même masse.

(ENS Lyon)

▷ **Solution.**

1. On va montrer que le déterminant de A est un entier impair. Pour cela on passe dans $\mathbb{Z}/2\mathbb{Z}$. La classe de $\det A$ modulo 2 est égale au déterminant de

$$\begin{pmatrix} 0 & 1 & \dots & 1 \\ 1 & 0 & \dots & 1 \\ \vdots & & \ddots & \vdots \\ 1 & \dots & 1 & 0 \end{pmatrix} \in \mathcal{M}_n(\mathbb{Z}/2\mathbb{Z}).$$

En additionnant les $n-1$ premières colonnes à la dernière, on obtient une n-ième colonne ne contenant que des $n-1$ et par multilinéarité

$$\det A \equiv -\det \begin{pmatrix} 0 & 1 & \dots & 1 \\ 1 & 0 & \dots & 1 \\ \vdots & & \ddots & \vdots \\ 1 & \dots & 1 & 1 \end{pmatrix} \ [2]$$

car n est pair. En retranchant cette dernière colonne à toute les autres, on obtient

$$\det A \equiv -\det \begin{pmatrix} -1 & 0 & \dots & 0 & 1 \\ 0 & -1 & \dots & 0 & 1 \\ \vdots & & \ddots & & \vdots \\ 0 & \dots & \dots & -1 & 1 \\ 0 & \dots & \dots & 0 & 1 \end{pmatrix} \equiv -(-1)^{n-1} \equiv 1 \ [2].$$

2. Notons $x_1, \dots, x_{2n+1}$ les masses des cailloux. Soit i fixé. On peut trouver deux sous-ensembles disjoints A_i, B_i de $[\![1, 2n+1]\!] \backslash \{i\}$ de cardinal n tels que $\sum_{k \in A_i} x_k = \sum_{k \in B_i} x_k$. Cela s'écrit aussi $\sum_{j=1}^{2n+1} a_{ij} x_j = 0$ où $a_{ii} = 0$, $a_{ij} = 1$ si $j \in A_i$ et $a_{ij} = -1$ si $j \in B_j$. On dispose donc d'une matrice

$A = (a_{ij}) \in M_{2n+1}(\mathbb{R})$ de diagonale nulle et dont les autres coefficients valent ± 1 telle que $AX = 0$, avec $X = (x_1, \ldots, x_{2n+1})$. Par ailleurs, chaque ligne de A contient exactement n coefficients égaux à 1 et n coefficients égaux à -1. Le vecteur $U = (1, 1, \ldots, 1)$ est donc aussi dans $\operatorname{Ker} A$. Or, d'après la question 1, le mineur principal de taille $2n$ est non nul. Donc le rang de A est $2n$ et $\dim \operatorname{Ker} A = 1$. Les vecteurs X et U sont donc proportionnels. ◁

Le dernier thème de ce chapitre est la décomposition LU. *On dit qu'une matrice* A *admet une décomposition* LU *si elle peut s'écrire comme produit d'une matrice* L *triangulaire inférieure (lower) et d'une matrice* U *triangulaire supérieure (upper). Dans l'exercice suivant, on cherche à quelle condition une telle décomposition existe pour une matrice inversible. La méthode exposée fournit un algorithme pour déterminer* L *et* U *basée sur la méthode du pivot de Gauss. La décomposition* LU *est utilisée dans la résolution numérique d'un système linéaire. Elle ramène la résolution de* $AX = Y$ *à la résolution successive de deux systèmes triangulaires :* $LZ = Y$ *et* $UX = Z$.

1.38. Décomposition LU avec pivot de Gauss

Soit $A = (a_{ij})_{1 \leqslant i,j \leqslant n} \in \mathcal{M}_n(\mathbb{R})$.

1. Donner la matrice qui, multipliée à gauche de A, va intervertir les lignes i et j. Quel est son déterminant ?

2. Donner la matrice qui, multipliée à gauche de A, va ajouter à la i-ième ligne la j-ième ligne multipliée par $\lambda \in \mathbb{R}$. Quel est son déterminant ? son inverse ?

3. On suppose dorénavant A inversible. Montrer qu'il existe $M \in GL_n(\mathbb{R})$ telle que $MA = U$ où U est triangulaire supérieure.

4. Préciser à quelles conditions on peut écrire $A = LU$, où U (resp. L) est triangulaire supérieure (resp. inférieure).

(École polytechnique)

▷ **Solution.**

1. Soit $1 \leqslant i, j \leqslant n$, $i \neq j$. Considérons la matrice $M_{i,j}$ obtenue en échangeant les lignes i et j de la matrice I_n, c'est-à-dire

$$M_{ij} = \sum_{k \neq i,j} E_{kk} + E_{ij} + E_{ji}.$$

Pour tout couple d'indice (p, q), on obtient, $M_{ij}E_{pq} = E_{pq}$ si $p \neq i, j$, $M_{ij}E_{iq} = E_{jq}$ et $M_{ij}E_{jq} = E_{iq}$. La multiplication à gauche par M_{ij}

échange les lignes i et j des matrices de base. Par linéarité de la multiplication par M_{ij}, il en est de même pour toute matrice de $\mathcal{M}_n(\mathbb{R})$.

Comme M_{ij} est obtenue en échangeant deux lignes de I_n, on a $\det M_{ij} = -\det I_n = -1$.

2. Soit $1 \leqslant i, j \leqslant n$, $i \neq j$. Posons $T_{ij}(\lambda) = I_n + \lambda E_{ij}$. Pour tout couple d'indice (p, q), on obtient

$$T_{ij}(\lambda)E_{pq} = E_{pq} \text{ si } p \neq j \quad \text{et} \quad T_{ij}(\lambda)E_{jq} = E_{jq} + \lambda E_{iq}.$$

Ainsi, pour les matrices de base, la multiplication à gauche par $T_{ij}(\lambda)$ ajoute à la i-ième ligne de la matrice sa j-ième ligne multipliée par λ. Par linéarité, il en est de même pour toute matrice de $\mathcal{M}_n(\mathbb{R})$.

La matrice $T_{ij}(\lambda)$ est triangulaire supérieure ou inférieure selon que $i < j$ ou $i > j$, avec des 1 sur la diagonale. On a donc $\det(T_{ij}(\lambda)) = 1$. La multiplication à gauche par $T_{ij}(-\lambda)$ ajoute à la i-ième ligne d'une matrice la j-ième multipliée par $-\lambda$. On a donc $T_{ij}(-\lambda)T_{ij}(\lambda)I_n = I_n$ et $(T_{ij}(\lambda))^{-1} = T_{ij}(-\lambda)$.

3. On applique la méthode du pivot de Gauss. La matrice $A = (a_{ij})$ étant inversible, sa première colonne n'est pas nulle. Si $a_{i1} \neq 0$, la matrice $M_{1i_1}A = (b_{ij})$ vérifie $b_{11} \neq 0$. En remplaçant, pour $2 \leqslant i \leqslant n$, la ligne L_i de cette matrice par $L_i - \frac{b_{i1}}{b_{11}}L_1$, *i.e.* en multipliant à gauche par les matrices $T_{i1}\left(-\frac{b_{i1}}{b_{11}}\right)$, on obtient une matrice $A^{(1)} = (a_{ij}^{(1)})$, dont les termes de la première colonne sont nuls à part le premier.

Si on note B_1 la matrice obtenue en supprimant la première ligne et la première colonne de $A^{(1)}$, on a $\det(A^{(1)}) = a_{11}^{(1)} \det B_1 \neq 0$. La matrice B_1 est donc inversible et sa deuxième colonne possède un terme non nul $a_{i_2 2}^{(1)}$, avec $i_2 \geqslant 2$. En multipliant $A^{(1)}$ à gauche par M_{2i_2} on obtient une matrice dont le terme d'indice $(2, 2)$ est non nul. En multipliant à gauche la matrice obtenue par des matrices de la forme $T_{i2}(\lambda)$ pour $3 \leqslant i \leqslant n$, on obtient une matrice $A^{(2)}$ de la forme

$$\begin{pmatrix} a_{11}^{(2)} & a_{12}^{(2)} & \dots & a_{1n} \\ 0 & a_{22}^{(2)} & \dots & a_{2n}^{(2)} \\ 0 & 0 & & \\ \vdots & \vdots & B_2 & \\ 0 & 0 & & \end{pmatrix},$$

où $a_{11}^{(2)}$ et $a_{22}^{(2)}$ ne sont pas nuls et où B_2 est inversible.

On poursuit le procédé. On obtient une suite de matrice $A^{(1)}$, $A^{(2)}$, $\dots$, $A^{(n-1)}$ dans laquelle la matrice $A^{(p)}$ a tous ses termes d'indice (i, j)

avec $j \leqslant p$ et $i > j$ qui sont nuls. La matrice $A^{(n-1)}$ est donc triangulaire supérieure. On pose $A^{(n-1)} = U$. La matrice U a été obtenue en multipliant à gauche A par de matrices de la forme M_{ij} ou $T_{ij}(\lambda)$. Le produit de ces matrices inversibles est une matrice inversible qu'on note M. On a donc $MA = U$.

4. Pour toute matrice $A \in \mathcal{M}_n(\mathbb{R})$ et $1 \leqslant p \leqslant n$, on note $A_p \in \mathcal{M}_p(\mathbb{R})$ la matrice obtenue en gardant les p premières lignes et les p premières colonnes de la matrice A.

• Supposons qu'il existe une matrice L triangulaire inférieure et une matrice U triangulaire supérieure telles que $A = LU$. On a alors pour tout $p \in [\![1, n]\!]$, $A_p = L_p U_p$. En effet, L est de la forme $\begin{pmatrix} L_p & 0 \\ \times & \times \end{pmatrix}$ et U de la forme $\begin{pmatrix} U_p & \times \\ 0 & \times \end{pmatrix}$. Par conséquent, A est de la forme $\begin{pmatrix} L_p U_p & \times \\ \times & \times \end{pmatrix}$ et $A_p = L_p U_p$.

La matrice A est inversible, donc les matrices L et U sont inversibles. Comme elles sont triangulaires, cela signifie que leurs termes diagonaux sont non nuls. Mais alors, pour $1 \leqslant p \leqslant n$, les matrices U_p et L_p, qui sont également triangulaires, ont leurs termes diagonaux non nuls. Elles sont donc inversibles et leur produit A_p est inversible. On a donc

$$\forall p \in [\![1, n]\!], \quad \det(A_p) \neq 0.$$

• Montrons que cette condition est suffisante. Soit donc $A \in \mathcal{M}_n(\mathbb{R})$ telle que, pour tout $p \in [\![1, n]\!]$, $\det A_p \neq 0$. Reprenons la construction effectuée dans la question 3.

On a $\det A_1 = a_{11} \neq 0$. D'emblée le coefficient d'indice $(1, 1)$ est non nul. Il n'est donc pas besoin de multiplier A par une matrice du type M_{1i}. On obtient $A^{(1)}$ en multipliant A par un produit de matrices de la forme $T_{i1}(\lambda)$ avec $2 \leqslant i \leqslant n$. La matrice $A^{(1)}$ s'obtient à partir de A en effectuant des opérations élémentaires sur les lignes. Il en est de même de $A_2^{(1)}$ à partir de A_2. On a donc $\det(A_2^{(1)}) = \det A_2 \neq 0$. Comme $A_2^{(1)} = \begin{pmatrix} a_{11} & * \\ 0 & a_{22}^{(1)} \end{pmatrix}$ est triangulaire supérieure, on a $a_{22}^{(1)} \neq 0$.

D'emblée le terme d'indice $(2, 2)$ de $A^{(1)}$ n'est pas nul. La matrice $A^{(2)}$ va s'obtenir à partir de $A^{(1)}$ en multipliant à gauche seulement par des matrices de la forme T_{i2}, avec $3 \leqslant i \leqslant n$. Le raisonnement peut se poursuivre. Si $A^{(p-1)}$ est construite, elle est déduite de A par des transformations élémentaires sur les lignes. Il en est de même de $A_p^{(p-1)}$ à partir de A_p. On obtient $\det A_p^{(p-1)} = \det A_p \neq 0$. Comme par construction $A_p^{(p-1)}$ est triangulaire supérieure, on a en particulier $a_{pp}^{(p-1)} \neq 0$ et on peut se servir de ce terme comme pivot.

Au total, la matrice M est le produit de matrices de la forme $T_{ij}(\lambda)$, avec $i > j$ (en effet, à la première étape, on a $j = 1$, $i \geqslant 2$, à la seconde $j = 2$, $i \geqslant 3$, ...). Toutes ces matrices sont triangulaires inférieures, avec des 1 sur la diagonale. Il en résulte que leur produit M est triangulaire inférieure, avec des 1 sur la diagonale. La matrice est inversible et $L = M^{-1}$ est triangulaire inférieure, avec des 1 sur la diagonale. De $MA = U$, on déduit $A = LU$.

Conclusion. La matrice $A \in GL_n(\mathbb{R})$ peut s'écrire sous la forme LU, avec L triangulaire inférieure et U triangulaire supérieure si et seulement si, pour tout $p \in [\![1, n]\!]$, on a $\det A_p \neq 0$. Quand la décomposition existe, on peut supposer de plus que la diagonale de L est constituée de 1. ◁

Lorsque A *est inversible, la décomposition* $A = LU$, *quand elle existe, est unique si on impose que* L *a une diagonale de* 1. *En effet, si on a deux telles décompositions* $A = LU = L'U'$, *on a alors* $L'^{-1}L = U'U^{-1}$. *Comme* $L'^{-1}L$ *est triangulaire inférieure avec une diagonale de* 1 *et* $U'U^{-1}$ *triangulaire supérieure, on a nécessairement* $L'^{-1}L = U'U^{-1} = I_n$ *et donc* $L = L'$, $U = U'$.

Pour démontrer l'existence de la décomposition LU, *d'autres méthodes que celle du pivot existent. On peut par exemple raisonner par récurrence sur la taille* n *de la matrice, la propriété étant évidente pour* $n = 1$. *Si elle est vérifiée au rang* $n - 1$, *il existe* L_1 *et* U_1 *respectivement triangulaire inférieure et supérieure, avec des* 1 *sur la diagonale de* L_1 *telles que* $A_{n-1} = L_1U_1$. *On écrit* A *sous la forme* $A = \begin{pmatrix} A_{n-1} & Y \\ {}^tX & a \end{pmatrix}$, *où* X *et* Y *sont des matrices de* $\mathcal{M}_{n-1,1}(\mathbb{R})$ *et* $a \in \mathbb{R}$ *et on cherche* L *et* U *sous la forme* $L = \begin{pmatrix} L_1 & 0 \\ {}^tX_1 & 1 \end{pmatrix}$ *et* $U = \begin{pmatrix} U_1 & Y_1 \\ 0 & a_1 \end{pmatrix}$, *avec* $a_1 \in \mathbb{K}^*$, X_1 *et* Y_1 *dans* $\mathcal{M}_{n-1,1}(\mathbb{R})$. *On obtient*

$$LU = \begin{pmatrix} L_1U_1 & L_1Y_1 \\ {}^tX_1U_1 & {}^tX_1Y_1 + a_1 \end{pmatrix}.$$

On veut $Y = L_1Y_1$, ${}^tX = {}^tX_1U_1$ *et* ${}^tX_1Y_1 + a_1 = a$, *soit* $Y_1 = L_1^{-1}Y$, $X_1 = ({}^tU_1)^{-1}X$ *et* $a_1 = a - {}^tXU_1^{-1}L^{-1}Y = a - {}^tXA_{n-1}^{-1}Y$. *Comme* $\det A = a_1 \det U_1 \det L_1$, *on a* $a_1 \neq 0$. *Les matrices ainsi construites conviennent donc.*

Notons pour finir qu'on peut démontrer que pour toute matrice A *inversible, on peut trouver une matrice de permutation* P *telle que* PA *admette une décomposition* LU. *Le lecteur intéressé pourra faire le lien avec la décomposition de Bruhat qui a été présentée dans l'exercice 7.16 du tome 1 d'Algèbre.*

L'exercice suivant regroupe deux énoncés sur la décomposition LU *des matrices tridiagonale.*

1.39. Décomposition LU d'une matrice tridiagonale

Soit $(a_1, \ldots, a_n) \in \mathbb{R}^n$, $(b_1, \ldots, b_{n-1}) \in \mathbb{R}^{n-1}$, $(c_1, \ldots, c_{n-1}) \in \mathbb{R}^{n-1}$ et $A = (a_{ij})_{1 \leqslant i,j \leqslant n}$ définie par

$\forall i \in [\![1, n]\!]$, $a_{ii} = a_i$, $\forall i \in [\![1, n-1]\!]$, $a_{i,i+1} = b_i$, $\forall i \in [\![2, n]\!]$, $a_{i,i-1} = c_i$

et $a_{ij} = 0$ sinon. Pour $k \in [\![1, n]\!]$, on pose $\delta_k = \det(a_{ij})_{1 \leqslant i,j \leqslant k}$.

1. Établir une relation de récurrence sur les δ_k.

2. On suppose que tous les δ_k sont non nuls. Établir l'existence d'une décomposition de A sous la forme LU :

$$A = \begin{pmatrix} 1 & 0 & 0 & \ldots & 0 \\ l_1 & 1 & 0 & \ldots & 0 \\ & \ddots & \ddots & & \\ & & \ddots & \ddots & \\ 0 & 0 & \ldots & l_{n-1} & 1 \end{pmatrix} \begin{pmatrix} \delta_1 & b_1 & 0 & \ldots & 0 \\ 0 & \frac{\delta_2}{\delta_1} & b_2 & \ldots & 0 \\ & & \ddots & \ddots & \\ & & & \ddots & b_{n-1} \\ 0 & 0 & 0 & \ldots & \frac{\delta_n}{\delta_{n-1}} \end{pmatrix}.$$

3. Déterminer les déterminants δ_k et la décomposition LU quand elle existe, dans le cas où

$$\forall i \in [\![1, n]\!], \ a_{ii} = 2b; \quad \forall i \in [\![1, n-1]\!], \ b_i = c_i = -1.$$

(École polytechnique)

Solution.

1. Pour $k \in [\![1, n]\!]$, on note $A_k \in \mathcal{M}_k(\mathbb{R})$ la matrice constituée des k premières lignes et colonnes de A et $\delta_k = \det A_k$. On a

$$A_k = \begin{pmatrix} a_1 & b_1 & 0 & \ldots & 0 \\ c_1 & a_2 & \ddots & & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & a_{k-1} & b_{k-1} \\ 0 & \ldots & 0 & c_{k-1} & a_k \end{pmatrix}.$$

On obtient $\delta_1 = a_1$, $\delta_2 = a_1 a_2 - b_1 c_1$ et pour $k \geqslant 3$, on obtient en développant selon la dernière ligne

$$\delta_k = a_k\delta_{k-1} - c_{k-1}\det\begin{pmatrix} a_1 & b_1 & 0 & \dots & 0 \\ c_1 & a_2 & \ddots & & \vdots \\ 0 & \ddots & \ddots & b_{k-2} & \vdots \\ \vdots & & \ddots & a_{k-2} & 0 \\ 0 & \dots & 0 & c_{k-2} & b_{k-1} \end{pmatrix},$$

et en développant par rapport à la dernière colonne

$$\boxed{\delta_k = a_k\delta_{k-1} - b_{k-1}c_{k-1}\delta_{k-2}}.$$

2. Soit $(l_1, \dots, l_{n-1}) \in \mathbb{R}^{n-1}$. On pose $L = (l_{ij})$ et $U = (u_{ij})$, avec

$$\begin{array}{llll} l_{ii} = 1 & \text{si } i \in [\![1,n]\!], & l_{i,i-1} = l_{i-1} & \text{si } i \in [\![2,n]\!], \\ u_{i,i} = \dfrac{\delta_i}{\delta_{i-1}} & \text{si } i \in [\![1,n]\!], & u_{i,i+1} = b_i & \text{si } i \in [\![1,n-1]\!], \end{array}$$

et $l_{ij} = u_{ij} = 0$ sinon, en posant $\delta_0 = 1$. On cherche à montrer l'existence de $(l_1, \dots, l_{n-1})$ tel que $A = LU$. On note que la matrice U est bien définie car les δ_k ne sont pas nuls. Il faut, pour tout $(i,j) \in [\![1,n]\!]$,

$$a_{ij} = \sum_{k=1}^{n} l_{ik}u_{kj} = u_{ij} + l_{i-1}u_{i-1,j},$$

car $l_{ik} = 0$ si $k \neq i$ et $k \neq i-1$. Cela équivaut à $a_1 = a_{11} = \dfrac{\delta_1}{\delta_0}$ et

$$\begin{array}{ll} \forall i \in [\![2,n]\!], & a_i = a_{ii} = \dfrac{\delta_i}{\delta_{i-1}} + l_{i-1}b_{i-1} \\ \forall i \in [\![1,n-1]\!], & b_i = a_{i,i+1} = u_{i,i+1} = b_i \\ \forall i \in [\![2,n]\!], & c_{i-1} = a_{i,i-1} = l_{i-1}u_{i-1,i-1} = l_{i-1}\dfrac{\delta_{i-1}}{\delta_{i-2}}, \end{array}$$

car pour $j \notin \{i-1, i, i+1\}$, les deux membres de l'égalité sont nuls. L'égalité $a_1 = \dfrac{\delta_1}{\delta_0}$ est vérifiée. On doit avoir, pour $2 \leqslant i \leqslant n$,

$$\boxed{l_{i-1} = \frac{c_{i-1}\delta_{i-2}}{\delta_{i-1}}},$$

ce qui définit $(l_1, \dots, l_{n-1})$. Il ne reste plus qu'à vérifier que, pour tout $i \in [\![2,n]\!]$,

$$a_i = \frac{\delta_i}{\delta_{i-1}} + l_{i-1}b_{i-1} = \frac{\delta_i}{\delta_{i-1}} + \frac{c_{i-1}b_{i-1}\delta_{i-2}}{\delta_{i-1}},$$

c'est-à-dire

$$\delta_i = a_i\delta_{i-1} - b_{i-1}c_{i-1}\delta_{i-2},$$

ce qui résulte de la première question. On remarque que cette décomposition est unique.

Il résulte de l'exercice 1.38 que la non-nullité de tous les δ_k est aussi une condition nécessaire pour l'existence d'une telle décomposition.

3. On a, toujours avec $\delta_0 = 1$, pour $2 \leqslant k \leqslant n$,

$$\delta_k = 2b\delta_{k-1} - \delta_{k-2}.$$

La suite $(\delta_k)_{k\geqslant 0}$ est récurrente linéaire d'ordre 2. Intéressons-nous aux racines de $r^2 - 2br + 1 = 0$ dont le discriminant est $\Delta = 4b^2 - 4$.

• Premier cas : $|b| > 1$.

Dans ces conditions, $\Delta > 0$. Les racines de $r^2 - 2br + 1 = 0$ sont réelles et distinctes $b + \sqrt{b^2-1}$ et $b - \sqrt{b^2-1}$. Il existe donc A et B réels tels que pour tout $k \in \mathbb{N}$,

$$\delta_k = \mathrm{A}(b + \sqrt{b^2-1})^k + \mathrm{B}(b - \sqrt{b^2-1})^k$$

En écrivant cette égalité pour $k = 0$ ou $k = 1$, on obtient le système

$$\begin{cases} \mathrm{A} + \mathrm{B} = 1 \\ \mathrm{A}(b + \sqrt{b^2-1}) + \mathrm{B}(b - \sqrt{b^2-1}) = 2b \end{cases}$$

qui donne $\mathrm{A} = \dfrac{b+\sqrt{b^2-1}}{2\sqrt{b^2-1}}$ et $\mathrm{B} = -\dfrac{b-\sqrt{b^2-1}}{2\sqrt{b^2-1}}$, ce qui permet d'écrire pour $k \geqslant 0$,

$$\boxed{\delta_k = \frac{1}{2\sqrt{b^2-1}}\left[(b + \sqrt{b^2-1})^{k+1} - (b - \sqrt{b^2-1})^{k+1}\right]}.$$

La condition $\delta_k = 0$ équivaut à $(b + \sqrt{b^2-1})^{k+1} = (b - \sqrt{b^2-1})^{k+1}$. Dans ces conditions,

$$b + \sqrt{b^2-1} = \pm(b - \sqrt{b^2-1}) \text{ et en élevant au carré } b\sqrt{b^2-1} = 0,$$

ce qui est impossible puisque $|b| > 1$.

Par conséquent, si $|b| > 1$, on a, pour tout $k \in [\![1, n]\!]$, $\delta_k \neq 0$. La matrice A peut s'écrire sous la forme LU.

Si $b > 1$, il existe $t > 0$ tel que $b = \operatorname{ch} t$. Dans ces conditions, on obtient $b + \sqrt{b^2-1} = e^t$, $b - \sqrt{b^2-1} = e^{-t}$ et pour tout $k \geqslant 0$,

$$\boxed{\delta_k = \frac{\operatorname{sh}(k+1)t}{\operatorname{sh} t}}.$$

Si $b < -1$, il existe $t > 0$ tel que $b = -\operatorname{ch} t$. Dans ces conditions, on obtient, pour tout $k \geqslant 0$,

$$\boxed{\delta_k = (-1)^k \frac{\operatorname{sh}(k+1)t}{\operatorname{sh} t}}.$$

• Deuxième cas : $b = \pm 1$.

Dans ces conditions, $\Delta = 0$ et b est la seule racine de $r^2 - 2br + 1 = 0$. Il existe alors A et B réels tels que pour tout $n \geqslant 0$, $\delta_k = b^k(Ak + B)$. On a $D_0 = 1 = B$ et $D_1 = 2b = b(A + 1)$, ce qui donne $A = B = 1$ et

$$\boxed{\delta_k = (k+1) \text{ si } b = 1 \quad \text{et} \quad \delta_k = (-1)^k(k+1) \text{ si } b = -1}.$$

Là encore, aucun δ_k n'est nul et A possède une décomposition LU.

• Troisième cas : $|b| < 1$.

Dans ces conditions, $\Delta < 0$. Les racines de $r^2 - 2br + 1 = 0$ sont complexes conjugués :

$$b + i\sqrt{1 - b^2} \quad \text{et} \quad b - i\sqrt{1 - b^2}.$$

Posons $\theta = \arccos b \in]0, \pi[$. On a alors $\sqrt{1 - b^2} = \sin\theta$ et les racines sont $e^{i\theta}$ et $e^{-i\theta}$. Il existe A et B dans $\mathbb{C}$ tels que pour tout $k \geqslant 0$ on ait

$$\delta_k = Ae^{ik\theta} + Be^{-ik\theta}.$$

On a $A + B = 1$ et $Ae^{i\theta} + Be^{-i\theta} = 2\cos\theta$. Il vient $A = \frac{e^{i\theta}}{2i\sin\theta}$ et $B = -\frac{e^{i\theta}}{2i\sin\theta}$ et $\delta_k = \frac{1}{2i\sin\theta}\left[e^{i(k+1)\theta} - e^{-i(k+1)\theta}\right]$ ou plus simplement

$$\boxed{\delta_k = \frac{\sin(k+1)\theta}{\sin\theta}}.$$

$\delta_k = 0$ équivaut à $(k+1)\theta \in \pi\mathbb{Z}$. Comme $\theta \in]0, \pi[$, $\delta_k = 0$ si, et seulement si, $\theta = \frac{l\pi}{k+1}$ avec $1 \leqslant l \leqslant k$. On conclut que la décomposition LU de la matrice A existe si et seulement si b n'est pas de la forme $b = \cos\frac{l\pi}{k+1}$ avec $1 \leqslant l \leqslant k \leqslant n$. ◁

Chapitre 2
Réduction

On doit à Camille Jordan (1838-1922) de nombreux résultats sur la réduction des endomorphismes, qu'il découvre notamment à travers l'étude des groupes. Dépassant la notion des groupes de permutations pour en atteindre une plus abstraite, il s'intéresse à la classification des groupes finis à travers leurs représentations linéaires (autrement dit les morphismes d'un groupe fini G *dans le groupe linéaire* $\mathrm{GL}(\mathrm{E})$ *d'un espace vectoriel* E*). Il va même jusqu'à donner la description des classes de similitude à l'aide des formes dites de Jordan.*

Le problème fondamental de la réduction est bien celui de caractériser les classes de similitude de l'algèbre $\mathcal{L}(\mathrm{E})$ *où* E *est un* K*-espace vectoriel de dimension finie ou, ce qui revient au même, les classes de similitude de l'algèbre* $\mathcal{M}_n(\mathrm{K})$*. La recherche d'une matrice la plus simple possible pour représenter un endomorphisme donné vise de multiples buts : calculer les puissances successives de cet endomorphisme, son commutant, résoudre des systèmes différentiels linéaires... Une idée naturelle pour essayer de « réduire » l'étude d'un endomorphisme* u *donné à des choses plus simples consiste à essayer de décomposer l'espace vectoriel* E *en une somme directe de sous-espaces non triviaux stables par* u*. Cela n'est évidemment pas toujours possible (le lecteur se reportera aux exercices 2.37 et 2.39 concernant la simplicité et la semi-simplicité). Les sous-espaces stables les plus simples sont ceux sur lesquels* u *coïncide avec une homothétie. On est ainsi naturellement amené à la notion de valeur propre. Si* λ *est un scalaire, on s'intéresse donc au sous-espace* $\mathrm{E}_\lambda = \mathrm{Ker}(u - \lambda\,\mathrm{Id}_\mathrm{E})$ *appelé sous-espace propre pour la valeur propre* λ *lorsqu'il n'est pas nul. Le théorème de décomposition des noyaux nous assure que les différents sous-espaces propres d'un endomorphisme sont en somme directe. Le cas où la somme remplit tout l'espace* E *mène à la notion d'endomorphisme diagonalisable : un tel endomorphisme peut être représenté par une matrice diagonale (il suffit de prendre une base formée de vecteurs propres). Pour les endomorphismes diagonalisables il est alors très facile de répondre à la question initiale de savoir quand ils sont semblables : il faut et suffit qu'ils aient les mêmes valeurs propres et que les espaces propres associés aient la même dimension. Il est aussi facile, en se rame-*

nant à une matrice diagonale, de calculer les puissances d'un tel endomorphisme, son exponentielle (si on travaille sur un sous-corps de $\mathbb{C}$), son commutant... Cependant il ne s'agit là que d'une réponse partielle à notre problème. Un pas supplémentaire vers sa résolution, lorsque le corps de base est algébriquement clos (par exemple sur $\mathbb{C}$), consiste à établir le théorème de décomposition de Dunford : tout endomorphisme u d'un $\mathbb{C}$-espace vectoriel de dimension finie s'écrit de manière unique sous la forme $u = d + n$ où d est diagonalisable, n est nilpotent et commute avec d. Ce résultat, qui n'est pas dans les programmes actuels des classes préparatoires, fait l'objet de l'exercice 2.30. Cette décomposition amène l'attention sur les classes de similitude des endomorphismes nilpotents. Avec un peu de travail il est alors possible d'obtenir le théorème de Jordan qui règle complètement la question de la détermination des classes de similitude sur un corps algébriquement clos. Dans le cadre général, l'outil fondamental pour résoudre cette question est la notion d'endomorphisme cyclique : voir les exercices 2.38, 2.42 ainsi que le commentaire qui suit le second exercice. Après l'étude de ces deux exercices, le lecteur pourra sans mal se lancer dans la lecture d'un bon ouvrage d'algèbre linéaire, pour avoir enfin la solution complète au problème posé en introduction !

Voici maintenant la liste des thèmes successifs selon lesquels nous avons regroupé les exercices. Le plan suit la progression logique que nous venons d'évoquer plus haut.

- *Nos premiers exercices concernent les valeurs propres. Un scalaire λ est valeur propre d'un endomorphisme u si et seulement si $u - \lambda \mathrm{Id}_E$ n'est pas injectif donc, comme on est en dimension finie, non bijectif, soit encore de déterminant nul. Autrement dit, les valeurs propres de u sont exactement les racines du polynôme caractéristique χ_u de u. Cela permet de définir la multiplicité d'une valeur propre comme son ordre de multiplicité en tant que racine de χ_u. Il est facile de montrer que cette multiplicité majore la dimension de l'espace propre associé à λ. Les énoncés des exercices concernent des déterminations de spectre, des questions de multiplicité, ou de localisation des valeurs propres. Il y a notamment plusieurs exercices sur les matrices à coefficients réels positifs.*

- *Le thème suivant sera consacré aux polynômes d'endomorphismes (ou de matrices). On sait que le critère de diagonalisation le plus puissant est l'existence d'un polynôme annulateur qui soit scindé à racines simples. Les exercices porteront sur la notion de polynôme minimal, le théorème de décomposition des noyaux...*

- *Viendront ensuite les exercices concernant la diagonalisation, la trigonalisation, puis la réduction de Dunford avec quelques applications notamment à l'exponentielle des matrices. Comme nous l'avons déjà dit,*

la décomposition $d+n$ ramène le problème de la réduction pour un corps algébriquement clos à la réduction des endomorphismes nilpotents. Plusieurs exercices sont alors consacrés à ces derniers.

• *Nous regrouperons ensuite les exercices concernant les sous-espaces stables, la notion d'endomorphismes cycliques.*

• *Viendront ensuite diverses applications : recherche de commutant, équations à base de crochet de Lie, équation de Sylvester* $AX - XB = C$.

• *Le dernier thème sera consacré à des exercices de nature topologique.*

Par convention, sauf mention contraire, K *désigne un corps commutatif quelconque. Selon l'usage, on identifie souvent les éléments de* K^n *avec les matrices unicolonnes et les matrices avec l'application linéaire canoniquement associée.*

Dans le premier exercice ci-après on connaît la dimension de l'espace propre et on cherche l'ordre de multiplicité de la valeur propre.

2.1. Valeur propre simple

Soit K un corps commutatif, $\lambda \in K$, $A \in \mathcal{M}_n(K)$, $X, Y \in \mathcal{M}_{n,1}(K)$ tels que

$$AX = \lambda X, \quad {}^tYA = \lambda {}^tY, \quad {}^tYX \neq 0, \quad \operatorname{rg}(A - \lambda I_n) = n - 1.$$

Montrer que λ est valeur propre simple de A.

(**École polytechnique**)

▷ **Solution.**

Les vecteurs X et Y sont différents de 0 puisque ${}^tYX \neq 0$. Le scalaire λ est valeur propre de A (et aussi de ${}^tA$) et par hypothèse l'espace propre associé à λ pour A est une droite (car $\operatorname{rg}(A - \lambda I_n) = n - 1$). Cette droite est donc dirigée par X. Comme on souhaite calculer l'ordre de multiplicité de λ en tant que racine du polynôme caractéristique de A, on va essayer de calculer ce polynôme mais en choisissant une matrice semblable à A qui soit plus simple. Il faut évidemment utiliser les informations que l'on a sur la transposée de A. Le vecteur Y est un vecteur propre de ${}^tA$ pour la valeur propre λ. Il définit une forme linéaire non nulle sur K^n (par $Z \mapsto {}^tYZ$) dont le noyau est un hyperplan H. L'hypothèse ${}^tYX \neq 0$ signifie que X n'est pas dans cet hyperplan H.

On remarque alors que, pour tout $Z \in H$, on a ${}^tYAZ = \lambda {}^tYZ = 0$ et $AZ \in H$. Ainsi l'hyperplan H est stable par A. Considérons une base de K^n constituée de X et d'une base de H. Dans cette base, l'endomorphisme canoniquement associé à A a une matrice de la forme

$$\mathrm{M} = \begin{pmatrix} \lambda & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & \mathrm{N} & \\ 0 & & & \end{pmatrix},$$

où $\mathrm{N} \in \mathcal{M}_{n-1}(\mathrm{K})$. On a donc

$$\chi_A = \chi_M = \det \begin{pmatrix} X-\lambda & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & XI_{n-1} - N & \\ 0 & & & \end{pmatrix} = (X-\lambda)\chi_N.$$

La matrice $\lambda I_n - M$ est semblable à $\lambda I_n - A$. Elle est donc de rang $n-1$. Comme $\lambda I_n - M = \begin{pmatrix} 0 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & \lambda I_{n-1} - N & \\ 0 & & & \end{pmatrix}$, $N - \lambda I_{n-1}$ est de rang $n-1$. Ainsi, λ n'est pas valeur propre de N et λ est bien une valeur propre simple de A. ◁

Il est important de retenir que si u est un endomorphisme d'un espace E *et f une forme linéaire non nulle sur* E*, l'hyperplan* $\operatorname{Ker} f$ *est stable par u si et seulement si f est un vecteur propre de ${}^t u$.*

L'énoncé suivant concerne encore des questions de multiplicité de valeur propre.

2.2. Existence d'une valeur propre double

Soient A, B et C trois matrices de $\mathcal{M}_2(\mathrm{K})$. Montrer qu'il existe un triplet $(\alpha, \beta, \gamma) \in \mathrm{K}^3 \backslash \{0\}$ tel que $\alpha A + \beta B + \gamma C$ ait une valeur propre double.

(ENS Lyon)

▷ **Solution.**

Bien entendu si la famille (A, B, C) est liée le résultat est clair, car on peut trouver une combinaison linéaire nulle des trois matrices avec des coefficients (α, β, γ) non tous nuls.

Supposons donc que la famille (A, B, C) est libre. On connaît des matrices simples qui ont une valeur propre double : toutes les matrices

triangulaires supérieures de la forme $\begin{pmatrix} \lambda & \mu \\ 0 & \lambda \end{pmatrix}$. Or, l'ensemble de ces matrices forme clairement un sous-espace vectoriel de $\mathcal{M}_2(\mathrm{K})$ de dimension 2. Ce sous-espace ne peut donc pas être en somme directe avec l'espace $\mathrm{Vect}(\mathrm{A}, \mathrm{B}, \mathrm{C})$ qui est de dimension 3 (car $\dim \mathcal{M}_2(\mathrm{K}) = 4$). Il y a donc une matrice non nulle de la forme $\alpha\mathrm{A} + \beta\mathrm{B} + \gamma\mathrm{C}$ qui admet une valeur propre double. ◁

2.3. Détermination de spectre

On définit une suite de matrices par $\mathrm{A}_1 = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$ et la relation de récurrence $\mathrm{A}_{n+1} = \begin{pmatrix} \mathrm{A}_n & \mathrm{A}_n \\ \mathrm{A}_n & -\mathrm{A}_n \end{pmatrix}$. Déterminer les valeurs propres de A_n.

(École polytechnique)

▷ **Solution.**

On commence naturellement par regarder le cas $n = 1$. Le polynôme caractéristique de A_1 est

$$\chi_1(\mathrm{X}) = \begin{vmatrix} \mathrm{X}-1 & -1 \\ -1 & \mathrm{X}+1 \end{vmatrix} = \mathrm{X}^2 - 2.$$

On a donc $\mathrm{Sp}\,\mathrm{A}_1 = \{\pm\sqrt{2}\}$.

Essayons maintenant de trouver une relation entre le polynôme caractéristique de A_n et celui de A_{n+1}. Notons que la matrice A_n est de taille 2^n. On a

$$\chi_{n+1}(\mathrm{X}) = \begin{vmatrix} \mathrm{X}\mathrm{I}_{2^n} - \mathrm{A}_n & -\mathrm{A}_n \\ -\mathrm{A}_n & \mathrm{X}\mathrm{I}_{2^n} + \mathrm{A}_n \end{vmatrix}.$$

Les quatre matrices qui interviennent commutent deux à deux. On a donc (cf. exercice 1.27)

$$\chi_{n+1}(\mathrm{X}) = \det(\mathrm{X}^2\mathrm{I}_{2^n} - 2\mathrm{A}_n^2) = \det(\mathrm{X}\mathrm{I}_{2^n} - \sqrt{2}\mathrm{A}_n)\det(\mathrm{X}\mathrm{I}_{2^n} + \sqrt{2}\mathrm{A}_n).$$

On obtient donc en factorisant $\sqrt{2}$

$$\chi_{n+1}(\mathrm{X}) = 2^{2^n}\chi_n\left(\frac{\mathrm{X}}{\sqrt{2}}\right)\chi_n\left(\frac{-\mathrm{X}}{\sqrt{2}}\right).$$

Les racines de χ_{n+1} sont donc obtenues en multipliant celles de χ_n par $\pm\sqrt{2}$.

Conclusion. On a pour tout n, $\mathrm{Sp}\,\mathrm{A}_n = \{\pm\sqrt{2}^n\}$. $\lhd$

Pour une matrice donnée (disons réelle ou complexe) de taille importante le calcul des valeurs propres n'est pas chose facile. D'une part le calcul du polynôme caractéristique n'est pas simple (rappelons que c'est un déterminant) et d'autre part on sait bien que la recherche des racines d'un polynôme est un problème difficile. Ces questions ayant une grande importance pratique on a été amené à chercher des algorithmes efficaces. On se préoccupe ici du calcul de χ_A où A est une matrice de $\mathcal{M}_n(\mathbb{C})$. Les coefficients de χ_A sont donnés par les fonctions symétriques élémentaires $\sigma_1, \ldots, \sigma_n$ des valeurs propres $\lambda_1, \ldots, \lambda_n$ (prises avec multiplicité). Une première idée est de calculer ces fonctions symétriques élémentaires à l'aide des formules de Newton qui relient les σ_k aux sommes de Newton $\mathrm{S}_k = \sum_{i=1}^{n} \lambda_i^k$. En effet, ces sommes de Newton sont plus faciles à trouver puisque $\mathrm{S}_k = \mathrm{Tr}\,\mathrm{A}^k$ pour tout k (trigonaliser A pour le voir). Cela constitue la méthode de Leverrier. L'exercice suivant présente l'algorithme de Faddeev qui consiste à remplacer la suite A^k par une autre suite de matrices avec le même objectif : obtenir les coefficients successifs du polynôme caractéristique. On y utilisera les formules de Newton que le lecteur trouvera dans l'exercice 5.26 du premier tome d'algèbre.

2.4. Algorithme de Faddeev

Soit $\mathrm{A} \in \mathcal{M}_n(\mathbb{C})$. On définit une suite $(\mathrm{A}_k)_{k\geqslant 0}$ de matrices en posant $\mathrm{A}_0 = \mathrm{A}$ et pour $k \geqslant 1$,

$$\mathrm{A}_k = \mathrm{A}\left(\mathrm{A}_{k-1} - \frac{1}{k}\,\mathrm{Tr}(\mathrm{A}_{k-1})\mathrm{I}_n\right).$$

Montrer que A_n est nulle.

(École polytechnique)

$\rhd$ **Solution.**

On commence par calculer les premiers termes de la suite $(\mathrm{A}_k)_{k\geqslant 0}$:

$$\mathrm{A}_0 = \mathrm{A}\ ; \quad \mathrm{A}_1 = \mathrm{A}^2 - \mathrm{Tr}(\mathrm{A}_0)\mathrm{A}\ ; \quad \mathrm{A}_2 = \mathrm{A}^3 - \mathrm{Tr}(\mathrm{A}_0)\mathrm{A}^2 - \frac{1}{2}\,\mathrm{Tr}(\mathrm{A}_1)\mathrm{A}.$$

Par une récurrence immédiate, on montre alors que pour tout $k \in [\![0, n]\!]$,

$$\mathrm{A}_k = \mathrm{A}^{k+1} - \mathrm{Tr}(\mathrm{A}_0)\mathrm{A}^k - \frac{1}{2}\,\mathrm{Tr}(\mathrm{A}_1)\mathrm{A}^{k-1} - \cdots - \frac{1}{k}\,\mathrm{Tr}(\mathrm{A}_{k-1})\mathrm{A}.$$

Pour calculer A_n, il faut calculer $\mathrm{Tr}(A_k)$ pour tout $k \in [\![0, n-1]\!]$. On note $\lambda_1, \ldots, \lambda_n$ les valeurs propres de A comptées avec multiplicité. Les valeurs propres de A^k sont $\lambda_1^k, \ldots, \lambda_n^k$ et $\mathrm{Tr}(A^k) = \sum_{i=1}^{n} \lambda_i^k$. On obtient $\mathrm{Tr}(A_0) = \mathrm{Tr}(A)$ puis

$$\mathrm{Tr}(A_1) = \mathrm{Tr}(A^2) - (\mathrm{Tr}(A))^2 = \sum_{i=1}^{n} \lambda_i^2 - \left(\sum_{i=1}^{n} \lambda_i\right)^2 = -2 \sum_{1 \leqslant i < j \leqslant n} \lambda_i \lambda_j.$$

En notant $\sigma_1, \ldots, \sigma_n$ les fonctions symétriques élémentaires des valeurs propres, on a donc $\mathrm{Tr}(A_0) = \sigma_1$ et $\mathrm{Tr}(A_1) = -2\sigma_2$.

Montrons par récurrence sur k que $\mathrm{Tr}(A_k) = (-1)^k(k+1)\sigma_{k+1}$. C'est vrai pour $k = 1$ et $k = 2$ comme on vient de le voir. Si la proposition est vérifiée jusqu'au rang $k-1$, on obtient, en posant $S_p = \sum_{i=1}^{n} \lambda_i^p$,

$$\begin{aligned}\mathrm{Tr}(A_k) &= \mathrm{Tr}(A^{k+1}) - \mathrm{Tr}(A_0)\,\mathrm{Tr}(A^k) - \frac{1}{2}\,\mathrm{Tr}(A_1)\,\mathrm{Tr}(A^{k-1}) - \\ &\qquad\qquad \cdots - \frac{1}{k}\,\mathrm{Tr}(A_{k-1})\,\mathrm{Tr}(A) \\ &= S_{k+1} - \sigma_1 S_k + \sigma_2 S_{k-1} - \sigma_3 S_{k-2} + \cdots + (-1)^k \sigma_k S_1.\end{aligned}$$

Les formules de Newton (cf exercice 5.26 du tome 1 d'algèbre) donnent

$$S_{k+1} - \sigma_1 S_k + \cdots + (-1)^k \sigma_k S_1 + (-1)^{k+1}(k+1)\sigma_{k+1} = 0$$

et on a donc $\mathrm{Tr}(A_k) = -(-1)^{k+1}(k+1)\sigma_{k+1} = (-1)^k(k+1)\sigma_{k+1}$. La proposition est vérifiée au rang $k+1$ et donc pour tout k. On obtient en particulier

$$A_n = A^{n+1} - \sigma_1 A^n + \sigma_2 A^{n-1} + \cdots + \sigma_n A.$$

Comme $\lambda_1, \ldots, \lambda_n$ sont les racines du polynôme caractéristique, on a

$$\chi_A = X^n + \sum_{k=0}^{n} (-1)^k \sigma_k X^{n-k}$$

et donc $A_n = A\chi_A(A)$. D'après le théorème de Cayley-Hamilton, $\chi_A(A) = 0$ donc $A_n = 0$. ◁

Les coefficients du polynôme caractéristique sont donc, au signe près, les $\frac{1}{k+1}\,\mathrm{Tr}(A_k)$ *(pour* $0 \leqslant k \leqslant n-1$*). La suite* (A_k) *définie dans l'énoncé fournit ainsi un algorithme de calcul du polynôme caractéristique appelé*

algorithme de Fadeev, bien plus rapide que celui qui consiste à calculer $\det(\mathrm{X}\mathrm{I}_n - \mathrm{A})$.

Notons à cette occasion que c'est le mathématicien allemand Frobenius (1849-1917) qui donna en 1878 la première démonstration du théorème de Cayley-Hamilton.

Le calcul des valeurs propres étant difficile, on peut être amené à simplement essayer de les localiser. Voici un résultat assez classique de ce type, basé sur le lemme d'Hadamard.

2.5. Lemme d'Hadamard, disques de Gershgörin

Soit $\mathrm{A} = (a_{ij})_{1 \leqslant i,j \leqslant n} \in \mathcal{M}_n(\mathbb{C})$. On pose $\mathrm{R}_i = \sum_{\substack{1 \leqslant j \leqslant n \\ j \neq i}} |a_{ij}|$ pour tout $i \in [\![1, n]\!]$.

1. On suppose que pour tout $i \in [\![1, n]\!]$, $|a_{ii}| > \mathrm{R}_i$. Montrer que A est inversible.

2. Montrer que $\mathrm{Sp}\,\mathrm{A} \subset \bigcup_{i=1}^{n} \{z \in \mathbb{C}, |z - a_{ii}| \leqslant \mathrm{R}_i\}$.

3. On suppose à nouveau que pour tout $i \in [\![1, n]\!]$. $|a_{ii}| > \mathrm{R}_i$. Montrer que

$$|\det \mathrm{A}| \geqslant \prod_{i=1}^{n} (|a_{ii}| - \mathrm{R}_i).$$

4. On suppose que $\mathrm{A} \in \mathcal{M}_n(\mathbb{R})$ et que pour tout $i \in [\![1, n]\!]$, $a_{ii} > \mathrm{R}_i$. Montrer que

$$\det \mathrm{A} \geqslant \prod_{i=1}^{n} (|a_{ii}| - \mathrm{R}_i).$$

(École polytechnique)

▷ **Solution.**

1. Raisonnons par l'absurde et supposons A non inversible. Il existe alors $\mathrm{X} = (x_1, \ldots, x_n)$ un vecteur colonne non nul tel que $\mathrm{AX} = 0$. Pour tout $1 \leqslant i \leqslant n$, on a $\sum_{j=1}^{n} a_{ij} x_j = 0$ puis, en isolant le terme d'indice i,

$$|a_{ii} x_i| = \left| \sum_{j \neq i} a_{ij} x_j \right| \leqslant \sum_{j \neq i} |a_{ij}||x_j| \leqslant \mathrm{R}_i \max_{1 \leqslant k \leqslant n} |x_k|.$$

Prenons $i_0 \in [\![1,n]\!]$ tel que $|x_{i_0}| = \max_{1\leqslant k\leqslant n} |x_k|$. On a $|x_{i_0}| > 0$ car $X \neq 0$. On obtient $|a_{i_0 i_0}||x_{i_0}| \leqslant \mathrm{R}_{i_0}|x_{i_0}|$ et donc $|a_{i_0 i_0}| \leqslant \mathrm{R}_{i_0}$. Cela contredit l'hypothèse. On conclut que $A \in \mathrm{GL}_n(\mathbb{C})$.

Ce résultat constitue le lemme d'Hadamard.

2. Soit $\lambda \in \mathrm{Sp}(A)$. Alors $A - \lambda I_n$ n'est pas inversible. Donc, il existe $i_0 \in [\![1,n]\!]$ tel que $|a_{i_0 i_0} - \lambda| \leqslant \mathrm{R}_{i_0}$ *i.e.* $\lambda \in \{z \in \mathbb{C}, |z - a_{i_0 i_0}| \leqslant \mathrm{R}_{i_0}\}$. Finalement,

$$\lambda \in \bigcup_{i=1}^{n} \{z \in \mathbb{C}, |z - a_{ii}| \leqslant \mathrm{R}_i\}.$$

Les disques de centre a_{ii} et de rayon R_i dont la réunion contient le spectre de A sont appelés disques de Gershgörin.

3. Notons A' la matrice obtenue à partir de A en multipliant, pour tout $1 \leqslant i \leqslant n$ la i-ième ligne par $\frac{1}{|a_{ii}| - \mathrm{R}_i}$ (le dénominateur est bien non nul). On obtient

$$\det A = \det A' \prod_{i=1}^{n} (|a_{ii}| - \mathrm{R}_i).$$

Il n'y a plus qu'à prouver que $\det A' \geqslant 1$. Posons $A' = (a'_{ij})_{1\leqslant i,j\leqslant n}$. On a pour tout $1 \leqslant i \leqslant n$,

$$|a'_{ii}| - \sum_{j\neq i} |a'_{ij}| = \frac{|a_{ii}| - \mathrm{R}_i}{\big||a_{ii}| - \mathrm{R}_i\big|} = \frac{|a_{ii}| - \mathrm{R}_i}{|a_{ii}| - \mathrm{R}_i} = 1.$$

Pour toute valeur propre λ de A', il existe $1 \leqslant i \leqslant n$ tel que

$$|\lambda - a'_{ii}| \leqslant \sum_{j\neq i} |a'_{ij}|$$

et donc, par l'inégalité triangulaire,

$$|\lambda| - |a'_{ii}| \geqslant -\sum_{j\neq i} |a'_{ij}| \quad \text{et} \quad |\lambda| \geqslant |a'_{ii}| - \sum_{j\neq i} |a'_{ij}| = 1.$$

Toute valeur propre de A' a un module supérieur ou égal à 1. Or le déterminant de A' est le produit de ses valeurs propres (distinctes ou confondues). Donc $|\det A'| \geqslant 1$. On en déduit que

$$|\det A| = |\det A'| \prod_{i=1}^{n} (|a_{ii}| - \mathrm{R}_i) \geqslant \prod_{i=1}^{n} (|a_{ii}| - \mathrm{R}_i).$$

4. D'après la question 3. nous avons seulement à établir que $\det A > 0$. Pour cela nous allons utiliser un argument de connexité. Notons $\mathcal{C}$ l'ensemble des matrices $(a_{ij})_{1\leqslant i,j\leqslant n} \in \mathcal{M}_n(\mathbb{R})$ vérifiant

$$\forall i \in [\![1,n]\!],\ \ a_{ii} > \sum_{\substack{1\leqslant j\leqslant n\\ j\neq i}} |a_{ij}|$$

et montrons que $\mathcal{C}$ est convexe. Soient $A = (a_{ij})$ et $B = (b_{ij})$ dans $\mathcal{C}$ et $t \in]0,1[$. Montrons que $(1-t)A + tB = ((1-t)a_{ij} + tb_{ij}) \in \mathcal{C}$.

Pour tout $i \in [\![1,n]\!]$, on a

$$\begin{aligned}\sum_{j\neq i} |(1-t)a_{ij} + tb_{ij}| &\leqslant \sum_{j\neq i}((1-t)|a_{ij}| + t|b_{ij}|)\\ &\leqslant (1-t)\sum_{j\neq i}|a_{ij}| + t\sum_{j\neq i}|b_{ij}|\\ &< (1-t)a_{ii} + tb_{ii}.\end{aligned}$$

Donc $(1-t)A + tB \in \mathcal{C}$ et $\mathcal{C}$ est convexe et en particulier connexe. Or la fonction $M \in \mathcal{C} \longmapsto \det M$ est continue et, d'après la question 1. ne s'annule pas sur $\mathcal{C}$. Son image est donc un intervalle de $\mathbb{R}$ ne contenant pas 0 : cette image est contenue dans $\mathbb{R}_+^*$ ou dans $\mathbb{R}_-^*$. Comme $I_n \in \mathcal{C}$ et $\det I_n = 1 > 0$, pour tout $M \in \mathcal{C}$, $\det M > 0$. La question 3 donne le résultat voulu. ◁

On peut appliquer le résultat obtenu dans la question 2 à une matrice compagnon pour en déduire un résultat de localisation des racines d'un polynôme quelconque.

Nous commençons maintenant une série d'exercices sur les matrices à coefficients réels positifs et tout d'abord sur les matrices stochastiques. Celles-ci interviennent naturellement en Probabilités dans l'étude des chaînes de Markov (le lecteur pourra se reporter au commentaire précédent l'exercice 7.14 dans notre premier tome d'exercices d'algèbre pour un peu plus de détails). Le premier énoncé montre que le spectre d'une matrice stochastique est entièrement contenu dans le disque unité du plan complexe, résultat que précisera l'énoncé d'après.

2.6. Matrices stochastiques (1)

Soit S l'ensemble des matrices réelles stochastiques, c'est-à-dire l'ensemble des matrices $P = (p_{ij}) \in \mathcal{M}_n(\mathbb{R})$ telles que

$$\forall (i,j) \in [\![1,n]\!]^2,\ p_{ij} \geqslant 0 \quad \text{et} \quad \forall i \in [\![1,n]\!],\ \sum_{j=1}^{n} p_{ij} = 1.$$

1. Montrer que tous les éléments de S ont une valeur propre commune.

2. Si P, Q sont dans S, en est-il de même de PQ ?

3. Soit $P \in S$ et λ une valeur propre complexe de P. Montrer que $|\lambda| \leqslant 1$.

(École polytechnique)

▷ **Solution.**

1. Si $P \in S$, et si U est le vecteur de $\mathbb{R}^n$ dont toutes les coordonnées valent 1, la seconde condition qui définit les éléments de S équivaut à $PU = U$. Donc tous les matrices stochastiques admettent 1 comme valeur propre et U comme vecteur propre associé.

2. Si P, Q sont dans S, les coefficients de PQ sont tous positifs. De plus $(PQ)U = PU = U$ de sorte que la seconde condition est aussi remplie. Donc S est stable pour le produit.

3. Soit $\lambda \in \operatorname{Sp} P$ et $X = {}^t(x_1, \ldots, x_n)$ un vecteur propre associé. Choisissons un indice i tel que $|x_i| = \max_{1 \leqslant k \leqslant n} |x_k|$. Comme $PX = \lambda X$ on a, en regardant la i-ième coordonnée,

$$p_{i1}x_1 + \cdots + p_{in}x_n = \lambda x_i$$

Et en passant au module, il vient

$$|\lambda x_i| = |\lambda||x_i| = |p_{i1}x_1 + \cdots + p_{in}x_n| \leqslant (p_{i1} + \cdots + p_{in})|x_i| = |x_i|.$$

On conclut que $|\lambda| \leqslant 1$. ◁

Dans l'énoncé suivant on prouve que les éventuelles valeurs propres de module 1 sont des racines de l'unité.

2.7. Matrices stochastiques (2)

Soit $A = (a_{ij}) \in \mathcal{M}_n(\mathbb{R})$ une matrice stochastique. Soit λ une valeur propre de A de module 1 et $X = (x_1, \ldots, x_n) \in \mathbb{C}^n$ un vecteur propre associé.

1. Si x_i est une composante de X de module maximal, montrer que λx_i est encore une composante de X de module maximal.

2. En déduire que λ est une racine m-ième de l'unité avec $m \leqslant n$.

3. On suppose que pour tout $i \in [\![1, n]\!]$, $a_{ii} \neq 0$. Montrer que la seule valeur propre de A de module 1 est 1.

(ENS Ulm)

▷ **Solution.**

1. On suppose $\lambda \neq 1$ sinon la question est triviale. La i-ième ligne de l'égalité $AX = \lambda X$ conduit à

$$\lambda - a_{ii} = \sum_{j \neq i} a_{ij} \frac{x_j}{x_i}. \tag{1}$$

On en déduit que

$$1 - a_{ii} = |\lambda| - a_{ii} \leqslant |\lambda - a_{ii}| \leqslant \sum_{j \neq i} a_{ij} \left|\frac{x_j}{x_i}\right| \leqslant \sum_{j \neq i} a_{ij} = 1 - a_{ii}.$$

Les inégalités sont donc toutes des égalités. On déduit de cela plusieurs choses.

• L'égalité $1 - a_{ii} = |\lambda - a_{ii}|$ montre que λ est sur le cercle de centre a_{ii} et de rayon $1 - a_{ii}$. Si $a_{ii} > 0$ ce cercle est tangent intérieurement au cercle unité de $\mathbb{C}$ en 1. On aurait alors $\lambda = 1$ et cela est exclu. On a donc nécessairement $a_{ii} = 0$. Il en résulte notamment que l'ensemble $I = \{j \neq i,\ a_{ij} \neq 0\}$ n'est pas vide.

• D'après le cas d'égalité dans l'inégalité triangulaire, les complexes $a_{ij}\frac{x_j}{x_i}$ pour $j \neq i$ sont tous sur une même demi-droite d'origine 0. Cela est intéressant uniquement si $a_{ij} \neq 0$ c'est-à-dire si $j \in I$.

• Pour $j \in I$ on a $a_{ij}\left|\frac{x_j}{x_i}\right| = a_{ij}$, c'est-à-dire $|x_j| = |x_i|$.

En combinant ces deux derniers points, on obtient l'existence d'un réel θ tel que, pour tout $j \in I$, $x_j = e^{i\theta} x_i$. En remplaçant dans la relation (1), on obtient

$$\lambda = e^{i\theta} \sum_{j \in I} a_{ij} = e^{i\theta} \sum_{j=1}^{n} a_{ij} = e^{i\theta}$$

Ainsi, $e^{i\theta} = \lambda$ et $x_j = \lambda x_i$ pour tout indice $j \in \mathrm{I}$. D'où le résultat demandé.

2. On suppose encore $\lambda \neq 1$. Soit $k \geqslant 1$ le nombre de coordonnées de X de module maximal. Parmi les $k+1$ complexes $x_i, \lambda x_i, \lambda^2 x_i, \ldots, \lambda^k x_i$, qui sont tous des composantes de X de module maximal, il y en a nécessairement deux qui sont égaux (principe des tiroirs). Il existe donc $r < s$ tels que $\lambda^r x_i = \lambda^s x_i$ et comme $x_i \neq 0$, on a $\lambda^m = 1$ avec $m \in [\![1, n]\!]$ en posant $m = s - r$.

3. La contraposée a été vue dans la question 1. ◁

L'énoncé suivant n'a rien à voir avec la réduction, mais comme il porte aussi sur le thème des matrices stochastiques nous avons choisi de le présenter ici.

2.8. Matrices stochastiques (3)

Soit S l'ensemble des matrices carrées d'ordre n réelles stochastiques.

1. Montrer que S est un convexe compact.

2. Déterminer le plus petit entier p tel qu'il existe un sous-espace affine de dimension p de $\mathcal{M}_n(\mathbb{R})$ contenant S.

(École polytechnique)

▷ **Solution.**

1. Rappelons qu'une matrice A est stochastique si ses coefficients sont tous positifs et si $\mathrm{AU} = \mathrm{U}$ où U désigne le vecteur dont toutes les composantes valent 1 (voir exercice 2.6). Si A, B sont dans S il est aisé de voir que

$$\forall t \in [0, 1], \quad (1-t)\mathrm{A} + t\mathrm{B} \in \mathrm{S}.$$

En effet, les coefficients sont positifs par convexité de l'intervalle $[0, 1]$ et on a bien $((1-t)\mathrm{A} + t\mathrm{B})\mathrm{U} = (1-t)\mathrm{U} + t\mathrm{U} = \mathrm{U}$. Donc S est convexe.

On choisit de munir $\mathcal{M}_n(\mathbb{R})$ de la norme définie par $\|\mathrm{A}\|_1 = \max_{i,j} |a_{ij}|$ (toutes les normes sont équivalentes). Alors S est borné par 1 et fermé comme intersection d'hyperplans affines et de demi-espaces fermés. Donc S est compact.

2. Cette question demande de trouver la dimension du sous-espace affine engendré par S. Notons f_i la forme linéaire sur $\mathcal{M}_n(\mathbb{R})$ qui à une matrice associe la somme des coefficients de sa i-ième ligne. Par définition S est inclus dans le sous-espace affine $\mathrm{V} = \bigcap_{i=1}^{n} f_i^{-1}(1)$. De manière plus

précise S est l'intersection de V avec la partie de $\mathcal{M}_n(\mathbb{R})$ formée des matrices à coefficients positifs. Comme les formes linéaires $(f_1, \ldots, f_n)$ sont clairement libres, le sous-espace affine V est de dimension $n^2 - n$. On va prouver que S engendre affinement V.

Soit $A = (a_{ij})$ une matrice de V. Alors, A peut s'écrire comme barycentre de n matrices de V dont la dernière ligne est positive :

$$A = \sum_{j=1}^{n} a_{nj} B_j$$

où B_j est obtenue en gardant les lignes $1, \ldots n-1$ de A et en remplaçant la dernière ligne par $(0, 0, \ldots, 0, 1, 0, \ldots, 0)$ avec le 1 en j-ième position (la somme des poids a_{nj} est égale à 1 puisque A est dans V).

De même, chacune des matrices B_j peut s'écrire comme barycentre de matrices de V dont les deux dernières lignes sont positives. Par associativité du barycentre, A est donc barycentre de matrices de V dont les deux dernières lignes sont positives. On itère cette idée pour obtenir finalement A comme barycentre de matrices stochastiques. ◁

En reprenant les calculs déjà menés dans l'exercice 2.7 on montre maintenant que pour une matrice strictement stochastique l'espace propre associé à la valeur propre 1 *est une droite.*

2.9. Matrices strictement stochastiques

Soit $A = (a_{ij})_{1 \leqslant i,j \leqslant n} \in \mathcal{M}_n(\mathbb{R})$ une matrice strictement stochastique, c'est-à-dire stochastique avec des coefficients strictement positifs.

1. Montrer que $1 \in \operatorname{Sp} A$ et que $\dim \operatorname{Ker}(A - I) = 1$.

2. Montrer que toute valeur propre complexe de A est de module inférieur à 1 et que 1 est la seule valeur propre de module 1.

(ENS Ulm)

▷ **Solution.**

1. Le vecteur U dont toutes les coordonnées valent 1 est un vecteur propre pour la valeur propre 1. Soit $X = (x_1, \ldots, x_n) \in \mathbb{C}^n$ un vecteur non nul tel que $AX = X$. On reprend la démonstration de la première question de l'exercice 2.7. On choisit un indice i tel que $|x_i|$ soit maximal. En regardant la i-ième coordonnée de l'égalité $AX = X$ il vient

$$(1-a_{ii})x_i = \left|\sum_{j\neq i} a_{ij}\frac{x_j}{x_i}\right| \leqslant \sum_{j\neq i} a_{ij}\left|\frac{x_j}{x_i}\right| \leqslant \sum_{j\neq i} a_{ij} = 1-a_{ii}.$$

Toutes les inégalités sont donc des égalités. Il en résulte d'une part que pour tout $j \in [\![1,n]\!]$, $|x_j| = |x_i|$ (car $a_{ij} \neq 0$) et d'autre part que tous les $a_{ij}\frac{x_j}{x_i}$ ont le même argument (cas d'égalité dans l'inégalité triangulaire). Comme a_{ij} est un réel strictement positif, les x_j sont donc tous égaux et X est colinéaire à U.

Le lecteur pourra montrer qu'en fait 1 *est une racine simple du polynôme caractéristique de* A *(ce qui est plus fort que le résultat prouvé ici).*

2. La première partie de la question a été résolue dans l'exercice 2.6 et la seconde est un cas particulier de la question 3 de l'exercice 2.7 puisque les coefficients diagonaux de A sont non nuls. ◁

Il y a en fait beaucoup de résultats concernant le spectre des matrices réelles dont les coefficients sont positifs. Le théorème de Perron Frobenius, qui est prouvé dans l'exercice suivant, généralise les résultats vus en 2.6 et 2.7 pour les matrices stochastiques. Celui-ci affirme qu'une matrice réelle A *à coefficients positifs qui est irréductible (notion définie dans l'énoncé) admet une valeur propre réelle positive* ρ *telle que* $|\lambda| \leqslant \rho$ *pour toute valeur propre (complexe) de* A*. De plus, on peut trouver pour* ρ *un vecteur propre à coordonnées positives (vecteur de Perron). Ce résultat a été obtenu par Perron en 1907 pour les matrices à coefficients strictement positifs, puis généralisé par Frobenius par la suite. L'approche utilisée dans l'exercice est due à Wielandt.*

2.10. Théorème de Perron-Frobenius (1907)

Soit $A \in \mathcal{M}_n(\mathbb{R})$ une matrice à coefficients positifs. On note u l'endomorphisme de $\mathbb{R}^n$ canoniquement associé à A et $(e_1, \ldots, e_n)$ la base canonique de $\mathbb{R}^n$. On suppose que A est irréductible : cela signifie qu'il n'existe aucune permutation σ de $[\![1,n]\!]$ telle que la matrice de u dans la base $(e_{\sigma(1)}, \ldots, e_{\sigma(n)})$ soit triangulaire par blocs.

1. Montrer que $(I_n + A)^{n-1}$ est à coefficients strictement positifs. Si $X \in \mathbb{R}^n$ est un vecteur non nul à coordonnées positives, on pourra comparer le nombre de coordonnées strictement positives de X et de $Y = (I_n + A)X$.

2. Si $X = (X_1, \ldots, X_n) \in \mathbb{R}_+^n$ est non nul, on pose $r(X) = \min_{X_i \neq 0} \frac{(AX)_i}{X_i}$ où $(AX)_i$ désigne la i-ième coordonnée du vecteur AX. Montrer que r admet un maximum ρ sur $\mathbb{R}_+^n \setminus \{0\}$.

3. Montrer que $\rho \in \mathrm{Sp}\, A$ et que $|\lambda| \leqslant \rho$ pour tout $\lambda \in \mathrm{Sp}\, A$.

(**ENS Cachan**)

▷ **Solution.**

1. Soit $X = (X_1, \ldots, X_n) \in \mathbb{R}_+^n$ et $Y = (I_n + A)X$. Posons $Y = (Y_1, \ldots, Y_n)$. On a pour tout i, $Y_i = X_i + \sum_{j=1}^{n} a_{ij} X_j$ et il est donc clair que Y est aussi à coordonnées positives. Comme nous y invite l'énoncé, on va comparer le nombre de coordonnées strictement positives de X et de Y. Si $X = 0$ on a évidemment $Y = 0$. Si $X > 0$ (on note ainsi le fait que $X_i > 0$ pour tout i), alors $Y > 0$. Supposons donc que X ait exactement p coordonnées strictement positives avec $1 \leqslant p < n$. Il est clair que si $X_i > 0$, alors $Y_i > 0$. Donc Y a au moins p coordonnées strictement positives. On va prouver qu'il en a au moins une de plus. Notons $i_1, ..., i_p$ les indices des p coordonnées strictement positives de X (les autres étant nulles). Supposons par l'absurde que $Y_i = 0$ pour $i \notin \{i_1, ..., i_p\}$. On a donc pour un tel indice i,

$$\sum_{j=1}^{n} a_{ij} X_j = \sum_{k=1}^{p} a_{ii_k} X_{i_k} = 0.$$

Comme les X_{i_k} sont non nuls, cela impose que $a_{ii_k} = 0$ pour tout $i \notin \{i_1, ..., i_p\}$ et tout $k \in [\![1, p]\!]$. Cela contredit l'irréductibilité de la matrice A. En effet, si on effectue sur A une permutation consistant à mettre les colonnes et les lignes d'indice $i_1, ..., i_p$ en premier, la matrice obtenue est triangulaire par blocs.

Il découle directement de ce résultat que si X est un vecteur non nul à coordonnées positives, alors $(A + I_n)^{n-1} X > 0$. C'est en particulier le cas en prenant pour X les vecteurs de la base canonique. Ainsi, toutes les colonnes de $(I_n + A)^{n-1}$ sont à coefficients strictement positifs.

2. Il est clair que pour tout $X = (X_1, \ldots, X_n) \in \mathbb{R}_+^n \setminus \{0\}$ et tout $t > 0$ on a $r(tX) = r(X)$. Cette homogénéité permet de ramener la recherche du maximum de r sur l'ensemble $K = \{X = (X_1, \ldots, X_n) \in \mathbb{R}_+^n, \sum_{i=1}^{n} X_i = 1\}$ qui a le mérite d'être compact (il est clairement fermé et borné dans $\mathbb{R}^n$). Il reste à étudier la continuité de r. Les fonctions $X \mapsto \frac{(AX)_i}{X_i}$ sont toutes continues sur l'ensemble $P = (\mathbb{R}_+^*)^n$. Par ailleurs

le minimum de deux fonctions continues (et donc, par récurrence, de n) est encore une fonction continue (rappelons que pour deux réels a et b on a $\min(a,b) = \frac{a+b}{2} - \frac{|b-a|}{2}$). Il en découle que r est continue sur P. Mais rien ne garantit que r reste continue en un point de la frontière de P, c'est-à-dire en un vecteur positif X qui a au moins une coordonnée nulle. Et d'ailleurs c'est faux en général ! Prenons par exemple la matrice $A = \begin{pmatrix} 2 & 1 \\ 0 & 1 \end{pmatrix}$. Prenons $X_\varepsilon = (1, \varepsilon)$ pour $\varepsilon > 0$. On a

$$r(X_\varepsilon) = \min\left(\frac{2+\varepsilon}{1}, \frac{\varepsilon}{\varepsilon}\right) = 1$$

pour tout $\varepsilon > 0$. Et $r(1,0) = 2$ n'est donc pas égal à la limite de $r(X_\varepsilon)$ lorsque $\varepsilon \to 0^+$. Certes A n'est pas irréductible mais ce n'est pas la cause du problème : le lecteur fabriquera sans mal un exemple de matrice irréductible pour laquelle r n'est pas continue en tout point de K.

On va utiliser la première question pour contourner ce problème. Posons $K' = (I_n + A)^{n-1}(K)$. Comme image continue d'un compact, K' est encore compact, et il est inclus dans P d'après la première question. Donc r est continue sur K' et par suite atteint un maximum ρ. Il ne reste plus qu'à prouver que ρ est le maximum de r sur K. Soit $X \in K$ et $Y = (I_n + A)^{n-1}X \in K'$. On va montrer que $r(X) \leqslant r(Y)$. Par définition de $r(X)$ on a $(AX)_i \geqslant r(X)X_i$ pour tout $i \in [\![1, n]\!]$ (même si $X_i = 0$). Autrement dit, le vecteur $AX - r(X)X$ est positif. Son image par $(I_n + A)^{n-1}$ est donc à coordonnées positives. Or, cette image vaut $AY - r(X)Y$ car A et $(I_n + A)^{n-1}$ commutent. On a donc $(AY)_i \geqslant r(X)Y_i$ pour tout i et par suite $r(Y) = \min\limits_{1 \leqslant i \leqslant n} \dfrac{(AY)_i}{Y_i} \geqslant r(X)$. Il en découle que r est majorée par ρ sur K. Mais cette valeur est atteinte dans K : en effet, il existe $Y \in K'$ tel que $r(Y) = \rho$ et il suffit de prendre l'unique vecteur de K colinéaire à Y.

Conclusion. La fonction r admet ρ pour maximum sur K et donc sur $\mathbb{R}_+^n \setminus \{0\}$.

3. Par définition on a $\rho \geqslant 0$. En fait ρ n'est pas nul, car il est clair que $r(1, 1, \ldots, 1) > 0$. Reprenons $Y \in K'$ tel que $r(Y) = \rho$. On va prouver qu'en fait $AY = \rho Y$. On a $AY - \rho Y \geqslant 0$. Supposons que ce vecteur ne soit pas nul. D'après la question 1 son image par $(I_n + A)^{n-1}$ est alors un vecteur strictement positif. Posons $Z = (I_n + A)^{n-1}Y$. On a alors $AZ - \rho Z > 0$. Pour $\varepsilon > 0$ assez petit le vecteur $AZ - \rho Z - \varepsilon Z$ reste strictement positif. On a alors $r(Z) \geqslant \rho + \varepsilon$, ce qui contredit la définition de ρ.

Montrons enfin que ρ majore le module de toute valeur propre complexe de A. Soit $\lambda \in \operatorname{Sp} A$ et $X \in \mathbb{C}^n$ un vecteur propre associé. On a

pour tout i, $\lambda X_i = \sum_{j=1}^{n} a_{ij} X_j$. En prenant le module il vient, par inégalité triangulaire,

$$|\lambda||X_i| \leqslant \sum_{j=1}^{n} a_{ij}|X_j|$$

pour tout i. Autrement dit, si on note $\tilde{X}$ le vecteur $(|X_1|, \ldots, |X_n|)$, on a $(A\tilde{X})_i \geqslant |\lambda|\tilde{X}_i$ pour tout i et donc $|\lambda| \leqslant r(\tilde{X})$. Or, $r(\tilde{X}) \leqslant \rho$. D'où le résultat.

Conclusion. ρ est une valeur propre de A et $|\lambda| \leqslant \rho$ pour tout $\lambda \in \mathrm{Sp}\, A$. On notera par ailleurs que A admet un vecteur propre à coordonnées strictement positives pour ρ. $\lhd$

Le lecteur pourra en fait prouver que l'espace propre associé à ρ est de dimension 1 *et même, ce qui est un peu plus difficile, que ρ est une racine simple du polynôme caractéristique de* A. *Indiquons que la seconde épreuve du concours de l'ENSAE en 1994 a porté sur ce sujet ainsi que le début de l'épreuve d'Agrégation Externe de 1995. Une très bonne référence pour toutes les questions relatives aux matrices positives est le livre de Henry Minc,* Nonnegative Matrices, *paru aux éditions Wiley.*

Avant d'aborder les exercices concernant la diagonalisation, nous regroupons ci-après des énoncés portant sur les polynômes d'endomorphismes. Rappelons que si u est un endomorphisme d'un K*-espace vectoriel* E *l'application ψ qui à* $P \in K[X]$ *associe* $P(u)$ *définit un morphisme d'algèbre de* $\mathbb{K}[X]$ *dans* $\mathcal{L}(E)$. *Son image* $K[u]$ *est donc une sous-algèbre commutative de* $\mathcal{L}(E)$. *Lorsque* E *est de dimension finie le morphisme ψ ne peut pas être injectif. Son noyau est un idéal non nul de* K[X] *dont le générateur unitaire est appelé polynôme minimal de u. En revanche, lorsque* E *est de dimension infinie,* $\mathrm{Ker}\, \psi$ *peut très bien être réduit à* $\{0\}$ *et l'existence d'un polynôme annulateur non nul n'est plus assurée.*

2.11. Polynôme annulateur

Soit E un espace vectoriel de dimension infinie et u, v deux endomorphismes de E qui commutent. On suppose que u (resp. v) admet un polynôme annulateur non nul P (resp. Q). Montrer qu'il en est de même de $u + v$.

(École polytechnique)

▷ **Solution.**

Notons que l'existence d'un polynôme annulateur non nul de u équivaut au caractère lié de la famille $(u^k)_{k\in\mathbb{N}}$. Notons n le plus grand entier tel que $(\mathrm{Id}, u, \ldots, u^{n-1})$ soit libre. La famille $(\mathrm{Id}, u, \ldots, u^{n-1}, u^n)$ est alors liée et on a donc $u^n \in \mathrm{Vect}(\mathrm{Id}, u, \ldots, u^{n-1})$. Il est alors facile de prouver par récurrence sur k que $u^k \in \mathrm{Vect}(\mathrm{Id}, u, \ldots, u^{n-1})$ pour tout $k \geqslant n$. Autrement dit, on a

$$\mathrm{Vect}(\mathrm{Id}, u, \ldots, u^{n-1}) = \mathrm{Vect}(u^k)_{k\geqslant 0}.$$

En fait, n n'est rien d'autre que le degré du polynôme minimal de u. Notons de même m le degré du polynôme minimal de v. Il est alors clair que

$$\mathrm{F} = \mathrm{Vect}(u^k v^l)_{(k,l)\in\mathbb{N}^2} = \mathrm{Vect}(u^i v^j)_{0\leqslant i\leqslant n-1,\, 0\leqslant j\leqslant m-1}.$$

En particulier, il s'agit d'un espace de dimension finie. Comme u et v commutent, on a $(u+v)^k = \sum_{i=0}^{k} \mathrm{C}_k^i u^i v^{k-i}$ pour tout k. Le sous-espace $\mathrm{Vect}((u+v)^k)_{k\geqslant 0}$ est donc contenu dans F et par suite il est de dimension finie. D'après la remarque initiale, cela prouve que $u+v$ admet un polynôme annulateur non nul. ◁

2.12. Noyaux et images de polynômes d'endomorphismes

Soient A et B dans K[X], f un endomorphisme d'un K-espace vectoriel. On pose $\mathrm{D} = \mathrm{pgcd}(\mathrm{A}, \mathrm{B})$ et $\mathrm{M} = \mathrm{ppcm}(\mathrm{A}, \mathrm{B})$. Exprimer $\mathrm{Ker\,D}(f)$, $\mathrm{Im\,D}(f)$, $\mathrm{Ker\,M}(f)$, $\mathrm{Im\,M}(f)$ à l'aide de $\mathrm{Ker\,A}(f)$, $\mathrm{Im\,A}(f)$, $\mathrm{Ker\,B}(f)$ et $\mathrm{Im\,B}(f)$.

(**École polytechnique**)

▷ **Solution.**

Nous noterons A′ et B′ les quotients de A et B par D (sans risque de confusion avec la dérivée). On a donc $\mathrm{A} = \mathrm{DA}'$, $\mathrm{B} = \mathrm{DB}'$ et $\mathrm{M} = \mathrm{DA'B'}$. Les polynômes A′ et B′ sont premiers entre eux et nous allons utiliser une relation de Bezout $\mathrm{A'P} + \mathrm{B'Q} = 1$ (et donc $\mathrm{D} = \mathrm{AP} + \mathrm{BQ}$). Nous utiliserons enfin le résultat évident suivant : si u, v sont deux endomorphismes, $\mathrm{Ker}\, u \subset \mathrm{Ker}(v \circ u)$ et $\mathrm{Im}(v \circ u) \subset \mathrm{Im}\, v$. Commençons par chercher le noyau et l'image de $\mathrm{D}(f)$.

• Comme $\mathrm{A}(f) = \mathrm{A}'(f) \circ \mathrm{D}(f)$, on a $\mathrm{Ker\,D}(f) \subset \mathrm{Ker\,A}(f)$. Il en est de même avec B et $\mathrm{Ker\,D}(f)$ est donc inclus dans $\mathrm{Ker\,A}(f) \cap \mathrm{Ker\,B}(f)$. On va montrer qu'il y a égalité. Pour cela, on utilise la relation de Bezout : $\mathrm{D}(f) = \mathrm{P}(f) \circ \mathrm{A}(f) + \mathrm{Q}(f) \circ \mathrm{B}(f)$. Il en découle directement que si

$\mathrm{A}(f)(x) = \mathrm{B}(f)(x) = 0$, alors $\mathrm{D}(f)(x) = 0$. On conclut

$$\boxed{\operatorname{Ker} \mathrm{D}(f) = \operatorname{Ker} \mathrm{A}(f) \cap \operatorname{Ker} \mathrm{B}(f)}.$$

• Considérons $\operatorname{Im} \mathrm{D}(f)$. Comme $\mathrm{A}(f) = \mathrm{D}(f) \circ \mathrm{A}'(f)$, on en déduit que $\operatorname{Im} \mathrm{A}(f) \subset \operatorname{Im} \mathrm{D}(f)$. De même, $\operatorname{Im} \mathrm{B}(f) \subset \operatorname{Im} \mathrm{D}(f)$ et on a donc l'inclusion $\operatorname{Im} \mathrm{A}(f) + \operatorname{Im} \mathrm{B}(f) \subset \operatorname{Im} \mathrm{D}(f)$. Inversement, comme $\mathrm{D}(f) = \mathrm{A}(f) \circ \mathrm{P}(f) + \mathrm{B}(f) \circ \mathrm{Q}(f)$, on a $\operatorname{Im} \mathrm{D}(f) \subset \operatorname{Im} \mathrm{A}(f) + \operatorname{Im} \mathrm{B}(f)$. On conclut que

$$\boxed{\operatorname{Im} \mathrm{D}(f) = \operatorname{Im} \mathrm{A}(f) + \operatorname{Im} \mathrm{B}(f)}.$$

• Regardons maintenant $\operatorname{Ker} \mathrm{M}(f)$. On a $\mathrm{M} = \mathrm{AB}'$ et donc $\mathrm{M}(f) = \mathrm{B}'(f) \circ \mathrm{A}(f)$. Ainsi $\operatorname{Ker} \mathrm{A}(f) \subset \operatorname{Ker} \mathrm{M}(f)$. De même, $\operatorname{Ker} \mathrm{B}(f) \subset \operatorname{Ker} \mathrm{M}(f)$ et par conséquent la somme $\operatorname{Ker} \mathrm{A}(f) + \operatorname{Ker} \mathrm{B}(f)$ est contenue dans $\operatorname{Ker} \mathrm{M}(f)$.

Inversement, soit $x \in \operatorname{Ker} \mathrm{M}(f)$. On a $\mathrm{Id} = \mathrm{A}'(f) \circ \mathrm{P}(f) + \mathrm{B}'(f) \circ \mathrm{Q}(f)$ et en appliquant en x,

$$x = (\mathrm{A}'(f) \circ \mathrm{P}(f))(x) + (\mathrm{B}'(f) \circ \mathrm{Q}(f))(x)$$

Or le premier vecteur est dans $\operatorname{Ker} \mathrm{B}(f)$, car $\mathrm{BA}'\mathrm{P} = \mathrm{MP}$ est un multiple de M, et pour des raisons analogues le second est dans $\operatorname{Ker} \mathrm{A}(f)$. Il s'ensuit que x est dans $\operatorname{Ker} \mathrm{A}(f) + \operatorname{Ker} \mathrm{B}(f)$. On conclut que

$$\boxed{\operatorname{Ker} \mathrm{M}(f) = \operatorname{Ker} \mathrm{A}(f) + \operatorname{Ker} \mathrm{B}(f)}.$$

• Considérons enfin $\operatorname{Im} \mathrm{M}(f)$. Comme $\mathrm{M} = \mathrm{AB}'$, on en déduit que $\operatorname{Im} \mathrm{M}(f) \subset \operatorname{Im} \mathrm{A}(f)$. Par symétrie, $\operatorname{Im} \mathrm{M}(f) \subset \operatorname{Im} \mathrm{B}(f)$ et finalement $\operatorname{Im} \mathrm{M}(f) \subset \operatorname{Im} \mathrm{A}(f) \cap \operatorname{Im} \mathrm{B}(f)$. Ici encore il y a égalité. En effet, on écrit toujours

$$x = (\mathrm{A}'(f) \circ \mathrm{P}(f))(x) + (\mathrm{B}'(f) \circ \mathrm{Q}(f))(x).$$

Si $x \in \operatorname{Im} \mathrm{B}(f)$ on peut l'écrire $x = \mathrm{B}(f)(y)$ et dans ce cas

$$(\mathrm{A}'(f) \circ \mathrm{P}(f))(x) = (\mathrm{A}'(f) \circ \mathrm{P}(f) \circ \mathrm{B}(f))(y) = (\mathrm{M}(f) \circ \mathrm{P}(f))(y) \in \operatorname{Im} \mathrm{M}(f).$$

De même, si $x \in \operatorname{Im} \mathrm{A}(f)$, le vecteur $(\mathrm{B}'(f) \circ \mathrm{Q}(f))(x)$ est dans $\operatorname{Im} \mathrm{M}(f)$. Donc si $x \in \operatorname{Im} \mathrm{A}(f) \cap \operatorname{Im} \mathrm{B}(f)$, il est dans $\operatorname{Im} \mathrm{M}(f)$. On conclut donc que

$$\boxed{\operatorname{Im} \mathrm{M}(f) = \operatorname{Im} \mathrm{A}(f) \cap \operatorname{Im} \mathrm{B}(f)}. \quad \lhd$$

Pour un endomorphisme u d'un espace de dimension finie, il est classique de montrer que la suite de sous-espaces $(\operatorname{Ker} u^k)_{k \geqslant 0}$ est stricte-

croissante jusqu'à un certain rang p puis stationnaire. Le lecteur trouvera cela dans l'exercice 6.14 du premier tome d'algèbre. Dans l'énoncé suivant on constate que cet indice p n'est autre que la valuation du polynôme minimal de u.

2.13. Valuation du polynôme minimal

Soit E un K-espace vectoriel de dimension n, u un endomorphisme de E. On note P le polynôme minimal de u et p la valuation de P.

1. On suppose $p = 0$. Que peut-on dire de u ?

2. On suppose $p = 1$. Montrer que $\mathrm{E} = \mathrm{Ker}\, u \oplus \mathrm{Im}\, u$.

3. Dans le cas général, montrer que $\mathrm{E} = \mathrm{Ker}\, u^p \oplus \mathrm{Im}\, u^p$ et que p est le plus petit entier pour lequel on a cette propriété.

(**École polytechnique**)

▷ **Solution.**

1. Dire que 0 n'est pas racine de P revient à dire que 0 n'est pas valeur propre de u et donc que u est inversible.

2. Si $p = 1$ on peut écrire $\mathrm{P} = \mathrm{XQ(X)}$ avec $\mathrm{Q}(0) \neq 0$. D'après le théorème de décomposition des noyaux, on a $\mathrm{Ker}\, u \oplus \mathrm{Ker}\, \mathrm{Q}(u) = \mathrm{E}$. D'autre part, $\mathrm{P}(u) = \mathrm{Q}(u) \circ u = 0$. Il s'ensuit que $\mathrm{Im}\, u \subset \mathrm{Ker}\, \mathrm{Q}(u)$. En vertu du théorème du rang, $\dim \mathrm{Im}\, u = \dim \mathrm{E} - \dim \mathrm{Ker}\, u$, ce qui vaut $\dim \mathrm{Ker}\, \mathrm{Q}(u)$ d'après ce qui précède. On a donc l'égalité $\mathrm{Im}\, u = \mathrm{Ker}\, \mathrm{Q}(u)$ et finalement :

$$\boxed{\mathrm{E} = \mathrm{Ker}\, u \oplus \mathrm{Im}\, u}.$$

3. On supposera dans toute cette question $p \geqslant 1$, *i.e.* u non inversible. On procède comme avant. Écrivons $\mathrm{P} = \mathrm{X}^p \mathrm{Q(X)}$ avec $\mathrm{Q}(0) \neq 0$. Par le théorème de décomposition des noyaux, on a $\mathrm{E} = \mathrm{Ker}\, u^p \oplus \mathrm{Ker}\, \mathrm{Q}(u)$. Comme $\mathrm{Q}(u) \circ u^p = 0$, on a $\mathrm{Im}\, u^p \subset \mathrm{Ker}\, \mathrm{Q}(u)$ et les deux espaces ayant la même dimension, ils sont égaux. On a donc

$$\boxed{\mathrm{E} = \mathrm{Ker}\, u^p \oplus \mathrm{Im}\, u^p}.$$

Notons que si $q \geqslant p$, le polynôme $\mathrm{X}^q \mathrm{Q(X)}$ annule aussi u et on a $\mathrm{E} = \mathrm{Ker}\, u^q \oplus \mathrm{Ker}\, \mathrm{Q}(u)$. Comme $\mathrm{Ker}\, u^p \subset \mathrm{Ker}\, u^q$, on a forcément égalité (les deux espaces ont la même dimension). Ainsi, la suite $(\mathrm{Ker}\, u^k)_{k \geqslant 0}$ est constante à partir du rang p. Comme $\mathrm{Im}\, u^q \subset \mathrm{Im}\, u^p$, le théorème du rang montre que la suite $(\mathrm{Im}\, u^k)_{k \geqslant 0}$ est aussi constante à partir du rang p. En particulier, on a $\mathrm{E} = \mathrm{Ker}\, u^q \oplus \mathrm{Im}\, u^q$ pour tout $q \geqslant p$. On va maintenant montrer que ce n'est plus le cas pour $q < p$.

Si $1 \leqslant q < p$, le polynôme $X^qQ(X)$ n'annule pas u (par définition du polynôme minimal) et on a donc $\operatorname{Ker} u^q \subsetneq \operatorname{Ker} Q(u) \neq E$. Ainsi, $\operatorname{Ker} u^q$ est strictement inclus dans $\operatorname{Ker} u^p$. Supposons par l'absurde qu'on ait tout de même $E = \operatorname{Ker} u^q \oplus \operatorname{Im} u^q$. On va voir qu'alors on a forcément $\operatorname{Ker} u^q = \operatorname{Ker} u^{q+1}$ et $\operatorname{Im} u^q = \operatorname{Im} u^{q+1}$. En itérant cela, on aura alors $\operatorname{Ker} u^q = \operatorname{Ker} u^p$, et on vient de voir à l'instant que ce n'est pas le cas.

L'inclusion $\operatorname{Ker} u^q \subset \operatorname{Ker} u^{q+1}$ étant triviale, prenons x dans $\operatorname{Ker} u^{q+1}$. On a alors $u^q(x) \in \operatorname{Im} u^q$ et $u(u^q(x)) = 0$. Donc $u^q(x)$ est dans $\operatorname{Im} u^q$ et dans $\operatorname{Ker} u$, donc aussi dans $\operatorname{Ker} u^q$. Ainsi $u^q(x) = 0$ et $x \in \operatorname{Ker} u^q$. On a donc $\operatorname{Ker} u^q = \operatorname{Ker} u^{q+1}$. D'autre part, $\operatorname{Im} u^{q+1} \subset \operatorname{Im} u^q$ et $\dim \operatorname{Im} u^{q+1} = \dim E - \dim \operatorname{Ker} u^{q+1} = \dim E - \operatorname{Ker} u^q = \dim \operatorname{Im} u^q$. Par conséquent, $\operatorname{Im} u^q = \operatorname{Im} u^{q+1}$. Comme dans ces conditions $E = \operatorname{Ker} u^{q+1} \oplus \operatorname{Im} u^{q+1}$, on a de même $\operatorname{Ker} u^{q+2} = \operatorname{Ker} u^{q+1} = \operatorname{Ker} u^q$ et $\operatorname{Im} u^{q+2} = \operatorname{Im} u^{q+1} = \operatorname{Im} u^q$. Et ainsi de proche en proche. Les suites $(\operatorname{Ker} u^k)_{k\geqslant 1}$ et $(\operatorname{Im} u^k)_{k\geqslant 1}$ sont donc stationnaires à partir de q. Or on a vu que $\operatorname{Ker} u^q \subsetneq \operatorname{Ker} u^p$. Contradiction.

Conclusion. L'entier p est le plus petit entier naturel tel que $E = \operatorname{Ker} u^p \oplus \operatorname{Im} u^p$; c'est aussi l'entier à partir duquel les suites $(\operatorname{Ker} u^k)_{k\geqslant 1}$ et $(\operatorname{Im} u^k)_{k\geqslant 1}$ sont stationnaires. ◁

Dans l'exercice suivant on montre que le polynôme minimal d'une matrice $A \in \mathcal{M}_n(K)$ a en fait ses coefficients dans le plus petits sous-corps de K qui contient tous les coefficients de la matrice A.

2.14. Invariance du polynôme minimal par extension de corps

Soit $M \in \mathcal{M}_n(\mathbb{Q})$. On appelle $\pi_M \in \mathbb{Q}[X]$ (resp. $\mu_M \in \mathbb{R}[X]$) le polynôme minimal de M en tant que matrice de $\mathcal{M}_n(\mathbb{Q})$ (resp. $\mathcal{M}_n(\mathbb{R})$). Montrer que $\pi_M = \mu_M$.

(ENS Lyon)

▷ **Solution.**

Si $P \in \mathbb{R}[X]$ est tel que $P(M) = 0$, alors μ_M divise P (l'ensemble des polynômes annulateurs dans $\mathbb{R}[X]$ de M est un idéal engendré par μ_M). Or $\pi_M \in \mathbb{R}[X]$ et $\pi_M(M) = 0$. Ainsi μ_M divise π_M. Pour prouver que ces deux polynômes sont égaux, il suffit de montrer qu'ils ont le même degré (ils sont tous deux unitaires). Cela découle de l'invariance du rang par extension de corps :

Lemme. *Soit $D \in \mathcal{M}_{n,p}(K)$, L un sur-corps commutatif de K. On note r_K (resp. r_L) le rang de A comme matrice à coefficients dans K (resp. dans L). On a alors $r_K = r_L$.*

Démonstration. Tout d'abord, on a l'équivalence suivante pour $B \in \mathcal{M}_m(K)$:

$$B \in \mathrm{GL}_m(K) \iff \det B \neq 0 \iff B \in \mathrm{GL}_m(L)$$

le déterminant étant le même, que le calcul se fasse dans K ou L. Par conséquent, toute sous-matrice B de D est inversible sur K si, et seulement si, elle est inversible sur L. Or, r_K est la taille maximale des sous-matrices de D inversibles sur K donc, d'après l'équivalence, la taille maximale des sous-matrices de D inversibles sur L. Ainsi $r_K = r_L$. ♢

Posons alors $d = \deg \pi_M$. Par minimalité de d, la famille $(I_n, M, M^2, \ldots, M^{d-1})$ est libre dans $\mathcal{M}_n(\mathbb{Q})$. Écrivons la matrice D de cette famille de vecteurs dans la base canonique $(E_{ij})_{1 \leqslant i,j \leqslant n}$ de $\mathcal{M}_n(\mathbb{Q})$. Elle possède d colonnes et n^2 lignes. Les coefficients de D sont tous dans $\mathbb{Q}$ puisque ce sont des coefficients des matrices M^k. Le rang de D sur $\mathbb{Q}$ est égal à d par hypothèse. Le lemme permet de dire qu'il s'agit aussi du rang de D sur $\mathbb{R}$. Il s'ensuit que les colonnes $(I, M, \ldots M^{d-1})$ sont libres dans $\mathcal{M}_n(\mathbb{R})$. Il n'existe donc pas de polynôme annulateur de M à coefficients réels, non nul, de degré strictement inférieur à d.

On en déduit $\deg \mu_M \geqslant d = \deg \pi_M$. Comme on a $\mu_M | \pi_M$ et comme les deux polynômes sont unitaires, on peut affirmer que $\boxed{\pi_M = \mu_M}$. ◁

Le résultat s'étend à un corps quelconque K *et une extension* L *de* K, *avec la même démonstration.*

Dans l'exercice suivant on montre que, deux matrices A *et* B *de* $\mathcal{M}_n(K)$ *étant données, si* A *et* B *sont semblables lorsqu'on se place dans un surcorps de* K, *alors elles sont semblables dans* $\mathcal{M}_n(K)$. *La preuve donnée ici est valable avec un corps infini, le cas général nécessitant le théorème de caractérisation des classes de similitude à l'aide des facteurs invariants. Notons que la première question est souvent posée comme exercice d'oral. Quant à la seconde, on reconnaîtra qu'elle n'est plus dans l'esprit des programmes actuels.*

2.15. Similitude et extension de corps

1. Montrer que deux matrices carrées réelles semblables dans $\mathcal{M}_n(\mathbb{C})$ sont également semblables dans $\mathcal{M}_n(\mathbb{R})$.

2. Soit K un corps commutatif infini, L un sur-corps de K, M et N deux matrices de $\mathcal{M}_n(K)$. On suppose M et N semblables dans $\mathcal{M}_n(L)$. Démontrer qu'elles le sont également dans $\mathcal{M}_n(K)$.

(École polytechnique)

▷ **Solution.**

1. Soit M et N deux matrices de $\mathcal{M}_n(\mathbb{R})$ semblables dans $\mathcal{M}_n(\mathbb{C})$: il existe donc $P \in GL_n(\mathbb{C})$ telle que $N = P^{-1}MP$, ce qu'on peut aussi écrire $PN = MP$. Écrivons $P = Q + iR$ avec $Q, R \in \mathcal{M}_n(\mathbb{R})$. On a donc

$$(Q + iR)N = M(Q + iR)$$

et en identifiant partie réelle et imaginaire, il vient

$$QN = MQ \quad \text{et} \quad RN = MR.$$

Si l'une des deux matrices, Q ou R est inversible, M et N sont bien semblables dans $\mathcal{M}_n(\mathbb{R})$. Seulement, il se peut qu'aucune de ces deux matrices ne soit inversible. Dans ce cas, considérons pour $\lambda \in \mathbb{C}$, la matrice $P_\lambda = Q + \lambda R \in \mathcal{M}_n(\mathbb{C})$. La fonction $\lambda \longmapsto \det(P_\lambda)$ est une fonction polynomiale en λ non identiquement nulle puisqu'en i elle vaut $\det P \neq 0$. Il existe donc $\lambda_0 \in \mathbb{R}$ tel que $\det(P_{\lambda_0}) \neq 0$. Comme $P_{\lambda_0}N = MP_{\lambda_0}$ et P_{λ_0} est inversible dans $\mathcal{M}_n(\mathbb{R})$, M et N sont bien semblables en tant que matrices réelles.

2. Une généralisation de ce qui précède pour des corps quelconques est exclue. Considérons le système linéaire homogène (S) : $PN - MP = 0$ où l'inconnue est $P = (p_{i,j})_{1 \leqslant i,j \leqslant n}$. C'est un système à n^2 inconnues et n^2 équations et on cherche en fait à montrer qu'il admet une solution $P \in GL_n(K)$ sachant qu'il en a une dans $GL_n(L)$. Notons r le rang de ce système c'est-à-dire le rang de la matrice de ce système qui est de taille n^2 et à coefficients dans K. Comme (S) a une solution non nulle dans L on a forcément $r < n^2$. En effet, que le rang de (S) est le même que l'on considère le système à coefficients dans K ou dans L (cf. exercice 2.14). Posons $m = n^2 - r$. L'espace des solutions du système dans K (resp. dans L) est donc un espace de dimension m sur K (resp. sur L).

Choisissons une base $(P_1, \ldots, P_m)$ de l'espace des K-solutions du système. L'invariance du rang par extension de corps que l'on vient de rappeler montre que cette famille $(P_1, \ldots, P_m)$ est encore de rang m dans l'espace $\mathcal{M}_n(L)$. Autrement dit, l'espace des K-solutions est $KP_1 \oplus \cdots \oplus KP_m$ et l'espace des L-solutions est $LP_1 \oplus \cdots \oplus LP_m$. On sait par hypothèse qu'il existe une matrice inversible $Q \in GL_n(L)$ qui est solution. Elle se décompose sous la forme $Q = \alpha_1 P_1 + \cdots + \alpha_m P_m$, les α_i étant dans L. On cherche à prouver qu'il existe des scalaires $\lambda_1, \ldots, \lambda_m$ de K tels que $P = \lambda_1 P_1 + \cdots + \lambda_m P_m$ soit encore inversible.

Considérons pour cela le polynôme $\det(X_1P_1 + \cdots + X_mP_m)$ de $K[X_1, \ldots, X_m]$. Ce polynôme est non nul puisque sa valeur en $(\alpha_1, \ldots, \alpha_m)$ est non nulle (car Q est inversible). Comme K est infini (c'est ici seulement que sert cette hypothèse), ce polynôme ne peut

être identiquement nul sur K^m : il existe donc $(\lambda_1, \ldots, \lambda_m) \in \mathrm{K}^m$ tel que $\mathrm{P} = \lambda_1 \mathrm{P}_1 + \cdots + \lambda_m \mathrm{P}_m$ ait un déterminant non nul. La matrice $\mathrm{P} \in \mathcal{M}_n(\mathrm{K})$ est inversible et vérifie $\mathrm{QN} = \mathrm{MQ}$: les matrices M et N sont donc semblables dans $\mathcal{M}_n(\mathrm{K})$. ◁

Le résultat est vrai en toutes généralités mais à notre connaissance il n'y en a pas de preuve élémentaire, c'est-à-dire qui ne passe pas par la détermination des classes de similitude.

Nous en venons maintenant aux exercices concernant la diagonalisation.

2.16. Diagonalisabilité d'une matrice

Soit K un sous-corps de $\mathbb{C}$ et $(a_1, \ldots, a_n) \in \mathrm{K}^n$. On considère la matrice

$$\mathrm{M} = \begin{pmatrix} 0 & 0 & \ldots & 0 & a_1 \\ 0 & \ddots & & 0 & a_2 \\ \vdots & & \ddots & \vdots & \vdots \\ 0 & 0 & \ldots & 0 & a_{n-1} \\ a_1 & a_2 & \ldots & a_{n-1} & a_n \end{pmatrix}.$$

À quelle condition A est-elle diagonalisable ?

(École Polytechnique)

▷ **Solution.**

Lorsque $a_1 = \cdots = a_{n-1} = 0$, la matrice A est diagonale. Nous supposerons désormais l'un des a_i pour $i \in [\![1, n-1]\!]$ non nul. La matrice A est de rang 2 et donc $\dim \operatorname{Ker} \mathrm{A} = n-2$. Regardons les valeurs propres non nulles : soit $\lambda \in \mathrm{K}^*$, $\mathrm{X} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ et le système (S) $\mathrm{AX} = \lambda \mathrm{X}$.

Explicitons (S) :

$$(\mathrm{S}) \iff (\forall i \in [\![1, n-1]\!],\ a_i x_n = \lambda x_i) \text{ et } (a_1 x_1 + \cdots + a_n x_n = \lambda x_n\ (*)).$$

Comme λ est non nul, on peut multiplier $(*)$ par λ qui devient

$$a_1^2 x_n + \cdots + a_{n-1}^2 x_n + \lambda a_n x_n = \lambda^2 x_n$$

ou encore

$$(\lambda^2 - a_n\lambda - (a_1^2 + \cdots + a_{n-1}^2))x_n = 0.$$

Le système n'est pas de Cramer si, et seulement si

$$\lambda^2 - a_n\lambda - (a_1^2 + \cdots + a_{n-1}^2) = 0$$

et dans ces conditions, le sous-espace propre associé à λ est de dimension 1 (le rang du système étant $n-1$). Pour que A soit diagonalisable, il faut et il suffit que la somme directe des sous-espaces propres soit de dimension n, autrement dit dans notre cas, il faut et il suffit que le polynôme $X^2 - a_nX - (a_1^2 + \cdots + a_{n-1}^2)$ possède deux racines distinctes non nulles, autrement dit

$$a_1^2 + \cdots + a_{n-1}^2 \neq 0 \quad \text{et} \quad a_n^2 + 4(a_1^2 + \cdots + a_{n-1}^2) \neq 0. \lhd$$

On remarquera que si $K = \mathbb{R}$, *ces conditions sont toujours réalisées, ce à quoi on pouvait s'attendre, puisqu'alors* A *est symétrique réelle, donc diagonalisable.*

2.17. Racine d'une matrice diagonalisable inversible

Soit $A \in GL_n(\mathbb{C})$. On suppose qu'il existe $p \in \mathbb{N}^*$ tel que A^p soit diagonalisable. Montrer que A est diagonalisable.

La propriété subsiste-t-elle si A n'est pas inversible ?

(**École polytechnique**)

▷ **Solution.**

Il existe $\lambda_1, \ldots, \lambda_r$ dans $\mathbb{C}$ deux à deux distincts tels que $(X - \lambda_1)\ldots(X - \lambda_r)$ annule A^p. Par conséquent, le polynôme $(X^p - \lambda_1)\ldots(X^p - \lambda_r)$ annule A. Ce polynôme est scindé à racines simples, car pour tout $\lambda \in \mathbb{C}^*$, $X^p - \lambda$ est scindé à racines simples (dans $\mathbb{C}^n$, tout nombre complexe non nul admet exactement p racines p-ièmes) et si $\lambda \neq \mu$, $X^p - \lambda$ et $X^p - \mu$ n'ont aucune racine en commun. Il s'ensuit que A est diagonalisable.

La propriété ne subsiste pas si A n'est pas inversible : par exemple, si A est nilpotente non nulle, $A^n = 0$ est diagonalisable sans que A le soit. ◁

Voici une étude plus détaillée du cas $p = 2$, *sur le corps* $\mathbb{C}$, *et aussi sur le corps* $\mathbb{R}$.

2.18. Diagonalisabilité de f dans le cas f^2 diagonalisable

1. Soit $f \in \mathcal{L}(\mathbb{C}^n)$. Montrer que f est diagonalisable si, et seulement si, f^2 est diagonalisable et $\operatorname{Ker} f = \operatorname{Ker} f^2$.

2. Soit $f \in \mathcal{L}(\mathbb{R}^n)$. On suppose f^2 diagonalisable. À quelle condition nécessaire et suffisante f est-elle diagonalisable ?

3. À quelle condition nécessaire et suffisante la matrice complexe (resp. réelle)

$$\mathrm{A} = \begin{pmatrix} 0 & 0 & \dots & a_n \\ 0 & 0 & \dots & 0 \\ 0 & a_2 & & 0 \\ a_1 & 0 & \dots & 0 \end{pmatrix}$$

est-elle diagonalisable ?

(École Polytechnique)

▷ **Solution.**

1. Si f est diagonalisable, il existe une base $\mathcal{B}$ de $\mathbb{C}^n$ telle que la matrice de f dans cette base soit $\operatorname{Diag}(\lambda_1, \dots, \lambda_r, 0, \dots, 0)$ avec λ_i valeur propre non nulle de f. La matrice de f^2 dans cette même base est $\operatorname{Diag}(\lambda_1^2, \dots, \lambda_r^2, 0 \dots, 0)$: l'endomorphisme f^2 est donc diagonalisable et $\operatorname{rg} f^2 = \operatorname{rg} f = r$. Comme on a $\operatorname{Ker} f \subset \operatorname{Ker} f^2$ il en découle que $\operatorname{Ker} f^2 = \operatorname{Ker} f$.

Réciproquement, supposons f^2 diagonalisable et $\operatorname{Ker} f = \operatorname{Ker} f^2$. Notons $\mu_1, \dots, \mu_r$ les valeurs propres non nulles de f^2. Pour $1 \leqslant i \leqslant p$, on prend $\lambda_i \in \mathbb{C}^*$ telle que $\lambda_i^2 = \mu_i$. Comme $\mathrm{X} - \lambda_i$ et $\mathrm{X} + \lambda_i$ sont premiers entre eux, le théorème de décomposition des noyaux donne $\operatorname{Ker}(f^2 - \mu_i \operatorname{Id}) = \operatorname{Ker}(f - \lambda_i \operatorname{Id}) \oplus \operatorname{Ker}(f + \lambda_i \operatorname{Id})$. Dans ces conditions, on peut écrire

$$\begin{aligned} \mathbb{C}^n &= \operatorname{Ker} f^2 \oplus \left(\bigoplus_{i=1}^{r} \operatorname{Ker}(f^2 - \mu_i \operatorname{Id}) \right) \\ &= \operatorname{Ker} f \oplus \left(\bigoplus_{i=1}^{r} (\operatorname{Ker}(f - \lambda_i \operatorname{Id}) \oplus \operatorname{Ker}(f + \lambda_i \operatorname{Id})) \right), \end{aligned}$$

si bien que $\mathbb{C}^n$ est somme de sous-espaces propres de f, *i.e.* f est diagonalisable.

2. Si f est diagonalisable, en écrivant la matrice de f dans une base de vecteurs propres, il apparaît que $\operatorname{Ker} f = \operatorname{Ker} f^2$ et que les valeurs propres de f^2 sont positives ou nulles.

Réciproquement, si $\operatorname{Ker} f = \operatorname{Ker} f^2$ et $\operatorname{Sp} f^2 \subset \mathbb{R}_+$, ce qui a été fait dans la première question peut être repris (chaque valeur propre μ non

nulle de f^2 possède bien une racine carrée λ dans $\mathbb{R}$), et on démontre de manière analogue que f est diagonalisable.

La condition nécessaire et suffisante est donc $\operatorname{Ker} f = \operatorname{Ker} f^2$ et $\operatorname{Sp} f^2 \subset \mathbb{R}_+$.

3. Soit $(e_1, \ldots, e_n)$ la base canonique de $\mathbb{C}^n$ (resp. $\mathbb{R}^n$). On a pour tout $1 \leqslant i \leqslant n$, $Ae_i = a_i e_{n-i+1}$ et donc $A^2 e_i = a_i a_{n-i+1} e_i$. Il s'ensuit que A^2 est diagonalisable. Pour que A le soit, il faut et il suffit que $\operatorname{Ker} A = \operatorname{Ker} A^2$ si $K = \mathbb{C}$ (resp. $\operatorname{Ker} A = \operatorname{Ker} A^2$ et les produits $a_i a_{n-i+1}$ positifs ou nuls si $K = \mathbb{R}$). Traduisons la condition $\operatorname{Ker} A = \operatorname{Ker} A^2$. Le sous-espace $\operatorname{Ker} A$ est engendré par les e_i tels que $a_i = 0$ et le sous-espace $\operatorname{Ker} A^2$ est engendré par les e_i tels que $a_i a_{n-i+1} = 0$. Pour que $\operatorname{Ker} A = \operatorname{Ker} A^2$, il faut et il suffit que pour tout i,

$$a_i a_{n-i+1} = 0 \implies a_i = 0,$$

ce qui est encore équivalent à, pour tout i, $a_i = 0 \iff a_{n-i+1} = 0$.

Conclusion. La matrice complexe (resp. réelle) A est diagonalisable si et seulement si pour tout $1 \leqslant i \leqslant n$, $a_i = 0 \iff a_{n-i+1} = 0$ (resp. a_i et a_{n-i+1} sont du même signe strict pour tout i). ◁

L'algèbre des matrices diagonales est commutative. Pour qu'on puisse diagonaliser simultanément plusieurs endomorphismes il est donc nécessaire que ceux-ci commutent. En fait, cela est suffisant et c'est ce résultat que propose l'exercice suivant.

2.19. Diagonalisation simultanée

Soit E un espace vectoriel de dimension finie.

1. Montrer que la restriction d'un endomorphisme diagonalisable à un sous-espace stable est diagonalisable.

2. Soit $A \subset \mathcal{L}(E)$ formée d'endomorphismes diagonalisables commutant deux à deux. Montrer que les éléments de A admettent une base commune de diagonalisation.

(**École polytechnique**)

▷ **Solution.**

1. Soit u un endomorphisme diagonalisable du K-espace vectoriel E. Il existe P polynôme de K[X], scindé à racines simples annulant u. Supposons que u laisse stable un sous-espace F. Alors P annule aussi $u_{|F}$ et $u_{|F}$ est diagonalisable.

2. La partie A engendre un sous-espace de $\mathcal{L}(\mathrm{E})$ qui est de dimension finie. Il existe alors une famille $(u_1, \ldots, u_p)$ de A constituant une base de ce sous-espace. Si on trouve une base diagonalisant tous les éléments de cette famille, tout endomorphisme de A s'exprimant comme combinaison linéaire des u_i , il sera diagonalisable dans cette base.

Montrons l'existence d'une base commune de diagonalisation des u_i par une récurrence sur $p \geqslant 1$. C'est trivial si $p = 1$. Supposons $p \geqslant 2$ et le résultat vrai au rang $p-1$. Comme u_p est diagonalisable, si $\lambda_1, \ldots, \lambda_k$ désignent les valeurs propres deux à deux distinctes de u_p, on peut écrire

$$\mathrm{E} = \bigoplus_{i=1}^{k} \mathrm{Ker}(u_p - \lambda_i \mathrm{Id}).$$

Comme chaque u_j, pour $1 \leqslant j \leqslant p-1$, commute avec u_p, u_j laisse stable chaque sous-espace propre $\mathrm{Ker}(u_p - \lambda_i \mathrm{Id})$. Notons u_{ji} l'endomorphisme induit par u_j sur $\mathrm{Ker}(u_p - \lambda_i \mathrm{Id})$. À i fixé dans $\{1, 2, \ldots, k\}$, les u_{ji} en tant que restrictions d'endomorphismes diagonalisables, sont eux-mêmes diagonalisables et commutent deux à deux. Il y en a $p-1$ et ils sont donc justifiables de l'hypothèse de récurrence : il existe $\mathcal{B}_i$, base de $\mathrm{Ker}(u_p - \lambda_i \mathrm{Id})$, qui diagonalise chaque u_{ji}. La base $\mathcal{B}_i$ est donc constituée de vecteurs propres pour u_j, $1 \leqslant j \leqslant p-1$, mais aussi pour u_p puisque les vecteurs de $\mathcal{B}_i$ sont dans $\mathrm{Ker}(u_p - \lambda_i \mathrm{Id})$. Comme E est somme directe des sous-espaces propres de u_p, $(\mathcal{B}_1, \mathcal{B}_2, \ldots, \mathcal{B}_k)$ est une base de E constituée de vecteurs propres pour $u_1, \ldots, u_p$: c'est une base commune de diagonalisation. ◁

Il est aussi possible de faire une récurrence (forte) sur la dimension de l'espace E. *Si tous les éléments de* A *sont des homothéties le résultat est trivial. Sinon, on choisit un élément* u *de* A *qui n'est une homothétie. Les sous-espaces propres de* u *sont stables par tous les éléments de* A *et de dimensions strictement inférieures à celle de* E. *Les endomorphismes induits sur ces sous-espaces commutent deux à deux ce qui permet d'appliquer l'hypothèse de récurrence. Le lecteur pourra aussi montrer par une récurrence du même type que, sur un* $\mathbb{C}$*-espace vectoriel de dimension finie* E, *toute famille commutative d'endomorphismes de* E *admet un vecteur propre commun.*

Voici un exemple où l'on codiagonalise, mais élémentairement, une sous-algèbre commutative de $\mathcal{M}_n(\mathbb{C})$.

2.20. Matrices circulantes (1)

Soit $\mathcal{A}$ l'ensemble des matrices de la forme

$$\begin{pmatrix} a_0 & a_{n-1} & \dots & a_1 \\ a_1 & a_0 & \dots & a_2 \\ \vdots & \vdots & \dots & \vdots \\ a_{n-1} & a_{n-2} & \dots & a_0 \end{pmatrix}$$

de $\mathcal{M}_n(\mathbb{C})$ (on les appelle des *matrices circulantes*).

1. Montrer que $\mathcal{A}$ est une sous-algèbre commutative de $\mathcal{M}_n(\mathbb{C})$.

2. Montrer que les éléments de $\mathcal{A}$ sont simultanément diagonalisables.

3. On remplace $\mathbb{C}$ par un corps K dans lequel $X^n - 1$ est scindé. Le résultat de la question 2 subsiste-t-il ? Discuter selon la caractéristique de K.

(ENS Ulm-Lyon-Cachan)

▷ **Solution.**

1. Chaque ligne d'une matrice $A \in \mathcal{A}$ s'obtient par permutation circulaire à partir de la précédente (ce qui explique le nom de matrices circulantes qu'on leur donne). Soit $J = \begin{pmatrix} 0 & \cdots & \cdots & 0 & 1 \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & 0 & \vdots \\ 0 & \cdots & 0 & 1 & 0 \end{pmatrix}$, c'est-à-dire $J = (m_{i,j})$ où $m_{i,j} = 1$ si $i \equiv j+1 \mod n$ et $m_{i,j} = 0$ sinon. C'est la matrice de permutation associée au n-cycle $(1, 2, \dots, n)$. Cette interprétation de J rend facile le calcul de ses puissances : J agit sur les vecteurs de la base canonique $(e_1, \dots, e_n)$ de $\mathbb{C}^n$ en augmentant l'indice d'une unité (modulo n). On obtient alors facilement que, pour tout $k \in \mathbb{N}$, $J^k = (m_{i,j}^k)$, où $m_{i,j}^k = 1$ si $i \equiv j+k \mod n$ et $m_{i,j}^k = 0$. On obtient en particulier que $J^n = I_n$.

Ainsi une matrice $A = \begin{pmatrix} a_0 & a_{n-1} & \dots & a_1 \\ a_1 & a_0 & \dots & a_2 \\ \vdots & \vdots & \dots & \vdots \\ a_{n-1} & a_{n-2} & \dots & a_0 \end{pmatrix}$ de $\mathcal{A}$ s'écrit tout simplement comme un polynôme en J : $A = \sum_{k=0}^{n-1} a_k J^k$. Donc $\mathcal{A} =$

$\mathrm{Vect}(\mathrm{I}, \mathrm{J}, \ldots, \mathrm{J}^{n-1})$ est un sous-espace vectoriel de $\mathcal{M}_n(\mathbb{C})$. La famille $(\mathrm{I}, \mathrm{J}, \ldots, \mathrm{J}^{n-1})$ est libre donc $\mathcal{A}$ est de dimension n. Comme le produit de deux vecteurs de cette base appartient encore à la base et que ce produit est commutatif, par linéarité $\mathcal{A}$ est stable pour la multiplication et la multiplication est commutative dans $\mathcal{A}$. Enfin $\mathcal{A}$ contient la matrice I_n, et est donc une sous-algèbre commutative de $\mathcal{M}_n(\mathbb{C})$.

2. Puisque $\mathrm{J}^n = \mathrm{I}_n$, la matrice J possède un polynôme annulateur scindé à racines simples $\mathrm{X}^n - 1$. Il existe donc D diagonale et $\mathrm{P} \in \mathrm{GL}_n(\mathbb{C})$ telles que $\mathrm{J} = \mathrm{PDP}^{-1}$. On obtient, pour tout $k \in \mathbb{N}$, $\mathrm{J}^k = \mathrm{PD}^k\mathrm{P}^{-1}$ et avec les notations précédentes, pour tout $\mathrm{A} \in \mathcal{A}$,

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

Comme $\displaystyle\sum_{k=0}^{n-1} a_k \mathrm{D}^k$ est diagonale, toutes les matrices de $\mathcal{A}$ sont diagonalisables dans la même base que J. Les valeurs propres de J étant les racines n-ièmes de l'unité, celles de A sont les $\displaystyle\sum_{k=0}^{n-1} a_k \omega^k$, où $\omega \in \mathrm{U}_n$.

3. Si K est un corps quelconque, on exprime de la même façon les matrices de $\mathcal{A}$ en fonction de J. On suppose $n \geqslant 2$ sinon toutes les matrices sont diagonales.

• Si K est de caractéristique nulle, le polynôme $\mathrm{P} = \mathrm{X}^n - 1$ est encore scindé à racines simples puisque $\mathrm{P}' = n\mathrm{X}^{n-1}$ a comme seule racine 0. Le résultat de la question 2 subsiste.

• Supposons que K soit de caractéristique p, nombre premier.

Si p ne divise pas n, la seule racine de P' est encore 0 qui n'est pas racine de P. Le résultat subsiste. Si p divise n, on pose $n = pq$. On a alors $\mathrm{X}^n - 1 = \mathrm{X}^{pq} - 1 = (\mathrm{X}^q - 1)^p$, car pour $1 \leqslant k \leqslant p-1$, p divise C_p^k. Les valeurs propres de J sont des racines de $\mathrm{X}^q - 1$. Si J est diagonalisable, on a $\mathrm{J}^q = \mathrm{I}_n$. C'est manifestement faux : le plus petit entier pour lequel $\mathrm{J}^k = \mathrm{I}_n$ est n. Donc J n'est pas diagonalisable et le résultat ne subsiste pas. ◁

On peut vérifier de suite que, pour toute racine n-ième de l'unité ω, le vecteur de coordonnées $(1, \omega, \ldots, \omega^{n-1})$ est une base de la droite propre de J *associée à la valeur propre ω.*

Si on se place dans l'espace quotient $\mathbb{C}[\mathrm{X}]/(\mathrm{X}^n - 1)$ muni de la base $(1, \mathrm{X}, \ldots, \mathrm{X}^{n-1})$, la multiplication par X *(plus précisément par la classe de* X*) induit un n-cycle sur les vecteurs de la base. Cela permet de donner une réalisation abstraite des matrices circulantes étudiées dans l'exercice*

précédent. C'est ce que propose l'énoncé suivant qui se place toutefois dans $\mathbb{C}_{n-1}[X]$ *pour éviter de parler de quotient...*

2.21. Matrices circulantes (2)

Soit $F \in \mathbb{C}_{n-1}[X]$. $F = \sum_{k=0}^{n-1} a_k X^k$ et $\Phi : \mathbb{C}_{n-1}[X] \longrightarrow \mathbb{C}_{n-1}[X]$ l'application qui à P associe le reste de la division euclidienne de PF par $X^n - 1$.

1. Vérifier que Φ est linéaire et écrire sa matrice dans la base canonique.

2. Trouver les valeurs et les vecteurs propres de l'application linéaire Φ. Est-elle diagonalisable ?

(École polytechnique)

▷ **Solution.**

1. Le degré de $\Phi(P)$ est inférieur à celui de $X^n - 1$, donc $\Phi(P)$ appartient à $\mathbb{C}_{n-1}[X]$. Montrons que Φ est linéaire. Si P_1 et P_2 sont deux éléments de $\mathbb{C}_{n-1}[X]$, R_1 et R_2 leurs images par Φ, il existe Q_1 et Q_2 dans $\mathbb{C}[X]$ tels que

$$FP_1 = (X^n - 1)Q_1 + R_1 \quad \text{et} \quad FP_2 = (X^n - 1)Q_2 + R_2.$$

Pour tout $(\lambda_1, \lambda_2) \in \mathbb{C}^2$, on a

$$F(\lambda_1 P_1 + \lambda_2 P_2) = (X^n - 1)(\lambda_1 Q_1 + \lambda_2 Q_2) + (\lambda_1 R_1 + \lambda_2 R_2).$$

Puisque $\lambda_1 R_1 + \lambda_2 R_2 \in \mathbb{C}_{n-1}[X]$, c'est le reste dans la division de $F(\lambda_1 P_1 + \lambda_2 P_2)$ par $X^n - 1$ et

$$\Phi(\lambda_1 P_1 + \lambda_2 P_2) = \lambda_1 R_1 + \lambda_2 R_2 = \lambda_1 \Phi(P_1) + \lambda_2 \Phi(P_2).$$

Déterminons $\Phi(X^j)$ pour $j \in [\![0, n-1]\!]$. On a $FX^j = \sum_{k=0}^{n-1} a_k X^{k+j}$. Si $k \geqslant n - j$, alors $n \leqslant k + j \leqslant 2n - 2$. On écrit $X^{k+j} = X^{k+j-n} X^n$, et donc $X^{k+j} \equiv X^{k+j-n}[X^n - 1]$ avec $0 \leqslant k + j - n \leqslant n - 2$. On obtient modulo $X^n - 1$

$$FX^j = \sum_{k=0}^{n-j-1} a_k X^{k+j} + \sum_{k=n-j}^{n-1} a_k X^{k+j} \equiv \sum_{k=0}^{n-j-1} a_k X^{k+j} + \sum_{k=n-j}^{n-1} a_k X^{k+j-n}.$$

Puisque ce dernier polynôme est dans $\mathbb{C}_{n-1}[X]$, on en déduit

$$\Phi(X^j) = \sum_{k=0}^{n-j-1} a_k X^{k+j} + \sum_{k=n-j}^{n-1} a_k X^{k+j-n} = \sum_{l=0}^{j-1} a_{l+n-j} X^l + \sum_{l=j}^{n-1} a_{l-j} X^l.$$

Nous en déduisons que la matrice M de Φ dans la base canonique de $\mathbb{C}_{n-1}[X]$ est

$$M = \begin{pmatrix} a_0 & a_{n-1} & \dots & a_1 \\ a_1 & a_0 & \dots & a_2 \\ \vdots & \vdots & \dots & \vdots \\ a_{n-1} & a_{n-2} & \dots & a_0 \end{pmatrix}.$$

2. La matrice de Φ dans la base canonique de $\mathbb{C}_{n-1}[X]$ est donc une matrice circulante. Nous n'allons pas utiliser les résultats de l'exercice 2.20 mais chercher directement les valeurs propres et vecteurs propres de F.

• Soit λ une valeur propre de Φ et P un vecteur propre associé. Par définition de Φ, il existe $Q \in \mathbb{C}[X]$ tel que $FP = (X^n - 1)Q + \lambda P$, c'est-à-dire

$$(F(X) - \lambda)P(X) = (X^n - 1)Q(X).$$

Si $\omega \in U_n$ (l'ensemble des racines n-ièmes de l'unité), on a l'égalité $P(\omega)(F(\omega) - \lambda) = 0$. On en déduit que si $\lambda \neq F(\omega)$, alors $P(\omega) = 0$, et donc que si, pour tout $\omega \in U_n$, $\lambda \neq F(\omega)$, le polynôme P est nul, puisqu'il appartient à $\mathbb{C}_{n-1}[X]$ et a n racines distinctes. Cela montre que $Sp(\Phi) \subset \{F(\omega), \omega \in U_n\}$.

• Réciproquement, soit $\omega \in U_n$. Alors ω est racine de $F(X) - F(\omega)$ ainsi que de $X^n - 1$. Il existe donc Q et P_ω dans $\mathbb{C}_{n-1}[X]$ tels que

$$Q(X) = \frac{F(X) - F(\omega)}{X - \omega} \quad \text{et} \quad P_\omega = \frac{X^n - 1}{X - \omega}.$$

On a alors $(F(X) - F(\omega))P_\omega(X) = (X^n - 1)Q(X)$, c'est-à-dire

$$F(X)P_\omega(X) = (X^n - 1)Q(X) + F(\omega)P_\omega(X).$$

On obtient donc $\Phi(P_\omega) = F(\omega)P_\omega$, et P_ω est vecteur propre de Φ pour la valeur propre $F(\omega)$. On en déduit que $Sp(\Phi = \{F(\omega), \omega \in U_n\}$.

• Montrons que la famille $(P_\omega)_{\omega \in U_n}$ est une famille libre. Supposons $\sum_{\omega \in U_n} \lambda_\omega P_\omega = 0$, où les λ_ω sont des complexes. En évaluant cela en tout $\omega_0 \in U_n$, il vient $\lambda_{\omega_0} P_{\omega_0}(\omega_0) = 0$. Comme $P_{\omega_0}(\omega_0) \neq 0$, car $X^n - 1$ n'a que des racines simples, on en déduit que $\lambda_{\omega_0} = 0$, et ceci pour tout

$\omega_0 \in \mathrm{U}_n$. La famille est libre et est donc une base de $\mathbb{C}_{n-1}[\mathrm{X}]$, puisqu'elle comporte n vecteurs.

L'endomorphisme Φ est donc diagonalisable, puisque $(\mathrm{P}_\omega)_{\omega\in\mathrm{U}_n}$ est une base de vecteurs propres. On notera que les $\mathrm{F}(\omega)$ ne sont pas nécessairement distincts. Pour $\lambda \in \mathrm{Sp}(\Phi)$, la dimension du sous-espace propre correspondant est égale à $\mathrm{Card}\{\omega \in \mathrm{U}_n, \lambda = \mathrm{F}(\omega)\}$. $\lhd$

Voici maintenant quelques exercices concernant l'étude de la diagonalisabilité de matrices définies par blocs. Les variantes sont nombreuses mais les techniques sont toujours un peu les mêmes.

2.22. Diagonalisabilité d'une matrice par blocs (1)

Soit $\mathrm{U} \in \mathcal{M}_n(\mathbb{C})$ et $\mathrm{V} = \begin{pmatrix} 0 & \mathrm{I}_n \\ \mathrm{U} & 0 \end{pmatrix} \in \mathcal{M}_{2n}(\mathbb{C})$.

1. Relier les sous-espaces propres de V à ceux de U.

2. Trouver une condition nécessaire et suffisante sur U pour que V soit diagonalisable.

(**École polytechnique**)

$\rhd$ **Solution.**

1. Soit X et Y dans $\mathbb{C}^n$, $\mathrm{Z} = \begin{pmatrix} \mathrm{X} \\ \mathrm{Y} \end{pmatrix} \in \mathbb{C}^{2n}$, $\lambda \in \mathbb{C}$. Le vecteur VZ est égal à $\begin{pmatrix} \mathrm{Y} \\ \mathrm{UX} \end{pmatrix}$ et on en déduit que

$$\begin{aligned} \mathrm{Z} \in \mathrm{Ker}(\mathrm{V} - \lambda \mathrm{I}_{2n}) &\iff \mathrm{Y} = \lambda \mathrm{X} \ \text{ et } \ \mathrm{UX} = \lambda \mathrm{Y} \\ &\iff \mathrm{Z} = \begin{pmatrix} \mathrm{X} \\ \lambda \mathrm{X} \end{pmatrix} \ \text{ et } \ \mathrm{X} \in \mathrm{Ker}(\mathrm{U} - \lambda^2 \mathrm{I}_n). \end{aligned}$$

On en déduit donc que les valeurs propres de V sont les racines carrées des valeurs propres de U. Comme de plus l'application $\mathrm{X} \mapsto \begin{pmatrix} \mathrm{X} \\ \lambda \mathrm{X} \end{pmatrix}$ de $\mathbb{C}^n$ dans $\mathbb{C}^{2n}$ est injective, si $\mu^2 = \lambda \in \mathrm{Sp}\,\mathrm{U}$, la dimension du sous-espace propre relatif à la valeur propre μ pour V est exactement la dimension du sous-espace propre relatif à λ pour U.

2. La solution la plus naturelle consiste à utiliser la question précédente. La matrice V est diagonalisable si et seulement si la somme des dimensions de ses espaces propres vaut $2n$. Tout complexe non nul admet deux racines carrées distinctes mais pas 0. On est obligé de distinguer le noyau de V. La question 1 montre que

$$\dim \operatorname{Ker} V + \sum_{\mu \in \operatorname{Sp} V \setminus \{0\}} \dim \operatorname{Ker}(V - \mu I_{2n}) = \dim \operatorname{Ker} U + 2 \sum_{\lambda \in \operatorname{Sp} U \setminus \{0\}} \dim \operatorname{Ker}(U - \mu I_n).$$

Cette somme vaut $2n$ si et seulement si $\operatorname{Ker} U = \{0\}$ et U est diagonalisable. Donc V est diagonalisable si et seulement si U est inversible et diagonalisable. ◁

Voici une autre approche possible qui ne fait pas appel à la question 1. On observe que $V^2 = \begin{pmatrix} U & 0 \\ 0 & U \end{pmatrix}$. *Il en découle que* V^2 *est diagonalisable si et seulement si* U *l'est. Si* V *est diagonalisable il en est de même de* V^2 *mais la réciproque est fausse en général (cf. exercice 2.17). Il faut en plus que* $\operatorname{Ker} V = \operatorname{Ker} V^2$ *ce qui ici n'a lieu que lorsque* U *est inversible.*

2.23. Diagonalisabilité d'une matrice par blocs (2)

Soit $A \in \mathcal{M}_n(\mathbb{C})$.

1. On pose $B = \begin{pmatrix} A & 0 \\ A & A \end{pmatrix} \in \mathcal{M}_{2n}(\mathbb{C})$. Donner une condition nécessaire et suffisante sur A pour B soit diagonalisable.

2. Même question avec $B = \begin{pmatrix} A & 0 \\ C & A \end{pmatrix} \in \mathcal{M}_{2n}(\mathbb{C})$ où $C \in \mathcal{M}_n(\mathbb{C})$ commute avec A. Que peut-on dire dans le cas où A et C ne commutent pas ?

(**École polytechnique**)

▷ **Solution.**

1. Regardons le cas $n = 1$ pour avoir des idées. Pour $a \in \mathbb{C}$, si $M = \begin{pmatrix} a & 0 \\ a & a \end{pmatrix}$ est diagonalisable, M est semblable à la matrice aI_2 car a est sa seule valeur propre. Mais alors $M = aI_2$, et nécessairement $a = 0$. Cela donne déjà une idée du résultat.

Supposons $B = \begin{pmatrix} A & 0 \\ A & A \end{pmatrix}$ diagonalisable. On va regarder les polynômes en B. On a $B^2 = \begin{pmatrix} A^2 & 0 \\ 2A^2 & A^2 \end{pmatrix}$ et $B^3 = \begin{pmatrix} A^3 & 0 \\ 3A^3 & A^3 \end{pmatrix}$ ce qui nous invite à penser que $B^k = \begin{pmatrix} A^k & 0 \\ kA^k & A^k \end{pmatrix}$ pour tout k. Cela se vérifie immédiatement par une récurrence sur k. Comme la matrice B est diagonalisable, elle est annulée par un polynôme scindé à racines simples P. On a alors

$P(A) = 0$ et aussi, en posant $P = a_k X^k + \cdots + a_0$,

$$ka_k A^k + (k-1)a_{k-1}A^{k-1} + a_1 A = 0, \quad i.e. \quad (XP')(A) = 0.$$

On en déduit déjà que A doit être diagonalisable (ce qu'on pouvait voir directement, car si $(e_1, \ldots, e_{2n})$ est la base canonique de $\mathbb{C}^{2n}$, le sous-espace $F = \mathrm{Vect}(e_{n+1}, \ldots, e_{2n})$ est stable par B et A est la matrice de cette restriction dans la base $(e_{n+1}, \ldots, e_{2n})$). Mais plus précisément, le polynôme minimal μ_A de A divise P et XP'. Comme les racines de P sont simples, P et P' sont premiers entre eux, et μ_A est donc égal à X. Ce qui signifie que $A = 0$.

Réciproquement, si $A = 0$, $B = 0$ est évidemment diagonalisable.

Conclusion. La matrice B est diagonalisable si, et seulement si $A = 0$.

2. • On utilise les mêmes idées. Si A et C commutent, on vérifie par récurrence que pour tout $k \geqslant 0$,

$$B^k = \begin{pmatrix} A^k & 0 \\ kCA^{k-1} & A^k \end{pmatrix}.$$

Supposons B diagonalisable. Son polynôme minimal μ_B est scindé à racines simples. Comme pour tout polynôme P on a $P(B) = \begin{pmatrix} P(A) & 0 \\ CP'(A) & P(A) \end{pmatrix}$, il vient d'une part $\mu_B(A) = 0$, et A est donc diagonalisable avec μ_A qui divise μ_B et d'autre part $C\mu'_B(A) = 0$. En fait on a $\chi_B = \chi_A^2$, donc A et B ont les mêmes valeurs propres. En particulier, on a $\mu_A = \mu_B$, et donc

$$C\mu'_A(A) = 0.$$

Or la matrice $\mu'_A(A)$ est inversible : en effet, comme μ_A est scindé à racines simples, μ_A et μ'_A sont premiers entre eux, si bien qu'il existe $P, Q \in \mathbb{C}[X]$ tels que $P\mu_A + Q\mu'_A = 1$, d'après le théorème de Bezout. En évaluant en A, il vient $Q(A)\mu'_A(A) = I_n$ et $\mu'_A(A)$ est bien inversible. Par conséquent $C = 0$.

Réciproquement, si A est diagonalisable et si $C = 0$, il existe P scindé à racines simples annulant A. Dans ces conditions, on a $P(B) = 0$ et B est diagonalisable.

• Dans le cas où C et A ne commutent pas, B peut être diagonalisable sans que C soit nul (par contre A doit être nécessairement diagonalisable). C'est ce que prouve l'exemple suivant traité avec Maple :

```
> with(linalg):

Warning, the protected names norm and trace have been redefined
and unprotected
```

```
> M:=[[1,1,0,0],[0,0,0,0],[0,1,1,1],[0,0,0,0]]:B:=evalm(M);
```

$$\mathrm{B} := \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

```
> P:=charpoly(B,x);mu:=minpoly(B,x);
```

$$\mathrm{P} := (x-1)^2 x^2$$

$$\mu := -x + x^2$$

Le polynôme minimal est scindé à racines simples, et donc la matrice est bien diagonalisable. ◁

Le résultat de l'exercice qui suit concerne la théorie algébrique des graphes. Il a aussi fait l'objet d'une partie de l'épreuve d'algorithmique du concours 1996 de l'ENS Lyon. On y utilise le fait qu'une matrice symétrique réelle est diagonalisable.

2.24. Théorème de Hoffman et Singleton (1960)

Soit A une matrice carrée réelle d'ordre n à coefficients dans $\{0,1\}$, de trace nulle, symétrique et telle qu'il existe un entier $d \geqslant 1$ vérifiant $\mathrm{A}^2 + \mathrm{A} - (d-1)\mathrm{I}_n = \mathrm{J}$ où J est la matrice carrée d'ordre n dont tous les coefficients valent 1. On note U le vecteur colonne dont tous les coefficients sont égaux à 1.

1. Montrer que $\mathrm{AU} = d\mathrm{U}$. En déduire que $n = d^2 + 1$.

2. Soit a, b les racines du polynôme $\mathrm{X}^2 + \mathrm{X} - (d-1)$. Montrer que le spectre de A est inclus dans $\{a, b, d\}$.

3. Montrer que $d \in \{1, 2, 3, 7, 57\}$. Déterminer A pour $d = 1$ ou $d = 2$.

(**École polytechnique**)

▷ **Solution.**

1. Notons $\mathrm{A}=(a_{ij})_{1\leqslant i,j\leqslant n}$ et $\mathrm{A}^2 = (b_{ij})_{1\leqslant i,j\leqslant n}$. La nullité de la trace de A implique la nullité de tous les coefficients diagonaux a_{ii}. Par ailleurs, on a

$$b_{ij} = \sum_{k=1}^{n} a_{ik}a_{kj} = \sum_{k=1}^{n} a_{ik}a_{jk},$$

car A est symétrique. Il s'agit du nombre de 1 communs à la ligne i et à la ligne j de A. En particulier, b_{ii} est égal au nombre de 1 qu'il y a

dans la ligne i. La relation vérifiée par A conduit à $b_{ii} + a_{ii} - (d-1) = 1$ c'est-à-dire $b_{ii} = d$. Or, AU est exactement le vecteur dont la i-ième coordonnée est la somme des termes de la i-ième ligne de A, somme qui vaut donc d. Ainsi, $AU = dU$.

Il vient alors $A^2U + AU - (d-1)U = JU = nU$, donc

$$d^2 + d - (d-1) = d^2 + 1 = n.$$

2. Soit λ une valeur propre de A, X un vecteur propre associé. On a alors $JX = (\lambda^2 + \lambda - (d-1))X$. Donc X est aussi un vecteur propre de J pour la valeur propre $\lambda^2 + \lambda - (d-1)$. La matrice J est de rang 1 et diagonalisable : ses valeurs propres sont 0 et n, l'espace propre associé à n étant la droite vectorielle dirigée par U. Il en résulte que, soit X est lié à D, et alors d'après la question 1 $\lambda = d$, soit $X \in \operatorname{Ker} J$, et alors $\lambda \in \{a, b\}$. C'est le résultat voulu.

3. Observons que a, b, d sont deux à deux distincts. La matrice A est symétrique réelle, donc diagonalisable. L'espace propre associé à d est la droite dirigée par U. Notons α et β les dimensions des espaces propres associés à a et b. On a $n = 1 + \alpha + \beta = d^2 + 1$ (1), car A est diagonalisable. Le fait que la trace de A est nulle conduit à la seconde relation $\alpha a + \beta b + d = 0$ (2). Les contraintes sur d vont provenir de ce que α et β doivent être *entiers*. Explicitons pour commencer a et b :

$$a = \frac{-1 + \sqrt{4d-3}}{2} \quad \text{et} \quad b = \frac{-1 - \sqrt{4d-3}}{2}.$$

En substituant dans (2) on obtient

$$-\frac{\alpha + \beta}{2} + \sqrt{4d-3}\frac{\alpha - \beta}{2} = -d$$

et en remplaçant $\alpha + \beta$ par d^2 : $\sqrt{4d-3}(\alpha - \beta) = d(d-2)$ (3).

- Si $\alpha = \beta$, on a nécessairement $d = 2$.
- Si $\alpha \neq \beta$. $\sqrt{4d-3}$ est un nombre rationnel. Donc $4d-3$ est un carré d'entier : $4d - 3 = p^2$. La relation (3) impose alors que p divise $d(d-2) = \dfrac{p^2+3}{4}\dfrac{p^2-5}{4}$ et donc que $16p$ divise $(p^2+3)(p^2-5)$. Cela impose $p|15$ c'est-à-dire $p \in \{1, 3, 5, 15\}$.

Les seules valeurs possibles de d sont donc $1, 2, 3, 7, 57$.

Si $d = 1$, alors $n = 2$. Les termes diagonaux de A sont nuls et chaque ligne de A contient un terme égal à 1. On trouve une seule solution $A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, dont on vérifie qu'elle convient.

Si $d = 2$, alors $n = 5$. Chaque ligne de A contient exactement deux termes égaux à 1 et les termes diagonaux sont nuls. Si on échange deux

lignes de A ainsi que les colonnes correspondantes, la matrice obtenue reste symétrique et il est clair, au vu des relations que doivent vérifier les coefficients, qu'elle est encore solution. On peut donc supposer qu'on a $a_{12} = a_{13} = 1$ et $a_{14} = a_{15} = 0$. Comme $b_{12} = 1 - a_{12} = 0$, les lignes 1 et 2 n'ont pas de 1 en commun et $a_{23} = 0$. On a donc $a_{32} = 0$. De $b_{23} = 1 - a_{23} = 1$, on déduit que les lignes 2 et 3 n'ont qu'un 1 en commun, le premier terme. On a donc $a_{24} = 0$ ou $a_{34} = 0$. Quitte à échanger les deux dernières lignes et colonnes, on peut supposer $a_{24} = 1$ et $a_{34} = 0$. Alors A est entièrement déterminée :

$$\mathrm{A} = \begin{pmatrix} 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 \end{pmatrix},$$

et cette matrice convient. Le raisonnement qui précède montre qu'une fois fixés les coefficients non nuls de la première ligne a_{1i} et a_{1j}, il y a deux possibilités pour les lignes i et j et qu'ensuite A est entièrement déterminé. On obtient donc $\mathrm{C}_4^2 \times 2 = 12$ matrices différentes qui s'obtiennent à partir de la précédente en permutant les lignes et les colonnes. $\lhd$

Exposons un peu de théorie des graphes pour bien comprendre le résultat que l'on vient de démontrer. Un graphe G *est un couple* (S, A)*, où* S *est un ensemble fini dont les éléments sont appelés les sommets et* A *un ensemble de paires d'éléments distincts de* S*, appelés les arêtes. Le cardinal de* S *est appelé l'ordre du graphe. Deux sommets x et y sont adjacents si $\{x, y\} \in \mathrm{A}$. Étant donné un sommet x, le nombre de sommets adjacents à x est appelé le degré de x. Un chemin de longueur p entre les sommets x et y est une suite $(x = s_0, s_1, \ldots, s_p = y)$ de sommets adjacents. La distance entre deux sommets x et y est la longueur minimale d'un chemin de x à y et le diamètre du graphe est la distance maximale entre deux sommets distincts.*

Un graphe d'ordre n peut être représenté par sa matrice d'adjacence $\mathrm{A} \in \mathcal{M}_n(\mathbb{R})$. On numérote les sommets $\mathrm{S} = \{x_i,\ i \in [\![1, n]\!]\}$ et on définit $\mathrm{A} = (a_{ij})_{1 \leqslant i,j \leqslant n}$ par $a_{ij} = 1$ si $\{x_i, x_j\} \in \mathrm{A}$ et $a_{ij} = 0$ sinon.

On considère un graphe d'ordre n, de diamètre 2 (deux sommets quelconques sont donc à une distance 1 ou 2) et de degré maximal d. On a alors $n \leqslant d^2 + 1$. En effet, si on fixe un sommet s, il possède au plus d voisins et chacun de ceux-ci possède au plus $d - 1$ voisins autre que s. Donc le graphe possède au plus $1 + d + d(d-1) = d^2 + 1$ sommets. S'il possède le nombre maximal de sommets, c'est-à-dire $d^2 + 1$, on dit que c'est un graphe de Moore.

Le raisonnement précédent montre qu'alors chaque sommet est de degré d (on dit que le graphe est d-régulier) et est relié à tout sommet distinct par exactement un chemin (d'ordre 1 ou 2).

La matrice d'adjacence A d'un graphe de Moore vérifie l'équation $A^2 + A - (d-1)I_n = J$. *En effet, avec les notations de l'énoncé, cela équivaut à* $b_{ij} + a_{ij} = 1 + (d-1)\delta_{ij}$. *On remarque que* $b_{ij} = \sum_{k=1}^{n} a_{ik}a_{ki}$ *est le nombre de chemins de longueur 2 du sommet i au sommet j.*

• *Pour $j = i$, on a $a_{ii} = 0$ et $b_{ii} = \sum_{k=1}^{n} a_{ik}^2 = d$, nombre de sommets voisins du i-ième et donc $a_{ii} + b_{ii} = d$.*

• *Si $i \neq j$, soit $a_{ij} = 0$ et alors les sommets i et j sont reliés par un chemin d'ordre 2 et $b_{ij} = 1$, soit $a_{ij} = 1$ et alors i et j ne sont pas relié par un chemin d'ordre 2 donc $b_{ij} = 0$. On a dans tous les cas $a_{ij} + b_{ij} = 1$.*

L'exercice montre que $d \in \{2, 3, 7, 57\}$ (le cas $d = 1$ doit être écarté car alors $n = 2$ et le diamètre est 1). Pour $d = 2$, $d = 3$ et $d = 7$, on a montré l'existence de graphes de Moore. Pour $d = 57$, on ignore encore s'il en existe un.

Voici maintenant une série d'exercices concernant la trigonalisation. Rappelons le résultat fondamental : une matrice est trigonalisable si et seulement si son polynôme caractéristique est scindé. Les exercices proposés concernent presque tous des questions de trigonalisation simultanée.

2.25. Trigonalisation simultanée (1)

Soit A, B deux matrices de $\mathcal{M}_n(\mathbb{C})$ qui commutent. Montrer que A et B possèdent un vecteur propre commun, puis établir qu'elles sont trigonalisables simultanément.

(École polytechnique)

▷ **Solution.**

• Comme $\mathbb{C}$ est algébriquement clos, le polynôme caractéristique de A qui est de degré $n \geqslant 1$ admet au moins une racine. Donc A admet au moins une valeur propre λ. L'espace propre $\mathrm{Ker}(A - \lambda I_n)$ est stable par B car A et B commutent. Le même argument montre que la restriction de B à cet espace propre admet une valeur propre μ. Si X est un vecteur propre associé on a alors $BX = \mu X$ et $AX = \lambda X$. Donc X répond au problème.

• On va montrer par récurrence sur n qu'il existe $S \in GL_n(\mathbb{C})$ tel que $S^{-1}AS$ et $S^{-1}BS$ soient triangulaires supérieures. C'est trivial si $n = 1$. Supposons $n \geqslant 2$ et le résultat vrai au rang $n-1$. D'après la première question, il existe $e \in \mathbb{C}^n$ non nul et λ et μ complexes vérifiant $Ae = \lambda e$ et $Be = \mu e$. Complétons e en une base $(e, \varepsilon_2, \ldots, \varepsilon_n)$ de $\mathbb{C}^n$. Si Q désigne la matrice de passage de la base canonique de $\mathbb{C}^n$ à $(e, \varepsilon_2, \ldots, \varepsilon_n)$, on a :

$$Q^{-1}AQ = \left(\begin{array}{c|c} \lambda & \times \cdots \times \\ \hline 0 & \\ \vdots & A' \\ 0 & \end{array}\right) \quad \text{et} \quad Q^{-1}BQ = \left(\begin{array}{c|c} \mu & \times \cdots \times \\ \hline 0 & \\ \vdots & B' \\ 0 & \end{array}\right).$$

Comme A et B commutent, il en va de même pour A' et B'. D'après l'hypothèse de récurrence, il existe $R \in GL_{n-1}(\mathbb{C})$ tel que $R^{-1}A'R = T$ et $R^{-1}B'R = U$ soient triangulaires supérieures. Posons $P = \left(\begin{array}{c|c} 1 & 0 \cdots 0 \\ \hline 0 & \\ \vdots & R \\ 0 & \end{array}\right)$. Alors $P \in GL_n(\mathbb{C})$ et $P^{-1} = \left(\begin{array}{c|c} 1 & 0 \cdots 0 \\ \hline 0 & \\ \vdots & R^{-1} \\ 0 & \end{array}\right)$.

Par conséquent :

$$P^{-1}Q^{-1}AQP = (QP)^{-1}A(QP) = \left(\begin{array}{c|c} \lambda & \times \cdots \times \\ \hline 0 & \\ \vdots & R^{-1}A'R \\ 0 & \end{array}\right)$$

$$= \left(\begin{array}{c|c} \lambda & \times \cdots \times \\ \hline 0 & \\ \vdots & T \\ 0 & \end{array}\right) \in T_n(\mathbb{C})$$

et de même,

$$P^{-1}Q^{-1}BQP = (QP)^{-1}B(QP) = \left(\begin{array}{c|c} \mu & \times \cdots \times \\ \hline 0 & \\ \vdots & R^{-1}B'R \\ 0 & \end{array}\right)$$

$$= \left(\begin{array}{c|c} \mu & \times \cdots \times \\ \hline 0 & \\ \vdots & U \\ 0 & \end{array}\right) \in T_n(\mathbb{C}).$$

La matrice inversible QP convient. ◁

On a déjà signalé plus haut (cf. page 93) que si $(A_i)_{i\in I}$ est une famille commutative quelconque de matrices de $\mathcal{M}_n(\mathbb{C})$ on peut trouver un vecteur propre commun à toutes les matrices de la famille. La même récurrence montre alors que la famille $(A_i)_{i\in I}$ est cotrigonalisable.

Bien entendu, comme deux matrices triangulaires ne commutent pas forcément, le fait que A et B commutent est suffisant mais n'est pas nécessaire pour que A et B soient cotrigonalisables. On va voir dans la suite plusieurs autres conditions suffisantes. La variante suivante est une première généralisation, puisqu'on retrouve l'exercice précédent si $C = 0$. Les mêmes idées interviennent dans la solution.

2.26. Trigonalisation simultanée (2)

Soient A, B et C trois matrices de $\mathcal{M}_n(\mathbb{C})$ telles que

$$AB - BA = C, \quad AC = CA, \quad BC = CB.$$

1. Montrer que ces trois matrices ont un vecteur propre commun.

2. Montrer que ces trois matrices sont trigonalisables simultanément.

(École polytechnique)

▷ **Solution.**

1. On remarque que si X est un vecteur propre pour A et B, alors X est forcément un vecteur du noyau de C. On va donc regarder ce qui se passe sur le noyau de C.

• Supposons que $\operatorname{Ker} C$ n'est pas réduit à $\{0\}$. Alors les endomorphismes définis par A et B laissent stable ce sous-espace puisque A et B commutent avec C. Leurs restrictions A′ et B′ à $\operatorname{Ker} C$ commutent car si $X \in \operatorname{Ker} C$,

$$ABX - BAX = CX = 0.$$

Comme $\mathbb{C}$ est algébriquement clos, A′ et B′ ont un vecteur propre commun (cf. exercice précédent).

• Il reste donc à montrer que $\operatorname{Ker} C \neq \{0\}$. Supposons par l'absurde C inversible. On a alors

$$C^{-1}AB - C^{-1}BA = I_n$$

Or, $\operatorname{Tr}(C^{-1}AB) = \operatorname{Tr}(BC^{-1}A) = \operatorname{Tr}(C^{-1}BA)$ car B et C^{-1} commutent. C'est absurde car I_n serait de trace nulle.

2. On identifie les matrices et les endomorphismes canoniquement associés. En prenant une nouvelle base $\mathcal{B}$ qui admet comme premier vecteur un vecteur propre commun à A, B et C, on peut trouver $P \in \mathrm{GL}_n(\mathbb{C})$ telle que

$$P^{-1}AP = \left(\begin{array}{c|c} \lambda & \times \cdots \times \\ \hline 0 & \\ \vdots & A' \\ 0 & \end{array}\right), \quad P^{-1}BP = \left(\begin{array}{c|c} \mu & \times \cdots \times \\ \hline 0 & \\ \vdots & B' \\ 0 & \end{array}\right),$$

$$P^{-1}CP = \left(\begin{array}{c|c} 0 & \times \cdots \times \\ \hline 0 & \\ \vdots & C' \\ 0 & \end{array}\right).$$

On a alors $A'B' - B'A' = C'$, $A'C' = C'A'$ et $B'C' = C'B'$. Une récurrence sur la taille de la matrice permet de conclure. $\triangleleft$

Voici encore une version différente.

2.27. Trigonalisation simultanée (3)

1. Soient $A, B \in \mathcal{M}_n(\mathbb{C})$ telles que $AB - BA = B$. Montrer que B est nilpotente.

2. Soient $A, B \in \mathcal{M}_n(\mathbb{C})$ telles qu'il existe $\lambda, \mu \in \mathbb{C}$ vérifiant $AB - BA = \lambda A + \mu B$. Montrer que A et B ont un vecteur propre commun, puis qu'elles sont simultanément trigonalisables.

(**École polytechnique**)

▷ **Solution.**

1. On va calculer les puissances de B à partir de la relation $AB - BA = B$. En multipliant par B à droite on a $AB^2 - BAB = B^2$ et en multipliant par B à gauche $BAB - B^2A = B^2$. Ainsi, $2B^2 = AB^2 - B^2A$. Montrons alors par récurrence sur $k \geqslant 1$ que $AB^k - B^kA = kB^k$. C'est vu pour $k = 1$ ou $k = 2$. Supposons $k \geqslant 3$ et le résultat vrai au rang $k - 1$. On a par hypothèse de récurrence

$$\begin{aligned} AB^k - B^kA &= (AB^{k-1} - B^{k-1}A)B + B^{k-1}(AB - BA) \\ &= (k-1)B^k + B^k = kB^k. \end{aligned}$$

D'où le résultat. Pour en déduire que B est nilpotente, nous vous proposons pas moins de trois arguments.

• On peut utiliser le résultat classique de l'exercice 2.33, puisque pour tout $k \geqslant 1$ on a $\mathrm{Tr}(k\mathrm{B}^k) = \mathrm{Tr}(\mathrm{AB}^k) - \mathrm{Tr}(\mathrm{B}^k\mathrm{A}) = 0$, et donc $\mathrm{Tr}(\mathrm{B}^k) = 0$.

• Il est aussi possible d'utiliser la notion de polynôme minimal. En effet, les relations précédentes montrent que pour tout polynôme P on a

$$\mathrm{AP(B)} - \mathrm{P(B)A} = \mathrm{BP'(B)}$$

Si on applique cela en prenant pour P le polynôme minimal μ de B, on voit que $\mathrm{X}\mu'(\mathrm{X})$ annule aussi B. Donc $\mu(\mathrm{X})$ divise $\mathrm{X}\mu'(\mathrm{X})$. Ces deux polynômes ayant le même degré, ils sont colinéaires. La seule racine de μ est alors 0 (l'ordre de multiplicité d'une autre racine serait le même dans μ et dans μ' ce qui est impossible). Donc $\mu = \mathrm{X}^p$ pour un certain entier p et $\mathrm{B}^p = 0$: B est nilpotente.

• Voici enfin un dernier argument très court : l'application $f : \mathrm{M} \mapsto \mathrm{AM} - \mathrm{MB}$ est un endomorphisme de $\mathcal{M}_n(\mathbb{C})$. Les égalités obtenues plus haut signifient que si B^k est non nulle, alors k est dans le spectre de f. Comme $\mathcal{M}_n(\mathbb{C})$ est de dimension finie, le spectre de f est fini et les puissances de B sont donc forcément nulles à partir d'un certain rang.

2. Si $\lambda = \mu = 0$, A et B commutent et on se retrouve avec la situation de l'exercice 2.25. On supposera donc $(\lambda, \mu) \neq 0$ et pour des raisons de symétrie, on peut supposer $\mu \neq 0$. Comme dans les deux exercices précédent il suffit de prouver que A et B admettent un vecteur propre commun, la cotrigonalisation s'obtenant alors facilement par récurrence sur la taille n (la relation se transmet bien au rang inférieur).

Pour cela, on va essayer de se ramener à une relation de la forme étudiée dans la question 1. Posons $\mathrm{B}' = \lambda\mathrm{A} + \mu\mathrm{B}$. On a alors $\mathrm{AB}' - \mathrm{B}'\mathrm{A} = \mu\mathrm{B}'$. La question précédente, appliquée avec $\frac{1}{\mu}\mathrm{A}$ et B' montre que B' est nilpotente. En particulier son noyau est non nul. Si $\mathrm{X} \in \mathrm{Ker}\,\mathrm{B}'$ on a $\mathrm{B}'\mathrm{AX} = 0$ ce qui montre que $\mathrm{Ker}\,\mathrm{B}'$ est stable par A. Comme on travaille sur $\mathbb{C}$, la restriction de A à $\mathrm{Ker}\,\mathrm{B}'$ admet une valeur propre α. Soit X un vecteur propre associé. On a alors $\mathrm{B}'\mathrm{X} = 0$, ce qui donne $\lambda\mathrm{AX} + \mu\mathrm{BX} = 0$, soit encore $\mathrm{BX} = -\frac{\lambda\alpha}{\mu}\mathrm{X}$. Donc X est un vecteur propre commun à A et B. ◁

Voici encore une question du même type, mais plus difficile.

2.28. Trigonalisation simultanée (4)

Soient A et B dans $\mathcal{M}_n(\mathbb{C})$ telles que $\mathrm{rg}(\mathrm{AB} - \mathrm{BA}) \leqslant 1$. Montrer que A et B sont cotrigonalisables.

(ENS Ulm)

▷ **Solution.**

Posons $C = AB - BA$. Si C est nulle, alors A et B commutent et on retrouve la situation de l'exercice 2.25. On va donc supposer C de rang 1. Une première idée est de chercher un vecteur propre commun à A et B pour faire ensuite une récurrence sur la taille n. Il est clair qu'un tel vecteur propre commun X de A et B doit être dans le noyau de C. Par hypothèse, ce noyau est un hyperplan de $\mathbb{C}^n$. Soit λ une valeur propre de A. Quitte à remplacer A par $A' = A - \lambda I_n$, on peut supposer que $\lambda = 0$ (on a toujours $A'B - BA' = C$). Deux cas se présentent :

• Dans le cas où $\operatorname{Ker} A \subset \operatorname{Ker} C$, tout se passe bien. En effet, pour tout X de $\operatorname{Ker} A$ on a $ABX = CX + BAX = 0$. Donc $\operatorname{Ker} A$ est stable par B. Un vecteur propre de la restriction de B à $\operatorname{Ker} A$ est alors vecteur propre commun à A et B.

• Regardons maintenant le cas où $\operatorname{Ker} A$ n'est pas inclus dans $\operatorname{Ker} C$. On peut donc trouver $X \in \mathbb{C}^n$ tel que $AX = 0$ et $CX \neq 0$. On a alors $CX = ABX \in \operatorname{Im} A$. Comme CX est un vecteur non nul de l'image de C et que $\operatorname{Im} C$ est une droite, on a $\operatorname{Im} C \subset \operatorname{Im} A$. Dans ce cas on voit que $\operatorname{Im} A$ est stable par B : en effet, soit $Y \in \operatorname{Im} A$ et Z tel que $Y = AZ$. Alors $ABZ - BAZ = CZ$ ce qui donne $BY = ABZ - CZ \in \operatorname{Im} A$. Problème : on ne voit pas pourquoi A et B auraient un vecteur propre commun dans $\operatorname{Im} A$.

Mais en fait, il n'est pas essentiel de trouver un vecteur propre commun à A et B, c'est-à-dire une droite stable par A et B : n'importe quel sous-espace strict de $\mathbb{C}^n$ stable par A et B permet de faire une récurrence sur la dimension ! Et on vient justement de montrer que si $\lambda \in \operatorname{Sp} A$, alors soit $\operatorname{Ker}(A - \lambda I_n)$ soit $\operatorname{Im}(A - \lambda I_n)$ est stable par B (et naturellement aussi par A). Comme A n'est pas scalaire (sinon $C = 0$), ces deux sous-espaces sont non nuls et strictement inclus dans $\mathbb{C}^n$.

Rédigeons alors la récurrence. Le résultat est trivial pour $n = 1$. Supposons le résultat vrai pour tout couple de matrices de taille $< n$ et prenons A, B dans $\mathcal{M}_n(\mathbb{C})$ telles que $\operatorname{rg}(AB - BA) \leqslant 1$. Si $AB = BA$ on sait que A et B sont cotrigonalisables. Si $C = AB - BA$ est de rang 1, on vient de prouver qu'il existe un sous-espace non trivial F de $\mathbb{C}^n$ stable par A et par B. En prenant une base de F que l'on prolonge en base de $\mathbb{C}^n$, les endomorphismes canoniquement associés à A et B admettent dans cette base de $\mathbb{C}^n$ des matrices de la forme

$$A' = \begin{pmatrix} \boxed{A_1} & \times \\ 0 & \boxed{A_2} \end{pmatrix} \text{ et } B' = \begin{pmatrix} \boxed{B_1} & \times \\ 0 & \boxed{B_2} \end{pmatrix} \text{ respectivement.}$$

On a alors

$$A'B' - B'A' = \begin{pmatrix} \boxed{A'_1B'_1 - B'_1A_1} & \times \\ 0 & \boxed{A'_2B'_2 - B'_2A_2} \end{pmatrix}.$$

Puisque $\mathrm{rg}(A'B' - B'A') = 1$, on a $\mathrm{rg}(A'_1B'_1 - B'_1A_1) \leqslant 1$ et $\mathrm{rg}(A'_2B'_2 - B'_2A_2) \leqslant 1$, et par hypothèse de récurrence il existe P et Q inversibles telles que $P^{-1}A_1P$ et $P^{-1}B_1P$ d'une part, $Q^{-1}A_1Q$ et $Q^{-1}B_1Q$ d'autre part, soient triangulaires supérieures. Si R est la matrice diagonale par blocs $\mathrm{Diag}(P, Q)$, $R^{-1}A'R$ et $R^{-1}B'R$ sont triangulaires supérieures et donc A et B sont simultanément trigonalisables. ◁

Après tous ces exercices le lecteur se demandera légitimement s'il y a une condition nécessaire et suffisante pour que deux matrices A *et* B *de* $\mathcal{M}_n(\mathbb{C})$ *soient cotrigonalisables. On voit de suite que si* A, B *sont cotrigonalisables, la matrice* AB − BA *se trigonalise en une matrice dont la diagonale est nulle. Elle est donc nilpotente. En fait, il en est de même de toute matrice de la forme* $P(A, B)(AB - BA)$ *où* $P(X, Y)$ *désigne un polynôme en deux variables* X, Y *non commutatives : par exemple si* $P(X, Y) = XY^2X^3$ *on a* $P(A, B) = AB^2A^3$. *Et bien un théorème de Mc Coy affirme que la réciproque est vraie : si* $P(A, B)(AB - BA)$ *est nilpotente pour tout polynôme en deux variables non commutatives* P, *alors* A *et* B *sont cotrigonalisables. On en déduit directement le résultat de l'exercice 2.25 (cas où* $AB - BA = 0$*) mais aussi celui de l'exercice 2.26. En effet, si* $C = AB - BA$ *commute avec* A *et* B *on a* $C^k = (AC^{k-1})B - B(AC^{k-1})$ *pour tout* $k \geqslant 1$ *et on en déduit que* $\mathrm{Tr}(C^k) = 0$. *Cela implique classiquement que* C *est nilpotente (voir exercice 2.33). Pour tout polynôme* P *en deux variables non commutatives on a alors* $[P(A, B)C]^n = P(A, B)^nC^n = 0$ *car* C *commute avec* $P(A, B)$.

Voici un exercice délicat où on l'on cherche un vecteur propre commun à une famille commutative de matrices.

2.29. Vecteur propre commun à une famille de matrices

Soit K un corps commutatif et S une partie de $\mathcal{M}_n(K)$ telle que les éléments de S commutent deux à deux. On suppose de plus que les éléments de S possèdent un vecteur propre commun. Démontrer l'existence d'un vecteur propre commun aux éléments de ${}^tS$.

(ENS Ulm)

▷ **Solution.**

Le résultat a déjà été évoqué lorsque K est $\mathbb{C}$ ou plus généralement un corps algébriquement clos (voir page 93).

Quitte à rajouter les matrices $\lambda \mathrm{I}_n$ pour $\lambda \in \mathrm{K}$, on peut supposer sans perte de généralité qu'elles appartiennent à S. On peut remarquer que les éléments du sous-espace engendré par S possèdent la même propriété que ceux de S : ils commutent deux à deux et possèdent un vecteur propre en commun. On peut donc supposer que S est un sous-espace vectoriel de $\mathcal{M}_n(\mathrm{K})$ et en prendre une base $\mathcal{B}$ dont le premier élément est I_n. Écrivons $\mathcal{B} = (\mathrm{I}_n, \mathrm{A}_1, \ldots, \mathrm{A}_p)$. Nous savons qu'il existe $\mathrm{X}_0 \in \mathrm{K}^n$ non nul tel que pour tout $i \in [\![1, p]\!]$, il existe $\lambda_i \in \mathrm{K}$ vérifiant $\mathrm{A}_i\mathrm{X}_0 = \lambda_i \mathrm{X}_0$. En changeant A_i en $\mathrm{A}_i - \lambda_i \mathrm{I}_n$, la famille $\mathcal{B}$ reste une base de S, et pour tout $i \in [\![1, n]\!]$, $\mathrm{A}_i\mathrm{X}_0 = 0$. Nous allons démontrer l'existence d'un vecteur non nul appartenant aux noyaux de chaque ${}^t\mathrm{A}_i$. Ce vecteur sera alors un vecteur propre commun à tous les éléments de ${}^t\mathrm{S}$.

Pour cela, opérons une récurrence sur l'entier n. Le résultat est trivial pour $n = 1$. Supposons $n \geqslant 2$ et le résultat vrai en taille $< n$.

Le polynôme caractéristique des matrices A_i peut se mettre sous la forme $\mathrm{X}^p\mathrm{Q}$ avec $p \geqslant 1$ et $\mathrm{Q}(0) \neq 0$. Supposons que pour l'une des matrices A_i, disons A_1 pour fixer les idées, on ait $p < n$ et $\deg \mathrm{Q} \geqslant 1$. Le théorème de décomposition des noyaux permet d'écrire $\mathrm{K}^n = \operatorname{Ker} \mathrm{A}_1^p \oplus \operatorname{Ker} \mathrm{Q}(\mathrm{A}_1)$, ces deux sous-espaces n'étant pas réduits à $\{0\}$. Considérons $\mathcal{B}_1$ une base de $\operatorname{Ker} \mathrm{A}_1^p$, $\mathcal{B}_2$ une base de $\operatorname{Ker} \mathrm{Q}(\mathrm{A}_1)$ et P la matrice de passage de la base canonique à la base $(\mathcal{B}_1, \mathcal{B}_2)$. Comme les A_i commutent deux à deux, ils laissent stables les sous-espaces $\operatorname{Ker} \mathrm{A}_1^p$ et $\operatorname{Ker} \mathrm{Q}(\mathrm{A}_1)$. Pour tout $1 \leqslant i \leqslant p$, la matrice $\mathrm{P}^{-1}\mathrm{A}_i\mathrm{P}$ est de la forme $\begin{pmatrix} \mathrm{A}_i' & 0 \\ 0 & \mathrm{B}_i \end{pmatrix}$ avec A_i' (resp. B_i) la matrice de l'endomorphisme induit par A_i sur $\operatorname{Ker} \mathrm{A}_1^p$ (resp. sur $\operatorname{Ker} \mathrm{Q}(\mathrm{A}_1)$). Les matrices A_i' commutent deux à deux et possèdent un vecteur non nul appartenant à tous les noyaux $\operatorname{Ker} \mathrm{A}_i'$. Comme la taille de ces matrices est strictement inférieure à n, par hypothèse de récurrence, il existe Y_0 non nul tel que ${}^t\mathrm{A}_i'\mathrm{Y}_0 = 0$ (la seule valeur propre de A_i' est 0 car A_i' est nilpotente). Dans ces conditions, le vecteur $\mathrm{Y}_0' = \begin{pmatrix} \mathrm{Y}_0 \\ 0 \end{pmatrix} \in \mathrm{K}^n$ est non nul et ${}^t(\mathrm{P}^{-1}\mathrm{A}_i\mathrm{P})\mathrm{Y}_0' = 0$ pour tout i. Comme ${}^t\mathrm{P}\,{}^t\mathrm{A}_i\,{}^t\mathrm{P}^{-1}\mathrm{Y}_0' = 0$, le vecteur non nul ${}^t\mathrm{P}^{-1}\mathrm{Y}_0$ est dans le noyau de chaque ${}^t\mathrm{A}_i$ et répond au problème.

Il nous reste à traiter le cas où pour chaque matrice A_i, le polynôme caractéristique est égal à X^n (ce qui correspond avec les notations utilisées plus haut à $p = n$ et $\mathrm{Q} = 1$). Les matrices A_i sont donc nilpotentes. Ainsi, la famille $({}^t\mathrm{A}_1, \ldots, {}^t\mathrm{A}_p)$ est composée de matrices nilpotentes commutant deux à deux. L'existence d'un vecteur propre commun à toutes les matrices ${}^t\mathrm{A}_i$ est assurée par le lemme suivant :

Lemme. *Soit* E *un* K*-espace vectoriel de dimension finie non nulle et* $(u_1, \ldots, u_p)$ *une famille d'endomorphismes nilpotents commutant deux*

à deux. Il existe alors $x \in \mathrm{E}$, vecteur propre commun à tous les endomorphismes u_i.

Démonstration. Procédons par récurrence sur l'entier p. Si $p = 1$, comme $\mathrm{E} \neq \{0\}$, $\operatorname{Ker} u_1$ n'est pas réduit à $\{0\}$ et un vecteur non nul de $\operatorname{Ker} u_1$ convient.

Supposons $p \geqslant 2$ et le résultat vrai au rang $p-1$. Si $u_p = 0$, on applique l'hypothèse de récurrence à la famille $u_1, \ldots, u_{p-1}$. Le vecteur x trouvé convient aussi pour u_p. Si $u_p \neq 0$, $\mathrm{F} = \operatorname{Ker} u_p$ est un sous-espace non réduit à $\{0\}$ laissé stable par les u_i. Pour $1 \leqslant i \leqslant p-1$, notons u_i' l'endomorphisme induit par u_i sur F. Les endomorphismes $u_1', \ldots, u_{p-1}'$ commutent deux à deux et sont nilpotents, et donc redevables de l'hypothèse de récurrence : il existe $x \in \mathrm{F}$ un vecteur propre commun à tous les u_i'. C'est donc un vecteur propre pour u_i avec $i \leqslant p-1$ et même pour u_p car $x \in \mathrm{F} = \operatorname{Ker} u_p$. La preuve du lemme est achevée. ♢ ◁

L'exercice suivant fait démontrer l'importante décomposition de Dunford.

2.30. Décomposition de Dunford

Soit E un $\mathbb{C}$-espace vectoriel de dimension $n \geqslant 1$ et u un endomorphisme de E. Montrer l'existence d'un unique couple (d, n) d'endomorphismes de E tel que

(*i*) $u = d + n$,

(*ii*) d et n commutent,

(*iii*) d est diagonalisable et n est nilpotent.

Vérifier en outre que d et n sont des polynômes en u.

(**École polytechnique**)

▷ **Solution.**

• Traitons d'abord l'existence du couple (d, n). Écrivons

$$\chi_u = (\mathrm{X} - \lambda_1)^{m_1} \ldots (\mathrm{X} - \lambda_r)^{m_r},$$

avec les $\lambda_i \in \mathbb{C}$ deux à deux distincts et les entiers $m_i \geqslant 1$. D'après le théorème de décomposition des noyaux et le théorème de Cayley-Hamilton, E est somme directe des sous-espaces caractéristiques $\mathrm{F}_i = \operatorname{Ker}(u - \lambda_i \operatorname{Id})^{n_i}$:

$$\mathrm{E} = \bigoplus_{i=1}^{r} \mathrm{F}_i.$$

Soit d l'endomorphisme de E dont la restriction à F_i est exactement $\lambda_i \operatorname{Id}_{\mathrm{F}_i}$: autrement dit, d est diagonalisable et admet chaque sous-espace

F_i comme espace propre pour la valeur propre λ_i. Posons $n = u - d$. Il ne reste plus qu'à vérifier que n est nilpotent et commute avec d. Comme chaque sous-espace caractéristique de E est stable par u et par d (donc aussi par n), il suffit de le vérifier pour les restrictions aux F_i. Notons avec un indice i les restrictions à F_i. On a $u_i = d_i + n_i$ et $d_i = \lambda_i id_{F_i}$. Or, par définition de F_i, $(u_i - \lambda_i \mathrm{Id}_{F_i})^{m_i} = 0$. Donc n_i est nilpotent et commute clairement avec l'homothétie d_i. D'où le résultat. Le couple (d, n) convient donc.

Avant de prouver l'unicité, vérifions que d et n sont des polynômes en u. Cela se voit dans la démonstration du théorème de décomposition des noyaux : pour tout i la projection π_i sur F_i parallèlement à la somme des autres sous-espaces caractéristiques est un polynôme en u. Or, par construction, on a simplement pris

$$d = \sum_{i=1}^{r} \lambda_i \pi_i \in \mathbb{C}[u].$$

Bien entendu $n = u - d$ est alors aussi dans $\mathbb{C}[u]$.

• Supposons l'existence d'un autre couple (d', n') répondant au problème. On a alors $d' - d = n - n'$. Comme d' commute avec n', il commute aussi avec u, donc avec tout polynôme en u. En particulier d' commute avec d. Ainsi d et d' sont codiagonalisables (cf. exercice 2.19) et donc $d' - d$ est diagonalisable. De même n commute avec n'. Il en découle que $n - n'$ est nilpotent. Le seul endomorphisme diagonalisable et nilpotent étant 0 on a $d = d'$ et $n = n'$. ◁

Cette décomposition peut être effectuée sur un corps quelconque K *dès lors que le polynôme caractéristique de u est scindé sur* K *i.e. dès lors que u est trigonalisable. C'est une condition nécessaire puisque si $u = d + n$, d diagonalisable, n nilpotente, $nd = dn$, alors d et n sont co-trigonalisables (cf. exercice 2.25), donc u est également trigonalisable.*

La décomposition de Dunford est très utile dans le calcul des puissances d'une matrice et intervient donc assez naturellement dans l'étude des exponentielles de matrices.

2.31. Image de l'exponentielle

On admet ici que toute matrice A de $\mathcal{M}_n(\mathbb{C})$ se décompose de manière unique sous la forme A = D + N avec D diagonalisable, N nilpotente et DN = ND (voir l'exercice précédent pour la preuve de ce fait). Il s'agit de la décomposition de Dunford.

1. Déterminer la décomposition de Dunford de e^A en fonction de celle de A.

2. En déduire toutes les matrices A de $\mathcal{M}_n(\mathbb{C})$ telles que $e^{\mathrm{A}} = \mathrm{I}_n$.

3. Quelle est l'image de l'application $\mathrm{M} \longmapsto e^{\mathrm{M}}$ de $\mathcal{M}_n(\mathbb{C})$ dans lui-même ?

(École polytechnique)

▷ **Solution.**

1. Posons $\mathrm{A} = \mathrm{D}+\mathrm{N}$ avec D diagonalisable, N nilpotente et $\mathrm{DN} = \mathrm{ND}$. Comme D et N commutent on a $e^{\mathrm{A}} = e^{\mathrm{D}}e^{\mathrm{N}}$. Posons $e^{\mathrm{N}} = \mathrm{I}_n + \mathrm{N}'$ avec

$$\mathrm{N}' = \sum_{k=1}^{n-1} \frac{\mathrm{N}^k}{k!}.$$

La matrice N' est nilpotente, puisque de la forme $\mathrm{NP(N)}$ avec P polynôme. La matrice e^{D} est diagonalisable car si $\mathrm{D} = \mathrm{P}^{-1}\,\mathrm{Diag}(d_1, \ldots, d_n)\mathrm{P}$ alors $e^{\mathrm{D}} = \mathrm{P}^{-1}\,\mathrm{Diag}(e^{d_1}, \ldots, e^{d_n})\mathrm{P}$. De plus, comme e^{D} et e^{N} commutent, N' commute avec e^{D}. L'écriture $e^{\mathrm{A}} = e^{\mathrm{D}} + e^{\mathrm{D}}\mathrm{N}'$ est donc la décomposition de Dunford de e^{A} car $e^{\mathrm{D}}\mathrm{N}'$ est nilpotente et commute avec e^{D}.

2. Soit $\mathrm{A} \in \mathcal{M}_n(\mathbb{C})$ vérifiant $e^{\mathrm{A}} = \mathrm{I}_n$. On reprend les notations de la question précédente. L'unicité de la décomposition de Dunford permet de dire que

$$e^{\mathrm{A}} = \mathrm{I}_n \iff (e^{\mathrm{D}} = \mathrm{I}_n \text{ et } e^{\mathrm{D}}\mathrm{N}' = 0).$$

Comme e^{D} est inversible, on doit donc avoir $\mathrm{N}' = 0$. Or, $\mathrm{N}' = \mathrm{NP(N)}$ avec $\mathrm{P(X)} = \sum_{k=1}^{n-1} \frac{\mathrm{X}^{k-1}}{k!}$. Le coefficient constant de P est égal à 1. Il en découle que la matrice $\mathrm{P(N)}$ est inversible : en effet, dans une base trigonalisant N on obtient une matrice unipotente (triangulaire supérieure avec des 1 sur la diagonale). Ainsi $\mathrm{N}' = 0$ et la matrice A est diagonalisable. En se plaçant dans une base de diagonalisation on voit que les valeurs propres de A doivent être dans $2i\pi\mathbb{Z}$. Réciproquement, tout matrice diagonalisable dont le spectre est contenu dans $2i\pi\mathbb{Z}$ a comme image I_n par l'exponentielle.

3. Nous savons que si $\mathrm{M} \in \mathcal{M}_n(\mathbb{C})$, alors $\exp(\mathrm{M})$ est inversible puisque, M et $-\mathrm{M}$ commutant,

$$\exp(\mathrm{M})\exp(-\mathrm{M}) = \exp(0) = \mathrm{I}_n.$$

On sait également que l'image de l'exponentielle sur $\mathbb{C}$ est $\mathbb{C}^*$: il est donc raisonnable de penser que l'image de $\mathcal{M}_n(\mathbb{C})$ par l'exponentielle est $\mathrm{GL}_n(\mathbb{C})$ tout entier. C'est ce que nous allons prouver. Soit B une matrice inversible dont on note $\mathrm{D}'+\mathrm{N}'$ la décomposition de Dunford. On cherche

A telle que $e^{\mathrm{A}} = \mathrm{B}$ par sa décomposition de Dunford $\mathrm{A} = \mathrm{D}+\mathrm{N}$. D'après la question 1. cela équivaut donc à $e^{\mathrm{D}} = \mathrm{D}'$ et $e^{\mathrm{D}}(e^{\mathrm{N}} - \mathrm{I}_n) = \mathrm{N}'$.

• On s'occupe d'abord de la première équation. Il est clair que D' est aussi inversible : cela se voit dans la preuve de l'exercice 2.30 mais peut aussi se retrouver directement ici puisque $\mathrm{D}' = \mathrm{B} - \mathrm{N}' = \mathrm{B}(\mathrm{I}_n - \mathrm{B}^{-1}\mathrm{N}')$. La matrice $\mathrm{B}^{-1}\mathrm{N}'$ est nilpotente, car N' et B^{-1} commutent, donc la matrice $\mathrm{I}_n - \mathrm{B}^{-1}\mathrm{N}'$ est inversible (elle se trigonalise en une matrice avec des 1 sur la diagonale).

Soit $\mathrm{P} \in \mathrm{GL}_n(\mathbb{C})$ telle que $\mathrm{D}' = \mathrm{P}^{-1}\,\mathrm{Diag}(d'_1, \ldots, d'_n)\mathrm{P}$. Les nombres complexes d'_k sont tous non nuls et peuvent donc s'écrire $d'_k = e^{d_k}$ où $d_k \in \mathbb{C}$. Si on pose $\mathrm{D} = \mathrm{P}^{-1}\,\mathrm{Diag}(d_1, \ldots, d_n)\mathrm{P}$ on a bien $e^{\mathrm{D}} = \mathrm{D}'$. Les nombres complexes d_k ne sont pas uniques : ils ne sont définis qu'à un élément de $2i\pi\mathbb{Z}$ près. On vient en fait d'expliciter toutes les solutions : en effet, comme D est diagonalisable et commute avec $e^{\mathrm{D}} = \mathrm{D}'$, les matrices D et D' se diagonalisent forcément dans une même base.

• Passons maintenant à la seconde équation qui s'écrit aussi $e^{\mathrm{N}} - \mathrm{I}_n = \mathrm{D}'^{-1}\mathrm{N}'$. Comme D'^{-1} et N' commutent, la matrice $\mathrm{N}'' = \mathrm{D}'^{-1}\mathrm{N}'$ est aussi nilpotente. La question posée se reformule donc comme suit : si $\mathcal{N}$ désigne l'ensemble des matrices nilpotentes de $\mathcal{M}_n(\mathbb{C})$, l'application ψ qui à N associe $\mathrm{N}'' = e^{\mathrm{N}} - \mathrm{I}_n = \sum_{k=1}^{n-1} \frac{\mathrm{N}^k}{k!}$ est-elle une surjection de $\mathcal{N}$ sur $\mathcal{N}$?

En fait, c'est même une bijection ! Pour écrire la bijection réciproque on pense bien entendu au logarithme et cela ne pose aucun problème particulier car on travaille avec des matrices nilpotentes. Considérons donc

$$\varphi : \mathrm{N} \mapsto \ln(\mathrm{I}_n + \mathrm{N}) = \sum_{k=1}^{n-1} (-1)^{k-1} \frac{\mathrm{N}^k}{k}.$$

Notons tout d'abord que ψ et φ sont bien des applications de $\mathcal{N}$ dans $\mathcal{N}$ car $\psi(\mathrm{N}) = \mathrm{NR}(\mathrm{N})$ où R est un polynôme et il en est de même pour φ. Il ne reste plus qu'à justifier que $\psi \circ \varphi = \varphi \circ \psi = \mathrm{Id}$. Posons $\mathrm{P}(\mathrm{X}) = \sum_{k=1}^{n-1} \frac{\mathrm{X}^k}{k!}$ et $\mathrm{Q}(\mathrm{X}) = \sum_{k=1}^{n-1} (-1)^{k-1} \frac{\mathrm{X}^k}{k}$. L'application ψ associe à N la matrice $\mathrm{P}(\mathrm{N})$ et l'application φ associe à N la matrice $\mathrm{Q}(\mathrm{N})$. Ainsi, $\psi \circ \varphi$ associe à N la matrice $\mathrm{P}(\mathrm{Q}(\mathrm{N})) = (\mathrm{P} \circ \mathrm{Q})(\mathrm{N})$. Pour la calculer on a seulement besoin du polynôme $\mathrm{P} \circ \mathrm{Q}$ tronqué à la puissance $n-1$. Or, comme $e^x - 1 = \mathrm{P}(x) + o(x^{n-1})$ et $\ln(1+x) = \mathrm{Q}(x) + o(x^{n-1})$ celui-ci est exactement la partie polynomiale du développement limité à l'ordre $n-1$ en 0 de la fonction $e^{\ln(1+x)} - 1 = x$, c'est-à-dire X ! On a donc $(\mathrm{P} \circ \mathrm{Q})(\mathrm{N}) = \mathrm{N}$ pour toute matrice nilpotente N. C'est pareil dans l'autre sens puisque $\ln(1 + e^x - 1) = x$.

On peut donc trouver une unique matrice nilpotente N telle que $e^{\mathrm{N}} - \mathrm{I}_n = \mathrm{N}''$.

• Ce n'est pas tout à fait fini : on vient de trouver une infinité de matrices diagonalisables D telles que $e^{\mathrm{D}} = \mathrm{D}'$ et une unique matrice nilpotente N telle que $e^{\mathrm{N}} - \mathrm{I}_n = \mathrm{N}'' = \mathrm{D}'^{-1}\mathrm{N}'$. La matrice $\mathrm{A} = \mathrm{D} + \mathrm{N}$ sera un antécédent de B si D et N *commutent* (n'oublions pas que $\mathrm{D}+\mathrm{N}$ est supposé être la décomposition de Dunford de A) ! Or ce n'est pas le cas *a priori*. Ce qui précède montre que N est un polynôme en $\mathrm{D}'^{-1}\mathrm{N}'$. En fait on peut très bien choisir pour D un polynôme en D' : il suffit, avec les notations ci-dessus, de prendre $d_i = d_j$ chaque fois que $d'_i = d'_j$. On a alors $\mathrm{D} = \mathrm{Q}(\mathrm{D}')$, où Q est un polynôme d'interpolation de Lagrange qui envoie les d'_i sur les d_i. Si on fait un tel choix de D, alors D et N commutent et on a $e^{\mathrm{D}+\mathrm{N}} = e^{\mathrm{D}}e^{\mathrm{N}} = \mathrm{D}'(\mathrm{I}_n + \mathrm{D}'^{-1}\mathrm{N}') = \mathrm{D}' + \mathrm{N}' = \mathrm{B}$.

Conclusion. L'image de $\mathcal{M}_n(\mathbb{C})$ par l'exponentielle est $\mathrm{GL}_n(\mathbb{C})$. ◁

On notera qu'il y a toujours une infinité d'antécédents à un élément $\mathrm{B} \in \mathrm{GL}_n(\mathbb{C})$ *donné. Lorsqu'on se place sur* $\mathbb{R}$ *la situation se complique et l'image n'est pas égale à* $\mathrm{GL}_n(\mathbb{R})$*, ce qui se voit déjà pour* $n = 1$*. On peut en fait démontrer que l'image est l'ensemble des carrés de* $\mathrm{GL}_n(\mathbb{R})$*. On peut aussi avoir un antécédent unique comme c'est le cas dans l'exercice suivant, dont nous donnons une solution autonome même s'il découle directement de ce qui précède.*

2.32. Exponentielle de matrices réelles diagonalisables

Soient A, B deux matrices diagonalisables de $\mathcal{M}_n(\mathbb{R})$. On suppose que $e^{\mathrm{A}} = e^{\mathrm{B}}$. Montrer que A = B.

(École polytechnique)

▷ **Solution.**

Si A et B sont toutes les deux diagonales le résultat est évident car la fonction exponentielle est injective sur $\mathbb{R}$. Il nous suffit dons de montrer que A et B sont codiagonalisables pour conclure. On sait pour cela qu'il suffit de prouver que A et B commutent (cf. exercice 2.19). Il est clair que A commute avec $e^{\mathrm{A}} = e^{\mathrm{B}}$. Or, B est un polynôme en e^{B}. Pour le voir il suffit de se placer dans une base diagonalisant B : si $\mathrm{B} = \mathrm{P}^{-1}\mathrm{D}\mathrm{P}$ avec P inversible et $\mathrm{D} = \mathrm{Diag}(d_1, \ldots, d_n)$, alors $e^{\mathrm{B}} = \mathrm{P}^{-1}\mathrm{D}'\mathrm{P}$ avec $\mathrm{D}' = \mathrm{Diag}(e^{d_1}, \ldots, e^{d_n})$. Or il est aisé de trouver un polynôme Q tel que $\mathrm{Q}(e^{d_k}) = d_k$ pour tout k (en utilisant un polynôme d'interpolation de Lagrange). On a alors $\mathrm{Q}(\mathrm{D}') = \mathrm{D}$ et donc $\mathrm{Q}(e^{\mathrm{B}}) = \mathrm{B}$. Ainsi A commute avec B. ◁

À partir de la décomposition de Dunford, on est fondamentalement ramené à l'étude des classes de similitude des matrices nilpotentes. Les exercices qui suivent sont consacrés aux éléments nilpotents.

2.33. Caractérisation des matrices nilpotentes avec la trace

Soit K un sous-corps de $\mathbb{C}$ et $A \in \mathcal{M}_n(K)$. On suppose que pour tout $k \geqslant 1$, la trace de A^k est nulle. Montrer que A est nilpotente.
(École polytechnique)

▷ **Solution.**

Nous allons proposer plusieurs solutions de cette question classique :

• Le polynôme caractéristique de A est scindé sur $\mathbb{C}$. Raisonnons par l'absurde et supposons A non nilpotente. Alors A possède des valeurs propres (complexes) non nulles. On va noter $\lambda_1, \ldots, \lambda_r$ ces valeurs propres non nulles de A ($r \geqslant 1$) et $n_1, \ldots, n_r$ leurs multiplicités respectives. Pour tout $k \geqslant 1$, on a

$$\mathrm{Tr}(A^k) = n_1\lambda_1^k + \cdots + n_r\lambda_r^k = 0.$$

Si on écrit ces relations pour k variant de 1 à r, on obtient que $(n_1, \ldots, n_r)$ est solution du système linéaire

$$\begin{pmatrix} \lambda_1 & \lambda_2 & \cdots & \lambda_r \\ \lambda_1^2 & \lambda_2^2 & \cdots & \lambda_r^2 \\ \vdots & & & \vdots \\ \lambda_1^r & \lambda_2^r & \cdots & \lambda_r^r \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_r \end{pmatrix} = 0.$$

Or, ce système est de Cramer puisque le déterminant de la matrice du système vaut

$$\lambda_1 \ldots \lambda_r \prod_{1 \leqslant i < j \leqslant r} (\lambda_j - \lambda_i) \neq 0 \quad \text{(cf. exercice 1.8)}$$

Nécessairement, on a $n_1 = \cdots = n_r = 0$, ce qui est exclu.

• Deuxième solution à l'aide des formules de Newton : on considère désormais $\lambda_1, \ldots, \lambda_n$, les racines complexes de χ_A comptées avec multiplicités. Dire que $\mathrm{Tr}(A^k) = 0$ pour tout $k \geqslant 1$, c'est exactement dire que les sommes de Newton relatives à $\lambda_1, \ldots, \lambda_n$ sont nulles :

$$\mathrm{Tr}(A^k) = \lambda_1^k + \cdots + \lambda_n^k = 0,$$

car A étant semblable à une matrice triangulaire supérieure à coefficients diagonaux $\lambda_1, \ldots, \lambda_n$, A^k est semblable à une matrice triangulaire supérieure à coefficients diagonaux $\lambda_1^k, \ldots, \lambda_n^k$.

Les formules de Newton (voir exercice 5.26, tome 1 d'algèbre) donnent alors que les fonctions symétriques élémentaires des λ_i sont nulles : $\sigma_1 = \cdots = \sigma_n = 0$. On en déduit que

$$\chi_A = X^n - \sigma_1 X^{n-1} + \cdots + (-1)^{n-1} X + (-1)^n \sigma_n = X^n,$$

et $A^n = 0$ d'après le théorème de Cayley-Hamilton.

• Voici enfin une troisième solution par récurrence sur la taille de la matrice. Si $n = 1$, comme $\mathrm{Tr}(A) = 0$, on a $A = 0$. Soit $n \geqslant 2$. Supposons le résultat vrai jusqu'au rang $n-1$. Écrivons le polynôme caractéristique de A :

$$\chi_A = X^n - \sigma_1 X^{n-1} + \cdots + (-1)^{n-1} X + (-1)^n \sigma_n.$$

D'après le théorème de Cayley-Hamilton, on a

$$\chi_A(A) = 0 = A^n - \sigma_1 A^{n-1} + \cdots + (-1)^{n-1} A + (-1)^n \sigma_n I_n,$$

et en exprimant la trace de cette somme, par hypothèse, il vient

$$(-1)^n \sigma_n n = 0,$$

ou encore $n \det A = 0$, soit $\det A = 0$. Donc 0 est valeur propre de A. Écrivons $\chi_A = X^p Q$ avec $Q(0) \neq 0$, $p \geqslant 1$.

D'après le théorème de décomposition des noyaux, on a

$$K^n = \mathrm{Ker}(\chi_A(A)) = \mathrm{Ker}\, A^p \oplus \mathrm{Ker}\, Q(A).$$

Supposons $\mathrm{Ker}\, Q(A) \neq \{0\}$. Prenons une base obtenue par réunion d'une base de $\mathrm{Ker}\, A^p$ et d'une base de $\mathrm{Ker}\, Q(A)$. L'endomorphisme A dans cette base admet une matrice de la forme

$$\begin{pmatrix} \boxed{A'} & 0 \\ 0 & \boxed{B'} \end{pmatrix}.$$

On a $A'^p = 0$, donc A' est nilpotente et, pour tout $k \geqslant 1$, $\mathrm{Tr}(A'^k) = 0$. Comme A^k est semblable à

$$\begin{pmatrix} \boxed{A'^k} & 0 \\ 0 & \boxed{B'^k} \end{pmatrix},$$

on a $\mathrm{Tr}(B'^k) = 0$. D'autre part, la restriction de A à $\mathrm{Ker}\, Q(A)$ est injective, si bien que B′ est inversible. Or par hypothèse de récurrence, B′ est aussi nilpotente, ce qui fournit la contradiction.

On peut donc affirmer que $\mathrm{Ker}\, Q(A) = \{0\}$ et $K^n = \mathrm{Ker}\, A^p$: la matrice A est bien nilpotente. ◁

La troisième solution est valable sur un corps commutatif de caractéristique nulle quelconque. Les autres aussi à condition d'admettre l'existence d'un surcorps dans lequel le polynôme caractéristique est scindé.

Les énoncés qui suivent concernent des questions de sous-espaces stables.

2.34. Sous-espaces stables par un endomorphisme nilpotent

Soit E un K-espace vectoriel de dimension n et u un endomorphisme nilpotent de E de rang $n-1$. Montrer que E admet exactement $n+1$ sous-espaces vectoriels stables par u et que ce sont les $\mathrm{Ker}\, u^k$ pour $k \in [\![0, n]\!]$.

(École polytechnique)

▷ **Solution.**

Bien entendu les sous-espaces $\mathrm{Ker}\, u^k$ pour $k \in [\![0, n]\!]$ sont tous stables par u. On va d'abord montrer que $\dim \mathrm{Ker}\, u^k = k$ pour tout $k \in [\![0, n]\!]$ ce qui prouvera que ces $n+1$ sous-espaces sont deux à deux distincts. Le théorème du rang appliqué à la restriction de u à $\mathrm{Im}\, u^k$ montre que

$$\mathrm{rg}\, u^k = \mathrm{rg}\, u^{k+1} + \dim(\mathrm{Ker}\, u \cap \mathrm{Im}\, u^k) \leqslant \mathrm{rg}\, u^{k+1} + 1.$$

Autrement dit $\dim \mathrm{Ker}\, u^{k+1} \leqslant \dim \mathrm{Ker}\, u^k + 1$: la dimension des noyaux itérés augmente d'une unité au plus à chaque fois. Mais comme $\mathrm{Ker}\, u^0 = \{0\}$ et $\mathrm{Ker}\, u^n = E$ on a nécessairement $\dim \mathrm{Ker}\, u^k = k$ pour tout $k \in [\![0, n]\!]$.

Considérons maintenant un sous-espace quelconque F de E stable par u. Notons p sa dimension. La restriction de u à F est encore nilpotente avec un indice forcément $\leqslant p$. On a donc $u|_F^p = 0$ ce qui veut dire que $F \subset \mathrm{Ker}\, u^p$. Comme les deux espaces ont la même dimension ils sont égaux. ◁

On vient de retrouver dans cet exercice un cas particulier d'un résultat bien connu sur la suite des noyaux itérés : elle est croissante, mais croit de moins en moins vite (cf. tome 1, exercice 6.14).

2.35. Endomorphisme nilpotent semi-simple

Soit E un espace vectoriel de dimension finie. Trouver les endomorphismes nilpotents de E tels que tout sous-espace stable admette un supplémentaire stable.

(École polytechnique)

▷ **Solution.**

Soit u un endomorphisme nilpotent de E tel que tout sous-espace stable admette un supplémentaire stable. Le noyau de u admet donc un supplémentaire stable F. Comme $\chi_u = X^n$ (n étant la dimension de E), le polynôme caractéristique de $u_{|F}$ est X^r avec $r = \dim F$. Or, comme $F \cap \operatorname{Ker} u = \{0\}$, $u_{|F}$ est injective *i.e.* 0 n'est pas valeur propre de $u_{|F}$. Mais alors, r est nul et $\operatorname{Ker} u = E$. L'endomorphisme u est donc nul. Réciproquement, l'endomorphisme nul convient. ◁

Il s'agissait de trouver les endomorphismes u nilpotents semi-simples (voir l'exercice 2.37). Comme χ_u est scindé, cela revient à trouver, d'après l'exercice mentionné, les endomorphismes à la fois nilpotents et diagonalisables : seul 0 convient.

L'exercice suivant est plus difficile. Il précise quels sont les sous-espaces stables par un endomorphisme nilpotent qui ont un supplémentaire stable.

2.36. Existence d'un supplémentaire stable par un endomorphisme nilpotent

Soit E un K-espace vectoriel de dimension finie, $f \in \mathcal{L}(E)$ nilpotent et V un sous-espace stable par f. Montrer que V admet un supplémentaire stable si et seulement si pour tout $k \in \mathbb{N}$, $f^k(V) = f^k(E) \cap V$.

(École polytechnique)

▷ **Solution.**

• La nécessité de la condition est facile et n'utilise pas le caractère nilpotent de f. Supposons que V admette un supplémentaire stable W. Soit k un entier naturel. On a toujours $f^k(V) \subset f^k(E)$ et comme V est stable on a aussi $f^k(V) \subset V$. On a donc $f^k(V) \subset f^k(E) \cap V$.

Soit $x \in f^k(E) \cap V$. Il existe un vecteur $y \in E$ tel que $x = f^k(y)$. Ce vecteur s'écrit $y = y_V + y_W$ avec $y_V \in V$ et $y_W \in W$. Ainsi, par stabilité

de V et W, il vient

$$x = f^k(y_V) + f^k(y_W) \quad \text{avec } f^k(y_V) \in V \text{ et } f^k(y_W) \in W.$$

Or comme V et W sont en somme directe et $x \in V$, nécessairement $f^k(y_W) = 0$. Il en résulte que $x = f^k(y_V) \in f^k(V)$. On conclut

$$\boxed{f^k(V) = f^k(E) \cap V}.$$

• Réciproquement, supposons que l'on ait $f^k(V) = f^k(E) \cap V$, pour tout $k \in \mathbb{N}$. Nous allons procéder par récurrence sur l'indice de nilpotence p de f (*i.e.* le plus petit entier $p \geqslant 1$ tel que $f^p = 0$).

Si $p = 1$, $f = 0$ et n'importe quel supplémentaire W de V est stable par f.

Supposons $p \geqslant 2$. Alors f induit un endomorphisme $\tilde{f}$ nilpotent de $f(E)$. De plus, l'indice de nilpotence de l'endomorphisme $\tilde{f}$ est $p - 1$ puisque $\tilde{f}^k(f(E)) = f^{k+1}(E)$ pour tout $k \geqslant 0$. Notons $E' = f(E)$ et $V' = f(V)$, qui est un sous-espace de E'. Nous allons appliquer l'hypothèse de récurrence au sous-espace V' de E'. Pour cela, vérifions que pour $k \geqslant 0$, $\tilde{f}^k(V') = \tilde{f}^k(E') \cap V'$, autrement dit que $f^{k+1}(V) = f^{k+1}(E) \cap f(V)$. L'inclusion $f^{k+1}(V) \subset f^{k+1}(E) \cap f(V)$ est immédiate. Si $x \in f^{k+1}(E) \cap f(V)$, alors $x \in f^{k+1}(E) \cap V = f^{k+1}(V)$ par hypothèse. L'égalité des sous-espaces est donc vérifiée.

Par hypothèse de récurrence, nous pouvons donc affirmer l'existence d'un sous-espace W de $E' = f(E)$, stable par f tel que

$$f(E) = f(V) \oplus W = (f(E) \cap V) \oplus W.$$

Les sous-espaces W et V sont en somme directe puisque si $x \in V \cap W$, $x \in f(E) \cap V$ et $x = 0$ car $(f(E) \cap V) \cap W = \{0\}$. Construisons à partir de W un sous-espace W' stable par f, supplémentaire de V. Il est naturel de penser à $W_1 = f^{-1}(W)$ qui a le mérite d'être stable par f. Malheureusement, il n'a aucune raison d'être un supplémentaire de V (il contient $\operatorname{Ker} f$ et $\operatorname{Ker} f \cap V \neq \{0\}$ si V n'est pas réduit à $\{0\}$ puisque $f_{|V}$ est nilpotent). Nous allons corriger cela : le sous-espace W est clairement contenu dans W_1. Les sous-espace V et W étant en somme directe, la somme $(V \cap W_1) + W$ est directe et contenue dans W_1. Prenons W_2 un supplémentaire de $(V \cap W_1) \oplus W$ dans W_1 : on a

$$W_1 = f^{-1}(W) = (V \cap W_1) \oplus W \oplus W_2 = (V \cap W_1) \oplus W',$$

en ayant posé $W' = W \oplus W_2$. Le sous-espace W' est stable par f puisque $f(W') \subset W \subset W'$.

Montrons que W′ est bien un supplémentaire de V. Soit $x \in V \cap W'$. Alors $x \in W_1 \cap V$ et comme $W_1 \cap V$ est en somme directe avec W′, $x = 0$: la somme $V \oplus W'$ est directe.

Soit $z \in E$. Alors $f(z) \in f(E) = f(V) \oplus W$ et il existe donc $v \in V$ et $w \in W$ tel que $f(z) = f(v) + w$. Dans ces conditions, $f(z - v) = w \in W$ et $z - v \in W_1$. Le vecteur $z - v$ s'écrit $v' + w'$ avec $v' \in V \cap W_1 \subset V$ et $w' \in W'$. Ainsi, on a $z = (v + v') + w' \in V + W'$. On conclut que $E = V \oplus W'$ et W′ constitue bien un supplémentaire de V stable par f. ◁

Le lecteur pourra ainsi vérifier que dans la situation de l'exercice 2.34 les seuls sous-espaces stables par u qui ont un supplémentaire stable sont $\{0\}$ *et* E.

On dit qu'un endomorphisme u de E *est semi-simple si tout sous-espace de* E *stable par u admet un supplémentaire stable. L'exercice suivant montre qu'un endomorphisme est diagonalisable s'il est semi-simple et si son polynôme caractéristique est scindé. Il en résulte qu'un endomorphisme d'un* $\mathbb{C}$*-espace vectoriel est semi-simple si et seulement si il est diagonalisable.*

2.37. Endomorphismes semi-simples

Soit E un K-espace vectoriel, $u \in \mathcal{L}(E)$. Montrer l'équivalence des deux conditions suivantes :

(*i*) u est diagonalisable ;

(*ii*) χ_u est scindé et tout sous-espace de E stable par u admet un supplémentaire stable par u.

(ENS Lyon)

▷ **Solution.**

• Supposons (*i*). Alors χ_u étant le polynôme caractéristique d'une matrice diagonale, il est scindé. Soit F un sous-espace stable pour u. On peut supposer F non nul et distinct de E. Posons $n = \dim E$ et $p = \dim F$. Choisissons une base $B_F = (f_1, \ldots, f_p)$ de F et $B = (e_1, \ldots, e_n)$ une base de vecteurs propres de E. Le théorème de la base incomplète permet de compléter la base B_F en une base de E en rajoutant des vecteurs de B disons $e_{i_1}, \ldots, e_{i_{n-p}}$. Il est clair que le sous-espace engendré par les vecteurs que l'on rajoute est un supplémentaire de F qui est stable par u.

• Supposons (ii). Considérons $\mathrm{F} = \bigoplus_{\lambda \in \mathrm{Sp}\, u} \mathrm{Ker}(u - \lambda \mathrm{Id}_\mathrm{E})$, la somme des espaces propres de u. C'est un sous-espace stable par u donc par hypothèse il admet un supplémentaire stable G. Supposons $\dim \mathrm{G} \geqslant 1$ et notons u_G l'endomorphisme induit par u sur G. Comme χ_u est scindé, il en va de même pour χ_{u_G}, qui admet donc une racine λ puisque son degré n'est pas nul. Il existe donc x non nul dans G tel que $u(x) = u_\mathrm{G}(x) = \lambda x$. Mais alors x est dans $\mathrm{F} \cap \mathrm{G}$, ce qui entraîne que $\mathrm{F} \cap \mathrm{G}$ n'est pas réduit à $\{0\}$. On aboutit à une contradiction.

On en déduit que la dimension de G ne peut valoir que 0 et que E est donc somme de ses sous-espaces propres pour u : u est diagonalisable. ◁

On peut démontrer que, dans le cas d'un corps quelconque, l'endomorphisme u est semi-simple si et seulement si son polynôme minimal est produit de polynômes irréductibles unitaires distincts. On en déduit qu'une matrice de $\mathcal{M}_n(\mathbb{R})$ est semi-simple si et seulement si elle est diagonalisable dans $\mathcal{M}_n(\mathbb{C})$.

Une manière très simple de construire des sous-espaces stables par un endomorphisme u d'un espace E *est de prendre un vecteur x quelconque de* E *et de considérer le sous-espace vectoriel E_x engendré par la famille $(u^k(x))_{k\in\mathbb{N}}$. Il s'agit bien entendu du plus petit sous-espace u-stable qui contient x. Le cas où l'on peut trouver un vecteur x tel que $\mathrm{E}_x = \mathrm{E}$ conduit à la notion importante d'endomorphisme cyclique. L'énoncé qui suit en donne une caractérisation simple à l'aide du polynôme minimal.*

2.38. Endomorphismes cycliques

Soit E un K-espace vectoriel de dimension finie, $u \in \mathcal{L}(\mathrm{E})$ et $x \in \mathrm{E}$. On définit

$$\mathrm{I} = \{\mathrm{P} \in \mathrm{K}[\mathrm{X}],\ \mathrm{P}(u) = 0\} \quad \text{et} \quad \mathrm{I}_x = \{\mathrm{P} \in \mathrm{K}[\mathrm{X}],\ \mathrm{P}(u)(x) = 0\}.$$

1. Montrer l'existence de polynômes unitaires non nuls μ et μ_x tels que $\mathrm{I} = \mu\mathrm{K}[\mathrm{X}]$ et $\mathrm{I}_x = \mu_x\mathrm{K}[\mathrm{X}]$. Montrer que μ_x divise μ.

2. Montrer qu'il existe $x \in \mathrm{E}$ tel que $\mu_x = \mu$.

3. On dit que u est *cyclique* s'il existe $x \in \mathrm{E}$ tel que $\mathrm{E} = \mathrm{Vect}(u^k(x))_{k\in\mathbb{N}}$. Montrer l'équivalence

$$u \text{ est cyclique} \iff \mu = \chi_u.$$

(ENS Ulm)

▷ **Solution.**

1. L'ensemble I est le noyau du morphisme d'algèbres

$$\varphi : \mathrm{P} \in \mathrm{K}[\mathrm{X}] \longmapsto \mathrm{P}(u) \in \mathcal{L}(\mathrm{E}).$$

C'est donc un idéal de E. L'algèbre $\mathcal{L}(\mathrm{E})$ étant de dimension finie, φ ne peut être injectif et $\mathrm{I} \neq \{0\}$. Nous savons alors qu'il existe un unique polynôme unitaire non nul μ, appelé polynôme minimal de u, tel que $\mathrm{I} = \mu\mathrm{K}[\mathrm{X}]$.

Montrons directement que I_x est un idéal. Soit P et Q dans I_x et A dans K[X]. Le polynôme nul est dans I_x, et

$$(\mathrm{P}+\mathrm{Q})(u)(x) = (\mathrm{P}(u)+\mathrm{Q}(u))(x) = \mathrm{P}(u)(x)+\mathrm{Q}(u)(x) = 0,$$
$$(\mathrm{AP})(u)(x) = \mathrm{A}(u)\circ\mathrm{P}(u)(x) = \mathrm{A}(u)(\mathrm{P}(u)(x)) = \mathrm{A}(u)(0) = 0,$$

et on a bien $\mathrm{P}+\mathrm{Q} \in \mathrm{I}_x$ et $\mathrm{AP} \in \mathrm{I}_x$. L'ensemble I_x est donc un idéal de K[X], non réduit à $\{0\}$ puisque $\mathrm{I} \subset \mathrm{I}_x$: il existe donc μ_x, polynôme unitaire non nul, tel que $\mathrm{I}_x = \mu_x\mathrm{K}[\mathrm{X}]$. Comme $\mu \in \mathrm{I}_x$, μ_x divise μ.

Le polynôme unitaire μ_x est appelé polynôme minimal ponctuel de u en x.

2. Supposons $n = \dim \mathrm{E} \geqslant 1$ (si $\mathrm{E} = \{0\}$, $x = 0$ convient : $\mu = \mu_0 = 1$). Dans ces conditions on a $\deg \mu \geqslant 1$. On va utiliser la décomposition de μ en facteurs irréductibles.

• Si μ est irréductible, alors pour $x \neq 0$, μ_x est un diviseur de μ distinct de 1, c'est nécessairement μ.

• Supposons que μ soit de la forme P^r avec P irréductible. Considérons une base $(e_1, \ldots, e_n)$ de E. Pour $1 \leqslant i \leqslant n$, μ_{e_i} est de la forme P^{r_i} avec $1 \leqslant r_i \leqslant n$: si on note $s = \max_{1\leqslant i\leqslant n} r_i$, alors $s \leqslant r$ et $\mathrm{P}^s(u)$ est nul en chaque e_i et donc $\mathrm{P}^s(u) = 0$. Autrement dit $\mu | \mathrm{P}^s$. Cela entraîne $r \leqslant s$, et finalement $r = s$: autrement dit, il existe $i \in [\![1, n]\!]$ tel que $\mu_{e_i} = \mathrm{P}^r = \mu$.

• Supposons μ quelconque : on peut écrire

$$\mu = \mathrm{P}_1^{r_1} \ldots \mathrm{P}_s^{r_s},$$

avec $s \geqslant 1$, les P_i irréductibles deux à deux distincts, et les r_i entiers naturels non nuls. Posons pour $1 \leqslant i \leqslant s$, $\mathrm{Q}_i = \mathrm{P}_i^{r_i}$ et $\mathrm{E}_i = \operatorname{Ker} \mathrm{Q}_i(u)$. Comme $\mu(u) = 0$, le théorème de décomposition des noyaux nous assure que

$$\mathrm{E} = \operatorname{Ker} \mathrm{Q}_1(u) \oplus \cdots \oplus \operatorname{Ker} \mathrm{Q}_s(u) = \mathrm{E}_1 \oplus \cdots \oplus \mathrm{E}_s.$$

Chaque E_i est stable par u ; nous noterons u_i l'endomorphisme induit sur E_i par u. Comme pour tout $x \in \mathrm{E}_i$, $\mathrm{Q}_i(u)(x) = 0$, on a $\mathrm{Q}_i(u_i) = 0$. Ainsi, le polynôme minimal de u_i, que nous noterons μ_i est un diviseur

de $Q_i = P_i^{r_i}$. Il est en fait égal à Q_i. En effet, si on avait $\mu_i = P_i^{r_i'}$ avec $r_i' < r_i$, alors le polynôme

$$P = P_1^{r_1} \cdots P_{i-1}^{r_{i-1}} P_i^{r_i'} P_{i+1}^{r_{i+1}} \cdots P_s^{r_s}$$

annulerait u (en effet $P(u)$ s'annule sur chaque E_j pour $j \neq i$, puisque si $x \in E_j$, $P_j^{r_j}(u)(x) = 0$ et s'annule aussi sur E_i. puisque si $x \in E_i$. $P_i^{r_i'}(u)(x) = 0$; comme E est somme directe des sous-espaces E_i, $P(u)$ est l'endomorphisme nul). Cela contredirait la minimalité de μ.

D'après ce qui précède, il existe $x_i \in E_i$ tel que le polynôme minimal ponctuel de u_i en x_i soit égal au polynôme minimal de u_i *i.e.* Q_i. Or le polynôme minimal ponctuel de u_i en x_i est aussi le polynôme minimal ponctuel de u en x_i puisque si $P \in K[X]$, $P(u)(x) = P(u_i)(x)$. On a donc $\mu_{x_i} = Q_i$.

Considérons $x = x_1 + \cdots + x_s$. On va montrer que $\mu_x = \mu$. Soit $P \in K[X]$ tel que $P(u)(x) = 0$. On a alors

$$0 = P(u)(x) = P(u)(x_1) + P(u)(x_2) + \cdots + P(u)(x_s).$$

Chaque E_i étant stable par u, on a $P(u)(x_i) \in E_i$ pour tout i et, comme la somme de ces sous-espaces est directe, $P(u)(x_i) = 0$ pour tout $1 \leqslant i \leqslant s$. On en déduit que $Q_i = \mu_{x_i}$ divise P et comme ces polynômes sont premiers entre eux deux à deux, $\mu = Q_1 \ldots Q_s$ divise P. En particulier, μ divise donc μ_x, et finalement $\mu_x = \mu$.

3. Pour x vecteur de E, on pose $E_x = \mathrm{Vect}(u^k(x))_{k\in\mathbb{N}}$. C'est le plus petit sous-espace de E stable par u, contenant x. Montrons pour commencer que la dimension de E_x n'est rien d'autre que le degré r du polynôme minimal ponctuel μ_x. En effet, par définition la famille $\mathrm{Vect}(x, u(x), \ldots, u^{r-1}(x))$ est libre. Notons W le sous-espace de dimension r qu'elle engendre. C'est un sous-espace de E_x. On va montrer en fait que $W = E_x$. Comme $\mu_x(u)(x) = 0$ la famille $(x, u(x), \ldots, u^r(x))$ est liée. Donc $u^r(x) \in \mathrm{Vect}(x, u(x), \ldots, u^{r-1}(x)) = W$. Cela prouve que W est stable par u (l'image d'un vecteur quelconque de la base de W est encore dans W). Comme W contient x on a donc $E_x \subset W$ et finalement $W = E_x$. En particulier E_x est bien de dimension r.

La solution en découle aisément :

• Si u est cyclique, il existe $x \in E$ tel que $E_x = E$ donc tel que $\deg \mu_x = n$ où $n = \dim E$. On a alors $\deg \mu \geqslant n$ d'après la première question. Mais on sait par le théorème de Cayley-Hamilton que μ divise le polynôme caractéristique χ_u de u. Ce dernier étant de degré n on a donc $\mu = \chi_u$.

• Réciproquement, supposons $\chi_u = \mu$. D'après la question précédente, il existe $x \in E$ tel que $\mu_x = \mu = \chi_u$. Pour un tel x on a $\dim E_x = \deg \mu_x = \deg \mu = \deg \chi_u = n$ et donc $E_x = E$. Ainsi u est cyclique. ◁

Il découle directement de cet exercice qu'un endomorphisme nilpotent d'indice de nilpotence égal à la dimension n de l'espace E *est cyclique. En effet, son polynôme minimal vaut* X^n *et est égal au polynôme caractéristique. C'est un exercice classique de démontrer cela directement : on sait que pour tout vecteur x en dehors de* $\operatorname{Ker} u^{n-1}$, *la famille* $(x, u(x), \ldots, u^{n-1}(x))$ *est une base de* E.

Voici une première application de la notion d'endomorphisme cyclique pour caractériser les endomorphismes simples, c'est-à-dire ceux qui n'ont aucun sous-espace stable non trivial.

2.39. Endomorphismes simples

Soit E un K-espace vectoriel de dimension finie $n \geqslant 1$, u un endomorphisme de E. Montrer l'équivalence des deux conditions suivantes :

(i) les seuls sous-espaces de E stables par u sont $\{0\}$ et E ;

(ii) le polynôme caractéristique χ_u de u est irréductible sur K.

(ENS Ulm)

▷ **Solution.**

• $(ii) \Longrightarrow (i)$.

Supposons χ_u irréductible. Soit F un sous-espace de E stable par u. Nous savons que $\chi_{u_{|\mathrm{F}}}$ divise χ_u. Par hypothèse, il en résulte que $\deg \chi_{u_{|\mathrm{F}}}$ vaut 0 ou n ($= \deg \chi_u$). Puisque $\deg \chi_{u_{|\mathrm{F}}} = \dim \mathrm{F}$, le sous-espace F est donc réduit à $\{0\}$ ou est égal à E tout entier.

• $(i) \Longrightarrow (ii)$.

Réciproquement, supposons que les seuls sous-espaces stables par u soient $\{0\}$ et E. Il est facile de voir que le polynôme minimal μ de u est irréductible. En effet, supposons que $\mu = \mathrm{PQ}$ avec P, Q de degré $\geqslant 1$. On a alors $\mathrm{E} = \operatorname{Ker} \mu(u) = \operatorname{Ker} \mathrm{P}(u) \oplus \operatorname{Ker} \mathrm{Q}(u)$. Les sous-espaces $\operatorname{Ker} \mathrm{P}(u)$ et $\operatorname{Ker} \mathrm{Q}(u)$ sont stables par E et non triviaux : en effet, si par exemple $\operatorname{Ker} \mathrm{P}(u) = \{0\}$ alors $\mathrm{P}(u)$ est inversible et on a alors $\mathrm{Q}(u) = 0$ ce qui contredit la définition de μ.

L'exercice 2.38 permet alors de conclure. En effet, si x est un vecteur non nul de E, le sous-espace $\mathrm{E}_x = \operatorname{Vect}(u^k(x))_{k \in \mathbb{N}}$ est stable par u et non nul. Il est donc égal à E, ce qui veut dire que u est cyclique. On sait alors que son polynôme minimal μ est égal à son polynôme caractéristique. En particulier χ_u est irréductible. ◁

Voici deux exercices sur la notion de polynôme minimal ponctuel. Lorsqu'on se place en dimension infinie, il peut arriver que tout vecteur x admette un polynôme minimal ponctuel non nul sans que l'endomorphisme u n'ait de polynôme minimal. Prenons par exemple la dérivation D *dans l'espace* $\mathbb{C}[X]$ *: pour tout polynôme* P *la famille* $(D^k(P))_{k\in\mathbb{N}}$ *est liée mais* D *n'a pas de polynôme annulateur non nul. En effet, le degré du polynôme minimal ponctuel* μ_P *d'un polynôme donné* P *est égal à* $1+\deg P$ *(c'est tout simplement* $X^{1+\deg P}$*). Un polynôme annulateur de* D *doit être multiple de tous les* X^k*,* $k\in\mathbb{N}$*, et est donc nul. L'exercice suivant montre que le caractère majoré de la famille des* $(\deg \mu_x)_{x\in E}$ *est en fait suffisant pour avoir l'existence d'un polynôme minimal global. Il est important de faire l'exercice 2.38 avant de l'aborder.*

2.40. Polynôme minimal ponctuel (1)

Soit V un espace vectoriel quelconque, f un endomorphisme de V et $n\in\mathbb{N}^*$. On suppose que, pour tout $x\in V$, la famille $(x, f(x), \ldots, f^n(x))$ est liée. Montrer que la famille $(\mathrm{Id}_V, f, \ldots, f^n)$ est liée.

(ENS Ulm)

▷ **Solution.**

On peut remarquer que lorsque n vaut 1, on retrouve le lemme célèbre qui affirme que si pour x dans V, x et $f(x)$ sont colinéaires, alors f est une homothétie.

Pour $x\in V$ on pose $I_x=\{Q\in K[X],\ Q(f)(x)=0\}$. Comme il a été vu dans l'exercice 2.38, I_x est un idéal de $K[X]$, non nul par hypothèse, dont le générateur unitaire sera noté μ_x (on l'appelle le polynôme minimal ponctuel de x). Par hypothèse on a $\deg\mu_x\leqslant n$ pour tout x de V. Quitte à prendre n plus petit on peut très bien supposer qu'il existe un vecteur a de E tel que $\deg\mu_a=n$.

Nous allons montrer que μ_a annule f ce qui répondra à la question. Prenons un vecteur x quelconque dans V. Le sous-espace

$$\begin{aligned} W &= \operatorname*{Vect}_{k\in\mathbb{N}} f^k(x) + \operatorname*{Vect}_{k\in\mathbb{N}} f^k(a) \\ &= \mathrm{Vect}(x, f(x), \ldots, f^n(x)) + \mathrm{Vect}(a, f(a), \ldots, f^n(a)). \end{aligned}$$

est stable par f (car somme de deux sous-espaces stables) et de dimension finie. Notons f' l'endomorphisme induit par f sur W. Par hypothèse le polynôme minimal ponctuel pour f' de tout vecteur y de W est de degré $\leqslant n$ (ce polynôme n'est autre que μ_y). Or nous savons d'après

la question 2 de l'exercice 2.38 qu'il existe $y \in W$ tel que $\mu' = \mu_y$ où μ' désigne le polynôme minimal de f'. Comme $a \in W$ le polynôme μ_a divise μ' et comme ils ont tous les deux le même degré, ils sont égaux. On a donc $\mu' = \mu_a$. En particulier, comme $x \in W$ le polynôme μ_x divise $\mu' = \mu_a$ et on a $\mu_a(f')(x) = \mu_a(f)(x) = 0$. Comme cela est valable pour tout $x \in V$, on a $\mu_a(f) = 0$ et $(\mathrm{Id}_V, f, \ldots, f^n)$ est liée. $\triangleleft$

Dans l'exercice suivant, qui fournit une caractérisation des matrices cycliques, on trouvera une preuve différente du fait que le polynôme minimal est le polynôme minimal ponctuel en un certain vecteur, preuve valable sur tout corps infini.

2.41. Polynôme minimal ponctuel (2)

Soit K un corps commutatif infini et $A \in \mathcal{M}_n(K)$. Montrer que le polynôme minimal de A est de degré n si et seulement si l'application

$$\begin{array}{rccc} & (K^n)^2 & \longrightarrow & K^n \\ \varphi : & (X, Y) & \longmapsto & ({}^tXY, {}^tXAY, \ldots, {}^tXA^{n-1}Y) \end{array}$$

est surjective.

(ENS Ulm)

▷ **Solution.**

• Si $\deg \mu_A < n$, il existe donc $a_0, a_1, \ldots, a_{n-1}$ dans K, non tous nuls tels que

$$a_0 I_n + a_1 A + \cdots + a_{n-1} A^{n-1} = 0.$$

Dans ces conditions, pour tout $(X, Y) \in (K^n)^2$, on a

$$\begin{aligned} 0 &= {}^tX(a_0 I_n + a_1 A + \cdots + a_{n-1} A^{n-1})Y \\ &= a_0({}^tXY) + a_1({}^tXAY) + \cdots + a_{n-1}({}^tXA^{n-1}Y). \end{aligned}$$

Par conséquent, tout élément $(\alpha_0, \ldots, \alpha_{n-1})$ dans l'image de Φ est dans l'hyperplan de K^n d'équation

$$a_0\alpha_0 + a_1\alpha_1 + \cdots + a_{n-1}\alpha_{n-1} = 0.$$

Il s'ensuit que Φ n'est pas surjective.

• Supposons $\deg \mu_A = n$. Comme pour tout $P \in K[X]$, $P(A) = 0$ équivaut à $P({}^tA) = 0$, le polynôme minimal de ${}^tA$ est μ_A. D'après l'exercice 2.38 il existe $X \in K^n$ tel que le polynôme minimal ponctuel de X pour ${}^tA$ soit μ_A. En particulier, la famille $(X, {}^tAX, \ldots, {}^tA^nX)$ est libre. La

famille des vecteurs transposés $({}^t\mathrm{X}, {}^t\mathrm{XA}, \ldots, {}^t\mathrm{XA}^{n-1})$ l'est tout autant. Les n formes linéaires $\mathrm{Y} \in \mathrm{K}^n \longmapsto {}^t\mathrm{XA}^k\mathrm{Y}$ pour $0 \leqslant k \leqslant n-1$ sont donc libres dans l'espace dual et le système (S) : ${}^t\mathrm{XA}^k\mathrm{Y} = \alpha_k$, $0 \leqslant k \leqslant n-1$ d'inconnue Y est un système de Cramer pour tout second membre $(\alpha_0, \alpha_1, \ldots, \alpha_{n-1}) \in \mathrm{K}^n$. Il en résulte que la fonction Φ est surjective. Nous allons pour finir donner une nouvelle preuve de l'existence de $\mathrm{X} \in \mathrm{K}^n$ tel que le polynôme minimal ponctuel de X pour ${}^t\mathrm{A}$ soit μ_{A}, preuve qui utilise l'hypothèse K infini. C'est sans doute cette preuve qu'aurait amené l'examinateur lors de l'oral. Plaçons-nous dans le cas abstrait. Soit E un K-espace vectoriel de dimension finie, K étant infini, et u un endomorphisme de E. Considérons l'ensemble $\mathcal{D}$ des diviseurs unitaires de μ_u. Cet ensemble est fini. Pour tout $x \in \mathrm{E}$, le polynôme minimal ponctuel μ_x de x pour u est dans $\mathcal{D}$. Le vecteur x est dans le noyau de $\operatorname{Ker} \mu_x(u)$, si bien que

$$\mathrm{E} = \bigcup_{x \in \mathrm{E}} \operatorname{Ker} \mu_x(u).$$

Par conséquent, E est réunion des noyaux $\operatorname{Ker} \mu_{u,x}(u)$. Or, ces sous-espaces sont en nombre fini puisque $\mathcal{D}$ est fini. Un résultat classique affirme que, lorsque le corps de base est infini, E ne peut être réunion finie de sous-espaces stricts (voir l'exercice 6.2 de notre tome 1 d'Algèbre). Il existe donc un vecteur x tel que $\operatorname{Ker} \mu_x(u) = \mathrm{E}$. Il s'ensuit que $\mu_u = \mu_x$, ce qui achève la preuve. $\lhd$

Notons que lorsqu'on se place sur $\mathbb{C}$, ou plus généralement sur un corps algébriquement clos, on peut aussi prouver ce dernier résultat par récurrence sur la dimension : l'existence d'une valeur propre pour ${}^t u$ donne un hyperplan stable par u qui permet de se ramener à une dimension inférieure. Le lecteur est invité à rédiger les détails. La preuve que nous venons de rédiger montre que, si le corps de base est $\mathbb{R}$ ou $\mathbb{C}$, l'ensemble des vecteurs x tels que $\mu = \mu_x$ est en fait très gros : c'est le complémentaire d'une réunion finie de sous-espaces stricts, donc clairement un ouvert dense de l'espace.

L'exercice suivant est un pas vers le théorème de décomposition de Frobenius qui règle la question des classes de similitude (voir le commentaire qui suit l'exercice).

2.42. Réduction d'un endomorphisme en somme d'endomorphismes cycliques

Soit E un K-espace vectoriel de dimension finie n et u un endomorphisme de E. On note π le polynôme minimal de u. Pour tout $x \in \mathrm{E}$, on note $\mathrm{E}_x = \{\mathrm{P}(u)(x),\ \mathrm{P} \in \mathrm{K}[\mathrm{X}]\}$.

1. On suppose π irréductible. Soit F un sous-espace de E stable par u et $x \in \mathrm{E}$. Montrer que $\mathrm{E}_x \subset \mathrm{F}$ ou $\mathrm{E}_x \cap \mathrm{F} = \{0\}$. Montrer qu'il existe des vecteurs $x_1, \ldots, x_p$ de E tels que $\mathrm{E} = \bigoplus_{i=1}^{p} \mathrm{E}_{x_i}$.

2. Montrer le même résultat en supposant que la décomposition de π en facteurs irréductibles est sans facteur carré.

3. Soit $\mathrm{A} \in \mathcal{M}_n(\mathbb{Q})$ telle que $\mathrm{A}^4 = \mathrm{I}_n$. Montrer qu'il existe $\mathrm{P} \in \mathrm{GL}_n(\mathbb{Q})$ tel que $\mathrm{P}^{-1}\mathrm{AP}$ appartienne à $\mathcal{M}_n(\mathbb{Z})$.

(ENS Lyon)

▷ **Solution.**

Nous supposerons connus les résultats de l'exercice 2.38.

1. • Supposons $\mathrm{E}_x \cap \mathrm{F} \neq \{0\}$ (en particulier $x \neq 0$). Nous allons montrer que $\mathrm{E}_x \subset \mathrm{F}$. Notons $v = u_{|\mathrm{E}_x}$ la restriction de u à E_x : v est alors cyclique. On a donc $\mu_v = \chi_v$ mais comme μ_v divise $\mu_u = \pi$ qui est supposé irréductible, on a $\chi_v = \mu_v = \pi$. Prenons alors un vecteur y non nul quelconque dans $\mathrm{E}_x \cap \mathrm{F}$. On a $\mathrm{E}_y \subset \mathrm{E}_x$ car E_y est le plus petit sous-espace contenant y qui est stable par u. Comme $\dim \mathrm{E}_y \geqslant 1$, $\chi_{v_{|\mathrm{E}_y}}$ est un polynôme de degré supérieur ou égal à 1 qui divise $\chi_v = \pi$. Comme celui-ci est irréductible, on a forcément $\chi_{v_{|\mathrm{E}_y}} = \pi$. En particulier, la dimension de E_y est égale à $\deg \pi = \dim \mathrm{E}_x$. Ainsi $\mathrm{E}_x = \mathrm{E}_y$. D'autre part $\mathrm{E}_y \subset \mathrm{F}$ puisque F est stable par u et contient y. Donc, $\mathrm{E}_x = \mathrm{E}_y \subset \mathrm{F}$.

• Montrons l'existence de $x_1, \ldots, x_p$ tels que $\mathrm{E} = \bigoplus_{i=1}^{p} \mathrm{E}_{x_i}$.

On considère les familles $(x_1, \ldots, x_p)$ d'éléments de E telles que la somme des E_{x_i} soit directe (il en existe, (0) par exemple). On en choisit une de telle manière que $\dim \bigoplus_{i=1}^{p} \mathrm{E}_{x_i}$ soit maximale (une telle famille existe car la dimension des sous-espaces de E est majorée par n). Imaginons un instant que $\mathrm{F} = \bigoplus_{i=1}^{p} \mathrm{E}_{x_i}$ soit distinct de E. Alors, il existe $x \in \mathrm{E}$ tel que $x \notin \mathrm{F}$. Par conséquent, E_x n'est pas contenu dans F, et d'après ce qui précède, $\mathrm{E}_x \cap \mathrm{F} = \{0\}$. Mais alors, la somme $\mathrm{F} + \mathrm{E}_x$ est directe et

$$\dim(\mathrm{E}_{x_1} \oplus \ldots \oplus \mathrm{E}_{x_p} \oplus \mathrm{E}_v) > \dim(\mathrm{E}_{x_1} \oplus \ldots \oplus \mathrm{E}_{x_p})$$

ce qui contredit le choix de la famille des x_i.

2. Écrivons $\pi = \mathrm{P}_1 \ldots \mathrm{P}_r$ où les P_i sont irréductibles et deux à deux distincts. D'après le théorème de décomposition des noyaux on a

$$\mathrm{E} = \operatorname{Ker} \pi = \bigoplus_{k=1}^{r} \operatorname{Ker} \mathrm{P}_k(u).$$

On note $\mathrm{F}_k = \operatorname{Ker} \mathrm{P}_k(u)$ pour $1 \leqslant k \leqslant r$. Alors u induit un endomorphisme u_k sur F_k et $\mu_{u_k} = \mathrm{P}_k$ (en effet, $\mathrm{P}_k(u_k) = 0$ d'où $\mu_{u_k} | \mathrm{P}_k$; comme $\mathrm{F}_k \neq \{0\}$, sinon π / P_k annulerait u, on a bien l'égalité). On sait alors que F_k est somme directe de sous-espaces du type :

$$\{\mathrm{P}(u_k)(x).\ \mathrm{P} \in \mathrm{K}[\mathrm{X}]\} = \{\mathrm{P}(u)(x).\ \mathrm{P} \in \mathrm{K}[\mathrm{X}]\} = \mathrm{E}_x$$

Par associativité de la somme directe, on a le résultat pour E tout entier.

3. On prend dans cette question $\mathrm{E} = \mathbb{Q}^n$ et u l'endomorphisme canoniquement associé à A. Le polynôme μ_A divise $\mathrm{X}^4 - 1 = (\mathrm{X} - 1)(\mathrm{X} + 1)(\mathrm{X}^2 + 1)$. Donc μ_A est produit d'irréductibles de $\mathbb{Q}[\mathrm{X}]$ deux à deux distincts. D'après ce qui précède $\mathbb{Q}^n$ est somme directe de sous-espaces cycliques du type E_x avec $\mu_{u_{|\mathrm{E}_x}} = \mathrm{X} - 1$, $\mathrm{X} + 1$ ou $\mathrm{X}^2 + 1$. Dans une base bien choisie, la matrice de $u_{|\mathrm{E}_x}$ est (1), (−1) ou $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$. Par conséquent, A est semblable dans $\mathcal{M}_n(\mathbb{Q})$ à une matrice B diagonale par blocs, avec sur la diagonale des blocs du type (1), (−1) ou $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$. On a alors $\mathrm{B} \in \mathcal{M}_n(\mathbb{Z})$. ◁

Le cas général est sans doute un peu long pour faire l'objet d'un exercice d'oral. Donnons tout de même le résultat qui répondra enfin à la question posée tout au début du chapitre de savoir quand deux endomorphismes sont semblables. On se donne E *un* K*-espace de dimension* $n \geqslant 1$ *et* u *un endomorphisme de* E *dont on note* μ *le polynôme minimal. Le but est de décomposer* E *en somme directe de sous-espaces stables par* u *dans lesquels* u *induit des endomorphismes cycliques. On sait qu'on peut trouver un vecteur* x *de* E *tel que* $\mu_x = \mu$ *où* μ_x *désigne le polynôme minimal ponctuel en* x *(voir l'exercice 2.38). Le lemme fondamental consiste à prouver que* $\mathrm{E}_x = \operatorname{Vect}(u^k(x))_{k \geqslant 0}$ *admet un supplémentaire* F *stable par* u. *Il suffit alors de recommencer le travail avec la restriction* v *de* u *à* F. *On notera que comme* μ *annule* v, *le polynôme minimal de* v *divise* μ. *On obtient donc le résultat suivant : l'espace* E *se décompose sous la forme* $\mathrm{E} = \mathrm{E}_{x_1} \oplus \mathrm{E}_{x_2} \oplus \cdots \oplus \mathrm{E}_{x_s}$ *et, en notant* P_k *le polynôme minimal (ou caractéristique) de la restriction de* u *à* E_{x_k}, $\mathrm{P}_s | \mathrm{P}_{s-1} | \ldots | \mathrm{P}_2 | \mathrm{P}_1 = \mu$.

On a donc $\chi_u = P_1P_2 \dots P_s$. *Les sous-espaces* E_{x_i} *ne sont pas uniques, mais on démontre que la suite* $(P_1, \dots, P_s)$ *l'est. On appelle cette suite de polynômes la suite des invariants de similitude de* u. *En effet, le résultat essentiel est alors que deux endomorphismes* u *et* v *sont semblables si et seulement si ils ont les mêmes invariants de similitude.*

Nous commençons maintenant une série d'exercices consacrés à diverses équations portant sur des endomorphismes (ou des matrices).

2.43. Équation matricielle

Trouver les matrices $X \in \mathcal{M}_2(\mathbb{R})$ telles que $X^2 + X = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$.
(École Polytechnique)

▷ **Solution.**

La matrice $A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$ est symétrique réelle, donc diagonalisable (en base orthonormée) et ses valeurs propres sont 0 et 2. Le vecteur $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$ étant dans le noyau de A, on en déduit que A est diagonalisable dans la base orthonormée $\left(\frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix}, \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix}\right)$ et si $P = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$, $P^{-1}AP = \mathrm{Diag}(0, 2)$.

Si X est une matrice solution, X commute avec A et laisse donc stable les sous-espaces propres de A : la matrice $X' = P^{-1}XP$ est donc diagonale. On recherche les matrices répondant au problème de la forme $X = P\,\mathrm{Diag}(\alpha, \beta)P^{-1}$. La matrice X répond au problème si et seulement si $\mathrm{Diag}(\alpha, \beta)^2 + \mathrm{Diag}(\alpha, \beta) = \mathrm{Diag}(0, 2)$, autrement dit si et seulement si $\alpha^2 + \alpha = 0$ et $\beta^2 + \beta = 2$. On trouve alors quatre couples solutions correspondant à $\alpha = -1$ ou 0 et $\beta = 1$ ou 2. En calculant $P\,\mathrm{Diag}(\alpha, \beta)P^{-1}$, on obtient les quatre solutions suivantes

$$\begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix}, \ \frac{1}{2}\begin{pmatrix} -3 & -1 \\ -1 & -3 \end{pmatrix}, \ \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \text{ et } \begin{pmatrix} -1 & -1 \\ -1 & -1 \end{pmatrix}. \ \triangleleft$$

Les deux exercices suivants sont consacrés au commutant : la recherche du commutant de u *correspond à la résolution de l'équation* $uv - vu = 0$.

2.44. Commutant d'une matrice carrée de taille 2

Soit $A \in \mathcal{M}_2(\mathbb{C})$. L'algèbre C des matrices de $\mathcal{M}_2(\mathbb{C})$ commutant avec A peut-elle être un corps ?

(École polytechnique)

▷ **Solution.**

La réponse est non ! Si A est une homothétie, son commutant est $\mathcal{M}_2(\mathbb{C})$ tout entier et ce n'est pas un corps. Sinon, le commutant de A, qui contient tous les polynômes en A, contient au moins le plan $\mathrm{Vect}(I_2, A)$. Comme le corps de base est $\mathbb{C}$, la matrice A admet au moins une valeur propre λ. Alors $A - \lambda I_2$ n'est pas inversible et est dans le commutant de A. ◁

Notons que le résultat n'est plus vrai sur le corps des réels par exemple. Si on prend $A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$, *il est facile de vérifier que le commutant de* A *est le plan* $\mathrm{Vect}(I_2, A)$ *formé des matrices* $\begin{pmatrix} a & -b \\ b & a \end{pmatrix}$. *Or, il est aisé de vérifier que ce commutant est un corps isomorphe à* $\mathbb{C}$.

On sait que le commutant d'une matrice A *contient la sous-algèbre* $K[A]$ *des polynômes en* A. *L'exercice suivant montre qu'il y a égalité si et seulement si* A *est cyclique.*

2.45. Dimension du commutant

Soit $A \in \mathcal{M}_n(K)$. On suppose A diagonalisable et on note $\lambda_1, \ldots, \lambda_r$ les valeurs propres deux à deux distinctes de A et $n_1, \ldots, n_r$ leurs multiplicités respectives. On note $\mathcal{C}(A)$ le commutant de A et $K[A] = \{P(A),\ P \in K[X]\}$.

1. Calculer les dimensions de $\mathcal{C}(A)$ et $K[A]$.

2. Montrer les équivalences suivantes :

$$\dim \mathcal{C}(A) = n \iff \dim K[A] = n \iff r = n \iff \mathcal{C}(A) = K[A].$$

3. On ne suppose plus A diagonalisable. Trouver une condition nécessaire et suffisante sur A pour que $K[A] = \mathcal{C}(A)$.

(ENS Ulm)

▷ **Solution.**

1. • Calculons la dimension de $\mathrm{K}[\mathrm{A}]$. Le polynôme minimal de A est $\mu_{\mathrm{A}} = (\mathrm{X} - \lambda_1) \dots (\mathrm{X} - \lambda_r)$ puisque A est diagonalisable. Si on considère l'application linéaire

$$\varphi : \mathrm{P} \in \mathrm{K}[\mathrm{X}] \longmapsto \mathrm{P}(\mathrm{A}) \in \mathrm{K}[\mathrm{A}],$$

elle est surjective et son noyau est $\mu_{\mathrm{A}}\mathrm{K}[\mathrm{X}]$. Sa restriction à $\mathrm{K}_{r-1}[\mathrm{X}]$ est donc injective, et même surjective puisque si $\mathrm{P} \in \mathrm{K}[\mathrm{X}]$ on a $\mathrm{P}(\mathrm{A}) = \mathrm{R}(\mathrm{A})$, où R est le reste de P modulo μ_{A}. Ainsi, la dimension de $\mathrm{K}[\mathrm{A}]$ est égale à celle de $\mathrm{K}_{r-1}[\mathrm{X}]$, c'est-à-dire à r.

• Identifions les matrices aux endomorphismes de K^n canoniquement associés et notons $\mathrm{E}_1, \dots, \mathrm{E}_r$ les sous-espaces propres de A associés respectivement à $\lambda_1, \dots, \lambda_r$. Si B commute avec A, alors B laisse stable chaque sous-espace E_i et induit sur E_i un endomorphisme B_i. Considérons l'application Ψ qui à B dans le commutant de A associe $(\mathrm{B}_1, \dots, \mathrm{B}_r)$ dans $\mathcal{L}(\mathrm{E}_1) \times \cdots \times \mathcal{L}(\mathrm{E}_r)$. Cette application est clairement linéaire et injective (car si B est nulle sur chaque E_i, $\mathrm{B} = 0$). Elle est également surjective : en effet, si on se donne $(\mathrm{B}_1, \dots, \mathrm{B}_r)$ dans $\mathcal{L}(\mathrm{E}_1) \times \cdots \times \mathcal{L}(\mathrm{E}_r)$, l'endomorphisme B qui coïncide avec B_i sur le sous-espace E_i est bien dans le commutant de B, puisque chaque $\mathrm{B}_{|\mathrm{E}_i} = \mathrm{B}_i$ commute avec $\mathrm{A}_{|\mathrm{E}_i} = \lambda_i \mathrm{Id}_{\mathrm{E}_i}$. Il en résulte que $\mathcal{C}(\mathrm{A})$ est isomorphe à $\mathcal{L}(\mathrm{E}_1) \times \cdots \times \mathcal{L}(\mathrm{E}_r)$ et

$$\dim \mathcal{C}(\mathrm{A}) = n_1^2 + n_2^2 + \cdots + n_r^2,$$

puisque n_i est la dimension de E_i.

2. Notons que comme pour tout $i \in [\![1, r]\!]$, on a $n_i^2 \geqslant n_i \geqslant 1$, on a l'inégalité

$$\dim \mathcal{C}(\mathrm{A}) = n_1^2 + \cdots + n_r^2 \geqslant n_1 + \cdots + n_r = n.$$

Rappelons aussi que $\mathrm{K}[\mathrm{A}] \subset \mathcal{C}(\mathrm{A})$. Pour avoir égalité il suffit donc que les deux espaces aient la même dimension.

• Supposons que $\dim \mathcal{C}(\mathrm{A}) = n$. L'inégalité précédente montre que nécessairement $n_i^2 = n_i$ pour tout i. On a donc $n_i = 1$ pour tout i et par suite $r = n$. Ainsi, $\dim \mathrm{K}[\mathrm{A}] = n = \dim \mathcal{C}(\mathrm{A})$ et $\mathcal{C}(\mathrm{A}) = \mathrm{K}[\mathrm{A}]$.

• Supposons que $\dim \mathrm{K}[\mathrm{A}] = n$ c'est-à-dire que $r = n$. Comme la somme $n_1 + \cdots + n_n$ vaut n, chaque n_i est nécessairement égal à 1. Donc $\dim \mathcal{C}(\mathrm{A}) = n_1^2 + \cdots + n_n^2 = n = \dim \mathrm{K}[\mathrm{A}]$ et comme on a l'inclusion $\mathrm{K}[\mathrm{A}] \subset \mathcal{C}(\mathrm{A})$, on a l'égalité des deux sous-espaces : $\mathcal{C}(\mathrm{A}) = \mathrm{K}[\mathrm{A}]$.

• Supposons enfin que $\mathrm{K}[\mathrm{A}] = \mathcal{C}(\mathrm{A})$. On a donc les inégalités suivantes

$$\dim \mathrm{K}[\mathrm{A}] = r \leqslant n = n_1 + \cdots + n_r \leqslant n_1^2 + \cdots + n_r^2 = \dim \mathcal{C}(\mathrm{A}).$$

Comme $\dim \mathrm{K}[\mathrm{A}] = \dim \mathcal{C}(\mathrm{A})$, on en déduit que $\mathcal{C}(\mathrm{A})$ est de dimension n.

3. La dimension de K[A] est toujours égale au degré de μ_A. Cette dimension est donc d'après le théorème de Cayley-Hamilton inférieure ou égale à n. Établissons le lemme suivant sur celle du commutant :

Lemme. *La dimension du commutant de* A *est toujours supérieure ou égale à* n.

Démonstration. [1] Il s'agit en fait de démontrer que $\dim S \geqslant n$, où S est le sous-espace des solutions du système linéaire

$$AX - XA = 0,$$

où l'inconnue est $X = (x_{ij})_{1 \leqslant i,j \leqslant n} \in \mathcal{M}_n(K)$. Plaçons-nous d'abord dans le cas où A est trigonalisable. Quitte à se placer dans une nouvelle base idoine, on peut supposer $A = (a_{ij})_{1 \leqslant i,j \leqslant n}$ triangulaire supérieure. Cherchons les solutions $X = (x_{ij})_{1 \leqslant i,j \leqslant n}$ triangulaires supérieures. Si on restreint le système à $T_n(K)$, il reste $\frac{n(n+1)}{2}$ inconnues. Comme $AX - XA$ est triangulaire supérieure, dire que X est solution revient à écrire $\frac{n(n+1)}{2}$ équations correspondant à la nullité des coefficients de $AX - XA$ dans la partie supérieure. Mais de ces équations, on peut en retirer n, puisque les équations traduisant la nullité des coefficients diagonaux sont triviales : $a_{ii}x_{ii} - x_{ii}a_{ii} = 0$ (pour $1 \leqslant i \leqslant n$). Ce système homogène a donc seulement $\frac{n(n+1)}{2} - n$ équations pour $\frac{n(n+1)}{2}$ inconnues. Nous savons alors que l'espace des solutions est au moins de dimension n.

Si $\dim S \cap T_n(K) \geqslant n$, *a fortiori*, la dimension de S est au moins n.

Le résultat reste valable lorsque A n'est pas trigonalisable. La dimension de l'espace des solutions du système $AX - XA = 0$ reste inchangée si on remplace K par un sur-corps commutatif L de K (c'est un corollaire direct du lemme de l'exercice 2.14). Il suffit alors de prendre L de telle manière que χ_A soit scindé. Dans ce cas-là, A est trigonalisable et la dimension de l'espace des solutions est supérieure ou égale à n, que ce soit sur L ou sur K. On conclut que $\dim \mathcal{C}(A) \geqslant n$. ♢

Dans le cas où $K[A] = \mathcal{C}(A)$, le lemme entraîne que la dimension de ce sous-espace est n. En particulier, on a $\dim K[A] = \deg \mu_A = n$. Comme μ_A est un diviseur de χ_A, on obtient que

$$\boxed{\mu_A = \chi_A}.$$

Réciproquement, supposons que $\chi_A = \mu_A$. On sait alors que A est cyclique (se reporter à l'exercice 2.38), autrement dit qu'il existe un

1. Le preuve que nous allons donner est celle de notre collègue Yves Duval. Elle a été publiée dans la RMS 114 de janvier 2004.

vecteur e de K^n tel que $(e, \mathrm{A}e, \mathrm{A}^2e, \ldots, \mathrm{A}^{n-1}e)$ soit une base de K^n. Considérons alors l'application

$$f : \mathrm{B} \in \mathcal{C}(\mathrm{A}) \longmapsto \mathrm{B}e \in \mathrm{K}^n.$$

Cette application est linéaire. Elle est également injective, car si $\mathrm{B}e = 0$, on a

$$\mathrm{B}\mathrm{A}^k e = \mathrm{A}^k \mathrm{B}e = 0 \text{ pour tout } k \geqslant 0.$$

L'endomorphisme B s'annule sur une base de K^n, donc est nul.

On en déduit que $\dim \mathcal{C}(\mathrm{A}) \leqslant \dim \mathrm{K}^n = n$. Compte tenu du lemme, on a donc $\dim \mathcal{C}(\mathrm{A}) = \dim \mathrm{K}[\mathrm{A}] = \deg \mu_{\mathrm{A}} = n$. Comme on a l'inclusion $\mathrm{K}[\mathrm{A}] \subset \mathcal{C}(\mathrm{A})$, on a l'égalité $\mathcal{C}(\mathrm{A}) = \mathrm{K}[\mathrm{A}]$.

Conclusion. $\boxed{\mathcal{C}(\mathrm{A}) = \mathrm{K}[\mathrm{A}] \iff \mathrm{A} \text{ cyclique} \iff \chi_{\mathrm{A}} = \mu_{\mathrm{A}}}$. ◁

Notons que le résultat de la dernière question est immédiat dès lors que l'on dispose du théorème de décomposition en sommes de sous-espaces cycliques exposé page 131. En effet, si u n'est pas cyclique, alors sa liste de facteurs invariants $(\mathrm{P}_1, \ldots, \mathrm{P}_s)$ contient au moins 2 polynômes. Notons $\mathrm{E} = \mathrm{E}_{x_1} \oplus \cdots \oplus \mathrm{E}_{x_s}$ une décomposition de E en sous-espaces cycliques correspondants aux polynômes P_i. Il est clair que la projection sur E_{x_s} parallèlement à la somme des autres sous-espaces commute avec u. Mais elle ne peut pas être un polynôme en u car P_s divise P_1.

Voici une équation toujours construite avec le crochet de Lie $[f, g] = fg - gf$ mais avec un second membre. Lorsqu'il y a des solutions (g étant l'inconnue), il s'agit bien entendu d'un sous-espace affine dirigé par le commutant de f. La solution est courte car la plupart des arguments ont déjà été rencontrés dans les exercices précédents.

2.46. Équation avec un crochet de Lie (1)

Soit E un K-espace vectoriel de dimension finie n (K sous-corps de $\mathbb{C}$) et f et g des endomorphismes de E tels que $fg - gf = f$.

1. Montrer que f est nilpotent.

2. On suppose que g est diagonalisable et que $\dim \operatorname{Ker} f = 1$. Déterminer g.

(École polytechnique)

▷ **Solution.**

1. Nous avons donné trois solutions différentes de cette question dans l'exercice 2.27. Notons que c'est pour cette question que sert l'hypothèse faite sur le corps K.

2. L'endomorphisme f est nilpotent de rang $n-1$. On a vu dans l'exercice 2.34 qu'on a alors $\dim \operatorname{Ker} f^k = k$ pour tout $k \in [\![0, n]\!]$ et, comme nous l'avons signalé page 126 dans la remarque qui suit l'exercice 2.38, f est cyclique : pour tout vecteur $e \notin \operatorname{Ker} f^{n-1}$ la famille $(e, f(e), \ldots, f^{n-1}(e))$ est une base de E.

On va expliciter g dans une telle base en choisissant pour e un vecteur propre de g : c'est possible puisque si g est diagonalisable elle possède forcément un vecteur propre e qui n'appartient pas à $\operatorname{Ker} f^{n-1}$. On note λ la valeur propre associée. On a alors

$$f(e) = fg(e) - gf(e) = \lambda f(e) - g(f(e))$$

de sorte que $g(f(e)) = (\lambda - 1)f(e)$. Autrement dit, $f(e)$ est vecteur propre de g pour la valeur propre $\lambda - 1$. En itérant, on voit de même que $f^k(e)$ est un vecteur propre de g relativement à la valeur propre $\lambda - k$ pour tout $k \in [\![0, n-1]\!]$. On en déduit que $(e, f(e), \ldots, f^{n-1}(e))$ est une base de vecteurs propres de g.

Réciproquement, si e est un vecteur quelconque de E qui n'appartient pas à $\operatorname{Ker} f^{n-1}$, la famille $(e, f(e), \ldots, f^{n-1}(e))$ est une base de E et pour tout $\lambda \in \mathrm{K}$, l'endomorphisme g défini par $g(f^k(e)) = (\lambda - k)f^k(e)$ pour tout $k \in [\![0, n-1]\!]$, est solution de $fg - gf = f$. ◁

Voici une équation du même type où il n'y a pas de solution.

2.47. Équation avec un crochet de Lie (2)

Soit E un K-espace vectoriel de dimension finie, $(f, h) \in \mathcal{L}(\mathrm{E})^2$. On suppose χ_f irréductible et $\operatorname{rg} h = 1$. Montrer que l'équation $fg - gf = h$ d'inconnue $g \in \mathcal{L}(\mathrm{E})$ n'a pas de solution.

(**École polytechnique**)

▷ **Solution.**

On raisonne bien entendu par l'absurde en supposant qu'il y a une solution g. On va essayer de trouver un sous-espace F stable par f et non trivial (*i.e.* de dimension comprise entre 1 et $\dim \mathrm{E} - 1$), ce qui contredira l'irréductibilité de χ_f : en effet, le polynôme caractéristique de la restriction de f à F sera un diviseur non trivial de χ_f (on utilise le sens facile de l'exercice 2.39). Soit e une base de $\operatorname{Im} h$ et F l'espace engendré par les vecteurs $f^k(e)$, $k \in \mathbb{N}$. C'est un sous-espace vectoriel de E stable par f, de dimension non nulle puisqu'il contient e. Montrons que $\mathrm{F} \subset \operatorname{Ker} h$, ce qui montrera que $\dim \mathrm{F} \leqslant n - 1$. Il suffit de montrer

que, pour tout $k \in \mathbb{N}$, $f^k(e) \in \operatorname{Ker} h$. Il existe $x \in \mathrm{E}$ tel que $e = h(x)$ et $f^k(e) = (f^k h)(x)$. On a $\operatorname{rg}(f^k h) \leqslant \operatorname{rg} h \leqslant 1$. Soit $f^k h = 0$ et il n'y a rien à démontrer, soit $\operatorname{rg}(f^k h) = \operatorname{rg} h = 1$. On considère alors une base $\mathcal{B} = (e_1, \ldots, e_n)$ de E telle que $\operatorname{Vect}(e_1, \ldots, e_{n-1}) = \operatorname{Ker}(f^k h)$ et soit A la matrice de $f^k h$ dans cette base. On a $\operatorname{Tr}(f^k h) = \operatorname{Tr} \mathrm{A} = a_{nn}$ et par ailleurs $\operatorname{Tr}(f^k h) = \operatorname{Tr}(f^{k+1} g) - \operatorname{Tr}(f^k g f) = 0$, donc $a_{nn} = 0$. Cela montre que $\operatorname{Im}(f^k h) \subset \operatorname{Ker}(f^k h)$. Il est clair que $\operatorname{Ker} h \subset \operatorname{Ker}(f^k h)$. Comme ces sous-espaces ont même dimension, ils sont égaux. On en déduit que $f^k(e) = (f^k h)(x) \in \operatorname{Ker} h$ et $\mathrm{F} \subset \operatorname{Ker} h$. $\triangleleft$

Encore un exercice portant sur le crochet de Lie.

2.48. Le crochet de Lie

Soit V un espace vectoriel de dimension finie sur $\mathbb{C}$. Pour a, b dans $\mathcal{L}(\mathrm{V})$, on pose $\varphi_a(b) = ab - ba$.

1. Calculer $\varphi_a^n(b)$ pour a, b dans $\mathcal{L}(\mathrm{V})$ et $n \in \mathbb{N}$.

2. Montrer que $a^{n+1} b - b a^{n+1} = \sum_{k=0}^{n} a^k (ab - ba) a^{n-k}$.

3. Montrer que si 0 est valeur propre de $a \in \mathcal{L}(\mathrm{V})$, alors a est nilpotent si et seulement si φ_a est nilpotent.

4. Montrer que si $a \in \mathcal{L}(\mathrm{V})$ n'a qu'une valeur propre, alors φ_a est nilpotent. Étudier la réciproque.

(École polytechnique)

$\triangleright$ **Solution.**

1. L'application φ_a se décompose sous la forme $\varphi_a = g - d$ où g (resp. d) est la composition à gauche (resp. à droite) par a. Comme g et d commutent, on peut appliquer la formule du binôme de Newton dans l'algèbre $\mathcal{L}(\mathcal{L}(\mathrm{V}))$. Il vient

$$\varphi_a^n = \sum_{k=0}^{n} \mathrm{C}_n^k g^{n-k} (-1)^k d^k.$$

Autrement dit, pour $b \in \mathcal{L}(\mathrm{V})$,

$$\varphi_a^n(b) = \sum_{k=0}^{n} \mathrm{C}_n^k (-1)^k a^{n-k} b a^k.$$

2. Pour a et b dans $\mathcal{L}(\mathrm{V})$, on peut écrire

$$\begin{aligned}\sum_{k=0}^{n} a^k(ab-ba)a^{n-k} &= \sum_{k=0}^{n} a^{k+1}ba^{n-k} - \sum_{k=0}^{n} a^k ba^{n+1-k}\\ &= \sum_{k=1}^{n+1} a^k ba^{n+1-k} - \sum_{k=0}^{n} a^k ba^{n+1-k}\\ &= a^{n+1}b - ba^{n+1}.\end{aligned}$$

3. • Supposons a nilpotent et prenons $n \geqslant 0$ tel que $a^n = 0$. Si $k \in [\![0, 2n]\!]$, $2n-k$ ou k est supérieur ou égal à n, si bien que pour tout $b \in \mathcal{L}(\mathrm{V})$,

$$\varphi_a^{2n}(b) = \sum_{k=0}^{2n} \mathrm{C}_{2n}^k (-1)^k a^{2n-k} b a^k = 0.$$

Donc φ_a est nilpotent.

• Réciproquement, on suppose que 0 est valeur propre de a et que a n'est pas nilpotent. Nous allons montrer que φ_a n'est pas nilpotent en construisant un vecteur propre de φ_a correspondant à une valeur propre non nulle.

Comme a n'est pas nilpotent, il existe $\lambda \in \mathbb{C}^*$ tel que $\mathrm{Ker}(a - \lambda\,\mathrm{Id}) \neq \{0\}$. Comme 0 est valeur propre de a, a n'est pas un isomorphisme et $\mathrm{Im}\, a \neq \mathrm{V}$. Soit $\mathrm{F} \neq \{0\}$ un supplémentaire de $\mathrm{Im}\, a$. On peut construire un endomorphisme b nul sur $\mathrm{Im}\, a$ et qui envoie F sur une droite de $\mathrm{Ker}(a - \lambda\,\mathrm{Id})$ (avec $b_{|\mathrm{F}} \neq 0$). Par construction, l'image de b est contenue dans $\mathrm{Ker}(a - \lambda\,\mathrm{Id})$. Si $x \in \mathrm{V}$, on a

$$(ab-ba)(x) = ab(x) - 0 = \lambda b(x), \text{ et donc } \varphi_a(b) = \lambda b.$$

4. Si λ est l'unique valeur propre de a, on constate que $\varphi_{a-\lambda\,\mathrm{Id}} = \varphi_a$ et $a - \lambda\,\mathrm{Id}$ est nilpotente. Donc φ_a est nilpotente d'après la question précédente.

Réciproquement, supposons φ_a nilpotent. Soit λ une valeur propre de a. Alors $\varphi_{a-\lambda\,\mathrm{Id}} = \varphi_a$ est nilpotent et 0 est valeur propre de $a - \lambda\,\mathrm{Id}$. D'après la question précédente, $a - \lambda\,\mathrm{Id}$ est nilpotent et finalement, a n'a qu'une valeur propre. La réciproque est donc vraie. ◁

Le thème des exercices suivants est l'équation de Sylvester $\mathrm{AX} - \mathrm{XB} = \mathrm{Y}$ *où* $\mathrm{A}, \mathrm{B}, \mathrm{Y}$ *sont des matrices carrées fixées et* X *l'inconnue. On retrouve le problème de la recherche du commutant en prenant* $\mathrm{A} = \mathrm{B}$ *et* $\mathrm{Y} = 0$.

2.49. Équation de Sylvester (1)

Soient A et B dans $\mathcal{M}_n(\mathbb{C})$. À quelle condition l'équation

$$AX - XB = Y$$

a-t-elle une solution $X \in \mathcal{M}_n(\mathbb{C})$, quelle que soit $Y \in \mathcal{M}_n(\mathbb{C})$?

(**École polytechnique**)

▷ **Solution.**

On cherche à quelle condition l'endomorphisme Φ de $\mathcal{M}_n(\mathbb{C})$ défini par $\Phi(X) = AX - XB$ est surjectif. Comme $\mathcal{M}_n(\mathbb{C})$ est de dimension finie, il faut que Φ soit injectif. Montrons que ceci est réalisé si et seulement si A et B n'ont pas de valeur propre commune.

• Supposons que A et B n'ont pas de valeur propre commune. Comme $\mathbb{C}$ est algébriquement clos, leurs polynômes caractéristiques χ_A et χ_B sont premiers entre eux. D'après le théorème de Bezout, il existe U et V dans $\mathbb{C}[X]$ tels que $U\chi_A + V\chi_B = 1$. On a alors $U(A)\chi_A(A) + V(A)\chi_B(A) = I_n$. Puisque $\chi_A(A) = 0$, d'après le théorème de Cayley-Hamilton, on en déduit que $V(A)\chi_B(A) = I_n$ et donc que $\chi_B(A)$ est inversible. Soit maintenant $X \in \mathcal{M}_n(\mathbb{C})$ tel que $\Phi(X) = 0$, c'est-à-dire $AX = XB$. On montre facilement par récurrence que pour tout $k \in \mathbb{N}$, on a : $A^kX = XB^k$. D'où l'on déduit, pour tout $P \in \mathbb{C}[X]$, $P(A)X = XP(B)$. En particulier, on a $\chi_B(A)X = X\chi_B(B)$ et donc $\chi_B(A)X = 0$, puisque $\chi_B(B) = 0$. Puisque $\chi_B(A)$ est inversible, on en déduit $X = 0$; Φ est donc injective.

• Supposons qu'au contraire, A et B ont une valeur propre commune λ. On va construire une matrice non nulle dans le noyau de Φ. L'idée est de trouver une matrice non nulle M telle que $AM = \lambda M$ et $MB = \lambda M$. La première égalité a lieu si toutes les colonnes de M sont des vecteurs propres de A pour la valeur propre λ et la seconde a lieu si toutes les lignes de M sont des vecteurs propres de B toujours pour la valeur propre λ. Il suffit pour fabriquer une telle matrice de se donner un vecteur propre X pour A, un vecteur propre Y pour B et de considérer la matrice $M = X^tY$: toutes ses colonnes sont colinéaires à X et toutes ses lignes sont colinéaires à Y. De plus elle n'est pas nulle, car son terme général est x_iy_j où $X = (x_1, \ldots, x_n)$ et $Y = (y_1, \ldots, y_n)$ et chacun des deux vecteurs a au moins une coordonnée non nulle. On vient donc de construire M non nulle telle que $\Phi(M) = 0$: Φ n'est pas injective et l'équivalence voulue est établie. ◁

Voici un second énoncé, plus détaillé, où l'on retrouve ce résultat en étudiant plus précisément le spectre de l'application $X \mapsto AX - XB$.

2.50. Équation de Sylvester (2)

Soit $(A, B) \in \mathcal{M}_n(\mathbb{C})^2$.

1. Déterminer les spectres des endomorphismes de $\mathcal{M}_n(\mathbb{C})$ définis par $\alpha : M \longmapsto AM$ et $\beta : M \longmapsto MB$. Décrire les sous-espaces propres de α et β et donner leur dimension en fonction des dimensions des sous-espaces propres de A et B.

2. Montrer que si A (resp. B) est diagonalisable, α (resp. B) l'est aussi.

3. Donner une condition nécessaire ou suffisante pour que l'endomorphisme φ de $\mathcal{M}_n(\mathbb{C})$ défini par $\varphi(M) = AM - MB$ soit surjectif.

4. Soient $A = \begin{pmatrix} A_1 & A_2 \\ 0 & A_4 \end{pmatrix}$, $B = \begin{pmatrix} B_1 & B_2 \\ B_3 & B_4 \end{pmatrix}$ telles que A et B commutent et $\operatorname{Sp} A_1 \cap \operatorname{Sp} A_4 = \emptyset$. Montrer que $B_3 = 0$.

(**École polytechnique**)

▷ **Solution.**

1. Si X est un vecteur propre de A pour la valeur propre λ, la matrice M composée de n colonnes égales à X vérifie $AM = \lambda M$: ainsi, λ est valeur propre de α.

Réciproquement, soit λ est une valeur propre de α et M une matrice non nulle de colonnes $(C_1, \ldots, C_n)$. On a

$$M \in \operatorname{Ker}(\alpha - \lambda \operatorname{Id}) \iff AM = \lambda M \iff \forall i \in [\![1, n]\!],\ AC_i = \lambda C_i.$$

Comme une des colonnes de M est non nulle, λ est une valeur propre de A. On a donc $\operatorname{Sp} A = \operatorname{Sp} \alpha$.

De plus, $\operatorname{Ker}(\alpha - \lambda \operatorname{Id})$ est composée des matrices $M = (C_1, \ldots, C_n)$ avec C_i dans $\operatorname{Ker}(A - \lambda I_n)$ pour tout $1 \leqslant i \leqslant n$. Ainsi, $\operatorname{Ker}(\alpha - \lambda \operatorname{Id})$ est isomorphe à $(\operatorname{Ker}(A - \lambda I_n))^n$ et on en déduit que

$$\dim \operatorname{Ker}(\alpha - \lambda \operatorname{Id}) = n \dim \operatorname{Ker}(A - \lambda I_n).$$

Il est inutile de refaire le même travail pour β. En effet, on a clairement

$$\beta(M) = \lambda M \iff MB = \lambda M \iff {}^t B\, {}^t M = \lambda\, {}^t M$$

et comme la transposition est un isomorphisme de $\mathcal{M}_n(\mathbb{C})$ sur lui-même il suffit d'appliquer le résultat précédent à ${}^t B$. On a donc

$$\operatorname{Sp} \beta = \operatorname{Sp}^t B = \operatorname{Sp} B$$

et pour toute valeur propre λ de B,

$$\dim \operatorname{Ker}(\beta - \lambda \operatorname{Id}) = n \dim \operatorname{Ker}({}^t\mathrm{B} - \lambda \operatorname{Id}) = n \dim \operatorname{Ker}(\mathrm{B} - \lambda \operatorname{Id}),$$

car une matrice et sa transposée ont même rang.

2. Si A est diagonalisable, on peut écrire

$$\begin{aligned}\sum_{\lambda \in \operatorname{Sp} \alpha} \dim \operatorname{Ker}(\alpha - \lambda \operatorname{Id}) &= \sum_{\lambda \in \operatorname{Sp} \mathrm{A}} n \dim \operatorname{Ker}(\mathrm{A} - \lambda \operatorname{Id}) \\ &= n \sum_{\lambda \in \operatorname{Sp} \mathrm{A}} \dim \operatorname{Ker}(\mathrm{A} - \lambda \operatorname{Id}) = n^2.\end{aligned}$$

On en déduit que α est diagonalisable. De même, on montre que si B est diagonalisable, β l'est aussi.

Ce calcul prouve qu'en fait il y a équivalence entre la diagonalisabilité de A *et celle de* α. *Pour le montrer, on peut également écrire la matrice de* α *dans la base* $(\mathrm{E}_{11}, \mathrm{E}_{21}, \ldots, \mathrm{E}_{n1}, \mathrm{E}_{12}, \ldots, \mathrm{E}_{nn})$ *: on constate qu'il s'agit d'une matrice diagonale par blocs où, sur la diagonale, on trouve* n *blocs* A. *Il en résulte que* $\mu_{\mathrm{A}} = \mu_\alpha$.

3. Remarquons que $\mathcal{M}_n(\mathbb{C})$ étant de dimension finie, dire que φ est surjective revient à dire que φ est bijective.

Les endomorphismes α et β commutent, si bien qu'il existe une base commune de trigonalisation (voir exercice 2.25). Dans une telle base la matrice de φ est triangulaire et ses coefficients diagonaux sont de la forme $\lambda - \mu$ avec $\lambda \in \operatorname{Sp} \mathrm{A}$ et $\mu \in \operatorname{Sp} \mathrm{B}$. Il s'ensuit que si le spectre de A et le spectre de B sont disjoints, les coefficients diagonaux sont tous non nuls et φ est donc inversible. Réciproquement, si A et B ont une valeur propre commune λ, on procède comme dans l'exercice précédent pour fabriquer une matrice propre à la fois pour α et β c'est-à-dire dans le noyau de φ : rappelons qu'il suffit de prendre la matrice $\mathrm{M} = \mathrm{X}\,{}^t\mathrm{Y} = (x_i y_j)_{1 \leqslant i,j \leqslant n}$ où $\mathrm{X} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ est un vecteur propre de A et $\mathrm{Y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$ un vecteur propre de ${}^t\mathrm{B}$.

Conclusion. L'endomorphisme φ est surjectif si et seulement si

$$\boxed{\operatorname{Sp} \mathrm{A} \cap \operatorname{Sp} \mathrm{B} = \emptyset}.$$

4. En écrivant les matrices par blocs AB et BA, on constate que $\mathrm{AB} = \mathrm{BA}$ entraîne $\mathrm{A}_4\mathrm{B}_3 = \mathrm{B}_3\mathrm{A}_4$. Autrement dit, B_3 est dans le noyau de $\mathrm{M} \longmapsto \mathrm{A}_4\mathrm{M} - \mathrm{M}\mathrm{A}_4$. Or cet endomorphisme est injectif puisque les spectres de A_4 et A_4 sont disjoints. On a donc $\mathrm{B}_3 = 0$. ◁

Dans la version suivante, formulée abstraitement, on se place sur un corps quelconque. L'existence d'une valeur propre commune n'est alors plus assurée. Elle est remplacée par une condition arithmétique sur les polynômes caractéristiques.

2.51. Équation de Sylvester (3)

Soit E un espace vectoriel de dimension finie $n \geqslant 1$ et $(f, g) \in \mathcal{L}(\mathrm{E})^2$.

1. On suppose f et g non inversibles. Montrer qu'il existe $h \in \mathcal{L}(\mathrm{E})$ non nul telle que $h \circ f = g \circ h$.

2. On suppose que f et g ont une valeur propre commune. Montrer qu'il existe $h \in \mathcal{L}(\mathrm{E})$ non nul tel que $h \circ f = g \circ h$.

3. Soit $h \in \mathcal{L}(\mathrm{E})$ tel que $h \circ f = g \circ h$ et $\mathrm{rg}(h) = r \geqslant 1$. Montrer que les polynômes caractéristiques de f et g ont un facteur commun de degré r.

4. Trouver f et g tels que χ_f et χ_g ont un facteur commun de degré r, mais tels qu'il n'existe aucun h de rang supérieur ou égal à r vérifiant $h \circ f = g \circ h$.

(École polytechnique)

▷ **Solution.**

1. On va construire h non nul dans $\mathcal{L}(\mathrm{E})$ tel que $h \circ f = g \circ h = 0$. Géométriquement la condition $h \circ f = 0$ signifie que le noyau de h doit contenir $\mathrm{Im}\, f$ et la condition $g \circ h = 0$ signifie que l'image de h doit être contenue dans $\mathrm{Ker}\, g$. Soit F un supplémentaire de $\mathrm{Im}\, f$ et h_1 une application linéaire non nulle de F dans $\mathrm{Ker}\, g$: cela est possible car ni F, ni $\mathrm{Ker}\, g$ ne sont nuls. L'application $h \in \mathcal{L}(\mathrm{E})$ dont la restriction à F est h_1 et la restriction à $\mathrm{Im}\, f$ est nulle vérifie bien $h \circ f = 0$ car $\mathrm{Im}\, f \subset \mathrm{Ker}\, h$, et $g \circ h = 0$ car $\mathrm{Im}\, h \subset \mathrm{Ker}\, g$.

2. Soit λ une valeur propre commune à f et g. Les endomorphismes $f - \lambda \mathrm{Id}_\mathrm{E}$ et $g - \lambda \mathrm{Id}_\mathrm{E}$ ne sont pas inversibles. D'après la question 1, il existe $h \in \mathcal{L}(\mathrm{E})$, non nul, tel que

$$h \circ (f - \lambda \mathrm{Id}_\mathrm{E}) = (g - \lambda \mathrm{Id}_\mathrm{E}) \circ h.$$

Par distributivité, on obtient $h \circ f - \lambda h = g \circ h - \lambda h$ et donc $h \circ f = g \circ h$.

3. Soit F, G, H les matrices de f, g, h dans une base quelconque. Il existe des matrices P et Q dans $\mathrm{GL}_n(\mathrm{K})$ telles que $\mathrm{H} = \mathrm{PJ}_r\mathrm{Q}$, où $\mathrm{J}_r = \begin{pmatrix} \mathrm{I}_r & 0 \\ 0 & 0 \end{pmatrix}$. L'égalité $h \circ f = g \circ h$ s'écrit $\mathrm{HF} = \mathrm{GH}$ soit

$J_r QFQ^{-1} = P^{-1}GPJ_r$. On écrit par blocs $QFQ^{-1} = \begin{pmatrix} F_1 & F_2 \\ F_3 & F_4 \end{pmatrix}$ et $P^{-1}GP = \begin{pmatrix} G_1 & G_2 \\ G_3 & G_4 \end{pmatrix}$. L'égalité précédente donne $\begin{pmatrix} F_1 & F_2 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} G_1 & 0 \\ G_3 & 0 \end{pmatrix}$ et donc $F_1 = G_1$, $F_2 = 0$, $G_3 = 0$. On en déduit que

$$\begin{aligned} \chi_f = \chi_F = \chi_{QFQ^{-1}} &= \det \begin{pmatrix} XI_r - F_1 & 0 \\ -F_3 & XI_{n-r} - F_4 \end{pmatrix} \\ &= \det(XI_r - F_1) \det(XI_{n-r} - F_4) \end{aligned}$$

et de même

$$\chi_g = \det(XI_r - G_1) \det(XI_{n-r} - G_4) = \det(XI_r - F_1) \det(XI_{n-r} - G_4).$$

Les polynômes caractéristiques de f et g ont un facteur de degré r en commun.

4. Soit f l'application nulle et g un endomorphisme dont la matrice dans une base quelconque est une matrice nilpotente de rang $n-1$, par exemple $\begin{pmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{pmatrix}$. Les endomorphismes f et g ont même polynôme caractéristique X^n. Si $h \in \mathcal{L}(E)$ vérifie $h \circ f = g \circ h$, alors $g \circ h = 0$ et $\operatorname{Im} h \subset \operatorname{Ker} g$. On en déduit que $\operatorname{rg} h \leqslant \dim(\operatorname{Ker} g) \leqslant 1$. Pour $r \in [\![2, n]\!]$, les polynômes caractéristiques de f et g ont un facteur de degré r en commun, mais il n'existe pas d'endomorphisme h de rang r tel que $h \circ f = g \circ h$. ◁

Voici une équation non linéaire. On s'intéresse aux cubes de $\mathcal{M}_3(\mathbb{R})$.

2.52. Matrices possédant une racine cubique

Déterminer l'image de $\Phi : A \in \mathcal{M}_3(\mathbb{R}) \longmapsto A^3 \in \mathcal{M}_3(\mathbb{R})$.
(ENS Ulm)

▷ **Solution.**

Commençons par quelques observations fondamentales. Si B est dans l'image de Φ, donc de la forme $B = A^3$, alors toute matrice semblable à B aussi : en effet, pour $P \in GL_3(\mathbb{R})$ on a

$$P^{-1}BP = P^{-1}A^3P = (P^{-1}AP)^3.$$

On cherche donc plutôt les classes de similitude incluses dans l'image de Φ. Une autre remarque qu'on peut faire dès maintenant est que si $B = A^3$, alors A et B commutent : en effet, $AB = A^4 = BA$.

Soit $B \in \mathcal{M}_3(\mathbb{R})$ donnée. On va discuter selon la manière dont on peut réduire B. L'examinateur s'attendra sûrement à ce que le candidat commence par regarder le cas simple où B est diagonalisable. Comme tout réel admet une racine cubique, il est clair que toute matrice diagonale $D = \mathrm{Diag}(a, b, c)$ est dans l'image de Φ : il suffit de constater que $D = \Delta^3$ avec $\Delta = \mathrm{Diag}(\sqrt[3]{a}, \sqrt[3]{b}, \sqrt[3]{b})$. Il en résulte donc que toute matrice diagonalisable est dans l'image de Φ. On va poursuivre la discussion selon le polynôme caractéristique χ de B. Celui-ci est de degré 3 donc admet au moins une racine réelle.

• Regardons pour commencer le cas où χ n'a qu'une seule racine réelle a et un facteur irréductible de degré 2 qui admet deux racines complexes conjuguées $u \pm iv$ avec $v \neq 0$. Dans ce cas B est diagonalisable dans $\mathcal{M}_3(\mathbb{C})$. En fait, B est semblable dans $\mathcal{M}_3(\mathbb{R})$ à la matrice $B' = \begin{pmatrix} a & 0 & 0 \\ 0 & u & -v \\ 0 & v & u \end{pmatrix}$. En effet, on a $\chi = (X-a)((X-u)^2+v^2)$ et, d'après le lemme de décomposition des noyaux,

$$\mathbb{R}^3 = \mathrm{Ker}(B - aI_3) \oplus \mathrm{Ker}((B - uI_3)^2 + v^2 I_3).$$

Prenons $X \in \mathrm{Ker}((B-uI_3)^2+v^2I_3)$ et $Y = \frac{1}{v}(BX-uX)$. Alors Y n'est pas colinéaire à X, sinon X serait vecteur propre de B. On a $BX = uX + vY$ et

$$BY - uY = (B - uI_3)Y = \frac{1}{v}(B - uI_3)^2X = -vX,$$

soit $BY = -vX+uY$. Dans une base constituée d'un vecteur propre pour la valeur propre a, de X et de Y, l'endomorphisme canoniquement associé à B a pour matrice B'. Il existe donc $P \in \mathrm{GL}_3(\mathbb{R})$ telle que $B = PB'P^{-1}$.

On peut aller un peu plus vite en utilisant l'exercice 2.15. Prenons une base quelconque du plan $\mathrm{Ker}((B-uI_3)^2+v^2I_3)$ *et notons* C *la matrice de la restriction de* B *à ce plan dans cette base. La matrice* C *est semblable sur* $\mathbb{C}$ *à la matrice* $\begin{pmatrix} u+iv & 0 \\ 0 & u-iv \end{pmatrix}$. *Mais il en est de même de la matrice* $\begin{pmatrix} u & -v \\ v & u \end{pmatrix}$ *car ses valeurs propres sont* $u \pm iv$. *Donc* C *est semblable sur* $\mathbb{C}$ *et donc aussi sur* $\mathbb{R}$, *à cette dernière matrice. Le fait que* B *est semblable à* B' *en découle de suite.*

Comme a admet une racine cubique, il nous suffit de regarder si la matrice $\begin{pmatrix} u & -v \\ v & u \end{pmatrix}$ est un cube de $\mathcal{M}_2(\mathbb{R})$. Or, si on identifie $\mathbb{C}$ et $\mathbb{R}^2$, l'endomorphisme de $\mathbb{R}^2$ de matrice $\begin{pmatrix} u & -v \\ v & u \end{pmatrix}$ s'identifie à la similitude $\mathbb{R}$-linéaire de $\mathbb{C}$: $z \longmapsto (u+iv)z$. Il est alors clair que si on se donne s et t deux réels tels que $(s+it)^3 = (u+iv)$, on a

$$\begin{pmatrix} u & -v \\ v & u \end{pmatrix} = \begin{pmatrix} s & -t \\ t & s \end{pmatrix}^3.$$

On a alors $B' = \Phi(A')$ où $A' = \begin{pmatrix} \sqrt[3]{a} & 0 & 0 \\ 0 & s & -t \\ 0 & t & s \end{pmatrix}$ et donc $B \in \operatorname{Im}\Phi$.

• Il nous reste à regarder le cas où χ est scindé sur $\mathbb{R}$. S'il admet 3 racines distinctes alors B est diagonalisable et cela a déjà été traité. Il reste donc deux cas à regarder : A admet une valeur propre simple a et une valeur propre double b en étant non diagonalisable, ou A admet une valeur propre triple b sans être diagonalisable (c'est-à-dire sans être l'homothétie bI_3). Commençons par le premier de ces deux cas. On a alors $\mathbb{R}^3 = \operatorname{Ker}(B-aI_3) \oplus \operatorname{Ker}(B-bI_3)^2$ et $\operatorname{Ker}(B-bI_3)^2 \neq \operatorname{Ker}(B-bI_3)$ car la dimension de l'espace propre associé à b est égale à 1. Prenons $X_1 \in \operatorname{Ker}(B-aI_3)$, $X_3 \in \operatorname{Ker}(B-bI_3)^2 \setminus \operatorname{Ker}(B-bI_3)$ et $X_2 = (B-bI_3)X_3$. Dans la base (X_1, X_2, X_3), l'endomorphisme canoniquement associé à B a pour matrice $B' = \begin{pmatrix} a & 0 & 0 \\ 0 & b & 1 \\ 0 & 0 & b \end{pmatrix}$. Il nous suffit de regarder si B' est un cube ou non. Si $A^3 = B'$, alors A et B' commutent de sorte que A stabilise forcément les sous-espaces $\operatorname{Ker}(B-aI_3)$ et $\operatorname{Ker}(B-bI_3)^2$. Elle doit donc être de la forme $A = \begin{pmatrix} \alpha & 0 & 0 \\ 0 & \beta & \gamma \\ 0 & \delta & \varepsilon \end{pmatrix}$. Comme précédemment on doit prendre $\alpha = \sqrt[3]{a}$ et le problème se ramène finalement à savoir si la matrice $\begin{pmatrix} b & 1 \\ 0 & b \end{pmatrix}$ est un cube de $\mathcal{M}_2(\mathbb{R})$. Essayons d'abord de chercher une solution triangulaire supérieure de la forme $\begin{pmatrix} \beta & \gamma \\ 0 & \beta \end{pmatrix}$. On est conduit aux relations $\beta^3 = b$ et $3\gamma\beta^2 = 1$. Si b n'est pas nul on a une solution obtenue en prenant $\beta = \sqrt[3]{b}$ et $\gamma = \frac{1}{3\beta^2}$. Montrons que dans le cas $b = 0$ il n'y a pas de solution. Supposons qu'il existe $M \in \mathcal{M}_2(\mathbb{R})$ telle que $M^3 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = J$. On a $J^2 = 0$ (J est nilpotente) donc

$M^6 = 0$ et M est forcément nilpotente. Mais son indice de nilpotence est forcément $\leqslant 2$ donc $M^2 = 0$ et par suite $M^3 = 0$: contradiction. Donc J n'est pas un cube.

• Terminons par le dernier cas : B admet une valeur propre triple b sans être diagonalisable. Quitte à remplacer B par une matrice semblable, on peut supposer que $B = bI_3 + N$ où N est triangulaire supérieure de diagonale nulle (N est nilpotente et non nulle). Comme avant on peut d'abord chercher une solution triangulaire supérieure de la forme $A = \beta I_3 + cN + dN^2$, où β est la racine cubique réelle de b. Un calcul facile donne $A^3 = bI_3 + 3\beta^2 cN + 3(\beta c^2 + \beta^2 d)N^2$. Si $b \neq 0$ alors $\beta \neq 0$ et pour avoir $A^3 = B$ il suffit de prendre $c = \frac{1}{3\beta^2}$ et $d = -\frac{c}{\beta}$. Montrons pour finir que lorsque $b = 0$, $B = N$ n'est pas un cube. C'est le même argument qu'avant. Si $N = M^3$, alors M est nilpotente et donc $M^3 = N = 0$: la seule matrice nilpotente de $\mathcal{M}_3(\mathbb{R})$ qui est un cube est la matrice nulle.

Conclusion. On constate que les seules matrices $B \in \mathcal{M}_3(\mathbb{R})$ qui ne sont pas des cubes sont celles qui admettent 0 pour valeur propre double ou triple et pour lesquelles la dimension de $\operatorname{Ker} B$ n'est pas égale à la multiplicité de la valeur propre. Une autre manière de dire la même chose : une matrice B est dans l'image de Φ si et seulement si $\operatorname{Ker} B = \operatorname{Ker} B^2$. ◁

Il est classique de montrer que la condition $\operatorname{Ker} B = \operatorname{Ker} B^2$ *équivaut aussi à* $\operatorname{Im} B = \operatorname{Im} B^2$ *ou encore à* $\mathbb{R}^3 = \operatorname{Im} B \oplus \operatorname{Ker} B$ *(le lecteur se reportera à l'exercice 6.15 du tome 1 d'algèbre).*

Nous allons terminer ce chapitre par des exercices de nature topologique. Il est remarquable de pouvoir lire certaines propriétés algébriques d'une matrice complexe sur des caractéristiques topologiques de sa classe de similitude. Le premier énoncé montre que les seules classes de similitude bornées sont celles des matrices scalaires (ce sont donc des singletons).

2.53. Classes de similitude bornées

Déterminer les matrices A à coefficients complexes dont la classe de similitude est bornée.

(**École polytechnique**)

▷ **Solution.**

Il est clair que si la matrice A est scalaire, sa classe de similitude ne contient qu'elle-même, donc est bornée.

Montrons que les matrices scalaires sont les seules matrices ayant cette propriété. Toutes les normes sur $\mathcal{M}_n(\mathbb{C})$ étant équivalentes, on va travailler avec la norme infinie. Soit $A \in \mathcal{M}_n(\mathbb{C})$ et f l'endomorphisme de $\mathbb{C}^n$ canoniquement associé à A. Si A n'est pas scalaire, f n'est pas une homothétie vectorielle. C'est un résultat classique qu'il existe alors un vecteur $u \in \mathbb{C}^n$ tel que la famille $(u, f(u))$ soit libre. Soit $\lambda \in \mathbb{C}^*$ et $\mathcal{B} = (e_1, \ldots, e_n)$ une base de $\mathbb{C}^n$ telle que $e_1 = \lambda u$, $e_2 = f(u)$. On a alors $f(e_1) = \lambda e_2$. La matrice B de f dans la base $\mathcal{B}$ est semblable à A. Sa première colonne est $\begin{pmatrix} 0 \\ \lambda \\ 0 \\ \vdots \\ 0 \end{pmatrix}$. On a donc $\|B\|_\infty \geqslant |\lambda|$. Comme λ est quelconque, la classe de similitude de A n'est pas bornée. $\lhd$

L'énoncé suivant montre que les matrices diagonalisables complexes se caractérisent par le fait que leur classe de similitude est fermée.

2.54. Classe de similitude d'une matrice diagonalisable

Si $A \in \mathcal{M}_n(\mathbb{C})$ on note S(A) la classe de similitude de A.

1. Montrer que si A est inversible, l'adhérence de S(A) est incluse dans $GL_n(\mathbb{C})$.

2. Montrer que A est diagonalisable si et seulement si S(A) est fermée.

(ENS Ulm)

$\triangleright$ **Solution.**

1. Soit $A \in GL_n(\mathbb{C})$ et (A_p) une suite d'éléments de S(A) qui converge vers une matrice B. On a alors $\det A_p = \det A$ pour tout $p \in \mathbb{N}$. Le déterminant est une application continue, donc $\det B = \lim_{p \to +\infty} \det A_p = \det A \neq 0$ et $B \in GL_n(\mathbb{C})$.

2. • On suppose que A est diagonalisable. On se donne encore (A_p) une suite de S(A) qui converge vers une matrice B. Comme A_p est semblable à A, on a $\chi_{A_p} = \chi_A$. Les coefficients du polynôme caractéristique sont des fonctions continues. On en déduit que $\chi_B = \lim_{p \to +\infty} \chi_{A_p} = \chi_A$. Le polynôme minimal μ_A de A est à racines simples car A est diagonalisable. Puisque A_p est semblable à A, on a $\mu_A(A_p) = 0$ pour tout p. Par passage à la limite, on en déduit que $\mu_A(B) = 0$. La matrice B possède un polynôme annulateur à racines simples, donc est diagonalisable. Les

matrices A et B sont diagonalisables et ont même polynôme caractéristique (donc mêmes valeurs propres comptées avec multiplicité) ; elles sont semblables. Cela montre que $B \in S(A)$ et donc que $S(A)$ est fermée.

• Montrons la réciproque. Soit $A = (a_{ij}) \in \mathcal{M}_n(\mathbb{C})$ telle que $S(A)$ soit fermée. Quitte à remplacer A par une matrice semblable, on peut supposer que A est triangulaire supérieure. Notons $\mathcal{B} = (e_1, \ldots, e_n)$ la base canonique de $\mathbb{C}^n$. Pour $p \in \mathbb{N}^*$, écrivons la matrice A_p de l'endomorphisme u canoniquement associé à A dans la base $(\frac{1}{p}e_1, \frac{1}{p^2}e_2, \ldots, \frac{1}{p^n}e_n)$. On a, pour tout $j \in [\![1, n]\!]$,

$$u\Big(\frac{1}{p^j}e_j\Big) = \frac{1}{p^j}u(e_j) = \frac{1}{p^j}\sum_{i=1}^{j} a_{i,j}e_i = \sum_{i=1}^{j} p^{i-j}a_{i,j}\frac{1}{p^i}e_i.$$

Le coefficient (i, j) de A_p est donc égal à $p^{i-j}a_{ij}$ si $i \leqslant j$ et est nul sinon. Autrement dit, A_p est aussi triangulaire supérieure et a la même diagonale que A, mais tous les coefficients strictement au-dessus de la diagonale sont divisés par une puissance de p. Il en découle que A_p converge vers la matrice diagonale $D = \mathrm{Diag}(a_{11}, \ldots, a_{nn})$ lorsque p tend vers $+\infty$. Or, A_p est semblable à A pour tout p. Comme $S(A)$ est fermée, D est semblable à A qui est donc diagonalisable. ◁

Notons qu'une classe de similitude ne peut pas être ouverte. Voici un argument simple : la trace étant constante sur une classe de similitude, celle-ci est incluse dans un hyperplan affine de $\mathcal{M}_n(\mathbb{C})$.

Dans l'exercice ci-après on montre que les matrices nilpotentes sont les matrices dont la classe de similitude est adhérente à la matrice nulle.

2.55. Classe de similitude d'une matrice nilpotente

Soit $n \in \mathbb{N}^*$ et $A \in \mathcal{M}_n(\mathbb{C})$. Montrer que A est nilpotente si et seulement si il existe une suite de matrices semblables à A qui converge vers la matrice nulle.

(École polytechnique)

▷ **Solution.**

Supposons qu'une suite $(A_p)_{p\in\mathbb{N}}$ composée de matrices semblables à A converge vers 0. Alors, comme le polynôme caractéristique de M est une fonction continue de la matrice M, χ_{A_p} converge dans $\mathbb{R}_n[X]$ vers $\chi_0 = X^n$. Or la suite χ_{A_p} est constante et égale à χ_A. On en déduit

que $\chi_A = X^n$ et $A^n = 0$ en vertu du théorème de Cayley-Hamilton : la matrice A est nilpotente.

Réciproquement, supposons A nilpotente. Identifions A et l'endomorphisme de K^n canoniquement associé à A. Il existe une base $(e_1, \ldots e_n)$ telle que la matrice de A dans cette base soit triangulaire supérieure à diagonale nulle. Notons $A_0 = (a_{ij})_{1 \leqslant i,j \leqslant n}$ la matrice de A dans cette base. Comme dans l'exercice 2.54, pour $p \geqslant 1$, on considère la base $\mathcal{B}_p = (e_1/p, e_2/p^2, \ldots, e_n/p^n)$. On a

$$A(e_k) = a_{k-1,k}e_{k-1} + a_{k-2,k}e_{k-2} + \cdots + a_{1,k}e_1,$$

et donc

$$A\left(\frac{e_k}{p^k}\right) = \frac{a_{k-1,k}}{p}\frac{e_{k-1}}{p^{k-1}} + \frac{a_{k-2,k}}{p^2}\frac{e_{k-2}}{p^{k-2}} + \cdots + \frac{a_{1,k}}{p^{k-1}}\frac{e_1}{p}.$$

Notons A_p la matrice de A dans la base $\mathcal{B}_p$, c'est une matrice semblable à A. La suite de matrices A_p converge bien vers 0 car chaque coefficient tend vers 0 quand p tend vers l'infini. ◁

2.56. Adhérence de l'ensemble des racines de l'identité

Trouver l'adhérence de $\mathcal{A} = \{M \in \mathcal{M}_n(\mathbb{C}),\ \exists p \in \mathbb{N}^*,\ M^p = I_n\}$.
(École polytechnique)

▷ **Solution.**

Si $M \in \mathcal{A}$ et vérifie $M^p = I_n$, les valeurs propres de M sont des racines p-ièmes de l'unité. De plus M est diagonalisable car le polynôme $X^p - 1$ annule A et est scindé à racines simples. On remarque que réciproquement, si M est une matrice diagonalisable de $\mathcal{M}_n(\mathbb{C})$ dont les valeurs propres sont toutes des racines p-ièmes de 1, on a $M^p = I_n$ et M appartient à $\mathcal{A}$.

Regardons le cas $n = 1$ qui va nous donner une idée du résultat général. En identifiant $\mathcal{M}_1(\mathbb{C})$ avec $\mathbb{C}$, $\mathcal{A}$ est l'ensemble des racines de l'unité. Son adhérence est alors le cercle unité S^1 de $\mathbb{C}$. En effet, $\mathcal{A}$ est l'image de $\pi\mathbb{Q}$ par le morphisme surjectif $f : \mathbb{R} \to S^1$ défini par $f(\theta) = e^{i\theta}$. Comme f est continu, et comme $\pi\mathbb{Q}$ est dense dans $\mathbb{R}$, on a $\overline{\mathcal{A}} = f(\mathbb{R}) = S^1$. Revenons au cas général. Nous allons montrer que l'adhérence de $\mathcal{A}$ est exactement l'ensemble F des matrices de $\mathcal{M}_n(\mathbb{C})$ dont les valeurs propres sont toutes de module 1. On a évidemment $\mathcal{A} \subset F$. Montrons que $F \subset \overline{\mathcal{A}}$. Soit $M \in F$. La matrice M peut s'écrire $P^{-1}TP$, avec P inversible et T triangulaire supérieure. Comme $\mathcal{A}$ est stable par conjugaison, il suffit

de prouver que T est dans l'adhérence de $\mathcal{A}$: si T_p est une suite de $\mathcal{A}$ qui converge vers T alors la suite $M_p = P^{-1}T_pP$ converge vers M et les matrices M_p sont toutes dans $\mathcal{A}$. Les termes diagonaux de $T = (t_{ij})$ sont de module 1. D'après le cas $n = 1$, pour tout $k \in [\![1, n]\!]$, on peut trouver une suite $t_{kk}^{(p)}$ de racines de l'unité qui converge vers t_{kk}. On considère alors la matrice T_p dont les termes non diagonaux sont les termes de T et les termes diagonaux sont les $t_{kk}^{(p)}$. La suite (T_p) converge vers T. Pour que les matrices T_p soient dans $\mathcal{A}$, il suffit de faire en sorte que les coefficients diagonaux soient deux à deux distincts, car dans ce cas T_p est diagonalisable et l'une de ses puissances vaut I_n. Or cela n'est pas très difficile à réaliser. Pour $1 \leqslant k \leqslant n$, considérons $\theta_k \in \mathbb{R}$ tel que $t_{kk} = e^{i\theta_k}$. Pour $p \in \mathbb{N}^*$, on prend $t_{kk}^{(p)} = e^{2i\pi\mu_k^{(p)}}$ avec $\mu_k^{(p)} = \dfrac{k}{p} + \dfrac{1}{p}\mathrm{E}\left(\dfrac{p\theta_k}{2\pi}\right)$. On a

$$|e^{2i\pi\mu_k^{(p)}} - e^{i\theta_k}| \leqslant |2\pi\mu_k^{(p)} - \theta_k| \leqslant \frac{2\pi k}{p} + \frac{2\pi}{p}$$

ce qui montre que $t_{kk}^{(p)}$ converge vers t_{kk} pour tout k. Vérifions que les termes diagonaux de T_p sont deux à deux distincts pour p assez grand. Soit k et l deux entiers distincts dans $[\![1, n]\!]$.

- Si $\theta_k \equiv \theta_l\ [2\pi]$, alors $2\pi\mu_k - 2\pi\mu_l \equiv \dfrac{2\pi(k-l)}{p}\ [2\pi]$.
- Si $\theta_k \not\equiv \theta_l\ [2\pi]$, alors $2\pi\mu_k - 2\pi\mu_l \underset{p\to\infty}{\sim} \theta_k - \theta_l$.

Ce qui prouve le résultat.

Pour montrer que F est l'adhérence de $\mathcal{A}$, il suffit de vérifier que F est fermé. Soit (M_p) une suite de matrices de F qui converge vers $M \in \mathcal{M}_n(\mathbb{C})$. On note $\lambda_{1,p}, \ldots, \lambda_{n,p}$ les valeurs propres de M_p (comptées avec multiplicité et rangées dans un ordre quelconque). Les coefficients du polynôme caractéristique étant des fonctions polynomiales des coefficients de la matrice, on a $\chi_M = \lim\limits_{p\to+\infty} \chi_{M_p}$ (*i.e.* les coefficients de χ_M sont les limites des coefficients de χ_{M_p}). D'autre part, la suite $(\lambda_{1,p}, \ldots, \lambda_{n,p})_{p\in\mathbb{N}}$ de $\mathbb{C}^n$ étant bornée, on peut en extraire une suite convergente. Soit $\varphi : \mathbb{N} \longrightarrow \mathbb{N}$ une telle extraction et

$$(\lambda_1, \ldots, \lambda_n) = \lim_{p\to+\infty} (\lambda_{1,\varphi(p)}, \ldots, \lambda_{n,\varphi(p)}).$$

Pour tout $p \in \mathbb{N}$, on a $\chi_{M_{\varphi(p)}} = \prod\limits_{i=1}^{n} (X - \lambda_{i,\varphi(p)})$. On en déduit par passage à la limite que $\chi_M = \prod\limits_{i=1}^{n} (X - \lambda_i)$. Les valeurs propres de M sont donc $\lambda_1, \ldots, \lambda_n$. Comme, pour tout $i \in [\![1, n]\!]$, $|\lambda_i| = \lim\limits_{p\to+\infty} |\lambda_{i,\varphi(p)}| = 1$,

les valeurs propres de M sont de module 1 et M appartient à F qui est fermé. ◁

Il est facile de généraliser le résultat de cet exercice. Si X *est une partie de* $\mathbb{C}$*, l'adhérence de l'ensemble des matrices* $A \in \mathcal{M}_n(\mathbb{C})$ *telles que* $\mathrm{Sp}\,A \subset X$ *est l'ensemble des matrices* A *telles que* $\mathrm{Sp}\,A \subset \overline{X}$.

Dans l'énoncé suivant on prouve que l'identité est isolée dans l'ensemble des racines q-ièmes de l'identité (par exemple dans l'ensemble des symétries pour $q = 2$*).*

2.57. Point isolé dans l'ensemble des racines q-ièmes de l'identité

Soit $q \in \mathbb{N}^*$. On note E_q l'ensemble des $A \in \mathrm{GL}_n(\mathbb{C})$ telles que $A^q = I_n$.

1. Que dire de $A \in E_q$ si 1 est la seule valeur propre de A ?

2. Montrer que I_n est un point isolé de E_q.

3. Soit $A_0 \in E_q$. Montrer l'existence de $\varepsilon > 0$ tel que si $A \in E_q$ et $\|A - A_0\| < \varepsilon$, alors A et A_0 sont semblables.

(ENS Lyon)

▷ **Solution.**

1. Toute matrice A de E_q est annulée par $X^q - 1$ qui est scindé à racines simples, donc est diagonalisable. Si 1 est sa seule valeur propre, la matrice A est semblable, donc égale, à la matrice identité I_n.

2. On munit $\mathcal{M}_n(\mathbb{C})$ d'une norme triple associée à une norme quelconque de $\mathbb{C}^n$. Si $A \in \mathcal{M}_n(\mathbb{C})$ et $\lambda \in \mathrm{Sp}(A)$, de vecteur propre associé X, on a $\|AX\| = |\lambda|\|X\| \leqslant |||A|||X\|$ et donc $|\lambda| \leqslant |||A|||$.

Soit alors $A \in E_q$ distincte de I_n. D'après la question 1. A admet une valeur propre λ différente de 1. Cette valeur propre λ est une racine q-ième de l'unité. Il existe donc $k \in \mathbb{N}$, $0 < |k| \leqslant \frac{q}{2}$ tel que $\lambda = e^{\frac{2ik\pi}{q}}$. On a alors $|\lambda - 1| = 2\left|\sin\frac{k\pi}{q}\right| \geqslant 2\sin\frac{\pi}{q}$ car la fonction sinus croît sur $\left[0, \frac{\pi}{2}\right]$. Comme $\lambda - 1$ est une valeur propre de $A - I_n$ on a

$$|||A - I_n||| \geqslant |\lambda - 1| \geqslant 2\sin\frac{\pi}{q}.$$

Autrement dit la boule ouverte de centre I_n et de rayon $2\sin\frac{\pi}{q}$ ne contient pas d'autre élément de E_q en dehors de l'identité. Cela montre que I_n est un point isolé de E_q.

3. Raisonnons par l'absurde et supposons que pour tout $k \in \mathbb{N}^*$, il existe $M_k \in E_q$ tel que $\|A - M_k\| < \frac{1}{k}$ sans que M_k et A soient semblables.

Comme $(M_k)_{k\geqslant 0}$ converge vers A, χ_{M_k} converge vers χ_A (dans $\mathbb{C}_n[X]$) puisque les coefficients de χ_M sont des fonctions continues de M. Or, l'ensemble des polynômes caractéristiques des matrices de E_q est fini : en effet, le spectre de ces matrices est contenu dans l'ensemble des racines q-ièmes de l'unité. Il s'ensuit que la suite χ_{M_k} est constante à partir d'un rang k_0 et égale à χ_A. Dans ces conditions, M_{k_0} est semblable à A car ces matrices sont diagonalisables et comme elles ont même polynôme caractéristique, elles ont même spectre et même ordre pour les valeurs propres. C'est la contradiction cherchée. ◁

Le résultat prouvé dans l'exercice suivant est important. Il montre que l'ensemble des matrices diagonalisables de $\mathcal{M}_n(\mathbb{C})$ est dense. Ainsi, lorsqu'on cherche à prouver une propriété qui dépend continûment d'une matrice $A \in \mathcal{M}_n(\mathbb{C})$, on peut se contenter de la prouver lorsque A est diagonalisable, ce qui est souvent beaucoup plus facile, et l'étendre par un argument de densité. On en verra un exemple d'application dans l'exercice 2.61.

2.58. Adhérence et intérieur

1. Trouver l'adhérence et l'intérieur de l'ensemble des matrices diagonalisables de $\mathcal{M}_n(\mathbb{C})$.

2. Trouver l'adhérence et l'intérieur de l'ensemble des matrices cycliques de $\mathcal{M}_n(\mathbb{C})$. On rappelle que A est dite cyclique si $\mu_A = \chi_A$, ce qui équivaut à l'existence d'un vecteur X tel que $\mathbb{C}^n = \mathrm{Vect}(X, AX, \ldots, A^{n-1}X)$ (se reporter à l'exercice 2.38).

(ENS Ulm)

▷ **Solution.**

1. Notons $\mathcal{D}$ l'ensemble des matrices diagonalisables de $\mathcal{M}_n(\mathbb{C})$. On va prouver que $\mathcal{D}$ est dense dans $\mathcal{M}_n(\mathbb{C})$, cet espace étant muni de sa topologie d'espace vectoriel normé. Pour cela, on utilise le fait que toute matrice A de $\mathcal{M}_n(\mathbb{C})$ est semblable à une matrice triangulaire T. Écrivons $A = PTP^{-1}$ avec T triangulaire supérieure et P inversible. Pour $n \in \mathbb{N}^*$, on considère la matrice $T_n = (t_{ij}^{(n)})$ dont les coefficients sont égaux à ceux de T, sauf les coefficients diagonaux qu'on définit par $t_{ii}^{(n)} = t_{ii} + \frac{i}{n}$. Pour i et j indices distincts, on a $t_{ii}^{(n)} \neq t_{jj}^{(n)}$ si $t_{ii} = t_{jj}$ et sinon $\lim\limits_{n\to+\infty} t_{ii}^{(n)} - t_{jj}^{(n)} = t_{ii} - t_{jj} \neq 0$. Pour n assez grand, tous les termes diagonaux de T_n sont distincts deux à deux. Autrement dit, à partir d'un certain rang T_n possède n valeurs propres distinctes et est

donc diagonalisable. Il en est de même de PT_nP^{-1}. Comme la suite (PT_nP^{-1}) converge vers A, on a le résultat voulu.

Notons que ce résultat ne reste pas vrai dans $\mathcal{M}_n(\mathbb{R})$. La matrice $A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ *n'est pas limite d'une suite (A_n) de matrices diagonalisables de $\mathcal{M}_2(\mathbb{R})$. Dans le cas contraire, son polynôme caractéristique χ_A serait limite de la suite χ_{A_n}. Pour tout $n \in \mathbb{N}$, le discriminant de χ_{A_n} est positif. Par passage à la limite, le discriminant de χ_A serait positif. C'est faux puisqu'il vaut -4.*

2. Soit A une matrice cyclique et X un vecteur tel que $\mathbb{C}^n = \mathrm{Vect}(X, AX, \ldots, A^{n-1}X)$. La famille $(X, AX, \ldots, A^{n-1}X)$ est alors une base de $\mathbb{C}^n$. Notons B la base canonique de $\mathbb{C}^n$. La fonction f qui à $M \in \mathcal{M}_n(\mathbb{C})$ associe $f(M) = \det_B(X, MX, \ldots, M^{n-1}X)$ est continue car polynomiale en les coefficients de M. Par hypothèse elle prend une valeur non nulle en A. Donc elle ne s'annule pas sur tout un voisinage V de A et toute matrice M de V est alors cyclique. Autrement dit, l'ensemble des matrices cycliques de $\mathcal{M}_n(\mathbb{C})$ est ouvert. On va voir que cet ouvert est dense. Cherchons d'abord des exemples de matrices cycliques simples. Une matrice diagonalisable n'est pas forcément cyclique : par exemple I_n n'est pas cyclique (pour $n \geqslant 2$). En fait, une matrice diagonalisable dont le spectre est de cardinal n est cyclique. En effet, notons $\lambda_1, \ldots, \lambda_n$ les valeurs propres distinctes d'une telle matrice A et $X_1, \ldots, X_n$ une base de vecteurs propres associés à ces valeurs propres. Considérons le vecteur $X = X_1 + X_2 + \cdots + X_n$. La matrice de la famille $(X, AX, \ldots, A^{n-1}X)$ dans la base $(X_1, \ldots, X_n)$ est une matrice de Vandermonde : elle est inversible et la famille $(X, AX, \ldots, A^{n-1}X)$ est une base de $\mathbb{C}^n$. Il est clair que toute matrice diagonale est limite d'une suite de matrices diagonales à coefficients diagonaux deux à deux distincts. Il en découle que toute matrice diagonalisable est limite d'une suite de matrices diagonalisables dont le spectre est de cardinal n. L'adhérence de l'ensemble des matrices cycliques contient donc l'ensemble $\mathcal{D}$ des matrices diagonalisables. Or celui-ci est dense dans $\mathcal{M}_n(\mathbb{C})$ d'après la question précédente. ◁

On revient maintenant sur la localisation du spectre d'une matrice complexe et notamment sur le plus grand module des valeurs propres qu'on appelle le rayon spectral. Il intervient notamment lorsqu'on étudie la convergence de la suite des puissances d'une matrice.

2.59. Rayon spectral

Pour $A \in \mathcal{M}_n(\mathbb{C})$, on note $\rho(A) = \max_{\alpha \in \mathrm{Sp}(A)} |\alpha|$ le *rayon spectral de* A.

1. Montrer que $\sum A^k$ converge si et seulement si $\rho(A) < 1$.

2. Soit A, B, C dans $\mathcal{M}_n(\mathbb{C})$. On suppose que $\rho(A)\rho(B) < 1$. Montrer qu'il existe une unique matrice $D \in \mathcal{M}_n(\mathbb{C})$ telle que

$$ADB - D = C.$$

(**École polytechnique**)

▷ **Solution.**

1. Supposons $\rho(A) \geqslant 1$ et soit $\lambda \in \mathrm{Sp}(A)$ une valeur propre telle que $|\lambda| \geqslant 1$. Soit $e \in \mathbb{C}^n$ un vecteur propre de A pour λ. Pour tout entier k, on a alors $A^k e = \lambda^k e$ et comme $|\lambda| \geqslant 1$, $A^k e$ ne tend pas vers 0 (car en ayant choisi une norme quelconque, $\|A^k e\| = |\lambda|^k \|e\| \geqslant \|e\|$). Ainsi, $(A^k)_{k \geqslant 0}$ ne converge pas vers 0 et nécessairement la série $\sum A^k$ diverge.

Supposons maintenant que $\rho(A) < 1$: les valeurs propres de A sont de module strictement inférieur à 1. En particulier, 1 n'est pas valeur propre et $B = I - A$ est inversible. On a pour $p \geqslant 0$,

$$B(I + A + A^2 + \cdots + A^p) = I - A^{p+1}$$

et donc

$$I + A + A^2 + \cdots + A^p = B^{-1}(I - A^{p+1}).$$

Nous allons démontrer que $\lim_{p \to +\infty} A^p = 0$ ce qui nous garantira la convergence de la série et donnera pour somme $B^{-1} = (I - A)^{-1}$. Pour cela nous allons utiliser la décomposition de Dunford (cf. exercice 2.30). On peut écrire $A = D + N$ dans $\mathcal{M}_n(\mathbb{C})$ avec D diagonalisable, N nilpotente et $ND = DN$. Soit $(e_1, \ldots, e_n)$ une base de vecteurs propres pour D. On considère la norme hermitienne de $\mathbb{C}^n$ qui fait de cette base une base orthonormale, autrement dit, si $X = x_1 e_1 + \cdots + x_n e_n$, $\|X\|^2 = |x_1|^2 + \cdots + |x_n|^2$. Dans ces conditions, en considérant la triple norme de $\mathcal{M}_n(\mathbb{C})$ associée à la norme que nous venons de définir, vérifions que la norme de D vaut $\rho(A) < 1$: si $\|X\| \leqslant 1$,

$$\begin{aligned}\|DX\| &= \|\lambda_1 x_1 e_1 + \cdots + \lambda_n x_n e_n\| = \left(|\lambda_1 x_1|^2 + \cdots + |\lambda_n x_n|^2\right)^{1/2} \\ &\leqslant \left(\max_{1 \leqslant i \leqslant n} |\lambda_i|^2\right)^{1/2} \|X\| = \rho(A)\|X\| \leqslant \rho(A),\end{aligned}$$

$\lambda_1, \ldots, \lambda_n$ désignant les valeurs propres respectivement associées à $e_1, \ldots e_n$. Si λ_{i_0} est une valeur propre de module maximal égal à $\rho(A)$, $\|De_{i_0}\| = \|\lambda_{i_0} e_{i_0}\| = \rho(A)$. Ainsi, il vient $|||D||| = \rho(A) < 1$. Les matrices D et N commutant, la formule du binôme de Newton permet d'écrire pour $p \geqslant n$

$$\begin{aligned} A^p &= (D+N)^p \\ &= \sum_{k=0}^{p} C_p^k D^{p-k} N^k = D^p + C_p^1 D^{p-1} N + \cdots + C_p^{n-1} D^{p-n+1} N^{n-1}, \end{aligned}$$

compte tenu du fait que $N^n = 0$. Ainsi, A^p apparaît comme la somme de n termes : $C_p^k D^{p-k} N^k$ pour $0 \leqslant k \leqslant n-1$. Il suffit donc de montrer que chacun de ces termes tend vers 0. Or, comme $C_p^k = \dfrac{p \ldots (p-k+1)}{k!} \leqslant \dfrac{p^k}{k!}$, on a pour k fixé dans $[\![0, n-1]\!]$

$$|||C_p^k D^{p-k} N^k||| \leqslant C_p^k |||D|||^{p-k} |||N|||^k \leqslant \frac{|||N|||^k}{k!} p^k |||D|||^{p-k} \xrightarrow[p \to \infty]{} 0,$$

car la suite géométrique l'emporte sur la puissance. On en déduit que A^p converge vers 0 lors que p tend vers l'infini.

Conclusion. $\sum A^k$ converge si, et seulement si $\rho(A) < 1$ et dans ce cas,

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

Comme pour tout $0 \leqslant k \leqslant n-1$, $|||C_p^k D^{p-k} N^k||| = O\big(p^k \rho(A)^p\big)$, on obtient $|||A^p||| = O\big(p^{n-1}\rho(A)^p\big)$. Ce résultat est d'ailleurs valable quelle que soit la valeur du rayon spectral.

2. Il suffit de prouver que l'application linéaire

$$\begin{array}{rccc} \Phi : & \mathcal{M}_n(\mathbb{C}) & \longrightarrow & \mathcal{M}_n(\mathbb{C}) \\ & D & \longmapsto & ADB - D \end{array}$$

est bijective. Comme nous sommes en dimension finie, il suffit de prouver que Φ est injective. Prenons D dans le noyau de Φ. Alors $ADB = D$ et par une récurrence immédiate, on obtient que pour tout $k \in \mathbb{N}$, $A^k D B^k = D$. En considérant la même norme subordonnée que précédemment, et en tenant compte de la remarque faite à la fin de la première question, on peut majorer ainsi

$$\begin{aligned} |||D||| = |||A^k D B^k||| &\leqslant |||A^k|||.|||D|||.|||B^k||| = O\big(k^{2n-2}\rho(A)^k \rho(B)^k\big) \\ &\leqslant O\big(k^{2n-2}(\rho(A)\rho(B))^k\big) \xrightarrow[k \to +\infty]{} 0, \end{aligned}$$

puisque $\rho(A)\rho(B) < 1$. Il en résulte que $|||D||| = 0$ et D est nulle. ◁

2.60. Suite des puissances bornée

Déterminer l'ensemble des matrices $A \in \mathcal{M}_n(\mathbb{C})$ telles que la suite $(A^k)_{k\geqslant 0}$ soit bornée.

(ENS Ulm)

▷ **Solution.**

On note $\rho(A) = \max_{\lambda\in \mathrm{Sp}\, A} |\lambda|$ le rayon spectral de A. Lorsque $n = 1$ on est ramené à une simple suite géométrique de $\mathbb{C}$ et il est bien connu que, pour $a \in \mathbb{C}$, la suite a^k est bornée si et seulement si $|a| \leqslant 1$. La position du rayon spectral par rapport à 1 joue donc un rôle important.

• Si $\rho(A) < 1$, la suite A^k converge vers 0 (voir l'exercice précédent) et en particulier les puissances de A sont bornées.

• Supposons $\rho(A) > 1$. Alors A possède une valeur propre $\lambda \in \mathbb{C}$ de module > 1. Si X est un vecteur propre associé, on a

$$\|A^k X\| = |\lambda|^k \|X\| \xrightarrow[k\to+\infty]{} +\infty,$$

et la suite $(\Lambda^k)_{k\in\mathbb{N}}$ ne saurait être bornée.

• Il reste à étudier le cas où $\rho(\Lambda) = 1$. Dans ce cas la suite (A^k) n'est pas forcément bornée : c'est le cas par exemple avec $A = I_n$ mais ce n'est pas le cas par exemple pour $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ car $A^k = \begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix}$ pour tout $k \in \mathbb{N}$. Commençons à nous ramener au cas où il n'y a qu'une seule valeur propre en utilisant la décomposition en sous-espaces caractéristiques. L'espace $\mathbb{C}^n$ est somme directe des sous-espaces caractéristiques $F_1, \ldots, F_r$ respectivement relatifs aux valeurs propres deux à deux distinctes $\lambda_1, \ldots, \lambda_r$. Si $\mathcal{B}$ est une base obtenue par recollement de bases de $F_1, \ldots, F_r$, la matrice de l'endomorphisme associé à A dans cette nouvelle base est diagonale par blocs, de la forme

$$\begin{pmatrix} \boxed{A_1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \boxed{A_r} \end{pmatrix}.$$

Les puissances de A sont bornées si et seulement si toutes les suites $(A_i^k)_{k\in\mathbb{N}}$ sont bornées (pour $1 \leqslant i \leqslant r$). Comme on l'a déjà dit, lorsque $|\lambda_i| < 1$ la suite $(A_i^k)_{k\in\mathbb{N}}$ est bornée. Regardons ce qui se passe lorsque $|\lambda_i| = 1$. En remarquant que $(A_i^k)_{k\in\mathbb{N}}$ borné équivaut à ce que les $(\overline{\lambda_i} A_i)^k$ soient bornées, on peut supposer $\lambda_i = 1$. Autrement dit, A_i est une matrice unipotente que l'on peut écrire $A_i = I + N$ avec I matrice identité

et N nilpotente. Soit p le plus petit entier naturel tel que $N^p = 0$. Nous savons alors que $(I, N, \ldots, N^{p-1})$ est une famille libre que nous pouvons compléter en une base $\mathcal{M}$ de l'espace des matrices complexes de même taille que A_i. Pour $k \geqslant p$, la formule du binôme donne

$$A_i^k = (I + N)^k = I + kN + C_k^2 N^2 + \cdots + C_k^{p-1} N^{p-1}.$$

On obtient alors les coordonnées de A_i^k dans la base $\mathcal{M}$ en lisant les coefficients de ce polynôme en N. Si $p \geqslant 1$, le coefficient de N est k qui n'est pas borné. Pour que $(A_i^k)_{k\in\mathbb{N}}$ soit borné, il est nécessaire que $p = 1$ *i.e.* A_i soit diagonale, ce qui signifie encore que le sous-espace propre relatif à λ_i est égal au sous-espace caractéristique : $\mathrm{Ker}(A - \lambda_i I) = F_i$. Réciproquement, si A_i est diagonale, ses puissances sont clairement bornées.

Conclusion. Les puissances de A sont bornées si, et seulement si $\rho(A) \leqslant 1$ et pour toute valeur propre complexe de module 1, $\mathrm{Ker}(A - \lambda I) = F_\lambda$, F_λ désignant le sous-espace caractéristique relatif à λ. ◁

L'exercice suivant utilise des arguments de densité pour étendre un résultat démontré dans un cas particulier : la densité de $\mathrm{GL}_n(\mathbb{C})$ dans $\mathcal{M}_n(\mathbb{C})$ et celle de l'ensemble des matrices diagonalisables de $\mathcal{M}_n(\mathbb{C})$ dans $\mathcal{M}_n(\mathbb{C})$.

Le premier résultat s'établit en remarquant que pour $A \in \mathcal{M}_n(\mathbb{C})$, la fonction polynôme $\lambda \longmapsto \det(A - \lambda I_n)$ n'est pas identiquement nulle donc a des zéros isolés. Pour λ assez petit non nul, $A - \lambda I_n$ est inversible. Pour $k \in \mathbb{N}$, la matrice $A_k = A - \frac{1}{k} I_n$ est inversible et la suite $(A_k)_{k\geqslant 1}$ converge vers A. Le second résultat a été vu dans l'exercice 2.58.

2.61. Fonctions polynomiales Φ sur $\mathcal{M}_n(\mathbb{C})$ vérifiant $\Phi(AB) = \Phi(BA)$

On désigne par $f_1, \ldots, f_n$ les n fonctions de $A \in \mathcal{M}_n(\mathbb{C})$ qui donnent les coefficients du polynôme caractéristique χ_A de A :

$$\chi_A = X^n + f_1(A)X^{n-1} + \cdots + f_{n-1}(A)X + f_n(A).$$

1. Montrer qu'on a $f_i(AB) = f_i(BA)$ pour tout $i \in [\![1, n]\!]$ et tous A et B dans $\mathcal{M}_n(\mathbb{C})$.

2. Soit Φ une fonction polynôme de $\mathcal{M}_n(\mathbb{C})$ dans $\mathbb{C}$, telle que pour tout couple $(A, B) \in \mathcal{M}_n(\mathbb{C})$, $\Phi(AB) = \Phi(BA)$. Montrer que Φ est un polynôme en $f_1, \ldots, f_n$.

(**École polytechnique**)

▷ **Solution.**

1. Il suffit en fait de prouver que $\chi_{AB} = \chi_{BA}$. Si B est inversible, on peut écrire

$$\chi_{AB} = \det(XI_n - AB) = \det((XB^{-1} - A)B) = \det(XB^{-1} - A)\det B$$
$$= \det(B(XB^{-1} - A)) = \det(XI_n - BA) = \chi_{BA}.$$

Pour A fixé, les fonctions $B \longmapsto f_i(BA) - f_i(AB)$ $(1 \leqslant i \leqslant n)$ sont des fonctions polynomiales des coefficients de $B \in \mathcal{M}_n(\mathbb{C})$: ce sont des fonctions continues de B. Comme $\chi_{AB} = \chi_{BA}$ pour B inversible, elles sont nulles sur $\mathrm{GL}_n(\mathbb{C})$ qui est dense dans $\mathcal{M}_n(\mathbb{C})$. Par continuité, elles sont nulles sur tout $\mathcal{M}_n(\mathbb{C})$.

On peut éviter de s'appuyer sur la densité topologique de $\mathrm{GL}_n(\mathbb{C})$ *par un argument algébrique valable sur un corps commutatif* K *quelconque : la matrice* $B - YI_n$, *à coefficients dans le corps* $K(Y)$ *des fractions rationnelles en* Y *est inversible dans* $\mathcal{M}_n(K(Y))$, *puisque son déterminant est un polynôme en* Y, *non nul de degré* n. *Ainsi, on a comme précédemment*

$$\chi_{A(B-YI_n)}(X) = \chi_{(B-YI_n)A}(X).$$

En faisant $Y = 0$, *il vient* $\chi_{AB} = \chi_{BA}$. *Une troisième solution est possible en utilisant un astucieux calcul de déterminants par blocs.*

2. Remarquons que Φ est constante sur les classes de similitude : si A, B sont dans $\mathcal{M}_n(\mathbb{C})$ et $P \in \mathrm{GL}_n(\mathbb{C})$ avec $B = P^{-1}AP$,

$$\Phi(B) = \Phi(P^{-1}AP) = \Phi(APP^{-1}) = \Phi(A).$$

On va naturellement s'intéresser au cas des matrices diagonalisables pour lesquelles les fonctions f_i sont au signe près les fonctions symétriques élémentaires des valeurs propres : si $\lambda_1, \ldots, \lambda_n$ sont les valeurs propres de A comptées avec multiplicité, on a

$$\chi_A = \prod_{i=1}^{n}(X - \lambda_i) = X^n + \sum_{k=0}^{n-1}(-1)^{n-k}\sigma_{n-k}(\lambda_1, \ldots, \lambda_n)X^k.$$

Considérons le polynôme en $(\lambda_1, \ldots, \lambda_n) \in \mathbb{C}^n$

$$P(\lambda_1, \ldots, \lambda_n) = \Phi(\mathrm{Diag}(\lambda_1, \ldots, \lambda_n)).$$

Si σ est une permutation de $[\![1, n]\!]$, $\mathrm{Diag}(\lambda_{\sigma(1)}, \ldots, \lambda_{\sigma(n)})$ est semblable à $\mathrm{Diag}(\lambda_1, \ldots, \lambda_n)$ (cela correspond à la permutation σ des vecteurs de la base canonique). Il s'ensuit que

$$P(\lambda_{\sigma(1)}, \ldots, \lambda_{\sigma(n)}) = P(\lambda_1, \ldots, \lambda_n).$$

Autrement dit, le polynôme P est symétrique en les λ_i. Il existe donc un polynôme R tel que

$$\begin{aligned} \mathrm{P}(\lambda_1, \ldots, \lambda_n) &= \mathrm{R}(\sigma_1(\lambda_1, \ldots, \lambda_n), \ldots, \sigma_n((\lambda_1, \ldots, \lambda_n)) \\ &= \mathrm{R}(-f_1(\mathrm{D}), f_2(\mathrm{D}), \ldots, (-1)^n f_n(\mathrm{D})) \end{aligned}$$

où $\mathrm{D} = \mathrm{Diag}(\lambda_1, \ldots, \lambda_n)$. Si $\mathrm{R}' = \mathrm{R}(-\mathrm{X}_1, \mathrm{X}_2, \ldots, (-1)^n \mathrm{X}_n)$, on a donc pour toute matrice diagonale D

$$\Phi(\mathrm{D}) = \mathrm{R}'(f_1(\mathrm{D}), \ldots, f_n(\mathrm{D})).$$

Comme Φ et les f_i prennent les mêmes valeurs sur deux matrices semblables, on a donc pour tout matrice A diagonalisable, semblable à D diagonale,

$$\Phi(\mathrm{A}) = \Phi(\mathrm{D}) = \mathrm{R}'(f_1(\mathrm{D}), \ldots, f_n(\mathrm{D})) = \mathrm{R}'(f_1(\mathrm{A}), \ldots, f_n(\mathrm{A})).$$

Nous savons que l'ensemble des matrices diagonalisables est dense dans $\mathcal{M}_n(\mathbb{C})$. Comme les fonctions Φ, R' et f_i sont continues car polynomiales, l'identité suivante valable pour A diagonalisable

$$\Phi(\mathrm{A}) = \mathrm{R}'(f_1(\mathrm{A}), \ldots, f_n(\mathrm{A})),$$

reste vraie par densité pour A quelconque dans $\mathcal{M}_n(\mathbb{C})$.
Conclusion. La fonction Φ est une fonction polynomiale en $f_1, \ldots, f_n$. ◁

Nous terminons ce chapitre avec un exercice relativement difficile mais très intéressant.

2.62. Familles de matrices anticommutantes

Déterminer le cardinal maximum d'une famille de matrices de $\mathrm{GL}_n(\mathbb{C})$ qui anticommutent deux à deux.

(ENS Ulm)

▷ **Solution.**

On considérera qu'une famille de $\mathrm{GL}_n(\mathrm{C})$ réduite à un seul élément répond au problème. On notera $r_n \in \mathbb{N}^* \cup \{\infty\}$ le cardinal maximal en dimension n.

Il est clair que $r_n = 1$ si n est impair : en effet, si A et B sont deux matrices de $\mathcal{M}_n(\mathbb{C})$ telles que $\mathrm{AB} = -\mathrm{BA}$, en passant au déterminant, il vient

$$\det A \det B = \det AB = \det(-AB) = (-1)^n \det A \det B = -\det A \det B,$$

d'où $\det A \det B = 0$ et A ou B n'est pas inversible.

Passons au cas n pair. Si A et B sont dans $GL_n(\mathbb{C})$ et anticommutent, A est semblable à son opposée puisque $AB = -BA$ entraîne $A = B(-A)B^{-1}$. Comme $\chi_{-A} = \det(XI_n + A) = (-1)^n \det(-XI_n - A) = \chi_A(-X)$ et puisque A est semblable à $-A$, $\chi_A = \chi_{-A} = \chi_A(-X)$ ce qui entraîne que si λ est valeur propre de A, alors $-\lambda$ est valeur propre avec la même multiplicité.

• Le candidat passant l'oral prouvera son bon sens scientifique en essayant de voir ce qu'il en est pour les petites dimensions. Étudions donc le cas $n = 2$.

Prenons une famille anticommutante $(A_1, \ldots, A_r)$ de $GL_2(\mathbb{C})$. On suppose $r \geqslant 2$. Soit $\lambda \in \operatorname{Sp} A_1$ ($\mathbb{C}$ est algébriquement clos). Alors $-\lambda \in \operatorname{Sp} A_1$. Comme λ est non nul, la matrice A_1 a deux valeurs propres distinctes et est donc diagonalisable. Soit $P \in GL_2(\mathbb{C})$ tel que $P^{-1}A_1P = \begin{pmatrix} \lambda & 0 \\ 0 & -\lambda \end{pmatrix}$. Alors la famille des $P^{-1}A_kP$ est aussi anticommutante et on a donc construit une nouvelle famille de même cardinal avec la première matrice diagonale. On la notera encore $(A_1, \ldots, A_r)$. Quitte à toutes les diviser par λ on peut supposer $A_1 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$.

Écrivons $A_k = \begin{pmatrix} a_k & b_k \\ c_k & d_k \end{pmatrix}$ $(2 \leqslant k \leqslant r)$. On a alors

$$A_1A_k = \begin{pmatrix} a_k & b_k \\ -c_k & -d_k \end{pmatrix} = -A_1A_k = \begin{pmatrix} -a_k & b_k \\ -c_k & d_k \end{pmatrix}$$

D'où $a_k = d_k = 0$ pour tout $k \geqslant 2$. Supposons $r \geqslant 4$. Les complexes b_k et c_k sont non nuls puisque les matrices A_k sont inversibles. On a

$$A_2A_3 = \begin{pmatrix} b_2c_3 & 0 \\ 0 & b_3c_2 \end{pmatrix} = -A_3A_2 = \begin{pmatrix} -b_3c_2 & 0 \\ 0 & -b_2c_3 \end{pmatrix}$$

D'où $b_2c_3 = -b_3c_2$ et $\frac{b_2}{c_2} = -\frac{b_3}{c_3}$. En écrivant $A_2A_4 = -A_4A_2$, il vient $\frac{b_2}{c_2} = -\frac{b_4}{c_4}$. De même avec $A_3A_4 = -A_4A_3$, $\frac{b_3}{c_3} = -\frac{b_4}{c_4}$. C'est impossible ! On a donc nécessairement $r \leqslant 3$. En fait, on peut exhiber une famille anticommutante de trois matrices :

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

Ainsi on vient de prouver que $r_2 = 3$.

• On va essayer de trouver une relation de récurrence entre r_n pour n pair et des valeurs r_k plus petites. Posons $n = 2p$ avec $p \in \mathbb{N}^*$. On fait l'hypothèse que le cardinal maximum d'une famille anticommutante de matrices de $\mathrm{GL}_p(\mathbb{C})$ est fini et vaut r_p. Soit $(\mathrm{A}_1, \ldots, \mathrm{A}_r)$ une famille anticommutante de $\mathrm{GL}_n(\mathbb{C})$ avec $r \geqslant 3$. Comme dans le cas $n = 2$, on va montrer qu'on peut se ramener à une famille de matrices diagonalisables. Comme on l'a vu plus haut pour $n = 2$, A_1 est semblable à $-\mathrm{A}_1$. Soit λ une valeur propre de A_1. Nous allons noter $\mathrm{F}_\lambda = \mathrm{Ker}(\mathrm{A}_1 - \lambda \mathrm{I}_n)^m$ et $\mathrm{F}_{-\lambda} = \mathrm{Ker}(\mathrm{A}_1 + \lambda \mathrm{I}_n)^m$ les sous-espaces caractéristiques correspondants de A_1.

Lemme. *Pour $k \geqslant 2$, $\mathrm{A}_k(\mathrm{F}_\lambda) \subset \mathrm{F}_{-\lambda}$ et $\mathrm{A}_k(\mathrm{F}_{-\lambda}) \subset \mathrm{A}_k(\mathrm{F}_\lambda)$.*

Démonstration.
Soit $\mathrm{X} \in \mathrm{F}_\lambda$, $\mathrm{A} = \mathrm{A}_1$ et $\mathrm{B} = \mathrm{A}_k$. On a

$$\begin{aligned}(\mathrm{A} - \lambda \mathrm{I}_n)^m \mathrm{BX} &= (\mathrm{A} - \lambda \mathrm{I}_n)^{m-1} \mathrm{B}(-\mathrm{A} - \lambda)\mathrm{X} = \ \ldots \ = \mathrm{B}(-\mathrm{A} - \lambda \mathrm{I}_n)^m \mathrm{X} \\ &= (-1)^m \mathrm{B}(\mathrm{A} + \lambda \mathrm{I}_n)^m \mathrm{X} = (-1)^m \mathrm{B}0 = 0.\end{aligned}$$

Par conséquent $\mathrm{BX} \in \mathrm{F}_{-\lambda}$. Cela montre la première inclusion. L'autre s'obtient en remplaçant λ par $-\lambda$. ♢

Lemme. *Pour $k \geqslant 2$, $\mathrm{E}_\lambda = \mathrm{F}_\lambda \oplus \mathrm{F}_{-\lambda}$ est stable par A_k.*

Démonstration. Pour $k = 1$, c'est parce que les sous-espaces caractéristiques de A_1 sont stables par A_1.

La stabilité par A_k pour $k \geqslant 2$ résulte du lemme précédent. Le fait que F_λ et $\mathrm{F}_{-\lambda}$ soient en somme directe est une conséquence du théorème de décomposition des noyaux puisque $(\mathrm{X} - \lambda)^m$ et $(\mathrm{X} + \lambda)^m$ sont premiers entre eux. ♢

Nous avons vu que χ_{A_1} pouvait s'écrire

$$\chi_{\mathrm{A}_1} = (\mathrm{X} - \lambda_1)^{n_1}(\mathrm{X} + \lambda_1)^{n_1}(\mathrm{X} - \lambda_2)^{n_2}(\mathrm{X} + \lambda_2)^{n_2} \ \ldots (\mathrm{X} - \lambda_s)^{n_s}(\mathrm{X} + \lambda_s)^{n_s},$$

où les λ_k et $-\lambda_k$ sont deux à deux distincts et les $n_k \in \mathbb{N}^*$. On obtient

$$\begin{aligned}\mathbb{C}^n &= \mathrm{F}_{\lambda_1} \oplus \mathrm{F}_{-\lambda_1} \oplus \mathrm{F}_{\lambda_2} \ominus \mathrm{F}_{-\lambda_2} \oplus \ \ldots \oplus \mathrm{F}_{\lambda_s} \oplus \mathrm{F}_{-\lambda_s} \\ &= \mathrm{E}_{\lambda_1} \oplus \mathrm{E}_{\lambda_2} \oplus \ \ldots \ominus \mathrm{E}_{\lambda_s}.\end{aligned}$$

Nous identifions les matrices et les endomorphismes canoniquement associés. D'après les lemmes précédents, pour $1 \leqslant l \leqslant s$, E_{λ_l} est stable par chaque A_k. Prenons $\mathcal{B}_l$ une base de F_{λ_l} et $\mathcal{B}'_l$ une base de $\mathrm{F}_{-\lambda_l}$. Pour

$k \geqslant 2$, la matrice de A_k restreinte à E_{λ_l} dans la base $(\mathcal{B}_l, \mathcal{B}'_l)$ est de la forme

$$\left(\begin{array}{c|c} 0 & C \\ \hline B & 0 \end{array}\right)$$

Puisque $\mathbb{C}^n$ est somme directe des E_{λ_l}, il existe un unique endomorphisme u de $\mathbb{C}^n$ dont la restriction à chaque E_{λ_l} admet comme matrice dans la base $(\mathcal{B}_l, \mathcal{B}'_l)$:

$$\left(\begin{array}{c|c} I & 0 \\ \hline 0 & -I \end{array}\right)$$

On remarque que

$$\left(\begin{array}{c|c} 0 & B \\ \hline C & 0 \end{array}\right)\left(\begin{array}{c|c} I & 0 \\ \hline 0 & -I \end{array}\right) = \left(\begin{array}{c|c} 0 & B \\ \hline -C & 0 \end{array}\right) = -\left(\begin{array}{c|c} 0 & B \\ \hline C & 0 \end{array}\right)\left(\begin{array}{c|c} I & 0 \\ \hline 0 & -I \end{array}\right)$$

Autrement dit, pour $k \geqslant 2$, A_k et u anticommutent sur chaque E_{λ_l} et donc sur $\mathbb{C}^n$ tout entier. Si on note D la matrice de u dans la base canonique, $(D, A_2, \ldots, A_r)$ est une famille anticommutante de $GL_n(\mathbb{C})$ de même cardinal dont la première matrice D est semblable à

$$\left(\begin{array}{c|c} I_p & 0 \\ \hline 0 & -I_p \end{array}\right)$$

En remplaçant A_1 par D et les matrices par des matrices semblables, on peut supposer que $A_1 = \left(\begin{array}{c|c} I & 0 \\ \hline 0 & -I \end{array}\right)$. Ainsi, s'il existe une famille anticommutante de cardinal $r \geqslant 2$, il en existe une que nous noterons encore $(A_1, A_2, \ldots, A_r)$ dont la première matrice est $A_1 = \left(\begin{array}{c|c} I_p & 0 \\ \hline 0 & -I_p \end{array}\right)$.

• Si $E_1 = \mathrm{Vect}(e_1, \ldots, e_p) = \mathrm{Ker}(A_1 - I) = \mathrm{Ker}(A_1 - I)^p$ et $E_2 = \mathrm{Vect}(e_{p+1}, \ldots, e_n) = \mathrm{Ker}(A_1 + I) = \mathrm{Ker}(A_1 + I)^p$, on a d'après le premier lemme $A_k(E_1) \subset E_2$ et $A_k(E_2) \subset E_1$ pour $2 \leqslant k \leqslant r$. Il existe donc B_k et C_k dans $GL_p(\mathbb{C})$ tels que

$$A_k = \left(\begin{array}{c|c} 0 & B_k \\ \hline C_k & 0 \end{array}\right)$$

Pour $2 \leqslant k, l \leqslant r$ distincts, on a

$$A_k A_l = \left(\begin{array}{c|c} B_k C_l & 0 \\ \hline 0 & C_k B_l \end{array}\right) = -A_l A_k = \left(\begin{array}{c|c} -B_l C_k & 0 \\ \hline 0 & -C_l B_k \end{array}\right)$$

On en déduit que $\begin{cases} B_k C_l = -B_l C_k \\ C_k B_l = -C_l B_k \end{cases}$. Comme les A_k sont inversibles, les B_k et les C_k sont inversibles. Posons, pour $k \geqslant 3$, $B = B_2$, $C = C_2$ et $D_k = B_k B^{-1}$.

Lemme. $D_3, \ldots, D_r$ *anticommutent.*

Démonstration.

Pour $3 \leqslant k, l \leqslant r$, distincts, on a $B_k C = -BC_k$ ou $B^{-1}B_k = -C_k C^{-1}$. Par conséquent,

$$\begin{aligned} D_k D_l &= B_k B^{-1} B_l B^{-1} = B_k (B^{-1} B_l) B^{-1} \\ &= -B_k (C_l C^{-1}) B^{-1} = -(B_k C_l) C^{-1} B^{-1} \\ &= B_l C_k C^{-1} B^{-1} = B_l (C_k C^{-1}) B^{-1} \\ &= -B_l B^{-1} B_k B^{-1} = -D_l D_k \end{aligned}$$

Par hypothèse, le cardinal d'une famille anticommutante de matrices inversibles de taille p est inférieur ou égal à r_p. On a donc $r - 2 \leqslant r_p$ et $r \leqslant r_p + 2$. On en déduit que r_n est fini et $r_n \leqslant r_p + 2$.

• Nous allons montrer qu'en fait $r_n = r_p + 2$. Par hypothèse, il existe une famille anticommutante de $\mathrm{GL}_p(\mathbb{C})$ de cardinal r_p. On la note $D_3, \ldots, D_r$ avec $r = r_p + 2$. On pose

$$A_1 = \left(\begin{array}{c|c} I_p & 0 \\ \hline 0 & -I_p \end{array}\right), \quad A_2 = \left(\begin{array}{c|c} 0 & I_p \\ \hline -I_p & 0 \end{array}\right) \quad \text{et} \quad A_k = \left(\begin{array}{c|c} 0 & D_k \\ \hline D_k & 0 \end{array}\right),$$

pour $3 \leqslant k \leqslant r$. On vérifie aisément que A_1 anticommute avec $A_2, \ldots, A_r$ et que A_2 anticommute avec tous les A_k pour $k \geqslant 3$. Enfin, pour $k, l \geqslant 3$, distincts

$$A_k A_l = \left(\begin{array}{c|c} D_k D_l & 0 \\ \hline 0 & D_k D_l \end{array}\right) = -\left(\begin{array}{c|c} D_l D_k & 0 \\ \hline 0 & D_l D_k \end{array}\right) = -A_l A_k.$$

On a donc construit une famille anticommutante de cardinal $r = r_p + 2$. Le cardinal maximal des familles anticommutantes de $\mathrm{GL}_n(\mathbb{C})$ existe et vaut donc $r_p + 2$.

• Pour terminer, montrons par récurrence sur n que le cardinal maximum des familles anticommutantes de $\mathrm{GL}_n(\mathbb{C})$ est fini et vaut $r_n = 2\nu_2(n) + 1$ où $\nu_2(n)$ est l'exposant de 2 dans la décomposition de n en produit de facteurs premiers.

∗ C'est vu si n est impair ou $n = 2$ (la formule convient bien).

∗ Si $n \geqslant 3$, n pair, on pose $p = \frac{n}{2}$. D'après l'hypothèse de récurrence, le cardinal maximum des familles anticommutantes de $\mathrm{GL}_p(\mathbb{C})$ est $2\nu_2(p)+1$. D'après ce qui précède, le cardinal maximum des familles anticommutantes de $\mathrm{GL}_n(\mathbb{C})$ est $2\nu_2(p)+1+2 = 2(\nu_2(p)+1)+1 = 2\nu_2(n)+1$.

Conclusion. Le cardinal maximal des familles de matrices de $\mathrm{GL}_n(\mathbb{C})$ qui anticommutent existe et vaut $\boxed{r_n = 2\nu_2(n) + 1}$. ◁

Chapitre 3
Le groupe linéaire

La connaissance d'une partie génératrice d'un groupe est fondamentale comme le lecteur pourra s'en rendre compte à travers la plupart des exercices de ce chapitre. Pour un groupe donné, on cherche à avoir des éléments générateurs les plus simples possibles. Lorsque ce groupe opère naturellement sur un ensemble (comme c'est le cas du groupe symétrique ou ici du groupe linéaire), on va donc regarder les éléments qui ont le plus de points fixes. Pour le groupe symétrique on considère alors les transpositions (qui bougent seulement deux éléments) et qui en forment bien une partie génératrice. Dans le cas du groupe linéaire d'un espace vectoriel de dimension finie, on regarde les applications linéaires inversibles qui admettent un hyperplan de points fixes. Cela mène alors aux transvections (cas non diagonalisable) et aux dilatations (qui sont diagonalisables).

Le premier exercice de ce chapitre se propose de démontrer que le groupe linéaire $\mathrm{GL}_n(\mathrm{K})$ *est bien engendré par l'ensemble des matrices de dilatation et des matrices de transvection. Il s'agit simplement de regarder convenablement l'algorithme du pivot de Gauss permettant de calculer l'inverse d'une matrice.*

3.1. Génération du groupe linéaire

Soit $n \geqslant 2$ et K un corps commutatif. On appelle matrice de transvection toute matrice de la forme $\mathrm{T}_{ij}(\lambda) = \mathrm{I}_n + \lambda \mathrm{E}_{ij}$ où $i \neq j$ et $\lambda \in \mathrm{K}$, et matrice de dilatation toute matrice diagonale $\mathrm{D}_i(\alpha) = \mathrm{I}_n + (\alpha - 1)\mathrm{E}_{ii}$ avec $\alpha \in \mathrm{K}^*$.

1. Montrer que l'ensemble des matrices de transvection engendre le groupe $\mathrm{SL}_n(\mathrm{K})$ et que l'ensemble des matrices de transvection et de dilatation engendre le groupe $\mathrm{GL}_n(\mathrm{K})$.

2. Déterminer les centres des groupes $\mathrm{GL}_n(\mathrm{K})$ et $\mathrm{SL}_n(\mathrm{K})$.

3. On suppose que $\mathrm{K} = \mathbb{R}$ ou $\mathbb{C}$. Démontrer que $\mathrm{SL}_n(\mathrm{K})$ est connexe par arcs.

(École Polytechnique)

▷ **Solution.**

1. Notons que les matrices de transvection sont toutes de déterminant 1. Le groupe qu'elles engendrent est donc inclus dans $\mathrm{SL}_n(\mathrm{K})$. La multiplication à gauche (resp. à droite) par une matrice de transvection $\mathrm{T}_{ij}(\lambda)$ revient à effectuer l'opération élémentaire $\mathrm{L}_i \leftarrow \mathrm{L}_i + \lambda \mathrm{L}_j$ (resp. $\mathrm{C}_j \leftarrow \mathrm{C}_j + \lambda \mathrm{C}_i$). Notons qu'il est possible de réaliser l'échange de deux lignes (ou de deux colonnes) uniquement à l'aide de transvections modulo un changement de signe : en effet, la matrice $\mathrm{T}_{ij}(1)\mathrm{T}_{ji}(-1)\mathrm{T}_{ij}(1)$ a pour effet (par multiplication à gauche) de remplacer L_i par L_j et L_j par $-\mathrm{L}_i$, les autres lignes étant invariantes. Il n'est évidemment pas possible de réaliser l'échange de deux lignes sans apparition de ce signe moins puisque cette opération change le signe du déterminant.

Soit $\mathrm{A} \in \mathrm{GL}_n(\mathrm{K})$. En appliquant l'algorithme du pivot de Gauss, nous allons transformer A en une matrice de dilatation mais en utilisant uniquement des transvections. Comme A est inversible, sa première colonne n'est pas nulle. Si $a_{i1} \neq 0$ avec $i \geqslant 2$, l'opération $\mathrm{L}_1 \leftarrow \mathrm{L}_1 - \frac{a_{11}-1}{a_{i1}}\mathrm{L}_i$ permet de mettre un coefficient 1 en position $(1,1)$. Si tous les coefficients a_{i1} pour $i \geqslant 2$ sont nuls, on effectue l'échange des lignes $\mathrm{L}_1 \leftarrow \mathrm{L}_2$ et $\mathrm{L}_2 \leftarrow -\mathrm{L}_1$ pour se ramener au cas précédent. En utilisant le coefficient $(1,1)$ comme pivot, une succession d'opérations sur les lignes puis sur les colonnes permet d'annuler tous les autres coefficients de la première ligne et de la première colonne. Autrement dit, il existe des matrices de transvections $\mathrm{M}_1, \ldots, \mathrm{M}_p$ et $\mathrm{N}_1, \ldots, \mathrm{N}_q$ telles que

$$\mathrm{M}_p \ldots \mathrm{M}_1 \mathrm{A} \mathrm{N}_1 \ldots \mathrm{N}_q = \begin{pmatrix} 1 & 0 \\ 0 & \mathrm{A}_1 \end{pmatrix}$$

où $\mathrm{A}_1 \in \mathrm{GL}_{n-1}(\mathrm{K})$.

On recommence le même algorithme sur la matrice A_1 et ainsi de suite. On aboutit à la fin de cet algorithme à une matrice diagonale $\mathrm{Diag}(1, 1, \ldots, 1, \alpha)$, où le scalaire α n'est autre que $\det \mathrm{A}$. On vient donc de montrer que pour toute matrice inversible A, il existe des matrices de transvection $\mathrm{U}_1, \ldots, \mathrm{U}_r$ et $\mathrm{V}_1, \ldots, \mathrm{V}_s$ telles que

$$\mathrm{A} = \mathrm{U}_r \ldots \mathrm{U}_1 \mathrm{D}_n(\det \mathrm{A}) \mathrm{V}_1 \ldots \mathrm{V}_s.$$

Cela permet de répondre à la question : toute matrice $\mathrm{A} \in \mathrm{SL}_n(\mathrm{K})$ s'écrit comme un produit de matrices de transvection et toute matrice de $\mathrm{GL}_n(\mathrm{K})$ est produit de matrices de transvection et de dilatation.

2. Soit A une matrice qui commute avec toutes les matrices de transvection. Alors A commute avec toutes les matrices E_{ij}, $i \neq j$. Cela impose que A est scalaire. On en déduit que le centre de $\mathrm{GL}_n(\mathrm{K})$ est l'ensemble des matrices scalaires $\lambda \mathrm{I}_n$ avec $\lambda \neq 0$. Il est isomorphe au

groupe $(K^*, \times)$. Et le centre de $SL_n(K)$ est l'ensemble des matrices scalaires λI_n avec $\lambda^n = 1$. Il est isomorphe au groupe des racines n-ième de l'unité dans le corps K. Son cardinal dépend bien entendu du corps K.

3. Montrons que toute matrice $A \in SL_n(K)$ est reliée par un arc continu à l'identité I_n. D'après la première question, il existe une partie X contenue dans l'ensemble des couples $(i,j) \in [\![1,n]\!]^2$ avec $i \neq j$ et une famille $(\lambda_C)_{C\in X}$ de K telles que A soit le produit des transvections $T_C(\lambda_C)$:

$$A = \prod_{C\in X} T_C(\lambda_C).$$

En posant $\varphi : t \in [0,1] \longmapsto A_t = \prod_{C\in X} T_C(t\lambda_C)$, on obtient un arc continu qui relie $\varphi(0) = I_n$ à $\varphi(1) = A$ et $SL_n(K)$ est donc connexe par arcs. ◁

Le résultat de la première question est très important et sera utilisé dans plusieurs des exercices de ce chapitre, dont le suivant.

3.2. Groupe engendré par les matrices diagonalisables inversibles

Quel est le sous-groupe de $GL_n(\mathbb{R})$ engendré par l'ensemble des matrices inversibles diagonalisables ?

(ENS Ulm)

▷ **Solution.**

On sait (voir l'exercice précédent) que $GL_n(\mathbb{R})$ est engendré par les matrices de dilatation et les matrices de transvection. Les premières sont des matrices diagonales. Montrons que toute matrice de transvection peut s'écrire comme un produit de matrices diagonalisables. Si $M = I_n + \lambda E_{ij}$ avec $i \neq j$ et $\lambda \in \mathbb{R}$ est une telle matrice, il suffit d'écrire $M = D^{-1}(DM)$ où D est une matrice diagonale dont les coefficients diagonaux sont non nuls et deux à deux distincts (par exemple $1, 2, \ldots, n$). La matrice D^{-1} est diagonale et la matrice DM est triangulaire avec la même diagonale que D : elle est donc diagonalisable.

Conclusion. L'ensemble des matrices diagonalisables inversibles engendre tout le groupe $GL_n(\mathbb{R})$. ◁

Les prochains exercices sont consacrés à des sous-groupes du groupe linéaire. Le premier demande ainsi de montrer que $GL_2(\mathbb{Q})$ *ne contient pas de sous-groupe cyclique de cardinal* 5.

3.3. Élément d'ordre 5 de $\mathrm{GL}_2(\mathbb{Q})$

Montrer que le groupe $\mathrm{GL}_2(\mathbb{Q})$ ne contient pas d'élément d'ordre 5.

(ENS Ulm)

▷ **Solution.**

Supposons par l'absurde qu'il existe $A \in \mathrm{GL}_2(\mathbb{Q})$ d'ordre 5. Le polynôme $P = X^5 - 1 = (X-1)(X^4 + X^3 + X^2 + X + 1)$ annule A. Montrons que le second facteur Q est irréductible dans $\mathbb{Q}[X]$. En effet, Q n'a pas de racine réelle (donc *a fortiori* pas de racine rationnelle) et si $Q = AB$ où A, B sont deux polynômes unitaires du second degré alors, par unicité de la décomposition dans $\mathbb{R}[X]$, ces polynômes sont $X^2 + \sqrt{2}X + 1$ et $X^2 - \sqrt{2}X + 1$. C'est absurde car $\sqrt{2} \notin \mathbb{Q}$. Donc Q est bien irréductible dans $\mathbb{Q}[X]$.

Le polynôme minimal μ de A est un diviseur de P dans $\mathbb{Q}[X]$. Comme μ est de degré inférieur à 2, on a $\mu = X - 1$, et donc $A = I$. D'où le résultat. ◁

Si deux groupes G *et* G′ *sont isomorphes, tout sous-groupe de* G *doit se retrouver dans* G′. *L'exercice suivant utilise cette idée pour montrer que deux groupes linéaires* $\mathrm{GL}_n(K)$ *et* $\mathrm{GL}_m(L)$ *ne peuvent pas être isomorphes si* $n \neq m$.

3.4. Isomorphismes entre groupes linéaires

Soient K et L deux corps commutatifs de caractéristique différente de 2.

1. Soit G un sous-groupe fini de $\mathrm{GL}_n(K)$ tel que pour tout $A \in G$, $A^2 = I$. Montrer que $\mathrm{Card}\, G \leqslant 2^n$.

2. On suppose qu'il existe un morphisme injectif du groupe $\mathrm{GL}_n(K)$ dans le groupe $\mathrm{GL}_m(L)$. Montrer que $n \leqslant m$.

3. Existe-t-il un isomorphisme entre $\mathrm{GL}_n(\mathbb{Q})$ et $\mathrm{GL}_n(\mathbb{R})$? entre $\mathrm{GL}_n(\mathbb{R})$ et $\mathrm{GL}_n(\mathbb{C})$?

(ENS Lyon)

▷ **Solution.**

1. Comme K est de caractéristique différente de 2, le polynôme $X^2 - 1 = (X-1)(X+1)$ est scindé à racines simples. Par hypothèse, ce polynôme annule toute matrice A de G qui est donc diagonalisable avec

un spectre inclus dans $\{\pm 1\}$. Géométriquement les éléments de G sont des symétries.

Par ailleurs G est abélien : en effet, pour tout $A \in G$, on a $A = A^{-1}$ et si $(A, B) \in G^2$,

$$AB = (AB)^{-1} = B^{-1}A^{-1} = BA.$$

Les éléments de G sont alors codiagonalisables (c'est un résultat classique que le lecteur pourra retrouver dans l'exercice 2.19). Il existe donc $P \in GL_n(K)$ telle que, pour tout $A \in G$,

$$P^{-1}AP = \begin{pmatrix} \pm 1 & 0 & \dots & 0 \\ 0 & \pm 1 & & \vdots \\ \vdots & & \ddots & \vdots \\ 0 & \dots & 0 & \pm 1 \end{pmatrix}.$$

Le cardinal de G est donc inférieur à celui des familles $(\varepsilon_1, \dots, \varepsilon_n) \in \{-1, 1\}^n$ qui est égal à 2^n.

2. Considérons le sous-groupe G de $GL_n(K)$ constitué des matrices diagonales avec des 1 ou -1 sur la diagonale :

$$G = \{M \in GL_n(K),\ \exists (\varepsilon_1, \dots, \varepsilon_n) \in \{-1, 1\}^n,\ M = \mathrm{Diag}(\varepsilon_1, \dots, \varepsilon_n)\}.$$

Le groupe G est de cardinal 2^n et pour tout $A \in G$, $A^2 = I_n$. On considère un morphisme injectif de $GL_n(K)$ dans $GL_m(L)$. Alors l'image G' de G par ce morphisme est un sous-groupe fini isomorphe à G. En particulier, pour tout $B \in G'$, $B^2 = I_m$ et $\mathrm{Card}\, G' = 2^n$. D'après la première question, on a $2^n \leqslant 2^m$, ce qui entraîne $n \leqslant m$.

3. La question précédente montre que si $GL_n(K)$ est isomorphe à $GL_m(L)$ alors nécessairement $n = m$. Le groupe $GL_n(\mathbb{Q})$ ne peut pas être isomorphe à $GL_n(\mathbb{R})$ (ou $GL_n(\mathbb{C})$) car il est dénombrable alors que ces deux derniers sont indénombrables. Montrons pour terminer que $GL_n(\mathbb{R})$ et $GL_n(\mathbb{C})$ ne sont pas isomorphes. Pour $n = 1$, $\mathbb{R}^*$ et $\mathbb{C}^*$ ne sont pas isomorphes : il y a des éléments d'ordre 3 dans $\mathbb{C}^*$ mais pas dans $\mathbb{R}^*$. Ce cas suffit à conclure puisque le centre de $GL_n(K)$ est isomorphe à K^* (voir l'exercice 3.1). Donc si $GL_n(\mathbb{R})$ et $GL_n(\mathbb{C})$ étaient isomorphes, leurs centres $\mathbb{R}^*$ et $\mathbb{C}^*$ le seraient aussi, ce qui n'est pas. $\lhd$

L'exercice suivant étudie encore l'existence d'un sous-groupe fini d'un certain type dans le groupe linéaire.

3.5. Sous-groupe de $\mathrm{GL}_n(\mathbb{R})$

Pour quelles valeurs de n existe-t-il un sous-groupe de $\mathrm{GL}_n(\mathbb{R})$ isomorphe au groupe additif $(\mathbb{Z}/4\mathbb{Z})^2$?

(ENS Cachan)

▷ **Solution.**

Une première remarque s'impose : si une valeur de n convient, toute valeur plus grande convient aussi puisque si $n \leqslant m$, $\mathrm{GL}_n(\mathbb{R})$ peut s'identifier au sous-groupe de $\mathrm{GL}_m(\mathbb{R})$ formé des matrices diagonales par blocs de la forme $\begin{pmatrix} \mathrm{M} & 0 \\ 0 & \mathrm{I}_{m-n} \end{pmatrix}$ où $\mathrm{M} \in \mathrm{GL}_n(\mathbb{R})$. On cherche donc, si elle existe, la plus petite valeur de n qui convient.

Analysons un peu ce que l'on cherche. Le groupe $\mathbb{Z}/4\mathbb{Z}$ est cyclique d'ordre 4. Trouver un sous-groupe isomorphe à $\mathbb{Z}/4\mathbb{Z}$ revient à trouver un élément d'ordre 4. On peut voir le groupe $(\mathbb{Z}/4\mathbb{Z})^2$ comme deux copies « indépendantes » de $\mathbb{Z}/4\mathbb{Z}$: il est engendré par deux éléments d'ordre 4 (à savoir $(1,0)$ et $(0,1)$) qui commutent, les deux sous-groupes engendrés par ces deux éléments n'ayant en commun que l'élément neutre. Réciproquement, si A, B sont deux matrices de $\mathrm{GL}_n(\mathbb{R})$ d'ordre 4, qui commutent et telles que $\langle \mathrm{A} \rangle \cap \langle \mathrm{B} \rangle = \{\mathrm{I}_n\}$ (où $\langle \mathrm{A} \rangle = \{\mathrm{I}_n, \mathrm{A}, \mathrm{A}^2, \mathrm{A}^3\}$ désigne le groupe engendré par A), alors le groupe engendré par A et B est isomorphe à $(\mathbb{Z}/4\mathbb{Z})^2$: en effet, il est facile de vérifier que $\langle \mathrm{A}, \mathrm{B} \rangle = \{\mathrm{A}^k\mathrm{B}^l,\ 0 \leqslant k, l \leqslant 3\}$ et que $\psi : (\mathbb{Z}/4\mathbb{Z})^2 \to (\mathrm{A}, \mathrm{B})$ qui à $(\bar{k}, \bar{l})$ associe $\mathrm{A}^k\mathrm{B}^l$ est bien définie et est un isomorphisme de groupes. On est donc ramené à chercher deux matrices A et B vérifiant les propriétés énumérées ci-dessus.

Pour commencer étudions à quelle condition une matrice M de $\mathrm{GL}_n(\mathbb{R})$ est d'ordre 4. Cela signifie que $\mathrm{M}^4 = \mathrm{I}$ et que $\mathrm{M}^2 \neq \mathrm{I}$. La matrice M est diagonalisable dans $\mathcal{M}_n(\mathbb{C})$ et son spectre est inclus dans l'ensemble $\{1, -1, i, -i\}$ des racines quatrièmes de l'unité. Son spectre contient nécessairement i ou $-i$ car sinon M serait d'ordre 2. Par ailleurs, comme M est à coefficients réels, si i est valeur propre alors $-i$ aussi et avec la même multiplicité. On doit donc déjà avoir $n \geqslant 2$. Dans le cas $n = 2$, la matrice réelle $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$, qui a pour polynôme caractéristique $\mathrm{X}^2 + 1$, est bien d'ordre 4, diagonalisable sur $\mathbb{C}$, et semblable à la matrice $\begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}$. À partir de cette matrice il est facile de trouver une solution dans le cas $n = 4$. Il suffit en effet de prendre les deux matrices

$$A = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{pmatrix} \text{ et } B = \begin{pmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Elles sont toutes les deux d'ordre 4, commutent, et vérifient bien la condition $\langle A \rangle \cap \langle B \rangle = \{I_4\}$.

Pour conclure l'exercice, on va prouver qu'on ne peut pas trouver de solution dans le cas $n = 3$. Supposons par l'absurde qu'existent A et B dans $GL_3(\mathbb{R})$ vérifiant les propriétés ci-dessus. Comme on vient de le voir i et $-i$ sont valeurs propres de A et comme elle doivent avoir le même ordre de multiplicité celui-ci vaut 1. Notons $\varepsilon = \pm 1$ la troisième valeur propre de A. Le polynôme minimal de A est donc $(X - \varepsilon)(X^2 + 1)$ et le lemme des noyaux montre que $\mathbb{R}^3$ est somme directe de la droite $\mathrm{Ker}(A - \varepsilon I_3)$ et du plan $\mathrm{Ker}(A^2 + I_3)$. Ce qui vient d'être dit pour A est aussi valable pour B. Notons $\varepsilon' \in \{\pm 1\}$ la valeur propre réelle de B, i et $-i$ étant ses deux autres valeurs propres. Comme A et B commutent, la droite $\mathrm{Ker}(A - \varepsilon I_3)$ est stable par B et il s'agit donc d'une droite propre, nécessairement pour la valeur propre réelle ε'. Donc $\mathrm{Ker}(A - \varepsilon I_3) = \mathrm{Ker}(B - \varepsilon' I_3)$. De même le plan $\mathrm{Ker}(A^2 + I_3)$ est stable par B. Les valeurs propres complexes de la restriction de B à ce plan sont forcément i et $-i$ de sorte que $\mathrm{Ker}(B^2 + I_3) = \mathrm{Ker}(A^2 + I_3)$. Comme $\varepsilon^2 = \varepsilon'^2 = 1$, cela implique que $A^2 = B^2$, ce qui contredit l'hypothèse $\langle A \rangle \cap \langle B \rangle = \{I_n\}$. ◁

Si G *est un groupe, on appelle exposant de* G *le ppcm des ordres des éléments de* G *si celui-ci est défini ou* $+\infty$ *si les ordres des éléments de* G *n'ont aucun multiple commun. Un groupe infini peut avoir un exposant fini : c'est par exemple le cas du groupe additif* $(\mathbb{Z}/2\mathbb{Z})^{\mathbb{N}}$ *qui est d'exposant* 2. *Un célèbre problème posé par Burnside en 1902 demande si un groupe d'exposant fini et de type fini (c'est-à-dire engendré par une partie finie) est nécessairement fini. La réponse est négative et un contre-exemple a été trouvé, mais seulement en 1975. L'exercice suivant consiste à prouver qu'un sous-groupe du groupe linéaire* $GL_n(\mathbb{C})$ *qui est d'exposant fini est fini.*

3.6. Un théorème de Burnside

1. Soit $A \in \mathcal{M}_n(\mathbb{C})$ telle que $\mathrm{Tr}(A^k) = 0$ pour tout $k \in \mathbb{N}^*$. Montrer que A est nilpotente.

2. Soit G un sous-groupe de $\mathrm{GL}_n(\mathbb{C})$, $(\mathrm{M}_i)_{1\leqslant i\leqslant n^2} \in \mathrm{G}^{n^2}$ une base de $\mathrm{Vect}(\mathrm{G})$ et $f : \mathrm{G} \to \mathbb{C}^{n^2}$ l'application qui à $\mathrm{A} \in \mathrm{G}$ associe $(\mathrm{Tr}(\mathrm{AM}_i))_{1\leqslant i\leqslant n^2}$. Montrer que si $f(\mathrm{A}) = f(\mathrm{B})$ alors $\mathrm{AB}^{-1} - \mathrm{I}$ est nilpotente.

3. On suppose que toutes les matrices de G sont diagonalisables. Montrer que f est injective.

4. En déduire qu'un sous-groupe de $\mathrm{GL}_n(\mathbb{C})$ d'exposant fini (c'est-à-dire qu'il existe un entier N tel que $\mathrm{A}^{\mathrm{N}} = \mathrm{I}$ pour toute matrice A du groupe) est fini. *C'est le théorème de Burnside.*

(ENS Lyon)

▷ **Solution.**

1. Ce résultat est classique. On en trouvera plusieurs preuves dans l'exercice 2.33 du chapitre réduction.

2. Posons $\mathrm{D} = \mathrm{AB}^{-1}$. Par linéarité de la trace, on a $\mathrm{Tr}(\mathrm{AM}) = \mathrm{Tr}(\mathrm{BM})$ pour toute matrice $\mathrm{M} \in \mathrm{Vect}(\mathrm{G})$ et en particulier pour toute matrice M de G. Soit $k \in \mathbb{N}^*$. On a $\mathrm{Tr}(\mathrm{D}^k) = \mathrm{Tr}(\mathrm{AB}^{-1}\mathrm{D}^{k-1}) = \mathrm{Tr}(\mathrm{BB}^{-1}\mathrm{D}^{k-1}) = \mathrm{Tr}(\mathrm{D}^{k-1})$. Il en résulte donc que pour tout $k \in \mathbb{N}$, $\mathrm{Tr}\,\mathrm{D}^k = \mathrm{Tr}\,\mathrm{I}_n = n$. Ainsi, pour $k \geqslant 1$,

$$\mathrm{Tr}(\mathrm{D}-\mathrm{I}_n)^k = \mathrm{Tr}\left(\sum_{j=0}^{k} \mathrm{C}_k^j (-1)^j \mathrm{D}^{k-j}\right) = n\sum_{j=0}^{k} \mathrm{C}_k^j (-1)^j = n(1-1)^k = 0.$$

La question précédente permet de conclure.

3. Supposons de plus que les éléments de G sont diagonalisables et reprenons les notations précédentes. La matrice $\mathrm{D} = \mathrm{AB}^{-1}$ est dans G. Elle est donc diagonalisable. Mais alors $\mathrm{D} - \mathrm{I}$ est aussi diagonalisable ! Comme elle est nilpotente d'après ce qui précède, elle est nulle. Donc $\mathrm{D} = \mathrm{I}$, c'est-à-dire $\mathrm{A} = \mathrm{B}$. Conclusion : f est injective.

4. Prenons pour G un sous-groupe de $\mathrm{GL}_n(\mathbb{C})$ d'exposant fini N. Toute matrice A de G est annulée par le polynôme $\mathrm{X}^{\mathrm{N}} - 1$ qui est scindé à racines simples. Donc toute matrice de G est diagonalisable. Le résultat précédent s'applique et l'application $f : \mathrm{G} \to \mathbb{C}^{n^2}$ définie dans la question 2 est injective. En réalité, l'image de f est incluse dans X^{n^2} où X est l'ensemble des traces des éléments de G. Pour conclure, il suffit donc de prouver que X est fini. Or, vu ce qui précède, les valeurs propres des éléments de G appartiennent à l'ensemble fini des racines N-ième de l'unité. Donc X est fini. ◁

L'exercice suivant prouve que le groupe linéaire $\mathrm{GL}_n(\mathbb{C})$ ne contient pas de sous-groupe borné non trivial arbitrairement petit.

3.7. Petits sous-groupes de $\mathbf{GL}_n(\mathbb{C})$

On munit $\mathcal{M}_n(\mathbb{C})$ de la norme triple subordonnée à une norme de $\mathbb{C}^n$. Soit G un sous-groupe borné de $\mathrm{GL}_n(\mathbb{C})$.

1. Soit $M \in G$. Montrer que les valeurs propres de M sont de module 1 et que M est diagonalisable.

2. On suppose que $G \subset B(I_n, \sqrt{2})$ (boule ouverte de centre l'identité et de rayon $\sqrt{2}$). Montrer que G est réduit à $\{I_n\}$.

3. Montrer que le résultat reste vrai si on remplace $\sqrt{2}$ par $\sqrt{3}$.

(**École Polytechnique**)

▷ **Solution.**

1. On notera $\|\ \|$ la norme de $\mathbb{C}^n$ utilisée et $|||\ |||$ la norme triple de $\mathcal{M}_n(\mathbb{C})$ associée. Rappelons pour commencer que si $A \in \mathcal{M}_n(\mathbb{C})$ est une matrice quelconque et λ une valeur propre de A, alors $|\lambda| \leqslant |||A|||$. En effet, si X est un vecteur propre associé à la valeur propre λ on a $AX = \lambda X$ donc $|\lambda|\|X\| = \|\lambda X\| = \|AX\| \leqslant |||A|||X\|$. Comme X n'est pas nul, on peut simplifier $\|X\|$ pour obtenir l'inégalité annoncée.

Comme G est un groupe borné, il en résulte que l'ensemble des valeurs propres des éléments de G est borné. Soit $M \in G$ et λ une valeur propre de G. Pour tout $p \in \mathbb{Z}$, M^p est dans G et a pour valeur propre λ^p. D'après ce qu'on vient de dire, la suite $(\lambda^p)_{p\in\mathbb{Z}}$ est bornée et cela n'a lieu que si λ est de module 1 (λ n'est pas nul puisque M est inversible). Les valeurs propres de M sont donc toutes de module 1. Montrons maintenant que M est diagonalisable. Cela va encore découler du fait que la suite (M^p) est bornée et de la décomposition de Dunford (voir l'exercice 2.30). La matrice M s'écrit (de manière unique) $M = D+N$ avec D diagonalisable, N nilpotente et $ND = DN$. Soit s l'indice de nilpotence de N. On va montrer que $N = 0$, c'est-à-dire que $s = 1$. Supposons $s \geqslant 2$. Alors $\mathrm{Ker}\, N$ est strictement inclus dans $\mathrm{Ker}\, N^2$. Prenons X dans $\mathrm{Ker}\, N^2 \setminus \mathrm{Ker}\, N$. Comme D et N commutent, on a pour tout $p \geqslant s$,

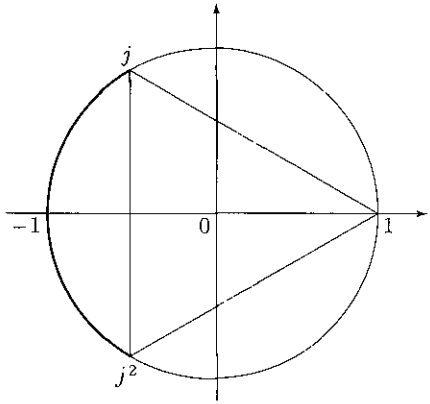
$$M^p = (D+N)^p = D^p + pD^{p-1}N + C_p^2 D^{p-2}N^2 + \cdots + C_p^{s-1} D^{p-s+1}N^{s-1}.$$

En particulier, $M^pX = D^pX + pD^{p-1}NX$. Il en découle que la suite $(M^pX)_{p\in\mathbb{N}}$ n'est pas bornée dans $\mathbb{C}^n$ ce qui est absurde car la suite $(M^p)_{p\in\mathbb{N}}$ est bornée.

2. On suppose que $G \subset B(I_n, \sqrt{2})$. Soit $M \in G$. On va montrer que M est égal à la matrice identité et pour cela, il nous suffit de prouver que sa seule valeur propre est 1 (car M est diagonalisable d'après la question précédente). Soit λ dans le spectre de M. On a $|\lambda - 1| \leqslant |||M - I_n||| < \sqrt{2}$ par hypothèse. Comme $M^p \in G$ pour tout $p \in \mathbb{Z}$, on doit aussi avoir

$|\lambda^p - 1| < \sqrt{2}$ pour tout $p \in \mathbb{Z}$. Il suffit de prouver que cela n'est le cas que si $\lambda = 1$. Géométriquement, cela consiste à prouver que l'une des puissances de λ au moins possède une partie réelle négative. Posons $\lambda = e^{i\theta}$ avec $0 < \theta < \frac{\pi}{2}$ (on peut se contenter de traiter ce cas quitte à prendre le conjugué de λ c'est-à-dire son inverse ce qui revient juste à remplacer M par M^{-1}). Il existe clairement un entier p tel que $\frac{\pi}{2} \leqslant p\theta < \pi$ car $\frac{\pi}{2\theta} > 1$. On a alors $\cos(p\theta) \leqslant 0$ et $|1 - \lambda^p| \geqslant \sqrt{2}$.

3. On raisonne comme dans la question précédente. Il faut juste changer l'argument final pour montrer que si λ est un complexe de module 1 distinct de 1, il existe au moins un entier p tel $|1 - \lambda^p| \geqslant \sqrt{3}$. Les deux complexes du cercle unité qui sont à distance $\sqrt{3}$ de 1 sont j et j^2, les racines cubiques non triviales de 1.



On doit donc montrer que l'une des puissances de λ tombe dans le petit arc de cercle délimité par j et j^2. Posons $\lambda = e^{i\theta}$ avec $0 < \theta \leqslant \pi$ (on peut se limiter à ce cas quitte à remplacer λ par son inverse). Si $\theta \geqslant \frac{2\pi}{3}$ c'est fini car on a directement $|\lambda - 1| \geqslant \sqrt{3}$. Supposons $0 < \theta < \frac{2\pi}{3}$. Il existe un entier p tel que $\frac{2\pi}{3} \leqslant p\theta < \frac{4\pi}{3}$ car $\frac{2\pi}{3\theta} > 1$. D'où le résultat. ◁

Si on remplace $\sqrt{3}$ par une valeur plus grande le résultat n'est plus forcément vrai comme le montre le sous-groupe de $\mathrm{GL}_1(\mathbb{C})$ égal à $\{1, j, j^2\}$. En effet, la norme sur $\mathbb{C}$ étant le module, ce groupe est inclus dans la boule fermée de centre 1 et de rayon $\sqrt{3}$.

Si G est un groupe, on appelle commutateur de G tout élément de la forme $ghg^{-1}h^{-1}$. Le sous-groupe engendré par les commutateurs est appelé le groupe dérivé de G.

3.8. Groupe dérivé de $GL_n(K)$

Soit $n \geqslant 2$ et K un corps possédant au moins 3 éléments. Montrer que le sous-groupe de $GL_n(K)$ engendré par tous les commutateurs $ABA^{-1}B^{-1}$ avec $(A, B) \in GL_n(K)^2$, est égal à $SL_n(K)$.

(École Polytechnique)

▷ **Solution.**

Notons tout d'abord que pour tout $(A, B) \in GL_n(K)^2$ on a $\det(ABA^{-1}B^{-1}) = 1$ par multiplicativité du déterminant. Le groupe engendré par les commutateurs est donc inclus dans $SL_n(K)$. On sait que ce dernier est engendré par les matrices de transvection (cf. exercice 3.1). On va donc essayer de voir si toute matrice de transvection est un commutateur. Pour $i \neq j$ et λ dans K on pose $T_{ij}(\lambda) = I_n + \lambda E_{ij}$. On a alors $T_{ij}(\lambda)^{-1} = T_{ij}(-\lambda)$. Calculons le commutateur défini par une matrice de dilatation $D_i(a)$ avec $a \notin \{0, 1\}$ (il en existe car K contient au moins 3 éléments) et une matrice de transvection $T_{ij}(b) = I_n + bE_{ij}$. On a

$$D_i(a)(I_n+bE_{ij})D_i(a)^{-1} = I_n+D_i(a)bE_{ij}D_i(a)^{-1} = I_n+abE_{ij} = T_{ij}(ab),$$

et donc

$$D_i(a)T_{ij}(b)D_i(a)^{-1}T_{ij}(b)^{-1} = T_{ij}(ab)T_{ij}(-b) = T_{ij}((a-1)b).$$

Lorsque b varie dans K, le scalaire $(a-1)b$ décrit aussi K. Donc toute matrice de transvection est un commutateur. Par suite, le groupe engendré par les commutateurs de $GL_n(K)$ est égal à $SL_n(K)$. ◁

Lorsque K *est le corps à deux éléments le résultat reste vrai pour* $n \geqslant 3$ *mais pas pour* $n = 2$. *On peut aussi montrer que le groupe dérivé de* $SL_n(K)$ *est* $SL_n(K)$ *sauf si* $n = 2$ *et* $\mathrm{Card}(K) \leqslant 3$.

3.9. Commutateur égal à $-I_2$

Trouver une condition nécessaire et suffisante sur le corps K pour qu'il existe $(A, B) \in SL_2(K)^2$ tel que $-I_2 = ABA^{-1}B^{-1}$.

(ENS Ulm)

▷ **Solution.**

Montrer qu'une matrice inversible M est un commutateur, c'est-à-dire de la forme $ABA^{-1}B^{-1}$ avec A, B inversibles, revient à prouver qu'elle s'écrit $M = AC$ avec C semblable à A^{-1}. Il est alors facile de montrer

que la matrice $-I_2$ est un commutateur de $GL_2(K)$ pour tout corps K puisqu'il suffit d'écrire $-I_2 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$ les deux matrices de ce produit étant clairement semblables entre elles et semblables à leurs inverses. La difficulté de l'exercice tient au fait qu'on veut un commutateur de $SL_2(K)$: il faut non seulement que A soit de déterminant 1 (ce qui n'est pas le cas ci-dessus) mais aussi que la matrice de passage B telle que $C = BA^{-1}B^{-1}$ soit de déterminant 1.

• On peut noter que les corps de caractéristique 2 conviennent. En effet, dans ce cas $-I_2 = I_2$ et il suffit de prendre $A = B = I_2$. On supposera dans la suite que la caractéristique de K n'est pas 2. On va chercher une condition nécessaire en supposant qu'il existe un couple $(A, B) \in SL_2(K)^2$ tel que $-I_2 = ABA^{-1}B^{-1}$. Cette relation est équivalente à $AB + BA = 0$, forme qui a le mérite de nous éviter de calculer les inverses. On peut aussi noter que $-A = B^{-1}AB$, c'est-à-dire que A et $-A$ sont semblables (et de même pour B). En particulier on a $\operatorname{Tr} A = -\operatorname{Tr} A$ et donc $\operatorname{Tr} A = 0$ puisque $\operatorname{caract}(K) \neq 2$. De même $\operatorname{Tr} B = 0$. Posons alors $A = \begin{pmatrix} a & b \\ c & -a \end{pmatrix}$ et $B = \begin{pmatrix} x & y \\ z & -x \end{pmatrix}$. La relation $AB + BA = 0$ est équivalente à l'égalité $2ax + bz + cy = 0$ (1). On a par ailleurs $a^2 + bc = -1$ (2) et $x^2 + yz = -1$ (3) car $\det A = \det B = 1$. Supposons $b \neq 0$. On extrait z de la relation (1) en fonction de x et y et on le remplace dans l'égalité (3). Il vient $x^2 - \frac{2a}{b}xy - \frac{c}{b}y^2 = -1$ et comme $\frac{c}{b} = -\frac{1}{b^2} - \frac{a^2}{b^2}$ par (2) on obtient finalement

$$x^2 - \frac{2a}{b}xy + \frac{a^2}{b^2}y^2 + \frac{1}{b^2} = \left(x - \frac{a}{b}y\right)^2 + \frac{1}{b^2} = -1.$$

Autrement dit, pour que A et B existent il est nécessaire que -1 soit une somme de deux carrés d'éléments de K. Dans le cas où $b = 0$ on a $-1 = a^2 = a^2 + 0^2$ qui est encore somme de deux carrés.

• Montrons maintenant que cette condition est suffisante. Supposons que $-1 = \alpha^2 + \beta^2$ où α, β sont deux éléments de K. On cherche une solution aux équations (1), (2) et (3) avec $b = c$ et $y = z$. Prenons $a = \alpha$ et $b = c = \beta$, c'est-à-dire $A = \begin{pmatrix} \alpha & \beta \\ \beta & -\alpha \end{pmatrix}$. La condition (1) s'écrit alors $2(\alpha x + \beta y) = 0$ et il suffit de prendre $x = \beta$ et $y = -\alpha$ pour la remplir. On a alors $B = \begin{pmatrix} \beta & -\alpha \\ -\alpha & -\beta \end{pmatrix}$. Ces deux matrices conviennent.

On peut noter qu'un corps de caractéristique 2 vérifie la condition trouvée puisqu'on a alors $-1 = 1 = 1^2 + 0^2$. On peut donc conclure que, dans tous les cas, $-I_2$ est un commutateur de $SL_2(K)$ si et seulement si -1 est somme de deux carrés. ◁

On peut montrer que si K *est un corps infini, toute matrice de* $\mathrm{SL}_n(\mathrm{K})$ *est un commutateur de* $\mathrm{GL}_n(\mathrm{K})$. *Le lecteur pourra trouver ce résultat sous forme de problème corrigé dans le recueil* GIANELLA, KRUST, TAIEB, TOSEL, *Problèmes choisis de mathématiques supérieures, Scopos 14, Springer-Verlag, 2001, p.192.*

Nous terminons cette série d'exercices sur les sous-groupes par un exercice très difficile issu de la structure de groupe de Lie des groupes linéaires. Pour une introduction en langue française à cette théorie le lecteur pourra consulter le livre de R. MNEIMNE, F. TESTARD, *Introduction à la théorie des groupes de Lie classiques, Hermann, 1986.*

3.10. Sous-groupe discret de $\mathbf{SL_2(\mathbb{R})}$

Soit G un sous-groupe fermé de $\mathrm{SL}_2(\mathbb{R})$ non commutatif tel que $\mathrm{G} = -\mathrm{G}$ et pour tout $\mathrm{M} \in \mathrm{G} \setminus \{\mathrm{I}_2, -\mathrm{I}_2\}$, $|\operatorname{Tr}(\mathrm{M})| > 2$.

Soit $\mathrm{G}_0 = \mathrm{G} \setminus \{\mathrm{I}_2, -\mathrm{I}_2\}$. Montrer que G_0 est fermé.

(**École Polytechnique**)

▷ **Solution.**

• Si $\mathrm{M} \in \mathrm{G}_0$, on a $\chi_\mathrm{M} = \mathrm{X}^2 - (\operatorname{Tr}\mathrm{M})\mathrm{X} + \det\mathrm{M} = \mathrm{X}^2 - (\operatorname{Tr}\mathrm{M})\mathrm{X} + 1$ qui est de discriminant $\Delta = (\operatorname{Tr}\mathrm{M})^2 - 4 > 0$. Donc χ_M est scindé à racines simples ce qui prouve que M est diagonalisable.

Pour montrer que G_0 est fermé, il suffit de prouver que si $(\mathrm{A}_n)_{n\geqslant 0}$ est une suite de G_0 qui converge vers A, alors $\mathrm{A} \in \mathrm{G}_0$. La suite $(\mathrm{A}_n)_{n\geqslant 0}$ étant dans G qui est fermé, nous savons que $\mathrm{A} \in \mathrm{G}$. Il s'agit donc de prouver que $\mathrm{A} \neq \mathrm{I}_2$ et $g \neq -\mathrm{I}_2$ c'est-à-dire qu'il n'existe pas de suite de $\mathrm{G}_0 = \mathrm{G} \setminus \{-\mathrm{I}_2, \mathrm{I}_2\}$ qui converge vers I_2 ou $-\mathrm{I}_2$. On dit alors que I_2 et $-\mathrm{I}_2$ sont des points isolés dans G.

Pour cela raisonnons par l'absurde et supposons qu'il existe une suite $(\mathrm{A}_n)_{n\geqslant 0}$ dans G_0 qui converge vers I_2 ou $-\mathrm{I}_2$. Comme $\mathrm{G} = -\mathrm{G}$, on peut même supposer $\lim\limits_{n\to+\infty} \mathrm{A}_n = \mathrm{I}_2$ quitte à remplacer $(\mathrm{A}_n)_{n\geqslant 0}$ par $(-\mathrm{A}_n)_{n\geqslant 0}$.

La théorie sous-jacente à ce problème est celle des groupes et des algèbres de Lie qui permet de relier des propriétés d'un sous-groupe de $\mathrm{GL}(\mathrm{E})$ à celles d'une sous-algèbre de Lie de $\mathcal{L}(\mathrm{E})$ par l'intermédiaire de la fonction exp. Nous n'allons bien sûr éviter d'en exposer tous les tenants et aboutissants pour ne garder que les maillons nécessaires à notre problème.

Quand on cherche des sous-groupes non triviaux de $\mathrm{SL}_2(\mathbb{R})$, on pense évidemment à $\mathrm{SO}(2)$ mais celui-ci est essentiellement composé d'endo-

morphismes non diagonalisables. Il vient ensuite

$$\mathrm{H} = \left\{ \begin{pmatrix} \lambda & 0 \\ 0 & 1/\lambda \end{pmatrix},\ \lambda \in \mathbb{R}_+^* \right\}.$$

• On va essayer de voir si G ne contiendrait pas un sous-groupe isomorphe au sous-groupe abélien H de $\mathrm{SL}_2(\mathbb{R})$. Ce groupe H est isomorphe à $(\mathbb{R}, +)$ par l'application

$$t \in \mathbb{R} \longmapsto \exp\left(t \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\right) = \begin{pmatrix} e^t & 0 \\ 0 & e^{-t} \end{pmatrix} \in \mathrm{H}.$$

Pour cela, il nous faut faire intervenir l'application logarithme définie sur les matrices. La situation est compliquée par le fait que l'exponentielle n'est pas injective d'où l'impossibilité de la définir comme réciproque. Utilisons plutôt la série définissant le logarithme. Comme nous le verrons plus loin, il peut être utile de se placer sur $\mathbb{C}$. On choisit une norme sur $\mathbb{C}^2$ et on pose pour $\mathrm{A} \in \mathcal{M}_2(\mathbb{C})$.

$$\|\mathrm{A}\| = \sup_{\|x\| \leqslant 1} \|\mathrm{A}x\|.$$

Si $(\mathrm{A}, \mathrm{B}) \in \mathcal{M}_2(\mathbb{C})^2$, $\|\mathrm{AB}\| \leqslant \|\mathrm{A}\|\|\mathrm{B}\|$. On vérifie aisément que le rayon spectral $\rho(\mathrm{A}) = \max_{\lambda \in \mathrm{sp}(\mathrm{A})} |\lambda|$ est majoré par $\|\mathrm{A}\|$. Pour $\mathrm{A} \in \mathrm{B}(0,1)$, on pose alors

$$\ln(\mathrm{I}_2 + \mathrm{A}) = \sum_{n=0}^{+\infty} (-1)^{n+1} \frac{\mathrm{A}^n}{n}.$$

Cette série converge car elle est absolument convergente (en effet, $\|\mathrm{A}\| < 1$ et $\left\|(-1)^{n+1}\frac{\mathrm{A}^n}{n}\right\| \leqslant \frac{1}{n}\|\mathrm{A}^n\| \leqslant \|\mathrm{A}\|^n$ pour $n \geqslant 1$) et $\mathcal{M}_2(\mathbb{C})$ est complet. On a donc défini le logarithme sur $\mathrm{B}(\mathrm{I}_2, 1)$.

Lemme. *Pour* $\mathrm{A} \in \mathcal{M}_2(\mathbb{C})$ *telle que* $\|\mathrm{A} - \mathrm{I}_2\| < 1$ *on a* $\exp(\ln \mathrm{A}) = \mathrm{A}$.

Démonstration. On imagine que le résultat est facile lorsque A est diagonale. Vérifions le résultat pour A diagonalisable. Supposons donc qu'il existe $\mathrm{P} \in \mathrm{GL}_2(\mathbb{C})$ et $(\lambda, \mu) \in \mathbb{C}^2$ tels que

$$\mathrm{A} = \mathrm{P}^{-1} \begin{pmatrix} \lambda & 0 \\ 0 & \mu \end{pmatrix} \mathrm{P}.$$

Alors $\mathrm{sp}(\mathrm{A} - \mathrm{I}_2) = \{\lambda - 1, \mu - 1\}$ et comme $\|\mathrm{A} - \mathrm{I}_2\| < 1$, $|\lambda - 1| < 1$ et $|\mu - 1| < 1$. On a ainsi pour $\mathrm{N} \geqslant 1$,

$$\sum_{n=1}^{N}(-1)^{n+1}\frac{(A-I_2)^n}{n} = P^{-1}\left(\sum_{n=1}^{N}\frac{(-1)^{n+1}}{n}\begin{pmatrix}\lambda-1 & 0\\ 0 & \mu-1\end{pmatrix}^n\right)P$$

$$= P^{-1}\begin{pmatrix}\sum_{n=1}^{N}(-1)^{n+1}\frac{(\lambda-1)^n}{n} & 0\\ 0 & \sum_{n=1}^{N}(-1)^{n+1}\frac{(\mu-1)^n}{n}\end{pmatrix}P$$

$$\xrightarrow[N\to+\infty]{} P^{-1}\begin{pmatrix}\ln\lambda & 0\\ 0 & \ln\mu\end{pmatrix}P$$

puisque la multiplication par P et P^{-1} est continue. Ainsi, on a

$$\exp\ln A = \exp\left[P^{-1}\begin{pmatrix}\ln\lambda & 0\\ 0 & \ln\mu\end{pmatrix}P\right] = P^{-1}\exp\left[\begin{pmatrix}\ln\lambda & 0\\ 0 & \ln\mu\end{pmatrix}\right]P$$

$$= P^{-1}\begin{pmatrix}\exp(\ln\lambda) & 0\\ 0 & \exp(\ln\mu)\end{pmatrix}P = P^{-1}\begin{pmatrix}\lambda & 0\\ 0 & \mu\end{pmatrix}P = A.$$

Il reste à vérifier le résultat pour A quelconque. Pour cela, nous allons utiliser la densité des matrices diagonalisables (voir l'exercice 2.58) : dans $\mathcal{M}_2(\mathbb{C})$, les matrices diagonalisables forment une partie dense (c'est pour cela que l'on a pris le soin de se placer sur $\mathbb{C}$, le résultat tombant en défaut sur $\mathbb{R}$). Il existe $(A_n)_{n\in\mathbb{N}}$ une suite de $\mathcal{M}_2(\mathbb{C})$ formée de matrices diagonalisables avec $\lim_{n\to+\infty} A_n = A$. Pour n assez grand, A_n est dans l'ouvert $B(I_2,1)$ qui contient A. Quitte à tronquer cette suite, on peut supposer que pour $n \geqslant 0$, $\|A_n - I_2\| < 1$. Soit $r \in]0,1[$. Montrons que $M \longmapsto \ln M$ est continue sur $B(I_2,r)$. Si $M \in B(I_2,r)$, $\left\|(-1)^{n+1}\frac{(M-I_2)^n}{n}\right\| \leqslant \frac{r^n}{n}$ qui est le terme général d'une série convergente. La série du logarithme est donc normalement convergente, donc la série est continue. La continuité étant une propriété locale, le logarithme est continu sur $B(I_2,1)$. De même, si $r>0$, et $\|M\| < r$, $\left\|\frac{M^n}{n!}\right\| \leqslant \frac{r^n}{n!}$ est le terme général d'une série convergente. La série de l'exponentielle est normalement convergente, donc elle est continue sur $B(0,r)$ et finalement sur $\mathcal{M}_2(\mathbb{C})$. Ainsi, on a

$$\ln A_n \xrightarrow[n\to+\infty]{} \ln A, \quad \text{et} \quad A_n = \exp\ln A_n \xrightarrow[n\to+\infty]{} \exp\ln A.$$

Par unicité de la limite $A = \exp\ln A$. Le lemme est prouvé.

• Établissons maintenant que si I_2 est un point d'accumulation de G, G contient un sous-groupe isomorphe à $(\mathbb{R},+)$.

Lemme. *Il existe* $H \in \mathcal{M}_2(\mathbb{R})$ *non nul tel que pour tout* $t \in \mathbb{R}$,

$$\exp(tH) \in G.$$

Démonstration. Pour n assez grand, on a $\|A_n - I_2\| < 1$. Dans ces conditions, $\ln A_n$ est défini et $\ln A_n \neq 0$ car $\exp \ln A_n = A_n \neq I_2$. On pose $H_n = \frac{\ln A_n}{\|\ln A_n\|}$. Étant donné que nous sommes en dimension finie et que $\|H_n\| = 1$, d'après le théorème de Bolzano-Weierstrass, il existe une sous-suite $(H_{\varphi(n)})_{n\geqslant 0}$ ($\varphi : \mathbb{N} \longrightarrow \mathbb{N}$ strictement croissante) convergente vers une matrice H. Montrons que H convient. Soit $t \in \mathbb{R}$. On a

$$\exp(tH) = \lim_{n\to+\infty} \exp(tH_{\varphi(n)}) = \lim_{n\to\infty} \exp\left[\frac{t}{\|\ln A_{\varphi(n)}\|} \ln A_{\varphi(n)}\right]$$

Soit $P_n = \exp\left[E\left(\frac{t}{\|\ln A_{\varphi(n)}\|}\right) \ln A_{\varphi(n)}\right]$. On a

$$\left\| \left(\frac{t}{\|\ln A_{\varphi(n)}\|} - E\left(\frac{t}{\|\ln A_{\varphi(n)}\|}\right)\right) \ln A_{\varphi(n)} \right\| \leqslant \|\ln A_{\varphi(n)}\| \xrightarrow[n\to+\infty]{} \|\ln I_2\| = 0$$

par continuité de ln. Ainsi par continuité de exp, on obtient

$$\exp\left[-\left(\frac{t}{\|\ln A_{\varphi(n)}\|} - E\left(\frac{t}{\|\ln A_{\varphi(n)}\|}\right)\right) \ln A_{\varphi(n)}\right] \xrightarrow[n\to+\infty]{} \exp 0 = I_2.$$

Ainsi, P_n s'écrit

$$P_n = \exp(tH_{\varphi(n)}) \exp\left(-\left(\frac{t}{\|\ln A_{\varphi(n)}\|} - E\left(\frac{t}{\|\ln A_{\varphi(n)}\|}\right)\right) \ln A_{\varphi(n)}\right)$$

et P_n converge vers $e^{tH} I_2 = e^{tH}$. Or $P_n = [A_{\varphi(n)}]^{E\left(\frac{t}{\|\ln A_{\varphi(n)}\|}\right)} \in G$ et comme G est fermé, $e^{tH} = \lim_{n\to+\infty} P_n \in G$. Le lemme est prouvé.♢

- Montrons que H est diagonalisable. Comme $\exp H \in G$, on a

$$1 = \det \exp H = \exp \operatorname{Tr} H,$$

et la trace de H étant réelle, elle est nulle. Il existe $k \in \mathbb{C}$ tel que le spectre complexe de H soit $\{k, -k\}$. En fait, k est nécessairement réel. En effet, pour tout t, le spectre de $\exp tH$ est $\{e^{tk}, e^{-tk}\} \subset \mathbb{R}$ car les éléments de G sont diagonalisables. En dérivant par rapport à t, e^{tk}, on a $ke^{tk} \in \mathbb{R}$ et k est nécessairement réel.

Si $k = 0$, alors H est nilpotente et $H^2 = 0$. Par conséquent, $\exp H = I_2 + H$ et $\operatorname{Tr} \exp H = \operatorname{Tr} I_2 + \operatorname{Tr} H = 2 + 0 = 2$. Comme $\exp H \in G$, $\exp H = +I_2$ i.e. $H = 0$ ou $H = -2I_2$. Ces deux cas sont manifestement exclus puisque H est de norme 1. Donc k est non nul et H est diagonalisable.

Quitte à changer G par son image par un morphisme de conjugaison, on peut supposer

$$\mathrm{H} = \begin{pmatrix} k & 0 \\ 0 & -k \end{pmatrix}$$

et dans ces conditions pour tout $t \in \mathbb{R}$

$$\exp t\mathrm{H} = \begin{pmatrix} e^{kt} & 0 \\ 0 & e^{-kt} \end{pmatrix} \subset \mathrm{G}$$

et finalement, pour tout $\lambda \in \mathbb{R}^*$,

$$\mathrm{A}_\lambda = \begin{pmatrix} \lambda & 0 \\ 0 & 1/\lambda \end{pmatrix} \in \mathrm{G}.$$

• Comme G est non commutatif, il existe $\mathrm{M} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \subset \mathrm{G}$ tel que $b \neq 0$ ou $c \neq 0$. Considérons une telle matrice. Montrons qu'alors $b \neq 0$ et $c \neq 0$. Imaginons un instant que $b = 0$. Alors :

$$\mathrm{M} = \begin{pmatrix} a & 0 \\ c & 1/a \end{pmatrix}$$

Mais alors le produit $\mathrm{MA}_{1/a} = \begin{pmatrix} 1 & 0 \\ c/a & 1 \end{pmatrix}$ est dans G, ce qui est impossible car cette matrice est distincte de I_2 et $-\mathrm{I}_2$ et pourtant sa trace vaut 2. On conclut que $b \neq 0$ et $c \neq 0$.

• Soit $\lambda > 0$. Un petit calcul nous donne

$$\mathrm{A}_\lambda \mathrm{M} \mathrm{A}_\lambda = \begin{pmatrix} a\lambda^2 & b \\ c & d/\lambda^2 \end{pmatrix}.$$

À l'aide d'une étude de fonction, on constate que si $ad \leqslant 0$, il existe $\lambda > 0$ avec $0 \leqslant |a\lambda^2 + d/\lambda^2| < 2$. Or cette inégalité est impossible pour $\mathrm{Tr}(\mathrm{A}_\lambda \mathrm{M} \mathrm{A}_\lambda) = a\lambda^2 + d/\lambda^2$ car $\mathrm{A}_\lambda \mathrm{M} \mathrm{A}_\lambda \subset \mathrm{G}$. Par conséquent $ad > 0$.

• Comme $\mathrm{G} = -\mathrm{G}$, on peut supposer $a > 0$. Alors $d > 0$. En prenant $\lambda = \sqrt{\sqrt{d/a}}$, on prouve que

$$\begin{pmatrix} \sqrt{ad} & b \\ c & \sqrt{ad} \end{pmatrix} \in \mathrm{G}.$$

On peut donc supposer $a = d > 0$. On a alors

$$\mathrm{M} = \begin{pmatrix} a & b \\ c & a \end{pmatrix}$$

et $\operatorname{Tr} M = 2a > 2$ d'où $a > 1$. Comme $1 = \det M = a^2 - bc$, on a $bc = a^2 - 1 > 0$.

• Calculons $M' = M^2 A_\lambda M^{-1}$ ($\lambda > 0$). On a

$$\begin{aligned} M' &= \begin{pmatrix} a^2+bc & 2ab \\ 2ac & a^2+bc \end{pmatrix} A_\lambda M^{-1} \\ &= \begin{pmatrix} 2a^2-1 & 2ba \\ 2ac & 2a^2-1 \end{pmatrix} \begin{pmatrix} \lambda & 0 \\ 0 & 1/\lambda \end{pmatrix} M^{-1} \\ &= \begin{pmatrix} (2a^2-1)\lambda & 2ab/\lambda \\ 2ac\lambda & (2a^2-1)/\lambda \end{pmatrix} \begin{pmatrix} a & -b \\ -c & a \end{pmatrix} \\ &= \begin{pmatrix} \times & -b(2a^2-1)\lambda + 2a^2 b/\lambda \\ 2a^2 c\lambda - c(2a^2-1)/\lambda & \times \end{pmatrix} \\ &= \begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix} \in \mathrm{G}. \end{aligned}$$

Pour que $c' = 0$, il suffit que $2a^2\lambda^2 = 2a^2 - 1 > 2 - 1 = 1$. Prenons $\lambda = \lambda_0$ avec $\lambda_0 = \sqrt{\frac{2a^2-1}{2a^2}} < 1$. Alors $c' = 0$. Mais alors $b' \neq 0$ (en effet $b' = 0$ impose $\lambda^2(2a^2-1) = 2a^2$, puisque $b \neq 0$, et $\lambda = \sqrt{2a^2/(2a^2-1)} > 1$ distinct de λ_0). Or nous avons déjà démontré un peu plus haut que si $M' = \begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix} \in \mathrm{G}$ avec $b' \neq 0$ alors, nécessairement $c' \neq 0$: et bien voilà la contradiction tant attendue ! ◁

Nous abordons maintenant une série d'exercices consacrés à la recherche de morphismes de groupes ayant pour but ou pour source un groupe linéaire.

3.11. Sous-groupes à un paramètre de $\mathrm{GL}_n(\mathbb{C})$

1. Trouver les morphismes dérivables de $(\mathbb{R}, +)$ dans $(\mathrm{GL}_n(\mathbb{C}), \times)$.

2. En déduire les morphismes continus de $(\mathbb{R}, +)$ dans $(\mathrm{GL}_n(\mathbb{C}), \times)$.

(École Polytechnique)

▷ **Solution.**

1. Soit $f : \mathbb{R} \to \mathrm{GL}_n(\mathbb{C})$ un morphisme dérivable. Pour $(s, t) \in \mathbb{R}^2$, on a

$$f(s+t) = f(s)f(t).$$

En dérivant par rapport à t, on obtient, pour $(s,t) \in \mathbb{R}^2$, $f'(s+t) = f(s)f'(t)$ puis $f'(s) = f(s)f'(0)$ en prenant $t = 0$. Posons $\mathrm{A} = f'(0)$ et considérons la fonction $g : \mathbb{R} \longrightarrow \mathcal{M}_n(\mathbb{C})$ définie par $g(s) = f(s)e^{-s\mathrm{A}}$. Les fonctions f et $s \longmapsto e^{-s\mathrm{A}}$ sont dérivables, donc g également. Pour tout $s \in \mathbb{R}$,

$$g'(s) = f'(s)e^{-s\mathrm{A}} + f(s)(-\mathrm{A}e^{-s\mathrm{A}}) = (f'(s) - f(s)\mathrm{A})e^{-s\mathrm{A}} = 0.$$

L'application g est donc constante. De plus, $g(0) = f(0) = \mathrm{I}_n$, donc, pour tout $s \in \mathbb{R}$,

$$g(s) = f(s)e^{-s\mathrm{A}} = \mathrm{I}_n.$$

On en déduit que pour tout $s \in \mathbb{R}$, $f(s) = e^{s\mathrm{A}}$.

Réciproquement, A étant une matrice quelconque de $\mathcal{M}_n(\mathbb{C})$, on vérifie que l'application

$$f : t \in \mathbb{R} \longmapsto e^{t\mathrm{A}} \in \mathrm{GL}_n(\mathbb{C})$$

est un morphisme dérivable de $(\mathbb{R}, +)$ dans $(\mathrm{GL}_n(\mathbb{C}), \times)$. Cela résulte immédiatement des propriétés de l'exponentielle.

2. Soit $f : \mathbb{R} \to \mathrm{GL}_n(\mathbb{C})$ un morphisme continu. On va montrer en fait que f est nécessairement dérivable. L'idée est d'utiliser une primitive de f. Posons pour $x \in \mathbb{R}$, $\mathrm{F}(x) = \int_0^x f(t)dt$. La fonction F est de classe $\mathcal{C}^1$ sur $\mathbb{R}$. Imaginons qu'il existe $a > 0$ tel que la matrice $\mathrm{F}(a)$ soit inversible. On a $f(x+u) = f(x)f(u)$ pour tout $(x,u) \in \mathbb{R}^2$. Intégrons cette égalité par rapport à u sur le segment $[0,a]$. Il vient

$$\int_0^a f(x+u)\mathrm{d}u = f(x)\int_0^a f(u)\mathrm{d}u = f(x)\mathrm{F}(a).$$

Or, $g(x) = \displaystyle\int_0^a f(x+u)\mathrm{d}u = \int_x^{x+a} f(t)\mathrm{d}t$ est une fonction dérivable de x car f est continue. Comme on a $f(x) = g(x)\mathrm{F}(a)^{-1}$, f est aussi dérivable et on peut conclure avec la question précédente.

Il ne reste plus qu'à prouver l'existence de a. Pour cela on observe que $\mathrm{F}'(0) = f(0) = \mathrm{I}_n$. Comme le taux d'accroissement $\frac{1}{t}\mathrm{F}(t)$ tend vers I_n lorsque t tend vers 0^+, et comme le groupe $\mathrm{GL}_n(\mathbb{C})$ est ouvert, on est certain que $\frac{1}{t}\mathrm{F}(t)$ est inversible pour t assez petit. Il en est alors de même de $\mathrm{F}(t)$.

Conclusion. Les morphismes continus de $(\mathbb{R}, +)$ dans $(\mathrm{GL}_n(\mathbb{C}), \times)$ sont les applications $t \longmapsto e^{t\mathrm{A}}$, où A est une matrice quelconque. ◁

Dans l'énoncé suivant on remplace le groupe additif $\mathbb{R}$ par le groupe des nombres complexes de module 1 qui est isomorphe au quotient $\mathbb{R}/\mathbb{Z}$. On pourra utiliser le résultat de l'exercice précédent.

3.12. Morphismes continus de S^1 dans $\mathrm{GL}_n(\mathbb{R})$

Soit $\varphi : (\mathrm{S}^1, \times) \to (\mathrm{GL}_n(\mathbb{R}), \times)$ un morphisme de groupes continu où S^1 désigne le groupe des nombres complexes de module 1.

1. Montrer que pour tout $z \in \mathrm{S}^1$, $\det(\varphi(z)) = 1$.

2. Montrer que pour tout $z \in \mathrm{S}^1$, les valeurs propres complexes de $\varphi(z)$ sont de module 1.

3. Achever la description de φ.

(ENS Ulm)

⌙ **Solution.**

1. Posons $\psi = \det \circ \varphi$. Il s'agit d'un morphisme continu de groupes de S^1 dans $\mathbb{R}^*$. On va montrer que ψ est trivial. Comme S^1 est connexe, que ψ est continu et que $\psi(1) = 1$, $\psi(\mathrm{S}^1)$ est un intervalle inclus dans $\mathbb{R}_+^*$. Comme S^1 est compact, $\psi(\mathrm{S}^1)$ est même un segment. De plus, comme ψ est morphisme, $\psi(\mathrm{S}^1)$ est un sous-groupe de $\mathbb{R}_+^*$. Or, le seul segment de $\mathbb{R}_+^*$ qui est un sous-groupe est le singleton $\{1\}$ (un sous-groupe non trivial de $\mathbb{R}_+^*$ n'est pas borné). On a donc $\varphi(\mathrm{S}^1) \subset \mathrm{SL}_n(\mathbb{R})$.

2. On munit $\mathcal{M}_n(\mathbb{C})$ d'une norme triple quelconque associée à une norme de $\mathbb{C}^n$. Comme S^1 est compact, et φ continu, le groupe image $\varphi(\mathrm{S}^1)$ est compact et en particulier borné. Il existe donc une constante $\mathrm{M} > 0$ telle que $\|\varphi(z)\| \leqslant \mathrm{M}$ pour tout $z \in \mathrm{S}^1$. Or si λ une valeur propre complexe de $\varphi(z)$ on a $|\lambda| \leqslant \|\varphi(z)\|$. Ainsi, l'ensemble des valeurs propres des éléments de $\varphi(\mathrm{S}^1)$ est borné. De plus, si λ est dans le spectre de $\varphi(z)$ alors, pour tout $p \in \mathbb{Z}$, λ^p est valeur propre de $\varphi(z)^p = \varphi(z^p) \in \varphi(\mathrm{S}^1)$. La suite $(\lambda^p)_{p\in\mathbb{Z}}$ est donc bornée et cela impose que $|\lambda| = 1$.

3. L'application $\psi : t \longmapsto \varphi(e^{it})$ est un morphisme continu de $(\mathbb{R}, +)$ dans $(\mathrm{GL}_n(\mathbb{R}), \times)$. D'après l'exercice 3.11, il existe $\mathrm{A} \in \mathcal{M}_n(\mathbb{C})$ telle que, pour tout $t \in \mathbb{R}$, $\psi(t) = e^{t\mathrm{A}}$. Comme ψ est à valeurs dans $\mathrm{GL}_n(\mathbb{R})$, la matrice $\mathrm{A} = \psi'(0)$ est réelle. Par construction, ψ est périodique de période 2π. On a donc, pour tout réel t,

$$e^{t\mathrm{A}} = e^{2\pi\mathrm{A}+t\mathrm{A}} = e^{2\pi\mathrm{A}}e^{t\mathrm{A}}$$

et donc $e^{2\pi\mathrm{A}} = \mathrm{I}_n$.

La résolution de l'équation $e^{\mathrm{M}} = \mathrm{I}_n$ a été menée dans l'exercice 2.31 mais nous allons la refaire ici. Les valeurs propres complexes de $e^{2\pi\mathrm{A}}$ sont les $e^{2\pi\lambda}$ avec $\lambda \in \mathrm{Sp}\,\mathrm{A}$. Il faut $e^{2\pi\lambda} = 1$ donc $\lambda \in i\mathbb{Z}$.

Montrons que A est diagonalisable dans $\mathbb{C}$. D'après la décomposition de Dunford, il existe D diagonalisable et N nilpotente telles que $\mathrm{A} = \mathrm{D} + \mathrm{N}$ et $\mathrm{DN} = \mathrm{ND}$. Puisque N et D commutent, on a $e^{2\pi\mathrm{A}} = e^{2\pi\mathrm{N}}e^{2\pi\mathrm{D}}$. La matrice $e^{2\pi\mathrm{D}}$ est diagonalisable. Comme les valeurs propres de D sont

égales à celles de A, les valeurs propres de $e^{2\pi D}$ sont celles de $e^{2\pi A}$ soit 1. Ainsi $e^{2\pi D} = I_n$ et $e^{2\pi N} = I_n$. Supposons que N n'est pas la matrice nulle. On a alors $\operatorname{Ker} N \neq \operatorname{Ker} N^2$. Soit $X \in \operatorname{Ker} N^2 \setminus \operatorname{Ker} N$. Il vérifie $e^{2\pi N}X = X + 2\pi NX \neq X$, ce qui contredit l'égalité $e^{2\pi N} = I_n$. On a donc $N = 0$ et A diagonalisable dans C.

Les valeurs propres de A non nulles sont deux à deux conjuguées. Il existe $k_1, \ldots, k_r$ dans $\mathbb{Z}^*$ et $P \in GL_n(\mathbb{C})$ tels que

$$A = P \operatorname{Diag}(ik_1, -ik_1, \ldots, ik_r, -ik_r, 0, \ldots, 0)P^{-1}.$$

On obtient pour tout réel t.

$$e^{tA} = P \operatorname{Diag}(e^{itk_1}, e^{-itk_1}, \ldots, e^{itk_r}, e^{-itk_r}, 1, \ldots, 1)P^{-1}.$$

On en déduit qu'il existe une matrice $Q \in GL_n(\mathbb{R})$, indépendante de t telle que

$$e^{tA} = Q \begin{pmatrix} R_{tk_1} & & & & & \\ & \ddots & & & & \\ & & R_{tk_r} & & & \\ & & & 1 & & \\ & & & & \ddots & \\ & & & & & 1 \end{pmatrix} Q^{-1},$$

où $R_\theta = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$.

Réciproquement toute application de la forme

$$e^{it} \longmapsto Q \begin{pmatrix} R_{tk_1} & & & & & \\ & \ddots & & & & \\ & & R_{tk_r} & & & \\ & & & 1 & & \\ & & & & \ddots & \\ & & & & & 1 \end{pmatrix} Q^{-1}$$

est un morphisme continu de S_1 dans $GL_n(\mathbb{R})$. D'une part une telle application est bien définie car pour $k \in \mathbb{Z}$, R_{tk} ne dépend que de t modulo 2π. De plus c'est un morphisme de groupes car $R_{(t+t')k} = R_{tk}R_{t'k}$ pour t, t' réels et $k \in \mathbb{Z}$. Il est continu car pour tout $(t, t') \in \mathbb{R}^2$ et tout $k \in \mathbb{Z}$, $|e^{kit} - e^{kit'}| \leqslant |k||e^{it} - e^{it'}|$ et donc $|\cos(kt) - \cos(kt')| \leqslant |k||e^{it} - e^{it'}|$ et $|\sin(kt) - \sin(kt')| \leqslant |k||e^{it} - e^{it'}|$. ◁

3.13. Morphismes de $(\mathbb{R}^*, \times)$ dans $(\mathrm{GL}_n(\mathbb{C}), \times)$

Déterminer les morphismes $f : t \longmapsto (f_{ij}(t))_{1\leqslant i,j\leqslant n}$ de $(\mathbb{R}^*, \times)$ dans $(\mathrm{GL}_n(\mathbb{C}), \times)$ tels que chaque f_{ij} soit une fraction rationnelle. **(ENS Ulm)**

▷ **Solution.**

• On va commencer par se ramener au cas où les fonctions f_{ij} sont polynomiales. Soit p un polynôme unitaire de degré minimal tel que, pour tout $(i,j) \in [\![1,n]\!]^2$, $g_{ij} = pf_{ij} \in \mathbb{C}[X]$. Soit $(i,j) \in [\![1,n]\!]^2$. On a, pour tout t et u réels non nuls,

$$f_{ij}(tu) = \sum_{k=1}^{n} f_{ik}(t) f_{kj}(u).$$

En dérivant par rapport à u, on obtient

$$t f'_{ij}(tu) = \sum_{k=1}^{n} f_{ik}(t) f'_{kj}(u)$$

puis en prenant $u = 1$,

$$t f'_{ij}(t) = \sum_{k=1}^{n} f_{ik}(t) f'_{kj}(1).$$

En remplaçant f_{ij} par $\frac{g_{ij}}{p}$, il vient

$$t\frac{g'_{ij}(t)}{p(t)} - t\frac{p'(t)g_{ij}(t)}{(p(t))^2} = \sum_{k=1}^{n} \frac{g_{ik}(t)}{p(t)} f'_{kj}(1)$$

et donc

$$t\frac{p'(t)g_{ij}(t)}{p(t)} = tg'_{ij}(t) - \sum_{k=1}^{n} g_{ik}(t) f'_{kj}(1).$$

On en déduit que p divise $Xp'g_{ij}$. Si α est une racine de p d'ordre de multiplicité $m \geqslant 1$, alors α est racine d'ordre $m-1$ de p'. Donc si α est non nul, il est racine de g_{ij} et cela pour tout couple (i,j). C'est impossible car $X - \alpha$ diviserait tous les polynômes g_{ij} et pour tout $(i,j) \in [\![1,n]\!]^2$, $\frac{p}{X-\alpha} f_{ij} \in \mathbb{C}[X]$, ce qui est contraire à la définition de p. Ainsi 0 est la seule racine (éventuelle) de p. Il existe donc $m \in \mathbb{N}$ tel que $p(t) = t^m$. On a alors, pour $(i,j) \in [\![1,n]\!]^2$, t et u dans $\mathbb{R}^*$,

$$g_{ij}(tu) = p(tu) f_{ij}(tu) = t^m u^m \sum_{k=1}^{n} f_{ik}(t) f_{kj}(u) = \sum_{k=1}^{n} g_{ik}(t) g_{kj}(u).$$

Autrement dit, l'application $g : t \mapsto (g_{ij}(t))_{1 \leqslant i,j \leqslant n}$ est un morphisme de $\mathbb{R}^*$ dans $\mathrm{GL}_n(\mathbb{C})$ dont les coefficients sont tous polynomiaux en t.

• Notons d le maximum des degrés des polynômes g_{ij}, il existe des matrices $\mathrm{A}_0, \ldots, \mathrm{A}_d$ dans $\mathcal{M}_n(\mathbb{C})$ telles que, pour tout $t \in \mathbb{R}^*$,

$$g(t) = \sum_{k=0}^{d} t^k \mathrm{A}_k.$$

Une telle écriture est unique car, en posant $\mathrm{A}_k = (a_{i,j}^k)$, elle équivaut pour tout $(i,j) \in [\![1,n]\!]^2$ à

$$g_{ij}(t) = \sum_{k-0}^{d} a_{ij}^k t^k,$$

pour tout $t \neq 0$ et donc à l'égalité formelle $g_{ij} = \sum_{k=0}^{d} a_{ij}^k \mathrm{X}^k$.

On a, pour t et u dans $\mathbb{R}^*$,

$$g(tu) = \sum_{k=0}^{d} t^k u^k \mathrm{A}_k = g(t)g(u) = \sum_{k=0}^{d} t^k \mathrm{A}_k g(u)$$

et donc, par unicité de l'écriture, pour $u \neq 0$ et $k \in [\![1,n]\!]$

$$u^k \mathrm{A}_k = \mathrm{A}_k g(u) = \sum_{l=0}^{d} u^l \mathrm{A}_k \mathrm{A}_l.$$

On en déduit que $\mathrm{A}_k^2 = \mathrm{A}_k$ et que $\mathrm{A}_k \mathrm{A}_l = 0$ si $l \neq k$. On a de plus $g(1) = \sum_{k=0}^{d} \mathrm{A}_k = \mathrm{I}_n$.

• Réciproquement, soit n et d des entiers naturels, $\mathrm{A}_0, \ldots, \mathrm{A}_d$ des éléments de $\mathcal{M}_n(\mathbb{C})$ tels que, pour tout $(k,l) \in [\![1,n]\!]^2$, $\mathrm{A}_k^2 = \mathrm{A}_k$ et $\mathrm{A}_k \mathrm{A}_l = 0$ si $k \neq l$, et $\sum_{k-0}^{d} \mathrm{A}_k = \mathrm{I}_n$. Considérons la fonction $f : \mathbb{R}^* \longrightarrow \mathcal{M}_n(\mathbb{C})$ définie par

$$f(t) = \frac{1}{t^n} \sum_{k=0}^{d} t^k \mathrm{A}_k.$$

On vérifie facilement que $f(tu) = f(t)f(u)$ pour tout t et u non nuls. Comme $f(1) = \mathrm{I}_n$ par hypothèse, on a, pour t non nul, $f(t)f\left(\frac{1}{t}\right) = f(1) = \mathrm{I}_n$ donc f est à valeurs dans $\mathrm{GL}_n(\mathbb{C})$ et elle a toutes les propriétés voulues.

• Il reste juste à expliquer comment caractériser les familles de matrices A_k ayant les propriétés ci-dessus. Si elles existent, les A_k sont

des matrices de projection. Pour tout $X \in \mathbb{C}^n$, on a $X = \sum_{k=0}^{d} A_k X$, donc $\mathbb{C}^n = \sum_{k=0}^{d} \operatorname{Im} A_k$. Si, de plus, il existe des éléments $X_0, \ldots, X_d$ de $\mathbb{C}^n$ tels que $X = \sum_{k=0}^{d} A_k X_k$, on a alors $A_l X = A_l X_l$, ce qui montre l'unicité de la décomposition. On a donc $\mathbb{C}^n = \bigoplus_{k=0}^{d} \operatorname{Im} A_k$.

Réciproquement, toute décomposition de $\mathbb{C}^n$ en somme directe de sous-espaces fournit des matrices A_k qui conviennent, chacune de ces matrices étant la matrice d'une projection sur un des sous-espaces, parallèlement à la somme des autres. ◁

Dans l'exercice suivant on s'intéresse à des morphismes dont le groupe linéaire est la source. Un morphisme étant caractérisé par sa restriction à une partie génératrice, on va à nouveau utiliser le résultat de l'exercice 3.1.

3.14. Morphismes de $\mathrm{GL}_n(\mathrm{K})$ dans un groupe abélien fini

Soit $n \geqslant 2$, G un groupe abélien fini et K un corps commutatif. Étudier les morphismes de $\mathrm{GL}_n(\mathrm{K})$ dans G. On étudiera le cas $\mathrm{K} = \mathbb{C}$ puis $\mathrm{K} = \mathbb{R}$ et enfin $\mathrm{K} = \mathbb{Z}/q\mathbb{Z}$ avec q premier.

(ENS Ulm)

▷ **Solution.**

Le groupe G sera noté multiplicativement et son neutre sera noté 1. On rappelle que le groupe linéaire $\mathrm{GL}_n(\mathrm{K})$ est engendré par l'ensemble des matrices de dilatation et des matrices de transvection. Plus précisément, toute matrice M de $\mathrm{GL}_n(\mathrm{K})$ s'écrit $M = B_1 \ldots B_k A B'_1 \ldots B'_l$, où A est la matrice de dilation $\mathrm{Diag}(1, 1, \ldots, 1, \det M)$ et $B_1, \ldots, B_k, B'_1, \ldots, B'_l$ des matrices de transvection (voir l'exercice 3.1).

Soit f un morphisme de $\mathrm{GL}_n(\mathrm{K})$ dans G. On va s'intéresser à l'image par f des matrices de transvection et de dilatation. Plusieurs idées sont possibles.

• Une première idée est d'utiliser le caractère *fini* de G. Si le cardinal de G est p, on a, pour tout $g \in \mathrm{G}$, $g^p = 1$. On en déduit que pour tout $M \in \mathrm{GL}_n(\mathrm{K})$, $f(M^p) = (f(M))^p = 1$. Autrement dit, toutes les puissances p-ièmes sont dans $\operatorname{Ker} f$. On peut alors remarquer que si $i \neq j$, on a

$$(\mathrm{I}_n + \lambda \mathrm{E}_{ij})(\mathrm{I}_n + \lambda' \mathrm{E}_{ij}) = \mathrm{I}_n + (\lambda + \lambda')\mathrm{E}_{ij}.$$

On en déduit que $(\mathrm{I}_n + \lambda \mathrm{E}_{ij})^p = \mathrm{I}_n + p\lambda \mathrm{E}_{ij}$, puis que, si le corps K est de caractéristique nulle, toute matrice de transvection $\mathrm{B} = \mathrm{I}_n + \lambda \mathrm{E}_{ij}$ est une puissance p-ième puisqu'on peut écrire $\mathrm{B} = \left(\mathrm{I}_n + \frac{\lambda}{p}\mathrm{E}_{ij}\right)^p$. On a donc $f(\mathrm{B}) = 1$.

Ainsi, si K est de caractéristique nulle, la restriction de f à $\mathrm{SL}_n(\mathrm{K})$ est triviale.

• En fait, cela est vrai plus généralement. Une seconde idée, est d'exploiter le caractère *abélien* de G. Si $\mathrm{M} = \mathrm{ABA}^{-1}\mathrm{B}^{-1}$ est un commutateur de $\mathrm{GL}_n(\mathrm{K})$ on a $f(\mathrm{M}) = f(\mathrm{A})f(\mathrm{B})f(\mathrm{A})^{-1}f(\mathrm{B})^{-1} = 1$ car G est commutatif. On a vu dans l'exercice 3.8 que si K n'est pas le corps réduit à deux éléments, toute transvection est un commutateur. On retrouve donc de cette manière le fait que la restriction de f à $\mathrm{SL}_n(\mathrm{K})$ est triviale.

• On va maintenant regarder l'image d'une matrice de dilatation. Comme le demande l'énoncé, on regarde d'abord le cas où $\mathrm{K} = \mathbb{C}$. Si $\mathrm{A} = \mathrm{Diag}(1, \ldots, 1, \alpha)$ est une matrice de dilatation et β une racine p-ième de α, on peut écrire $\mathrm{A} = (\mathrm{Diag}(1, \ldots, 1, \beta))^p$. On a donc $f(\mathrm{A}) = 1$ d'après ce qui a été dit plus haut. Comme les matrices de transvection et de dilatation engendrent $\mathrm{GL}_n(\mathbb{C})$, f est le morphisme trivial $f : \mathrm{M} \longmapsto 1$.

• Supposons maintenant que $\mathrm{K} = \mathbb{R}$.

Si p est impair, la démonstration précédente s'applique, puisqu'alors tout nombre réel possède une racine p-ième. On trouve pour seul morphisme le morphisme trivial.

Supposons que p est pair. Si $\alpha > 0$, la matrice $\mathrm{A} = \mathrm{Diag}(1, \ldots, 1, \alpha)$ possède une racine p-ième et $f(\mathrm{A}) = 1$. Toute matrice M de $\mathrm{GL}_n(\mathbb{R})$ s'écrit $\mathrm{M} = \mathrm{B}_1 \ldots \mathrm{B}_k \mathrm{A} \mathrm{B}'_1 \ldots \mathrm{B}'_l$, où A est une matrice de dilation et $\mathrm{B}_1, \ldots, \mathrm{B}_k$, $\mathrm{B}'_1, \ldots, \mathrm{B}'_l$ des matrices de transvection. D'après ce qui précède, on a $f(\mathrm{M}) = f(\mathrm{A})$. Comme $\det \mathrm{M} = \det \mathrm{A} = \alpha$, si $\det \mathrm{M} > 0$, on a $f(\mathrm{M}) = 1$.

Si M et N sont deux matrices de $\mathrm{GL}_n(\mathbb{R})$ de déterminant négatif, on a, puisque M^2 et NM ont un déterminant positif,

$$f(\mathrm{N}) = f(\mathrm{N})f(\mathrm{M}^2) = f(\mathrm{NM}^2) = f(\mathrm{NM})f(\mathrm{M}) = f(\mathrm{M}).$$

Il existe donc $a \in \mathrm{G}$ tel que

$$f(\mathrm{M}) = \begin{cases} 1 & \text{si } \det \mathrm{M} > 0 \\ a & \text{si } \det \mathrm{M} < 0. \end{cases}$$

L'élément a de G vérifie $a^2 = 1$, car si $\det \mathrm{M} < 0$, $1 = f(\mathrm{M}^2) = a^2$.

Réciproquement si a est un élément de G qui vérifie $a^2 = 1$, toute application définie ainsi est un morphisme de $\mathrm{GL}_n(\mathbb{R})$ dans G.

Notons que puisque p est pair, il résulte du lemme de Cauchy (cf. exercice 2.10 du tome 1 d'algèbre) que G possède des éléments a d'ordre 2. Ainsi, il existe des morphismes non triviaux de $\mathrm{GL}_n(\mathbb{R})$ dans G.

• Si K est un corps quelconque distinct de $\mathbb{Z}/2\mathbb{Z}$, on a toujours $f(\mathrm{A}) = 1$ pour toute matrice de transvection A. On obtient donc

$$f(\mathrm{M}) = f(\mathrm{Diag}(1, \ldots, 1, \det \mathrm{M})),$$

de sorte qu'il existe un morphisme φ de $(\mathrm{K}^*, \times)$ dans $(\mathrm{G}, \times)$ tel que $f = \varphi \circ \det$. L'existence d'un morphisme non trivial φ dépend de K et G.

• Traitons encore l'exemple de $\mathrm{K} = \mathbb{Z}/q\mathbb{Z}$ où q est un nombre premier impair. On est amené à chercher les morphismes de $(\mathbb{Z}/q\mathbb{Z})^*$ dans G. Or on sait que le groupe $(\mathbb{Z}/q\mathbb{Z})^*$ est cyclique (voir l'exercice 4.12 du tome 1 d'algèbre). Si u est un générateur de ce groupe, un morphisme φ est parfaitement déterminé par l'élément $a = \varphi(u)$. En effet, pour tout $x \in (\mathbb{Z}/q\mathbb{Z})^*$, il existe $k \in \mathrm{N}$ tel que $x = u^k$ et on a alors $\varphi(x) = a^k$.

Voyons quels sont les $a \in \mathrm{G}$ qui peuvent convenir. Comme $u^{q-1} = 1$ on doit avoir $a^{q-1} = \varphi(u^{q-1}) = \varphi(1) = 1$.

Si p est premier avec $q-1$, on a forcément $a = 1$. En effet, l'ordre de a dans G doit diviser $q-1$ et p. Dans ce cas, le seul morphisme entre $(\mathbb{Z}/q\mathbb{Z})^*$ et G est donc le morphisme trivial.

Si $q-1$ et p ne sont pas premiers entre eux, a doit avoir pour ordre un diviseur du pgcd de p et $q-1$. On peut trouver de tels éléments non triviaux dans G : si m un nombre premier qui divise p et $q-1$, le lemme de Cauchy assure l'existence d'un élément $a \in \mathrm{G}$ d'ordre m. Il vérifie $a^{q-1} = 1$ et l'application φ définie par $\varphi(u^k) = a^k$ est un morphisme non trivial de $(\mathbb{Z}/q\mathbb{Z})^*$ dans G et $f = \varphi \circ \det$, un morphisme non trivial de $\mathrm{GL}_n(\mathrm{K})$ dans G.

• Regardons enfin le cas particulier $\mathrm{K} = \mathbb{Z}/2\mathbb{Z}$. Dans ce cas il n'y a pas de matrice de dilatation (hormis l'identité) et $\mathrm{GL}_n(\mathbb{Z}/2\mathbb{Z}) = \mathrm{SL}_n(\mathbb{Z}/2\mathbb{Z})$. Nous avons indiqué dans la remarque qui suit l'exercice 3.8 que pour $n \geqslant 3$, $\mathrm{GL}_n(\mathbb{Z}/2\mathbb{Z})$ est égal à son groupe dérivé. Il en découle que le seul morphisme $f : \mathrm{GL}_n(\mathbb{Z}/2\mathbb{Z}) \to \mathrm{G}$ est le morphisme trivial. Il reste encore le cas $n = 2$. On a, outre I_n, deux matrices de transvection dans $\mathrm{GL}_2(\mathbb{Z}/2\mathbb{Z})$ à savoir $\mathrm{T} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ et $\mathrm{U} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$. Les autres éléments de $\mathrm{GL}_2(\mathbb{Z}/2\mathbb{Z})$ sont I_2, $\mathrm{TU} = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$, $\mathrm{UT} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$ et $\mathrm{UTU} = \mathrm{TUT} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$. On pose $f(\mathrm{T}) = a$ et $f(\mathrm{U}) = b$. Comme $\mathrm{T}^2 = \mathrm{U}^2 = \mathrm{I}_2$, on a $a^2 = b^2 = 1$. Si p est impair, cela impose $a = b = 1$ et f est trivial. Supposons p pair dans la suite. De

$f(\mathrm{TUT}) = f(\mathrm{UTU})$, on tire $aba = bab$ et donc $a = b$ (a et b commutent). Si $a \in \mathrm{G}$ et $a^2 = 1$, on vérifie que l'application de $\mathrm{GL}_2(\mathrm{K})$ définie par $f(\mathrm{T}) = f(\mathrm{U}) = f(\mathrm{TUT}) = a$ et $f(\mathrm{UT}) = f(\mathrm{TU}) = f(1) = 1$ est un morphisme de $\mathrm{GL}_2(\mathrm{K})$ dans G. Comme p est pair, G possède des éléments d'ordre 2, donc on peut trouver des morphismes non triviaux.

En fait $\mathrm{SL}_2(\mathbb{Z}/2\mathbb{Z})$ *est isomorphe au groupe symétrique* $\mathcal{S}_3$ *et si on identifie le groupe* $\{1, a\}$ *engendré par l'élément d'ordre 2 choisi au groupe multiplicatif* $\{\pm 1\}$, *le morphisme construit n'est autre que la signature.* ◁

Un cas particulier qui a aussi été posé directement à l'oral des ENS consiste à déterminer les morphismes de $\mathrm{GL}_n(\mathrm{K})$ *dans* K^* *lorsque* K *est un corps fini. Comme dans l'exercice on voit qu'un tel morphisme est de la forme* $\varphi(\det)$ *où* φ *est un morphisme de* K^* *dans lui-même. Comme* K^* *est cyclique,* φ *est de la forme* $x \mapsto x^k$ *où* k *est un entier qu'on peut choisir entre* 0 *et* $\mathrm{Card}\,\mathrm{K} - 1$.

Les exercices qui suivent concernent les matrices inversibles à coefficients entiers. Rappelons que $\mathrm{GL}_n(\mathbb{Z})$ *désigne le groupe des inversibles de l'anneau* $\mathcal{M}_n(\mathbb{Z})$. *Il est facile de voir que* $\mathrm{A} \in \mathcal{M}_n(\mathbb{Z})$ *est dans* $\mathrm{GL}_n(\mathbb{Z})$ *si et seulement si* $\det \mathrm{A} = \pm 1$. *On note enfin* $\mathrm{SL}_n(\mathbb{Z})$ *le sous-groupe de* $\mathrm{GL}_n(\mathbb{Z})$ *formé des matrices de déterminant* 1.

On sait que sur un corps K, *le groupe linéaire* $\mathrm{GL}_n(\mathrm{K})$ *est engendré par les matrices de transvection et les matrices de dilatation, et que le groupe* $\mathrm{SL}_n(\mathrm{K})$ *est engendré par les matrices de transvection. L'exercice suivant étudie la situation sur l'anneau* $\mathbb{Z}$ *dans le cas* $n = 2$.

3.15. Génération de $\mathrm{SL}_2(\mathbb{Z})$

1. Montrer que le groupe $\mathrm{SL}_2(\mathbb{Z})$ est engendré par les deux matrices $\mathrm{U} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ et $\mathrm{V} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$.

2. En déduire qu'il est aussi engendré par U et $\mathrm{T} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$.

(ENS Cachan)

▷ **Solution.**

1. Notons G le sous-groupe engendré par U et V. On vérifie par récurrence, que $\mathrm{U}^k = \begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix}$ pour tout $k \in \mathbb{Z}$. Par suite, en prenant la transposée, on a $\mathrm{V}^k = \begin{pmatrix} 1 & 0 \\ k & 1 \end{pmatrix}$. Le groupe G contient donc toutes

les matrices de transvection (à coefficients dans $\mathbb{Z}$). En particulier, si A est une matrice de G, toute matrice obtenue à partir de A en faisant des opérations élémentaires de la forme $L_i \leftarrow L_i + kL_j$ ou $C_i \leftarrow C_i + kC_j$, avec $\{i,j\} = \{1,2\}$ et $k \in \mathbb{Z}$, est encore dans G. L'idée de la preuve est alors la suivante. Soit $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL_2(\mathbb{Z})$. On va effectuer des opérations du type précédent sur A jusqu'à tomber sur une matrice A′ de transvection. Cela prouve que A est dans G puisqu'elle s'écrit alors comme un produit de matrices de transvection.

• Supposons d'abord que $ac \neq 0$. Comme $ad - bc = 1$, a et c sont premiers entre eux. Effectuons la division euclidienne de a par c : $a = q_1c + r_1$, avec $0 \leqslant r_1 < |c|$. En effectuant l'opération $L_1 \leftarrow L_1 - q_1L_2$ on obtient une matrice de la forme : $\begin{pmatrix} r_1 & \times \\ c & \times \end{pmatrix}$. Si r_1 n'est pas nul, on continue en effectuant la division euclidienne de c par r_1 : $c = q_2r_1 + r_2$ avec $0 \leqslant r_2 < r_1$. Par une manipulation on arrive alors à la matrice $\begin{pmatrix} r_1 & \times \\ r_2 & \times \end{pmatrix}$. On continue l'algorithme d'Euclide jusqu'au premier reste nul. L'avant dernier reste est égal à $1 = \text{pgcd}(a,c)$ et on tombe sur une matrice de la forme $\begin{pmatrix} 1 & b' \\ 0 & d' \end{pmatrix}$ ou de la forme $\begin{pmatrix} 0 & b' \\ 1 & d' \end{pmatrix}$. Dans le premier cas $d' = 1$, puisque le déterminant ne change pas au cours des manipulations, de sorte qu'il s'agit d'une matrice de G. Dans le second cas, on rajoute la deuxième ligne à la première, puis on retranche la première ligne à la deuxième pour se ramener au cas précédent.

• Si l'un des coefficients a ou c est nul, l'autre ne l'est pas (car $ad-bc = 1$) et il suffit d'ajouter la ligne qui le contient à l'autre ligne pour se ramener au cas précédent.

2. Il suffit de constater que $UTU = V$. ◁

La résultat de cet exercice a aussi été obtenu dans le tome 1 comme corollaire de l'étude du groupe modulaire (exercice 2.17). Dans l'exercice 7.19 de ce même tome il est montré que les matrices de transvection $I_n + \lambda E_{ij}$ *avec* $\lambda \in \mathbb{Z}$ *et* $i \neq j$ *et les matrices de dilatation* $I_n - 2E_{ii}$ *engendrent le groupe* $GL_n(\mathbb{Z})$. *En adaptant un peu la solution de cet exercice on peut montrer que les matrices de transvection engendrent tout le groupe* $SL_n(\mathbb{Z})$. *En fait, comme* $(I_n + E_{ij})^k = I_n + kE_{ij}$ *pour tout* $k \in \mathbb{Z}$, *les matrices* $I_n + E_{ij}$ *avec* $i \neq j$ *suffisent. Cela généralise donc le résultat de cet exercice. Le fait que* $SL_n(\mathbb{Z})$ *est de type fini, c'est-à-dire engendré par un nombre fini d'éléments, est utilisé dans l'exercice suivant.*

3.16. Endomorphismes surjectifs de $\mathrm{SL}_n(\mathbb{Z})$

Soit G un groupe de type fini (*i.e.* possédant un système fini de générateurs) et $f : \mathrm{G} \to \mathrm{G}$ un morphisme surjectif. Soit H un groupe fini et g un morphisme de G dans H. On veut montrer que $\mathrm{Ker}\, f \subset \mathrm{Ker}\, g$.

1. Montrer que l'ensemble des morphismes de G dans H est fini.

2. Soit $a \in \mathrm{Ker}\, f$. Montrer l'existence d'une suite (b_n) de G telle que $f^n(b_n) = a$.

3. On pose $g_n = g \circ f^n$. Montrer que pour $m > n$, $g_m(b_n) = e$.

4. Montrer que si $a \notin \mathrm{Ker}\, g$, alors $g_m \neq g_n$ pour $m > n$ et conclure.

5. Application. Montrer que $\mathrm{SL}_2(\mathbb{Z})$ est engendré par $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ et $\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$. Montrer que pour tout $\mathrm{A} \in \mathrm{SL}_2(\mathbb{Z})$, $\mathrm{A} \neq \mathrm{I}_2$, on peut trouver un groupe fini H et un morphisme f de $\mathrm{SL}_2(\mathbb{Z})$ dans H tels que $f(\mathrm{A})$ n'est pas l'élément neutre de H. Montrer que tout endomorphisme surjectif de $\mathrm{SL}_2(\mathbb{Z})$ est bijectif. Généraliser à $\mathrm{SL}_n(\mathbb{Z})$.

(ENS Ulm)

▷ **Solution.**

1. Soit S une partie génératrice finie de G. L'application qui à un morphisme $f : \mathrm{G} \to \mathrm{H}$ associe sa restriction à S est injective car deux morphismes qui coïncident sur une partie génératrice sont forcément égaux. Comme $\mathcal{F}(\mathrm{S}, \mathrm{H})$ est fini car S et H sont finis, il n'y a donc qu'un nombre fini de morphismes de G dans H.

2. Comme f^n est surjectif pour tout n, il existe $b_n \in \mathrm{G}$ tel que $f^n(b_n) = a$.

3. Pour $m > n$, $g_m(b_n) = g(f^m(b_n)) = g(f^{m-n}(f^n(b_n))) = g(f^{m-n}(a)) = g(e) = e$ car $m - n > 0$.

4. Par ailleurs, $g_m(b_m) = g(a)$. Donc si $g(a) \neq e$, les morphismes g_m sont deux à deux distincts car pour tout m la suite $(g_m(b_n))_{n \geqslant 0}$ est stationnaire à partir du rang $m + 1$. Cela est impossible à cause du résultat de la question **1**. Il en résulte que $g(a) = e$ et donc que $\mathrm{Ker}\, f \subset \mathrm{Ker}\, g$.

5. Le fait que $\mathrm{SL}_2(\mathbb{Z})$ est engendré par $\mathrm{U} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ et $\mathrm{V} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$ est l'objet de l'exercice 3.15. Il s'agit donc d'un groupe de type fini. Soit A une matrice de $\mathrm{SL}_2(\mathbb{Z})$ distincte de l'identité. Si A n'est diagonale, on choisit p un nombre premier qui ne divise pas l'un des coef-

ficients non diagonal de A et H le groupe fini $\mathrm{Sl}_2(\mathbb{Z}/p\mathbb{Z})$. La réduction modulo p est un morphisme de groupes qui envoie A sur une matrice qui n'est pas l'identité. Si A est diagonale, forcément $A = -I$ et on prend par exemple la réduction modulo 3. Il découle alors des question précédente que tout endomorphisme surjectif de $\mathrm{SL}_2(\mathbb{Z})$ a un noyau trivial donc est bijectif.

La même argumentation s'applique au groupe $\mathrm{SL}_n(\mathbb{Z})$ qui est aussi de type fini d'après les remarques qui précèdent l'exercice. ◁

Dans l'énoncé suivant on montre que les sous-groupes finis de $\mathrm{SL}_2(\mathbb{Z})$ *sont tous cycliques.*

3.17. Sous-groupes finis de $\mathbf{SL_2(\mathbb{Z})}$

1. Soit G un sous-groupe fini de $\mathrm{GL}_2(\mathbb{Z})$. Montrer qu'il existe un produit scalaire sur $\mathbb{R}^2$, stable par tous les éléments de G, ou si l'on préfère, tel que les éléments de G soient des endomorphismes orthogonaux pour ce produit scalaire.

2. En déduire que les sous-groupes finis de $\mathrm{SL}_2(\mathbb{Z})$ sont cycliques. **(ENS Ulm)**

▷ **Solution.**

1. Notons $\langle\ ,\ \rangle$ le produit scalaire canonique sur $\mathbb{R}^2$. On utilise un argument classique consistant à moyenner les formes bilinéaires $(x,y) \mapsto \langle g(x), g(y)\rangle$ pour g parcourant G. Posons donc $\langle x,y\rangle_{\mathrm{G}} = \dfrac{1}{|\mathrm{G}|}\displaystyle\sum_{g\in\mathrm{G}} \langle g(x), g(y)\rangle$ pour x, y dans $\mathbb{R}^2$. L'application $\langle\ ,\ \rangle_{\mathrm{G}}$ est clairement une forme bilinéaire symétrique. On a $\langle x,x\rangle_{\mathrm{G}} = \dfrac{1}{|\mathrm{G}|}\displaystyle\sum_{g\subset\mathrm{G}} \|g(x)\|^2$ qui est positif, et nul seulement lorsque $x = 0$ car $\mathrm{Id} \in \mathrm{G}$. Il en résulte que $\langle\ .\ \rangle_{\mathrm{G}}$ est un produit scalaire sur $\mathbb{R}^2$. Montrons que tous les éléments de G sont orthogonaux pour ce produit scalaire. Soit $g_0 \in \mathrm{G}$. On a, pour x, y dans $\mathbb{R}^2$.

$$\begin{aligned}\langle g_0(x), g_0(y)\rangle_{\mathrm{G}} &= \frac{1}{|\mathrm{G}|}\sum_{g\in\mathrm{G}} \langle (g\circ g_0)(x), (g\circ g_0)(y)\rangle \\ &= \frac{1}{|\mathrm{G}|}\sum_{h\in\mathrm{G}} \langle h(x), h(y)\rangle = \langle x,y\rangle_{\mathrm{G}}\end{aligned}$$

car l'application $g \mapsto g\circ g_0$ est une permutation des éléments de G.

2. Le groupe orthogonal de $\mathbb{R}^2$ pour le produit scalaire $\langle\ ,\ \rangle_G$ est isomorphe à $\mathrm{O}_2(\mathbb{R})$ et comme les éléments de G sont de déterminant 1, G est isomorphe à un sous-groupe fini de $\mathrm{SO}_2(\mathbb{R})$ lui-même isomorphe au groupe multiplicatif S^1 des nombres complexes de module 1. Pour conclure, il nous suffit donc de prouver que tous les sous-groupes finis de S^1 sont cycliques. Soit H un tel sous-groupe et d son cardinal. Par le théorème de Lagrange, on a $z^d = 1$ pour tout $z \in \mathrm{H}$. Donc H est inclus dans le groupe U_d des racines d-ièmes de l'unité. Comme H possède d éléments, on a égalité $\mathrm{H} = \mathrm{U}_d$. Et il est bien connu que U_d est un groupe cyclique. ◁

L'exercice suivant montre qu'il n'y a qu'un nombre fini d'ordres possibles pour les éléments de $\mathrm{GL}_2(\mathbb{Z})$ *et donc un nombre fini de cardinaux possibles pour les sous-groupes cycliques de* $\mathrm{SL}_2(\mathbb{Z})$.

3.18. Ordres des éléments de $\mathrm{GL}_2(\mathbb{Z})$

Montrer que l'ordre de toute matrice A de $\mathrm{GL}_2(\mathbb{Z})$ est infini ou égal à $1, 2, 3, 4$, ou 6. Montrer que dans chaque cas hormis 1, il existe une infinité de matrices de cet ordre.

(ENS Ulm)

▷ **Solution.**

Soit $\mathrm{A} \in \mathrm{GL}_2(\mathbb{Z})$ que l'on suppose distincte de l'identité. Dire que A est d'ordre fini revient à dire qu'il existe un entier $p \geqslant 2$ tel que $\mathrm{A}^p = \mathrm{I}_2$, ou si l'on préfère que A est annulée par le polynôme $\mathrm{X}^p - 1$. Cela est le cas si et seulement si $\mathrm{X}^p - 1$ est divisible par le polynôme minimal de A (en tant que matrice de $\mathcal{M}_2(\mathbb{Q})$). On sait que A est annulée par son polynôme caractéristique $\chi = \mathrm{X}^2 - t\mathrm{X} + d$ où $t = \mathrm{Tr}(\mathrm{A}) \in \mathbb{Z}$ et $d = \det(\mathrm{A}) = \pm 1$. Le polynôme χ est le polynôme minimal de A sauf dans le cas où A est annulée par un polynôme de degré 1, c'est-à-dire est une matrice scalaire. Comme on doit avoir $\det \mathrm{A} = \pm 1$, cela ne se produit que pour $\mathrm{A} = \mathrm{I}_2$, cas qu'on a exclu, et $\mathrm{A} = -\mathrm{I}_2$ qui est une matrice d'ordre 2. Supposons à partir de maintenant que $\mathrm{A} \neq \pm \mathrm{I}_2$. Notons que le polynôme χ divise $\mathrm{X}^p - 1$ si et seulement si il a deux racines distinctes qui sont des racines p-ièmes de l'unité. On va donc regarder les racines de χ en discutant selon les valeurs de d et t. Le discriminant de χ est $\Delta = t^2 - 4d$.

• Supposons pour commencer $d = -1$. Le discriminant Δ est alors strictement positif. Les racines de χ sont donc réelles. Or les seules racines de l'unité réelle sont -1 et 1. On a $\chi(-1) = t$. Donc si la trace t de A est non nulle, A est d'ordre infini. Si $t = 0$ alors $\chi = \mathrm{X}^2 - 1$ et A

est d'ordre 2. Montrons qu'il existe une infinité de matrices d'ordre 2 de ce type. On choisit le coefficient $(1,1)$ égal à un entier n quelconque. Le coefficient $(2,2)$ doit valoir $-n$ pour que la trace soit nulle. On complète pour avoir un déterminant égal à -1 en prenant par exemple la matrice $\begin{pmatrix} n & 1+n \\ 1-n & -n \end{pmatrix}$.

• Supposons maintenant $d=1$ et $|t|\geqslant 3$. On a encore $\Delta>0$ et les racines de χ sont réelles. Comme $\chi(-1)\neq 0$, la matrice A est d'ordre infini.

• Supposons $d=1$ et $|t|=2$. On a $\Delta=0$ et χ admet une racine double. La matrice A est encore d'ordre infini. C'est le cas par exemple de toutes les matrices triangulaires $\begin{pmatrix} 1 & n \\ 0 & 1 \end{pmatrix}$ pour $n\geqslant 1$.

• Supposons $d=1$ et $t=-1$. On a $\chi=X^2+X+1$. Ce polynôme divise X^3-1 et A est d'ordre 3. C'est le cas des matrices $\begin{pmatrix} n & n^2+n+1 \\ -1 & -n-1 \end{pmatrix}$ pour tout $n\in\mathbb{Z}$.

• Supposons $d=1$ et $t=0$. On a $\chi=X^2+1$, polynôme qui divise X^4-1 et A est d'ordre 4 (car $A^2=-I_2\neq I_2$). C'est le cas des matrices $\begin{pmatrix} n & n^2+1 \\ -1 & -n \end{pmatrix}$ pour tout $n\in\mathbb{Z}$.

• Supposons enfin $d=1$ et $t=1$. On a $\chi=X^2-X+1$, diviseur de X^6-1 et A est d'ordre 6 (car χ ne divise ni X^3-1 ni X^2-1). C'est le cas des matrices $\begin{pmatrix} n & n^2+n-1 \\ -1 & -n+1 \end{pmatrix}$ pour tout $n\in\mathbb{Z}$.

Ainsi, l'ordre d'une matrice de $\mathrm{GL}_2(\mathbb{Z})$ appartient à l'ensemble $\{1,2,3,4,6,\infty\}$ et dans chacun des cas (hormis 1) on a bien une infinité de matrices ayant cet ordre. ◁

Avec quelques connaissances sur les polynômes cyclotomiques (leur irréductibilité sur $\mathbb{Q}$ et la factorisation sur $\mathbb{Q}$ des polynômes X^n-1) il est possible de caractériser tous les ordres possibles des éléments de $\mathrm{GL}_n(\mathbb{Z})$.

L'exercice suivant va généraliser ce qui précède. En effet, il démontre que l'ensemble des cardinaux des sous-groupes finis de $\mathrm{GL}_n(\mathbb{Z})$ est majoré (on parle de restriction cristallographique). En particulier, il existe un entier N_n tel que l'ordre de toute matrice d'ordre fini de $\mathrm{GL}_n(\mathbb{Z})$ soit inférieur à N_n. L'exercice ci-dessus montre que pour $n=2$ l'entier $N_2=6$ convient.

3.19. Sous-groupes finis de $\mathbf{GL}_n(\mathbb{Z})$

1. Soit $A \in \mathcal{M}_n(\mathbb{Z})$. On suppose que A est annulée par un polynôme complexe non nul P dont les racines sont simples et de module < 1. Montrer que $A = 0$.

2. Soit G un sous-groupe fini de $\mathrm{GL}_n(\mathbb{Z})$. Montrer que si $p \geqslant 3$ est un nombre premier, la réduction modulo p de $\mathrm{GL}_n(\mathbb{Z})$ dans $\mathrm{GL}_n(\mathbb{Z}/p\mathbb{Z})$ restreinte à G est injective. Que peut-on en déduire ?

(**ENS Lyon**)

▷ **Solution.**

1. Comme le polynôme P est à racines simples, la matrice A est diagonalisable dans $\mathcal{M}_n(\mathbb{C})$. Soit D diagonale et Q inversible telles que $A = Q^{-1}DQ$. Comme toutes les valeurs propres de A sont de module < 1, la suite $(A^p)_{p\geqslant 0}$ tend vers 0 lorsque p tend vers l'infini. En particulier, il existe un entier p tel que $\|A^p\| < 1$ où $\|(m_{ij})\| = \max_{1\leqslant i,j\leqslant n} |m_{ij}|$. Mais la matrice A^p étant à coefficients dans $\mathbb{Z}$, cela impose que $A^p = 0$. On a alors $D^p = 0$ et donc $D = 0$ car D est diagonale. Ainsi, A est nulle.

2. L'application $f : G \to \mathrm{GL}_n(\mathbb{Z}/p\mathbb{Z})$ qui à une matrice $M = (m_{ij})$ de G associe $\overline{M} = (\overline{m_{ij}})$, sa réduction modulo p, est un morphisme de groupes. Pour montrer qu'il est injectif, on montre que son noyau est réduit à l'identité. Soit $M \in \operatorname{Ker} f$. On a donc $M \equiv I_n$ modulo p de sorte qu'on peut écrire $M = I_n + pA$ avec $A \in \mathcal{M}_n(\mathbb{Z})$. Comme G est fini, M est d'ordre fini $q \geqslant 1$. L'égalité $M^q = I_n$ signifie que le polynôme $P(X) = (1 + pX)^q - 1$ annule la matrice A. Or les racines de P sont les nombres complexes $z_k = \dfrac{w_k - 1}{p}$ où $w_1, \ldots, w_q$ sont les racines q-ièmes de l'unité. Ces racines sont deux à deux distinctes et comme $p \geqslant 3$ on a $|z_k| < 1$ pour tout k. La question précédente permet d'affirmer que A est nulle et donc que M est la matrice identité.

De ce qui précède, on déduit que les sous-groupes finis de $\mathrm{GL}_n(\mathbb{Z})$ ont un cardinal majoré par le cardinal de $\mathrm{GL}_n(\mathbb{Z}/3\mathbb{Z})$, et donc (majoration rapide) par 3^{n^2}. En particulier, l'ordre de tout élément de $\mathrm{GL}_n(\mathbb{Z})$ est, soit infini, soit $\leqslant 3^{n^2}$. D'autre part, il n'y a, à isomorphisme près, qu'un nombre fini de sous-groupes finis dans $\mathrm{GL}_n(\mathbb{Z})$ (car il n'y a qu'un nombre fini de groupes de cardinal fini donné, à isomorphisme près bien entendu). ◁

Voici maintenant quelques exercices où le corps de base est un corps fini.

3.20. Un calcul de signature

Soit $p \geqslant 3$ un nombre premier et M une matrice de $\mathrm{GL}_2(\mathbb{Z}/p\mathbb{Z})$. L'application $X \mapsto MX$ est une permutation de $(\mathbb{Z}/p\mathbb{Z})^2$. Quelle est sa signature ? *On pourra utiliser sans le redémontrer le fait que* $(\mathbb{Z}/p\mathbb{Z})^*$ *est un groupe cyclique.*

(ENS Ulm)

▷ **Solution.**

• La signature induit bien entendu un morphisme de groupes de $\mathrm{GL}_2(\mathbb{Z}/p\mathbb{Z})$ dans $\{\pm 1\}$, que l'on notera ε. Pour connaître un morphisme de groupes, il suffit de le connaître sur une partie génératrice. Or, on sait que le groupe linéaire est engendré par les matrices de transvection $\begin{pmatrix} 1 & \lambda \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ \lambda & 1 \end{pmatrix}$ avec $\lambda \in \mathbb{Z}/p\mathbb{Z}$ et les matrices de dilatation $\begin{pmatrix} 1 & 0 \\ 0 & \alpha \end{pmatrix}, \begin{pmatrix} \alpha & 0 \\ 0 & 1 \end{pmatrix}$ avec $\alpha \in (\mathbb{Z}/p\mathbb{Z})^*$. Le lecteur se reportera à l'exercice 3.1 en notant que le résultat est rapide à prouver dans le cas des matrices 2×2 (ce qu'il faudrait peut être faire le jour de l'oral).

• Montrons que toutes les matrices de transvection ont une signature égale à 1. En effet, c'est clair pour la matrice identité, et si λ n'est pas nul, on observe que $T_\lambda = \begin{pmatrix} 1 & \lambda \\ 0 & 1 \end{pmatrix}$ est d'ordre p dans le groupe $\mathrm{GL}_2(\mathbb{Z}/p\mathbb{Z})$. Par conséquent, on a $\varepsilon(T_\lambda)^p = \varepsilon(T_\lambda^p) = \varepsilon(Id) = 1$. Comme p est impair, on a forcément $\varepsilon(T_\lambda) = 1$. Le raisonnement est identique pour la transposée de T_λ.

On peut aussi utiliser le fait que les transvections sont des carrés du groupe $\mathrm{GL}_2(\mathbb{Z}/p\mathbb{Z})$. *Comme les matrices de transvection engendrent le groupe spécial linéaire* $\mathrm{SL}_2(\mathbb{Z}/p\mathbb{Z})$, *la signature restreinte à ce sous-groupe est donc triviale. Cette remarque n'est toutefois pas indispensable pour poursuivre.*

• Les matrices de dilatation $D_\alpha = \begin{pmatrix} \alpha & 0 \\ 0 & 1 \end{pmatrix}$ avec $\alpha \in (\mathbb{Z}/p\mathbb{Z})^*$ forment un sous-groupe isomorphe à $(\mathbb{Z}/p\mathbb{Z})^*$. Comme $(\mathbb{Z}/p\mathbb{Z})^*$ est cyclique (l'énoncé le rappelle et et le lecteur en trouvera une preuve dans l'exercice 4.12 du tome 1 d'algèbre) il suffit ici encore de s'intéresser à la signature d'un générateur.

Soit donc α qui engendre $(\mathbb{Z}/p\mathbb{Z})^*$. On va compter les orbites de $(\mathbb{Z}/p\mathbb{Z})^2$ sous l'action de D_α. Il y a tout d'abord les vecteurs $(0, x)$ qui donnent des orbites réduites à des singletons (et il y en a p) puis les orbites des vecteurs $(1, x)$ qui contiennent les $p - 1$ vecteurs (y, x) avec

$y \neq 0$ (et on en a encore p). Au total il y a donc $2p$ orbites et la signature est donc $(-1)^{p^2-2p} = -1$ car p est impair.

• On aimerait tout de même avoir une vision globale plus claire de cette signature. On connaît un morphisme de groupes important sur $\mathrm{GL}_2(\mathbb{Z}/p\mathbb{Z})$ à savoir le déterminant. Il est assez naturel d'essayer de relier notre signature au déterminant.

Si x est dans $\mathbb{Z}/p\mathbb{Z}$ on sait que $x^{p-1} = 1$ (c'est le petit théorème de Fermat) et donc $x^{\frac{p-1}{2}}$ vaut 1 ou -1 puisque son carré fait 1 (et le polynôme $\mathrm{X}^2 - 1$ n'a que 1 et -1 pour racines dans $\mathbb{Z}/p\mathbb{Z}$, qui est un corps). Pour notre générateur α de tout à l'heure on a forcément $\alpha^{\frac{p-1}{2}} = -1$, car sinon on aurait $x^{\frac{p-1}{2}} = 1$ pour tout x non nul et le polynôme $\mathrm{X}^{\frac{p-1}{2}} - 1$ aurait $p-1$, racines ce qu'interdit son degré.

Ainsi, on a $\varepsilon(\mathrm{D}_\alpha) = -1 = \alpha^{\frac{p-1}{2}} = \det(\mathrm{D}_\alpha)^{\frac{p-1}{2}}$. Cette relation est alors vraie pour toute les matrices de dilatation et aussi pour les matrices de transvection (le déterminant vaut 1). Deux morphismes de groupes qui coïncident sur une partie génératrice étant égaux, on a donc démontré la jolie formule

$$\boxed{\forall \mathrm{M} \in \mathrm{GL}_2(\mathbb{Z}/p\mathbb{Z}), \quad \varepsilon(\mathrm{M}) = \det \mathrm{M}^{\frac{p-1}{2}}}. \quad \lhd$$

Les groupes linéaires sur les corps finis fournissent des exemples importants de groupes finis. L'exercice suivant propose l'étude du groupe $\mathrm{SL}_2(\mathbb{Z}/3\mathbb{Z})$.

3.21. Étude de $\mathrm{SL}_2(\mathbb{Z}/3\mathbb{Z})$

1. Quel est le cardinal de $\mathrm{GL}_2(\mathbb{Z}/3\mathbb{Z})$? de $\mathrm{SL}_2(\mathbb{Z}/3\mathbb{Z})$?

2. Montrer qu'il n'existe aucun morphisme surjectif de groupes de $\mathrm{SL}_2(\mathbb{Z}/3\mathbb{Z})$ dans $\mathbb{Z}/2\mathbb{Z}$.

3. Montrer qu'il n'existe aucun sous-groupe de $\mathrm{SL}_2(\mathbb{Z}/3\mathbb{Z})$ de cardinal 12.

4. $\mathrm{SL}_2(\mathbb{Z}/3\mathbb{Z})$ est-il isomorphe à $\mathcal{S}_4$?

5. Montrer qu'il existe un morphisme de groupes surjectif de $\mathrm{SL}_2(\mathbb{Z}/3\mathbb{Z})$ sur $\mathcal{A}_4$.

(ENS Ulm)

▷ **Solution.**

1. La donnée d'une matrice de $\mathrm{GL}_2(\mathbb{Z}/3\mathbb{Z})$ équivaut à la donnée d'une famille libre de deux vecteurs du plan $(\mathbb{Z}/3\mathbb{Z})^2$. Il y a $3^2 - 1 = 8$ choix

pour le premier vecteur, qui doit simplement être non nul, et pour chacun de ces choix, il y a $9-3=6$ manières de la compléter en une famille libre puisqu'il faut choisir le second vecteur en dehors de la droite vectorielle engendrée par le premier. Ainsi, $|\mathrm{GL}_2(\mathbb{Z}/3\mathbb{Z})| = 8.6 = 48$.

Comme le déterminant est un morphisme surjectif de groupes de $\mathrm{GL}_2(\mathbb{Z}/3\mathbb{Z})$ dans $(\mathbb{Z}/3\mathbb{Z})^*$, dont le noyau est $\mathrm{SL}_2(\mathbb{Z}/3\mathbb{Z})$, on a $|\mathrm{SL}_2(\mathbb{Z}/3\mathbb{Z})| = 24$.

2. Soit ψ un morphisme de groupes de $\mathrm{SL}_2(\mathbb{Z}/3\mathbb{Z})$ dans $\mathbb{Z}/2\mathbb{Z}$. On se propose de prouver que ψ est trivial. Nous savons qu'un morphisme de groupes est uniquement déterminé par l'image d'une partie génératrice. Or, nous savons que les matrices de transvection $\begin{pmatrix} 1 & \lambda \\ 0 & 1 \end{pmatrix}$ et $\begin{pmatrix} 1 & 0 \\ \lambda & 1 \end{pmatrix}$, où $\lambda \in \{1,2\}$, engendrent $\mathrm{SL}_2(\mathbb{Z}/3\mathbb{Z})$ (voir l'exercice 3.1).

On a pour tout $A \in \mathrm{SL}_2(\mathbb{Z}/3\mathbb{Z})$, $\psi(A^2) = 2\psi(A) = 0$, c'est-à-dire que ψ est nul sur les carrés. Or, $\begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}^2$ et $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}^2$. Donc ψ est nul sur toutes les matrices de transvection, et par suite identiquement nul.

3. Supposons par l'absurde qu'il existe un sous-groupe H de $\mathrm{SL}_2(\mathbb{Z}/3\mathbb{Z})$ de cardinal 12 et soit $a \notin \mathrm{H}$. On a alors $\mathrm{SL}_2(\mathbb{Z}/3\mathbb{Z}) = \mathrm{H} \cup a\mathrm{H}$, la réunion étant disjointe. Le produit de deux éléments de H est dans H, le produit d'un élément de H et d'un élément de aH est dans aH et le produit de deux éléments de aH est dans H (car l'égalité $ahah' = ah''$ avec h, h', h'' dans H implique que $a \in \mathrm{H}$). Il en découle directement que l'application qui envoie les éléments de H sur 0 et ceux de aH sur 1 est un morphisme de groupes surjectif de $\mathrm{SL}_2(\mathbb{Z}/3\mathbb{Z})$ sur $\mathbb{Z}/2\mathbb{Z}$, ce qui contredit le résultat de la question précédente.

4. Comme $\mathcal{S}_4$ (qui est aussi de cardinal 24) admet un sous-groupe d'ordre 12, à savoir le groupe alterné $\mathcal{A}_4$, la question précédente permet d'affirmer que $\mathrm{SL}_2(\mathbb{Z}/3\mathbb{Z})$ n'est pas isomorphe à $\mathcal{S}_4$.

5. On va utiliser une opération du groupe $\mathrm{SL}_2(\mathbb{Z}/3\mathbb{Z})$ sur un ensemble à 4 éléments. On prend l'ensemble E des droites vectorielles du plan $(\mathbb{Z}/3\mathbb{Z})^2$. Il y a 4 droites qui sont $\mathrm{D}_1 = \mathrm{Vect}(1,0)$, $\mathrm{D}_2 = \mathrm{Vect}(0,1)$, $\mathrm{D}_3 = \mathrm{Vect}(1,1)$ et $\mathrm{D}_4 = \mathrm{Vect}(1,2)$. Une matrice $A \in \mathrm{SL}_2(\mathbb{Z}/3\mathbb{Z})$ envoie une droite sur une droite et cela de manière injective (deux vecteurs libres sont envoyés sur deux vecteurs libres). Il en résulte que $\mathrm{SL}_2(\mathbb{Z}/3\mathbb{Z})$ opère sur E. Notons $\varphi : \mathrm{SL}_2(\mathbb{Z}/3\mathbb{Z}) \to \mathcal{S}_\mathrm{E}$ le morphisme de groupes correspondant à cette opération. Soit $A \in \mathrm{Ker}\,\varphi$. On a $A(\mathrm{D}_1) = \mathrm{D}_1$ et $A(\mathrm{D}_2) = \mathrm{D}_2$ donc A est une homothétie, *i.e.* $A = \pm \mathrm{Id}$ (car $A \in \mathrm{SL}_2(\mathbb{Z}/3\mathbb{Z})$). Inversement, Id et $-\mathrm{Id}$ opèrent de manière tri-

viale sur E. Donc $|\operatorname{Ker}\varphi| = 2$ et par suite $|\operatorname{Im}\varphi| = 12$. On observe que les transvections vues ci-dessus qui engendrent $\mathrm{SL}_2(\mathbb{Z}/3\mathbb{Z})$ donnent des 3-cycles de E. Par exemple $\varphi\left(\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}\right) = (\mathrm{D}_2, \mathrm{D}_3, \mathrm{D}_1)$. Il en résulte que $\varphi(\mathrm{SL}_2(\mathbb{Z}/3\mathbb{Z})) = \mathcal{A}_\mathrm{E} \simeq \mathcal{A}_4$. D'où le résultat. ◁

Les trois exercices suivant étudient encore des groupes linéaires finis mais sur des anneaux quotients $\mathbb{Z}/m\mathbb{Z}$. Les techniques sont les mêmes avec des complications supplémentaires liées à la non intégrité de l'anneau.

3.22. Étude de $\mathrm{SL}_2(\mathbb{Z}/2^n\mathbb{Z})$

Soit A un anneau commutatif et $\mathrm{G} = \mathrm{SL}_2(\mathrm{A})$ l'ensemble des matrices $(2,2)$ à coefficients dans A et de déterminant 1.

1. Montrer que G est un groupe pour le produit matriciel.

2. On prend $\mathrm{A} = \mathbb{Z}/p\mathbb{Z}$ avec p premier. Calculer $\operatorname{Card}\mathrm{G}$. Dans le cas où $p = 2$, étudier l'ordre des éléments de G : reconnaître un groupe connu isomorphe à G.

3. On prend $\mathrm{A} = \mathbb{Z}/2^n\mathbb{Z}$. Quel est le cardinal de G ? Montrer que tout élément de G a un ordre inférieur ou égal à $3 \times 2^{n-1}$.

(ENS Lyon)

▷ **Solution.**

1. L'ensemble $\mathcal{M}_2(\mathrm{A})$ des matrices carrées de taille (2.2) à coefficients dans A est un anneau pour les opérations usuelles. Le déterminant d'une matrice $\mathrm{M} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathcal{M}_2(\mathrm{A})$ étant défini par $\det\mathrm{M} = ad - bc$, il est facile de vérifier que $\det(\mathrm{MN}) = \det\mathrm{M}\det\mathrm{N}$ pour tout couple de matrices (M, N). Par suite, si M est inversible dans $\mathcal{M}_2(\mathrm{A})$, $\det\mathrm{M}$ est un inversible de l'anneau A. Réciproquement, soit $\mathrm{M} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathcal{M}_2(\mathrm{A})$ telle que $\det\mathrm{M}$ soit inversible dans A. Il est facile de vérifier que la matrice $(\det\mathrm{M})^{-1}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$ est l'inverse de M. On a donc montré que le groupe des inversibles de l'anneau $\mathcal{M}_2(\mathrm{A})$, notons-le $\mathrm{GL}_2(\mathrm{A})$, est exactement l'ensemble des matrices dont le déterminant appartient à $\mathrm{U}(\mathrm{A})$, le groupe des inversibles de A. De plus, le déterminant établit un morphisme de groupes entre $\mathrm{GL}_2(\mathrm{A})$ et $\mathrm{U}(\mathrm{A})$ dont $\mathrm{G} = \mathrm{SL}_2(\mathrm{A})$ n'est autre que le noyau. Par conséquent G est bien un groupe pour le produit matriciel.

2. Notons que le morphisme det : $\mathrm{GL}_2(\mathrm{A}) \to \mathrm{U}(\mathrm{A})$ est surjectif car si α est un inversible de A la matrice diagonale $\begin{pmatrix} \alpha & 0 \\ 0 & 1 \end{pmatrix}$ est inversible et a pour déterminant α. Il en découle que $\mathrm{Card}\,\mathrm{G} \nearrow \mathrm{Card}\,\mathrm{U}(\mathrm{A}) = \mathrm{Card}\,\mathrm{GL}_2(\mathrm{A})$: c'est un corollaire direct du lemme des bergers et le lecteur trouvera une preuve détaillée de ce fait dans la solution de l'exercice 2.4 du premier tome d'algèbre (page 37).

Lorsque $\mathrm{A} = \mathbb{Z}/p\mathbb{Z}$ avec p premier il s'agit d'un corps, et $\mathrm{U}(\mathrm{A})$ est de cardinal $p-1$. Le calcul du cardinal de $\mathrm{GL}_2(\mathbb{Z}/p\mathbb{Z})$ est facile : pour définir une matrice inversible, on choisit d'abord sa première colonne qui doit être non nulle (il y a donc p^2-1 possibilités) puis la seconde qui ne doit pas être colinéaire à la première (et il y a donc p^2-p possibilités). Le cardinal de $\mathrm{GL}_2(\mathbb{Z}/p\mathbb{Z})$ est donc $(p^2-1)(p^2-p)$ et celui de $\mathrm{G} = \mathrm{SL}_2(\mathbb{Z}/p\mathbb{Z})$ est $p(p^2-1)$. Le lecteur trouvera dans l'exercice 1.7 du tome 1 d'algèbre le calcul du cardinal de $\mathrm{GL}_n(\mathrm{K})$ pour n quelconque et K corps fini quelconque.

En particulier, avec $p = 2$, on a $\mathrm{Card}\,\mathrm{SL}_2(\mathbb{Z}/2\mathbb{Z}) = 6$. Les 6 éléments de ce groupe sont l'identité, les matrices $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$, $\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$ et $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, qui sont d'ordre 2, et les matrices $\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$ et $\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$, qui sont d'ordre 3. On a la même répartition des ordres que dans le groupe symétrique $\mathcal{S}_3$ qui, outre l'identité, contient 3 transpositions et deux 3-cycles. Et effectivement ces deux groupes sont isomorphes. Pour le voir, il suffit d'observer qu'une matrice M de $\mathrm{SL}_2(\mathbb{Z}/2\mathbb{Z})$ induit une permutation des 3 vecteurs non nuls du plan $(\mathbb{Z}/2\mathbb{Z})^2$ et que cette permutation n'est l'identité que si $\mathrm{M} = \mathrm{I}_2$.

À isomorphisme près il n'y a que deux groupes de cardinal 6 : $\mathbb{Z}/6\mathbb{Z}$ qui est cyclique et $\mathcal{S}_3$. Comme $\mathrm{SL}_2(\mathbb{Z}/2\mathbb{Z})$ n'est pas commutatif, ce fait permet de dire directement qu'il est isomorphe à $\mathcal{S}_3$.

3. L'étude détaillée du cas $n = 1$ menée dans la question précédente incite à procéder par récurrence sur l'entier n. Et effectivement, comme l'anneau $\mathbb{Z}/2^n\mathbb{Z}$ n'est pas intègre pour $n \geqslant 2$, il n'est pas évident de trouver directement le cardinal de $\mathrm{GL}_2(\mathbb{Z}/2^n\mathbb{Z})$. On dispose d'un morphisme d'anneau naturel π_n entre $\mathbb{Z}/2^{n+1}\mathbb{Z}$ et $\mathbb{Z}/2^n\mathbb{Z}$ qui à la classe d'un entier k modulo 2^{n+1} associe sa classe modulo 2^n (qui bien entendu ne dépend pas du représentant k choisi dans sa classe modulo 2^{n+1}). Ce morphisme est surjectif et son noyau est de cardinal 2 (il contient les classes de 0 et de 2^n). On a alors un morphisme de groupes entre $\mathrm{SL}_2(\mathbb{Z}/2^{n+1}\mathbb{Z})$ et $\mathrm{SL}_2(\mathbb{Z}/2^n\mathbb{Z})$ en considérant l'application φ_n qui à la matrice $\mathrm{M} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z}/2^{n+1}\mathbb{Z})$ associe la

matrice $\begin{pmatrix} \pi_n(a) & \pi_n(b) \\ \pi_n(c) & \pi_n(d) \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z}/2^n\mathbb{Z})$. Le noyau de φ_n est de cardinal 8. Vérifions que φ_n est surjectif. Soit M une matrice de $\mathrm{SL}_2(\mathbb{Z}/2^n\mathbb{Z})$. Comme π_n est surjectif on peut trouver quatre entiers a, b, c, d tels que $\varphi_n \begin{pmatrix} \bar{a} & \bar{b} \\ \bar{c} & \bar{d} \end{pmatrix} = \mathrm{M}$ où $\bar{k}$ désigne la classe de l'entier k dans $\mathbb{Z}/2^{n+1}\mathbb{Z}$. Mais cette matrice n'est pas forcément de déterminant 1 ! Par hypothèse on a $ad - bc \equiv 1\ [2^n]$ donc, modulo 2^{n+1}, on a soit $ad - bc \equiv 1$, soit $ad - bc \equiv 1 + 2^n$. Dans le premier cas, c'est bon. Dans le second cas, l'un au moins des quatre entiers a, b, c, d est impair, disons par exemple a pour fixer les idées. On remplace alors d par $d + 2^n$. On obtient alors un antécédent de M par φ_n qui est bien dans $\mathrm{SL}_2(\mathbb{Z}/2^{n+1}\mathbb{Z})$.

Ce travail montre donc que $\operatorname{Card} \mathrm{SL}_2(\mathbb{Z}/2^{n+1}\mathbb{Z}) = 8 \times \operatorname{Card} \mathrm{SL}_2(\mathbb{Z}/2^n\mathbb{Z})$ et, compte tenu de la question précédente, on peut conclure que

$$\boxed{\operatorname{Card} \mathrm{SL}_2(\mathbb{Z}/2^n\mathbb{Z}) = 3 \times 2^{3n-2}}.$$

Comme $\mathrm{U}(\mathbb{Z}/2^n\mathbb{Z})$ *est de cardinal* 2^{n-1} *on en déduit que* $\mathrm{GL}_2(\mathbb{Z}/2^n\mathbb{Z})$ *est de cardinal* $3 \times 2^{4n-3}$. *Notons qu'on aurait aussi pu étudier le morphisme induit par* π_n *entre* $\mathrm{GL}_2(\mathbb{Z}/2^{n+1}\mathbb{Z})$ *et* $\mathrm{GL}_2(\mathbb{Z}/2^n\mathbb{Z})$. *Le noyau est le même et la surjectivité est immédiate. On en aurait déduit le cardinal de* $\mathrm{GL}_2(\mathbb{Z}/2^n\mathbb{Z})$ *puis celui de* $\mathrm{SL}_2(\mathbb{Z}/2^n\mathbb{Z})$.

Pour la question sur les ordres des éléments de G on procède encore par récurrence sur n, le cas $n = 1$ étant traité dans la question 2. Supposons le résultat vrai au rang n et soit $\mathrm{M} \in \mathrm{SL}_2(\mathbb{Z}/2^{n+1}\mathbb{Z})$. Notons p l'ordre de $\varphi_n(\mathrm{M})$. Par hypothèse de récurrence, $p \leqslant 3 \times 2^{n-1}$. On a $\varphi_n(\mathrm{M}^p) = \varphi_n(\mathrm{M})^p = \mathrm{I}_2$, donc $\mathrm{M}^p \in \operatorname{Ker} \varphi_n$. On constate que les 7 matrices du noyau de φ_n distinctes de l'identité sont toutes d'ordre 2. On a donc $\mathrm{M}^{2p} = \mathrm{I}_2$ et l'ordre de M est inférieur à $2p$ donc à 3×2^n. Cela termine la récurrence.

Le lecteur montrera sans mal que ce résultat est aussi vrai pour les éléments de $\mathrm{GL}_2(\mathbb{Z}/2^n\mathbb{Z})$. ◁

On trouvera dans la solution de l'exercice suivant le calcul du cardinal de $\mathrm{SL}_2(\mathbb{Z}/d\mathbb{Z})$ *pour tout entier d. L'exercice est difficile, surtout la dernière question.*

3.23. Réduction modulo n

Si A est un anneau commutatif, $\mathrm{SL}_2(\mathrm{A})$ désigne le groupe des matrices à coefficients dans A et de déterminant 1. Pour $n \in \mathbb{N}^*$ on considère $\psi_n : \mathrm{SL}_2(\mathbb{Z}) \to \mathrm{SL}_2(\mathbb{Z}/n\mathbb{Z})$, l'application qui à toute matrice A associe la matrice $\bar{\mathrm{A}}$ dont les coefficients sont les classes modulo n des coefficients de A.

1. Montrer que ψ_n est un morphisme de groupes surjectif.

2. Cela est-il encore vrai avec $\varphi_n : \mathrm{GL}_2(\mathbb{Z}) \to \mathrm{GL}_2(\mathbb{Z}/n\mathbb{Z})$?

3. Si $d = \mathrm{pgcd}(n, m)$ montrer que $\mathrm{Ker}\,\psi_d = \mathrm{Ker}\,\psi_n \mathrm{Ker}\,\psi_m$ (ensemble des produits d'un élément de $\mathrm{Ker}\,\psi_n$ par un élément de $\mathrm{Ker}\,\psi_m$).

(ENS Ulm, Lyon)

▷ **Solution.**

1. Notons que le fait que $\mathrm{SL}_2(\mathrm{A})$ est un groupe est prouvé dans l'exercice précédent. Si $\mathrm{A} \in \mathrm{SL}_2(\mathbb{Z})$, alors $\det(\psi_n(\mathrm{A})) = \overline{\det \mathrm{A}} = \bar{1}$, donc $\psi_n(\mathrm{A})$ appartient à $\mathrm{SL}_2(\mathbb{Z}/n\mathbb{Z})$. L'application ψ_n est clairement un morphisme de groupes. Il reste à montrer qu'il est surjectif. Rappelons que si K est un corps, $\mathrm{SL}_2(\mathrm{K})$ est engendré par les transvections (cf. exercice 3.1). Nous allons démontrer qu'il en est de même de $\mathrm{SL}_2(\mathbb{Z}/n\mathbb{Z})$, pour $n \in \mathbb{N}^*$ quelconque. On notera $\overline{m}$ la classe modulo n d'un entier m.

Soit $\mathrm{M} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z}/n\mathbb{Z})$. On va essayer de décomposer M comme un produit de transvections en suivant la preuve déjà utilisée avec un corps. Il y a une complication supplémentaire liée au fait qu'un élément non nul de $\mathbb{Z}/n\mathbb{Z}$ n'est pas forcément inversible. On distingue deux cas.

• Supposons pour commencer que a est inversible dans $\mathbb{Z}/n\mathbb{Z}$. Comme $ad - bc = \bar{1}$, on obtient

$$\begin{pmatrix} \bar{1} & \bar{0} \\ -a^{-1}c & \bar{1} \end{pmatrix} \mathrm{M} = \begin{pmatrix} a & b \\ \bar{0} & d - a^{-1}bc \end{pmatrix} = \begin{pmatrix} a & b \\ \bar{0} & a^{-1} \end{pmatrix}$$

puis

$$\begin{pmatrix} \bar{1} & -ab \\ \bar{0} & \bar{1} \end{pmatrix} \begin{pmatrix} \bar{1} & \bar{0} \\ -a^{-1}c & \bar{1} \end{pmatrix} \mathrm{M} = \begin{pmatrix} a & \bar{0} \\ \bar{0} & a^{-1} \end{pmatrix}.$$

On montre que $\begin{pmatrix} a & \bar{0} \\ \bar{0} & a^{-1} \end{pmatrix}$ est un produit de matrices de transvection. On commence par faire apparaître un $\bar{1}$ en haut à gauche en écrivant

$$\begin{pmatrix} \bar{1} & -\bar{1} + a^{-1} \\ \bar{0} & \bar{1} \end{pmatrix} \begin{pmatrix} \bar{1} & \bar{0} \\ \bar{1} & \bar{1} \end{pmatrix} \begin{pmatrix} a & \bar{0} \\ \bar{0} & a^{-1} \end{pmatrix} = \begin{pmatrix} \bar{1} & a^{-1}(a^{-1} - \bar{1}) \\ a & a^{-1} \end{pmatrix}.$$

Ensuite on écrit

$$\begin{pmatrix} \overline{1} & \overline{0} \\ -a & \overline{1} \end{pmatrix}\begin{pmatrix} \overline{1} & a^{-1}(a^{-1}-\overline{1}) \\ a & a^{-1} \end{pmatrix} = \begin{pmatrix} \overline{1} & a^{-1}(a^{-1}-\overline{1}) \\ \overline{0} & \overline{1} \end{pmatrix}.$$

L'inverse d'une matrice de transvection étant une matrice de transvection, on obtient que M est un produit de matrices de transvection.

• Dans le cas où a n'est pas inversible, on considère

$$\begin{pmatrix} \overline{1} & \lambda \\ \overline{0} & \overline{1} \end{pmatrix} \mathrm{M} = \begin{pmatrix} a+c\lambda & b+d\lambda \\ c & d \end{pmatrix}.$$

On va essayer de choisir λ de sorte que $a+\lambda c$ soit inversible pour se ramener au cas précédent. Notons $p_1, \ldots, p_k$ les diviseurs premiers de n. On considère des entiers α, β, γ, δ tels que $a = \overline{\alpha}$, $b = \overline{\beta}$, $c = \overline{\gamma}$, $d = \overline{\gamma}$. Puisque a n'est pas inversible, α n'est pas premier avec n ; ils ont des facteurs premiers en commun. Supposons par exemple que α est divisible par $p_1, \ldots p_r$ $(r \leqslant k)$ et n'est pas divisible par $p_{r+1}, \ldots, p_k$. De $ad-bc=1$, on tire $\alpha\delta - \beta\gamma \equiv 1 \pmod{p_i}$, pour tout i, donc $-\beta\gamma \equiv 1 \pmod{p_i}$ pour $1 \leqslant i \leqslant r$. On en déduit que γ est premier avec p_i pour $1 \leqslant i \leqslant r$. Posons $\mu = \prod_{i=r+1}^{k} p_i$ et $\lambda = \overline{\mu}$. On vérifie que $\alpha + \gamma\mu$ n'est divisible par aucun des entiers p_i donc $a+\lambda c$ est inversible dans $\mathbb{Z}/n\mathbb{Z}$. Ce qui précède montre que $\begin{pmatrix} \overline{1} & \lambda \\ \overline{0} & \overline{1} \end{pmatrix}$ M est produit de matrices de transvection. Il en est donc de même de M.

On peut alors conclure quant à la surjectivité de ψ_n. En effet, le groupe $\operatorname{Im}\psi_n$ contient toutes les matrices de transvection puisque si $\mu \in \mathbb{Z}$, $\begin{pmatrix} \overline{1} & \overline{\mu} \\ \overline{0} & \overline{1} \end{pmatrix}$ est l'image de $\begin{pmatrix} 1 & \mu \\ 0 & 1 \end{pmatrix}$, et de même pour les transposées. Ce qui précède montre donc que $\operatorname{Im}\psi_n = \mathrm{SL}(\mathbb{Z}/n\mathbb{Z})$.

2. En général, l'application $\varphi_n : \mathrm{GL}_2(\mathbb{Z}) \longrightarrow \mathrm{GL}_2(\mathbb{Z}/n\mathbb{Z})$ n'est pas surjective. En effet les matrices inversibles de $\mathrm{GL}_2(\mathbb{Z})$ sont les matrices dont le déterminant est inversible dans $\mathbb{Z}$ c'est-à-dire égal à ± 1. On a donc pour tout $\mathrm{A} \in \mathrm{GL}_2(\mathbb{Z})$, $\det(\varphi_n(\mathrm{A})) = \pm\overline{1}$. Or l'image de $\mathrm{GL}_2(\mathbb{Z}/n\mathbb{Z})$ par le déterminant est l'ensemble des éléments inversibles de $\mathbb{Z}/n\mathbb{Z}$ et cet ensemble contient strictement $\{\pm\overline{1}\}$ pour $n \geqslant 5$.

Pour $n=2$, on a $\mathrm{SL}_2(\mathbb{Z}/2\mathbb{Z}) = \mathrm{GL}_2(\mathbb{Z}/2\mathbb{Z})$ et l'application φ_2 est surjective.

Pour $n=3$ ou 4, l'ensemble $\varphi_n(\mathrm{GL}_2(\mathbb{Z}))$ contient $\mathrm{SL}_2(\mathbb{Z}/n\mathbb{Z})$ et $\begin{pmatrix} -\overline{1} & \overline{0} \\ \overline{0} & 1 \end{pmatrix} = \varphi_n\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$. Toute matrice de déterminant $-\overline{1}$ est

produit de $\begin{pmatrix} -\overline{1} & \overline{0} \\ \overline{0} & \overline{1} \end{pmatrix}$ et d'un élément de $\mathrm{SL}_2(\mathbb{Z}/n\mathbb{Z})$ donc appartient à $\varphi_n(\mathrm{GL}_2(\mathbb{Z})$ de sorte que l'application est surjective.

Conclusion. L'application $\varphi_n : \mathrm{GL}_2(\mathbb{Z}) \longrightarrow \mathrm{GL}_2(\mathbb{Z}/n\mathbb{Z})$ est surjective si et seulement si $n \leqslant 4$.

3. Il est évident que $\mathrm{Ker}\,\psi_n \subset \mathrm{Ker}\,\psi_d$: si une matrice $\mathrm{A} \in \mathrm{SL}_2(\mathbb{Z})$ donne l'identité lorsqu'on la réduit modulo n, c'est *a fortiori* aussi le cas lorsqu'on la réduit modulo d puisque d divise n. On a de même $\mathrm{Ker}\,\psi_m \subset \mathrm{Ker}\,\psi_d$ et par suite $\mathrm{Ker}\,\psi_n\,\mathrm{Ker}\,\psi_m \subset \mathrm{Ker}\,\psi_d$. L'autre inclusion est beaucoup moins évidente (il n'est même pas évident que $\mathrm{Ker}\,\psi_n\,\mathrm{Ker}\,\psi_m$ soit un sous-groupe !). Si on travaillait avec des groupes finis on pourrait utiliser un argument de cardinalité. C'est ce qu'on va faire en introduisant[1] un groupe fini intermédiaire, à savoir $\mathbb{Z}/e\mathbb{Z}$ où e désigne le ppcm de n et m. Notons ρ_n (resp. ρ_m et ρ_d) le morphisme naturel de $\mathrm{SL}_2(\mathbb{Z}/e\mathbb{Z})$ dans $\mathrm{SL}_2(\mathbb{Z}/n\mathbb{Z})$ (resp. $\mathrm{SL}_2(\mathbb{Z}/m\mathbb{Z})$ et $\mathrm{SL}_2(\mathbb{Z}/d\mathbb{Z})$). On a bien entendu $\psi_n = \rho_n \circ \psi_e$ et de même pour les autres. En particulier ρ_n est un morphisme surjectif puisque ψ_n l'est d'après la première question.

• On a $\mathrm{Ker}\,\psi_n = \psi_e^{-1}(\mathrm{Ker}\,\rho_n)$, et de même pour les autres noyaux. Il suffit alors de prouver que $\mathrm{Ker}\,\rho_d = \mathrm{Ker}\,\rho_n\,\mathrm{Ker}\,\rho_m$. En effet, supposons que cela soit établi et montrons que $\mathrm{Ker}\,\psi_n\,\mathrm{Ker}\,\psi_m = \mathrm{Ker}\,\psi_d$. On a déjà vu l'inclusion triviale ci-dessus. Soit $\mathrm{A} \in \mathrm{Ker}\,\psi_d$. Alors $\psi_e(\mathrm{A}) \in \mathrm{Ker}\,\rho_d$, donc on peut écrire $\psi_e(\mathrm{A}) = \mathrm{B}'\mathrm{C}'$ avec $\mathrm{B}' \in \mathrm{Ker}\,\rho_n$ et $\mathrm{C}' \in \mathrm{Ker}\,\rho_m$. Comme ψ_e est surjectif, on peut écrire $\mathrm{B}' = \psi_e(\mathrm{B})$ avec $\mathrm{B} \in \mathrm{Ker}\,\psi_n$ et $\mathrm{C}' = \psi_e(\mathrm{C})$ avec $\mathrm{C} \in \mathrm{Ker}\,\psi_m$. On a alors $\psi_e(\mathrm{A}) = \psi_e(\mathrm{BC})$, et donc $\mathrm{A} = \mathrm{BCD}$ avec $\mathrm{D} \in \mathrm{Ker}\,\psi_e$. C'est une décomposition de la forme voulue car $\mathrm{CD} \in \mathrm{Ker}\,\psi_m$ puisque m divise e.

• On va donc s'attacher à prouver que $\mathrm{Ker}\,\rho_d = \mathrm{Ker}\,\rho_n\,\mathrm{Ker}\,\rho_m$, égalité qui a lieu dans le groupe fini $\mathrm{SL}_2(\mathbb{Z}/e\mathbb{Z})$. Pour cela, on note que l'inclusion $\mathrm{Ker}\,\rho_n\,\mathrm{Ker}\,\rho_m \subset \mathrm{Ker}\,\rho_d$ est toujours triviale. Pour obtenir l'égalité on va montrer que les deux parties ont le même cardinal ! Notons tout d'abord que $\mathrm{Ker}\,\rho_n \cap \mathrm{Ker}\,\rho_m = \{\psi_e(\mathrm{I}_2)\}$ de sorte que le cardinal de $\mathrm{Ker}\,\rho_n\,\mathrm{Ker}\,\rho_m$ est égal au produit $|\mathrm{Ker}\,\rho_n| \times |\mathrm{Ker}\,\rho_m|$. En effet, considérons un élément de $\mathrm{Ker}\,\rho_n \cap \mathrm{Ker}\,\rho_m$ qu'on écrit $\psi_e(\mathrm{A})$ où $\mathrm{A} \in \mathrm{SL}_2(\mathbb{Z})$. On a alors $\mathrm{A} \equiv \mathrm{I}_2\ [n]$ et $\mathrm{A} \equiv \mathrm{I}_2\ [m]$ de sorte que $\mathrm{A} \equiv \mathrm{I}_2\ [e]$ et finalement $\psi_e(\mathrm{A}) = \psi_e(\mathrm{I}_2)$. Il est alors facile d'en déduire que l'application qui à $(\mathrm{U}, \mathrm{V}) \in \mathrm{Ker}\,\rho_n \times \mathrm{Ker}\,\rho_m$ associe le produit UV est injective et établit donc une bijection entre $\mathrm{Ker}\,\rho_n \times \mathrm{Ker}\,\rho_m$ et $\mathrm{Ker}\,\rho_n\,\mathrm{Ker}\,\rho_m$.

• Les cardinaux des noyaux ci-dessus se calculent : par surjectivité de ρ_n on a $|\mathrm{Ker}\,\rho_n| \times |\mathrm{SL}_2(\mathbb{Z}/n\mathbb{Z})| = |\mathrm{SL}_2(\mathbb{Z}/e\mathbb{Z})|$. Ainsi, on doit, pour conclure, démontrer la formule suivante :

1. Nous en profitons pour remercier M. Franck Taïcb qui nous a suggéré cette idée.

$$|\mathrm{SL}_2(\mathbb{Z}/n\mathbb{Z})| \times |\mathrm{SL}_2(\mathbb{Z}/m\mathbb{Z})| = |\mathrm{SL}_2(\mathbb{Z}/d\mathbb{Z})| \times |\mathrm{SL}_2(\mathbb{Z}/e\mathbb{Z})|.$$

Cela nous amène donc au problème combinatoire suivant : quel est le cardinal de $\mathrm{SL}_2(\mathbb{Z}/n\mathbb{Z})$, $n \geqslant 1$ étant quelconque ? Le lecteur a déjà rencontré un cas particulier de cette question dans l'exercice 3.22 où n était une puissance de 2. On va ici aussi se ramener aux puissances de nombres premiers. Notons $n = p_1^{\alpha_1} \dots p_s^{\alpha_s}$ la factorisation de n en nombres premiers. Comme précédemment, on dispose d'un morphisme naturel f de $\mathrm{SL}_2(\mathbb{Z}/n\mathbb{Z})$ dans le groupe produit $\mathrm{SL}_2(\mathbb{Z}/p_1^{\alpha_1}\mathbb{Z}) \times \cdots \times \mathrm{SL}_2(\mathbb{Z}/p_s^{\alpha_s}\mathbb{Z})$ qui à la classe d'une matrice A modulo n associe le r-uplet formé de ses classes modulo $p_1^{\alpha_1}, \dots, p_r^{\alpha_r}$. Ce morphisme est injectif, car si $\mathrm{A} \equiv \mathrm{I}_2\ [p_i^{\alpha_i}]$ alors $\mathrm{A} \equiv \mathrm{I}_2\ [n]$. Il est aussi surjectif par le théorème chinois. En effet, si $\mathrm{A}_1, \dots, \mathrm{A}_r$ sont des matrices données dans $\mathrm{SL}_2(\mathbb{Z})$ il existe A telle que $\mathrm{A} \equiv \mathrm{A}_i\ [p_i^{\alpha_i}]$ pour tout i. On a alors $\det \mathrm{A} \equiv \det \mathrm{A}_i \equiv 1\ [p_i^{\alpha_i}]$ pour tout i et donc $\det \mathrm{A} \equiv 1\ [n]$, de sorte que la réduction modulo n de A est dans $\mathrm{SL}_2(\mathbb{Z}/n\mathbb{Z})$ et a pour image le r-uplet formé des réductions des matrices A_i. Il en découle que f est un isomorphisme et en particulier que

$$|\mathrm{SL}_2(\mathbb{Z}/n\mathbb{Z})| = |\mathrm{SL}_2(\mathbb{Z}/p_1^{\alpha_1}\mathbb{Z})| \times \cdots \times |\mathrm{SL}_2(\mathbb{Z}/p_s^{\alpha_s}\mathbb{Z})|.$$

Il nous reste donc à trouver le cardinal de $\mathrm{SL}_2(\mathbb{Z}/p^\alpha\mathbb{Z})$ où p est un nombre premier. On procède comme dans l'exercice 3.22 avec $p = 2$. Considérons le morphisme naturel de $\mathrm{SL}_2(\mathbb{Z}/p^\alpha\mathbb{Z})$ dans $\mathrm{SL}_2(\mathbb{Z}/p\mathbb{Z})$. Il est surjectif et son noyau est de cardinal $(p^{\alpha-1})^3$. Comme

$$|\mathrm{SL}_2(\mathbb{Z}/p\mathbb{Z})| = \frac{|\mathrm{GL}_2(\mathbb{Z}/p\mathbb{Z})|}{p-1} = \frac{(p^2-1)(p^2-p)}{p-1} = p(p-1)(p+1),$$

on obtient finalement $|\mathrm{SL}_2(\mathbb{Z}/p^\alpha\mathbb{Z})| = (p-1)(p+1)p^{3\alpha-2}$. Pour conclure, il suffit d'observer que pour tout nombre premier p on a $\nu_p(n)+\nu_p(m) = \nu_p(d)+\nu_p(e)$ où ν_p désigne la valuation p-adique. Cela provient simplement de ce que $nm = de$! ◁

3.24. Surjection de $\mathrm{GL}_2(\mathbb{Z}/m\mathbb{Z})$ dans $\mathrm{GL}_2(\mathbb{Z}/n\mathbb{Z})$

Soient m, n dans $\mathbb{N}^*$ avec $n|m$. Trouver une surjection naturelle de $\mathrm{GL}_2(\mathbb{Z}/m\mathbb{Z})$ dans $\mathrm{GL}_2(\mathbb{Z}/n\mathbb{Z})$. On pourra d'abord étudier le cas où m et n ont les mêmes facteurs premiers.

(ENS Lyon)

▷ **Solution.**

Il y a une surjection naturelle de $\mathbb{Z}/m\mathbb{Z}$ sur $\mathbb{Z}/n\mathbb{Z}$ à savoir l'application qui, pour tout entier a, associe à la classe de a modulo m sa classe modulo n. Cette application est bien définie car si $a \equiv b \pmod m$ alors $a \equiv b \pmod n$. Elle induit naturellement une application π de $\mathcal{M}_2(\mathbb{Z}/m\mathbb{Z})$ dans $\mathcal{M}_2(\mathbb{Z}/n\mathbb{Z})$. On va commencer par montrer que π envoie $\mathrm{GL}_2(\mathbb{Z}/m\mathbb{Z})$ dans $\mathrm{GL}_2(\mathbb{Z}/n\mathbb{Z})$.

Étant donnée une matrice A à coefficients dans $\mathbb{Z}$ on notera $\pi_n(\mathrm{A})$ la matrice de $\mathcal{M}_2(\mathbb{Z}/n\mathbb{Z})$ obtenue par réduction des coefficients de A modulo n.

Considérons une matrice de $\mathrm{GL}_2(\mathbb{Z}/m\mathbb{Z})$ que l'on écrit $\pi_m(\mathrm{A})$ avec $\mathrm{A} \in \mathcal{M}_2(\mathbb{Z})$. Il existe $\mathrm{B} \in \mathcal{M}_2(\mathbb{Z})$ tel que $\pi_m(\mathrm{A})\pi_m(\mathrm{B}) = \pi_m(\mathrm{AB}) = \pi_m(\mathrm{I}_2)$. On a alors $\pi_n(\mathrm{A})\pi_n(\mathrm{B}) = \pi_n(\mathrm{AB}) = \pi_n(\mathrm{I}_2)$ car n divise m et $\pi_n(\mathrm{A})$ est inversible. Ainsi π va de $\mathrm{GL}_2(\mathbb{Z}/m\mathbb{Z})$ dans $\mathrm{GL}_2(\mathbb{Z}/n\mathbb{Z})$. On va démontrer qu'elle est surjective.

• On suppose pour commencer que m et n ont les mêmes facteurs premiers, chaque facteur apparaissant avec un exposant plus élevé dans m. Ainsi m apparaît comme le produit de n et de nombres premiers qui divisent n. En fait, il suffit de démontrer le résultat pour $m = np$. avec p premier divisant n et de réitérer le procédé.

Considérons une matrice de $\mathrm{GL}_2(\mathbb{Z}/n\mathbb{Z})$, que l'on écrit $\pi_n(\mathrm{A})$ avec $\mathrm{A} \in \mathcal{M}_2(\mathbb{Z})$. Il existe $\mathrm{B} \in \mathcal{M}_2(\mathbb{Z})$ tel que $\pi_n(\mathrm{A})\pi_n(\mathrm{B}) = \pi_n(\mathrm{AB}) = \pi_n(\mathrm{I}_2)$. Autrement dit, il existe $\mathrm{U} \in \mathcal{M}_2(\mathbb{Z})$ tel que $\mathrm{AB} = \mathrm{I}_2 + n\mathrm{U}$. On cherche $\mathrm{A}' \in \mathcal{M}_2(\mathbb{Z})$ telle que $\pi_m(\mathrm{A}') \in \mathrm{GL}_2(\mathbb{Z}/m\mathbb{Z})$ et $\pi_n(\mathrm{A}') = \pi_n(\mathrm{A})$. On cherche donc A′ sous la forme $\mathrm{A}' = \mathrm{A} + n\mathrm{V}$, avec $\mathrm{V} \in \mathcal{M}_2(\mathbb{Z})$. On a alors $\mathrm{A}'\mathrm{B} = \mathrm{AB} + n\mathrm{AV} = \mathrm{I}_2 + n\mathrm{U} + n\mathrm{AV}$ et $\pi_m(\mathrm{B})$ est l'inverse de $\pi_m(\mathrm{A}')$ si et seulement si $\mathrm{A}'\mathrm{B} \equiv \mathrm{I}_2\ [m]$, c'est-à-dire si $n\mathrm{U} + n\mathrm{AV} \equiv 0\ [np]$ ou encore $\mathrm{U} + \mathrm{AV} \equiv 0\ [p]$. Par hypothèse, p divise n donc $\mathrm{AB} = \mathrm{I}_2 + n\mathrm{U}$ implique $\mathrm{AB} \equiv \mathrm{I}_2\ [p]$. On en déduit $\mathrm{ABU} \equiv \mathrm{U}\ [p]$, ce qui montre que l'on peut choisir $\mathrm{V} = -\mathrm{BU}$.

• Examinons maintenant le cas où m possède des facteurs premiers qui n'apparaissent pas dans n. Compte tenu de ce qui précède, il suffit de démontrer la propriété pour $m = pn$, avec p premier ne divisant pas n et de réitérer le procédé.

On reprend la démonstration précédente. Elle est inchangée jusqu'à $\mathrm{U} + \mathrm{AV} \equiv 0\ [p]$. Si $\pi_p(\mathrm{A})$ est inversible dans $\mathcal{M}_2(\mathbb{Z}/p\mathbb{Z})$ d'inverse $\pi_p(\mathrm{C})$, on peut prendre $\mathrm{V} = -\mathrm{CU}$. Montrons qu'on peut se ramener à ce cas. Il suffit de montrer, dans le cas où $\pi_p(\mathrm{A})$ n'est pas inversible, qu'on peut trouver $\mathrm{S} \in \mathcal{M}_2(\mathbb{Z})$ telle que $\pi_p(\mathrm{A} + n\mathrm{S})$ soit inversible. On remarque que $\overline{n}$ est inversible dans le corps $\mathbb{Z}/p\mathbb{Z}$; on note n' un entier tel que $nn' \equiv 1\ [p]$. Si $\pi_p(\mathrm{A}) = 0$, la matrice $\mathrm{S} = n'\mathrm{I}_2$ convient. Reste à traiter le cas où $\pi_p(\mathrm{A})$ est de rang 1. Parmi les deux vecteurs colonnes C_1 et C_2

de A, l'un n'est pas nul modulo p. Supposons que c'est C_1. On a alors, pour $\mathrm{X} \in \mathbb{Z}^2$,

$$\det(\mathrm{C}_1, \mathrm{C}_2 + n\mathrm{X}) \equiv n \det(\mathrm{C}_1, \mathrm{X}) \quad [p].$$

Si on choisit X tel que tel que $\pi_p(\mathrm{X})$ ne soit pas colinéaire à $\pi_p(\mathrm{C}_1)$ (c'est possible : dans $(\mathbb{Z}/p\mathbb{Z})^2$, il y a $p^2 - p \geqslant 2$ vecteurs non colinéaires à $\pi_p(\mathrm{C}_1)$), la matrice dont les colonnes sont C_1 et $\mathrm{C}_2 + n\mathrm{X}$ est inversible, car son déterminant n'est pas nul. Si S est la matrice dont les colonnes sont 0 et X, $\pi_p(\mathrm{A} + n\mathrm{S})$ est inversible, ce qui permet de conclure la démonstration. $\triangleleft$

Index

Achevé d'imprimer sur les presses de l'Imprimerie BARNÉOUD
B.P. 44 - 53960 BONCHAMP-LÈS-LAVAL
Dépôt légal : janvier 2006 - N° d'imprimeur : 512.061
Imprimé en France

Ce livre est le second tome d'algèbre, et le quatrième volume dans l'ordre de parution, du recueil d'exercices résolus de S. Francinou, H. Gianella et S. Nicolas.

Comme dans les volumes précédents, les auteurs ont tenu à présenter les solutions les plus pédagogiques possible, essayant « d'exposer clairement les idées et démarches de raisonnement, sans pour autant escamoter les détails ou calculs qui peuvent paraître évidents. »

Le premier chapitre de ce tome est consacré au déterminant et peut être abordé dès la première année en classe préparatoire. Le second chapitre sur la réduction des endomorphismes constitue le cœur du programme d'algèbre de seconde année, et il est le plus riche des trois. Le dernier chapitre, aux exercices plus difficiles, est dédié à l'étude du groupe linéaire. Dans chacun de ces trois chapitres la difficulté est en général croissante : on commence par des questions techniques ou des savoir-faire indispensables (calculs de déterminants, recherche de valeurs propres, générateurs du groupe linéaire...) et on termine par des exercices plus théoriques.

Aux concours de l'X et des ENS, les énoncés proposés aux candidats sont souvent des résultats intéressants par eux-mêmes. Les auteurs ont fait l'effort de les identifier et de les présenter, autant que possible, dans leur contexte naturel. C'est ainsi qu'à l'intérieur d'un chapitre les exercices sont regroupés par thèmes et insérés dans un texte de présentation qui tour à tour dégage des idées générales, apporte des prolongements ou effectue quelques rappels de cours.

Le troisième et dernier tome d'algèbre portera sur les espaces euclidiens, les espaces hermitiens, les formes quadratiques et la géométrie.

Collection enseignement des mathématiques

ISBN 2-84225-091-5

9 782842 250911

Graphisme : Massin