

Épreuve orale d'analyse de document scientifique en Mathématiques

1. Extraits du rapport du concours 2013 (la consultation du texte intégral du rapport disponible sur le site www.polytechnique.edu est fortement recommandée)

L'épreuve orale proprement dite qui débute après deux heures de préparation, dure quarante minutes et se compose d'un exposé de quinze à vingt minutes suivi d'une séance de questions menée par l'examinateur. Les sujets sont des textes mathématiques d'origines variées qui incluent tous de vrais résultats mathématiques accessibles à tous les taupins de la filière MP maîtrisant tous les aspects du programme officiel. Il n'y a pas de sujet type. Les textes présentent cependant un point commun primordial, celui de pouvoir se prêter à une analyse scientifique avec les seules connaissances du programme officiel. Celle-ci doit avoir pour base le dossier mis à la disposition du candidat, lequel s'attachera à développer son analyse de façon cohérente. L'analyse du document scientifique doit se focaliser en premier lieu strictement sur le texte fourni et notamment sur ce qui apparaît être son thème principal. Des développements complémentaires ne se justifieront qu'à la condition expresse qu'ils ont un rapport direct avec le texte. Chaque dossier est accompagné d'une notice contenant des remarques générales communes à tous les sujets et des remarques particulières dont le candidat tiendra compte pour la préparation de son analyse.

2. Les textes présentés

ADS-M-1

Titre : "3-cycle implique chaos".

Source : Jean-Yves Briend, *Le théorème de Sharkovskii*, Le journal de maths des élèves de l'ENS Lyon, Volume 1 (1995), No. 3.

Sujet : Il s'agit de démontrer le théorème de Sharkovskii qui affirme qu'une fonction continue du segment $[0, 1]$ dans lui même admet des points périodiques de toutes périodes dès qu'elle admet un point de période 3.

ADS-M-2

Titre : "Moyennes géométriques de matrices positives".

Source : R. Bhatia and J. Holbrook, *Noncommutative geometric means*, Math. Intelligencer, 28 (2006), 32-39.

Sujet : Introduire dans le cadre des matrices $n \times n$ définies positives une moyenne géométrique pour deux matrices, généralisant la moyenne géométrique usuelle \sqrt{AB} quand

$n = 1$. Généraliser cette construction au cas de trois matrices.

ADS-M-3

Titre : “Une conjecture sur les zéros d’un polynôme et de ses dérivées”.

Source : J. Draisma and J. P. de Jong, *On the Casas-Alvaro conjecture*, Newsletter of the EMS, 80 (2011), 29-33.

Sujet : Point sur la conjecture de Casas-Alvero, qui n’est résolue que dans quelques cas. Elle affirme qu’un polynôme unitaire à coefficients complexes de degré n qui a un zéro commun avec chacune de ses dérivées d’ordre au plus $n - 1$ est nécessairement de la forme $(X - a)^n$. Présentation de méthodes *ad hoc* en petite dimension, contre-exemples, méthode p -adique.

ADS-M-4

Titre : “Somme non entière”.

Source : M. Müller and D. Schleicher, *How to add a noninteger number of terms : from axioms to new identities*, Amer. Math. Monthly 118 (2011), 136-152.

Sujet : Étendre la définition du symbole $\sum_{\nu=x}^y f(\nu)$, classique pour des entiers $x \leq y$ et une fonction f définie sur \mathbb{C} à valeurs complexes, au cas où x et y sont des nombres complexes.

ADS-M-5

Titre : “Pavages”

Source : Pierre de La Harpe, *Pavages*, publié par l’Université de Neuchâtel, 2011.

Sujet : Le texte expose quelques problèmes, certains encore ouverts, concernant les pavages du plan. Les pavages apériodiques de Penrose sont aussi discutés.

ADS-M-6

Titre : “Théorème de Perron-Frobenius”

Source : Jean-Étienne Rombaldi, *Matrices positives et irréductibles*, extrait d’un texte publié par l’Université Joseph Fourier, 2007.

Sujet : Étude du rayon spectral des matrices à coefficients strictement positifs.

ADS-M-1

3-cycle implique chaos

Travail demandé :

Il vous est demandé d'étudier puis de présenter le(s) texte(s) joint(s) à travers un exposé de synthèse d'une durée comprise entre 15 et 20 minutes.

Si l'étude de la totalité du dossier et la préparation d'un exposé cohérent dans la durée impartie ne vous paraît pas possible, vous pouvez décider de vous limiter à une partie du dossier.

Remarques générales :

1. Les textes proposés, quelle que soit leur origine, peuvent présenter des défauts (coquilles typographiques, négligences ou sous-entendus de l'auteur, voire erreurs...) qui, sauf exception, n'ont pas été corrigés.

2. Les textes proposés peuvent contenir des exercices qu'il n'est pas demandé de résoudre. Néanmoins, vous pouvez vous aider des énoncés de ces exercices pour enrichir votre exposé.

3. Vous pouvez annoter les documents qui vous sont fournis. Vos annotations ne seront pas regardées par l'examineur.

3-cycle implique chaos

Travaux pratiques

Au pays des publications mathématiques, il est rare de commencer un article par une section bricolage. C'est pourtant ce que nous allons faire en vous proposant de réaliser vous-mêmes un montage facile et peu onéreux. Tout d'abord, il vous faut vous procurer deux récipients cylindriques, de même hauteur mais l'un de volume huit fois plus grand que l'autre. Placez alors le plus petit au centre du plus grand de manière à confondre leurs axes. On a ainsi réalisé un récipient annulaire. Voir la figure 1.

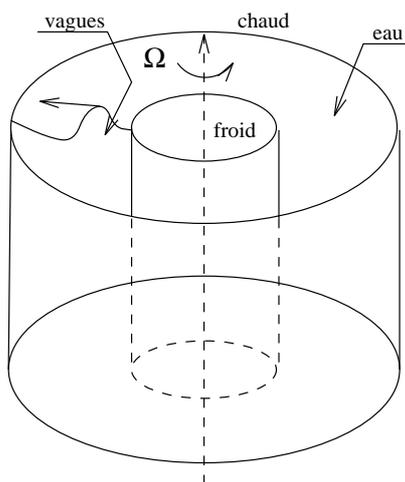


Figure 1: L'expérience de Lorentz

Une fois ce contenant réalisé, on peut le remplir d'eau et faire tourner le tout suivant son axe, mis vertical, sur un tourne-disque en position 78 tours par minutes. Il ne vous reste plus qu'à imposer entre le centre et la périphérie une forte différence de température, le tout continuant de tourner. Vous obtenez ainsi l'expérience réalisée par Lorentz dans les années 60 (voir [1]–[3]) dans son étude des phénomènes météorologiques. Si vous avez beaucoup de chance, vous observerez comme lui que des vagues d'aspect évoluant sans loi apparente se forment à un bord du récipient pour aller mourir sur l'autre. Lorentz ne se laissa pas impressionner par cet état de fait semble-t-il impénétrable à toute mise en loi et effectua des manipulations numériques sur les équations du système. Il

trouva alors pour l'évolution des vagues le modèle suivant : l'énergie cinétique maximale K_{n+1} de la $(n+1)$ -ème vague est fonction de celle de la précédente sous la forme

$$K_{n+1} = f(K_n)$$

où f est une application continue donnée de l'intervalle $[0, 450]$ dans lui-même. Ainsi Lorentz fut amené à étudier l'itération de la fonction f , c'est-à-dire à comprendre le comportement des orbites des points x de l'intervalle, définies par

$$\mathcal{O}(x) = \{f^n(x), n \geq 0\}$$

où $f^0 = id$ et $f^n = f^{n-1} \circ f$ est la n -ième itérée de f . On se retrouve donc en partant de deux boîtes de choucroute à commencer l'étude des *systèmes dynamiques en dimension 1*, théorie qui connaît depuis deux décennies des développements considérables. Dans cet article, nous allons donner la démonstration d'un théorème dû à Šarkovskiĭ en 1964 (voir [4]) qui a le mérite d'être tout à la fois simple à énoncer, d'avoir une démonstration s'appuyant sur des outils élémentaires, et cependant d'être un résultat profond et très plaisant.

Enoncés des résultats

En 1975, Li et Yorke (voir [5]) publient un article qui allait avoir un certain retentissement, sous le titre accrocheur de « Période 3 Implique Chaos ». En fait, leur résultat était un corollaire du théorème général de Šarkovskiĭ cité plus haut, et que nous nous proposons de montrer ici.

THÉORÈME 1 (PÉRIODE 3 IMPLIQUE CHAOS). *Si $f : [0, 1] \rightarrow [0, 1]$ est une application continue ayant un point périodique de période 3, alors elle a des points périodiques de toutes périodes.*

Dans la suite, on notera $I = [0, 1]$ et on considérera une application continue $f : I \rightarrow I$. En général, si $A \subseteq I$, on notera $fA = f(A)$ l'image de A par f .

DÉFINITION 1. *Un point $x \in I$ est dit périodique de période $n > 1$ si $f^n x = x$ et $f^i x \neq x$ pour $i = 1, \dots, n-1$. On dit que f est un point fixe si $fx = x$.*

Ainsi quand un point est périodique, son orbite est finie mais en plus f agit sur celle-ci bijectivement, comme un cycle.

Nous pouvons maintenant passer à l'énoncé général du théorème de Šarkovskiĭ. Pour cela, nous avons besoin de définir un ordre spécial sur l'ensemble \mathbf{N}^* des nombres entiers strictement positifs.

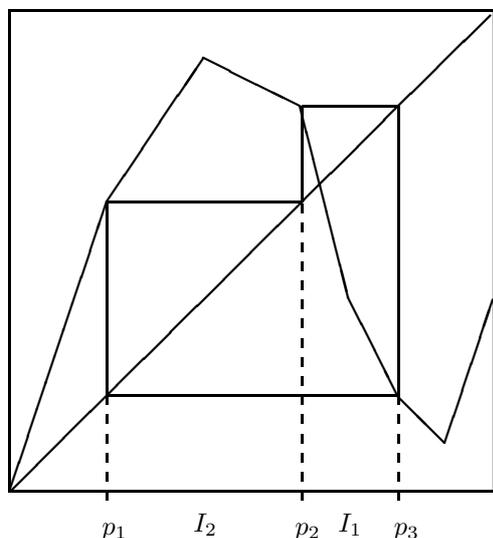


Figure 2: Le point p_1 est de période 3

DÉFINITION 2. On appelle ordre de Šarkovskii sur \mathbf{N}^* l'ordre \succ défini comme suit :

$$3 \succ 5 \succ 7 \succ 9 \succ \dots \succ 2.3 \succ 2.5 \succ 2.7 \succ \dots \succ 2^n.3 \succ 2^n.5 \succ \dots \succ 2^{n+1}.3 \succ 2^{n+1}.5 \succ \dots \succ 2^n \succ 2^{n-1} \succ \dots \succ 4 \succ 2 \succ 1$$

et c'est un ordre total.

THÉORÈME 2 (ŠARKOVSKIĬ). Soit $f : I \rightarrow I$ une application continue ayant un point périodique de période n . Alors pour tout m vérifiant $n \succ m$, f admet un point périodique de période m .

Le premier théorème est donc conséquence immédiate de ce dernier.

Un lemme facile

Dans la suite on se donne une application continue f de I dans lui-même. Commençons par faire une remarque cruciale pour la suite, et dont la démonstration est laissée en exercice. Supposons qu'il existe dans I deux intervalles fermés J_0 et J_1 tels que $f(J_0) \supseteq J_1$. Alors il existe un intervalle $J'_0 \subseteq J_0$ tel que $f(J'_0) = J_1$. Ce résultat acquis, nous pouvons énoncer le lemme 1, à la base de toute la démonstration du théorème de Šarkovskii.

LEMME 1. On a les résultats suivants :

1. soit $J \subseteq I$ un sous-intervalle de I tel que $J \subseteq f(J)$. Alors f a un point fixe dans J .

2. Soit $(I_i)_{i \in \mathbf{N}}$ une suite éventuellement finie de sous-intervalles fermés de I tels que $f(I_i) \supseteq I_{i+1}$. Alors il existe une suite d'intervalles fermés emboîtés J_n telle que

$$J_n \subseteq I_0 \quad \text{et} \quad f^n(J_n) = I_n.$$

De plus, il existe $x \in I_0$ tel que $f^n x \in I_n$ pour tout n .

DÉMONSTRATION. Posons $\bar{J} = [a, b]$, par hypothèse il existe un $z \in \bar{J}$ tel que $f(z) \leq a$ et un $w \in \bar{J}$ tel que $f(w) \geq b$. Posons alors $g(x) = f(x) - x$. On a

$$g(z) = f(z) - z \leq f(z) - a \leq 0$$

$$g(w) = f(w) - w \geq f(w) - b \geq 0$$

et donc par le théorème des valeurs intermédiaires, il existe un $x \in \bar{J}$ tel que $g(x) = 0$ i.e. $f(x) = x$. La démonstration du deuxième point se fait par récurrence sur n .

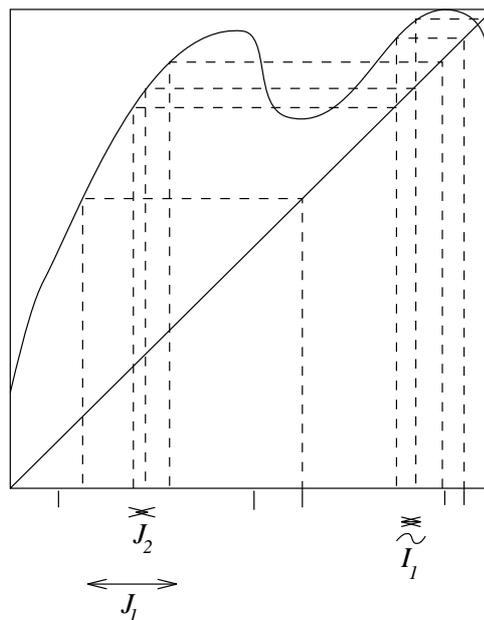


Figure 3: La récurrence

Pour $n = 1$, $fI_0 \supseteq I_1$ et donc il existe un intervalle $J_1 \subseteq I_0$ tel que $fJ_1 = I_1$. Supposons maintenant J_1, \dots, J_n déjà construits. Puisque $fI_n \supseteq I_{n+1}$ il existe un intervalle $\tilde{I}_n \subseteq I_n$ tel que $f\tilde{I}_n = I_{n+1}$. Or par hypothèse de récurrence, $f^n J_n = I_n \supseteq \tilde{I}_n$ et par le même procédé, on trouve $J_{n+1} \subseteq J_n$ tel que $f^n J_{n+1} = \tilde{I}_n$. Maintenant il vient $f^{n+1} J_{n+1} = f\tilde{I}_n = I_{n+1}$ (voir la figure 3).

Ensuite, pour avoir le point x du lemme, il suffit de prendre un point dans

$$\bigcap_{n \geq 0} J_n$$

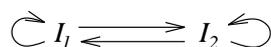
qui est non-vide par compacité. Ceci achève la démonstration du lemme 1. ■

Le Graphe de Markov

Comme le montre le lemme 1, il y a un lien fort entre l'orbite d'un point et son cheminement dans une famille de sous-intervalles (son *itinéraire*). Cela motive la définition suivante :

DÉFINITION 3. Soit $\{I_k\}_k$ un ensemble fini de sous intervalles fermés de I d'intérieurs disjoints deux à deux (on dira dans la suite une partition). Alors le graphe de Markov associé à la partition $\{I_k\}$ et à la fonction f est le graphe orienté dont les sommets sont les éléments de la partition $\{I_k\}$ et qui comporte une arête $I_i \rightarrow I_j$ si et seulement si $f(I_i) \supseteq I_j$.

EXEMPLE 1. L'application tente avec la partition donnée sur la figure a pour graphe de Markov :



et l'on voit qu'il peut y avoir des arêtes allant d'un sommet à lui même.

Cette définition permet de relire le lemme 1 : si l'on a un chemin dans le graphe de Markov, alors il existe un point dont l'itinéraire dans la partition est donné précisément par ce chemin.

L'idée qui sous-tend la définition du graphe de Markov est centrale dans l'étude des systèmes dynamiques. Elle consiste à découper notre système en un nombre fini de morceaux pour ensuite étudier la dynamique comme agissant sur l'ensemble de ces morceaux. On a « discrétisé » la dynamique. Pris de myopie, nous devons diviser l'image que nous avons en pièces assez grosses, et comme à la télévision, regarder le monde à travers un damier simplificateur. Ce qui est remarquable, c'est qu'un tel damier, même grossier, peut nous donner des informations très profondes sur le système dynamique étudié.

Démonstration du premier théorème

En fait, toute la démonstration du théorème de Šarkovskii repose sur l'étude du graphe de Markov

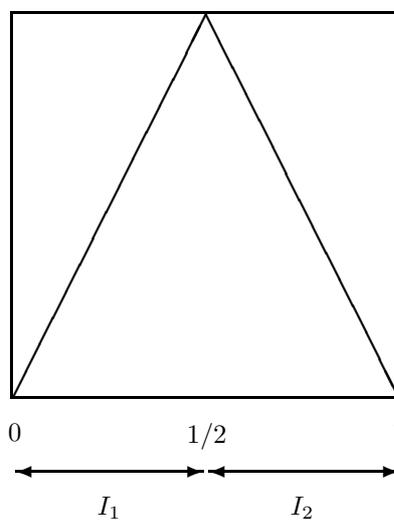


Figure 4: L'application « tente »

de f associé à une bonne partition. Cette étude est faite dans les lemmes 2–7. Cependant, nous pouvons dès maintenant démontrer le premier théorème dit « période trois implique chaos ».

DÉMONSTRATION. Supposons donc que f ait un point p_1 de période 3. Nous ne traitons que le cas où l'orbite périodique s'ordonne en $p_1 < p_2 < p_3$ avec $fp_1 = p_2$, $fp_2 = p_3$ et $fp_3 = p_1$. L'autre cas se fait de la même manière. Regardons la partition associée à cette orbite : $I_1 = [p_2, p_3]$, $I_2 = [p_1, p_2]$ (voir la figure 2). Comme alors $fI_1 \supseteq I_1 \cup I_2$ et que $fI_2 \supseteq I_1$, on a un graphe de Markov donné par



Choisissons maintenant de bons chemins dans ce graphe pour produire des orbites périodiques de toutes les périodes.

Soit d'abord $m > 3$ et montrons que f admet un point de période m . Regardons en effet dans le graphe de Markov le chemin suivant

$$\underbrace{I_1 \rightarrow I_1 \rightarrow \dots \rightarrow I_1}_{m-2} \rightarrow I_2 \rightarrow I_1$$

L'intervalle $J = J_m$ donné par la deuxième partie du lemme 1 appliqué à la famille $(I_k)_k$ vérifie

$$\begin{aligned} f^i J &\subseteq I_1 & 0 \leq i \leq m-2 \\ f^{m-1} J &\subseteq I_2 \\ f^m J &= I_1 \end{aligned}$$

et par la première partie de ce même lemme, il existe un point x de J tel que $f^m x = x$.

Montrons que x est de période exactement m . Supposons par l'absurde que $f^i x = x$, avec i compris entre 1 et $m - 1$. Alors $I_2 \ni f^{m-1} x = f^{i-1} x \in I_1$ et donc $f^{m-1} x = p_2$ qui est l'unique point de $I_1 \cap I_2$. Finalement $x = f^m x = p_3$ ce qui est impossible car $f p_3 = p_1 \notin I_1$ ce qui contredit l'itinéraire de x .

Pour montrer que f a un point de période 2, on regarde le chemin

$$I_2 \rightarrow I_1 \rightarrow I_2$$

et l'on conclut de la même manière. ■

La manière dont on a construit notre partition est ici très importante. L'ensemble des extrémités des intervalles constituant celle-ci est en effet une orbite périodique $\mathcal{O}(x)$. Celle-ci est stable par f . Mais surtout, si le point périodique construit grâce à un chemin dans le graphe de Markov n'est pas de la période désirée, une de ses images par f va être 'à cheval' sur deux intervalles de la partition, et donc son orbite rencontre $\mathcal{O}(x)$ ce qui sera une contradiction.

Énoncé des lemmes préparatoires

Nous allons maintenant donner l'énoncé des lemmes qui, en étudiant des graphes de Markov, permettent de démontrer le théorème de Šarkovskii. Nous démontrerons ces lemmes plus loin dans l'article.

Soit donc f une application continue de I dans lui-même, ayant un point x de période $n \geq 2$ et rangeons dans I les points de l'orbite de x par ordre croissant

$$x_0 < x_1 < \dots < x_{n-1}$$

Cela nous donne une partition en $n - 1$ intervalles fermés, définie par les $[x_i, x_{i+1}]$, $i = 1, \dots, n - 1$. Nous allons décrire les propriétés du graphe de Markov Γ associé à cette partition.

LEMME 2. *Il existe un intervalle I_1 de la partition tel que l'on ait une flèche de I_1 vers lui-même.*

En d'autre terme, on a dans le graphe Γ un sommet tel que

$$\curvearrowright I_1$$

ce qui sera bien pratique pour faire des chemins de longueur quelconque dans le graphe, en bouclant autant de fois qu'il le faut sur I_1 . Dans toute la suite, I_1

désignera un tel intervalle. Regardons maintenant ce qui se passe si l'on cherche à aller de I_1 vers un autre sommet :

LEMME 3. *Pour tout sommet K du graphe de Markov il existe un chemin dans ce graphe partant de I_1 et aboutissant en K .*

Autrement dit, en partant de I_1 , on peut aller n'importe où. Que se passe-t-il en revanche lorsque l'on souhaite arriver en I_1 ? Le lemme suivant nous donne la réponse :

LEMME 4. *Supposons qu'il n'existe pas de sommet du graphe de Markov distinct de I_1 , duquel on puisse partir pour aboutir par un chemin du graphe en I_1 . Alors n est pair, et f envoie tous les points de l'orbite de x qui sont à gauche de I_1 sur ceux qui sont à droite, et vice-versa. Enfin, il existe un point de période 2.*

Cela veut dire que si l'on ne peut aboutir en I_1 , n est pair et l'on a une description combinatoire simple de l'orbite de x .

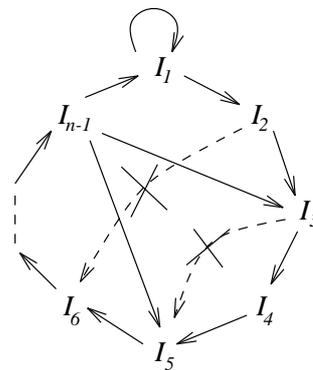


Figure 5: Le lemme à la boucle

LEMME 5. *Supposons que f ait un point de période impaire différente de 1 et supposons que x soit un tel point, de période $n \neq 1$ prise minimale. Alors avec I_1 comme dans le lemme 2, on a*

1. Γ contient le cycle

$$I_1 \rightarrow I_2 \rightarrow \dots \rightarrow I_{n-1} \rightarrow I_1$$

2. Γ contient les arêtes de la forme

$$I_{n-1} \rightarrow I_{2i-1}, 2i - 1 < n$$

3. il n'existe pas d'arête de la forme

$$I_j \rightarrow I_{j+k}$$

pour $k > 1, j \geq 1$.

Tout devient plus clair en regardant la figure 5.

On peut immédiatement déduire de ces lemmes un corollaire simple

COROLLAIRE 1. *Si f a un point de période impaire $n \neq 1$, alors f a des points périodiques de toutes les périodes plus grandes que n et de toutes les périodes paires plus petites que n .*

La démonstration en est simple, au vu du lemme 5, en prenant un n minimal et en choisissant dans le graphe Γ dessiné dans la figure 5 le bon chemin. Pour les points de période plus grande que n , on boucle autant de fois qu'il le faut sur I_1 puis on fait la grande boucle. Pour les points de période paire plus petites que n , on prend un raccourci $I_{n-1} \rightarrow I_{2i+1}$. Ensuite, la remarque concluant la démonstration du premier théorème permet de conclure.

Énonçons maintenant les derniers lemmes, qui n'ont pas à proprement parler rapport avec le graphe de Markov.

LEMME 6. *Si f a un point de période paire, elle a un point de période 2.*

LEMME 7. *Soit c un point de période n pour f , et h un entier positif quelconque. Alors c est un point de période $n/(n,h)$ pour f^h , où (n,h) est le plus grand diviseur commun de h et n . Réciproquement, si c est un point de période m pour f^h , alors c est un point de période mh/d pour f , avec d qui divise h mais est premier avec m .*

Démonstration du théorème de Šarkovskii

Nous pouvons maintenant procéder à la preuve du deuxième théorème. Supposons donc que f ait une orbite de période n .

1. Regardons d'abord le cas où n est une puissance de 2, i.e. $n = 2^k$. Soit alors $m \prec n$. m est de la forme 2^e , avec $0 \leq e \leq k$. Par le corollaire 1 et le lemme 6, on peut supposer $e > 0$. Maintenant $g = f^{m/2}$ a un point de période 2^{k-e+1} par le lemme 7, et donc un point de période 2 par le lemme 6. Par le lemme 7, ce point est de période m pour f .

2. Reste maintenant le cas où n est de la forme $p \cdot 2^k$ où $p > 1$ est impair. Soit $m \prec n$. Il y a trois cas possibles :

- 1er cas $m = q \cdot 2^k$ $q > 0$ est pair
- 2ème cas $m = q \cdot 2^k$ $q > p$ impair
- 3ème cas $m = 2^l$ $l \leq k$

et ces trois cas sont les seuls. Comme dans la première partie de la démonstration, nous allons regarder la bonne itérée de f pour lui appliquer soit le corollaire 1 soit la première étape de la démonstration.

Dans les deux premiers cas, regardons $g = f^{2^k}$. Elle a un point de période p impaire. Comme soit $q > p$, soit q est pair, le corollaire 1 nous donne un point de période q pour g . Dans le premier cas, ce point est de période $q2^k$ pour f , d'après le lemme 7.

Dans le second cas, ce point est de période $q2^e$ pour f , pour un certain $e \leq k$, encore par le lemme 7. Si $e = k$ c'est fini et sinon on remplace n par $q2^e$. Alors comme $m = 2^e(q2^{k-e})$, on peut appliquer le premier cas, déjà démontré.

Dans le troisième cas, d'après le premier cas ci-dessus démontré, f a un point de période $2 \cdot 2^k = 2^{k+1}$. On peut alors appliquer la première étape de la démonstration car $l < k + 1$.

Cela achève la démonstration du théorème de Šarkovskii. Reste maintenant à démontrer les lemmes.

Démonstration des lemmes 2–7

Le lemme 2

Remarquons que l'orbite de x_0 est contenue dans l'intervalle $[x_0, x_{n-1}]$. Comme $n > 1$, on a

$$fx_0 > x_0 \text{ et } fx_{n-1} < x_{n-1},$$

on peut donc considérer

$$x_a = \max\{x_i : fx_i > x_i\}$$

et l'intervalle $I_1 = [x_a, x_{a+1}]$ convient. cqfd.

Dans la suite, c'est cet intervalle que nous prendrons égal à I_1 .

Le lemme 3

Si l'on part de I_1 , peut-on aboutir n'importe où ? Considérons pour un entier i l'ensemble V_i des sommets de Γ extrémités d'un chemin de longueur i partant de I_1 . Il nous faut démontrer que pour i assez grand, V_i est l'ensemble de tous les sommets. Remarquons d'abord que $V_i \neq \emptyset$ et que $V_{i+1} \supseteq V_i$, car on peut boucler sur I_1 . Posons

$$U_i = \bigcup_{K' \in V_i} K' \subseteq I$$

On a clairement $U_i \subseteq U_{i+1}$.

Affirmation : s'il existe $K' \in V_i$ tel que $f(\partial K') \not\subseteq U_i$ alors $V_{i+1} \neq V_i$.

En effet nous remarquons d'abord que l'ensemble des extrémités des intervalles constituant la partition forme une orbite, et donc est stable par f . Soit maintenant $w \in \partial K'$ tel que $fw \notin U_i$. Comme $K' \in V_i$, fK' contient un intervalle I_j . en particulier il en contient un qui a fw dans son bord, et celui-ci est bien dans $V_{i+1} - V_i$. L'affirmation est bien démontrée.

La suite V_i est croissante et finie donc il existe un i compris entre 0 et $n - 1$ tel que $V_{i+1} = V_i$. D'après l'affirmation ci-dessus, $U_i \cap \mathcal{O}(x_0)$ est stable par f , et donc égal à toute l'orbite de x_0 . Cela montre que V_i est égal à l'ensemble de tous les sommets du graphe. cqfd.

Le lemme 4

Rappelons que $I_1 = [x_a, x_{a+1}]$ comme dans la démonstration du lemme 2. Supposons par l'absurde qu'il existe $x_i < x_a$ tel que $fx_i \leq x_a$. Soit x_b le plus grand des $x_j < x_a$ tels que $fx_j \leq x_a$. Alors par définition $fx_b \leq x_a$ et $fx_{b+1} \geq x_{a+1}$. Il en découle une flèche dans le graphe de Markov :

$$[x_b, x_{b+1}] \rightarrow I_1.$$

Comme $x_b \neq x_a$ c'est impossible par hypothèse. On vérifie de la même manière que pour $x_i \geq x_{a+1}$, on a $fx_i \leq x_a$. Tout cela implique clairement la première partie du lemme, et que n est pair. Soit maintenant $J_0 = [x_0, x_a]$ et $J_1 = [x_{a+1}, x_{n-1}]$. Comme $fJ_0 \supseteq J_1$ et $fJ_1 \supseteq J_0$, et que $J_0 \cap J_1 = \emptyset$, il existe un x dans J_0 tel que $f^2x = x$, et comme $fx \in J_1$, x est bien de période 2. cqfd.

Le lemme 5

Par le lemme 2, on a l'intervalle I_1 . Par le lemme 3 et le lemme 4, comme n est impair, il existe $k > 1$ et des intervalles I_2, \dots, I_k de la partition tels que l'on ait dans le graphe le chemin

$$I_1 \rightarrow I_2 \rightarrow \dots \rightarrow I_k \rightarrow I_1$$

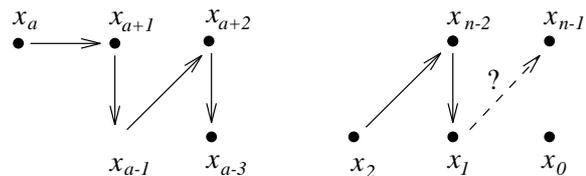
et soit prenons k minimal. Montrons que $k = n - 1$ Pour cela supposons par l'absurde que $k < n - 1$. Alors si k est impair le chemin $I_1 \rightarrow I_2 \rightarrow \dots \rightarrow I_k \rightarrow I_1$ fournit un point de période k ce qui contredit la minimalité de n . Et si k est pair, le même chemin en lui rajoutant une boucle sur I_1 à la fin donne un point de période impaire $k + 1 < n$ ce qui est encore une contradiction. Donc $k = n - 1$ et cela implique les première et troisième assertions du lemme. Reste à montrer la deuxième.

D'après le lemme 2, $fx_a \geq x_{a+1}$ et $fx_{a+1} \leq x_a$, et comme x_a n'est pas de période 2, l'une au moins des deux inégalités est stricte. Supposons par exemple que $fx_{a+1} < x_a$. Montrons qu'alors $fx_a = x_{a+1}$ et $fx_{a+1} = x_{a-1}$. En effet sinon on peut trouver un $j > 2$ tel que $fI_1 \supseteq I_j$, ce qui est interdit par la troisième assertion du lemme.

Donc $I_2 = [x_{a-1}, x_a]$. Enfin par définition de x_a et comme $fx_a = x_{a+1}$, que $fx_{a-1} > x_{a-1}$ et que fx_{a-1} ne peut être égal à x_a (par minimalité de $k = n - 1$) ni à x_{a+1} , on a $fx_{a-1} \geq x_{a+2}$. Alors par la deuxième assertion, on a en fait égalité et $I_3 = [x_{a+1}, x_{a+2}]$. On peut ainsi répéter cet argument, et finalement les I_i se rangent de la manière suivante :

$$\underbrace{I_{n-1} \dots I_4 I_2}_{\text{pairs}} \quad I_1 \quad \underbrace{I_3 I_5 \dots I_{n-2}}_{\text{impairs}}$$

et en fait on a une description complète de l'orbite périodique :



Comme

$$fI_{n-3} = f([x_1, x_2]) \supseteq I_{n-2} = [x_{n-2}, x_{n-1}]$$

et aucun autre intervalle, il vient $fx_1 = x_{n-1}$. Comme $fI_{n-1} \supseteq I_1$, on a $fI_{n-1} \supseteq [x_a, x_{n-1}]$ et donc pour $2i + 1 = 1, \dots, n - 2$, on a $fI_{n-1} \supseteq I_{2i+1}$, ce qui est la deuxième assertion du lemme. cqfd.

Le lemme 6

Soit x un point de période $n > 2$ minimale. Supposons par l'absurde que f n'ait pas de point de période 2. D'après le corollaire, si n était impair, il y aurait un tel point, et donc n est pair. D'après le lemme 4, il existe un sommet I_k de la partition, différent de I_1 , et tel que l'on ait $I_k \rightarrow I_1$. Soit maintenant $k > 1$ le plus petit entier tel qu'il existe un cycle $I_1 \rightarrow \dots \rightarrow I_k \rightarrow I_1$ injectif de longueur k dans le graphe de Markov. Comme dans la démonstration du lemme 5, on vérifie que $k = n - 1$ et qu'il n'y a pas d'arête de la forme $I_j \rightarrow I_j + l$, pour $l > 1$. On vérifie ensuite qu'il existe des arêtes reliant I_{n-1} à tous les I_{2i} , et on a en particulier le chemin $I_{n-1} \rightarrow I_{n-2} \rightarrow I_{n-1}$ qui nous donne un point de période 2, ce qui est impossible par hypothèse. cqfd.

Le lemme 7

Soit c un point de période n pour f , et posons $m = n/(h, n)$ où (h, n) désigne le plus grand commun diviseur de h et n . Si $f^{kh}c = c$ alors n divise kh et donc m divise k , et c'est fini.

Pour la deuxième partie du lemme, considérons c un point de période m pour f^h . Alors c est de période n pour f , où n divise mh . Donc nous pouvons écrire $n = mh/d$. Or d'après la première partie du lemme, $n/(h, n) = nd/h$ et donc $(h, n) = h/d$ ce qui nous donne $h = de$, et $(de, me) = e$. cqfd.

Conclusion

Voilà, nous sommes maintenant en possession d'un théorème qui donne une description complète de l'ensemble des périodes d'un endomorphisme de l'intervalle. Bien sûr, cela n'épuise pas les questions que l'on peut se poser sur l'itération des fonctions.

Tout d'abord, le théorème de Šarkovskii est-il optimal? La réponse est positive. Pour cela, on peut se reporter à [6], où il est montré que pour tout n il existe une application continue de l'intervalle dans lui-même ayant un point de période n mais pas de point de période m pour $m \succ n$. De même, il existe une application ayant des points de périodes toutes les puissances de 2, et aucune autre période. C'est donc un résultat optimal. On peut cependant aller plus loin dans l'analyse des orbites périodiques, et ajouter à la période une caractéristique combinatoire, i.e. la manière dont f chamboule l'ordre de l'orbite. C'est la théorie du forçage, qui cherche quels sont les types combinatoires nécessairement présents en présence d'un autre. Voir encore [6].

On peut ensuite se poser la question de savoir si le graphe de Markov ne pourrait pas nous donner d'autres informations sur la dynamique. Par exemple, dans [7], il est démontré que la plus grande valeur propre λ de la matrice d'adjacence du graphe de Markov est reliée à l'entropie topologique de l'application f par

$$h_{\text{top}}(f) \geq \log \lambda$$

Ainsi, l'application tente a une entropie supérieure ou égale à $\log 2$.

Enfin, existe-t-il des résultats de type Šarkovskii pour des applications continues sur d'autres espaces que des intervalles, par exemple des variétés compactes. Dans le cadre de la dimension 1, dans [7] il est démontré des

théorèmes de ce type pour des applications du cercle. Citons enfin [8], où est démontrée une conjecture de Misiurewicz sur les application continues des graphes : si f est une application continue d'un graphe Y dans lui-même, alors il existe un entier $L = L(Y)$ tel que si f a des points de période $1, \dots, L$, alors f a des points de toutes les périodes.

Pour avoir une vue d'ensemble moderne sur les systèmes dynamiques en dimension 1, on peut consulter le livre de W. de Melo et S. van Strien [9].

Bibliographie

- [1] E.N. Lorentz *Deterministic non periodic flows* J. Atmosph. Sci. **20**, 1963, 130–141
- [2] E.N. Lorentz *The mechanics of vacillation* J. Atmosph. Sci. **20**, 1963, 448–464
- [3] E.N. Lorentz *The predictability of hydrodynamic flow* Trans. N-Y. Acad. Sci. Ser. II **25**, 1963, 409–432
- [4] A.N. Šarkovskii *Coexistence of cycles of a continuous mapping of the line into itself* Ukrain. Math. Ž. **16**, 1964, 61–71 (en Russe, avec résumé en Anglais)
- [5] T.Y. Li & J.A. Yorke *Period three implies chaos* Amer. Math. Monthly **82**, 1975, 985–992
- [6] L.S. Block & W.A. Coppel *Dynamics in one dimension* Springer Lect. Notes in Math. **1513**, 1992
- [7] L. Block, J. Guckenheimer, M. Misiurewicz & L.S. Young *Periodic points and topological entropy of one dimensional maps* in Global theory of dynamical systems Z. Nitecki & C. Robinson ed. Springer Lect. Notes in Math. **819**, 1980
- [8] A.M. Blokh *Spectral decomposition, periods of cycles and a conjecture of M. Misiurewicz for graph maps* in Ergodic theory and related topics III U. Krengel, K. Richter & V. Warstat ed. Springer Lect. Notes in Math. **1514**, 1992
- [9] W. de Melo & S. van Strien *One-dimensional dynamics* Springer 1993
- [10] R.L. Devaney *An introduction to chaotic dynamical systems* Addison-Wesley 1987

ADS-M-2

Moyennes géométriques de matrices positives

Travail demandé :

Il vous est demandé d'étudier puis de présenter le(s) texte(s) joint(s) à travers un exposé de synthèse d'une durée comprise entre 15 et 20 minutes.

Si l'étude de la totalité du dossier et la préparation d'un exposé cohérent dans la durée impartie ne vous paraît pas possible, vous pouvez décider de vous limiter à une partie du dossier.

Remarques générales :

1. Les textes proposés, quelle que soit leur origine, peuvent présenter des défauts (coquilles typographiques, négligences ou sous-entendus de l'auteur, voire erreurs...) qui, sauf exception, n'ont pas été corrigés.

2. Les textes proposés peuvent contenir des exercices qu'il n'est pas demandé de résoudre. Néanmoins, vous pouvez vous aider des énoncés de ces exercices pour enrichir votre exposé.

3. Vous pouvez annoter les documents qui vous sont fournis. Vos annotations ne seront pas regardées par l'examineur.

Remarque particulière :

Les figures annoncées dans le texte sont regroupées à la fin du document.

Moyennes géométriques de matrices positives

Soit \mathbb{R}^+ l'ensemble des nombres positifs¹. Soient a et b dans \mathbb{R}^+ , une *moyenne* $m(a, b)$ peut être définie de plusieurs façons. Il est raisonnable d'espérer qu'une telle opération binaire m sur \mathbb{R}^+ ait les propriétés suivantes :

- (i) $m(a, b) = m(b, a)$.
- (ii) $\min(a, b) \leq m(a, b) \leq \max(a, b)$.
- (iii) $m(\alpha a, \alpha b) = \alpha m(a, b)$ pour tout $\alpha > 0$.
- (iv) $m(a, b)$ est une fonction croissante de a et b .
- (v) $m(a, b)$ est une fonction continue de a et b .

Les trois moyennes familières - arithmétique, géométrique et harmonique - satisfont ces cinq conditions. Plus généralement, on définit les *moyennes binômiales*, pour $-\infty \leq p \leq +\infty$ par

$$m_p(a, b) = \left(\frac{a^p + b^p}{2} \right)^{1/p} \quad \text{pour } -\infty \leq p \leq \infty,$$

où, pour les valeurs particulières p dans $\{0, -\infty, \infty\}$, on définit $m_p(a, b)$ par

$$\begin{aligned} m_0(a, b) &= \lim_{p \rightarrow 0} m_p(a, b) = \sqrt{ab}, \\ m_\infty(a, b) &= \lim_{p \rightarrow \infty} m_p(a, b) = \max(a, b), \\ m_{-\infty}(a, b) &= \lim_{p \rightarrow -\infty} m_p(a, b) = \min(a, b). \end{aligned}$$

Les moyennes arithmétiques et harmoniques correspondent respectivement aux cas $p = 1$ et $p = -1$. On peut comparer ces différentes moyennes : pour a et b donnés, la fonction $p \mapsto m_p(a, b)$ est croissante.

On a de même une théorie des moyennes assez développée pour les matrices définies positives. Introduisons quelques notations et définitions. Soit $\mathbb{M}_n = \mathbb{M}_n(\mathbb{R})$ l'ensemble des matrices $n \times n$ à coefficients réels, \mathbb{S}_n le sous-ensemble des matrices symétriques et \mathbb{P}_n celui des matrices

1. Dans cet article, l'adjectif "positif" signifie "strictement positif" lorsqu'il qualifie un réel.

définies positives². On note que \mathbb{S}_n est un sous-espace vectoriel de \mathbb{M}_n et \mathbb{P}_n est un cône³ ouvert de \mathbb{S}_n . Soient deux matrices définies positives A et B : on dit que $A \geq B$ (*resp.* $A > B$) si $A - B$ est semi-définie positive (*resp.* définie positive). On note de façon équivalente $A \geq B$ et $B \leq A$. On remarquera qu'il se peut que l'on n'ait ni $A \geq B$ ni $B \geq A$. À toute matrice inversible⁴ P de taille $n \times n$, on associe l'application de *conjugaison* Γ_P de \mathbb{M}_n dans lui-même qui associe à une matrice A sa matrice *congruente* $\Gamma_P(A) = {}^tPAP$; quand P est une matrice orthogonale, on dit que A et tPAP sont *orthogonalement congruentes*, ou *orthogonalement semblables*. On note \mathbb{O}_n le groupe des matrices orthogonales de taille $n \times n$.

Nous pouvons maintenant préciser les conditions que doit satisfaire une moyenne sur l'ensemble des matrices définies positives. L'analogie avec les propriétés (i) - (v) du cas des nombres réels suggère que l'on demande à une moyenne M d'associer à deux matrices A et B définies positives, une matrice définie positive $M(A, B)$ telle que l'on ait les propriétés

- (I) $M(A, B) = M(B, A)$.
- (II) Si $A \leq B$, alors $A \leq M(A, B) \leq B$.
- (III) $\forall P \in GL_n : M({}^tPAP, {}^tPBP) = {}^tPM(A, B)P$.
- (IV) $M(A, B)$ est une fonction croissante de A et B , c'est-à-dire que si on a $A_1 \leq A_2$ et $B_1 \leq B_2$, alors $M(A_1, B_1) \leq M(A_2, B_2)$.
- (V) $M(A, B)$ est une fonction continue de A et B .

La condition de monotonie (IV) est source de difficulté dans la construction de moyennes sur l'ensemble des matrices, comme le montre l'exemple suivant.

$$\text{Pour } A = \begin{pmatrix} 3 & 1 \\ 1 & 1 \end{pmatrix} \text{ et } B = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}, \text{ on a } A \geq B \text{ mais } A^2 \not\geq B^2.$$

Une façon classique d'étendre aux matrices définies positives une fonction $f : \mathbb{R}_+^* \rightarrow \mathbb{R}$ est la suivante. Pour une matrice définie positive A , il existe une matrice orthogonale $P \in \mathbb{O}_n$ telle que ${}^tPAP = \text{Diag}(\lambda_1, \dots, \lambda_n)$, où les λ_i sont positifs ; on pose alors $\tilde{f}(A) = {}^tP \text{Diag}(f(\lambda_1), \dots, f(\lambda_n)) P$. On remarquera que la matrice $\tilde{f}(A)$ est symétrique, mais pas nécessairement positive. Une telle application \tilde{f} est dite *matriciellement monotone* si l'inégalité $A \geq B$ dans \mathbb{P}_n implique $\tilde{f}(A) \geq \tilde{f}(B)$, au sens où la matrice $\tilde{f}(A) - \tilde{f}(B)$ est semi-définie positive. On peut montrer que la fonction $x \mapsto x^p$ est monotone si et seulement si $0 \leq p \leq 1$, que l'application \ln est monotone, mais que \exp ne l'est pas. Dans la suite, on notera simplement f

2. On dit qu'une matrice $A \in \mathbb{S}_n$ est *semidéfinie positive* si elle est symétrique et si pour tout $X \in \mathbb{R}^n$ on a ${}^tXAX \geq 0$ et qu'elle est *définie positive* si elle est symétrique et si pour tout $X \in \mathbb{R}^n \setminus \{0\}$ on a ${}^tXAX > 0$.

3. C'est-à-dire que pour tout $\lambda > 0$ et toute matrice A définie positive, alors λA est définie positive.

4. On note GL_n l'ensemble des matrices inversibles $n \times n$

l'extension de \tilde{f} de f aux matrices définies positives.

Le lecteur vérifiera que les moyennes arithmétiques et harmoniques, qui associent à deux matrices définies positives A et B les matrices respectives $\frac{1}{2}(A+B)$ et $2(A^{-1}+B^{-1})^{-1}$ satisfont bien les conditions (I) - (V).

La notion de moyenne géométrique est plus délicate. Toute matrice définie positive A admet une unique racine carrée définie positive, notée $A^{1/2}$, mais le produit de deux telles racines carrées n'est en général même pas symétrique! En revanche, si les matrices définies positives A et B commutent, alors le produit $A^{1/2}B^{1/2}$ est symétrique et défini positif. On cherche donc d'autres expressions qui se réduisent à $A^{1/2}B^{1/2}$ quand les matrices définies positives commutent. Un choix plausible est la quantité

$$\exp\left(\frac{\ln A + \ln B}{2}\right), \quad \text{qui vaut également} \quad \lim_{p \rightarrow 0} \left(\frac{A^p + B^p}{2}\right)^{1/p}. \quad (1)$$

Malheureusement, ce candidat n'est pas non plus monotone : prendre deux matrices définies positives telles que $X \geq Y$ et $\exp X \not\geq \exp Y$, et prendre A et B telles que $X = \frac{1}{2}(\ln A + \ln B)$ et $Y = \frac{1}{2} \ln B$.

La condition (III) n'est pas non plus innocente ; en particulier, notre candidat malheureux ne la satisfait pas non plus.

L'analogie de \sqrt{ab} qui a les propriétés souhaitées est en fait l'expression

$$A\sharp B = A^{1/2} \left(A^{-1/2} B A^{-1/2} \right)^{1/2} A^{1/2}, \quad (2)$$

introduite par Pusz et Woronowicz en 1975. Il n'est pas évident *a priori* que cette expression est symétrique en A et B , mais nous le verrons bientôt. La monotonie en B provient de ce que la congruence conserve l'ordre (c'est-à-dire que $(B_1 \geq B_2) \Rightarrow ({}^t X B_1 X \geq {}^t X B_2 X)$) et de ce que la racine carrée est monotone.

La symétrie de \sharp est plus évidente si l'on utilise la caractérisation suivante, due à Ando. On a

$$A\sharp B = \max \left\{ X / X = {}^t X \text{ et } \begin{pmatrix} A & X \\ X & B \end{pmatrix} \geq 0 \right\}. \quad (3)$$

Parmi ses autres caractérisations, on peut définir $A\sharp B$ comme l'unique solution de l'équation de Riccati

$$X A^{-1} X = B. \quad (4)$$

On appellera la quantité $A\sharp B$ la *moyenne géométrique* de A et B . Elle satisfait aux conditions (I) - (V) requises. Pour vérifier la propriété (III), on pourra utiliser (3) et (4). On vérifiera qu'elle satisfait également la relation

$$\left(\frac{A^{-1} + B^{-1}}{2}\right)^{-1} \leq A\sharp B \leq \frac{A + B}{2}. \quad (5)$$

Deux matrices positives A et B peuvent être diagonalisées simultanément dans une base orthonormale (c'est-à-dire par une conjugaison orthogonale Γ_U avec $U \in \mathbb{O}_n$) si et seulement si elles commutent. En l'absence de commutation, on peut utiliser la conjugaison non unitaire définies en deux étapes de la façon suivante

$$(A, B) \xrightarrow{\Gamma_{A^{-1/2}}} (I, A^{-1/2}BA^{-1/2}) \xrightarrow{\Gamma_U} (I, D), \quad (6)$$

où U est une matrice orthogonale telle la matrice ${}^tP(A^{-1/2}BA^{-1/2})P$ est une matrice diagonale, notée D . Cela ôte quelque mystère à la construction (2). En fait, à toute moyenne m définie sur les nombres positifs, on peut associer l'application M qui à tout couple de matrices définies positives (A, B) associe $M(A, B) = \Gamma_{A^{-1/2}}m(I, D)$; la fonction M est monotone dès que la fonction $f(x) = m(1, x)$ est matriciellement monotone. La formule (2) correspond au cas où $m(a, b) = (ab)^{1/2}$.

L'argument que nous avons introduit pour obtenir la symétrie de la moyenne géométrique n'est pas indispensable. Soit en effet m une moyenne sur les réels positifs, $f(x) = m(1, x)$ et

$$M(A, B) = A^{1/2}f\left(A^{-1/2}BA^{-1/2}\right)A^{1/2}. \quad (7)$$

Contrairement à son apparence, cette expression est symétrique en A et B . Montrons qu'on a effectivement

$$f\left(A^{-1/2}BA^{-1/2}\right) = A^{-1/2}B^{1/2}f\left(B^{-1/2}AB^{-1/2}\right)B^{1/2}A^{-1/2}.$$

En utilisant la décomposition $A^{-1/2}B^{1/2} = QU$, où Q est définie positive et U est orthogonale, il suffit de montrer que l'on a

$$f(Q^2) = QUf({}^tUQ^{-2}U){}^tUQ = Qf(Q^2)Q.$$

Cela équivaut à dire que pour toute valeur propre λ de Q , on a

$$m(1, \lambda^2) = \lambda m(1, \lambda^{-2}) \lambda,$$

ce qui découle des relations (i) et (iii) de la moyenne m . Un argument similaire montre que M satisfait la relation (III).

Un corollaire simple de cette construction est que des inégalités du type (5) se conservent lorsque l'on passe de moyennes sur les réels positifs à des moyennes sur les matrices définies positives.

Que se passe-t-il lorsqu'on considère trois matrices définies positives ? Les moyennes arithmétiques et harmoniques définies respectivement par $\frac{1}{3}(A + B + C)$ et $(\frac{1}{3}(A^{-1} + B^{-1} + C^{-1}))^{-1}$ ne posent pas de problème particulier. En revanche, la moyenne géométrique soulève de nouveau des questions intéressantes.

On aimerait disposer d'une moyenne géométrique $G(A, B, C)$ qui se réduise à $A^{1/3}B^{1/3}C^{1/3}$ quand les matrices A, B et C commutent deux à deux. En outre, on voudrait avoir les propriétés suivantes

- (a) $G(A, B, C) = G(\pi(A, B, C))$ pour toute permutation π .
- (b) $G({}^tPAP, {}^tPBP, {}^tPCP) = {}^tPG(A, B, C)P$ pour tout $P \in GL_n$.
- (c) $G(A, B, C)$ est une fonction croissante de A, B et C .
- (d) $G(A, B, C)$ est une fonction continue de A, B et C .

Aucun des moyens présentés ci-dessus dans le cas de deux matrices ne s'étend réellement au cas de trois matrices : les expressions (2), (3) et (4) n'ont pas de généralisation utile. La définition d'une moyenne géométrique pour trois matrices définies positives a été un problème titillant pendant de nombreuses années. Nous présentons maintenant quelques progrès obtenus récemment sur cette question.

Une géométrie ne peut pas être plus vraie qu'une autre ; elle peut seulement être plus commode. (Henri Poincaré)

On munit l'espace \mathbb{M}_n du produit scalaire naturel $\langle A, B \rangle = \text{tr}({}^tAB)$ et on note $\|\cdot\|_2$ la norme qui lui est associée. Soient A et B deux éléments de \mathbb{P}_n . Un chemin γ joignant A et B est une application continûment dérivable γ de $[0, 1]$ dans $\mathbb{P}_n \subset \mathbb{M}_n$ telle que $\gamma(0) = A$ et $\gamma(1) = B$. La longueur d'un chemin est la quantité

$$L(\gamma) = \int_0^1 \|\gamma^{-1/2}(t)\gamma'(t)\gamma^{-1/2}(t)\|_2 dt. \quad (8)$$

Parmi tous les chemins joignant A et B il en existe un unique dont la longueur est minimale ; on l'appelle *géodésique* joignant A et B ; on le note $[A, B]$ et on appelle *distance* de A à B la quantité $\delta_2(A, B) = L([A, B])$; il s'agit effectivement d'une distance, car on a pour tout A, B et C dans \mathbb{P}_n :

- $\delta_2(A, B) = \delta_2(B, A)$,
- $\delta_2(A, B) = 0 \Leftrightarrow A = B$,
- $\delta_2(A, B) \leq \delta_2(A, C) + \delta_2(C, B)$.

Puisque deux matrices semblables ont même trace, pour tout $P \in GL_n$, on a $\delta_2(\Gamma_P(A), \Gamma_P(B)) = \delta_2(A, B)$ pour tout A et B dans \mathbb{P}_n . On énonce ce fait en disant qu'une conjugaison par une matrice inversible est une *isométrie* pour la distance δ_2 sur \mathbb{P}_n .

L'application qui associe à une matrice symétrique A son exponentielle $\exp(A)$ est une bijection de \mathbb{S}_n sur \mathbb{P}_n , et on a pour toutes matrices symétriques S et T la relation

$$\|S - T\|_2 \leq \delta_2(\exp(S), \exp(T)). \quad (9)$$

Nous donnons en appendice quelques éléments de démonstration de cette relation.

Lorsque les matrices S et T commutent, la géodésique joignant $A = \exp(S)$ et $B = \exp(T)$ est le chemin

$$\gamma(t) = \exp((1-t)S + tT) = \exp((1-t)S)\exp(tT) = A^{1-t}B^t, \quad \text{pour } 0 \leq t \leq 1.$$

Pour $t \in [0, 1]$, on a en outre $\delta_2(A, \gamma(t)) = t\delta_2(A, B)$.

Dans le cas où les matrices A et B ne commutent pas, on peut se ramener au cas commutatif en utilisant le fait que la conjugaison $\Gamma_{A^{-1/2}}$ est une isométrie. La géodésique $[I, A^{-1/2}BA^{-1/2}]$ joignant la matrice I à la matrice $A^{-1/2}BA^{-1/2}$ est paramétrisée par $\gamma_0(t) = (A^{-1/2}BA^{-1/2})^t$ et donc la géodésique $[A, B] = [\Gamma_{A^{-1/2}}(I), \Gamma_{A^{-1/2}}(A^{-1/2}BA^{-1/2})]$ est paramétrisée par

$$\gamma(t) = A^{1/2}(A^{-1/2}BA^{-1/2})^t A^{1/2}, \quad 0 \leq t \leq 1. \quad (10)$$

Cela montre que la moyenne géométrique $A\sharp B$ définie par la formule (2) n'est autre que le milieu de la géodésique joignant A et B dans \mathbb{P}_n . Ainsi, tandis que (2), (3) et (4) ont pu sembler être des variantes non commutatives de la moyenne géométrique, des considérations géométriques naturelles conduisent à la définition (2). On notera également que pour tout $t \in [0, 1]$, on peut voir la quantité $\gamma(t)$ définie en (10) comme une moyenne de A et B , au sens de la formule (7), associée à la fonction $f(x) = x^t$. Mais ces moyennes ne sont pas symétriques et la condition (I) est en défaut, à l'exception du cas où $t = 1/2$.

Cette discussion conduit également à une formule explicite pour la distance δ_2 . On a en effet $\delta_2(A, B) = \delta_2(I, A^{-1/2}BA^{-1/2}) = \|\ln(I) - \ln(A^{-1/2}BA^{-1/2})\|_2 = \|\ln(A^{-1/2}BA^{-1/2})\|_2$. Puisque les matrices $A^{-1/2}BA^{-1/2}$ et $A^{-1}B$ ont mêmes valeurs propres, on a

$$\delta_2(A, B) = \left(\sum_{i=1}^n (\ln(\lambda_i(A^{-1}B)))^2 \right)^{1/2}. \quad (11)$$

L'inégalité (9) traduit une caractéristique essentielle de \mathbb{P}_n , à savoir sa *courbure positive*. Pour donner une idée de ce que cela signifie, considérons dans \mathbb{S}_n un triangle de sommets O, S, T . Par l'application exponentielle, il est envoyé sur un "triangle" d'extrémités $I, \exp(S), \exp(T)$ dans \mathbb{P}_n . Les longueurs des côtés $[O, S]$ et $[O, T]$, mesurés par la norme $\|\cdot\|_2$, sont égales à celles de leurs images, les géodésiques $[I, \exp(S)]$ et $[I, \exp(T)]$, mesurées par δ_2 . Par la relation (9), la longueur du troisième côté $[\exp(S), \exp(T)]$ est supérieure (ou égale) à $\|S - T\|_2$. Le cas d'un triangle général de sommets A, B, C dans \mathbb{P}_n s'y ramène par la conjugaison isométrique $\Gamma_{A^{-1/2}}$ qui envoie A sur I . Une façon d'énoncer ce phénomène est de dire que deux géodésiques émanant d'un point de \mathbb{P}_n se séparent plus rapidement que leurs images réciproques (pour la fonction exponentielle) dans \mathbb{S}_n . Ce phénomène est illustré dans la Figure 1, qui illustre également la situation opposée de contraction qu'implique la considération de l'exponentielle imaginaire et des matrices anti-symétriques, non détaillée ici.

Revenons à la moyenne géométrique dans \mathbb{P}_n . Pour trois matrices A, B, C de \mathbb{P}_n , on a l'inégalité

$$\delta_2(A\#B, A\#C) \leq \frac{1}{2}\delta_2(B, C). \quad (12)$$

Cette inégalité nous dit que dans tout triangle géodésique de \mathbb{P}_n de sommets A, B, C , la longueur de la géodésique joignant les milieux de deux côtés est au plus égale à la moitié de la longueur du troisième côté; dans un espace euclidien, la relation (12) est une égalité. La Figure 2 illustre ce fait.

Nous avons vu que la moyenne géométrique $A\#B$ est le milieu de la géodésique $[A, B]$. Cela suggère de définir la moyenne géométrique des trois matrices A, B, C de \mathbb{P}_n comme "centre de gravité" du triangle géodésique $\Delta(A, B, C)$.

Dans un espace vectoriel euclidien, le centre de gravité \mathbf{g} du triangle de sommets $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3$ est défini par $\mathbf{g} = \frac{1}{3}(\mathbf{x}_1 + \mathbf{x}_2 + \mathbf{x}_3)$. Cette définition ne peut pas s'étendre telle quelle au cône \mathbb{P}_n , mais le centre de gravité peut être caractérisé par d'autres propriétés susceptibles de s'étendre à \mathbb{P}_n . En voici trois :

(M1) \mathbf{g} est le point d'intersection des trois médianes du triangle $\Delta(A, B, C)$;

(M2) \mathbf{g} est l'unique point où la fonction

$$\mathbf{x} \mapsto \|\mathbf{x} - \mathbf{x}_1\|^2 + \|\mathbf{x} - \mathbf{x}_2\|^2 + \|\mathbf{x} - \mathbf{x}_3\|^2$$

atteint son minimum ;

(M3) \mathbf{g} est l'unique point d'intersection de la suites de triangles emboîtés $(\Delta_n)_n$ et où $\Delta_1 = \Delta$ et Δ_{j+1} s'obtient en joignant les milieux des côtés de Δ_j .

Pour définir la moyenne géométrique de A, B, C dans \mathbb{P}_n , on peut tenter d'utiliser l'une ou l'autre de ces caractérisations, dont la formulation peut permettre l'extension à \mathbb{P}_n . On va noter ici la différence fondamentale entre la géométrie (dite *hyperbolique*) de \mathbb{P}_n et celle d'un espace vectoriel euclidien.

En fait, la caractérisation (M1) ne peut pas s'étendre comme telle : en général, les médianes d'un triangle (comprises comme les géodésiques joignant un côté et le milieu du côté opposé - ce qui a bien un sens -) peuvent très bien ne pas s'intersecter, même deux à deux. La Figure 3 illustre ce fait. La Figure 4 illustre le fait que si l'on voit le triangle de \mathbb{P}_2 de sommets A, B, C comme l'enveloppe convexe de ses sommets (on comprend ici les segments comme des segments géodésiques), ce n'est pas un objet bidimensionnel.

La caractérisation (M2) s'étend à \mathbb{P}_n . Élie Cartan a montré qu'étant donnés trois points A, B, C de \mathbb{P}_n , il existe un unique point, notons le $G_2 = G_2(A, B, C)$ de \mathbb{P}_n , où la fonction

$$X \mapsto \delta_2^2(A, X) + \delta_2^2(B, X) + \delta_2^2(C, X)$$

atteint son minimum. Cet abord a été récemment étudié, mais on ne sait toujours pas si G_2 est une fonction monotone. Plus précisément, soit A, A', B, C des matrices définies positives avec $A' \geq A$; a-t-on $G_2(A', B, C) \geq G_2(A, B, C)$? L'expérimentation numérique suggère une réponse positive.

La caractérisation itérative (M3) se prête bien à une généralisation à \mathbb{P}_n : on construit par récurrence $\Delta(A_{j+1}, B_{j+1}, C_{j+1}) = \Delta(B_j \# C_j, C_j \# A_j, A_j \# B_j)$, avec $A_1 = A, B_1 = B, C_1 = C$. L'inégalité (12) assure la convergence de chaque suite $(A_j)_j, (B_j)_j, (C_j)_j$ vers une même limite que l'on note $G_3(A, B, C)$. Il est très aisé de vérifier que G_3 satisfait les conditions requises (a), (b) et (d); la propriété (c) provient du fait que pour toute matrice inversible P , la conjugaison Γ_P est une isométrie.

APPENDICE

Éléments de démonstration de l'inégalité (9)

Soient S et T deux matrices symétriques. On considère $D_T \exp(S)$, la dérivée de la fonction exponentielle au point S selon la direction T , définie par

$$D_T \exp(S) = \lim_{t \rightarrow 0} \frac{\exp(S + tT) - \exp(S)}{t}. \quad (13)$$

Si on se place dans une base orthonormale dans laquelle $S = \text{Diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$, on a

$$D_T \exp(S) = \left(\frac{e^{\lambda_i} - e^{\lambda_j}}{\lambda_i - \lambda_j} t_{ij} \right)_{1 \leq i, j \leq n}. \quad (14)$$

Il s'ensuit que l'on a

$$\exp(-S/2) D_T \exp(S) \exp(-S/2) = \left(\frac{\sinh(\lambda_i - \lambda_j)/2}{(\lambda_i - \lambda_j)/2} t_{ij} \right)_{1 \leq i, j \leq n}. \quad (15)$$

Puisqu'on a pour tout x : $\sinh(x) \geq x$, on a l'inégalité

$$\|S\|_2 \leq \|\exp(-S/2) D_T \exp(S) \exp(-S/2)\|_2. \quad (16)$$

L'inégalité (9) se déduit alors de la formule (8) et de la relation infinitésimale (16).

ANNEXE : FIGURES

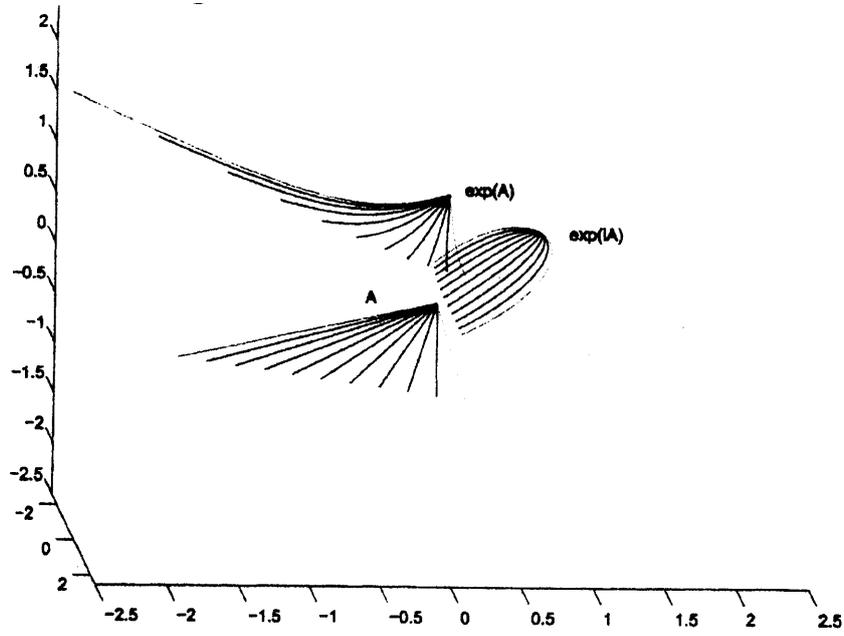


Figure 1 : Écartement et contraction de géodésiques

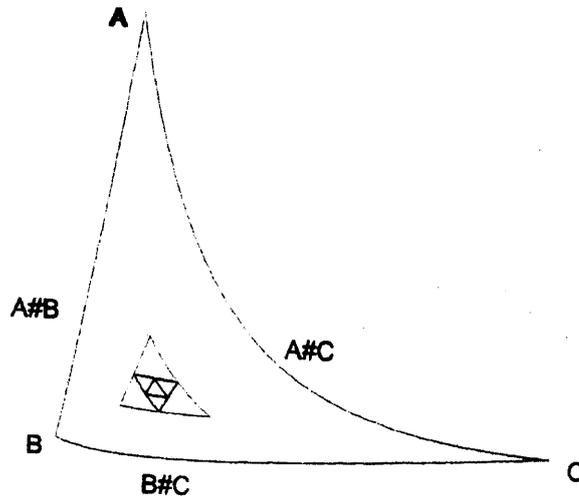


Figure 2 : Distance de deux moyennes

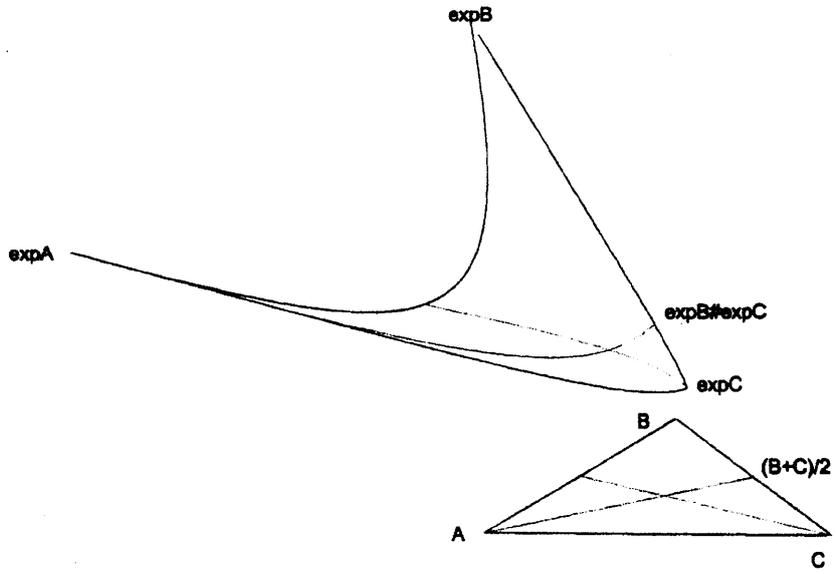


Figure 3 : Les médianes ne se rencontrent pas nécessairement

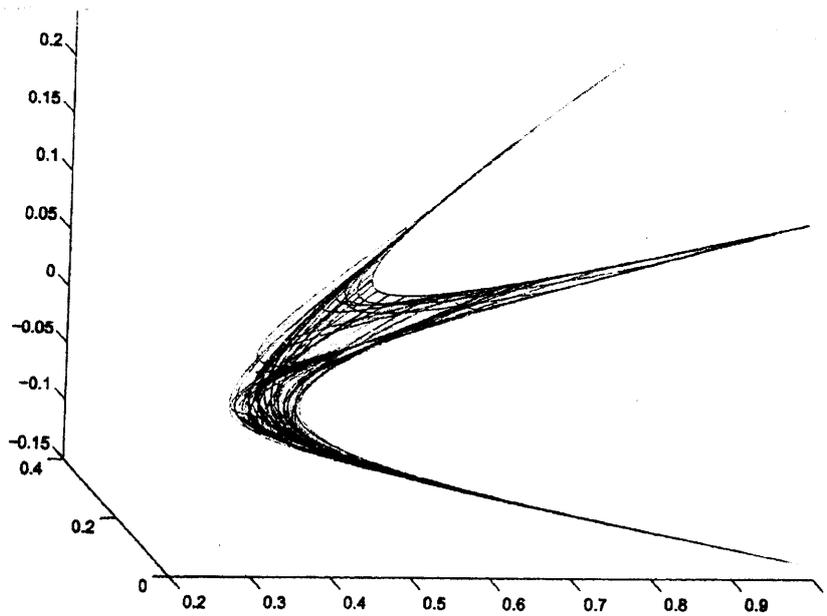


Figure 4 : L'enveloppe convexe de trois points n'est pas bi-dimensionnelle

ADS-M-3

Une conjecture sur les zéros d'un polynôme et de ses dérivées

Travail demandé :

Il vous est demandé d'étudier puis de présenter le(s) texte(s) joint(s) à travers un exposé de synthèse d'une durée comprise entre 15 et 20 minutes.

Si l'étude de la totalité du dossier et la préparation d'un exposé cohérent dans la durée impartie ne vous paraît pas possible, vous pouvez décider de vous limiter à une partie du dossier.

Remarques générales :

1. Les textes proposés, quelle que soit leur origine, peuvent présenter des défauts (coquilles typographiques, négligences ou sous-entendus de l'auteur, voire erreurs...) qui, sauf exception, n'ont pas été corrigés.

2. Les textes proposés peuvent contenir des exercices qu'il n'est pas demandé de résoudre. Néanmoins, vous pouvez vous aider des énoncés de ces exercices pour enrichir votre exposé.

3. Vous pouvez annoter les documents qui vous sont fournis. Vos annotations ne seront pas regardées par l'examineur.

Remarque particulière :

Une partie A , éventuellement vide, de \mathbb{C} est dite convexe si pour tout couple d'éléments a, b de A , le segment $\{ta + (1-t)b / t \in [0, 1]\}$ est inclus dans A . Une intersection de parties convexes est convexe. Pour une partie B de \mathbb{C} , l'intersection de toutes les parties convexes qui contiennent B s'appelle l'enveloppe convexe de B .

UNE CONJECTURE SUR LES ZÉROS D'UN POLYNÔME ET DE SES DÉRIVÉES

1. LA CONJECTURE

En 2001, le mathématicien Eduardo Casas-Alvero a formulé la conjecture suivante, d'autant plus remarquable qu'elle n'implique que des connaissances du niveau des études secondaires, et n'est toujours pas démontrée, douze ans après avoir été énoncée.

Conjecture 1. *Soit $f \in \mathbb{C}[x]$ un polynôme unitaire de degré $n > 0$. On suppose que pour tout entier $k \in [1, n - 1]$, le polynôme f et sa dérivée k -ième $f^{(k)}$ ont un zéro commun dans \mathbb{C} . Alors, il existe un nombre complexe a tel que $f(x) = (x - a)^n$.*

Dans la suite, on appellera polynôme de Casas-Alvero, ou polynôme C-A, un polynôme satisfaisant la condition énoncée. On remarquera également que cette condition est équivalente au fait que pour tout k , le PGCD des polynômes f et $f^{(k)}$ est non constant. On notera $f' = f^{(1)}$ et $f'' = f^{(2)}$.

La conjecture ne peut pas s'étendre à n'importe quel corps commutatif. En effet, pour $\mathbb{K} = \mathbb{Z}/7\mathbb{Z}$, le polynôme $f(x) = x^4 + x^2 + 2x$ est un contre-exemple à la conjecture.

La conjecture n'est pas résolue ; même le cas où l'on suppose que f est scindé sur \mathbb{R} n'est pas résolu : le fait que les zéros de $f^{(i)}$ et ceux de $f^{(i+1)}$ soient entrelacés ne permet pas à lui seul de démontrer la conjecture pour un degré supérieur ou égal à 5 (cf. Fig. 1).

Dans cette note, on passe en revue quelques travaux menés sur la conjecture C-A. En particulier, dans le paragraphe 6, on présentera l'abord algébrique de Graf von Bothmer, Labs, Schicho et van de Woestijne qui démontrent la conjecture pour une infinité de valeurs de n , en nous limitant aux cas où n est le double d'un nombre premier. Leur méthode s'applique plus généralement aux entiers de la forme $n'p^e$, où n' est un entier petit (au plus 4) et p un nombre premier. Par essence, il est peu probable que leur méthode permette de résoudre complètement la conjecture C-A ; pour cette raison, nous présentons également d'autres abords plus analytiques. Dans le deuxième paragraphe, nous indiquons quelques symétries de la conjecture et une voie d'attaque via l'inversion.

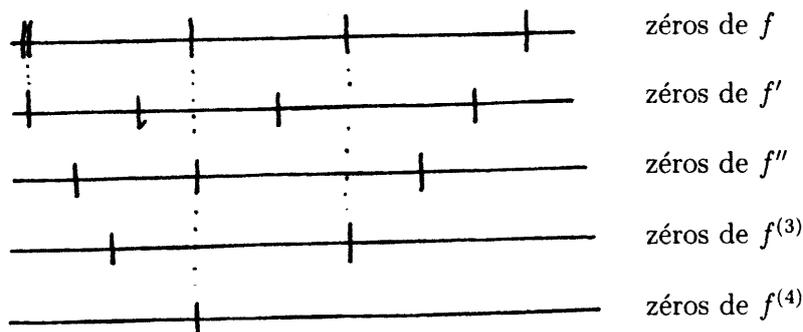


Figure 1 Zéros potentiels de f de degré 5

Dans le troisième paragraphe, nous présentons le théorème de Gauss-Lucas sur la localisation des zéros d'un polynôme et de ses dérivées et nous en déduisons la conjecture C-A en petit degré. Dans le paragraphe suivant, nous montrons qu'on ne peut trop relâcher la condition de la conjecture en présentant des "contre-exemples" si on omet une des dérivées. Dans le paragraphe 5, nous établissons que l'ensemble des zéros commun à un polynôme C-A f et à ses dérivées ne peut être exactement 2.

2. LES SYMÉTRIES

L'observation élémentaire suivante décrit les symétries de l'ensemble des polynômes C-A : leurs zéros peuvent être simultanément transformés par une similitude.

Lemme 1. *Si f est un polynôme C-A de degré n , alors, pour tout $a \in \mathbb{C}^*$ et tout $b \in \mathbb{C}$, le polynôme $a^{-n}f(ax + b)$ est C-A.*

La conjecture C-A implique que l'ensemble des polynômes C-A f de degré n donné pour lesquels $f(0) \neq 0$ est stable pour l'inversion $f(x) \mapsto x^n f(1/x)/f(0)$, et donc par toute transformation, dite transformation de Möbius, qui est produit d'un nombre fini de similitudes et d'inversions. La proposition suivante montre que cette condition nécessaire est également suffisante.

Proposition 1. *Si l'ensemble des polynômes C-A de degré n qui ne s'annulent pas à l'origine est stable par inversion, alors la conjecture C-A est vraie en degré n .*

Esquisse de démonstration. Supposons par l'absurde que f est un polynôme C-A de degré n avec au moins deux zéros distincts. Quitte à appliquer une transformation affine, on peut supposer que 1 est un zéro de f ; soit d sa multiplicité : on a $1 \leq d < n$. On peut

alors trouver une famille de transformations de Möbius, $\lambda(t)$, dépendant d'un paramètre $t \in \mathbb{C}^*$ telles que

$$\forall t : \lambda(t)(1) = 1, \text{ et } \forall z \in \mathbb{C} \setminus \{1\} : \lim_{t \rightarrow 0} \lambda(t)(z) = -1.$$

Alors, quand t tend vers 0, $\lambda(t)(f(x))$ tend vers $(x-1)^d(x+1)^{n-d}$. Mais alors ce polynôme devrait également être C-A, ce qu'il n'est pas. \square

D'autres transformations peuvent être imaginées. Par exemple, si on sait montrer que l'ensemble des polynômes C-A est stable par la transformation $\prod(x - \alpha_i) \mapsto \prod(x - \alpha_i^2)$, alors on peut en déduire la conjecture C-A. On ne sait malheureusement pas comment exploiter ces observations.

3. LE THÉORÈME DE GAUSS-LUCAS

On rappelle le théorème classique de Gauss-Lucas

Théorème 1. *Pour tout polynôme non constant $f \in \mathbb{C}[x]$, les zéros de sa dérivée f' sont situés dans l'enveloppe convexe des zéros de f . Plus précisément, si α est un zéro de f' , alors, ou bien c'est un zéro de f , ou bien il est situé dans l'intérieur relatif de l'enveloppe convexe des zéros de f .*

Ici, l'intérieur relatif d'un point est ce point, celui d'un intervalle d'extrémité a et b est l'intervalle ouvert d'extrémités a et b , et celui d'un polygone est ce polygone privé de ses bords. Le théorème de Gauss-Lucas permet de montrer qu'un polynôme C-A a au moins un de ses zéros dans l'intérieur relatif de l'enveloppe convexe, notée C , de l'ensemble de ses zéros. Soit en effet z_1, \dots, z_k les zéros de f , distincts deux à deux, qui ne sont pas dans l'intérieur relatif de C , et soit m_1, \dots, m_k leurs multiplicités respectives ; on pose $m = \max_i m_i$. Si $m = \deg(f)$, alors f est la puissance m -ième d'un polynôme linéaire et l'assertion est manifestement vraie. Sinon, on a $m < n$ et considère le polynôme $f^{(m)}$: tous ses zéros sont dans l'intérieur relatif de C et donc l'un des zéros de f y est.

Cette remarque permet de démontrer la conjecture C-A pour les degrés $n \leq 4$. Supposons par exemple que f est un polynôme C-A de degré 4 et soient z_1, z_2, z_3, z_4 ses zéros (non nécessairement distincts). Puisque f et f' ont un zéro en commun, f a un zéro double et on peut, sans perte de généralité, supposer que $z_1 = z_2$. D'après ce qui précède, ou bien les quatre zéros sont confondus (la conjecture est vérifiée) ou bien les trois zéros z_1, z_3 , et z_4 sont deux à deux distincts et alignés. On peut alors effectuer une similitude qui envoie ces trois points alignés sur la droite réelle. Il y a deux possibilités :

ou bien z_1 est entre z_3 et z_4 ; alors f'' a un zéro dans l'intervalle ouvert d'extrémités z_1 et z_3 et un zéro dans l'intervalle ouvert d'extrémités z_1 et z_4 , et alors f'' n'a aucun zéro

en commun avec f , ce qui est contradictoire,

ou bien z_3 est entre z_1 et z_4 ; alors, f'' doit avoir un zéro en z_3 et un zéro dans l'intervalle ouvert d'extrémités z_1 et z_3 ; dans ce cas, $f^{(3)}$ n'a pas de zéro en commun avec f , ce qui est contradictoire.

De tels arguments sont insuffisants pour traiter le cas de degrés supérieurs à 4 : en effet, des arguments de nature purement topologique ne permettent pas d'exclure une configuration comme celle de la figure 1.

4. "CONTRE-EXEMPLES APPROCHÉS" AVEC UNIQUEMENT DES ZÉROS RÉELS

Nous montrons ici que la conclusion de la conjecture n'est pas valide si on suppose seulement que f a un zéro commun avec chacune de ses dérivées, sauf une ; plus précisément on a le résultat suivant

Théorème 2. *Pour tout $n > 1$ et $\ell \in \{1, \dots, n-1\}$, il existe un polynôme $f \in \mathbb{R}[x]$ avec $f(0) = f(1) = 0$ dont tous les zéros sont réels et inclus dans l'intervalle $[0, 1]$ et qui possède un zéro commun avec chacune de ses dérivées $f^{(k)}$, pour $k \in \{1, \dots, n-1\} \setminus \{\ell\}$.*

Esquisse de démonstration. Si f est un polynôme réel qui n'a que des zéros réels, il en va de même pour toutes ses dérivées. Pour $1 \leq k \leq n$, on note $\alpha_{k,1} \leq \alpha_{k,2} \leq \dots \leq \alpha_{k,n-k}$ les zéros de $f^{(k)}$. Nous allons démontrer un résultat un peu plus précis. Pour $k \in E(\ell) = \{1, \dots, n-1\} \setminus \{\ell\}$, choisissons m_k dans $[1, n-k]$. Pour un $(n-2)$ -uplet $\beta = (\beta_k)_{k \in E(\ell)}$, on note $g(\beta)$ le polynôme $x(x-1) \prod_{k \in E(\ell)} (x - \beta_k)$. L'application $\Phi : [0, 1]^{n-2} \rightarrow [0, 1]^{n-2}$ qui à β associe $(\alpha_{k,m_k})_{k \in E(\ell)}(g(\beta))$ est continue ; d'après le théorème de Brouwer (résultat admis), elle admet un point fixe. On vérifie facilement que le polynôme $g(\beta)$ associé à un tel point fixe a les propriétés requises. \square

Un exemple. Pour $n = 5$ et les paires $(1, 2)$, $(2, 3)$ et $(4, 4)$, on trouve $\beta_1, \beta_4 \approx 0,629099$ et $\beta_2 \approx 0,887298$. Le graphe des dérivées de f est donné dans la Figure 2.

5. DEUX ZÉROS NE CONVIENNENT PAS

On démontre dans ce paragraphe le résultat suivant.

Proposition 2. *Soit f un polynôme C-A de degré n et soit a_i un zéro commun à f et $f^{(i)}$. On a $\text{card}\{a_i / 1 \leq i \leq n-1\} \neq 2$.*

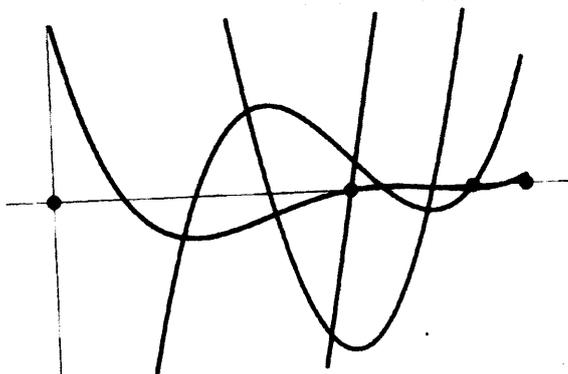


Figure 2 Contre-exemple approché en degré 5

On remarquera que cet énoncé est plus fort que le fait que f ne peut pas être de la forme $(x-a)^k(x-b)^{n-k}$ avec $0 < k < n$, établi dans la démonstration de la proposition 1.

Démonstration. On raisonne par l'absurde, en supposant que le cardinal des a_i est 2. Quitte à effectuer une transformation affine, on peut supposer que $\{a_i / 1 \leq i \leq n-1\} = \{0, 1\}$. On va montrer par induction descendante sur k que $f^{(k)}$ est un polynôme totalement positif ou totalement négatif sur l'intervalle ouvert $]0, 1[$. C'est vrai pour $k = n$: on a $f^{(n)} = n!$. Supposons notre hypothèse satisfaite au rang k . Alors, il existe un nombre complexe c_{k-1} tel que pour tout réel x on a $f^{(k-1)}(x) = \int_0^x f^{(k)}(t)dt + c_{k-1}$; puisque $f^{(k)}(t)$ est réel pour tout réel t et que $f^{(k-1)}$ s'annule en 0 ou en 1, le nombre c_{k-1} est réel et donc le polynôme $f^{(k-1)}$ est à coefficients réels. Il est strictement monotone sur l'intervalle $[0, 1]$ et il s'annule en une des extrémités de cet intervalle. Il est donc strictement positif ou strictement négatif sur $]0, 1[$. Cette propriété est donc également vraie pour le polynôme $f = f^{(0)}$; mais alors le polynôme f ne peut pas s'annuler en 0 et en 1 ; cette contradiction achève la démonstration de la proposition. \square

6. LA RÉDUCTION MODULO p

Le théorème suivant est actuellement le résultat positif principal sur la conjecture de Casas-Alvero. Moyennant deux résultats admis concernant les valuations p -adiques sur \mathbb{C} , nous en donnerons la démonstration, en nous restreignant au cas où n est le double d'un nombre premier, cas techniquement plus simple mais qui illustre les ressorts de la démonstration générale.

Théorème 3. *On considère un entier $n' \in \{1, 2, 3, 4\}$, un nombre premier p supérieur à n' et différent de 7 si $n' = 4$, un entier $e \geq 0$ et on pose $n = n'p^e$. Alors la conjecture de Casas-Alvero est vraie pour le degré n .*

Valuation. Une valuation sur un corps commutatif \mathbb{K} est une application $v : \mathbb{K} \rightarrow \mathbb{R} \cup \{\infty\}$ telle que l'on a

- (1) $v^{-1}(\{\infty\}) = \{0\}$,
- (2) $v(ab) = v(a) + v(b)$,
- (3) $v(a + b) \geq \min(v(a), v(b))$,

pour tout couple (a, b) d'éléments de \mathbb{K} .

Valuation p -adique sur \mathbb{C} . Pour un nombre premier p , il existe une unique valuation v sur \mathbb{Q} qui est telle que pour tout entier positif m on a $v(m) = \max\{e/p^e \text{ divise } m\}$; on l'appelle la valuation p -adique sur \mathbb{Q} et on la note v_p . Nous admettons d'une part qu'il existe une valuation v sur \mathbb{C} qui coïncide avec v_p sur \mathbb{Q} (on continuera de la noter v_p). On note $V^+ = \{z \in \mathbb{C} / v_p(z) \geq 0\}$. Nous admettons d'autre part qu'il existe un corps commutatif \mathbb{K} et une application *surjective* $\sigma : V^+ \rightarrow \mathbb{K}$ telle que

- (1) $\sigma(1_{\mathbb{C}}) = 1_{\mathbb{K}}$,
- (2) $\forall a, b \in V^+ : \sigma(a + b) = \sigma(a) + \sigma(b)$,
- (3) $\forall a, b \in V^+ : \sigma(ab) = \sigma(a)\sigma(b)$,
- (4) $\forall a \in V^+ : v_p(a) > 0 \Rightarrow \sigma(a) = 0$.

Notation. Pour $a \in V^+$, on notera $\tilde{a} = \sigma(a)$. Bien que $\tilde{p} = \tilde{0}$, nous suivrons l'usage et noterons simplement a l'élément \tilde{a} dans le cas où $a \in \mathbb{Z}$. En particulier, 0 et 1 dénotent respectivement le neutre de \mathbb{K} pour l'addition et le neutre de \mathbb{K} pour la multiplication.

Coefficients du binôme. L'image par σ du coefficient binomial $\binom{2p}{k}$ est nulle si et seulement si $k \notin \{0, p, 2p\}$. En particulier, on a dans $\mathbb{K}[x]$ la relation $(f + g)^p = f^p + g^p$, dite le rêve du bizuth.

Dérivées de Hasse. Pour un polynôme $f = \alpha_n x^n + \alpha_{n-1} x^{n-1} + \dots + \alpha_1 x + \alpha_0$ de $\mathbb{C}[x]$ ou de $\mathbb{K}[x]$, et un entier $i \leq n$, on définit sa i -ème dérivée de Hasse par

$$H_i(f) = \binom{n}{i} \alpha_n x^{n-i} + \binom{n-1}{i} \alpha_{n-1} x^{n-i-1} + \dots + \binom{i}{i} \alpha_i.$$

On notera que pour un polynôme $f \in \mathbb{C}[x]$, on a $H_i(f) = f^{(i)}/i!$, mais que pour $f \in \mathbb{K}[x]$, on peut avoir $H_i(f) \neq 0$, alors que $f^{(i)} = 0$, comme le montre l'exemple du polynôme $f = x^{2p}$ pour $i = p$.

Condition C-A dans $\mathbb{K}[x]$. Dans le cas de $\mathbb{C}[x]$, un polynôme est C-A s'il est unitaire de degré positif et s'il a au moins un zéro commun avec chacune de ses dérivées de Hasse. Dans la suite, on dira qu'un polynôme de $\mathbb{K}[x]$ est C-A s'il est unitaire de degré positif et s'il a au moins un zéro commun avec chacune de ses dérivées de Hasse.

Démonstration du théorème 3. Comme indiqué dans l'introduction de ce paragraphe, nous nous limitons au cas où $n' = 2$ et $e = 1$, c'est-à-dire au cas où $n = 2p$, avec p premier impair. On raisonne par l'absurde en supposant que $f \in \mathbb{C}[x]$ est un polynôme C-A avec au moins deux zéros distincts. En effectuant une translation, on peut supposer que l'un de ses zéros est 0, c'est-à-dire que $f(x) = x^h(x - x_0) \cdots (x - x_{2p-h-1})$, où $1 \leq h < n$ et les x_i , non nécessairement distincts deux à deux, sont tous non nuls ; on peut supposer que parmi les x_i , le zéro x_0 est de valuation minimale. Une homothétie de rapport $1/x_0$ permet de remplacer x_0 par 1 et tous les autres x_i par des complexes de valuation positive. En bref, on s'est ramené au cas où

$$\begin{aligned} f(x) &= x^h(x-1)(x-x_1) \cdots (x-x_{2p-h-1}) \\ &= x^{2p} + c_{2p-1}x^{2p-1} + \cdots + c_h x^h, \\ &\text{avec } \forall i \in [1, 2p-h-1] : v_p(x_i) \geq 0. \end{aligned}$$

Puisque tous les x_i ont une valuation positive, il en va de même des c_i et on peut considérer le polynôme de $\mathbb{K}[x]$

$$\begin{aligned} \tilde{f}(x) &= x^h(x-1)(x-\tilde{x}_1) \cdots (x-\widetilde{x_{2p-h-1}}) \\ &= x^{2p} + \widetilde{c_{2p-1}}x^{2p-1} + \cdots + \tilde{c}_h x^h. \end{aligned}$$

Si le zéro 0, 1 ou x_i est commun à f et $H_i(f)$, alors le zéro 0, 1 ou \tilde{x}_i est commun à \tilde{f} et $H_i(\tilde{f})$. Il s'ensuit que le polynôme \tilde{f} est C-A.

Nous montrons maintenant que \tilde{c}_i est nul si p ne divise pas i . Supposons d'abord qu'il existe $k \in [p+1, 2p-1]$ tel que $\tilde{c}_k \neq 0$, et choisissons un tel k maximal. On a

$$H_k(\tilde{f}) = \binom{2p}{k} x^{2p-k} + \tilde{c}_k = \tilde{c}_k.$$

Donc $H_k(\tilde{f})$ est un polynôme constant. Puisqu'il a un zéro commun avec \tilde{f} , il est nul et donc $\tilde{c}_k = 0$.

Si $h \geq p$, on peut écrire

$$\tilde{f}(x) = x^{2p} + \tilde{c}_p x^p. \quad (1)$$

Si $h < p$, on peut écrire $\tilde{f}(x) = x^{2p} + \tilde{c}_p x^p + \widetilde{c_{p-1}}x^{p-1} + \cdots + \tilde{c}_h x^h$. Supposons maintenant qu'il existe $k \in [h, p-1]$ tel que $\tilde{c}_k \neq 0$, et choisissons un tel k maximal. On a

$$H_k(\tilde{f}) = \binom{2p}{k} x^{2p-k} + \binom{p}{k} \widetilde{c_{p-k}} x^k + \tilde{c}_k = \tilde{c}_k,$$

et on montre comme précédemment que $\tilde{c}_k = 0$.

La relation (1) est donc établie dans tous les cas.

On peut trouver un nombre complexe z_p tel que $(z_p)^p = c_p$; on a alors $(\tilde{z}_p)^p = \tilde{c}_p$. En utilisant le rève du bizuth, la relation (1) devient

$$\tilde{f}(x) = x^{2p} + \tilde{z}_p^p x^p = (x^2 + \tilde{z}_p x)^p. \quad (2)$$

On a

$$\begin{aligned} H_p(\tilde{f}) &= \binom{2p}{p} x^p + \binom{p}{p} \tilde{c}_p = \\ &= 2x^p + \tilde{c}_p = (1+1)x^p + \tilde{c}_p = \\ &= (1+1)^p x^p + \tilde{c}_p = (2x + \tilde{z}_p)^p. \end{aligned}$$

Considérons le polynôme $g(x) = x^2 + \tilde{z}_p x$. Sa dérivée est le polynôme $g'(x) = 2x + \tilde{z}_p$. Puisque \tilde{f} et $H_p(\tilde{f})$ ont un zéro commun, il en va de même de g et g' , donc g a un zéro double et donc $\tilde{z}_p = 0$, d'où $\tilde{f}(x) = x^{2p}$, ce qui contredit notre hypothèse que f a deux zéros distincts. Cela établit le Théorème 3 dans le cas où $n = 2p$, avec p premier impair. \square

7. EXERCICES

Exercice 1 Montrer que $V^+ = \{z \in \mathbb{C} / v_p(z) \geq 0\}$ est un anneau et que $V^{++} = \{z \in \mathbb{C} / v_p(z) > 0\}$ est un idéal de cet anneau. Montrer qu'en outre tout idéal de V^+ qui contient strictement V^{++} est V^+ . Montrer que V^{++} est un idéal intègre.

Exercice 2 Soit p un nombre premier supérieur à 3. Montrer que la conjecture C-A est vraie pour les polynômes de degré 3 de $\mathbb{K}[x]$. En déduire que la conjecture C-A est vraie pour les polynômes de degré $3p$ sur \mathbb{C} .

Exercice 3 Montrer que le polynôme $x^4 + x \in \mathbb{Z}/7\mathbb{Z}$ ne satisfait pas la conjecture C-A, où les dérivées sont prises au sens de Hasse.

Exercice 4 Démontrer la conjecture C-A pour les polynômes de degré premier de $\mathbb{C}[x]$.

Exercice 5 Compléter la démonstration de la Proposition 1.

Exercice 6 Démontrer l'existence de la valuation p -adique sur \mathbb{Q} .

Exercice 7 Démontrer que le corps \mathbb{K} du texte est algébriquement clos, c'est-à-dire que tout polynôme non constant de $\mathbb{K}[x]$ a au moins un zéro dans \mathbb{K} .

ADS-M-4

Sommation non entière

Travail demandé :

Il vous est demandé d'étudier puis de présenter le(s) texte(s) joint(s) à travers un exposé de synthèse d'une durée comprise entre 15 et 20 minutes.

Si l'étude de la totalité du dossier et la préparation d'un exposé cohérent dans la durée impartie ne vous paraît pas possible, vous pouvez décider de vous limiter à une partie du dossier.

Remarques générales :

1. Les textes proposés, quelle que soit leur origine, peuvent présenter des défauts (coquilles typographiques, négligences ou sous-entendus de l'auteur, voire erreurs...) qui, sauf exception, n'ont pas été corrigés.

2. Les textes proposés peuvent contenir des exercices qu'il n'est pas demandé de résoudre. Néanmoins, vous pouvez vous aider des énoncés de ces exercices pour enrichir votre exposé.

3. Vous pouvez annoter les documents qui vous sont fournis. Vos annotations ne seront pas regardées par l'examineur.

Sommation non entière

1. INTRODUCTION. Les mathématiques sont l'art de l'abstraction et de la généralisation. Historiquement, les "nombres" ont d'abord été les entiers naturels ; ensuite, les nombres rationnels, négatifs, réels, complexes ont été introduits. De même, le concept de dérivation a été généralisé aux dérivations successives, d'ordre 2, 3, ... puis au "calcul fractionnaire" avec les dérivations d'ordre non entier ; il y a de même des travaux sur les itérations d'ordre non entier.

Cependant, lorsqu'on additionne des quantités, le nombre de termes considéré est généralement un entier naturel : on sait additionner 2 nombres, 3 nombres, mais comment additionner -7 nombres, ou bien les π premiers termes de la série harmonique ?

Dans cette note, on montre qu'il y a une façon très naturelle d'étendre les sommations aux cas où le "nombre de termes" est réel, voire complexe. On pourrait penser que cette méthode a été découverte il y a au moins deux cents ans - comme nous le pensions initialement. À notre grande surprise, cette méthode ne semble pas avoir été étudiée dans la littérature, ni être connue des experts, à l'exception de quelques remarques sporadiques dans les œuvres d'Euler (cf. la relation (13)). Bien évidemment, certaines méthodes classiques d'introduction de la fonction Γ sont un exemple de sommation avec un nombre complexe de termes, comme en témoigne l'équation (5) dans le paragraphe 1.2.

Dans cette note de présentation, nous ne donnons pas toutes les démonstrations.

1.1. Les Axiomes On commence par donner des conditions naturelles pour les sommations avec un nombre complexe de termes ; dans la suite, x, y, z et s désignent des nombres complexes et f et g des fonctions à valeurs complexes définies sur \mathbb{C} ou des parties de \mathbb{C} , éventuellement soumises à des conditions spécifiées ultérieurement.

(S1) Relation de Chasles.

$$\sum_{\nu=x}^y f(\nu) + \sum_{\nu=y+1}^z f(\nu) = \sum_{\nu=x}^z f(\nu).$$

(S2) Invariance par translation.

$$\sum_{\nu=x+s}^{y+s} f(\nu) = \sum_{\nu=x}^y f(\nu + s)$$

(S3) Linéarité. Pour λ et μ dans \mathbb{C} ,

$$\sum_{\nu=x}^y (\lambda f(\nu) + \mu g(\nu)) = \lambda \sum_{\nu=x}^y f(\nu) + \mu \sum_{\nu=x}^y g(\nu).$$

(S4) Cohérence.

$$\sum_{\nu=1}^1 f(\nu) = f(1)$$

(S5) Monômes. Pour tout $d \in \mathbb{N}$, l'application

$$z \mapsto \sum_{\nu=x}^z \nu^d$$

est polynômiale dans \mathbb{C} .

(S6) Continuité de la translation à droite. Si pour tout $z \in \mathbb{C}$, on a $\lim_{n \rightarrow +\infty} f(z+n) = 0$, alors

$$\lim_{n \rightarrow +\infty} \sum_{\nu=x}^y f(\nu+n) = 0; \quad (1)$$

plus généralement, pour toute suite de polynômes de degré fixé, $(p_n)_{n \in \mathbb{N}}$ telle que

$$\forall z \in \mathbb{C} : \lim_{n \rightarrow +\infty} |f(z+n) - p_n(z+n)| = 0,$$

on demande d'avoir

$$\lim_{n \rightarrow +\infty} \left| \sum_{\nu=x}^y f(\nu+n) - \sum_{\nu=x}^y p_n(\nu+n) \right| = 0. \quad (2)$$

Les quatre premiers axiomes (S1) - (S4) sont si naturels qu'il est difficile d'imaginer une méthode de sommation qui ne les satisfasse pas. Ils impliquent facilement que l'on a pour tout $n \in \mathbb{N}^*$: $\sum_{\nu=1}^n f(\nu) = f(1) + f(2) + \dots + f(n)$ et on assure une généralisation de la notion classique de sommation.

L'axiome (S5) est motivé par les relations classiques

$$\sum_{\nu=1}^n \nu = \frac{n(n+1)}{2}, \sum_{\nu=1}^n \nu^2 = \frac{n(n+1)(2n+1)}{6}, \sum_{\nu=1}^n \nu^3 = \left(\frac{n(n+1)}{2}\right)^2$$

et les formules similaires pour les puissances supérieures. On montrera ci-dessous que nos axiomes impliquent la validité de ces formules¹.

Enfin, l'axiome ($\overrightarrow{S6}$) est également naturel. Le premier cas, (1), exprime le fait que si f tend vers 0 à l'infini, la somme sur le domaine de longueur constante $[x+n, y+n]$ doit également tendre vers 0. La relation (2) exprime un phénomène semblable, à cela près qu'une approximation polynômiale a été introduite (cf. la discussion après la Proposition 1).

1.2. Des Axiomes à une Définition. Pour voir comment ces conditions déterminent une méthode de sommation de façon univoque, on commence par sommer des fonctions polynômiales. Le cas le plus simple est celui de la somme $\sum_{\nu=1}^{1/2} c$, où $c \in \mathbb{C}$. D'après l'axiome (S1), on doit avoir

$$\sum_{\nu=1}^{1/2} c + \sum_{\nu=3/2}^1 c = \sum_{\nu=1}^1 c.$$

En appliquant les axiomes (S2) au membre de gauche et (S4) au membre de droite, on obtient

$$\sum_{\nu=1}^{1/2} c + \sum_{\nu=1}^{1/2} c = c.$$

Il en résulte que $\sum_{\nu=1}^{1/2} c = c/2$. Ce calcul élémentaire peut être étendu pour couvrir toute somme d'une fonction polynômiale avec un nombre rationnel de termes.

Proposition 1. *Soit $p \in \mathbb{C}[X]$ et P l'unique polynôme tel que $P(0) = 0$ et $P(z) - P(z-1) = p(z)$, pour tout z complexe. Alors :*

1. Avec le même résultat, on peut remplacer les conditions que les sommes de monômes soient polynômiales par la condition *a priori* plus faible que ce soient des fonctions \mathcal{C}^∞ qui sont somme de leur série de Taylor.

– La définition

$$\sum_{\nu=x}^y p(\nu) = P(y) - P(x-1) \quad (3)$$

satisfait les axiomes (S1)- $(\vec{S6})$, dans le cas où f est un polynôme.

- Réciproquement, toute sommation qui satisfait les axiomes (S1)-(S4) satisfait également (3) pour tout polynôme $p \in \mathbb{Q}[X]$ et tout couple de complexes x et y dont la différence est dans \mathbb{Q} .
- Toute sommation qui satisfait les axiomes (S1)-(S5) satisfait également (3) pour tout polynôme $p \in \mathbb{C}[X]$ et tout couple de complexes x et y .

Démonstration. Pour démontrer la première assertion, on considère (3) comme une définition. On vérifie aisément les propriétés (S1), (S3), (S4) et (S5).

Pour vérifier (S2), on associe au polynôme p les polynômes q et r définis par $q(x) = p(x+s)$ et $r(x) = P(x+s) - P(s)$. On montre d'abord que “ $r = Q$ ” : on a $r(0) = 0$ et $r(z) - r(z-1) = P(z+s) - P(z+s-1) = q(z)$; on pourra donc appliquer (3) au couple (q, r) . On a

$$\begin{aligned} \sum_{\nu=x}^y p(\nu+s) &= \sum_{\nu=x}^y r(\nu) = q(y) - q(x-1) \\ &= P(y+s) - P(x+s-1) = \sum_{\nu=x+s}^{y+s} p(\nu). \end{aligned}$$

Montrons finalement que $(\vec{S6})$ est satisfait. On considère l'espace vectoriel V_d constitué des polynômes complexes de degré au plus égal d , muni de la norme $\|\cdot\|$ définie par $\|q\| = \sum_{i=0}^d |q(i)|$. La forme linéaire s_x^y sur V_d qui associe à un polynôme q le complexe $s_x^y(q) = \sum_{\nu=x}^y q(\nu)$ est continue et si une suite $(q_n)_n$ tend vers 0, alors la suite $s_x^y(q_n)$ tend vers 0. Pour montrer la relation (2), on définit q_n par $q_n(z) = p(x+n) - p_n(x+n)$. Par hypothèse, la suite q_n tend vers 0 au sens de la convergence simple ; elle tend donc vers 0 au sens de la norme $\|\cdot\|$ et on a donc (3).

Pour démontrer le deuxième point, on va développer l'idée utilisée pour démontrer que $\sum_{\nu=1}^{1/2} c = c/2$. D'après (S1), on a pour tout entier $r \geq 1$

$$\sum_{\nu=1}^r \nu^d = \sum_{\nu=1}^{r/s} \nu^d + \sum_{\nu=r/s+1}^{2r/s} \nu^d + \cdots + \sum_{\nu=(s-1)r/s+1}^r \nu^d.$$

Comme nous l'avons déjà remarqué, le membre de gauche garde son sens

classique. En utilisant (S2) et (S3), on en déduit

$$\sum_{\nu=1}^r \nu^d = \sum_{k=0}^{s-1} \sum_{\nu=1}^{r/s} (\nu + kr/s)^d = s \sum_{\nu=1}^{r/s} \nu^d + \sum_{k=0}^{s-1} \sum_{\nu=1}^{r/s} q_{d-1,k}(\nu),$$

où $q_{d-1,k}(\nu) = (\nu + kr/s)^d - \nu^d$ est un polynôme nul si $d = 0$ et de degré $d - 1$ sinon.

Cette relation étant établie, on raisonne par récurrence. Si $d = 0$, l'équation précédente détermine $\sum_{\nu=1}^{r/s} 1$ et par linéarité la somme $\sum_{\nu=1}^{r/s} c$, pour n'importe quel nombre complexe c . Une fois qu'on a évalué la somme $\sum_{\nu=1}^{r/s} f(\nu)$ sur tous les polynômes de degré $d-1$, la relation que l'on vient d'établir détermine $\sum_{\nu=x}^y \nu^d$, et par linéarité, cela détermine $\sum_{\nu=1}^{r/s} f(\nu)$ pour tous les polynômes de degré d . Si p est un polynôme et $x - y$ est dans \mathbb{Q} , les axiomes (S1)-(S4) impliquent que $\sum_{\nu=x}^y p(\nu) = \sum_{\nu=1}^{y-x+1} p(\nu + x - 1)$. Mais on a déjà montré que l'équation (3) est une définition admissible pour une méthode de sommation sur les polynômes; ainsi, si une méthode de sommation satisfait les relations (S1)-(S4), pour tout polynôme p et tout couple de complexes (x, y) de différence rationnelle, la relation (3) est satisfaite.

Le troisième point résulte du fait que deux polynômes qui coïncident sur les rationnels sont égaux. \square

Soit maintenant f une fonction arbitraire de \mathbb{C} dans \mathbb{C} . Si on cherche à définir $\sum_{\nu=x}^y f(\nu)$ pour x et y dans \mathbb{C} , on peut écrire

$$\sum_{\nu=x}^y f(\nu) + \sum_{\nu=y+1}^{y+n} f(\nu) = \sum_{\nu=x}^{x+n-1} f(\nu) + \sum_{\nu=x+n}^{y+n} f(\nu),$$

où $n \in \mathbb{N}^*$ est un entier positif arbitraire. On en déduit

$$\begin{aligned} \sum_{\nu=x}^y f(\nu) &= \sum_{\nu=x}^{x+n-1} f(\nu) - \sum_{\nu=y+1}^{y+n} f(\nu) + \sum_{\nu=x+n}^{y+n} f(\nu) \\ &= \sum_{\nu=1}^n (f(\nu + x - 1) - f(\nu + y)) + \sum_{\nu=x}^y f(\nu + n). \end{aligned} \quad (4)$$

À quoi nous sert ce réarrangement élémentaire? Dans la dernière ligne, la première somme du membre de droite porte sur un nombre entier de termes : elle peut être évaluée de façon classique. Pour ce qui est de la seconde somme, l'avantage est que nous avons translaté la somme par n vers la droite. Puisque

(4) a lieu pour tout n , on peut utiliser $(\vec{S6})$ pour évaluer la limite quand n tend vers l'infini : si pour tout z , la suite $(f(n+z))_n$ tend vers 0, alors $(\vec{S6})$ implique que la limite de la dernière somme est nulle quand n tend vers l'infini, et on obtient dans ce cas

$$\sum_{\nu=x}^y f(\nu) = \sum_{\nu=1}^{\infty} (f(\nu+x-1) - f(\nu+y)).$$

C'est bien sûr une condition particulière à imposer à f , mais cette idée peut être généralisée. Considérons par exemple la fonction $f = \ln$; pour $\nu \in [x, y] \subset \mathbb{R}^+$, les valeurs $f(\nu+n)$ sont bien approchées par la fonction constante $f(n)$, avec une erreur qui tend vers 0 quand n tend vers l'infini ; on dira dans ce cas que f est "approximativement constante". En utilisant (S3), on obtient pour tout $n \in \mathbb{N}^*$

$$\sum_{\nu=x}^y \ln(\nu+n) = \sum_{\nu=x}^y \ln(n) + \sum_{\nu=x}^y (\ln(\nu+n) - \ln n).$$

Mais, par $(\vec{S6})$, la dernière somme s'annule quand n tend vers l'infini, tandis que le sommant de la première somme est constant et on peut donc évaluer celle-ci par la Proposition 1. En passant à la limite en n , on obtient nécessairement

$$\sum_{\nu=x}^y \ln \nu = \lim_{n \rightarrow \infty} \left(\sum_{\nu=1}^n (\ln(\nu+x-1) - \ln(\nu+y)) + (y-x+1) \ln n \right).$$

Avant de généraliser encore plus notre définition, observons que notre méthode interpole la fonction factorielle de façon classique. On définit

$$\prod_{\nu=x}^y f(\nu) := \exp \left(\sum_{\nu=x}^y \ln f(\nu) \right)$$

et on obtient

$$\begin{aligned} z! = \prod_{\nu=1}^z \nu &= \lim_{n \rightarrow \infty} \exp \left(\sum_{\nu=1}^n \ln \left(\frac{\nu}{\nu+z} \right) + z \ln n \right) \\ &= \lim_{n \rightarrow \infty} \left(n^z \prod_{\nu=1}^n \frac{\nu}{\nu+z} \right) = \Gamma(z+1), \end{aligned} \quad (5)$$

en utilisant une représentation classique de la fonction Γ .

Nous pouvons maintenant nous appuyer sur le calcul heuristique (4), la Proposition 1 et l'axiome $(\vec{S6})$ pour donner une définition générale : tout ce dont nous avons besoin est de pouvoir approximer $f(n+z)$ par une suite de valeurs $p_n(n+z)$ de polynômes p_n de degré fixé.

Les domaines de définition nécessitent quelque soin : l'exemple du logarithme montre que l'on ne doit pas se limiter à des fonctions définies sur \mathbb{C} . Ce dont nous avons besoin, c'est de considérer des domaines de définition U stables par la translation $u \mapsto u+1$. Cela nous conduit à la définition suivante (le degré du polynôme nul est par convention $-\infty$).

Définition 1 (Fonctions fractionnairement sommables). *Soit U une partie non vide de \mathbb{C} et $d \in \mathbb{N} \cup \{-\infty\}$. Une fonction $f : U \rightarrow \mathbb{C}$ est dite fractionnairement sommable de degré d si les conditions suivantes sont satisfaites.*

- L'ensemble U est stable par la translation $u \mapsto u+1$.
- Il existe une suite de polynômes de degré d telle que

$$\text{pour tout } x \in U \text{ on a } \lim_{n \rightarrow \infty} |f(n+x) - p_n(n+x)| = 0.$$

- Pour tout x dans U et tout y tel que $y+1$ est dans U , la limite

$$\lim_{n \rightarrow \infty} \left(\sum_{\nu=n+x}^{n+y} p_n(\nu) + \sum_{\nu=1}^n (f(\nu+x-1) - f(\nu+y)) \right)$$

existe, où la première somme est définie en (3).

Dans ce cas, on note cette limite

$$\sum_{\nu=x}^y f(\nu) \text{ ou plus simplement } \sum_x^y f.$$

Si la fonction $\ln f$ est fractionnairement sommable, on pose

$$\prod_{\nu=x}^y f(\nu) = \exp \left(\sum_{\nu=x}^y f(\nu) \right).$$

On notera que cette définition ne dépend pas du choix des polynômes approximatifs. Soit en effet $(q_n)_n$ une autre famille de polynômes approximatifs ; pour tout $x \in U$ on a $\lim_{n \rightarrow \infty} (p_n(n+x) - q_n(n+x)) = 0$; puisque $\mathbb{C}_d[X]$ est de dimension finie $d+1$, cette dernière relation est valable pour tout x de \mathbb{C} . Comme on l'a montré dans la Proposition 1, les sommes de polynômes satisfont l'axiome $(\vec{S6})$; en substituant 0 à f et $q_n - p - n$ à p_n dans $(\vec{S6})$, on obtient

$$\lim_{n \rightarrow \infty} \left(\sum_{\nu=n+x}^{n+y} p_n(\nu) - \sum_{\nu=n+x}^{n+y} q_n(\nu) \right) = 0.$$

En outre, cette définition est la seule qui satisfasse les axiomes (S1)-($\vec{S6}$), comme on l'établit dans le théorème suivant

Théorème 1. *La Définition 1 satisfait les axiomes (S1)-($\vec{S6}$) (pour des domaines de définition convenables) et c'est la seule définition avec cette propriété, pour les classes de fonctions considérées.*

Démonstration. On a déjà montré l'unicité en dérivant la Définition 1 des axiomes (S1)-($\vec{S6}$). Il nous reste à vérifier que la Définition 1 induit les axiomes (S1)-($\vec{S6}$). Les axiomes (S3) et (S5) sont automatiquement satisfaits. Les axiomes (S1), (S2) et (S4) sont facilement vérifiés. Pour vérifier ($\vec{S6}$), on utilise la définition et les autres axiomes, en particulier (S1), et on calcule

$$\begin{aligned} \Delta &:= \lim_{n \rightarrow \infty} \left(\sum_{\nu=x}^y f(\nu+n) - \sum_{\nu=x}^y p_n(\nu+n) \right) \\ &= \lim_{n \rightarrow \infty} \left(\sum_{\nu=x}^y f(\nu+n) - \sum_{\nu=x}^y f(\nu) + \sum_{\nu=x}^y (f(\nu+x-1) - f(\nu+y)) \right) \\ &= \lim_{n \rightarrow \infty} \left(\sum_{\nu=x+n}^{y+n} f(\nu) - \sum_{\nu=x}^y f(\nu) + \sum_{\nu=x}^{n+x-1} f(\nu) - \sum_{\nu=y+1}^{y+n} f(\nu) \right) = 0. \end{aligned}$$

Cela montre que la Définition satisfait tous les axiomes. \square

2. PROPRIÉTÉS DES SOMMES FRACTIONNAIRES. Maintenant que nous avons une définition de sommes avec un nombre non entier de termes, il est intéressant de rechercher les propriétés des sommes finies classiques qui se concervent dans ce cadre, et d'en trouver de nouvelles.

2.1. Propriétés classiques généralisées. Une des identités les plus fondamentales pour les sommes finies est la sommation des progressions géométriques. Pour simplifier, on considère un réel $q \in [0, 1[$. La fonction $\nu \mapsto q^\nu$ est approximativement nulle et la définition donne

$$\sum_{\nu=0}^x q^\nu = \sum_{\nu=1}^{\infty} (q^{\nu-1} - q^{\nu+x}) = (1 - q^{x+1}) \sum_{\nu=1}^{\infty} q^{\nu-1} = \frac{1 - q^{x+1}}{1 - q}. \quad (6)$$

La formule de sommation des séries géométriques est donc valable pour tout $x \in \mathbb{C}$.

Un calcul similaire montre que la formule du binôme reste valable dans le nouveau cadre : pour tout $c \in \mathbb{C} \setminus \{-1, -2, \dots\}$ et $x \in \mathbb{C}$ avec $|x| < 1$, on a

$$(1+x)^c = \sum_{\nu=0}^c \binom{c}{\nu} x^\nu. \quad (7)$$

On peut étendre la formule (6) au cas $q > 1$ et la formule (7) au cas $|x| > 1$, mais cela implique la “sommation à gauche” que nous introduirons dans la troisième partie.

La multiplication des sommes fractionnaires fournit un exemple d’identité avec une structure plus compliquée. Pour $x \in \mathbb{C}$, et sous réserve que toutes les sommes non entières existent, on a

$$\begin{aligned} \left(\sum_{\nu=1}^x f(\nu) \right) \cdot \left(\sum_{\nu=1}^x g(\nu) \right) \\ = \sum_{\nu=1}^x \left(f(\nu)g(\nu) + f(\nu) \sum_{k=1}^{\nu-1} g(k) + g(\nu) \sum_{k=1}^{\nu-1} f(k) \right), \end{aligned} \quad (8)$$

ce qui généralise les formules du type $(a_1 + a_2)(b_1 + b_2) = a_1b_1 + a_2b_2 + a_2b_1 + b_2a_1$.

2.2. Propriétés nouvelles et fonctions spéciales. Comme nous l’avons vu dans le paragraphe 1.2, notre définition interpole la fonction factorielle par la fonction Γ ,

$$z! \equiv \prod_{\nu=1}^z = \Gamma(z + 1). \quad (9)$$

Une conséquence amusante est la relation

$$\prod_{\nu=1}^{-1/2} (\nu^2 + 1) = \tanh \pi; \quad (10)$$

que l’on obtient en utilisant la relation classique $\Gamma(z)\Gamma(1-z) = \pi/\sin(\pi z) = i\pi/\sin(\pi iz)$:

$$\begin{aligned} \prod_{\nu=1}^{-1/2} (\nu^2 + 1) &= \prod_{\nu=1}^{-1/2} (\nu + i) \prod_{\nu=1}^{-1/2} (\nu - i) = \frac{\Gamma(1/2 + i)\Gamma(1/2 - i)}{\Gamma(1 + i)\Gamma(1 - i)} \\ &= \frac{\Gamma(1/2 + i)\Gamma(1/2 - i)}{i\Gamma(i)\Gamma(1 - i)} = \frac{\sin(\pi i)}{i \sin(\pi(1/2 + i))} = \frac{\sin i\pi}{i \cos i\pi} \\ &= \frac{\sinh \pi}{\cosh \pi} = \tanh \pi. \end{aligned} \quad (11)$$

Plusieurs sommes non entières sont liées aux fonctions spéciales. Nous prendrons la série harmonique comme exemple. Puisque $\nu \mapsto \nu^{-1}$ est approximativement nulle, on a

$$\sum_{\nu=1}^x \frac{1}{\nu} = \sum_{\nu=1}^{\infty} \left(\frac{1}{\nu} - \frac{1}{\nu + x} \right), \quad (12)$$

et en particulier, on a

$$\sum_{\nu=1}^{-1/2} \frac{1}{\nu} = -2 \left(1 - \frac{1}{2} + \frac{1}{3} \cdots \right) = -2 \ln 2, \quad (13)$$

ce qui avait déjà été remarqué par Euler.

La fonction ζ d'Hurwitz est une fonction définie sur $\mathbb{C} \setminus \{1\} \times \mathbb{C} \setminus \{0, -1, -2, \dots\}$ et qui vérifie

$$\zeta(s, x) = \sum_{\nu=0}^{\infty} \frac{1}{(\nu + x)^s} \text{ pour } \Re(s) > 1 \text{ et } x > 0. \quad (14)$$

On note

$$\zeta'_s(s, x) = \frac{\partial}{\partial s} \zeta(s, x).$$

On peut montrer que l'on a

$$\sum_{\nu=1}^x \nu^a = \zeta(-a, 1) - \zeta(-a, x+1) \text{ pour } x \in \mathbb{C} \setminus \{-1, -2, -3, \dots\} \text{ et } a \neq -1, \quad (15)$$

avec le cas particulier

$$\sum_{\nu=1}^{-1/2} \nu^a = (2 - 2^{-a}) \zeta(-a, 1), \text{ pour } a \neq -1. \quad (16)$$

En dérivant (15) b fois (où b est un entier naturel), on obtient

$$\sum_{\nu=1}^x \nu^a (\ln \nu)^b = (-1)^b (\zeta^{(b)}(-a, 1) - \zeta^{(b)}(-a, x+1)) \quad (17)$$

ce qui conduit à quelques valeurs particulières inattendues, comme

$$\sum_{\nu=1}^{-1/2} \nu \ln \nu = -\frac{\ln 2}{24} - \frac{3}{2} \zeta'(-1, 1). \quad (18)$$

2.3. Séries miroir et sommation à gauche. L'identité

$$f(-10) + f(-9) + f(-8) + f(-7) = \sum_{\nu=-10}^{-7} f(\nu) = \sum_{\nu=7}^{10} f(-\nu)$$

semble si triviale qu'elle n'est presque jamais mentionnée. S'étend-elle aux sommes fractionnaires ? Il y a là un problème crucial puisqu'avec notre définition, la seconde somme implique de connaître le comportement de f au voisinage de $-\infty$. Nous traitons ce point dans le paragraphe suivant.

3. UN AUTRE AXIOME ET LA SOMMATION À GAUCHE. On peut modifier l'axiome $(\overrightarrow{S6})$ en s'intéressant aux limites quand $n \rightarrow -\infty$. On obtient ainsi l'axiome de "continuité de la translation à gauche" :

$(\overleftarrow{S6})$ **Continuité de la translation à gauche.** Si pour tout $z \in \mathbb{C}$, on a $\lim_{n \rightarrow +\infty} f(z - n) = 0$, alors

$$\lim_{n \rightarrow +\infty} \sum_{\nu=x}^y f(\nu - n) = 0; \quad (19)$$

plus généralement, pour toute suite de polynômes de degré fixé, $(p_n)_{n \in \mathbb{N}}$ telle que

$$\forall z \in \mathbb{C} : \lim_{n \rightarrow +\infty} |f(z - n) - p_n(z - n)| = 0,$$

on demande d'avoir

$$\lim_{n \rightarrow +\infty} \left| \sum_{\nu=x}^y f(\nu - n) - \sum_{\nu=x}^y p_n(\nu - n) \right| = 0. \quad (20)$$

En répétant les calculs du sous-paragraphe 1.2. on obtient une définition alternative que l'on ne donne pas ici formellement : c'est exactement la Définition 1, à cela près que toute limite $n \rightarrow \infty$ est remplacée par $n \rightarrow -\infty$.

On montre de même que cette définition est la seule qui satisfasse (S1)-(S5) et $(\overleftarrow{S6})$. On remarque qu'en général l'existence de $\overrightarrow{\sum}$ et $\overleftarrow{\sum}$ sont indépendantes ; en outre, si l'une et l'autre existent, elles peuvent avoir des valeurs différentes. Par exemple, pour tout z avec $\Re(z) > 0$, on a

$$\overrightarrow{\sum}_{\nu=1}^{-1/2} \nu^z = (-1)^{z+1} (2 - 2^{-z}) \zeta(-z, 1), \quad (21)$$

tandis que

$$\overleftarrow{\sum}_{\nu=1}^{-1/2} \nu^z = (2 - 2^{-z}) \zeta(-z, 1). \quad (22)$$

En revanche, on a la relation

$$\overrightarrow{\sum}_{\nu=a}^b f(\nu) = \overleftarrow{\sum}_{\nu=-b}^{-a} f(-\nu). \quad (23)$$

4. UN PRODUIT INFINI. Les sommes non entières ne sont pas qu'un nouvel environnement peuplé d'énoncés dénués de sens dans le contexte classique. En fait, quelques unes des formules obtenues sont bien connues et d'autres semblent nouvelles. Bien sûr, toutes les identités obtenues peuvent en principe être calculées sans recours aux sommes non entières. Mais les démonstrations fondées sur la sommation non entière peuvent être assez intuitives et simples, puisque la plupart des étapes ne sont que des généralisations de propriétés de sommation très classiques. Nous en donnons un exemple en donnant une forme close pour le produit infini

$$P(x) := \lim_{n \rightarrow \infty} \prod_{k=1}^{2n} \left(1 + \frac{2x}{k}\right)^{-k(-1)^k} \quad (24)$$

pour $x > -1/2$. Ce produit a été étudié en premier par Borwein et Dykshoorn en 1993. En prenant les logarithmes, on obtient

$$\ln P(x) = - \sum_{k=1}^{\infty} \left(2k \ln \left(1 + \frac{2x}{k}\right) - 2 \left(k - \frac{1}{2}\right) \ln \left(1 + \frac{2x}{2(k - 1/2)}\right)\right).$$

On considère la fonction $\nu \mapsto 2\nu \ln(1 + x/\nu)$ qui tend vers $2x$ quand $\nu \rightarrow \infty$. D'après la Définition 1, on a

$$\begin{aligned} & \sum_{\nu=1}^{-1/2} 2\nu \ln \left(1 + \frac{x}{\nu}\right) \\ &= \lim_{n \rightarrow \infty} \left[-\frac{1}{2} \cdot 2x + \sum_{k=1}^n 2k \ln \left(1 + \frac{x}{k}\right) - 2 \left(k - \frac{1}{2}\right) \ln \left(1 + \frac{x}{k - 1/2}\right) \right]. \end{aligned}$$

On obtient donc

$$\begin{aligned}
\ln P(x) &= -x - \sum_{\nu=1}^{-1/2} 2\nu \ln \left(1 + \frac{x}{\nu}\right) = -x - 2 \sum_{\nu=1}^{-1/2} \nu \ln \left(\frac{\nu+x}{\nu}\right) \\
&= -x - 2 \sum_{\nu=1}^{-1/2} \nu \ln(\nu+x) + 2 \sum_{\nu=1}^{-1/2} \nu \ln \nu \\
&= -x - 2 \sum_{\nu=1+x}^{-1/2+x} (\nu-x) \ln \nu - \frac{\ln 2}{12} - 3\zeta'_s(-1, 1) \\
&= -x - 2 \sum_{\nu=1+x}^{-1/2+x} \nu \ln \nu + 2x \sum_{\nu=1+x}^{-1/2+x} \ln \nu - \frac{\ln 2}{12} - 3\zeta'_s(-1, 1) \\
&= -x - 2 \left(\zeta'_s \left(-1, x + \frac{1}{2}\right) - \zeta'_s(-1, x+1) \right) \\
&\quad + 2x \left(\ln \left(\left(x - \frac{1}{2}\right)! \right) - \ln(x!) \right) - \frac{\ln 2}{12} - 3\zeta'_s(-1, 1),
\end{aligned}$$

où l'on a utilisé l'équation (18), le décalage d'indices, l'équation (17), l'équation (5) et la sommation continue. En prenant l'exponentielle des deux membres extrêmes, on obtient finalement

$$\begin{aligned}
P(x) &= 2^{-1/12} \left(\frac{\Gamma(x + \frac{1}{2})}{\Gamma(x+1)} \right)^{2x} \times \\
&\quad \times \exp \left(-x - 2\zeta'_s(-1, x + \frac{1}{2}) + 2\zeta'_s(-1, x+1) - 3\zeta'_s(-1, 1) \right).
\end{aligned}$$

[...]

ADS-M-5

PAVAGES

Travail demandé :

Il vous est demandé d'étudier puis de présenter le(s) texte(s) joint(s) à travers un exposé de synthèse d'une durée comprise entre 15 et 20 minutes.

Si l'étude de la totalité du dossier et la préparation d'un exposé cohérent dans la durée impartie ne vous paraît pas possible, vous pouvez décider de vous limiter à une partie du dossier.

Remarques générales :

1. Les textes proposés, quelle que soit leur origine, peuvent présenter des défauts (coquilles typographiques, négligences ou sous-entendus de l'auteur, voire erreurs...) qui, sauf exception, n'ont pas été corrigés.

2. Les textes proposés peuvent contenir des exercices qu'il n'est pas demandé de résoudre. Néanmoins, vous pouvez vous aider des énoncés de ces exercices pour enrichir votre exposé.

3. Vous pouvez annoter les documents qui vous sont fournis. Vos annotations ne seront pas regardées par l'examineur.

PAVAGES

Chapitres.

1. Quelques pavages du plan, des très classiques aux inconnus.
2. Les puzzles à solutions multiples, source de problèmes pédagogiques.
3. Une partie du 18e problème de Hilbert.
4. Une question de périodicité.
5. Pavages de Penrose.
6. Quelques autres développements.

En 1951, le mathématicien Hermann Weyl (1885-1955) donna quatre leçons à l'intention d'un public non spécialisé, à Princeton. Elles ont été rédigées sous forme d'un livre superbe, *Symmetry* [Weyl-52]. On y lit page 103 :

The art of ornament contains in implicit form the oldest piece of higher mathematics known to us.

Cela pour noter d'entrée que le thème de mon exposé est absolument classique et qu'il se rattache à de très anciennes traditions ; il reste néanmoins plein d'énigmes et de surprises, d'où son importance dans la recherche mathématique actuelle. En particulier, le sujet est vaste, comme en témoigne par exemple la référence encyclopédique obligée [GrSh-86]. Mon choix pour ce qui suit est donc hautement subjectif, nécessairement. Mais j'aimerais faire passer au moins une idée :

L'histoire de la géométrie en dimensions 2 et 3 ne se limite pas à Euclide (les "Éléments", vers 300 av. J.-C.), loin s'en faut : d'une part elle remonte à des temps bien plus anciens (allusion de Weyl, voir aussi par exemple [Jone-56] et [Ham1-16]), d'autre part elle a occupé tout le XXe siècle, et elle est encore aujourd'hui un domaine actif de recherche.

1. Quelques pavages du plan, des très classiques aux inconnus

Etant donné un polygone plan P , quand peut-on paver le plan tout entier par des copies isométriques de P , sans chevauchement ?

1.1. Pavages réguliers. Si P est un polygone régulier, il y a trois exemples classiques : pavage par hexagones réguliers, ou nids d'abeille, pavage par carrés, ou quadrillage, et pavage par triangles équilatéraux.

Convenons qu'un pavage par polygones est *strict* si, lorsque deux de ses polygones se touchent, ils ont en commun ou bien un seul point qui est un sommet de chacun d'entre eux, ou bien un côté entier de chacun d'entre eux. Il est alors facile de montrer que tout pavage strict du plan par une seule espèce de polygone régulier est l'un des trois exemples classiques : hexagonal, carré, ou triangulaire, comme ci-dessus. On attribue parfois la démonstration de ce résultat au mathématicien grec Pappus, du III^e siècle après J.-C. Le résultat a de très nombreuses illustrations : autant dans de magnifiques mosaïques antiques que dans divers pavages actuels de sols, par exemple dans nos salles de bain.

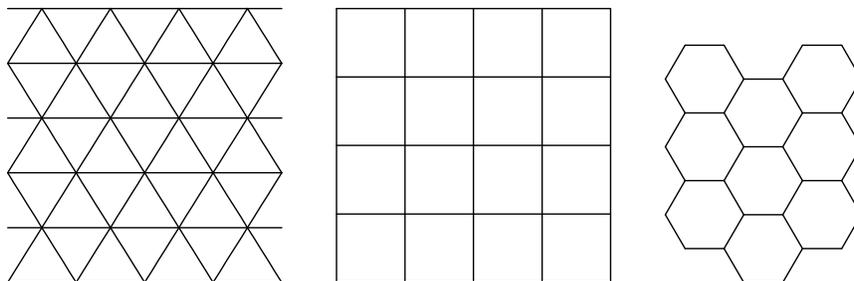


Figure 1 : pavages réguliers.

1.2. Pavages par triangles, hexagones et quadrilatères. Préparons-nous à évoquer un *problème ouvert*, c'est-à-dire un problème auquel de nombreux mathématiciens ont déjà réfléchi, mais dont la solution n'est toujours pas connue. On se donne un polygone P , disons convexe (sans "angle rentrant"), et on se demande s'il existe un pavage du plan par des copies isométriques de P . Mais on ne suppose plus que P est nécessairement régulier !

Si P est un triangle, il existe toujours un pavage par copies de P . En effet, deux triangles égaux s'assemblent naturellement pour former un parallélogramme, une suite de parallélogrammes pave une bande infinie à côtés parallèles, et une suite de bandes parallèles pave le plan.

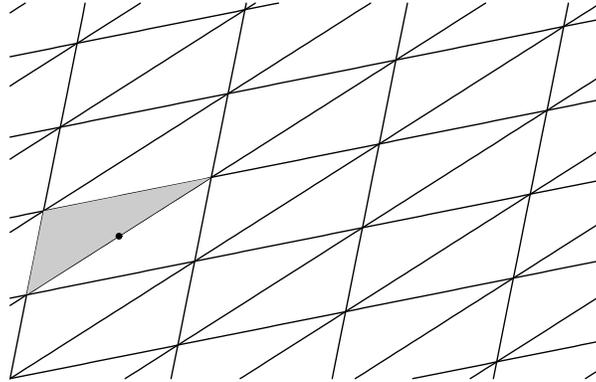


Figure 2 : pavage triangulaire.

Si P est un hexagone convexe, la réponse est un peu plus longue à formuler, mais elle est connue depuis la thèse [Rein-18] soutenue en 1918 par Karl Reinhardt (1895-1941), sous la direction de Bieberbach. En fait, il existe trois familles d'hexagones convexes, chaque hexagone de l'une de ces familles pave le plan (ce qu'on vérifie très facilement), et aucun autre hexagone convexe ne pave le plan (ce qui constitue la partie du théorème la plus délicate à montrer). Décrivons *l'une* de ces trois familles : il s'agit des hexagones convexes ayant une paire de côtés parallèles et de même longueur.

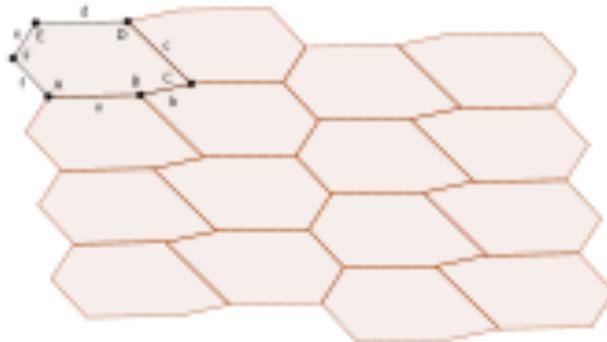


Figure 3 : pavage par certains hexagones.

Si P est un quadrilatère, il existe toujours un pavage par copies de P . La démonstration ressemble à celle de l'énoncé analogue pour les triangles : deux quadrilatères égaux s'assemblent naturellement pour former un hexagone avec côtés opposés parallèles et de même longueur, de tels hexagones pavent le plan, donc P aussi.

Si P est un polygone convexe à 7 côtés ou plus, la réponse est facile à formuler : P ne pave pas le plan. La démonstration n'est pas difficile, mais tout de même trop longue pour ce texte.

1.3. Le cas des pentagones, un problème ouvert. Reste le cas des pentagones convexes. Il est facile de montrer que certains pentagones pavent le plan (par exemple la réunion d'un carré avec un triangle rectangle dont l'hypothénuse coïncide avec un côté du carré – une “maison”) et que certains pentagones ne pavent pas (par exemple un pentagone régulier). Il est même facile de décrire d'assez grandes familles qui pavent la plan. Mais voici un *problème ouvert* :

trouver un algorithme qui, étant donné un pentagone convexe P décrit par ses angles et les longueurs de ses côtés, permette de savoir si oui ou non P pave le plan.

2. Les puzzles à solutions multiples source de problèmes pédagogiques

2.1. Un énoncé général. Soient F' et F'' deux figures planes ; pour simplifier la discussion, je supposerai que F' et F'' sont soit des polygones, soit des réunions de polygones (exemple : F' est constitué de deux carrés disjoints). Existe-t-il un puzzle dont les pièces $P_1^{(0)}, \dots, P_k^{(0)}$ peuvent s'assembler pour former F' , et également pour former F'' ? En d'autres termes, F' et F'' sont-ils *équidécomposables* ? Ou, encore en d'autres termes, cette fois plus techniques, existe-t-il deux familles (P'_1, \dots, P'_k) et (P''_1, \dots, P''_k) de polygones¹ du plan satisfaisant aux trois conditions suivantes ?

- (a) P'_j et P''_j sont isométriques à $P_j^{(0)}$, pour $j = 1, \dots, k$;
- (b) $\bigcup_{1 \leq j \leq k} P'_j = F'$ et $\bigcup_{1 \leq j \leq k} P''_j = F''$;
- (c) P'_i et P'_j sont d'intérieurs disjoints, et de même P''_i et P''_j sont d'intérieurs disjoints, pour $i, j \in \{1, \dots, k\}$ avec $i \neq j$.

Pour qu'un tel puzzle existe, il faut et il suffit que les aires de F' et F'' soient égales². Si F' et F'' sont donnés et ont la même aire, il est

¹Précisions concernant nos notations. Quand un pavage est décrit par une égalité du type $F' = \bigcup_{i \in I} P'_i$, aussi bien F' que les P'_i désignent des portions bien précises du plan, et il ne s'agirait pas de confondre un P'_i avec un translaté de ce P'_i . Au contraire, les pièces $P_1^{(0)}, \dots, P_k^{(0)}$ d'un puzzle sont des polygones dont seuls nous importent les angles et les longueurs des côtés, donc pas la position dans le plan ; ceci s'applique également aux pièces des collections \mathcal{P} apparaissant plus bas.

²L'affirmation est parfois connue sous le nom de *théorème de Wallace-Bolyai-Gerwien* (formulé et démontré vers 1830). Elle est à la base de la théorie de Hilbert des aires des polygones plans, elle est aussi sous-jacente à plusieurs constructions dans les *Eléments* d'Euclide.

naturel de chercher un puzzle comme ci-dessus avec un nombre k de pièces minimum.

2.2. Une infinité d'exemples. A titre de premier exemple, montrons une figure qui est (presque ...) une démonstration du théorème de Pythagore. En référence à un triangle rectangle T , dont on note c la longueur de l'hypothénuse et a, b les longueurs des cathètes, on prend pour F' un carré de côté c et pour F'' la réunion de deux carrés de côtés a et b .

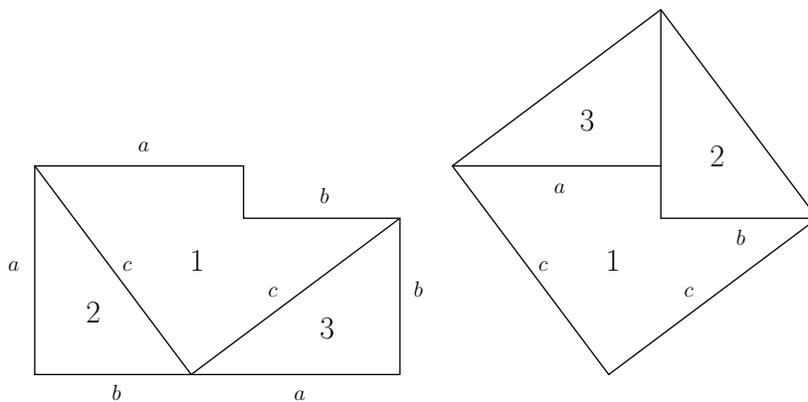


Figure 4 : le théorème de Pythagore (voir aussi [Peri] et [Cant–11]).

Cette “démonstration” indique que F' et F'' sont équidécomposables, avec $k = 3$ selon la notation utilisée plus haut.

Comme deuxième exemple, demandons comment il est possible d'équidécomposer un rectangle de côtés $A + a, B$ en un rectangle de côtés $A, B + b$, lorsque $(A + a)B = A(B + b)$ et $0 < a < A$. (La solution fait partie de la démonstration usuelle du théorème de Wallace-Bolyai-Gerwien évoqué plus haut. Indication en fin d'article.)

Ces exemples font partie d'une liste qui peut être allongée à l'infini ; voir par exemple [Rous]. Bien que n'ayant aucune pratique de l'enseignement secondaire, j'imagine que ces problèmes d'équidécomposition ont une vertu appréciable : on en distingue facilement les données des solutions !

3. Une partie du 18e problème de Hilbert

En 1900, Hilbert formula 23 problèmes qui ont joué depuis un rôle important dans le développement des mathématiques [Hilb–00]. Le 18e problème, relatif à la géométrie usuelle, s'articule en trois parties,

dont la deuxième est un problème en dimension trois, la dimension de la cristallographie ; rappelons que c'est de la fin du XIXe siècle que date la classification des groupes de symétrie des cristaux³. Avant de formuler l'analogie en dimension deux de cette partie du 18e problème, quelques préliminaires.

Soit Γ un groupe de transformations du plan formé d'isométries, c'est-à-dire de translations, rotations, symétries axiales, et symétries glissantes⁴ ; dire que Γ est un groupe, c'est dire qu'il contient avec toute isométrie γ l'isométrie inverse γ^{-1} , et avec toute paire d'isométries γ_1, γ_2 la composition $\gamma_1\gamma_2$. On suppose qu'il existe un polygone plan $P^{(0)}$ qui est un *domaine fondamental* pour l'action de Γ , c'est-à-dire qui est tel que les images $\gamma(P^{(0)})$ de $P^{(0)}$ par les éléments γ de Γ recouvrent le plan, et que les intérieurs des images $\gamma(P^{(0)})$ de $P^{(0)}$ soient disjointes deux à deux. Un polygone $P^{(0)}$ est dit *isoèdre* s'il apparaît de cette manière, et *anisoèdre* sinon. Il résulte de la définition qu'un polygone isoèdre pave le plan. Le problème de Hilbert peut s'énoncer comme suit⁵ :

existe-t-il un polygone anisoèdre qui pave le plan ?

La réponse pour la dimension 3 fut trouvée en 1928 par Reinhardt, déjà cité ; elle est positive. La réponse pour la dimension 2, due à Heesch [Hees-35], est également positive ; voir la figure 5, et [Hees-35].

³Quand le temps d'une découverte scientifique arrive, on remarque souvent (a posteriori !) qu'elle est due simultanément et indépendamment à plusieurs auteurs. La classification des groupes cristallographiques illustre bien ce fait, puisqu'elle est due essentiellement indépendamment au Russe Evgraf Stepanovich Fedorov (1853-1919), à l'Allemand Arthur Moritz Schönflies (1853-1928) et à l'Anglais William Barlow (1845-1934). Notons encore que le minéralogiste Fedorov et le mathématicien Schönflies ont d'abord obtenu leurs résultats indépendamment, puis qu'ils les ont comparés avant publication et ont ainsi pu apporter chacun quelques corrections au travail de l'autre.

⁴Rappelons qu'une *symétrie glissante* du plan est une isométrie de la forme $\gamma = \sigma_d\tau_a = \tau_a\sigma_d$, composition de σ_d et τ_a (l'ordre n'importe pas), où σ_d est une symétrie relativement à un axe (= une droite) d du plan et où τ_a est une translation d'amplitude a dans la direction de d . Les symétries glissantes sont caractérisées parmi les isométries du plan par deux propriétés : elles n'ont aucun point fixe et elles renversent l'orientation.

⁵C'est plus exactement l'analogie en dimension 2 du problème formulé par Hilbert en dimension 3 : existe-t-il un polyèdre anisoèdre qui pave l'espace ? Ou plus exactement, copié de la page 99 de [Hilb-fr] : *Existe-t-il aussi des polyèdres qui NE se présentent PAS comme régions fondamentales de groupes de déplacement, et au moyen desquels cependant on peut, en juxtaposant convenablement les exemplaires congruents, arriver à remplir sans lacunes l'espace tout entier ?*

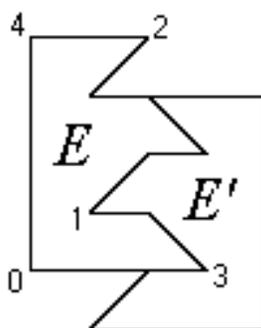


Figure 5 : emboîtement de deux décagones de Heesch.

On a trouvé depuis d'autres polygones anisoédres qui pavent le plan, dont certains pentagones convexes.

4. Une question de périodicité

4.1. **\mathcal{P} -pavages périodiques.** Dans les pavages discutés aux chapitres 1 et 3, tous les pavés sont isométriques. Abordons un cas plus général : étant donné une collection finie

$$\mathcal{P} = \{P_1^{(0)}, \dots, P_k^{(0)}\}$$

de polygones, convenons qu'un \mathcal{P} -pavage est un recouvrement sans chevauchement du plan

$$\mathbf{R}^2 = \bigcup_{i \in I} P_i, \quad \text{avec} \quad \overset{\circ}{P}_i \cap \overset{\circ}{P}_j = \emptyset \text{ si } i \neq j,$$

où chacun des polygones P_i est isométrique à l'un des $P_\ell^{(0)}$ et où $\overset{\circ}{P}_i$ désigne l'intérieur de P_i . La figure 6 indique un exemple de \mathcal{P} -pavage bien répandu, avec $k = 2$, la collection \mathcal{P} étant constituée d'un carré $P_1^{(0)}$ et d'un octogone régulier $P_2^{(0)}$.

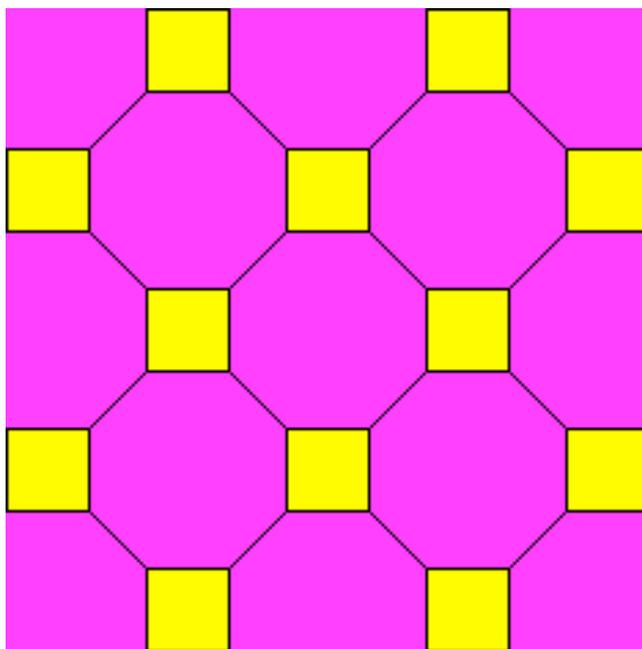


Figure 6 : pavage par carrés et octogones réguliers.

Un tel pavage est dit *périodique* s'il existe deux translations τ_1, τ_2 de directions différentes qui laissent le pavage *invariant*, c'est-à-dire telles que $\tau_1(P_i)$ soit un des P_j , et de même $\tau_2(P_i)$ soit un des P_j (un autre !), pour tout $i \in I$. Il est facile d'exhiber des collections \mathcal{P} pour lesquelles il existe d'une part des pavages périodiques et d'autre part des pavages non périodiques ; par exemple, si \mathcal{P} consiste en un seul polygone qui est un rectangle de côtés 1 et 2, on obtient

- un pavage périodique en subdivisant horizontalement chaque carré d'un quadrillage par carrés de côté 2,
- un pavage non périodique en subdivisant aléatoirement, horizontalement ou verticalement, chaque carré d'un quadrillage par carrés de côté 2,
- un autre pavage non périodique en subdivisant un carré horizontalement et tous les autres verticalement.

4.2. Le problème de Wang. Le problème suivant remonte à un travail du logicien Hao Wang, de 1961 ; il est lié à la question de savoir s'il existe un algorithme (ce qui est à peu près équivalent à la question de savoir s'il existe un programme d'ordinateur) qui, étant donné une famille finie \mathcal{P} de polygones, permet de savoir s'il existe un \mathcal{P} -pavage du plan. Voici l'énoncé du problème de Wang :

existe-t-il une famille $\mathcal{P} = \{P_1^{(0)}, \dots, P_k^{(0)}\}$ telle que
– il existe des \mathcal{P} -pavages du plan,
– il n'existe aucun \mathcal{P} -pavage périodique du plan ?

La réponse a sans doute surpris plus d'un expert : *oui, de telles collections \mathcal{P} existent.* Dans la première réponse (Berger, 1966) on avait $k > 1000$. Plus tard, on trouva d'autres réponses avec k plus petit, par exemple avec $k = 6$ (Robinson, 1971), ce qui permet de faire des dessins. La réponse aujourd'hui la plus célèbre est celle de Penrose, pour laquelle $k = 2$; voir [Penr-74] et [Gard-77]. (Une réponse essentiellement équivalente fut obtenue simultanément et indépendamment par Robert Ammann [Sene-04].) On peut en décrire plusieurs variantes, dont l'une est évoquée ci-dessous.

On ne connaît pas la réponse si on impose la condition $k = 1$ (tous les pavés isométriques),

ceci constitue un autre problème ouvert

captivant et important.

(Notons toutefois que, en dimension 3, une variante du problème est résolue depuis 1988. En effet, Conway et Schmitt ont décrit un polyèdre $P^{(0)}$ qui a les propriétés suivantes : (i) il existe un pavage de \mathbf{R}^3 par des polyèdres dont chacun est une image de $P^{(0)}$ par une isométrie *préservant l'orientation*, (ii) il n'existe aucun tel pavage de \mathbf{R}^3 qui soit invariant par une translation. Voir la description dans [Sene-95], ainsi que [Radi-95].)

5. Sur les pavages de Penrose

5.1. Définition des \mathcal{P}_A -pavages. Nous allons décrire une *variation* des pavages de Penrose. “Notre” famille \mathcal{P}_A est constituée de deux triangles isocèles à sommets coloriés, notés L_A et S_A (avec “L” pour “Large” et “S” pour “Small”), définis comme suit :

- les angles de L_A sont $\theta, 2\theta, 2\theta$, ses côtés $1, \varphi, \varphi$,
un des sommets de la base (d'angle 2θ) est blanc,
les deux autres sommets sont noirs ;
- les angles de S_A sont $3\theta, \theta, \theta$, ses côtés $\varphi, 1, 1$,
un des sommets de la base (d'angle θ) est noir,
les deux autres sommets sont blancs ;

où θ désigne l'angle $\pi/5$ (= un dixième de tour = 36°) et où

$$\varphi = \frac{1 + \sqrt{5}}{2} \approx 1,618033987\dots$$

désigne le *nombre d'or*. Nos descriptions de L_A et S_A illustrent les égalités trigonométriques

$$2 \cos(\theta) = \varphi \quad \text{et} \quad 2 \cos(2\theta) = \varphi^{-1} = \varphi - 1.$$

Notons pour la suite que chacun de ces triangles a un seul côté ayant deux extrémités de la même couleur, et nous l'appellerons le *côté distingué* du triangle. Notons aussi que les angles aux deux extrémités d'un côté distingué sont distincts, c'est-à-dire qu'il y en a un grand et un petit ; les côtés distingués sont *orientés*, du sommet de petit angle vers le sommet de grand angle.

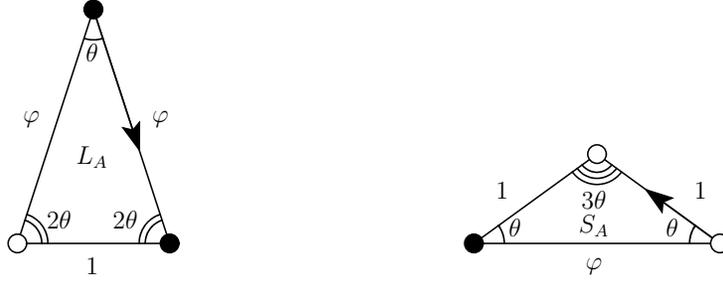


Figure 7 : les triangles L_A et S_A .

Convenons qu'un \mathcal{P}_A -pavage désigne ici un pavage strict $\mathbf{R}^2 = \bigcup_{i \in I} P_i$ dans lequel chaque P_i est un triangle à sommets coloriés, isométrique à l'un de L_A , S_A par une isométrie respectant les couleurs des sommets, deux tels triangles P_i et P_j étant toujours d'intérieurs disjoints, et avec de plus les conditions suivantes :

- (C1) chaque fois qu'un sommet d'un P_i coïncide avec un sommet d'un P_j , ces deux sommets sont de la même couleur,
- (C2) chaque fois qu'un côté distingué de P_i coïncide avec un côté distingué de P_j , l'orientation de ce côté dans P_i coïncide avec son orientation dans P_j .

A priori, il n'est nullement évident qu'il existe un seul \mathcal{P}_A -pavage du plan. Nous allons toutefois tenter d'indiquer pourquoi

- il existe des \mathcal{P}_A -pavages (il en existe même une infinité non dénombrable),
- mais il n'existe aucun \mathcal{P}_A -pavage périodique.

5.2. Sur l'apériodicité des \mathcal{P}_A -pavages. Introduisons une second famille \mathcal{P}_B , constituée de deux triangles isocèles à sommets coloriés, notés L_B et S_B , définis comme suit :

- les angles de L_B sont $3\theta, \theta, \theta$, ses côtés $\varphi^2, \varphi, \varphi$, un des sommets de la base (d'angle θ) est blanc, les deux autres sommets sont noirs ;
- les angles de S_B sont $\theta, 2\theta, 2\theta$, ses côtés $1, \varphi, \varphi$, un des sommets de la base (d'angle 2θ) est blanc, les deux autres sommets sont noirs.

Les triangles L_B et S_B ont également un côté distingué, qui est orienté.

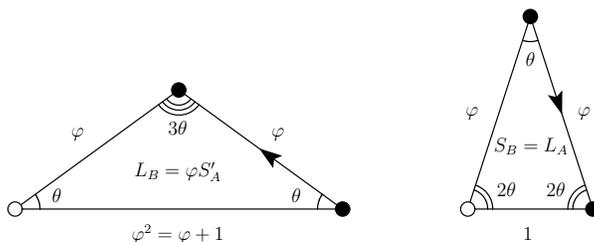


Figure 8 : les triangles L_B et S_B .

Il y a trois observations fondamentales à ne pas manquer :

- $S_B = L_A$,
- L_B est une juxtaposition d'une copie de L_A et d'une copie de S_A , la juxtaposition respectant les conditions (C1) et (C2),
- dans tout \mathcal{P}_A -pavage, le petit côté à extrémités noire et blanche d'une copie P' de S_A est nécessairement adjacent à la base d'une copie P'' de L_A et la réunion de P' et P'' est précisément une copie de L_B .

Il en résulte que, à partir de tout \mathcal{P}_A -pavage $\bigcup_{i \in I} P_i$ (supposé exister !), on peut définir de manière unique un \mathcal{P}_B -pavage $\bigcup_{j \in J} Q_j$ par copies isométriques de L_B et S_B , et satisfaisant aux conditions (C1) et (C2).

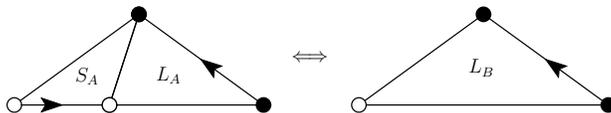


Figure 9 : illustration de l'observation (ii).

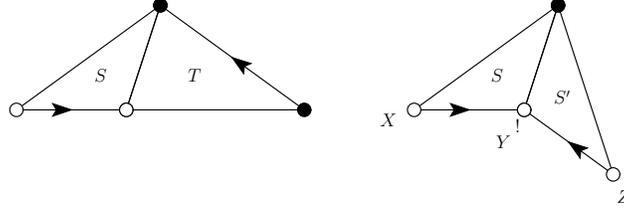


Figure 10 : illustration de l'observation (iii).

Avec un peu plus de temps, on montre de manière élémentaire que ce *processus d'inflation* peut se répéter, et fournit ainsi une suite infinie de pavages pour des familles $\mathcal{P}_A, \mathcal{P}_B, \mathcal{P}_C, \mathcal{P}_D, \dots$, suite qu'il est d'ailleurs préférable de noter désormais $(\mathcal{P}_{(n)})_{n \geq 1}$, avec $\mathcal{P}_{(n)} = \{L_{(n)}, P_{(n)}\}$, et $\mathcal{P}_{(0)} = \mathcal{P}_A$, $\mathcal{P}_{(1)} = \mathcal{P}_B$, etc. De plus, on constate que la famille $\mathcal{P}_{(n+4)}$ est la famille constituée des deux images des triangles du niveau n par une homothétie de rapport φ^2 , ce qui s'écrit aussi

$$L_{(n+4)} = \varphi^2 L_{(n)}, \quad S_{(n+4)} = \varphi^2 S_{(n)}$$

pour tout $n \geq 0$. (Si on ne s'occupe que des triangles, en oubliant la coloration des sommets et l'orientation des côtés distingués, on a $S_{(n+1)} = L_{(n)}$ et $L_{(n+1)} = S_{(n)} \cup L_{(n)}$. En revanche, si on tient compte des colorations et des orientations, la première valeur de $m > n$ pour laquelle $L_{(m)}$ et $S_{(m)}$ sont des copies homothétiques de $L_{(n)}$ et $S_{(n)}$ est bien $m = n + 4$.)

Voici l'argument qui exclut l'existence de $\mathcal{P}_{(0)}$ -pavage périodique. S'il en existait un, invariant par une translation τ , le $\mathcal{P}_{(n)}$ -pavage obtenu par n inflations serait également invariant par τ , pour tout $n \geq 1$. Pour n assez grand, les dimensions caractéristiques de tous les triangles de ce $\mathcal{P}_{(n)}$ -pavage seraient supérieures à l'amplitude de la translation τ , et ceci est absurde.

5.3. Sur l'existence des \mathcal{P}_A -pavages. L'argument est semblable à l'argument précédent, à la différence importante près qu'il convient de d'abord *subdiviser* au lieu de *recoller*.

Plus précisément, à partir de la paire $\mathcal{P}_{(0)} = \{L_{(0)} = L_A, S_{(0)} = S_A\}$, on définit une paire $\mathcal{P}_{(-1)} = \{L_{(-1)}, S_{(-1)}\}$ avec

- $L_{(-1)} = S_{(0)}$,
- $S_{(-1)}$ tel que $S_{(-1)}$ et $L_{(-1)}$ se recollent pour former $L_{(0)}$.

Il en résulte que, à partir d'un $\mathcal{P}_{(0)}$ -pavage, même partiel, c'est-à-dire un pavage d'une partie F du plan, on définit d'abord de manière unique un $\mathcal{P}_{(-1)}$ -pavage de cette même partie F , puis en itérant le procédé une famille de $\mathcal{P}_{(-n)}$ -pavages de F .

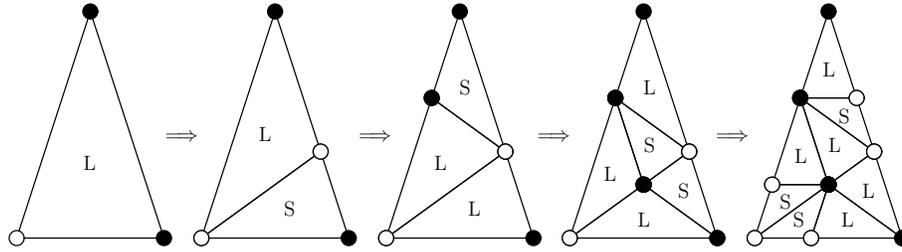


Figure 11 : $\mathcal{P}_{(-1)}$ -pavage, ..., $\mathcal{P}_{(-4)}$ -pavage de $F = L_A$.

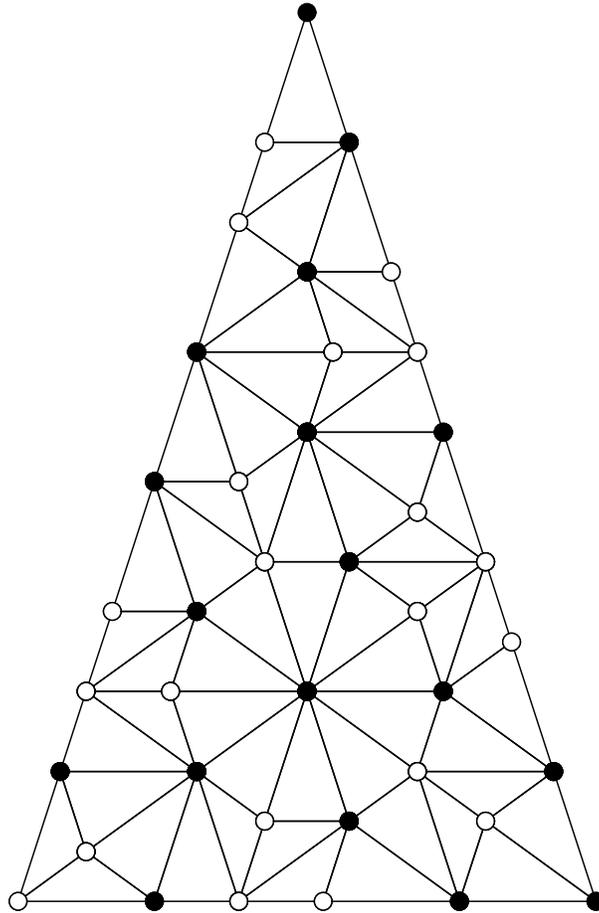


Figure 12 : $\mathcal{P}_{(-8)}$ -pavage du même triangle L_A .

On constate à nouveau que

$$L_{(-n-4)} = \varphi^{-2}L_{(-n)}, \quad S_{(-n-4)} = \varphi^{-2}S_{(-n)}$$

pour tout $n \geq 0$. Donc, en dilatant le $\mathcal{P}_{(-4n)}$ -pavage de L_A par le rapport convenable, qui est φ^{2n} , on obtient un $\mathcal{P}_{(0)}$ -pavage d'une copie homothétique $F_n := \varphi^n L_A$ de L_A , de plus en plus grande lorsque n croît. En observant qu'on peut ajuster les pavages de F_n et de F_{n+4} , on arrive à montrer qu'il existe un $\mathcal{P}_{(0)}$ -pavage du plan tout entier.

5.4. Quelques propriétés des \mathcal{P}_A -pavages. En exploitant les règles d'inflation et de décomposition, on peut montrer que :

- (a) il existe une infinité non dénombrable de \mathcal{P}_A -pavages non isométriques deux à deux⁶,
- (b) un \mathcal{P}_A -pavage ne possède jamais de symétrie de translation,
- (c) toute partie finie de tout \mathcal{P}_A -pavage se retrouve une infinité de fois dans tout autre \mathcal{P}_A -pavage,
- (d) dans tout \mathcal{P}_A -pavage, il existe des parties finies arbitrairement grandes qui possèdent *une symétrie de rotation d'ordre 5*,
- (e) à deux exceptions près, les \mathcal{P}_A -pavages ne possèdent globalement pas de symétrie d'ordre 5.

Il est frappant de comparer la condition (c) ci-dessus à un résultat fondamental de toute la cristallographie classique :

*pour tout pavage périodique du plan ou de l'espace,
l'ordre d'une rotation de l'espace laissant le pavage invariant
n'est jamais 5, plus précisément
l'ordre d'une telle rotation est l'un des entiers 1, 2, 3, 4, 6.*

Remarque.

Dans un \mathcal{P}_A -pavage du plan, les copies de L_A apparaissent nécessairement par paires adjacentes le long de leurs côtés distingués, et constituent ainsi des quadrilatères appelés *cerfs-volants* ; de même, les copies de S_A apparaissent nécessairement par paires le long de leurs côtés distingués, et constituent des *fléchettes*. Ainsi, tout pavage du plan par copies de L_A et S_A fournit un pavage du plan par cerfs-volants et fléchettes. On obtient aussi un pavage par deux types de losanges. Pour en voir plus, regarder par exemple

<http://www.ams.org/samplings/feature-column/fcarc-penrose>
ou un nombre presque infini de pages web en tapant sur google "Penrose tiling".

⁶Deux \mathcal{P}_A -pavages $\bigcup_{i \in I} P_i$ et $\bigcup_{j \in J} Q_j$ du plan sont *isométriques* s'il existe une isométrie ψ du plan et une bijection $\nu : I \rightarrow J$ telle que $\psi(P_i) = Q_{\nu(i)}$ pour tout $i \in I$, l'isométrie ψ respectant les couleurs des sommets des polygones.

5.5. Usage des \mathcal{P} -pavages en physique des quasi-cristaux. Certains alliages métalliques peuvent prendre des formes *quasi-cristallines*, vues pour la première fois au microscope électronique par Dan Shechtman au matin du 8 avril 1982. La découverte était si révolutionnaire qu'elle fut incomprise et combattue pendant assez longtemps, de sorte que la première publication la relatant date de novembre 1984 [SBGC–84]. Les choses ont bien changé depuis 1982, puisque le prix Nobel de chimie 2011 a été décerné à Dan Shechtman pour sa découverte ; voir la remarquable page officielle [Nobel].

Pour obtenir un quasi-cristal, on peut par exemple refroidir très brutalement une goutte d'aluminium et de manganèse en proportions convenables. A l'observation par rayons X ou au microscope électronique, le matériau obtenu révèle des symétries d'ordre 5, en contradiction flagrante avec toute la cristallographie classique, aussi bien expérimentale que théorique. Pour lever la contradiction, il faut admettre que le matériau obtenu n'a pas de structure cristalline ; il s'agit d'un nouveau type de structure, dite quasi-cristalline, qui a enrichi la théorie de la physique du solide.

Les modèles utilisés depuis pour comprendre ces quasi-cristaux font un large appel à des pavages non périodiques, qui partagent avec les quasi-cristaux la propriété d'exhiber un *ordre qui est à la fois à grande distance et non périodique*.

Les pavages de type Penrose sont encore aujourd'hui l'objet de nombreuses recherches et de vulgarisations diverses.

6. Quelques autres développements

Jusqu'ici dans cet exposé, quand il s'agissait de comparer deux polygones (ou polyèdres) ou deux pavages, la comparaison se faisait toujours à l'aide d'isométries, c'est-à-dire d'éléments d'un groupe G , qui est le *groupe Iso(E^2) de toutes les isométries⁷ du plan euclidien E^2* , ou le *groupe Iso(E^3) de toutes les isométries de l'espace euclidien E^3* . Or il est intéressant de varier l'espace et le groupe !

⁷Voici une liste, complète et sans répétition, des éléments de ce groupe :

- (id) l'identité,
- (tr) les translations (autres que l'identité), une par vecteur non nul du plan
- (rot) les rotations (autres que l'identité), une par couple (C, ψ) formé d'un centre C dans le plan et d'un angle $\psi \in]0, 2\pi[$,
- (sim) les symétries axiales, une par droite du plan,
- (sgl) les symétries glissantes, une par couple (d, a) formé d'une droite orientée d et d'une amplitude $a > 0$.

Pour les symétries glissantes, voir la note 4.

Depuis la découverte de la géométrie hyperbolique au cours du XIXe siècle (Gauss, Lobatchevski, Bolyai, Poincaré, Klein, ...) ⁸, on peut étudier des pavages de l'espace hyperbolique H^n de dimension n , dont le groupe d'isométrie $G = \text{Iso}(H^n)$ a des propriétés bien différentes de celles de $\text{Iso}(E^n)$. En particulier, l'artiste hollandais Maurits Cornelis Escher (1898–1972) s'est rendu célèbre en créant (notamment) de multiples pavages du plan hyperbolique.



Figure 13 : gravure sur bois (Escher, 1958).

Le sujet a aussi vu de belles contributions du mathématicien Harold Scott MacDonald Coxeter (1907–2003).

Ne serait-ce que pour mentionner qu'il existe une activité récente dans ce domaine, citons un article [MaMo–98] dans lequel les auteurs décrivent une méthode (il en existe d'autres) pour construire des

⁸Et ce n'est pas fini ! Thurston a changé notre compréhension de la géométrie hyperbolique, notamment par son cours à Princeton de la fin des années 1970. Plus récemment, la "conjecture de géométrisation" de Thurston a été démontrée dans trois textes rendus publics en 2002 et 2003 par Grisha Perelman. Cette conjecture – maintenant un théorème – implique que toute variété de dimension 3 est constituée de pièces munies de géométries, parmi lesquelles la géométrie hyperbolique joue le premier rôle.

pavages apériodiques du plan hyperbolique. Un pavage de H^2 est dit périodique s'il existe un groupe d'isométries $\Gamma \subset \text{Iso}(E^2)$ qui d'une part laisse le pavage invariant et pour lequel d'autre part il existe une partie compacte K de H^2 telle que $\bigcup_{\gamma \in \Gamma} \gamma(K) = H^2$.

Par exemple : soit T un triangle hyperbolique isocèle d'angles α, β, β tels que $5\alpha + 8\beta = 2\pi$, et tels que α ne soit pas un multiple rationnel de π . Il existe un pavage du plan hyperbolique par copies isométriques de T , mais il n'existe aucun tel pavage qui soit de plus périodique.

Le groupe G de transformation de l'espace à paver X n'a pas besoin d'être un groupe d'isométries ! Parmi les exemples standards, mentionnons le groupe des *transformations affines* de l'espace \mathbf{R}^n , ou le groupe des *transformations projectives* de l'espace projectif $\mathbf{P}_{\mathbf{R}}^n$. Ce dernier cas a fait l'objet de plusieurs travaux récents d'Yves Benoist, dont on peut trouver la liste sur sa page web [Benoist] ; voir en particulier son article de vulgarisation [Beno-01].

REFERENCES

- [Beno-01] Yves BENOIST, *Pavages du plan*, Journées X-UPS des 14 & 15 mai 2001,
<http://www.math.polytechnique.fr/xups/xups01.01.pdf>
- [Benoist] Yves BENOIST, *Page web*,
<http://www.math.u-psud.fr/~benoist/>
- [Cant-11] Serge CANTAT, *Découpage d'Airy et théorème de Pythagore*, Site "images des mathématiques",
<http://images.math.cnrs.fr/Decoupage-d-Airy-et-theoreme-de.html>
- [Gard-77] Martin GARDNER, *Extraordinary nonperiodic tiling that enriches the theory of tiles*, Scientific American **236** (1977) 110–121.
- [Ghys-11] Etienne GHYS, *L'énigme des pentagones*, Images des mathématiques, 28 octobre 2011
<http://images.math.cnrs.fr/L-enigme-des-pentagones.html>
- [GrSh-86] Branko GRÜNBAUM et Geoffrey C. SHEPHARD, *Tilings and patterns*, Freeman and Co, 1987.
- [Haml-16] Alfred Dwight Foster HAMLIN, *A history of ornament, ancient and medieval*, New York, The Century Co., 1916.
- [HaKw-11] Pierre DE LA HARPE, *Prix Nobel de chimie, quasi-cristaux, périodicité et pavages de Penrose*, Images des mathématiques, 28 octobre 2011
<http://images.math.cnrs.fr/Prix-Nobel-de-chimie-quasi.html>
- [Hees-35] H. HEESCH, *Aufbau der Ebene aus kongruenten Bereichen*, Nachr. Ges. Wiss. Göttingen (2) **1** (1935), 115–117. Traduction anglaise en ligne :
<http://www.angelfire.com/mn3/anisohedral/heesch35.html>
- [Hilb-00] David HILBERT, *Mathematische Probleme*, deuxième congrès international des mathématiciens, Paris, 1900. Göttingen Nachrichten, 1900. En ligne (original et traduction anglaise) :
<http://www.mathematik.uni-bielefeld.de/~kersten/hilbert/rede.html>
<http://aleph0.clarku.edu/~djoyce/hilbert/problems.html#prob18>

- [Hilb–fr] David HILBERT, *Sur les problèmes futurs des Mathématiques*, traduction de M.L. Laugel, Compte Rendu du deuxième congrès international des mathématiciens tenu à Paris du 6 au 12 août 1900, Gauthier-Villars (1902), 58–114 (traduction de [Hilb–00]).
- [MaMo–98] Gennadi MARGULIS and Shahar MOZES, *Aperiodic tilings of the hyperbolic plane by convex polygons*, Israel Journal of Mathematics **107** (1998), 319–325.
- [Jone–56] Owen JONES, *The grammar of ornament*, 1856 (et van Nostrand, 1982).
- [Nobel] NOBELPRIZE.ORG, *The Nobel Prize in Chemistry 2011*, http://www.nobelprize.org/nobel_prizes/chemistry/laureates/2011/info.html
- [Penr–74] Roger PENROSE, *The rôle of aesthetics in pure and applied mathematical research*, Bull. Inst. Math. Appl. **10** (1974), 266–271. Voir aussi, du même : *Set of tiles for converging a surface*, brevet, 9 janvier 1979, <http://v3.espacenet.com/publicationDetails/biblio?CC=US&NR=4133152&KC=&FT=E>
- [Peri] Perigal’s Proof of the Pythagorean Theorem, <http://www.cut-the-knot.org/pythagoras/Perigal.shtml>
- [Radi–95] Charles RADIN, *Aperiodic tilings in higher dimensions*, Proc. Amer. Math. Soc. **123** (1995), 3543–3548.
- [Rein–18] Karl REINHARDT, *Über die Zerlegung der Ebene in Poygone*, Dissertation der Naturwiss. Fakultät, Universität Frankfurt/Main, Borna, 1918. Voir aussi Béla BOLLOBÁS, *Filling the plane with congruent convex hexagons without overlapping*, Ann. Univ. Sci. Budapest. Eötvös Sect. Math. **6** (1963), 117–123.
- [Rous] *Dissections de polygones, réguliers ou non réguliers*, dossier proposé par Alain ROUSSEAU, en particulier la “liste des animations”, <http://www.cabri.net/abracadabri/abraJava/Dissection/ListDiss.html>
- [SBGC–84] D. SHECHTMAN, I. BLECH, D. GRATIAS, and J. W. CAHN, *Metallic Phase with Long-Range Orientational Order and No Translational Symmetry*, Phys. Rev. Lett. **53** (1984), 1951–1953.
- [Sene–95] Marjorie SENECHAL, *Quasicrystals and geometry*, Cambridge University Press, 1995.
- [Sene–04] Marjorie SENECHAL, *The mysterious Mr. Ammann*, Mathematical Intelligencer **26** (2004), 10–21
<http://www.springerlink.com/content/ct873r41m7605p44/fulltext.pdf>
http://tilings.math.uni-bielefeld.de/substitution_rules/ammann_beenker
- [Weyl–52] Hermann WEYL, *Symmetry*, Princeton University Press, 1952.

ADS-M-6

Théorème de Perron-Frobenius

Travail demandé :

Il vous est demandé d'étudier puis de présenter le(s) texte(s) joint(s) à travers un exposé de synthèse d'une durée comprise entre 15 et 20 minutes.

Si l'étude de la totalité du dossier et la préparation d'un exposé cohérent dans la durée impartie ne vous paraît pas possible, vous pouvez décider de vous limiter à une partie du dossier.

Remarques générales :

1. Les textes proposés, quelle que soit leur origine, peuvent présenter des défauts (coquilles typographiques, négligences ou sous-entendus de l'auteur, voire erreurs...) qui, sauf exception, n'ont pas été corrigés.

2. Les textes proposés peuvent contenir des exercices qu'il n'est pas demandé de résoudre. Néanmoins, vous pouvez vous aider des énoncés de ces exercices pour enrichir votre exposé.

3. Vous pouvez annoter les documents qui vous sont fournis. Vos annotations ne seront pas regardées par l'examineur.

Théorème de Perron-Frobenius

- Sections :*
1. Rappels sur les normes matricielles et le rayon spectral
 2. Matrices positives
 3. Matrices strictement positives et théorème de Perron-Frobenius
 4. Matrices de permutation
 5. Matrices irréductibles

On désigne par \mathbb{K} le corps des réels ou des complexes.

Si n, m sont deux entiers naturels non nuls, on désigne par $\mathcal{M}_{n,m}(\mathbb{K})$ l'espace vectoriel des matrices à n lignes et m colonnes à coefficients dans \mathbb{K} .

Pour $n = m$, on note $\mathcal{M}_n(\mathbb{K})$ pour $\mathcal{M}_{n,n}(\mathbb{K})$.

Un vecteur de \mathbb{K}^n est identifié à un élément de $\mathcal{M}_{n,1}(\mathbb{K})$.

Pour $A = ((a_{i,j}))_{\substack{1 \leq i \leq n \\ 1 \leq j \leq m}}$ dans $\mathcal{M}_{n,m}(\mathbb{K})$ et $x = (x_i)_{1 \leq i \leq m}$ dans \mathbb{K}^m , on note $(Ax)_i$ la composante numéro i du vecteur Ax .

Pour toute matrice $A = ((a_{i,j}))_{\substack{1 \leq i \leq n \\ 1 \leq j \leq m}}$ dans $\mathcal{M}_{n,m}(\mathbb{K})$, on note

$$|A| = ((|a_{i,j}|))_{\substack{1 \leq i \leq n \\ 1 \leq j \leq m}}.$$

Pour toute matrice $A = ((a_{i,j}))_{1 \leq i,j \leq n}$ dans $\mathcal{M}_n(\mathbb{K})$ et pour $1 \leq p < q \leq n$, on note $A_{p,q} = ((a_{ij}))_{p \leq i,j \leq q}$ une sous-matrice principale de A .

1. Rappels sur les normes matricielles et le rayon spectral

On rappelle les résultats suivants (voir [2]).

Pour toute norme $x \mapsto \|x\|$ sur \mathbb{K}^n , l'application :

$$A \mapsto \|A\| = \sup_{\|x\|=1} \|Ax\|$$

définit une norme sur $\mathcal{M}_n(\mathbb{K})$.

La norme matricielle induite par $\|\cdot\|_\infty$ est définie par :

$$\forall A \in \mathcal{M}_n(\mathbb{K}), \quad \|A\|_\infty = \max_{1 \leq i \leq n} \sum_{j=1}^n |a_{ij}|.$$

La norme matricielle induite par $\|\cdot\|_1$ est définie par :

$$\forall A \in \mathcal{M}_n(\mathbb{K}), \quad \|A\|_1 = \max_{1 \leq j \leq n} \sum_{i=1}^n |a_{ij}| = \|^t A\|_\infty.$$

Si A est une matrice carrée d'ordre n à coefficients complexes, alors son rayon spectral est le réel :

$$\rho(A) = \max_{\lambda \in \text{Sp}(A)} |\lambda|,$$

où $\text{Sp}(A)$ désigne l'ensemble des valeurs propres complexes de A .

Pour toute norme matricielle induite par une norme vectorielle, on a :

$$\rho(A) \leq \|A\|.$$

Quelle que soit la norme choisie sur $\mathcal{M}_n(\mathbb{C})$, on a :

$$\rho(A) = \lim_{k \rightarrow +\infty} \left(\|A^k\|^{\frac{1}{k}} \right). \quad (1.1)$$

L'application ρ est continue de $\mathcal{M}_n(\mathbb{C})$ dans \mathbb{R} (voir [2], théorème 4.15).

2. Matrices positives

Définition 2.1 Une matrice A dans $\mathcal{M}_{n,m}(\mathbb{R})$ est dite *positive* [resp. *strictement positive*] et on note $A \geq 0$ [resp. $A > 0$], si tous ses coefficients sont positifs ou nuls [resp. strictement positifs].

Si A, B sont deux matrices dans $\mathcal{M}_{n,m}(\mathbb{R})$ la notation $A \geq B$ [resp. $A > B$, ou $A \leq B$, ou $A < B$] signifie que la matrice $A - B$ est positive [resp. $A - B$ est strictement positive, ou $B - A$ est positive ou $B - A$ est strictement positive].

Remarque 2.1 En considérant la matrice $A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, on voit que les conditions $A \geq 0$ et $A \neq 0$ n'entraînent pas $A > 0$.

Lemme 2.1 Si z_1, \dots, z_n sont des nombres complexes tels que :

$$\left| \sum_{k=1}^n z_k \right| = \sum_{i=1}^n |z_k|, \quad (1.2)$$

alors il existe un réel θ tels que :

$$\forall k \in \{1, \dots, n\}, \quad z_k = |z_k| e^{i\theta}.$$

Démonstration. Si tous les z_k ($1 \leq k \leq n$) sont nuls, alors le résultat est trivial avec n'importe quel réel θ .

On suppose donc qu'il existe au moins un vecteur z_k non nul et on note I l'ensemble des indices k compris entre 1 et n tels que $z_k \neq 0$. Pour tout $k \in I$, chaque z_k peut s'écrire $z_k = \rho_k e^{i\theta_k}$ avec $\rho_k = |z_k| > 0$ et $\theta_k \in]-\pi, \pi]$ et on a :

$$\begin{cases} \left| \sum_{k \in I} z_k \right|^2 = \sum_{k \in I} |z_k|^2 + \sum_{j \neq k} \rho_j \rho_k \cos(\theta_j - \theta_k), \\ \left(\sum_{k \in I} |z_k| \right)^2 = \sum_{k \in I} |z_k|^2 + \sum_{j \neq k} \rho_j \rho_k \end{cases}$$

et l'égalité (1.2) équivaut à :

$$\sum_{j, k \in I, j \neq k} \rho_j \rho_k (1 - \cos(\theta_j - \theta_k)) = 0.$$

Tous les termes de cette somme étant positifs ou nuls avec $\rho_j \rho_k > 0$, on en déduit que $\cos(\theta_j - \theta_k) = 1$ pour $j \neq k$ dans I , avec $-\pi < \theta_j - \theta_k \leq \pi$, ce qui équivaut à $\theta_j = \theta_k$. En notant θ cette valeur commune on peut prendre $\theta_k = \theta$ pour les indices k tels que $\rho_k = 0$ et on a alors $z_k = \rho_k e^{i\theta} = |z_k| e^{i\theta}$ pour tout entier k compris entre 1 et n . ■

Avec le résultat qui suit on résume quelques propriétés élémentaires des matrices positives.

Théorème 2.1

- (i) Pour toute matrice A dans $\mathcal{M}_{n,m}(\mathbb{K})$, la matrice $|A|$ est positive et $|A| = 0$ si et seulement si $A = 0$.
- (ii) Pour toute matrice A dans $\mathcal{M}_{n,m}(\mathbb{K})$ et tout scalaire λ , on a $|\lambda A| = |\lambda| |A|$.
- (iii) Pour toutes matrices A, B dans $\mathcal{M}_{n,m}(\mathbb{K})$, on a $|A + B| \leq |A| + |B|$.
- (iv) Pour toutes matrices A dans $\mathcal{M}_{n,m}(\mathbb{K})$ et B dans $\mathcal{M}_{m,r}(\mathbb{K})$, on a $|AB| \leq |A| |B|$.
- (v) Si $A \in \mathcal{M}_{n,m}(\mathbb{R})$ est positive et $B \in \mathcal{M}_{m,r}(\mathbb{R})$ strictement positive, alors $AB = 0$ entraîne $A = 0$.
- (vi) Si $A \in \mathcal{M}_{n,m}(\mathbb{R})$ est strictement positive et $x \in \mathbb{R}^m$ est positif non nul, alors Ax est strictement positif.
- (vii) Si A, B dans $\mathcal{M}_{n,m}(\mathbb{R})$ et A', B' dans $\mathcal{M}_{m,r}(\mathbb{R})$ sont telles que $0 \leq A \leq B$ et $0 \leq A' \leq B'$, alors $0 \leq AA' \leq BB'$.
- (viii) Si A, B dans $\mathcal{M}_n(\mathbb{K})$ sont telles que $|A| \leq B$, alors $|A^k| \leq |A|^k \leq B^k$ pour tout entier naturel k .
- (ix) Si $A \in \mathcal{M}_{n,m}(\mathbb{R})$ strictement positive et $B \in \mathcal{M}_{m,r}(\mathbb{K})$ sont telles que $|AB| = A|B|$, alors il existe des réels $\theta_1, \dots, \theta_r$ tels que $B = |B| \Delta$, où Δ est la matrice diagonale de termes diagonaux $e^{i\theta_1}, \dots, e^{i\theta_r}$.
- (x) Si $A \in \mathcal{M}_{n,m}(\mathbb{K})$ est telle qu'il existe un vecteur x strictement positif dans \mathbb{R}^m tel que $Ax = |A|x$, alors $A = |A|$.

Démonstration. Les points (i) et (ii) sont évidents.

Le point (iii) se déduit des inégalités $|a_{i,j} + b_{i,j}| \leq |a_{i,j}| + |b_{i,j}|$ pour $1 \leq i \leq n$ et $1 \leq j \leq m$.

Les inégalités :

$$\left| \sum_{k=1}^m a_{i,k} b_{k,j} \right| \leq \sum_{k=1}^m |a_{i,k}| |b_{k,j}|$$

pour $1 \leq i \leq n$ et $1 \leq j \leq r$, où $\left| \sum_{k=1}^m a_{i,k} b_{k,j} \right|$ est coefficient d'indice (i, j) de $|AB|$ et $\sum_{k=1}^m |a_{i,k}| |b_{k,j}|$ celui de $|A| |B|$, signifient que $|AB| \leq |A| |B|$.

L'égalité $AB = 0$ équivaut à :

$$\sum_{k=1}^m a_{i,k} b_{k,j} = 0$$

pour $1 \leq i \leq n$ et $1 \leq j \leq r$, et pour A, B positives cela équivaut à $a_{i,k} b_{k,j} = 0$ pour tout k compris entre 1 et m , équivalent à $a_{i,k} = 0$ si de plus B est strictement positive, c'est-à-dire que $A = 0$.

Si x est un vecteur positif non nul alors il existe un entier k compris entre 1 et m tel que $x_k > 0$ et pour tout i compris entre 1 et n , on a :

$$(Ax)_i = \sum_{j=1}^m a_{i,j} x_j \geq a_{i,k} x_k > 0$$

si la matrice A est strictement positive. On a donc $Ax > 0$ si $0 < A$, $0 \leq x$ et $x \neq 0$.

Si les matrices A et A' sont positives alors la matrice AA' est également positive et en écrivant que :

$$BB' - AA' = B(B' - A') + (B - A)A',$$

on déduit que si $A \leq B$, $A' \leq B'$, avec A', B positives, alors $AA' \leq BB'$.

L'assertion (viii) se montre facilement par récurrence avec (iv) et (vii).

L'égalité $|AB| = A|B|$ avec A strictement positive équivaut à :

$$\left| \sum_{k=1}^m a_{pk} b_{kq} \right| = \sum_{k=1}^m a_{pk} |b_{kq}|$$

pour tout p compris entre 1 et n et q compris entre 1 et r . Pour p, q fixés, la suite $(z_k)_{1 \leq k \leq m}$ de nombres complexes définie par $z_k = a_{pk} b_{kq}$ est donc telle que $\left| \sum_{k=1}^m z_k \right| = \sum_{k=1}^m |z_k|$, ce qui équivaut à l'existence d'un réel θ_{pq} tel que $z_k = e^{i\theta_{pq}} |z_k|$ pour tout k compris entre 1 et m (lemme 1.1). On a donc :

$$a_{pk} b_{kq} = e^{i\theta_{pq}} |a_{pk} b_{kq}| = a_{pk} e^{i\theta_{pq}} |b_{kq}|$$

ce qui avec $a_{pk} > 0$ équivaut à :

$$b_{kq} = e^{i\theta_{pq}} |b_{kq}|.$$

En fixant p et en notant θ_q pour θ_{pq} , cela s'écrit $B = |B| \Delta$, où Δ est la matrice diagonale de termes diagonaux $e^{i\theta_1}, \dots, e^{i\theta_r}$.

Soient $A = ((a_{i,j}))_{\substack{1 \leq i \leq n \\ 1 \leq j \leq m}} \in \mathcal{M}_n(\mathbb{K})$ et $x \in \mathbb{R}^m$ tels que $Ax = |A|x$. On a alors pour tout i compris entre 1 et n :

$$\sum_{j=1}^m (|a_{i,j}| - a_{i,j}) x_j = 0$$

et si de plus x est strictement positif, cela équivaut à :

$$|a_{i,j}| = \Re(a_{i,j}), \quad \Im(a_{i,j}) = 0,$$

c'est-à-dire à $A = |A|$. ■

Du point (viii) du lemme précédent, on déduit le résultat suivant.

Théorème 2.2 *Si A, B dans $\mathcal{M}_n(\mathbb{C})$ sont telles que $|B| \leq A$, alors :*

$$\rho(B) \leq \rho(|B|) \leq \rho(A).$$

Démonstration. Pour $|B| \leq A$, on a $|B|^k \leq |B|^k \leq A^k$ pour tout $k \geq 1$ et en conséquence :

$$\|B^k\|_\infty \leq \left\| |B|^k \right\|_\infty \leq \|A^k\|_\infty,$$

soit avec la croissance de $t \mapsto t^{\frac{1}{k}}$ sur \mathbb{R}^+ :

$$\forall k \geq 1, \quad \|B^k\|_\infty^{\frac{1}{k}} \leq \left\| |B|^k \right\|_\infty^{\frac{1}{k}} \leq \|A^k\|_\infty^{\frac{1}{k}}.$$

Il suffit alors de faire tendre k vers l'infini en utilisant l'égalité (1.1). ■

Corollaire 2.1 *Si A, B dans $\mathcal{M}_n(\mathbb{R})$ sont telles que $0 \leq A \leq B$, alors :*

$$\rho(A) \leq \rho(B).$$

Corollaire 2.2 Si A est une matrice réelle positive d'ordre n , alors pour toute sous-matrice principale $A_{p,q} = ((a_{ij}))_{p \leq i, j \leq q}$ de A , on a $\rho(A_{p,q}) \leq \rho(A)$.
En particulier, on a :

$$\max_{1 \leq i \leq n} a_{ii} \leq \rho(A).$$

Démonstration. On a :

$$0 \leq B = \begin{pmatrix} 0 & 0 & 0 \\ 0 & A_{p,q} & 0 \\ 0 & 0 & 0 \end{pmatrix} \leq A$$

et donc $\rho(B) = \rho(A_{p,q}) \leq \rho(A)$. ■

Remarque 2.2 Les inégalités $a_{ii} \leq \rho(A)$ ne sont pas nécessairement vérifiées si la matrice A n'est pas positive. Par exemple toute matrice $A = \begin{pmatrix} \alpha & \beta \\ \gamma & -\alpha \end{pmatrix}$ semblable à $J = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ a un rayon spectral nul et une telle matrice s'écrit, en notant $P = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ une matrice inversible :

$$A = P^{-1}JP = \frac{1}{ad - bc} \begin{pmatrix} dc & d^2 \\ -c^2 & -dc \end{pmatrix}.$$

En prenant $dc > 0$, on a $a_{11} > \rho(A) = 0$.

Corollaire 2.3 Si A, B dans $\mathcal{M}_n(\mathbb{R})$ sont telles que $0 \leq A < B$, alors :

$$\rho(A) < \rho(B).$$

Démonstration. Si $\rho(A) = 0$, alors $\rho(B) \geq \max_{1 \leq i \leq n} b_{ii} > 0$ du fait que B est strictement positive.

Pour tous i, j compris entre 1 et n , on a $0 \leq \frac{a_{ij}}{b_{ij}} < 1$, on peut donc trouver un réel $\lambda > 0$ tel que $\frac{a_{ij}}{b_{ij}} < \lambda < 1$, soit $\frac{1}{\lambda}a_{ij} < b_{ij}$. On a donc $0 \leq \frac{1}{\lambda}A < B$ et $\rho\left(\frac{1}{\lambda}A\right) \leq \rho(B)$ ce qui entraîne pour $\rho(A) > 0$:

$$\rho(A) < \frac{1}{\lambda}\rho(A) = \rho\left(\frac{1}{\lambda}A\right) \leq \rho(B),$$

et donc $\rho(A) < \rho(B)$. ■

Théorème 2.3 Soit A une matrice positive dans $\mathcal{M}_n(\mathbb{R})$ telle que la somme des termes de chaque ligne [resp. colonne] est constante égale à α . Le réel α est alors une valeur propre de A et :

$$\rho(A) = \alpha = \|A\|_\infty \quad [\text{resp. } \rho(A) = \alpha = \|A\|_1]$$

Démonstration. Les égalités $\sum_{j=1}^n a_{ij} = \alpha$ pour tout i compris entre 1 et n reviennent à dire

que le vecteur $e = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \in \mathbb{R}^n$ est vecteur propre de la matrice A associé à la valeur propre

α . La matrice A et le réel α étant positifs, on a alors :

$$\alpha \leq \rho(A) \leq \|A\|_\infty = \max_{1 \leq i \leq n} \sum_{j=1}^n a_{ij} = \alpha,$$

soit $\alpha = \rho(A) = \|A\|_\infty$.

En raisonnant avec la transposée de la matrice A et en utilisant le fait qu'une matrice et sa transposée ont mêmes valeurs propres, on obtient le deuxième résultat en considérant que $\|{}^t A\|_\infty = \|A\|_1$. ■

Corollaire 2.4 *Pour toute matrice positive A dans $\mathcal{M}_n(\mathbb{R})$, on a :*

$$\inf_{1 \leq i \leq n} \left(\sum_{j=1}^n a_{ij} \right) \leq \rho(A) \leq \sup_{1 \leq i \leq n} \left(\sum_{j=1}^n a_{ij} \right),$$

$$\inf_{1 \leq j \leq n} \left(\sum_{i=1}^n a_{ij} \right) \leq \rho(A) \leq \sup_{1 \leq j \leq n} \left(\sum_{i=1}^n a_{ij} \right).$$

Démonstration. Pour tout i compris entre 1 et n , on note $\alpha_i = \sum_{j=1}^n a_{ij}$ et :

$$\alpha = \inf_{1 \leq i \leq n} (\alpha_i), \quad \beta = \sup_{1 \leq i \leq n} (\alpha_i).$$

On montre tout d'abord qu'on peut construire une matrice $B = ((b_{i,j}))_{1 \leq i,j \leq n}$ dans $\mathcal{M}_n(\mathbb{R})$ telle que $0 \leq B \leq A$ et :

$$\forall i \in \{1, \dots, n\}, \quad \sum_{j=1}^n b_{ij} = \alpha.$$

Si $\alpha = 0$, en prenant $B = 0$, on a bien $0 \leq B \leq A$ et $\sum_{j=1}^n b_{ij} = 0 = \alpha$ pour tout i compris entre 1 et n .

Si $\alpha = \inf_{1 \leq i \leq n} (\alpha_i) > 0$, alors tous les α_i sont strictement positifs et en posant :

$$b_{ij} = \frac{\alpha}{\alpha_i} a_{ij} \quad (1 \leq i, j \leq n)$$

on a $0 \leq b_{ij} \leq a_{ij}$ pour tous i, j compris entre 1 et n , soit $0 \leq B \leq A$ et pour tout i compris entre 1 et n :

$$\sum_{j=1}^n b_{ij} = \frac{\alpha}{\alpha_i} \sum_{j=1}^n a_{ij} = \alpha.$$

On en déduit alors que

$$\alpha = \rho(B) \leq \rho(A) \leq \|A\|_\infty = \beta.$$

En raisonnant avec ${}^t A$, considérant que $\rho({}^t A) = \rho(A)$ et $\|{}^t A\|_\infty = \|A\|_1$, on obtient le deuxième encadrement. ■

Corollaire 2.5 *Si A est une matrice positive dans $\mathcal{M}_n(\mathbb{R})$ telle $\sum_{j=1}^n a_{ij} > 0$ [resp. $\sum_{i=1}^n a_{ij} > 0$] pour tout i [resp. j] compris entre 1 et n , alors $\rho(A) > 0$. En particulier une matrice strictement positive a son rayon spectral strictement positif.*

Remarque 2.3 *Si $\rho(A) = 0$, alors toutes les valeurs propres de A sont nulles et la trace de A qui est la somme des valeurs propres est également nulle. Pour A strictement positive, on a donc $\text{Tr}(A) = \sum_{i=1}^n a_{ii} > 0$ et $\rho(A) > 0$.*

Corollaire 2.6 Si A est une matrice positive dans $\mathcal{M}_n(\mathbb{R})$ telle que A^k soit strictement positive pour un entier $k \geq 1$, alors $\rho(A) > 0$.

Démonstration. Du théorème de trigonalisation des matrices complexes, on déduit que $\rho(A^k) = (\rho(A))^k$, puis $A^k > 0$ donne $\rho(A^k) > 0$ et $\rho(A) > 0$.

■

Théorème 1.4 Pour toute matrice positive A dans $\mathcal{M}_n(\mathbb{R})$ et tout vecteur x strictement positif dans \mathbb{R}^n , on a :

$$\inf_{1 \leq i \leq n} \frac{(Ax)_i}{x_i} \leq \rho(A) \leq \sup_{1 \leq i \leq n} \frac{(Ax)_i}{x_i}.$$

Démonstration. On note :

$$r(A, x) = \inf_{1 \leq i \leq n} \frac{(Ax)_i}{x_i}, \quad s(A, x) = \sup_{1 \leq i \leq n} \frac{(Ax)_i}{x_i},$$

On désigne par D_x la matrice diagonale de termes diagonaux x_1, \dots, x_n et on utilise le corollaire 1.4 avec la matrice positive $D_x^{-1}AD_x$. La multiplication à droite par D_x a pour effet de multiplier chaque colonne j de la matrice A par x_j et la multiplication à gauche par D_x^{-1} a pour effet de diviser chaque ligne i de la matrice A par x_i , de sorte que :

$$D_x^{-1}AD_x = \left(\frac{x_j}{x_i} a_{ij} \right),$$

ce qui entraîne :

$$r(A, x) = \inf_{1 \leq i \leq n} \sum_{j=1}^n \frac{x_j}{x_i} a_{ij} \leq \rho(D_x^{-1}AD_x) = \rho(A) \leq \sup_{1 \leq i \leq n} \sum_{j=1}^n \frac{x_j}{x_i} a_{ij} = s(A, x).$$

■

Remarque 2.4 On a aussi :

$$\inf_{1 \leq j \leq n} x_j \sum_{i=1}^n \frac{a_{ij}}{x_i} \leq \rho(A) \leq \sup_{1 \leq j \leq n} x_j \sum_{i=1}^n \frac{a_{ij}}{x_i}.$$

Corollaire 2.7 Soit A une matrice positive dans $\mathcal{M}_n(\mathbb{R})$. Si il existe un vecteur x strictement positif et deux constantes réelles positives ou nulles α, β telles que $\alpha x \leq Ax \leq \beta x$ [resp. $\alpha x < Ax < \beta x$], alors $\alpha \leq \rho(A) \leq \beta$ [Resp. $\alpha < \rho(A) < \beta$].

Démonstration. L'inégalité $\alpha x \leq Ax \leq \beta x$ équivaut à $\alpha x_i \leq (Ax)_i \leq \beta x_i$ pour tout i compris entre 1 et n , ce qui entraîne :

$$\alpha \leq \inf_{1 \leq i \leq n} \frac{(Ax)_i}{x_i} \leq \rho(A) \leq \sup_{1 \leq i \leq n} \frac{(Ax)_i}{x_i} \leq \beta.$$

On procède de même pour les inégalités strictes.

■

Corollaire 2.8 Soit A une matrice positive dans $\mathcal{M}_n(\mathbb{R})$. Si la matrice A admet un vecteur propre strictement positif, alors la valeur propre associée est $\rho(A)$ et :

$$\begin{aligned}\rho(A) &= \sup_{x>0} \inf_{1 \leq i \leq n} \frac{(Ax)_i}{x_i} \\ &= \inf_{x>0} \sup_{1 \leq i \leq n} \frac{(Ax)_i}{x_i}.\end{aligned}$$

Démonstration. On note, pour tout vecteur x strictement positif :

$$r(A, x) = \inf_{1 \leq i \leq n} \frac{(Ax)_i}{x_i}, \quad s(A, x) = \sup_{1 \leq i \leq n} \frac{(Ax)_i}{x_i},$$

et on a $r(A, x) \leq \rho(A) \leq s(A, x)$ (théorème 1.4), ce qui entraîne l'existence de $\sup_{x>0} \inf_{1 \leq i \leq n} \frac{(Ax)_i}{x_i}$

et $\inf_{x>0} \sup_{1 \leq i \leq n} \frac{(Ax)_i}{x_i}$.

Si, de plus x est un vecteur propre strictement positif de A , alors la valeur propre α associée est réelle positive et avec :

$$\alpha = \inf_{1 \leq i \leq n} \frac{(Ax)_i}{x_i} \leq \rho(A) \leq \sup_{1 \leq i \leq n} \frac{(Ax)_i}{x_i} = \alpha,$$

on déduit que $\alpha = \rho(A) = \sup_{x>0} \inf_{1 \leq i \leq n} \frac{(Ax)_i}{x_i} = \inf_{x>0} \sup_{1 \leq i \leq n} \frac{(Ax)_i}{x_i}$. ■

3. Matrices strictement positives et théorème de Perron-Frobénius

Lemme 3.1 Soit A une matrice strictement positive dans $\mathcal{M}_n(\mathbb{R})$. Si x est un vecteur propre non nul de A associé à une valeur propre λ telle que $|\lambda| = \rho(A)$, alors $\rho(A)$ est valeur propre de A avec $|x|$ comme vecteur propre associé. Le vecteur $|x|$ est strictement positif et il existe un réel θ tel que $x = e^{i\theta} |x|$.

Démonstration. On a $\rho(A) > 0$ du fait que $A > 0$.

De $Ax = \lambda x$ avec $|\lambda| = \rho(A)$, on déduit que :

$$\rho(A) |x| = |\lambda x| = |Ax| \leq |A| |x| = A |x|, \tag{1.3}$$

ce qui entraîne que le vecteur $y = A |x| - \rho(A) |x|$ est positif. Si ce vecteur est non nul, alors $Ay > 0$ (point (vi) du théorème 1.1), ce qui signifie en notant $z = A |x|$ que $\rho(A) z < Az$ avec $z > 0$ (le vecteur x est non nul) qui entraîne $\rho(A) < \rho(A)$ (corollaire 1.7), soit une impossibilité. On a donc $y = 0$, c'est-à-dire $A |x| = \rho(A) |x|$, ce qui signifie que $\rho(A)$ est valeur propre de A avec $|x|$ comme vecteur propre associé. De plus avec $|x| = \frac{1}{\rho(A)} A |x|$, on déduit que $|x| > 0$.

Enfin de (1.3), on déduit que $A |x| = |Ax|$ et donc qu'il existe un réel θ tel que $x = e^{i\theta} |x|$ (point (ix) du théorème 1.1 avec $B = x$). ■

Théorème 3.1 (Perron-Frobenius I) Si A est une matrice strictement positive dans $\mathcal{M}_n(\mathbb{R})$ alors $\rho(A)$ est l'unique valeur propre de A de module maximum et l'espace propre associé est une droite vectorielle engendrée par un vecteur strictement positif.

Démonstration. Dans la démonstration du lemme précédent on a vu que si λ est une valeur propre de la matrice A telle que $|\lambda| = \rho(A)$ et si x est un vecteur propre non nul associé, alors $x = e^{i\theta} |x|$ avec $A|x| = \rho(A)|x|$. Le rayon spectral $\rho(A)$ est donc valeur propre de A . De plus, avec :

$$\lambda x = Ax = A(e^{i\theta} |x|) = e^{i\theta} A|x| = e^{i\theta} \rho(A)|x| = \rho(A)x,$$

on déduit que $\lambda x = \rho(A)x$ avec $x \neq 0$, et $\lambda = \rho(A)$. Donc $\rho(A)$ est l'unique valeur propre de A de module maximal.

En notant $E_{\rho(A)}$ l'espace propre associé à la valeur propre $\rho(A)$, tout vecteur non nul x dans $E_{\rho(A)}$ est tel que $|x| > 0$ et aucune des composantes de x n'est nulle. S'il existe deux vecteurs x, y linéairement indépendants dans $E_{\rho(A)}$, alors le vecteur $z = x_1 y - y_1 x$ est non nul dans $E_{\rho(A)}$ avec $z_1 = 0$, ce qui est impossible. On a donc $\dim(E_{\rho(A)}) = 1$. ■

Corollaire 3.1 *Si A est une matrice strictement positive dans $\mathcal{M}_n(\mathbb{R})$ alors il existe une unique vecteur propre associé à la valeur propre $\rho(A)$ dans le compact $F = \{x \in \mathbb{R}^n \mid x \geq 0, \|x\|_1 = 1\}$.*

Démonstration. On a vu que $E_{\rho(A)} = \mathbb{R}x$ avec $x > 0$. Le vecteur $v = \frac{1}{\|x\|_1}x$ est alors l'unique élément de $F \cap E_{\rho(A)}$. ■

Le vecteur $x \in F \cap E_{\rho(A)}$ est appelé, le vecteur de Perron de la matrice strictement positive A .

En utilisant la densité de l'ensemble des matrices strictement positives dans l'ensemble des matrices positives, on déduit le résultat suivant.

Corollaire 3.2 *Si A est une matrice positive dans $\mathcal{M}_n(\mathbb{R})$ alors $\rho(A)$ est valeur propre de A et il existe un vecteur propre associé non nul positif.*

Démonstration. Pour tout entier naturel non nul k , on pose :

$$A_k = \left(\left(a_{ij} + \frac{1}{k} \right) \right)_{1 \leq i, j \leq n}$$

et on désigne par x_k le vecteur de Perron de la matrice strictement positive A_k . On a $\lim_{k \rightarrow +\infty} A_k = A$ et avec la continuité du rayon spectral, on déduit que $\lim_{k \rightarrow +\infty} \rho(A_k) = \rho(A)$. D'autre part, la suite $(x_k)_{k \geq 1}$ étant dans le compact F , on peut en extraire une sous suite $(x_{\varphi(k)})_{k \geq 1}$ convergente vers un vecteur $x \geq 0$ et on a :

$$Ax = \lim_{k \rightarrow +\infty} A_{\varphi(k)} x_{\varphi(k)} = \lim_{k \rightarrow +\infty} \rho(A_{\varphi(k)}) x_{\varphi(k)} = \rho(A)x,$$

c'est-à-dire que x est un vecteur propre non nul (puisque $\|x\|_1 = 1$) positif de A associé à la valeur propre $\rho(A)$. ■

Corollaire 3.3 *Si A est une matrice positive dans $\mathcal{M}_n(\mathbb{R})$ alors :*

$$\rho(I_n + A) = 1 + \rho(A).$$

Démonstration. Pour toute matrice A dans $\mathcal{M}_n(\mathbb{K})$, on a :

$$\text{Sp}(I_n + A) = \{1 + \lambda \mid \lambda \in \text{Sp}(A)\}$$

et donc $\rho(I_n + A) \leq 1 + \rho(A)$.

Si de plus A est positive, alors $\rho(A)$ est valeur propre de A , donc $1 + \rho(A)$ est valeur propre de $I_n + A$ et $1 + \rho(A) \leq \rho(I_n + A)$, d'où l'égalité. ■

Le théorème de Perron-Frobenius associé au théorème de Gerschgorin et Hadamard (voir [2], théorème 2.35), nous permet d'obtenir le résultat suivant de localisation des valeurs propres d'une matrice.

Pour tout nombre complexe a et tout réel positif r , on note :

$$D(a, r) = \{z \in \mathbb{C} \mid |z - a| \leq r\}$$

le disque fermé de centre a et de rayon r .

Corollaire 3.4 Soient A une matrice complexe dans $\mathcal{M}_n(\mathbb{C})$ et B une matrice positive dans $\mathcal{M}_n(\mathbb{R})$ telle que $|A| \leq B$. Toutes les valeurs propres de A sont dans la réunion des disques de centre a_{ii} et de rayon $\rho(B) - b_{ii}$, soit :

$$\text{Sp}(A) \subset \bigcup_{i=1}^n D(a_{ii}, \rho(B) - b_{ii}).$$

Démonstration. Supposons dans un premier temps que la matrice B soit strictement positive.

Si x est le vecteur de Perron de B , alors pour tout entier i compris entre 1 et n , on a :

$$\sum_{\substack{j=1 \\ j \neq i}}^n |a_{ij}| x_j \leq \sum_{\substack{j=1 \\ j \neq i}}^n b_{ij} x_j = (\rho(B) - b_{ii}) x_i,$$

soit :

$$\frac{1}{x_i} \sum_{\substack{j=1 \\ j \neq i}}^n |a_{ij}| x_j \leq \rho(B) - b_{ii}.$$

D'autre part, le théorème de Gerschgorin-Hadamard nous dit que si $\lambda \in \mathbb{C}$ est une valeur propre de la matrice C , alors il existe un indice i compris entre 1 et n tel que $|\lambda - c_{ii}| \leq \sum_{j=1, j \neq i}^n |c_{ij}|$.

La matrice A ayant les mêmes valeurs propres que la matrice $C_x = D_x^{-1} A D_x$, où D_x est la matrice diagonale de termes diagonaux x_1, \dots, x_n (c'est-à-dire que $C_x = \begin{pmatrix} x_j & \\ & a_{ij} \\ & & x_i \end{pmatrix}$), on déduit que pour toute valeur propre $\lambda \in \mathbb{C}$ de A , il existe un indice i tel que :

$$|\lambda - a_{ii}| = |\lambda - c_{ii}| \leq \sum_{\substack{j=1 \\ j \neq i}}^n |c_{ij}| = \frac{1}{x_i} \sum_{\substack{j=1 \\ j \neq i}}^n |a_{ij}| x_j \leq \rho(B) - b_{ii},$$

ce qui donne le résultat annoncé.

Si la matrice B est positive, non strictement positive, on désigne pour tout réel $\varepsilon > 0$ par B_ε la matrice $B_\varepsilon = ((b_{ij} + \varepsilon))$ et on a $B_\varepsilon > A$, donc :

$$\text{Sp}(A) \subset \bigcup_{i=1}^n D(a_{ii}, \rho(B_\varepsilon) - b_{ii} - \varepsilon).$$

Avec la continuité du rayon spectral, on en déduit alors le résultat. ■

Du théorème de Perron-Frobenius on déduit également le résultat suivant.

Corollaire 3.5 Une matrice strictement positive dans $\mathcal{M}_n(\mathbb{R})$ ne peut avoir deux vecteurs propres positifs linéairement indépendants.

Démonstration. Soit y un vecteur propre positif non nul associé à une valeur propre λ de A .

On a $Ay > 0$ et il existe un indice i compris entre 1 et n tel que $(Ay)_i = \lambda y_i$ avec $y_i > 0$, ce qui entraîne que $\lambda > 0$ et $y = \frac{1}{\lambda}Ay > 0$.

Du corollaire 1.8, on déduit que nécessairement $\lambda = \rho(A)$ et le théorème de Perron-Frobenius nous permet alors de conclure. ■

Lemme 3.2 Soit A une matrice strictement positive dans $\mathcal{M}_n(\mathbb{R})$. Si x est un vecteur propre strictement positif de A associé à la valeur propre $\rho(A)$, alors il existe alors un unique vecteur y strictement positif qui est vecteur propre de tA associé à la valeur propre $\rho({}^tA) = \rho(A)$ tel que ${}^tyx = 1$.

Démonstration. La matrice tA est strictement positive avec $\rho({}^tA) = \rho(A)$, l'espace propre associé à la valeur propre $\rho(A)$ de tA est donc une droite vectorielle dirigée par un vecteur $z > 0$. En posant $y = \frac{1}{{}^tzx}z$, on a $y > 0$, ${}^tAy = \rho(A)y$ et ${}^tyx = 1$.

Réciproquement, si ${}^tAy = \rho(A)y$ avec $y > 0$, ${}^tyx = 1$, alors $y = \alpha z$ avec $\alpha = \frac{1}{{}^tzx}$, ce qui prouve l'unicité du vecteur y . ■

Avec les notations et hypothèses du lemme précédent, on désigne par L la matrice d'ordre n définie par :

$$L = x{}^ty = ((x_i y_j))_{1 \leq i, j \leq n}.$$

On désigne par $E_{\rho(A)}$ l'espace propre associé à la valeur propre $\rho(A)$ de A (matrice strictement positive) et par $E'_{\rho(A)}$ l'espace propre associé à la valeur propre $\rho({}^tA) = \rho(A)$ de tA .

Avec le lemme qui suit, on résume quelques propriétés de cette matrice L .

- Lemme 3.3** (i) La matrice L est indépendante du choix du vecteur x .
(ii) La matrice L est strictement positive de rang 1.
(iii) $Lx = x$, ${}^tLy = y$.
(iv) Pour tout entier naturel k non nul, on a :

$$L^k = L, A^k L = L A^k = \rho(A)^k L, \left(\frac{1}{\rho(A)} A - L \right)^k = \left(\frac{1}{\rho(A)} A \right)^k - L.$$

- (v) L est la matrice de la projection sur la droite vectorielle $E_{\rho(A)}$ parallèlement à l'hyperplan :

$$H = \{z \in \mathbb{C}^n \mid {}^tyz = 0\}.$$

Démonstration. (i) On a $E_{\rho(A)} = \mathbb{C}x$, $E'_{\rho(A)} = \mathbb{C}y$, avec $x > 0$, $y > 0$ et ${}^tyx = 1$.

Si (x', y') est un autre couple de vecteurs vérifiant ces propriétés, alors $x' = \alpha x$, $y' = \beta y$, avec $\alpha > 0$, $\beta > 0$, $1 = {}^ty'x' = \alpha\beta{}^tyx = \alpha\beta$ et :

$$x'{}^ty' = \alpha\beta x{}^ty = x{}^ty = L.$$

(ii) Les vecteurs x et y étant strictement positifs, il en est de même de la matrice L . De plus, pour j compris entre 1 et n , la colonne numéro j de L est $y_j x > 0$, il en résulte que L est de rang 1 avec $\text{Im}(L) = \mathbb{C}x = E_{\rho(A)}$.

(iii) Avec l'associativité du produit matriciel, on peut écrire :

$$\begin{aligned} Lx &= x \begin{pmatrix} t y x \end{pmatrix} = \begin{pmatrix} t y x \end{pmatrix} x = x \\ {}^t Ly &= y \begin{pmatrix} t x y \end{pmatrix} = \begin{pmatrix} t x y \end{pmatrix} y = {}^t \begin{pmatrix} t y x \end{pmatrix} y = y. \end{aligned}$$

(iv) On a $L^2 = x \begin{pmatrix} t y x \end{pmatrix} {}^t y = \begin{pmatrix} t y x \end{pmatrix} (x {}^t y) = L$ et par récurrence $L^k = L$ pour tout $k \geq 1$. De $Ax = \rho(A)x$, on déduit que $A^k x = \rho(A)^k x$ pour tout $k \geq 1$ et :

$$A^k L = A^k x {}^t y = \rho(A)^k x {}^t y = \rho(A)^k L.$$

De même avec ${}^t Ay = \rho(A)y$, on obtient :

$$LA^k = x {}^t y A^k = x {}^t \left(({}^t A)^k y \right) = x {}^t \left(\rho(A)^k y \right) = \rho(A)^k L.$$

La dernière relation peut se montrer par récurrence sur $k \geq 1$. Le résultat est évident pour $k = 1$ et en le supposant acquis pour $k \geq 1$, on a :

$$\begin{aligned} \left(\frac{1}{\rho(A)} A - L \right)^{k+1} &= \left(\frac{1}{\rho(A)} A - L \right) \left(\left(\frac{1}{\rho(A)} A \right)^k - L \right) \\ &= \left(\frac{1}{\rho(A)} A \right)^{k+1} - \frac{1}{\rho(A)^k} LA^k - \frac{1}{\rho(A)} AL + L^2 \\ &= \left(\frac{1}{\rho(A)} A \right)^{k+1} - \frac{1}{\rho(A)^k} \rho(A)^k L - \frac{1}{\rho(A)} \rho(A) L + L \\ &= \left(\frac{1}{\rho(A)} A \right)^{k+1} - L. \end{aligned}$$

(v) On a $L^2 = L$, $\text{Im}(L) = E_{\rho(A)}$ et le noyau de L est donné par :

$$\text{Ker}(L) = \{z \in \mathbb{C}^n \mid x {}^t y z = 0\} = \{z \in \mathbb{C}^n \mid ({}^t y z) x = 0\} = H,$$

c'est-à-dire que L est la projection sur $E_{\rho(A)}$ parallèlement à H . ■

Théorème 3.2 Avec les notations et hypothèses du lemme 1.3, on a :

$$\lim_{k \rightarrow +\infty} \left(\frac{1}{\rho(A)} A \right)^k = L = x {}^t y.$$

Démonstration. Avec $\left(\frac{1}{\rho(A)} A \right)^k - L = \left(\frac{1}{\rho(A)} A - L \right)^k$ pour $k \geq 1$, il suffit de montrer que la matrice $\frac{1}{\rho(A)} A - L$ a un rayon spectral strictement inférieur à 1, ce qui revient à dire que $\rho(A - \rho(A)L) < \rho(A)$.

Si λ est une valeur propre non nulle de $B = A - \rho(A)L$ et z est un vecteur propre non nul associé, alors $(A - \rho(A)L)z = \lambda z$ et avec $L(A - \rho(A)L) = 0$, on déduit que $\lambda Lz = 0$ et $Lz = 0$ pour $\lambda \neq 0$, ce qui entraîne :

$$\lambda z = Az - \rho(A)Lz = Az,$$

c'est-à-dire que λ est valeur propre de A avec z pour vecteur propre associé. On a donc montré que toute valeur propre λ non nulle de B est aussi valeur propre de A et donc $|\lambda| \leq \rho(A)$.

Si λ est une valeur propre non nulle de B telle que $|\lambda| = \rho(A)$, alors $\lambda = \rho(A)$ (Perron-Frobénius) et tout vecteur propre associé z est aussi vecteur propre de A , on a donc $z = \alpha x$ avec $\alpha \in \mathbb{C}^*$ pour $z \neq 0$, ce qui entraîne :

$$\begin{aligned} \rho(A) z &= (A - \rho(A)L) z = (A - \rho(A)L) \alpha x \\ &= \alpha Ax - \alpha \rho(A)Lx = \alpha \rho(A)x - \alpha \rho(A)x = 0 \end{aligned}$$

en contradiction avec $\rho(A) > 0$ et $z \neq 0$.

On a donc $|\lambda| < \rho(A)$ pour toute valeur propre de B et $\rho(B) < \rho(A)$. D'où le résultat. ■

Remarque 3.1 Avec ce résultat, on retrouve le fait que la matrice L ne dépend que de la matrice strictement positive A et pas du choix de x .

On est maintenant en mesure de préciser le théorème de Perron-Frobénius.

Théorème 3.3 (Perron-Frobénius II) Si A est une matrice strictement positive dans $\mathcal{M}_n(\mathbb{R})$ alors $\rho(A)$ est l'unique valeur propre de A de module maximum et cette valeur propre est simple (l'espace propre associé est donc une droite vectorielle).

Démonstration. On sait déjà que $\rho(A)$ est l'unique valeur propre de A de module maximum. Notons p sa multiplicité algébrique.

Le théorème de trigonalisation sur \mathbb{C} nous dit qu'il existe une matrice inversible P telle que $T = P^{-1}AP$ soit triangulaire supérieure de diagonale :

$$(\rho(A), \dots, \rho(A), \lambda_{p+1}, \dots, \lambda_n)$$

avec $|\lambda_i| < \rho(A)$ pour i compris entre $p+1$ et n (si $p < n$). En écrivant, pour tout entier naturel non nul k , que :

$$\left(\frac{1}{\rho(A)}T\right)^k = P^{-1} \left(\frac{1}{\rho(A)}A\right)^k P$$

et en utilisant la continuité du produit matriciel, on déduit du théorème précédent que :

$$\lim_{k \rightarrow +\infty} \left(\frac{1}{\rho(A)}T\right)^k = P^{-1}LP = L',$$

avec L' triangulaire supérieure de diagonale $(1, \dots, 1, 0, \dots, 0)$.

On a donc $\text{rang}(L') \geq p$ et avec $\text{rang}(L') = \text{rang}(L) = 1$, on déduit que nécessairement $p = 1$. ■

Si la matrice A est positive, on a vu que $\rho(A)$ est valeur propre de A , mais cette valeur propre n'est pas nécessairement simple (prendre par exemple la matrice identité). Mais s'il existe un entier naturel r tel que A^r soit strictement positive, alors $\rho(A)$ est valeur propre simple de A .

Corollaire 3.6 Si A est une matrice positive dans $\mathcal{M}_n(\mathbb{R})$ et s'il existe un entier naturel r tel que A^r soit strictement positive, alors $\rho(A)$ est valeur propre simple de A (l'espace propre associé est donc une droite vectorielle).

Démonstration. On sait déjà que $\rho(A)$ est valeur propre de A (corollaire 1.10). Notons p sa multiplicité algébrique.

Le théorème de trigonalisation sur \mathbb{C} nous dit qu'il existe une matrice inversible P telle que $T = P^{-1}AP$ soit triangulaire supérieure de diagonale :

$$(\rho(A), \dots, \rho(A), \lambda_{p+1}, \dots, \lambda_n)$$

avec $|\lambda_i| \leq \rho(A)$ pour tout entier i compris entre $p+1$ et n (si $p < n$). La matrice $T^p = P^{-1}A^pP$ est alors triangulaire supérieure de diagonale :

$$(\rho(A)^p, \dots, \rho(A)^p, \lambda_{p+1}^p, \dots, \lambda_n^p)$$

et $\rho(A)^p = \rho(A^p)$ est alors valeur propre de A^p de multiplicité supérieure ou égale à p . Mais A^p étant strictement positive cette multiplicité vaut 1, on a donc $p = 1$. ■

Une matrice vérifiant les hypothèses du corollaire est un cas particulier de matrice positive irréductible.

4. Matrices de permutation

Si n est un entier naturel supérieur ou égal à 2, on désigne par \mathfrak{S}_n le groupe des permutations de l'ensemble $\{1, \dots, n\}$ et par $\mathcal{B} = \{e_1, \dots, e_n\}$ la base canonique de \mathbb{K}^n ($\mathbb{K} = \mathbb{R}$ ou \mathbb{C}).

Pour tout couple (i, j) d'entiers naturels, on note $\delta_{i,j}$ le symbole de Kronecker ($\delta_{ii} = 1$ et $\delta_{i,j} = 0$ pour $i \neq j$).

Définition 4.1 Si $\sigma \in \mathfrak{S}_n$, on appelle matrice de permutation associée à σ , la matrice de passage P_σ de la base canonique de \mathbb{K}^n à la base $\mathcal{B}_\sigma = \{e_{\sigma(1)}, \dots, e_{\sigma(n)}\}$.

On a donc, si P_σ est une matrice de permutation, $P_\sigma e_j = e_{\sigma(j)}$ pour tout entier j compris entre 1 et n , ce qui revient à dire que :

$$P_\sigma = ((\delta_{i,\sigma(j)}))_{1 \leq i, j \leq n}.$$

En particulier, on a $P_{Id} = I_n$.

Une telle matrice de permutation étant la matrice de passage de la base canonique de \mathbb{C}^n à une base orthonormée est unitaire et donc $P_\sigma^{-1} = {}^t P_\sigma$.

De plus il est facile de vérifier que pour toutes permutations σ, τ dans \mathfrak{S}_n , on a $P_\sigma P_\tau = P_{\sigma \circ \tau}$ (il suffit d'écrire que pour tout i compris entre 1 et n , on a $P_\sigma P_\tau e_i = P_\sigma e_{\tau(i)} = e_{\sigma \circ \tau(i)} = P_{\sigma \circ \tau} e_i$). On en déduit alors que $P_\sigma^{-1} = P_{\sigma^{-1}}$.

Pour toute permutation $\sigma \in \mathfrak{S}_n$ et tout vecteur $x = (x_i)_{1 \leq i \leq n} \in \mathbb{K}^n$, on a :

$$P_\sigma^{-1} x = \sum_{j=1}^n x_j e_{\sigma^{-1}(j)} = \sum_{i=1}^n x_{\sigma(i)} e_i$$

en faisant le changement d'indice $i = \sigma^{-1}(j)$.

C'est-à-dire que $P_\sigma^{-1} x = (x_{\sigma(i)})_{1 \leq i \leq n}$ se déduit de x en faisant agir la permutation σ sur les composantes de x .

On en déduit alors que pour toute matrice $A = ((a_{i,j}))_{1 \leq i, j \leq n} \in \mathcal{M}_n(\mathbb{K})$:

$$P_\sigma^{-1} A = ((a_{\sigma(i),j}))_{1 \leq i, j \leq n}$$

se déduit de A en faisant agir la permutation σ sur les lignes de A .

La multiplication à droite d'une matrice A par une matrice de permutation P_σ va faire agir la permutation σ sur les colonnes de A . En effet, pour tout j compris entre 1 et n , on a :

$$A P_\sigma e_j = A e_{\sigma(j)} = \sum_{i=1}^n a_{i,\sigma(j)} e_i$$

et donc $AP_\sigma = ((a_{i,\sigma(j)}))_{1 \leq i,j \leq n}$.

On a donc pour toute permutation $\sigma \in \mathfrak{S}_n$ et toute matrice $A = ((a_{i,j}))_{1 \leq i,j \leq n}$ dans $\mathcal{M}_n(\mathbb{K})$:

$$P_\sigma^{-1}AP_\sigma = ((a_{\sigma(i),\sigma(j)}))_{1 \leq i,j \leq n},$$

c'est-à-dire que $P_\sigma^{-1}AP_\sigma$ se déduit de A en faisant agir la permutation σ sur les lignes et les colonnes de A .

Remarque 4.1 Une matrice de permutation étant unitaire dans $\mathcal{M}_n(\mathbb{C})$ est diagonalisable sur \mathbb{C} . Mais sur \mathbb{R} ce résultat n'est plus valable. En effet si σ est un cycle d'ordre 3, alors P_σ a pour polynôme minimale $X^3 - 1$ qui n'est pas scindé sur \mathbb{R} et cette matrice n'est pas diagonalisable sur \mathbb{R} .

5. Matrices irréductibles

Définition 5.1 Une matrice $A \in \mathcal{M}_n(\mathbb{K})$ est dite réductible s'il existe une matrice de permutation P_σ telle que :

$$P_\sigma^{-1}AP_\sigma = \begin{pmatrix} B & 0 \\ C & D \end{pmatrix} \quad (1.4)$$

où $B \in \mathcal{M}_p(\mathbb{K})$ avec $1 \leq p \leq n-1$.

Une matrice non réductible est dite irréductible.

Exemple 5.1 Une matrice ayant tous ses coefficients non nuls est irréductible.

Exemple 5.2 Une matrice ayant une ligne (ou une colonne) nulle est réductible. En effet si la ligne numéro i est nulle en transposant la ligne 1 avec la ligne i et la colonne 1 avec la colonne i , on obtient une matrice avec la première ligne nulle.

Exemple 5.3 Une matrice $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ dans $\mathcal{M}_2(\mathbb{K})$ est irréductible si et seulement si $bc \neq 0$.

En effet $\mathfrak{S}_2 = \{\tau_{12}, I_d\}$ et $P_{I_d}^{-1}AP_{I_d} = A$ ou $P_{\tau_{12}}^{-1}AP_{\tau_{12}} = \begin{pmatrix} a & c \\ b & d \end{pmatrix}$ est de la forme (1.4) si et seulement si $b = 0$ ou $c = 0$.

Une condition suffisante d'irréductibilité est donnée par le résultat suivant.

Lemme 5.1 Soit $A \in \mathcal{M}_n(\mathbb{K})$. S'il existe un entier naturel p tel que A^p ait tous ses coefficients non nuls, alors A est irréductible.

Démonstration. Si A est réductible, il existe alors une matrice de permutation P_σ telle que $P_\sigma^{-1}AP_\sigma$ soit de la forme $\begin{pmatrix} B & 0 \\ C & D \end{pmatrix}$ et pour tout entier $p \geq 1$ on a $P_\sigma^{-1}A^pP_\sigma = \begin{pmatrix} B^p & 0 \\ C_p & D^p \end{pmatrix}$, ce qui signifie que la matrice A^p est également réductible, elle a donc au moins un coefficient nul. ■

Exemple 5.4 La matrice $A = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix}$ est irréductible puisque $A^2 > 0$.

Cette condition n'est pas nécessaire comme le montre l'exemple de la matrice $A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$.

Une définition équivalente de la notion de matrice réductible est donnée par le résultat suivant.

Théorème 5.1 *Une matrice $A \in \mathcal{M}_n(\mathbb{K})$ est réductible si et seulement si il existe une partition non triviale (I, J) de $\{1, \dots, n\}$ telle que $a_{ij} = 0$ pour tout $(i, j) \in I \times J$.*

Démonstration. Soit $A \in \mathcal{M}_n(\mathbb{K})$ réductible. Il existe une permutation σ telle que $a_{\sigma(i), \sigma(j)} = 0$ pour $1 \leq i \leq p$ et $p+1 \leq j \leq n$, où p est un entier compris entre 1 et $n-1$. En posant $I = \{\sigma(1), \dots, \sigma(p)\}$, $J = \{\sigma(p+1), \dots, \sigma(n)\}$, on définit une partition non triviale de $\{1, \dots, n\}$ telle que $a_{ij} = 0$ pour tout $(i, j) \in I \times J$.

Réciproquement supposons qu'il existe une partition non triviale (I, J) de $\{1, \dots, n\}$ telle que $a_{ij} = 0$ pour tout $(i, j) \in I \times J$. En notant p le cardinal de I , on a $1 \leq p \leq n-1$ et il existe $\sigma \in \mathfrak{S}_n$ telle que $I = \{\sigma(1), \dots, \sigma(p)\}$, $J = \{\sigma(p+1), \dots, \sigma(n)\}$, ce qui donne :

$$P_\sigma^{-1} A P_\sigma = ((a_{\sigma(i), \sigma(j)}))_{1 \leq i, j \leq n} = \begin{pmatrix} B & 0 \\ C & D \end{pmatrix},$$

c'est-à-dire que A est réductible. ■

De ce théorème, on peut déduire qu'une matrice est irréductible si et seulement si sa transposée l'est.

Une autre définition équivalente de la notion de matrice réductible est donnée par le résultat suivant.

Théorème 5.2 *Une matrice $A \in \mathcal{M}_n(\mathbb{K})$ est réductible si et seulement si il existe une partie non triviale J de $\{1, \dots, n\}$ telle que le sous-espace vectoriel de \mathbb{K}^n , $V_J = \text{Vect}\{e_j \mid j \in J\}$, soit stable par A .*

Démonstration. Dire que $A \in \mathcal{M}_n(\mathbb{K})$ est réductible équivaut à dire qu'il existe une partition non triviale (I, J) de $\{1, \dots, n\}$ telle que $a_{ij} = 0$ pour tout $(i, j) \in I \times J$, ce qui revient à dire que pour tout $j \in J$ on a :

$$Ae_j = \sum_{i=1}^n a_{ij} e_i = \sum_{i \in J} a_{ij} e_i \in V_J,$$

encore équivalent à dire que V_J est stable par A . ■

Les propriétés élémentaires suivantes nous seront utiles.

Lemme 5.2 *Si A est réductible, alors $|A|$ et $I_n + A$ sont réductibles.*

Démonstration. Cela résulte de $P_\sigma^{-1} A P_\sigma = \begin{pmatrix} B & 0 \\ C & D \end{pmatrix}$ entraîne :

$$\begin{cases} P_\sigma^{-1} |A| P_\sigma = \begin{pmatrix} |B| & 0 \\ |C| & |D| \end{pmatrix}, \\ P_\sigma^{-1} (I_n + A) P_\sigma = \begin{pmatrix} I_p + B & 0 \\ C & I_{n-p} + D \end{pmatrix}. \end{cases}$$

puisque l'action de P_σ est seulement de permuter des lignes et des colonnes. ■

On a vu que si A est une matrice strictement positive dans $\mathcal{M}_n(\mathbb{R})$ alors $\rho(A)$ est l'unique valeur propre de A de module maximum, que cette valeur propre est simple, que l'espace propre

associé est une droite vectorielle engendrée par un vecteur strictement positif et qu'une telle matrice ne peut avoir deux vecteurs propres positifs linéairement indépendants.

Ces résultats ne s'étendent pas au cas des matrices positives. Mais pour les matrices positives qui sont de plus irréductibles, on a des résultats analogues.

Lemme 5.3 *Soit $A \in \mathcal{M}_n(\mathbb{R})$ positive et irréductible. Si $y \in \mathbb{R}^n$ un vecteur positif non nul, alors soit y est strictement positif et il en est de même du vecteur $z = (I_n + A)y$, soit y a au moins une composante nulle et le nombre de coordonnées nulles de z est strictement inférieur au nombre de coordonnées nulles de y . Dans tous les cas, le vecteur $(I_n + A)^{n-1}y$ est strictement positif.*

Démonstration. Les composantes du vecteur $z = (I_n + A)y$ sont données par :

$$z_i = \sum_{j=1}^n a_{ij}y_j + y_i \quad (1 \leq i \leq n).$$

Si la matrice A et le vecteur y sont positifs, on a alors $z_i \geq y_i \geq 0$ et z est positif.

Avec $z \geq y$, on déduit que z est strictement positif si y l'est.

En supposant que y a au moins une composante nulle, de $0 \leq y_i \leq z_i$, on déduit que $z_i = 0$ entraîne $y_i = 0$. Le nombre de coordonnées nulles du vecteur z est donc inférieur ou égal à celui de y .

Supposons que z et y ont le même nombre de composantes nulles. En notant J_y l'ensemble des indices compris entre 1 et n tels que $y_i = 0$, on a $z_i > 0$ pour $i \notin J_y$ et en conséquence $z_i = y_i = 0$ pour tout $i \in J_y$, avec $z_i = \sum_{j \notin J_y} a_{ij}y_j$ et $y_j > 0$ pour $j \notin J_y$. On a donc en tenant compte du fait que les coefficients a_{ij} sont positifs ou nuls, $a_{ij} = 0$ pour $i \in J_y$ et $j \notin J_y$ avec J_y de cardinal compris entre 1 et $n - 1$ (y a au moins une composante nulle et n'est pas le vecteur nul) ce qui revient à dire que la matrice A est réductible. En conclusion le nombre de composantes nulles de z est strictement inférieur à celui de y , si la matrice positive A est irréductible.

Si le vecteur y est positif non nul, il a alors au moins une coordonnée strictement positive et ce qui précède nous dit que le vecteur $(I_n + A)y$ a au moins deux coordonnées strictement positives. Par récurrence on déduit alors que le vecteur $(I_n + A)^{n-1}y$ a au moins n coordonnées strictement positives, ce qui revient à dire qu'il est strictement positif. ■

Théorème 5.2 *Soit $A \in \mathcal{M}_n(\mathbb{R})$ positive. Cette matrice est irréductible si et seulement si la matrice $(I_n + A)^{n-1}$ est strictement positive.*

Démonstration. Du lemme 1.7, on déduit que si A est une matrice positive irréductible, alors pour tout j compris entre 1 et n le vecteur $(I_n + A)^{n-1}e_j$ est strictement positif, ce qui équivaut à dire que la matrice $(I_n + A)^{n-1}$ est strictement positive.

Réciproquement si la matrice $(I_n + A)^{n-1}$ est strictement positive, alors elle est irréductible ainsi que A (lemme 1.6). ■

Remarque 5.1 *On peut en fait montrer que $A \in \mathcal{M}_n(\mathbb{K})$ est irréductible si et seulement si $(I_n + |A|)^{n-1}$ est strictement positive (voir [1], théorème 6.2.23).*

Théorème 5.3 (Perron-Frobenius III) *Si A est positive et irréductible dans $\mathcal{M}_n(\mathbb{R})$ alors $\rho(A)$ est strictement positif, c'est une valeur propre simple de A et l'espace propre associé est une droite vectorielle engendrée par un vecteur strictement positif.*

Démonstration. La matrice A étant irréductible n'a pas de ligne nulle, on a donc puisqu'elle est positive $\sum_{j=1}^n a_{ij} > 0$ pour tout i compris entre 1 et n , ce qui entraîne $\rho(A) > 0$ (corollaire 1.5).

Avec le théorème de trigonalisation sur \mathbb{C} on voit que si $\rho(A)$ est valeur propre de multiplicité supérieure ou égal à 2 alors il en est de même de $1 + \rho(A)$ comme valeur propre de $I_n + A$. Mais $I_n + A$ positive et $(I_n + A)^{n-1}$ strictement positive entraîne $\rho(I_n + A) = 1 + \rho(A)$ (corollaire 1.11) est valeur propre simple de $I_n + A$ (corollaire 1.14). En conséquence $\rho(A)$ est valeur propre simple de A .

L'espace propre associé est donc de dimension 1 et on sait qu'il peut être engendré par un vecteur positif x (corollaire 1.10). De $Ax = \rho(A)x$, on déduit que $(I_n + A)^{n-1}x = (1 + \rho(A))^{n-1}x$ et avec $(I_n + A)^{n-1}x > 0$ (théorème 1.1, point (vi)), $(1 + \rho(A))^{n-1} > 0$, on déduit que $x > 0$.

■