Développements pour l'agrégation de mathématiques

Elouan Renault

17/03/2024

Sommaire

Introduction

Ce document regroupe des développements possibles pour les oraux d'agrégation de mathématiques. Il s'agit des développements que j'avais préparés durant ma préparation à ce concours (2023-2024).

⚠ Note importante avant de lire ce document : Si vous cherchez un document rédigé synthétiquement, donnant seulement les éléments essentiels de chaque preuve, alors vous êtes au mauvais endroit. Bien au contraire, ce document a été écrit avec l'ambition de détailler le plus précisément possible chaque preuve. En particulier, les rédactions proposées dans ce document ne sont pas adaptées à une exposition orale d'une durée de 15 minutes.

Le document est divisé en 4 sections. Les deux premières, intitulées "Algèbre et géométrie" et "Analyse et probabilités", contiennent dans l'ordre suivant :

- un tableau donnant pour chaque développement la liste des leçons proposées pour présenter ce développement;
- un tableau donnant pour chaque leçon la liste des développements qui pourraient être présentés pour cette leçon;
- les développements;
- éventuellement un ou plusieurs compléments conseillés;
- des commentaires personnels;
- les références bibliographiques;
- les références détaillées (chaque développement est référencé par un lien cliquable qui redirige vers cette sous-section, par exemple [DP]);

La troisième section contient les résultats annexes. Enfin, la dernière section contient toutes les figures de ce document.

Bien sûr, mes commentaires sur ces développements sont purement subjectifs, notamment lorsque je parle de difficulté. Vous trouverez dans ces derniers une très forte occurrence de l'adjectif "technique". Inversement, il m'arrive de trouver certains développements faciles. Mais gardez en tête que la difficulté est une notion relative à chacun, et que certains lecteurs trouveront faciles des développements que j'estime difficiles, tandis que d'autres jugeront difficiles des développements qui me semblent faciles. Pour savoir si un développement vous convient (en termes de difficulté et de temps de présentation), il est impératif de s'entraîner à l'exposer dans les conditions des oraux d'agrégation. Cela vous permettra aussi de vous approprier vos développements, et éventuellement d'y ajouter votre touche personnelle dans leur présentation.

Les épreuves orales de l'agrégation permettant aux candidats d'utiliser des livres de mathématiques, j'ai essayé de donner une liste (non exhaustive) de livres qui pourraient vous aider à revoir votre développement le jour de l'oral. Néanmoins, la rédaction de la preuve diffère souvent de celle du livre donné en référence. Par conséquent, si vous souhaitez revoir vos développements dans des livres durant ces épreuves, je vous conseille de travailler en premier lieu vos développements sur des livres.

Toujours dans le cas où vous souhaiteriez revoir vos développements, ce document ne doit donc pas être votre support principal pour travailler vos développements, mais seulement un support secondaire pour éventuellement mieux comprendre certaines preuves. En effet, n'étant pas le couteau le plus aiguisé du tiroir, j'ai la fâcheuse tendance à m'éterniser sur la rédaction des preuves, en donnant toujours plus de détails que la rédaction dont je m'inspire.

Par ailleurs, qui dit plus de détails, dit aussi plus de coquilles, de maladresses, de raisonnements tirés par les cheveux, ou d'erreurs monumentales. Chaque fois que je m'éloigne de la référence originale pour compléter des passages qui ne sont pas suffisamment détaillés pour mon intellect limité, je prends le risque d'écrire une erreur, ou d'embrouiller le lecteur en l'inondant d'explications superflues. Aussi, je vous serai infiniment reconnaissant de me signaler toute erreur à l'adresse e-mail suivante : erratаouquestions@gmail.com (vous pouvez aussi me poser des questions comme son nom l'indique).

Toutefois, par souci d'honnêteté, je me dois d'avouer que, durant aucune de ces épreuves orales, je n'ai eu le temps de revoir mes développements dans les livres, faute de temps de préparation. J'ai toujours été un étudiant très lent, et je savais que je n'aurai pas le temps d'écrire ma leçon et de revoir mes deux développements en seulement trois heures. D'autant plus que j'avais préparé peu de leçons, utilisant par conséquent la quasi-totalité de mon temps de préparation pour construire et rédiger ma leçon.

C'est pourquoi j'ai préféré rédiger une rédaction personnelle de mes développements plutôt que d'imiter celles des livres, de telle sorte que je comprenne bien chaque preuve. C'est tout l'enjeu de cet exercice : il ne s'agit pas de réciter une preuve que l'on aurait apprise par cœur, mais bien de s'approprier un résultat pour être le plus clair possible lors de son exposé. Je rappelle à toutes fins utiles qu'il s'agit d'un concours pour devenir enseignant, et non d'un concours sélectionnant les étudiants les plus forts en mathématiques. Aussi, un développement simple exposé avec clarté et pédagogie sera bien plus valorisé qu'un développement compliqué exposé superficiellement ou, pire encore, exposé de façon confuse et désordonnée.

Ce dernier conseil ne provient pas de moi, mais de mes enseignants qui ont encadré la préparation à l'agrégation de mathématiques de l'université d'Aix-Marseille durant la session 2024. Je tiens à ce propos à les remercier pour leur aide précieuse durant cette année. Je tiens également à remercier :

- mes camarades de promotion, qui m'ont également été d'une grande aide en présentant leurs développements tout au long de l'année;
- tous les agrégatifs qui ont mis en ligne leurs plans ou leurs développements;
- les créateurs du site agreg-maths.fr : votre site m'a été d'une grande utilité.

Enfin, je vous souhaite bon courage pour votre préparation à ce concours.

Repetitio est mater studiorum

(la répétition est la mère de l'apprentissage)

Algèbre et géométrie

Liste des développements d'algèbre et géométrie

Numéro	Développement	leçons proposées
1	Automorphismes de $\mathfrak{S}_n$	101,104,105,108,190
2	Construction des corps finis	123,125,141
3	Critère d'Eisenstein	122,141,142
4	Cyclicité de $(\mathbb{Z}/p^\alpha\mathbb{Z})^\star$	104,108,120,121
5	Décomposition de Dunford	150,151,152,154,155,156
6	Décomposition polaire	106,152,154,157,158
7	Dénombrement des polynômes irréductibles unitaires sur un corps fini	123,125,141,190
8	Déterminant de Gram	149,161,191
9	Enveloppe convexe des matrices orthogonales	157,158,159,161,181
10	Générateurs de $GL_n(\mathbb{K})$ et $SL_n(\mathbb{K})$	106,108,162
11	Générateurs du groupe orthogonal d'un espace vectoriel euclidien	108,148,158,161,191
12	Générateurs du groupe orthogonal d'une forme quadratique définie	108,148,170,171,191
13	Loi d'inertie de Sylvester	148,170,171
14	Probabilité que deux éléments commutent dans un groupe	101,103,104,190
15	Simplicité de $\mathfrak{A}_n$	103,104,105,108
16	Surjectivité de l'exponentielle matricielle	150,155,204,214
17	Théorème de Burnside	106,153,156
18	Théorème de Cauchy	101,104,105
19	Théorème de d'Alembert-Gauss	144
20	Théorème de Dirichlet	102,120,121
21	Théorème de Kronecker	102,144
22	Théorème des deux carrés de Fermat	121,122,127
23	Théorème des restes chinois	120,122,142
24	Théorème spectral	148,151,152,157,158
25	Une suite de polygones	127,149,150,152,153,181,191,226

Liste des leçons d'algèbre et géométrie

Numéro	Leçon	Développements
101	Groupe opérant sur un ensemble. Exemples et applications.	Automorphismes de $\mathfrak{S}_n$ Probabilité que deux éléments commutent dans un groupe Théorème de Cauchy
102	Groupe des nombres complexes de module 1. Racines de l'unité. Applications.	Théorème de Dirichlet Théorème de Kronecker
103	Conjugaison dans un groupe. Exemples de sous-groupes distingués et de groupes quotients. Applications.	Probabilité que deux éléments commutent dans un groupe Simplicité de $\mathfrak{A}_n$
104	Groupes finis. Exemples et applications.	Automorphismes de $\mathfrak{S}_n$ Cyclicité de $(\mathbb{Z}/p^\alpha\mathbb{Z})^\star$ Probabilité que deux éléments commutent dans un groupe Simplicité de $\mathfrak{A}_n$ Théorème de Cauchy
105	Groupe des permutations d'un ensemble fini. Applications.	Automorphismes de $\mathfrak{S}_n$ Simplicité de $\mathfrak{A}_n$ Théorème de Cauchy
106	Groupe linéaire d'un espace vectoriel de dimension finie E, sous-groupes de $GL(E)$. Applications.	Décomposition polaire Générateurs de $GL_n(\mathbb{K})$ et $SL_n(\mathbb{K})$ Théorème de Burnside
108	Exemples de parties génératrices d'un groupe. Applications.	Automorphismes de $\mathfrak{S}_n$ Cyclicité de $(\mathbb{Z}/p^\alpha\mathbb{Z})^\star$ Générateurs de $GL_n(\mathbb{K})$ et $SL_n(\mathbb{K})$ Générateurs du groupe orthogonal d'un espace vectoriel euclidien Générateurs du groupe orthogonal d'une forme quadratique définie Simplicité de $\mathfrak{A}_n$
120	Anneaux $\mathbb{Z}/n\mathbb{Z}$. Applications.	Cyclicité de $(\mathbb{Z}/p^\alpha\mathbb{Z})^\star$ Théorème de Dirichlet Théorème des restes chinois
121	Nombres premiers. Applications.	Cyclicité de $(\mathbb{Z}/p^\alpha\mathbb{Z})^\star$ Théorème de Dirichlet Théorème des deux carrés de Fermat Série des inverses des nombres premiers
122	Anneaux principaux. Exemples et applications.	Critère d'Eisenstein Théorème des deux carrés de Fermat Théorème des restes chinois
123	Corps finis. Applications.	Construction des corps finis Dénombrement des polynômes irréductibles unitaires sur un corps fini
125	Extensions de corps. Exemples et applications.	Construction des corps finis Dénombrement des polynômes irréductibles unitaires sur un corps fini

Numéro	Leçon	Développements
127	Exemples de nombres remarquables. Exemples d'anneaux de nombres remarquables. Applications.	Théorème des deux carrés de Fermat Une suite de polygones
141	Polynômes irréductibles à une indéterminée. Corps de rupture. Exemples et applications.	Construction des corps finis Critère d'Eisenstein Dénombrement des polynômes irréductibles unitaires sur un corps fini
142	PGCD et PPCM, algorithmes de calcul. Applications.	Critère d'Eisenstein Théorème des restes chinois
144	Racines d'un polynôme. Fonctions symétriques élémentaires. Exemples et applications.	Théorème de d'Alembert-Gauss Théorème de Kronecker
148	Dimension d'un espace vectoriel. Rang. Exemples et applications.	Générateurs du groupe orthogonal d'un espace vectoriel euclidien Générateurs du groupe orthogonal d'une forme quadratique définie Loi d'inertie de Sylvester Théorème spectral
149	Déterminant. Exemples et applications.	Déterminant de Gram Une suite de polygones
150	Polynômes d'endomorphisme en dimension finie. Réduction d'un endomorphisme en dimension finie. Applications.	Décomposition de Dunford Surjectivité de l'exponentielle matricielle Une suite de polygones
151	Sous-espaces stables par un endomorphisme ou une famille d'endomorphismes d'un espace vectoriel de dimension finie. Applications.	Décomposition de Dunford Théorème spectral
152	Endomorphismes diagonalisables en dimension finie.	Décomposition de Dunford Décomposition polaire Théorème spectral Une suite de polygones
153	Valeurs propres, vecteurs propres. Calculs exacts ou approchés d'éléments propres. Applications.	Théorème de Burnside Une suite de polygones
154	Exemples de décompositions de matrices. Applications.	Décomposition de Dunford Décomposition polaire
155	Exponentielle de matrices. Applications.	Décomposition de Dunford Surjectivité de l'exponentielle matricielle Équation de Sylvester
156	Endomorphismes trigonalisables. Endomorphismes nilpotents.	Décomposition de Dunford Théorème de Burnside Équation de Sylvester
157	Matrices symétriques réelles, matrices hermitiennes.	Décomposition polaire Enveloppe convexe des matrices orthogonales Théorème spectral
158	Endomorphismes remarquables d'un espace vectoriel euclidien.	Décomposition polaire Enveloppe convexe des matrices orthogonales Générateurs du groupe orthogonal d'un espace vectoriel euclidien Théorème spectral

Numéro	Leçon	Développements
159	Formes linéaires et dualité en dimension finie. Exemples et applications.	Enveloppe convexe des matrices orthogonales Théorème des extrema liés
161	Espaces vectoriels et espaces affines euclidiens : distances, isométries.	Déterminant de Gram Enveloppe convexe des matrices orthogonales Générateurs du groupe orthogonal d'un espace vectoriel euclidien
162	Systèmes d'équations linéaires ; opérations élémentaires, aspects algorithmiques et conséquences théoriques.	Équation de Sylvester Générateurs de $GL_n(\mathbb{K})$ et $SL_n(\mathbb{K})$
170	Formes quadratiques sur un espace vectoriel de dimension finie. Orthogonalité. Applications.	Générateurs du groupe orthogonal d'une forme quadratique définie Loi d'inertie de Sylvester
171	Formes quadratiques réelles. Coniques. Exemples et applications.	Générateurs du groupe orthogonal d'une forme quadratique définie Loi d'inertie de Sylvester
181	Convexité dans $\mathbb{R}^n$. Applications en algèbre et en géométrie.	Enveloppe convexe des matrices orthogonales Une suite de polygones
190	Méthodes combinatoires, problèmes de dénombrement.	Automorphismes de $\mathfrak{S}_n$ Dénombrement des polynômes irréductibles unitaires sur un corps fini Probabilité que deux éléments commutent dans un groupe Nombres de Bell
191	Exemples d'utilisation de techniques d'algèbre en géométrie.	Déterminant de Gram Générateurs du groupe orthogonal d'un espace vectoriel euclidien Générateurs du groupe orthogonal d'une forme quadratique définie Une suite de polygones

Automorphismes de $\mathfrak{S}_n$

[DP] **Lemme 1.** Soit $\varphi \in \mathrm{Aut}(\mathfrak{S}_n)$. Si l'image de toute transposition par φ est une transposition alors $\varphi \in \mathrm{Int}(\mathfrak{S}_n)$.

Preuve. Si $n = 2$, $\mathrm{Aut}(\mathfrak{S}_n) = \mathrm{Int}(\mathfrak{S}_n) = \{Id_{\mathfrak{S}_n}\}$. Supposons $n \geq 3$. Pour tout $i \in [\![2, n]\!]$, notons $\tau_i = (1, i)$. Soient $i, j \in [\![2, n]\!]$. Si $i \neq j$, alors $\tau_i\tau_j \neq \tau_j\tau_i$, puis $\varphi(\tau_i)\varphi(\tau_j) \neq \varphi(\tau_j)\varphi(\tau_i)$ car $\varphi \in \mathrm{Aut}(\mathfrak{S}_n)$, et donc $\varphi(\tau_i)$ et $\varphi(\tau_j)$ ont des supports non disjoints. Montrons ensuite qu'il existe $a_1, \ldots, a_n \in [\![1, n]\!]$ deux à deux distincts tels que,

$$\forall i \in [\![2, n]\!], \ \varphi(\tau_i) = (a_1, a_i).$$

Tout d'abord, $\varphi(\tau_2)$ et $\varphi(\tau_3)$ étant des transpositions à supports non disjoints, il existe $a_1 \in [\![1, n]\!]$ et $a_2, a_3 \in [\![1, n]\!]$ distincts de a_1 tels que,

$$\varphi(\tau_2) = (a_1, a_2), \ \varphi(\tau_3) = (a_1, a_3).$$

et $a_2 \neq a_3$ car $\tau_2 \neq \tau_3$ et φ est injective. Si $n = 3$, il n'y a plus rien à montrer. Sinon, montrons que, pour tout $i \in [\![4, n]\!]$, $a_1 \in \mathrm{Supp}(\varphi(\tau_i))$. On raisonne par l'absurde en supposant qu'il existe $i \in [\![4, n]\!]$ tel que a_1 n'appartienne pas au support de $\varphi(\tau_i)$. En particulier,

$$\begin{cases} \varnothing \neq \mathrm{Supp}(\varphi(\tau_2)) \cap \mathrm{Supp}(\varphi(\tau_i)) \subset \{a_2\} \\ \varnothing \neq \mathrm{Supp}(\varphi(\tau_3)) \cap \mathrm{Supp}(\varphi(\tau_i)) \subset \{a_3\} \end{cases}$$

ce qui implique $a_2, a_3 \in \mathrm{Supp}(\varphi(\tau_i))$. Sachant que $\varphi(\tau_i)$ est une transposition et que $a_2 \neq a_3$, il vient donc $\varphi(\tau_i) = (a_2, a_3)$. Par suite,

$$\varphi(\tau_2\tau_3\tau_i) = \varphi(\tau_2)\varphi(\tau_3)\varphi(\tau_i) = (a_1, a_2)(a_1, a_3)(a_2, a_3) = (a_1, a_3) = \varphi(\tau_3),$$

et donc, en composant par φ^{-1}, $\tau_2\tau_3\tau_i = \tau_3$, ce qui contredit $(\tau_2\tau_3\tau_i)(1) = i$ et $\tau_3(1) = 3$. Ainsi, pour tout $i \in [\![4, n]\!]$, $a_1 \in \mathrm{Supp}(\varphi(\tau_i))$. Or, $\varphi(\tau_4), \ldots, \varphi(\tau_n)$ sont des transpositions. Par conséquent, il existe $a_4, \ldots, a_n \in [\![4, n]\!]$ distincts de a_1 tels que,

$$\forall i \in [\![2, n]\!], \ \varphi(\tau_i) = (a_1, a_i).$$

et $a_2, \ldots, a_n$ sont deux à deux distincts car $\tau_2, \ldots, \tau_n$ sont deux à deux distincts et φ est injective. Posons,

$$\sigma = \begin{pmatrix} 1 & 2 & \cdots & n \\ a_1 & a_2 & \cdots & a_n \end{pmatrix} \in \mathfrak{S}_n$$

Alors,

$$\forall i \in [\![2, n]\!], \ \varphi(\tau_i) = \sigma\tau_i\sigma^{-1}.$$

Ainsi, les applications $\mathfrak{S}_n \to \mathfrak{S}_n$, $\rho \mapsto \sigma\rho\sigma^{-1}$ et φ sont des morphismes de groupes coïncidant sur $\{\tau_i, \ 1 \leq i \leq n\}$, qui est une partie génératrice de $\mathfrak{S}_n$, et donc ces derniers sont égaux. En particulier, $\varphi \in \mathrm{Int}(\mathfrak{S}_n)$. □

[DP] **Lemme 2.** Soit $\sigma \in \mathfrak{S}_n$. On suppose que σ est un produit de k transpositions à supports disjoints. Alors, en notant $c(\sigma)$ le centralisateur de σ, il vient,

$$|c(\sigma)| = k!(n - 2k)!2^k.$$

Preuve. Par hypothèse, il existe $\tau_1, \ldots, \tau_k$ des transpositions à supports deux à deux disjoints telles que $\sigma = \tau_1 \circ \cdots \circ \tau_k$. Soit $\rho \in \mathfrak{S}_n$. Alors,

$$\rho\sigma\rho^{-1} = (\rho\tau_1\rho^{-1}) \cdots (\rho\tau_k\rho^{-1}). \tag{1}$$

Soit $i \in [\![1, n]\!]$. Soit $\{a_i, b_i\} \subset [\![1, n]\!]$ tel que $\tau_i = (a_i, b_i)$. Alors,

$$\rho(a_i, b_i)\rho^{-1} = (\rho(a_i), \rho(b_i)). \tag{2}$$

En particulier, $\rho\tau_1\rho^{-1}, \ldots, \rho\tau_k\rho^{-1}$ sont des transpositions. De plus, elles sont à supports disjoints car les τ_i le sont et ρ est injective. L'égalité (1) est donc la décomposition de $\rho\sigma\rho^{-1}$ en produit de cycles à supports disjoints. Par unicité de celle-ci, il s'en suit que ρ et σ commutent si et seulement si,

$$\{\rho\tau_1\rho^{-1}, \ldots, \rho\tau_k\rho^{-1}\} = \{\tau_1, \ldots, \tau_k\}.$$

À présent, supposons que ρ commute avec σ, puis dénombrons les possibilités pour chaque valeur de ρ.

- Tout d'abord, pour $\rho\tau_1\rho^{-1}$, il y a k possibilités. L'application $\eta \mapsto \rho\eta\rho^{-1}$ étant injective, il reste alors $k - 1$ possibilités pour $\rho\tau_2\rho^{-1}$. En itérant le raisonnement, on obtient donc qu'il y a $k!$ possibilités pour l'ensemble $\{\rho\tau_1\rho^{-1}, \ldots, \rho\tau_k\rho^{-1}\}$. D'autre part, par hypothèse sur ρ, il existe $f \in \mathfrak{S}_k$ telle que, pour tout $i \in [\![1, k]\!]$, on ait $\rho\tau_i\rho^{-1} = \tau_{f(i)}$, c'est-à-dire,

  $$(\rho(a_i), \rho(b_i)) = (a_{f(i)}, b_{f(i)})$$

  en utilisant l'égalité (2). Il reste alors deux possibilités, soit $\rho(a_i) = a_{f(i)}$ et $\rho(b_i) = b_{f(i)}$, soit $\rho(a_i) = b_{f(i)}$ et $\rho(b_i) = a_{f(i)}$. Il y a donc $k!2^k$ valeurs possibles pour le $2k$-uplet $(\rho(a_1), \ldots, \rho(a_k), \rho(b_1), \ldots, \rho(b_k))$.

- Finalement, notre hypothèse sur ρ n'impose rien sur ses valeurs restantes. Puisqu'il en reste $n - 2k$, il y a donc $(n - 2k)!$ possibilités.

Cela fait donc $k!(n - 2k)!2^k$ possibilités pour ρ. D'où le résultat. □

Remarque. De façon plus formelle (mais moins intuitive de mon point de vue), en notant $F = [\![1,n]\!] \setminus \mathrm{Supp}(\sigma) = [\![1,n]\!] \setminus \{a_1,\dots,a_k,b_1,\dots,b_k\}$, on peut en fait montrer que l'application,

$$\begin{array}{rccc} \Psi & : & \mathfrak{S}_k \times \{0,1\}^k \times \mathfrak{S}(F) & \longrightarrow & c(\sigma) \\ & & (f,\varepsilon_1,\dots,\varepsilon_k,g) & \longmapsto & \Psi(f,\varepsilon_1,\dots,\varepsilon_k,g) \end{array}$$

est une bijection, où l'on a noté $\Psi(f,\varepsilon_1,\dots,\varepsilon_k,g)$ l'application définie par,

$$\forall i \in [\![1,n]\!],\ \Psi(f,\varepsilon_1,\dots,\varepsilon_k,g)(i) = \begin{cases} a_{f(i)} & \text{si } i \in \mathrm{Supp}(\sigma) \text{ et } \varepsilon_i = 0, \\ b_{f(i)} & \text{si } i \in \mathrm{Supp}(\sigma) \text{ et } \varepsilon_i = 1, \\ g(i) & \text{si } i \notin \mathrm{Supp}(\sigma). \end{cases}$$

Le lemme suivant n'est pas nécessaire pour prouver le théorème. Il s'agit d'une généralisation du lemme précédent.

[DP] **Lemme 3.** Soit $\sigma \in \mathfrak{S}_n$. On suppose que $k_2,\dots,k_n$ sont des entiers naturels tels que σ soit le produit de $k_2+\cdots+k_n$ cycles à supports disjoints, dont k_2 cycles d'ordre 2, k_3 cycles d'ordre $3,\dots,k_n$ cycles d'ordre n. Alors, en notant $k_1 = n-(2k_2+\cdots+nk_n)$ et $c(\sigma)$ le centralisateur de σ, il vient,

$$|c(\sigma)| = \prod_{i=1}^{n} k_i! i^{k_i}.$$

Preuve. Notons $r = k_2+\cdots+k_n$. Par hypothèse, il existe $\gamma_1,\dots,\gamma_r$ des cycles à supports disjoints tels que $\sigma = \gamma_1\dots\gamma_r$. Soit $\rho \in \mathfrak{S}_n$. Alors,

[DP]
$$\rho\sigma\rho^{-1} = (\rho\gamma_1\rho^{-1})\cdots(\rho\gamma_r\rho^{-1}). \tag{3}$$

Or, pour tout cycle $\gamma = (x_1,\dots,x_p)$,

$$\rho(x_1,\dots,x_p)\rho^{-1} = (\rho(x_1),\dots,\rho(x_p)). \tag{4}$$

En particulier, pour tout $i \in [\![1,r]\!]$, $\rho\gamma_i\rho^{-1}$ est un cycle de même ordre que γ_i. De plus, ces cycles sont à supports disjoints car les γ_i le sont et ρ est injective. Il ressort de ce qui précède que l'égalité (3) est la décomposition de $\rho\sigma\rho^{-1}$ en produit de cycles à supports disjoints. Par unicité de celle-ci, il s'en suit que ρ et σ commutent si et seulement si,

$$\{\rho\gamma_1\rho^{-1},\dots,\rho\gamma_k\rho^{-1}\} = \{\gamma_1,\dots,\gamma_k\}. \tag{5}$$

À présent, supposons que ρ commute avec σ, puis dénombrons les possibilités pour chaque valeur de ρ.

- Fixons $i \in [\![2,n]\!]$. Notons Γ_i le sous-ensemble de $\{\gamma_1,\dots,\gamma_r\}$ formé de ses éléments d'ordre i. Il est de cardinal k_i par hypothèse. Notons,

$$\rho\Gamma_i\rho^{-1} = \{\rho\gamma\rho^{-1},\ \gamma \in \Gamma_i\}.$$

Or, on a vu précédemment que, pour tout $k \in [\![1,r]\!]$, $\rho\gamma_k\rho^{-1}$ est un cycle de même ordre que γ_k. Par conséquent, l'égalité (5) impose $\rho\Gamma_i\rho^{-1} = \Gamma_i$. Si on choisit un premier élément quelconque dans $\rho\Gamma_i\rho^{-1}$, il y a donc k_i possibilités. Lorsqu'on en choisit un deuxième, il ne reste plus que k_i-1 possibilités, par injectivité de $\eta \mapsto \rho\eta\rho^{-1}$. En itérant le raisonnement, on obtient donc qu'il y a $k_i!$ possibilités pour l'ensemble $\rho\Gamma_i\rho^{-1}$. Fixons ensuite $\gamma = (x_1,\dots,x_i) \in \Gamma_i$. Alors, $\rho\gamma\rho^{-1} = \gamma$, c'est-à-dire, d'après l'égalité (4),

$$(\rho(x_1),\dots,\rho(x_i)) = (x_1,\dots,x_i)$$

ce qui laisse donc i possibilités pour $\rho(x_1),\dots,\rho(x_i)$. Puisque Γ_i est de cardinal k_i, il y a donc $k_i!i^{k_i}$ possibilités pour le $(n-k_1)$-uplet,

$$\left(\rho(x),\ x \in \bigcup_{\gamma\in\Gamma_i} \mathrm{Supp}(\gamma)\right).$$

- Finalement, notre hypothèse sur ρ n'impose rien sur ses valeurs restantes. Puisqu'il en reste $n-(2k_2+\cdots+nk_n) = k_1$, il y a donc $k_1!$ possibilités.

En tenant compte des résultats ci-dessus, on obtient donc qu'il y a,

$$k_1!\prod_{i=2}^{n} k_i! i^{k_i} = \prod_{i=1}^{n} k_i! i^{k_i},$$

possibilités pour ρ. D'où le résultat. □

Théorème 4. Si $n \neq 6$, les automorphismes de $\mathfrak{S}_n$ sont tous intérieurs.

Preuve. Soit $\varphi \in \mathrm{Aut}(\mathfrak{S}_n)$. Soit τ une transposition. Montrons que $\varphi(\tau)$ est une transposition, ce qui suffira pour conclure d'après le lemme 1. Tout d'abord, $\varphi(\tau)$ est d'ordre 2 dans $\mathfrak{S}_n$. On écrit ensuite,

$$\varphi(\tau) = \tau_1\dots\tau_k$$

la décomposition de $\varphi(\tau)$ en produit de cycles $\tau_1,\dots,\tau_k$ à supports disjoints. En particulier,

$$2 = \mathrm{ordre}(\varphi(\tau)) = \mathrm{ppcm}(\mathrm{ordre}(\tau_1),\dots,\mathrm{ordre}(\tau_k)),$$

ce qui impose que $\tau_1,\dots,\tau_k$ sont des transpositions. Pour conclure que $\varphi(\tau)$ est une transposition, il suffit donc de montrer que $k=1$. Pour cela, on commence par remarquer que $\varphi(c(\tau)) = c(\varphi(\tau))$. En effet, si $\sigma \in \mathfrak{S}_n$ alors,

$$\begin{array}{rcl} \sigma \in \varphi(c(\tau)) & \iff & (\exists \rho \in \mathfrak{S}_n \ /\ (\sigma = \varphi(\rho)) \wedge (\rho\tau = \tau\rho)) \\ & \iff & (\exists \rho \in \mathfrak{S}_n \ /\ (\sigma = \varphi(\rho)) \wedge (\varphi(\rho)\varphi(\tau) = \varphi(\tau)\varphi(\rho))) \\ & \iff & \sigma\varphi(\tau) = \varphi(\tau)\sigma \\ \sigma \in \varphi(c(\tau)) & \iff & \sigma \in c(\varphi(\tau)). \end{array}$$

D'où l'égalité. En particulier,

$$|c(\varphi(\tau))| = |\varphi(c(\tau))| = |c(\tau)|$$

car φ est bijective. Or, d'après le lemme 2,

$$\begin{cases} |c(\tau)| = 2(n-2)! \\ |c(\varphi(\tau))| = k!(n-2k)!2^k \end{cases}$$

et donc,

$$2(n-2)! = k!(n-2k)!2^k.$$

Par ailleurs,

$$\begin{aligned} 2(n-2)! = k!(n-2k)!2^k &\iff \frac{(n-2)!}{k!(n-2k)!2^{k-1}} = 1 \\ &\iff \frac{\binom{n-2}{2k-2}(2k-2)!}{k!2^{k-1}} = 1 \\ 2(n-2)! = k!(n-2k)!2^k &\iff \frac{\binom{n-2}{2k-2}(2k-3)(2k-5)\times\cdots\times 3\times 1}{k} = 1 \end{aligned}$$

la dernière équivalence étant due à l'égalité,

$$(2k-2)! = \left(\prod_{i=1}^{k-1}(2i)\right)\left(\prod_{i=1}^{k-2}(2i+1)\right) = 2^{k-1}(k-1)!\prod_{i=1}^{k-2}(2i+1).$$

Or, si $k > 3$ alors $2k-3 > k$ et donc,

$$\frac{\binom{n-2}{2k-2}(2k-3)(2k-5)\times\cdots\times 3\times 1}{k} > 1$$

ce qui est impossible d'après ce qui précède. Si $k = 3$ alors,

$$2(n-2)! = k!(n-2k)!2^k \iff (n-2)(n-3)(n-4)(n-5) = 24$$

ce qui équivaut à $n = 6$. En effet,

- si $n \in \{2,3,4,5\}$ alors $(n-2)(n-3)(n-4)(n-5) = 0$,
- si $n \geq 7$ alors $(n-2)(n-3)(n-4)(n-5) \geq 5\times 4\times 3\times 2 = 120 > 24$.

Puisqu'on a supposé $n \neq 6$, le cas $k = 3$ est donc impossible. Enfin, si $k = 2$,

$$2(n-2)! = k!(n-2k)!2^k \iff n^2 - 3n - 2 = 0$$

qui n'admet pas de solution dans $\mathbb{N}$. Finalement, on en conclut que $k = 1$. En particulier, $\varphi(\tau)$ est une transposition et donc $\varphi \in \mathrm{Int}(\mathfrak{S}_n)$ d'après le premier lemme. □

Commentaires personnels. Développement à la fois trop technique et trop long pour mon niveau. J'ai finalement renoncé à le présenter pour ces raisons. Bien qu'il puisse être exposé dans de nombreuses leçons, les autres développements que je propose pour ces leçons sont, à mon avis, bien plus accessibles.

Construction des corps finis

Soit p un nombre premier. À isomorphisme près, il existe un unique corps de cardinal p. On le note $\mathbb{F}_p$.

[IG] **Théorème 1.** Soit $n \in \mathbb{N}^*$. On note $q = p^n$.

(i) Si $\mathbb{K}$ est un corps de décomposition du polynôme $X^q - X$ sur $\mathbb{F}_p$, alors $\mathbb{K}$ est de cardinal q.

(ii) Tout corps de cardinal q est $\mathbb{F}_p$-isomorphe à $\mathbb{K}$.

Autrement dit, il existe un unique corps de cardinal q à $\mathbb{F}_p$-isomorphisme près.

Preuve. On sépare la preuve de l'existence et de l'unicité.

Existence. Considérons l'application suivante,

$$\begin{array}{rccc} \mathcal{F} & : & \mathbb{K} & \longrightarrow & \mathbb{K} \\ & & x & \longmapsto & x^p \end{array}$$

Il s'agit d'un morphisme de corps. Vérifions cela.

(i) Tout d'abord, $\mathcal{F}(1) = 1$.

(ii) Montrons ensuite que, pour tous $x, y \in \mathbb{K}$, $(x+y)^p = x^p + y^p$. D'après la formule du binôme de Newton,

$$(x+y)^p = \sum_{k=0}^{p} \binom{p}{k} x^k y^{p-k}.$$

Or, pour tout $k \in [\![1, p-1]\!]$, p divise $\binom{p}{k}$. En effet, p divise,

$$k!\binom{p}{k} = p(p-1)\dots(p-k+1)$$

et donc p divise $\binom{p}{k}$ d'après le lemme de Gauss (car p et $k!$ sont premiers entre eux), ce qui entraîne $\binom{p}{k} = 0$ dans $\mathbb{F}_p$. D'où,

$$(x+y)^p = x^p + y^p.$$

(iii) Pour tous $x, y \in \mathbb{K}$, $\mathcal{F}(xy) = \mathcal{F}(x)\mathcal{F}(y)$.

Ainsi, $\mathcal{F}$ est un morphisme de corps, appelé morphisme de Frobenius. Notons ensuite $\mathbb{L}$ l'ensemble des points fixes du morphisme de corps,

$$\mathcal{F}^n = \underbrace{\mathcal{F} \circ \dots \circ \mathcal{F}}_{n \text{ fois}}.$$

Alors, $\mathbb{L}$ est un corps, en tant qu'ensemble des points fixes d'un morphisme de corps. De plus, les éléments de $\mathbb{L}$ sont exactement les racines du polynôme $X^q - X$ dans $\mathbb{K}$. Or, ce polynôme est à racines simples dans $\mathbb{K}$, car $X^q - X$ est premier avec son polynôme dérivé, et scindé dans $\mathbb{K}$, car $\mathbb{K}$ est un corps de décomposition de $X^q - X$. Par conséquent, ce polynôme admet exactement q racines dans $\mathbb{K}$, et donc $|\mathbb{L}| = q$. Par ailleurs, $\mathbb{L}$ est un corps de décomposition de $X^q - X$ sur $\mathbb{F}_p$. En effet, si l'on note $\alpha_1, \dots, \alpha_q$ les éléments de $\mathbb{L}$, il vient,

$$X^q - X = \prod_{k=1}^{q} (X - \alpha_k)$$

et donc $X^q - X$ est scindé sur $\mathbb{L}$, avec $\mathbb{L} = \mathbb{F}_p[\alpha_1, \dots, \alpha_k]$, ce qui prouve que $\mathbb{L}$ est un corps de décomposition du polynôme $X^q - X$ sur $\mathbb{F}_p$. Par unicité (à $\mathbb{F}_p$-isomorphismes près) des corps de décomposition, on en déduit que $\mathbb{K}$ et $\mathbb{L}$ sont $\mathbb{F}_p$-isomorphes. En particulier, $|\mathbb{K}| = |\mathbb{L}| = q$.

Unicité. Soit $\mathbb{M}$ un corps de cardinal q. Alors, $\mathbb{M}$ est de caractéristique p (car c'est un diviseur premier de q). En particulier, le sous-corps premier de $\mathbb{M}$ est de cardinal p, et donc il s'agit de $\mathbb{F}_p$ (à isomorphisme près !). Puisque le groupe multiplicatif $\mathbb{M}^*$ est d'ordre $q-1$, le théorème de Lagrange assure que,

$$\forall x \in \mathbb{M}^*,\ x^p = x,$$

et donc,

$$\forall x \in \mathbb{M},\ x^q = x$$

c'est-à-dire que tout élément de $\mathbb{M}$ est racine de $X^q - X$. Puisque $\mathbb{M}$ est de cardinal q, et que $X^q - X$ possède au plus q racines distinctes, on en déduit que,

$$X^q - X = \prod_{k=1}^{q} (X - \beta_k)$$

où l'on a noté $\beta_1, \dots, \beta_q$ les éléments de $\mathbb{M}$. En particulier, $X^q - X$ est scindé sur $\mathbb{M}$, et l'on a bien $\mathbb{M} = \mathbb{F}_p[\beta_1, \dots, \beta_q]$, ce qui prouve que $\mathbb{M}$ est un corps de décomposition de $X^q - X$ sur $\mathbb{F}_p$. Par unicité de ce dernier, il s'en suit que $\mathbb{M}$

et $\mathbb{K}$ sont $\mathbb{F}_p$-isomorphes.

En toute rigueur, il faudrait distinguer le sous-corps premier de $\mathbb{K}$ et celui de $\mathbb{M}$ (qui sont isomorphes mais pas nécessairement égaux), et raisonner avec un isomorphisme entre ces deux corps. Mais la preuve est alors plus lourde en notations. La justification ci-dessus est donc couramment employée. □

Remarque. En particulier, $[\mathbb{K} : \mathbb{F}_p] = n$. En effet, si on note $m = [\mathbb{K} : \mathbb{F}_p]$ alors $\mathbb{K}$ est un $\mathbb{F}_p$-espace vectoriel de dimension m, et il est donc isomorphe à $\mathbb{F}_p^m$, ce qui implique $p^n = |\mathbb{K}| = p^m$, puis $m = n$.

Notons $\mathbb{F}_q$ l'unique corps de cardinal q (à isomorphisme près). Notons ensuite $\mathcal{I}_n$ l'ensemble des polynômes irréductibles de degré n sur $\mathbb{F}_p$.

[IG] **Théorème 2.** Avec ces notations, il vient,

(i) $\mathcal{I}_n \neq \varnothing$;

(ii) pour tout $\pi \in \mathcal{I}_n$, $\mathbb{F}_p[X]/(\pi)$ est un corps de cardinal q.

Preuve.

(i) Puisque le groupe multiplicatif $\mathbb{F}_q^*$ est cyclique, il admet un générateur ξ. Alors, $\mathbb{F}_q = \mathbb{F}_p[\xi]$. En effet, on a,

$$\mathbb{F}_q = \{1, \xi, \xi^2, \dots, \xi^{q-1}\} \subset \{P(\xi),\ P \in \mathbb{F}_p[X]\} = \mathbb{F}_p[\xi]$$

et $\mathbb{F}_p[\xi] \subset \mathbb{F}_q[\xi] = \mathbb{F}_q$, car $\mathbb{F}_p \subset \mathbb{F}_q$ et $\xi \in \mathbb{F}_q$. D'où l'égalité, $\mathbb{F}_q = \mathbb{F}_p[\xi]$. Notons π le polynôme minimal de ξ sur $\mathbb{F}_p$. Alors, π est irréductible sur $\mathbb{F}_p$ et $\deg(\pi) = [\mathbb{F}_p[\xi] : \mathbb{F}_p] = [\mathbb{F}_q : \mathbb{F}_p] = n$, ce qui prouve que $\pi \in \mathcal{I}_n$.

(ii) Soit $\pi \in \mathcal{I}_n$. Alors, $\mathbb{F}_p[X]/(\pi)$ est un corps de rupture de π sur $\mathbb{F}_p$. En particulier, c'est un $\mathbb{F}_p$-espace vectoriel de degré n, et donc $\mathbb{F}_p[X]/(\pi)$ est isomorphe à $\mathbb{F}_p^n$. D'où, $|\mathbb{F}_p[X]/(\pi)| = p^n = q$.

□

[IG] **Corollaire 3.** Si $\pi \in \mathcal{I}_n$, alors π divise $X^q - X$ sur $\mathbb{F}_p$, et donc π est scindé sur $\mathbb{F}_q$.

Preuve. Soit $\pi \in \mathcal{I}_n$. Soit $\mathbb{M}$ un corps de décomposition de π sur $\mathbb{F}_p$. Soit θ une racine de π dans $\mathbb{M}$. Notons a_n le coefficient dominant de π. Puisque $\pi \in \mathcal{I}_n$, $\frac{1}{a_n}\pi$ est le polynôme minimal de θ sur $\mathbb{F}_p$, ce qui fournit $[\mathbb{F}_p(\theta) : \mathbb{F}_p] = n$. Par suite, $\mathbb{F}_p(\theta)$ est de cardinal $p^n = q$ (car il est isomorphe à $\mathbb{F}_p^n$). Or, $\theta \neq 0$ car π est irréductible sur $\mathbb{F}_p$. D'après le théorème de Lagrange, on a donc $\theta^{q-1} = 1$ (car θ appartient au groupe multiplicatif $\mathbb{F}_p(\theta)^*$ qui est de cardinal $q-1$), puis $\theta^q = \theta$. Par conséquent, $X^q - X$ est un polynôme annulateur de θ sur $\mathbb{F}_p$, ce qui implique que $\frac{1}{a_n}\pi$ divise $X^q - X$ sur $\mathbb{F}_p$ (car $\frac{1}{a_n}\pi$ est le polynôme minimal de θ sur $\mathbb{F}_p$), et donc que π divise $X^q - X$ sur $\mathbb{F}_p$. A fortiori, π divise $X^q - X$ sur $\mathbb{F}_q$ et donc π est scindé sur $\mathbb{F}_q$ car $X^q - X$ est scindé sur $\mathbb{F}_q$. □

Remarque. En particulier, tout corps de rupture de π sur $\mathbb{F}_p$ est aussi un corps de décomposition de π sur $\mathbb{F}_p$. En effet, si $\mathbb{L}$ est un corps de rupture de π sur $\mathbb{F}_p$, alors c'est un $\mathbb{F}_p$-espace vectoriel de degré n, et donc $\mathbb{L}$ est isomorphe à $\mathbb{F}_q$. Le fait que π soit scindé sur $\mathbb{F}_q$ assure alors que π est aussi scindé sur $\mathbb{L}$.

Commentaires personnels. Étant donné l'importance de ces résultats dans l'étude des corps finis, je pense qu'il est préférable de savoir les prouver (ou du moins avoir une idée de preuve). Cela est d'autant plus important si vous souhaitez présenter le dénombrement des polynômes irréductibles unitaires sur un corps fini (théorème 1), puisque ce dernier utilise le théorème 1. Et quitte à savoir le prouver, autant en faire un développement. D'autre part, comme c'est souvent le cas en théorie des corps, il vaut mieux perdre en rigueur au profit d'une rédaction plus lisible. C'est pourquoi on identifie tous les corps de cardinal p sous la notation $\mathbb{F}_p$. Toutefois, gardez bien en tête qu'il s'agit d'un abus de notation.

Critère d'Eisenstein

Soit $\mathbb{A}$ un anneau factoriel. On note $\mathbb{K}$ le corps des fractions de $\mathbb{A}$. Dans [JR], les résultats qui suivent sont énoncés et démontrés pour $\mathbb{A} = \mathbb{Z}$.

Définition 1. Soit $P \in \mathbb{A}[X]$. On appelle contenu du polynôme P, et on note $c(P)$, le pgcd de ses coefficients.

Remarque. Le contenu d'un polynôme est seulement défini modulo $\mathbb{A}^\star$.

Définition 2. Un polynôme de $\mathbb{A}[X]$ est dit primitif si son contenu vaut 1.

[JR] **Lemme 3 (de Gauss).** Soient $P, Q \in \mathbb{A}[X]$. Alors, $c(PQ) = c(P)c(Q)$.

Preuve. On procède en deux étapes.

Étape n°1. Montrons que le produit de deux polynômes primitifs P et Q de $\mathbb{A}[X]$ est aussi primitif. On raisonne par contradiction en supposant que PQ n'est pas primitif. Puisque $\mathbb{A}$ est factoriel, il existe donc un élément irréductible p de $\mathbb{A}$ divisant tous les coefficients de PQ. En notant π la surjection canonique de $\mathbb{A}$ sur $\mathbb{B} := \mathbb{A}/(p)$, il s'en suit que, $\pi(P)\pi(Q) = \pi(PQ) = 0$. Or, $\mathbb{B}$ est intègre. En effet, p est premier car irréductible dans un anneau factoriel (cf. théorème 7.6 de [JR]) et donc $\mathbb{A}/(p)$ est intègre (cf. théorème 7.5 de [JR]). D'où, $\pi(P) = 0$ ou $\pi(Q) = 0$, puis p divise $c(P)$ ou $c(Q)$, ce qui contredit le fait que P et Q soient primitifs.

Étape n°2. Montrons l'égalité souhaitée. Par définition du contenu, il existe $\widetilde{P}$ et $\widetilde{Q}$ des polynômes primitifs de $\mathbb{A}[X]$ tels que $P = c(P)\widetilde{P}$ et $Q = c(Q)\widetilde{Q}$, et donc $PQ = c(P)c(Q)R$ où l'on a noté $R = \widetilde{P}\widetilde{Q}$. Or, R est primitif d'après l'étape précédente. Par homogénéité du pgcd, il s'en suit que,

$$c(PQ) = c(c(P)c(Q)R) = c(P)c(Q)c(R) = c(P)c(Q).$$

□

[XG] **Lemme 4.** Soit $P \in \mathbb{A}[X]$. Si P est réductible dans $\mathbb{K}[X]$, alors il existe des polynômes $Q, R \in \mathbb{A}[X]$ non constants tels que $P = QR$.

Preuve. Soient $Q, R \in \mathbb{K}[X]$ deux polynômes non constants tels que $P = QR$, qui existent car P est réductible sur $\mathbb{K}[X]$. Soit $\alpha \in \mathbb{A}$ tel que $\alpha Q, \alpha R \in \mathbb{A}[X]$ (on peut prendre par exemple le produit des dénominateurs des coefficients de Q et R). Soient $\widetilde{Q}, \widetilde{R} \in \mathbb{A}[X]$ des polynômes primitifs tels que $\alpha Q = c(\alpha Q)\widetilde{Q}$ et $\alpha R = c(\alpha R)\widetilde{R}$. Avec ces notations, il vient,

$$\alpha^2 P = (\alpha Q)(\alpha R) = c(\alpha Q)c(\alpha R)\widetilde{Q}\widetilde{R}.$$

D'après le lemme de Gauss, cette égalité fournit $\alpha^2 c(P) = c(\alpha Q)c(\alpha R)$. D'où, en utilisant les deux égalités précédentes, $\alpha^2 P = \alpha^2 c(P)\widetilde{Q}\widetilde{R}$, et donc,

$$P = c(P)\widetilde{Q}\widetilde{R}$$

avec $\widetilde{Q}$ et $\widetilde{R}$ des polynômes non constants de $\mathbb{A}[X]$ (ils sont de même degré que Q et R). □

Remarque. A fortiori, P est réductible dans $\mathbb{A}[X]$.

La preuve ci-dessous est inspirée de celle d'Antoine Barrier (qu'on peut trouver ici).

[JR] **Théorème 5 (Critère d'Eisenstein).** Soit $P(X) = a_n X^n + \cdots + a_1 X + a_0$ un polynôme à coefficients dans $\mathbb{A}$. On suppose qu'il existe un nombre premier p tel que,

- p divise $a_0, a_1, \ldots, a_{n-1}$;
- p ne divise pas a_n;
- p^2 ne divise pas a_0.

Alors, P est irréductible dans $\mathbb{K}[X]$. Si de plus P est primitif, alors il est irréductible dans $\mathbb{A}[X]$.

Preuve. On raisonne par contradiction en supposant que P soit réductible dans $\mathbb{K}[X]$. D'après le lemme précédent, il existe donc des polynômes $Q, R \in \mathbb{A}[X]$ non constants tels que $P = QR$. Notons q et r les degrés respectifs de Q et R, et $b_0, \ldots, b_q, c_0, \ldots, c_r$ leurs coefficients respectifs. On a donc,

$$Q = \sum_{k=0}^{q} b_k X^k, \quad R = \sum_{l=0}^{r} c_l X^l.$$

Notons π la surjection canonique de $\mathbb{A}[X]$ sur $\mathbb{A}[X]/(p)$. De l'égalité $P = QR$ dans $\mathbb{A}$, on en déduit que,

$$\pi(P) = \pi(a_n)X^n = \pi(Q)\pi(R)$$

car $a_0, \ldots, a_{n-1}$ sont divisibles par p. Montrons ensuite que,

$$\begin{cases} \forall k \in [\![0, q-1]\!], \ \pi(b_k) = 0, \\ \forall l \in [\![0, r-1]\!], \ \pi(c_l) = 0. \end{cases}$$

Dans le cas contraire, on peut noter k_0 le plus petit entier compris entre 0 et $q-1$ tel que $\pi(b_{k_0})$ soit non nul, et l_0 le plus petit entier compris entre 0 et $r-1$ tel que $\pi(c_{l_0})$ soit non nul. On a donc,

$$\pi(Q) = \sum_{k=k_0}^{q} \pi(b_k)X^k,\ \pi(R) = \sum_{l=l_0}^{r} \pi(c_l)X^l.$$

Par suite, le terme de degré l_0+k_0 de $\pi(Q)\pi(Q)$ est $\pi(b_{k_0})\pi(c_{l_0})$. En identifiant les coefficients de degré l_0+k_0 dans l'égalité $\pi(a_n)X^n = \pi(Q)\pi(R)$, on obtient donc $\pi(b_{k_0})\pi(c_{l_0}) = 0$ avec $\pi(b_{k_0}) \neq 0$ et $\pi(c_{l_0}) \neq 0$, ce qui contredit l'intégrité de $\mathbb{A}/(\pi)$ (ce dernier est intègre pour les mêmes raisons que dans la preuve du lemme de Gauss). D'où,

$$\begin{cases} \forall k \in [\![0, q-1]\!],\ \pi(b_k) = 0, \\ \forall l \in [\![0, r-1]\!],\ \pi(c_l) = 0. \end{cases}$$

En particulier, $\pi(b_0) = \pi(c_0) = 0$, c'est-à-dire que p divise b_0 et c_0, et donc p^2 divise a_0, ce qui est impossible. Ainsi, P est irréductible sur $\mathbb{K}[X]$.

Supposons ensuite que P est primitif, et montrons que P est irréductible sur $\mathbb{A}[X]$. Soient $Q, R \in \mathbb{A}[X]$ tels que $P = QR$. A fortiori, $P = QR$ dans $\mathbb{K}[X]$. Or, P est irréductible dans $\mathbb{K}[X]$. Donc, Q ou R est constant. Les rôles de Q et R étant symétriques, on peut supposer sans perte de généralité que Q est constant. Et puisque $Q \in \mathbb{A}[X]$, il existe donc $a \in \mathbb{A}$ tel que $Q = a$. D'après le lemme de Gauss, l'égalité $P = QR$ donne $1 = c(P) = ac(R)$, car P est primitif. Cette égalité est seulement vraie modulo $\mathbb{A}^\star$, mais cela suffit pour conclure que $a \in \mathbb{A}^\star$. Ainsi, $Q \in \mathbb{A}[X]^\star$ (on a $(\mathbb{A}[X])^\star = \mathbb{A}^\star$ car $\mathbb{A}$ est intègre), ce qui prouve finalement que P est irréductible dans $\mathbb{A}[X]$. □

Remarque. Le fait que $(\mathbb{A}[X])^\star = \mathbb{A}^\star$ lorsque $\mathbb{A}$ est intègre résulte de la formule des degrés : pour tous $P, Q \in \mathbb{A}[X]$ non nuls, $\deg(PQ) = \deg(P) + \deg(Q)$. Cette formule est vraie lorsque l'anneau $\mathbb{A}$ est intègre, mais n'est pas garantie pour un anneau non intègre (par exemple $(2X)^2 = 0$ dans $(\mathbb{Z}/4\mathbb{Z})[X]$ n'est pas de degré 2). De même l'égalité $\mathbb{A}[X]^\star = \mathbb{A}^\star$ n'est pas nécessairement vraie pour un anneau non intègre (par exemple $(2X+1)^2 = 1$ dans $(\mathbb{Z}/4\mathbb{Z})[X]$ est un polynôme non constant inversible).

Commentaires personnels. Développement technique. On peut également le présenter pour $\mathbb{A} = \mathbb{Z}$, ce qui facilite certains passages, mais le rend moins pertinent dans la leçon n°122 (sur les anneaux principaux). Je pense que c'est un bon choix de développement, puisqu'il est accessible et rentre dans plusieurs leçons. De plus, le critère d'Eisenstein figure sur le programme du concours.

Cyclicité de $(\mathbb{Z}/p^\alpha\mathbb{Z})^\star$

[JR] **Théorème 1.** Soient p un nombre premier impair et α un entier supérieur ou égal à 2. Alors, le groupe multiplicatif $(\mathbb{Z}/p^\alpha\mathbb{Z})^*$ est cyclique.

Preuve. On procède en plusieurs étapes.

Étape n°1. Soit $k \in [\![1, p-1]\!]$. Montrons que p divise $\binom{p}{k}$. Tout d'abord, p divise $k!\binom{p}{k}$ car $k!\binom{p}{k} = p(p-1)\dots(p-k+1)$. Or, p est premier avec $k!$. En effet, si p divise $k!$ alors le lemme d'Euclide assure l'existence de $j \in [\![1, k]\!]$ tel que p divise j, ce qui est impossible car $j < p$. Ainsi, p divise $k!\binom{p}{k}$ et p est premier avec $k!$. D'après le lemme de Gauss, il s'en suit que p divise $\binom{p}{k}$.

Étape n°2. Il existe une suite $(\lambda_k)_{k\in\mathbb{N}}$ d'entiers naturels non nuls tels que, pour tout $k \in \mathbb{N}$, λ_k et p soient premiers entre eux et,

$$(1+p)^{p^k} = 1 + \lambda_k p^{k+1}.$$

Montrons ce résultat par récurrence. Pour $k = 0$, $\lambda_1 = 1$ convient. Pour $k = 1$, la formule du binôme de Newton donne,

$$(1+p)^p = \sum_{j=0}^{p} \binom{p}{j} p^j = 1 + p^2 + \sum_{j=2}^{p} \binom{p}{j} p^j = 1 + \lambda_1 p^2.$$

où l'on a noté $\nu = \displaystyle\sum_{j=2}^{p} \binom{p}{j} p^{j-2}$ et $\lambda_1 = 1 + \nu$. De plus, p divise ν car, pour tout $j \in [\![2, p]\!]$, p divise $\binom{p}{j}$. Par suite, λ_1 est premier avec p, car si d divise p et λ_1 alors d divise ν et λ_1, puis d divise $1 = \lambda_1 - \nu$, et donc $d = \pm 1$. Supposons ensuite le résultat établi pour un rang $k \in \mathbb{N}^*$ fixé. En utilisant de nouveau la formule du binôme de Newton, ainsi que l'hypothèse de récurrence, il vient,

$$(1+p)^{p^{k+1}} = (1 + \lambda_k p^{k+1})^p = \sum_{j=0}^{p} \binom{p}{j} \lambda_k^j p^{j(k+1)},$$

et donc,

$$(1+p)^{p^{k+1}} = 1 + \lambda_k p^{k+2} + \sum_{j=2}^{p} \binom{p}{j} \lambda_k^j p^{j(k+1)}.$$

Or, pour tout $j \in [\![2, p]\!]$, p^{k+3} divise $\binom{p}{j} p^{j(k+1)}$ (car $k \geq 1$), d'où l'existence de $\mu \in \mathbb{N}^*$ tel que,

$$\sum_{j=2}^{p} \binom{p}{j} \lambda_k^j p^{j(k+1)} = \mu p^{k+3}$$

En notant $\lambda_{k+1} = \lambda_k + \mu p$, on obtient donc,

$$(1+p)^{p^{k+1}} = 1 + \lambda_{k+1} p^{k+2}.$$

avec p qui est premier avec λ_{k+1}. En effet, si d divise p et λ_{k+1} alors d divise $\lambda_k = \lambda_{k+1} - \mu p$, et donc $d = \pm 1$ (car λ_k est premier avec p par hypothèse de récurrence). Ceci achève la récurrence.

Étape n°3. Notons $x = 1 + p + p^\alpha\mathbb{Z}$. Montrons que x est un élément d'ordre $p^{\alpha-1}$ du groupe multiplicatif $(\mathbb{Z}/p^\alpha\mathbb{Z})^*$. Tout d'abord,

$$(1+p)^{p^{\alpha-1}} = 1 + \lambda_{\alpha-1} p^\alpha \equiv 1 \ [p^\alpha].$$

et donc l'ordre de x dans $(\mathbb{Z}/p^\alpha\mathbb{Z})^*$ divise $p^{\alpha-1}$, c'est-à-dire qu'il est de la forme p^k pour $0 \leq k \leq \alpha - 1$. Or, pour tout $k \in [\![0, \alpha-2]\!]$,

$$(1+p)^{p^k} = 1 + \lambda_k p^{k+1} \not\equiv 1 \ [p^\alpha].$$

car sinon, p^α diviserait $\lambda_k p^{k+1}$, et donc p^α diviserait p^{k+1} d'après le lemme de Gauss (car p est premier avec λ_k), ce qui est impossible car $k + 1 < \alpha$. Ainsi, x est d'ordre $p^{\alpha-1}$ dans $(\mathbb{Z}/p^\alpha\mathbb{Z})^*$.

Étape n°4. Fixons $g_0 = k + p\mathbb{Z}$ un générateur du groupe cyclique $(\mathbb{Z}/p\mathbb{Z})^*$. Notons $y = k^{p^{\alpha-1}} + p^\alpha\mathbb{Z}$. Montrons que y est un élément d'ordre $p-1$ du groupe multiplicatif $(\mathbb{Z}/p^\alpha\mathbb{Z})^*$. Tout d'abord, $k^{p^{\alpha-1}} + p\mathbb{Z}$ est d'ordre $p-1$ dans $(\mathbb{Z}/p\mathbb{Z})^*$ car $p^{\alpha-1}$ et $p-1$ sont premiers entre eux (cf. théorème 1.11 de [JR]). Remarquons ensuite que,

- $(k^{p^{\alpha-1}})^{p-1} = k^{p^{\alpha-1}(p-1)} \equiv 1 \ [p^\alpha]$ d'après le théorème de Lagrange, car le groupe multiplicatif $(\mathbb{Z}/p^\alpha\mathbb{Z})^*$ est d'ordre $p^\alpha(p-1)$;
- si $l \in \mathbb{N}^*$ vérifie $(k^{p^{\alpha-1}})^l \equiv 1 \ [p^\alpha]$, a fortiori $(k^{p^{\alpha-1}})^l \equiv 1 \ [p]$, et donc $p-1$ divise l car $k^{p^{\alpha-1}} + p\mathbb{Z}$ est d'ordre $p-1$ dans $(\mathbb{Z}/p\mathbb{Z})^*$.

Des points précédents, on en déduit que y est d'ordre $p-1$ dans $(\mathbb{Z}/p^\alpha\mathbb{Z})^*$.

Étape n°5. Concluons. Puisque x et y sont d'ordre respectif $p^{\alpha-1}$ et $p-1$ (qui sont premiers entre eux) dans le groupe $(\mathbb{Z}/p^\alpha\mathbb{Z})^*$ (qui est commutatif), on en déduit que xy est d'ordre $p^{\alpha-1}(p-1)$ dans $(\mathbb{Z}/p^\alpha\mathbb{Z})^*$ (cf. lemme 2). Or,

$$|(\mathbb{Z}/p^\alpha\mathbb{Z})^*| = p^\alpha(p-1)$$

Donc xy est un générateur du groupe multiplicatif $(\mathbb{Z}/p^\alpha\mathbb{Z})^*$, ce qui prouve finalement que ce dernier est cyclique. □

Complément conseillé.

- Ordre d'un élément dans un groupe : lemme 2.

Commentaires personnels. Pour être tout à fait honnête, je n'avais pas eu le temps de bien préparer ce développement, car je l'ai choisi assez tardivement. Mais je pense que c'est un choix de développement raisonnable.

Décomposition de Dunford

Soient E un espace vectoriel de dimension finie $n \in \mathbb{N}^*$ sur un corps commutatif $\mathbb{K}$. Soit $u \in L(E)$. On suppose que χ_u est scindé sur $\mathbb{K}$. Notons $\lambda_1, \dots, \lambda_p$ ses racines deux à deux distinctes et $\alpha_1, \dots, \alpha_p$ leur multiplicité respective, de telle sorte que,

$$\chi_u = \prod_{k=1}^{p} (X - \lambda_k)^{\alpha_k}.$$

Pour tout $k \in [\![1, p]\!]$, on note $N_k = \mathrm{Ker}((u - \lambda_k Id_E)^{\alpha_k})$. L'espace vectoriel N_k est appelé le sous-espace caractéristique de u associé à la valeur propre λ_k.

[JR] **Lemme 1.** Avec ces notations, $E = \bigoplus_{k=1}^{p} N_k$. De plus, pour tout $k \in [\![1, p]\!]$,

(i) le projecteur π_k sur N_k parallèlement à $\bigoplus_{\substack{j=1 \\ j \neq k}}^{p} N_j$ est un polynôme en u,

(ii) N_k est stable par u et l'endomorphisme induit par $u - \lambda_k Id_E$ sur N_k est nilpotent.

Preuve. Tout d'abord, d'après le théorème de décomposition des noyaux,

$$E = \bigoplus_{k=1}^{p} N_k.$$

De plus, pour tout $k \in [\![1, p]\!]$, le projecteur π_k est un polynôme en u, d'où le point (i). Soit $k \in [\![1, p]\!]$. Tout d'abord, pour tout $P \in \mathbb{K}[X]$, $\mathrm{Ker}(P(u))$ est stable par u car $P(u)$ commute avec u, et donc, pour tout $x \in \mathrm{Ker}(P(u))$,

$$(P(u))(u(x)) = (P(u) \circ u)(x) = (u \circ P(u))(x) = u(P(u)(x)) = 0$$

c'est-à-dire $u(x) \in \mathrm{Ker}(P(u))$. En particulier, N_k est donc stable par u. Notons v_k l'endomorphisme induit par $u - \lambda_k Id_E$ sur N_k. Alors, pour tout $x \in N_k$,

$$v_k^{\alpha_k}(x) = (u - \lambda_k Id_E)^{\alpha_k}(x) = 0$$

et donc $v_k^{\alpha_k} = 0$. Ainsi, v_k est nilpotent. □

Remarque. On peut montrer que l'indice de nilpotence de v_k est la multiplicité de λ_k en tant que racine du polynôme minimal de u (cf. [JR]).

[JR] **Théorème 2 (Décomposition de Dunford).** Il existe un unique couple (d, v) d'endomorphismes de E tel que d soit diagonalisable, v soit nilpotent, d et v commutent et $u = d + v$. De plus, d et v sont des polynômes en u.

Preuve. On sépare la preuve de l'existence et de l'unicité.

Existence. Pour tout $k \in [\![1, p]\!]$, on note $d_k = \lambda_k Id_{N_k}$. On note ensuite,

$$d = \sum_{k=1}^{p} d_k \circ \pi_k.$$

Alors, d est un polynôme en u (car les π_k sont des polynômes en u d'après le lemme 1). De plus, d est diagonalisable. En effet, si $\mathcal{B}_1, \dots, \mathcal{B}_p$ sont des bases respectives de $N_1, \dots, N_p$ alors leur concaténation $\mathcal{B}$ est une base de E dans laquelle la matrice de d est la matrice diagonale,

$$\mathrm{Diag}(\lambda_1 I_{n_1}, \dots, \lambda_p I_{n_p})$$

où l'on a noté, pour tout $k \in [\![1, p]\!]$, $n_k = \dim(N_k)$. Notons ensuite $v = u - d$, de telle sorte que $u = d + v$. Puisque d est un polynôme en u, v est également un polynôme en u. En particulier, d et v commutent. Il reste à vérifier que v est nilpotent. Notons $\alpha = \max_{1 \leq k \leq n} \alpha_k$. Alors, pour tout $x \in E$,

$$v^{\alpha}(x) = (v^{\alpha}) \left(\sum_{k=1}^{p} \pi_k(x) \right) = \sum_{k=1}^{p} (v^{\alpha})(\pi_k(x)) = \sum_{k=1}^{p} (v_k^{\alpha})(\pi_k(x)) = 0$$

car, pour tout $k \in [\![1, p]\!]$, $v_k^{\alpha} = v_k^{\alpha - \alpha_k} \circ v_k^{\alpha_k} = 0$ (d'après la preuve du lemme 1). D'où $v^{\alpha} = 0$, ce qui prouve que v est nilpotent.

Unicité. Soit (d', v') un couple d'endomorphismes de E vérifiant les mêmes conditions que (d, u), c'est-à-dire que d' est diagonalisable, v' est nilpotent, d' et v' commutent et $u = d' + v'$. Alors, d' et u commutent car,

$$d' \circ u = d' \circ d' + d' \circ v' = d' \circ d' + v' \circ d' = u \circ d'$$

et donc d' commute avec d (car d est un polynôme en u). De même, v' et u commutent car,

$$v' \circ u = v' \circ d' + v' \circ v' = d' \circ v' + v' \circ v' = u \circ v'$$

et donc v' commute avec v (car v est un polynôme en u). Par suite,

- $d - d'$ est diagonalisable en tant que somme d'endomorphismes diagonalisables qui commutent. En effet, en prenant une base commune $\mathcal{B}$ de diagonalisation de d et d', on obtient que la matrice de $d - d'$ dans cette base est diagonale, et donc $d - d'$ est bien diagonalisable.

- $v - v'$ est nilpotent en tant que somme d'endomorphismes nilpotent qui commutent. En effet, si $v^a = 0$ et $(v')^b = 0$ alors la formule du binôme de Newton fournit que,

$$(v - v')^{a+b} = \sum_{k=0}^{a+b} \binom{a+b}{k} v^k (v')^{a+b-k} = 0$$

car pour tout $k \in [\![0, a+b]\!]$, si $k \geq a$ alors $v^k = 0$ et sinon $a + b - k > b$ et donc $(v')^{a+b-k} = 0$ (ainsi, tous les termes de la somme ci-dessus sont nuls).

Or, $d - d' = (u - v) - (u - v') = v' - v$. Donc $d - d'$ est diagonalisable et nilpotent, ce qui implique $d - d' = 0$. D'où $d = d'$, ce qui entraîne $v = v'$. □

Remarque. On dit que $u = d + v$ est la décomposition de Dunford de u.

Supposons désormais $\mathbb{K} = \mathbb{C}$.

[JR] **Théorème 3.** Si $u = d+v$ est la décomposition de Dunford de u alors celle de e^u est donnée par $e^u = e^d + e^d(e^v - Id_E)$ avec e^d diagonalisable et $e^d(e^v - Id_E)$ nilpotent.

Preuve. Tout d'abord, puisque d et v commutent, on a bien,

$$e^u = e^d e^v = e^d + e^d(e^v - Id_E).$$

De plus,

- e^d est diagonalisable car d est diagonalisable et e^d est un polynôme en d,

- e^d et $e^d(e^v - Id_E)$ commutent car ce sont des polynômes en u.

Il reste à vérifier que $e^d(e^v - Id_E)$ est nilpotent. Notons q l'indice de nilpotence de v. Alors,

$$e^d(e^v - Id_E) = e^d \sum_{k=1}^{q-1} \frac{v^k}{k!} = e^d v \sum_{k=1}^{q-1} \frac{v^{k-1}}{k!} = v\left(e^d \sum_{k=1}^{q-1} \frac{v^{k-1}}{k!}\right)$$

car les endomorphismes ci-dessus commutent entre eux (ce sont des polynômes en u). De nouveau, puisque ces endomorphismes commutent entre eux, on a,

$$\left[e^d(e^v - Id_E)\right]^q = v^q \left(e^d \sum_{k=1}^{q-1} \frac{v^{k-1}}{k!}\right)^q = 0$$

et donc $e^d(e^v - Id_E)$ est bien nilpotent. □

Application 4. L'endomorphisme u est diagonalisable si et seulement si e^u l'est. [JR]

Preuve. Reprenons les notations précédentes.

- Si u est diagonalisable alors $v = 0$, puis,

$$e^u = e^d + e^d(e^0 - Id_E) = e^d$$

et donc e^u est diagonalisable.

- Réciproquement, supposons que e^u soit diagonalisable. Par unicité de la décomposition de Dunford, on en déduit que $e^d(e^v - Id_E) = 0$, et donc $e^v - Id_E = 0$ en multipliant par e^{-d}, c'est-à-dire $e^v = Id_E$. Or, en notant q l'indice de nilpotence de v, il vient,

$$Id_E = e^v = \sum_{k=0}^{q} \frac{v^k}{k!}$$

(en fait la somme s'arrête à $q - 1$), et donc $P(v) = 0$ où l'on a noté,

$$P = \sum_{k=1}^{q} \frac{X^k}{k!}.$$

Autrement dit, P est un polynôme annulateur de v. Par suite, X^q divise P (car X^q est le polynôme minimal de v), ce qui implique $q = 1$. D'où $v = 0$ (car le polynôme minimal de v est X). Dans la décomposition de Dunford de u, il reste alors $u = d$, ce qui prouve que u est diagonalisable.

□

Compléments conseillés.

- Calcul effectif de la décomposition de Dunford : [JR], Algèbre et géométrie page 614.

- Diagonalisation simultanée : théorème 1.

Commentaires personnels. C'est un développement très (trop ?) classique. Mais il n'est pas nécessaire d'être original pour être admis au concours, donc ce n'est pas un inconvénient majeur. De mon point de vue, c'est un bon choix de développement, puisqu'il est accessible et rentre dans de nombreuses leçons. De plus, la décomposition de Dunford figure sur le programme du concours. Pour la leçon n°155 (sur l'exponentielle de matrices), il vaut mieux admettre le lemme 1 pour avoir le temps de prouver le théorème 3 et l'application 4. Pour les autres leçons, on peut seulement prouver le lemme 1 et le théorème 2.

Décomposition polaire

[JR] **Lemme 1.** Soit $A \in S_n^+(\mathbb{R})$. Il existe une unique matrice $B \in S_n^+(\mathbb{R})$ telle que $A = B^2$. De plus, si A est définie positive, il en est de même pour B.

Preuve. On montre tout d'abord l'existence, puis l'unicité.

Existence. D'après le théorème spectral, il existe $\lambda_1, \dots, \lambda_n \in \mathbb{R}$ et $P \in O_n(\mathbb{R})$ tels que $A = PD^tP$, où $D = \mathrm{Diag}(\lambda_1, \dots, \lambda_n)$. Puisque $A \in S_n^+(\mathbb{R})$, ses valeurs propres sont positives, et on peut donc poser,

$$\Delta = \mathrm{Diag}\Big(\sqrt{\lambda_1}, \dots \sqrt{\lambda_n}\Big), \ B = P\Delta^tP.$$

Alors, $B \in S_n^+(\mathbb{R})$, car B est symétrique et ses valeurs propres sont positives. De plus, $B^2 = P(\Delta^2)^tP = PD^tP = A$, d'où l'existence.

Unicité. Soit $C \in S_n^+(\mathbb{R})$ une matrice telle que $B^2 = C^2$. Soit $L \in \mathbb{R}[X]$ tel que, pour tout $i \in [\![1, n]\!]$, on ait $L(\lambda_i) = \sqrt{\lambda_i}$. On peut construire un tel polynôme par interpolation de Lagrange. Alors,

$$L(D) = \mathrm{Diag}(L(\lambda_1), \dots, L(\lambda_n)) = \mathrm{Diag}\Big(\sqrt{\lambda_1}, \dots \sqrt{\lambda_n}\Big) = \Delta.$$

Puis,

$$B = P\Delta^tP = PL(D)^tP = L(PD^tP) = L(A) = L(C^2) \in \mathbb{R}[C].$$

En particulier, B et C commutent. Sachant que ces deux matrices sont diagonalisables (en tant que matrices symétriques à coefficients réels), le théorème 1 assure qu'elles sont donc simultanément diagonalisables. Autrement dit, il existe $Q \in GL_n(\mathbb{R})$ et $D_1, D_2 \in D_n(\mathbb{R})$ tels que $B = QD_1Q^{-1}$ et $C = QD_2Q^{-1}$. Avec ces notations, il vient :

$$B^2 = QD_1^2Q^{-1}, \ C^2 = QD_2^2Q^{-1}.$$

Or, $B^2 = C^2$. Donc, $QD_1^2Q^{-1} = QD_2^2Q^{-1}$, puis $D_1^2 = D_2^2$ en multipliant par Q^{-1} à gauche et par Q à droite. En notant $\alpha_1, \dots, \alpha_n$ et $\beta_1, \dots, \beta_n$ les coefficients diagonaux respectifs de D_1 et D_2, il vient alors,

$$\mathrm{Diag}(\alpha_1^2, \dots, \alpha_n^2) = \mathrm{Diag}(\beta_1^2, \dots, \beta_n^2),$$

ce qui implique que, pour tout $i \in \{1, \dots, n\}$, on a $\alpha_i = \pm\beta_i$. Or, les réels $\alpha_1, \dots, \alpha_n, \beta_1, \dots, \beta_n$ sont tous positifs. En effet, leurs coefficients diagonaux sont respectivement les valeurs propres de B et C, qui sont bien positives car B et C sont des matrices symétriques positives. Donc, pour tout $i \in \{1, \dots, n\}$, $\alpha_i = \beta_i$, puis $D_1 = D_2$, ce qui prouve finalement que $B = C$. □

[JR] **Théorème 2 (Décomposition polaire).** Notons $S_n^{++}(\mathbb{R})$ l'ensemble des matrices symétriques définies positives. Alors,

$$\begin{array}{ccc} O_n(\mathbb{R}) \times S_n^{++}(\mathbb{R}) & \longrightarrow & GL_n(\mathbb{R}) \\ (M, S) & \longmapsto & MS \end{array}$$

est un homéomorphisme.

Preuve. Tout d'abord, φ est continue en tant qu'application bilinéaire sur un produit d'espaces vectoriels normés de dimension finie. Soit $M \in M_n(\mathbb{R})$.

Étape n°1. Montrons qu'il existe $(O, S) \in O_n(\mathbb{R}) \times S_n^{++}(\mathbb{R})$ tel que $OS = M$. Tout d'abord, ${}^tMM \in S_n(\mathbb{R})$. De plus, pour tout $X \in M_{n,1}(\mathbb{R})$ non nul,

$${}^tX{}^tMMX = {}^t(MX)MX = \|MX\|^2 > 0$$

car $M \in GL_n(\mathbb{R})$. Donc, ${}^tMM \in S_n^{++}(\mathbb{R})$. D'après le lemme précédent, il existe donc $S \in S_n^{++}(\mathbb{R})$ tel que ${}^tMM = S^2$. Posons $O = MS^{-1}$. Alors,

$${}^tOO = {}^t(S^{-1})^tMMS^{-1} = [({}^tS)^{-1}]^tMMS^{-1} = [S^{-1}]^tMMS^{-1} = I_n,$$

car ${}^tMM = S^2$. Donc $(O, S) \in O_n(\mathbb{R}) \times S_n^{++}(\mathbb{R})$ et $M = OS$.

Étape n°2. Montrons ensuite l'unicité d'un tel couple. Pour cela, considérons $(O', S') \in O_n(\mathbb{R}) \times S_n^{++}(\mathbb{R})$ tel que $O'S' = M$. Alors,

$$S^2 = {}^tMM = {}^t(O'S')(O'S') = [{}^tS']{}^tO'O'S' = S'I_nS' = (S')^2$$

ce qui implique $S = S'$ d'après le lemme. Ainsi, φ est bijective.

Étape n°3. Montrons finalement que φ^{-1} est continue. Pour cela, nous allons utiliser la caractérisation séquentielle de la continuité. Soit $M \in GL_n(\mathbb{R})$. On note $(O, S) = \varphi^{-1}(M)$. Soit $(M_p)_{p\in\mathbb{N}}$ une suite de matrices inversibles qui converge vers M dans $GL_n(\mathbb{R})$. Pour tout $p \in \mathbb{N}$, on note,

$$(O_p, S_p) = \varphi^{-1}(M_p).$$

Nous allons montrer que $(O_p)_{p\in\mathbb{N}}$ et $(S_p)_{p\in\mathbb{N}}$ convergent respectivement vers O et S. Sachant que $(O_p)_{p\in\mathbb{N}}$ est une suite à termes dans $O_n(\mathbb{R})$ qui est compact, il existe donc une application φ strictement croissante de $\mathbb{N}$ dans $\mathbb{N}$ telle que $(O_{\varphi(p)})_{p\in\mathbb{N}}$ converge vers une matrice O' de $O_n(\mathbb{R})$. Or, pour tout $p \in \mathbb{N}$,

$$S_p = {}^tO_pM_p.$$

Par continuité du produit matriciel et de la transposée, on en déduit que $(S_{\varphi(p)})_{p\in\mathbb{N}}$ converge vers la matrice S' définie par $S' = {}^tO'M$. Par ailleurs, S' est inversible en tant que produit de matrices inversibles, et S' appartient à $\overline{S_n^{++}(\mathbb{R})}$ en tant que limite d'une suite à termes dans $S_n^{++}(\mathbb{R})$. Or,

$$\overline{S_n^{++}(\mathbb{R})} = S_n^{+}(\mathbb{R}).$$

Donc, $S' \in GL_n(\mathbb{R}) \cap S_n^+(\mathbb{R})$. Sachant que, $GL_n(\mathbb{R}) \cap S_n^+(\mathbb{R}) = S_n^{++}(\mathbb{R})$, il s'en suit que $S' \in S_n^{++}(\mathbb{R})$. On a donc,

$$O' \in O_n(\mathbb{R}),\ S' \in S_n^{++}(\mathbb{R}),\ M = O'S',$$

ce qui prouve que $(O', S') = \varphi^{-1}(M) = (O, S)$. Finalement, $(O_p)_{p\in\mathbb{N}}$ est une suite à termes dans un compact possédant une unique valeurs d'adhérence qui est O, donc elle converge vers O par la proposition 2. Or, pour tout $p \in \mathbb{N}$,

$$S_p = O_p^{-1}M_p = {}^tO_pM_p.$$

Par continuité du produit matriciel et de la transposée, on en conclut que la suite $(S_{\varphi(p)})_{p\in\mathbb{N}}$ converge vers ${}^tOM = S$. Ainsi, $(O_p)_{p\in\mathbb{N}}$ et $(S_p)_{p\in\mathbb{N}}$ convergent respectivement vers O et S, ce qui prouve que φ^{-1} est continue en M. □

Compléments conseillés.

- Compacité : proposition 2.
- Compacité de $O_n(\mathbb{R})$: théorème 3.
- Décomposition polaire sur $M_n(\mathbb{R})$: théorème 1.
- Diagonalisation simultanée : théorème 1.
- Matrices symétriques : sous-section 3.18.
- Théorème spectral : corollaire 4.

Commentaires personnels. C'est un développement très (trop ?) classique. Mais il n'est pas nécessaire d'être original pour être admis au concours, donc ce n'est pas un inconvénient majeur. On utilise beaucoup de résultats au programme, donc vous devez être irréprochable sur ces derniers. À noter qu'il se recase dans de nombreuses leçons. De plus, la décomposition polaire figure sur le programme du concours.

Dénombrement des polynômes irréductibles unitaires sur un corps fini

[JR] **Théorème 1.** Soient p un nombre premier et $n \in \mathbb{N}^*$. Pour tout $d \in \mathbb{N}^*$, on note $\mathcal{U}_d(p)$ l'ensemble des polynômes unitaires irréductibles de degré d dans $\mathbb{F}_p[X]$ et $I_d(p)$ le cardinal de $\mathcal{U}_d(p)$. Alors,

(i) $X^{p^n} - X = \prod_{d|n} \prod_{P \in \mathcal{U}_d(p)} P$;

(ii) $I_n(p) = \frac{1}{n} \sum_{d|n} \mu\left(\frac{n}{d}\right) p^d$;

(iii) $I_n(p) \underset{n \to +\infty}{\sim} \frac{p^n}{n}$.

Preuve. Montrons tout d'abord le point (i). Pour cela, on montre tout d'abord le résultat suivant : pour tout $d \in \mathbb{N}^*$, pour tout $P \in \mathcal{U}_d(p)$,

$$P \mid X^{p^n} - 1 \iff d \mid n.$$

Soient $d \in \mathbb{N}^*$ et $P \in \mathcal{U}_d(p)$. On procède par double implication.

• Supposons que P divise $X^{p^n} - 1$. Soit $\mathbb{K}$ un corps de décomposition de $X^{p^n} - 1$ sur $\mathbb{F}_p$. Alors, P est scindé sur $\mathbb{K}$. Soit x une racine de P dans $\mathbb{K}$. D'après la formule de multiplicativité des degrés,

$$[\mathbb{K} : \mathbb{F}_p] = [\mathbb{K} : \mathbb{F}_p(x)][\mathbb{F}_p(x) : \mathbb{F}_p].$$

De plus, $\mathbb{F}_p(x)$ étant un corps de rupture de P sur $\mathbb{F}_p$, il vient,

$$[\mathbb{F}_p(x) : \mathbb{F}_p] = \deg(P) = d.$$

D'autre part, $[\mathbb{K} : \mathbb{F}_p] = n$ (d'après le théorème 1). Donc, $n = [\mathbb{K} : \mathbb{F}_p(x)]d$, ce qui prouve que d divise n.

• Réciproquement, on suppose que d divise n, puis on note $\mathbb{L} = \mathbb{F}_p[X]/(P)$. Alors, $\mathbb{L}$ est un corps de rupture de P sur $\mathbb{F}_p$. En particulier, $\mathbb{L}$ est un espace vectoriel de dimension d sur $\mathbb{F}_p$, et donc $\mathbb{L}$ est de cardinal p^d. Par conséquent, le groupe multiplicatif $\mathbb{L}^*$ est d'ordre $p^d - 1$, ce qui entraîne,

$$\forall x \in \mathbb{L}^*,\ x^{p^d - 1} = 1$$

d'après le théorème de Lagrange. En particulier, en notant $\overline{X}$ la classe de X modulo P, il vient $\overline{X}^{p^d} = \overline{X}$. Par récurrence immédiate, il s'en suit que, pour tout $k \in \mathbb{N}$, $\overline{X}^{p^{kd}} = \overline{X}$, ce qui implique $\overline{X}^{p^n} = \overline{X}$, car d divise n. Autrement dit, P divise $X^{p^n} - X$ dans $\mathbb{F}_p[X]$.

Ainsi, pour tout $d \in \mathbb{N}^*$, pour tout $P \in \mathcal{U}_d(p)$,

$$P \mid X^{p^n} - 1 \iff d \mid n.$$

En particulier,

$$\prod_{d|n} \prod_{P \in \mathcal{U}_d(p)} P \mid X^{p^n} - 1.$$

En effet, ces polynômes sont deux à deux premiers entre eux (puisqu'ils sont unitaires, irréductibles et deux à deux distincts) et divisent tous $X^{p^n} - 1$ (voir le lemme 1 pour plus de détails). D'autre part, $X^{p^n} - 1$ est sans facteur carré, puisqu'il est premier avec son polynôme dérivé. Sa décomposition en produit d'irréductibles dans $\mathbb{F}_p[X]$ peut donc s'écrire :

$$X^{p^n} - 1 = \prod_{k=1}^{r} P_k$$

avec $P_1, \ldots, P_r \in \mathbb{F}_p[X]$ des polynômes unitaires, irréductibles et deux à deux distincts. Et d'après ce qui précède, pour tout $k \in [\![1, r]\!]$, le degré de P_k divise d. Par suite,

$$X^{p^n} - 1 = \prod_{k=1}^{r} P_k \mid \prod_{d|n} \prod_{P \in \mathcal{U}_d(p)} P$$

Les deux polynômes ci-dessus étant unitaires et associés, il vient finalement,

$$X^{p^n} - 1 = \prod_{d|n} \prod_{P \in \mathcal{U}_d(p)} P.$$

Passons au point (ii). En passant aux degrés dans l'égalité précédente, il vient,

$$p^n = \sum_{d|n} \sum_{P \in \mathcal{U}_d(p)} d = \sum_{d|n} d I_d(p).$$

Et donc, d'après la formule d'inversion de Möbius (théorème 2),

$$n I_n(p) = \sum_{d|n} \mu\left(\frac{n}{d}\right) p^d.$$

Passons finalement au point (iii). Supposons $n \geq 2$, et notons,

$$r_n = \sum_{\substack{d|n \\ d \neq n}} \mu\left(\frac{n}{d}\right) p^d.$$

Par inégalité triangulaire,

$$|r_n| \leq \sum_{\substack{d|n \\ d \neq n}} p^d$$

car μ est à valeurs dans $\{-1, 0, 1\}$. Or, les diviseurs stricts de n sont inférieurs ou égaux à $\lfloor \frac{n}{2} \rfloor$, puisque si d est un diviseur strict de n, il existe $l \geq 2$ tel que $n = ld \geq 2d$. Par suite,

$$|r_n| \leq \sum_{k=1}^{\lfloor \frac{n}{2} \rfloor} p^k = p \frac{p^{\lfloor \frac{n}{2} \rfloor} - 1}{p-1} \leq p^{\lfloor \frac{n}{2} \rfloor + 1} \leq p^{\frac{n}{2}+1}$$

car $p - 1 \geq 1$. Finalement,

$$\left| \frac{nI_n(p)}{p^n} - 1 \right| = \frac{|r_n|}{p^n} \leq \frac{1}{p^{\frac{n}{2}-1}} \underset{n \to +\infty}{\longrightarrow} 0.$$

Autrement dit,

$$I_n(p) \underset{n \to +\infty}{\sim} \frac{p^n}{n}.$$

□

Compléments conseillés.

- Construction des corps finis : théorème 1.
- Fonction de Möbius : théorème 2.
- Un lemme d'arithmétique bien utile : lemme 1.

Commentaires personnels. Développement technique, qui illustre bien les liens étroits entre la théorie des corps et l'arithmétique des polynômes, ce qui permet de le présenter dans de nombreuses leçons. Je le recommande.

Déterminant de Gram

Soit E un espace préhilbertien réel.

Définition 1. Soient $x_1, \dots, x_n \in E$. Notons $\mathcal{M}_G(x_1, \dots, x_n)$ la matrice de terme général $(x_i \mid x_j)$. On appelle matrice de Gram de $x_1, \dots, x_n$ la matrice $\mathcal{M}_G(x_1, \dots, x_n)$ et déterminant de Gram son déterminant, noté $G(x_1, \dots, x_n)$.

Proposition 2. Si $x_1, \dots, x_n \in E$ alors $G(x_1, \dots, x_n) \neq 0$ si et seulement si la famille $(x_1, \dots, x_n)$ est libre.

Preuve. Soit $(e_1, \dots, e_r)$ une base orthonormée de $\mathrm{Vect}(x_1, \dots, x_n)$. Notons,

$$A = ((x_j \mid e_i))_{\substack{1 \leq i \leq r \\ 1 \leq j \leq n}} \in M_{r,n}(\mathbb{R})$$

Autrement dit, A est la matrice dont la j-ième colonne contient les coordonnées du vecteur x_j dans la base $(e_1, \dots, e_r)$. Alors, pour tout $(i,j) \in [\![1,n]\!] \times [\![1,r]\!]$,

$$({}^tAA)_{i,j} = \sum_{k=1}^{r} (x_i \mid e_k)(x_j \mid e_k) = \left(x_i \,\middle|\, \sum_{k=1}^{r} (x_j \mid e_k) e_k \right) = (x_i \mid x_j)$$

et donc ${}^tAA = \mathcal{M}_G(x_1, \dots, x_n)$. Donc, $\mathcal{M}_G(x_1, \dots, x_n) \in S_n(\mathbb{R})$ et,

$$\forall X \in M_{n,1}(\mathbb{R}),\ {}^tX {}^tAAX = {}^t(AX)AX = \|AX\|^2 \geq 0.$$

D'où, $\mathcal{M}_G(x_1, \dots, x_n) \in S_n^+(\mathbb{R})$. De plus, $\mathcal{M}_G(x_1, \dots, x_n) \in S_n^{++}(\mathbb{R})$, si et seulement si l'inégalité précédente est stricte dès que $X \neq 0$, ce qui équivaut à $\mathrm{Ker}(A) = \{0\}$, ou encore $A \in GL_n(\mathbb{R})$. Sachant que,

$$G(x_1, \dots, x_n) \neq 0 \iff \mathcal{M}_G(x_1, \dots, x_n) \in S_n^{++}(\mathbb{R})$$

et $\mathrm{rg}(A) = \mathrm{rg}(x_1, \dots, x_n)$, on en déduit que $G(x_1, \dots, x_n) \neq 0$ si et seulement si la famille $(x_1, \dots, x_n)$ est libre. □

Remarque. En fait, on a même $\mathrm{rg}(\mathcal{M}_G(x_1, \dots, x_n)) = \mathrm{rg}(x_1, \dots, x_n)$.

[XG] **Théorème 3.** Soit F un sous-espace vectoriel de E de dimension finie, muni d'une base $(e_1, \dots, e_n)$. Soit $x \in E$. Alors,

$$d(x,F)^2 = \frac{G(e_1, \dots, e_n, x)}{G(e_1, \dots, e_n)}.$$

Preuve. Puisque l'espace F est de dimension finie, le projeté orthogonal de x sur F existe. Notons le y. Alors,

(i) $d(x,F) = \|x - y\|$,

(ii) pour tout $k \in [\![1,n]\!]$,

$$(x \mid e_k) = (x - y + y \mid e_k) = (x - y \mid e_k) + (y \mid e_k) = (y \mid e_k)$$

car $x - y \in F^\perp$,

(iii) $\|x\|^2 = \|y + x - y\|^2 = \|y\|^2 + \|x - y\|^2$, car $x - y$ et y sont orthogonaux.

Ainsi,

$$\mathcal{M}_G(e_1, \dots, e_n, x) = \begin{pmatrix} (e_1 \mid e_1) & \dots & (e_1 \mid e_n) & (e_1 \mid y) \\ \vdots & & \vdots & \vdots \\ (e_n \mid e_1) & \dots & (e_n \mid e_n) & (e_n \mid y) \\ (y \mid e_1) & \dots & (y \mid e_n) & \|y\|^2 + \|x - y\|^2 \end{pmatrix}.$$

En utilisant la linéarité du déterminant par rapport à la dernière colonne de cette matrice, il vient $G(e_1, \dots, e_n, x) = \alpha + \beta$ où l'on a noté,

$$\begin{cases} \alpha = \begin{vmatrix} (e_1 \mid e_1) & \dots & (e_1 \mid e_n) & (e_1 \mid y) \\ \vdots & & \vdots & \vdots \\ (e_n \mid e_1) & \dots & (e_n \mid e_n) & (e_n \mid y) \\ (y \mid e_1) & \dots & (y \mid e_n) & \|y\|^2 \end{vmatrix} \\ \beta = \begin{vmatrix} (e_1 \mid e_1) & \dots & (e_1 \mid e_n) & 0 \\ \vdots & & \vdots & \vdots \\ (e_n \mid e_1) & \dots & (e_n \mid e_n) & 0 \\ (y \mid e_1) & \dots & (y \mid e_n) & \|x - y\|^2 \end{vmatrix}. \end{cases}$$

Or, $\alpha = G(e_1, \dots, e_n, y) = 0$ car la famille $(e_1, \dots, e_n, y)$ est liée. De plus, en développant le déterminant par rapport à la dernière colonne de β, on obtient,

$$\beta = \|x - y\|^2 G(e_1, \dots, e_n) = d(x,F)^2 G(e_1, \dots, e_n).$$

Sachant que $G(e_1, \dots, e_n) \neq 0$, car $(e_1, \dots, e_n)$ est libre, on en conclut que,

$$d(x,F)^2 = \frac{G(e_1, \dots, e_n, x)}{G(e_1, \dots, e_n)}.$$

□

Dans [XG], la preuve de l'inégalité de Hadamard (théorème ci-dessous) n'utilise pas le théorème ci-dessus. Par contre, dans la preuve d'Antoine Barrier (qu'on peut trouver ici), les inégalités de Hadamard sont déduites de ce théorème.

Théorème 4 (Inégalité de Hadamard).

(i) Soient $x_1, \dots, x_n \in E$. Alors, $G(x_1, \dots, x_n) \leq \prod_{k=1}^{n} \|x_k\|^2$.

(ii) Soient $x_1, \dots, x_n \in \mathbb{R}^n$. Alors, $|\det(x_1, \dots, x_n)| \leq \prod_{k=1}^{n} \|x_k\|_2$, où $\|\cdot\|_2$ désigne la norme euclidienne usuelle sur $\mathbb{R}^n$.

Dans ces deux points, on a égalité si et seulement si la famille $(x_k)_{1\leq k\leq n}$ est orthogonale ou si l'un des vecteurs est nul.

Preuve.

(i) Procédons par récurrence sur n. Pour $n = 1$, l'égalité est toujours vérifiée et la famille (x_1) est toujours orthogonale. Supposons ensuite le résultat établi pour un rang $n \in \mathbb{N}^*$ fixé, et fixons $x_1, \dots, x_{n+1} \in E$. Si la famille $(x_1, \dots, x_{n+1})$ est liée alors,

$$G(x_1, \dots, x_{n+1}) = 0 \leq \prod_{k=1}^{n+1} \|x_k\|_2.$$

Sinon, $(x_1, \dots, x_n, x_{n+1})$ est libre. Notons ensuite $F = \text{Vect}(x_1, \dots, x_n)$. Puisque $(x_1, \dots, x_n)$ est libre (en tant que famille extraite d'une famille libre), c'est donc une base de F. Ainsi, d'après le théorème précédent,

$$G(x_1, \dots, x_n, x_{n+1}) = \|x_{n+1} - y\|^2 G(x_1, \dots, x_n).$$

où l'on a noté y le projeté orthogonal de x_{n+1} sur F. Par hypothèse de récurrence, il s'en suit que,

$$G(x_1, \dots, x_n, x_{n+1}) \leq \|x_{n+1} - y\|^2 \prod_{k=1}^{n} \|x_k\|^2,$$

avec égalité si et seulement si la famille $(x_1, \dots, x_n)$ est orthogonale. Or, puisque $x_{n+1} - y$ et y sont orthogonaux, il vient,

$$\|x_{n+1}\|^2 = \|y + x_{n+1} - y\|^2 = \|y\|^2 + \|x_{n+1} - y\|^2 \geq \|x_{n+1} - y\|^2$$

avec égalité si et seulement si $y = 0$, c'est-à-dire $x_{n+1} \in F^\perp$. Ainsi,

$$G(x_1, \dots, x_n, x_{n+1}) \leq \prod_{k=1}^{n+1} \|x_k\|^2,$$

avec égalité si et seulement si la famille $(x_1, \dots, x_{n+1})$ est orthogonale.

(ii) Notons A la matrice dont la j-ième colonne contient les coordonnées du vecteur x_j dans la base canonique de $\mathbb{R}^n$. On a vu dans la preuve de la proposition 2 que ${}^tAA = \mathcal{M}_G(x_1, \dots, x_n)$. En particulier,

$$G(x_1, \dots, x_n) = \det({}^tAA) = \det({}^tA)\det(A) = \det(A)^2$$

et donc, en appliquant le point précédent,

$$\det(A)^2 = G(x_1, \dots, x_n) \leq \prod_{k=1}^{n} \|x_k\|_2^2.$$

Finalement, on en conclut que,

$$|\det(x_1, \dots, x_n)| \leq \prod_{k=1}^{n} \|x_k\|_2.$$

□

Remarque. Ces théorèmes restent vrais pour un espace préhilbertien complexe.

Complément conseillé.

- Document de Jérôme Von Buhren sur le déterminant de Gram.

Commentaires personnels. Pour être tout à fait honnête, je n'avais pas eu le temps de bien préparer ce développement, car je l'ai choisi assez tardivement. Mais je pense que c'est un choix de développement raisonnable.

Enveloppe convexe des matrices orthogonales

Les lemmes suivants sont fortement inspirés de la version de Louis Ducongé (qu'on peut trouver ici).

[IP] **Lemme 1.** Soient H un $\mathbb{R}$-espace de Hilbert et C une partie convexe fermée non vide. Pour tout $x \in H \setminus C$, il existe une application $f : H \to \mathbb{R}$ linéaire continue telle que $f(x) > \sup_{y \in C} f(y)$.

Preuve. Soit $x \in H \setminus C$. Notons a le projeté orthogonal de x sur C (qui existe car C est une partie convexe fermée non vide), puis considérons l'application,

$$\begin{array}{rccc} f & : & H & \longrightarrow & \mathbb{R} \\ & & y & \longmapsto & (x-a \mid y) \end{array}$$

Montrons que f convient à notre problème. Tout d'abord, f est linéaire, avec pour tout $y \in H$, en vertu de l'inégalité de Cauchy-Schwarz,

$$|f(y)| \leq \|x-a\|\|y\|$$

ce qui prouve que f est aussi continue. Notons ensuite que, d'après le théorème de projection sur un convexe fermé, pour tout $y \in C$, $(x-a \mid y-a) \leq 0$. Or, pour tout $y \in C$,

$$(x-a \mid y-a) = (x-a \mid y) - (x-a \mid a) = f(y) - f(a)$$

Donc, pour tout $y \in C$, $f(y) \leq f(a)$. Par suite, pour tout $y \in C$,

$$f(x) - f(y) \geq f(x) - f(a) = (x-a \mid x) - (x-a \mid a) = \|x-a\|^2 > 0$$

(la dernière inégalité est stricte car $x \notin C$), et donc $f(x) > \sup_{y \in C} f(y)$. □

Remarque. Il s'agit d'une version affaiblie du théorème de Hahn-Banach.

Lemme 2. Soient H un $\mathbb{R}$-espace de Hilbert, A une partie non vide de H et $x \in H$. Alors,

$$x \in \overline{\mathrm{Conv}(A)} \iff \left(\forall f \in H^*,\ f(x) \leq \sup_{y \in \mathrm{Conv}(A)} f(y) \right).$$

Preuve. Supposons que $x \in \overline{\mathrm{Conv}(A)}$. Soit $f \in H^*$. Montrons que,

$$f(x) \leq \sup_{y \in \mathrm{Conv}(A)} f(y).$$

Tout d'abord, puisque $x \in \overline{\mathrm{Conv}(A)}$, il existe une suite $(x_n)_{n \in \mathbb{N}}$ d'éléments de $\mathrm{Conv}(A)$ telle que $(x_n)_{n \in \mathbb{N}}$ converge vers x. Alors, pour tout $n \in \mathbb{N}$,

$$f(x_n) \leq \sup_{y \in \mathrm{Conv}(A)} f(y)$$

car $x_n \in \mathrm{Conv}(A)$. Or, $\lim_{n \to +\infty} f(x_n) = f(x)$ puisque f est continue. D'où, en faisant tendre n vers l'infini dans l'inégalité précédente :

$$f(x) \leq \sup_{y \in \mathrm{Conv}(A)} f(y).$$

Pour le sens réciproque, on prouve la contraposée. Supposons $x \notin \overline{\mathrm{Conv}(A)}$. Alors, le lemme précédent assure qu'il existe $f \in H^*$ telle que,

$$f(x) > \sup_{y \in \overline{\mathrm{Conv}(A)}} f(y).$$

Or,

$$\sup_{y \in \overline{\mathrm{Conv}(A)}} f(y) \geq \sup_{y \in \mathrm{Conv}(A)} f(y)$$

car $\mathrm{Conv}(A) \subset \overline{\mathrm{Conv}(A)}$. D'où,

$$f(x) > \sup_{y \in \mathrm{Conv}(A)} f(y).$$

□

Soit $n \in \mathbb{N}^*$. Notons $\langle \cdot \mid \cdot \rangle_2$ le produit scalaire sur $M_{n,1}(\mathbb{R})$ défini par,

$$\forall X, Y \in M_{n,1}(\mathbb{R}),\ \langle X \mid Y \rangle_2 = {}^tXY.$$

Rappelons à toutes fins utiles que ce dernier vérifie, pour tout $M \in M_n(\mathbb{R})$ et pour tous $X, Y \in M_{n,1}(\mathbb{R})$:

$$\langle MX \mid Y \rangle_2 = \langle X \mid {}^tMY \rangle_2$$

car $\langle AX \mid Y \rangle_2 = {}^t(AX)Y = {}^tX{}^tAY = \langle X \mid {}^tAY \rangle_2$.

Notons ensuite $\|\cdot\|_2$ la norme sur $M_{n,1}(\mathbb{R})$ induite par ce produit scalaire, puis $|||\cdot|||$ la norme subordonnée à cette norme, c'est-à-dire la norme définie par,

$$\forall M \in M_n(\mathbb{R}),\ |||M||| = \sup \left\{ \frac{\|MX\|_2}{\|X\|_2} \;\middle|\; X \in M_{n,1}(\mathbb{R}),\ X \neq 0 \right\}.$$

[IP] **Théorème 3.** L'enveloppe convexe de $O_n(\mathbb{R})$ est la boule unité de l'espace vectoriel normé $(M_n(\mathbb{R}), |||\cdot|||)$.

Preuve. Notons B la boule unité de $O_n(\mathbb{R})$. Montrons que $\mathrm{Conv}(O_n(\mathbb{R})) = B$.

• Montrons que $\mathrm{Conv}(O_n(\mathbb{R})) \subset B$. Tout d'abord, rappelons que,

$$\forall O \in O_n(\mathbb{R}),\ |||O||| = 1.$$

En effet, si $O \in O_n(\mathbb{R})$ alors, pour tout $X \in M_{n,1}(\mathbb{R})$,

$$\|OX\|_2 = {}^t(OX)OX = {}^tX{}^tOOX = {}^tXI_nX = {}^tXX = \|X\|_2$$

et donc,

$$|||O||| = \sup\left\{\frac{\|OX\|_2}{\|X\|_2}\ \middle|\ X \in M_{n,1}(\mathbb{R}),\ X \neq 0\right\} = 1.$$

Considérons ensuite $M \in \mathrm{Conv}(O_n(\mathbb{R}))$. Alors, il existe $\lambda_1, \dots, \lambda_p \in \mathbb{R}_+$ et $O_1, \dots, O_p \in O_n(\mathbb{R})$ tels que,

$$\begin{cases} M = \lambda_1 O_1 + \cdots + \lambda_p O_p, \\ \lambda_1 + \cdots + \lambda_p = 1. \end{cases}$$

Avec ces notations, il vient,

$$|||M||| = \left|\left|\left|\sum_{i=1}^{p} \lambda_i O_i\right|\right|\right| \leq \sum_{i=1}^{p} |\lambda_i| \cdot |||O_i||| = \sum_{i=1}^{p} \lambda_i = 1$$

ce qui prouve que $M \in B$. Ainsi, $\mathrm{Conv}(O_n(\mathbb{R})) \subset B$.

• Montrons ensuite que $B \subset \mathrm{Conv}(O_n(\mathbb{R}))$. Soit $M \in B$. D'après le corollaire 4, $\mathrm{Conv}(O_n(\mathbb{R}))$ est compact. En particulier, $\mathrm{Conv}(O_n(\mathbb{R}))$ est fermé et donc,

$$\overline{\mathrm{Conv}(O_n(\mathbb{R}))} = \mathrm{Conv}(O_n(\mathbb{R})).$$

Pour conclure, il suffit donc de montrer que $M \in \overline{\mathrm{Conv}(O_n(\mathbb{R}))}$. Et pour cela, le lemme 2 assure qu'il suffit de montrer que,

$$\forall f \in (M_n(\mathbb{R}))^*,\ f(M) \leq \sup_{O \in O_n(\mathbb{R})} f(O).$$

Soit $f \in (M_n(\mathbb{R}))^*$. Notons $\langle \cdot \mid \cdot \rangle$ le produit scalaire défini par,

$$\forall A, B \in M_n(\mathbb{R}),\ \langle A \mid B \rangle = \mathrm{Tr}({}^tAB).$$

Avec ces notations, $(M_n(\mathbb{R}), \langle \cdot \mid \cdot \rangle)$ forme un espace euclidien. Par conséquent, le théorème de Riesz (son énoncé restreint à un espace euclidien suffit) assure l'existence de $A \in M_n(\mathbb{R})$ tel que,

$$\forall B \in M_n(\mathbb{R}),\ f(B) = \mathrm{Tr}({}^tAB).$$

Écrivons alors la décomposition polaire de A : soit $(O_0, S_0) \in O_n(\mathbb{R}) \times S_n^+(\mathbb{R})$ tel que $A = O_0S_0$ (un tel couple existe d'après le théorème 1). En particulier, $S_0 \in S_n(\mathbb{R})$, et donc le théorème spectral assure l'existence de $\lambda_1, \dots, \lambda_n \in \mathbb{R}$ et $X_1, \dots, X_n \in M_{n,1}(\mathbb{R})$ tels que,

- $\lambda_1, \dots, \lambda_n$ sont les valeurs propres de S;
- $X_1, \dots, X_n$ sont des vecteurs propres de S, respectivement associés aux valeurs propres $\lambda_1, \dots, \lambda_n$;
- $(X_1, \dots, X_n)$ est une base orthonormée de $(M_{n,1}(\mathbb{R}), \|\cdot\|_2)$.

De plus, $\lambda_1, \dots, \lambda_n \in \mathbb{R}_+$ car $S \in S_n^+(\mathbb{R})$. Montrons finalement que,

$$\mathrm{Tr}({}^tAM) \leq \sup_{O \in O_n(\mathbb{R})} \mathrm{Tr}({}^tAO).$$

Puisque,

$$\sup_{O \in O_n(\mathbb{R})} \mathrm{Tr}({}^tAO) \geq \mathrm{Tr}({}^tAO_0) = \mathrm{Tr}({}^tS_0{}^tO_0O_0) = \mathrm{Tr}(S_0) = \sum_{i=1}^{n} \lambda_i$$

il suffit donc de montrer que $\mathrm{Tr}({}^tAM) \leq \sum_{i=1}^{n} \lambda_i$.

Puisque $(X_1, \dots, X_n)$ est une base orthonormée, le lemme 1 fournit,

$$\mathrm{Tr}({}^tAM) = \sum_{i=1}^{n} \langle {}^tAMX_i \mid X_i \rangle_2.$$

Ensuite, une propriété classique de la transposée donne,

$$\mathrm{Tr}({}^tAM) = \sum_{i=1}^{n} \langle MX_i \mid AX_i \rangle_2$$

Sachant que $X_1, \dots, X_n$ sont des vecteurs propres de S_0 respectivement associés aux valeurs propres $\lambda_1, \dots, \lambda_n$ et que $A = O_0S_0$, il vient,

$$\mathrm{Tr}({}^tAM) = \sum_{i=1}^{n} \langle MX_i \mid O_0S_0X_i \rangle_2 = \sum_{i=1}^{n} \langle MX_i \mid \lambda_i O_0X_i \rangle_2.$$

On applique ensuite l'inégalité de Cauchy-Schwarz, ce qui donne,

$$\mathrm{Tr}({}^tAM) \leq \sum_{i=1}^{n} \lambda_i \|MX_i\|_2 \|OX_i\|_2$$

(car $\lambda_1, \dots, \lambda_n \in \mathbb{R}_+$). Or, par définition de $|||\cdot|||$, pour tout $i \in [\![1, n]\!]$,

$$\begin{cases} \|MX_i\|_2 \leq |||M|||\|X_i\|_2 \leq \|X_i\|_2 = 1, \\ \|O_0X_i\|_2 \leq |||O_0|||\|X_i\|_2 \leq \|X_i\|_2 = 1. \end{cases}$$

car $M \in B$ et $O_0 \in O_n(\mathbb{R})$. D'où,

$$\mathrm{Tr}({}^tAM) \leq \sum_{i=1}^{n} \lambda_i$$

ce qui prouve finalement que $M \in \mathrm{Conv}(O_n(\mathbb{R}))$. D'où $B \subset \mathrm{Conv}(O_n(\mathbb{R}))$.

Par double inclusion, on en conclut que $\mathrm{Conv}(O_n(\mathbb{R})) = B$. □

Remarque. L'énoncé du théorème de Riesz restreint à un espace euclidien se démontre de façon plus élémentaire (voir le théorème 1).

Compléments conseillés.

- Décomposition polaire : théorème 2.
- Points extrémaux de l'enveloppe convexe de $O_n(\mathbb{R})$ (Zuily-Queffélec page 223).
- Théorème de Carathéodory : corollaire 4.
- Théorème de représentation de Riesz (en dimension finie) : théorème 1.
- Trace et base orthonormée : proposition 1.

Commentaires personnels. Pour être tout à fait honnête, je n'avais pas eu le temps de bien préparer ce développement, car je l'ai choisi assez tardivement. Je pense que c'est un développement très (trop ?) long, qui utilise beaucoup de résultats, dont un qui est hors-programme me semble-t-il (le théorème de Carathéodory). C'est un développement qui vous demandera probablement beaucoup de temps de préparation, mais ce temps peut être rentabilisé par le nombre important de leçons dans lesquelles rentre ce développement. Mais si ce développement vous semble trop difficile, n'hésitez pas à l'abandonner.

Générateurs de $GL_n(\mathbb{K})$ et $SL_n(\mathbb{K})$

Soient $\mathbb{K}$ un corps et $n \in \mathbb{N}^*$. Pour tout $(i,j) \in [\![1,n]\!]^2$, on note $E_{i,j}$ la matrice de $M_n(\mathbb{K})$ dont tous les termes sont nuls sauf le terme de position (i,j).

Définition 1. Pour tout $(i,j) \in [\![1,n]\!]^2$ tel que $i \neq j$, et pour tout $\lambda \in \mathbb{K}$, on note :

$$T_{i,j}(\lambda) = I_n + \lambda E_{i,j}$$

On appelle matrice de transvection les matrices de la forme précédente.

Définition 2. Pour tout $\lambda \in \mathbb{K}$, on note :

$$S(\lambda) = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & 1 & 0 \\ 0 & \cdots & 0 & \lambda \end{pmatrix}$$

On appelle matrice de dilatation les matrices de la forme précédente.

[JR] **Théorème 3.** Le groupe $GL_n(\mathbb{K})$ est engendré par les matrices de transvection et les matrices de dilatation. Le groupe $SL_n(\mathbb{K})$ est engendré par les matrices de transvection et les matrices de dilatation.

Ce développement n'est pas encore rédigé (c'est le seul du document à ne pas l'être).

Générateurs du groupe orthogonal d'un espace vectoriel euclidien

Soit E un espace euclidien. Notons $n = \dim(E)$.

Lemme 1. Soient $u, v \in E$ tels que $u \neq v$ et $\|u\| = \|v\|$. Alors, il existe une unique réflexion σ de E telle que $\sigma(u) = v$.

Preuve. Pour le démontrer, procédons par analyse-synthèse.

Analyse. Supposons qu'il existe un hyperplan H de E telle que la réflexion σ par rapport à H vérifie $\sigma(u) = v$. Soit $(a, b) \in H \times H^\perp$ tel que $u = a + b$. Alors, $v = \sigma(u) = a - b$. Donc, $u - v = 2b \in H^\perp$. Or, $\dim(H^\perp) = 1$ et $u - v \neq 0$. Donc, $H^\perp = \text{Vect}(u - v)$, puis $H = \text{Vect}(u - v)^\perp$.

Synthèse. Réciproquement, notons $H = \text{Vect}(u - v)^\perp$, puis σ la réflexion de E par rapport à H. Alors,

$$(u + v \mid u - v) = (u \mid u) - (u \mid v) + (v \mid u) - (v \mid x) = \|u\|^2 - \|v\|^2 = 0$$

et donc $u + v \in H$. Par suite,

$$u = \underbrace{\frac{u+v}{2}}_{\in H} + \underbrace{\frac{u-v}{2}}_{\in H^\perp}$$

ce qui prouve que,

$$\sigma(u) = \frac{u+v}{2} - \frac{u+v}{2} = v.$$

Ainsi, σ est l'unique réflexion de E qui envoie u sur v. □

Remarque. Cette réflexion est représentée en dimension 2 et 3 sur les figures 1 et 2.

Lemme 2. Soient $H_1, \dots, H_m$ des hyperplans de E. Alors,

$$\dim\left(\bigcap_{k=1}^{m} H_k\right) \geq n - m.$$

Preuve. On procède par récurrence sur $m \in \mathbb{N}^*$. Pour $m = 1$, il s'agit d'une égalité. Supposons ensuite le résultat établi pour un rang $m \in \mathbb{N}^*$ fixé. Alors, d'après la formule de Grassmann,

$$\dim\left(\bigcap_{k=1}^{m+1} H_k\right) = \dim\left(\bigcap_{k=1}^{m} H_k\right) + \dim(H_{m+1}) - \dim\left(\left(\bigcap_{k=1}^{m} H_k\right) + H_{m+1}\right).$$

Et donc, par hypothèse de récurrence,

$$\dim\left(\bigcap_{k=1}^{m+1} H_k\right) \geq (n - m) + (n - 1) - n = n - m - 1.$$

car, $\dim(H_{m+1}) = n - 1$,

$$\dim\left(\bigcap_{k=1}^{m} H_k\right) \geq n - m, \text{ et } \dim\left(\left(\bigcap_{k=1}^{m} H_k\right) + H_{m+1}\right) \leq \dim(E) = n.$$

D'où l'hérédité. Par récurrence, on en conclut que, pour tout $m \in \mathbb{N}^*$,

$$\dim\left(\bigcap_{k=1}^{m} H_k\right) \geq n - m.$$

□

Théorème 3. Si $\phi \in O(E)$ alors ϕ est produit de $\text{rg}(\phi - Id_E)$ réflexions mais pas moins. [FG]

Preuve. Montrons par récurrence sur $r \in [\![0, n]\!]$ que tout élément ϕ de $O(E)$ tel que $\text{rg}(\phi - Id_E) \leq r$ est produit d'au plus r réflexions.

Initialisation. Pour $r = 0$, $\phi = Id_E$ et donc ϕ est le produit de 0 réflexion par convention.

Hérédité. Supposons maintenant le résultat établi pour un rang $r \in [\![0, n-1]\!]$ fixé. Considérons $\phi \in O(E)$ tel que $\text{rg}(\phi - Id_E) \leq r + 1$. Si $\text{rg}(\phi - Id_E) \leq r$, le résultat est immédiat par hypothèse de récurrence. On peut donc supposer que $\text{rg}(\phi - Id_E) = r + 1$. En particulier, $\text{Im}(\phi - Id_E) \neq \{0_E\}$, c'est-à-dire qu'il existe $x_0 \in E \setminus \{0_E\}$ tel que $\phi(x_0) \neq x_0$. Notons $y_0 = \phi(x_0)$, qui vérifie $\|y_0\| = \|x_0\|$, car $\phi \in O(E)$. Afin d'appliquer notre hypothèse de récurrence, nous allons composer ϕ par une réflexion σ de E envoyant y_0 sur x_0, de telle sorte que $\text{rg}(\sigma \circ \phi - Id_E) < \text{rg}(\phi - Id_E)$. D'après le lemme 1, cette réflexion existe, puisque x_0 et y_0 sont de même norme. Il s'agit de la réflexion σ par rapport à l'hyperplan $H = \text{Vect}(y_0 - x_0)^\perp$. Montrons que,

$$\text{rg}(\sigma \circ \phi - Id_E) \leq r,$$

ce qui permettra d'appliquer l'hypothèse de récurrence. Pour cela, il suffit de montrer que $\text{Ker}(\phi - Id_E)$ est strictement inclus dans $\text{Ker}(\sigma \circ \phi - Id_E)$. Soit $x \in \text{Ker}(\phi - Id_E)$. Alors,

$$(x \mid y_0 - x_0) = (x \mid y_0) - (x \mid x_0) = (\phi(x) \mid \phi(x_0)) - (x \mid x_0) = 0,$$

puisque $\phi \in O(E)$. Donc, $x \in H$. Par suite, $\sigma(\phi(x)) = \sigma(x) = x$, c'est-à-dire $x \in \mathrm{Ker}(\sigma \circ \phi - Id_E)$. Ainsi,

$$\mathrm{Ker}(\phi - Id_E) \subset \mathrm{Ker}(\sigma \circ \phi - Id_E)$$

et cette inclusion est stricte car $x_0 \in \mathrm{Ker}(\sigma \circ \phi - Id_E) \setminus \mathrm{Ker}(\phi - Id_E)$. D'où,

$$\dim(\mathrm{Ker}(\phi - Id_E)) < \dim(\mathrm{Ker}(\sigma \circ \phi - Id_E))$$

et donc, $\mathrm{rg}(\sigma \circ \phi - Id_E) < \mathrm{rg}(\phi - Id_E) = r + 1$ avec le théorème du rang. D'où, $\mathrm{rg}(\sigma \circ \phi - Id_E) \leq r$. Par hypothèse de récurrence, il existe donc $m \in \mathbb{N}$ et $\rho_1, \dots, \rho_m$ des réflexions de E tels que $m \leq \mathrm{rg}(\sigma \circ \phi - Id_E)$ et,

$$\sigma \circ \phi = \rho_1 \circ \cdots \circ \rho_m.$$

Par suite, $m + 1 \leq \mathrm{rg}(\sigma \circ \phi - Id_E) + 1 \leq r + 1$ et,

$$\phi = \sigma \circ \rho_1 \circ \cdots \circ \rho_m$$

ce qui prouve que ϕ est produit d'au plus $r+1$ réflexions de E. Par récurrence, on en conclut que tout élément $\phi \in O(E)$ tel que $\mathrm{rg}(\phi - Id_E) \leq r$ est produit d'au plus r réflexions. Il reste maintenant à prouver qu'on ne peut pas faire mieux. Le résultat est clair pour Id_E. Soit $\phi \in O(E) \setminus \{Id_E\}$. Supposons qu'il existe $m \in \mathbb{N}^*$ et $H_1, \dots, H_m$ des hyperplans de E tels que,

$$\phi = \rho_1 \circ \cdots \circ \rho_m.$$

où l'on a noté, pour tout $k \in [\![1, m]\!]$, ρ_k la réflexion de E par rapport à H_k. Montrons que $m \geq r$. Tout d'abord,

$$\bigcap_{k=1}^{m} H_k \subset \mathrm{Ker}(\phi - Id_E).$$

En effet, si $x \in H_1 \cap \cdots \cap H_m$ alors, pour tout $k \in [\![1, m]\!]$, $\rho_k(x) = x$, puis,

$$\phi(x) = (\rho_1 \circ \cdots \circ \rho_m)(x) = (\rho_1 \circ \cdots \circ \rho_{m-1})(\rho_m(x)) = (\rho_1 \circ \cdots \circ \rho_{m-1})(x),$$

et donc $\phi(x) = x$ par récurrence finie. D'où,

$$\bigcap_{k=1}^{m} H_k \subset \mathrm{Ker}(\phi - Id_E).$$

En particulier,

$$\dim\left(\bigcap_{k=1}^{m} H_k\right) \leq \dim(\mathrm{Ker}(\phi - Id_E)).$$

Et donc, $\dim(\mathrm{Ker}(\phi - Id_E)) \geq n - m$ en utilisant le lemme 2. Or, d'après le théorème du rang,

$$\dim(\mathrm{Ker}(\phi - Id_E)) = n - \mathrm{rg}(\phi - Id_E).$$

D'où, $m \geq \mathrm{rg}(\phi - Id_E)$, ce qui achève la preuve. □

Figures. 1, 2.

Complément conseillé.

- Générateurs du groupe orthogonal d'une forme quadratique définie : théorème 2.

Commentaires personnels. Je pense qu'il vaut mieux présenter un seul des deux lemmes en fonction de la leçon présentée. Si vous présentez la leçon 148, choisissez plutôt le lemme 2. Sinon, choisissez plutôt le lemme 1. Vous pouvez aussi ne prouver que le théorème, si vous préférez avoir plus de temps pour le présenter. Par ailleurs, ce développement est assez facile et se recase dans de nombreuses leçons. Je le recommande. De plus, ce dernier possède l'immense avantage de mêler algèbre et géométrie, et vous offre l'opportunité de dessiner, ce qui sera apprécié par le jury.

Générateurs du groupe orthogonal d'une forme quadratique définie

Soient E un espace vectoriel de dimension finie $n \in \mathbb{N}^*$ sur un corps commutatif $\mathbb{K}$. Soit q une forme quadratique anisotrope (définie) sur E. Tout sous-espace vectoriel F de E est donc non isotrope (et même anisotrope), et on peut ainsi considérer la symétrie orthogonale par rapport à F, car $F \oplus F^\perp = E$. Notons φ la forme polaire de q.

Lemme 1. Soient $u, v \in E$ tels que $u \neq v$ et $q(u) = q(v)$. Alors, il existe une unique réflexion σ de E telle que $\sigma(u) = v$.

Preuve. Pour le démontrer, procédons par analyse-synthèse.

Analyse. Supposons qu'il existe un hyperplan H de E telle que la réflexion σ par rapport à H vérifie $\sigma(u) = v$. Soit $(a, b) \in H \times H^\perp$ tel que $u = a + b$. Alors, $v = \sigma(u) = a - b$. Donc, $u - v = 2b \in H^\perp$. Or, $\dim(H^\perp) = 1$ (car q est non dégénérée) et $u - v \neq 0$. Donc, $H^\perp = \text{Vect}(u - v)$, puis $H = \text{Vect}(u - v)^\perp$.

Synthèse. Réciproquement, notons $H = \text{Vect}(u - v)^\perp$, puis σ la réflexion de E par rapport à H. Alors,

$$\varphi(u + v, u - v) = \varphi(u, u) - \varphi(u, v) + \varphi(v, u) - \varphi(v, v) = q(u) - q(v) = 0$$

et donc $u + v \in H$. Par suite,

$$u = \underbrace{\frac{u+v}{2}}_{\in H} + \underbrace{\frac{u-v}{2}}_{\in H^\perp}$$

($H^\perp = \text{Vect}(u - v)$ car q est non dégénérée), ce qui prouve que,

$$\sigma(u) = \frac{u+v}{2} - \frac{u+v}{2} = v.$$

Ainsi, σ est l'unique réflexion de E qui envoie u sur v. □

Théorème 2. Si $\psi \in O(q)$ alors ψ est produit de $\text{rg}(\psi - Id_E)$ réflexions mais pas moins.

Preuve. Montrons par récurrence sur $r \in [\![0, n]\!]$ que tout élément ψ de $O(q)$ tel que $\text{rg}(\psi - Id_E) \leq r$ est produit d'au plus r réflexions.

Initialisation. Pour $r = 0$, $\psi = Id_E$ et donc ψ est le produit de 0 réflexion par convention.

Hérédité. Supposons maintenant le résultat établi pour un rang $r \in [\![0, n-1]\!]$ fixé. Considérons $\psi \in O(q)$ tel que $\text{rg}(\psi - Id_E) \leq r + 1$. Si $\text{rg}(\psi - Id_E) \leq r$, le résultat est immédiat par hypothèse de récurrence. On peut donc supposer que $\text{rg}(\psi - Id_E) = r + 1$. En particulier, $\text{Im}(\psi - Id_E) \neq \{0_E\}$, c'est-à-dire qu'il existe $x_0 \in E \setminus \{0_E\}$ tel que $\psi(x_0) \neq x_0$. Notons $y_0 = \psi(x_0)$, qui vérifie $q(x_0) = q(y_0)$, car $\psi \in O(q)$. Afin d'appliquer notre hypothèse de récurrence, nous allons composer ψ par une réflexion σ de E envoyant y_0 sur x_0, de telle sorte que $\text{rg}(\sigma \circ \psi - Id_E) < \text{rg}(\psi - Id_E)$. D'après le lemme 1, cette réflexion existe, car $q(x_0) = q(y_0)$. Il s'agit de la réflexion σ par rapport à l'hyperplan $H = \text{Vect}(y - x_0)^\perp$. Montrons que,

$$\text{rg}(\sigma \circ \psi - Id_E) \leq r,$$

ce qui permettra d'appliquer l'hypothèse de récurrence. Pour cela, il suffit de montrer que $\text{Ker}(\psi - Id_E)$ est strictement inclus dans $\text{Ker}(\sigma \circ \psi - Id_E)$. Soit $x \in \text{Ker}(\psi - Id_E)$. Alors,

$$\varphi(x, y_0 - x_0) = \varphi(x, y_0) - \varphi(x, x_0) = \varphi(\psi(x), \psi(x_0)) - \varphi(x, x_0) = 0,$$

puisque $\psi \in O(q)$. Donc, $x \in H$. Par suite, $\sigma(\psi(x)) = \sigma(x) = x$, c'est-à-dire $x \in \text{Ker}(\sigma \circ \psi - Id_E)$. Ainsi,

$$\text{Ker}(\psi - Id_E) \subset \text{Ker}(\sigma \circ \psi - Id_E)$$

et cette inclusion est stricte car $x_0 \in \text{Ker}(\sigma \circ \psi - Id_E) \setminus \text{Ker}(\psi - Id_E)$. D'où,

$$\dim(\text{Ker}(\psi - Id_E)) < \dim(\text{Ker}(\sigma \circ \psi - Id_E))$$

et donc, $\text{rg}(\sigma \circ \psi - Id_E) < \text{rg}(\psi - Id_E) = r + 1$ avec le théorème du rang. D'où, $\text{rg}(\sigma \circ \psi - Id_E) \leq r$. Par hypothèse de récurrence, il existe donc $m \in \mathbb{N}$ et $\rho_1, \ldots, \rho_m$ des réflexions de E tels que $m \leq \text{rg}(\sigma \circ \psi - Id_E)$ et,

$$\sigma \circ \psi = \rho_1 \circ \cdots \circ \rho_m.$$

Par suite, $m + 1 \leq \text{rg}(\sigma \circ \psi - Id_E) + 1 \leq r + 1$ et,

$$\psi = \sigma \circ \rho_1 \circ \cdots \circ \rho_m$$

ce qui prouve que ψ est produit d'au plus $r+1$ réflexions de E. Par récurrence, on en conclut que tout élément $\psi \in O(q)$ tel que $\text{rg}(\psi - Id_E) \leq r$ est produit d'au plus r réflexions. Il reste maintenant à prouver qu'on ne peut pas faire mieux. Le résultat est clair pour Id_E. Soit $\psi \in O(q) \setminus \{Id_E\}$. Supposons qu'il existe $m \in \mathbb{N}^*$ et $H_1, \ldots, H_m$ des hyperplans de E tels que,

$$\psi = \rho_1 \circ \cdots \circ \rho_m.$$

où l'on a noté, pour tout $k \in [\![1, m]\!]$, ρ_k la réflexion de E par rapport à H_k. Montrons que $m \geq r$. Tout d'abord,

$$\bigcap_{k=1}^{m} H_k \subset \mathrm{Ker}(\psi - Id_E).$$

En effet, si $x \in H_1 \cap \cdots \cap H_m$ alors, pour tout $k \in [\![1, m]\!]$, $\rho_k(x) = x$, puis,

$$\psi(x) = (\rho_1 \circ \cdots \circ \rho_m)(x) = (\rho_1 \circ \cdots \circ \rho_{m-1})(\rho_m(x)) = (\rho_1 \circ \cdots \circ \rho_{m-1})(x),$$

et donc $\psi(x) = x$ par récurrence finie. D'où,

$$\bigcap_{k=1}^{m} H_k \subset \mathrm{Ker}(\psi - Id_E).$$

En particulier,

$$\dim\left(\bigcap_{k=1}^{m} H_k\right) \leq \dim(\mathrm{Ker}(\psi - Id_E)).$$

Et donc, $\dim(\mathrm{Ker}(\psi - Id_E)) \geq n - m$ en utilisant le lemme 2. Or, d'après le théorème du rang,

$$\dim(\mathrm{Ker}(\psi - Id_E)) = n - \mathrm{rg}(\psi - Id_E).$$

D'où, $m \geq \mathrm{rg}(\psi - Id_E)$, ce qui achève la preuve. □

Remarque. Dans cette preuve, l'hypothèse que q soit définie est nécessaire car l'espace vectoriel $\mathrm{Vect}(y - x_0)^\perp$ peut être isotrope si on suppose seulement que q est non dégénérée. Toutefois, une autre preuve permet d'établir que si q est non dégénérée alors tout élément de $O(q)$ est produit d'au plus n réflexions.

Figures. 1, 2.

Complément conseillé.

- Générateurs du groupe orthogonal d'un espace vectoriel euclidien : théorème 3.

Commentaires personnels. Pour être tout à fait honnête, je ne suis pas sûr de la pertinence d'un tel développement. À quelques détails près, la preuve est identique à celle du théorème 3. Par conséquent, si vous avez choisi de présenter les générateurs du groupe orthogonal d'un espace vectoriel euclidien, ça ne vous demandera pas beaucoup d'efforts de présenter ce développement. Mais d'un autre côté, cette grande ressemblance avec la preuve du théorème 3 soulève un problème : cette preuve utilise très peu les formes quadratiques. On peut donc douter de sa pertinence dans une leçon sur les formes quadratiques. Dans l'idéal, il faudrait plutôt présenter sa version pour une forme quadratique non dégénérée, qui utilise la théorie des formes quadratiques de façon bien plus cruciale. Mais cette dernière est plus difficile.

Loi d'inertie de Sylvester

Soit $\mathbb{K}$ un corps commutatif de caractéristique différente de 2. Soit E un $\mathbb{K}$-espace vectoriel de dimension finie $n \in \mathbb{N}^*$. Soit q une forme quadratique sur E. On note φ sa forme polaire.

Définition 1. Une base $\mathcal{B} = (e_1, \dots, e_n)$ de E est dite q-orthogonale si,

$$\forall (i,j) \in [\![1,n]\!]^2, \quad i \neq j \Longrightarrow \varphi(e_i, e_j) = 0.$$

Autrement dit, si on note, pour tout $i \in [\![1,n]\!]$, $a_i = q(e_i)$, alors $(e_1, \dots, e_n)$ est une base q-orthogonale de E si, pour tout $x = \sum_{i=1}^{n} x_i e_i \in E$,

$$q(x) = \sum_{i=1}^{n} a_i x_i^2$$

ou d'une manière équivalente :

$$\mathrm{Mat}_{\mathcal{B}}(q) = \begin{pmatrix} a_1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & a_n \end{pmatrix}.$$

[JG] **Théorème 2.** Il existe une base q-orthogonale de E.

Preuve. Dans cette preuve, E et q ne sont plus fixés. Montrons, par récurrence sur $n \in \mathbb{N}^*$ que, pour tout $\mathbb{K}$-espace vectoriel E de dimension n, et pour toute forme quadratique q sur E, il existe une base q-orthogonale de E.

Initialisation. Soient E une droite vectorielle sur $\mathbb{K}$ et q une forme quadratique sur E. Alors, toute base de E est une base q-orthogonale, d'où le résultat.

Hérédité. Supposons ensuite le résultat établi pour un rang $n \in \mathbb{N}^*$ fixé. Soient E un $\mathbb{K}$-espace vectoriel de dimension $n+1$ et q une forme quadratique sur E.

$*$ Si $q = 0$, toutes les bases de E sont q-orthogonales, d'où le résultat.

$*$ Supposons ensuite $q \neq 0$. Soit $x \in E$ tel que $q(x) \neq 0$. Notons,

$$F = \{y \in E \mid \varphi(x,y) = 0\}.$$

Autrement dit, F est le noyau de la forme linéaire ω, définie pour tout $y \in E$ par $\omega(y) = \varphi(x,y)$. Il s'agit donc d'un sous-espace vectoriel de E (en tant que noyau d'une forme linéaire). De plus, d'après le théorème du rang,

$$\dim(F) = \dim(\mathrm{Ker}(\omega)) = \dim(E) - \dim(\mathrm{Im}(\omega)).$$

Or, $\mathrm{Im}(\omega) = \mathbb{K}$ en tant que sous-espace vectoriel de $\mathbb{K}$ non réduit à $\{0_E\}$ (car $q(x) \in \mathrm{Im}(\omega)$). D'où, $\dim(F) = n$. Par hypothèse de récurrence, il existe donc $(v_1, \dots, v_n)$ une base q_F-orthogonale de F (où l'on a noté q_F la restriction de q à F). D'autre part, $x \notin F$ (car $\omega(x) = q(x) \neq 0$). Puisque F est un hyperplan de E, ceci assure que $E = \mathrm{Vect}(x) \oplus F$. Par conséquent, $(x, v_1, \dots, v_n)$ est une base de E. Il reste à montrer que cette base est q-orthogonale de E. Notons :

- $\mathcal{B} = (x, v_1, \dots, v_n)$;
- $\mathcal{B}' = (v_1, \dots, v_n)$.

Avec ces notations, il suffit de prouver que $\mathrm{Mat}_{\mathcal{B}}(q)$ est une matrice diagonale pour conclure que $\mathcal{B}$ est une base q-orthogonale de E. Tout d'abord,

$$\mathrm{Mat}_{\mathcal{B}}(q) = \begin{pmatrix} \varphi(x,x) & \varphi(x,v_1) & \cdots & \varphi(x,v_n) \\ \varphi(v_1,x) & \varphi(v_1,v_1) & \cdots & \varphi(v_1,v_n) \\ \vdots & \vdots & & \vdots \\ \varphi(v_n,x) & \varphi(v_n,v_1) & \cdots & \varphi(v_n,v_n) \end{pmatrix}.$$

Or, pour tout $i \in [\![1,n]\!]$, $\varphi(v_i,x) = \varphi(x,v_i) = 0$ (car $v_i \in F$), et :

$$\begin{pmatrix} \varphi(v_1,v_1) & \cdots & \varphi(v_1,v_n) \\ \vdots & & \vdots \\ \varphi(v_n,v_1) & \cdots & \varphi(v_n,v_n) \end{pmatrix} = \mathrm{Mat}_{\mathcal{B}'}(q_F) = \begin{pmatrix} 0 & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & 0 \end{pmatrix}$$

car $(v_1, \dots, v_n)$ est une base q_F-orthogonale de F. Donc,

$$\mathrm{Mat}_{\mathcal{B}}(q) = \begin{pmatrix} q(x) & 0 & \cdots & 0 \\ 0 & q(v_1) & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & q(v_n) \end{pmatrix}$$

ce qui prouve que $\mathcal{B}$ est une base q-orthogonale de E. Ainsi, E possède une base q-orthogonale.

Finalement, par récurrence, on en conclut que, pour tout $\mathbb{K}$-espace vectoriel E de dimension finie et pour toute forme quadratique q sur E, il existe des bases q-orthogonales de E. □

Remarque. Pour construire une telle base, on peut appliquer l'algorithme de réduction de Gauss.

Notons r le rang de la forme quadratique q.

[JG] **Théorème 3.** Supposons $\mathbb{K} = \mathbb{R}$. Alors, il existe une base $\mathcal{B} = (e_1, \dots, e_n)$ de E et $p \in [\![0, n]\!]$ tels que,

$$\forall x = \sum_{k=1}^{n} x_k e_k \in E, \quad q(x) = \sum_{k=1}^{p} x_k^2 - \sum_{k=p+1}^{r} x_k^2$$

c'est-à-dire que :

$$\mathrm{Mat}_{\mathcal{B}}(q) = \left(\begin{array}{c|c|c} I_p & \mathbf{0} & \mathbf{0} \\ \hline \mathbf{0} & -I_{r-p} & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{0} & \mathbf{0} \end{array}\right).$$

De plus, cet entier p ne dépend que de la forme quadratique (il ne dépend pas de la base $\mathcal{B}$). Le couple $(p, r-p)$ est appelé la signature de q.

Preuve. Soit $\mathcal{B}_0 = (f_1, \dots, f_n)$ une base q-orthogonale de E. Alors,

$$\mathrm{Mat}_{\mathcal{B}_0}(q) = \begin{pmatrix} a_1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & a_n \end{pmatrix}.$$

où l'on a noté, pour tout $k \in [\![1, n]\!]$, $a_k = q(f_k)$. En particulier,

$$r = \mathrm{rg}(q) = \mathrm{rg}(\mathrm{Mat}_{\mathcal{B}_0}(q)) = \mathrm{card}(\{k \in [\![1, n]\!] \mid a_k \neq 0\}).$$

Quitte à permuter les vecteurs de la base, on peut donc supposer que :

$$a_1, \dots, a_r \in \mathbb{R}^*, \ a_{r+1} = \cdots = a_n = 0.$$

Notons ensuite,

$$p = \mathrm{card}(\{k \in [\![1, r]\!] \mid a_k > 0\}).$$

De nouveau, quitte à permuter les vecteurs de la base, on peut donc supposer que :

$$a_1, \dots, a_p \in \mathbb{R}_+^*, \ a_{p+1}, \dots, a_r \in \mathbb{R}_-^*.$$

Pour tout $k \in [\![1, n]\!]$, notons :

$$e_k = \begin{cases} \dfrac{1}{\sqrt{a_k}} f_k & \text{si } 1 \leq k \leq p, \\ \dfrac{1}{\sqrt{-a_k}} f_k & \text{si } p+1 \leq k \leq r, \\ f_k & \text{sinon.} \end{cases}$$

puis $\mathcal{B} = (e_1, \dots, e_n)$. Avec ces notations, il vient, pour tout $k \in [\![1, n]\!]$,

- si $k \in [\![1, p]\!]$, $q(e_k) = \left(\dfrac{1}{\sqrt{a_k}}\right)^2 q(f_k) = \dfrac{1}{a_k} a_k = 1$;
- si $k \in [\![p+1, r]\!]$, $q(e_k) = \left(\dfrac{1}{\sqrt{-a_k}}\right)^2 q(f_k) = \dfrac{1}{-a_k} a_k = -1$;
- si $k \in [\![r, n]\!]$, $q(e_k) = q(f_k) = 0$.

D'où,

$$\mathrm{Mat}_{\mathcal{B}}(q) = \left(\begin{array}{c|c|c} I_p & \mathbf{0} & \mathbf{0} \\ \hline \mathbf{0} & -I_{r-p} & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{0} & \mathbf{0} \end{array}\right).$$

Il reste à montrer que cet entier p ne dépend pas de la base $\mathcal{B}$.

Soient $\mathcal{B}' = (e'_1, \dots, e'_n)$ une base de E et $p' \in [\![1, n]\!]$ tels que :

$$\mathrm{Mat}_{\mathcal{B}'}(q) = \left(\begin{array}{c|c|c} I_{p'} & \mathbf{0} & \mathbf{0} \\ \hline \mathbf{0} & -I_{r-p'} & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{0} & \mathbf{0} \end{array}\right).$$

Notons :

$$\begin{cases} F = \mathrm{Vect}(e_1, \dots, e_p), & G = \mathrm{Vect}(e_{p+1}, \dots, e_n), \\ F' = \mathrm{Vect}(e'_1, \dots, e'_{p'}), & G' = \mathrm{Vect}(e'_{p'+1}, \dots, e'_n) \end{cases}$$

de telle sorte que, pour tout $x \in E \setminus \{0_E\}$,

$$x \in F \cup F' \Longrightarrow q(x) > 0, \quad x \in G \cup G' \Longrightarrow q(x) \leq 0$$

car, pour tout $x = (x_1, \dots, x_n)_{\mathcal{B}} = (x'_1, \dots, x'_n)_{\mathcal{B}'} \in E$,

$$q(x) = \sum_{k=1}^{p} x_k^2 - \sum_{k=p+1}^{r} x_k^2 = \sum_{k=1}^{p'} (x'_k)^2 - \sum_{k=p+1}^{r} (x'_k)^2.$$

Par conséquent, $F \cap G' = \{0_E\}$ et $F' \cap G = \{0_E\}$. Autrement dit,

- F et G' sont en somme directe,
- F' et G sont en somme directe,

et donc,

- $\dim(F + G') = \dim(F) + \dim(G') = p + n - p'$,
- $\dim(F' + G) = \dim(F') + \dim(G) = p' + n - p$.

Or, $\dim(F+G') \leq n$ et $\dim(F'+G) \leq n$ car $F+G'$ et $F'+G$ sont des sous-espaces vectoriels de E. D'où, $p+n-p' \leq n$ et $p'+n-p \leq n$, c'est-à-dire $p \leq p'$ et $p' \leq p$, ce qui prouve finalement que $p = p'$. □

[JG] **Théorème 4.** Supposons $\mathbb{K} = \mathbb{C}$. Alors, il existe une base $\mathcal{B} = (e_1, \dots, e_n)$ de E et $p \in [\![0, n]\!]$ tels que,

$$\forall x = \sum_{k=1}^{n} x_k e_k \in E, \quad q(x) = \sum_{k=1}^{r} x_k^2$$

c'est-à-dire que :

$$\mathrm{Mat}_{\mathcal{B}}(q) = \left(\begin{array}{c|c} I_r & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{0} \end{array}\right).$$

Preuve. Soit $\mathcal{B}_0 = (f_1, \dots, f_n)$ une base q-orthogonale de E. Alors,

$$\mathrm{Mat}_{\mathcal{B}_0}(q) = \begin{pmatrix} a_1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & a_n \end{pmatrix}.$$

où l'on a noté, pour tout $k \in [\![1, n]\!]$, $a_k = q(f_k)$. En particulier,

$$r = \mathrm{rg}(q) = \mathrm{rg}(\mathrm{Mat}_{\mathcal{B}_0}(q)) = \mathrm{card}(\{k \in [\![1, n]\!] \mid a_k \neq 0\}).$$

Quitte à permuter les vecteurs de la base, on peut donc supposer que :

$$a_1, \dots, a_r \in \mathbb{C}^*, \ a_{r+1} = \cdots = a_n = 0.$$

Pour tout $k \in [\![1, r]\!]$, il existe $\lambda_k \in \mathbb{C}^*$ tel que $a_k = \lambda_k^2$ (car $a_k \in \mathbb{C}^*$). On note alors :

$$e_k = \begin{cases} \dfrac{1}{\lambda_k} f_k & \text{si } 1 \leq k \leq r, \\ f_k & \text{sinon.} \end{cases}$$

puis $\mathcal{B} = (e_1, \dots, e_n)$. Avec ces notations, il vient, pour tout $k \in [\![1, n]\!]$,

- si $k \in [\![1, r]\!]$, $q(e_k) = \left(\dfrac{1}{\lambda_k}\right)^2 q(f_k) = \dfrac{1}{a_k} a_k = 1$;
- si $k \in [\![r, n]\!]$, $q(e_k) = q(f_k) = 0$.

D'où,

$$\mathrm{Mat}_{\mathcal{B}}(q) = \left(\begin{array}{c|c} I_r & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{0} \end{array}\right).$$

□

Complément conseillé.

- Formes quadratiques : sous-section 2.

Commentaires personnels. Développement non rédigé durant ma préparation au concours. Je m'abstiendrai donc de vous conseiller sur ce dernier.

Probabilité que deux éléments commutent dans un groupe

[FG] **Théorème 1 (Théorème de Dixon).** Soit G un groupe fini non abélien. On note $n(G)$ le nombre de couples d'éléments de G qui commutent divisé par le nombre total de couples de G. Alors, $n(G) \leq \frac{5}{8}$.

Preuve. Pour tout $a \in G$, on note C_a le centralisateur de a. Alors, l'ensemble des couples d'éléments de G qui commutent est l'union disjointe de $\{a\} \times C_a \subset G^2$ pour $a \in G$. En particulier, son cardinal est égal à la somme des cardinaux de C_a pour $a \in G$. D'où,

$$|G|^2 n(G) = \sum_{a \in G} |C_a|.$$

Par ailleurs, pour tout $a \in G$, C_a est le stabilisateur de a pour l'action par conjugaison de G sur lui-même, c'est-à-dire l'action,

$$\begin{array}{rccc} \cdot & : & G \times G & \longrightarrow & G \\ & & (g,a) & \longmapsto & gag^{-1} \end{array}$$

Considérons alors $\mathcal{O}_1, \ldots, \mathcal{O}_k$ les classes de conjugaison deux à deux distinctes de G, c'est-à-dire ses orbites pour l'action par conjugaison. Puisque celles-ci forment une partition de G, il vient,

$$|G|^2 n(G) = \sum_{i=1}^{k} \sum_{a \in \mathcal{O}_i} |C_a| = \sum_{i=1}^{k} \sum_{a \in \mathcal{O}_i} \frac{|G|}{|\mathcal{O}_i|} = |G| \sum_{i=1}^{k} \frac{1}{|\mathcal{O}_i|} \sum_{a \in \mathcal{O}_i} 1.$$

Puis,

$$|G|^2 n(G) = |G| \sum_{i=1}^{k} \frac{1}{|\mathcal{O}_i|} |\mathcal{O}_i| = |G| k.$$

Et donc $n(G) = \dfrac{k}{|G|}$. Remarquons ensuite que l'orbite d'un élément a de G pour l'action par conjugaison est réduite au point $\{a\}$ si et seulement si $a \in Z(G)$. Sachant que les orbites forment une partition de G, il vient,

$$|G| = \sum_{i=1}^{k} |\mathcal{O}_i|.$$

Or, pour tout $a \in G$, $|C_a(G)| = 1$ si et seulement si $a \in Z(G)$. Il y a donc exactement $|Z(G)|$ classes de conjugaison de cardinal 1, et les $k - |Z(G)|$ classes de conjugaison restantes sont de cardinal supérieur ou égal à 2. D'où, en isolant les classes de conjugaison de cardinal 1 dans la somme précédente,

$$|G| \geq |Z(G)| + 2(k - |Z(G)|) = 2k - |Z(G)|$$

En divisant l'inégalité précédente par $|G|$, on obtient,

$$n(G) = \frac{k}{|G|} \leq \frac{1}{2} + \frac{|Z(G)|}{2|G|}.$$

Pour montrer que $n(G) \leq \frac{5}{8}$, il suffit donc de montrer que $\frac{1}{2} + \frac{|Z(G)|}{2|G|} \leq \frac{5}{8}$, c'est-à-dire que l'indice de $Z(G)$ dans G est supérieur ou égal à 4. Pour cela, on montre que les autres cas sont impossibles.

- Si $[G : Z(G)] = 1$ alors $Z(G) = G$ et donc G est abélien, ce qui contredit notre hypothèse.
- Si $[G : Z(G)] = 2$ ou $[G : Z(G)] = 3$ alors $G/Z(G)$ est cyclique, et donc G est abélien par le lemme 3, ce qui contredit notre hypothèse.

Ainsi, $[G : Z(G)] \geq 4$, et donc $n(G) \leq \frac{5}{8}$. □

Remarque. Le nombre $n(G)$ correspond à la probabilité que deux éléments de G tirés uniformément et indépendamment commutent.

Proposition 2. Notons D_8 le groupe diédral d'ordre 8. Alors, $n(D_8) = \frac{5}{8}$. [LM]

Preuve. Le groupe diédral D_8 est composé (voir figure 3) de quatre rotations, d'angles respectifs 0, $\frac{\pi}{2}$, π et $\frac{3\pi}{2}$ et de quatre réflexions. Déterminons alors les couples d'éléments de D_8 qui commutent. On distingue trois cas.

- **Cas n°1.** Produit de deux rotations. Si r_1 et r_2 sont deux rotations de D_8 d'angles respectifs θ_1 et θ_2 alors $r_1 \circ r_2$ est une rotation d'angle $\theta_1 + \theta_2$. En particulier, les rotations de D_8 commutent toutes entre elles.
- **Cas n°2.** Produit de deux réflexions. Soient s_1 et s_2 deux réflexions de D_8, d'axes respectifs Δ_1 et Δ_2. Si $s_1 = s_2$, alors s_1 et s_2 commutent. Sinon, les axes Δ_1 et Δ_2 ne sont pas parallèles (il n'y a que quatre axes possibles dans D_8), et admettent donc un unique point d'intersection M. Avec ces notations, $s_1 \circ s_2$ est une rotation affine de centre M, et d'angle

$2(\Delta_2, \Delta_1)$, où (Δ_2, Δ_1) désigne l'angle orienté entre les droites Δ_2 et Δ_1. En particulier,

$$\begin{aligned} s_1 \circ s_2 = s_2 \circ s_1 &\iff 2(\Delta_1, \Delta_2) \equiv 2(\Delta_2, \Delta_1)\ [2\pi] \\ &\iff 4(\Delta_1, \Delta_2) \equiv 0\ [2\pi] \\ s_1 \circ s_2 = s_2 \circ s_1 &\iff (\Delta_1, \Delta_2) \equiv 0\ \left[\frac{\pi}{2}\right]. \end{aligned}$$

Ainsi, deux réflexions de D_8 commutent si et seulement si leurs axes sont perpendiculaires. Par suite, toute réflexion de D_8 d'axe Δ commute avec elle-même et avec la réflexion d'axe $\Delta^\perp$, mais ne commute pas avec les autres réflexions.

- **Cas n°3.** Produit d'une rotation et d'une réflexion. Soient r une rotation et s une réflexion d'axe Δ de D_8. Alors, $r \circ s \circ r^{-1}$ est une réflexion d'axe $r(\Delta)$. En effet, $\det(r \circ s \circ r^{-1}) = -1$ et,

$$(r \circ s \circ r^{-1})(r(\Delta)) = r(s(\Delta)) = r(\Delta)$$

et donc $r \circ s \circ r^{-1}$ est une réflexion d'axe $r(\Delta)$. Or, r et s commutent si et seulement si $r \circ s \circ r^{-1} = s$. D'après ce qui précède, on en déduit que r et s commutent si et seulement si $r(\Delta) = \Delta$, donc si et seulement si $r = -Id$ ou $r = Id$. Par conséquent, une réflexion de D_8 commute avec $-Id$ et Id, mais ne commute pas avec les autres rotations.

Dénombrons finalement le nombre de couples de D_8 qui commutent, en distinguant suivant leur premier élément.

- Les éléments $-Id$ et Id commutent avec tout autre élément de D_8. Il y a donc $2 \times 8 = 16$ couples de D_8 qui commutent et dont le premier élément est $-Id$ ou Id.
- Les rotations d'angle $\frac{\pi}{2}$ et $\frac{3\pi}{2}$ commutent avec toute rotation, mais ne commutent avec aucune réflexion. Il y a donc $2 \times 4 = 8$ couples de D_8 correspondant.
- Chaque réflexion commute uniquement avec deux rotations ($-Id$ et Id) et deux réflexions (elle-même et la réflexion d'axe perpendiculaire au sien). Il y a donc $4 \times 4 = 16$ couples de D_8 correspondant.

Finalement, il y a donc $16+8+16 = 40$ couples de D_8 qui commutent. Sachant que D_8 possède huit éléments, on en conclut que $n(D_8) = \dfrac{40}{8^2} = \dfrac{5}{8}$. □

Figures. 3, 4, 5, 6.

Compléments conseillés.

- Centre d'un groupe : lemme 3.
- Probabilité que deux permutations commutent : proposition 1.

Commentaires personnels. Je ne vais pas passer par quatre chemins : c'est mon développement préféré. Et par chance, je suis passé sur ce développement lors de mon oral d'algèbre et géométrie. La preuve du théorème de Dixon est technique, mais le calcul de $n(D_8)$ est assez intuitif. De plus, ce développement possède l'immense avantage de mêler algèbre et géométrie, et vous offre l'opportunité de dessiner, ce qui sera apprécié par le jury.

Simplicité de $\mathfrak{A}_n$

[JR] **Théorème 1.** Si $n \geq 3$, le groupe $\mathfrak{A}_n$ est engendré par les 3-cycles.

Preuve. Montrons tout d'abord que le produit de deux transpositions est un produit de 3-cycles. Soient τ_1 et τ_2 des transpositions. On distingue trois cas.

- **Cas n°1.** Si $|\text{Supp}(\tau_1) \cap \text{Supp}(\tau_2)| = 2$, alors $\tau_1 = \tau_2$ puis $\tau_1\tau_2 = Id = \gamma^3$ pour tout 3-cycle γ.
- **Cas n°2.** Si $|\text{Supp}(\tau_1) \cap \text{Supp}(\tau_2)| = 1$, alors il existe $x, y, z \in [\![1, n]\!]$ deux à deux distincts tels que $\tau_1 = (x, y)$ et $\tau_2 = (y, z)$. Par suite,

$$\tau_1\tau_2 = (x, y)(y, z) = (x, y, z).$$

- **Cas n°2.** Si $\text{Supp}(\tau_1) \cap \text{Supp}(\tau_2) = \varnothing$, alors il existe $a, b, c, d \in [\![1, n]\!]$ deux à deux distincts tels que $\tau_1 = (a, b)$ et $\tau_2 = (c, d)$. Par suite,

$$\tau_1\tau_2 = (a, b)(c, d) = (a, b)(b, c)(b, c)(c, d) = (a, b, c)(b, c, d).$$

Montrons désormais que les 3-cycles engendrent $\mathfrak{A}_n$. Tout d'abord, les 3-cycles appartiennent bien à $\mathfrak{A}_n$, car pour tout entier $r \geq 2$, la signature d'un r-cycle est $(-1)^{r+1}$. Soit $\sigma \in \mathfrak{A}_n$. On sait qu'il existe $\tau_1, \ldots, \tau_d$ des transpositions de $\mathfrak{S}_n$ telles que,

$$\sigma = \tau_1 \circ \cdots \circ \tau_d.$$

En particulier, $\varepsilon(\sigma) = (-1)^d$. Or, $\varepsilon(\sigma) = 1$. Donc d est pair. On peut alors regrouper les facteurs du produit précédent par deux :

$$\sigma = \prod_{i=1}^{\frac{d}{2}} \tau_i\tau_{i+1}.$$

Sachant que les permutations $\tau_i\tau_{i+1}$ sont des produits de 3-cycles (d'après ce qui précède), on en déduit que σ est un produit de 3-cycles. □

Remarque. Si $n = 3$, le cas n°2 est impossible.

[JR] **Lemme 2.** Si $n \geq 5$, les 3-cycles sont conjugués dans $\mathfrak{A}_n$.

Preuve. Soient $\alpha = (x_1, x_2, x_3)$ et $\beta = (y_1, y_2, y_3)$ deux 3-cycles. Considérons σ la permutation de $[\![1, n]\!]$ qui envoie x_1 sur y_1, x_2 sur y_2, x_3 sur y_3 et laisse fixe tous les autres éléments. Alors,

$$\sigma\alpha\sigma^{-1} = \sigma(x_1, x_2, x_3)\sigma^{-1} = (\sigma(x_1), \sigma(x_2), \sigma(x_3)) = (y_1, y_2, y_3) = \beta.$$

Si $\varepsilon(\sigma) = 1$, alors $\sigma \in \mathfrak{A}_n$, et donc l'égalité ci-dessus fournit que α et β sont conjugués dans $\mathfrak{A}_n$. Sinon, on fixe $x_4, x_5 \in [\![1, n]\!] \setminus \{x_1, x_2, x_3\}$ (qui existent car $n \geq 5$), et on note $\rho = \tau \circ \sigma$. Alors,

$$\rho\alpha\rho^{-1} = (\rho(x_1), \rho(x_2), \rho(x_3)) = (y_1, y_2, y_3) = \beta.$$

avec $\varepsilon(\rho) = \varepsilon(\tau)\varepsilon(\sigma) = (-1)^2 = 1$, et donc l'égalité ci-dessus fournit que α et β sont conjugués dans $\mathfrak{A}_n$. □

Remarque. On a utilisé le résultat suivant : pour tout cycle $(x_1, \ldots, x_r) \in \mathfrak{S}_n$, et pour tout $\sigma \in \mathfrak{S}_n$,

$$\sigma \circ (x_1, \ldots, x_r) \circ \sigma^{-1} = (\sigma(x_1), \ldots, \sigma(x_r)).$$

Théorème 3. Si $n = 3$ ou $n \geq 5$, le groupe $\mathfrak{A}_n$ est simple. [JR]

Preuve. Pour $n = 3$, $\mathfrak{A}_3$ est de cardinal 3, et le théorème de Lagrange assure donc que ses seuls sous-groupes sont $\{Id\}$ et $\mathfrak{A}_3$. A fortiori, $\mathfrak{A}_3$ est simple.

Supposons maintenant $n \neq 5$. Soit H un sous-groupe distingué de $\mathfrak{A}_n$ distinct de $\{Id\}$. Montrons que $H = \mathfrak{A}_n$. Si H contient un 3-cycle, alors il contient ses conjugués, et donc il contient tous les 3-cycles par le lemme précédent. Or, les 3-cycles engendrent $\mathfrak{A}_n$ d'après le premier lemme. Par suite, si H contient un 3-cycle, il contient $\mathfrak{A}_n$ tout entier, c'est-à-dire que $H = \mathfrak{A}_n$. Pour conclure, il suffit donc de construire un 3-cycle appartenant à H. Soit $\sigma \in H \setminus \{Id\}$. Soit $x \in [\![1, n]\!]$ tel que $\sigma(x) \neq x$. Notons $y = \sigma(x)$. Soit $z \in [\![1, n]\!] \setminus \{x, y, \sigma(y)\}$ (qui existe car $n \geq 4$). Notons $\gamma = (x, y, z)$. Notons $\rho = \sigma\gamma\sigma^{-1}\gamma^{-1}$, qui appartient à H car H est un sous-groupe distingué de $\mathfrak{A}_n$. De plus,

$$\rho = \sigma(x, y, z)\sigma^{-1}\gamma^{-1} = (\sigma(x), \sigma(y), \sigma(z))(z, y, x) = (y, \sigma(y), \sigma(z))(z, y, x).$$

En particulier, le support de σ est contenu dans l'ensemble $\{x, y, z, \sigma(y), \sigma(z)\}$, qui est de cardinal au plus 5. D'autre part, $\rho \neq Id$, car $\rho(z) = \sigma(y) \neq z$. D'où l'existence de $c_1, \ldots, c_r$ des cycles à supports deux à deux disjoints, telles que,

$$\rho = c_1 \circ \cdots \circ c_r,$$

avec,

$$\bigcup_{k=1}^{r} \text{Supp}(c_k) = \text{Supp}(\rho).$$

Notons ensuite $l_1, \dots, l_r$ les longueurs respectives des cycles $c_1, \dots, c_r$. Quitte à les renommer, on peut supposer sans perte de généralité que $l_1 \leq \cdots \leq l_r$. D'après l'égalité précédente,

$$l_1 + \cdots + l_r = \left| \bigcup_{k=1}^{r} \mathrm{Supp}(c_k) \right| = |\mathrm{Supp}(\rho)| \leq 5.$$

Sachant que $l_1, \dots, l_r \geq 2$, ceci implique $r \leq 2$. On distingue

- **Cas n°1.** Si $r = 1$, alors $\varepsilon(\rho) = (-1)^{l_1+1} = 1$, car $\rho \in H \subset \mathfrak{A}_n$, ce qui implique que l_1 est impair. D'où, $l_1 = 3$ ou $l_1 = 5$, c'est-à-dire que ρ est soit un 3-cycle, soit un 5-cycle.
- **Cas n°2.** Si $r = 2$, alors $l_1 = 2$, et donc $\varepsilon(\rho) = (-1)^{3+l_2} = 1$, ce qui implique que l_2 est pair. D'où, $l_1 = l_2 = 2$, c'est-à-dire que ρ est un produit de deux transpositions.

Ainsi, σ est soit un 3-cycle, soit un 5-cycle, soit un produit de deux transpositions. On distingue de nouveau plusieurs cas.

- **Cas n°1.** Si ρ est un 3-cycle, H contient un 3-cycle.
- **Cas n°2.** Si $\rho = (x_1, x_2, x_3, x_4, x_5)$ est un 5-cycle, alors en considérant $\widetilde{\gamma} = (x_1, x_2, x_3)$, on obtient que $\widetilde{\rho} = \rho\widetilde{\gamma}\rho^{-1}\widetilde{\gamma}^{-1} \in H$ est un 3-cycle. Plus précisément,

  $$\widetilde{\rho} = (\rho(x_1), \rho(x_2), \rho(x_3))(x_3, x_2, x_1) = (x_2, x_3, x_4)(x_3, x_2, x_1)$$

  et donc $\widehat{\rho} = (x_1, x_4, x_2)$.
- **Cas n°3.** Si $\rho = (x_1, x_2)(x_3, x_4)$ est un produit de deux transpositions, alors en fixant $x_5 \in \{x_1, x_2, x_3, x_4\}$ (qui existe car $n \geq 5$), puis en notant $\widehat{\gamma} = (x_1, x_2, x_5)$, on obtient que $\widehat{\rho} = \rho\widehat{\gamma}\rho^{-1}\widehat{\gamma}^{-1} \in H$ est un 3-cycle. Plus précisément,

  $$\widehat{\rho} = (\rho(x_1), \rho(x_2), \rho(x_5))(x_5, x_2, x_1) = (x_2, x_1, x_5)(x_5, x_2, x_1)$$

  et donc $\widehat{\rho} = (x_1, x_2, x_5)$.

Dans tous les cas, H contient un 3-cycle, ce qui achève la preuve. □

Remarque. Le cycle γ construit ci-dessus n'est pas le même que celui de [JR].

Si vous êtes rapide, vous pouvez également démontrer le résultat suivant (dans le cas contraire, le jury sera susceptible de vous interroger sur le cas $n = 4$).

Proposition 4. Le groupe $\mathcal{A}_4$ n'est pas simple.

Preuve. Le groupe $\mathcal{A}_4$ possède 12 éléments :

- l'identité : Id;
- 8 éléments d'ordre 3, qui sont les 3-cycles :

  $$(1,2,3), (1,3,2), (1,2,4), (1,4,2), (1,3,4), (1,4,3), (2,3,4), (2,4,3);$$

- 3 éléments d'ordre 2, qui sont les doubles-transpositions (produit de deux transpositions à supports disjoints) :

  $$(1,2)(3,4), (1,3)(2,4), (1,4)(2,3)$$

  (ces éléments sont bien d'ordre 2, car si τ_1 et τ_2 sont deux transpositions à supports disjoints alors $\mathrm{ordre}(\tau_1\tau_2) = \mathrm{ppcm}(\mathrm{ordre}(\tau_1), \mathrm{ordre}(\tau_2)) = 2$).

Notons,

$$V_4 = \{Id, (1,2)(3,4), (1,3)(2,4), (1,4)(2,3)\}$$

(V est l'initiale de *Vierergruppe*, qui signifie groupe de quatre en allemand). On peut montrer que c'est un sous-groupe distingué de $\mathcal{A}_4$, ce qui prouve que le groupe $\mathfrak{A}_4$ n'est pas simple. □

Remarque. Le groupe $\mathfrak{A}_4$ possède 4 classes de conjugaison :

- la classe de conjugaison de l'identité : $\{Id\}$;
- la classe de conjugaison des doubles transpositions :

  $$\{(1,2)(3,4), (1,3)(2,4), (1,4)(2,3)\};$$

- une première classe constituée de 3-cycles :

  $$\{(1,2,3), (1,4,2), (1,3,4), (2,4,3)\};$$

- une deuxième classe constituée de 3-cycles :

  $$\{(1,3,2), (1,2,4), (1,4,3), (2,3,4)\}.$$

Commentaires personnels. C'est un développement très (trop ?) classique. Mais il n'est pas nécessaire d'être original pour être admis au concours, donc ce n'est pas un inconvénient majeur. C'est un développement technique, que j'ai dû beaucoup travailler pour bien le maîtriser. Malgré cela, je pense que c'était un investissement rentable puisqu'il se recase dans de nombreuses leçons.

Surjectivité de l'exponentielle matricielle

[IP] **Théorème 1.** Pour tout $A \in GL_n(\mathbb{C})$, il existe $B \in \mathbb{C}[A]$ tel que $A = \exp(B)$. En particulier, l'application,

$$\begin{array}{rccc} \exp & : & M_n(\mathbb{C}) & \longrightarrow & GL_n(\mathbb{C}) \\ & & A & \longmapsto & \exp(A) \end{array}$$

est surjective.

Preuve. Soit $A \in M_n(\mathbb{C})$. On procède en plusieurs étapes.

Étape n°1. Montrons que, pour tout $M \in M_n(\mathbb{C})$, $\exp(M) \in \mathbb{C}[M]$. Soit $M \in M_n(\mathbb{C})$. Par définition de l'exponentielle matricielle, on a,

$$\exp(M) = \sum_{n=0}^{+\infty} \frac{M^n}{n!} = \lim_{n \to +\infty} \underbrace{\sum_{k=0}^{n} \frac{M^k}{k!}}_{\in \mathbb{C}[M]}.$$

Or, $\mathbb{C}[M]$ est une partie fermée de $M_n(\mathbb{C})$ en tant que sous-espace vectoriel de dimension finie de $M_n(\mathbb{C})$ (car $M_n(\mathbb{C})$ est lui-même de dimension finie, égale à n^2). Donc $\exp(M) \in \mathbb{C}[M]$ en tant que limite d'une suite d'éléments de $\mathbb{C}[M]$.

Remarque. Puisque $M_n(\mathbb{C})$ est de dimension finie, toutes les normes sur $M_n(\mathbb{C})$ sont équivalentes, et la convergence ci-dessus est donc valable pour toute norme sur $M_n(\mathbb{C})$.

Étape n°2. Notons $\mathbb{C}[A]^\star$ le groupe des inversibles de l'anneau $(\mathbb{C}[A], +, \times)$. Montrons que $\mathbb{C}[A]^\star = \mathbb{C}[A] \cap GL_n(\mathbb{C})$. L'inclusion $\mathbb{C}[A]^\star \subset \mathbb{C}[A] \cap GL_n(\mathbb{C})$ est immédiate. Réciproquement, considérons $M \in \mathbb{C}[A] \cap GL_n(\mathbb{C})$. Notons π_M le polynôme minimal de M sur $\mathbb{C}$, et notons $a_0, \dots, a_n$ ses coefficients. Alors,

$$\pi_M(M) = M^n + a_{n-1}M^{n-1} + \cdots + a_1 M + a_0 I_n = 0.$$

Or, $a_0 \neq 0$ car $M \in GL_n(\mathbb{C})$. Par suite,

$$M\left(M^{n-1} + \left(-\frac{a_{n-1}}{a_0}\right)M^{n-2} + \cdots + \left(-\frac{a_1}{a_0}\right)\right) = I_n.$$

et donc,

$$M^{-1} = M^{n-1} + \left(-\frac{a_{n-1}}{a_0}\right)M^{n-2} + \cdots + \left(-\frac{a_1}{a_0}\right) \in \mathbb{C}[M]$$

ce qui prouve que $M \in \mathbb{C}[M]^*$. D'où, $\mathbb{C}[A]^\star = \mathbb{C}[A] \cap GL_n(\mathbb{C})$. Or, si M et N sont deux matrices qui commutent alors $\exp(M+N) = \exp(M)\exp(N)$. En particulier,

$$\forall (M, N) \in \mathbb{C}[A],\ \exp(M+N) = \exp(M)\exp(N)$$

(car M et N commutent en tant que polynôme en A). Par suite, l'application,

$$\begin{array}{ccc} (\mathbb{C}[A], +) & \longrightarrow & (\mathbb{C}[A]^\star, \times) \\ M & \longmapsto & \exp(M) \end{array}$$

est un morphisme de groupes ($\exp(M)$ est inversible d'inverse $\exp(-M)$).

Étape n°3. Montrons que $\mathbb{C}[A]^\star$ est connexe. Pour cela, montrons qu'il est connexe par arcs. Soient $M, N \in \mathbb{C}[A]^\star$. La fonction,

$$\begin{array}{ccc} \mathbb{C} & \longrightarrow & \mathbb{C} \\ z & \longmapsto & \det((1-z)M + zN) \end{array}$$

étant polynomiale, elle admet donc un nombre fini de zéros, notés $z_1, \dots, z_m$. De plus, ces derniers sont distincts de 0 et de 1 car $\det(M) \neq 0$ et $\det(N) \neq 0$. Puisque $\mathbb{C} \setminus \{z_1, \dots, z_m\}$ est connexe par arcs, il existe donc une application continue $\gamma : [0,1] \to \mathbb{C} \setminus \{z_1, \dots, z_m\}$ tels que $\gamma(0) = 0$ et $\gamma(1) = 1$. Montrons que l'application,

$$\begin{array}{rccc} f & : & [0,1] & \longrightarrow & \mathbb{C}[A]^\star \\ & & t & \longmapsto & (1-\gamma(t))M + \gamma(t)N \end{array}$$

est une application continue telle que $f(0) = M$ et $f(1) = N$. Soit $t \in [0,1]$. Alors, $\det(f(t)) \neq 0$ car $\gamma(t) \notin \{z_1, \dots, z_m\}$, et donc $f(t) \in GL_n(\mathbb{C})$. D'où, $f(t) \in \mathbb{C}[A] \cap GL_n(\mathbb{C}) = \mathbb{C}[A]^\star$. La fonction f est donc bien définie. De plus, f est continue en tant que composée d'applications continues, et vérifie bien $f(0) = f(1)$. Ainsi, $\mathbb{C}[A]^\star$ est connexe par arcs, donc connexe.

Étape n°4. Montrons que $\exp(\mathbb{C}[A])$ est une partie ouverte de $\mathbb{C}[A]^\star$. Pour cela, nous allons appliquer le théorème d'inversion locale à l'application,

$$\begin{array}{ccc} \mathbb{C}[A] & \longrightarrow & \mathbb{C}[A] \\ M & \longmapsto & \exp(M) \end{array}$$

En effet, cette application est différentiable (et même de classe C^∞) et sa différentielle au point $0_{\mathcal{M}_n(\mathbb{C})}$ (la matrice nulle) est $Id_{\mathbb{C}[A]}$, qui est bien inversible. D'après le théorème d'inversion locale il existe donc,

- U un ouvert de $\mathbb{C}[A]$ contenant $0_{\mathcal{M}_n(\mathbb{C})}$;

- V un ouvert de $\mathbb{C}[A]$ contenant $\exp(0_{\mathcal{M}_n(\mathbb{C})}) = I_n$;

tels que l'application $U \to V,\ B \mapsto \exp(B)$ soit un C^1-difféomorphisme. Soit $M \in \exp(\mathbb{C}[A])$. Montrons qu'il existe un ouvert de $\mathbb{C}[A]$ contenant M. Alors, MV est un ouvert de $\mathbb{C}[A]$ contenant M. En effet, $M \in MV$ car $I_n \in V$ et MV est un ouvert de $\mathbb{C}[A]$ en tant qu'image de V (qui est un ouvert de $\mathbb{C}[A]$) par l'application,

$$\begin{array}{ccc} \mathbb{C}[A] & \longrightarrow & \mathbb{C}[A] \\ B & \longmapsto & MB \end{array}$$

qui est un homéomorphisme, car $M \in GL_n(\mathbb{C})$. Ainsi, MV est un ouvert de $\mathbb{C}[A]$ contenant M. Ceci valant pour tout $M \in \exp(\mathbb{C}[A])$, on en déduit que $\exp(\mathbb{C}[A])$ est ouvert dans $\mathbb{C}[A]$. Or, $\exp(\mathbb{C}[A]) \subset \mathbb{C}^*[A]$. Donc $\exp(\mathbb{C}[A])$ est un ouvert de $\mathbb{C}^*[A]$.

Étape n°5. Montrons que $\exp(\mathbb{C}[A])$ est une partie fermée de $\mathbb{C}[A]^\star$. Tout d'abord, montrons que,

$$\mathbb{C}[A]^\star \setminus \exp(\mathbb{C}[A]) = \bigcup_{M \in \mathbb{C}[A]^\star \setminus \exp(\mathbb{C}[A])} M \exp(\mathbb{C}[A]).$$

Si $M \in \mathbb{C}[A]^\star \setminus \exp(\mathbb{C}[A])$ alors $M \in M\exp(\mathbb{C}[A])$ car $I_n = \exp(0) \in \exp(\mathbb{C}[A])$, d'où l'inclusion $\subset$. Réciproquement, considérons $M \in \mathbb{C}[A]^\star \setminus \exp(\mathbb{C}[A])$ et $N \in M\exp(\mathbb{C}[A])$. Alors, $N \notin \exp(\mathbb{C}[A])$ car sinon on aurait $M \in \exp(\mathbb{C}[A])$ (car $\exp(\mathbb{C}[A])$ est un sous-groupe de $\mathbb{C}[A]^\star$). D'où l'inclusion $\supset$, puis l'égalité. Or, pour tout $M \in \mathbb{C}[A]^\star \setminus \exp(\mathbb{C}[A])$, $M\exp(\mathbb{C}[A])$ est une partie ouverte de $\mathbb{C}^*[A]$, en tant qu'image de $\exp(\mathbb{C}[A])$ par l'homéomorphisme $B \mapsto MB$. Par suite, $\mathbb{C}[A]^\star \setminus \exp(\mathbb{C}[A])$ est un ouvert de $\mathbb{C}^*[A]$ en tant que réunion d'ouverts de $\mathbb{C}^*[A]$, et donc $\exp(\mathbb{C}[A])$ est fermé dans $\mathbb{C}[A]^\star$.

Étape n°6. Finalement, $\exp(\mathbb{C}[A])$ est une partie ouverte et fermée de $\mathbb{C}[A]^\star$. Par connexité de $\mathbb{C}[A]^\star$, on en déduit que $\exp(\mathbb{C}[A]) = \mathbb{C}[A]^\star$. En particulier, si $A \in GL_n(\mathbb{C})$, il existe $B \in \mathbb{C}[A]$ tel que $A = \exp(B)$ et donc l'application,

$$\begin{array}{ccccc} \exp & : & M_n(\mathbb{C}) & \longrightarrow & GL_n(\mathbb{C}) \\ & & A & \longmapsto & \exp(A) \end{array}$$

est surjective. □

Remarque. En revanche, elle n'est pas injective puisque,

$$\exp(\mathrm{Diag}(2i\pi, \dots, 2i\pi)) = \mathrm{Diag}(e^{2i\pi}, \dots, e^{2i\pi}) = I_n = \exp(0_{M_n(\mathbb{C})}).$$

Compléments conseillés.

- Connexité du plan complexe privé d'un nombre fini de points;
- Étude de l'exponentielle sur $M_n(\mathbb{R})$.
- Différentielle de l'exponentielle matricielle.

Commentaires personnels. Développement technique, mais qui se présente dans de nombreuses leçons. On y exploite des résultats d'analyse pour démontrer une propriété algébrique, ce qui sera apprécié par le jury, et permettra de le présenter en algèbre comme en analyse.

Théorème de Burnside

[JR] **Lemme 1.** Une matrice $N \in M_n(\mathbb{C})$ est nilpotente si et seulement si, pour tout $k \in \mathbb{N}^*$, $\mathrm{Tr}(N^k) = 0$.

Preuve. Rappelons que N est nilpotente si et seulement si sa seule valeur propre est 0. En effet, si N est nilpotente d'indice $p \in \mathbb{N}^*$, alors son polynôme minimal est X^p et donc 0 est sa seule valeur propre. Réciproquement, si 0 est la seule valeur propre de N alors son polynôme minimal est de la forme X^k pour $k \in \mathbb{N}^*$, et donc N est nilpotente.

Supposons que N soit nilpotente. Alors, pour tout $k \in \mathbb{N}^*$, N^k est nilpotente, et donc 0 est sa seule valeur propre complexe, ce qui implique $\mathrm{Tr}(N^k) = 0$, car la trace d'une matrice à coefficients complexes est égale à la somme de ses valeurs propres (complexes) comptées avec multiplicité.

Réciproquement, on suppose que, pour tout $k \in \mathbb{N}^*$, $\mathrm{Tr}(N^k) = 0$. Raisonnons par l'absurde en supposant que N ne soit pas nilpotente. Puisque χ_N est scindé (d'après le théorème de d'Alembert-Gauss), il existe $p \in \mathbb{N}^*$, $\lambda_1, \dots, \lambda_p \in \mathbb{C}^*$ deux à deux distincts et $\alpha_1, \dots, \alpha_p \in \mathbb{N}^*$ tels que,

$$\chi_N = (X - \lambda_1)^{\alpha_1} \dots (X - \lambda_p)^{\alpha_p} X^{n-(\alpha_1+\dots+\alpha_p)}.$$

Or, N est trigonalisable, car χ_N est scindé. Il existe donc $P \in GL_n(\mathbb{C})$ tel que,

$$N = PTP^{-1}$$

avec T une matrice triangulaire supérieure dont les coefficients diagonaux sont les λ_i, répétés α_i fois, puis 0 répétés $n - (\alpha_1 + \dots + \alpha_p)$ fois. En particulier, pour tout $k \in \mathbb{N}$, T^k est donc une matrice triangulaire supérieure de coefficients diagonaux λ_i^k, répétés α_i fois, puis 0 répétés $n - (\alpha_1 + \dots + \alpha_p)$ fois, d'où,

$$0 = \mathrm{Tr}(N^k) = \mathrm{Tr}(T^k) = \sum_{j=1}^{p} \alpha_j \lambda_j^k$$

car la trace est invariante par similitude. Par conséquent, $\alpha_1, \dots, \alpha_p$ vérifient le système suivant :

$$\begin{pmatrix} \lambda_1 & \cdots & \lambda_p \\ \vdots & \ddots & \vdots \\ \lambda_1^p & \cdots & \lambda_p^p \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_p \end{pmatrix} = 0.$$

Or, par multilinéarité du déterminant, on peut faire apparaître une matrice de Vandermonde (ou sa transposée, selon la convention choisie),

$$\begin{vmatrix} \lambda_1 & \cdots & \lambda_p \\ \vdots & \ddots & \vdots \\ \lambda_1^p & \cdots & \lambda_p^p \end{vmatrix} = \prod_{j=1}^{p} \lambda_j \begin{vmatrix} 1 & \cdots & 1 \\ \vdots & \ddots & \vdots \\ \lambda_1^{p-1} & \cdots & \lambda_p^{p-1} \end{vmatrix} = \prod_{j=1}^{p} \lambda_j \prod_{1 \le l < k \le p} (\lambda_k - \lambda_l) \neq 0$$

car $\lambda_1, \dots, \lambda_p$ sont deux à deux distincts. Le système précédent admet donc pour unique solution le vecteur nul, ce qui contredit $\alpha_1, \dots, \alpha_p \in \mathbb{N}^*$. □

[JR] **Théorème 2 (Théorème de Burnside).** Tout sous-groupe de $GL_n(\mathbb{C})$ d'exposant fini est fini.

Preuve. Soit G un sous-groupe de $GL_n(\mathbb{C})$ d'exposant fini $m \in \mathbb{N}^*$. Notons,

$$\mathrm{Tr}(G) = \{\mathrm{Tr}(A),\ A \in G\}.$$

Étape n°1. Montrons que tous les éléments de G sont diagonalisables et que l'ensemble $\mathrm{Tr}(G)$ est fini. Puisque G est d'exposant m, on a, pour tout $A \in G$, $A^m = I_n$, et donc A est diagonalisable et ses valeurs propres appartiennent à $\mathbb{U}_m$, car $X^m - I_n = \prod_{k=1}^{m} (X - e^{\frac{2ik\pi}{m}})$ est un polynôme annulateur de A scindé à racines simples. En particulier, l'ensemble $\mathrm{Tr}(G)$ est fini. En effet, la trace d'une matrice à coefficients complexes étant égale à la somme de ses valeurs propres (complexes) comptées avec multiplicité, et les valeurs propres de tout élément de G étant des racines m-ièmes de l'unité, il vient,

$$\mathrm{Tr}(G) \subset \left\{ \sum_{k=1}^{n} \lambda_k,\ \lambda_1, \dots, \lambda_n \in \mathbb{U}_m \right\}$$

et donc $\mathrm{Tr}(G)$ est fini.

Étape n°2. Notons $F = \mathrm{Vect}(G)$. D'après le théorème de la base incomplète, on peut extraire de G des matrices $M_1, \dots, M_p \in G$ telles que $(M_1, \dots, M_p)$ soit une base de F. On considère ensuite l'application,

$$\begin{array}{rccc} \varphi & : & G & \longrightarrow & \mathbb{C}^p \\ & & M & \longmapsto & (\mathrm{Tr}(MM_1), \dots, \mathrm{Tr}(MM_p)) \end{array}$$

Tout d'abord, $\varphi(G) \subset \mathrm{Tr}(G)^p$, et donc $\varphi(G)$ est fini d'après l'étape précédente. Montrons que φ est injective, ce qui suffira pour conclure, puisque l'image d'un ensemble infini par une application injective ne saurait être fini.

Soient $A, B \in G$ telles que $\varphi(A) = \varphi(B)$, c'est-à-dire que, pour tout $k \in [\![1, p]\!]$,

$$\mathrm{Tr}(AM_k) = \mathrm{Tr}(BM_k).$$

Puisque $(M_1, \dots, M_p)$ est une base de F, la linéarité de la trace entraîne que,

$$\forall M \in G,\ \mathrm{Tr}(AM) = \mathrm{Tr}(BM). \tag{1}$$

Montrons ensuite que $AB^{-1} - I_n$ est nilpotente. Pour cela, on utilise le lemme précédent. Pour tout $k \in \mathbb{N}^*$, en appliquant l'égalité (1) à $B^{-1}(AB^{-1})^k \in G$, il vient,

$$\mathrm{Tr}((AB^{-1})^{k+1}) = \mathrm{Tr}(AB^{-1}(AB^{-1})^k) = \mathrm{Tr}(BB^{-1}(AB^{-1})^k)$$

puis $\mathrm{Tr}((AB^{-1})^{k+1}) = \mathrm{Tr}((AB^{-1})^k)$. Par récurrence immédiate, on en déduit que, pour tout $k \in \mathbb{N}^*$,

$$\mathrm{Tr}((AB^{-1})^k) = \mathrm{Tr}(AB^{-1}) = \mathrm{Tr}(BB^{-1}) = \mathrm{Tr}(I_n) = n.$$

en appliquant de nouveau l'égalité (1) à $B^{-1} \in G$. Fixons $k \in \mathbb{N}^*$. Alors,

$$(AB^{-1} - I_n)^k = \sum_{j=0}^{n} \binom{k}{j} (-1)^{k-j} (AB^{-1})^j$$

Puis, par linéarité de la trace,

$$\mathrm{Tr}((AB^{-1} - I_n)^k) = \sum_{j=0}^{n} \binom{k}{j} (-1)^{k-j} \mathrm{Tr}((AB^{-1})^j)$$

Et donc, d'après ce qui précède,

$$\mathrm{Tr}((AB^{-1} - I_n)^k) = n \sum_{j=0}^{n} \binom{k}{j} (-1)^{k-j} = n(1-1)^j = 0.$$

Le lemme précédent permet d'en déduire que $AB^{-1} - I_n$ est nilpotente. Or, $AB^{-1} - I_n \in G$ est diagonalisable car tout élément de G l'est. Donc, $AB^{-1} - I_n$ est semblable à la matrice nulle, ce qui implique $AB^{-1} - I_n = 0$, c'est-à-dire $A = B$. Ainsi, φ est injective, et donc G est fini car $\varphi(G)$ est fini. □

Commentaires personnels. Pour être tout à fait honnête, je l'avais présenté en oral blanc et j'avais complètement foiré la fin. Je ne le maîtrisais pas encore suffisamment bien, et j'avais écrit $\mathrm{Tr}((AB^{-1})^j) = (\mathrm{Tr}(AB^{-1}))^j$, ce qui donnait ensuite :

$$\mathrm{Tr}((AB^{-1} - I_n)^k) = \sum_{j=0}^{n} \binom{k}{j} (-1)^{k-j} n^j$$

m'empêchant donc de conclure. Or, l'égalité $\mathrm{Tr}((AB^{-1})^j) = (\mathrm{Tr}(AB^{-1}))^j$ est fausse a priori (la trace n'est pas multiplicative). Par exemple, si $A = B = I_n$ alors $\mathrm{Tr}((AB^{-1})^j) = \mathrm{Tr}(I_n) = n$ et $(\mathrm{Tr}(AB^{-1}))^j = (\mathrm{Tr}(I_n))^j = n^j$. Malgré cette mésaventure, j'apprécie ce développement, qui se comprend et se retient facilement, et peut être présenté dans plusieurs leçons. Je le recommande.

Théorème de Cauchy

Définition 1. Soit p un nombre premier. Un groupe fini est appelé un p-groupe si son ordre est une puissance de p.

Proposition 2. Soit p un nombre premier. Soit G un p-groupe agissant sur un ensemble X. Notons,

$$X^G = \{x \in X,\ \forall g \in G,\ g \cdot x = x\}.$$

Alors,

$$\left|X^G\right| \equiv |X| \ [p].$$

Preuve. Soit $x_1, \dots, x_r \in E$ tels que $G \cdot x_1, \dots, G \cdot x_r$ soient les orbites deux à deux distinctes pour l'action de G sur X. D'après la formule des classes,

$$|X| = \sum_{i=1}^{r} [G : G_{x_i}].$$

En isolant les orbites de cardinal supérieur ou égal à 1, on obtient donc,

$$|X| = \sum_{i=1}^{|X^G|} 1 + \sum_{i=1}^{r-|X^G|} |G \cdot x_i| = \left|X^G\right| + \sum_{i=1}^{r-|X^G|} |G \cdot x_i|$$

Or, d'après le théorème de Lagrange, pour tout $i \in [\![1, r]\!]$, $|G \cdot x_i|$ divise $|G|$ (car $G \cdot x_i$ est un sous-groupe de G), et donc p divise $|G \cdot x_i|$ (car $|G|$ est une puissance de p et $|G \cdot x_i| \neq 1$). Donc,

$$\left|X^G\right| \equiv |X| \ [p].$$

□

[JR] **Théorème 3 (de Cauchy).** Soit G un groupe fini de cardinal $n \in \mathbb{N}^*$. Si p est un diviseur premier de n, alors il existe un élément d'ordre p.

Preuve. Notons H le sous-groupe de $\mathfrak{S}_p$ engendré par le p-cycle $(1, 2, \dots, p)$. Le groupe symétrique $\mathfrak{S}_p$ agit sur G^p par permutation des composantes :

$$\begin{array}{ccc} \mathfrak{S}_p \times G^p & \longrightarrow & G^p \\ (\sigma, (g_1, \dots, g_p)) & \longmapsto & (g_{\sigma(1)}, \dots, g_{\sigma(p)}) \end{array}$$

En particulier, H agit également sur G^p. Notons,

$$E = \{(g_1, \dots, g_p) \in G^p,\ g_1 \dots g_p = 1\}.$$

Remarquons que s'il existe un élément $g \in G$ d'ordre p, alors $(g, \dots, g) \in E$ (d'où l'intérêt d'étudier cet ensemble). Pour tout $(g_1, \dots, g_p) \in G^p$,

$$g_1 \dots g_{p-1} g_p = 1 \Longrightarrow g_p^{-1} = g_1 \dots g_{p-1} \Longrightarrow g_p g_1 \dots g_{p-1} = 1$$

c'est-à-dire que si $(g_1, \dots, g_p) \in E$ alors $(1, \dots, p) \cdot (g_1, \dots, g_p) \in E$, et donc l'ensemble E est stable sous action de H. D'autre part, en remarquant que l'application,

$$\begin{array}{ccc} G^{p-1} & \longrightarrow & E \\ (g_1, \dots, g_{p-1}) & \longmapsto & (g_1, \dots, g_{p-1}, (g_1 \dots g_{p-1})^{-1}) \end{array}$$

est bijective, on obtient que $|E| = n^{p-1}$. En particulier, p divise $|E|$. Or, d'après la proposition précédente,

$$E^G \equiv |E| \ [p]$$

donc p divise également E^G. Or, E^G est non vide (car $(e, \dots, e) \in E^G$). Donc E^G possède au moins p éléments. Soit $(g_1, \dots, g_p) \in E^G \setminus \{(e, \dots, e)\}$. Alors,

$$(g_1, \dots, g_{p-1}, g_p) = (1, \dots, p-1, p) \cdot (g_1, \dots, g_{p-1}, g_p) = (g_2, \dots, g_p, g_1)$$

ce qui implique,

$$\begin{cases} g_1 &= g_2, \\ &\vdots \\ g_{p-1} &= g_p, \\ g_p &= g_1, \end{cases}$$

et donc $g_1 = \dots = g_p$. Ainsi, g_1 est un élément de $G \setminus \{e\}$ tel que $g_1^p = e$, ce qui prouve que g_1 est d'ordre p (car p est premier). Finalement, on en conclut que G admet un élément d'ordre p. □

Application 4 (Classification des groupes d'ordre 6). Soit G un groupe fini d'ordre 6. Alors, G est isomorphe à $\mathbb{Z}/6\mathbb{Z}$ ou à $\mathfrak{S}_3$.

Preuve. Si G admet un élément d'ordre 6, alors G est cyclique d'ordre 6, donc isomorphe à $\mathbb{Z}/6\mathbb{Z}$. Sinon, on considère $a \in G$ d'ordre 2 et $b \in G$ d'ordre 3 (qui existent d'après le théorème de Cauchy). Montrons que G possède exactement trois éléments d'ordre 2 et deux éléments d'ordre 3.

• Tout d'abord, b et b^{-1} sont d'ordre 3 et $b \neq b^{-1}$ (car sinon on aurait $b^2 = e$). Donc G possède au moins deux éléments d'ordre 3.

• Ensuite, le lemme 1 assure que bab^{-1} et $b^{-1}ab$ sont d'ordre 2. Or, $ab \neq ba$ car sinon ab serait d'ordre 6 (d'après le lemme 2). Donc $bab^{-1} \neq a$ et $b^{-1}ab \neq a$. De plus, $bab^{-1} \neq b^{-1}ab$ car sinon on aurait $b^2a = ab^2$, et donc,

$$ab = ab^4 = ab^2b^2 = b^2ab^2 = b^2b^2a = b^4a = ba$$

ce qui contredirait $ab \neq ba$. Il y a donc au moins trois éléments d'ordre 2.

Ainsi, dans G il y a :

- un élément d'ordre 1 (l'élément neutre e);
- trois éléments d'ordre 2 (qui sont a, bab^{-1} et $b^{-1}ab$);
- deux éléments d'ordre 3 (qui sont b et b^{-1}).

De ces informations, nous allons déduire la table de Cayley de G (sa table de multiplication). Ceci nous permettra de déterminer G à isomorphisme près.

• D'après le lemme 1, aba est d'ordre 3 (car c'est le conjugué de b par a, a^{-1} étant égal à a). De plus, $aba \neq b$ (sinon on aurait $ab = ba$ en multipliant par a à droite, ce qui n'est pas le cas). Donc, $aba = b^{-1}$. Cette égalité va nous permettre de calculer ab et ba.

• Montrons que $ab = bab^{-1}$. Tout d'abord, puisque $aba = b^{-1}$, il vient,

$$(ab)^2 = (aba)b = b^{-1}b = e$$

De plus $ab \neq e$ (car sinon on aurait $a = b^{-1}$, ce qui est impossible étant donné que a et b^{-1} sont d'ordre différent). Donc ab est d'ordre 2. Or, $ab \neq b^{-1}ab$ car $b^{-1} \neq e$. De même, $ab \neq a$ car $b \neq e$. Ainsi, ab est un élément de G d'ordre 2, distinct de a et $b^{-1}ab$, ce qui implique $ab = bab^{-1}$ (car c'est le seul élément d'ordre 2 restant).

• De façon analogue, on obtient $ba = b^{-1}ab$.

• Des égalités $aba = b^{-1}$, $ab = bab^{-1}$ et $ba = b^{-1}ab$, on en déduit que,

$$\begin{cases} bab^{-1}a = aba = b^{-1}, \\ b^{-1}aba = ba^2 = b, \\ b^{-1}a = aba^2 = ab = bab^{-1}, \\ abab^{-1} = a^2b = b, \\ b^{-1}abbab^{-1} = ba^2b = b^2 = b^{-1}, \\ bbab^{-1} = bb^{-1}abb^{-1} = a, \\ b^{-1}bab^{-1} = ab^{-1} = a^2ba = ba = b^{-1}ab, \\ ab^{-1}ab = aba = b^{-1}, \end{cases}$$

$$\begin{cases} bab^{-1}b^{-1}ab = babab = bb^{-1}b = b, \\ bb^{-1}ab = ab = bab^{-1}, \\ b^{-1}b^{-1}ab = b^{-1}ba = a, \\ bab^{-1}b = ba = b^{-1}ab, \\ b^{-1}abb = b^{-1}bab^{-1}b = a, \\ ab^{-1} = b^3ab^{-1} = b^2ab = b^{-1}ab, \\ bab^{-1}b^{-1} = bb^{-1}abb^{-1} = a, \\ b^{-1}abb^{-1} = b^{-1}a = bab^{-1}. \end{cases}$$

Finalement, on obtient que la table de Cayley de G est la table suivante :

$\cdot$	e	a	bab^{-1}	$b^{-1}ab$	b	b^{-1}
e	e	a	bab^{-1}	$b^{-1}ab$	b	b^{-1}
a	a	e	b	b^{-1}	bab^{-1}	$b^{-1}ab$
bab^{-1}	bab^{-1}	b^{-1}	e	b	$b^{-1}ab$	a
$b^{-1}ab$	$b^{-1}ab$	b	b^{-1}	e	a	bab^{-1}
b	b	$b^{-1}ab$	a	bab^{-1}	b^{-1}	e
b^{-1}	b^{-1}	bab^{-1}	$b^{-1}ab$	a	e	b

Il s'agit de la même table de Cayley que celle de $\mathfrak{S}_3$, donnée ci-dessous :

$\circ$	Id	$(1,2)$	$(2,3)$	$(1,3)$	$(1,2,3)$	$(1,3,2)$
Id	Id	$(1,2)$	$(2,3)$	$(1,3)$	$(1,2,3)$	$(1,3,2)$
$(1,2)$	$(1,2)$	Id	$(1,2,3)$	$(1,3,2)$	$(2,3)$	$(1,3)$
$(2,3)$	$(2,3)$	$(1,3,2)$	Id	$(1,2,3)$	$(1,3)$	$(1,2)$
$(1,3)$	$(1,3)$	$(1,2,3)$	$(1,3,2)$	Id	$(1,2)$	$(2,3)$
$(1,2,3)$	$(1,2,3)$	$(1,3)$	$(1,2)$	$(2,3)$	$(1,3,2)$	Id
$(1,3,2)$	$(1,3,2)$	$(2,3)$	$(1,3)$	$(1,2)$	Id	$(1,2,3)$

Par conséquent, G et $\mathfrak{S}_3$ sont isomorphes. Un exemple d'isomorphisme de G vers $\mathfrak{S}_3$ est le morphisme de groupes $\varphi : G \to \mathfrak{S}_3$ défini par,

$$\begin{cases} \varphi(e) = Id, \\ \varphi(a) = (1,2), \\ \varphi(bab^{-1}) = (2,3), \\ \varphi(b^{-1}ab) = (1,3), \\ \varphi(b) = (1,2,3), \\ \varphi(b^{-1}) = (1,3,2). \end{cases}$$

□

Complément conseillé.

- Ordre d'un élément dans un groupe : lemme 2.

Commentaires personnels. Pour être tout à fait honnête, je n'avais pas eu le temps de bien préparer ce développement, car je l'avais choisi tardivement, sous les recommandations d'un ami (coucou Max). La preuve du théorème de Cauchy étant assez courte, j'ai choisi de rajouter la classification des groupes d'ordre 6. Lors de l'exposé, n'écrivez pas tous les calculs ainsi que la table de multiplication de G. Montrez seulement les égalités $aba = b^{-1}$, $ab = bab^{-1}$ et $ba = b^{-1}ab$, puis expliquez que ces dernières permettent d'obtenir la table de multiplication de G, puis de conclure que G est isomorphe à $\mathfrak{S}_3$. Je pense que c'est un bon choix de développement.

Théorème de d'Alembert-Gauss

[JR] **Théorème 1 (d'Alembert-Gauss).** Le corps $\mathbb{C}$ est algébriquement clos.

Preuve. Soit $P \in \mathbb{C}[X]$ de degré $n \geq 1$. Notons $a_0, \ldots, a_n$ ses coefficients. On procède en plusieurs étapes.

Étape n°1. Montrons que $\lim\limits_{|z|\to+\infty} |P(z)| = +\infty$. Pour tout $z \in \mathbb{C}^*$,

$$|P(z)| = |z|^n \left| \sum_{k=0}^{n} a_k z^{k-n} \right|.$$

Or,

$$\sum_{k=0}^{n} a_k z^{k-n} = 1 + \sum_{k=0}^{n-1} a_k z^{k-n} \underset{|z|\to+\infty}{\longrightarrow} 1.$$

car, pour tout $k \in [\![0, n-1]\!]$, $\lim\limits_{|z|\to+\infty} a_k z^{k-n} = 0$. Donc,

$$|P(z)| \underset{|z|\to+\infty}{\longrightarrow} +\infty.$$

Étape n°2. Montrons qu'il existe $z_0 \in \mathbb{C}$ tel que $|P(z_0)| = \min\limits_{z\in\mathbb{C}} |P(z)|$. D'après l'étape précédente, il existe $R > 0$ tel que, pour tout $z \in \mathbb{C}$,

$$|z| > R \Longrightarrow |P(z)| > |P(0)|.$$

D'autre part, puisque la fonction $z \mapsto P(z)$ est continue sur $\overline{D(0,R)}$ qui est compact, il existe $z_0 \in \overline{D(0,R)}$ tel que $|P(z_0)| = \min\limits_{z\in\overline{D(0,R)}} |P(z)|$. Par suite,

$$|P(z_0)| = \min_{z\in\mathbb{C}} |P(z)|.$$

En effet, pour tout $z \in \mathbb{C}$,

- si $z \in \overline{D(0,R)}$ alors $|P(z)| \geq |P(z_0)|$,
- sinon $|P(z)| > |P(0)| \geq |P(z_0)|$,

et donc $|P(z)| \geq |P(z_0)|$ dans tous les cas.

Étape n°3. Montrons finalement que z_0 est une racine de P. Pour cela, on raisonne par l'absurde en supposant que $P(z_0) \neq 0$. Notre stratégie consiste à réaliser une étude asymptotique de P afin de trouver un nombre complexe z "proche" de z_0 tel que $|P(z)| < |P(z_0)|$. Tout d'abord, d'après la formule de Taylor (pour les polynômes),

$$P(z + z_0) = \sum_{k=0}^{n} \frac{P^{(k)}(z_0)}{k!} z^k.$$

Pour tout $k \in [\![0, n]\!]$, notons,

$$b_k = \frac{P^{(k)}(z_0)}{k!}.$$

Puisque $P(z_0) \neq 0$, on a $b_0 \neq 0$, ce qui permet d'écrire, pour tout $z \in \mathbb{C}$,

$$P(z + z_0) = b_0 \left(1 + \sum_{k=1}^{n} \frac{b_k}{b_0} z^k \right).$$

On cherche donc $z \in \mathbb{C}$ de telle sorte que,

$$\left| 1 + \sum_{k=1}^{n} \frac{b_k}{b_0} z^k \right| < 1.$$

Lorsque z est proche de 0, le terme prépondérant dans le module ci-dessus est le terme de degré le plus petit. C'est pourquoi on note p le plus petit entier de $[\![1, n]\!]$ tel que b_p soit non nul. Avec ces notations, il vient, pour tout $z \in \mathbb{C}$,

$$1 + \sum_{k=1}^{n} \frac{b_k}{b_0} z^k = 1 + \sum_{k=p}^{n} \frac{b_k}{b_0} z^k = 1 + \frac{b_p}{b_0} z^p \sum_{k=p}^{n} \frac{b_k}{b_p} z^{k-p} = 1 + \frac{b_p}{b_0} z^p (1 + \varepsilon(z))$$

où l'on a noté,

$$\varepsilon(z) = \begin{cases} 0 & \text{si } p = n, \\ \displaystyle\sum_{k=p}^{n} \frac{b_k}{b_p} z^{k-p} & \text{sinon.} \end{cases}$$

Soit ω une racine p-ième de $-\dfrac{b_p}{b_0}$. Alors, pour tout $z \in \mathbb{C}$,

$$P(z + z_0) = b_0[1 - (\omega z)^p (1 + \varepsilon(z))].$$

avec $\lim\limits_{z\to 0} \varepsilon(z) = 0$. Et donc, pour tout $t \in \mathbb{R}$,

$$P\left(z_0 + \frac{t}{\omega}\right) = b_0 \left[1 - t^p \left(1 + \varepsilon\left(\frac{t}{\omega} \right) \right) \right].$$

Puisque $\lim_{z\to 0} \varepsilon(z) = 0$, il existe $\alpha \in]0,1[$ tel que,

$$\forall t \in]0,\alpha],\ \left|\varepsilon\left(\frac{t}{\omega}\right)\right| \leq \frac{1}{2}$$

Par inégalité triangulaire, il s'en suit que, pour tout $t \in]0,\alpha]$,

$$\left|P\left(z_0 + \frac{t}{\omega}\right)\right| \leq |b_0|\left[|1-t^p| + |t|^p\left|\varepsilon\left(\frac{t}{\omega}\right)\right|\right] \leq |b_0|\left[1 - t^p + \frac{1}{2}t^p\right]$$

(on peut enlever les valeurs absolues car $t \in [0,1]$). D'où, pour tout $t \in]0,\alpha]$,

$$\left|P\left(z_0 + \frac{t}{\omega}\right)\right| \leq |b_0|\left[1 - \frac{1}{2}t^p\right] < |b_0| = |P(z_0)|$$

ce qui contredit $|P(z_0)| = \min_{z\in\mathbb{C}} |P(z)|$. Ainsi, $P(z_0) = 0$. □

Remarque. En réalité, nous avons démontré que si $P(z_0) \neq 0$, alors on peut trouver z aussi proche que l'on veut de z_0 tel que $|P(z)| < |P(z_0)|$.

Commentaires personnels. Développement un peu technique, mais qui se comprend facilement. On exploite des résultats d'analyse pour démontrer une propriété algébrique, ce qui sera apprécié par le jury. De plus, le théorème de d'Alembert-Gauss figure sur le programme du concours.

Théorème de Dirichlet

Lemme 1. Soient $n \in \mathbb{N}^*$ et p un nombre premier. Supposons qu'il existe $a \in \mathbb{Z}$ tel que :

(i) p divise $\Phi_n(a)$;

(ii) pour tout $d \in \mathbb{N}^*$ tel que $d \mid n$ et $d \neq n$, p ne divise pas $\Phi_d(a)$.

Alors, $p \equiv 1\ [n]$.

Preuve. D'après le lemme 6,

$$X^n - 1 = \prod_{d|n} \Phi_d. \tag{1}$$

En particulier,

$$a^n - 1 = \prod_{d|n} \Phi_d(a).$$

Or, $\Phi_n(a) \in p\mathbb{Z}$ par hypothèse. Donc $a^n - 1 \in p\mathbb{Z}$, c'est-à-dire $[a]_p^n = 1$ dans $\mathbb{Z}/p\mathbb{Z}$. Par conséquent, l'ordre de $[a]_p$ dans $(\mathbb{Z}/p\mathbb{Z})^*$ divise n. Montrons que ce dernier est même égal à n. Soit l un diviseur strict de n. D'après l'égalité (1),

$$[a^l - 1]_p = \left[\prod_{d|l} \Phi_d(a)\right]_p = \prod_{d|l} [\Phi_d(a)]_p$$

dans $\mathbb{Z}/p\mathbb{Z}$. Or, tout diviseur d de l est un diviseur strict de n, et donc, par hypothèse, $[\Phi_d(a)]_p \neq [0]_p$ dans $\mathbb{Z}/p\mathbb{Z}$. Par intégrité de $\mathbb{Z}/p\mathbb{Z}$, il s'en suit que,

$$\prod_{d|l} [\Phi_d(a)]_p \neq [0]_p$$

et donc que $[a^l]_p \neq [1]_p$. Ainsi, $[a]_p$ est d'ordre n dans le groupe multiplicatif $(\mathbb{Z}/p\mathbb{Z})^*$. D'après le théorème de Lagrange, il s'en suit que n divise $p - 1$, c'est-à-dire que $p \equiv 1\ [n]$. □

[FG] **Théorème 2 (Théorème de Dirichlet).** Soit $n \in \mathbb{N}^*$. Alors, il existe une infinité de nombres premiers congrus à 1 modulo n.

Preuve. On raisonne par l'absurde en supposant qu'il existe un nombre fini de nombres premiers $p_1, \ldots, p_m$ congrus à 1 modulo n. Nous allons montrer qu'il existe en fait un autre nombre premier p congru à 1 modulo n. Le lemme précédent fournit une condition suffisante pour que p soit congru à 1 modulo n, mais ne garantit pas que p soit distinct de $p_1, \ldots, p_m$. C'est pourquoi on pose, [FG]

$$N = np_1 \ldots p_m.$$

Alors, pour tout $i \in \{1, \ldots, m\}$, p_i n'est pas congru à 1 modulo N. En effet, sinon on aurait N qui diviserait $p_i - 1$, puis p_i qui diviserait $p_i - 1$ car p_i divise N, et donc p_i diviserait $p_i - (p_i - 1) = 1$, ce qui contredirait le fait que p_i soit premier. Par conséquent, si p est un nombre premier congru à 1 modulo N, alors p est un nombre premier congru à 1 modulo n distinct de $p_1, \ldots, p_m$, et donc p convient à notre problème. Posons ensuite,

$$P = \prod_{\substack{d|N \\ d \neq N}} \Phi_d.$$

Nous allons donc montrer qu'il existe un nombre premier p et un entier relatif a tels que p divise $\Phi_N(a)$ mais ne divise pas $P(a)$, ce qui suffira pour conclure que p est congru à 1 modulo N d'après le lemme précédent. D'après le lemme 6, Φ_N et P sont premiers entre eux. Par conséquent, le théorème de Bézout assure l'existence de $U, V \in \mathbb{Q}[X]$ tels que,

$$\Phi_N U + PV = 1. \tag{2}$$

Puisque l'ensemble des entiers relatifs k tels que $kU \in \mathbb{Z}[X]$ et $kV \in \mathbb{Z}[X]$ est infini (car il contient tous les multiples du PPCM des dénominateurs des coefficients de U et V) et que le polynôme Φ_N est non constant, on peut considérer un entier relatif a tel que $aU \in \mathbb{Z}[X]$, $aV \in \mathbb{Z}[X]$, $\Phi_N(a) \neq 0$ et $\Phi_N(a) \neq \pm 1$. Les deux derniers points sont cruciaux si l'on espère trouver un nombre premier p divisant $\Phi_N(a)$. Posons ensuite,

$$S = aU \in \mathbb{Z}[X], \quad T = aV \in \mathbb{Z}[X].$$

De plus, d'après le lemme 6, $\Phi_N \in \mathbb{Z}[X]$ et $P \in \mathbb{Z}[X]$. En multipliant l'égalité (2) par a, il vient donc $\Phi_N S + PT = a$ dans $\mathbb{Z}[X]$. En particulier,

$$\Phi_N(a)S(a) + P(a)T(a) = a \tag{3}$$

dans $\mathbb{Z}$. D'autre part, d'après le lemme 6,

$$a^N - 1 = \Phi_N(a)P(a)$$

dans $\mathbb{Z}$. Soit p un nombre premier divisant $\Phi_N(a)$. Alors, d'après l'égalité ci-dessus, p divise $a^N - 1$. Par suite, p ne divise pas a, car sinon p diviserait

a^N, et donc p diviserait $a^N - (a^N - 1) = 1$, ce qui contredirait le fait que p soit premier. Finalement, on en déduit que p ne divise pas $P(a)$, car sinon p diviserait a en conséquence de l'égalité (3). Ainsi, p est un nombre premier divisant $\Phi_N(a)$ mais ne divisant pas $P(a)$, ce qui achève la preuve. □

Remarque. Autre référence, JR, chapitre 12 théorème 12.38.

Compléments conseillés.

- Racines primitives de l'unité et polynômes cyclotomiques : sous-section 3.22.
- Connaître les preuves des cas particuliers [JR] exercice 11.3.

Commentaires personnels. Développement très sympathique, qui se comprend et se retient facilement, ce qui en fait un choix stratégique même s'il rentre dans peu de leçons.

Théorème de Kronecker

[JR] **Théorème 1 (Théorème de Kronecker).** Soit $P \in \mathbb{Z}[X]$ unitaire. Si les racines complexes de P sont non nulles et de module inférieur ou égal à 1, alors ce sont des racines de l'unité.

Preuve. Pour tout $n \in \mathbb{N}^*$, on note E_n l'ensemble des polynômes unitaires de degré n à coefficients dans $\mathbb{Z}$ dont toutes les racines complexes sont non nulles et de module inférieur ou égal à 1, puis R_n l'ensemble des racines complexes des polynômes appartenant à E_n. Soit $n \in \mathbb{N}^*$.

Étape n°1. Montrons que R_n est fini. Soit $P \in E_n$. Notons $a_0, \dots, a_n \in \mathbb{Z}$ ses coefficients et $\alpha_1, \dots, \alpha_n$ ses racines complexes. On a donc,

$$P(X) = \sum_{k=0}^{n} a_k X^k = \prod_{k=1}^{n} (X - \alpha_k)$$

avec $a_n = 1$ car P est unitaire. D'après les relations entre coefficients et racines, pour tout $k \in [\![0, n-1]\!]$,

$$a_k = (-1)^k \sigma_{n-k,n}(\alpha_1, \dots, \alpha_n) = (-1)^k \sum_{1 \leq i_1 < \dots < i_{n-k} \leq n} \alpha_{i_1} \dots \alpha_{i_{n-k}}.$$

En particulier, par inégalité triangulaire, pour tout $k \in [\![0, n-1]\!]$,

$$|a_k| \leq \sum_{1 \leq i_1 < \dots < i_{n-k} \leq n} |\alpha_{i_1}| \dots |\alpha_{i_{n-k}}| \leq \binom{n}{n-k}$$

car, pour tout $j \in [\![1, n]\!]$, $|\alpha_j| \leq 1$. Sachant que $a_0, \dots, a_n$ sont entiers, il y a donc un nombre fini de possibilités pour chaque coefficient a_k, et donc pour P. Ainsi, E_n est fini. Sachant que tout polynôme de degré n admet au plus n racines distinctes, on en déduit que $\mathrm{card}(R_n) \leq n \,\mathrm{card}(E_n)$, et donc R_n est fini.

Étape n°2. Soit $\alpha \in R_n$, c'est-à-dire qu'il existe $P \in E_n$ tel que $P(\alpha) = 0$. Montrons que $\alpha^2 \in R_n$. Notons ensuite $a_0, \dots, a_n \in \mathbb{Z}$ les coefficients de P et $\alpha_2, \dots, \alpha_n$ ses racines complexes restantes. Notons $\alpha_1 = \alpha$ et,

$$Q = \prod_{k=1}^{n} (X - \alpha_k^2),$$

puis montrons que $Q \in E_n$. Tout d'abord, Q est unitaire, et ses racines complexes sont non nulles et de module inférieur ou égal à 1. Il reste à montrer que ses coefficients sont entiers. Pour cela, on remarque que,

$$P(-X) = (-1)^n \prod_{k=1}^{n} (X + \alpha_k).$$

Par suite,

$$(-1)^n P(-X)P(X) = \prod_{k=1}^{n} (X + \alpha_k)(X - \alpha_k) = \prod_{k=1}^{n} (X^2 - \alpha_k^2) = Q(X^2),$$

et donc $Q(X^2) \in \mathbb{Z}[X]$. Par ailleurs, en notant $b_0, \dots, b_n$ les coefficients de Q, on obtient que,

$$Q(X^2) = \sum_{k=0}^{n} b_k (X^2)^k = \sum_{k=0}^{n} b_k X^{2k}.$$

Sachant que $Q(X^2) \in \mathbb{Z}[X]$, il vient alors $b_0, \dots, b_n \in \mathbb{Z}$, puis $Q \in \mathbb{Z}[X]$, et donc $\alpha^2 \in R_n$. Par récurrence immédiate, il s'en suit que,

$$\left\{ \alpha^{2^k},\ k \in \mathbb{N} \right\} \subset R_n.$$

Étape n°3. Montrons finalement que α est une racine de l'unité. Puisque R_n est fini, l'inclusion précédente fournit que l'ensemble,

$$\left\{ \alpha^{2^k},\ k \in \mathbb{N} \right\}$$

est fini, ce qui assure l'existence de $k, l \in \mathbb{N}$ tels que $l \neq k$ et $\alpha^{2^k} = \alpha^{2^l}$. Or, $\alpha \neq 0$ car $P(0) \neq 0$. D'où, $\alpha^{|2^l - 2^k|} = 1$ avec $|2^l - 2^k| \in \mathbb{N}^*$, ce qui prouve que α est une racine de l'unité. Ainsi, tout élément de R_n est une racine de l'unité, ce qui achève la preuve. □

Corollaire 2. Soit $P \in \mathbb{Z}[X]$ unitaire et irréductible sur $\mathbb{Q}$. Si les racines complexes de P sont de module inférieur ou égal à 1, alors $P = X$ ou $P = \Phi_k$ pour $k \in \mathbb{N}^*$.

Preuve. Soit α une racine complexe de P. Si 0 est racine de P, alors X divise P, et donc $P = X$ car P est unitaire et irréductible sur $\mathbb{Q}$. Sinon, le théorème précédent assure que α est une racine de l'unité. Donc, en notant k l'ordre de α dans le groupe multiplicatif $\mathbb{U}$, il s'en suit que α est racine de Φ_k. Or, P est le polynôme minimal de α sur $\mathbb{Q}$, car P est un polynôme annulateur de α, unitaire, et irréductible sur $\mathbb{Q}$. Donc Φ_k divise P, ce qui implique $P = \Phi_k$ car ces deux polynômes sont unitaires et irréductibles sur $\mathbb{Q}$. □

Remarque. Le théorème est faux si P n'est pas unitaire. Par exemple, $2X - 1$ est à coefficients entiers, et son unique racine est $\frac{1}{2}$ qui est bien non nulle et de module inférieur ou égal à 1, mais ce n'est pas une racine de l'unité. De même, le théorème est faux si P n'est pas à coefficients entiers. Par exemple, $X - \frac{1}{2}$.

Commentaires personnels. Développement très sympathique, qui se comprend et se retient facilement, ce qui en fait un choix stratégique même s'il rentre dans peu de leçons.

Théorème des deux carrés de Fermat

Théorème 1 (Théorème de Wilson). Soit $p \geq 2$ un entier naturel. Alors, p est premier si et seulement si $(p-1)! \equiv -1\,[p]$.

Preuve. Supposons que p soit premier. Remarquons tout d'abord que $[(p-1)!]_p$ est le produit des éléments de $(\mathbb{Z}/p\mathbb{Z})^*$. En particulier, pour tout élément a de $(\mathbb{Z}/p\mathbb{Z})^*$, a et a^{-1} sont tout deux présents dans ce produit. Ainsi, si $a \neq a^{-1}$, ces deux termes se simplifieront dans ce produit. Il reste ensuite à voir à quelle condition $a \neq a^{-1}$ (et c'est là qu'intervient l'hypothèse que p est premier). En effet, $a = a^{-1}$ si et seulement si $a^2 = 1$ et donc si et seulement si $a = \pm 1$ car $\mathbb{Z}/p\mathbb{Z}$ est intègre. Par suite, -1 et 1 sont les seuls éléments de $(\mathbb{Z}/p\mathbb{Z})^*$ égaux à leur inverse. Si $p = 2$, on a $(p-1)! = 1 \equiv -1\,[2]$ et si $p = 3$ on a $(p-1)! = 2 \equiv -1\,[3]$. Sinon, il existe $a_1, \ldots, a_k \in (\mathbb{Z}/p\mathbb{Z})^*$ (avec $k = \frac{p-3}{2}$) tels que $-1, 1, x_1, \ldots, x_k, x_1^{-1}, \ldots, x_k^{-1}$ soient les éléments deux à deux distincts de $(\mathbb{Z}/p\mathbb{Z})^*$. En particulier,

$$[(p-1)!]_p = \prod_{x \in (\mathbb{Z}/p\mathbb{Z})^*} x = -\prod_{l=1}^{k} x_l x_l^{-1} = -1$$

c'est-à-dire $(p-1)! \equiv -1\,[p]$. Dans tous les cas, le résultat est donc vérifié.

Réciproquement, supposons que $(p-1)! \equiv -1\,[p]$. On raisonne par l'absurde en supposant qu'il existe un diviseur a de p tel que $1 < a < p$. Alors, $a \mid (p-1)!$, car a apparaît dans le produit,

$$(p-1)! = 1 \times 2 \times \cdots \times (p-2) \times (p-1).$$

De plus, a divise p qui divise $(p-1)!+1$ par hypothèse, d'où a divise $(p-1)!+1$ par transitivité. Par suite, a divise $1 = (p-1)!+1-(p-1)!$ ce qui est absurde. Ainsi, p est bien un nombre premier. □

Théorème 2. L'anneau $\mathbb{Z}[i]$ est euclidien.

Preuve. Montrons que $\mathbb{Z}[i]$ est euclidien pour le stathme,

$$\begin{array}{rccc} N & : & \mathbb{Z}[i] & \longrightarrow & \mathbb{N} \\ & & \alpha = m + ni & \longmapsto & |m+ni|^2 \end{array}$$

Soient $\alpha \in \mathbb{Z}[i]$ et $\beta \in \mathbb{Z}[i] \setminus \{0\}$. On cherche donc $\gamma, \rho \in \mathbb{Z}[i]$ tels que,

$$\begin{cases} \alpha = \beta\gamma + \rho, \\ N(\rho) < N(\gamma). \end{cases}$$

Pour cela, il suffit de construire $\gamma \in \mathbb{Z}[i]$ tel que,

$$N\left(\frac{\alpha}{\beta} - \gamma\right) < 1.$$

En effet, le nombre complexe $\rho = \alpha - \beta\gamma \in \mathbb{Z}[i]$ vérifiera alors,

$$N(\rho) = N(\beta) N\left(\frac{\alpha}{\beta} - \gamma\right) < N(\beta)$$

et donc γ et ρ conviendront à notre problème. On souhaite donc construire un entier de Gauss γ qui soit à une distance du complexe $\frac{\alpha}{\beta}$ strictement inférieure à 1. Un tel entier de Gauss existe toujours mais il n'est pas unique comme le montre la figure 7 (sur celle-ci, $\gamma_1, \gamma_2, \gamma_3$ et γ_4 conviennent). Pour le prouver, on note,

$$x = \Re\left(\frac{\alpha}{\beta}\right),\ y = \Im\left(\frac{\alpha}{\beta}\right) \in \mathbb{R}.$$

Soit $\gamma = m + in \in \mathbb{Z}[i]$. Alors,

$$N\left(\frac{\alpha}{\beta} - \gamma\right) = (x-m)^2 + (y-n)^2.$$

Pour que cette quantité soit strictement inférieure à 1, il suffit que,

$$0 \leq |x-m| \leq \frac{1}{2},\ 0 \leq |y-n| \leq \frac{1}{2}$$

Et pour que ces inégalités soient vérifiées, on prend,

$$m = \begin{cases} \lfloor x \rfloor & \text{si } x \leq \lfloor x \rfloor + \frac{1}{2} \\ \lfloor x \rfloor + 1 & \text{sinon.} \end{cases}$$

et,

$$n = \begin{cases} \lfloor y \rfloor & \text{si } y \leq \lfloor y \rfloor + \frac{1}{2} \\ \lfloor y \rfloor + 1 & \text{sinon.} \end{cases}$$

puis $\gamma = m + in$ (sur la figure 7, γ correspond à γ_1). En effet, on a alors,

$$N\left(\frac{\alpha}{\beta} - \gamma\right) = (x-m)^2 + (y-n)^2 \leq \left(\frac{1}{2}\right)^2 + \left(\frac{1}{2}\right)^2 = \frac{1}{2} < 1.$$

Finalement, en notant $\rho = \alpha - \beta\gamma \in \mathbb{Z}[i]$ il vient,

$$\begin{cases} \alpha = \beta\gamma + \rho, \\ N(\rho) < N(\gamma), \end{cases}$$

ce qui prouve finalement que $\mathbb{Z}[i]$ est euclidien pour le stathme N. □

Lemme 3. Soit $\alpha \in \mathbb{Z}[i]$. Alors, α est inversible dans $\mathbb{Z}[i]$ si et seulement si $N(\alpha) = 1$. En particulier, $\mathbb{Z}[i]^* = \{\pm 1, \pm i\}$.

Preuve. Si α est inversible alors il existe $\beta \in \mathbb{Z}[i]$ tel que $\alpha\beta = 1$ et donc

$$1 = N(1) = N(\alpha\beta) = N(\alpha)N(\beta)$$

dans $\mathbb{N}$, et donc $N(\alpha) = 1$. Réciproquement, si $N(\alpha) = 1$ alors $\alpha\overline{\alpha} = N(\alpha) = 1$ avec $\overline{\alpha} \in \mathbb{Z}[i]$ (car $\Re(\overline{\alpha}) = \Re(\alpha) \in \mathbb{Z}$ et $\Im(\overline{\alpha}) = -\Im(\alpha) \in \mathbb{Z}$), ce qui prouve que α est inversible. Or, $\mathbb{U} \cap \mathbb{Z}[i] = \{\pm 1, \pm i\}$. D'où, $\mathbb{Z}[i]^* = \{\pm 1, \pm i\}$. □

Théorème 4. Soit p un nombre premier tel que $p \geq 3$. Alors, il existe $x, y \in \mathbb{N}$ tels que $p = x^2 + y^2$ si et seulement si $p \equiv 1$ [4].

Preuve. Supposons que p soit congru à 1 modulo 4.

Étape n°1. Montrons tout d'abord que -1 est un carré modulo p. D'après le théorème de Wilson,

$$(p-1)! \equiv -1 \ [p].$$

Or, puisque p est impair,

$$(p-1)! = \left(\prod_{k=1}^{\frac{p-1}{2}} k\right)\left(\prod_{k=1}^{\frac{p-1}{2}} (p-k)\right) \equiv \left(\prod_{k=1}^{\frac{p-1}{2}} k\right)\left(\prod_{k=1}^{\frac{p-1}{2}} (-k)\right) \ [p]$$

et donc en notant $N = (\frac{p-1}{2})!$, il vient $(p-1)! \equiv (-1)^{\frac{p-1}{2}} N^2$ $[p]$. Et puisque $p \equiv 1$ [4], ceci permet de conclure que $(p-1)! \equiv N^2$ $[p]$, et donc que -1 est un carré modulo p. Autrement dit, p divise $N^2 + 1 = (N - i)(N + i)$ dans $\mathbb{Z}[i]$.

Étape n°2. Montrons que p n'est pas irréductible dans $\mathbb{Z}[i]$. Puisque $\mathbb{Z}[i]$ est euclidien et que p divise $(N - i)(N + i)$ dans $\mathbb{Z}[i]$, le lemme d'Euclide est vrai (c'est le cas dans tout anneau factoriel). Si p était irréductible dans $\mathbb{Z}[i]$, le lemme d'Euclide donnerait donc p divise $N - i$ ou p divise $N + i$ dans $\mathbb{Z}[i]$, ce qui est impossible.

Étape n°3. Concluons. Puisque,

- $p \neq 0$,
- $p \notin \mathbb{Z}[i]^*$,
- et que p n'est pas irréductible dans $\mathbb{Z}[i]$,

on en déduit que p est réductible dans $\mathbb{Z}[i]$, c'est-à-dire qu'il existe $\alpha, \beta \in \mathbb{Z}[i]$ tout deux non inversibles tels que $p = \alpha\beta$. En particulier,

$$p^2 = N(p) = N(\alpha)N(\beta)$$

ce qui implique que $N(\alpha)$ divise p^2, et donc $N(\alpha) \in \{1, p, p^2\}$. Or, $N(\alpha) \neq 1$ et $N(\beta) \neq 1$ car $\alpha, \beta \notin \mathbb{Z}[i]^*$, ce qui entraîne $N(\alpha) \notin \{1, p^2\}$. D'où, $N(\alpha) = p$. Finalement, en notant $x = |\Re(\alpha)| \in \mathbb{N}$ et $y = |\Im(\alpha)| \in \mathbb{N}$, il vient donc,

$$p = N(\alpha) = x^2 + y^2.$$

Réciproquement, supposons qu'il existe $x, y \in \mathbb{N}$ tels que $p = x^2 + y^2$. Montrons que x et y sont de parité différente.

- Si $x \equiv 0$ [2] et $y \equiv 0$ [2] alors $p \equiv 0$ [2], ce qui contredit $p \geq 3$.
- Si $x \equiv 1$ [2] et $y \equiv 1$ [2] alors $p \equiv 2 \equiv 0$ [2], ce qui contredit $p \geq 3$.

Ainsi, x et y sont de parité différente. D'où l'existence de $k, q \in \mathbb{N}$ tels que,

$$p = (2k)^2 + (2q+1)^2 = 4(k^2 + q^2) + 4q + 1 \equiv 1 \ [4]$$

ce qui prouve finalement que $p \equiv 1$ [4]. □

Remarque. Dans la preuve ci-dessus, on a montré que si $p \equiv 1$ [4] alors -1 est un carré modulo p. Il s'agit en fait d'une équivalence : -1 est un carré modulo p si et seulement si $p \equiv 1$ [4].

Figure. 7.

Complément conseillé.

- Théorème des deux carrés généralisé (pour un entier naturel quelconque). On peut trouver une preuve dans [JR] Algèbre.

Commentaires personnels. Je ne vais pas passer par quatre chemins : c'est mon deuxième développement préféré. Il est technique, mais se comprend et se retient bien. De plus, il possède l'immense avantage de mêler arithmétique, algèbre et géométrie, et vous offre l'opportunité de dessiner, ce qui sera apprécié par le jury.

Théorème des restes chinois

Soit $\mathbb{A}$ un anneau principal. Soient $n \geq 2$ et $a_1, \dots, a_n$ des éléments non nuls et non inversibles de $\mathbb{A}$, supposés deux à deux premiers entre eux. Notons,

- $a = a_1 \dots a_n$,
- pour tout $j \in [\![1, n]\!]$, $b_j = \dfrac{a}{a_j} = \displaystyle\prod_{\substack{i=1 \\ i \neq j}}^{n} a_i$,
- pour tout $k \in [\![1, n]\!]$, π_k la surjection canonique de $\mathbb{A}$ sur $\mathbb{A}/(a_k)$,
- π la surjection canonique de $\mathbb{A}$ sur $\mathbb{A}/(a)$.

[JR] **Lemme 1.** Sous ces hypothèses, $b_1, \dots, b_n$ sont premiers entre eux dans leur ensemble, c'est-à-dire que tout diviseur commun à $b_1, \dots, b_n$ est inversible.

Preuve. On raisonne par contradiction en supposant qu'il existe un diviseur non inversible d commun à $b_1, \dots, b_n$. Notons que d est non nul, car b_1 est non nul. Sachant que $\mathbb{A}$ est factoriel (car principal), d admet alors un diviseur irréductible p, qui est donc un diviseur commun à $b_1, \dots, b_n$. En particulier, p divise b_1, et le lemme d'Euclide assure alors l'existence de $i \in [\![2, n]\!]$ tel que p divise a_i. Or, p divise aussi b_i et donc le lemme d'Euclide assure l'existence de $j \in [\![1, n]\!]$ distinct de i tel que p divise a_j. Ainsi, p est un diviseur irréductible commun à a_i et a_j (et $i \neq j$), ce qui contredit le fait que a_i et a_j soient premiers entre eux. Ainsi, $b_1, \dots, b_n$ sont premiers entre eux dans leur ensemble. □

[JR] **Théorème 2 (des restes chinois).** L'application,

$$\begin{array}{rccc} \Phi & : & \mathbb{A}/(a) & \longrightarrow & \mathbb{A}/(a_1) \times \cdots \times \mathbb{A}/(a_n) \\ & & \pi(x) & \longmapsto & (\pi_1(x), \dots, \pi_n(x)) \end{array}$$

est un isomorphisme d'anneaux

Preuve. On procède en plusieurs étapes.

Étape n°1. Tout d'abord, l'application,

$$\begin{array}{rccc} \varphi & : & \mathbb{A} & \longrightarrow & \mathbb{A}/(a_1) \times \cdots \times \mathbb{A}/(a_n) \\ & & x & \longmapsto & (\pi_1(x), \dots, \pi_n(x)) \end{array}$$

est un morphisme d'anneaux, car les surjections canoniques $\pi_1, \dots, \pi_n$ sont des morphismes d'anneaux (par définition des opérations sur le quotient).

Étape n°2. Montrons que $\mathrm{Ker}(\varphi) = (a)$. Tout d'abord, $(a) \subset \mathrm{Ker}(\varphi)$ puisque tout multiple de a est un multiple de chaque a_i. Réciproquement, considérons $x \in \mathrm{Ker}(\varphi)$. Alors, pour tout $k \in [\![1, n]\!]$, a_k divise x. Autrement dit, x est un multiple commun à $a_1, \dots, a_n$. Par suite, $\mathrm{ppcm}(a_1, \dots, a_n)$ divise x. Or, $\mathrm{ppcm}(a_1, \dots, a_n) = a$, car $a_1, \dots, a_n$ sont deux à deux premiers entre eux. Donc a divise x, c'est-à-dire $x \in (a)$. D'où, $\mathrm{Ker}(\varphi) = (a)$.

Étape n°3. Montrons que φ est surjective. D'après le théorème de Bézout, il existe $u_1, \dots, u_n \in \mathbb{A}$ tels que,

$$b_1 u_1 + \cdots + b_n u_n = 1.$$

car $b_1, \dots, b_n$ sont premiers entre eux dans leur ensemble (d'après le lemme précédent). En particulier, pour tout $i \in [\![1, n]\!]$, $b_i u_i \equiv 1\ [a_i]$, c'est-à-dire,

$$\pi_i(b_i u_i) = 1$$

(car, pour tout $j \in [\![1, n]\!]$ distinct de i, a_i divise b_j). Ainsi,

$$\forall (i, j) \in [\![1, n]\!]^2,\ \pi_i(b_j u_j) = \begin{cases} 1 & \text{si } i = j, \\ 0 & \text{sinon.} \end{cases}$$

Soient $y_1, \dots, y_n \in \mathbb{A}$. Notons,

$$x = \sum_{j=1}^{n} b_j u_j y_j.$$

Alors, pour tout $i \in [\![1, n]\!]$,

$$\pi_i(x) = \sum_{j=1}^{n} \pi_i(b_j u_j) \pi_i(y_j) = \pi_i(y_i)$$

(car π_i est un morphisme d'anneaux). D'où,

$$\varphi(x) = (\pi_1(y_1), \dots, \pi_n(y_n))$$

Finalement,

$$\mathbb{A}/(a_1) \times \cdots \times \mathbb{A}/(a_n) = \pi_1(\mathbb{A}) \times \cdots \times \pi_n(\mathbb{A}) = \mathrm{Im}(\varphi)$$

ce qui prouve que φ est surjective.

Étape n°4. D'après le premier théorème d'isomorphisme, on en conclut

que Φ est un isomorphisme. De plus, on a même démontré que sa bijection réciproque est,

$$\begin{array}{rccc} \Phi^{-1} & : & \mathbb{A}/(a_1)\times\cdots\times\mathbb{A}/(a_n) & \longrightarrow & \mathbb{A}/(a) \\ & & (\pi_1(x),\dots,\pi_n(x)) & \longmapsto & \pi\left(\sum_{j=1}^{n} b_j u_j y_j\right) \end{array}$$

□

[JR] **Application 3.** Le système d'équations diophantiennes,

$$\begin{cases} k \equiv 2\ [4] \\ k \equiv 3\ [5] \\ k \equiv 1\ [9] \end{cases}$$

admet pour ensemble de solutions $118 + 180\mathbb{Z}$.

Preuve. Tout d'abord, puisque $\text{pgcd}(4,5,9) = 1$, le théorème des restes chinois assure que ce système admet une unique solution modulo $4 \times 5 \times 9 = 180$. De plus, dans la preuve de ce théorème, on a vu que cette solution s'obtenait en déterminant une relation de Bézout entre les entiers suivants :

$$b_1 = 5 \times 9 = 45,\ b_2 = 4 \times 9 = 36,\ b_3 = 4 \times 5 = 20.$$

Pour cela, on cherche tout d'abord une relation de Bézout entre b_2 et b_3. En effet, $\text{pgcd}(b_2, b_3) = 4$ et l'algorithme d'Euclide permet d'obtenir la relation de Bézout suivante :

$$\text{pgcd}(b_2, b_3) = 4 = 2 \times 20 - 36 = 2b_3 - b_2.$$

Ici, on peut bien sûr la trouver sans utiliser l'algorithme. Mais pour de grands entiers, l'algorithme d'Euclide peut s'avérer utile pour trouver une relation de Bézout entre b_2 et b_3 (on rappelle qu'il n'y a pas unicité lorsque les entiers ne sont pas premiers entre eux). On cherche ensuite une relation de Bézout entre $\text{pgcd}(b_2, b_3)$ et b_1. On a,

$$1 = 45 - 11 \times 4 = b_1 - 11 \times \text{pgcd}(b_2, b_3).$$

Finalement, les deux égalités précédentes fournissent,

$$1 = b_1 - 11(2b_3 - b_2) = b_1 + 11b_2 - 22b_3.$$

(par rapport aux notations de la preuve précédente, cela correspond à $u_1 = 1$, $u_2 = 11$ et $u_3 = -22$). Ainsi, l'unique solution modulo $4 \times 5 \times 9 = 180$ est,

$$2u_1b_1 + 3u_2b_2 + 1u_3b_3 = 838 \equiv 118\ [180].$$

Autrement dit, l'ensemble des solutions est $118 + 180\mathbb{Z}$. □

Commentaires personnels. Développement très sympathique, qui se comprend et se retient facilement, et peut être présenté dans plusieurs leçons. De plus, le théorème des restes chinois figure sur le programme du concours.

Théorème spectral

Soient E un espace euclidien de dimension finie, $\mathcal{B}$ une base orthonormée de E et f un endomorphisme symétrique de E. Notons A sa matrice dans la base $\mathcal{B}$ et $n = \dim(E)$.

[JG] **Lemme 1.** L'endomorphisme f admet au moins une valeur propre réelle.

Preuve. Soit λ une valeur propre complexe de A, c'est-à-dire qu'il existe une matrice colonne $X \in M_{n,1}(\mathbb{C})$ non nulle telle que $AX = \lambda X$. Alors,

$$(\lambda - \overline{\lambda})\,{}^t\overline{X}X = {}^t\overline{X}(\lambda X) - {}^t(\overline{\lambda X})X = {}^t\overline{X}AX - {}^t\overline{X}\,{}^t\overline{A}X = 0$$

car A est symétrique à coefficients réels. Puisque X est non nul, ceci implique $\lambda = \overline{\lambda}$, c'est-à-dire $\lambda \in \mathbb{R}$. Ainsi, λ est une valeur propre réelle de A, donc une valeur propre réelle de f (car λ est racine réelle de χ_A qui est égal à χ_f). □

[JG] **Lemme 2.** Soit F un sous-espace vectoriel de E stable par f. Alors, $F^\perp$ est stable par f.

Preuve. Soit $x \in F^\perp$. Soit $y \in F$. Alors, $f(y) \in F$ car F est stable par f, puis $(f(x) \mid y) = (x \mid f(y)) = 0$ (la première égalité provient du fait que f est symétrique, et la seconde de $x \in F^\perp$ et $f(y) \in F$). D'où, $f(x) \in F^\perp$. Ainsi, $F^\perp$ est stable par f. □

[JG] **Théorème 3 (Théorème spectral).** Il existe une base orthonormée de E formée de vecteurs propres de f.

Preuve. Montrons par récurrence sur $n \in \mathbb{N}^*$ que, pour tout espace euclidien E de dimension n, et pour tout endomorphisme symétrique f de E, il existe une base orthonormée formée de vecteurs propres de f.

Initialisation. Soit E un espace euclidien de dimension 1. Alors, tout endomorphisme f de E est une homothétie, et donc n'importe quelle base orthonormée de E est formée de vecteurs propres de f (car tout vecteur propre non nul est vecteur propre de f).

Hérédité. Supposons ensuite le résultat établi pour un rang $n \in \mathbb{N}^*$ fixé. Soient E un espace euclidien de dimension $n+1$ et f un endomorphisme symétrique de E. D'après le lemme 1, f admet au moins une valeur propre réelle λ. Soit e_1 un vecteur propre de f associé à λ. Notons $F = \mathbb{R}e_1$. Alors, F est stable par f, et donc $F^\perp$ est stable par f en conséquence du lemme 2. Notons g l'endomorphisme induit par f sur $F^\perp$. Il s'agit d'un endomorphisme symétrique de $F^\perp$. De plus,

$$\dim(F^\perp) = \dim(E) - \dim(F) = n$$

Par hypothèse de récurrence, il existe donc $e_2, \ldots, e_{n+1}$ des vecteurs propres de g (qui sont a fortiori des vecteurs propres de f) tels que $(e_2, \ldots, e_{n+1})$ soit une base orthonormée de $F^\perp$. Puisque $F^\perp \oplus F = E$, il s'en suit que $(e_1, \ldots, e_{n+1})$ est une base orthonormée de E, formée de vecteurs propres de f, ce qui achève l'hérédité. □

Corollaire 4. Soit $A \in S_n(\mathbb{R})$. Il existe $P \in O_n(\mathbb{R})$ et $D \in M_n(\mathbb{R})$ diagonale telles que $A = PD\,{}^tP$.

Preuve. Considérons l'endomorphisme de $M_{n,1}(\mathbb{R})$ suivant,

$$\begin{array}{rccc} f & : & M_{n,1}(\mathbb{R}) & \longrightarrow & M_{n,1}(\mathbb{R}) \\ & & X & \longmapsto & AX \end{array}$$

Il s'agit de l'endomorphisme de $M_{n,1}(\mathbb{R})$ canoniquement associé à la matrice A. On munit $M_{n,1}(\mathbb{R})$ de son produit scalaire canonique, c'est-à-dire,

$$\forall X, Y \in M_{n,1}(\mathbb{R}),\ (X \mid Y) = {}^tXY.$$

Alors, f est un endomorphisme symétrique de $(M_{n,1}(\mathbb{R}), (\cdot \mid \cdot))$. En effet,

$$\forall X, Y \in M_{n,1}(\mathbb{R}),\ (f(X) \mid Y) = {}^t(AX)Y = {}^tXAY = (X \mid f(Y))$$

car $A \in S_n(\mathbb{R})$. D'après le théorème spectral, il existe donc $X_1, \ldots, X_n$ des vecteurs propres de f, respectivement associés à des valeurs propres $\lambda_1, \ldots, \lambda_n$ de f, tels que $\mathcal{B} := (X_1, \ldots, X_n)$ soit une base orthonormée de $M_{n,1}(\mathbb{R})$. En notant P la matrice de passage de la base canonique de $M_{n,1}(\mathbb{R})$ vers la base $\mathcal{B}$ (qui appartient à $O_n(\mathbb{R})$ car ces deux bases sont orthonormées), il vient,

$$A = P\,\mathrm{Diag}(\lambda_1, \ldots, \lambda_n)P^{-1} = P\,\mathrm{Diag}(\lambda_1, \ldots, \lambda_n)\,{}^tP.$$

□

Théorème 5 (Orthogonalisation simultanée). Soient q_1 et q_2 deux formes quadratiques sur E. Supposons que q_1 soit définie positive. Alors, il existe une base de E qui soit q_1-orthonormée et q_2-orthogonale.

Preuve. Notons φ_1 la forme polaire de q_1. Alors, φ_1 est un produit scalaire sur E (car q_1 est définie positive) et donc (E, φ_1) est un espace euclidien. Soit $\mathcal{B}$ une base orthonormée de (E, φ_1). Notons S_2 la matrice de q_2 dans cette base. Puisque $S_2 \in S_n(\mathbb{R})$, le corollaire précédent assure l'existence de $P \in O_n(\mathbb{R})$ et $D \in M_n(\mathbb{R})$ diagonale telles que,

$$S_2 = PD\,{}^tP.$$

Autrement dit, D est la matrice de q_2 dans une base $\mathcal{B}'$ de E ($\mathcal{B}'$ est la base formée des vecteurs dont les coordonnées dans la base $\mathcal{B}$ sont les coefficients des colonnes de P). Or, $\mathcal{B}$ étant une base orthonormée de (E, φ_1), il vient,

$$\mathrm{Mat}_{\mathcal{B}'}(q_1) = {}^tP\,\mathrm{Mat}_{\mathcal{B}}(q_1)P = {}^tPI_nP = I_n$$

Ainsi, $\mathcal{B}'$ est une base de E qui est q_1-orthonormée et q_2-orthogonale. □

Corollaire 6. Soient $S_1 \in S_n^{++}(\mathbb{R})$ et $S_2 \in S_n(\mathbb{R})$. Il existe $P \in GL_n(\mathbb{R})$ et $D \in M_n(\mathbb{R})$ diagonale telles que ${}^tPS_1P = I_n$ et ${}^tPS_2P = D$.

Preuve. Considérons les applications suivantes,

$$\begin{array}{rccc} q_1 & : & M_{n,1}(\mathbb{R}) & \longrightarrow & M_{n,1}(\mathbb{R}) \\ & & X & \longmapsto & {}^tXS_1X \end{array}$$

et,

$$\begin{array}{rccc} q_2 & : & M_{n,1}(\mathbb{R}) & \longrightarrow & M_{n,1}(\mathbb{R}) \\ & & X & \longmapsto & {}^tXS_2X \end{array}$$

Alors, q_1 et q_2 sont des formes quadratiques sur $M_{n,1}(\mathbb{R})$. De plus, en notant $\mathcal{C}$ la base canonique de $M_{n,1}(\mathbb{R})$, il vient,

$$\mathrm{Mat}_{\mathcal{C}}(q_1) = S_1,\ \mathrm{Mat}_{\mathcal{C}}(q_2) = S_2.$$

Or, q_1 est définie positive, car S_1 l'est. D'après le théorème précédent, il existe donc une base $\mathcal{B}$ de E qui soit q_1-orthonormée et q_2-orthogonale, c'est-à-dire qu'il existe $D \in M_n(\mathbb{R})$ diagonale telle que,

$$\mathrm{Mat}_{\mathcal{B}}(q_1) = I_n,\ \mathrm{Mat}_{\mathcal{B}}(q_2) = D.$$

En notant P la matrice de passage de $\mathcal{C}$ vers $\mathcal{B}$, les formules de changement de base pour les formes quadratiques fournissent alors,

$$\begin{cases} {}^tPS_1P^t = {}^tP\,\mathrm{Mat}_{\mathcal{C}}(q_1)P = \mathrm{Mat}_{\mathcal{B}}(q_1) = I_n, \\ {}^tPS_2P^t = {}^tP\,\mathrm{Mat}_{\mathcal{C}}(q_2)P = \mathrm{Mat}_{\mathcal{B}}(q_2) = D. \end{cases}$$

Ainsi, ${}^tPS_1P = I_n$ et ${}^tPS_2P = D$. □

Remarque. La matrice P n'est pas nécessairement orthogonale (car $\mathcal{C}$ n'est pas nécessairement q_1-orthonormée).

Commentaires personnels. Ce développement contient trop d'items. Cette preuve du théorème spectral étant trop courte, elle ne peut pas faire seule l'objet d'un développement. Il est inutile de démontrer le corollaire 4, la preuve ne présente pas de difficulté particulière. Pour être tout à fait honnête, je ne l'ai pas vraiment testé. Je pensais présenter les lemmes 1 et 2, les théorèmes 3 et 5, et le corollaire 6. À noter qu'il peut se présenter dans de nombreuses leçons. De plus, le théorème spectral figure sur le programme du concours.

Une suite de polygones

[XG] **Lemme 1.** Soient $a_1, \dots, a_n \in \mathbb{C}$. Notons $L = \sum_{k=0}^{n-1} a_{k+1} X^k$ et,

$$A = \begin{pmatrix} a_1 & a_2 & \dots & a_n \\ a_n & a_1 & \dots & a_{n-1} \\ \vdots & \vdots & & \vdots \\ a_2 & a_3 & \dots & a_1 \end{pmatrix}.$$

Alors, $\chi_A = \prod_{\lambda \in \mathbb{U}_n} (X - L(\lambda))$.

Preuve. Notons $(e_1, \dots, e_n)$ la base canonique de $\mathbb{R}^n$. Nous allons exprimer A comme un polynôme en la matrice,

$$J = \begin{pmatrix} 0 & 1 & 0 & \dots & \dots & 0 \\ \vdots & 0 & \ddots & \ddots & & \vdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \vdots & & \ddots & \ddots & 0 \\ 0 & 0 & \cdots & \cdots & 0 & 1 \\ 1 & 0 & \cdots & \cdots & \cdots & 0 \end{pmatrix} = \begin{pmatrix} 0 & I_{n-1} \\ 1 & 0 \end{pmatrix} \in M_n(\mathbb{R}).$$

Soit f l'endomorphisme de $\mathbb{R}^n$ canoniquement associé à la matrice J. Alors,

$$\begin{cases} f(e_1) = e_n, \\ \forall k \in [\![2, n]\!],\ f(e_k) = e_{k-1}. \end{cases}$$

Montrons, par récurrence finie sur $p \in [\![1, n-1]\!]$ que,

$$\begin{cases} \forall k \in [\![1, p]\!],\ f^p(e_k) = e_{n-p+k}, \\ \forall k \in [\![p+1, n]\!],\ f^p(e_k) = e_{k-p}. \end{cases}$$

L'initialisation a déjà été prouvée. Supposons le résultat établi pour un rang $p \in [\![1, n-2]\!]$ (on suppose $n \geq 3$, sinon l'initialisation suffit). Soit $k \in [\![1, n]\!]$. On distingue deux cas.

- **Cas n°1.** Supposons $k \in [\![1, p+1]\!]$. Alors,

$$f^{p+1}(e_k) = f^p(f(e_k)) = \begin{cases} f^p(e_n) & \text{si } k = 1, \\ f^p(e_{k-1}) & \text{sinon.} \end{cases}$$

Puis, par hypothèse de récurrence,

$$f^{p+1}(e_k) = \begin{cases} e_{n-p} & \text{si } k = 1, \\ e_{n-p+k-1} & \text{sinon,} \end{cases}$$

et donc $f^{p+1}(e_k) = e_{n-(p+1)+k}$.

- **Cas n°2.** Supposons $k \in [\![p+2, n]\!]$. Alors,

$$f^{p+1}(e_k) = f^p(f(e_k)) = f^p(e_{k-1}) = e_{k-p-1}$$

et donc $f^{p+1}(e_k) = e_{k-(p+1)}$.

Par récurrence finie, il s'en suit que,

$$\forall p \in [\![1, n-1]\!],\ \begin{cases} \forall k \in [\![1, p]\!],\ f^p(e_k) = e_{n-p+k}, \\ \forall k \in [\![p+1, n]\!],\ f^p(e_k) = e_{k-p}. \end{cases}$$

D'où,

$$\forall p \in [\![0, n-1]\!],\ J^p = \begin{pmatrix} 0 & I_{n-p} \\ I_p & 0 \end{pmatrix}.$$

Par suite, $A = L(J)$. Calculons ensuite le polynôme caractéristique de J. Soit $\lambda \in \mathbb{C}$. Alors,

$$\chi_J(\lambda) = \begin{vmatrix} \lambda & -1 & 0 & \dots & \dots & 0 \\ \vdots & \lambda & \ddots & \ddots & & \vdots \\ \vdots & 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & 0 & \cdots & 0 & \lambda & -1 \\ -1 & 0 & \cdots & \cdots & \cdots & \lambda \end{vmatrix}.$$

Puis, en développant ce déterminant par rapport à la première colonne,

$$\chi_J(\lambda) = \lambda \begin{vmatrix} \lambda & -1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \lambda & -1 \\ 0 & \cdots & \cdots & 0 & \lambda \end{vmatrix} + (-1)^{n+2} \begin{vmatrix} -1 & 0 & \cdots & \cdots & 0 \\ \lambda & \ddots & \ddots & & \vdots \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda & -1 \end{vmatrix}.$$

Et donc, $\chi_J(\lambda) = \lambda\lambda^{n-1} + (-1)^{n+2}(-1)^{n-1} = \lambda^n - 1$. Ainsi, $\chi_J = X^n - 1$. En particulier, J est diagonalisable sur $\mathbb{C}$ (car son polynôme caractéristique est

scindé à racines simples sur $\mathbb{C}$) et ses valeurs propres sont les racines n-ièmes de l'unité. D'où l'existence de $P \in GL_n(\mathbb{C})$ telle que $A = PDP^{-1}$ avec,

$$D = \mathrm{Diag}\left(1, e^{\frac{2i\pi}{n}}, \dots, e^{\frac{2i(n-1)\pi}{n}}\right).$$

Par récurrence immédiate, il vient alors,

$$\forall k \in \mathbb{N},\ J^k = PD^kP^{-1}.$$

Et donc,

$$A = L(J) = PL(D)P^{-1} = P\,\mathrm{Diag}\left(L(1), L\left(e^{\frac{2i\pi}{n}}\right), \dots, L\left(e^{\frac{2i(n-1)\pi}{n}}\right)\right)P^{-1}.$$

En particulier,

$$\chi_A = \prod_{k=0}^{n-1}\left(X - L\left(e^{\frac{2ik\pi}{n}}\right)\right) = \prod_{\lambda\in\mathbb{U}_n}(X - L(\lambda)).$$

□

[IP] **Proposition 2.** Soit $(P^{(k)})_{k\in\mathbb{N}} = ((z_1^{(k)}, \dots, z_n^{(k)}))_{k\in\mathbb{N}}$ une suite d'éléments de $\mathbb{C}^n$ vérifiant la relation de récurrence, pour tout $k \in \mathbb{N}$,

$$P^{(k+1)} = \left(\frac{z_1^{(k)} + z_2^{(k)}}{2}, \dots, \frac{z_{n-1}^{(k)} + z_n^{(k)}}{2}, \frac{z_n^{(k)} + z_1^{(k)}}{2}\right).$$

Alors, $(P^{(k)})_{k\in\mathbb{N}}$ converge vers le n-uplet $(g, \dots, g)$, où g désigne l'isobarycentre des points $z_1^{(0)}, \dots, z_n^{(0)}$.

Preuve. On procède en plusieurs étapes.

Étape n°1. Montrons que la suite $(P^{(k)})_{k\in\mathbb{N}}$ converge. Notons $M = \frac{1}{2}J + \frac{1}{2}I$. Alors,

$$\forall k \in \mathbb{N},\ P^{(k+1)} = MP^{(k)}.$$

Par récurrence immédiate, on en déduit que,

$$\forall k \in \mathbb{N},\ P^{(k)} = M^kP^{(0)}.$$

D'autre part, d'après le lemme précédent,

$$\chi_M = \prod_{\lambda\in\mathbb{U}_n}\left[X - \frac{1}{2}(1+\lambda)\right]$$

et donc M est diagonalisable sur $\mathbb{C}$ (car son polynôme caractéristique est scindé à racines simples sur $\mathbb{C}$) avec $\mathrm{Sp}_{\mathbb{C}}(M) = \{\frac{1}{2}(1+\lambda),\ \lambda \in \mathbb{U}_n\}$. D'où l'existence de $Q \in GL_n(\mathbb{C})$ telle que $M = QDQ^{-1}$, où l'on a noté, pour tout $l \in [\![2, n]\!]$,

$$\lambda_l = \frac{1}{2}\left(1 + e^{\frac{2il\pi}{n}}\right) \text{ et } D = \mathrm{Diag}(1, \lambda_2, \dots, \lambda_n).$$

Par récurrence immédiate, il vient alors,

$$\forall k \in \mathbb{N},\ M^k = QD^kQ^{-1}.$$

Or, pour tout $z = e^{i\theta} \in \mathbb{U}\setminus\{1\}$, on a,

$$\left|1 + e^{i\theta}\right| = \left|e^{i\frac{\theta}{2}}\left(e^{-i\frac{\theta}{2}} + e^{i\frac{\theta}{2}}\right)\right| = \left|e^{i\frac{\theta}{2}}\right|\left|2\cos\left(\frac{\theta}{2}\right)\right| = 2\left|\cos\left(\frac{\theta}{2}\right)\right| < 2$$

car $\theta \not\equiv 0\ [2\pi]$ (voir figure 9). Et donc,

$$D^k = \mathrm{Diag}\left(1, \lambda_2^k, \dots, \lambda_n^k\right) \underset{k\to+\infty}{\longrightarrow} \mathrm{Diag}(1, 0, \dots, 0)$$

car, pour tout $l \in [\![2, n]\!]$, $|\lambda_l| < 1$. Par continuité du produit matriciel, il s'en suit que,

$$P^{(k)} = QD^kQ^{-1}P^{(0)} \underset{k\to+\infty}{\longrightarrow} P^{(\infty)}.$$

où l'on a noté $P^{(\infty)} = Q\,\mathrm{Diag}(1, 0, \dots, 0)Q^{-1}P^{(0)}$.

Étape n°2. Montrons que $P^{(\infty)} \in \mathrm{Vect}((1, \dots, 1))$. Notons $z_1^{(\infty)}, \dots, z_n^{(\infty)}$ les coordonnées du vecteur $P^{(\infty)}$ dans la base canonique de $\mathbb{C}^n$. De nouveau, par continuité du produit matriciel, il vient,

$$P^{(k+1)} = MP^{(k)} \underset{k\to+\infty}{\longrightarrow} MP^{(\infty)}.$$

D'où, $MP^{(\infty)} = P^{(\infty)}$ par unicité de la limite, c'est-à-dire,

$$\begin{cases} \dfrac{z_1^{(\infty)} + z_2^{(\infty)}}{2} &= z_1^{(\infty)} \\ \vdots & \\ \dfrac{z_{n-1}^{(\infty)} + z_n^{(\infty)}}{2} &= z_{n-1}^{(\infty)} \\ \dfrac{z_n^{(\infty)} + z_1^{(\infty)}}{2} &= z_n^{(\infty)} \end{cases}$$

ce qui implique, $z_1^{(\infty)} = \dots = z_{n-1}^{(\infty)} = z_n^{(\infty)}$. Donc,

$$P^{(\infty)} = \left(z_1^{(\infty)}, \dots, z_1^{(\infty)}\right) \in \mathrm{Vect}((1, \dots, 1)).$$

On aurait aussi pu remarquer que $P^{(\infty)} \in E_1(M)$ et $E_1(M) = \mathrm{Vect}((1, \dots, 1))$. En effet, $\dim(E_1(M)) = 1$ (puisque 1 est racine simple de χ_M), et du fait que $(1, \dots, 1) \in E_1(M)$, on en déduit que $E_1(M) = \mathrm{Vect}((1, \dots, 1))$.

Étape n°3. Montrons que $P^{(\infty)} = (g, \dots, g)$. Considérons l'application,

$$\begin{array}{rccc} \text{Isobar} & : & \mathbb{C}^n & \longrightarrow & \mathbb{C} \\ & & (a_1, \dots, a_n) & \longmapsto & \dfrac{a_1 + \cdots + a_n}{n} \end{array}$$

qui est continue, car elle est polynomiale (c'est la fonction polynomiale associée au polynôme $\frac{1}{n}(X_1 + \cdots + X_n) \in \mathbb{C}[X_1, \dots, X_n]$). Il s'agit de application qui à n points du plan complexe associe leur isobarycentre. Soit $k \in \mathbb{N}$. Alors,

$$\text{Isobar}(P^{(k+1)}) = \frac{1}{n}\left[z_1^{(k+1)} + \cdots + z_{n-1}^{(k+1)} + z_n^{(k+1)}\right].$$

Puis,

$$\text{Isobar}(P^{(k+1)}) = \frac{1}{n}\left[\frac{z_1^{(k)} + z_2^{(k)}}{2} + \cdots + \frac{z_{n-1}^{(k)} + z_n^{(k)}}{2} + \frac{z_n^{(k)} + z_1^{(k)}}{2}\right].$$

Dans la somme ci-dessus, chacun des termes $z_l^{(k)}$ (pour $1 \leq l \leq n$) apparaît exactement deux fois. D'où,

$$\text{Isobar}(P^{(k+1)}) = \frac{1}{n}\left[\frac{2z_1^{(k)} + \cdots + 2z_n^{(k)}}{2}\right].$$

Et donc,

$$\text{Isobar}(P^{(k+1)}) = \frac{1}{n}\left[z_1^{(k)} + \cdots + z_{n-1}^{(k)} + z_n^{(k)}\right] = \text{Isobar}(P^{(k)}).$$

Ainsi, la suite $(\text{Isobar}(P^{(k)}))_{k\in\mathbb{N}}$ est constante, avec,

$$\forall k \in \mathbb{N},\ \text{Isobar}(P^{(k)}) = \text{Isobar}(P^{(0)}) = g.$$

Or, par continuité de l'application Isobar,

$$\text{Isobar}(P^{(k)}) \underset{k\to+\infty}{\longrightarrow} \text{Isobar}(P^{(\infty)}).$$

D'où, par unicité de la limite,

$$g = \text{Isobar}(P^{(\infty)}) = \text{Isobar}\left(z_1^{(\infty)}, \dots, z_1^{(\infty)}\right) = z_1^{(\infty)}$$

et donc $P^{(\infty)} = (g, \dots, g)$. □

Figure. 8, 9.

Commentaires personnels. Développement long, mais qui se comprend et se retient facilement, et qui se recase dans de très nombreuses leçons. De plus, ce développement possède l'immense avantage de mêler algèbre et géométrie, et vous offre l'opportunité de dessiner, ce qui sera apprécié par le jury.

Références

[BM] : Objectif agrégation de V. Beck, J. Malick et G. Peyré (H & K).

[DP] : Cours d'algèbre de D. Perrin (ellipses).

[FG] : Exercices de mathématiques oraux x-ens, algèbre 1 de S. Francinou, H. Gianella et S. Nicolas (Cassini).

[FG] : Exercices de mathématiques oraux x-ens, algèbre 2 de S. Francinou, H. Gianella et S. Nicolas (Cassini).

[FG] : Exercices de mathématiques oraux x-ens, algèbre 3 de S. Francinou, H. Gianella et S. Nicolas (Cassini).

[IG] : Théorie de Galois de I. Gozard (ellipses).

[IP] : L'oral à l'agrégation de mathématiques - Une sélection de développements de L. Isenmann et T. Pecatte.

[JG] : Algèbre linéaire de J. Grifone (Cépaduès).

[JR] : Mathématiques pour l'agrégation, algèbre et géométrie de J. Rombaldi (De Boeck Supérieur).

[JR] : Mathématiques pour l'agrégation, analyse et probabilités de J. Rombaldi (De Boeck Supérieur).

[LM] : 131 développements pour l'oral de D. Lesesvre, P. Montagnon, P. Le Barbenchon et T. Pierron (Dunod).

[XG] : Algèbre et probabilités de X. Gourdon (ellipses).

[XG] : Analyse de X. Gourdon (ellipses).

Références détaillées

Développement ou résultat utile	Références
Automorphismes de $\mathfrak{S}_n$	[DP] : Chapitre I, proposition 8.8, lemme 8.9, théorème 8.7.
Construction des corps finis	[IG] : Chapitre VII, théorèmes VII.17 et VII.24, corollaire VII.25.
Polynômes réductibles	[XG] : Algèbre, chapitre 2, section 1, exercice 4, question 2.
Critère d'Eisenstein	[JR] : Algèbre, chapitre 12, lemme 12.10, théorème 12.33.
Cyclicité de $(\mathbb{Z}/p^\alpha\mathbb{Z})^\star$	[JR] : Algèbre, chapitre 1, théorème 1.11, chapitre 10, théorème 10.15.
Décomposition de Dunford	[JR] : Algèbre, chapitre 19, théorèmes 19.9, 19.10 et 19.19, corollaire 19.7.
Décomposition polaire	[JR] : Algèbre, chapitre 22, théorèmes 22.30, 22.31 et 22.33.
Dénombrement des polynômes irréductibles unitaires sur un corps fini	[JR] : Algèbre, chapitre 13, lemme 13.7, théorèmes 13.6 et 13.7, corollaire 13.1.
Déterminant de Gram	[XG] : Algèbre, chapitre 5, section 3, théorème 8.
Enveloppe convexe des matrices orthogonales	[IP] : Développement 26.
Générateurs de $GL_n(\mathbb{K})$ et $SL_n(\mathbb{K})$	[XG] : Algèbre, chapitre 3, section 6, problème 13.
Générateurs du groupe orthogonal d'un espace vectoriel euclidien	[FG] : Algèbre 3, chapitre 1, exercice 1.33.
Loi d'inertie de Sylvester	[JG] : Chapitre 9, théorèmes 9.12 et 9.16.
Probabilité que deux éléments commutent dans un groupe	[FG] : Algèbre 1, chapitre 2, exercice 2.12.
Probabilité que deux éléments commutent dans D_8	[LM] : Développement 2.
Simplicité de $\mathfrak{A}_n$	[JR] : Algèbre, chapitre 2, théorèmes 2.10 et 2.12, exercice 2.24.
Surjectivité de l'exponentielle matricielle	[IP] : Développement 63.
Théorème de Burnside	[JR] : Algèbre, chapitre 5, lemmes 5.3 et 5.4, théorème 5.7.
Théorème de Cauchy	[JR] : Algèbre, chapitre 1, théorèmes 1.30 et 1.33, corollaire 1.5.
Théorème de d'Alembert-Gauss	[JR] : Algèbre, chapitre 12, théorème 12.30.
Théorème de Dirichlet	[FG] : Algèbre 1, chapitre 4, exercice 4.18.
Théorème de Kronecker	[JR] : Algèbre, chapitre 12, exercice 12.3.
Théorème des deux carrés de Fermat	[???] : ???
Théorème des restes chinois	[JR] : Algèbre, chapitre 8, lemme 8.9, théorème 8.13, exemple 10.2.
Théorème spectral	[JG] : Chapitre 7, théorème 7.38.
Déterminant circulant	[XG] : Algèbre, chapitre 4, section 2, exercice 4.
Une suite de polygones	[IP] : Développement 62.

Analyse et probabilités

Liste des développements d'analyse et probabilités

Numéro	Développement	leçons proposées
1	Algorithme du gradient à pas optimal	219,226,229,253
2	Base hilbertienne de polynômes orthogonaux	201,209,213,234,239,245,250
3	Équation de la chaleur	209,235,239,241,246
4	Équation de Sylvester	155,156,162,221
5	Équivalence des normes en dimension finie et théorème de Riesz	203,206,208
6	Espace de Bergman	201,205,208,213,234,245
7	Expression des $\zeta(2k)$	230,235,243,244,246
8	Formule de Stirling (par les intégrales de Wallis)	223,224,236
9	Formule de Stirling (par le théorème central limite)	223,224,235,236,261,262,264,266
10	Méthode de Laplace	218,224,228,236,239
11	Méthode de Newton	218,223,224,226,229
12	Nombres de Bell	190,241,243
13	Projection sur un convexe fermé	205,208,213,219,253
14	Prolongement de la fonction Gamma d'Euler	235,236,239,241,244,245
15	Série des inverses des nombres premiers	121,230,244,264,266
16	Système de Lotka-Volterra	220,229
17	Théorème central limite	218,250,261,262,266
18	Théorème d'Ascoli	201,203,205,228
19	Théorème d'Hadamard-Lévy	204,214,215
20	Théorème d'inversion locale	214,215
21	Théorème de Cauchy-Lipschitz	205,220
22	Théorème de Cauchy-Lipschitz linéaire	205,206,221
23	Théorème de réarrangement de Riemann	230
24	Théorème des extrema liés	159,206,214,215,219

Liste des leçons d'analyse et probabilités

Numéro	Leçon	Développements
201	Espaces de fonctions. Exemples et applications.	Base hilbertienne de polynômes orthogonaux Espace de Bergman Théorème d'Ascoli
203	Utilisation de la notion de compacité.	Équivalence des normes en dimension finie et théorème de Riesz Théorème d'Ascoli
204	Connexité. Exemples d'applications.	Surjectivité de l'exponentielle matricielle Théorème d'Hadamard-Lévy
205	Espaces complets. Exemples et applications.	Espace de Bergman Projection sur un convexe fermé Théorème d'Ascoli Théorème de Cauchy-Lipschitz Théorème de Cauchy-Lipschitz linéaire
206	Exemples d'utilisation de la notion de dimension finie en analyse.	Équivalence des normes en dimension finie et théorème de Riesz Théorème de Cauchy-Lipschitz linéaire Théorème des extrema liés
208	Espaces vectoriels normés, applications linéaires continues. Exemples.	Équivalence des normes en dimension finie et théorème de Riesz Espace de Bergman Projection sur un convexe fermé
209	Approximation d'une fonction par des fonctions régulières. Exemples d'applications.	Base hilbertienne de polynômes orthogonaux Équation de la chaleur
213	Espaces de Hilbert. Exemples d'applications.	Base hilbertienne de polynômes orthogonaux Espace de Bergman Projection sur un convexe fermé
214	Théorème d'inversion locale, théorème des fonctions implicites. Illustrations en analyse et en géométrie.	Surjectivité de l'exponentielle matricielle Théorème d'Hadamard-Lévy Théorème d'inversion locale Théorème des extrema liés
215	Applications différentiables définies sur un ouvert de $\mathbb{R}^n$. Exemples et applications.	Théorème d'Hadamard-Lévy Théorème d'inversion locale Théorème des extrema liés
218	Formules de Taylor. Exemples et applications.	Méthode de Laplace Méthode de Newton Théorème central limite
219	Extremums : existence, caractérisation, recherche. Exemples et applications.	Algorithme du gradient à pas optimal Projection sur un convexe fermé Théorème des extrema liés
220	Illustrer par des exemples la théorie des équations différentielles ordinaires.	Système de Lotka-Volterra Théorème de Cauchy-Lipschitz

Numéro	Leçon	Développements
221	Équations différentielles linéaires. Systèmes d'équations différentielles linéaires. Exemples et applications.	Équation de Sylvester Théorème de Cauchy-Lipschitz linéaire
223	Suites numériques. Convergence, valeurs d'adhérence. Exemples et applications.	Formule de Stirling (par les intégrales de Wallis) Formule de Stirling (par le théorème central limite) Méthode de Newton
224	Exemples de développements asymptotiques de suites et de fonctions.	Formule de Stirling (par les intégrales de Wallis) Formule de Stirling (par le théorème central limite) Méthode de Laplace Méthode de Newton
226	Suites vectorielles et réelles définies par une relation de récurrence $u_{n+1} = f(u_n)$. Exemples. Applications à la résolution approchée d'équations.	Une suite de polygones Algorithme du gradient à pas optimal Méthode de Newton
228	Continuité, dérivabilité des fonctions réelles d'une variable réelle. Exemples et applications.	Méthode de Laplace Théorème d'Ascoli
229	Fonctions monotones. Fonctions convexes. Exemples et applications.	Algorithme du gradient à pas optimal Méthode de Newton Système de Lotka-Volterra
230	Séries de nombres réels ou complexes. Comportement des restes ou des sommes partielles des séries numériques. Exemples.	Expression des $\zeta(2k)$ Série des inverses des nombres premiers Théorème de réarrangement de Riemann
234	Fonctions et espaces de fonctions Lebesgue-intégrables.	Base hilbertienne de polynômes orthogonaux Espace de Bergman
235	Problèmes d'interversion de symboles en analyse.	Équation de la chaleur Expression des $\zeta(2k)$ Formule de Stirling (par le théorème central limite) Prolongement de la fonction Gamma d'Euler
236	Illustrer par des exemples quelques méthodes de calcul d'intégrales de fonctions d'une ou plusieurs variables.	Formule de Stirling (par les intégrales de Wallis) Formule de Stirling (par le théorème central limite) Méthode de Laplace Prolongement de la fonction Gamma d'Euler
239	Fonctions définies par une intégrale dépendant d'un paramètre. Exemples et applications.	Base hilbertienne de polynômes orthogonaux Équation de la chaleur Méthode de Laplace Prolongement de la fonction Gamma d'Euler
241	Suites et séries de fonctions. Exemples et contre-exemples.	Équation de la chaleur Nombres de Bell Prolongement de la fonction Gamma d'Euler
243	Séries entières, propriétés de la somme. Exemples et applications.	Expression des $\zeta(2k)$ Nombres de Bell

Numéro	Leçon	Développements
244	Exemples d'études et d'applications de fonctions usuelles et spéciales.	Expression des $\zeta(2k)$ Prolongement de la fonction Gamma d'Euler Série des inverses des nombres premiers
245	Fonctions holomorphes et méromorphes sur un ouvert de $\mathbb{C}$. Exemples et applications.	Base hilbertienne de polynômes orthogonaux Espace de Bergman Prolongement de la fonction Gamma d'Euler
246	Séries de Fourier. Exemples et applications.	Équation de la chaleur Expression des $\zeta(2k)$
250	Transformation de Fourier. Applications.	Base hilbertienne de polynômes orthogonaux Théorème central limite
253	Utilisation de la notion de convexité en analyse.	Algorithme du gradient à pas optimal Projection sur un convexe fermé
261	Loi d'une variable aléatoire: caractérisations, exemples, applications.	Formule de Stirling (par le théorème central limite) Théorème central limite
262	Convergences d'une suite de variables aléatoires. Théorèmes limite. Exemples et applications.	Formule de Stirling (par le théorème central limite) Théorème central limite
264	Variables aléatoires discrètes. Exemples et applications.	Formule de Stirling (par le théorème central limite) Série des inverses des nombres premiers
266	Utilisation de la notion d'indépendance en probabilités.	Formule de Stirling (par le théorème central limite) Série des inverses des nombres premiers Théorème central limite

Algorithme du gradient à pas optimal

[FG] **Théorème 1.** Soit $f : \mathbb{R}^n \to \mathbb{R}$ une fonction elliptique. Alors, f passe par un minimum local en unique point $x^\star$, et ce minimum est global. De plus, la suite $(x_n)_{n\in\mathbb{N}}$ définie par $x_0 \in \mathbb{R}^n$ et,

$$x_{n+1} = \begin{cases} x^\star & \text{si } x_n = x^\star, \\ x_n + s_n \nabla f(x_n) & \text{sinon.} \end{cases}$$

où $s_n := \arg\min_{t\in\mathbb{R}} f(x_n + t\nabla f(x_n))$, est bien définie et converge vers $x^\star$.

Preuve. Tout d'abord, f étant strictement convexe et coercive, elle passe bien par un minimum local en unique point $x^\star$, et ce minimum est global. Montrons ensuite que $(x_n)_{n\in\mathbb{N}}$ est bien définie et converge vers $x^\star$.

Étape n°1. Soit $x \in \mathbb{R}^n$ tel que $\nabla f(x) \neq 0$. Montrons que la fonction,

$$\begin{array}{rccc} \varphi_x & : & \mathbb{R} & \longrightarrow & \mathbb{R} \\ & & t & \longmapsto & f(x + t\nabla f(x)) \end{array}$$

passe par un minimum global en un unique point $s \in \mathbb{R}$. En premier lieu, on remarque que cette fonction est coercive. En effet, pour tout $x \in \mathbb{R}^n$,

$$\|x + t\nabla f(x)\| \geq |t|\|\nabla f(x)\| - \|x\| \underset{|t|\to+\infty}{\longrightarrow} +\infty$$

car $\nabla f(x) \neq 0$, ce qui implique $\lim_{|t|\to+\infty} \|x + t\nabla f(x)\| = +\infty$, puis,

$$\lim_{|t|\to+\infty} f(x + t\nabla f(x)\| = +\infty$$

par composition de limites, puisque f est coercive. Autrement dit, la fonction φ_x est coercive, et donc φ_x passe par un minimum global en un point $s \in \mathbb{R}$ (cf. proposition 3). D'autre part, φ_x est dérivable, avec, pour tout $t \in \mathbb{R}$,

$$\varphi_x'(t) = Df(x + t\nabla f(x))(\nabla f(x)) = (\nabla f(x + t\nabla f(x)) \mid \nabla f(x)).$$

En particulier, $\varphi_x'(s) = (\nabla f(x + s\nabla f(x)) \mid \nabla f(x))$. Or, $\varphi_x'(s) = 0$, puisque φ_x passe par un minimum en s. Par conséquent, les gradients de f aux points $x + s\nabla f(x)$ et $\nabla f(x)$ sont orthogonaux. Or, pour tout $t \in \mathbb{R}$,

$$\begin{aligned} &\varphi_x(t) - \varphi_x(s) = f(x + t\nabla f(x)) - f(x + s\nabla f(x)) \\ &\geq (t - s)(\nabla f(x + s\nabla f(x)) \mid \nabla f(x)) + \frac{\alpha}{2}|t - s|^2\|\nabla f(x)\|^2 \end{aligned}$$

puisque f est α-convexe (on a appliqué le corollaire 8 aux points $x + t\nabla f(x)$ et $x + s\nabla f(x)$). Et donc, pour tout $t \in \mathbb{R} \setminus \{s\}$,

$$\varphi_x(t) - \varphi_x(s) > \frac{\alpha}{2}|t - s|^2\|\nabla f(x)\|^2$$

car $(\nabla f(x + s\nabla f(x)) \mid \nabla f(x)) = 0$ et $\nabla f(x) \neq 0$. De cette inégalité, on en déduit que φ_x passe par son minimum global uniquement au point s.

Étape n°2. Soit $x \in \mathbb{R}^n$ tel que $\nabla f(x) = 0$. Alors, pour tout $y \in \mathbb{R}^n \setminus \{x\}$,

$$f(y) - f(x) \geq (\nabla f(x) \mid y - x) + \frac{\alpha}{2}\|y - x\|^2 = \frac{\alpha}{2}\|y - x\|^2 > 0$$

puisque f est α-convexe (cf. corollaire 8). De cette inégalité, on en déduit que f passe par son minimum global uniquement au point x, et donc que $x = x^\star$. Ainsi, pour tout $x \in \mathbb{R}^n$,

$$\nabla f(x) = 0 \Longrightarrow x = x^\star$$

(en fait la réciproque est vraie puisque $x^\star$ est un point de minimum global de f), c'est-à-dire en écrivant la contraposée de cette assertion,

$$x \neq x^\star \Longrightarrow \nabla f(x) \neq 0.$$

D'après l'étape n°1, l'assertion ci-dessus permet de conclure que $(x_n)_{n\in\mathbb{N}}$ est bien définie (c'est-à-dire que s_n a bien un sens dès que $x_n \neq x^\star$).

Étape n°3. À présent, montrons que $(x_n)_{n\in\mathbb{N}}$ converge vers $x^\star$. Tout d'abord, s'il existe $n \in \mathbb{N}$ tel que $x_n = x^\star$, alors $(x_n)_{n\in\mathbb{N}}$ est stationnaire en $x^\star$. On peut donc supposer que, pour tout $n \in \mathbb{N}$, $x_n \neq x^\star$, et donc $\nabla f(x_n) \neq 0$ (d'après l'étape n°2). Alors, pour tout $n \in \mathbb{N}$,

$$f(x_n) - f(x_{n+1}) \geq (\nabla f(x_{n+1}) \mid x_n - x_{n+1}) + \frac{\alpha}{2}\|x_n - x_{n+1}\|^2$$

car f est α-convexe (on a appliqué le corollaire 8 aux points x_n et x_{n+1}). Or,

$$(\nabla f(x_{n+1}) \mid x_n - x_{n+1}) = (\nabla f(x_n + s_n\nabla f(x_n)) \mid \nabla f(x_n)) = 0$$

car $\nabla f(x_n) \neq 0$ (cf. étape n°1). D'où, pour tout $n \in \mathbb{N}$,

$$f(x_n) - f(x_{n+1}) \geq \frac{\alpha}{2}\|x_n - x_{n+1}\|^2 \geq 0$$

et donc $(f(x_n))_{n\in\mathbb{N}}$ est décroissante. Or, cette suite est également minorée (par $f(x^\star)$ notamment). Donc, $(f(x_n))_{n\in\mathbb{N}}$ converge. D'après l'inégalité précédente, ceci entraîne que,

$$\lim_{n\to+\infty} x_{n+1} - x_n = 0.$$

D'autre part, f étant coercive, il existe $A > 0$ tel que,

$$\forall x \in \mathbb{R}^n,\ \|x\| > A \Longrightarrow f(x) > f(x_0).$$

Puisque la suite $(f(x_n))_{n\in\mathbb{N}}$ est décroissante, ceci entraîne que,

$$\{x_n,\ n \in \mathbb{N}\} \subset \overline{B(0,A)}$$

(où $\overline{B(0,A)}$ désigne la boule fermée de centre 0 et de rayon A dans $\mathbb{R}^n$). Par le théorème de Bolzano-Weierstrass, il existe donc une suite $(x_{\varphi(n)})_{n\in\mathbb{N}}$ extraite de $(x_n)_{n\in\mathbb{N}}$ qui converge vers un certain x_∞. Montrons que $x_\infty = x^\star$. Puisque $(x_{\varphi(n)+1} - x_{\varphi(n)})_{n\in\mathbb{N}}$ converge vers 0, la suite $(x_{\varphi(n)+1})_{n\in\mathbb{N}}$ converge aussi vers $x^\star$. Sachant que f est de classe C^1, il s'en suit que,

$$\lim_{n\to+\infty} (\nabla f(x_{\varphi(n)+1}) \mid \nabla f(x_{\varphi(n)})) = \|\nabla f(x_\infty)\|$$

par continuité du gradient. Or, d'après l'étape n°1, pour tout $n \in \mathbb{N}$,

$$(\nabla f(x_{\varphi(n)+1}) \mid \nabla f(x_{\varphi(n)})) = (\nabla f(x_{\varphi(n)}) + s_{\varphi(n)} \nabla f(x_{\varphi(n)}) \mid \nabla f(x_{\varphi(n)})) = 0.$$

Et donc,

$$\lim_{n\to+\infty} (\nabla f(x_{\varphi(n)+1}) \mid \nabla f(x_{\varphi(n)})) = 0.$$

Par unicité de la limite, on en déduit que $\|\nabla f(x_\infty)\| = 0$, puis $\nabla f(x_\infty) = 0$, et donc $x_\infty = x^\star$ (d'après l'étape n°2). Ainsi, $(x_n)_{n\in\mathbb{N}}$ possède une unique valeur d'adhérence qui est $x^\star$. Puisque $(x_n)_{n\in\mathbb{N}}$ est à termes dans un compact, ceci permet de conclure que $(x_n)_{n\in\mathbb{N}}$ converge vers $x^\star$. □

Compléments conseillés.

- Compacité : proposition 2.
- Fonctions convexes et fonctions α-convexes : sous-section 3.15.
- Algorithme du gradient à pas optimal pour une forme quadratique : développement 24 de [JB].
- Mise en œuvre de l'algorithme de gradient à pas optimal pour une fonctionnelle quadratique.

Commentaires personnels. Développement technique, nécessitant quelques prérequis sur les fonctions α-convexes, mais qui se présente dans beaucoup de leçons différentes. De plus, l'algorithme de gradient à pas optimal pour une fonctionnelle quadratique figure sur le programme d'option calcul scientifique. Je vous recommande ce développement si vous avez choisi cette option.

Base hilbertienne de polynômes orthogonaux

Soit I un intervalle de $\mathbb{R}$.

Définition 1. On appelle fonction poids une fonction $\rho : I \to \mathbb{R}$ mesurable et strictement positive telle que,

$$\forall n \in \mathbb{N},\ \int_I |x|^n \rho(x)dx < +\infty.$$

Notons $L^2(\rho, I)$ l'espace de Hilbert $L^2_{\mathbb{C}}(I, \mathcal{B}(I), \mu)$, où μ désigne la mesure de densité ρ par rapport à la mesure de Lebesgue sur I.

Remarque. Le produit scalaire sur cet espace est donc défini par,

$$\forall f, g \in L^2(\rho, I),\ (f \mid g)_\rho = \int_I f(x)\overline{g(x)}d\mu(x) = \int_I f(x)\overline{g(x)}\rho(x)dx.$$

De plus, pour tout $n \in \mathbb{N}$, la fonction,

$$\begin{array}{rccc} e_n & : & I & \longrightarrow & \mathbb{C} \\ & & x & \longmapsto & x^n \end{array}$$

appartient à $L^2(\rho, I)$ par construction.

[FG] **Théorème 2.** Soit $\rho : I \to \mathbb{R}$ une fonction poids. Notons $(P_n)_{n\in\mathbb{N}}$ la famille obtenue en appliquant le procédé d'orthonormalisation de Gram-Schmidt à la famille $(e_n)_{n\in\mathbb{N}}$ dans $L^2(\rho, I)$. Supposons qu'il existe $\alpha > 0$ tel que,

$$\int_I e^{\alpha|x|}\rho(x)dx < +\infty.$$

Alors, la famille $(P_n)_{n\in\mathbb{N}}$ est une base hilbertienne de $(L^2(\rho, I), (\cdot \mid \cdot)_\rho)$.

Preuve. Par construction, $(P_n)_{n\in\mathbb{N}}$ est une famille orthonormée de l'espace de Hilbert $(L^2(\rho, I), (\cdot \mid \cdot)_\rho)$. Pour conclure, il suffit donc de montrer que,

$$\{P_n,\ n \in \mathbb{N}\}^\perp = \{0\}.$$

ce qui équivaut à $\{e_n,\ n \in \mathbb{N}\}^\perp = \{0\}$ puisque $\{P_n,\ n \in \mathbb{N}\}^\perp = \{e_n,\ n \in \mathbb{N}\}^\perp$. Soit $f \in L^2(\rho, I)$. On introduit alors la fonction suivante,

$$\begin{array}{rccc} \phi & : & \mathbb{R} & \longrightarrow & \mathbb{C} \\ & & x & \longmapsto & \begin{cases} \rho(x)f(x) & \text{si } x \in I, \\ 0 & \text{sinon.} \end{cases} \end{array}$$

Montrons que $\phi \in L^1(\mathbb{R})$. Pour tout $t \in \mathbb{R}_+$,

$$t \leq \frac{1}{2}(1 + t^2)$$

car $(1 + t)^2 = 1 + t^2 - 2t \geq 0$. Donc, pour tout $x \in I$,

$$|\phi(x)| = |f(x)|\rho(x) \leq \frac{1}{2}\left(1 + |f(x)|^2\right)\rho(x).$$

Puis,

$$\int_{\mathbb{R}} |\phi(x)|dx = \int_I |f(x)|\rho(x)dx \leq \frac{1}{2}\int_I \rho(x)dx + \frac{1}{2}\int_I |f(x)|^2\rho(x)dx.$$

Or, $\displaystyle\int_I \rho(x)dx < +\infty$, car ρ est une fonction poids, et $\displaystyle\int_I |f(x)|^2\rho(x)dx < +\infty$, car $f \in L^2(\rho, I)$. D'où, $\displaystyle\int_{\mathbb{R}} |\phi(x)|dx < +\infty$, c'est-à-dire $\phi \in L^1(\mathbb{R})$. On peut donc considérer sa transformée de Fourier $\widehat{\phi}$, c'est-à-dire la fonction définie pour tout $\xi \in \mathbb{R}$ par,

$$\widehat{\phi}(\xi) = \int_{\mathbb{R}} \phi(x)e^{-i\xi x}dx.$$

Montrons que $\widehat{\phi}$ se prolonge en une fonction G holomorphe sur,

$$B_\alpha := \left\{z \in \mathbb{C},\ |\mathrm{Im}(z)| < \frac{\alpha}{2}\right\}.$$

Et pour ce faire, nous allons appliquer le théorème d'holomorphie sous le signe intégral à la fonction,

$$\begin{array}{rccc} g & : & B_\alpha \times I & \longrightarrow & \mathbb{C} \\ & & (z, x) & \longmapsto & e^{-izx}f(x)\rho(x) \end{array}$$

(i) Pour tout $z \in B_\alpha$, la fonction $I \to \mathbb{C},\ t \mapsto g(z, t)$ est mesurable.

(ii) Pour tout $t \in I$, la fonction $B_\alpha \to \mathbb{C},\ z \mapsto g(z, t)$ est holomorphe.

(iii) Soit $(z, t) \in B_\alpha \times I$. Alors,

$$e^{-izx} = e^{-i(\Re(z)+i\,\mathrm{Im}(z))x} = e^{\mathrm{Im}(z)x}e^{-i\Re(z)x}$$

et donc $|e^{-izx}| = e^{\mathrm{Im}(z)x} \leq e^{\frac{\alpha|x|}{2}}$. Par suite,

$$|g(z, x)| \leq e^{\frac{\alpha|x|}{2}}|f(x)|\rho(x).$$

Or,

$$e^{\frac{\alpha|x|}{2}}|f(x)|\rho(x) = \left(e^{\frac{\alpha|x|}{2}}\sqrt{\rho(x)}\right)\left(|f(x)|\sqrt{\rho(x)}\right)$$

avec $\displaystyle\int_I e^{\alpha|x|}\rho(x)dx < +\infty$ par hypothèse sur ρ et $\displaystyle\int_I |f(x)|^2\rho(x)dx < +\infty$, puisque $f \in L^2(\rho, I)$. D'après l'inégalité de Cauchy-Schwarz, la fonction $x \mapsto e^{\frac{\alpha|x|}{2}}|f(x)|\rho(x)$ est donc intégrable sur I avec,

$$\int_I e^{\frac{\alpha|x|}{2}}|f(x)|\rho(x)dx \leq \left(\int_I e^{\alpha|x|}\rho(x)dx\right)^{\frac{1}{2}}\left(\int_I |f(x)|^2\rho(x)dx\right)^{\frac{1}{2}}.$$

Ainsi,

$$\forall (z,t) \in B_\alpha \times I,\ |g(z,t)| \leq e^{\frac{\alpha|x|}{2}}|f(x)|\rho(x)$$

avec $x \mapsto e^{\frac{\alpha|x|}{2}}|f(x)|\rho(x)$ qui est intégrable sur I.

D'après le théorème d'holomorphie sous le signe intégral, la fonction,

$$\begin{array}{rccc} G & : & B_\alpha & \longrightarrow & \mathbb{C} \\ & & z & \longmapsto & \displaystyle\int_I g(z,t)dt \end{array}$$

est donc bien définie et holomorphe sur B_α, avec,

$$\forall n \in \mathbb{N},\ \forall z \in B_\alpha,\ G^{(n)}(z) = (-i)^n \int_I x^n e^{-izx} f(x)\rho(x)dx.$$

En particulier,

$$\forall n \in \mathbb{N},\ G^{(n)}(0) = (-i)^n \int_I x^n f(x)\rho(x)dx = (-i)^n (f \mid e_n)_\rho.$$

Supposons à présent que $f \in \{e_n,\ n \in \mathbb{N}\}^\perp$. Alors, d'après l'égalité ci-dessus, pour tout $n \in \mathbb{N}$, $G^{(n)}(0) = 0$. Or, la fonction G étant holomorphe sur B_α, on a donc,

$$\forall z \in D\Big(0, \frac{\alpha}{2}\Big),\ G(z) = \sum_{n=0}^{\infty} \frac{G^{(n)}(0)}{n!},$$

(puisque $d(0, \mathbb{C}\setminus B_\alpha) = \frac{\alpha}{2}$). Par suite, G est nulle sur $D(0, \frac{\alpha}{2})$. Sachant que B_α est un ouvert connexe, le principe du prolongement analytique assure alors que G est nulle sur B_α tout entier. En particulier, $\widehat{\phi}$ est nulle sur $\mathbb{R}$. Par injectivité de la transformation de Fourier sur $L^1(\mathbb{R})$, on en déduit que $\phi = 0$. Et puisque ρ est strictement positive, ceci entraîne que f est nulle dans $L^2(\rho, I)$. Ainsi,

$$\{P_n,\ n \in \mathbb{N}\}^\perp = \{e_n,\ n \in \mathbb{N}\}^\perp = \{0\},$$

ce qui prouve finalement que $(P_n)_{n\in\mathbb{N}}$ est une base hilbertienne de $L^2(\rho, I)$. □

Compléments conseillés.

- Donner un contre-exemple si ρ ne vérifie par la condition "il existe $\alpha > 0$ tel que,
$$\int_I e^{\alpha|x|}\rho(x)dx < +\infty".$$
- Donner un exemple de familles de polynômes vérifiant ce théorème.
- Déduire de ce théorème une base hilbertienne de $L^2(\mathbb{R})$ (lien ici).

Commentaires personnels. Développement technique, exploitant de nombreux théorèmes d'analyse au programme : théorème d'holomorphie sous le signe intégral, principe du prolongement analytique et injectivité de la transformée de Fourier sur $L^1(\mathbb{R})$. Je pense qu'il est important de bien citer toutes les hypothèses requises pour ces théorèmes (à l'oral). En revanche, ce développement se comprend et se retient facilement. De plus, il se recase dans de très nombreuses leçons. Je le recommande.

Équation de la chaleur

[IP] **Théorème 1.** Soit $u_0 \in L^2_{\mathbb{R}}([0,2\pi])$. Notons $(d_n)_{n\in\mathbb{Z}}$ la suite de ses coefficients de Fourier. Il existe une unique fonction $u : \mathbb{R}_+^* \times \mathbb{R} \to \mathbb{R}$ telle que,

(i) pour tout $t > 0$, la fonction $u(t,\cdot)$ est 2π-périodique;

(ii) les fonctions $\partial_t u$ et $\partial_x^2 u$ sont bien définies et continues sur $\mathbb{R}_+^* \times \mathbb{R}$;

(iii) $\partial_t u = \partial_x^2 u$ sur $\mathbb{R}_+^* \times \mathbb{R}$;

(iv) $\|u(t,\cdot) - u_0\|_2 \underset{t\to 0}{\longrightarrow} 0$.

De plus, la fonction u est de classe C^∞ sur $\mathbb{R}_+^* \times \mathbb{R}$.

Preuve. Procédons par analyse-synthèse.

Analyse. Supposons qu'il existe une fonction $u : \mathbb{R}_+^* \times \mathbb{R} \to \mathbb{R}$ vérifiant les points (i), (ii), (iii) et (iv). Alors, pour tout $t \in \mathbb{R}_+^*$, $u(t,\cdot) \in L^2_{\mathbb{R}}([0,2\pi])$, car $u(t,\cdot) \in C^0([0,2\pi])$ en conséquence du point (ii), ce qui nous permet de poser, pour tout $n \in \mathbb{Z}$,

$$c_n(t) = \frac{1}{2\pi}\int_0^{2\pi} u(t,x)e^{-inx}dx.$$

Il s'agit des coefficients de Fourier de la fonction 2π-périodique $u(t,\cdot)$. Fixons $n \in \mathbb{Z}$. Nous allons appliquer le théorème de dérivation sous le signe intégral à la fonction $(t,x) \mapsto u(t,x)e^{-inx}$ pour montrer que c_n est dérivable sur $\mathbb{R}_+^*$.

- Pour tout $t \in \mathbb{R}_+^*$, la fonction $u(t,\cdot)$ est intégrable sur $[0,2\pi]$ (car elle est continue).
- Pour tout $x \in [0,2\pi]$, la fonction $u(\cdot,x)$ est dérivable sur $\mathbb{R}_+^*$.
- Soit I un segment inclus dans $\mathbb{R}_+^*$. Alors,

$$\forall (t,x) \in I \times [0,2\pi],\ |u(t,x)e^{-inx}| = |u(t,x)| \leq \sup_{I\times[0,2\pi]} |u|$$

avec $(t,x) \mapsto \sup\limits_{I\times[0,2\pi]} |u|$ qui est bien intégrable sur $[0,2\pi]$.

Ainsi, d'après le théorème de dérivation sous le signe intégral, c_n est dérivable sur $\mathbb{R}_+^*$ avec,

$$\forall t \in \mathbb{R}_+^*,\ c_n'(t) = \int_0^{2\pi} \partial_t u(t,x)e^{-inx}dx.$$

Puis, en utilisant le point (iii),

$$\forall t \in \mathbb{R}_+^*,\ c_n'(t) = \frac{1}{2\pi}\int_0^{2\pi} \partial_x^2 u(t,x)e^{-inx}dx$$

c'est-à-dire que $c_n'(t)$ est le n-ième coefficient de Fourier de la fonction $\partial_x^2 u(t,\cdot)$. Les propriétés de calcul des coefficients de Fourier de la dérivée d'une fonction fournissent alors,

$$\forall t \in \mathbb{R}_+^*,\ c_n'(t) = (ni)^2 c_n(t) = -n^2 c_n(t).$$

Ainsi, c_n est solution de l'équation différentielle linéaire $y' = -n^2 y$, d'inconnue $y : \mathbb{R}_+^* \to \mathbb{R}$ dérivable. D'où l'existence de $\alpha_n \in \mathbb{R}$ tel que,

$$\forall t \in \mathbb{R}_+^*,\ c_n(t) = \alpha_n \exp(-n^2 t).$$

Il reste à déterminer α_n. D'après la formule de Parseval, pour tout $t \in \mathbb{R}_+^*$,

$$\|u(t,\cdot) - u_0\|_2^2 = \sum_{k=-\infty}^{+\infty} |c_k(t) - d_k|^2 = \sum_{k=-\infty}^{+\infty} |\alpha_k e^{-k^2 t} - d_k|^2$$

car $u(t,\cdot) - u_0 \in L^2_{\mathbb{R}}([0,2\pi])$. Par suite, pour tout $t \in \mathbb{R}_+^*$,

$$0 \leq |\alpha_n e^{-n^2 t} - d_n|^2 \leq \sum_{k=-\infty}^{+\infty} |\alpha_k e^{-k^2 t} - d_k|^2 = \|u(t,\cdot) - u_0\|_2^2 \underset{t\to 0}{\longrightarrow} 0.$$

D'où, par encadrement de limites,

$$|\alpha_n e^{-n^2 t} - d_n|^2 \underset{t\to 0}{\longrightarrow} 0.$$

Sachant que,

$$|\alpha_n e^{-n^2 t} - d_n|^2 \underset{t\to 0}{\longrightarrow} |\alpha_n - d_n|^2$$

il vient $\alpha_n = d_n$, et donc,

$$\forall t \in \mathbb{R}_+^*,\ c_n(t) = d_n \exp(-n^2 t).$$

Par ailleurs, d'après le théorème de Dirichlet, pour tout $t \in \mathbb{R}_+^*$,

$$\forall x \in \mathbb{R},\ u(t,x) = \sum_{n=-\infty}^{+\infty} c_n(t)e^{inx}$$

car $u(t,\cdot)$ est 2π-périodique et de classe C^1 sur $\mathbb{R}$. Finalement,

$$\forall (t,x) \in \mathbb{R}_+^* \times \mathbb{R},\ u(t,x) = \sum_{n=-\infty}^{+\infty} d_n e^{-n^2 t} e^{inx}$$

Ainsi, si une solution existe, elle est unique, et son expression est donnée par la formule ci-dessus.

Synthèse. On souhaite définir $u : \mathbb{R}_+^* \times \mathbb{R} \to \mathbb{C}$ par,

$$\forall (t,x) \in \mathbb{R}_+^* \times \mathbb{R},\ u(t,x) = \sum_{n=-\infty}^{+\infty} d_n e^{-n^2 t} e^{inx}.$$

Montrons qu'une telle fonction est bien définie et de classe C^∞ sur $\mathbb{R}_+^* \times \mathbb{R}$. Pour cela, on applique le théorème de dérivation sous le signe somme à la suite de fonctions $(f_n)_{n\in\mathbb{Z}}$ définie par,

$$\begin{array}{rccc} f_n & : & \mathbb{R}_+^* \times \mathbb{R} & \longrightarrow & \mathbb{C} \\ & & (t,x) & \longmapsto & d_n e^{-n^2 t} e^{inx} \end{array}$$

Le théorème de dérivation sous le signe somme évoqué ci-dessus, correspond au théorème de dérivation sous le signe intégral, dans le cas particulier où l'espace mesuré est $(\mathbb{Z}, \mathcal{P}(\mathbb{Z}), \text{card})$.

- Pour tout $n \in \mathbb{Z}$, la fonction f_n est de classe C^∞ avec, pour tous $a, b \in \mathbb{N}$, et pour tout $(t,x) \in \mathbb{R}_+^* \times \mathbb{R}$,

$$\partial_t^a \partial_x^b f_n(t,x) = (-1)^a i^b d_n n^{2a+b} e^{-n^2 t} e^{inx}$$

- Soient $a, b \in \mathbb{N}$, $n \in \mathbb{Z}$ et $t_0 \in \mathbb{R}_+^*$. Alors, pour tout $(t,x) \in \mathbb{R}_+^*$,

$$|\partial_t^a \partial_x^b f_n(t,x)| = |d_n||n|^{2a+b} e^{-n^2 t} \leq |d_n||n|^{2a+b} e^{-n^2 t_0}.$$

  D'autre part, d'après la formule de Parseval,

$$\|u_0\|_2^2 = \sum_{k=-\infty}^{+\infty} |d_k|^2 \geq |d_n|^2.$$

  et donc $|d_n| \leq \|u_0\|_2$. Par suite, pour tout $n \in \mathbb{Z}$ et pour tout $(t,x) \in \mathbb{R}_+^*$,

$$|\partial_t^a \partial_x^b f_n(t,x)| \leq \|u_0\|_2 |n|^{2a+b} e^{-n^2 t_0}.$$

  avec $\displaystyle\sum_{n=-\infty}^{+\infty} \|u_0\|_2 |n|^{2a+b} e^{-n^2 t_0} < +\infty$ car,

$$\|u_0\|_2 |n|^{2a+b} e^{-n^2 t_0} \underset{|n|\to+\infty}{=} o\left(\frac{1}{n^2}\right).$$

D'après le théorème de dérivation sous le signe intégral, la fonction,

$$\begin{array}{rccc} u & : & \mathbb{R}_+^* \times \mathbb{R} & \longrightarrow & \mathbb{C} \\ & & (t,x) & \longmapsto & \displaystyle\sum_{n=-\infty}^{+\infty} f_n(t,x) \end{array}$$

est bien définie et de classe C^∞, avec, pour tous $a, b \in \mathbb{N}$, et,

$$\forall (t,x) \in \mathbb{R}_+^* \times \mathbb{R},\ \partial_t^a \partial_x^b u(t,x) = (-1)^a i^b \sum_{n=-\infty}^{+\infty} n^{2a+b} e^{-n^2 t} e^{inx}.$$

En particulier, $\partial_t u$ et ∂_x^2 sont bien définies et continues sur $\mathbb{R}_+^* \times \mathbb{R}$, avec, pour tout $(t,x) \in \mathbb{R}_+^* \times \mathbb{R}$,

$$\partial_t u(t,x) = -\sum_{n=-\infty}^{+\infty} n^2 e^{-n^2 t} e^{inx} = \partial_x^2 u(t,x).$$

Donc u vérifie les points (ii) et (iii). De plus, pour tout $t > 0$, $u(t,\cdot)$ est bien 2π-périodique, et donc u vérifie aussi le point (i). Par ailleurs, il faut vérifier que u est à valeurs réelles (ce qui n'est pas immédiat car les f_n sont à valeurs complexes !). Notons $(a_n)_{n\in\mathbb{N}}$ et $(b_n)_{n\in\mathbb{N}^*}$ la suite des coefficients de Fourier trigonométriques de u_0. Soit $(t,x) \in \mathbb{R}_+^* \times \mathbb{R}$. Alors,

$$u(t,x) = d_0 + \sum_{n=1}^{+\infty} (d_{-n} e^{-inx} + d_n e^{inx}) e^{-n^2 t}.$$

Or, $d_0 = a_0$ et, pour tout $n \in \mathbb{N}$,

$$d_{-n} e^{-inx} + d_n e^{inx} = a_n \cos(nx) + b_n \sin(nx).$$

D'où,

$$u(t,x) = a_0 + \sum_{n=1}^{+\infty} (a_n \cos(nx) + b_n \sin(nx)) e^{-n^2 t}.$$

Sachant que $(a_n)_{n\in\mathbb{N}}$ et $(b_n)_{n\in\mathbb{N}^*}$ sont des suites réelles (car la fonction u_0 est à valeurs réelles), on en déduit que $u(t,x) \in \mathbb{R}$. Finalement, il reste à prouver le point (iv). Soit $t > 0$. Pour tout $n \in \mathbb{Z}$, le n-ième coefficient de Fourier de la fonction $u(t,\cdot)$ est $d_n e^{-n^2 t}$. En effet, pour tout $n \in \mathbb{Z}$,

$$\frac{1}{2\pi} \int_0^{2\pi} u(t,x) e^{-inx} dx = \sum_{k=-\infty}^{+\infty} d_k e^{-k^2 t} \left(\frac{1}{2\pi} \int_0^{2\pi} e^{i(k-n)x} \right.$$

(la permutation série-intégrale étant permise par la convergence uniforme de la série $\sum_{n\in\mathbb{N}^*}(f_{-n}+f_n)$ sur $[0,2\pi]$, et cette série converge bien uniformément sur $[0,2\pi]$ car elle converge normalement sur $[0,2\pi]$), et donc,

$$\frac{1}{2\pi}\int_0^{2\pi} u(t,x)e^{-inx}dx = \sum_{k=-\infty}^{+\infty} d_k e^{-k^2t}\delta_{k,n} = d_n e^{-n^2t}.$$

D'où, en utilisant la formule de Parseval,

$$\|u(t,\cdot)-u_0\|_2^2 = \sum_{n=-\infty}^{+\infty} |d_n|^2|1-e^{-n^2t}|^2. \tag{1}$$

Pour déterminer la limite de cette quantité lorsque t tend vers 0, nous allons appliquer le théorème de convergence dominée (à paramètre) à la fonction,

$$\begin{array}{ccc} \mathbb{R}_+^*\times\mathbb{Z} & \longrightarrow & \mathbb{R} \\ (t,n) & \longmapsto & |d_n|^2|1-e^{-n^2t}|^2 \end{array}$$

sur l'espace mesuré $(\mathbb{Z},\mathcal{P}(\mathbb{Z}),\text{card})$.

- Pour tout $t\in\mathbb{R}$, la fonction $\mathbb{N}\to\mathbb{R},\ n\mapsto|d_n|^2|1-e^{-n^2t}|^2$ est mesurable de $(\mathbb{Z},\mathcal{P}(\mathbb{Z}))$ vers $(\mathbb{R},\mathcal{B}(\mathbb{R}))$ (en fait, c'est toujours le cas lorsqu'on munit l'ensemble de départ de sa tribu discrète, comme c'est le cas ici).
- Pour tout $n\in\mathbb{Z}$, $|d_n|^2|1-e^{-n^2t}|^2 \underset{t\to0}{\longrightarrow} 0$.
- Pour tout $n\in\mathbb{Z}$, et pour tout $t\in\mathbb{R}_+^*$,

  $$|d_n|^2|1-e^{-n^2t}|^2 \le |d_n|^2.$$

  De plus, d'après la formule de Parseval,

  $$\sum_{n=-\infty}^{+\infty}|d_n|^2 = \|u_0\|_2^2 < +\infty$$

  car $u_0\in L^2_{\mathbb{R}}([0,2\pi])$.

Ainsi, d'après le théorème de convergence dominée (à paramètre),

$$\lim_{t\to0}\sum_{n=-\infty}^{+\infty}|d_n|^2|1-e^{-n^2t}|^2 = 0.$$

En utilisant l'égalité (1), il en résulte que u vérifie le point (iv), ce qui achève la preuve. □

Remarque. Traditionnellement, l'énoncé du théorème de convergence dominée est restreint aux suites de fonctions. Cependant, la caractérisation séquentielle de la limite permet d'étendre ce résultat aux fonctions à paramètre.

Commentaires personnels. Développement technique, exploitant de nombreux théorèmes d'analyse au programme : théorème de dérivation sous le signe intégral, permutation série-intégrale (en cas de convergence uniforme), formule de Parseval, théorème de convergence dominée (à paramètre). Je pense qu'il est important de bien citer toutes les hypothèses requises pour ces théorèmes (à l'oral). De plus, l'équation de la chaleur figure sur le programme d'option calcul scientifique. Je vous recommande ce développement si vous avez choisi cette option.

Équation de Sylvester

Soit $n \in \mathbb{N}^*$. Soit $\|\cdot\|$ une norme d'algèbre sur $M_n(\mathbb{C})$. Par exemple, si $\|\cdot\|$ est une norme quelconque sur $M_{n,1}(\mathbb{C})$, alors la norme $\|\cdot\|$ sur $M_n(\mathbb{C})$ définie pour tout $A \in M_n(\mathbb{C})$ par,

$$\|A\| = \sup\left\{\frac{\|AX\|}{\|X\|},\ X \in M_{n,1}(\mathbb{C}) \setminus \{0\}\right\}$$

est une norme d'algèbre sur $M_n(\mathbb{C})$.

[XG] **Lemme 1.** Soit $A \in M_n(\mathbb{C})$ telle que, pour tout $\lambda \in \mathrm{Sp}_{\mathbb{C}}(A)$, $\Re(\lambda) < 0$. Alors, il existe $\alpha > 0$ et $K > 0$ tels que, pour tout $t \geq 0$, $\|e^{tA}\| \leq Ke^{-\alpha t}$.

Preuve. Puisque χ_A est scindé sur $\mathbb{C}$ (par le théorème de d'Alembert-Gauss) la décomposition de Dunford assure l'existence de $D \in M_n(\mathbb{C})$ diagonalisable et de $N \in M_n(\mathbb{C})$ nilpotente telles que $A = D + N$ et $DN = ND$. Par suite,

$$\forall t \in \mathbb{R},\ e^{tA} = e^{tD}e^{tN}.$$

car les matrices tD et tN commutent. Sachant que $\|\cdot\|$ est une norme d'algèbre, il vient donc, pour tout $t \in \mathbb{R}$,

$$\|e^{tA}\| \leq \|e^{tD}\|\|e^{tN}\|.$$

Fixons ensuite $t \in \mathbb{R}_+$, et trouvons une majoration de $\|e^{tD}\|$ et $\|e^{tN}\|$. Notons $\lambda_1, \ldots, \lambda_n$ les valeurs propres de A. Soit $P \in GL_n(\mathbb{C})$ tel que,

$$D = P\,\mathrm{Diag}(\lambda_1, \ldots, \lambda_n)P^{-1}.$$

Alors,

$$e^{tD} = P\,\mathrm{Diag}(e^{t\lambda_1}, \ldots, e^{t\lambda_n})P^{-1}.$$

De nouveau, puisque $\|\cdot\|$ est une norme d'algèbre, il vient,

$$\|e^{tD}\| \leq \|P\|\|\mathrm{Diag}(e^{t\lambda_1}, \ldots, e^{t\lambda_n})\|\|P^{-1}\|.$$

Notons ensuite $\|\cdot\|_\infty$ la norme sur $M_n(\mathbb{C})$ définie par,

$$\forall M = (m_{i,j})_{1\leq i,j\leq n} \in M_n(\mathbb{C}),\ \|M\|_\infty = \max_{1\leq i,j\leq n}|m_{i,j}|.$$

Alors, en notant $m = \max_{1\leq i\leq n} \Re(\lambda_i)$ (qui est indépendant de t) il vient,

$$\|\mathrm{Diag}(e^{t\lambda_1}, \ldots, e^{t\lambda_n})\|_\infty = e^{mt}.$$

Or, les normes $\|\cdot\|$ et $\|\cdot\|_\infty$ sont équivalentes (puisque $M_n(\mathbb{C})$ est de dimension finie). En particulier, il existe $c \in \mathbb{R}_+$ tel que, pour tout $M \in M_n(\mathbb{C})$, on ait $\|M\| \leq c\|M\|_\infty$, et donc,

$$\|e^{tD}\| \leq \|P\|\|P^{-1}\|ce^{mt}.$$

Par ailleurs, puisque N est nilpotente,

$$e^{tN} = \sum_{k=0}^{n} \frac{t^k}{k!}N^k,$$

et donc,

$$\|e^{tN}\| \leq \sum_{k=0}^{n} \frac{t^k}{k!}\left\|N^k\right\| \leq \sum_{k=0}^{n} \frac{t^k}{k!}\|N\|^k.$$

car $\|\cdot\|$ est une norme d'algèbre. D'où,

$$\|e^{tA}\| \leq \|e^{tD}\|\|e^{tN}\| \leq \|P\|\|P^{-1}\|ce^{mt}\sum_{k=0}^{n} \frac{t^k}{k!}\|N\|^k.$$

Or, par croissance comparée,

$$\|P\|\|P^{-1}\|ce^{mt}\sum_{k=0}^{n} \frac{t^k}{k!}\|N\|^k \underset{t\to+\infty}{=} o\left(e^{\frac{m}{2}t}\right)$$

(car $m < 0$). En particulier, il existe $K > 0$ tel que,

$$\forall t \in \mathbb{R}_+,\ \|P\|\|P^{-1}\|ce^{mt}\sum_{k=0}^{n} \frac{t^k}{k!}\|N\|^k \leq Ke^{\frac{m}{2}}.$$

Ainsi, en notant $\alpha = -\frac{m}{2} > 0$, il en résulte que,

$$\forall t \in \mathbb{R}_+,\ \|e^{tA}\| \leq Ke^{-\alpha t}.$$

□

[XG] **Théorème 2.** Soient $A, B \in M_n(\mathbb{R})$ telles que, pour toute valeur complexe λ de A ou de B, on ait $\Re(\lambda) < 0$. Alors, pour tout $C \in M_n(\mathbb{R})$, l'équation matricielle $AX + XB = C$ admet une unique solution.

Preuve. On sépare la preuve de l'existence et de l'unicité.

Existence. Étudions tout d'abord l'équation différentielle suivante,

$$Y'(t) = AY(t) + Y(t)B,\ t \in \mathbb{R} \tag{E}$$

d'inconnue une fonction $Y : \mathbb{R} \to M_n(\mathbb{R})$ dérivable. Notons que cette équation différentielle est linéaire. En effet, si on note, pour tout $(i,j) \in [\![1,n]\!]^2$, et pour tout $t \in \mathbb{R}$, $y_{i,j}(t)$ le coefficient en i-ième et j-ième colonne de la matrice $Y(t)$, il vient que Y est solution de (E) si et seulement si, les $y_{i,j}$ sont solution du système linéaire suivant,

$$y'_{i,j}(t) = \sum_{k=1}^{n} a_{i,k} y_{k,j}(t) + \sum_{k=1}^{n} b_{k,j} y_{i,k}(t),\ t \in \mathbb{R}.$$

Ainsi, d'après le théorème de Cauchy-Lipschitz linéaire, il existe une unique solution (globale) au problème de Cauchy,

$$\begin{cases} Y'(t) = AY(t) + Y(t)B,\ t \in \mathbb{R}, \\ Y(0) = C \end{cases} \tag{C}$$

En fait, ici le théorème de Cauchy-Lipschitz linéaire n'est pas nécessaire, car on peut expliciter une solution (globale) à ce problème de Cauchy. Il s'agit de l'application,

$$\begin{array}{rccc} Y & : & \mathbb{R} & \longrightarrow & M_n(\mathbb{R}) \\ & & t & \longmapsto & e^{tA} C e^{tB} \end{array}$$

En effet, rappelons que, pour tout $M \in M_n(\mathbb{C})$, et pour tout $t \in \mathbb{R}$,

$$(e^{tM})' = M e^{tM} = e^{tM} M$$

(M et e^{tM} commutent car e^{tM} est un polynôme en M). Et donc,

$$\forall t \in \mathbb{R},\ Y'(t) = (e^{tA})' C e^{tB} + e^{tA} C (e^{tB})' = AY(t) + Y(t)B.$$

De plus, $Y(0) = C$, et donc Y est bien solution du problème de Cauchy de (C). Autrement dit,

$$\forall t \in \mathbb{R},\ Y(t) = C + \int_0^t (AY(s) + Y(s)B) ds,$$

c'est-à-dire,

$$\forall t \in \mathbb{R},\ Y(t) = C + A\left(\int_0^t Y(s)\right) + \left(\int_0^t Y(s)\right) B. \tag{1}$$

Or, d'après le lemme précédent, il existe $K_1, K_2, \alpha_1, \alpha_2$ des réels strictement positifs tels que,

$$\forall t \in \mathbb{R},\ \|e^{tA}\| \leq K_1 e^{-\alpha_1 t},\ \|e^{tB}\| \leq K_2 e^{-\alpha_2 t}$$

car pour toute valeur complexe λ de A ou de B, on a $\Re(\lambda) < 0$. Et donc,

$$\forall t \in \mathbb{R},\ \|Y(t)\| \leq \|e^{tA}\| \|C\| \|e^{tB}\| \leq \|C\| K_1 K_2 e^{-(\alpha_1+\alpha_2)t}.$$

car $\|\cdot\|$ est une norme d'algèbre. De cette inégalité, on en déduit que,

(i) $\lim\limits_{t \to +\infty} Y(t) = 0$;

(ii) la fonction $t \mapsto Y(t)$ est intégrable sur $\mathbb{R}_+$.

On peut donc faire tendre t vers l'infini dans l'égalité (1), ce qui donne,

$$0 = C + A\left(\int_0^{+\infty} Y(s)\right) + \left(\int_0^{+\infty} Y(s) ds\right) B.$$

Ainsi,

$$X = -\int_0^{+\infty} Y(s) ds \in M_n(\mathbb{R})$$

vérifie $AX + XB = C$.

Unicité. Considérons l'application,

$$\begin{array}{rccc} \Phi & : & M_n(\mathbb{R}) & \longrightarrow & M_n(\mathbb{R}) \\ & & X & \longmapsto & AX + XB \end{array}$$

Étant donné que, pour tout $C \in M_n(\mathbb{R})$ l'équation matricielle $AX + XB = C$ d'inconnue $X \in M_n(\mathbb{R})$ admet au moins une solution, l'application Φ est donc surjective. Puisque Φ est un endomorphisme de $M_n(\mathbb{C})$ qui est de dimension finie, on en déduit que Φ est un isomorphisme de $M_n(\mathbb{C})$. En particulier, Φ est injectif, ce qui fournit l'unicité de l'équation matricielle $AX + XB = C$. □

Commentaires personnels. Développement technique, mais qui se comprend et se retient facilement. De plus, il se recase dans de nombreuses leçons. Je le recommande.

Équivalence des normes en dimension finie et théorème de Riesz

Soit $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$. Pour ce développement, il est nécessaire de connaître les deux propriétés qui caractérisent les compacts des espaces métriques : celle de Borel-Lebesgue et celle de Bolzano-Weierstrass. Nous aurons également besoin du résultat suivant :

Proposition 1. Soit E un espace vectoriel de dimension finie. Soit $(e_1, \dots, e_n)$ une base de E. Notons N_∞ la norme sur E définie par,

$$\forall x = \sum_{k=1}^{n} x_k e_k \in E,\ N_\infty(x) = \max_{1\leq k\leq n} |x_k|.$$

Les parties compactes de (E, N_∞) sont ses parties fermées bornées.

Preuve. Nous savons déjà que toute partie compacte d'un espace métrique est fermée et bornée. Réciproquement, considérons K une partie fermée bornée de (E, N_∞). Soit $(x_k)_{k\in\mathbb{N}}$ une suite d'éléments de K. Pour tout $k \in \mathbb{N}$ et pour tout $l \in [\![1, n]\!]$, on note $x_k^{(l)}$ la j-ième cordonnée du vecteur x_k dans la base $(e_1, \dots, e_n)$. Puisque K est une partie bornée de (E, N_∞), il existe $M \in \mathbb{R}_+$ tel que, pour tout $x \in K$, on ait $N_\infty(x) \leq M$. La définition de N_∞ donne alors,

$$\forall k \in \mathbb{N},\ \forall l \in [\![1, n]\!],\ |x_k^{(l)}| \leq N_\infty(x_k) \leq M.$$

Ainsi, pour tout $l \in [\![1, n]\!]$, la suite $(x_k^{(l)})_{k\in\mathbb{N}}$ est bornée. D'après le théorème de Bolzano-Weierstrass, il existe donc $\varphi_1, \dots, \varphi_n$ des applications strictement croissantes de $\mathbb{N}$ dans $\mathbb{N}$ et $x_\infty^{(1)}, \dots, x_\infty^{(n)} \in E$ telles que,

$$\forall l \in [\![1, n]\!],\ |x^{(l)}_{\varphi_l(k)} - x_\infty^{(l)}| \underset{k\to+\infty}{\longrightarrow} 0.$$

Par suite, en notant $\varphi = \varphi_1 \circ \cdots \circ \varphi_n$, il vient,

$$\forall l \in [\![1, n]\!],\ |x^{(l)}_{\varphi(k)} - x_\infty^{(l)}| \underset{k\to+\infty}{\longrightarrow} 0$$

car la suite $(x^{(l)}_{\varphi(k)})_{k\in\mathbb{N}}$ est extraite de $(x^{(l)}_{\varphi_l(k)})_{k\in\mathbb{N}}$. D'où, par définition de N_∞,

$$N_\infty(x_{\varphi(k)} - x_\infty) \underset{k\to+\infty}{\longrightarrow} 0$$

c'est-à-dire que $(x_{\varphi(k)})_{k\in\mathbb{N}}$ converge vers $x_\infty := \sum_{l=1}^{n} x_\infty^{(l)} e_l$ dans (E, N_∞). Et puisque K est une partie fermée de (E, N_∞), on a bien $x_\infty \in K$. Ainsi, K est une partie compacte de (E, N_∞). □

Les théorèmes qui suivent permettront de généraliser ce résultat à tout espace vectoriel normé de dimension finie.

Théorème 2. Dans un espace vectoriel normé de dimension finie, toutes les normes sont équivalentes. [XG]

Preuve. Soit (E, N) un espace vectoriel normé. Soit $\mathcal{B} = (e_1, \dots, e_n)$ une base de E. On note N_∞ la norme définie, pour tout $x \in E$ par,

$$N_\infty(x) = \max_{1\leq k\leq n} |x_k|.$$

où $x_1, \dots, x_n$ désigne les coordonnées du vecteur x dans la base $\mathcal{B}$. Montrons que les normes N et N_∞ sont équivalentes, c'est-à-dire qu'il existe $\alpha, \beta \in \mathbb{R}_+^*$ tels que, pour tout $x \in E$,

$$\alpha N_\infty(x) \leq N(x) \leq \beta N_\infty(x).$$

Commençons par prouver l'existence de β. Pour tout $x = (x_1, \dots, x_n)_\mathcal{B} \in E$,

$$N(x) = N\left(\sum_{k=1}^{n} x_k e_k\right) \leq \sum_{k=1}^{n} |x_k| N(e_k) \leq N_\infty(x) \sum_{k=1}^{n} N(e_k).$$

Ainsi, en notant $\beta = \sum_{k=1}^{n} N(e_k) \in \mathbb{R}_+^*$, il vient, pour tout $x \in E$,

$$N(x) \leq \beta N_\infty(x).$$

Il reste à prouver l'existence de α. Pour cela, on note,

$$S = \{x \in E,\ N_\infty(x) = 1\}$$

qui est une partie compacte de (E, N_∞) en tant que partie fermée bornée de cet espace. D'autre part, l'application $N : (E, N_\infty) \to (\mathbb{R}, |\cdot|)$ est continue, car elle est β-lipschitzienne. En effet, pour tous $x, y \in E$,

$$|N(x) - N(y)| \leq N(x - y) \leq \beta N_\infty(x - y).$$

Puisque K est compact, on en déduit que N est bornée et atteint ses bornes sur K. En particulier, il existe $x_0 \in K$ tel que,

$$\inf_{x\in K} N(x) = N(x_0).$$

Alors, pour tout $x \in E$,

$$N(x) = N\left(N_\infty(x)\frac{1}{N_\infty(x)}x\right) = N_\infty(x)N\left(\frac{1}{N_\infty(x)}x\right) \geq N_\infty(x)N(x_0)$$

(la deuxième égalité résulte de l'homogénéité de N, et l'inégalité du fait que $\frac{1}{N_\infty(x)}x \in K$). Ainsi, en notant $\alpha = N(x_0) \in \mathbb{R}_+^*$, il vient, pour tout $x \in E$,

$$N(x) \geq \alpha N_\infty(x).$$

Finalement, on en conclut que N et N_∞ sont équivalentes. En particulier, si N_1 et N_2 sont deux normes sur E alors elles sont toutes deux équivalentes à N_∞, et donc N_1 et N_2 sont équivalentes (par transitivité). □

Corollaire 3. Soit (E, N) un espace vectoriel normé de dimension finie. Les parties compactes de (E, N) sont ses parties fermées bornées.

Preuve. On sait déjà que les parties compactes de (E, N_∞) sont ses parties fermées bornées. Sachant que les normes N et N_∞ sont équivalentes, il s'en suit que les parties compactes de (E, N) sont ses parties fermées bornées. □

La preuve du théorème ci-dessous diffère légèrement de celle de [XG]. Cette preuve alternative provient du cours d'Emmanuel Russ, dispensé en master de préparation à l'agrégation de mathématiques, à l'université d'Aix-Marseille.

[XG] **Théorème 4 (Théorème de Riesz).** Soit $(E, \|\cdot\|)$ un espace vectoriel normé. La boule $\overline{B_E(0,1)}$ est compacte si et seulement si E est de dimension finie.

Preuve. Si E est de dimension finie, c'est une conséquence du corollaire précédent. Réciproquement, supposons que $\overline{B_E(0,1)}$ soit compacte. On raisonne par l'absurde en supposant que E soit de dimension infinie. Puisque,

$$\overline{B_E(0,1)} \subset \bigcup_{x \in E} B_E\left(x, \frac{1}{2}\right)$$

la compacité de $\overline{B_E(0,1)}$ assure l'existence de $x_1, \ldots, x_n \in E$ tels que,

$$\overline{B_E(0,1)} \subset \bigcup_{k=1}^{n} B_E\left(x_k, \frac{1}{2}\right).$$

Notons ensuite $F = \mathrm{Vect}(x_1, \ldots, x_n)$. Puisque E est de dimension infinie, on peut considérer $x \in E \setminus F$. Notons $d = d(x, F)$. Alors, $d > 0$ car F est fermé dans E (en tant que sous-espace vectoriel de dimension finie de E) et $x \notin F$. Soit $y \in F$ tel que $\|x - y\| \leq d + \frac{d}{2} = \frac{3d}{2}$. Notons,

$$x_0 = \frac{1}{\|x - y\|}(x - y).$$

Puisque $x_0 \in \overline{B_E(0,1)}$, il existe $k \in [\![1, n]\!]$ tel que $x_0 \in B_E(x_k, 1)$, c'est-à-dire,

$$\|x_0 - x_k\| < \frac{1}{2}.$$

En multipliant cette inégalité par $\|x - y\|$, l'homogénéité de la norme,

$$\|x - (\underbrace{y + \|x - y\|x_k}_{\in F})\| < \frac{1}{2}\|x - y\| = \frac{3d}{4} < d$$

ce qui contredit $d = d(x, F)$. Finalement, on en conclut que E est de dimension finie. □

Commentaires personnels. Développement technique, mais qui se comprend et se retient facilement. De plus, l'équivalence des normes en dimension finie est un résultat très important, donc je pense qu'il est préférable de savoir le prouver.

Espace de Bergman

Notons $\mathbb{D} = \{z \in \mathbb{C},\ |z| < 1\}$. On note λ la mesure de Lebesgue sur $\mathbb{C}$ identifié à $\mathbb{R}^2$, $L^2(\mathbb{D})$ l'espace $L^2_{\mathbb{C}}(\mathbb{D}, \mathcal{B}(\mathbb{D}), \lambda)$, $H(\mathbb{D})$ l'espace des fonctions holomorphes sur $\mathbb{D}$ (identifié à son quotient par la relation d'égalité presque-partout).

Définition 1. On appelle espace de Bergman de $\mathbb{D}$, et on note $B(\mathbb{D})$, l'espace $H(\mathbb{D}) \cap L^2(\mathbb{D})$. Il s'agit d'un sous-espace vectoriel de $L^2(\mathbb{D})$.

[BB] **Lemme 2.** Soient $f \in B(\mathbb{D})$ et K un compact inclus dans $\mathbb{D}$. Alors,

$$\max_{z\in K}|f(z)| \leq \frac{1}{\sqrt{\pi}d(K,\mathbb{U})}\|f\|_2.$$

où $d(K, \mathbb{U})$ désigne la distance du compact K au cercle unité $\mathbb{U}$.

Preuve. Soit $z_0 \in K$. Fixons alors $r \in]0, 1-|z_0|[$ de telle sorte que $\overline{D(z_0, r)} \subset \mathbb{D}$. D'après la formule de la moyenne (corollaire immédiat de celle de Cauchy),

$$\forall \rho \in [0, r],\ f(z_0) = \frac{1}{2\pi}\int_0^{2\pi} f(z_0 + \rho e^{i\theta})d\theta,$$

car f est holomorphe sur $\mathbb{D}$. Par un changement de variable en coordonnées polaires, il vient alors,

$$\int_{D(z_0,r)} f(\omega)d\lambda(\omega) = \int_0^r \left(\int_0^{2\pi} f(z_0 + \rho e^{i\theta})d\theta\right)\rho d\rho = \pi r^2 f(z_0).$$

Or, d'après l'inégalité de Cauchy-Schwarz,

$$\left|\int_{D(z_0,r)} f(\omega)d\lambda(\omega)\right| = |(f \mid 1_{D(z_0,r)})_2| \leq \|1_{D(z_0,r)}\|_2\|f\|_2$$

Sachant que $\|1_{D(z_0,r)}\|_2^2 = \lambda(D(z_0, r)) = \pi r^2$, il s'en suit que,

$$|f(z_0)| \leq \frac{1}{\sqrt{\pi}r}\|f\|_2.$$

D'où, en faisant tendre r vers $1 - |z_0|$,

$$|f(z_0)| \leq \frac{1}{\sqrt{\pi}(1-|z_0|)}\|f\|_2.$$

Finalement, on en conclut que,

$$\max_{z\in K}|f(z)| \leq \frac{1}{\sqrt{\pi}d(K,\mathbb{U})}\|f\|_2$$

car $d(K, \mathbb{U}) \leq d(z_0, \mathbb{U}) = 1 - |z_0|$. □

Théorème 3. Pour tout $n \in \mathbb{N}$, on note $e_n : \mathbb{D} \to \mathbb{C},\ z \mapsto \sqrt{\frac{n+1}{\pi}}z^n$. Alors, [BB]

(i) $(B(\mathbb{D}), (\cdot \mid \cdot)_2)$ est un espace de Hilbert;

(ii) $(e_n)_{n\in\mathbb{N}}$ est une base hilbertienne de $(B(\mathbb{D}), (\cdot \mid \cdot)_2)$;

(iii) toute application f de $B(\mathbb{D})$ s'écrit :

$$f = \int_{\mathbb{D}} f(\omega)K(\cdot, \omega)d\lambda(\omega),$$

où, l'on a posé, pour tout $(z, \omega) \in \mathbb{D}^2$, $K(z, \omega) = \frac{1}{\pi(1-z\overline{\omega})^2}$.

Preuve. Pour le point (i), il suffit de montrer $(B(\mathbb{D}), \|\ \|_2)$ est complet. Soit $(f_n)_{n\in\mathbb{N}}$ une suite de Cauchy dans $(B(\mathbb{D}), \|\cdot\|_2)$. D'après le lemme précédent, pour tout compact K contenu dans $\mathbb{D}$, et pour tous $p, q \in \mathbb{N}$,

$$\|f_p - f_q\|_{\infty,K} \leq \frac{1}{\sqrt{\pi}d(K,\mathbb{U})}\|f_p - f_q\|_2.$$

Par suite, pour tout $\varepsilon > 0$, il existe $N \in \mathbb{N}$ tel que, pour tous $p, q \geq N$,

$$\|f_p - f_q\|_{\infty,K} \leq \varepsilon$$

D'après le théorème de Weierstrass, il existe donc $f \in H(\mathbb{D})$ telle que $(f_n)_{n\in\mathbb{N}}$ converge uniformément vers f sur tout compact contenu dans $\mathbb{D}$. D'autre part, puisque la suite $(f_n)_{n\in\mathbb{N}}$ est de Cauchy dans $(L^2(\mathbb{D}), \|\cdot\|_2)$ qui est complet, elle converge vers une fonction $g \in L^2(\mathbb{D})$ dans cet espace, c'est-à-dire,

$$\|f_n - g\|_2 \underset{n\to+\infty}{\longrightarrow} 0.$$

En particulier, il existe une suite $(f_{\varphi(n)})_{n\in\mathbb{N}}$, extraite de $(f_n)_{n\in\mathbb{N}}$, qui converge λ-presque partout vers g sur $\mathbb{D}$. Or, $(f_{\varphi(n)})_{n\in\mathbb{N}}$ converge simplement vers f sur $\mathbb{D}$ (puisque $(f_n)_{n\in\mathbb{N}}$ converge uniformément vers f sur tout compact inclus dans $\mathbb{D}$). Donc $g = f$ λ-presque partout, ce qui implique $f \in L^2(\mathbb{D})$. Ainsi, $f \in H(\mathbb{D}) \cap L^2(\mathbb{D}) = B(\mathbb{D})$. Sachant que la suite $(f_n)_{n\in\mathbb{N}}$ converge vers f dans $(L^2(\mathbb{D}), \|\cdot\|_2)$, elle converge donc vers f dans $(B(\mathbb{D}), \|\cdot\|_2)$.

Passons ensuite au point (ii). Soient $n, m \in \mathbb{N}$. De nouveau par un changement de variable en coordonnées polaires, il vient, pour tout $z \in \mathbb{D}$,

$$\int_{\mathbb{D}} z^n \overline{z^m} d\lambda(z) = \int_0^1 \left(\int_0^{2\pi} \rho^{n+m} e^{i(n-m)\theta} \rho d\theta \right) d\rho = \frac{1}{n+m+2} 2\pi \delta_{n,m}.$$

En particulier,

$$\int_{\mathbb{D}} z^n \overline{z^n} d\lambda(z) = \frac{1}{n+1}\pi.$$

ce qui prouve que $(e_n)_{n\in\mathbb{N}}$ est une famille orthonormée de $(B(\mathbb{D}), \|\cdot\|_2)$. Pour obtenir son caractère Hilbertien, il suffit désormais de montrer que,

$$\{e_n,\ n \in \mathbb{N}\}^\perp = \{0\}.$$

Soit $f \in B(\mathbb{D})$. Puisque f est holomorphe sur $\mathbb{D}$, il existe une suite de nombres complexes $(a_n)_{n\in\mathbb{N}}$ telle que, pour tout $z \in \mathbb{D}$, on ait,

$$f(z) = \sum_{k=0}^{+\infty} a_k z^k.$$

De plus, cette suite vérifie, pour tout $r \in [0, 1[$, et pour tout $k \in \mathbb{N}$,

$$a_k = \frac{1}{2\pi r^k} \int_0^{2\pi} f(re^{i\theta}) e^{-ik\theta} d\theta.$$

Soit $n \in \mathbb{N}$. Alors,

$$\begin{aligned}
(f \mid e_n)_2 &= \sqrt{\frac{n+1}{\pi}} \int_{\mathbb{D}} f(z)\overline{z^n} d\lambda(z) \\
&= \sqrt{\frac{n+1}{\pi}} \int_0^1 \left(\int_0^{2\pi} f(re^{i\theta}) \overline{(re^{i\theta})^n} d\theta \right) r dr \\
&= \sqrt{\frac{n+1}{\pi}} \int_0^1 (2\pi r^n a_n) r^{n+1} dr \\
&= \sqrt{\frac{n+1}{\pi}} \frac{2\pi}{2n+2} \\
(f \mid e_n)_2 &= \sqrt{\frac{\pi}{n+1}} a_n.
\end{aligned}$$

Ainsi, pour tout $n \in \mathbb{N}$

$$(f \mid e_n)_2 = 2\sqrt{\frac{\pi}{n+1}} a_n.$$

Par conséquent, si $f \in \{e_n,\ n \in \mathbb{N}\}^\perp$ alors, pour tout $n \in \mathbb{N}$, $a_n = 0$ et donc $f = 0$. D'où, $\{e_n,\ n \in \mathbb{N}\}^\perp = \{0\}$, ce qui prouve finalement que $(e_n)_{n\in\mathbb{N}}$ est une base hilbertienne de $B(\mathbb{D})$. □

Point (iii) À faire.

Compléments conseillés.

- Formules de changement de variable (coordonnées polaires) : théorème 7.29 [GH].
- Topologie de la convergence uniforme sur tout compact : chapitre 12, page 155 [PT].

Commentaires personnels. Développement long et très technique, nécessitant de nombreux prérequis. Malgré cela, il s'agit d'un de mes développements favoris. Il rentre dans de nombreuses leçons, mais ce sont des leçons où j'avais déjà beaucoup de développements. Donc ce n'était peut-être pas un choix si judicieux que ça.

Expression des $\zeta(2k)$

[FG] **Proposition 1.** Pour tout $z \in D(0,2\pi)$ non nul,

$$\frac{z}{e^z-1} = 1 - \frac{z}{2} + 2\sum_{k=1}^{+\infty} \frac{(-1)^{k-1}}{(2\pi)^{2k}} \zeta(2k) z^{2k}.$$

Preuve. Fixons $z \in D(0,2\pi)$ non nul, et considérons la fonction 2π-périodique φ définie pour tout $x \in]-\pi,\pi]$ par,

$$\varphi(x) = \exp\left(\frac{zx}{2\pi}\right).$$

Cette fonction étant continue par morceaux sur $\mathbb{R}$, on peut donc calculer ses coefficients de Fourier. Soit $n \in \mathbb{Z}$. Alors,

$$c_n(\varphi) = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{\frac{zt}{2\pi}} e^{-int} dt = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{(\frac{z}{2\pi}-in)t} dt.$$

Or, $z \neq 2i\pi n$ car $|z| < 2\pi$ et $z \neq 0$. D'où,

$$\int_{-\pi}^{\pi} e^{(\frac{z}{2\pi}-in)t} dt = \frac{2\pi}{z-2i\pi n}\left[e^{(\frac{z}{2\pi}-in)t}\right]_{-\pi}^{\pi} = \frac{2\pi}{z-2i\pi n}\left(e^{\frac{z}{2}-in\pi} - e^{-\frac{z}{2}+in\pi}\right).$$

Et donc,

$$c_n(\varphi) = \frac{(-1)^n}{z-2i\pi n}(e^{\frac{z}{2}} - e^{-\frac{z}{2}}),$$

car $e^{\frac{z}{2}-in\pi} = e^{\frac{z}{2}}(e^{-i\pi})^n = (-1)^n e^{\frac{z}{2}}$ et $e^{-\frac{z}{2}+in\pi} = e^{-\frac{z}{2}}(e^{i\pi})^n = (-1)^n e^{-\frac{z}{2}}$. Or, la fonction φ étant de classe C^1 par morceaux, le théorème de Dirichlet assure la convergence simple de sa série de Fourier sur $\mathbb{R}$ vers la fonction,

$$x \mapsto \frac{1}{2}(\varphi(x^+) + \varphi(x^-)).$$

où l'on a noté, pour tout $x_0 \in \mathbb{R}$,

$$\varphi(x_0^+) = \lim_{\substack{x\to x_0 \\ x>x_0}} \varphi(x),\ \varphi(x_0^-) = \lim_{\substack{x\to x_0 \\ x<x_0}} \varphi(x).$$

Autrement dit, pour tout $x \in \mathbb{R}$,

$$\frac{1}{2}(\varphi(x^+) + \varphi(x^-)) = (e^{\frac{z}{2}} - e^{-\frac{z}{2}}) \sum_{n=-\infty}^{+\infty} \frac{(-1)^n}{z-2i\pi n} e^{inx}.$$

En particulier, pour $x = \pi$,

$$\frac{1}{2}(e^{\frac{z}{2}} + e^{-\frac{z}{2}}) = (e^{\frac{z}{2}} - e^{-\frac{z}{2}}) \sum_{n=-\infty}^{+\infty} \frac{1}{z-2i\pi n},$$

c'est-à-dire, en divisant par $e^{\frac{z}{2}} - e^{-\frac{z}{2}}$ (qui est non nul car $z \notin 2i\pi\mathbb{Z}$),

$$\frac{e^{\frac{z}{2}} + e^{-\frac{z}{2}}}{2(e^{\frac{z}{2}} - e^{-\frac{z}{2}})} = \sum_{n=-\infty}^{+\infty} \frac{1}{z-2i\pi n}. \tag{1}$$

Transformons cette somme pour qu'elle soit indicée sur $\mathbb{N}$. Tout d'abord,

$$\begin{aligned}
\sum_{n=-\infty}^{+\infty} \frac{1}{z-2i\pi n} &= \sum_{n=-\infty}^{+\infty} \frac{z+2i\pi n}{z^2+4\pi^2n^2} \\
&= \frac{1}{z} + \sum_{n=1}^{+\infty}\left[\frac{z+2i\pi n}{z^2+4\pi^2n^2} + \frac{z+2i\pi(-n)}{z^2+4\pi^2(-n)^2}\right] \\
\sum_{n=-\infty}^{+\infty} \frac{1}{z-2i\pi n} &= \frac{1}{z} + 2\sum_{n=1}^{+\infty} \frac{z}{z^2+4\pi^2n^2}.
\end{aligned}$$

Donc,

$$\sum_{n=-\infty}^{+\infty} \frac{1}{z-2i\pi n} = \frac{1}{z} + 2\sum_{n=1}^{+\infty} \frac{z}{z^2+4\pi^2n^2}. \tag{2}$$

D'autre part, en multipliant par $e^{\frac{z}{2}}$ au numérateur et au dénominateur,

$$\frac{e^{\frac{z}{2}} + e^{-\frac{z}{2}}}{2(e^{\frac{z}{2}} - e^{-\frac{z}{2}})} = \frac{e^z+1}{2(e^z-1)} = \frac{e^z-1+2}{2(e^z-1)} = \frac{e^z-1}{2(e^z-1)} + \frac{2}{2(e^z-1)}.$$

Et donc,

$$\frac{e^{\frac{z}{2}} + e^{-\frac{z}{2}}}{2(e^{\frac{z}{2}} - e^{-\frac{z}{2}})} = \frac{1}{2} + \frac{1}{e^z-1}. \tag{3}$$

Ainsi, en combinant les égalités (1), (2) et (3), il vient,

$$\frac{z}{e^z-1} = 1 - \frac{z}{2} + 2\sum_{n=1}^{+\infty} \frac{z^2}{z^2+4\pi^2n^2}. \tag{4}$$

Pour obtenir l'égalité souhaitée, il reste à développer en série entière le terme général de la somme ci-dessus. Soit $n \in \mathbb{N}^*$. Alors,

$$\frac{z^2}{z^2+4\pi^2n^2} = \left(\frac{z}{2\pi n}\right)^2 \frac{1}{(\frac{z}{2\pi n})^2+1} = \left(\frac{z}{2\pi n}\right)^2 \sum_{k=0}^{+\infty} (-1)^k \left(\frac{z}{2\pi n}\right)^{2k}.$$

Puis,

$$\frac{z^2}{z^2+4\pi^2n^2}=\sum_{k=0}^{+\infty}(-1)^k\left(\frac{z}{2\pi n}\right)^{2(k+1)}=\sum_{k=1}^{+\infty}(-1)^{k+1}\left(\frac{z}{2\pi n}\right)^{2k}.$$

Et donc,

$$\sum_{n=1}^{+\infty}\frac{z^2}{z^2+4\pi^2n^2}=\sum_{n=1}^{+\infty}\sum_{k=1}^{+\infty}(-1)^{k+1}\left(\frac{z}{2\pi n}\right)^{2k}.$$

À présent, nous allons utiliser le théorème de Fubini pour permuter ces deux sommes. Tout d'abord, d'après le théorème de Fubini-Tonelli,

$$\sum_{n=1}^{+\infty}\sum_{k=1}^{+\infty}\left|(-1)^k\left(\frac{z}{2\pi n}\right)^{2k}\right|=\sum_{n=1}^{+\infty}\sum_{k=1}^{+\infty}\left(\frac{|z|}{2\pi n}\right)^{2k}=\sum_{n=1}^{+\infty}\left(\frac{|z|^2}{4\pi^2n^2}\frac{1}{\frac{|z|^2}{4\pi^2n^2}+1}\right),$$

et donc,

$$\sum_{n=1}^{+\infty}\sum_{k=1}^{+\infty}\left|(-1)^k\left(\frac{z}{2\pi n}\right)^{2k}\right|=\sum_{n=1}^{+\infty}\frac{|z|^2}{|z|^2+4\pi^2n^2}<+\infty$$

car $\frac{|z|^2}{|z|^2+4\pi^2n^2}\underset{n\to+\infty}{\sim}\frac{1}{4\pi^2n^2}$. Ainsi, d'après le théorème de Fubini-Lebesgue,

$$\begin{aligned}\sum_{n=1}^{+\infty}\frac{z^2}{z^2+4\pi^2n^2}&=\sum_{n=1}^{+\infty}\sum_{k=1}^{+\infty}(-1)^{k+1}\left(\frac{z}{2\pi n}\right)^{2k}\\&=\sum_{k=1}^{+\infty}\frac{(-1)^{k+1}}{(2\pi)^{2k}}\left(\sum_{n=1}^{+\infty}\frac{1}{n^{2k}}\right)z^{2k}\\\sum_{n=1}^{+\infty}\frac{z^2}{z^2+4\pi^2n^2}&=\sum_{k=1}^{+\infty}\frac{(-1)^{k+1}}{(2\pi)^{2k}}\zeta(2k)z^{2k}.\end{aligned}$$

Finalement, en utilisant l'égalité (4), on en conclut que,

$$\frac{z}{e^z-1}=1-\frac{z}{2}+2\sum_{k=1}^{+\infty}\frac{(-1)^{k-1}}{(2\pi)^{2k}}\zeta(2k)z^{2k}.$$

□

[FG] **Théorème 2 (Calcul des $\zeta(2k)$).** Notons $(b_n)_{n\in\mathbb{N}}$ la suite définie par $b_0=1$ et, pour tout $n\in\mathbb{N}^*$, $b_n=-\frac{1}{n+1}\sum_{k=0}^{n-1}\binom{n+1}{k}b_k$. Pour tout $k\in\mathbb{N}^*$,

$$\zeta(2k)=(-1)^{k+1}\frac{2^{2k-1}b_{2k}}{(2k)!}\pi^{2k}.$$

Preuve. Notons f la somme de la série entière,

$$1-\frac{z}{2}+2\sum_{k=1}^{+\infty}\frac{(-1)^{k-1}}{(2\pi)^{2k}}\zeta(2k)z^{2k}$$

sur $D(0,2\pi)$. Pour tout $n\in\mathbb{N}$, notons a_n le n-ième coefficient de cette série entière, multiplié par $n!$. En particulier, pour tout $k\in\mathbb{N}^*$,

$$2\frac{(-1)^{k-1}}{(2\pi)^{2k}}\zeta(2k)=\frac{a_{2k}}{(2k)!}$$

c'est-à-dire,

$$\zeta(2k)=(-1)^{k+1}\frac{2^{2k-1}a_{2k}}{(2k)!}\pi^{2k}.$$

Pour conclure, il suffit donc de montrer que $(a_n)_{n\in\mathbb{N}}=(b_n)_{n\in\mathbb{N}}$. Tout d'abord,

$$\forall z\in D(0,2\pi),\ f(z)=\sum_{n=0}^{+\infty}\frac{a_n}{n!}z^n.$$

Puis, pour tout $z\in D(0,2\pi)$, on réalise le produit de Cauchy suivant,

$$z=f(z)(e^z-1)=\left(\sum_{n=0}^{+\infty}\frac{a_n}{n!}z^n\right)\left(\sum_{n=1}^{+\infty}\frac{z^n}{n!}\right)=\sum_{n=1}^{+\infty}\left(\sum_{k=0}^{n-1}\frac{a_k}{k!(n-k)!}\right)z^n.$$

Par identification des coefficients de ces séries entières, il s'en suit que,

$$\forall n\geq 2,\ \sum_{k=0}^{n-1}\frac{a_k}{k!(n-k)!}=0.$$

c'est-à-dire, en changeant d'indice, et en réarrangeant les termes,

$$\forall n\in\mathbb{N}^*,\ a_n=-\sum_{k=0}^{n-1}\frac{n!}{k!(n+1-k)!}a_k==-\frac{1}{n+1}\sum_{k=0}^{n-1}\frac{(n+1)!}{k!(n+1-k)!}a_k,$$

Ainsi,

$$\forall n\in\mathbb{N}^*,\ a_n=-\frac{1}{n+1}\sum_{k=0}^{n-1}\binom{n+1}{k}a_k.$$

Sachant que $a_0=1=b_0$, on en déduit que $(a_n)_{n\in\mathbb{N}}=(b_n)_{n\in\mathbb{N}}$, ce qui achève la preuve. □

Commentaires personnels. Développement très calculatoire. Je trouvais le résultat joli, mais cette preuve est infâme, et inintéressante de mon point de vue. Étant extrêmement nul en calculs, j'ai dû le répéter un très grand nombre de fois avant de la maîtriser. Et puisque je n'arrivais pas à la faire en quinze minutes, j'étais obligé de court-circuiter certaines étapes de calcul pour finir dans les temps. Si comme moi vous détestez les calculs, fuyez ce développement.

Formule de Stirling (par les intégrales de Wallis)

⚠ Les preuves ci-dessous sont différentes de celles de [JR].

[JR] **Proposition 1 (Intégrales de Wallis).** Soit $(I_n)_{n\in\mathbb{N}}$ la suite définie pour tout $n\in\mathbb{N}$ par, $I_n=\int_0^{\frac{\pi}{2}}\sin^n(t)dt$. Pour tout $n\in\mathbb{N}$, $I_{n+2}=\frac{n+1}{n+2}I_n$. De plus,

(i) $I_n \underset{n\to+\infty}{\sim} \sqrt{\frac{\pi}{2n}}$;

(ii) pour tout $n\in\mathbb{N}$, $I_{2n}=\frac{1}{2^{2n}}\binom{2n}{n}\frac{\pi}{2}$.

Preuve. Fixons $n\in\mathbb{N}$. Par intégration par parties, il vient,

$$\begin{aligned}
I_{n+2} &= \int_0^{\frac{\pi}{2}}\sin(t)\sin^{n+1}(t)dt\\
&= \left[-\cos(t)\sin^{n+1}(t)\right]_0^{\frac{\pi}{2}}+(n+1)\int_0^{\frac{\pi}{2}}\cos^2(t)\sin^n(t)dt\\
&= (n+1)\int_0^{\frac{\pi}{2}}(1-\sin^2(t))\sin^n(t)dt\\
&= (n+1)\left(\int_0^{\frac{\pi}{2}}\sin^n(t)dt-\int_0^{\frac{\pi}{2}}\sin^{n+2}(t)dt\right)\\
I_{n+2} &= (n+1)(I_n-I_{n+2}).
\end{aligned}$$

D'où, la formule de récurrence $I_{n+2}=\frac{n+1}{n+2}I_n$. D'autre part,

$$\forall t\in\left[0,\frac{\pi}{2}\right],\ \sin^{n+1}(t)\leq\sin^n(t)$$

d'où $I_{n+1}\leq I_n$ en intégrant cette inégalité sur $[0,\frac{\pi}{2}]$. La suite $(I_n)_{n\in\mathbb{N}}$ est donc décroissante. De plus, ses termes sont strictement positifs puisque,

$$\forall n\in\mathbb{N},\ I_n\geq\int_{\frac{\pi}{6}}^{2\pi}\sin^n(t)dt\geq\int_{\frac{\pi}{6}}^{2\pi}\frac{1}{2^n}dt=\frac{1}{2^n}\frac{11}{6}\pi>0.$$

On peut également dire que $t\mapsto\sin^n(t)$ étant continue et strictement positive sur $[\frac{\pi}{6},2\pi]$, son intégrale sur ce segment est donc strictement positive. Avec ce qui précède, il vient ensuite,

$$\forall n\in\mathbb{N},\ 0<\frac{n+1}{n+2}=\frac{I_{n+2}}{I_n}\leq\frac{I_{n+1}}{I_n}\leq 1$$

ce qui prouve que $I_{n+1}\underset{n\to+\infty}{\sim}I_n$ par encadrement de limites. Or,

$$\forall n\in\mathbb{N},\ (n+1)I_{n+1}I_n=(n+1)\frac{n}{n+1}I_{n-1}I_n=nI_nI_{n-1}.$$

Donc, la suite $((n+1)I_{n+1}I_n)_{n\in\mathbb{N}}$ est constante égale à I_1I_0, c'est-à-dire $\frac{\pi}{2}$. De ce fait, on en déduit que,

$$nI_n^2\underset{n\to+\infty}{\sim}(n+1)I_{n+1}I_n=\frac{\pi}{2}$$

c'est-à-dire $I_n\underset{n\to+\infty}{\sim}\sqrt{\frac{\pi}{2n}}$. D'où le point (i).

Montrons ensuite le point (ii) par récurrence. Tout d'abord, on a bien $I_0=\frac{\pi}{2}$. Supposons ensuite le résultat établi pour un rang $n\in\mathbb{N}$ fixé. Alors,

$$I_{2n+2}=\frac{2n+1}{2n+2}I_{2n}=\frac{(2n+1)(2n+2)}{(2n+2)^2}\frac{1}{2^{2n}}\frac{(2n)!}{(n!)^2}\frac{\pi}{2}=\frac{1}{2^{2n+2}}\frac{(2n+2)!}{[(n+1)!]^2}\frac{\pi}{2},$$

et donc $I_{2n+2}=\frac{1}{2^{2n+2}}\binom{2n+2}{n+1}\frac{\pi}{2}$. Par récurrence, il s'en suit le point (ii). □

Théorème 2 (Formule de Stirling). On a $n!\underset{n\to+\infty}{\sim}\sqrt{2\pi n}\left(\frac{n}{e}\right)^n$. [JR]

Preuve. Soit $(u_n)_{n\in\mathbb{N}^*}$ la suite définie pour tout $n\in\mathbb{N}$ par,

$$u_n=\frac{\sqrt{n}}{n!}\left(\frac{n}{e}\right)^n>0.$$

Montrons que la suite $(u_n)_{n\in\mathbb{N}}$ converge. Pour tout $n\in\mathbb{N}^*$,

$$\frac{u_{n+1}}{u_n}=\frac{\sqrt{n+1}}{(n+1)!}\frac{(n+1)^{(n+1)}}{e^{n+1}}\frac{n!}{\sqrt{n}}\frac{e^n}{n^n}=\frac{1}{e}\left(1+\frac{1}{n}\right)^{n+\frac{1}{2}}.$$

Et donc, pour tout $n\in\mathbb{N}$,

$$\ln\left(\frac{u_{n+1}}{u_n}\right)=\left(n+\frac{1}{2}\right)\ln\left(1+\frac{1}{n}\right)-1.$$

Or, le développement limité usuel du logarithme assure l'existence d'une suite $(\varphi(n))_{n\in\mathbb{N}^*}$ bornée telle que, pour tout $n\in\mathbb{N}^*$, on ait,

$$\ln\left(1+\frac{1}{n}\right)=\frac{1}{n}-\frac{1}{2n^2}+\frac{\varphi(n)}{n^3}.$$

D'où, en utilisant les deux égalités précédentes,

$$\ln\left(\frac{u_{n+1}}{u_n}\right)=1-\frac{1}{2n}+\frac{\varphi(n)}{n^2}+\frac{1}{2n}-\frac{1}{4n^2}+\frac{\varphi(n)}{2n^3}=\frac{\psi(n)}{n^2}.$$

où l'on a noté, pour tout $n \in \mathbb{N}^*$, $\psi(n) = \varphi(n) + \frac{\varphi(n)}{2n} - \frac{1}{4}$. De plus, la suite $(\psi(n))_{n\in\mathbb{N}^*}$ est bornée car, pour tout $n \in \mathbb{N}^*$,

$$|\psi(n)| \leq 2|\varphi(n)| + \frac{1}{4} \leq 2\sup_{\mathbb{N}}|\varphi| + \frac{1}{4} < +\infty.$$

Des points précédents, on en déduit que,

$$|\ln(u_{n+1}) - \ln(u_n)| = \left|\ln\left(\frac{u_{n+1}}{u_n}\right)\right| \underset{n\to+\infty}{=} O\left(\frac{1}{n^2}\right).$$

Par comparaison de séries à termes positifs, on en déduit que la série de terme général $|\ln(u_{n+1}) - \ln(u_n)|$ converge, ce qui implique que la série,

$$\sum_{n\in\mathbb{N}^*} (\ln(u_{n+1}) - \ln(u_n))$$

converge, ce qui équivaut à la convergence de la suite $(\ln(u_n))_{n\in\mathbb{N}^*}$. Notons l la limite de cette suite. Par continuité de la fonction exponentielle, il vient,

$$u_n = \exp(\ln(u_n)) \underset{n\to+\infty}{\longrightarrow} e^l.$$

Autrement dit,

$$n! \underset{n\to+\infty}{\sim} C\sqrt{n}\left(\frac{n}{e}\right)^n$$

où l'on a noté $C = e^l$. En particulier, avec le point (ii) de la proposition 1,

$$I_{2n} = \frac{1}{2^{2n}}\frac{(2n)!}{(n!)^2}\frac{\pi}{2} \underset{n\to+\infty}{\sim} \frac{1}{2^{2n}}C\sqrt{2n}\left(\frac{2n}{e}\right)^{2n}\frac{1}{C^2 n}\left(\frac{e}{n}\right)^{2n}\frac{\pi}{2} = \frac{1}{C}\sqrt{\frac{2}{n}}\frac{\pi}{2}.$$

Or, d'après le point (i) de la proposition 1,

$$I_{2n} \underset{n\to+\infty}{\sim} \frac{1}{2}\sqrt{\frac{\pi}{n}}.$$

Donc,

$$\frac{1}{C}\sqrt{\frac{2}{n}}\frac{\pi}{2} \underset{n\to+\infty}{\sim} \frac{1}{2}\sqrt{\frac{\pi}{n}},$$

ce qui implique $C = \sqrt{2\pi}$. Finalement, on en conclut que,

$$n! \underset{n\to+\infty}{\sim} \sqrt{2\pi n}\left(\frac{n}{e}\right)^n.$$

□

Commentaires personnels. Développement assez calculatoire, mais les arguments théoriques utilisés sont élémentaires (c'est niveau licence). Étant extrêmement nul en calculs, j'ai dû le répéter un grand nombre de fois avant de le maîtriser. Malheureusement, il est difficile de couvrir toutes les leçons sans aucun développement calculatoire.

Formule de Stirling (par le théorème central limite)

[LL] **Théorème 1 (Formule de Stirling).** On a $n! \underset{n\to+\infty}{\sim} \sqrt{2\pi n}\left(\frac{n}{e}\right)^n$.

Preuve. Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires i.i.d (indépendantes et identiquement distribuées), suivant une loi de Poisson de paramètre 1. Notons ensuite, pour tout $n \in \mathbb{N}^*$,

$$S_n = \sum_{k=1}^{n} X_k.$$

Notons également, pour tout $n \in \mathbb{N}^*$,

$$Z_n = \frac{S_n - n}{\sqrt{n}}.$$

Étape n°1. On fixe une variable aléatoire Z suivant une loi normale centrée réduite. Puisque,

(i) les variables aléatoires de la suite $(X_n)_{n\in\mathbb{N}^*}$ sont i.i.d;

(ii) $X_1 \in L^2_{\mathbb{N}}(\Omega, \mathcal{A}, \mathbb{P})$;

(iii) $\mathbb{E}[X_1] = 1$ et $\mathbb{V}[X_1] = 1$;

le théorème central limite assure que $(Z_n)_{n\in\mathbb{N}^*}$ converge en loi vers Z. Or, F_Z est continue. Donc, $(F_{Z_n})_{n\in\mathbb{N}}$ converge simplement vers F_Z sur $\mathbb{R}$.

Étape n°2. Nous allons ensuite appliquer le théorème de convergence dominée à la suite de fonctions $(t \mapsto \mathbb{P}(Z_n > t))_{n\in\mathbb{N}^*}$.

(i) Pour tout $n \in \mathbb{N}^*$, $t \mapsto \mathbb{P}(Z_n > t)$ est mesurable.

(ii) Pour tout $t \in \mathbb{R}_+$, la suite $(\mathbb{P}(Z_n > t))_{n\in\mathbb{N}^*} = (1 - F_{Z_n}(t))_{n\in\mathbb{N}^*}$ converge vers $\mathbb{P}(Z > t) = 1 - F_Z(t)$ (d'après l'étape n°1).

(iii) Soient $n \in \mathbb{N}^*$ et $t \in \mathbb{R}_+$. Tout d'abord,

$$\mathbb{P}(Z_n > t) \leq \mathbb{P}(Z_n^2 > t^2)$$

car $\{Z_n > t\} \subset \{Z_n^2 > t^2\}$. Or, d'après l'inégalité de Markov,

$$\mathbb{P}(Z_n^2 > t^2) \leq \frac{\mathbb{E}[Z_n^2]}{t^2} = \frac{1}{t^2}$$

En effet, S_n suivant une loi de poisson de paramètre n, on a,

$$\mathbb{E}[Z_n] = \frac{1}{\sqrt{n}}(\mathbb{E}[S_n] - n) = 0,\ \mathbb{V}[Z_n] = \frac{1}{(\sqrt{n})^2}\mathbb{V}[S_n] = 1,$$

et donc $\mathbb{E}[Z_n^2] = \mathbb{V}[Z_n] + (\mathbb{E}[Z_n])^2 = 1$. D'où,

$$\mathbb{P}(Z_n > t) \leq \mathbb{P}(Z_n^2 > t^2) \leq \frac{\mathbb{E}[Z_n^2]}{t^2} = \frac{1}{t^2}.$$

Ainsi, en notant $\varphi : \mathbb{R}_+ \to \mathbb{R},\ t \mapsto 1_{[0,1[}(t) + \frac{1}{t^2}1_{[1,+\infty[}(t)$, il vient, pour tout $n \in \mathbb{N}^*$ et pour tout $t \in \mathbb{R}_+$,

$$\mathbb{P}(Z_n > t) \leq \varphi(t)$$

avec φ qui est intégrable sur $\mathbb{R}_+$.

D'après le théorème de convergence dominée, il s'en suit que,

$$\int_0^{+\infty} \mathbb{P}(Z_n > t)dt \underset{n\to+\infty}{\longrightarrow} \int_0^{+\infty} \mathbb{P}(Z > t)dt.$$

Étape n°3. Calculons $\int_0^{+\infty} \mathbb{P}(Z > t)dt$. Tout d'abord,

$$\int_0^{+\infty} \mathbb{P}(Z > t)dt = \frac{1}{\sqrt{2\pi}} \int_0^{+\infty} \left(\int_0^{+\infty} \exp\left(-\frac{u^2}{2}\right) 1_{u\geq t} du \right) dt.$$

Puis, d'après le théorème de Fubini-Tonelli,

$$\begin{aligned}
\int_0^{+\infty} \mathbb{P}(Z > t)dt &= \frac{1}{\sqrt{2\pi}} \int_0^{+\infty} \left(\int_0^{+\infty} \exp\left(-\frac{u^2}{2}\right) 1_{u\geq t} dt \right) du \\
&= \frac{1}{\sqrt{2\pi}} \int_0^{+\infty} \exp\left(-\frac{u^2}{2}\right) \left(\int_0^{u} dt \right) du \\
&= \frac{1}{\sqrt{2\pi}} \int_0^{+\infty} u \exp\left(-\frac{u^2}{2}\right) du \\
&= \frac{1}{\sqrt{2\pi}} \lim_{A\to+\infty} \left[-e^{-\frac{u^2}{2}} \right]_0^A \\
\int_0^{+\infty} \mathbb{P}(Z > t)dt &= \frac{1}{\sqrt{2\pi}} \lim_{A\to+\infty} \left[1 - e^{-\frac{A^2}{2}} \right].
\end{aligned}$$

Et donc,

$$\int_0^{+\infty} \mathbb{P}(Z > t)dt = \frac{1}{\sqrt{2\pi}}.$$

Étape n°4. Soit $n \in \mathbb{N}^*$. Calculons $\int_0^{+\infty} \mathbb{P}(Z_n > t)dt$. Tout d'abord,

$$\{Z_n > t\} = \{S_n > \sqrt{n}t + n\}.$$

Donc,

$$\int_0^{+\infty} \mathbb{P}(Z_n > t)dt = \int_0^{+\infty} \mathbb{P}(S_n > \sqrt{n}t + n)dt.$$

Or, puisque S_n est une variable aléatoire discrète, on a,

$$\mathbb{P}(S_n > \sqrt{n}t + n) = \sum_{k=0}^{+\infty} \mathbb{P}(S_n = k)1_{k>\sqrt{n}t+n}.$$

D'où, par le théorème de Fubini-Tonelli,

$$\begin{aligned}
\int_0^{+\infty} \mathbb{P}(Z_n > t)dt &= \int_0^{+\infty} \left(\sum_{k=0}^{+\infty} \mathbb{P}(S_n = k)1_{k>\sqrt{n}t+n} \right) dt \\
&= \sum_{k=0}^{+\infty} \left[\mathbb{P}(S_n = k) \int_0^{+\infty} 1_{t<\frac{k-n}{\sqrt{n}}} dt \right] \\
\int_0^{+\infty} \mathbb{P}(Z_n > t)dt &= \sum_{k=n+1}^{+\infty} \frac{k-n}{\sqrt{n}} \mathbb{P}(S_n = k).
\end{aligned}$$

Et puisque S_n suit une loi de Poisson de paramètre n, il vient,

$$\begin{aligned}
\int_0^{+\infty} \mathbb{P}(Z_n > t)dt &= \frac{e^{-n}}{\sqrt{n}} \sum_{k=n+1}^{+\infty} (k-n)\frac{n^k}{k!} \\
&= \frac{e^{-n}}{\sqrt{n}} \left[\sum_{k=n+1}^{+\infty} k\frac{n^k}{k!} - \sum_{k=n+1}^{+\infty} n\frac{n^k}{k!} \right] \\
&= \frac{e^{-n}}{\sqrt{n}} n \left[\sum_{k=n+1}^{+\infty} \frac{n^{k-1}}{(k-1)!} - \sum_{k=n+1}^{+\infty} \frac{n^k}{k!} \right] \\
&= e^{-n}\sqrt{n} \left[\sum_{k=n}^{+\infty} \frac{n^k}{k!} - \sum_{k=n+1}^{+\infty} \frac{n^k}{k!} \right] \\
\int_0^{+\infty} \mathbb{P}(Z_n > t)dt &= e^{-n}\sqrt{n}\frac{n^n}{n!}.
\end{aligned}$$

Ainsi, pour tout $n \in \mathbb{N}^*$,

$$\int_0^{+\infty} \mathbb{P}(Z_n > t)dt = \frac{\sqrt{n}}{n!} \left(\frac{n}{e}\right)^n.$$

Finalement, des étapes n°2,3 et 4, on en déduit que,

$$\frac{\sqrt{n}}{n!} \left(\frac{n}{e}\right)^n \underset{n\to+\infty}{\longrightarrow} \int_0^{+\infty} \frac{1}{\sqrt{2\pi}}$$

c'est-à-dire que,

$$n! \underset{n\to+\infty}{\sim} \sqrt{2\pi n} \left(\frac{n}{e}\right)^n.$$

□

Remarque. Si X est une variable aléatoire réelle alors,

$$\int_0^{+\infty} \mathbb{P}(X > t)dt = \mathbb{E}[X1_{X\geq 0}].$$

En effet, pour tout $t \in \mathbb{R}_+$,

$$\mathbb{P}(X > t) = \int_\Omega 1_{X(\omega)\geq 0} d\mathbb{P}(\omega),$$

puis,

$$\int_0^{+\infty} \mathbb{P}(X > t)dt = \int_0^{+\infty} \left(\int_\Omega 1_{X(\omega)\geq t} d\mathbb{P}(\omega) \right) dt,$$

et donc, d'après le théorème de Fubini-Tonelli,

$$\int_0^{+\infty} \mathbb{P}(X > t)dt = \int_\Omega \left(\int_0^{+\infty} 1_{X(\omega)\geq t} dt \right) d\mathbb{P}(\omega).$$

Or,

$$\int_0^{+\infty} 1_{X(\omega)\geq t} dt = X(\omega)1_{X(\omega)\geq 0}$$

D'où,

$$\int_0^{+\infty} \mathbb{P}(X > t)dt = \int_\Omega X(\omega)1_{X(\omega)\geq 0} d\mathbb{P}(\omega) = \mathbb{E}[X1_{X\geq 0}].$$

Compléments conseillés.

- Convergence en loi : théorème 3.
- Variables aléatoires mutuellement indépendantes et identiquement distribuées : lemme 5.

Commentaires personnels. Développement assez calculatoire, mais utilisant aussi beaucoup de résultats théoriques (théorème central limite, caractérisation de la convergence en loi par les fonctions de répartition, théorème de convergence dominée et théorème de Fubini). Étant extrêmement nul en calculs, j'ai dû le répéter un grand nombre de fois avant de le maîtriser. Je suis passé sur ce développement lors de mon oral d'analyse et probabilités.

Méthode de Laplace

[FR] **Théorème 1.** Soient $a < b \leq \infty$, $\varphi : [a, b[\to \mathbb{R}$ une fonction de classe C^2 avec $\varphi' > 0$ sur $]a, b[$ et $\varphi'(a) \geq 0$. Soit $f : [a, b[\to \mathbb{C}$ une fonction continue en a telle que $f(a) \neq 0$. On suppose qu'il existe $t_0 \in \mathbb{R}$ tel que $e^{-t_0\varphi} f \in L^1(]a, b[)$. Alors, pour tout $t \geq t_0$, l'intégrale,

$$F(t) = \int_a^b e^{-t\varphi(x)} f(x) dx$$

est bien définie. De plus,

(i) si $\varphi'(a) > 0$, alors $F(t) \underset{t\to+\infty}{\sim} \dfrac{1}{\varphi'(a)} \dfrac{e^{-t\varphi(a)} f(a)}{t}$,

(ii) si $\varphi'(a) = 0$ et $\varphi''(a) > 0$, alors $F(t) \underset{t\to+\infty}{\sim} \sqrt{\dfrac{\pi}{2\varphi''(a)}} \dfrac{e^{-t\varphi(a)} f(a)}{\sqrt{t}}$.

Preuve. Soit $t \geq t_0$. Alors, pour tout $x \in [a, b[$,

$$|e^{-t\varphi(x)} f(x)| = e^{-(t-t_0)\varphi(x)} e^{-t_0\varphi(x)} |f(x)| \leq e^{-(t-t_0)\varphi(a)} e^{-t_0\varphi(x)} |f(x)|$$

car φ est croissante (et même strictement croissante) sur $[a, b[$. D'où,

$$\int_a^b |e^{-t\varphi(x)} f(x)| dx \leq e^{-(t-t_0)\varphi(a)} \int_a^b e^{-t_0\varphi(x)} |f(x)| < +\infty$$

car $e^{-t_0\varphi} f \in L^1(]a, b[)$. Ainsi, F est bien définie sur $[t_0, +\infty[$. D'autre part, f est bornée au voisinage de 0. En effet, la continuité de f en 0 assure l'existence de $\alpha > 0$ tel que, pour tout $x \in [0, \alpha]$, on ait $|f(x) - f(0)| \leq 1$ et donc, par la seconde inégalité triangulaire,

$$\forall x \in [0, \alpha],\ |f(x)| \leq M$$

où l'on a noté $M = 1 + |f(0)|$.

Étape n°1. Montrons tout d'abord le point (i) sous les hypothèses $a = 0$ et, pour tout $x \in [0, b[$, $\varphi(x) = x$. Pour exploiter la majoration précédente, nous allons découper l'intégrale étudiée en deux morceaux :

$$\forall t \geq t_0,\ \int_0^b te^{-tx} f(x) dx = \int_0^\alpha te^{-tx} f(x) dx + \int_\alpha^b te^{-tx} f(x) dx.$$

Montrons que,

$$\lim_{t\to+\infty} \int_0^\alpha te^{-tx} f(x) dx = f(0), \quad \lim_{t\to+\infty} \int_\alpha^b te^{-tx} f(x) dx = 0.$$

Soit $t \geq t_0$. Tout d'abord, le changement de variable $u = tx$ donne,

$$\int_0^\alpha te^{-tx} f(x) dx = \int_0^{\alpha t} te^{-u} f\left(\frac{u}{t}\right) \frac{1}{t} du = \int_0^{+\infty} e^{-u} f\left(\frac{u}{t}\right) 1_{[0,\alpha t]}(u) du.$$

Nous allons ensuite appliquer le théorème de convergence dominée à l'intégrale ci-dessus.

- Pour tout $t \geq t_0$, la fonction $u \mapsto e^{-u} f\left(\dfrac{u}{t}\right) 1_{[0,\alpha t]}(u)$ est mesurable.
- Soit $u \in \mathbb{R}_+$. Par continuité de f en 0,

$$\lim_{t\to+\infty} e^{-u} f\left(\frac{u}{t}\right) 1_{[0,\alpha t]}(u) = f(0) e^{-u}.$$

- Pour tout $(t, u) \in [t_0, +\infty[\times\mathbb{R}_+$,

$$\left| e^{-u} f\left(\frac{u}{t}\right) 1_{[0,\alpha t]}(u) \right| \leq e^{-u} M$$

avec $u \mapsto Me^{-u}$ qui est bien intégrable sur $\mathbb{R}_+$.

Ainsi, par le théorème de convergence dominée (à paramètre),

$$\lim_{t\to+\infty} \int_0^{+\infty} e^{-u} f\left(\frac{u}{t}\right) 1_{[0,\alpha t]}(u) du = f(0) \int_0^{+\infty} e^{-u} du = f(0).$$

Et donc,

$$\lim_{t\to+\infty} \int_0^\alpha te^{-tx} f(x) dx = f(0).$$

Passons au deuxième morceau de l'intégrale. Soit $t \geq t_0$.

$$\forall x \in [\alpha, b[,\ |e^{-tx} f(x)| = e^{-(t-t_0)x} e^{-t_0 x} |f(x)| \leq e^{-(t-t_0)\alpha} e^{-t_0 x} |f(x)|$$

et donc,

$$\left| \int_\alpha^b te^{-tx} f(x) dx \right| \leq te^{-(t-t_0)\alpha} \int_\alpha^b e^{-t_0 x} |f(x)| dx \underset{t\to+\infty}{\longrightarrow} 0$$

car $te^{-(t-t_0)\alpha} \underset{t\to+\infty}{\longrightarrow} 0$ (par croissance comparée) et,

$$\int_\alpha^b e^{-t_0 x} |f(x)| dx < +\infty.$$

Finalement,

$$tF(t) = \int_0^\alpha te^{-tx}f(x)dx + \int_\alpha^b te^{-tx}f(x)dx \underset{t\to+\infty}{\longrightarrow} f(0)$$

c'est-à-dire $F(t) \underset{t\to+\infty}{\sim} \dfrac{f(0)}{t}$ (car $f(0) \neq 0$).

Étape n°2. À présent, démontrons le point (i) sous l'hypothèse de l'énoncé, c'est-à-dire $\varphi'(a) > 0$. Pour cela, on se ramène au cas précédent en utilisant un changement de variable. On note,

$$c = \sup_{x\in]a,b[} \varphi(x) - \varphi(a) \in \mathbb{R} \cup \{+\infty\}.$$

Puis, on considère l'application,

$$\begin{array}{rccc} \psi & : & [a,b[& \longrightarrow & [0,c[\\ & & x & \longmapsto & \varphi(x) - \varphi(a) \end{array}$$

Tout d'abord, d'après le théorème des valeurs intermédiaires, $\psi([a,b[) = [0,c[$. De plus, ψ est de classe C^1 avec $\psi' = \varphi' > 0$ sur $[a,b[$, et donc ψ est bijective et sa bijection réciproque, notée ψ^{-1}, est de classe C^1 sur $[0,c[$. En particulier, l'application $]a,b[\to]0,c[,\ x \mapsto \psi(x)$ est un C^1-difféomorphisme. On peut donc effectuer le changement de variable $y = \psi(x)$, ce qui donne,

$$F(t) = e^{-t\varphi(a)} \int_0^c e^{-ty}g(y)dy,$$

où l'on a noté, pour tout $y \in [0,c[$, $g(y) = f(\psi^{-1}(y))(\psi^{-1})'(y)$. Or,

- g est continue sur $[0,c[$, car $f \in C^0([a,b[)$ et $\psi^{-1} \in C^1([0,c[)$;
- $g(0) = f(\psi^{-1}(0))(\psi^{-1})'(0) = f(a)\dfrac{1}{(\psi'\circ\psi^{-1})(0)} = \dfrac{f(a)}{\varphi'(a)} \neq 0$;
- $e^{-t_0 y}g \in L^1(]0,c[)$ (par le théorème de changement de variable).

On peut donc appliquer les résultats de l'étape n°1 à g, ce qui donne,

$$F(t) = e^{-t\varphi(a)} \int_0^c e^{-ty}g(y)dy \underset{t\to+\infty}{\sim} e^{-t\varphi(a)}\frac{g(0)}{t} = \frac{1}{\varphi'(a)}\frac{e^{-t\varphi(a)}f(a)}{t}.$$

D'où le point (i).

Étape n°3. Montrons tout d'abord le point (ii) sous les hypothèses $a = 0$ et, pour tout $x \in [0,b[$, $\varphi(x) = x^2$. Nous allons de nouveau découper l'intégrale étudiée en deux morceaux. Montrons que,

$$\lim_{t\to+\infty} \int_0^\alpha \sqrt{t}e^{-tx^2}f(x)dx = \frac{\sqrt{\pi}}{2}f(0), \quad \lim_{t\to+\infty} \int_\alpha^b \sqrt{t}e^{-tx^2}f(x)dx = 0.$$

Soit $t \geq t_0$. Tout d'abord, le changement de variable $u = x\sqrt{t}$ donne,

$$\int_0^\alpha \sqrt{t}e^{-tx^2}f(x)dx = \int_0^{\alpha\sqrt{t}} \sqrt{t}e^{-u^2}f\left(\frac{u}{\sqrt{t}}\right)\frac{1}{\sqrt{t}}du$$

et donc,

$$\int_0^\alpha \sqrt{t}e^{-tx^2}f(x)dx = \int_0^{+\infty} e^{-u^2}f\left(\frac{u}{\sqrt{t}}\right)1_{[0,\alpha\sqrt{t}]}(u)du.$$

Nous allons ensuite appliquer le théorème de convergence dominée à l'intégrale ci-dessus.

- Pour tout $t \geq t_0$, la fonction $u \mapsto e^{-u^2}f\left(\dfrac{u}{\sqrt{t}}\right)1_{[0,\alpha\sqrt{t}]}(u)$ est mesurable.
- Soit $u \in \mathbb{R}_+$. Par continuité de f en 0,

$$\lim_{t\to+\infty} e^{-u^2}f\left(\frac{u}{\sqrt{t}}\right)1_{[0,\alpha\sqrt{t}]}(u) = f(0)e^{-u^2}.$$

- Pour tout $(t,u) \in [t_0,+\infty[\times\mathbb{R}_+$,

$$\left|e^{-u^2}f\left(\frac{u}{\sqrt{t}}\right)1_{[0,\alpha\sqrt{t}]}(u)\right| \leq e^{-u^2}M$$

avec $u \mapsto Me^{-u^2}$ qui est bien intégrable sur $\mathbb{R}_+$.

Ainsi, par le théorème de convergence dominée (à paramètre),

$$\lim_{t\to+\infty} \int_0^{+\infty} e^{-u^2}f\left(\frac{u}{\sqrt{t}}\right)1_{[0,\alpha\sqrt{t}]}(u)du = f(0)\int_0^{+\infty} e^{-u^2}du = \frac{\sqrt{\pi}}{2}f(0).$$

Et donc,

$$\lim_{t\to+\infty} \int_0^\alpha \sqrt{t}e^{-tx^2}f(x)dx = \frac{\sqrt{\pi}}{2}f(0).$$

Passons au deuxième morceau de l'intégrale. Soit $t \geq t_0$.

$$\forall x \in [\alpha,b[,\ |e^{-tx^2}f(x)| = e^{-(t-t_0)x^2}e^{-t_0x^2}|f(x)| \leq e^{-(t-t_0)\alpha^2}e^{-t_0x^2}|f(x)|$$

et donc,

$$\left|\int_\alpha^b \sqrt{t}e^{-tx^2}f(x)dx\right| \leq \sqrt{t}e^{-(t-t_0)\alpha^2}\int_\alpha^b e^{-t_0x^2}|f(x)|dx \underset{t\to+\infty}{\longrightarrow} 0$$

car $\sqrt{t}e^{-(t-t_0)\alpha^2} \underset{t\to+\infty}{\longrightarrow} 0$ (par croissance comparée) et,

$$\int_\alpha^b e^{-t_0x^2}|f(x)|dx < +\infty.$$

Finalement,

$$\sqrt{t}F(t) = \int_0^\alpha \sqrt{t}e^{-tx^2}f(x)dx + \int_\alpha^b \sqrt{t}e^{-tx^2}f(x)dx \underset{t\to+\infty}{\longrightarrow} \frac{\sqrt{\pi}}{2}f(0).$$

c'est-à-dire $F(t) \underset{t\to+\infty}{\sim} \frac{\sqrt{\pi}}{2}\frac{f(0)}{\sqrt{t}}$ (car $f(0) \neq 0$).

Étape n°4. À présent, démontrons le point (ii) sous l'hypothèse de l'énoncé, c'est-à-dire $\varphi'(a) = 0$ et $\varphi''(a) > 0$. Pour cela, on se ramène au cas précédent en utilisant un changement de variable. On note,

$$d = \sup_{x\in]a,b[} \sqrt{\varphi(x) - \varphi(a)} \in \mathbb{R} \cup \{+\infty\}.$$

Puis, on considère l'application,

$$\begin{array}{rccc} \sigma & : & [a,b[& \longrightarrow & [0,d[\\ & & x & \longmapsto & \sqrt{\varphi(x) - \varphi(a)} \end{array}$$

Tout d'abord, d'après le théorème des valeurs intermédiaires, $\sigma([a,b[) = [0,c[$. De plus, σ est de classe C^1 sur $]a,b[$ avec,

$$\forall x \in]a,b[,\ \sigma'(x) = \frac{\varphi'(x)}{2\sqrt{\varphi(x) - \varphi(a)}} > 0.$$

Par suite, l'application $]a,b[\to]0,c[,\ x \mapsto \sigma(x)$ est un C^1-difféomorphisme. On peut donc effectuer le changement de variable $y = \sigma(x)$, ce qui donne,

$$F(t) = e^{-t\varphi(a)} \int_0^d e^{-ty^2} f(\sigma^{-1}(y))(\sigma^{-1})'(y)dy, \tag{1}$$

A priori, l'application $y \mapsto f(\sigma^{-1}(y))(\sigma^{-1})'(y)$ est seulement définie sur $]0,d[$. Pour appliquer les résultats de l'étape n°3, il faut donc vérifier qu'elle est bien définie sur $[0,d[$. Pour cela, on montre que l'application σ est de classe C^1 sur $[a,b[$. D'une part,

$$\frac{\varphi'(x)}{x-a} = \frac{\varphi'(x) - \varphi'(a)}{x-a} \underset{x\to a}{\longrightarrow} \varphi''(a).$$

Et donc $\varphi'(x) \underset{x\to a}{\sim} (x-a)\varphi''(a)$ (puisque $\varphi''(a) \neq 0$). D'autre part, d'après la formule de Taylor-Young, il existe une fonction $\varepsilon : [a,d[\to \mathbb{R}$ telle que,

$$\begin{cases} \forall x \in [a,d[,\ \varphi(x) = \varphi(a) + \varphi'(a)(x-a) + \dfrac{\varphi''(a)}{2}(x-a)^2 + \varepsilon(x)(x-a)^2, \\ \lim\limits_{x\to a} \varepsilon(x) = 0. \end{cases}$$

Sachant que $\varphi'(a) = 0$, il s'en suit que,

$$\frac{\varphi(x) - \varphi(a)}{(x-a)^2} = \frac{\varphi''(a)}{2} + \varepsilon(x) \underset{x\to a}{\longrightarrow} \frac{\varphi''(a)}{2}.$$

Et donc $\varphi(x) - \varphi(a) \underset{x\to a}{\sim} \frac{\varphi''(a)}{2}(x-a)^2$ (puisque $\varphi''(a) \neq 0$). Finalement,

$$\sigma'(x) = \frac{\varphi'(x)}{2\sqrt{\varphi(x) - \varphi(a)}} \underset{x\to a}{\sim} \frac{(x-a)\varphi''(a)}{2\sqrt{\frac{1}{2}\varphi''(a)(x-a)^2}} = \sqrt{\frac{\varphi''(a)}{2}}.$$

Ainsi,

- σ est continue sur $[a,b[$;
- σ est dérivable sur $]a,b[$;
- $\sigma'(x) \underset{x\to a}{\longrightarrow} \sqrt{\frac{\varphi''(a)}{2}}$.

D'après le théorème de la limite de la dérivée, on en déduit que σ est de classe C^1 sur $[a,b[$, avec $\sigma'(a) = \sqrt{\frac{\varphi''(a)}{2}}$. Sa dérivée est donc strictement positive sur cet intervalle, ce qui entraîne que σ^{-1} est de classe C^1 sur $[0,c[$. Par ailleurs, d'après l'égalité (1),

$$F(t) = e^{-t\varphi(a)} \int_0^d e^{-ty^2} h(y)dy,$$

où l'on a noté, pour tout $y \in [0,d[$, $h(y) = f(\sigma^{-1}(y))(\sigma^{-1})'(y)$. Or,

- h est continue sur $[0,d[$, car $f \in C^0([a,b[)$ et $\sigma^{-1} \in C^1([0,d[)$;
- $h(0) = f(\sigma^{-1}(0))(\sigma^{-1})'(0) = f(a)\dfrac{1}{(\sigma' \circ \sigma^{-1})(0)} = f(a)\sqrt{\dfrac{2}{\varphi''(a)}} \neq 0$;
- $e^{-t_0y^2}h \in L^1(]0,d[)$ (par le théorème de changement de variable).

On peut donc appliquer les résultats de l'étape n°1 à g, ce qui donne,

$$F(t) \underset{t\to+\infty}{\sim} e^{-t\varphi(a)}\frac{\sqrt{\pi}}{2}\frac{h(0)}{\sqrt{t}} = \sqrt{\frac{\pi}{2\varphi''(a)}}\frac{e^{-t\varphi(a)}f(a)}{\sqrt{t}}.$$

D'où le point (ii).

□

Remarque. Traditionnellement, l'énoncé du théorème de convergence dominée est restreint aux suites de fonctions. Cependant, la caractérisation séquentielle de la limite permet d'étendre ce résultat aux fonctions à paramètre.

Compléments conseillés.

- Formule de Stirling généralisée : théorème 1.
- Théorème de convergence dominée à paramètre : théorème 1.

Commentaires personnels. Développement très technique et très long si on écrit tous les détails. En revanche, il se recase dans de très nombreuses leçons. Pour être tout à fait honnête, je ne maîtrisais pas suffisamment ce développement pour l'exposer en seulement 15 minutes. Si c'est aussi votre cas, je vous conseille de prouver seulement le point (ii).

Méthode de Newton

[FR] **Théorème 1 (Méthode de Newton).** Soit $f : [c,d] \to \mathbb{R}$ une fonction de classe C^2 telle que $f(c) < 0 < f(d)$ et $f' > 0$ sur $[c,d]$. On considère la suite récurrente définie par $x_0 \in [c,d]$ et, pour tout $n \in \mathbb{N}$, $x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$ (sous réserve que $x_n \in [c,d]$). Alors, f s'annule en un unique point a, et,

(i) il existe $\alpha > 0$ tel que, pour tout $x_0 \in [a-\alpha, a+\alpha]$, la suite $(x_n)_{n\in\mathbb{N}}$ est bien définie et converge quadratiquement vers a;

(ii) si $f'' > 0$ sur $[a,d]$, alors, pour tout $x_0 \in]a,d]$, la suite $(x_n)_{n\in\mathbb{N}}$ est strictement décroissante et $x_{n+1} - a \underset{n\to+\infty}{\sim} \frac{f''(a)}{2f'(a)}(x_n - a)^2$.

Preuve. Tout d'abord, d'après le théorème des valeurs intermédiaires, la fonction f s'annule bien en au moins un point a de $[c,d]$, car f est continue sur cet intervalle avec $f(c) < 0 < f(d)$. De plus, ce point est bien unique, puisque f est strictement croissante. Considérons ensuite l'application,

$$\begin{array}{rccc} F & : & [c,d] & \longrightarrow & \mathbb{R} \\ & & x & \longmapsto & x - \dfrac{f(x)}{f'(x)} \end{array}$$

Alors, un point de $[c,d]$ est un zéro de f si et seulement si c'est un point fixe de F. D'où l'idée d'étudier la suite $(x_n)_{n\in\mathbb{N}}$, qui est la suite des itérés de x_0 par F. En effet, si $(x_n)_{n\in\mathbb{N}}$ converge vers un réel x_∞, alors,

$$x_{n+1} = F(x_n) \underset{n\to+\infty}{\longrightarrow} F(x_\infty)$$

par continuité de F, et donc $F(x_\infty) = x_\infty$ par unicité de la limite, c'est-à-dire que x_∞ est un point fixe de F, donc un zéro de f par ce qui précède. Ainsi, si $(x_n)_{n\in\mathbb{N}}$ converge, c'est nécessairement vers a.

Étape n°1. Montrons le point (i). Soit $x \in [c,d]$. Puisque f est de classe C^1 sur $[c,d]$ et deux fois dérivable sur $]c,d[$, la formule de Taylor-Lagrange assure l'existence de z strictement compris entre a et x tel que,

$$f(a) = f(x) + f'(x)(a-x) + \frac{f''(z)}{2}(a-x)^2.$$

Or, $f(a) = 0$. En divisant par $f'(x)$ (qui est non nul), puis en réarrangeant les termes de cette égalité, on obtient donc,

$$F(x) - a = \frac{f''(z)}{2f'(x)}(a-x)^2. \tag{1}$$

En particulier, en posant $C = \dfrac{\max_{[c,d]} |f''|}{2\min_{[c,d]} f'}$ (qui existe car f' et f'' sont continues sur $[c,d]$ qui est compact et f' est strictement positive sur $[c,d]$), il vient,

$$|F(x) - a| \leq C|a-x|^2.$$

Soit $\alpha > 0$ tel que $\alpha C < 1$. Alors, pour tout $x \in [a-\alpha, a+\alpha]$,

$$|F(x) - a| \leq C\alpha^2 \leq \alpha$$

et donc l'intervalle $[a-\alpha, a+\alpha]$ est stable par la fonction F. En particulier, si $x_0 \in [a-\alpha, a+\alpha]$, la suite $(x_n)_{n\in\mathbb{N}}$ est bien définie. De plus, pour tout $n \in \mathbb{N}$,

$$|x_{n+1} - a| \leq C|x_n - a|^2.$$

Et donc, par récurrence immédiate, on obtient que, pour tout $n \in \mathbb{N}$,

$$C|x_n - a| \leq (C\alpha)^{2^n}$$

ce qui prouve finalement que $(x_n)_{n\in\mathbb{N}}$ converge quadratiquement vers a.

Étape n°2. Prouvons le point (ii). Supposons que f'' soit strictement positive sur $[a,d]$. En reprenant l'égalité (1), on obtient donc, pour tout $x \in [a,d]$,

$$F(x) - a \geq 0.$$

D'autre part, pour tout $x \in]a,d]$,

$$F(x) = x - \frac{f(x)}{f'(x)} < x \leq d$$

car $f(x) > f(a) = 0$ et $f' > 0$ sur $[a,d]$. Ainsi, l'intervalle $[c,d]$ est stable par la fonction F. En particulier, si $x_0 \in]a,d]$, la suite $(x_n)_{n\in\mathbb{N}}$ est alors bien définie, et à valeurs dans $]a,d]$. De plus, la suite $(x_n)_{n\in\mathbb{N}}$ est décroissante, car pour tout $n \in \mathbb{N}$,

$$x_{n+1} = F(x_n) = x_n - \frac{f(x_n)}{f'(x_n)} < x_n$$

car $f(x_n) > f(a) = 0$ et $f' > 0$ sur $[a,d]$. Ainsi, la suite $(x_n)_{n\in\mathbb{N}}$ est décroissante et minorée. D'après le théorème de convergence monotone, elle converge donc vers un certain réel x_∞. Mais alors, par continuité de F,

$$x_{n+1} = F(x_n) \underset{n\to+\infty}{\longrightarrow} F(x_\infty).$$

Or, $(x_{n+1})_{n\in\mathbb{N}}$ converge aussi vers x_∞. Par unicité de la limite, on en déduit que $x_\infty = F(x_\infty)$, c'est-à-dire $f(x_\infty) = 0$, ce qui implique $x_\infty = a$ (car a est l'unique point où s'annule f). Ainsi, $(x_n)_{n\in\mathbb{N}}$ converge vers a.

Démontrons finalement l'équivalent recherché. Pour tout $n \in \mathbb{N}$, la formule de Taylor-Lagrange appliquée à f entre a et x_n assure l'existence de $z_n \in]a, x_n[$ tel que,

$$\frac{x_{n+1} - a}{(x_n - a)^2} = \frac{f''(z_n)}{2f'(x_n)}.$$

Par encadrement de limites, la suite $(z_n)_{n\in\mathbb{N}}$ converge vers a. D'où, en faisant tendre n vers l'infini dans l'égalité précédente,

$$\frac{x_{n+1} - a}{(x_n - a)^2} \underset{n\to+\infty}{\longrightarrow} \frac{f''(a)}{2f'(a)},$$

car f' et f'' sont continues au point a. Finalement, on en conclut que,

$$x_{n+1} - a \underset{n\to+\infty}{\sim} \frac{f''(a)}{2f'(a)}(x_n - a)^2.$$

□

Complément conseillé.

- Exemple. On fixe $y > 0$ et on prend $f(x) = x^2 - y$. Résoudre alors la relation de récurrence, et donner une estimation de l'erreur $x_n - a$ avec $a = \sqrt{y}$. Exercice 49 [FR].

Commentaires personnels. C'est un développement très (trop ?) classique. Mais il n'est pas nécessaire d'être original pour être admis au concours, donc ce n'est pas un inconvénient majeur. De mon point de vue, c'est un bon choix de développement, puisqu'il est accessible et rentre dans de nombreuses leçons. De plus, la méthode de Newton figure sur le programme du concours.

Nombres de Bell

[BB] **Proposition 1 (Nombres de Bell).** Pour tout $n \in \mathbb{N}^*$, notons B_n le nombre de partitions de l'ensemble $[\![1,n]\!]$. Notons $B_0 = 1$. Alors, pour tout $n \in \mathbb{N}^*$,

$$B_n = \frac{1}{e}\sum_{p=0}^{+\infty}\frac{p^n}{p!}.$$

Preuve. Pour tout $n \in \mathbb{N}^*$ et pour tout $k \in [\![0,n]\!]$, on note,

- $\mathcal{P}_k([\![1,n]\!])$ l'ensemble des parties à k éléments de l'ensemble $[\![1,n]\!]$,
- $\mathcal{Q}_{n+1,k}$ l'ensemble des partitions de l'ensemble $[\![1,n+1]\!]$ dont l'élément contenant $n+1$ est de cardinal $k+1$,
- $\mathcal{R}_n$ l'ensemble des partitions de l'ensemble $[\![1,n]\!]$.

Étape n°1. Soit $n \in \mathbb{N}^*$. Montrons que,

$$B_{n+1} = \sum_{k=0}^{n}\binom{n}{k}B_k.$$

Soit $k \in [\![0,n]\!]$. Tout élément P de $\mathcal{Q}_{n+1,k}$, s'écrit $P = \{A \cup \{n+1\}\} \cup B$ où A est une partie à k éléments de $[\![1,n]\!]$ et B est une partition de $[\![1,n]\!] \setminus A$. Plus précisément, en notant, pour toute partie finie E de $\mathbb{N}$,

$$\begin{array}{rccl}\phi_E & : & [\![1,\mathrm{card}(E)]\!] & \longrightarrow & E \\ & & k & \longmapsto & \begin{cases}\min(E) & \text{si } k=1,\\ \min(E\setminus\{\phi(1),\dots,\phi(k)\}) & \text{sinon.}\end{cases}\end{array}$$

c'est-à-dire la suite finie qui énumère les éléments de E dans l'ordre strictement croissant, on obtient que l'application,

$$\begin{array}{ccc}\mathcal{P}_k([\![1,n]\!])\times\mathcal{R}_{n-k} & \longrightarrow & \mathcal{Q}_{n+1,k} \\ (A,P) & \longmapsto & \{A\cup\{n+1\}\}\cup\phi_{[\![1,n]\!]\setminus A}(P)\end{array}$$

est bien définie et bijective. En particulier,

$$\mathrm{card}(\mathcal{Q}_{n+1,k}) = \mathrm{card}(\mathcal{P}_k([\![1,n]\!]))\,\mathrm{card}(\mathcal{R}_{n-k}) = \binom{n}{k}B_{n-k}.$$

Or, $\mathcal{R}_{n+1}$ est l'union disjointe des $\mathcal{Q}_{n+1,k}$ pour $k \in [\![0,n]\!]$. Donc,

$$B_{n+1} = \sum_{k=0}^{n}\binom{n}{k}B_{n-k} = \sum_{k=0}^{n}\binom{n}{n-k}B_k = \sum_{k=0}^{n}\binom{n}{k}B_k.$$

Étape n°2. Montrons par récurrence sur $n \in \mathbb{N}^*$ que $B_n \leq n!$. L'initialisation est claire, et si $B_n \leq n!$ pour un entier $n \in \mathbb{N}^*$ fixé alors,

$$B_{n+1} \leq \sum_{k=0}^{n}\binom{n}{k}k! \leq \sum_{k=0}^{n}n! = (n+1)!$$

en utilisant l'égalité précédente. *Autre preuve.* Soit $n \in \mathbb{N}^*$. Pour toute partie E de $[\![1,n]\!]$, notons γ_E le cycle $(\phi_E(1),\dots,\phi_E(\mathrm{card}(E)))$. Alors, l'application,

$$\begin{array}{ccc}\mathcal{R}_n & \longrightarrow & \mathfrak{S}_n \\ \{P_1,\dots,P_r\} & \longmapsto & \{\gamma_{P_1},\dots,\gamma_{P_r}\}\end{array}$$

est injective, et donc $B_n = \mathrm{card}(\mathcal{R}_n) \leq \mathrm{card}(\mathfrak{S}_n) \leq n!$.

Étape n°3. D'après l'étape n°2, la série entière $\sum_{n\in\mathbb{N}}\frac{B_n}{n!}t^n$ possède un rayon de convergence R supérieur ou égal à 1. Notons S sa somme. D'après le théorème de dérivation des séries entières, S est dérivable sur $]-R,R[$, avec,

$$\forall t \in]-R,R[,\ S'(t) = \sum_{n=1}^{+\infty}\frac{B_n}{(n-1)!}t^{n-1} = \sum_{n=0}^{+\infty}\frac{B_{n+1}}{n!}t^n.$$

Soit $t \in]-R,R[$. En utilisant l'étape n°1, l'égalité précédente fournit,

$$S'(t) = \sum_{n=0}^{+\infty}\sum_{k=0}^{n}\binom{n}{k}B_k\frac{1}{n!}t^n = \sum_{n=0}^{+\infty}\sum_{k=0}^{n}\frac{B_k}{k!}t^k\frac{t^{n-k}}{(n-k)!}.$$

On reconnaît alors le produit de Cauchy des séries entières $\sum_{n\in\mathbb{N}}\frac{B_n}{n!}t^n$ et $\sum_{n\in\mathbb{N}}\frac{t^n}{n!}$, dont les rayons de convergence sont respectivement R et $+\infty$. Donc,

$$S'(t) = \left(\sum_{n=0}^{+\infty}\frac{t^n}{n!}\right)\left(\sum_{n=0}^{+\infty}\frac{B_n}{n!}t^n\right) = \exp(t)S(t).$$

Étape n°4. D'après l'étape précédente, S est solution du problème de Cauchy,

$$\begin{cases}y'(t) = \exp(t)y(t),\ t \in]-R,R[\\ y(0) = 1.\end{cases}$$

Mais ce problème de Cauchy (linéaire) admet pour unique solution,

$$\begin{array}{ccc}]-R,R[& \longrightarrow & \mathbb{R} \\ t & \longmapsto & \frac{1}{e}\exp(\exp(t)) \end{array}$$

Donc, pour tout $t \in]-R,R[$, $S(t) = \frac{1}{e}\exp(\exp(t))$. Soit $t \in \mathbb{R}$. Alors,

$$\exp(\exp(t)) = \sum_{p=0}^{+\infty} \frac{\exp(t)^p}{p!} = \sum_{p=0}^{+\infty} \frac{\exp(pt)}{p!} = \sum_{p=0}^{+\infty}\sum_{n=0}^{+\infty} \frac{p^n t^n}{p!n!}.$$

En particulier, en remplaçant t par $|t|$, on obtient que la famille $(\frac{p^n t^n}{p!n!})_{(p,n)\in\mathbb{N}^2}$ est sommable. D'après le théorème de Fubini, il s'en suit que,

$$\exp(\exp(t)) = \sum_{n=0}^{+\infty} \left(\sum_{p=0}^{+\infty} \frac{p^n}{p!} \right) \frac{t^n}{n!}.$$

D'autre part, pour tout $t \in]-R,R[$,

$$\exp(\exp(t)) = eS(t) = e\sum_{n=0}^{+\infty} \frac{B_n}{n!} t^n.$$

Par unicité du développement en série entière de $\exp\circ\exp$, on en déduit que $R = +\infty$, $S = \frac{1}{e}\exp\circ\exp$ et, pour tout $n \in \mathbb{N}^*$

$$B_n = \frac{1}{e}\sum_{p=0}^{+\infty} \frac{p^n}{p!}.$$

□

Commentaires personnels. Développement très sympathique, qui se comprend et se retient facilement. Je recommande ce développement.

Projection sur un convexe fermé

Soit $(H, (\cdot \mid \cdot))$ un espace de Hilbert réel.

[GH] **Théorème 1 (Théorème de projection sur un convexe fermé).** Soit $C \subset H$ un convexe fermé non vide. Soit $x \in H$. Il existe un unique $y \in C$ tel que $\|x-y\| = d(x,C)$, que l'on note $P_C(x)$. De plus, $P_C(x)$ est caractérisé par la propriété : $P_C(x) \in C$ et pour tout $z \in C$, $\Re((x - P_C(x) \mid z - P_C(x))) \leq 0$. On appelle projeté orthogonal de x sur C le vecteur $P_C(x)$.

Preuve. Commençons par montrer l'existence et l'unicité d'un tel élément.

Existence. Notons $d = d(x,C)$. Soit $(y_n)_{n\in\mathbb{N}}$ une suite d'éléments de C telle que,

$$\lim_{n\to+\infty} \|x - y_n\| = d. \tag{1}$$

Remarquons qu'il en existe au moins une, d'après la caractérisation de la borne inférieure. Montrons que la suite $(y_n)_{n\in\mathbb{N}}$ est de Cauchy. Soit $\varepsilon > 0$. Soient $p, q \in \mathbb{N}$. D'après l'identité du parallélogramme,

$$2\|y_p - x\|^2 + 2\|y_q - x\|^2 = \|y_p - y_q\|^2 + \|y_p + y_q - 2x\|^2,$$

c'est-à-dire,

$$\|y_p - y_q\|^2 = 2\|y_p - x\|^2 + 2\|y_q - x\|^2 - 4\left\|\frac{y_p + y_q}{2} - x\right\|^2.$$

D'où, par définition de d,

$$\|y_p - y_q\|^2 \leq 2\|y_p - x\|^2 + 2\|y_q - x\|^2 - 4d^2$$

car $\frac{y_p+y_q}{2} \in C$, par convexité de C. Or, d'après (1), il existe $N \in \mathbb{N}$ tel que, pour tout $n \geq N$, on ait $\|y_n - x\|^2 \leq d^2 + \frac{\varepsilon}{4}$. Par suite, pour tous $p, q \geq N$,

$$\|y_p - y_q\|^2 \leq \varepsilon$$

ce qui prouve finalement que la suite $(y_n)_{n\in\mathbb{N}}$ est de Cauchy. Puisque C est complet, en tant que partie fermée de H qui est complet, on en déduit que $(y_n)_{n\in\mathbb{N}}$ converge vers un certain $y_\infty \in C$. Alors, par continuité de la norme,

$$\lim_{n\to+\infty} \|x - y_n\| = \|x - y_\infty\|.$$

D'où, en utilisant (1), $\|x - y_\infty\| = d$ par unicité de la limite.

Unicité. Soient $y, z \in C$ tels que $\|x-y\| = \|x-z\| = d$. Considérons alors $(u_n)_{n\in\mathbb{N}}$ la suite définie par,

$$u_n = \begin{cases} y & \text{si } n \text{ est pair,} \\ z & \text{sinon.} \end{cases}$$

Cette suite possède deux valeurs d'adhérence, qui sont y et z. D'autre part, elle vérifie,

$$\lim_{n\to+\infty} \|x - u_n\| = d,$$

et donc elle converge d'après l'étape précédente. A fortiori, $(u_n)_{n\in\mathbb{N}}$ possède une unique valeur d'adhérence, ce qui implique $y = z$.

Caractérisation. Montrons finalement que le vecteur $P_C(x)$ est caractérisé par la propriété $P_C(x) \in C$ et,

$$\forall z \in C,\ \Re((x - P_C(x) \mid z - P_C(x))) \leq 0.$$

Soit $z \in C$. Pour tout $t \in [0,1]$, notons $z_t = tz + (1-t)P_C(x)$, qui est un élément de C car C est convexe. Soit $t \in]0,1]$. Alors,

$$\|z_t - x\|^2 \geq \|P_c(x) - x\|^2. \tag{2}$$

Or,

$$\|z_t - x\|^2 = \|t(z - P_C(x)) + (P_C(x) - x)\|^2$$

puis, en développant avec le produit hermitien,

$$\|z_t - x\|^2 = t^2\|z - P_C(x)\|^2 + 2t\Re((z - P_C(x) \mid P_C(x) - x)) + \|P_C(x) - x\|^2.$$

D'après l'inégalité (2), il s'en suit que,

$$t^2\|z - P_C(x)\|^2 + 2t\Re((z - P_C(x) \mid P_C(x) - x)) \geq 0$$

et donc, en divisant par t qui est non nul,

$$t\|z - P_C(x)\|^2 + 2\Re((z - P_C(x) \mid P_C(x) - x)) \geq 0.$$

Cette inégalité étant vraie pour tout $t \in]0,1]$, on peut donc faire tendre t vers 0 dans celle-ci, ce qui donne,

$$\Re((z - P_C(x) \mid P_C(x) - x)) \geq 0,$$

c'est-à-dire,

$$\Re((z - P_C(x) \mid x - P_C(x))) \leq 0.$$

Donc, $P_C(x)$ vérifie bien cette propriété. Il reste à voir que c'est le seul élément de C à la vérifier. Soit $y \in C$ tel que,

$$\forall z \in C,\ \Re((x-y \mid z-y)) \leq 0.$$

Alors, pour tout $z \in C$,

$$\begin{aligned} \|x-z\|^2 &= \|x-y+y-z\|^2 \\ &= \|x-y\|^2 + 2\Re((x-y, y-z)) + \|y-z\|^2 \\ \|x-z\|^2 &\geq \|x-y\|^2 \end{aligned}$$

car $\Re((x-y, y-z)) = -\Re(x-y, z-y) \geq 0$. Donc, $\|x-y\| = d(x,C)$, ce qui implique $P_C(x) = y$ d'après le début de cette preuve. □

Remarque. La figure 10 illustre cette caractérisation du projeté orthogonal.

Figures. 10.

Compléments conseillés.

- Distance entre un point et une partie fermée d'un espace vectoriel normé de dimension finie : théorème 1.
- Théorème de représentation de Riesz dans un espace de Hilbert : théorème 1.

Commentaires personnels. Développement très sympathique, qui se comprend et se retient facilement, et qui se présente dans de nombreuses leçons. De plus, le théorème de projection sur un convexe fermé figure sur le programme du concours.

Prolongement de la fonction Gamma d'Euler

Notons $\mathcal{H}_+ = \{z \in \mathbb{C},\ \Re(z) > 0\}$.

Proposition 1. La fonction,

$$\begin{array}{rccc} \Gamma & : & \mathcal{H}_+ & \longrightarrow & \mathbb{C} \\ & & z & \longmapsto & \displaystyle\int_0^{+\infty} t^{z-1}e^{-t}dt \end{array}$$

est bien définie et holomorphe sur $\mathcal{H}_+$.

Preuve. Considérons la fonction,

$$\begin{array}{rccc} f & : & \mathcal{H}_+ \times \mathbb{R}_+^* & \longrightarrow & \mathbb{C} \\ & & (z,t) & \longmapsto & t^{z-1}e^{-t} \end{array}$$

Nous appliquer le théorème d'holomorphie sous le signe intégral à f.

(i) Pour tout $z \in \mathcal{H}_+$, la fonction $\mathbb{R}_+^* \to \mathbb{C},\ t \mapsto f(z,t)$ est mesurable.

(ii) Pour tout $t \in \mathbb{R}_+^*$, la fonction $\mathcal{H}_+ \to \mathbb{C},\ z \mapsto f(z,t)$ est holomorphe.

(iii) Soit K un compact de $\mathbb{C}$ inclus dans $\mathcal{H}_+$. Notons,

$$\alpha = \min_{z\in K} \Re(z),\ \beta = \max_{z\in K} \Re(z).$$

qui sont bien définis et strictement positifs, puisque $\Re$ est une fonction continue de K qui est compact vers $\mathbb{R}_+^*$. Soit $(z,t) \in K \times \mathbb{R}_+$. Alors,

$$f(z,t) = \exp((z-1)\ln(t) - t) = e^{i\Im(z)\ln(t)}e^{(\Re(z)-1)\ln(t)-t}.$$

En particulier, $|f(z,t)| = e^{(\Re(z)-1)\ln(t)-t}$, et donc,

$$|f(z,t)| \leq \begin{cases} e^{(\alpha-1)\ln(t)-t} = t^{\alpha-1}e^{-t} & \text{si } t \leq 1, \\ e^{(\beta-1)\ln(t)-t} = t^{\beta-1}e^{-t} & \text{si } t \geq 1. \end{cases}$$

(car la fonction ln est négative sur $]0,1]$ et positive sur $[1,+\infty[$). Ainsi,

$$|f(z,t)| \leq \varphi(t)$$

où l'on a posé,

$$\varphi(t) = \begin{cases} t^{\alpha-1}e^{-t} & \text{si } t \leq 1, \\ t^{\beta-1}e^{-t} & \text{si } t \geq 1. \end{cases}$$

Or, φ est intégrable sur $]0,1]$, car elle est positive avec $t^{\alpha-1}e^{-t} \underset{t\to 0}{\sim} t^{\alpha-1}$ et $\alpha - 1 > -1$, et φ est intégrable sur $[1,+\infty[$, car $t^{\beta-1}e^{-t} \underset{t\to+\infty}{=} o(\frac{1}{t^2})$. Finalement, φ est intégrable sur $]0,+\infty[$.

Ainsi, d'après le théorème d'holomorphie sous le signe intégral, la fonction Γ est bien définie et holomorphe sur $\mathcal{H}_+$. □

Théorème 2. La fonction Γ admet un unique prolongement holomorphe sur $\mathbb{C} \setminus (-\mathbb{N})$. [BM]

Preuve. On procède en plusieurs étapes.

Étape n°1. Montrons que,

$$\forall z \in \mathcal{H}_+,\ \Gamma(z) = \sum_{n=0}^{+\infty} \frac{(-1)^n}{n!(z+n)} + \int_1^{+\infty} t^{z-1}e^{-t}dt. \tag{1}$$

Soit $z \in \mathcal{H}_+$. En développant l'exponentielle en série entière, il vient,

$$\int_0^1 t^{z-1}e^{-t}dt = \int_0^1 \sum_{n=0}^{+\infty} \frac{(-1)^n}{n!}t^{n+z-1}dt.$$

Appliquons ensuite le théorème de Fubini pour permuter la somme et l'intégrale ci-dessus. Tout d'abord, d'après le théorème de Fubini-Tonelli,

$$\int_0^1 \sum_{n=0}^{+\infty} \left|\frac{(-1)^n}{n!}t^{n+z-1}\right| dt = \int_0^1 \sum_{n=0}^{+\infty} \frac{1}{n!}t^{n+\Re(z)-1}dt.$$

car, pour tout $\zeta \in \mathbb{C}$, $\left|t^\zeta\right| = \left|e^{i\Im(\zeta)\ln(t)}e^{\Re(\zeta)\ln(t)}\right| = t^{\Re(\zeta)}$. Par suite,

$$\int_0^1 \sum_{n=0}^{+\infty} \left|\frac{(-1)^n}{n!}t^{n+z-1}\right| dt = \int_0^1 t^{\Re(z)-1}e^t dt < +\infty$$

car $t^{\Re(z)-1}e^t \underset{t\to 0}{\sim} t^{\Re(z)-1}$ et $\Re(z)-1 > -1$. On peut donc appliquer le théorème de Fubini-Lebesgue, ce qui donne,

$$\int_0^1 \sum_{n=0}^{+\infty} \frac{(-1)^n}{n!}t^{n+z-1}dt = \sum_{n=0}^{+\infty} \frac{(-1)^n}{n!}\int_0^1 t^{n+z-1}dt = \sum_{n=0}^{+\infty} \frac{(-1)^n}{n!}\left[\frac{t^{n+z}}{n+z}\right]_0^1$$

et donc,

$$\int_0^1 \sum_{n=0}^{+\infty} \frac{(-1)^n}{n!}t^{n+z-1}dt = \sum_{n=0}^{+\infty} \frac{(-1)^n}{n!(n+z)}.$$

D'où l'égalité (1).

Étape n°2. Pour tout $n \in \mathbb{N}$, notons,

$$\begin{array}{rccc} g_n & : & \mathbb{C} \setminus \{-n\} & \longrightarrow & \mathbb{C} \\ & & z & \longmapsto & \dfrac{(-1)^n}{n!(n+z)} \end{array}$$

Montrons que la somme g de la série de fonctions $\sum_{n\in\mathbb{N}} g_n$ est holomorphe sur $\mathbb{C} \setminus (-\mathbb{N})$.

(i) Pour tout $n \in \mathbb{N}$, g_n est holomorphe sur $\mathbb{C} \setminus (-\mathbb{N})$.

(ii) Soit K un compact de $\mathbb{C}$ inclus dans $\mathbb{C} \setminus (-\mathbb{N})$. Notons,

$$N = \left\lfloor \max_{z\in K} |z| \right\rfloor + 1.$$

Pour tout $z \in K$, et pour tout $n \geq N+1$,

$$|g_n(z)| = \frac{1}{n!|n+z|} \leq \frac{1}{n!(n-N)}$$

puisque $|n+z| \geq n - |z| \geq n - N$. Or, la série $\sum_{n>N} \dfrac{1}{n!(n-N)}$ converge.

Donc, $\sum_{n\in\mathbb{N}} g_n$ converge normalement sur K.

Des points précédents, on en déduit que g est holomorphe sur $\mathbb{C} \setminus (-\mathbb{N})$.

Étape n°3. Par ailleurs, en appliquant de nouveau le théorème d'holomorphie sous le signe intégral à f, cette fois-ci sur $[1, +\infty[$, on obtient que l'application,

$$\begin{array}{rccc} h & : & \mathbb{C} & \longrightarrow & \mathbb{C} \\ & & z & \longmapsto & \displaystyle\int_1^{+\infty} t^{z-1}e^{-t}dt \end{array}$$

est holomorphe sur $\mathbb{C}$, et donc la fonction $g+h$ est holomorphe sur $\mathbb{C} \setminus (-\mathbb{N})$. Or, d'après l'égalité (1),

$$\forall z \in \mathcal{H}_+,\ \Gamma(z) = g(z) + h(z).$$

Ainsi, $g+h$ est un prolongement holomorphe de Γ sur $\mathbb{C}\setminus(-\mathbb{N})$. De plus, un tel prolongement est unique par le principe du prolongement analytique, puisque $\mathbb{C} \setminus (-\mathbb{N})$ est un ouvert connexe. □

Compléments conseillés.

- Étude de la fonction Gamma d'Euler sur $\mathbb{R}_+$: Analyse de X. Gourdon (ellipses). Chapitre 4, sujet d'étude 1.
- Formule de Gauss, formule des compléments, formule de duplication, formule de Weierstrass (voir ce document)

Commentaires personnels. Développement assez technique, exploitant de nombreux théorèmes d'analyse au programme : théorème d'holomorphie sous le signe intégral, holomorphie de la somme d'une série de fonctions holomorphes (en cas de convergence normale), principe du prolongement analytique. Je pense qu'il est important de bien citer toutes les hypothèses requises pour ces théorèmes (à l'oral). Il se recase dans de nombreuses leçons. Je recommande ce développement.

Série des inverses des nombres premiers

Soit $(p_n)_{n\in\mathbb{N}^*}$ la suite des nombres premiers rangés dans l'ordre croissant.

Théorème 1. Soit $\alpha > 1$. La suite $\left(\prod_{k=1}^{n} \frac{1}{1-p_k^{-\alpha}}\right)_{n\in\mathbb{N}^*}$ converge vers $\zeta(\alpha)$.

Preuve. Soit $\alpha > 1$. On procède en plusieurs étapes.

Étape n°1. Soit $\mathbb{P}$ la mesure de probabilité sur $(\mathbb{N}^*, \mathcal{P}(\mathbb{N}^*))$ telle que :

$$\forall n \in \mathbb{N}^*,\ \mathbb{P}(\{n\}) = \frac{1}{\zeta(\alpha)}\frac{1}{n^\alpha}.$$

Il s'agit de la mesure de probabilité sur $(\mathbb{N}^*, \mathcal{P}(\mathbb{N}^*))$ définie par,

$$\forall A \subset \mathbb{N}^*,\ \mathbb{P}(A) = \frac{1}{\zeta(\alpha)} \sum_{n\in A} \frac{1}{n^\alpha}.$$

(cf. proposition 1).

Remarque. La loi de probabilité $\mathbb{P}$ est appelée loi ζ de paramètre α.

Étape n°2. Pour tout $n \in \mathbb{N}^*$, notons $A_n = \bigcap_{k=1}^{n} (\mathbb{N}^* \setminus p_k\mathbb{N}^*)$. Montrons que,

(i) pour tout $m \in \mathbb{N}^*$, $\mathbb{P}(m\mathbb{N}^*) = \frac{1}{m^\alpha}$,

(ii) $\mathbb{P}\left(\bigcap_{n\in\mathbb{N}^*} A_n\right) = \frac{1}{\zeta(\alpha)}$,

(iii) les évènements de la famille $(p_n\mathbb{N}^*)_{n\in\mathbb{N}^*}$ sont mutuellement indépendants.

Tout d'abord, pour tout $m \in \mathbb{N}^*$,

$$\mathbb{P}(m\mathbb{N}^*) = \frac{1}{\zeta(\alpha)} \sum_{k=1}^{+\infty} \frac{1}{(km)^\alpha} = \frac{1}{m^\alpha}\frac{1}{\zeta(\alpha)} \sum_{k=1}^{+\infty} \frac{1}{k^\alpha} = \frac{1}{m^\alpha}.$$

D'où le point (i). Ensuite, on a,

$$\bigcap_{n\in\mathbb{N}^*} A_n = \bigcap_{n\in\mathbb{N}^*} (\mathbb{N}^* \setminus p_n\mathbb{N}^*) = \{m \in \mathbb{N}^*,\ \forall n \in \mathbb{N}^*,\ p_n \nmid m\} = \{1\}.$$

En particulier, en utilisant le point (i),

$$\mathbb{P}\left(\bigcap_{n\in\mathbb{N}^*} A_n\right) = \mathbb{P}(\{1\}) = \frac{1}{\zeta(\alpha)}. \qquad \text{[JR]}$$

D'où le point (ii). Enfin, si I est une partie finie de $\mathbb{N}^*$ alors,

$$\bigcap_{i\in I} p_i\mathbb{N}^* = (\operatorname*{ppcm}_{i\in I} p_i)\mathbb{N}^* = \left(\prod_{i\in I} p_i\right)\mathbb{N}^*$$

(la première égalité est vraie en toute généralité, et la seconde est vraie car les p_i sont des nombres premiers deux à deux distincts), et donc,

$$\mathbb{P}\left(\bigcap_{i\in I} p_i\mathbb{N}^*\right) = \prod_{i\in I} \frac{1}{p_i^\alpha} = \prod_{i\in I} \mathbb{P}(p_i\mathbb{N}^*)$$

en utilisant le point (i), et donc les évènements de la famille $(p_n\mathbb{N}^*)_{n\in\mathbb{N}^*}$ sont mutuellement indépendants.

Étape n°3. Montrons que le produit $\prod_{n\in\mathbb{N}^*} \frac{1}{1-p_n^{-\alpha}}$ est convergent avec,

$$\prod_{n=1}^{+\infty} \frac{1}{1-p_n^{-\alpha}} = \zeta(\alpha).$$

Puisque $(A_n)_{n\in\mathbb{N}^*}$ est une suite d'évènements décroissante (pour l'inclusion), le théorème de convergence monotone assure que,

$$\lim_{n\to+\infty} \mathbb{P}(A_n) = \mathbb{P}\left(\bigcap_{n\in\mathbb{N}^*} A_n\right) = \frac{1}{\zeta(\alpha)}.$$

Or, les évènements de la famille $(p_n\mathbb{N}^*)_{n\in\mathbb{N}^*}$ étant mutuellement indépendants, ceux de la famille $(\mathbb{N}^* \setminus p_n\mathbb{N}^*)_{n\in\mathbb{N}^*}$ le sont aussi. D'où, pour tout $n \in \mathbb{N}^*$,

$$\mathbb{P}(A_n) = \prod_{k=1}^{n} \mathbb{P}(\mathbb{N}^* \setminus p_n\mathbb{N}^*) = \prod_{k=1}^{n} (1 - \mathbb{P}(\mathbb{N}^* \setminus p_n\mathbb{N}^*)) = \prod_{k=1}^{n} (1 - p_k^{-\alpha}).$$

Finalement, on en conclut que,

$$\prod_{k=1}^{n} (1 - p_k^{-\alpha}) \underset{n\to+\infty}{\longrightarrow} \frac{1}{\zeta(\alpha)}$$

c'est-à-dire,

$$\prod_{k=1}^{n} \frac{1}{1-p_k^{-\alpha}} \underset{n\to+\infty}{\longrightarrow} \zeta(\alpha).$$

□

Pour démontrer que la série $\sum_{n\in\mathbb{N}^*} \frac{1}{p_n}$ est divergente, nous aurons besoin du fait que $\lim\limits_{\substack{\alpha\to 1\\ \alpha>1}} \zeta(\alpha) = +\infty$. Dans la proposition suivante, on énonce un résultat plus précis.

Proposition 2. On a $\zeta(\alpha) \underset{\substack{\alpha\to 1\\ \alpha>1}}{\sim} \frac{1}{\alpha-1}$. En particulier, $\lim\limits_{\substack{\alpha\to 1\\ \alpha>1}} \zeta(\alpha) = +\infty$.

Preuve. Nous allons effectuer une comparaison série-intégrale. Fixons $\alpha > 1$. Puisque l'application $t \mapsto \frac{1}{t^\alpha}$ est décroissante sur $[1, +\infty[$, on a,

$$\zeta(\alpha) - 1 = \sum_{k=2}^{+\infty} \frac{1}{k^\alpha} = \sum_{k=1}^{+\infty} \frac{1}{(k+1)^\alpha} \leq \int_1^{+\infty} \frac{dt}{t^\alpha} \leq \sum_{k=1}^{+\infty} \frac{1}{k^\alpha} = \zeta(\alpha).$$

Or, $\int_1^{+\infty} \frac{dt}{t^\alpha} = \frac{1}{\alpha-1}$. D'où,

$$\frac{1}{\alpha-1} \leq \zeta(\alpha) \leq \frac{1}{\alpha-1} + 1$$

puis, en multipliant par $\alpha - 1$,

$$1 \leq (\alpha-1)\zeta(\alpha) \leq 1 + \alpha - 1 = \alpha \underset{\substack{\alpha\to 1\\ \alpha>1}}{\longrightarrow} 1$$

et donc par encadrement de limites, $\lim\limits_{\substack{\alpha\to 1\\ \alpha>1}} (\alpha-1)\zeta(\alpha) = 1$, c'est-à-dire,

$$\zeta(\alpha) \underset{\substack{\alpha\to 1\\ \alpha>1}}{\sim} \frac{1}{\alpha-1}.$$

En particulier, $\lim\limits_{\substack{\alpha\to 1\\ \alpha>1}} \zeta(\alpha) = +\infty$. □

[JR] **Théorème 3.** La série $\sum_{n\in\mathbb{N}^*} \frac{1}{p_n}$ est divergente.

Preuve. Soit $\alpha > 1$. Dans le théorème 1, on a vu que,

$$\lim_{n\to+\infty} \prod_{k=1}^{n} \left(\frac{1}{1-p_k^{-\alpha}}\right) = \zeta(\alpha).$$

Par les propriétés de calcul et par continuité de ln, il s'en suit que,

$$\lim_{n\to+\infty} -\sum_{k=1}^{n} \ln(1-p_k^{-\alpha}) = \lim_{n\to+\infty} \ln\left(\prod_{k=1}^{n}\left(\frac{1}{1-p_k^{-\alpha}}\right)\right) = \ln(\zeta(\alpha)).$$

Ainsi, la série $\sum_{n\in\mathbb{N}^*} \ln(1-p_n^{-\alpha})$ est convergente avec,

$$-\sum_{n=1}^{+\infty} \ln(1-p_n^{-\alpha}) = \ln(\zeta(\alpha)). \tag{1}$$

D'autre part,

$$-\ln(1-p_n^{-1}) \underset{n\to+\infty}{\sim} \frac{1}{p_n}$$

et,

$$\forall n \in \mathbb{N}^*,\ -\ln(1-p_n^{-1}) \geq 0.$$

Ainsi, la série $-\sum_{n\in\mathbb{N}^*} \ln(1-p_n^{-1})$ est divergente si et seulement si $\sum_{n\in\mathbb{N}^*} \frac{1}{p_n}$ est divergente. Montrons que la première série est divergente. Soit $A > 0$. Alors, il existe $\alpha > 1$ tel que $\ln(\zeta(\alpha)) \geq A+1$, car $\lim\limits_{\substack{\alpha\to 1\\ \alpha>1}} \ln(\zeta(\alpha)) = +\infty$. D'autre part, puisque la série de terme général $\ln(1-p_n^{-\alpha})$ est convergente, il existe $N \in \mathbb{N}$ tel que, pour tout $n \geq N$,

$$\left|\sum_{k=n+1}^{+\infty} \ln(1-p_k^{-\alpha})\right| \leq 1.$$

Or, pour tout $n \in \mathbb{N}^*$,

$$-\sum_{k=1}^{n} \ln(1-p_k^{-1}) \geq -\sum_{k=1}^{n} \ln(1-p_k^{-\alpha})$$

et d'après l'égalité (1),

$$-\sum_{k=1}^{n} \ln(1-p_k^{-\alpha}) = \ln(\zeta(\alpha)) + \sum_{k=n+1}^{+\infty} \ln(1-p_k^{-\alpha}).$$

En utilisant la seconde inégalité triangulaire, il s'en suit que, pour tout $n \geq N$,

$$-\sum_{k=1}^{n} \ln(1-p_k^{-1}) \geq |\ln(\zeta(\alpha))| - \left|\sum_{k=n+1}^{+\infty} \ln(1-p_k^{-\alpha})\right| \geq A + 1 - 1 = A.$$

Ainsi, pour tout $A > 0$, il existe $n \geq N$ tel que, pour tout $n \geq N$,

$$-\sum_{k=1}^{n} \ln(1 - p_k^{-1}) \geq A$$

ce qui signifie que la série $-\sum_{n\in\mathbb{N}^*} \ln(1 - p_n^{-1})$ est divergente. Finalement, en utilisant ce qui précède, on en conclut que la série $\sum_{n\in\mathbb{N}^*} \frac{1}{p_n}$ est divergente. □

Complément conseillé.

- Mesure de probabilité associée à une série : proposition 1.

Commentaires personnels. Développement très sympathique, qui se recase dans de nombreuses leçons. Par contre, le développement est assez long. Ne démontrez surtout pas la proposition 2 durant votre exposé. Écrivez directement que $\zeta(\alpha) \underset{\substack{\alpha\to 1\\ \alpha>1}}{\sim} \frac{1}{\alpha - 1}$, et précisez à l'oral que cet équivalent se démontre en effectuant une comparaison série intégrale. En revanche, entraînez-vous à rédiger proprement cette comparaison série-intégrale : c'est niveau licence donc vous devez être irréprochable.

Système de Lotka-Volterra

Cette version est fortement inspirée de celle d'EWna (qu'on peut trouver ici).

[FB] **Théorème 1 (Système de Lotka-Volterra).** Soient $a, b, c, d \in \mathbb{R}_+^*$. Soient $t_0 \in \mathbb{R}$ et $x_0, y_0 \in \mathbb{R}_+^*$. Considérons l'équation différentielle,

$$\begin{cases} x' = ax - bxy, \\ y' = -cy + dxy, \end{cases} \tag{E}$$

puis notons (C_0) le problème de Cauchy associé à (E) en $(t_0, (x_0, y_0))$. Alors, il existe une unique solution globale (sur $\mathbb{R}$) de (C_0), et celle-ci est périodique. De plus, on peut prédire l'allure du portrait de phase de (E).

Preuve. On procède en plusieurs étapes.

Étape n°1. Vérifions tout d'abord que (C_0) est bien un problème de Cauchy. Considérons l'application,

$$\begin{array}{rccc} F & : & \mathbb{R} \times \mathbb{R}^2 & \longmapsto & \mathbb{R}^2 \\ & & (t, (x_1, x_2)) & \longmapsto & (ax_1 - bx_1x_2, -cx_2 + dx_1x_2) \end{array}$$

Alors, si x, y sont des fonctions dérivables de $\mathbb{R}$ dans $\mathbb{R}$, x et y sont solutions de (C_0) si et seulement si $Y = (x, y)$ est solution du problème de Cauchy,

$$\begin{cases} Y' = F(Y), \\ Y(t_0) = (x_0, y_0). \end{cases}$$

Puisque F est de classe C^1, le théorème de Cauchy-Lipschitz assure l'existence d'une unique solution maximale $(I, (x, y))$.

Étape n°2. Montrons que,

$$\forall t \in I,\ x(t) \neq 0.$$

On raisonne par l'absurde en supposant qu'il existe $t_1 \in I$ tel que $x(t_1) = 0$. Montrons que x est alors identiquement nulle sur I. Notons (C_1) le problème de Cauchy associé à (E) sur l'intervalle I en $(t_1, (0, y(t_1)))$. Avec ces notations, (x, y) est solution globale de (C_1). Pour montrer que x est nulle sur I, nous allons exhiber une autre solution globale de (C_1) de la forme $(0, v)$. En effet, en cherchant une solution sous cette forme, on obtient que,

$$(u, v) := (t \mapsto 0,\ t \mapsto y(t_1)e^{-c(t-t_1)})$$

est solution globale de (C_1). En particulier, (x, y) et (u, v) sont deux solutions maximales de (C_1). Or, d'après le théorème de Cauchy-Lipschitz, il existe une unique solution maximale au problème de Cauchy (C_1). Donc, $x = 0$ sur I, ce qui contredit $x(t_0) = x_0 > 0$. Ainsi, x ne s'annule pas sur I. Puisque x est continue sur I, le théorème des valeurs intermédiaires assure que,

$$\forall t \in I,\ x(t) > 0.$$

De façon analogue, on obtient que,

$$\forall t \in I,\ y(t) > 0.$$

Étape n°3. Montrons que x et y sont bornées, ce qui permettra de conclure que $I = \mathbb{R}$ en utilisant le théorème de sortie de tout compact. Notons,

$$\begin{array}{rccc} \gamma & : & I & \longrightarrow & (\mathbb{R}_+^*)^2 \\ & & t & \longmapsto & (x(t), y(t)) \end{array}$$

On cherche une fonction $H : (\mathbb{R}_+^*)^2 \to \mathbb{R}$ telle que $H \circ \gamma$ soit constante. Tout d'abord, pour tout $t \in I$,

$$\begin{cases} \dfrac{x'(t)}{x(t)} = a - by(t) \\ \dfrac{y'(t)}{y(t)} = -c + dx(t) \end{cases}$$

car $(I, (x, y))$ est solution de (C_0). Par suite, pour tout $t \in I$,

$$\frac{x'(t)}{x(t)}(-c + dx(t)) = (a - by(t))\frac{y'(t)}{y(t)}.$$

Notons ensuite,

$$\begin{array}{rccc} H & : & (\mathbb{R}_+^*)^2 & \longrightarrow & \mathbb{R} \\ & & (x_1, x_2) & \longmapsto & dx_1 - c\ln(x_1) + bx_2 - a\ln(x_2) \end{array}$$

Montrons que $H \circ \gamma$ est constante. Pour tout $t \in I$,

$$(H \circ \gamma)(t) = dx(t) - c\ln(x(t)) + by(t) - a\ln(y(t)).$$

D'où, pour tout $t \in I$,

$$(H \circ \gamma)'(t) = \frac{x'(t)}{x(t)}(dx(t) - c) + \frac{y'(t)}{y(t)}(b - ay(t)) = 0$$

et donc $H \circ \gamma$ est constante. Notons ensuite, pour tous $\alpha, \beta \in \mathbb{R}_+^*$,

$$\begin{array}{rccc} \mathcal{F}_{\alpha,\beta} & : & \mathbb{R}_+^* & \longrightarrow & \mathbb{R} \\ & & s & \longmapsto & \beta s - \alpha \ln(s) \end{array}$$

de telle sorte que,

$$\forall (x_1, x_2) \in (\mathbb{R}_+^*)^2,\ H(x_1, x_2) = \mathcal{F}_{c,d}(x_1) + \mathcal{F}_{a,b}(x_2).$$

En dérivant $\mathcal{F}_{\alpha,\beta}$, on obtient le tableau de variations suivant.

s	0		$\frac{\alpha}{\beta}$		$+\infty$
$\mathcal{F}'_{\alpha,\beta}(s)$	‖	$-$	0	$+$	
$\mathcal{F}_{\alpha,\beta}$	$+\infty$	↘	$\mathcal{F}_{\alpha,\beta}(\frac{\alpha}{\beta})$	↗	$+\infty$

En particulier, $\mathcal{F}_{c,d}$ et $\mathcal{F}_{a,b}$ admettent tout deux un minimum global. Notons $K = (H \circ \gamma)(0)$. Puisque $H \circ \gamma$ est constante, il vient,

$$\forall t \in I,\ \mathcal{F}_{c,d}(x(t)) = K - \mathcal{F}_{a,b}(y(t)) \leq K - \min \mathcal{F}_{a,b}$$

Or, puisque $\lim\limits_{s \to +\infty} \mathcal{F}_{c,d}(s) = +\infty$, il existe $M \in \mathbb{R}_+^*$ tel que,

$$\forall s > M,\ \mathcal{F}_{c,d}(s) > K - \min \mathcal{F}_{a,b}.$$

Des deux inégalités précédentes, on en déduit que,

$$\forall t \in I,\ x(t) \leq M$$

(voir figure 12). Ainsi,

$$\forall t \in I,\ 0 < x(t) \leq M,$$

et donc x est bornée sur I. De façon analogue, on obtient que y est bornée sur I. D'après le théorème de sortie de tout compact, on en déduit que $I = \mathbb{R}$, et donc (x, y) est une solution globale (sur $\mathbb{R}$) de (C_0).

Remarque. En fait, puisque $\lim\limits_{s \to 0} \mathcal{F}_{c,d}(s) = +\infty$, le même argument donne aussi l'existence de $m \in \mathbb{R}_+^*$ tel que, pour tout $t \in I$, $x(t) \geq m$.

Étape n°3. Passons ensuite à une étude qualitative de l'équation différentielle (E). Tout d'abord, les isoclines horizontales et verticales de F sont :

$$\begin{cases} I_0 = \{(x_1, x_2) \in \mathbb{R}^2,\ x_1(a - bx_2) = 0\} = (\{0\} \times \mathbb{R}) \cup (\mathbb{R} \times \{\frac{a}{b}\}), \\ I_\infty = \{(x_1, x_2) \in \mathbb{R}^2,\ (-c + dx_1)x_2 = 0\} = (\{\frac{c}{d}\} \times \mathbb{R}) \cup (\mathbb{R} \times \{0\}). \end{cases}$$

Les points d'intersection $(0,0)$ et $(\frac{c}{d}, \frac{a}{b})$ sont les points d'équilibre de (E). On partitionne alors le quart de plan $\mathbb{R}_+^* \times \mathbb{R}_+^*$ en quatre régions (voir figure 13) :

$$\begin{cases} A = \left]0, \frac{c}{d}\right[\times \left]0, \frac{a}{b}\right], \\ B = \left[\frac{c}{d}, +\infty\right[\times \left]0, \frac{a}{b}\right[, \\ C = \left]\frac{c}{d}, +\infty\right[\times \left[\frac{a}{b}, +\infty\right[, \\ D = \left]0, \frac{c}{d}\right] \times \left]\frac{a}{b}, +\infty\right[. \end{cases}$$

Dans ces régions, on connaît la direction globale du champ de vecteur γ, c'est-à-dire le sens de variation de x et y. En effet, pour tout $t \in I$,

$$\begin{cases} x'(t) = x(t)(a - by(t)) \\ y'(t) = (-c + dx(t))y(t) \end{cases}$$

et donc si $t \in \mathbb{R}$ est tel que $\gamma(t) \in A \setminus I_0$ alors $x'(t) > 0$ et $y'(t) < 0$, ce qui entraîne que x est strictement croissante et y est strictement décroissante sur un voisinage de t (puisque x' et y' sont continues sur I). De façon analogue, on obtient la direction globale de γ en distinguant neuf cas, présentés sur la figure 13 (le neuvième cas est le point d'équilibre $(\frac{c}{d}, \frac{a}{b})$).

Étape n°4. Supposons $(x_0, y_0) \in A$ et montrons qu'il existe alors $t_1 > t_0$ tel que $\gamma(t_1) \in B$. On raisonne par l'absurde en supposant que, pour tout $t > t_0$, $\gamma(t) \notin B$. Les sens de variation de x et y fournissent alors,

$$\forall t > t_0,\ \gamma(t) \in A$$

(car si γ sort de A, ça ne peut être que vers B, cf. les directions du champ de vecteur γ sur la figure 13). Par suite,

- x est (strictement) croissante et majorée sur $[t_0, +\infty[$ (car $x' > 0$),
- y est (strictement) décroissante et minorée sur $[t_0, +\infty[$ (car $y' < 0$).

D'après le théorème de convergence monotone, x et y admettent tout deux une limite en $+\infty$, que l'on note respectivement x_∞ et y_∞, et l'on a donc (x_∞, y_∞) qui est un point d'équilibre de (E). Or, par croissance de x,

$$\forall t > t_0,\ x_\infty \geq x_0 > 0.$$

Donc $(x_\infty, y_\infty) = (\frac{c}{d}, \frac{a}{b})$. D'autre part, par décroissance de y,

$$\forall t > t_0,\ \frac{a}{b} = y_\infty \leq y_0 \leq \frac{a}{b}.$$

Donc y est stationnaire en t_0, ce qui contredit le fait que y soit strictement décroissante sur $[t_0, +\infty[$. D'où l'existence de $t_1 > t_0$ tel que $\gamma(t_1) \in B$.

Étape n°5. En itérant le raisonnement, on obtient alors l'existence de réels t_2, t_3, t_4, t_5 tels que $t_5 > t_4 > t_3 > t_2 > t_1$ et,

$$\gamma(t_2) \in C,\ \gamma(t_3) \in D,\ \gamma(t_4) \in A,\ \gamma(t_5) \in B.$$

Le théorème des valeurs intermédiaires appliqué à x assure qu'on peut choisir t_1 et t_5 de telle sorte que,

$$x(t_1) = x(t_5) = \frac{c}{d},$$

c'est-à-dire que $\gamma(t_1), \gamma(t_5) \in B \cap I_\infty$. Montrons que $\gamma(t_1) = \gamma(t_5)$, c'est-à-dire que $y(t_1) = y(t_5)$. Puisque $H \circ \gamma$ est constante, on a,

$$\mathcal{F}_{a,b}(y(t_1)) = K - \mathcal{F}_{c,d}(x(t_1)) = K - \mathcal{F}_{c,d}(x(t_5)) = \mathcal{F}_{a,b}(y(t_5)).$$

Or, $y(t_1), y(t_5) \in]0, \frac{a}{b}[$ et $\mathcal{F}_{a,b}$ est injective sur $]0, \frac{a}{b}[$ (car strictement décroissante sur cet intervalle). Donc $y(t_1) = y(t_5)$, ce qui prouve que $\gamma(t_1) = \gamma(t_5)$.

Étape n°6. Montrons finalement que γ est périodique. Notons $T = t_5 - t_1$ (qui est strictement positif), et considérons l'application,

$$\begin{array}{rccc} \widetilde{\gamma} & : & \mathbb{R} & \longrightarrow & (\mathbb{R}_+^*)^2 \\ & & t & \longmapsto & \gamma(t+T) \end{array}$$

Alors, γ et $\widetilde{\gamma}$ sont deux solutions globales du problème de Cauchy (C_1) associé à (E) en $(t_1, \gamma(t_1))$. Or, d'après le théorème de Cauchy-Lipschitz, ce problème de Cauchy admet une unique solution, ce qui implique que,

$$\forall t \in \mathbb{R},\ \gamma(t+T) = \gamma(t)$$

c'est-à-dire que γ est T-périodique. Ainsi, γ est périodique. Toutefois, depuis l'étape n°4, on a supposé $(x_0, y_0) \in A$. En fait, si $(x_0, y_0) \in A \cup B \cup C \cup D$, le même raisonnement permet de conclure que γ est périodique. En effet, même si les autres régions ne sont pas bornées (contrairement à A) nous n'avons pas utilisé cette hypothèse sur A. Enfin, si $(x_0, y_0) = (\frac{c}{d}, \frac{a}{b})$ alors γ est constante (a fortiori périodique). □

Figures. 12, 13.

Compléments conseillés.

- Théorème de Cauchy-Lipschitz : théorème 2.
- Théorème de sortie de tout compact : théorème 2.

Commentaires personnels. Développement assez technique et très long si on écrit tous les détails. De plus, il se recase dans peu de leçons. Mais beaucoup de ces détails peuvent être exposés oralement. Pour être tout à fait honnête, je ne maîtrisais pas suffisamment ce développement pour l'exposer en seulement 15 minutes. Malgré cela, je pense que c'est un bon choix de développement. Il vous offre l'opportunité de dessiner, ce qui sera apprécié par le jury.

Théorème central limite

On se place sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$. Soient $(X_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires réelles et X une variable aléatoire réelle (sur $(\Omega, \mathcal{A})$).

Pour tout $n \in \mathbb{N}^*$, on note $S_n = \sum_{k=1}^{n} X_k$.

[BB] **Théorème 1 (Théorème central limite).** On suppose que les variables aléatoires de $(X_n)_{n\in\mathbb{N}^*}$ sont indépendantes et identiquement distribuées et que $X_1 \in L^2_{\mathbb{R}}(\Omega, \mathcal{A}, \mathbb{P})$. On note ensuite $m = \mathbb{E}[X_1]$ et $\sigma = \sqrt{\mathbb{V}[X_1]}$. Alors, la suite de variables aléatoires $(\frac{S_n - nm}{\sqrt{n}})_{n\in\mathbb{N}^*}$ converge en loi vers toute variable aléatoire dont la loi est $\mathcal{N}(0, \sigma^2)$.

Preuve. Pour tout $n \in \mathbb{N}^*$, notons $\psi_n = \varphi_{\frac{S_n - nm}{\sqrt{n}}}$ et $\psi = \varphi_{(X_1 - m)}$ (qui est la fonction caractéristique commune à toutes les variables aléatoires $X_n - m$, car celles-ci sont identiquement distribuées). Montrons que,

$$\frac{S_n - nm}{\sqrt{n}} \xrightarrow[n\to+\infty]{\mathcal{L}} \mathcal{N}(0, \sigma^2).$$

D'après le théorème de Lévy, il suffit de montrer que la suite $(\psi_n)_{n\in\mathbb{N}^*}$ converge simplement sur $\mathbb{R}$, vers la fonction caractéristique de la loi normale d'espérance m et d'écart type σ, c'est-à-dire la fonction $t \mapsto e^{-\frac{\sigma^2 t^2}{2}}$.

Étape n°1. Soit $n \in \mathbb{N}^*$. Calculons ψ_n. Les variables aléatoires $X_1, \dots, X_n$ étant mutuellement indépendantes, on a donc, pour tout $t \in \mathbb{R}$,

$$\psi_n(t) = \prod_{k=1}^{n} \varphi_{\frac{X_k - m}{\sqrt{n}}}(t) = \left(\varphi_{\frac{X_1 - m}{\sqrt{n}}}(t)\right)^n = \left(\psi\left(\frac{t}{\sqrt{n}}\right)\right)^n.$$

Étape n°2. Pour étudier la convergence simple de la suite de fonctions précédente, on prouve tout d'abord le résultat suivant :

$$\forall a, b \in \overline{D(0,1)},\ |a^n - b^n| \leq n|a - b|.$$

En effet, si $a, b \in \overline{D(0,1)}$ alors,

$$|a^n - b^n| = \left|(a-b)\sum_{k=0}^{n-1} a^k b^{n-1-k}\right| \leq |a-b| \sum_{k=0}^{n-1} |a|^k |b|^{n-1-k} \leq n|a-b|.$$

Étape n°3. Fixons finalement $t \in \mathbb{R}$ et étudions la convergence de la suite $(\psi(\frac{t}{\sqrt{n}}))_{n\in\mathbb{N}^*}$. Pour cela, nous allons effectuer un développement limité de ψ, en utilisant la formule de Taylor-Young. En effet, ψ est de classe C^2 sur $\mathbb{R}$ car $X_1 \in L^2_{\mathbb{R}}(\Omega, \mathcal{A}, \mathbb{P})$, avec notamment,

$$\begin{cases} \psi(0) = \mathbb{E}[1_\Omega] = 1, \\ \psi'(0) = i\mathbb{E}[X_1 - m] = 0, \\ \psi''(0) = i^2\mathbb{E}[(X_1 - m)^2] = -\mathbb{V}[X_1 - m] + (\mathbb{E}[X_1 - m])^2 = -\sigma^2. \end{cases}$$

D'après la formule de Taylor-Young, il existe donc une fonction $\varepsilon : \mathbb{R} \to \mathbb{R}$ telle que, $\lim_{h\to 0} \varepsilon(h) = 0$, et, pour tout $h \in \mathbb{R}$,

$$\psi(h) = 1 - \frac{\sigma^2 h^2}{2} + \varepsilon(h)h^2.$$

Fixons ensuite $n \in \mathbb{N}^*$. D'après l'étape n°2,

$$\left|\left(\psi\left(\frac{t}{\sqrt{n}}\right)\right)^n - \left(1 - \frac{\sigma^2 t^2}{2n}\right)^n\right| \leq n\left|\psi\left(\frac{t}{\sqrt{n}}\right) + \frac{\sigma^2 t^2}{2n} - 1\right|.$$

Or, d'après l'égalité précédente,

$$\psi\left(\frac{t}{\sqrt{n}}\right) + \frac{\sigma^2 t^2}{2n} - 1 = \varepsilon\left(\frac{t}{\sqrt{n}}\right)\frac{t^2}{n}.$$

Donc,

$$\left|\left(\psi\left(\frac{t}{\sqrt{n}}\right)\right)^n - \left(1 - \frac{\sigma^2 t^2}{2n}\right)^n\right| \leq \left|\varepsilon\left(\frac{t}{\sqrt{n}}\right)\right| t^2 \xrightarrow[n\to+\infty]{} 0.$$

Sachant que $\left(1 - \frac{\sigma^2 t^2}{2n}\right)^n \xrightarrow[n\to+\infty]{} e^{-\frac{\sigma^2 t^2}{2}}$, on en conclut que,

$$\psi_n(t) = \left(\psi\left(\frac{t}{\sqrt{n}}\right)\right)^n \xrightarrow[n\to+\infty]{} e^{-\frac{\sigma^2 t^2}{2}},$$

ce qui achève la preuve. □

Compléments conseillés.

- Convergence en loi : théorème 2.
- Fonction caractéristique (probabilités) : proposition 3.

Commentaires personnels. Développement très sympathique, qui se comprend et se retient facilement, et qui se présente dans de nombreuses leçons. De plus, le théorème central limite figure sur le programme du concours. En revanche, ce développement est un peu court. Il faudrait peut-être ajouter une application. Je vous laisse y réfléchir (désolé, je n'aime pas les probabilités).

Théorème d'Ascoli

[BB] **Théorème 1 (Théorème d'Ascoli).** Soit $\mathcal{A}$ une partie de $C(X,\mathbb{K})$. Alors, les assertions suivantes sont équivalentes :

(i) la partie $\mathcal{A}$ est équicontinue et bornée;

(ii) la partie $\mathcal{A}$ est relativement compacte.

Preuve. On procède par double implication.

• (i) $\Rightarrow$ (ii). Supposons que $\mathcal{A}$ soit équicontinue et bornée. Tout d'abord, K est séparable car compact, c'est-à-dire qu'il existe une suite $(x_p)_{p\in\mathbb{N}}$ d'éléments de K telle que l'ensemble $\{x_p,\ p\in\mathbb{N}\}$ soit dense dans K. Soit $(f_n)_{n\in\mathbb{N}}$ une suite d'éléments de $\mathcal{A}$. Nous allons construire une application φ strictement croissante de $\mathbb{N}$ dans $\mathbb{N}$ telle que $(f_{\varphi(n)})_{n\in\mathbb{N}}$ converge uniformément sur K. On remarque tout d'abord que la suite $(f_n)_{n\in\mathbb{N}}$ est bornée, car $\mathcal{A}$ est borné, et donc il existe $M\in\mathbb{R}$ tel que, pour tout $n\in\mathbb{N}$, $\|f_n\|_\infty\leq M$. En particulier, pour tout $p\in\mathbb{N}$, l'ensemble,

$$\{f_n(x_p),\ n\in\mathbb{N}\}$$

est relativement compact dans $\mathbb{C}$, car inclus dans $\overline{D(0,M)}=\{z\in\mathbb{C},\ |z|\leq M\}$ qui est compact. Par le procédé d'extraction diagonale, on peut donc construire une application $\varphi:\mathbb{N}\to\mathbb{N}$ strictement croissante telle que, pour tout $p\in\mathbb{N}$, la suite $(f_{\varphi(n)}(x_p))_{n\in\mathbb{N}}$ converge dans $\mathbb{C}$.

Montrons que $(f_{\varphi(n)})_{n\in\mathbb{N}}$ est de Cauchy dans $(C(K,\mathbb{C}),\|\cdot\|_\infty)$. Soit $\varepsilon>0$. D'après le théorème de Heine, $\mathcal{A}$ est uniformément équicontinue, c'est-à-dire qu'il existe $\alpha>0$ tel que,

$$\forall f\in\mathcal{A},\ \forall(x_1,x_2)\in X^2,\ \left(d(x_1,x_2)\leq\alpha\Longrightarrow|f(x_1)-f(x_2)|\leq\frac{\varepsilon}{3}\right). \tag{1}$$

Or, l'ensemble $\{x_p,\ p\in\mathbb{N}\}$ est dense dans K. Donc,

$$K\subset\bigcup_{p\in\mathbb{N}}B(x_p,\alpha).$$

Par compacité de K, il existe alors $y_1,\dots,y_m\in\{x_p,\ p\in\mathbb{N}\}$ tels que,

$$K\subset\bigcup_{k=1}^{m}B(y_k,\alpha). \tag{2}$$

Soit $k\in\{1,\dots,m\}$. La suite $(f_{\varphi(n)}(y_k))_{n\in\mathbb{N}}$ est de Cauchy dans $\mathbb{C}$ car convergente. D'où l'existence de $N_k\in\mathbb{N}$ tel que,

$$\forall p,q\geq N_k,\ |f_{\varphi(p)}(y_k)-f_{\varphi(q)}(y_k)|\leq\frac{\varepsilon}{3}. \tag{3}$$

Posons alors $N=\max\limits_{1\leq k\leq m}(N_k)$. Montrons que,

$$\forall p,q\geq N,\ \forall x\in K,\ |f_{\varphi(p)}(x)-f_{\varphi(q)}(x)|\leq\varepsilon.$$

Soient $p,q\geq N$ et $x\in X$. D'après l'inclusion (2), il existe $k\in\{1,\dots,m\}$ tel que $x\in B(y_k,\alpha)$. Par suite, d'après l'assertion (1),

$$\forall f\in\mathcal{A},\ |f(x)-f(y_k)|\leq\frac{\varepsilon}{3}.$$

En particulier,

$$|f_{\varphi(p)}(x)-f_{\varphi(p)}(y_k)|\leq\frac{\varepsilon}{3},\quad|f_{\varphi(q)}(x)-f_{\varphi(q)}(y_k)|\leq\frac{\varepsilon}{3}.$$

Or, par inégalité triangulaire,

$$\begin{aligned}|f_{\varphi(p)}(x)-f_{\varphi(q)}(x)|\ &\leq\ |f_{\varphi(p)}(x)-f_{\varphi(p)}(y_k)|+|f_{\varphi(p)}(y_k)-f_{\varphi(q)}(y_k)|\\&\quad+|f_{\varphi(q)}(y_k)-f_{\varphi(q)}(x)|.\end{aligned}$$

D'où,

$$|f_{\varphi(p)}(x)-f_{\varphi(q)}(x)|\leq\frac{2\varepsilon}{3}+|f_{\varphi(p)}(y_k)-f_{\varphi(q)}(y_k)|.$$

Puis, en utilisant l'assertion (3),

$$|f_{\varphi(p)}(x)-f_{\varphi(q)}(x)|\leq\frac{2\varepsilon}{3}+\frac{\varepsilon}{3}=\varepsilon.$$

car $p,q\geq N\geq N_k$. Finalement,

$$\forall\varepsilon>0,\ \exists N\in\mathbb{N},\ \forall p,q\geq N,\ \|f_{\varphi(p)}-f_{\varphi(q)}\|_\infty\leq\varepsilon$$

Ainsi, $(f_{\varphi(n)})_{n\in\mathbb{N}}$ est de Cauchy dans $(C(K,\mathbb{C}),\|\cdot\|_\infty)$. Ce dernier étant complet, il en résulte que $(f_{\varphi(n)})_{n\in\mathbb{N}}$ converge dans $(C(K,\mathbb{C}),\|\cdot\|_\infty)$. On en conclut que $\mathcal{A}$ est relativement compacte.

• (ii) $\Rightarrow$ (i). Réciproquement, on suppose que $\mathcal{A}$ est relativement compacte, ce qui équivaut à ce que $\overline{\mathcal{A}}$ soit compacte. En particulier, $\overline{\mathcal{A}}$ est bornée et donc $\mathcal{A}$ est également bornée. Soit $\varepsilon>0$. Alors,

$$\overline{\mathcal{A}}\subset\bigcup_{f\in\overline{\mathcal{A}}}B\left(f,\frac{\varepsilon}{3}\right).$$

Par compacité de $\overline{\mathcal{A}}$, il existe $f_1, \dots, f_n \in \overline{\mathcal{A}}$ tels que,

$$\overline{\mathcal{A}} \subset \bigcup_{k=1}^{n} B\Big(f_k, \frac{\varepsilon}{3}\Big). \tag{4}$$

Soit $k \in \{1, \dots, n\}$. D'après le théorème de Heine, f_k est uniformément continue sur K en tant que fonction continue sur un compact. Autrement dit, il existe $\alpha_k > 0$ tel que,

$$\forall (x,y) \in K^2,\ \Big(d(x,y) \leq \alpha_k \Longrightarrow |f_k(x) - f_k(y)| \leq \frac{\varepsilon}{3}\Big).$$

Posons alors $\alpha = \min\limits_{1 \leq k \leq n} (\alpha_k)$. Montrons que,

$$\forall f \in \mathcal{A},\ \forall (x,y) \in K^2,\ (d(x,y) \leq \alpha \Longrightarrow |f(x) - f(y)| \leq \varepsilon).$$

Soient $f \in \mathcal{A}$ et $x, y \in K$ tels que $d(x,y) \leq \alpha$. D'après l'inclusion (4), il existe $k \in \{1, \dots, n\}$ tel que, $f \in B\big(f_k, \frac{\varepsilon}{3}\big)$. Par inégalité triangulaire, il s'en suit que,

$$|f(x) - f(y)| \leq |f(x) - f_k(x)| + |f_k(x) - f_k(y)| + |f_k(y) - f(y)|$$

et donc,

$$|f(x) - f(y)| \leq 2\|f - f_k\|_\infty + |f_k(x) - f_k(y)| \leq \frac{2\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon$$

car $f \in B\big(f_k, \frac{\varepsilon}{3}\big)$ et $d(x,y) \leq \alpha \leq \alpha_k$. Ainsi,

$$\forall \varepsilon > 0,\ \exists \alpha > 0,\ \forall f \in \mathcal{A},\ \forall (x,y) \in K^2,\ (d(x,y) \leq \alpha \Longrightarrow |f(x) - f(y)| \leq \varepsilon)$$

c'est-à-dire que $\mathcal{A}$ est uniformément équicontinue. A fortiori, $\mathcal{A}$ est équicontinue. □

Remarque. Il existe un énoncé plus général du théorème d'Arzelà-Ascoli, caractérisant les parties relativement compactes de $C(X,E)$, où (X, d_X) est un espace métrique compact et (E, d_E) est un espace métrique complet.

Application 2. Soient $M, c \in \mathbb{R}_+^*$ et $\alpha \in]0,1]$. Considérons,

$$\mathcal{A} = \left\{ f \in C([0,1], \mathbb{C}) \;\middle|\; \begin{array}{l} \forall (x,y) \in [0,1]^2,\ |f(x) - f(y)| \leq M|x-y|^\alpha \\ |f(0)| \leq c \end{array} \right\}.$$

Alors, $\mathcal{A}$ est une partie compacte de $(C([0,1], \mathbb{C}), \|\cdot\|_\infty)$.

Preuve. Montrons tout d'abord que $\mathcal{A}$ est bornée. Soit $f \in \mathcal{A}$. Alors, d'après la seconde inégalité triangulaire,

$$\forall x \in X,\ |f(x)| \leq M|x|^\alpha + |f(0)|$$

car $|x|^\alpha = \exp(\alpha \ln(|x|)) \leq \exp(0) = 1$. D'où, $\|f\|_\infty \leq M + c$, ce qui prouve que $\mathcal{A}$ est bornée. Montrons ensuite que $\mathcal{A}$ est équicontinue. Soit $\varepsilon > 0$. Posons $\eta = \exp\big(\frac{1}{\alpha} \ln\big(\frac{\varepsilon}{M}\big)\big)$. Alors, pour tout $f \in \mathcal{A}$ et pour tous $x, y \in X$ tels que $|x - y| \leq \eta$,

$$|f(x) - f(y)| \leq M|x-y|^\alpha = M\left(\exp\left(\frac{1}{\alpha} \ln\Big(\frac{\varepsilon}{M}\Big)\right)\right)^\alpha = \varepsilon.$$

Donc $\mathcal{A}$ est équicontinue (et même uniformément équicontinue). D'après le théorème d'Ascoli, on en déduit que $\mathcal{A}$ est une partie relativement compacte de $C([0,1], \mathbb{C})$, et donc que $\overline{\mathcal{A}}$ est compact dans $(C([0,1], \mathbb{C}), \|\cdot\|_\infty)$. Pour conclure, il suffit donc de montrer que $\mathcal{A} = \overline{\mathcal{A}}$, c'est-à-dire que $\mathcal{A}$ est fermé dans $(C([0,1], \mathbb{C}), \|\cdot\|_\infty)$. Soit $(f_n)_{n \in \mathbb{N}}$ une suite de fonctions à termes dans $\mathcal{A}$ qui converge vers f dans $(C([0,1], \mathbb{C}), \|\cdot\|_\infty)$. Soient $x, y \in [0,1]$. Alors, pour tout $n \in \mathbb{N}$,

$$|f_n(x) - f_n(y)| \leq M|x-y|^\alpha,$$

et $|f_n(0)| \leq c$. D'où, en faisant tendre n vers l'infini,

$$|f(x) - f(y)| \leq M|x-y|^\alpha,$$

et $|f(0)| \leq c$, car $(f_n)_{n \in \mathbb{N}}$ converge simplement vers f sur $[0,1]$. Donc, $f \in \mathcal{A}$. Ainsi, $\mathcal{A}$ est fermée, ce qui achève la preuve. □

Compléments conseillés.

- Compacité relative : sous-section 3.4.
- Complétude et espaces de fonctions : théorème 1.
- Équicontinuité : sous-section 3.12.
- Séparabilité : proposition 2.

Commentaires personnels. Développement long et très technique, nécessitant de nombreux prérequis. Malgré cela, il s'agit d'un de mes développements favoris. Je vous le recommande si vous appréciez la topologie et l'analyse fonctionnelle. De plus, le théorème d'Ascoli figure sur le programme du concours.

Théorème d'Hadamard-Lévy

[BB] **Théorème 1 (Théorème d'Hadamard-Lévy).** Soit $f : \mathbb{R}^n \to \mathbb{R}^n$ une application de classe C^2. Les assertions suivantes sont équivalentes :

(i) f est un difféomorphisme de classe C^1 de $\mathbb{R}^n$ dans $\mathbb{R}^n$,

(ii) pour tout $x \in \mathbb{R}^n$, $Df(x)$ est inversible, et $\lim\limits_{\|x\|\to+\infty} \|f(x)\| = +\infty$.

Preuve. On procède par double implication.

• (i) $\Rightarrow$ (ii). Soit $x \in \mathbb{R}^n$. En utilisant la formule de composée d'applications différentiables, il vient,

$$Id_{\mathbb{R}^n} = D(Id_{\mathbb{R}^n})(x) = D(f^{-1} \circ f)(x) = Df^{-1}(f(x)) \circ Df(x)$$

ce qui prouve que $Df(x)$ est inversible, d'inverse $Df^{-1}(f(x))$. Fixons $A > 0$. L'ensemble,

$$\{x \in \mathbb{R}^n,\ \|f(x)\| \leq A\}$$

est compact, car c'est l'image réciproque du compact $\overline{B(0,A)}$ par l'application f^{-1} qui est continue par hypothèse. En particulier, il existe $R > 0$ tel que,

$$\{x \in \mathbb{R}^n,\ \|f(x)\| \leq A\} \subset \overline{B(0,R)}.$$

En passant au complémentaire dans l'inclusion précédente, il s'en suit que,

$$\forall x \in \mathbb{R}^n,\ (\|x\| > R \Longrightarrow \|f(x)\| > A)\,.$$

Et donc, $\lim\limits_{\|x\|\to+\infty} \|f(x)\| = +\infty$.

• (ii) $\Rightarrow$ (i). Quitte à remplacer f par la fonction $\mathbb{R}^n \to \mathbb{R}^n$, $x \mapsto f(x) - f(0)$, on peut supposer sans perte de généralité que $f(0) = 0$. Nous allons construire une application $s : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$ telle que,

$$\forall (t,x) \in \mathbb{R} \times \mathbb{R}^n,\ (f \circ s)(t,x) = tx. \tag{1}$$

Supposons que $s : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$ soit une application différentiable vérifiant l'égalité (1). En dérivant cette égalité par rapport à t, on obtient,

$$\forall (t,x) \in \mathbb{R} \times \mathbb{R}^n,\ Df(s(t,x))(\partial_t s(t,x)) = x$$

Et donc,

$$\forall (t,x) \in \mathbb{R} \times \mathbb{R}^n,\ \partial_t s(t,x) = (Df(s(t,x)))^{-1}(x)$$

Cette dernière égalité nous amène à étudier le problème de Cauchy autonome, dépendant d'un paramètre $x \in \mathbb{R}^n$, suivant,

$$(C_x) : \begin{cases} y'(t) = F(x, y(t)), & t \in \mathbb{R} \\ y(0) = 0 \end{cases}$$

où l'on a noté,

$$\begin{array}{rccc} F & : & \mathbb{R}^n \times \mathbb{R}^n & \longrightarrow & \mathbb{R} \\ & & (x,z) & \longmapsto & (Df(z))^{-1}(x) \end{array}$$

L'application F est de classe C^1. En effet, l'application

$$\begin{array}{rccc} b & : & GL(\mathbb{R}^n) \times \mathbb{R}^n & \longrightarrow & \mathbb{R}^n \\ & & (\varphi, x) & \longmapsto & \varphi(x) \end{array}$$

est de classe C^∞ en tant qu'application bilinéaire. De plus, on sait que,

$$\begin{array}{rccc} \Lambda & : & GL(\mathbb{R}^n) & \longrightarrow & GL(\mathbb{R}^n) \\ & & \varphi & \longmapsto & \varphi^{-1} \end{array}$$

est de classe C^∞. Les applications,

$$\begin{array}{rccc} \pi_1 & : & \mathbb{R}^n \times \mathbb{R}^n & \longrightarrow & \mathbb{R}^n \\ & & (x,z) & \longmapsto & x \end{array} \quad \Big| \quad \begin{array}{rccc} \pi_2 & : & \mathbb{R}^n \times \mathbb{R}^n & \longrightarrow & \mathbb{R}^n \\ & & (x,z) & \longmapsto & z \end{array}$$

sont aussi de classe C^∞ en tant qu'applications linéaires. Enfin, l'application,

$$\begin{array}{rccc} Df & : & \mathbb{R}^n & \longrightarrow & GL(\mathbb{R}^n) \\ & & z & \longmapsto & Df(z) \end{array}$$

est de classe C^1 car f est de classe C^2. Ainsi,

$$F = b \circ (\Lambda \circ Df \circ \pi_2, \pi_1)$$

est de classe C^1. En particulier, pour tout $x \in \mathbb{R}^n$, $F(x,\cdot)$ est de classe C^1, et le théorème de Cauchy-Lipschitz assure donc l'existence d'une unique solution maximale $(]t^-(x), t^+(x)[, s_x)$ au problème de Cauchy (C_x).

Soit $x_0 \in \mathbb{R}^n$. Montrons tout d'abord que,

$$\forall t \in]t^-(x_0), t^+(x_0)[,\ f(s_{x_0}(t)) = tx_0. \tag{2}$$

Notons $u :]t^-(x_0), t^+(x_0)[\to \mathbb{R}^n,\ t \mapsto f(s_{x_0}(t)) - tx_0$. Puisque s_{x_0} est solution de (C_x), u est dérivable avec, pour tout $t \in]t^-(x_0), t^+(x_0)[$,

$$u'(t) = Df(s_{x_0}(t))([s_{x_0}]'(t)) - x_0 = 0.$$

Donc u est constante. Sachant que $u(0) = 0$, il s'en suit l'égalité (2). Montrons ensuite que $t^+(x_0) = +\infty$. On raisonne par contradiction en supposant que $t^+(x_0) < +\infty$. D'après le théorème de sortie de tout compact, il vient donc,

$$\lim_{t \to t^+(x_0)} \|s(x_0, t)\| = +\infty.$$

Et puisque $\lim_{\|y\| \to +\infty} \|f(y)\| = +\infty$, ceci implique,

$$\lim_{t \to t^+(x_0)} \|f(s(x_0, t))\| = +\infty.$$

Or, d'après l'inégalité (2),

$$\lim_{t \to t^+(x_0)} \|f(s(x_0, t))\| = \lim_{t \to t^+(x_0)} \|tx_0\| = |t^+(x_0)| \cdot \|x_0\| < +\infty.$$

D'où, $t^+(x_0) = +\infty$. De même, on obtient $t^-(x_0) = -\infty$. Ainsi, l'application,

$$\begin{array}{rccc} s & : & \mathbb{R} \times \mathbb{R}^n & \longrightarrow & \mathbb{R} \\ & & (t, x) & \longmapsto & s_x(t) \end{array}$$

est bien définie et vérifie l'égalité (1), d'après l'égalité (2). De plus, le théorème de Cauchy-Lipschitz à paramètre assure que s est de classe C^1, car F est de classe C^1. En particulier, la fonction,

$$\begin{array}{rccc} g & : & \mathbb{R}^n & \longrightarrow & \mathbb{R}^n \\ & & x & \longmapsto & s(1, x) \end{array}$$

est de classe C^1 et vérifie,

$$\forall x \in \mathbb{R}^n,\ f(g(x)) = x.$$

En particulier, f est surjective. Montrons ensuite que g est surjective. Pour cela, nous allons utiliser un argument de connexité.

- Montrons que $g(\mathbb{R}^n)$ est fermé. Soit $(x_n)_{n \in \mathbb{N}}$ une suite à termes dans $\mathbb{R}^n$ telle que $(g(x_n))_{n \in \mathbb{N}}$ converge vers y. Par continuité de f, il vient,

  $$\lim_{n \to +\infty} x_n = \lim_{n \to +\infty} f(g(x_n)) = f(y).$$

  Puis, par continuité de g,

  $$y = \lim_{n \to +\infty} g(x_n) = g(f(y)).$$

  En particulier, $y \in g(\mathbb{R}^n)$, ce qui prouve que $g(\mathbb{R}^n)$ est fermé.

- Montrons que $g(\mathbb{R}^n)$ est ouvert. Soit $x_0 \in \mathbb{R}^n$. Puisque f est de classe C^1 et que $Df(x_0)$ est inversible, le théorème d'inversion locale assure l'existence de,

  - U un ouvert de $\mathbb{R}^n$ contenant x_0,
  - V un ouvert de $\mathbb{R}^n$ contenant $f(x_0)$,

  tels que l'application,

  $$\begin{array}{rccc} \phi & : & U & \longrightarrow & V \\ & & x & \longmapsto & f(x) \end{array}$$

  soit un difféomorphisme de classe C^1. D'autre part, par continuité de g, il existe un ouvert U' de U tel que $g(U') \subset U$. Or,

  $$g(U') = \phi^{-1}(\phi(g(U'))) = \phi^{-1}(U').$$

  Mais $\phi^{-1}(U')$, image directe de U' par ϕ^{-1}, est aussi l'image réciproque de l'ouvert U' par ϕ qui est continue. C'est donc un ouvert de $\mathbb{R}^n$. Ainsi, $g(U')$ est un ouvert de $\mathbb{R}^n$, contenu dans $g(\mathbb{R}^n)$ et contenant $g(x_0)$, ce qui prouve que $g(\mathbb{R}^n)$ est ouvert.

Ainsi, $g(\mathbb{R}^n)$ est une partie non vide, fermée et ouverte de $\mathbb{R}^n$. D'où,

$$g(\mathbb{R}^n) = \mathbb{R}^n$$

par connexité de $\mathbb{R}^n$. Ainsi, f est bijective de bijection réciproque g. Sachant que f et g sont de classe C^1, on en conclut que f est un difféomorphisme de classe C^1 de $\mathbb{R}^n$ dans $\mathbb{R}^n$. □

Remarque. Le théorème reste vrai si f est supposée seulement de classe C^1.

Complément conseillé.

- Équations différentielles dépendant d'un paramètre : chapitre XI, [JD].

Commentaires personnels. Développement à la fois trop technique et trop long pour mon niveau. J'avais voulu le remplacer par d'autres développements mais je n'avais pas pu, faute de temps. Je vous le déconseille si vous souhaitez uniquement être admis au concours (sans objectif de classement).

Théorème d'inversion locale

Soient $n, p \in \mathbb{N}^*$. Soit U un ouvert de $\mathbb{R}^n$.

[FR] **Théorème 1 (Théorème d'inversion locale).** Soient $f : U \to \mathbb{R}^p$ une application de classe C^1 et $x_0 \in U$. On suppose que $Df(x_0)$ est inversible. Alors, il existe un ouvert V de $\mathbb{R}^n$ contenu dans U et contenant x_0 ainsi qu'un ouvert W de $\mathbb{R}^p$ contenant $f(x_0)$ tel que f soit un difféomorphisme de classe C^1 de V dans W.

Preuve. Tout d'abord, $Df(x_0)$ étant un isomorphisme de $\mathbb{R}^n$ vers $\mathbb{R}^p$, on a nécessairement $n = p$ par le théorème du rang. De plus, quitte à remplacer f par la fonction,

$$\begin{array}{ccc} U - x_0 & \longrightarrow & \mathbb{R}^n \\ x & \longmapsto & [Df(x_0)]^{-1} \cdot [f(x + x_0) - f(x_0)] \end{array}$$

on peut supposer sans perte de généralité,

$$x_0 = 0, \quad f(x_0) = 0, \quad Df(x_0) = Id_{\mathbb{R}^n},$$

Étape n°1. Posons, pour tout $x \in U$, $g(x) = x - f(x)$. Montrons qu'il existe $r > 0$ tel que,

(*i*) $\overline{B(0,r)} \subset U$;

(*ii*) pour tout $x \in B(0,r)$, $Df(x) \in GL(\mathbb{R}^n)$;

(*iii*) pour tous $x, x' \in \overline{B(0,r)}$, $\|g(x) - g(x')\| \leq \frac{1}{2}\|x - x'\|$.

Tout d'abord, puisque $GL(\mathbb{R}^n)$ est ouvert dans $L(\mathbb{R}^n)$, et que $Df(0) \in GL(\mathbb{R}^n)$, il existe $\alpha > 0$ tel que, pour tout $\phi \in L(\mathbb{R}^n)$,

$$\|\phi - Df(0)\| < \alpha \Longrightarrow \phi \in GL(\mathbb{R}^n).$$

Quitte à réduire α, on peut même supposer $2\alpha < 1$, ce qui servira à démontrer le point (*iii*). Or, l'application $x \mapsto Df(x)$ est continue en 0, car f est de classe C^1 sur U. D'où l'existence de $r > 0$ tel que, $B(0, \frac{3r}{2}) \subset U$, et pour tout $x \in B(0, \frac{3r}{2})$,

$$\|Df(x) - Df(0)\| < \alpha$$

et donc $Df(x) \in GL(\mathbb{R}^n)$. En particulier, pour tout $x \in B(0, \frac{3r}{2})$,

$$\|Dg(x)\| = \|Df(0) - Df(x)\| < \alpha \leq \frac{1}{2},$$

car $Dg(x) = Id_{\mathbb{R}^n} - Df(x) = Df(0) - Df(x)$ et $2\alpha < 1$. Or, $B(0, \frac{3r}{2})$ est un ouvert convexe de $\mathbb{R}^n$. D'après l'inégalité des accroissements finis, on a donc, pour tous $x, x' \in B(0, \frac{3r}{2})$,

$$\|g(x) - g(x')\| \leq \frac{1}{2}\|x - x'\|.$$

Puisque $\overline{B(0,r)} \subset B(0, \frac{3r}{2})$, on en déduit les point (*i*), (*ii*) et (*iii*).

Étape n°2. On fixe $y \in B(0, \frac{r}{2})$. Montrons qu'il existe un unique $x \in B(0,r)$ tel que $f(x) = y$. Pour cela, on considère l'application,

$$\begin{array}{ccccc} g_y & : & \overline{B(0,r)} & \longrightarrow & \overline{B(0,r)} \\ & & x & \longmapsto & y + g(x) \end{array}$$

Tout d'abord, g_y est bien définie car, pour tout $x \in \overline{B(0,r)}$,

$$\|g(x)\| = \|g(x) - g(0)\| \leq \frac{1}{2}\|x\| \leq \frac{r}{2}$$

d'après (*iii*), et donc,

$$\|y + g(x)\| \leq \|y\| + \|g(x)\| < \frac{r}{2} + \frac{r}{2},$$

c'est-à-dire $g_y(x) \in B(0,r)$. De plus, pour tous $x, x' \in \overline{B(0,r)}$,

$$\|g_y(x) - g_y(x')\| = \|g(x) - g(x')\| \leq \frac{1}{2}\|x - x'\|.$$

Donc g_y est contractante. Or, $\overline{B(0,r)}$ est complet, en tant que partie fermée de $\mathbb{R}^n$ qui est complet. D'après le théorème du point fixe de Banach, il existe donc un unique $x \in \overline{B(0,r)}$ tel que $g_y(x) = x$. En particulier, $x \in B(0,r)$ car $g_y(x) \in B(0,r)$, et $f(x) = y$ en utilisant la définition de $g_y(x)$.

Étape n°3. Considérons les ouverts de $\mathbb{R}^n$ suivants,

$$W = B\left(0, \frac{r}{2}\right), \quad V = f^{-1}(W) \cap B(0,r).$$

Alors, l'application,

$$\begin{array}{ccccc} f & : & V & \longrightarrow & W \\ & & x & \longmapsto & f(x) \end{array}$$

est une bijection de classe C^1 sur V. Pour conclure, il reste à montrer que sa bijection réciproque, notée f^{-1}, est de classe C^1 sur W. Montrons que f^{-1}

est lipschitzienne de rapport 2. Soient $y, y' \in W$. Soient $x, x' \in V$ tels que $y = f(x)$ et $y' = f(x')$. Par suite,

$$\|f^{-1}(y) - f^{-1}(y')\| = \|g(x) - g(x') + y - y'\| \leq \|g(x) - g(x')\| + \|y - y'\|.$$

Or, d'après (iii),

$$\|g(x) - g(x')\| \leq \frac{1}{2}\|x - x'\| = \frac{1}{2}\|f^{-1}(y) - f^{-1}(y')\|.$$

D'où,

$$\|f^{-1}(y) - f^{-1}(y')\| \leq 2\|y - y'\|.$$

Ainsi, f^{-1} est lipschitzienne de rapport 2.

Étape n°4. Soit $y_0 \in W$. Notons $x_0 = f^{-1}(y_0)$, puis considérons,

$$\begin{array}{rccc} R & : & V & \longrightarrow & \mathbb{R}^n \\ & & x & \longmapsto & f(x) - f(x_0) - Df(x_0)(x - x_0) \end{array}$$

et,

$$\begin{array}{rccc} S & : & W & \longrightarrow & \mathbb{R}^n \\ & & y & \longmapsto & f^{-1}(y) - f^{-1}(y_0) - [Df(x_0)]^{-1} \cdot [y - y_0] \end{array}$$

Montrons que f^{-1} est différentiable au point y_0, avec $D(f^{-1})(y_0) = Df(x_0)^{-1}$. Puisque $Df(x_0)^{-1}$ est linéaire continue, il suffit de montrer que,

$$\lim_{\substack{y \to y_0 \\ y \neq y_0}} \frac{1}{\|y - y_0\|} S(y) = 0.$$

Notons tout d'abord que, pour tout $y \in W$,

$$S(y) = -\left[Df(x_0)\right]^{-1} \cdot \left[R(f^{-1}(y))\right].$$

En particulier, pour tout $y \in W$,

$$\|S(y)\| = \left\|Df(x_0)^{-1}(R(f^{-1}(y)))\right\| \leq \|Df(x_0)^{-1}\| \times \left\|R(f^{-1}(y))\right\|.$$

Et puisque f^{-1} est 2-lipschitzienne, il vient alors, pour tout $y \in W$,

$$\frac{\|S(y)\|}{\|y - y_0\|} \leq 2\|Df(x_0)^{-1}\| \frac{\left\|R(f^{-1}(y))\right\|}{\|f^{-1}(y) - x_0\|},$$

De plus, f^{-1} est continue sur W car lipschitzienne, et donc,

$$\lim_{y \to y_0} 2\|Df(x_0)^{-1}\| \frac{\left\|R(f^{-1}(y))\right\|}{\|f^{-1}(y) - x_0\|} = \lim_{x \to x_0} 2\|Df(x_0)^{-1}\| \frac{\|R(x)\|}{\|x - x_0\|} = 0.$$

Finalement, l'inégalité précédente permet de conclure que,

$$\lim_{y \to y_0} \frac{\|S(y)\|}{\|y - y_0\|} = 0,$$

et donc que f^{-1} est différentiable en y_0, avec $D(f^{-1})(y_0) = Df(x_0)^{-1}$.

Étape n°5. Finalement, f^{-1} est de classe C^1 sur W car,

$$\begin{array}{rccc} \Lambda & : & GL(\mathbb{R}^n) & \longrightarrow & GL(\mathbb{R}^n) \\ & & \varphi & \longrightarrow & \varphi^{-1} \end{array}$$

est continue, et donc,

$$D(f^{-1}) = \Lambda \circ Df \circ f^{-1}$$

est continue sur W en tant que composée d'applications continues sur W. □

Figure. 15.

Compléments conseillés.

- Accroissements finis : sous-section 3.1.
- Régularité de l'inverse dans $GL(\mathbb{R}^n)$: lemme 1.
- Théorème du point fixe de Banach-Picard : théorème 1.

Commentaires personnels. Développement à la fois trop technique et trop long pour mon niveau. J'ai finalement renoncé à le présenter pour ces raisons. Néanmoins, c'est un théorème au programme du concours, donc je pense qu'il est préférable d'étudier cette preuve classique, qui repose sur le théorème du point fixe de Banach.

Théorème de Cauchy-Lipschitz

Soient $n, m \in \mathbb{N}^*$, I un intervalle ouvert de $\mathbb{R}$, Ω un ouvert de $(\mathbb{R}^m)^n$ et F une application de $I \times \Omega$ dans $\mathbb{R}^m$. On note $\|\cdot\|$ la norme usuelle sur $\mathbb{R}^m$.

Considérons l'équation différentielle,

$$y'(t) = F(t, y(t)), \quad t \in I. \tag{E}$$

Soit $(t_0, y_0) \in I \times \Omega$. Notons (C) le problème de Cauchy de (E) en (t_0, y_0).

[JD] **Lemme 1.** Supposons que F soit continue. Soient $J \subset I$ un intervalle ouvert et $y : J \to \mathbb{R}^m$. Alors, (J, y) est solution de (C) si et seulement si,

(i) y est continue et, pour tout $t \in J$, $(t, y(t)) \in \Omega$,

(ii) pour tout $t \in J$, $y(t) = y_0 + \displaystyle\int_{t_0}^{t} F(s, y(s))ds$.

Preuve. C'est une conséquence du théorème fondamental de l'analyse. □

[JD] **Théorème 2 (Théorème de Cauchy-Lipschitz).** Supposons que F soit continue et localement lipschitzienne par rapport à son deuxième argument. Alors, le problème de Cauchy (C) admet une unique solution maximale.

Preuve. On procède en deux étapes.

Étape n°1. Montrons tout d'abord l'existence et l'unicité locale. Puisque la fonction F est localement lipschitzienne, il existe $r > 0$ et $L > 0$ tels que,

$$[t_0 - r, t_0 + r] \times \overline{B(y_0, r)} \subset I \times \Omega,$$

et, pour tous $(t, u), (t, v) \in [t_0 - r, t_0 + r] \times \overline{B(y_0, r)}$,

$$\|F(t, u) - F(t, v)\| \leq L\|u - v\|.$$

De plus, puisque F est continue sur $[t_0 - r, t_0 + r] \times \overline{B(y_0, r)}$ qui est compact, on peut poser,

$$M = \sup\left\{\|F(t, x)\|,\ (t, x) \in [t_0 - r, t_0 + r] \times \overline{B(y_0, r)}\right\}.$$

Soit $\alpha \in]0, r[$. Notons C_α l'ensemble des fonctions continues de $[t_0 - \alpha, t_0 + \alpha]$ dans $\overline{B(y_0, r)}$, que l'on munit de la distance induite par la norme $\|\cdot\|_\infty$. Pour tout $\varphi \in C_\alpha$, on considère l'application,

$$\begin{array}{rccc} \Lambda(\varphi) & : & [t_0 - \alpha, t_0 + \alpha] & \longrightarrow & \mathbb{R}^d \\ & & t & \longmapsto & y_0 + \displaystyle\int_{t_0}^{t} F(u, \varphi(u))du \end{array}$$

Soient $\varphi, \psi \in C_\alpha$. Soit $t \in [t_0, t_0 + \alpha]$. Par inégalité triangulaire,

$$\|\Lambda(\varphi)(t) - y_0\| \leq \int_{t_0}^{t} \|F(u, \varphi(u))\|du \leq M(t - t_0).$$

En procédant de même sur $[t_0 - \alpha, t_0]$, il vient, pour tout $t \in [t_0 - \alpha, t_0 + \alpha]$,

$$\|\Lambda(\varphi)(t) - y_0\| \leq M|t - t_0| \leq M\alpha.$$

En particulier, si $\alpha \leq \frac{r}{M}$ alors $\Lambda(\varphi) \in C_\alpha$. D'autre part,

$$\begin{array}{rcl} \|\Lambda(\varphi)(t) - \Lambda(\varphi)(t)\| & \leq & \displaystyle\int_{t_0}^{t} \|F(u, \varphi(u)) - F(u, \psi(u))\|du \\ & \leq & \displaystyle\int_{t_0}^{t} L\|\varphi(u) - \psi(u)\|du \\ \|\Lambda(\varphi)(t) - \Lambda(\varphi)(t)\| & \leq & L\|\varphi - \psi\|_\infty(t - t_0). \end{array}$$

En procédant de même sur $[t_0 - \alpha, t_0]$, il vient, pour tout $t \in [t_0 - \alpha, t_0 + \alpha]$,

$$\|\Lambda(\varphi)(t) - \Lambda(\varphi)(t)\| \leq L\|\varphi - \psi\|_\infty|t - t_0| \leq \alpha L\|\varphi - \psi\|_\infty.$$

En particulier, si $\alpha < \frac{1}{L}$ alors $\|\Lambda(\varphi) - \Lambda(\varphi)\|_\infty \leq \alpha L\|\varphi - \psi\|_\infty$ avec $\alpha L < 1$. On choisit donc $\alpha \in]0, r[$ tel que $\alpha \leq \frac{r}{M}$ et $\alpha < \frac{1}{L}$, puis on considère l'application,

$$\begin{array}{rccc} \Lambda & : & C_\alpha & \longrightarrow & C_\alpha \\ & & \varphi & \longmapsto & \Lambda(\varphi) \end{array}$$

qui est bien définie et contractante d'après ce qui précède. Or, C_α est complet car $[t_0 - \alpha, t_0 + \alpha]$ est compact et $\overline{B(y_0, r)}$ est complet en tant que partie fermée de $\mathbb{R}^m$ qui est complet. D'après le théorème du point fixe, il existe donc une unique application $y \in C_\alpha$ tel que $F(y) = y$, qui est l'unique solution de (E) sur $]t_0 - \alpha, t_0 + \alpha[$ d'après le lemme précédent.

Étape n°2. Montrons désormais qu'il existe une unique solution maximale au problème de Cauchy (C). On note S l'ensemble des intervalles ouverts $J \subset I$ tels qu'il existe une solution de (C) sur J, puis $\widetilde{J}$ la réunion de ces intervalles. Soient (J_1, y_1) et (J_2, y_2) des solutions de (C). Notons,

$$V = \{t \in J_1 \cap J_2,\ y_1(t) = y_2(t)\}.$$

Montrons que V est une partie non vide, fermée et ouverte de $J_1 \cap J_2$.

- V est non vide car $t_0 \in J_1 \cap J_2$ et $y_1(t_0) = y_0 = y_2(t_0)$.
- V est une partie fermée de $J_1 \cap J_2$ en tant qu'image réciproque de $\{0\}$, qui est fermé, par l'application continue $J_1 \cap J_2 \to \mathbb{R}^m,\ t \mapsto y_1(t) - y_2(t)$.
- Montrons ensuite que V est une partie ouverte de $J_1 \cap J_2$. Soit $s_0 \in V$, c'est-à-dire que $y_1(s_0) = y_2(s_0)$. Notons alors z_0 cette valeur commune. D'après l'étape n°1, il existe $r > 0$ tel que, $[s_0 - r, s_0 + r]$ soit inclus dans I, et le problème de Cauchy,

$$\begin{cases} z'(t) = F(t, z(t)), \\ z(s_0) = z_0. \end{cases}$$

  possède une unique solution sur $]s_0 - r, s_0 + r[$. Quitte à réduire r, on peut supposer que $]s_0 - r, s_0 + r[\subset J_1 \cap J_2$ (car $J_1 \cap J_2$ est ouvert), et le problème de Cauchy précédent possédera toujours une unique solution sur cet intervalle. Avec cette hypothèse supplémentaire, les fonctions y_1 et y_2 sont solutions de ce problème de Cauchy, ce qui implique qu'elles coïncident sur $]s_0 - r, s_0 + r[$, c'est-à-dire $]s_0 - r, s_0 + r[\subset V$.

Or, $J_1 \cap J_2$ est convexe, en tant qu'intersection de convexes non disjoints. En particulier, $J_1 \cap J_2$ est connexe. Donc, $V = J_1 \cap J_2$, car V est une partie non vide, fermée et ouverte de $J_1 \cap J_2$, ce qui prouve que y_1 et y_2 coïncident sur $J_1 \cap J_2$. Pour tout $t \in \widetilde{J}$, il existe $J \in S$ et $y_J : J \to \mathbb{R}^m$ tel que (J, y_J) soit solution de (C) et on pose alors,

$$\widetilde{y}(t) = y_J(t).$$

D'après ce qui précède, l'expression ci-dessus dépend uniquement de t (elle ne dépend pas du choix de (J, y_J)). Par conséquent, l'application,

$$\begin{array}{rccc} \widetilde{y} & : & \widetilde{J} & \longrightarrow & \mathbb{R}^m \\ & & t & \longmapsto & \widetilde{y}(t) \end{array}$$

est bien définie. De plus, $\widetilde{y}$ est l'unique solution maximale de (C). En effet,

- $\widetilde{y}(t_0) = y_1(t_0) = y_0$,
- Si $t \in \widetilde{J}$, il existe (J, y_J) solution de (C) tel que $t \in J$, et par construction, $\widetilde{y}$ et y_J coïncide alors sur l'intervalle ouvert J, et donc $\widetilde{y}$ est dérivable au point t avec,

$$(\widetilde{y})'(t) = (y_J)'(t) = F(t, y_J(t)) = F(t, \widetilde{y}(t)).$$

  Donc $\widetilde{y}$ est bien solution de (C).
- Si (J, y) une solution de (C) alors, $J \subset \widetilde{J}$ et la restriction de $\widetilde{y}$ à J est y, ce qui prouve que $(\widetilde{J}, \widetilde{y})$ est un prolongement de (J, y).

Ainsi, $(\widetilde{J}, \widetilde{y})$ est l'unique solution maximale de (C). □

Complément conseillé.

- Théorème du point fixe de Banach-Picard : théorème 1.

Commentaires personnels. Développement technique, qui se présente dans peu de leçons. Étant donné que le théorème de Cauchy-Lipschitz figure sur le programme du concours, je pense qu'il est préférable de savoir le prouver (ou du moins avoir une idée de preuve). Et quitte à savoir le prouver, autant en faire un développement.

Théorème de Cauchy-Lipschitz linéaire

Soient $d \in \mathbb{N}^*$, I un intervalle ouvert de $\mathbb{R}$, et A, B des applications continues de I dans $M_d(\mathbb{R})$. On note $\|\cdot\|$ la norme usuelle sur $M_{d,1}(\mathbb{R})$.

Considérons l'équation différentielle,

$$Y'(t) = A(t)Y(t) + B(t), \quad t \in I. \tag{E}$$

d'inconnue une fonction $Y : I \to M_{d,1}(\mathbb{R})$ dérivable. Soit $(t_0, y_0) \in I \times \Omega$. On note (C) le problème de Cauchy de (E) en (t_0, y_0).

[AP] **Théorème 1 (Théorème de Cauchy-Lipschitz linéaire).** Le problème de Cauchy (C) admet une unique solution globale.

Preuve. On sépare la preuve de l'existence et de l'unicité.

Existence. Considérons $(Y_n)_{n\in\mathbb{N}}$ la suite de fonctions définie pour tout $t \in I$ par, $Y_0(t) = Y_0$, et pour tout $n \in \mathbb{N}$ par,

$$Y_{n+1}(t) = Y_0 + \int_{t_0}^{t} (A(s)Y_n(s) + B(s))ds.$$

Soit J un segment inclus dans I. Puisque A et B sont continues sur J qui est compact, on peut donc poser,

$$\alpha = \sup_{s\in J}\|A(s)\|, \ \beta = \sup_{s\in J}\|B(s)\|,$$

où la norme utilisée ci-dessus est la norme d'opérateur sur $M_d(\mathbb{R})$, c'est-à-dire la norme définie pour tout $M \in M_d(\mathbb{R})$ par,

$$\|M\| := \sup\left\{\frac{\|MX\|}{\|X\|}, \ X \neq 0\right\}.$$

Montrons alors, par récurrence sur $n \in \mathbb{N}$, que, pour tout $t \in J$,

$$\|Y_{n+1}(t) - Y_n(t)\| \leq (\alpha\|Y_0\| + \beta)\frac{\alpha^n}{(n+1)!}|t - t_0|^{n+1}.$$

Initialisation. Soit $t \in J$. Si $t \geq t_0$ alors,

$$\begin{aligned}
\|Y_1(t) - Y_0\| &\leq \int_{t_0}^{t} \|A(s)Y_0 + B(s)\|ds \\
&\leq \int_{t_0}^{t} (\|A(s)\|\|Y_0\| + \|B(s)\|)ds \\
\|Y_1(t) - Y_0\| &\leq (\alpha\|Y_0\| + \beta)(t - t_0).
\end{aligned}$$

De même, si $t \leq t_0$,

$$\|Y_1(t) - Y_0\| \leq \int_{t}^{t_0} \|A(s)Y_0 + B(s)\|ds \leq (\alpha\|Y_0\| + \beta)(t_0 - t).$$

D'où le résultat.

Hérédité. Supposons le résultat établi pour un rang $n \in \mathbb{N}$ fixé. Soit $t \in J$ tel que $t \geq t_0$. Alors, par deux inégalités triangulaires, et par définition de α,

$$\|Y_{n+2}(t) - Y_{n+1}(t)\| \leq \alpha \int_{t_0}^{t} \|Y_{n+1}(s) - Y_n(s)\|ds.$$

Or, par hypothèse de récurrence, pour tout $s \in [t_0, t]$,

$$\|Y_{n+1}(s) - Y_n(s)\| \leq (\alpha\|Y_0\| + \beta)\frac{\alpha^n}{(n+1)!}(s - t_0)^{n+1}.$$

Des deux inégalités précédentes, on en déduit que,

$$\|Y_{n+2}(t) - Y_{n+1}(t)\| \leq (\alpha\|Y_0\| + \beta)\frac{\alpha^{n+1}}{(n+1)!}\int_{t_0}^{t} (s - t_0)^{n+1}ds.$$

Et donc,

$$\|Y_{n+2}(t) - Y_{n+1}(t)\| \leq (\alpha\|Y_0\| + \beta)\frac{\alpha^{n+1}}{(n+2)!}(t - t_0)^{n+2}$$

car,

$$\int_{t_0}^{t} (s - t_0)^{n+1}ds = \frac{(t - t_0)^{n+2}}{n+2}.$$

En procédant de même pour $t \leq t_0$, on prouve ainsi l'hérédité. Par récurrence, il s'en suit l'inégalité recherchée. En particulier, en notant,

$$L = \sup_{t\in J}|t - t_0|$$

il vient, pour tout $n \in \mathbb{N}$,

$$\sup_J\|Y_{n+1} - Y_n\| \leq (\alpha\|Y_0\| + \beta)\frac{\alpha^n}{(n+1)!}L^{n+1}.$$

Sachant que l'expression de droite est le terme général d'une série convergente (par critère de d'Alembert), on en déduit que la série,

$$\sum_{n\in\mathbb{N}}(Y_{n+1}-Y_n)$$

converge normalement sur J, puis que la suite de fonctions $(Y_n)_{n\in\mathbb{N}}$ converge uniformément sur tout segment contenu dans I, vers une fonction $Y : I \to \mathbb{R}^d$ continue. Soit $t \in I$. Puisque $(AY_n+B)_{n\in\mathbb{N}}$ est une suite de fonctions continues qui converge uniformément sur $[t, t_0]$ vers $AY + B$, il vient,

$$\lim_{n\to+\infty}\int_{t_0}^{t}(A(s)Y_n(s)+B(s))ds=\int_{t_0}^{t}(A(s)Y(s)+B(s))ds.$$

Et donc,

$$\lim_{n\to+\infty}Y_{n+1}(t)=Y_0+\int_{t_0}^{t}(A(s)Y(s)+B(s))ds.$$

Or, $\lim\limits_{n\to+\infty}Y_{n+1}(t)=Y(t)$. Par unicité de la limite, on en déduit que,

$$Y(t)=Y_0+\int_{t_0}^{t}(A(s)Y(s)+B(s))ds.$$

En particulier, $Y(0) = Y_0$ et, d'après le théorème fondamental de l'analyse, Y est dérivable sur I, avec, pour tout $t \in I$,

$$Y'(t)=A(t)Y(t)+B(t).$$

Autrement dit, Y est solution de (C).

Unicité. Passons à l'unicité. Soient Y et Z deux solutions de (C). Soit J un segment contenu dans I. De nouveau, on note,

$$\alpha=\sup_{s\in J}\|A(s)\|.$$

Montrons par récurrence sur $n \in \mathbb{N}$, que,

$$\forall t\in J,\ \|Y(t)-Z(t)\|\leq\frac{\alpha^n}{n!}|t-t_0|^n\|Y-Z\|_{\infty,J}.$$

Initialisation. Soit $t \in J$. Alors,

$$Y(t)=Y_0+\int_{t_0}^{t}(A(s)Y(s)+B(s))ds,\ Z(t)=Y_0+\int_{t_0}^{t}(A(s)Z(s)+B(s))ds.$$

Et donc,

$$\|Y(t)-Z(t)\|\leq\alpha|t-t_0|\|Y-Z\|_{\infty,J}.$$

Hérédité. Supposons ensuite le résultat établi pour un rang $n \in \mathbb{N}$ fixé. Soit $t \in J$ tel que $t \geq t_0$. Alors,

$$\|Y(t)-Z(t)\|\leq\frac{\alpha^{n+1}}{n!}\|Y-Z\|_{\infty,J}\int_{t_0}^{t}(s-t_0)^n ds.$$

Et donc,

$$\|Y(t)-Z(t)\|\leq\frac{\alpha^{n+1}}{(n+1)!}\|Y-Z\|_{\infty,J}(t-t_0)^{n+1}$$

car,

$$\int_{t_0}^{t}(s-t_0)^n ds=\frac{(t-t_0)^{n+1}}{(n+1)!}.$$

En procédant de même pour $t \leq t_0$, on prouve ainsi l'hérédité. Par récurrence, il s'en suit l'inégalité recherchée. En faisant tendre n vers l'infini dans celle-ci, on en conclut que, pour tout $t \in J$, $Y(t) = Z(t)$. Ainsi, Y et Z coïncident sur tout segment inclus dans I, ce qui prouve finalement que $Y = Z$. □

Commentaires personnels. Développement très sympathique, qui se comprend et se retient facilement. Selon moi, ce dernier apporte une valeur ajoutée par rapport à la preuve du théorème de Cauchy-Lipschitz général, puisqu'on se passe du théorème du point fixe. De plus, le théorème de Cauchy-Lipschitz linéaire figure sur le programme du concours.

Théorème de réarrangement de Riemann

[FG] **Théorème 1 (Réarrangement de Riemann).** Soit $(a_n)_{n\in\mathbb{N}}$ une suite de nombres réels telle que la série de terme général a_n soit semi-convergente. Alors, pour tout $\alpha \in \mathbb{R}$, il existe $\sigma \in \mathfrak{S}(\mathbb{N})$ telle que la série de terme général $a_{\sigma(n)}$ soit convergente de somme α.

Preuve. Soit $\alpha \in \mathbb{R}$.

Étape n°1. Tout d'abord, notons,

$$E^+ = \{n \in \mathbb{N},\ a_n \geq 0\},\ E^- = \{n \in \mathbb{N},\ a_n < 0\}.$$

Alors, E^+ et E^- sont infinis, car sinon la suite $(a_n)_{n\in\mathbb{N}}$ serait de signe constant à partir d'un certain rang et donc la convergence absolue de la série $\sum_{n\in\mathbb{N}} a_n$ serait équivalente à sa convergence, ce qui est impossible puisque cette série est semi-convergente. De plus, puisque pour tout $n \in \mathbb{N}$,

$$\min(0, a_n) = \frac{a_n - |a_n|}{2},\ \max(0, a_n) = \frac{a_n + |a_n|}{2}$$

les séries $\sum_{n\in\mathbb{N}} \min(0, a_n)$ et $\sum_{n\in\mathbb{N}} \max(0, a_n)$ divergent en tant que somme d'une série convergente et d'une série divergente.

Étape n°2. On construit ensuite la permutation σ. Posons $\sigma(0) = 0$, et pour tout $n \in \mathbb{N}^*$,

- Si $\sum_{k=0}^{n-1} a_{\sigma(k)} \leq \alpha$, on va ajouter un terme positif à cette somme. Pour cela, on choisit pour $\sigma(n)$ le plus petit entier appartenant à E^+ qui soit distinct de $\sigma(0), \ldots, \sigma(n-1)$ (un tel entier existe car E^+ est infini).
- Si $\sum_{k=0}^{n-1} a_{\sigma(k)} > \alpha$, on va ajouter un terme strictement négatif à cette somme. Pour cela, on choisit pour $\sigma(n)$ le plus petit entier appartenant à E^- qui soit distinct de $\sigma(0), \ldots, \sigma(n-1)$ (un tel entier existe car E^- est infini).

On définit ainsi une application,

$$\begin{array}{rccc} \sigma & : & \mathbb{N} & \longrightarrow & \mathbb{N} \\ & & n & \longmapsto & \sigma(n) \end{array}$$

qui est injective par construction. Il reste à montrer que σ est surjective. Tout d'abord, montrons que $E^+ \subset \sigma(\mathbb{N})$. On raisonne par l'absurde en supposant qu'il existe $N \in E^+$ tel que $N \notin \sigma(\mathbb{N})$. Pour tout $k \in \mathbb{N}$, si $\sigma(k) \in E^+$ alors $N \notin \{\sigma(0), \ldots, \sigma(k-1)\}$ (car $N \notin \sigma(\mathbb{N})$) et donc, par construction de σ,

$$N \geq \min\{q \in E^+,\ q \notin \{\sigma(0), \ldots, \sigma(k-1)\}\} = \sigma(k).$$

Cette inégalité est même stricte car $N \notin \sigma(\mathbb{N})$. D'où $\sigma(\mathbb{N}) \cap E^+ \subset [\![0, N-1]\!]$. En particulier, $\sigma(\mathbb{N}) \cap E^+$ est fini. Par injectivité de σ, il existe donc $n_0 \in \mathbb{N}$ tel que, pour tout $n \geq n_0$, $\sigma(n) \notin E^+$. En effet, sinon on aurait,

$$\forall n_0 \in \mathbb{N}, \exists \varphi(n_0) \geq n_0,\ \sigma(\varphi(n_0)) \in \sigma(\mathbb{N}) \cap E^+$$

et $\varphi(\mathbb{N})$ serait donc une partie infinie de $\mathbb{N}$ dont l'image par σ serait finie, ce qui contredirait alors l'injectivité de σ. Puisque $\mathbb{N} \setminus E^+ = E^-$, il s'en suit que,

$$\forall n \geq n_0,\ \sigma(n) \in E^- \tag{1}$$

et donc,

$$\forall n \geq n_0,\ \sum_{k=0}^{n-1} a_{\sigma(k)} > \alpha. \tag{2}$$

Ainsi, $\sum_{k\in\mathbb{N}} a_{\sigma(k)}$ est une série à termes négatifs à partir d'un certain rang dont la suite de ses sommes partielles est minorée, et donc elle converge. Montrons alors que la série $\sum_{n\in\mathbb{N}} \min(0, a_n)$ converge, ce qui contredira les résultats de l'étape n°1. Notons ϕ l'unique application strictement croissante de $\mathbb{N}$ dans E^-. Il s'agit de la suite définie par $\phi(0) = \min E^-$ et, pour tout $k \in \mathbb{N}$,

$$\phi(k+1) = \min E^- \setminus \{\phi(0), \ldots, \phi(k)\}.$$

On a montré précédemment que $\sigma(\mathbb{N}) \cap E^+$ était fini. Notons p son cardinal. Par construction de σ, et d'après l'assertion (1), l'ensemble $\sigma([\![0, n_0 - 1]\!])$ est constitué des p éléments de $\sigma(\mathbb{N}) \cap E^+$ et de $n_0 - p$ éléments de E^-. Or, par construction de σ, et d'après l'inégalité (2), la suite $(\sigma(n))_{n\geq n_0}$ énumère dans l'ordre strictement croissant les éléments de $E^- \setminus \sigma([\![0, n_0 - 1]\!])$. De ces points, on en déduit que :

$$\forall n \geq n_0,\ \sigma(n) = \phi(n - p).$$

En particulier, les séries $\sum_{k\in\mathbb{N}} a_{\sigma(k)}$ et $\sum_{k\in\mathbb{N}} a_{\phi(k)}$ sont donc de même nature, puisqu'elles diffèrent d'un nombre fini de termes. Or, pour tout $k \in \mathbb{N}$,

$$\sum_{k=0}^{n} a_{\phi(k)} \leq \sum_{k=0}^{n} \min(0, a_k) \underset{n\to+\infty}{\longrightarrow} -\infty.$$

Donc, la série $\sum_{k\in\mathbb{N}} a_{\phi(k)}$ diverge, ce qui contredit le fait que les séries $\sum_{k\in\mathbb{N}} a_{\sigma(k)}$ et $\sum_{k\in\mathbb{N}} a_{\phi(k)}$ soient de même nature. Ainsi, $E^+ \subset \sigma(\mathbb{N})$. De façon analogue, on montre que $E^- \subset \sigma(\mathbb{N})$, ce qui permet de conclure que σ est surjective.

Étape n°3. Montrons que la série $\sum_{n\in\mathbb{N}} a_{\sigma(n)}$ est convergente de somme α. Soit $\varepsilon > 0$. Puisque la série $\sum_{n\in\mathbb{N}} a_n$ converge, la suite $(a_n)_{n\in\mathbb{N}}$ converge vers 0. D'où, l'existence de $n_1 \in \mathbb{N}$ tel que, pour tout $n \geq n_1$, $|a_n| \leq \varepsilon$. Or, σ étant bijective, l'ensemble $\{k \in \mathbb{N},\ \sigma(k) < n_1\}$ est donc une partie non vide et finie de $\mathbb{N}$. En particulier, cet ensemble admet un maximum. Soit $n_0 > \max\{k \in \mathbb{N},\ \sigma(k) < n_1\}$. Alors, pour tout $n \geq n_0$, $\sigma(n) \geq n_1$ et donc $|a_{\sigma(n)}| \leq \varepsilon$.

Montrons l'existence de $N \geq n_0$ tel que $\sigma(N) \in E^+$ et $\sigma(N+1) \in E^-$. Supposons le contraire. On aurait alors, pour tout $N \geq n_0$,

$$\sigma(N) \in E^+ \Longrightarrow \sigma(N+1) \notin E^-.$$

Or, puisque σ est surjective et que E^+ est infini, il existe $N \geq n_0$ tel que $\sigma(N) \in E^+$. L'implication ci-dessus entraîne, par récurrence immédiate, que $\sigma([\![N+1, +\infty[\![) \subset \mathbb{N} \setminus E^-$. Sachant que σ est surjective, il s'en suit que $E^- \subset \sigma([\![1, N]\!])$, ce qui contredit le fait que E^- soit infini.

D'où l'existence de $N \geq n_0$ tel que $\sigma(N) \in E^+$ et $\sigma(N+1) \in E^-$. Pour tout $n \in \mathbb{N}$, posons,

$$S_n = \sum_{k=0}^{n} a_{\sigma(k)}.$$

Alors, $S_{N-1} \leq \alpha$ car $\sigma(N) \in E^+$ et $S_N > \alpha$ car $\sigma(N+1) \in E^-$. En particulier, $S_N > S_{N-1}$. D'où,

$$S_N - S_{N-1} = |S_N - S_{N-1}| = |a_{\sigma(N)}| \leq \varepsilon,$$

car $N \geq n_0$. A fortiori,

$$\alpha - \varepsilon \leq \alpha < S_N < S_{N-1} + \varepsilon \leq \alpha + \varepsilon$$

et donc $S_N \in [\alpha-\varepsilon, \alpha+\varepsilon]$. Soit $n > N$. Montrons que $S_n \leq \alpha+\varepsilon$. Supposons le contraire. Alors, $|a_{\sigma(n)}| \leq \varepsilon$ car $n \geq N \geq n_0$, et donc,

$$S_{n-1} = S_n - a_{\sigma(n)} \geq S_n - \varepsilon > \alpha.$$

Par suite, $\sigma(n) \in E^-$, et donc $S_{n-1} = S_n - a_{\sigma(n)} > S_n > \alpha + \varepsilon$ car $a_{\sigma(n)} < 0$. Ainsi, $n-1$ est un entier tel que $S_{n-1} > \alpha + \varepsilon$. Sous réserve que $n-1$ soit toujours strictement supérieur à N, on peut appliquer ce résultat à $n-1$ pour obtenir que $S_{n-2} > \alpha + \varepsilon$. Par récurrence descendante, ce résultat permet d'obtenir que $S_N > \alpha+\varepsilon$, ce qui contredit $S_N \in [\alpha-\varepsilon, \alpha+\varepsilon]$. D'où, $S_n \leq \alpha+\varepsilon$.

Soit $n > N$. Montrons que $S_n \geq \alpha - \varepsilon$. Supposons le contraire. Alors, $|a_{\sigma(n)}| \leq \varepsilon$ car $n \geq N \geq n_0$, et donc,

$$S_{n-1} = S_n - a_{\sigma(n)} \leq S_n + \varepsilon < \alpha.$$

Par suite, $\sigma(n) \in E^+$, et donc $S_{n-1} = S_n - a_{\sigma(n)} < S_n < \alpha - \varepsilon$ car $a_{\sigma(n)} > 0$. Ainsi, $n-1$ est un entier tel que $S_{n-1} < \alpha - \varepsilon$. Sous réserve que $n-1$ soit toujours strictement supérieur à N, on peut appliquer ce résultat à $n-1$ pour obtenir que $S_{n-2} < \alpha - \varepsilon$. Par récurrence descendante, ce résultat permet d'obtenir que $S_N < \alpha-\varepsilon$, ce qui contredit $S_N \in [\alpha-\varepsilon, \alpha+\varepsilon]$. D'où, $S_n \geq \alpha-\varepsilon$.

Finalement, on obtient donc que, pour tout $n \geq N$, $|S_n - \alpha| \leq \varepsilon$. Ceci valant pour tout $\varepsilon > 0$, on en conclut que la série $\sum_{n\in\mathbb{N}} a_{\sigma(n)}$ est convergente de somme α. □

Commentaires personnels. Développement à la fois trop technique et trop long pour mon niveau. J'ai finalement renoncé à le présenter pour ces raisons. Je trouvais ce résultat formidable mais ce n'était pas raisonnable de préparer un développement aussi difficile pour une seule leçon.

Théorème des extrema liés

La preuve du théorème suivant provient du cours de Stéphane Rigat, dispensé en master de préparation à l'agrégation de mathématiques, à l'université d'Aix-Marseille.

[FR] **Théorème 1 (Théorème des extrema liés).** Soient U un ouvert de $\mathbb{R}^n$, f, $g_1, \dots, g_p$ des fonctions réelles de classe C^1 de U dans $\mathbb{R}$ et,

$$X = \{x \in U,\ g_1(x) = \cdots = g_p(x) = 0\}.$$

Si la restriction de f à X admet un extremum local en $x_0 \in X$ et si les différentielles $Dg_1(x_0), \dots, Dg_p(x_0)$ sont des formes linéaires indépendantes sur $\mathbb{R}^n$, alors il existe des réels $\lambda_1, \dots, \lambda_p$, appelés multiplicateurs de Lagrange, tels que :

$$Df(x_0) = \lambda_1 Dg_1(x_0) + \cdots + \lambda_p Dg_p(x_0).$$

Preuve. Remarquons tout d'abord que si $p = n$, alors $(Dg_1(x_0), \dots, Dg_n(x_0))$ est une base de $(\mathbb{R}^n)^*$ (pour des raisons de cardinalité), et le résultat est donc vrai sans supposer que la restriction de f à X admet un extremum local.

Par conséquent, on peut supposer $p < n$. Pour prouver le théorème, on raisonne par contradiction en supposant que $Df(x_0) \notin \text{Vect}(Dg_1(x_0), \dots, Dg_p(x_0))$. La famille $(Dg_1(x_0), \dots, Dg_p(x_0), Df(x_0))$ est donc libre. D'après le théorème de la base incomplète, il existe alors $\varphi_{p+2}, \dots, \varphi_n \in (\mathbb{R}^n)^*$ tels que,

$$\mathcal{B} := (Dg_1(x_0), \dots, Dg_p(x_0), Df(x_0), \varphi_{p+2}, \dots, \varphi_n)$$

soit une base de $(\mathbb{R}^n)^*$. D'autre part, l'hypothèse sur f assure l'existence de $\alpha > 0$ tel que la restriction de f à $B(x_0, \alpha) \cap X$ passe par un extremum global en x_0. Considérons ensuite l'application,

$$\begin{array}{rccc} F & : & U & \longrightarrow & \mathbb{R}^n \\ & & x & \longmapsto & (g_1(x), \dots, g_p(x), f(x), \varphi_{p+2}(x), \dots, \varphi_n(x)) \end{array}$$

Nous allons appliquer le théorème d'inversion locale à F. Tout d'abord, F est de classe C^1 sur U avec, pour tout $x \in U$,

$$DF(x) = (Dg_1(x), \dots, Dg_p(x), Df(x), \varphi_{p+2}, \dots, \varphi_n).$$

En particulier, $J_F(x_0)$ est la matrice de passage de la base canonique de $(\mathbb{R}^n)^*$ à la base $\mathcal{B}$, c'est-à-dire la matrice dont les colonnes contiennent les coordonnées des vecteurs de $\mathcal{B}$ dans la base canonique de $(\mathbb{R}^n)^*$. En effet, en notant $(e_1^*, \dots, e_n^*)$ la base duale de $(\mathbb{R}^n)^*$, il vient,

$$\begin{cases} \forall j \in [\![1, n]\!],\ Dg_j(x_0) = \displaystyle\sum_{i=1}^n \partial_i g_j(x_0) e_i^*, \\ Df(x_0) = \displaystyle\sum_{i=1}^n \partial_i f(x_0) e_i^*, \\ \forall j \in [\![p+2, n]\!],\ \varphi_j = D\varphi_j(x_0) = \displaystyle\sum_{i=1}^n \partial_i \varphi_j(x_0) e_i^*. \end{cases}$$

Ainsi, $J_F(x_0)$ est bien la matrice de passage de la base canonique de $(\mathbb{R}^n)^*$ à la base $\mathcal{B}$. En particulier, $J_F(x_0) \in GL_n(\mathbb{R})$. Par suite, le théorème d'inversion locale (appliquée à la restriction de F à $U \cap B(x_0, \alpha)$), assure l'existence d'un ouvert $V \subset U \cap B(x_0, \alpha)$ de $\mathbb{R}^n$ contenant x_0 tel que,

$$\begin{array}{ccc} V & \longrightarrow & F(V) \\ x & \longmapsto & F(x) \end{array}$$

soit un C^1-difféomorphisme. En particulier, $F(V)$ est un ouvert de $\mathbb{R}^n$ qui contient $F(x_0)$, et donc il existe $\varepsilon > 0$ tel que $B(F(x_0), \varepsilon) \subset F(V)$. En notant e_{p+1} le $p+1$-ième vecteur de la base canonique de $\mathbb{R}^n$, il vient ensuite,

$$F(x_0) + \frac{\varepsilon}{2} e_{p+1} \in B(F(x_0), \varepsilon) \subset F(V)$$

c'est-à-dire qu'il existe $b \in V$ tel que $F(b) = F(x_0) + \frac{\varepsilon}{2} e_{p+1}$. Or,

$$\begin{cases} F(b) = (g_1(b), \dots, g_p(b), f(b), \varphi_{p+2}(b), \dots, \varphi_n(b)), \\ F(x_0) + \frac{\varepsilon}{2} e_{p+1} = (0, \dots, 0, f(x_0) + \frac{\varepsilon}{2}, \varphi_{p+2}(x_0), \dots, \varphi_n(x_0)). \end{cases}$$

Donc,

$$\begin{cases} g_1(b) = \cdots = g_p(b) = 0 \\ f(b) = f(x_0) + \frac{\varepsilon}{2} \end{cases}$$

ce qui prouve que $b \in X$ et $f(b) > f(x_0)$. De même, puisque,

$$F(x_0) - \frac{\varepsilon}{2} e_{p+1} = (0, \dots, 0, f(x_0) - \frac{\varepsilon}{2}, \varphi_{p+2}(x_0), \dots, \varphi_n(x_0)) \in F(V)$$

on obtient l'existence de $a \in X \cap V$ tel que $f(a) = f(x_0) - \frac{\varepsilon}{2}$. En particulier,

$$f(a) < f(x_0) < f(b)$$

avec $a, b \in X \cap V \subset X \cap B(x_0, \alpha)$, ce qui contredit le fait que la restriction de f à $X \cap B(x_0, \alpha)$ passe par un extremum global en x_0. □

Remarque. Le théorème est énoncé sans preuve dans [FR].

Commentaires personnels. Développement très sympathique, qui se comprend et se retient facilement, et qui se présente dans de nombreuses leçons. De plus, le théorème des extrema liés figure sur le programme du concours.

Références

[AP] : Cours d'analyse de A. Pommellet (ellipses).

[BB] : Analyse pour l'agrégation de mathématiques de J. Bernis et L. Bernis (ellipses).

[BM] : Objectif agrégation de V. Beck, J. Malick et G. Peyré (H & K).

[FB] : Équations différentielles de F. Berthelin (Cassini).

[FG] : Exercices de mathématiques oraux x-ens, analyse 1 de S. Francinou, H. Gianella et S. Nicolas (Cassini).

[FG] : Exercices de mathématiques oraux x-ens, analyse 2 de S. Francinou, H. Gianella et S. Nicolas (Cassini).

[FG] : Exercices de mathématiques oraux x-ens, analyse 3 de S. Francinou, H. Gianella et S. Nicolas (Cassini).

[FG] : Exercices de mathématiques oraux x-ens, analyse 4 de S. Francinou, H. Gianella et S. Nicolas (Cassini).

[FR] : Petit guide de calcul différentiel à l'usage de la licence et de l'agrégation de F. Rouvière (Cassini).

[GH] : Mesure, intégration, probabilités de T. Gallouët et R. Herbin (ellipses).

[GK] : De l'intégration aux probabilités de O. Garet et A. Kurtzman (ellipses).

[IP] : L'oral à l'agrégation de mathématiques - Une sélection de développements de L. Isenmann et T. Pecatte.

[JD] : Analyse numérique et équations différentielles de J. Demailly (EDP Sciences).

[JR] : Mathématiques pour l'agrégation, algèbre et géométrie de J. Rombaldi (De Boeck Supérieur).

[JR] : Mathématiques pour l'agrégation, analyse et probabilités de J. Rombaldi (De Boeck Supérieur).

[LL] : 131 développements pour l'oral de P. Le Barbenchon, D. Lesesvre, P. Montagnon et T. Pierron (Dunod).

[XG] : Algèbre et probabilités de X. Gourdon (ellipses).

[XG] : Analyse de X. Gourdon (ellipses).

Références détaillées

Développement ou résultat utile	Références
Algorithme du gradient à pas optimal	[FG] : Analyse 4, chapitre 1, exercice 21.
Base hilbertienne de polynômes orthogonaux	[BM] : Chapitre 3, théorème 3.49 et exercice 3.7.
Équation de la chaleur	[IP] : Développement 27.
Équation de Sylvester	[XG] : Analyse, chapitre 6, problème 7.
Équivalence des normes en dimension finie	[XG] : Analyse, chapitre 1, section 5, théorème 3.
Théorème de Riesz	[XG] : Analyse, chapitre 1, section 5, exercice 9.
Espace de Bergman	[BB] : Chapitre II, développement 18.
Expression des $\zeta(2k)$	[FG] : Analyse 2, chapitre 4, exercice 4.18.
Intégrales de Wallis	[JR] : Analyse, chapitre 8, exercice 8.11.
Formule de Stirling (par les intégrales de Wallis)	[JR] : Analyse, chapitre 8, section 8.10.
Formule de Stirling (par le théorème central limite)	[LL] : Développement 95.
Méthode de Laplace	[FR] : Chapitre 6, exercice 113.
Méthode de Newton	[FR] : Chapitre 4, exercice 49.
Nombres de Bell	[BB] : Chapitre II, développement 39.
Projection sur un convexe fermé	[GH] : Chapitre 6, théorème 6.47 et proposition 6.49.
Prolongement de la fonction Gamma d'Euler	[BM] : Chapitre 2, exercice 2.10.
Série des inverses des nombres premiers	[JR] : Algèbre, chapitre 11, exercice 11.8 et exercice 11.9.
Système de Lotka-Volterra	[FB] : Chapitre 5, section 3.2.
Système de Lotka-Volterra	[FG] : Analyse 4, chapitre 3, exercice 3.32.
Théorème central limite	[BB] : Développement 30.
Théorème d'Ascoli	[BB] : Chapitre II, développement 1.
Théorème d'Hadamard-Lévy	[BB] : Chapitre II, développement 6.
Théorème d'inversion locale	[FR] : Chapitre 5, théorème 5.1.
Théorème de Cauchy-Lipschitz	[JD] : Chapitre V, section 3.2. et section 3.3.
Théorème de Cauchy-Lipschitz linéaire	[AP] : Chapitre 35, théorème 35.2.1.
Théorème de réarrangement de Riemann	[FG] : Analyse 1, chapitre 3, exercice 3.48.
Théorème des extrema liés	∅

Annexe

Accroissements finis

Théorème 1 (Théorème de Rolle). Soient $a, b \in \mathbb{R}$ tels que $a < b$. Soit $g : [a, b] \to \mathbb{R}$ une fonction continue sur $[a, b]$ et dérivable sur $]a, b[$ telle que $g(a) = g(b)$. Alors, il existe $c \in]a, b[$ tel que $g'(c) = 0$.

Preuve. Puisque g est continue sur un compact, g est bornée et atteint ses bornes, c'est-à-dire qu'il existe $x_1, x_2 \in [a, b]$ tels que, pour tout $x \in [a, b]$,

$$g(x_1) \leq g(x) \leq g(x_2).$$

Si $(x_1, x_2) \in \{a, b\}^2$ alors $g(x_1) = g(x_2)$, et donc g est constante d'après l'inégalité ci-dessus. En particulier, $g' = 0$ sur $]a, b[$ et donc n'importe quel choix de c dans $]a, b[$ convient. Sinon, $x_1 \in]a, b[$ ou $x_2 \in]a, b[$. On distingue deux cas analogues.

• Cas n°1 : $x_1 \in]a, b[$. Alors, pour tout $x \in]a, x_1[$,

$$\frac{g(x) - g(x_1)}{x - x_1} \leq 0$$

car $g(x) - g(x_1) \geq 0$ et $x - x_1 \leq 0$. En faisant tendre x vers x_1, on obtient donc $g'(x_1) \leq 0$. D'autre part, pour tout $x \in]x_1, b[$,

$$\frac{g(x) - g(x_1)}{x - x_1} \geq 0$$

car $g(x) - g(x_1) \geq 0$ et $x - x_1 \geq 0$. En faisant tendre x vers x_1, on obtient donc $g'(x_1) \geq 0$. D'où $g'(x_1) = 0$, et donc $c = x_1$ convient à notre problème.

• Cas n°2 : $x_2 \in]a, b[$. Alors, pour tout $x \in]a, x_2[$,

$$\frac{g(x) - g(x_2)}{x - x_2} \geq 0$$

car $g(x) - g(x_2) \leq 0$ et $x - x_2 \leq 0$. En faisant tendre x vers x_2, on obtient donc $g'(x_2) \geq 0$. D'autre part, pour tout $x \in]x_2, b[$,

$$\frac{g(x) - g(x_2)}{x - x_2} \geq 0$$

car $g(x) - g(x_2) \geq 0$ et $x - x_2 \geq 0$. En faisant tendre x vers x_2, on obtient donc $g'(x_2) \geq 0$. D'où $g'(x_2) = 0$, et donc $c = x_2$ convient à notre problème. □

Théorème 2 (Théorème des accroissements finis). Soient $a, b \in \mathbb{R}$ tels que $a < b$. Soit $f : [a, b] \to \mathbb{R}$ une fonction continue sur $[a, b]$ et dérivable sur $]a, b[$. Alors, il existe $c \in \mathbb{R}$ tel que,

$$f(b) - f(a) = f'(c)(b - a).$$

Preuve. L'idée de la preuve est d'appliquer le théorème de Rolle à une fonction bien choisie. Plutôt que de sortir cette fonction de notre chapeau, nous allons esquisser un raisonnement pour parvenir à cette fonction. On cherche une fonction $g : [a, b] \to \mathbb{R}$ continue sur $[a, b]$ et dérivable sur $]a, b[$ telle que $g(a) = g(b)$, et vérifiant, pour tout $x \in]a, b[$,

$$g'(x) = \frac{f(b) - f(a)}{b - a} + f'(x).$$

En effet, en appliquant le théorème de Rolle à g, on obtiendra alors le résultat souhaité. Dès lors, on peut penser à la fonction g définie pour tout $x \in [a, b]$ par,

$$g(x) = \frac{f(b) - f(a)}{b - a}(x - a) + f(x) - f(a)$$

qui est continue sur $[a, b]$ et dérivable sur $]a, b[$ telle que $g(a) = g(b) = 0$, et, pour tout $x \in]a, b[$,

$$g'(x) = \frac{f(b) - f(a)}{b - a} + f'(x).$$

□

Remarque. On aurait pu choisir n'importe quel primitive de,

$$x \mapsto \frac{f(b) - f(a)}{b - a} + f'(x).$$

En effet, toute primitive de cette fonction diffère de g d'une constante, et prend donc aussi la même valeur en a et en b.

Proposition 3 (Inégalité des accroissements finis). Soient $a, b \in \mathbb{R}$ tels que $a < b$. Soit $f : [a, b] \to \mathbb{R}$ une fonction continue sur $[a, b]$ et dérivable sur $]a, b[$. On suppose qu'il existe $M \in \mathbb{R}$ tel que, pour tout $x \in]a, b[$, $|f'(x)| \leq M$. Alors,

$$|f(b) - f(a)| \leq M|b - a|.$$

Preuve. D'après le théorème des accroissements finis, il existe $c \in]a,b[$ tel que,

$$f(b) - f(a) = f'(c)(b-a).$$

En particulier,

$$|f(b) - f(a)| \leq |f'(c)||b-a| \leq M|b-a|.$$

□

On démontre ensuite l'inégalité des accroissements finis pour une fonction d'une variable réelle et à valeurs vectorielles, ce qui permettra finalement de démontrer l'inégalité des accroissements finis pour une fonction d'une variable vectorielle et à valeurs vectorielles.

Soient $(E, \|\cdot\|_E)$ et $(F, \|\cdot\|_F)$ deux espaces vectoriels normés.

Théorème 4 (Inégalité des accroissements finis). Soient $a, b \in \mathbb{R}$ tels que $a < b$. Soient $f : [a,b] \to F$ et $g : [a,b] \to \mathbb{R}$ deux fonctions continues sur $[a,b]$ et dérivables sur $]a,b[$. On suppose que, pour tout $t \in]a,b[$, $\|f'(t)\|_F \leq g'(t)$. Alors,

$$\|f(b) - f(a)\|_F \leq g(b) - g(a).$$

En particulier, s'il existe $M \in \mathbb{R}$ tel que, pour tout $t \in]a,b[$, $\|f'(t)\| \leq M$ alors,

$$\|f(b) - f(a)\|_F \leq M|b-a|.$$

Preuve. Soit $\varepsilon > 0$. Montrons que,

$$\|f(b) - f(a)\|_F \leq g(b) - g(a) + \varepsilon(b-a) + \varepsilon.$$

Posons,

$$E = \{t \in [a,b],\ \|f(t) - f(a)\|_F \leq g(t) - g(a) + (t-a)\varepsilon + \varepsilon\}.$$

Tout d'abord, E est non vide car $a \in E$. De plus, E est une partie fermée de $\mathbb{R}$ en tant qu'image réciproque de $]-\infty, 0[$, qui est une partie fermée de $\mathbb{R}$, par l'application continue,

$$\begin{array}{rccc} u & : & [a,b] & \longrightarrow & \mathbb{R} \\ & & t & \longmapsto & \|f(t) - f(a)\|_F - [g(t) - g(a) + (t-a)\varepsilon + \varepsilon] \end{array}$$

Notons $c = \sup E \leq b$. Puisque E est fermé, $c \in E$. Montrons que $c > a$. Par continuité de u, il existe $r > 0$ tel que, pour tout $t \in [a, a+r]$,

$$|u(t) - u(0)| \leq \frac{\varepsilon}{2}$$

et donc,

$$u(t) \leq u(0) + \frac{\varepsilon}{2} = -\frac{\varepsilon}{2} \leq 0.$$

D'où, $[a, a+r] \subset E$, ce qui implique $c \geq a + r > a$. Montrons que $c = b$. On raisonne par l'absurde en supposant que $c < b$. Alors, $c \in]a,b[$, et donc f et g sont dérivables au point c. En particulier, il existe $\delta > 0$ tel que, pour tout $h \in]0, \delta]$,

$$(i)\ \left\| \frac{f(c+h) - f(c)}{h} - f'(c) \right\|_F \leq \frac{\varepsilon}{2},$$

$$(ii)\ \left| \frac{g(c+h) - g(c)}{h} - g'(c) \right| \leq \frac{\varepsilon}{2}.$$

Soit $h \in]0, \delta]$. Montrons que $c + h \in E$, ce qui contredira $c = \sup(E)$. Par inégalité triangulaire,

$$\|f(c+h) - f(a)\|_F \leq \|f(c+h) - f(c)\|_F + \|f(c) - f(a)\|_F. \tag{1}$$

D'après (i), puis par hypothèse sur f et g,

$$\|f(c+h) - f(c)\|_F \leq h\left(\|f'(c)\|_F + \frac{\varepsilon}{2}\right) \leq hg'(c) + \frac{\varepsilon}{2}h.$$

De plus, d'après (ii),

$$h|g'(c)| \leq |g(c+h) - g(c)| + h\frac{\varepsilon}{2}.$$

Or, pour tout $t \in]a,b[$, $g'(c) \geq \|f'(c)\|_F \geq 0$. En particulier, $g'(c) \geq 0$ et g est croissante sur $]a,b[$, et donc $g(c+h) - g(c) \geq 0$. On peut donc enlever les valeurs absolues dans l'inégalité ci-dessus, ce qui donne,

$$hg'(c) \leq g(c+h) - g(c) + h\frac{\varepsilon}{2}.$$

Et donc,

$$\|f(c+h) - f(c)\|_F \leq g(c+h) - g(c) + h\varepsilon.$$

D'autre part,

$$\|f(c) - f(a)\|_F \leq g(c) - g(a) + \varepsilon(c-a) + \varepsilon$$

car $c \in E$. En utilisant les deux inégalités précédentes dans l'inégalité (1), on obtient alors,

$$\|f(c+h) - f(a)\|_F \leq g(c+h) - g(a) + \varepsilon(c+h-a) + \varepsilon,$$

c'est-à-dire $c+h \in E$, ce qui contredit $c = \sup(E)$. D'où $c = b$. En particulier, $b \in E$, c'est-à-dire,

$$\|f(b) - f(a)\|_F \leq g(b) - g(a) + \varepsilon(b-a) + \varepsilon.$$

Mais ceci vaut pour tout $\varepsilon > 0$. Finalement, en faisant tendre ε vers 0, il vient,

$$\|f(b) - f(a)\|_F \leq g(b) - g(a).$$

En particulier, si pour tout $t \in]a,b[$, $\|f'(t)\|_F \leq M$, en posant, pour tout $t \in [a,b]$, $g(t) = Mt$, on obtient $\|f(b) - f(a)\|_F \leq M|b-a|$ □

On note $|||\cdot|||$ la norme subordonnée à $\|\cdot\|_E$ et $\|\cdot\|_F$, c'est-à-dire la norme définie pour toute application linéaire $L : E \to F$ par,

$$|||L||| = \sup\left\{\frac{\|L(x)\|_F}{\|x\|_E},\ x \in E \setminus \{0_E\}\right\}.$$

Théorème 5 (Inégalité des accroissements finis généralisée). Soient U un ouvert convexe de E et $f : U \to F$ une fonction continue et différentiable. On suppose qu'il existe un réel $M > 0$ tel que, pour tout $x \in U$, $|||Df(x)||| \leq M$, où $Df(x)$ désigne la différentielle de f au point x. Alors, pour tous $a, b \in U$,

$$\|f(b) - f(a)\|_F \leq M\|b - a\|_E.$$

Preuve. Soient $a, b \in U$. Considérons la fonction φ définie pour tout $t \in [0, 1]$ par,

$$\varphi(t) = f(a(1 - t) + tb).$$

Tout d'abord, φ est bien définie, car U est convexe. Ensuite, φ est différentiable sur $[0, 1]$, en tant que composée de $t \mapsto a(1 - t) + tb$ qui est de classe C^∞ sur $[0, 1]$ (c'est une fonction polynomiale en t), par la fonction f qui est différentiable sur U. De plus, pour tout $(t_0, h) \in [a, b] \times \mathbb{R}$,

$$D\varphi(t_0) \cdot h = Df(a(1 - t_0) + t_0b) \cdot ((a - b)h) = hDf(a(1 - t_0) + t_0b) \cdot (a - b).$$

En particulier, pour $h = 1$ il vient,

$$\varphi'(t) = Df(a(1 - t_0) + t_0b) \cdot (a - b)$$

puis,

$$\|\varphi'(t)\|_F \leq |||Df(a(1 - t_0) + t_0b)||| \times \|a - b\|_E.$$

Or, d'après le théorème 4,

$$\|\varphi(1) - \varphi(0)\|_F \leq |||Df(a(1 - t_0) + t_0b)||| \times \|a - b\|_E \leq M\|a - b\|_E.$$

Sachant que,

$$\|\varphi(1) - \varphi(0)\|_F = \|f(b) - f(a)\|_F,$$

il s'en suit finalement le résultat recherché. □

Centre d'un groupe

Définition 1. On appelle centre de G, et on note $Z(G)$, l'ensemble des éléments de G qui commutent avec tous les éléments de G. Autrement dit :

$$Z(G) = \{z \in G,\ \forall g \in G,\ gz = zg\}.$$

Proposition 2. Pour tout $h \in G$, on note $\mathrm{Int}_h : G \to G,\ g \mapsto hgh^{-1}$. Alors, l'application,

$$\begin{array}{ccc} G & \longrightarrow & \mathrm{Aut}(G) \\ h & \longmapsto & \mathrm{Int}_h \end{array}$$

est un morphisme de groupes, de noyau $Z(G)$. En particulier, $Z(G)$ est un sous-groupe distingué de G.

Lemme 3. Soit G un groupe. Si $G/Z(G)$ est monogène, alors G est abélien.

Preuve. Notons π la surjection canonique de G vers $G/Z(G)$. Soient $a, b \in G$. Par hypothèse, il existe $g_0 \in G$ tel que $G/Z(G) = \langle \pi(g_0) \rangle$. En particulier, il existe $m, n \in \mathbb{Z}$ tels que $\pi(a) = \pi(g_0)^n$ et $b = \pi(g_0)^m$. Autrement dit, il existe $x, y \in Z(G)$ tels que $a = g_0^n x$ et $b = g_0^m y$. Par suite,

$$ab = g_0^n x g_0^m y = x g_0^n g_0^m y = x y g_0^{n+m} = y x g_0^m g_0^n = y g_0^m x g_0^n = g_0^m y g_0^n x = ba.$$

On en conclut que G est abélien. □

Compacité

Proposition 1. Soient $(X_1, d_1), \dots, (X_p, d_p)$ des espaces métriques. Soient $(x_1^{(n)})_{n\in\mathbb{N}}, \dots, (x_p^{(n)})_{n\in\mathbb{N}}$ des suites telles que,

(i) pour tout $i \in [\![1, p]\!]$, $(x_i^{(n)})_{n\in\mathbb{N}}$ est une suite à termes dans X_i;

(ii) pour tout $i \in [\![1, p]\!]$, l'ensemble $\{x_i^{(n)},\ n \in \mathbb{N}\}$ est inclus dans une partie compacte de X_i.

Alors, il existe $\varphi : \mathbb{N} \to \mathbb{N}$ strictement croissante telle que, pour tout $i \in [\![1, p]\!]$, la suite $(x_i^{(\varphi(n))})_{n\in\mathbb{N}}$ converge dans (X_i, d_i).

Preuve. Pour tout $i \in [\![1, p]\!]$, fixons K_i une partie compacte de X_i telle que,

$$\{x_i^{(n)},\ n \in \mathbb{N}\} \subset K_i.$$

Nous allons construire des applications $\varphi_1, \dots, \varphi_p$ strictement croissantes de $\mathbb{N}$ dans $\mathbb{N}$ telles que, pour tout $i \in [\![1, p]\!]$, la suite $(x_i^{((\varphi_1\circ\cdots\circ\varphi_i)(n)))})_{n\in\mathbb{N}}$ converge dans (X_i, d_i). Pour cela, on procède par récurrence.

Initialisation. Puisque la suite $(x_1^{(n)})_{n\in\mathbb{N}}$ est à termes dans K_1 qui est compact, il existe $\varphi_1 : \mathbb{N} \to \mathbb{N}$ strictement croissante telle que $(x_1^{(\varphi_1(n))})_{n\in\mathbb{N}}$ converge dans (K_1, d_1). A fortiori, $(x_1^{(\varphi_1(n))})_{n\in\mathbb{N}}$ converge aussi dans (X_1, d_1).

Hérédité. Supposons ensuite avoir construit, pour un rang $q \in [\![1, p-1]\!]$ fixé, des applications $\varphi_1, \dots, \varphi_q$ strictement croissantes de $\mathbb{N}$ dans $\mathbb{N}$ telles que, pour tout $i \in [\![1, q]\!]$, la suite $(x_i^{((\varphi_1\circ\cdots\circ\varphi_i)(n)))})_{n\in\mathbb{N}}$ converge dans (X_i, d_i). Notons,

$$(u^{(n)})_{n\in\mathbb{N}} := (x_{q+1}^{((\varphi_1\circ\cdots\circ\varphi_q)(n)))})_{n\in\mathbb{N}}$$

Alors,

$$\{u^{(n)},\ n \in \mathbb{N}\} \subset \{x_{q+1}^{(n)},\ n \in \mathbb{N}\} \subset K_{q+1}.$$

Puisque K_{q+1} est compact, on en déduit qu'il existe $\varphi_{q+1} : \mathbb{N} \to \mathbb{N}$ strictement croissante telle que $(u^{(\varphi_{q+1}(n))})_{n\in\mathbb{N}}$ converge dans (X_{q+1}, d_{q+1}). Or,

$$\forall n \in \mathbb{N},\ u^{(\varphi_{q+1}(n))} = x_{q+1}^{((\varphi_1\circ\cdots\circ\varphi_q)(\varphi_{q+1}(n))))} = x_{q+1}^{((\varphi_1\circ\cdots\circ\varphi_{q+1})(n)))}.$$

Donc, $(x_{q+1}^{((\varphi_1\circ\cdots\circ\varphi_{q+1})(n)))})_{n\in\mathbb{N}}$ converge dans (X_{q+1}, d_{q+1}).

Nous avons donc construit $\varphi_1, \dots, \varphi_p$ des applications strictement croissantes de $\mathbb{N}$ dans $\mathbb{N}$ telles que, pour tout $i \in [\![1, p]\!]$, la suite $(x_i^{((\varphi_1\circ\cdots\circ\varphi_i)(n)))})_{n\in\mathbb{N}}$ converge dans (X_i, d_i). Notons alors,

$$\varphi = \varphi_1 \circ \cdots \circ \varphi_p$$

qui est strictement croissante par composition d'applications strictement croissantes. Alors, pour tout $i \in [\![1, p]\!]$, la suite $(x_i^{((\varphi(n)))})_{n\in\mathbb{N}}$ converge en tant que suite extraite de $(x_i^{((\varphi_1\circ\cdots\circ\varphi_i)(n)))})_{n\in\mathbb{N}}$ qui est convergente (l'extractrice étant $\varphi_{i+1} \circ \cdots \circ \varphi_p$ si $i < p$, l'identité sinon). □

Remarque. Ce résultat reste vrai pour un nombre dénombrable de suites (voir proposition 3).

Proposition 2. Soient (X, d) un espace métrique compact et $(x_p)_{p\in\mathbb{N}}$ une suite à termes dans X possédant une unique valeur d'adhérence. Alors, $(x_p)_{p\in\mathbb{N}}$ converge vers cette valeur d'adhérence.

Preuve. Notons x_∞ l'unique valeur d'adhérence de $(x_p)_{p\in\mathbb{N}}$. Montrons que $(x_p)_{p\in\mathbb{N}}$ converge vers x_∞. On raisonne par l'absurde en supposant qu'il existe $\varepsilon_0 > 0$ tel que, pour tout $p \in \mathbb{N}$, il existe un entier naturel k_p, dépendant de p, tel que $k_p \geq p$ et $d(x_{k_p}, x_\infty) > \varepsilon_0$. Nous allons construire par récurrence une suite $(\varphi(p))_{p\in\mathbb{N}}$ strictement croissante d'entiers naturels telle que,

$$\forall p \in \mathbb{N},\ d(x_{\varphi(p)}, x_\infty) > \varepsilon_0.$$

Pour l'initialisation, on choisit $\varphi(0) = k_0$. Supposons ensuite, pour $p \in \mathbb{N}$ fixé, avoir construit des entiers naturels $\varphi(0) < \cdots < \varphi(p)$ tels que, pour tout $l \in [\![0, p]\!]$, on ait $d(x_{\varphi(l)}, x_\infty) > \varepsilon_0$. On pose alors $N = \varphi(p) + 1$, puis $\varphi(p+1) = k_N$. Alors, $\varphi(p+1) \geq N > \varphi(p)$ et $d(x_{\varphi(p+1)}, x_\infty) > \varepsilon_0$, ce qui termine la récurrence. Ainsi, $(x_{\varphi(p)})_{p\in\mathbb{N}}$ est une suite extraite de $(x_p)_{p\in\mathbb{N}}$ vérifiant, pour tout $p \in \mathbb{N}$,

$$d(x_{\varphi(p)}, x_\infty) > \varepsilon_0. \tag{1}$$

Et puisque $(x_{\varphi(p)})_{p\in\mathbb{N}}$ est une suite à termes dans X qui est compact, cette suite possède au moins une valeur d'adhérence, disons y_∞. A fortiori, y_∞ est une valeur d'adhérence de $(x_p)_{p\in\mathbb{N}}$. Or, la suite $(x_p)_{p\in\mathbb{N}}$ possède une unique valeur d'adhérence. D'où, $x_\infty = y_\infty$, et donc x_∞ est une valeur d'adhérence de $(x_{\varphi(p)})_{p\in\mathbb{N}}$, ce qui contredit l'inégalité (1). Ainsi, $(x_p)_{p\in\mathbb{N}}$ converge vers x_∞. □

Compacité relative

Définition 1. Soient (X,d) un espace métrique et $Y \subset X$. On dit que Y est relativement compact dans X si Y est inclus dans une partie compacte de X.

Remarque. On peut également élargir cette définition au cadre plus général des espaces topologiques.

Proposition 2. Soient (X,d) un espace métrique et $Y \subset X$. Les assertions suivantes sont équivalentes :

(i) Y est relativement compact dans (X,d);

(ii) de toute suite d'éléments de Y, on peut en extraire une sous-suite convergente dans (X,d);

(iii) $\overline{Y}$ est compact dans (X,d).

Preuve. On montre que $(i) \Longrightarrow (ii) \Longrightarrow (iii) \Longrightarrow (i)$.

• $\boxed{(i) \Longrightarrow (ii)}$. Supposons que Y soit relativement compact dans X, c'est-à-dire qu'il existe un compact K de X contenant Y. Soit $(y_n)_{n\in\mathbb{N}}$ une suite à termes dans Y. A fortiori, cette suite est à termes dans K. Puisque K est compact, il existe donc une application $\varphi : \mathbb{N} \to \mathbb{N}$ strictement croissante telle que $(y_{\varphi(n)})_{n\in\mathbb{N}}$ converge dans K muni de la distance induite par d. A fortiori, $(y_{\varphi(n)})_{n\in\mathbb{N}}$ converge dans (X,d).

Remarque. En revanche, cette limite n'appartient pas nécessairement à Y. En effet, prenons par exemple, $X = [0,1]$, $Y = [0,1[$, et $(y_n)_{n\in\mathbb{N}}$ la suite définie pour tout $n \in \mathbb{N}$ par,

$$y_n = 1 - \frac{1}{n+1}.$$

Alors, $(y_n)_{n\in\mathbb{N}}$ converge vers 1. En particulier, elle possède une unique valeur d'adhérence qui est 1, mais celle-ci n'appartient pas à Y.

• $\boxed{(ii) \Longrightarrow (iii)}$. Supposons que toute suite d'éléments de Y possède une valeur d'adhérence. Soit $(x_n)_{n\in\mathbb{N}}$ une suite à termes dans $\overline{Y}$. Pour tout $n \in \mathbb{N}$, il existe donc $y_n \in Y$ tel que,

$$d(x_n, y_n) \leq \frac{1}{2^n}.$$

En particulier, $(y_n)_{n\in\mathbb{N}}$ est une suite à termes dans Y. D'après notre hypothèse, il existe donc une application $\varphi : \mathbb{N} \to \mathbb{N}$ strictement croissante telle que $(y_{\varphi(n)})_{n\in\mathbb{N}}$ converge dans (X,d). Notons l sa limite. Alors, $(x_{\varphi(n)})_{n\in\mathbb{N}}$ converge aussi vers l, puisque, par inégalité triangulaire, on a pour tout $n \in \mathbb{N}$,

$$\begin{aligned} d(x_{\varphi(n)}, l) &\leq d(x_{\varphi(n)}, y_{\varphi(n)}) + d(y_{\varphi(n)}, l) \\ &\leq \frac{1}{2^{\varphi(n)}} + d(y_{\varphi(n)}, l) \\ d(x_{\varphi(n)}, l) &\leq \frac{1}{2^n} + d(y_{\varphi(n)}, l) \underset{n\to+\infty}{\longrightarrow} 0. \end{aligned}$$

Pour la dernière inégalité, on a utilisé l'inégalité $\varphi(n) \geq n$ qui est une propriété des fonctions strictement croissantes de $\mathbb{N}$ dans $\mathbb{N}$ se démontrant par récurrence. Ainsi, $\overline{Y}$ est compact.

• $\boxed{(iii) \Longrightarrow (i)}$. Si $\overline{Y}$ est compact, alors il s'agit d'une partie compacte de X contenant Y, et donc Y est bien relativement compact dans (X,d). □

Proposition 3. Soient $(X_p, d_p)_{p\in\mathbb{N}}$ une suite d'espaces métriques et pour tout $p \in \mathbb{N}$, $(x_{n,p})_{n\in\mathbb{N}}$ une suite à termes dans X_p. On suppose que pour tout $p \in \mathbb{N}$, l'ensemble $\{x_{n,p},\ n \in \mathbb{N}\}$ est relativement compact dans (X_p, d_p). Sous ces hypothèses, il existe une application $\varphi : \mathbb{N} \to \mathbb{N}$ strictement croissante telle que, pour tout $p \in \mathbb{N}$, la suite $(x_{\varphi(n),p})_{n\in\mathbb{N}}$ converge dans (X_p, d_p).

Preuve. Montrons, par récurrence sur $p \in \mathbb{N}$, qu'il existe $\varphi_0, \ldots, \varphi_p$ des applications strictement croissantes telles que, pour tout $k \in [\![0,p]\!]$, la suite $(x_{\varphi_0\circ\cdots\circ\varphi_p(n),k})_{n\in\mathbb{N}}$ converge dans (X_k, d_k).

Initialisation. Puisque $\{x_{n,0},\ n \in \mathbb{N}\}$ est une partie relativement compacte de (X_0, d_0), d'après la proposition 2, il existe une application φ_0 strictement croissante de $\mathbb{N}$ dans $\mathbb{N}$ telle que la suite $(x_{\varphi_0(n),0})_{n\in\mathbb{N}}$ converge dans (X_0, d_0).

Hérédité. Supposons le résultat établi pour un rang $p \in \mathbb{N}$ fixé. Pour tout $k \in [\![0,p+1]\!]$, on note,

$$(y_{n,k})_{n\in\mathbb{N}} = (x_{\varphi_0\circ\cdots\circ\varphi_p(n),k})_{n\in\mathbb{N}}.$$

Alors, $\{y_{n,p+1},\ n \in \mathbb{N}\}$ est relativement compact dans (X_{p+1}, d_{p+1}) car,

$$\{y_{n,p+1},\ n \in \mathbb{N}\} \subset \{x_{n,p+1},\ n \in \mathbb{N}\}$$

et $\{x_{n,p+1},\ n \in \mathbb{N}\}$ est relativement compact dans (X_{p+1}, d_{p+1}). D'après la proposition 2, il existe donc une application φ_{p+1} strictement croissante de $\mathbb{N}$ dans $\mathbb{N}$ telle que la suite $(y_{\varphi_{p+1}(n),p+1})_{n\in\mathbb{N}}$ converge dans (X_{p+1}, d_{p+1}). Or, d'après notre hypothèse de récurrence, pour tout $k \in [\![0,p]\!]$, la suite $(y_{n,k})_{n\in\mathbb{N}}$ converge dans (X_k, d_k). A fortiori, pour tout $k \in [\![0,p]\!]$, la suite $(y_{\varphi_{p+1}(n),k})_{n\in\mathbb{N}}$

converge dans (X_k, d_k) en tant que suite extraite de $(y_{n,k})_{n\in\mathbb{N}}$, ce qui termine la récurrence car,

$$(y_{\varphi_{p+1}(n),k})_{n\in\mathbb{N}} = (x_{\varphi_0\circ\cdots\circ\varphi_p(\varphi_{p+1}(n)),k})_{n\in\mathbb{N}} = (x_{\varphi_0\circ\cdots\circ\varphi_p\circ\varphi_{p+1}(n),k})_{n\in\mathbb{N}}.$$

Posons finalement, pour tout $n \in \mathbb{N}$,

$$\varphi(n) = (\varphi_0 \circ \cdots \circ \varphi_n)(n).$$

Montrons que φ convient à notre problème. On rappelle ensuite que toute fonction f strictement croissante de $\mathbb{N}$ dans $\mathbb{N}$ vérifie, pour tout $n \in \mathbb{N}$, $f(n) \geq n$. Cette propriété se démontre facilement par récurrence. Soit $n \in \mathbb{N}$. Montrons que $\varphi(n+1) > \varphi(n)$. D'après la propriété rappelée ci-dessus, $\varphi_{n+1}(n+1) \geq n+1 > n$. D'où,

$$\varphi(n+1) = (\varphi_0 \circ \cdots \circ \varphi_n)(\varphi_{n+1}(n+1)) > (\varphi_0 \circ \cdots \circ \varphi_n)(n) = \varphi(n),$$

car $\varphi_0 \circ \cdots \circ \varphi_n$ est strictement croissante. Ainsi, φ est strictement croissante. D'autre part, pour tout $p \in \mathbb{N}$, $(x_{\varphi(n),p})_{n\in\mathbb{N}}$ converge dans (X_p, d_p) en tant que suite extraite de $(x_{(\varphi_0\circ\cdots\circ\varphi_p)(n),p})_{n\in\mathbb{N}}$ à partir du rang p, qui est convergente dans (X_p, d_p) par construction. D'où la proposition. □

Complétude et espaces de fonctions

Théorème 1. Soient (X, d_X) un espace métrique compact et (E, d_E) un espace métrique complet. Notons $C(X, E)$ l'ensemble des fonctions continues de X dans E. On munit cet ensemble d'une structure d'espace métrique par la distance d définie pour tous $f, g \in C(X, E)$ par,

$$d(f, g) = \sup_{x \in X} d_E(f(x), g(x)).$$

Alors, $(C(X, E), d)$ est un espace métrique complet.

Autrement dit, si $(f_n)_{n\in\mathbb{N}}$ est une suite de fonctions continues de X dans E telle que,

$$\forall \varepsilon > 0,\ \exists N \in \mathbb{N},\ \forall p, q \geq N,\ \forall x \in X,\ d_E(f_p(x), f_q(x)) \leq \varepsilon,$$

alors $(f_n)_{n\in\mathbb{N}}$ converge uniformément sur X.

Preuve. Soit $(f_n)_{n\in\mathbb{N}}$ une suite de Cauchy dans $(C(X, E), d)$, c'est-à-dire que,

$$\forall \varepsilon > 0,\ \exists N \in \mathbb{N},\ \forall p, q \geq N,\ \forall x \in X,\ d_E(f_p(x), f_q(x)) \leq \varepsilon.$$

Pour prouver ce théorème, on procède en trois étapes. On construit tout d'abord la limite simple f de la suite $(f_n)_{n\in\mathbb{N}}$. On prouve ensuite que f est continue sur X. Enfin, on démontre que $(f_n)_{n\in\mathbb{N}}$ converge uniformément vers f sur X.

Étape n°1. Montrons que $(f_n)_{n\in\mathbb{N}}$ converge simplement sur X. Soit $x \in X$. D'après l'assertion ci-dessus, la suite $(f_n(x))_{n\in\mathbb{N}}$ est de Cauchy dans (E, d_E). Ce dernier étant complet, cette suite converge donc dans (E, d_E) vers un élément y_x (dépendant de x). Considérons alors la fonction,

$$\begin{array}{rccc} f & : & X & \longrightarrow & E \\ & & x & \longmapsto & y_x \end{array}$$

Étape n°2. Montrons ensuite que f est continue sur X. Soient $x_0 \in X$ et $\varepsilon > 0$. Puisque la suite $(f_n)_{n\in\mathbb{N}}$ est de Cauchy dans $(C(X, E), d)$, il existe $N \in \mathbb{N}$ tel que,

$$\forall p, q \geq N,\ \forall x \in X,\ d_E(f_p(x), f_q(x)) \leq \frac{\varepsilon}{3}.$$

Soit $x \in X$. D'après l'assertion ci-dessus, pour tout $p \geq N$, $d_E(f_p(x), f_N(x)) \leq \frac{\varepsilon}{3}$. D'où, en faisant tendre p vers l'infini, $d_E(f(x), f_N(x)) \leq \frac{\varepsilon}{3}$ car $(f_p(x))_{p\in\mathbb{N}}$ converge vers $f(x)$ et $y \mapsto d_E(y, f_N(x))$ est continue sur X (d'après la seconde inégalité triangulaire). On a donc,

$$\forall x \in X,\ d_E(f_N(x), f(x)) \leq \frac{\varepsilon}{3}.$$

D'autre part, la fonction f_N étant continue sur X, il existe donc $\alpha > 0$ tel que,

$$\forall x \in X,\ \left(d_X(x, x_0) \leq \alpha \Longrightarrow d_E(f_N(x), f_N(x_0)) \leq \frac{\varepsilon}{3}\right).$$

Soit $x \in X$ tel que $d_X(x, x_0) \leq \alpha$. Alors, par inégalité triangulaire,

$$d_E(f(x), f(x_0)) \leq d_E(f(x), f_N(x)) + d_E(f_N(x), f_N(x_0)) + d_E(f_N(x_0), f(x_0))$$

et donc,

$$d_E(f(x), f(x_0)) \leq \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon.$$

Ainsi, f est continue sur X.

Étape n°3. Montrons finalement que $(f_n)_{n\in\mathbb{N}}$ converge uniformément vers f sur X. Soit $\varepsilon > 0$. Puisque la suite $(f_n)_{n\in\mathbb{N}}$ est de Cauchy dans $(C(X, E), d)$, il existe $N \in \mathbb{N}$ tel que,

$$\forall n, p \geq N,\ \forall x \in X,\ d_E(f_n(x), f_p(x)) \leq \varepsilon.$$

Soient $n \geq N$ et $x \in X$. Alors, pour tout $p \geq N$, $d_E(f_n(x), f_p(x)) \leq \varepsilon$. En faisant tendre p vers l'infini, la continuité de $y \mapsto d_E(f_n(x), y)$ assure que $d_E(f_n(x), f(x)) \leq \varepsilon$, car $(f_p(x))_{p\in\mathbb{N}}$ converge vers $f(x)$. Ainsi,

$$\forall \varepsilon > 0,\ \exists N \in \mathbb{N},\ \forall n \geq N,\ \forall x \in X,\ d_E(f_n(x), f(x)) \leq \varepsilon,$$

c'est-à-dire que $(f_n)_{n\in\mathbb{N}}$ converge uniformément vers f sur X. □

Convergence en loi

On se place sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$. Soient $(X_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires réelles et X une variable aléatoire réelle (sur $(\Omega, \mathcal{A})$). On note $C_b(\mathbb{R}, \mathbb{R})$ l'ensemble des fonctions de $\mathbb{R}$ dans $\mathbb{R}$ continues et bornées.

Définition 1. On dit que la suite $(X_n)_{n\in\mathbb{N}^*}$ converge en loi vers X, et on note $X_n \xrightarrow[n\to+\infty]{\mathcal{L}} X$ si, pour tout $\varphi \in C_b(\mathbb{R}, \mathbb{R})$,

$$\int_\Omega \varphi(X_n(\omega))d\mathbb{P}(\omega) \underset{n\to+\infty}{\longrightarrow} \int_\Omega \varphi(X(\omega))d\mathbb{P}(\omega).$$

Si tel est le cas, $(X_n)_{n\in\mathbb{N}^*}$ converge en loi vers toute variable aléatoire qui suit la même loi que X. C'est pourquoi on écrit également,

$$X_n \xrightarrow[n\to+\infty]{\mathcal{L}} \mathbb{P}_X.$$

[GH] **Théorème 2 (Théorème de Lévy).** On a $X_n \xrightarrow[n\to+\infty]{\mathcal{L}} X$ si et seulement si la suite de fonctions φ_{X_n} converge simplement vers φ_X sur $\mathbb{R}$.

[GK] **Théorème 3.** Soit $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé. Soient $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires réelles et X une variable aléatoire réelle, toutes définies sur $(\Omega, \mathcal{A}, \mathbb{P})$. Les assertions suivantes sont équivalentes :

(i) $(X_n)_{n\in\mathbb{N}}$ converge en loi vers X;

(ii) Pour tout point $x \in \mathbb{R}$, si F_X est continue au point x alors $(F_{X_n}(x))_{n\in\mathbb{N}}$ converge vers $F_X(x)$.

En particulier, si F_X est continue, alors $(X_n)_{n\in\mathbb{N}}$ converge en loi vers X si et seulement si $(F_{X_n})_{n\in\mathbb{N}}$ converge simplement vers F_X sur $\mathbb{R}$.

Remarque. La fonction de répartition d'une variable aléatoire réelle à densité est continue (résultat dont on se servira pour prouver la formule de Stirling).

Décomposition polaire sur $M_n(\mathbb{R})$

Soit $n \in \mathbb{N}^*$.

Théorème 1. Soit $M \in M_n(\mathbb{R})$. Alors, il existe $O \in O_n(\mathbb{R})$ et $S \in S_n^+(\mathbb{R})$ tel que $M = OS$.

Preuve. Soit $M \in M_n(\mathbb{R})$. D'après le théorème 1, il existe une suite $(M_k)_{k\in\mathbb{N}}$ de matrices inversibles qui converge vers M. Pour tout $k \in \mathbb{N}$, le théorème 2 assure l'existence de $(O_k, S_k) \in O_n(\mathbb{R}) \times S_n^{++}(\mathbb{R})$ tel que $M_k = O_k S_k$. Sachant que $O_n(\mathbb{R})$ est compact (d'après le théorème 3), il existe donc $O \in O_n(\mathbb{R})$ et $\varphi : \mathbb{N} \to \mathbb{N}$ une extractrice telle que $(O_{\varphi(k)})_{n\in\mathbb{N}}$ converge vers O. Or,

$$\forall k \in \mathbb{N}, \quad S_{\varphi(k)} = O_{\varphi(k)}^{-1} M_{\varphi(k)} = {}^tO_{\varphi(k)} M_{\varphi(k)}$$

Par continuité de la transposée et du produit matriciel, on en déduit que :

$$\lim_{k\to+\infty} S_{\varphi(k)} = \lim_{k\to+\infty} {}^tO_{\varphi(k)} M_{\varphi(k)} = {}^tOM$$

Notons $S = {}^tOM$. Puisque $\overline{S_n^{++}(\mathbb{R})} = S_n^+(\mathbb{R})$ (d'après le théorème 2), il vient $S \in S_n^+(\mathbb{R})$. D'autre part $M = I_n M = O\,{}^tOM = OS$. Ainsi, $M = OS$ avec $O \in O_n(\mathbb{R})$ et $S \in S_n^+(\mathbb{R})$. □

Remarque. Si $M \notin GL_n(\mathbb{R})$, cette décomposition n'est pas unique.

Densité de $GL_n(\mathbb{R})$ dans $M_n(\mathbb{R})$

Soit $n \in \mathbb{N}^*$.

Théorème 1. Notons $\overline{GL_n(\mathbb{R})}$ l'adhérence de $GL_n(\mathbb{R})$ dans $M_n(\mathbb{R})$. Alors, $\overline{GL_n(\mathbb{R})} = M_n(\mathbb{R})$.

Preuve. Soit $M \in M_n(\mathbb{R})$. Pour tout $k \in \mathbb{N}^*$, notons,

$$M_k = M - \frac{1}{k} I_n.$$

Avec ces notations, pour tout $k \in \mathbb{N}^*$,

$$\begin{aligned} M_k \in GL_n(\mathbb{R}) &\iff \det\left(M - \frac{1}{k} I_n\right) \neq 0 \\ &\iff \mathrm{Ker}\left(M - \frac{1}{k} I_n\right) = \{0\} \\ M_k \in GL_n(\mathbb{R}) &\iff \frac{1}{k} \notin \mathrm{Sp}(M). \end{aligned}$$

D'où,

$$\{k \in \mathbb{N} \mid M_k \notin GL_n(\mathbb{R})\} \subset \left\{ k \in \mathbb{N} \;\middle|\; \frac{1}{k} \in Sp(M) \right\}.$$

Or, $\mathrm{Sp}(M)$ est un ensemble fini (et même de cardinal inférieur ou égal à n, en tant qu'ensemble des racines du polynôme à coefficients réels χ_M qui est de degré n). Par conséquent, l'ensemble,

$$E := \{k \in \mathbb{N} \mid M_k \notin GL_n(\mathbb{R})\}$$

est fini. Notons φ la bijection strictement croissante qui énumère les éléments de E, c'est-à-dire l'application définie par :

$$\varphi(0) = \min E, \quad \forall k \in \mathbb{N},\ \varphi(k+1) = \min(E \setminus \{\varphi(0), \ldots, \varphi(k)\}).$$

Avec ces notations, $(M_{\varphi(k)})_{k \in \mathbb{N}}$ est une suite de matrices inversibles. D'autre part, pour tout $k \in \mathbb{N}$, $\varphi(k) \geq k$ (propriété classique des fonctions strictement croissante de $\mathbb{N}$ dans $\mathbb{N}$), donc,

$$\lim_{k \to +\infty} \frac{1}{\varphi(k)} = 0$$

ce qui implique,

$$\lim_{k \to +\infty} M_{\varphi(k)} = M$$

Ainsi, M est limite d'une suite de matrices inversibles. □

Dérivation

Théorème 1. Soient I un intervalle de $\mathbb{R}$ et $f : I \to \mathbb{R}$ une fonction de classe C^1 telle que $f' > 0$ sur I. Alors, l'application,

$$\begin{array}{rccc} g & : & I & \longrightarrow & f(I) \\ & & x & \longmapsto & f(x) \end{array}$$

est bijective et sa bijection réciproque, notée g^{-1}, est de classe C^1 sur J.

Preuve. Tout d'abord, g est surjective par construction, et g est injective, car elle est strictement croissante. Donc g est bijective. De plus, puisque g est dérivable sur I avec $g' = f' > 0$, g^{-1} est dérivable sur J avec,

$$(g^{-1})' = \frac{1}{g' \circ g^{-1}}$$

Or, g' est continue sur I, car f est de classe C^1 sur I, et g^{-1} est continue sur J car dérivable sur J. Finalement, $(g^{-1})'$ est continue en tant que composée d'applications continues, et donc g^{-1} est de classe C^1 sur J. □

Théorème 2 (Théorème de la limite de la dérivée). Soient I un intervalle de $\mathbb{R}$, $a \in I$, $f : I \to \mathbb{R}$ une fonction continue sur I et dérivable sur $I \setminus \{a\}$. On suppose que f' admet une limite $l \in \overline{\mathbb{R}}$. Alors,

$$\lim_{x \to a} \frac{f(x) - f(a)}{x - a} = l.$$

En particulier, si $l \in \mathbb{R}$, f est dérivable au point a et f' est continue en a avec $f'(a) = l$.

Preuve. Soit $x \in I \setminus \{a\}$. Notons J_x l'intervalle d'extrémités a et x. Puisque f est continue sur $\overline{J_x}$ et dérivable sur J_x, le théorème des accroissements finis assure l'existence de $c_x \in J_x$ tel que,

$$f(x) - f(a) = f'(c_x)(x - a).$$

Ainsi, pour tout $x \in I \setminus \{a\}$,

$$\frac{f(x) - f(a)}{x - a} = f'(c_x)$$

avec $c_x \underset{x \to a}{\longrightarrow} a$ (car $|c_x - a| \leq |x - a|$). D'où, en faisant tendre x vers a,

$$\lim_{x \to a} \frac{f(x) - f(a)}{x - a} = l$$

(par composition de limites). Supposons ensuite $l \in \mathbb{R}$. Alors, l'assertion ci-dessus prouve que f est dérivable au point a avec $f'(a) = l$. □

Diagonalisation simultanée

Théorème 1. Soient $\mathbb{K}$ un corps commutatif, E un espace vectoriel sur $\mathbb{K}$, puis f et g deux endomorphismes diagonalisables qui commutent. Alors, f et g sont simultanément diagonalisables, c'est-à-dire qu'il existe une base de E telle que les matrices de f et g dans cette base soient toutes les deux diagonales. En particulier, si $A, B \in M_n(\mathbb{K})$ sont deux matrices diagonalisables qui commutent, alors il existe $P \in GL_n(\mathbb{K})$ et $D_1, D_2 \in D_n(\mathbb{K})$ tels que $A = PD_1P^{-1}$ et $B = PD_2P^{-1}$.

Preuve. Montrons tout d'abord que les sous-espaces propres de f sont stables par g. Soit $\lambda \in f$. Soit $x \in E_\lambda(f)$. Alors,

$$f(g(x)) = g(f(x)) = g(\lambda x) = \lambda g(x)$$

c'est-à-dire $g(x) \in E_\lambda(f)$. Donc, $E_\lambda(f)$ est bien stable par g. Notons ensuite $\lambda_1, \dots, \lambda_r$ les valeurs propres de f, puis $n_1, \dots, n_r$ les dimensions respectives des sous-espaces propres $E_{\lambda_1}(f), \dots, E_{\lambda_r}(f)$. Pour tout $k \in \{1, \dots, r\}$, notons f_k et g_k les endomorphismes respectivement induits par f et g sur $E_{\lambda_k}(f)$. Alors, pour tout $k \in \{1, \dots, r\}$,

(i) $f_k = \lambda_k Id_{E_k}$ où l'on a noté $E_k = E_{\lambda_k}(f)$;

(ii) g_k est diagonalisable en tant qu'endomorphisme induit par un endomorphisme diagonalisable.

Le point (ii) assure l'existence de $\mathcal{B}_1, \dots, \mathcal{B}_r$ des bases respectives de $E_{\lambda_1}(f), \dots, E_{\lambda_r}(f)$ telles que les matrices respectives de $g_1, \dots, g_r$ dans ces bases soient diagonales. Notons $\mathcal{B}$ la concaténation des bases $\mathcal{B}_1, \dots, \mathcal{B}_r$. Puisque f est diagonalisable, il s'agit d'une base de E. Nous allons montrer que $\mathcal{B}$ est une base qui diagonalise simultanément f et g.

Pour tout $(M_1, \dots, M_r) \in M_{n_1}(\mathbb{K}) \times \cdots \times M_{n_r}(\mathbb{K})$, on note $\mathrm{Diag}(M_1, \dots, M_r)$ la matrice de $M_n(\mathbb{K})$ diagonale par blocs de coefficients diagonaux $M_1, \dots, M_r$. Alors,

$$\begin{cases} \mathrm{Mat}_{\mathcal{B}}(f) = \mathrm{Diag}(\mathrm{Mat}_{\mathcal{B}_1}(f_1), \dots, \mathrm{Mat}_{\mathcal{B}_r}(f_r)), \\ \mathrm{Mat}_{\mathcal{B}}(g) = \mathrm{Diag}(\mathrm{Mat}_{\mathcal{B}_1}(g_1), \dots, \mathrm{Mat}_{\mathcal{B}_r}(g_r)). \end{cases}$$

Or, d'après le point (i), pour tout $k \in \{1, \dots, r\}$, $\mathrm{Mat}_{\mathcal{B}_k}(f_k) = \lambda_k I_{n_k}$. D'où,

$$\mathrm{Mat}_{\mathcal{B}}(f) = \mathrm{Diag}(\mathrm{Mat}_{\mathcal{B}_1}(f_1), \dots, \mathrm{Mat}_{\mathcal{B}_r}(f_r)) = \mathrm{Diag}(\lambda_1 I_{n_1}, \dots, \lambda_r I_{n_r}).$$

En particulier, $\mathrm{Mat}_{\mathcal{B}}(f)$ est diagonale. De plus, par le point (ii), $\mathrm{Mat}_{\mathcal{B}}(g)$ est une matrice diagonale en tant que matrice diagonale par blocs, dont les blocs diagonaux sont des matrices diagonales. Ainsi, $\mathcal{B}$ est une base qui diagonalise simultanément f et g. Ce résultat s'étend aux matrices A et B en considérant leurs endomorphismes de $\mathbb{K}^n$ canoniquement associés. □

Distance entre un point et une partie fermée d'un e.v.n de dimension finie

Théorème 1. Soient E un espace vectoriel normé de dimension finie et F une partie non vide et fermée de E. Soit $x \in E$. Alors, il existe $y \in F$ tel que,

$$\|x - y\| = d(x, F).$$

Preuve. Notons $d = d(x, F)$. Par définition de d, pour tout $n \in \mathbb{N}$, il existe $y_n \in F$ tel que,

$$\|x - y_n\| \leq d + \frac{1}{n+1}$$

Ensuite, pour tout $n \in \mathbb{N}$,

$$\|x - y_n\| \leq d + 1$$

c'est-à-dire $y_n \in \overline{B_E(x, d+1)}$. Or, $\overline{B_E(x, d+1)}$ est une partie compacte de E en tant que partie fermée bornée de E qui est de dimension finie. Par suite, il existe une suite $(y_{\varphi(n)})_{n\in\mathbb{N}}$ extraite de $(y_n)_{n\in\mathbb{N}}$ qui converge vers $y_\infty \in E$. Et puisque F est fermé, on a même $y_\infty \in F$. D'autre part, pour tout $n \in \mathbb{N}$,

$$d \leq \|x - y_{\varphi(n)}\| \leq d + \frac{1}{\varphi(n)+1} \leq d + \frac{1}{n+1}$$

car $\varphi(n) \geq n$ en tant qu'application strictement croissante de $\mathbb{N}$ dans $\mathbb{N}$. D'où, par encadrement de limites,

$$\lim_{n\to+\infty} \|x - y_{\varphi(n)}\| = d.$$

Or, par continuité de la norme,

$$\lim_{n\to+\infty} \|x - y_{\varphi(n)}\| = \|x - y_\infty\|.$$

Ainsi, y_∞ est un élément de F tel que $\|x - y_\infty\| = d$. □

Remarque. En revanche, un tel y n'est pas unique (voir la figure 11).

Équicontinuité

Définition 1. Soient (X, d_X) et (Y, d_Y) deux espaces métriques. On note $C(X, Y)$ l'ensemble des fonctions continues de X dans Y. Soient $\mathcal{A}$ une partie de $C(X, Y)$. On dit que $\mathcal{A}$ est équicontinue si,

$$\begin{cases} \forall x_0 \in X,\ \forall \varepsilon > 0,\ \exists \alpha > 0,\ \forall f \in \mathcal{A},\ \forall x \in X, \\ (d_X(x, x_0) \leq \alpha \Longrightarrow d_Y(f(x), f(x_0)) \leq \varepsilon). \end{cases}$$

On dit que $\mathcal{A}$ est uniformément équicontinue si,

$$\begin{cases} \forall \varepsilon > 0,\ \exists \alpha > 0,\ \forall f \in \mathcal{A}, \forall (x_1, x_2) \in X^2, \\ (d_X(x_1, x_2) \leq \alpha \Longrightarrow d_Y(f(x_1), f(x_2)) \leq \varepsilon). \end{cases}$$

Théorème 2 (Théorème de Heine). Soient (X, d_X) un espace métrique compact et (Y, d_Y) un espace métrique. Soit $\mathcal{A}$ une partie de $C(X, Y)$. Si $\mathcal{A}$ est équicontinue, alors $\mathcal{A}$ est uniformément équicontinue.

Preuve. Soit $\varepsilon > 0$. Pour tout $x \in X$, il existe $\alpha_x > 0$ tel que,

$$\forall f \in \mathcal{A},\ \forall x' \in X,\ \left(d_X(x, x') \leq \alpha \Longrightarrow d_Y(f(x), f(x')) \leq \frac{\varepsilon}{2}\right) \tag{1}$$

car $\mathcal{A}$ est équicontinue. Mais alors,

$$X \subset \bigcup_{x \in X} B\left(x, \frac{\alpha_x}{2}\right).$$

Par compacité de X, il existe donc $x_1, \ldots, x_n \in X$ tels que,

$$X \subset \bigcup_{k=1}^{n} B\left(x_k, \frac{\alpha_k}{2}\right). \tag{2}$$

où l'on a noté, pour tout $k \in \{1, \ldots, n\}$, $\alpha_k = \alpha_{x_k}$. Posons $\alpha = \min\limits_{1 \leq k \leq n}\left(\frac{\alpha_k}{2}\right)$ (qui est un réel strictement positif). Montrons que,

$$\forall f \in \mathcal{A},\ \forall (x, x') \in X^2,\ (d_X(x, x') \leq \alpha \Longrightarrow d_Y(f(x), f(x')) \leq \varepsilon).$$

Soient $f \in \mathcal{A}$ et $x, x' \in X$ tels que $d_X(x, x') \leq \alpha$. D'après l'inclusion (2), il existe $k \in \{1, \ldots, n\}$ tel que $x \in B\left(x_k, \frac{\alpha_k}{2}\right)$. Par inégalité triangulaire, il vient alors,

$$d_X(x', x_k) \leq d_X(x', x) + d_X(x, x_k) < \alpha + \frac{\alpha_k}{2} \leq \frac{\alpha_k}{2} + \frac{\alpha_k}{2} = \alpha_k,$$

et donc $x' \in B(x_k, \alpha_k)$. D'où, d'après l'assertion (1),

$$d_Y(f(x), f(x_k)) \leq \frac{\varepsilon}{2}, \quad d_Y(f(x_k), f(x')) \leq \frac{\varepsilon}{2},$$

et donc, par inégalité triangulaire,

$$d_Y(f(x), f(x')) \leq d_Y(f(x), f(x_k)) + d_Y(f(x_k), f(x')) \leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

ce qui prouve finalement que $\mathcal{A}$ est uniformément équicontinue. □

Fonction caractéristique (probabilités)

Soit X une variable aléatoire sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$.

Définition 1. Soit X une variable aléatoire réelle sur Ω. On appelle fonction caractéristique de X, et on note φ_X, la fonction définie pour tout $t \in \mathbb{R}$ par,

$$\varphi_X(t) = \mathbb{E}\left[e^{itX}\right] = \int_\Omega e^{itX(\omega)} d\mathbb{P}(\omega).$$

[GH] **Proposition 2.** Soit X une variable aléatoire réelle sur Ω. Alors,

(i) φ_X est continue et bornée.

(ii) La loi de X est entièrement déterminée par φ_X, c'est-à-dire que si Y est une variable aléatoire réelle sur Ω telle que $\varphi_X = \varphi_Y$ alors $\mathbb{P}_X = \mathbb{P}_Y$.

(iii) Si $\varphi_X \in L^1_{\mathbb{C}}(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda)$ alors X est à densité, et sa densité f vérifie, pour λ-presque tout x, $f(x) = \frac{1}{2\pi} \int_{\mathbb{R}} e^{-ixt} \varphi_X(t) dt$.

Proposition 3. Supposons que $X \in L^p_{\mathbb{R}}(\Omega, \mathcal{A}, \mathbb{P})$. Alors, φ_X est de classe C^p sur $\mathbb{R}$. De plus, pour tout $k \in [\![0, p]\!]$, $\varphi_X^{(k)}(0) = i^k \mathbb{E}[X^k]$.

Preuve. Ceci résulte du théorème de dérivation sous le signe intégral. □

Fonction de Möbius

Définition 1. On appelle fonction de Möbius, et on note μ, la fonction définie pour tout $n \in \mathbb{N}^*$ par,

$$\mu(n) = \begin{cases} 1 & \text{si } n = 1, \\ (-1)^r & \text{si } n \text{ est sans facteur carré}, \\ 0 & \text{sinon.} \end{cases}$$

On admet le lemme suivant durant l'oral.

[JR] **Théorème 2 (Formule d'inversion de Möbius).** Si $(a(n))_{n\in\mathbb{N}^*}$ et $(b(n))_{n\in\mathbb{N}^*}$ sont deux suites réelles telles que, pour tout $n \in \mathbb{N}^*$, on ait $a(n) = \sum_{d|n} b(d)$, alors, pour tout $n \in \mathbb{N}^*$, $b(n) = \sum_{d|n} \mu\left(\frac{n}{d}\right) a(d)$.

Preuve. Soit $n \in \mathbb{N}^*$. Montrons que,

$$\sum_{d|n} \mu(d) = \begin{cases} 1 & \text{si } n = 1, \\ 0 & \text{sinon.} \end{cases}$$

Si $n = 1$, le résultat est clair. Supposons $n \geq 2$. On écrit,

$$n = \prod_{k=1}^{r} p_k^{\alpha_k}$$

la décomposition de n en facteurs premiers, les p_k étant des nombres premiers deux à deux distincts, et les α_k étant des entiers strictement positifs. Dès lors, les diviseurs strictement positifs de n sont de la forme,

$$d = \prod_{k=1}^{r} p_k^{\beta_k}.$$

avec, pour tout $k \in [\![1, r]\!]$, $0 \leq \beta_k \leq \alpha_k$. De plus, si l'un des β_k est supérieur ou égal à 2 alors $\mu(d) = 0$. Donc,

$$\sum_{d|n} \mu(d) = \sum_{(\beta_1,\dots,\beta_r)\in\{0,1\}^r} \mu\left(\prod_{k=1}^{r} p_k^{\beta_k}\right) = \sum_{(\beta_1,\dots,\beta_r)\in\{0,1\}^r} (-1)^{\beta_1+\cdots+\beta_r}.$$

Pour $l \in [\![0, r]\!]$, il y a $\binom{r}{l}$ façons de choisir un r-uplet $(\beta_1, \dots, \beta_r)$ de $\{0, 1\}$ formé de l termes égaux à 1 et $r-l$ termes égaux à 0, et dans ce cas $\beta_1 + \cdots + \beta_r = l$. En effet, pour tout $l \in [\![0, r]\!]$, l'application,

$$\begin{array}{ccc} \mathcal{P}_l([\![1, r]\!]) & \longrightarrow & E_l \\ X & \longmapsto & (1_X(1), \dots, 1_X(r)) \end{array}$$

est une bijection de l'ensemble $\mathcal{P}_l([\![1, r]\!])$ des parties à l éléments de $[\![1, r]\!]$, qui est de cardinal $\binom{r}{l}$, vers l'ensemble E_l des r-uplets $(\beta_1, \dots, \beta_r)$ de $\{0, 1\}$ formés de l termes égaux à 1 et $r - l$ termes égaux à 0. Par suite,

$$\sum_{d|n} \mu(d) = \sum_{l=1}^{r} \sum_{(\beta_1,\dots,\beta_r)\in E_l} (-1)^{\beta_1+\cdots+\beta_r} = \sum_{l=1}^{r} \sum_{(\beta_1,\dots,\beta_r)\in E_l} (-1)^l,$$

car $(E_l)_{0\leq l\leq r}$ est une partition de $\{0, 1\}^r$. Et donc,

$$\sum_{d|n} \mu(d) = \sum_{l=1}^{r} (-1)^l \operatorname{card}(E_l) = \sum_{l=1}^{r} \binom{r}{l} (-1)^l = (1-1)^r = 0,$$

l'avant-dernière égalité étant due à la formule du binôme de Newton. D'où,

$$\forall n \in \mathbb{N}^*, \ \sum_{d|n} \mu(d) = \begin{cases} 1 & \text{si } n = 1, \\ 0 & \text{sinon.} \end{cases} \tag{1}$$

On prouve désormais le lemme. Soit $n \in \mathbb{N}^*$. D'après notre hypothèse,

$$\sum_{d|n} \mu\left(\frac{n}{d}\right) a(d) = \sum_{d|n} \mu(d) a\left(\frac{n}{d}\right) = \sum_{d|n} \mu(d) \left(\sum_{k|\frac{n}{d}} b(k)\right).$$

Or, pour tous $k, d \in \mathbb{N}^*$, d divise n et k divise $\frac{n}{d}$ si et seulement si k divise n et d divise $\frac{n}{k}$. Par conséquent,

$$\sum_{d|n} \mu\left(\frac{n}{d}\right) a(d) = \sum_{k|n} \sum_{d|\frac{n}{k}} \mu(d) b(k) = \sum_{k|n} b(k) \sum_{d|\frac{n}{k}} \mu(d).$$

De plus, d'après l'égalité (1), tous les termes de cette somme sont nuls hormis celui pour $n = k$ qui vaut $b(n)$. D'où,

$$\sum_{d|n} \mu\left(\frac{n}{d}\right) a(d) = b(n).$$

□

Remarque. La réciproque est vraie, mais n'est pas nécessaire pour prouver le théorème 1.

Fonctions convexes et fonctions α-convexes

Proposition 1. Soient X un convexe de $\mathbb{R}^n$ et $f : X \to \mathbb{R}$ une fonction. Soit $x^\star$ un point de minimum local de f sur X.

(i) Si f est convexe, alors $x^\star$ est un point de minimum global de f sur X.

(ii) Si f est strictement convexe, alors $x^\star$ est l'unique point de minimum global de f sur X.

Définition 2. Soit $f : \mathbb{R}^n \to \mathbb{R}$. On dit que f est coercive si,

$$\lim_{\|x\| \to +\infty} f(x) = +\infty$$

c'est-à-dire,

$$\forall A > 0,\ \exists R > 0,\ \forall x \in \mathbb{R}^n,\ (\|x\| \geq R \Longrightarrow f(x) \geq A).$$

Proposition 3. Soit $f : \mathbb{R}^n \to \mathbb{R}$. Si f est coercive, alors elle passe par un minimum global.

Corollaire 4. Soit $f : \mathbb{R}^n \to \mathbb{R}$. Si f est strictement convexe et coercive, alors f passe par un minimum local en unique point, et ce minimum est global.

Proposition 5. Soient $f : \mathbb{R}^n \to \mathbb{R}$ une fonction différentiable. Alors, f est convexe si et seulement si, pour tous $x, y \in \mathbb{R}^n$,

$$f(y) \geq f(x) + (\nabla f(x) \mid y - x).$$

Définition 6. Soient $f : \mathbb{R}^n \to \mathbb{R}$ et $\alpha > 0$. On dit que f est α-convexe si, pour tous $x, y \in \mathbb{R}^n$, pour tout $t \in [0, 1]$,

$$f(tx + (1-t)y) \leq tf(x) + (1-t)f(y) - \frac{\alpha}{2} t(1-t)\|x - y\|^2.$$

Remarque. Toute fonction α-convexe est strictement convexe.

Proposition 7. Soient $f : \mathbb{R}^n \to \mathbb{R}$ une fonction et $\alpha > 0$. Notons,

$$\begin{array}{rccc} g_\alpha & : & \mathbb{R}^n & \longrightarrow & \mathbb{R} \\ & & x & \longmapsto & f(x) - \frac{\alpha}{2}\|x\|^2 \end{array}$$

Alors, f est α-convexe si et seulement si g_α est convexe.

Preuve. Soient $t \in [0, 1]$ et $x, y \in \mathbb{R}^n$. Alors,

$$\begin{aligned} &g_\alpha(tx + (1-t)y) - [tg_\alpha(x) + (1-t)g_\alpha(y)] = \\ &f(tx + (1-t)y) - [tf(x) + (1-t)f(y)] + \frac{\alpha}{2} t(1-t)\|x - y\|^2 \end{aligned} \tag{1}$$

(les détails de ce calcul sont donnés plus loin). Par conséquent,

$$\begin{aligned} &g_\alpha(tx + (1-t)y) \leq [tg_\alpha(x) + (1-t)g_\alpha(y)] \Longleftrightarrow \\ &f(tx + (1-t)y) \leq tf(x) + (1-t)f(y) - \frac{\alpha}{2} t(1-t)\|x - y\|^2 \end{aligned}$$

et donc f est α-convexe si et seulement si g_α est convexe. □

Corollaire 8. Soient $f : \mathbb{R}^n \to \mathbb{R}$ une fonction différentiable et $\alpha > 0$. Alors, f est α-convexe si et seulement si, pour tous $x, y \in \mathbb{R}^n$,

$$f(y) \geq f(x) + (\nabla f(x) \mid y - x) + \frac{\alpha}{2}\|y - x\|^2.$$

Preuve. Reprenons la fonction g_α définie à la proposition précédente. Puisque f est différentiable, g_α est également différentiable, avec,

$$\forall x \in \mathbb{R}^n,\ \nabla g_\alpha(x) = \nabla f(x) - \alpha x.$$

D'où, pour tous $x, y \in \mathbb{R}^n$,

$$\begin{aligned} &g_\alpha(y) - g_\alpha(x) - (\nabla g_\alpha(x) \mid y - x) = \\ &f(y) - f(x) - (\nabla f(x) \mid y - x) - \frac{\alpha}{2}\|x - y\|^2 \end{aligned} \tag{2}$$

(de nouveau, les détails de ce calcul sont donnés plus loin). Par conséquent,

$$\begin{aligned} &g_\alpha(y) \geq g_\alpha(x) + (\nabla g_\alpha(x) \mid y - x) \Longleftrightarrow \\ &f(y) \geq f(x) + (\nabla f(x) \mid y - x) + \frac{\alpha}{2}\|y - x\|^2. \end{aligned}$$

En utilisant les propositions 5 et 7, il s'en suit alors le résultat recherché. □

Corollaire 9. Soient $\alpha > 0$ et $f : \mathbb{R}^n \to \mathbb{R}$ une fonction. Si f est différentiable et α-convexe, alors f est coercive.

Preuve. Soit $y \in \mathbb{R}^n$. D'après la proposition précédente,

$$f(y) \geq f(0) + (\nabla f(0) \mid y) + \frac{\alpha}{2}\|y\|^2.$$

Or, d'après l'inégalité de Cauchy-Schwarz,

$$|(\nabla f(0) \mid y)| \leq \|\nabla f(0)\|\|y\|.$$

Donc,

$$f(y) \geq f(0) - \|\nabla f(0)\|\|y\| + \frac{\alpha}{2}\|y\|^2 = f(0) + \|y\| \left[-\|\nabla f(0)\| + \frac{\alpha}{2}\|y\|\right].$$

Finalement, pour tout $y \in \mathbb{R}^n$,

$$f(y) \geq f(0) + \|y\| \left[-\|\nabla f(0)\| + \frac{\alpha}{2}\|y\|\right] \underset{\|y\|\to+\infty}{\longrightarrow} +\infty,$$

ce qui implique $\lim\limits_{\|y\|\to+\infty} f(y) = +\infty$. □

Définition 10. Une fonction $f : \mathbb{R}^n \to \mathbb{R}$ est dite elliptique si elle est de classe C^1 et s'il existe $\alpha > 0$ tel que f soit α-convexe.

Remarque. Toute fonction elliptique est strictement convexe et coercive.

Détails des calculs pour obtenir l'égalité (1). Tout d'abord,

$$g_\alpha(tx+(1-t)y)-[tg_\alpha(x)+(1-t)g_\alpha(y)]=f(tx+(1-t)y)-\frac{\alpha}{2}\|tx+(1-t)y\|^2-\left[t\left(f(x)-\frac{\alpha}{2}\|x\|^2\right)+(1-t)\left(f(y)-\frac{\alpha}{2}\|y\|^2\right)\right].$$

Puis, en réarrangeant les termes de cette égalité,

$$g_\alpha(tx+(1-t)y)-[tg_\alpha(x)+(1-t)g_\alpha(y)]=f(tx+(1-t)y)-[tf(x)+(1-t)f(y)]-\frac{\alpha}{2}\|tx+(1-t)y\|^2+\frac{\alpha}{2}t\|x\|^2+\frac{\alpha}{2}(1-t)\|y\|^2.$$

Et donc,

$$g_\alpha(tx+(1-t)y)-[tg_\alpha(x)+(1-t)g_\alpha(y)]=f(tx+(1-t)y)-[tf(x)+(1-t)f(y)]+\frac{\alpha}{2}\left[-\|tx+(1-t)y\|^2+t\|x\|^2+(1-t)\|y\|^2\right]. \tag{3}$$

De plus,

$$\|tx+(1-t)y\|^2=\|t(x-y)+y\|^2=t^2\|x-y\|^2+2t(x-y\mid y)+\|y\|^2=t^2\|x-y\|^2+2t(x\mid y)-2t\|y\|^2+\|y\|^2.$$

Puis,

$$-\|tx+(1-t)y\|^2+t\|x\|^2+(1-t)\|y\|^2=-t^2\|x-y\|^2-2t(x\mid y)+t\|x\|^2+t\|y\|^2=t\left[-t\|x-y\|^2-2(x\mid y)+\|x\|^2+\|y\|^2\right]=t[-t\|x-y\|^2+\|x-y\|^2].$$

Et donc,

$$-\|tx+(1-t)y\|^2+t\|x\|^2+(1-t)\|y\|^2=t(1-t)\|x-y\|^2.$$

Ainsi, en reportant l'égalité précédente dans l'égalité (3), il en résulte que,

$$g_\alpha(tx+(1-t)y)-[tg_\alpha(x)+(1-t)g_\alpha(y)]=f(tx+(1-t)y)-[tf(x)+(1-t)f(y)]+\frac{\alpha}{2}t(1-t)\|x-y\|^2.$$

Détails des calculs pour obtenir l'égalité (2). Tout d'abord,

$$g_\alpha(y)-g_\alpha(x)-(\nabla g_\alpha(x)\mid y-x)=f(y)-\frac{\alpha}{2}\|y\|^2+f(x)+\frac{\alpha}{2}\|x\|^2-(\nabla f(x)-\alpha x\mid y-x).$$

Puis,

$$g_\alpha(y)-g_\alpha(x)-(\nabla g_\alpha(x)\mid y-x)=f(y)-f(x)-(\nabla f(x)\mid y-x)+(\alpha x\mid y-x)-\frac{\alpha}{2}\|y\|^2+\frac{\alpha}{2}\|x\|^2.$$

Or,

$$(\alpha x\mid y-x)-\frac{\alpha}{2}\|y\|^2+\frac{\alpha}{2}\|x\|^2=\alpha(x\mid y)-\frac{\alpha}{2}\|y\|^2-\frac{\alpha}{2}\|x\|^2=-\frac{\alpha}{2}\left[\|x\|^2-2(x\mid y)+\|y\|^2\right]=-\frac{\alpha}{2}\|x-y\|^2.$$

Donc,

$$g_\alpha(y)-g_\alpha(x)-(\nabla g_\alpha(x)\mid y-x)=f(y)-f(x)-(\nabla f(x)\mid y-x)-\frac{\alpha}{2}\|x-y\|^2.$$

Formes quadratiques

Soit $\mathbb{K}$ un corps de caractéristique différente de 2. Soit E un $\mathbb{K}$-espace vectoriel de dimension finie $n \in \mathbb{N}^*$.

Définition 1. Soit $q : E \to \mathbb{K}$. On dit que q est une forme quadratique sur E s'il existe une forme bilinéaire symétrique $\varphi : E \times E \to \mathbb{K}$ telle que, pour tout $x \in E$, $q(x) = \varphi(x, x)$. Si tel est le cas, on dit que φ est une forme bilinéaire associée à q.

Proposition 2 (Identité de polarisation). Soit q une forme quadratique sur E. Considérons l'application,

$$\begin{array}{rccc} \varphi & : & E \times E & \longrightarrow & \mathbb{K} \\ & & (x, y) & \longmapsto & \dfrac{1}{2}(q(x+y) - q(x) - q(y)) \end{array}$$

Alors, φ est l'unique forme bilinéaire associée à q. On l'appelle la forme polaire de q.

Preuve. Par définition d'une forme quadratique, il existe au moins une forme bilinéaire symétrique $\psi : E \times E \to \mathbb{K}$ telle que, pour tout $x \in E$, $q(x) = \psi(x, x)$. Alors, pour tout $(x, y) \in E^2$,

$$q(x+y) = \psi(x+y, x+y) = \psi(x, x) + 2\psi(x, y) + \psi(y, y) = q(x) + 2\psi(x, y) + q(y)$$

car ψ est une forme bilinéaire symétrique, et donc,

$$\psi(x, y) = \frac{1}{2}(q(x + y) - q(x) - q(y)) = \varphi(x, y).$$

En particulier, φ est bien une forme bilinéaire associée à q, car elle est égale à β qui en est une. De plus, φ l'unique forme bilinéaire associée à q, puisqu'on vient de montrer que toute forme bilinéaire ψ associée à q est égale à φ. □

Remarque. Dans la preuve ci-dessus, on a utilisé le fait que $2 \in \mathbb{K}^*$, ce qui est vrai car $\mathbb{K}$ est de caractéristique différente de 2.

Définition 3. Soit $\mathcal{B}$ une base de E. On appelle matrice de q dans la base $\mathcal{B}$, et on note $\mathrm{Mat}_{\mathcal{B}}(q)$, la matrice de sa forme polaire.

Définition 4. Soit q une forme quadratique sur E. On appelle noyau de q, et on note $\mathrm{Ker}(q)$, le noyau de sa forme polaire.

Définition 5. Soit q une forme quadratique sur E. On appelle cône isotrope de q, et on note $\mathcal{C}_q$ la partie de E définie par $\mathcal{C}_q = \{x \in E,\ q(x) = 0_{\mathbb{K}}\}$. Notons que $\mathrm{Ker}(q) \subset \mathcal{C}_q$.

Remarque. La terminologie provient du fait que le cône isotrope est stable par multiplication par un scalaire. En revanche, ce n'est pas nécessairement un sous-espace vectoriel de E.

Formule de Stirling généralisée

[FR] **Théorème 1 (Formule de Stirling).** On a $\Gamma(t+1) \underset{t\to+\infty}{\sim} \sqrt{2\pi t}\left(\frac{t}{e}\right)^t$.

Preuve. Rappelons que la fonction Γ est définie pour tout $t \in \mathbb{R}_+^*$ par,

$$\Gamma(t+1) = \int_{-1}^{+\infty} x^t e^{-x} dx.$$

Fixons $t \in \mathbb{R}_+^*$. Tout d'abord, le changement de variable $x = (1+u)t$ donne,

$$\Gamma(t+1) = t^{t+1} \int_{-1}^{+\infty} (1+u)^t e^{-(1+u)t} du.$$

De plus, pour tout $u \in \mathbb{R}$,

$$(1+u)^t e^{-(1+u)t} = e^{t\ln(1+u)} e^{-(1+u)t} = e^{-t(1+u-\ln(1+u))}.$$

En notant,

$$\begin{array}{rccc} \varphi & : &]-1,+\infty[& \longmapsto & \mathbb{R} \\ & & u & \longmapsto & 1+u-\ln(1+u) \end{array}$$

il vient donc,

$$\Gamma(t+1) = t^{t+1} \int_{-1}^{+\infty} e^{-t\varphi(u)} du.$$

Nous allons maintenant appliquer la méthode de Laplace (théorème 1) pour déterminer un équivalent de l'intégrale,

$$\int_0^{+\infty} e^{-t\varphi(u)} du.$$

La fonction φ est de classe C^2 (et même de classe C^∞) sur $]-1,+\infty[$ avec,

$$\forall u \in]-1,+\infty[,\ \varphi'(u) = \frac{u}{1+u},\ \varphi''(u) = \frac{1}{(1+u)^2}.$$

En particulier, $\varphi' > 0$ sur $]0,+\infty[$, $\varphi'(0) = 0$ et $\varphi''(0) = 1 > 0$. En appliquant le point (ii) du théorème précédent (en choisissant pour f la fonction constante égale à 1, et $b = +\infty$), il en résulte que,

$$\int_0^{+\infty} e^{-t\varphi(u)} du \underset{t\to+\infty}{\sim} \sqrt{\frac{\pi}{2}} \frac{e^{-t}}{\sqrt{t}}.$$

Et donc,

$$e^t\sqrt{t} \int_0^{+\infty} e^{-t\varphi(u)} du \underset{t\to+\infty}{\longrightarrow} \sqrt{\frac{\pi}{2}}.$$

D'autre part, le changement de variable $v = -u$ donne,

$$\int_{-1}^{0} e^{-t\varphi(u)} du = \int_0^1 e^{-t\psi(v)} dv.$$

où l'on a noté, pour tout $v \in [0,1[$, $\psi(v) = \varphi(-v)$. La fonction ψ est de classe C^2 (et même de classe C^∞) sur $[0,1[$ avec,

$$\forall v \in]1,+\infty[,\ \psi'(v) = \frac{v}{1-v},\ \psi''(v) = \frac{1}{(1-v)^2}.$$

En particulier, $\psi' > 0$ sur $[0,1[$, $\psi'(0) = 0$ et $\psi''(0) = 1 > 0$. En appliquant le point (ii) du théorème précédent (en choisissant pour f la fonction constante égale à 1, et $b = 1$), il en résulte que,

$$\int_0^1 e^{-t\psi(v)} dv \underset{t\to+\infty}{\sim} \sqrt{\frac{\pi}{2}} \frac{e^{-t}}{\sqrt{t}}.$$

Et donc,

$$e^t\sqrt{t} \int_{-1}^{0} e^{-t\varphi(u)} du \underset{t\to+\infty}{\longrightarrow} \sqrt{\frac{\pi}{2}}.$$

Pour tout $t \in \mathbb{R}_+^*$, notons $g(t) = \sqrt{t}\left(\frac{t}{e}\right)^t$. Alors,

$$\frac{\Gamma(t+1)}{g(t)} = e^t\sqrt{t} \int_{-1}^{0} e^{-t\varphi(u)} du + e^t\sqrt{t} \int_0^{+\infty} e^{-t\varphi(u)} du \underset{t\to+\infty}{\longrightarrow} 2\sqrt{\frac{\pi}{2}}.$$

Et donc $\frac{\Gamma(t+1)}{g(t)} \underset{t\to+\infty}{\longrightarrow} \sqrt{2\pi}$, c'est-à-dire,

$$\Gamma(t+1) \underset{t\to+\infty}{\sim} \sqrt{2\pi t}\left(\frac{t}{e}\right)^t.$$

□

Remarque. L'heuristique du changement de variable est expliqué dans [FR].

Matrices symétriques

Soit $n \in \mathbb{N}^*$. On rappelle que l'application,

$$\begin{array}{rccc} \langle \cdot, \cdot \rangle & : & M_{n,1}(\mathbb{R}) \times M_{n,1}(\mathbb{R}) & \longrightarrow & \mathbb{R} \\ & & (X, Y) & \longmapsto & {}^tXY \end{array}$$

est un produit scalaire sur $M_{n,1}(\mathbb{R})$. La norme induite par ce produit scalaire est,

$$\begin{array}{rccc} \|\cdot\| & : & M_{n,1}(\mathbb{R}) & \longrightarrow & \mathbb{R} \\ & & (X, Y) & \longmapsto & \sqrt{{}^tXX} \end{array}$$

Pour tous $d_1, \dots, d_n \in \mathbb{R}$, on note $\mathrm{Diag}(d_1, \dots, d_n)$ la matrice diagonale de coefficients diagonaux $d_1, \dots, d_n$.

Lemme 1. Soit $S \in S_n(\mathbb{R})$. Alors, les valeurs propres de S dans $\mathbb{C}$ sont toutes réelles. De plus,

(i) $S \in S_n^+(\mathbb{R})$ si et seulement si toutes ses valeurs propres sont positives;

(ii) $S \in S_n^{++}(\mathbb{R})$ si et seulement si toutes ses valeurs propres sont strictement positives.

Preuve. Le premier point résulte du théorème spectral. Les points suivants se démontrent de façon similaire, en mettant des inégalités larges dans (i) et des inégalités strictes dans (ii).

(i) On procède par double implication.

• $\boxed{\Longrightarrow}$. Supposons que $S \in S_n^+(\mathbb{R})$. Soit λ une valeur propre de S, c'est-à-dire qu'il existe $X \in M_{n,1}(\mathbb{R})$ non nul tel que $SX = \lambda X$. Alors,

$$ {}^tXSX = \lambda {}^tXX = \lambda \|X\|^2.$$

Or, ${}^tXSX \in \mathbb{R}_+$, car $S \in S_n^+(\mathbb{R})$. Puisque $\|X\| > 0$, ceci implique $\lambda \in \mathbb{R}_+$.

• $\boxed{\Longleftarrow}$. Réciproquement, supposons que les valeurs propres de S soient toutes positives. D'après le théorème spectral, il existe $D \in D_n(\mathbb{R})$ et $P \in O_n(\mathbb{R})$ tels que $S = {}^tPDP$. Notons ensuite $\lambda_1, \dots, \lambda_n$ les coefficients de D. Ce sont des réels positifs, car ce sont les valeurs propres de S. Posons,

$$\Delta = \mathrm{Diag}\left(\sqrt{\lambda_1}, \dots \sqrt{\lambda_n}\right).$$

Alors, ${}^t\Delta = \Delta$ et $\Delta^2 = D$. D'où, pour tout $X \in M_{n,1}(\mathbb{R})$,

$$ {}^tXSX = {}^tX{}^tPDPX = {}^t(\Delta PX)\Delta PX = \|\Delta PX\|^2 \geq 0$$

ce qui prouve que $S \in S_n^+(\mathbb{R})$.

(ii) On procède par double implication.

• $\boxed{\Longrightarrow}$. Supposons que $S \in S_n^{++}(\mathbb{R})$. Soit λ une valeur propre de S, c'est-à-dire qu'il existe $X \in M_{n,1}(\mathbb{R})$ non nul tel que $SX = \lambda X$. Alors,

$$ {}^tXSX = \lambda {}^tXX = \lambda \|X\|^2.$$

Or, ${}^tXSX \in \mathbb{R}_+^*$, car $S \in S_n^{++}(\mathbb{R})$. Puisque $\|X\| > 0$, ceci implique $\lambda \in \mathbb{R}_+^*$.

• $\boxed{\Longleftarrow}$. Réciproquement, supposons que les valeurs propres de S soient toutes strictement positives. D'après le théorème spectral, il existe $D \in D_n(\mathbb{R})$ et $P \in O_n(\mathbb{R})$ tels que $S = {}^tPDP$. Notons ensuite $\lambda_1, \dots, \lambda_n$ les coefficients de D. Ce sont des réels strictement positifs, car ce sont les valeurs propres de S. Posons,

$$\Delta = \mathrm{Diag}\left(\sqrt{\lambda_1}, \dots \sqrt{\lambda_n}\right).$$

Alors, $\Delta \in GL_n(\mathbb{R})$ car $\det(\Delta) = \sqrt{\lambda_1 \dots \lambda_n} > 0$, ${}^t\Delta = \Delta$ et $\Delta^2 = D$. D'où, pour tout $X \in M_{n,1}(\mathbb{R})$ non nul,

$$ {}^tXSX = {}^tX{}^tPDPX = {}^t(\Delta PX)\Delta PX = \|\Delta PX\|^2 > 0$$

car $\Delta P \in GL_n(\mathbb{R})$ et $X \neq 0$, ce qui prouve que $S \in S_n^{++}(\mathbb{R})$.

□

Théorème 2. Notons $\overline{S_n^{++}(\mathbb{R})}$ l'adhérence de $S_n^{++}(\mathbb{R})$ dans $M_n(\mathbb{R})$. Alors,

$$\overline{S_n^{++}(\mathbb{R})} = S_n^+(\mathbb{R}).$$

Preuve. Montrons tout d'abord que $S_n^+(\mathbb{R})$ est fermé dans l'espace vectoriel $S_n(\mathbb{R})$. Soit $(S_p)_{p\in\mathbb{N}}$ une suite de matrices symétriques positives convergente vers une matrice S dans $M_n(\mathbb{R})$. Montrons que $S \in S_n^+(\mathbb{R})$. Tout d'abord, $S \in S_n(\mathbb{R})$ car $S_n(\mathbb{R})$ est une partie fermée de $M_n(\mathbb{R})$, en tant que sous-espace vectoriel de dimension finie de $M_n(\mathbb{R})$. Considérons ensuite $X \in M_{n,1}(\mathbb{R})$. L'application $M \mapsto {}^tXMX$ est continue, car linéaire sur $M_n(\mathbb{R})$ qui est de dimension finie. Par conséquent, la suite $({}^tXS_pX)_{p\in\mathbb{N}}$ converge vers ${}^tXSX$. Or, il s'agit d'une suite de réels positifs, car $(S_p)_{p\in\mathbb{N}}$ est une suite de matrices

symétriques positives. Donc ${}^tXSX \geq 0$ en tant que limite d'une suite de réels positifs, ce qui prouve que $S \in S_n^+(\mathbb{R})$. Et puisque $S_n^{++}(\mathbb{R}) \subset S_n^+(\mathbb{R})$, il vient,

$$\overline{S_n^{++}(\mathbb{R})} \subset \overline{S_n^+(\mathbb{R})} = S_n^+(\mathbb{R}).$$

Il reste à montrer l'inclusion réciproque. Soit $S \in S_n^+(\mathbb{R})$. D'après le théorème spectral, il existe $D \in D_n(\mathbb{R})$ et $P \in O_n(\mathbb{R})$ tels que $S = {}^tPDP$. Notons ensuite $\lambda_1, \ldots, \lambda_n$ les coefficients de D. Il s'agit des valeurs propres de S. Par le lemme 1, ce sont des réels positifs, car $S \in S_n^+(\mathbb{R})$. Soit $p \in \mathbb{N}$. Posons,

$$D_p = \mathrm{Diag}\left(\lambda_1 + \frac{1}{p}, \ldots, \lambda_n + \frac{1}{p}\right),$$

puis $S_p = {}^tPD_pP$. Alors, $S_p \in S_n(\mathbb{R})$ car,

$${}^tS_p = {}^tP{}^tD_p{}^t({}^tP) = {}^tPD_pP = S_p.$$

De plus, les valeurs propres de S_p sont $\lambda_1 + \frac{1}{p}, \ldots, \lambda_n + \frac{1}{p}$, qui sont des réels strictement positifs. D'où $S_p \in S_n^{++}(\mathbb{R})$ d'après le lemme 1. D'autre part, $(D_p)_{p\in\mathbb{N}}$ converge vers D dans $M_n(\mathbb{R})$. En effet, en considérant la norme $\|\cdot\|_\infty$, définie pour tout $A = (a_{i,j})_{1\leq i,j\leq n} \in M_n(\mathbb{R})$ par $\|A\|_\infty = \max\limits_{1\leq i,j\leq n} |a_{i,j}|$, il vient,

$$\forall p \in \mathbb{N},\ \|D_p - D\|_\infty = \frac{1}{p} \underset{p\to+\infty}{\longrightarrow} 0.$$

et donc,

$$D_p \underset{p\to+\infty}{\longrightarrow} D.$$

Or, l'application $M \mapsto {}^tPMP$ est continue, car linéaire sur $M_n(\mathbb{R})$ qui est de dimension finie. D'où,

$$S_p = {}^tPD_pP \underset{p\to+\infty}{\longrightarrow} {}^tPDP = S$$

ce qui prouve finalement que $S \in \overline{S_n^{++}(\mathbb{R})}$. □

Ordre d'un élément dans un groupe

Lemme 1. Soient $\varphi : G \to H$ un isomorphisme de groupes et $g \in G$. Alors, g est d'ordre fini si et seulement si $\varphi(g)$ est d'ordre fini, et si tel est le cas, g et $\varphi(g)$ ont le même ordre. En particulier, deux éléments conjugués de G ont le même ordre.

Preuve. Pour tout $l \in \mathbb{N}^*$,

$$\varphi(g)^l = e \iff \varphi(g^l) = e \iff g^l = e$$

(la première équivalence résulte du fait que φ est un morphisme de groupes, la seconde du fait que φ est bijective). De cette équivalence, on en déduit que g est d'ordre fini si et seulement si $\varphi(g)$ est d'ordre fini, et que si tel est le cas, g et $\varphi(g)$ ont le même ordre. En particulier, si $g \in G$, alors en appliquant ce résultat à l'automorphisme $\varphi : h \mapsto ghg^{-1}$, on obtient que si $h \in G$ est d'ordre fini alors ghg^{-1} est aussi d'ordre fini et son ordre est égal à celui de h. □

[JR] **Lemme 2.** Soit G un groupe. Soient $x, y \in G$ d'ordre fini tels que $xy = yx$. Notons a et b l'ordre respectif de x et y dans G. Supposons que a et b soient premiers entre eux. Alors, xy est d'ordre ab dans G.

Preuve. Puisque x et y commutent, on a, pour tout $k \in \mathbb{N}$, $(xy)^k = x^k y^k$. En particulier, $(xy)^{ab} = x^{ab}y^{ab} = (x^a)^b(y^b)^a = (e_G)^b(e_G)^a = e_G$. De plus, si $(xy)^k = e_G$ alors $(xy)^{ak} = e_G$, puis,

$$e_G = (xy)^{ka} = (x^a)^k y^{ka} = y^{ka}.$$

Par suite, b divise ka, et donc b divise k d'après le lemme de Gauss (car b est premier avec a). De façon analogue, on obtient que a divise k. Puisque a et b sont premiers entre eux, il s'en suit que ab divise k. Finalement, xy est donc d'ordre ab dans G. □

Probabilité que deux permutations commutent

Proposition 1. La suite $(n(\mathfrak{S}_k))_{k\in\mathbb{N}^*}$ converge vers 0.

Preuve. Pour tout $k \in \mathbb{N}^*$, notons m_k le cardinal de l'ensemble des partitions de l'entier k. Soit $k \in \mathbb{N}^*$. Dans $\mathfrak{S}_k$, il y a m_k classes de conjugaison distinctes. Or, dans la preuve du théorème de Dixon, nous avons montré que, pour tout groupe G, le nombre $n(G)$ est égal au nombre de classes de conjugaison de G divisé par le cardinal de G. Par conséquent,

$$n(\mathfrak{S}_k) = \frac{m_k}{|\mathfrak{S}_k|} = \frac{m_k}{k!}.$$

Or, $m_k \leq 2^k$. Donc, pour tout $k \in \mathbb{N}^*$,

$$0 \leq n(\mathfrak{S}_k) \leq \frac{2^k}{k!} \underset{k\to+\infty}{\longrightarrow} 0.$$

De cet encadrement, on en déduit que $\lim\limits_{k\to+\infty} n(\mathfrak{S}_k) = 0$. □

Mesure de probabilité associée à une série

La proposition suivante est un rappel sur les probabilités discrètes.

Proposition 1. Soit $(a_n)_{n\in\mathbb{N}^*}$ une suite de réels positifs telle que $\sum_{n=1}^{+\infty} a_n = 1$. Alors, il existe une unique mesure de probabilité $\mathbb{P}$ sur $(\mathbb{N}^*, \mathcal{P}(\mathbb{N}^*))$ telle que,

$$\forall n \in \mathbb{N},\ \mathbb{P}(\{n\}) = a_n.$$

Preuve. Si $\mathbb{P}$ est une mesure de probabilité sur $(\mathbb{N}^*, \mathcal{P}(\mathbb{N}^*))$ telle que,

$$\forall n \in \mathbb{N},\ \mathbb{P}(\{n\}) = a_n,$$

alors, pour tout $A \subset \mathbb{N}^*$, la σ-additivité de $\mathbb{P}$ donne,

$$\mathbb{P}(A) = \mathbb{P}\left(\bigcup_{n\in A} \{n\}\right) = \sum_{n\in A} \mathbb{P}(\{n\}) = \sum_{n\in A} a_n.$$

D'où l'unicité. Montrons ensuite que l'application,

$$\begin{array}{rccc} \mathbb{P} & : & \mathcal{P}(\mathbb{N}^*) & \longrightarrow & [0,1] \\ & & A & \longmapsto & \displaystyle\sum_{n\in A} a_n \end{array}$$

est une mesure de probabilité sur $(\mathbb{N}^*, \mathcal{P}(\mathbb{N}^*))$ convenant à notre problème.

- $\mathbb{P}$ est bien définie de $\mathcal{P}(\mathbb{N}^*)$ dans $[0,1]$ car si $A \subset \mathbb{N}^*$

$$0 \leq \sum_{n\in A} a_n \leq \sum_{n\in\mathbb{N}} a_n = 1$$

(par positivité des a_n).

- $\mathbb{P}(\varnothing) = 0$ (par convention, toute somme indicée par $\varnothing$ est vide).
- Soit $(A_k)_{k\in\mathbb{N}}$ une suite de parties deux à deux disjointes de $\mathbb{N}^*$. Notons,

$$A = \bigcup_{k\in\mathbb{N}^*} A_k.$$

Alors,

$$\mathbb{P}(A) = \sum_{n\in A} a_n = \sum_{n\in\mathbb{N}} a_n 1_A(n) = \sum_{n\in\mathbb{N}} \sum_{k\in\mathbb{N}} a_n 1_{A_k}(n)$$

(la dernière égalité résulte du fait que $(A_k)_{k\in\mathbb{N}^*}$ est une partition de A). Puis, d'après le théorème de Fubini-Tonelli,

$$\mathbb{P}(A) = \sum_{k\in\mathbb{N}} \sum_{n\in\mathbb{N}} a_n 1_{A_k}(n) = \sum_{k\in\mathbb{N}} \sum_{n\in A_k} a_n = \sum_{k\in\mathbb{N}} \mathbb{P}(A_k).$$

Ainsi, $\mathbb{P}$ est une mesure de probabilité sur $(\mathbb{N}^*, \mathcal{P}(\mathbb{N}^*))$. De plus, cette mesure de probabilité vérifie bien, pour tout $n \in \mathbb{N}$, $\mathbb{P}(\{n\}) = a_n$. D'où l'existence. □

Racines primitives de l'unité et polynômes cyclotomiques

Soit $n \in \mathbb{N}^*$. On note U_n l'ensemble des racines n-ièmes de l'unité du corps des nombres complexes. Alors,

$$U_n = \{z \in \mathbb{C},\ z^n = 1\} = \left\{e^{\frac{2ik\pi}{n}},\quad 0 \leq k \leq n-1\right\}.$$

De plus, U_n est un sous-groupe du groupe multiplicatif $\mathbb{C}^*$ de cardinal n. Donc, U_n est cyclique en tant que sous-groupe fini du groupe des inversibles d'un corps commutatif.

Définition 1. Une racine n-ième de l'unité z est dite primitive si elle est d'ordre exactement n dans le groupe multiplicatif $\mathbb{C}^*$, c'est-à-dire que,

$$n = \min\left\{k \in \mathbb{N}^*,\ z^k = 1\right\}.$$

On note P_n l'ensemble des racines primitives n-ième de l'unité.

Proposition 2. Les racines primitives n-ièmes de l'unité sont les $e^{\frac{2ik\pi}{n}}$ pour $0 \leq k \leq n-1$ tels que $\text{pgcd}(k,n) = 1$.

Preuve. Soit $k \in \{0,\dots,n-1\}$ tel que $\text{pgcd}(k,n) = 1$. Notons ω l'ordre de $e^{\frac{2ik\pi}{n}}$ dans le groupe multiplicatif $\mathbb{C}^*$. Montrons que $\omega = n$. Tout d'abord, ω divise n par le théorème de Lagrange. Par ailleurs,

$$e^{\frac{2ik\omega\pi}{n}} = \left(e^{\frac{2ik\pi}{n}}\right)^{\omega} = 1$$

et donc $\frac{2k\omega\pi}{n} \in 2\pi\mathbb{Z}$, c'est-à-dire $n \mid \omega k$. Or, $\text{pgcd}(k,n) = 1$. D'après le lemme de Gauss, n divise donc ω. D'où, $\omega = n$, c'est-à-dire que $e^{\frac{2ik\pi}{n}}$ est une racine primitive n-ième de l'unité.

Réciproquement, supposons que ζ soit une racine primitive n-ième de l'unité. Puisque ζ est une racine n-ième de l'unité, il existe $k \in \{0,\dots,n-1\}$ tel que $\zeta = e^{\frac{2ik\pi}{n}}$. Soit $d \in \mathbb{N}^*$ un diviseur commun à k et n, c'est-à-dire qu'il existe $(l,m) \in \mathbb{N}^*$ tels que $k = ld$ et $n = md$. Alors, $\zeta = e^{\frac{2il\pi}{m}}$, et donc $\zeta^m = 1$. Par suite, m divise n car ζ est d'ordre n dans le groupe multiplicatif $\mathbb{C}^*$, puis $m = n$ et donc $d = 1$, ce qui prouve finalement que $\text{pgcd}(k,n) = 1$. □

Proposition 3. L'ensemble $\{P_d,\ d \mid n\}$ est une partition de U_n.

Preuve. On montre que $\{P_d,\ d \mid n\}$ est un ensemble de parties non vides et deux à deux disjointes de U_n dont la réunion est égale à U_n.

(i) Si $d \in \mathbb{N}^*$ tel que $d \mid n$ alors $P_d \neq \varnothing$ car, d'après la proposition 2, $e^{\frac{i\pi}{d}} \in P_d$.

(ii) Soient $d, d' \in \mathbb{N}^*$ tels que d et d' divisent n. Si $P_d \cap P_{d'} \neq \varnothing$ alors en considérant $\zeta \in P_d \cap P_{d'}$ il s'en suit que ζ est à la fois d'ordre d et d' dans le groupe multiplicatif $\mathbb{C}^*$, d'où $d = d'$.

(iii) Montrons que,

$$U_n = \bigcup_{d|n} P_d.$$

- $\boxed{\subset}$. Soit $\zeta \in U_n$. Notons d l'ordre de ζ dans le groupe multiplicatif $\mathbb{C}^*$. Alors, $\zeta \in P_d$. De plus, d'après le théorème de Lagrange, d divise n. Donc,
$$\zeta \in \bigcup_{d|n} P_d.$$
- $\boxed{\supset}$. Réciproquement, soit $\zeta \in P_d$ pour $d \in \mathbb{N}^*$ divisant n. Alors, il existe $m \in \mathbb{N}^*$ tel que $n = md$ et que $\zeta^d = 1$, puis $\zeta^n = \zeta^{md} = (\zeta^d)^m = 1$, et donc $\zeta \in U_n$.

Par double inclusion, il en résulte que,

$$U_n \subset \bigcup_{d|n} P_d.$$

Ainsi, $\{P_d,\ d \mid n\}$ est une partition de U_n. □

Définition 4. Pour tout $n \in \mathbb{N}^*$, on appelle n-ième polynôme cyclotomique, et on note Φ_n, le polynôme de $\mathbb{C}[X]$ défini par :

$$\Phi_n = \prod_{\zeta \in P_n} (X - \zeta).$$

Remarque. D'après la proposition 2, on a aussi,

$$\Phi_n = \prod_{\substack{1 \leq k \leq n-1 \\ \text{pgcd}(k,n)=1}} \left(X - e^{\frac{2ik\pi}{n}}\right).$$

et donc $\deg(\Phi_n) = \text{card}(\{1 \leq k \leq n,\ \text{pgcd}(k,n) = 1\}) = \varphi(n)$ où $\varphi(n)$ désigne l'indicatrice d'Euler de n.

Lemme 5. Soient $A, B \in \mathbb{Z}[X]$. Supposons que B soit un polynôme unitaire. Alors, il existe $Q, R \in \mathbb{Z}[X]$ tels que $A = BQ + R$ avec $\deg(R) < \deg(B)$ ou $R = 0$.

Preuve. Montrons par récurrence sur $n \in \mathbb{N}$ que si $A \in \mathbb{Z}[X]$ vérifie $\deg(A) \leq n$ ou $A = 0$ alors il existe $Q, R \in \mathbb{Z}[X]$ tels que $A = BQ + R$.

Initialisation. Soit $A \in \mathbb{Z}[X]$ tel que le polynôme A soit constant. On distingue deux cas.

- Si B est constant alors $B = 1$ et donc $Q = A$ et $R = 0$ conviennent.
- Sinon, $Q = 0$ et $R = A$ conviennent car $\deg(B) > \deg(A)$.

Hérédité. Supposons ensuite le résultat établi pour un rang $n \in \mathbb{N}$ fixé. Soit $A \in \mathbb{Z}[X]$ un polynôme tel que $\deg(A) \leq n + 1$. Si $\deg(A) < n$, l'hypothèse de récurrence permet de conclure directement. On peut donc supposer $\deg(A) = n + 1$. Notons $p = \deg(B)$. On distingue ensuite deux cas.

- Cas n°1 : $p \leq n + 1$. Alors, $A - a_{n+1}X^{n+1-p}B$ est un polynôme de degré inférieur ou égal à n. Par hypothèse de récurrence, il existe donc $P, S \in \mathbb{Z}[X]$ tels que,

  $$A - a_{n+1}X^{n+1-p}B = PB + S,$$

  avec $\deg(R) < \deg(B)$ ou $R = 0$. Par suite,

  $$A = (a_{n+1}X^{n+1-p} + P)B + S$$

  avec $\deg(R) < \deg(B)$ ou $R = 0$, et donc $Q = a_{n+1}X^{n+1-p} + P$ et $R = S$ conviennent.
- Cas n°2 : $p > n + 1$. Alors, $Q = 0$ et $R = A$ conviennent car $\deg(B) > \deg(A)$.

Dans tous les cas, le résultat est donc vrai au rang $n + 1$, ce qui termine la récurrence. □

[FG] **Lemme 6.** Soit $n \in \mathbb{N}^*$. Alors,

(i) $X^n - 1 = \prod_{d|n} \Phi_d$,

(ii) $\Phi_n \in \mathbb{Z}[X]$,

(iii) Φ_n et $\prod_{d|n} \Phi_d$ sont premiers entre eux dans $\mathbb{Q}[X]$.

Preuve.

(i) Cette égalité résulte du fait que les racines n-ièmes de l'unité d'ordre d pour $d \mid n$ partitionnent l'ensemble des racines n-ièmes de l'unité.

(ii) Montrons par récurrence sur $n \in \mathbb{N}$ que, pour tout $k \in \{1, \dots, n\}$, $\Phi_k \in \mathbb{Z}[X]$.

Initialisation. On a $\Phi_1 = X - 1 \in \mathbb{Z}[X]$.

Hérédité. Supposons le résultat établi pour un rang $n \in \mathbb{N}^*$ fixé. Alors,

$$X^{n+1} - 1 = \Phi_{n+1} \prod_{\substack{d|n+1 \\ d \neq n+1}} \Phi_d \tag{1}$$

avec $X^{n+1} - 1 \in \mathbb{Z}[X]$ et $\prod_{\substack{d|n+1 \\ d \neq n+1}} \Phi_d \in \mathbb{Z}[X]$ en conséquence de l'hypothèse de récurrence. Par ailleurs, puisque $\prod_{\substack{d|n+1 \\ d \neq n+1}} \Phi_d$ est un polynôme unitaire, il existe $Q, R \in \mathbb{Z}[X]$ tels que,

$$X^{n+1} - 1 = Q \prod_{\substack{d|n+1 \\ d \neq n+1}} \Phi_d + R \tag{2}$$

avec $\deg(R) < \deg(\Phi_{n+1})$ ou $R = 0$. Les égalités (1) et (2) fournissent tout deux la division euclidienne de $X^{n+1} - 1$ par $\prod_{\substack{d|n+1 \\ d \neq n+1}} \Phi_d$ dans $\mathbb{C}[X]$. De l'unicité de cette décomposition, on en déduit que $\Phi_{n+1} = Q$ et $R = 0$. En particulier, $\Phi_{n+1} \in \mathbb{Z}[X]$, ce qui termine la récurrence.

(iii) Soit $d \in \mathbb{N}^*$ un diviseur strict de n. Notons tout d'abord que Φ_n et Φ_d sont sans racines communes dans $\mathbb{C}$, car les racines respectives de Φ_n et Φ_d sont les éléments de U_n d'ordre n et d. A fortiori, Φ_n et $\prod_{d|n} \Phi_d$ sont sans racines communes dans $\mathbb{C}$, donc premiers entre eux dans $\mathbb{Q}[X]$. En effet, sinon il existerait un diviseur non constant D commun à ces deux polynômes dans $\mathbb{Q}$, et donc D posséderait au moins une racine dans $\mathbb{C}$, qui serait alors une racine commune de ces deux polynômes dans $\mathbb{C}$.

□

Régularité de l'inverse dans $GL_n(\mathbb{R}^n)$

Soit $n \in \mathbb{N}^*$.

Lemme 1. L'application,

$$\begin{array}{rccc} \Lambda & : & GL(\mathbb{R}^n) & \longrightarrow & GL(\mathbb{R}^n) \\ & & \varphi & \longrightarrow & \varphi^{-1} \end{array}$$

est de classe C^∞ sur $GL(\mathbb{R}^n)$.

Preuve. Soit $\mathscr{B}$ une base de E. L'application,

$$\begin{array}{rccc} \Phi & : & GL(E) & \longrightarrow & GL_n(\mathbb{R}) \\ & & \varphi & \longrightarrow & \mathrm{Mat}_{\mathscr{B}}(\varphi) \end{array}$$

est un isomorphisme. En particulier, Φ et Φ^{-1} sont de classe C^∞, comme toute application linéaire sur un espace de dimension finie. Montrons ensuite que, l'application,

$$\begin{array}{rccc} \Theta & : & GL_n(\mathbb{R}) & \longrightarrow & GL_n(\mathbb{R}) \\ & & M & \longrightarrow & M^{-1} \end{array}$$

est de classe C^∞ sur $GL_n(\mathbb{R})$. Pour tout $M \in GL_n(\mathbb{R})$,

$$M^{-1} = \frac{1}{\det(M)}\, {}^t\mathrm{Comat}(M).$$

Or, $M \mapsto \det(M)$ et $M \mapsto {}^t\mathrm{Comat}(M)$ sont des polynômes en les coefficients de M, et $M \mapsto \det(M)$ ne s'annule pas sur $GL_n(\mathbb{R})$. Par conséquent, chaque coefficient de M^{-1} est une fraction rationnelle en les coefficients de M, dont le dénominateur ne s'annule pas. En particulier, Θ est donc de classe C^∞ sur $GL_n(\mathbb{R})$. Finalement, il en résulte que $\Lambda = \Phi^{-1} \circ \Theta \circ \Phi$ est de classe C^∞ sur $GL(E)$. □

Séparabilité

Définition 1. Un espace métrique est dit séparable s'il contient une partie dense et dénombrable.

Proposition 2. Tout espace métrique compact est séparable.

Preuve. Soit (K, d) un espace métrique compact. Soit $p \in \mathbb{N}^*$. Alors,

$$K \subset \bigcup_{x \in K} B\left(x, \frac{1}{p}\right).$$

Par compacité de K, il existe donc $x_1^{(p)}, \ldots, x_{m_p}^{(p)} \in X$ tels que,

$$K = \bigcup_{k=1}^{m_p} B\left(x_k^{(p)}, \frac{1}{p}\right). \tag{1}$$

Posons ensuite,

$$N = \bigcup_{p \in \mathbb{N}^*} \{x_k^{(p)},\ 1 \leq k \leq m_p\}.$$

Tout d'abord, N est dénombrable en tant qu'union dénombrable d'ensembles finis. Montrons que N est dense dans K. Soient $x \in K$ et $\varepsilon > 0$. Considérons $p \in \mathbb{N}^*$ tel que $\frac{1}{p} \leq \varepsilon$. L'inclusion (1) assure l'existence de $k \in \{1, \ldots, m_p\}$ tel que $d(x, x_k) \leq \frac{1}{p}$. Finalement, x_k est un élément de N tel que $d(x, x_k) \leq \varepsilon$, ce qui prouve que N est dense dans K, et donc que K est séparable. □

Théorème de Carathéodory

Soit $n \in \mathbb{N}^*$. Soient E un $\mathbb{R}$-espace vectoriel de dimension n et A une partie non vide de E. On note $\mathrm{Conv}(A)$ l'enveloppe convexe de A.

Théorème 1 (Théorème de Carathéodory). L'enveloppe convexe de A est l'ensemble des combinaisons convexes de famille de $n+1$ points de A.

Preuve. On sait déjà que,

$$\mathrm{Conv}(A) = \left\{ \lambda_1 a_1 + \cdots + \lambda_p a_p \;\middle|\; \begin{array}{l} p \in \mathbb{N}^*, \\ \lambda_1, \ldots, \lambda_p \in \mathbb{R}_+, \\ \lambda_1 + \cdots + \lambda_p = 1, \\ a_1, \ldots, a_p \in A. \end{array} \right\}. \tag{1}$$

(ce résultat reste vrai en dimension infinie). Notons ensuite B l'ensemble des combinaisons convexes de famille de $n+1$ points de A, c'est-à-dire :

$$B = \left\{ \lambda_1 a_1 + \cdots + \lambda_{n+1} a_{n+1} \;\middle|\; \begin{array}{l} \lambda_1, \ldots, \lambda_{n+1} \in \mathbb{R}_+ \\ \lambda_1 + \cdots + \lambda_{n+1} = 1, \\ a_1, \ldots, a_{n+1} \in A. \end{array} \right\}.$$

Avec ces notations, il suffit de montrer que $B = \mathrm{Conv}(A)$. Puisque $\mathrm{Conv}(A)$ est une partie convexe de A contenant A, on a bien l'inclusion $B \subset \mathrm{Conv}(A)$ (là encore, ce résultat reste vrai en dimension infinie). Réciproquement, fixons $x \in \mathrm{Conv}(A)$. D'après l'égalité (1), l'ensemble,

$$P := \left\{ p \in \mathbb{N}^* \;\middle|\; \begin{array}{l} \exists \lambda_1, \ldots, \lambda_p \in \mathbb{R}_+, \\ \exists a_1, \ldots, a_p \in A \mid \\ \lambda_1 + \cdots + \lambda_p = 1, \\ x = \lambda_1 a_1 + \cdots + \lambda_p a_p. \end{array} \right\}$$

est une partie non vide de $\mathbb{N}^*$. Par conséquent, P admet un minimum que l'on note p. Puisque $p \in P$, il existe $\lambda_1, \ldots, \lambda_p \in \mathbb{R}_+$ et $a_1, \ldots, a_p \in A$ tels que,

$$\begin{cases} \lambda_1 + \cdots + \lambda_p = 1, \\ x = \lambda_1 a_1 + \cdots + \lambda_p a_p. \end{cases}$$

Pour conclure que $x \in B$, il suffit de montrer que $p \leq n+1$. Pour cela, on raisonne par l'absurde en supposant que $p \geq n+2$. Sous-cette hypothèse, l'application linéaire suivante :

$$\begin{array}{rccc} \phi & : & \mathbb{R}^p & \longrightarrow & E \times \mathbb{R} \\ & & (\mu_1, \ldots, \mu_p) & \longmapsto & \left(\displaystyle\sum_{i=1}^p \mu_i, \sum_{i=1}^p \mu_i a_i \right) \end{array}$$

est non injective, car d'après le théorème du rang,

$$\dim(\mathrm{Ker}(\phi)) = \dim(\mathbb{R}^p) - \dim(\mathrm{Im}(\phi))$$

ce qui implique,

$$\dim(\mathrm{Ker}(\phi)) \geq \dim(\mathbb{R}^p) - \dim(E \times \mathbb{R}) = p - (n+1) \geq 1.$$

On peut donc considérer $(\mu_1, \ldots, \mu_p) \in \mathbb{R}^p \setminus \{(0, \ldots, 0)\}$ tel que,

$$\begin{cases} \mu_1 + \cdots + \mu_p = 0, \\ \mu_1 a_1 + \cdots + \mu_p a_p = 0. \end{cases}$$

Avec ces notations, il vient, pour tout $\nu \in \mathbb{R}$,

$$\begin{cases} \displaystyle\sum_{i=1}^p (\lambda_i + \nu \mu_i) = 1, \\ \displaystyle x = \sum_{i=1}^p (\lambda_i + \nu \mu_i) a_i. \end{cases}$$

Soit $j \in [\![1, p]\!]$ tel que,

$$-\frac{\lambda_j}{\mu_j} = \min_{\mu_i > 0} -\frac{\lambda_i}{\mu_i}$$

(notons qu'il existe bien au moins un $\mu_i > 0$ car ils sont non tous nuls et leur somme vaut 0). On choisit alors $\nu = -\dfrac{\lambda_j}{\mu_j} \leq 0$ de telle sorte que,

(i) $\lambda_j + \nu \mu_j = 0$;

(ii) pour tout $i \in [\![1, p]\!]$, $\lambda_i + \nu \mu_i \geq 0$.

En effet,

$$\lambda_j + \nu \mu_j = \lambda_j + \left(-\frac{\lambda_j}{\mu_j} \right)$$

d'où le point (i), et pour tout $i \in [\![1, p]\!]$,

- si $\mu_i > 0$ alors $\lambda_i + \nu \mu_i \geq \lambda_i + \left(-\frac{\lambda_i}{\mu_i} \right) \mu_i = 0$;

- si $\mu_i \leq 0$ alors $\nu\mu_i \geq 0$ et donc $\lambda_i + \nu\mu_i \geq 0$ (car $\lambda_i \geq 0$);

d'où le point (ii). Ainsi,

$$\begin{cases} 1 = \displaystyle\sum_{i=1}^{p}(\lambda_i + \nu\mu_i) = \sum_{\substack{i\neq j \\ i=1}}^{p}(\lambda_i + \nu\mu_i), \\ x = \displaystyle\sum_{i=1}^{p}(\lambda_i + \nu\mu_i)a_i = \sum_{\substack{i\neq j \\ i=1}}^{p}(\lambda_i + \nu\mu_i)a_i, \end{cases}$$

ce qui contredit la minimalité de p. D'où $p \leq n+1$. Si $p = n+1$, il est clair que $x \in B$. Sinon, pour tout $i \in [\![p, n+1]\!]$, on note $\lambda_i = 0$ et $a_i = a_1$. Avec ces notations, il vient,

$$\begin{cases} \lambda_1 + \cdots + \lambda_{n+1} = \lambda_1 + \cdots + \lambda_p = 1, \\ x = \lambda_1 a_1 + \cdots + \lambda_p a_p = \lambda_1 a_1 + \cdots + \lambda_{n+1} a_{n+1}. \end{cases}$$

ce qui prouve que $x \in B$. Finalement, $\mathrm{Conv}(A) = B$ (par double inclusion). □

Corollaire 2. Si A est compact, alors $\mathrm{Conv}(A)$ est compact.

Preuve. Soit $(x^{(k)})_{k\in\mathbb{N}}$ une suite d'éléments de $\mathrm{Conv}(A)$. D'après le théorème de Carathéodory, pour tout $k \in \mathbb{N}$, il existe $\lambda_1^{(k)}, \ldots, \lambda_{n+1}^{(k)}$ des réels positifs et $a_1^{(k)}, \ldots, a_{n+1}^{(k)} \in A$ tels que,

$$\begin{cases} x^{(k)} = \displaystyle\sum_{p=1}^{n+1} \lambda_p^{(k)} a_p^{(k)}, \\ \displaystyle\sum_{p=1}^{n+1} \lambda_p^{(k)} = 1. \end{cases}$$

En particulier, pour tout $k \in \mathbb{N}$, et pour tout $p \in [\![1, n+1]\!]$,

$$0 \leq \lambda_p^{(k)} \leq \lambda_1^{(k)} + \cdots + \lambda_{n+1}^{(k)} = 1$$

ce qui prouve que les suites $(\lambda_1^{(k)})_{k\in\mathbb{N}}, \ldots, (\lambda_{n+1}^{(k)})_{k\in\mathbb{N}}$ sont bornées. De plus, les suites $(a_1^{(k)})_{k\in\mathbb{N}}, \ldots, (a_{n+1}^{(k)})_{k\in\mathbb{N}}$ sont à termes dans A qui est compact. D'après la proposition 1, il existe donc :

- une application $\varphi : \mathbb{N} \to \mathbb{N}$ strictement croissante,
- $\lambda_1^{(\infty)}, \ldots, \lambda_{n+1}^{(\infty)} \in \mathbb{R}_+$,
- et $a_1^{(\infty)}, \ldots, a_{n+1}^{(\infty)} \in A$,

tels que, pour tout $p \in [\![1, n+1]\!]$, on ait,

$$\begin{cases} \lim\limits_{k\to+\infty} a_p^{(\varphi(k))} = a_p^{(\infty)}, \\ \lim\limits_{k\to+\infty} \lambda_p^{(\varphi(k))} = \lambda_p^{(\infty)}. \end{cases}$$

Par suite,

$$1 = \sum_{p=1}^{n+1} \lambda_p^{(\varphi(k))} \underset{k\to+\infty}{\longrightarrow} \sum_{p=1}^{n+1} \lambda_p^{(\infty)}$$

et,

$$x^{(\varphi(k))} = \sum_{p=1}^{n+1} \lambda_p^{(\varphi(k))} a_p^{(\varphi(k))} \underset{k\to+\infty}{\longrightarrow} \sum_{p=1}^{n+1} \lambda_p^{(\infty)} a_p^{(\infty)}.$$

Ainsi, $(x^{(k)})_{k\in\mathbb{N}}$ converge vers $\displaystyle\sum_{p=1}^{n+1} \lambda_p^{(\infty)} a_p^{(\infty)}$ qui appartient bien à $\mathrm{Conv}(A)$ car $\lambda_1^{(\infty)}, \ldots, \lambda_{n+1}^{(\infty)} \in \mathbb{R}_+$ et $\displaystyle\sum_{p=1}^{n+1} \lambda_p^{(\infty)} = 1$.

Autrement dit, toute suite d'éléments de $\mathrm{Conv}(A)$ admet au moins une valeur d'adhérence dans $\mathrm{Conv}(A)$, c'est-à-dire que $\mathrm{Conv}(A)$ est compact. □

Théorème 3. L'ensemble $O_n(\mathbb{R})$ est une partie compacte de $M_n(\mathbb{R})$.

Preuve. Puisque $M_n(\mathbb{R})$ est un espace vectoriel de dimension finie, il suffit de montrer que $O_n(\mathbb{R})$ est une partie fermée et bornée de $M_n(\mathbb{R})$.

- Montrons que $O_n(\mathbb{R})$ est borné. Puisque $M_n(\mathbb{R})$ est de dimension finie, cette propriété ne dépend pas du choix de la norme. On choisit la norme $\|\cdot\|$ définie pour tout $M \in M_n(\mathbb{R})$ par $\|M\| = \mathrm{Tr}({}^tMM)$. Alors,

  $$\forall M \in O_n(\mathbb{R}),\ \|M\| = \mathrm{Tr}({}^tMM) = \mathrm{Tr}(I_n) = n.$$

  Donc, $O_n(\mathbb{R})$ est une partie bornée de $(M_n(\mathbb{R}), \|\cdot\|)$. Or, toutes les normes sur $M_n(\mathbb{R})$ sont équivalentes (car $M_n(\mathbb{R})$ est de dimension finie), et donc $O_n(\mathbb{R})$ est une partie bornée de $M_n(\mathbb{R})$ quelque soit la norme choisie.

- Montrons que $O_n(\mathbb{R})$ est fermé. Considérons l'application,

  $$\begin{array}{rccc} f & : & M_n(\mathbb{R}) & \longrightarrow & M_n(\mathbb{R}) \\ & & M & \longmapsto & {}^tMM \end{array}$$

  Alors, f est continue en tant que composée de l'application,

  $$\begin{array}{ccc} M_n(\mathbb{R}) & \longrightarrow & M_n(\mathbb{R}) \times M_n(\mathbb{R}) \\ M & \longmapsto & ({}^tM, M) \end{array}$$

qui est continue (car linéaire en dimension finie) par l'application,

$$\begin{array}{ccc} M_n(\mathbb{R}) & \longrightarrow & M_n(\mathbb{R}) \times M_n(\mathbb{R}) \\ (A, B) & \longmapsto & AB \end{array}$$

qui est continue (car bilinéaire en dimension finie). Par suite, $O_n(\mathbb{R})$ est fermé en tant qu'image réciproque de $\{I_n\}$ (qui est fermé) par f qui est continue.

Ainsi, $O_n(\mathbb{R})$ est une partie fermée et bornée de $M_n(\mathbb{R})$. Puisque $M_n(\mathbb{R})$ est de dimension finie, il en résulte que $O_n(\mathbb{R})$ est une partie compacte de $M_n(\mathbb{R})$. □

Corollaire 4. L'ensemble $\mathrm{Conv}(O_n(\mathbb{R}))$ est une partie compacte de $M_n(\mathbb{R})$.

Preuve. D'après le théorème 3, $O_n(\mathbb{R})$ est une partie compacte de $M_n(\mathbb{R})$. Sachant que $M_n(\mathbb{R})$ est un espace vectoriel de dimension finie, la proposition 2 assure que $\mathrm{Conv}(O_n(\mathbb{R}))$ est également une partie compacte de $M_n(\mathbb{R})$. □

Théorème de convergence dominée à paramètre

Théorème 1 (Théorème de convergence dominée à paramètre). Soient T une partie de $\mathbb{R}$, $(X, \mathcal{A}, m)$ un espace mesuré, $f : T \times X \to \mathbb{C}$ une application et $a \in \overline{T}$. On suppose que,

(i) pour tout $t \in T$, la fonction $f(t, \cdot)$ est mesurable.

(ii) pour tout $x \in X$, la fonction $f(\cdot, x)$ admet une limite en a, que l'on note $f_a(x)$.

(iii) il existe une application $g : X \to \mathbb{C}$ intégrable sur X telle que,

$$\forall t \in T, \forall x \in X, \quad |f(t,x)| \leq g(x).$$

Alors, f_a est intégrable sur X et,

$$\lim_{t \to a} \int_X f(t,x) dm(x) = \int_X f_a(x) dm(x).$$

Preuve. Il suffit d'utiliser la caractérisation séquentielle de la limite, puis d'appliquer le théorème de convergence dominée. □

Théorème de représentation de Riesz (en dimension finie)

En dimension finie, le théorème de Riesz possède une démonstration bien plus élémentaire qu'il est bon de connaître. C'est l'objet de cette sous-section.

Théorème 1. Soit E un espace euclidien. Pour tout $f \in E^*$, il existe un unique $x \in E$ tel que, pour tout $y \in E$, $f(y) = (x \mid y)$.

Preuve. Pour tout $x \in E$, notons,

$$\begin{array}{rccc} \phi_x & : & E & \longrightarrow & \mathbb{R} \\ & & y & \longmapsto & (x \mid y) \end{array}$$

Alors, l'application,

$$\begin{array}{rccc} \Phi & : & E & \longrightarrow & E^* \\ & & x & \longmapsto & \phi_x \end{array}$$

est linéaire et injective, car $\mathrm{Ker}(\varphi) = E^\perp = \{0\}$. Or, $\dim(E^*) = \dim(E)$. Donc Φ est un isomorphisme. En particulier, Φ est bijective, ce qui achève la preuve. □

Remarque. C'est un cas particulier du théorème de Riesz.

Théorème de représentation de Riesz dans un espace de Hilbert

[GH] **Corollaire 1.** Si F est un sous-espace vectoriel fermé de H alors,

$$H = F \oplus F^\perp.$$

Preuve. Tout d'abord, $F \cap F^\perp = \{0\}$ (cette égalité est vraie pour un espace vectoriel normé quelconque E et un sous-espace vectoriel F quelconque de E). En effet, si $x \in F \cap F^\perp$ alors $(x \mid x) = 0$ ce qui implique $x = 0$. Montrons ensuite que $H = F + F^\perp$. Soit $x \in H$. Notons $P_F(x)$ le projeté orthogonal de x sur F. Soit $y \in F$. D'après la caractérisation de $P_F(x)$,

$$(x - P_F(x) \mid y) = (x - P_F(x) \mid P_F(x) + y - P_F(x)) \leq 0$$

car $P_F(x) + y \in F$. D'où, $(x - P_F(x) \mid y) \leq 0$. De même,

$$-(x - P_F(x) \mid y) = (x - P_F(x) \mid P_F(x) - y - P_F(x)) \leq 0$$

car $P_F(x) - y \in F$. D'où, $(x - P_F(x) \mid y) \geq 0$. Par double inégalité, il vient,

$$(x - P_F(x) \mid y) = 0.$$

Ceci valant pour tout $y \in F$, il s'en suit que $x - P_F(x) \in F^\perp$. Ainsi,

$$x = \underbrace{P_F(x)}_{\in F} + \underbrace{x - P_F(x)}_{\in F^\perp}$$

ce qui prouve que $x \in F + F^\perp$. Ainsi, $H = F \oplus F^\perp$. □

Remarque. Dans la preuve ci-dessus, on a montré que $x - P_F(x) \in F^\perp$. On peut en fait montrer que cette propriété caractérise $P_F(x)$ (cf. proposition 6.51 de [GH]).

Notons H' l'ensemble des formes linéaires continues sur H. On munit H' de la norme subordonnée aux normes $\|\cdot\|$ et $|\cdot|$, c'est-à-dire la norme $\|\cdot\|$ définie pour tout $\phi \in H'$ par,

$$\|\phi\| = \sup_{\substack{x \in H \\ x \neq 0}} \frac{|f(x)|}{\|x\|}.$$

Pour tout $a \in H$, notons,

$$\begin{array}{rccc} \phi_a & : & H & \longrightarrow & \mathbb{R} \\ & & x & \longmapsto & (x \mid a) \end{array}$$

Remarquons que, pour tout $a \in H$, $\phi_a \in H'$ car, d'après l'inégalité de Cauchy-Schwarz,

$$\forall x \in H,\ |(a \mid x)| \leq \|a\|\|x\|$$

d'où $\phi_a \in H'$ avec $\|\phi_a\| \leq \|a\|$ (on a même $\|\phi_a\| = \|a\|$ puisque $\phi_a(a) = \|a\|^2$).

[GH] **Théorème 2 (Représentation de Riesz).** Pour tout $\phi \in H'$, il existe un unique vecteur a de H tel que,

$$\forall x \in H,\ \phi(x) = (x \mid a).$$

Ainsi, l'application,

$$\begin{array}{rccc} \Phi & : & H & \longrightarrow & H' \\ & & a & \longmapsto & \phi_a \end{array}$$

est un isomorphisme. De plus, c'est une isométrie.

Preuve. Si $\phi = 0$ alors $a = 0$ est bien le seul vecteur qui convient. Supposons $\phi \neq 0$, et notons F son noyau. Alors, F est un sous-espace vectoriel fermé de E (en tant qu'image réciproque de $\{0\}$ qui est fermé par ϕ qui est continue). Par ailleurs, $F^\perp \neq \{0\}$ (car $F \neq H$ et $H = F \oplus F^\perp$), ce qui permet de considérer $x_0 \in F^\perp \setminus \{0\}$. En particulier, $x_0 \notin F$, c'est-à-dire $\phi(x_0) \neq 0$. Montrons que,

$$x = \underbrace{x - \frac{\phi(x)}{\phi(x_0)} x_0}_{\in F} + \underbrace{\frac{\phi(x)}{\phi(x_0)} x_0}_{\in F^\perp}.$$

Tout d'abord, l'égalité est immédiate. Ensuite, la linéarité de ϕ donne,

$$\phi\left(x - \frac{\phi(x)}{\phi(x_0)} x_0\right) = \phi(x) - \frac{\phi(x)}{\phi(x_0)} \phi(x_0) = 0$$

et donc $x - \dfrac{\phi(x)}{\phi(x_0)} x_0 \in F$. Enfin, $\dfrac{\phi(x)}{\phi(x_0)} x_0$ car $x_0 \in F^\perp$. D'où,

$$x = \underbrace{x - \frac{\phi(x)}{\phi(x_0)} x_0}_{\in F} + \underbrace{\frac{\phi(x)}{\phi(x_0)} x_0}_{\in F^\perp}.$$

En particulier,

$$(x \mid x_0) = \frac{\phi(x)}{\phi(x_0)} \|x_0\|$$

c'est-à-dire,

$$\phi(x) = \frac{\phi(x_0)}{\|x_0\|} (x \mid x_0) = \left(x \;\middle|\; \frac{\phi(x_0)}{\|x_0\|} x_0\right).$$

Ainsi, en notant $a = \dfrac{\phi(x_0)}{\|x_0\|} x_0$, il vient $\phi = \phi_a$. De plus, a est unique, car si $b \in H$ vérifie, pour tout $x \in H$, $\phi(x) = (x \mid b)$, alors,

$$(b-a \mid b-a) = (b-a \mid b) - (b-a \mid a) = \phi(b-a) - \phi(b-a) = 0$$

et donc $a = b$. Ainsi, Φ est bijective. Sachant que Φ est linéaire, on en déduit que Φ est un isomorphisme. De plus, Φ est une isométrie puisqu'on a montré précédemment que, pour tout $a \in H$, $\|\phi_a\| = \|a\|$. □

Théorème de sortie de tout compact

Lemme 1. Soit $g :]\alpha, \beta[\to \mathbb{R}^m$. Les assertions suivantes sont équivalentes :

(i) g admet une limite en β,

(ii) pour toute suite $(t_n)_{n\in\mathbb{N}}$ à termes dans $]\alpha, \beta[$ qui converge vers β, la suite $(g(t_n))_{n\in\mathbb{N}}$ converge.

Preuve. L'implication (i) $\Rightarrow$ (ii) est immédiate. Montrons que (ii) $\Rightarrow$ (i). Soient $(t_n)_{n\in\mathbb{N}}$ et $(s_n)_{n\in\mathbb{N}}$ deux suites à termes dans $]\alpha, \beta[$ qui convergent vers β. Par hypothèse, les suites $(g(t_n))_{n\in\mathbb{N}}$ et $(g(s_n))_{n\in\mathbb{N}}$ convergent respectivement vers $l_1, l_2 \in \mathbb{R}^m$. Montrons que $l_1 = l_2$. Pour tout $n \in \mathbb{N}$, posons,

$$r_n = \begin{cases} t_n & \text{si } n \in 2\mathbb{N}, \\ s_n & \text{si } n \in 2\mathbb{N}+1. \end{cases}$$

Alors, $(r_n)_{n\in\mathbb{N}}$ converge vers β. Par hypothèse, $(g(r_n))_{n\in\mathbb{N}}$ converge donc vers $l \in \mathbb{R}^m$. Or, puisque $(g(r_n))_{n\in\mathbb{N}}$ est une suite extraite de $(g(t_n))_{n\in\mathbb{N}}$, on en déduit que $l = l_1$. De même, $l = l_2$. D'où $l_1 = l_2$. Ainsi, pour toute suite à termes dans $]\alpha, \beta[$ qui converge vers β, la suite de ses images par g converge vers l. Montrons désormais que g admet une limite en β qui vaut l. Sinon,

$$\exists \varepsilon_0 > 0,\ \forall \delta > 0,\ \exists t_\delta \in]\beta - \delta, \beta[,\ \|g(t_\delta) - l\| > \varepsilon_0.$$

En prenant $\delta_n = \frac{1}{n+1}$, on construit alors une suite $(u_n)_{n\in\mathbb{N}}$ à termes dans $]\alpha, \beta[$ qui converge vers β, et donc $(g(u_n))_{n\in\mathbb{N}}$ converge vers l d'après ce qui précède. Mais ceci est impossible puisque, pour tout $n \in \mathbb{N}$, $\|g(u_n) - l\| > \varepsilon_0$. Ainsi, g admet une limite en β qui vaut l. □

[JD] **Théorème 2 (Théorème de sortie de tout compact).** Soit $(]a, b[, y)$ une solution de (E) telle que $b < +\infty$. S'il existe $\varepsilon > 0$ et K un compact de $I \times \Omega$ tels que l'ensemble,

$$\{(t, y(t)),\ t \in]b - \varepsilon, b[\} \subset K$$

alors il existe $c > b$ tel que y se prolonge en une solution de (E) sur $]a, c[$.

Preuve. Montrons tout d'abord que y possède une limite l en b et que $(b, l) \in K$. La fonction f étant continue sur K qui est compact, on peut poser,

$$L = \sup_{(t,x)\in K} \|F(t, x)\|.$$

Fixons $(t_n)_{n\in\mathbb{N}}$ une suite à termes dans $]b - \varepsilon, b[$ qui converge vers b. D'après l'inégalité des accroissements finis, pour tous $t, s \in]b - \varepsilon, b[$,

$$\|y(t) - y(s)\| \leq L|t - s|.$$

En particulier, pour tous $p, q \in \mathbb{N}$,

$$\|y(t_p) - y(t_q)\| \leq L|t_p - t_q|$$

ce qui implique que $(y(t_n))_{n\in\mathbb{N}}$ est une suite de Cauchy car $(t_n)_{n\in\mathbb{N}}$ est elle-même de Cauchy. Puisque la suite $(y(t_n))_{n\in\mathbb{N}}$ est à termes dans K, et que K est complet en tant que partie fermée de $\mathbb{R}^{m+1}$, on en conclut que $(y(t_n))_{n\in\mathbb{N}}$ converge vers un certain $y_b \in K$. D'après le premier lemme,

$$\lim_{t\to b} y(t) = y_b.$$

De plus, $(b, y_b) \in K \subset I \times \Omega$. D'après le théorème de Cauchy-Lipschitz (local), il existe donc $]c, d[\subset I$ tel que $(]c, d[, z)$ soit solution du problème de Cauchy de (E) en (b, y_b). En particulier, $b \in]c, d[$. Notons ensuite,

$$\begin{array}{rccc} \varphi & : &]a, d[& \longrightarrow & \mathbb{R}^m \\ & & t & \longmapsto & \begin{cases} y(t) & \text{si } t \in]a, b[, \\ z(t) & \text{si } t \in [b, d[. \end{cases} \end{array}$$

Montrons désormais que φ est une solution de (E) qui prolonge strictement y. Tout d'abord, φ est dérivable sur $]a, b[\cup]b, d[$ avec, pour tout $t \in]a, b[\cup]b, d[$,

$$\varphi'(t) = F(t, \varphi(t)).$$

De plus, par continuité de F au point (b, y_b),

$$\lim_{\substack{t\to b \\ t<b}} \varphi'(t) = \lim_{\substack{t\to b \\ t<b}} F(t, y(t)) = F(b, y_b) = \lim_{\substack{t\to b \\ t>b}} F(t, z(t)) = \lim_{\substack{t\to b \\ t>b}} \varphi'(t)$$

car $\lim_{t\to b} y(t) = y_b$. D'après le théorème de la limite de la dérivée (théorème 2), on en conclut que φ est solution de (E). De plus, φ est un prolongement strict de y car $]a, d[$ contient strictement $]a, d[$. □

Théorème du point fixe de Banach-Picard

Théorème 1 (Théorème du point fixe de Banach-Picard). Soit (E, d) un espace métrique complet. Soit $f : E \to E$ une application contractante, c'est-à-dire qu'il existe $k \in]0, 1[$ tel que, pour tous $x, y \in E$,

$$d(f(x), f(y)) \leq kd(x, y).$$

Alors, f admet un unique point fixe.

Preuve. Soit $x_0 \in E$. Considérons la suite $(x_n)_{n\in\mathbb{N}}$ définie pour tout $n \in \mathbb{N}$ par $x_{n+1} = f(x_n)$. Soit $p \in \mathbb{N}$. Montrons par récurrence sur $n \in \mathbb{N}$ que,

$$d(x_{n+p}, x_n) \leq k^n d(x_p, x_0).$$

Pour $n = 0$, il s'agit d'une égalité. Supposons ensuite le résultat établi pour un rang $n \in \mathbb{N}$ fixé. Alors, par hypothèse sur f,

$$d(x_{n+p+1}, x_{n+1}) = d(f(x_{n+p}), f(x_n)) \leq kd(x_{n+p}, x_n)$$

et donc, par hypothèse de récurrence,

$$d(x_{n+p+1}, x_{n+1}) \leq k^{n+1} d(x_p, x_0).$$

Ainsi, pour tous $n, p \in \mathbb{N}$,

$$d(x_{n+p}, x_n) \leq k^n d(x_p, x_0).$$

Puisque $k < 1$, cette dernière inégalité entraîne que la suite $(x_n)_{n\in\mathbb{N}}$ est de Cauchy. La complétude de (E, d) assure alors l'existence de $x_\infty \in E$ tel que $(x_n)_{n\in\mathbb{N}}$ converge vers x_∞ dans (E, d). Puisque f est contractante, f est continue en x_∞, et donc,

$$x_{n+1} = f(x_n) \underset{n\to+\infty}{\longrightarrow} f(x_\infty).$$

Par unicité de la limite, on en conclut que $f(x_\infty) = x_\infty$. Donc f admet un point fixe. Il reste à montrer son unicité. Soient $a, b \in E$ tels que $f(a) = a$ et $f(b) = b$. Alors,

$$d(a, b) = d(f(a), f(b)) < kd(a, b)$$

avec $k \in]0, 1[$, ce qui n'est possible que si $d(a, b) = 0$, c'est-à-dire $a = b$. Ainsi, f admet un unique point fixe. □

Trace et base orthonormée

Soit $n \in \mathbb{N}^*$. Notons $\langle \cdot \mid \cdot \rangle_2$ le produit scalaire sur $M_{n,1}(\mathbb{R})$ défini par,

$$\forall X, Y \in M_{n,1}(\mathbb{R}),\ \langle X \mid Y \rangle_2 = {}^tXY.$$

Notons $\|\cdot\|_2$ la norme induite par ce produit scalaire. Rappelons à toutes fins utiles que, pour tout $M \in M_n(\mathbb{R})$, et pour tous $X, Y \in M_{n,1}(\mathbb{R})$,

$$\langle AX \mid Y \rangle_2 = \langle X \mid {}^tAY \rangle_2$$

car $\langle AX \mid Y \rangle_2 = {}^t(AX)Y = {}^tX{}^tAY = \langle X \mid {}^tAY \rangle_2$.

Proposition 1. Soient $A \in M_n(\mathbb{R})$ et $(X_1, \dots, X_n)$ une base orthonormée de $(M_{n,1}(\mathbb{R}), \|\cdot\|_2)$. Alors,

$$\mathrm{Tr}(A) = \sum_{i=1}^{n} \langle AX_i \mid X_i \rangle_2.$$

Preuve. Commençons par prouver ce résultat pour la base canonique. Notons $a_{i,j}$ le terme général de A et $(C_1, \dots, C_n)$ la base canonique. Rappelons que, pour tout $M \in M_n(\mathbb{R})$ et pour tout $i \in [\![1, n]\!]$, MC_i est égal à la première colonne de M et ${}^tC_iM$ est égal à la première ligne de M. Par conséquent,

$$\forall i \in [\![1, n]\!],\ {}^tC_iAC_i = a_{i,j}$$

et donc,

$$\mathrm{Tr}(A) = \sum_{i=1}^{n} a_{i,i} = \sum_{i=1}^{n} \langle AC_i \mid C_i \rangle_2.$$

Ainsi,

$$\forall M \in M_n(\mathbb{R}),\ \mathrm{Tr}(M) = \sum_{i=1}^{n} \langle MC_i \mid C_i \rangle_2. \tag{1}$$

Notons ensuite P la matrice de passage de $(C_1, \dots, C_n)$ à $(X_1, \dots, X_n)$. Alors, $P \in O_n(\mathbb{R})$. De plus, pour tout $i \in [\![1, n]\!]$, $PX_i = C_i$ et ${}^tPC_i = X_i$. Par suite,

$$\begin{aligned}
\sum_{i=1}^{n} \langle AX_i \mid X_i \rangle_2 &= \sum_{i=1}^{n} \langle A{}^tPC_i \mid {}^tPC_i \rangle_2 \\
&= \sum_{i=1}^{n} \langle PA{}^tPC_i \mid C_i \rangle_2 \\
&= \mathrm{Tr}(PA{}^tP) \\
&= \mathrm{Tr}(PAP^{-1}) \\
\sum_{i=1}^{n} \langle AX_i \mid X_i \rangle_2 &= \mathrm{Tr}(A)
\end{aligned}$$

où la deuxième égalité résulte du rappel préliminaire et la troisième égalité de l'égalité (1). Ainsi,

$$\mathrm{Tr}(A) = \sum_{i=1}^{n} \langle AX_i \mid X_i \rangle_2.$$

□

Un lemme d'arithmétique bien utile

Lemme 1. Soit $\mathbb{K}$ un corps. Soient $Q_1, \dots, Q_n \in \mathbb{K}[X]$ des polynômes deux à deux premiers entre eux et $P \in \mathbb{K}[X]$. On suppose que, pour tout $k \in [\![1, n]\!]$, Q_k divise P. Alors, leur produit $Q_1 \cdots Q_n$ divise P.

Preuve. Montrons, par récurrence sur $n \in \mathbb{N}^*$, que si $Q_1, \dots, Q_n \in \mathbb{K}[X]$ sont des polynômes deux à deux premiers entre eux divisant chacun P, alors leur produit divise aussi P.

Initialisation. Soient A et B deux polynômes premiers entre eux divisant P. Alors, AB est un diviseur commun de PB et PA, et donc,

$$AB \mid \operatorname{pgcd}(PB, PA) = P \operatorname{pgcd}(A, B) = P.$$

Hérédité. Supposons maintenant le résultat établi pour un rang $n \geq 2$ fixé. Soient $Q_1, \dots, Q_{n+1}$ des polynômes deux à deux premiers entre eux divisant chacun P. Alors, les polynômes Q_1 et $Q_2 \cdots Q_{n+1}$ sont premiers entre eux. En effet, si D est un diviseur irréductible de Q_1 et $Q_2 \cdots Q_{n+1}$, d'après le lemme d'Euclide, il existe $i \in [\![2, n+1]\!]$ tel que D divise Q_i, ce qui contredit le fait que Q_1 et Q_i soient premiers entre eux. Ainsi, Q_1 et $Q_2 \cdots Q_{n+1}$ sont premiers entre eux. Or, par hypothèse de récurrence,

$$Q_2 \cdots Q_{n+1} \mid P.$$

On peut donc appliquer le résultat de l'initialisation aux polynômes $A = Q_1$ et $B = Q_2 \cdots Q_{n+1}$, ce qui donne finalement,

$$Q_1 \cdots Q_{n+1} \mid P.$$

D'où l'hérédité. Par récurrence, il s'en suit le résultat souhaité. □

Variables aléatoires mutuellement indépendantes et identiquement distribuées

Définition 1. Soit $(\mathcal{A}_i)_{i\in I}$ une famille de tribus de Ω incluses dans $\mathcal{A}$. On dit que la famille $(\mathcal{A}_i)_{i\in I}$ est indépendante si, pour tout $(A_i)_{i\in I} \in \prod_{i\in I} \mathcal{A}_i$, la famille $(A_i)_{i\in I}$ est indépendante.

Proposition 2. Soit $X : (\Omega, \mathcal{A}) \to (E, \mathcal{T})$ une variable aléatoire. On note,

$$\sigma(X) = \{X^{-1}(A),\ A \in \mathcal{T}\}.$$

Alors, $\sigma(X)$ est une tribu sur Ω. On dit que c'est la tribu engendrée par une variable aléatoire X.

Définition 3. Soit $(X_i : (\Omega, \mathcal{A}) \to (E_i, \mathcal{T}_i))_{i\in I}$ une famille de variables aléatoires. On dit que la famille $(X_i)_{i\in I}$ est indépendante (ou que ces variables aléatoires sont mutuellement indépendantes) si la famille de tribus $(\sigma(X_i))_{i\in I}$ est indépendante, c'est-à-dire si, pour toute partie finie J de I,

$$\forall (A_j)_{j\in J} \in \prod_{j\in J} \mathcal{T}_j,\ \mathbb{P}\left(\bigcap_{j\in J}(X_j \in A_j)\right) = \prod_{j\in J} \mathbb{P}(X_j \in A_j).$$

Définition 4. Des variables aléatoires sont dites i.i.d (indépendantes et identiquement distribuées) lorsqu'elles sont mutuellement indépendantes et qu'elles suivent toutes la même loi de probabilité.

Je remercie Pierre Mathieu, enseignant à l'université d'Aix-Marseille, pour ces éclaircissements concernant le résultat ci-dessous.

Lemme 5. Soit $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé. Il existe une suite $(X_n)_{n\in\mathbb{N}^*}$ de variables aléatoires telle que,

(i) les variables aléatoires de $(X_n)_{n\in\mathbb{N}^*}$ sont mutuellement indépendantes;

(ii) pour tout $n \in \mathbb{N}^*$, $X_n \hookrightarrow \mathbb{P}$.

Autrement dit, $(X_n)_{n\in\mathbb{N}^*}$ est une suite de variables aléatoires i.i.d (indépendantes et identiquement distribuées) suivant la loi $\mathbb{P}$.

Preuve. Pour tout $n \in \mathbb{N}^*$, notons $\Omega_n = \Omega$, $\mathcal{A}_n = \mathcal{A}$ et $\mathbb{P}_n = \mathbb{P}$. Notons,

$$E = \prod_{n\in\mathbb{N}^*} \Omega_n,\ \mathcal{B} = \bigotimes_{n\in\mathbb{N}^*} \mathcal{A}_n,\ \mu = \bigotimes_{n\in\mathbb{N}^*} \mathbb{P}_n$$

Pour tout $n \in \mathbb{N}^*$, notons X_n la projection de E sur Ω_n.

- Pour tout $n \in \mathbb{N}^*$, X_n est une variable aléatoire de $(E, \mathcal{B})$ vers $(\Omega, \mathcal{A})$.
- Soit $n \in \mathbb{N}^*$. Montrons que $X_n \hookrightarrow \mathbb{P}$. Pour tout $A \in \mathcal{A}$,

  $$X_n^{-1}(A) = \left(\prod_{k=1}^{n-1} \Omega\right) \times A \times \left(\prod_{k=n+1}^{+\infty} \Omega\right)$$

  d'où, par définition de la mesure produit,

  $$\mu(X_n \in A) = \left(\prod_{k=1}^{n-1} \mathbb{P}(\Omega)\right) \mathbb{P}(A) \left(\prod_{k=n+1}^{+\infty} \mathbb{P}(\Omega)\right) = \mathbb{P}(A).$$

  Ainsi, $\mu_{X_n} = \mathbb{P}$, c'est-à-dire $X_n \hookrightarrow \mathbb{P}$.
- Soient I une partie finie de $\mathbb{N}^*$ et $(A_i)_{i\in I} \in \prod_{i\in I} \mathcal{A}_i$. Tout d'abord,

  $$\forall i \in I,\ X_i^{-1}(A_i) = \left(\prod_{n=1}^{i-1} \Omega\right) \times A_i \times \left(\prod_{n=i+1}^{+\infty} \Omega\right).$$

  D'où,

  $$\bigcap_{i\in I}(X_i \in A_i) = \prod_{n\in\mathbb{N}^*} B_n$$

  où l'on a noté, pour tout $n \in \mathbb{N}^*$,

  $$B_n = \begin{cases} A_n & \text{si } n \in I, \\ \Omega & \text{sinon.} \end{cases}$$

  Puis, par définition de la mesure produit,

  $$\mu\left(\bigcap_{i\in I}(X_i \in A_i)\right) = \mu\left(\prod_{n\in\mathbb{N}^*} B_n\right) = \prod_{n\in\mathbb{N}^*} \mathbb{P}(B_n) = \prod_{i\in I} \mathbb{P}(A_i).$$

  Or,

  $$\forall i \in I, \mu(X_i \in A_i) = \mathbb{P}(A_i)$$

  car $X_i \hookrightarrow \mathbb{P}$. Des égalités précédentes, on en déduit que,

  $$\mu\left(\bigcap_{i\in I}(X_i \in A_i)\right) = \prod_{i\in I} \mu(X_i \in A_i),$$

  ce qui prouve finalement que les variables aléatoires de la suite $(X_n)_{n\in\mathbb{N}^*}$ sont mutuellement indépendantes.

Finalement, $(X_n)_{n\in\mathbb{N}^*}$ est bien une suite de variables aléatoires i.i.d (indépendantes et identiquement distribuées) suivant la loi $\mathbb{P}$. □

Remarque. En particulier, si $\Omega = \mathbb{N}^*$, $\mathcal{A} = \mathcal{P}(\mathbb{N}^*)$, et si $\mathbb{P}$ est la loi de Poisson de paramètre 1, on a donc construit une suite $(X_n)_{n\in\mathbb{N}^*}$ de variables aléatoires i.i.d (indépendantes et identiquement distribuées), suivant une loi de Poisson de paramètre 1. On se servira de l'existence d'une telle suite pour prouver la formule de Stirling (théorème 1).

Références

[AP] : Cours d'analyse de A. Pommellet (ellipses).

[BB] : Analyse pour l'agrégation de mathématiques de J. Bernis et L. Bernis (ellipses).

[BM] : Objectif agrégation de V. Beck, J. Malick et G. Peyré (H & K).

[FB] : Équations différentielles de F. Berthelin (Cassini).

[FG] : Exercices de mathématiques oraux x-ens, analyse 1 de S. Francinou, H. Gianella et S. Nicolas (Cassini).

[FG] : Exercices de mathématiques oraux x-ens, analyse 2 de S. Francinou, H. Gianella et S. Nicolas (Cassini).

[FG] : Exercices de mathématiques oraux x-ens, analyse 3 de S. Francinou, H. Gianella et S. Nicolas (Cassini).

[FG] : Exercices de mathématiques oraux x-ens, analyse 4 de S. Francinou, H. Gianella et S. Nicolas (Cassini).

[FR] : Petit guide de calcul différentiel à l'usage de la licence et de l'agrégation de F. Rouvière (Cassini).

[GH] : Mesure, intégration, probabilités de T. Gallouët et R. Herbin (ellipses).

[GK] : De l'intégration aux probabilités de O. Garet et A. Kurtzman (ellipses).

[IP] : L'oral à l'agrégation de mathématiques - Une sélection de développements de L. Isenmann et T. Pecatte.

[JD] : Analyse numérique et équations différentielles de J. Demailly (EDP Sciences).

[JR] : Mathématiques pour l'agrégation, algèbre et géométrie de J. Rombaldi (De Boeck Supérieur).

[JR] : Mathématiques pour l'agrégation, analyse et probabilités de J. Rombaldi (De Boeck Supérieur).

[LL] : 131 développements pour l'oral de P. Le Barbenchon, D. Lesesvre, P. Montagnon et T. Pierron (Dunod).

[XG] : Algèbre et probabilités de X. Gourdon (ellipses).

[XG] : Analyse de X. Gourdon (ellipses).

Références détaillées

Développement ou résultat utile	Références
Décomposition polaire sur $M_n(\mathbb{R})$	[JR] : Algèbre, chapitre 22, théorèmes 22.32.
Théorème de Lévy	[GH] : Chapitre 10, théorème 10.22.
Convergence en loi et fonction de répartition	[GK] : Chapitre 11, corollaire 11.10.
Fonction caractéristique (probabilités)	[GH] : Chapitre 10, proposition 10.20.
Formule d'inversion de Möbius	[JR] : Algèbre, chapitre 11, théorème 11.17.
Théorème de représentation de Riesz	[GH] : Chapitre 6, théorèmes 6.53 et 6.56.

Figures

Ne pas dessiner quand on fait des maths, c'est criminel.

Générateurs du groupe orthogonal d'un espace vectoriel euclidien

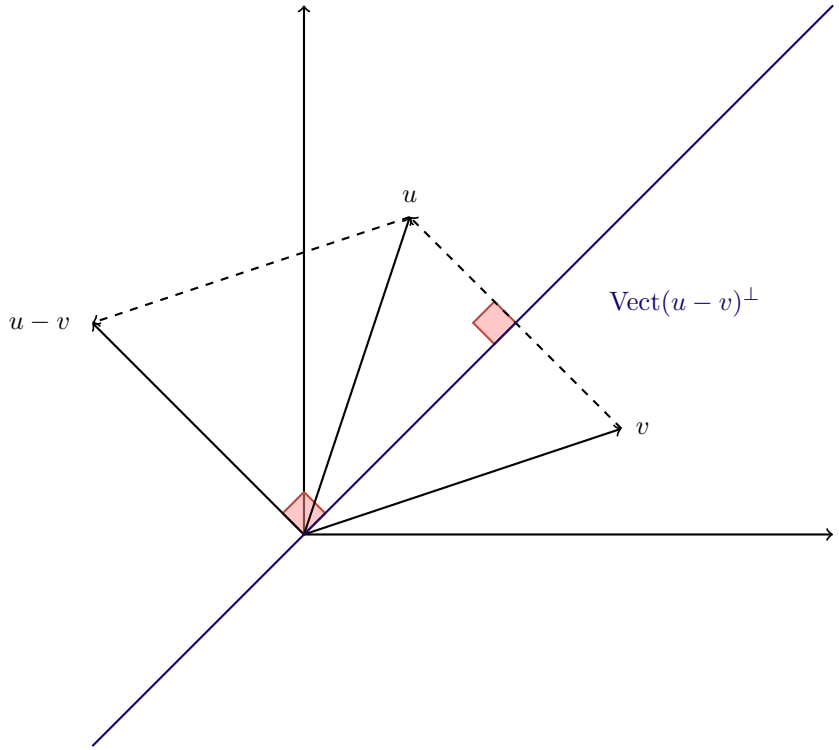

Figure 1: Réflexion envoyant u sur v en dimension 2

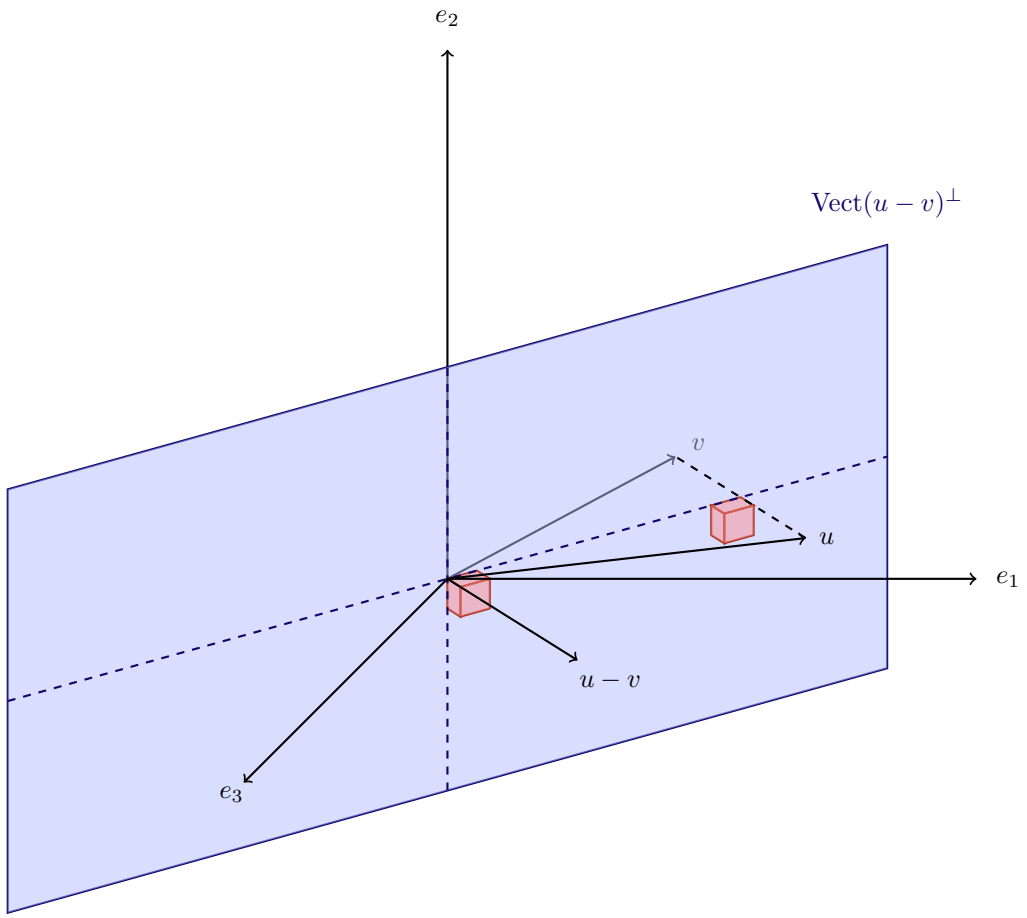

Figure 2: Réflexion envoyant u sur v en dimension 3

Groupe diédral

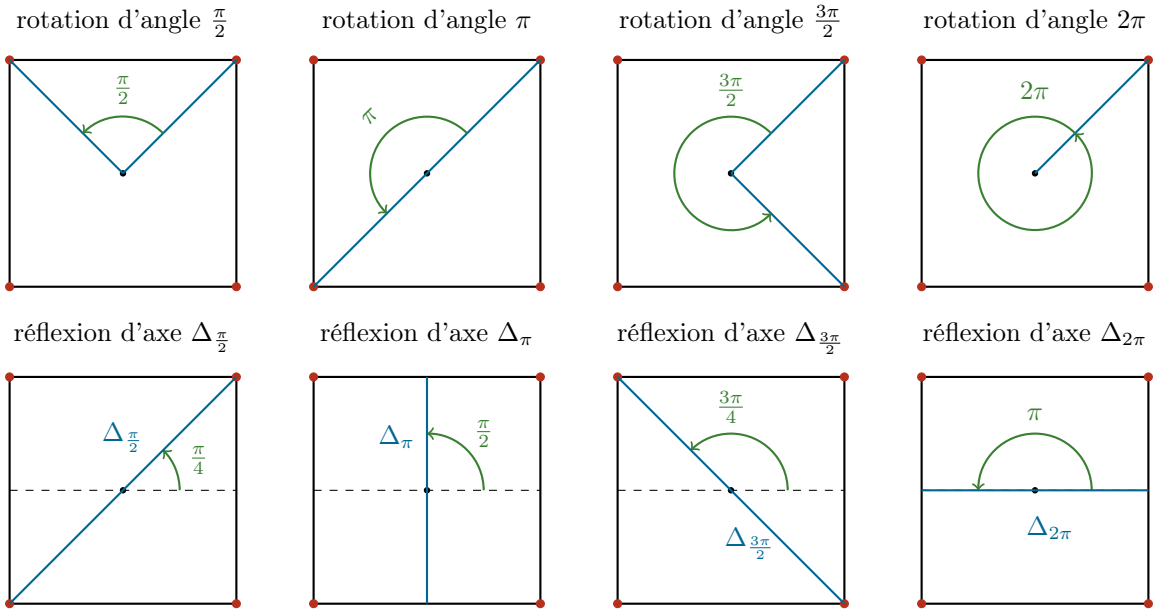

Figure 3: Groupe diédral

Composée de deux rotations

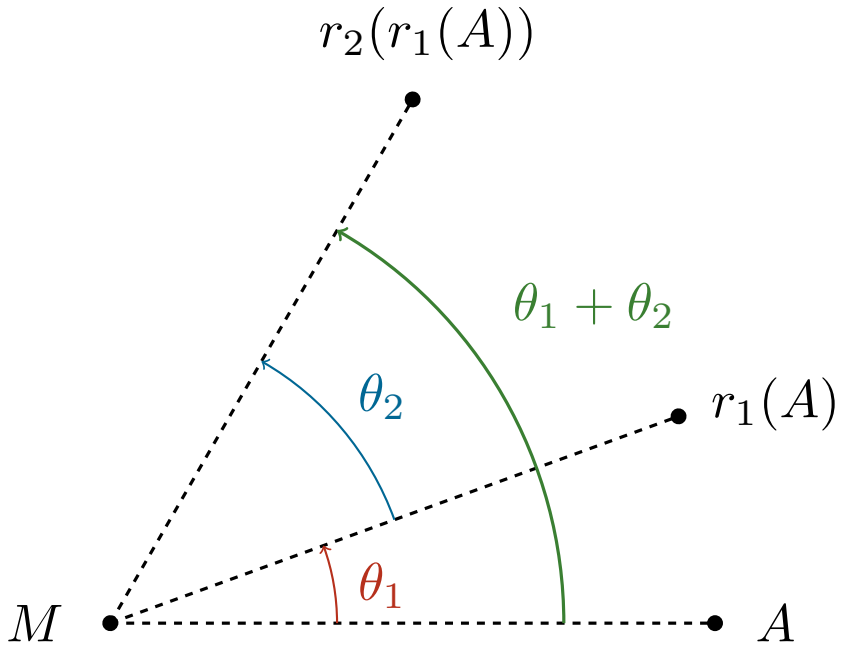

Figure 4: Composée de deux rotations

Composée de deux réflexions d'axes non parallèles

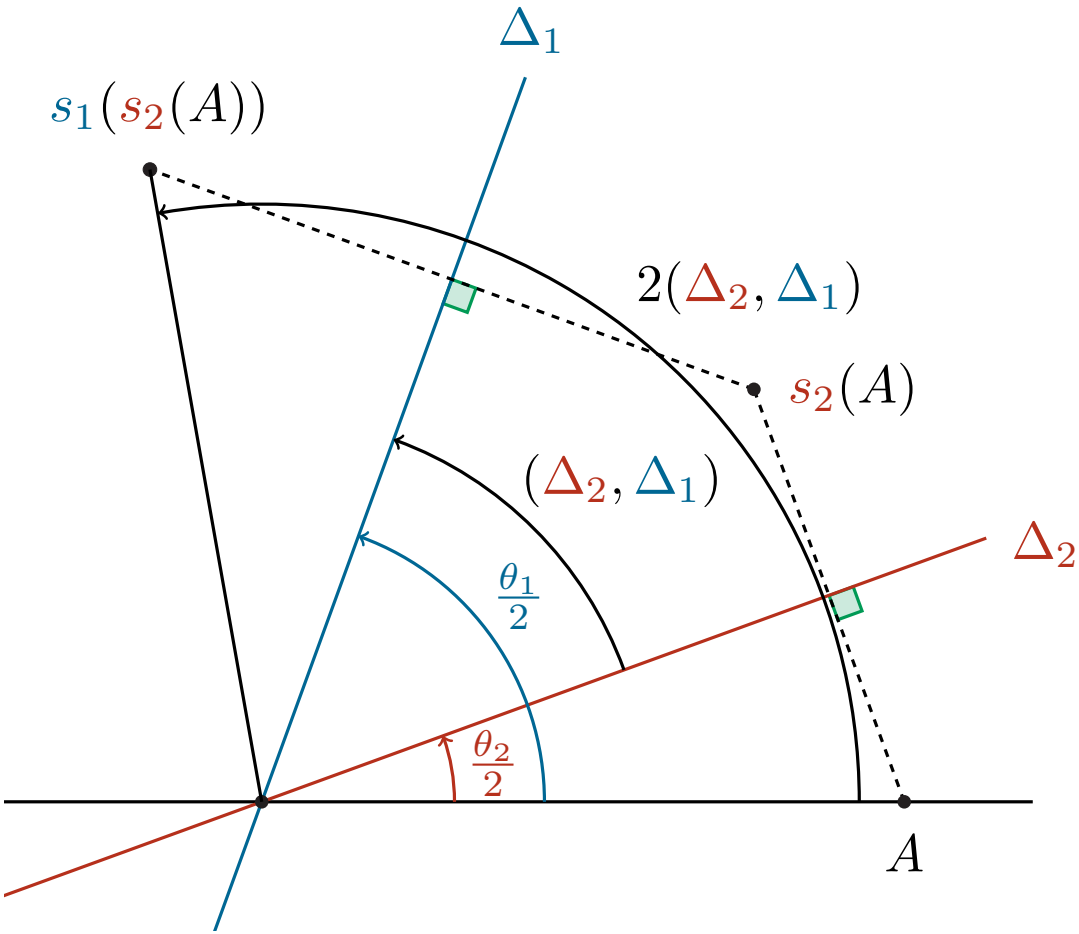

Figure 5: Composée de deux réflexions d'axes non parallèles

Conjugué d'une réflexion par une rotation

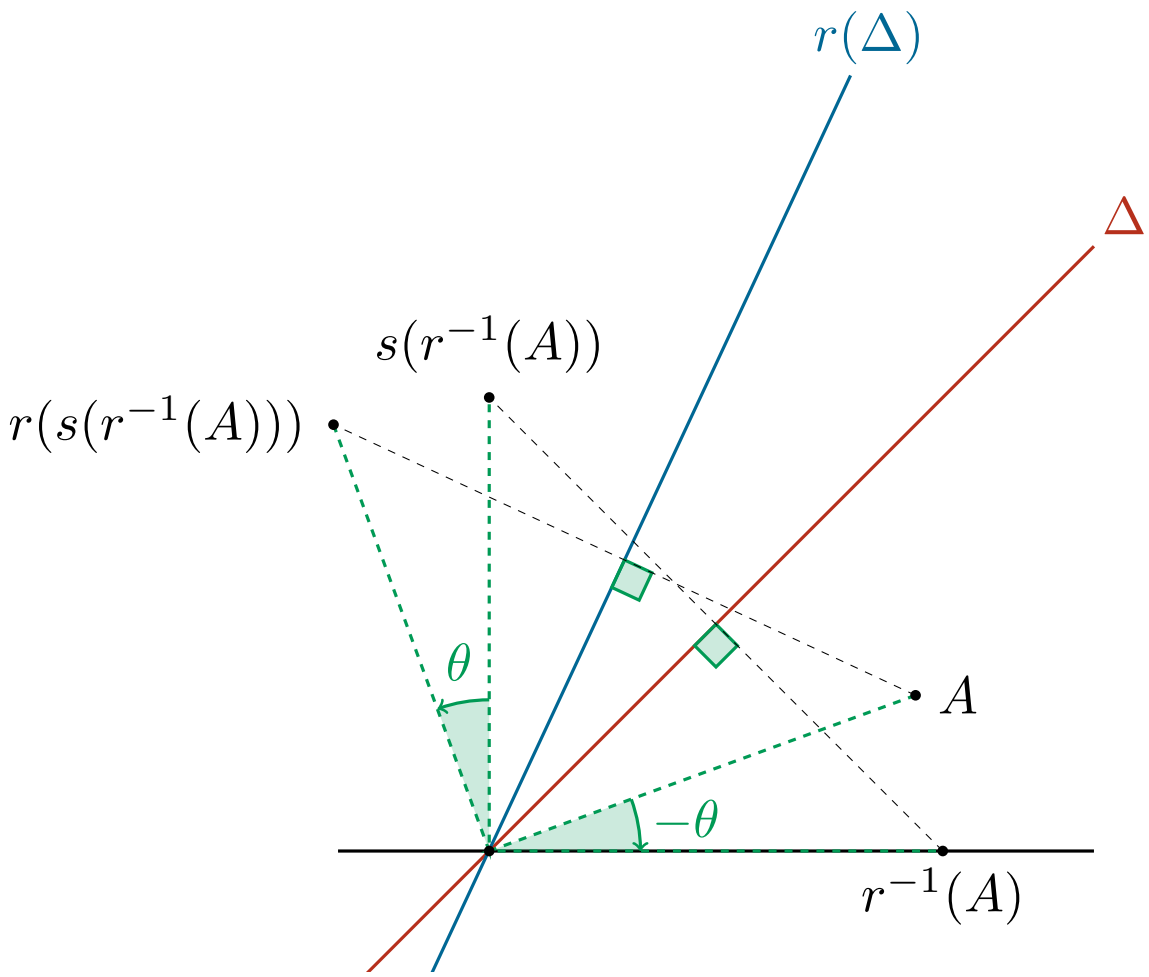

Figure 6: Conjugué d'une réflexion par une rotation

Entiers de Gauss

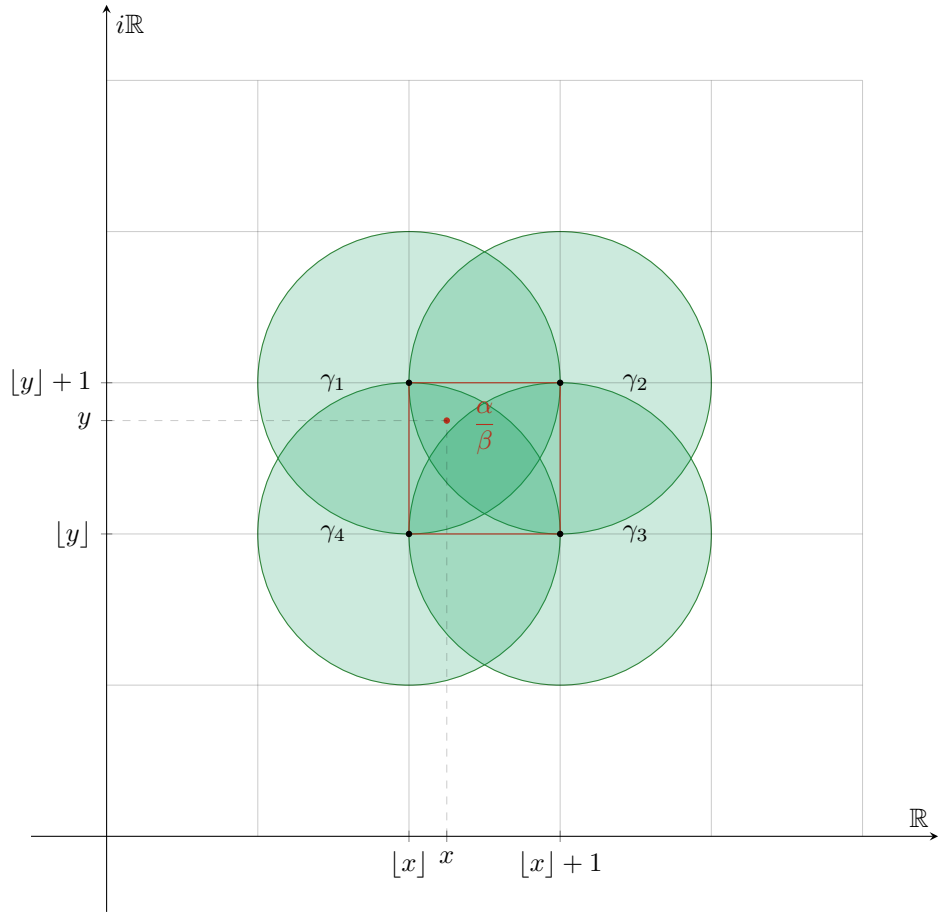

Figure 7: Entiers de Gauss

Une suite de polygones

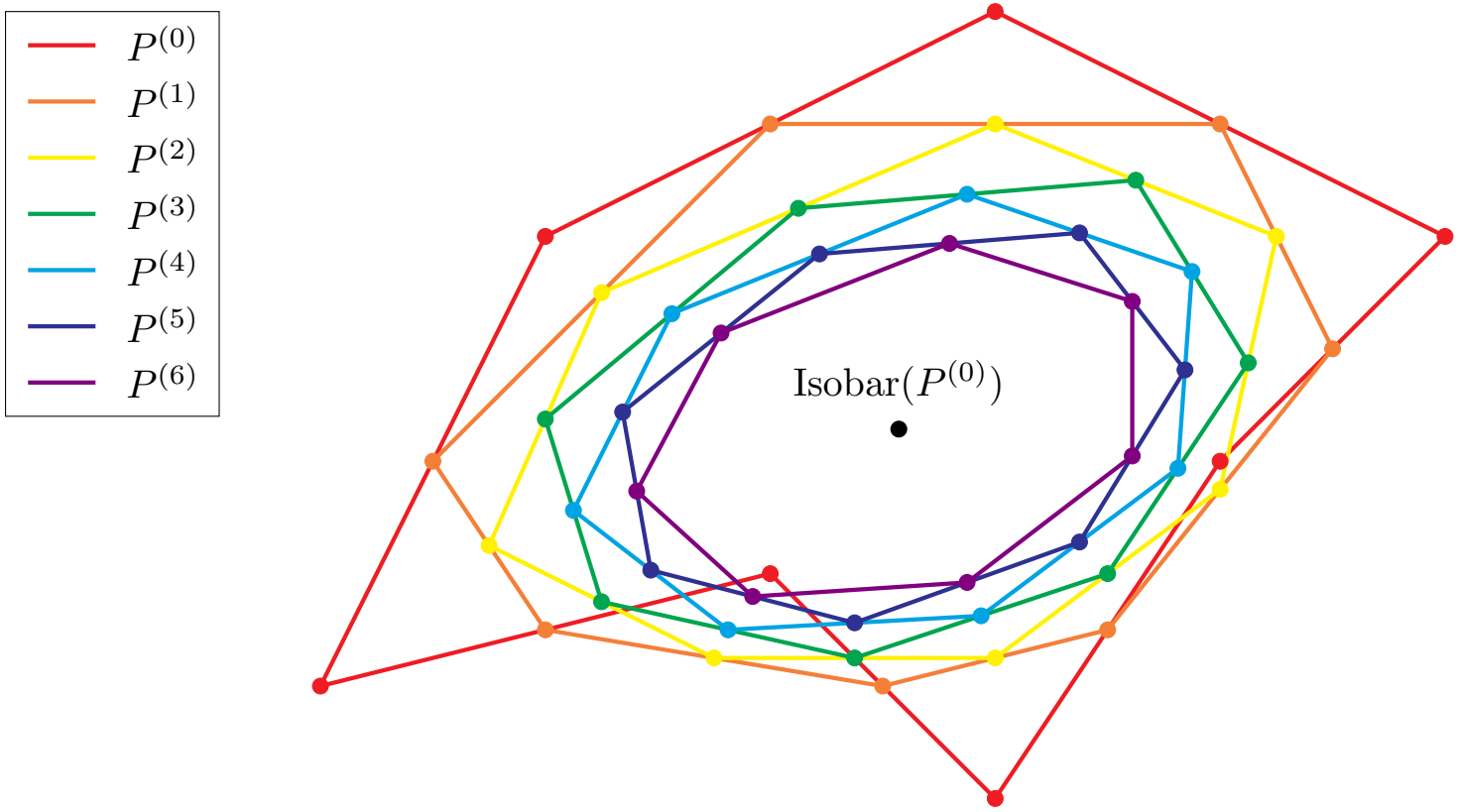

Figure 8: Convergence de la suite $(P^{(k)})_{k\in\mathbb{N}}$ vers $(g,\ldots,g)$

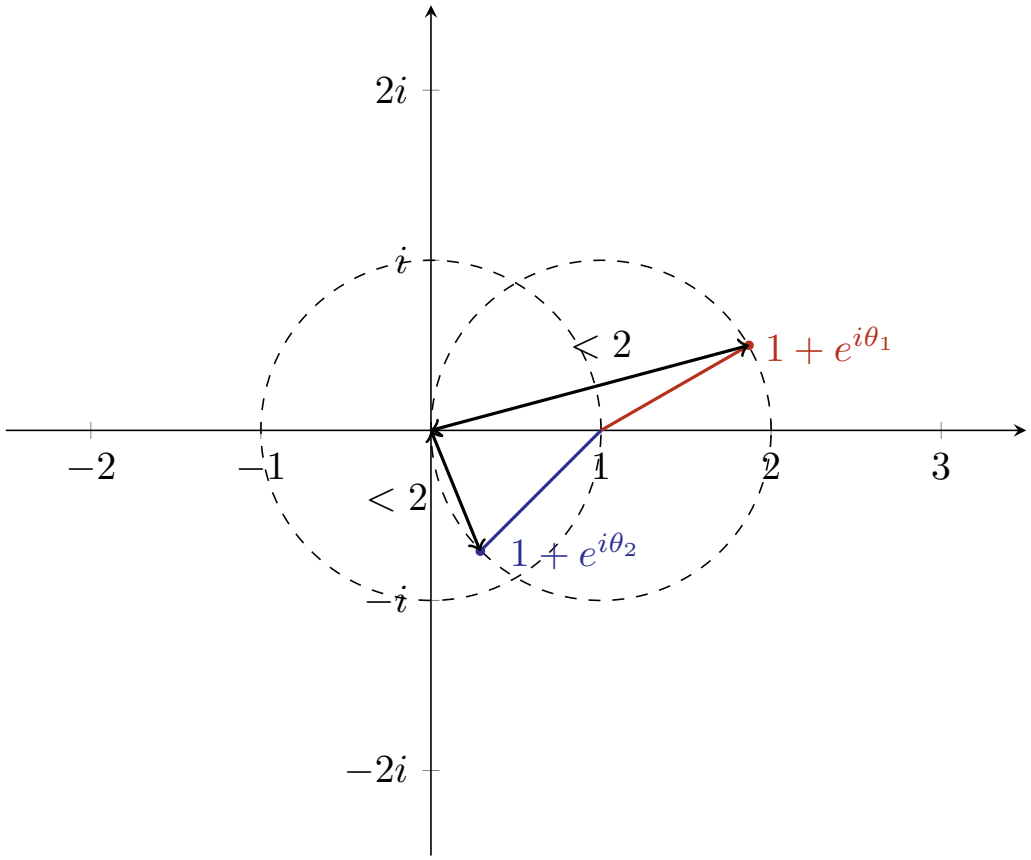

Figure 9: Un dessin pour justifier une majoration

Projection sur un convexe fermé

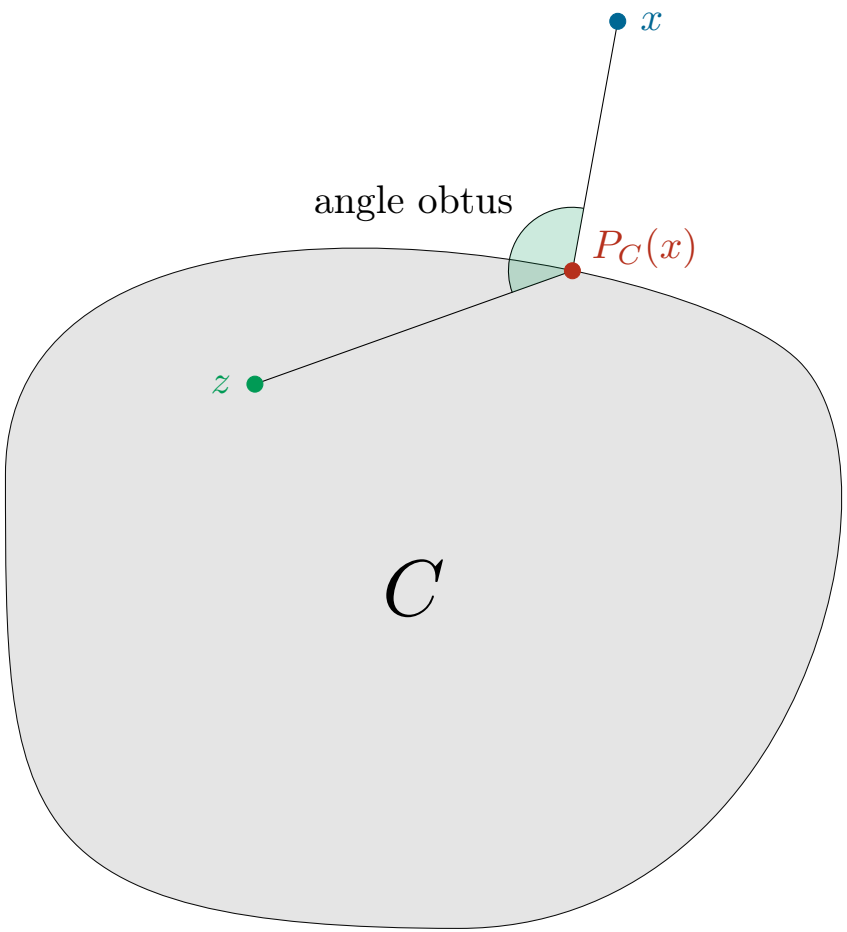

Figure 10: Caractérisation du projeté orthogonal sur un convexe fermé

Distance entre un point et une partie fermée d'un e.v.n de dimension finie

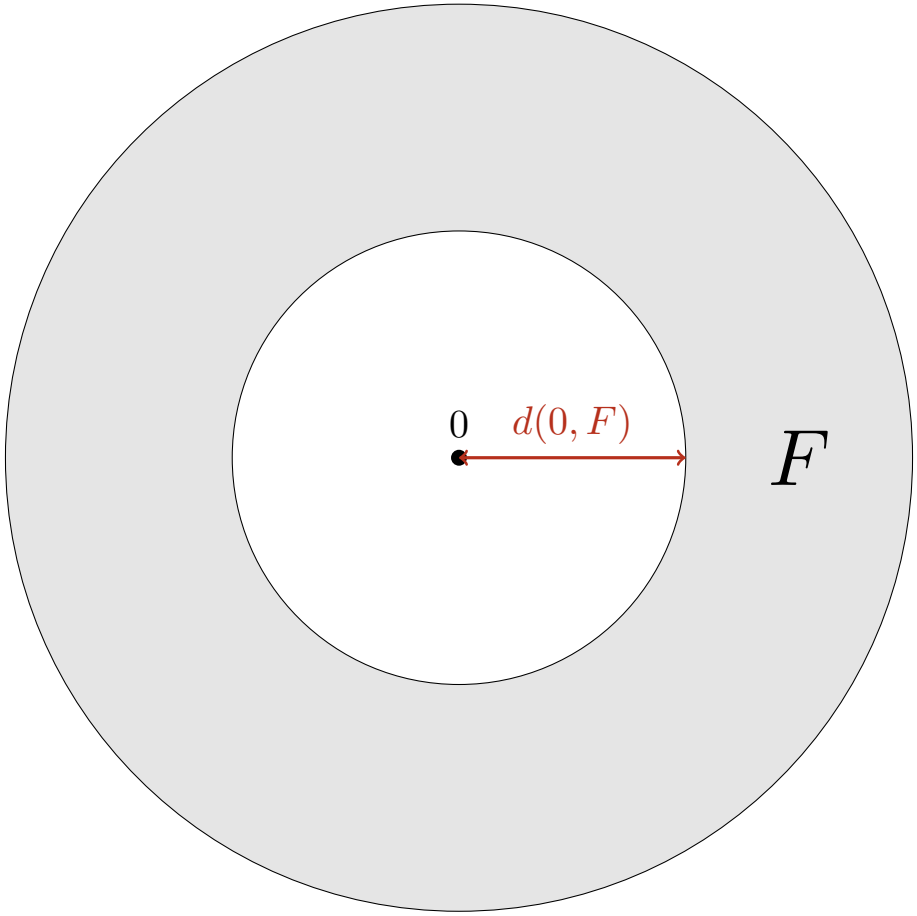

Figure 11: Distance entre un point et une partie fermée d'un espace vectoriel de dimension finie

Soit $F = \{x \in \mathbb{R}^2 \mid 1 \leq \|x\|_2 \leq 2\}$ (où $\|\cdot\|_2$ désigne la norme usuelle sur $\mathbb{R}^2$). Alors, F est une partie non vide et fermée de $\mathbb{R}^2$ et,

$$\forall y \in \mathbb{R}^2,\ (\|y\|_2 = 1 \Longrightarrow d(0, F) = \|y - 0\|).$$

Ainsi, la distance entre 0 et F est atteinte par une infinité de points de F.

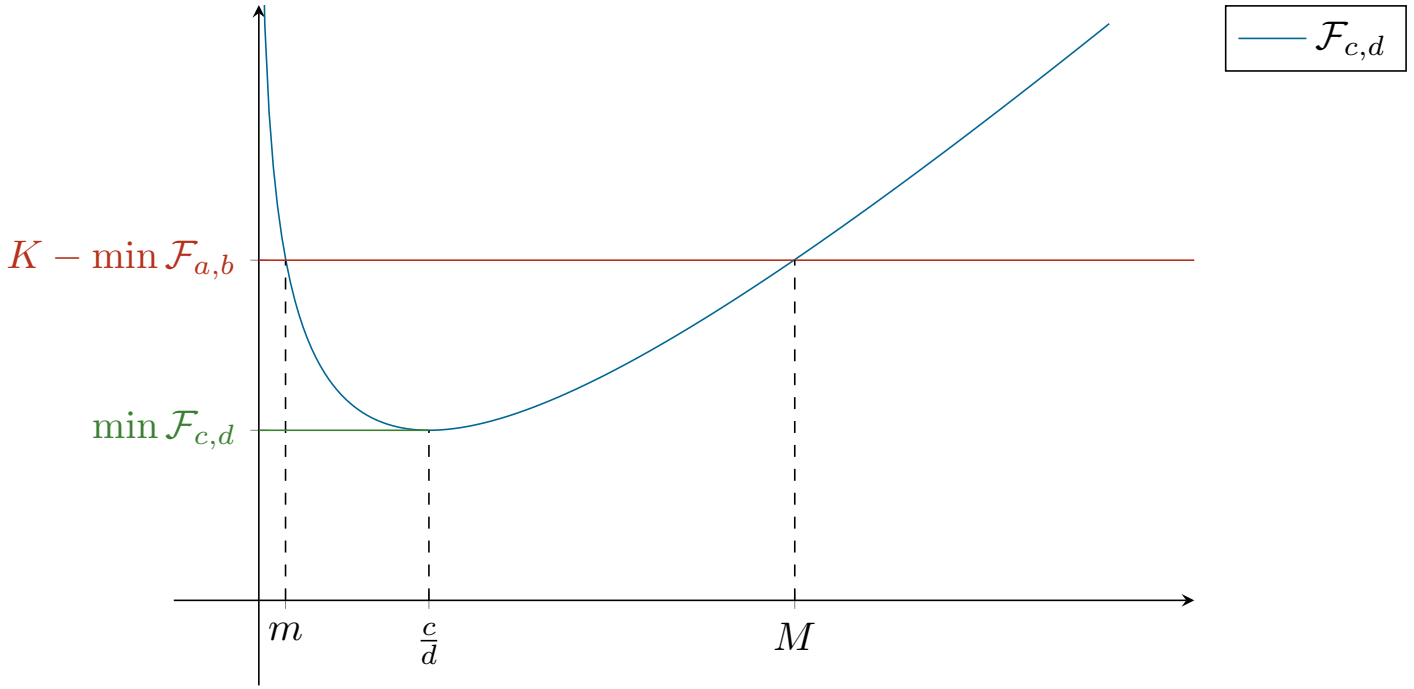

Figure 12: Allure du graphe de $\mathcal{F}_{c,d}$

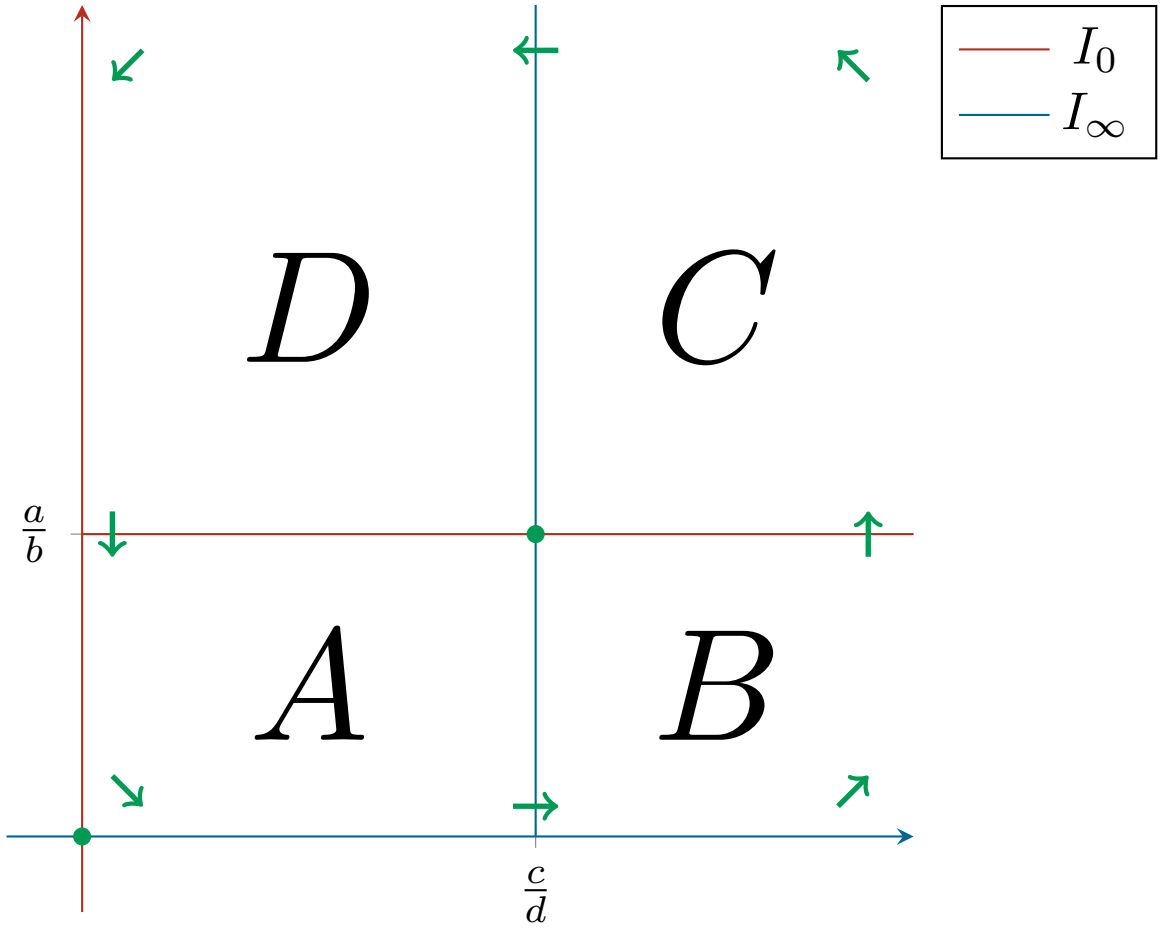

Figure 13: Découpage du quart de plan en régions

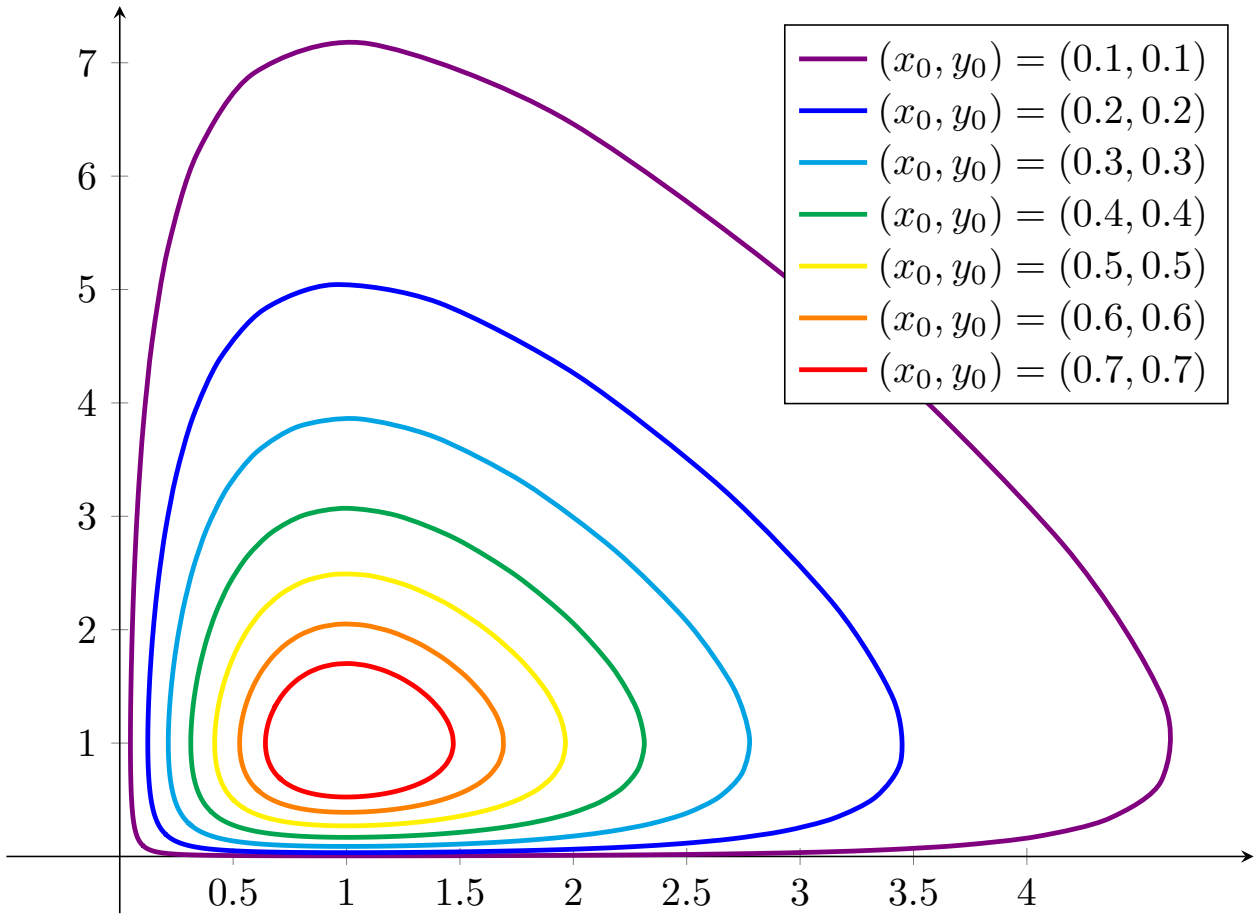

Figure 14: Portrait de phase de (E) pour $a = b = 1$, $c = d = 2$ et $(x_0, y_0) \in \{(k \cdot 10^{-1}, k \cdot 10^{-1}),\ 1 \leq k \leq 7\}$.

Théorème d'inversion locale

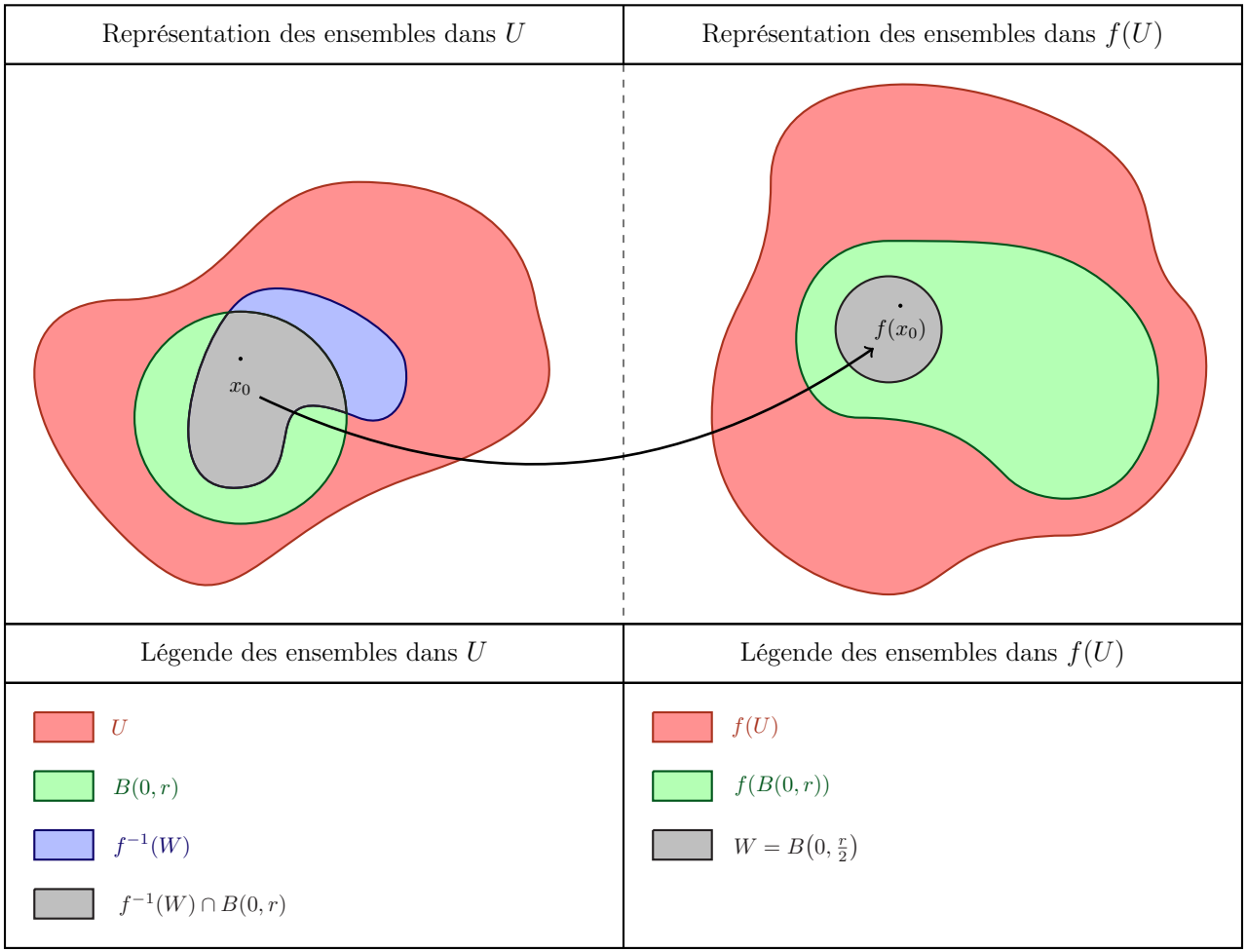

Figure 15: Panorama des ensembles intervenant dans la construction du difféomorphisme

Index des notations

Ensembles

- Ensemble des entiers compris entre a et b : $[\![a, b]\!]$
- Ensemble des entiers naturels : $\mathbb{N}$
- Ensemble des entiers relatifs : $\mathbb{Z}$
- Ensemble des nombres rationnels : $\mathbb{Q}$
- Ensemble des nombres réels : $\mathbb{R}$
- Ensemble des nombres complexes : $\mathbb{C}$
- Disque unité ouvert (des nombres complexes) : $\mathbb{D}$
- Cercle unité (des nombres complexes) : $\mathbb{U}$
- Ensemble des racines n-ièmes de l'unité : U_n
- Ensemble des éléments non nuls d'un anneau $\mathbb{A}$: $\mathbb{A}^*$

  ⚠ ne pas confondre avec $\mathbb{A}^\star$
- Ensemble des éléments respectivement positifs, strictement positifs, négatifs, strictement négatifs d'un anneau totalement ordonné $\mathbb{A}$: $\mathbb{A}_+$, $\mathbb{A}_+^*$, $\mathbb{A}_-$, $\mathbb{A}_-^*$
- Partie réelle d'un nombre complexe z : $\Re(z)$
- Partie imaginaire d'un nombre complexe z : $\Im(z)$
- Application identité d'un ensemble E : Id_E
- Fonction indicatrice d'une partie A d'un ensemble E : 1_A
- Cardinal d'un ensemble E : $|E|$ ou $\mathrm{card}(E)$ (varie selon mon humeur)

Fonctions usuelles

- Fonction zêta de Riemann : ζ
- Fonction Gamma d'Euler : Γ

Groupes

- Centre d'un groupe G : $Z(G)$
- Ordre d'un élément g dans un groupe G : $\mathrm{ordre}(g)$
- Groupe des automorphismes d'un groupe G : $\mathrm{Aut}(G)$
- Groupe des automorphismes intérieurs d'un groupe G : $\mathrm{Int}(G)$
- Groupe symétrique d'ordre n : $\mathfrak{S}_n$
- Groupe alterné d'ordre n : $\mathfrak{A}_n$
- Support d'une permutation σ : $\mathrm{Supp}(\sigma)$

Arithmétique (dans les anneaux)

- Groupe des inversibles d'un anneau $\mathbb{A}$: $\mathbb{A}^\star$

  ⚠ ne pas confondre avec $\mathbb{A}^*$
- Idéal engendré par un élément a dans un anneau $\mathbb{A}$: (a)
- Plus grand diviseur commun : pgcd (défini modulo $\mathbb{A}^\star$)
- Plus petit multiple commun : ppcm (défini modulo $\mathbb{A}^\star$)
- Anneau des polynômes à coefficients dans un anneau $\mathbb{A}$: $\mathbb{A}[X]$
- Degré d'un polynôme P à coefficients dans un anneau : $\deg(P)$
- Contenu d'un polynôme P à coefficients dans un anneau factoriel : $c(P)$ (défini modulo $\mathbb{A}^\star$)
- Anneau des entiers de Gauss : $\mathbb{Z}[i]$
- Polynôme cyclotomique d'ordre n : Φ_n

Corps

- Degré d'une extension $\mathbb{L}$ d'un corps $\mathbb{K}$: $[\mathbb{L} : \mathbb{K}]$

Espaces vectoriels

- Dimension d'un $\mathbb{K}$-espace vectoriel E : $\dim(E)$
- Espace vectoriel engendré par une partie A d'un espace vectoriel E : $\mathrm{Vect}(A)$
- Noyau d'une application linéaire f entre deux espaces vectoriels : $\mathrm{Ker}(f)$
- Image d'une application linéaire f entre deux espaces vectoriels : $\mathrm{Im}(f)$
- Rang d'une application linéaire f entre deux espaces vectoriels : $\mathrm{rg}(f)$
- Orthogonal d'une partie A d'un espace préhilbertien : $A^\perp$
- Dual algébrique d'un espace vectoriel E : E^*
- Dual topologique d'un espace vectoriel normé E : E'
- Enveloppe convexe d'une partie A d'un $\mathbb{R}$-espace vectoriel E : $\mathrm{Conv}(A)$
- Spectre d'un endomorphisme f d'un espace vectoriel E de dimension finie : $\mathrm{Sp}(f)$
- Polynôme caractéristique d'un endomorphisme f d'un espace vectoriel E de dimension finie : χ_f
- Polynôme minimal d'un endomorphisme f d'un espace vectoriel E de dimension finie : π_f

Matrices

- Espace vectoriel des matrices à n lignes et p colonnes sur un corps $\mathbb{K}$: $M_{n,p}(\mathbb{K})$
- Espace vectoriel des matrices carrées sur un corps $\mathbb{K}$: $M_n(\mathbb{K})$
- Matrice identité de taille n : I_n
- Matrice nulle (de taille non spécifiée) : $\mathbf{0}$
- Matrice d'un endomorphisme f dans une base $\mathcal{B}$: $\mathrm{Mat}_{\mathcal{B}}(f)$
- Matrice d'une forme quadratique q dans une base $\mathcal{B}$: $\mathrm{Mat}_{\mathcal{B}}(q)$
- Rang d'une matrice A à coefficients dans un corps : $\mathrm{rg}(A)$
- Trace d'une matrice carrée A à coefficients dans un corps : $\mathrm{Tr}(A)$
- Déterminant d'une matrice carrée A à coefficients dans un corps : $\det(A)$
- Groupe linéaire sur un corps $\mathbb{K}$: $GL_n(\mathbb{K})$
- Groupe spécial linéaire sur un corps $\mathbb{K}$: $SL_n(\mathbb{K})$
- Transposée d'une matrice A sur un corps $\mathbb{K}$: ${}^tA$
- Espace vectoriel des matrices symétriques sur un corps $\mathbb{K}$: $S_n(\mathbb{K})$
- Espace vectoriel des matrices antisymétriques sur un corps $\mathbb{K}$: $A_n(\mathbb{K})$
- Groupe orthogonal d'ordre n : $O_n(\mathbb{R})$
- Groupe spécial orthogonal d'ordre n : $SO_n(\mathbb{R})$
- Espace vectoriel des matrices symétriques positives : $S_n^+(\mathbb{R})$
- Espace vectoriel des matrices symétriques définies positives : $S_n^{++}(\mathbb{R})$
- Polynôme caractéristique d'une matrice carrée A à coefficients dans un corps : χ_A
- Polynôme minimal d'une matrice carrée A à coefficients dans un corps : π_A
- Exponentielle d'une matrice carrée A à coefficients réels ou complexes : $\exp(A)$ ou e^A
- Spectre d'une matrice carrée A à coefficients dans un corps $\mathbb{K}$: $\mathrm{Sp}_{\mathbb{K}}(A)$ ($\mathrm{Sp}(A)$ en l'absence d'ambiguïté)

Calcul différentiel

- Différentielle d'une fonction f au point x_0 : $Df(x_0)$
- Différentielle d'une fonction f au point x_0, évaluée en h : $Df(x_0)(h)$
- Gradient d'une fonction f au point x : $\nabla f(x)$
- Dérivée partielle d'une fonction f par rapport à sa k-ième variable au point x : $\partial_k f(x)$

  ⚠ Dans le développement sur l'équation de la chaleur, $\partial_t u(t,x)$ et $\partial_x u(t,x)$ désigne les dérivées partielles respectives de u par rapport à sa première et deuxième variable au point (t,x).

- Dérivée partielle d'ordre n d'une fonction f par rapport à sa k-ième variable au point x : $\partial_k^n f(x)$

  ⚠ Dans le développement sur l'équation de la chaleur, $\partial_x^2 u(t,x)$ désigne la dérivée partielle seconde de u par rapport à sa deuxième variable.

Espaces métriques

- Adhérence d'une partie Y d'un espace métrique X : $\overline{Y}$

- Boule ouverte centrée en x de rayon r d'un espace métrique X : $B_X(x, r)$ ($B(x, r)$ en l'absence d'ambiguïté)
- Distance d'un point x à une partie non vide Y d'un espace métrique X : $d(x, Y)$
- Distance entre deux parties non vides Y, Z d'un espace métrique X : $d(Y, Z)$

Espaces probabilisés

- Tribu des boréliens d'un espace topologique X : $\mathcal{B}(X)$
- Espace de Lebesgue sur un espace mesuré (X, T, m) à valeurs dans $\mathbb{K} \in \{\mathbb{N}, \mathbb{R}, \mathbb{C}\}$: $L_{\mathbb{K}}(X, T, m)$
- Espérance d'une variable aléatoire X d'un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ vers $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$: $\mathbb{E}(X)$
- Variance d'une variable aléatoire X d'un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ vers $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$: $\mathbb{V}(X)$
- Loi normale d'espérance μ et d'écart-type σ : $\mathcal{N}(\mu, \sigma^2)$
- La variable aléatoire X suit une loi de probabilité μ : $X \hookrightarrow \mu$