Développements pour l'agrégation externe

Florian LEMONNIER

Année 2014 – 2015

Avant-propos

Cher agrégatif,

Avant toute chose, bon courage.

Voici, dans ce document, tous les développements que j'ai rédigés durant mon année de préparation à l'agrégation externe de mathématiques, en 2014/2015. J'ai essayé, le plus souvent possible, d'y ajouter des annotations, des commentaires, ou des idées à retenir pour les questions.

Évidemment, ce document ne peut pas être exempt de coquilles, de mauvaise foi, d'explications trop elliptiques, ou même d'erreurs mathématiques grossières. Gardez en tête cet avertissement ! Il va également de soi que l'oral de l'agrégation est quelque chose qu'on doit préparer de façon personnelle ; vous et moi, nous ne pouvons pas avoir les mêmes goûts, ni les mêmes aptitudes. Ce document ne peut pas être plus qu'une boîte à idées : un développement, ça se prépare et ça se comprend au tableau. Vous ne travaillerez pas votre rythme de présentation en lisant une feuille de papier !

Je ne peux absolument pas assumer seul la paternité de ce document, car beaucoup d'idées m'ont été inspirées par d'autres personnes que je tiens tout particulièrement à remercier :
- ceux et celles qui, par le passé, ont mis en ligne leurs développements, ou leurs plans résumés (et tout particulièrement Arnaud GIRAND, Ophélie ROUBY et Hélène HIVERT) ;
- ceux et celles qui, durant l'année, ont vu et/ou lu certains de mes développements, et qui m'ont permis de les corriger et/ou de les enrichir de commentaires, d'idées, ou de questions (et parmi eux, notamment, Arnaud STOCKER, Laura GAY, Maud JOUBAUD et Caroline ROBET) ;
- les enseignants qui ont encadré la préparation à l'ENS Rennes et à l'université Rennes 1 durant l'année 2014/2015, et qui nous ont posé des questions "classiques".

Je me suis efforcé de mettre des références aussi précises et récentes que possible ; cependant, quelques développements ont été considérablement modifiés entre la référence utilisée et le développement final. C'est comme ça !

Chaque développement est précédé d'une liste de leçons : celles écrites en droit correspondent à mes couplages en fin d'année ; celles écrites en penché correspondent à d'autres couplages possibles, mais qui me plaisaient moins (ou à une cruelle présence de mauvaise foi).

Bien sûr, je reste ouvert à toute remarque de votre part.

Diffusion à titre gratuit uniquement.

Table des matières

Diffusion à titre gratuit uniquement.

Diffusion à titre gratuit uniquement.

Automorphismes de $K(X)$

Leçons : 140, *141*

[X-ENS Al1], exercice 5.54

Théorème

Soit K un corps quelconque.

Les automorphismes de K-algèbres de $K(X)$ sont les applications de la forme $\left|\begin{array}{ccc} K(X) & \to & K(X) \\ G & \mapsto & G\left(\frac{aX+b}{cX+d}\right) \end{array}\right.$

où $a, b, c, d \in K^4$ vérifient $ad - bc \neq 0$.

Démonstration :

Étape 1 : Déterminons les endomorphismes de K-algèbres de $K(X)$.

Soit Φ un endomorphisme de K-algèbres de $K(X)$. Posons $F = \Phi(X)$.

Soit $P = \sum_{k\in\mathbb{N}} p_k X^k \in K[X]$, on a : $\Phi(P) = \sum_{k\in\mathbb{N}} p_k \Phi(X^k) = \sum_{k\in\mathbb{N}} p_k F^k = P \circ F$.

Ainsi, pour tous $P \in K[X]$, $Q \in K[X]\backslash\{0\}$ et $G = \dfrac{P}{Q}$, où, on a : $\Phi(G) = \dfrac{\Phi(P)}{\Phi(Q)} = \dfrac{P \circ F}{Q \circ F} = G \circ F$.

Réciproquement, pour $F \in K(X)\backslash K$, $\Phi_F : \left|\begin{array}{ccc} K(X) & \to & K(X) \\ G & \mapsto & G \circ F \end{array}\right.$ est bien un morphisme de K-algèbres. [1]

Remarquons que si $F = a \in K$, alors Φ_F n'est pas bien défini : en effet $\dfrac{1}{X-a}$ n'a pas d'image par Φ_F.

Ainsi, l'ensemble des endomorphismes de K-algèbres de $K(X)$ est l'ensemble des Φ_F, où F parcourt $K(X)\backslash K$.

Étape 2 : Cherchons à quelle condition sur F, Φ_F est un automorphisme.

Supposons que Φ_F soit un automorphisme.

Alors Φ_F est surjectif et donc : $\exists G \in K(X), \Phi_F(G) = G \circ F = X$.

Soient $F = \dfrac{A}{B}$ et $G = \dfrac{P}{Q}$ des représentations irréductibles de ces fractions rationnelles.

On écrit $P = \sum_{j=0}^{d_P} p_j X^j$ et $Q = \sum_{k=0}^{d_Q} q_k X^k$ où d_P et d_Q sont les degrés respectifs de P et Q.

On a : $G \circ F = X \Leftrightarrow P \circ F = X(Q \circ F)$

$$\Leftrightarrow \sum_{j=0}^{d_P} p_j F^j = X \sum_{k=0}^{d_Q} q_k F^k$$

$$\Leftrightarrow \sum_{j=0}^{d_P} p_j \frac{A^j}{B^j} = X \sum_{k=0}^{d_Q} q_k \frac{A^k}{B^k}$$

$$\Leftrightarrow \sum_{j=0}^{d_P} p_j A^j B^{m-j} = X \sum_{k=0}^{d_Q} q_k A^k B^{m-k}, \text{ où on a noté } m = \max\left\{d_P, d_Q\right\}.$$

– D'une part, $A | (p_0 - q_0 X)\, B^m$.
Or A et B sont premiers entre eux, donc A et B^m sont également premiers entre eux, d'où $A | p_0 - q_0 X$.
Aussi, comme P et Q sont premiers entre eux, on a $(p_0, q_0) \neq (0,0)$.
Donc $p_0 - q_0 X$ est de degré 0 ou 1, donc A est de degré 0 ou 1.

1. Il s'agit de vérifier :
– $\Phi_F(1) = 1$.
– $\Phi_F(G)$ est bien défini pour tout $G \in K(X)$.
Pour cela, on montre que $G \circ F = \dfrac{B^m \times (P \circ F)}{B^m \times (Q \circ F)}$, où $F = \dfrac{A}{B}$ est une écriture irréductible et $m = \max\{\deg P, \deg Q\}$ est une écriture de $G \circ F$ en fraction de polynômes et que le polynôme $B^m \times (Q \circ F)$ n'a qu'un nombre fini de racines.
– Φ_F est K-linéaire.
– $\Phi_F(G_1 G_2) = \Phi_F(G_1)\, \Phi_F(G_2)$ pour tous $G_1, G_2 \in K(X)$.

Diffusion à titre gratuit uniquement.

– D'autre part, $B \Big| p_{d_P} A^{d_P} B^{m-d_P} - q_{d_Q} X A^{d_Q} B^{m-d_Q}$.
  – Si $m = d_P = d_Q$, alors $B|(p_m - q_m X)\, A^m$ et donc $B|p_m - q_m X$ car B et A^m sont premiers entre eux. Or $q_m \neq 0$ car Q est de degré $m = d_Q$, donc $\deg B = 0$ ou 1.
  – Si $m = d_Q > d_P$, alors $B|q_m X A^m$ donc $B|q_m X$ donc $\deg B = 0$ ou 1.
  – Si $m = d_P > d_Q$, alors $B|p_m A^m$ donc $B|p_m$ donc $\deg B = 0$.
  Toujours est-il que $\deg B = 0$ ou 1.

Par conséquent, il existe $(a,b,c,d) \in K^4, F = \dfrac{aX+b}{cX+d}$.

Et F ne pouvant évidemment pas être constant, on doit imposer $ad - bc \neq 0$.

Étape 3 : Réciproquement, montrons que cette condition nécessaire est suffisante.

Pour $(a,b,c,d) \in K^4$ vérifiant $ad - bc \neq 0$, on note $\Phi_{\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)}$ le morphisme de K-algèbres défini par :

$$\Phi_{\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)}(X) = \frac{aX+b}{cX+d}.$$

Montrons que : $\Phi : \left| \begin{array}{ccc} \mathrm{GL}_2(K) & \to & \mathrm{End}_{K\text{-alg.}}(K(X)) \\ M & \mapsto & \Phi_{M^{-1}} \end{array} \right.$ est un morphisme de groupes.

Cela découle de l'égalité :

$$\Phi_{\left(\begin{smallmatrix} a' & b' \\ c' & d' \end{smallmatrix}\right)}\left(\Phi_{\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)}(X)\right) = \frac{aX+b}{cX+d} \circ \frac{a'X+b'}{c'X+d'} = \frac{a\frac{a'X+b'}{c'X+d'}+b}{c\frac{a'X+b'}{c'X+d'}+d} = \Phi_{\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)\left(\begin{smallmatrix} a' & b' \\ c' & d' \end{smallmatrix}\right)}(X)$$

On reconnaît effectivement les coefficients du produit matriciel $\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix} = \begin{pmatrix} aa'+bc' & ab'+bd' \\ ca'+dc' & cb'+dd' \end{pmatrix}$.

Donc $\forall M, N \in \mathrm{GL}_2(K), \Phi(MN) = \Phi_{(MN)^{-1}} = \Phi_{N^{-1}M^{-1}} = \Phi_{M^{-1}} \circ \Phi_{N^{-1}} = \Phi(M) \circ \Phi(N)$.

On en déduit alors que Φ_M est inversible, d'inverse $\Phi_{M^{-1}}$, pour tout $M \in \mathrm{GL}_2(K)$.

On en déduit finalement que l'ensemble des automorphismes de K-algèbres de $K(X)$ est l'ensemble des applications de la forme $\left| \begin{array}{ccc} K(X) & \to & K(X) \\ G & \mapsto & G\left(\frac{aX+b}{cX+d}\right) \end{array} \right.$ avec $(a,b,c,d) \in K^4$ vérifiant $ad - bc \neq 0$.[2] ■

Références

[X-ENS Al1] S. FRANCINOU, H. GIANELLA et S. NICOLAS – *Oraux X-ENS Algèbre 1*, 2e éd., Cassini, 2007.

2. Le 2nd théorème d'isomorphisme peut même nous donner un résultat supplémentaire.
On a montré au cours de la démonstration que $\mathrm{Im}(\Phi) = \mathrm{Aut}_{K\text{-alg.}}(K(X))$.
De plus

$$\Phi_{a,b,c,d}(X) = X \Leftrightarrow aX + b = cX^2 + dX \Leftrightarrow b = c = 0 \text{ et } a = d$$

Donc $\mathrm{Ker}(\Phi) = \{\lambda I_2 | \lambda \in K^\times\}$.
Et on en déduit donc :

$$\mathrm{Aut}_{K\text{-alg.}}(K(X)) \simeq {}^{\mathrm{GL}_2(K)}\!/_{\mathrm{Ker}(\Phi)} = {}^{\mathrm{GL}_2(K)}\!/_{\{\lambda I_2 | \lambda \in K^\times\}} = \mathrm{PGL}_2(K)$$

Diffusion à titre gratuit uniquement.

Borne de Bézout [3]

Leçons : 143, 180 [4], *142, 144*

Merci Arnaud ! [5]

Théorème

Soient $A, B \in \mathbb{K}[X,Y]$ de degrés totaux respectifs m et n ; on suppose que A et B sont premiers entre eux et que $\mathbb{K}$ est de cardinal infini [6].
On note $V(A) = \left\{(x,y) \in \mathbb{K}^2 \middle| A(x,y) = 0\right\}$ et $V(B) = \left\{(x,y) \in \mathbb{K}^2 \middle| B(x,y) = 0\right\}$.
Alors on a : $\#(V(A) \cap V(B)) \leqslant mn$.

Démonstration :

Évidemment, on suppose $V(A) \cap V(B) \neq \varnothing$, car sinon, il n'y a rien à montrer.

Étape 1 : Soit $(x,y) \in V(A) \cap V(B)$; on note $R_Y := \mathrm{Res}_Y(A,B)$, et on a $R_Y(x) = 0$.
Comme $A \wedge B = 1$, on sait que $R_Y \in \mathbb{K}[X]\backslash\{0\}$ et donc R_Y admet au plus $\deg R_Y$ racines.
Ainsi, pour un point de $V(A) \cap V(B)$, il y a au plus $\deg R_Y$ abscisses possibles, et similairement, au plus $\deg R_X$ ordonnées possibles. [7]
Dès lors, $\#(V(A) \cap V(B)) \leqslant \deg R_X \deg R_Y < \infty$.

Étape 2 : Obtenons désormais une borne sur $\deg R_Y$.

On note $A(X,Y) = \sum_{k=0}^{p} a_k(X)Y^k$ et $B(X,Y) = \sum_{k=0}^{q} b_k(X)Y^k$, avec $\deg a_k \leqslant m-k$, $\deg b_k \leqslant n-k$, $a_p \neq 0$ et $b_q \neq 0$.

Par conséquent, $R_Y = \det\left(\mathrm{Syl}_Y(A,B)\right) =$

$$\left|\begin{array}{cccccc} a_p & \cdots & \cdots & \cdots & a_0 & 0 \\ & \ddots & & & & \ddots \\ 0 & & a_p & \cdots & \cdots & a_0 \\ b_q & \cdots & \cdots & b_0 & & \\ & \ddots & & & \ddots & 0 \\ 0 & & \ddots & & & \ddots \\ & & & b_q & \cdots & b_0 \end{array}\right| \begin{array}{l} \left.\vphantom{\begin{array}{c}a\\a\\a\end{array}}\right\} q \text{ lignes} \\ \left.\vphantom{\begin{array}{c}a\\a\\a\end{array}}\right\} p \text{ lignes} \end{array} .$$

Notons $\mathrm{Syl}_Y(A,B) = \left(c_{i,j}\right)_{1 \leqslant i,j \leqslant p+q}$.

Soit $i \in [\![1,q]\!]$, on a : $\deg c_{i,j} = \begin{cases} \deg a_{p-(j-i)} & \text{si } i \leqslant j \leqslant p+i \\ -\infty & \text{sinon} \end{cases} \leqslant m-p+j-i.$

Soit $i \in [\![q+1, q+p]\!]$, on a : $\deg c_{i,j} = \begin{cases} \deg b_{q-(j-(i-q))} & \text{si } i-q \leqslant j \leqslant i \\ -\infty & \text{sinon} \end{cases} \leqslant n-i+j.$ [8]

3. Je sais : vous vous posez très probablement la même question que moi. Et je vous propose de clore ici immédiatement le débat sur l'accent du 'e' du nom de famille d'Étienne machin (1730-1783). Si on se réfère à une thèse de Liliane ALFONSI, qui a écrit *Étienne Bézout, mathématicien des Lumières* en 2011, l'accent apparaît ou non selon les documents, mais il est mis systématiquement à partir d'une certaine date (1765 pour les manuscrits, 1770 pour les imprimés) par Bézout lui-même. En conséquence, respectons son choix de mettre un accent ; même si l'académie de Créteil le lui a retiré en donnant le nom d'Étienne Bezout à un lycée de Nemours, en Seine-et-Marne.

4. Dans la 180, je prends un des deux polynômes de degré total égal à 2. Okay, c'est moche, mais qui veut faire les coniques devant le jury ?

5. Un lien vers la page personnelle d'Arnaud STOCKER. On trouvera néanmoins une piste de démonstration au théorème 10.111 de A. SZPIRGLAS – *Algèbre pour la L3*, Pearson Éducation, 2009.

6. Cette hypothèse est dispensable : si $\mathbb{K}$ est de cardinal fini, alors $\mathbb{K}$ s'injecte dans sa clôture algébrique $\overline{\mathbb{K}}$ qui est de cardinal infini. Alors les courbes algébriques définies par A et B sur $\overline{\mathbb{K}}^2$ s'intersectent en au plus mn points de $\overline{\mathbb{K}}^2$. En conséquence, les courbes algébriques définies par A et B sur $\mathbb{K}^2$ s'intersectent en au plus mn points de $\mathbb{K}^2$.

7. Vous l'aurez compris, $R_X = \mathrm{Res}_X(A,B) \in \mathbb{K}[Y]\backslash\{0\}$.

8. C'est LE passage difficile du développement, parce qu'il faut être capable d'expliquer ceci clairement au tableau. Quand $i \in [\![1,q]\!]$, on remarque que $c_{i,i} = a_p$, puis, quand $0 \leqslant j-i \leqslant p$, passer de $c_{i,i}$ à $c_{i,j}$ revient à faire $(j-i)$ pas vers la droite, donc à passer de a_p à $a_{p-(j-i)}$. Quand $i \in [\![q+1, q+p]\!]$, on opère similairement : $c_{i,i-q} = b_q$, puis, en décalant de $(j-(i-q))$ cases vers la droite, où $0 \leqslant j-(i-q) \leqslant q$, on obtient $c_{i,j} = b_{q-(j-(i-q))}$.

Diffusion à titre gratuit uniquement.

Par la formule du déterminant en fonction des coefficients de la matrice : $R_Y = \sum_{\sigma\in\mathfrak{S}_{p+q}} \underbrace{\varepsilon(\sigma)\prod_{i=1}^{p+q} c_{i,\sigma(i)}}_{=:F_\sigma}$.

Et pour tout $\sigma \in \mathfrak{S}_{p+q}$, on a :

$$\deg F_\sigma = \sum_{i=1}^{p+q} \deg c_{i,\sigma(i)} \leqslant \sum_{i=1}^{q} m - p + \sigma(i) - i + \sum_{i=q+1}^{q+p} n - i + \sigma(i) \underset{9}{=} mq - pq + np = mn + \underbrace{(m-p)(q-n)}_{\leqslant 0}$$
$$\leqslant mn.$$

En conséquence, $\deg R_Y \leqslant mn$.
Similairement, $\deg R_X \leqslant mn$ et donc $\#(V(A)\cap V(B)) \leqslant (mn)^2$.

Étape 3 : On numérote alors les éléments de $V(A)\cap V(B)$: $V(A)\cap V(B) = \{(x_i,y_i) | i \in [\![1,r]\!]\}$.

Soit $\mathcal{E} := \left\{ \dfrac{x_i - x_j}{y_j - y_i} \middle| y_i \neq y_j, i,j \in [\![1,r]\!] \right\}$, alors $\#\mathcal{E} < \infty = \#\mathbb{K}^\times$.

Donc $\exists u \in \mathbb{K}^\times \setminus \mathcal{E}$ et ensuite $\forall i,j \in [\![1,r]\!], x_i - x_j \neq u\left(y_j - y_i\right) \Leftrightarrow x_i + uy_i \neq x_j + uy_j$.

On effectue alors le changement de variables : $\begin{cases} X' = X + uY \\ Y' = Y \end{cases}$ et on pose $\begin{cases} \widetilde{A}(X',Y') = A(X,Y) \\ \widetilde{B}(X',Y') = B(X,Y) \end{cases}$.

Soit alors Φ : $\left|\begin{array}{ccc} V(A)\cap V(B) & \to & \mathcal{R}ac\left(\mathrm{Res}_{Y'}\left(\widetilde{A},\widetilde{B}\right)\right) \\ (x,y) & \mapsto & x+uy \end{array}\right.$.

– Φ est bien définie :

$$(x,y) \in V(A)\cap V(B) \Rightarrow A(x,y) = B(x,y) = 0 \Rightarrow \widetilde{A}(x+uy,y) = \widetilde{B}(x+uy,y) = 0$$
$$\Rightarrow \mathrm{Res}_{Y'}\left(\widetilde{A}(x+uy,Y'),\widetilde{B}(x+uy,Y')\right) = 0 \Rightarrow \left(\mathrm{Res}_{Y'}\left(\widetilde{A},\widetilde{B}\right)\right)(x+uy) = 0$$

– Φ est injective :
Si (x,y) et $(x',y') \in V(A)\cap V(B)$ sont distincts, alors $u \notin \mathcal{E}$ impose $x+uy \neq x'+uy'$.

D'où : $\#(V(A)\cap V(B)) \leqslant \#\mathcal{R}ac\left(\mathrm{Res}_{Y'}\left(\widetilde{A},\widetilde{B}\right)\right) \leqslant \deg R_{Y'}\left(\widetilde{A},\widetilde{B}\right) \underset{10}{\leqslant} mn$. ■

9. En effet, σ étant une bijection de $[\![1,p+q]\!]$, on a : $\sum_{i=1}^{p+q} \sigma(i) - i = 0$.

10. Car les degrés de $\widetilde{A}$ et $\widetilde{B}$ sont inférieurs ou égaux à ceux de A et B.

Diffusion à titre gratuit uniquement.

Décomposition de Dunford

Leçons : 153, 154, 155, 157

[Gou Al], partie 4.4.2

Théorème

Soit E un $\mathbb{K}$-espace vectoriel de dimension finie, et $u \in \mathcal{L}(E)$ tel que χ_u soit scindé sur $\mathbb{K}$.
Alors, il existe un unique couple $(d, n) \in \mathcal{L}(E)^2$, tel que d soit diagonalisable, n nilpotent, $u = d + n$ et $d \circ n = n \circ d$.
De plus, $(d, n) \in \mathbb{K}[u]^2$.

On commence par montrer le lemme qui suit.

Lemme

Soit $u \in \mathcal{L}(E)$ et $F \in \mathbb{K}[X]$ tel que $F(u) = 0$.
Soit $F = \beta \prod_{i=1}^{s} M_i^{\alpha_i}$ la décomposition en facteurs irréductibles de F dans $\mathbb{K}[X]$.
Pour $i \in [\![1, s]\!]$, on note $N_i = \operatorname{Ker} M_i^{\alpha_i}(u)$.
On a alors :

- $E = \bigoplus_{i=1}^{s} N_i$;
- $\forall i \in [\![1, s]\!]$, le projecteur sur N_i parallèlement à $\bigoplus_{j \neq i} N_j$ est un polynôme en u.

Démonstration du lemme :

L'égalité $E = \bigoplus_{i=1}^{s} N_i$ découle directement du lemme des noyaux.

Pour $i \in [\![1, s]\!]$, on note $Q_i = \prod_{j \neq i} M_j^{\alpha_j}$; alors les Q_i, où $i \in [\![1, s]\!]$, sont premiers entre eux dans leur ensemble.

D'après Bézout, on obtient donc :

$$\exists U_1, \dots, U_s \in \mathbb{K}[X], \sum_{i=1}^{s} U_i Q_i = 1.$$

Pour $i \in [\![1, s]\!]$, on note $P_i = U_i Q_i$ et $p_i = P_i(u) \in \mathbb{K}[u]$.
L'égalité de Bézout nous fournit :

$$\sum_{i=1}^{s} p_i = \mathrm{Id}_E \tag{1}$$

Par ailleurs, si $i \neq j$, alors $F | Q_i Q_j$, de sorte que

$$\forall j \neq i \in [\![1, s]\!], p_i \circ p_j = Q_i Q_j(u) \circ U_i U_j(u) = 0 \tag{2}$$

On en déduit, par (1) : $\forall j \in [\![1, s]\!], p_j = \sum_{i=1}^{s} p_i \circ p_j$; puis, par (2) : $\forall j \in [\![1, s]\!], p_j^2 = p_j$.

En conséquence, les applications p_i, où $i \in [\![1, s]\!]$, sont des projecteurs.
Reste à montrer que ce sont bien ceux dont traite l'énoncé du lemme !
Montrons que : $\forall i \in [\![1, s]\!], N_i = \operatorname{Im} p_i$.

$\subset$ **:** Soit $y = p_i(x) \in \operatorname{Im} p_i$.
On a : $\left(M_i^{\alpha_i}(u)\right)(y) = \left(M_i^{\alpha_i} P_i\right)(u)(x) = 0$, car $F | M_i^{\alpha_i} P_i$.
Et donc $x \in \operatorname{Ker} M_i^{\alpha_i} = N_i$.

$\supset$ **:** Soit $x \in N_i$.
D'après (1) : $x = \sum_{j=1}^{s} p_j(x)$. Or pour $j \neq i$, $p_j(x) = P_j(u)(x) = 0$ car $M_i^{\alpha_i} | P_j$.
Et donc $x = p_i(x) \in \operatorname{Im} p_i$.

Diffusion à titre gratuit uniquement.

Montrons que : $\forall i \in [\![1,s]\!], \text{Ker } p_i = \bigoplus_{j\neq i} N_j$.

$\subset$ **:** Soit $x \in \text{Ker } p_i$.

Par (1), on a donc : $x = \sum_{j=1}^{s} p_j(x) \in \bigoplus_{j\neq i} \text{Im } p_j = \bigoplus_{j\neq i} N_j$, vu que $p_i(x) = 0$.

$\supset$ **:** Soit $j \neq i$ et $x \in N_j$; alors $x = p_j(t)$ avec $t \in E$.
Ainsi, d'après (2) : $p_i(x) = p_i \circ p_j(t) = 0$.
Donc $\forall j \neq i, N_j \subset \text{Ker } p_i$ donc $\bigoplus_{j\neq i} N_j \subset \text{Ker } p_i$. ■

Démonstration du théorème :

Existence : On écrit $\chi_u = \prod_{i=1}^{s} (X-\lambda_i)^{\alpha_i}$ et pour $i \in [\![1,s]\!]$, on note $N_i = \text{Ker}\left((u-\lambda_i \text{Id}_E)^{\alpha_i}\right)$.
Grâce au théorème de Cayley-Hamilton, on aurait pu rédiger le lemme précédent avec χ_u à la place de F ; on reprend alors les notations introduites au cours de la démonstration du lemme.
On pose alors $d := \sum_{i=1}^{s} \lambda_i p_i$; d est diagonalisable [11] et $d \in \mathbb{K}[u]$.
Puis, $n := u - d = \sum_{i=1}^{s} (u-\lambda_i \text{Id}_E)\, p_i \in \mathbb{K}[u]$. Reste à montrer que n est nilpotent.
On a, pour $q \in \mathbb{N}^*$, $n^q = \sum_{i=1}^{s} (u-\lambda_i \text{Id}_E)^q\, p_i$. [12]
Or $(u-\lambda_i \text{Id}_E)^{\alpha_i}\, p_i = \left((X-\lambda_i)^{\alpha_i} P_i\right)(u) = 0$ car $\chi_u \mid (X-\lambda_i)^{\alpha_i} P_i$.
Notant $q = \max_{1\leqslant i\leqslant s} \alpha_i$, on a donc : $n^q = 0$.

Unicité : Soit $(d', n') \in \mathcal{L}(E)^2$ tel que d' diagonalisable, n' nilpotent, $u = d' + n'$ et $d' \circ n' = n' \circ d'$.
On a donc : $d - d' = n' - n$.
Or d' et n' commutent avec u donc aussi avec d et n (vu qu'ils sont dans $\mathbb{K}[u]$).
Donc d et d' sont simultanément diagonalisables donc $d - d'$ est diagonalisable.
Mais aussi, $n' - n$ est nilpotent donc $d - d' = n' - n = 0$. ■

Références

[Gou Al] X. GOURDON – *Les maths en tête : Algèbre*, 2e éd., Ellipses, 2009.

11. Deux façons de le voir :
   – On prend des bases de $N_1, \ldots, N_s$, qu'on concatène pour obtenir une base de E qu'on appelle $\mathcal{B}$. Alors dans la base $\mathcal{B}$, d a pour matrice $\text{diag}(\underbrace{\lambda_1, \ldots, \lambda_1}_{\dim N_1}, \ldots, \underbrace{\lambda_s, \ldots, \lambda_s}_{\dim N_s})$.
   – Chacun des p_i est diagonalisable, vu que $X^2 - X$ en est un polynôme annulateur scindé à racines simples. Comme ce sont des polynômes en u, ils commutent tous. On utilise alors le théorème de diagonalisation simultanée et on en déduit que d est diagonalisable.

12. Cela se montre aisément par récurrence, en utilisant que les p_i commutent en tant que polynômes en u (et donc on peut utiliser le binôme de Newton), que $p_i^2 = p_i$ et que $p_i \circ p_j = 0$.

Diffusion à titre gratuit uniquement.

Dual de $\mathcal{M}_n(\mathbb{K})$

Leçons : 159

[X-ENS Al1], exercices 7.8, 7.9 et 7.11

Théorème

L'application $f : \left| \begin{array}{ccc} \mathcal{M}_n(\mathbb{K}) & \to & \mathcal{M}_n(\mathbb{K})^* \\ A & \mapsto & f_A : X \mapsto \mathrm{tr}(AX) \end{array} \right.$ est un isomorphisme entre $\mathcal{M}_n(\mathbb{K})$ et son dual.

Démonstration :

On note $\left(E_{i,j}\right)_{1\leqslant i,j\leqslant n}$ la base canonique de $\mathcal{M}_n(\mathbb{K})$.

La linéarité de la trace et la bilinéarité du produit matriciel impliquent la linéarité de f ; par ailleurs, $\mathcal{M}_n(\mathbb{K})$ et $\mathcal{M}_n(\mathbb{K})^*$ sont de même dimension n^2. Il suffit donc de montrer que f est injectif.

Soit $A \in \mathcal{M}_n(\mathbb{K})$ telle que $f_A = 0$.

Alors, pour tous $i, j \in [\![1, n]\!]$, on a :

$$0 = \mathrm{tr}\left(AE_{i,j}\right) = \sum_{k=1}^{n} \left(AE_{i,j}\right)_{k,k} = \sum_{k=1}^{n}\sum_{l=1}^{n} a_{k,l}\delta_{i,l}\delta_{j,k} = a_{j,i}$$

Finalement, $A = 0$, f est injectif : c'est un isomorphisme. ■

Corollaire (*Caractérisation de la trace*)

Soit $g \in \mathcal{M}_n(\mathbb{K})^*$ vérifiant $\forall (X, Y) \in \mathcal{M}_n(\mathbb{K})^2, g(XY) = g(YX)$.
Alors g est proportionnel à la trace : $\exists \lambda \in \mathbb{K}, \forall X \in \mathcal{M}_n(\mathbb{K}), g(X) = \lambda \mathrm{tr}(X)$.

Démonstration :

D'après le théorème précédent, $\exists A \in \mathcal{M}_n(\mathbb{K}), \forall X \in \mathcal{M}_n(\mathbb{K}), g(X) = \mathrm{tr}(AX)$.

L'hypothèse sur g nous fournit donc : $\forall (X, Y) \in \mathcal{M}_n(\mathbb{K})^2, \mathrm{tr}(AXY) = \mathrm{tr}(AYX) = \mathrm{tr}(XAY)$.

On en déduit alors : $\forall X \in \mathcal{M}_n(\mathbb{K}), \forall Y \in \mathcal{M}_n(\mathbb{K}), \mathrm{tr}((AX - XA)Y) = 0$.

L'isomorphisme précédent nous donne alors : $\forall X \in \mathcal{M}_n(\mathbb{K}), AX - XA = 0$, c'est-à-dire : $A \in \mathrm{Z}\left(\mathcal{M}_n(\mathbb{K})\right)$.

En conséquence, A est une homothétie. [13]

■

Corollaire

Si $n \geqslant 2$,
Alors tout hyperplan de $\mathcal{M}_n(\mathbb{K})$ rencontre $\mathrm{GL}_n(\mathbb{K})$.

Démonstration :

Soit H un hyperplan de $\mathcal{M}_n(\mathbb{K})$, et soit φ une forme linéaire de noyau H.

Il existe donc $A \in \mathcal{M}_n(\mathbb{K})$, telle que $\forall X \in \mathcal{M}_n(\mathbb{K}), \varphi(X) = \mathrm{tr}(AX)$.

On cherche donc une matrice $X \in \mathrm{GL}_n(\mathbb{K})$, telle que $\mathrm{tr}(AX) = 0$.

Pour simplifier, on note r le rang de A et A est équivalente à la matrice $J_r = \left(\begin{array}{c|c} \mathrm{I}_r & 0 \\ \hline 0 & 0 \end{array}\right) \in \mathcal{M}_n(\mathbb{K})$, c'est-à-dire : $\exists P, Q \in \mathrm{GL}_n(\mathbb{K}), A = PJ_rQ$.

On a donc, pour tout $X \in \mathcal{M}_n(\mathbb{K})$: $\mathrm{tr}(AX) = \mathrm{tr}\left(PJ_rQX\right) = \mathrm{tr}\left(J_rQXP\right)$.

Si on trouve $Y \in \mathrm{GL}_n(\mathbb{K})$ telle que $\mathrm{tr}\left(J_rY\right) = 0$, on a gagné : on posera $X = Q^{-1}YP^{-1}$ qui sera à la fois dans $\mathrm{GL}_n(\mathbb{K})$ et dans H.

Par exemple, on peut poser $Y = \begin{pmatrix} 0 & \dots & \dots & 0 & 1 \\ 1 & \ddots & & & 0 \\ 0 & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & 1 & 0 \end{pmatrix} \in \mathcal{M}_n(\mathbb{K})$.

En effet, Y est inversible, car de déterminant $(-1)^{n+1}$ et J_rY est de trace nulle (car de diagonale nulle). ■

13. Pour le montrer remplacer X par une matrice $E_{i,j}$; la réciproque est une trivialité.

Diffusion à titre gratuit uniquement.

Corollaire

Soit $(A, B) \in \mathcal{M}_n(\mathbb{K})^2$. Alors s'équivalent :

1. $\exists X \in \mathcal{M}_n(\mathbb{K}), AX + XA = B$
2. $\forall C \in \mathcal{M}_n(\mathbb{K}), AC + CA = 0 \Rightarrow \text{tr}(BC) = 0$

Démonstration :

Soit $h : \left| \begin{array}{ccc} \mathcal{M}_n(\mathbb{K}) & \to & \mathcal{M}_n(\mathbb{K}) \\ X & \mapsto & AX + XA \end{array} \right.$.

Alors 1. $\Leftrightarrow B \in \text{Im } h$ et 2. $\Leftrightarrow \forall C \in \text{Ker } h, f_C(B) = 0 \Leftrightarrow B \in f(\text{Ker } h)^\circ$.
Mais $\dim f(\text{Ker } h)^\circ = n^2 - \dim f(\text{Ker } h) = n^2 - \dim \text{Ker } h = \dim \text{Im } h$.
Il suffit donc de montrer que $\text{Im } h \subset f(\text{Ker } h)^\circ$.
Soit $D \in \text{Im } h$, disons $D = AY + YA$, avec $Y \in \mathcal{M}_n(\mathbb{K})$.
Alors $\forall C \in \text{Ker } h$,

$$f_C(D) = \text{tr}(CD) = \text{tr}(C(AY + YA)) = \text{tr}(CAY) + \text{tr}(CYA) = \text{tr}(CAY) + \text{tr}(ACY) = \text{tr}((CA + AC)Y) = 0.$$

Donc $D \in f(\text{Ker } h)^\circ$.
D'où $f(\text{Ker } h)^\circ = \text{Im } h$ et 1. $\Leftrightarrow$ 2.. ■

Références

[X-ENS Al1] S. FRANCINOU, H. GIANELLA et S. NICOLAS – *Oraux X-ENS Algèbre 1*, 2e éd., Cassini, 2007.

Diffusion à titre gratuit uniquement.

Ellipse de Steiner

Leçons : 180, 181, 182

[CAPES11], d'après le sujet 4
[Dbm1], proposition 5.II.5

Théorème

Soient $M_1\,(r_1)$, $M_2\,(r_2)$ et $M_3\,(r_3)$ trois points distincts non-alignés du plan affine $\mathcal{P}$.
On note $P = (X - r_1)\,(X - r_2)\,(X - r_3)$ et ω, ω' les zéros de P'.
Alors les points $F(\omega)$ et $F'(\omega')$ sont les foyers d'une ellipse tangente aux trois côtés du triangle $M_1M_2M_3$, en leurs milieux (qu'on notera A, B et C).

Démonstration :

On va d'abord énoncer et montrer deux lemmes.

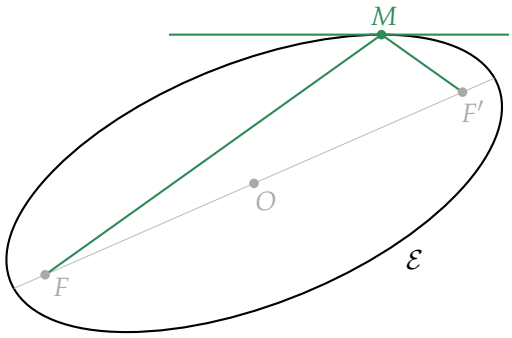
Lemme 1

Soient $\mathcal{E}$ une ellipse non-plate et $M \in \mathcal{E}$.
On note F et F' les foyers de $\mathcal{E}$.
Alors la tangente à $\mathcal{E}$ en M est la bissectrice extérieure de $\widehat{FMF'}$.

Démonstration du lemme 1 :

Si O est le centre de l'ellipse, on a la paramétrisation $\overrightarrow{OM(t)} = a\cos t\,\overrightarrow{e_1} + b\sin t\,\overrightarrow{e_2}$ quand $M(t)$ parcourt $\mathcal{E}$.
On dérive l'expression $\left\|\overrightarrow{M(t)F}\right\| + \left\|\overrightarrow{M(t)F'}\right\| = 2a$ et on obtient :

$$\left\langle \frac{\overrightarrow{M(t)F}}{\left\|\overrightarrow{M(t)F}\right\|}, \frac{\mathrm{d}\overrightarrow{M(t)}}{\mathrm{d}t} \right\rangle + \left\langle \frac{\overrightarrow{M(t)F'}}{\left\|\overrightarrow{M(t)F'}\right\|}, \frac{\mathrm{d}\overrightarrow{M(t)}}{\mathrm{d}t} \right\rangle = 0.$$



Ceci se raccourcit en : $\left\langle \frac{\overrightarrow{M(t)F}}{\left\|\overrightarrow{M(t)F}\right\|} + \frac{\overrightarrow{M(t)F'}}{\left\|\overrightarrow{M(t)F'}\right\|}, \frac{\mathrm{d}\overrightarrow{M(t)}}{\mathrm{d}t} \right\rangle = 0$. Mais le premier membre du produit scalaire est le vecteur directeur de la bissectrice intérieure de $\widehat{FMF'}$!
En conséquence : la bissectrice $\left|\begin{array}{l}\text{intérieure}\\ \text{extérieure}\end{array}\right.$ est la $\left|\begin{array}{l}\text{normale}\\ \text{tangente}\end{array}\right.$ à $\mathcal{E}$ en M. ■

Lemme 2 (*Poncelet* [14])

Soit $\mathcal{E}$ une ellipse non-plate de foyers F et F'.
Soit P un point extérieur à $\mathcal{E}$, par lequel passent deux tangentes à $\mathcal{E}$, aux points notés T_1 et T_2.
Alors on a : $\left(\overrightarrow{PT_1}, \overrightarrow{PF}\right) \equiv \left(\overrightarrow{PF'}, \overrightarrow{PT_2}\right)\ [\pi]$.

Démonstration du lemme 2 :

Si Δ désigne une droite de $\mathcal{P}$, on note σ_Δ la symétrie axiale d'axe Δ.
Ainsi, $\sigma_{(PF)} \circ \sigma_{(PT_1)}$ et $\sigma_{(PT_2)} \circ \sigma_{(PF')}$ sont des rotations de centre de P ; notre but va être de montrer qu'elles sont égales.
On désigne par N_1 et N_2 les symétriques de F par les symétries axiales d'axes (PT_1) et (PT_2).
D'une part, $\sigma_{(PF)} \circ \sigma_{(PT_1)}\,(N_1) = \sigma_{(PF)}(F) = F$.
D'autre part, comme (PT_1) est tangente à l'ellipse en T_1, c'est la bissectrice extérieure de $\widehat{FT_1F'}$, donc F', T_1 et N_1 sont alignés dans cet ordre.
De même, comme (PT_2) est tangente à l'ellipse en T_2, c'est la bissectrice extérieure de $\widehat{FT_2F'}$, donc F', T_2 et N_2 sont alignés dans cet ordre.

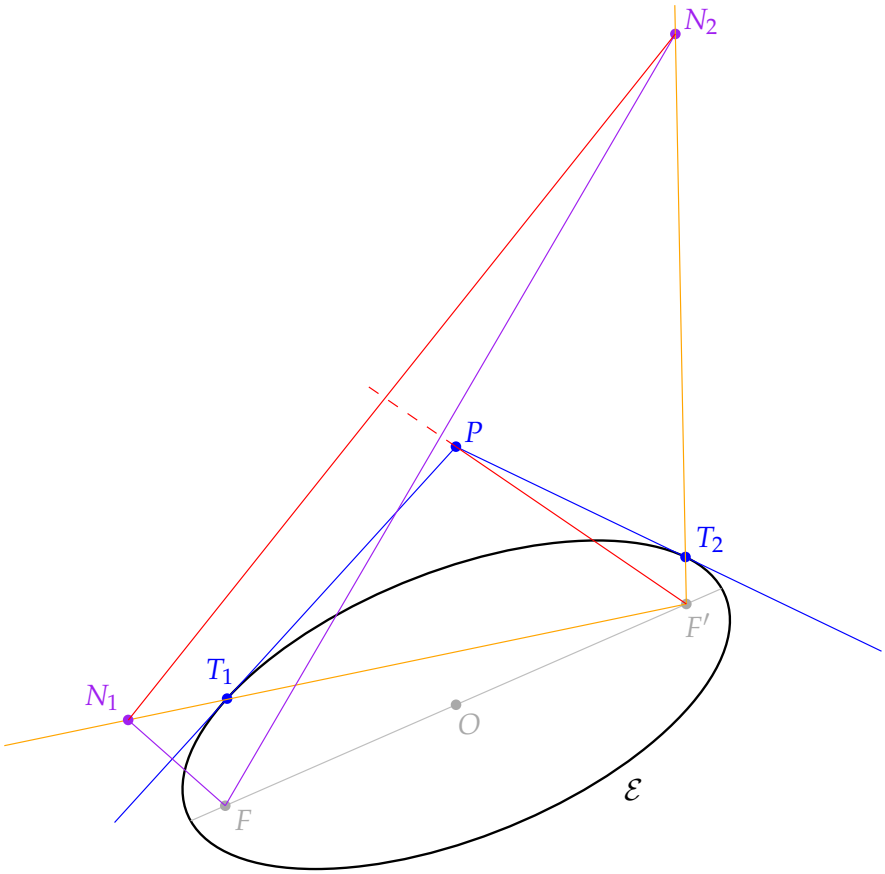
14. On trouvera une référence pour la démonstration de ce lemme avec ce document de Florian BOUGUET.

Diffusion à titre gratuit uniquement.

Ainsi, $F'N_1 = F'T_1 + T_1N_1 = F'T_1 + FT_1 = 2a = F'T_2 + FT_2 = F'T_2 + T_2N_2 = F'N_2$.
Aussi, on a : $PN_1 = PF = PN_2$, donc (PF') est la médiatrice de $[N_1N_2]$.
Ainsi, $\sigma_{(PT_2)} \circ \sigma_{(PF')}(N_1) = \sigma_{(PT_2)}(N_2) = F$.
En conséquence, $\sigma_{(PF)} \circ \sigma_{(PT_1)} = \sigma_{(PT_2)} \circ \sigma_{(PF')}$, puis, en regardant les angles de ces rotations :

$$\left(\overrightarrow{PT_1}, \overrightarrow{PF}\right) \equiv \left(\overrightarrow{PF'}, \overrightarrow{PT_2}\right) \ [\pi].$$

■



Démontrons désormais le théorème.

1. On va commencer par vouloir traiter le cas où F et F' sont confondus.
   On a : $P' = 3X^2 - 2(r_1 + r_2 + r_3)X + (r_1r_2 + r_2r_3 + r_3r_1)$.
   Ainsi, on a les équivalences :

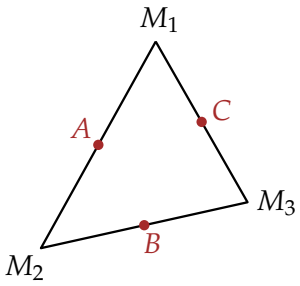
$$\begin{aligned}
P' \text{ a une racine double} &\Leftrightarrow (r_1 + r_2 + r_3)^2 - 3(r_1r_2 + r_2r_3 + r_3r_1) = 0 \\
&\Leftrightarrow \frac{r_1 - r_2}{r_3 - r_2} = \frac{r_2 - r_3}{r_1 - r_3} \\
&\Leftrightarrow \begin{cases} M_1M_2.M_1M_3 = (M_2M_3)^2 \\ \left(\overrightarrow{M_2M_3}, \overrightarrow{M_2M_1}\right) \equiv \left(\overrightarrow{M_3M_1}, \overrightarrow{M_3M_2}\right) \ [2\pi] \end{cases} \\
&\Leftrightarrow \begin{cases} M_1M_2.M_1M_3 = (M_2M_3)^2 \\ M_1M_2 = M_1M_3 \text{ (isocélisme en } M_1) \end{cases} \\
&\Leftrightarrow M_1M_2 = M_2M_3 = M_3M_1 \\
&\Leftrightarrow M_1M_2M_3 \text{ est équilatéral}
\end{aligned}$$

   Mais le cas du triangle équilatéral est trivial : le cercle inscrit répond à notre problème (il est tangent aux 3 côtés en leurs milieux car les médianes se confondent avec les bissectrices).

2. Désormais, on suppose que $M_1M_2M_3$ n'est pas équilatéral, en conséquence, $F \neq F'$.
   On va montrer que $\mathcal{E} = \{M \in \mathcal{P} | MF + MF' = AF + AF'\}$ répond à notre problème.

Diffusion à titre gratuit uniquement.

→ On a bien $A \in \mathcal{E}$, mais $\mathcal{E}$ est-elle une vraie ellipse, c'est-à-dire non-plate ?
Pour cela, il suffit de montrer que $A \notin [FF']$.
Par l'absurde, on suppose que $A \in [FF']$.
Alors A est barycentre à coefficients positifs de F et F'.
Mais, par le théorème de Lucas, F et F' sont barycentres à coefficients positifs de M_1, M_2 et M_3 (car P et P' n'ont pas de racine commune).
Par associativité du barycentre, A est dans l'intérieur strict de $M_1M_2M_3$. Contradiction !



→ On va montrer que (M_1M_2) est tangente à $\mathcal{E}$.
On note a l'affixe de A, et on calcule $P'(a) = 3(a-\omega)(a-\omega') = (a-r_1)(a-r_2) + (a-r_3)\underbrace{(2a-r_1-r_2)}_{=0}$.
Ainsi $3(\omega-a)(\omega'-a) = \dfrac{r_2-r_1}{2}\dfrac{r_1-r_2}{2}$, puis $12\dfrac{\omega-a}{r_1-r_2} = \dfrac{r_2-r_1}{\omega'-a}$.
En passant aux arguments, il vient : $\left(\overrightarrow{M_2M_1}, \overrightarrow{AF}\right) \equiv \left(\overrightarrow{AF'}, \overrightarrow{M_1M_2}\right)\ [2\pi]$.
Ainsi, $\left(\overrightarrow{M_1M_2}, \overrightarrow{AF}\right) \equiv \left(\overrightarrow{AF'}, \overrightarrow{M_1M_2}\right)\ [\pi]$ donc (M_1M_2) est une bissectrice des droites (AF) et (AF'), nécessairement extérieure à AFF', car sinon (M_1M_2) couperait $[FF']$.
D'après le lemme 1, (M_1M_2) est tangente à $\mathcal{E}$, en A.

→ On va montrer que (M_1M_3) est tangente à $\mathcal{E}$.[15]
On a : $P'(r_1) = 3(r_1-\omega)(r_1-\omega') = (r_1-r_2)(r_1-r_3)$.
Ainsi, $3\dfrac{\omega'-r_1}{r_3-r_1} = \dfrac{r_2-r_1}{\omega-r_1}$, donc $\left(\overrightarrow{M_1M_3}, \overrightarrow{M_1F'}\right) \equiv \left(\overrightarrow{M_1F}, \overrightarrow{M_1M_2}\right)\ [2\pi]$.
On note (M_1T) l'autre tangente à $\mathcal{E}$ issue de M_1.[16]
Le lemme de Poncelet nous donne : $\left(\overrightarrow{M_1M_2}, \overrightarrow{M_1F}\right) \equiv \left(\overrightarrow{M_1F'}, \overrightarrow{M_1T}\right)\ [\pi]$.
Donc $\left(\overrightarrow{M_1F'}, \overrightarrow{M_1M_3}\right) \equiv \left(\overrightarrow{M_1M_2}, \overrightarrow{M_1F}\right) \equiv \left(\overrightarrow{M_1F'}, \overrightarrow{M_1T}\right)\ [\pi]$.
En conséquence, $(M_1M_3) = (M_1T)$ est tangente à $\mathcal{E}$.

→ Enfin, montrons que $C \in \mathcal{E}$.[17]
On note C' le point de tangence de (M_1M_3) et $\mathcal{E}$.
Par le lemme 1, (M_1M_3) est bissectrice extérieure de $\widehat{FC'F'}$ en C'.
Donc G, symétrique de F par rapport à (M_1M_3), vérifie : $G \in (F'C')$, autrement dit $C' \in (F'G)$.
Ainsi, $C' \in (F'G) \cap (M_1M_3)$.
Pour montrer que $C = C'$, on va montrer que $C \in (F'G)$.
En évaluant P' en c, comme on l'avait fait en a précédemment, on obtient que (M_1M_3) est la bissectrice extérieure de l'angle $\widehat{FCF'}$.
Et donc aussi $C \in (F'G)$. Ayé. ■

Références

[CAPES11] D.-J. MERCIER et J.-É. ROMBALDI – *Annales du CAPES externe de mathématiques (2009 à 2011)*, Publibook, 2011.
[Dbm1] G. DEBEAUMARCHÉ – *Manuel de mathématiques (volume 1)*, Ellipses, 2004.

15. On montrerait similairement que (M_2M_3) est tangente à $\mathcal{E}$.
16. Comprenez "autre que (M_1M_2)".
17. On montrerait similairement que $B \in \mathcal{E}$.

Diffusion à titre gratuit uniquement.

Étude de l'anneau $\mathbb{Z}\left[\frac{1+i\sqrt{19}}{2}\right]$

Leçons : 122

[Per], partie II.5

Théorème

On note $\alpha = \dfrac{1+i\sqrt{19}}{2}$ et $A = \mathbb{Z}[\alpha]$.
A est un anneau principal, non-euclidien.

Démonstration :

Étape 1 : α est racine de $P = T^2 - T + 5$, car $\alpha + \overline{\alpha} = 1$ et $\alpha\overline{\alpha} = 5$.
Ainsi, $A = \left\{a + b\alpha \middle| (a,b) \in \mathbb{Z}^2\right\}$ est un sous-anneau de $\mathbb{C}$.[18]
Donc A est intègre ; et comme $\overline{\alpha} = 1 - \alpha$, A est stable par conjugaison.
Pour $z = a + b\alpha \in A$, on définit la norme :

$$N(z) = z\overline{z} = (a+b\alpha)(a+b\overline{\alpha}) = a^2 + ab(\alpha + \overline{\alpha}) + b^2\alpha\overline{\alpha} = a^2 + ab + 5b^2.$$

Alors $N(z) \in \mathbb{N}$, et $N(zz') = N(z)N(z')$.
De plus, $N(z) = 0 \Rightarrow \left(a + \dfrac{b}{2}\right)^2 + \dfrac{19}{4}b^2 = 0 \Rightarrow a = b = 0 \Rightarrow z = 0$.
Soit $z \in A^\times$, alors $N(z)N\left(z^{-1}\right) = 1$ donc $N(z) = 1$.
Alors $\left(a + \dfrac{b}{2}\right)^2 + \underbrace{\dfrac{19}{4}}_{>1} b^2 = 1$, donc $b = 0$ et $a = \pm 1$. Ainsi, $A^\times = \{\pm 1\}$.

Étape 2 : Supposons A euclidien, alors $\exists x \in A\backslash A^\times, \pi_{A/(x)}\big|_{A^\times \cup \{0\}}$ est surjective.[19]
En particulier, $A/(x)$ est un corps et $\#A/(x) \leqslant 3$, donc $A/(x) = K$, où $K \simeq \mathbb{F}_2$ ou $\mathbb{F}_3$.
On en déduit l'existence d'un morphisme d'anneaux surjectif $\varphi : A \to K$.
Alors $\beta = \varphi(\alpha)$ vérifie $\beta^2 - \beta + 5 = 0$.
Mais cette équation ne possède de solution ni dans $\mathbb{F}_2$, ni dans $\mathbb{F}_3$.[20]
On aboutit à une contradiction, et A n'est donc pas euclidien.

Étape 3 : On introduit une "pseudo-division euclidienne".

Lemme

Soient $a, b \in A\backslash\{0\}$.
Alors il existe $(q,r) \in A^2$, tels que :

1. $N(r) < N(b)$;
2. $a = bq + r$ ou $2a = bq + r$.

Démonstration :
Soit $x = \dfrac{a}{b} = \dfrac{a\overline{b}}{N(b)} \in \mathbb{C}$, qu'on écrit aussi $x = u + v\alpha$, où $u, v \in \mathbb{Q}$. On note $n = \lfloor v \rfloor$.

– Supposons que $v \notin \left]n + \dfrac{1}{3}, n + \dfrac{2}{3}\right[$; soient s et t les plus proches entiers de u et v.
Ainsi, $|s-u| \leqslant \dfrac{1}{2}$ et $|t-v| \leqslant \dfrac{1}{3}$.
On pose $q = s + t\alpha \in A$ et :

$$(x-q)\overline{(x-q)} = (s-u)^2 + (s-u)(t-v) + 5(t-v)^2 \leqslant \frac{1}{4} + \frac{1}{6} + \frac{5}{9} = \frac{9+6+20}{36} = \frac{35}{36} < 1.$$

On pose $r = a - bq = b(x-q)$ et on a $N(r) < N(b)$.

18. Car A est un sous-groupe de $\mathbb{C}$, contient 1 et est stable par multiplication.
19. La démonstration est dans le rappel sur les anneaux, en page 136.
20. Cela se démontre facilement en cherchant de façon exhaustive.

Diffusion à titre gratuit uniquement.

– Supposons désormais que $v \in \left]n+\frac{1}{3}, n+\frac{2}{3}\right[$, alors $2x = 2u + 2v\alpha$ et $2v \in \left]2n+\frac{2}{3}, 2n+1+\frac{1}{3}\right[$ et on est ramené au cas précédent : on peut écrire $2a = bq + r$, avec $N(r) < N(b)$. ■

Étape 4 : Montrons que A est principal.

On a : $A \simeq {}^{\mathbb{Z}[T]}\!/_{(P)}$, donc ${}^{A}\!/_{(2)} \underset{21}{\simeq} {}^{\mathbb{Z}[T]}\!/_{(2,P)} \underset{22}{\simeq} {}^{\mathbb{F}_2[T]}\!/_{(P)}$.

Mais $T^2 - T + 5$ est irréductible sur $\mathbb{F}_2$ car de degré 2 sans racine ; donc ${}^{A}\!/_{(2)}$ est un corps et (2) est maximal dans A.
Soit $I \neq (0)$ un idéal de A, et soit $a \in I\backslash\{0\}$ de norme $N(a)$ minimale.
Soit $x \in I\backslash(a)$;

→ Si $x = aq + r$ avec $N(r) < N(a)$ ou $r = 0$, alors comme $r \in I$, par minimalité de $N(a)$, on a $r = 0$.
Ainsi $x \in (a)$: contradiction.

→ Ainsi, $2x = aq + r$, et même $2x = aq$ en répétant le procédé qu'on vient à peine de faire.
Comme (2) est maximal, l'idéal (2) est premier, d'où $a \in (2)$ ou $q \in (2)$.
Si $q \in (2)$, alors $q = 2q'$ et $x = aq'$ donc $x \in (a)$. Contradiction.
Donc $a \in (2)$, c'est-à-dire : $a = 2a'$.
Comme $q \notin (2)$ et (2) est maximal, on a : $(2, q) = A$, donc $\exists \lambda, \mu \in A, 2\lambda + q\mu = 1$.
Donc $a' = 2\lambda a' + q\mu a' = \lambda a + \mu x \in I$.
Or $0 < N(a') < N(a)$. Contradiction.

Ainsi, $I = (a)$ et A est principal. ■

Références

[Per] D. PERRIN – *Cours d'algèbre*, Ellipses, 1996.

21. Notons $\pi_P : \mathbb{Z}[T] \to {}^{\mathbb{Z}[T]}\!/_{(P)}$ et $\pi_{\overline{2}} : {}^{\mathbb{Z}[T]}\!/_{(P)} \to \left({}^{\mathbb{Z}[T]}\!/_{(P)}\right)\!/_{(\overline{2})}$ les projections canoniques.
Ker $\pi_{\overline{2}} \circ \pi_P = \left\{f \in \mathbb{Z}[T] \middle| \exists u \in \mathbb{Z}[T], \overline{f} = \overline{2u}\right\} = \{f \in \mathbb{Z}[T] | \exists u, v \in \mathbb{Z}[T], f = 2u + Pv\} = (2, P)$.
Ainsi $\pi_{\overline{2}} \circ \pi_P$ induit un isomorphisme ${}^{\mathbb{Z}[T]}\!/_{(2,P)} \simeq \left({}^{\mathbb{Z}[T]}\!/_{(P)}\right)\!/_{(\overline{2})} \simeq {}^{A}\!/_{(2)}$.

22. Notons $\pi_2 : \mathbb{Z}[T] \to {}^{\mathbb{Z}[T]}\!/_{(2)}$ et $\pi_{\overline{P}} : {}^{\mathbb{Z}[T]}\!/_{(2)} \to \left({}^{\mathbb{Z}[T]}\!/_{(2)}\right)\!/_{(\overline{P})}$ les projections canoniques.
Ker $\pi_{\overline{P}} \circ \pi_2 = \left\{f \in \mathbb{Z}[T] \middle| \exists u \in \mathbb{Z}[T], \overline{f} = \overline{Pu}\right\} = \{f \in \mathbb{Z}[T] | \exists u, v \in \mathbb{Z}[T], f = Pu + 2v\} = (2, P)$.
Ainsi $\pi_{\overline{P}} \circ \pi_2$ induit un isomorphisme ${}^{\mathbb{Z}[T]}\!/_{(2,P)} \simeq \left({}^{\mathbb{Z}[T]}\!/_{(2)}\right)\!/_{(\overline{P})} \simeq {}^{\mathbb{F}_2[T]}\!/_{(P)}$.

Diffusion à titre gratuit uniquement.

Étude du groupe $\mathrm{O}(p,q)$

Leçons : 156, 158, *106, 170, 171*

[H2G2], partie VI.2

Notations : pour $p,q \in \mathbb{N}$, $\mathrm{O}(p,q)$ désigne le sous-groupe de $\mathrm{GL}_{p+q}(\mathbb{R})$ formé des isométries de la forme quadratique standard sur $\mathbb{R}^{p+q}$, de signature (p,q), c'est-à-dire $x_1^2 + \ldots x_p^2 - x_{p+1}^2 - \ldots - x_{p+q}^2$, dont on note $\mathrm{I}_{p,q}$ la matrice dans la base canonique. On notera également $\mathrm{O}(p)$ le groupe $\mathrm{O}(p,\mathbb{R})$.

Théorème

Soient $p,q \in \mathbb{N}^*$.
Il existe un homéomorphisme : $\mathrm{O}(p,q) \simeq \mathrm{O}(p) \times \mathrm{O}(q) \times \mathbb{R}^{pq}$.

On commence par montrer le lemme qui suit.

Lemme

L'application $\exp : \mathcal{S}_n(\mathbb{R}) \to \mathcal{S}_n^{++}(\mathbb{R})$ est un homéomorphisme.

Démonstration du lemme :

→ Soit $S \in \mathcal{S}_n(\mathbb{R})$; par théorème spectral, $S = P\mathrm{diag}\left(\lambda_1,\ldots,\lambda_n\right)P^{-1}$, où $P \in \mathrm{O}(n)$ et $\forall i \in [\![1,n]\!], \lambda_i \in \mathbb{R}$.
Alors $\exp(S) = P\mathrm{diag}\left(\mathrm{e}^{\lambda_1},\ldots,\mathrm{e}^{\lambda_n}\right)P^{-1} = P\mathrm{diag}\left(\mathrm{e}^{\lambda_1},\ldots,\mathrm{e}^{\lambda_n}\right){}^{\mathrm{t}}P \in \mathcal{S}_n^{++}(\mathbb{R})$.
Et par restriction, $\exp : \mathcal{S}_n(\mathbb{R}) \to \mathcal{S}_n^{++}(\mathbb{R})$ est continue.

→ Pour la surjectivité, soit $B \in \mathcal{S}_n^{++}(\mathbb{R})$,
$B = P\mathrm{diag}\left(\lambda_1,\ldots,\lambda_n\right)P^{-1} = \exp\underbrace{\left(P\mathrm{diag}\left(\ln\lambda_1,\ldots,\ln\lambda_n\right)P^{-1}\right)}_{\in\mathcal{S}_n(\mathbb{R})}$, où $P \in \mathrm{O}(n)$ et $\forall i \in [\![1,n]\!], \lambda_i > 0$.

→ L'injectivité de $\exp : \mathcal{S}_n(\mathbb{R}) \to \mathcal{S}_n^{++}(\mathbb{R})$ découle de son injectivité sur $\mathcal{D}_n(\mathbb{R})$. [23]

→ Reste à montrer la bicontinuité ; soit $\left(B_p\right)_p = \left(\exp A_p\right)_p$ une suite de $\mathcal{S}_n^{++}(\mathbb{R})$ qui converge vers $B = \exp A \in \mathcal{S}_n^{++}(\mathbb{R})$: montrons que $A_p \underset{p\to\infty}{\longrightarrow} A$.
$\left(B_p\right)_p$ converge donc est bornée pour $|||.|||_2$, et par continuité [24] de l'inverse sur $\mathrm{GL}_n(\mathbb{R})$, la suite $\left(B_p^{-1}\right)_p$ converge (vers B^{-1}), et est également bornée pour $|||.|||_2$.
Or, $\forall M \in \mathcal{S}_n(\mathbb{R}), |||M|||_2 = \sqrt{\rho\left({}^{\mathrm{t}}MM\right)} = \sqrt{\rho\left(M^2\right)} = \rho(M)$. [25]
Donc l'union des spectres des matrices B_p, $p \in \mathbb{N}$, est à la fois majorée par $C \in \mathbb{R}$ et minorée par $C' \in \mathbb{R}$ (puisque les spectres des matrices B_p^{-1} sont eux-mêmes majorés).
En conséquence, $\bigcup_{n\in\mathbb{N}} \mathrm{Sp}\left(B_p\right) \subset [C',C] \subset \mathbb{R}^{+*}$ et $\bigcup_{n\in\mathbb{N}} \mathrm{Sp}\left(A_p\right) \subset [\ln C', \ln C]$ qui est compact.
On a ainsi démontré que la suite $\left(A_p\right)$ est bornée pour $|||.|||_2$.
Soit alors $\left(A_{p_k}\right)_k$ une sous-suite, dont on note $\tilde{A} \in \mathcal{S}_n(\mathbb{R})$ la limite.
Ainsi : $\exp\left(\tilde{A}\right) \underset{k\to\infty}{\longleftarrow} \exp\left(A_{p_k}\right) = B_{p_k} \underset{k\to\infty}{\longrightarrow} B = \exp(A)$.
Puis, par unicité de la limite et par injectivité de l'exponentielle sur $\mathcal{S}_n(\mathbb{R})$, il vient : $A = \tilde{A}$.
La suite $\left(A_p\right)_p$ est donc bornée et ne possède qu'une seule valeur d'adhérence : elle converge. [26] ■

23. D'accord c'est un peu expéditif, mais le développement est long. Je le détaille ici dans le cas de $\mathcal{S}_n(\mathbb{R})$ mais c'est pareil dans $\mathcal{D}_n(\mathbb{R})$. Soient $A, A' \in \mathcal{S}_n(\mathbb{R})$, tels que $\exp(A) = \exp(A')$; on note $\mathrm{e}^{\lambda_1},\ldots,\mathrm{e}^{\lambda_n}$ les valeurs propres de $\exp(A)$ (vu qu'elles sont dans $\mathbb{R}^{+*}$). Soit Q un polynôme interpolateur tel que : $\forall i \in [\![1,n]\!], Q\left(\mathrm{e}^{\lambda_i}\right) = \lambda_i$. Alors, on obtient : $A = Q(\exp A) = Q(\exp A') \in \mathbb{C}[A']$. Ainsi, A et A' commutent, et par diagonalisation simultanée $A = PDP^{-1}$ et $A' = PD'P^{-1}$ avec $P \in \mathrm{GL}_n(\mathbb{R})$ et D, D' deux matrices diagonales. Puis $\exp(A) = \exp(A') \Leftrightarrow \exp(D) = \exp(D') \Leftrightarrow D = D' \Leftrightarrow A = A'$.

24. Cela vient de la continuité du déterminant, et de sa non-annulation sur $\mathrm{GL}_n(\mathbb{R})$.

25. Cela découle de la décomposition polaire.

26. Par l'absurde, soit (u_n) une suite bornée n'admettant qu'une valeur d'adhérence, notée l, et ne convergeant pas vers l. Alors $\exists\varepsilon > 0, \forall N \in \mathbb{N}, \exists n \geqslant N, |u_n - l| > \varepsilon$. On peut extraire une sous-suite $\left(u_{\sigma(n)}\right)$ vérifiant : $\forall n \in \mathbb{N}, \left|u_{\sigma(n)} - l\right| > \varepsilon$. Par Bolzano-Weierstrass, $\left(u_{\sigma(n)}\right)$ étant bornée, on peut réextraire une sous-suite convergente, disons vers l'. Mais $|l' - l| \geqslant \varepsilon$, et (u_n) possède deux valeurs d'adhérence distinctes. Contradiction : (u_n) converge donc vers l.

Diffusion à titre gratuit uniquement.

Démonstration du théorème :

→ Désormais, $n = p+q$; soit $M \in \mathrm{O}(p,q) \subset \mathrm{GL}_n(\mathbb{R})$.
Par décomposition polaire, on peut écrire $M = OS$, avec $O \in \mathrm{O}(n)$ et $S \in \mathcal{S}_n^{++}(\mathbb{R})$.
Notre objectif est de montrer que O et S sont dans $\mathrm{O}(p,q)$; on note $T = {}^tMM = S^2$.
Cependant, $M \in \mathrm{O}(p,q) \Leftrightarrow M\mathrm{I}_{p,q}{}^tM = \mathrm{I}_{p,q} \Leftrightarrow {}^tM^{-1}\mathrm{I}_{p,q}^{-1}M^{-1} = \mathrm{I}_{p,q}^{-1} \Leftrightarrow {}^tM^{-1}\mathrm{I}_{p,q}M^{-1} = \mathrm{I}_{p,q}$

$$\Leftrightarrow {}^tM^{-1} \in \mathrm{O}(p,q) \Leftrightarrow {}^tM \in \mathrm{O}(p,q).$$

Ainsi, $T \in \mathrm{O}(p,q)$. Mais $T \in \mathcal{S}_n^{++}(\mathbb{R})$, donc $\exists U \in \mathcal{S}_n(\mathbb{R}), T = \exp U$.
Mais on a : $T \in \mathrm{O}(p,q) \Leftrightarrow T\mathrm{I}_{p,q}{}^tT = \mathrm{I}_{p,q}$

$$\begin{aligned}
&\Leftrightarrow {}^tT = \mathrm{I}_{p,q}^{-1}T^{-1}\mathrm{I}_{p,q} \\
&\Leftrightarrow \exp\left({}^tU\right) = \mathrm{I}_{p,q}\exp(-U)\mathrm{I}_{p,q}^{-1} \\
&\Leftrightarrow \exp\left({}^tU\right) = \exp\left(-\mathrm{I}_{p,q}U\mathrm{I}_{p,q}^{-1}\right) \\
&\Leftrightarrow \underbrace{{}^tU}_{=U} = -\mathrm{I}_{p,q}U\mathrm{I}_{p,q}^{-1} \qquad \text{(par bijectivité de exp)} \\
&\Leftrightarrow \exp\left(\frac{{}^tU}{2}\right) = \exp\left(-\mathrm{I}_{p,q}\frac{U}{2}\mathrm{I}_{p,q}^{-1}\right) \\
&\Leftrightarrow {}^t\exp\left(\frac{U}{2}\right) = \mathrm{I}_{p,q}\exp\left(\frac{U}{2}\right)^{-1}\mathrm{I}_{p,q}^{-1} \\
&\Leftrightarrow \exp\left(\frac{U}{2}\right) \in \mathrm{O}(p,q)
\end{aligned}$$

Or $\exp\left(\frac{U}{2}\right) \in \mathcal{S}_n(\mathbb{R})$ et $\exp\left(\frac{U}{2}\right)^2 = \exp(U) = T = S^2$. Donc $S = \exp\left(\frac{U}{2}\right) \in \mathrm{O}(p,q)$.[27]
Également, $O = MS^{-1} \in \mathrm{O}(p,q)$, et comme la décomposition polaire est une bijection bicontinue, on a l'homéomorphisme :

$$\mathrm{O}(p,q) \simeq (\mathrm{O}(p,q) \cap \mathrm{O}(n)) \times \left(\mathrm{O}(p,q) \cap \mathcal{S}_n^{++}(\mathbb{R})\right).$$

→ Soit $\left(\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right) \in \mathrm{O}(p,q) \cap \mathrm{O}(n)$.[28]
On a alors :

$$\left(\begin{array}{c|c} \mathrm{I}_p & 0 \\ \hline 0 & -\mathrm{I}_q \end{array}\right) = \left(\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right)\left(\begin{array}{c|c} \mathrm{I}_p & 0 \\ \hline 0 & -\mathrm{I}_q \end{array}\right)\left(\begin{array}{c|c} {}^tA & {}^tC \\ \hline {}^tB & {}^tD \end{array}\right) = \left(\begin{array}{c|c} A{}^tA - B{}^tB & A{}^tC - B{}^tD \\ \hline C{}^tA - D{}^tB & C{}^tC - D{}^tD \end{array}\right)$$

$$\left(\begin{array}{c|c} \mathrm{I}_p & 0 \\ \hline 0 & \mathrm{I}_q \end{array}\right) = \left(\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right)\left(\begin{array}{c|c} {}^tA & {}^tC \\ \hline {}^tB & {}^tD \end{array}\right) = \left(\begin{array}{c|c} A{}^tA + B{}^tB & A{}^tC + B{}^tD \\ \hline C{}^tA + D{}^tB & C{}^tC + D{}^tD \end{array}\right)$$

Ainsi, $B{}^tB = 0$, donc $\sum_{i,j} b_{i,j}^2 = \mathrm{tr}\left(B{}^tB\right) = 0$, donc $B = 0$.
De même, $C = 0$, puis $A \in \mathrm{O}(p)$ et $D \in \mathrm{O}(q)$.
Conséquemment, $\mathrm{O}(p,q) \cap \mathrm{O}(n) = \left\{ \left(\begin{array}{c|c} A & 0 \\ \hline 0 & D \end{array}\right) \middle| A \in \mathrm{O}(p), D \in \mathrm{O}(q) \right\} \simeq \mathrm{O}(p) \times \mathrm{O}(q)$.

→ Définissons l'ensemble $L = \left\{U \in \mathcal{M}_n(\mathbb{R}) \middle| U\mathrm{I}_{p,q} + \mathrm{I}_{p,q}U = 0\right\}$.
Alors exp : $L \cap \mathcal{S}_n(\mathbb{R}) \to \mathrm{O}(p,q) \cap \mathcal{S}_n^{++}(\mathbb{R})$ est un homéomorphisme :
– Si $U \in L \cap \mathcal{S}_n(\mathbb{R})$, alors $U = {}^tU = -\mathrm{I}_{p,q}U\mathrm{I}_{p,q}^{-1}$, donc $\exp(U) \in \mathrm{O}(p,q)$.
Mais aussi, $\exp(U) \in \mathcal{S}_n^{++}(\mathbb{R})$. Cette application est donc bien définie.
– L'injectivité découle de celle vue dans le lemme.
– Pour la surjectivité : soit $T \in \mathrm{O}(p,q) \cap \mathcal{S}_n^{++}(\mathbb{R})$, alors $\exists U \in \mathcal{S}_n(\mathbb{R}), T = \exp(U)$.
Et comme $U \in \mathcal{S}_n(\mathbb{R})$, $T \in \mathrm{O}(p,q) \Rightarrow {}^tU = -\mathrm{I}_{p,q}U\mathrm{I}_{p,q}^{-1} \Rightarrow U \in L$.
Et donc $\exists U \in L \cap \mathcal{S}_n(\mathbb{R}), T = \exp(U)$.
– La bicontinuité découle de celle vue dans le lemme.
Mais $L \cap \mathcal{S}_n(\mathbb{R})$ a la particularité d'être un $\mathbb{R}$-espace vectoriel, cherchons sa dimension !
On note $U = \left(\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right) \in \mathcal{M}_n(\mathbb{R})$; on a : $U \in \mathcal{S}_n(\mathbb{R}) \Leftrightarrow {}^tA = A, {}^tD = D, {}^tB = C$.

27. C'est l'unicité de la racine carrée matricielle dans $\mathcal{S}_n(\mathbb{R})$.
28. Je fais appel au bon sens de tous pour comprendre quelle est la taille de chaque bloc.

Diffusion à titre gratuit uniquement.

Puis $UI_{p,q} + I_{p,q}U = \left(\begin{array}{c|c} 2A & 0 \\ \hline 0 & 2D \end{array}\right)$ et donc $U \in L \Leftrightarrow A = D = 0$.

Donc $L \cap \mathcal{S}_n(\mathbb{R}) = \left\{ \left(\begin{array}{c|c} 0 & {}^tB \\ \hline B & 0 \end{array}\right) \middle| B \in \mathcal{M}_{p,q}(\mathbb{R}) \right\} \simeq \mathcal{M}_{p,q}(\mathbb{R}) \simeq \mathbb{R}^{pq}$.[29]

$\rightarrow$ En fin de compte on a bien démontré que :

$$O(p,q) \simeq O(p) \times O(q) \times \mathbb{R}^{pq}.$$

■

Références

[H2G2] P. CALDERO et J. GERMONI – *Histoires hédonistes de groupes et de géométries*, Calvage & Mounet, 2013.

29. Remarquez que ce n'est pas grave de parler d'isomorphisme ici : un isomorphisme d'espaces vectoriels de dimension finie est un homéomorphisme !

Diffusion à titre gratuit uniquement.

Groupes d'isométries du tétraèdre et du cube

Leçons : 161, 183 [30]**, *101, 104, 105***

[H2G2], partie XII.3

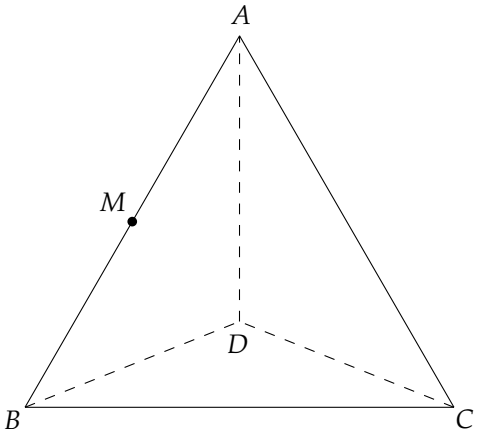
Théorème

On va montrer les résultats suivants :

1. Les groupes d'isométries du tétraèdre régulier Δ_4 sont : $\mathrm{Isom}(\Delta_4) \simeq \mathfrak{S}_4$ et $\mathrm{Isom}^+(\Delta_4) \simeq \mathfrak{A}_4$;
2. Les groupes d'isométries du cube C_6 sont [31] : $\mathrm{Isom}^+(C_6) \simeq \mathfrak{S}_4$ et $\mathrm{Isom}(C_6) \simeq \mathfrak{S}_4 \times \mathbb{Z}/2\mathbb{Z}$.

Démonstration :

1. On fait agir $\mathrm{Isom}(\Delta_4)$ sur l'ensemble des sommets $\mathcal{S} = \{A, B, C, D\}$; on obtient donc un morphisme de groupes $\varphi : \left| \begin{array}{ccc} \mathrm{Isom}(\Delta_4) & \to & \mathfrak{S}(\mathcal{S}) \simeq \mathfrak{S}_4 \\ g & \mapsto & g_{|\mathcal{S}} \end{array} \right.$.

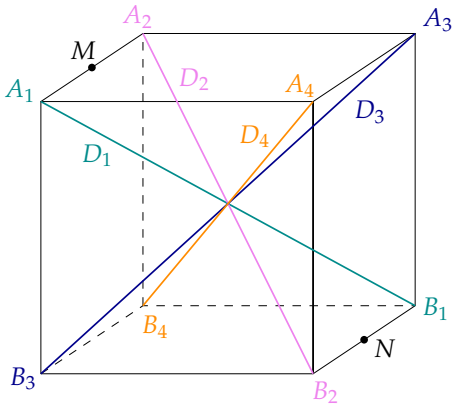


   → φ est injective : si $\varphi(g) = \mathrm{Id}_{\mathcal{S}}$, alors g stabilise $\mathcal{S}$, qui est un repère affine de $\mathbb{R}^3$, d'où $g = \mathrm{Id}_{\mathbb{R}^3}$.

   → φ est surjective : soit M le milieu de $[AB]$, la réflexion par rapport au plan (MCD) réalise la transposition $(A\ B)$. Similairement, toutes les transpositions sont dans $\varphi(\mathrm{Isom}(\Delta_4))$ et elles engendrent $\mathfrak{S}(\mathcal{S})$, donc $\varphi(\mathrm{Isom}(\Delta_4)) = \mathfrak{S}(\mathcal{S})$. En conséquence, φ est un isomorphisme et $\mathrm{Isom}(\Delta_4) \simeq \mathfrak{S}_4$.
   Comme $\mathrm{Isom}^+(\Delta_4)$ est d'indice 2 dans $\mathrm{Isom}(\Delta_4)$, on a : $\mathrm{Isom}^+(\Delta_4) \simeq \mathfrak{A}_4$. [32]

2. Les grandes diagonales du cube, qui relient deux sommets opposés, sont au nombre de 4. Ce sont les plus grandes distances existant entre 2 points du cube, donc les isométries du cube stabilisent l'ensemble des grandes diagonales du cube.
   Pour $i \in [\![1,4]\!]$, on note donc D_i la diagonale (A_iB_i) et $\mathcal{D}$ l'ensemble de ces grandes diagonales. On fait donc agir $\mathrm{Isom}^+(C_6)$ sur $\mathcal{D}$, d'où le morphisme de groupes $\varphi : \left| \begin{array}{ccc} \mathrm{Isom}^+(C_6) & \to & \mathfrak{S}(\mathcal{D}) \simeq \mathfrak{S}_4 \\ g & \mapsto & g_{|\mathcal{D}} \end{array} \right.$.

   → φ est injective ; en effet, soit g tel que $\varphi(g) = \mathrm{Id}_{\mathcal{D}}$.
   Alors pour tout $i \in [\![1,4]\!]$, $g(A_i) = A_i$ et $g(B_i) = B_i$ ou $g(A_i) = B_i$ et $g(B_i) = A_i$.
   Supposons qu'il existe $i \in [\![1,4]\!]$ tel que g fixe A_i et B_i ; et sans perdre en généralité, disons $i = 1$.
   Comme $A_1A_2 \neq A_1B_2$ et que $g \in \mathrm{Isom}(C_6)$, nécessairement, $g(A_2) = A_2$.
   Similairement, $g(A_4) = A_4$ et $g(B_3) = B_3$.
   g fixe donc un repère affine de l'espace, d'où $g = \mathrm{Id}_{\mathbb{R}^3}$.
   Désormais, on suppose que $\forall i \in [\![1,4]\!]$, $g(A_i) = B_i$.
   Notons O le centre du cube, alors, on a : pour tout $i \in [\![1,4]\!]$, $s_O g(A_i) = A_i$, d'où $s_O g = \mathrm{Id}_{\mathbb{R}^3}$.
   Ceci contredit alors la positivité de g et ce cas n'est donc pas possible. φ est donc injective.



   → φ est surjective, car la transposition $(D_1\ D_2)$ est l'image par φ du retournement d'axe (MN), où M et N sont les milieux des segments $[A_1A_2]$ et $[B_1B_2]$.

30. Oui, c'est de la mauvaise foi, et je vous laisse ainsi intact le bonheur de travailler sur la leçon 183.
31. Et on s'en sert dans un autre développement : la table de caractères de $\mathfrak{S}_4$, qu'on peut lire en page 37.
32. Soit H d'indice 2 dans $\mathfrak{S}_n$, alors $\forall g \in \mathfrak{S}_n, \overline{g}^2 = 1$ dans $\mathfrak{S}_n/H$ et donc $g^2 \in H$; autrement dit, H contient tous les carrés d'éléments de $\mathfrak{S}_n$. En particulier, H contient tous les 3-cycles qui engendrent $\mathfrak{A}_n$. Donc, pour des raisons de cardinalité $H = \mathfrak{A}_n$.

Diffusion à titre gratuit uniquement.

Ainsi, $\mathrm{Isom}^+(\mathrm{C}_6) \simeq \mathfrak{S}_4$.
O étant l'isobarycentre de C_6, on a : $\forall g \in \mathrm{Isom}(\mathrm{C}_6), g(O) = O$.
Par conséquent, pour tout $g \in \mathrm{Isom}(\mathrm{C}_6)$, on a : $gs_O = s_O g$, vu que $\mathrm{L}(s_O) = -\mathrm{Id}_{\mathbb{R}^3}$ et que g et s_O ont un point fixe commun. [33]
Ainsi, l'application $F : \left| \begin{array}{ccc} \mathrm{Isom}(\mathrm{C}_6) & \to & \mathrm{Isom}^+(\mathrm{C}_6) \times \mathbb{Z}/2\mathbb{Z} \\ g & \mapsto & \begin{cases} (g,0) & \text{si } g \in \mathrm{Isom}^+(\mathrm{C}_6) \\ (gs_O,1) & \text{sinon} \end{cases} \end{array} \right.$ est un isomorphisme de groupes.
Donc $\mathrm{Isom}(\mathrm{C}_6) \simeq \mathfrak{S}_4 \times \mathbb{Z}/2\mathbb{Z}$. ■

Références

[H2G2] P. CALDERO et J. GERMONI – *Histoires hédonistes de groupes et de géométries*, Calvage & Mounet, 2013.

33. Soient f et g deux applications affines vérifiant $\mathrm{L}(f) \circ \mathrm{L}(g) = \mathrm{L}(g) \circ \mathrm{L}(f)$ et $\exists O \in \mathbb{R}^3, f(O) = O = g(O)$.
Alors $\forall M \in \mathbb{R}^3, f \circ g(M) = f \circ g(O) + \mathrm{L}(f \circ g)\left(\overrightarrow{OM}\right) = g \circ f(O) + \mathrm{L}(g \circ f)\left(\overrightarrow{OM}\right) = g \circ f(M)$.

Diffusion à titre gratuit uniquement.

Irréductibilité des polynômes cyclotomiques sur $\mathbb{Z}$

Leçons : 102, 141, *120, 121, 144*

[Per], théorème 4.10

Théorème

Soit $n \in \mathbb{N}^*$; le polynôme cyclotomique Φ_n (qui est dans $\mathbb{Z}[X]$) est irréductible sur $\mathbb{Z}$, donc sur $\mathbb{Q}$.

Démonstration :

Étape 1 : Par récurrence forte, on va commencer par montrer que $\forall n \in \mathbb{N}^*, \Phi_n \in \mathbb{Z}[X]$.
- Si $n = 1$, on a bien $\Phi_1 = X - 1 \in \mathbb{Z}[X]$.
- Soit $n > 1$, on suppose : $\forall k \in [\![1, n-1]\!], \Phi_k \in \mathbb{Z}[X]$.
  Soit $F = \prod_{\substack{d|n \\ d \neq n}} \Phi_d$; on a $F \in \mathbb{Z}[X]$ et F unitaire. D'une part, $\Phi_n F = X^n - 1$. [34]
  Et comme F est unitaire, on peut faire la division euclidienne de $X^n - 1$ par F dans $\mathbb{Z}[X]$:
  $$X^n - 1 = PF + R \text{ avec } P, R \in \mathbb{Z}[X] \text{ et } \deg R < \deg F.$$
  Cette division euclidienne est aussi vraie dans $\mathbb{C}[X]$; elle y est même unique, et on obtient : $P = \Phi_n$ et $R = 0$.
  Conséquemment, $\Phi_n \in \mathbb{Z}[X]$.

Étape 2 : Soit $\xi \in \mathbb{C}$ une racine primitive n^{e} de l'unité ; soit p premier, avec $p \nmid n$.
Ainsi, ξ^p est aussi une racine primitive n^{e} de l'unité, on note $\omega = \xi^p$.

Étape 3 : Soit π_ξ (respectivement π_ω) le polynôme minimal de ξ (resp. ω) sur $\mathbb{Q}$.
On va montrer que π_ξ et π_ω sont des éléments de $\mathbb{Z}[X]$.
$\mathbb{Z}$ est un anneau euclidien, donc a fortiori, $\mathbb{Z}$ est factoriel, donc $\mathbb{Z}[X]$ est un anneau factoriel. [35]
Donc on peut écrire $\Phi_n = \prod_{i=1}^{r} F_i^{\alpha_i}$, où les F_i sont irréductibles dans $\mathbb{Z}[X]$ et les α_i sont dans $\mathbb{N}^*$.
Quitte à multiplier les F_i par -1, comme Φ_n est unitaire, on peut supposer que les F_i sont unitaires.
Comme $\Phi_n(\xi) = \Phi_n(\omega) = 0$; $\exists i_0 \in [\![1, r]\!], F_{i_0}(\xi) = 0$ et $\exists i_1 \in [\![1, r]\!], F_{i_1}(\omega) = 0$.
Mais comme F_{i_0} et F_{i_1} sont irréductibles dans $\mathbb{Z}[X]$ (donc dans $\mathbb{Q}[X]$) et unitaires, ce sont les polynômes minimaux de ξ et ω, sur $\mathbb{Q}$.
Conséquemment, $F_{i_0} = \pi_\xi$ et $F_{i_1} = \pi_\omega$.
Ainsi, on a montré que $\pi_\xi \big| \Phi_n$ et $\pi_\omega | \Phi_n$ dans $\mathbb{Z}[X]$.

Étape 4 : On va montrer désormais que $\pi_\xi = \pi_\omega$; supposons par l'absurde que $\pi_\xi \neq \pi_\omega$.
Comme π_ξ et π_ω sont irréductibles dans $\mathbb{Z}[X]$ et distincts, on a : $\pi_\xi \pi_\omega \big| \Phi_n$ dans $\mathbb{Z}[X]$.
Par ailleurs, comme $\pi_\omega(\xi^p) = 0$, ξ est racine de $\pi_\omega(X^p)$; ainsi, π_ξ étant le polynôme minimal de ξ sur $\mathbb{Q}$, on a : $\pi_\xi(X) \big| \pi_\omega(X^p)$ dans $\mathbb{Q}[X]$.
Autrement dit, $\exists Q \in \mathbb{Q}[X], \pi_\omega(X^p) = \pi_\xi(X)Q(X)$.
Mais $\pi_\xi \in \mathbb{Z}[X]$ est unitaire ; on a donc la division euclidienne dans $\mathbb{Z}[X]$:
$$\pi_\omega(X^p) = \pi_\xi(X)S(X) + R(X) \text{ avec } S, R \in \mathbb{Z}[X] \text{ et } \deg R < \deg \pi_\xi.$$
Par unicité de la division euclidienne dans $\mathbb{Q}[X]$, on a : $Q = S$ et $R = 0$.
En particulier, $\pi_\xi(X) \big| \pi_\omega(X^p)$ dans $\mathbb{Z}[X]$.
Dans la suite, on notera $\overline{P} \in \mathbb{F}_p[X]$ la réduction modulo p du polynôme $P \in \mathbb{Z}[X]$.
Le morphisme de Frobenius fournit alors : $\overline{\pi_\xi(X)} \Big| \overline{\pi_\omega(X^p)} = \overline{\pi_\omega(X)}^p$.
Soit A un facteur irréductible de $\overline{\pi_\xi}$ dans $\mathbb{F}_p[X]$.
Ainsi, $A | \overline{\pi_\omega}^p$, et par le lemme d'Euclide : $A | \overline{\pi_\omega}$.
Comme $\pi_\xi \pi_\omega \big| \Phi_n$ dans $\mathbb{Z}[X]$, $\overline{\pi_\xi \pi_\omega} \big| \overline{\Phi_n}$ dans $\mathbb{F}_p[X]$ et donc $A^2 \big| \overline{\Phi_n}$ dans $\mathbb{F}_p[X]$.
Ensuite, $A^2 \big| \overline{X^n - 1}$, donc $\exists B \in \mathbb{F}_p[X], \overline{X^n - 1} = A^2 B$, puis en dérivant $\overline{nX^{n-1}} = A(2A'B + AB')$.
Donc, $A \Big| \overline{nX^{n-1}}$ dans $\mathbb{F}_p[X]$.
Mais $A \Big| \overline{nX^n - n}$ dans $\mathbb{F}_p[X]$, donc $A | \overline{n} \neq 0$.

34. Cela se justifie bien en disant que les racines n^{es} de l'unité sont les racines primitives d^{es} de l'unité, avec $d | n$.
35. « Je veux bien, je les connais ces résultats-là, mais si on me demande comment ça se démontre ?
– Eh bien, tu n'as qu'à aller voir à la page 136. »

Diffusion à titre gratuit uniquement.

En conséquence, $\deg A = 0$, ce qui contredit l'irréductibilité de A (car $\mathbb{F}_p$ est un corps donc les polynômes de degré nul sont inversibles dans $\mathbb{F}_p[X]$).
On obtient donc (enfin) une contradiction : ainsi $\pi_{\xi^{\bar{k}}} = \pi_\omega$.

Étape 5 : Montrons, par récurrence sur $s \in \mathbb{N}^*$ que :

$$\forall \alpha \text{ racine de } \pi_{\xi^{\bar{k}}}, \forall p_1, \dots, p_s \text{ premiers tels que } (p_1 \dots p_s) \wedge n = 1, \pi_{\xi^{\bar{k}}}\left(\alpha^{p_1 \cdots p_s}\right) = 0.$$

- Si $s = 1$, on a déjà vu ça dans l'étape précédente.
- Soit $s > 1$, α une racine de $\pi_{\xi^{\bar{k}}}$, et $k = p_1 \dots p_s$ où les p_i sont des nombres premiers tels que $k \wedge n = 1$.
  On a $p_1 \dots p_{s-1} \wedge n = 1$ donc $\alpha^{p_1 \cdots p_{s-1}}$ est racine de $\pi_{\xi^{\bar{k}}}$ par hypothèse de récurrence.
  Or, aussi $p_s \wedge n = 1$, donc $\pi_{\xi^{\bar{k}}}\left((\alpha^{p_1 \cdots p_{s-1}})^{p_s}\right) = 0$, d'où $\pi_{\xi^{\bar{k}}}\left(\alpha^k\right) = 0$.

Étape 6 : Ainsi, tous les éléments de μ_n^* sont racines de $\pi_{\xi^{\bar{k}}}$, donc $\deg \pi_{\xi^{\bar{k}}} \geqslant \varphi(n)$.
Et comme $\pi_{\xi^{\bar{k}}} \big| \Phi_n$, on a $\deg \pi_{\xi^{\bar{k}}} = \varphi(n)$; $\pi_{\xi^{\bar{k}}}$ et Φ_n étant unitaires, on obtient $\Phi_n = \pi_{\xi^{\bar{k}}}$. ■

Références

[Per] D. PERRIN – *Cours d'algèbre*, Ellipses, 1996.

Diffusion à titre gratuit uniquement.

Polygones réguliers constructibles [36]

Leçons : 102, 121, 125, 182, 183

[MerCdG], théorème 318

Théorème (*Gauss-Wantzel*)

Soit p un nombre premier impair, $\alpha \in \mathbb{N}^*$.
Alors l'angle $\widehat{\frac{2\pi}{p^\alpha}}$ est constructible $\Leftrightarrow \alpha = 1$ et p est un nombre premier de Fermat (c'est-à-dire que p est un nombre premier qui s'écrit sous la forme $1 + 2^{2^\beta}$, où $\beta \in \mathbb{N}$).

Démonstration :

$\Rightarrow$ On pose $q = p^\alpha$ et $\omega = \exp\left(\frac{2i\pi}{q}\right)$.
On suppose que l'angle $\widehat{\frac{2\pi}{q}}$ est constructible, id est, que $\cos\left(\frac{2\pi}{q}\right)$ est un nombre constructible.
Alors, par le théorème de Wantzel [37], on obtient : $\left[\mathbb{Q}\left(\cos\frac{2\pi}{q}\right) : \mathbb{Q}\right] = 2^m$, où $m \in \mathbb{N}$.
Aussi, le polynôme cyclotomique Φ_q étant le polynôme minimal de ω, on a :

$$[\mathbb{Q}(\omega) : \mathbb{Q}] = \deg \Phi_q = \varphi(q) = p^{\alpha-1}(p-1).$$

Comme $\omega^2 - 2\omega\cos\frac{2\pi}{q} + 1 = 0$, on a $\cos\frac{2\pi}{q} \in \mathbb{Q}(\omega)$ et même $\left[\mathbb{Q}(\omega) : \mathbb{Q}\left(\cos\frac{2\pi}{q}\right)\right] = 2$.
Par multiplicativité du degré, on obtient $2^{m+1} = p^{\alpha-1}(p-1)$.
Comme p est impair, il vient $\alpha = 1$, puis $p = 1 + 2^{m+1}$; montrons que $m+1$ est une puissance de 2.
On écrit alors $m + 1 = \lambda 2^\beta$, avec $\beta \in \mathbb{N}$ et $\lambda \in \mathbb{N}^*$ impair ; on a alors $p = 1 + \left(2^{2^\beta}\right)^\lambda$.
Or, λ étant impair, on a dans $\mathbb{Z}[X]$: $1 + X \big| 1 + X^\lambda$ et donc $1 + 2^{2^\beta} \Big| p$ et donc, comme p est premier, on en déduit $\lambda = 1$.
Donc p est un nombre premier de Fermat.

$\Leftarrow$ On note $n = 2^\beta$, de sorte que $p = 1 + 2^n$, et $\xi = \exp\left(\frac{2i\pi}{p}\right)$.
On a : $[\mathbb{Q}(\xi) : \mathbb{Q}] = \deg \Phi_p = \varphi(p) = p - 1$.

36. On peut ajouter quelques résultats autour de ce développement pour détailler son utilité.

Lemme

- Les angles de la forme $\widehat{\frac{2\pi}{2^\alpha}}$ sont constructibles, où $\alpha \in \mathbb{N}^*$.
- Soient $m, n \in \mathbb{N}^*$, avec $m \wedge n = 1$,
  Alors l'angle $\widehat{\frac{2\pi}{mn}}$ est constructible $\Leftrightarrow \widehat{\frac{2\pi}{m}}$ et $\widehat{\frac{2\pi}{n}}$ le sont.

En conséquence, les polygones réguliers constructibles sont ceux qui possèdent $2^{\alpha_2}\prod_{p\in\mathcal{F}} p^{\alpha_p}$ côtés, où $\mathcal{F}$ est l'ensemble des nombres premiers de Fermat, et où les α_i sont des entiers naturels.

En effet :
- C'est immédiat, puisque par récurrence, il suffit de savoir tracer des bissectrices à la règle et au compas.
- $\Leftarrow$ Il est facile de construire le multiple d'un nombre constructible (en reportant avec le compas le bon nombre de fois la corde formée par l'angle sur le cercle unité).

  $\Rightarrow$ Par Bézout, $\exists \lambda, \mu \in \mathbb{Z}, \lambda m + \mu n = 1$; dès lors $\widehat{\frac{2\pi}{mn}} = \lambda\widehat{\frac{2\pi}{m}} + \mu\widehat{\frac{2\pi}{n}}$.
  Et on construit sans peine la somme de deux angles constructibles en traçant des représentants de ces angles avec un côté adjacent.

37. Le théorème de Pierre-Laurent Wantzel, énoncé en 1837, donne une condition nécessaire et suffisante pour qu'un nombre soit constructible à la règle et au compas : il faut et il suffit que ce nombre appartienne à une extension de $\mathbb{Q}$ qui soit le terme d'une suite d'extensions quadratiques.

Diffusion à titre gratuit uniquement.

On note $G = \mathrm{Aut}_{\mathbb{Q}}(\mathbb{Q}(\xi))$; et si $g \in G$, alors g fixe $\mathbb{Q}$ et est entièrement déterminé par $g(\xi)$.
g étant un morphisme d'anneaux, on a : $0 = g(0) = g\left(\Phi_p(\xi)\right) = \Phi_p(g(\xi))$.
Donc $g(\xi)$ est nécessairement une racine de Φ_p, donc $g(\xi) \in \left\{\xi, \xi^2, \dots, \xi^{p-1}\right\}$.
Il faudrait alors montrer qu'on définit bien ainsi des automorphismes du corps $\mathbb{Q}(\xi)$; et alors

$$G = \left\{g_k : \xi \mapsto \xi^k \middle| k \in [\![1, p-1]\!]\right\} \simeq \left(\mathbb{Z}/p\mathbb{Z}\right)^\times \underset{38}{\simeq} \mathbb{Z}/(p-1)\mathbb{Z}.$$

Désormais, g désignera un générateur de G.
Pour $i \in [\![0, n]\!]$, on note $K_i = \mathrm{Ker}\left(g^{2^i} - \mathrm{Id}\right)$; c'est un sous-corps de $\mathbb{Q}(\xi)$.
De plus, $\forall i \in [\![0, n-1]\!], g^{2^{i+1}} = \left(g^{2^i}\right)^2$ implique $K_i \subseteq K_{i+1}$.
Comme g génère G, $\left(g^i(\xi)\right)_{0 \leqslant i \leqslant p-2}$ est une $\mathbb{Q}$-base de $\mathbb{Q}(\xi)$.
Soit $z \in K_0$, $\exists \lambda_0, \dots, \lambda_{p-2} \in \mathbb{Q}$, $z = \sum_{i=0}^{p-2} \lambda_i g^i(\xi)$; mais $z = g(z) = \lambda_{p-2}\xi + \sum_{i=1}^{p-2} \lambda_{i-1} g^i(\xi)$.
Tous les scalaires λ_i sont donc égaux et $z = \lambda_0 \sum_{i=0}^{p-2} g^i(\xi) = \lambda_0 \sum_{j=1}^{p-1} \xi^j = -\lambda_0 \in \mathbb{Q}$. Donc $K_0 = \mathbb{Q}$.
Pour montrer que $\forall i \in [\![0, n-1]\!], K_i \neq K_{i+1}$, il faudrait considérer l'élément $z = \sum_{h=0}^{2^{n-i-1}-1} g^{2^{i+1}h}(\xi)$. [39]
On en déduit alors qu'on a la suite d'extensions :

$$\mathbb{Q} = K_0 \underset{\neq}{\subset} K_1 \underset{\neq}{\subset} \dots \underset{\neq}{\subset} K_n = \mathbb{Q}(\xi).$$

Mais $2^n = [\mathbb{Q}(\xi) : \mathbb{Q}] = \prod_{i=0}^{n-1} \underbrace{[K_{i+1} : K_i]}_{\geqslant 2}$.
Ainsi, $\forall i \in [\![0, n-1]\!], [K_{i+1} : K_i] = 2$.
Par le théorème de Wantzel, tous les éléments de $\mathbb{Q}(\xi)$ sont donc constructibles ; mais $\cos\frac{2\pi}{p} = \frac{\xi + \xi^{-1}}{2}$ en fait partie. ■

Références

[MerCdG] [40] D.-J. MERCIER – *Cours de géométrie, préparation au CAPES et à l'agrégation*, Publibook, 2008.

38. Pour la cyclicité de $\mathbb{F}_p^\times$, on renvoie à la page 139.

39. Okay, je le fais, mais c'est vraiment parce que c'est vous. On a : $g^{2^i}(z) = \sum_{h=0}^{2^{n-i-1}-1} g^{2^{i+1}h+2^i}(\xi) \neq z$ car les vecteurs de base intervenant dans la décomposition ne sont pas les mêmes (on a décalé les coordonnées de 2^i, alors qu'entre deux coordonnées non-nulles, il y a $2^{i+1}-1$ zéros). Et aussi : $g^{2^{i+1}}(z) = \sum_{h=0}^{2^{n-i-1}-1} g^{2^{i+1}(h+1)}(\xi) = \sum_{h=1}^{2^{n-i-1}-1} g^{2^{i+1}h}(\xi) + \underbrace{g^{2^{i+1}2^{n-i-1}}(\xi)}_{=\xi} = z$.

40. On trouvera également dans cette référence une façon de construire le pentagone régulier à la règle et au compas.

Diffusion à titre gratuit uniquement.

Polynômes irréductibles sur $\mathbb{F}_q$

Leçons : 125, 141, 190, *123, 144*

[FG], exercices 3.11 et 5.10

Théorème

Soit p un nombre premier et $r \in \mathbb{N}^*$; on note $q = p^r$, soit $n \in \mathbb{N}^*$.
Soit $\mathcal{A}(n,q)$ l'ensemble des polynômes de $\mathbb{F}_q[X]$ irréductibles unitaires de degré n et $I(n,q) = \#\mathcal{A}(n,q)$.
Alors :

1. On a la factorisation suivante dans $\mathbb{F}_q[X]$: $X^{q^n} - X = \prod_{d|n} \prod_{P \in \mathcal{A}(d,q)} P$.
2. Notant μ la fonction de Möbius [41], on a : $I(n,q) = \frac{1}{n} \sum_{d|n} \mu\left(\frac{n}{d}\right) q^d$.
3. On en déduit alors l'équivalent : $I(n,q) \underset{n\to\infty}{\sim} \frac{q^n}{n}$.

Démonstration :

On commence bien évidemment par fixer une bonne fois pour toutes une clôture algébrique $\overline{\mathbb{F}_q}$ de $\mathbb{F}_q$, dans laquelle on travaillera tout le long de ce développement.

1. – Soit $d|n$ et $P \in \mathcal{A}(d,q)$; fixons x une racine de P dans $\overline{\mathbb{F}_q}$.
   Alors $K := \mathbb{F}_q(x)$ est un corps de rupture de P et $[K : \mathbb{F}_q] = \deg P = d$ et donc, par unicité des corps finis : $K = \mathbb{F}_{q^d}$.
   Comme $\mathbb{F}_{q^d}$ est l'ensemble des racines de $X^{q^d} - X$ et que ce polynôme divise $X^{q^n} - X$, x est racine de $X^{q^n} - X$. [42]
   Donc $P \Big| X^{q^n} - X$, car P étant irréductible sur un corps fini, il est à racines simples dans $\overline{\mathbb{F}_q}$. [43]
   Puis, par irréductibilité, on en déduit : $\left(\prod_{d|n} \prod_{P \in \mathcal{A}(d,q)} P\right) \Big| X^{q^n} - X$.
   – Soit P un facteur irréductible de $X^{q^n} - X$, de degré $d \geqslant 1$.
   Comme $X^{q^n} - X$ est scindé sur $\mathbb{F}_{q^n}$, P l'est aussi.
   Soit x racine de P, qui est donc dans $\mathbb{F}_{q^n}$; $K := \mathbb{F}_q(x)$ est un corps intermédiaire entre $\mathbb{F}_q$ et $\mathbb{F}_{q^n}$.
   On a alors : $[\mathbb{F}_{q^n} : K][K : \mathbb{F}_q] = [\mathbb{F}_{q^n} : \mathbb{F}_q] = n$ et donc $d = [K : \mathbb{F}_q]$ divise n.

41. Pour $k \in \mathbb{N}^*$, on définit μ par $\mu(k) = 0$ si k possède un facteur carré et $\mu(k) = (-1)^r$ sinon, où r désigne alors le nombre de facteurs premiers distincts de k.

42. En effet, $x^{q^n} = x^{(q^d)^{\frac{n}{d}}} = x^{q^d (q^d)^{\frac{n}{d}-1}} = \left(x^{q^d}\right)^{(q^d)^{\frac{n}{d}-1}} = x^{(q^d)^{\frac{n}{d}-1}} = \ldots = x^{(q^d)^0} = x$.

43. Démontrons ce fait général :

Lemme

Si K est un corps fini ou de caractéristique nulle, et si $P \in K[X]$ est un polynôme irréductible,
Alors P est à racines simples dans la clôture algébrique $\overline{K}$ de K.

En effet, soit α une racine multiple de P, alors $\exists Q \in \overline{K}[X], P = (X-\alpha)^2 Q$.
En dérivant, on obtient ensuite $P' = (X-\alpha)(2Q + (X-\alpha)Q')$, donc $(X-\alpha)|P'$ dans $\overline{K}[X]$.
Et finalement, $(X-\alpha)|P \wedge P'$ dans $\overline{K}[X]$.
Or $P \wedge P'|P$ dans $K[X]$ et $\deg P \wedge P' \geqslant 1$, donc, par irréductibilité de P, on a : $P = P \wedge P'$, et comme $\deg P' < \deg P$: $P' = 0$.
Si K est de caractéristique nulle, alors P est une constante donc nul ou inversible donc non-irréductible. On a donc une contradiction dans ce cas.
Sinon, K est de caractéristique $p > 0$ et $P \in K[X^p]$, d'où $P(X) = R(X^p) = R(X)^p$, avec $R \in K[X]$, car le morphisme de Frobenius est un automorphisme quand le corps est fini.
Pour terminer sur ce sujet, citons un contre-exemple sur un corps infini de caractéristique positive.
Soit $P = T^2 + X \in \mathbb{F}_2(X)[T]$.
P est irréductible, car il n'admet pas de racine dans $\mathbb{F}_2(X)$ (pour des problèmes de parité) et qu'il est de degré 2.
P admet une racine $\alpha = \sqrt{-X} \in \overline{\mathbb{F}_2(X)}$ et donc $P = (T-\alpha)(T+\alpha) = (T-\alpha)^2$ car $\overline{\mathbb{F}_2(X)}$ est un corps de caractéristique 2.

Diffusion à titre gratuit uniquement.

Les racines de $X^{q^n} - X$ dans $\mathbb{F}_{q^n}$ sont simples donc tous les facteurs irréductibles de $X^{q^n} - X$ dans $\mathbb{F}_q[X]$ interviennent avec une multiplicité égale à 1.

Ainsi $X^{q^n} - X \left| \left(\prod_{d|n} \prod_{P \in \mathcal{A}(d,q)} P \right) \right.$.

– Les deux polynômes considérés étant unitaires, on en déduit $X^{q^n} - X = \prod_{d|n} \prod_{P \in \mathcal{A}(d,q)} P$.

2. **Lemme (*Formule d'inversion de Möbius*)**

Soit $f : \mathbb{N}^* \to \mathbb{R}$ et $g : \left| \begin{array}{ccc} \mathbb{N}^* & \to & \mathbb{R} \\ n & \mapsto & \sum_{d|n} f(d) \end{array} \right.$.

Alors $\forall n \in \mathbb{N}^*, f(n) = \sum_{d|n} \mu\left(\frac{n}{d}\right) g(d) = \sum_{d|n} \mu(d) g\left(\frac{n}{d}\right)$.

Démonstration :

Soit $n \in \mathbb{N}^*$, on a : $\sum_{d|n} \mu(d) g\left(\frac{n}{d}\right) = \sum_{d|n} \sum_{d' \left| \frac{n}{d} \right.} \mu(d) f(d') = \sum_{dd'|n} \mu(d) f(d') = \sum_{d'|n} f(d') \left(\sum_{d \left| \frac{n}{d'} \right.} \mu(d) \right)$.

Soit désormais $k \in \mathbb{N}^*$, $k > 1$.

On écrit $k = p_1^{\alpha_1} \dots p_r^{\alpha_r}$ sa décomposition en facteurs premiers, avec $r > 0$.

On a ensuite : $\sum_{d|k} \mu(d) = \mu(1) + \sum_{i=1}^{r} \sum_{1 \leqslant \gamma_1 < \dots < \gamma_i \leqslant r} \mu\left(p_{\gamma_1} \dots p_{\gamma_i}\right) + 0 = \sum_{i=0}^{r} \binom{r}{i} (-1)^i = (1-1)^r = 0$.

Et si $k = 1$, $\sum_{d|1} \mu(d) = \mu(1) = 1$, ce qui nous donne donc $\sum_{d|n} \mu(d) g\left(\frac{n}{d}\right) = f(n)$. ■

En regardant les degrés dans la factorisation précédente, on obtient : $q^n = \sum_{d|n} dI(d,q)$.

En appliquant la formule d'inversion de Möbius à la fonction $f : n \mapsto nI(n,q)$, on obtient :

$$\forall n \in \mathbb{N}^*, nI(n,q) = \sum_{d|n} \mu\left(\frac{n}{d}\right) q^d.$$

3. On pose $r_n = \sum_{\substack{d|n \\ d<n}} \mu\left(\frac{n}{d}\right) q^d$ et alors : $|r_n| \leqslant \sum_{i=1}^{\left\lfloor \frac{n}{2} \right\rfloor} q^d = q \frac{q^{\left\lfloor \frac{n}{2} \right\rfloor} - 1}{q-1} = \mathcal{O}_{n\to\infty}\left(q^{\left\lfloor \frac{n}{2} \right\rfloor}\right)$.

Donc r_n est négligeable devant q^n, d'où : $I(n,q) = \frac{q^n + r_n}{n} \underset{n\to\infty}{\sim} \frac{q^n}{n}$. ■

Références

[FG] S. FRANCINOU et H. GIANELLA – *Exercices de mathématiques pour l'agrégation (Algèbre 1)*, 2ème éd., Masson, 1997.

Diffusion à titre gratuit uniquement.

Réduction des endomorphismes normaux [44]

Leçons : 150, 151, 153, 154, 155, 160, *101*, *158*

[Gou Al], section 5.3.2

Théorème

Soit E un $\mathbb{R}$-espace vectoriel de dimension $n < \infty$, et $f \in \mathcal{L}(E)$ un endomorphisme normal.
Alors, il existe une base orthonormale $\mathcal{B}$ de E, dans laquelle la matrice de f est une matrice de $\mathcal{M}_n(\mathbb{R})$ de type $(*)$:

$$\begin{pmatrix} \lambda_1 & & & & & & \\ & \ddots & & & & 0 & \\ & & \lambda_r & & & & \\ & & & \boxed{\begin{matrix} a_1 & -b_1 \\ b_1 & a_1 \end{matrix}} & & & \\ & & & & \ddots & & \\ & 0 & & & & \boxed{\begin{matrix} a_s & -b_s \\ b_s & a_s \end{matrix}} & \end{pmatrix}$$

Démonstration :

On va d'abord énoncer et montrer deux lemmes.

Lemme 1

1. Soit F un sous-espace vectoriel stable par $u \in \mathcal{L}(E)$. Alors $F^\perp$ est stable par u^*.
2. Soit E_λ un sous-espace propre de $u \in \mathcal{L}(E)$, endomorphisme normal.
   Alors E_λ et $E_\lambda^\perp$ sont stables par u et u^*.

Démonstration du lemme 1 :

1. On a : $\forall x \in F, \forall y \in F^\perp, \langle x, u^*(y)\rangle = \langle u(x), y\rangle = 0$, car $u(x) \in F$. Donc $\forall y \in F^\perp, u^*(y) \in F^\perp$.
2. D'abord E_λ est stable par u, donc on vient de montrer que $E_\lambda^\perp$ est stable par u^*.
   On a : $\forall x \in E_\lambda, u(u^*(x)) = u^*(u(x)) = u^*(\lambda x) = \lambda u^*(x)$, donc $u^*(x) \in E_\lambda$.
   Donc E_λ est stable par u^*, donc par la première partie du lemme : $E_\lambda^\perp$ est stable par $u^{**} = u$. ■

Lemme 2

Soit E un $\mathbb{R}$-espace vectoriel de dimension 2, $u \in \mathcal{L}(E)$ un endomorphisme normal, avec $\mathrm{Sp}_{\mathbb{R}}(u) = \varnothing$.
Alors dans toute base orthonormale $\mathcal{B}$ de E, $\mathcal{M}at_{\mathcal{B}}(u)$ est de la forme $\begin{pmatrix} a & -b \\ b & a \end{pmatrix}$, avec $b \neq 0$.

Démonstration du lemme 2 :

Écrivons $M = \mathcal{M}at_{\mathcal{B}}(u) = \begin{pmatrix} a & c \\ b & d \end{pmatrix}$; comme $\mathrm{Sp}_{\mathbb{R}}(u) = \varnothing$, on a nécessairement $b \neq 0$.

Par normalité, on a : $MM^* = M^*M$, on a : $\begin{pmatrix} a^2+c^2 & ab+cd \\ ab+cd & a^2+c^2 \end{pmatrix} = \begin{pmatrix} a^2+b^2 & ac+bd \\ ac+bd & c^2+d^2 \end{pmatrix}$.

Donc $b = \pm c$.

↪ Si $b = c$, alors M est symétrique réelle ; mais $\mathrm{Sp}_{\mathbb{R}}(u) = \varnothing$. Contradiction. [45]

↪ Donc $b = -c$, et alors : $ab + cd = ac + bd$ donne $2(a-d)b = 0$ donc $a = d$. ■

44. On ne prendra pas le temps d'écrire l'énoncé au tableau. Dans le plan de la leçon 150, on énoncera le résultat ainsi : "Soit $M \in \mathcal{M}_n(\mathbb{R})$ une matrice normale ; alors, sous l'action par conjugaison induite par $\mathrm{O}_n(\mathbb{R})$, l'orbite de M contient une matrice de type $(*)$". On ne démontrera pas le lemme 2.

45. Soit $S \in \mathcal{S}_n(\mathbb{R})$, alors S admet une valeur propre réelle. En effet, S est une matrice complexe, donc admet une valeur propre $\lambda \in \mathbb{C}$, associée à un vecteur propre $x \in \mathbb{C}^n \backslash \{0\}$. De l'égalité $Sx = \lambda x$, on déduit :

$$\lambda \sum_{k=1}^{n} |x_k|^2 = \lambda {}^t x \overline{x} = ({}^t x S)\,\overline{x} = {}^t x\,(S\overline{x}) = \overline{\lambda}\,{}^t x \overline{x} = \overline{\lambda} \sum_{k=1}^{n} |x_k|^2.$$

Et comme $x \neq 0$, on en déduit que $\lambda = \overline{\lambda}$, autrement dit $\lambda \in \mathbb{R}$.

Diffusion à titre gratuit uniquement.

On procède par récurrence forte sur $n = \dim E$.
- Pour $n = 1$, le résultat est trivial.
- Soit $n \geqslant 2$. On a deux cas.

Cas 1 : $\mathrm{Sp}_{\mathbb{R}}(f) \neq \varnothing$. Soit alors $\lambda \in \mathrm{Sp}_{\mathbb{R}}(f)$.
Par le lemme 1, $E_\lambda^\perp$ est stable par f et f^*, donc $f|_{E_\lambda^\perp}$ est normal [46].
Comme $\dim E_\lambda^\perp \leqslant n-1$, par récurrence, il existe $\mathcal{B}_2$ une base orthonormale de $E_\lambda^\perp$ telle que $\mathcal{M}at_{\mathcal{B}_2}\left(f|_{E_\lambda^\perp}\right)$ soit de la forme $(*)$.
Soit alors $\mathcal{B}_1$ une base orthonormale de E_λ [47].
La concaténation $\mathcal{B} = (\mathcal{B}_1, \mathcal{B}_2)$ est une base orthonormale de E telle que $\mathcal{M}at_{\mathcal{B}}(f)$ soit de la forme $(*)$.

Cas 2 : $\mathrm{Sp}_{\mathbb{R}}(f) = \varnothing$.
Soit $Q = X^2 + \alpha X + \beta$ un facteur irréductible de χ_f dans $\mathbb{R}[X]$. Notons $N = \mathrm{Ker}\, Q(f)$.
Montrons que : $N \neq \{0\}$; effectivement, écrivons $Q = (X-\lambda)(X-\overline{\lambda})$ dans $\mathbb{C}[X]$.
Ainsi, $\lambda \in \mathrm{Sp}_{\mathbb{C}}(f)$ et $\det(f - \lambda \mathrm{Id}) = 0$, donc

$$\det Q(f) = \det(f - \lambda \mathrm{Id}) \det(f - \overline{\lambda}\mathrm{Id}) = 0,$$

donc $N \neq \{0\}$.
Comme $Q(f)$ commute avec f et f^*, alors N est stable par f et f^*.
Posons $g = f|_N$ alors $g^* = f^*|_N$ et $g^* g = (f^* f)\,|_N$ est symétrique réel.
Soient alors $\mu \in \mathrm{Sp}_{\mathbb{R}}(g^* g)$ et $x \in N \backslash \{0\}$, tels que $g^* g(x) = \mu x$.
Soit $F = \mathrm{Vect}\{x, f(x)\}$; on a $\dim F = 2$ car $\mathrm{Sp}_{\mathbb{R}}(f) = \varnothing$.
Comme $f^2(x) = -\alpha f(x) - \beta x$, alors F est stable par f ; même : $F = \mathrm{Vect}\left\{f(x), f^2(x)\right\}$ car $\beta \neq 0$.
On a $f^*(f(x)) = \mu x \in F$ et $f^*\left(f^2(x)\right) = f\left(f^* f(x)\right) = f(\mu x) = \mu f(x) \in F$.
Donc F est stable par f^* ; ainsi, $f|_F$ est un endomorphisme normal.
Par le lemme 2, il existe $\mathcal{B}_2$ une base orthonormale de F avec

$$\mathcal{M}at_{\mathcal{B}_2}\left(f|_F\right) = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}, \text{ où } b \neq 0.$$

Comme F est stable par f et f^*, par le début du lemme 1, $F^\perp$ est stable par f^* et $f^{**} = f$.
Donc $f|_{F^\perp}$ est un endomorphisme normal.
Comme $F^\perp$ est de dimension $n-2$, par récurrence, il existe une base orthonormale $\mathcal{B}_1$ de $F^\perp$, telle que $\mathcal{M}at_{\mathcal{B}_1}\left(f|_{F^\perp}\right)$ soit de type $(*)$.
Et la concaténation $\mathcal{B} = (\mathcal{B}_1, \mathcal{B}_2)$ est une base orthonormale de E, telle que $\mathcal{M}at_{\mathcal{B}}(f)$ soit de type $(*)$. ■

Références

[Gou Al] X. GOURDON – *Les maths en tête : Algèbre*, 2e éd., Ellipses, 2009.

46. En effet, $f^*|_{E_\lambda^\perp}$ existe ; par unicité de l'adjoint, on montre que : $f^*|_{E_\lambda^\perp} = \left(f|_{E_\lambda^\perp}\right)^*$, puis on a que $f|_{E_\lambda^\perp}$ et $\left(f|_{E_\lambda^\perp}\right)^*$ commutent car f et f^* commutent.
47. Ben oui : E_λ est de dimension finie donc admet une base et on l'orthonormalise par Schmidt !

Diffusion à titre gratuit uniquement.

Simplicité de $\mathfrak{A}_n$ pour $n \geqslant 5$

Leçons : 101, 103, 104, 105, 108

[Ulm], exercices 7.10-11
[Per], théorème 8.1

Théorème

Soit $n \in \mathbb{N}$, $n \geqslant 5$.
$\mathfrak{A}_n$ est simple.

Démonstration :

Étape 1 : Montrons que $\mathfrak{A}_5$ est simple.
Soit $H \triangleleft \mathfrak{A}_5$, H est réunion de classes de conjugaison dans $\mathfrak{A}_5$.
On connait les classes de conjugaison dans $\mathfrak{S}_5$, déduisons-en celles de $\mathfrak{A}_5$.

Lemme

Soit $n \in \mathbb{N}$, $G = \mathfrak{A}_n$ ou $\mathfrak{S}_n$.
On note $\sigma^G = \left\{\gamma\sigma\gamma^{-1} \middle| \gamma \in G\right\}$ et $Z_G(\sigma) = \left\{\gamma \in G \middle| \gamma\sigma\gamma^{-1} = \sigma\right\}$ où $\sigma \in \mathfrak{A}_n$.
On a deux cas :
- Soit $\#\sigma^{\mathfrak{A}_n} = \dfrac{1}{2}\#\sigma^{\mathfrak{S}_n}$ et $Z_{\mathfrak{A}_n}(\sigma) = Z_{\mathfrak{S}_n}(\sigma)$;
- Soit $\sigma^{\mathfrak{A}_n} = \sigma^{\mathfrak{S}_n}$ et $\exists \alpha \in \mathfrak{S}_n \backslash \mathfrak{A}_n, \alpha\sigma\alpha^{-1} = \sigma$.

Démonstration :
On considère $\tilde{\varepsilon} : \left|\begin{array}{ccc} Z_{\mathfrak{S}_n}(\sigma) & \to & \{\pm 1\} \\ \gamma & \mapsto & \varepsilon(\gamma) \end{array}\right.$. On a donc deux cas :
- Soit $\tilde{\varepsilon}$ est trivial, alors $Z_{\mathfrak{S}_n}(\sigma) \subseteq \mathfrak{A}_n$ et par suite $Z_{\mathfrak{S}_n}(\sigma) = Z_{\mathfrak{A}_n}(\sigma)$.
  Par la relation orbite-stabilisateur pour l'action de conjugaison :
  $$\#\mathfrak{S}_n = \#Z_{\mathfrak{S}_n}(\sigma) \times \#\sigma^{\mathfrak{S}_n} \text{ et } \#\mathfrak{A}_n = \#Z_{\mathfrak{A}_n}(\sigma) \times \#\sigma^{\mathfrak{A}_n} \text{ d'où } \#\sigma^{\mathfrak{A}_n} = \frac{1}{2}\#\sigma^{\mathfrak{S}_n}.$$
- Soit $\tilde{\varepsilon}$ est non-trivial et $Z_{\mathfrak{A}_n}(\sigma) = \operatorname{Ker}\tilde{\varepsilon}$ est d'indice 2 dans $Z_{\mathfrak{S}_n}(\sigma)$.
  Ainsi, il existe $\alpha \in \mathfrak{S}_n \backslash \mathfrak{A}_n$, tel que $\alpha\sigma\alpha^{-1} = \sigma$.
  De plus, la relation orbite-stabilisateur donne ici : $\#\sigma^{\mathfrak{A}_n} = \#\sigma^{\mathfrak{S}_n}$ d'où $\sigma^{\mathfrak{A}_n} = \sigma^{\mathfrak{S}_n}$. ■

D'après le lemme, au passage de $\mathfrak{S}_5$ à $\mathfrak{A}_5$, les classes de conjugaison sont soit conservées, soit coupées en deux. On construit alors le tableau suivant :

Classes dans $\mathfrak{S}_5$	Cardinaux dans $\mathfrak{S}_5$	Passage dans $\mathfrak{A}_5$
Id	1	Conservée car de cardinal impair
$(a\ b\ c)$	20	Conservée car $(d\ e)(a\ b\ c)(d\ e)^{-1} = (a\ b\ c)$
$(a\ b)(c\ d)$	15	Conservée car de cardinal impair
$(a\ b\ c\ d\ e)$	24	Coupée car $24 \nmid 60$ (relation orbite-stabilisateur)

Or H est réunion de classes de conjugaison, contient Id et vérifie $\#H \mid 60$ (théorème de Lagrange).
Ainsi, on conclut assez rapidement que $\#H \in \{1, 60\}$ et donc $\mathfrak{A}_5$ est simple. [48]

Étape 2 : Déduisons-en que $\mathfrak{A}_n$ est simple pour tout $n \geqslant 5$.
Soit $H \triangleleft \mathfrak{A}_n$, $H \neq \{\mathrm{Id}\}$.
Soit $\sigma \in H \backslash \{\mathrm{Id}\}$ et $a \in [\![1, n]\!]$ tel que : $a \neq \sigma(a) =: b$.

48. On peut faire autrement ici (on ne le fera pas, parce que c'est plus long même si c'est plus rigolo). $\mathfrak{A}_5$ admet 5 classes de conjugaison donc 5 caractères irréductibles : de degrés 1 (pour la représentation triviale), d_2, d_3, d_4 et d_5. On a : $60 = 1 + d_2^2 + d_3^2 + d_4^2 + d_5^2$, et on en déduit (en essayant de manière exhaustive, c'est bon hein, 59 c'est pas la mort non plus), que $d_2 = d_3 = 3$, $d_4 = 4$ et $d_5 = 5$. Soit $\mathcal{C}$ une classe de conjugaison de $\mathfrak{A}_5$. $\mathbb{1}_{\mathcal{C}}$ étant une fonction centrale, on obtient :
$$\mathbb{1}_{\mathcal{C}} = \sum_{i=1}^5 \langle \mathbb{1}_{\mathcal{C}}, \chi_i\rangle \chi_i = \frac{1}{\#G}\sum_{i=1}^5 \sum_{g\in G} \mathbb{1}_{\mathcal{C}}(g)\overline{\chi_i(g)}\chi_i = \frac{\#\mathcal{C}}{\#G}\sum_{i=1}^5 \overline{\chi_i(\mathcal{C})}\chi_i.$$
Et donc $1 = \frac{\#\mathcal{C}}{\#G}\sum_{i=1}^5 |\chi_i(\mathcal{C})|^2$. Finalement, $\sum_{i=1}^5 |\chi_i(\mathcal{C})|^2 = \frac{\#G}{\#\mathcal{C}}$. Supposons maintenant que $\mathcal{C} \neq \{\mathrm{Id}\}$, et donc $\frac{\#G}{\#\mathcal{C}} \leqslant 5$, donc $\forall i \in [\![2, 5]\!]$, $|\chi_i(\mathcal{C})| < 3 \leqslant \chi_i(\mathrm{Id})$. Donc $\forall i \in [\![2, 5]\!]$, $\operatorname{Ker}\chi_i = \{\mathrm{Id}\}$. Donc les seuls sous-groupes distingués de $\mathfrak{A}_5$ sont $\mathfrak{A}_5$ et $\{\mathrm{Id}\}$, donc $\mathfrak{A}_5$ est simple.

Diffusion à titre gratuit uniquement.

Soit $c \notin \{a, b, \sigma(b)\}$, et $\tau = (a\ c\ b)$ et $\tau^{-1} = (a\ b\ c)$.
On définit $\rho = \underbrace{\tau\sigma\tau^{-1}}_{\in H}\underbrace{\sigma^{-1}}_{\in H} = \tau\left(\sigma\tau^{-1}\sigma^{-1}\right) = (a\ b\ c)(\sigma(a)\ \sigma(b)\ \sigma(c)) \in H$.
Donc supp $\rho = \{a, b, c, \sigma(b), \sigma(c)\}$, d'où # supp $\rho \leqslant 5$.
Aussi $\rho \neq \mathrm{Id}$ car $\rho(b) = \tau\sigma\tau^{-1}(a) = \tau\sigma(b) \neq b$ car $\sigma(b) \neq c = \tau^{-1}(b)$.
Soit alors $E \subseteq [\![1, n]\!]$, avec $E \supseteq$ supp ρ et $\#E = 5$.
On définit l'injection $i : \left|\begin{array}{ccc} \mathfrak{A}(E) & \to & \mathfrak{A}_n \\ u & \mapsto & \overline{u} \end{array}\right.$ où $\overline{u}_{|E} = u$ et $\overline{u}_{|E^C} = \mathrm{Id}$.
Soit $H' = i^{-1}(H)$; i est un morphisme de groupes et donc $H' \triangleleft \mathfrak{A}(E)$.
Or $\rho_{|E} \neq \mathrm{Id}$ et $\rho_{|E} \in H'$ car $\rho \in H$.
Ainsi, $H' = \mathfrak{A}(E)$ car $\mathfrak{A}(E) \simeq \mathfrak{A}_5$ est simple.
Soit alors v un 3-cycle de $\mathfrak{A}(E)$; $\overline{v} \in H$ et $\overline{v}$ est un 3-cycle de $\mathfrak{A}_n$.
Comme on l'a vu précédemment, pour $n \geqslant 5$, les 3-cycles sont conjugués dans $\mathfrak{A}_n$.
Donc H contient tous les 3-cycles ; mais ils engendrent $\mathfrak{A}_n$!
Finalement, $H = \mathfrak{A}_n$ et donc $\mathfrak{A}_n$ est simple. ■

Références

[Ulm] F. ULMER – *Théorie des groupes*, Ellipses, 2012.
[Per] D. PERRIN – *Cours d'algèbre*, Ellipses, 1996.

Diffusion à titre gratuit uniquement.

Sous-groupes distingués et caractères

Leçons : 107, 109, *101, 103, 104, 106*

[Ulm], partie 17.3
[Pey], partie VIII.1.3

Théorème

Soit G un groupe fini de caractères irréductibles $\chi_1, \dots, \chi_m$.
Alors les sous-groupes distingués de G sont les $\bigcap_{j\in J} \operatorname{Ker} \chi_j$, où $J \subset [\![1, m]\!]$.

On commence par montrer le lemme qui suit.

Lemme

Soit $\rho : G \to \mathrm{GL}(V)$ une représentation de G de caractère χ.
Alors $\operatorname{Ker} \chi = \operatorname{Ker} \rho$.

Démonstration du lemme :

Soit $g \in G$, par Lagrange, $\rho(g)$ est d'ordre fini divisant $\#G$, donc est diagonalisable ; on peut même dire que ses valeurs propres, qu'on nomme λ_j, où $j \in [\![1, \dim V]\!]$, sont des racines de l'unité et donc de module 1.

On en déduit $\chi(g) = \sum_{i=1}^{\dim V} \lambda_j$, ce qui fournit : $|\chi(g)| \leqslant \sum_{j=1}^{\dim V} |\lambda_j| = \dim V = \chi(\mathrm{e})$.

Mais le cas d'égalité dans cette inégalité triangulaire est $\lambda_1 = \ldots = \lambda_{\dim V}$.

Ainsi : $\chi(g) = \chi(\mathrm{e}) \Leftrightarrow \lambda_1 = \ldots = \lambda_{\dim V} = 1 \Leftrightarrow \rho(g) = \mathrm{Id}_V \Leftrightarrow g \in \operatorname{Ker} \rho$. ■

Démonstration du théorème :

Étape 1 : Soit $H \triangleleft G$.

On considère l'action par translation à gauche de G sur ${}^{G}/_{H}$ dont on note $\varphi : G \to \mathfrak{S}_{{}^{G}/_{H}}$ le morphisme structurel.

Soit alors χ le caractère associé à la représentation par permutations $\rho_\varphi : G \to \mathrm{GL}(V)$, où V est un $\mathbb{C}$-espace vectoriel de dimension $\#{}^{G}/_{H}$.

Le lemme précédent nous donne que : $\operatorname{Ker} \chi = \operatorname{Ker} \rho_\varphi = \operatorname{Ker} \varphi = H$.

Les sous-groupes distingués de G sont donc les noyaux des caractères de G.

Étape 2 : On va exprimer les noyaux des caractères de G en fonction des noyaux des caractères irréductibles de G.

Soit χ un caractère de G, associé à la représentation $\rho : G \to \mathrm{GL}(V)$.

On décompose alors V en somme directe de sous-représentations irréductibles : $V = \bigoplus_{i=1}^{s} \underbrace{V_i \oplus \ldots \oplus V_i}_{a_i \text{ fois}}$,

avec $V_1, \dots, V_s$ représentations irréductibles "distinctes" de G, de caractères associés $\chi_1, \dots, \chi_s$.

On a alors : $g \in \operatorname{Ker} \chi \Leftrightarrow g \in \operatorname{Ker} \rho \Leftrightarrow \forall i \in [\![1, s]\!], g \in \operatorname{Ker} \rho_{|V_i} \Leftrightarrow \forall i \in [\![1, s]\!], g \in \operatorname{Ker} \chi_i$.

Donc $\operatorname{Ker} \chi = \bigcap_{i=1}^{s} \operatorname{Ker} \chi_i$. ■

On va appliquer ce résultat à $\mathcal{D}_6$. Commençons par déterminer sa table de caractères.

Étape 1 : On va chercher d'abord tous ses caractères irréductibles de degré 1 ; ce sont les morphismes de groupes de $\mathcal{D}_6 \to \mathbb{C}^*$.

Nécessairement, si χ est un caractère de degré 1 de $\mathcal{D}_6$, on a : $\chi(s)^2 = \chi(\mathrm{e}) = 1$ et donc $\chi(s) = \pm 1$.

Mais aussi, $\chi(sr)^2 = 1$ et donc $\chi(s)\chi(r)\chi(s)\chi(r) = 1$, d'où $\chi(r) = \pm 1$.

Il y a donc 4 candidats possibles pour être une représentation parmi ces applications, il resterait à vérifier que ce sont bien des morphismes de groupes de $\mathcal{D}_6 \to \mathbb{C}^*$. [49]

49. Kaa vous suggère d'avoir confiance...

Diffusion à titre gratuit uniquement.

Étape 2 : Les autres caractères de $\mathcal{D}_6$ sont nécessairement de degré supérieur ou égal à 2.
La relation reliant degrés des caractères irréductibles et cardinal du groupe, $\sum_{i\in I} n_i^2 = \#G$, nous indique qu'il nous reste à trouver 2 caractères irréductibles de degré 2.
On aura alors la table de $\mathcal{D}_6$.

Étape 3 : On pose $\omega = \mathrm{e}^{\frac{\mathrm{i}\pi}{3}}$, et pour $h \in \{1,2\}$: $\rho_h\left(r^k\right) = \begin{pmatrix} \omega^{hk} & 0 \\ 0 & \omega^{-hk} \end{pmatrix}$ et $\rho_h\left(sr^k\right) = \begin{pmatrix} 0 & \omega^{-hk} \\ \omega^{hk} & 0 \end{pmatrix}$.
On devrait vérifier que $\rho_h : \mathcal{D}_6 \to \mathrm{GL}\left(\mathbb{C}^2\right)$ est bien un morphisme de groupes, ie, une représentation.
On note alors χ_h, pour $h \in \{1,2\}$, le caractère de ρ_h.

$$\text{On a : } \#\mathcal{D}_6\langle\chi_h,\chi_h\rangle = \sum_{k=0}^{5}\left(\chi_h\left(r^k\right)^2 + \chi_h\left(sr^k\right)^2\right) = \sum_{k=0}^{5}\left(\omega^{hk}+\omega^{-hk}\right)^2 = \sum_{k=0}^{5}\omega^{2hk} + 2\times 6 + \sum_{k=0}^{5}\omega^{-2hk}$$
$$= 12 + 4\left(1+\mathrm{j}+\mathrm{j}^2\right) = 12.$$

Donc χ_1 et χ_2 sont bien irréductibles.
Voici la table des caractères irréductibles de $\mathcal{D}_6$.

	r^k	sr^k
ψ_1	1	1
ψ_2	1	-1
ψ_3	$(-1)^k$	$(-1)^k$
ψ_4	$(-1)^k$	$(-1)^{k+1}$
χ_1	$2\cos\left(\frac{k\pi}{3}\right)$	0
χ_2	$2\cos\left(\frac{2k\pi}{3}\right)$	0

Les sous-groupes distingués de $\mathcal{D}_6$ sont donc : $\mathcal{D}_6$, $\langle r\rangle$, $\langle s, r^2\rangle$, $\langle sr, r^2\rangle$, $\{\mathrm{e}\}$, $\langle r^3\rangle$ et $\langle r^2\rangle$ (intersection des noyaux de ψ_2 et ψ_3).

Références

[Ulm] F. ULMER – *Théorie des groupes*, Ellipses, 2012.
[Pey] G. PEYRÉ – *L'algèbre discrète de la transformée de Fourier*, Ellipses, 2004.

Diffusion à titre gratuit uniquement.

Table de caractères de $\mathcal{D}_n$

Leçons : 108, *102, 104, 107, 109, 154*

Librement inspiré de [Pey], paragraphe VIII.1.3

On va construire la table de caractères de $\mathcal{D}_n$.

On commence par chercher tous les caractères irréductibles de degré 1 ; ce sont les morphismes de groupes de $\mathcal{D}_n \to \mathbb{C}^*$.
Nécessairement, si χ est un caractère irréductible de degré 1 de $\mathcal{D}_n$, alors $\chi(s)^2 = \chi(\mathrm{e}) = 1$, donc $\chi(s) = \pm 1$.
Aussi, $\chi(sr)^2 = 1$, donc $\chi(s)\chi(r)\chi(s)\chi(r) = 1$, d'où $\chi(r) = \pm 1$.
La dernière relation que doivent vérifier les générateurs est $r^n = \mathrm{e}$.

$\to$ Si n est pair, aucun problème.

$\to$ Si n est impair, cela impose $\chi(r) = 1$.

On obtient donc deux cas :

Cas où n est pair :

	r^k	sr^k
ψ_1	1	1
ψ_2	1	-1
ψ_3	$(-1)^k$	$(-1)^k$
ψ_4	$(-1)^k$	$(-1)^{k+1}$

Les autres caractères irréductibles sont de degré $\geqslant 2$.

On pose $\omega_n = \exp\left(\dfrac{2\mathrm{i}\pi}{n}\right)$ et pour $h \in \mathbb{N}$, $\rho_h : \left| \begin{array}{ccc} r^k & \mapsto & \begin{pmatrix} \omega_n^{hk} & 0 \\ 0 & \omega_n^{-hk} \end{pmatrix} \\ sr^k & \mapsto & \begin{pmatrix} 0 & \omega_n^{-hk} \\ \omega_n^{hk} & 0 \end{pmatrix} \end{array} \right.$.

On devrait[50] ici vérifier que $\rho_h : \mathcal{D}_n \to \mathrm{GL}\left(\mathbb{C}^2\right)$ est un morphisme de groupes, donc une représentation de $\mathcal{D}_n$ de degré 2. On note χ_h son caractère.
Maintenant, on remarque que :
– il suffit de prendre $h \in [\![0, n-1]\!]$, car $\omega_n^n = 1$;
– $\forall h \in [\![0, n-1]\!]$, ρ_h et ρ_{n-h} sont isomorphes, car $\forall g \in \mathcal{D}_n, \rho_h(g) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \rho_{n-h}(g) \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}^{-1}$; et on peut se restreindre à $h \in \left[\!\left[0, \dfrac{n}{2}\right]\!\right]$;
– $\chi_0 : \left| \begin{array}{ccc} r^k & \mapsto & 2 \\ sr^k & \mapsto & 0 \end{array} \right.$ donc $\chi_0 = \psi_1 + \psi_2$;
– $\chi_{\frac{n}{2}} : \left| \begin{array}{ccc} r^k & \mapsto & 2(-1)^k \\ sr^k & \mapsto & 0 \end{array} \right.$ donc $\chi_{\frac{n}{2}} = \psi_3 + \psi_4$.

On se restreint donc à $h \in \left[\!\left[1, \dfrac{n}{2} - 1\right]\!\right]$.

Soit $h \in \left[\!\left[1, \dfrac{n}{2} - 1\right]\!\right]$, supposons ρ_h réductible ; alors ρ_h admettrait une sous-représentation non-triviale.
Mais les seules droites stables par $\rho_h(r)$ sont $\mathbb{C}e_1$ et $\mathbb{C}e_2$ — où (e_1, e_2) désigne la base canonique de $\mathbb{C}^2$ — car $\omega_n^h \neq \omega_n^{-h}$.
Or, $\mathbb{C}e_1$ et $\mathbb{C}e_2$ ne sont pas stables par $\rho_h(s)$. Contradiction.
On obtient donc $\dfrac{n}{2} - 1$ nouvelles représentations irréductibles, de caractères :

	r^k	sr^k
χ_h	$2\cos\left(\frac{2\pi hk}{n}\right)$	0

De plus, les caractères χ_h, où h parcourt $\left[\!\left[1, \dfrac{n}{2} - 1\right]\!\right]$, sont bel et bien différents.

Vérifions qu'on a toutes les représentations irréductibles ; on calcule : $4 \times 1^2 + \left(\dfrac{n}{2} - 1\right) \times 2^2 = 2n = \#\mathcal{D}_n$.

50. Je n'aime pas ce développement parce qu'il est répétitif, ennuyeux, et parce que si on veut le faire rigoureusement du début à la fin, on doit s'empêtrer dans de longues considérations triviales... Mais il bouche un trou.

Diffusion à titre gratuit uniquement.

Cas où n est impair : Cette fois-ci, on n'a que 2 représentations de degré 1.

	r^k	sr^k
ψ_1	1	1
ψ_2	1	-1

On définit, pour $h \in \left[\!\left[1, \frac{n-1}{2}\right]\!\right]$, les représentations ρ_h comme précédemment.
Elles sont toujours irréductibles et de caractères distincts (donc non-isomorphes). On obtient donc de nouveaux caractères irréductibles :

	r^k	sr^k
χ_h	$2\cos\left(\frac{2\pi hk}{n}\right)$	0

Enfin, on vérifie qu'on a bien fini le travail : $2 \times 1^2 + \dfrac{n-1}{2} \times 2^2 = 2n = \#\mathcal{D}_n$.

Références

[Pey] G. PEYRÉ – *L'algèbre discrète de la transformée de Fourier*, Ellipses, 2004.

Diffusion à titre gratuit uniquement.

Table de caractères de $\mathfrak{S}_4$

Leçons : 105, 107, 109, *104*

Librement inspiré de [Pey], paragraphe VIII.1.4

On va construire la table de caractères de $\mathfrak{S}_4$.

Étape 1 : On détermine les classes de conjugaison dans $\mathfrak{S}_4$. C'est facile : deux éléments sont conjugués si, et seulement si, ils ont le même type. On obtient donc :
- Type [1] : l'identité, 1 élément.
- Type [2] : les transpositions, $6 = \binom{4}{2}$ éléments.
- Type [2,2] : les doubles-transpositions, $3 = \frac{\binom{4}{2}}{2}$ éléments.
- Type [3] : les 3-cycles, $8 = 2 \times \binom{4}{3}$ éléments.
- Type [4] : les 4-cycles, $6 = 3!$ éléments.

Il y a donc 5 classes de conjugaison dans $\mathfrak{S}_4$, donc $\mathfrak{S}_4$ admet 5 caractères irréductibles.

	[1] 1	[2] 6	[2,2] 3	[3] 8	[4] 6

Étape 2 : Les deux premières lignes sont faciles à remplir : il s'agit des caractères associés à la représentation triviale et au morphisme signature.

	[1] 1	[2] 6	[2,2] 3	[3] 8	[4] 6
χ_1	1	1	1	1	1
χ_ε	1	-1	1	1	-1

Étape 3 : Intéressons maintenant à la représentation par permutations.
On note $\mathcal{B} = (e_1, \ldots, e_4)$ la base canonique de $\mathbb{C}^4$ et on définit la représentation par permutation par :

$$\rho_P : \left| \begin{array}{ccc} \mathfrak{S}_4 & \to & \mathrm{GL}\left(\mathbb{C}^4\right) \\ \sigma & \mapsto & \left(e_i \mapsto e_{\sigma(i)}\right) \end{array} \right. .$$

Cette représentation laisse stable $\mathrm{Vect}\{(1,1,1,1)\}$, dont $H := \left\{x \in \mathbb{C}^4 \middle| x_1 + \ldots + x_4 = 0\right\}$ est un supplémentaire stable ; elle induit une représentation ρ_S (appelée "représentation standard") sur H.
Et comme ρ_P induit la représentation triviale sur $\mathrm{Vect}\{(1,1,1,1)\}$, on a la relation : $\chi_P = \chi_1 + \chi_S$.
Pour savoir si χ_S est irréductible, on doit calculer son carré scalaire ; il est facile de calculer χ_P, car on voit que $\chi_P(\sigma)$ compte le nombre de 1 sur la diagonale de la matrice de la permutation σ, c'est-à-dire le nombre de points fixes de σ.
On a $\#\mathfrak{S}_4\langle\chi_S, \chi_S\rangle = 1 \times 3^2 + 6 \times 1^2 + 3 \times (-1)^2 + 8 \times 0^2 + 6 \times (-1)^2 = 24$ donc $\langle\chi_S, \chi_S\rangle = 1$.
Donc χ_S est irréductible, on le rajoute à la table.

	[1] 1	[2] 6	[2,2] 3	[3] 8	[4] 6
χ_1	1	1	1	1	1
χ_ε	1	-1	1	1	-1
χ_S	3	1	-1	0	-1

Diffusion à titre gratuit uniquement.

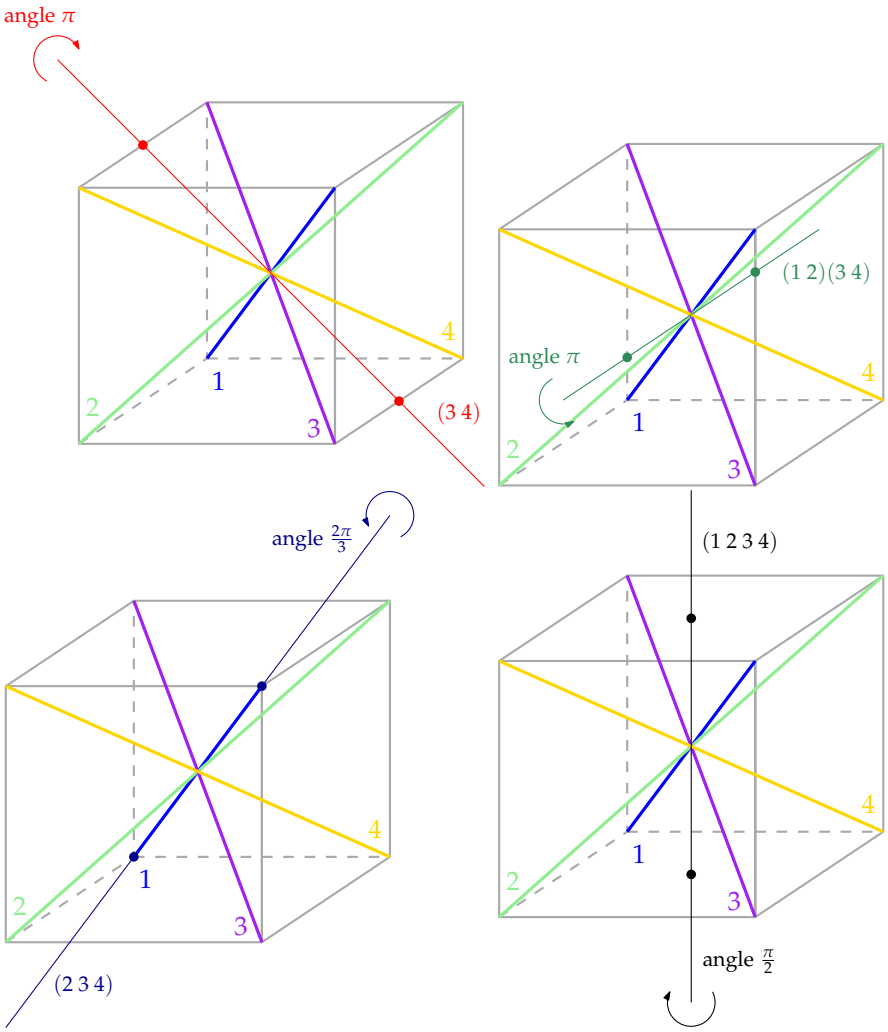
Étape 4 : Il nous reste deux caractères irréductibles, mais on peut déjà obtenir leur degré.

En effet, on a : $\sum_{i=1}^{5} n_i^2 = \#\mathfrak{S}_4$, où n_i désigne le degré du $i^{\text{ème}}$ caractère irréductible.

On en déduit que $n_4^2 + n_5^2 = 13$, d'où $n_4 = 2$ et $n_5 = 3$. [51]

	[1]	[2]	[2,2]	[3]	[4]
	1	6	3	8	6
χ_1	1	1	1	1	1
χ_ε	1	-1	1	1	-1
χ_S	3	1	-1	0	-1
	2				
	3				

Étape 5 : Faisons un peu de géométrie. On sait que $\mathfrak{S}_4 \simeq \text{Isom}^+(\text{Cube})$. [52]



Ces figures permettent[a], sachant que, dans $\mathbb{R}^3$, la trace d'une rotation vectorielle d'angle θ est $1 + 2\cos\theta$, d'obtenir les valeurs de χ_C, caractère associé à la représentation de $\mathfrak{S}_4$ comme groupe des isométries positives du cube :

- $\chi_C([1]) = 3$;
- $\chi_C([2]) = -1$;
- $\chi_C([2,2]) = -1$;
- $\chi_C([3]) = 0$;
- $\chi_C([4]) = 1$.

Pour montrer que χ_C est irréductible, il suffit de montrer que $\langle \chi_C, \chi_C \rangle = 1$, de la même manière qu'on l'a fait pour χ_S précédemment.

a. Je trouve que c'est mieux de faire un patron de cube pendant la préparation plutôt que de vouloir faire des dessins au tableau. Vous avez pensé que vous n'aurez peut-être pas toutes les couleurs que vous voulez pour votre dessin ? L'État n'a plus les moyens d'acheter massivement des craies colorées pour l'organisation de l'agrégation de mathématiques... Pour ma main, un cube de 5 cm d'arête, avec des petits rabats à coller, c'est nickel.

51. Ou l'inverse, mais ça on s'en fout.

52. Ou pas ! Les grandes diagonales du cube, qui relient deux sommets diamétralement opposés, sont les plus grandes distances qu'on puisse trouver entre deux points d'un cube. Nécessairement, les isométries conservant les distances, l'ensemble $\mathcal{D}$ des grandes diagonales est laissé stable par n'importe quelle isométrie. On obtient une action de $\text{Isom}^+(\text{Cube})$ sur $\mathcal{D}$ d'où un morphisme $\varphi : \left| \begin{array}{ccc} \text{Isom}^+(\text{Cube}) & \to & \mathfrak{S}_4 \\ g & \mapsto & g_{|\mathcal{D}} \end{array} \right.$. Il s'agit maintenant de montrer que φ est injectif ; soit g une isométrie du cube fixant chaque diagonale. On appelle $A_1, \dots, A_4$ les sommets de face du bas du cube, et pour $i \in [\![1,4]\!]$, B_i est le sommet du cube diamétralement opposé à A_i. Supposons que g fixe un sommet, sans perte de généralité, disons : $g(A_1) = A_1$. Comme $A_1A_2 \neq A_1B_2$, et comme g fixe (A_2B_2), nécessairement, $g(A_2) = A_2$ et $g(B_2) = B_2$. Similairement, $g(A_3) = A_3$. Mais (A_1, A_2, A_3, B_2) est un repère affine de $\mathbb{R}^3$. Donc $g = \text{Id}$. En revanche, si $\forall i \in [\![1,4]\!], g(A_i) = B_i$, alors on montre comme tout de suite que $s_0 g$ est l'identité, où s_0 désigne la symétrie centrale du cube. Mais on a alors une contradiction avec le fait que g soit une isométrie positive. En conclusion, φ est injective. Pour la surjectivité, il suffit d'exhiber la réalisation de chaque transposition (ce qui est un peu plus clair avec le premier des quatre dessins : il faut faire une rotation d'angle π autour d'axes rejoignant deux milieux d'arêtes opposées).

Diffusion à titre gratuit uniquement.

	[1]	[2]	[2,2]	[3]	[4]
	1	6	3	8	6
χ_1	1	1	1	1	1
χ_ε	1	-1	1	1	-1
χ_S	3	1	-1	0	-1
	2				
χ_C	3	-1	-1	0	1

Étape 6 : Enfin, on remplit la dernière ligne en utilisant : $\sum_{i=1}^{5} n_i \chi_i(s) = 0$ pour $s \neq \mathrm{Id}$.

	[1]	[2]	[2,2]	[3]	[4]
	1	6	3	8	6
χ_1	1	1	1	1	1
χ_ε	1	-1	1	1	-1
χ_S	3	1	-1	0	-1
χ_4	2	0	2	-1	0
χ_C	3	-1	-1	0	1

Références

[Pey] G. PEYRÉ – *L'algèbre discrète de la transformée de Fourier*, Ellipses, 2004.

Diffusion à titre gratuit uniquement.

Théorème de Burnside [53]

Leçons : 104, 106, 157, *153*

[X-ENS Al2], exercices 3.8, 2.28 et 1.10

Théorème

Soit G un sous-groupe de $\mathrm{GL}_n(\mathbb{C})$.
Si G est d'exposant fini (c'est-à-dire : $\exists N \in \mathbb{N}^*, \forall A \in G, A^N = \mathrm{I}_n$),
Alors G est fini.

Démonstration :

Étape 1 : On commence par montrer un lemme.

Lemme

Soit $A \in \mathcal{M}_n(\mathbb{C})$, telle que : $\forall k \in \mathbb{N}^*, \mathrm{tr}\left(A^k\right) = 0$;
Alors A est nilpotente.

Démonstration :
χ_A est scindé sur $\mathbb{C}$; par l'absurde, on suppose que A n'est pas nilpotente.
Alors A possède des valeurs propres non-nulles, qu'on note $\lambda_1, \dots, \lambda_r$, et de multiplicités respectives $n_1, \dots, n_r$, évidemment toutes non-nulles. Ainsi :
$\exists P \in \mathrm{GL}_n(\mathbb{C}), A = PTP^{-1}$, où $T \in \mathcal{M}_n(\mathbb{C})$ est triangulaire de diagonale $(\underbrace{\lambda_1, \dots, \lambda_1}_{n_1}, \dots, \underbrace{\lambda_r, \dots, \lambda_r}_{n_r}, 0, \dots, 0)$.

De plus, $\forall k \in \mathbb{N}^*, A^k = PT^kP^{-1}$, d'où $\mathrm{tr}\left(A^k\right) = \sum_{i=1}^{r} n_i \lambda_i^k = 0$.

Ainsi, $\begin{pmatrix} n_1 \\ n_2 \\ \vdots \\ n_r \end{pmatrix}$ est solution du système linéaire : $\begin{pmatrix} \lambda_1 & \lambda_2 & \dots & \lambda_r \\ \lambda_1^2 & \lambda_2^2 & \dots & \lambda_r^2 \\ \vdots & \vdots & & \vdots \\ \lambda_1^r & \lambda_2^r & \dots & \lambda_r^r \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_r \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$.

53. En 1902, William Burnside pose le problème suivant "Est-ce que les groupes de torsion de type fini sont tous finis ?" (c'est le problème de Burnside).

Un rappel : un groupe de torsion est un groupe dont tous les éléments sont d'ordre fini. Par Lagrange, les groupes finis sont de torsion ; mais il existe des groupes infinis qui soient de torsion, comme $\mathbb{Q}/\mathbb{Z}$ par exemple.

La question de Burnside se réécrit alors "Est ce que les groupes possédant un nombre fini de générateurs et dont tous les éléments sont d'ordre fini sont tous finis ?"

Conscient de la difficulté de sa conjecture, il en émet une plus faible, dite "bornée" : "Est-ce que les groupes d'exposant fini possédant un nombre fini de générateurs sont tous finis ?"

Remarque : $\mathbb{Q}/\mathbb{Z}$ est de torsion, mais il n'est pas d'exposant fini.

En 1905, il démontre que tout sous-groupe d'exposant fini de $\mathrm{GL}_n(\mathbb{C})$ est fini, sans supposer qu'il soit de type fini (c'est ce développement).

Un énoncé équivalent est que toute représentation d'un groupe d'exposant fini dans un espace vectoriel complexe de dimension finie est d'image finie. Ceci met en évidence la difficulté de construire un contre-exemple à sa conjecture : il faut, d'après son théorème, qu'un tel groupe n'ait aucune représentation fidèle de degré fini.

En 1911, Issai Schur va même plus loin en montrant que tout sous-groupe de torsion de type fini de $\mathrm{GL}_n(\mathbb{C})$ est fini, donnant un résultat analogue à celui de 1905 pour la version "non-bornée" de la conjecture. C'est en 1964 que la conjecture de Burnside est réfutée pour sa version "non-bornée", puis en 1968 pour la version "bornée". Ces résultats des années 1960 ont apporté des contre-exemples uniquement pour des groupes d'exposant impair ; c'est en 1992 qu'un contre-exemple d'exposant pair a été construit. On sait aujourd'hui construire pour tout entier n impair supérieur ou égal à 665 un groupe infini de type fini et d'exposant n, et on sait qu'il existe un entier n supérieur ou égal à 2^{48} et divisible par 2^9, tel qu'il existe un groupe infini de type fini et d'exposant n.

D'autres résultats ont été montré à propos de cette conjecture, mais il reste encore aujourd'hui des problèmes ouverts à ce sujet.

En conclusion, l'intérêt du théorème de 1905 est d'avoir apporté une condition nécessaire que doit vérifier un contre-exemple à la conjecture bornée de 1902.

Diffusion à titre gratuit uniquement.

Par déterminant de Vandermonde [54], le déterminant de ce système vaut $\prod_{i=1}^{r}\lambda_i \prod_{1\leqslant i<j\leqslant r}(\lambda_j-\lambda_i)\neq 0$.

On a donc unicité de la solution, et nécessairement : $n_1=n_2=\ldots=n_r=0$. ■

Étape 2 : Soit $(M_i)_{1\leqslant i\leqslant m}$ une base de $\mathrm{Vect}(G)$ et $f : \left|\begin{array}{ccc} G & \to & \mathbb{C}^m \\ A & \mapsto & (\mathrm{tr}\,(AM_i))_{1\leqslant i\leqslant m}\end{array}\right.$.

On va montrer que si $f(A)=f(B)$, alors $AB^{-1}-\mathrm{I}_n$ est nilpotente.
Supposons donc $f(A)=f(B)$; par linéarité de la trace on a : $\forall M\in\mathrm{Vect}(G), \mathrm{tr}(AM)=\mathrm{tr}(BM)$.
Soit $D=AB^{-1}\in G$, $k\in\mathbb{N}^*$; on a $B^{-1}D^{k-1}\in G$, d'où :

$$\mathrm{tr}\left(D^k\right)=\mathrm{tr}\left(AB^{-1}D^{k-1}\right)=\mathrm{tr}\left(BB^{-1}D^{k-1}\right)=\mathrm{tr}\left(D^{k-1}\right).$$

On en déduit : $\forall k\in\mathbb{N}, \mathrm{tr}\left(D^k\right)=\mathrm{tr}\,(\mathrm{I}_n)=n$.
Puis, pour $k\in\mathbb{N}^*$,

$$\mathrm{tr}\left((D-\mathrm{I}_n)^k\right)=\mathrm{tr}\left(\sum_{j=0}^{k}\binom{j}{k}(-1)^jD^{k-j}\right)=\sum_{j=0}^{k}\binom{j}{k}(-1)^j\mathrm{tr}\left(D^{k-j}\right)=n\sum_{j=0}^{k}\binom{j}{k}(-1)^j=n(1-1)^k=0.$$

Le lemme précédent permet alors de montrer que $D-\mathrm{I}_n$ est nilpotente.

Étape 3 : Montrons désormais que f est injective, ie $A=B$.
Soit $N\in\mathbb{N}^*$ l'exposant de G, ainsi, X^N-1 annule toutes les matrices de G.
C'est un polynôme scindé à racines simples dans $\mathbb{C}$, et donc : $\forall M\in G, M$ est diagonalisable.
Donc D est diagonalisable, mais $\forall P\in\mathrm{GL}_n(\mathbb{C}), P\,(D-\mathrm{I}_n)\,P^{-1}=PDP^{-1}-\mathrm{I}_n$.
Donc $D-\mathrm{I}_n$ est diagonalisable.
Mais $D-\mathrm{I}_n$ est déjà nilpotente, dès lors : $D-\mathrm{I}_n=0$. D'où $A=B$.

Étape 4 : Reste à conclure.
On note $X=\{\mathrm{tr}A|A\in G\}$; on a : $f(G)\subset X^m$.
Comme les valeurs propres des éléments de G sont des racines $N^{\text{èmes}}$ de l'unité, on obtient que X est fini.
Puis, par injectivité de f, on obtient que G est fini. ■

Références

[X-ENS Al2] S. FRANCINOU, H. GIANELLA et S. NICOLAS – *Oraux X-ENS Algèbre 2*, 2ème éd., Cassini, 2009.

54. Calculons le déterminant $V(x_1,\ldots,x_n)=\left|\begin{array}{ccccc} 1 & x_1 & x_1^2 & \ldots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \ldots & x_2^{n-1} \\ \vdots & \vdots & \vdots & & \vdots \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_n & x_n^2 & \ldots & x_n^{n-1}\end{array}\right|$.

Si deux des x_i sont égaux, alors $V(x_1,\ldots,x_n)=0$, car deux lignes sont identiques.
Supposons donc les x_i deux à deux distincts, on pose : $P(X)=V(x_1,\ldots,x_{n-1},X)$.
C'est un polynôme de degré au plus $n-1$; en développant par rapport à la dernière ligne, on obtient que le coefficient en X^{n-1} est $V(x_1,\ldots,x_{n-1})$.
On a : $\forall i\in[\![1,n-1]\!], P(x_i)=0$, car alors deux lignes sont égales.
Les x_i étant deux à deux distincts, on a : $\prod_{i=1}^{n-1}(X-x_i)\Big|P$, et $\prod_{i=1}^{n-1}(X-x_i)$ est un polynôme unitaire de degré $n-1$.
Il en résulte que : $P(X)=V(x_1,\ldots,x_{n-1})\prod_{i=1}^{n-1}(X-x_i)$; en particulier : $P(x_n)=V(x_1,\ldots,x_{n-1})\prod_{i=1}^{n-1}(x_n-x_i)$.
Par récurrence, on en tire la formule de Vandermonde : $V(x_1,\ldots,x_n)=\prod_{1\leqslant i<j\leqslant n}\left(x_j-x_i\right)$.

Diffusion à titre gratuit uniquement.

Théorème de Frobenius-Zolotarev [55]

Leçons : 103, 106, 123, 152, *105*

[OA], exercice 5.4

Théorème

Soient p premier impair, $n \geqslant 1$ un entier.
Alors on a :

$$\forall u \in \mathrm{GL}_n\left(\mathbb{F}_p\right), \varepsilon(u) = \left(\frac{\det u}{p}\right)$$

On rappelle que :
- $\varepsilon(u)$ est la signature de u, vu comme permutation de $\mathbb{F}_p^n$;
- le symbole de Legendre est désigné par : $\left(\frac{a}{p}\right) = \begin{cases} 0 & \text{si } a \equiv 0\ [p] \\ 1 & \text{si } a \text{ est un carré dans } \mathbb{F}_p \\ -1 & \text{sinon} \end{cases}$.

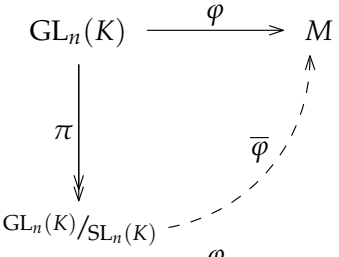
Démonstration :
On va montrer que $\varepsilon = \left(\frac{\cdot}{p}\right) \circ \det$ est une factorisation de la signature.

Lemme 1

Soit K un corps et M un groupe abélien. On suppose $K \neq \mathbb{F}_2$ ou $n \neq 2$.
Alors tout morphisme de groupes $\varphi : \mathrm{GL}_n(K) \to M$ se factorise par le déterminant, c'est-à-dire : il existe un unique morphisme de groupes $\delta : K^\times \to M$ tel que $\varphi = \delta \circ \det$.

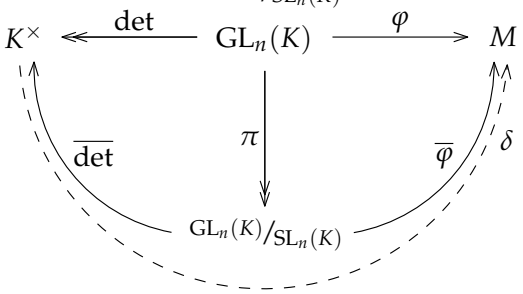
Démonstration du lemme 1 :
Comme $K \neq \mathbb{F}_2$ ou $n \neq 2$, on a : $\mathcal{D}(\mathrm{GL}_n(K)) = \mathrm{SL}_n(K)$.

Pour $x, y \in \mathrm{GL}_n(K)$, $\varphi([x,y]) = [\varphi(x), \varphi(y)] = \mathrm{e}$ car M est abélien.
Or, $\mathcal{D}(\mathrm{GL}_n(K))$ est engendré par les commutateurs ; on en déduit : $\mathcal{D}(\mathrm{GL}_n(K)) \subset \mathrm{Ker}\varphi$.
On a donc la factorisation suivante, où $\overline{\varphi}$ est l'unique morphisme tel que $\varphi = \overline{\varphi} \circ \pi$:



Et comme $\det : \mathrm{GL}_n(K) \to K^\times$ est un morphisme surjectif de noyau $\mathrm{SL}_n(K)$, on peut compléter ce diagramme commutatif pour le suivant, où $\overline{\det}$ est un isomorphisme (d'après le 1er théorème d'isomorphisme).

On a : $\varphi = \delta \circ \det$, où $\delta = \overline{\varphi} \circ \left(\overline{\det}\right)^{-1}$, et δ est l'unique morphisme de groupes de $K^\times$ vers M, car $\det$ est surjectif.



55. Donnons de ce résultat une application : le calcul du symbole de Legendre $\left(\frac{2}{p}\right)$. Posons u : $\left|\begin{array}{ccc} \mathbb{F}_p & \to & \mathbb{F}_p \\ x & \mapsto & 2x \end{array}\right.$. On a $\det u = 2$, calculons désormais la signature de la permutation u de $\mathbb{F}_p$. On construit le tableau suivant :

x	0	1	2	…	$\frac{p-1}{2}$	$\frac{p+1}{2}$	…	$p-2$	$p-1$
$u(x)$	0	2	4	…	$p-1$	1	…	$p-4$	$p-2$

Il s'agit de calculer le nombre d'inversions engendrées par cette permutation ; soit $k \geqslant \frac{p+1}{2}$, l'élément k voit sa position relative à $p-k$ éléments inversée par u. Le nombre total d'inversions est alors : $\sum_{k=\frac{p+1}{2}}^{p-1} p-k = \sum_{j=1}^{\frac{p-1}{2}} j = \frac{\frac{p-1}{2}\frac{p+1}{2}}{2} = \frac{p^2-1}{8}$. Et donc, $\varepsilon(u) = (-1)^{\frac{p^2-1}{8}}$, d'où, par Frobenius-Zolotarev : $\left(\frac{2}{p}\right) = (-1)^{\frac{p^2-1}{8}}$. La même méthode permettrait de calculer $\left(\frac{-1}{p}\right)$, mais dans ce cas, il y a plus efficace.

Diffusion à titre gratuit uniquement.

■

Dans le cadre de ce théorème, ce lemme dit que : il existe un unique morphisme de groupes $\delta : \mathbb{F}_p^\times \to \{\pm 1\}$, tel que $\varepsilon = \delta \circ \det$.

Lemme 2

Soit p premier impair.
Le symbole de Legendre est l'unique morphisme de groupes non-trivial de $\mathbb{F}_p^\times$ dans $\{\pm 1\}$.

Démonstration du lemme 2 :

Comme p est premier impair $\left(\frac{a}{p}\right) = a^{\frac{p-1}{2}}$ dans $\mathbb{F}_p$, d'où $\left(\frac{\cdot}{p}\right)$ est un morphisme de groupes.

C'est un morphisme non-trivial car $\Phi : \left| \begin{array}{ccc} \mathbb{F}_p^\times & \to & \mathbb{F}_p^\times \\ x & \mapsto & x^2 \end{array} \right.$ est non-injective, car $1^2 = (-1)^2$ et $1 \neq -1$ (comme $p \geqslant 3$), donc non-surjective.

Soit $\alpha : \mathbb{F}_p^\times \to \{\pm 1\}$ un morphisme de groupes non-trivial.

Nécessairement, on a : $\left(\mathbb{F}_p^\times : \operatorname{Ker} \alpha\right) = \#\operatorname{Im} \alpha = 2$ par le 1er théorème d'isomorphisme.

Or $\mathbb{F}_p^\times$ est cyclique [56] donc possède un unique sous-groupe d'indice 2 qu'on appelle H.

On a ainsi la partition $\mathbb{F}_p^\times = H \sqcup xH$, où $x \notin H$ et $\alpha(y) = \left\{ \begin{array}{ll} 1 & \text{si } y \in H \\ -1 & \text{sinon} \end{array} \right.$ donc α est entièrement déterminé.

Donc il existe un unique morphisme de groupes non-trivial de $\mathbb{F}_p^\times$ dans $\{\pm 1\}$: le symbole de Legendre. ■

Il reste alors à montrer que ε est non-trivial : $\varepsilon = \delta \circ \det$ impliquera que δ est non-trivial, puis que $\delta = \left(\frac{\cdot}{p}\right)$.

Notons $q = p^n$. Comme $\mathbb{F}_p$-espaces vectoriels, $\mathbb{F}_q$ et $\mathbb{F}_p^n$ sont isomorphes.

Il suffit donc d'exhiber une bijection $\mathbb{F}_p$-linéaire de $\mathbb{F}_q$ de signature -1.

$\mathbb{F}_q^\times$ est cyclique, notons g un de ses générateurs. La bijection $x \mapsto gx$ de $\mathbb{F}_q$ fixe 0 et donc agit sur $\mathbb{F}_q^\times$ comme la permutation $\left(g \;\; g^2 \;\; \dots \;\; g^{q-1}\right)$.

Sa signature est donc $(-1)^q = -1$, car q est impair. Et donc ε n'est pas trivial, d'où $\varepsilon = \left(\frac{\cdot}{p}\right) \circ \det$. ■

Références

[OA] V. BECK, J. MALICK et G. PEYRÉ – *Objectif Agrégation*, 2e éd., H&K, 2005.

56. Pour la cyclicité de $\mathbb{F}_q^\times$, on renvoie à la page 139.

Diffusion à titre gratuit uniquement.

Théorème de Kronecker et application

Leçons : 143, 144 [57], *102*

[X-ENS Al1], exercice 5.33
[Szp], théorème 10.80

Théorème

Soit $P \in \mathbb{Z}[X]$, unitaire, et dont les racines sont de module $\leqslant 1$.
On suppose que $P(0) \neq 0$.
Alors les racines de P sont des racines de l'unité.

Démonstration :

→ On note $z_1, \dots, z_n \in \mathbb{C}$ les racines de P, comptées avec leur multiplicité, où $n = \deg P \geqslant 1$.
On note $\sigma_1, \dots, \sigma_n$ les fonctions symétriques élémentaires des z_i.
On a donc : $P = X^n - \sigma_1 X^{n-1} + \dots + (-1)^n \sigma_n$.
Comme $P \in \mathbb{Z}[X]$, on en déduit que $\forall i \in [\![1, n]\!], \sigma_i \in \mathbb{Z}$.
Mais on sait que $\forall i \in [\![1, n]\!], |z_i| \leqslant 1$, et donc, par inégalité triangulaire :

$$\forall p \in [\![1, n]\!], |\sigma_p| = \left| \sum_{I \in \mathcal{P}_p([\![1,n]\!])} \prod_{i \in I} z_i \right| \leqslant \sum_{I \in \mathcal{P}_p([\![1,n]\!])} 1 = \#\mathcal{P}_p([\![1, n]\!]) = \binom{p}{n}.$$

Ainsi, pour tout $n \in \mathbb{N}^*$, l'ensemble $\Omega_n := \left\{ P \in \mathbb{Z}[X] \middle| P \text{ unitaire, } \deg P = n \text{ et } \mathcal{R}ac(P) \subset \overline{\mathcal{D}(0,1)} \right\}$ est fini.

→ On pose, pour $k \in \mathbb{N}^*$, $P_k = \prod_{i=1}^{n} \left(X - z_i^k \right) \in \mathbb{C}[X]$ et $Q_k = X^k - Y \in \mathbb{Z}[X, Y]$.
P_k est alors un polynôme unitaire de degré n, à racines de module $\leqslant 1$; on veut montrer que $P_k \in \Omega_n$, montrons donc que $P_k \in \mathbb{Z}[X]$.
On pose $R_k(Y) = \mathrm{Res}_X \left(P(X), Q_k(X, Y) \right) \in \mathbb{Z}[Y]$.
La formule du résultant à l'aide des racines fournit :

$$R_k(Y) = 1^k \prod_{i=1}^{n} Q_k \left(z_i, Y \right) = \prod_{i=1}^{n} \left(z_i^k - Y \right) = (-1)^n P_k(Y).$$

Ainsi, P_k est à coefficients entiers et donc $P_k \in \Omega_n$.

→ Or, $\Omega_n < \infty$ et $\forall P \in \Omega_n, \#\mathcal{R}ac(P) < \infty$ donc l'ensemble E des racines des éléments de Ω_n est fini.
En conséquence, $\forall i \in [\![1, n]\!]$, l'application $\left| \begin{array}{ccc} \mathbb{N}^* & \to & E \\ k & \mapsto & z_i^k \end{array} \right.$ ne peut pas être injective.
Donc $\exists k \neq l \in \mathbb{N}^*, z_i^k = z_i^l$.
Comme $z_i \neq 0$, on a $z_i^{k-l} = 1$ et donc z_i est une racine de l'unité. ■

57. Dans la leçon 144, le second tiers du théorème évolue un peu.
On pose, pour $k \in \mathbb{N}^*$, $P_k = \prod_{i=1}^{n} \left(X - z_i^k \right) \in \mathbb{C}[X]$.
P_k est alors un polynôme unitaire de degré n, à racines de module $\leqslant 1$; on veut montrer que $P_k \in \Omega_n$, montrons donc que $P_k \in \mathbb{Z}[X]$.
On note $\Sigma_1, \dots, \Sigma_n$ les fonctions symétriques élémentaires des z_i^k.
Le coefficient de X^{n-r} dans P_k est alors $(-1)^r \Sigma_r$, qui est un polynôme symétrique en $z_1, \dots, z_n$ à coefficients dans $\mathbb{Z}$; c'est donc aussi un polynôme de $\mathbb{Z}[\sigma_1, \dots, \sigma_n]$.
Ainsi, P_k est à coefficients entiers et donc $P_k \in \Omega_n$.

Diffusion à titre gratuit uniquement.

Corollaire

Soit $P \in \mathbb{Z}[X]$, unitaire et irréductible, à racines de module $\leqslant 1$.
Alors $P = X$ ou P est un polynôme cyclotomique.

Démonstration :

Supposons que $P \neq X$.
P étant irréductible, on en déduit que $P(0) \neq 0$.
On peut donc appliquer le théorème de Kronecker : les racines de P sont des racines de l'unité.
Ainsi, il existe $N \in \mathbb{N}^*$, tel que : $\forall z \in \mathcal{R}ac(P), z^N - 1 = 0$.
Mais P est à racines simples ; en effet, dans le cas contraire, $P \wedge P'$ serait un polynôme non-constant divisant P, ce qui contredirait l'irréductibilité de P.
Ainsi, $P \big| X^N - 1$.
Or $X^N - 1 = \prod_{d|N} \Phi_d$ est la décomposition en facteurs irréductibles de $X^N - 1$ dans $\mathbb{Z}[X]$.
P étant irréductible, P est un polynôme cyclotomique. ■

Corollaire

Soit $P \in \mathbb{Z}[X]$, unitaire, à racines de module $\leqslant 1$.
Alors P est produit d'une puissance de X et de polynômes cyclotomiques.

Références

[X-ENS Al1] S. FRANCINOU, H. GIANELLA et S. NICOLAS – *Oraux X-ENS Algèbre 1*, 2e éd., Cassini, 2007.
[Szp] A. SZPIRGLAS – *Algèbre pour la L3*, Pearson Éducation, 2009.

Diffusion à titre gratuit uniquement.

Théorème de Molien [58]

Leçons : 101, 124, 142, *104, 106, 107, 140, 151, 152, 154, 155*

Très librement inspiré de [Lei], page 95

Théorème

On note $M = \mathbb{C}[X_1, \dots, X_n]$ et $M_k \subset M$ le sous-espace des polynômes homogènes de degré k.
Soit G un sous-groupe fini de $\mathrm{GL}_n(\mathbb{C})$.
Pour $g \in \mathrm{GL}_n(\mathbb{C})$, on définit : $\sigma_g : \left| \begin{array}{ccc} M & \to & M \\ P(\underline{X}) & \mapsto & P(\underline{X}g) \end{array} \right.$, où $\underline{X} = (X_1 \ \dots \ X_n)$ est le vecteur-ligne des indéterminées.
Alors σ_g induit un automorphisme sur chaque M_k, on note $m_k = \dim M_k$ et $m_k(G) = \dim M_k^G$, où $M_k^G = \left\{P \in M_k \middle| \forall g \in G, \sigma_g(P) = P\right\}$.
Et on a l'égalité dans $\mathbb{C}[\![Z]\!]$: [59]

$$\frac{1}{\#G} \sum_{g \in G} \frac{1}{\det(\mathrm{Id} - gZ)} = \sum_{k=0}^{\infty} m_k(G) Z^k$$

Démonstration :

Étape 1 : σ_g induit un automorphisme sur chaque M_k.

Soient $g, g' \in \mathrm{GL}_n(\mathbb{C})$, on a : $\sigma_{gg'} = \sigma_g \sigma_{g'}$ et $\sigma_{\mathrm{Id}} = \mathrm{Id}$, donc $\sigma_g \in \mathrm{GL}(M)$.
Soit $k \in \mathbb{N}$, on remarque que l'image par σ_g d'un monôme de degré total k est un polynôme homogène de degré k.
Par linéarité, on en déduit que $\sigma_g(M_k) \subset M_k$.
Or $\sigma_{g|_{M_k}}$ est injectif car σ_g l'est et $\dim M_k < \infty$, donc $\sigma_{g|_{M_k}} \in \mathrm{GL}(M_k)$.

Étape 2 : Montrons l'égalité de séries formelles $\dfrac{1}{\det(\mathrm{Id} - gZ)} = \displaystyle\sum_{k=0}^{\infty} \mathrm{tr}\left(\sigma_{g|_{M_k}}\right) Z^k$ pour $g \in G$.

Par le théorème de Lagrange, on a : $g^{\#G} = \mathrm{Id}$.
Or $X^{\#G} - 1$ est scindé à racines simples, donc g est diagonalisable.
Donc $\exists u \in \mathrm{GL}_n(\mathbb{C}), g = udu^{-1}$, avec $d = \mathrm{diag}(\lambda_1, \dots, \lambda_n)$.
Ainsi, $\sigma_{g|_{M_k}}$ et $\sigma_{d|_{M_k}}$ sont semblables, d'où $\mathrm{tr}\left(\sigma_{g|_{M_k}}\right) = \mathrm{tr}\left(\sigma_{d|_{M_k}}\right)$.
Mais aussi :

$$\frac{1}{\det(\mathrm{Id} - gZ)} = \frac{1}{\det(\mathrm{Id} - dZ)} = \prod_{i=1}^{n} \frac{1}{1 - \lambda_i Z} = \prod_{i=1}^{n} \sum_{k=0}^{\infty} \lambda_i^k Z^k = \sum_{k=0}^{\infty} v_k Z^k$$

où on a posé $v_k = \displaystyle\sum_{\substack{r_1, \dots r_n \in \mathbb{N} \\ r_1 + \dots + r_n = k}} \lambda_1^{r_1} \dots \lambda_n^{r_n}$.

Or :

$$\sigma_{d|_{M_k}}\left(X_1^{r_1} \dots X_n^{r_n}\right) = (\lambda_1 X_1)^{r_1} \dots (\lambda_n X_n)^{r_n} = \left(\lambda_1^{r_1} \dots \lambda_n^{r_n}\right) X_1^{r_1} \dots X_n^{r_n}$$

58. Le jury veut que vous soyez capables d'en parler si vous présentez ce développement ! Un lien vers un document rédigé par Arnaud STOCKER pour vous donner une idée de ce à quoi ce théorème peut servir.

59. Même si je le déconseille parce qu'on n'a pas besoin d'arguments analytiques dans ce développement, quand on veut travailler avec des séries entières, on doit ajouter le lemme suivant, qui justifie que toutes les séries entières en jeu ont des rayons de convergence supérieurs ou égaux à 1 :

Lemme

Si $|z| < 1$, alors $\left(\dfrac{1}{1-z}\right)^n = \displaystyle\sum_{k=0}^{\infty} m_k z^k$.

En effet, M_k admet pour base $\left(X_1^{r_1} \dots X_n^{r_n}\right)_{\substack{r_1, \dots, r_n \in \mathbb{N} \\ r_1 + \dots + r_n = k}}$; donc $m_k = \dim M_k = \#\{(r_1, \dots, r_n) \in \mathbb{N}^n | r_1 + \dots + r_n = k\}$.

Et pour $|z| < 1$: $\left(\dfrac{1}{1-z}\right)^n = \left(\displaystyle\sum_{k=0}^{\infty} z_k\right)^n = \sum_{k=0}^{\infty} \left(\sum_{\substack{r_1, \dots, r_n \in \mathbb{N} \\ r_1 + \dots + r_n = k}} 1\right) z^k = \sum_{k=0}^{\infty} m_k z^k$.

Diffusion à titre gratuit uniquement.

Et comme $\left(X_1^{r_1}\dots X_n^{r_n}\right)_{\substack{r_1,\dots,r_n\in\mathbb{N}\\ r_1+\dots+r_n=k}}$ est une base de M_k, on obtient : $v_k = \operatorname{tr}\left(\sigma_{d\,|M_k}\right)$.

Donc $\dfrac{1}{\det(\mathrm{Id}-gZ)} = \displaystyle\sum_{k=0}^{\infty} \operatorname{tr}\left(\sigma_{g\,|M_k}\right) Z^k$ pour tout $g \in G$.

Étape 3 : Concluons.

Lemme

Soit V un $\mathbb{C}$-ev de dimension $n < \infty$, $\varphi : G \to \mathrm{GL}(V)$ un morphisme de groupes et G un groupe fini.
On note $V^G = \{v \in V | \forall g \in G, \varphi(g)(v) = v\}$.
On a alors :

$$\frac{1}{\#G}\sum_{g\in G} \operatorname{tr}(\varphi(g)) = \dim V^G$$

Démonstration :

On pose $p_G = \dfrac{1}{\#G}\displaystyle\sum_{g\in G}\varphi(g)$. On a : $\forall h \in G, \varphi(h)\circ p_G = \dfrac{1}{\#G}\displaystyle\sum_{g\in G}\varphi(h)\circ\varphi(g) = \dfrac{1}{\#G}\displaystyle\sum_{g\in G}\varphi(hg) = p_G$.

Donc $\forall v \in V, \forall h \in G, \varphi(h)\,(p_G(v)) = p_G(v)$, donc $p_G(V) \subset V^G$.

Soit $v \in V^G$, $p_G(v) = \dfrac{1}{\#G}\displaystyle\sum_{g\in G} v = v$ et donc $p_G(V) = V^G$.

Soit $v \in V$, $p_G\,(p_G(v)) = p_G(v)$, car $p_G(v) \in p_G(V) = V^G$ et donc p_G est un projecteur.

Dès lors : $\dim V^G = \operatorname{rg} p_G = \operatorname{tr} p_G = \dfrac{1}{\#G}\displaystyle\sum_{g\in G}\operatorname{tr}(\varphi(g))$. ■

Ici, $\forall k \in \mathbb{N}, \sigma^{(k)} : \left|\begin{array}{ccc} G & \to & \mathrm{GL}\,(M_k) \\ g & \mapsto & \sigma_{g\,|M_k}\end{array}\right.$ est un morphisme de groupes.

Par le lemme, on a : $\dim M_k^G = m_k(G) = \dfrac{1}{\#G}\displaystyle\sum_{g\in G}\operatorname{tr}\left(\sigma_{g\,|M_k}\right)$.

Par conséquent :

$$\frac{1}{\#G}\sum_{g\in G}\frac{1}{\det(\mathrm{Id}-gZ)} = \frac{1}{\#G}\sum_{g\in G}\sum_{k=0}^{\infty}\operatorname{tr}\left(\sigma_{g\,|M_k}\right)Z^k = \sum_{k=0}^{\infty}\left(\frac{1}{\#G}\sum_{g\in G}\operatorname{tr}\left(\sigma_{g\,|M_k}\right)Z^k\right) = \sum_{k=0}^{\infty} m_k(G)Z^k$$ ■

Références

[Lei] É. LEICHTNAM – *Exercices corrigés de mathématiques Polytechnique-ENS (Tome algèbre et géométrie)*, Ellipses, 1999.

Diffusion à titre gratuit uniquement.

Théorème des deux carrés

Leçons : 120, 121, 122, 126

[Per], partie II.6
[Duv], partie 6.1

Théorème

Soit p un nombre premier impair, on note $\Sigma = \left\{n \in \mathbb{N} \middle| \exists a, b \in \mathbb{N}, n = a^2 + b^2\right\}$.
On a : $p \in \Sigma \Leftrightarrow p \equiv 1\,[4]$.

Démonstration :

Pour commencer, quelques mots sur $\mathbb{Z}[\mathrm{i}]$: on définit la "norme" N : $\left|\begin{array}{ccc} \mathbb{Z}[\mathrm{i}] & \to & \mathbb{N} \\ z = a + \mathrm{i}b & \mapsto & z\overline{z} = a^2 + b^2 \end{array}\right.$;
alors N est multiplicative, ce qui signifie que $N(zz') = N(z)N(z')$ pour tous $z, z' \in \mathbb{Z}[\mathrm{i}]$.

Lemme 1

On a : $\mathbb{Z}[\mathrm{i}]^\times = \{\pm 1, \pm \mathrm{i}\}$.

Démonstration du lemme 1 :

$\subset$ Si $z \in \mathbb{Z}[\mathrm{i}]^\times$, alors $N(z)N\left(z^{-1}\right) = N(1) = 1$, donc $N(z) = 1$.
Or $z = a + \mathrm{i}b$, avec $a, b \in \mathbb{Z}$, donc $a^2 + b^2 = 1$ et on a ($a = 0$ et $b = \pm 1$) ou ($a = \pm 1$ et $b = 0$).

$\supset$ Cette vérification est immédiate... ■

Lemme 2

On a l'équivalence : $p \in \Sigma \Leftrightarrow p$ est réductible dans $\mathbb{Z}[\mathrm{i}]$.

Démonstration du lemme 2 :

$\Rightarrow$ Si $p = a^2 + b^2$, alors dans $\mathbb{Z}[\mathrm{i}]$, $p = (a + \mathrm{i}b)(a - \mathrm{i}b)$.
Comme $N(a + \mathrm{i}b) = N(a - \mathrm{i}b) = p > 1$, on sait que $a + \mathrm{i}b$, $a - \mathrm{i}b \notin \mathbb{Z}[\mathrm{i}]^\times$ et donc p est réductible.

$\Leftarrow$ Si $p = zz'$ dans $\mathbb{Z}[\mathrm{i}]$ avec $z, z' \notin \mathbb{Z}[\mathrm{i}]^\times$, on a : $N(p) = N(z)N(z') = p^2$.
Mais on sait que $N(z) \neq 1 \neq N(z')$, donc $N(z) = p$. ■

Comme $\mathbb{Z}[\mathrm{i}]$ est factoriel [60], par le lemme d'Euclide, on a :

$$p \text{ réductible dans } \mathbb{Z}[\mathrm{i}] \Leftrightarrow (p) \text{ non-premier dans } \mathbb{Z}[\mathrm{i}] \Leftrightarrow {}^{\mathbb{Z}[\mathrm{i}]}\!/_{(p)} \text{ non-intègre}$$

Mais comme $\mathbb{Z}[\mathrm{i}] \simeq {}^{\mathbb{Z}[X]}\!/_{\left(X^2+1\right)}$, on a les isomorphismes suivants :

$$\mathbb{Z}[\mathrm{i}]/_{(p)} \underset{61}{\simeq} \mathbb{Z}[X]/_{\left(X^2+1,\,p\right)} \simeq \left(\mathbb{Z}[X]/_{(p)}\right)/_{\left(\overline{X^2+1}\right)} \simeq \mathbb{F}_p[X]/_{\left(X^2+1\right)}$$

En conséquence, p est réductible dans $\mathbb{Z}[\mathrm{i}] \Leftrightarrow {}^{\mathbb{F}_p[X]}\!/_{\left(X^2+1\right)}$ non-intègre
$\Leftrightarrow X^2 + 1$ réductible dans $\mathbb{F}_p[X]$
$\Leftrightarrow X^2 + 1$ a une racine dans $\mathbb{F}_p$
$\Leftrightarrow -1$ est un carré dans $\mathbb{F}_p$
$\Leftrightarrow (-1)^{\frac{p-1}{2}} = \left(\dfrac{-1}{p}\right) = 1 \Leftrightarrow p \equiv 1\,[4]$ ■

60. Le plus simple pour montrer la factorialité, c'est de montrer que $\mathbb{Z}[\mathrm{i}]$ est euclidien pour la norme N, puis de dire que les anneaux euclidiens sont factoriels (voir en page 136).

61. Je tape les explications pour un isomorphisme, adaptez ceci pour trouver les autres. Ce passage me semble absolument indispensable à savoir rédiger pour pouvoir présenter ce développement.
Notons $\pi_{X^2+1} : \mathbb{Z}[X] \to {}^{\mathbb{Z}[X]}\!/_{\left(X^2+1\right)}$ et $\pi_{\overline{p}} : {}^{\mathbb{Z}[X]}\!/_{\left(X^2+1\right)} \to \left({}^{\mathbb{Z}[X]}\!/_{\left(X^2+1\right)}\right)/_{(\overline{p})}$ les projections canoniques.
Alors $\mathrm{Ker}\,\pi_{\overline{p}} \circ \pi_{X^2+1} = \left\{f \in \mathbb{Z}[X] \middle| \exists u \in \mathbb{Z}[X], \overline{f} = \overline{pu}\right\} = \left\{f \in \mathbb{Z}[X] \middle| \exists u, v \in \mathbb{Z}[X], f = pu + \left(X^2+1\right)v\right\} = \left(p, X^2+1\right)$.
En conséquence, ${}^{\mathbb{Z}[X]}\!/_{\left(p,\,X^2+1\right)} \simeq \left({}^{\mathbb{Z}[X]}\!/_{\left(X^2+1\right)}\right)/_{(\overline{p})} \simeq \mathbb{Z}[\mathrm{i}]/_{(p)}$.

Diffusion à titre gratuit uniquement.

Corollaire

Soit $n \in \mathbb{N}^*$, qu'on décompose en facteurs premiers : $n = \prod_{p\in\mathcal{P}} p^{\nu_p(n)}$ (où $\mathcal{P}$ désigne l'ensemble des nombres premiers).
On a l'équivalence : $n \in \Sigma \Leftrightarrow \big(\forall p \in \mathcal{P}, p \equiv 3\,[4] \Rightarrow \nu_p(n) \equiv 0\,[2]\big)$.

Démonstration :

Lemme 3

Σ est stable par multiplication.

Démonstration du lemme 3 :

En effet, on sait déjà que $n \in \Sigma \Leftrightarrow \exists z \in \mathbb{Z}[\mathrm{i}], n = N(z)$.
En conséquence, si $n, n' \in \Sigma$, alors $nn' = N(z)N(z') = N(zz') \in \Sigma$. ■

$\Leftarrow$ On décompose n de la façon suivante :

$$n = \underbrace{\left(\prod_{\substack{p\in\mathcal{P}\\ p\equiv 3\,[4]}} p^{\frac{\nu_p(n)}{2}}\right)^2}_{\text{Carré parfait}} \underbrace{\left(\prod_{\substack{p\in\mathcal{P}\\ p\not\equiv 3\,[4]}} p^{\nu_p(n)}\right)}_{\substack{\text{Somme de 2 carrés}\\ \text{(lemme 3)}}}$$

$\Rightarrow$ Soit $n = a^2 + b^2 \in \Sigma$, on note $\delta = a \wedge b$, $a' = \dfrac{a}{\delta}$ et $b' = \dfrac{b}{\delta}$.
Ainsi, $a' \wedge b' = 1$ et $n = \delta^2\left(a'^2 + b'^2\right)$.
Soit p un diviseur premier impair de $a'^2 + b'^2$; alors dans $\mathbb{Z}[\mathrm{i}]$, on a : $p|(a' + \mathrm{i}b')(a' - \mathrm{i}b')$.

- Par l'absurde, supposons p irréductible dans $\mathbb{Z}[\mathrm{i}]$.
  Le lemme d'Euclide nous indique que $p|(a' + \mathrm{i}b')$ ou que $p|(a' - \mathrm{i}b')$; mais par passage au conjugué, si p divise l'un, alors p divise l'autre.
  Donc p divise les deux, puis par somme et différence, on obtient : $p|2a'$ et $p|2\mathrm{i}b'$ dans $\mathbb{Z}[\mathrm{i}]$.
  En passant à la norme, on en déduit : $p^2\big|4a'^2$ et $p^2\big|4b'^2$, dans $\mathbb{Z}$.
  Mais on sait que p est impair, et donc $p|a'$ et $p|b'$.
  Contradiction !
- On peut donc écrire $p = xy$ dans $\mathbb{Z}[\mathrm{i}]$, avec en plus $N(x) \neq 1 \neq N(y)$ (ce qui signifie, rappelons-le, que x et y peuvent être pris non-inversibles).
  En passant à la norme, on obtient : $p^2 = N(x)N(y)$; puis, p étant premier, on obtient : $p = N(x)$.
  En conséquence, $p \in \Sigma$, d'où $p \equiv 1\,[4]$.

Ainsi, on a montré que les facteurs premiers congrus à 3 modulo 4 sont "dans" le δ^2, c'est-à-dire d'exposant pair. ■

Références

[Per] D. PERRIN – *Cours d'algèbre*, Ellipses, 1996.
[Duv] D. DUVERNEY – *Théorie des nombres*, 2e éd., Dunod, 2007.

Diffusion à titre gratuit uniquement.

Théorèmes de Chevalley-Warning et d'Erdös-Ginzburg-Ziv

Leçons : 120 [62], 123, 142, 144, *121, 126, 190*

[Ser], paragraphe 1.2
[Zav], problème 7.II

Théorème (*Chevalley-Warning*)

Soient p un nombre premier, $r \in \mathbb{N}^*$; on note $q = p^r$.
Soient $f_1, \dots, f_s \in \mathbb{F}_q[X_1, \dots, X_n]$, tels que $\sum_{i=1}^{s} \deg f_i < n$.
On note $V = \left\{ (x_1, \dots, x_n) \in \mathbb{F}_q^n \middle| \forall i \in [\![1, s]\!], f_i(x_1, \dots, x_n) = 0 \right\}$.
Alors on a : $\#V \equiv 0 \,[p]$.

Démonstration :

Posons $P = \prod_{i=1}^{s} \left(1 - f_i^{q-1}\right)$ et soit $\underline{x} = (x_1, \dots, x_n) \in \mathbb{F}_q^n$.

- Si $\underline{x} \in V$, alors $\forall i \in [\![1, s]\!], f_i(\underline{x}) = 0$ et donc $P(\underline{x}) = 1$.
- Si $\underline{x} \notin V$, alors $\exists i_0 \in [\![1, s]\!], f_{i_0}(\underline{x}) \neq 0$ puis $f_{i_0}(\underline{x})^{q-1} = 1$ d'où $P(\underline{x}) = 0$.

Ainsi, en posant $S(f) = \sum_{\underline{x} \in \mathbb{F}_q^n} f(\underline{x})$ pour $f \in \mathbb{F}_q[X_1, \dots, X_n]$, on a : $S(P) = \sum_{\underline{x} \in V} 1 + 0 \equiv \#V \,[p]$.

On veut désormais montrer que $S(P) = 0$.

Lemme

Soit $u \in \mathbb{N}$, vérifiant $u = 0$ ou $(q-1) \nmid u$.
On pose $s(X^u) = \sum_{x \in \mathbb{F}_q} x^u$, et on a alors $s(X^u) = 0$, avec la convention $0^0 = 1$. [63]

Démonstration :

Si $u = 0$, alors $\forall x \in \mathbb{F}_q, x^u = 1$ et $s(X^u) = 0$.
Si $(q-1) \nmid u$, on écrit la division euclidienne $u = (q-1)k + r$, où $k \in \mathbb{N}$ et $0 < r < q-1$.
Soit y un générateur de $\mathbb{F}_q^\times$, qui est cyclique [64].
On a donc $y^u = \left(y^{q-1}\right)^k y^r = y^r \neq 1$ car y est d'ordre $(q-1)$ et $0 < r < q-1$.
Ainsi on a :

$$s(X^u) = \sum_{x \in \mathbb{F}_q} x^u = \sum_{x \in \mathbb{F}_q^\times} x^u = \sum_{x \in \mathbb{F}_q^\times} (xy)^u = y^u \sum_{x \in \mathbb{F}_q^\times} x^u = y^u s(X^u)$$

Donc $(1 - y^u) s(X^u) = 0$, puis par intégrité de $\mathbb{F}_q$, $s(X^u) = 0$ car $1 - y^u \neq 0$. ■

On a $\deg P = \sum_{i=1}^{s} (q-1) \deg f_i < n(q-1)$ et donc $P = \sum_{|\underline{u}| < n(q-1)} \alpha_{\underline{u}} \underline{X}^{\underline{u}}$ où $|\underline{u}| = \sum_{j=1}^{n} u_j$ et $\alpha_{\underline{u}} \in \mathbb{F}_q$.
D'où $S(P) = \sum_{\underline{x} \in \mathbb{F}_q^n} \sum_{|\underline{u}| < n(q-1)} \alpha_{\underline{u}} \underline{x}^{\underline{u}} = \sum_{|\underline{u}| < n(q-1)} \alpha_{\underline{u}} S(\underline{X}^{\underline{u}})$.

Or, pour $|\underline{u}| < n(q-1)$, $S(\underline{X}^{\underline{u}}) = \sum_{(x_1, \dots, x_n) \in \mathbb{F}_q^n} x_1^{u_1} \dots x_n^{u_n} = \sum_{x_1 \in \mathbb{F}_q} x_1^{u_1} \dots \sum_{x_n \in \mathbb{F}_q} x_n^{u_n} = \prod_{j=1}^{n} s(X^{u_i})$.

Mais $\sum_{j=1}^{n} u_j < n(q-1)$ impose $\exists j_0 \in [\![1, n]\!], u_{j_0} < q-1$ donc $(q-1) \nmid u_{j_0}$ ou $u_{j_0} = 0$.
Donc $s\left(X^{u_{j_0}}\right) = 0$, d'où $S(P) = 0$ puis $\#V \equiv 0 \,[p]$. ■

62. On passera le lemme permettant de démontrer le théorème de Chevalley-Warning et on détaillera le cas non-premier du théorème d'Erdös-Ginzburg-Ziv.

63. On peut montrer que si $u \geqslant 1$ et $(q-1) | u$, alors $s(X^u) = -1$.
En effet, $\exists k \in \mathbb{N}^*, u = (q-1)k$ et pour $x \in \mathbb{F}_q^\times$, $x^u = \left(x^{q-1}\right)^k = 1^k = 1$.
En outre $0^u = 0$, donc $s(X^u) = (q-1) \times 1 + 0 = -1$ dans $\mathbb{F}_q$.

64. Pour la cyclicité de $\mathbb{F}_q^\times$, on renvoie à la page 139.

Diffusion à titre gratuit uniquement.

Théorème (*Erdös-Ginzburg-Ziv*)

Soit p un nombre premier, et soient $a_1, \dots, a_{2p-1} \in \mathbb{Z}$. [65]
Parmi ces $(2p-1)$ nombres entiers, on peut en trouver p dont la somme est divisible par p.

Démonstration :

Pour $a \in \mathbb{Z}$, on note $\overline{a}$ sa classe modulo p.

On considère les polynômes $P_1 = \sum_{k=1}^{2p-1} X_k^{p-1}, P_2 = \sum_{k=1}^{2p-1} \overline{a_k} X_k^{p-1} \in \mathbb{F}_p\left[X_1, \dots, X_{2p-1}\right]$.

On a : $\deg P_1 + \deg P_2 = 2p-2 < 2p-1$, et $(0, \dots, 0)$ est une racine commune à ces deux polynômes ; donc, par le théorème de Chevalley-Warning, ils admettent une autre racine commune $\left(x_1, \dots, x_{2p-1}\right) \in \mathbb{F}_p^{2p-1}$.

De $P_1\left(x_1, \dots, x_{2p-1}\right) = 0$, il vient que parmi $x_1, \dots, x_{2p-1}$, exactement p d'entre eux sont non-nuls, et on les note $x_{n_1}, \dots, x_{n_p}$.

De $P_2\left(x_1, \dots, x_{2p-1}\right) = 0$, il vient ensuite que $\sum_{i=1}^{p} \overline{a_{n_i}} = 0$.

On a donc trouvé p éléments $a_{n_1}, \dots, a_{n_p}$ dont la somme est divisible par p. ■

Références

[Ser] J.-P. SERRE – *Cours d'arithmétique*, 1e éd., Presses Universitaires de France, 1970.
[Zav] M. ZAVIDOVIQUE – *Un max de maths*, Calvage & Mounet, 2013.

65. Ce résultat reste vrai pour n'importe quel $n \in \mathbb{N}^*$. On opère par récurrence forte.
- Si $n = 1$, le résultat est trivial.
- Soit $n > 1$, on suppose le résultat jusqu'au rang $(n-1)$. Soient $a_1, \dots, a_{2n-1} \in \mathbb{Z}$.
  - Si n est premier, c'est l'objet du développement.
  - Sinon, on écrit $n = pn'$, avec p premier et $n' \in \mathbb{N}^*$.
    On a alors : $2n-1 = 2n'p - 1 = (2n'-1)p + p - 1$.
    Pour i allant de 1 à $2n'-1$, on construit par récurrence les ensembles suivants appelés E_i :

    E_i est un ensemble de p éléments parmi $\left\{a_j \middle| j \in [\![1, (i+1)p-1]\!]\right\} \setminus \bigcup_{k=1}^{i-1} E_k$, dont la somme est divisible par p.

    La construction de ces ensembles utilise le résultat démontré dans le développement, car $\#\left\{a_j \middle| j \in [\![1, (i+1)p-1]\!]\right\} \setminus \bigcup_{k=1}^{i-1} E_k = 2p-1$.

    Puis, pour $i \in [\![1, 2n'-1]\!]$, S_i désigne la somme des éléments de E_i et $S'_i = \dfrac{S_i}{p} \in \mathbb{Z}$.

    Par hypothèse de récurrence, parmi les $(2n'-1)$ entiers S'_i, il en existe n' dont la somme est divisible par n', et on les note $S'_{k_1}, \dots, S'_{k_{n'}}$.

    On pose alors $E = \bigsqcup_{i=1}^{n'} E_{k_i}$, et $\#E = n'p = n$ et $\sum_{x \in E} x = \sum_{i=1}^{n'} S_{k_i} = p \sum_{i=1}^{n'} S'_{k_i}$.

    Or $n' \left| \sum_{i=1}^{n'} S'_{k_i} \right.$ donc $n = pn' \left| \sum_{x \in E} x \right.$.

On peut même montrer un résultat d'optimalité : prenons un ensemble de $(2n-2)$ entiers composé de $(n-1)$ fois 0, et de $(n-1)$ fois 1. Un sous-ensemble de n entiers parmi ceux-ci sera de somme comprise entre 1 et $n-1$, donc non-divisible par n.

Diffusion à titre gratuit uniquement.

Algorithme du gradient à pas optimal

Leçons : 232, *215, 219, 226, 229*

[HU], exercice II.8
[X-ENS Al3], exercice 2.35

Théorème

Soient $A \in \mathcal{S}_n^{++}(\mathbb{R})$ et $b \in \mathbb{R}^n$; on veut minimiser $f : x \mapsto \frac{1}{2}\langle Ax, x\rangle + \langle b, x\rangle$, quand x parcourt $\mathbb{R}^n$.
Il existe une unique solution à ce problème, et elle est caractérisée par $\nabla f(\overline{x}) = 0$.
De plus, l'algorithme défini par $\begin{cases} x_0 \in \mathbb{R}^n \\ \forall k \in \mathbb{N}, x_{k+1} = x_k + t_k d_k \end{cases}$, où $d_k = -\nabla f(x_k)$ et où t_k est l'unique réel minimisant la fonction $t \mapsto f(x_k + t d_k)$, converge vers $\overline{x}$.

Démonstration :

1. Soit $\overline{x}$ un point minimal, alors nécessairement $\nabla f(\overline{x}) = 0$.
   Or, pour tous $x, h \in \mathbb{R}^n$: $f(x+h) = \frac{1}{2}\langle A(x+h), x+h\rangle + \langle b, x+h\rangle$
   $$= \frac{1}{2}\langle Ax, x\rangle + \frac{1}{2}\langle Ah, x\rangle + \frac{1}{2}\langle Ax, h\rangle + \frac{1}{2}\langle Ah, h\rangle + \langle b, x\rangle + \langle b, h\rangle$$
   $$= f(x) + \langle Ax, h\rangle + \langle b, h\rangle + \frac{1}{2}\langle Ah, h\rangle \quad \text{(car } A \text{ est symétrique)}$$
   $$= f(x) + \langle Ax + b, h\rangle + o(\|h\|)$$
   De ce calcul, il vient notamment que f est différentiable [66] et que $\forall x \in \mathbb{R}^n, \nabla f(x) = Ax + b$.
   Mais A étant symétrique définie positive, f est strictement convexe et on en déduit : $\overline{x} = -A^{-1}b$.
   Procédons alors au calcul de la valeur optimale :
   $$\overline{f} := f(\overline{x}) = \frac{1}{2}\langle -b, -A^{-1}b\rangle + \langle b, -A^{-1}b\rangle = \frac{1}{2}\langle A^{-1}b, b\rangle - \langle A^{-1}b, b\rangle = -\frac{1}{2}\langle A^{-1}b, b\rangle$$
2. Soit $k \in \mathbb{N}$; on suppose que $d_k \neq 0$, car sinon $Ax_k = -b$ et alors l'algorithme a convergé en temps fini, et on n'a plus rien à dire.
3. On va maintenant calculer t_k.
   Pour $t \in \mathbb{R}$, on pose $g(t) := f(x_k + t d_k) = f(x_k) + \langle \underbrace{Ax_k + b}_{=-d_k}, t d_k\rangle + \frac{1}{2}\langle A t d_k, t d_k\rangle$.
   Et donc, $g(t) = f(x_k) - t\|d_k\|^2 + \frac{t^2}{2}\langle Ad_k, d_k\rangle$.
   Ainsi [67], g atteint son minimum en $t_k = \frac{\|d_k\|^2}{\langle Ad_k, d_k\rangle}$ (on rappelle que $d_k \neq 0$, assurant que $\langle Ad_k, d_k\rangle \neq 0$, étant donné que A est symétrique définie positive).
4. Calculons l'erreur commise entre $f(x_k)$ et $\overline{f}$.
   $$f(x_{k+1}) = f(x_k + t_k d_k) = f(x_k) - \frac{\|d_k\|^4}{\langle Ad_k, d_k\rangle} + \frac{1}{2}\frac{\|d_k\|^4}{\langle Ad_k, d_k\rangle} = f(x_k) - \frac{1}{2}\frac{\|d_k\|^4}{\langle Ad_k, d_k\rangle}$$
   $$f(x_{k+1}) - \overline{f} = \left(f(x_k) - \overline{f}\right) - \frac{1}{2}\frac{\|d_k\|^4}{\langle Ad_k, d_k\rangle} = \left(f(x_k) - \overline{f}\right)\left(1 - \frac{1}{2\left(f(x_k) - \overline{f}\right)}\frac{\|d_k\|^4}{\langle Ad_k, d_k\rangle}\right)$$
   Mais en fait, $\langle A^{-1}d_k, d_k\rangle = \langle A^{-1}(Ax_k + b), Ax_k + b\rangle = \langle x_k, Ax_k\rangle + \langle x_k, b\rangle + \underbrace{\langle A^{-1}b, Ax_k\rangle}_{=\langle b, x_k\rangle} + \langle A^{-1}b, b\rangle$
   $$= 2\left(\frac{1}{2}\langle Ax_k, x_k\rangle + \langle b, x_k\rangle - \overline{f}\right) = 2\left(f(x_k) - \overline{f}\right)$$

66. En même temps, f est polynomiale, donc la différentiabilité était déjà évidente.
67. Je trouvais que le fait d'invoquer directement la formule bien connue des polynômes du 2nd degré était préférable à la dérivation, car cela permet d'aller plus vite, ce qui n'est pas négligeable sur ce développement.

Diffusion à titre gratuit uniquement.

On en déduit alors que : $f(x_{k+1}) - \overline{f} = \left(f(x_k) - \overline{f}\right)\left(1 - \dfrac{\|d_k\|^4}{\langle A^{-1}d_k, d_k\rangle\langle Ad_k, d_k\rangle}\right)$.

Lemme (*Inégalité de Kantorovitch*[68])

Soit $A \in \mathcal{S}_n^{++}(\mathbb{R})$, dont λ_1 et λ_n sont les plus petite et grande valeurs propres.

Alors $\forall x \in \mathbb{R}^n, \langle Ax, x\rangle\langle A^{-1}x, x\rangle \leqslant \dfrac{1}{4}\left(\sqrt{\dfrac{\lambda_1}{\lambda_n}} + \sqrt{\dfrac{\lambda_n}{\lambda_1}}\right)\|x\|^4$.[69]

Démonstration du lemme :

En notant $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ dans une base orthonormée de vecteurs propres de A[70], on fait les calculs suivants :

$$\langle Ax, x\rangle\langle A^{-1}x, x\rangle = \left(\sum_{i=1}^n \lambda_i x_i^2\right)\left(\sum_{i=1}^n \frac{1}{\lambda_i}x_i^2\right) \underset{\substack{\text{Cauchy-}\\ \text{Schwarz}}}{\geqslant} \left(\sum_{i=1}^n \frac{\sqrt{\lambda_i}}{\sqrt{\lambda_i}}x_i^2\right)^2 = \left(\sum_{i=1}^n x_i^2\right)^2 \geqslant 0$$

Mais comme on sait[71] que $\sqrt{ab} \leqslant \dfrac{1}{2}(a+b)$:

$$\sqrt{\langle Ax, x\rangle\langle A^{-1}x, x\rangle} = \sqrt{\frac{\lambda_1}{\lambda_n}}\sqrt{\left(\sum_{i=1}^n \frac{\lambda_i}{\lambda_1}x_i^2\right)\left(\sum_{i=1}^n \frac{\lambda_n}{\lambda_i}x_i^2\right)} \leqslant \frac{1}{2}\sqrt{\frac{\lambda_1}{\lambda_n}}\sum_{i=1}^n\left(\frac{\lambda_i}{\lambda_1} + \frac{\lambda_n}{\lambda_i}\right)x_i^2$$

On peut montrer que $\alpha : x \mapsto \dfrac{x}{\lambda_1} + \dfrac{\lambda_n}{x}$ est décroissante sur $\left(\lambda_1, \sqrt{\lambda_1\lambda_n}\right)$ et croissante sur $\left(\sqrt{\lambda_1\lambda_n}, \lambda_n\right)$.[72]

Toujours est-il que α admet son maximum en λ_1 ou en λ_n ; mais $\alpha(\lambda_1) = \alpha(\lambda_n) = 1 + \dfrac{\lambda_n}{\lambda_1}$.

D'où : $\sqrt{\langle Ax, x\rangle\langle A^{-1}x, x\rangle} \leqslant \dfrac{1}{2}\sqrt{\dfrac{\lambda_1}{\lambda_n}}\displaystyle\sum_{i=1}^n\left(1 + \frac{\lambda_n}{\lambda_1}\right)x_i^2 = \frac{1}{2}\left(\sqrt{\frac{\lambda_1}{\lambda_n}} + \sqrt{\frac{\lambda_n}{\lambda_1}}\right)\sum_{i=1}^n x_i^2$.

Ce qui donne finalement, en élevant au carré : $\langle Ax, x\rangle\langle A^{-1}x, x\rangle \leqslant \dfrac{1}{4}\left(\sqrt{\dfrac{\lambda_1}{\lambda_n}} + \sqrt{\dfrac{\lambda_n}{\lambda_1}}\right)^2\|x\|^4$. ■

Utilisons l'inégalité de Kantorovitch :

$$\begin{aligned} f(x_{k+1}) - \overline{f} &\leqslant \left(f(x_k) - \overline{f}\right)\left(1 - \frac{4}{\left(\sqrt{c(A)} + \frac{1}{\sqrt{c(A)}}\right)^2}\right) = \left(f(x_k) - \overline{f}\right)\left(1 - \frac{4c(A)}{(c(A)+1)^2}\right) \\ &\leqslant \left(f(x_k) - \overline{f}\right)\left(\frac{c(A)-1}{c(A)+1}\right)^2 \end{aligned}$$

Et donc $\forall k \in \mathbb{N}, f(x_k) - \overline{f} \leqslant \left(f(x_0) - \overline{f}\right)\left(\dfrac{c(A)-1}{c(A)+1}\right)^{2k}$.

68. Dites juste que vous admettez l'inégalité de Kantorovitch et utilisez-la sans l'énoncer. Gardez un œil sur le chrono...

69. Comme $A \in \mathcal{S}_n^{++}(\mathbb{R})$, $\|A\|_2 = \sqrt{\rho({}^tAA)} = \rho(A) = \lambda_n$ (dites "décomposition polaire") et aussi $\left\|A^{-1}\right\|_2 = \dfrac{1}{\lambda_1}$. Ainsi, on fait apparaître le conditionnement en norme 2 de A : $\text{cond}_2(A) = \dfrac{\lambda_n}{\lambda_1}$, qu'on notera par la suite (pour plus de simplicité) $c(A)$.

70. Rappelez-vous le théorème spectral : toute matrice symétrique est diagonalisable dans une base orthonormée.

71. En fait, c'est une application toute bête des identités remarquables : comme $(a-b)^2 \geqslant 0$, on sait que $a^2 + b^2 \geqslant 2ab$. Ainsi, $\left(\dfrac{1}{2}(a+b)\right)^2 = \dfrac{1}{4}\left(a^2 + b^2\right) + \dfrac{1}{2}ab \geqslant ab$.

72. On peut, mais là, j'ai la flemme.

Diffusion à titre gratuit uniquement.

5. Pour finir, on va calculer l'erreur sur $\|x_k - \overline{x}\|$.
   On a : $\|x_k - \overline{x}\|^2 \underset{73}{\leqslant} \frac{1}{\lambda_1}\langle A\left(x_k - \overline{x}\right), x_k - \overline{x}\rangle = \frac{1}{\lambda_1}\left(\langle Ax_k, x_k\rangle - \langle Ax_k, \overline{x}\rangle - \langle A\overline{x}, x_k\rangle + \langle A\overline{x}, \overline{x}\rangle\right)$

$$= \frac{1}{\lambda_1}\left(\langle Ax_k, x_k\rangle - 2\langle x_k, A\overline{x}\rangle - 2\overline{f}\right) = \frac{1}{\lambda_1}\left(2f\left(x_k\right) - 2\overline{f}\right)$$
$$= \frac{2}{\lambda_1}\left(f\left(x_k\right) - \overline{f}\right)$$

   En fin de compte, $\|x_k - \overline{x}\| \leqslant \sqrt{\frac{2}{\lambda_1}}\sqrt{f\left(x_k\right) - \overline{f}} \leqslant \sqrt{\frac{2}{\lambda_1}\left(f\left(x_0\right) - \overline{f}\right)}\left(\frac{c(A)-1}{c(A)+1}\right)^k$.
   Comme $\left|\frac{c(A)-1}{c(A)+1}\right| < 1$, on en déduit que la suite $(x_k)_{k\in\mathbb{N}}$ converge vers $\overline{x}$.[74] ■

Références

[HU] J.-B. HIRIART-URRUTY – *Optimisation et analyse convexe*, EDP Sciences, 2009.
[X-ENS Al3] S. FRANCINOU, H. GIANELLA et S. NICOLAS – *Oraux X-ENS Algèbre 3*, 2e éd., Cassini, 2013.

73. Vous souvenez-vous du quotient de Rayleigh ?
74. Ouf !

Diffusion à titre gratuit uniquement.

Densité des fonctions continues nulle part dérivables

Leçons : 201, 202, 205, 208, 228, *213*

[ZQ], Section VIII.I.4.e

Théorème

L'ensemble A des fonctions continues sur $[0,1]$ qui ne sont dérivables en aucun point de $[0,1]$ est dense dans $\left(\mathcal{C}^0([0,1]), \|.\|_\infty\right)$.[75]

Démonstration :

Comme $E = \mathcal{C}^0([0,1])$ est un espace de Banach, on va utiliser le lemme de Baire pour montrer $\overline{A} = E$.
Plus précisément, on va exhiber une suite $(F_n)_{n\in\mathbb{N}}$ telle que :

- $\forall n \in \mathbb{N}$, F_n est fermé ;
- $\forall n \in \mathbb{N}$, F_n est d'intérieur vide ;
- $A^c \subset \bigcup_{n\in\mathbb{N}} F_n$.

Alors on aura montré que A^c est d'intérieur vide dans E.
On pose $F_n = \{f \in E \mid \exists x \in I, \forall y \in I, |f(y) - f(x)| \leqslant n|y-x|\}$, où I désigne l'intervalle $[0,1]$.

Étape 1 : Montrons d'abord que $A^c \subset \bigcup_{n\in\mathbb{N}} F_n$.

Soit $f \in A^c$, alors f est dérivable en au moins un point $x_0 \in I$.
Ainsi la quantité $\dfrac{f(x_0) - f(y)}{x_0 - y}$ est bornée quand $y \to x_0$.
Et f étant continue sur I, $\exists N \in \mathbb{N}, \forall y \in I, \left|\dfrac{f(x_0) - f(y)}{x_0 - y}\right| \leqslant N$ et donc $f \in F_N$.
Ainsi, $A^c \subset \bigcup_{n\in\mathbb{N}} F_n$.

Étape 2 : Soit $n \in \mathbb{N}$, on va maintenant montrer que F_n est fermé.

Soit $(f_k)_{k\in\mathbb{N}}$, une suite dans F_n qui converge vers $f \in E$.
Comme les f_k sont dans F_n, on a :

$$\forall k \in \mathbb{N}, \exists x_k \in I, \forall y \in I, |f_k(x_k) - f_k(y)| \leqslant n|x_k - y| \tag{3}$$

Et comme I est compact, il existe une extraction $\left(x_{\varphi(k)}\right)_{k\in\mathbb{N}}$ qui converge vers $x \in I$.
Soit $k \in \mathbb{N}$,

$$\left|f_{\varphi(k)}\left(x_{\varphi(k)}\right) - f(x)\right| \leqslant \underbrace{\left|f_{\varphi(k)}\left(x_{\varphi(k)}\right) - f\left(x_{\varphi(k)}\right)\right|}_{\leqslant \|f_{\varphi(k)} - f\|_\infty \underset{k\to\infty}{\longrightarrow} 0} + \underbrace{\left|f\left(x_{\varphi(k)}\right) - f(x)\right|}_{\substack{\underset{k\to\infty}{\longrightarrow} 0 \\ (f \text{ est continue})}}$$

Donc, $\lim_{k\to\infty} f_{\varphi(k)}\left(x_{\varphi(k)}\right) = f(x)$.
Dès lors, par passage à la limite dans (3), on obtient : $\forall y \in I, |f(x) - f(y)| \leqslant n|x-y|$, donc $f \in F_n$, puis F_n est fermé.

Étape 3 : Pour tout $n \in \mathbb{N}$, on va finalement montrer que F_n est d'intérieur vide.

Comme E est métrique, on va montrer qu'il n'existe pas de boule ouverte $\mathcal{B}(f,\varepsilon) \subset F_n$ avec $f \in F_n$ et $\varepsilon > 0$, ie :

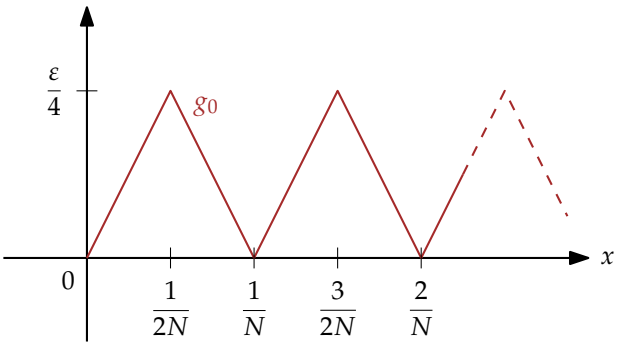
$$\forall f \in F_n, \forall \varepsilon > 0, \mathcal{B}(f,\varepsilon) \cap F_n^c \neq \varnothing$$

75. Quand on fait ce développement, il faut absolument être capable de donner un exemple d'une telle fonction ; allez, c'est le cadeau de la maison : $\forall x \in \mathbb{R}, f(x) = \sum_{n=0}^{+\infty} \dfrac{\{2^n x\}}{2^n}$, où $\{x\}$ désigne la distance de x à son entier le plus proche. Par ailleurs, une fois qu'on a montré que cette fonction est continue et nulle part dérivable, le théorème de Weierstrass implique le résultat du développement. Une application de ce théorème est de montrer que toute fonction continue s'écrit comme somme de deux fonctions continues nulle part dérivables.

Diffusion à titre gratuit uniquement.

Soit donc $f \in F_n$, et $\varepsilon > 0$.
Par le théorème de Weierstrass : $\exists P \in \mathbb{R}[X], \|P - f\|_\infty \leqslant \dfrac{\varepsilon}{2}$.
Soit $N \in \mathbb{N}^*$ (à qui on imposera une condition par la suite) ; on définit g_0, fonction périodique de période $\dfrac{1}{N}$ telle que :

$$g_0(x) = \begin{cases} \dfrac{\varepsilon N}{2} x & \text{si } x \in \left[0, \dfrac{1}{2N}\right] \\ \dfrac{\varepsilon}{2} - \dfrac{\varepsilon N}{2} x & \text{si } x \in \left[\dfrac{1}{2N}, \dfrac{1}{N}\right] \end{cases}$$



g_0 est continue et $\|g_0\|_\infty = \dfrac{\varepsilon}{4}$.
On pose $g = P + g_0$; on a : $\|f - g\|_\infty = \|f - P - g_0\|_\infty \leqslant \|f - P\|_\infty + \|g_0\|_\infty \leqslant \dfrac{\varepsilon}{2} + \dfrac{\varepsilon}{4} < \varepsilon$.
Donc $g \in \mathcal{B}(f, \varepsilon)$.
On a, pour tous $x, y \in I$: $|g(x) - g(y)| \geqslant |g_0(x) - g_0(y)| - |P(x) - P(y)|$.
De plus, pour $x \in I$, il existe $k \in [\![0, 2N-1]\!]$, tel que $x \in \left[\dfrac{k}{2N}, \dfrac{k+1}{2N}\right]$.
Puis, soit $y_x \in \left[\dfrac{k}{2N}, \dfrac{k+1}{2N}\right]$, alors on a : $|g_0(x) - g_0(y_x)| = \dfrac{\varepsilon N}{2} |x - y_x|$.
Et par le théorème des accroissements finis, on a : $|P(x) - P(y_x)| \leqslant \|P'\|_\infty |x - y_x|$.
Par conséquent : $\forall x \in I, \exists y_x \in I, |g(x) - g(y_x)| \geqslant \left(\dfrac{\varepsilon N}{2} - \|P'\|_\infty\right) |x - y_x|$.
On se rend alors compte qu'il suffit d'imposer $\dfrac{\varepsilon N}{2} - \|P'\|_\infty > n$, ie $N > \dfrac{2(n + \|P'\|_\infty)}{\varepsilon}$ dès le début pour avoir :

$$\forall x \in I, \exists y_x \in I, |g(x) - g(y_x)| > n |x - y_x|$$

Ainsi, $g \in F_n^c$, et finalement $\mathcal{B}(f, \varepsilon) \cap F_n^c \neq \varnothing$. ■

Références

[ZQ] H. QUEFFÉLEC et C. ZUILY – *Analyse pour l'agrégation*, 4e éd., Dunod, 2013.

Diffusion à titre gratuit uniquement.

Densité des polynômes orthogonaux [76]

Leçons : 209, 213, 239, 245, *201, 202, 207, 240, 244*

[OA], exercice 3.7

Rappels :

1. Pour I un intervalle de $\mathbb{R}$, on appelle fonction poids une fonction $\rho : I \to \mathbb{R}$ mesurable, strictement positive, et telle que $\forall n \in \mathbb{N}$, $\int_I |x|^n \rho(x)\,\mathrm{d}x < \infty$.
2. On note $\mathrm{L}^2(I,\rho)$ l'espace des fonctions (où on a identifié celles qui sont égales presque partout), de carré intégrable pour la mesure de densité ρ par rapport à la mesure de Lebesgue ; il est muni du produit scalaire $\langle f,g\rangle_\rho = \int_I f(x)\overline{g(x)}\rho(x)\,\mathrm{d}x$. C'est un espace de Hilbert, qui contient les polynômes.
3. Il existe une unique famille $(P_n)_{n\in\mathbb{N}}$ de polynômes unitaires orthogonaux deux à deux, et vérifiant $\deg P_n = n$. On l'appelle "famille des polynômes orthogonaux associée à ρ". On l'obtient en appliquant le procédé de Gram-Schmidt à la famille $(X^n)_{n\in\mathbb{N}}$.

Théorème

Soient I un intervalle [77] de $\mathbb{R}$ et ρ une fonction poids.
On suppose : $\exists a > 0$, $\int_I \mathrm{e}^{a|x|}\rho(x)\,\mathrm{d}x < \infty$. [78]
Alors les polynômes orthogonaux associés à ρ forment une base hilbertienne de $\mathrm{L}^2(I,\rho)$.

Démonstration :

Étape 1 : $(P_n)_{n\in\mathbb{N}}$ est une famille orthonormée ; il suffit de montrer qu'elle est totale.

On va montrer que $\mathrm{Vect}\left((P_n)_{n\in\mathbb{N}}\right)$ est dense dans $\mathrm{L}^2(I,\rho)$, c'est-à-dire que $\mathrm{Vect}\left((P_n)_{n\in\mathbb{N}}\right)^\perp = \{0\}$.
Par construction, on a : $\mathrm{Vect}\left((P_n)_{n\in\mathbb{N}}\right) = \mathrm{Vect}\left((X^n)_{n\in\mathbb{N}}\right)$.
On pose, pour tout $n \in \mathbb{N}$, $g_n : x \mapsto x^n$.

76. À quoi servent les polynômes orthogonaux ? Ils apparaissent en intégration numérique (révisez la méthode de Gauss) et pour la diagonalisation des opérateurs auto-adjoints compacts dans une base hilbertienne.

77. Si I est compact, l'hypothèse sur ρ est inutile et le théorème est trivialisé par le théorème de Weierstrass.

78. Regardons un contre-exemple (à connaître, mais mieux vaut prendre le temps de bien faire le développement et se garder le contre-exemple sous le coude pour les questions).
Prenons $I = \mathbb{R}^{+*}$ et $w(x) = x^{-\ln x}$ une fonction poids sur $\mathbb{R}^{+*}$; en fait, w ne vérifie pas l'hypothèse d'écrasement, et on va voir que les polynômes orthogonaux pour w ne constitue pas une base hilbertienne de $\mathrm{L}^2(I,w)$.
On pose $f : \left|\begin{array}{ccc} \mathbb{R}^{+*} & \to & \mathbb{R} \\ x & \mapsto & \sin(2\pi\ln x)\end{array}\right.$, et on fixe $n \in \mathbb{N}$.

$$\begin{aligned}
\langle f, g_n\rangle &= \int_{\mathbb{R}^{+*}} x^n \sin(2\pi\ln x) x^{-\ln x}\,\mathrm{d}x \\
&= \int_{\mathbb{R}} \mathrm{e}^{ny}\sin(2\pi y)\,(\mathrm{e}^y)^{-y}\,\mathrm{e}^y\,\mathrm{d}y \quad (\text{on a posé } y = \ln x \text{ et } \mathrm{d}x = \mathrm{e}^y\mathrm{d}y) \\
&= \int_{\mathbb{R}} \mathrm{e}^{(n+1)y-y^2}\sin(2\pi y)\,\mathrm{d}y \quad \left(\text{on utilise } (n+1)y - y^2 = \frac{(n+1)^2}{4} - \left(y - \frac{n+1}{2}\right)^2 \text{ pour la ligne suivante}\right) \\
&= \exp\left(\frac{(n+1)^2}{4}\right)\int_{\mathbb{R}} \exp\left(-\left(y-\frac{n+1}{2}\right)^2\right)\sin(2\pi y)\,\mathrm{d}y \\
&= \exp\left(\frac{(n+1)^2}{4}\right)\int_{\mathbb{R}} \mathrm{e}^{-t^2}\sin(2\pi t + (n+1)\pi)\,\mathrm{d}t \quad \left(\text{on a posé } t = y - \frac{n+1}{2} \text{ et } \mathrm{d}t = \mathrm{d}y\right) \\
&= (-1)^{n+1}\exp\left(\frac{(n+1)^2}{4}\right)\int_{\mathbb{R}} \mathrm{e}^{-t^2}\sin(2\pi t)\,\mathrm{d}t \\
&= 0 \quad (\text{intégrale d'une fonction impaire intégrable})
\end{aligned}$$

Ainsi, f est dans l'orthogonal de l'espace vectoriel engendré par les polynômes orthogonaux pour w, sans pour autant être nulle dans $\mathrm{L}^2(I,w)$.

Diffusion à titre gratuit uniquement.

On va donc montrer que $\forall f \in \mathrm{L}^2(I,\rho)$, $\left[\forall n \in \mathbb{N}, \langle f, g_n\rangle_\rho = 0\right] \Rightarrow [f = 0]$.
Dans la suite, on fixe $f \in \mathrm{L}^2(I,\rho)$ et vérifiant $\forall n \in \mathbb{N}, \langle f, g_n\rangle_\rho = 0$.

Étape 2 : On pose $\varphi = f\rho\mathbb{1}_I$; montrons que $\varphi \in \mathrm{L}^1(\mathbb{R})$.

Rappelons que : $\forall t \in \mathbb{R}^+, t \leqslant \dfrac{1+t^2}{2}$. Ainsi, $\forall x \in I, |\varphi(x)| \leqslant \dfrac{1+|f(x)|^2}{2}\rho(x)$.
Mais ρ est intégrable sur I (car c'est une fonction poids) et $\rho|f|^2$ aussi (car $f \in \mathrm{L}^2(I,\rho)$).
En conséquence, $\varphi \in \mathrm{L}^1(\mathbb{R})$.

Étape 3 : On peut donc considérer sa transformée de Fourier $\hat{\varphi}$ définie par : $\forall \omega \in \mathbb{R}, \hat{\varphi}(\omega) = \displaystyle\int_I f(x)\mathrm{e}^{-\mathrm{i}\omega x}\rho(x)\,\mathrm{d}x$.

On va montrer que $\hat{\varphi}$ se prolonge en une fonction holomorphe F sur la bande $B_a = \left\{z \in \mathbb{C}\,\middle|\,|\mathrm{Im}\ z| < \dfrac{a}{2}\right\}$.

On pose, pour $z \in B_a$, $g(z,x) := \mathrm{e}^{-\mathrm{i}zx}f(x)\rho(x)$.
Déjà, pour tout $z \in B_a$, on a :

$$\int_I |g(z,x)|\,\mathrm{d}x = \int_I \mathrm{e}^{\mathrm{Im}\ z.x}|f(x)|\rho(x)\,\mathrm{d}x \leqslant \int_I \mathrm{e}^{\frac{a|x|}{2}}|f(x)|\rho(x)\,\mathrm{d}x \underset{\substack{\text{Cauchy-}\\\text{Schwarz}}}{\leqslant} \sqrt{\int_I \mathrm{e}^{a|x|}\rho(x)\,\mathrm{d}x}\sqrt{\int_I |f(x)|^2\rho(x)\,\mathrm{d}x} < \infty$$

La fonction $F : \left|\begin{array}{ccc} B_a & \to & \mathbb{C} \\ z & \mapsto & \int_I g(z,x)\,\mathrm{d}x \end{array}\right.$ est donc bien définie.
On va utiliser le théorème d'holomorphie sous le signe intégrale :
– $\forall z \in B_a$, $x \mapsto g(z,x)$ est mesurable.
– Pour presque tout $x \in I$, $z \mapsto g(z,x)$ est holomorphie.
– $\forall z \in B_a$, $|g(z,x)| \leqslant \mathrm{e}^{\frac{a|x|}{2}}|f(x)|\rho(x)$ qui est une fonction de x, indépendante de z et qui n'a pas oublié d'être intégrable sur $\mathbb{R}$.

Donc F est holomorphe sur B_a et coïncide sur $\mathbb{R}$ avec $\hat{\varphi}$.

Étape 4 : Le théorème précédent nous permet également de calculer les dérivées de F :

$$\forall n \in \mathbb{N}, \forall z \in B_a, F^{(n)}(z) = (-\mathrm{i})^n \int_I x^n \mathrm{e}^{-\mathrm{i}zx}f(x)\rho(x)\,\mathrm{d}x$$

Ainsi, on obtient, pour tout $n \in \mathbb{N}$:

$$F^{(n)}(0) = (-\mathrm{i})^n \int_I x^n f(x)\rho(x)\,\mathrm{d}x = (-\mathrm{i})^n\langle f, g_n\rangle_\rho = 0$$

Par unicité du développement en série entière d'une fonction holomorphe, il existe un voisinage de 0 sur lequel $F \equiv 0$.
Par le théorème de prolongement analytique, on en déduit que $F \equiv 0$ sur B_a.
Conséquemment, $\hat{\varphi} = F_{|\mathbb{R}} \equiv 0$.
Comme φ est intégrable, on peut invoquer l'injectivité de la transformée de Fourier pour obtenir la nullité presque partout de φ sur $\mathbb{R}$.
Par stricte positivité de ρ, on en déduit que f est nulle presque partout sur I, autrement dit, $f = 0$ dans $\mathrm{L}^2(I,\rho)$.
Ce qui prouve le théorème. ■

Références

[OA] V. BECK, J. MALICK et G. PEYRÉ – *Objectif Agrégation*, 2e éd., H&K, 2005.

Diffusion à titre gratuit uniquement.

Équation de la chaleur sur un anneau

Leçons : 222, 246, *235, 241*

[X-ENS An4], exercice 1.28

Théorème

Soit $u_0 : \mathbb{R} \to \mathbb{R}$ une fonction non-nulle, continue, $\mathcal{C}^1$ par morceaux, et 2π-périodique.
Alors il existe une unique solution au problème $\begin{cases} \frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} \\ \forall x \in \mathbb{R}, u(0,x) = u_0(x) \end{cases}$ qui soit 2π-périodique par rapport à x et pour tout $t \in \mathbb{R}^+$, continue sur $\mathbb{R}^+ \times \mathbb{R}$ et $\mathcal{C}^\infty$ sur $\mathbb{R}^{+*} \times \mathbb{R}$.

Démonstration :

Analyse : Soit u une solution du problème posé.
Pour $t \in \mathbb{R}^{+*}$, $u_t : x \mapsto u(t,x)$ est 2π-périodique et $\mathcal{C}^\infty$, et donc elle est somme de sa série de Fourier !

$$\forall t \in \mathbb{R}^{+*}, \forall x \in \mathbb{R}, u(t,x) = \sum_{n=-\infty}^{+\infty} c_n(t) \mathrm{e}^{\mathrm{i}nx} \text{ avec } c_n(t) = \frac{1}{2\pi} \int_0^{2\pi} u(t,x) \mathrm{e}^{-\mathrm{i}nx} \,\mathrm{d}x, n \in \mathbb{Z}.$$

Montrons que $\dfrac{\partial u}{\partial t}$ s'obtient par dérivation formelle de cette expression.
Comme pour u, on peut écrire :

$$\forall t \in \mathbb{R}^{+*}, \forall x \in \mathbb{R}, \frac{\partial u}{\partial t}(t,x) = \sum_{n=-\infty}^{+\infty} \tilde{c}_n(t) \mathrm{e}^{\mathrm{i}nx} \text{ avec } \tilde{c}_n(t) = \frac{1}{2\pi} \int_0^{2\pi} \frac{\partial u}{\partial t}(t,x) \mathrm{e}^{-\mathrm{i}nx} \,\mathrm{d}x, n \in \mathbb{Z}.$$

Or $(t,x) \mapsto u(t,x)\mathrm{e}^{\mathrm{i}nx}$ admet une dérivée partielle par rapport à t qui soit continue sur $\mathbb{R}^{+*} \times [0, 2\pi]$.
Donc c_n est $\mathcal{C}^1$ sur $\mathbb{R}^{+*}$ et $c_n'(t) = \tilde{c}_n(t)$, d'où :

$$\forall t \in \mathbb{R}^{+*}, \forall x \in \mathbb{R}, \frac{\partial u}{\partial t}(t,x) = \sum_{n=-\infty}^{+\infty} c_n'(t) \mathrm{e}^{\mathrm{i}nx}.$$

Similairement, pour $t \in \mathbb{R}^{+*}$, $x \mapsto \dfrac{\partial^2 u}{\partial x^2}(t,x)$ est somme de sa série de Fourier, et par double intégration par parties :

$$\forall t \in \mathbb{R}^{+*}, \forall x \in \mathbb{R}, \frac{\partial^2 u}{\partial x^2}(t,x) = -\sum_{n=-\infty}^{+\infty} n^2 c_n(t) \mathrm{e}^{\mathrm{i}nx}.$$

À t fixé, toutes les séries considérées convergent normalement (séries de Fourier de fonctions continues et $\mathcal{C}^1$ par morceaux).
Désormais le problème s'écrit :

$$\forall t \in \mathbb{R}^{+*}, \forall x \in \mathbb{R}, 0 = \sum_{n=-\infty}^{+\infty} \left(c_n'(t) + n^2 c_n(t) \right) \mathrm{e}^{\mathrm{i}nx}.$$

Ainsi, pour $p \in \mathbb{Z}$, $\forall t \in \mathbb{R}^{+*}$:

$$\begin{aligned} 0 &= \int_0^{2\pi} \mathrm{e}^{-\mathrm{i}px} \sum_{n=-\infty}^{+\infty} \left(c_n'(t) + n^2 c_n(t) \right) \mathrm{e}^{\mathrm{i}nx} \,\mathrm{d}x \underset{\text{(cvn)}}{=} \sum_{n=-\infty}^{+\infty} \left(c_n'(t) + n^2 c_n(t) \right) \int_0^{2\pi} \mathrm{e}^{\mathrm{i}(n-p)x} \,\mathrm{d}x \\ &= 2\pi \left(c_p'(t) + p^2 c_p(t) \right). \end{aligned}$$

En d'autres termes : $\forall n \in \mathbb{Z}, \forall t \in \mathbb{R}^{+*}, c_n'(t) + n^2 c_n(t) = 0$.
Ce qui nous donne donc : $\forall n \in \mathbb{Z}, \exists \alpha_n \in \mathbb{C}, \forall t \in \mathbb{R}^{+*}, c_n(t) = \alpha_n \mathrm{e}^{-n^2 t}$.
Mais u_0 est 2π-périodique, continue et $\mathcal{C}^1$ par morceaux, donc somme de sa série de Fourier.

Diffusion à titre gratuit uniquement.

On écrit : $\forall x \in \mathbb{R}, u_0(x) = \sum_{n=-\infty}^{+\infty} C_n \mathrm{e}^{\mathrm{i}nx}$.
En appliquant Parseval à la fonction $u_0 - u_t$, on obtient :

$$\forall t \in \mathbb{R}^{+*}, \sum_{n=-\infty}^{+\infty} |C_n - c_n(t)|^2 = \frac{1}{2\pi} \int_0^{2\pi} |u_0(x) - u_t(x)|^2 \, \mathrm{d}x.$$

Notamment, $\forall n \in \mathbb{Z}, \forall t \in \mathbb{R}^{+*}, |C_n - c_n(t)|^2 \leqslant \frac{1}{2\pi} \int_0^{2\pi} |u(0,x) - u(t,x)|^2 \, \mathrm{d}x$.
On va utiliser le théorème de convergence dominée :
– d'une part, $(t,x) \mapsto |u(0,x) - u(t,x)|^2$ est continue sur le compact $[0,1] \times [0,2\pi]$, donc bornée, donc majorée par une fonction intégrable sur $[0,2\pi]$ et indépendante de t ;
– d'autre part, $\forall x \in [0,2\pi], |u(0,x) - u(t,x)|^2 \underset{t\to 0}{\longrightarrow} 0$.

Ainsi, on obtient : $\int_0^{2\pi} |u(0,x) - u(t,x)|^2 \, \mathrm{d}x \underset{t\to 0}{\longrightarrow} 0$, ce qui fournit ensuite, en utilisant la continuité de la fonction c_n sur $\mathbb{R}^+$:

$$C_n = \lim_{t\to 0} c_n(t) = c_n(0) = \alpha_n.$$

Ainsi, si u est solution du problème posé, alors on a : $\forall t \in \mathbb{R}^+, \forall x \in \mathbb{R}, u(t,x) = \sum_{n=-\infty}^{+\infty} C_n \mathrm{e}^{-n^2 t} \mathrm{e}^{\mathrm{i}nx}$, où les C_n sont les coefficients de Fourier de u_0.

Synthèse : Montrons que la fonction $u : (t,x) \mapsto \sum_{n=-\infty}^{+\infty} C_n \mathrm{e}^{-n^2 t} \mathrm{e}^{\mathrm{i}nx}$ convient.

On a : $\forall t \in \mathbb{R}^+, \forall x \in \mathbb{R}, \left| C_n \mathrm{e}^{-n^2 t} \mathrm{e}^{\mathrm{i}nx} \right| \leqslant |C_n|$ et $\sum_{n=-\infty}^{+\infty} |C_n| < +\infty$ car u_0 est 2π-périodique, continue et $\mathcal{C}^1$ par morceaux.
Donc la série définissant u converge normalement, donc u est bien définie et continue sur $\mathbb{R}^+ \times \mathbb{R}$, car $\forall n \in \mathbb{Z}, (t,x) \mapsto C_n \mathrm{e}^{-n^2 t} \mathrm{e}^{\mathrm{i}nx}$ est continue.
Aussi, $\forall t \in \mathbb{R}^+, u_t$ est 2π-périodique.
Enfin, $\forall n \in \mathbb{Z}, \frac{\partial^{k+l}}{\partial t^k \partial x^l} \left(C_n \mathrm{e}^{-n^2 t} \mathrm{e}^{\mathrm{i}nx} \right) = (-1)^k \mathrm{i}^l n^{2k+l} C_n \mathrm{e}^{-n^2 t} \mathrm{e}^{\mathrm{i}nx}$.
Soit $t_0 > 0, \forall t > t_0, \left| (-1)^k \mathrm{i}^l n^{2k+l} C_n \mathrm{e}^{-n^2 t} \mathrm{e}^{\mathrm{i}nx} \right| \leqslant n^{2k+l} K \mathrm{e}^{-n^2 t_0} = o\left(\frac{1}{n^2}\right)$, où on a utilisé l'inégalité :

$$|C_n| = \frac{1}{2\pi} \left| \int_0^{2\pi} \sum_{k=-\infty}^{+\infty} C_k \mathrm{e}^{\mathrm{i}(k-n)x} \, \mathrm{d}x \right| = \frac{1}{2\pi} \left| \int_0^{2\pi} u_0(x) \mathrm{e}^{-\mathrm{i}nx} \, \mathrm{d}x \right| \leqslant \frac{1}{2\pi} \int_0^{2\pi} |u_0(x)| \, \mathrm{d}x =: K.$$

Ceci est donc le terme d'une série normalement convergente sur $]t_0, +\infty[$; la dérivation formelle y est donc autorisée ; mais t_0 étant arbitraire, on obtient que u est $\mathcal{C}^\infty$ sur $\mathbb{R}^{+*} \times \mathbb{R}$ et que :

$$\forall k, l \in \mathbb{N}, \frac{\partial^{k+l} u}{\partial t^k \partial x^l}(t,x) = \sum_{n=-\infty}^{+\infty} (-1)^k \mathrm{i}^l n^{2k+l} C_n \mathrm{e}^{-n^2 t} \mathrm{e}^{\mathrm{i}nx}.$$

D'où, $\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}$; u est donc solution du problème posé. ■

Références

[X-ENS An4] S. FRANCINOU, H. GIANELLA et S. NICOLAS – *Oraux X-ENS Analyse 4*, 1e éd., Cassini, 2012.

Estimateur du maximum de vraisemblance pour le paramètre d'une loi $\mathcal{U}([0,\theta])$

Leçons : 263, *260, 262*

Merci Caroline ![79]

Théorème

Soient $X_1, \dots, X_n$ des variables aléatoires indépendantes et identiquement distribuées selon la loi $\mathbb{P}_\theta = \mathcal{U}([0,\theta])$, avec $\theta > 0$.
On note $\widehat{\theta_n}$ l'estimateur du maximum de vraisemblance du paramètre θ.
Alors :

1. $\widehat{\theta_n} = \max_{1\leqslant i\leqslant n} X_i$;
2. $\widehat{\theta_n}$ est biaisé car $\mathbb{E}_\theta\left[\widehat{\theta_n}\right] = \dfrac{n}{n+1}\theta$;
3. $\widehat{\theta_n}$ est fortement consistant, ce qui signifie $\widehat{\theta_n} \xrightarrow[n\to\infty]{\mathbb{P}_\theta\text{-ps}} \theta$;
4. $\widehat{\theta_n}$ est de vitesse $\dfrac{1}{n}$, car $n\left(\widehat{\theta_n} - \theta\right) \xrightarrow[n\to\infty]{\mathbb{P}_\theta\text{-}\mathcal{L}} -\mathcal{E}\left(\dfrac{1}{\theta}\right)$;
5. Le risque quadratique de $\widehat{\theta_n}$ vaut $\mathbb{E}_\theta\left[\left(\widehat{\theta_n} - \theta\right)^2\right] = \dfrac{2\theta^2}{(n+1)(n+2)}$.

Démonstration :

1. Relativement à la mesure de Lebesgue sur $(\mathbb{R}^+)^n$, le modèle statistique $\left((\mathbb{R}^+)^n, \left\{\mathbb{P}_\theta^{\otimes n}\right\}_{\theta>0}\right)$ est dominé est donc admet une vraisemblance qui s'écrit :

$$\forall \theta > 0, L_n(x_1,\dots,x_n;\theta) = \begin{cases} \theta^{-n} & \text{si } 0 \leqslant x_1,\dots,x_n \leqslant \theta \\ 0 & \text{sinon} \end{cases}.$$

   Donc l'estimateur recherché vaut $\widehat{\theta_n} = \max_{1\leqslant i\leqslant n} X_i$.[80]

2. On va d'abord calculer la loi de $\widehat{\theta_n}$; soit $t \in \mathbb{R}^+$.

$$F_{\widehat{\theta_n}}(t) = \mathbb{P}_\theta\left(\widehat{\theta_n} \leqslant t\right) = \mathbb{P}_\theta\left(\forall i \in [\![1,n]\!], X_i \leqslant t\right) = \mathbb{P}_\theta\left(X_1 \leqslant t\right)^n = \begin{cases} \left(\dfrac{t}{\theta}\right)^n & \text{si } t \leqslant \theta \\ 1 & \text{sinon} \end{cases}.$$

   On en déduit que $\widehat{\theta_n}$ est à densité et qu'elle vaut :

$$f_{\widehat{\theta_n}}(t) = \frac{n}{\theta}\left(\frac{t}{\theta}\right)^{n-1} \mathbb{1}_{[0,\theta]}(t).$$

   On peut désormais calculer $\mathbb{E}_\theta\left[\widehat{\theta_n}\right]$:

$$\mathbb{E}_\theta\left[\widehat{\theta_n}\right] = \int_{\mathbb{R}^+} t f_{\widehat{\theta_n}}(t)\,\mathrm{d}t = \frac{n}{\theta^n}\int_0^\theta t^n\,\mathrm{d}t = \frac{n}{\theta^n}\frac{\theta^{n+1}}{n+1} = \frac{n}{n+1}\theta.$$

   Donc $\widehat{\theta_n}$ est un estimateur biaisé[81], mais asymptotiquement sans biais car $\lim_{n\to\infty} \mathbb{E}_\theta\left[\widehat{\theta_n}\right] = \theta$.

79. Un lien vers la page personnelle de Caroline ROBET. On trouvera néanmoins une partie de la démonstration dans le livre de B. CADRE et C. VIAL – *Statistique mathématique*, Ellipses, 2012.
80. Vous m'autorisez à ne pas tracer la courbe sur cette page ? Promis, je la ferai au tableau.
81. C'était prévisible car $\widehat{\theta_n} \leqslant \theta$ $\mathbb{P}_\theta$-ps.

Diffusion à titre gratuit uniquement.

3. Soit $\varepsilon > 0$, on a :

$$\mathbb{P}_\theta\left(\left|\widehat{\theta_n} - \theta\right| \geqslant \varepsilon\right) = 1 - \mathbb{P}_\theta\left(\theta - \varepsilon < \widehat{\theta_n} < \theta + \varepsilon\right) \underset{82}{=} 1 - \left(F_{\widehat{\theta_n}}(\theta+\varepsilon) - F_{\widehat{\theta_n}}(\theta-\varepsilon)\right) = 1 - \left(1 - \left(\frac{\theta-\varepsilon}{\theta}\right)^n\right)$$
$$= \left(\frac{\theta-\varepsilon}{\theta}\right)^n$$

Pour ε suffisamment petit, $\left|\frac{\theta-\varepsilon}{\theta}\right| < 1$ donc $\sum_{n\geqslant 1} \mathbb{P}_\theta\left(\left|\widehat{\theta_n} - \theta\right| \geqslant \varepsilon\right)$ converge.

On va utiliser le lemme de Borel-Cantelli, notons A_n l'événement $\left\{\left|\widehat{\theta_n} - \theta\right| \geqslant \varepsilon\right\}$.

Comme $\sum_{n=1}^{\infty} \mathbb{P}_\theta(A_n) < \infty$, on a : $\mathbb{P}_\theta\left(\liminf_{n\to\infty} A_n^{\mathrm{c}}\right) = 1$.

En d'autres termes : $\forall \varepsilon > 0, \mathbb{P}_\theta\text{-ps}, \exists n \in \mathbb{N}^*, \forall k \geqslant n, \left|\widehat{\theta_k} - \theta\right| < \varepsilon$.

En particulier : $\forall p \in \mathbb{N}^*, \exists N_p\ \mathbb{P}_\theta\text{-négligeable}, \forall \omega \in N_p^{\mathrm{c}}, \exists n \in \mathbb{N}^*, \forall k \geqslant n, \left|\widehat{\theta_k}(\omega) - \theta\right| < \frac{1}{p}$.

On pose alors $N = \bigcup_{p\in\mathbb{N}^*} N_p$; N est un événement $\mathbb{P}_\theta$-négligeable et :

$$\forall \omega \in N^{\mathrm{c}}, \forall p \in \mathbb{N}^*, \exists n \in \mathbb{N}^*, \forall k \geqslant n, \left|\widehat{\theta_k}(\omega) - \theta\right| < \frac{1}{p}.$$

Autrement dit, on a : $\widehat{\theta_n} \underset{n\to\infty}{\overset{\mathbb{P}_\theta\text{-ps}}{\longrightarrow}} \theta$.

4. Soit $t \in \mathbb{R}^+$ et $n \geqslant \frac{t}{\theta}$:

$$\mathbb{P}_\theta\left(n\left(\theta - \widehat{\theta_n}\right) \geqslant t\right) = \mathbb{P}_\theta\left(\theta - \frac{t}{n} \geqslant \widehat{\theta_n}\right) = F_{\widehat{\theta_n}}\left(\theta - \frac{t}{n}\right) = \left(\frac{\theta - \frac{t}{n}}{\theta}\right)^n = \left(1 - \frac{t}{n\theta}\right)^n \underset{n\to\infty}{\longrightarrow} \mathrm{e}^{-\frac{t}{\theta}}.$$

Donc $\mathbb{P}_\theta\left(n\left(\theta - \widehat{\theta_n}\right) \leqslant t\right) \underset{n\to\infty}{\longrightarrow} 1 - \mathrm{e}^{-\frac{t}{\theta}}$, d'où $n\left(\theta - \widehat{\theta_n}\right) \underset{n\to\infty}{\overset{\mathbb{P}_\theta\text{-}\mathcal{L}}{\longrightarrow}} \mathcal{E}\left(\frac{1}{\theta}\right)$ et donc $n\left(\widehat{\theta_n} - \theta\right) \underset{n\to\infty}{\overset{\mathbb{P}_\theta\text{-}\mathcal{L}}{\longrightarrow}} -\mathcal{E}\left(\frac{1}{\theta}\right)$.

5. Par le lemme de transfert, on a :

$$\mathbb{E}_\theta\left[\widehat{\theta_n}^2\right] = \int_{\mathbb{R}^+} t^2 f_{\widehat{\theta_n}}(t)\,\mathrm{d}t = \frac{n}{\theta^n}\int_0^\theta t^{n+1}\,\mathrm{d}t = \frac{n}{\theta^n}\frac{\theta^{n+2}}{n+2} = \frac{n}{n+2}\theta^2.$$

Donc le risque quadratique vaut :

$$\mathbb{E}_\theta\left[\left(\widehat{\theta_n} - \theta\right)^2\right] = \mathbb{E}_\theta\left[\widehat{\theta_n}^2\right] - 2\theta\mathbb{E}_\theta\left[\widehat{\theta_n}\right] + \theta^2 = \left(\frac{n}{n+2} - 2\frac{n}{n+1} + 1\right)\theta^2 = \frac{2}{(n+1)(n+2)}\theta^2.$$

■

82. On utilise ici que $\widehat{\theta_n}$ est sans atome, car $\widehat{\theta_n}$ possède une densité.

Diffusion à titre gratuit uniquement.

Étude de $\mathrm{vp}\left(\frac{1}{x}\right)$

Leçons : 254, 255 [83]

[Zui], 2.3.iv-3.1.2.iii-3.2.3.(3)-10.3.6.11)

Théorème

La fonction $x \mapsto \frac{1}{x}$ n'est pas localement intégrable sur $\mathbb{R}$.

Cependant, on peut quand même lui associer une distribution appelée valeur principale de $\frac{1}{x}$ et notée $\mathrm{vp}\left(\frac{1}{x}\right)$, qu'on définit par :

$$\forall \varphi \in \mathcal{D}(\mathbb{R}), \left\langle \mathrm{vp}\left(\frac{1}{x}\right), \varphi \right\rangle = \lim_{\varepsilon \to 0} \int_{|x| \geqslant \varepsilon} \frac{\varphi(x)}{x} \, \mathrm{d}x$$

On va montrer les résultats suivants :

1. $\mathrm{vp}\left(\frac{1}{x}\right)$ est une distribution d'ordre 1 ;
2. $x\mathrm{vp}\left(\frac{1}{x}\right) = 1$ et donc $\mathrm{supp}\left(\mathrm{vp}\left(\frac{1}{x}\right)\right) = \mathbb{R}$;
3. $\frac{\partial}{\partial x}(\ln|x|) = \mathrm{vp}\left(\frac{1}{x}\right)$;
4. $\mathrm{vp}\left(\frac{1}{x}\right)$ est une distribution tempérée ;
5. $\widehat{\mathrm{vp}\left(\frac{1}{x}\right)} = -2\mathrm{i}\pi\mathrm{H} + \mathrm{i}\pi$.

Démonstration :

1. $\mathrm{vp}\left(\frac{1}{x}\right)$ est bien une forme linéaire sur $\mathcal{D}(\mathbb{R})$; montrons qu'elle vérifie la propriété de continuité.
   Soit K un compact de $\mathbb{R}$, et M tel que $K \subset [-M, M]$.
   Soit $\varphi \in \mathcal{D}(K)$, on a :
   $$\left\langle \mathrm{vp}\left(\frac{1}{x}\right), \varphi \right\rangle = \lim_{\varepsilon \to 0} \int_{\varepsilon \leqslant |x| \leqslant M} \frac{\varphi(x)}{x} \, \mathrm{d}x$$
   Par la formule de Taylor avec reste intégral :
   $$\varphi(x) = \varphi(0) + \int_0^1 \frac{(1-t)^0}{0!} \varphi'(tx) x \, \mathrm{d}t = \varphi(0) + x \int_0^1 \varphi'(tx) \, \mathrm{d}t$$
   On pose $\psi(x) = \int_0^1 \varphi'(tx) \, \mathrm{d}t$, et $\psi \in \mathcal{C}^\infty(\mathbb{R})$; de plus : $\forall x \in \mathbb{R}, |\psi(x)| \leqslant \|\varphi'\|_\infty$.
   Ainsi :
   $$\int_{\varepsilon \leqslant |x| \leqslant M} \frac{\varphi(x)}{x} \, \mathrm{d}x = \varphi(0) \underbrace{\int_{\varepsilon \leqslant |x| \leqslant M} \frac{\mathrm{d}x}{x}}_{I_1} + \underbrace{\int_{\varepsilon \leqslant |x| \leqslant M} \psi(x) \, \mathrm{d}x}_{I_2}$$
   La fonction $x \mapsto \frac{1}{x}$ étant intégrable sur $[\varepsilon, M]$ et impaire, on obtient que $I_1 = 0$.
   D'autre part, on a : $\left| \mathbb{1}_{]-\infty,-\varepsilon] \cup [\varepsilon,+\infty[}(x)\psi(x) \right| \leqslant |\psi(x)| \leqslant \|\varphi'\|_\infty$ intégrable sur $[-M, M]$.

83. On ne démontrera que les points 1, 2 et 3 dans la leçon 255 ; les points 2, 4 et 5 dans la leçon 254.

Diffusion à titre gratuit uniquement.

Donc, par convergence dominée, on a : $\lim_{\varepsilon\to 0} I_2 = \int_{|x|\leqslant M} \psi(x)\,\mathrm{d}x$.
Donc $\left\langle \mathrm{vp}\left(\frac{1}{x}\right), \varphi\right\rangle = \int_{|x|\leqslant M} \psi(x)\,\mathrm{d}x$, d'où : $\left|\left\langle \mathrm{vp}\left(\frac{1}{x}\right), \varphi\right\rangle\right| \leqslant \underbrace{2M}_{C_K} \|\varphi'\|_\infty$.

Donc $\mathrm{vp}\left(\frac{1}{x}\right)$ est une distribution d'ordre au plus 1. On veut montrer qu'elle est d'ordre 1 exactement : par l'absurde, on va supposer qu'elle est d'ordre 0.
Alors, pour tout compact $K \subset \mathbb{R}$:

$$\exists C_K > 0, \forall \varphi \in \mathcal{D}(K), \left|\left\langle \mathrm{vp}\left(\frac{1}{x}\right), \varphi\right\rangle\right| \leqslant C_K \|\varphi\|_\infty$$

Soit donc $K = [0,2]$; pour $n \in \mathbb{N}^*$, on pose $\varphi_n \in \mathcal{D}(K)$ telle que : $\begin{cases} 0 \leqslant \varphi_n \leqslant 1 \text{ sur } \mathbb{R} \\ \varphi_n \equiv 1 \text{ sur } \left[\frac{1}{n}, 1\right] \\ \varphi_n \equiv 0 \text{ sur } \left]-\infty, \frac{1}{2n}\right] \cup [2, +\infty[\end{cases}$.

Alors, pour $\varepsilon \leqslant \dfrac{1}{2n}$, on a :

$$\int_{|x|\geqslant\varepsilon} \frac{\varphi_n(x)}{x}\,\mathrm{d}x = \int_{\frac{1}{2n}}^{2} \frac{\varphi_n(x)}{x}\,\mathrm{d}x \geqslant \int_{\frac{1}{n}}^{1} \frac{1}{x}\,\mathrm{d}x = \ln n$$

Donc $\left|\left\langle \mathrm{vp}\left(\frac{1}{x}\right), \varphi_n\right\rangle\right| \geqslant \ln n \underbrace{\|\varphi_n\|_\infty}_{=1}$, ainsi $\forall n \in \mathbb{N}^*, C_K \geqslant \ln n$ et on aboutit à une contradiction.

2. Soit $\varphi \in \mathcal{D}(\mathbb{R})$, on a :

$$\left\langle x\mathrm{vp}\left(\frac{1}{x}\right), \varphi\right\rangle = \left\langle \mathrm{vp}\left(\frac{1}{x}\right), x\varphi\right\rangle = \lim_{\varepsilon\to 0} \int_{|x|\geqslant\varepsilon} \varphi(x)\,\mathrm{d}x = \int_{\mathbb{R}} \varphi(x)\,\mathrm{d}x = \langle 1, \varphi\rangle$$

On a utilisé au passage la convergence dominée car $\left|\mathbb{1}_{]-\infty,-\varepsilon]\cup[\varepsilon,+\infty[}(x)\varphi(x)\right| \leqslant |\varphi(x)|$, qui est intégrable sur $\mathbb{R}$.
Et comme $x\mathrm{vp}\left(\frac{1}{x}\right) = 1$, le support de la distribution $\mathrm{vp}\left(\frac{1}{x}\right)$ contient celui de la fonction constante et égale à 1, autrement dit $\mathbb{R}$.
Par conséquent, $\mathrm{supp}\left(\mathrm{vp}\left(\frac{1}{x}\right)\right) = \mathbb{R}$.

3. Soit $f : x \mapsto \ln|x|$; $f \in \mathrm{L}^1_{\mathrm{loc}}(\mathbb{R})$ donc définit une distribution. [84]
Soit $\varphi \in \mathcal{D}(\mathbb{R})$, on a :

$$\langle f', \varphi\rangle = -\langle f, \varphi'\rangle = -\int_{\mathbb{R}} \ln|x|\varphi'(x)\,\mathrm{d}x = -\lim_{\varepsilon\to 0} \int_{|x|\geqslant\varepsilon} \ln|x|\varphi'(x)\,\mathrm{d}x$$

Là encore [85], on a utilisé la convergence dominée, car : $\left|\mathbb{1}_{]-\infty,-\varepsilon]\cup[\varepsilon,+\infty[}(x)\ln|x|\varphi'(x)\right| \leqslant |\ln|x|\varphi'(x)|$ qui est intégrable, car φ est continue à support compact et f est localement intégrable.
On pose :

$$\begin{aligned} I_\varepsilon &= \int_{-\infty}^{-\varepsilon} \ln(-x)\varphi'(x)\,\mathrm{d}x + \int_{\varepsilon}^{+\infty} \ln(x)\varphi'(x)\,\mathrm{d}x \\ &= [\ln(-x)\varphi(x)]_{-\infty}^{-\varepsilon} - \int_{-\infty}^{-\varepsilon} \frac{\varphi(x)}{x}\,\mathrm{d}x + [\ln(x)\varphi(x)]_{\varepsilon}^{+\infty} - \int_{\varepsilon}^{+\infty} \frac{\varphi(x)}{x}\,\mathrm{d}x \\ &= \ln\varepsilon\varphi(-\varepsilon) - \int_{|x|\geqslant\varepsilon} \frac{\varphi(x)}{x}\,\mathrm{d}x - \ln\varepsilon\varphi(\varepsilon) \\ &= \ln\varepsilon[\varphi(-\varepsilon) - \varphi(\varepsilon)] - \int_{|x|\geqslant\varepsilon} \frac{\varphi(x)}{x}\,\mathrm{d}x \end{aligned}$$

84. C'est une fonction continue hors de 0, et il ne se pose de problème qu'en 0. On utilise le critère de Riemann ; en effet : $\sqrt{x}\ln x \underset{x\to 0^+}{\longrightarrow} 0$, donc $\ln x = o_{0^+}\left(\frac{1}{\sqrt{x}}\right)$ est donc intégrable au voisinage de 0.

85. Si vous vous ennuyez en lisant ce développement, dites-vous bien que je me suis aussi ennuyé à l'écrire. Bah oui.

Diffusion à titre gratuit uniquement.

Or, on a, pour $x \in \mathbb{R}$, $\varphi(x) = \varphi(0) + x\psi(x)$, où $\psi \in \mathcal{C}^\infty(\mathbb{R})$, donc :

$$I_\varepsilon = \underbrace{-\varepsilon \ln \varepsilon}_{\underset{\varepsilon\to 0}{\longrightarrow} 0} \underbrace{[\psi(-\varepsilon) + \psi(\varepsilon)]}_{\underset{\varepsilon\to 0}{\longrightarrow} 2\psi(0)} - \int_{|x|\geqslant\varepsilon} \frac{\varphi(x)}{x}\,\mathrm{d}x \underset{\varepsilon\to 0}{\longrightarrow} -\left\langle \mathrm{vp}\left(\frac{1}{x}\right), \varphi \right\rangle$$

D'où $\langle f', \varphi\rangle = \left\langle \mathrm{vp}\left(\frac{1}{x}\right), \varphi \right\rangle$.

4. Soit $\varphi \in \mathcal{S}(\mathbb{R})$.
   Soit $\varepsilon \in]0,1[$, on a : $\displaystyle\int_{|x|\geqslant\varepsilon} \frac{\varphi(x)}{x}\,\mathrm{d}x = \int_{\varepsilon\leqslant|x|\leqslant 1} \frac{\varphi(x)}{x}\,\mathrm{d}x + \int_{|x|\geqslant 1} \frac{\varphi(x)}{x}\,\mathrm{d}x$.
   Or $x \mapsto \dfrac{\varphi(0)}{x}$ étant intégrable sur $[\varepsilon, 1]$ et impaire : $\displaystyle\int_{\varepsilon\leqslant|x|\leqslant 1} \frac{\varphi(0)}{x}\,\mathrm{d}x = 0$.
   Ainsi :

$$\begin{aligned}\int_{|x|\geqslant\varepsilon} \frac{\varphi(x)}{x}\,\mathrm{d}x &= \int_{\varepsilon\leqslant|x|\leqslant 1} \frac{\varphi(x)-\varphi(0)}{x}\,\mathrm{d}x + \int_{|x|\geqslant 1} \frac{\varphi(x)}{x}\,\mathrm{d}x \\ &= \int_{\varepsilon\leqslant|x|\leqslant 1} \psi(x)\,\mathrm{d}x + \int_{|x|\geqslant 1} \frac{\varphi(x)}{x}\,\mathrm{d}x\end{aligned}$$

   Or, par convergence dominée [86], on obtient :

$$\left\langle \mathrm{vp}\left(\frac{1}{x}\right), \varphi \right\rangle = \int_{|x|\leqslant 1} \psi(x)\,\mathrm{d}x + \int_{|x|\geqslant 1} \frac{\varphi(x)}{x}\,\mathrm{d}x$$

   Et on en déduit alors :

$$\left|\left\langle \mathrm{vp}\left(\frac{1}{x}\right), \varphi \right\rangle\right| \leqslant 2\|\psi\|_\infty + \int_{|x|\geqslant 1} \frac{\mathrm{d}x}{x^2}\,\|x \mapsto x\varphi(x)\|_\infty \leqslant 2\left\|\varphi'\right\|_\infty + 2\,\|x \mapsto x\varphi(x)\|_\infty$$

   Ce qui montre que $\mathrm{vp}\left(\dfrac{1}{x}\right)$ est tempérée.

5. Pour plus de commodité, je vous propose de noter désormais $T = \mathrm{vp}\left(\dfrac{1}{x}\right)$.
   De l'égalité $xT = 1$, on déduit $\widehat{xT} = \widehat{1} = 2\pi\delta_0$.
   En conséquence, $-\dfrac{1}{\mathrm{i}}\dfrac{\mathrm{d}}{\mathrm{d}\xi}\left(\widehat{T}\right) = 2\pi\delta_0$, d'où $\widehat{T} = -2\mathrm{i}\pi\mathrm{H} + C$, avec $C \in \mathbb{C}$ à déterminer.
   Soit $\varphi \in \mathcal{S}(\mathbb{R})$, $\langle \widehat{T \circ (-\mathrm{Id})}, \varphi\rangle = \langle T \circ (-\mathrm{Id}), \widehat{\varphi}\rangle = \langle T, \widehat{\varphi} \circ (-\mathrm{Id})\rangle$. [87]
   Mais on a : $\widehat{\varphi}(-x) = \dfrac{1}{2\pi}\displaystyle\int_{-\infty}^{+\infty} \varphi(t)\mathrm{e}^{\mathrm{i}xt}\,\mathrm{d}t = \frac{1}{2\pi}\int_{-\infty}^{+\infty} \varphi(-u)\mathrm{e}^{-\mathrm{i}xu}\,\mathrm{d}u = \widehat{\varphi \circ (-\mathrm{Id})}(x)$.
   Donc $\langle \widehat{T \circ (-\mathrm{Id})}, \varphi\rangle = \langle T, \widehat{\varphi \circ (-\mathrm{Id})}\rangle = \langle \widehat{T}, \varphi \circ (-\mathrm{Id})\rangle = \langle \widehat{T} \circ (-\mathrm{Id}), \varphi\rangle$.

$$\begin{aligned}\text{Mais } \langle T \circ (-\mathrm{Id}), \varphi\rangle &= \lim_{\varepsilon\to 0}\left(\int_{-\infty}^{-\varepsilon} \frac{\varphi(-x)}{x}\,\mathrm{d}x + \int_{\varepsilon}^{+\infty} \frac{\varphi(-x)}{x}\,\mathrm{d}x\right) \\ &= \lim_{\varepsilon\to 0}\left(\int_{+\infty}^{+\varepsilon} \frac{\varphi(x)}{x}\,\mathrm{d}x + \int_{-\varepsilon}^{-\infty} \frac{\varphi(x)}{x}\,\mathrm{d}x\right) = -\langle T, \varphi\rangle.\end{aligned}$$

   On en déduit donc que $\widehat{T} \circ (-\mathrm{Id}) = \widehat{T \circ (-\mathrm{Id})} = -\widehat{T}$.
   On prend alors $\varphi \in \mathcal{S}(\mathbb{R})$, avec $\operatorname{supp}\varphi \subset \mathbb{R}^+$, et $\displaystyle\int_{\mathbb{R}} \varphi = 1$.
   Ainsi : $C = \langle \widehat{T} \circ (-\mathrm{Id}), \varphi\rangle = -\langle \widehat{T}, \varphi\rangle = 2\mathrm{i}\pi - C$, d'où $C = \mathrm{i}\pi$. ■

Références

[Zui] C. ZUILY – *Éléments de distributions et d'équations aux dérivées partielles*, Dunod, 2002.

86. On va peut-être pouvoir arrêter de détailler maintenant...
87. Notez qu'on a ici utilisé le fait que le jacobien de $(-\mathrm{Id})$ vaut 1.

Diffusion à titre gratuit uniquement.

Formule des compléments

Leçons : 236, 245, *207, 235, 239*

[AM], section 8.4.4

Théorème

On rappelle qu'on définit la fonction Gamma d'Euler par :

$$\forall z \in \{s \in \mathbb{C} | \Re(s) > 0\}, \Gamma(z) = \int_0^{+\infty} t^{z-1} \mathrm{e}^{-t} \, \mathrm{d}t$$

On a l'égalité suivante :

$$\forall z \in \{s \in \mathbb{C} | 0 < \Re(s) < 1\}, \Gamma(z)\Gamma(1-z) = \frac{\pi}{\sin \pi z}$$

On commence par montrer le lemme qui suit.

Lemme

On a l'égalité suivante :

$$\forall \alpha \in]0,1[, \int_0^{+\infty} \frac{\mathrm{d}t}{t^\alpha (1+t)} = \frac{\pi}{\sin \pi\alpha}$$

Démonstration du lemme :

$\forall \alpha \in]0,1[$, on définit $I_\alpha := \displaystyle\int_0^{+\infty} \frac{\mathrm{d}t}{t^\alpha(1+t)}$.

I_α est bien définie car c'est l'intégrale d'une fonction mesurable positive ; on a même $I_\alpha < +\infty$. En effet :

- $t \mapsto \dfrac{1}{t^\alpha(1+t)}$ est continue sur $]0,+\infty[$ (donc localement intégrable) ;
- En 0 : $\dfrac{1}{t^\alpha(1+t)} \underset{t\to 0}{\sim} \dfrac{1}{t^\alpha}$, qui est intégrable car $0 < \alpha < 1$;
- En $+\infty$: $\dfrac{1}{t^\alpha(1+t)} \underset{t\to+\infty}{\sim} \dfrac{1}{t^{\alpha+1}}$, qui est intégrable car $\alpha + 1 > 1$.

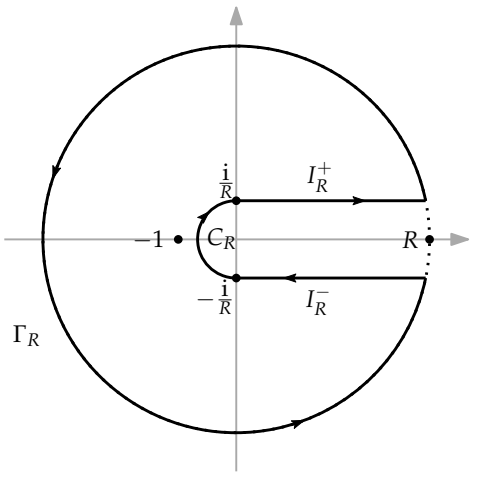
On note $\Omega = \mathbb{C}\backslash\mathbb{R}^+$ et $f : \left| \begin{array}{ccc} \Omega\backslash\{-1\} & \to & \mathbb{C} \\ z & \mapsto & \frac{1}{z^\alpha(1+z)} \end{array} \right.$, où l'on convient $z^\alpha = r^\alpha \mathrm{e}^{\mathrm{i}\alpha\theta}$ quand $z = r\mathrm{e}^{\mathrm{i}\theta}$, où $\theta \in]0, 2\pi[$.

La fonction f est holomorphe sur $\Omega\backslash\{-1\}$ et possède un pôle simple en -1 avec :

$$\mathrm{Res}(f,-1) = \lim_{z\to -1} (1+z)f(z) = \frac{1}{(-1)^\alpha} = \mathrm{e}^{-\mathrm{i}\pi\alpha}$$

Pour $R > 1$, on définit le chemin $\gamma_R = C_R \cup I_R^+ \cup \Gamma_R \cup I_R^-$, où :

- $C_R = \left\{ \dfrac{1}{R}\mathrm{e}^{\mathrm{i}\theta} \middle| \theta \in \left[\dfrac{\pi}{2}, \dfrac{3\pi}{2}\right] \right\}$;
- $I_R^\pm = \left[\pm\dfrac{\mathrm{i}}{R}, \pm\dfrac{\mathrm{i}}{R} + \sqrt{R^2 - \dfrac{1}{R^2}} \right]$;
- $\Gamma_R = \left\{ R\mathrm{e}^{\mathrm{i}\theta} \middle| \theta \in [\theta_R, 2\pi - \theta_R] \right\}$, avec $\theta_R = \arcsin \dfrac{\frac{1}{R}}{R}$.



Le théorème des résidus donne donc :

$$\forall R > 1, \int_{\gamma_R} f(z) \, \mathrm{d}z = 2\mathrm{i}\pi \mathrm{e}^{-\mathrm{i}\pi\alpha}$$

Diffusion à titre gratuit uniquement.

On va passer à la limite quand $R \to +\infty$.
Tout d'abord :

$$\left|\int_{C_R} f(z)\,\mathrm{d}z\right| = \left|\int_{\frac{\pi}{2}}^{\frac{3\pi}{2}} f\left(\frac{1}{R}\mathrm{e}^{\mathrm{i}\theta}\right)\mathrm{i}\frac{1}{R}\mathrm{e}^{\mathrm{i}\theta}\,\mathrm{d}\theta\right| \leqslant \int_{\frac{\pi}{2}}^{\frac{3\pi}{2}} \frac{\frac{1}{R}}{\left(\frac{1}{R}\right)^{\alpha}\left|1+\frac{1}{R}\mathrm{e}^{\mathrm{i}\theta}\right|}\,\mathrm{d}\theta \leqslant \pi\frac{\left(\frac{1}{R}\right)^{1-\alpha}}{1-\frac{1}{R}} \underset{R\to+\infty}{\longrightarrow} 0$$

Aussi :

$$\left|\int_{\Gamma_R} f(z)\,\mathrm{d}z\right| = \left|\int_0^{2\pi} \mathbb{1}_{[\theta_R, 2\pi-\theta_R]}(\theta)\frac{\mathrm{i}R\mathrm{e}^{\mathrm{i}\theta}}{R^{\alpha}\mathrm{e}^{\mathrm{i}\alpha\theta}\left(1+R\mathrm{e}^{\mathrm{i}\theta}\right)}\,\mathrm{d}\theta\right| \leqslant \int_0^{2\pi}\frac{R}{R^{\alpha}\left|1+R\mathrm{e}^{\mathrm{i}\theta}\right|}\,\mathrm{d}\theta \leqslant 2\pi\frac{R^{1-\alpha}}{R-1} \underset{R\to+\infty}{\longrightarrow} 0$$

De plus : $\displaystyle\int_{I_R^+} f(z)\,\mathrm{d}z = \int_0^{\sqrt{R^2-\frac{1}{R^2}}} f\left(\frac{\mathrm{i}}{R}+t\right)\,\mathrm{d}t = \int_0^{\sqrt{R^2-\frac{1}{R^2}}} \frac{1}{\left(t+\frac{\mathrm{i}}{R}\right)^{\alpha}\left(1+t+\frac{\mathrm{i}}{R}\right)}\,\mathrm{d}t.$

Comme $\left(t+\dfrac{\mathrm{i}}{R}\right)^{\alpha} = \left(\sqrt{t^2+\dfrac{1}{R^2}}\exp\left(\mathrm{i}\arctan\dfrac{\frac{1}{R}}{t}\right)\right)^{\alpha} \underset{R\to+\infty}{\longrightarrow} t^{\alpha}$, on a :

- $\mathbb{1}_{\left]0,\sqrt{R^2-\frac{1}{R^2}}\right]}(t) f\left(\dfrac{\mathrm{i}}{R}+t\right) \underset{R\to+\infty}{\longrightarrow} \mathbb{1}_{\mathbb{R}^{+*}}(t)\dfrac{1}{t^{\alpha}(1+t)}$;
- $\left|\mathbb{1}_{\left]0,\sqrt{R^2-\frac{1}{R^2}}\right]}(t) f\left(\dfrac{\mathrm{i}}{R}+t\right)\right| \leqslant \mathbb{1}_{\mathbb{R}^{+*}}(t)\dfrac{1}{t^{\alpha}(1+t)}$ qui est intégrable.

Par théorème de convergence dominée, on déduit :

$$\lim_{R\to+\infty}\int_{I_R^+} f(z)\,\mathrm{d}z = I_{\alpha}$$

Enfin, de la même façon, en utilisant le fait que $\left(t-\dfrac{\mathrm{i}}{R}\right)^{\alpha} \underset{R\to+\infty}{\longrightarrow} t^{\alpha}\mathrm{e}^{2\mathrm{i}\pi\alpha}$, on a :

$$\lim_{R\to+\infty}\int_{I_R^-} f(z)\,\mathrm{d}z = -\mathrm{e}^{-2\mathrm{i}\pi\alpha} I_{\alpha}$$

Donc $\left(1-\mathrm{e}^{-2\mathrm{i}\pi\alpha}\right) I_{\alpha} = 2\mathrm{i}\pi\mathrm{e}^{-\mathrm{i}\pi\alpha}$, c'est-à-dire :

$$I_{\alpha} = \frac{\pi}{\sin\pi\alpha}$$

■

Démonstration du théorème :
D'après le théorème des zéros isolés, il suffit de prouver l'égalité pour $z = \alpha \in]0,1[$. Soit donc $\alpha \in]0,1[$.
En utilisant le théorème de Fubini, on obtient :

$$\begin{aligned}\Gamma(\alpha)\Gamma(1-\alpha) &= \left(\int_0^{+\infty} t^{\alpha-1}\mathrm{e}^{-t}\,\mathrm{d}t\right)\left(\int_0^{+\infty} s^{-\alpha}\mathrm{e}^{-s}\,\mathrm{d}s\right) = \int_0^{+\infty}\int_0^{+\infty} s^{-\alpha}t^{\alpha-1}\mathrm{e}^{-t-s}\,\mathrm{d}t\,\mathrm{d}s \\ &= \int_0^{+\infty}\int_0^{+\infty}\left(\frac{t}{s}\right)^{\alpha}\mathrm{e}^{-(s+t)}\,\mathrm{d}s\,\frac{\mathrm{d}t}{t}\end{aligned}$$

On réalise le changement de variables donné par le système $\left\{\begin{array}{rcl} u &=& s+t \\ v &=& \frac{s}{t}\end{array}\right.$ et dont le jacobien vaut

$$\left|\det\begin{pmatrix} 1 & 1 \\ \frac{1}{t} & \frac{-s}{t^2}\end{pmatrix}\right| = \frac{1}{t}+\frac{s}{t^2} = \frac{1}{t}+\frac{v}{t} = \frac{v+1}{t}$$

On en déduit donc :

$$\begin{aligned}\Gamma(\alpha)\Gamma(1-\alpha) &= \int_0^{+\infty}\int_0^{+\infty} v^{-\alpha}\mathrm{e}^{-u}\,\frac{\mathrm{d}u\,\mathrm{d}v}{v+1} = \int_0^{+\infty}\frac{1}{v^{\alpha}(v+1)}\int_0^{+\infty}\mathrm{e}^{-u}\,\mathrm{d}u\,\mathrm{d}v = \int_0^{+\infty}\frac{\mathrm{d}v}{v^{\alpha}(v+1)} \\ &= \frac{\pi}{\sin\pi\alpha}\end{aligned}$$

■

Références

[AM] É. AMAR et É. MATHERON – *Analyse complexe*, Cassini, 2004.

Diffusion à titre gratuit uniquement.

Formule sommatoire de Poisson

Leçons : 240, 246, 254, 255, *228, 235, 241* [88]

[Gou An], problème 4.4 et exercice 3.4
[Wil], exemple 7.29.f)

Théorème

Soit $f \in \mathcal{S}(\mathbb{R})$.
Alors la série $\sum_{n\in\mathbb{Z}} f(.+n)$ converge normalement sur tout compact de $\mathbb{R}$ et :

$$\forall x \in \mathbb{R}, \sum_{n\in\mathbb{Z}} f(x+n) = \sum_{n\in\mathbb{Z}} \widehat{f}(n)\mathrm{e}^{2\mathrm{i}\pi nx},$$

où on a noté $\widehat{f}(x) = \int_{-\infty}^{+\infty} f(t)\mathrm{e}^{-2\mathrm{i}\pi xt}\,\mathrm{d}t$, pour $x \in \mathbb{R}$.

Démonstration :

1. Comme $f \in \mathcal{S}(\mathbb{R})$, en particulier, $\forall k \in \mathbb{N}, f^{(k)}(x) = \mathcal{O}_{+\infty}\left(\frac{1}{x^2}\right)$.
   Ainsi, $\exists M > 0, \forall x \in \mathbb{R}, |x| \geqslant 1 \Rightarrow |f(x)| \leqslant \frac{M}{x^2}$.
   D'où $\forall K > 0, \forall x \in [-K, K], \forall n \in \mathbb{Z}, |n| > K+1 \Rightarrow |f(x+n)| \leqslant \frac{M}{(x+n)^2} \leqslant \frac{M}{(|n|-K)^2}$.
   Donc la série $\sum_{n\in\mathbb{Z}} f(.+n)$ converge normalement sur tout compact. [89]
   On note F sa limite simple.
   De façon similaire, on montre que $\sum_{n\in\mathbb{Z}} f'(.+n)$ converge normalement sur tout compact, donc uniformément sur tout segment de $\mathbb{R}$.
   On peut donc appliquer le théorème de dérivation des séries de fonctions (on rappelle que f est $\mathcal{C}^1$) : F est donc de classe $\mathcal{C}^1$ sur $\mathbb{R}$ (en fait le théorème dit d'abord "sur tout segment de $\mathbb{R}$") et $\forall x \in \mathbb{R}, F'(x) = \sum_{n\in\mathbb{Z}} f'(x+n)$.

2. Par ailleurs, soit $x \in \mathbb{R}$: $\forall N \in \mathbb{N}, \sum_{n=-N}^{N} f(x+1+n) = \sum_{n=-N+1}^{N+1} f(x+n)$.
   Donc en faisant tendre N vers l'infini, on obtient $F(x+1) = F(x)$; F est donc 1-périodique.
   On va calculer ses coefficients de Fourier, pour $N \in \mathbb{Z}$:

$$\begin{aligned} c_N(F) &= \int_0^1 F(t)\mathrm{e}^{-2\mathrm{i}\pi Nt}\,\mathrm{d}t \underset{\text{(cvu)}}{=} \sum_{n\in\mathbb{Z}} \int_0^1 f(t+n)\mathrm{e}^{-2\mathrm{i}\pi Nt}\,\mathrm{d}t = \sum_{n\in\mathbb{Z}} \int_n^{n+1} f(t)\mathrm{e}^{-2\mathrm{i}\pi Nt}\mathrm{e}^{-2\mathrm{i}\pi Nn}\,\mathrm{d}t \\ &\underset{90}{=} \int_{-\infty}^{+\infty} f(t)\mathrm{e}^{-2\mathrm{i}\pi Nt}\,\mathrm{d}t = \widehat{f}(N) \end{aligned}$$

   Comme F est $\mathcal{C}^1$ sur $\mathbb{R}$, sa série de Fourier converge normalement vers F sur $\mathbb{R}$ et :

$$\forall x \in \mathbb{R}, F(x) = \sum_{n\in\mathbb{Z}} \widehat{f}(n)\mathrm{e}^{2\mathrm{i}\pi nx}.$$ [91]

■

88. Dans les leçons sur les distributions, on fait le corollaire 1, dans les autres, on fait le corollaire 2.
89. Car à partir de $|n| > K+1$, on a : $\left\| f(.+n)_{|[-K,K]} \right\|_\infty = \mathcal{O}\left(\frac{1}{n^2}\right)$.
90. L'intégrale converge car $f(t) = \mathcal{O}_{+\infty}\left(\frac{1}{t^2}\right)$.
91. Rappelons que la période de F vaut 1.

Diffusion à titre gratuit uniquement.

Corollaire (*Une distribution invariante par transformation de Fourier*)

On note $\delta_{\mathbb{Z}} = \sum_{k\in\mathbb{Z}} \delta_k$.
Dans $\mathcal{S}'(\mathbb{R})$, on a : $\delta_{\mathbb{Z}} = \widehat{\delta_{\mathbb{Z}}} = \sum_{k\in\mathbb{Z}} \mathrm{e}^{2\mathrm{i}\pi k}$.

Démonstration :

1. Montrons que $\delta_{\mathbb{Z}}$ définit bien un élément de $\mathcal{S}'(\mathbb{R})$.
   D'après ce qu'on vient de faire : $\forall \varphi \in \mathcal{S}(\mathbb{R}), \langle \delta_{\mathbb{Z}}, \varphi\rangle = \sum_{k\in\mathbb{Z}} \varphi(k)$ est bien défini.
   Reste à montrer que $\delta_{\mathbb{Z}}$ est une distribution tempérée.
   On rappelle que $\|.\|_{n,p} : \varphi \mapsto \sup_{x\in\mathbb{R}} \left|x^n \varphi^{(p)}(x)\right|$, où $n, p \in \mathbb{N}$, sont les semi-normes qui définissent la topologie de $\mathcal{S}(\mathbb{R})$.

$$\begin{aligned}\forall \varphi \in \mathcal{S}(\mathbb{R}), |\langle \delta_{\mathbb{Z}}, \varphi\rangle| &\leqslant \sum_{k\in\mathbb{Z}} |\varphi(k)| = |\varphi(0)| + \sum_{k\in\mathbb{Z}^*} \frac{1}{k^2}\left|k^2\varphi(k)\right| \\ &\leqslant \|\varphi\|_{0,0} + \sum_{k\in\mathbb{Z}^*} \frac{1}{k^2}\|\varphi\|_{2,0} \leqslant \frac{\pi^2}{3}\left(\|\varphi\|_{0,0} + \|\varphi\|_{2,0}\right)\end{aligned}$$

   Ainsi, $\delta_{\mathbb{Z}} \in \mathcal{S}'(\mathbb{R})$.
2. On peut donc calculer sa transformée de Fourier, en utilisant la transformée de Fourier en 0 :

$$\forall \varphi \in \mathcal{S}(\mathbb{R}), \langle \widehat{\delta_{\mathbb{Z}}}, \varphi\rangle = \langle \delta_{\mathbb{Z}}, \widehat{\varphi}\rangle = \sum_{n\in\mathbb{Z}} \widehat{\varphi}(n) = \sum_{n\in\mathbb{Z}} \varphi(n) = \langle \delta_{\mathbb{Z}}, \varphi\rangle.$$

   Donc $\widehat{\delta_{\mathbb{Z}}} = \delta_{\mathbb{Z}}$.
   D'autre part, on a, pour $n \in \mathbb{Z}$, $\varphi \in \mathcal{S}(\mathbb{R})$:

$$\langle \widehat{\delta_n}, \varphi\rangle = \langle \delta_n, \widehat{\varphi}\rangle = \widehat{\varphi(n)} = \int_{-\infty}^{+\infty} \varphi(x)\mathrm{e}^{-2\mathrm{i}\pi n x}\,\mathrm{d}x = \langle \mathrm{e}^{-2\mathrm{i}\pi n x}, \varphi\rangle.$$

   Ainsi, dans $\mathcal{S}'(\mathbb{R})$, $\widehat{\delta_{\mathbb{Z}}} = \sum_{k\in\mathbb{Z}} \mathrm{e}^{-2\mathrm{i}\pi k x} = \sum_{k\in\mathbb{Z}} \mathrm{e}^{2\mathrm{i}\pi k x}$. ■

Corollaire (*Une égalité entre deux sommes*)

$$\forall s > 0, \sum_{n\in\mathbb{Z}} \mathrm{e}^{-\pi n^2 s} = \frac{1}{\sqrt{s}} \sum_{n\in\mathbb{Z}} \mathrm{e}^{-\frac{\pi n^2}{s}}.$$

Démonstration :

Soit $\alpha > 0$, on va appliquer la formule sommatoire de Poisson à $f : x \mapsto \mathrm{e}^{-\alpha x^2}$.
Soit $n \in \mathbb{Z}$, $\widehat{f}(n) = \int_{-\infty}^{+\infty} \mathrm{e}^{-\alpha t^2}\mathrm{e}^{-2\mathrm{i}\pi n t}\,\mathrm{d}t = \frac{1}{\sqrt{\alpha}}\int_{-\infty}^{+\infty} \mathrm{e}^{-u^2}\mathrm{e}^{-2\mathrm{i}\pi n\frac{u}{\sqrt{\alpha}}}\,\mathrm{d}u = \frac{1}{\sqrt{\alpha}} I(n)$, où on a posé $u = \sqrt{\alpha}t$ et $\forall x \in \mathbb{R}, I(x) := \int_{-\infty}^{+\infty} \mathrm{e}^{-u^2}\mathrm{e}^{-2\mathrm{i}\pi x\frac{u}{\sqrt{\alpha}}}\,\mathrm{d}u$.
On va chercher une équation différentielle vérifiée par I ;

→ I est dérivable, en effet : posons $h : (x,u) \mapsto \mathrm{e}^{-u^2}\mathrm{e}^{-2\mathrm{i}\pi x\frac{u}{\sqrt{\alpha}}}$.
- $\forall x \in \mathbb{R}, h(x,.)$ est $\mathcal{C}^1$ et intégrable (par comparaison avec l'intégrale de Gauss) ;
- $\forall x, u \in \mathbb{R}, \frac{\partial h}{\partial x}(x,u) = -2\mathrm{i}\pi\frac{u}{\sqrt{\alpha}}\mathrm{e}^{-u^2}\mathrm{e}^{-2\mathrm{i}\pi x\frac{u}{\sqrt{\alpha}}}$;
- $\frac{\partial h}{\partial x}$ est continue sur $\mathbb{R}^2$ et $\forall x, u \in \mathbb{R}, \left|\frac{\partial h}{\partial x}(x,u)\right| \leqslant \frac{2\pi u}{\sqrt{\alpha}}\mathrm{e}^{-u^2}$, fonction majorante à la fois intégrable et indépendante de x.

On en déduit en plus que $I'(x) = \frac{-2\mathrm{i}\pi}{\sqrt{\alpha}}\int_{-\infty}^{+\infty} u\mathrm{e}^{-u^2}\mathrm{e}^{-2\mathrm{i}\pi x\frac{u}{\sqrt{\alpha}}}\,\mathrm{d}u$.

Diffusion à titre gratuit uniquement.

$\rightarrow$ Par une intégration par parties :

$$\begin{aligned} I(x) &= \left[\mathrm{e}^{-u^2}\frac{-\sqrt{\alpha}}{2\mathrm{i}\pi x}\mathrm{e}^{-2\mathrm{i}\pi x\frac{u}{\sqrt{\alpha}}}\right]_{-\infty}^{+\infty} - \int_{-\infty}^{+\infty} -2u\mathrm{e}^{-u^2}\frac{-\sqrt{\alpha}}{2\mathrm{i}\pi x}\mathrm{e}^{-2\mathrm{i}\pi x\frac{u}{\sqrt{\alpha}}}\,\mathrm{d}u \\ &= 0 - \frac{\sqrt{\alpha}}{\mathrm{i}\pi x}\int_{-\infty}^{+\infty} u\mathrm{e}^{-u^2}\mathrm{e}^{-2\mathrm{i}\pi x\frac{u}{\sqrt{\alpha}}}\,\mathrm{d}u \\ &= -\frac{\sqrt{\alpha}}{\mathrm{i}\pi x}\frac{\sqrt{\alpha}}{-2\mathrm{i}\pi}I'(x) = -\frac{\alpha}{2\pi^2 x}I'(x) \end{aligned}$$

Donc $I(x) = I(0)\exp\left(-\dfrac{\pi^2x^2}{\alpha}\right) = \sqrt{\pi}\exp\left(-\dfrac{\pi^2x^2}{\alpha}\right)$.

Ainsi, $\widehat{f}(n) = \sqrt{\dfrac{\pi}{\alpha}}\exp\left(-\dfrac{\pi^2n^2}{\alpha}\right)$.

On applique la formule de Poisson en 0 : $\displaystyle\sum_{n\in\mathbb{Z}}\mathrm{e}^{-\alpha n^2} = \sum_{n\in\mathbb{Z}}\sqrt{\frac{\pi}{\alpha}}\mathrm{e}^{-\frac{\pi^2n^2}{\alpha}}$.

On pose enfin $s = \dfrac{\pi}{\alpha}$, et alors : $\displaystyle\sum_{n\in\mathbb{Z}}\mathrm{e}^{-\frac{n^2\pi}{s}} = \sqrt{s}\sum_{n\in\mathbb{Z}}\mathrm{e}^{-\pi n^2 s}$. ■

Références

[Gou An] X. GOURDON – *Les maths en tête : Analyse*, 2e éd., Ellipses, 2008.
[Wil] M. WILLEM – *Analyse harmonique réelle*, Hermann, 1997.

Diffusion à titre gratuit uniquement.

Harmonicité et propriété de la moyenne

Leçons : 222, *215, 228*

[X-ENS An4], exercices 1.23 et 1.24

Théorème

1. Soient U un ouvert de $\mathbb{R}^2$ et $f \in \mathcal{C}^2(U, \mathbb{R})$ vérifiant la propriété de la moyenne, c'est-à-dire telle que :
$$\forall (x_0, y_0) \in U, \exists r_0 > 0, \forall r \in]0, r_0[, f(x_0, y_0) = \frac{1}{2\pi} \int_0^{2\pi} f(x_0 + r\cos\theta, y_0 + r\sin\theta)\, \mathrm{d}\theta$$
Alors $\Delta f \equiv 0$ sur U (on dit que f est harmonique sur U).
2. Réciproquement, si U est un ouvert de $\mathbb{R}^2$ et si $f \in \mathcal{C}^2(U, \mathbb{R})$ vérifie $\Delta f \equiv 0$,
Alors f vérifie la propriété de la moyenne sur U.

Démonstration :

1. Soient $(x_0, y_0) \in U$ et r_0 donné par l'hypothèse.
On pose $m : r \mapsto \displaystyle\int_0^{2\pi} f(x_0 + r\cos\theta, y_0 + r\sin\theta)\, \mathrm{d}\theta$ pour $r \in [0, r_0[$.
Par hypothèse, m est constante, égale à $2\pi f(x_0, y_0)$.
Pour simplifier, on note $g : (r, \theta) \mapsto f(x_0 + r\cos\theta, y_0 + r\sin\theta)$; g est continue sur $[0, r_0[\times [0, 2\pi]$ et admet des dérivées partielles $\dfrac{\partial g}{\partial r}$ et $\dfrac{\partial^2 g}{\partial r^2}$ sont continues sur $[0, 2\pi]$ à r fixé dans $[0, r_0[$.
Par théorème de dérivation, on obtient les dérivées de m sur $[0, r_0[$:
$$\forall r \in [0, r_0[, 0 = m'(r) = \int_0^{2\pi} \left(\cos\theta \frac{\partial f}{\partial x}(x_0 + r\cos\theta, y_0 + r\sin\theta) + \sin\theta \frac{\partial f}{\partial y}(x_0 + r\cos\theta, y_0 + r\sin\theta) \right) \mathrm{d}\theta$$
Puis en redérivant, en utilisant le théorème de Schwarz (les dérivées partielles sont toutes prises en $(x_0 + r\cos\theta, y_0 + r\sin\theta)$ pour plus de simplicité) :
$$\forall r \in [0, r_0[, 0 = m''(r) = \int_0^{2\pi} \left(\cos^2\theta \frac{\partial^2 f}{\partial x^2} + 2\cos\theta\sin\theta \frac{\partial^2 f}{\partial x \partial y} + \sin^2\theta \frac{\partial^2 f}{\partial y^2} \right) \mathrm{d}\theta$$
En particulier, quand on prend $r = 0$ dans cette égalité, on obtient :
$$\begin{aligned} 0 &= \frac{\partial^2 f}{\partial x^2}(x_0, y_0) \underbrace{\int_0^{2\pi} \frac{\cos 2\theta + 1}{2}\, \mathrm{d}\theta}_{=\pi} + \frac{\partial^2 f}{\partial x \partial y}(x_0, y_0) \underbrace{\int_0^{2\pi} \sin 2\theta\, \mathrm{d}\theta}_{=0} + \frac{\partial^2 f}{\partial y^2}(x_0, y_0) \underbrace{\int_0^{2\pi} \frac{1 - \cos 2\theta}{2}\, \mathrm{d}\theta}_{=\pi} \\ &= \pi \Delta f(x_0, y_0) \end{aligned}$$
Et f est donc harmonique sur U.
2. Par translation, on se ramène au cas où $(0, 0) \in U$ et il suffit de ne montrer la propriété de la moyenne qu'en ce point.
Comme U est ouvert, il existe $R > 0$ telle que $\mathcal{B}(0, R) \subset U$.
La fonction $g : (r, \theta) \mapsto f(r\cos\theta, r\sin\theta)$ est continue sur $[0, R[\times [0, 2\pi]$ et admet une dérivée partielle $\dfrac{\partial g}{\partial r}$ est continue sur $[0, 2\pi]$ à r fixé dans $[0, R[$.
On en déduit que $F : r \mapsto \displaystyle\int_0^{2\pi} g(r, \theta)\, \mathrm{d}\theta$ est de classe $\mathcal{C}^1$ sur $[0, R[$.
De plus, $\forall r \in [0, R[, F'(r) = \displaystyle\int_0^{2\pi} \left(\cos\theta \frac{\partial f}{\partial x}(r\cos\theta, r\sin\theta) + \sin\theta \frac{\partial f}{\partial y}(r\cos\theta, r\sin\theta) \right) \mathrm{d}\theta$.
On considère la forme différentielle $\omega = -\dfrac{\partial f}{\partial y}(x, y)\mathrm{d}x + \dfrac{\partial f}{\partial x}(x, y)\mathrm{d}y$ et soit $\Gamma = \mathcal{C}(0, r)$, paramétré par $\theta \mapsto (r\cos\theta, r\sin\theta)$, où $r \in [0, R[$.

Diffusion à titre gratuit uniquement.

Ainsi, $\displaystyle\int_\Gamma \omega = \int_0^{2\pi} \left(r\sin\theta \frac{\partial f}{\partial y}(r\cos\theta, r\sin\theta) + r\cos\theta \frac{\partial f}{\partial x}(r\cos\theta, r\sin\theta) \right) \mathrm{d}\theta = rF'(r)$.

Mais f est harmonique, donc $\displaystyle\frac{\partial}{\partial y}\left(-\frac{\partial f}{\partial y}\right) = -\frac{\partial^2 f}{\partial y^2} = \frac{\partial^2 f}{\partial x^2} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right)$ et ω est une forme différentielle fermée... et comme $\mathcal{D}(0,R)$ est étoilé, par le théorème de Poincaré [92], ω est exacte.

Ainsi, comme Γ est un arc fermé, on en déduit que $\forall r \in [0,R[, rF'(r) = \displaystyle\int_\Gamma \omega = 0$.

Donc $\forall r \in]0,R[, F'(r) = 0$, et par continuité en 0, F est donc constante sur $[0,R[$ et demeure égale à $F(0) = 2\pi f(0,0)$.

Donc, finalement : $\exists R > 0, \forall r \in [0,R[, f(0,0) = \dfrac{1}{2\pi}\displaystyle\int_0^{2\pi} f(r\cos\theta, r\sin\theta)\,\mathrm{d}\theta$. ■

Références

[X-ENS An4] S. FRANCINOU, H. GIANELLA et S. NICOLAS – *Oraux X-ENS Analyse 4*, 1re éd., Cassini, 2012.

92. Comme on l'utilise rarement, un petit rappel sur le théorème de Poincaré. Dans $\mathbb{R}^n$, on écrit $\omega = \displaystyle\sum_{k=1}^n \alpha_k e_k^*$ (où $\left(e_k^*\right)_{1\leqslant k\leqslant n}$ est la base duale de la base canonique de $\mathbb{R}^n$) et on souhaiterait réussir à montrer que ω possède pour "primitive" la fonction $f : x \mapsto \displaystyle\int_0^1 \omega((1-t)a + tx)(x-a)\,\mathrm{d}t$, où a est au cœur de l'ouvert étoilé U ; en pratique, pour tout $i \in [\![1,n]\!]$, on va montrer l'égalité $\dfrac{\partial f}{\partial x_i}(x) = \alpha_i(x)$. Pour cela, on utilise le théorème de dérivation des intégrales à paramètre.

Diffusion à titre gratuit uniquement.

Inégalité de Hoeffding et application [93]

Leçons : 253, 260, 261, 262, *229*

[Ouv2], exercice 10.11

Théorème

Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires réelles, indépendantes et centrées.
De plus, on suppose que $\forall n \in \mathbb{N}^*$, X_n est ps bornée par c_n, où $c_n > 0$.
On note : $\forall n \in \mathbb{N}^*, S_n = \sum_{j=1}^{n} X_j$ et $a_n = \sum_{j=1}^{n} c_j^2$.
Alors $\forall \varepsilon > 0, \forall n \in \mathbb{N}^*, \mathbb{P}(|S_n| > \varepsilon) \leqslant 2\exp\left(-\frac{\varepsilon^2}{2a_n}\right)$.

Démonstration :

→ Il va s'agir de démontrer le lemme qui suit.

Lemme

Soit X une variable aléatoire réelle, centrée, et ps bornée par 1.
On note L_X sa transformée de Laplace.
On a : $\forall t \in \mathbb{R}, L_X(t) \leqslant \exp\left(\frac{t^2}{2}\right)$.

Démonstration :

Soit $t \in \mathbb{R}$, on a : $\forall x \in [-1,1], tx = \frac{1-x}{2}(-t) + \frac{1+x}{2}t$.

Donc, par convexité de l'exponentielle, on en déduit : $\forall x \in [-1,1], \mathrm{e}^{tx} \leqslant \frac{1-x}{2}\mathrm{e}^{-t} + \frac{1+x}{2}\mathrm{e}^{t}$.

Mais on sait que $|X| \leqslant 1$ ps ; en particulier, e^{tX} est bornée ps donc admet un moment d'ordre 1.

Ainsi, L_X est bien définie en t et :

$$L_X(t) \leqslant \frac{\mathbb{E}[1-X]}{2}\mathrm{e}^{-t} + \frac{\mathbb{E}[1+X]}{2}\mathrm{e}^{t} = \frac{1}{2}\left(\mathrm{e}^{-t} + \mathrm{e}^{t}\right) = \operatorname{ch} t = \sum_{n=0}^{\infty}\frac{t^{2n}}{(2n)!} \leqslant \sum_{n=0}^{\infty}\frac{t^{2n}}{n!2^n} = \exp\left(\frac{t^2}{2}\right). \quad \blacksquare$$

→ Fixons $n \in \mathbb{N}^*$.

On applique le lemme aux variables $\frac{X_j}{c_j}$, où $j \in [\![1,n]\!]$: $\forall t \in \mathbb{R}, L_{X_j}(t) = L_{\frac{X_j}{c_j}}(tc_j) \leqslant \exp\left(\frac{t^2}{2}c_j^2\right)$.

Mais par indépendance des $\exp\left(tX_j\right)$ pour $j \in [\![1,n]\!]$, il vient :

$$\forall t \in \mathbb{R}, L_{S_n}(t) = \prod_{j=1}^{n} L_{X_j}(t) \leqslant \exp\left(\frac{t^2}{2}a_n\right).$$

→ Soient $t > 0$ et $\varepsilon > 0$; on a : $S_n > \varepsilon \Leftrightarrow \mathrm{e}^{tS_n} > \mathrm{e}^{t\varepsilon}$.

Ainsi, par l'inégalité de Markov, on obtient :

$$\mathbb{P}(S_n > \varepsilon) = \mathbb{P}\left(\mathrm{e}^{tS_n} > \mathrm{e}^{t\varepsilon}\right) \leqslant \frac{\mathbb{E}\left[\mathrm{e}^{tS_n}\right]}{\mathrm{e}^{t\varepsilon}} = \mathrm{e}^{-t\varepsilon}L_{S_n}(t) \leqslant \exp\left(-t\varepsilon + \frac{t^2}{2}a_n\right).$$

Cette inégalité est vraie pour tout $t > 0$, donc en particulier pour $t = \frac{\varepsilon}{a_n}$, où $-t\varepsilon + \frac{t^2}{2}a_n$ réalise son minimum, d'où : $\mathbb{P}(S_n > \varepsilon) \leqslant \exp\left(\frac{\varepsilon^2}{a_n^2}\frac{a_n}{2} - \frac{\varepsilon}{a_n}\varepsilon\right) = \exp\left(-\frac{\varepsilon^2}{2a_n}\right)$.

→ Soit $\varepsilon > 0$ quelconque.

On a : $\mathbb{P}(|S_n| > \varepsilon) = \mathbb{P}(S_n > \varepsilon) + \mathbb{P}(S_n < -\varepsilon)$.

Mais on aurait pu appliquer tout ce qu'on vient de faire aux variables $-X_j$, où $j \in [\![1,n]\!]$, et donc :

$\mathbb{P}(S_n < -\varepsilon) = \mathbb{P}(-S_n > \varepsilon) \leqslant \exp\left(-\frac{\varepsilon^2}{2a_n}\right)$.

Et finalement, $\mathbb{P}(|S_n| > \varepsilon) \leqslant 2\exp\left(-\frac{\varepsilon^2}{2a_n}\right)$. $\blacksquare$

Diffusion à titre gratuit uniquement.

Corollaire

Soit $\alpha > 0$; on ajoute l'hypothèse supplémentaire : $\exists \beta > 0, \forall n \in \mathbb{N}^*, a_n \leqslant n^{2\alpha-\beta}$.
Alors : $\dfrac{S_n}{n^\alpha} \xrightarrow[n\to\infty]{\text{ps}} 0$.

Démonstration :

Soit $\varepsilon > 0$, on a, par Hoeffding : $\forall n \in \mathbb{N}^*, \mathbb{P}\left(|S_n| > n^\alpha \varepsilon\right) \leqslant 2\exp\left(-\dfrac{n^{2\alpha}\varepsilon^2}{2a_n}\right) \leqslant 2\exp\left(-\dfrac{n^\beta \varepsilon^2}{2}\right)$.

Mais la série $\displaystyle\sum_{n\geqslant 1} \exp\left(-\frac{n^\beta\varepsilon^2}{2}\right)$ converge (par le critère de Riemann), car $\exists N \in \mathbb{N}^*, \forall n \leqslant N, \dfrac{\varepsilon^2}{2} n^\beta \geqslant 2\ln n$

et donc $\exists N \in \mathbb{N}^*, \forall n \leqslant N, 0 \leqslant \exp\left(-\dfrac{n^\beta\varepsilon^2}{2}\right) \geqslant \dfrac{1}{n^2}$.

Ainsi, la série $\displaystyle\sum_{n\geqslant 1} \mathbb{P}\left(|S_n| > n^\alpha\varepsilon\right)$ converge, d'où, par Borel-Cantelli :

$$\forall \varepsilon > 0, \mathbb{P}\left(\limsup_{n\to\infty} \left\{|S_n| > n^\alpha\varepsilon\right\}\right) = 0, \text{ ie } \mathbb{P}\left(\liminf_{n\to\infty} \left\{|S_n| \leqslant n^\alpha\varepsilon\right\}\right) = 1.$$

En particulier, $\forall p \in \mathbb{N}^*, \mathbb{P}\left(\displaystyle\bigcup_{n\in\mathbb{N}^*} \bigcap_{k\geqslant n} \left\{\left|\frac{S_k}{k^\alpha}\right| \leqslant \frac{1}{p}\right\}\right) = 1$.

C'est-à-dire : $\forall p \in \mathbb{N}^*, \exists N_p$ négligeable, $\forall \omega \in N_p^c, \exists n \in \mathbb{N}, \forall k \geqslant n, \left|\dfrac{S_k(\omega)}{k^\alpha}\right| \leqslant \dfrac{1}{p}$.

On pose alors $N = \displaystyle\bigcup_{p\in\mathbb{N}^*} N_p$, alors N est négligeable et :

$$\forall \omega \in N^c, \forall p \in \mathbb{N}^*, \exists n \in \mathbb{N}, \forall k \geqslant n, \left|\frac{S_k(\omega)}{k^\alpha}\right| \leqslant \frac{1}{p}.$$

Ou, en d'autres termes : $\dfrac{S_n}{n^\alpha} \xrightarrow[n\to\infty]{\text{ps}} 0$. ■

Références

[Ouv2] J.-Y. OUVRARD – *Probabilités 2*, 3e éd., Cassini, 2009.

93. Allez, une autre application, si vous aimez les statistiques ; elle provient de la partie 3.2 du livre *Statistique mathématique*, de B. CADRE et C. VIAL, paru en 2012 aux éditions Ellipses. Plaçons-nous dans un modèle statistique $\left(\mathcal{H}^n, \{P_\theta\}_{\theta\in\Theta}\right)$ avec $\mathcal{H} \subset \mathbb{R}$ et $\Theta \subset \mathbb{R}^d$. Le paramètre d'intérêt est $g(\theta)$ avec $g : \Theta \to \mathbb{R}$ une fonction continue On suppose que $X_1, \ldots, X_n$ sont indépendantes et identiquement distribuées, de loi P_θ, bornées P_θ-ps et avec $\mathbb{E}_\theta[X_1] = g(\theta)$. Soit c une borne P_θ-presque sûre de $X_1 - g(\theta)$. On veut un intervalle de confiance pour $g(\theta)$.

Par Hoeffding, $P_\theta\left(\left|\overline{X_n} - g(\theta)\right| > \varepsilon\right) = P_\theta\left(\left|\displaystyle\sum_{j=1}^n \left(X_j - g(\theta)\right)\right| > n\varepsilon\right) \leqslant 2\exp\left(-\dfrac{n^2\varepsilon^2}{2nc^2}\right) = 2\exp\left(-\dfrac{n\varepsilon^2}{2c^2}\right)$. Soit $\alpha \in]0,1[$, on choisit ε de sorte que $\alpha = 2\exp\left(-\dfrac{n\varepsilon^2}{2c^2}\right)$, ie : $\varepsilon = c\sqrt{\dfrac{2}{n}\ln\dfrac{2}{\alpha}}$. Dès lors, on obtient l'intervalle de confiance par excès au niveau $1-\alpha$, pour $g(\theta)$: $I_\alpha = \left[\overline{X_n} - c\sqrt{\dfrac{2}{n}\ln\dfrac{2}{\alpha}}, \overline{X_n} + c\sqrt{\dfrac{2}{n}\ln\dfrac{2}{\alpha}}\right]$.

Diffusion à titre gratuit uniquement.

Intégrale de Fresnel

Leçons : 236, 239, *235*

[Gou An], exercice 5.4.5

Théorème

On a l'égalité suivante :

$$\int_0^{+\infty} \mathrm{e}^{\mathrm{i}x^2}\,\mathrm{d}x = \frac{\sqrt{\pi}}{2}\mathrm{e}^{\mathrm{i}\frac{\pi}{4}}$$

Démonstration :

On pose $\Phi = \displaystyle\int_0^{+\infty} \mathrm{e}^{\mathrm{i}x^2}\,\mathrm{d}x$.

Pour $t, T \geqslant 0$, on pose : $f(t) = \displaystyle\int_0^t \mathrm{e}^{\mathrm{i}x^2}\,\mathrm{d}x$, $F(t) = \displaystyle\iint_{[0,t]^2} \mathrm{e}^{\mathrm{i}\left(x^2+y^2\right)}\,\mathrm{d}x\mathrm{d}y$ et $I(T) = \dfrac{1}{T}\displaystyle\int_0^T F(t)\,\mathrm{d}t$.

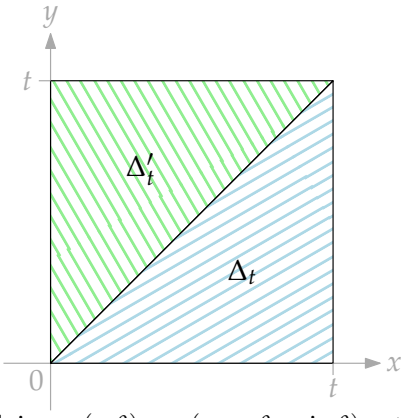
On va exprimer $F(t)$ de deux manières.

D'une part, la fonction $(x,y) \mapsto \mathrm{e}^{\mathrm{i}\left(x^2+y^2\right)}$ est continue sur le pavé compact $[0,t]^2$. En appliquant Fubini, on obtient : $F(t) = f(t)^2$.

D'autre part, notons :

$$\Delta_t = \left\{(x,y) \in \mathbb{R}^2 \,\middle|\, x \in [0,t] \text{ et } y \in [0,x]\right\}$$

$$\Delta'_t = \left\{(x,y) \in \mathbb{R}^2 \,\middle|\, y \in [0,t] \text{ et } x \in [0,y]\right\}$$



La fonction $(x,y) \mapsto (y,x)$ est un $\mathcal{C}^1$-difféomorphisme de Δ_t sur Δ'_t.
Par changement de variable, on obtient :

$$F(t) = \iint_{\Delta_t} \mathrm{e}^{\mathrm{i}\left(x^2+y^2\right)}\,\mathrm{d}x\mathrm{d}y + \iint_{\Delta'_t} \mathrm{e}^{\mathrm{i}\left(x^2+y^2\right)}\,\mathrm{d}x\mathrm{d}y = 2\iint_{\Delta_t} \mathrm{e}^{\mathrm{i}\left(x^2+y^2\right)}\,\mathrm{d}x\mathrm{d}y$$

La forme de l'intégrande suggère alors un passage en coordonnées polaires : $(r,\theta) \mapsto (r\cos\theta, r\sin\theta)$ est un $\mathcal{C}^1$-difféomorphisme de $K_t := \left\{(r,\theta) \in \mathbb{R}^2 \,\middle|\, \theta \in \left[0, \dfrac{\pi}{4}\right] \text{ et } r \in \left[0, \dfrac{t}{\cos\theta}\right]\right\}$ sur Δ_t.

On obtient alors :

$$F(t) = 2\iint_{K_t} \mathrm{e}^{\mathrm{i}r^2} r\,\mathrm{d}r\mathrm{d}\theta = \frac{1}{\mathrm{i}}\int_0^{\frac{\pi}{4}}\int_0^{\frac{t}{\cos\theta}} 2\mathrm{i}r\mathrm{e}^{\mathrm{i}r^2}\,\mathrm{d}r\mathrm{d}\theta = -\mathrm{i}\int_0^{\frac{\pi}{4}} \left(\exp\left(\frac{\mathrm{i}t^2}{\cos^2\theta}\right) - 1\right)\mathrm{d}\theta = \frac{\mathrm{i}\pi}{4} - \mathrm{i}\int_0^{\frac{\pi}{4}} \exp\left(\frac{\mathrm{i}t^2}{\cos^2\theta}\right)\mathrm{d}\theta$$

Puis, en injectant dans $I(T)$, on obtient :

$$I(T) = \frac{\mathrm{i}\pi}{4} - \frac{\mathrm{i}}{T}\int_0^T\int_0^{\frac{\pi}{4}} \exp\left(\frac{\mathrm{i}t^2}{\cos^2\theta}\right)\mathrm{d}\theta\mathrm{d}t$$

En appliquant Fubini, puis le changement de variable $u = \dfrac{t}{\cos\theta}$, on trouve :

$$I(T) = \frac{\mathrm{i}\pi}{4} - \frac{\mathrm{i}}{T}\int_0^{\frac{\pi}{4}}\int_0^T \exp\left(\frac{\mathrm{i}t^2}{\cos^2\theta}\right)\mathrm{d}t\mathrm{d}\theta = \frac{\mathrm{i}\pi}{4} - \frac{\mathrm{i}}{T}\int_0^{\frac{\pi}{4}}\int_0^{\frac{T}{\cos\theta}} \mathrm{e}^{\mathrm{i}u^2}\cos\theta\,\mathrm{d}u\mathrm{d}\theta = \frac{\mathrm{i}\pi}{4} - \frac{\mathrm{i}}{T}\int_0^{\frac{\pi}{4}} f\left(\frac{T}{\cos\theta}\right)\cos\theta\,\mathrm{d}\theta$$

On souhaite désormais faire tendre T vers $+\infty$. On va montrer que f est bornée au voisinage de $+\infty$. Il suffit de montrer que l'intégrale de Fresnel est convergente. Étudions $\displaystyle\int_1^t \mathrm{e}^{\mathrm{i}x^2}\,\mathrm{d}x$.

Par changement de variable $u = x^2$ et en faisant une intégration par parties :

$$\int_1^t \mathrm{e}^{\mathrm{i}x^2}\,\mathrm{d}x = \int_1^{t^2} \mathrm{e}^{\mathrm{i}u}\frac{\mathrm{d}u}{2\sqrt{u}} = \left[\frac{\mathrm{e}^{\mathrm{i}u}}{\mathrm{i}}\frac{1}{2\sqrt{u}}\right]_1^{t^2} - \int_1^{t^2}\frac{\mathrm{e}^{\mathrm{i}u}}{\mathrm{i}}\frac{-1}{4u^{\frac{3}{2}}}\,\mathrm{d}u = \underbrace{\frac{\mathrm{e}^{\mathrm{i}t^2}}{2\mathrm{i}t}}_{\substack{\longrightarrow 0 \\ t\to\infty}} \underbrace{-\frac{\mathrm{e}^{\mathrm{i}}}{2\mathrm{i}} + \frac{1}{4\mathrm{i}}\int_1^{t^2}\frac{\mathrm{e}^{\mathrm{i}u}}{u^{\frac{3}{2}}}\,\mathrm{d}u}_{\substack{\text{quantité bornée par} \\ \text{convergence dominée}}}$$

Diffusion à titre gratuit uniquement.

Donc f est bornée au voisinage de $+\infty$, donc $\lim\limits_{T\to+\infty} I(T) = \mathrm{i}\dfrac{\pi}{4}$.
Par ailleurs, $\lim\limits_{t\to+\infty} f(t) = \Phi$ et $\lim\limits_{t\to+\infty} F(t) = \Phi^2$. Montrons que $\lim\limits_{T\to\infty} I(T) = \Phi^2$.
Soit $\varepsilon > 0$; $\exists A > 0, \forall t \geqslant A, \left|F(t) - \Phi^2\right| < \dfrac{\varepsilon}{2}$. Soit $T > A$, on a :

$$I(T) = \frac{1}{T}\int_0^A F(t)\,\mathrm{d}t + \frac{1}{T}\int_A^T F(t)\,\mathrm{d}t \text{ puis } \left|I(T) - \Phi^2\right| \leqslant \underbrace{\frac{1}{T}\int_0^A \left|F(t) - \Phi^2\right|\,\mathrm{d}t}_{<\frac{\varepsilon}{2} \text{ pour } T\geqslant T_0} + \underbrace{\frac{1}{T}\int_A^T \left|F(t) - \Phi^2\right|\,\mathrm{d}t}_{<\frac{\varepsilon}{2}\frac{T-A}{T}}$$

Donc pour $T \geqslant \max\{A, T_0\}$, $\left|I(T) - \Phi^2\right| < \varepsilon$, d'où $\lim\limits_{T\to\infty} I(T) = \Phi^2$.
Par unicité de la limite : $\Phi^2 = \mathrm{i}\dfrac{\pi}{4}$, d'où $\Phi = \pm\dfrac{\sqrt{\pi}}{2}\mathrm{e}^{\mathrm{i}\frac{\pi}{4}}$.
Pour déterminer le bon signe, on regarde le signe de Im (Φ) :

$$\begin{aligned}\text{Im}\,(\Phi) &= \text{Im}\left(\int_0^\infty \mathrm{e}^{\mathrm{i}x^2}\,\mathrm{d}x\right) = \text{Im}\left(\int_0^\infty \mathrm{e}^{\mathrm{i}u}\,\frac{\mathrm{d}u}{2\sqrt{u}}\right) = \int_0^\infty \frac{\sin u}{2\sqrt{u}}\,\mathrm{d}u = \sum_{k=0}^{\infty}\int_{2k\pi}^{2(k+1)\pi}\frac{\sin u}{2\sqrt{u}}\,\mathrm{d}u \\ &= \sum_{k=0}^{\infty}\int_{2k\pi}^{(2k+1)\pi}\frac{\sin u}{2\sqrt{u}}\,\mathrm{d}u + \int_{(2k+1)\pi}^{2(k+1)\pi}\frac{\sin u}{2\sqrt{u}}\,\mathrm{d}u = \sum_{k=0}^{\infty}\int_{2k\pi}^{(2k+1)\pi}\frac{\sin u}{2}\left(\frac{1}{\sqrt{u}} - \frac{1}{\sqrt{u+\pi}}\right)\mathrm{d}u \geqslant 0\end{aligned}$$

Ainsi, $\Phi = \dfrac{\sqrt{\pi}}{2}\mathrm{e}^{\mathrm{i}\frac{\pi}{4}}$. ■

Références

[Gou An] X. GOURDON – *Les maths en tête : Analyse*, 2e éd., Ellipses, 2008.

Méthode de Newton [94]

Leçons : 226, 229, 232, 253, *206, 218, 219, 223, 224, 228*

[Rou], exercice 49

Théorème

Soit $f : [c,d] \to \mathbb{R}$ une fonction de classe $\mathcal{C}^2$.
On suppose $f(c) < 0 < f(d)$ et $f' > 0$ sur $[c,d]$; f s'annule donc en un unique point de $]c,d[$, noté a.
Pour $x_0 \in [c,d]$, on pose tant qu'on peut $x_{n+1} = F(x_n)$, où $F : x \mapsto x - \dfrac{f(x)}{f'(x)}$.

1. Il existe $C > 0$, tel que si $|x_0 - a| < \dfrac{1}{C}$, alors $\forall n \in \mathbb{N}, x_n$ est bien défini et $|x_n - a| \leqslant |x_0 - a|$.
   Dans ce cas, (x_n) converge vers a à vitesse quadratique.
2. Si $f'' > 0$ sur $[c,d]$ et $x_0 > a$, alors la suite (x_n) est bien définie, $\forall n \in \mathbb{N}, x_n > a$.
   Dans ce cas, on a l'équivalent : $x_{n+1} - a \underset{n\to\infty}{\sim} \dfrac{f''(a)}{2f'(a)}(x_n - a)^2$.

Démonstration :

Étape 1 : Soit $x \in [c,d]$, comme $f(a) = 0$, on a :

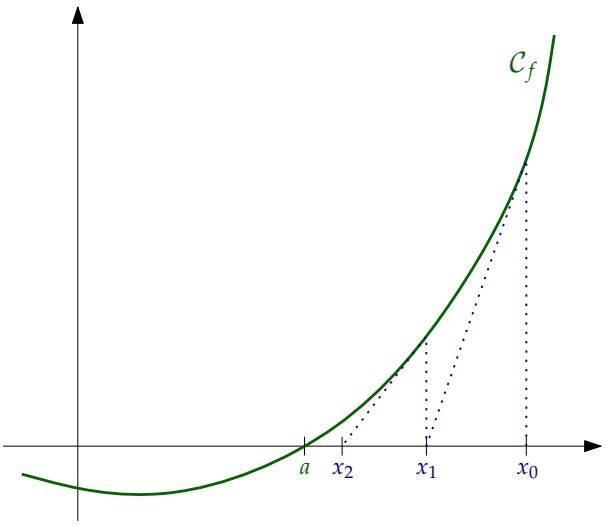
$$\begin{aligned} F(x) - a &= x - \frac{f(x)}{f'(x)} - a \\ &= x - a - \frac{f(x) - f(a)}{f'(x)} \\ &= \frac{f(a) - f(x) - (a-x)f'(x)}{f'(x)} \end{aligned}$$

Et par l'égalité de Taylor-Lagrange, il existe z compris strictement entre a et x, tel que :

$$f(a) - f(x) - (a-x)f'(x) = \frac{(a-x)^2}{2} f''(z)$$

On en déduit donc :

$$F(x) - a = \frac{f''(z)}{2f'(x)}(x-a)^2$$



Étape 2 : Montrons la première partie du théorème.

On pose $C = \dfrac{\max|f''|}{2\min|f'|}$; C est bien défini car f'' et f' sont continues sur le segment $[c,d]$, intervalle sur lequel $f' > 0$.
On a alors : $|F(x) - a| \leqslant C|x-a|^2$.
Soit $\alpha \in \left]0, \dfrac{1}{C}\right[$, tel que $I := [a-\alpha, a+\alpha] \subset [c,d]$;
Alors $\forall x \in I, |F(x) - a| \leqslant C\alpha^2 < \alpha$, donc $F(x) \in I$ puis $F(I) \subset I$.

94. On dispose d'une application. Soit $y \in \mathbb{R}^{+*}$; imaginons qu'on veuille estimer $a = \sqrt{y}$. Soit $f : x \mapsto x^2 - y$, définie sur un intervalle $[c,d]$, avec $0 < c < d$ et $c^2 < y < d^2$. Pour approcher a, on doit itérer la fonction $F(x) = x - \dfrac{x^2-y}{2x}$.
On a alors : $F(x) - a = \dfrac{(x-a)^2}{2x}$ et $F(x) + a = \dfrac{(x+a)^2}{2x}$.
Donc, en prenant $x_0 \in]a,d]$ et en posant $x_n = F^n(x_0)$, on obtient : $\dfrac{x_n + a}{x_n - a} = \left(\dfrac{x_0 + a}{x_0 - a}\right)^{2^n}$.
Par conséquent : $1 + \dfrac{2a}{x_n - a} = \left(1 + \dfrac{2a}{x_0 - a}\right)^{2^n} \geqslant 1 + \left(\dfrac{2a}{x_0 - a}\right)^{2^n}$.
On obtient donc un encadrement de l'erreur : $0 < x_n - a \leqslant 2a\left(\dfrac{x_0 - a}{2a}\right)^{2^n}$.

Diffusion à titre gratuit uniquement.

Par conséquent, si $x_0 \in I$, alors $\forall n \in \mathbb{N}, x_n \in I$ et $|x_{n+1} - a| = |F(x_n) - a| \leqslant C|x_n - a|^2$.
Et par une récurrence immédiate, il vient :

$$C|x_n - a| \leqslant (C|x_{n-1} - a|)^2 \leqslant \ldots \leqslant (C|x_0 - a|)^{2^n} \leqslant (C\alpha)^{2^n}$$

Ce qui prouve la convergence d'ordre 2 de (x_n) vers a car $C\alpha < 1$, dans le cas où $|x_0 - a| \leqslant \alpha$.

Étape 3 : Utilisons désormais l'hypothèse supplémentaire : $f'' > 0$ sur le segment $[c, d]$.

Pour $x \in]a, d]$, on a : $F(x) = x - \dfrac{f(x)}{f'(x)} < x$ car $f' > 0$ et f s'annule en a.

D'autre part : $\exists z \in]a, x[, F(x) - a = \dfrac{f''(z)}{2f'(x)}(x - a)^2 > 0$ car f'' et f' sont strictement positives.

Ainsi, $\forall x \in]a, d], a < F(x) < x \leqslant d$, donc $]a, d]$ est stable par F.
On a même : si $x_0 \in]a, d]$, alors $\forall n \in \mathbb{N}, x_n \in]a, d]$ et la suite (x_n) décroît.
Comme (x_n) est également minorée par a, la suite (x_n) converge ; on note $l \in [a, d]$ sa limite.
l est un point fixe de F donc par conséquent $f(l) = 0$ et donc $l = a$.
Comme dans le cas précédent, $|x_{n+1} - a| \leqslant C|x_n - a|^2$ et donc la convergence est quadratique.

Étape 4 : Enfin, cette inégalité est essentiellement optimale.

Si $a < x_0 \leqslant d$, alors $\forall n \in \mathbb{N}, x_n \in]a, d]$ et $\exists z_n \in]a, x_n[$, $\dfrac{x_{n+1} - a}{(x_n - a)^2} = \dfrac{1}{2}\dfrac{f''(z_n)}{f'(x_n)}$.

Par continuité de f' et f'', on déduit $\dfrac{f''(z_n)}{2f'(x_n)} \underset{n\to\infty}{\longrightarrow} \dfrac{f''(a)}{2f'(a)}$, d'où finalement :

$$x_{n+1} - a \underset{n\to\infty}{\sim} \frac{f''(a)}{2f'(a)}(x_n - a)^2$$

■

Références

[Rou] F. ROUVIÈRE – *Petit guide de calcul différentiel*, 4e éd., Cassini, 2014.

Diffusion à titre gratuit uniquement.

Méthode des petits pas

Leçons : 224

[Rom], partie 8.5.4

Théorème

Soit $f :]-1,1[\to \mathbb{R}$ une fonction continue, et admettant un développement limité en 0 de la forme : $f(x) = x - \alpha x^{p+1} + \beta x^{2p+1} + o\left(x^{2p+1}\right)$, avec $\alpha > 0$, $\beta \in \mathbb{R}^*$ et $p > 0$.
Sous de bonnes conditions initiales, la suite définie par $x_{n+1} = f(x_n)$ vérifie :

$$x_n = \frac{1}{\sqrt[p]{pn\alpha}} - \frac{\gamma}{p^2\alpha^2\sqrt[p]{p\alpha}} \frac{\ln n}{n\sqrt[p]{n}} + o\left(\frac{\ln n}{n\sqrt[p]{n}}\right), \text{ quand } \gamma := \frac{(1+p)\alpha^2}{2} - \beta \neq 0.$$

Démonstration :

Étape 1 : On peut écrire $f(x) = xg(x)$ et $x - f(x) = \alpha x^{p+1}h(x)$, avec $g(x)$ et $h(x)$ qui tendent vers 1 quand x tend vers 0.
Ainsi, $\exists \eta > 0, \forall x \in [-\eta, \eta], g(x) > 0$ et $h(x) > 0$.
En particulier, $\forall x \in]0, \eta], 0 < f(x) < x \leqslant \eta$.
$]0, \eta]$ est stable par f donc si $x_0 \in]0, \eta]$, (x_n) est bien définie et à valeurs dans $]0, \eta]$.

Étape 2 : De plus, $\forall n \in \mathbb{N}, 0 < x_{n+1} < x_n$, donc (x_n) est décroissante et minorée par 0.
Donc elle converge vers un point fixe l de f, car f est continue, avec $l \in [0, \eta]$.
De l'inégalité $f(x) < x$ valable sur $]0, \eta]$, on déduit $l = 0$.

Étape 3 : Soit $\lambda \in \mathbb{R}$; on note $y_n = x_{n+1}^\lambda - x_n^\lambda$. Dès lors :

$$y_n = x_n^\lambda \left(\left(1 - \alpha x_n^p + \beta x_n^{2p} + o\left(x_n^{2p}\right)\right)^\lambda - 1\right) = x_n^\lambda \left(-\lambda\alpha x_n^p + \lambda\beta x_n^{2p} + \frac{\lambda(\lambda-1)}{2}\left(\alpha x_n^p\right)^2 + o\left(x_n^{2p}\right)\right)$$
$$= -\lambda\alpha x_n^{\lambda+p} + \left(\lambda\beta + \frac{\lambda(\lambda-1}{2}\alpha^2\right) x_n^{2p} + o\left(x_n^{2p}\right)$$

Cette relation ne nous intéresse vraiment que pour $\lambda = -p$:

$$y_n = p\alpha + \left(-p\beta + \frac{-p(-p-1)}{2}\alpha^2\right) x_n^p + o\left(x_n^p\right) = p\alpha + p\underbrace{\left(\frac{p+1}{2}\alpha^2 - \beta\right)}_{=\gamma} x_n^p + o\left(x_n^p\right) \underset{n\to\infty}{\longrightarrow} p\alpha$$

(Remarquons qu'on a ici utilisé le fait que $x_n \underset{n\to\infty}{\longrightarrow} 0$ et $p > 0$.)
Par téléscopage, et par le lemme de Cesàro (qui sera démontré par la suite) :

$$\frac{1}{n}\left(x_n^{-p} - x_0^{-p}\right) = \frac{1}{n}\sum_{k=0}^{n-1} y_k \underset{n\to\infty}{\longrightarrow} p\alpha$$

Il s'ensuit alors facilement : $\frac{1}{n}x_n^{-p} \underset{n\to\infty}{\longrightarrow} p\alpha$, d'où $x_n \underset{n\to\infty}{\sim} \frac{1}{\sqrt[p]{np\alpha}}$.

Étape 4 : On suppose désormais $\gamma \neq 0$; comme $y_n = p\alpha + p\gamma x_n^p(1 + o(1))$, on a :

$$\sum_{k=0}^{n-1} y_k = np\alpha + p\gamma \sum_{k=0}^{n-1} x_k^p(1 + o(1)).$$

On note $S_n = \dfrac{\sum_{k=0}^{n-1} y_k - np\alpha}{p\gamma}$, de sorte que $S_n = \sum_{k=0}^{n-1} x_k^p(1 + o(1))$.
Mais $x_n^p \underset{n\to\infty}{\sim} \frac{1}{np\alpha} \underset{n\to\infty}{\sim} \frac{1}{(n+1)p\alpha}$, et comme $\sum_{n\geqslant 0} \frac{1}{n+1}$ diverge, on a :

$$S_n \underset{n\to\infty}{\sim} \frac{1}{p\alpha}\sum_{k=0}^{n-1}\frac{1}{k+1} \underset{n\to\infty}{\sim} \frac{1}{p\alpha}\sum_{k=0}^{n-1}\ln\left(1 + \frac{1}{k+1}\right) \underset{n\to\infty}{\sim} \frac{1}{p\alpha}\ln\left(\prod_{k=0}^{n-1}\frac{k+2}{k+1}\right) \underset{n\to\infty}{\sim} \frac{1}{p\alpha}\ln(n+1) \underset{n\to\infty}{\sim} \frac{\ln n}{p\alpha}$$

Diffusion à titre gratuit uniquement.

D'où $\sum_{k=0}^{n-1} y_k = np\alpha + \frac{\gamma}{\alpha}\ln n(1+o(1))$.

Puis $x_n^{-p} = np\alpha + \frac{\gamma}{\alpha}\ln n + o(\ln n)$.

Enfin $x_n = (np\alpha)^{-\frac{1}{p}}\left(1+\frac{\gamma}{\alpha}\frac{\ln n}{n}+o\left(\frac{\ln n}{n}\right)\right)^{-\frac{1}{p}} = \frac{1}{\sqrt[p]{np\alpha}}\left(1-\frac{\gamma}{p^2\alpha^2}\frac{\ln n}{n}+o\left(\frac{\ln n}{n}\right)\right)$.

Ce qui revient à conclure : $x_n = \frac{1}{\sqrt[p]{pn\alpha}} - \frac{\gamma}{p^2\alpha^2\sqrt[p]{p\alpha}}\frac{\ln n}{n\sqrt[p]{n}} + o\left(\frac{\ln n}{n\sqrt[p]{n}}\right)$. ■

Lemme (*Cesàro*)

Soit $(a_n) \in \mathbb{C}^{\mathbb{N}}$, une suite qui converge vers $l \in \mathbb{C}$.
Alors $\lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} a_k = l$.

Démonstration :

Soit $\varepsilon > 0$; $\exists N \in \mathbb{N}^*, \forall n > N, |a_n - l| < \varepsilon$.
Alors, pour $n > N$:

$$\left|\sum_{k=1}^{n} a_k - nl\right| \leqslant \underbrace{\left|\sum_{k=1}^{N} a_k - Nl\right|}_{=:K} + \left|\sum_{k=N+1}^{n} a_k - (n-N)l\right| \leqslant K + \sum_{k=N+1}^{n} |a_k - l| \leqslant K + (n-N)\varepsilon \leqslant K + n\varepsilon$$

Puis $\forall n > N, \left|\frac{1}{n}\sum_{k=1}^{n} a_k - l\right| \leqslant \frac{K}{n} + \varepsilon$.

Or, $\exists N_1 \geqslant N, \forall n \geqslant N_1, \frac{K}{n} < \varepsilon$ et alors $\forall n > N_1, \left|\frac{1}{n}\sum_{k=1}^{n} a_k - l\right| \leqslant 2\varepsilon$. ■

Références

[Rom] J.-E. ROMBALDI – *Éléments d'analyse réelle : CAPES et agrégation de mathématiques*, EDP Sciences, 2004.

Diffusion à titre gratuit uniquement.

Nombre de zéros des solutions d'une équation différentielle

Leçons : 220, 221, 224

[ZQ], théorème X.VI.3

Théorème

Soient $a \in \mathbb{R}$ et $q : [a, +\infty[\to \mathbb{R}^{+*}$ une fonction de classe $\mathcal{C}^1$.
On suppose que $\displaystyle\int_a^{+\infty} \sqrt{q(u)}\,\mathrm{d}u = +\infty$ et que $q'(x) = o_{x\to+\infty}\left(q(x)^{\frac{3}{2}}\right)$. [95]
Soit y une solution réelle non-nulle de $y'' + qy = 0$ sur $[a, +\infty[$ et soit $N(x)$ le nombre de zéros de y sur $[a, x]$.
Alors $N(x) \underset{x\to+\infty}{\sim} \dfrac{1}{\pi}\displaystyle\int_a^x \sqrt{q(u)}\,\mathrm{d}u$.

Démonstration :

Étape 1 : On va faire un "changement de temps" : posons, pour $x \in [a, +\infty[$, $\tau(x) = \displaystyle\int_a^x \sqrt{q(u)}\,\mathrm{d}u$.

Alors τ est $\mathcal{C}^1$ sur $[a, +\infty[$ et $\forall x \in [a, +\infty[$, $\tau'(x) = \sqrt{q(x)} > 0$.
Donc τ est une bijection $\mathcal{C}^1$ croissante de $[a, +\infty[$ sur $\mathbb{R}^+$ et donc τ^{-1} est une bijection $\mathcal{C}^1$ croissante de $\mathbb{R}^+$ sur $[a, +\infty[$. [96]
On pose alors $Y = y \circ \tau^{-1}$, ie $\forall x \in [a, +\infty[, y(x) = Y(\tau(x))$.
En conséquence, pour $x \in [a, +\infty[$:

$$y'(x) = \tau'(x)Y'(\tau(x)) = \sqrt{q(x)}Y'(\tau(x)) \text{ et } y''(x) = \frac{q'(x)}{2\sqrt{q(x)}}Y'(\tau(x)) + q(x)Y''(\tau(x)).$$

L'équation devient alors : $q(x)Y''(\tau(x)) + \dfrac{q'(x)}{2\sqrt{q(x)}}Y'(\tau(x)) + q(x)Y(\tau(x)) = 0$.

En posant $t = \tau(x)$ et $\varphi(t) = \dfrac{q'(x)}{2q(x)^{\frac{3}{2}}}$, on obtient finalement :

$$Y''(t) + \varphi(t)Y'(t) + Y(t) = 0, \text{ où } t \in \mathbb{R}^+.$$

Étape 2 : On rappelle le lemme de relèvement. [97]

Lemme (*Relèvement*)

Soient $y_1, y_2 : [a, +\infty[\to \mathbb{R}$ de classe $\mathcal{C}^1$ et sans zéro commun ; on note $w = y_1y_2' - y_2y_1'$.
Si $y_1(a) + \mathrm{i}y_2(a) = r_0\mathrm{e}^{\mathrm{i}\theta_0}$, alors on peut écrire : $y_1 = r\cos\theta$ et $y_2 = r\sin\theta$, avec $r = \sqrt{y_1^2 + y_2^2}$ et $\theta : x \mapsto \displaystyle\int_a^x \frac{w(t)}{r(t)^2}\,\mathrm{d}t$.

Démonstration :

On pose $\varphi = y_1 + \mathrm{i}y_2$ (par hypothèse, φ ne s'annule jamais) et $\psi : x \mapsto \displaystyle\int_a^x \frac{\varphi'(t)}{\varphi(t)}\,\mathrm{d}t + \ln r_0 + \mathrm{i}\theta_0$.

Alors $\left(\varphi\mathrm{e}^{-\psi}\right)' = \left(\varphi' - \psi'\varphi\right)\mathrm{e}^{-\psi} = \left(\varphi' - \dfrac{\varphi'}{\varphi}\varphi\right)\mathrm{e}^{-\psi} = 0$.

D'où $\forall x \in [a, +\infty[, \varphi(x)\mathrm{e}^{-\psi(x)} = \varphi(a)\mathrm{e}^{-\psi(a)} = r_0\mathrm{e}^{\mathrm{i}\theta_0}\exp\left(-\ln r_0 - \mathrm{i}\theta_0\right) = 1$.

95. Cette hypothèse est indispensable : exhibons un contre-exemple. Prenons $a = 1$ et $q(x) = \frac{1}{4x^2}$. On a $q'(x) = \frac{-1}{2x^3}$ puis $\frac{q'(x)}{q(x)^{\frac{3}{2}}} = \frac{\frac{-1}{2x^3}}{\frac{1}{8x^3}} = -4$. Résolvons $y'' + \frac{1}{4x^2}y = 0$. On commence par chercher $y(x)$ sous la forme x^α ; alors on trouve $\alpha = \frac{1}{2}$. Ainsi, $\sqrt{x}$ est solution de l'équation (on pourrait déjà conclure notre contre-exemple). Par la méthode de Liouville, on va chercher $y(x)$ sous la forme $\sqrt{x}z(x)$; on se ramène alors à $z'' + \frac{1}{x}z' = 0$; et on obtient $y(x) = \sqrt{x}\ln x$.

96. Rappelons que $\left(f \circ f^{-1}\right)(x) = x$ fournit $\left(f^{-1}\right)'(x) = \dfrac{1}{\left(f' \circ f^{-1}\right)(x)}$.

97. Qu'on n'aura jamais le temps de démontrer au tableau, ni même d'énoncer.

Diffusion à titre gratuit uniquement.

Donc $y_1 + \mathrm{i}y_2 = \mathrm{e}^{\psi} = r\mathrm{e}^{\mathrm{i}\theta}$ où l'on convient $r = \sqrt{y_1^2 + y_2^2}$ et $\theta = \operatorname{Im}\psi$.
Mais pour tout $x \geqslant a$,

$$\psi(x) = \ln r_0 + \mathrm{i}\theta_0 + \int_a^x \frac{y_1'(t) + \mathrm{i}y_2'(t)}{y_1(t) + \mathrm{i}y_2(t)}\,\mathrm{d}t = \ln r_0 + \mathrm{i}\theta_0 + \int_a^x \frac{\left(y_1'(t) + \mathrm{i}y_2'(t)\right)\left(y_1(t) + \mathrm{i}y_2(t)\right)}{r(t)^2}\,\mathrm{d}t.$$

Donc $\theta(x) = \theta_0 + \displaystyle\int_a^x \frac{y_1(t)y_2'(t) - y_2(t)y_1'(t)}{r(t)^2}\,\mathrm{d}t = \theta_0 + \int_a^x \frac{w(t)}{r(t)^2}\,\mathrm{d}t.$ ■

Supposons que Y et Y' aient un zéro commun en t_0.
Comme $\varphi(t) \underset{t\to+\infty}{\longrightarrow} 0$ et comme φ est continue sur $\mathbb{R}^+$, elle est bornée.
Par le théorème de Cauchy-Lipschitz, le système différentiel $\begin{pmatrix} Y''(t) \\ Y'(t) \end{pmatrix} = \begin{pmatrix} -\varphi(t) & -1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} Y'(t) \\ Y(t) \end{pmatrix}$
admet une unique solution quand on lui rajoute la condition $\begin{pmatrix} Y'(t_0) \\ Y(t_0) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$.
Ainsi, $Y \equiv 0$, et donc $y \equiv 0$. Mais on avait exclu cette éventualité : contradiction.
Donc Y et Y' n'ont aucun zéro commun.
Par le lemme de relèvement, on obtient $Y = r\sin\theta$ et $Y' = r\cos\theta$, où $r, \theta : \mathbb{R}^+ \to \mathbb{R}$ sont $\mathcal{C}^1$.

Étape 3 : En dérivant ces égalités, on obtient :

$$\left\{\begin{array}{rcccl} Y' &=& r'\sin\theta + r\theta'\cos\theta &=& r\cos\theta \quad (L_1) \\ Y'' &=& r'\cos\theta - r\theta'\sin\theta &=& -\varphi r\cos\theta - r\sin\theta \quad (L_2) \end{array}\right. .$$

Alors $\cos\theta\,(L_1) - \sin\theta\,(L_2)$ fournit : $r\theta' = r + \varphi r\cos\theta\sin\theta$.
Et sachant que r ne s'annule pas, on a : $\theta' = 1 + \varphi\dfrac{\sin 2\theta}{2}$.
Ainsi, $\forall t \in \mathbb{R}^+, |\theta'(t) - 1| = |\varphi(t)|\dfrac{|\sin 2\theta(t)|}{2} \leqslant \dfrac{|\varphi(t)|}{2}$.
Or $\varphi(t) \underset{t\to+\infty}{\longrightarrow} 0$, donc $\theta'(t) \underset{t\to+\infty}{\longrightarrow} 1$.
Comme $\displaystyle\int_0^{+\infty} \mathrm{d}t$ diverge, on a : $\theta(t) - \theta(0) = \int_0^t \theta'(s)\,\mathrm{d}s \underset{t\to+\infty}{\sim} \int_0^t \mathrm{d}s = t$.
En conséquence, on obtient : $\theta(t) \underset{t\to+\infty}{\sim} t$.

Étape 4 : Notons M_a^b le nombre de zéros de Y sur $[a, b]$; on va montrer que $M_0^t \underset{t\to+\infty}{\sim} \dfrac{t}{\pi}$.
Soit $t_0 \in \mathbb{R}^+$, tel que $\forall t \geqslant t_0, \theta'(t) > 0$, alors dès que $t \geqslant t_0$, on a :

$$\begin{aligned} M_{t_0}^t &= \#\{u \in [t_0, t] | \sin\theta(u) = 0\} = \#\{v \in [\theta(t_0), \theta(t)] | \sin v = 0\} = \#\{k \in \mathbb{Z} | \theta(t_0) \leqslant k\pi \leqslant \theta(t)\} \\ &= \left\lfloor \frac{\theta(t)}{\pi} \right\rfloor - \left\lceil \frac{\theta_0}{\pi} \right\rceil + 1 \underset{t\to+\infty}{\sim} \frac{\theta(t)}{\pi} \underset{t\to+\infty}{\sim} \frac{t}{\pi}. \end{aligned}$$

On va montrer que $M_0^{t_0}$ est fini.
Par l'absurde, si on suppose $M_0^{t_0} = \infty$, alors l'ensemble $\{u \in [0, t_0] | Y(u) = 0\}$ possède un point d'accumulation u.
Soit (u_n) une suite dans cet ensemble qui tend vers u.
Comme Y est $\mathcal{C}^1$, on a : $Y'(u) = \lim\limits_{n\to\infty} \dfrac{Y(u_n) - Y(u)}{u_n - u} = 0$.
Mais on a déjà vu que c'est impossible... donc $M_0^{t_0} < \infty$.
En conséquence : $M_0^t \underset{t\to+\infty}{\sim} M_{t_0}^t \underset{t\to+\infty}{\sim} \dfrac{t}{\pi}$.

Étape 5 : Enfin, rattachons cela à la quantité $N(x)$.

$$\begin{aligned} N(x) &= \#\{s \in [a, x] | y(s) = 0\} = \#\{s \in [a, x] | Y(\tau(s)) = 0\} = \#\{t \in [0, \tau(x)] | Y(t) = 0\} \\ &= M_0^{\tau(x)} \underset{x\to+\infty}{\sim} \frac{\tau(x)}{\pi} = \frac{1}{\pi}\int_a^x \sqrt{q(u)}\,\mathrm{d}u. \end{aligned}$$

■

Références

[ZQ] H. QUEFFÉLEC et C. ZUILY – *Analyse pour l'agrégation*, 4e éd., Dunod, 2013.

Diffusion à titre gratuit uniquement.

Processus de Galton-Watson [98]

Leçons : 223, 226 [99], 229, 243, 260, 264, *206, 241, 244, 253*

Merci Laura ! [100]

Soit X une variable aléatoire intégrable à valeurs dans $\mathbb{N}$.
On note, pour $n \in \mathbb{N}$, $p_n = \mathbb{P}(X = n)$ et $m = \mathbb{E}[X] = \sum_{n=0}^{\infty} n p_n < \infty$.
Soit $\left(X_{i,j}\right)_{i,j\in\mathbb{N}}$ une famille de variables aléatoires indépendantes et identiquement distribuées, suivant la loi $\mathbb{P}_X$.
On définit la suite $(Z_n)_{n\in\mathbb{N}}$ de la façon suivante :

$$\begin{cases} Z_0 = 1 \\ \forall n \in \mathbb{N}, Z_{n+1} = \sum_{i=1}^{Z_n} X_{i,n} \end{cases}$$

L'idée est alors de modéliser avec (Z_n) la taille d'une population ; plus précisément, Z_n symbolisera le nombre d'individus à la $n^{\text{ème}}$ génération, et pour $i \in [\![1, Z_n]\!]$, $X_{i,n}$ représentera le nombre de descendants que l'individu de la $n^{\text{ème}}$ génération portant le numéro i aura engendré (les individus de la population qu'on considère génèrent des enfants tous seuls).
On va étudier la suite (Z_n), et particulièrement, on va répondre à la question "Que vaut $\mathbb{P}(\exists n \in \mathbb{N}, Z_n = 0)$? (Quelle est la probabilité que la population considérée s'éteigne ?)" [101]

Lemme

On a : $\forall n \in \mathbb{N}^*, \forall i \in \mathbb{N}, Z_n \perp\!\!\!\perp X_{i,n}$.

Démonstration :
Soit $n \in \mathbb{N}^*$; Z_n ne dépend que de Z_{n-1} et de la famille $(X_{i,n-1})_{i\in\mathbb{N}}$.
Ainsi, par une récurrence immédiate, il vient : Z_n ne dépend que de la famille $\left(X_{i,j}\right)_{i\geqslant 0, j<n}$.
Et, par indépendance des variables $X_{i,j}$, on obtient que $\forall i \in \mathbb{N}, Z_n \perp\!\!\!\perp X_{i,n}$. ■

98. On sautera l'aspect modélisation et le premier lemme. Dans la 229, on ne montre pas que π_∞ est le plus petit point fixe de G, mais on détaille la stricte convexité par la suite. Dans le dernier théorème, on ne construit que les tableaux de variations, et oralement.
99. Attention à bien justifier ce développement dans cette leçon. C'est la suite $(\pi_n)_{n\in\mathbb{N}}$ qui permet de ne pas être de mauvaise foi, évitez l'argument "On regarde la suite $(Z_n, n)_{n\in\mathbb{N}}$" qui est quand même moins élégant ; ou encore pire : "$Z_{n+1} = f_n(Z_n)$" qui est carrément hors-sujet.
100. Un lien vers la page personnelle de Laura GAY.
101. On aurait pu se poser la question "Que vaut $\mathbb{E}[Z_n]$? (Quel est le nombre moyen d'individus à la $n^{\text{ème}}$ génération ?)"

Théorème

On a : $\forall n \in \mathbb{N}, \mathbb{E}[Z_n] = m^n$.

En effet, soit $n \in \mathbb{N}$;

$$\begin{aligned}
\mathbb{E}[Z_{n+1}] &= \mathbb{E}\left[\sum_{i=1}^{Z_n} X_{i,n}\right] = \mathbb{E}\left[\mathbb{E}\left[\sum_{i=1}^{Z_n} X_{i,n}\middle| Z_n\right]\right] = \mathbb{E}\left[\mathbb{E}\left[\sum_{i=1}^{\infty} \mathbb{1}_{i\leqslant Z_n} X_{i,n}\middle| Z_n\right]\right] \\
&= \mathbb{E}\left[\sum_{i=1}^{\infty} \mathbb{1}_{i\leqslant Z_n} \mathbb{E}\left[X_{i,n}\middle| Z_n\right]\right] \text{ (par Fubini-Tonelli, car } \mathbb{1}_{i\leqslant Z_n} X_{i,n} \geqslant 0 \text{ ps et par } Z_n\text{-mesurabilité de } \mathbb{1}_{i\leqslant Z_n}) \\
&= \mathbb{E}\left[\sum_{i=1}^{Z_n} \mathbb{E}\left[X_{i,n}\right]\right] \text{ (d'après le lemme, on a l'indépendance entre } X_{i,n} \text{ et } Z_n) \\
&= \mathbb{E}\left[\sum_{i=1}^{Z_n} m\right] = m\mathbb{E}[Z_n]
\end{aligned}$$

On conclut par récurrence, en utilisant le fait que $Z_0 = 1$.

Diffusion à titre gratuit uniquement.

On note, pour $n \in \mathbb{N}$, $\pi_n = \mathbb{P}(Z_n = 0)$; et $\pi_\infty = \mathbb{P}(\exists n \in \mathbb{N}, Z_n = 0)$ la probabilité d'extinction.
Comme $Z_n = 0 \Rightarrow Z_{n+1} = 0$, la suite d'événements $(\{Z_n = 0\})_{n\in\mathbb{N}}$ est croissante et on a bien :

$$\pi_\infty = \mathbb{P}\left(\bigcup_{n\in\mathbb{N}} \{Z_n = 0\}\right) = \lim_{n\to\infty} \pi_n$$

Si $p_0 = 0$, alors on a $\forall n \in \mathbb{N}^*, Z_n \geqslant 1$ ps et $\pi_\infty = 0$.
Si $p_0 = 1$, alors on a $\forall n \in \mathbb{N}^*, Z_n = 0$ ps et $\pi_\infty = 1$.
On suppose donc désormais $p_0 \in]0,1[$.

Proposition

On définit la série génératrice de X par $G : s \mapsto \mathbb{E}\left[s^X\right] = \sum_{k=0}^{\infty} p_k s^k$.

On a les résultats suivants :

1. G est bien définie sur $[0,1]$ et y est de classe $\mathcal{C}^1$.
2. (a) G est strictement croissante sur $]0,1[$.
   (b) G est convexe sur $]0,1[$.
   (c) G est strictement convexe sur $]0,1[\Leftrightarrow p_0 + p_1 < 1$.

Démonstration :

1. $\forall k \in \mathbb{N}, s \mapsto p_k s^k$ est de classe $\mathcal{C}^1$ sur $[0,1]$, la série $\sum_{k\geqslant 0} p_k 1^k$ converge (vers 1), et la série de fonctions $\sum_{k\geqslant 1} k p_k s^{k-1}$ converge normalement (car X est intégrable) donc uniformément sur $[0,1]$.
   Par conséquent, la série $\sum_{k\geqslant 0} p_k s^k$ converge uniformément vers G, de classe $\mathcal{C}^1$ sur $[0,1]$.
2. La série entière $\sum_{k\geqslant 0} p_k s^k$ ayant un rayon de convergence $\geqslant 1$, on a :

$$\forall s \in [0,1[, G'(s) = \sum_{k=1}^{\infty} k p_k s^{k-1} \text{ et } G''(s) = \sum_{k=2}^{\infty} k(k-1) p_k s^{k-2}$$

   Comme $p_0 < 1$, on a : $\exists k_0 > 0, p_{k_0} > 0$.
   (a) Ainsi : $\forall s \in]0,1[, G'(s) \geqslant k_0 p_{k_0} s^{k_0-1} > 0$ et G est strictement croissante sur $]0,1[$.
   (b) Aussi : $\forall s \in]0,1[, G''(s) \geqslant k_0 (k_0-1) p_{k_0} s^{k_0-2} \geqslant 0$ et G est convexe sur $]0,1[$.
   (c) Si $p_0 + p_1 = 1$, alors on a $k_0 = 1$ et G est affine donc n'est pas strictement convexe sur $]0,1[$.
   Si $p_0 + p_1 < 1$, alors on peut avoir $k_0 > 1$ et $G'' > 0$ sur $]0,1[$ d'où la stricte convexité. ■

Proposition

Pour $n \in \mathbb{N}$, on définit la série génératrice de Z_n par $G_n : s \mapsto \mathbb{E}\left[s^{Z_n}\right] = \sum_{k=0}^{\infty} \mathbb{P}(Z_n = k) s^k$.

Comme précédemment, on peut montrer que G_n est bien définie sur $[0,1]$.
On a : $\forall n \in \mathbb{N}^*, G_n = \underbrace{G \circ \ldots \circ G}_{n \text{ fois}}$ sur $[0,1]$.

Diffusion à titre gratuit uniquement.

Démonstration :
Soit $n \in \mathbb{N}$, $s \in [0,1]$,

$$G_{n+1}(s) = \mathbb{E}\left[s^{Z_{n+1}}\right] = \mathbb{E}\left[s^{\sum_{i=1}^{Z_n} X_{i,n}}\right] = \mathbb{E}\left[\prod_{i=1}^{Z_n} s^{X_{i,n}}\right] = \mathbb{E}\left[\sum_{j=0}^{\infty} \mathbb{1}_{Z_n=j} \prod_{i=1}^{j} s^{X_{i,n}}\right]$$

$$= \sum_{j=0}^{\infty} \mathbb{E}\left[\mathbb{1}_{Z_n=j} \prod_{i=1}^{j} s^{X_{i,n}}\right] \text{(par Fubini-Tonelli, car les termes sont tous positifs)}$$

$$= \sum_{j=0}^{\infty} \mathbb{E}\left[\mathbb{1}_{Z_n=j}\right] \prod_{i=1}^{j} \mathbb{E}\left[s^{X_{i,n}}\right] \text{(car les variables en jeu ici sont toutes indépendantes)}$$

$$= \sum_{j=0}^{\infty} \mathbb{P}(Z_n = j)\, \mathbb{E}\left[s^X\right]^j = \sum_{j=0}^{\infty} \mathbb{P}(Z_n = j)\, G(s)^j = G_n(G(s))$$

On conclut par récurrence, en utilisant le fait que $Z_1 = X_{0,0}$ suit la loi de X. ■

Proposition

La probabilité d'extinction π_∞ est le plus petit point fixe de G sur l'intervalle $[0,1]$.

Démonstration :
La proposition précédente donne : $\forall n \in \mathbb{N}, \forall s \in [0,1], G_{n+1}(s) = G(G_n(s))$.
En évaluant en 0, on obtient la relation : $\pi_{n+1} = G(\pi_n)$, puis par continuité de G sur $[0,1]$, on obtient que π_∞ est un point fixe de G. Reste à montrer que c'est le plus petit.
Soit $u \in [0,1]$ un point fixe de G. On va montrer par récurrence que $\forall n \in \mathbb{N}^*, \pi_n \leqslant u$.
- On a : $\pi_1 = G(\pi_0) = G(\mathbb{P}(Z_0 = 0)) = G(0) \leqslant G(u) = u$, car G est croissante.
- Si $\pi_n \leqslant u$, alors $\pi_{n+1} = G(\pi_n) \leqslant G(u) = u$, toujours par croissance de G.

Par conséquent, $\forall n \in \mathbb{N}, \pi_n \leqslant u$, puis par passage à la limite : $\pi_\infty \leqslant u$. ■

Théorème

Si $m \leqslant 1$, alors $\pi_\infty = 1$.
Si $m > 1$, alors π_∞ est l'unique point fixe de G sur $]0,1[$.

Démonstration :
On rappelle qu'on a deux cas :
- Si $p_0 + p_1 = 1$, et alors G est une fonction affine ; comme $p_0 > 0$, la droite représentative de G coupe en un unique point la droite d'équation $y = x$. Nécessairement, ce point d'intersection a pour coordonnées $(1,1)$.
- Sinon, G est strictement convexe sur $]0,1[$; il en va alors de même pour $x \mapsto G(x) - x$ qui s'annule donc au plus deux fois [102].

Dans tous les cas, on a : $G(x) - x$ s'annule au plus 2 fois sur $[0,1]$; aussi $G'(0) = p_1$ et $G'(1) = \sum_{n=1}^{\infty} n p_n = m$.

Supposons $m > 1$.
Alors $G' - 1$ est une fonction croissante de $p_1 - 1 < 0$ (car $p_0 > 0$) à $m - 1 > 0$, donc elle s'annule en un point $\alpha \in]0,1[$.
La fonction $G - \mathrm{Id}$ est alors décroissante sur $[0,\alpha]$ puis croissante sur $[\alpha,1]$.
Comme $G(0) - 0 = p_0 > 0$ et $G(1) - 1 = 0$, il existe un point dans l'intervalle $]0,\alpha]$ où $G - \mathrm{Id}$ s'annule.
π_∞ est donc l'unique point fixe de G sur l'intervalle $]0,1[$ (car G en a au plus 2).

x	0	π_∞	α		1
$G'(x) - 1$	$p_1 - 1$	$-$	0	$+$	$m - 1$
$G(x) - x$	p_0	$\searrow$	0	$\nearrow$	0

102. Effectivement, si cette fonction s'annule en trois points distincts, par le théorème de Rolle, sa dérivée s'annulera en deux points distincts, $a < b$. Mais $x \mapsto G(x) - x$ est convexe, donc sa dérivée est croissante, donc nulle sur $[a,b]$. Ceci implique que $G' - 1$ n'est pas strictement croissante, et donc que $x \mapsto G(x) - x$ n'est pas strictement convexe. Attention à l'idée reçue qui consiste à penser que "strictement convexe" équivaut à "dérivée seconde strictement positive" (pour les fonctions deux fois dérivables) – pour vous en convaincre, considérez $x \mapsto x^4$.

Diffusion à titre gratuit uniquement.

Supposons $m \leqslant 1$.

Alors $G' - 1$ est une fonction croissante sur $[0,1]$, négative ou nulle en 1 ; donc négative sur $[0,1]$.
Donc $G - \mathrm{Id}$ est décroissante sur $[0,1]$, et s'annule en 1. Comme cette fonction admet au plus 2 annulations, elle ne s'annule qu'en 1 (car sinon elle s'annulerait sur un intervalle non-réduit à un singleton).
Par conséquent, $\pi_\infty = 1$.

x	0		1
$G'(x) - 1$	$p_1 - 1$	$-$	$m - 1$
$G(x) - x$	p_0 ↘		0

■

Diffusion à titre gratuit uniquement.

Théorème Central Limite [103]

Leçons : 218, 249 [104], 261, 262, 263, *223, 245, 260*

[ZQ], section XIII.II.3.d)
[Nou], proposition 1.19.10

Théorème

Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de vaiid L^2, $S_n = \sum_{i=1}^{n} X_i$, $m = \mathbb{E}[X_1]$ et $\sigma^2 = \text{Var}\ (X_1) > 0$.

Alors on a : $\dfrac{S_n - nm}{\sqrt{n\sigma^2}} \xrightarrow[n\to\infty]{\mathcal{L}} \mathcal{N}(0,1)$.

Démonstration :

Quitte à travailler avec $\dfrac{X_i - m}{\sigma}$ en lieu et place de X_i, on peut supposer que $m = 0$ et $\sigma^2 = 1$.

On va donc montrer que $\dfrac{S_n}{\sqrt{n}}$ converge en loi vers une gaussienne centrée réduite.

On va utiliser le théorème de Lévy et montrer que : $\forall t \in \mathbb{R}, \varphi_{\frac{S_n}{\sqrt{n}}}(t) \xrightarrow[n\to\infty]{} \mathrm{e}^{-\frac{t^2}{2}}$.

En effet, on dispose du lemme suivant.

Lemme 1

Si $X \sim \mathcal{N}(0,1)$,
Alors $\forall t \in \mathbb{R}, \varphi_X(t) = \mathrm{e}^{-\frac{t^2}{2}}$.

Démonstration du lemme 1 :

Par le lemme de transfert, $\varphi_X(t) = \dfrac{1}{\sqrt{2\pi}} \displaystyle\int_{\mathbb{R}} \mathrm{e}^{\mathrm{i}tx}\mathrm{e}^{-\frac{x^2}{2}}\,\mathrm{d}x$.

En fait, on va calculer $G(z) = \dfrac{1}{\sqrt{2\pi}} \displaystyle\int_{\mathbb{R}} \underbrace{\mathrm{e}^{zx}\mathrm{e}^{-\frac{x^2}{2}}}_{f(z,x)}\,\mathrm{d}x$.

G est bien définie sur $\mathbb{C}$, on va appliquer le théorème d'holomorphie sous le signe intégrale pour montrer que G est holomorphe.

Plus précisément, on va l'appliquer sur tout disque $\mathcal{D}(0,R)$ où $R > 0$:

- $\forall z \in \mathcal{D}(0,R), x \mapsto f(z,x)$ est intégrable sur $\mathbb{R}$;
- $\forall x \in \mathbb{R}, z \mapsto f(z,x)$ est holomorphe sur $\mathcal{D}(0,R)$;

103. On dispose de l'application suivante.
Soit une pièce qu'on lance n fois ; on symbolise les résultats des lancers par une suite $(X_i)_{i\in\mathbb{N}^*}$ de vaiid de loi $b(p)$. On cherche à estimer le paramètre p.
D'après le TCL, on a : $\dfrac{S_n - np}{\sqrt{np(1-p)}} \overset{b(p)-\mathcal{L}}{\underset{n\to\infty}{\longrightarrow}} \mathcal{N}(0,1)$.
En notant $\overline{X_n} = \dfrac{S_n}{n}$, cela revient à dire $\sqrt{\dfrac{n}{p(1-p)}}\,(\overline{X_n} - p) \overset{b(p)-\mathcal{L}}{\underset{n\to\infty}{\longrightarrow}} \mathcal{N}(0,1)$.
Mais, par la loi faible des grands nombres : $\sqrt{\overline{X_n}\,(1-\overline{X_n})} \overset{b(p)-\mathbb{P}}{\underset{n\to\infty}{\longrightarrow}} \sqrt{p(1-p)}$.
Donc, en utilisant le lemme de Slutsky : $\sqrt{\dfrac{n}{\overline{X_n}\,(1-\overline{X_n})}}\,(\overline{X_n} - p) \overset{b(p)-\mathcal{L}}{\underset{n\to\infty}{\longrightarrow}} \mathcal{N}(0,1)$.
Soit $N \sim \mathcal{N}(0,1)$, q le quantile d'ordre $1 - \dfrac{\alpha}{2}$ de $\mathcal{N}(0,1)$, on a :

$$\mathbb{P}_p\left(-q \leqslant \sqrt{\frac{n}{\overline{X_n}\,(1-\overline{X_n})}}\,(\overline{X_n} - p) \leqslant q\right) \simeq \mathbb{P}(N \leqslant q) - \mathbb{P}(N \leqslant -q) = 2\mathbb{P}(N \leqslant q) - 1 = 1 - \alpha.$$

Donc, pour n assez grand, on dispose de l'intervalle de confiance asymptotique $\left[\overline{X_n} - \dfrac{q}{\sqrt{n}}\sqrt{\overline{X_n}\,(1-\overline{X_n})}, \overline{X_n} + \dfrac{q}{\sqrt{n}}\sqrt{\overline{X_n}\,(1-\overline{X_n})}\right]$ qui contient p avec une probabilité proche de $1 - \alpha$.

104. Pour la leçon 249, on peut ajouter une comparaison de calculs d'intervalles de confiance asymptotiques, entre l'inégalité de Tchebychev et le théorème central limite ; pour cela, on peut prendre exemple sur le document de Laura GAY.

Diffusion à titre gratuit uniquement.

– $\forall x \in \mathbb{R}, \forall z \in \mathcal{D}(0,R), |f(z,x)| = \left|e^{zx}e^{-\frac{x^2}{2}}\right| = e^{x\mathrm{Re}z}e^{-\frac{x^2}{2}} \leqslant e^{Rx}e^{-\frac{x^2}{2}}$, majorant qui a eu la bonne idée d'être à la fois intégrable et indépendant de z.

Donc $\forall R > 0$, G est holomorphe sur $\mathcal{D}(0,R)$ et donc sur $\mathbb{C}$.

Soit $u \in \mathbb{R}$, $G(u) = \dfrac{1}{\sqrt{2\pi}}\displaystyle\int_{\mathbb{R}} e^{ux-\frac{x^2}{2}}\,dx = \dfrac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} e^{\frac{u^2}{2}}e^{-\frac{(x-u)^2}{2}}\,dx = e^{\frac{u^2}{2}}$.

G et $z \mapsto e^{\frac{z^2}{2}}$ sont des fonctions holomorphes qui coïncident sur $\mathbb{R}$, donc, par prolongement analytique, elles sont égales sur $\mathbb{C}$.

En particulier, $\forall t \in \mathbb{R}, \varphi_X(t) = G(it) = e^{-\frac{t^2}{2}}$. ■

Comme $X_1 \in L^2$, φ_{X_1} est de classe $\mathcal{C}^2$; de plus :

$$\varphi'_{X_1}(0) = \mathbb{E}[iX_1] = 0 \text{ et } \varphi''_{X_1}(0) = \mathbb{E}\left[-X_1^2\right] = -\mathrm{Var}\ (X_1) - \mathbb{E}[X_1]^2 = -1.$$

D'autre part, $\forall t \in \mathbb{R}, \varphi_{\frac{S_n}{\sqrt{n}}}(t) = \varphi_{S_n}\left(\dfrac{t}{\sqrt{n}}\right) = \left(\varphi_{X_1}\left(\dfrac{t}{\sqrt{n}}\right)\right)^n$, car les variables X_i sont iid.

φ_{X_1} étant de classe $\mathcal{C}^2$, on a le développement limité :

$$\varphi_{X_1}\left(\frac{t}{\sqrt{n}}\right) = \varphi_{X_1}(0) + \frac{t}{\sqrt{n}}\varphi'_{X_1}(0) + \frac{t^2}{2n}\varphi''_{X_1}(0) + \frac{\varepsilon_n}{n} = 1 - \frac{t^2}{2n} + \frac{\varepsilon_n}{n}, \text{ où } \varepsilon_n \underset{n\to\infty}{\longrightarrow} 0.$$

On en déduit alors que $\varphi_{\frac{S_n}{\sqrt{n}}}(t) = \left(1 - \dfrac{t^2}{2n} + \dfrac{\varepsilon_n}{n}\right)^n$.

Lemme 2

Si $(z_n)_{n\in\mathbb{N}^*}$ est une suite de nombres complexes qui tend vers $z \in \mathbb{C}$, Alors $\lim\limits_{n\to\infty}\left(1 + \dfrac{z_n}{n}\right)^n = e^z$.

Démonstration du lemme 2 :

On a : $e^{z_n} - \left(1 + \dfrac{z_n}{n}\right)^n = \displaystyle\sum_{k=0}^{\infty} a_{k,n}z_n^k$ avec $a_{k,n} = \begin{cases} \frac{1}{k!}\left(1 - \frac{n(n-1)\ldots(n-k+1)}{n^k}\right) & \text{si } k \leqslant n \\ \frac{1}{k!} & \text{sinon} \end{cases}$.

Les $a_{k,n}$ sont donc tous positifs, on en déduit :

$$\begin{aligned}
\left|e^{z_n} - \left(1 + \frac{z_n}{n}\right)^n\right| &\leqslant \sum_{k=0}^{\infty} a_{k,n}|z_n|^k \\
&= e^{|z_n|} - \left(1 + \frac{|z_n|}{n}\right)^n \\
&= e^{|z_n|} - \exp\left(n\ln\left(1 + \frac{|z_n|}{n}\right)\right) \\
&\leqslant e^{|z_n|} - \exp\left(n\left(\frac{|z_n|}{n} - \frac{|z_n|^2}{2n^2}\right)\right) \text{ car } \ln(1+x) \geqslant x - \frac{x^2}{2} \text{ (pour } x > 0) \\
&\leqslant e^{|z_n|}\left(1 - \exp\left(-\frac{|z_n|^2}{2n}\right)\right) \\
&\leqslant e^{|z_n|}\frac{|z_n|^2}{2n} \text{ car } 1 - e^x \leqslant x \text{ (pour } x \in \mathbb{R})
\end{aligned}$$

Donc $\left|e^z - \left(1 + \dfrac{z_n}{n}\right)^n\right| \leqslant |e^z - e^{z_n}| + e^{|z_n|}\dfrac{|z_n|^2}{2n} \underset{n\to\infty}{\longrightarrow} 0$. ■

On applique le lemme 2 avec $z_n = -\dfrac{t^2}{2} + \varepsilon_n$, d'où $\varphi_{\frac{S_n}{\sqrt{n}}}(t) \underset{n\to\infty}{\longrightarrow} e^{-\frac{t^2}{2}}$. ■

Références

[ZQ] H. QUEFFÉLEC et C. ZUILY – *Analyse pour l'agrégation*, 4e éd., Dunod, 2013.
[Nou] I. NOURDIN – *Agrégation de mathématique - Épreuve orale*, 2e éd., Dunod, 2006.

Diffusion à titre gratuit uniquement.

Théorème de Bernstein (sur les séries entières)

Leçons : 244, *218, 224, 241, 243*

[Gou An], exercice 4.4.8

Théorème

Soit $a > 0$, $f :]-a, a[\to \mathbb{R}$ une fonction de classe $\mathcal{C}^\infty$.
On suppose que $\forall k \in \mathbb{N}, \forall x \in]-a, a[, f^{(2k)}(x) \geqslant 0$.
Alors f est développable en série entière sur $]-a, a[$, ie analytique sur $]-a, a[$.

Démonstration :

Soit $b \in]0, a[$, on va d'abord montrer que f est développable en série entière en 0 avec un rayon de convergence supérieur ou égal à b.

Étape 1 : On va d'abord montrer un résultat similaire pour une fonction auxiliaire.

Soit $F : \left| \begin{array}{ccc}]-a, a[& \to & \mathbb{R} \\ x & \mapsto & f(x) + f(-x) \end{array} \right.$.

F est paire et donc : $\forall k \in \mathbb{N}, F^{(2k+1)}$ est impaire, d'où $F^{(2k+1)}(0) = 0$.

Par la formule de Taylor avec reste intégral, on obtient : $\forall n \in \mathbb{N}, \forall x \in [0, b]$,

$$F(x) = F(0) + \frac{x^2}{2!}F''(0) + \ldots + \frac{x^{2n}}{(2n)!}F^{(2n)}(0) + R_n(x) \text{ où } R_n(x) = \int_0^x \frac{(x-t)^{2n+1}}{(2n+1)!}F^{(2n+2)}(t)\,\mathrm{d}t$$

Or $\forall k \in \mathbb{N}, F^{(2k)}(0) = f^{(2k)}(0) + (-1)^{2k}f^{(2k)}(0) = 2f^{(2k)}(0) \geqslant 0$.
Donc $\forall n \in \mathbb{N}, \forall x \in [0, b], R_n(x) \leqslant F(x)$.
On a finalement :

$$0 \leqslant R_n(x) = \int_0^x \left(\frac{x-t}{b-t}\right)^{2n+1} \frac{(b-t)^{2n+1}}{(2n+1)!}F^{(2n+2)}(t)\,\mathrm{d}t \underset{105}{\leqslant} \left(\frac{x}{b}\right)^{2n+1} \int_0^x \frac{(b-t)^{2n+1}}{(2n+1)!}F^{(2n+2)}(t)\,\mathrm{d}t$$
$$\leqslant \left(\frac{x}{b}\right)^{2n+1} R_n(b) \leqslant \left(\frac{x}{b}\right)^{2n+1} F(b)$$

Donc, pour $x \in [0, b[$, $\dfrac{x}{b} \in [0, 1[$ et $\lim\limits_{n\to\infty} R_n(x) = 0$.

Dès lors : $\forall x \in [0, b[, F(x) = \sum\limits_{n=0}^{\infty} \dfrac{x^{2n}}{(2n)!}F^{(2n)}(0)$, puis, par parité : $\forall x \in]-b, b[, F(x) = \sum\limits_{n=0}^{\infty} \dfrac{x^{2n}}{(2n)!}F^{(2n)}(0)$.

Étape 2 : Montrons désormais le résultat pour f.

Par la formule de Taylor avec reste intégral, on a : $\forall x \in]-b, b[, \forall n \in \mathbb{N}$,

$$f(x) = f(0) + xf'(0) + \ldots + \frac{x^{2n+1}}{(2n+1)!}f^{(2n+1)}(0) + r_n(x) \text{ où } r_n(x) = \int_0^x \frac{(x-t)^{2n+1}}{(2n+1)!}f^{(2n+2)}(t)\,\mathrm{d}t$$

Or $\forall t \in [0, x], F^{(2n+2)}(t) = f^{(2n+2)}(t) + (-1)^{2n+2}f^{(2n+2)}(-t) \geqslant f^{(2n+2)}(t) \geqslant 0$.
Si $x \geqslant 0$, alors :

$$|r_n(x)| \leqslant \int_0^x \frac{|x-t|^{2n+1}}{(2n+1)!}\left|f^{(2n+2)}(t)\right|\,\mathrm{d}t \leqslant \int_0^x \frac{(x-t)^{2n+1}}{(2n+1)!}F^{(2n+2)}(t)\,\mathrm{d}t = R_n(x)$$

Et si $x < 0$, alors :

$$|r_n(x)| \leqslant \int_x^0 \frac{|x-t|^{2n+1}}{(2n+1)!}\left|f^{(2n+2)}(t)\right|\,\mathrm{d}t = \int_x^0 \frac{(t-x)^{2n+1}}{(2n+1)!}F^{(2n+2)}(t)\,\mathrm{d}t$$
$$\leqslant \int_{-x}^0 \frac{(-u-x)^{2n+1}}{(2n+1)!}F^{(2n+2)}(-u)\,(-\mathrm{d}u) = \int_0^{-x} \frac{(-x-u)^{2n+1}}{(2n+1)!}\underbrace{F^{(2n+2)}}_{\text{paire}}(u)\,\mathrm{d}u$$
$$\leqslant R_n(-x)$$

105. On a $\dfrac{x-t}{b-t} \leqslant \dfrac{x}{b}$ car $\dfrac{x-t}{b-t} = 1 + \dfrac{x-b}{b-t}$ est décroissante en $t \in [0, x]$.

Diffusion à titre gratuit uniquement.

Donc $\forall x \in]-b,b[, \lim\limits_{n\to\infty} r_n(x) = 0$.

Pour $p \in \mathbb{N}$, on définit $S_p : x \mapsto \sum\limits_{k=0}^{p} \frac{x^k}{k!} f^{(k)}(0)$; on a montré : $\forall x \in]-b,b[, \lim\limits_{n\to\infty} S_{2n+1}(x) = f(x)$.

Or $S_{2n}(x) - S_{2n-1}(x) = \frac{x^{2n}}{(2n)!} f^{(2n)}(0) = \frac{1}{2}\frac{x^{2n}}{(2n)!} F^{(2n)}(0) \underset{n\to\infty}{\longrightarrow} 0$ car F est développable en série entière en 0 avec un rayon supérieur ou égal à b.
Dès lors : $\lim\limits_{n\to\infty} S_{2n}(x) = f(x)$ pour $x \in]-b,b[$.
Et donc f est développable en série entière en 0 avec un rayon de convergence supérieur ou égal à b.

On a donc montré pour l'instant que f est développable en série entière en 0 avec un rayon de convergence égal à a.

Soit alors $x_0 \in]-a,a[$, on pose alors : $g : \left| \begin{array}{ccc}]|x_0|-a, a-|x_0|[& \to & \mathbb{R} \\ x & \mapsto & f(x_0+x) \end{array} \right.$.
Alors g est de classe $\mathcal{C}^\infty$ sur $]|x_0|-a, a-|x_0|[$, avec $a-|x_0| > 0$ et $\forall k \in \mathbb{N}, g^{(2k)}(x) = f^{(2k)}(x_0+x) \geqslant 0$.
Ce qu'on vient de faire permet de montrer que g est développable en série entière en 0 avec un rayon de convergence égal à $a-|x_0|$.
On en déduit alors que f est analytique sur $]-a,a[$. [106] ■

Références

[Gou An] X. GOURDON – *Les maths en tête : Analyse*, 2e éd., Ellipses, 2008.

106. On dispose d'une application : tan est développable en série entière sur $\left]-\frac{\pi}{2}, \frac{\pi}{2}\right[$. En effet, tan est $\mathcal{C}^\infty$ sur cet intervalle, $\tan' = 1 + \tan^2$ et par Leibniz, on obtient : $\tan^{(n+1)} = \sum\limits_{k=0}^{n} \binom{k}{n} \tan^{(k)} \tan^{(n-k)}$, pour $n \in \mathbb{N}^*$. Par récurrence, on montre alors que pour tout $n \in \mathbb{N}$, $\tan^{(n)}$ est positive sur $\left[0, \frac{\pi}{2}\right[$. Or, tan est impaire, donc les dérivées d'ordre impair de tan sont paires ; on en déduit qu'elles sont positives sur $\left]-\frac{\pi}{2}, \frac{\pi}{2}\right[$. Par le théorème de Bernstein, tan′ est développable en série entière sur $\left]-\frac{\pi}{2}, \frac{\pi}{2}\right[$, et donc tan l'est aussi.

Diffusion à titre gratuit uniquement.

Théorème de Cauchy-Lipschitz [107]

Leçons : 203, 206, 220, 221, *205, 208*

[Rou], exercice 60

Théorème

Soient $\|.\|$ une norme sur $\mathbb{R}^m$, I un intervalle de $\mathbb{R}$ et $f : I \times \mathbb{R}^m \to \mathbb{R}^m$ une application continue et globalement lipschitzienne en la 2e variable au sens suivant :

$$\forall K \subset I \text{ intervalle compact}, \exists k > 0, \forall t \in K, \forall y, z \in \mathbb{R}^m, \|f(t,y) - f(t,z)\| \leqslant k\|y - z\|. \text{ [108]}$$

Alors, si $t_0 \in I$ et $x \in \mathbb{R}^m$ sont donnés, le problème de Cauchy (P) : $\begin{cases} y'(t) = f(t, y(t)) \text{ [109]} \\ y(t_0) = x \end{cases}$ admet une unique solution définie sur I tout entier.

Démonstration :

→ **Cas 1 :** Supposons que I soit compact.

1. On va se ramener à un problème de point fixe.
   Dire que y est solution de (P) signifie que y est dérivable sur I, et même $\mathcal{C}^1$, vu que f est continue.
   Ainsi, on a : $\forall t \in I, y(t) = x + \int_{t_0}^{t} f(s, y(s))\,\mathrm{d}s$. [110]
   Réciproquement, si y est continue et vérifie cette égalité, alors y est $\mathcal{C}^1$ et c'est une solution de (P).
   Ainsi, y est une solution de $(P) \Leftrightarrow y = F(y)$ où $F : \left| \begin{array}{ccc} \mathcal{C}(I, \mathbb{R}^m) & \to & \mathcal{C}(I, \mathbb{R}^m) \\ y & \mapsto & \left(t \mapsto x + \int_{t_0}^{t} f(s, y(s))\,\mathrm{d}s\right) \end{array} \right.$.
2. Montrons que F possède un unique point fixe.
   Soit k la constante de Lipschitz de f associée à l'intervalle compact I [111] et l la longueur de I.
   On munit $\mathcal{C}(I, \mathbb{R}^m)$ de la norme $N_k(y) = \max_{t \in I} \left(\mathrm{e}^{-k|t-t_0|}\|y(t)\|\right)$.
   On vérifie facilement que c'est une norme sur $\mathcal{C}(I, \mathbb{R}^m)$, car I est compact, $\|.\|$ est une norme, et exp est à valeurs strictement positives.
   De plus, $\forall y \in \mathcal{C}(I, \mathbb{R}^m)$, $\mathrm{e}^{-kl}\|y\|_\infty \leqslant N_k(y) \leqslant \|y\|_\infty$; ainsi N_k et $\|.\|_\infty$ sont équivalentes.
   Comme $\mathcal{C}(I, \mathbb{R}^m)$ est complet pour $\|.\|_\infty$, il l'est aussi pour N_k.
   Et comme $\mathcal{C}(I, \mathbb{R}^m)$ est stable par F, pour pouvoir appliquer le théorème de Picard, on veut montrer que F est contractante.
   Soient $y, z \in \mathcal{C}(I, \mathbb{R}^m)$, $t \in I$ avec $t \geqslant t_0$:

$$F(y)(t) - F(z)(t) = \int_{t_0}^{t} (f(s, y(s)) - f(s, z(s)))\,\mathrm{d}s.$$

   En conséquence, on obtient la suite d'inégalités :

$$\begin{aligned} \mathrm{e}^{-k(t-t_0)}\|F(y)(t) - F(z)(t)\| &\leqslant \mathrm{e}^{-k(t-t_0)} \int_{t_0}^{t} \|f(s, y(s)) - f(s, z(s))\|\,\mathrm{d}s \\ &\leqslant \mathrm{e}^{-k(t-t_0)} \int_{t_0}^{t} k\|y(s) - z(s)\|\,\mathrm{d}s \\ &\leqslant \mathrm{e}^{-k(t-t_0)} N_k(y - z) \int_{t_0}^{t} k\mathrm{e}^{k(s-t_0)}\,\mathrm{d}s = \mathrm{e}^{-k(t-t_0)} N_k(y - z)\left(\mathrm{e}^{k(t-t_0)} - 1\right) \\ &\leqslant N_k(y - z)\left(1 - \mathrm{e}^{-k(t-t_0)}\right) \end{aligned}$$

107. S'il reste du temps, on peut utiliser le théorème pour montrer l'existence d'une unique solution à l'équation du pendule.
108. Dans le cas linéaire, on oublie f et on remplace par $A \in \mathcal{C}(I, \mathcal{M}_m(\mathbb{R}))$ et $b \in \mathcal{C}(I, \mathbb{R}^m)$.
109. Ou $y'(t) = A(t)y(t) + b(t)$.
110. A et b sont continues, aucun problème dans le cas linéaire. Ici, et dans la suite, on remplacera $f(s, y(s))$ par l'expression $A(s)y(s) + b(s)$.
111. Dans le cas linéaire, on pose $k = \max_{t \in I} |||A(t)|||$, où $|||.|||$ est la norme subordonnée à $\|.\|$.

Diffusion à titre gratuit uniquement.

Similairement, on aurait pu montrer, dans le cas où $t \leqslant t_0$, que :

$$\mathrm{e}^{-k(t_0-t)}\|F(y)(t)-F(z)(t)\| \leqslant N_k(y-z)\left(1-\mathrm{e}^{-k(t_0-t)}\right).$$

En définitive, on a :

$$\forall t \in I, \mathrm{e}^{-k|t-t_0|}\|F(y)(t)-F(z)(t)\| \leqslant N_k(y-z)\left(1-\mathrm{e}^{-k|t-t_0|}\right).$$

On prend alors le maximum sur I, et on obtient :

$$N_k(F(y)-F(z)) \leqslant \underbrace{\left(1-\mathrm{e}^{-kl}\right)}_{<1} N_k(y-z).$$

F est donc contractante sur $\left(\mathcal{C}\left(I,\mathbb{R}^m\right), N_k\right)$ donc admet un unique point fixe.

→ **Cas 2 :** Reste donc à traiter le cas général où I est intervalle quelconque.

On peut écrire l'intervalle I comme union croissante d'intervalles compacts : $I = \bigcup_{j\in\mathbb{N}} I_j$, avec la contrainte $t_0 \in I_j$, pour tout $j \in \mathbb{N}$.

Par ce qui précède, on peut définir y_j, la solution de (P) sur I_j.

– Soit alors y une solution de (P) sur I ; par unicité sur I_j, on a donc : $y_{|I_j} \equiv y_j$.
On obtient ainsi l'unicité de y sur I.

– Réciproquement, les y_j se raccordent, par unicité sur I_j, et donc $y : t \mapsto y_j(t)$ si $t \in I_j$ est bien définie.

Le problème sur un intervalle I quelconque admet donc une unique solution définie sur I tout entier. ■

Références

[Rou] F. ROUVIÈRE – *Petit guide de calcul différentiel*, 4e éd., Cassini, 2014.

Diffusion à titre gratuit uniquement.

Théorème de réarrangement de Riemann

Leçons : 202 [112]**, 230,** ***223***

[X-ENS An1], exercice 3.48

Théorème

Soit $\sum_{n\geqslant 0} a_n$ une série réelle semi-convergente et $\alpha \in \mathbb{R}$.

Alors il existe $\sigma \in \mathfrak{S}(\mathbb{N})$ telle que $\sum_{n=0}^{\infty} a_{\sigma(n)} = \alpha$.

Démonstration :

Étape 1 : Partitionnons $\mathbb{N}$ en deux ensembles infinis ; on note $E^+ = \{n \in \mathbb{N} | a_n \geqslant 0\}$ et $E^- = \{n \in \mathbb{N} | a_n < 0\}$.
On a donc : $\mathbb{N} = E^+ \sqcup E^-$.
Supposons par l'absurde $\#E^- < \infty$.
Alors la suite $(a_n)_{n\in\mathbb{N}}$ est à termes positifs à partir d'un certain rang ; ainsi la convergence de la série $\sum_{n\geqslant 0} a_n$ équivaut à sa convergence absolue.
Et ceci contredit alors la semi-convergence de $\sum_{n\geqslant 0} a_n$.
Ainsi $\#E^- = \infty$ et, similairement, $\#E^+ = \infty$.

Étape 2 : Pour $n \in \mathbb{N}$, on a : $\max\{0, a_n\} = \dfrac{a_n + |a_n|}{2}$ et $\min\{0, a_n\} = \dfrac{a_n - |a_n|}{2}$.
En conséquence, $\sum_{n\geqslant 0} \max\{0, a_n\}$ et $\sum_{n\geqslant 0} \min\{0, a_n\}$ divergent.

Étape 3 : On va construire $\sigma(n)$ par récurrence sur $n \in \mathbb{N}$.
On pose $\sigma(0) = 0$ et pour $n \in \mathbb{N}^*$, deux cas se présentent :
– Si $\sum_{k=0}^{n-1} a_{\sigma(k)} \leqslant \alpha$, alors on va ajouter un terme positif, et $\sigma(n) = \min E^+ \backslash (E^+ \cap \{\sigma(k) | k \in [\![0, n-1]\!]\})$.
– Si $\sum_{k=0}^{n-1} a_{\sigma(k)} > \alpha$, alors on va ajouter un terme négatif, et $\sigma(n) = \min E^- \backslash (E^- \cap \{\sigma(k) | k \in [\![0, n-1]\!]\})$.

Étape 4 : La construction de σ implique son injectivité ; montrons sa surjectivité.
Par l'absurde, soit $N \in \mathbb{N}$, avec $N \notin \sigma(\mathbb{N})$; imaginons que $N \in E^+$.
Alors $E^+ \cap \sigma(\mathbb{N})$ est fini donc $\exists n_0 \in \mathbb{N}, \forall n \geqslant n_0, \sigma(n) \in E^-$.
En conséquence, dès que $n \geqslant n_0$, $\sum_{k=0}^{n-1} a_{\sigma(k)} > \alpha$.
Comme la série $\sum_{n\geqslant 0} a_{\sigma(n)}$ est à termes négatifs à partir d'un certain rang et que ses sommes partielles sont minorées, elle converge.
Soit $\varphi : \mathbb{N} \to E^-$ une application bijective strictement croissante [113] ; les séries $\sum_{n\geqslant 0} a_{\sigma(n)}$ et $\sum_{n\geqslant 0} a_{\varphi(n)}$ diffèrent d'un nombre fini de termes donc sont de même nature.
Mais les sommes partielles de $\sum_{n\geqslant 0} a_{\varphi(n)}$ sont majorées par celles de $\sum_{n\geqslant 0} \min\{0, a_n\}$ qui divergent vers $-\infty$.
La série $\sum_{n\geqslant 0} a_{\varphi(n)}$ est donc divergente : contradiction.
On opérerait similairement si on avait $N \in E^-$; on a ainsi montré que σ est bijective.

Étape 5 : Il reste à montrer que $\sum_{n\geqslant 0} a_{\sigma(n)}$ converge et que sa somme vaut α.
Soit $\varepsilon > 0$; comme $\sum_{n\geqslant 0} a_n$ converge, on a : $a_n \underset{n\to\infty}{\longrightarrow} 0$, d'où $\exists n_1 \in \mathbb{N}, \forall n \geqslant n_1, |a_n| < \varepsilon$.

112. Dans la 202, on énoncera la conclusion du théorème comme suit "l'ensemble $\left\{\sum n = 0^N a_{\sigma(n)} \middle| N \in \mathbb{N}, \sigma \in \mathfrak{S}(\mathbb{N})\right\}$ est dense dans $\mathbb{R}$".

113. Pour la construire, il suffit de poser $\varphi(0) = \min E^-$ et pour $n \in \mathbb{N}^*$, $\varphi(n) = \min\left(E^- \backslash \{\varphi(k) | k \in [\![0, n-1]\!]\}\right)$.

Diffusion à titre gratuit uniquement.

Par injectivité, l'ensemble $\{n \in \mathbb{N} | \sigma(n) \leqslant n_1\}$ est fini ; on pose $n_0 = \max\{k \in \mathbb{N} | \sigma(k) \leqslant n_1\}$.
Si $n > n_0$, alors $\sigma(n) > n_1$ donc $\left|a_{\sigma(n)}\right| < \varepsilon$.
On sait que $\exists N \geqslant n_0, \sigma(N) \in E^+$ et $\sigma(N+1) \in E^-$, car les $\sigma(n)$ ne peuvent pas rester dans E^+ ou dans E^-.
On pose alors, pour $n \in \mathbb{N}$, $S_n = \sum_{k=0}^{n} a_{\sigma(k)}$; on va montrer que $\forall n \geqslant N, |S_n - \alpha| \leqslant \varepsilon$.
Comme $\sigma(N) \in E^+$, on a : $S_{N-1} \leqslant \alpha$; et comme $\sigma(N+1) \in E^-$, aussi $S_N > \alpha$.
Or $|S_N - S_{N-1}| = \left|a_{\sigma(N)}\right| < \varepsilon$ donc $\alpha + \varepsilon \geqslant S_{N-1} + \varepsilon \geqslant S_N > \alpha$.
Soit $n > N$, imaginons que $S_n > \alpha + \varepsilon$.
Comme on passe de S_{n-1} à S_n par un saut de longueur $\left|a_{\sigma(n)}\right| < \varepsilon$, on en déduit que $S_{n-1} > \alpha$.
Cela entraîne que $a_{\sigma(n)} < 0$ et donc $S_{n-1} \leqslant S_n$.
Mais alors de proche en proche : $\alpha + \varepsilon < S_n \leqslant S_{n-1} \leqslant S_{n-2} \leqslant \ldots \leqslant S_N$. D'où une contradiction.
Ainsi, $\forall n > N, S_n \leqslant \alpha + \varepsilon$.
Similairement, on montre que $\forall n > N, S_n \geqslant \alpha - \varepsilon$. Ainsi, $\forall n > N, |S_n - \alpha| \leqslant \varepsilon$.
On a donc construit $\sigma \in \mathfrak{S}(\mathbb{N})$ telle que $\sum_{n \geqslant 0} a_{\sigma(n)}$ converge et soit de somme α. ■

Références

[X-ENS An1] S. FRANCINOU, H. GIANELLA et S. NICOLAS – *Oraux X-ENS Analyse 1*, 3e éd., Cassini, 2014.

Diffusion à titre gratuit uniquement.

Théorème de Riesz-Fischer

Leçons : 201, 205, 208, 234

[Bré], théorème IV.8
[Rud], théorème 3.11

Théorème

Soit X un espace mesuré, muni d'une mesure positive μ.
Pour tout $p \in [1, +\infty]$, $\mathrm{L}^p(\mu)$ est un espace de Banach. [114]

Démonstration :

Étape 1 : Montrons que $\mathrm{L}^\infty(\mu)$ est complet.

Soit $(f_n)_{n\in\mathbb{N}}$ une suite de Cauchy dans $\mathrm{L}^\infty(\mu)$.

Ainsi, $\forall k \in \mathbb{N}^*, \exists N_k \in \mathbb{N}, \forall m, n \geqslant N_k, \|f_n - f_m\|_\infty \leqslant \dfrac{1}{k}$.

Donc il existe une famille $(E_k)_{k\in\mathbb{N}^*}$ d'ensembles μ-négligeables, vérifiant :

$$\forall k \in \mathbb{N}^*, \exists N_k \in \mathbb{N}, \forall m, n \geqslant N_k, \forall x \in X\backslash E_k, |f_n(x) - f_m(x)| \leqslant \frac{1}{k}$$

On définit $E = \bigcup_{k\in\mathbb{N}^*} E_k$; et comme l'union est dénombrable : $\mu(E) = 0$.

Et donc, on obtient :

$$\forall k \in \mathbb{N}^*, \exists N_k \in \mathbb{N}, \forall m, n \geqslant N_k, \forall x \in X\backslash E, |f_n(x) - f_m(x)| \leqslant \frac{1}{k} \tag{4}$$

Puis $\forall x \in X\backslash E, \forall k \in \mathbb{N}^*, \exists N_k \in \mathbb{N}, \forall m, n \geqslant N_k, |f_n(x) - f_m(x)| \leqslant \dfrac{1}{k}$,
autrement dit : $\forall x \in X\backslash E$, $(f_n(x))_{n\in\mathbb{N}}$ est une suite de Cauchy dans $\mathbb{R}$.
Et comme $\mathbb{R}$ est complet, cette suite converge ; on note $f(x)$ sa limite.
On va montrer que la fonction f ainsi définie μ-presque partout est bien dans $\mathrm{L}^\infty(\mu)$ et que c'est bien la limite de la suite $(f_n)_{n\in\mathbb{N}}$ en norme $\|.\|_\infty$.
Par passage à la limite dans (4), on obtient :

$$\forall k \in \mathbb{N}^*, \exists N_k \in \mathbb{N}, \forall n \geqslant N_k, \forall x \in X\backslash E, |f_n(x) - f(x)| \leqslant \frac{1}{k}$$

Ainsi, $\exists N_1 \in \mathbb{N}, \|f\|_\infty \leqslant 1 + \left\|f_{N_1}\right\|_\infty < \infty$, donc $f \in \mathrm{L}^\infty(\mu)$.

Enfin, $\forall k \in \mathbb{N}^*, \exists N_k \in \mathbb{N}, \forall n \geqslant N_k, \|f_n - f\|_\infty \leqslant \dfrac{1}{k}$.

Donc : $\lim\limits_{n\to\infty} \|f_n - f\|_\infty = 0$.

Étape 2 : Pour $p \in [1, +\infty[$, $\mathrm{L}^p(\mu)$ est également complet.

Soit $(f_n)_{n\in\mathbb{N}}$ une suite de Cauchy dans $\mathrm{L}^p(\mu)$.

On va en fait montrer qu'une sous-suite converge dans $\mathrm{L}^p(\mu)$; en effet, prenons $\varepsilon > 0$.

Si d'une part : $\exists K_0 \in \mathbb{N}, \forall k \geqslant K_0, \left\|f_{n_k} - f\right\|_p \leqslant \dfrac{\varepsilon}{2}$,

et d'autre part : $\exists N_0 \in \mathbb{N}, \forall m, n \geqslant N_0, \|f_n - f_m\|_p \leqslant \dfrac{\varepsilon}{2}$,

Alors finalement, en posant $N = \max\left\{n_{K_0}, N_0\right\}$ on obtient : $\forall n \geqslant N, \|f - f_n\|_p \leqslant \varepsilon$, et le théorème est prouvé.

Soit donc $(n_k)_{k\in\mathbb{N}}$ une suite strictement croissante d'entiers telle que :

$$\forall k \in \mathbb{N}, \forall m, n \geqslant n_k, \|f_n - f_m\|_p \leqslant \frac{1}{2^k}$$

114. Une application de la complétude des espaces L^p est le prolongement de la transformée de Fourier à L^2, qui utilise le fait que toute suite de Cauchy dans L^2 converge dans L^2 (voir le développement consacré en page 103).

Diffusion à titre gratuit uniquement.

Ainsi, $\forall k \in \mathbb{N}, \left\|f_{n_{k+1}} - f_{n_k}\right\|_p \leqslant \dfrac{1}{2^k}$.
Notons désormais $\hat{f}_k$ pour f_{n_k}, ainsi $\forall k \in \mathbb{N}, \left\|\hat{f}_{k+1} - \hat{f}_k\right\|_p \leqslant \dfrac{1}{2^k}$.
Posons alors, pour $n \in \mathbb{N}$: $g_n = \displaystyle\sum_{k=0}^{n} \left|\hat{f}_{k+1} - \hat{f}_k\right|$, on a donc :

$$\|g_n\|_p \leqslant \sum_{k=0}^{n} \left\|\hat{f}_{k+1} - \hat{f}_k\right\|_p \leqslant \sum_{k=0}^{n} \frac{1}{2^k} \leqslant 2$$

Par convergence monotone, g_n converge ponctuellement sur $X\backslash E$, où E est un ensemble μ-négligeable, vers une fonction g définie presque partout.
Les fonctions g_n étant à valeurs dans $[0, +\infty]$, on a :

$$\|g\|_p^p = \int_X |g(x)|^p \,\mathrm{d}x = \int_X \liminf_{n\to\infty} |g_n(x)|^p \,\mathrm{d}x \leqslant \liminf_{n\to\infty} \int_X |g_n(x)|^p \,\mathrm{d}x \leqslant 2^p.$$

Donc $g \in \mathrm{L}^p(\mu)$. [115]
Soit $x \in X\backslash E$:

$$\forall m \geqslant n \geqslant 1, \left|\hat{f}_m(x) - \hat{f}_n(x)\right| \leqslant \left|\hat{f}_m(x) - \hat{f}_{m-1}(x)\right| + \ldots + \left|\hat{f}_{n+1}(x) - \hat{f}_n(x)\right| = g_{m-1}(x) - g_{n-1}(x)$$

Donc $\forall x \in X\backslash E, \left(\hat{f}_n(x)\right)_n$ est une suite de Cauchy dans $\mathbb{R}$ qui est complet donc elle converge ; on note $\hat{f}(x)$ sa limite.
De plus, $\forall m \geqslant n \geqslant 1, \forall x \in X\backslash E, \left|\hat{f}_m(x) - \hat{f}_n(x)\right| \leqslant g_{m-1}(x)$.
Par passage à la limite : $\forall n \geqslant 1, \forall x \in X\backslash E, \left|\hat{f}(x) - \hat{f}_n(x)\right| \leqslant g(x)$.
En particulier, $\left\|\hat{f}\right\|_p \leqslant \|g\|_p + \left\|\hat{f}_1\right\|_p < \infty$, donc $\hat{f} \in \mathrm{L}^p(\mu)$.
Or $\forall x \in X\backslash E, \left|\hat{f}(x) - \hat{f}_n(x)\right|^p \leqslant (g(x) - g_{n-1}(x))^p \underset{n\to\infty}{\longrightarrow} 0$ et $\left|\hat{f}(x) - \hat{f}_n(x)\right|^p \leqslant g(x)^p$ avec $g \in \mathrm{L}^p(\mu)$.
Donc, par convergence dominée, il vient :

$$\left\|\hat{f} - \hat{f}_n\right\|_p \underset{n\to\infty}{\longrightarrow} 0$$

■

Références

[Bré] H. BRÉZIS – *Analyse fonctionnelle*, 2e éd., Dunod, 2005.
[Rud] W. RUDIN – *Analyse réelle et complexe*, 3e éd., Dunod, 2009.

115. C'est pour montrer que $g \in \mathrm{L}^p(\mu)$ qu'on a besoin d'utiliser [Rud].

Diffusion à titre gratuit uniquement.

Théorème de Stampacchia [116]

Leçons : 213, 219, *205, 206, 208*

[Bré], partie V.3

Théorème

Soit H un espace de Hilbert, et $a : H \times H \to \mathbb{R}$ une forme bilinéaire continue coercive, c'est-à-dire :

$$\exists C \in \mathbb{R}, \forall u, v \in H, |a(u,v)| \leqslant C\|u\|\|v\| \text{ et } \exists \alpha \in \mathbb{R}^{+*}, \forall u \in H, a(u,u) \geqslant \alpha\|u\|^2.$$

Soit K un convexe fermé non-vide de H.
Alors $\forall \varphi \in H', \exists! u \in K, \forall v \in K, a(u, v-u) \geqslant \varphi(v-u)$.
De plus, si a est symétrique, alors u se caractérise par : $\begin{cases} u \in K \\ \frac{1}{2}a(u,u) - \varphi(u) = \min\limits_{v \in K} \left\{ \frac{1}{2}a(v,v) - \varphi(v) \right\} \end{cases}$.

116. Ce qui suit provient directement de la page d'Adrien FONTAINE.

Corollaire (*Lax-Milgram*)

Soit a une forme bilinéaire, continue et coercive. Alors : $\forall \varphi \in H', \exists! u \in H, \forall v \in H, a(u,v) = \varphi(v)$.
De plus, si a est symétrique, alors u se caractérise par : $\begin{cases} u \in H \\ \frac{1}{2}a(u,u) - \varphi(u) = \min\limits_{v \in H} \left\{ \frac{1}{2}a(v,v) - \varphi(v) \right\} \end{cases}$.

Le théorème de Lax-Milgram est un outil fondamental dans l'étude de certaines équations aux dérivées partielles (souvent utilisé dans sa version symétrique). Le théorème de Stampacchia généralise le théorème de Lax-Milgram. Il peut aussi s'obtenir directement à partir de la théorie hilbertienne. Donnons-en maintenant quelques applications.

Soit $I =]0,1[$. On définit l'espace de Sobolev : $\mathrm{H}^1(I) = \left\{ u \in \mathrm{L}^2(I) \middle| \exists g \in \mathrm{L}^2(I), \forall \varphi \in \mathcal{C}_0^\infty(I), \int_I u\varphi' = -\int_I g\varphi \right\}$ et $\mathrm{H}_0^1(I) = \left\{ u \in \mathrm{L}^2(I) \middle| u(0) = u(1) = 0 \text{ et } \exists g \in \mathrm{L}^2(I), \forall \varphi \in \mathcal{C}_0^\infty(I), \int_I u\varphi' = -\int_I g\varphi \right\}$.

Désormais, on veut résoudre le problème (où f est une fonction donnée, par exemple dans $\mathcal{C}(\overline{I})$ ou dans $\mathrm{L}^2(I)$) :

$$\begin{cases} -u'' + u = f \text{ sur } I \\ u(0) = u(1) = 0 \end{cases} \tag{5}$$

La condition aux limites $u(0) = u(1) = 0$ s'appelle condition de Dirichlet homogène.

Une solution classique de (5) est une fonction $u \in \mathcal{C}^2(\overline{I})$ vérifiant (5) au sens usuel. Une solution faible de (5) est une fonction $u \in \mathrm{H}_0^1(I)$ qui vérifie

$$\forall \varphi \in \mathrm{H}_0^1(I), \int_0^1 u'\varphi' + \int_0^1 u\varphi = \int_0^1 f\varphi \tag{6}$$

On remarque que toute solution classique est une solution faible (intégration par parties). Pour établir l'existence de solution au problème (5), on montre l'existence d'une solution au sens faible, puis on montre que cette solution faible est de classe $\mathcal{C}^2$ et qu'une solution faible de classe $\mathcal{C}^2$ est une solution classique. Pour montrer l'existence d'une solution faible, on utilise le théorème de Lax-Milgram.

En effet, dans le cas où $f \in \mathcal{C}(\overline{I})$, on applique Lax-Milgram dans $H = \mathrm{H}_0^1(I)$, avec la forme bilinéaire $a(u,v) = \int_I u'v' + \int_I uv$ et la forme linéaire $\varphi(v) = \int_I fv$, et on montre qu'il existe une unique solution $u \in \mathcal{C}^1(\overline{I})$ de (6).

Un problème se pose lorsque les conditions de Dirichlet ne sont plus homogènes :

$$\begin{cases} -u'' + u = f \text{ sur } I \\ u(0) = \alpha \text{ et } u(1) = \beta \end{cases} \tag{7}$$

En effet, on ne peut plus se placer dans $\mathrm{H}_0^1(I)$. C'est ici qu'on va devoir utiliser le théorème de Stampacchia, qui est plus général que celui de Lax-Milgram.

On prend $f \in \mathrm{L}^2(I)$ et $\alpha, \beta \in \mathbb{R}$. Dans l'espace $\mathrm{H}^1(I)$, on introduit le convexe fermé $K = \{v \in \mathrm{H}^1(I) | v(0) = \alpha \text{ et } v(1) = \beta\}$. Si u est une solution classique de (7), on a : $\forall v \in K, \int_I u'(v-u)' + \int_I u(v-u) = \int_I f(v-u)$. Donc, on a, en particulier :

$$\forall v \in K, \int_I u'(v-u)' + \int_I u(v-u) \geqslant \int_I f(v-u) \tag{8}$$

On utilise alors le théorème de Stampacchia : il existe $u \in K$ unique solution de (8).

Diffusion à titre gratuit uniquement.

Démonstration :

Étape 1 : Laissons de côté l'éventuelle symétrie de a pour commencer.

Soit $\varphi \in H'$; par le théorème de Riesz : $\exists! f \in H, \forall v \in H, \varphi(v) = \langle f, v\rangle$.

D'autre part, à $u \in H$ fixé, $v \mapsto a(u,v)$ est une forme linéaire continue sur H, donc aussi par Riesz : $\exists! A(u) \in H, \forall v \in H, a(u,v) = \langle A(u), v\rangle$.

La bilinéarité de a entraîne que $A \in \mathcal{L}(H,H)$ et on notera désormais Au pour $A(u)$, saluant ainsi cette formidable nouvelle.

De plus, $\forall u \in H, \|Au\|^2 = \langle Au, Au\rangle = a(u,Au) \leqslant C\|u\|\|Au\|$ donc $\|Au\| \leqslant C\|u\|$.[117]

Aussi, $\forall u \in H, \langle Au, u\rangle = a(u,u) \geqslant \alpha\|u\|^2$.

Revenons à ce qu'on veut montrer. Il s'agit de voir que : $\exists! u \in K, \forall v \in K, \langle Au, v-u\rangle \geqslant \langle f, v-u\rangle$.

Soit alors $\rho > 0$, on ajoutera une contrainte sur ρ par la suite.

Ce qu'on veut montrer équivaut à : $\exists! u \in K, \forall v \in K, \langle \rho f - \rho Au + u - u, v-u\rangle \leqslant 0$, ou encore à $\exists! u \in K, \forall v \in K, u = p_K(\rho f - \rho Au + u)$, où p_K désigne la projection sur le convexe fermé K.

On va montrer que pour ρ convenablement choisi, $S : v \mapsto p_K(\rho f - \rho Av + v)$ est une application contractante de K dans K.

p_K étant 1-lipschitzienne, on a : $\forall v_1, v_2 \in K$,

$$\begin{aligned}\|S(v_1) - S(v_2)\|^2 &\leqslant \|(v_1 - v_2) - \rho(Av_1 - Av_2)\|^2 \\ &= \|v_1 - v_2\|^2 - 2\rho\langle v_1 - v_2, Av_1 - Av_2\rangle + \rho^2\|Av_1 - Av_2\|^2 \\ &\leqslant \|v_1 - v_2\|^2\left(1 - 2\rho\alpha + \rho^2 C^2\right)\end{aligned}$$

On fixe donc $\rho > 0$ tel que $1 - 2\rho\alpha + \rho^2C^2 < 1$, c'est-à-dire $0 < \rho < \dfrac{2\alpha}{C^2}$.

Alors S est contractante, et comme K est complet (car fermé dans un Hilbert) ; S admet un unique point fixe u, qui est la solution du problème.

Étape 2 : Supposons désormais que a soit symétrique, et profitons-en pour aller un peu plus loin.

a définit alors un nouveau produit scalaire sur H (comme a est coercive, elle est définie positive), et la norme associée $u \mapsto \sqrt{a(u,u)}$ est équivalente à la norme $\|.\|$ (la continuité et la coercivité donnent chacune une inégalité).

En conséquence, H, muni du produit scalaire a, demeure un espace de Hilbert.

Ainsi, par le théorème de Riesz : $\exists! g \in H, \forall v \in H, \varphi(v) = a(g,v)$.

On a ainsi :

$$\begin{aligned}u \in K \text{ et } \forall v \in K, a(u, v-u) \geqslant \varphi(v-u) &\Leftrightarrow u \in K \text{ et } \forall v \in K, a(g-u, v-u) \leqslant 0 \\ &\Leftrightarrow u = p_K(g) \text{ au sens du produit scalaire } a \\ &\Leftrightarrow u \in K \text{ et } a(g-u, g-u) = \min_{v\in K}\{a(g-v, g-v)\}\end{aligned}$$

Mais $a(g-v, g-v) = a(g,g) - 2a(g,v) + a(v,v)$, et donc minimiser $a(g-v,g-v)$ revient à minimiser $\dfrac{1}{2}a(v,v) - a(g,v) = \frac{1}{2}a(v,v) - \varphi(v)$.

D'où la caractérisation annoncée. ∎

Références

[Bré] H. BRÉZIS – *Analyse fonctionnelle*, 2e éd., Dunod, 2005.

117. Pour être précis, on suppose d'abord que $Au \neq 0$, on peut alors diviser par $\|Au\|$. Évidemment, l'inégalité obtenue tient toujours quand $Au = 0$.

Diffusion à titre gratuit uniquement.

Théorème de Weierstrass par les polynômes de Bernstein [118]

Leçons : 209, 228, 241, 249, 264, *201, 202, 260, 262*

[ZQ], parties VII.IV.1 et XIII.II.1.c

Théorème

Soit $f : [0,1] \to \mathbb{C}$ une fonction continue, et ω son module de continuité uniforme, défini par :

$$\omega(h) = \sup\{|f(u) - f(v)| \mid |u - v| \leqslant h\}$$

Pour $n \in \mathbb{N}^*$, on définit le $n^{\text{ème}}$ polynôme de Bernstein de f :

$$B_n = \sum_{k=0}^{n} \binom{k}{n} X^k (1-X)^{n-k} f\left(\frac{k}{n}\right)$$

Alors :

1. $(B_n)_{n\in\mathbb{N}}$ converge uniformément vers f sur $[0,1]$ et, plus précisément,
   $\exists C \in \mathbb{R}, \forall n \in \mathbb{N}^*, \|f - B_n\|_\infty \leqslant C\omega\left(\frac{1}{\sqrt{n}}\right)$;
2. Cette majoration est essentiellement optimale, au sens où il existe f une fonction lipschitzienne, telle que : $\exists \delta > 0, \forall n \in \mathbb{N}^*, \|f - B_n\|_\infty \geqslant \delta\omega\left(\frac{1}{\sqrt{n}}\right)$.

Démonstration :

Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de vaiid suivant la loi de Bernoulli $b(x)$ où $x \in [0,1]$.

Dès lors, $S_n = \sum_{k=1}^{n} X_k \sim \mathcal{B}(n,x)$, nous fournit, par la loi des grands nombres :

$$\underbrace{\mathbb{E}\left[f\left(\frac{S_n}{n}\right)\right]}_{\underset{n\to\infty}{\longrightarrow}\mathbb{E}[f(x)]=f(x)} = \sum_{k=0}^{n} \binom{k}{n} x^k (1-x)^{n-k} f\left(\frac{k}{n}\right) = B_n(x)$$

Il va s'agir de préciser cette intuition

1. On va avoir besoin du lemme suivant.
   Lemme
   Soit $h \in [0,1]$, $\lambda > 0$, tels que $\lambda h \in [0,1]$.
   Alors $\omega(\lambda h) \leqslant (\lambda+1)\omega(h)$.

   Démonstration :
   Montrons d'abord que ω est sous-additive ; soient $\delta, \varepsilon > 0$.
   On définit $F : \left|\begin{array}{ccc} A_{\delta+\varepsilon} & \to & \mathbb{R} \\ (x,y) & \mapsto & |f(x) - f(y)| \end{array}\right.$, où $A_{\delta+\varepsilon} = \{(x,y) \in [0,1] \mid |x-y| \leqslant \delta + \varepsilon\}$.
   F est continue sur $A_{\delta+\varepsilon}$ compact donc $\exists (x_0, y_0) \in A_{\delta+\varepsilon}, \omega(\delta+\varepsilon) = |f(x_0) - f(y_0)|$.
   Soit $z \in [0,1]$, tel que $|x_0 - z| \leqslant \delta$ et $|y_0 - z| \leqslant \varepsilon$; ainsi

   $$\omega(\delta+\varepsilon) \leqslant |f(x_0) - f(z)| + |f(z) - f(y_0)| \leqslant \omega(\delta) + \omega(\varepsilon)$$

   Ainsi, par une récurrence immédiate, on en déduit : $\forall N \in \mathbb{N}, \omega(Nh) \leqslant N\omega(h)$.
   Et finalement : $\omega(\lambda h) \leqslant \omega\left((\lfloor\lambda\rfloor + 1)h\right) \leqslant (\lfloor\lambda\rfloor + 1)\omega(h) \leqslant (\lambda+1)\omega(h)$. ■

118. Il faut être capable d'en déduire le théorème usuel énoncé habituellement sur l'intervalle $[a,b]$. La première partie du sujet de Mathématiques générales du concours cycle master des ENS de Cachan et Rennes en 2014 utilise le théorème de Weierstrass pour montrer que $\mathbb{Z}[X]$ est dense dans l'ensemble des fonctions continues de $[a,b]$ dans $\mathbb{R}$ si, et seulement si, $\mathbb{Z} \cap [a,b] = \varnothing$.

Diffusion à titre gratuit uniquement.

En particulier, on a : $\omega\left(\left|x-\frac{S_n}{n}\right|\right) \leqslant \left(\sqrt{n}\left|x-\frac{S_n}{n}\right|+1\right)\omega\left(\frac{1}{\sqrt{n}}\right)$.
De plus, en utilisant l'inégalité de Hölder :

$$\begin{aligned}|f(x)-B_n(x)| &\leqslant \mathbb{E}\left[\left|f(x)-f\left(\frac{S_n}{n}\right)\right|\right] \leqslant \mathbb{E}\left[\omega\left(\left|x-\frac{S_n}{n}\right|\right)\right] \leqslant \mathbb{E}\left[\sqrt{n}\left|x-\frac{S_n}{n}\right|+1\right]\omega\left(\frac{1}{\sqrt{n}}\right)\\ &\leqslant \left(\sqrt{n}\left\|x-\frac{S_n}{n}\right\|_1+1\right)\omega\left(\frac{1}{\sqrt{n}}\right) \leqslant \left(\sqrt{n}\left\|x-\frac{S_n}{n}\right\|_2+1\right)\omega\left(\frac{1}{\sqrt{n}}\right)\end{aligned}$$

Or $\left\|x-\frac{S_n}{n}\right\|_2^2 = \mathbb{E}\left[\left(x-\frac{S_n}{n}\right)^2\right] = \operatorname{Var}\left(x-\frac{S_n}{n}\right)+\mathbb{E}\left[x-\frac{S_n}{n}\right]^2 = \frac{\operatorname{Var} S_n}{n^2} = \frac{nx(1-x)}{n^2} = \frac{x(1-x)}{n}$.
Donc on obtient : $|f(x)-B_n(x)| \leqslant \left(\sqrt{x(1-x)}+1\right)\omega\left(\frac{1}{\sqrt{n}}\right) \leqslant \frac{3}{2}\omega\left(\frac{1}{\sqrt{n}}\right)$.

2. Soit $f : x \mapsto \left|x-\frac{1}{2}\right|$; on a :

$$\omega(h) = \sup\left\{\left|\left|x-\frac{1}{2}\right|-\left|y-\frac{1}{2}\right|\right| \middle| |x-y| \leqslant h\right\} \leqslant \sup\{|x-y| \,|\, |x-y| \leqslant h\} = h$$

Désormais $(X_n)_{n\in\mathbb{N}^*}$ est une suite de vaiid suivant la loi $b\left(\frac{1}{2}\right)$.
On a : $\|f-B_n\|_\infty \geqslant \left|f\left(\frac{1}{2}\right)-B_n\left(\frac{1}{2}\right)\right| = B_n\left(\frac{1}{2}\right) = \mathbb{E}\left[f\left(\frac{S_n}{n}\right)\right] = \mathbb{E}\left[\left|\frac{S_n}{n}-\frac{1}{2}\right|\right] = \frac{1}{2n}\mathbb{E}\left[|2S_n-n|\right]$.
Donc $\|f-B_n\|_\infty \geqslant \frac{1}{2n}\mathbb{E}\left[|\varepsilon_1+\ldots+\varepsilon_n|\right]$, avec $\varepsilon_j = 2X_j-1$ variable de Rademacher, (ε_j) famille iid.
On pose [119] $Y = \prod_{j=1}^{n}\left(1+\frac{\mathrm{i}}{\sqrt{n}}\varepsilon_j\right)$, et alors, presque sûrement, on a :

$$|Y| = \prod_{j=1}^{n}\sqrt{1+\frac{\varepsilon_j^2}{n}} = \prod_{j=1}^{n}\sqrt{1+\frac{1}{n}} \leqslant \prod_{j=1}^{n}\exp\left(\frac{1}{n}\right) = \exp\left(\frac{1}{2}\sum_{j=1}^{n}\frac{1}{n}\right) = \sqrt{\mathrm{e}}.$$

Par indépendance des ε_j pour $j \in [\![1,n]\!]$, on obtient :

$$\mathbb{E}\left[\varepsilon_j Y\right] = \mathbb{E}\left[\varepsilon_j\left(1+\frac{\mathrm{i}}{\sqrt{n}}\varepsilon_j\right)\prod_{k\neq j}\left(1+\frac{\mathrm{i}}{\sqrt{n}}\varepsilon_k\right)\right] = \left(\mathbb{E}\left[\varepsilon_j\right]+\frac{\mathrm{i}}{\sqrt{n}}\mathbb{E}\left[\varepsilon_j^2\right]\right)\prod_{k\neq j}\mathbb{E}\left[1+\frac{\mathrm{i}}{\sqrt{n}}\varepsilon_k\right] = \frac{\mathrm{i}}{\sqrt{n}}.$$

Ainsi on obtient :

$$\left|\mathbb{E}\left[\sum_{j=1}^{n}\varepsilon_j Y\right]\right| = \left|\sum_{j=1}^{n}\mathbb{E}\left[\varepsilon_j Y\right]\right| = \left|\sum_{j=1}^{n}\frac{\mathrm{i}}{\sqrt{n}}\right| = \sqrt{n}.$$

Alors que d'autre part :

$$\left|\mathbb{E}\left[\sum_{j=1}^{n}\varepsilon_j Y\right]\right| \leqslant \mathbb{E}\left[|\varepsilon_1+\ldots+\varepsilon_n|\,|Y|\right] \leqslant \sqrt{\mathrm{e}}\mathbb{E}\left[|\varepsilon_1+\ldots+\varepsilon_n|\right].$$

Finalement : $\mathbb{E}\left[|\varepsilon_1+\ldots+\varepsilon_n|\right] \geqslant \sqrt{\frac{n}{\mathrm{e}}}$ puis $\|f-B_n\|_\infty \geqslant \frac{1}{2\sqrt{\mathrm{e}n}} \geqslant \frac{1}{2\sqrt{\mathrm{e}}}\omega\left(\frac{1}{\sqrt{n}}\right)$. ■

Références

[ZQ] H. QUEFFÉLEC et C. ZUILY – *Analyse pour l'agrégation*, Dunod, 2013.

119. Ce "parachutage" correspond en fait à ce qu'on doit faire quand on veut démontrer l'inégalité de Khintchine, à la manière de ce qui est fait dans le livre de C. ZUILY et H. QUEFFÉLEC.

Diffusion à titre gratuit uniquement.

Théorèmes d'Abel angulaire et taubérien faible

Leçons : 207, 223, 230, 235, 241, 243, 244

[Gou An], exercices 4.4.10-11

Théorème (*Abel angulaire*)

Soit $\sum_{n\geqslant 0} a_n z^n$ une série entière de rayon de convergence $\geqslant 1$ et telle que $\sum_{n\geqslant 0} a_n$ converge.
On note f sa somme sur $\mathcal{D}(0,1)$, on fixe $\theta_0 \in \left[0, \frac{\pi}{2}\right[$ et on pose :

$$\Delta_{\theta_0} = \{z \in \mathbb{C} \,|\, |z| < 1\} \cap \left\{1 - \rho e^{i\theta} \,\middle|\, \rho \in \mathbb{R}^{+*}, \theta \in [-\theta_0, \theta_0]\right\}.$$

Alors on a : $\lim\limits_{\substack{z\to 1 \\ z\in\Delta_{\theta_0}}} f(z) = \sum\limits_{n=0}^{\infty} a_n$. [120]

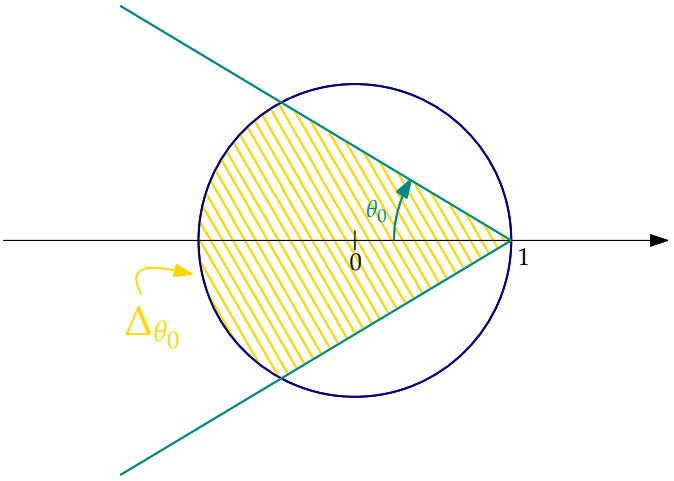
Démonstration :

→ Soient $S = \sum\limits_{n=0}^{\infty} a_n$, $S_N = \sum\limits_{n=0}^{N} a_n$ et $R_N = S - S_N$, pour $N \in \mathbb{N}$.
Soit $z \in \mathbb{C}^*$, avec $|z| < 1$ et $N \in \mathbb{N}^*$,

$$\begin{aligned}
&\sum_{n=0}^{N} a_n z^n - S_N \\
&\quad = \sum_{n=1}^{N} (R_{n-1} - R_n)(z^n - 1) + a_0 - a_0 \\
&\quad = \sum_{n=0}^{N-1} R_n \left(z^{n+1} - 1\right) - \sum_{n=1}^{N} R_n (z^n - 1) \\
&\quad = \sum_{n=0}^{N-1} R_n \left(z^{n+1} - z^n\right) - R_N \left(z^N - 1\right) + 0R_0 \\
&\quad = (z-1) \sum_{n=0}^{N-1} R_n z^n - R_N \left(z^N - 1\right)
\end{aligned}$$

Donc, par passage à la limite $N \to +\infty$, on obtient :

$$f(z) - S = (z-1) \sum_{n=0}^{\infty} R_n z^n \qquad (9)$$



→ Soit $\varepsilon > 0$, et $N \in \mathbb{N}$ tel que $\forall n > N, |R_n| < \varepsilon$.
Alors, pour $|z| < 1$, d'après (9), on a :

$$|f(z) - S| \leqslant |z-1| \left|\sum_{n=0}^{N} R_n z^n\right| + |z-1|\varepsilon \left(\sum_{n=N+1}^{\infty} |z|^n\right) \leqslant |z-1| \left(\sum_{n=0}^{N} |R_n|\right) + \varepsilon \frac{|z-1|}{1-|z|} \qquad (10)$$

→ Soit $z = 1 - \rho e^{i\varphi} \in \Delta_{\theta_0}$, où $\rho > 0$ et $|\varphi| \leqslant \theta_0$.
Alors $|z|^2 = \left(1 - \rho e^{i\varphi}\right)\left(1 - \rho e^{-i\varphi}\right) = 1 - 2\rho\cos\varphi + \rho^2$.
Mais quand $z \to 1$, $\rho \to 0$ et pour $\rho \leqslant \cos\theta_0$, on a :

$$\frac{|z-1|}{1-|z|} = (1+|z|)\frac{|z-1|}{1-|z|^2} \leqslant 2\frac{\rho}{2\rho\cos\varphi - \rho^2} \leqslant \frac{2}{2\cos\varphi - \rho} \leqslant \frac{2}{2\cos\theta_0 - \cos\theta_0} = \frac{2}{\cos\theta_0}$$

120. On a quelques applications de ce résultat : $\sum\limits_{n=0}^{\infty} \frac{(-1)^n}{2n+1} = \frac{\pi}{4}$ et $\sum\limits_{n=0}^{\infty} \frac{(-1)^{n-1}}{n} = \ln 2$.

Diffusion à titre gratuit uniquement.

→ Soit $\alpha > 0$, tel que $\alpha \sum_{n=0}^{N} |R_n| < \varepsilon$.

Alors, quand $z \in \Delta_{\theta_0}$ est tel que $|z-1| \leqslant \min\{\alpha, \cos\theta_0\}$, par (10) :

$$|f(z) - S| \leqslant \varepsilon + \varepsilon \frac{2}{\cos\theta_0}$$

■

Théorème (*Taubérien faible*)

Soit $\sum_{n\geqslant 0} a_n z^n$ une série entière de rayon de convergence $\geqslant 1$, on note f sa somme sur $\mathcal{D}(0,1)$.
On suppose : $\exists S \in \mathbb{C}, \lim_{\substack{x\to 1\\ x<1}} f(x) = S$.
Si $a_n = o\left(\frac{1}{n}\right)$ [121], alors $\sum_{n\geqslant 0} a_n$ converge et $S = \sum_{n=0}^{\infty} a_n$. [122]

Démonstration :

→ Pour $N \in \mathbb{N}^*$, on pose $S_N := \sum_{n=0}^{N} a_n$ et pour $x \in]0,1[$, on a :

$$S_N - f(x) = \sum_{n=0}^{N} a_n (1-x^n) - \sum_{n=N+1}^{\infty} a_n x^n$$

Et pour $n \in \mathbb{N}^*$, $(1-x^n) = (1-x)\left(1 + x + \ldots + x^{n-1}\right) \leqslant (1-x)n$.
Ainsi :

$$|S_N - f(x)| \leqslant (1-x)\sum_{n=1}^{N} n|a_n| + \sum_{n=N+1}^{\infty} \frac{n}{N}|a_n| x^n \leqslant (1-x)MN + \sup_{n>N}(n|a_n|)\frac{1}{N(1-x)},$$

où M est un majorant de la suite $(n|a_n|)_{n\in\mathbb{N}}$, suite qui tend vers 0.

→ Soit $\varepsilon > 0$, avec $\varepsilon < 1$. On a :

$$\forall N \in \mathbb{N}^*, \left|S_N - f\left(1 - \frac{\varepsilon}{N}\right)\right| \leqslant \varepsilon M + \sup_{n>N}(n|a_n|)\frac{1}{\varepsilon}.$$

→ Dès lors, si N_0 est tel que $\sup_{n>N_0}(n|a_n|) < \varepsilon^2$, alors :

$$\forall N \geqslant N_0, \left|S_N - f\left(1 - \frac{\varepsilon}{N}\right)\right| \leqslant M\varepsilon + \varepsilon = (M+1)\varepsilon$$

Par hypothèse, on a : $\lim_{N\to\infty} f\left(1 - \frac{\varepsilon}{N}\right) = S$, donc $\exists N_1 \in \mathbb{N}, \forall N \geqslant N_1, \left|f\left(1 - \frac{\varepsilon}{N}\right) - S\right| \leqslant \varepsilon$.
En fin de compte, par inégalité triangulaire, il vient :

$$\forall N \geqslant \max\{N_0, N_1\}, |S_N - S| \leqslant (M+2)\varepsilon.$$

■

Références

[Gou An] X. GOURDON – *Les maths en tête : Analyse*, 2e éd., Ellipses, 2008.

121. Cette condition s'appelle condition taubérienne.

122. Il s'agit d'une réciproque partielle au théorème d'Abel angulaire. Il reste vrai en supposant seulement $a_n = \mathcal{O}\left(\frac{1}{n}\right)$, et constitue alors le théorème taubérien fort, dont la démonstration est très différente. La réciproque totale du théorème d'Abel angulaire est fausse, comme le montre le contre-exemple suivant : $\lim_{\substack{z\to 1\\ |z|<1}} \sum_{n=0}^{\infty} (-1)^n z^n = \lim_{\substack{z\to 1\\ |z|<1}} \frac{1}{1+z} = \frac{1}{2}$ alors que $\sum_{n\geqslant 0}(-1)^n$ diverge.

Diffusion à titre gratuit uniquement.

Transformée de Fourier-Plancherel [123]

Leçons : 207, 234, 235, 240, *201, 202, 208, 241*

[Far], section IX.2
[Rud], lemme 4.16

Théorème

Soit $f \in L^1 \cap L^2$. On rappelle que $\forall x \in \mathbb{R}, \widehat{f}(x) = \int_{\mathbb{R}} f(t)e^{-ixt}\,dt$.
Alors $\widehat{f} \in L^2$ et $\|f\|_2 = \frac{1}{\sqrt{2\pi}} \left\|\widehat{f}\right\|_2$.

Démonstration :

Étape 1 : On définit $\widetilde{f} : x \to \overline{f(-x)}$ et $g = f \star \widetilde{f}$.

Ainsi, $\forall x \in \mathbb{R}, g(x) = \int_{\mathbb{R}} f(y)\widetilde{f}(x-y)\,dy = \int_{\mathbb{R}} f(x+y)\overline{f(y)}\,dy$ et $g(0) = \int_{\mathbb{R}} |f(y)|^2\,dy = \|f\|_2^2$.

On a : $\forall x \in \mathbb{R}, \widehat{\widetilde{f}}(x) = \int_{\mathbb{R}} \overline{f(-t)}e^{-ixt}\,dt = \overline{\int_{\mathbb{R}} f(u)e^{-ixu}\,du} = \overline{\widehat{f}(x)}$ et $\widehat{g} = \widehat{f \star \widetilde{f}} = \widehat{f}\widehat{\widetilde{f}} = \left|\widehat{f}\right|^2$.

Par inégalité de Young [124], comme $f, \widetilde{f} \in L^1$, alors $g \in L^1$ et $\|g\|_1 \leqslant \|f\|_1 \left\|\widetilde{f}\right\|_1$.

Et par propriété de régularisation [125], comme $f, \widetilde{f} \in L^2$, alors $g \in \mathcal{C}^0$ et $\|g\|_\infty \leqslant \|f\|_2 \left\|\widetilde{f}\right\|_2$.

Étape 2 : Introduisons une partition de l'unité.

On pose, pour $n \in \mathbb{N}^*$ et $t \in \mathbb{R}$: $\Phi_n(t) = e^{-\frac{|t|}{n}}$; pour $x \in \mathbb{R}$:

$$\begin{aligned}\varphi_n(x) = \frac{1}{2\pi}\widehat{\Phi_n}(x) &= \frac{1}{2\pi}\int_{\mathbb{R}} e^{-\frac{|t|}{n}-ixt}\,dt = \frac{1}{2\pi}\int_{\mathbb{R}^-} e^{\frac{t}{n}-ixt}\,dt + \frac{1}{2\pi}\int_{\mathbb{R}^+} e^{-\frac{t}{n}-ixt}\,dt \\ &= \frac{1}{2\pi}\left(\frac{1}{\frac{1}{n}-ix} - \frac{1}{-\frac{1}{n}-ix}\right) = \frac{n}{2\pi}\left(\frac{1}{1-inx} + \frac{1}{1+inx}\right) = \frac{1}{\pi}\frac{n}{1+n^2x^2}\end{aligned}$$

Or $(\varphi_n \star g)(0) = \int_{\mathbb{R}} \varphi_n(y)g(-y)\,dy = \frac{1}{2\pi}\int_{\mathbb{R}}\int_{\mathbb{R}} \Phi_n(t)e^{-iyt}\,dt g(-y)\,dy$.

On va utiliser Fubini car $\int_{\mathbb{R}}\int_{\mathbb{R}} |\Phi_n(t)| \left|e^{-iyt}\right|\,dt\,|g(-y)|\,dy = \|\Phi_n\|_1\|g\|_1 < +\infty$:

$$(\varphi_n \star g)(0) = \frac{1}{2\pi}\int_{\mathbb{R}} \Phi_n(t)\int_{\mathbb{R}} e^{-iyt}g(-y)\,dy\,dt = \frac{1}{2\pi}\int_{\mathbb{R}} \Phi_n(t)\overline{\widehat{g}(t)}\,dt = \frac{1}{2\pi}\int_{\mathbb{R}} \Phi_n(t)\left|\widehat{f}(t)\right|^2\,dt$$

Étape 3 : Montrons que $\lim_{n\to\infty}(\varphi_n \star g)(0) = g(0)$.

Soit $\varepsilon > 0$, comme g est continue : $\exists \eta > 0, \forall x \in \mathbb{R}, |x| < \eta \Rightarrow |g(x) - g(0)| < \varepsilon$.
Ainsi :

$$\begin{aligned}|(\varphi_n \star g)(0) - g(0)| &= \left|\int_{\mathbb{R}} \varphi_n(y)g(-y)\,dy - \int_{\mathbb{R}} \varphi_n(y)g(0)\,dy\right| \\ &\leqslant \underbrace{\int_{-\eta}^{\eta} \varphi_n(y)|g(-y)-g(0)|\,dy}_{\leqslant\varepsilon} + \underbrace{\int_{|y|>\eta} \varphi_n(y)|g(-y)-g(0)|\,dy}_{\leqslant 2\|g\|_\infty \int_{|y|>\eta}\varphi_n(y)\,dy \xrightarrow[n\to\infty]{} 0}\end{aligned}$$

Donc pour n assez grand, $|(\varphi_n \star g)(0) - g(0)| \leqslant 2\varepsilon$.

123. On n'hésitera pas à sauter les calculs qui donnent $\varphi_n(x)$; on admettra l'inversion de la transformée de Fourier dans $\mathcal{S}(\mathbb{R})$, et la densité de $L^1 \cap L^2$ dans L^2 (sauf dans la 235).

124. Inégalité de Young : Soient $p, q \in [1, +\infty]$, tels que $\frac{1}{p} + \frac{1}{q} \geqslant 1$. Soit r vérifiant $1 + \frac{1}{r} = \frac{1}{p} + \frac{1}{q}$. Si $f \in L^p$ et $g \in L^q$, alors $f \star g \in L^r$ et $\|f \star g\|_r \leqslant \|f\|_p\|g\|_q$.

125. Soient $p, p' \in [1, +\infty]$ deux exposants conjugués. Si $f \in L^p$ et $g \in L^{p'}$, alors $f \star g$ est uniformément continue et bornée et $\|f \star g\|_\infty \leqslant \|f\|_p\|g\|_{p'}$.

Diffusion à titre gratuit uniquement.

Étape 4 : Concluons !

Par convergence monotone, puisque $0 \leqslant \Phi_n \left|\widehat{f}\right|^2 \leqslant \Phi_{n+1} \left|\widehat{f}\right|^2$, on a : $\lim\limits_{n\to\infty} (\varphi_n \star g)(0) = \dfrac{1}{2\pi} \displaystyle\int_{\mathbb{R}} \left|\widehat{f}(t)\right|^2 \mathrm{d}t$.

D'où $\|f\|_2^2 = \dfrac{1}{2\pi} \left\|\widehat{f}\right\|_2^2$, montrant également que $\widehat{f} \in \mathrm{L}^2$. ■

Théorème

La transformée de Fourier, définie sur $\mathrm{L}^1 \cap \mathrm{L}^2$, se prolonge en un isomorphisme, proportionnel à une isométrie, de L^2 sur L^2.

Démonstration :

Étape 1 : $\mathrm{L}^1 \cap \mathrm{L}^2$ est dense dans L^2.

Si $f \in \mathrm{L}^2$, alors $f_n = \mathbb{1}_{[-n,n]} f \in \mathrm{L}^1 \cap \mathrm{L}^2$ et par convergence dominée : $\|f - f_n\|_2 \underset{n\to\infty}{\longrightarrow} 0$.

Étape 2 : Montrons qu'on prolonge ainsi la transformée de Fourier à L^2 tout entier.

Soit (f_n) une suite de $\mathrm{L}^1 \cap \mathrm{L}^2$ qui tend vers f dans L^2.

Comme $\left\|\widehat{f_n} - \widehat{f_m}\right\|_2 = \sqrt{2\pi}\, \|f_n - f_m\|_2$, donc $\left(\widehat{f_n}\right)$ est de Cauchy dans L^2 qui est complet donc converge ; on note $\widehat{f}$ sa limite.

Soit (g_n) une autre suite de $\mathrm{L}^1 \cap \mathrm{L}^2$ qui tend vers f dans L^2.

On pose $h_{2n} = f_n$ et $h_{2n+1} = g_n$, pour $n \in \mathbb{N}$; ainsi $h_n \underset{n\to\infty}{\overset{\|\cdot\|_2}{\longrightarrow}} f$.

Ainsi $\left(\widehat{h_n}\right)$ est de Cauchy donc converge dans L^2, donc $\left(\widehat{f_n}\right)$ et $(\widehat{g_n})$ ont même limite.

Cela montre donc que $\widehat{f}$ est indépendant de la suite convergeant vers f.

Et par passage à la limite dans $\|f_n\|_2 = \dfrac{1}{\sqrt{2\pi}} \left\|\widehat{f_n}\right\|_2$, il vient $\|f\|_2 = \dfrac{1}{\sqrt{2\pi}} \left\|\widehat{f}\right\|_2$.

Étape 3 : Montrons que la transformée de Fourier est une bijection de $\mathcal{S}(\mathbb{R})$ sur lui-même.

Par intégrations par parties successives, on montre que :

$$\forall k, p \in \mathbb{N}, (\mathrm{i}t)^k \widehat{f^{(p)}}(t) = \int_{\mathbb{R}} \mathrm{e}^{-\mathrm{i}tx} \frac{\partial^k}{\partial x^k} \left((-\mathrm{i}x)^p f(x)\right) \mathrm{d}x$$

Cela fournit donc : $f \in \mathcal{S}(\mathbb{R}) \Rightarrow \widehat{f} \in \mathcal{S}(\mathbb{R})$.

Ainsi, si $f \in \mathcal{S}(\mathbb{R})$, alors $\widehat{f} \in \mathrm{L}^1$, et la formule d'inversion de la transformée de Fourier fournit : $f(x) = \dfrac{1}{2\pi} \widehat{\widehat{f}(-x)}$.

Étape 4 : On en déduit que l'image de L^2 par $\widehat{\ }$ est dense et fermée dans L^2.

On a : $\underbrace{\mathcal{C}_{\mathrm{C}}^\infty(\mathbb{R})}_{\substack{\text{dense dans} \\ \mathrm{L}^2(\mathbb{R})}} \subset \mathcal{S}(\mathbb{R}) \subset \mathrm{Im}\,(\widehat{\ }) \subset \mathrm{L}^2(\mathbb{R})$, donc $\mathrm{Im}\,(\widehat{\ })$ est dense dans $\mathrm{L}^2(\mathbb{R})$.

Soit $g \in \mathrm{L}^2(\mathbb{R})$, comme $\mathrm{Im}\,(\widehat{\ })$ est dense dans $\mathrm{L}^2(\mathbb{R})$: $\exists\, (f_n) \in \mathrm{L}^2(\mathbb{R})^{\mathbb{N}}, \widehat{f_n} \underset{n\to\infty}{\longrightarrow} g$.

Donc $\left(\widehat{f_n}\right)$ est de Cauchy, alors (f_n) est de Cauchy dans L^2 donc converge vers $f \in \mathrm{L}^2(\mathbb{R})$.

Par continuité de $\widehat{\ }$: $f_n \underset{n\to\infty}{\overset{\mathrm{L}^2}{\longrightarrow}} f$ donc $\widehat{f_n} \underset{n\to\infty}{\overset{\mathrm{L}^2}{\longrightarrow}} \widehat{f}$ d'où $g = \widehat{f}$.

Donc $\mathrm{Im}\,(\widehat{\ }) = \mathrm{L}^2(\mathbb{R})$.

■

Références

[Far] J. FARAUT – *Calcul intégral*, EDP Sciences, 2006.

[Rud] W. RUDIN – *Analyse réelle et complexe*, 3e éd., Dunod, 2009.

Diffusion à titre gratuit uniquement.

Ellipsoïde de John-Loewner [126]

Leçons : 152, 160, 170, 171, *158, 181, 203, 219, 229, 253*

[X-ENS Al3], exercice 3.37

Théorème

Soit K un compact d'intérieur non-vide de $\mathbb{R}^n$.
Il existe un unique ellipsoïde centré en 0, de volume minimal, contenant K.

Notations : Q (resp. Q^+, Q^{++}) désigne l'ensemble des formes quadratiques (resp. positives, définies positives) de $\mathbb{R}^n$. Si $q \in Q$, on désigne par $D(q)$ le déterminant d'une matrice de q dans une base orthonormale (on montrera dans le lemme 1 que $D(q)$ ne dépend pas de la base orthonormale choisie).

Démonstration :

On munit $\mathbb{R}^n$ de sa structure euclidienne usuelle.
Un ellipsoïde centré en 0 de $\mathbb{R}^n$ a une équation du type $q(x) \leqslant 1$, où $q \in Q^{++}$, et on pose alors :

$$\mathcal{E}_q = \{x \in \mathbb{R}^n | q(x) \leqslant 1\}$$

Lemme 1

Le volume de $\mathcal{E}_q$ est $V_q = \dfrac{V_0}{\sqrt{D(q)}}$, où V_0 désigne le volume de la boule unité pour la norme euclidienne canonique.

Démonstration du lemme 1 :

Il existe une base orthonormale $\mathcal{B} = (e_1, \dots, e_n)$ dans laquelle q s'écrit $q(x) = \sum_{i=1}^{n} a_i x_i^2$, avec tous les $a_i > 0$, car $q \in Q^{++}$.
Ainsi : $V_q = \int \dots \int_{\sum_{i=1}^n a_i x_i^2 \leqslant 1} \mathrm{d}x_1 \dots \mathrm{d}x_n$. [127]
Le changement de variables défini par $x_i = \dfrac{t_i}{\sqrt{a_i}}$ est un $\mathcal{C}^1$-difféomorphisme de jacobien $\dfrac{1}{\sqrt{a_1 \dots a_n}}$.
Dès lors : $V_q = \int \dots \int_{\sum_{i=1}^n t_i^2 \leqslant 1} \dfrac{\mathrm{d}t_1 \dots \mathrm{d}t_n}{\sqrt{a_1 \dots a_n}}$.
Soit S la matrice de q dans une base orthonormale de $\mathbb{R}^n$ quelconque.
$S \in \mathcal{S}_n(\mathbb{R})$ donc $\exists P \in \mathcal{O}_n(\mathbb{R}), S = P \mathrm{diag}\,(a_1, \dots, a_n)\,{}^tP$, d'où $\det S = a_1 \dots a_n$.
Donc $D(q) = a_1 \dots a_n$ est indépendant de la base orthonormale choisie et $V_q = \dfrac{V_0}{\sqrt{D(q)}}$. ■

126. Comme d'habitude, il faut savoir déterminer cet ellipsoïde sur quelques exemples : triangle équilatéral centré en l'origine, carré dont un des sommets est l'origine... Et il faut également savoir dire quelques mots sur son utilité. Je vous propose ici de démontrer le résultat du développement sur les sous-groupes compacts de $\mathrm{GL}_n(\mathbb{R})$ (voir en page 114).

On désigne par $\mathcal{B}$ la boule unité de $\mathbb{R}^n$, muni du produit scalaire usuel. Soit G un sous-groupe compact de $\mathrm{GL}_n(\mathbb{R})$, on pose $K = \bigcup_{A \in G} A(\mathcal{B})$ et on définit $\varphi : \left| \begin{array}{ccc} G \times \mathcal{B} & \to & \mathbb{R}^n \\ (A, x) & \mapsto & Ax \end{array} \right.$ de sorte que $K = \mathrm{Im}\,\varphi$.

Comme G est compact et $\mathcal{B}$ aussi (on est en dimension finie !), $G \times \mathcal{B}$ est compact, et par continuité de φ, K aussi. De plus, $0 \in \mathcal{B} \subset K \Rightarrow 0 \in \overset{\circ}{K}$; par le théorème de John-Loewner, il existe une unique matrice $S \in \mathcal{S}_n^{++}(\mathbb{R})$, telle que $K \subset \mathcal{E}_{q_S}$, où $\mathcal{E}_{q_S}$ est de volume minimal et $q_S : x \mapsto {}^txSx$.

On va montrer que $G \subset \mathcal{O}(S) = \left\{M \in \mathcal{M}_n(\mathbb{R}) \middle| {}^tMSM = S\right\}$. Soit $M \in G$. Soit $y = Ax \in K$, avec $x \in \mathcal{B}$; alors $y = M\underbrace{M^{-1}Ax}_{\in K}$, donc $K \subset M(K)$. Réciproquement, si $x \in M(K)$, alors $x = MAy$, avec $A \in G$ et $y \in \mathcal{B}$; comme $MA \in G$, on a $x \in K$, d'où $K = M(K)$. G étant compact, les suites $(M^p)_{p \in \mathbb{N}}$ et $\left(M^{-p}\right)_{p \in \mathbb{N}}$ sont bornées ; on en déduit alors que $|\det M| = 1$.

Posons $R = {}^tMSM$; alors $R \in \mathcal{S}_n^{++}(\mathbb{R})$, car ${}^txRx = {}^t(Mx)S(Mx) > 0$ dès que $Mx \neq 0$ (ie $x \neq 0$), vu que $S \in \mathcal{S}_n^{++}(\mathbb{R})$. Aussi, si $x \in K$, alors ${}^txRx = {}^t(Mx)S(Mx) \leqslant 1$, car $Mx \in M(K) = K \subset \mathcal{E}_{q_S}$. Mais comme $|\det M| = 1$, on a : $\det R = \det S$, d'où : $V_{q_S} = V_{q_R}$. Par unicité, on a donc $R = S$ et ${}^tMSM = S$, c'est-à-dire $M \in \mathcal{O}(S)$.

127. On peut se demander si le volume est ainsi bien défini. En effet, il semble dépendre du choix de la base orthonormée choisie. Mais il n'en est rien. Si on change de base orthonormale, alors la matrice de passage est une matrice orthogonale. Le jacobien du changement de variables vaut alors 1, et donc le volume ne dépend finalement pas de la base orthonormée choisie !

Diffusion à titre gratuit uniquement.

On reformule alors le problème : montrons qu'il existe un unique $q \in Q^{++}$, tel que $D(q)$ soit maximal et tel que $\forall x \in K, q(x) \leqslant 1$.
On munit Q de la norme $N : q \mapsto \sup_{\|x\| \leqslant 1} |q(x)|$ et on pose $\mathcal{A} = \{q \in Q^+ | \forall x \in K, q(x) \leqslant 1\}$; on va chercher à maximiser $D(q)$ sur $\mathcal{A}$.
Montrons que $\mathcal{A}$ est un compact convexe non-vide de Q.

$\mathcal{A}$ **est convexe :** soient $q, q' \in \mathcal{A}$, $\lambda \in [0,1]$;
D'une part : $\forall x \in \mathbb{R}^n, (\lambda q + (1-\lambda)q')(x) = \lambda q(x) + (1-\lambda)q'(x) \geqslant 0$.
Et d'autre part : $\forall x \in K, (\lambda q + (1-\lambda)q')(x) \leqslant \lambda + (1-\lambda) = 1$.
Donc $\lambda q + (1-\lambda)q' \in \mathcal{A}$.

$\mathcal{A}$ **est fermé :** soit $(q_n)_{n\in\mathbb{N}}$ une suite dans $\mathcal{A}$ qui converge vers q dans (Q, N).
Alors $\forall x \in \mathbb{R}^n, |q(x) - q_n(x)| \leqslant N(q - q_n)\|x\|^2$ donc $\lim_{n\to\infty} q_n(x) = q(x)$.
Conséquemment, $\forall x \in \mathbb{R}^n, q(x) \geqslant 0$ et $\forall x \in K, q(x) \leqslant 1$ et donc $q \in \mathcal{A}$.

$\mathcal{A}$ **est borné :** soit $q \in \mathcal{A}$.
K est d'intérieur non-vide donc $\exists a \in K, \exists r > 0, \mathcal{B}(a, r) \subset K$.
Soit $x \in \mathbb{R}^n$, si $\|x\| \leqslant r$, alors $a + x \in K$ donc $q(a + x) \leqslant 1$.
D'autre part $q(-a) = q(a) \leqslant 1$ et par l'inégalité de Minkowski [128] :

$$\sqrt{q(x)} = \sqrt{q(x + a - a)} \leqslant \sqrt{q(x + a)} + \sqrt{q(-a)} \leqslant 2 \text{ donc } q(x) \leqslant 4$$

Maintenant, si $\|x\| \leqslant 1$, $q(x) = \frac{1}{r^2} q(rx) \leqslant \frac{4}{r^2}$ et donc $N(q) \leqslant \frac{4}{r^2}$.

$\mathcal{A}$ **est non-vide :** comme K est compact, $\exists M > 0, \forall x \in K, \|x\| \leqslant M$.
Alors $q : x \mapsto \frac{\|x\|^2}{M^2}$ est dans $\mathcal{A}$.

Comme det est continue, $q \mapsto D(q)$ est continue sur $\mathcal{A}$, qui est compact.
Donc D atteint son maximum sur $\mathcal{A}$ en un point noté q_0.
Et comme $\left(x \mapsto \frac{\|x\|^2}{M^2}\right) \in \mathcal{A}$ et est définie positive, on a : $D(q_0) > 0$, ie $q_0 \in Q^{++}$.
Il existe donc un ellipsoïde $\mathcal{E}_{q_0}$ de volume minimal contenant K ; reste à montrer qu'il est unique, c'est-à-dire, montrer que q_0 est unique.

Lemme 2 (*Stricte convexité logarithmique du déterminant*)

Soient $A, B \in \mathcal{S}_n^{++}(\mathbb{R})$, $\alpha, \beta \in \mathbb{R}^+$ vérifiant $\alpha + \beta = 1$.
Alors $\det(\alpha A + \beta B) \geqslant (\det A)^\alpha (\det B)^\beta$.
Et si $A \neq B$, l'inégalité est stricte. [129]

Démonstration du lemme 2 :
Comme $A \in \mathcal{S}_n^{++}(\mathbb{R})$ et $B \in \mathcal{S}_n(\mathbb{R})$, par pseudo-réduction simultanée [130] :

$$\exists P \in \mathrm{GL}_n(\mathbb{R}), A = {}^tPP \text{ et } B = {}^tP\,\mathrm{diag}(\lambda_1, \ldots, \lambda_n)\, P$$

On note $D = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$.

128. On part de l'inégalité de Cauchy-Schwarz : $\varphi(x, y)^2 \leqslant q(x)q(y)$.
Ainsi $\frac{q(x+y)-q(x)-q(y)}{2} \leqslant \sqrt{q(x)q(y)}$ donc $q(x + y) \leqslant q(x) + 2\sqrt{q(x)q(y)} + q(y)$, d'où : $\sqrt{q(x + y)} \leqslant \sqrt{q(x)} + \sqrt{q(y)}$.

129. On ne démontrera le lemme 2 que dans les leçons ayant trait à la convexité des fonctions : 229 et 253 ; dans les autres, on démontrera plutôt le lemme 1.

130. Matriciellement, le théorème spectral dit que si $A \in \mathcal{S}_n(\mathbb{R})$, il existe $P \in \mathcal{O}_n(\mathbb{R})$, telle que $D = P^{-1}AP$ soit diagonale. Mais comme $P^{-1} = {}^tP$, cela signifie également que A et D sont congruentes. Aussi, on peut dire que, dans la base de $\mathbb{R}^n$ définie par les colonnes de P, la matrice de la forme quadratique $q_A : X \mapsto {}^tXAX$ est diagonale. Cette nouvelle base est donc à la fois orthonormée pour le produit scalaire canonique de $\mathbb{R}^n$ et orthogonale pour q_A. Ce fait se généralise en le théorème qui suit.

Théorème (*Pseudo-réduction simultanée*)

- Si q et q' sont deux formes quadratiques sur un $\mathbb{R}$-espace vectoriel E de dimension finie, et si q est définie positive, Alors, il existe une base de E orthonormée pour q et orthogonale pour q'.
- Si $A, B \in \mathcal{S}_n(\mathbb{R})$ et si A est définie positive, Alors, il existe $P \in \mathrm{GL}_n(\mathbb{R})$ telle que : ${}^tPAP = \mathrm{I}_n$ et ${}^tPBP$ est diagonale.

Diffusion à titre gratuit uniquement.

Et comme $B \in \mathcal{S}_n^{++}(\mathbb{R})$, $\forall i \in [\![1,n]\!], \lambda_i > 0$.[131]
Ainsi : $(\det A)^\alpha (\det B)^\beta = (\det P)^{2\alpha} (\det P)^{2\beta} (\det D)^\beta = (\det P)^2 (\det D)^\beta$.
Et $\det(\alpha A + \beta B) = (\det P)^2 \det(\alpha \mathrm{I}_n + \beta D)$.
On veut alors montrer que :

$$\det(\alpha \mathrm{I}_n + \beta D) \geqslant (\det D)^\beta \Leftrightarrow \prod_{i=1}^n (\alpha + \beta\lambda_i) \geqslant \left(\prod_{i=1}^n \lambda_i\right)^\beta \Leftrightarrow \sum_{i=1}^n \ln(\alpha + \beta\lambda_i) \geqslant \beta \sum_{i=1}^n \ln \lambda_i$$

Or, par concavité du logarithme : $\forall i \in [\![1,n]\!], \ln(\alpha + \beta\lambda_i) \geqslant \alpha \ln(1) + \beta \ln(\lambda_i) = \beta \ln(\lambda_i)$.
On obtient le résultat souhaité en sommant sur $i \in [\![1,n]\!]$.
Et si $A \neq B$, un des λ_i est différent de 1, et donc, par stricte convexité du logarithme, on obtient une inégalité stricte. ■

Par l'absurde, soit $q \in \mathcal{A} \setminus \{q_0\}$, tel que $D(q) = D(q_0)$.
Soient S et S_0 les matrices de q et q_0 dans la base canonique de $\mathbb{R}^n$.
Comme $\mathcal{A}$ est convexe, $\frac{1}{2}(q + q_0) \in \mathcal{A}$, et par stricte convexité logarithmique du déterminant sur $\mathcal{S}_n^{++}(\mathbb{R})$:

$$D\left(\frac{1}{2}(q+q_0)\right) = \det\left(\frac{1}{2}(S+S_0)\right) > (\det S)^{\frac{1}{2}} (\det S_0)^{\frac{1}{2}} = \det S_0 = D(Q_0)$$

Ce qui contredit la maximalité de q_0. D'où l'unicité. ■

Références

[X-ENS Al3] S. FRANCINOU, H. GIANELLA et S. NICOLAS – *Oraux X-ENS Algèbre 3*, Cassini, 2010.

131. Alors là, je dis attention ! Contrairement à ce que peut laisser penser la notation, les λ_i NE SONT PAS les valeurs propres de B. En effet, on n'a pas diagonalisé la matrice B, puisque P est dans $\mathrm{GL}_n(\mathbb{R})$ et pas dans $\mathcal{O}_n(\mathbb{R})$. Et à vrai dire, P ne peut pas être dans $\mathcal{O}_n(\mathbb{R})$ à partir du moment où $A \neq \mathrm{I}_n$, vu que $A = {}^tPP$. Pour montrer que les λ_i sont tous positifs, il suffit de supposer, par exemple, que λ_r soit négatif ou nul. Alors on prend $X = P^{-1}e_r \in \mathbb{R}^n$ et ${}^tXBX = \lambda_r \leqslant 0$ ce qui contredit le fait que $B \in \mathcal{S}_n^{++}(\mathbb{R})$. Et vous ne pourrez pas dire qu'on ne vous avait pas prévenu.

Diffusion à titre gratuit uniquement.

Lemme de Morse

Leçons : 158, 170, 171, 214, 215, 218

[Rou], exercice 114

Théorème

Soit U un ouvert de $\mathbb{R}^n$ avec $0 \in U$ et $f : U \to \mathbb{R}$ de classe $\mathcal{C}^3$.
On suppose que $\mathrm{D}f(0) = 0$ et que $\mathrm{D}^2 f(0)$ est non-dégénérée, de signature $(p, n-p)$.
Alors il existe un $\mathcal{C}^1$-difféomorphisme φ entre deux voisinages de l'origine dans $\mathbb{R}^n$, tel que $\varphi(0) = 0$ et $f(x) - f(0) = u_1^2 + \ldots + u_p^2 - u_{p+1}^2 - \ldots - u_n^2$, où $u = \varphi(x)$.

Démonstration :

On applique la formule de Taylor avec reste intégral à l'ordre 1 :

$$f(x) - f(0) - \mathrm{D}f(0).x = \int_0^1 \frac{(1-t)^1}{1!} \mathrm{D}^2 f(tx).(x,x)\,\mathrm{d}t$$

Ainsi $f(x) - f(0) = {}^{\mathrm{t}}xQ(x)x$, où $Q : x \mapsto \displaystyle\int_0^1 (1-t)\mathrm{D}^2 f(tx)\,\mathrm{d}t$ est une fonction $\mathcal{C}^1$. [132]

On va avoir besoin du lemme suivant :

Lemme

Soit $A_0 \in \mathrm{GL}_n(\mathbb{R}) \cap \mathcal{S}_n(\mathbb{R})$.
Alors il existe V, voisinage ouvert de A_0 dans $\mathcal{S}_n(\mathbb{R})$ et $\rho : V \to \mathrm{GL}_n(\mathbb{R})$, de classe $\mathcal{C}^1$, telle que :

$$\forall A \in V, A = {}^{\mathrm{t}}\rho(A) A_0 \rho(A)$$

Démonstration :

Étape 1 : Considérons $\chi : \left| \begin{array}{ccc} \mathcal{M}_n(\mathbb{R}) & \to & \mathcal{S}_n(\mathbb{R}) \\ M & \mapsto & {}^{\mathrm{t}}MA_0M \end{array} \right.$; c'est une application polynomiale, donc $\mathcal{C}^1$.

Pour $H \in \mathcal{M}_n(\mathbb{R})$, utilisant que A_0 est symétrique,

$$\begin{aligned} \chi(\mathrm{I}_n + H) - \chi(\mathrm{I}_n) &= {}^{\mathrm{t}}(\mathrm{I}_n + H) A_0 (\mathrm{I}_n + H) - A_0 = {}^{\mathrm{t}}HA_0 + A_0H + {}^{\mathrm{t}}HA_0H \\ &= {}^{\mathrm{t}}(A_0H) + A_0H + o\left(\|H\|^2\right) \end{aligned}$$

Ainsi $\mathrm{D}\chi(\mathrm{I}_n).H = {}^{\mathrm{t}}(A_0H) + A_0H$ donc $H \in \mathrm{Ker}\,(\mathrm{D}\chi(\mathrm{I}_n)) \Leftrightarrow A_0H \in \mathcal{A}_n(\mathbb{R})$.

Étape 2 : On aimerait appliquer le théorème d'inversion locale à χ... mais on ne peut pas.
On pose $F = \{H \in \mathcal{M}_n(\mathbb{R}) | A_0H \in \mathcal{S}_n(\mathbb{R})\}$.
Soit $\psi = \chi_{|F} : F \to \mathcal{S}_n(\mathbb{R})$. $\mathrm{I}_n \in F$ et $\mathrm{Ker}\,(\mathrm{D}\psi(\mathrm{I}_n)) = \mathrm{Ker}\,(\mathrm{D}\chi(\mathrm{I}_n)) \cap F = \{0\}$.
Comme $\dim F = \dim \mathcal{S}_n(\mathbb{R})$, $\mathrm{D}\psi(\mathrm{I}_n)$, restriction de $\mathrm{D}\chi(\mathrm{I}_n)$ à F, est bijective.
Et comme ψ est de classe $\mathcal{C}^1$, par le théorème d'inversion locale, il existe un voisinage ouvert U de I_n dans F tel que ψ soit un $\mathcal{C}^1$-difféomorphisme de U sur $V = \psi(U)$.
On peut supposer $U \subset \mathrm{GL}_n(\mathbb{R})$, quitte à prendre $U \cap U'$ où U' est un voisinage ouvert de I_n dans $\mathrm{GL}_n(\mathbb{R})$; un tel U' existant par continuité de det.
Ainsi, V est un voisinage ouvert de $A_0 = \psi(\mathrm{I}_n)$ dans $\mathcal{S}_n(\mathbb{R})$, et :

$$\forall A \in V, A = {}^{\mathrm{t}}\psi^{-1}(A) A_0 \psi^{-1}(A)$$

Et il suffit alors de poser $\rho = \psi^{-1}$. ∎

Ici, $Q(x)$ est toujours symétrique et $Q(0) = \dfrac{1}{2}\mathrm{D}^2 f(0)$ est inversible.
Par le lemme, il existe V, voisinage de $Q(0)$ dans $\mathcal{S}_n(\mathbb{R})$ et $\rho : V \to \mathrm{GL}_n(\mathbb{R})$ de classe $\mathcal{C}^1$, telle que :

$$\forall A \in V, A = {}^{\mathrm{t}}\rho(A) Q(0) \rho(A)$$

132. Si on vous demande pourquoi, dites qu'il y a une dérivation sous le signe intégrale.

Diffusion à titre gratuit uniquement.

Et comme Q est continue, il existe W, voisinage de 0 dans $\mathbb{R}^n$, tel que :

$$\forall x \in W, Q(x) \in V \text{ et } Q(x) = {}^t\rho(Q(x))Q(0)\rho(Q(x))$$

On pose alors $M(x) = \rho(Q(x))$ et $y = M(x)x$, on obtient : $f(x) - f(0) = {}^tyQ(0)y$.
Par le théorème d'inertie de Sylvester, $\exists A \in \mathrm{GL}_n(\mathbb{R}), {}^tAQ(0)A = \left(\begin{array}{c|c} \mathrm{I}_p & 0 \\ \hline 0 & -\mathrm{I}_{n-p} \end{array}\right)$, car $Q(0)$ est de signature $(p, n-p)$.
Alors, en posant $y = Au$, on obtient :

$$f(x) - f(0) = {}^tyQ(0)y = {}^tu{}^tAQ(0)Au = u_1^2 + \ldots + u_p^2 - u_{p+1}^2 - \ldots - u_n^2$$

Soit alors $\varphi : x \mapsto u = A^{-1}M(x)x$; on a bien $\varphi(0) = 0$ et φ est $\mathcal{C}^1$ sur W.
Puis, pour $h \in W$, $\varphi(h) - \varphi(0) = A^{-1}M(h)h - A^{-1}M(0)0 = A^{-1}(M(0) + o(1))h = A^{-1}M(0)h + o(\|h\|)$, donc $\mathrm{D}\varphi(0) = A^{-1}M(0)$ qui est inversible.
On applique le théorème d'inversion locale à φ, qui est donc un $\mathcal{C}^1$-difféomorphisme entre deux voisinages de 0 dans $\mathbb{R}^n$. [133] ■

Références

[Rou] F. ROUVIÈRE – *Petit guide de calcul différentiel*, 4e éd., Cassini, 2014.

133. Il n'est cependant pas évident qu'il soit nécessaire d'avoir M de classe $\mathcal{C}^1$ (ie f de classe $\mathcal{C}^3$) pour que φ soit de classe $\mathcal{C}^1$. Clarifions donc tout cela.
Soient $x \in W$ et $h \in \mathbb{R}^n$ tel que $x + h \in W$:

$$\begin{aligned} \varphi(x+h) - \varphi(x) &= A^{-1}M(x+h)(x+h) - A^{-1}M(x)x \\ &= A^{-1}(M(x) + \mathrm{D}M(x).h + o(\|h\|))(x+h) - A^{-1}M(x)x \\ &= A^{-1}M(x)h + A^{-1}(\mathrm{D}M(x).h)x + o(\|h\|) \end{aligned}$$

Et c'est la continuité de M et de DM qui rend continue l'application :

$$\mathrm{D}\varphi : x \longmapsto \Big(h \mapsto A^{-1}(M(x)h + (\mathrm{D}M(x).h)x)\Big)$$

Diffusion à titre gratuit uniquement.

Partitions d'un entier en parts fixées

Leçons [134] **: 124, 126, 140, 190, *102, 224***

[X-ENS An2], exercice 3.15

Théorème

Soient $a_1, \dots, a_k \in \mathbb{N}^*$ premiers entre eux dans leur ensemble.
Pour $n \in \mathbb{N}^*$, on note

$$u_n = \#\left\{(x_1, \dots, x_k) \in \mathbb{N}^k \middle| a_1x_1 + \dots + a_kx_k = n\right\}$$

Alors on a :

$$u_n \underset{n\to\infty}{\sim} \frac{1}{a_1a_2\dots a_k}\frac{n^{k-1}}{(k-1)!}$$

Démonstration :

Étape 1 : Considérons le produit de Cauchy des séries formelles $\sum_{x_i=0}^{\infty} X^{a_ix_i}$, pour $i \in [\![1,k]\!]$.

On note $f(X)$ ce produit de Cauchy ; on a les égalités suivantes :

$$f(X) = \prod_{i=1}^{k}\left(\sum_{x_i=0}^{\infty} X^{a_ix_i}\right) = \prod_{i=1}^{k}\frac{1}{1-X^{a_i}}$$
$$= \sum_{n=0}^{\infty}\left(\sum_{\substack{(x_1,\dots,x_k)\in\mathbb{N}^k \\ a_1x_1+\dots+a_kx_k=n}} 1\right)X^n = \sum_{n=0}^{\infty} u_nX^n$$

Étape 2 : Décomposons la fraction rationnelle $f(X)$ en éléments simples.

La série formelle $f(X)$ est la série génératrice de la suite $(u_n)_{n\in\mathbb{N}}$; c'est une fraction rationnelle dont les pôles sont les racines $a_1^{\text{èmes}}, \dots, a_k^{\text{èmes}}$ de l'unité.
Le pôle 1 est de multiplicité k.

Soit $\omega \neq 1$ un pôle de f. Comme $\frac{1}{1-X^{a_i}}$, où $i \in [\![1,k]\!]$, n'a que des pôles simples, ω est de multiplicité inférieure ou égale à k. Par l'absurde, on suppose que ω soit un pôle de multiplicité k.
On aurait alors : $\forall i \in [\![1,k]\!], \omega^{a_i} = 1$.
Or, d'après le théorème de Bézout, les $(a_i)_{i\in[\![1,k]\!]}$ étant premiers entre eux dans leur ensemble :

$$\exists\,(u_1, \dots, u_k) \in \mathbb{Z}^k, a_1u_1 + \dots + a_ku_k = 1$$

Alors $\omega = \omega^{\sum_{i=1}^{k} a_iu_i} = \prod_{i=1}^{k}(\omega^{a_i})^{u_i} = 1$.
Contradiction ! On en déduit donc que la multiplicité du pôle $\omega \neq 1$ est strictement inférieure à k.
Notons $\mathcal{P} = \{\omega_1, \dots, \omega_p\}$ l'ensemble des pôles de $f(X)$, avec $\omega_1 = 1$.
Par décomposition en éléments simples, il existe $c_{i,j} \in \mathbb{C}$ pour $i \in [\![1,p]\!]$, $j \in [\![1,k-1]\!]$ et $\alpha \in \mathbb{C}$, tels que :

$$f(X) = \frac{\alpha}{(1-X)^k} + \sum_{\substack{1\leqslant i\leqslant p \\ 1\leqslant j\leqslant k-1}}\frac{c_{i,j}}{(\omega_i - X)^j}$$

Étape 3 : Développons en série formelle les éléments simples de $f(X)$.

En effet, pour $\omega \in \mathcal{P}$ et $j \in [\![1,k]\!]$, $\frac{1}{(\omega-X)^j}$ est développable en série formelle. Ses coefficients

134. Pas besoin d'arguments d'analyse dans ce développement, les séries formelles suffisent. Il est donc tout à fait artificiel de faire ce développement avec des séries entières pour le caser en analyse.

Diffusion à titre gratuit uniquement.

s'obtiennent en dérivant $(j-1)$ fois le développement en série formelle de $\dfrac{1}{\omega - X}$.
On a :

$$\frac{1}{\omega - X} = \frac{1}{\omega}\frac{1}{1-\frac{X}{\omega}} = \frac{1}{\omega}\sum_{n=0}^{\infty}\left(\frac{X}{\omega}\right)^n = \sum_{n=0}^{\infty}\frac{X^n}{\omega^{n+1}}$$

Puis, pour $j \in [\![1,k]\!]$,

$$\frac{(j-1)!}{(\omega - X)^j} = \sum_{n=j-1}^{\infty} n(n-1)\dots(n-j+2)\frac{X^{n-j+1}}{\omega^{n+1}}$$

Par conséquent,

$$\frac{1}{(\omega - X)^j} = \sum_{n=j-1}^{\infty}\frac{n!}{(j-1)!(n-j+1)!}\frac{X^{n-j+1}}{\omega^{n+1}} = \sum_{n=0}^{\infty}\frac{(n+j-1)!}{(j-1)!n!}\frac{X^n}{\omega^{n+j}} = \sum_{n=0}^{\infty}\binom{n+j-1}{n}\frac{X^n}{\omega^{n+j}}$$

Ainsi :

$$f(X) = \alpha\sum_{n=0}^{\infty}\binom{n+k-1}{n}X^n + \sum_{\substack{1\leqslant i\leqslant p\\ 1\leqslant j\leqslant k-1}} c_{i,j}\left(\sum_{n=0}^{\infty}\binom{n+j-1}{n}\frac{X^n}{\omega^{n+j}}\right)$$

Étape 4 : Déduisons-en un équivalent de u_n en l'infini.
La dernière expression de $f(X)$ nous fournit, par unicité du développement en série formelle :

$$u_n = \alpha\binom{n+k-1}{n} + \sum_{\substack{1\leqslant i\leqslant p\\ 1\leqslant j\leqslant k-1}} c_{i,j}\binom{n+j-1}{n}\frac{1}{\omega^{n+j}}$$

Or, pour $r \in \mathbb{N}^*$, on a : $\dbinom{n+r-1}{n} = \dfrac{(n+r-1)\dots(n+1)}{(r-1)!} \underset{n\to\infty}{\sim} \dfrac{n^{r-1}}{(r-1)!}$.
Ainsi, $\alpha\dbinom{n+r-1}{n} \underset{n\to\infty}{\sim} \alpha\dfrac{n^{r-1}}{(r-1)!}$
Et $\forall i \in [\![1,p]\!], \forall j \in [\![1-k-1]\!], c_{i,j}\dbinom{n+j-1}{n}\dfrac{1}{\omega^{n+j}} = o\left(n^{k-1}\right)$, car les pôles de $\mathcal{P}$ sont des racines de l'unité, donc de module 1.
Par conséquent,

$$u_n \underset{n\to\infty}{\sim} \alpha\frac{n^{k-1}}{(k-1)!}$$

Reste à calculer α.
Pour cela, on multiplie $f(X)$ par $(1-X)^k$ et on substitue 1 à X :[135]

$$(1-X)^k f(X) = \prod_{i=1}^{k}\frac{1-X}{1-X^{a_i}} = \prod_{i=1}^{k}\frac{1}{1+X+\ldots+X^{a_i-1}}$$

$$\alpha + 0 = \prod_{i=1}^{k}\frac{1}{a_i}$$

D'où :

$$u_n \underset{n\to\infty}{\sim} \frac{1}{a_1\dots a_k}\frac{n^{k-1}}{(k-1)!}$$

■

Références

[X-ENS An2] S. FRANCINOU, H. GIANELLA et S. NICOLAS – *Oraux X-ENS Analyse 2*, 2e éd., Cassini, 2009.

135. Pour le membre de gauche, on effectue la substitution dans l'expression obtenue à la fin de l'étape 2. Dans le cadre des séries formelles, on rappelle qu'on ne peut substituer à X que des séries formelles de valuation non-nulle ; on ne peut donc pas se placer dans ce cadre pour effectuer cette substitution.

Diffusion à titre gratuit uniquement.

$\mathrm{SO}_3(\mathbb{R})$ est simple, mais pas seulement

Leçons : 161, 204, *101, 103, 106, 108, 160, 203*

[H2G2], partie VII.A
[Per], partie VI.2

Théorème

$\mathrm{SO}_3(\mathbb{R})$ est un groupe simple, connexe et compact.

Démonstration :

→ Montrons que $\mathrm{SO}_3(\mathbb{R})$ est compact. [136]

On a $\mathrm{SO}_3(\mathbb{R}) = \psi^{-1}(\{\mathrm{I}_3\}) \cap \det^{-1}(\{1\})$ est fermé dans $\mathcal{M}_3(\mathbb{R})$, où $\psi : \left| \begin{array}{ccc} \mathcal{M}_3(\mathbb{R}) & \to & \mathcal{M}_3(\mathbb{R}) \\ M & \mapsto & {}^tMM \end{array} \right.$ est une application continue.
Aussi, $\mathrm{O}_3(\mathbb{R})$ est borné car ses éléments sont des isométries ; donc $\mathrm{SO}_3(\mathbb{R})$ est borné.
Ainsi, comme $\mathcal{M}_3(\mathbb{R})$ est de dimension finie, $\mathrm{SO}_3(\mathbb{R})$ est compact.

→ Aussi, $\mathrm{SO}_3(\mathbb{R})$ est connexe (par arcs) ; on va montrer qu'on peut relier continûment ses éléments à I_3.
Soit $M \in \mathrm{SO}_3(\mathbb{R})$, on dispose du résultat de réduction :

$$\exists P \in \mathrm{O}_3(\mathbb{R}), M = PU_\theta P^{-1}, \text{ où } U_\theta = \left(\begin{array}{c|cc} 1 & 0 & 0 \\ \hline 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{array}\right) \text{ avec } \theta \in \mathbb{R}.$$

Soit alors $\gamma : \left| \begin{array}{ccc} [0,1] & \to & \mathrm{SO}_3(\mathbb{R}) \\ t & \mapsto & PU_{t\theta}P^{-1} \end{array} \right.$; γ est un chemin continu reliant M à I_3, et restant dans $\mathrm{SO}_3(\mathbb{R})$.
Donc $\mathrm{SO}_3(\mathbb{R})$ est connexe (par arcs).

→ On va maintenant montrer que $\mathrm{SO}_3(\mathbb{R})$ est simple ; pour cela, soit $H \triangleleft \mathrm{SO}_3(\mathbb{R})$, non-réduit à $\{\mathrm{I}_3\}$. Voilà ce qu'on va faire : on va montrer que les retournements de $\mathbb{R}^3$ sont tous conjugués dans $\mathrm{SO}_3(\mathbb{R})$, puis que H en contient un ; ainsi H les contiendra tous, et comme ils engendrent $\mathrm{SO}_3(\mathbb{R})$ [137], on aura $H = \mathrm{SO}_3(\mathbb{R})$, d'où la simplicité de $\mathrm{SO}_3(\mathbb{R})$. Commençons par le lemme suivant.

Lemme

$\mathrm{SO}_3(\mathbb{R})$ agit transitivement sur l'ensemble des droites de $\mathbb{R}^3$.

Démonstration :
Soient D, D' deux droites de $\mathbb{R}^3$, engendrées par les vecteurs unitaires d et d'.

136. Notons que la connexité et la compacité de $\mathrm{SO}_n(\mathbb{R})$ se montre exactement de la même façon, pour tout $n \in \mathbb{N}^*$.

137. Soit $n \in \mathbb{N}^*$, on va d'abord montrer que les réflexions orthogonales engendrent $\mathrm{O}_n(\mathbb{R})$.
Un rappel : une réflexion orthogonale, c'est une matrice diagonalisable de spectre $\{-1, 1\}$, avec -1 de multiplicité 1.
Soit $u \in \mathrm{O}_n(\mathbb{R})$ et $F_u = \{x \in E | u(x) = x\}$ l'espace de ses points fixes. On pose $p_u = n - \dim F_u$ et on va en fait montrer que u est produit d'au plus p_u réflexions orthogonales.
On raisonne par récurrence sur $p_u \in \mathbb{N}$. Le cas $p_u = 0$ est trivial puisqu'il correspond à $u = \mathrm{I}_n$.
Supposons donc $p_u > 0$; soit $x \in F_u^\perp \backslash \{0\}$, et soit $y = u(x)$. Comme $x \notin F_u$, $y \neq x$; et comme F_u et $F_u^\perp$ sont u-stables, $y \in F_u^\perp$. De plus, $\langle x - y, x + y\rangle = \|x\|^2 - \|y\|^2 = 0$ (car u est une isométrie), donc $x - y$ et $x + y$ sont orthogonaux. Soit alors τ la réflexion orthogonale associée au vecteur $x - y$ (ie telle que $E_{-1} = \mathrm{Vect}\{x - y\}$). On a donc : $\tau(x - y) = y - x$ et $\tau(x + y) = x + y$, d'où, par demi-différence : $\tau(u(x)) = \tau(y) = x$. Aussi, $x - y \in F_u^\perp$, ce qui implique sur $\tau_{|F_u} = \mathrm{Id}_{F_u}$. En conséquence, $F_u \subset F_{\tau u}$; mais $x \in F_{\tau u} \backslash F_u$, donc $p_u > p_{\tau u}$. On utilise donc notre hypothèse de récurrence sur τu : $\tau u = \tau_1 \dots \tau_r$, où les τ_i sont des réflexions orthogonales et $r \leqslant p_{\tau u}$. Mais alors on a $u = \tau\tau_1 \dots \tau_r$ et $r + 1 \leqslant p_u$, ce qui achève la récurrence.
Désormais, soit $u \in \mathrm{SO}_n(\mathbb{R})$.
Pour $n = 3$, on conclut alors que les retournements engendrent $\mathrm{SO}_3(\mathbb{R})$: si $u \neq I_3$, alors $u = \tau_1\tau_2 = (-\tau_1)(-\tau_2)$ et les opposés des réflexions orthogonales sont ici des retournements.
Quand $n \geqslant 3$, il y a encore un peu de travail ; on peut déjà écrire $u = \tau_1 \dots \tau_{2p}$ avec $2p \leqslant n$, les τ_i étant des réflexions orthogonales. Prenons une paire de réflexions orthogonales τ_1, τ_2 ; alors on peut trouver une paire de retournements σ_1, σ_2 telle que : $\tau_1\tau_2 = \sigma_1\sigma_2$. En effet : soient H_1 et H_2 les hyperplans laissés fixes par τ_1 et τ_2 et soit V un sev de $H_1 \cap H_2$ qui soit de dimension $n - 3$. Alors $\tau_1\tau_{2|V} = \mathrm{Id}$ et donc $\tau_1\tau_2\left(V^\perp\right) \subset V^\perp$. Mais d'après le cas $n = 3$, on peut écrire $\tau_1\tau_{2|V^\perp} = \sigma_1\sigma_2$, où les σ_i sont des retournements de $V^\perp$. Il ne reste qu'à les prolonger par l'identité sur V, et on a gagné.

Diffusion à titre gratuit uniquement.

Soient (e_1, e_2) et (e_1', e_2') des bases orthonormales de $D^\perp$ et de $D'^\perp$.
On a donc construit deux bases orthonormales de $\mathbb{R}^3$; la matrice de passage P de l'une à l'autre est donc dans $\mathrm{O}_3(\mathbb{R})$.
Mais quitte à changer d' en $-d'$, on peut supposer que $P \in \mathrm{SO}_3(\mathbb{R})$. ■

Soient R_D et $R_{D'}$ deux retournements de $\mathbb{R}^3$ d'axes respectifs D et D'.
Par le lemme, il existe $S \in \mathrm{SO}_3(\mathbb{R})$ envoyant D sur D'.
Alors $SR_DS^{-1} \in \mathrm{SO}_3(\mathbb{R})$ est semblable à R_D ; c'est donc un retournement de $\mathbb{R}^3$.
Soit $x \in D'$, on a : $SR_D \underbrace{S^{-1}x}_{\in D} = SS^{-1}x = x$; ainsi l'axe de SR_DS^{-1} est D', id est : $SR_DS^{-1} = R_{D'}$. On a ainsi montré que tous les retournements sont conjugués dans $\mathrm{SO}_3(\mathbb{R})$.
Reste à trouver un retournement dans H.
Soit $h \in H$ avec $h \neq \mathrm{I}_3$. On pose : $\varphi : \left| \begin{array}{ccc} \mathrm{SO}_3(\mathbb{R}) & \to & \mathbb{R} \\ g & \mapsto & \mathrm{tr}\left(ghg^{-1}h^{-1}\right) \end{array} \right.$.
φ est une application continue, donc $\varphi\left(\mathrm{SO}_3(\mathbb{R})\right)$ est un compact connexe de $\mathbb{R}$, un segment.
$\forall g \in \mathrm{SO}_3(\mathbb{R}), ghg^{-1}h^{-1} \in \mathrm{SO}_3(\mathbb{R})$ donc $\varphi(g)$ est de la forme $1 + 2\cos\theta$, avec $\theta \in \mathbb{R}$ (par le théorème de réduction des éléments de $\mathrm{SO}_3(\mathbb{R})$).
Comme en plus $\varphi\left(\mathrm{I}_3\right) = 3$, on en déduit que $\varphi\left(\mathrm{SO}_3(\mathbb{R})\right) = [a, 3]$, pour un certain réel a.
Par l'absurde, supposons que $a = 3$.
Alors $\forall g \in \mathrm{SO}_3(\mathbb{R}), \mathrm{tr}\left(ghg^{-1}h^{-1}\right) = 3$, et donc $ghg^{-1}h^{-1} = \mathrm{I}_3$.
En conséquence, $h \in \mathrm{Z}\left(\mathrm{SO}_3(\mathbb{R})\right) = \{\mathrm{I}_3\}$, ce qui est exclu. [138]
On a donc bien $a < 3$.
La suite $\left(1 + 2\cos\dfrac{\pi}{n}\right)_{n\in\mathbb{N}}$ ayant 3 pour limite, on peut prendre $n \in \mathbb{N}^*$ tel que $a < 1 + 2\cos\dfrac{\pi}{n} < 3$.
Soit alors $g_n \in \mathrm{SO}_3(\mathbb{R})$ tel que $\varphi\left(g_n\right) = 1 + 2\cos\dfrac{\pi}{n}$.
On pose $h_n = g_n h g_n^{-1} h^{-1}$; $h_n \in H$, car H est conjugué dans $\mathrm{SO}_3(\mathbb{R})$ et car $h \in H$.
Comme $\mathrm{tr}\ h_n = 1 + 2\cos\dfrac{\pi}{n}$, on obtient que h_n est une rotation d'angle $\pm\dfrac{\pi}{n}$; ainsi $h_n^n \in H$ est une rotation d'angle π, autrement dit, un retournement. Ce qui conclut la preuve. ■

Références

[H2G2] P. CALDERO et J. GERMONI – *Histoires hédonistes de groupes et de géométries*, Calvage & Mounet, 2013.
[Per] D. PERRIN – *Cours d'algèbre*, Ellipses, 1996.

138. En effet, soit $n \geqslant 2$, et $h \in \mathrm{Z}\left(\mathrm{SO}_n(\mathbb{R})\right)$; on va montrer que h est une homothétie.
Soit D une droite de $\mathbb{R}^n$; on a : $R_{h(D)} = hR_Dh^{-1} = R_D$, car h est central.
Donc h laisse stables toutes les droites de $\mathbb{R}^n$, c'est donc une homothétie (c'est facile : on prend deux vecteurs non-colinéaires, ce sont des vecteurs propres de h, leur somme également, et on montre que tous les vecteurs sont de même valeur propre.)

Diffusion à titre gratuit uniquement.

Sous-groupes compacts de $\mathrm{GL}_n(\mathbb{R})$

Leçons : 150, 181, 203, 206, *101, 106, 208*

[Ale], problème III.III.A.1

Théorème

Soit V un $\mathbb{R}$-espace vectoriel de dimension finie, et soit K un convexe compact non-vide de V.
Soit G un sous-groupe compact de $\mathrm{GL}(V)$ vérifiant : $\forall u \in G, u(K) \subset K$.
Alors $\exists x \in K, \forall u \in G, u(x) = x$.

Démonstration :

Étape 1 : Soit N une norme euclidienne sur V.
Pour tout $x \in V$, on pose

$$\nu(x) = \max_{u \in G} N(u(x)).$$

On va montrer que ν définit une norme G-invariante sur V.
Comme G est compact : pour tout $x \in V$, $\{u(x) | u \in G\}$ est compact.
On en déduit que ν est bien définie sur V.
Par ailleurs, on a : $\forall x \in V, \forall u \in G, \nu(x) = \nu(u(x))$, car la composition par u est une bijection du groupe G.
De plus :
– ν est à valeurs dans $\mathbb{R}^+$ car N l'est ;
– $\forall x \in V, \nu(x) = 0 \Rightarrow N(x) = 0 \Rightarrow x = 0$;
– $\forall x \in V, \forall \lambda \in \mathbb{R}, \nu(\lambda x) = |\lambda| \nu(x)$ car les éléments de G sont linéaires et car N est une norme.
Il nous reste à montrer que ν vérifie l'inégalité triangulaire ; soient $x, y \in V$.
$\nu(x+y)$ étant définie à partir d'un maximum, on a :

$$\exists u_0 \in G, \nu(x+y) = N\left(u_0(x+y)\right).$$

Alors $\nu(x+y) \leqslant N\left(u_0(x)\right) + N\left(u_0(y)\right) \leqslant \nu(x) + \nu(y)$.
Et si on a l'égalité $\nu(x+y) = \nu(x) + \nu(y)$, alors $u_0(x)$ et $u_0(y)$ sont positivement liés, donc x et y aussi (car u_0 est inversible).

Étape 2 : Comme ν est continue sur K, elle y admet un minimum, disons en $a \in K$.
Soit $u \in G$, par argument d'invariance, on a : $\nu(u(a)) = \nu(a)$.
Comme ν est convexe (c'est une norme !), ses ensembles de niveau sont convexes, et donc, on obtient : $\nu\left(\dfrac{u(a)+a}{2}\right) = \nu(a)$.
Ainsi, $\nu(u(a)+a) = 2\nu(a) = \nu(a) + \nu(u(a))$.
Le cas d'égalité dans l'inégalité triangulaire pour ν fournit : $u(a) = \lambda a$, pour un certain $\lambda > 0$.
Mais $\nu(u(a)) = \nu(a)$, donc $\lambda = 1$ ou $a = u(a) = 0$.
Finalement, $\forall u \in G, u(a) = a$. ■

Corollaire

Soit G un sous-groupe compact de $\mathrm{GL}_n(\mathbb{R})$.
Alors il existe une forme quadratique q définie positive sur $\mathbb{R}^n$ telle que $G \subset O(q)$. [139]

Démonstration :
On munit G d'une nouvelle structure de groupe $(G, \Diamond)$ par : $\forall A, B \in G, A \Diamond B := BA$.
On pose :

$$\rho : \left| \begin{array}{ccc} (G, \Diamond) & \to & \mathrm{GL}\left(\mathcal{S}_n(\mathbb{R})\right) \\ A & \mapsto & (S \mapsto {}^tASA) \end{array} \right. .$$

– ρ est bien définie car $\forall A \in G, \rho(A) \in \mathcal{L}\left(\mathcal{S}_n(\mathbb{R})\right)$ est inversible, d'inverse $\rho\left(A^{-1}\right)$;
– ρ est un morphisme de groupes (pour la loi $\Diamond$) ;

139. Ce résultat peut-être démontré de façon différente en utilisant l'ellipsoïde de John-Loewner (voir en page 105).

Diffusion à titre gratuit uniquement.

– ρ est continue, car $\rho = (b \circ \Delta)|_G$, où $b : \left| \begin{array}{ccc} \mathcal{M}_n(\mathbb{R})^2 & \to & \mathcal{L}(\mathcal{S}_n(\mathbb{R})) \\ (A,B) & \mapsto & (S \mapsto {}^tASB) \end{array} \right.$ est continue (par bilinéarité et dimension finie) et $\Delta : \left| \begin{array}{ccc} \mathcal{M}_n(\mathbb{R}) & \to & \mathcal{M}_n(\mathbb{R})^2 \\ A & \mapsto & (A,A) \end{array} \right.$ est continue (par linéarité et dimension finie).

$\rho(G)$ est un sous-groupe (car ρ est un morphisme de groupes et G un groupe) compact (car ρ est continue et G compact) de $\mathrm{GL}(\mathcal{S}_n(\mathbb{R}))$.
On pose $H = \{{}^tMM | M \in G\}$ et K l'enveloppe convexe de H.
La compacité de G implique celle de H puis celle de K.[140]
De plus, comme $G \subset \mathrm{GL}_n(\mathbb{R})$, on obtient $H \subset \mathcal{S}_n^{++}(\mathbb{R})$; et comme $\mathcal{S}_n^{++}(\mathbb{R})$ est convexe, on a $K \subset \mathcal{S}_n^{++}(\mathbb{R})$.
Enfin,

$$\forall A \in G, \forall M \in G, \rho(A)\left({}^tMM\right) = {}^tA{}^tMMA = {}^t(MA)(MA) \in H \subset K ;$$

et donc par linéarité de $\rho(A)$, K est stable par $\rho(A)$.
On applique le résultat précédent pour obtenir :

$$\exists S \in K, \forall A \in G, S = \rho(A)(S) = {}^tASA.$$

Et comme $K \subset \mathcal{S}_n^{++}(\mathbb{R})$, on a bien $G \subset O(q_S)$, où $q_S : x \mapsto {}^txSx$ est une forme quadratique définie positive sur $\mathbb{R}^n$. ■

Références

[Ale] M. ALESSANDRI – *Thèmes de géométrie, Groupes en situation géométrique*, Dunod, 1999.

140. On dispose du lemme suivant, conséquence du théorème de Carathéodory :

Lemme

Soit $\mathcal{E}$ un espace affine réel de dimension $n < \infty$, $\mathcal{A} \subset \mathcal{E}$, on suppose que $\mathcal{A} \neq \varnothing$ et que $\mathcal{A}$ est compact.
Alors son enveloppe convexe $\mathrm{Cv}(\mathcal{A})$ est compacte.

En effet, on pose $K = \left\{(t_1, \ldots, t_{n+1}) \in [0,1]^{n+1} | t_1 + \ldots + t_{n+1} = 1\right\}$, c'est un compact de $\mathbb{R}^{n+1}$.
On définit $f : \left| \begin{array}{ccc} K \times \mathcal{E}^{n+1} & \to & \mathcal{E} \\ (t_1, \ldots, t_{n+1}, A_1, \ldots, A_{n+1}) & \mapsto & t_1A_1 + \ldots + t_{n+1}A_{n+1} \end{array} \right.$.
D'après Carathéodory, $f\left(K \times \mathcal{A}^{n+1}\right) = \mathrm{Cv}(\mathcal{A})$; or f est continue et $K \times \mathcal{A}^{n+1}$ est compact donc $\mathrm{Cv}(\mathcal{A})$ l'est aussi.

Diffusion à titre gratuit uniquement.

Surjectivité de l'exponentielle

Leçons : 156, 204, *153, 214, 215*

[Zav], problème 9.II

Théorème

Soit $n \in \mathbb{N}^*$, on a : $\forall A \in \mathrm{GL}_n(\mathbb{C}), \exists P \in \mathbb{C}[X], A = \exp(P(A))$.
Ce résultat implique la surjectivité de la fonction exp : $\mathcal{M}_n(\mathbb{C}) \to \mathrm{GL}_n(\mathbb{C})$.

Démonstration :

Étape 1 : Rappelons que $\forall A \in \mathcal{M}_n(\mathbb{C}), \exp(A) \in \mathbb{C}[A]$.[141]

Désormais, fixons $A \in \mathcal{M}_n(\mathbb{C})$.

Étape 2 : On va montrer que $\mathbb{C}[A]^\times = \mathbb{C}[A] \cap \mathrm{GL}_n(\mathbb{C})$.

$\subset$ **:** Trivial.

$\supset$ **:** Soit $B \in \mathbb{C}[A] \cap \mathrm{GL}_n(\mathbb{C})$, dont on note μ_B le polynôme minimal.
Si $X | \mu_B$, alors $\mu_B = XQ$, avec $Q \in \mathbb{C}[X]$, puis $0 = BQ(B)$; mais $B \in \mathrm{GL}_n(\mathbb{C})$ donc $Q(B) = 0$.
Ceci contredit la minimalité de μ_B.
Donc $\mu_B = \alpha + XQ$, avec $\alpha \in \mathbb{C}^*$ et $Q \in \mathbb{C}[X]$; alors $0 = \alpha + BQ(B)$ et donc

$$B^{-1} = \frac{-Q(B)}{\alpha} \in \mathbb{C}[B] \subset \mathbb{C}[A].$$

Étape 3 : On a donc :
- $\exp(\mathbb{C}[A]) \subset \mathbb{C}[A] \cap \mathrm{GL}_n(\mathbb{C}) = \mathbb{C}[A]^\times$.
- $\forall M, N \in \mathbb{C}[A], \exp(M+N) = \exp(M)\exp(N)$ car M et N commutent.

Étape 4 : Montrons que $\mathbb{C}[A]^\times$ est un ouvert connexe de $\mathbb{C}[A]$.
$\mathrm{GL}_n(\mathbb{C})$ étant ouvert[142] dans $\mathcal{M}_n(\mathbb{C})$, on déduit de $\mathbb{C}[A]^\times = \mathbb{C}[A] \cap \mathrm{GL}_n(\mathbb{C})$ que $\mathbb{C}[A]^\times$ est ouvert dans $\mathbb{C}[A]$.
On va montrer que $\mathbb{C}[A]^\times$ est connexe par arcs ; soient $M, N \in \mathbb{C}[A]^\times$.[143]

141. En effet, $\exp(A) = \lim_{n\to\infty} \underbrace{\sum_{k=0}^{n} \frac{A^k}{k!}}_{\in \mathbb{C}[A]}$ et $\mathbb{C}[A]$ est un sous-espace vectoriel de $\mathcal{M}_n(\mathbb{C})$ donc de dimension finie et donc fermé.

142. $\mathrm{GL}_n(\mathbb{C}) = \det^{-1}(\mathbb{R}^*)$ et le déterminant est continu.

143. Petit rappel sur la connexité.

Lemme

1. L'image continue d'un connexe de (E,d) dans (F,d') est connexe.
2. Un espace métrique (E,d) est connexe si et seulement si toute application continue f de (E,d) dans $(\{0,1\},\delta)$ où δ est la distance discrète est constante.
3. Un espace métrique connexe par arcs est connexe.

En effet :

1. Soit A une partie ouverte et fermée de $f(E)$. Comme f est continue, $f^{-1}(A)$ est ouverte et fermée dans E. Par connexité de E, on a :
   - soit $f^{-1}(A) = \varnothing$ et alors il n'existe pas de $x \in E$ tel que $f(x) \in A$ et donc $A = f(E) \cap A = \varnothing$;
   - soit $f^{-1}(A) = E$ et alors $\forall x \in E, f(x) \in A$ donc $A = f(E)$.

   Donc $f(E)$ est connexe.
2. - Si E est connexe, et $f : E \to \{0,1\}$ est continue, alors $f(E)$ est connexe, donc $f(E) \neq \{0,1\}$ (car $\{0,1\}$ est une réunion de deux singletons, donc de deux fermés). Ainsi, f est constante.
   - Supposons que E ne soit pas connexe : $E = F_0 \sqcup F_1$, où F_0 et F_1 sont deux fermés non-vides. On définit f par $f|_{F_0} \equiv 0$ et $f|_{F_1} \equiv 1$. On va montrer que f est continue ; soit F un fermé de $\{0,1\}$, on va montrer que $f^{-1}(F)$ est fermé dans E. Si $F = \{0,1\}$, alors $f^{-1}(F) = E$ est fermé. Si $F = \{0\}$, alors $f^{-1}(F) = F_0$ est fermé. Si $F = \{1\}$, alors $f^{-1}(F) = F_1$ est fermé. Si $F = \varnothing$, alors $f^{-1}(F) = \varnothing$ est fermé.
3. Soit $f : E \to \{0,1\}$ une fonction continue ; on va montrer que f est constante. Soient $x, y \in E$. Comme E est connexe par arcs, il existe un chemin continu $\gamma : [0,1] \to E$ tel que $\gamma(0) = x$ et $\gamma(1) = y$. L'application $f \circ \gamma : [0,1] \to \{0,1\}$ est alors continue ; mais $[0,1]$ étant connexe, $f \circ \gamma$ est donc constante. Alors : $f(y) = f(\gamma(1)) = f(\gamma(0)) = f(x)$.

Diffusion à titre gratuit uniquement.

On a : $\det(zM + (1-z)N) \in \mathbb{C}[z]\backslash\{0\}$, car le déterminant est polynomial et M inversible.
On note Z l'ensemble des racines de ce polynôme ; c'est un ensemble fini, donc $\mathbb{C}\backslash Z$ est connexe par arcs, contient 0 et 1 ; donc il existe un chemin γ continu reliant 0 et 1 dans $\mathbb{C}\backslash Z$.
Alors le chemin $t \mapsto \gamma(t)M + (1-\gamma(t))N$ est continu et relie M et N dans $\mathbb{C}[A]^\times$.

Étape 5 : Montrons que $\exp(\mathbb{C}[A])$ est ouvert.
On sait que exp est de classe $\mathcal{C}^1$ sur $\mathbb{C}[A]$ et que $\mathrm{D}(\exp)(0) = \mathrm{Id}_{\mathbb{C}[A]}$ est bijective.
Par le théorème d'inversion locale, il existe $\mathcal{U}$ un voisinage ouvert de 0 dans $\mathbb{C}[A]$, $\mathcal{V}$ un voisinage ouvert de I_n dans $\mathbb{C}[A]^\times$ tels que : $\exp : \mathcal{U} \to \mathcal{V}$ soit un $\mathcal{C}^1$-difféomorphisme.
En particulier, $\mathcal{V} \subset \exp(\mathbb{C}[A])$.
Soit $B \in \mathbb{C}[A]$, on a : $\exp(B+\mathcal{U}) = \exp(B)\mathcal{V}$. [144]
Or $\exp(B) \in \mathrm{GL}_n(\mathbb{C})$ donc la multiplication à gauche par $\exp(B)$ est bicontinue, par conéquent $\exp(B)\mathcal{V}$ est un ouvert, donc un voisinage de $\exp(B) \in \mathbb{C}[A]^\times$.
Mais $\exp(B)\mathcal{V} = \exp(U+\mathcal{U}) \subset \exp(\mathbb{C}[A])$, donc $\exp(\mathbb{C}[A])$ est voisinage de chacun de ses points, donc ouvert.

Étape 6 : Montrons que $\exp(\mathbb{C}[A])$ est fermé dans $\mathbb{C}[A]^\times$.
On a : $\mathbb{C}[A]^\times \setminus \exp(\mathbb{C}[A]) = \bigcup_{M \in \mathbb{C}[A]^\times \setminus \exp(\mathbb{C}[A])} M\exp(\mathbb{C}[A])$. [145]
Comme dans l'étape précédente, on montre que $\forall M \in \mathbb{C}[A]^\times \setminus \exp(\mathbb{C}[A]), M\exp(\mathbb{C}[A])$ est ouvert (car M est inversible et $\exp(\mathbb{C}[A])$ est ouvert).
Donc $\exp(\mathbb{C}[A])$ est fermé dans $\mathbb{C}[A]^\times$.

Comme $\mathbb{C}[A]^\times$ est connexe et comme $\exp(\mathbb{C}[A])$ est ouvert, fermé, non-vide dans $\mathbb{C}[A]^\times$, on a :

$$\mathbb{C}[A]^\times = \exp(\mathbb{C}[A]).$$

Conséquemment, $\forall A \in \mathrm{GL}_n(\mathbb{C}), \exists P \in \mathbb{C}[X], A = \exp(P(A))$. D'où la surjectivité de l'exponentielle. [146] ■

Références

[Zav] M. ZAVIDOVIQUE – *Un max de maths*, Calvage & Mounet, 2013.

144. Cela se fait très bien par double inclusion. Si $M \in \mathcal{U}$, $\exp(B+M) = \exp(B)\exp(M) \in \exp(B)\mathcal{V}$. Si $N \in \mathcal{V}$, alors $\exists M \in \mathcal{U}, N = \exp(M)$ et $\exp(B)N = \exp(B)\exp(M) = \exp(B+M) \in \exp(B+\mathcal{U})$.

145. Là encore, cette égalité se vérifie simplement par double-inclusion. L'inclusion "$\subset$" découle du fait que $\mathrm{I}_n \in \exp(\mathbb{C}[A])$. L'inclusion "$\supset$" demande un peu plus de rédaction : si $M \in \mathbb{C}[A]^\times \setminus \exp(\mathbb{C}[A])$ et si $N \in M\exp(\mathbb{C}[A])$, alors $M \in N\exp(\mathbb{C}[A])$. Supposons que $N \in \exp(\mathbb{C}[A])$, alors on aurait $M \in \exp(\mathbb{C}[A])$, ce qui est exclu.

146. Citons un corollaire pour terminer.

Corollaire

On a l'égalité ensembliste : $\exp(\mathcal{M}_n(\mathbb{R})) = \left\{A^2 \middle| A \in \mathrm{GL}_n(\mathbb{R})\right\}$.

En effet :

$\subset$ **:** Soit $M \in \mathcal{M}_n(\mathbb{R})$, on a : $\exp(M) = \underbrace{\exp\left(\frac{M}{2}\right)^2}_{\in \mathrm{GL}_n(\mathbb{R})}$.

$\supset$ **:** Soit $A \in \mathcal{M}_n(\mathbb{R})$, telle que $A = C^2$ où $C \in \mathrm{GL}_n(\mathbb{R})$; par le théorème : $\exists P \in \mathbb{C}[X], C = \exp(P(C))$.
Et comme $C \in \mathcal{M}_n(\mathbb{R}), C = \overline{C} = \exp\left(\overline{P}(C)\right)$. (Non, là, franchement, débrouillez-vous pour le détail, j'ai la flemme.)
Puis : $A = C^2 = C\overline{C} = \exp\left(P(C) + \overline{P}(C)\right)$ car $P(C)$ et $\overline{P}(C)$ commutent.
Or $P + \overline{P} \in \mathbb{R}[X]$ donc $P(C) + \overline{P}(C) \in \mathcal{M}_n(\mathbb{R})$.

Diffusion à titre gratuit uniquement.

Théorème des extrema liés [147, 148]

Leçons : 151, 159, 214, 215, 219

[Gou An], partie 5.3.2

Théorème

Soient $f, g_1, \dots, g_r : U \to \mathbb{R}$ des fonctions de classe $\mathcal{C}^1$ sur un ouvert U de $\mathbb{R}^n$.
On pose $\Gamma = \{x \in U | \forall i \in [\![1, r]\!], g_i(x) = 0\}$.
On suppose que :
- $f|_\Gamma$ admet un extremum local en $a \in \Gamma$;
- les formes linéaires $\mathrm{D}g_1(a), \dots, \mathrm{D}g_r(a)$ sont linéairement indépendantes.

Alors il existe $\lambda_1, \dots, \lambda_r \in \mathbb{R}$, appelés multiplicateurs de Lagrange, tels que : $\mathrm{D}f(a) = \sum_{i=1}^{r} \lambda_i \mathrm{D}g_i(a)$.

147. On peut aussi écrire "extremums" ou "extrémums", mais pas "extréma", "extremas" ou "extrémas".
148. Donnons de ce théorème quelques applications.

Le théorème spectral : Soit E un espace euclidien, $u \in \mathcal{L}(E)$ un endomorphisme symétrique (c'est-à-dire tel que $u^* = u$) ; alors il existe une base orthonormée de E formée de vecteurs propres de u. On considère les applications différentiables :

$$f : \left| \begin{array}{ccc} E & \to & \mathbb{R} \\ x & \mapsto & \langle u(x), x \rangle \end{array} \right. \quad \text{et } g : \left| \begin{array}{ccc} E & \to & \mathbb{R} \\ x & \mapsto & \langle x, x \rangle \end{array} \right. .$$

On note aussi $\mathcal{S} = \{x \in E | g(x) = 1\}$ la sphère unité ; on est en dimension finie, donc elle est compacte. L'application f étant continue sur E, elle atteint son maximum sur $\mathcal{S}$, en un point noté e_1. Par ailleurs, pour $x \in E$ et $h \in E$:

$$\mathrm{D}f(x).h = 2\langle u(x), h \rangle \text{ et } \mathrm{D}g(x).h = 2\langle x, h \rangle.$$

D'après le théorème des extrema liés, $\exists \lambda_1 \in \mathbb{R}, \mathrm{D}f(e_1) = \lambda_1 \mathrm{D}g(e_1)$. Autrement dit, $u(e_1) = \lambda_1 e_1$. On peut à présent raisonner par récurrence ; soit $F = e_1^\perp$. Il suffit de montrer que $u_{|F}$ est symétrique et appliquer la récurrence.

L'inégalité entre les moyennes arithmétique et géométrique : On considère les applications :

$$f : \left| \begin{array}{ccc} (\mathbb{R}^+)^n & \to & \mathbb{R} \\ x & \mapsto & \sqrt[n]{x_1 \dots x_n} \end{array} \right. \quad \text{et } g : \left| \begin{array}{ccc} \mathbb{R}^n & \to & \mathbb{R} \\ x & \mapsto & \frac{1}{n}\sum_{i=1}^n x_i - 1 \end{array} \right. .$$

L'ensemble $K = \{x \in \mathbb{R}^n | x_1 \geqslant 0, \dots x_n \geqslant 0, g(x) = 0\}$ est un compact de $\mathbb{R}^n$. La fonction f atteint son maximum sur K en un point $a = (a_1, \dots, a_n)$; notamment $f(a) \geqslant f(1, \dots, 1) = 1$. En conséquence, $a \in (\mathbb{R}^{+*})^n$, ouvert sur lequel f et g sont de classe $\mathcal{C}^1$. Par ailleurs, pour $x \in (\mathbb{R}^{+*})^n$, $\mathrm{D}g(x) = \frac{1}{n}(1, \dots, 1)$. D'après le théorème des extrema liés, $\exists \lambda \in \mathbb{R}, \mathrm{D}f(a) = \lambda \mathrm{D}g(a)$. On montre alors que $\frac{\partial f}{\partial x_i}(a) = \frac{f(a)}{n a_i}$ pour $i \in [\![1, n]\!]$. En conséquence, $f(a) = \lambda a_1 = \dots = \lambda a_n$, puis $a_1 = \dots = a_n = 1$. Ainsi, $\forall x \in K, f(x) \leqslant 1$; puis, par homogénéité : $\forall x_1, \dots, x_n \in \mathbb{R}^+, \sqrt[n]{x_1 \dots x_n} \leqslant \frac{x_1 + \dots + x_n}{n}$.

Une utilisation en statistiques : Soit $X_1, \dots, X_n$ un n-échantillon de moyenne ν et de variance σ^2.

Quand $\sum_{i=1}^n a_i = 1$, on sait que $\sum_{i=1}^n a_i X_i$ est un estimateur non-biaisé de ν. Parmi tous les estimateurs de cette forme, on recherche celui de variance minimale.

Par le calcul, la variance de l'estimateur $\sum_{i=1}^n a_i X_i$ vaut $\sigma^2 \sum_{i=1}^n a_i^2$.

On pose :

$$f : \left| \begin{array}{ccc} \mathbb{R}^n & \to & \mathbb{R} \\ x & \mapsto & x_1^2 + \dots + x_n^2 \end{array} \right. \quad \text{et } g : \left| \begin{array}{ccc} \mathbb{R}^n & \to & \mathbb{R} \\ x & \mapsto & x_1 + \dots + x_n - 1 \end{array} \right. .$$

On veut minimiser f sur l'ensemble où g s'annule.
Par le théorème des extrema liés, si f admet un extremum en $(a_1, \dots, a_n)$, alors

$$\exists \lambda \in \mathbb{R}, \mathrm{D}f(a_1, \dots, a_n) = \lambda \mathrm{D}g(a_1, \dots, a_n).$$

Mais $\frac{\partial f}{\partial x_i}(a_1, \dots, a_n) = 2a_i$ et $\frac{\partial g}{\partial x_i}(a_1, \dots, a_n) = 1$.
Donc $\forall i \in [\![1, n]\!], 2a_i = 1$.
On montre alors que $a_1 = \dots = a_n = \frac{1}{n}$; c'est bien un minimum, car $f(1, 0, \dots, 0) = 1 > n\frac{1}{n^2} = \frac{1}{n}$.

Diffusion à titre gratuit uniquement.

Démonstration :

Étape 1 : Commençons par quelques petites remarques.
On voit déjà que nécessairement $r \leqslant n$ car les formes linéaires $\mathrm{D}g_1(a), \dots, \mathrm{D}g_r(a)$ forment une famille libre de $(\mathbb{R}^n)^*$.
De plus, si $r = n$, le résultat est trivial, car alors $\mathrm{D}g_1(a), \dots, \mathrm{D}g_r(a)$ forment une base de $(\mathbb{R}^n)^*$.
On peut donc supposer désormais que $r < n$.
Soit $s = n - r \geqslant 1$; on procède à l'identification entre $(x, y) \in \mathbb{R}^s \times \mathbb{R}^r$ et $(x_1, \dots, x_s, y_1, \dots, y_r) \in \mathbb{R}^n$.
On écrit alors $a = (\alpha, \beta)$, où $\alpha \in \mathbb{R}^s$ et $\beta \in \mathbb{R}^r$.

Étape 2 : Intéressons-nous à une matrice qui va nous être utile pour la suite. [149]
Soit $A = \begin{pmatrix} \frac{\partial g_1}{\partial x_1} & \dots & \frac{\partial g_1}{\partial x_s} & \frac{\partial g_1}{\partial y_1} & \dots & \frac{\partial g_1}{\partial y_r} \\ \vdots & & \vdots & \vdots & & \vdots \\ \frac{\partial g_r}{\partial x_1} & \dots & \frac{\partial g_r}{\partial x_s} & \frac{\partial g_r}{\partial y_1} & \dots & \frac{\partial g_r}{\partial y_r} \end{pmatrix} (a) \in \mathcal{M}_{r,n}(\mathbb{R})$. [150]
Comme $(\mathrm{D}g_i(a))_{1 \leqslant i \leqslant r}$ est une famille libre, on a : $\operatorname{rg} A = r$. [151]
On peut donc extraire de A une sous-matrice inversible de format $r \times r$; quitte à renuméroter les variables... on peut supposer que $\det \left(\frac{\partial g_i}{\partial y_j}(a) \right)_{1 \leqslant i,j \leqslant r} \neq 0$!
En notant $g = (g_1, \dots, g_r)$, ceci se reformule en "$\mathrm{D}_y g(a)$ est inversible".

Étape 3 : On applique le théorème des fonctions implicites à g au voisinage de a. [152]
Il nous fournit ici :
- U', voisinage ouvert de α dans $\mathbb{R}^s$,
- Ω, voisinage ouvert de a dans $\mathbb{R}^n$ et
- $\varphi = (\varphi_1, \dots, \varphi_r) : U' \to \mathbb{R}^r$ de classe $\mathcal{C}^1$, tels que :
$$(x \in U', (x, y) \in \Omega \text{ et } g(x, y) = 0) \Leftrightarrow (x \in U' \text{ et } y = \varphi(x)).$$

En d'autres termes, les éléments de $\Gamma \cap \Omega$ s'écrivent $(x, \varphi(x))$.

Étape 4 : On pose $h : \left| \begin{array}{ccc} U' & \to & \mathbb{R} \\ x & \mapsto & f(x, \varphi(x)) \end{array} \right.$.
Comme $h(\alpha) = f(a)$ et $\forall x \in U', (x, \varphi(x)) \in \Gamma$, ensemble où f admet un extremum local en a ; on obtient que h admet un extremum local en α.
On note $\psi = (\mathrm{Id}_{\mathbb{R}^s}, \varphi)$.

149. Si elle ne servait à rien, on n'en parlerait pas...
150. Cette notation, bâtarde, signifie
$$A = \begin{pmatrix} \frac{\partial g_1}{\partial x_1}(a) & \dots & \frac{\partial g_1}{\partial x_s}(a) & \frac{\partial g_1}{\partial y_1}(a) & \dots & \frac{\partial g_1}{\partial y_r}(a) \\ \vdots & & \vdots & \vdots & & \vdots \\ \frac{\partial g_r}{\partial x_1}(a) & \dots & \frac{\partial g_r}{\partial x_s}(a) & \frac{\partial g_r}{\partial y_1}(a) & \dots & \frac{\partial g_r}{\partial y_r}(a) \end{pmatrix}.$$
Elle a le mérite de prendre moins de place au tableau.
151. Cela se démontre par l'absurde ; supposons que $\operatorname{rg} A < r$. Alors il existe une famille de scalaires non-nulle $(\mu_1, \dots, \mu_r)$ telle que : $\forall j \in [\![1, n]\!], \sum_{i=1}^{r} \mu_i \frac{\partial g_i}{\partial z_j}(a) = 0$, où on a désigné les variables $(x_1, \dots, x_s, y_1, \dots, y_r)$ par $(z_1, \dots, z_n)$. En conséquence, $\sum_{i=1}^{r} \mu_i \mathrm{D}g_i(a) = 0$; contredisant ainsi la liberté de la famille $(\mathrm{D}g_i(a))_{1 \leqslant i \leqslant r}$.
152. Si je rappelle son énoncé, c'est pas pour vous offenser, c'est juste que ça me fait du bien.

Théorème (*des fonctions implicites*)

Soient U un ouvert de $\mathbb{R}^s \times \mathbb{R}^r$, (a, b) un point de U, et $f \in \mathcal{C}^1(U, \mathbb{R}^r)$.
On suppose que $f(a, b) = 0$ et que $\det \mathrm{D}_y f(a, b) \neq 0$.
Alors l'équation $f(x, y) = 0$ peut être résolue localement par rapport aux variables y, c'est-à-dire : il existe un voisinage ouvert V de a dans $\mathbb{R}^s$, un voisinage ouvert W de b dans $\mathbb{R}^r$, avec $V \times W \subset U$ et une unique application $\varphi : V \to W$ de classe $\mathcal{C}^1$ telle que :
$$(x \in V, y \in W \text{ et } f(x, y) = 0) \Leftrightarrow (x \in V \text{ et } y = \varphi(x)).$$
De plus, $\mathrm{D}_y f(x_0, y_0)$ est inversible pour tout $(x_0, y_0) \in V \times W$.

Diffusion à titre gratuit uniquement.

Ainsi : $\forall i \in [\![1,s]\!], 0 = \dfrac{\partial h}{\partial x_i}(\alpha) = \dfrac{\partial (f\circ\psi)}{\partial x_i}(\alpha) = \displaystyle\sum_{j=1}^{s} \frac{\partial f}{\partial x_j}(\psi(\alpha))\frac{\partial \psi_j}{\partial x_i}(\alpha) + \sum_{j=1}^{r} \frac{\partial f}{\partial y_j}(\psi(\alpha))\frac{\partial \psi_{s+j}}{\partial x_i}(\alpha)$.

Cependant, $\forall j \in [\![1,s]\!], \dfrac{\partial \psi_j}{\partial x_i} = \delta_{i,j}$ et $\forall j \in [\![1,r]\!], \dfrac{\partial \psi_{s+j}}{\partial x_i} = \dfrac{\partial \varphi_j}{\partial x_i}$.

Dès lors : $\forall i \in [\![1,s]\!], 0 = \dfrac{\partial f}{\partial x_i}(a) + \displaystyle\sum_{j=1}^{r} \frac{\partial f}{\partial y_j}(a)\frac{\partial \varphi_j}{\partial x_i}(\alpha)$.

De plus, $g\circ\psi$ est nulle sur U' donc sa $k^{\text{ème}}$ composante, $g_k\circ\psi$ (où $k \in [\![1,r]\!]$), aussi ; en conséquence :
$\forall i \in [\![1,s]\!], 0 = \dfrac{\partial g_k}{\partial x_i}(a) + \displaystyle\sum_{j=1}^{r} \frac{\partial g_k}{\partial y_i}(a)\frac{\partial \varphi_j}{\partial x_i}(\alpha)$.

Posons alors :

$$M = \left(\begin{array}{ccccccc} \frac{\partial f}{\partial x_1}(a) & \dots & \frac{\partial f}{\partial x_s}(a) & \frac{\partial f}{\partial y_1}(a) & \dots & \frac{\partial f}{\partial y_r}(a) \\ \hline & & & A & & \end{array}\right) \in \mathcal{M}_{r+1,n}(\mathbb{R}).$$

Les s premières colonnes de M sont combinaisons linéaires des r dernières ; ainsi $\operatorname{rg} M \leqslant r$.
Les $r+1$ lignes de M sont alors liées !

Ainsi, $\exists\,(\mu_0,\dots,\mu_r) \in \mathbb{R}^{r+1}\backslash\{0\}, \mu_0 \mathrm{D}f(a) + \displaystyle\sum_{i=1}^{r} \mu_i \mathrm{D}g_i(a) = 0$.

Mais $(\mathrm{D}g_i(a))_{1\leqslant i\leqslant r}$ est une famille libre, donc $\mu_0 \neq 0$ (car sinon, tous les μ_i devraient être nuls).

On obtient alors le résultat souhaité en posant : $\forall i \in [\![0,r]\!], \lambda_i = \dfrac{\mu_i}{\mu_0}$. ■

Références

[Gou An] X. GOURDON – *Les maths en tête : Analyse*, 2e éd., Ellipses, 2008.

Diffusion à titre gratuit uniquement.

Développement asymptotique de la série harmonique

Leçons : *223, 224, 230*

[X-ENS An1], exercice 3.18

On pose, pour tout $n \geqslant 1$, $H_n = \sum_{k=1}^{n} \frac{1}{k}$; cherchons le développement asymptotique de H_n quand n tend vers l'infini.

1. Posons, pour $n \in \mathbb{N}^*$, $u_n = H_n - \ln n$ et $v_n = u_n - \frac{1}{n}$; on va montrer que (u_n) et (v_n) sont adjacentes.
   En effet :
   – Déjà, $\forall n \in \mathbb{N}^*$, $u_n - v_n = \frac{1}{n} > 0$ et $u_n - v_n \underset{n\to\infty}{\longrightarrow} 0$.
   – D'une part, l'inégalité $\ln(1+x) \leqslant x$ (valable pour $x > -1$) fournit la décroissance de (u_n) :
   $$u_n - u_{n-1} = \frac{1}{n} - \ln n + \ln(n-1) = \frac{1}{n} + \ln\left(1 - \frac{1}{n}\right) \leqslant 0.$$
   – D'autre part, la suite (v_n) croît, car $v_{n+1} - v_n = \frac{1}{n} - \ln(n+1) + \ln n = \frac{1}{n} - \ln\left(1 + \frac{1}{n}\right) \geqslant 0$.
   Donc, en tant que suites adjacentes, (u_n) et (v_n) convergent, vers la même limite ; cette limite, qu'on notera γ, s'appelle la constante d'Euler [153].
2. Du coup, on a montré que $H_n = \ln n + \gamma + o(1)$ quand $n \to \infty$.
   On pose, pour $n \in \mathbb{N}^*$, $t_n = u_n - \gamma$.
   Lorsque $n \to \infty$, on a : $t_n - t_{n-1} = \ln\left(1 - \frac{1}{n}\right) + \frac{1}{n} = -\frac{1}{n} - \frac{1}{2n^2} + o\left(\frac{1}{n^2}\right) + \frac{1}{n} \underset{n\to\infty}{\sim} \frac{-1}{2n^2}$.
   Ainsi, la série $\sum_{k\geqslant 2} (t_k - t_{k-1})$ converge.
   Par théorème de sommation des équivalents, on obtient :
   $$t_n = -\sum_{k=n+1}^{\infty} (t_k - t_{k-1}) \underset{n\to\infty}{\sim} -\sum_{k=n+1}^{\infty} \frac{-1}{2k^2} = \frac{1}{2}\sum_{k=n+1}^{\infty} \frac{1}{k^2}$$
3. On va justement chercher un équivalent simple de la quantité $\sum_{k=n+1}^{\infty} \frac{1}{k^\alpha}$, où $\alpha > 1$.
   Comme $t \mapsto \frac{1}{t^\alpha}$ est décroissante et intégrable sur $[1, +\infty[$, on a :
   $$\forall k \geqslant 2, \forall t \in [k, k+1], \frac{1}{t^\alpha} \leqslant \frac{1}{k^\alpha} \leqslant \frac{1}{(t-1)^\alpha}.$$
   En intégrant, on en déduit que : $\forall k \geqslant 2$, $\int_k^{k+1} \frac{\mathrm{d}t}{t^\alpha} \leqslant \frac{1}{k^\alpha} \leqslant \int_{k-1}^{k} \frac{\mathrm{d}t}{t^\alpha}$.
   D'où, en sommant entre $n+1$ et N, puis en faisant tendre N vers l'infini (on a vu que les quantités convergent) :
   $$\int_{n+1}^{\infty} \frac{\mathrm{d}t}{t^\alpha} \leqslant \sum_{k=n+1}^{\infty} \frac{1}{k^\alpha} \leqslant \int_n^{\infty} \frac{\mathrm{d}t}{t^\alpha}.$$
   Comme les deux intégrales sont équivalentes à $\frac{1}{\alpha-1}\frac{1}{n^{\alpha-1}}$, le cas $\alpha = 2$ fournit alors : $t_n \underset{n\to\infty}{\sim} \frac{1}{2n}$.
   Désormais, on a montré que $H_n = \ln n + \gamma + \frac{1}{2n} + o\left(\frac{1}{n}\right)$.
4. Continuons, et posons désormais, pour $n \in \mathbb{N}^*$, $w_n = u_n - \gamma - \frac{1}{2n}$; on a donc $w_n \underset{n\to\infty}{\longrightarrow} 0$.
   En conséquence, la somme $\sum_{k=n+1}^{\infty} (w_k - w_{k-1})$ vaut $-w_n$.

153. Valeur approchée par défaut : $\gamma \simeq 0,577215$.

Diffusion à titre gratuit uniquement.

Or $\forall n \in \mathbb{N}^*$, $w_n - w_{n-1} = u_n - u_{n-1} + \dfrac{1}{2n} - \dfrac{1}{2n-2} = \ln\left(1 - \dfrac{1}{n}\right) + \dfrac{1}{n} + \dfrac{1}{2n} - \dfrac{1}{2n-2}$.
Donc, quand $n \to \infty$,

$$\begin{aligned} w_n - w_{n-1} &= -\frac{1}{n} - \frac{1}{2n^2} - \frac{1}{3n^3} + \frac{1}{n} + \frac{1}{2n}\left(1 - \frac{1}{1-\frac{1}{n}}\right) + o\left(\frac{1}{n^3}\right) \\ &= -\frac{1}{2n^2} - \frac{1}{3n^3} + \frac{1}{2n}\left(\frac{1}{n} + \frac{1}{n^2} + o\left(\frac{1}{n^2}\right)\right) + o\left(\frac{1}{n^3}\right) \\ &\underset{n\to\infty}{\sim} \frac{1}{6n^3} \end{aligned}$$

Ensuite, par sommation des équivalents, $w_n \underset{n\to\infty}{\sim} -\dfrac{1}{6}\displaystyle\sum_{k=n+1}^{\infty} \frac{1}{k^3} \underset{n\to\infty}{\sim} \frac{-1}{6}\frac{1}{2n^2} = \frac{-1}{12n^2}$.
Donc, on va s'arrêter [154] avec le développement suivant : $H_n = \ln n + \gamma + \dfrac{1}{2n} - \dfrac{1}{12n^2} + o\left(\dfrac{1}{n^2}\right)$.

5. Pour finir, notons pour $n \in \mathbb{N}^*$, $k_n = \min\{k \in \mathbb{N} | H_k \geqslant n\}$ (le rang auquel la série harmonique dépasse la valeur n).
On va déduire du développement asymptotique de H_n la valeur de $\lim\limits_{n\to\infty} \dfrac{k_{n+1}}{k_n}$. [155]
On pose $H_n = \ln n + \gamma + \varepsilon_n$, où $\varepsilon_n \underset{n\to\infty}{\longrightarrow} 0$.
Par définition de k_n, on a : $\ln k_n + \gamma + \varepsilon_{k_n} \geqslant n$ et $\ln(k_n - 1) + \gamma + \varepsilon_{k_n-1} < n$.
Puis $k_n \geqslant \exp(n - \gamma - \varepsilon_{k_n})$ et $k_n - 1 < \exp(n - \gamma - \varepsilon_{k_n-1})$.
D'où l'encadrement $\exp(n - \gamma - \varepsilon_{k_n}) \leqslant k_n < \exp(n - \gamma - \varepsilon_{k_n-1}) + 1$.
Ainsi, $k_n \underset{n\to\infty}{\longrightarrow} \mathrm{e}^{n-\gamma}$ puis $\lim\limits_{n\to\infty} \dfrac{k_{n+1}}{k_n} = \mathrm{e}$.

Références

[X-ENS An1] S. FRANCINOU, H. GIANELLA et S. NICOLAS – *Oraux X-ENS Analyse 1*, 3e éd., Cassini, 2014.

154. C'est déjà assez chiant comme ça...
155. On s'occupe comme on peut !

Diffusion à titre gratuit uniquement.

Distributions à support ponctuel

Leçons : ***254, 255***

[Zui], théorème 2.4.5

Théorème

Soit Ω un ouvert de $\mathbb{R}^n$ et $x_0 \in \Omega$.

1. Soit $\alpha \in \mathbb{N}^n$, on a : $\operatorname{supp} \partial^\alpha \delta_{x_0} = \{x_0\}$.
2. Soit $T \in \mathcal{D}'(\Omega)$, telle que $\operatorname{supp} T \subset \{x_0\}$.
   Alors $\exists k \in \mathbb{N}, \exists \left(a_\alpha\right)_{|\alpha| \leqslant k} \subset \mathbb{C}, T = \sum_{|\alpha| \leqslant k} a_\alpha \partial^\alpha \delta_{x_0}$.

Démonstration :

1. Soit $\varphi \in \mathcal{D}\left(\Omega \backslash \{x_0\}\right)$, on a : $\langle \partial^\alpha \delta_{x_0}, \varphi \rangle = 0$ donc $\operatorname{supp} \partial^\alpha \delta_{x_0} \subset \{x_0\}$.
   Réciproquement, soit V_{x_0} un voisinage de x_0 dans Ω et $\chi \in \mathcal{D}\left(V_{x_0}\right)$ avec $\chi \equiv 1$ au voisinage de x_0.
   On pose $\varphi : x \mapsto \dfrac{(x-x_0)^\alpha}{\alpha!}\chi(x)$, alors $\varphi \in \mathcal{D}\left(V_{x_0}\right)$.
   De plus, par la formule de Leibniz :

$$\partial^\alpha \varphi\left(x_0\right) = \frac{1}{\alpha!} \sum_{\beta \leqslant \alpha} \binom{\alpha}{\beta} \partial^\beta \left[(x-x_0)^\alpha\right] \big|_{x=x_0} \partial^{\alpha-\beta} \chi\left(x_0\right) = \frac{1}{\alpha!} \binom{\alpha}{\alpha} \partial^\alpha \left[(x-x_0)^\alpha\right] \big|_{x=x_0} \chi\left(x_0\right) = \frac{1}{\alpha!} 1\alpha! 1 = 1$$

2. On aura besoin des deux lemmes suivants.

   Lemme 1

   Soit $T \in \mathcal{D}'(\Omega)$, $\varphi \in \mathcal{D}(\Omega)$ tels que $\operatorname{supp} T \cap \operatorname{supp} \varphi = \varnothing$.
   Alors $\langle T, \varphi \rangle = 0$.

   Lemme 2

   Soit $k \in \mathbb{N}$, $T \in \mathcal{D}'^{(k)}(\Omega)$ et $\varphi \in \mathcal{C}^k_{\mathrm{C}}(\Omega)$.
   On suppose que : $\forall x \in \operatorname{supp} T, \forall \alpha \in \mathbb{N}^n, |\alpha| \leqslant k \Rightarrow \partial^\alpha \varphi(x) = 0$.
   Alors $\langle T, \varphi \rangle = 0$.

   Soit ω un ouvert contenant x_0 avec $\overline{\omega}$ compact dans Ω.
   Comme $\overline{\omega}$ est compact, on sait que $T \in \mathcal{D}'(\omega)$ est d'ordre fini, noté $k \in \mathbb{N}$.
   Soit $\chi \in \mathcal{D}(\omega)$, tel que $\chi \equiv 1$ sur un voisinage de x_0, soit $\varphi \in \mathcal{D}(\Omega)$.
   On a : $\langle T, (1-\chi)\varphi \rangle = 0$ car $\operatorname{supp} T \cap \operatorname{supp}(1-\chi) = \varnothing$ donc $\langle T, \chi\varphi \rangle = \langle T, \varphi \rangle$.
   On pose : $\psi : x \mapsto \chi(x) \left[\varphi(x) - \sum_{|\alpha| \leqslant k} \dfrac{(x-x_0)^\alpha}{\alpha!} \partial^\alpha \varphi\left(x_0\right)\right]$; alors $\psi \in \mathcal{D}(\omega)$ et pour $|\beta| \leqslant k$:

$$\begin{aligned}
\partial^\beta \psi\left(x_0\right) &= \sum_{\gamma \leqslant \beta} \binom{\beta}{\gamma} \partial^\gamma \chi\left(x_0\right) \partial^{\beta-\gamma} \left[\varphi(x) - \sum_{|\alpha| \leqslant k} \frac{(x-x_0)^\alpha}{\alpha!} \partial^\alpha \varphi\left(x_0\right)\right] \Bigg|_{x=x_0} \\
&= \binom{\beta}{0} 1 \left(\partial^\beta \varphi\left(x_0\right) - 0 - \frac{\beta!}{\beta!} \partial^\beta \varphi\left(x_0\right)\right) = 0
\end{aligned}$$

   Et comme $\operatorname{supp} T = \{x_0\}$, par le lemme 2, on a : $\langle T, \varphi \rangle = 0$.
   Par conséquent,

$$\langle T, \varphi \rangle = \langle T, \chi\varphi \rangle = \sum_{|\alpha| \leqslant k} \frac{\partial^\alpha \varphi\left(x_0\right)}{\alpha!} \langle T, \chi_\alpha \rangle = \sum_{|\alpha| \leqslant k} \frac{\langle T, \chi_\alpha \rangle}{\alpha!} \partial^\alpha \varphi\left(x_0\right)$$

   où $\chi_\alpha : x \mapsto \chi(x)(x-x_0)^\alpha$.

Diffusion à titre gratuit uniquement.

Démonstration du lemme 1 :
Comme $\operatorname{supp}\varphi \subset (\operatorname{supp} T)^c$, on a :

$$\forall x \in \operatorname{supp}\varphi, \exists V_x \text{ un voisinage ouvert de } x, \forall \psi \in \mathcal{D}(V_x), \langle T, \psi\rangle = 0$$

Alors on obtient : $\operatorname{supp}\varphi \subset \bigcup_{x\in\operatorname{supp}\varphi} V_x$.

Et comme $\operatorname{supp}\varphi$ est compact, par Borel-Lebesgue : $\exists x_1, \dots, x_N \in \operatorname{supp}\varphi, \operatorname{supp}\varphi \subset \bigcup_{i=1}^{N} V_{x_i}$.

Soit $(\chi_i)_{1\leqslant i\leqslant N}$ une famille d'éléments tels que :
– $\forall i \in [\![1, N]\!], \chi_i \in \mathcal{D}(V_{x_i})$;
– $\sum_{i=1}^{N} \chi_i(x) = 1$ pour $x \in \operatorname{supp}\varphi$.

Alors $\varphi = \sum_{i=1}^{N} \chi_i\varphi$ et donc $\langle T, \varphi\rangle = \sum_{i=1}^{N} \langle T, \chi_i\varphi\rangle = 0$ car $\chi_i\varphi \in \mathcal{D}(V_{x_i})$. ■

Démonstration du lemme 2 :
On note $K = \operatorname{supp} T \cap \operatorname{supp}\varphi$. Si $K = \varnothing$, on renvoie au lemme 1. Supposons donc désormais que K est non-vide.
Pour $\varepsilon > 0$, on pose $K_\varepsilon = \{x \in \mathbb{R}^n | d(x, K) \leqslant \varepsilon\}$. On a : $\exists \varepsilon_0 \in]0, 1], \forall \varepsilon \in]0, \varepsilon_0], K_\varepsilon \subset \Omega$.
Soit $\chi_\varepsilon \in \mathcal{D}(K_\varepsilon)$, telle que $\chi_\varepsilon \equiv 1$ sur $K_{\frac{\varepsilon}{2}}$ et telle que quand $|\alpha| \leqslant k$, on ait : $|\partial^\alpha \chi_\varepsilon| \leqslant C_\alpha \varepsilon^{-|\alpha|}$.
On a $\langle T, (1-\chi_\varepsilon)\varphi\rangle = 0$ car $\operatorname{supp} T \cap \operatorname{supp}(1-\chi_\varepsilon)\varphi \subset K \cap K_{\frac{\varepsilon}{2}}^c = \varnothing$.
Donc $\langle T, \varphi\rangle = \langle T, \chi_\varepsilon\rangle$, puis :

$$|\langle T, \varphi\rangle| \leqslant C \sum_{|\alpha|\leqslant k} \sup_{K_\varepsilon} |\partial^\alpha(\chi_\varepsilon\varphi)|, \text{ où } C \in \mathbb{R}^{+*} \text{ est indépendant de } \varepsilon$$

Or, par Leibniz : $\partial^\alpha(\chi_\varepsilon\varphi) = \sum_{\beta\leqslant\alpha} \binom{\alpha}{\beta} \partial^{\alpha-\beta}\chi_\varepsilon \partial^\beta\varphi$.
D'où : $|\langle T, \varphi\rangle| \leqslant C \sum_{|\alpha|\leqslant k} \sum_{\beta\leqslant\alpha} \binom{\alpha}{\beta} C_{\alpha-\beta} \varepsilon^{|\beta|-|\alpha|} \sup_{K_\varepsilon} \left|\partial^\beta\varphi\right|$.
Et comme $|\beta| - |\alpha| \geqslant |\beta| - k$ et $\varepsilon \leqslant 1$, on a :

$$|\langle T, \varphi\rangle| \leqslant C \sum_{|\beta|\leqslant k} \sum_{\substack{|\alpha|\leqslant k \\ \alpha\geqslant\beta}} \binom{\alpha}{\beta} C_{\alpha-\beta} \varepsilon^{|\beta|-k} \sup_{K_\varepsilon} \left|\partial^\beta\varphi\right|$$

On pose $C' = C \max_{|\beta|\leqslant k} \left(\sum_{\substack{|\alpha|\leqslant k \\ \alpha\geqslant\beta}} \binom{\alpha}{\beta} C_{\alpha-\beta} \right)$ et on obtient :

$$|\langle T, \varphi\rangle| \leqslant C' \sum_{|\beta|\leqslant k} \varepsilon^{|\beta|-k} \sup_{K_\varepsilon} \left|\partial^\beta\varphi\right|$$

On va montrer que $\lim_{\varepsilon\to 0} \varepsilon^{|\beta|-k} \sup_{K_\varepsilon} \left|\partial^\beta\varphi\right| = 0$ pour $|\beta| \leqslant k$; on aura ainsi terminé la preuve.
On sait que : $\exists x_\varepsilon \in K_\varepsilon, \sup_{K_\varepsilon} \left|\partial^\beta\varphi\right| = \left|\partial^\beta\varphi(x_\varepsilon)\right|$ et $\exists x_0 \in K, |x_\varepsilon - x_0| \leqslant \varepsilon$.
$\partial^\beta\varphi$ étant continue à support compact, elle est uniformément continue. Soit $\delta > 0$;

$$\exists \eta > 0, \forall x, x' \in \Omega, |x - x'| < \eta \Rightarrow \left|\partial^\beta\varphi(x) - \partial^\beta\varphi(x')\right| \leqslant \delta$$

Choisissons $\varepsilon \leqslant \eta$, on obtient $\left|\partial^\beta\varphi(x_\varepsilon) - \partial^\beta\varphi(x_0)\right| \leqslant \delta$ et comme $x_0 \in K \subset \operatorname{supp} T$: $\left|\partial^\beta\varphi(x_\varepsilon)\right| \leqslant \delta$.
On en déduit, pour $|\beta| = k$, $\lim_{\varepsilon\to 0} \varepsilon^{|\beta|-k} \sup_{K_\varepsilon} \left|\partial^\beta\varphi\right| = 0$.

Diffusion à titre gratuit uniquement.

Supposons désormais $|\beta| < k$; par la formule de Taylor avec reste intégral :

$$\partial^\beta \varphi(x_\varepsilon) = \sum_{|\gamma| \leqslant k-1-|\beta|} \frac{(x_\varepsilon - x_0)^\gamma}{\gamma!} \partial^{\gamma+\beta}\varphi(x_0) + \sum_{|\gamma|=k-|\beta|} \int_0^1 \frac{(1-t)^{k-|\beta|-1}}{(k-|\beta|-1)!} \partial^{\gamma+\beta}\varphi(x_0 + t(x_\varepsilon - x_0))(x_\varepsilon - x_0)^\gamma \, \mathrm{d}t$$

$$\left|\partial^\beta \varphi(x_\varepsilon)\right| \leqslant \sum_{|\gamma| \leqslant k-1-|\beta|} \frac{\left|(x_\varepsilon - x_0)^\gamma\right|}{\gamma!} \left|\partial^{\gamma+\beta}\varphi(x_0)\right| + \frac{1}{(k-|\beta|-1)!} \sum_{|\gamma|=k-|\beta|} \int_0^1 (1-t)^{k-|\beta|-1} \left|\partial^{\gamma+\beta}\varphi((1-t)x_0 + tx_\varepsilon)\right| \left|(x_\varepsilon - x_0)^\gamma\right| \, \mathrm{d}t$$

Pour tout $t \in [0,1]$, on a : $|(1-t)x_0 + tx_\varepsilon - x_0| = t\,|x_\varepsilon - x_0| \leqslant \varepsilon$ donc $(1-t)x_0 + tx_\varepsilon \in K_\varepsilon$.
D'autre part, $x_0 \in K \subset \operatorname{supp} T$ donc pour $|\gamma| + |\beta| \leqslant k-1$, on a : $\partial^{\gamma+\beta}\varphi(x_0) = 0$.
Par conséquent :

$$\left|\partial^\beta \varphi(x_\varepsilon)\right| \leqslant \frac{1}{(k-|\beta|-1)!} \sum_{|\gamma|=k-|\beta|} \int_0^1 1 \sup_{K_\varepsilon} \left|\partial^{\gamma+\beta}\varphi\right| \varepsilon^{|\gamma|} \, \mathrm{d}t = \frac{\varepsilon^{k-|\beta|}}{(k-|\beta|-1)!} \sum_{|\gamma|+|\beta|=k} \sup_{K_\varepsilon} \left|\partial^{\gamma+\beta}\varphi\right|$$

Donc $\varepsilon^{|\beta|-k}\left|\partial^\beta \varphi(x_\varepsilon)\right| \leqslant \frac{1}{(k-|\beta|-1)!} \sum_{|\alpha|=k} \underbrace{\sup_{K_\varepsilon} |\partial^\alpha \varphi|}_{\xrightarrow[\varepsilon\to 0]{} 0}$.

Et finalement : $\lim_{\varepsilon\to 0} \varepsilon^{|\beta|-k} \sup_{K_\varepsilon} \left|\partial^\beta \varphi\right| = 0$. ■

■

Références

[Zui] C. ZUILY – *Éléments de distributions et d'équations aux dérivées partielles*, Dunod, 2002.

Diffusion à titre gratuit uniquement.

Inversion de la transformée de Fourier

Leçons : ***234, 235, 239, 240, 261***

[Ouv2], section 12.3

Rappels :

1. Si μ est une mesure bornée sur $\mathbb{R}^d$, on définit : $\widehat{\mu} : \left| \begin{array}{ccc} \mathbb{R}^d & \to & \mathbb{C} \\ t & \mapsto & \int_{\mathbb{R}^d} \mathrm{e}^{\mathrm{i}\langle x,t\rangle}\,\mathrm{d}\mu(x) \end{array} \right.$.
2. Soit $y \in \mathbb{R}^d$, si $g(y-.)$ est μ-intégrable, on définit $(g \star \mu)(y) = \displaystyle\int_{\mathbb{R}^d} g(y-x)\,\mathrm{d}\mu(x)$.

Théorème

Soit μ une mesure bornée sur $\mathbb{R}^d$ et telle que $\widehat{\mu} \in \mathrm{L}^1(\lambda)$.

Alors $\mu \ll \lambda$ et sa densité (au sens de Radon-Nikodym) est : $h(x) = \left(\dfrac{1}{2\pi}\right)^d \displaystyle\int_{\mathbb{R}^d} \widehat{\mu}(t)\mathrm{e}^{-\mathrm{i}\langle x,t\rangle}\,\mathrm{d}t$.

Prérequis

- Pour $\sigma > 0$, $g_\sigma : x \mapsto \left(\dfrac{1}{\sqrt{2\pi}\sigma}\right) \exp\left(-\dfrac{\|x\|^2}{2\sigma^2}\right)$ est une densité de probabilité.
- On a : $\forall t \in \mathbb{R}^d, \widehat{g_1}(t) = \left(\sqrt{2\pi}\right)^d g_1(t)$.
- Par convergence dominée : $\forall f \in \mathcal{C}_\mathrm{b}\left(\mathbb{R}^d\right), \forall x \in \mathbb{R}^d, \left(f \star g_\sigma\right)(x) \underset{\sigma\to 0}{\longrightarrow} f(x)$.

Démonstration :

Étape 1 : On va montrer que : $\forall y \in \mathbb{R}^d, \left(g_\sigma \star \mu\right)(y) = \left(\dfrac{1}{\sqrt{2\pi}}\right)^d \displaystyle\int_{\mathbb{R}^d} \widehat{\mu}(v) g_1(\sigma v)\mathrm{e}^{-\mathrm{i}\langle y,v\rangle}\,\mathrm{d}v$.

Comme g_σ et μ sont bornées, on a bien : $\forall y \in \mathbb{R}^d, g_\sigma(y-.) \in \mathrm{L}^1(\mu)$.

De plus, $g_\sigma(y-x) = g_\sigma(x-y) = \left(\dfrac{1}{\sqrt{2\pi}\sigma}\right)^d \exp\left(-\dfrac{\|x-y\|^2}{2\sigma^2}\right) = \left(\dfrac{1}{\sqrt{2\pi}\sigma}\right)^d g_1\left(\dfrac{x-y}{\sigma}\right)$.

Par un changement de variable, on obtient ensuite :

$$g_\sigma(y-x) = \left(\frac{1}{\sqrt{2\pi}}\right)^d \int_{\mathbb{R}^d} g_1(z) \exp\left(\mathrm{i}\left\langle \frac{x-y}{\sigma}, z\right\rangle\right) \frac{\mathrm{d}z}{\sigma^d} = \left(\frac{1}{\sqrt{2\pi}}\right)^d \int_{\mathbb{R}^d} g_1(\sigma v) \exp(\mathrm{i}\langle x-y, v\rangle)\,\mathrm{d}v$$

Donc $\left(g_\sigma \star \mu\right)(y) = \displaystyle\int_{\mathbb{R}^d} \left(\frac{1}{\sqrt{2\pi}}\right)^d \int_{\mathbb{R}^d} g_1(\sigma v) \exp(\mathrm{i}\langle x-y, v\rangle)\,\mathrm{d}v\mathrm{d}\mu(x)$.

Or $|g_1(\sigma v)\exp(\mathrm{i}\langle x-y, v\rangle)| = \left(\dfrac{1}{\sqrt{2\pi}}\right)^d \exp\left(-\dfrac{\sigma^2\|v\|2}{2}\right)$ est intégrable par rapport à $\lambda \otimes \mu$ car μ est bornée et $v \mapsto \exp\left(-\dfrac{\sigma^2\|v\|^2}{2}\right) \in \mathrm{L}^1(\lambda)$.

On applique Fubini :

$$\left(g_\sigma \star \mu\right)(y) = \left(\frac{1}{\sqrt{2\pi}}\right)^d \int_{\mathbb{R}^d} g_1(\sigma v)\mathrm{e}^{-\mathrm{i}\langle y,v\rangle} \int_{\mathbb{R}^d} \mathrm{e}^{\mathrm{i}\langle x,v\rangle}\,\mathrm{d}\mu(x)\mathrm{d}v = \left(\frac{1}{\sqrt{2\pi}}\right)^d \int_{\mathbb{R}^d} g_1(\sigma v)\mathrm{e}^{-\mathrm{i}\langle y,v\rangle}\widehat{\mu}(v)\mathrm{d}v$$

Étape 2 : Montrons que $\mu \ll \lambda$.

Soit $f \in \mathcal{C}_\mathrm{K}\left(\mathbb{R}^d\right)$, on a : $\displaystyle\int_{\mathbb{R}^d} f(x)\,\mathrm{d}\mu(x) = \int_{\mathbb{R}^d} \lim_{\sigma\to 0} \left(f \star g_\sigma\right)(x)\,\mathrm{d}\mu(x)$.

On applique le théorème de convergence dominée, car :

$$\left|\left(f \star g_\sigma\right)(x)\right| = \left|\int_{\mathbb{R}^d} f(x-y)g_\sigma(y)\,\mathrm{d}y\right| \leqslant \int_{\mathbb{R}^d} |f(x-y)|\,|g_\sigma(y)|\,\mathrm{d}y \leqslant \|f\|_\infty \in \mathrm{L}^1(\mu) \text{ car } \mu \text{ est bornée.}$$

Diffusion à titre gratuit uniquement.

On a donc : $\int_{\mathbb{R}^d} f(x)\,\mathrm{d}\mu(x) = \lim_{\sigma\to 0}\int_{\mathbb{R}^d}(f\star g_\sigma)(x)\,\mathrm{d}\mu(x) = \lim_{\sigma\to 0}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} f(y)g_\sigma(x-y)\,\mathrm{d}y\mathrm{d}\mu(x)$.
On applique Fubini, car $|f(y)g_\sigma(x-y)| \leqslant |f(y)|\,\|g_\sigma\|_\infty$ est $\mu\otimes\lambda$-intégrable.
Ainsi

$$\begin{aligned}\int_{\mathbb{R}^d} f(x)\,\mathrm{d}\mu(x) &= \lim_{\sigma\to 0}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} f(y)g_\sigma(x-y)\,\mathrm{d}\mu(x)\mathrm{d}y = \lim_{\sigma\to 0}\int_{\mathbb{R}^d} f(y)\int_{\mathbb{R}^d} g_\sigma(y-x)\,\mathrm{d}\mu(x)\mathrm{d}y\\ &= \lim_{\sigma\to 0}\int_{\mathbb{R}^d} f(y)\,(g_\sigma\star\mu)(y)\,\mathrm{d}y\end{aligned}$$

Mais on a :

$$|f(y)\,(g_\sigma\star\mu)(y)| = |f(y)|\left|\int_{\mathbb{R}^d}\left(\frac{1}{\sqrt{2\pi}}\right)^d \widehat{\mu}(v)g_1(\sigma v)\mathrm{e}^{-\mathrm{i}\langle y,v\rangle}\,\mathrm{d}v\right| \leqslant \frac{|f(y)|}{(2\pi)^d}\,\|\widehat{\mu}\|_1$$

qui est intégrable par rapport à la mesure de Lebesgue sur $\mathbb{R}^d$.
Ainsi, par convergence dominée :

$$\forall f\in\mathcal{C}_{\mathrm{K}}\left(\mathbb{R}^d\right),\ \int_{\mathbb{R}^d} f(x)\,\mathrm{d}\mu(x) = \int_{\mathbb{R}^d} f(y)\lim_{\sigma\to 0}(g_\sigma\star\mu)(y)\,\mathrm{d}y$$

On a donc bien $\mu\ll\lambda$.

Étape 3 : La relation précédente nous donne : $h(x) = \lim_{\sigma\to 0}(g_\sigma\star\mu)(x)$.

Ainsi : $h(x) = \lim_{\sigma\to 0}\left(\frac{1}{2\pi}\right)^d\int_{\mathbb{R}^d}\widehat{\mu}(v)g_1(\sigma v)\mathrm{e}^{-\mathrm{i}\langle x,v\rangle}\,\mathrm{d}v$.

Or $\left|\widehat{\mu}(v)g_1(\sigma v)\mathrm{e}^{-\mathrm{i}\langle x,v\rangle}\right| \leqslant \left(\frac{1}{\sqrt{2\pi}}\right)^d|\widehat{\mu}(v)| \in \mathrm{L}^1\left(\mathbb{R}^d\right)$.

Donc $h(x) = \left(\frac{1}{2\pi}\right)^d\int_{\mathbb{R}^d}\widehat{\mu}(v)\mathrm{e}^{-\mathrm{i}\langle x,v\rangle}\,\mathrm{d}v$. ■

Références

[Ouv2] J.-Y. OUVRARD – *Probabilités 2*, 3e éd., Cassini, 2009.

Diffusion à titre gratuit uniquement.

Ruine du joueur

Leçons : *223, 226, 249, 264*

[GS], partie 3.9

Théorème

Soit $(Y_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires indépendantes et identiquement distribuées de loi $p\delta_1 + (1-p)\delta_{-1}$; on note $S_n = a + \sum_{i=1}^{n} Y_i$, où $a \in \mathbb{N}$.
Soit $b \in \mathbb{N}^*$, on note $T = \inf\{n \in \mathbb{N} | S_n \in \{0, a+b\}\}$ et $\rho = \mathbb{P}(S_T = a+b)$.[156]
On a deux cas (on note $q = 1-p$) :

- Soit $p \neq \frac{1}{2}$, dans ce cas $\rho = \dfrac{1-\left(\frac{q}{p}\right)^a}{1-\left(\frac{q}{p}\right)^{a+b}}$ et $\mathbb{E}[T] = \dfrac{(a+b)\rho - a}{p-q}$;
- Soit $p = \frac{1}{2}$, dans ce cas $\rho = \dfrac{a}{a+b}$ et $\mathbb{E}[T] = ab$.

Démonstration :

Étape 1 : On va commencer par obtenir une formule de récurrence.
On note $\mathbb{P}_k = \mathbb{P}(.|S_0 = k)$ et $A = \{S_T = a+b\}$. On calcule $\mathbb{P}_k(A)$.
Par la formule des probabilités totales :

$$\mathbb{P}_k(A) = \mathbb{P}_k(A|Y_1 = 1)\,\mathbb{P}_k(Y_1 = 1) + \mathbb{P}_k(A|Y_1 = -1)\,\mathbb{P}_k(Y_1 = -1) = \mathbb{P}_{k+1}(A)p + \mathbb{P}_{k-1}(A)q.$$

On note p_k pour désigner $\mathbb{P}_k(A)$, et on obtient la récurrence linéaire :

$$\begin{cases} p_k = pp_{k+1} + qp_{k-1}, \text{ pour } 1 \leqslant k \leqslant a+b-1 \\ p_0 = 1 \text{ et } p_{a+b} = 0 \end{cases}.$$

Étape 2 : Déterminons la valeur de ρ.

- Dans le cas où $p \neq \frac{1}{2}$, alors l'équation caractéristique de cette récurrence linéaire est : $x = px^2 + q$, dont les solutions sont 1 et $\frac{q}{p}$.
  On en déduit alors : $\exists \alpha, \beta \in \mathbb{R}, \forall k \in [\![0, a+b]\!], p_k = \alpha + \beta\left(\frac{q}{p}\right)^k$.
  Mais on dispose des valeurs de p_0 et de p_{a+b} ; cela nous fournit : $0 = \alpha + \beta$ et $1 = \alpha + \beta\left(\frac{q}{p}\right)^{a+b}$.
  Ainsi, $\alpha = -\beta$ et $\beta = \left(-1 + \left(\frac{q}{p}\right)^{a+b}\right)^{-1}$.
  Par conséquent, $\rho = p_a = \dfrac{1-\left(\frac{q}{p}\right)^a}{1-\left(\frac{q}{p}\right)^{a+b}}$.
- Dans le cas où $p = \frac{1}{2}$, alors l'équation caractéristique devient : $x = \frac{1}{2}x^2 + \frac{1}{2}$, dont 1 est l'unique solution.
  On en déduit alors : $\exists \alpha, \beta \in \mathbb{R}, \forall k \in [\![0, a+b]\!], p_k = \alpha + \beta k$.
  Mais on dispose des valeurs de p_0 et de p_{a+b} ; cela nous fournit : $0 = \alpha$ et $1 = \alpha + \beta(a+b)$.

156. Un point sur l'interprétation de l'énoncé. Les variables Y_n désignent le résultat d'un jeu qui peut se solder par le gain ou la perte d'un euro pour le joueur ; la probabilité de gagner un euro valant p. La valeur S_n désigne la fortune du joueur après n répétitions du jeu ; ainsi $S_0 = a$ est l'argent que possède le joueur avant de commencer à jouer. La banque, quant à elle, possède b euros ; quand le joueur gagne, elle perd, et réciproquement. Ainsi, le jeu doit s'arrêter quand l'un ou l'autre des acteurs (le joueur ou la banque) fait faillite : cela se produit à l'instant T. La probabilité que le joueur gagne finalement contre la banque est donc ρ. On veut connaître cette probabilité, ainsi que le temps moyen de jeu, en fonction des valeurs de a, b et p.

Diffusion à titre gratuit uniquement.

Ainsi, $\alpha = 0$ et $\beta = \dfrac{1}{a+b}$.
Par conséquent, $\rho = p_a = \dfrac{a}{a+b}$.

Étape 3 : On va suivre la même tactique pour calculer l'espérance de T : d'abord obtenir une formule de récurrence, puis en profiter, en distinguant les cas selon la valeur de p.

– La formule des probabilités totales nous donne :

$$\mathbb{E}_k[T] = \mathbb{E}_k\left[T|Y_1=1\right]\mathbb{P}\left(Y_1=1\right) + \mathbb{E}_k\left[T|Y_1=-1\right]\mathbb{P}\left(Y_1=-1\right) = \left(1+\mathbb{E}_{k+1}[T]\right)p + \left(1+\mathbb{E}_{k-1}[T]\right)q.$$

On note e_k pour désigner $\mathbb{E}_k[T]$, et on obtient la récurrence linéaire :

$$\begin{cases} e_k = 1 + pe_{k+1} + qe_{k-1}, \text{ pour } 1 \leqslant k \leqslant a+b-1 \\ e_0 = 0 \text{ et } e_{a+b} = 0 \end{cases}.$$

On note, pour $k \geqslant 1$, $d_k = e_k - e_{k-1}$; la relation de récurrence devient : $0 = 1 + pd_{k+1} - qd_k$.

– Dans le cas où $p \neq \dfrac{1}{2}$, on pose l le réel vérifiant : $0 = 1 + pl - ql$, c'est-à-dire que $l = \dfrac{1}{q-p}$.
La relation de conséquence devient donc : $0 = p\left(d_{k+1} - l\right) - q\left(d_k - l\right)$.
Par conséquent, $d_{k+1} - l = \dfrac{q}{p}\left(d_k - l\right)$, et donc $d_k - l = \left(\dfrac{q}{p}\right)^{k-1}\left(d_1 - l\right)$.
Donc, pour $n \geqslant 1$, on a :

$$e_n = \sum_{k=1}^{n} d_k + e_0 = \sum_{k=1}^{n}\left(l + \left(\frac{q}{p}\right)^{k-1}\left(d_1 - l\right)\right) = nl + \left(d_1 - l\right)\sum_{k=0}^{n-1}\left(\frac{q}{p}\right)^{k} = nl + \left(d_1 - l\right)\frac{1-\left(\frac{q}{p}\right)^{n}}{1-\frac{q}{p}}.$$

Or $e_{a+b} = 0 = (a+b)l + \left(d_1 - l\right)\dfrac{1-\left(\frac{q}{p}\right)^{a+b}}{1-\frac{q}{p}}$, ce qui nous donne : $\dfrac{d_1 - l}{1-\frac{q}{p}} = -\dfrac{(a+b)l}{1-\left(\frac{q}{p}\right)^{a+b}}$.

Donc $\mathbb{E}[T] = e_a = al - \left(1 - \left(\dfrac{q}{p}\right)^{a}\right)\dfrac{(a+b)l}{1-\left(\frac{q}{p}\right)^{a+b}} = \dfrac{1}{q-p}\left(a - (a+b)\dfrac{1-\left(\frac{q}{p}\right)^{a}}{1-\left(\frac{q}{p}\right)^{a+b}}\right) = \dfrac{(a+b)\rho - a}{p-q}$.

– Dans le cas où $p = \dfrac{1}{2}$, alors $\dfrac{1}{2}d_{k+1} = \dfrac{1}{2}d_k - 1$, donc $d_{k+1} = d_k - 2$, d'où $d_k = d_1 - 2(k-1)$.
Alors, pour $n \geqslant 1$, $e_n = \displaystyle\sum_{k=1}^{n} d_k = nd_1 - 2\sum_{k=0}^{n-1} k = nd_1 - 2\frac{n(n-1)}{2} = n\left(d_1 - (n-1)\right)$.
Or $e_{a+b} = 0 = (a+b)\left(d_1 + (a+b-1)\right)$, d'où $d_1 = a+b-1$.
Donc $\mathbb{E}[T] = e_a = a(a+b-1-(a-1)) = ab$. ■

Références

[GS] G. R. GRIMMETT et D. R. STIRZAKER – *Probability and Random Processes*, 3e éd., Oxford University Press, 2001.

Diffusion à titre gratuit uniquement.

Théorème de Carathéodory

Leçons : ***181***

[TauGéo], résultats 4.3.5-4.3.6

Théorème

Soient $\mathcal{E}$ un espace affine de dimension finie, d'espace vectoriel associé E, et $\mathcal{A} \subset \mathcal{E}$, avec $\mathcal{A} \neq \varnothing$.
Tout élément de $\text{Conv}(\mathcal{A})$ s'écrit comme combinaison convexe de k points de $\mathcal{A}$, avec $k \leqslant 1 + \dim \mathcal{E}$.

Démonstration :

Soit $M \in \text{Conv}(\mathcal{A})$; par définition de l'enveloppe convexe, M est combinaison convexe d'un nombre fini de points de $\mathcal{A}$, notés $A_1, \dots, A_k$.

On a donc : $M = t_1 A_1 + \dots + t_k A_k$, avec $0 \leqslant t_1, \dots, t_k \leqslant 1$ et $\sum_{i=1}^{k} t_i = 1$.

On suppose que $k > 1 + \dim \mathcal{E}$, puisque sinon, on est content.

La famille $\left(\overrightarrow{A_1A_2}, \dots, \overrightarrow{A_1A_k}\right)$ est liée, car elle possède au moins $(1 + \dim E)$ vecteurs de E.

En conséquence, il existe $\lambda_2, \dots, \lambda_k \in \mathbb{R}$ non-tous nuls, tels que : $\lambda_2 \overrightarrow{A_1A_2} + \dots + \lambda_k \overrightarrow{A_1A_k} = \overrightarrow{0}$.

On pose alors $\mu_1 = \lambda_2 + \dots + \lambda_k$ et pour $i \in [\![2, k]\!]$, $\mu_i = -\lambda_i$.

On a alors : $\mu_1 \overrightarrow{OA_1} + \dots + \mu_k \overrightarrow{OA_k} = \overrightarrow{0}$, où O est un point quelconque, fixé, de $\mathcal{E}$.

Comme $\mu_1 + \dots + \mu_k = 0$, et que les μ_i ($i \in [\![1, k]\!]$) sont non-tous nuls, on sait que : $\exists j \in [\![1, k]\!], \mu_j > 0$.

On pose alors $\lambda = \min\left\{ \dfrac{t_i}{\mu_i} \middle| \mu_i > 0 \right\}$, puis, pour $i \in [\![1, k]\!]$, $v_i = t_i - \lambda \mu_i$.

De cette façon, $v_1, \dots, v_k \geqslant 0$ et $\sum_{i=1}^{k} v_i = \sum_{i=1}^{k} t_i - 0 = 1$.

Aussi, $\exists q \in [\![1, k]\!], \lambda = \dfrac{t_q}{\mu_q}$, d'où $v_q = 0$.

Ainsi, $\overrightarrow{OM} = \sum_{i=1}^{k} t_i \overrightarrow{OA_i} = \sum_{i=1}^{k} v_i \overrightarrow{OA_i} + \lambda \sum_{i=1}^{k} \mu_i \overrightarrow{OA_i} = \sum_{i=1}^{k} v_i \overrightarrow{OA_i} + \overrightarrow{0} = \sum_{\substack{i=1 \\ i \neq q}}^{k} v_i \overrightarrow{OA_i}$.

Donc $M = \sum_{\substack{i=1 \\ i \neq q}}^{k} v_i A_i$, donc M est combinaison convexe de $(k-1)$ points de $\mathcal{A}$.

Donc M peut s'écrire (en itérant) comme combinaison convexe d'au plus $(1 + \dim \mathcal{E})$ points. ■

Corollaire

Sous les mêmes hypothèses :

1. si $\mathcal{A}$ est compact, alors $\text{Conv}(\mathcal{A})$ est compact ;
2. si $\mathcal{A}$ est borné, alors $\text{Conv}(\mathcal{A})$ est borné et de même diamètre que $\mathcal{A}$: $\delta(\mathcal{A}) = \delta(\text{Conv}(\mathcal{A}))$.

Démonstration :

1. On pose $n = \dim \mathcal{E}$, et $K = \left\{ (t_1, \dots, t_{n+1}) \in [0,1]^{n+1} \middle| t_1 + \dots + t_{n+1} = 1 \right\}$.
   K est compact ; on définit : $f : \left| \begin{array}{ccc} K \times \mathcal{E}^{n+1} & \to & \mathcal{E} \\ (t_1, \dots, t_{n+1}, A_1, \dots, A_{n+1}) & \mapsto & t_1 A_1 + \dots + t_{n+1} A_{n+1} \end{array} \right.$.
   D'après le théorème de Carathéodory, $f\left(K \times \mathcal{A}^{n+1}\right) = \text{Conv}(\mathcal{A})$.
   Or f est continue, et $K \times \mathcal{A}^{n+1}$ est compact, donc $\text{Conv}(\mathcal{A})$ est compact.
2. Comme $\mathcal{A} \subseteq \text{Conv}(\mathcal{A})$, on a : $\delta(\mathcal{A}) \leqslant \delta(\text{Conv}(\mathcal{A}))$.
   Comme $\mathcal{A}$ est borné, il existe $A \in \mathcal{A}$ et $r > 0$, tels que $\mathcal{A} \subset \overline{\mathcal{B}}(A, r)$.
   Comme $\overline{\mathcal{B}}(A, r)$ est convexe, on a : $\text{Conv}(\mathcal{A}) \subset \overline{\mathcal{B}}(A, r)$.
   Ainsi, $\text{Conv}(\mathcal{A})$ est borné.

   Soit $M \in \text{Conv}(\mathcal{A})$, $M = t_1 A_1 + \dots + t_k A_k$, où $A_1, \dots, A_k \in \mathcal{A}$, $t_1, \dots, t_k \geqslant 0$ et $\sum_{i=1}^{k} t_i = 1$.

Diffusion à titre gratuit uniquement.

Soit $N \in \mathcal{A}$, on a : $MN \leqslant t_1 A_1 N + \ldots + t_k A_k N \leqslant \delta(\mathcal{A})\,(t_1 + \ldots + t_k) = \delta(\mathcal{A})$.
Ainsi, la distance d'un point quelconque de $\mathcal{A}$ à un point quelconque de $\mathrm{Conv}(\mathcal{A})$ est inférieure à $\delta(\mathcal{A})$.
Soit alors $P \in \mathrm{Conv}(\mathcal{A})$, $MP \leqslant t_1 A_1 P + \ldots + t_k A_k P \leqslant \delta(\mathcal{A})\,(t_1 + \ldots + t_k) = \delta(\mathcal{A})$.
Puis, par passage à la borne supérieure, $\delta(\mathrm{Conv}(\mathcal{A})) \leqslant \delta(\mathcal{A})$ donc $\delta(\mathrm{Conv}(\mathcal{A})) = \delta(\mathcal{A})$. ■

Références

[TauGéo] P. TAUVEL – *Géométrie*, 2e éd., Dunod, 2005.

Diffusion à titre gratuit uniquement.

Théorème de Sophie Germain

Leçons : ***120, 121, 123, 126***

[X-ENS Al1], exercices 4.39

Théorème

Soit p un nombre premier de Sophie Germain, c'est-à-dire un nombre premier impair tel que $q = 2p + 1$ soit un nombre premier.
Alors il n'existe pas de triplet $(x, y, z) \in \mathbb{Z}^3$ tel que $xyz \not\equiv 0\,[p]$ et $x^p + y^p + z^p = 0$.[157]

Démonstration :

On notera ici $\mathcal{P}$ l'ensemble des nombres premiers.
On raisonne par l'absurde ; soit $(x, y, z) \in \mathbb{Z}^3$, tel que $xyz \not\equiv 0\,[p]$ et $x^p + y^p + z^p = 0$.
Soit $d = \text{pgcd}\,(x, y, z)$, quitte à poser $x' = \dfrac{x}{d}$, $y' = \dfrac{y}{d}$ et $z' = \dfrac{z}{d}$, on peut supposer que $d = 1$.

Étape 1 : Montrons qu'alors x, y et z sont premiers entre eux deux à deux.

Par l'absurde, soit r un facteur premier de x et y.
Alors $r|x^p + y^p$, puis $r|z^p$ et donc, par le lemme d'Euclide : $r|z$.
On contredit alors l'hypothèse selon laquelle x, y et z sont premiers entre eux dans leur ensemble.
Désormais, on a donc : $x \wedge y = x \wedge z = y \wedge z = 1$.

Étape 2 : Montrons que $\exists (a, \alpha) \in \mathbb{Z}^2, y + z = a^p$ et $\displaystyle\sum_{k=0}^{p-1} (-z)^{p-1-k} y^k = \alpha^p$.

On va appliquer le lemme suivant.

Lemme

Soit $u, v \in \mathbb{Z}$, avec $u \wedge v = 1$ et $\exists w \in \mathbb{Z}, uv = w^k$, où $k \geqslant 2$.
Alors u et v sont tous les deux des puissances $k^{\text{èmes}}$.

Démonstration :
On écrit $u = \displaystyle\prod_{p\in\mathcal{P}} p^{\alpha_p}$, $v = \prod_{p\in\mathcal{P}} p^{\beta_p}$ et $w = \prod_{p\in\mathcal{P}} p^{\gamma_p}$, où $\alpha, \beta, \gamma \in \mathbb{N}^{(\mathcal{P})}$.
Et comme $uv = w^k$, on a : $\forall p \in \mathcal{P}, \alpha_p + \beta_p = k\gamma_p$.
Mais, α_p et β_p ne peuvent pas être simultanément non-nuls, puisqu'on a $u \wedge v = 1$.
Conséquemment, $\forall p \in \mathcal{P}$, $k|\alpha_p$ et $k|\beta_p$.
Donc u et v sont des puissances $k^{\text{èmes}}$. ■

Ici, on a $(y + z)\left(\displaystyle\sum_{k=0}^{p-1} (-z)^{p-1-k} y^k\right) = y^p + z^p = -x^p = (-x)^p$.

Il serait donc intéressant de montrer que $y + z$ et $\displaystyle\sum_{k=0}^{p-1} (-z)^{p-1-k} y^k$ sont premiers entre eux.

Par l'absurde, supposons qu'il existe un nombre premier, appelons-le r, qui les divise tous les deux.
Alors, de l'égalité précédente, il vient que $r^2|x^p$, donc $r|x$.
Comme $y \equiv -z\,[r]$, on a : $\underbrace{\displaystyle\sum_{k=0}^{p-1} (-z)^{p-1-k} y^k}_{\equiv 0\,[r]} \equiv \sum_{k=0}^{p-1} y^{p-1} \equiv p y^{p-1}\,[r]$.

157. Il y a beaucoup de choses intéressantes à dire à propos de ce résultat. Sophie Germain (1776-1831) est quasiment la seule femme mathématicienne de son temps. Elle suivit les cours de l'École polytechnique par correspondance, car les femmes n'y étaient pas admises et c'est sous le pseudonyme masculin de Maurice Leblanc qu'elle écrivait à Gauss pour lui faire part de ses découvertes arithmétiques. En 2001, le plus grand nombre de Sophie Germain qu'on connaissait était $109433307 \times 2^{66452} - 1$, possédant 20013 chiffres. À l'heure actuelle, on conjecture qu'il en existe une infinité. Le théorème de Sophie Germain, démontré en 1823, est une résolution partielle du grand théorème de Fermat — mais si, vous savez : pour $n \geqslant 3$, il n'existe pas de solution non-triviale dans $\mathbb{Z}^3$ à l'équation $x^n + y^n = z^n$ — que Fermat mentionnait dans une annotation marginale, sans la prouver "par manque de place". On est certain aujourd'hui qu'il ne pouvait pas en avoir une démonstration complète (bon, en même temps, quand tu t'appelles Fermat, ton prof de maths va avoir du mal à te reprocher de bluffer dans tes copies, non ?).

Diffusion à titre gratuit uniquement.

Donc $r|py^{p-1}$, et donc, par le lemme de Gauss :
– soit $r|p$, et alors, ces deux nombres étant premiers, on obtient $r = p$ et donc $p|x$, contredisant l'hypothèse $xyz \not\equiv 0\,[p]$;
– soit $r|y$, mais c'est impossible puisque $r|x$ et $x \wedge y = 1$.

On obtient ainsi une contradiction ; et on en déduit $(y+z) \wedge \left(\sum_{k=0}^{p-1} (-z)^{p-1-k} y^k\right) = 1$.

Puis, par le lemme, on obtient :

$$\exists (a,\alpha) \in \mathbb{Z}^2, y+z = a^p \text{ et } \sum_{k=0}^{p-1} (-z)^{p-1-k} y^k = \alpha^p.$$

Similairement, on montrerait $x+z = b^p$ et $y+z = c^p$, avec $b, c \in \mathbb{Z}$.

Étape 3 : Un (et un seul, vu qu'ils sont premiers entre eux deux à deux) des trois entiers x, y et z est divisible par q.
Soit $m \in \mathbb{Z}$, tel que $q \nmid m$.
Alors, par le petit théorème de Fermat, on obtient : $(m^p)^2 = m^{q-1} \equiv 1\,[q]$ et donc, comme $\mathbb{Z}/q\mathbb{Z}$ est un corps [158], on a : $m^p \equiv \pm 1\,[q]$.
Par l'absurde, on suppose $q \nmid x$, $q \nmid y$ et $q \nmid z$.
Alors $0 = x^p + y^p + z^p$ est congru à 3, 1, -1 ou -3 modulo q. Ce qui est absurde puisque $q > 5$.
Sans perte de généralité, disons que $q|x$, et qu'incidemment : $q \nmid y$ et $q \nmid z$.

Étape 4 : Tels Jean-Claude Dusse, cherchons à conclure.
On a : $b^p + c^p - a^p = x + z + x + y - y - z = 2x \equiv 0\,[q]$.
Et comme $x \equiv 0\,[q]$, on a : $y \equiv c^p\,[q]$; mais $q \nmid y$ donc $q \nmid c$, d'où $y \equiv \pm 1\,[q]$. Similairement, $z \equiv \pm 1\,[q]$.
Donc $a^p = y + z$ est congru à 2, 0 ou -2 modulo q ; mais une puissance $p^{\text{ème}}$ est congrue à 1, 0 ou -1 modulo q.
Donc $y + z \equiv 0[q]$.
Comme dans l'étape 2, on obtient : $\alpha^p = \sum_{k=0}^{p-1} (-z)^{p-1-k} y^k \equiv p y^{p-1}\,[q]$.
Or $p-1$ est pair et $y \equiv \pm 1\,[q]$ et donc $\alpha^p \equiv p\,[q]$; mais aussi α^p est congru à 1, 0 ou -1 modulo q.
Contradiction !
Il n'y a donc pas de triplet $(x,y,z) \in \mathbb{Z}^3$ tel que : $xyz \not\equiv 0\,[q]$ et $x^p + y^p + z^p = 0$. ■

Références

[X-ENS Al1] S. FRANCINOU, H. GIANELLA et S. NICOLAS – *Oraux X-ENS Algèbre 1*, 3ème éd., Cassini, 2014.

158. Le polynôme $X^2 - 1 \in \mathbb{Z}/q\mathbb{Z}[X]$ admet au plus deux racines puisqu'il est de degré 2 sur un corps ; on vérifie facilement qu'il s'agit de 1 et -1.

Diffusion à titre gratuit uniquement.

Théorème des événements rares de Poisson

Leçons : ***218, 241, 249, 261, 262, 264***

[Ouv 1], théorème 7.1
[Ouv 2], théorème 14.20

Théorème (*Événements rares*)

Soit, pour tout $n \in \mathbb{N}^*$, une famille finie $\left\{A_{n,j} | 1 \leqslant j \leqslant M_n\right\}$ d'événements indépendants.
On pose : $p_{n,j} = \mathbb{P}\left(A_{n,j}\right)$ et $S_n = \sum_{j=1}^{M_n} \mathbb{1}_{A_{n,j}}$.
On suppose [159] : $\max_{1\leqslant j\leqslant M_n} p_{n,j} \underset{n\to\infty}{\longrightarrow} 0$ et $\sum_{j=1}^{M_n} p_{n,j} \underset{n\to\infty}{\longrightarrow} \lambda > 0$.
Alors $(S_n)_{n\in\mathbb{N}^*}$ converge en loi vers la loi $\mathcal{P}(\lambda)$.

On commence par montrer le théorème de Poisson, dont le théorème des événements rares est une généralisation.

Théorème (*Poisson*)

On considère $(S_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires de lois $\mathcal{B}(n, p_n)$.
Si $\lim_{n\to\infty} np_n = \lambda > 0$,
Alors $(S_n)_{n\in\mathbb{N}^*}$ converge en loi vers la loi $\mathcal{P}(\lambda)$.

Démonstration du théorème de Poisson :

Comme $np_n \underset{n\to\infty}{\longrightarrow} \lambda$, on a : $p_n = \frac{\lambda}{n} + o\left(\frac{1}{n}\right)$.

Soit $k \in \mathbb{N}$, pour $n \geqslant k$:

$$\mathbb{P}(S_n = k) = \binom{n}{k} p_n^k (1-p_n)^{n-k} = \frac{n(n-1)\dots(n-k+1)}{k!}\left[\frac{\lambda}{n} + o\left(\frac{1}{n}\right)\right]^k \left[1 - \frac{\lambda}{n} + o\left(\frac{1}{n}\right)\right]^{n-k}$$

Or $n(n-1)\dots(n-k+1)\left[\frac{\lambda}{n} + o\left(\frac{1}{n}\right)\right]^k = \frac{n}{n}\frac{n-1}{n}\dots\frac{n-k+1}{n}[\lambda + o(1)]^k \underset{n\to\infty}{\longrightarrow} \lambda^k$

Et $\left[1 - \frac{\lambda}{n} + o\left(\frac{1}{n}\right)\right]^{n-k} = \exp\left[(n-k)\ln\left(1 - \frac{\lambda}{n} + o\left(\frac{1}{n}\right)\right)\right] = \exp\left[(n-k)\left(-\frac{\lambda}{n} + o\left(\frac{1}{n}\right)\right)\right] \underset{n\to\infty}{\longrightarrow} \mathrm{e}^{-\lambda}$

Par conséquent, $\mathbb{P}(S_n = k) \underset{n\to\infty}{\longrightarrow} \frac{\lambda^k}{k!}\mathrm{e}^{-\lambda}$, d'où $S_n \underset{n\to\infty}{\xrightarrow{\mathcal{L}}} \mathcal{P}(\lambda)$. ■

Démonstration du théorème des événements rares :

On va utiliser le théorème de Lévy.

Soit $n \in \mathbb{N}^*$, par indépendance des $A_{n,j}$ pour $1 \leqslant j \leqslant M_n$, on a, pour $t \in \mathbb{R}$:

$$\varphi_{S_n}(t) = \prod_{j=1}^{M_n} \varphi_{\mathbb{1}_{A_{n,j}}}(t) = \prod_{j=1}^{M_n}\left(p_{n,j}\mathrm{e}^{\mathrm{i}t} + 1 - p_{n,j}\right) = \prod_{j=1}^{M_n}\left(1 + p_{n,j}\left(\mathrm{e}^{\mathrm{i}t} - 1\right)\right)$$

D'où, en posant $z = \mathrm{e}^{\mathrm{i}t} - 1$:

$$\mathrm{Log}\left(\varphi_{S_n}(t)\right) = \sum_{j=1}^{M_n} \mathrm{Log}\left(1 + p_{n,j}z\right)$$

Par la formule de Taylor avec reste intégral, pour $|z| < 1$:

$$\mathrm{Log}(1+z) = z + \int_1^{1+z} \frac{(1+z-v)^1}{1!}\frac{-1}{v^2}\,\mathrm{d}v = z + \int_0^1 (1+z-1-zu)\frac{-1}{(1+zu)^2}z\,\mathrm{d}u = z - z^2\int_0^1 \frac{1-u}{(1+zu)^2}\,\mathrm{d}u$$

159. Attention, certaines hypothèses dans le livre de Ouvrard sont inutiles.

Diffusion à titre gratuit uniquement.

Comme $\max\limits_{1\leqslant j\leqslant M_n} p_{n,j} \underset{n\to\infty}{\longrightarrow} 0 : \exists N \in \mathbb{N}, \forall n \geqslant N, \max\limits_{1\leqslant j\leqslant M_n} \left|p_{n,j}z\right| < \dfrac{1}{2}$.
Soit alors $n \geqslant N$,

$$\operatorname{Log}\varphi_{S_n}(t) = \sum_{j=1}^{M_n}\left[p_{n,j}z - p_{n,j}^2 z^2 \int_0^1 \frac{1-u}{\left(1+p_{n,j}zu\right)^2}\,\mathrm{d}u\right] = z\sum_{j=1}^{M_n} p_{n,j} - z^2\sum_{j=1}^{M_n} p_{n,j}^2 \int_0^1 \frac{1-u}{\left(1+p_{n,j}zu\right)^2}\,\mathrm{d}u$$

Or, pour $u \in [0,1]$, par inégalité triangulaire : $\left|1+p_{n,j}zu\right| \geqslant 1 - p_{n,j}|z|u \geqslant 1 - p_{n,j}|z| \geqslant \dfrac{1}{2}$.

$$\text{Donc } \left|\sum_{j=1}^{M_n} p_{n,j}^2 \int_0^1 \frac{1-u}{\left(1+p_{n,j}zu\right)^2}\,\mathrm{d}u\right| \leqslant \sum_{j=1}^{M_n} p_{n,j}^2 \int_0^1 \frac{|1-u|}{\left|1+p_{n,j}zu\right|^2}\,\mathrm{d}u \leqslant \sum_{j=1}^{M_n} p_{n,j}^2 4\int_0^1 (1-u)\,\mathrm{d}u = 2\sum_{j=1}^{M_n} p_{n,j}^2$$
$$\leqslant 2\left(\max_{1\leqslant j\leqslant M_n} p_{n,j}\right)\left(\sum_{j=1}^{M_n} p_{n,j}\right)$$

Ainsi $\lim\limits_{n\to\infty}\sum\limits_{j=1}^{M_n} p_{n,j}^2 \int_0^1 \dfrac{1-u}{\left(1+p_{n,j}zu\right)^2}\,\mathrm{d}u = 0$ donc $\lim\limits_{n\to\infty} \operatorname{Log}\varphi_{S_n}(t) = \lim\limits_{n\to\infty} z\sum\limits_{j=1}^{M_n} p_{n,j} = \lambda z$.

Donc $\lim\limits_{n\to\infty} \varphi_{S_n}(t) = \exp\left(\lambda\left(\mathrm{e}^{\mathrm{i}t}-1\right)\right)$ d'où $S_n \xrightarrow[n\to\infty]{\mathcal{L}} \mathcal{P}(\lambda)$. ■

Références

[Ouv 1] J.-Y. OUVRARD – *Probabilités 1*, 2e éd., Cassini, 2007.
[Ouv 2] J.-Y. OUVRARD – *Probabilités 2*, 3e éd., Cassini, 2009.

Diffusion à titre gratuit uniquement.

Anneaux euclidiens, principaux, factoriels

Développements : Étude de l'anneau $\mathbb{Z}\left[\frac{1+i\sqrt{19}}{2}\right]$ (page 16), Irréductibilité des polynômes cyclotomiques (page 23)

[Per], parties II.3,4,5
[RDO], partie 3.3.3

Définition

Soit A un anneau.
- A est euclidien $\Leftrightarrow$ A est intègre et
$$\exists \nu : A\backslash\{0\} \to \mathbb{N}, \forall a, b \in A\backslash\{0\}, \exists q, r \in A, a = bq + r \text{ et } (r = 0 \text{ ou } \nu(r) < \nu(b)).$$
- A est principal $\Leftrightarrow$ A est intègre et tout idéal de A est principal.
- P est un système de représentants des irréductibles (sri) de A $\Leftrightarrow$
$$\forall p \in A \text{ irréductible}, \exists! q \in P, p \text{ et } q \text{ sont associés.}$$
- Soit P un sri de A.
A est factoriel $\Leftrightarrow$ A est intègre et $\forall a \in A\backslash\{0\}, \exists! u \in A^\times, \exists! \left(v_p(a)\right)_{p\in P} \in \mathbb{N}^{(P)}, a = u\prod_{p\in P} p^{v_p(a)}$.

Proposition

Un anneau euclidien est principal.

Démonstration :
Soit I un idéal de l'anneau euclidien A, avec $I \neq (0)$.
Soit alors $b \in I$, tel que $\nu(b)$ soit minimal.
Soit $a \in I$, on a la division euclidienne $a = bq + r$, avec $r = 0$ ou $\nu(r) < \nu(b)$.
Or $r \in I$, donc par minimalité de $\nu(b)$ on a $r = 0$, puis $a \in (b)$ et enfin $I = (b)$. ■

Par exemple, $\mathbb{Z}$, muni de $\nu(n) = |n|$, est euclidien.

Proposition

Un anneau principal est factoriel.

Démonstration :
On fixe P un sri de l'anneau principal A.

Étape 1 : Dans un anneau principal, il n'existe pas de suite d'idéaux qui soit strictement croissante (pour l'inclusion).
En effet, par l'absurde, considérons une suite d'idéaux strictement croissante $(J_n)_{n\in\mathbb{N}}$.
Alors, on montre sans effort que $J = \bigcup_{n\in\mathbb{N}} J_n$ est aussi un idéal.
Soit alors $a \in A$, tel que $J = (a)$ (on a le droit car A est principal).
Alors $\exists p \in \mathbb{N}, a \in J_p$, d'où $J \subset J_p$, puis $J = J_p$.
Dès lors $\forall n \geqslant p, J_n = J_p$. Contradiction avec la stricte croissance de la suite.

Étape 2 : On pose alors $\mathcal{E} = \{\text{idéaux non-nuls de } A \text{ dont les générateurs n'admettent pas de décomposition}\}$.
- On montre d'abord par l'absurde que $\mathcal{E} \neq \varnothing$.
Soit $J_0 \in \mathcal{E}$, on construit une suite *finie* strictement croissante d'éléments de $\mathcal{E}$, la plus longue possible.
Soit alors J_p son plus grand élément, et $a \in A$ tel que $(a) = J_p$; a n'est donc pas décomposable.
En conséquence, a n'est pas irréductible, donc $\exists b, c \in A\backslash A^\times, a = bc$.
Donc $(b), (c) \supset (a) = J_p$ et $(b), (c) \neq (a)$; par maximalité de J_p, on a que $(b), (c) \notin \mathcal{E}$.
Donc b et c sont décomposables, et donc $a = bc$ l'est aussi. Contradiction.
- Pour l'unicité, on suppose que $u\prod_{p\in P} p^{v_p} = u'\prod_{p\in P} p'^{v'_p}$.
Soit $p_0 \in P$, on a donc $p_0^{v_{p_0}} \left| p_0^{v'_{p_0}}\ u'\prod_{\substack{p\in P\\ p\neq p_0}} p^{v'_p}\right.$; par le lemme de Gauss, on obtient que $p_0^{v_{p_0}} \left| p_0^{v'_{p_0}}\right.$.
Donc $v_{p_0} \leqslant v'_{p_0}$, et par symétrie $v_{p_0} = v'_{p_0}$, puis $v = v'$ et $u = u'$.

Diffusion à titre gratuit uniquement.

■

Proposition

$A[X]$ est principal $\Leftrightarrow$ A est un corps.

Démonstration :

$\Leftarrow$ Si A est un corps, alors $A[X]$ est un anneau euclidien pour le degré.

$\Rightarrow$ Si $A[X]$ est principal, alors $A[X]$ est intègre donc A est intègre (cela se montre facilement).
On montre alors que X est irréductible (en utilisant le fait que $\deg PQ = \deg P + \deg Q$, quand l'anneau des coefficients est intègre).
Donc l'idéal (X) est premier dans l'anneau $A[X]$ qui est factoriel donc il est maximal.
Mais $A[X]$ est principal donc $A \simeq {}^{A[X]}/_{(X)}$ est un corps. ■

Proposition

Si A est un anneau factoriel, alors $A[X]$ est un anneau factoriel.

Démonstration :

Lemme (*Gauss*)

On note, pour $P \in A[X]$, $c(P)$ le pgcd des coefficients de P (on utilise ici la factorialité de A), qu'on appelle contenu de P.
Alors on a $c(PQ) = c(P)c(Q)$ modulo $A^\times$.

Démonstration :

On se ramène d'abord au cas où P et Q sont primitifs puis on obtient le cas général.
On note $P = a_n X^n + \ldots + a_0$ et $Q = b_m X^m + \ldots + b_0$.
Si $c(PQ) \neq 1$, il existe $p \in A$ irréductible divisant $c(PQ)$.
P et Q étant primitifs, on note i_0 et j_0 les rangs des coefficients de P et Q de plus petits degrés tels que $p \nmid a_{i_0}$ et $p \nmid b_{j_0}$.
Mais, par hypothèse : $p | c_{i_0+j_0} = a_0 b_{i_0+j_0} + \ldots + a_{i_0} b_{j_0} + \ldots + a_{i_0+j_0} b_0$ donc $p | a_{i_0}$ ou b_{j_0}. Impossible. ■

Lemme

Les irréductibles de $A[X]$ sont les :
- éléments irréductibles de A ;
- polynômes primitifs de $A[X]$ qui sont irréductibles dans $\mathrm{Frac}(A)[X]$.

Démonstration :

Il faut utiliser le lemme de Gauss mais c'est pas trop difficile. ■

Soit $P \in A[X]$ un polynôme primitif ; par factorialité de $\mathrm{Frac}(A)[X]$, on peut écrire $P = \prod_{i \in I} P_i^{\alpha_i}$ dans $\mathrm{Frac}(A)[X]$, où les P_i sont des irréductibles de $\mathrm{Frac}(A)[X]$.
On peut écrire $P_i = \frac{a_i}{b_i} Q_i$ avec $Q_i \in A[X]$ un polynôme primitif ; l'irréductibilité de Q_i dans $\mathrm{Frac}(A)[X]$ implique que Q_i est irréductible dans $A[X]$.
On a : $\left(\prod_{i \in I} b_i^{\alpha_i}\right) P = \left(\prod_{i \in I} a_i^{\alpha_i}\right) \left(\prod_{i \in I} Q_i^{\alpha_i}\right)$, et donc, au contenu, on a, modulo $A^\times$, $\prod_{i \in I} b_i^{\alpha_i} \equiv \prod_{i \in I} a_i^{\alpha_i}$.
Donc $P = u \prod_{i \in I} Q_i^{\alpha_i}$, où $u \in A^\times$.
Si P n'est pas primitif, on le divise par son contenu pour se ramener au cas primitif.

Lemme

Soit B un anneau intègre, dans lequel on a l'existence d'une décomposition qui le rendrait factoriel, sans en avoir l'unicité. On a équivalence entre :

1. B est factoriel ;
2. Si p est irréductible et $p | ab$, alors $p | a$ ou $p | b$;
3. Si $a | bc$ et si a et b sont premiers entre eux, alors $a | c$.

Diffusion à titre gratuit uniquement.

Démonstration :

3 ⇒ 2 : évident.

2 ⇒ 1 : on suppose que $u\prod_{p\in P} p^{v_p} = u'\prod_{p\in P} p'^{v'_p}$.

Soit $p_0 \in P$, on a donc $p_0^{v_{p_0}} \left| p_0^{v'_{p_0}}\ u'\prod_{\substack{p\in P\\ p\neq p_0}} p^{v'_p}\right.$; par le lemme de Gauss, on obtient que $p_0^{v_{p_0}} \left| p_0^{v'_{p_0}}\right.$.

Donc $v_{p_0} \leqslant v'_{p_0}$, et par symétrie $v_{p_0} = v'_{p_0}$, puis $v = v'$ et $u = u'$. On obtient donc l'unicité de la décomposition.

1 ⇒ 3 : on décompose a, b et c en facteurs premiers, et on veut montrer que $v_p(a) \leqslant v_p(c)$ pour tout $p \in B$ irréductible.
Sinon, pour un p, $v_p(a) > v_p(c)$, mais comme $a|bc$, $v_p(b) \geqslant v_p(a) - v_p(c)$ donc $v_p(b)$ et $v_p(a)$ sont non-nuls.
Dès lors $p|a$ et $p|b$, d'où une contradiction. ■

On prend P un polynôme irréductible, et on montre que ${}^{A[X]}\!/_{(P)}$ est intègre, ce qui donnera que (P) est premier.
Si $P = p \in A$, alors ${}^{A[X]}\!/_{(p)} \simeq {}^{A}\!/_{(p)}[X]$ est intègre car ${}^{A}\!/_{(p)}$ l'est.
Sinon, on a i : ${}^{A[X]}\!/_{(P)} \to {}^{\mathrm{Frac}(A)[X]}\!/_{(P)}$ et ${}^{\mathrm{Frac}(A)[X]}\!/_{(P)}$ est intègre ; on va montrer que i est injectif, autrement dit : $(P\,\mathrm{Frac}(A)[X]) \cap A[X] = P\,A[X]$.
Soit $Q \in A[X]$ et $Q = PR$ avec $R \in \mathrm{Frac}(A)[X]$; on écrit $R = \frac{a}{b}R'$ et $Q = cQ'$ avec R' et Q' dans $A[X]$ et primitifs ; on a : $cbQ' = aPR'$.
Donc en passant au contenu, $b|a$, donc $R \in A[X]$.
L'autre inclusion est triviale, et le lemme précédent permet de montrer le théorème. ■

$\mathbb{Z}\left[\frac{1+i\sqrt{19}}{2}\right]$ et ${}^{\mathbb{R}[X,Y]}\!/_{(X^2+Y^2+1)}$ sont des exemples d'anneaux principaux non-euclidiens (mais c'est dur à montrer).

On termine avec un critère pour montrer qu'un anneau n'est pas euclidien.

Proposition

Soit A un anneau euclidien.
$\exists x \in A\backslash A^\times$, $\pi_{|A^\times\cup\{0\}}$ est surjective sur ${}^{A}\!/_{(x)}$, où $\pi : A \to {}^{A}\!/_{(x)}$ est la projection canonique.

Démonstration :

Si A est un corps, alors $x = 0$ convient.
Sinon, soit $x \in A\backslash\left(A^\times \cup \{0\}\right)$ tel que $\nu(x)$ soit minimal.
Soit $a \in A, a = xq + r$ avec $r = 0$ ou $\nu(r) < \nu(x)$.
Si $r \neq 0$, par minimalité de $\nu(x)$, on a : $r \in A^\times$ et donc, modulo (x), $a \equiv r \in A^\times \cup \{0\}$. ■

Il en découle notamment que ${}^{A}\!/_{(x)}$ est un corps, donc (x) est maximal. Ce critère permet ainsi de montrer que $\mathbb{Z}\left[\frac{1+i\sqrt{19}}{2}\right]$ n'est pas euclidien. [160]

Références

[Per] D. PERRIN – *Cours d'algèbre*, Ellipses, 1996.
[RDO] E. RAMIS, C. DESCHAMPS et J. ODOUX – *Cours de mathématiques spéciales (Algèbre 1)*, 2ème éd., Masson, 1993.

160. Voir le développement consacré en page 16.

Diffusion à titre gratuit uniquement.

Groupe multiplicatif d'un corps fini

Développements : Polygones réguliers constructibles (page 25), Frobenius-Zolotarev (page 42), Chevalley-Warning et Erdös-Ginzburg-Ziv (page 50)

[Ser], partie 1.2

Théorème

Soit p un nombre premier et $r \in \mathbb{N}^*$; on note $q = p^r$.
Le groupe multiplicatif $\mathbb{F}_q^\times$ du corps fini $\mathbb{F}_q$ est cyclique d'ordre $(q-1)$.

Démonstration :

On a, pour $d \geqslant 1$, $\varphi(d) = \#\{x \in [\![1,d]\!] | x \wedge d = 1\} = \#\left\{x \in \mathbb{Z}/d\mathbb{Z} \middle| < x >= \mathbb{Z}/d\mathbb{Z}\right\}$.

Lemme 1

Soit $n \in \mathbb{N}^*$.
On a l'égalité suivante : $n = \sum_{d|n} \varphi(d)$.

Démonstration du lemme 1 :

Soit, pour $d|n$, C_d l'unique sous-groupe de $\mathbb{Z}/n\mathbb{Z}$ d'ordre d.
On note Φ_d l'ensemble des générateurs de C_d.
Tout élément de $\mathbb{Z}/n\mathbb{Z}$ engendre un des C_d, où $d|n$. Ainsi on a :

$$n = \#\mathbb{Z}/n\mathbb{Z} = \#\left(\bigsqcup_{d|n} \Phi_d\right) = \sum_{d|n} \#\,\Phi_d = \sum_{d|n} \varphi(d)$$

■

Lemme 2

Soit H un groupe d'ordre $n < \infty$. On suppose que : $\forall d|n, \#\left\{x \in H \middle| x^d = 1\right\} \leqslant d$.
Alors H est cyclique.

Démonstration du lemme 2 :

Soit $d|n$, et $x \in H$ d'ordre d (s'il en existe). Alors $< x >= \left\{1, x, \ldots, x^{d-1}\right\}$ est cyclique d'ordre d.
Par hypothèse, si $y \in H$ vérifie $y^d = 1$, alors $y \in < x >$.
Précisément, les seuls éléments de H qui soient d'ordre d sont les générateurs de $< x >$. Il y en a $\varphi(d)$.
Résumons :

- Soit il existe $x \in H$ d'ordre d, et il y a $\varphi(d)$ éléments d'ordre d ;
- Soit il n'en existe pas et il y a 0 élément d'ordre d.

Supposons que pour un $d|n$, il y ait 0 élement d'ordre d.
Alors on aurait $\#H < \sum_{d|n} \varphi(d) = n$ d'après le lemme 1.
On aboutit donc à une contradiction avec l'hypothèse $\#H = n$.
En particulier, il existe $x \in H$ qui soit d'ordre n.
Ainsi $H =< x >$ est cyclique. ■

On applique le lemme 2 à $H = \mathbb{F}_q^\times$, $n = q-1$.
$\mathbb{F}_q$ étant un anneau intègre, on a que, pour tout $d|n$, le polynôme $X^d - 1$ possède au plus d racines dans $\mathbb{F}_q$, par conséquent, $\#\left\{x \in \mathbb{F}_q^\times \middle| x^d = 1\right\} \leqslant d$.
Ainsi, $\mathbb{F}_q^\times$ est cyclique :

$$\mathbb{F}_q^\times \simeq \mathbb{Z}/(q-1)\mathbb{Z}$$

■

Références

[Ser] J.-P. SERRE – *Cours d'arithmétique*, 1e éd., Presses Universitaires de France, 1970.

Diffusion à titre gratuit uniquement.