École Normale Supérieure Paris-Saclay

AGRÉGATION – DÉVELOPPEMENTS

Antoine BARRIER

antoinebarrier@free.fr

http://antoinebarrier.free.fr/fr/documents/#agregation

2018/2019 – Dernière mise à jour : 21 novembre 2024

Ce document regroupe mes développements, préparés en 2018/2019 au cours de mon année de préparation de l'Agrégation de Mathématiques (master M2 FESup de l'ÉNS Paris-Saclay).

À l'été 2024, j'ai repris l'ensemble des développements afin de corriger les erreurs et d'améliorer la rédaction et la mise en page. En utilisant ce document, veuillez à garder à l'esprit que le format qui en découle n'est pas nécessairement celui attendu le jour de l'épreuve : les développements sont rédigés pour être les plus complets possibles, avec la contrainte d'une page maximum chacun, et leur nouvelle version est écrite cinq ans après avoir passé le concours, avec le recul de plus de quatre années d'enseignement.

Préparer, travailler et répéter ses développements est primordial pour bien aborder les épreuves orales. Il est très important de s'approprier les développements afin d'être dans les meilleurs conditions le jour J. En particulier, il faut avoir ses propres notations et s'assurer que celles-ci concordent avec toutes les leçons concernées.

Sommaire

Première partie

Développements de Mathématiques Générales

COUPLAGES

#	Développement	Leçon
1	Caractérisation des endomorphismes semi-simples	122, 154
2	Cardinal de $\mathcal{D}_n(\mathbb{F}_q)$	101, 106, 150, 155, 190
3	Critère d'EISENSTEIN	141, 142
4	Décomposition de DUNFORD et calcul de l'exponentielle d'une matrice	153, 155, 156 , 157
5	Degré de $\mathbb{Q}[\{\sqrt{p_i}\}_{1\leq i\leq n}]$ sur $\mathbb{Q}$	125, 151
6	Déterminant de GRAM et inégalité de HADAMARD	152, 161
7	Factorisations LU et de CHOLESKY	162
8	Formes de HANKEL	144, 170, 171
9	Formule de POISSON discrète	110
10	Homéomorphisme de l'exponentielle	156, 158, 160
11	Irréductibilité de Φ_n	102, 120, 121, 123, 125, 141
12	Isométries du cube	105, 161, 183
13	Isomorphisme $\mathrm{SU}_2(\mathbb{C})/\{\pm I_2\} \simeq \mathrm{SO}_3(\mathbb{R})$	106, 108, 182, 183
14	Lemme de MORSE	158
15	Loi de réciprocité quadratique	120, 121, 123, 126 , 170, 171, 190
16	Méthode du gradient à pas optimal pour la fonctionnelle quadratique	162
17	Réduction de JORDAN	151, 153, 154, 157 , 159
18	Réduction des endomorphismes normaux	150, 160
19	Simplicité de $\mathfrak{A}_n$ pour $n \geq 5$	103, 104, 105, 108
20	Sous-groupes distingués et caractères. Table de $\mathfrak{S}_4$	103, 107
21	Structure des groupes abéliens finis	107, 110
22	Suites de polygones	152, 181, 182
23	Théorème de CARATHÉODORY	181
24	Théorème de KRONECKER	102, 144
25	Théorème de Sophie GERMAIN	142
26	Théorème de SYLOW	101, 104
27	Théorème des deux carrés	122, 126
28	Théorème des extrema liés	159

1 CARACTÉRISATION DES ENDOMORPHISMES SEMI-SIMPLES

[Gou09, §4.5, p224–226]

ÉNONCÉ

Soient $\mathbb{K}$ un corps commutatif puis E un $\mathbb{K}$-espace vectoriel de dimension finie.

Théorème. *Un endormorphisme $f \in \mathcal{L}(E)$ est semi-simple si et seulement si son polynôme minimal π_f est sans facteur carré.*

DÉVELOPPEMENT

Soit $f \in \mathcal{L}(E)$. Par factorialité, écrivons la décomposition de son polynôme minimal en facteurs irréductibles de $\mathbb{K}[X]$: il existe $r \in \mathbb{N}$, des polynômes $(P_k)_{1\leq k\leq r} \in \mathbb{K}[X]$ distincts, unitaires et irréductibles, et des multiplicités $(m_k)_{1\leq k\leq r} \in (\mathbb{N}^*)^r$ tels que

$$\pi_f = \prod_{k=1}^{r} P_k^{m_k}\,, \qquad \text{et alors} \qquad E = \bigoplus_{k=1}^{r} \ker P_k^{m_k}(f)$$

d'après le lemme des noyaux.

Montrons le lemme suivant, où $N_k = \ker P_k^{m_k}(f)$ pour $k \in [\![1\,;r]\!]$.

Lemme. *Soit F un sous-espace stable de E. Posons $F_k = F \cap N_k$ pour $k \in [\![1\,;r]\!]$. Alors*

$$F = \bigoplus_{k=1}^{r} F_k\,.$$

Soit F un sous-espace stable de E. Déjà, il est clair que les $(F_k)_{1\leq k\leq r}$ sont en somme directe puis, comme les $(F_k)_{1\leq k\leq r}$ sont tous inclus dans F, que $\bigoplus_{k=1}^{r} F_k \subset F$.

Réciproquement, soit $k \in [\![1\,;r]\!]$. La projection p_k sur N_k parallèlement à $\bigoplus_{\ell\neq k} N_\ell$ est un polynôme en f, d'où $p_k(F) \subset F$ puisque F est f-stable. Par ailleurs $p_k(F) \subset \mathrm{Im}(p_k) = N_k$, et alors $p_k(F) \subset F \cap N_k = F_k$. Comme $\sum_{k=1}^{r} p_k = \mathrm{Id}_E$, on en déduit que :

$$F = \sum_{k=1}^{r} p_k(F) \subset \bigoplus_{k=1}^{r} F_k\,.$$

Lemme. *Supposons π_f irréductible. Alors f est semi-simple.*

Soit F un sous-espace f-stable distinct de E. Prenons $x_1 \notin F$ et soit $E_1 = \mathbb{K}[f](x_1)$. C'est un sous-espace f-stable. Comme le polynôme minimal local $\pi_{f,\,x_1}$ est le polynôme minimal de $f_{|E_1}$, $\pi_{f,\,x_1} \mid \pi_f$ at donc $\pi_{f,\,x_1} = \pi_f$ par irréductibilité.

Supposons qu'il existe $y = P(f)(x_1) \in E_1 \cap F$ non nul. Comme $\pi_{f,\,x_1} \nmid P$ et $\pi_{f,\,x_1}$ est irréductible, P et $\pi_{f,\,x_1}$ sont premiers entre eux et par l'identité de Bézout il existe $U, V \in \mathbb{K}[X]$ tels que $UP + V\pi_{f,\,x_1} = 1$, et alors $x_1 = U(f)(y) \in F$ puisque F est f-stable, ce qui est absurde.

Ainsi E_1 et F sont en somme directe. Si $E_1 \oplus F = E$, on a trouvé un supplémentaire f-stable. Sinon on choisit $x_2 \in E \setminus (E_1 \oplus F)$ et on recommence. En un nombre fini d'itérations, on obtient un sous-espace $E_1 \oplus \cdots \oplus E_L$ qui est f-stable et supplémentaire de F.

Montrons désormais le théorème par double implication.

- Supposons π_f sans facteur carré.
  Soit F un sous-espace vectoriel f-stable. Écrivons, d'après le premier lemme,

  $$F = \bigoplus_{k=1}^{r} F \cap N_k\,, \qquad \text{où} \quad N_k = \ker P_k(f) \text{ pour } k \in [\![1\,;r]\!]\,.$$

  Fixons $k \in [\![1\,;r]\!]$. Comme N_k est f-stable, $f_{|N_k}$ admet pour polynôme minimal P_k qui est irréductible. Le second lemme assure qu'il existe un sous-espace G_k supplémentaire de $F_k = F \cap N_k$ dans N_k et qui est f-stable. Alors :

  $$E = \bigoplus_{k=1}^{r}(F_k \oplus G_k) = \Big(\bigoplus_{k=1}^{r} F_k\Big) \bigoplus \Big(\bigoplus_{k=1}^{r} G_k\Big) = F \oplus G\,,$$

  où $G = \bigoplus_{k=1}^{r} G_k$ est f-stable puisque les $(G_k)_{1\leq k\leq r}$ le sont tous.
  Donc f est semi-simple.
- Supposons f semi-simple et l'un des $(m_k)_{1\leq k\leq r}$ strictement supérieur à 1, disons m_K.
  Posons $F = \ker P_K(f)$ qui est f-stable.
  L'endormorphisme f étant semi-simple, F admet un supplémentaire f-stable G.
  Soit alors $M_K = \frac{\pi_f}{P_K^2}$. On a, d'une part, $P_K M_K(f)(G) \subset F$ puisque

  $$P_K(f)(P_K M_K)(f)(G) = \pi_f(f)(G) = 0\,,$$

  et, d'autre part, $P_K M_K(f)(G) \subset G$ du fait que G est f-stable. Ainsi

  $$P_K M_K(f)(G) \subset F \cap G = \{0\}\,.$$

  Comme par ailleurs

  $$P_K M_K(f)(F) \subset P_K(f)(F) = \{0\}\,,$$

  le polynôme $P_K M_K$ annule f par supplémentarité de F et G, ce qui est absurde par définition du polynôme minimal π_f, étant donné que

  $$\deg(P_K M_K) < \deg\big(P_K^2 M_K\big) = \deg(\pi_f)\,.$$

  Ainsi tous les $(m_k)_{1\leq k\leq r}$ sont égaux à 1. Donc π_f est sans facteur carré.

2 CARDINAL DE $\mathcal{D}_n(\mathbb{F}_q)$

[Rom17, §5.6, p148–151]

ÉNONCÉ

APPLICATION. [CARDINAL DE $\mathcal{D}_n(\mathbb{F}_q)$]
Soient $n \in \mathbb{N}$ et $q = p^r$ où p est un nombre premier et $r \in \mathbb{N}^*$. Soit $\mathcal{D}_n(\mathbb{F}_q)$ l'ensemble des matrices diagonalisables de taille n sur $\mathbb{F}_q$. Alors si par convention $|\mathcal{GL}_0(\mathbb{F}_q)| = 1$, on a :

$$|\mathcal{D}_n(\mathbb{F}_q)| = \sum_{\substack{(n_k)_{1\leq k\leq q}\in\mathbb{N}^q \\ n_1+\cdots+n_q=n}} \frac{|\mathcal{GL}_n(\mathbb{F}_q)|}{\prod_{i=1}^q |\mathcal{GL}_{n_i}(\mathbb{F}_q)|}.$$

DÉVELOPPEMENT

Soit $E = (\mathbb{F}_q)^n$ muni de sa base canonique (ou tout autre $\mathbb{F}_q$-espace vectoriel de dimension n). Dans tout ce qui suit, E_k désignera un sous-espace vectoriel de E. Notons :

$$\mathcal{F} = \left\{ (E_k)_{1\leq k\leq q} : E = \bigoplus_{k=1}^q E_k \right\},$$

puis, pour $n_1, \dots, n_q$ entiers tels que $\sum_{k=1}^q n_k = n$:

$$\mathcal{F}_{(n_1,\dots,n_q)} = \left\{ (E_k)_{1\leq k\leq q} \in \mathcal{F} : \forall k \in [\![1\,;q]\!]\ \dim(E_k) = n_k \right\}.$$

On remarque que les $(\mathcal{F}_{(n_1,\dots,n_q)})_{n_1+\cdots+n_q=n}$ forment une partition de $\mathcal{F}$. On va montrer que $|\mathcal{D}_n(\mathbb{F}_q)| = |\mathcal{F}|$ puis on calculera le cardinal de chaque $\mathcal{F}_{(n_1,\dots,n_q)}$.

LEMME. *Soit $M \in \mathcal{M}_n(\mathbb{F}_q)$. Alors $M \in \mathcal{D}_n(\mathbb{F}_q)$ si et seulement si $M^q = M$.*

Soit M diagonalisable. Son polynôme minimal est scindé à racines simples : écrivons

$$\pi_M(X) = \prod_{\lambda\in \mathrm{Sp}(M)} X - \lambda.$$

Il est clair que $\pi_M(X) \mid X^q - X = \prod_{\lambda\in\mathbb{F}_q} X - \lambda$, et donc $M^q = M$.
Réciproquement, si $M^q = M$, alors M admet un polynôme annulateur scindé à racines simples, donc est diagonalisable.

LEMME. *On a $|\mathcal{D}_n(\mathbb{F}_q)| = |\mathcal{F}|$.*

Pour $M \in \mathcal{D}_n(\mathbb{F}_q)$, appliquons le lemme des noyaux à $X^q - X = \prod_{\lambda\in\mathbb{F}_q} X - \lambda$:

$$E = \bigoplus_{\lambda\in\mathbb{F}_q} \ker(M - \lambda I_n) = \bigoplus_{k=1}^q \underbrace{\ker(M - \lambda_k I_n)}_{E_k}.$$

Il est alors clair que

$$\begin{array}{rccc} \phi : & \mathcal{D}_n(\mathbb{F}_q) & \longrightarrow & \mathcal{F} \\ & M & \longmapsto & (E_k)_{1\leq k\leq q} \end{array}$$

est une bijection (la surjectivité est immédiate, et pour l'injectivité $\phi(M) = \phi(M')$ implique que M et M' coïncident sur chacun des $(E_k)_{1\leq k\leq q}$, donc par somme directe $M = M'$).

Calculons désormais $|\mathcal{F}_N|$, où $N = (n_1, \dots, n_q) \in \mathbb{N}^q$ est fixé tel que $\sum_{k=1}^q n_k = n$.
Considérons l'application

$$\begin{array}{ccc} \mathcal{GL}_n(\mathbb{F}_q) \times \mathcal{F}_N & \longrightarrow & \mathcal{F}_N \\ \big(M, (E_k)_{1\leq k\leq q}\big) & \longmapsto & (ME_k)_{1\leq k\leq q}. \end{array}$$

Elle est bien à valeurs dans $\mathcal{F}_N$ car pour $M \in \mathcal{GL}_n(\mathbb{F}_q)$ et $(E_k)_{1\leq k\leq q} \in \mathcal{F}_N$, on a $E = ME = \sum_{k=1}^q ME_k$ avec $\dim(ME_k) = n_k$ donc $E = \bigoplus_{k=1}^q ME_k$ et ainsi $(ME_k)_{1\leq k\leq q} \in \mathcal{F}_N$.
C'est une action de groupe (les axiomes étant vérifiés grâce aux propriétés des matrices).

Notons de plus que cette action est transitive. En effet, si $(E_k^1)_{1\leq k\leq q}, (E_k^2)_{1\leq k\leq q} \in \mathcal{F}_N$, munissons chaque E_k^i d'une base $\mathcal{B}_k^i$ et considérons M la matrice qui envoie la base $\mathcal{B}_k^1$ sur la base $\mathcal{B}_k^2$ (ce qui est possible puisque les deux bases ont même cardinal). Alors M est entièrement déterminée et on a clairement $ME_k^1 = E_k^2$ pour tout $k \in [\![1\,;r]\!]$.

Contemplons le stabilisateur de $(E_k)_{1\leq k\leq q} \in \mathcal{F}_N$: on a que $M \in \mathrm{Stab}\big((E_k)_{1\leq k\leq q}\big)$ si et seulement si $ME_k \subset E_k$ pour tout $k \in [\![1\,;r]\!]$ (avec en fait égalité par bijectivité). Donc $\mathrm{Stab}\big((E_k)_{1\leq k\leq q}\big)$ est en bijection avec les matrices diagonales par blocs ayant des blocs de tailles $n_1, \dots, n_q$ et pour lesquelles chaque bloc est une matrice inversible (en faisant un changement de base). On a donc $\big|\mathrm{Stab}\big((E_k)_{1\leq k\leq q}\big)\big| = \prod_{k=1}^q |\mathcal{GL}_{n_i}(\mathbb{F}_q)|$.

L'action étant transitive, l'équation aux classes donne

$$|\mathcal{F}_N| = \frac{|\mathcal{GL}_n(\mathbb{F}_q)|}{\prod_{i=1}^q |\mathcal{GL}_{n_i}(\mathbb{F}_q)|}.$$

Il ne reste plus qu'à utiliser le deuxième lemme et le partitionnement de $\mathcal{F}$ pour obtenir :

$$|\mathcal{D}_n(\mathbb{F}_q)| = \sum_{\substack{(n_k)_{1\leq k\leq q}\in\mathbb{N}^q \\ n_1+\cdots+n_q=n}} \frac{|\mathcal{GL}_n(\mathbb{F}_q)|}{\prod_{i=1}^q |\mathcal{GL}_{n_i}(\mathbb{F}_q)|}.$$

COMMENTAIRES

Pour éviter de s'embrouiller sur l'isomorphisme ϕ, une bonne manière de procéder est de voir M comme la matrice d'une application de $\mathrm{GL}(E)$ *dans la base canonique* de E. Fixant la base, on associe une unique application à une matrice, et l'isomorphisme devient plus clair.

3 CRITÈRE D'EISENSTEIN

[FGN07a, §5.16, p188–190]

ÉNONCÉ

Soit A un anneau factoriel.

PROPOSITION. [CRITÈRE D'EISENSTEIN]
Soient $n \in \mathbb{N}^*$, $(a_i)_{0\leq i\leq n} \in A^{n+1}$ puis $P(X) = \sum_{i=0}^n a_i X^i \in A[X]$.
Supposons qu'il existe $p \in A$ premier tel que :

- $p \nmid a_n$,
- $p \mid a_i$ *pour tout* $i < n$,
- $p^2 \nmid a_0$.

Alors P est irréductible dans $\mathrm{Frac}(A)[X]$.

DÉVELOPPEMENT

Pour $Q \in A[X]$, on note $c(Q)$ un PGCD de ses coefficients.

Vérifions d'abord que si $Q, R \in A[X]$, alors $c(QR) = c(Q)\cdot c(R)$.

- Traitons le cas $c(Q) = c(R) = 1$. Supposons par l'absurde que $c(QR) \neq 1$.
  On peut considérer p irréductible divisant $c(QR)$. Alors p divise tous les coefficients du polynôme QR, donc $0 = \overline{QR} = \overline{Q}\cdot\overline{R}$ dans $A/(p)[X]$.
  Or p est irréductible donc est premier (puisque l'anneau A est factoriel), ou encore (p) est premier. En particulier $A/(p)$ puis $A/(p)[X]$ sont intègres. Ainsi $\overline{Q} = 0$ ou $\overline{R} = 0$, c'est-à-dire $p \mid c(Q)$ ou $p \mid c(R)$, ce qui est absurde.
- Dans le cas général, on remarque que $c(\alpha S) = \alpha\cdot c(S)$ pour tout $\alpha \in A$ et $S \in A[X]$.
  Écrivons $Q = c(Q)\cdot Q'$, $R = c(R)\cdot R'$ avec $c(Q') = c(R') = 1$. Alors :

$$c(QR) = c(Q)\cdot c(R)\cdot c(Q'R') = c(Q)\cdot c(R)\,.$$

Soit désormais P comme dans l'énoncé. Supposons-le non irréductible sur $\mathrm{Frac}(A)[X]$.

Écrivons $P = c(P)P'$ avec P' primitif puis $P' = Q'R'$ avec $Q', R' \in \mathrm{Frac}(A)[X]$ de degrés strictement inférieurs à $\deg(P)$. Notons q (respectivement r) le produit des dénominateurs des coefficients de Q' (respectivement R'). Alors $Q = qQ'$ et $R = rR'$ sont à coefficients dans A et $qrP' = QR$. En passant aux PGCD, on a $qr = c(Q)\cdot c(R)$, donc :

$$P = c(P)\cdot\frac{1}{c(Q)}\cdot Q\cdot\frac{1}{c(R)}\cdot R = \left(\frac{c(P)}{c(Q)}\cdot Q\right)\left(\frac{1}{c(R)}\cdot R\right),$$

et P s'écrit comme produit de deux polynômes de $A[X]$ de degrés inférieurs à P.

Écrivons $P = QR$ dans $A[X]$, où, en oubliant les notations de la partie précédente, Q et R sont des polynômes de $A[X]$ de degrés respectifs ℓ et m strictement inférieurs à $\deg(P)$.

Dans $A/(p)[X]$, on a d'après la seconde hypothèse

$$\overline{a_n}X^n = \overline{Q(X)}\cdot\overline{R(X)}\,.$$

Ainsi, si $Q(X) = \sum_{j=0}^{\ell} q_j X^j$ et $R(X) = \sum_{k=0}^{m} r_k X^k$, alors $\overline{q_\ell} \neq 0$ et $\overline{r_m} \neq 0$ puisque par hypothèse $\overline{a_n} \neq 0$.
On peut alors considérer

$$j_0 = \min\Big(\{j \in [\![0\,;\ell]\!] : \overline{q_j} \neq 0\}\Big) \qquad \text{et} \qquad k_0 = \min\Big(\{k \in [\![0\,;m]\!] : \overline{r_k} \neq 0\}\Big)\,.$$

Si $j_0 + k_0 < n$, le monôme de degré $i_0 + j_0$ dans $\overline{QR}$ a pour coefficient $\overline{q_{j_0} r_{k_0}} \neq 0$ par intégrité, ce qui est faux.

Ainsi $j_0 + k_0 = n$ et donc nécessairement $j_0 = \ell$ et $k_0 = m$. Autrement dit, Q et R sont des monômes dans $A/(p)[X]$.

En particulier, $\overline{q_0} = \overline{r_0} = 0$, c'est-à-dire $p \mid q_0$ et $p \mid r_0$, et donc $p^2 \mid q_0 r_0 = a_0$, ce qui contredit la dernière hypothèse sur P.

Ainsi P est irréductible.

COMMENTAIRES

Certaines leçons nécessitent de considérer spécifiquement l'anneau $A = \mathbb{Z}$. La proposition s'écrit alors comme suit, et il faut adapter (et travailler !) les notations du développement.

PROPOSITION. [CRITÈRE D'EISENSTEIN]
Soient $n \in \mathbb{N}^*$, $(a_i)_{0\leq i\leq n} \in \mathbb{Z}^{n+1}$ puis $P(X) = \sum_{i=0}^n a_i X^i \in \mathbb{Z}[X]$.
Supposons qu'il existe p nombre premier tel que :

- $p \nmid a_n$,
- $p \mid a_i$ *pour tout* $i < n$,
- $p^2 \nmid a_0$.

Alors P est irréductible dans $\mathbb{Q}[X]$.

4 DÉCOMPOSITION DE DUNFORD ET CALCUL DE L'EXPONENTIELLE D'UNE MATRICE

[Gou09, §4.4, p193]

ÉNONCÉ

Soit E un $\mathbb{K}$-espace vectoriel de dimension finie, où $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$.

THÉORÈME. [DÉCOMPOSITION DE DUNFORD]
Soit $f \in \mathcal{L}(E)$ de polynôme caractéristique scindé. Il existe un unique couple $(d, n) \in \mathcal{L}(E)^2$ tel que d est diagonalisable, n est nilpotent, $f = d + n$ et d commute avec n.
De plus, d et n sont des polynômes en f.

APPLICATION. Soient $n \in \mathbb{N}^*$ et $M \in \mathcal{M}_n(\mathbb{K})$ de polynôme caractéristique scindé, de racines notées $(\mu_k)_{1\leq k\leq n}$ ou $(\lambda_i)_{1\leq i\leq r}$ de multiplicités $(\alpha_i)_{1\leq i\leq r}$. Soient (D, N) la décomposition de DUNFORD de M et $P \in \mathcal{GL}_n(\mathbb{K})$ telle que $P^{-1}DP = \Delta = \operatorname{diag}(\mu_1, \dots, \mu_n)$. Pour $i \in [\![1\,;r]\!]$, soit P_i le projecteur sur $\ker(M - \lambda_i I_n)^{\alpha_i}$ parallèlement à $\bigoplus_{j\neq i} \ker(M - \lambda_j I_n)^{\alpha_j}$. Alors :

$$\exp(M) = P \operatorname{diag}(\mathrm{e}^{\mu_1}, \dots, \mathrm{e}^{\mu_n}) P^{-1} \sum_{k=0}^{n-1} \frac{N^k}{k!} = \sum_{i=1}^{r} \mathrm{e}^{\lambda_i} \left[\sum_{p=0}^{\alpha_i - 1} \frac{(M - \lambda_i I_n)^p}{p!} \right] P_i \,.$$

DÉVELOPPEMENT

Le polynôme χ_f étant scindé, écrivons-le sous la forme $\chi_f = \prod_{i=1}^{r} (X - \lambda_i)^{\alpha_i}$. Par le lemme des noyaux, on a que $E = \bigoplus_{i=1}^{r} N_i$ où $N_i = \ker(f - \lambda_i \operatorname{id}_E)^{\alpha_i}$ pour $i \in [\![1\,;r]\!]$.

LEMME. *Pour $i \in [\![1\,;r]\!]$, soit p_i la projection sur N_i parallèlement à $\bigoplus_{j\neq i} N_j$. Alors*

$$\sum_{i=1}^{r} p_i = \operatorname{id}_E \,, \quad \forall i \in [\![1\,;r]\!] \quad p_i \in \mathbb{K}[f] \quad \text{et} \quad \forall (i,j) \in [\![1\,;r]\!]^2 \quad i \neq j \;\Rightarrow\; p_i \circ p_j = 0 \,.$$

Notons $M_i = X - \lambda_i$, puis $Q_i = \prod_{j\neq i} M_j^{\alpha_j}$ pour $i \in [\![1\,;r]\!]$. Les $(Q_i)_{1\leq i\leq r}$ sont premiers dans leur ensemble, donc d'après l'identité de BÉZOUT il existe $(U_i)_{1\leq i\leq r} \in \mathbb{K}[X]^r$ tels que $\sum_{i=1}^{r} U_i(f) \circ Q_i(f) = \operatorname{id}_E$. On va montrer que $p_i = U_i(f) \circ Q_i(f)$ pour tout $i \in [\![1\,;r]\!]$.

Avant cela, notons que si cette relation est vérifiée, on a $\sum_{i=1}^{r} p_i = \operatorname{id}_E$ et, pour $i \in [\![1\,;r]\!]$, $p_i \in \mathbb{K}[f]$ puis, si $j \neq i$, comme $\chi_f \mid Q_i\, Q_j$ et puisque les polynômes en f commutent

$$p_i \circ p_j = U_i(f) \circ Q_i(f) \circ U_j(f) \circ Q_j(f) = Q_i(f) \circ Q_j(f) \circ U_i(f) \circ U_j(f) = 0 \,.$$

Soit $i \in [\![1\,;r]\!]$. Reste à montrer que p_i définie ci-dessus est bien la projection annoncée. Déjà, c'est une projection puisque $p_i = \sum_{j=1}^{r} p_i \circ p_j = p_i^2$.

Vérifions que $\operatorname{Im}(p_i) = N_i$.

- Si $y = p_i(x) \in \operatorname{Im}(p_i)$, alors $y \in N_i$ puisque

$$M_i^{\alpha_i}(f)(y) = M_i^{\alpha_i}(f) \circ p_i(x) = U_i(f) \circ \chi_f(f)(x) = 0 \,.$$

- Réciproquement si $x \in N_i$, alors $x = \sum_{i=1}^{r} p_i(x)$ mais $p_j(x) = U_j(f) \circ Q_j(f)(x) = 0$ pour $j \neq i$ puisque $M_i^{\alpha_i} \mid Q_j$, donc $x = p_i(x) \in \operatorname{Im}(p_i)$.

Enfin, assurons-nous que $\ker(p_i) = \bigoplus_{j\neq i} N_j$.

- Pour tout $j \neq i$, on a $N_j \subset \ker(p_i)$ par ce qui précède et donc $\bigoplus_{j\neq i} N_j \subset \ker(p_i)$.
- Réciproquement, si $x \in \ker(p_i)$, alors $x = \sum_{j\neq i} p_j(x) \in \bigoplus_{j\neq i} N_j$.

Revenons à la décomposition de DUNFORD. Montrons l'existence de cette décomposition.
Posons $d = \sum_{i=1}^{r} \lambda_i\, p_i \in \mathbb{K}[f]$. Dans une base concaténée des bases des $(N_i)_{1\leq i\leq r}$, la matrice de d est diagonale donc d est diagonalisable. Reste à poser $n = f - d = \sum_{i=1}^{r} (f - \lambda_i) \circ p_i \in \mathbb{K}[f]$ et vérifier qu'il est nilpotent. Par récurrence, on montre que $n^q = \sum_{i=1}^{r} (f - \lambda_i)^q \circ p_i$ pour tout entier q. Prenant $q = \max_{1\leq i\leq r} \alpha_i$, on obtient que $n^q = 0$, et donc n est nilpotent.

Pour l'unicité, soient $f = d + n = d' + n'$ deux décompositions. On suppose uniquement que d et n sont des polynômes en f. d' et n' commutent avec $f = d' + n'$ et donc avec d et n qui sont des polynômes en f. On a $d - d' = n' - n$ avec $d - d'$ diagonalisable car d et d' commutent et sont diagonalisables, et $n' - n$ est nilpotent (en utilisant la formule du binôme de NEWTON à la puissance égale à la somme des indices de nilpotence). Un endomorphisme diagonalisable et nilpotent est nécessairement nul, donc $d = d'$ et $n = n'$.

Passons à l'application. La première formule découle du calcul suivant :

$$\exp(M) = \exp(D + N) = \exp(D) \exp(N) = P \exp(\Delta) P^{-1} \sum_{k=0}^{n-1} \frac{N^k}{k!} \,,$$

où la deuxième égalité est vraie par commutativité de D et N, et la dernière par continuité du produit matriciel en utilisant que $D = P\Delta P^{-1}$.

Par ailleurs, la preuve ci-dessus assure que $D = \sum_{i=1}^{r} \lambda_i P_i$ et $N = \sum_{i=1}^{r} (M - \lambda_i I_n) P_i$, d'où :

$$\exp(D) = \sum_{p=0}^{+\infty} \sum_{i=1}^{r} \frac{\lambda_i^p}{p!} P_i = \sum_{i=1}^{r} \sum_{p=0}^{+\infty} \frac{\lambda_i^p}{p!} P_i = \sum_{i=1}^{r} \mathrm{e}^{\lambda_i} P_i \,,$$

$$\exp(N) = \sum_{p=0}^{+\infty} \sum_{i=1}^{r} \frac{(M - \lambda_i I_n)^p}{p!} P_i = \sum_{i=1}^{r} \sum_{p=0}^{\alpha_i - 1} \frac{(M - \lambda_i I_n)^p}{p!} P_i \,,$$

et on obtient finalement la deuxième formule.

COMMENTAIRES

Il faut connaître la méthode effective de calcul de la décomposition de DUNFORD.

5 DEGRÉ DE $\mathbb{Q}[\{\sqrt{p_i}\}_{1\leq i\leq n}]$ SUR $\mathbb{Q}$

[Cog00, §2.1, p60–62]

ÉNONCÉ

Théorème. *Pour tout entier $n \in \mathbb{N}^*$ et toute famille $(p_i)_{1\leq i\leq n}$ d'entiers supérieurs ou égaux à 2, tous sans facteur carré, et premiers deux à deux, on a :*

$$\left[\mathbb{Q}\left[\{\sqrt{p_i}\}_{1\leq i\leq n}\right] : \mathbb{Q}\right] = 2^n .$$

DÉVELOPPEMENT

Lorsque les $(p_i)_{1\leq i\leq n}$ sont fixés, on note $\mathbb{Q}_k = \mathbb{Q}\left[\{\sqrt{p_i}\}_{1\leq i\leq k}\right]$ pour $k \in [\![0\,;n]\!]$.

Lemme. *Soient $n \in \mathbb{N}^*$ et $(p_i)_{1\leq i\leq n}$ des entiers satisfaisants.*
Pour $J \subset [\![1\,;n]\!]$, on note $P_J = \prod_{j\in J}\sqrt{p_j}$, puis on pose $K_n = \left\{P_J : J \subset [\![1\,;n]\!]\right\}$.
Alors $\mathbb{Q}_n = \mathrm{Vect}_{\mathbb{Q}}(K_n)$.

En effet, on vérifie que si $P_J, P_{J'} \in K_n$, alors $P_J P_{J'} \in \mathrm{Vect}_{\mathbb{Q}}(K_n)$ puisque

$$P_J P_{J'} = P_{J\oplus J'} \prod_{i\in J\cap J'} p_i\,, \qquad \text{où } J\oplus J' = (J\cup J')\setminus(J\cap J')\,.$$

Ainsi $\mathrm{Vect}_{\mathbb{Q}}(K_n)$ est stable par produit, donc est une algèbre contenant $\mathbb{Q}$ et les $(\sqrt{p_i})_{1\leq i\leq n}$. C'est de plus la plus petite algèbre possédant cette propriété, donc $\mathbb{Q}_n = \mathrm{Vect}_{\mathbb{Q}}(K_n)$.

Montrons maintenant par récurrence forte sur $n \in \mathbb{N}^*$ que pour toute famille $(p_i)_{1\leq i\leq n}$ d'entiers satisfaisants, on a $[\mathbb{Q}_n : \mathbb{Q}_{n-1}] = 2$.

Le résultat attendu découlera alors du théorème de la base télescopique :

$$[\mathbb{Q}_n : \mathbb{Q}] = \prod_{k=1}^{n} [\mathbb{Q}_k : \mathbb{Q}_{k-1}] = 2^n .$$

Dans la suite, on notera $D_k = [\mathbb{Q}_k : \mathbb{Q}_{k-1}]$ pour $k \in [\![1\,;n]\!]$.

- Pour $n = 1$, soit $p_1 \geq 2$ sans facteur carré. $X^2 - p_1$ est un polynôme annulateur de $\sqrt{p_1}$. Si $\sqrt{p_1} \notin \mathbb{Q}$, c'est en fait un polynôme irréductible sur $\mathbb{Q}$, et donc $\mathbb{Q}_1$ est le corps de rupture de ce polynôme. Ainsi $D_1 = \deg(X^2 - p_1) = 2$. Supposons que ce ne soit pas le cas, et écrivons $\sqrt{p_1} = \frac{p}{q} \in \mathbb{Q}$ avec $p \wedge q = 1$, $q \in \mathbb{N}^*$. Alors $p_1 q^2 = p^2$ et donc $p \neq 0$ et $p^2 \mid p_1$, ce qui est faux par hypothèse sur p_1. Ainsi $D_1 = 2$.
- Pour l'hérédité, fixons alors un $n \in \mathbb{N}$ tel que la propriété est vraie pour tout rang inférieur ou égal à n. Soient $p_1, \dots, p_{n+1}$ satisfaisants.
  Comme précédemment, il est clair que $D_{n+1} \leq 2$ car $X^2 - p_{n+1}$ est un polynôme annulateur de $\sqrt{p_{n+1}}$ sur $\mathbb{Q}_n$. Pour que l'extension soit de degré 2, il faut et il suffit donc que $\sqrt{p_{n+1}} \notin \mathbb{Q}_n$. Supposons que ce n'est pas le cas.

  Par hypothèse de récurrence, on a que $\sqrt{p_{n+1}} \in \mathbb{Q}_n = \mathbb{Q}_{n-1}[\sqrt{p_n}]$.
  Or, le théorème de la base télescopique donne $[\mathbb{Q}_n : \mathbb{Q}] = 2^n$ et, en reprenant les notations du Lemme, on a que K_n est une famille de cardinal 2^n engendrant $\mathbb{Q}_n$, donc c'est en fait une $\mathbb{Q}$-base de $\mathbb{Q}_n$. En particulier, on a

  $$\mathbb{Q}_n = \mathbb{Q}_{n-1} \oplus \sqrt{p_n}\mathbb{Q}_{n-1}\,.$$

  Soient donc $a, b \in \mathbb{Q}_{n-1}$ tels que $\sqrt{p_{n+1}} = a + b\sqrt{p_n}$. On a alors :

  $$p_{n+1} = \left(a^2 + b^2 p_n\right) + 2ab\sqrt{p_n}\,.$$

  Par unicité de la décomposition d'un élément de $\mathbb{Q} \subset \mathbb{Q}_n$ en $\mathbb{Q}_{n-1} \oplus \sqrt{p_n}\mathbb{Q}_{n-1}$, on a donc nécessairement $2ab = 0$.
  - Si $a = 0$, alors $\sqrt{p_n p_{n+1}} = b\,p_n \in \mathbb{Q}_{n-1}$, ce qui contredit l'hypothèse de récurrence appliquée à la famille $p_1, \dots, p_{n-1}$ et $p_n p_{n+1}$ (ce dernier entier est bien sans facteur carré et premier avec les autres).
  - Sinon, $b = 0$ et alors $\sqrt{p_{n+1}} = a \in \mathbb{Q}_{n-1}$, ce qui contredit à nouveau l'hypothèse de récurrence appliquée à la famille $p_1, \dots, p_{n-1}$ et p_{n+1}.

  On a donc une contradiction. Ainsi $D_{n+1} = 2$ et on a le résultat au rang $n+1$.

On conclut par principe de récurrence.

COMMENTAIRES

A-t-on $\mathbb{Q}_n = \mathbb{Q}[\sqrt{p_1 \cdots p_n}]$?

Le lemme est là pour rallonger un peu le développement, mais à la place il est possible de montrer le théorème de la base télescopique.

6 DÉTERMINANT DE GRAM ET INÉGALITÉ DE HADAMARD

[Gou09, §5.3, p262–263]

ÉNONCÉ

Soit $(E, \langle \cdot \mid \cdot \rangle)$ un espace préhilbertien (réel ou complexe).

THÉORÈME. [DISTANCE À UN SOUS-ESPACE VECTORIEL]
Soit F un sous-espace vectoriel de E de dimension finie n et muni d'une base $(e_1, \dots, e_n)$. Alors pour tout $x \in E$, on a :

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

THÉORÈME. [INÉGALITÉS DE HADAMARD]
(i) Soient $x_1, \dots, x_n$ des vecteurs de E. Alors $G(x_1,\dots,x_n) \leq \prod_{i=1}^n \|x_i\|^2$.
(ii) Soient $x_1, \dots, x_n$ des vecteurs de $\mathbb{C}^n$. Alors $\left|\det(x_1,\dots,x_n)\right| \leq \prod_{i=1}^n \|x_i\|_2$, où $\|\cdot\|_2$ désigne la norme hermitienne standard sur $\mathbb{C}^n$.
Dans les deux points, on a égalité si et seulement si la famille $(x_i)_{1\leq i\leq n}$ est orthogonale ou si l'un des vecteurs est nul.

DÉVELOPPEMENT

Remarquons que le déterminant de GRAM d'une famille de vecteurs liée est nul, par linéarité du produit scalaire. Réciproquement, si le déterminant est nul, les vecteurs colonnes des produits scalaires sont liés, donc il existe $k \in [\![1\,;n]\!]$ et $(\lambda_\ell)_{1\leq\ell\leq n,\ell\neq k}$ coefficients non tous nuls tels que :

$$\forall i \in [\![1\,;n]\!] \qquad \langle e_i \mid e_k\rangle = \sum_{\substack{\ell=1\\ \ell\neq k}}^n \lambda_\ell \langle e_i \mid e_\ell\rangle = \langle e_i \mid \sum_{\substack{\ell=1\\ \ell\neq k}}^n \overline{\lambda_\ell}\, e_\ell\rangle\,,$$

et donc $e_k - \sum_{\ell\neq k} \overline{\lambda_\ell}\, e_\ell \in \mathrm{Vect}(e_1,\dots,e_n)^\perp$. Mais $e_k - \sum_{\ell\neq k} \overline{\lambda_\ell}\, e_\ell \in \mathrm{Vect}(e_1,\dots,e_n)$, donc $e_k - \sum_{\ell\neq k} \overline{\lambda_\ell}\, e_\ell = 0$ et la famille de vecteurs est liée.

Montrons alors le premier théorème.

1. F étant de dimension finie, on a que $d(x,F)$ est atteint en la projection $f \in F$ de x, ainsi $d(x,F) = \|x-f\|$. Par définition de f, notons que :

$$\forall i \in [\![1\,;n]\!] \quad \langle x \mid e_i\rangle = \langle f \mid e_i\rangle \qquad \text{et} \qquad \|x\|^2 = \|f\|^2 + \|x-f\|^2\,.$$

2. On a $M_G(e_1,\dots,e_n,x) = \left(\begin{array}{ccc|c} & \mathcal{M}_G(e_1,\dots,e_n) & & \begin{array}{c}\langle e_1 \mid x\rangle\\ \vdots \\ \langle e_n \mid x\rangle\end{array} \\ \hline \langle x \mid e_1\rangle & \cdots & \langle x \mid e_n\rangle & \|x\|^2 \end{array}\right)$.

D'où par linéarité par rapport à la dernière colonne :

$$\begin{aligned} G(e_1,\dots,e_n,x) &= \left|\begin{array}{ccc|c} & \mathcal{M}_G(e_1,\dots,e_n) & & \begin{array}{c}\langle e_1 \mid f\rangle\\ \vdots \\ \langle e_n \mid f\rangle\end{array} \\ \hline \langle f \mid e_1\rangle & \cdots & \langle f \mid e_n\rangle & \|f\|^2 \end{array}\right| \\ &\quad + \left|\begin{array}{ccc|c} & \mathcal{M}_G(e_1,\dots,e_n) & & \begin{array}{c}0\\ \vdots \\ 0\end{array} \\ \hline \langle f \mid e_1\rangle & \cdots & \langle f \mid e_n\rangle & \|x-f\|^2 \end{array}\right| \\ &= G(e_1,\dots,e_n,f) + \|x-f\|^2 \cdot G(e_1,\dots,e_n) \\ G(e_1,\dots,e_n,x) &= d(x,F)^2 \cdot G(e_1,\dots,e_n)\,, \end{aligned}$$

car $f \in F = \mathrm{Vect}(e_1,\dots,e_n)$.

Montrons désormais le second théorème.

1. Si la famille $(x_1,\dots,x_n)$ est liée, alors $G(x_1,\dots,x_n) = 0$ et l'inégalité est évidente. On montre par récurrence sur $n \in \mathbb{N}^*$ qu'une famille de n vecteurs libres de E vérifie l'inégalité, avec égalité si et seulement si ils sont orthogonaux.
   - Si $n=1$, on a $G(x_1) = \|x_1\|^2$.
   - Supposons la propriété vraie au rang n. Soient $(x_i)_{1\leq i\leq n+1}$ des vecteurs libres de E. Notons $F = \mathrm{Vect}(x_1,\dots,x_n)$ et considérons f la projection orthogonale de x_{n+1} sur F. Alors :

$$\begin{aligned} G(x_1,\dots,x_{n+1}) &= G(x_1,\dots,x_n)\cdot\|x_{n+1}-f\|^2 && \text{par le premier théorème} \\ &\overset{(1)}{\leq} \prod_{i=1}^n \|x_i\|^2 \cdot \|x_{n+1}-f\|^2 && \text{par hypothèse de récurrence} \\ G(x_1,\dots,x_{n+1}) &\overset{(2)}{\leq} \prod_{i=1}^n \|x_i\|^2 \cdot \|x_{n+1}\|^2\,, \end{aligned}$$

   comme $\|x_{n+1}-f\|^2 \leq \|x_{n+1}-f\|^2 + \|f\|^2 = \|x_{n+1}\|^2$.
   On a de plus égalité dans (1) si et seulement si la famille $(x_i)_{1\leq i\leq n}$ est orthogonale (par hypothèse de récurrence) et dans (2) si et seulement si $\|x_{n+1}-f\|^2 = \|x_{n+1}\|^2$, c'est-à-dire $\|f\|^2 = 0$, ou encore x_{n+1} est orthogonal aux $(x_i)_{1\leq i\leq n}$.
   On a donc montré l'hypothèse de récurrence au rang $n+1$.

   D'où le résultat par principe de récurrence.
2. Notons que $M_G(x_1,\dots,x_n) = \overline{N^\top} N$ où N est la matrice de vecteurs colonnes les $(x_i)_{1\leq i\leq n}$. On applique alors le point précédent avec $G(x_1,\dots,x_n) = \left|\det(N)\right|^2$.

7 FACTORISATIONS LU ET DE CHOLESKY

[AK02, §6.2/3, p113–124]

ÉNONCÉ

THÉORÈME. [FACTORISATION LU]
Soit $A \in \mathcal{M}_n(\mathbb{K})$ une matrice dont tous les mineurs principaux sont non nuls. Alors il existe un unique couple (L,U) de matrices tel que L est triangulaire inférieure avec diagonale de 1, U est triangulaire supérieure et $A = LU$.

THÉORÈME. [FACTORISATION DE CHOLESKY]
Soit A une matrice symétrique réelle définie positive. Alors il existe une unique matrice B triangulaire inférieure telle que tous ses éléments diagonaux soient positifs et $A = BB^*$.

DÉVELOPPEMENT

Commençons par démontrer la factorisation LU. Soit $A \in \mathcal{M}_n(\mathbb{K})$ satisfaisante.

- Existence. On effectue l'algorithme du pivot de GAUSS sur A en vérifiant qu'aucune matrice de permutation n'est nécessaire, *i.e.*, que les pivots sont tous non nuls. Pour cela, on crée par récurrence une suite $(A^k)_{1\leq k\leq n}$ de matrices en posant $A^1 = A$ puis $A^{k+1} = E^k A^k$ pour $k \in [\![1\,;n-1]\!]$, tant que $a^k_{k,k} \neq 0$, avec :

$$E^k = \begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & 1 & & \\ & & & -\ell_{k+1,k} & 1 & \\ & & & \vdots & & \ddots \\ & & & -\ell_{n,k} & & 1 \end{pmatrix}, \qquad \text{où } \ell_{i,k} = \frac{a^k_{i,k}}{a^k_{k,k}} \text{ pour } i \in [\![1\,;n]\!].$$

Supposons tous les $(a^k_{k,k})_{1\leq k\leq n}$ non nuls[1]. Alors la suite $(A_k)_{1\leq k\leq n}$ est bien définie, et

$$\forall k \in [\![1\,;n]\!] \qquad A^k = \begin{pmatrix} a^1_{1,1} & \star & & \cdots & & \star \\ & \ddots & & & & \\ & & a^k_{k,k} & \ddots & & \vdots \\ & & \star & \star & & \\ & & \vdots & \ddots & \ddots & \star \\ & & \star & \cdots & \star & \star \end{pmatrix}.$$

On peut alors poser $U = A^n$, $L = (E^1)^{-1}\cdots(E^{n-1})^{-1}$ et vérifier, en obtenant

$$L = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ \ell_{2,1} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ \ell_{n,1} & \cdots & \ell_{n,n-1} & 1 \end{pmatrix},$$

que l'on a bien $A = LU$, $U \in \mathcal{TS}_n(\mathbb{K})$ et $L \in \mathcal{TI}^1_n(\mathbb{K})$.

Reste à voir que les pivots $(a^k_{k,k})_{1\leq k\leq n}$ sont tous non nuls. Procédons par récurrence forte :
- $a^1_{1,1} = \Delta^1 \neq 0$ car le premier mineur principal est non nul.
- Soit $k \in [\![2\,;n-1]\!]$. Supposons les $k-1$ premiers pivots non nuls.
  Cette hypothèse assure que la formule $A^k = E^{k-1}\cdots E^1 A$ est valable. En considérant des produits par blocs on voit que Δ^k la matrice extraite des k premières lignes et colonnes de A se factorise en $\Delta^k = L^k_{1,1}U^k_{1,1}$ où $L^k_{1,1} \in \mathcal{TI}^1_k(\mathbb{K})$ et $U^k_{1,1} \in \mathcal{TS}_k(\mathbb{K})$. Δ^k et $L^k_{1,1}$ étant inversibles, $U^k_{1,1}$ l'est aussi et donc $a^k_{k,k} \neq 0$ puisque c'est le dernier coefficient diagonal de $U^k_{1,1}$.

- Unicité. Soient $(L_1,U_1),(L_2,U_2)$ deux couples convenants.
  On a $L_2 \in \mathcal{TI}^1_n(\mathbb{K})$ et donc $L_2^{-1} \in \mathcal{TI}^1_n(\mathbb{K})$. Comme A est inversible, on en déduit que $U_1 \in \mathcal{TS}_n(\mathbb{K})$ est inversible, d'où

$$L_2^{-1}L_1 = U_2U_1^{-1} \in \mathcal{TI}^1_n(\mathbb{K}) \cap \mathcal{TS}_n(\mathbb{K}) = \{I_n\}\,.$$

Ainsi $L_1 = L_2$ et $U_1 = U_2$, d'où l'unicité.

Passons à la factorisation de CHOLESKY. Soit A satisfaisante.

- Existence. A est définie positive donc vérifie les hypothèses de factorisation LU : écrivons $A = LU$ et posons $D = \operatorname{diag}\left(\sqrt{u_{1,1}},\ldots,\sqrt{u_{n,n}}\right)$, ce qui est possible puisque pour tout $k \in [\![1\,;n]\!]$, on a $\prod_{i=1}^k u_{i,i} = \det\Delta^k > 0$ (les $(u_{i,i})_{1\leq i\leq n}$ sont donc strictement positifs).
  On pose $B = LD$ et $C = D^{-1}U$ qui vérifient toujours $A = BC$. Comme $A = A^*$, il en découle que $C^*B^* = BC$ ou encore

$$B^{-1}C^* = C(B^*)^{-1} \in \mathcal{TI}_n(\mathbb{K}) \cap \mathcal{TS}_n(\mathbb{K})\,,$$

qui est donc une matrice diagonale. Comme B et C ont même coefficients diagonaux, on en déduit que c'est en fait l'identité et on a $C = B^*$.

- Unicité. Si B_1 et B_2 conviennent, on a $B_2^{-1}B_1 = B_2^*(B_1^*)^{-1}$ et cette matrice, nommée D, est diagonale par les mêmes arguments que précèdemment. Comme $B_1 = B_2D$, on remarque que $A = B_2B_2^* = B_2(DD^*)B_2^*$, d'où $D^2 = I_n$ puisque B_2 est inversible. Les coefficients diagonaux de la décomposition de CHOLESKY étant positifs, on a forcément $D = I_n$ et donc $B_1 = B_2$.

1. en fait il n'y a besoin de cette hypothèse que pour les $n-1$ premiers pivots

8 FORMES DE HANKEL

[CG13, §V.D, p197]

ÉNONCÉ

PROPOSITION. [FORME DE HANKEL]
Soit $P \in \mathbb{R}[X]$ de degré n. Notons $x_1, \dots, x_t$ ses racines distinctes et $m_1, \dots, m_t$ leur multiplicités respectives. On définit, pour $k \in \mathbb{N}$, $s_k = \sum_{\ell=1}^{t} m_\ell x_\ell^k$, puis on pose

$$\forall (X_0, \dots, X_{n-1}) \in \mathbb{C}^n \qquad s(X_0, \dots, X_{n-1}) = \sum_{0 \leq i,j \leq n-1} s_{i+j} X_i X_j \, .$$

Alors la restriction $s_{\mathbb{R}}$ de s à $\mathbb{R}^n$ est une forme quadratique sur $\mathbb{R}^n$, de signature (p,q) où $p+q=t$ et $p-q$ est le nombre de racines réelles de P.

DÉVELOPPEMENT

1. Dans un premier temps, vérifions que s est une forme quadratique sur $\mathbb{R}^n$.
s est un polynôme homogène de degré 2, donc une forme quadratique sur $\mathbb{C}^n$. Il suffit donc de s'assurer que $s(\mathbb{R}^n) \subset \mathbb{R}$. Pour cela, on vérifie que les $(s_k)_{k\in\mathbb{N}}$ sont réels. Or, pour tout $k \in \mathbb{N}$, on peut écrire puisque x_i et $\overline{x_i}$ ont même multiplicité dans P à coefficients réels :
$$s_k = \sum_{1 \leq \ell \leq t \,:\, x_\ell \in \mathbb{R}} m_\ell \, x_\ell^k + \sum_{1 \leq \ell \leq t \,:\, \mathrm{Im}(x_\ell) > 0} m_\ell \underbrace{\left(x_\ell^k + \overline{x_\ell}^k\right)}_{\in \mathbb{R}} \quad \in \mathbb{R} \, .$$
Ainsi s est une forme quadratique sur $\mathbb{R}^n$ dont on note (p,q) la signature.
2. Pour $\ell \in [\![1\,;t]\!]$, définissons $\phi_\ell(X_0, \dots, X_{n-1}) = \sum_{i=0}^{n-1} x_\ell^i X_i$ de sorte que :
$$\begin{aligned} \sum_{\ell=1}^{t} m_\ell \, \phi_\ell^2 &= \sum_{\ell=1}^{t} m_\ell \sum_{0 \leq i,j \leq n-1} x_\ell^{i+j} X_i X_j = \sum_{0 \leq i,j \leq n-1} \left(\sum_{\ell=1}^{t} m_\ell x_\ell^{i+j} \right) X_i X_j \\ &= \sum_{0 \leq i,j \leq n-1} s_{i+j} X_i X_j = s \, . \end{aligned}$$
Remarquons que les $(\phi_\ell)_{1\leq\ell\leq t}$ sont des formes linéaires indépendantes. En effet, si $\mathcal{B}$ est la base duale de la base canonique de $\mathbb{C}^n$, on a :
$$M_{\mathcal{B}}(\phi_1, \dots, \phi_t) = \begin{pmatrix} 1 & \cdots & 1 \\ x_1 & \cdots & x_t \\ \vdots & \ddots & \vdots \\ x_1^{n-1} & \cdots & x_t^{n-1} \end{pmatrix} ,$$
matrice dont le mineur principal d'ordre t est inversible (déterminant de VANDERMONDE). La matrice est donc de rang t, tout comme la famille $(\phi_\ell)_{1\leq\ell\leq t}$, qui est donc composée de formes indépendantes.

Le rang étant invariant par extension de corps, on a que
$$p+q = \mathrm{rg}(s_{\mathbb{R}}) = \mathrm{rg}(s) = \mathrm{rg}\left(\sum_{\ell=1}^{t} m_\ell \phi_\ell^2 \right) = t \, ,$$
puisque les $(m_\ell)_{1\leq\ell\leq t}$ sont tous strictement positifs. Ainsi $p+q=t$.
3. Fixons $\ell \in [\![1\,;t]\!]$ et regardons la signature de $\psi_\ell = \phi_\ell^2 + \overline{\phi_\ell}^2$ lorsque $\mathrm{Im}(x_\ell) > 0$.
On remarque que $\psi_\ell = 2\,\mathrm{Re}(\phi_\ell^2)$ est une forme quadratique réelle et on calcule que
$$\phi_\ell^2 = \mathrm{Re}(\phi_\ell)^2 - \mathrm{Im}(\phi_\ell)^2 + 2i\,\mathrm{Re}(\phi_\ell)\,\mathrm{Im}(\phi_\ell) \, , \quad \text{d'où} \quad \psi_\ell = 2\,\mathrm{Re}(\phi_\ell)^2 - 2\,\mathrm{Im}(\phi_\ell)^2 \, .$$
Par ailleurs, comme $x_\ell \neq \overline{x_\ell}$, il est facile de vérifier que ϕ_ℓ et $\overline{\phi_\ell}$ sont indépendantes. Ainsi ψ_ℓ est de rang 2 sur $\mathbb{C}$, donc sur $\mathbb{R}$, et nécessairement ψ_ℓ est de signature $(1,1)$: en effet on a écrit sa réduction de GAUSS (si $\mathrm{Re}(\phi_\ell)$ et $\mathrm{Im}(\phi_\ell)$ étaient liées, ϕ_ℓ serait de rang 1).
4. Reste à écrire s sous la forme
$$s = \sum_{1 \leq \ell \leq t \,:\, x_\ell \in \mathbb{R}} m_\ell \, \phi_\ell^2 + \sum_{1 \leq \ell \leq t \,:\, \mathrm{Im}(x_\ell) > 0} m_\ell \, \psi_\ell^2 \, .$$
C'est la décomposition de GAUSS de s. Notant r le nombre de racines réelles de P, s est alors de signature
$$(r,0) + \left(\tfrac{t-r}{2}, \tfrac{t-r}{2}\right) = \left(\tfrac{t+r}{2}, \tfrac{t-r}{2}\right) .$$
Par unicité de la signature d'une forme quadratique réelle, on a alors $p-q=r$.

COMMENTAIRES

Il faut bien maitriser l'utilisation du rang qui intervient très régulièrement dans ce développement, en particulier l'invariance du rang par extension de corps.

A quoi servent les formes de HANKEL ? On peut calculer le nombre de racines réelles/complexes d'un polynôme sans les connaître. En effet :
- on peut calculer les $(s_k)_{k\in\mathbb{N}}$ par récurrence en utilisant les polynômes symétriques élémentaires, sans avoir besoin des racines (voir par exemple [Gou09, §2.5, p80]),
- puis on a un algorithme pour trouver la réduction de GAUSS, donc la signature, d'une forme quadratique.

Il faut au moins savoir faire ces calculs sur des exemples simples.

9 FORMULE DE POISSON DISCRÈTE

[Pey04, §II.3, p44]

ÉNONCÉ

Soit G un groupe abélien fini et H un sous-groupe de G.

LEMME. *$H^\# \simeq \widehat{G/H}$ est de cardinal $\frac{|G|}{|H|}$.*

PROPOSITION. [FORMULE DE POISSON DISCRÈTE]
Pour toute fonction $f : G \longrightarrow \mathbb{C}$, et pour tout $g \in G$, on a :

$$\sum_{h\in H} f(gh) = \frac{|H|}{|G|} \sum_{\chi\in H^\#} \widehat{f}(\overline{\chi})\,\chi(g)\,.$$

En particulier ($g = 1$), on a $\displaystyle\sum_{h\in H} f(h) = \frac{|H|}{|G|} \sum_{\chi\in H^\#} \widehat{f}(\overline{\chi})$.

DÉVELOPPEMENT

Commençons par le lemme. Regardons l'application :

$$\begin{array}{rccc} \varphi : & \widehat{G/H} & \longrightarrow & H^\# \\ & \chi & \longmapsto & \widetilde{\chi} : g \longmapsto \chi(gH)\,. \end{array}$$

Pour tout $\chi \in \widehat{G/H}$, il est clair que $\widetilde{\chi}$ est bien un caractère. Comme de plus $\widetilde{\chi}(h) = \chi(H) = 1$ pour tout $h \in H$, ceci justifie que l'application φ est bien définie.
De plus, il est immédiat de vérifier que φ est un morphisme de groupes.

Vérifions que φ est bijective.

- Pour l'injectivité, il est clair que $\widetilde{\chi_1} = \widetilde{\chi_2}$ implique $\chi_1 = \chi_2$ par surjectivité de l'application $G \longrightarrow G/H, g \longmapsto gH$, et puisque $\chi_1(gH) = \widetilde{\chi_1}(g) = \widetilde{\chi_2}(g) = \chi_2(gH)$ pour tout $g \in G$.
- Pour la surjectivité, fixons $\gamma \in H^\#$ et définissons $\chi \in \widehat{G/H}$ par $\chi(gH) = \gamma(g)$ pour tout $g \in G$. Cette application est bien définie puisque si $gH = g'H$, alors $g(g')^{-1} \in H$ et donc $\gamma\big(g(g')^{-1}\big) = 1$, ou encore $\gamma(g) = \gamma(g')$. C'est bien un caractère puisque par passage au quotient χ hérite de la structure de morphisme.

Ainsi $H^\# \simeq \widehat{G/H}$, d'où l'on déduit immédiatement

$$\big|H^\#\big| = \frac{|G|}{|H|}\,.$$

Passons maintenant à la formule de POISSON.
On note S un système de représentants de G/H dans G[1].
Soit donc $f : G \longrightarrow \mathbb{C}$. Regardons l'application :

$$\begin{array}{rccc} \widetilde{f} : & G & \longrightarrow & \mathbb{C} \\ & g & \longmapsto & \sum_{h\in H} f(gh)\,. \end{array}$$

Remarquons que $\widetilde{f}$ est invariante par H, puisque si $g \in G$ et $h \in H$:

$$\widetilde{f}(gh) = \sum_{h'\in H} f(ghh') = \sum_{h''\in H} f(gh'') = \widetilde{f}(g)\,.$$

L'application $\widetilde{f}$ passe donc au quotient et définit une application[2] $\widetilde{f} : G/H \longrightarrow \mathbb{C}$.
On peut donc décomposer $\widetilde{f}$ en série de FOURIER[3] :

$$\forall g \in G \qquad \widetilde{f}(\overline{g}) = \sum_{\chi\in\widehat{G/H}} \langle \widetilde{f} \mid \chi\rangle\,\chi(\overline{g})\,.$$

Fixons $\chi \in \widehat{G/H}$. Comme $S \times H \longrightarrow G, (g,h) \longmapsto gh$ est bijective, il vient :

$$\begin{aligned} \langle \widetilde{f} \mid \chi\rangle &= \frac{1}{|G/H|} \sum_{g\in S} \widetilde{f}(\overline{g})\,\overline{\chi(\overline{g})} = \frac{|H|}{|G|} \sum_{g\in S}\sum_{h\in H} f(gh)\,\overline{\chi(\overline{g})} \\ &= \frac{|H|}{|G|} \sum_{g'\in G} f(g')\,\overline{\chi(\overline{g'})} \qquad \text{puisque } \chi\big(\overline{gh}\big) = \chi(\overline{g}) \\ \langle \widetilde{f} \mid \chi\rangle &= \frac{|H|}{|G|} \sum_{g'\in G} f(g')\,\overline{\widetilde{\chi}(g')} = \frac{|H|}{|G|}\,\widehat{f}(\overline{\chi})\,. \end{aligned}$$

D'où finalement, pour tout $g \in G$, en utilisant le lemme :

$$\sum_{h\in H} f(gh) = \sum_{\chi\in\widehat{G/H}} \frac{|H|}{|G|}\,\widehat{f}(\overline{\widetilde{\chi}})\,\chi(\overline{g}) = \frac{|H|}{|G|} \sum_{\widetilde{\chi}\in H^\#} \widehat{f}(\overline{\widetilde{\chi}})\,\widetilde{\chi}(g)\,.$$

COMMENTAIRES

Il faut connaître la formule de POISSON vue en analyse, qui s'écrit presque de la même manière !

La formule de POISSON discrète peut notamment être utilisée en théorie des codes correcteurs.

1. $S \longrightarrow G/H, g \longmapsto gH$ est bijective
2. toujours appelée $\widetilde{f}$
3. c'est en fait la formule d'inversion, qui découle du fait que $\hat{G}$ est une base orthonormée de $\mathcal{F}(G,\mathbb{C})$

10 HOMÉOMORPHISME DE L'EXPONENTIELLE

[CG13, §VI.2.5, p357–358]

ÉNONCÉ

THÉORÈME. $\exp : \mathcal{S}_n(\mathbb{R}) \longrightarrow \mathcal{S}_n^{++}(\mathbb{R})$ *est un homéomorphisme.*

DÉVELOPPEMENT

On procède en plusieurs étapes.

1. Vérifions que cette application est bien définie, *i.e.*, à valeurs dans $\mathcal{S}_n^{++}(\mathbb{R})$.
Soit en effet $A \in \mathcal{S}_n(\mathbb{R})$. S est diagonalisable dans une base orthonormée : on peut écrire
$$A = P \cdot \operatorname{diag}(\lambda_1, \dots, \lambda_n) \cdot P^{-1}\,,$$
où $P \in \mathcal{O}_n(\mathbb{R})$ et les $(\lambda_i)_{1\leq i\leq n}$ sont les valeurs propres (réelles) de A. Il vient alors
$$\exp(A) = P \cdot \operatorname{diag}\big(\mathrm{e}^{\lambda_1}, \dots, \mathrm{e}^{\lambda_n}\big) \cdot P^{-1} \quad \in \mathcal{S}_n^{++}(\mathbb{R})\,.$$

2. Montrons maintenant que l'application est surjective.
Choisissons pour cela $B \in \mathcal{S}_n^{++}(\mathbb{R})$ et écrivons
$$B = \tilde{P} \cdot \operatorname{diag}(\mu_1, \dots, \mu_n) \cdot \tilde{P}^{-1}\,,$$
où $\tilde{P} \in \mathcal{O}_n(\mathbb{R})$ et les $(\mu_i)_{1\leq i\leq n} \in (\mathbb{R}_+^*)^n$ sont les valeurs propres de B. En posant
$$A = \tilde{P} \cdot \operatorname{diag}\big(\ln(\mu_1), \dots, \ln(\mu_n)\big) \cdot \tilde{P}^{-1}\,,$$
il est clair que $A \in \mathcal{S}_n(\mathbb{R})$ et que $\exp(A) = B$.

3. Passons à l'injectivité.
Prenons $A, A' \in \mathcal{S}_n(\mathbb{R})$ telles que $\exp(A) = \exp(A')$ et notons $\lambda_1, \dots, \lambda_n$ les valeurs propres de A. Choisissons aussi $P \in \mathcal{O}_n(\mathbb{R})$ telle que $A = P \cdot \operatorname{diag}(\lambda_1, \dots, \lambda_n) \cdot P^{-1}$.
Par interpolation de LAGRANGE, il existe un polynôme $Q \in \mathbb{R}[X]$ tel que
$$\forall i \in [\![1\,;n]\!] \qquad Q(\mathrm{e}^{\lambda_i}) = \lambda_i\,.$$
A' commute alors avec
$$\begin{aligned} Q\big(\exp(A')\big) = Q\big(\exp(A)\big) &= Q\Big(P \cdot \operatorname{diag}\big(\mathrm{e}^{\lambda_1}, \dots, \mathrm{e}^{\lambda_n}\big) \cdot P^{-1}\Big) \\ &= P \cdot Q\Big(\operatorname{diag}\big(\mathrm{e}^{\lambda_1}, \dots, \mathrm{e}^{\lambda_n}\big)\Big) \cdot P^{-1} = P \cdot \operatorname{diag}(\lambda_1, \dots, \lambda_n) \cdot P^{-1} = A\,. \end{aligned}$$
Ainsi A et A' sont diagonalisables et commutent, donc sont co-diagonalisables.
Écrivons[1] $A = R \cdot \operatorname{diag}(\lambda_1, \dots, \lambda_n) \cdot R^{-1}$ et $A' = R \cdot \operatorname{diag}(\mu_1, \dots, \mu_n) \cdot R^{-1}$. Alors
$$\operatorname{diag}(\mathrm{e}^{\lambda_1}, \dots, \mathrm{e}^{\lambda_n}) = R^{-1} \cdot \exp(A') \cdot R = R^{-1} \cdot \exp(A) \cdot R = \operatorname{diag}(\mathrm{e}^{\mu_1}, \dots, \mathrm{e}^{\mu_n})\,,$$
et donc nécessairement $\lambda_i = \mu_i$ pour tout $i \in [\![1\,;n]\!]$, puis $A = A'$. D'où l'injectivité.

1. quitte à réindicer les $(\lambda_i)_{1\leq i\leq n}$

4. Montrons que l'application et sa réciproque sont continues.
Comme $\exp$ est continue sur $\mathcal{M}_n(\mathbb{R})$, il en est de même de sa restriction à $\mathcal{S}_n(\mathbb{R})$. Il suffit donc de montrer que la réciproque est continue.

Choisissons donc une suite $(B_p)_{p\in\mathbb{N}} \in \mathcal{S}_n^{++}(\mathbb{R})^{\mathbb{N}}$ qui converge vers $B \in \mathcal{S}_n^{++}(\mathbb{R})$ et définissons $(A_p)_{p\in\mathbb{N}} \in \mathcal{S}_n(\mathbb{R})^{\mathbb{N}}$ et $A \in \mathcal{S}_n(\mathbb{R})$ telles que $B_p = \exp(A_p)$ pour tout $p \in \mathbb{N}$ et $B = \exp(A)$. On veut montrer que $A_p \underset{p\to+\infty}{\longrightarrow} A$.

Notons que $(B_p)_{p\in\mathbb{N}}$ est bornée pour $|||\cdot|||_2$ et remarquons que $(B_p^{-1})_{p\in\mathbb{N}}$ converge vers B^{-1}, donc est également bornée pour $|||\cdot|||_2$. Or, pour $M \in \mathcal{S}_n^{++}(\mathbb{R})$, on a
$$|||M|||_2 = \max\big(\operatorname{Sp}(M)\big)\,.$$
En effet, si $M \in \mathcal{S}_n^{++}(\mathbb{R})$, M est diagonalisable en base orthonormée donc
$$|||M|||_2 = |||\operatorname{diag}(\lambda_1, \dots, \lambda_n)|||_2\,,$$
où les $(\lambda_i)_{1\leq i\leq n}$ sont les valeurs propres réelles positives de M. On a alors
$$\forall x \in \mathbb{R}^n \qquad \|Mx\|_2 \leq \max\big(\operatorname{Sp}(M)\big) \cdot \|x\|_2\,,$$
avec égalité pour $x = (\delta_{i\,i_{\max}})_{1\leq i\leq n}$ où $i_{\max} \in [\![1\,;n]\!]$ est tel que $\lambda_{i_{\max}} = \max\big(\operatorname{Sp}(M)\big)$, ce qui assure que $|||M|||_2 = \max\big(\operatorname{Sp}(M)\big)$.

Ainsi on peut majorer les spectres des $(B_p)_{p\in\mathbb{N}}$ et de B par une constante C, puis ceux des $(B_p^{-1})_{p\in\mathbb{N}}$ et de B^{-1} par une constante $\frac{1}{C'}$.
Les valeurs propres des $(B_p)_{p\in\mathbb{N}}$ sont donc incluses dans $[C', C]$ compact de $\mathbb{R}_+$, puis celles de $(A_p)_{p\in\mathbb{N}}$ dans $[\ln(C), \ln(C')]$ compact de $\mathbb{R}$.
La suite $(A_p)_{p\in\mathbb{N}}$ est donc bornée pour $|||\cdot|||_2$, donc admet une valeur d'adhérence, qui ne peut être que A. En effet, si $(A_p)_{p\in\mathbb{N}}$ converge (quitte à extraire) vers $\tilde{A}$, alors, comme $B_p = \exp(A_p)$ et par continuité de l'exponentielle,
$$B_p \underset{p\to+\infty}{\longrightarrow} B = \exp\big(\tilde{A}\big)\,, \quad \text{et donc par injectivité} \quad A = \tilde{A}\,.$$
La suite n'ayant qu'une seule valeur d'adhérence A, elle converge vers A.
Ce qui conclut quant à la continuité de l'application réciproque.

11 IRRÉDUCTIBILITÉ DE Φ_n

[Per96, §3.4, p81–83]

ÉNONCÉ

Soit $n \in \mathbb{N}^*$.

PROPOSITION. $\Phi_n \in \mathbb{Z}[X]$.

THÉORÈME. *Φ_n est irréductible sur $\mathbb{Z}$ et donc sur $\mathbb{Q}$.*

COROLLAIRE. *On a $\left[\mathbb{Q}[e^{2i\pi/n}] : \mathbb{Q}\right] = \varphi(n)$.*

DÉVELOPPEMENT

Pour montrer que $\Phi_n \in \mathbb{Z}[X]$, on procède par récurrence forte sur $n \in \mathbb{N}^*$.

- On a que $\Phi_1 = X - 1 \in \mathbb{Z}[X]$.
- Si $n > 1$ et le résultat est vrai pour tout $d \mid n$ avec $d < n$, considérons

$$F(X) = \prod_{d|n, d\neq n} \Phi_d(X)\,.$$

Alors $F(X) \in \mathbb{Z}[X]$ par hypothèse de récurrence et est unitaire.
On peut effectuer la division euclidienne de $X^n - 1$ par $F(X)$ dans $\mathbb{Z}[X]$ et on a

$$X^n - 1 = F(X)P(X) + R(X)\,, \quad \text{avec } P, R \in \mathbb{Z}[X] \text{ et } \deg(R) < \deg(F)\,.$$

Or on sait que $X^n - 1 = \Phi_n(X)F(X)$ dans $\mathbb{Q}[X]$, donc $F(X)\cdot\big(\Phi_n(X) - P(X)\big) = R(X)$, ce qui implique, en considérant les degrés, que $\Phi_n = P \in \mathbb{Z}[X]$.

Passons maintenant au théorème.

LEMME. *Soit ζ une racine n-ième primitive de l'unité. Soit p un nombre premier tel que $p \nmid n$. On sait que ζ^p est une autre racine n-ième primitive de l'unité. Notons $Q \in \mathbb{Q}[X]$ (respectivement R) le polynôme minimal[a] de ζ (respectivement ζ^p) sur $\mathbb{Q}$.*
Alors $Q \in \mathbb{Z}[X]$, $Q \mid \Phi_n$ et $Q = R$.

a. le polynôme minimal est défini sur un anneau principal, donc pas sur $\mathbb{Z}[X]$!

Vérifions que ce lemme permet de conclure. On veut montrer que $Q = \Phi_n$.
Soit ζ une racine n-ième primitive de l'unité. On sait que les racines n-ièmes primitives de l'unité sont exactement les ζ^k telles que $k \wedge n = 1$. Soit ζ^k l'une d'entre elles, et écrivons $k = \prod_{i=1}^{\ell} p_i^{\alpha_i}$ où les $(p_i)_{1\leq i\leq \ell}$ sont premiers et ne divisent pas n.
Par le lemme, ζ a même polynôme minimal que ζ^{p_1}, puis que $(\zeta^{p_1})^{p_1}$, puis que $\zeta^{p_1^{\alpha_1}}$, puis que $(\zeta^{p_1^{\alpha_1}})^{p_2^{\alpha_2}} = \zeta^{p_1^{\alpha_1}p_2^{\alpha_2}}$, et finalement ζ et ζ^k ont même polynôme minimal.

Ainsi toutes les racines n-ièmes primitives de l'unité ont même polynôme minimal. Q admet donc $\varphi(n)$ racines au moins. Comme $Q \mid \Phi_n$, qui a exactement $\varphi(n)$ racines, on a que $Q = \Phi_n$. Alors Φ_n est irréductible.

Finalement, il ne reste qu'à montrer le lemme.
Vérifions déjà que $Q, R \in \mathbb{Z}[X]$. $\mathbb{Z}[X]$ est factoriel, donc on peut écrire $\Phi_n = \prod_{i=1}^{s} P_i$ où les $(P_i)_{1\leq i\leq s}$ sont des polynômes irréductibles de $\mathbb{Z}[X]$. Φ_n étant unitaire, on peut supposer que les $(P_i)_{1\leq i\leq s}$ le sont aussi. ζ étant racine de Φ_n, on a que ζ annule l'un des $(P_i)_{1\leq i\leq s}$, qui est unitaire et irréductible sur $\mathbb{Z}$ *donc*[1] sur $\mathbb{Q}$. Ainsi $Q = P_i \in \mathbb{Z}[X]$ pour un $i \in [\![1\,;s]\!]$, et de même $R = P_j \in \mathbb{Z}[X]$ pour un $j \in [\![1\,;s]\!]$.
Notons de plus que Q et R divisent Φ_n dans $\mathbb{Z}[X]$, et donc que divise aussi Φ_n si $Q \neq R$.

On veut montrer que $Q = R$. Supposons que ce n'est pas le cas.
On a que ζ est racine de $R(X^p)$, donc $Q(X) \mid R(X^p)$ dans $\mathbb{Q}[X]$[2].
Écrivons $R(X^p) = Q(X)H(X)$ puis $H(X) = \frac{a}{b}H'(X)$ où $H' \in \mathbb{Z}[X]$ est de contenu 1. Alors

$$1 = c\big(R(X^p)\big) = c(Q)\cdot\frac{a}{b}\cdot c(H') = \frac{a}{b}\,, \quad \text{et donc } H \in \mathbb{Z}[X]\,.$$

Si $R(X) = \sum_{i=0}^{r} a_i X^i$, on a alors $\overline{a_i} = \overline{a_i}^p$ dans $\mathbb{F}_p$ pour tout $i \in [\![0\,;r]\!]$, d'où

$$\overline{R}(X^p) = \left(\sum_{i=0}^{r}\overline{a_i}X^{ip}\right) = \left(\sum_{i=0}^{r}\left(\overline{a_i}X^i\right)^p\right) = \left(\sum_{i=0}^{r}\overline{a_i}X^i\right)^p = \overline{R}(X)^p\,,$$

l'avant-dernière égalité provenant du morphisme de FROBENIUS. Ainsi $\overline{R}(X)^p = \overline{Q}(X)\cdot\overline{H}(X)$.

Considérons un facteur irréductible T de $\overline{Q}$ sur $\mathbb{F}_p$. On a nécessairement que T divise $\overline{R}$ ou $\overline{R}^{p-1}$, et alors par récurrence $T \mid \overline{R}$.

On a vu que $QR \mid \Phi_n$, et donc dans $\mathbb{F}_p$ on a $T^2 \mid \overline{\Phi_n} \mid X^n - 1$. Ainsi, dans son corps de décomposition, $X^n - 1$ a une racine double, ce qui est faux puisque $p \nmid n$: en effet on vérifie que $(X^n - 1)' = \overline{n}X^{n-1}$ n'a pour racine que 0 qui n'est pas une racine de $X^n - 1$.

Ainsi $Q = R$. Ce qui démontre le lemme.

Enfin, pour le corollaire, il suffit de remarquer que $\zeta = \mathrm{e}^{2i\pi/n}$ est une racine n-ième primitive de l'unité. Son polynôme minimal sur $\mathbb{Q}$ est donc Φ_n, de degré $\varphi(n)$.

COMMENTAIRES

Une argumentation parfois plus détaillée est proposée dans le [Gou09, §2.5, p91], ce qui peut être une bonne lecture pour se remettre les idées en place.

Il faut savoir justifier que la division euclidienne d'un polynôme de $\mathbb{Z}[X]$ par un polynôme de $\mathbb{Z}[X]$ à coefficient dominant inversible donne des polynômes dans $\mathbb{Z}[X]$.

1. il faut en être convaincu
2. les divisions euclidiennes se font dans $\mathbb{Q}[X]$ euclidien. Lorsque le diviseur est unitaire, on montre en fait que l'on peut faire la division dans $\mathbb{Z}[X]$

12 ISOMÉTRIES DU CUBE

[Rom17, §3.4.4, p88] [CG13, §XII.3, p365]

ÉNONCÉ

Théorème. $\mathrm{Isom}(\mathcal{C}) \simeq \mathfrak{S}_4 \times \mathbb{Z}/2\mathbb{Z}$ *et* $\mathrm{Isom}^+(\mathcal{C}) \simeq \mathfrak{S}_4$.

DÉVELOPPEMENT

On se place sur $\mathbb{R}^3$ et on note $\mathcal{S} = \{-1,1\}^3$ et $\mathcal{C}$ le cube de sommets $\mathcal{S}$. On nomme les sommets de $\mathcal{S}$ comme sur la figure ci-dessous.

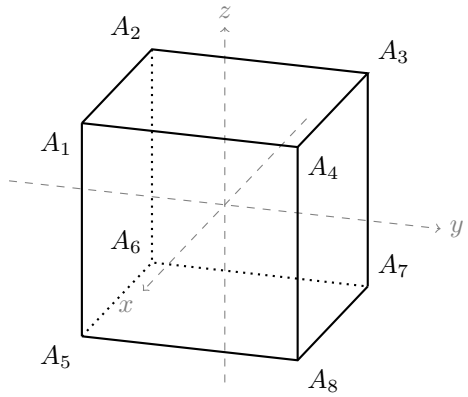

FIGURE 12.1 – Le cube $\mathcal{C}$ et la disposition des sommets de $\mathcal{S}$.

On sait que $\mathrm{Isom}(\mathcal{C}) = \mathrm{Isom}(\mathcal{S})$. De plus, on a l'isomorphisme $\mathrm{Isom}(\mathcal{S}) \simeq \mathrm{Isom}^+(\mathcal{S}) \times \mathbb{Z}/2\mathbb{Z}$ puisque l'on a une bijection

$$\begin{array}{ccc} \mathrm{Isom}^+(\mathcal{S}) & \longrightarrow & \mathrm{Isom}^-(\mathcal{S}) \\ \rho & \longmapsto & \rho \circ s \end{array}, \quad \text{où } s : x \longmapsto -x \in \mathrm{Isom}^-(\mathcal{S}).$$

On veut donc montrer que $\mathrm{Isom}^+(\mathcal{S}) \simeq \mathfrak{S}_4$. On va utiliser le lemme suivant :

Lemme. *$f \in \mathrm{Isom}^+(\mathcal{S})$ est entièrement déterminée par la donnée de $f(A_i)$ et $f(A_j)$ pour un couple $(i,j) \in [\![1\,;8]\!]^2$ tel que $[A_i\,A_j]$ est une arête du cube.*

En effet, soient $f, g \in \mathrm{Isom}^+(\mathcal{S})$ telles que[1] $f(A_1) = g(A_1)$ et $f(A_2) = g(A_2)$.
Regardons $f^{-1} \circ g \in \mathrm{Isom}^+(\mathcal{S})$: il s'agit d'une rotation fixant A_1 et A_2, donc le plan $(O\,A_1\,A_2)$: c'est donc l'identité puisqu'une rotation non triviale de $\mathbb{R}^3$ fixe au plus une droite.

1. on se place sur l'arête $[A_1\,A_2]$ sans perte de généralité, quitte à réindicer les sommets

Montrons désormais que $\mathrm{Isom}^+(\mathcal{S}) \simeq \mathfrak{S}_4$.

Soit $\mathcal{D} = \{D_k : k \in [\![1\,;4]\!]\} = \{[A_1\,A_7], [A_2\,A_8], [A_3\,A_5], [A_4\,A_6]\}$ l'ensemble des (grandes) diagonales du cube. Si $f \in \mathrm{Isom}^+(S)$, alors f induit une permutation de $\mathcal{D}$, puisque f conserve les distances, les segments, et que chaque point de $\mathcal{S}$ est envoyé sur un point de $\mathcal{S}$. On peut donc définir le morphisme de groupes $\Phi : \mathrm{Isom}^+(S) \longrightarrow \mathfrak{S}(\mathcal{D}), f \longmapsto \sigma_f$ où σ_f est telle que

$$\forall k \in [\![1\,;4]\!] \qquad f(D_k) = D_{\sigma_f(k)}\,.$$

- Vérifions que c'est un morphisme injectif.
  Soit $f \in \mathrm{Isom}^+(S)$ telle que $f(D_k) = D_k$ pour tout $k \in [\![1\,;4]\!]$.
  En particulier $[f(A_1)\,f(A_7)] = f\big([A_1\,A_7]\big) = [A_1\,A_7]$. On a deux cas :
  - Soit $f(A_1) = A_1$. D'une part, comme $f\big([A_1\,A_2]\big) = [A_1\,f(A_2)]$, la conservation des distances impose que $f(A_2) \in \{A_2, A_4, A_5\}$. D'autre part, la diagonale $[A_2\,A_8]$ est fixée, donc $f(A_2) \in \{A_2, A_8\}$. Ainsi $f(A_2) = A_2$ et le lemme assure que $f = \mathrm{id}$.
  - Soit $f(A_1) = A_7 = -A_1$. Puisque $f\big([A_1\,A_2]\big) = [A_7\,f(A_2)]$, il découle par conservation des distances que $f(A_2) \in \{A_3, A_6, A_8\}$. La diagonale $[A_2\,A_8]$ étant fixée, on a $f(A_2) = A_8 = -A_2$. De même, on vérifie que $f(A_4) = -A_4$ et $f(A_5) = -A_5$. Comme f conserve les diagonales on a en fait $f(A_i) = -A_i$ pour tout $i \in [\![1\,;8]\!]$, autrement dit $s \circ f = \mathrm{id}$ soit $f = s \in \mathrm{Isom}^-(\mathcal{S})$, ce qui contredit la définition de f.
  Ainsi $\ker(\Phi) = \{\mathrm{id}\}$ et Φ est bien un morphisme injectif de $\mathrm{Isom}^+(\mathcal{S})$ dans $\mathfrak{S}(\mathcal{D})$.
- Reste à montrer qu'il est surjectif, ou encore que $\mathrm{Isom}^+(\mathcal{S})$ contient 24 éléments. Soit

$$\begin{array}{ccc} \mathrm{Isom}^+(\mathcal{S}) \times \mathcal{S} & \longrightarrow & \mathcal{S} \\ (f, A_i) & \longmapsto & f(A_i)\,. \end{array}$$

  On va montrer que cette action est transitive et que le stabilisateur d'un sommet contient 3 isométries. L'équation aux classes conclura alors que :

$$\mathrm{card}\big(\mathrm{Isom}^+(\mathcal{S})\big) = \mathrm{card}(\mathcal{S}) \cdot \mathrm{card}\big(\mathrm{Stab}(A_1)\big) = 8 \times 3 = 24\,.$$

  - Montrons d'abord que l'action n'a qu'une orbite.
    Prenons par exemple A_1, et regardons $\rho_z \in \mathrm{Isom}^+(\mathcal{S})$ la rotation d'axe $(O\,z)$ et d'angle $-\frac{\pi}{2}$ ainsi que $\rho_y \in \mathrm{Isom}^+(\mathcal{S})$ la rotation d'axe $(O\,y)$ et d'angle $\frac{\pi}{2}$. On a

$$\begin{array}{llll} \mathrm{id}(A_1) = A_1\,, & \rho_z(A_1) = A_2\,, & \rho_z^2(A_1) = A_3\,, & \rho_z^3(A_1) = A_4\,, \\ \rho_y(A_1) = A_5\,, & \rho_y^2(A_1) = A_6\,, & \rho_z^2\rho_y(A_1) = A_7\,, & \rho_z^3\rho_y(A_1) = A_8\,, \end{array}$$

    et ainsi $\mathrm{Isom}^+(\mathcal{S}) \cdot A_1 = \mathcal{S}$: l'action est transitive.
  - Montrons maintenant que $\mathrm{Stab}(A_1)$ possède 3 éléments.
    Si $f \in \mathrm{Isom}^+(\mathcal{S})$ fixe A_1, alors f permute $\{A_2, A_4, A_5\}$ par conservation des distances. Le lemme permet alors d'assurer que $\mathrm{card}\big(\mathrm{Stab}(A_1)\big) \leq 3$.
    Par ailleurs, l'identité et les rotations d'axe $(O\,A_1)$ et d'angles $\pm\frac{2\pi}{3}$ sont bien des isométries distinctes de $\mathrm{Stab}(A_1)$.

13 ISOMORPHISME $\mathrm{SU}_2(\mathbb{C})/\{\pm I_2\} \simeq \mathrm{SO}_3(\mathbb{R})$

[CG13, §VII.3, p232]

ÉNONCÉ

THÉORÈME. *On a l'isomorphisme* $\mathrm{SU}_2(\mathbb{C})/\{\pm I_2\} \simeq \mathrm{SO}_3(\mathbb{R})$.

DÉVELOPPEMENT

RAPPEL. On identifie $\mathbb{H}$ et $\mathbb{R}_+^* \times \mathrm{SU}_2(\mathbb{C})$ via

$$h = x + y\mathsf{i} + z\mathsf{j} + t\mathsf{k} \quad \longmapsto \quad \begin{pmatrix} a & -\overline{b} \\ b & \overline{a} \end{pmatrix} \text{ où } a = x + iy \text{ et } b = -z + it\,.$$

Si $h = x + y\mathsf{i} + z\mathsf{j} + t\mathsf{k} \in \mathbb{H}$, on pose $\overline{h} = x - y\mathsf{i} - z\mathsf{j} - t\mathsf{k} \in \mathbb{H}$, puis $N(h) = h\overline{h} = x^2 + y^2 + z^2 + t^2$, qui définit une norme sur $\mathbb{H}$, de forme bilinéaire symétrique associée :

$$\forall (h, h') \in \mathbb{H}^2 \qquad \langle h \mid h' \rangle = \tfrac{1}{2}\big(h\overline{h'} + h'\overline{h}\big)\,.$$

Enfin, on note $\mathbb{I} = \mathbb{R}\cdot\mathsf{i} + \mathbb{R}\cdot\mathsf{j} + \mathbb{R}\cdot\mathsf{k}$ et on remarque que $\mathbb{I} = (\mathbb{R}\cdot 1)^\perp$.

Pour $h \in \mathrm{SU}_2(\mathbb{C})$, soit $\varphi_h : \mathbb{H} \longrightarrow \mathbb{H}, u \longmapsto huh^{-1} = hu\overline{h}$. C'est un automorphisme de $\mathbb{H}$. L'application

$$\begin{array}{rccc} \varphi : & \mathrm{SU}_2(\mathbb{C}) & \longrightarrow & \mathrm{Aut}(\mathbb{H}) \\ & h & \longmapsto & \varphi_h \end{array}$$

est alors bien définie et est un morphisme de groupes, car $\varphi_1 = \mathrm{id}$ et pour $h_1, h_2 \in \mathbb{H}$ on a :

$$\forall u \in \mathbb{H} \qquad \varphi_{h_1h_2}(u) = h_1h_2u(h_1h_2)^{-1} = h_1h_2uh_2^{-1}h_1^{-1} = \varphi_{h_1} \circ \varphi_{h_2}(u)\,.$$

Fixons $h \in \mathrm{SU}_2(\mathbb{C})$. L'application φ_h est linéaire et conserve la norme puisque :

$$\forall u \in \mathbb{H} \qquad N\big(\varphi_h(u)\big) = \det\big(huh^{-1}\big) = \det(u) = N(u)\,.$$

Ainsi $\varphi_h \in \mathcal{O}(\mathbb{H}) \simeq \mathcal{O}_4(\mathbb{R})$. Comme φ_h stabilise $\mathbb{R}\cdot 1$, elle stabilise $(\mathbb{R}\cdot 1)^\perp = \mathbb{I}$. Posons alors $\phi_h = \varphi_{h|\mathbb{I}} \in \mathcal{O}(\mathbb{I}) \simeq \mathcal{O}_3(\mathbb{R})$ puis $\phi : h \longmapsto \phi_h$, qui définit une action de $\mathrm{SU}_2(\mathbb{C})$ sur $\mathbb{I} \simeq \mathbb{R}^3$.

- Déterminons $\ker(\phi)$.
  Si $h \in \mathrm{SU}_2(\mathbb{C})$ est tel que $\phi_h = \mathrm{Id}$, on a $hu = uh$ pour tout $u \in \mathbb{I}$, donc $h \in Z(\mathbb{H}) = \mathbb{R}\cdot 1$. Comme $h \in \mathrm{SU}_2(\mathbb{C})$, il en découle que $h = \pm 1$. Ainsi $\ker(\phi) = \{\pm I_2\}$.
- Vérifions maintenant que $\mathrm{Im}(\phi) \subset \mathrm{SO}_3(\mathbb{R})$.
  L'application ϕ est continue car $\phi_h(u)$ est une multiplication en h, h^{-1} et u. $\mathrm{SU}_2(\mathbb{C})$ étant isomorphe[1] à $\mathbb{S}^3$, elle est connexe et alors par continuité $\mathrm{Im}(\phi)$ l'est aussi. En conséquence, $\mathrm{Im}(\phi)$ est incluse dans la composante connexe de $\phi_1 = \mathrm{id}$ qui est donc $\mathrm{SO}_3(\mathbb{R})$.
- Enfin montrons que ϕ est surjective dans $\mathrm{SO}_3(\mathbb{R})$.
  Pour cela, rappelons-nous que $\mathrm{SO}_3(\mathbb{R})$ est engendrée par les retournements (*i.e.*, les symétries orthogonales par rapport à un sous-espace vectoriel de dimension $n-2$).
  Prenons donc $h \in \mathbb{I} \cap \mathrm{SU}_2(\mathbb{C})$ et montrons que $\phi_h = r_h$ le retournement d'axe $\mathbb{R}\cdot h$.
  En effet, ϕ_h est orthogonal et vérifie
  - $\phi_h(h) = h$ donc $\phi_{h|\mathbb{R}\cdot h} = \mathrm{id}$,
  - si h' est orthogonal à h, alors $h\overline{h'} + h'\overline{h} = 0$ ou encore $h(-h') + h'(-h) = 0$, d'où $hh' = -h'h$ puis $\phi_h(h') = hh'h^{-1} = -h'$. Ainsi $\phi_{h|(\mathbb{R}\cdot h)^\perp} = -\mathrm{id}$.

En passant au quotient, on obtient donc l'isomorphisme $\mathrm{SU}_2(\mathbb{C})/\{\pm I_2\} \simeq \mathrm{SO}_3(\mathbb{R})$.

Explicitons cet isomorphisme.
Soient $(x, y, z) \in \mathbb{R}^3 \setminus \{(0,0,0)\}$ tels que $x^2 + y^2 + z^2 = 1$ et $\theta \in [0, 2\pi[$.
Posons $q' = x\mathsf{i} + y\mathsf{j} + z\mathsf{k} \in \mathrm{SU}_2(\mathbb{C})$ puis $q = \cos\frac{\theta}{2}\cdot 1 + \sin\frac{\theta}{2}\cdot q'$.
Alors $N(q) = \cos^2\frac{\theta}{2} + \sin^2\frac{\theta}{2}\cdot N(q') = 1$ par orthogonalité de $\mathbb{R}\cdot 1$ et $\mathbb{I}$, donc $q \in \mathrm{SU}_2(\mathbb{C})$.

On a vu que $\phi_{q'}$ est la rotation d'axe (x, y, z) et d'angle π.
On va montrer que ϕ_q est la rotation d'axe de vecteur directeur (x, y, z) et d'angle θ.
En effet, pour $u \in \mathbb{I} \cap \mathrm{SU}_2(\mathbb{C})$, on a :

$$\phi_q(u) = qu\overline{q} = \cos^2\tfrac{\theta}{2}\cdot u + \cos\tfrac{\theta}{2}\cdot\sin\tfrac{\theta}{2}\cdot(q'u - uq') - \sin^2\tfrac{\theta}{2}\cdot q'uq'\,.$$

Donc, d'une part, puisque $(q')^2 = -q'\overline{q'} = -1$:

$$\phi_q(q') = \big(\cos^2\tfrac{\theta}{2} - \sin^2\tfrac{\theta}{2}\cdot(q')^2\big)\cdot q' = \big(\cos^2\tfrac{\theta}{2} + \sin^2\tfrac{\theta}{2}\big)\cdot q' = q'\,,$$

et d'autre part, si $\langle u \mid q' \rangle = 0$, comme on a vu que $q'u = -uq'$:

$$\phi_q(u) = \big(\cos^2\tfrac{\theta}{2} - \sin^2\tfrac{\theta}{2}\big)\cdot u + 2\cos\tfrac{\theta}{2}\cdot\sin\tfrac{\theta}{2}\cdot q'u = \cos\theta\cdot u + \sin\theta\cdot q'u\,.$$

En remarquant que $q'u$ correspond au produit vectoriel $\vec{q'} \wedge \vec{u}$, on a que $(q', u, q'u)$ forme une base orthonormée directe, et donc ϕ_q est bien la rotation annoncée.

COMMENTAIRES

Il faut être à l'aise avec la manipulation des quaternions, notamment la non commutativité, les calculs de carrés selon la norme, … Noter par exemple que le produit scalaire doit être symétrique, donc l'ordre des éléments dans chaque produit le définissant est important. Il faut aussi savoir retrouver le lien entre produit dans $\mathbb{H}$ et produit vectoriel.

La dernière partie n'est pas référencée. Elle peut être condensée avec la preuve de la surjectivité, mais on peut l'omettre si l'on manque de temps ou si l'on ne s'en souvient plus. On peut aussi mentionner le résultat pour aiguiller les questions du jury.

1. topologiquement, c'est-à-dire que l'on a une bijection continue d'inverse continu

14 LEMME DE MORSE

[Rou99, §5.5, p327]

ÉNONCÉ

THÉORÈME. [LEMME DE MORSE]
Soit $f : U \longrightarrow \mathbb{R}$ une fonction de classe $\mathcal{C}^3$ définie sur un ouvert U de $\mathbb{R}^n$ contenant 0. On suppose que $df(0) = 0$ et $d^2f(0)$ est non dégénérée, de signature $(p, n-p)$.
Alors il existe un $\mathcal{C}^1$-difféomorphisme φ entre deux voisinages de l'origine de $\mathbb{R}^n$ tel que $\varphi(0) = 0$ et, au voisinage de 0,

$$f(x) - f(0) = \varphi_1(x)^2 + \cdots + \varphi_p(x)^2 - \varphi_{p+1}(x)^2 - \cdots - \varphi_n(x)^2 .$$

DÉVELOPPEMENT

La fonction f est de classe $\mathcal{C}^3$. La formule de TAYLOR avec reste intégral donne, pour $x \in U$:

$$f(x) - f(0) = \int_0^1 (1-t) d^2f(tx)(x,x)\,\mathrm{d}t = x^\top \int_0^1 (1-t) d^2f(tx)\,\mathrm{d}t\, x = x^\top Q(x)\, x ,$$

où $Q : x \in U \longmapsto \int_0^1 (1-t) d^2f(tx)\,\mathrm{d}t$. Par dérivation sous le signe intégral, on a que Q est de classe $\mathcal{C}^1$. Par ailleurs, $Q(0) = \frac{1}{2} D^2f(0)$ est symétrique inversible de signature $(p, n-p)$.

LEMME. *Soit $A_0 \in \mathcal{S}_n(\mathbb{R}) \cap \mathcal{GL}_n(\mathbb{R})$. Alors il existe un voisinage $V \in \mathcal{V}(A_0) \cap \mathcal{S}_n(\mathbb{R})$ et une application $g \in \mathcal{C}^1\big(V, \mathcal{GL}_n(\mathbb{R})\big)$ telles que*

$$\forall A \in V \qquad A = g(A)^\top A_0 g(A) .$$

En effet, soit

$$\varphi : \mathcal{M}_n(\mathbb{R}) \longrightarrow \mathcal{S}_n(\mathbb{R}), M \longmapsto M^\top A_0 M .$$

On a que φ est différentiable et

$$\forall M \in \mathcal{M}_n(\mathbb{R}) \quad \forall H \in \mathcal{M}_n(\mathbb{R}) \qquad d\varphi(M)(H) = H^\top A_0 M + M^\top A_0 H ,$$

d'où, en particulier,

$$\forall H \in \mathcal{M}_n(\mathbb{R}) \qquad d\varphi(I_n)(H) = H^\top A_0 + A_0 H = (A_0 H)^\top + A_0 H .$$

On a donc $\ker(d\varphi(I_n)) = A_0^{-1}\mathcal{A}_n(\mathbb{R})$ où $\mathcal{A}_n(\mathbb{R})$ est l'ensemble des matrices antisymétriques. De plus $d\varphi(I_n)$ est surjective dans $\mathcal{S}_n(\mathbb{R})$ puisque

$$\forall A \in \mathcal{S}_n(\mathbb{R}) \qquad d\varphi(I_n)\Big(\frac{1}{2} A_0^{-1} A\Big) = A .$$

Soit $F = A_0^{-1}\mathcal{S}_n(\mathbb{R})$. Notons que $F \oplus \ker\big(d\varphi(I_n)\big) = \mathcal{M}_n(\mathbb{R})$ et $I_n \in F$. Si $\psi = \varphi_{|F}$, on a que $d\psi(I_n)$ est bijective. Le théorème d'inversion locale donne que ψ est un $\mathcal{C}^1$-difféomorphisme local d'un voisinage W de I_n dans[1] $\mathcal{GL}_n(\mathbb{R})$ sur un voisinage V de $A_0 = \psi(I_n)$ dans $\mathcal{S}_n(\mathbb{R})$.

Pour $A \in V$, il existe donc une unique matrice inversible $M \in W$ telle que $A = M^\top A_0 M$. On a que $M = \psi^{-1}(A)$, et donc $g = \psi^{-1}$ convient.

Revenons à la preuve du lemme de MORSE.
Appliquons le résultat du lemme précédent à $Q(0) \in \mathcal{S}_n(\mathbb{R})$.
Si $M(x) = g\big(Q(x)\big)$ pour $x \in U$, on a, au voisinage de 0,

$$Q(x) = \big(M(x)\big)^\top Q(0)\, M(x) ,$$

$$\text{et donc} \quad f(x) - f(0) = x^\top Q(x)\, x = \big(M(x)\, x\big)^\top Q(0) \underbrace{M(x)\, x}_{=y} = y^\top Q(0)\, y .$$

Comme $Q(0) = \frac{1}{2} D^2f(0)$ est de signature $(p, n-p)$, il existe, par classification des formes quadratiques, une matrice $A \in \mathcal{GL}_n(\mathbb{R})$ telle que

$$A^\top Q(0) A = \mathrm{diag}(\underbrace{1, \ldots, 1}_{p \text{ termes}}, \underbrace{-1, \ldots, -1}_{n-p \text{ termes}}) ,$$

et le changement linéaire de coordonnées $y = Au$ donne alors :

$$y^\top Q(0)\, y = u^\top A^\top Q(0)\, A\, u = u_1^2 + \cdots + u_p^2 - u_{p+1}^2 - \cdots - u_n^2 .$$

Posons donc $\varphi : x \in U \longmapsto A^{-1} M(x)\, x$. L'application φ est de classe $\mathcal{C}^1$ et on a :

$$\forall h \in \mathbb{R}^n \qquad d\varphi(0)(h) = A^{-1}\big(dM(0)(h) \times 0 + M(0) \times h\big) = A^{-1} M(0)(h) .$$

Donc $d\varphi(0) = A^{-1} M(0)$ est inversible.

Par le théorème d'inversion locale, on a que φ est un $\mathcal{C}^1$-difféomorphisme local entre deux voisinages de 0 car $\varphi(0) = 0$, ce qui conclut puisqu'alors dans ce voisinage de 0 :

$$f(x) - f(0) = \varphi_1(x)^2 + \cdots + \varphi_p(x)^2 - \varphi_{p+1}(x)^2 - \cdots - \varphi_n(x)^2 .$$

COMMENTAIRES

Que se passe-t-il sans l'hypothèse que f est $\mathcal{C}^3$? On n'a plus l'assurance que φ est $\mathcal{C}^1$!
En effet, si f est de classe $\mathcal{C}^3$, alors M est $\mathcal{C}^1$ et on calcule :

$$\forall x \in U \quad \forall h \in \mathbb{R}^n \qquad d\varphi(x)(h) = A^{-1}\big(dM(x)(h) \times 0 + M(x) \times h\big) ,$$

et l'on voit que $x \longmapsto d\varphi(x)$ ne serait pas continue si M n'était pas $\mathcal{C}^1$.

1. quitte à restreindre, $\mathcal{GL}_n(\mathbb{R})$ étant ouvert

15 LOI DE RÉCIPROCITÉ QUADRATIQUE

[Rom17, §13.6-7, p429–435]

ÉNONCÉ

THÉORÈME. [LOI DE RÉCIPROCITÉ QUADRATIQUE]
Soient $p \neq q$ des nombres premiers impairs. Alors :

$$\left(\frac{p}{q}\right)\left(\frac{q}{p}\right) = (-1)^{\frac{(p-1)(q-1)}{4}}, \quad \text{où } \left(\frac{x}{p}\right) = \begin{cases} 1 & \text{si } x \text{ est un carré dans } \mathbb{F}_p^*, \\ 0 & \text{si } x = 0, \\ -1 & \text{sinon.} \end{cases}$$

DÉVELOPPEMENT

Soient p et q des entiers premiers impairs. Montrons d'abord le lemme suivant.

LEMME. *Pour $a \in \mathbb{F}_q^*$, on a $\overline{\left(\frac{a}{q}\right)} = a^{\frac{q-1}{2}}$ dans $\mathbb{F}_q^*$ et $\left|\{x \in \mathbb{F}_q^* : ax^2 = 1\}\right| = 1 + \left(\frac{a}{q}\right)$.*

Soit $a \in \mathbb{F}_q^*$. Si $a = b^2 \in \mathbb{F}_q^*$ est un carré, alors nécessairement $a^{\frac{q-1}{2}} = b^{q-1} = \overline{1} = \overline{\left(\frac{a}{q}\right)}$.
Comme $\mathbb{F}_q^*$ a $\frac{q-1}{2}$ carrés[1], et que $X^{\frac{q-1}{2}} - 1$ a au plus $\frac{q-1}{2}$ solutions dans $\mathbb{F}_q^*$, on en déduit que si a n'est pas un carré, alors $a^{\frac{q-1}{2}} = \overline{-1} = \overline{\left(\frac{a}{q}\right)}$.
Ensuite, si $a = b^2$ est un carré, alors $ax^2 = 1 \iff (bx)^2 = 1 \iff x = \pm b^{-1}$ et donc, puisque $q \neq 2$, l'équation $ax^2 = 1$ a deux solutions. Sinon, l'équation n'a pas de solution puisque le produit d'un carré c par un non carré d est un non carré : en effet

$$\overline{\left(\frac{cd}{q}\right)} = (cd)^{\frac{q-1}{2}} = c^{\frac{q-1}{2}} \cdot d^{\frac{q-1}{2}} = \overline{\left(\frac{c}{q}\right)} \cdot \overline{\left(\frac{d}{q}\right)} = \overline{1} \times \overline{-1} = \overline{-1},$$

d'où $\left(\frac{cd}{q}\right) = -1$ puisque $q \neq 2$. Ceci termine la démonstration du lemme.

Soit $X = \left\{x \in \mathbb{F}_q^p : \sum_{i=1}^p x_i^2 = 1\right\}$. Dénombrons X modulo p de deux manières différentes.

- Considérons d'abord l'action suivante, où les indices sont définis modulo p :

$$\begin{array}{ccc} \mathbb{F}_p \times X & \longrightarrow & X \\ (\overline{k}, (x_1, \dots, x_p)) & \longmapsto & (x_{k+1}, \dots, x_{k+p}). \end{array}$$

  Le cardinal de l'orbite d'un élément divise $p = |\mathbb{F}_p|$, donc vaut 1 ou p. L'orbite de $x \in X$ est réduite à lui-même si et seulement si $x_1 = \dots = x_p$. Le nombre de tels x est le nombre de solutions de $px_1^2 = 1$, c'est-à-dire $1 + \left(\frac{p}{q}\right)$ d'après le lemme. Ainsi

$$|X| \equiv 1 + \left(\frac{p}{q}\right) \mod p.$$

- Par ailleurs, remarquons que $X = \left\{x \in \mathbb{F}_q^p : f(x) = 1\right\}$ où f est la forme quadratique associée à Id_p dans la base canonique de $\mathbb{F}_q^p$. Posons :

$$M = \mathrm{diag}(\underbrace{J, J, \dots, J}_{d = \frac{p-1}{2} \text{ fois}}, a), \quad \text{où } J = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \text{ et } a = (-1)^d.$$

  Comme $\mathrm{rg}(M) = p$ et $\det(M) = a\det(J)^d = (-1)^d(-1)^d = 1$, la forme quadratique g associée à M dans la base canonique de $\mathbb{F}_q^p$ est non dégénérée et, par classification des formes quadratiques[2], f et g sont congruentes sur $\mathbb{F}_q$. Ainsi $|X| = |X'|$ où

$$X' = \left\{x \in \mathbb{F}_q^p : g(x) = 1\right\} = \left\{x \in \mathbb{F}_q^p : 2\sum_{k=1}^{d} x_{2k}x_{2k-1} + ax_p^2 = 1\right\}.$$

  Dénombrons les éléments x de X'. Si $x_{2k+1} = 0$ pour tout $k \in [\![1\,;d]\!]$, alors $ax_p^2 = 1$ et il y a $1 + \left(\frac{a}{q}\right)$ choix pour x_p et q^d pour les $(x_{2k})_{1\leq k\leq d}$. Sinon, on a $q(q^d - 1)$ possibilités pour les $(x_{2k+1})_{1\leq k\leq d}$, puis on choisit les $(x_{2k})_{1\leq k\leq d}$ satisfaisant $2\sum_{k=1}^{d} x_{2k-1}x_{2k} = 1 - ax_p^2$, équation d'un hyperplan affine de cardinal q^{d-1}. Finalement :

$$|X| = |X'| = q^d\left(1 + \left(\frac{a}{q}\right)\right) + q^d(q^d - 1) = q^d\left(\left(\frac{a}{q}\right) + q^d\right).$$

En utilisant une nouvelle fois le lemme, il vient

$$\begin{aligned} & 1 + \left(\frac{p}{q}\right) \equiv \left(\frac{q}{p}\right)\left(\left(\frac{(-1)^{\frac{p-1}{2}}}{q}\right) + \left(\frac{q}{p}\right)\right) \mod p \\ \iff & \left(\frac{q}{p}\right) + \left(\frac{q}{p}\right)\left(\frac{p}{q}\right) \equiv \left(\frac{q}{p}\right) + \left((-1)^{\frac{p-1}{2}}\right)^{\frac{q-1}{2}} \mod p \\ \iff & \left(\frac{q}{p}\right)\left(\frac{p}{q}\right) \equiv (-1)^{\frac{(p-1)(q-1)}{2}} \mod p \end{aligned}$$

Les deux membres étant égaux à ± 1 dans $\mathbb{Z}$, le résultat en découle puisque $p \neq 2$.

COMMENTAIRES

Il faut maitriser de nombreuses notions sur les formes quadratiques : classification, égalité des cardinaux de X et X', expression de la forme quadratique à partir de sa matrice, …

Il faut connaître le cas $p = 2$. La loi de réciprocité quadratique sert notamment à résoudre des équations diophantiennes. Savoir si un élément est un carré dans $\mathbb{F}_q$ permet aussi de classifier une forme quadratique : connaissant un déterminant, il suffit de savoir si c'est un carré ou non.

1. car $\phi : \mathbb{F}_q^* \longrightarrow \mathbb{F}_q^*, a \longmapsto a^2$ est un morphisme de groupes et $|\mathrm{Im}(\phi)| = |\mathbb{F}_q^*|/|\ker(\phi)| = \frac{q-1}{2}$

2. valable pour un corps de caractéristique différente de 2 : c'est bien le cas ici puisque $q \geq 3$

16 MÉTHODE DU GRADIENT À PAS OPTIMAL POUR LA FONCTIONNELLE QUADRATIQUE

[FGN12, §1.21, p39–41]

ÉNONCÉ

THÉORÈME. [MÉTHODE DE GRADIENT À PAS OPTIMAL]
Soit $p \in \mathbb{N}^*$. Soient $A \in \mathcal{S}_n^{++}(\mathbb{R})$, $b \in \mathbb{R}^p$ et $c \in \mathbb{R}$. Définissons

$$\begin{array}{rccc} f : & \mathbb{R}^p & \longrightarrow & \mathbb{R} \\ & x & \longmapsto & \frac{1}{2}\langle Ax \mid x\rangle - \langle b \mid x\rangle + c\,. \end{array}$$

Alors la suite $(x_n)_{n\in\mathbb{N}}$ définie par $x_0 \in \mathbb{R}^p$ et la relation de récurrence

$$\forall n \in \mathbb{N} \qquad x_{n+1} = x_n - \rho_n \nabla f(x_n) \quad \text{où} \quad \rho_n = \operatorname{argmin}_{\rho>0} f\big(x_n - \rho\nabla f(x_n)\big)\,,$$

converge vers l'unique minimum global de f.

DÉVELOPPEMENT

On calcule pour commencer que $\nabla f(x) = Ax - b$ pour tout $x \in \mathbb{R}^p$.

1. L'application f est elliptique. En effet, on a si $\alpha = \inf \mathrm{Sp}(A)$:

$$\forall (x,y) \in \big(\mathbb{R}^p\big)^2 \qquad \big\langle \nabla f(x) - \nabla f(y) \mid x-y\big\rangle = \big\langle A(x-y) \mid x-y\big\rangle \geq \alpha \cdot \|x-y\|^2\,.$$

2. Montrons que f satisfait

$$\forall (x,y) \in \big(\mathbb{R}^p\big)^2 \qquad f(y) - f(x) \geq \big\langle \nabla f(x) \mid y-x\big\rangle + \frac{\alpha}{2}\cdot\|x-y\|^2\,.$$

En effet, soient $x, y \in \mathbb{R}^p$. On a :

$$\begin{aligned} f(y) - f(x) &= \frac{1}{2}\big(\langle Ay \mid y\rangle - \langle Ax \mid x\rangle\big) - \langle b \mid y-x\rangle \\ &= \langle Ax - b \mid y-x\rangle + \frac{1}{2}\big(\langle Ay \mid y\rangle + \langle Ax \mid x\rangle\big) - \langle Ax \mid y\rangle \\ &= \langle Ax-b \mid y-x\rangle + \frac{1}{2}\big\langle A(x-y) \mid (x-y)\big\rangle \\ f(y) - f(x) &\geq \big\langle \nabla f(x) \mid y-x\big\rangle + \frac{\alpha}{2}\cdot\|x-y\|^2\,. \end{aligned}$$

3. L'application f est continue et coercive puisque, pour $x \in \mathbb{R}^p$,

$$f(x) \geq \frac{\alpha}{2}\cdot\|x\|^2 - \|b\|\cdot\|x\| + c \underset{\|x\|\to+\infty}{=} \frac{\alpha}{2}\cdot\|x\|^2 + O\big(\|x\|\big) \underset{\|x\|\to+\infty}{\longrightarrow} +\infty\,.$$

Ainsi f atteint son minimum global en un point $x_\star$ vérifiant $\nabla f(x_\star) = 0$: c'est donc nécessairement $x_\star = A^{-1}b$. Notons en particulier que $x_\star$ est le seul minimum local de f.

4. Soit $x \in \mathbb{R}^p$ tel que $\nabla f(x) \neq 0$. Notons qu'alors $x \neq x_\star$. Définissons

$$\begin{array}{rccc} \varphi_x : & \mathbb{R}_+ & \longrightarrow & \mathbb{R} \\ & \rho & \longmapsto & f\big(x - \rho\nabla f(x)\big)\,. \end{array}$$

C'est une fonction de classe $\mathcal{C}^1$ qui tend vers $+\infty$ en $+\infty$. Elle atteint donc son minimum en un point ρ_x annulant la dérivée de φ_x. Calculons :

$$0 = \varphi_x'(\rho_x) = -\Big\langle \nabla f\big(x + \rho_x\nabla f(x)\big) \mid \nabla f(x)\Big\rangle\,.$$

Donc, pour tout $\rho \neq \rho_x$:

$$\varphi_x(\rho) - \varphi_x(\rho_x) \geq \underbrace{\Big\langle \nabla f\big(x+\rho_x\nabla f(x)\big) \mid (\rho-\rho_x)\nabla f(x)\Big\rangle}_{=0} + \frac{\alpha}{2}\cdot\big\|(\rho-\rho_x)\nabla f(x)\big\|^2 > 0\,.$$

On en déduit que ρ_x est l'unique minimum de φ_x.

5. La suite $(x_n)_{n\in\mathbb{N}}$ est ainsi bien définie, stationnaire en $x_\star$ si elle l'atteint. Supposons qu'elle n'atteigne pas $x_\star$, et montrons que $(x_n)_{n\in\mathbb{N}}$ converge tout de même vers $x_\star$.
La suite $\big(f(x_n)\big)_{n\in\mathbb{N}}$ est strictement décroissante et minorée, donc converge.
Soit $n \in \mathbb{N}$. Par ce qui a été démontré à l'étape précédente, remarquons que $\nabla f(x_n)$ et $\nabla f(x_{n+1})$ sont orthogonaux, et comme $\nabla f(x_n)$ et $x_{n+1} - x_n$ sont colinéaires, on obtient :

$$f(x_n) - f(x_{n+1}) \geq \frac{\alpha}{2}\cdot\|x_{n+1} - x_n\|^2\,.$$

Il en découle que $\|x_{n+1} - x_n\| \underset{n\to+\infty}{\longrightarrow} 0$.

Or, par coercivité de f, la suite $(x_n)_{n\in\mathbb{N}}$ vit dans un compact de $\mathbb{R}^p$, donc admet une valeur d'adhérence x. Soit ψ une extractrice telle que $x_{\psi(n)} \longrightarrow_{n\to+\infty} x$. Par continuité de ∇f, et puisque $(x_{\psi(n)+1})_{n\in\mathbb{N}}$ converge aussi vers x puisque $\|x_{n+1} - x_n\| \longrightarrow_{n\to+\infty} 0$, on a

$$\nabla f(x_{\psi(n)}) \underset{n\to+\infty}{\longrightarrow} \nabla f(x) \qquad \text{et} \qquad \nabla f(x_{\psi(n)+1}) \underset{n\to+\infty}{\longrightarrow} \nabla f(x)\,.$$

D'où

$$0 = \big\langle \nabla f(x_{\psi(n)}) \mid \nabla f(x_{\psi(n)+1})\big\rangle \underset{n\to+\infty}{\longrightarrow} \big\|\nabla f(x)\big\|^2\,,$$

de sorte que nécessairement $\nabla f(x) = 0$ et donc $x = x_\star$.

COMMENTAIRES

Bien penser à faire un dessin en annexe de la leçon concernée.

En ouverture, on peut rappeler que la propriété importante est l'ellipticité de f : l'algorithme de gradient à pas optimal converge en fait pour toute fonction elliptique !

17 RÉDUCTION DE JORDAN

[Rom17, Ch21, p672–675] [MM16, ChX, p111]

ÉNONCÉ

Soit E un $\mathbb{K}$-espace vectoriel de dimension finie $n \in \mathbb{N}^*$.

Théorème. [Décomposition de Jordan]
Soit $f \in \mathcal{L}(E)$ de polynôme caractéristique scindé. Notons $\lambda_1, \dots, \lambda_r$ ses valeurs propres. Alors il existe des entiers $(d_{j,k})_{1\leq j\leq r, 1\leq k\leq \ell_j}$ satisfaisants $d_{j,1} \geq \cdots \geq d_{j,\ell_j} \geq 1$ pour tout $j \in [\![1\,;r]\!]$ et tels que, dans une certaine base $\mathcal{B}$ de E, $M_{\mathcal{B}}(f)$ soit diagonale par blocs avec les blocs $(B_{j,k})_{1\leq j\leq r, 1\leq k\leq \ell_j}$, où

$$\forall j \in [\![1\,;r]\!] \quad \forall k \in [\![1\,;\ell_j]\!] \qquad B_{j,k} = \lambda_j I_{d_{j,k}} + J_{d_{j,k}}\,,$$

en notant, pour $d \in \mathbb{N}^*$,

$$J_d = \begin{pmatrix} 0 & & \cdots & & 0 \\ 1 & 0 & & & \\ 0 & 1 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & 0 & \\ 0 & \cdots & 0 & 1 & 0 \end{pmatrix} \in \mathcal{M}_d(\mathbb{K})\,.$$

DÉVELOPPEMENT

Montrons d'abord le lemme suivant.

Lemme. *Soit $f \in \mathcal{L}(E)$ nilpotent d'indice $q \in [\![1\,;n]\!]$. Soit $x \in E$ tel que $f^{q-1}(x) \neq 0$ et $\mathcal{B}_x = \big(x, f(x), \dots, f^{q-1}(x)\big)$. Alors $E_{f,x} = \mathrm{Vect}(\mathcal{B}_x)$ est f-stable et $\mathcal{B}_x$ en est une base.*

En effet, le plus petit sous-espace f-stable contenant x est $\mathbb{K}[f](x)$, et en remarquant que $f^n(x) = 0$ pour $n \geq q$, on a que $E_{f,x} = \mathbb{K}[f](x)$. Donc $E_{f,x}$ est f-stable et de dimension au plus q. Vérifions que la famille $\mathcal{B}_x$ est libre.
Soit $(\lambda_i)_{0\leq i\leq q-1} \in \mathbb{K}^q$ tels que $\sum_{i=0}^{q-1} \lambda_i f^i(x) = 0$. Alors, par linéarité,

$$\forall k \in \mathbb{N} \qquad 0 = f^k\left(\sum_{i=0}^{q-1} \lambda_i f^i(x)\right) = \sum_{i=0}^{q-1} \lambda_i f^{i+k}(x)\,.$$

Prenant $k = q-1$, il vient $\lambda_0 f^{q-1}(x) = 0$ et donc $\lambda_0 = 0$. Puis avec $k = q-2$, on obtient de même $\lambda_1 = 0$. En allant jusqu'à $k = 0$, on obtient que les $(\lambda_i)_{0\leq i\leq q-1}$ sont tous nuls, et donc que la famille $\mathcal{B}_x$ est libre. Comme elle est génératrice, c'est une base de $E_{f,x}$.

Intéressons-nous maintenant au cas où f est nilpotent.

Lemme. [Décomposition de Jordan d'un endomorphisme nilpotent]
Soit $f \in \mathcal{L}(E)$ nilpotent. Il existe des entiers $d_1 \geq \cdots \geq d_\ell$ tels que dans une certaine base $\mathcal{B}$ de E, $M_{\mathcal{B}}(f)$ soit diagonale par blocs avec les blocs $(J_{d_k})_{1\leq k\leq \ell}$.

Pour montrer ce résultat, on procède par récurrence forte sur $n = \dim(E)$.

- Pour $n = 1$, on a que f est l'endomorphisme nul et donc le résultat est vrai.
- Soient $n \geq 2$ tel que le résultat est vrai en dimensions inférieures, et f nilpotent d'ordre q. Le résultat est vérifié si $q = 1$, puisqu'alors f est nul, et si $q = n$, puisque si x satisfait le lemme, alors $E_{f,x} = E$ et la matrice de f dans la base $(x, f(x), \dots, f^{n-1}(x))$ est J_n.

  Si $1 < q < n$, on cherche une décomposition de E en sous-espaces stricts stables. Considérons $x \in E$ comme dans le lemme et cherchons un supplémentaire f-stable de $E_{f,x}$. Complétons $(x, f(x), \dots, f^{q-1}(x))$ en une base, notée[1] $(e_1, \dots, e_n)$, et posons

$$F = \Big\{y \in E : \forall j \in \mathbb{N}, e_q^*\big(f^j(y)\big) = 0\Big\} = \bigcap_{j=0}^{q-1} \ker\big(e_q^* \circ f^j\big)\,,$$

  où l'on a utilisé que $f^q = 0$ dans la seconde égalité. Il s'ensuit que F est un sous-espace vectoriel f-stable de E de dimension au moins $n-q$. Vérifions que $E_{f,x} \cap F = \{0\}$.
  Soit $y = \sum_{i=0}^{q-1} \lambda_i f^i(x) \in E_{f,x} \cap F$. On a $0 = e_q^*(y) = \lambda_{q-1}$, puis $0 = e_q^*(f(y)) = \lambda_{q-2}$, et ainsi de suite jusqu'à $0 = e_q^*(f^{q-1}(y)) = \lambda_0$. Ainsi $y = 0$ et $E_{f,x} \cap F = \{0\}$.
  Ceci implique que $\dim(F) \leq n-q$ et donc, par ce qui précède, $\dim(F) = n-q$.
  On a alors la décomposition en sous-espaces stables $E = E_{f,x} \oplus F$. La matrice de $f_{E_{f,x}}$ dans la base $(x, \dots, f^{q-1}(x))$ est J_q. En appliquant l'hypothèse de récurrence à f_F, on obtient la décomposition souhaitée en concaténant les bases.

D'où finalement le résultat par récurrence forte.

Revenons au cas général où f est de polynôme caractéristique scindé.
Écrivons $\chi_f(X) = \prod_{k=1}^r (X - \lambda_k)^{\alpha_k}$. Par le lemme des noyaux, on a que

$$E = \bigoplus_{k=1}^r \ker(f - \lambda_k \,\mathrm{id})^{\alpha_k} = \bigoplus_{k=1}^r N_k$$

est une décomposition en sous-espaces f-stables de E.
En appliquant la décomposition démontrée précédemment à la famille d'endomorphismes nilpotents $(f_{N_k} - \lambda_k \,\mathrm{id}_{N_k})_{1\leq k\leq r}$, on obtient la décomposition souhaitée dans la base de E concaténant les bases $(\mathcal{B}_k)_{1\leq k\leq r}$ trouvées.

1. de sorte que $e_q = f^{q-1}(x)$

18 RÉDUCTION DES ENDOMORPHISMES NORMAUX

[Gou09, §5.3, p258] [Rom17, §22.10, p728]

ÉNONCÉ

Soit $u \in \mathcal{L}(E)$ un endomorphisme normal, où E est un $\mathbb{R}$-espace vectoriel de dimension finie.

PROPOSITION. *Supposons $n = \dim(E) = 2$. Alors :*

- *soit u a une valeur propre réelle, et alors u est diagonalisable dans une b.o.n.,*
- *soit la matrice de u dans une b.o.n. est de la forme $\begin{pmatrix} a & -b \\ b & a \end{pmatrix}$.*

THÉORÈME. [RÉDUCTION DES ENDOMORPHISMES NORMAUX]
Il existe une base orthonormée $\mathcal{B}$ de E telle que

$$M_{\mathcal{B}}(u) = \begin{pmatrix} \operatorname{diag}(\lambda_1, \dots, \lambda_p) & & & \\ & R_1 & & \\ & & \ddots & \\ & & & R_r \end{pmatrix}$$

où $p + 2r = \dim(E)$, $(\lambda_i)_{1 \leq i \leq p} \in \mathbb{R}^p$ et, pour $1 \leq k \leq r$, $R_k = \begin{pmatrix} a_k & -b_k \\ b_k & a_k \end{pmatrix}$ avec $b_k \neq 0$.

DÉVELOPPEMENT

- Montrons d'abord que si F est u-stable alors $F^\perp$ est u^*-stable.
  Soit en effet $x \in F^\perp$. On a :
  $$\forall y \in F \qquad \langle y \mid u^*(x) \rangle = \langle u(y) \mid x \rangle = 0$$
  puisque $u(y) \in F$. Donc $u^*(x) \in F^\perp$.
- Si maintenant $F = E_\lambda$ est un sous-espace propre de u, alors $E_\lambda^\perp$ est u-stable.
  En effet, E_λ est stable par u, donc par u^* puisque ces deux endomorphismes commutent. Mais alors $E_\lambda^\perp$ est stable par $(u^\perp)^\perp = u$ par ce qui précède.
- Montrons désormais le proposition, soit le cas $\dim(E) = 2$.
  - Si u a une valeur propre réelle λ, soit e_1 un vecteur propre normé. Alors si $u \neq \lambda \operatorname{id}$, $\mathbb{R}e_1 = E_\lambda$ donc $(\mathbb{R}e_1)^\perp$ est une droite engendrée par un vecteur e_2 normé stable par u par le point précédent donc un sous-espace propre, et alors $M_{(e_1,e_2)}(u)$ est diagonale, avec (e_1, e_2) orthonormée.
  - Si u n'a pas de valeur propre réelle, écrivons $M = \begin{pmatrix} a & c \\ b & d \end{pmatrix}$ sa matrice dans une base orthonormée de E. L'endomorphisme u est normal donc $M^\top M = MM^\top$, ce qui donne $a^2 + c^2 = a^2 + b^2$ et $ab + cd = ac + bd$.
    Si $b = c$, alors M est symétrique donc admet une valeur propre réelle[1], ce qui est absurde. Donc $c = -b$, avec $b \neq 0$ sinon M serait diagonale, et donc $d = a$. Finalement
    $$M = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}.$$
- Passons maintenant à la preuve du théorème, par récurrence forte sur $n = \dim(E)$:
  - Si $n = 1$, le résultat est évident et on vient de traiter le cas $n = 2$.
  - Supposons $n \geq 3$ et le résultat vrai en dimension inférieure à n.
    Si u admet une valeur propre réelle λ_1, alors $E_{\lambda_1}^\perp$ est stable par u. On peut donc appliquer l'hypothèse de récurrence à $u_{|E_{\lambda_1}^\perp}$, et on obtient le résultat en concaténant une base orthonormée de E_{λ_1} avec la base obtenue, qui reste bien orthonormée.
    Sinon, soit Q un facteur irréductible de χ_u. Ce facteur Q est nécessairement de degré 2 et s'écrit sous la forme
    $$Q = (X - \lambda)(X - \overline{\lambda}), \quad \text{où} \lambda \in \mathbb{C} \setminus \mathbb{R}.$$
    Le noyau $\ker Q(u)$ est alors non nul, puisque si M est la matrice de u dans une base, alors λ est valeur propre de M et
    $$\det\big(Q(M)\big) = \det(M - \lambda I_n) \det(M - \overline{\lambda} I_n) = 0,$$
    donc $Q(u)$ est non inversible.
    Soit $v = u_{|\ker Q(u)}$. Alors $v^* v$ est symétrique donc[2] admet une valeur propre $\mu \in \mathbb{R}$. Soit x un vecteur propre associé et
    $$F = \operatorname{Vect}\big(x, u(x)\big) = \operatorname{Vect}\big(u(x), u^2(x)\big).$$
    Le sous-espace F est stable par u et de dimension 2. Vérifions que F est également stable par u^*. D'une part :
    $$u^*\big(u(x)\big) = v^*\big(v(x)\big) = \mu\, x \in F,$$
    et, d'autre part :
    $$u^*\big(u^2(x)\big) = u\big(u^*(u(x))\big) = u(\mu\, x) = \mu\, u(x) \in F.$$
    On peut donc appliquer l'hypothèse de récurrence à $u_{|F^\perp}$ et le cas $n = 2$ à $u_{|F}$, ce qui donne le résultat voulu dans la base concaténée.

1. notons que dans ce cas, pas besoin d'invoquer la réduction des endomorphismes symétriques : il suffit de calculer le polynôme caractéristique de la matrice et de voir que ces deux racines sont réelles
2. là encore on n'utilise seulement que les valeurs propres d'une matrice symétrique sont réelles

19 SIMPLICITÉ DE $\mathfrak{A}_n$ POUR $n \geq 5$

[Per96, §I.8, p28–29] [Rom17, §2.9, p67]

ÉNONCÉ

PROPOSITION. *Pour $n \geq 5$, le groupe alterné $\mathfrak{A}_n$ est simple.*

DÉVELOPPEMENT

Fixons $n \geq 5$ et commençons par démontrer le lemme suivant :

LEMME. *Le groupe alterné satisfait les propriétés suivantes :*

i) $\mathfrak{A}_n$ est le sous-groupe engendré par les trois cycles,

ii) Les 3-cycles sont conjugués dans $\mathfrak{A}_n$.

i) Notons G le sous-groupe engendré par les 3-cycles.

On vérifie que $G \subset \mathfrak{A}_n$ car tous les éléments de G ont pour signature 1.

Réciproquement, fixons $\sigma \in \mathfrak{A}_n$, et considérons une décomposition de σ en produit de transpositions :

$$\sigma = \prod_{i=1}^{p} \tau_i \,.$$

Comme $\mathcal{E}(\sigma) = 1$, notons que p est nécessairement pair.
Or, on remarque que pour $i, j, k, \ell \in [\![1\,;n]\!]$ tous distincts, on a :

$$(i\,j)\,(k\,\ell) = (i\,j\,k)\,(j\,k\,\ell) \quad \text{et} \quad (i\,j)\,(k\,i) = (i\,k\,j)\,.$$

Ainsi, le produit de deux transpositions distinctes est un 3-cycle. En particulier, σ s'écrit comme produit de $p/2$ (au plus) 3-cycles, et donc $\sigma \in G$.

Finalement on a bien $G = \mathfrak{A}_n$.

ii) Soit $a, b, c \in [\![1\,;n]\!]$ distincts et $d, e, f \in [\![1\,;n]\!]$ distincts. Montrons que $(a\,b\,c)$ et $(d\,e\,f)$ sont conjugués dans $\mathfrak{A}_n$. Comme les 3-cycles forment une classe de conjugaison dans $\mathfrak{S}_n$, il existe $\sigma \in \mathfrak{S}_n$ tel que

$$\sigma\,(a\,b\,c)\,\sigma^{-1} = (d\,e\,f)\,.$$

Si $\sigma \in \mathfrak{A}_n$, le résultat est vérifié.

Sinon, comme $n \geq 5$, on peut choisir $i, j \in [\![1\,;n]\!] \setminus \{a, b, c\}$ distincts et alors $\sigma' = \sigma\,(i\,j)$ est un élément de $\mathfrak{S}_n$ tel que

$$\sigma'\,(a\,b\,c)\,(\sigma')^{-1} = (d\,e\,f)\,.$$

Ainsi, les 3-cycles sont bien conjugués dans $\mathfrak{A}_n$.

Passons à la proposition. Soit H un sous-groupe distingué de $\mathfrak{A}_n$ distinct de $\{\mathrm{Id}\}$. Montrons que $H = \mathfrak{A}_n$. Par le lemme, il suffit de montrer que H possède un 3-cycle.

Soit $\sigma \in H \setminus \{\mathrm{Id}\}$. Choisissons $a \in [\![1\,;n]\!]$ tel que $b = \sigma(a) \neq a$. Fixons $c \in [\![1\,;b]\!] \setminus \{a, b, \sigma(b)\}$ (ce qui est possible car $n \geq 5$) et considérons le 3-cycle $\gamma = (a\,b\,c)$ puis

$$\sigma_2 = \sigma\,\gamma\,\sigma^{-1}\,\gamma^{-1}\,.$$

On a $\sigma_2 \in H$ car $\sigma \in H$ et $\gamma\,\sigma^{-1}\,\gamma^{-1} \in H$, et on vérifie que

$$\sigma_2 = \big(b\,\sigma(b)\,\sigma(c)\big)\,(a\,c\,b)\,.$$

On va décomposer σ_2 en produit de cycles à support disjoints, en remarquant que

$$\mathrm{Supp}(\sigma_2) \subset \{a, b, c, \sigma(b), \sigma(c)\}$$

a au plus 5 éléments. En raisonnant sur le type de σ_2, seuls les 4 cas suivants se produisent :

(5) Si $\sigma_2 = (i\,j\,k\,\ell\,m)$, alors, comme H est distingué dans $\mathfrak{A}_n$,

$$\sigma_3 = (i\,j\,k)\,\sigma_2\,(i\,j\,k)^{-1}\sigma_2^{-1} \in H\,.$$

On obtient donc un 3-cycle de $\mathfrak{A}_n$ puisque

$$\sigma_3 = (i\,j\,k)\,(j\,\ell\,k) = (i\,j\,\ell)\,.$$

(3, 1, 1) Si σ_2 est un 3-cycle, il n'y a rien à montrer.

(2, 2, 1) Si $\sigma_2 = (i\,j)\,(k\,\ell)$, alors de même

$$\sigma_3 = (i\,j\,k\,\ell\,m)\,\sigma_2\,(i\,j\,k\,\ell\,m)^{-1}\,\sigma_2^{-1} \in H\,.$$

On vérifie que $\sigma_3 = (i\,j\,k\,\ell\,m)\,(j\,i\,\ell\,k\,m)^{-1} = (i\,k\,m\,\ell\,j)$, et on conclut comme dans le cas où le type est (5).

(1, 1, 1, 1, 1) Si $\sigma_2 = \mathrm{Id}$, σ et γ commutent, ce qui est faux puisque, par construction de c :

$$\sigma\,\gamma(a) = \sigma(b) \neq c = \gamma\,\sigma(a)\,.$$

Dans tous les cas possibles, on a bien trouvé un 3-cycle dans H. Ainsi $H = \mathfrak{A}_n$.

COMMENTAIRES

Que se passe-t-il pour $n < 5$?

- $\mathfrak{A}_3 = \{\mathrm{Id}, (1\,2\,3), (1\,3\,2)\} \simeq \mathbb{Z}/3\mathbb{Z}$ est simple,
- $\mathfrak{A}_4 = \{\mathrm{Id}, \text{3-cycles}, \text{bi-transpositions}\}$ n'est pas simple, le groupe des bi-transpositions étant distingué dans $\mathfrak{A}_4$. De plus, les 3-cycles ne sont pas conjugués, sinon le cardinal de $\mathfrak{A}_4$ serait divisible par 8, ce qui n'est pas le cas.

20 SOUS-GROUPES DISTINGUÉS ET CARACTÈRES. TABLE DE $\mathfrak{S}_4$

[Pey04, §VIII.1.4/VIII.2.1, p228–232]

ÉNONCÉ

Soit G un groupe fini. Soient $\chi_1, \dots, \chi_r$ ses caractères irréductibles.

LEMME. *Soit χ associé à une représentation (ρ, V) de G. Alors*

$$\ker(\rho) = \ker(\chi) = \big\{g \in G : \chi(g) = \chi(1)\big\}\,.$$

PROPOSITION. *Les sous-groupes distingués de G sont les $\big(\bigcap_{i\in I} \ker\chi_i\big)_{I\subset[\![1\,;r]\!]}$.*

APPLICATION. Table des caractères de $\mathfrak{S}_4$.

DÉVELOPPEMENT

- Commençons par le lemme. On procède par double inclusion :
  - Si $\rho(g) = \rho(1) = \mathrm{id}$, alors $\chi(g) = \mathrm{Tr}(\mathrm{id}) = \dim(V) = \chi(1)$. Donc $\ker(\rho) \subset \ker(\chi)$.
  - Réciproquement, soit $g \in G$ tel que $\chi(g) = \chi(1) = \dim(V)$.
    $\rho(g)$ est diagonalisable, de valeurs propres des racines de l'unité $\zeta_1, \dots, \zeta_d$ où $d = \dim(V)$. Comme $\chi(g)$ en est la somme, on a s'il existe $i \in [\![1\,;d]\!]$ tel que $\zeta_i \neq 1$:
    $$\Re\big(\chi(g)\big) = \sum_{i=1}^{d} \Re(\zeta_i) < d = \Re\big(\chi(1)\big)\,,$$
    ce qui est absurde. Donc $\rho(g)$ est diagonalisable avec pour seule valeur propre 1 : c'est donc l'identité et ainsi $g \in \ker(\rho)$. D'où l'inclusion réciproque.
- Passons à la proposition. Une intersection de groupes distingués étant distinguée, tout sous-groupe de la forme $\bigcap_{i\in I} \ker\chi_i$, où I est une partie $[\![1\,;r]\!]$, est distingué.
  Réciproquement, soit N distingué dans G. Notons $\mathbb{C}G/N = \mathrm{Vect}\big((e_{\overline{g}})_{g\in G}\big)$, où $\overline{g} = gN$ pour $g \in G$, et considérons la représentation de G sur $\mathbb{C}G/N$:
  $$\rho : G \longrightarrow \mathcal{GL}\big(\mathbb{C}G/N\big), g \longmapsto \big(e_{\overline{h}} \longmapsto e_{\overline{gh}}\big)\,, \quad \text{qui vérifie}$$
  $$\ker(\rho) = \big\{g \in G : \forall h \in G, \overline{gh} = \overline{h}\big\} = \big\{g \in G : \forall h \in G, h^{-1}gh \in N\big\} = N\,.$$
  Décomposons $\mathbb{C}G/N$ en somme de représentations irréductibles $\oplus_{j=1}^{s} V_j$.
  Notant θ_j le caractère de V_j pour $j \in [\![1\,;s]\!]$, puis $a_i = \mathrm{card}\big(\{j \in [\![1\,;s]\!] : \theta_j = \chi_i\}\big)$ pour $i \in [\![1\,;r]\!]$, le caractère associé à ρ se décompose en $\chi = \sum_{i=1}^{r} a_i\chi_i = \sum_{i\in I} a_i\chi_i$ où $I = \{i \in [\![1\,;r]\!] : a_i > 0\}$. Finalement :
  $$N = \ker(\rho) = \bigcap_{j=1}^{r} \ker(\rho_{|V_j}) = \bigcap_{i\in I} \ker(\chi_i)\,.$$
- Regardons maintenant la table de $\mathfrak{S}_4$ afin d'en déduire ses sous-groupes distingués.
  Il y a 5 classes de conjugaison : l'identité, 6 transpositions, 8 3-cycles, 3 bi-transpositions et 6 4-cycles. On sait donc qu'il y a 5 caractères irréductibles. Parmi eux, les caractères irréductibles de degré 1 sont le caractère trivial χ_1 et la signature χ_ε.
  De plus, $|\mathfrak{S}_4| = 24 = 1+1+4+9+9$ est la seule décomposition possible de $|\mathfrak{S}_4|$ en somme de 5 carrés dont deux exactement valent 1, si bien que les caractères irréductibles restants sont de degré 2, 3 et 3. Déterminons-les.
  Regardons la représentation par permutation $\rho_{\text{perm}} : \mathfrak{S}_4 \longrightarrow \mathrm{GL}_4(\mathbb{C})$ qui possède pour sous-représentation $\mathrm{Vect}\big((1,1,1,1)\big)$. Ainsi, $\chi_{\text{perm}} = \chi_1 + \chi_3$ pour un caractère χ_3 et on peut calculer, pour toute classe C :
  $$\forall \sigma \in C \qquad \chi_3(\sigma) = \mathrm{card}\big(\mathrm{Fix}(\sigma)\big) - 1\,.$$
  En vérifiant que $\langle \chi_3 \mid \chi_3\rangle = 1$, on sait que χ_3 est irréductible.
  Soit maintenant la représentation naturelle de $\mathfrak{S}_4$ dans $\mathrm{Isom}^+(\mathcal{C})$ le groupe des isométries positives du cube. Calculons son caractère χ_c associé :
  - l'identité est envoyée sur l'identité, de trace 3,
  - une transposition est envoyée sur une rotation d'angle π d'axe passant par le milieu d'une arête, de trace -1,
  - un 4-cycle est envoyé sur une rotation d'angle $\frac{\pi}{2}$ d'axe $(O\,x)$ par exemple, de trace 1,
  - une bi-transposition est envoyée sur une rotation d'angle π d'axe $(O\,x)$ par exemple, de trace -1,
  - un 3-cycle est envoyé sur une rotation d'angle $\frac{2\pi}{3}$ d'axe passant par un sommet, de trace 0.

  Finalement, on a déterminé χ_c et on vérifie qu'il est irréductible.
  En utilisant l'orthogonalité sur les colonnes, on obtient χ_2 le caractère irréductible de degré 2, et on peut dresser la table des caractères de $\mathfrak{S}_4$.

TABLE 20.1 – Table des caractère de $\mathfrak{S}_4$

Type	$(1,1,1,1)$	$(2,1,1)$	$(3,1)$	(4)	$(2,2)$
χ_1	1	1	1	1	1
χ_ε	1	-1	1	-1	1
χ_3	3	1	0	-1	-1
χ_c	3	-1	0	1	-1
χ_2	2	0	-1	0	2

En appliquant la proposition, les sous-groupes distingués de $\mathfrak{S}_4$ non triviaux sont donc le groupe alterné $\mathfrak{A}_4$ ainsi que le groupe des bi-transpositions.

21 STRUCTURE DES GROUPES ABÉLIENS FINIS

[Col11, §I.2.4, p250–252]

ÉNONCÉ

Soit G un groupe abélien fini.

DÉFINITION. On appelle exposant de G le PPCM des ordres de ses éléments.

THÉORÈME. [STRUCTURE DES GROUPES ABÉLIENS FINIS]
Il existe un unique entier ℓ et une unique suite $d_1 \geq d_2 \geq \cdots \geq d_\ell$ d'entiers supérieurs ou égaux à 2 tels que d_1 est l'exposant de G, $d_{i+1} \mid d_i$ pour tout $i \in [\![1\,;\ell-1]\!]$ et $G \simeq \prod_{i=1}^{\ell} \mathbb{Z}/d_i\mathbb{Z}$.

DÉVELOPPEMENT

La preuve nécessite trois résultats préliminaires.

1. L'application $\iota : G \longrightarrow \hat{\hat{G}}, g \longmapsto (\chi \longmapsto \chi(g))$ est un isomorphisme de groupes.
   Pour $g \in G$, on vérifie déjà que $\iota(g) \in \hat{\hat{G}}$. En effet, pour $g \in G$ et $\chi_1, \chi_2 \in \hat{G}$, on a
   $$\iota(g)(\chi_{\text{triv}}) = 1 \qquad \text{et} \qquad \iota(g)(\chi_1\chi_2) = (\chi_1\chi_2)(g) = \chi_1(g)\chi_2(g) = \iota(g)(\chi_1)\iota(g)(\chi_2)\,.$$
   Ensuite, ι est bien un morphisme de groupes puisque pour $g, h \in G$ et $\chi \in \hat{G}$:
   $$\iota(e)(\chi) = \chi(e) = 1 \qquad \text{et} \qquad \iota(gh)(\chi) = \chi(gh) = \chi(g)\chi(h) = \iota(g)(\chi)\iota(h)(\chi)\,.$$
   On sait que $G, \hat{G}$ puis $\hat{\hat{G}}$ ont même cardinal, donc il reste à montrer que ι est injective. Soit donc $g \in G$ tel que $\iota(g) = 1$. Les caractères de $\hat{G}$ formant une base de l'ensemble des fonctions centrales de G dans $\mathbb{C}$[1], écrivons :
   $$\mathbb{1}_g = \sum_{\chi \in \hat{G}} \langle \mathbb{1}_g \mid \chi \rangle \, \chi\,, \qquad \text{où } \langle \mathbb{1}_g \mid \chi \rangle = \frac{1}{|G|} \sum_{h \in G} \mathbb{1}_g(h)\overline{\chi(h)} = \frac{\overline{\chi(g)}}{|G|} = \frac{1}{|G|} \text{ pour } \chi \in \hat{G}\,.$$
   En évaluant en e, il vient $\mathbb{1}_g(e) = \sum_{\chi \in \hat{G}} \frac{1}{|G|}\chi(e) = \sum_{\chi \in \hat{G}} \frac{1}{|G|} = 1$, soit $g = e$.
2. Il existe $g \in G$ tel que g est d'ordre égal à l'exposant de G.
   Vérifions pour cela que l'ensemble des ordres des éléments de G est stable par PPCM, puisqu'alors en un nombre fini d'itérations (G est fini), on trouvera un élément satisfaisant.
   Soient $x, y \in G$ d'ordre respectif a et b. Trouvons un élément d'ordre $a \vee b$. Écrivons
   $$k = \prod_{p \in \mathcal{P} \,:\, v_p(a) > v_p(b)} p^{v_p(a)} \qquad \text{et} \qquad \ell = \prod_{p \in \mathcal{P} \,:\, v_p(a) \leq v_p(b)} p^{v_p(b)}\,.$$
   de sorte que $a \vee b = k\ell$ et que k et ℓ sont premiers entre eux.
   Posant $x' = x^{a/k}$ et $y' = y^{b/\ell}$ d'ordre respectif k et ℓ, le produit $x'y'$ est d'ordre $k\ell = a \vee b$.
3. G et $\hat{G}$ ont même exposant.
   Soit en effet H un groupe commutatif fini, d'exposant d. Alors si $\chi \in \hat{H}$, on a
   $$\forall h \in H \qquad \chi^d(h) = \chi(h)^d = \chi(h^d) = \chi(e) = 1\,,$$
   donc $\chi^d = 1$. Ainsi l'exposant de $\hat{H}$ est plus petit que celui de H. Appliquant à $H = G$ puis à $H = \hat{G}$, on a $d(\hat{\hat{G}}) \leq d(\hat{G}) \leq d(G)$ et, puisque $G \simeq \hat{\hat{G}}$, ce sont en fait des égalités.

Passons maintenant à la preuve du théorème. On procède par récurrence sur $|G|$.

- Avec la convention $\prod_{i=1}^{0} = \{e\}$, le résultat est évident si $|G| = 1$ en prenant $\ell = 0$.
- Supposons donc $|G| > 1$ et le résultat vrai pour tout groupe H tel que $|H| < |G|$.
  Soit d l'exposant de G. On sait que d est aussi l'exposant de $\hat{G}$, donc on peut trouver $\chi_1 \in \hat{G}$ tel que χ_1 est d'ordre d. Comme $\chi(g)$ est une racine d-ième de l'unité pour $\chi \in \hat{G}$ et $g \in G$, on remarque que $\chi_1(G)$ est un sous-groupe de $\mathbb{U}_d$.
  Par ailleurs, si $\big|\chi_1(G)\big| = d' < d$, alors $\chi_1^{d'}(g) = 1$ pour tout $g \in G$, donc $o(\chi_1) \leq d' < d$, ce qui est faux. Ainsi $\chi_1(G) = \mathbb{U}_d$.
  Choisissons $g_1 \in G$ tel que $\chi_1(g_1) = \mathrm{e}^{\frac{2i\pi}{d}}$. Il s'ensuit que g_1 est d'ordre d (on savait déjà que $o(g_1) \mid d$), et ainsi $H_1 = \langle g_1 \rangle$ est cyclique d'ordre d donc isomorphe à $\mathbb{Z}/d\mathbb{Z}$.
  Vérifions que $G = G_1 \times H_1$ où $G_1 = \ker(\chi_1)$.
  - D'une part, on a l'isomorphisme $\chi_1(H_1) \simeq \mathbb{U}_d$, ce qui assure que $H_1 \cap \ker(\chi_1) = \{e\}$,
  - D'autre part, si $g \in G$, on peut choisir $h \in H_1$ tel que $\chi_1(h) = \chi_1(g)$ et alors
    $$gh^{-1} \in \ker(\chi_1) = G_1 \qquad \text{d'où} \qquad g = (gh^{-1})h \in G_1H_1\,.$$
    Ainsi $G = G_1H_1$.

  En appliquant l'hypothèse de récurrence à G_1, de cardinal strictement inférieur à celui de G, on obtient que $G \simeq \prod_{i=2}^{\ell} \mathbb{Z}/d_i\mathbb{Z}$ pour un ℓ et $d_2 \geq \cdots \geq d_\ell$. En vérifiant que $d \mid d_2$, c'est-à-dire que d_2 l'exposant de G_1 divise d l'exposant de G (ce qui est clair), on obtient l'isomorphisme souhaité avec $d_1 = d$.
  D'où l'hypothèse de récurrence au rang $|G|$.

On conclut par principe de récurrence.

1. en fait l'ensemble des fonctions de G dans $\mathbb{C}$ puisque G est abélien

COMMENTAIRES

Ce développement un peu long peut être raccourci en admettant le premier point dans le plan.

Attention, on ne fait que la preuve d'existence dans ce développement. On pourra trouver une preuve de l'unicité dans [Rom17, §1.9, p29–30].

22 SUITES DE POLYGONES

[Gou09, §4.2, p180]

ÉNONCÉ

LEMME. [DÉTERMINANT CIRCULANT]
Soient $n \in \mathbb{N}^*$ puis $a_1, \dots, a_n \in \mathbb{C}$. En posant $Q(X) = \sum_{i=0}^{n-1} a_{i+1} X^i$ et $\omega = \mathrm{e}^{\frac{2i\pi}{n}}$, on a :

$$\begin{vmatrix} a_1 & a_2 & \cdots & a_n \\ a_n & a_1 & \cdots & a_{n-1} \\ \vdots & \vdots & \ddots & \vdots \\ a_2 & a_3 & \cdots & a_1 \end{vmatrix} = \prod_{\ell=0}^{n-1} Q(\omega^\ell)\,.$$

PROPOSITION. [CONVERGENCE D'UNE SUITE DE POLYGONES VERS L'ISOBARYCENTRE]
On définit par récurrence une suite $\left(P^{(k)}\right)_{k\in\mathbb{N}}$ par $P^{(0)} = \left(z_1^{(0)}, \dots, z_n^{(0)}\right) \in \mathbb{C}^n$ et, pour $k \in \mathbb{N}$, $P^{(k+1)} = \left(\frac{z_1^{(k)}+z_2^{(k)}}{2}, \dots, \frac{z_{n-1}^{(k)}+z_n^{(k)}}{2}, \frac{z_n^{(k)}+z_1^{(k)}}{2}\right)$. Alors

$$P^{(k)} \underset{k\to+\infty}{\longrightarrow} (g, g, \dots, g)\,, \quad \text{où } g = \mathrm{Isobar}\left(z_1^{(0)}, \dots, z_n^{(0)}\right).$$

DÉVELOPPEMENT

Commençons par le lemme. Posons $\omega = \mathrm{e}^{2i\pi/n}$ puis :

$$J = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & 0 & \ddots & 0 \\ 0 & & \ddots & \ddots & 1 \\ 1 & 0 & \cdots & 0 & 0 \end{pmatrix} \quad \text{et} \quad \Omega = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ 1 & \omega & \cdots & \omega^{n-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \omega^{n-1} & \cdots & \omega^{(n-1)^2} \end{pmatrix} = \left(\omega^{(i-1)(j-1)}\right)_{\substack{1\le i\le n \\ 1\le j\le n}}.$$

Remarquons que Ω est une matrice de VANDERMONDE de déterminant non nul puisque les $(\omega^i)_{0\le i\le n-1}$ sont tous distincts. On cherche le déterminant de $A = Q(J)$. Calculons :

$$A\,\Omega = \begin{pmatrix} Q(1) & Q(\omega) & \cdots & Q(\omega^{n-1}) \\ Q(1) & \omega\,Q(\omega) & \cdots & \omega^{n-1}\,Q(\omega^{n-1}) \\ \vdots & \vdots & \ddots & \vdots \\ Q(1) & \omega^{n-1}\,Q(\omega) & \cdots & \omega^{(n-1)^2}\,Q(\omega^{n-1}) \end{pmatrix}$$

En utilisant la linéarité du déterminant par rapport à chaque colonne, on a donc :

$$\det(A\,\Omega) = Q(1)\,Q(\omega)\cdots Q\left(\omega^{n-1}\right)\,\det(\Omega)\,.$$

Puisque $\det(\Omega) \neq 0$, on obtient alors $\det(A) = \prod_{\ell=0}^{n-1} Q\left(\omega^\ell\right)$.

Passons à la proposition.
Remarquons que la relation de récurrence se réécrit $P^{(k+1)} = A\,P^{(k)}$ où $A = \frac{1}{2}I_n + \frac{1}{2}J$. Il est alors immédiat de vérifier que $P^{(k)} = A^k P^{(0)}$ pour tout $k \in \mathbb{N}$.
Étudions donc la matrice A en cherchant ses valeurs propres. Pour $\lambda \in \mathbb{C}$, le lemme donne :

$$\chi_A(\lambda) = \det(\lambda I_n - A) = \det\left(\left(\lambda - \frac{1}{2}\right)I - \frac{1}{2}J\right) = \prod_{\ell=0}^{n-1}\left(\left(\lambda-\frac{1}{2}\right) - \frac{1}{2}\omega^\ell\right) = \prod_{\ell=0}^{n-1}\left(\lambda - \frac{1+\omega^\ell}{2}\right).$$

Les valeurs propres de A sont donc les $(\lambda_\ell)_{0\le\ell\le n-1} = \left(\frac{1+\omega^\ell}{2}\right)_{0\le\ell\le n-1}$. Comme elles sont toutes distinctes, A est diagonalisable. Écrivons $A = R\cdot D\cdot R^{-1}$ où $D = \mathrm{diag}(\lambda_0, \dots, \lambda_{n-1})$. En remarquant que $|\lambda_\ell| < 1$ pour $\ell \in [\![1\,;n-1]\!]$, on en déduit que :

$$A^k = R\cdot D^k\cdot R^{-1} \underset{k\to+\infty}{\longrightarrow} R\cdot\mathrm{diag}(1,0,\dots,0)\cdot R^{-1}\,, \quad \text{puis}$$
$$P^{(k)} \underset{k\to+\infty}{\longrightarrow} \underbrace{R\cdot\mathrm{diag}(1,0,\dots,0)\cdot R^{-1}\cdot P^{(0)}}_{P^{(\infty)}}\,.$$

$P^{(\infty)}$ est alors un point fixe de A, c'est-à-dire un vecteur propre associé à la valeur propre 1. Comme le sous-espace propre de A associé à 1 est de dimension 1 et contient $(1,\dots,1)$:

$$\exists g \in \mathbb{C} \qquad P^{(\infty)} = (g, \dots, g)\,.$$

Reste à remarquer que les isobarycentres des $\left(P^{(k)}\right)_{k\in\mathbb{N}}$ sont tous égaux par associativité des barycentres, et donc que

$$g = \mathrm{Isobar}\left((g,\dots,g)\right) = \mathrm{Isobar}\left(P^{(\infty)}\right) = \lim_{k\to+\infty}\mathrm{Isobar}\left(P^{(k)}\right) = \mathrm{Isobar}\left(P^{(0)}\right),$$

le passage à la limite étant justifié par continuité de l'application $(z_1,\dots,z_n) \longmapsto \frac{1}{n}\sum_{i=1}^n z_i$.

COMMENTAIRES

On pourra insister sur l'intuition géométrique, en assimilant la suite $\left(P^{(k)}\right)_{k\in\mathbb{N}}$ à une suite de polygones définis par leurs sommets, via l'identification entre $\mathbb{C}$ et $\mathbb{R}^2$.

Bien que circulante, la matrice A utilisée pour démontrer la proposition est creuse : on peut donc calculer son déterminant par un simple développement, sans connaître la formule du déterminant circulant. Toutefois, appliquer la formule ici à l'avantage de fournir directement une forme factorisée du déterminant, et redémontrer cette formule est aussi pertinent.

Attention, seul le lemme est dans le [Gou09]. La proposition n'est pas référencée, elle doit donc être bien maitrisée. Un exercice similaire est cependant proposé dans le [FGN07b, §1.22, p45].

23 THÉORÈME DE CARATHÉODORY

[Gou08, §1.5, p54]

ÉNONCÉ

THÉORÈME. [THÉORÈME DE CARATHÉODORY]
Soit X un espace affine de dimension finie n. Alors pour tout $S \subset X$, $\mathrm{conv}(S)$ est l'ensemble des barycentres à coefficients positifs d'au plus $n+1$ points de S. Autrement dit :

$$\mathrm{conv}(S) = \left\{ \sum_{i=1}^{n+1} \lambda_i x_i : (x_i, \lambda_i)_{1\leq i\leq n+1} \in (S \times \mathbb{R}_+)^{n+1} \right\}.$$

COROLLAIRE. *Soit S un compact de X euclidien. Alors $\mathrm{conv}(S)$ est compact.*

APPLICATION. Soit $M \in \mathcal{M}_n(\mathbb{Z})$. Notons $(C_j)_{1\leq j\leq n}$ ses colonnes. L'équation diophantienne $MX = 0$ admet une solution non nulle dans $\mathbb{N}^n$ si et seulement si $0 \in \mathrm{conv}\big(\{C_1, \dots, C_n\}\big)$.

DÉVELOPPEMENT

1. Soit $S \subset X$ et $x \in \mathrm{conv}(S)$. Écrivons $x = \sum_{i=1}^p \lambda_i x_i$ pour des $(x_i)_{1\leq i\leq p} \in S^p$ et des $(\lambda_i)_{1\leq i\leq p} \in (\mathbb{R}_+^*)^p$ tels que $\sum_{i=1}^p \lambda_i = 1$, et de sorte que p soit minimal.
   Supposons par l'absurde que $p \geq n+2$. Les vecteurs $\{x_1, \dots, x_p\}$ étant liés dans X :

   $$\exists (\alpha_2, \dots, \alpha_p) \in \mathbb{R}^{p-1} \setminus \{0_{\mathbb{R}^{p-1}}\} \qquad \sum_{i=2}^p \alpha_i \overrightarrow{x_1 x_i} = \vec{0}.$$

   En posant $\alpha_1 = -\sum_{i=2}^p \alpha_i$, on a ainsi $\sum_{i=1}^p \alpha_i x_i = 0$ et $\sum_{i=1}^p \alpha_i = 0$. En particulier,

   $$\forall t \in \mathbb{R} \qquad x = \sum_{i=1}^p (\lambda_i + t\alpha_i) x_i \quad \text{et} \quad \sum_{i=1}^p \lambda_i + t\alpha_i = 1.$$

   Considérons alors $F = \big\{ t \in \mathbb{R} : \forall i \in [\![1\,;p]\!], \lambda_i + t\alpha_i \geq 0 \big\}$. L'ensemble F n'est pas vide (il contient 0) et n'est pas $\mathbb{R}$ tout entier puisque l'un au moins des $(\alpha_i)_{1\leq i\leq p}$ est non nul. Ainsi, F admet une borne inférieure et/ou[1] supérieure notée t_0 et de la forme $-\lambda_{i_0}/\alpha_{i_0}$ pour un certain $i_0 \in [\![1\,;p]\!]$. F étant fermé, il est clair que $t_0 \in F$, ce qui implique que :

   $$x = \sum_{i=1}^p (\lambda_i - t_0\alpha_i) x_i = \sum_{\substack{i=1 \\ i\neq i_0}}^p (\lambda_i - t_0\alpha_i) x_i \qquad \text{avec} \qquad \forall i \in [\![1\,;p]\!] \quad \lambda_i - t_0\alpha_i \geq 0.$$

   Ceci contredit la minimalité de p. On en déduit que $p \leq n+1$, ce qui permet de conclure.

2. Soit S un compact de X. Définissons :

   $$\begin{array}{rccc} f : & S^{n+1} \times \big\{(\lambda_i)_{1\leq i\leq n+1} \in \mathbb{R}_+^{n+1} : \sum_{i=1}^{n+1} \lambda_i = 1\big\} & \longrightarrow & \mathrm{conv}(S) \\ & (x_1, \dots, x_{n+1}, \lambda_1, \dots, \lambda_{n+1}) & \longmapsto & \displaystyle\sum_{i=1}^{n+1} \lambda_i x_i. \end{array}$$

   D'après le théorème de CARATHÉODORY, cette application est bien définie et est surjective. Comme elle est continue et définie sur un compact, son image $\mathrm{conv}(S)$ est compacte.
3. Soit $M \in \mathcal{M}_n(\mathbb{Z})$.
   - S'il existe $X \in \mathbb{N}^n$ non nul tel que $MX = 0$, alors :

     $$\sum_{i=1}^n x_i C_i = 0 \qquad \text{ou encore} \qquad \sum_{i=1}^n \frac{x_i}{\sum_{j=1}^n x_j} C_i = 0,$$

     et donc $0 \in \mathrm{conv}(C_1, \dots, C_n)$.
   - Réciproquement, supposons que $0 \in \mathrm{conv}(C_1, \dots, C_n)$.
     Si l'une des colonnes est nulle, le résultat est évident.
     Sinon, soit $p \in \mathbb{N}^*$ minimal tel que l'on puisse écrire $0 = \sum_{j=1}^p \lambda_j C_{i_j}$ pour des $(\lambda_j)_{1\leq j\leq p} \in (\mathbb{R}_+^*)^p$ et p colonnes distinctes $(C_{i_j})_{1\leq j\leq p}$.
     Notons $r = \mathrm{rg}(C_{i_1}, \dots, C_{i_p})$. D'une part, on a $r < p$ puisque les colonnes sont liées. D'autre part, $\mathrm{conv}(C_{i_1}, \dots, C_{i_p}) \subset \mathrm{Vect}(C_{i_1}, \dots, C_{i_p})$ espace de dimension r, donc le théorème de CARATHÉODORY donne $p \leq r+1$ par minimalité de p. D'où $r = p-1$.
     Les sous-espaces $\ker_{\mathbb{Q}}(C_{i_1}, \dots, C_{i_p}) \subset \ker_{\mathbb{R}}(C_{i_1}, \dots, C_{i_p})$ sont donc de dimension[2] 1, dirigés par un vecteur $\Lambda = (\lambda_1, \dots, \lambda_p)$ à coefficients réels. Ainsi,

     $$\exists \alpha \in \mathbb{R}_+^* \qquad \alpha\Lambda = (\mu_1, \dots, \mu_p) \in (\mathbb{Q}_+^*)^p \quad \text{et} \quad \sum_{j=1}^p \mu_j M_{i_j} = 0.$$

     En multipliant par le produit d des dénominateurs de $\alpha\Lambda$, on a donc :

     $$\sum_{j=1}^p d\,\mu_j M_{i_j} = 0, \qquad \text{où} \qquad \forall j \in [\![1\,;p]\!] \quad d\,\mu_j \in \mathbb{N}.$$

     Autrement dit, si $X = (x_1, \dots, x_n)$ est défini par $x_{i_j} = d\,\mu_j$ pour $j \in [\![1\,;p]\!]$ et $x_i = 0$ pour $i \notin \{i_1, \dots, i_p\}$, alors $X \in \mathbb{N}^n$ est non nul (puisque $p \geq 1$) et tel que $MX = 0$.

COMMENTAIRES

Il faut être à l'aise sur l'invariance du rang par extension de corps.

1. en fait *et* puisque la somme des $(\alpha_i)_{1\leq i\leq p}$ est nulle
2. la dimension est la même par invariance du rang par extension de corps

24 THÉORÈME DE KRONECKER

[FGN07a, §5.33, p213] [Gou09, §2.5, p89]

ÉNONCÉ

Théorème. [Théorème de Kronecker]
Soit $P \in \mathbb{Z}[X]$ unitaire tel que ses racines sont toutes de module inférieur ou égal à 1 et tel que $P(0) \neq 0$. Alors toutes les racines de P sont des racines de l'unité.

Corollaire. [Théorème de Kronecker]
Soit $P \in \mathbb{Z}[X]$ unitaire et irréductible sur $\mathbb{Q}$. Si toutes les racines de P sont de module inférieur ou égal à 1, alors $P = X$ ou P est un polynôme cyclotomique : $P = \Phi_k$ pour un $k \in \mathbb{N}^*$.

DÉVELOPPEMENT

Procédons à la démonstration du théorème en 3 étapes.

1. Soit Ω l'ensemble des $P \in \mathbb{Z}[X]$ de degré n unitaires dont toutes les racines sont de module inférieur ou égal à 1 et tels que $P(0) \neq 0$.

   Vérifions que Ω est fini. Soit $P \in \Omega$ et soient $z_1, \dots, z_n$ ses racines comptées sans leur multiplicité. Pour $r \in [\![1\,;n]\!]$, on pose

   $$\sigma_r(z_1, \dots, z_n) = \sum_{I \in \mathcal{P}_r([\![1\,;n]\!])} \prod_{i \in I} z_i$$

   le r-ième polynôme symétrique élémentaire, de sorte que :

   $$P(X) = X^n + \sum_{r=1}^{n} (-1)^r \sigma_r(z_1, \dots, z_n) X^{n-r}\,.$$

   Notons que les $\big(\sigma_r(z_1, \dots, z_n)\big)_{1 \leq r \leq n}$ sont dans $\mathbb{Z}$. Comme $|z_i| \leq 1$ pour tout $i \in [\![1\,;n]\!]$:

   $$\forall r \in [\![1\,;n]\!] \qquad \big|\sigma_r(z_1, \dots, z_n)\big| \leq \sum_{I \in \mathcal{P}_r([\![1\,;n]\!])} \prod_{i \in I} |z_i| \leq \operatorname{card}\big(\mathcal{P}_r([\![1\,;n]\!])\big) = \binom{n}{r}\,.$$

   Autrement dit, le nombre de polynômes appartenant à Ω est nécessairement fini.
2. Pour k entier naturel non nul, définissons le polynôme

   $$P_k = \prod_{i=1}^{n} (X - z_i^k)\,.$$

   Montrons que les $(P_k)_{k \in \mathbb{N}^*}$ sont des éléments de Ω.
   Notons déjà qu'ils sont tous de degré n, unitaires, et que leurs racines sont toutes de module dans $]0, 1]$. Reste donc à vérifier qu'ils sont des polynômes de $\mathbb{Z}[X]$.

   Soit $k \in \mathbb{N}^*$. Posant $\sigma_r^{(k)} = \sigma_r\big(z_1^k, \dots, z_n^k\big)$ pour $r \in [\![1\,;n]\!]$, on décompose P_k en

   $$P_k = X^n + \sum_{r=1}^{n} (-1)^r \sigma_r^{(k)} X^{n-r}\,.$$

   Fixons $r \in [\![1\,;n]\!]$. On remarque que :

   $$\sigma_r^{(k)} = \sum_{I \in \mathcal{P}_r([\![1\,;n]\!])} \prod_{i \in I} z_i^k = Q_r(z_1, \dots, z_n)\,, \quad \text{où } Q_r(X_1, \dots, X_n) = \sum_{I \in \mathcal{P}_r([\![1\,;n]\!])} \prod_{i \in I} X_i^k$$

   est un polynôme symétrique. Par le théorème de décomposition en polynômes symétriques, il existe un polynôme T_r à coefficients dans $\mathbb{Z}$ tel que

   $$Q_r(X_1, \dots, X_n) = T_r\big(\sigma_1(X_1, \dots, X_n), \dots, \sigma_n(X_1, \dots, X_n)\big) = T_r\big(\sigma_1^{(1)}, \dots, \sigma_n^{(1)}\big)\,.$$

   Les $\big(\sigma_p^{(1)}\big)_{1 \leq p \leq n}$ étant des entiers, il en découle que $\sigma_r^{(k)} \in \mathbb{Z}$. Finalement, ceci étant valable pour tout $r \in [\![1\,;n]\!]$, on obtient que $P_k \in \mathbb{Z}[X]$.

   Ainsi, tous les $(P_k)_{k \in \mathbb{N}^*}$ sont des éléments de Ω.
3. L'ensemble Ω est fini et contient des polynômes de degré n, donc l'ensemble des racines des polynômes de Ω est également fini. Par définition de Ω, cet ensemble ne contient pas 0.

   Soit $i \in [\![1\,;n]\!]$ fixé. Les $(z_i^k)_{k \in \mathbb{N}^*}$ sont des éléments d'un ensemble fini, il existe deux entiers strictement positifs $k_1 < k_2$ tels que $z_i^{k_1} = z_i^{k_2}$, soit $z_i^{k_2 - k_1} = 1$ puisque $z_i \neq 0$.
   Autrement dit, z_i est une racine de l'unité.

Montrons alors le corollaire.

Si $P(0) = 0$, alors on obtient que $P = X$.

Sinon, en appliquant le théorème et en gardant les notations de la démonstrations, on obtient que les $(z_i)_{1 \leq i \leq \deg(P)}$ sont des racines de l'unité.
Ainsi, par exemple, il existe $k \in \mathbb{N}^*$ tel que z_1 est une racine k-ième de l'unité. Comme Φ_k est le polynôme minimal de z_1, nécessairement $\Phi_k \mid P$. Les polynômes P et Φ_k étant irréductibles et unitaires, ceci implique que $P = \Phi_k$.

COMMENTAIRES

Il faut bien maitriser le théorème de décomposition en polynômes symétriques élémentaires.

25 THÉORÈME DE SOPHIE GERMAIN

[FGN07a, §4.39, p167]

ÉNONCÉ

Théorème. [Théorème de Sophie Germain]
Soit p un nombre premier impair tel que $q = 2p+1$ est premier. Alors il n'existe pas de triplet $(x,y,z) \in \mathbb{Z}^3$ tel que $p \nmid xyz$ et $x^p + y^p + z^p = 0$.

DÉVELOPPEMENT

Supposons qu'il existe un triplet (x,y,z) solution.

Quitte à diviser par $d = \mathrm{PGCD}(x,y,z)$, on peut supposer que $\mathrm{PGCD}(x,y,z) = 1$.

1. Montrons que x, y, z sont premiers deux à deux.
   Si $\mathrm{PGCD}(x,y) > 1$, soit p_0 un diviseur premier de ce PGCD.
   Alors $z^p = -(x^p + y^p)$ est divisible par p_0, donc $p_0 \mid z$ et $\mathrm{PGCD}(x,y,z) \geq p_0$, ce qui est contradictoire. Ainsi $\mathrm{PGCD}(x,y) = 1$, et on obtient de même que
   $$\mathrm{PGCD}(x,y) = \mathrm{PGCD}(x,z) = \mathrm{PGCD}(y,z) = 1\,.$$

2. Montrons le lemme suivant :
   $$\forall m \in \mathbb{Z} \qquad q \nmid m \quad \Longrightarrow \quad m^p \equiv \pm 1 \mod q\,.$$
   En effet, soit $m \in \mathbb{Z}$ tel que $q \nmid m$. D'après le petit théorème de Fermat, on sait que
   $$(m^p)^2 = m^{2p} = m^{q-1} \equiv 1 \mod q\,, \qquad \text{et donc} \qquad m^p \equiv \pm 1 \mod q\,.$$

3. Montrons qu'un (seul) des trois éléments x, y, z est divisible par q.
   Si q ne divise ni x, ni y, ni z, alors $x^p, y^p, z^p \equiv \pm 1 \mod q$ d'après le lemme, et donc $0 = x^p + y^p + z^p$ est congru à $-3, -1, 1$ ou 3 modulo q. Puisque $q \geq 7$, ceci est absurde.
   Dans la suite, quitte à intervertir les rôles de x, y et z, on suppose que $q \mid x$.
   Comme $\mathrm{PGCD}(x,y) = \mathrm{PGCD}(x,z) = 1$, notons que $q \nmid yz$.

4. Comme p est impair, écrivons
   $$(-x)^p = y^p + z^p = y^p - (-z)^p = (y+z)\underbrace{\sum_{k=0}^{p-1} (-z)^{p-1-k} y^k}_{r}\,.$$
   Supposons que $\mathrm{PGCD}(y+z, r) > 1$, et soit p_0 un diviseur premier de ce PGCD.
   Alors $p_0^2 \mid x^p$, donc $p_0 \mid x$ puis, étant donné que $y \equiv -z \mod p_0$:
   $$0 \equiv r \equiv p\, y^{p-1} \mod p_0\,, \qquad \text{ou encore} \qquad p_0 \mid p\, y^{p-1}\,.$$
   D'après le lemme de Gauss, deux cas se présentent :
   - soit $p_0 \mid p$, et alors $p_0 = p$. Ce n'est pas possible puisque $p \nmid x$,
   - soit $p_0 \mid y$, et alors $p_0 \mid \mathrm{PGCD}(x,y) = 1$, ce qui est impossible.

   Par l'absurde, on vient de montrer que $\mathrm{PGCD}(y+z, r) = 1$. Comme $(y+z)r = x^p$, on en déduit[1] que
   $$\exists a \in \mathbb{Z} \quad y + z = a^p \qquad \text{et} \qquad \exists \alpha \in \mathbb{Z} \quad r = \alpha^p\,.$$
   Le même raisonnement assure que
   $$\exists b \in \mathbb{Z} \quad x + z = b^p \qquad \text{et} \qquad \exists c \in \mathbb{Z} \quad x + y = c^p\,.$$

5. D'une part, on remarque, en utilisant que $q \mid x$ pour la première équation ainsi que le lemme du second point pour les deux suivantes, que
   $$\begin{cases} b^p + c^p - a^p = 2x \equiv 0 \mod q \\ c^p \equiv y \equiv \pm 1 \mod q \\ b^p \equiv z \equiv \pm 1 \mod q \end{cases}$$
   D'autre part, si $q \nmid a$, alors $a^p \equiv \pm 1 \mod q$ par le lemme, si bien que $b^p + c^p - a^p$ est congru à $-3, -1, 1$ ou 3 modulo q, ce qui est à nouveau contradictoire.
   Ainsi $q \mid a$, et alors $y \equiv -z \mod q$. En raisonnant comme à l'étape 4, il en découle que
   $$\alpha^p = r \equiv p\, y^{p-1} \mod q\,.$$
   Or, comme $y \equiv \pm 1 \mod q$, il vient
   $$\alpha^p \equiv p \mod q\,.$$
   Mais α^p est congru à $-1, 0$ ou 1 d'après le lemme, ce qui est contradictoire.

Finalement, il n'existe pas de triplet satisfaisant.

1. si le produit de deux entiers premiers entre eux est une puissance k-ième pour un $k \in \mathbb{N}^*$, alors ces deux entiers sont aussi des puissances k-ième, ce que l'on peut vérifier à partir des décompositions en produit de facteurs premiers

26 THÉORÈME DE SYLOW

[Per96, §I.5, p18–19] [Szp09, §3.VIII.2, p273–274]

ÉNONCÉ

Soit G un groupe fini d'ordre $n \in \mathbb{N}^*$. Soit p un nombre premier. On note $n = p^a m$ où $p \nmid m$.

THÉORÈME. [THÉORÈME DE SYLOW 1]
Il existe au moins un p-SYLOW, i.e., un sous-groupe d'ordre p^a, dans G.

COROLLAIRE. *Il existe au moins un sous-groupe de G d'ordre p^i pour tout $i \in [\![1\,;a]\!]$.*

DÉVELOPPEMENT

Pour montrer le premier théorème de SYLOW, on va d'abord étudier l'exemple de $\mathcal{GL}_n(\mathbb{F}_p)$, auquel on se ramènera ensuite dans le cas général.

1. Le cas de $G = \mathcal{GL}_n(\mathbb{F}_p)$. Commençons par calculer son cardinal :

$$\begin{aligned}\operatorname{card}\big(\mathcal{GL}_n(\mathbb{F}_p)\big) &= \operatorname{card}\big(\text{bases de } \mathbb{F}_p\big)\\ &= \big(p^n-1\big)\times\big(p^n-p\big)\times\cdots\times\big(p^n-p^{n-1}\big)\\ &= p^{n(n-1)/2}m\,,\end{aligned}$$

où $m \in \mathbb{N}^*$ est tel que $m \wedge p = 1$. Un p-SYLOW de G a donc $p^{n(n-1)/2}$ éléments.

Regardons T l'ensemble des matrices triangulaires supérieures de coefficients diagonaux 1. On vérifie que T est un sous-groupe de G dont le cardinal est p puissance le nombre de coefficients strictement supérieurs, c'est-à-dire

$$\operatorname{card}(T) = p^{n(n-1)/2}\,.$$

Autrement dit, T est un p-SYLOW de $\mathcal{GL}_n(\mathbb{F}_p)$.

2. Soit maintenant G un groupe quelconque d'ordre n.
Montrons que G est isomorphe à un sous-groupe de $\mathcal{GL}_n(\mathbb{F}_p)$.
Par le théorème de CAYLEY[1], on sait que G est isomorphe à un sous-groupe de $\mathfrak{S}_n$.
De plus, l'application

$$\begin{array}{ccc}\mathfrak{S}_n & \longrightarrow & \mathcal{GL}_n(\mathbb{F}_p)\\ \sigma & \longmapsto & P_\sigma\,,\end{array}$$

où P_σ est la matrice de permutation associée à σ, est un morphisme injectif.
On obtient ainsi que G est isomorphe à un sous-groupe de $\mathcal{GL}_n(\mathbb{F}_p)$.

3. Pour achever la démonstration, montrons que si G est un sous-groupe d'un groupe H possédant un p-SYLOW S, alors G possède un p-SYLOW.
Faisons agir H sur H/S par translation à gauche. On a

$$\forall a \in H \qquad \operatorname{Stab}_H(aS) = aSa^{-1}\,,$$

et, en restreignant l'action à G,

$$\forall a \in H \qquad \operatorname{Stab}_G(aS) = aSa^{-1} \cap G\,.$$

Notons que les $(aSa^{-1}\cap G)_{a\in H}$ sont des p-groupes. Montrons que l'un d'entre eux est un p-SYLOW de G, c'est-à-dire qu'il existe $a \in H$ tel que $|G|/|aSa^{-1}\cap G|$ est premier avec p.
Supposons que ce n'est pas le cas, *i.e.*, que p divise $|G|/|\operatorname{Stab}_G(aS)|$ pour tout $a \in H$.
D'après l'équation aux classes

$$|H/S| = \sum_{O\in\mathcal{O}_G}|O| = \sum_{O\in\mathcal{O}_G}\frac{|G|}{|\operatorname{Stab}_G(a_O S)|}$$

où, pour toute orbite $O \in \mathcal{O}_G$, a_O désigne un élément de H tel que $a_O S \in O$.
Alors $|H/S|$ est divisible par p, ce qui est impossible puisque S est un p-SYLOW de H.
Ainsi, le sous-groupe G admet un p-SYLOW.

Montrons désormais le corollaire. Par le théorème, on peut supposer que G est un p-groupe d'ordre p^a. On procède par récurrence sur $a \in \mathbb{N}$:

- Si $a = 0$ ou $a = 1$, il n'y a rien à montrer.
- Si $a \geq 2$, supposons le résultat vrai pour $a-1$.
  Le centre $Z(G)$ est un p-groupe non trivial de G. Choisissons $x \neq e \in Z(G)$. L'ordre de x divisant l'ordre de $Z(G)$, il existe $b \in [\![1\,;a]\!]$ tel que $o(x) = p^b$. En posant $y = x^{p^{b-1}}$, on a $o(y) = p$ et on peut considérer le sous-groupe $H = G/\langle y\rangle$ qui est d'ordre p^{a-1}.
  Soit $i \in [\![1\,;a]\!]$. Par hypothèse de récurrence, H admet un sous-groupe H_{i-1} d'ordre p^{i-1}.
  Posant $G_i = \pi^{-1}(H_{i-1})$ où $\pi : G \longrightarrow G/H$ est la surjection canonique, il vient

$$G_i/\ker(\pi) \simeq H_{i-1}\,, \qquad \text{donc} \qquad |G_i| = p\,|H_{i-1}| = p^i\,.$$

  Ainsi, on a trouvé un sous-groupe de G d'ordre p^i pour $i \in [\![1\,;a]\!]$. Bien sûr, G possède un groupe d'ordre p^0. L'hypothèse de récurrence est donc vraie au rang a.

On conclut par principe de récurrence.

1. on utilise l'action par translation à gauche pour le montrer

27 THÉORÈME DES DEUX CARRÉS

[Per96, §II.6, p56–58]

ÉNONCÉ

On note $\mathbb{Z}[\mathrm{i}] = \{a + \mathrm{i}b : a, b \in \mathbb{Z}\}$ l'ensemble des entiers de Gauss, muni de la norme N définie par $N(a + \mathrm{i}b) = a^2 + b^2$, et on définit

$$\Sigma = \left\{ n \in \mathbb{N} : \exists (a, b) \in \mathbb{N}^2, n = a^2 + b^2 \right\}.$$

Lemme. *Σ est stable par produit.*

Lemme. *Si $p \in \mathcal{P}$, alors $p \in \Sigma \iff p \equiv 1, 2 \mod 4$.*

Théorème. [théorème des deux carrés de Fermat]
$n \in \Sigma$ si et seulement si pour tout $p \in \mathcal{P}$ tel que $p \mid n$ et $p \equiv 3 \mod 4$, alors $2 \mid v_p(n)$.

DÉVELOPPEMENT

- Commençons par le premier lemme.
Remarquons d'abord que

$$n \in \Sigma \qquad \iff \qquad \exists z \in \mathbb{Z}[\mathrm{i}] \quad n = N(z).$$

Soient donc $n, n' \in \Sigma$. Choisissons $z = a + \mathrm{i}b, z' = c + \mathrm{i}d \in \mathbb{Z}[\mathrm{i}]$ tels que $n = N(z)$ et $n' = N(z')$. Alors $nn' = N(zz') \in \Sigma$ puisque $zz' \in \mathbb{Z}[\mathrm{i}]$. Ainsi, Σ est stable par produit.
Pour la suite, notons que l'on a obtenu que :

$$\forall (a, b, c, d) \in \mathbb{Z}^4 \qquad (a^2 + b^2)(c^2 + d^2) = (ac - bd)^2 + (ad + bc)^2.$$

- Ce premier lemme invite à s'intéresser aux nombres premiers appartenant à Σ, ce qui est l'objet du second. On sait déjà que $2 = 1^2 + 1^2 \in \Sigma$. Considérons alors p premier impair.

S'il existe $a, b \in \mathbb{Z}$ tels que $p = a^2 + b^2$, alors a ou b (exclusivement) est impair égal à $2k+1$ pour un $k \in \mathbb{Z}$, et alors

$$p \equiv (2k + 1)^2 \equiv 1 \mod 4.$$

Réciproquement, notons que $p \in \Sigma$ si et seulement si p est non irréductible dans $\mathbb{Z}[\mathrm{i}]$.
En effet, d'une part si $p = a^2 + b^2 \in \Sigma$, alors $a \neq 0, b \neq 0$ et $p = (a + \mathrm{i}b)(a - \mathrm{i}b)$ avec $a \pm \mathrm{i}b \notin \mathbb{Z}[\mathrm{i}]^\times$. D'autre part, si maintenant $p = zz'$ avec $z, z' \in \mathbb{Z}[\mathrm{i}]$ non inversibles, alors nécessairement $N(z) = N(z') = p$, donc $p \in \Sigma$.

Soit $p \equiv 1 \mod 4$. On souhaite donc montrer que p est non irréductible dans l'anneau principal $\mathbb{Z}[\mathrm{i}]$, *i.e.*, que $(p) = p\,\mathbb{Z}[\mathrm{i}]$ est non premier ou encore que $\mathbb{Z}[\mathrm{i}]/(p)$ est non intègre. Or, on sait que $\mathbb{Z}[\mathrm{i}] \simeq \mathbb{Z}[X]/(X^2 + 1)$. Il s'ensuit que

$$\mathbb{Z}[\mathrm{i}]/(p) \simeq \mathbb{Z}[X]/(X^2 + 1, p) \simeq \big(\mathbb{Z}[X]/(p)\big)/(X^2 + 1) \simeq \mathbb{F}_p[X]/(X^2 + 1).$$

Ainsi, dire que (p) est non premier, c'est dire que $X^2 + 1$ n'est pas irréductible sur $\mathbb{F}_p$, *i.e.*, que $X^2 + 1$ admet une racine dans $\mathbb{F}_p$, ou encore que $-1 \in \left(\mathbb{F}_p^*\right)^2$, ce qui équivaut à $(-1)^{\frac{p-1}{2}} = 1$, soit $p \equiv 1 \mod 4$. Ceci conclut la preuve du lemme[1].

- Venons-en alors au théorème.
Écrivons la décomposition de n en facteurs premiers :

$$n = \prod_{p \in \mathcal{P}} p^{v_p(n)}.$$

Si $v_p(n)$ est pair pour tout $p \in \mathcal{P}$ tel $p \equiv 3 \mod 4$, alors $n \in \Sigma$ en utilisant les deux lemmes et le fait que tout carré est dans $\mathcal{P}$.

Réciproquement, soit $p \equiv 3 \mod 4$. On sait par ce qui précède que p est irréductible dans $\mathbb{Z}[\mathrm{i}]$. Ainsi, si p divise $n = a^2 + b^2 = (a + \mathrm{i}b)(a - \mathrm{i}b)$, alors p divise soit $a + \mathrm{i}b$ soit $a - \mathrm{i}b$, ce qui implique que $p \mid a$ et $p \mid b$, et donc que $p^2 \mid n$. On procède de même avec $n' = n/p^2$ tant que $p \mid n'$, et on obtient que $v_p(n)$ est pair.

COMMENTAIRES

Lors de la démonstration du second lemme, on a utilisé le résultat suivant, qu'il peut être intéressant d'écrire en fin de développement s'il reste un peu de place au tableau :

Proposition. *Soit A un anneau commutatif et I, J des idéaux de A. Alors*

$$(A/I)/\pi_I(J) \simeq A/(I + J) \simeq (A/J)/\pi_J(I),$$

où π_I et π_J sont les projections de A dans A/I et A/J.

La preuve de ce résultat consiste à montrer que $p \circ \pi_I$, où $p : A/I \longrightarrow (A/I)/\pi_I(J)$, est une application surjective de noyau $I + J$.

1. notons que le raisonnement réciproque justifie en fait l'équivalence

28 THÉORÈME DES EXTREMA LIÉS

[Gou08, §5.2/3, p317/327]

ÉNONCÉ

THÉORÈME. [THÉORÈME DES EXTREMA LIÉS]
Soient $r, n \in \mathbb{N}^*$ et $f, g_1, \dots, g_r : U \longrightarrow \mathbb{R}$ des fonctions de classe $\mathcal{C}^1$ définies sur un ouvert U de $\mathbb{R}^n$. Notons $\Gamma = \{x \in U : g_1(x) = \cdots = g_r(x) = 0\}$. Si $f_{|\Gamma}$ admet un extremum local en $a \in \Gamma$ et si la famille de formes linéaires $\big(dg_1(a), \dots, dg_r(a)\big)$ est libre, alors

$$\exists!(\lambda_i)_{1\leq i\leq r} \in \mathbb{R}^r \qquad df(a) = \sum_{i=1}^{r} \lambda_i \, dg_i(a)\,.$$

Les réels $(\lambda_i)_{1\leq i\leq r}$ sont appelés multiplicateurs de LAGRANGE.

DÉVELOPPEMENT

Soit $a \in \Gamma$ un extremum local de $f_{|\Gamma}$ tel que $\big(dg_1(a), \dots, dg_r(a)\big)$ soit libre.

Notons déjà que, nécessairement, $r \leq n$, et qu'en cas d'égalité, la famille est une base du dual de $\mathbb{R}^n$ et donc le résultat est évident.

Supposons donc $r < n$ et posons $s = n - r \geq 1$. En identifiant $\mathbb{R}^n$ et $\mathbb{R}^s \times \mathbb{R}^r$, on écrira ses éléments sous la forme $(x, y) = (x_1, \dots, x_s, y_1, \dots, y_r)$. On note $a = (\alpha, \beta)$. La matrice

$$M_a = \begin{pmatrix} \partial_{x_1} g_1(a) & \cdots & \partial_{x_s} g_1(a) & \partial_{y_1} g_1(a) & \cdots & \partial_{y_r} g_1(a) \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \partial_{x_1} g_r(a) & \cdots & \partial_{x_s} g_r(a) & \partial_{y_1} g_r(a) & \cdots & \partial_{y_r} g_r(a) \end{pmatrix}$$

est de rang r par hypothèse de liberté de $\big(dg_1(a), \dots, dg_r(a)\big)$. On peut en extraire une matrice carrée de taille r inversible, qui, quitte à renommer les variables, est la matrice

$$M_{a,2} = \begin{pmatrix} \partial_{y_1} g_1(a) & \cdots & \partial_{y_r} g_1(a) \\ \vdots & \ddots & \vdots \\ \partial_{y_1} g_r(a) & \cdots & \partial_{y_r} g_r(a) \end{pmatrix}.$$

On peut appliquer le théorème des fonctions implicites à $g = (g_1, \dots, g_r)$ en a puisque, par ce qui précède, g est $\mathcal{C}^1$, $g(a) = g\big((\alpha, \beta)\big) = 0$ et $\partial_2 g(\alpha, \beta)$ est inversible. Il existe donc φ de classe $\mathcal{C}^1$ telle que $g\big((x, y)\big) = 0$ au voisinage de $a = (\alpha, \beta)$ si et seulement si $y = \varphi(x)$, c'est-à-dire localement $(x, y) \in \Gamma \iff y = \varphi(x)$.

Posons $h : x \longmapsto f(x, \varphi(x))$ au voisinage de α. L'application h admet un extremum local en α puisque $\big(\alpha, \varphi(\alpha)\big) = a$ et $\big(x, \varphi(x)\big) \in \Gamma$ au voisinage de α. Si $u : x \longmapsto \big(x, \varphi(x)\big)$, alors h est différentiable en α par composition et $0 = dh(\alpha) = df\big(u(\alpha)\big) \circ du(\alpha)$, soit matriciellement :

$$0 = \begin{pmatrix} \partial_{x_1} f(a) & \cdots & \partial_{x_s} f(a) & \partial_{y_1} f(a) & \cdots & \partial_{y_r} f(a) \end{pmatrix} \begin{pmatrix} 1 & & & \\ & \ddots & & \\ & & 1 & \\ \partial_{x_1}\varphi_1(\alpha) & \cdots & \partial_{x_s}\varphi_1(\alpha) \\ \vdots & \ddots & \vdots \\ \partial_{x_1}\varphi_r(\alpha) & \cdots & \partial_{x_s}\varphi_r(\alpha) \end{pmatrix}$$

$$= \begin{pmatrix} \partial_{x_1} f(a) + \sum_{j=1}^{r} \partial_{x_1}\varphi_j(\alpha)\, \partial_{y_j} f(a) \\ \vdots \\ \partial_{x_s} f(a) + \sum_{j=1}^{r} \partial_{x_s}\varphi_j(\alpha)\, \partial_{y_j} f(a) \end{pmatrix}.$$

Ainsi, on obtient que $0 = \partial_{x_i} h(\alpha) = \partial_{x_i} f(a) + \sum_{j=1}^{r} \partial_{x_i}\varphi_j(\alpha)\, \partial_{y_j} f(a)$ pour tout $i \in [\![1\,;s]\!]$.

Par ailleurs, comme $g_k\big(x, \varphi(x)\big) = 0$ pour tout $k \in [\![1\,;r]\!]$, on a de même :

$$\forall k \in [\![1\,;r]\!] \quad \forall i \in [\![1\,;s]\!] \qquad 0 = \partial_{x_i} g_k(a) + \sum_{j=1}^{r} \partial_{x_i}\varphi_j(\alpha)\, \partial_{y_j} g_k(a)\,.$$

Considérons alors la matrice

$$M = \begin{pmatrix} \partial_{x_1} f(a) & \cdots & \partial_{x_s} f(a) & \partial_{y_1} f(a) & \cdots & \partial_{y_r} f(a) \\ \partial_{x_1} g_1(a) & \cdots & \partial_{x_s} g_1(a) & \partial_{y_1} g_1(a) & \cdots & \partial_{y_r} g_1(a) \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \partial_{x_1} g_r(a) & \cdots & \partial_{x_s} g_r(a) & \partial_{y_1} g_r(a) & \cdots & \partial_{y_r} g_r(a) \end{pmatrix},$$

dont on note les colonnes $(C_k)_{1\leq k\leq n}$ et les lignes $(L_i)_{0\leq i\leq r}$. Par ce qui précède, les s premières colonnes de M sont des combinaisons linéaires des r dernières ($C_k = \sum_{j=1}^{r} \partial_{x_k}\varphi_j(\alpha)\, C_{s+j}$ pour $k \in [\![1\,;s]\!]$), donc M est de rang au plus r.

Les lignes de M sont alors liées. Comme par hypothèse les r dernières lignes sont libres, la première est combinaison linéaire des autres et

$$\exists(\lambda_i)_{1\leq i\leq r} \in \mathbb{R}^r \qquad L_0 = \sum_{i=1}^{r} \lambda_i L_i \quad \text{ou encore} \quad df(a) = \sum_{i=1}^{r} \lambda_i \, dg_i(a)\,,$$

ce qui conclut l'existence des multiplicateurs de LAGRANGE. Par ailleurs, leur unicité est claire puisque $\big(dg_1(a), \dots, dg_r(a)\big)$ est une famille libre.

COMMENTAIRES

Il faut faire un joli dessin pour expliquer l'intuition, en prenant par exemple pour Γ la sphère unité et pour f une application linéaire.

BIBLIOGRAPHIE MATHÉMATIQUES GÉNÉRALES

[AK02] G. Allaire et S.-M. Kaber : *Algèbre linéaire numérique*. Ellipses, 2002.

[CG13] P. Caldero et J. Germoni : *Histoires hédonistes de groupes et de géométries - Tome 1*. Calvage et Mounet, 2013.

[Cog00] M. Cognet : *Algèbre linéaire*. Bréal, 2000.

[Col11] P. Colmez : *Éléments d'analyse et d'algèbre*. Les éditions de l'École Polytechnique, 2 ème édition, 2011.

[FGN07a] S. Francinou, H. Gianella et S. Nicolas : *Oraux X-ENS - Algèbre 1*. Cassini, 2007.

[FGN07b] S. Francinou, H. Gianella et S. Nicolas : *Oraux X-ENS - Analyse 1*. Cassini, 2007.

[FGN12] S. Francinou, H. Gianella et S. Nicolas : *Oraux X-ENS - Analyse 4*. Cassini, 2012.

[Gou08] X. Gourdon : *Les maths en tête - Analyse*. Ellipses, 2ème édition, 2008.

[Gou09] X. Gourdon : *Les maths en tête - Algèbre*. Ellipses, 2ème édition, 2009.

[MM16] R. Mansuy et R. Mneimné : *Algèbre linéaire : Réduction des endomorphismes*. De Boeck, 2 ème édition, 2016.

[Per96] D. Perrin : *Cours d'algèbre*. Ellipses, 1996.

[Pey04] G. Peyré : *L'algèbre discrète de la transformée de* Fourier. Ellipses, 2004.

[Rom17] J.-E. Rombaldi : *Mathématiques pour l'agrégation : Algèbre et géométrie*. De Boeck, 2017.

[Rou99] F. Rouvière : *Petit guide de calcul différentiel*. Cassini, 1999.

[Szp09] A. Szpriglas : *Algèbre L3*. Pearson Education, 2ème édition, 2009.

Deuxième partie

Développements d’Analyse et de Probabilités

COUPLAGES

#	Développement	Leçons
29	Calcul d'une intégrale par le théorème des résidus	236, 245
30	Complétude de $L^p(E, \mathcal{A}, \mu)$	201, 205, 234
31	Connexité des valeurs d'adhérence d'une suite et lemme de la grenouille	203, 204
32	Densité des polynômes orthogonaux	202, 213, 250
33	Équation de BURGERS	222
34	Équation de la chaleur périodique	222, 235
35	Espérance conditionnelle	234, 260
36	Étude de deux suites récurrentes	223, 224
37	Factorisations LU et de CHOLESKY	233
38	Formule d'EULER-MACLAURIN et application à la série harmonique	224, 230
39	Inégalité de HOEFFDING	262
40	Injectivité de la fonction caractéristique et application	250, 265
41	Intégrale de DIRICHLET	228, 235, 236, 239
42	Lemme de MORSE	214, 215
43	Méthode de NEWTON	223, 226
44	Méthode du gradient à pas optimal	219, 229, 233, 253
45	Processus de branchement de GALTON-WATSON	226, 229, 253, 264
46	Projection sur un convexe fermé et théorème de RIESZ-FRÉCHET	208, 213
47	Prolongement holomorphe de Γ	207, 239, 245, 265
48	Stabilité de LIAPOUNOV	220, 221
49	Théorème central limite et intervalle de confiance	261, 262
50	Théorème de BANACH-STEINHAUS et série de FOURIER divergente	205, 208, 246
51	Théorème de BERNSTEIN	243
52	Théorème de CAUCHY-LIPSCHITZ linéaire	221
53	Théorème de CAUCHY-LIPSCHITZ (globalement lipschitzien)	220
54	Théorème de FEJÉR	209, 241, 246
55	Théorème de SARD	204
56	Théorème de STONE-WEIERSTRASS	201, 202, 203
57	Théorème de WEIERSTRASS	209, 228, 241, 260 , 261, 264
58	Théorème des extrema liés	214, 215, 219
59	Théorèmes d'ABEL et taubérien faible	207, 230, 243

29 CALCUL D'UNE INTÉGRALE PAR LE THÉORÈME DES RÉSIDUS

[Tau06, §15.2, p192]

ÉNONCÉ

Exemple. Soit $\alpha \in]-1,1[$. Alors $I_\alpha = \displaystyle\int_0^{+\infty} \frac{x^\alpha \ln(x)}{x^2-1}\,\mathrm{d}x = \frac{\pi^2}{4\cos^2\left(\frac{\pi}{2}\alpha\right)}$.

DÉVELOPPEMENT

Fixons $\alpha \in]-1,1[$.

- Vérifions d'abord que I_α est bien définie.
  La fonction intégrande est, d'une part, intégrable en 0 puisque
  $$\frac{x^\alpha \ln(x)}{x^2-1} \underset{x\to 0}{\sim} x^\alpha \ln(x) \underset{x\to 0}{=} o\Big(\frac{1}{x^\varepsilon}\Big), \qquad \text{où } \varepsilon \in]-\alpha\,;1[\,,$$
  et, d'autre part, intégrable en $+\infty$ puisque
  $$\frac{x^\alpha \ln(x)}{x^2-1} \underset{x\to+\infty}{\sim} x^{\alpha-2}\ln(x) \underset{x\to+\infty}{=} o\Big(\frac{1}{x^{2-(\alpha+\varepsilon)}}\Big), \qquad \text{où } \varepsilon \in]0\,;1-\alpha[\,.$$
  Enfin, comme $\ln(x) = \ln(1+x-1) \sim_{x\to 1} x-1$, l'intégrande est continue en 1.
- Considérons Log une détermination du logarithme complexe sur $U = \mathbb{C}\setminus \mathrm{i}\mathbb{R}_-$ telle que $\mathrm{Log}(1)=0$. Pour $z\in U$, définissons
  $$z^\alpha = \exp\big(\alpha\,\mathrm{Log}(z)\big) \qquad \text{puis} \qquad f(z) = \frac{z^\alpha\,\mathrm{Log}(z)}{z^2-1}\,.$$
  La fonction f est méromorphe et admet deux pôles, -1 et 1, simples. D'après le théorème des résidus, l'intégrale de f sur le chemin γ ci-dessous, où $\varepsilon \in]0\,;\frac{1}{2}[$ et $R>2$ est nulle.

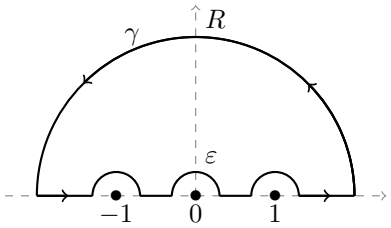

Figure 29.1 – Le chemin γ

Dans la suite, pour $a\in\mathbb{C}$ et $r\in\mathbb{R}_+^*$, on pose $\gamma_{a,r} : t\in[0\,;\pi] \longmapsto a + r\,\mathrm{e}^{\mathrm{i}t}$.

- Soit $r\in\mathbb{R}_+^*$. Pour $t\in[0\,;\pi]$, on a les inégalités $\big|\mathrm{Log}(r\,\mathrm{e}^{\mathrm{i}t})\big| = \big|\ln(r)+\mathrm{i}t\big| \le \big|\ln(r)\big| + \pi$ et, d'après la seconde inégalité triangulaire, $\big|r^2\,\mathrm{e}^{2\mathrm{i}t}-1\big| \ge \big||r^2\,\mathrm{e}^{2\mathrm{i}t}|-|-1|\big| = |r^2-1|$, d'où
  $$\begin{aligned}\left|\int_{\gamma_{0,r}} f(z)\,\mathrm{d}z\right| &= \left|\int_0^\pi \mathrm{i}r\,\mathrm{e}^{\mathrm{i}t}\,f\big(r\,\mathrm{e}^{\mathrm{i}t}\big)\,\mathrm{d}t\right| \le \int_0^\pi r\frac{r^\alpha\big|\mathrm{Log}(r\,\mathrm{e}^{\mathrm{i}t})\big|}{\big|r^2\,\mathrm{e}^{2\mathrm{i}t}-1\big|}\,\mathrm{d}t \\ &\le \int_0^\pi \frac{r^{\alpha+1}}{|r^2-1|}\big(|\ln(r)|+\pi\big)\,\mathrm{d}t = \pi\frac{r^{\alpha+1}}{|r^2-1|}\big(|\ln(r)|+\pi\big)\,.\end{aligned}$$
  En faisant tendre r vers 0 et $+\infty$, on obtient que
  $$\lim_{\varepsilon\to 0}\int_{\gamma_{0,\varepsilon}} f(z)\,\mathrm{d}z = 0 \qquad \text{et} \qquad \lim_{R\to+\infty}\int_{\gamma_{0,R}} f(z)\,\mathrm{d}z = 0\,.$$
- Remarquons que f est en fait holomorphe en 1, donc continue, puisque
  $$\mathrm{Res}(f,1) = \lim_{z\to 1}\frac{z^\alpha\,\mathrm{Log}(z)}{z+1} = \frac{1^\alpha\,\mathrm{Log}(1)}{2} = 0\,, \qquad \text{d'où} \qquad \lim_{\varepsilon\to 0}\int_{\gamma_{1,\varepsilon}} f(z)\,\mathrm{d}z = 0\,.$$
- Calculons
  $$\mathrm{Res}(f,-1) = \lim_{z\to -1}\frac{z^\alpha\,\mathrm{Log}(z)}{z+1} = \frac{(-1)^\alpha\times\mathrm{i}\pi}{-2} = -\frac{\mathrm{i}\pi}{2}\,\mathrm{e}^{\mathrm{i}\pi\alpha}\,.$$
  Étant donné que l'application $z\longmapsto f(z) - \frac{\mathrm{Res}(f,-1)}{z+1}$ est holomorphe en -1, on obtient
  $$\lim_{\varepsilon\to 0}\int_{\gamma_{-1,\varepsilon}} f(z)\,\mathrm{d}z = \lim_{\varepsilon\to 0}\int_{\gamma_{-1,\varepsilon}} \frac{\mathrm{Res}(f,-1)}{z+1}\,\mathrm{d}z = \mathrm{i}\pi\,\mathrm{Res}(f,-1) = \frac{\pi^2}{2}\,\mathrm{e}^{\mathrm{i}\pi\alpha}\,.$$
- En appliquant le théorème des résidus, et en faisant tendre $\varepsilon\to 0$ et $R\to+\infty$, on obtient :
  $$\int_{-\infty}^{+\infty}\frac{x^\alpha\,\mathrm{Log}(x)}{x^2-1}\,\mathrm{d}x - \frac{\pi^2}{2}\,\mathrm{e}^{\mathrm{i}\pi\alpha} = 0\,.$$
  Comme
  $$\begin{aligned}\int_{-\infty}^0 \frac{x^\alpha\,\mathrm{Log}(x)}{x^2-1}\,\mathrm{d}x &= (-1)^\alpha\int_{-\infty}^0\frac{|x|^\alpha\ln(|x|)}{x^2-1}\,\mathrm{d}x + (-1)^\alpha\mathrm{i}\pi\int_{-\infty}^0\frac{|x|^\alpha}{x^2-1}\,\mathrm{d}x \\ &= \mathrm{e}^{\mathrm{i}\pi\alpha}\,I_\alpha + \mathrm{e}^{\mathrm{i}\pi\alpha}\,\mathrm{i}\pi\int_0^{+\infty}\frac{x^\alpha}{x^2-1}\,\mathrm{d}x\,,\end{aligned}$$
  il vient
  $$\frac{1+\mathrm{e}^{\mathrm{i}\pi\alpha}}{\mathrm{e}^{\mathrm{i}\pi\alpha}}I_\alpha + \mathrm{i}\pi\int_0^{+\infty}\frac{x^\alpha}{x^2-1}\,\mathrm{d}x = \frac{\pi^2}{2}\,.$$
  En prenant la partie réelle dans cette dernière équation, il en découle que
  $$I_\alpha = \frac{\pi^2}{2}\frac{1}{\Re(\mathrm{e}^{-\mathrm{i}\pi\alpha}+1)}\,, \qquad \text{soit finalement} \qquad I_\alpha = \frac{\pi^2}{4}\frac{1}{\cos^2\left(\frac{\pi}{2}\alpha\right)}\,,$$
  puisque $\mathrm{e}^{-\mathrm{i}\pi\alpha}+1 = \mathrm{e}^{-\mathrm{i}\frac{\pi}{2}\alpha}\,2\cos\left(\frac{\pi}{2}\alpha\right)$ et donc $\Re(\mathrm{e}^{-\mathrm{i}\pi\alpha}+1) = 2\cos^2\left(\frac{\pi}{2}\alpha\right)$.

30 COMPLÉTUDE DE $L^p(E,\mathcal{A},\mu)$

[Bre99, §IV.2, p57–58]

ÉNONCÉ

THÉORÈME. [THÉORÈME DE RIESZ-FICHER]
Soit $(E,\mathcal{A},\mu)$ un espace mesuré. Pour tout $p\in[1,+\infty]$, l'espace $\left(L^p(E,\mathcal{A},\mu),\|\cdot\|_p\right)$ est un espace de BANACH.

COROLLAIRE. *Toute suite convergente dans L^p admet une sous-suite qui converge μ-p.p..*

DÉVELOPPEMENT

On admet que l'espace $\left(L^p(E,\mathcal{A},\mu),\|\cdot\|_p\right)$ est un espace vectoriel normé.

Pour la complétude, on distingue les cas p fini et infini :

- Cas $p=+\infty$. Soit $(f_n)_{n\in\mathbb{N}}$ une suite de CAUCHY dans L^∞ :

$$\forall\varepsilon>0\quad\exists N_\varepsilon\in\mathbb{N}\quad\forall m,n\geq N_\varepsilon\qquad\|f_n-f_m\|_\infty\leq\varepsilon\,.$$

Pour $n,m\in\mathbb{N}$, définissons

$$A_{n,m}=\left\{x\in E:\left|f_n(x)-f_m(x)\right|>\|f_n-f_m\|_\infty\right\}$$
$$\text{et}\qquad A_n=\left\{x\in E:\left|f_n(x)\right|>\|f_n\|_\infty\right\}$$
$$\text{puis posons}\qquad A=\bigcup_{(n,m)\in\mathbb{N}^2}A_{n,m}\bigcup\bigcup_{n\in\mathbb{N}}A_n\,.$$

L'ensemble A est un mesurable de $\mathcal{A}$ qui est de mesure nulle. De plus, pour tout $x\in A^{\mathsf{c}}$, la suite $\left(f_n(x)\right)_{n\in\mathbb{N}}$ est une suite de CAUCHY de $\mathbb{K}$ complet ($\mathbb{R}$ ou $\mathbb{C}$) donc converge vers une limite notée $f(x)\in\mathbb{K}$. Posant $f(x)=0$ pour $x\in A$, on définit alors une application $f:E\longrightarrow\mathbb{K}$ mesurable (limite simple de fonctions mesurables).
Vérifions que $f\in L^\infty$. Soit $x\in A^{\mathsf{c}}$. On sait que

$$\forall m,n\geq N_1\qquad\left|f_n(x)-f_m(x)\right|\leq 1\,.$$

Ainsi, en prenant $n=N_1$ et en faisant tendre $m\to+\infty$, il vient

$$\left|f(x)\right|\leq\left|f(x)-f_{N_1}(x)\right|+\left|f_{N_1}(x)\right|\leq 1+\|f_{N_1}\|_\infty\,.$$

Comme $\mu(A)=0$, on a obtenu que $|f|\leq 1+\|f_{N_1}\|_\infty$ μ-p.p. et ainsi $f\in L^\infty$.
Enfin, $\lim_{n\to+\infty}\|f_n-f\|_\infty=0$, puisque, par définition de A :

$$\forall\varepsilon>0\quad\exists N_\varepsilon\in\mathbb{N}\quad\forall n\geq N_\varepsilon\quad\forall x\in A^{\mathsf{c}}\qquad\left|f_n(x)-f(x)\right|\leq\varepsilon\,.$$

Ainsi, la suite $(f_n)_{n\in\mathbb{N}}$ admet une limite f, et la convergence a lieu μ-p.p..

- Cas $p<+\infty$. Soit $(f_n)_{n\in\mathbb{N}}$ une suite de CAUCHY dans L^p.
Montrons que $(f_n)_{n\in\mathbb{N}}$ possède une valeur d'adhérence.
Pour $k\in\mathbb{N}$, on pose $\varepsilon_k=2^{-k}$ puis on choisit $N_k\in\mathbb{N}$ tel que

$$\forall m,n\geq N_k\qquad\|f_n-f_m\|_p\leq\varepsilon_k\,.$$

Définissons une suite d'entiers $(n_k)_{k\in\mathbb{N}}$ par récurrence :

$$n_0=N_0\qquad\text{puis}\qquad\forall k\in\mathbb{N}^*\quad n_k=\max(N_k,n_{k-1}+1)\,,$$

puis posons $\tilde{f}_k=f_{n_k}$ pour $k\in\mathbb{N}$. La suite $(\tilde{f}_k)_{k\in\mathbb{N}}$ est une suite extraite de $(f_n)_{n\in\mathbb{N}}$ satisfaisant :

$$\sum_{k=0}^{+\infty}\|\tilde{f}_{k+1}-\tilde{f}_k\|_p\leq\sum_{k=0}^{+\infty}\varepsilon_k\leq 2<+\infty\,.$$

Considérons alors la suite d'applications $(g_\ell)_{\ell\in\mathbb{N}}$ définie par

$$\forall\ell\in\mathbb{N}\qquad g_\ell=\sum_{k=0}^{\ell}\left|\tilde{f}_{k+1}-\tilde{f}_k\right|\,.$$

La suite $(g_\ell)_{\ell\in\mathbb{N}}$ est croissante μ-p.p., on peut donc définir sa limite g. Comme $\|g_\ell\|_p\leq 2$ pour tout $\ell\in\mathbb{N}$, le théorème de convergence monotone assure que

$$\|g\|_p\leq 2\,.$$

En particulier g est finie μ-p.p. : il existe $A\in\mathcal{A}$ de mesure nulle tel que g est finie sur A^{c}.
Soit $x\in A^{\mathsf{c}}$. Pour $i,j\in\mathbb{N}$ tels que $i<j$, on a :

$$\left|\tilde{f}_j(x)-\tilde{f}_i(x)\right|\leq\sum_{k=i}^{j-1}\left|\tilde{f}_{k+1}(x)-\tilde{f}_k(x)\right|\leq g(x)-g_{i-1}(x)\quad\underset{i\to+\infty}{\longrightarrow}\quad 0\,.$$

Ainsi la suite $\left(\tilde{f}_k(x)\right)_{k\in\mathbb{N}}$ est de CAUCHY. Notons $f(x)$ sa limite.
En posant $f(x)=0$ pour $x\in A$, on définit une fonction f mesurable.
Soit $i\in\mathbb{N}$. Faisant tendre $j\to+\infty$ ci-dessus, $\left|f-\tilde{f}_i\right|\leq g$ sur A^{c}, soit $|f|\leq g+\left|\tilde{f}_i\right|$ μ-p.p..
Autrement dit, $f\in L^p$, et par convergence dominée on obtient

$$\lim_{k\to+\infty}\|f-\tilde{f}_k\|_p=0\,.$$

Ainsi $(f_n)_{n\in\mathbb{N}}$ est une suite de CAUCHY admettant une valeur d'adhérence $f\in L^p$.
La preuve suivie assure que la sous-suite $\left(\tilde{f}_k\right)_{k\in\mathbb{N}}$ converge vers f μ-p.p..

31 CONNEXITÉ DES VALEURS D'ADHÉRENCE D'UNE SUITE ET LEMME DE LA GRENOUILLE

[Gou08, §1.4, p46] [FGN07, §2.19, p86]

ÉNONCÉ

PROPOSITION. *Soient (E,d) un espace métrique compact et $(u_n)_{n\in\mathbb{N}} \in E^{\mathbb{N}}$ telle que*

$$d(u_n, u_{n+1}) \underset{n\to+\infty\to}{\longrightarrow} 0\,.$$

Alors l'ensemble Γ des valeurs d'adhérence de $(u_n)_{n\in\mathbb{N}}$ est connexe.

APPLICATION. [LEMME DE LA GRENOUILLE]
Soient $f : [0,1] \longrightarrow [0,1]$ une fonction continue et $(x_n)_{n\in\mathbb{N}} \in [0,1]^{\mathbb{N}}$ une suite définie par

$$x_0 \in [0,1] \qquad \text{et} \qquad \forall n \in \mathbb{N} \quad x_{n+1} = f(x_n)\,.$$

Alors $(x_n)_{n\in\mathbb{N}}$ converge si et seulement si $\lim_{n\to+\infty} x_{n+1} - x_n = 0$.

DÉVELOPPEMENT

Commençons par la proposition en raisonnant par l'absurde.
Supposons que l'ensemble fermé $\Gamma = \cap_{p\in\mathbb{N}} \overline{\{u_n : n \geq p\}}$ est non connexe. Il existe alors A, B deux fermés non vides disjoints de Γ tels que $\Gamma = A \sqcup B$.
Les ensembles A et B sont compacts en tant que fermés dans des compacts. Puisqu'ils sont disjoints, cela implique que $\alpha = d(A,B) > 0$. Considérons alors[1] :

$$A' = \left\{x \in E : d(x,A) < \frac{\alpha}{3}\right\} = A + \mathbb{B}\left(0, \frac{\alpha}{3}\right) \qquad \text{et} \qquad B' = \left\{x \in E : d(x,B) < \frac{\alpha}{3}\right\}.$$

Étant donné que A' et B' sont ouverts, l'ensemble $K = \left(A' \cup B'\right)^{\mathrm{c}}$ est fermé, donc compact.

Construisons une sous-suite de $(u_n)_{n\in\mathbb{N}}$ à valeurs dans K.
Fixons $N_0 \in \mathbb{N}$ tel que $d(u_n, u_{n+1}) < \frac{\alpha}{3}$ pour tout $n \geq N_0$.
Prenons $a \in A$ et $b \in B$. Comme a et b sont valeurs d'adhérence de $(u_n)_{n\in\mathbb{N}}$, les ensembles

$$V_a = \left\{n \geq N_0 : d(u_n, a) < \frac{\alpha}{3}\right\} \qquad \text{et} \qquad V_b = \left\{n \geq N_0 : d(u_n, b) < \frac{\alpha}{3}\right\}$$

sont infinis.

- Soit $N_a \in V_a$. Comme V_b est infini, on peut choisir $N_b \in V_b$ tel que $N_b > N_a$. Alors

$$\exists k_0 \in [\![N_a + 1\,;\, N_b - 1]\!] \qquad u_{k_0} \in K\,.$$

  En effet, supposons que ce ne soit pas le cas : en choisissant $n \in [\![N_a\,;N_b - 1]\!]$ tel que $u_n \in A'$ et $u_{n+1} \in B'$, on a $d(u_n, u_{n+1}) \geq \frac{\alpha}{3}$, ce qui est absurde puisque $n \geq N_0$.

- Soit $r \in \mathbb{N}$. Supposons construits $k_0 < k_1 < \cdots < k_r$ tels que $u_{k_j} \in K$ pour $j \in [\![0\,;r]\!]$. En réitérant le processus précédent avec N_a choisi tel que $N_a > k_r$ (ce qui est possible puisque V_a est infini), on obtient l'existence de $k_{r+1} \in \mathbb{N}$ tel que $k_{r+1} > k_r$ et $u_{k_{r+1}} \in K$.

Par récurrence, on crée ainsi une sous-suite $(u_{k_r})_{r\in\mathbb{N}}$ à valeurs dans K. Comme K est compact, cette sous-suite admet une valeur d'adhérence $\ell \in K$, qui est aussi une valeur d'adhérence de la suite $(u_n)_{n\in\mathbb{N}}$. Ainsi :

$$\ell \in \Gamma \cap K = (A \cup B) \cap K \subset \left(A' \cup B'\right) \cap K = \varnothing\,.$$

Ce qui est absurde. Finalement, Γ est bien un ensemble connexe.

Passons à l'application.

$\Longrightarrow$ Le sens direct est évident.

$\Longleftarrow$ Réciproquement, supposons que $\lim_{n\to+\infty} x_{n+1} - x_n = 0$.

Notons Γ l'ensemble des valeurs d'adhérence de $(x_n)_{n\in\mathbb{N}}$.
D'après la proposition précédente, l'ensemble Γ est un intervalle fermé de $[0,1]$.

Vérifions de plus que Γ est constitué de points fixes de f. Soient $a \in \Gamma$ et $\varphi : \mathbb{N} \longrightarrow \mathbb{N}$ une fonction strictement croissante telle que $x_{\varphi(n)} \longrightarrow_{n\to+\infty} a$. Alors :

$$\begin{aligned} a &= \lim_{n\to+\infty} x_{\varphi(n)} = \lim_{n\to+\infty} x_{\varphi(n)+1} && \text{puisque } x_{n+1} - x_n \underset{n\to+\infty}{\longrightarrow} 0 \\ &= \lim_{n\to+\infty} f(x_{\varphi(n)}) \\ a &= f(a) && \text{par continuité de } f\,. \end{aligned}$$

Supposons alors que $(x_n)_{n\in\mathbb{N}}$ possède au moins deux valeurs d'adhérences, disons $\ell < \ell'$. Dans ce cas, $[\ell, \ell'] \subset \Gamma$, et on peut donc[2] trouver $N \in \mathbb{N}$ tel que $x_N \in [\ell, \ell']$. Mais alors x_N est un point fixe de f et la suite $(x_n)_{n\in\mathbb{N}}$ est stationnaire en x_N, donc ne possède qu'une seule valeur d'adhérence, ce qui est absurde.

En conclusion, $(x_n)_{n\in\mathbb{N}}$ n'a qu'une valeur d'adhérence, donc converge.

COMMENTAIRES

Les deux résultats peuvent être illustrés par des dessins au tableau. S'il reste un peu de temps, on peut égaleemnt expliquer l'appellation « lemme de la grenouille », en reprenant l'idée de la proposition dans le cadre de l'application.

1. on pourra faire dessiner la figure du [Gou08]

2. puisque par exemple $\frac{\ell+\ell'}{2}$ est valeur d'adhérence

32 DENSITÉ DES POLYNÔMES ORTHOGONAUX

[BMP05, §3.4, p140]

ÉNONCÉ

THÉORÈME. [BASE HILBERTIENNE DE POLYNÔMES ORTHOGONAUX]
Soit $\rho : I \longrightarrow \mathbb{R}$ une fonction de poids où I est un intervalle de $\mathbb{R}$. Soit $(P_n)_{n\in\mathbb{N}}$ la famille de polynômes unitaires, orthogonaux pour $L^2(I,\rho)$, et tels que $\deg(P_n) = n$ pour $n \in \mathbb{N}$. Si $\int_I \mathrm{e}^{\alpha|x|}\,\rho(x)\,\mathrm{d}x < +\infty$ pour un $\alpha > 0$, alors $(P_n)_{n\in\mathbb{N}}$ est une base hilbertienne de $L^2(I,\rho)$.

EXEMPLE. Contre-exemple si l'hypothèse du théorème n'est pas vérifiée : soit la fonction de poids $w : x \longmapsto x^{-\ln(x)}$ définie sur $I = \mathbb{R}_+$, alors $(P_n)_{n\in\mathbb{N}}$ ne forme pas une base de $L^2(I,w)$.

DÉVELOPPEMENT

La famille $(P_n)_{n\in\mathbb{N}}$ est bien définie. En effet, comme $x \in I \longmapsto |x|^n\,\mathrm{e}^{-\alpha|x|}$ est bornée, on a[1] $x^n \in L^1(I,\rho)$ pour tout $n \in \mathbb{N}$. Notons que, pour la même raison, $x^n \in L^2(I,\rho)$ pour $n \in \mathbb{N}$.

Soit alors $f \in L^2(I,\rho)$. Définissons la fonction $\varphi : x \in I \longmapsto f(x)\,\rho(x)$. Comme on a la majoration $\big|\varphi(x)\big| = \big|f(x)\big|\,\rho(x) \leq \frac{1}{2}\big(1 + |f(x)|^2\big)\,\rho(x)$ pour tout $x \in I$, on déduit que

$$\int_I \big|\varphi(x)\big|\,\mathrm{d}x \leq \frac{1}{2}\left(\int_I \rho(x)\,\mathrm{d}x + \int_I \big|f(x)\big|^2\rho(x)\,\mathrm{d}x\right) < +\infty\,.$$

Autrement dit $\varphi \in L^1(I)$ et on peut définir sa transformée de FOURIER

$$\hat{\varphi} : \mathbb{R} \longrightarrow \mathbb{C}, \omega \longmapsto \int_I \mathrm{e}^{-ix\omega}\,f(x)\,\rho(x)\,\mathrm{d}x\,.$$

On veut montrer que $\hat{\varphi}$ est prolongeable sur $B_\alpha = \big\{z \in \mathbb{C} : \big|\mathrm{Im}(z)\big| < \frac{\alpha}{2}\big\}$ en une application holomorphe. Considérons pour cela la fonction $g : B_\alpha \times I \longrightarrow \mathbb{C}, (z,x) \longmapsto \mathrm{e}^{-ixz}\,f(x)\,\rho(x)$ et vérifions qu'elle satisfait les hypothèses du théorème d'holomorphie sous le signe intégral :

- pour tout $z \in B_\alpha$, l'application $g(z,\,\cdot\,)$ est mesurable,
- pour tout $x \in I$, l'application $g(\,\cdot\,,x)$ est holomorphe,
- pour tout $z \in B_\alpha$, on a $\big|g(z,x)\big| \leq \mathrm{e}^{\alpha|x|/2}\,|f(x)|\,\rho(x) = h(x)$ pour $x \in I$ et on vérifie que la dominatrice h est intégrable d'après l'inégalité de CAUCHY-SCHWARZ :

$$\int_I h(x)\,\mathrm{d}x \leq \left(\int_I \mathrm{e}^{\alpha|x|}\,\rho(x)\,\mathrm{d}x\right)^{1/2}\left(\int_I \big|f(x)\big|^2\rho(x)\,\mathrm{d}x\right)^{1/2} < +\infty\,.$$

Ainsi $F : z \in B_\alpha \longmapsto \displaystyle\int_I \mathrm{e}^{-ixz}\,f(x)\,\rho(x)\,\mathrm{d}x$ est holomorphe et coïncide avec $\hat{\varphi}$ sur $\mathbb{R}$.

1. on fait l'abus de notation x^n pour désigner l'application $x \in I \longmapsto x^n$

Supposons que $f \in (x^n)^\perp$ pour tout $n \in \mathbb{N}$.
Étant donné que F est holomorphe, il vient, pour tout $n \in \mathbb{N}$:

$$F^{(n)}(0) = \int_I \partial_z^n\big(\mathrm{e}^{-ixz}\,f(x)\,\rho(x)\big)_{|0}\,\mathrm{d}x = (-i)^n\int_I x^n f(x)\,\rho(x)\,\mathrm{d}x = (-i)^n\,\langle x^n \mid f\rangle_\rho = 0\,.$$

Ainsi $F = 0$ au voisinage de 0, ce qui implique, par unicité du prolongement analytique, que F est nulle sur B_α. En particulier, $\hat{\varphi} = 0$, et donc $\varphi = 0$ par injectivité de la transformée de FOURIER sur L^1. Comme $\rho > 0$ p.p., il s'ensuit que $f = 0$ p.p..

Finalement, puisque $(P_n)_{n\in\mathbb{N}}$ est une base de $\mathbb{R}[X]$:

$$\mathrm{Vect}\big((P_n)_{n\in\mathbb{N}}\big)^\perp = \mathrm{Vect}\big((x^n)_{n\in\mathbb{N}}\big)^\perp = \{0\}\,,$$

donc $(P_n)_{n\in\mathbb{N}}$ est dense dans $L^2(I,\rho)$. Comme de plus $(P_n)_{n\in\mathbb{N}}$ est orthogonale, la famille $(P_n)_{n\in\mathbb{N}}$ est une base hilbertienne de $L^2(I,\rho)$.

Passons au contre-exemple. On vérifie préalablement, par changement de variable $y = \ln(x)$, que $x^n \in L^2(I,w)$ pour tout $n \in \mathbb{N}$. Considérant $f : x \in \mathbb{R}_+ \longmapsto \sin\big(2\pi\ln(x)\big)$, on calcule

$$\begin{aligned}
\forall n \in \mathbb{N} \qquad \langle f \mid x^n\rangle_\rho &= \int_{\mathbb{R}_+} x^n \sin\big(2\pi\ln(x)\big)\,x^{-\ln(x)}\,\mathrm{d}x \\
&= \int_{\mathbb{R}_+} \mathrm{e}^{(n+1)\ln(x)}\sin\big(2\pi\ln(x)\big)\,\mathrm{e}^{-\ln(x)^2}\,\frac{\mathrm{d}x}{x} \\
&= \int_{\mathbb{R}} \mathrm{e}^{(n+1)y}\sin(2\pi y)\,\mathrm{e}^{-y^2}\,\mathrm{d}y && \text{par CDV } y = \ln(x) \\
&= \mathrm{e}^{\frac{(n+1)^2}{4}}\int_{\mathbb{R}} \sin(2\pi y)\,\mathrm{e}^{-(y-\frac{n+1}{2})^2}\,\mathrm{d}y \\
&= \mathrm{e}^{\frac{(n+1)^2}{4}}\int_{\mathbb{R}} \sin\big(2\pi t + \pi(n+1)\big)\,\mathrm{e}^{-t^2}\,\mathrm{d}t && \text{par CDV } t = y - \frac{n+1}{2} \\
&= (-1)^{n+1}\,\mathrm{e}^{\frac{(n+1)^2}{4}}\int_{\mathbb{R}} \sin(2\pi t)\,\mathrm{e}^{-t^2}\,\mathrm{d}t \\
\forall n \in \mathbb{N} \qquad \langle f \mid x^n\rangle_\rho &= 0 && \text{par imparité}
\end{aligned}$$

Ainsi, $f \in \mathrm{Vect}\big((x^n)_{n\in\mathbb{N}}\big)^\perp = \mathrm{Vect}\big((P_n)_{n\in\mathbb{N}}\big)^\perp$.

Puisque $f \neq 0$, la famille des polynômes orthogonaux n'est pas totale.

COMMENTAIRES

Pour ne pas perdre le jury, il faut prendre le temps d'expliquer le raisonnement que l'on va suivre. Il faut aussi connaître des exemples de familles de polynômes orthogonaux !

33 ÉQUATION DE BURGERS

[BMP05, §1.6, p36]

ÉNONCÉ

THÉORÈME. [ÉQUATION DE BURGERS GÉNÉRALISÉE]
Soient $a, u_0 : \mathbb{R} \longrightarrow \mathbb{R}$ des fonctions de classe $\mathcal{C}^1$, avec u_0 et u_0' bornées. Il existe $T \in]0\,;+\infty]$ tel que le système suivant admette une unique solution de classe $\mathcal{C}^1$ sur $]-T\,;T[\times\mathbb{R}$:

$$\begin{cases} \partial_t u + a(u)\,\partial_x u = 0 \\ u(0,\,\cdot\,) = u_0\,. \end{cases}$$

DÉVELOPPEMENT

On va procéder par analyse/synthèse.
Supposons qu'il existe une solution u sur $]-T\,;T[\times\mathbb{R}$ où $T \in]0\,;+\infty]$ est à déterminer.

1. Soit $x_0 \in \mathbb{R}$ et soit le problème de CAUCHY :

$$\forall t \in]-T\,;T[\qquad x'(t) = a\big(u(t,x(t))\big) \qquad \text{avec} \qquad x(0) = x_0\,.$$

L'application $a \circ u$ étant $\mathcal{C}^1$, le théorème de CAUCHY-LIPSCHITZ assure l'existance d'une solution maximale. Étant donné que la solution x n'explose pas sur tout compact (x' y étant bornée), le lemme des bouts assure qu'elle est définie globalement sur $]-T\,;T[$. De plus, comme u est solution de l'équation de BURGERS, on a pour tout $t \in]-T\,;T[$:

$$\frac{\mathrm{d}}{\mathrm{d}t} u\big(t,x(t)\big) = \partial_t u\big(t,x(t)\big) + \partial_x u\big(t,x(t)\big)\,x'(t) = \big(\partial_t u + a(u)\,\partial_x u\big)\big(t,x(t)\big) = 0\,.$$

Ainsi, u est constante sur la courbe $\big(t,x(t)\big)_{t\in]-T\,;T[}$. En reprenant l'équation satisfaite par la solution x, on observe que x' est constante, égale à $a\big(u_0(x_0)\big)$, d'où :

$$\forall t \in]-T\,;T[\qquad x(t) = x_0 + t\,a\big(u_0(x_0)\big)\,.$$

2. On voudrait trouver une fonction $\varphi \in \mathcal{C}^1\big(]-T\,;T[\times\mathbb{R},\mathbb{R}\big)$ vérifiant[1]

$$\forall (t,x) \in]-T\,;T[\times\mathbb{R} \qquad u(t,x) = u_0\big(\varphi(t,x)\big) \quad \text{et} \quad \varphi(0,x) = x\,.$$

Posons $c = a \circ u_0$ et considérons $M = \sup_{\mathbb{R}} |c'| < +\infty$. En effet, c' est bornée puisque $c' = a'(u_0)\,u_0'$ avec par hypothèse u_0, u_0' bornées et a' continue.
Quitte à réduire T, on peut supposer $T \leq \frac{1}{M} \in]0\,;+\infty]$. Étudions alors l'application

$$\begin{array}{rccc} F : &]-T\,;T[\times\mathbb{R} & \longrightarrow &]-T\,;T[\times\mathbb{R} \\ & (t,x) & \longmapsto & \big(t, x + t\,c(x)\big)\,. \end{array}$$

- F est surjective. En effet, soit $t \in]-T\,;T[$ et $\psi_t : x \in \mathbb{R} \longmapsto x + t\,c(x)$. Pour $x \in \mathbb{R}$, $|t\,c'(x)| \leq |t| \cdot M < 1$, d'où uniformément $\psi_t' = 1 + t\,c' > 0$, et ainsi ψ_t est d'image $\mathbb{R}$.
- F est injective. En effet si $F(t_1,x_1) = F(t_2,x_2)$ pour $(t_1,x_1),(t_2,x_2) \in]-T\,;T[\times\mathbb{R}$, alors $t_1 = t_2$, puis $x_2 - x_1 = t_1\big(c(x_1) - c(x_2)\big) = t_1 \int_{x_2}^{x_1} c'(x)\,\mathrm{d}x$, d'où l'on tire $|x_2 - x_1| \leq |t_1| \cdot M \cdot |x_2 - x_1|$ avec $|t_1| \cdot M < 1$, ce qui n'est possible que si $x_1 = x_2$.
- F est $\mathcal{C}^1$ et, pour tout $(t,x) \in]-T\,;T[\times\mathbb{R}$, comme $\big|t\,c'(x)\big| \leq |t| \cdot M < 1$, on sait que $DF(t,x) = \begin{pmatrix} 1 & 0 \\ c(x) & 1 + t\,c'(x) \end{pmatrix}$ est inversible.

D'après le théorème d'inversion globale, justifié par les deux derniers points,

$$\exists G \in \mathcal{C}^1\big(]-T\,;T[\times\mathbb{R}\big) \qquad \begin{cases} \forall (t,x) \in]-T\,;T[\times\mathbb{R} & G\circ F(t,x) = (t,x) \\ \forall (t,y) \in]-T\,;T[\times\mathbb{R} & F\circ G(t,y) = (t,y)\,. \end{cases}$$

Pour $(t,y) \in]-T\,;T[\times\mathbb{R}$, on remarque que $G(t,y)$ est de la forme $\big(t,\varphi(t,y)\big)$ pour une certaine fonction φ de classe $\mathcal{C}^1$ et uniquement déterminée par a et u_0. Comme pour $(t,x) \in]-T\,;T[\times\mathbb{R}$, on a $u\big(F(t,x)\big) = u_0(x)$, il vient :

$$\begin{aligned} \big(0,\varphi(0,x)\big) &= G(0,x) = G\big(F(0,x)\big) = (0,x) \\ \text{et} \qquad u(t,x) &= u\big(F\big(G(t,x)\big)\big) = u\big(F\big(t,\varphi(t,x)\big)\big) = u_0\big(\varphi(t,x)\big)\,. \end{aligned}$$

Ainsi u est uniquement déterminée sur $]-T\,;T[\times\mathbb{R}$.

Réciproquement, fixons $T \leq \frac{1}{M}$ et définissons $u = u_0 \circ \varphi$.
Notons que φ vérifie $\varphi\big(F(t,x)\big) = \varphi\big(t, x + t\,c(x)\big) = x$ pour $(t,x) \in]-T\,;T[\times\mathbb{R}$, d'où

$$\partial_t \varphi\big(F(t,x)\big) + c(x)\,\partial_x \varphi\big(F(t,x)\big) = 0\,,$$

soit, F étant surjective, $\partial_t\varphi(t,y) + c(\varphi(t,y))\,\partial_x\varphi(t,y) = 0$ pour $(t,y) \in]-T\,;T[\times\mathbb{R}$. Ainsi :

$$\partial_t u + a(u)\,\partial_x u = \big(u_0' \circ \varphi\big)\,\big(\partial_t\varphi + a(u)\,\partial_x\varphi\big) = \big(u_0'\circ\varphi\big)\,\big(\partial_t\varphi + (c\circ\varphi)\,\partial_x\varphi\big) = 0\,.$$

Ainsi u vérifie l'équation de BURGERS généralisée, et clairement $u(0,\,\cdot\,) = u_0$.

COMMENTAIRES

L'équation de BURGERS (non visqueuse) est $\partial_t u + u\,\partial_x u = 0$. Ici on considère une équation généralisée du type $\partial_t u + \partial_x A(u) = 0$, qui se réécrit $\partial_t u + A'(u)\,\partial_x u = 0$.

Il faut bien expliquer le raisonnement suivi pour l'unicité : l'idée consiste à chercher des courbes $\big(t,x(t)\big)_{t\in]-T\,;T[}$ sur lesquelles u vérifie une équation différentielle ordinaire : si l'on impose à $x(t)$ de varier à la vitesse de la propagation, on s'attend à ce que u soit constante sur la courbe.

Il faut savoir ce qu'il se passe lorsque $T > \frac{1}{M}$ ou lorsque l'on se place uniquement sur $x \in \mathbb{R}_+$.

1. $\varphi(t,x)$ est le point d'origine d'une potentielle courbe caractéristique passant par un point $(t,x) \in]-T\,;T[\times\mathbb{R}$

34 ÉQUATION DE LA CHALEUR PÉRIODIQUE

[FGN12, §1.28, p49] [QZ13, §IV.VI.3, p105]

ÉNONCÉ

THÉORÈME. [ÉQUATION DE LA CHALEUR PÉRIODIQUE]
Soit $u_0 : \mathbb{R} \longrightarrow \mathbb{R}$ non identiquement nulle, 2π-périodique continue et $\mathcal{C}^1_{pm}$.
Il existe une unique solution $u \in \mathcal{C}^0(\mathbb{R}_+ \times \mathbb{R}) \cap \mathcal{C}^\infty(\mathbb{R}_+^* \times \mathbb{R})$ au système :

$$\begin{cases} \partial_t u = \partial^2_{xx} u \\ u(0,\,\cdot\,) = u_0\,. \end{cases}$$

DÉVELOPPEMENT

Procédons par analyse-synthèse. Supposons d'abord u solution. Pour $t > 0$ et $x \in \mathbb{T}$, on a :

$$\begin{cases} u(t,x) = \displaystyle\sum_{n\in\mathbb{Z}} c_n(t)\,\mathrm{e}^{\mathrm{i}nx}\,, & \text{où } c_n(t) = \dfrac{1}{2\pi}\displaystyle\int_0^{2\pi} u(t,x)\,\mathrm{e}^{-\mathrm{i}nx}\,\mathrm{d}x\,, \\ \partial_t u(t,x) = \displaystyle\sum_{n\in\mathbb{Z}} \tilde{c}_n(t)\,\mathrm{e}^{\mathrm{i}nx}\,, & \text{où } \tilde{c}_n(t) = \dfrac{1}{2\pi}\displaystyle\int_0^{2\pi} \partial_t u(t,x)\,\mathrm{e}^{-\mathrm{i}nx}\,\mathrm{d}x = c_n'(t)\,, \\ \partial^2_{xx} u(t,x) = -\displaystyle\sum_{n\in\mathbb{Z}} n^2 c_n(t)\,\mathrm{e}^{\mathrm{i}nx}\,. \end{cases}$$

En effet, les fonctions considérées sont de classe $\mathcal{C}^\infty$ en la variable spatiale : leur série de FOURIER converge donc en tout point. La relation $\tilde{c}_n = c_n'$ entre coefficients de FOURIER découle du théorème de dérivation sous le signe intégral. L'expression donnée pour ∂^2_{xx} s'obtient soit par une double intégration par parties des coefficients de FOURIER, soit par une double dérivation terme à terme, les séries de fonctions de $\partial_x u$ et $\partial^2_{xx} u$ convergeant normalement (fonctions de classe $\mathcal{C}^1$). Ces trois séries convergent normalement en la variable spatiale.

Soit $t > 0$. Comme u est solution, la somme $S(x) = \sum_{n\in\mathbb{Z}} \big(c_n'(t) + n^2 c_n(t)\big)\,\mathrm{e}^{\mathrm{i}nx}$ est nulle pour tout $x \in \mathbb{T}$, avec convergence normale de la série. Ainsi

$$\begin{aligned} \forall p \in \mathbb{Z} \qquad 0 &= \frac{1}{2\pi}\int_0^{2\pi} \sum_{n\in\mathbb{Z}} \big(c_n'(t) + n^2 c_n(t)\big)\,\mathrm{e}^{\mathrm{i}(n-p)x}\,\mathrm{d}x \\ &= \frac{1}{2\pi}\sum_{n\in\mathbb{Z}} \big(c_n'(t) + n^2 c_n(t)\big)\int_0^{2\pi} \mathrm{e}^{\mathrm{i}(n-p)x}\,\mathrm{d}x = c_p'(t) + p^2 c_p(t) \end{aligned}$$

Ceci étant vrai pour tout $t > 0$, on en déduit que

$$\forall p \in \mathbb{Z} \quad \exists C_p \in \mathbb{C} \quad \forall t > 0 \qquad c_p(t) = C_p\,\mathrm{e}^{-p^2 t}\,.$$

Fixons $t \in]0,1]$ et appliquons l'égalité de PARSEVAL à $u(0,\,\cdot\,) - u(t,\,\cdot\,)$. En notant, pour $n \in \mathbb{Z}$, $c_n(0)$ le n-ième coefficient de FOURIER de u_0, il vient par convergence dominée

$$\sum_{n\in\mathbb{Z}} \big|c_n(0) - c_n(t)\big|^2 = \frac{1}{2\pi}\int_0^{2\pi} \big|u(0,x) - u(t,x)\big|^2\,\mathrm{d}x \underset{t\to 0}{\longrightarrow} 0\,,$$

puisque $\big|u(0,\,\cdot\,) - u(t,\,\cdot\,)\big|$ est bornée uniformément pour $t \in]0,1]$. Ainsi, pour tout $p \in \mathbb{Z}$, $c_p(t) \longrightarrow_{t\to 0} c_p(0)$, soit $C_p = c_p(0)$. Une potentielle solution u est donc unique, donnée par

$$\forall t \in \mathbb{R}_+ \quad \forall x \in \mathbb{T} \qquad u(t,x) = \sum_{n\in\mathbb{Z}} C_n\,\mathrm{e}^{-n^2 t}\,\mathrm{e}^{\mathrm{i}nx}\,.$$

Vérifions que u définie ainsi est solution de l'équation de la chaleur.

Soit $(t,x) \in \mathbb{R}_+^* \times \mathbb{T}$. Comme, pour $n \in \mathbb{Z}$, on a la majoration $|C_n\,\mathrm{e}^{-n^2 t}\,\mathrm{e}^{\mathrm{i}nx}| \leq |C_n|$, et que $|C_n|$ est le terme général d'une série convergente puisque $u_0 \in \mathcal{C}^1_{\text{pm}}$, on sait que u conserge normalement. En particulier u est bien définie et $u \in \mathcal{C}^0(\mathbb{R}_+ \times \mathbb{R})$.
De plus, u est clairement 2π-périodique et satisfait la condition initiale $u(0,\,\cdot\,) = u_0$.

Fixons $k, \ell \in \mathbb{N}$. Pour tout $n \in \mathbb{Z}$ et $(t,x) \in \mathbb{R}_+ \times \mathbb{T}$, il vient

$$\frac{\partial^{k+\ell}}{\partial t^k \partial x^\ell}\big(C_n\,\mathrm{e}^{-n^2 t}\,\mathrm{e}^{\mathrm{i}nx}\big) = C_n\,(-1)^k\,\mathrm{i}^\ell\,n^{2k+\ell}\,\mathrm{e}^{-n^2 t}\,\mathrm{e}^{\mathrm{i}nx}\,,$$

puis, si $t \geq t_0$ où $t_0 > 0$:

$$\left|\frac{\partial^{k+\ell}}{\partial t^k \partial x^\ell}\big(C_n\,\mathrm{e}^{-n^2 t}\,\mathrm{e}^{\mathrm{i}nx}\big)\right| \leq |C_n|\cdot|n|^{2k+\ell}\,\mathrm{e}^{-n^2 t_0} \leq \|u_0\|_1 \cdot |n|^{2k+\ell}\,\mathrm{e}^{-n^2 t_0} \underset{n\to+\infty}{=} o\Big(\frac{1}{n^2}\Big)\,.$$

Ainsi on obtient que $u \in \mathcal{C}^\infty(\mathbb{R}_+^* \times \mathbb{R})$ par convergence normale, avec :

$$\forall (k,\ell) \in \mathbb{N}^2 \quad \forall (t,x) \in \mathbb{R}_+^* \times \mathbb{T} \qquad \frac{\partial^{k+\ell}}{\partial t^k \partial x^\ell} u(t,x) = \sum_{n\in\mathbb{Z}} C_n\,(-1)^k\,\mathrm{i}^\ell\,n^{2k+\ell}\,\mathrm{e}^{-n^2 t}\,\mathrm{e}^{\mathrm{i}nx}\,.$$

Enfin, l'application u vérifie l'équation de la chaleur en prenant $(k,\ell) = (0,2)$ puis $(1,0)$.

COMMENTAIRES

Il faut être clair sur les convergences normales (interversions série-intégrale) et sur les hypothèses de domination (vite vérifiées ici par périodicité et continuité). Il serait trop long d'écrire toutes les hypothèses, donc il faut les mentionner à l'oral et montrer une certaine aisance.

Par ailleurs, on peut s'attendre à des questions autour de l'équation de la chaleur dans d'autres contextes (par exemple sans la périodicité, voir [Gou08, p348]), ou autour des solutions faibles dans le monde des distributions.

35 ESPÉRANCE CONDITIONNELLE

[BL07, §VI.2, p156–157]

ÉNONCÉ

PROPOSITION. [ESPÉRANCE CONDITIONNELLE]
Soient $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé et $\mathcal{B}$ une sous-tribu de $\mathcal{A}$. Soit de plus $X \in L^1(\Omega, \mathcal{A}, \mathbb{P})$ une variable aléatoire réelle intégrable. Alors il existe une unique (presque sûrement) variable aléatoire Z telle que :
(i) Z est $\mathcal{B}$-mesurable,
(ii) $\mathbb{E}[X\,\mathbb{1}_B] = \mathbb{E}[Z\,\mathbb{1}_B]$ pour tout $B \in \mathcal{B}$.
De plus, Z est intégrable.

DÉFINITION. [ESPÉRANCE CONDITIONNELLE]
Ainsi définie, Z est appelée espérance conditionnelle de X sachant $\mathcal{B}$ et notée $\mathbb{E}[X \mid \mathcal{B}]$.

DÉVELOPPEMENT

Commençons par montrer l'unicité.

Soient Z_1 et Z_2 deux variables aléatoires satisfaisant les propriétés (i) et (ii).
Étant donné que Z_1 et Z_2 sont $\mathcal{B}$-mesurables, $W = \mathbb{1}_{Z_1 > Z_2}$ l'est aussi et alors

$$\mathbb{E}[Z_1\,W] = \mathbb{E}[X\,W] = \mathbb{E}[Z_2\,W] \qquad \text{soit} \qquad \mathbb{E}\big[(Z_1 - Z_2)\,\mathbb{1}_{Z_1 > Z_2}\big] = 0\,.$$

Comme par ailleurs $(Z_1 - Z_2)\,\mathbb{1}_{Z_1 > Z_2} \geq 0$ p.s., on en déduit que

$$(Z_1 - Z_2)\,\mathbb{1}_{Z_1 > Z_2} = 0 \text{ p.s.} \qquad \text{ou encore} \qquad Z_1 \leq Z_2 \text{ p.s..}$$

En raisonnement de manière symétrique, on obtient finalement $Z_1 = Z_2$ p.s., d'où l'unicité.

Passons à l'existence. On procède en trois étapes selon l'espace dans lequel X vit.

1. Supposons $X \in L^2(\Omega, \mathcal{A}, \mathbb{P})$.
Rappelons que $L^2(\Omega, \mathcal{A}, \mathbb{P})$ est un espace de HILBERT, dont $L^2(\Omega, \mathcal{B}, \mathbb{P})$ est un sous-espace vectoriel convexe fermé (puisqu'une limite de fonctions $\mathcal{B}$-mesurables est $\mathcal{B}$-mesurable). On peut donc considérer la projection Z de X sur $L^2(\Omega, \mathcal{B}, \mathbb{P})$. Alors Z satisfait la propriété (i) et comme $\mathbb{1}_B \in L^2(\Omega, \mathcal{B}, \mathbb{P})$ pour $B \in \mathcal{B}$, il vient :
$$\forall B \in \mathcal{B} \qquad \mathbb{E}[X\,\mathbb{1}_B] = \mathbb{E}\big[(X - Z)\,\mathbb{1}_B\big] + \mathbb{E}[Z\,\mathbb{1}_B] = \mathbb{E}[Z\,\mathbb{1}_B]\,.$$
Ainsi, $Z = \mathbb{E}[X \mid \mathcal{B}]$ satisfait les deux propriétés démandées, et Z est intégrable car L^2.
Avant de poursuivre, notons que $\mathbb{E}[X \mid \mathcal{B}]$ satisfait les propriétés suivantes sur $L^2(\Omega, \mathcal{A}, \mathbb{P})$.
   - Si $X \in L^2$ est positive p.s., alors $\mathbb{E}[X \mid \mathcal{B}] \geq 0$ p.s..
   En effet, la propriété (ii) avec $B = \big\{\mathbb{E}[X \mid \mathcal{B}] < 0\big\} \in \mathcal{B}$ donne que $\mathbb{1}_B = 0$ p.s..
   - Si $X_1, X_2 \in L^2$, alors $\mathbb{E}[X_1 + X_2 \mid \mathcal{B}] = \mathbb{E}[X_1 \mid \mathcal{B}] + \mathbb{E}[X_2 \mid \mathcal{B}]$.
   C'est une conséquence de la linéarité de la projection.
   - Si $X \in L^2$, alors $\mathbb{E}[X] = \mathbb{E}\big[\mathbb{E}[X \mid \mathcal{B}]\big]$ en prenant $B = \Omega$ dans la propriété (ii).
2. Supposons $X \in L^1$ et positive p.s..
Pour $n \in \mathbb{N}$, on définit la variable aléatoire $X_n = \min(X, n) \in L^2$.
Par le point précédent, les $\big(\mathbb{E}[X_n \mid \mathcal{B}]\big)_{n\in\mathbb{N}}$ existent et constituent une suite est croissante puisque $(X_n)_{n\in\mathbb{N}}$ l'est. Posons alors
$$Z = \lim_{n\to+\infty} \uparrow \mathbb{E}[X_n \mid \mathcal{B}]\,.$$
La variable Z est $\mathcal{B}$-mesurable et
$$\begin{aligned}
\forall B \in \mathcal{B} \qquad \mathbb{E}[X\,\mathbb{1}_B] &= \lim_{n\to+\infty} \uparrow \mathbb{E}[X_n\,\mathbb{1}_B] && \text{par convergence monotone}\\
&= \lim_{n\to+\infty} \uparrow \mathbb{E}\big[\mathbb{E}[X_n \mid \mathcal{B}]\,\mathbb{1}_B\big] && \text{par la propriété (ii) dans le cas } L^2\\
\forall B \in \mathcal{B} \qquad \mathbb{E}[X\,\mathbb{1}_B] &= \mathbb{E}[Z\,\mathbb{1}_B] && \text{par convergence monotone}
\end{aligned}$$
Donc $Z = \mathbb{E}[X \mid \mathcal{B}]$ vérifie les deux propriétés, et satisfait de plus $\mathbb{E}[X \mid \mathcal{B}] \geq 0$ p.s..
Prenant $B = \Omega$ ci-dessus, on obtient que $\mathbb{E}[X] = \mathbb{E}\big[\mathbb{E}[X \mid \mathcal{B}]\big] \in [0\,;+\infty]$ dans ce cas.
3. Supposons enfin que $X \in L^1$.
On décompose X sous la forme $X = X^+ - X^-$ avec X^+, X^- des variables aléatoires L^1 et positives p.s.. Par le cas précédent, on peut considérer la variable aléatoire $\mathcal{B}$-mesurable
$$Z = \mathbb{E}[X^+ \mid \mathcal{B}] - \mathbb{E}[X^- \mid \mathcal{B}]\,.$$
Remarquons que Z est intégrable puisque
$$\mathbb{E}\big[|Z|\big] = \mathbb{E}\big[\mathbb{E}[X^+ \mid \mathcal{B}]\big] + \mathbb{E}\big[\mathbb{E}[X^- \mid \mathcal{B}]\big] = \mathbb{E}[X^+] + \mathbb{E}[X^-] = \mathbb{E}\big[|X|\big]\,.$$
Un calcul similaire assure alors que $\mathbb{E}[Z] = \mathbb{E}[X]$. De plus, par linéarité de l'espérance et par la propriété (ii) dans le cas L^1 positif :
$$\begin{aligned}
\forall B \in \mathcal{B} \qquad \mathbb{E}[X\,\mathbb{1}_B] &= \mathbb{E}\big[(X^+ - X^-)\,\mathbb{1}_B\big]\\
&= \mathbb{E}[X^+\,\mathbb{1}_B] - \mathbb{E}[X^-\,\mathbb{1}_B]\\
&= \mathbb{E}\big[\mathbb{E}[X^+ \mid \mathcal{B}]\,\mathbb{1}_B\big] - \mathbb{E}\big[\mathbb{E}[X^- \mid \mathcal{B}]\,\mathbb{1}_B\big]\\
&= \mathbb{E}\big[\big(\mathbb{E}[X^+ \mid \mathcal{B}] - \mathbb{E}[X^- \mid \mathcal{B}]\big)\,\mathbb{1}_B\big]\\
\forall B \in \mathcal{B} \qquad \mathbb{E}[X\,\mathbb{1}_B] &= \mathbb{E}[Z\,\mathbb{1}_B]\,.
\end{aligned}$$
Finalement $Z = \mathbb{E}[X \mid \mathcal{B}]$ satisfait les propriétés demandées, et on a aussi démontré que
$$\mathbb{E}\big[\mathbb{E}[X \mid \mathcal{B}]\big] = \mathbb{E}[X]\,.$$

36 ÉTUDE DE DEUX SUITES RÉCURRENTES

[FGN07, §2.25, p99–103]

ÉNONCÉ

Application. Soit f une application continue définie au voisinage de 0^+ admettant un développement asymptotique en 0 de la forme $f(x) =_{x\to 0} x - ax^\alpha + o(x^\alpha)$, où $a > 0$ et $\alpha > 1$. Pour $u_0 > 0$ assez petit, la suite $(u_n)_{n\in\mathbb{N}}$ définie par la relation de récurrence $u_{n+1} = f(u_n)$ pour $n \in \mathbb{N}$ vérifie

$$u_n \underset{n\to+\infty}{\sim} \frac{1}{\big(na(\alpha-1)\big)^{\frac{1}{\alpha-1}}}\,.$$

Exemples de $f = \sin$ et de $f = \log(1+\,\cdot\,)$.

Application. Soit $(u_n)_{n\in\mathbb{N}}$ la suite définie par la condition initiale $u_0 \in \mathbb{R}$ et par la relation de récurrence $u_{n+1} = u_n + \mathrm{e}^{-u_n}$ pour $n \in \mathbb{N}$. On a le développement asymptotique d'ordre 2

$$u_n \underset{n\to+\infty}{=} \ln n + \frac{\ln n}{2n} + o\Big(\frac{\ln n}{n}\Big)\,.$$

DÉVELOPPEMENT

Commençons par la première application.
Par continuité de f, notons que $f(0) = 0$. Aussi, comme $f(x) =_{x\to 0^+} x - ax^\alpha\big(1+o(1)\big)$

$$\exists \eta \in \mathbb{R}_+^* \quad \forall x \in]0\,;\eta] \qquad f(x) < x\,.$$

Choisissons $u_0 \in]0\,;\eta]$. La suite $(u_n)_{n\in\mathbb{N}}$ est alors décroissante et minorée par 0 : elle converge vers un point fixe de f sur $[0\,;\eta]$, c'est-à-dire vers 0. Ainsi $u_n \longrightarrow_{n\to+\infty} 0$.

Afin d'appliquer le théorème de Cesàro, on voudrait trouver un réel β tel que $\big(u_{n+1}^\beta - u_n^\beta\big)_{n\in\mathbb{N}}$ converge vers une limite non nulle. Or

$$\begin{aligned} f(x)^\beta - x^\beta &\underset{x\to 0}{=} \big(x - ax^\alpha + o(x^\alpha)\big)^\beta - x^\beta \underset{x\to 0}{=} x^\beta\Big(\big(1 - ax^{\alpha-1} + o(x^{\alpha-1})\big)^\beta - 1\Big) \\ &\underset{x\to 0}{=} x^\beta\big(-a\beta x^{\alpha-1} + o(x^{\alpha-1})\big) \underset{x\to 0}{\sim} -a\beta x^{\alpha+\beta-1}\,. \end{aligned}$$

Prenant $\beta = 1-\alpha$, on a $f(x)^{1-\alpha} - x^{1-\alpha} \longrightarrow_{x\to 0^+} a(\alpha-1)$. Comme $u_n \longrightarrow_{n\to+\infty} 0$, il vient

$$u_{n+1}^{1-\alpha} - u_n^{1-\alpha} \underset{n\to+\infty}{\longrightarrow} a(\alpha-1)\,, \qquad \text{et donc} \qquad u_n^{1-\alpha} - u_0^{1-\alpha} \underset{n\to+\infty}{\sim} na(\alpha-1)$$

d'après le théorème de Cesàro. Ainsi

$$u_n^{1-\alpha} \underset{n\to+\infty}{\sim} na(\alpha-1)\,, \qquad \text{soit finalement} \qquad u_n \underset{n\to+\infty}{\sim} \frac{1}{\big(na(\alpha-1)\big)^{\frac{1}{\alpha-1}}}\,.$$

On peut appliquer cette formule avec, par exemple, les fonctions :

- $f(x) = \sin(x) \underset{x\to 0}{=} x - \frac{x^3}{6} + o(x^3)$. En prenant $a = \frac{1}{6}$ et $\alpha = 3$, il vient :

$$u_n \underset{n\to+\infty}{\sim} \frac{1}{\big(\frac{n}{3}\big)^{\frac{1}{2}}} = \sqrt{\frac{3}{n}}\,,$$

- $f(x) = \ln(1+x) \underset{x\to 0}{=} x - \frac{x^2}{2} + o(x^2)$. En prenant $a = \frac{1}{2}$ et $\alpha = 2$, on obtient :

$$u_n \underset{n\to+\infty}{\sim} \frac{1}{\big(\frac{n}{2}\big)^{1}} = \frac{2}{n}\,.$$

Passons à la deuxième application.
Remarquons d'abord que la suite $(u_n)_{n\in\mathbb{N}}$ est croissante et que, si elle converge vers une limite finie ℓ, alors par continuité $\ell = \ell + \mathrm{e}^{-\ell}$, ce qui est impossible. Ainsi $(u_n)_{n\in\mathbb{N}}$ diverge vers $+\infty$.

Posons alors $v_n = \mathrm{e}^{u_n}$ pour $n \in \mathbb{N}$. On a que

$$\forall n \in \mathbb{N} \qquad v_{n+1} = \mathrm{e}^{u_{n+1}} = \mathrm{e}^{u_n}\exp(\mathrm{e}^{-u_n}) = v_n \exp(1/v_n)\,.$$

Comme $(v_n)_{n\in\mathbb{N}}$ diverge aussi vers $+\infty$, le développement limité en 0 de l'exponentielle donne

$$v_{n+1} \underset{n\to+\infty}{=} v_n\left(1 + \frac{1}{v_n} + \frac{1}{2v_n^2} + o\Big(\frac{1}{v_n^2}\Big)\right) \underset{n\to+\infty}{=} v_n + 1 + \frac{1}{2v_n} + o\Big(\frac{1}{v_n}\Big)\,,$$

et ainsi $v_{n+1} - v_n \underset{n\to\infty}{\longrightarrow} 1$. D'après le théorème de Cesàro,

$$v_n \underset{n\to+\infty}{\sim} n\,, \qquad \text{d'où} \qquad v_{n+1} - v_n - 1 \underset{n\to+\infty}{\sim} \frac{1}{2v_n} \underset{n\to+\infty}{\sim} \frac{1}{2n}\,.$$

Étant donné que $\frac{1}{2n}$ est le terme général de signe positif d'une série divergente, le théorème de sommation des équivalents permet d'écrire

$$v_{n+1} - v_1 - n = \sum_{k=1}^{n} v_{k+1} - v_k - 1 \underset{n\to+\infty}{\sim} \sum_{k=1}^{n} \frac{1}{2k} \underset{n\to+\infty}{\sim} \frac{\ln n}{2}\,.$$

Autrement dit, $v_n =_{n\to+\infty} n + \frac{\ln n}{2} + o(\ln n)$. En passant au logarithme, on obtient que

$$u_n \underset{n\to+\infty}{=} \ln n + \ln\left(1 + \frac{\ln n}{2n} + o\Big(\frac{\ln n}{n}\Big)\right) \qquad \text{ou encore} \qquad u_n \underset{n\to+\infty}{=} \ln n + \frac{\ln n}{2n} + o\Big(\frac{\ln n}{n}\Big)\,,$$

où dans la dernière étape on a utilisé le développement limité de $\ln(1+\,\cdot\,)$ en 0.

37 FACTORISATIONS LU ET DE CHOLESKY

[AK02, §6.2/3, p113–124]

ÉNONCÉ

THÉORÈME. [FACTORISATION LU]
Soit $A \in \mathcal{M}_n(\mathbb{K})$ une matrice dont tous les mineurs principaux sont non nuls. Alors il existe un unique couple (L,U) de matrices tel que L est triangulaire inférieure avec diagonale de 1, U est triangulaire supérieure et $A = LU$.

THÉORÈME. [FACTORISATION DE CHOLESKY]
Soit A une matrice symétrique réelle définie positive. Alors il existe une unique matrice B triangulaire inférieure telle que tous ses éléments diagonaux soient positifs et $A = BB^*$.

DÉVELOPPEMENT

Commençons par démontrer la factorisation LU. Soit $A \in \mathcal{M}_n(\mathbb{K})$ satisfaisante.

- Existence. On effectue l'algorithme du pivot de GAUSS sur A en vérifiant qu'aucune matrice de permutation n'est nécessaire, *i.e.*, que les pivots sont tous non nuls. Pour cela, on crée par récurrence une suite $(A^k)_{1\leq k\leq n}$ de matrices en posant $A^1 = A$ puis $A^{k+1} = E^k A^k$ pour $k \in [\![1\,;n-1]\!]$, tant que $a^k_{k,k} \neq 0$, avec :

$$E^k = \begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & 1 & & \\ & & & -\ell_{k+1,k} & 1 & \\ & & & \vdots & & \ddots \\ & & & -\ell_{n,k} & & 1 \end{pmatrix}, \qquad \text{où } \ell_{i,k} = \frac{a^k_{i,k}}{a^k_{k,k}} \text{ pour } i \in [\![1\,;n]\!].$$

Supposons tous les $(a^k_{k,k})_{1\leq k\leq n}$ non nuls[1]. Alors la suite $(A_k)_{1\leq k\leq n}$ est bien définie, et

$$\forall k \in [\![1\,;n]\!] \qquad A^k = \begin{pmatrix} a^1_{1,1} & \star & & \cdots & & \star \\ & \ddots & & & & \\ & & a^k_{k,k} & \ddots & & \vdots \\ & & \star & \star & & \\ & & \vdots & \ddots & \ddots & \star \\ & & \star & \cdots & \star & \star \end{pmatrix}.$$

1. en fait il n'y a besoin de cette hypothèse que pour les $n-1$ premiers pivots

On peut alors poser $U = A^n$, $L = (E^1)^{-1}\cdots(E^{n-1})^{-1}$ et vérifier, en obtenant

$$L = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ \ell_{2,1} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ \ell_{n,1} & \cdots & \ell_{n,n-1} & 1 \end{pmatrix},$$

que l'on a bien $A = LU$, $U \in \mathcal{TS}_n(\mathbb{K})$ et $L \in \mathcal{TI}^1_n(\mathbb{K})$.

Reste à voir que les pivots $(a^k_{k,k})_{1\leq k\leq n}$ sont tous non nuls. Procédons par récurrence forte :
- $a^1_{1,1} = \Delta^1 \neq 0$ car le premier mineur principal est non nul.
- Soit $k \in [\![2\,;n-1]\!]$. Supposons les $k-1$ premiers pivots non nuls.
  Cette hypothèse assure que la formule $A^k = E^{k-1}\cdots E^1 A$ est valable. En considérant des produits par blocs on voit que Δ^k la matrice extraite des k premières lignes et colonnes de A se factorise en $\Delta^k = L^k_{1,1}U^k_{1,1}$ où $L^k_{1,1} \in \mathcal{TI}^1_k(\mathbb{K})$ et $U^k_{1,1} \in \mathcal{TS}_k(\mathbb{K})$. Δ^k et $L^k_{1,1}$ étant inversibles, $U^k_{1,1}$ l'est aussi et donc $a^k_{k,k} \neq 0$ puisque c'est le dernier coefficient diagonal de $U^k_{1,1}$.

- Unicité. Soient $(L_1,U_1),(L_2,U_2)$ deux couples convenants.
  On a $L_2 \in \mathcal{TI}^1_n(\mathbb{K})$ et donc $L_2^{-1} \in \mathcal{TI}^1_n(\mathbb{K})$. Comme A est inversible, on en déduit que $U_1 \in \mathcal{TS}_n(\mathbb{K})$ est inversible, d'où

$$L_2^{-1}L_1 = U_2U_1^{-1} \in \mathcal{TI}^1_n(\mathbb{K}) \cap \mathcal{TS}_n(\mathbb{K}) = \{I_n\}\,.$$

  Ainsi $L_1 = L_2$ et $U_1 = U_2$, d'où l'unicité.

Passons à la factorisation de CHOLESKY. Soit A satisfaisante.

- Existence. A est définie positive donc vérifie les hypothèses de factorisation LU : écrivons $A = LU$ et posons $D = \operatorname{diag}\left(\sqrt{u_{1,1}},\dots,\sqrt{u_{n,n}}\right)$, ce qui est possible puisque pour tout $k \in [\![1\,;n]\!]$, on a $\prod_{i=1}^k u_{i,i} = \det\Delta^k > 0$ (les $(u_{i,i})_{1\leq i\leq n}$ sont donc strictement positifs).
  On pose $B = LD$ et $C = D^{-1}U$ qui vérifient toujours $A = BC$. Comme $A = A^*$, il en découle que $C^*B^* = BC$ ou encore

$$B^{-1}C^* = C(B^*)^{-1} \in \mathcal{TI}_n(\mathbb{K}) \cap \mathcal{TS}_n(\mathbb{K})\,,$$

  qui est donc une matrice diagonale. Comme B et C ont même coefficients diagonaux, on en déduit que c'est en fait l'identité et on a $C = B^*$.

- Unicité. Si B_1 et B_2 conviennent, on a $B_2^{-1}B_1 = B_2^*(B_1^*)^{-1}$ et cette matrice, nommée D, est diagonale par les mêmes arguments que précèdemment. Comme $B_1 = B_2D$, on remarque que $A = B_2B_2^* = B_2(DD^*)B_2^*$, d'où $D^2 = I_n$ puisque B_2 est inversible. Les coefficients diagonaux de la décomposition de CHOLESKY étant positifs, on a forcément $D = I_n$ et donc $B_1 = B_2$.

38 FORMULE D'EULER-MACLAURIN ET APPLICATION À LA SÉRIE HARMONIQUE

[Gou08, §4.7, p301]

ÉNONCÉ

Pour $n \in \mathbb{N}^*$, soient B_n et b_n les n^{e} polynôme et nombre de BERNOULLI.

THÉORÈME. [FORMULE D'EULER-MACLAURIN]
Soient $m, n \in \mathbb{Z}$ avec $m < n$, $r \in \mathbb{N}^*$ et $f : [m\,;n] \longrightarrow \mathbb{C}$ une fonction de classe $\mathcal{C}^r$. Alors :

$$\sum_{k=m}^{n} f(k) = \int_m^n f(t)\,\mathrm{d}t + \frac{f(m)+f(n)}{2} + \sum_{\ell=2}^{r} \frac{b_\ell}{\ell!}\big[f^{(\ell-1)}(n) - f^{(\ell-1)}(m)\big] + R_r\,,$$

où $R_r = \frac{(-1)^{r+1}}{r!}\int_m^n \tilde{B}_r(t) f^{(r)}(t)\,\mathrm{d}t$, avec $\tilde{B}_r : t \in \mathbb{R} \longmapsto B_r(t - \lfloor t \rfloor)$.

APPLICATION. [SÉRIE HARMONIQUE]
Soit $r \in \mathbb{N}^*$. On a $H_n =_{n\to+\infty} \ln n + \gamma + \frac{1}{2n} - \sum_{\ell=1}^{r-1} \frac{b_{2\ell}}{2\ell}\frac{1}{n^{2\ell}} + O\big(\frac{1}{n^{2r}}\big)$.

DÉVELOPPEMENT

RAPPEL. On rappelle que $B_1 = X - \frac{1}{2}$ et que pour $n \in \mathbb{N}^*$, $B_n' = nB_{n-1}$, $B_n(0) = b_n$ et $b_{2n+1} = 0$. De plus, $B_n(0) = B_n(1)$ si $n \geq 2$.

Soient $m, n \in \mathbb{Z}$ avec $m < n$. Procédons par récurrence sur $r \in \mathbb{N}^*$, en définissant la propriété

$$\mathcal{H}_r : \text{« la formule est vraie pour toute fonction } f : [m\,;n] \longrightarrow \mathbb{C} \text{ de classe } \mathcal{C}^r \text{ »}.$$

- Initialisation. Soit $f : [m\,;n] \longrightarrow \mathbb{C}$ une fonction de classe $\mathcal{C}^1$. Montrons que

$$R_1 = \int_m^n \tilde{B}_1(t) f'(t)\,\mathrm{d}t = \sum_{k=m}^{n} f(k) - \int_m^n f(t)\,\mathrm{d}t - \frac{f(m)+f(n)}{2}\,.$$

  Comme $B_1 = X - \frac{1}{2}$, on obtient en intégrant par parties pour $k \in [\![m\,;n-1]\!]$:

$$\int_k^{k+1} \tilde{B}_1(t) f'(t)\,\mathrm{d}t = \big[(t-k-\tfrac{1}{2})f(t)\big]_k^{k+1} - \int_k^{k+1} f(t)\,\mathrm{d}t = \frac{f(k+1)+f(k)}{2} - \int_k^{k+1} f(t)\,\mathrm{d}t\,.$$

  En sommant sur $k \in [\![m\,;n-1]\!]$, on obtient la formule ci-dessus et $\mathcal{H}_1$ est vraie.
- Hérédité. Soit $r \geq 2$ et supposons $\mathcal{H}_{r-1}$ vraie. Soit $f : [m\,;n] \longrightarrow \mathbb{C}$ de classe $\mathcal{C}^r$. On a

$$\sum_{k=m}^{n} f(k) = \int_m^n f(t)\,\mathrm{d}t + \frac{f(m)+f(n)}{2} + \sum_{\ell=2}^{r-1} \frac{b_\ell}{\ell!}\big[f^{(\ell-1)}(n) - f^{(\ell-1)}(m)\big] + R_{r-1}$$

  par hypothèse de récurrence. Vérifions que $R_{r-1} = \frac{b_r}{r!}\big[f^{(r-1)}(n) - f^{(r-1)}(m)\big] + R_r$.

  En intégrant par parties pour $k \in [\![m\,;n-1]\!]$, les propriétés de B_r en préambule donnent

$$\begin{aligned}\int_k^{k+1} \tilde{B}_r(t) f^{(r)}(t)\,\mathrm{d}t &= \big[\tilde{B}_r(t) f^{(r-1)}(t)\big]_k^{k+1} - r\int_k^{k+1} \tilde{B}_{r-1}(t) f^{(r-1)}(t)\,\mathrm{d}t \\ &= b_r\big[f^{(r-1)}(k+1) - f^{(r-1)}(k)\big] - r\int_k^{k+1} \tilde{B}_{r-1}(t) f^{(r-1)}(t)\,\mathrm{d}t\,,\end{aligned}$$

  d'où $$R_r = \frac{(-1)^{r+1} b_r}{r!}\big[f^{(r-1)}(n) - f^{(r-1)}(m)\big] + R_{r-1}$$

  en sommant sur $k \in [\![m\,;n-1]\!]$ et en multipliant par $\frac{(-1)^{r+1}}{r!}$.
  Finalement, comme $b_r = 0$ si r est impair, on a $(-1)^r b_r = b_r$ et $\mathcal{H}_r$ est vraie.

D'où le résultat par récurrence.

Appliquons la formule à l'ordre $2r$, où $r \in \mathbb{N}^*$, avec la fonction $f : t \longmapsto \frac{1}{t}$ de classe $\mathcal{C}^\infty$ sur $[1\,;n]$ pour un $n \geq 2$. On calcule que $f^{(\ell-1)}(t) = \frac{(-1)^{\ell-1}(\ell-1)!}{t^\ell}$ pour $\ell \in \mathbb{N}^*$ et $t \geq 1$, d'où :

$$\begin{aligned}H_n &= \int_1^n \frac{\mathrm{d}t}{t} + \frac{1}{2}\Big(1+\frac{1}{n}\Big) + \sum_{\ell=2}^{2r} \frac{(-1)^{\ell-1} b_\ell}{\ell}\Big(\frac{1}{n^\ell} - 1\Big) + \frac{(-1)^{2r+1}}{(2r)!}\int_1^n \tilde{B}_{2r}(t)\frac{(-1)^{2r}(2r)!}{t^{2r+1}}\,\mathrm{d}t \\ &= \ln n + \underbrace{\frac{1}{2} + \sum_{\ell=2}^{2r} \frac{(-1)^\ell b_\ell}{\ell} - \int_1^{+\infty} \frac{\tilde{B}_{2r}(t)}{t^{2r+1}}\,\mathrm{d}t}_{c_r} + \frac{1}{2n} + \sum_{\ell=2}^{2r} \frac{(-1)^{\ell-1} b_\ell}{\ell}\frac{1}{n^\ell} + \int_n^{+\infty} \frac{\tilde{B}_{2r}(t)}{t^{2r+1}}\,\mathrm{d}t\,.\end{aligned}$$

Comme d'une part $\big|\int_n^{+\infty} \frac{\tilde{B}_{2r}(t)}{t^{2r+1}}\,\mathrm{d}t\big| \leq \|\tilde{B}_{2r}\|_\infty \cdot \int_n^{+\infty} \frac{\mathrm{d}t}{t^{2r+1}} =_{n\to+\infty} O(n^{-2r})$ puisque $\tilde{B}_{2r}$ est bornée, et d'autre part $c_r = \lim_{n\to+\infty} H_n - \log(n) = \gamma$, il vient

$$H_n \underset{n\to+\infty}{=} \ln n + \gamma + \frac{1}{2n} - \sum_{\ell=2}^{2r} \frac{(-1)^\ell b_\ell}{\ell}\frac{1}{n^\ell} + O\Big(\frac{1}{n^{2r}}\Big)\,.$$

Vu que $b_\ell = 0$ pour ℓ impair, on obtient la formule annoncée.

COMMENTAIRES

Ce développement calculatoire peut mener à des questions sur les propriétés et le calcul des polynômes et nombres de BERNOULLI et sur l'application de la formule à diverses fonctions, comme retrouver, en utilisant $\ln$, la formule de STIRLING avec reste à un ordre donné, ou écrire $\frac{\pi^2}{6}$ en fonction des $(b_n)_{n\in\mathbb{N}^*}$ en considérant la série $\sum_{k\in\mathbb{N}^*} \frac{1}{k^2}$ (voir notamment ce document).

Le calcul des premiers nombres de BERNOULLI donne :

$$H_n \underset{n\to+\infty}{=} \ln n + \gamma + \frac{1}{2n} - \frac{1}{12n^2} + \frac{1}{120n^4} - \frac{1}{252n^6} + \frac{1}{240n^8} - \frac{1}{132n^{10}} + O\Big(\frac{1}{n^{13}}\Big)\,.$$

39 INÉGALITÉ DE HOEFFDING

[Ouv09, Ch10, p132]

ÉNONCÉ

PROPOSITION. [INÉGALITÉ DE HOEFFDING]
Soient $n \in \mathbb{N}^*$ et $(X_i)_{1\leq i\leq n}$ des variables aléatoires réelles, indépendantes, centrées, et telles que, pour tout $i \in [\![1\,;n]\!]$, $|X_i| \leq c_i$ p.s. pour une constante réelle c_i. Alors si $S_n = \sum_{i=1}^n X_i$:

$$\forall t \in \mathbb{R}_+ \qquad \mathbb{P}\big(|S_n| \geq t\big) \leq 2\exp\left(-\frac{t^2}{2\sum_{i=1}^n c_i^2}\right).$$

COROLLAIRE. *Supposons maintenant qu'une suite $(X_i)_{i\in\mathbb{N}}$ infinie satisfasse les mêmes hypothèses. Si de plus il existe $\alpha, \beta > 0$ tels que $\sum_{i=1}^n c_i^2 \leq n^{2\alpha-\beta}$ pour tout $n \in \mathbb{N}^*$, alors*

$$\frac{S_n}{n^\alpha} \xrightarrow[n\to+\infty]{\text{p.s.}} 0\,.$$

DÉVELOPPEMENT

Commençons par l'inégalité de HOEFFDING. On procède en 3 étapes.

1. Soit d'abord X est une variable aléatoire centrée et bornée par 1 p.s.. Montrons que

$$\forall \lambda \in \mathbb{R} \qquad \mathbb{E}\big[\mathrm{e}^{\lambda X}\big] \leq \mathrm{e}^{\frac{\lambda^2}{2}}\,.$$

Soit en effet $\lambda \in \mathbb{R}$. En remarquant que tout élément $x \in [-1\,;1]$ se décompose sous la forme $x = \underbrace{\frac{1-x}{2}}_{\geq 0}\times(-1) + \underbrace{\frac{1+x}{2}}_{\geq 0}\times 1$, on obtient en utilisant la convexité de $\exp(\lambda\,\cdot\,)$ que

$$\forall x \in [-1\,;1] \qquad \mathrm{e}^{\lambda x} \leq \frac{1-x}{2}\,\mathrm{e}^{-\lambda} + \frac{1+x}{2}\,\mathrm{e}^{\lambda}\,.$$

Comme X est centrée, il s'ensuit que

$$\mathbb{E}\big[\mathrm{e}^{\lambda X}\big] \leq \frac{1}{2}\big(\mathrm{e}^{-\lambda}+\mathrm{e}^{\lambda}\big) = \cosh(\lambda) = \sum_{n=0}^{+\infty}\frac{\lambda^{2n}}{(2n)!} \leq \sum_{n=0}^{+\infty}\frac{\lambda^{2n}}{n!\,2^n} = \mathrm{e}^{\frac{\lambda^2}{2}}\,,$$

où l'on a utilisé que $(2n)! \geq n!\,2^n$ pour tout $n \in \mathbb{N}$ (ce qui peut être vérifié par récurrence).

2. Soit $i \in [\![1\,;n]\!]$. Comme $\frac{X_i}{c_i}$ est centrée et bornée par 1 p.s., on obtient

$$\forall \lambda \in \mathbb{R} \qquad \mathbb{E}\big[\mathrm{e}^{\lambda X_i}\big] = \mathbb{E}\left[\mathrm{e}^{\lambda c_i \frac{X_i}{c_i}}\right] \leq \mathrm{e}^{\frac{\lambda^2 c_i^2}{2}}\,.$$

Autrement dit, pour tout $i \in [\![1\,;n]\!]$, X_i est c_i-sous-gaussienne.

3. Fixons $t \in \mathbb{R}_+$. Pour tout $\lambda \geq 0$, on calcule

$$\begin{aligned}
\mathbb{P}(S_n \geq t) = \mathbb{P}\big(\mathrm{e}^{\lambda S_n} \geq \mathrm{e}^{\lambda t}\big) &\leq \mathrm{e}^{-\lambda t}\,\mathbb{E}\Big[\mathrm{e}^{\lambda\sum_{i=1}^n X_i}\Big] && \text{par l'inégalité de MARKOV}\\
&\leq \mathrm{e}^{-\lambda t}\prod_{i=1}^n \mathbb{E}\big[\mathrm{e}^{\lambda X_i}\big] && \text{par indépendance des } (X_i)_{1\leq i\leq n}\\
\mathbb{P}(S_n \geq t) &\leq \mathrm{e}^{-\lambda t}\prod_{i=1}^n \mathrm{e}^{\frac{\lambda^2 c_i^2}{2}} = \exp\left(-\lambda t + \frac{\sum c_i^2}{2}\lambda^2\right) && \text{par le deuxième point.}
\end{aligned}$$

Le choix optimal de $\lambda = \frac{t}{\sum_{i=1}^n c_i^2}$ permet d'aboutir à la borne

$$\mathbb{P}(S_n \geq t) \leq \exp\left(-\frac{t^2}{2\sum_{i=1}^n c_i^2}\right).$$

Les $(-X_i)_{1\leq i\leq n}$ satisfaisant les mêmes hypothèses que les $(X_i)_{1\leq i\leq n}$, il vient de même

$$\mathbb{P}(S_n \leq -t) \leq \exp\left(-\frac{t^2}{2\sum_{i=1}^n c_i^2}\right).$$

Le résultat en découle en écrivant que $\mathbb{P}\big(|S_n| \geq t\big) \leq \mathbb{P}(S_n \geq t) + \mathbb{P}(S_n \leq -t)$.

Passons au corollaire. Fixons $\varepsilon > 0$. D'après l'inégalité de HOEFFDING et l'hypothèse énoncée :

$$\sum_{n=0}^{+\infty}\mathbb{P}\big(|S_n| > n^\alpha\varepsilon\big) \leq 2\sum_{n=0}^{+\infty}\exp\left(-\frac{n^{2\alpha}\varepsilon^2}{2n^{2\alpha-\beta}}\right) = 2\sum_{n=0}^{+\infty}\exp\left(-\frac{\varepsilon^2}{2}n^\beta\right) < +\infty\,,$$

la convergence de la série ayant lieu car $\exp\big(-\frac{\varepsilon^2}{2}n^\beta\big) =_{n\to+\infty} o\big(\frac{1}{n^2}\big)$. Ainsi, le lemme de BOREL-CANTELLI assure que $\mathbb{P}\big(\limsup_{n\to+\infty}\big\{|S_n| > n^\alpha\varepsilon\big\}\big) = 0$. Par dénombrabilité, il en découle

$$\mathbb{P}\left(\bigcup_{k\in\mathbb{N}^*}\limsup_{n\to+\infty}\left\{|S_n| > \frac{n^\alpha}{k}\right\}\right) = 0\,, \qquad \text{et donc} \qquad \mathbb{P}\left(\bigcap_{k\in\mathbb{N}^*}\liminf_{n\to+\infty}\left\{|S_n| \leq \frac{n^\alpha}{k}\right\}\right) = 1\,.$$

En d'autres termes,

$$\mathbb{P}\left(\lim_{n\to+\infty}\frac{|S_n|}{n^\alpha} = 0\right) = 1\,, \qquad \text{soit} \qquad \frac{S_n}{n^\alpha} \xrightarrow[n\to+\infty]{\text{p.s.}} 0\,.$$

COMMENTAIRES

On rappelle qu'en probabilités, si $(A_n)_{n\in\mathbb{N}}$ est une suite d'évènements, $\limsup_{n\to+\infty} A_n$ désigne l'événement « A_n est réalisé pour une infinité de n » tandis que $\liminf_{n\to+\infty} A_n$ est l'événement « A_n est réalisé pour tout n à partir d'un certain rang ».

40 INJECTIVITÉ DE LA FONCTION CARACTÉRISTIQUE ET APPLICATION

[Ouv09, §12.2, p197]

ÉNONCÉ

PROPOSITION. *Soient $d \geq 1$ et X une variable aléatoire à valeurs dans $\mathbb{R}^d$. La fonction caractéristique ϕ_X caractérise la loi $\mathbb{P}_X$.*

APPLICATION. [LOI MULTINOMIALE POISSONNIFIÉE]
Soient $d \geq 1$ et $(p_j)_{1\leq j\leq d}$ des réels positifs tels que $\sum_{j=1}^d p_j = 1$. Soient $(Y^k)_{k\geq 1}$ des variables aléatoires discrètes, i.i.d., telles que $\mathbb{P}(Y^1 = j) = p_j$ pour $j \in [\![1\,;d]\!]$. Soit $N \sim \mathcal{P}(\lambda)$, pour un $\lambda > 0$, indépendante des $(Y^k)_{k\geq 1}$. Posons enfin $X^k = (\mathbb{1}_{Y^k=1}, \dots, \mathbb{1}_{Y^k=d})$ pour $k \in \mathbb{N}^*$.
Alors $S = \sum_{k=1}^N X^k$ suit la loi $\mathcal{P}(\lambda p_1) \otimes \cdots \otimes \mathcal{P}(\lambda p_d)$.

DÉVELOPPEMENT

RAPPEL. Pour $\sigma > 0$, on admet que $\phi_{\mathcal{N}(0,\sigma^2)} : \zeta \in \mathbb{R} \longmapsto \mathrm{e}^{-\frac{\sigma^2\zeta^2}{2}}$.

Faisons la preuve pour $d = 1$.
Soient X_1, X_2 des v.a. réelles telles que $\phi_{X_1} = \phi_{X_2}$. Pour $\sigma > 0$, on pose

$$g_\sigma : x \in \mathbb{R} \longmapsto \frac{1}{\sqrt{2\pi\sigma^2}}\,\mathrm{e}^{-\frac{x^2}{2\sigma^2}} = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{\mathbb{R}} g_{\frac{1}{\sigma}}(t)\,\mathrm{e}^{-\mathrm{i}tx}\,\mathrm{d}t\,,$$

puis, pour $j \in [\![1\,;2]\!]$, $f_{\sigma,j} : x \in \mathbb{R} \longmapsto \int_{\mathbb{R}} g_\sigma(x-y)\,\mathrm{d}\mu_j(y)$, et enfin $\mu_{\sigma,j} = f_{\sigma,j}\,\mathrm{Leb}$.
D'après le théorème de FUBINI-LEBESGUE, puis comme $\phi_{X_1} = \phi_{X_2}$:

$$\begin{aligned}
\forall x \in \mathbb{R} \qquad f_{\sigma,1}(x) &= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{\mathbb{R}} \left(\int_{\mathbb{R}} g_{\frac{1}{\sigma}}(t)\,\mathrm{e}^{-\mathrm{i}t(x-y)}\,\mathrm{d}t \right) \mathrm{d}\mu_1(y) \\
&= \frac{1}{2\pi} \int_{\mathbb{R}} \left(\int_{\mathbb{R}} \mathrm{e}^{\mathrm{i}ty}\,\mathrm{d}\mu_1(y) \right) \mathrm{e}^{-\frac{\sigma^2t^2}{2}}\,\mathrm{e}^{-\mathrm{i}tx}\,\mathrm{d}t \\
&= \frac{1}{2\pi} \int_{\mathbb{R}} \phi_{X_1}(t)\,\mathrm{e}^{-\frac{\sigma^2t^2}{2}}\,\mathrm{e}^{-\mathrm{i}tx}\,\mathrm{d}t \\
\forall x \in \mathbb{R} \qquad f_{\sigma,1}(x) &= \frac{1}{2\pi} \int_{\mathbb{R}} \phi_{X_2}(t)\,\mathrm{e}^{-\frac{\sigma^2t^2}{2}}\,\mathrm{e}^{-\mathrm{i}tx}\,\mathrm{d}t = \cdots = f_{\sigma,2}(x)\,.
\end{aligned}$$

Ainsi $\mu_{\sigma,1} = \mu_{\sigma,2}$. Soit $h \in \mathcal{C}_c(\mathbb{R})$. Pour $j \in [\![1\,;2]\!]$, on a :

$$\begin{aligned}
\left| \int_{\mathbb{R}} h(x)\,\mathrm{d}\mu_{\sigma,j}(x) - \int_{\mathbb{R}} h(y)\,\mathrm{d}\mu_j(y) \right| &= \left| \int_{\mathbb{R}} h(x) \int_{\mathbb{R}} g_\sigma(x-y)\,\mathrm{d}\mu_j(y)\,\mathrm{d}x - \int_{\mathbb{R}} h(y)\,\mathrm{d}\mu_j(y) \right| \\
&\leq \int_{\mathbb{R}} \underbrace{\left(\int_{\mathbb{R}} |h(x)-h(y)|\,g_\sigma(x-y)\,\mathrm{d}x \right)}_{u_\sigma(y)}\,\mathrm{d}\mu_j(y) \underset{\sigma\to 0^+}{\longrightarrow} 0\,,
\end{aligned}$$

l'inégalité provenant du théorème de FUBINI-TONELLI et du fait que $\int_{\mathbb{R}} g_\sigma = 1$, et la limite étant obtenue par convergence dominée puisque

- pour tout $\sigma > 0$, on a la domination $u_\sigma \leq 2\|h\|_\infty \int_{\mathbb{R}} g_\sigma = 2\|h\|_\infty$ intégrable selon μ_j ;
- $u_\sigma \underset{\sigma\to 0^+}{\longrightarrow} 0$ simplement : en effet, étant donné que $(g_\sigma)_{\sigma\to 0^+}$ est une approximation de l'unité, on peut écrire, pour $y \in \mathbb{R}$, que $u_\sigma(y) = \ell_y \star g_\sigma(0) \underset{\sigma\to 0^+}{\longrightarrow} \ell_y(0) = 0$, où $\ell_y = \big|h(y+\,\cdot\,) - h(y)\big|$ est continue à support compact.

On a ainsi obtenu que

$$\forall h \in \mathcal{C}_c(\mathbb{R}) \qquad \mathbb{E}\big[h(X_1)\big] = \lim_{\sigma\to 0} \int_{\mathbb{R}} h(x) f_{\sigma,1}(x)\,\mathrm{d}x = \lim_{\sigma\to 0} \int_{\mathbb{R}} h(x) f_{\sigma,2}(x)\,\mathrm{d}x = \mathbb{E}\big[h(X_2)\big]\,.$$

Ceci assure que X_1 et X_2 ont même loi, c'est-à-dire $\mathbb{P}_{X_1} = \mathbb{P}_{X_2}$.
Si $d \geq 2$, on procède de même avec $g^{(d)} : (x_1, \dots, x_d) \in \mathbb{R}^d \longmapsto \prod_{i=1}^d g_\sigma(x_i)$.

Passons à l'application. Soit $t = (t_1, \dots, t_d) \in \mathbb{R}^d$. On calcule :

$$\begin{aligned}
\phi_S(t) &= \mathbb{E}\Big[\mathrm{e}^{\mathrm{i}t\cdot\sum_{k=1}^N X^k}\Big] = \mathbb{E}\Big[\sum_{p=0}^{+\infty} \mathrm{e}^{\mathrm{i}t\cdot\sum_{k=1}^p X^k}\,\mathbb{1}_{N=p}\Big] \\
&= \sum_{p=0}^{+\infty} \mathbb{E}\Big[\mathrm{e}^{\mathrm{i}\sum_{k=1}^p t\cdot X^k}\,\mathbb{1}_{N=p}\Big] && \text{par théorème de FUBINI-LEBESGUE} \\
&= \sum_{p=0}^{+\infty} \prod_{k=1}^p \mathbb{E}\big[\mathrm{e}^{\mathrm{i}t\cdot X^k}\big]\mathbb{P}(N=p) && \text{par indépendance} \\
\phi_S(t) &= \sum_{p=0}^{+\infty} \mathbb{E}\big[\mathrm{e}^{\mathrm{i}t\cdot X^1}\big]^p\,\mathbb{P}(N=p) = g_N\Big(\mathbb{E}\big[\mathrm{e}^{\mathrm{i}t\cdot X^1}\big]\Big) && \text{puisque les } (X^k)_{k\in\mathbb{N}^*} \text{ sont i.i.d..}
\end{aligned}$$

Or, $\mathbb{E}\big[\mathrm{e}^{\mathrm{i}t\cdot X^1}\big] = \mathbb{E}\Big[\mathrm{e}^{\mathrm{i}\sum_{j=1}^d t_j \mathbb{1}_{Y^1=j}}\Big] = \sum_{j=1}^d p_j\,\mathrm{e}^{\mathrm{i}t_j}$, d'où, en utilisant que $\sum_{j=1}^d p_j = 1$:

$$\phi_S(t) = \mathrm{e}^{\lambda\left(\sum_{j=1}^d p_j\,\mathrm{e}^{\mathrm{i}t_j} - 1\right)} = \mathrm{e}^{\lambda\sum_{j=1}^d p_j(\mathrm{e}^{\mathrm{i}t_j}-1)} = \prod_{j=1}^d \mathrm{e}^{\lambda p_j(\mathrm{e}^{\mathrm{i}t_j}-1)} = \prod_{j=1}^d \phi_{\mathcal{P}(\lambda p_j)}(t_j)\,.$$

Par injectivité de la fonction caractéristique, on obtient que $S \sim \mathcal{P}(\lambda p_1) \otimes \cdots \otimes \mathcal{P}(\lambda p_d)$.

COMMENTAIRES

Il faut s'attendre à des questions sur les arguments supplémentaires en dimension quelconque.

Attention, ce développement technique diffère beaucoup de la référence. Il faut aussi être vigilant avec l'ordre des propriétés dans le plan : l'application utilise la fonction génératrice d'une loi de POISSON ainsi que la caractérisation de l'indépendance par la fonction caractéristique.

41 INTÉGRALE DE DIRICHLET

[Far00, §VII.2, p98] [FGN14, §4.27, p214]

ÉNONCÉ

APPLICATION. $\displaystyle\int_0^{+\infty} \frac{\sin(x)}{x}\,\mathrm{d}x = \frac{\pi}{2}$.

DÉVELOPPEMENT

Définissons les applications

$$f : \begin{array}{ccc} {\mathbb{R}_+}^2 & \longrightarrow & \mathbb{R} \\ (t,x) & \longmapsto & \begin{cases} \frac{\sin(x)}{x}\,\mathrm{e}^{-tx} & \text{si } x>0 \\ 1 & \text{sinon} \end{cases} \end{array} \quad \text{puis} \quad F : \begin{array}{ccc} \mathbb{R}_+ & \longrightarrow & \mathbb{R} \\ t & \longmapsto & \displaystyle\int_0^{+\infty} f(t,x)\,\mathrm{d}x\,. \end{array}$$

Remarquons que $F(t)$ est bien défini pour tout $t \in \mathbb{R}_+^*$ puisque $f(t,\,\cdot\,)$ est continue sur $\mathbb{R}_+$ et intégrable comme $\big|f(t,x)\big| =_{x\to+\infty} O(\mathrm{e}^{-tx})$.

Vérifions que l'intégrale de DIRICHLET $F(0)$ est semi-convergente. En intégrant par parties :

$$\forall A \in \mathbb{R}_+ \qquad \int_0^A \frac{\sin(x)}{x}\,\mathrm{d}x = \left[\frac{1-\cos(x)}{x}\right]_0^A + \int_0^A \frac{1-\cos(x)}{x^2}\,\mathrm{d}x\,.$$

Faisant tendre A vers $+\infty$, le crochet tend vers 0 tandis que l'intégrale de droite converge du fait que l'intégrande continue soit un $O\big(\frac{1}{x^2}\big)$ lorsque $x \to +\infty$. Ainsi, $F(0)$ est bien définie.

Étudions l'application F afin de calculer l'intégrale de DIRICHLET.

1. Montrons que F est continue et dérivable sur $\mathbb{R}_+^*$. L'application f vérifie :
   - pour tout $t>0$, $f(t,\,\cdot\,)$ est intégrable (comme vu précédemment),
   - pour tout $x \geq 0$, $f(\,\cdot\,,x)$ est dérivable, de dérivée $\partial_t f(\,\cdot\,,x) = -\sin(x)\,\mathrm{e}^{-\,\cdot\,x}$,
   - si $\alpha>0$, alors on a la domination intégrable suivante pour tout $t>\alpha$:

   $$\forall x \in \mathbb{R}_+ \qquad \big|\partial_t f(t,x)\big| \leq \big|\sin(x)\big|\,\mathrm{e}^{-\alpha x} \leq \mathrm{e}^{-\alpha x}\,.$$

   D'après le théorème de dérivabilité sous le signe intégral, il en découle que F est continue et dérivable sur $\mathbb{R}_+^*$, et pour tout $t>0$:

   $$\begin{aligned} F'(t) &= \int_0^{+\infty} \partial_t f(t,x)\,\mathrm{d}x = -\int_0^{+\infty} \sin(x)\,\mathrm{e}^{-tx}\,\mathrm{d}x = -\operatorname{Im}\left(\int_0^{+\infty} \mathrm{e}^{(\mathrm{i}-t)x}\,dx\right) \\ &= -\operatorname{Im}\left(\left[\frac{\mathrm{e}^{(\mathrm{i}-t)x}}{\mathrm{i}-t}\right]_0^{+\infty}\right) = \operatorname{Im}\left(\frac{1}{\mathrm{i}-t}\right) = -\frac{1}{1+t^2}\,. \end{aligned}$$

2. Il découle de l'expression de F' et de la continuité de F sur $\mathbb{R}_+^*$ que

   $$\exists C \in \mathbb{R} \quad \forall t \in \mathbb{R}_+^* \qquad F(t) = \arctan(t) + C\,.$$

   Pour déterminer C, calculons la limite de F en $+\infty$. D'après le théorème de convergence dominée (appliqué avec la domination $|f(t,\,\cdot\,)| \leq \mathrm{e}^{-\alpha\,\cdot}$ pour tout $t \geq \alpha$ où $\alpha>0$), on obtient, du fait que $f(t,\,\cdot\,) \longrightarrow_{t\to+\infty} 0$ simplement, que $\lim_{t\to+\infty} F(t) = 0$. Ainsi, $C = \frac{\pi}{2}$.
3. Pour conclure il suffit de montrer que F est continue en 0, puisqu'on aura alors

   $$\int_0^{+\infty} \frac{\sin(x)}{x}\,\mathrm{d}x = F(0) = -\arctan(0) + \frac{\pi}{2} = \frac{\pi}{2}\,.$$

   Soient $t \in \mathbb{R}_+^*$ et $A \in \mathbb{R}_+^*$. On a :

   $$\begin{aligned} \big|F(t)-F(0)\big| &= \left|F(t) \pm \int_0^A f(t,x)\,\mathrm{d}x \pm \int_0^A f(0,x)\,\mathrm{d}x - F(0)\right| \\ &\leq \left|\int_A^{+\infty} \mathrm{e}^{-tx}\,\frac{\sin(x)}{x}\,\mathrm{d}x\right| + \left|\int_0^A (\mathrm{e}^{-tx}-1)\,\frac{\sin(x)}{x}\,\mathrm{d}x\right| + \left|\int_A^{+\infty} \frac{\sin(x)}{x}\,\mathrm{d}x\right|. \end{aligned}$$

   Majorons la première intégrale. Pour $B>A$, on a :

   $$\begin{aligned} \int_A^B \mathrm{e}^{-tx}\,\frac{\sin(x)}{x}\,\mathrm{d}x &= \operatorname{Im}\left(\int_A^B \frac{\mathrm{e}^{(\mathrm{i}-t)x}}{x}\,\mathrm{d}x\right) = \operatorname{Im}\left(\left[\frac{\mathrm{e}^{(\mathrm{i}-t)x}}{(\mathrm{i}-t)x}\right]_A^B + \int_A^B \frac{\mathrm{e}^{(\mathrm{i}-t)x}}{(\mathrm{i}-t)x^2}\,\mathrm{d}x\right) \\ &= \operatorname{Im}\left(\frac{\mathrm{e}^{(\mathrm{i}-t)B}}{(\mathrm{i}-t)B} - \frac{\mathrm{e}^{(\mathrm{i}-t)A}}{(\mathrm{i}-t)A} + \int_A^B \frac{\mathrm{e}^{(\mathrm{i}-t)x}}{(\mathrm{i}-t)x^2}\,\mathrm{d}x\right), \end{aligned}$$

   d'où, en passant à la limite lorsque B tend vers $+\infty$:

   $$\begin{aligned} \left|\int_A^{+\infty} \mathrm{e}^{-tx}\,\frac{\sin(x)}{x}\,\mathrm{d}x\right| &\leq \left|\operatorname{Im}\left(-\frac{\mathrm{e}^{(\mathrm{i}-t)A}}{(\mathrm{i}-t)A} + \int_A^{+\infty} \frac{\mathrm{e}^{\mathrm{i}x-tx}}{(\mathrm{i}-t)x^2}\,\mathrm{d}x\right)\right| \\ &\leq \frac{1}{|\mathrm{i}-t|A} + \frac{1}{|\mathrm{i}-t|}\int_A^{+\infty} \frac{1}{x^2}\,\mathrm{d}x \leq \frac{2}{A} \qquad \text{car } |\mathrm{i}-t| \geq 1\,. \end{aligned}$$

   Soit $\varepsilon>0$. Choisissons $A \in \mathbb{R}_+^*$ tel que $\frac{2}{A} \leq \frac{\varepsilon}{3}$ et $\big|\int_A^{+\infty} \frac{\sin(x)}{x}\,\mathrm{d}x\big| \leq \frac{\varepsilon}{3}$. Par ce qui précède :

   $$\forall t \in \mathbb{R}_+^* \qquad \big|F(t)-F(0)\big| \leq \frac{2\varepsilon}{3} + \left|\int_0^A (\mathrm{e}^{-tx}-1)\,\frac{\sin(x)}{x}\,\mathrm{d}x\right|.$$

   Par convergence dominée, cette dernière intégrale tend vers 0 lorsque $t \to 0$, et ainsi $\big|F(t)-F(0)\big| \leq \varepsilon$ pour t assez petit.

   Finalement, F est continue en 0, et on a bien obtenu que

   $$\int_0^{+\infty} \frac{\sin(x)}{x}\,\mathrm{d}x = \frac{\pi}{2}\,.$$

42 LEMME DE MORSE

[Rou99, §5.5, p327]

ÉNONCÉ

THÉORÈME. [LEMME DE MORSE]
Soit $f : U \longrightarrow \mathbb{R}$ une fonction de classe $\mathcal{C}^3$ définie sur un ouvert U de $\mathbb{R}^n$ contenant 0. On suppose que $df(0) = 0$ et $d^2f(0)$ est non dégénérée, de signature $(p, n-p)$.
Alors il existe un $\mathcal{C}^1$-difféomorphisme φ entre deux voisinages de l'origine de $\mathbb{R}^n$ tel que $\varphi(0) = 0$ et, au voisinage de 0,

$$f(x) - f(0) = \varphi_1(x)^2 + \cdots + \varphi_p(x)^2 - \varphi_{p+1}(x)^2 - \cdots - \varphi_n(x)^2 .$$

DÉVELOPPEMENT

La fonction f est de classe $\mathcal{C}^3$. La formule de TAYLOR avec reste intégral donne, pour $x \in U$:

$$f(x) - f(0) = \int_0^1 (1-t) d^2f(tx)(x,x)\,\mathrm{d}t = x^\top \int_0^1 (1-t) d^2f(tx)\,\mathrm{d}t\, x = x^\top Q(x)\, x ,$$

où $Q : x \in U \longmapsto \int_0^1 (1-t) d^2f(tx)\,\mathrm{d}t$. Par dérivation sous le signe intégral, on a que Q est de classe $\mathcal{C}^1$. Par ailleurs, $Q(0) = \frac{1}{2} D^2f(0)$ est symétrique inversible de signature $(p, n-p)$.

LEMME. *Soit $A_0 \in \mathcal{S}_n(\mathbb{R}) \cap \mathcal{GL}_n(\mathbb{R})$. Alors il existe un voisinage $V \in \mathcal{V}(A_0) \cap \mathcal{S}_n(\mathbb{R})$ et une application $g \in \mathcal{C}^1\big(V, \mathcal{GL}_n(\mathbb{R})\big)$ telles que*

$$\forall A \in V \qquad A = g(A)^\top A_0 g(A) .$$

En effet, soit

$$\varphi : \mathcal{M}_n(\mathbb{R}) \longrightarrow \mathcal{S}_n(\mathbb{R}), M \longmapsto M^\top A_0 M .$$

On a que φ est différentiable et

$$\forall M \in \mathcal{M}_n(\mathbb{R}) \quad \forall H \in \mathcal{M}_n(\mathbb{R}) \qquad d\varphi(M)(H) = H^\top A_0 M + M^\top A_0 H ,$$

d'où, en particulier,

$$\forall H \in \mathcal{M}_n(\mathbb{R}) \qquad d\varphi(I_n)(H) = H^\top A_0 + A_0 H = (A_0 H)^\top + A_0 H .$$

On a donc $\ker(d\varphi(I_n)) = A_0^{-1}\mathcal{A}_n(\mathbb{R})$ où $\mathcal{A}_n(\mathbb{R})$ est l'ensemble des matrices antisymétriques. De plus $d\varphi(I_n)$ est surjective dans $\mathcal{S}_n(\mathbb{R})$ puisque

$$\forall A \in \mathcal{S}_n(\mathbb{R}) \qquad d\varphi(I_n)\Big(\frac{1}{2} A_0^{-1} A\Big) = A .$$

Soit $F = A_0^{-1}\mathcal{S}_n(\mathbb{R})$. Notons que $F \oplus \ker\big(d\varphi(I_n)\big) = \mathcal{M}_n(\mathbb{R})$ et $I_n \in F$. Si $\psi = \varphi_{|F}$, on a que $d\psi(I_n)$ est bijective. Le théorème d'inversion locale donne que ψ est un $\mathcal{C}^1$-difféomorphisme local d'un voisinage W de I_n dans[1] $\mathcal{GL}_n(\mathbb{R})$ sur un voisinage V de $A_0 = \psi(I_n)$ dans $\mathcal{S}_n(\mathbb{R})$.

Pour $A \in V$, il existe donc une unique matrice inversible $M \in W$ telle que $A = M^\top A_0 M$. On a que $M = \psi^{-1}(A)$, et donc $g = \psi^{-1}$ convient.

Revenons à la preuve du lemme de MORSE.
Appliquons le résultat du lemme précédent à $Q(0) \in \mathcal{S}_n(\mathbb{R})$.
Si $M(x) = g\big(Q(x)\big)$ pour $x \in U$, on a, au voisinage de 0,

$$Q(x) = \big(M(x)\big)^\top Q(0)\, M(x) ,$$

$$\text{et donc} \quad f(x) - f(0) = x^\top Q(x)\, x = \big(M(x)\, x\big)^\top Q(0) \underbrace{M(x)\, x}_{=y} = y^\top Q(0)\, y .$$

Comme $Q(0) = \frac{1}{2} D^2f(0)$ est de signature $(p, n-p)$, il existe, par classification des formes quadratiques, une matrice $A \in \mathcal{GL}_n(\mathbb{R})$ telle que

$$A^\top Q(0) A = \mathrm{diag}(\underbrace{1, \ldots, 1}_{p \text{ termes}}, \underbrace{-1, \ldots, -1}_{n-p \text{ termes}}) ,$$

et le changement linéaire de coordonnées $y = Au$ donne alors :

$$y^\top Q(0)\, y = u^\top A^\top Q(0)\, A\, u = u_1^2 + \cdots + u_p^2 - u_{p+1}^2 - \cdots - u_n^2 .$$

Posons donc $\varphi : x \in U \longmapsto A^{-1} M(x)\, x$. L'application φ est de classe $\mathcal{C}^1$ et on a :

$$\forall h \in \mathbb{R}^n \qquad d\varphi(0)(h) = A^{-1}\big(dM(0)(h) \times 0 + M(0) \times h\big) = A^{-1} M(0)(h) .$$

Donc $d\varphi(0) = A^{-1} M(0)$ est inversible.

Par le théorème d'inversion locale, on a que φ est un $\mathcal{C}^1$-difféomorphisme local entre deux voisinages de 0 car $\varphi(0) = 0$, ce qui conclut puisqu'alors dans ce voisinage de 0 :

$$f(x) - f(0) = \varphi_1(x)^2 + \cdots + \varphi_p(x)^2 - \varphi_{p+1}(x)^2 - \cdots - \varphi_n(x)^2 .$$

COMMENTAIRES

Que se passe-t-il sans l'hypothèse que f est $\mathcal{C}^3$? On n'a plus l'assurance que φ est $\mathcal{C}^1$!
En effet, si f est de classe $\mathcal{C}^3$, alors M est $\mathcal{C}^1$ et on calcule :

$$\forall x \in U \quad \forall h \in \mathbb{R}^n \qquad d\varphi(x)(h) = A^{-1}\big(dM(x)(h) \times 0 + M(x) \times h\big) ,$$

et l'on voit que $x \longmapsto d\varphi(x)$ ne serait pas continue si M n'était pas $\mathcal{C}^1$.

1. quitte à restreindre, $\mathcal{GL}_n(\mathbb{R})$ étant ouvert

43 MÉTHODE DE NEWTON

[Rou99, Ch4, p140]

ÉNONCÉ

THÉORÈME. [MÉTHODE DE NEWTON]
Soient deux réels $c < d$ et une fonction $f : [c\,;d] \longrightarrow \mathbb{R}$ de classe $\mathcal{C}^2$, telle que $f(c) < 0 < f(d)$ et $f' > 0$ sur $[c\,;d]$. On considère la suite récurrente définie, lorsque c'est possible, par $x_0 \in [c\,;d]$ puis

$$\forall n \in \mathbb{N} \qquad x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}\,.$$

Alors en notant a l'unique 0 de f, on a :

(i) il existe un réel $\alpha > 0$ tel que $[a \pm \alpha] \subset [c\,;d]$ et, pour $x_0 \in [a \pm \alpha]$, la suite $(x_n)_{n\in\mathbb{N}}$ est bien définie et converge vers a de manière quadratique :

$$\exists C \in \mathbb{R}_+^* \quad \forall n \in \mathbb{N} \qquad |x_{n+1} - a| \leq C \cdot |x_n - a|^2\,.$$

(ii) si de plus $f'' > 0$ sur $[a\,;d]$, alors pour $x_0 \in]a\,;d]$, la suite $(x_n)_{n\in\mathbb{N}}$ décroît strictement et

$$x_{n+1} - a \underset{n\to+\infty}{\sim} \frac{1}{2}\frac{f''(a)}{f'(a)}(x_n - a)^2\,.$$

DÉVELOPPEMENT

La définition de a découle de la stricte croissance de f et du théorème des valeurs intermédiaires. Dans la suite, soit

$$\begin{array}{rccc} \varphi : & [c\,;d] & \longrightarrow & \mathbb{R} \\ & x & \longmapsto & x - \frac{f(x)}{f'(x)}\,, \end{array}$$

qui est de classe $\mathcal{C}^1$ vérifie $\varphi(a) = a$ et $\varphi'(a) = 0$.

(i) Soit $x \in [c\,;d]$. Écrivons

$$\varphi(x) - a = x - a - \frac{f(x)}{f'(x)} = x - a - \frac{f(x) - f(a)}{f'(x)} = \frac{f(a) - f(x) - (a-x)f'(x)}{f'(x)}$$

En appliquant la formule de TAYLOR-LAGRANGE à l'ordre 2 à f, il vient[1]

$$\exists z_x \in [c\,;d] \qquad f(a) = f(x) + f'(x)(a-x) + \frac{f''(z_x)}{2}(a-x)^2\,,$$

et ainsi

$$\varphi(x) - a = \frac{1}{2}\frac{f''(z_x)}{f'(x)}(x-a)^2\,.$$

1. en fait z_x est entre x et a, mais cela ne sera pas utile

On a obtenu que

$$\forall x \in [c\,;d] \quad \big|\varphi(x) - a\big| \leq C(x-a)^2 \qquad \text{où} \qquad C = \frac{1}{2}\frac{\max_{[c,d]} f''}{\min_{[c,d]} f''} > 0\,.$$

Choisissant $\alpha > 0$ assez petit de sorte que $I_\alpha = [a \pm \alpha] \subset [c\,;d]$ et $\alpha < \frac{1}{C}$, il vient que I_α est φ-stable puisque

$$\forall x \in I_\alpha \qquad \big|\varphi(x) - a\big| \leq C\alpha^2 \leq \alpha\,.$$

Ainsi, si $x_0 \in I_\alpha$, la suite $(x_n)_{n\in\mathbb{N}}$ est bien définie et à valeurs dans I_α, d'où

$$\forall n \in \mathbb{N} \qquad |x_{n+1} - a| \leq C \cdot |x_n - a|^2\,.$$

Par récurrence, il en découle que, pour tout $n \in \mathbb{N}$,

$$C \cdot |x_n - a| \leq \big(C \cdot |x_0 - a|\big)^{2^n} \leq (C\alpha)^{2^n} \underset{n\to+\infty}{\longrightarrow} 0$$

puisque $\alpha < \frac{1}{C}$. On a donc obtenu la convergence quadratique de $(x_n)_{n\in\mathbb{N}}$ vers a.

(ii) Si $x \in]a\,;d]$, on a $f(x) > 0$, donc $\varphi(x) < x$, et par ce qui précède :

$$\varphi(x) - a = \frac{1}{2}\frac{f''(z_x)}{f'(x)}(x-a)^2 > 0\,.$$

Ainsi $I_2 =]a\,;d]$ est φ-stable et, pour $x_0 \in I_2$, la suite $(x_n - a)_{n\in\mathbb{N}}$ est décroissante positive, donc converge vers une limite $\ell \in [a\,;d]$ satisfaisant $\varphi(\ell) = \ell$, c'est-à-dire $f(\ell) = 0$, soit donc $\ell = a$. Par stabilité de I_2, ce que l'on a vu précédemment s'applique et

$$\forall n \in \mathbb{N} \qquad 0 \leq x_{n+1} - a \leq C(x_n - a)^2\,.$$

Comme $x_n \in]a\,;d]$ pour tout $n \in \mathbb{N}$, il vient alors

$$\frac{x_{n+1} - a}{(x_n - a)^2} = \frac{1}{2}\frac{f''(z_n)}{f'(x_n)} \underset{n\to+\infty}{\longrightarrow} \frac{1}{2}\frac{f''(a)}{f'(a)}\,,$$

ce qui donne l'équivalent annoncé.

COMMENTAIRES

On peut faire des graphiques pour expliquer les deux cas : pour le premier, avec une fonction qui nécessite x_0 très proche de I_α pour que la méthode de NEWTON converge, et pour le second, avec une fonction pour laquelle il n'y a pas de contrainte sur x_0 du fait que $f'' > 0$.

On peut aussi lire [Dem96, §IV.2.3, p98–100].

44 MÉTHODE DU GRADIENT À PAS OPTIMAL

[FGN12, §1.21, p39–41]

ÉNONCÉ

THÉORÈME. [MÉTHODE DE GRADIENT À PAS OPTIMAL]
Soit $p \in \mathbb{N}^*$. Soit $f : \mathbb{R}^p \longrightarrow \mathbb{R}$ elliptique, c'est-à-dire une application de classe $\mathcal{C}^1$ telle que

$$\exists \alpha \in \mathbb{R}_+ \quad \forall (x,y) \in \left(\mathbb{R}^p\right)^2 \qquad \big\langle \nabla f(x) - \nabla f(y) \mid x - y \big\rangle \geq \alpha \cdot \|x-y\|^2 .$$

Alors la suite $(x_n)_{n\in\mathbb{N}}$ définie par $x_0 \in \mathbb{R}^p$ et la relation de récurrence

$$\forall n \in \mathbb{N} \qquad x_{n+1} = x_n - \rho_n \nabla f(x_n) \quad \text{où} \quad \rho_n = \operatorname{argmin}_{\rho>0} f\big(x_n - \rho \nabla f(x_n)\big) ,$$

converge vers l'unique minimum global de f.

DÉVELOPPEMENT

Soit f satisfaisant les hypothèses du théorème. On procède en plusieurs étapes.

1. Montrons d'abord que

$$\forall (x,y) \in \left(\mathbb{R}^p\right)^2 \qquad f(y) - f(x) \geq \big\langle \nabla f(x) \mid y - x \big\rangle + \frac{\alpha}{2} \cdot \|x-y\|^2 .$$

En effet, soient $x, y \in \mathbb{R}^p$. D'après la formule de TAYLOR avec reste intégral :

$$\begin{aligned} f(y) - f(x) &= \int_0^1 \big\langle \nabla f\big(x + t(y-x)\big) \mid y - x \big\rangle \, \mathrm{d}t \\ &= \int_0^1 \big\langle \nabla f(x) \mid y - x \big\rangle \, \mathrm{d}t + \int_0^1 \big\langle \nabla f\big(x + t(y-x)\big) - \nabla f(x) \mid y - x \big\rangle \, \mathrm{d}t \\ &= \big\langle \nabla f(x) \mid y - x \big\rangle + \int_0^1 \frac{1}{t} \cdot \big\langle \nabla f\big(x + t(y-x)\big) - \nabla f(x) \mid t(y-x) \big\rangle \, \mathrm{d}t \\ &\geq \big\langle \nabla f(x) \mid y - x \big\rangle + \int_0^1 \frac{\alpha}{t} \cdot \|t(y-x)\|^2 \, \mathrm{d}t \\ f(y) - f(x) &\geq \big\langle \nabla f(x) \mid y - x \big\rangle + \frac{\alpha}{2} \cdot \|y-x\|^2 . \end{aligned}$$

2. L'application f est continue et coercive puisque, pour $x \in \mathbb{R}^p$,

$$f(x) \geq f(0) + \big\langle \nabla f(0) \mid x \big\rangle + \frac{\alpha}{2} \cdot \|x\|^2 \underset{\|x\|\to+\infty}{=} O\big(\|x\|\big) + \frac{\alpha}{2} \cdot \|x\|^2 \underset{\|x\|\to+\infty}{\longrightarrow} +\infty .$$

Ainsi f atteint son minimum global en un point $x_\star$ vérifiant $\nabla f(x_\star) = 0$.
Si $x_\star^1, x_\star^2 \in \mathbb{R}^p$ sont deux minimums locaux de f, alors $0 \geq \langle 0 \mid x_\star^1 - x_\star^2 \rangle + \frac{\alpha}{2} \cdot \|x_\star^1 - x_\star^1\|^2$, ce qui implique $x_\star^1 = x_\star^2$. Ainsi, $x_\star$ est le seul minimum global et local de f.

Par ailleurs, si $x \in \mathbb{R}^p$ est tel que $\nabla f(x) = 0$, alors

$$0 \geq f(x_\star) - f(x) \geq \frac{\alpha}{2} \cdot \|x_\star - x\|^2 \geq 0 , \qquad \text{et donc} \qquad x = x_\star .$$

3. Soit $x \in \mathbb{R}^p$ tel que $\nabla f(x) \neq 0$. Notons qu'alors $x \neq x_\star$. Soit la fonction de classe $\mathcal{C}^1$

$$\begin{array}{rccc} \varphi_x : & \mathbb{R}_+ & \longrightarrow & \mathbb{R} \\ & \rho & \longmapsto & f\big(x - \rho \nabla f(x)\big) , \end{array}$$

qui tend vers $+\infty$ en $+\infty$. Son minimum est donc atteint en un point ρ_x vérifiant :

$$0 = \varphi_x'(\rho_x) = -\Big\langle \nabla f\big(x + \rho_x \nabla f(x)\big) \mid \nabla f(x) \Big\rangle .$$

Donc, pour tout $\rho \neq \rho_x$:

$$\varphi_x(\rho) - \varphi_x(\rho_x) \geq \underbrace{\Big\langle \nabla f\big(x + \rho_x \nabla f(x)\big) \mid (\rho - \rho_x) \nabla f(x) \Big\rangle}_{=0} + \frac{\alpha}{2} \cdot \big\|(\rho - \rho_x) \nabla f(x)\big\|^2 > 0 .$$

On en déduit que ρ_x est l'unique minimum de φ_x.

4. La suite $(x_n)_{n\in\mathbb{N}}$ est ainsi bien définie, stationnaire en $x_\star$ si elle l'atteint. Supposons qu'elle n'atteigne pas $x_\star$, et montrons que $(x_n)_{n\in\mathbb{N}}$ converge tout de même vers $x_\star$.
La suite $\big(f(x_n)\big)_{n\in\mathbb{N}}$ est strictement décroissante et minorée, donc converge.
Soit $n \in \mathbb{N}$. Par ce qui a été démontré à l'étape précédente, remarquons que $\nabla f(x_n)$ et $\nabla f(x_{n+1})$ sont orthogonaux, et comme $\nabla f(x_n)$ et $x_{n+1} - x_n$ sont colinéaires, on obtient :

$$f(x_n) - f(x_{n+1}) \geq \frac{\alpha}{2} \cdot \|x_{n+1} - x_n\|^2 .$$

Il en découle que $\|x_{n+1} - x_n\| \underset{n\to+\infty}{\longrightarrow} 0$.

Or, par coercivité de f, la suite $(x_n)_{n\in\mathbb{N}}$ vit dans un compact de $\mathbb{R}^p$, donc admet une valeur d'adhérence x. Soit ψ une extractrice telle que $x_{\psi(n)} \longrightarrow_{n\to+\infty} x$. Par continuité de ∇f, et puisque $(x_{\psi(n)+1})_{n\in\mathbb{N}}$ converge aussi vers x puisque $\|x_{n+1} - x_n\| \longrightarrow_{n\to+\infty} 0$, on a

$$\nabla f(x_{\psi(n)}) \underset{n\to+\infty}{\longrightarrow} \nabla f(x) \qquad \text{et} \qquad \nabla f(x_{\psi(n)+1}) \underset{n\to+\infty}{\longrightarrow} \nabla f(x) .$$

D'où

$$0 = \big\langle \nabla f(x_{\psi(n)}) \mid \nabla f(x_{\psi(n)+1}) \big\rangle \underset{n\to+\infty}{\longrightarrow} \big\|\nabla f(x)\big\|^2 ,$$

de sorte que nécessairement $\nabla f(x) = 0$ et donc $x = x_\star$.

COMMENTAIRES

Bien penser à faire un dessin en annexe de la leçon concernée.

Notons que ce résultat est vérifié quelque soit le produit scalaire munissant $\mathbb{R}^p$.

45 PROCESSUS DE BRANCHEMENT DE GALTON-WATSON

[CDGM16, ChIII, p68–72]

ÉNONCÉ

PROPOSITION. *Soient $(X_i^j)_{i,j\in\mathbb{N}^*}$ des variables aléatoires i.i.d. à valeurs dans $\mathbb{N}$, de loi μ et d'espérance m. On définit le processus $(Z_n)_{n\in\mathbb{N}}$ par $Z_0 = 1$ et la relation $Z_{n+1} = \sum_{i=1}^{Z_n} X_i^{n+1}$ pour $n \in \mathbb{N}$, puis on s'intéresse à $\rho = \mathbb{P}(\exists n \in \mathbb{N} \quad Z_n = 0)$. Si $\mu \neq \delta_1$, on a :*

- *soit $m \leq 1$, et alors $\rho = 1$ et il a extinction du processus presque sûrement,*
- *soit $m > 1$, et alors $\rho < 1$ et il y a une probabilité strictement positive de survie.*

DÉVELOPPEMENT

1. Intéressons-nous d'abord aux propriétés de la fonction génératrice g_μ.
   Posons $p_k = \mu(\{k\})$ pour $k \in \mathbb{N}$. Les $(p_k)_{k\in\mathbb{N}}$ étant positifs, g_μ est croissante sur $[0\,;1]$. Comme g_μ est une série entière convergente en 1, elle est de classe $\mathcal{C}^\infty$ sur $[0\,;1[$ et

   $$\forall s \in [0\,;1[\qquad g''(s) = \sum_{k=2}^{+\infty} k(k-1)p_k\, s^{k-2} \geq 0\,,$$

   ainsi g_μ est convexe sur $[0\,;1]$. De plus, cette inégalité est stricte pour $s \neq 0$ dès que $p_k > 0$ pour un $k \geq 2$, et donc g_μ est strictement convexe sur $[0\,;1]$ si $\mu([2\,;+\infty[) > 0$.

2. Montrons par récurrence que $g_{Z_n} = g_\mu^{\circ n}$ sur $[-1\,;1]$ pour tout entier $n \in \mathbb{N}$.
   - Initialisation. On a $g_{Z_0}(s) = s^1 = s$ pour $s \in [-1\,;1]$ donc $g_{Z_0} = \mathrm{id}$.
   - Hérédité. Supposons que $g_{Z_n} = g_\mu^{\circ n}$ pour un $n \in \mathbb{N}$. Alors pour tout $s \in [-1\,;1]$:

   $$\begin{aligned}
   g_{Z_{n+1}}(s) &= \mathbb{E}\left[s^{\sum_{i=1}^{Z_n} X_i^{n+1}}\right] \\
   &= \mathbb{E}\left[\sum_{z=0}^{+\infty} s^{\sum_{i=1}^{z} X_i^{n+1}} \mathbb{1}_{Z_n=z}\right] \\
   &= \sum_{z=0}^{+\infty} \mathbb{E}\left[s^{\sum_{i=1}^{z} X_i^{n+1}} \mathbb{1}_{Z_n=z}\right] && \text{par le thm. de FUBINI-LEBESGUE} \\
   &= \sum_{z=0}^{+\infty} \left(\prod_{i=1}^{z} \mathbb{E}\left[s^{X_i^{n+1}}\right]\right) \mathbb{P}(Z_n = z) && \text{par indépendance} \\
   &= \sum_{z=0}^{+\infty} g_\mu(s)^z\, \mathbb{P}(Z_n = z) \\
   &= \mathbb{E}\left[g_\mu(s)^{Z_n}\right] \\
   g_{Z_{n+1}}(s) &= g_{Z_n} \circ g_\mu(s) = g_\mu^{\circ n+1}(s) && \text{par hypothèse de récurrence}
   \end{aligned}$$

3. Puisque $(Z_n = 0)_{n\in\mathbb{N}}$ est une suite croissante d'événements, remarquons que

   $$\rho = \lim_{n\to+\infty} \uparrow \mathbb{P}(Z_n = 0)\,.$$

   Étant donné que $\mathbb{P}(Z_n = 0) = g_{Z_n}(0) = g_\mu^{\circ n}(0)$, il vient :

   $$\rho = \lim_{n\to+\infty} g_\mu^{\circ n}(0) = \lim_{n\to+\infty} g_\mu^{\circ n+1}(0) = g_\mu\Big(\lim_{n\to+\infty} g_\mu^{\circ n}(0)\Big) = g_\mu(\rho)\,,$$

   la troisième égalité provenant de la continuité de g_μ. Ainsi, ρ est un point fixe de g_μ.
   Si z est un autre point fixe de g_μ sur $[0\,;1]$, alors par croissance de g_μ :

   $$\rho = \lim_{n\to+\infty} g_\mu^{\circ n}(0) \leq \lim_{n\to+\infty} g_\mu^{\circ n}(z) = z\,,$$

   et donc ρ est le plus petit point fixe de g_μ sur $[0\,;1]$.

4. Pour conclure, distinguons plusieurs cas selon la valeur de $m = g'_\mu(1)$:

   $m > 1$ Comme $g_\mu(1) = 1$ et $g'_\mu(1) > 1$

   $$\exists \eta \in [0\,;1] \quad \forall s \in [\eta\,;1[\qquad g_\mu(s) < s\,.$$

   Vu que $g_\mu(0) \geq 0$, le théorème des valeurs intermédiaires assure l'existence d'un point fixe de g_μ sur $[0\,;\eta[$. Autrement dit, $\rho < 1$.

   $m < 1$ L'application génératrice g_μ est convexe donc reste au dessus de sa tangente en 1, ce qui implique que 1 est le seul point fixe de g_μ sur $[0\,;1]$, c'est-à-dire $\rho = 1$.

   $m = 1$ Puisque $\mu \neq \delta_1$ et $m = 1$, on sait que $\mu([2\,;+\infty[) > 0$, et donc que g_μ est strictement convexe. Ainsi, g_μ reste au dessus de sa tangente en 1 et ne la touche qu'en 1. Comme dans le cas précédent, $\rho = 1$.

COMMENTAIRES

Il est utile de faire un dessin représentant chacun des trois cas. Il faut être à l'aise avec (et savoir démontrer) les propriétés de convexité : fonction au dessus (strictement) de sa tangente, … La convexité ne sert pas à conclure dans le cas $m > 1$, cependant elle permet d'assurer dans ce cas que ρ est l'unique point fixe de g_μ sur $[0\,;1[$.

Le résultat ne suppose pas que la loi μ admette une espérance finie : si $m = +\infty$, l'argument du premier cas considéré ci-dessus reste valable.

46 PROJECTION SUR UN CONVEXE FERMÉ ET THÉORÈME DE RIESZ-FRÉCHET

[Bre99, Ch5, p79–82]

ÉNONCÉ

THÉORÈME. [PROJECTION SUR UN CONVEXE FERMÉ]
Soit C un convexe fermé d'un espace de HILBERT *H sur $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$. Alors :*

$$\forall x \in H \quad \exists! p \in C \qquad \|x - p\| = d(x, C)\,.$$

De plus, p est l'unique élément de C satisfaisant

$$\forall c \in C \qquad \Re\big(\langle x - p \mid c - p\rangle\big) \leq 0\,.$$

COROLLAIRE. *Soit F un sous-espace vectoriel fermé de H. Alors $H = F \oplus F^\perp$.*

THÉORÈME. [THÉORÈME DE RIESZ-FRÉCHET]
Soit H un espace de HILBERT. *Alors pour toute application $\phi \in H'$:*

$$\exists! f_\phi \in H \quad \forall v \in H \qquad \phi(v) = \langle f_\phi \mid v\rangle\,.$$

De plus $\phi \in H' \longmapsto f_\phi \in H$ est une isométrie : $\|f_\phi\|_H = \|\phi\|_{H'}$ pour tout $\phi \in H'$.

DÉVELOPPEMENT

Commençons par montrer le théorème de projection. Fixons $x \in H$ et notons $d = d(x, C)$.

- Existence. Considérons une suite $(c_n)_{n\in\mathbb{N}}$ de C telle que $d_n = \|x - c_n\| \longrightarrow_{n\to+\infty} d$. On va montrer que c'est une suite de CAUCHY. Pour $p, q \in \mathbb{N}$, on a :

$$\begin{aligned}
\|c_p - c_q\|^2 &= \big\|(c_p - x) - (c_q - x)\big\|^2 \\
&= 2\left(d_p^2 + d_q^2\right) - \big\|(c_p - x) + (c_q - x)\big\|^2 \quad \text{par l'in. du parallèlogramme} \\
&= 2\left(d_p^2 + d_q^2\right) - 4\left\|\frac{c_p + c_q}{2} - x\right\|^2 \\
\|c_p - c_q\|^2 \leq 2\left(d_p^2 + d_q^2\right) - 4\,d^2 &\underset{p,\,q\to+\infty}{\longrightarrow} 0 \qquad \text{car } \frac{c_p + c_q}{2} \in C \text{ par convexité.}
\end{aligned}$$

  Ainsi, $(c_n)_{n\in\mathbb{N}}$ est une suite de CAUCHY. Comme C est complet, elle converge vers un point $p \in C$ et $\|x - p\| = d$ par continuité de la norme.
- Unicité. Soient $c_1, c_2 \in C$ satisfaisants. Alors en posant $c = \frac{c_1+c_2}{2} \in C$:

$$d^2 \leq \|x - c\|^2 = 2\left(\left\|\frac{x - c_1}{2}\right\|^2 + \left\|\frac{x - c_2}{2}\right\|^2\right) - \left\|\frac{c_1 - c_2}{2}\right\|^2 = d^2 - \left\|\frac{c_1 - c_2}{2}\right\|^2,$$

  si bien que $c_1 = c_2$.

- Soit $a \in C$. Pour $c \in C$ et $t \in [0, 1]$, on a $(1 - t)\,a + t\,c \in C$ donc :

$$\begin{aligned}
&\forall c \in C \quad \forall t \in]0\,;1] \quad \big\|x - \big((1 - t)\,a + t\,c\big)\big\|^2 \geq \|x - a\|^2 \\
\iff\ &\forall c \in C \quad \forall t \in]0\,;1] \quad \big\|x - a - t\,(c - a)\big\|^2 \geq \|x - a\|^2 \\
\iff\ &\forall c \in C \quad \forall t \in]0\,;1] \quad \|x - a\|^2 - 2t\Re\big(\langle x - a \mid c - a\rangle\big) + t^2\|c - a\|^2 \geq \|x - a\|^2 \\
\iff\ &\forall c \in C \quad \forall t \in]0\,;1] \quad -2\Re\big(\langle x - a \mid c - a\rangle\big) + t\|c - a\|^2 \geq 0\,.
\end{aligned}$$

  Si $a = p$, alors en faisant tendre t vers 0 dans la dernière inégalité, on obtient bien

$$\forall c \in C \qquad \Re\big(\langle x - p \mid c - p\rangle\big) \leq 0\,.$$

  Réciproquement, si $a \in C$ satisfait $\Re\big(\langle x - a \mid c - a\rangle\big) \leq 0$, alors on peut remonter dans les équivalences et en prenant $c = p$ et $t = 1$ dans la première inégalité, il vient

$$d = \|x - p\| \geq \|x - a\|\,, \qquad \text{soit} \qquad a = p\,.$$

Passons au corollaire. Soit $x \in H$. Par le théorème, la projection p sur F de x vérifie

$$\forall c \in C \quad \forall t \in \mathbb{K} \qquad \Re\big(\langle x - p \mid t\,c - p\rangle\big) \leq 0\,.$$

En prenant $t \in \mathbb{R}$ (puis aussi $t \in i\mathbb{R}$ si $\mathbb{K} = \mathbb{C}$), on obtient $\langle x - p \mid c\rangle = 0$. Ainsi, on a la décomposition $x = p + (x - p) \in F + F^\perp$. D'où $H = F \oplus F^\perp$.

Reste à montrer le théorème de RIESZ-FRÉCHET.

- Existence. Soit $\phi \in H'$. Remarquons d'abord que $f_\phi = 0$ convient pour $\phi = 0$.
  Sinon, on peut considérer $F = \ker(\phi) = \phi^{-1}(\{0\})$. C'est un hyperplan strict de H puisque $\dim(F^\perp) = 1$, du fait que $\phi_{|F^\perp}$ est linéaire, injective et non nulle.
  On peut alors considérer $f \in F^\perp$ tel que $\phi(f) = 1$. En posant $f_\phi = \frac{f}{\|f\|_H^2}$, il vient

$$\forall \lambda \in \mathbb{K} \quad \phi(\lambda \cdot f) = \lambda = \langle f_\phi \mid \lambda \cdot f\rangle \qquad \text{et} \qquad \forall v \in F \quad \phi(v) = 0 = \langle f_\phi \mid v\rangle\,.$$

  Étant donné que ϕ est continue, F est fermé, donc $H = F \oplus F^\perp$ d'après le corollaire, ce qui assure finalement que $\phi = \langle f_\phi \mid \cdot\rangle$.
- Unicité. Il suffit de remarquer que $f \in H \longmapsto \langle f \mid \cdot\rangle$ est linéaire et de noyau nul.

Soit $\phi \in H'$. De l'inégalité de CAUCHY-SCHWARZ, on tire $\|\phi\|_{H'} \leq \|f_\phi\|_H$ puis (cas d'égalité)

$$\|\phi\|_{H'} \geq \frac{\|\langle f_\phi \mid f_\phi\rangle\|}{\|f_\phi\|} = \|f_\phi\|\,.$$

Ainsi, $\|f_\phi\|_H = \|\phi\|_{H'}$.

47 PROLONGEMENT HOLOMORPHE DE Γ

[BMP05, §2.7, p82] [QZ13, §IX.II.1, p312]

ÉNONCÉ

Pour $z \in \mathbb{C}$, on définit, si cela à un sens, $\Gamma(z) = \int_0^{+\infty} t^{z-1}\,\mathrm{e}^{-t}\,\mathrm{d}t$.

PROPOSITION. *L'application Γ est holomorphe sur $\{z \in \mathbb{C} : \Re(z) > 0\}$.*

APPLICATION. Γ admet un unique prolongement méromorphe, et holomorphe sur $\mathbb{C} \setminus \mathbb{Z}_-$.

APPLICATION. [FORMULE DE GAUSS]

$$\forall z \in \Omega_0 \qquad \Gamma(z) = \lim_{n\to+\infty} \frac{n!\,n^z}{z(z+1)\cdots(z+n)}\,.$$

DÉVELOPPEMENT

Commençons par montrer que Γ est définie et holomorphe sur $\Omega_0 = \{z \in \mathbb{C} : \Re(z) > 0\}$.
Soit la fonction

$$\begin{array}{rccc} f : & \Omega_0 \times \mathbb{R}_+ & \longrightarrow & \mathbb{C} \\ & (z,t) & \longmapsto & \mathrm{e}^{(z-1)\ln(t)}\,\mathrm{e}^{-t} \end{array}$$

- Pour tout $z \in \Omega_0$, l'application $f(z,\cdot\,)$ est mesurable (car continue) sur $\mathbb{R}_+$,
- Pour tout $t \in \mathbb{R}_+$, l'application $f(\cdot\,,t)$ est holomorphe sur Ω_0,
- Soit $\varepsilon, M \in \mathbb{R}_+^*$ avec $\varepsilon < M$. Pour $z \in \Omega_0$ tel que $\Re(z) \in [\varepsilon\,;M]$, on a la domination

$$\forall t \in \mathbb{R}_+ \qquad \big|f(z,t)\big| = \mathrm{e}^{(\Re(z)-1)\ln(t)}\,\mathrm{e}^{-t} \leq \begin{cases} \mathrm{e}^{(\varepsilon-1)\ln(t)} = \frac{1}{t^{1-\varepsilon}} & \text{si } t \leq 1 \\ \mathrm{e}^{(M-1)\ln(t)}\,\mathrm{e}^{-t} = t^{M-1}\,\mathrm{e}^{-t} & \text{si } t > 1\,. \end{cases}$$

Ainsi, on a une domination par une fonction intégrable sur tout compact de Ω_0.

D'après le théorème d'holomorphie sous le signe intégral, Γ est holomorphe sur Ω_0.

Passons à l'application en prolongeant Γ sur $\mathbb{C} \setminus \mathbb{Z}_-$.
Soit $z \in \Omega_0$. Une intégration par parties donne :

$$\Gamma(z+1) = \int_0^{+\infty} t^z\,\mathrm{e}^{-t}\,\mathrm{d}t = \big[-t^z\,\mathrm{e}^{-t}\big]_0^{+\infty} + \int_0^{+\infty} z\,t^{z-1}\,\mathrm{e}^{-t}\,\mathrm{d}t = z\int_0^{+\infty} t^{z-1}\,\mathrm{e}^{-t}\,\mathrm{d}t = z\,\Gamma(z)\,.$$

Fixons $n \in \mathbb{N}$. Par une récurrence immédiate, il vient :

$$\Gamma(z+n) = (z+n-1)\,\Gamma(z+n-1) = \cdots = (z+n-1)\,\cdots\,z\,\Gamma(z)\,.$$

Définissons alors

$$\Gamma_n : z \longmapsto \frac{\Gamma(z+n)}{(z+n-1)\cdots z} \quad \text{où} \quad \Omega_n = \{z \in \mathbb{C} : \mathrm{Re}(z) > -n\} \setminus \{0\,;-1\,;\ldots\,;-(n-1)\}\,.$$

L'application Γ_n est bien définie et holomorphe sur Ω_n. Comme de plus Γ_n coïncide avec Γ sur Ω_0, c'est un prolongement holomorphe de Γ.

Notons, d'après le théorème de prolongement analytique, que Γ_m et Γ_n coïncident sur $\Omega_m \cap \Omega_n$ pour tout $m,n \in \mathbb{N}$. En posant, pour $z \in \mathbb{C} \setminus \mathbb{Z}_-$, $\tilde{\Gamma}(z) = \Gamma_n(z)$ où $n \in \mathbb{N}$ est tel que $z \in \Omega_n$, on définit ainsi un unique prolongement analytique $\tilde{\Gamma}$ de Γ sur $\mathbb{C} \setminus \mathbb{Z}_-$. Il vient alors

$$\forall n \in \mathbb{N} \quad \forall z \in \Omega_{n+1} \qquad (z+n)\,\tilde{\Gamma}(z) = \frac{\tilde{\Gamma}(z+n+1)}{(z+n-1)\cdots z} \xrightarrow[z\to-n]{} \frac{\Gamma(1)}{(-1)^n\,n!} = \frac{(-1)^n}{n!}\,,$$

et $\tilde{\Gamma}$ est méromorphe, de pôles les $(-n)_{n\in\mathbb{N}}$ tous simples ($\mathrm{Res}\big(\tilde{\Gamma},-n\big) = \frac{(-1)^n}{n!}$ pour $n \in \mathbb{N}$).

Passons à la formule de GAUSS.
Soit $z \in \Omega_0$. Par le théorème de convergence dominée :

$$\Gamma(z) = \lim_{n\to+\infty} I_n \qquad \text{où} \qquad I_n = \int_0^n t^{z-1}\Big(1-\frac{t}{n}\Big)^n\,\mathrm{d}t\,.$$

Soit $n \in \mathbb{N}$. En intégrant n fois par parties :

$$\begin{aligned} I_n &= \Big[\frac{t^z}{z}\Big(1-\frac{t}{n}\Big)^n\Big]_0^n + \frac{n}{n\,z}\int_0^n t^z\Big(1-\frac{t}{n}\Big)^{n-1}\,\mathrm{d}t = \frac{n}{n\,z}\frac{n-1}{n\,(z+1)}\int_0^n t^{z+1}\Big(1-\frac{t}{n}\Big)^{n-2}\,\mathrm{d}t \\ &= \frac{n}{n\,z}\frac{n-1}{n\,(z+1)}\cdots\frac{1}{n\,(z+n-1)}\int_0^n t^{z+n-1}\,\mathrm{d}t = \frac{n}{n\,z}\frac{n-1}{n\,(z+1)}\cdots\frac{1}{n\,(z+n-1)}\frac{n^{z+n}}{z+n} \\ I_n &= \frac{n!\,n^z}{z(z+1)\cdots(z+n)}\,. \end{aligned}$$

Le résultat en découle.

COMMENTAIRES

Attention, ce développement diffère beaucoup de la référence.

Il est conseillé de faire un dessin de Ω_n. Le jury insiste particulièrement sur la longueur de ce développement : il faut démontrer tout ce qui est annoncé en gérant bien le temps. Par ailleurs, il faut maîtriser les questions autour de la convergence de séries holomorphes et méromorphes.

Une application possible de la formule de GAUSS consiste à montrer que le prolongement $\tilde{\Gamma}$ ne s'annule pas et que de plus $\frac{1}{\tilde{\Gamma}}$ est une fonction entière.

48 STABILITÉ DE LIAPOUNOV

[Ber17, §6.2, p239] [Rou99, §3.3, p127]

ÉNONCÉ

THÉORÈME. [STABILITÉ DE LIAPOUNOV]
Soient $n \in \mathbb{N}^*$ et $f \in \mathcal{C}^1(\mathbb{R}^n, \mathbb{R}^n)$ telle que $f(0) = 0$. Si les valeurs propres de $Df(0)$ ont leur partie réelle strictement négative, 0 est un point d'équilibre attractif du système $y' = f(y)$.

DÉVELOPPEMENT

Posons $A = Df(0)$. On procède en trois étapes.

1. Soit z une solution de $z' = Az$. En posant $a = \min_{\lambda \in \mathrm{Sp}(A)} -\mathrm{Re}(\lambda) > 0$, montrons que
$$\exists C \in \mathbb{R}_+^* \quad \forall t \in \mathbb{R} \qquad \big\|z(t)\big\| \leq C\,\mathrm{e}^{-a\frac{t}{2}}\big\|z(0)\big\|\,.$$
D'après le lemme des noyaux, $\mathbb{C}^n = \bigoplus_{\lambda \in \mathrm{Sp}(A)} E_\lambda$ où, pour une valeur propre $\lambda \in \mathrm{Sp}(A)$, $E_\lambda = \ker(A - \lambda I_n)^{m_\lambda}$ avec m_λ la multiplicité de λ dans A. Le vecteur $z(0)$ se décompose donc de manière unique en $z(0) = \sum_{\lambda \in \mathrm{Sp}(A)} z_\lambda$ où $z_\lambda \in E_\lambda$ pour $\lambda \in \mathrm{Sp}(A)$. Alors :
$$\begin{aligned}\forall t \in \mathbb{R} \qquad z(t) = \mathrm{e}^{tA}\,z(0) &= \sum_{\lambda \in \mathrm{Sp}(A)} \mathrm{e}^{tA}\,z_\lambda = \sum_{\lambda \in \mathrm{Sp}(A)} \mathrm{e}^{t\lambda}\,\mathrm{e}^{t(A-\lambda I_n)}\,z_\lambda \\ &= \sum_{\lambda \in \mathrm{Sp}(A)} \mathrm{e}^{t\lambda}\left(\sum_{p=0}^{m_\lambda} \frac{t^p}{p!}(A-\lambda I_n)^p\right) z_\lambda \qquad \text{car } z_\lambda \in E_\lambda\,.\end{aligned}$$
Fixons $\lambda \in \mathrm{Sp}(A)$. Il est facile de vérifier que $\mathrm{e}^{tA}\,z_\lambda \in E_\lambda$ pour tout $t \in \mathbb{R}$, puis que
$$\exists C_\lambda \in \mathbb{R}_+^* \quad \forall t \in \mathbb{R} \qquad \big|\!\big|\!\big|\mathrm{e}^{tA}_{|E_\lambda}\big|\!\big|\!\big| \leq \mathrm{e}^{t\,\mathrm{Re}(\lambda)}\,C_\lambda\big(1+|t|\big)^{m_\lambda - 1} \leq \mathrm{e}^{-at}\,C_\lambda\big(1+|t|\big)^{n-1}\,.$$
Par l'inégalité triangulaire, il en découle, si $C = \max_{\lambda \in \mathrm{Sp}(A)} C_\lambda$, que pour tout $t \in \mathbb{R}$
$$\big\|z(t)\big\| \leq \sum_{\lambda \in \mathrm{Sp}(A)} \big\|\mathrm{e}^{tA}\,z_\lambda\big\| \leq \sum_{\lambda \in \mathrm{Sp}(A)} \big|\!\big|\!\big|\mathrm{e}^{tA}_{|E_\lambda}\big|\!\big|\!\big| \cdot \|z_\lambda\| \leq C\big(1+|t|\big)^{n-1}\,\mathrm{e}^{-at} \sum_{\lambda \in \mathrm{Sp}(A)} \|z_\lambda\|\,.$$
Le résultat en découle, quitte à modifier la valeur de C, par équivalence des normes ($z \in \mathbb{R}^n \longmapsto \sum_{\lambda \in \mathrm{Sp}(A)} \|z_\lambda\|$ étant une norme). Ainsi, les solutions de $z' = Az$ convergent vers 0, point d'équilibre attractif du système linéaire.
2. Montrons le même comportement localement pour le système non linéaire. Posons
$$\forall (u,v) \in \big(\mathbb{C}^n\big)^2 \qquad b(u,v) = \int_0^{+\infty} \big\langle \mathrm{e}^{tA}\,u \mid \mathrm{e}^{tA}\,v\big\rangle\,\mathrm{d}t \quad \text{puis} \quad q(u) = b(u,u)\,.$$
Les applications b et u sont bien définies puisque, d'après l'inégalité de CAUCHY-SCHWARZ,
$$\forall (u,v) \in \big(\mathbb{C}^n\big)^2 \qquad \big|\big\langle \mathrm{e}^{tA}\,u \mid \mathrm{e}^{tA}\,v\big\rangle\big| \leq \big\|\mathrm{e}^{tA}\,u\big\| \cdot \big\|\mathrm{e}^{tA}\,v\big\| \leq C^2\,\mathrm{e}^{-at}\,\|u\| \cdot \|v\|\,,$$
ce qui assure l'absolue convergence de l'intégrale.

Ainsi définies, b est une forme bilinéaire symétrique et q une forme quadratique définie positive. Il en découle que b est donc un produit scalaire sur $\mathbb{R}^n$. Remarquons également que q est différentiable, telle que $dq(u)(v) = 2\,b(u,v)$ pour $u, v \in \mathbb{C}^n$.
Soit y une solution de l'équation différentielle $y' = f(y)$. Posons $r(u) = f(u) - Au$ pour $u \in \mathbb{R}^n$. Puisque f est de classe $\mathcal{C}^1$ et $f(0) = 0$, on a $r(u) =_{u \to 0} o\big(\|u\|\big)$.
Montrons qu'il existe $\beta \in \mathbb{R}_+^*$ telle que $q(y)' \leq -\beta q(y)$ dès que y est assez proche de 0.
Par bilinéarité de b, on a :
$$q(y)' = dq(y)(y') = 2\,b(y,y') = 2\,b\big(y, f(y)\big) = 2\,b(y,Ay) + 2\,b\big(y, r(y)\big)\,.$$
Or, pour tout $u \in \mathbb{R}^n$:
$$2\,b(u,Au) = \big\langle \nabla q(u) \mid Au\big\rangle = \int_0^{+\infty} 2\big\langle \mathrm{e}^{tA}\,u \mid \mathrm{e}^{tA}\,Au\big\rangle\,\mathrm{d}t = \Big[\big\|\mathrm{e}^{tA}\,u\big\|^2\Big]_0^{+\infty} = -\|u\|^2\,.$$
Comme $\sqrt{q}$ est une norme sur $\mathbb{R}^n$, elle est équivalente à $\|\cdot\|$ et
$$\exists C' \in \mathbb{R}_+^* \qquad C'q(y) \leq \|y\|^2 \quad \text{d'où} \quad q(y)' \leq -C'q(y) + 2\,b\big(y, r(y)\big)\,.$$
Par ailleurs, l'équivalence des normes donne aussi que $r(u) =_{u \to 0} o\big(\sqrt{q(u)}\big)$, et donc :
$$\forall \varepsilon \in \mathbb{R}_+^* \quad \exists \eta \in \mathbb{R}_+^* \quad \forall u \in \mathbb{R}^n \qquad \sqrt{q(u)} \leq \eta \quad \Longrightarrow \quad \sqrt{q\big(r(u)\big)} \leq \varepsilon\sqrt{q(u)}\,.$$
Fixons $\varepsilon > 0$, et soit $\eta > 0$ associé. L'inégalité de CAUCHY-SCHWARZ donne :
$$\begin{aligned}\forall u \in \mathbb{R}^n \qquad q(u) \leq \eta^2 \quad &\Longrightarrow \quad \big|b\big(u, r(u)\big)\big| \leq \sqrt{q(u)} \cdot \sqrt{q\big(r(u)\big)} \leq \varepsilon\,q(u)\,, \\ \text{donc} \qquad q(y) \leq \eta^2 \quad &\Longrightarrow \quad q(y)' = -(C' - 2\varepsilon)\,q(y) = -\beta\,q(y)\,,\end{aligned}$$
où β est strictement positif dès que ε assez petit (rappelons que C' ne dépend par de ε).
3. Pour finir, notons que $q(y) \leq \eta^2$ sur $\mathbb{R}_+$ si $q\big(y(0)\big) \leq \eta^2$.
En effet, si ce n'est pas le cas, il existe $t_0, \delta > 0$ tels que $q(y) \leq \eta^2$ sur $[0\,;t_0]$ et $q(y) > \eta^2$ sur $]t_0\,;t_0+\delta]$, d'où $q\big(y(t_0)\big) = \eta^2$ et $q(y)'(t_0) \leq -\beta\,q\big(y(t_0)\big) < 0$, ce qui est absurde.
Ainsi, si $q\big(y(0)\big) \leq \eta^2$, on a $q(y)' \leq -\beta\,q(y)$, soit $\big(\mathrm{e}^{\beta t}\,q(y)\big)' \leq 0$ sur $\mathbb{R}_+$, et alors
$$q\big(y(t)\big) \leq \mathrm{e}^{-\beta t}\,q\big(y(0)\big) \quad \underset{t \to +\infty}{\longrightarrow} \quad 0\,.$$
La solution y converge donc vers 0. Ainsi 0 est stable.

COMMENTAIRES

La preuve consiste à construire une norme N telle que $N\big(y(t)\big) \leq C\,\mathrm{e}^{-\beta t}$ pour $t > 0$.

C'est un développement assez long, dont la fin technique requiert de ne pas être trop pressé.

49 THÉORÈME CENTRAL LIMITE ET INTERVALLE DE CONFIANCE

[BL07, §V.5, p136–137] [RS12, §3.4, p25]

ÉNONCÉ

THÉORÈME. [THÉORÈME CENTRAL LIMITE]
Soient $(X_i)_{i\in\mathbb{N}^*}$ des variables aléatoires i.i.d. réelles de carré intégrable, de moyenne m et de variance σ^2. Alors

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}(X_i-m) \quad \xrightarrow[n\to+\infty]{\mathcal{L}} \quad Z\sim\mathcal{N}(0,\sigma^2)\,.$$

APPLICATION. [INTERVALLE DE CONFIANCE ASYMPTOTIQUE]
Soient $(X_i)_{i\in\mathbb{N}^*}$ des réalisations i.i.d. de $\mathcal{B}(p)$ pour un $p\in[0\,;1]$ inconnu. Soient $\alpha\in]0\,;1]$ et $n\in\mathbb{N}$. Si $\hat{p}_n=\frac{1}{n}\sum_{k=1}^{n}X_i$ et si q_t est le quantile d'ordre $t\in[0\,;1]$ de $\mathcal{N}(0,1)$, alors

$$\mathrm{IC}_n^{(\alpha)}(X_1,\dots,X_n)=\left[\hat{p}_n\pm\frac{q_{1-\frac{\alpha}{2}}}{2\sqrt{n}}\right]$$

est un intervalle de confiance asymptotique de niveau α pour p.

DÉVELOPPEMENT

Si les $(X_i)_{i\in\mathbb{N}^*}$ sont constantes p.s., le résultat est clair (avec la convention $\mathcal{N}(0,0)=\delta_0$).

Sinon, supposons, quitte à considérer $\left(\frac{X_i-m}{\sigma}\right)_{i\in\mathbb{N}^*}$ plutôt que $(X_i)_{i\in\mathbb{N}^*}$, que $m=0$ et $\sigma=1$. En posant $S_n=\sum_{i=1}^{n}X_i$ pour $n\in\mathbb{N}^*$, montrons que

$$\forall t\in\mathbb{R} \qquad \phi_{\frac{S_n}{\sqrt{n}}}(t) \quad \xrightarrow[n\to+\infty]{} \quad \phi_{\mathcal{N}(0,1)}(t)=\mathrm{e}^{-\frac{t^2}{2}}\,.$$

Fixons $t\in\mathbb{R}$. Puisque les $(X_i)_{i\in\mathbb{N}^*}$ sont i.i.d. :

$$\phi_{\frac{S_n}{\sqrt{n}}}(t)=\mathbb{E}\left[\mathrm{e}^{\mathrm{i}t\frac{S_n}{\sqrt{n}}}\right]=\prod_{i=1}^{n}\mathbb{E}\left[\mathrm{e}^{\mathrm{i}t\frac{X_i}{\sqrt{n}}}\right]=\phi_{X_1}\left(\frac{t}{\sqrt{n}}\right)^n.$$

Comme X_1 est de carré intégrable, les théorèmes de dérivation (appliqué deux fois) et de continuité sous le signe intégral assurent que ϕ_{X_1} est de classe $\mathcal{C}^2$ sur $\mathbb{R}$, avec

$$\phi'_{X_1}(0)=\mathrm{i}\,\mathbb{E}[X]=0 \qquad \text{et} \qquad \phi''_{X_1}(0)=-\mathbb{E}[X^2]=-1\,.$$

En appliquant à ϕ_{X_1} la formule de TAYLOR-YOUNG à l'ordre 2 en 0, il vient

$$\phi_{X_1}(u)\underset{u\to0}{=}1-\frac{u^2}{2}+o(u^2)\,, \qquad \text{d'où}$$

$$\phi_{\frac{S_n}{\sqrt{n}}}(t)\underset{n\to+\infty}{=}\left(1-\frac{t^2}{2n}+o\left(\frac{1}{n}\right)\right)^n\underset{n\to+\infty}{=}\left(1+\frac{z_n}{n}\right)^n, \qquad \text{où } z_n\underset{n\to+\infty}{=}-\frac{t^2}{2}(1+o(1))\in\mathbb{C}\,.$$

LEMME. $\forall z\in\mathbb{C} \qquad \left|\mathrm{e}^z-\left(1+\frac{z}{n}\right)^n\right|\le\mathrm{e}^{|z|}-\left(1+\frac{|z|}{n}\right)^n.$

En effet, on obtient en posant $\alpha_n^k=1-\frac{n!}{(n-k)!\,n^k}\in[0\,;1]$ pour $k,n\in\mathbb{N}$ tels que $k\le n$:

$$\mathrm{e}^z-\left(1+\frac{z}{n}\right)^n=\sum_{k=0}^{+\infty}\frac{z^k}{k!}-\sum_{k=0}^{n}\binom{n}{k}\frac{z^k}{n^k}=\sum_{k=0}^{n}\frac{z^k}{k!}\,\alpha_n^k\,,$$

$$\text{d'où} \qquad \left|\mathrm{e}^z-\left(1+\frac{z}{n}\right)^n\right|\le\sum_{k=0}^{n}\frac{|z|^k}{k!}\alpha_n^k=\mathrm{e}^{|z|}-\left(1+\frac{|z|}{n}\right)^n.$$

Revenons à la suite de complexes $(z_n)_{n\in\mathbb{N}}$: elle vérifie $\mathrm{e}^{|z_n|}\longrightarrow_{n\to+\infty}\mathrm{e}^{\frac{t^2}{2}}$ et

$$\left(1+\frac{|z_n|}{n}\right)^n=\exp\left(n\ln\left(1+\frac{|z_n|}{n}\right)\right)\underset{n\to+\infty}{=}\exp\Big(|z_n|\,(1+o(1))\Big) \quad \xrightarrow[n\to+\infty]{} \quad \exp\left(\frac{t^2}{2}\right).$$

D'après le lemme, il vient alors que

$$\phi_{\frac{S_n}{\sqrt{n}}}(t)\underset{n\to+\infty}{\sim}\mathrm{e}^{z_n} \quad \xrightarrow[n\to+\infty]{} \quad \mathrm{e}^{-\frac{t^2}{2}}\,.$$

Le théorème de LÉVY permet de conclure la convergence en loi annoncée.

Passons à l'application. Comme les $(X_i)_{i\in\mathbb{N}^*}$ satisfont les hypothèses du théorème central limite avec $m=\mathbb{E}[X_1]=p$ et $\sigma=\mathrm{Var}(X_1)=p\,(1-p)$, on a la convergence :

$$\frac{\sqrt{n}}{\sqrt{p\,(1-p)}}(\hat{p}_n-p)=\frac{1}{\sqrt{n}\,\sqrt{p\,(1-p)}}\sum_{i=1}^{n}(X_i-p) \quad \xrightarrow[n\to+\infty]{\mathcal{L}} \quad Z\sim\mathcal{N}(0,1)\,.$$

La fonction de répartition de $\mathcal{N}(0,1)$ est continue donc, d'après le théorème de PORTMANTEAU :

$$\forall a<b\in\mathbb{R} \qquad \mathbb{P}\left(a\le\frac{\sqrt{n}}{\sqrt{p\,(1-p)}}(\hat{p}_n-p)\le b\right) \quad \xrightarrow[n\to+\infty]{} \quad \mathbb{P}(a\le Z\le b)\,.$$

Prenant $a=q_{\frac{\alpha}{2}}$ et $b=q_{1-\frac{\alpha}{2}}$, on a $b=-a>0$ par symétrie de $\mathcal{N}(0,1)$, d'où :

$$\mathbb{P}\left(p\in\left[\hat{p}_n\pm q_{1-\frac{\alpha}{2}}\frac{\sqrt{p\,(1-p)}}{\sqrt{n}}\right]\right)=\mathbb{P}\left(\left|\frac{\sqrt{n}}{\sqrt{p\,(1-p)}}(\hat{p}_n-p)\right|\le q_{1-\frac{\alpha}{2}}\right) \quad \xrightarrow[n\to+\infty]{} \quad 1-\alpha\,.$$

Comme $p\,(1-p)\le\frac{1}{4}$, on en déduit l'intervalle de confiance asymptotique annoncé.

COMMENTAIRES

Attention, le lemme est nécessaire du fait que le petit o définissant z_n est complexe.

On peut trouver un meilleur intervalle de confiance en utilisant le lemme de SLUTSKY à la fin.

50 THÉORÈME DE BANACH-STEINHAUS ET SÉRIE DE FOURIER DIVERGENTE

[Bre99, II.1, p16–17] [Gou08, An. A, p404–405]

ÉNONCÉ

THÉORÈME. [THÉORÈME DE BANACH-STEINHAUS]
Soient E un espace de BANACH *et F un espace vectoriel normé. Soit $(u_i)_{i\in I}$ une famille, indicée par un ensemble I quelconque, d'applications de $\mathcal{L}_c(E,F)$ simplement bornées (c'est-à-dire telles que $\sup_{i\in I}\|u_i(x)\| < +\infty$ pour tout $x \in E$). Alors*

$$\sup_{i\in I} |||u_i||| < +\infty\,.$$

APPLICATION. Il existe f continue, 2π-périodique et dont la série de FOURIER diverge en 0.

DÉVELOPPEMENT

Commençons par démontrer le théorème. Introduisons, pour $k \in \mathbb{N}$,

$$E_k = \bigcap_{i\in I}\Big\{x\in E : \|u_i(x)\|_F \le k\Big\} = \bigcap_{i\in I}\|u_i\|_F^{-1}\big([0\,;k]\big)\,,$$

qui est fermé en tant qu'intersection de fermés, les $(u_i)_{i\in I}$ étant continues.
Par hypothèse, chaque élément de E appartient à l'un des $(E_k)_{k\in\mathbb{N}}$, soit $E = \cup_{k\in\mathbb{N}} E_k$.
L'union des $(E_k)_{k\in\mathbb{N}}$ est donc d'intérieure non vide. Daprès le théorème de BAIRE, cela implique que l'un au moins des $(E_k)_{k\in\mathbb{N}}$ est d'intérieur non vide, notons-le E_K.
Soit $\overline{\mathbb{B}}(a,r)$ une boule incluse dans E_K et supposée (quitte à réduire r) fermée.
Soit $x \in E$. Comme $a \in \overline{\mathbb{B}}(a,r)$ et $y = a + r\frac{x}{\|x\|_E} \in \overline{\mathbb{B}}(a,r)$, il vient :

$$\begin{aligned}\forall i\in I \qquad \|u_i(x)\|_F &= \left\|u_i\left(\frac{\|x\|_E}{r}(y-a)\right)\right\|_F = \frac{\|x\|_E}{r}\|u_i(y-a)\|_F\\ &\le \frac{\|x\|_E}{r}\Big(\|u_i(y)\|_F + \|u_i(a)\|_F\Big) \le \frac{2K}{r}\|x\|_E\,.\end{aligned}$$

Le théorème de BANACH-STEINHAUS en découle par définition de la norme d'opérateur :

$$\sup_{i\in I}|||u_i||| = \sup_{i\in I}\sup_{x\in E\setminus\{0\}}\frac{\|u_i(x)\|_F}{\|x\|_E} \le \frac{2K}{r}\,.$$

Passons à l'application. Soit $\mathbb{T} = \mathbb{R}/2\pi\mathbb{Z}$. Étudions, pour $N \in \mathbb{N}^*$, l'application

$$\begin{aligned}\ell_N :\ \mathcal{C}(\mathbb{T}) &\longrightarrow \mathbb{C}\\ f &\longmapsto \sum_{n=-N}^{N} c_n(f) = S_N(f)(0) = D_N \star f(0)\end{aligned} \qquad \text{où} \quad D_N = \sum_{k=-n}^{n} \mathrm{e}^{\mathrm{i}k\cdot}\,.$$

- Déjà, ℓ_N est linéaire par linéarité des coefficients de FOURIER, et

$$\forall f\in\mathcal{C}(\mathbb{T}) \qquad \big|\ell_N(f)\big| = \big|D_N\star f(0)\big| \le \frac{1}{2\pi}\int_0^{2\pi}\big|D_N(t)\big|\cdot\big|f(-t)\big|\,\mathrm{d}t \le \|f\|_\infty\cdot\|D_N\|_1\,,$$

si bien que ℓ_N est continue et vérifie $|||\ell_N||| \le \|D_N\|_1$.
- Pour $\varepsilon > 0$, soit $f_\varepsilon = \frac{D_N}{|D_N|+\varepsilon} \in \mathcal{C}(\mathbb{T})$. Alors $\ell_N(f_\varepsilon) \longrightarrow_{\varepsilon\to 0} \|D_N\|_1$ puisque

$$D_N\star f_\varepsilon(0) = \int_0^{2\pi}\frac{\big|D_N(t)\big|^2}{D_N(t)^2+\varepsilon}\,\mathrm{d}t \quad\underset{\varepsilon\to 0}{\longrightarrow}\quad \int_0^{2\pi}\big|D_N(t)\big|\,\mathrm{d}t\,,$$

par parité de D_N puis par convergence dominée (avec la domination $|D_N|$).
Comme $\|f_\varepsilon\|_\infty \le 1$ pour tout $\varepsilon > 0$, on obtient ainsi $|||\ell_N||| = \|D_N\|_1$.
- Montrons alors que $\|D_N\|_1 \longrightarrow_{N\to+\infty} +\infty$. On a :

$$\|D_N\|_1 = \frac{1}{2\pi}\int_0^{2\pi}\left|\frac{\sin(\frac{2N+1}{2}t)}{\sin\frac{t}{2}}\right|\mathrm{d}t \ge \frac{1}{\pi}\int_0^{\pi}\left|\frac{\sin(\frac{2N+1}{2}t)}{\frac{t}{2}}\right|\mathrm{d}t = \frac{2}{\pi}\int_0^{(2N+1)\frac{\pi}{2}}\frac{\big|\sin(u)\big|}{u}\,\mathrm{d}u\,,$$

l'inégalité s'obtenant par π-périodicité puis puisque $\big|\sin\frac{t}{2}\big| \le \big|\frac{t}{2}\big|$ pour $t \in [0\,;\pi]$, tandis que la dernière égalité provenant du changement de variable $u = \frac{(2N+1)t}{2}$. Il s'ensuit que

$$\begin{aligned}\|D_N\|_1 \quad\underset{N\to+\infty}{\longrightarrow}\quad \int_0^{+\infty}\frac{\big|\sin(u)\big|}{u}\,\mathrm{d}u &= \sum_{n\in\mathbb{N}}\int_{n\pi}^{(n+1)\pi}\frac{\big|\sin(u)\big|}{u}\,\mathrm{d}u\\ &\ge \sum_{n\in\mathbb{N}}\frac{1}{(n+1)\pi}\int_0^{\pi}\big|\sin(u)\big|\,\mathrm{d}u = \frac{2}{\pi}\sum_{n\in\mathbb{N}^*}\frac{1}{n} = +\infty\,.\end{aligned}$$

- Pour conclure, rappelons que $\mathcal{C}(\mathbb{T})$ est complet puisque fermé dans $\mathcal{B}(\mathbb{R},\mathbb{C})$ complet. D'après le théorème de BANACH-STEINHAUS, puisque $\sup_{N\in\mathbb{N}}|||\ell_N||| = +\infty$

$$\exists f\in\mathcal{C}(\mathbb{T}) \qquad \sup_{N\in\mathbb{N}}\big|\ell_N(f)\big| = +\infty \quad \text{soit} \quad \sup_{N\in\mathbb{N}}\big|S_N(f)(0)\big| = +\infty\,,$$

Autrement dit, la série de FOURIER de f diverge en 0. En particulier, la série de FOURIER de f diffère de f.

COMMENTAIRES

Attention à introduire la bonne fonction ℓ_N, à valeurs dans $\mathbb{C}$ et non dans $\mathcal{C}(\mathbb{T})$: elle égale à $S_N(f)$ évaluée en 0. Attention également à la constante $\frac{1}{2\pi}$ dans les calculs, et à ne pas omettre que $\mathcal{C}(\mathbb{T})$ est complet afin d'appliquer le théorème.

51 THÉORÈME DE BERNSTEIN

[Gou08, §4.4, p250–251]

ÉNONCÉ

THÉORÈME. [THÉORÈME DE BERNSTEIN]
Soit $a > 0$ et $f :]-a\,;a[\longrightarrow \mathbb{R}$ une fonction de classe $\mathcal{C}^\infty$ telle que, pour tout entier k, $f^{(2k)} \geq 0$ sur $]-a\,;a[$. Alors f admet un développement en série entière sur $]-a\,;a[$.

APPLICATION. L'application $\tan$ est développement en série entière sur $]-\frac{\pi}{2}\,;\frac{\pi}{2}[$.

DÉVELOPPEMENT

Notons d'abord qu'il suffit de montrer le résultat sur tout intervalle de la forme $]-b\,;b[$ où $b \in]0\,;a[$ (par unicité du développement en série entière sur un voisinage de 0, les coefficients du développement ne dépendront alors pas de b).

Fixons alors $b \in]0\,;a[$ et introduisons[1] $F : x \in [-b\,;b] \longmapsto f(x) + f(-x)$.
Montrons que F est développable en série entière sur $]-b\,;b[$. La fonction f étant paire, ses dérivées d'ordre impairs s'annulent en 0. Par ailleurs,

$$\forall n \in \mathbb{N} \quad \forall x \in [-b\,;b] \qquad F^{(2n)}(x) = f^{(2n)}(x) + f^{(2n)}(-x) \geq 0 \quad \text{et} \quad F^{(2n)}(0) = 2f^{(2n)}(0)\,.$$

Soit $n \in \mathbb{N}$. Comme F est de classe $\mathcal{C}^{2n+2}$, la formule de TAYLOR avec reste intégral donne

$$\forall x \in [0\,;b] \qquad F(x) = \sum_{k=0}^{n} \frac{F^{(2k)}(0)}{(2k)!}x^{2k} + R_n(x) \quad \text{où } R_n(x) = \int_0^x \frac{(x-t)^{2n+1}}{(2n+1)!}F^{(2n+2)}(t)\,\mathrm{d}t\,.$$

Par positivité des dérivées d'ordre pair de F, on sait que $0 \leq R_n(b) \leq F(b)$.
Remarquons que $R_n \longrightarrow_{n\to+\infty} 0$ simplement sur $[0\,;b[$, puisque

$$\begin{aligned}
\forall x \in [0\,;b[\qquad 0 \leq R_n(x) &= \int_0^x \frac{(x-t)^{2n+1}}{(b-t)^{2n+1}}\frac{(b-t)^{2n+1}}{(2n+1)!}F^{(2n+2)}(t)\,\mathrm{d}t \\
&\leq \left(\frac{x}{b}\right)^{2n+1} \int_0^x \frac{(b-t)^{2n+1}}{(2n+1)!}F^{(2n+2)}(t)\,\mathrm{d}t \\
\forall x \in [0\,;b[\qquad 0 \leq R_n(x) &\leq \left(\frac{x}{b}\right)^{2n+1} R_n(b) \leq \left(\frac{x}{b}\right)^{2n+1} F(b) \underset{n\to+\infty}{\longrightarrow} 0\,,
\end{aligned}$$

la première inégalité provenant de la décroissance de $t \longmapsto \frac{x-t}{b-t}$ sur $[0\,;x]$.
En utilisant la parité de F, on obtient donc le développement en série entière

$$\forall x \in]-b\,;b[\qquad F(x) = \sum_{k=0}^{+\infty} \frac{F^{(2k)}(0)}{(2k)!}x^{2k}\,.$$

1. travailler sur $[-b\,;b]$ permet de s'assurer que $F(b)$ est fini, alors que $F(a)$ n'est pas défini *a priori*

Procédons de même pour f. Fixons $x \in]-b\,;b[$. On écrit, pour $n \in \mathbb{N}$,

$$f(x) = \sum_{k=0}^{2n+1} \frac{f^{(k)}(0)}{k!}x^k + r_n(x) \qquad \text{où } r_n(x) = \int_0^x \frac{(x-t)^{2n+1}}{(2n+1)!}f^{(2n+2)}(t)\,\mathrm{d}t\,.$$

Comme $f^{(2n+2)}(t) \leq f^{(2n+2)}(t) + f^{(2n+2)}(-t) \leq F^{(2n+2)}(t)$ pour $t \in [-b\,;b]$, il vient que

$$0 \leq r_n(x) \leq R_n(x) \underset{n\to+\infty}{\longrightarrow} 0\,, \qquad \text{et donc} \qquad S_{2n+1}(x) \underset{n\to+\infty}{\longrightarrow} f(x)$$

où l'on a posé $S_p(x) = \sum_{k=0}^{p} \frac{f^{(k)}(0)}{k!}x^k$ pour $p \in \mathbb{N}$.

Pour conclure, il reste à montrer que $S_{2n}(x) \longrightarrow_{n\to+\infty} f(x)$. Puisque la série $\sum_{k\in\mathbb{N}} \frac{F^{(2k)}(0)}{(2k)!}x^{2k}$ est convergeante (de somme $F(x)$) :

$$S_{2n}(x) - S_{2n-1}(x) = \frac{f^{(2n)}(0)}{(2n)!}x^{2n} = \frac{1}{2}\frac{F^{(2n)}(0)}{(2n)!}x^{2n} \underset{n\to+\infty}{\longrightarrow} 0\,.$$

Ainsi, $(S_{2n})_{n\in\mathbb{N}}$ et $(S_{2n+1})_{n\in\mathbb{N}}$ sont adjacentes, ce qui implique que $S_{2n}(x) \longrightarrow_{n\to+\infty} f(x)$.
Finalement, $S_n(x) \longrightarrow_{n\to+\infty} f(x)$, et l'on a obtenu que

$$\forall x \in]-b\,;b[\qquad f(x) = \sum_{k=0}^{+\infty} \frac{f^{(k)}(0)}{k!}x^k\,.$$

Appliquons ce résultat à l'application $f = \tan'$, en vérifiant que les dérivées d'ordre pair de f sont positives sur $I =]-\frac{\pi}{2}\,;\frac{\pi}{2}[$. Pour cela, montrons par récurrence sur $k \in \mathbb{N}$ que

$$\exists P_k \in \mathbb{R}_+[X] \quad \forall x \in I \qquad f^{(2k)}(x) = P_k\big(\tan^2(x)\big) \geq 0\,.$$

- Pour $k = 0$, le polynôme $P_0 = 1 + X$ convient puisque $f = 1 + \tan^2$.
- Soit $k \in \mathbb{N}$ tel qu'il existe un polynôme $P_k \in \mathbb{R}[X]$ satisfaisant la propriété. Soit $x \in I$. Comme $f^{(2k)}(x) = P_k\big(\tan^2(x)\big)$, on obtient en dérivant deux fois, si $v = \tan(x)$,

$$\begin{aligned}
f^{(2k+1)}(x) &= 2\,(1+v^2)\,v\,P_k'(v^2) = 2\,(v+v^3)\,P_k'(v^2)\,, \\
f^{(2k+2)}(x) &= \underbrace{2\left(1+v^2+3\,(1+v^2)\,v^2\right)P_k'(v^2) + 2\,(v+v^3)\times 2\,(1+v^2)\,v\,P_k''(v^2)}_{P_{k+1}(v^2)}
\end{aligned}$$

Le polynôme P_{k+1} est bien à coefficients positifs (P_k' et P_k'' l'étant), ce qui justifie l'hérédité.

Par principe de récurrence, $\tan'$, et donc $\tan$, sont développables en série entière sur $]-\frac{\pi}{2}\,;\frac{\pi}{2}[$.

52 THÉORÈME DE CAUCHY-LIPSCHITZ LINÉAIRE

[Ber17]

ÉNONCÉ

THÉORÈME. [THÉORÈME DE CAUCHY-LIPSCHITZ (CAS LINÉAIRE)]
Soient I un intervalle de $\mathbb{R}$, $A \in \mathcal{C}\big(I, \mathcal{M}_n(\mathbb{K})\big)$ et $B \in \mathcal{C}(I, \mathbb{K}^n)$, où $n \in \mathbb{N}^*$ et $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$. À $(t_0, y_0) \in I \times \mathbb{K}^n$ fixé, il existe une unique solution y globale au problème de CAUCHY

$$\begin{cases} \forall t \in I \qquad y'(t) = A(t)\, y(t) + B(t) \\ \qquad\qquad\quad y(t_0) = y_0\,. \end{cases}$$

DÉVELOPPEMENT

Dans un premier temps, supposons que I est un intervalle compact de $\mathbb{R}$.
Soient $E = \mathcal{C}(I, \mathbb{K}^n)$ et $\|\cdot\|$ une norme quelconque sur $\mathbb{K}^n$. Notons d'une part, par compacité de I et par continuité de A, que $\alpha = \sup_{t\in I} \big|\!\big|\!\big|A(t)\big|\!\big|\!\big|$ est fini, et, d'autre part, que l'espace E est complet pour la norme infinie $\|\cdot\|_\infty$ associée à $\|\cdot\|$. Considérons l'application :

$$\begin{array}{rccc} \Phi : & E & \longrightarrow & E \\ & y & \longmapsto & \Big(t \in I \longmapsto y_0 + \displaystyle\int_{t_0}^t A(s)\, y(s) + B(s)\, \mathrm{d}s\Big)\,. \end{array}$$

et montrons qu'une itérée de Φ est contractante. Vérifions par récurrence sur $p \in \mathbb{N}$ que :

$$\forall (y_1, y_2) \in E^2 \quad \forall t \in I \qquad \big\|\Phi^p(y_1)(t) - \Phi^p(y_2)(t)\big\| \leq \frac{\alpha^p\, |t - t_0|^p}{p!}\, \|y_1 - y_2\|_\infty\,.$$

- Pour $p = 0$, il n'y a rien à montrer.
- Supposons le résultat vrai au rang $p \in \mathbb{N}$. Pour $y_1, y_2 \in E$, il vient :

$$\begin{aligned} \forall t \in I \qquad \big\|\Phi^{p+1}(y_1)(t) - \Phi^{p+1}(y_2)(t)\big\| &= \left\| \int_{t_0}^t A(s) \left[\Phi^p(y_1)(s) - \Phi^p(y_2)(s)\right] \mathrm{d}s \right\| \\ &\leq \int_{t_0}^t \big|\!\big|\!\big|A(s)\big|\!\big|\!\big| \cdot \big\|\Phi^p(y_1)(s) - \Phi^p(y_2)(s)\big\|\, \mathrm{d}s \\ &\leq \int_{t_0}^t \alpha \times \frac{\alpha^p\, |s - t_0|^p}{p!}\, \|y_1 - y_2\|_\infty\, \mathrm{d}s \\ &\leq \frac{\alpha^{p+1}}{p!}\, \|y_1 - y_2\|_\infty \int_{t_0}^t |s - t_0|^p\, \mathrm{d}s \\ \forall t \in I \qquad \big\|\Phi^{p+1}(y_1)(t) - \Phi^{p+1}(y_2)(t)\big\| &\leq \frac{\alpha^{p+1}\, |t - t_0|^{p+1}}{(p+1)!}\, \|y_1 - y_2\|_\infty\,, \end{aligned}$$

où l'on a utilisé l'inégalité triangulaire dans la première inégalité et l'hypothèse d'hérédité dans la seconde. On obtient ainsi la propriété au rang $p+1$, ce qui justifie l'hérédité.

Le résultat en découle par principe de récurrence, et on en déduit immédiatement :

$$\forall p \in \mathbb{N} \quad \forall (y_1, y_2) \in E^2 \qquad \big\|\Phi^p(y_1) - \Phi^p(y_2)\big\|_\infty \leq \frac{\alpha^p\, \ell(I)^p}{p!}\, \|y_1 - y_2\|_\infty\,,$$

où $\ell(I)$ est la longueur (supposée finie) de I.

Comme $\frac{\alpha^p \ell(I)^p}{p!} \longrightarrow_{p\to+\infty} 0$, il existe $p \in \mathbb{N}^*$ tel que Φ^p est strictement contractante. Comme E est complet, un théorème de point fixe assure que Φ admet un unique point fixe, ou encore que le problème de CAUCHY admet une unique solution globale.

Supposons maintenant I quelconque. On sépare la preuve de l'existence et de l'unicité.

- Existence. Pour tout intervalle $J \subset I$ compact et contenant t_0, il existe, par ce qui précède, une unique solution y_J au problème de CAUCHY sur J. De plus, si J_1 et J_2 sont deux tels intervalles, alors y_{J_1} et y_{J_2} sont solutions du problème de CAUCHY sur l'intervalle compact $J_1 \cap J_2$ contenant t_0, donc coïncident par unicité sur $J_1 \cap J_2$.

  On peut alors définir une application $y : I \longrightarrow \mathbb{K}$ en posant $y(t) = y_J(t)$ pour $t \in I$, où J est un intervalle compact contenant t et t_0 (le choix de J étant arbitraire). La fonction y est solution de $y' = A\, y + B$ sur I (puisqu'elle coïncide localement avec une solution) et est telle que $y(t_0) = y_0$.

  Ceci assure l'existence d'une solution au problème de CAUCHY.
- Unicité. Supposons que y_1 et y_2 soient deux solutions du problème de CAUCHY sur I. Soit $t \in \mathbb{R}$. Choisissant un intervalle $J \subset I$ compact et contenant t et t_0, les fonctions y_1 et y_2 sont solutions du problème de CAUCHY sur J, donc $y_1(t) = y_2(t)$ par unicité de la solution sur J. Finalement, $y_1 = y_2$ et on conclut l'unicité de la solution sur I.

COMMENTAIRES

Il faut savoir ce qu'il se passe dans des cas plus généraux où y est solution de $y'(t) = f\big(t, y(t)\big)$ pour tout $t \in I$, avec $f : I \times \mathbb{K}^n \longrightarrow \mathbb{K}^n$ une fonction continue, et globalement ou localement lipschitzienne par rapport à la seconde variable. Dans le cas global, le résultat persiste, tandis que dans le cas local, il n'y a plus existence d'une solution globale, mais juste locale, et cette solution locale reste unique.

53 THÉORÈME DE CAUCHY-LIPSCHITZ (GLOBALEMENT LIPSCHITZIEN)

[Ber17, §3.2, p84]

ÉNONCÉ

THÉORÈME. [THÉORÈME DE CAUCHY-LIPSCHITZ (CAS GLOBALEMENT LIPSCHITZIEN)]
Soient I un intervalle de $\mathbb{R}$ et $f : I \times \mathbb{K}^n \longrightarrow \mathbb{K}^n$, où $n \in \mathbb{N}$ une fonction continue, et globalement lipchitzienne par rapport à la seconde variable pour une norme $\|\cdot\|$ sur $\mathbb{K}^n$.
À $(t_0, y_0) \in I \times \mathbb{K}^n$ fixé, il existe une unique solution y globale au problème de CAUCHY

$$\begin{cases} \forall t \in I & y'(t) = f\big(t, y(t)\big) \\ & y(t_0) = y_0 \,. \end{cases}$$

DÉVELOPPEMENT

Dans un premier temps, supposons que I est un intervalle compact de $\mathbb{R}$.
Soit $E = \mathcal{C}(I, \mathbb{K}^n)$. Notons que E est complet pour la norme infinie $\|\cdot\|_\infty$ associée à $\|\cdot\|$. Soit $k \geq 0$ la constante de lipschitziannité de f :

$$\forall t \in I \quad \forall (y_1, y_2) \in \big(\mathbb{K}^n\big)^2 \qquad \big\|f(t, y_1) - f(t, y_2)\big\| \leq k \, \|y_1 - y_2\| \,.$$

Enfin, soit N la norme définie par $N(y) = \sup_{t \in I} \big\|y(t)\big\| \, \mathrm{e}^{-k|t-t_0|}$ pour $y \in E$. L'espace E reste complet pour la norme N par équivalence de N et $\|\cdot\|_\infty$, puisque si $\ell(I)$ est la longueur de I :

$$\forall y \in E \qquad \mathrm{e}^{-k\ell(I)} \|y\|_\infty \leq N(y) \leq \|y\|_\infty \,.$$

Considérons l'application :

$$\begin{array}{rccc} \Phi : & E & \longrightarrow & E \\ & y & \longmapsto & \left(t \in I \longmapsto y_0 + \displaystyle\int_{t_0}^t f\big(s, y(s)\big) \, \mathrm{d}s\right), \end{array}$$

et montrons qu'elle est contractante pour N. Soient $y_1, y_2 \in E$ et $t \in I$. On a si $t \geq t_0$:

$$\begin{aligned} \big\|\Phi(y_1)(t) - \Phi(y_2)(t)\big\| \, \mathrm{e}^{-k|t-t_0|} &\leq \int_{t_0}^t \big\|f\big(s, y_1(s)\big) - f\big(s, y_2(s)\big)\big\| \, \mathrm{d}s \; \mathrm{e}^{-k|t-t_0|} \\ &\leq k \int_{t_0}^t \big\|y_1(s) - y_2(s)\big\| \, \mathrm{d}s \; \mathrm{e}^{-k|t-t_0|} \\ &\leq k \, N(y_1 - y_2) \int_{t_0}^t \mathrm{e}^{k|s-t_0| - k|t-t_0|} \, \mathrm{d}s \\ &\leq k \, N(y_1 - y_2) \int_{t_0}^t \mathrm{e}^{-k(t-s)} \, \mathrm{d}s \\ \big\|\Phi(y_1)(t) - \Phi(y_2)(t)\big\| \, \mathrm{e}^{-k|t-t_0|} &\leq N(y_1 - y_2)\big(1 - \mathrm{e}^{-k|t-t_0|}\big) \end{aligned}$$

On obtient le même résultat si $t \leq t_0$. Il en découle que :

$$\forall (y_1, y_2) \in E^2 \qquad N\big(\phi(y_1) - \phi(y_2)\big) \leq \big(1 - \mathrm{e}^{-k\ell(I)}\big) N(y_1 - y_2) < N(y_1 - y_2) \,.$$

Ainsi, Φ est strictement contractante pour (E, N).
Par un théorème de point fixe, Φ admet alors un unique point fixe.
En d'autres termes, le problème de CAUCHY admet une unique solution sur I.

Supposons maintenant I quelconque. On sépare la preuve de l'existence et de l'unicité.

- Existence. Pour tout intervalle $J \subset I$ compact et contenant t_0, il existe, par ce qui précède, une unique solution y_J au problème de CAUCHY sur J. De plus, si J_1 et J_2 sont deux tels intervalles, alors y_{J_1} et y_{J_2} sont solutions du problème de CAUCHY sur l'intervalle compact $J_1 \cap J_2$ contenant t_0, donc coïncident par unicité sur $J_1 \cap J_2$.

  On peut alors définir une application $y : I \longrightarrow \mathbb{K}$ en posant $y(t) = y_J(t)$ pour $t \in I$, où J est un intervalle compact contenant t et t_0 (le choix de J étant arbitraire). La fonction y est solution de $y'(t) = f\big(t, y(t)\big)$ pour tout $t \in I$ (puisqu'elle coïncide localement avec une solution) et est telle que $y(t_0) = y_0$.

  Ceci assure l'existence d'une solution au problème de CAUCHY.
- Unicité. Supposons que y_1 et y_2 soient deux solutions du problème de CAUCHY sur I. Soit $t \in \mathbb{R}$. Choisissant un intervalle $J \subset I$ compact et contenant t et t_0, les fonctions y_1 et y_2 sont solutions du problème de CAUCHY sur J, donc $y_1(t) = y_2(t)$ par unicité de la solution sur J. Finalement, $y_1 = y_2$ et on conclut l'unicité de la solution sur I.

COMMENTAIRES

Il faut savoir ce qu'il se passe dans le cas plus général d'une fonction f localement lipschitzienne en la seconde variable. Dans ce cadre, on doit introduire la notion de cylindre de sécurité pour montrer l'existence locale de solutions. L'existence d'une solution globale n'est plus garantie, mais l'unicité de la solution perdure : grâce au résultat local, on vérifie que deux solutions définies sur un même intervalle coïncident sur cet intervalle.

54 THÉORÈME DE FEJÉR

[QZ13, §IV.III.3, p84–85]

ÉNONCÉ

Soit $\mathbb{T} = \mathbb{R}/2\pi\mathbb{Z}$. Pour $n \in \mathbb{N}$, on pose D_n le noyau de DIRICHLET d'ordre n. Puis pour $N \in \mathbb{N}^*$, on définit $K_N = \frac{1}{N}\sum_{n=0}^{N} D_n$ le noyau de FEJÉR d'ordre N, et enfin on considèren, pour une fonction $f : \mathbb{T} \longrightarrow \mathbb{C}$ et lorsque cela à un sens, $\sigma_N(f) = f \star K_N$.

THÉORÈME. [THÉORÈME DE FEJÉR]

- *Si $f \in \mathcal{C}(\mathbb{T})$, alors $\|\sigma_N(f)\|_\infty \leq \|f\|_\infty$ pour $N \in \mathbb{N}^*$ et $\lim_{N\to+\infty}\|\sigma_N(f) - f\|_\infty = 0$.*
- *Si $f \in L^p(\mathbb{T})$ pour un $p \in [1\,;+\infty[$, alors $\|\sigma_N(f)\|_p \leq \|f\|_p$ pour $N \in \mathbb{N}^*$ et on a que $\lim_{N\to+\infty}\|\sigma_N(f) - f\|_p = 0$.*

DÉVELOPPEMENT

Soit $f \in \mathcal{C}(\mathbb{T})$. Pour $N \in \mathbb{N}^*$, on a $\|\sigma_N(f)\|_\infty = \|f\star K_N\|_\infty \leq \|f\|_\infty \cdot \|K_N\|_1 = \|f\|_\infty$ comme

$$\|K_N\|_1 = \frac{1}{2\pi}\int_{\mathbb{T}} K_N(t)\,\mathrm{d}t = \frac{1}{2\pi}\frac{1}{N}\sum_{n=0}^{N-1}\int_0^{2\pi} D_n(t)\,\mathrm{d}t = 1\,,$$

où l'on a utilisé la positivité de K_N et les propriétés des noyaux de DIRICHLET.

Introduisons l'application, appelée module de continuité de f, définie par

$$\omega : \delta \in \mathbb{R}_+ \longmapsto \sup\Big\{\big|f(u) - f(v)\big| : (u,v) \in \mathbb{R}^2 \text{ et } |v-u| \leq \delta\Big\}\,.$$

L'application ω est à valeurs dans $\mathbb{R}_+$ par uniforme continuité de f.

Soit $\delta > 0$. Pour $x \in \mathbb{T}$ et $N \in \mathbb{N}^*$, on a, en utilisant l'égalité $\int_{\mathbb{T}} K_N(t)\,\mathrm{d}t = 1$, l'inégalité triangulaire et le fait que $\sup_{|t|\in[\delta\,;\pi]} K_N(t) \leq \frac{1}{N\sin^2(\frac{\delta}{2})}$:

$$\begin{aligned}
\Big|\big(f - \sigma_N(f)\big)(x)\Big| &= \left|\frac{1}{2\pi}\int_0^{2\pi}\big(f(x) - f(x-t)\big)\,K_N(t)\,\mathrm{d}t\right| \\
&\leq \frac{1}{2\pi}\left(\int_{|t|\leq\delta}\big|f(x) - f(x-t)\big|\,K_N(t)\,\mathrm{d}t + \int_{\delta<|t|\leq\pi}\big|f(x) - f(x-t)\big|\,K_N(t)\,\mathrm{d}t\right) \\
&\leq \frac{1}{2\pi}\left(\omega(\delta)\int_{|t|\leq\delta} K_N(t)\,\mathrm{d}t + 2\,\|f\|_\infty\int_{\delta<|t|\leq\pi} K_N(t)\,\mathrm{d}t\right) \\
\Big|\big(f - \sigma_N(f)\big)(x)\Big| &\leq \omega(\delta) + 2\,\frac{\|f\|_\infty}{N\sin^2(\frac{\delta}{2})}\,.
\end{aligned}$$

Il vient finalement

$$\big\|f - \sigma_N(f)\big\|_\infty \leq \omega(\delta) + 2\,\frac{\|f\|_\infty}{N\sin^2(\frac{\delta}{2})}\,.$$

Ce raisonnement étant valable pour tout $\delta > 0$, on obtient en passant à la limite supérieure :

$$\limsup_{N\to+\infty}\big\|f - \sigma_N(f)\big\|_\infty \leq \inf_{\delta\in\mathbb{R}_+^*}\omega(\delta)\,.$$

Il s'ensuit que $\lim_{N\to+\infty}\|f - \sigma_N(f)\|_\infty = 0$ puisque $\omega(0) = 0$ et que, par uniforme continuité de f, ω est continue en 0. En effet[1] :

$$\forall\varepsilon > 0 \quad \exists\delta > 0 \qquad \omega(\delta) < \varepsilon \quad \text{et donc} \quad \omega \leq \varepsilon \text{ sur }]0\,;\delta]\,, \quad \text{d'où} \quad \omega(0^+) \leq \varepsilon\,.$$

Soit désormais $f \in L^p(\mathbb{T})$. Soit $N \in \mathbb{N}^*$.
D'après l'inégalité de JENSEN appliquée à la densité de probabilité $\frac{K_N}{2\pi}$:

$$\forall x \in \mathbb{T} \qquad \big|\sigma_N(f)(x)\big|^p = \left|\frac{1}{2\pi}\int_0^{2\pi} f(x-t)\,K_N(t)\,\mathrm{d}t\right|^p \leq \frac{1}{2\pi}\int_0^{2\pi}\big|f(x-t)\big|^p\,K_N(t)\,\mathrm{d}t\,.$$

Le théorème de FUBINI-TONELLI assure alors que :

$$\big\|\sigma_N(f)\big\|_p^p \leq \frac{1}{4\pi^2}\int_0^{2\pi}\left(\int_0^{2\pi}\big|f(x-t)\big|^p\,\mathrm{d}x\right)K_N(t)\,\mathrm{d}t = \|f\|_p^p\,\frac{1}{2\pi}\int_0^{2\pi}K_N(t)\,\mathrm{d}t = \|f\|_p^p\,.$$

En posant $g : t \in \mathbb{R} \longmapsto \|f - \tau_t f\|_p^p$, où τ_t est l'opération translatée par t, on obtient de même :

$$\begin{aligned}
\big\|f - \sigma_N(f)\big\|_p^p &\leq \frac{1}{4\pi^2}\int_0^{2\pi}\left(\int_0^{2\pi}\big|f(x) - f(x-t)\big|^p\,\mathrm{d}x\right)K_N(t)\,\mathrm{d}t \\
&\leq \frac{1}{2\pi}\int_0^{2\pi}K_N(t)\,g(t)\,\mathrm{d}t = \sigma_N(g)(0) \qquad \text{par parité de } K_N.
\end{aligned}$$

En admettant que g est continue, le premier point donne $\lim_{N\to+\infty}\|\sigma_N(f) - f\|_p = 0$.

COMMENTAIRES

Pour prouver la continuité de l'application g, on procède en trois étapes :

- d'abord, on justifie qu'il suffit de montrer la continuité 0,
- ensuite, on utilise le théorème de convergence dominée pour le montrer pour $f \in \mathcal{C}(\mathbb{T})$,
- enfin, on invoque la densité de $\mathcal{C}(\mathbb{T})$ dans L^p pour conclure[2].

1. $\omega(0^+)$ est bien définie puisque ω est croissante
2. plus généralement, le résultat est vrai sur $L^p(\mathbb{R}^d)$ et on passe alors par des fonctions $\mathcal{C}_c(\mathbb{R}^d)$

55 THÉORÈME DE SARD

[FGN07, §4.28, p253]

ÉNONCÉ

LEMME 1. *Pour tout ouvert U de $\mathbb{R}$, il existe une famille dénombrable d'intervalles ouverts $(U_i)_{i\in I}$ disjoints et tels que*

$$U = \bigcup_{i\in I} U_i\,.$$

THÉORÈME. [THÉORÈME DE SARD]
Soit $f : \mathbb{R} \longrightarrow \mathbb{R}$ une fonction de classe $\mathcal{C}^1$.
Si C désigne l'ensemble des zéros de f', alors $f(C)$ est de mesure nulle.

DÉVELOPPEMENT

Commençons par le lemme.
Soit U un ouvert de $\mathbb{R}$. On peut considérer la décomposition de U en composante connexes : il existe un ensemble I et des ensembles connexes $(U_i)_{i\in I}$ tels que

$$U = \bigcup_{i\in I} U_i\,.$$

Vérifions que la famille $(U_i)_{i\in I}$ convient.

- Déjà, pour tout $i \in I$, l'ensemble U_i est connexe donc est un intervalle de $\mathbb{R}$.
- Ensuite, comme U est ouvert, les $(U_i)_{i\in I}$ sont ouverts.
  En effet, soient $I \in I$ et $x \in U_i$. Il existe $\varepsilon > 0$ tel que $]x-\varepsilon\,;x+\varepsilon[\subset U$. L'ensemble $U_i \cup]x-\varepsilon\,;x+\varepsilon[$ est un connexe (non vide) de U contenant U_i. Étant donné que U_i est le plus grand connexe de U contenant x, cet ensemble est inclus dans U_i, ou encore $]x-\varepsilon\,;x+\varepsilon[\subset U_i$. Ainsi, U_i est ouvert.
- Enfin, si $i \in I$, alors U_i ouvert non vide contient un rationnel q_i. L'application $i \in I \longmapsto q_i$ étant injective puisque les $(U_i)_{i\in I}$ sont disjoints, l'ensemble I est (au plus) dénombrable.

Ceci prouve le lemme.

Passons au théorème de SARD. On procède en deux étapes.

1. Fixons $N \in \mathbb{N}^*$ et posons $I_N = [-N\,;N]$.
   Étudions f sur I_N en montrant que $f(C\cap I_N)$ est de mesure nulle.
   Soit $\varepsilon > 0$. On définit :
   $$A = \big\{x \in I_N : \big|f'(x)\big| < \varepsilon\big\}\,.$$
   Par continuité de f', l'ensemble A est un ouvert de I_N. On peut donc considérer la décomposition en intervalles ouverts disjoints, notée $(A_i)_{i\in I}$:
   $$A = \bigcup_{i\in I} A_i\,.$$
   Soit $i \in I$. D'après l'inégalité des accroissements finis, l'application f est ε-lipschitzienne sur A_i, si bien que $f(A_i)$ est un intervalle de longueur au plus $\varepsilon\,\mathrm{Leb}(A_i)$.
   Comme $C\cap I_N \subset A \subset I_N$, il vient alors :
   $$\mathrm{Leb}\big(f(C\cap I_N)\big) \leq \mathrm{Leb}\Big(f\Big(\bigcup_{i\in I} A_i\Big)\Big) \leq \sum_{i\in I} \mathrm{Leb}\big(f(A_i)\big) \leq \varepsilon\,\mathrm{Leb}(A) \leq \varepsilon\,\mathrm{Leb}(I_N)\,.$$
   Faisant tendre ε vers 0, on obtient que $f(C\cap I_N)$ est de mesure nulle.
2. Reste à en déduire que $f(C)$ est de mesure nulle.
   La suite de mesurables $(C\cap I_N)_{N\in\mathbb{N}^*}$ étant une suite croissante, il en est de même de $\big(f(C\cap I_N)\big)_{N\in\mathbb{N}^*}$, et on obtient par continuité croissante de la mesure de LEBESGUE :
   $$\mathrm{Leb}\big(f(C)\big) = \mathrm{Leb}\Big(\lim_{N\to+\infty} \uparrow f\big(C\cap I_N\big)\Big) = \lim_{N\to+\infty} \uparrow \mathrm{Leb}\big(f(C\cap I_N)\big) = 0\,.$$

COMMENTAIRES

Ce développement est un peu court, il laisse le temps pour faire des dessins explicatifs au tableau, notamment pour l'idée principale du développement qui consiste à montrer qu'une dérivée nulle en un point aplatit tout voisinage de ce point lorsque l'on prend son image par f.

On peut conclure en annonçant que le résultat est vrai lorsque f' n'est pas continue, mais c'est plus difficile à montrer (voir [GT98]).

56 THÉORÈME DE STONE-WEIERSTRASS

[HL09, Ch1, p26–30]

ÉNONCÉ

Soit (X,d) un compact non vide.

THÉORÈME. [THÉORÈME DE STONE-WEIERSTRASS]
Soit H une sous-algèbre de $\mathcal{C}(X,\mathbb{R})$ séparante et unitaire. Alors H est dense dans $\mathcal{C}(X,\mathbb{R})$.

DÉVELOPPEMENT

On élimine directement le cas où X ne possède qu'un élément, le résultat étant alors clair puisque H contient les fonctions constantes (H étant unitaire). Supposons donc $|X| \geq 2$.

1. Commençons par montrer que
$$\exists (P_n)_{n\in\mathbb{N}} \in \mathbb{R}[X]^{\mathbb{N}} \qquad P_n \underset{n\to+\infty}{\longrightarrow} |\cdot| \text{ uniformément sur } [-1\,;1]\,.$$
En effet, définissons $(P_n)_{n\in\mathbb{N}}$ par récurrence par
$$P_0 = 0 \qquad \text{et} \qquad \forall n\in\mathbb{N} \quad P_{n+1} = P_n + \frac{1}{2}\big(X^2 - P_n^2\big)\,.$$
Vérifions par récurrence que pour tout entier n, $0 \leq P_n \leq P_{n+1} \leq |\cdot|$ sur $[-1\,;1]$.
   - Initialisation. Pour $n=0$, on a $P_1 = \frac{X^2}{2}$ donc $0 = P_0 \leq P_1 \leq |\cdot|$.
   - Hérédité. Si le résultat vrai au rang n, de $P_{n+1} \leq |\cdot|$ on tire $P_{n+2} - P_{n+1} \geq 0$ et
$$\begin{aligned}\forall x\in[-1\,;1] \qquad |x| - P_{n+2}(x) &= \big(|x| - P_{n+1}(x)\big) - \tfrac{1}{2}\big(|x|^2 - P_{n+1}(x)^2\big)\\ &= \big(|x| - P_{n+1}(x)\big)\Big(1 - \tfrac{1}{2}\big(|x| + P_{n+1}(x)\big)\Big) \geq 0\,.\end{aligned}$$
   La suite $(P_n)_{n\in\mathbb{N}}$ est ainsi croissante et majorée. Elle converge et sa limite h vérifie, pour
$$\forall x\in[-1\,;1] \qquad 0 \leq h(x) \leq |x| \quad \text{et} \quad h(x) = h(x) + \frac{1}{2}\big(x^2 - h(x)^2\big)\,,$$
   soit en fait $h = |\cdot|$. On a obtenu que $(P_n)_{n\in\mathbb{N}}$ converge vers $|\cdot|$ sur $[-1\,;1]$, et de plus cette convergence est uniforme d'après le théorème de DINI.

2. Montrons que si $f,g\in\overline{H}$, alors $\min(f,g)$ et $\max(f,g)$ sont dans $\overline{H}$.
Déjà, si $h\in\overline{H}$, alors $|h|\in\overline{H}$. En effet, si $h\neq 0$, alors $P_n\big(\frac{h}{\|h\|_\infty}\big) \in \overline{H}$ pour tout $n\in\mathbb{N}$ puisque H (et donc $\overline{H}$) est une sous-algèbre. Par la convergence uniforme obtenue ci-dessus, on en déduit que $\frac{|h|}{\|h\|_\infty}\in\overline{H}$, puis $|h|\in\overline{H}$.
Revenons à $f,g\in H$. Le résultat découle alors de l'écriture
$$\max(f,g) = \frac{f+g}{2} + \frac{|f-g|}{2} \in \overline{H} \quad \text{et} \quad \min(f,g) = \frac{f+g}{2} - \frac{|f-g|}{2} \in \overline{H}\,.$$

3. Montrons ensuite que pour $x_1,x_2\in X$ distincts et $\alpha_1,\alpha_2\in\mathbb{R}$ fixés
$$\exists u\in H \qquad u(x_1) = \alpha_1 \quad \text{et} \quad u(x_2) = \alpha_2\,.$$
En effet, comme H est séparante, il existe $u_0\in H$ tel que $u_0(x_1)\neq u_0(x_2)$. Le système
$$\begin{cases}\lambda\, u_0(x_1) + \mu = \alpha_1\\ \lambda\, u_0(x_2) + \mu = \alpha_2\end{cases}$$
est donc de CRAMER : il admet une solution $(\lambda_\star,\mu_\star)$ et $u = \lambda_\star\, u_0 + \mu_\star \in H$ convient.

4. Montrons finalement que H est dense dans $\mathcal{C}(X,\mathbb{R})$.
Soit $f\in\mathcal{C}(X,\mathbb{R})$. Soit $\varepsilon>0$. Montrons qu'il existe $v\in\overline{H}$ tel que $\|v-f\|_\infty\leq\varepsilon$.
Soit $x\in X$. D'après le point précédent
$$\forall y\in X \quad \exists u_y\in H \qquad u_y(x) = f(x) \quad \text{et} \quad u_y(y) = f(y)\,.$$
Considérons alors, pour $y\in X$, l'ouvert (par continuité de u_y et de f)
$$\mathcal{O}_y = \big\{x'\in X : u_y(x') > f(x') - \varepsilon\big\}$$
contenant x et y. On peut ainsi écrire que $X = \bigcup_{y\in X}\mathcal{O}_y$, d'où, par compacité de X,
$$\exists r\in\mathbb{N} \quad \exists (y_i)_{1\leq i\leq r}\in X^r \qquad X = \bigcup_{i=1}^{r}\mathcal{O}_{y_i}\,.$$
Posons alors $v_x = \max_{1\leq i\leq r} u_{y_i}$, qui est un élément de $\overline{H}$ par le second point, satisfaisant $v_x(x) = f(x)$ et $v_x(x') > f(x') - \varepsilon$ pour tout $x'\in X$. Soit alors l'ouvert
$$\Omega_x = \big\{x'\in X : v_x(x') < f(x') + \varepsilon\big\}$$
contenant x. On en déduit, toujours par compacité et comme $X = \cup_{x\in X}\Omega_x$, que
$$\exists m\in\mathbb{N} \quad \exists (x_j)_{1\leq j\leq m}\in X^m \qquad X = \bigcup_{j=1}^{m}\Omega_{x_j}\,.$$
Reste à poser $v = \min_{1\leq j\leq m} v_{x_j}$, qui est encore un élément de $\overline{H}$ et qui vérifie :
$$\forall x\in X \quad \big|v(x) - f(x)\big| \leq \varepsilon \qquad \text{ou encore} \qquad \|v-f\|_\infty \leq \varepsilon\,.$$
Le raisonnement étant valable pour tout $\varepsilon>0$, on en déduit que $f\in\overline{H}$.
Autrement dit, $\overline{H} = \mathcal{C}(X,\mathbb{R})$ et H est dense.

COMMENTAIRES

Dans la première partie, on utilise le théorème de DINI, qu'il faut donc savoir démontrer.

Une extension de ce résultat dans le cas complexe se trouve dans [HL09].

57 THÉORÈME DE WEIERSTRASS

[QZ13, §XIII.II.1, p518–519]

ÉNONCÉ

THÉORÈME. [THÉORÈME DE WEIERSTRASS]
Soient $a, b \in \mathbb{R}$ avec $a < b$. Alors $\mathbb{R}[X]$ est dense dans $\mathcal{C}\big([a\,;b], \mathbb{C}, \|\cdot\|_\infty\big)$.

DÉVELOPPEMENT

On suppose que $[a\,;b] = [0\,;1]$, quitte à dilater par $t \in [0\,;1] \longmapsto a + (b-a)\,t$.
Soit $f \in \mathcal{C}([0\,;1], \mathbb{R})$. Introduisons ω_f le module de continuité de f défini par :

$$\omega_f : \delta \in \mathbb{R}_+ \longmapsto \sup\Big\{ \big|f(u) - f(v)\big| : (u,v) \in \mathbb{R}^2 \text{ et } |v-u| \leq \delta \Big\}.$$

L'application ω_f est à valeurs dans $\mathbb{R}_+$ par uniforme continuité de f, et vérifie les propriétés :

(i) ω_f est croissante, vérifie $\omega_f(0) = 0$ et est continue en 0 ;
(ii) Pour tout $\lambda > 0$ et tout $\delta > 0$, on a $\omega_f(\lambda\,\delta) \leq (\lambda+1)\,\omega_f(\delta)$.

Utilisons ces propriétés pour montrer le résultat. Pour $x \in [0\,;1]$, soit $(X_{i,x})_{i\in\mathbb{N}}$ des variables aléatoires i.i.d. de loi $\mathcal{B}(x)$, puis posons $S_{n,x} = \sum_{i=1}^n X_{i,x}$ pour $n \in \mathbb{N}^*$. D'après le lemme de transfert, on peut définir, pour $n \in \mathbb{N}^*$, le n^{e} polynôme de BERNSTEIN par

$$\forall x \in [0\,;1] \qquad B_n(f)(x) = \mathbb{E}\Big[f\Big(\frac{S_{n,x}}{n}\Big)\Big] = \sum_{k=0}^n \binom{n}{k} x^k\,(1-x)^{n-k}\,f\Big(\frac{k}{n}\Big),$$

Montrons que $B_n(f) \longrightarrow_{n\to+\infty} f$ uniformément sur $[0\,;1]$ en majorant $\big\|f - B_n(f)\big\|_{\infty,\,[0\,;1]}$ par le module de continuité w. Soient $x \in [0\,;1]$ et $n \in \mathbb{N}^*$. Il vient d'abord

$$\begin{aligned}\big|f - B_n(f)\big|(x) &= \Big|\mathbb{E}\Big[f(x) - f\Big(\frac{S_{n,x}}{n}\Big)\Big]\Big| \leq \mathbb{E}\Big[\Big|f(x) - f\Big(\frac{S_{n,x}}{n}\Big)\Big|\Big] \leq \mathbb{E}\Big[\omega_f\Big(\Big|x - \frac{S_{n,x}}{n}\Big|\Big)\Big] \\ &\leq \mathbb{E}\Big[\Big(1 + \sqrt{n}\Big|x - \frac{S_{n,x}}{n}\Big|\Big)\,\omega_f\Big(\frac{1}{\sqrt{n}}\Big)\Big] \leq \Big(1 + \sqrt{n}\,\mathbb{E}\Big[\Big|x - \frac{S_{n,x}}{n}\Big|\Big]\Big)\,\omega_f\Big(\frac{1}{\sqrt{n}}\Big)\end{aligned}$$

par les propriétés de ω_f. Or, d'après l'inégalité de CAUCHY-SCHWARZ :

$$\mathbb{E}\Big[\Big|x - \frac{S_{n,x}}{n}\Big|\Big] \leq \mathbb{E}\big[1^2\big]^{\frac12} \times \mathbb{E}\Big[\Big(x - \frac{S_{n,x}}{n}\Big)^2\Big]^{\frac12} = \mathbb{E}\Big[\Big(x - \frac{S_{n,x}}{n}\Big)^2\Big]^{\frac12},$$

puis, étant donné que $x - \frac{S_{n,x}}{n}$ est centrée et que les $(X_{i,x})_{i\in\mathbb{N}}$ sont indépendantes, on calcule :

$$\mathbb{E}\Big[\Big(x - \frac{S_{n,x}}{n}\Big)^2\Big] = \mathrm{Var}\Big(\frac{S_{n,x}}{n}\Big) = \frac{1}{n^2}\sum_{i=1}^n \mathrm{Var}(X_{i,x}) = \frac{x\,(1-x)}{n} \leq \frac{1}{4n},$$

si bien que finalement $\big\|f - B_n(f)\big\|_{\infty,\,[0\,;1]} \leq \frac32\,\omega_f\big(\frac{1}{\sqrt{n}}\big)$. Le résultat en découle par la propriété (ii), avec également un contrôle de la vitesse de convergence par le module de continuité.

Vérifions les propriétés de ω_f.

(ii) La limite $\omega_f(0^+)$ est bien définie puisque ω_f est croissante. De plus :

$$\forall \varepsilon > 0 \quad \exists \delta > 0 \qquad \omega_f(\delta) < \varepsilon \quad \text{et donc} \quad \omega_f \leq \varepsilon \text{ sur }]0\,;\delta], \quad \text{d'où} \quad \omega_f(0^+) \leq \varepsilon\,.$$

(iii) Soient $\lambda > 0$ et $\delta > 0$. Si $u, v \in \mathbb{R}$ sont tels que $|v-u| \leq \lambda\,\delta$ et $u \leq v$, alors il existe[1] $m \leq \lambda + 1$ et une subdivision $u = x_0 < x_1 < \cdots < x_m = v$ de pas inférieur à δ, et il vient

$$\big|f(v) - f(u)\big| \leq \sum_{i=0}^{m-1} \big|f(x_{i+1}) - f(x_i)\big| \leq m\,\omega_f(\delta) \leq \big(\lfloor\lambda\rfloor + 1\big)\,\omega_f(\delta) \leq (\lambda+1)\,\omega_f(\delta)\,.$$

Notons que si f est k-lipschitzienne, alors $\omega_f \leq \mathrm{id}$ et la vitesse de convergence des polynômes de BERNSTEIN est au pire en $\frac{1}{\sqrt{n}}$. Vérifions que cette vitesse est optimale en toute généralité.
Considérons l'application 1-lipschitzienne $f : x \in [0\,;1] \longmapsto \big|x - \frac12\big|$, qui vérifie $\omega_f = \mathrm{id}$. On a :

$$\big\|f - B_n(f)\big\|_{\infty,\,[0\,;1]} \geq \big|f - B_n(f)\big|\Big(\frac12\Big) = \Big|B_n(f)\Big(\frac12\Big)\Big| = \mathbb{E}\Big[\Big|\frac{S_{n,\frac12}}{n} - \frac12\Big|\Big] = \frac{\mathbb{E}\Big[\big|2S_{n,\frac12} - n\big|\Big]}{2n}\,.$$

Notons que $2S_{n,\frac12} - n = \sum_{i=1}^n 2X_{i,\frac12} - 1 = \sum_{i=1}^n \varepsilon_i$, où, pour $i \in \mathbb{N}^*$, $\varepsilon_i = 2X_{i,\frac12} - 1 \sim \mathcal{R}\big(\frac12\big)$.
Définissant la variable aléatoire $Y = \prod_{j=1}^n \big(1 + \mathrm{i}\frac{\varepsilon_j}{\sqrt{n}}\big)$, on calcule comme $\varepsilon_j^2 = 1$ pour $j \in [\![1\,;n]\!]$:

$$|Y| = \Big(\sqrt{1 + \frac{1}{n}}\Big)^n \leq \Big(\sqrt{\mathrm{e}^{\frac1n}}\Big)^n \leq \sqrt{\mathrm{e}} \qquad \text{d'où} \qquad \mathbb{E}\Big[\big|2S_{n,\frac12} - n\big|\Big] \geq \frac{\mathbb{E}\Big[\big|2S_{n,\frac12} - n\big| \cdot |Y|\Big]}{\sqrt{\mathrm{e}}}\,.$$

Or, comme les $(\varepsilon_j)_{j\in[\![1\,;n]\!]}$ sont i.i.d.,

$$\mathbb{E}\Big[\big(2S_{n,\frac12} - n\big)Y\Big] = n\,\mathbb{E}\Big[\varepsilon_1\Big(1 + \mathrm{i}\frac{\varepsilon_1}{\sqrt{n}}\Big)\prod_{j=2}^n\Big(1 + \mathrm{i}\frac{\varepsilon_j}{\sqrt{n}}\Big)\Big] = n\left(\mathbb{E}[\varepsilon_1] + \mathrm{i}\frac{\mathbb{E}[\varepsilon_1^2]}{\sqrt{n}}\right)\prod_{j=2}^n\Big(1 + \mathrm{i}\frac{\mathbb{E}[\varepsilon_j]}{\sqrt{n}}\Big)\,.$$

Les $(\varepsilon_j)_{j\in[\![1\,;n]\!]}$ étant centrées et de carré 1, il vient $\mathbb{E}\Big[\big(2S_{n,\frac12} - n\big)Y\Big] = \mathrm{i}\sqrt{n}$, puis finalement

$$\big\|f - B_n(f)\big\|_{\infty,\,[0\,;1]} \geq \frac{1}{2n}\sqrt{\frac{n}{\mathrm{e}}} = \frac{1}{2\sqrt{\mathrm{e}}\sqrt{n}} = \frac{1}{2\sqrt{\mathrm{e}}}\,\omega_f\Big(\frac{1}{\sqrt{n}}\Big)\,.$$

COMMENTAIRES

Le résultat ne peut être étendu à un intervalle non borné : en effet si $(P_n)_{n\in\mathbb{N}} \in \mathbb{R}[X]^{\mathbb{N}}$ converge uniformément sur $\mathbb{R}$ vers f continue, alors c'est une suite de CAUCHY pour la norme infinie, donc à $\varepsilon > 0$ fixé, il existe un rang N_0 tel que $\|P_{n+1} - P_n\|_\infty \leq \varepsilon$ pour $n \geq N_0$, et alors $P_{n+1} - P_n$ est un polynôme constant. Il s'ensuite que $f = P_{N_0} + \sum_{n=N_0}^{+\infty} P_{n+1} - P_n$ est aussi un polynôme.

1. faire un graphique au tableau

58 THÉORÈME DES EXTREMA LIÉS

[Gou08, §5.2/3, p317/327]

ÉNONCÉ

THÉORÈME. [THÉORÈME DES EXTREMA LIÉS]
Soient $r, n \in \mathbb{N}^*$ et $f, g_1, \dots, g_r : U \longrightarrow \mathbb{R}$ des fonctions de classe $\mathcal{C}^1$ définies sur un ouvert U de $\mathbb{R}^n$. Notons $\Gamma = \{x \in U : g_1(x) = \cdots = g_r(x) = 0\}$. Si $f_{|\Gamma}$ admet un extremum local en $a \in \Gamma$ et si la famille de formes linéaires $\big(dg_1(a), \dots, dg_r(a)\big)$ est libre, alors

$$\exists!(\lambda_i)_{1\leq i\leq r} \in \mathbb{R}^r \qquad df(a) = \sum_{i=1}^{r} \lambda_i \, dg_i(a)\,.$$

Les réels $(\lambda_i)_{1\leq i\leq r}$ sont appelés multiplicateurs de LAGRANGE.

DÉVELOPPEMENT

Soit $a \in \Gamma$ un extremum local de $f_{|\Gamma}$ tel que $\big(dg_1(a), \dots, dg_r(a)\big)$ soit libre.

Notons déjà que, nécessairement, $r \leq n$, et qu'en cas d'égalité, la famille est une base du dual de $\mathbb{R}^n$ et donc le résultat est évident.

Supposons donc $r < n$ et posons $s = n - r \geq 1$. En identifiant $\mathbb{R}^n$ et $\mathbb{R}^s \times \mathbb{R}^r$, on écrira ses éléments sous la forme $(x, y) = (x_1, \dots, x_s, y_1, \dots, y_r)$. On note $a = (\alpha, \beta)$. La matrice

$$M_a = \begin{pmatrix} \partial_{x_1} g_1(a) & \cdots & \partial_{x_s} g_1(a) & \partial_{y_1} g_1(a) & \cdots & \partial_{y_r} g_1(a) \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \partial_{x_1} g_r(a) & \cdots & \partial_{x_s} g_r(a) & \partial_{y_1} g_r(a) & \cdots & \partial_{y_r} g_r(a) \end{pmatrix}$$

est de rang r par hypothèse de liberté de $\big(dg_1(a), \dots, dg_r(a)\big)$. On peut en extraire une matrice carrée de taille r inversible, qui, quitte à renommer les variables, est la matrice

$$M_{a,2} = \begin{pmatrix} \partial_{y_1} g_1(a) & \cdots & \partial_{y_r} g_1(a) \\ \vdots & \ddots & \vdots \\ \partial_{y_1} g_r(a) & \cdots & \partial_{y_r} g_r(a) \end{pmatrix}.$$

On peut appliquer le théorème des fonctions implicites à $g = (g_1, \dots, g_r)$ en a puisque, par ce qui précède, g est $\mathcal{C}^1$, $g(a) = g\big((\alpha, \beta)\big) = 0$ et $\partial_2 g(\alpha, \beta)$ est inversible. Il existe donc φ de classe $\mathcal{C}^1$ telle que $g\big((x, y)\big) = 0$ au voisinage de $a = (\alpha, \beta)$ si et seulement si $y = \varphi(x)$, c'est-à-dire localement $(x, y) \in \Gamma \iff y = \varphi(x)$.

Posons $h : x \longmapsto f(x, \varphi(x))$ au voisinage de α. L'application h admet un extremum local en α puisque $\big(\alpha, \varphi(\alpha)\big) = a$ et $\big(x, \varphi(x)\big) \in \Gamma$ au voisinage de α. Si $u : x \longmapsto \big(x, \varphi(x)\big)$, alors h est différentiable en α par composition et $0 = dh(\alpha) = df\big(u(\alpha)\big) \circ du(\alpha)$, soit matriciellement :

$$0 = \begin{pmatrix} \partial_{x_1} f(a) & \cdots & \partial_{x_s} f(a) & \partial_{y_1} f(a) & \cdots & \partial_{y_r} f(a) \end{pmatrix} \begin{pmatrix} 1 & & & \\ & \ddots & & \\ & & 1 & \\ \partial_{x_1}\varphi_1(\alpha) & \cdots & \partial_{x_s}\varphi_1(\alpha) \\ \vdots & \ddots & \vdots \\ \partial_{x_1}\varphi_r(\alpha) & \cdots & \partial_{x_s}\varphi_r(\alpha) \end{pmatrix}$$

$$= \begin{pmatrix} \partial_{x_1} f(a) + \sum_{j=1}^{r} \partial_{x_1}\varphi_j(\alpha)\, \partial_{y_j} f(a) \\ \vdots \\ \partial_{x_s} f(a) + \sum_{j=1}^{r} \partial_{x_s}\varphi_j(\alpha)\, \partial_{y_j} f(a) \end{pmatrix}.$$

Ainsi, on obtient que $0 = \partial_{x_i} h(\alpha) = \partial_{x_i} f(a) + \sum_{j=1}^{r} \partial_{x_i}\varphi_j(\alpha)\, \partial_{y_j} f(a)$ pour tout $i \in [\![1\,;s]\!]$.

Par ailleurs, comme $g_k\big(x, \varphi(x)\big) = 0$ pour tout $k \in [\![1\,;r]\!]$, on a de même :

$$\forall k \in [\![1\,;r]\!] \quad \forall i \in [\![1\,;s]\!] \qquad 0 = \partial_{x_i} g_k(a) + \sum_{j=1}^{r} \partial_{x_i}\varphi_j(\alpha)\, \partial_{y_j} g_k(a)\,.$$

Considérons alors la matrice

$$M = \begin{pmatrix} \partial_{x_1} f(a) & \cdots & \partial_{x_s} f(a) & \partial_{y_1} f(a) & \cdots & \partial_{y_r} f(a) \\ \partial_{x_1} g_1(a) & \cdots & \partial_{x_s} g_1(a) & \partial_{y_1} g_1(a) & \cdots & \partial_{y_r} g_1(a) \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \partial_{x_1} g_r(a) & \cdots & \partial_{x_s} g_r(a) & \partial_{y_1} g_r(a) & \cdots & \partial_{y_r} g_r(a) \end{pmatrix},$$

dont on note les colonnes $(C_k)_{1\leq k\leq n}$ et les lignes $(L_i)_{0\leq i\leq r}$. Par ce qui précède, les s premières colonnes de M sont des combinaisons linéaires des r dernières ($C_k = \sum_{j=1}^{r} \partial_{x_k}\varphi_j(\alpha)\, C_{s+j}$ pour $k \in [\![1\,;s]\!]$), donc M est de rang au plus r.

Les lignes de M sont alors liées. Comme par hypothèse les r dernières lignes sont libres, la première est combinaison linéaire des autres et

$$\exists(\lambda_i)_{1\leq i\leq r} \in \mathbb{R}^r \qquad L_0 = \sum_{i=1}^{r} \lambda_i L_i \quad \text{ou encore} \quad df(a) = \sum_{i=1}^{r} \lambda_i \, dg_i(a)\,,$$

ce qui conclut l'existence des multiplicateurs de LAGRANGE. Par ailleurs, leur unicité est claire puisque $\big(dg_1(a), \dots, dg_r(a)\big)$ est une famille libre.

COMMENTAIRES

Il faut faire un joli dessin pour expliquer l'intuition, en prenant par exemple pour Γ la sphère unité et pour f une application linéaire.

59 THÉORÈMES D'ABEL ET TAUBÉRIEN FAIBLE

[Gou08, §4.4, p252–254]

ÉNONCÉ

Soit $(a_n)_{n\in\mathbb{N}} \in \mathbb{C}^{\mathbb{N}}$.

THÉORÈME. [THÉORÈME D'ABEL]
Soit la série entière $f : z \longmapsto \sum_{n\geq 0} a_n z^n$. On suppose que le rayon de convergence de f est 1 et que la série de terme général a_n converge. Alors, pour tout $\alpha \in \left[0\,;\,\frac{\pi}{2}\right[$,

$$\lim_{z\to 1,\, z\in\Delta_\alpha} f(z) = \sum_{n=0}^{+\infty} a_n\,, \qquad \text{où} \quad \Delta_\alpha = \left\{1-\rho\,\mathrm{e}^{\mathrm{i}\theta} \in \mathbb{D}(0,1) : \rho>0, \theta\in[-\alpha\,;\,\alpha]\right\}.$$

THÉORÈME. [THÉORÈME TAUBÉRIEN FAIBLE]
Soit la série entière $f : z \longmapsto \sum_{n\geq 0} a_n z^n$. On suppose que le rayon de convergence de f est 1, que $a_n =_{n\to+\infty} o\left(\frac{1}{n}\right)$ et qu'il existe $S\in\mathbb{C}$ tel que $f(x) \longrightarrow_{x\to 1^-} S$.
Alors la série de terme général a_n converge et $S = \sum_{n=0}^{+\infty} a_n$.

DÉVELOPPEMENT

Commençons par montrer le théorème d'ABEL. Notons $S = \sum_{n=0}^{+\infty} a_n$ et fixons $\alpha \in \left[0\,;\,\frac{\pi}{2}\right[$. Pour $n\in\mathbb{N}$, on pose $S_n = \sum_{k=0}^{n} a_k$ et $R_n = S - S_n$. Soit $z\in\mathbb{D}(0,1)$. Alors pour tout $N\in\mathbb{N}^*$

$$\begin{aligned}\sum_{n=0}^{N} a_n z^n - S_N &= \sum_{n=1}^{N}(R_{n-1}-R_n)(z^n-1) = \sum_{n=0}^{N-1} R_n(z^{n+1}-1) - \sum_{n=1}^{N} R_n(z^n-1)\\ &= \sum_{n=0}^{N-1} R_n(z^{n+1}-z^n) - R_N(z^N-1) = (z-1)\sum_{n=0}^{N-1} R_n z^n - R_N(z^N-1)\end{aligned}$$

d'où l'on obtient en passant à la limite $f(z) - S = (z-1)\sum_{n=0}^{+\infty} R_n z^n$.

Fixons alors $\varepsilon>0$ puis $N\in\mathbb{N}$ tel que $|R_n|<\varepsilon$ pour $n>N$.

$$\forall z\in\mathbb{D}(0,1) \qquad \left|f(z)-S\right| \leq |z-1|\left(\sum_{n=0}^{N}|R_n z^n| + \varepsilon\sum_{n=N+1}^{+\infty}|z|^n\right) \leq |z-1|\sum_{n=0}^{N}|R_n| + \varepsilon\frac{|z-1|}{1-|z|}\,.$$

Prenons désormais $z = 1-\rho\,\mathrm{e}^{\mathrm{i}\theta}\in\Delta_\alpha$ avec $\rho\leq\cos(\alpha)$[1]. Puisque $|z|^2 = 1-2\rho\cos(\theta)+\rho^2$,

$$\frac{|z-1|}{1-|z|} = (1+|z|)\frac{|z-1|}{1-|z|^2} = \frac{2\rho}{2\rho\cos(\theta)-\rho^2} \leq \frac{2}{2\cos(\theta)-\rho} \leq \frac{2}{2\cos(\alpha)-\cos(\alpha)} \leq \frac{2}{\cos(\alpha)}\,.$$

1. la restriction à Δ_α est ici cruciale : on n'aurait pas cette majoration dans tout $\mathbb{D}(0,1)$ entier (prendre $\theta\to\pm\frac{\pi}{2}$)

Choisissons $\rho_\varepsilon>0$ tel que $\rho_\varepsilon\sum_{n=0}^{N}|R_n|\leq\varepsilon$. Si $\rho\leq\rho_\varepsilon$, on obtient $\left|f(z)-S\right|\leq\varepsilon+\varepsilon\frac{2}{\cos(\alpha)}$. Autrement dit, on a obtenu que

$$\forall z\in\Delta_\alpha \qquad |z-1|\leq\rho_\varepsilon \quad\Longrightarrow\quad \left|f(z)-S\right|\leq\varepsilon+\varepsilon\frac{2}{\cos(\alpha)}\,.$$

Le raisonnement étant valable pour tout $\varepsilon>0$, le résultat en découle.

Passons au théorème taubérien faible. On note encore $S_n = \sum_{k=0}^{n} a_k$ pour $n\in\mathbb{N}$. Remarquons que pour $k\in\mathbb{N}$ et $x\in]0\,;1[$, on a $(1-x^k) = (1-x)\sum_{i=0}^{k-1} x^i \leq k(1-x)$, d'où en posant $M = \sup_{k>0} k|a_k|$ (fini par hypothèse) :

$$\begin{aligned}\forall n\in\mathbb{N}\quad \forall x\in]0\,;1[\qquad S_n - f(x) &= \sum_{k=1}^{n} a_k(1-x^k) - \sum_{k=n+1}^{+\infty} a_k x^k\\ \left|S_n-f(x)\right| &\leq \sum_{k=1}^{n}|a_k|k(1-x) + \sum_{k=n+1}^{+\infty}\frac{k}{n}|a_k|x^k\\ &\leq (1-x)Mn + \frac{\sup_{k>n} k|a_k|}{n}\sum_{k=n+1}^{+\infty} x^k\end{aligned}$$

$$\forall n\in\mathbb{N}\quad \forall x\in]0\,;1[\qquad \left|S_n-f(x)\right| \leq (1-x)Mn + \frac{\sup_{k>n} k|a_k|}{n(1-x)}\,.$$

Fixons $\varepsilon\in]0\,;1[$. En appliquant en $x = 1-\frac{\varepsilon}{n}$ à $n\in\mathbb{N}^*$ fixé, il vient

$$\forall n\in\mathbb{N}^* \qquad \left|S_n - f\left(1-\frac{\varepsilon}{n}\right)\right| \leq M\varepsilon + \frac{\sup_{k>n} k|a_k|}{\varepsilon}\,,$$

d'où il découle, puisque $a_n =_{n\to+\infty} o\left(\frac{1}{n}\right)$,

$$\exists N_0\in\mathbb{N}^*\quad \forall n\geq N_0 \qquad \left|S_n - f\left(1-\frac{\varepsilon}{n}\right)\right| \leq \varepsilon(M+1)\,.$$

Or, $f(x)\longrightarrow_{x\to1^-} S$ donc on peut choisir $N_1\in\mathbb{N}^*$ tel que $\left|S - f\left(1-\frac{\varepsilon}{n}\right)\right|<\varepsilon$ pour $n\geq N_1$. Ainsi, d'après l'inégalité triangulaire,

$$\forall n\geq\max(N_0,N_1) \qquad |S_n - S|\leq\varepsilon(M+2)\,.$$

Ceci étant valable pour tout $\varepsilon\in]0\,;1[$, on obtient que $(S_n)_{n\in\mathbb{N}}$ converge et que sa limite est S.

COMMENTAIRES

Penser à faire un dessin de Δ_α en début de développement pour expliquer les passages importants de la démonstration du théorème d'ABEL.

BIBLIOGRAPHIE ANALYSE ET PROBABILITÉS

[AK02] G. Allaire et S.-M. Kaber : *Algèbre linéaire numérique*. Ellipses, 2002.

[Ber17] F. Berthelin : *Équations différentielles*. Cassini, 2017.

[BL07] P. Barbe et M. Ledoux : *Probabilité*. EDP Sciences, 2007.

[BMP05] V. Beck, J. Malick et G. Peyré : *Objectif Agrégation*. H&K, 2ème édition, 2005.

[Bre99] H. Brezis : *Analyse fonctionnelle : théorie et applications*. Dunod, 1999.

[CDGM16] M. Cottrell, C. Duhamel, V. Genon-Catalot et T. Meyre : *Exercices de probabilités*. Cassini, 2016.

[Dem96] J.-P. Demailly : *Analyse numérique et équations différentielles*. Collection Grenoble Sciences, 1996.

[Far00] J. Faraut : *Calcul intégral*. Belin, 2000.

[FGN07] S. Francinou, H. Gianella et S. Nicolas : *Oraux X-ENS - Analyse 1*. Cassini, 2007.

[FGN12] S. Francinou, H. Gianella et S. Nicolas : *Oraux X-ENS - Analyse 4*. Cassini, 2012.

[FGN14] S. Francinou, H. Gianella et S. Nicolas : *Oraux X-ENS - Analyse 3*. Cassini, 2ème édition, 2014.

[Gou08] X. Gourdon : *Les maths en tête - Analyse*. Ellipses, 2ème édition, 2008.

[GT98] A. Gonnord et N. Tosel : *Thèmes d'analyse pour l'agrégation*. Ellipses, 1998.

[HL09] F. Hirsch et G. Lacombe : *Eléments d'analyse fonctionnelle*. Dunod, 2009.

[Ouv09] J.-Y. Ouvrard : *Probabilités : Tome 2*. Cassini, 3ème édition, 2009.

[QZ13] H. Queffélec et C. Zuily : *Analyse pour l'agrégation*. Dunod, 4ème édition, 2013.

[Rou99] F. Rouvière : *Petit guide de calcul différentiel*. Cassini, 1999.

[RS12] V. Rivoirard et G. Stoltz : *Statistique mathématique en action*. Vuibert, 2 ème édition, 2012.

[Tau06] P. Tauvel : *Analyse complexe pour la Licence 3*. Dunod, 2006.