Vrac

Table des matières

1 Pendule approché

Référence : Petit guide de calcul différentiel, François Rouvier

Proposition 1.1

Pour $a \in]0,1[$, on considère le système différentiel, et son système linéarisé :

$$(\mathcal{P}) : \begin{cases} \theta'' &= -\sin(\theta) \\ \theta(0) &= a \\ \theta'(0) &= 0 \end{cases} \quad \text{et } (\tilde{\mathcal{P}}) \begin{cases} \psi'' &= -\psi \\ \psi(0) &= a \\ \psi'(0) &= 0 \end{cases}$$

Alors

$$\forall t \geqslant 0, |\theta(t) - \psi(t)| \leqslant \frac{t|a|^3}{6}$$

Autrement dit, on dispose d'une majoration de l'erreur entre le système approché pour les petits angles, et le système "réel". Application numérique : si on se contente de regarder un pendule pendant 24 heures, et qu'on se contente d'une erreur absolue de 10^{-1} radians, l'approximation est satisfaisante pour $a \simeq 1.1$ degrés. Pour un temps de l'ordre de la minute, on bascule à $a \simeq 12.3$ degrés.

Démonstration : 1. Montrons que le système $\mathcal{P}$ définit une fonction θ bornée. Pour tout $t \geqslant 0$,

$$\theta''(t) = -\sin(\theta(t))$$

On multiplie par $\theta'(t)$ pour obtenir l'énergie du système :

$$\theta'(t)\theta''(t) = -\theta'(t)\sin(\theta(t))$$

On peut alors intégrer cette égalité :

$$\exists \kappa \in \mathbb{R}, \forall t \geqslant 0, \theta'(t)^2 = 2\cos(\theta(t)) + \kappa$$

La condition initiale donne $\kappa = -2\cos(a)$, donc

$$\theta'(t)^2 = 2(\cos(\theta(t) - \cos(a))$$

Ainsi,

$$\forall t \geqslant 0, \cos(\theta(t)) \geqslant \cos(a)$$

Il suit que

$$\forall t \geqslant 0, \theta(t) \in \bigsqcup_{n\in\mathbb{Z}} [-a + 2n\pi, a + 2n\pi]$$

Puisque $\theta(0) = a \in [-a, a]$, il suit par continuité de θ que $\theta(t) \in]-a, a[$ pour tout $t \geqslant 0$, donc

$$\forall t \geqslant 0, |\theta(t)| \leqslant a$$

2. On considère

$$\forall t \geqslant 0, Z(t) \overset{\text{déf.}}{=} \begin{bmatrix} \theta(t) - \psi(t) \\ \theta'(t) - \psi'(t) \end{bmatrix}$$

Alors Z vérifie l'équation différentielle linéaire :

$$Z'(t) = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} Z(t) + \begin{pmatrix} 0 \\ \theta(t) - \sin(\theta(t)) \end{pmatrix}$$

donc, en notant A la matrice antisymétrique, et $B(t)$ le deuxième terme, Z est alors donné par :

$$\forall t \geqslant 0, Z(t) = e^{tA}Z(0) + \int_0^t \exp((t-s)A)B(s)\mathrm{d}s$$

Or, A admet pour exponentielle :

$$e^{tA} = \begin{pmatrix} \cos(t) & \sin(t) \\ -\sin(t) & \cos(t) \end{pmatrix}$$

Ainsi, avec $Z(0) = 0$, la première ligne de $Z(t)$ vaut alors :

$$|\theta(t) - \psi(t)| = \left| \int_0^t \sin(t-s)\left[\theta(s) - \sin(\theta(s))\right] \mathrm{d}s \right|$$

Or,

$$\forall X \in \mathbb{R}, |\sin(X) - X| \leqslant \frac{|X|^3}{6}$$

Cette inégalité s'obtient en intégrant quatre fois $|\cos| \leqslant 1$. Par inégalité triangulaire, et en utilisant que $|\sin| \leqslant 1$, il reste :

$$|\theta(t) - \psi(t)| \leqslant \int_0^t \frac{|\theta(s)|^3}{6}$$

Et d'après **1.**, on aboutit au résultat escompté :

$$\forall t \geqslant 0, |\theta(t) - \psi(t)| \leqslant \frac{t|a|^3}{6}$$

□

2 Théorème de Lyapunov

Référence : Petit guide de calcul différentiel, François Rouvier

Théorème 2.1 : Théorème de Lyapunov

Soit $f : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ une application de classe C^1 telle que $f(0) = 0$. On note $A \overset{\text{déf.}}{=} \partial_X f(0)$. On considère le système differentiel :

$$\begin{cases} y' &= f(y) \\ y(0) &= x \in \mathbb{R}^n \end{cases}$$

On suppose que

$$\forall \lambda \in \sigma(A), \Re\lambda < 0$$

Alors

* 0 est un point d'équilibre attractif du système.
* Il existe un voisinage de 0 pour lequel dans toute condition initiale fait converger $y(t)$ exponentiellement vers 0.

Pour montrer ce théorème, on a d'abord besoin du lemme suivant.

Lemme 2.1

Soit $A \in \mathcal{M}_n(\mathbb{C})$. Il existe un polynôme $P \in \mathbb{R}[X]$ tel que

$$\forall x \in \mathbb{C}^n, \forall t \in \mathbb{R}, \left\|e^{tA}x\right\| \leqslant P(|t|)\left[\sum_{\lambda\in\sigma(A)} e^{t\Re\lambda}\right]\|x\|$$

Démonstration du lemme : Soient $x \in \mathbb{C}^n$ et $t \in \mathbb{R}$.

1. D'après le théorème de Cayley-Hamilton, χ_A annule A, et d'après le lemme des noyaux, il suit donc que

$$\mathbb{C}^n = \bigoplus_{\lambda\in\sigma(A)} \ker\left((A-\lambda I_n)^{m(\lambda)}\right)$$

On décompose x dans cette somme directe :

$$x = \sum_{\lambda\in\sigma(A)} x_\lambda$$

2. Notons

$$\Gamma_\lambda \overset{\text{déf.}}{=} \ker\left((A-\lambda I_n)^{m(\lambda)}\right)$$

Dans Γ_λ, on a :

$$e^{tA}x_\lambda = e^{t\lambda I}e^{t(A-\lambda I)}x_\lambda$$

Développons la deuxième exponentielle en série entière :

$$e^{tA}x_\lambda = e^{t\lambda I}\sum_{\eta=0}^{+\infty}\frac{t^\eta}{\eta!}(A-\lambda I)^\eta x_\lambda$$

Or, par définition de $x_\lambda \in \Gamma_\lambda$:

$$\forall \eta \geqslant m(\lambda), (A-\lambda I)^\eta x_\lambda = 0$$

donc finalement,

$$e^{tA}x_\lambda = e^{t\lambda I}\sum_{\eta=0}^{m(\lambda)-1}\frac{t^\eta}{\eta!}(A-\lambda I)^\eta x_\lambda$$

Ainsi,

$$e^{tA}x = \sum_{\lambda\in\sigma(A)} e^{t\lambda I}\sum_{\eta=0}^{m(\lambda)-1}\frac{t^\eta}{\eta!}(A-\lambda I)^\eta x_\lambda$$

3. Il reste plus qu'à contrôler la norme de cette somme.

$$\left\|e^{tA}x\right\| \leqslant \sum_{\lambda\in\sigma(A)}\left|e^{t\lambda}\right|\sum_{\eta=0}^{m(\lambda)-1}\frac{|t|^\eta}{\eta!}\|(A-\lambda I)^\eta x_\lambda\|$$

En considérant la norme subordonnée associée à $\|\cdot\|$, il existe alors une constante $C_\eta^\lambda > 0$ tel que

$$\|(A-\lambda I)^\eta x_\lambda\| \leqslant C_\eta^\lambda\|x_\lambda\|$$

Ensuite, en notant C^λ le plus grand des C_η^λ, on a :

$$\sum_{\eta=0}^{m(\lambda)-1}\frac{|t|^\eta}{\eta!}\|(A-\lambda I)^\eta x_\lambda\| \leqslant C^\lambda\sum_{\eta=0}^{m(\lambda)-1}\frac{|t|^\eta}{\eta!}\|x_\lambda\|$$

Puisque $\frac{(m(\lambda)-1-\eta)!}{(m(\lambda)-1)!} \leqslant 1$, on se retrouve avec :

$$\sum_{\eta=0}^{m(\lambda)-1}\frac{|t|^\eta}{\eta!}\|(A-\lambda I)^\eta x_\lambda\| \leqslant C^\lambda\sum_{\eta=0}^{m(\lambda)-1}\binom{m(\lambda)-1}{\eta}|t|^\eta\|x_\lambda\|$$

Par la formule du binôme :

$$\sum_{\eta=0}^{m(\lambda)-1} \frac{|t|^\eta}{\eta!} \|(A-\lambda I)^\eta x_\lambda\| \leqslant C^\lambda \left(1+|t|\right)^{m(\lambda)-1} \|x_\lambda\|$$

Puisque $1+|t| \geqslant 1$ et $m(\lambda) \leqslant n$, on a alors

$$\sum_{\eta=0}^{m(\lambda)-1} \frac{|t|^\eta}{\eta!} \|(A-\lambda I)^\eta x_\lambda\| \leqslant C^\lambda \left(1+|t|\right)^{n} \|x_\lambda\|$$

Démonstration : 1. Si z est la solution du système approché

$$\left\{ \begin{array}{rcl} z' & = & Az \\ z(0) & = & x \end{array} \right.$$

alors $z(t) = e^{tA}x$, et par la majoration du lemme, puisque chaque λ de partie réelle strictement négative,

$$z(t) \xrightarrow[t\to+\infty]{} 0$$

2. On introduit l'application $b : \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R}$:

$$\forall x, y \in \mathbb{R}^n, b(x,y) \overset{\text{déf.}}{=} \int_0^{+\infty} \langle e^{tA}x, e^{tA}y\rangle \mathrm{d}t$$

Montrons que b est bien définie. Pour cela, observons que

$$t \longmapsto \langle e^{tA}x, e^{tA}y\rangle$$

est continue sur $\mathbb{R}_+$, et que d'après l'inégalité de Cauchy-Schwarz,

$$\left|\langle e^{tA}x, e^{tA}y\rangle\right| \leqslant \left\|e^{tA}x\right\| \left\|e^{tA}y\right\|$$

D'après le lemme,

$$\left|\langle e^{tA}x, e^{tA}y\rangle\right| \leqslant P(|t|)^2 \left[\sum_{\lambda\in\sigma(A)} e^{t\Re\lambda}\right] \|x\| \, \|y\|$$

Or,

$$t \longmapsto P(|t|)^2 \left[\sum_{\lambda\in\sigma(A)} e^{t\Re\lambda}\right]$$

est intégrable sur $\mathbb{R}_+$, car le spectre de A est de partie réelle incluse dans $\mathbb{R}_-^*$, donc pour chaque $\lambda \in \sigma(A)$,

$$\forall \eta \in \mathbb{N}, t \longmapsto t^\eta e^{t\Re\lambda}$$

est intégrable sur $\mathbb{R}_+$. b est donc bien définie. C'est une application bilinéaire, symétrique, définie et positive. Montrons qu'elle est en effet définie, les autres affirmations étant immédiates. Si $b(x,x) = 0$ alors

Notons alors

$$P(t) \overset{\text{déf.}}{=} \max_{\lambda\in\sigma(A)} C^\lambda (1+t)^n$$

On a alors réussi à montrer que

$$\left\|e^{tA}x\right\| \leqslant P(|t|) \sum_{\lambda\in\sigma(A)} e^{t\Re\lambda} \|x_\lambda\|$$

Par inégalité triangulaire, on a $\|x_\lambda\| \leqslant \|x\|$, et donc on obtient la majoration souhaitée. □

$$\int_0^{+\infty} \left\|e^{tA}x\right\|^2 \mathrm{d}t = 0$$

donc, par continuité, $e^{tA}x = 0$. Ainsi, $x \in \ker(e^{tA})$. Or, e^{tA} est inversible, d'inverse e^{-tA}, donc $x = 0$. b est alors définie positive. On définit alors la *fonction de Lyapunov* :

$$q(x) \overset{\text{déf.}}{=} b(x,x) = \int_0^{+\infty} \left\|e^{tA}x\right\|^2 \mathrm{d}t$$

3. Montrons que

$$\left\langle \vec{\nabla} q(x), Ax \right\rangle = 2b(x, Ax) = -\|x\|^2$$

Pour cela, il suffit de remarquer que pour $k \in [\![1,n]\!]$, en notant e_k le k-ème vecteur de la base canonique de $\mathbb{R}^n$,

$$\partial_k q(x) = 2b(x, e_k)$$

Ainsi,

$$\left\langle \vec{\nabla} q(x), Ax \right\rangle = \sum_{k=1}^n 2b(x,e_k)[Ax]_k$$

donc par linéarité à droite de b,

$$\left\langle \vec{\nabla} q(x), Ax \right\rangle = 2b\left(x, \sum_{k=1}^n [Ax]_k e_k\right)$$

d'où

$$\left\langle \vec{\nabla} q(x), Ax \right\rangle = 2b\left(x, Ax\right)$$

Or, par définition de b,

$$b\left(x, Ax\right) = \int_0^{+\infty} \left\langle e^{tA}x, e^{tA}Ax\right\rangle \mathrm{d}t$$

Or, on constante que

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[\langle e^{tA}x, e^{tA}x\rangle\right] = 2\left\langle e^{tA}x, e^{tA}Ax\right\rangle$$

donc cette intégrale vaut finalement

$$b\left(x, Ax\right) = \frac{1}{2}\left[\left\|e^{tA}\right\|^2\right]_0^{+\infty}$$

ce qui montre bien que

$$\left\langle \vec{\nabla} q(x), Ax \right\rangle = 2b(x, Ax) = -\|x\|^2$$

4. On note

$$r(y) = f(y) - Ay$$

Soit y l'unique solution maximale (sur $I = [0, T^+)$) de ce problème de Cauchy, qui existe car f est de classe C^1, donc est localement lipschitzienne. On constate que

$$\forall t \geqslant [0, T^+), \frac{\mathrm{d}}{\mathrm{d}t}\left[q(y(t))\right] = 2b(y'(t), y(t))$$

Ainsi, puisque $y' = f(y)$, pour tout $t \in [0, T^+)$:

$$\begin{aligned} \tfrac{\mathrm{d}}{\mathrm{d}t}\left[q(y(t))\right] = & \quad 2b(f(y(t)) - Ay(t), y(t)) \\ & + \quad 2b(Ay(t), y(t)) \end{aligned}$$

D'après le point précédent, il suit donc que

$$\forall t \geqslant [0, T^+), \frac{\mathrm{d}}{\mathrm{d}t}\left[q(y(t))\right] = -\|y(t)\|^2 + 2b(y(t), r(y(t)))$$

5. Montrons que s'il existe $\alpha > 0$ tel que $q(y) \leqslant \alpha$, alors il existe $\beta > 0$ tel que

$$-\|y\|^2 + 2b(y, r(y)) \leqslant -\beta q(y)$$

Sur $\mathbb{R}^n$, q et la norme euclidienne sont équivalentes, on note $C_1, C_2 > 0$ tels que

$$\forall x \in \mathbb{R}^n, C_1\|x\|^2 \leqslant q(x) \leqslant C_2\|x\|^2$$

f est différentiable en 0, donc pour tout $\varepsilon > 0$, il existe $\delta > 0$ tel que si $q(x) < \delta$ alors

$$q(r(x)) = q(f(x) - Ax) \leqslant \varepsilon q(x)$$

Soit $\varepsilon > 0$. D'après l'inégalité de Cauchy-Schwarz,

$$-\|y\|^2 + 2b(y, r(y)) \leqslant -\|y\|^2 + 2\sqrt{q(y)q(r(y))}$$

Pour $q(y) \leqslant \delta$, on a alors

$$-\|y\|^2 + 2b(y, r(y)) \leqslant -\|y\|^2 + 2\varepsilon q(y)$$

Par équivalence des normes, on a $\|y\|^2 \geqslant \frac{1}{C_2} q(y)$, donc

$$-\|y\|^2 + 2b(y, r(y)) \leqslant \frac{-q(y)^2}{C_2} + 2\varepsilon q(y)$$

Pour $\varepsilon > 0$ assez petit,

$$-\beta \overset{\text{déf.}}{=} \frac{-1}{C_2} + 2\varepsilon < 0$$

Ainsi, on a bien montré que si $q(y) < \delta$ alors

$$-\|y\|^2 + 2b(y, r(y)) \leqslant -\beta q(y)$$

6. On peut alors enfin conclure. Supposons que $q(x) < \alpha$. Montrons que pour tout $t \geqslant 0$, $q(y(t)) < \alpha$. Par l'absurde, si

$$\tau \overset{\text{déf.}}{=} \inf\left\{t \geqslant 0, q(y(t)) \geqslant \alpha\right\} < +\infty$$

Alors d'après le point précédent,

$$\frac{\mathrm{d}}{\mathrm{d}t}[q(y)](\tau) \leqslant -\beta q(y)(\tau) = -\beta\alpha < 0$$

Alors l'application $t \longmapsto q(y(t))$ est strictement décroissante au voisinage de τ, donc il existe $\delta > 0$ tel que

$$\forall t \in]\tau - \delta, \tau[, q(y(t)) > q(y(\tau)) = \alpha$$

Cela contredit le fait que τ est le plus petit temps pour lequel $q(y(t)) \geqslant \alpha$. Ainsi, pour tout $t \geqslant 0$, $q(y(t)) < \alpha$. On dispose alors de l'inéquation différentielle suivante :

$$\forall t \geqslant 0, \frac{\mathrm{d}}{\mathrm{d}t}[q(y)](t) \leqslant -\beta q(y(t))$$

qui se résout par multiplication par $e^{\beta t}$ pour avoir finalement la décroissance exponentielle voulue :

$$q(y(t)) \leqslant e^{-\beta t} q(x)$$

qui est équivalente à celle avec la norme euclidienne par équivalence des normes. □

3 Théorème fondamental des polynômes symétriques

Référence : Les maths en tête, algèbre, Xavier Gourdon

Définition 3.1

Dans $\mathbb{Z}[X_1, \cdots, X_n]$, on définit le k-ème *polynôme symétrique élémentaire* :

$$\sigma_k(X_1, \cdots, X_n) \overset{\text{déf.}}{=} \sum_{I \in \mathcal{P}_k(\llbracket 1,n \rrbracket)} \left[\prod_{i \in I} X_i \right] \in \mathbb{Z}[X_1, \cdots, X_n]$$

On peut aussi les écrire de la sorte :

$$\sigma_k(X_1, \cdots, X_n) = \sum_{1 \leqslant i_1 < \cdots < i_k \leqslant n} X_{i_1} \cdots X_{i_n} \in \mathbb{Z}[X_1, \cdots, X_n]$$

Proposition 3.1 : Relation coefficients-racines

Pour $\sigma_k^X \in \mathbb{Z}[X_1, \cdots, X_n]$, on a sur $\mathbb{Z}[T]$:

$$\prod_{i=1}^{n} (T - X_i) = \sum_{k=0}^{n} (-1)^k \sigma_k^X T^{n-k}$$

Définition 3.2

Soit A un anneau commutatif unitaire. Un polynôme $P \in A[X_1, \cdots, X_n]$ est dit *symétrique* si :

$$\forall \sigma \in \mathfrak{S}_n, P\left(X_{\sigma(1)}, \cdots, X_{\sigma(n)}\right) = P(X_1, \cdots, X_n)$$

On note $A_{\mathsf{sym}}[X_1, \cdots, X_n]$ l'ensemble des polynômes symétriques. C'est une sous-algèbre de $A[X_1, \cdots, X_n]$.

Théorème 3.1 : Théorème fondamental des polynômes symétriques

Soit A un anneau commutatif unitaire. La sous-algèbre $A_{\mathsf{sym}}[X_1, \cdots, X_n]$ des polynômes symétriques de $A[X_1, \cdots, X_n]$ est engendrée par les polynômes symétriques fondamentaux notés σ_k^X :

$$A_{\mathsf{sym}}[X_1, \cdots, X_n] = A\left[\sigma_1^X, \cdots, \sigma_n^X\right]$$

Référence pour la démonstration : Serge Lang, *Algebra* + Wikipédia

Démonstration : On définit le *poids* d'un polynôme à plusieurs indéterminées : si $X_1, \cdots, X_n$ sont des indéterminées, on définit le *poids* pour un monôme par :

$$w\left(X_1^{\nu_1} \cdots X_n^{\nu_n}\right) \overset{\text{déf.}}{=} \sum_{k=1}^{n} k\nu_k$$

Si $P \in A[X_1, \cdots, X_n]$, son *poids* est le maximum des poids de ses monômes. Constatons alors que si un monôme $X_1^{\nu_1} \cdots X_n^{\nu_n}$ est de poids d alors le monôme

$$\left[\sigma_1^X\right]^{\nu_1} \cdots \left[\sigma_n^X\right]^{\nu_n}$$

est de degré d. En outre, si P est de poids d alors $P(\sigma_1^X, \cdots, \sigma_n^X)$ est de degré d.

On montre le résultat suivant, qui implique le théorème fondamental des polynômes symétriques :

Soit $P \in A_{\mathsf{sym}}[X_1, \cdots, X_n]$ un polynôme symétrique de degré d. Il existe alors un polynôme $g \in A[X_1, \cdots, X_n]$ de poids inférieur ou égal à d tel que

$$P(X_1, \cdots, X_n) = g\left(\sigma_1^X, \ldots, \sigma_n^X\right)$$

On démontre ce résultat par récurrence sur n, le nombre d'indéterminées. Si $n = 1$, alors $X = \sigma_1(X)$, et le théorème est vrai avec $g = P$. Soit $n \geqslant 1$. Supposons le théorème vérifié pour tout polynôme symétrique à $n-1$ indéterminées. Notons σ_k^X les polynômes symétriques élémentaires les $n-1$ indéterminées $X_1, \cdots, X_{n-1}$:

$$\sigma_k^X \overset{\text{déf.}}{=} \sigma_k(X_1, \cdots, X_{n-1})$$

et on note τ_k^X les polynômes symétriques élémentaires pour les n indéterminées $X_1, \cdots, X_n$:

$$\tau_k^X \overset{\text{déf.}}{=} \sigma_k(X_1, \cdots, X_{n-1}, X_n)$$

Alors on démontre par récurrence forte sur d le résultat suivant :

Pour tout polynôme symétrique P à n indéterminées de degré d à coefficients dans A, il existe un polynôme à n indéterminées g de poids inférieur ou égal à d tel que

$$P(X_1, \cdots, X_n) = g(\tau_1^X, \cdots, \tau_n^X)$$

- Pour $d \leqslant n$, les polynômes considérés sont des polynômes à n indéterminées symétriques de degré inférieur à n. Ils s'expriment par conséquent comme produit des polynômes symétriques élémentaires, qui ont un degré exactement égal au poids des polynômes symétriques ; donc l'initialisation est vérifiée.
- Soit $d \geqslant n$. On suppose que la propriété est vraie pour tout polynôme symétrique de degré $d' \leqslant d-1$. Soit $P \in A_{\text{sym}}[X_1, \cdots, X_n]$ de degré d. Alors $P(X_1, \cdots, X_{n-1}, 0)$ est un polynôme à $n-1$ indéterminées, donc par hypothèse de récurrence sur n, il existe $g_1 \in A\left[\sigma_1^X, \cdots, \sigma_{n-1}^X\right]$ de poids inférieur ou égal à d tel que

$$P(X_1, \cdots, X_{n-1}, 0) = g_1\left(\sigma_1^X, \cdots, \sigma_{n-1}^X\right)$$

Par le constat sur le poids vu avant le début de la récurrence, on constate que $g(\tau_1^X, \cdots, \tau_{n-1}^X)$ est de degré plus petit que d. On définit

$$P_1(X_1, \cdots, X_n) \overset{\text{déf.}}{=} P(X_1, \cdots, X_n) - g_1(\tau_1^X, \cdots, \tau_{n-1}^X)$$

Alors P_1 est de degré au plus d, et est symétrique. De plus, remarquons que

$$\sigma_k(X_1, \cdots, X_{n-1}, 0) = \sigma_k(X_1, \cdots, X_{n-1})$$

Donc, on a :

$$P_1(X_1, \cdots, X_{n-1}, 0) = 0$$

donc P_1 admet X_n comme diviseur. P_1 étant symétrique, il admet en fait $X_1 \cdots X_n$ comme diviseur, donc τ_n^X. Ainsi, il existe $P_2 \in A_{\text{sym}}[X_1, \cdots, X_n]$ tel que

$$P_1(X_1, \cdots, X_n) = \tau_n^X P_2(X_1, \cdots, X_n)$$

Or, P_2 est de degré plus petit que $d-n < d$, à n indéterminées, donc par hypothèse de récurrence sur d, il existe $g_2 \in A[X_1, \cdots, X_n]$ de poids inférieur ou égal à d tel que

$$P_2(X_1, \cdots, X_n) = g_2(\tau_1^X, \cdots, \tau_n^X)$$

Ainsi,

$$P(X_1, \cdots, X_n) = P_1(X_1, \cdots, X_n) - g_1(\tau_1^X, \cdots, \tau_{n-1}^X)$$

Soit encore

$$P(X_1, \cdots, X_n) = \tau_n^X g_2(\tau_1^X, \cdots, \tau_n^X) - g(\tau_1^X, \cdots, \tau_{n-1}^X)$$

Donc, en définissant

$$g(X_1, \cdots, X_n) = X_n g_2(X_1, \cdots, X_n) - g_1(X_1, \cdots, X_{n-1})$$

On a montré que

$$P(X_1, \cdots, X_n) = g\left(\tau_1^X, \cdots, \tau_n^X\right)$$

avec $g \in A[X_1, \cdots, X_n]$ de poids inférieur ou égal à d. Ce qui conclut la récurrence. □

Référence : Théorie algébrique des nombres, Pierre SAMUEL

Théorème 3.2 : Théorème de d'ALEMBERT-GAUSS

$\mathbb{C}$ est algébriquement clos : tout polynôme non constant de $\mathbb{C}[X]$ admet une racine dans $\mathbb{C}$.

Démonstration : On suppose acquises les propriétés suivantes :

(a) Tout polynôme réel de degré impair admet une racine dans $\mathbb{R}$.

(b) Tout polynôme complexe de degré 2 admet ses deux racines dans $\mathbb{C}$.

(c) Pour tout corps $\mathbb{K}$ et $P \in \mathbb{K}[X]$, il existe un corps de décomposition pour $\mathbb{P}$ par rapport à $\mathbb{K}$.

(d) Les relations racines-coefficients avec les polynômes symétriques élémentaires.

(e) Le théorème fondamental des polynômes symétriques.

Soit $P \in \mathbb{C}[X]$. Si $a \in \mathbb{C}$ est racine de $P\bar{P} \in \mathbb{R}[X]$, alors $P(a) = 0$ ou $\bar{P}(a) = 0$, auquel cas $P(\bar{a}) = 0$ donc P admet une racine. Ainsi, on peut se contenter de montrer le résultat sur les polynômes de $\mathbb{R}[X]$. Pour $d \in \mathbb{N}$, on note

$d = 2^n q$, avec q impair. On fixe donc $q \in \mathbb{N}$ impair. On se propose de montrer le théorème par récurrence sur n, la valuation 2-adique de d.

- Pour $n = 0$, d est impair, et d'après **(a)**, tout polynôme réel de degré impair admet une racine, réelle qui plus est. L'initialisation est donc vraie.
- Soit $n \in \mathbb{N}$. On suppose que tout polynôme réel de degré $2^{n'} q$ où $n' \leqslant n$ admet une racine dans $\mathbb{C}$. Soit $P \in \mathbb{R}[X]$ de degré $d = 2^n q$, que l'on suppose unitaire. D'après **(c)**, il existe une extension $\mathbb{K}$ de $\mathbb{C}$ dans laquelle il existe $x_1, \cdots, x_n \in \mathbb{K}$ tels que

$$P = \prod_{k=1}^{n} (X - x_k)$$

Montrons que pour tout $k \in [\![1, n]\!]$, $x_k \in \mathbb{C}$. Pour tout $c \in \mathbb{R}$, on considère

$$\forall i \leqslant j \in [\![1, n]\!], y_{i,j}^{[c]} \overset{\text{déf.}}{=} x_i + x_j + c x_i x_j$$

Il y a $\frac{d(d-1)}{2}$ éléments $y_{i,j}^{[c]}$, comptés avec multiplicité (*id est* on peut compter plusieurs fois compter le même élément dans ce décompte). Considérons :

$$G^{[c]} \overset{\text{déf.}}{=} \prod_{1 \leqslant i \leqslant j \leqslant n} (X - y_{i,j}^{[c]})$$

C'est un polynôme de $\mathbb{K}[X]$ de degré $\frac{d(d-1)}{2} = 2^{n-1} q$. Par les relations coefficients-racines **(d)**, les coefficients de $G^{[c]}$ sont fonctions polynomiales symétriques des $y_{i,j}^{[c]}$, donc des x_k. Or, d'après le théorème fondamental des polynômes symétriques **(e)**, tout polynôme réel symétrique en les x_i est polynôme réel en les fonctions symétriques élémentaires de x_i. Or, les x_i sont racines de $P \in \mathbb{R}[X]$, donc par les relations coefficients racines, les polynômes symétriques élémentaires de x_i sont des polynômes réels, donc $G^{[c]} \in \mathbb{R}[X]$. $G^{[c]}$ est alors un polynôme réel de degré $2^{n-1} q$. Par hypothèse de récurrence, l'une de ses racines est dans $\mathbb{C}$. Donc, on a montré que pour tout $c \in \mathbb{R}$, il existe deux indices $i(c)$ et $j(c)$ tels que

$$y_{i(c),j(c)}^{[c]} = x_{i(c)} + x_{j(c)} + c x_{i(c)} x_{j(c)} \in \mathbb{C}$$

Or, $c \longmapsto (i(c), j(c))$ envoie un réel sur $[\![1, n]\!]^2$, donc cette application ne peut pas être injective, puisque $\mathbb{R}$ est infini. Il existe donc deux éléments c et $c' \in \mathbb{R}$ tels que $i(c) = i(c') \overset{\text{déf.}}{=} r$ et $j(c) = j(c') \overset{\text{déf.}}{=} s$. Ainsi,

$$y_{r,s}^{[c]} \in \mathbb{C} \text{ et } y_{r,s}^{[c']} \in \mathbb{C}$$

Ce qui donne les égalités en les $x_r + x_s$ et $x_r x_s$ suivantes :

$$\begin{cases} x_r + x_s + c x_r x_s & \in & \mathbb{C} \\ x_r + x_s + c' x_r x_s & \in & \mathbb{C} \end{cases}$$

Par combinaisons linéaires, il suit que $x_r + x_s \in \mathbb{C}$ et $x_r x_s \in \mathbb{C}$. Or, x_r et x_s sont racines de

$$z^2 - (x_r + x_s) z + x_r x_s = 0$$

C'est un polynôme de degré 2 à coefficients complexes, donc admet ses deux racines dans $\mathbb{C}$, *via* **(b)**. Il suit que x_r et $x_s \in \mathbb{C}$. On a alors réussi à montré que P admet une racine complexe. Cela achève la récurrence. □

4 Théorème de Frobenius - Zolotarev

Référence : Objectif Agrégation, Vincent Beck, Jérôme Malick, Gabriel Peyré

Définition 4.1

Soit $p \in \mathbb{N}$ un nombre premier. On définit le *symbole de Legendre* comme étant l'application de $\mathbb{Z}$ dans $\{-1, 0, 1\}$ telle que

$$\forall a \in \mathbb{Z}, \left(\frac{a}{p}\right) \overset{\text{déf.}}{=} \begin{cases} 0 & \text{si} & a \in p\mathbb{Z} \\ 1 & \text{si} & a \text{ est un carré modulo } p \\ -1 & \text{sinon} & \end{cases}$$

Théorème 4.1 : Théorème de Frobenius-Zolotarev

Soient $p \geqslant 3$ un nombre premier, $\mathbb{F}_p$ le corps à p éléments et V un $\mathbb{F}_p$-espace vectoriel de dimension finie. Alors pour tout endomorphisme inversible $u \in \mathrm{GL}(V) \subset \mathfrak{S}(V)$, on a la relation :

$$\varepsilon(u) = \left(\frac{\det(u)}{p}\right)$$

Référence : Daniel Perrin, *Cours d'algèbre*

Lemme 4.1

Si $\mathbb{K}$ est un corps et $n \in \mathbb{N}^*$ alors :

- Deux transvections sont conjuguées dans $\mathsf{GL}_n(\mathbb{K})$.
- $\mathsf{SL}_n(\mathbb{K})$ est engendrée par les transvections.

Lemme 4.2

Si $\mathbb{K}$ est un corps différent de $\mathbb{F}_2$ et si $n \geqslant 3$ alors

$$D\left(\mathsf{GL}_n(\mathbb{K})\right) = \mathsf{SL}_n(\mathbb{K})$$

où $D(G)$ désigne le groupe dérivé de G, c'est-à-dire le groupe engendré par les commutateurs de G.

Démonstration du lemme : • Si $A, B \in \mathsf{GL}_n(\mathbb{K})$ alors

$$\det([A,B]) = \det(ABA^{-1}B^{-1}) = 1$$

donc $D(\mathsf{GL}_n(\mathbb{K})) \subset \mathsf{SL}_n(\mathbb{K})$.

• Soit $U \in \mathsf{SL}_n(\mathbb{K})$ une transvection, *id est* une matrice du type $I_n + \lambda E_{i,j}$ avec $i \neq j$. Alors U^2 est aussi une transvection, puisque $\mathbb{K}$ est de caractéristique différente de 2 :

$$U^2 = I_n + 2\lambda E_{i,j}$$

Or, deux transvections sont conjuguées dans $\mathsf{GL}_n(\mathbb{K})$: il existe $S \in \mathsf{GL}_n(\mathbb{K})$ telle que

$$U^2 = SUS^{-1}$$

d'où $U = SUS^{-1}U^{-1} = [S,U]$, donc $U \in D(\mathsf{GL}_n(\mathbb{R})$.

• Les transvections engendrent $\mathsf{SL}_n(\mathbb{K})$. On a montré que toutes les transvections étaient étaient incluses dans $D(\mathsf{GL}_n(\mathbb{K}))$. Par conséquent, $\mathsf{SL}_n(\mathbb{K}) \subset D(\mathsf{GL}_n(\mathbb{K}))$. D'où $D\left(\mathsf{GL}_n(\mathbb{K})\right) = \mathsf{SL}_n(\mathbb{K})$. □

Lemme 4.3

Soit $\mathbb{K}$ un corps différent de $\mathbb{F}_2$. Soit M un groupe abélien. Pour $n \geqslant 3$, on considère

$$\varphi : \mathsf{GL}_n(\mathbb{K}) \longrightarrow M$$

un morphisme de groupes. Il existe un unique morphisme de groupe

$$\delta : \mathbb{K}^\times \longrightarrow M$$

tel que $\varphi = \delta \circ \det$:

$$\begin{tikzcd}
\mathsf{GL}_n(\mathbb{K}) \arrow[rr, "\varphi"] \arrow[dr, "\det"'] & & M \\
& \mathbb{K}^\times \arrow[ur, "\delta"'] &
\end{tikzcd}$$

Démonstration du lemme : D'après le lemme **4.2**, $D(\mathrm{GL}_n(\mathbb{K})) = \mathrm{SL}_n(\mathbb{K})$. Par propriété universelle du groupe dérivée, ${}^{\mathrm{GL}_n(\mathbb{K})}/_{\mathrm{SL}_n(\mathbb{K})}$ est un groupe abélien. De plus, $\mathrm{SL}_n(\mathbb{K}) \subset \ker(\varphi)$. En effet, Si $A, B \in \mathrm{GL}_n(\mathbb{K})$,

$$\varphi([A,B]) = [\varphi(A), \varphi(B)] = 1_M$$

car M est commutatif. Puisque $\mathrm{SL}_n(\mathbb{K})$ est engendré par les commutateurs de $\mathrm{GL}_n(\mathbb{K})$, $\mathrm{SL}_n(\mathbb{K}) \subset \ker(\varphi)$. Par propriété universelle du groupe quotient, il existe alors un unique morphisme $\tilde{\varphi} : {}^{\mathrm{GL}_n(\mathbb{K})}/_{\mathrm{SL}_n(\mathbb{K})} \longrightarrow M$ tel qu $\varphi = \tilde{\varphi} \circ \pi$ où π désigne la surjection canonique :

$$\begin{tikzcd} \mathrm{GL}_n(\mathbb{K}) \arrow[rr, "\varphi"] \arrow[dr, "\pi"'] & & M \\ & {}^{\mathrm{GL}_n(\mathbb{K})}/_{\mathrm{SL}_n(\mathbb{K})} \arrow[ur, "\tilde{\varphi}"'] & \end{tikzcd}$$

Or, d'après le premier théorème d'isomorphisme appliqué à $\det$, il existe un unique isomorphisme $\overline{\det}$: ${}^{\mathrm{GL}_n(\mathbb{K})}/_{\mathrm{SL}_n(\mathbb{K})} \longrightarrow \mathbb{K}^\times$ tel que $\det = \pi \circ \overline{\det}$, qui fait commuter le diagramme commutatif suivant :

$$\begin{tikzcd} \mathrm{GL}_n(\mathbb{K}) \arrow[rr, "\det"] \arrow[dr, "\pi"'] & & \mathbb{K}^\times \\ & {}^{\mathrm{GL}_n(\mathbb{K})}/_{\mathrm{SL}_n(\mathbb{K})} \arrow[ur, "\overline{\det}"'] & \end{tikzcd}$$

Ainsi, on se retrouve dans la situation suivante :

$$\begin{tikzcd} \mathrm{GL}_n(\mathbb{K}) \arrow[r, "\varphi"] \arrow[d, "\pi"'] \arrow[dr, bend left=60, "\det"] & M \\ {}^{\mathrm{GL}_n(\mathbb{K})}/_{\mathrm{SL}_n(\mathbb{K})} \arrow[r, "\overline{\det}"] \arrow[ur, "\tilde{\varphi}"] & \mathbb{K}^\times \arrow[u, dashed, "\delta"'] \end{tikzcd}$$

On définit donc $\delta \overset{\text{déf.}}{=} \tilde{\varphi} \circ \overline{\det}^{-1}$, qui est unique par unicité de $\overline{\det}$ et $\tilde{\varphi}$. □

Lemme 4.4

Soit $p \geqslant 3$ premier, et $\mathbb{F}_p$ le corps à p éléments. On note δ le symbole de LEGENDRE par rapport à p. δ est l'unique morphisme non trivial de $\mathbb{F}_p^\times \longrightarrow \{-1, 1\}$.

Démonstration du lemme : • Le symbole de Legendre est non trivial, car il existe des éléments qui ne sont pas des carrés dans $\mathbb{F}_p^\times$.

• Soit $\alpha : \mathbb{F}_p^\times \longrightarrow \{-1, 1\}$ un morphisme non trivial. Son noyau définit un sous-groupe d'indice 2. De plus, $\mathbb{F}_p^\times$ est un groupe cyclique, de cardinal pair. Ainsi, d'après un théorème, $\ker\alpha$ est l'unique sous-groupe d'indice 2 de $\mathbb{F}_p^\times$. Il vérifie en particulier

$$\mathbb{F}_p^\times = \ker\alpha \sqcup x\ker\alpha$$

où $\alpha(x) = -1$. Ainsi, $\alpha(a)$ est entière déterminé selon l'appartenance de a à $\ker\alpha$ ou non. α est alors unique, et δ réalise exactement cette propriété. □

Démonstration : On considère

$$\varepsilon : \mathfrak{S}(V) \longrightarrow \{-1, 1\}$$

la signature que l'on va restreindre à $\mathrm{GL}(V)$, et on notera ε cette restriction. ε est alors un morphisme de groupe de $\mathrm{GL}_n(\mathbb{F}_p) \simeq \mathrm{GL}(V)$, où n est la dimension de V, à valeurs dans un groupe commutatif $\{-1, 1\}$. Par le lemme **4.3**, il existe $\delta : \mathbb{F}_p^\times \longrightarrow \{-1, 1\}$ tel que

$$\varepsilon = \delta \circ \det$$

Montrons que δ n'est pas le morphisme trivial. Grâce à cela, le lemme **4.4** nous permettra de conclure que δ est le symbole de Legendre.

Pour cela, on constate que $\sharp V = p^n$. Il existe une extension de $\mathbb{F}_p$ de cardinal $p^n \overset{\text{déf.}}{=} q$, on le note $\mathbb{F}_q$. Alors V et $\mathbb{F}_q$ sont isomorphes en tant qu'espaces vectoriels. On considère g un générateur du groupe cyclique $\mathbb{F}_p^\times$, et l'application

$$u : \left(\begin{array}{ccc} \mathbb{F}_q & \longrightarrow & \mathbb{F}_q \\ x & \longmapsto & gx \end{array} \right)$$

Cette application laisse 0 fixe, et agit comme la permutation $(1 \ \cdots \ p-1)$ sur les autres éléments. Ainsi,

$$\varepsilon(u) = (-1)^q = -1$$

en étendant la définition de ε sur $\mathbb{F}_q \simeq V$, et en remarquant que $q = p^n$ est impair. Ainsi,

$$\delta \circ \det(u) = -1$$

donc δ n'est pas le morphisme trivial. On a alors réussi à montrer que δ était bien le symbole de Legendre. □

5 Générateurs du groupe linéaire et du groupe spécial linéaire

Référence : Daniel PERRIN, *Cours d'algèbre*

Théorème 5.1

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

* $\mathsf{GL}(E)$, l'ensemble des endomorphismes inversibles de E est engendré par les transvections et les dilatations.
* $\mathsf{SL}(E)$, l'ensemble des endomorphisme de déterminant 1, est engendré par les transvections.

Proposition 5.1 : Dilatations

Soient $u \in \mathsf{GL}(E)$ et H un hyperplan de E tel que $u_{|_H} = \mathrm{id}_H$ (donc $H \subset \ker(u - \mathrm{id})$). Les propositions suivantes sont équivalentes :

(i) $\det u \overset{\text{déf.}}{=} \lambda \neq 1$.

(ii) u admet une valeur propre $\lambda \neq 1$, et u est diagonalisable.

(iii) L'image de $u - \mathrm{id}$ n'est pas incluse dans H : il existe $x \in E$ tel que $u(x) - x \notin H$.

(iv) Il existe $\lambda \in \mathbb{K}^*$ différent de 1, et une base $\mathcal{B}$ de E dans laquelle

$$\mathsf{Mat}_{\mathcal{B}}(u) = \begin{pmatrix} 1 & 0 & \cdots & \cdots & 0 \\ 0 & 1 & \ddots & & \vdots \\ \vdots & & \ddots & & \\ 0 & \cdots & & 1 & \\ 0 & & & & \lambda \end{pmatrix}$$

Dans le cas où l'une des ces quatre propositions est vérifiée, u est une *dilatation* de rapport λ d'hyperplan H (ou de direction D, lorsque D représente la direction propre pour la valeur propre $\lambda \neq 1$). On a alors $H = \ker(u - \mathrm{id})$ et $D = \mathsf{Im}(u - \mathrm{id})$.

Démonstration : [(i) $\Longrightarrow$ (ii)] Supposons que $\det(u) \overset{\text{déf.}}{=} \lambda \neq 1$. Alors $\ker(u - \mathrm{id})$ est de dimension au plus $n-1$. Or, $H \subset \ker(u - \mathrm{id})$, et H est de dimension $n-1$, donc u admet 1 comme valeur propre de multiplicité $n-1$, de sous-espace propre de dimension $n-1$. Puisque le produit des valeur propres vaut $\det(u)$, il suit que $\lambda \neq 1$ est la n-ème valeur propre, qui est simple pour u. Ainsi, u est alors diagonalisable.

[(ii) $\Longrightarrow$ (iii)] On suppose que $\lambda \neq 1$ est valeur propre pour u diagonalisable. Soit $x \in E$ vecteur propre pour la valeur propre λ. Alors

$$u(x) - x = (\lambda - 1)x \neq 0$$

Donc

$$u(x) - x \in \mathsf{Im}(u - \mathrm{id}) \cap (\complement_E H)$$

[(iii) $\Longrightarrow$ (iv)] Soit $x \in E$ tel que $u(x) - x \notin H$. Alors $u(x) \neq x$, donc $\ker(u - \mathrm{id}) \neq E$, donc par égalité de dimensions,

$$H = \ker(u - \mathrm{id})$$

D'après le théorème du rang,

$$\mathsf{rg}(u - \mathrm{id}) = 1$$

On considère $(e_1, \cdots, e_{n-1})$ une base de H, et $e_n \in \mathsf{Im}(u - \mathrm{id})$, n'appartenant pas à H. Alors $(e_1, \cdots, e_n)$ est une base de E pour la quelle $u(e_i) = e_i$ lorsque $i \neq n$. De plus, $u(e_n) - e_n \neq 0$, et e_n engendre $\mathsf{Im}(u - \mathrm{id})$, donc il existe $\mu \in \mathbb{K}$ non nul tel que

$$u(e_n) - e_n = \mu e_n$$

donc

$$u(e_n) = (1+\mu)e_n = \lambda e_n$$

où $\lambda \neq 1$.

[(iv) $\implies$ (i)] Le déterminant de cette matrice est $\lambda \neq 1$. □

Proposition 5.2 : Transvections

Soient $H = \ker \varphi$ un hyperplan de E, avec $\varphi \in E^*$ non nulle, et $u \in \mathsf{GL}(E)$ tel que $u_{|_H} = \mathrm{id}_H$ (*id est* $H \subset \ker(u - \mathrm{id})$), et différente de id_E. Les propositions suivantes sont équivalentes :

(i) $u \in \mathsf{SL}(E)$, *id est* $\det u = 1$.

(ii) u n'est pas diagonalisable.

(iii) L'image de $u - \mathrm{id}$ est contenue dans H.

(iv) L'endomorphisme induit par u sur ${}^{E}/_{H}$ est l'identité.

(v) Il existe $a \in H$ non nul tel que

$$\forall x \in E, u(x) = x + \varphi(x)a$$

(vi) Il existe une base $\mathcal{B}$ de E dans laquelle

$$\mathsf{Mat}_{\mathcal{B}}(u) = \begin{pmatrix} 1 & 0 & \cdots & \cdots & 0 \\ 0 & 1 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 1 \\ 0 & & \cdots & 0 & 1 \end{pmatrix}$$

Dans le cas où l'une de ces six propositions est vérifiée, u est une *transvection* d'hyperplan H, et de droite $D = \mathsf{Im}(u - \mathrm{id})$. De plus, $D = \mathsf{Vect}_{\mathbb{K}}(a) \subset H$.

Démonstration : Notons au préalable que l'hypothèse $u \neq \mathrm{id}$ implique par égalité de dimensions que

$$H = \ker(u - \mathrm{id})$$

[(i) $\implies$ (ii)] Supposons que $\det u = 1$. Dans ce cas, puisque u contient au moins $n-1$ fois la valeur propre 1, le fait que $\det u = 1$ donne que sa seule valeur propre est 1, de multiplicité algébrique n, mais de multiplicité géométrique $n-1$: supposons par l'absurde que $\dim(\ker(u - \mathrm{id})) = n$, alors $u = \mathrm{id}_E$, ce qui contredit l'hypothèse $u \neq \mathrm{id}_E$. u n'est donc pas diagonalisable.

[(ii) $\implies$ (iii)] Supposons que u ne soit pas diagonalisable. Soit $y = u(x) - x \in \mathsf{Im}(u - \mathrm{id})$. ??

[(iii) $\iff$ (iv)] On suppose que l'image de $u - \mathrm{id}$ est contenue dans H. On note $\bar{u} : {}^{E}/_{H} \longrightarrow {}^{E}/_{H}$ l'endomorphisme induit par u. Par hypothèse,

$$\forall x \in E, u(x) - x \in H$$

si et seulement si

$$\bar{u}(\bar{x}) - \bar{x} = 0$$

si et seulement si

$$\forall \bar{x} \in {}^{E}/_{H}, \bar{u}(\bar{x}) = \bar{x}$$

donc on a bien montré l'équivalence.

[(iv) $\implies$ (v)] Pour pour tout $x \in E$, $\bar{u}(\bar{x}) - \bar{x} = 0$ dans l'espace quotient, donc $u(x) - x \in H = \ker(\varphi)$. Puisque φ n'est pas nulle, c'est une forme surjective. Ainsi, il existe $x_0 \in E$ tel que $\varphi(x_0) = 1$. Alors, on note

$$a \overset{\text{déf.}}{=} u(x_0) - x_0 \in H$$

et $a \neq 0$ sans quoi $x_0 \in \ker(u - \mathrm{id}) = H$, et donc $\varphi(x_0) = 0$, ce qui contredit $\varphi(x_0) = 1$. On dispose de plus de :

$$E = H \oplus \mathbb{K}x_0$$

Soit $x = x_H + \varphi(x)x_0 \in E$. Alors

$$u(x) = u(x_H) + \varphi(x)u(x_0)$$

Puisque $x_H \in H$ et $u(x_0) = a + x_0$ par définition de a, il suit que

$$u(x) = x_H + \varphi(x)a + \varphi(x)x_0$$

donc

$$\forall x \in E, u(x) = x + \varphi(x)a$$

[(v) $\Longrightarrow$ (vi)] On considère $(e_1, \cdots, e_{n-2}, a)$ base de $H = \ker(u - \mathrm{id})$ (bien défini parce que $a \neq 0$, et théorème de la base incomplète). φ étant surjective, il existe $x_0 \in E$ tel que $\varphi(x_0) = 1$. Ainsi, d'après **(v)**, on a $a = u(x_0) - x_0$, et $x_0 \notin H$. Alors $(e_1, \cdots, e_{n-1}, x_0)$ est une base de E dans laquelle $u(e_i) = e_i$, $u(a) = a$ et

$$u(x_0) = x_0 + a$$

Dans la base $\mathcal{B} \overset{\text{déf.}}{=} (e_1, \cdots, e_{n-2}, a, x_0)$, u a donc bien la forme matricielle attendue.

[(vi) $\Longrightarrow$ (i)] Le déterminant de cette matrice triangulaire est bien 1. □

On note la suite $\tau(\varphi, a)$ la transvection définie par :

$$\forall x \in E, \tau(\varphi, a)(x) \overset{\text{déf.}}{=} x + \varphi(x)a$$

Référence additionnelle : Matthieu Romagny, *$SL(E)$ est engendré par les transvections*

Lemme 5.1

Soit $x \in E \backslash \{0\}$ et soient H_1, H_2 deux hyperplans distincts tels que $x \notin H_1 \cup H_2$. Alors il existe une transvection τ telle que $\tau(x) = x$ et $\tau(H_1) = H_2$.

Démonstration du lemme : On note

$$H \overset{\text{déf.}}{=} H_1 \cap H_2 + \mathbb{K}x$$

Alors $H \neq H_k$, car $x \in H \backslash H_k$. Par dimension, l'un n'est pas inclus dans l'autre, donc en particulier, puisque $x \notin H_1 \cup H_2$, on a (en complétant une base de H_k par exemple) :

$$E = H + H_1 = H + H_2$$

donc pour $z \in H_2 \backslash H$ (qui est non vide car $x \in H \backslash H_2$ et H_2 n'est pas inclus dans H), on note

$$z \overset{\text{déf.}}{=} y + a \in H_1$$

avec $y \in H_1$ et $a \in H$. Puisque $y \notin H$, sans quoi $z \in H$, il existe $\varphi \in E^*$ telle que $H = \ker(\varphi)$ et $\varphi(y) = 1$. Il ne reste plus qu'à définir

$$\forall t \in E, \tau(t) \overset{\text{déf.}}{=} t + \varphi(t)a$$

Alors $\tau(y) = z$, donc τ envoie H_1 sur H_2 et $\tau(x) = x$. □

Lemme 5.2

Supposons que $\dim E \geqslant 2$. Soient $x, y \in E$. Alors :

- $*$ Soit il existe τ une transvection telle que $\tau(x) = y$.
- $*$ Soit il existe deux transvections τ_1 et τ_2 telles que $y = \tau_1\tau_2(x)$.

Démonstration du lemme : • Si x et y ne sont pas colinéaires, la famille $(x, y - x)$ est libre donc en notant $a \overset{\text{déf}}{=} y - x$, il existe une forme linéaire $\varphi \in E^*$ telle que $\varphi(a) = 0$ et $\varphi(x) = 1$. On définit alors

$$\forall t \in E, \tau(t) \overset{\text{déf.}}{=} t + \varphi(t)a$$

et on constate bien que $\tau(x) = y$.

• Si x et y sont colinéaires, alors puisque $\dim E \geqslant 2$, il existe z tel que (x, z) et (y, z) ne le soient pas. Par le point précédent, il existe τ_1 une transvection telle que $\tau_1(z) = y$ et τ_2 une transvection telle que $\tau_2(x) = z$, si bien que

$$y = \tau_1 \circ \tau_2(x)$$

□

Démonstration : On démontre que $\mathrm{SL}(E)$ est engendré par les transvections par récurrence sur n.

- Pour $n=1$, $\mathrm{SL}(E)$ est réduit à l'identité, et l'ensemble des transvections est vide, donc le résultat est vrai pour $n=1$.

- Supposons que pour tout espace E de dimension n, $\mathrm{SL}(E)$ est engendré par les transvections. Soit E un $\mathbb{K}$-espace vectoriel de dimension $n+1$. Soit $v \in \mathrm{SL}(E)$. Soit $x \in E$ non nul. D'après le lemme **5.2**, quitte composer v par une ou deux transvections (qui dépendent de x), supposons que $v(x)=x$. Soit H un hyperplan tel que $x \notin H$. Ainsi, $x=v(x) \notin H$. Donc $x \notin H \cup v(H)$, qui sont deux hyperplans puisque $v \in \mathrm{SL}(E)$. D'après le lemme **5.1**, quitte à composer par une transvection (qui dépend de x), on peut supposer que $H=v(H)$. Alors $v_{|_H} \in \mathrm{SL}(H)$ avec $\dim H = n$. Par hypothèse de récurrence, il existe un produit de transvections $\tau_{i,H}$ tels que $v_{|_H} = \prod_{i \in I} \tau_{i,H}$. Alors chacune de ces transvections s'étendent en une unique transvection sur E qui fixe x : τ_i. Puisque $v(x)=x$, v est donc produit des τ_i. Ainsi, on a bien montré que $\mathrm{SL}(E)$ est engendré par les transvections.

- Soit $u \in \mathrm{GL}(E)$ de déterminant $\lambda \neq 1$. Soit v_λ une dilatation de rapport $\frac{1}{\lambda}$. Alors $v_\lambda \circ u \in \mathrm{SL}(E)$ est engendré par les transvections. Ainsi, u est engendré par les transvections et v_λ. On a finalement prouvé que $\mathrm{GL}(E)$ est bien engendré par les transvections. □

6 Nombre de Liouville

Référence : Petit guide du calcul différentiel, François Rouvière.

Proposition 6.1 : Existence d'un nombre transcendant

On pose

$$x \overset{\text{déf.}}{=} \sum_{n=0}^{+\infty} \frac{1}{2^{n!}}$$

Alors x est transcendant : il n'existe pas de polynôme $P \in \mathbb{Z}[X]$ tel que $P(x)=0$.

Lemme 6.1

Soit x un nombre irrationnel algébrique de degré au plus n. Alors

$$\exists C>0, \forall p \in \mathbb{Z}, \forall q \in \mathbb{N}^*, \left| x - \frac{p}{q} \right| \geqslant \frac{C}{q^n}$$

Démonstration du lemme : On va chercher à appliquer l'inégalité dite *de la moyenne*, ou l'inégalité des accroissements finis. Soit f le polynôme de $\mathbb{Z}[X]$ qui annule x. Il existe $a>0$ tel que sur $[x-a,x+a]$, $f(x) \neq 0$. Si $\frac{p}{q} \in [x-a,x+a]$, on a d'après l'inégalité des accroissements finis :

$$\left| f\left(\frac{p}{q}\right) \right| \leqslant \sup_{[x-a,x+a]} |f'| \left| x - \frac{p}{q} \right|$$

De plus, $f\left(\frac{p}{q}\right)$ est non nul, et se présente sous la forme $\frac{A}{q^n}$, avec $A \neq 0$ entier. Ainsi, $|A| \geqslant 1$. Il reste donc

$$\left| x - \frac{p}{q} \right| \geqslant \frac{|A|}{q^n \sup_{[x-a,x+a]} |f'|} \geqslant \frac{C_1}{q^n}$$

où $C_1 \overset{\text{déf.}}{=} \frac{1}{\sup_{[x-a,x+a]} |f'|}$. Cette inégalité est vraie pour $|x-\frac{p}{q}| \leqslant a$. Dans l'autre cas, on a $|x-\frac{p}{q}| \geqslant a$, donc $|x-\frac{p}{q}| \geqslant \frac{a}{q^n}$. En conclusion,

$$\left| x - \frac{p}{q} \right| \geqslant \frac{C}{q^n}$$

où $C \overset{\text{déf.}}{=} \min\{a, C_1\}$. □

Démonstration : On note

$$x_N \overset{\text{déf.}}{=} \sum_{n=0}^{N} \frac{1}{2^{n!}}$$

- Montrons que

$$0 < x - x_N \leqslant \left(\frac{1}{2^{k!}} \right)^k$$

Pour cela, on constate que

$$x - x_N = \sum_{n=N+1}^{+\infty} \frac{1}{2^{n!}}$$

Soit, par un petit changement de variables,

$$x - x_N = \sum_{n=1}^{+\infty} \frac{1}{2^{(n+N)!}}$$

Or, on dispose d'une majoration (grossière) :

$$(k+l)! \geqslant k!(k+l) \geqslant k! \cdot k + l$$

On l'applique ici :

$$x - x_N \leqslant \frac{1}{2^{N! \cdot N}} \sum_{n=1}^{+\infty} \frac{1}{2^n}$$

donc

$$x - x_N \leqslant \left(\frac{1}{2^{N!}}\right)^N$$

- Montrons que x est irrationnel. On a :

$$x_N \overset{\text{déf.}}{=} \frac{p_N}{q_N}$$

avec $q_N = 2^{N!}$ donc

$$0 < x - \frac{p_N}{q_N} \leqslant \left(\frac{1}{q_N}\right)^N$$

Si $x = \frac{p}{q}$ alors

$$0 < pq_N - p_N q \leqslant \frac{q}{q_N^{N-1}}$$

Pour N assez grand, le membre de gauche est strictement inférieur à 1, ce qui contredit $pq_N - p_N q \in \mathbb{Z}$. Donc x est irrationnel.

- Montrons que x n'est pas algébrique. Si tel était le cas, le lemme **6.1** nous donnerait, avec n le degré de x :

$$\frac{C}{q_N^n} \leqslant x - \frac{p_N}{q_N} \leqslant \left(\frac{1}{q_N}\right)^N$$

Ce qui contredirait la croissance de la fonction $\alpha \longmapsto x^\alpha$ pour N assez grand. x est donc un nombre transcendant. $\square$

7 Formule de Poincaré

Référence : Probabilité Tome 1, Jean-Yves Ouvrard

Proposition 7.1 : Formule de Poincaré

Soit $n \geqslant 2$. On considère un espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$ et $(A_i)_{i \in I}$ une suite finie de cardinal n d'éléments de $\mathcal{F}$. Alors

$$\mathbb{P}\left(\bigcup_{i \in I} A_i\right) = \sum_{k=1}^{n} \left[(-1)^{k-1} \sum_{J \in \mathcal{P}_k(I)} \mathbb{P}\left(\bigcap_{j \in J} A_j\right)\right]$$

où $\mathcal{P}_k(I)$ désigne les parties de I de cardinal k. Pour $n = 2$, cela est simplement

$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$$

Démonstration : C'est une démonstration sur récurrence sur n, le cardinal de I.

- Pour $n = 2$, il s'agit de constater que si $A, B \in \mathcal{F}$, alors

$$A \cup B = A \sqcup [B \backslash (A \cap B)]$$

Si bien que

$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B \backslash (A \cap B)) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$$

- Soit $n \geqslant 2$. Supposons que pour tout ensemble I de cardinal n, et pour toute suite finie $(A_i)_{i \in I} \in \mathcal{F}^I$, on dispose de la formule de Poincaré.

On considère donc I de cardinal $n+1$, $I = [\![1, n+1]\!]$ par souci de simplicité d'écriture, et $(A_i)_{i \in I}$ une suite de $\mathcal{F}^I$, composé donc de $n+1$ éléments. Alors, d'après le cas $n = 2$:

$$\mathbb{P}\left(\bigcup_{i \in I} A_i\right) = \mathbb{P}\left(\left[\bigcup_{i \in I \backslash \{n+1\}} A_i\right] \cup A_{n+1}\right)$$

permet d'avoir, avec I' l'ensemble $I \backslash \{n+1\}$ de cardinal

n :

$$\mathbb{P}\left(\bigcup_{i\in I} A_i\right) \overset{(*)}{=} \mathbb{P}\left(\bigcup_{i\in I'} A_i\right) + \mathbb{P}(A_{n+1}) - \mathbb{P}\left(\bigcup_{i\in I'} A_i \cap A_{n+1}\right)$$

Constatons que

$$\bigcup_{i\in I'} A_i \cap A_{n+1} = \bigcup_{i\in I'} (A_i \cap A_{n+1})$$

afin d'appliquer deux fois l'hypothèse de récurrence sur I'. Pour le premier terme, il s'agit de

$$\mathbb{P}\left(\bigcup_{i\in I'} A_i\right) \overset{(1)}{=} \sum_{k=1}^{n} \left[(-1)^{k-1} \sum_{J\in\mathcal{P}_k([\![1,n]\!])} \mathbb{P}\left(\bigcap_{j\in J} A_j\right)\right]$$

Pour le troisième terme, c'est essentiellement la même chose :

$$\mathbb{P}\left(\bigcup_{i\in I'} A_i \cap A_{n+1}\right) \overset{(3)}{=} \sum_{k=1}^{n} \left[(-1)^{k-1} \sum_{J\in\mathcal{P}_k([\![1,n]\!])} \mathbb{P}\left(\bigcap_{j\in J} A_j \cap A_{n+1}\right)\right]$$

Constatons que le premier terme se réécrit de la sorte, en faisant apparaître les parties de $[\![1,n+1]\!]$:

$$(1) = \sum_{k=1}^{n} \left[(-1)^{k-1} \sum_{\substack{J\in\mathcal{P}_k([\![1,n+1]\!])\\ n+1\notin J}} \mathbb{P}\left(\bigcap_{j\in J} A_j\right)\right]$$

Concentrons-nous sur le troisième terme. On voit que les intersections du type

$$\bigcap_{j\in J} A_j \cap A_{n+1}$$

où $J \in \mathcal{P}_k([\![1,n]\!])$ peuvent aussi être vus de la sorte :

$$\bigcap_{j\in J'} A_j$$

où $J' \in \mathcal{P}_{k+1}([\![1,n+1]\!])$ contient $n+1$. Ainsi, le troisième terme se réécrit

$$(3) = \sum_{k=1}^{n} \left[(-1)^{k-1} \sum_{\substack{J\in\mathcal{P}_{k+1}([\![1,n+1]\!])\\ n+1\in J}} \mathbb{P}\left(\bigcap_{j\in J} A_j\right)\right]$$

Afin de tenter un regroupement avec (1), on fait le changement de variable $k \leftarrow k+1$:

$$(3) = -\sum_{k=2}^{n+1} \left[(-1)^{k-1} \sum_{\substack{J\in\mathcal{P}_k([\![1,n+1]\!])\\ n+1\in J}} \mathbb{P}\left(\bigcap_{j\in J} A_j\right)\right]$$

A ce stade, on a :

$$\mathbb{P}\left(\bigcup_{i\in I} A_i\right) \overset{(*)}{=} (1) + \underbrace{\mathbb{P}(A_{n+1}) - (3)}_{\overset{\text{déf.}}{=} (2)}$$

Constatons que $\mathbb{P}(A_{n+1})$ est en fait le terme d'ordre 1 dans la somme de (3), car :

$$\{J \in \mathcal{P}_1([\![1,n+1]\!]), n+1 \in J\} = \{n+1\}$$

Si bien que

$$(2) = \mathbb{P}(A_{n+1}) - (3)$$

donne en faisant attention aux signes :

$$(2) = \sum_{k=1}^{n+1} \left[(-1)^{k-1} \sum_{\substack{J\in\mathcal{P}_k([\![1,n+1]\!])\\ n+1\in J}} \mathbb{P}\left(\bigcap_{j\in J} A_j\right)\right]$$

On aboutit alors à notre expression $(*)$:

$$\begin{aligned}(*) = \quad & \sum_{k=1}^{n} \left[(-1)^{k-1} \sum_{\substack{J\in\mathcal{P}_k([\![1,n+1]\!])\\ n+1\notin J}} \mathbb{P}\left(\bigcap_{j\in J} A_j\right)\right] \\ + & \sum_{k=1}^{n+1} \left[(-1)^{k-1} \sum_{\substack{J\in\mathcal{P}_k([\![1,n+1]\!])\\ n+1\in J}} \mathbb{P}\left(\bigcap_{j\in J} A_j\right)\right]\end{aligned}$$

Il reste donc qu'à examiner le terme d'ordre $n+1$ du premier terme. Il s'agit de constater qu'en fait

$$\{J \in \mathcal{P}_{n+1}([\![1,n+1]\!]), n+1 \notin J\} = \varnothing$$

et qu'on fait la convention usuelle que $\sum_\varnothing X = 0$. Le terme d'ordre $n+1$ du premier terme est donc nul, et il est alors possible de regrouper la somme. La dernière chose à constater est alors que pour tout $k \in [\![1,n+1]\!]$:

$$\begin{aligned}\mathcal{P}_k([\![1,n+1]\!]) = \quad & \{J \in \mathcal{P}_k([\![1,n+1]\!]), n+1 \notin J\} \\ \sqcup \ & \{J \in \mathcal{P}_k([\![1,n+1]\!]), n+1 \in J\}\end{aligned}$$

donc on a bien

$$(*) = \sum_{k=1}^{n+1} \left[(-1)^{k-1} \sum_{J\in\mathcal{P}_k([\![1,n+1]\!])} \mathbb{P}\left(\bigcap_{j\in J} A_j\right)\right]$$

Ce qui conclut la récurrence, et la démonstration de cette formule. □

8 Contre-exemple de du Bois-Reymond

Références :

1. *Analyse pour l'agrégation*, Hervé Queffélec, Claude Zuily
2. *Les Contre-exemples en mathématiques*, Bertrand Hauchecorne

Définition 8.1

Soit $f \in \mathcal{C}^0(\mathbb{R})$ 2π-périodique. On définit sa N-ème *somme de Fourier* par :

$$\forall x \in \mathbb{R}, S_N(f,x) \overset{\text{déf.}}{=} \sum_{n=-N}^{N} c_n(f) e^{\imath nt}$$

où

$$\forall n \in \mathbb{Z}, c_n(f) \overset{\text{déf.}}{=} \int_0^{2\pi} f(t) e^{-\imath nt} \, \mathrm{d}t$$

Théorème 8.1 : Contre-exemple de du Bois - Reymond

Il existe une fonction $f \in \mathcal{C}^0(\mathbb{R})$ 2π-périodique telle que

$$\sup_{N \geqslant 1} |S_N(f,0)| = +\infty$$

En particulier, la suite des sommes de Fourier $(S_N(f))_{N \in \mathbb{N}^*}$ diverge en 0.

Lemme 8.1 : Noyau de Gibbs, version ZQ

Soit $N \in \mathbb{N}^*$. On définit le *noyau de Gibbs* par :

$$\forall t \in \mathbb{R}, G_N(t) \overset{\text{déf.}}{=} \sum_{n=1}^{N} \frac{\sin(nt)}{n}$$

Alors

$$\forall N \in \mathbb{N}^*, \|G_N\|_\infty \leqslant \frac{\pi}{2} + 1$$

D'ailleurs, on dispose de l'égalité (qui ne nous servira pas ici) :

$$\forall t \in]0, 2\pi[, \lim_{N \to +\infty} G_N(t) = \frac{\pi - t}{2}$$

Démonstration du lemme : • Pour $r \in [0,1[$, on définit

$$\forall t \in \mathbb{R}, f_r(t) \overset{\text{déf.}}{=} \sum_{n=1}^{+\infty} \frac{r^n \sin(nt)}{n}$$

Alors,

$$\forall t \in \mathbb{R}, f_r(t) = \arctan\left(\frac{r \sin(t)}{1 - r\cos(t)}\right)$$

En effet, pour cela, il suffit de dériver les deux fonctions, de s'apercevoir qu'il s'agit de la même dérivée :

$$f_r'(t) = \frac{r\cos(t) - r^2}{1 - 2r\cos(t) + r^2}$$

et que les deux fonctions coïncident en 0.

• Le théorème d'Abel radial permet de passer à la limite lorsque $[r \to 1^-]$ pour avoir l'égalité

$$\sum_{n=1}^{+\infty} \frac{\sin(nt)}{n} = \frac{\pi - t}{2}$$

Plus précisément, la fonction

$$f \longmapsto f_r(t)$$

est continue sur $[0,1]$, le théorème d'Abel radial donne donc :

$$\sum_{n=1}^{+\infty} \frac{\sin(nt)}{t} = \arctan\left(\frac{\sin(t)}{1-\cos(t)}\right)$$

Les formules de trigonométrie

$$\sin(\theta) = 2\cos\left(\frac{\theta}{2}\right)\sin\left(\frac{\theta}{2}\right)$$

et

$$1-\cos(\theta) = 2\sin^2\left(\frac{\theta}{2}\right)$$

permettent alors d'avoir

$$\sum_{n=1}^{+\infty} \frac{\sin(nt)}{t} = \arctan\left(\frac{\cos\left(\frac{t}{2}\right))}{\sin\left(\frac{t}{2}\right)}\right)$$

Donc, pour tout $t \in \mathbb{R}$,

$$\sum_{n=1}^{+\infty} \frac{\sin(nt)}{t} = \arctan\left(\tan\left(\frac{\pi}{2} - \frac{t}{2}\right)\right)$$

Donc, on peut simplifier ceci pour

$$0 < \frac{\pi}{2} - \frac{t}{2} < \pi$$

d'où, pour $0 < t < 2\pi$, on a bien

$$\sum_{n=1}^{+\infty} \frac{\sin(nt)}{t} = \frac{\pi}{2} - \frac{t}{2}$$

- Revenons à :

$$\forall t \in \mathbb{R}, f_r(t) = \arctan\left(\frac{r\sin(t)}{1-r\cos(t)}\right)$$

Alors, on a immédiatement que

$$\|f_r\|_\infty \leqslant \frac{\pi}{2}$$

L'astuce est ici : on fait la convolution de f_r avec K_N, le noyau de Féjèr, et on a :

$$f_r * K_N(t) = \sum_{n=1}^{N} \left(1 - \frac{n}{N}\right) \frac{r^n \sin(nt)}{n}$$

Et donc, on dispose de l'inégalité suivante :

$$\sum_{n=1}^{N} \left(1 - \frac{n}{N}\right) \frac{r^n \sin(nt)}{n} \leqslant \|f_r\|_\infty \|K_N\|_1 \leqslant \frac{\pi}{2}$$

Or, la somme composée des termes qu'on ne souhaite pas pour Gibbs se majore aisément :

$$\sum_{n=1}^{N} \frac{n}{N} \frac{r^n \sin(nt)}{n} \leqslant \frac{1}{N} \sum_{n=1}^{N} r^n \leqslant 1$$

On a alors la sympathique inégalité

$$\sum_{n=1}^{N} \frac{r^n \sin(nt)}{t} \leqslant \frac{\pi}{2} + 1$$

Par continuité, (c'est une somme finie) on peut alors faire $[r \to 1^-]$, pour finalement avoir

$$\|G_N\|_\infty \leqslant \frac{\pi}{2} + 1$$

□

Lemme 8.2 : Noyau de GIBBS, version HOCHECORNE

On a :

$$\forall n \in \mathbb{N}^*, \forall x \in \mathbb{R}, \left|\sum_{k=1}^{n} \frac{\sin(kx)}{k}\right| \leqslant \pi + 1$$

Démonstration du lemme : Au vu de la parité et de la périodicité de la fonction

$$x \longmapsto \sum_{k=1}^{n} \frac{\sin(kx)}{k}$$

on peut se soustraire à le démontrer uniquement pour $x \in]0, \pi[$ (cette majoration étant vraie pour $x \in \{0, \pi\}$. On note

$$q \overset{\text{déf.}}{=} \left\lfloor \frac{\pi}{x} \right\rfloor$$

Par définition de la partie entière,

$$qx \leqslant \pi < (q+1)x$$

- Supposons $n \leqslant q$. Alors, en utilisant l'inégalité

$\sin(X) \leqslant X$ pour $X \geqslant 0$, on a alors

$$0 \leqslant \sum_{k=1}^{n} \frac{\sin(kx)}{k} \leqslant \sum_{k=1}^{q} \frac{\sin(kx)}{k}$$

d'où

$$\sum_{k=1}^{n} \frac{\sin(kx)}{k} \leqslant qx \leqslant \pi + 1$$

- Supposons que $n > q$.

1. Déjà, par le même calcul qu'au point précédent,

$$\left| \sum_{k=1}^{q} \frac{\sin(kx)}{k} \right| \leqslant \pi$$

2. Reste à s'occuper de la somme entre $q+1$ et n. On pose pour cela :

$$\forall k \in [\![q+1, n]\!], u_k(x) \stackrel{\text{déf.}}{=} \sum_{p=q+1}^{n} \sin(px)$$

et $u_q(x) = 0$. Alors, puisque

$$\sum_{p=q+1}^{k} e^{\imath p\pi} = \frac{\sin\left(\frac{(k-p)x}{2}\right)}{\sin\left(\frac{x}{2}\right)} e^{\imath y}$$

où

$$y = \frac{k+q+1}{2} x \in \mathbb{R}$$

il suit que pour $q-1 \leqslant k \leqslant n$:

$$|u_k(x)| \leqslant \left| \frac{\sin\left(\frac{(k-p)x}{2}\right)}{\sin\left(\frac{x}{2}\right)} \right| \leqslant \frac{1}{\sin\left(\frac{x}{2}\right)}$$

Cela reste vrai pour $k = q$.

3. On fait alors une sommation d'Abel pour faire apparaître $u_k(x)$ dans notre problème, en constatant que pour $k \in [\![q+1, n]\!]$,

$$\sin(kx) = u_k(x) - u_{k-1}(x)$$

Si bien que

$$\sum_{k=q+1}^{n} \frac{\sin(kx)}{k} = \sum_{k=q+1}^{n} \frac{u_k(x) - u_{k-1}(x)}{k}$$

On coupe en deux sommes :

$$\sum_{k=q+1}^{n} \frac{\sin(kx)}{k} = \sum_{k=q+1}^{n} \frac{u_k(x)}{k} - \sum_{k=q+1}^{n} \frac{u_{k-1}(x)}{k}$$

On réindexe la deuxième somme :

$$\sum_{k=q+1}^{n} \frac{\sin(kx)}{k} = \sum_{k=q+1}^{n} \frac{u_k(x)}{k} - \sum_{k=q}^{n-1} \frac{u_k(x)}{k+1}$$

On regroupe les deux sommes en isolant les termes d'ordre n et q :

$$\sum_{k=q+1}^{n} \frac{\sin(kx)}{k} = \sum_{k=q+1}^{n-1} u_k(x) \left[\frac{1}{k} - \frac{1}{k+1}\right] + \frac{u_n(x)}{n} + \frac{u_q(x)}{q+1}$$

Avec $u_q = 0$.

4. On majore le tout avec l'inégalité obtenue en **2.** :

$$\left| \sum_{k=q+1}^{n} \frac{\sin(kx)}{k} \right| \leqslant \frac{1}{\sin\left(\frac{x}{2}\right)} \left[\frac{1}{n} \sum_{k=q+1}^{n} \left(\frac{1}{k} - \frac{1}{k+1}\right)\right]$$

C'est une somme télescopique :

$$\left| \sum_{k=q+1}^{n} \frac{\sin(kx)}{k} \right| \leqslant \frac{1}{\sin\left(\frac{x}{2}\right)} \left[\frac{1}{n} + \frac{1}{q+1} - \frac{1}{n}\right]$$

Enfin, utilisons l'inégalité de concavité de sinus sur $\left[0, \frac{\pi}{2}\right]$:

$$\sin(u) \geqslant \frac{2}{\pi} x$$

Pour aboutir à :

$$\left| \sum_{k=q+1}^{n} \frac{\sin(kx)}{k} \right| \leqslant \frac{\pi}{x(q+1)}$$

Puisque q est la partie entière de $\frac{\pi}{x}$, il suit que

$$\left| \sum_{k=q+1}^{n} \frac{\sin(kx)}{k} \right| \leqslant 1$$

5. Il reste à faire le bilan. On a montré l'inégalité pour $n \leqslant q$. Pour $n > q$, on a :

$$\left| \sum_{k=1}^{n} \frac{\sin(kx)}{k} \right| \leqslant \left| \sum_{k=1}^{q} \frac{\sin(kx)}{k} \right| + \left| \sum_{k=q+1}^{n} \frac{\sin(kx)}{k} \right|$$

donc, on a l'inégalité souhaitée :

$$\left| \sum_{k=1}^{n} \frac{\sin(kx)}{k} \right| \leqslant \pi + 1$$

□

Démonstration : L'argument rapide mais non constructif consiste en l'utilisation du théorème de Banach-Steinhaus avec D_N. En effet, on a :

$$\|D_N\|_{L^1} \xrightarrow[N\to+\infty]{} +\infty$$

On se place dans $\mathcal{C}^0(\mathbb{T})$, l'ensemble des fonctions continues 2π-périodiques, muni de la norme $\|\cdot\|_\infty$. On considère l'application

$$\Lambda_{N,x} : \begin{pmatrix} \mathcal{C}^0(\mathbb{T}) & \longrightarrow & \mathbb{C} \\ f & \longmapsto & S_N(f,x) \end{pmatrix}$$

Alors

$$\|\Lambda_{N,x}\|_{\mathcal{C}^0(\mathcal{T})'} = \|D_N\|_{L^1}$$

Puisque

$$\|D_N\|_{L^1} \xrightarrow[N\to+\infty]{} +\infty$$

Il suit que

$$\forall x \in \mathbb{R}, \sup_{N\in\mathbb{N}} \|\Lambda_{N,x}\|_{\mathcal{C}^0(\mathcal{T})'} = +\infty$$

Donc, d'après la contraposée du théorème de Banach-Steinhaus, puisque l'espace $\left(\mathcal{C}^0(\mathbb{T}), \|\cdot\|_\infty\right)$ est complet, pour tout $x \in \mathbb{R}$, il existe $A(x)$ un ensemble G_δ dense (*id est* qui est une intersection dénombrable d'ouverts) de la forme

$$A(x) = \bigcap_{p\in\mathbb{N}} A_p(x)$$

où $A_p(x)$ est un ouvert dense tel que

$$\forall f \in A(x), \sup_{N\in\mathbb{N}} |\Lambda_{N,x}(f)| = \sup_{N\in\mathbb{N}} |S_N(f,x)| = +\infty$$

Donc pour $f \in A(x)$, $S_N(f,x)$ ne converge pas vers $f(x)$. Si on considère $\{x_k\}_{k\in\mathbb{N}}$ une partie dense dans $\mathbb{R}$, le théorème de Baire assure que

$$A \overset{\text{déf.}}{=} \bigcap_{k\in\mathbb{N}} A(x_k)$$

est dense, et pour tout $f \in A$,

$$\forall k \in \mathbb{N}, \sup_{N\in\mathbb{N}} |S_N(f,x_k)| = +\infty$$

□

Démonstration : On cherche donc ici à expliciter une fonction continue 2π-périodique qui admet une somme de Fourier diergente en 0, à l'aide du noyau de Gibbs. On va la construire pas à pas.

- On considère

$$\forall k \in \mathbb{N}^*, n_k \overset{\text{déf.}}{=} 2^{k^3} \text{ et } \alpha_k \overset{\text{déf.}}{=} \frac{1}{k^2}$$

Alors, on dispose des trois propriétés suivantes :

(i) Pour tout $k \in \mathbb{N}^*$, $n_{k+1} > 3n_k$;

(ii) $\sum_{k\geqslant 1} \alpha_k$ converge ;

(iii) $\alpha_k \ln(n_k) \xrightarrow[k\to+\infty]{} +\infty$

- On pose pour tout $k \in \mathbb{N}^*$,

$$\forall x \in \mathbb{R}, P_k(x) = e^{2\imath n_k x} G_{n_k}(x)$$

où G_{n_k} est le noyau de Gibbs. Alors, on peut développer le sinus en exponentielle, pour avoir

$$P_k(x) = \frac{1}{2\imath} \sum_{j=1}^{n_k} \frac{1}{j} \left[e^{\imath(2n_k+j)x} - e^{\imath(2n_k-j)x}\right]$$

Si bien que

$$\sigma(P_k) \overset{\text{déf.}}{=} \{n \in \mathbb{N}, c_n(P_k) \neq 0\} = [\![n_k, 3n_k]\!]$$

Par la propriété (i), on a alors

$$\sup(\sigma(P_k)) = 3n_k < n_{k+1} = \inf(\sigma(P_{k+1}))$$

- Pour tout $x \in \mathbb{R}$, on définit

$$f(x) \overset{\text{déf.}}{=} \sum_{k=1}^{+\infty} \alpha_k P_k(x)$$

Constatons que f est bien définie, et est même continue *via* le lemme **8.1** :

$$\forall k \in \mathbb{N}^*, \forall x \in \mathbb{R}, |\alpha_k P_k(x)| \leqslant \left(\frac{\pi}{2}+1\right) \alpha_k$$

second membre sommable d'après la propriété (ii) indépendamment de x. Ainsi, $\sum \alpha_k P_k$ converge normalement donc uniformément sur $\mathbb{R}$. f est donc une fonction continue, et 2π-périodique. Montrons que sa série Fourier diverge en 0. Pour cela, la convergence uniforme montre que

$$c_n(f) = \sum_{k=1}^{+\infty} \alpha_k c_n(P_k)$$

Or, $c_n(P_k) = 0$ sauf si $n \in \sigma(P_k)$, par définition de $\sigma(P_k)$, donc, par disjonction des $\sigma(P_k)$, il existe au plus un unique $k \in \mathbb{N}^*$ tel que

$$c_n(f) = \alpha_k c_n(P_k)$$

avec $k \in \mathbb{N}^*$ tel que $n \in [\![n_k, 3n_k]\!]$. Si un tel $k \in \mathbb{N}^*$ n'existe pas, $c_n(f) = 0$.

• Soit $k \in \mathbb{N}^*$. Observons alors $S_{2n_k}(f,x)$ pour $x \in \mathbb{R}$:

$$S_{2n_k}(f,x) = \sum_{n=-2n_k}^{2n_k} c_n(f)e^{\imath nx}$$

Par ce qu'on a montré sur $c_n(f)$, il suit que

$$S_{2n_k}(f,x) = \sum_{n=n_k}^{2n_k} \alpha_k c_n(P_k)e^{\imath nx}$$

On fait apparaître P_k, qui est un polynôme trigonométrique, donc vérifie $P_k = S(P_k)$, qui est une somme finie :

$$S_{2n_k}(f,x) = \sum_{n=n_k}^{3n_k} \alpha_k c_n(P_k)e^{\imath nx} - \sum_{n=2n_k+1}^{3n_k} \alpha_k c_n(P_k)e^{\imath nx}$$

D'où

$$S_{2n_k}(f,x) = \alpha_k P_k(x) - \sum_{n=2n_k+1}^{3n_k} \alpha_k c_n(P_k)e^{\imath nx}$$

L'expression développé de P_k permet alors d'avoir que

$$c_n(P_k) = \frac{1}{2\imath(n-2n_k)}$$

Puisque $P_k(0) = 0$, alors en évaluant en $x = 0$, on obtient :

$$S_{2n_k}(f,0) = \frac{-\alpha_k}{2\imath} \sum_{n=2n_k+1}^{3n_k} \frac{1}{n-2n_k}$$

Soit, par un petit changement de variables :

$$S_{2n_k}(f,0) = \frac{-\alpha_k}{2} \sum_{j=1}^{n_k} \frac{1}{j}$$

• Puisque $n_k \xrightarrow[k\to+\infty]{} +\infty$, on dipose alors de l'équivalent :

$$|S_{2n_k}(f,0)| = \frac{\alpha_k}{2} \sum_{j=1}^{n_k} \frac{1}{j} \sim \frac{\alpha_k}{4}\ln(n_k)$$

Et par la propriété (iii) :

$$\frac{\alpha_k}{2}\ln(n_k) \xrightarrow[k\to+\infty]{} +\infty$$

donc

$$|S_{2n_k}(f,0)| \xrightarrow[k\to+\infty]{} +\infty$$

donc la série de Fourier associée à f diverge en 0 ! □

9 Norme 1 du noyau de Dirichlet

Référence : Analyse pour l'agrégation, Hervé Queffélec, Claude Zuily

Définition 9.1

Soit $N \in \mathbb{N}^*$. On définit le *noyau de Dirichlet d'ordre N* par :

$$\forall t \in \mathbb{R}, D_N(t) = \sum_{k=-n}^{n} e^{\imath kt} = \frac{\sin\left(N+\frac{1}{2}\right)t}{\sin\left(\frac{t}{2}\right)}$$

Notons au passage que si $f \in L^1(\mathbb{R})$, alors

$$\forall x \in \mathbb{R}, D_N * f(x) = \sum_{n=-N}^{N} c_n(f)e^{\imath nt}$$

Proposition 9.1 : Noyau de Dirichlet

On dispose du développement asymptotique suivant :

$$\|D_N\|_{L^1(\mathbb{T})} \underset{N\to+\infty}{=} \frac{4}{\pi^2}\ln(N) + O(1)$$

Démonstration : **1.** Commençons par montrer que la fonction

$$\rho : x \longmapsto \text{cotan}(x) - \frac{1}{x}$$

est une fonction continue sur $]0,1]$, qui se prolonge par continuité en 0, par 0. Il s'agit simplement de faire un développement limité au voisinage de 0. ρ est en effet déjà continue sur $]0,\pi[$. En 0, on a :

$$\tan(x) \underset{x\to 0}{=} x + \frac{x^3}{3} + o(x^3)$$

Donc,

$$\text{cotan}(x) \underset{x\to 0}{=} \frac{1}{x + \frac{x^3}{3} + o(x^3)}$$

Par composition avec $u \longmapsto \frac{1}{1-u}$, il suit :

$$\text{cotan}(x) \underset{x\to 0}{=} \frac{1}{x}\left[1 - \frac{x^2}{3} + o(x^2)\right]$$

donc

$$\rho(x) \underset{x\to 0}{=} \frac{-x}{3} + o(x)$$

donc ρ se prolonge bien par continuité en 0. De même en π. On prolonge enfin ρ par imparité sur $[-\pi,\pi]$.

2. Soit $N \in \mathbb{N}^*$. Montrons qu'il existe $r_N \in \mathcal{C}^0([0,\pi])$ telle que

$$\forall x \in [-\pi,\pi], D_N(x) = \frac{2\sin(Nx)}{x} + r_N(x)$$

avec

$$\sup_{\substack{x\in[-\pi,\pi]\\ N\in\mathbb{N}}} |r_N(x)| < +\infty$$

Pour cela, on part du fait que pour tout $x \in \mathbb{R}\backslash 2\pi\mathbb{Z}$,

$$D_N(x) = \frac{\sin\left[\left(N+\frac{1}{2}\right)x\right]}{\sin\left(\frac{x}{2}\right)}$$

On développe le sinus au numérateur :

$$\sin\left(\left[N + \frac{1}{2}\right]x\right) = \sin(Nx)\cos\left(\frac{x}{2}\right) + \cos(Nx)\sin\left(\frac{x}{2}\right)$$

Ainsi, pour $x \in]0,\pi[$:

$$D_N(x) = \sin(Nx)\ \text{cotan}\left(\frac{x}{2}\right) + \cos(Nx)$$

Ainsi, par **1.**, on a :

$$D_N(x) = \sin(Nx)\left[\rho\left(\frac{x}{2}\right) + \frac{2}{x}\right] + \cos(Nx)$$

D'où pour $x \in]0,\pi[$:

$$D_N(x) = \frac{2\sin(Nx)}{x} + r_N(x)$$

où

$$r_N(x) \overset{\text{déf.}}{=} \sin(Nx)\rho\left(\frac{x}{2}\right) + \cos(Nx)$$

On prolonge cette égalité par continuité en 0, et par imparité sur $[-\pi,\pi]$. De plus, pour tout $x \in [-\pi,\pi]$, pour tout $N \in \mathbb{N}^*$:

$$|r_N(x)| \leqslant 1 + \sup_{x\in[-\pi,\pi]|\rho(x)|} < +\infty$$

3. On n'a plus qu'à effectivement se coller à ce calcul d'intégrale. On a :

$$\|D_N\|_{L^1(\mathbb{T})} = \frac{1}{2\pi}\int_{-\pi}^{\pi} |D_N(t)|\ \mathrm{d}t$$

Par la question précédente, il suit que lorsque $[N \to +\infty]$

$$\|D_N\|_{L^1(\mathbb{T})} = \frac{1}{\pi}\int_{-\pi}^{\pi} \left|\frac{\sin(Nt)}{t}\right| \mathrm{d}t + O(1)$$

On utilise la parité de la partie à l'intérieure de l'intégrale pour réduire les bornes à $[0,\pi]$:

$$\|D_N\|_{L^1(\mathbb{T})} = \frac{2}{\pi}\int_{0}^{\pi} \left|\frac{\sin(Nt)}{t}\right| \mathrm{d}t + O(1)$$

On fait le changement de variables $y = Nt$:

$$\|D_N\|_{L^1(\mathbb{T})} = \frac{2}{\pi}\int_{0}^{N\pi} \left|\frac{\sin(y)}{y}\right| \mathrm{d}y + O(1)$$

Puis, on "jette" dans le $O(1)$ l'intégrale entre 0 et 1 :

$$\|D_N\|_{L^1(\mathbb{T})} = \frac{2}{\pi}\int_{1}^{N\pi} \left|\frac{\sin(y)}{y}\right| \mathrm{d}y + O(1)$$

On introduit la fonction φ :

$$\forall x \in \mathbb{R}, \varphi(x) = \int_0^x |\sin(t)|\ \mathrm{d}t - \left[\frac{1}{2\pi}\int_0^{2\pi} |\sin(t)|\ \mathrm{d}t\right] x$$

φ est de classe C^1 sur $\mathbb{R}$, et est 2π-périodique :

$$\varphi(x+2\pi) - \varphi(x) = \int_x^{x+2\pi} |\sin(t)|\ \mathrm{d}t - 2\pi\left[\frac{1}{2\pi}\int_0^{2\pi} |\sin(t)|\ \mathrm{d}t\right]$$

Or,

$$\int_x^{x+2\pi} |\sin(t)|\ \mathrm{d}t = \int_0^{2\pi} |\sin(t)|\ \mathrm{d}t$$

d'où

$$\varphi(x+2\pi)-\varphi(x)=0$$

Ainsi, on peut dériver φ :

$$\varphi'(x)=|\sin(x)|-\int_0^{2\pi}|\sin(t)|\,\mathrm{d}t$$

avec $\int_0^{2\pi}|\sin(t)|\,\mathrm{d}t=\frac{2}{\pi}$. Ainsi, cela donne :

$$\|D_N\|_{L^1(\mathbb{T})}=\frac{2}{\pi}\int_1^{N\pi}\frac{\varphi'(y)}{y}\mathrm{d}y-\frac{4}{\pi^2}\int_1^{N\pi}\frac{\mathrm{d}y}{y}+O(1)$$

On reconnait $\frac{4}{\pi^2}\ln(N\pi)$ au deuxième terme. Quant au premier, on fait une intégration par parties :

$$\int_1^{N\pi}\frac{\varphi'(y)}{y}\mathrm{d}y=\int_1^{N\pi}\frac{\varphi(y)}{y^2}\,\mathrm{d}y+O(1)$$

Puisque φ est 2π-périodique et continue, elle est bornée, donc

$$\left|\int_1^{N\pi}\frac{\varphi(y)}{y^2}\,\mathrm{d}y\right|\leqslant\|\varphi\|_{L^\infty}$$

donc on a réussi à montrer que

$$\|D_N\|_{L^1(\mathbb{T})}=\frac{4}{\pi^2}\ln(N)+O(1)$$

□

10 Théorème de Féjèr - Théorème de Weierstrass

Référence : Analyse pour l'agrégation, Hervé QUEFFÉLEC, Claude ZUILY

Définition 10.1

Soit $f\in L^1(\mathbb{T})$. On définit pour $N\geqslant 1$ la N-ème *somme de FÉJÈR* :

$$\sigma_N(f)\overset{\text{déf.}}{=}K_N*f$$

où K_N est le *noyau de FÉJÈR* :

$$\forall t\in\mathbb{R},K_N(t)\overset{\text{déf}}{=}\frac{1}{N}\sum_{n=0}^{N-1}D_N(t)$$

Proposition 10.1 : Sur le noyau de FÉJÈR

(i) On a, pour $N\in\mathbb{N}^*$:

$$K_N=\sum_{n=-N}^{N}\left(1-\frac{|n|}{N}\right)e_n\text{ et }K_N(x)=\frac{1}{N}\left[\frac{\sin\left(\frac{Nx}{2}\right)}{\sin\left(\frac{x}{2}\right)}\right]^2$$

(ii) $\|K_N\|_{L^1(\mathbb{T})}=1$

(iii) Si $0<\delta\leqslant\pi$,

$$\lim_{N\to+\infty}\int_\delta^\pi K_N(t)\,\mathrm{d}t=0$$

(iv) Pour tout $f\in L^1(\mathbb{T})$, pour tout $N\in\mathbb{N}^*$:

$$\sigma_N(f)=\sum_{n=-N}^{N}\left(1-\frac{|n|}{N}\right)c_n(f)e_n$$

Théorème 10.1 : Théorème de convergence de FÉJÈR

(i) Soit $f \in \mathcal{C}^0(\mathbb{T})$ continue 2π-périodique. Alors

$$\forall N \geqslant 1, \|\sigma_N(f)\|_\infty \leqslant \|f\|_\infty$$

Et :

$$\|\sigma_N(f) - f\|_\infty \xrightarrow[N\to+\infty]{} 0$$

(ii) Soient $p \in [1, +\infty[$ et $f \in L^p(\mathbb{T})$. Alors

$$\forall N \geqslant 1, \|\sigma_N(f)\|_{L^p} \leqslant \|f\|_{L^p}$$

Et :

$$\|\sigma_N(f) - f\|_{L^p} \xrightarrow[N\to+\infty]{} 0$$

Démonstration : [i] • Pour la première majoration, on observe que pour tout $x \in \mathbb{R}$,

$$|[\sigma_N f](x)| \leqslant \|K_N\|_{L^1} \|f\|_\infty = \|f\|_\infty$$

d'où

$$\|\sigma_N f\|_\infty \leqslant \|f\|_\infty$$

• Par définition du produit de convolution, pour $x \in [0, \pi]$, on a :

$$\begin{aligned} [\Delta_N(f)](x) &\overset{\text{déf.}}{=} \sigma_N f(x) - f(x) \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi} [f(x-t) - f(x)]\, K_N(t) \text{ d}t \end{aligned}$$

• On considère $\delta > 0$. On coupe en deux intégrales :

$$\begin{aligned} [\Delta_N(f)](x) = &\quad \frac{1}{2\pi}\int_{-\delta}^{\delta} [f(x-t) - f(x)]\, K_N(t) \text{ d}t \\ &+ \frac{1}{2\pi}\int_{|t|\geqslant\delta} [f(x-t) - f(x)]\, K_N(t) \text{ d}t \end{aligned}$$

On note

$$\omega(\delta) \overset{\text{déf.}}{=} \sup\{|f(u) - f(v)|, |u-v| \leqslant \delta\}$$

Alors $\omega(\delta) < +\infty$ par continuité uniforme de f sur $\mathbb{R}$. On peut donc majorer la première intégrale :

$$\frac{1}{2\pi}\int_{-\delta}^{\delta} [f(x-t) - f(x)]\, K_N(t) \text{ d}t \leqslant \frac{\omega(\delta)}{2\pi}\|K_N\|_{L^1}$$

Quant à la deuxième intégrale, observons que sur $[-\pi, \pi]\backslash[-\delta, \delta]$:

$$K_N(t) = \frac{1}{N}\left[\frac{\sin\left(\frac{Nx}{2}\right)}{\sin\left(\frac{x}{2}\right)}\right]^2 \leqslant \frac{1}{N\sin^2\left(\frac{\delta}{2}\right)}$$

Le tout donne alors :

$$|[\Delta_N(f)](x)| \leqslant \frac{\omega(\delta)}{2\pi} + \frac{2\|f\|_\infty(\pi-\delta)}{\pi N \sin^2\left(\frac{\delta}{2}\right)}$$

Ainsi, on a obtenu une majoration indépendante de x, donc

$$\|\Delta_N(f)\|_\infty \leqslant \frac{\omega(\delta)}{2\pi} + \frac{2\|f\|_\infty}{N\sin^2\left(\frac{\delta}{2}\right)}$$

Et on peut passer à la limite supérieure dans cette inégalité :

$$\limsup_{N\to+\infty} \|\Delta_N(f)\|_\infty \leqslant \frac{\omega(\delta)}{2\pi}$$

• Montrons que

$$\omega(\delta) \xrightarrow[\delta\to 0]{} 0$$

L'argument principal est la continuité uniforme de f sur $\mathbb{R}$: Soit $\varepsilon > 0$. Il existe $\delta_0 > 0$ tel que pour tout $x, y \in \mathbb{R}$, si $|x-y| < \delta_0$ alors

$$|f(x) - f(y)| < \frac{\varepsilon}{2}$$

Ainsi,

$$\omega(\delta_0) \leqslant \frac{\varepsilon}{2} < \varepsilon$$

De plus, pour tout $0 < \delta < \delta_0$, si $|x-y| < \delta$ alors en en particulier $|x-y| < \delta_0$, et donc

$$\forall 0 < \delta < \delta_0, \omega(\delta) < \varepsilon$$

C'est donc la définition de $\lim_{\delta \to 0} \omega(\delta) = 0$.

- Par conséquent, on peut passer à la limite lorsque $[\delta \to 0]$ pour avoir

$$0 \leqslant \limsup_{N \to +\infty} \|\Delta_N(f)\|_\infty \leqslant 0$$

D'où

$$\lim_{N \to +\infty} \|\sigma_N f - f\|_\infty = 0$$

[ii] Soit $p \in [1, +\infty[$.

- Pour tout $u \in \mathbb{R}$,

$$|[\sigma_N f](u)|^p = \left| \frac{1}{2\pi} \int_{-\pi}^{\pi} f(u-t) K_N(t) \, \mathrm{d}t \right|^p$$

En écrivant

$$f(u-t)K_N(t) = f(u-t)K_N(t)^{\frac{1}{p}} K_N(t)^{1-\frac{1}{p}}$$

On utilise l'inégalité de Hölder :

$$\begin{aligned} |[\sigma_N f](u)|^p \leqslant \quad & \frac{1}{(2\pi)^p} \left(\int_{-\pi}^{\pi} |f(u-t)|^p K_N(t) \, \mathrm{d}t \right) \\ \times & \left(\int_{-\pi}^{\pi} K_N(t) \, \mathrm{d}t \right)^{p(1-\frac{1}{p})} \end{aligned}$$

Donc,

$$|[\sigma_N f](u)|^p \leqslant \frac{1}{2\pi} \int_{-\pi}^{\pi} |f(u-t)|^p K_N(t) \, \mathrm{d}t$$

Ainsi,

$$\|\sigma_N f\|_{L^p}^p = \frac{1}{2\pi} \int_{-\pi}^{\pi} |[\sigma_N f](u)|^p \, \mathrm{d}u$$

se majore :

$$\|\sigma_N f\|_{L^p}^p \leqslant \frac{1}{4\pi^2} \int_{-\pi}^{\pi} \left[\int_{-\pi}^{\pi} |f(u-t)|^p K_N(t) \, \mathrm{d}t \right] \, \mathrm{d}u$$

Ce qui se situe sous l'intégrale est positif, d'après le théorème de Fubini-Tonelli :

$$\|\sigma_N f\|_{L^p}^p \leqslant \frac{1}{4\pi^2} \int_{-\pi}^{\pi} K_N(t) \left[\int_{-\pi}^{\pi} |f(u-t)|^p \, \mathrm{d}u \right] \, \mathrm{d}t$$

$f \in L^1(\mathbb{T})$, donc

$$\int_{-\pi}^{\pi} |f(u-t)|^p \, \mathrm{d}u = \int_{-\pi}^{\pi} |f(u)|^p \, \mathrm{d}u = 2\pi \|f\|_{L^p}^p$$

Ainsi,

$$\|\sigma_N f\|_{L^p}^p \leqslant \frac{1}{2\pi} \int_{-\pi}^{\pi} K_N(t) \|f\|_{L^p}^p \mathrm{d}t$$

Soit

$$\|\sigma_N f\|_{L^p}^p \leqslant \|K_N\|_{L^1} \|f\|_{L^p}^p$$

avec $\|K_N\|_{L^1} = 1$, d'où

$$\|\sigma_N f\|_{L^p}^p \leqslant \|f\|_{L^p}^p$$

- On note comme dans **(i)** :

$$\Delta_N f \overset{\text{déf.}}{=} \sigma_N f - f$$

Le même type d'astuce qu'au premier point conduit à l'inégalité

$$\|\Delta_N f\|_{L^p}^p \leqslant \frac{1}{4\pi^2} \int_{-\pi}^{\pi} \left[\int_{-\pi}^{\pi} |f(u-t) - f(u)|^p K_N(t) \, \mathrm{d}t \right] \, \mathrm{d}u$$

D'après le théorème de Fubini-Tonelli :

$$\|\Delta_N f\|_{L^p}^p \leqslant \frac{1}{4\pi^2} \int_{-\pi}^{\pi} \left[\int_{-\pi}^{\pi} |f(u-t) - f(u)|^p \, \mathrm{d}u \right] K_N(t) \, \mathrm{d}t$$

On introduit la fonction définie sur $[-\pi, \pi]$ par :

$$g(t) \overset{\text{déf.}}{=} \frac{1}{2\pi} \int_{-\pi}^{\pi} |f(u+t) - f(u)|^p \, \mathrm{d}u$$

g est une fonction continue et 2π-périodique, d'après le résultat suivant (qui se montre par densité de $\mathcal{C}^0$ dans L^p muni de sa norme usuelle) :

$$\lim_{a \to 0} \|\tau_a f - f\|_{L^p} = 0$$

Ainsi, d'après le point **(i)**,

$$\|\sigma_N g - g\|_\infty \xrightarrow[N \to +\infty]{} 0$$

Or,

$$\|\Delta_N f\|_{L^p}^p \leqslant \frac{1}{2\pi} \int_{-\pi}^{\pi} g(-t) K_N(t) \, \mathrm{d}t$$

donc

$$\|\Delta_N f\|_{L^p}^p \leqslant \sigma_N(g, 0)$$

Il suit donc que

$$\|\sigma_N f - f\|_{L^p} \xrightarrow[N \to +\infty]{} 0$$

□

Remarque : Si $f \in L^1(\mathbb{R})$, T-périodique presque partout, et $a \in \mathbb{R}$:

$$\int_a^{a+T} f(t)\ \mathrm{d}t = \int_a^0 f(t)\ \mathrm{d}t + \int_0^T f(t)\ \mathrm{d}t + \int_T^{a+T} f(t)\ \mathrm{d}t$$

On fait un changement de variables dans la troisième intégrale :

$$\int_a^{a+T} f(t)\ \mathrm{d}t = \int_a^0 f(t)\ \mathrm{d}t + \int_0^T f(t)\ \mathrm{d}t + \int_0^a f(t+T)\ \mathrm{d}t$$

Puis f est T-périodique, donc

$$\int_a^{a+T} f(t)\ \mathrm{d}t = \int_a^0 f(t)\ \mathrm{d}t + \int_0^T f(t)\ \mathrm{d}t + \int_0^a f(t)\ \mathrm{d}t$$

donc

$$\int_a^{a+T} f(t)\ \mathrm{d}t = \int_0^T f(t)\ \mathrm{d}t$$

Corollaire 10.1 : Théorème de WEIERSTRASS

Toute fonction continue sur un compact $[a, b]$ est limite uniforme sur $[a, b]$ d'une suite de polynômes algébriques.

Démonstration : On le montre uniquement sur $[-1, 1]$, une petite translation permet de la montrer sur $[a, b]$.

- Soit $F \in \mathcal{C}([-1, 1])$. On définit

$$\forall t \in \mathbb{R}, f(t) \overset{\text{déf.}}{=} F(\cos(t))$$

f est 2π-périodique. f est paire, donc $c_n(f) = c_{-n}(f)$. Soit $N \in \mathbb{N}^*$. Ainsi, sa somme de Féjèr qui se calcule par :

$$\sigma_N f = \sum_{n=-N}^{N} \left(1 - \frac{|n|}{N}\right) c_n(f) e_n$$

se simplifie en :

$$\sigma_N f = c_0(f) e_0 + \sum_{n=1}^{N} \left(1 - \frac{n}{N}\right) c_n(f)(e_n + e_{-n})$$

Ainsi, pour tout $t \in \mathbb{R}$,

$$\left[\sigma_N f\right](t) = c_0(f) + 2 \sum_{n=1}^{N} \left(1 - \frac{n}{N}\right) c_n(f) \cos(nt)$$

- Or, il existe un unique polynôme $T_n \in \mathbb{R}_n[X]$ tel que

$$\forall t \in \mathbb{R}, \cos(nt) = T_n(\cos t)$$

C'est la famille des polynômes de Tchebychev de première espèce. Par conséquent, on a :

$$\forall t \in \mathbb{R}, \left[\sigma_N f\right](t) = P_N(\cos t)$$

où

$$P_N \overset{\text{déf.}}{=} c_0(f) + 2 \sum_{n=1}^{N} \left(1 - \frac{n}{N}\right) c_n(f) T_n$$

- C'est cette suite qui nous permettra de conclure. On a par surjectivité du cosinus dans $[-1, 1]$:

$$\sup_{x \in [-1,1]} |F(x) - P_N(x)| = \sup_{t \in \mathbb{R}} |F(\cos t) - P_N(\cos t)|$$

Autrement dit,

$$\sup_{x \in [-1,1]} |F(x) - P_N(x)| = \sup_{t \in \mathbb{R}} |f(t) - \left[\sigma_N f\right](t)|$$

Donc, pour retrouver un visage connu :

$$\sup_{x \in [-1,1]} |F(x) - P_N(x)| = \|f - \sigma_N f\|_\infty$$

f étant continue et 2π-périodique, le théorème de Féjèr permet de conclure que

$$\|f - \sigma_N f\|_\infty \xrightarrow[N \to +\infty]{} 0$$

D'où

$$\sup_{x \in [-1,1]} |F(x) - P_N(x)| \xrightarrow[N \to +\infty]{} 0$$

□

11 Formule sommatoire de Poisson

Référence : Analyse pour l'agrégation, Hervé QUEFFÉLEC, Claude ZUILY
Dans cette section, pour $F \in L^1(\mathbb{R})$, on définit sa *transformée de Fourier* par :

$$\forall x \in \mathbb{R}, \hat{F}(x) \overset{\text{déf.}}{=} \int_R e^{2\imath\pi x\nu} F(\nu) \, \mathrm{d}\nu$$

Théorème 11.1 : Formule sommatoire de POISSON

Soit $F \in L^1(\mathbb{R}) \cap \mathcal{C}^0(\mathbb{R})$. On suppose :

(i) Il existe $M > 0$ et $\alpha > 1$ tels que

$$\forall x \in \mathbb{R}, |F(x)| \leqslant \frac{M}{(1+|x|)^\alpha}$$

(ii) $\sum_{n\in\mathbb{Z}} \left|\hat{F}(n)\right| < +\infty$

Alors

$$\sum_{n\in\mathbb{Z}} F(n) = \sum_{n\in\mathbb{Z}} \hat{F}(n)$$

Remarque : Si $F \in \mathcal{S}(\mathbb{R})$, les hypothèses (i) et (ii) sont satisfaites.

Démonstration : On définit :

$$\forall x \in \mathbb{R}, f(x) \overset{\text{déf.}}{=} \sum_{n\in\mathbb{Z}} F(x+n)$$

• Montrons que f est bien définie, en montrant que la série converge simplement sur $\mathbb{R}$. Soit $x \in \mathbb{R}$. Pour cela, on constate que d'une part,

$$\forall n \in \mathbb{Z}\backslash\{0\}, n \geqslant |x|, |F(x+n)| \leqslant \frac{M}{(1+n-|x|)^\alpha}$$

Donc, puisque $|x| \geqslant \lfloor |x| \rfloor$:

$$\forall n \in \mathbb{Z}\backslash\{0\}, n \geqslant |x|, |F(x+n)| \leqslant \frac{M}{(1+|n|-\lfloor |x| \rfloor)^\alpha}$$

avec $\alpha > 1$, donc $\sum_{n\geqslant \lfloor |x| \rfloor} F(x+n)$ et $\sum_{n\leqslant -\lfloor |x| \rfloor} F(x+n)$ convergent, donc $\sum_{n\in\mathbb{Z}} F(x+n)$ converge simplement sur $\mathbb{R}$.

• Soit K un compact de $\mathbb{R}$. Alors

$$\forall x \in K, \forall n \in \mathbb{Z}, |F(x+n)| \leqslant \sup_{x\in K} \frac{M}{(1+|n+x|)^\alpha}$$

K étant compact, cette borne supérieure est atteinte pour un certain $x_0 \in K$:

$$\forall x \in K, \forall n \in \mathbb{Z}, |F(x+n)| \leqslant \frac{M}{(1+|n+x_0|)^\alpha}$$

C'est le terme général d'une série convergente, donc $\sum_{n\in\mathbb{Z}} F(x+n)$ converge normalement sur tout compact, donc uniformément sur tout compact ; il suit que f est continue sur $\mathbb{R}$.

• De plus, pour tout $x \in \mathbb{R}$,

$$f(x+1) = \sum_{n\in\mathbb{Z}} F(x+n+1)$$

donc par un petit changement de variables

$$f(x+1) = \sum_{n\in\mathbb{Z}} F(x+n)$$

donc

$$f(x+1) = f(x)$$

donc f est continue et 1-périodique. Pour qu'elle soit développable en série de Fourier, il reste à montrer que $\sum_{n\in\mathbb{Z}} |c_n(f)| < \infty$.

• On a :

$$c_n(f) = \int_0^1 \left[\sum_{k\in\mathbb{Z}} F(t+k)\right] e^{2\imath\pi tn} \, \mathrm{d}t$$

Par convergence normale sur tout compact,

$$c_n(f) = \sum_{k\in\mathbb{Z}} \int_0^1 F(t+n) e^{2\imath\pi nt} \, \mathrm{d}t$$

On fait un petit changement de variables :

$$c_n(f) = \sum_{n\in\mathbb{Z}} \int_n^{n+1} F(u)e^{2\imath\imath n(u-k)}\ \mathrm{d}u$$

Ainsi, avec $e^{2\pi\imath kn} = 1$, il suit

$$c_n(f) = \int_{\mathbb{R}} F(u)e^{2\imath\pi nu}\ \mathrm{d}u$$

Or, par l'hypothèse (ii), $\sum_{n\in\mathbb{Z}} |\hat{F}(n)|^2 < \infty$, donc on a convergence normale de la série de Fourier associée à f.

On peut alors développer en série de Fourier f : pour tout $x \in \mathbb{R}$

$$f(x) = \sum_{n\in\mathbb{Z}} \hat{F}(n)e^{2\imath\pi nx}$$

On évalue en 0 pour finalement alors

$$\sum_{n\in\mathbb{Z}} F(n) = \sum_{n\in\mathbb{Z}} \hat{F}(n)$$

□

12 Inégalité de Bernstein

Référence : Analyse pour l'agrégation, Hervé Queffélec, Claude Zuily

Théorème 12.1 : Inégalité de Bernstein

Soit $\lambda > 0$. On considère $\lambda_1, \cdots, \lambda_n \in \mathbb{R}$ distincts tels que $\max_{1\leqslant k\leqslant n} |\lambda_k| \leqslant \lambda$. On considère $a_1, \cdots, a_n \in \mathbb{C}$, et l'application

$$h : \left(\begin{array}{ccc} \mathbb{R} & \longrightarrow & \mathbb{C} \\ t & \longmapsto & \sum_{k=1}^n a_k e^{\imath\lambda_k t} \end{array}\right)$$

Alors h vérifie

$$\|h'\|_\infty \leqslant \lambda \|h\|_\infty$$

Remarque : h n'est pas en général périodique !

Remarque : Ce théorème généralise l'observation suivante :

$$\frac{\mathrm{d}}{\mathrm{d}t}[\sin(\lambda t)] = \lambda \sin\left(\lambda\left(t + \frac{\pi}{2\lambda}\right)\right)$$

Démonstration : L'objectif de cette démonstration est d'écrire :

$$h'(t) = \sum_{k\in\mathbb{Z}} c_k h(t + t_k)$$

avec $(t_k)_{k\in\mathbb{Z}}$ une suite réelle et $\sum_{k\in\mathbb{Z}} |c_k| = \lambda$.

- On a :

$$h'(t) = \sum_{j=1}^n \imath\lambda_j a_j e^{\imath\lambda_j t}$$

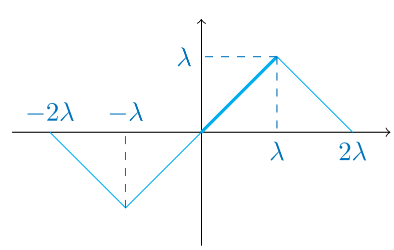
L'astuce réside ici à noter :

$$h'(t) = \sum_{j=1}^n \imath\varphi(\lambda_j) a_j e^{\imath\lambda_j t}$$

où :

$$\varphi(x) = \left\{\begin{array}{rcl} x & \text{si} & x \in [0, \lambda] \\ 2\lambda - x & \text{si} & \in [\lambda, 2\lambda] \end{array}\right.$$

On prolonge alors φ par imparité sur $[-2\lambda, 2\lambda]$. Enfin, on prolonge φ par 4λ-périodicité sur $\mathbb{R}$. φ est une fonction continue, C^1 par morceaux et périodique. On va la développer en série de Fourier pour déterminer la forme recherchée pour h'.



- On pose pour cela

$$S(x) \overset{\text{déf.}}{=} \varphi\left(\frac{2\lambda}{\pi}x\right)$$

On note $\omega \overset{\text{déf.}}{=} \frac{\pi}{2\lambda}$. S est 2π-périodique, tout aussi continue et C^1 par morceaux que φ. Pour $n \in \mathbb{Z}$, son n-ème coefficient de Fourier est donné par :

$$c_n(S) = \frac{1}{2\pi}\int_{-\pi}^{\pi} S(x)e^{\imath nx}\,\mathrm{d}x$$

Par définition de S, cela donne

$$c_n(S) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \varphi\left(\frac{x}{\omega}\right)e^{\imath nx}\,\mathrm{d}x$$

Un petit changement de variables donne alors

$$c_n(S) = \frac{\omega}{2\pi}\int_{-2\lambda}^{2\lambda} \varphi(\tau)\,e^{\imath n\omega\tau}\,\mathrm{d}\tau$$

On découpe ainsi cette intégrale en trois :

$$\begin{aligned} c_n(S) = \quad & \frac{\omega}{2\pi}\int_{-\lambda}^{\lambda} \tau e^{\imath n\omega\tau}\,\mathrm{d}\tau \\ + \quad & \frac{\omega}{2\pi}\int_{-2\lambda}^{-\lambda} (-2\lambda-\tau)e^{\imath n\omega\tau}\,\mathrm{d}\tau \\ + \quad & \frac{\omega}{2\pi}\int_{\lambda}^{2\lambda} (2\lambda-\tau)e^{\imath n\omega\tau}\,\mathrm{d}\tau \end{aligned}$$

On notera de façon plus concise :

$$c_n(S) = (1)+(2)+(3)$$

- Pour le terme (2), on fait le changement de variable $-\tau \to \tau$ pour avoir

$$(2) = \frac{\omega}{2\pi}\int_{\lambda}^{2\lambda} (-2\lambda+\tau)e^{-\imath n\omega\tau}\,\mathrm{d}\tau$$

Ainsi,

$$(2)+(3) = \frac{\omega}{2\pi}\int_{\lambda}^{2\lambda} \left[(-2\lambda+\tau)e^{-\imath n\omega\tau} + (2\lambda-\tau)e^{\imath n\omega\tau}\right]\,\mathrm{d}\tau$$

On notera par la suite $(2') = (2)+(3)$.

$$(2') = \frac{\omega}{2\pi}\int_{\lambda}^{2\lambda} \left[2\lambda\cdot 2\imath\sin(n\omega\tau) - 2\imath\tau\sin(n\omega\tau)\right]\,\mathrm{d}\tau$$

$(2')$ se redécoupe donc en deux intégrales $(2'')+(3'')$, où

$$(2'') = \frac{\omega}{2\pi}\int_{\lambda}^{2\lambda} 4\imath\lambda\imath\sin(n\omega\tau)\,\mathrm{d}\tau$$

donc, on peut calculer l'intégrale

$$(2'') = \frac{2\imath\omega\lambda}{n\pi\omega}\left[\cos(\lambda n\omega) - \cos(2\lambda n\omega)\right]$$

Soit encore, avec $2\lambda\omega = \pi$:

$$(2'') = \frac{2\imath\lambda}{n\pi}\left[\cos\left(\frac{n\pi}{2}\right) - \cos(n\pi)\right]$$

La formule

$$\cos(p) - \cos(q) = -2\sin\left(\frac{p+q}{2}\right)\sin\left(\frac{p-q}{2}\right)$$

donne

$$(2'') = \frac{-4\imath\lambda}{n\pi}\sin\left(\frac{3n\pi}{4}\right)\sin\left(\frac{-n\pi}{2}\right)$$

- Le terme $(3'')$ est donné par :

$$(3'') = \frac{-\imath\omega}{\pi}\int_{\lambda}^{2\lambda} \tau\sin(n\omega\tau)\,\mathrm{d}\tau$$

Une fait alors une bonne intégration par parties des familles

$$\begin{aligned} (3'') = \quad & \frac{-\imath\omega}{\pi}\left(\frac{-2\lambda\cos(2n\omega\lambda)}{n\omega} - \frac{-\lambda\cos(n\omega\lambda)}{n\omega}\right) \\ + \quad & \frac{-\imath\omega}{\pi}\int_{\lambda}^{2\lambda} \frac{-\cos(n\omega\tau)}{n\omega}\,\mathrm{d}\tau \end{aligned}$$

Soit, après réorganisation et calcul d'intégrale :

$$\begin{aligned} (3'') = \quad & \frac{-\imath\lambda}{n\pi}\left(-2\cos(\pi n) + \cos\left(n\frac{\pi}{2}\right)\right) \\ + \quad & \frac{\imath}{n^2\omega\pi}\left(\sin(2n\omega\lambda) - \sin(n\omega\lambda)\right) \end{aligned}$$

Puis, avec un coup de $2\lambda\omega = \pi$, on a finalement

$$\begin{aligned} (3'') = \quad & \frac{-\imath\lambda}{n\pi}\left(-2\cos(\pi n) + \cos\left(n\frac{\pi}{2}\right)\right) \\ + \quad & \frac{\imath}{n^2\omega\pi}\left(\sin(\pi n) - \sin\left(\frac{n\pi}{2}\right)\right) \end{aligned}$$

Puisque $\sin(n\pi) = 0$, on se contente donc d'obtenir :

$$\begin{aligned} (3'') = \quad & \frac{-\imath\lambda}{n^2\omega\pi}\left(-2\cos(\pi n) + \cos\left(n\frac{\pi}{2}\right)\right) \\ + \quad & \frac{-\imath}{n^2\omega\pi}\sin\left(\frac{n\pi}{2}\right) \end{aligned}$$

- Occupons-nous du terme (1), où on va gentiment refaire une intégration par parties :

$$(1) = \frac{\omega}{2\pi}\int_{-\lambda}^{\lambda} \tau e^{\imath n\omega\tau}\,\mathrm{d}\tau$$

L'intégration par parties donne

$$\begin{aligned} (1) = \quad & \frac{\omega}{2\pi}\left[\frac{\lambda e^{\imath n\omega\lambda}}{\imath n\omega} + \frac{\lambda e^{-\imath n\omega\lambda}}{\imath n\omega}\right] \\ - \quad & \frac{\omega}{2\pi}\int_{-\lambda}^{\lambda} \frac{e^{\imath n\omega\tau}}{\imath n\omega}\,\mathrm{d}\tau \end{aligned}$$

On réorganise, on intègre :

$$\begin{aligned} (1) = \quad & \frac{\lambda}{2\imath n\pi}\cdot 2\cos(n\omega\lambda) \\ - \quad & \frac{1}{2\imath n\pi}\left[\frac{e^{\imath n\omega\lambda}}{\imath n\omega} - \frac{e^{-\imath n\omega\lambda}}{\imath n\omega}\right] \end{aligned}$$

On utilise $2\lambda\omega = \pi$:

$$(1) = \frac{\lambda}{\imath n\pi}\cos\left(\frac{n\pi}{2}\right) - \frac{1}{-2n^2\omega\pi}\cdot 2\imath\sin(n\omega\lambda)$$

Soit encore

$$(1) = \frac{-\imath\lambda}{n\pi}\cos\left(\frac{n\pi}{2}\right) + \frac{\imath}{n^2\omega\pi}\sin\left(\frac{n\pi}{2}\right)$$

- Reste alors à tout remettre en place. On a

$$c_n(S) = (1) + (2'') + (3'')$$

On va aligner tout ce qu'on a eu, et mixer le tout :

$$\begin{aligned} c_n(S) = & \frac{-\imath\lambda}{n\pi}\cos\left(\frac{n\pi}{2}\right) \\ & + \frac{\imath}{n^2\omega\pi}\sin\left(\frac{n\pi}{2}\right) \\ & + \frac{-\imath\lambda}{n\pi}\left(-2\cos(\pi n) + \cos\left(n\frac{\pi}{2}\right)\right) \\ & + \frac{-\imath}{n\pi}\sin\left(\frac{n\pi}{2}\right) \\ & + \frac{2\imath\lambda}{n^2\omega\pi}\left[\cos\left(\frac{n\pi}{2}\right) - \cos(n\pi)\right] \end{aligned}$$

On trie par puissance de n.

$$\begin{aligned} c_n(S) = & \tfrac{\imath}{n\pi}\left(2\lambda\cos(n\pi) - 2\lambda\cos\left(\tfrac{n\pi}{2}\right) - \sin\left(\tfrac{n\pi}{2}\right)\right) \\ & + \tfrac{\imath}{n^2\omega\pi}\left[2\lambda\cos\left(\tfrac{n\pi}{2}\right) - 2\lambda\cos(n\pi) + \sin\left(\tfrac{n\pi}{2}\right)\right] \end{aligned}$$

D'où enfin,

$$c_n(S) = \left[2\lambda\cos\left(\frac{n\pi}{2}\right) - 2\lambda\cos(n\pi) + \sin\left(\frac{n\pi}{2}\right)\right]\left(\frac{\imath}{n^2\omega\pi} - \frac{\imath}{n\pi}\right)$$

- Ainsi, pour revenir h', on a :

$$h'(t) = \sum_{j=1}^{n} \imath a_j \varphi(\lambda_j) e^{\imath\lambda_j t}$$

Alors, d'après le développement précédent,

$$h'(t) = \sum_{j=1}^{n} \imath a_j \left(\sum_{k\in\mathbb{Z}} c_k(S) e^{\imath k\omega\lambda_j}\right) e^{\imath\lambda_j t}$$

Puisque $\sum_{n\in\mathbb{Z}} c_k(S)$, la série de Fourier de S converge normalement, donc

$$h'(t) = \sum_{k\in\mathbb{Z}} \imath c_k(S) \sum_{j=1}^{n} a_j e^{\imath\lambda_j(t+\omega k)}$$

On pose alors $t_k = t + k\omega$, et $c_k = \imath c_k(S)$, pour avoir

$$h'(t) = \sum_{k\in\mathbb{Z}} c_k h(t + t_k)$$

- La conclusion est donc aisée :

$$|h'(t)| \leqslant \|h\|_\infty \sum_{k\in\mathbb{Z}} |c_k| = \lambda \|h\|_\infty$$

D'où l'inégalité de Bernstein :

$$\|h'\|_\infty \leqslant \lambda \|h\|_\infty$$

□

13 Théorème de Baire

Références :

1. *Analyse pour l'agrégation*, Hervé QUEFFÉLEC, Claude ZUILY
2. *Functionnal Analysis*, Walter RUDIN
3. *Cours d'analyse fonctionnelle*, Daniel LI

Théorème 13.1 : Théorème de BAIRE (1899)

Soit (E, d) un espace métrique complet.

(i) Si $(\Omega_n)_{n\in\mathbb{N}^*}$ est une suite d'ouvert denses dans E alors $\bigcap_{n\in\mathbb{N}^*} \Omega_n$ est dense dans E.

(ii) Si $(F_n)_{n\in\mathbb{N}^*}$ est une suite de fermés d'intérieur vide dans E alors $\bigcup_{n\in\mathbb{N}^*} F_n$ est d'intérieur vide dans E.

Démonstration : Soit $(\Omega_n)_{n\in\mathbb{N}^*}$ une suite d'ouverts denses dans E. Soit U un ouvert non vide de E. Montrons que

$$U \cap \left(\bigcap_{n\in\mathbb{N}^*} \Omega_n \right) \neq \varnothing$$

(1.) Montrons qu'il existe une suite $(x_n)_{n\in\mathbb{N}}$ de E, et une suite de réels strictement positifs $(r_n)_{n\in\mathbb{N}}$ tels que $B(x_0, r_0[\subset U$, et

$$\forall n \in \mathbb{N}, \bar{B}(x_{n+1}, r_{n+1}] \subset B(x_n, r_n) \cap \Omega_{n+1}$$

et

$$\forall n \in \mathbb{N}, 0 < r_{n+1} < \frac{r_n}{2}$$

• Puisque U est ouvert, pour $x_0 \in U$, il existe $r_0 > 0$ tel que $\bar{B}(x_0, r_0] \subset U$. De plus, Ω_1 est dense dans X, donc

$$\Omega_1 \cap B(x_0, r_0[\neq \varnothing$$

et est un ouvert de X. Il existe donc $x_1 \in X$ tel que $x_1 \in \Omega_1 \cap B(x_0, r_0[$, par densité de Ω_1. L'ensemble étant ouvert, il existe $r_1 > 0$ tel que

$$\bar{B}(x_1, r_1] \subset B(x_1, 2r_1[\subset \Omega_1 \cap B(x_0, r_0[$$

avec $0 < r_1 < \frac{r_0}{2}$. Ce qui conclut l'initialisation pour $n = 1$.

• Supposons construits $x_0, x_1, \cdots, x_n \in X$ et $r_0, \cdots, r_n$. Par densité de Ω_{n+1} dans X, $\Omega_{n+1} \cap B(x_n, r_n[$ est non vide, et est ouvert. Il existe donc $x_{n+1} \in B(x_n, r_n[\cap \Omega_n$ et $r_{n+1} > 0$ tels que

$$\bar{B}\,(x_{n+1}, r_{n+1}] \subset B(x_{n+1}, 2r_{n+1}[\subset \Omega_{n+1} \cap B(x_n, r_n[\subset \bigcap_{k=1}^{n+1} \Omega_k$$

et, quitte à réduire r_{n+1} :

$$0 < r_{n+1} < \frac{r_n}{2}$$

(2.) Par récurrence, il suit que

$$\forall n \in \mathbb{N}, 0 < r_n < \frac{r_0}{2^n}$$

donc $r_n \xrightarrow[n\to+\infty]{} 0$. Ainsi, puisque

$$d(x_{n+p}, x_n) \leqslant r_n$$

la suite $(x_n)_{n\in\mathbb{N}}$ est de Cauchy dans E complet. Elle converge alors vers un certain $x \in E$. Pour tout $p, n \in \mathbb{N}$,

$$x_{n+p} \in B(x_n, r_n[$$

donc

$$\forall n \in \mathbb{N}, x \in \bar{B}(x_n, r_n]$$

donc

$$x \in \bigcap_{n\in\mathbb{N}} \bar{B}(x_n, r_n] \subset U \cap \left(\bigcap_{n\in\mathbb{N}^*} \Omega_n \right)$$

donc pour tout U ouvert de E, $U \cap \left(\bigcap_{n\in\mathbb{N}^*} \Omega_n\right) \neq \varnothing$, donc on a bien montré que $\left(\bigcap_{n\in\mathbb{N}^*} \Omega_n\right)$ est dense dans E. □

14 Théorème de Banach-Steinhaus

Référence : Cours d'analyse fonctionnelle, Daniel Li

Théorème 14.1 : Théorème de Banach-Steinhaus (1927)

Soient E un espace de Banach et F un espace vectoriel normé. On considère pour tout $i \in I$, des applications

$$T_i : E \longrightarrow F$$

linéaires, continues, où I est un ensemble quelconque, telles que :

$$\forall x \in E, C_x \overset{\text{déf.}}{=} \sup_{i\in I} \|T_i x\|_F < +\infty$$

Alors

$$C \overset{\text{déf.}}{=} \sup_{i\in I} \|T_i\|_{\mathcal{L}(E,F)} < +\infty$$

Remarque : Ce théorème affirme que si

$$\forall x \in E, \exists C_x > 0, \forall i \in I, \|T_i x\|_F \leqslant C_x \|x\|_E$$

Alors :

$$\exists C > 0, \forall x \in E, \forall i \in I, \|T_i x\|_F \leqslant C \|x\|_E$$

Remarque 2 : La démonstration proposée ici, basée sur le théorème de BAIRE, n'est pas directement issue de BANACH et STEINHAUS qui l'avaient démontré autrement, mais issu d'une proposition de SAKS.

Démonstration : On note pour tout $n \in \mathbb{N}^*$:

$$\Phi_n \overset{\text{déf.}}{=} \{x \in E, \forall i \in I, \|T_i x\|_F \leqslant n\}$$

On peut le réécrire de la sorte :

$$\Phi_n \overset{\text{déf.}}{=} \bigcap_{i \in I} \{x \in E, \|T_i x\|_F \leqslant n\}$$

• Φ_n est un fermé de E (intersection quelconque de fermés, qui sont eux-même fermés car image réciproque d'un fermé par une application continue). De plus, par hypothèse,

$$\bigcup_{n \in \mathbb{N}} \Phi_n = E$$

donc n'est pas d'intérieur vide. Par la contraposée du théorème de Baire, il existe $N \in \mathbb{N}^*$ tel que Φ_N soit d'intérieur non vide.

• Pour ce $N \in \mathbb{N}^*$, on se donne $x_0 \in E$, et $r_0 > 0$ tels que

$$B_E(x_0, r_0[\subset \overset{\circ}{\Phi}_N \subset \Phi_N$$

Ainsi,

$$\forall x \in B_E(x_0, r_0[, \forall i \in I, \|T_i x\|_F \leqslant N$$

Soit $y \in B_E(0, 1[$. Alors $x_0 + r_0 y \in B_E(x_0, r_0[$. Par conséquent,

$$\forall i \in I, \|T_i(x_0 + r_0 y)\|_F \leqslant N$$

donc, par inégalité triangulaire inversée

$$\forall i \in I, r_0 \|T_i y\|_F - \|T_i x_0\|_F \leqslant N$$

donc, en isolant $\|T_i y\|_F$

$$\forall i \in I, \|T_i y\|_F \leqslant \frac{1}{r_0} \left[N + \|T_i x_0\|_F\right]$$

donc, pour tout $y \in B_E(0, 1)$:

$$\forall i \in I, \|T_i y\|_F \leqslant \frac{1}{r_0} \left[N + \sup_{i \in I} \|T_i x_0\|_F\right]$$

donc, par définition de la norme subordonnée

$$\forall i \in I, \|T_i\|_{\mathcal{L}(E,F)} \leqslant \frac{1}{r_0} \left[N + \sup_{i \in I} \|T_i x_0\|_F\right]$$

donc, puisque cela est pour tout $i \in I$

$$\sup_{i \in I} \|T_i\|_{\mathcal{L}(E,F)} \leqslant \frac{1}{r_0} \left[N + \sup_{i \in I} \|T_i x_0\|_F\right] < +\infty$$

□

Corollaire 14.1

Soient E un espace de BANACH, F un espace vectoriel normé, et $(T_n)_{n\in\mathbb{N}^*}$ une suite d'applications linéaires continues telles que

$$\forall x \in E, \lim_{n\to+\infty} T_n x \overset{\text{déf.}}{=} Tx$$

existe. Alors

(i) On a :

$$\sup_{n\in\mathbb{N}^*} \|T_n\|_{\mathcal{L}(E,F)} < +\infty$$

(ii) $T : E \longrightarrow F$ est une application linéaire continue qui vérifie :

$$\|T\|_{\mathcal{L}(E,F)} \leqslant \liminf_{n\to+\infty} \|T_n\|_{\mathcal{L}(E,F)}$$

Démonstration : **(i)** La convergence de $(T_n x)_{n\in\mathbb{N}}$ dans F, pour tout $x \in E$, donne lieu à l'existence de C_x :

$$\forall x \in E, C_x \overset{\text{déf.}}{=} \sup_{n\in\mathbb{N}^*} \|T_n x\|_F < +\infty$$

Le théorème de Banach-Steinhaus (la complétude de E autorise son utilisation) permet alors de donner

$$C \overset{\text{déf.}}{=} \sup_{n\in\mathbb{N}^*} \|T_n\|_{\mathcal{L}(E,F)} < +\infty$$

(ii) Par continuité de T_n, on a :

$$\forall n \in \mathbb{N}^*, \forall x \in E, \|T_n x\|_F \leqslant \|T_n\|_{\mathcal{L}(E,F)} \|x\|_E$$

Soit, d'après ce qui précède :

$$\forall n \in \mathbb{N}^*, \forall x \in E, \|T_n x\|_F \leqslant C\|x\|_E$$

Par continuité de la norme, il suit que

$$\forall x \in E, \|Tx\|_F \leqslant C\|x\|_E$$

donc T est continue et linéaire, par convergence simple sur E. Enfin, en passant à la limite inférieure dans :

$$\forall n \in \mathbb{N}^*, \forall x \in E, \|T_n x\|_F \leqslant \|T_n\|_{\mathcal{L}(E,F)} \|x\|_E$$

on obtient :

$$\forall x \in E, \lim_{n\to+\infty} \|T_n x\|_F \leqslant \left[\liminf_{n\to+\infty} \|T_n\|_{\mathcal{L}(E,F)}\right] \|x\|_E$$

Soit,

$$\forall x \in E, \|Tx\|_F \leqslant \left[\liminf_{n\to+\infty} \|T_n\|_{\mathcal{L}(E,F)}\right] \|x\|_E$$

Ce qui montre que

$$\|T\|_{\mathcal{L}(E,F)} \leqslant \liminf_{n\to+\infty} \|T_n\|_{\mathcal{L}(E,F)}$$

□

15 Théorème de l'application ouverte - Théorème d'isomorphisme de Banach - Théorème du graphe fermé

Référence : Cours d'analyse fonctionnelle, Daniel LI

Théorème 15.1 : Théorème de l'application ouverte (SCHAUDER 1930)

Soient E et F deux espaces de BANACH et $T : E \longrightarrow F$ une application linéaire continue surjective. Alors il existe $c > 0$ tel que

$$B_F(0, c) \subset T\left[B_E(0, 1)\right]$$

On dit que T est une *application ouverte*.

Remarque : T est une application ouverte si et seulement si pour tout ouvert U de E, $T(U)$ est un ouvert de F. En effet, si U est un ouvert de E, on considère $x_0 \in U$ et $r_0 > 0$ tel que $B(x_0, r_0[\subset U$. Alors

$$T(x_0) + r_0 T\left[B_E(0,1)\right] = T\left[B(x_0, r_0)\right] \subset T(U)$$

De plus, en considérant $c > 0$ tel que $B_F(0,c) \subset T\left[B_E(0,1)\right]$, on a :

$$B_F(Tx_0, r_0 c) = Tx_0 + r_0 B_F(0,c) \subset T(x_0) + r_0 T\left[B_E(0,1)\right] \subset T(U)$$

Il suit que pour tout $y \in T(U)$, il existe $r' > 0$ tel que $B_F(y, r') \subset T(U)$, donc $T(U)$ est ouvert. La réciproque est vraie, en considérant $U = B_E(0,1)$.

Référence supplémentaire : Cours ANAF de KPS, ANAF 2

Lemme 15.1

Soient E un espace vectoriel normé, F un espace de BANACH et $T : E \longrightarrow F$ une application linéaire surjective. Alors

$$\exists r > 0, B_F(0,r) \subset \overline{T\left[B_E(0,1)\right]}$$

Démonstration du lemme : • On considère la suite de fermés :

$$\forall n \in \mathbb{N}, F_n \overset{\text{déf.}}{=} \overline{T\left[B_E(0,n)\right]}$$

Par surjectivité de T, on a

$$F = \bigcup_{n \in \mathbb{N}} T[B_E(0,n)]$$

donc

$$F \subset \bigcup_{n \in \mathbb{N}} \overline{T[B_E(0,n)]} \subset F$$

Autrement dit,

$$F = \bigcup_{n \in \mathbb{N}} F_n$$

Par la contraposée du théorème de Baire, l'un de ces fermés est d'intérieur non vide, considérons qu'il s'agit d'un certain F_{n_0}.

• Ainsi, il existe $y_0 \in F_{n_0}$ et $r_0 > 0$ tels que

$$B_F(y_0, r_0) \subset F_{n_0} = n_0 \overline{T\left[B_E(0,1)\right]}$$

On veut désormais centrer cette boule en 0. Pour cela, constatons que

$$B_F(0,r) = -y_0 + B(y_0, r)$$

• Montrons que $-y_0 \in \overline{T\left[B_E(0,1)\right]}$. Si $x_0 \in B_E(0,1)$, alors $-x_0 \in B_E(0,1)$. Ainsi,

$$-T\left[B_E(0,1)\right] = T\left[-B_E(0,1)\right] = T\left[B_E(0,1)\right]$$

donc

$$y \in T\left[B_E(0,1)\right] \iff -y \in T\left[B_E(0,1)\right]$$

Par conséquent,

$$-\overline{T\left[B_E(0,1)\right]} = \overline{-T\left[B_E(0,1)\right]} = \overline{T\left[B_E(0,1)\right]}$$

donc $-y_0 \in \overline{T\left[B_E(0,1)\right]}$. Ainsi,

$$B_F(0, r_0) \subset \overline{T\left[B_E(0,1)\right]} + \overline{T\left[B_E(0,1)\right]}$$

• Montrons que $\overline{T\left[B_E(0,1)\right]} + \overline{T\left[B_E(0,1)\right]} = \overline{T\left[B_E(0,1)\right]}$ en montrant que $\overline{T\left[B_E(0,1)\right]}$ est convexe. Pour cela, on montre que $T\left[B_E(0,1)\right]$ est convexe, l'adhérence d'un convexe étant convexe. Soient $y_1, y_2 \in T\left[B_E(0,1)\right]$. Il existe $x_1, x_2 \in B_E(0,1)$ tels que $y_1 = Tx_1$ et $y_2 = Tx_2$. Soit $t \in [0,1]$. Alors

$$ty_1 + (1-t)y_2 = T(tx_1 + (1-t)x_2)$$

avec

$$\|tx_1 + (1-t)x_2\|_E \leqslant 1$$

donc $ty_1 + (1-t)y_2 \in T\left[B_E(0,1)\right]$, donc $\overline{T\left[B_E(0,1)\right]}$ est convexe.

• On peut alors enfin conclure :

$$B_F\left(0, \frac{r_0}{2}\right) \subset \frac{1}{2}\overline{T\left[B_E(0,1)\right]} + \frac{1}{2}\overline{T\left[B_E(0,1)\right]}$$

Donc, par convexité,

$$B_F\left(0, \frac{r_0}{2}\right) \subset \overline{T\left[B_E(0,1)\right]}$$

Ce qui conclut la démonstration de ce lemme. □

Référence : KPS + *Eléments d'analyse pour l'agrégation*, Zuily-Queffelec.

Lemme 15.2

Soient E un espace de Banach, F un esapce vectoriel normé, $T : E \longrightarrow F$ linéaire continue. On suppose qu'il existe $0 < \alpha < 1$ et $r_0 > 0$ tels que

$$\forall y \in B_F(0,1), \exists x \in B_E(0, r_0), \|y - Tx\|_F < \alpha$$

On dit que T est *presque surjective*. Alors

$$\forall y \in B_F(0,1), \exists x \in B_E\left(0, \frac{2r_0}{1-\alpha}\right), y = T(x)$$

Démonstration du lemme : • Construisons par récurrence une suite $(x_k)_{n\in\mathbb{N}^*} \in B(0, r_0)^{\mathbb{N}}$ telle que

$$\forall n \in \mathbb{N}^*, \left\| y - \sum_{k=1}^{n} \alpha^{k-1} T x_k \right\|_F < \alpha^n$$

▷ Pour $n = 1$, c'est la condition de presque surjectivité qui nous permet de se donner $x_1 \in B(0, r_0)$ tel que

$$\|y - Tx_1\|_F < \alpha$$

▷ Supposons construits $x_1, \cdots, x_n \in B(0, r_0)$ tels que

$$\left\| y - \sum_{k=1}^{n} \alpha^{k-1} T x_k \right\|_F < \alpha^n$$

Alors, en particulier,

$$\left\| \frac{y - \sum_{k=1}^{n} \alpha^{k-1} T x_k}{\alpha^n} \right\|_F < 1$$

D'après la condition de presque surjectivité, il existe $x_{n+1} \in B(0, r_0)$ telle que

$$\left\| \frac{y - \sum_{k=1}^{n} \alpha^{k-1} T x_k}{\alpha^n} - T x_{n+1} \right\|_F < \alpha$$

Donc,

$$\left\| y - \sum_{k=1}^{n+1} \alpha^{k-1} T x_k \right\|_F < \alpha^{n+1}$$

Ce qui conclut la récurrence.

• La série $\sum \alpha^{k-1} x_k$ est absolument convergente dans E. En effet :

$$\sum_{k=1}^{+\infty} \|\alpha_{k-1} x_k\|_E \leqslant \frac{r_0}{1-\alpha} < +\infty$$

Puisque E est complet, toute série absolument convergente est convergente, donc on peut considérer la somme de cette série :

$$x \overset{\text{déf.}}{=} \sum_{k=1}^{+\infty} \alpha^{k-1} x_k$$

• On vient de montrer que

$$x \in \overline{B_E}\left(0, \frac{r_0}{1-\alpha}\right) \subset B_E\left(0, \frac{2r_0}{1-\alpha}\right)$$

Montrons que $y = T(x)$. Pour cela, il suffit de voir que pour tout $n \in \mathbb{N}$:

$$\|y - Tx\|_F \leqslant \left\| y - \sum_{k=1}^{n} \alpha^{k-1} x_k \right\|_F + \left\| \sum_{k=n+1}^{+\infty} \alpha^{k-1} x_k \right\|_F$$

Donc,

$$\|y - Tx\|_F \leqslant \alpha^n + \sum_{k=n+1}^{+\infty} \left\|\alpha^{k-1} x_k\right\|_F$$

Ce sont deux termes qui tendent vers 0 lorsque $[n \to +\infty]$, donc on a bien $y = Tx$. □

Remarque : Ainsi, si E est un espace de Banach et F est un espace vectoriel normé, si T est presque surjective, elle est en fait surjective.

Démonstration : D'après le lemme **15.1**, il existe $r > 0$ tel que

$$B_F(0,r) \subset \overline{T\left[B_E(0,1)\right]}$$

Montrons que T est presque surjective. Pour cela, on observe que

$$B_F(0,1) = \frac{1}{r} B_F(0,r)$$

donc

$$B_F(0,1) \subset \frac{1}{r}\overline{T\left[B_E(0,1)\right]}$$

Soit encore

$$B_F(0,1) \subset \overline{T\left[B_E\left(0,\frac{1}{r}\right)\right]}$$

Ainsi, pour tout $y \in B_F(0,1)$, il existe $x \in B_E\left(0,\frac{1}{r}\right)$ et $\alpha \in]0,1[$ tels que

$$\|y - Tx\|_F < \alpha$$

T est donc bien presque surjective. D'après le lemme **15.2**, pour tout $y \in B_F(0,1)$, il existe $x \in B_E\left(0,\frac{2}{r(1-\alpha)}\right)$ tel que

$$y = Tx$$

donc, on a montré que

$$B_F(0,1) \subset T\left[B_E\left(0,\frac{2}{r(1-\alpha)}\right)\right]$$

Par linéarité, on a bien montré que

$$B_F\left(0,\frac{r(1-\alpha)}{2}\right) \subset T\left[B_E(0,1)\right]$$

Ce qui montre que T est ouverte. □

Corollaire 15.1 : Théorème d'isomorphisme de BANACH

Soient E et F deux espaces de BANACH, et $f : E \longrightarrow F$ une application linéaire continue bijective. Alors $f^{-1} : F \longrightarrow E$ est continue.

Démonstration : On applique le théorème de l'application ouverte à T, application linéaire continue surjective : il existe $c > 0$ tel que

$$B_F(0,c) \subset T\left[B_E(0,1)\right]$$

Puisque T est bijective, il suit que

$$T^{-1}\left[B_F(0,c)\right] \subset T^{-1} \circ T\left[B_E(0,1)\right] = B_E(0,1)$$

Ainsi, pour tout $y \in F$ tel que $\|y\| < c$:

$$\left\|T^{-1}y\right\|_E \leqslant 1$$

Pour tout $y \in F$ non nul, $\frac{cy}{2\|y\|_F} \in B_F(0,c)$, donc :

$$\left\|T^{-1}\frac{cy}{2\|y\|_F}\right\|_E \leqslant 1$$

Soit, pour tout $y \in F$ non nul,

$$\left\|T^{-1}1\right\|_E \leqslant \frac{2}{c}\|y\|_F$$

L'inégalité reste vraie pour $y = 0$. Cela montre la continuité de T^{-1}. □

Corollaire 15.2

Soit E un espace vectoriel normé muni de deux normes $\|\cdot\|$ et N telles que E soit de BANACH pour chacune de ces deux normes. On suppose que

$$\exists C > 0, \forall x \in E, N(x) \leqslant C\|x\|$$

Alors ces deux normes sont en fait équivalentes.

Démonstration : L'application

$$\mathrm{id}_E : (E, \|x\|) \longrightarrow (E, N)$$

est une application linéaire bijective. De plus, l'hypothèse

$$\exists C > 0, \forall x \in E, N(x) \leqslant C\|x\|$$

rend cette application continue, sachant que E muni des deux normes en question est un espace de Banach. Le théorème d'isomorphisme de Banach permet alors de conclure que l'application inverse est continue : autrement dit, il existe $c > 0$ tel que

$$\forall x \in E, \|x\| \leqslant cN(x) \leqslant cC\|x\|$$

Ces normes sont donc bien équivalentes. □

Référence supplémentaire : Cours d'ANAF de KPS, ANAF 2

Corollaire 15.3 : Théorème du graphe fermé (BANACH 1929)

Soient E et F deux espaces de BANACH, et $T : E \longrightarrow F$ linéaire. On définit le *graphe* de T :

$$G(T) \overset{\text{déf.}}{=} \{(x, Tx) \in E \times F, x \in E\}$$

Les deux propositions suivantes sont équivalentes :

(i) T est une application linéaire continue.

(ii) $G(T)$ est fermé dans $E \times F$.

Démonstration : [(i) $\Longrightarrow$ (ii)] Si T est une application linéaire continue alors pour toute suite $(x_n)_n \in E^{\mathbb{N}}$ qui converge vers $x \in E$, la suite $(Tx_n)_n \in F^{\mathbb{N}}$ converge vers $Tx \in F$. Ainsi, si on considère $\{(x_n, Tx_n)\}_{n\in\mathbb{N}} \in G(T)^{\mathbb{N}}$ qui converge vers $(x, y) \in E \times F$, l'unicité de la limite dans F donne $Tx = y$, donc $(x, y) = (x, Tx) \in G(T)$, donc le graphe de G est bien fermé.

[(ii) $\Longrightarrow$ (i)] Supposons que $G(T)$ soit fermé dans $E \times F$. Alors $G(T)$ est un sous-espace vectoriel fermé du complet $E \times F$, pour la norme produit. On définit sur E la norme suivante :

$$\forall x \in E, N_E(x) \overset{\text{déf.}}{=} \|x\|_E + \|Tx\|_F$$

Montrons que (E, N_E) est complet. Soit $(x_n)_{n\in\mathbb{N}}$ une suite de Cauchy dans E pour la norme N_E : pour tout $\varepsilon > 0$, il existe $N_0 \in \mathbb{N}$ tel que

$$\forall n \geqslant N_0, \forall p \in \mathbb{N}, \|x_{n+p} - x_n\|_E + \|Tx_{x+p} - Tx_n\|_F \leqslant \varepsilon$$

Par conséquent, $(x_n)_n$ est de Cauchy dans $(E, \|\cdot\|_E)$ complet, donc converge vers un certain $x \in E$. De même, $(Tx_n)_{n\in\mathbb{N}}$ est de Cauchy dans $(F, \|\cdot\|_F)$ complet, donc converge vers un certain $y \in F$. Ainsi,

$$(x_n, Tx_n) \xrightarrow[n\to+\infty]{\|\cdot\|_{E\times F}} (x, y)$$

où

$$\|(x, y)\|_{E\times F} \overset{\text{déf.}}{=} \|x\|_E + \|y\|_F$$

est la norme produit. Or, G est supposé fermé dans $\left(E \times F, \|\cdot\|_{E\times F}\right)$, donc $(x, y) \in G(T)$, donc $y = Tx$. Ainsi défini,

$$N_E(x_n - x) = \|x_n - x\|_E + \|Tx_n - Tx\|_F$$

Soit,

$$N_E(x_n - x) = \|x_n - x\|_E + \|Tx_n - y\|_F$$

donc,

$$N_E(x_n - x) \xrightarrow[n\to+\infty]{} 0$$

donc $(x_n)_n$ converge vers x dans (E, N_E), ce qui en fait un espace de Banach. De plus,

$$\forall x \in E, \|x\|_E \leqslant N_E(x)$$

avec (E, N_E) et $(E, \|\cdot\|_E)$ deux espaces de Banach. Par le corollaire **15.2**, ces deux normes sont équivalentes : il existe $M > 0$ tel que

$$\forall x \in E, N_E(x) \leqslant M\|x\|_E$$

D'où finalement,

$$\forall x \in E, \|T(x)\|_F \leqslant N_E(x) \leqslant M\|x\|_E$$

donc T est continue. □

16 Théorème de prolongement de Tietze

Référence : Analyse pour l'agrégation, Hervé QUEFFÉLEC, Claude ZUILY

Théorème 16.1 : Théorème de prolongement de TIETZE

Soit (X, d) un espace métrique. Soit Y un fermé de X. Si $g_0 : Y \longrightarrow \mathbb{R}$ est une application continue, alors il existe $f_0 : X \longrightarrow \mathbb{R}$ continue qui prolonge g_0.

Ce théorème est une conséquence du lemme **15.2**, que l'on rappelle selon les termes exacts de l'énoncé de ZUILY-QUEFFELEC.

Lemme 16.1

Soient E un espace de BANACH, F un esapce vectoriel normé, $T : E \longrightarrow F$ linéaire continue. On suppose qu'il existe $0 < \alpha < 1$ et $r_0 > 0$ tels que pour tout $y \in F$ vérifiant $\|y\|_F \leqslant 1$,

$$\exists x \in E, \|x\|_E \leqslant r_0, \|y - Tx\|_F \leqslant \alpha$$

On dit que T est *presque surjective*. Alors pour tout $y \in F$, avec $\|y\|_F \leqslant 1$,

$$\exists x \in E, \|x\|_E \leqslant \frac{r_0}{1-\alpha}, y = T(x)$$

Démonstration : 1. Sur $\mathcal{C}_b^0(X, \mathbb{R})$, l'espace des fonctions continues bornées sur Y, muni de la norme uniforme :

$$\forall f \in \mathcal{C}_b^0(X, \mathbb{R}), \|f\|_\infty \overset{\text{déf.}}{=} \sup_{x \in X} |f(x)|$$

et sur $\mathcal{C}_b^0(Y, \mathbb{R})$, muni de la norme induite, on définit

$$T : \begin{pmatrix} \mathcal{C}^0(X, \mathbb{R}) & \longrightarrow & \mathcal{C}^0(Y, \mathbb{R}) \\ f & \longmapsto & f_{|Y} \end{pmatrix}$$

T est une application linéaire continue (Y est fermé). Montrons que T est presque surjective : il existe $0 < \alpha < 1$ et $r_0 > 0$ tels que pour tout $g \in \mathcal{C}^0(Y, \mathbb{R})$ vérifiant $\|g\|_\infty \leqslant 1$, il existe $f \in \mathcal{C}^0(X, \mathbb{R})$ tel que $\|f\|_\infty \leqslant r_0$ et

$$\|g - f_{|Y}\|_\infty \leqslant \alpha$$

On va montrer que c'est le cas pour $\alpha = \frac{2}{3}$ et $r_0 = \frac{1}{3}$. Soit $g \in \mathcal{C}^0(Y, \mathbb{R})$ telle que $\|g\|_\infty \leqslant 1$. On considère les deux ensembles suivants :

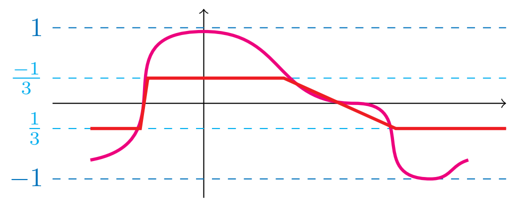
$$Y^+ \overset{\text{déf.}}{=} \left\{ x \in Y \,\middle|\, \frac{1}{3} \leqslant g(x) \leqslant 1 \right\}$$

et

$$Y^- \overset{\text{déf.}}{=} \left\{ x \in Y \,\middle|\, -1 \leqslant g(x) \leqslant \frac{-1}{3} \right\}$$

On note enfin

$$\forall x \in X, f(x) \overset{\text{déf.}}{=} \frac{1}{3} \frac{d(x, Y^-) - d(x, Y^+)}{d(x, Y^-) + d(x, Y^+)}$$



Alors $f \in \mathcal{C}_b^0(X, \mathbb{R})$ avec $\|f\|_\infty \leqslant \frac{1}{3}$. Montrons que $\|g - Tf\|_\infty \leqslant \frac{2}{3}$. On considère trois situations différentes :

- Si $x \in Y^+$, alors $f(x) = \frac{1}{3}$, donc

$$g(x) - f(x) = g(x) - \frac{1}{3} \in \left[0, \frac{2}{3}\right]$$

- Si $x \in Y^-$, alors $f(x) = \frac{-1}{3}$, donc

$$g(x) - f(x) = g(x) + \frac{1}{3} \in \left[\frac{-2}{3}, 0\right]$$

- Si $x \in Y \backslash (Y^+ \cup Y^-)$, alors

$$|g(x) - f(x)| \leqslant \frac{1}{3} + \frac{1}{3} \leqslant \frac{2}{3}$$

Dans tous les cas, on a alors

$$\|g - Tf\|_\infty \leqslant \frac{2}{3}$$

avec $\|g\|_\infty \leqslant 1$ et $\|f\|_\infty \frac{1}{3}$. Ainsi, T est bien presque surjective. Or, $\mathcal{C}^0(X, \mathbb{R})$ est un espace de Banach. D'après le lemme **16.1**, T est alors surjective, et vérifie de plus que si $g \in \mathcal{C}^0(Y, \mathbb{R})$ est tel que $\|g\|_\infty \leqslant 1$ alors il existe $f \in \mathcal{C}^0(X, \mathbb{R})$ telle que

$$\|f\|_\infty \leqslant \frac{\frac{1}{3}}{1 - \frac{2}{3}} = 1$$

et $g = f_{|Y}$.

2. Supposons que $g \in \mathcal{C}^0(Y,\mathbb{R})$ soit telle que $\|g\|_\infty < 1$. Montrons qu'il existe $f \in \mathcal{C}^0(X,\mathbb{R})$ tel que $\|f\|_\infty < 1$ et f prolonge g. Par le point précédent, il existe $h \in \mathcal{C}^0(X,\mathbb{R})$ tel que $\|h\|_\infty \leqslant 1$ et $g = h_{|Y}$. On considère

$$Z \overset{\text{déf.}}{=} \{x \in X, |h(x)| = 1\}$$

On définit

$$\forall x \in X, u(x) \overset{\text{déf.}}{=} \frac{d(x,Z)}{d(x,Z)+d(x,Y)}$$

et $f = hu$. Montrons que f est continue, prolonge g, et $\|f\|_\infty < 1$. u est bien définie, car Y et Z sont fermés et $Y \cap Z = \varnothing$.

u est la composées d'applications 1-lipschitzienne, donc u est continue. C'est donc alors aussi le cas de f. De plus, $u = 1$ sur Y, donc $f = h = g$ sur Y, donc f prolonge g. Enfin, en dehors de Y, $u < 1$, donc $|f| < |h| \leqslant 1$ en dehors de Y, et sur $|f| = |h| = |g| < 1$ sur Y. f répond donc bien à la question.

3. Soit

$$\varphi : \mathbb{R} \longrightarrow]-1,1[$$

un homéomorphisme. On considère

$$g \overset{\text{déf.}}{=} \varphi \circ g_0 \in \mathcal{C}^0(Y,\mathbb{R})$$

D'après **2.**, il existe $f \in \mathcal{C}^0(X,\mathbb{R})$ tel que $g = f_{|Y}$ et $\|f\|_\infty < 1$. On note alors

$$f_0 \overset{\text{déf.}}{=} \varphi^{-1} \circ f \in \mathcal{C}^0(X,\mathbb{R})$$

f_0 est continue sur X, et prolonge g_0. □

17 Méthode de Laplace

Référence : Analyse pour l'agrégation, Hervé QUEFFÉLEC, Claude ZUILY

Théorème 17.1 : Méthode de LAPLACE

Soit $I =]a,b[$ un intervalle de $\mathbb{R}$ (avec $a \in \{-\infty\} \cup \mathbb{R}$ et $b \in \mathbb{R} \cup \{+\infty\}$). Soient $\varphi \in \mathcal{C}^2(I)$ et $f \in \mathcal{C}^0(I,\mathbb{C})$. On définit

$$\forall t \in I, F(t) \overset{\text{déf.}}{=} \int_a^b e^{t\varphi(x)} f(x)\, \mathrm{d}x$$

On suppose :

(i) Pour tout $t > 0$,

$$\int_a^b e^{t\varphi(x)} |f(x)|\, \mathrm{d}x < +\infty$$

(ii) Il existe un unique point $x_0 \in I$ tel que $\varphi'(x_0) = 0$. On suppose de plus que ce point vérifie $\varphi''(x_0) < 0$, donc que x_0 est un unique maximum global de φ.

(iii) $f(x_0) \neq 0$

Alors on dipose de l'équivalent suivant lorsque $[t \to +\infty]$:

$$F(t) \sim \sqrt{\frac{2\pi}{t}} \frac{f(x_0)}{\sqrt{|\varphi''(x_0)|}} e^{t\varphi(x_0)}$$

Idée : On donne ici l'idée globale sous forme grossière, la démonstration établira tous les points techniques. On mettra en couleur rouge tous les raisonnements et autre passage à la limite hasardeux.

Déjà, on se donne θ une fonction lise telle que l'intégrale se passe sur $\mathbb{R}$. φ est C^2, donc on peut la développer jusqu'à l'ordre 2 en x_0 :

$$\varphi(x) = \varphi(x_0) + (x - x_0)^2\psi(x)$$

avec ψ de classe C^1 sur I. On reporte dans l'intégrale :

$$F(t) = \int_{\mathbb{R}} \theta(x)e^{t\varphi(x_0)}e^{t(x-x_0)^2\psi(x)}f(x)\ \mathrm{d}x$$

On fait le changement de variables $z^2 = (x - x_0)^2\psi(x)$. En particulier, on note $x = g(z)$.

$$F(t) = e^{t\varphi(x_0)} \int_{\mathbb{R}} \theta(g(z))e^{tz^2}f(g(z))g'(z)\ \mathrm{d}z = e^{t\varphi(x_0)} \int_{\mathbb{R}} e^{tz^2}h(z)\ \mathrm{d}z$$

On fait le changement de variables $y = \sqrt{t}z$:

$$F(t) = \frac{e^{t\varphi(x_0)}}{\sqrt{t}} \int_{\mathbb{R}} h\left(\frac{y}{\sqrt{t}}\right) e^{-y^2}\ \mathrm{d}y$$

On prend l'équivalent de cette intégrale :

$$F(t) \underset{t\to+\infty}{\sim} \frac{h(0)e^{t\varphi(x_0)}}{\sqrt{t}} \int_{\mathbb{R}} e^{-y^2}\ \mathrm{d}y = \frac{\sqrt{\pi}h(0)e^{t\varphi(x_0)}}{\sqrt{t}}$$

Ce qui est le résultat escompté.

Démonstration : • D'après la formule de Taylor avec reste intégral, pour tout $x \in I$,

$$\varphi(x) = \varphi(x_0) + (x - x_0)^2\psi(x)$$

ψ est continue sur I, et C^1 sur $I\backslash\{x_0\}$. De plus,

$$\psi(x_0) = \frac{\varphi''(x_0)}{2} < 0$$

Par continuité de ψ, il existe $\delta_1 > 0$ tel que $]x_0-\delta, x_0+\delta[\subset I$ et

$$\forall x \in]x_0 - \delta_1, x_0 + \delta_1[, \psi(x) < 0$$

• Afin d'effectuer le premier changement de variables, assurons nous qu'on utilise bien un C^1-difféomorphisme sur un certain intervalle. Définissons sur $]x_0-\delta, x_0+\delta[$ la fonction u définie par :

$$\forall x \in]x_0 - \delta, x_0 + \delta[, u(x) \overset{\text{déf.}}{=} (x - x_0)\sqrt{-\psi(x)}$$

Sur $]x_0 - \delta, x_0 + \delta[\backslash\{x_0\}$, u est dérivable, avec

$$u'(x) = \sqrt{-\psi(x)} - \frac{\psi'(x)}{2\sqrt{-\psi(x)}}(x - x_0)$$

On note alors

$$u'(x_0) = \sqrt{-\psi(x_0)} = \sqrt{\frac{-\varphi''(x_0)}{2}} > 0$$

Montrons que u' est continue au voisinage de x_0. Pour cela, on constate que pour $x \neq x_0$:

$$\psi(x) = \frac{\varphi(x) - \varphi(x_0)}{x - x_0}$$

D'où

$$(x-x_0)\psi'(x) = \frac{\varphi'(x) - \varphi'(x_0)}{x - x_0} - 2\psi(x) \xrightarrow[x\to x_0]{} \varphi''(x_0) - 2\psi(x_0) = 0$$

d'où

$$\sqrt{-\psi(x)} - \frac{\psi'(x)}{2\sqrt{-\psi(x)}}(x - x_0) \xrightarrow[x\to x_0]{} \sqrt{-\psi(x_0)}$$

Autrement dit,

$$u'(x) \xrightarrow[\substack{x\to x_0\\ x\neq x_0}]{} u'(x_0)$$

donc u est bien prolongeable en une fonction de classe C^1 sur $]x_0 - \delta_1, x_0 + \delta_1[$. Puisque $u'(x_0) > 0$, et que u' est continue, il existe $\delta > 0$ tel que u' ne s'annule pas sur $]x_0 - \delta, x_0 + \delta[$. Sur cet intervalle, u est donc strictement monotone, donc réalise un C^1-difféomorphisme sur son image.

• Considérons $\theta \in \mathcal{C}_0^\infty(\mathbb{R})$ telle que θ soit nulle sur $\mathbb{R}\backslash]x_0 - \delta, x_0 + \delta[$ et θ vaille 1 sur $\left]x_0 - \frac{\delta}{2}, x_0 + \frac{\delta}{2}\right[$. Alors u est un $\mathcal{C}^1$-difféomorphisme sur le support de θ. Réécrivons l'intégrale que l'on souhaite travailler (signalons que (i) permet d'affirmer que F est bien définie) :

$$F(t) = \int_a^b \theta(x)e^{t\varphi(x)}f(x)\ \mathrm{d}x + \int_a^b (1-\theta(x))e^{t\varphi(x)}f(x)\ \mathrm{d}x$$

On note $F_1(t)$ la première intégrale et $F_2(t)$ le deuxième.

• Étudions d'abord $F_1(t)$. Constatons que pusique θ est à support compact, l'intégrale peut se prendre sur $\mathbb{R}$. De plus, u réalise un difféomorphisme sur le support de θ. On note g l'application inverse de u sur $]x_0-\delta, x_0+\delta[$. On a :

$$F_1(t) = \int_{\mathbb{R}} \theta(x)e^{t\varphi(x_0)}e^{t(x-x_0)^2\psi(x)}f(x)\ \mathrm{d}x$$

On fait alors le changement de variables $z = u(x)$, équivalent à $x = g(z)$, possible grâce à notre travail précédent sur I et le choix de θ :

$$F_1(t) = e^{t\varphi(x_0)}\int_{\mathbb{R}} \theta(g(z))e^{tz^2}f(g(z))\,g'(z)\ \mathrm{d}z$$

On note $h(z) \overset{\text{déf.}}{=} \theta(g(z))f(g(z))g'(z)$. Alors

$$h(0) = \theta(x_0)f(x_0)g'(0)$$

Par théorème de dérivation de fonctions composées,

$$g'(0) = \frac{1}{u'(g(0))} = \frac{1}{u'(x_0)} = \sqrt{\frac{-2}{\varphi''(x_0)}}$$

d'où

$$h(0) = f(x_0)\sqrt{\frac{2}{-\varphi''(x_0)}}$$

A ce stade, on a :

$$F_1(t) = e^{t\varphi(x_0)}\int_{\mathbb{R}} h(z)e^{tz^2}\ \mathrm{d}z$$

On fait le changement de variables $y = \sqrt{t}z$ pour avoir :

$$F_1(t) = \frac{e^{t\varphi(x_0)}}{\sqrt{t}}\int_{\mathbb{R}} h\left(\frac{y}{\sqrt{t}}\right)e^{y^2}\ \mathrm{d}y$$

h étant à support compact, le théorème de convergence dominée permet alors d'avoir

$$\int_{\mathbb{R}} h\left(\frac{y}{\sqrt{t}}\right)e^{y^2}\ \mathrm{d}y \xrightarrow[t\to+\infty]{} h(0)\int_{\mathbb{R}} e^{-y^2}\ \mathrm{d}y = h(0)\sqrt{\pi}$$

D'où

$$F_1(t) \underset{t\to+\infty}{\sim} f(x_0)\sqrt{\frac{2\pi}{|\varphi''(x_0)|}}\frac{e^{t\varphi(x_0)}}{\sqrt{t}}$$

Ce qui est le résultat voulu pour F. (Cet équivalent est non nul *via* l'hypothèse (iii)).

• Montrons que $F_2(t) = o(F_1(t))$ en $+\infty$. Il s'agit de montrer que

$$\left|\frac{F_2(t)}{\sqrt{t}\kappa_0 e^{t\varphi(x_0)}}\right| \xrightarrow[t\to+\infty]{} 0$$

avec κ_0 la constante (par rapport à t) dans l'équivalent de $F_1(t)$. Montrons qu'il existe $\mu > 0$ tel que

$$\forall x \in I\backslash\left[x_0-\frac{\delta}{2}, x_0+\frac{\delta}{2}\right], \varphi(x_0)-\varphi(x) \geqslant \mu > 0$$

Notons que $I\backslash\left[x_0-\frac{\delta}{2}, x_0+\frac{\delta}{2}\right]$ est l'endroit où $1-\theta$ vaut 1, donc là où l'intégrale qui définit F_2 est non nulle. Si $x < x_0-\frac{\delta}{2}$, $\varphi'(x) > 0$, puisque φ est strictement croissante sur $I\cap]-\infty, x_0[$. Ainsi,

$$\varphi(x_0)-\varphi(x) = \varphi(x_0)-\varphi\left(x_0-\tfrac{\delta}{2}\right) + \varphi\left(x_0-\tfrac{\delta}{2}\right)-\varphi(x)$$

Donc, par croissance de φ,

$$\varphi(x_0)-\varphi(x) \geqslant \varphi(x_0)-\varphi\left(x_0-\frac{\delta}{2}\right)$$

qui est une minoration indépendante de x. Pour $x > x_0+\frac{\delta}{2}$, on a cette fois

$$\varphi(x_0)-\varphi(x) \geqslant \varphi(x_0)-\varphi\left(x_0+\frac{\delta}{2}\right)$$

Ce qui montre bien qu'il existe $\mu > 0$ tel que

$$\forall x \in I\backslash\left[x_0-\frac{\delta}{2}, x_0+\frac{\delta}{2}\right], \varphi(x_0)-\varphi(x) \geqslant \mu > 0$$

Pour $t > 1$, on fait alors la majoration astucieuse suivante :

$$t\varphi(x) = \varphi(x)+(t-1)\varphi(x)$$

donc

$$t\varphi(x) \leqslant \varphi(x)+(t-1)(\varphi(x_0)-\mu)$$

Ainsi, on peut enfin majorer $F_2(t)$:

$$|F_2(t)| = \left|\int_a^b (1-\theta(x))e^{t\varphi(x)}f(x)\ \mathrm{d}x\right|$$

Par inégalité triangulaire, et par notre petite travail précédent :

$$|F_2(t)| \leqslant e^{t\varphi(x_0)}e^{-\varphi(x_0)}e^{\mu(1-t)}\int_a^b (1-\theta(x))e^{\varphi(x)}\,|f(x)|\ \mathrm{d}x$$

Autrement dit, par l'hypothèse (i), on a alors en $+\infty$

$$F_2(t) = O\left(e^{-t\mu}e^{t\varphi(x_0)}\right)$$

donc

$$\left|\frac{F_2(t)\sqrt{t}}{\kappa_0 e^{t\varphi(x_0)}}\right| = O\left(e^{-t\mu}\sqrt{t}\right) \xrightarrow[t\to+\infty]{} 0$$

donc $F_2(t) = o\left(F_1(t)\right)$.

• On peut alors enfin conclure, puisque $F = F_1 + o(F_1)$, il suit bien que

$$F(t) \underset{t\to+\infty}{\sim} f(x_0)\sqrt{\frac{2\pi}{-\varphi''(x_0)}}\frac{e^{t\varphi(x_0)}}{\sqrt{t}}$$

□

Corollaire 17.1 : Formule de Stirling

$$n! \sim \sqrt{2\pi n}\left(\frac{n}{e}\right)^n$$

Démonstration : On cherche à utiliser la méthode de Laplace pour la fonction Γ

$$\forall z > 0, \Gamma(z) \overset{\text{déf.}}{=} \int_0^{+\infty} t^{z-1}e^{-t}\ \mathrm{d}t$$

On exprime pour cela t^{z-1} sous forme exponentielle :

$$\Gamma(z+1) = \int_0^{+\infty} e^{-t+z\ln(t)}\ \mathrm{d}t$$

On fait le changement de variables $t = zs$:

$$\Gamma(z+1) = ze^{z\ln(z)}\int_0^{+\infty} e^{-sz+z\ln(s)}\ \mathrm{d}s$$

On se retrouve alors ici sous les hypothèses de la méthode de Laplace avec $f(z) = 1$ et $\varphi(s) = \ln(s) - s$. En effet,

$$\varphi'(s) = \frac{1}{s} - 1$$

et s'annule seulement en 1, où $\varphi''(s) = -1 < 0$. Par conséquent,

$$\Gamma(z+1) \underset{z\to+\infty}{\sim} zz^z\sqrt{2\pi}\frac{e^{-z}}{\sqrt{z}}$$

Soit encore

$$\Gamma(z+1) \underset{z\to+\infty}{\sim} \sqrt{2\pi z}\left(\frac{z}{e}\right)^z$$

D'où en particulier la formule de Stirling. □

18 Théorème de Riesz-Fischer

Référence : Haïm Brézis, *Analyse fonctionnelle*
Juste pour faire cool, on va l'écrire avec des notations probabilistes.

Théorème 18.1 : de Riesz-Fischer

Soit $p \in [1, +\infty[\cup\{+\infty\}$ et $(\Omega, \mathcal{F}, \mu)$ un espace mesuré.

1. L'espace $L^p_\mu(\Omega)$ muni de sa norme $\|\cdot\|_{L^p}$ définie par :

$$\forall X \in L^p, \|X\|^p_{L^p} \overset{\text{déf.}}{=} \int_\Omega |X(\omega)|^p\ \mathrm{d}\mu(\omega)$$

pour $p < +\infty$ et par

$$\|X\|_{L^\infty} \overset{\text{déf.}}{=} \inf\left\{c > 0, |X| \leqslant c \quad \mu - \text{ps}\right\}$$

est complet ;

2. Si $(X_n)_n$ est une suite de fonctions mesurables qui converge vers une fonction mesurable X au sens L^p, alors il existe une extraction ϕ telle que la sous-suite $\left(X_{\phi(n)}\right)_n$ converge μ-presque sûrement vers X.

On suppose déjà acquis le caractère normé de L^p.

Démonstration : On constatera que **2.** est une conséquence de la démonstration de **1.**, car une suite convergente est avant tout une suite de Cauchy.

[(i)] Traitons d'abord le cas où $p = +\infty$. Soit $(X_n)_n$ une suite de Cauchy dans $L^\infty_\mu(\Omega)$. Alors

$$\forall k \in \mathbb{N}^*, \exists \phi(k) \in \mathbb{N}, \begin{matrix} \forall n \geqslant \phi(k) \\ \forall p \in \mathbb{N} \end{matrix}, \|X_{n+p} - X_n\|_{L^\infty} \leqslant \frac{1}{k}$$

Pour k, n, p fixés comme au-dessus, cela signifie qu'il existe $N_{k,n,p}$ négligeable tel que

$$\forall \omega \in \Omega \backslash N_{k,n,p}, |X_{n+p}(\omega) - X_n(\omega)| \leqslant \frac{1}{k}$$

On considère alors

$$N \overset{\text{déf.}}{=} \bigcup_{k \in \mathbb{N}^*} \bigcup_{n \geqslant \phi(k)} \bigcup_{p \in \mathbb{N}} N_{k,n,p}$$

Alors puisque N est union dénombrable de négligeables, N est lui-même négligeable. De plus, cela signifie que l'on peut remonter le "$\forall \omega \in \Omega \backslash N$" en tête de liste pour avoir : pour tout $\omega \in \Omega \backslash N$,

$$\forall k \in \mathbb{N}^*, \begin{matrix} \forall n \geqslant \phi(k) \\ \forall p \in \mathbb{N} \end{matrix}, |X_{n+p}(\omega) - X_n(\omega)| \leqslant \frac{1}{k}$$

Ce qui signifie que pour tout $\omega \in \Omega \backslash N$, la suite réelle $(X_n(\omega))_n$ est une suite de Cauchy. On dit aussi que $(X_n)_n$ est presque sûrement de Cauchy. Puisque $\mathbb{R}$ est complet, il suit que X_n converge presque sûrement vers une fonction mesurable X. Montrons que $X \in L^\infty$ et que $(X_n)_n$ converge vers X dans L^∞. Pour cela, on intercale le "$\forall \omega \in \Omega \backslash N$" entre le "$\forall k$" et le "$\forall n, \forall p$" : pour tout $k \in \mathbb{N}^*$:

$$\forall \omega \in \Omega \backslash N, \begin{matrix} \forall n \geqslant \phi(k) \\ \forall p \in \mathbb{N} \end{matrix}, |X_{n+p}(\omega) - X_n(\omega)| \leqslant \frac{1}{k}$$

Un passage à la limite lorsque $[p \to +\infty]$ donne alors, pour $k \in \mathbb{N}^*$:

$$\forall \omega \in \Omega \backslash N, \forall n \geqslant \phi(k), |X(\omega) - X_n(\omega)| \leqslant \frac{1}{k}$$

Une interversion entre les "$\forall \omega$" et "$\forall n$" donne alors

$$\forall k \in \mathbb{N}^*, \forall n \geqslant \phi(k), \|X - X_n\|_{L^\infty} \leqslant \frac{1}{k}$$

Si on se fixe k et n, cela donne alors

$$\|X\|_{L^\infty} \leqslant \|X_n\|_{L^{+\infty}} + \|X - X_n\|_{L^\infty} < \infty$$

donc $X \in L^\infty$, et on a de plus la convergence en norme L^∞ de $(X_n)_n$ vers X. Le point **2.** du théorème est ici donné par le fait que la convergence uniforme implique la convergence simple.

[(ii)] Pour $p < \infty$. Soit $(X_n)_n$ une suite de Cauchy. Alors, il existe une extraction ϕ telle que

$$\forall k \in \mathbb{N}^*, \forall q \in \mathbb{N}, \left\|X_{\phi(k)+q} - X_{\phi(k)}\right\|_{L^p} \leqslant \frac{1}{2^k}$$

Considérons alors la suite de fonctions mesurables suivante :

$$\forall n \in \mathbb{N}, \forall \omega \in \Omega, Y_n(\omega) \overset{\text{déf.}}{=} \sum_{k=0}^{n-1} |X_{k+1}(\omega) - X_k(\omega)|$$

Alors par construction de ϕ :

$$\forall n \in \mathbb{N}, \|Y_n\|_{L^p} \leqslant 1$$

De plus, $(Y_n)_n$ est une suite croissante. Par théorème de convergence monotone, il existe une fonction mesurable Y et un négligeable N tel que

$$\forall \omega \in \Omega \backslash N, Y_n(\omega) \xrightarrow[n \to +\infty]{} Y(\omega)$$

Alors, pour tout $q \in \mathbb{N}$, $n \in \mathbb{N}$, on a sur $\Omega \backslash N$:

$$\left|X_{\phi(n+q)} - X_{\phi(n)}\right| \leqslant \sum_{k=n}^{n+q-1} \left|X_{\phi(k+1)} - X_{\phi(k)}\right|$$

Ainsi, on a alors

$$\left|X_{\phi(n+q)} - X_{\phi(n)}\right| \leqslant Y - Y_n$$

Par croissance de $(Y_n)_n$ vers Y, il suit que pour n assez grand, si $\varepsilon > 0$, on a alors

$$\forall q \in \mathbb{N}, \forall \omega \in \Omega \backslash N, \left|X_{\phi(n+q)}(\omega) - X_{\phi(n)}(\omega)\right| \leqslant \varepsilon$$

Ainsi, la suite $\left(X_{\phi(n)}\right)_n$ est presque sûrement de Cauch sur $\mathbb{R}$, donc converge presque sûrement. On a donc bien montré le point **2.**. On note X la limite presque sûre de $(X_{\phi(n)})_n$. Montrons que $X \in L^p$ et que $X_{\phi(n)}$ converge vers X dans L^p. Pour cela, rappelons nous que par hypothèse

$$\forall k \in \mathbb{N}^*, \forall q \in \mathbb{N}, \left\|X_{\phi(k)+q} - X_{\phi(k)}\right\|_{L^p} \leqslant \frac{1}{2^k}$$

Donc, lorsque $[q \to +\infty]$:

$$\forall k \in \mathbb{N}^*, \left\|X - X_{\phi(k)}\right\|_{L^p} \leqslant \frac{1}{2^k}$$

Il suit qu'à k fixé :

$$\|X\|_{L^p} \leqslant \|X_{\phi(n)} - X\|_{L^p} + \|X_{\phi(n)}\|_{L^p} < \infty$$

donc $X \in L^p$. De plus, on a alors la convergence dans L^p de $(X_{\phi(k)})_k$ vers X. Par conséquent, la suite $(X_n)_n$ est une suite de Cauchy admettant une sous-suite convergente □

19 Méthode de la phase stationnaire

Référence : Analyse pour l'agrégation, Hervé QUEFFÉLEC, Claude ZUILY

Théorème 19.1 : Méthode de la phase stationnaire - cas où φ' ne s'annule pas

Soient $\varphi \in \mathcal{C}^\infty(\mathbb{R})$ réelle et $a \in \mathcal{C}_0^\infty(\mathbb{R}, \mathbb{C})$. On définit

$$\forall t \in \mathbb{R}, F(t) \overset{\text{déf.}}{=} \int_{\mathbb{R}} e^{\imath t\varphi(x)} a(x) \,\mathrm{d}x$$

On suppose que φ' ne s'annule pas sur le supporte de a. Alors

$$\forall N \in \mathbb{N}, \exists C_N > 1, \forall t \geqslant 1, |F(t)| \leqslant \frac{C_N}{t^N}$$

On dit que F est à *décroissance rapide*.

Démonstration : L'astuce cruciale ici est la suivante :

$$\left(e^{\imath t\varphi}\right)'(x) = \imath t\varphi'(x) e^{\imath t\varphi(x)}$$

D'où

$$e^{\imath t\varphi(x)} = \frac{1}{\imath t\varphi'(x)} \left(e^{\imath t\varphi}\right)'(x)$$

On montre le théorème par récurrence, à l'aide d'intégrations par parties grâce à cette observation. Plus précisément, on montre qu'il existe $(a_N)_{N\in\mathbb{N}}$ une suite de fonctions C^∞ à support compact telle que pour tout $N \in \mathbb{N}$,

$$\forall t \geqslant 1, F(t) = \frac{1}{t^N} \int_{\mathbb{R}} e^{\imath t\varphi(x)} a_N(x) \,\mathrm{d}x$$

- Pour $N = 0$, il suffit de poser $a_0 = a$.
- Soit $N \in \mathbb{N}$. Supposons qu'il existe $a_N \in \mathcal{C}_0^\infty(\mathbb{R})$ telle que

$$\forall t \geqslant 1, F(t) = \frac{1}{t^N} \int_{\mathbb{R}} e^{\imath t\varphi(x)} a_N(x) \,\mathrm{d}x$$

On a alors :

$$\forall t \geqslant 1, F(t) = \frac{1}{t^N} \int_{\mathbb{R}} \frac{1}{\imath t\varphi'(x)} \left(e^{\imath t\varphi}\right)'(x) a_N(x) \,\mathrm{d}x$$

On fait alors une intégration par parties (les termes de bord sont nuls par a_N est à support compact) :

$$\forall t \geqslant 1, F(t) = \frac{-1}{\imath t^{N+1}} \int_{\mathbb{R}} e^{\imath t\varphi(x)} \left(\frac{a_N}{\varphi'}\right)'(x) \,\mathrm{d}x$$

Il suffit alors de poser

$$a_{N+1} \overset{\text{déf.}}{=} \imath \left(\frac{a_N}{\varphi'}\right)' \in \mathcal{C}_0^\infty(\mathbb{R})$$

Pour avoir

$$\forall t \geqslant 1, F(t) = \frac{1}{t^{N+1}} \int_{\mathbb{R}} e^{\imath t\varphi(x)} a_{N+1}(x) \,\mathrm{d}x$$

Il suit alors que pour tout $N \in \mathbb{N}$:

$$\forall t \geqslant 1, |F(t)| \leqslant \frac{|\mathrm{supp}(a_N)| \|a_N\|_{L^\infty}}{t^N}$$

□

Théorème 19.2 : Méthode de la phase stationnaire - cas où φ' s'annule

Soient $\varphi \in \mathcal{C}^\infty(\mathbb{R})$ réelle et $a \in \mathcal{C}_0^\infty(\mathbb{R}, \mathbb{C})$. On définit

$$\forall t \in \mathbb{R}, F(t) \overset{\text{déf.}}{=} \int_{\mathbb{R}} e^{\imath t\varphi(x)} a(x) \, \mathrm{d}x$$

On suppose qu'il existe un unique point x_0 dans le support de a tel que $\varphi'(x_0) = 0$, et $\varphi''(x_0) \neq 0$.
Alors il existe une suite $(A_n)_{n\in\mathbb{N}} \in \mathbb{C}^{\mathbb{N}}$ tel que pour tout $N \in \mathbb{N}$ et $t \geqslant 1$,

$$F(t) = e^{\imath t\varphi(x_0)} \sum_{n=0}^{N} \frac{A_n}{t^n\sqrt{t}} + R_N(t)$$

où, en notant ε le signe de $\varphi''(x_0)$:

$$\begin{cases} A_0 &= \sqrt{\dfrac{2\pi}{|\varphi''(x_0)|}} e^{\imath\varepsilon\frac{\pi}{4}} a(x_0) \\ |R_N(t)| &\leqslant \dfrac{C_N}{t^{N+1}\sqrt{t}} \end{cases}$$

D'où si $a(x_0) \neq 0$,

$$F(t) \underset{t\to+\infty}{\sim} \frac{A_0}{\sqrt{t}}$$

Remarque : La démonstration présentée dans la référence utilise la *racine carrée d'un complexe de partie réelle strictement positive*. Pour $z \in \mathbb{C}$ avec $\Re z > 0$, on note $\sqrt{z}$ l'unique complexe de partie réelle strictement positive tel que son carré vaut z. Si $z = \rho e^{\imath\theta}$ avec $\theta \in \left]\frac{-\pi}{2}, \frac{\pi}{2}\right[$, alors

$$\sqrt{z} \overset{\text{déf.}}{=} \sqrt{\rho} e^{\imath\frac{\theta}{2}} \in \{\Re > 0\}$$

z possède une autre racine carrée dans $\mathbb{C}$, mais de partie réelle négative donnée par $\sqrt{\rho} e^{\imath\left(\frac{\theta}{2}+\pi\right)}$.

Lemme 19.1

Pour tout $\varepsilon > 0$, pour tout $\lambda \in \mathbb{R}$,

$$\forall \xi \in \mathbb{R}, \mathcal{F}\left[e^{-\varepsilon x^2 + \imath\lambda x^2}\right](\xi) = \frac{\sqrt{\pi}}{\sqrt{\varepsilon - \imath\lambda}} \exp\left(\frac{-\xi^2}{4(\varepsilon - \imath\lambda)}\right)$$

Remarque : La convention choisie pour la transformée de Fourier est la suivante :

$$\hat{f}(\xi) \overset{\text{déf.}}{=} \int_{\mathbb{R}} e^{\imath x\xi} f(x) \, \mathrm{d}x$$

La formule de Parseval donne dans ce cas :

$$\int_{\mathbb{R}} f(x)\overline{g(x)} \, \mathrm{d}x = \frac{1}{2\pi} \int_{\mathbb{R}} \hat{f}(\xi)\overline{\hat{g}(\xi)} \, \mathrm{d}\xi$$

Et la transformée de Fourier inverse est :

$$f(x) = \frac{1}{2\pi} \int_{\mathbb{R}} \hat{f}(\nu) e^{\imath x\nu} \, \mathrm{d}\nu$$

Lemme 19.2

Pour tout $N \in \mathbb{N}$, pour tout $y \in \mathbb{R}$,

$$e^{\imath y} = \sum_{n=0}^{N} \frac{(\imath y)^n}{n!} + A_N(y)$$

avec

$$|A_N(y)| \leqslant \frac{|y|^{N+1}}{(N+1)!}$$

Lemme 19.3 : Cas particulier du théorème

Soit $b \in \mathcal{C}_0^\infty(\mathbb{R}, \mathbb{C})$. On définit

$$\forall \lambda \in \mathbb{R}^*, G(\lambda) \overset{\text{déf.}}{=} \int_{\mathbb{R}} e^{\imath \lambda x^2} \overline{b(x)} \, \mathrm{d}x$$

Alors

$$G(\lambda) = \sqrt{\frac{\pi}{|\lambda|}} e^{\imath \frac{\pi}{4} \mathsf{sgn}(\lambda)} \sum_{n=0}^{N} \frac{1}{\lambda^n} \left[\frac{\imath^n \bar{b}^{(2n)}(0)}{4^n n!} \right] + R_N(\lambda)$$

où

$$|R_N(\lambda)| \leqslant \frac{1}{|\lambda|^{N+1} \sqrt{|\lambda|}} \left[C_N \int_{\mathbb{R}} |\xi|^{2N+2} \left| \hat{b}(\xi) \right| \, \mathrm{d}\xi \right]$$

Démonstration du lemme : On cherche à utiliser une relation du type Parseval, mais $x \longmapsto e^{\imath \lambda x^2}$ n'est pas L^2. On introduit $\varepsilon > 0$, et une fonction G_ε qui converge vers G, dans un sens que l'on va préciser. On considère alors

$$G_\varepsilon(\lambda) \overset{\text{déf.}}{=} \int_{\mathbb{R}} e^{\imath \lambda x^2 - \varepsilon x^2} \bar{b}(x) \, \mathrm{d}x$$

Alors

$$\begin{cases} e^{\imath \lambda x^2 - \varepsilon x^2} \bar{b}(x) & \xrightarrow[\varepsilon \to 0^+]{} & e^{\imath \lambda x^2} \bar{b}(x) \\ \left| e^{\imath \lambda x^2 - \varepsilon x^2} \bar{b}(x) \right| & \leqslant & |b(x)| \end{cases}$$

avec $b \in L^1$. D'après le théorème de convergence dominée,

$$\forall \lambda \in \mathbb{R}^*, G_\varepsilon(\lambda) \xrightarrow[\varepsilon \to 0^+]{} G(\lambda)$$

Or, d'après la formule de Parseval, et le lemme **19.1** :

$$G_\varepsilon(\lambda) = \frac{1}{2\pi} \int_{\mathbb{R}} \frac{\sqrt{\pi}}{\sqrt{\varepsilon - \imath \lambda}} \exp\left(\frac{-\xi^2}{4(\varepsilon - \imath \lambda)} \right) \hat{\bar{b}}(\xi) \, \mathrm{d}\xi$$

Le théorème de convergence dominée s'applique encore ici pour tous les termes dans l'intégrale. De plus,

$$\sqrt{\varepsilon - \imath \lambda} \xrightarrow[\varepsilon \to 0^+]{} \sqrt{|\lambda|} e^{\frac{\imath \pi}{4} \mathsf{sg}(\lambda)}$$

Par unicité de la limite, il suit que

$$G(\lambda) = \frac{1}{2\pi} \int_{\mathbb{R}} \frac{e^{\frac{-\imath \pi}{4} \mathsf{sg}(\lambda)} \sqrt{\pi}}{\sqrt{|\lambda|}} \exp\left(\frac{-\imath \xi^2}{4\lambda} \right) \hat{\bar{b}}(\xi) \, \mathrm{d}\xi$$

D'après le lemme **19.2**, on a pour tout $N \in \mathbb{N}$:

$$\exp\left(\frac{-\imath \xi^2}{4\lambda} \right) = \sum_{n=0}^{N} \left[\frac{(-1)^n \imath^n}{4^n n!} \right] \frac{\xi^{2n}}{\lambda^n} + A_N \left(\frac{-\xi^2}{4\lambda} \right)$$

Avec

$$\left| A_N \left(\frac{-\xi^2}{4\lambda} \right) \right| \leqslant \left[\frac{|\xi|^{2N+2}}{4^{N+1}(N+1)!} \right] \frac{1}{\lambda^{N+1}}$$

On pose alors

$$R_N(\lambda) \overset{\text{déf.}}{=} \frac{e^{\frac{-\imath \pi}{4} \mathsf{sg}(\lambda)}}{2\pi} \sqrt{\frac{\pi}{|\lambda|}} \int_{\mathbb{R}} A_N \left(\frac{-\xi^2}{4\lambda} \right) \hat{\bar{b}}(\xi) \, \mathrm{d}\xi$$

D'où l'existence d'une constante $C_N > 0$ telle que

$$|R_N(\lambda)| \leqslant \frac{C_N}{|\lambda|^{N+1}\sqrt{\lambda}} \int_{\mathbb{R}} |\xi|^{2N+2} \left|\hat{\bar{b}}(\xi)\right| \,\mathrm{d}\xi$$

On a alors :

$$G(\lambda) = \frac{e^{\frac{-\imath\pi}{4}\mathrm{sg}(\lambda)}}{2\pi}\sqrt{\frac{\pi}{|\lambda|}}\sum_{n=0}^{N}\frac{1}{\lambda^n}\left[\frac{(-1)^n\imath^n}{4^n n!}\right]\int_{\mathbb{R}}\xi^{2n}\hat{\bar{b}}(\xi)\,\mathrm{d}\xi + R_N(\lambda)$$

Enfin, d'après la formule d'inversion de Fourier :

$$\bar{b}(x) = \frac{1}{2\pi}\int_{\mathbb{R}} e^{\imath x\xi}\hat{\bar{b}}(\xi)\,\mathrm{d}\xi$$

$\bar{b}$ est de classe C^∞, et à support compact, donc on peut dériver sous l'intégrale pour avoir :

$$\bar{b}^{(2n)}(0) = \frac{\imath^{2n}}{2\pi}\int_{\mathbb{R}}\xi^{2n}\hat{\bar{b}}(\xi)\,\mathrm{d}\xi$$

Ainsi, on a compensation des $(-1)^n$ et des 2π, pour finalement aboutir au résultat souhaité :

$$G(\lambda) = \frac{e^{\frac{-\imath\pi}{4}\mathrm{sg}(\lambda)}}{2\pi}\sqrt{\frac{\pi}{|\lambda|}}\sum_{n=0}^{N}\frac{1}{\lambda^n}\left[\frac{\bar{b}^{(2n)}(0)\imath^n}{4^n n!}\right] + R_N(\lambda)$$

□

Démonstration : • D'après le formule de Taylor avec reste intégrale, pour tout $x \in \mathbb{R}$:

$$\varphi(x) = \varphi(x_0) + (x-x_0)^2\psi(x)$$

avec $\psi \in \mathcal{C}^\infty(\mathbb{R})$ telle que

$$\psi(x_0) = 2\varphi''(x_0)$$

Ainsi,

$$F(t) = e^{\imath t\varphi(x_0)}\int_{\mathbb{R}} e^{\imath t y^2\psi(y+x_0)}a(y+x_0)\,\mathrm{d}y$$

après avoir fait le petit changement de variables $y = x - x_0$. Pour se ramener au cas du lemme **19.3**, on souhaite faire le changement de variables

$$z^2 = y^2\psi(y+x_0)$$

• Montrons que ce changement de variables est possible pour un certain $y \in [-\delta,\delta]$. On pose $\varepsilon \overset{\text{déf.}}{=} \mathrm{sgn}(\varphi''(x_0))$, et

$$h(y) = y\sqrt{\varepsilon\psi(x_0+y)}$$

Puisque $\psi(x_0) = 2\varphi''(x_0) \neq 0$, et que h est continue, il existe un voisinage $]-\delta_1,\delta_1[$ sur lequel $\psi(x_0+y)$ ne s'annule pas, donc où h est dérivable. Sur $]-\delta_1,\delta_1[$,

$$h'(y) = \sqrt{\varepsilon\psi(y+x_0)} + \frac{\sqrt{\varepsilon}y\psi'(y+x_0)}{2\sqrt{\psi(x_0+y)}}$$

Alors $h'(0) = \sqrt{2|\varphi''(x_0)|} > 0$. Ainsi, par continuité de h', il existe $[-\delta,\delta]$ sur lequel h' ne s'annule pas. h réalise alors un C^1-difféomorphisme sur $]-\delta,\delta[$ dans son image. On considère $\chi \in \mathcal{C}_0^\infty(\mathbb{R})$ telle que $\chi = 1$ sur $\left[\frac{-\delta}{2},\frac{\delta}{2}\right]$, et $\chi = 0$ en dehors de $\left[\frac{-3\delta}{4},\frac{3\delta}{4}\right]$.

• On découpe alors F en deux morceaux :

$$F(t) = e^{\imath t\varphi(x_0)}\int_{\mathbb{R}}\chi(y)e^{\imath t\varepsilon h(y)^2}a(y+x_0)\,\mathrm{d}y + e^{\imath t\varphi(x_0)}\int_{\mathbb{R}}(1-\chi(y))e^{\imath t\varepsilon h(y)^2}a(y+x_0)\,\mathrm{d}y$$

On note $F_1(t)$ la première intégrale, et $F_2(t)$ la seconde.

• Pour F_2, constatons que sur le support de $(1-\chi)$, la dérivée de $y \longmapsto h(y)^2$ ne s'annule pas (φ' ne s'annule qu'une seule fois, en x_0). On se retrouve alors dans le contexte du théorème **19.1** (« cas où φ' ne s'annule pas »). Ainsi, on peut directement conclure que F_2 est à décroissance rapide : pour tout $M \in \mathbb{N}$, il existe une constante C_M telle que

$$\forall t \geqslant 1, |F_2(t)| \leqslant \frac{C_M}{t^M}$$

En particulier, pour un $N \in \mathbb{N}$ fixé, $F_2 = O\left(\frac{1}{t^{N+1}\sqrt{t}}\right)$.

• Pour F_1, on peut effectuer le changement de variables $z = h(y)$, ou encore $y = g(z)$, avec g l'inverse de f sur $]-\delta,\delta[$. On obtient

$$F_1(t) = e^{\imath t\varphi(x_0)}\int_{\mathbb{R}} e^{\imath t\varepsilon z^2}a(g(z)+x_0)\chi(g(z))g'(z)\,\mathrm{d}z$$

On note $b(z) = a(g(z)+x_0)\chi(g(z))g'(z)$. b est C^∞ à support compact. On retrouve alors enfin dans le cas du lemme **19.3**. On a alors :

$$F_1(t) = e^{\imath t\varphi(x_0)}\left[\sqrt{\frac{\pi}{t}}e^{\frac{-\imath\pi\varepsilon}{4}}\sum_{n=0}^{N}\frac{B_n}{t^n} + R_N(t)\right]$$

avec

$$R_N(t) = O\left(\frac{1}{t^{N+1}\sqrt{t}}\right)$$

• Bilan : on a bien

$$F(t) = e^{\imath t\varphi(x_0)}\sqrt{\frac{\pi}{t}}e^{\frac{-\imath\pi\varepsilon}{4}}\sum_{n=0}^{N}\frac{B_n}{t^n} + S_N(t)$$

où $S_N(t) = O\left(\frac{1}{t^{N+1}\sqrt{t}}\right)$. Reste à calculer B_0. D'après le lemme **19.3** :

$$B_0 = b(0) = a(x_0)\chi(0)g'(0)$$

avec $g'(0) = \frac{1}{h'(0)} = \frac{1}{\sqrt{2|\varphi''(x_0)|}}$. Donc

$$B_0 = \frac{a(x_0)}{\sqrt{|\varphi''(x_0)|}}$$

- Finalement, on a bien montré que

$$F(t) = e^{\imath t\varphi(x_0)} \sum_{n=0}^{N} \frac{A_n}{t^n\sqrt{t}} + O\left(\frac{1}{t^{N+1}\sqrt{t}}\right)$$

avec

$$A_0 = \sqrt{\pi} e^{\frac{-\imath\pi\varepsilon}{4}} B_0$$

□

20 Polynôme caractéristique d'un endomorphisme nilpotent

Référence : Les maths en tête, Algèbre, Xavier Gourdon

Proposition 20.1

Soit E un $\mathbb{K}$-espace vectoriel de dimension finie n. Soit $f \in \mathcal{L}(E)$ un endomorphisme nilpotent d'indice q. Avec la convention

$$\chi_f(X) \overset{\text{déf.}}{=} \det(f - X\mathrm{id}_E)$$

On a alors

$$\chi_f(X) = (-1)^n X^n$$

Il existe deux façons de le montrer : l'une avec une considération d'extension de corps, l'autre par récurrence plus fastidieuse.

Démonstration : On note $\hat{K}$ le corps de décomposition de χ_f dans $\mathbb{K}$. On fixe une base $\mathcal{B}$ de E, et on considère

$$A \overset{\text{déf.}}{=} \mathrm{Mat}_{\mathcal{B}}(f) \in \mathcal{M}_n(\mathbb{K})$$

Sur $\hat{\mathbb{K}}$, χ_f s'écrit

$$\chi_f = (-1)^n \prod_{i=1}^{n} (X - \lambda_i)$$

avec $\lambda \in \hat{\mathbb{K}}$. On considère A comme étant un élément de $\mathcal{M}_n\left(\hat{\mathbb{K}}\right)$. Pour tout $i \in [\![1, n]\!]$, λ_i est un vecteur propre de A, que l'on associe à un vecteur propre $X_i \in \hat{\mathbb{K}}^n$. Alors en particulier, on a :

$$A^q X_i = \lambda_i^q X_i = 0$$

Puisque X_i est un vecteur propre de A, il est en particulier non nul, donc cela implique que pour tout $i \in [\![1, n]\!]$, $\lambda_i^q = 0$, donc $\lambda_i = 0$ par intégrité du corps $\hat{\mathbb{K}}$. D'où

$$\chi_f = (-1)^n X^n$$

□

Cette démonstration est concise, et se comprend aisément sous le prisme de $\mathbb{C}$ comme extension algébriquement close de $\mathbb{R}$. La deuxième démonstration est plus visuelle, et n'utilise pas les extensions de corps, mais prend plus de temps à rédiger.

Démonstration : On démontre ce résultat par récurrence sur n la dimension de E.

- Pour $n = 1$, tout endomorphisme est de la forme $f(x) = \lambda x$, donc le seul endomorphisme nilpotent est l'endomorphisme nul. En particulier, le polynôme caractéristique (qui est de degré 1) vaut $-X$.
- Soit $n \geqslant 1$. Supposons que pour tout $\mathbb{K}$-espace vectoriel de dimension n, et pour tout endomorphisme nilpotent f,

$$\chi_f(X) = (-1)^n X^n$$

Soit E un $\mathbb{K}$-espace vectoriel de dimension $n+1$. Soit $f \in \mathcal{L}(E)$ nilpotent d'indice q. Constatons que

$$\det(f)^q = \det(f^q) = 0$$

donc par intégrité de $\mathbb{K}$, $\det(f) = 0$. Par conséquent, f est non inversible, donc non injective (E est de dimension finie). Soit $e_1 \in \ker(f)$ non nul. On considère $\mathcal{B}$ une base de E qui complète la famille libre (e_1), et on note

$$A \overset{\text{déf.}}{=} \mathrm{Mat}_{\mathcal{B}}(f) \in \mathcal{M}_{n+1}(\mathbb{K})$$

Alors A se présente sous la forme :

$$A = \begin{pmatrix} 0 & * \\ 0 & M \end{pmatrix}$$

avec $M \in \mathcal{M}_n(\mathbb{K})$. Alors, par propriété de multiplication de matrices triangulaire par blocs, on a :

$$A^q = \begin{pmatrix} 0 & * \\ 0 & M^q \end{pmatrix}$$

Or, A est nilpotente d'indice q, donc on a

$$M^q = 0_{n,n}$$

M est donc une matrice nilpotente d'ordre q de taille n. On peut alors appliquer l'hypothèse de récurrence sur M :

$$\chi_M(X) = (-1)^n X^n$$

Ainsi,

$$\chi_A(X) = \begin{vmatrix} -X & * \\ 0 & \chi_M(X) \end{vmatrix}$$

Donc

$$\chi_A(X) = -X\chi_M(X) = (-1)^{n+1}X^{n+1}$$

ce qui conclut la récurrence. □

21 Trigonalisation et polynôme caractéristique

Référence : Les maths en tête, Algèbre, Xavier Gourdon

Théorème 21.1

Soit E un $\mathbb{K}$-espace vectoriel de dimension finie n. Soit $f \in \mathcal{L}(E)$. Les deux propositions suivantes sont équivalentes :

(i) f est trigonalisable dans E ;

(ii) Le polynôme caractéristique χ_f de f est scindé dans $\mathbb{K}$.

On utilise la convention

$$\chi_f(X) \overset{\text{déf.}}{=} \det(f - X\mathrm{id}_E)$$

L'implication (ii) $\Longrightarrow$ (i) utilise deux méthodes. Montrons d'abord l'implication directe.

Démonstration : (i) $\longrightarrow$ (ii) Soit $\mathcal{B}$ une base dans laquelle la matrice de f est triangulaire supérieure. On note A cette matrice, et λ_i les coefficients diagonaux de A. χ_A se présente donc sous la forme

$$\chi_A(X) = \begin{vmatrix} \lambda_1 - X & & * & * \\ 0 & \lambda_2 - X & & * \\ \vdots & & \ddots & \\ 0 & \cdots & 0 & \lambda_n - X \end{vmatrix}$$

D'où

$$\chi_A(X) = (-1)^n \prod_{i=1}^{n} (X - \lambda_i)$$

donc χ_f est scindé sur $\mathbb{K}$. □

Pour l'implication réciproque, une première méthode utilise l'application transposée d'un endomorphisme. Si $f \in \mathcal{L}(E)$, on définit

$${}^t f : \begin{pmatrix} E^* & \longrightarrow & E^* \\ u & \longmapsto & u \circ f \end{pmatrix}$$

Si $\mathcal{B}$ est une base de E, et $\mathcal{B}^*$ sa base duale alors

$$\mathrm{Mat}_{\mathcal{B}^*}\left({}^t f\right) = {}^t\mathrm{Mat}_{\mathcal{B}}(f)$$

Démonstration : (ii) ⟶ (i) On démontre cette implication par récurrence sur n, le dimension de E.

• Pour $n = 1$, $\chi_f(X) = -\lambda X$, avec $\lambda \neq 0$. Il est donc scindé.

• Soit $n \in \mathbb{N}^*$. Supposons que pour tout endomorphisme sur un $\mathbb{K}$-espace vectoriel de dimension n de polynôme caractéristique scindé dans $\mathbb{K}$, cet endomorphisme est trigonalisable. Soit E un $\mathbb{K}$-espace vectoriel de dimension $n+1$. Soit $f \in \mathcal{L}(E)$ possédant un polynôme caractéristique scindé. On dispose de la relation suivante pour le polynôme caractéristique de ${}^t f$:

$$\chi_{{}^t f} = \chi_f$$

En conséquence, $\chi_{{}^t f}$ est scindé dans $\mathbb{K}$. ${}^t f$ possède un vecteur propre $x \in E^*$. $\mathbb{K}x$ est alors stable par ${}^t f$. On considère H l'orthogonal de $\mathbb{K}x$ dans $\mathbb{K}^n$. C'est un hyperplan stable par f (on utilise la propriété F est stable par f si et seulement si ${}^\perp F$ est stable par ${}^t f$). Le polynôme caractéristique de $f_{|_H}$ divise χ_f. Par hypothèse de récurrence, il existe une base $\mathcal{B}_n$ de H dans laquelle la matrice de $f_{|_H}$ est triangulaire supérieure. On complète $\mathcal{B}_n$ en $\mathcal{B}$ base de E. Ainsi, si

$$A \stackrel{\text{déf.}}{=} \mathrm{Mat}_{\mathcal{B}}(f)$$

et

$$B \stackrel{\text{déf.}}{=} \mathrm{Mat}_{\mathcal{B}_n}\left(f_{|_H}\right)$$

On obtient :

$$A = \begin{pmatrix} B & * \\ 0 & \alpha \end{pmatrix}$$

avec B triangulaire supérieure. Ainsi, A est aussi triangulaire supérieure. Cela conclut la récurrence. □

La deuxième méthode est plus élémentaire.

Démonstration : (ii) ⟶ (i) On démontre ici aussi l'implication par récurrence sur n. Pour $n = 1$, l'implication est vérifiée. Supposons que pour tout endomorphisme sur un $\mathbb{K}$-espace vectoriel de dimension n de polynôme caractéristique scindé dans $\mathbb{K}$, cet endomorphisme est trigonalisable. Soit E un $\mathbb{K}$-espace vectoriel de dimension $n+1$. Soit $f \in \mathcal{L}(E)$ possédant un polynôme caractéristique scindé. Il existe e_1 un vecteur propre de f. On complète la famille libre (e_1) (le vecteur est propre) en une base $\mathcal{B}$ de E. On note

$$A \stackrel{\text{déf.}}{=} \mathrm{Mat}_{\mathcal{B}}(f) \in \mathcal{M}_{n+1}(\mathbb{K})$$

Alors A se présente sous la forme

$$A = \begin{pmatrix} \alpha & * \\ 0 & B \end{pmatrix}$$

avec $\alpha \in \mathbb{K}$ et $B \in \mathcal{M}_n(\mathbb{K})$. B est la matrice d'un endomorphisme g sur H, où

$$H \stackrel{\text{déf.}}{=} \mathrm{Vect}_{\mathbb{K}}(e_2, \cdots, e_{n+1})$$

est un supplémentaire de (e_1) dans E. Alors

$$\chi_A(X) = (X - \alpha)\chi_B(X)$$

χ_A étant scindé, il suit que $\chi_B = \chi_g$ est scindé sur $\mathbb{K}$. g étant définie sur H de dimension n, l'hypothèse de récurrence nous permet d'obtenir une base $\mathcal{B}_H$ dans laquelle la matrice de g est triangulaire supérieure. La base

$$\mathcal{B}' = e_1 \sqcup \mathcal{B}_H$$

est alors une base qui trigonalise f :

$$\mathrm{Mat}_{\mathcal{B}}(f) = \begin{pmatrix} \alpha & * \\ & \mathrm{Mat}_{\mathcal{B}_H}(g) \end{pmatrix}$$

Cela conclut la récurrence. □

22 Réduction pratique

Référence : Les maths en tête, Algèbre, Xavier Gourdon

Dans $\mathbb{C}$, on se retrouve face à 4 de figures quand on cherche à réduire une matrice :

— Le polynôme caractéristique est scindé à racines simples : tout roule.

— χ est scindé dans $\mathbb{R}$, mais pas à racines simples. Si chaque sous-espace propre est de dimension la multiplicité dans χ de λ, la matrice se diagonalise, tout roule.

— Si les multiplicités ne concordent pas, et qu'on ne trouve que $n-1$ vecteurs propres pour A, on la trigonalise.

— Pire des cas, pas de concordance des multiplicités géométriques, et on ne dispose qu'un « faible »nombre de vecteurs propres. On utilise alors l'une des démonstrations du théorème **21.1**.

On va illustrer ces trois derniers cas par des exemples.

Proposition 22.1 : Cas plutôt favorable

Soit

$$M \overset{\text{déf.}}{=} \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix}$$

M est diagonalisable dans $\mathbb{C}$ (donc possède quatre directions propres).

Démonstration : Le polynôme caractéristique de M s'écrit

$$\chi_M(X) = (X^2+1)^2 = (X-\imath)^2(X+\imath)^2$$

Déterminons les sous-espaces propres de M en $\imath$ et $-\imath$. On peut soit observer la matrice $M - \imath I_4$:

$$M - \imath I_4 = \begin{pmatrix} -\imath & 0 & 0 & 1 \\ 0 & -\imath & -1 & 0 \\ 0 & 1 & -\imath & 0 \\ -1 & 0 & 0 & -\imath \end{pmatrix}$$

et trouver les relations $C_1 + \imath C_4 = 0$ et $\imath C_2 + C_3 = 0$, ou gentiment résoudre le système $MX = \imath X$, qui donne ici simplement

$$\begin{cases} t & = & \imath x \\ -z & = & \imath y \\ y & = & \imath z \\ -x & = & \imath t \end{cases}$$

Ainsi, on trouve

$$E_\imath(M) = \text{Vect}_\mathbb{C}\left\{\begin{pmatrix} 1 \\ 0 \\ 0 \\ \imath \end{pmatrix}, \begin{pmatrix} 0 \\ \imath \\ 1 \\ 0 \end{pmatrix}\right\}$$

De même, on trouve

$$E_{-\imath}(M) = \text{Vect}_\mathbb{C}\left\{\begin{pmatrix} \imath \\ 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ \imath \\ 0 \end{pmatrix}\right\}$$

C'est donc seulement ici que l'on peut conclure que M se retrouve diagonalisable avec :

$$M = P \begin{pmatrix} -\imath & 0 & 0 & 0 \\ 0 & -\imath & 0 & 0 \\ 0 & 0 & \imath & 0 \\ 0 & 0 & 0 & \imath \end{pmatrix} P^{-1}$$

où

$$P = \begin{pmatrix} \imath & 0 & 1 & 0 \\ 0 & 1 & 0 & \imath \\ 0 & \imath & 0 & 1 \\ 1 & 0 & \imath & 0 \end{pmatrix}$$

C'est donc ici un cas plutôt favorable. □

Proposition 22.2 : Cas pas trop favorable

Soit

$$M \overset{\text{déf.}}{=} \begin{pmatrix} 1 & 4 & -2 \\ 0 & 6 & -3 \\ -1 & 4 & 0 \end{pmatrix}$$

M est trigonalisable dans $\mathbb{C}$ et possède deux directions propres.

Démonstration : Le polynôme caractéristique de M est :

$$\chi_M(X) = -(X-3)(X-2)^2$$

La résolution de $MX = 3X$ donne

$$E_3(M) = \mathrm{Vect}_{\mathbb{C}} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

Quant à la résolution de $MX = 2X$:

$$E_2(M) = \mathrm{Vect}_{\mathbb{C}} \begin{pmatrix} 4 \\ 3 \\ 4 \end{pmatrix}$$

Puisque la dimension du sous-espace propre $E_2(M)$ est différente de la multiplicité de 2 dans χ_M, M n'est pas diagonalisable. On cherche donc ici à la trigonaliser. Les vecteurs $e_1 \overset{\text{déf.}}{=} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ et $e_2 \overset{\text{déf.}}{=} \begin{pmatrix} 4 \\ 3 \\ 4 \end{pmatrix}$ sont indépendants, et on complète cette famille libre avec $e_3 \overset{\text{déf.}}{=} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$ pour avoir une base de $\mathbb{C}^3$. On a $Me_1 = 3e_1$ et $Me_2 = 2e_2$. On note P la matrice de passage de la base canonique à cette nouvelle base :

$$P = \begin{pmatrix} 1 & 4 & 0 \\ 1 & 3 & 0 \\ 1 & 4 & 1 \end{pmatrix}$$

Alors

$$P^{-1}(Me_3) = \begin{pmatrix} -6 \\ 1 \\ 2 \end{pmatrix}$$

donc on a finalement

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

où

$$T = \begin{pmatrix} 3 & 0 & -6 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}$$

On remarque que cela se corse un peu dans le cas non diagonalisable, mais le cas suivant est pire. □

Proposition 22.3 : Cas pas favorable du tout

Soit

$$M \overset{\text{déf.}}{=} \begin{pmatrix} 2 & 2 & -3 \\ 5 & 1 & -5 \\ -3 & 4 & 0 \end{pmatrix}$$

M est trigonalisable dans $\mathbb{C}$ et ne possède qu'une seule direction propre.

Démonstration : Le calcul du polynôme caractéristique donne (somme des colonnes) :

$$\chi_M(X) = (1-X)^3$$

La résolution du système $MX = X$ donne alors l'unique direction propre de M :

$$E_1 = \mathrm{Vect}_{\mathbb{C}} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

Contrairement au cas précédent, on ne dipose qu'un seul vecteur propre, et il s'agit de trouver deux autres vecteurs pour former une base où M est semblable à une matrice triangulaire supérieure. On s'inspire pour cela du théorème **21.1**, et de sa démonstration, pour proposer deux méthodes pour trigonaliser M.

1. La première démonstration du théorème utilise les applications transposées. On se donne f l'endomorphisme de $\mathcal{L}(\mathbb{C}^3)$ de matrice M dans la base canonique de $\mathbb{C}^3$. Puisque

$$\chi_f = \chi_{{}^t f}$$

1 est aussi unique valeur propre de ${}^t f$. Si B^* est la base duale de la base canonique de $\mathbb{C}^3$, la matrice de ${}^t f$ dans cette base est ${}^t M$. Soit $u \in \left(\mathbb{C}^3\right)^*$ un vecteur propre de ${}^t f$ pour la valeur propre 1. Matriciellement, cela s'écrit

$${}^t MU = U$$

Alors, on obtient

$$E_1\left({}^t M\right) = \mathrm{Vect}_{\mathbb{C}} \begin{pmatrix} -2 \\ 1 \\ 1 \end{pmatrix}$$

On considère alors H, l'orthogonal de $E_1\left({}^t M\right)$, défini par :

$$H = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{C}^3, -2x + y + z = 0 \right\}$$

Par la propriété suivante : F est stable par f si et seulement si $F^\perp$ est stable par ${}^t f$, il suit que H est stable par f. Une base de H est donnée par $e_1 = \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix}$ et $e_2 = \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}$.

On a

$$\left\{\begin{array}{rcl} f(e_1) & = & -4e_1 - 5e_2 \\ f(e_2) & = & 5e_1 + 6e_2 \end{array}\right.$$

Dans la base (e_1, e_2), $f_{|H}$ admet pour matrice

$$M_1 = \begin{pmatrix} -4 & -5 \\ 5 & 6 \end{pmatrix}$$

Comme dans la démonstration du théorème, on a su se ramener *via* l'application transposée se ramener à de la trigonalisation en dimension inférieure. Trigonalisons M_1. Pour cela, on cherche le sous-espace propre de M_1 associé à 1 : soit $\mathbf{y} \in \mathbb{R}^2$ tel que $M_1\mathbf{y} = \mathbf{y}$. La résolution du système donne

$$E_1(M_1) = \text{Vect}_{\mathbb{C}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

Autrement dit,

$$f(e_1 + e_2) = e_1 + e_2$$

On pose alors $e'_1 = e_1 + e_2 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$, et $e'_2 = e_2 = \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}$. (e'_1, e'_2) est une base de H. Cette fois, on dispose des relations :

$$\left\{\begin{array}{rcl} f(e'_1) & = & e'_1 \\ f(e'_2) & = & 5e'_1 + e'_2 \end{array}\right.$$

Donc, restreinte à H, on dispose d'une forme triangulaire supérieure pour f. Reste à compléter la famille libre (e'_1, e'_2) en une base de $\mathbb{C}^3$, et on pose $e'_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$. Alors

$$f(e'_3) = -3e_1 + 5e_2 + e'_3$$

Soit

$$f(e'_3) = -3e'_1 - 2e'_2 + e'_3$$

Dans la base $\mathcal{B}' = (e'_1, e'_2, e'_3)$, il suit

$$T = \text{Mat}_{\mathcal{B}'}(f) = \begin{pmatrix} 1 & 5 & -3 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{pmatrix}$$

On a alors réussi à trouver une base de trigonalisation pour f :

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

où

$$P = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & -1 & 0 \end{pmatrix}$$

2. La deuxième méthode consiste, comme dans la deuxième démonstration du théorème, à partir d'un vecteur propre que l'on connaît de déterminer pas à pas une base où l'on peut trigonaliser. Ici, on a :

$$E_1 = \text{Vect}_{\mathbb{C}} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

On pose $u_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$, $u_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$ et $u_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$. (u_1, u_2, u_3) forme une base $\mathcal{B}_u$ de $\mathbb{C}^3$ dans laquelle f, l'endomorphisme associé à M pour la base canonique, admet pour matrice :

$$A = \text{Mat}_{\mathcal{B}_u}(f) = \begin{pmatrix} 1 & 2 & -3 \\ 0 & -1 & -2 \\ 0 & 2 & 3 \end{pmatrix}$$

(car $u_1 = e_1 + e_2 + e_3$ somme des trois vecteurs de la base canonique). On note $H = \text{Vect}_{\mathbb{C}}(u_2, u_3)$, et p la projection sur H parallèlement à $\mathbb{C}u_1$. On note $g = p \circ f_{|H}$ On considère le bloc inférieur droit de A :

$$A_1 = \text{Mat}_{(u_2, u_3)}(g) \begin{pmatrix} 2 & -3 \\ -1 & -2 \end{pmatrix}$$

On cherche à trigonaliser A_1. On résout pour cela $A_1Y = Y$ sur $\mathbb{C}^2$, ce qui donne

$$E_1(A_1) = \text{Vect}_{\mathbb{C}} \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

On pose $v_2 = -u_2 + u_3$ et $v_3 = u_2$. (v_2, v_3) forme une base $\mathcal{B}'_v$ de H dans laquelle

$$\text{Mat}_{\mathcal{B}_v}(g) = \begin{pmatrix} 1 & -2 \\ 0 & 1 \end{pmatrix}$$

Ainsi, dans la base $\mathcal{B}_0 = (u_1, v_2, v_3)$, f s'écrit

$$\text{Mat}_{\mathcal{B}_0}(f) = \begin{pmatrix} 1 & 5 & 2 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{pmatrix}$$

On a bien trouvé une base (différente de la première méthode) dans laquelle f est triangulaire supérieure :

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

avec

$$P = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & -1 & 0 \end{pmatrix}$$

□

23 Foire aux calculs de déterminants

Référence : Les maths en tête, Algèbre, Xavier GOURDON

Proposition 23.1 : Déterminant de VANDERMONDE

Soient $n \geqslant 2$, et $a_1, \cdots, a_n \in \mathbb{K}$. On définit le *déterminant de VANDERMONDE* par :

$$V(a_1, \cdots, a_n) \stackrel{\text{déf.}}{=} \begin{vmatrix} 1 & a_1 & a_1^2 & \cdots & a_1^{n-1} \\ 1 & a_2 & a_2^2 & \cdots & a_2^{n-1} \\ \vdots & & & & \vdots \\ \vdots & & & & \vdots \\ 1 & a_n & a_n^2 & \cdots & a_n^{n-1} \end{vmatrix}$$

Alors

$$V(a_1, \cdots, a_n) = \prod_{1 \leqslant i < j \leqslant n} (a_j - a_i)$$

Démonstration : On le montre par récurrence sur $n \geqslant 2$. Pour $n = 2$, on a bien pour $a, b \in \mathbb{K}$:

$$\begin{vmatrix} 1 & a \\ 1 & b \end{vmatrix} = b - a$$

Soit $n \geqslant 2$. Supposons vraie la formule du déterminant de Vandermonde pour n éléments de $\mathbb{K}$. Soient $a_1, \cdots, a_{n+1} \in \mathbb{K}$. Par définition :

$$V(a_1, \cdots, a_{n+1}) \stackrel{\text{déf.}}{=} \begin{vmatrix} 1 & a_1 & a_1^2 & \cdots & a_1^n \\ 1 & a_2 & a_2^2 & \cdots & a_2^n \\ \vdots & & & & \vdots \\ \vdots & & & & \vdots \\ 1 & a_{n+1} & a_{n+1}^2 & \cdots & a_{n+1}^n \end{vmatrix}_{[n+1]}$$

On fait l'opération élémentaire

$$C_{n+1} \leftarrow C_{n+1} - a_1 C_n$$

On obtient (on note V le déterminant en question) :

$$V = \begin{vmatrix} 1 & a_1 & a_1^2 & \cdots & a_1^{n-1} & 0 \\ 1 & a_2 & a_2^2 & \cdots & a_2^{n-1} & a_2^n - a_1 a_2^{n-1} \\ \vdots & & & & & \vdots \\ \vdots & & & & & \vdots \\ 1 & a_{n+1} & a_{n+1}^2 & \cdots & a_{n+1}^{n-1} & a_{n+1}^n - a_1 a_{n+1}^{n-1} \end{vmatrix}_{[n+1]}$$

On fait les opérations élémentaires (c'est la même que C_{n+1}) :

$$\forall i \in [\![2, n]\!], C_i \leftarrow C_i - a_1 C_{i-1}$$

On obtient :

$$V = \begin{vmatrix} 1 & 0 & 0 & \cdots & 0 \\ 1 & a_2 - a_1 & a_2^2 - a_1 a_2 & \cdots & a_2^n - a_1 a_2^{n-1} \\ \vdots & & & & \vdots \\ \vdots & & & & \vdots \\ 1 & a_{n+1} - a_1 & a_{n+1}^2 - a_1 a_{n+1} & \cdots & a_{n+1}^n - a_1 a_{n+1}^{n-1} \end{vmatrix}_{[n+1]}$$

On développe par rapport à la première ligne (on a fait toutes ces opérations pour) :

$$V = \begin{vmatrix} a_2 - a_1 & a_2^2 - a_1 a_2 & \cdots & a_2^n - a_1 a_2^{n-1} \\ \vdots & & & \vdots \\ \vdots & & & \vdots \\ a_{n+1} - a_1 & a_{n+1}^2 - a_1 a_{n+1} & \cdots & a_{n+1}^n - a_1 a_{n+1}^{n-1} \end{vmatrix}_{[n]}$$

Chaque ligne L_i possède le facteur commun $a_i - a_1$. On a alors la factorisation suivante :

$$V = \prod_{i=2}^{n} (a_i - a_1) \begin{vmatrix} 1 & a_2 & \cdots & a_2^{n-1} \\ \vdots & & & \vdots \\ \vdots & & & \vdots \\ 1 & a_{n+1} & \cdots & a_{n+1}^{n-1} \end{vmatrix}_{[n]}$$

D'où

$$V(a_1, \cdots, a_n, a_{n+1}) = \prod_{i=2}^{n} (a_i - a_1) V(a_2, \cdots, a_{n+1})$$

L'hypothèse de récurrence permet alors de conclure cette démonstration. □

Proposition 23.2

Soient $n \geqslant 2$ et $x_1, \cdots, x_n \in \mathbb{K}$. On considère pour tout $k \in [\![1, n-1]\!]$,

$$P_k(X) \overset{\text{déf.}}{=} X^k + \sum_{j=0}^{k-1} a_{j,k} X^j$$

Alors

$$\begin{vmatrix} 1 & P_1(x_1) & P_2(x_1) & \cdots & P_{n-1}(x_1) \\ 1 & P_1(x_2) & P_2(x_2) & \cdots & P_{n-1}(x_2) \\ \vdots & & & & \vdots \\ \vdots & & & & \vdots \\ 1 & P_1(x_n) & P_2(x_n) & \cdots & P_{n-1}(x_n) \end{vmatrix}_{[n]} = V(x_1, \cdots, x_n)$$

Démonstration : On note Δ le déterminant recherché. Par souci de compréhension des degrés, on note dans cette démonstration

$$M = (C_0 \quad | \quad C_1 \quad | \quad \cdots \quad | \quad C_{n-1})$$

pour désigner les colonnes de M (on comme la numérotation à zéro). On a :

$$P_1(X) = X + a_{0,1}$$

On fait l'opération élémentaire :

$$C_1 \leftarrow C_1 - a_{0,1} C_0$$

On obtient :

$$\Delta = \begin{vmatrix} 1 & x_1 & P_2(x_1) & \cdots & P_{n-1}(x_1) \\ 1 & x_2 & P_2(x_2) & \cdots & P_{n-1}(x_2) \\ \vdots & & & & \vdots \\ \vdots & & & & \vdots \\ 1 & x_n & P_2(x_n) & \cdots & P_{n-1}(x_n) \end{vmatrix}_{[n]}$$

On a :

$$P_2(X) = X^2 + a_{1,2} X + a_{0,2}$$

On fait l'opération élémentaire

$$C_2 \leftarrow C_2 - a_{1,2} C_1 - a_{0,2} C_0$$

Et on obtient cette fois-ci

$$\Delta = \begin{vmatrix} 1 & x_1 & x_1^2 & P_3(x_1) & \cdots & P_{n-1}(x_1) \\ 1 & x_2 & x_2^2 & P_3(x_2) & \cdots & P_{n-1}(x_2) \\ \vdots & & & & \vdots & \\ \vdots & & & & \vdots & \\ 1 & x_n & x_n^2 & P_3(x_{n-1}) & \cdots & P_{n-1}(x_n) \end{vmatrix}_{[n]}$$

Par récurrence (finie), *via* l'opération élémentaire (pour $k \geqslant 1$)

$$L_k \leftarrow L_k - \sum_{j=0}^{k-1} a_{j,k} L_k$$

On obtient donc

$$\Delta = \begin{vmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \cdots & x_2^{n-1} \\ \vdots & & & & \vdots \\ \vdots & & & & \vdots \\ 1 & x_n & x_n^2 & \cdots & x_n^{n-1} \end{vmatrix}_{[n]}$$

D'où

$$\Delta = V(x_1, \cdots, x_n)$$

Il est remarquable que ce déterminant ne dépende en fait pas des $a_{k,j}$. □

Proposition 23.3

Soient $a, b \in \mathbb{K}$, $n \geqslant 2$ et $x_1, \cdots, x_n$. Alors :

* Si $a \neq b$

$$\begin{vmatrix} x_1 & a & \cdots & \cdots & a \\ b & x_2 & a & \cdots & a \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ b & & \ddots & x_{n-1} & a \\ b & \cdots & \cdots & b & x_n \end{vmatrix}_{[n]} = \frac{b\prod_{i=1}^n (x_i - a) - a\prod_{i=1}^n (x_i - b)}{b-a}$$

$*$ Si $a = b$

$$\begin{vmatrix} x_1 & a & \cdots & \cdots & a \\ a & x_2 & a & \cdots & a \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ a & & \ddots & x_{n-1} & a \\ a & \cdots & \cdots & a & x_n \end{vmatrix}_{[n]} = \prod_{i=1}^n (x_i - a) + a\left[\sum_{i=1}^n \prod_{j\neq i} (x_j - a)\right]$$

Démonstration : 1. Supposons $a \neq b$. On note Δ_0 le déterminant recherché. Pour parvenir au calcul de ce déterminant, on introduit $\Delta(x)$ pour $x \in \mathbb{K}$ donné par :

$$\Delta(x) \overset{\text{déf.}}{=} \begin{vmatrix} x_1+x & a+x & \cdots & \cdots & a+x \\ b+x & x_2+x & a+x & \cdots & a+x \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ b+x & & \ddots & x_{n-1}+x & a+x \\ b+x & \cdots & \cdots & b+x & x_n+x \end{vmatrix}_{[n]}$$

Montrons que $\Delta(x) = Ax + B$. Cela impliquera que $\Delta_0 = \Delta(0) = B$. On fait l'opération élémentaire

$$\forall i \in [\![1, n-1]\!], C_i \leftarrow C_i - C_n$$

$$\Delta(x) = \begin{vmatrix} x_1-a & 0 & \cdots & \cdots & a+x \\ b-a & x_2-a & 0 & \cdots & a+x \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ b-a & & b-a & x_{n-1}-a & a+x \\ b-x_n & \cdots & \cdots & b-x_n & x_n+x \end{vmatrix}_{[n]}$$

Un développement par rapport à la dernière colonne permet alors d'affirmer que $\Delta(x)$ est en effet de la forme recherchée :

$$\Delta(x) = Ax + B$$

Il suffit de déterminer explicitement pour $x \neq y$ $\Delta(x)$ et $\Delta(y)$ indépendamment de A et B pour connaître complètement $\Delta(x)$. Dans notre cas, la connaissance de B permet de conclure la démonstration. On le fait ici en $x = -a$ et $y = -b$. En effet,

$$\Delta(-a) = \begin{vmatrix} x_1-a & 0 & \cdots & \cdots & 0 \\ b-a & x_2-a & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ b-a & & \ddots & x_{n-1}-a & 0 \\ b-a & \cdots & \cdots & b-a & x_n-a \end{vmatrix}_{[n]}$$

C'est une belle matrice triangulaire, d'où :

$$\Delta(-a) = \prod_{i=1}^n (x_i - a)$$

De même en $-b$:

$$\Delta(-b) = \begin{vmatrix} x_1-b & a-b & \cdots & \cdots & a-b \\ 0 & x_2-b & a-b & \cdots & a-b \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & & \ddots & x_{n-1}-b & a-b \\ 0 & \cdots & \cdots & 0 & x_n-b \end{vmatrix}_{[n]}$$

D'où

$$\Delta(-b) = \prod_{i=1}^n (x_i - b)$$

Disposant des deux relations

$$\begin{cases} \prod_{i=1}^n (x_i - a) & = & -aA + \Delta_0 \\ \prod_{i=1}^n (x_i - b) & = & -bA + \Delta_0 \end{cases}$$

$b(1) - a(2)$ donne alors

$$\Delta_0 = \frac{b\prod_{i=1}^n (x_i - a) - a\prod_{i=1}^n (x_i - b)}{b-a}$$

2. Supposons que $a = b$. On note Δ le déterminant recherché. On note

$$P(X) \overset{\text{déf.}}{=} \prod_{i=1}^{n}(x_i - X)$$

Et on considère

$$D(X) \overset{\text{déf.}}{=} \frac{XP(a) - aP(X)}{X - a}$$

C'est le déterminant cherché lorsque $a \neq b$. On note

$$Q(X) \overset{\text{déf.}}{=} XP(a) - aP(X)$$

Alors

$$Q(a) = 0$$

Il suit qu'il existe $R \in \mathbb{K}[X]$ tel que

$$Q(X) = (X - a)R(X)$$

donc D est polynomiale ! On pose $D(a) = R(a)$. De plus,

$$D(a) = \Delta$$

Reste à déterminer $R(a)$. Or,

$$Q'(X) = (X - a)R(X) + R'(X)$$

d'où $R(a) = Q'(a)$. D'où encore

$$R(a) = P(a) - aP'(a)$$

Cela permet de conclure que :

$$\Delta = \prod_{i=1}^{n}(x_i - a) + a\left[\sum_{i=1}^{n}\prod_{j\neq i}(x_j - a)\right]$$

3. On peut rédiger plus légèrement cette démonstration **2.** dans le cas où $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$, où on dispose d'une topologie, et de la notion de limite. On note toujours Δ le déterminant recherché. On considère l'application qui donne ce déterminant lorsque $a \neq b$:

$$\forall x \neq a, f(x) \overset{\text{déf.}}{=} \frac{xP(a) - aP(x)}{x - a}$$

avec

$$P(x) \overset{\text{déf.}}{=} \prod_{i=1}^{n}(x_i - x)$$

Il s'agit ici de faire tendre $[x \to a]$ ici, la continuité du déterminant impliquera que

$$f(a) = \Delta$$

L'expression de $f(x)$ fait intervenir un taux d'accroissement : posons

$$Q(x) = xP(a) - aP(x)$$

Alors

$$f(x) = \frac{Q(x) - Q(a)}{x - a}$$

Q est polynomiale, donc on peut faire tendre $[x \to a]$ (ce qui prolonge f par continuité en a), pour avoir

$$f(a) = Q'(a) = P(a) - aP'(a)$$

Ce qui permet de conclure sur l'expression de Δ. □

Proposition 23.4

Soient $n \geqslant 2$, $\alpha_1, \cdots, \alpha_n \in \mathbb{K}$ et $\beta_1, \cdots, \beta_n \in \mathbb{K}$. On note

$$\Delta_{\alpha,\beta} \overset{\text{déf.}}{=} \begin{vmatrix} (\alpha_1+\beta_1)^{n-1} & (\alpha_1+\beta_2)^{n-1} & \cdots & (\alpha_1+\beta_n)^{n-1} \\ (\alpha_2+\beta_1)^{n-1} & (\alpha_2+\beta_2)^{n-1} & \cdots & (\alpha_2+\beta_n)^{n-1} \\ \vdots & \vdots & \cdots & \vdots \\ (\alpha_n+\beta_1)^{n-1} & (\alpha_n+\beta_2)^{n-1} & \cdots & (\alpha_n+\beta_n)^{n-1} \end{vmatrix}$$

Alors

$$\Delta_{\alpha,\beta} = \left[\prod_{i=1}^{n}\binom{n-1}{i}\right](-1)^{\frac{n(n-1)}{2}}\prod_{1\leqslant i<j\leqslant n}(\alpha_i - \alpha_j)(\beta_i - \beta_j)$$

Démonstration : Malgré son aspect complexe, c'est un calcul assez simple quand on pense à la bonne astuce. Ici, l'astuce consiste à voir que lorsque $\alpha, \beta \in \mathbb{K}$,

$$(\alpha + \beta)^{n-1} = \sum_{k=0}^{n-1}\binom{n-1}{k}\alpha^k\beta^{n-1-k}$$

On réindexe, et on va nommer les choses :

$$(\alpha+\beta)^{n-1} = \sum_{k=1}^{n} \binom{n-1}{k-1} \alpha^{k-1}\beta^{n-k}$$

On pose pour $1 \leqslant i, j \leqslant n$:

$$p_{i,k} = \binom{n-1}{k-1}\alpha_i^{k-1} \text{ et } q_{k,j} = \beta_j^{n-k}$$

Enfin, on pose $P = (p_{i,j})$ et $Q = (q_{i,j})$. Alors, avec tout ce qu'on a écrit, on a :

$$M = PQ$$

Calculons alors $\det(P) \overset{=}{\pi}$ et $\det(Q) \overset{\text{déf.}}{=} \chi$.
D'une part,

$$\chi = \begin{vmatrix} \beta_1^{n-1} & \beta_2^{n-1} & \cdots & \beta_n^{n-1} \\ \beta_1^{n-2} & \beta_2^{n-2} & \cdots & \beta_n^{n-2} \\ \vdots & & & \vdots \\ 1 & \cdots & & 1 \end{vmatrix}_{[n]}$$

donc, après permutation (il y en a $\frac{n(n-1)}{2}$),

$$\chi = (-1)^{\frac{n(n-1)}{2}} V(\beta_1, \cdots, \beta_n)$$

Occupons-nous de π. On a :

$$\pi = \begin{vmatrix} \binom{n-1}{0} & \binom{n-1}{1}\alpha_1 & \binom{n-1}{2}\alpha_1^2 & \cdots & \binom{n-1}{n-1}\alpha_1^{n-1} \\ \binom{n-1}{0} & \binom{n-1}{1}\alpha_2 & \binom{n-1}{2}\alpha_2^2 & \cdots & \binom{n-1}{n-1}\alpha_2^{n-1} \\ \vdots & & & & \vdots \\ \binom{n-1}{0} & \binom{n-1}{1}\alpha_n & \binom{n-1}{2}\alpha_n^2 & \cdots & \binom{n-1}{n-1}\alpha_n^{n-1} \end{vmatrix}_{[n]}$$

On peut factoriser à chaque ligne par $\binom{n-1}{i}$, et on voit apparaître le déterminant de Vandermonde :

$$\pi = \prod_{i=1}^{n-1} \binom{n-1}{i} V(\alpha_1, \cdots, \alpha_n)$$

D'où

$$\Delta_{\alpha,\beta} = (-1)^{\frac{n(n-1)}{2}} \prod_{i=1}^{n-1} \binom{n-1}{i} V(\alpha_1, \cdots, \alpha_n) V(\beta_1, \cdots, \beta_n)$$

□

Corollaire 23.1

Soient $p \in \mathbb{N}$ et $n \geqslant p+1$. Alors

* Si $n = p+1$

$$\begin{vmatrix} 1 & 2^p & \cdots & n^p \\ 2^p & 3^p & \cdots & (n+1)^p \\ \vdots & \vdots & & \vdots \\ n^p & (n+1)^p & \cdots & (2n-1)^p \end{vmatrix} = (-1)^{\frac{n(n-1)}{2}} \left[(n-1)!\right]^n$$

* Si $n > p+1$

$$\begin{vmatrix} 1 & 2^p & \cdots & n^p \\ 2^p & 3^p & \cdots & (n+1)^p \\ \vdots & \vdots & & \vdots \\ n^p & (n+1)^p & \cdots & (2n-1)^p \end{vmatrix} = 0$$

Démonstration : • Supposons $p = n-1$. Alors on se retrouve dans le cas de la proposition**23.4**, avec $\alpha_i = i$ et $\beta_j = j-1$. En notant Δ le déterminant recherché,

$$\Delta = (-1)^{\frac{n(n-1)}{2}} \prod_{i=1}^{n-1} \binom{n-1}{i} \prod_{1\leqslant i<j\leqslant n} (j-i)^2$$

Chaque facteur se calcule de la sorte :

$$\prod_{i=1}^{n-1} \binom{n-1}{i} = \frac{[(n-1)!]^2}{\left[\prod_{i=1}^{n} i!\right]^2}$$

et

$$\prod_{1\leqslant i<j\leqslant n} (j-i) = \prod_{j=1}^{n-1} j!$$

D'où

$$\Delta = (-1)^{\frac{n(n-1)}{2}} [(n-1)!]^n$$

• Supposons $n > p+1$. Montrons que la famille concernée par ce déterminant est lieé dans $\mathbb{C}$. On note

$$\forall i \in [\![1, n]\!], P_i(X) = (X+i)^p$$

De sorte que

$$\Delta = \begin{vmatrix} P_1(0) & P_1(1) & \cdots & P_1(n-1) \\ P_2(0) & P_2(1) & \cdots & P_2(n-1) \\ \vdots & & & \vdots \\ P_n(0) & P_n(1) & \cdots & P_n(n-1) \end{vmatrix}_{[n]}$$

La famille $(P_1, \cdots, P_n)$ est une famille de $\mathbb{R}_p[X]$, de dimension $p+1$ composée de n éléments. Cette famille est alors liée. Il existe un polynôme P_j de cette famille et des scalaires $\lambda_1, \cdots, \lambda_n \in \mathbb{R}$ non tous nuls tels que

$$P_j = \sum_{k \neq j} \lambda_k P_k$$

En conséquence,

$$P_j(x) = \sum_{k \neq j} \lambda_k P_k(x)$$

On fait l'opération élémentaire

$$C_j \leftarrow C_j - \sum_{k \neq j} \lambda_k C_k$$

Pour obtenir une colonne nulle dans ce déterminant. D'où $\Delta = 0$.

□

Proposition 23.5

Soient $n \in \mathbb{N}^*$ et $p \leqslant n$. Alors

$$\begin{vmatrix} 1 & \binom{n}{1} & \binom{n}{2} & \cdots & \binom{n}{p} \\ 1 & \binom{n+1}{1} & \binom{n+1}{2} & \cdots & \binom{n+1}{p} \\ \vdots & & & & \vdots \\ 1 & \binom{p+n}{1} & \binom{p+n}{2} & \cdots & \binom{p+n}{p} \end{vmatrix}_{[p+1]} = 1$$

Ce calcul de déterminant fait intervenir la *formule du triangle de* Pascal :

$$\binom{k}{l} + \binom{k}{l+1} = \binom{k+1}{l+1}$$

Qui se démontre simplement en réduisant au même dénominateur les expressions factorielles :

$$\frac{k!}{l!(k-l)!} + \frac{k!}{(l+1)!(k-l-1)!} = \frac{k!(l+1) + k!(k-l)}{(l+1)!(k-l-1)!} = \frac{(k+1)!}{(l+1)!(k-l-1)!}$$

Démonstration : On note Δ_p le déterminant recherché. On cherche une relation de récurrence pour Δ_p. En constatant que pour tout $k \in \mathbb{N}$ et $l \in [\![1, k-1]\!]$,

$$\binom{k}{l} = \binom{k+1}{l+1} - \binom{k}{l+1}$$

On se décide de faire l'opération élémentaire

$$\forall i \in [\![2, p]\!], L_i \leftarrow L_i - L_{i-1}$$

Ce qui donne :

$$\Delta_p = \begin{vmatrix} 1 & \binom{n}{1} & \binom{n}{2} & \cdots & \binom{n}{p} \\ 0 & \binom{n}{0} & \binom{n}{1} & \cdots & \binom{n}{p-1} \\ \vdots & \vdots & & & \vdots \\ 0 & \binom{n+p-1}{0} & \binom{n+p-1}{1} & \cdots & \binom{n+p-1}{p-1} \end{vmatrix}_{[p+1]}$$

On développe par rapport à la première colonne pour finalement avoir

$$\Delta_p = \begin{vmatrix} 1 & \binom{n}{1} & \binom{n}{2} & \cdots & \binom{n}{p-1} \\ 1 & \binom{n+1}{1} & \binom{n+1}{2} & \cdots & \binom{n+1}{p-1} \\ \vdots & \vdots & & & \vdots \\ 1 & \binom{n+p-1}{1} & \binom{n+p-1}{2} & \cdots & \binom{n+p-1}{p-1} \end{vmatrix}_{[p]}$$

Donc $\Delta_p = \Delta_{p-1}$. Finalement, $\Delta_p = \Delta_1 = 1$.

□

Proposition 23.6

Soient $n \geqslant 2$, $a_1, \cdots, a_n \in \mathbb{K}$ et $x_1, \cdots, x_n \in \mathbb{K}$. Alors

$$\begin{vmatrix} a_1+x_1 & a_1 & \cdots & \cdots & a_1 \\ a_2 & a_2+x_2 & & \cdots & a_2 \\ \vdots & & \ddots & & \\ \vdots & & & \ddots & \\ a_n & a_n & \cdots & & a_n+x_n \end{vmatrix}_{[n]} = \prod_{k=1}^{n} x_k + \sum_{j=1}^{n} \left(a_j \prod_{k \neq j} x_k \right)$$

Démonstration : On cherche une relation de récurrence sur Δ_n, le déterminant recherché. On commence par utiliser la linéarité du déterminant grâce à la relation

$$\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n+x_n \end{pmatrix} = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ \vdots \\ x_n \end{pmatrix}$$

Donc Δ_n est alors somme de deux déterminants :

$$\Delta_n = \begin{vmatrix} a_1+x_1 & a_1 & \cdots & \cdots & a_1 \\ a_2 & a_2+x_2 & & \cdots & a_2 \\ \vdots & & \ddots & & \\ \vdots & & & \ddots & \\ a_n & a_n & \cdots & & a_n \end{vmatrix}_{[n]} + \begin{vmatrix} a_1+x_1 & a_1 & \cdots & \cdots & 0 \\ a_2 & a_2+x_2 & & \cdots & 0 \\ \vdots & & \ddots & & \\ \vdots & & & \ddots & \\ a_n & a_n & \cdots & & x_n \end{vmatrix}_{[n]}$$

On note D_1 et D_2 les premier et deuxième termes. Sur D_1, on effectue les opérations élémentaires :

$$\forall i [\![1, n-1]\!], L_i \leftarrow L_i - L_n$$

Ainsi,

$$D_1 = \begin{vmatrix} x_1 & 0 & \cdots & \cdots & a_1 \\ 0 & x_2 & & \cdots & a_2 \\ \vdots & & \ddots & & \\ \vdots & & & \ddots & \\ 0 & 0 & \cdots & & a_n \end{vmatrix}_{[n]}$$

C'est le déterminant d'une matrice triangulaire, d'où

$$D_1 = a_n \prod_{i=1}^{n-1} x_i$$

Quant à D_2, on peut développer par rapport à la dernière colonne :

$$D_2 = x_n \begin{vmatrix} a_1+x_1 & a_1 & \cdots & \cdots & a_1 \\ a_2 & a_2+x_2 & & \cdots & a_2 \\ \vdots & & \ddots & & \\ \vdots & & & \ddots & \\ a_{n-1} & a_{n-1} & \cdots & & a_{n-1}+x_{n-1} \end{vmatrix}_{[n-1]}$$

D'où

$$D_2 = x_n \Delta_{n-1}$$

Finalement, on obtient

$$\Delta_n = a_n \prod_{i=1}^{n-1} x_i + x_n \Delta_{n-1}$$

On conclut par récurrence que

$$\Delta_n = \sum_{j=1}^{n} a_j \prod_{i \neq j} x_i + \prod_{j=1}^{n} x_j$$

□

Proposition 23.7

Soient $n \geqslant 2$, $a_1, \cdots, a_n \in \mathbb{K}$ et $x \in \mathbb{K}$. Alors

$$\begin{vmatrix} x & a_1 & a_2 & \cdots & a_n \\ a_1 & x & a_2 & \cdots & a_n \\ \vdots & & \ddots & & \vdots \\ \vdots & & & \ddots & \vdots \\ a_1 & \cdots & & & x \end{vmatrix}_{[n+1]} = \left(x + \sum_{i=1}^{n} a_i\right) \prod_{i=1}^{n} (x - a_i)$$

Démonstration : On note Δ le déterminant recherché. Ce calcul repose sur deux opérations élémentaires. Le premier est

$$C_{n+1} \leftarrow C_{n+1} + \sum_{j=1}^{n} C_j$$

qui donne

$$\Delta = \begin{vmatrix} x & a_1 & a_2 & \cdots & x + \sum_{j=1}^{n} a_j \\ a_1 & x & a_2 & \cdots & x + \sum_{j=1}^{n} a_j \\ \vdots & & \ddots & & \vdots \\ \vdots & & & \ddots & \vdots \\ a_1 & \cdots & & & x + \sum_{j=1}^{n} a_j \end{vmatrix}_{[n+1]}$$

Donc, on peut factoriser :

$$\Delta = \left(x + \sum_{j=1}^{n} a_j\right) \begin{vmatrix} x & a_1 & a_2 & \cdots & 1 \\ a_1 & x & a_2 & \cdots & 1 \\ \vdots & & \ddots & & \vdots \\ \vdots & & & \ddots & \vdots \\ a_1 & \cdots & & & 1 \end{vmatrix}_{[n+1]}$$

La deuxième opération élémentaire est la suivante :

$$\forall j \in [\![1, n-1]\!], C_j \leftarrow C_j - a_j C_n$$

qui donne

$$\Delta = \left(x + \sum_{j=1}^{n} a_j\right) \begin{vmatrix} x - a_1 & a_1 - a_2 & a_2 - a_3 & \cdots & 1 \\ 0 & x - a_2 & a_2 - a_3 & \cdots & 1 \\ \vdots & & \ddots & & \vdots \\ \vdots & & & \ddots & \vdots \\ 0 & & \cdots & & 1 \end{vmatrix}_{[n+1]}$$

C'est le déterminant d'une matrice triangulaire, d'où le résultat

$$\Delta = \left(x + \sum_{j=1}^{n} a_j\right) \prod_{j=1}^{n} (x - a_j)$$

□

Proposition 23.8 : Déterminant de Cauchy

Soient $n \geqslant 2$, $a_1, \cdots, a_n \in \mathbb{K}$, $b_1, \cdots, b_n \in \mathbb{K}$ tels que $a_i + b_j \neq 0$ pour tout $i, j \in [\![1, n]\!]$. Alors

$$\begin{vmatrix} \frac{1}{a_1+b_1} & \frac{1}{a_1+b_2} & \cdots & \frac{1}{a_1+b_n} \\ \frac{1}{a_2+b_1} & \frac{1}{a_2+b_2} & \cdots & \frac{1}{a_2+b_n} \\ \vdots & & & \vdots \\ \frac{1}{a_n+b_1} & \frac{1}{a_n+b_2} & \cdots & \frac{1}{a_n+b_n} \end{vmatrix} = \frac{\prod_{i<j}(a_j - a_i) \prod_{i<j}(b_j - b_i)}{\prod_{1 \leqslant i,j \leqslant n}(a_i + b_j)}$$

Démonstration : L'idée générale derrière est de former en dernière ligne une combinaison linéaire des $\left(\frac{1}{a_i+b_j}\right)_j$ à la colonne C_i qui s'annule, sauf pour $j = n$ pour trouver une relation de récurrence entre ces déterminants de Cauchy. On note

$$\Gamma_n \overset{\text{déf.}}{=} \begin{vmatrix} \frac{1}{a_1+b_1} & \frac{1}{a_1+b_2} & \cdots & \frac{1}{a_1+b_n} \\ \frac{1}{a_2+b_1} & \frac{1}{a_2+b_2} & \cdots & \frac{1}{a_2+b_n} \\ \vdots & & & \vdots \\ \frac{1}{a_n+b_1} & \frac{1}{a_n+b_2} & \cdots & \frac{1}{a_n+b_n} \end{vmatrix}_n$$

On veut faire les opérations élémentaires

$$L_n \leftarrow \lambda_n L_n$$

(de déterminant $\frac{1}{\lambda_n}$), puis

$$L_n \leftarrow L_n - \sum_{i=1}^{n-1} \lambda_i L_i$$

avec $\lambda_j \in \mathbb{K}$ que l'on va déterminer pour avoir

$$\forall j \in [\![1, n-1]\!], \sum_{i=1}^{n} \frac{\lambda_i}{a_i + b_j} = 0$$

Autrement dit, on cherche $(\lambda_j)_{1 \leqslant j \leqslant n}$ tels que

$$\frac{\prod_{j=1}^{n-1}(X - b_j)}{\prod_{i=1}^{n}(X + a_i)} = \sum_{i=1}^{n} \frac{\lambda_i}{a_i + X}$$

Il suffit pour cela de multiplier par $a_i + X$ et d'évaluer en $X = -a_i$ pour obtenir les expressions suivantes :

$$\lambda_i = \frac{\prod_{j=1}^{n-1}(a_i + b_j)}{\prod_{k \neq i} a_i - a_k} \neq 0$$

Pour assurer la bonne définition de λ_i, supposons que les a_i sont deux à deux distincts. On fait d'abord

$$L_n \leftarrow \lambda_n L_n$$

pour avoir

$$\Gamma_n = \frac{1}{\lambda_n} \begin{vmatrix} \frac{1}{a_1+b_1} & \frac{1}{a_1+b_2} & \cdots & \frac{1}{a_1+b_n} \\ \frac{1}{a_2+b_1} & \frac{1}{a_2+b_2} & \cdots & \frac{1}{a_2+b_n} \\ \vdots & & & \vdots \\ \frac{\lambda_n}{a_n+b_1} & \frac{\lambda_n}{a_n+b_2} & \cdots & \frac{\lambda_n}{a_n+b_n} \end{vmatrix}_n$$

Et on fait l'opération élémentaire souhaitée :

$$L_n \leftarrow L_n - \sum_{i=1}^{n-1} \lambda_i L_i$$

On obtient :

$$\Gamma_n = \frac{1}{\lambda_n} \begin{vmatrix} \frac{1}{a_1+b_1} & \frac{1}{a_1+b_2} & \cdots & \frac{1}{a_1+b_n} \\ \frac{1}{a_2+b_1} & \frac{1}{a_2+b_2} & \cdots & \frac{1}{a_2+b_n} \\ \vdots & & & \vdots \\ 0 & 0 & \cdots & \sum_{i=1}^{n} \frac{\lambda_i}{a_i+b_n} \end{vmatrix}_n$$

On développe par rapport à cette dernière ligne :

$$\Gamma_n = \frac{R(b_n)}{\lambda_n} \begin{vmatrix} \frac{1}{a_1+b_1} & \frac{1}{a_1+b_2} & \cdots & \frac{1}{a_1+b_{n-1}} \\ \frac{1}{a_2+b_1} & \frac{1}{a_2+b_2} & \cdots & \frac{1}{a_2+b_{n-1}} \\ \vdots & & & \vdots \\ \frac{1}{a_{n-1}+b_1} & \frac{1}{a_{n-1}+b_2} & \cdots & \frac{1}{a_{n-1}+b_{n-1}} \end{vmatrix}_{[n-1]}$$

C'est un déterminant de Cauchy d'ordre $n-1$, où on a noté

$$R(b_n) = \sum_{i=1}^{n} \frac{\lambda_i}{a_i + b_n} = \frac{\prod_{j=1}^{n-1}(b_n - b_j)}{\prod_{i=1}^{n}(b_n + a_i)}$$

On a alors

$$\Gamma_n = \frac{\left[\prod_{j=1}^{n-1}(b_n - b_j)\right]\left[\prod_{i=1}^{n-1}(a_n - a_i)\right]}{\left[\prod_{i=1}^{n}(b_n + a_i)\right]\left[\prod_{j=1}^{n}(b_j + a_n)\right]} \Gamma_{n-1}$$

Avec $\Gamma_1 = \frac{1}{a_1+b_1}$, on peut alors conclure sur le déterminant de Cauchy :

$$\Gamma_n = \frac{\prod_{i<j}(a_j - a_i)\prod_{i<j}(b_j - b_i)}{\prod_{1 \leqslant i,j \leqslant n}(a_i + b_j)}$$

L'égalité reste vraie si on ne suppose plus que les a_i sont deux à deux distincts, ce qui donne un déterminant de Cauchy nul. □

Proposition 23.9 : Cofacteurs de la matrice de Vandermonde

Soient $n \geqslant 2$ et $\alpha_1, \cdots, \alpha_n \in \mathbb{K}$. Alors

$$\begin{vmatrix} 1 & \alpha_1 & \cdots & \alpha_1^{k-1} & \alpha_1^{k+1} & \cdots \alpha_1^n \\ 1 & \alpha_2 & \cdots & \alpha_2^{k-1} & \alpha_2^{k+1} & \cdots \alpha_2^n \\ \vdots & & & & & \vdots \\ 1 & \alpha_n & \cdots & \alpha_n^{k-1} & \alpha_n^{k+1} & \cdots \alpha_n^n \end{vmatrix} = \sigma_{n-k}(\alpha_1, \cdots, \alpha_n) \prod_{1 \leqslant i < j \leqslant n} (\alpha_j - \alpha_i)$$

où

$$\sigma_{n-k}(X_1, \ldots, X_n) \overset{\text{déf.}}{=} \sum_{I \in \mathcal{P}_{n-k}([\![1,n]\!])} \left[\prod_{i \in I} X_i\right]$$

Proposition 23.10 : Déterminant circulant

Soient $n \geqslant 2$, et $a_1, \cdots, a_n \in \mathbb{C}$. On considère

$$P(X) \overset{\text{déf.}}{=} a_1 + a_2 X + \cdots + a_n X^{n-1} \in \mathbb{C}[X]$$

En notant $\omega \overset{\text{déf.}}{=} e^{\frac{2\imath\pi}{n}}$, on a alors

$$\begin{vmatrix} a_1 & a_2 & \cdots & a_n \\ a_n & a_1 & \cdots & a_{n-1} \\ \vdots & & & \vdots \\ a_2 & a_3 & \cdots & a_1 \end{vmatrix} = \prod_{j=0}^{n-1} P\left(\omega^j\right)$$

Corollaire 23.2 : Application du déterminant circulant

Soit $n \geqslant 2$ et $\theta \in \mathbb{R}$. Alors

$$\begin{vmatrix} \cos\theta & \cos 2\theta & \cdots & \cos n\theta \\ \cos n\theta & \cos\theta & \cdots & \cos(n-1)\theta \\ \vdots & & & \vdots \\ \cos 2\theta & \cos 3\theta & \cdots & \cos\theta \end{vmatrix} = 2^{n-2} \sin^{n-2}\left(\frac{n\theta}{2}\right)\left[\sin^n\left(\frac{n+2}{2}\theta\right) - \sin^n\left(\frac{n\theta}{2}\right)\right]$$

Proposition 23.11 : Déterminant tridiagonal

Soient $n \geqslant 2$, et $a, b \in \mathbb{K}$.

$$\begin{vmatrix} a+b & ab & & & & \\ 1 & a+b & ab & & (0) & \\ & \ddots & \ddots & \ddots & & \\ & & \ddots & \ddots & \ddots & \\ & (0) & & 1 & a+b & ab \\ & & & & 1 & a+b \end{vmatrix}_{[n]} = \frac{a^{n+1} - b^{n+1}}{a-b}$$

24 Théorème de Cayley-Hamilton

Référence : Les maths en tête, Algèbre, Xavier Gourdon

Théorème 24.1 : Théorème de Cayley-Hamilton

Soient E un $\mathbb{K}$-espace vectoriel de dimension finie et $f \in \mathcal{L}(E)$. Alors

$$\chi_f(f) = 0 \in \mathcal{L}(E)$$

On propose plusieurs démonstrations de ce résultat. La première démonstration utilise le corps de décomposition de χ_f.

Démonstration : On note $M \in \mathcal{M}_n(\mathbb{K})$ la matrice associée à f dans la base canonique de $\mathbb{K}^n$. On considère $\mathbb{L}$ le corps de décomposition de χ_M dans $\mathbb{K}$. On considère alors que $M \in \mathcal{M}_n(\mathbb{L})$. χ_M est scindé dans $\mathbb{L}$, donc M se trigonalise dans ce corps : il existe $P \in \mathrm{GL}_n(\mathbb{L})$ et $T \in \mathcal{M}_n(\mathbb{L})$ telles que

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

Si bien que

$$\chi_M(X) = (-1)^n \prod_{i=1}^{n} (X - t_{i,i})$$

On considère $(\ell_1, \cdots, \ell^n)$ la base canonique de $\mathbb{L}$. On pose pour $k \in [\![1, n]\!]$:

$$P_k(X) \overset{\text{déf.}}{=} \prod_{i=1}^{k} (X - t_{i,i})$$

Montrons que

$$\forall k \in [\![1, n]\!], \forall j \in [\![1, k]\!], [P_k(T)]\, \ell_j = 0$$

On le montre par récurrence. Rappelons que

$$T \overset{\text{déf.}}{=} \begin{pmatrix} t_{1,1} & & & \\ & t_{2,2} & t_{i,j} & \\ & 0 & \ddots & \\ 0 & \cdots & & t_{n,n} \end{pmatrix}$$

- Pour $k = 1$, on a

$$T\ell_1 = t_{1,1}\ell_1$$

Donc

$$[P_1(T)]\ell_1 = (T - t_{1,1}I_n)\ell_1 = 0$$

- Soit $k \in [\![1, n-1]\!]$. Supposons que l'on a

$$\forall j \in [\![1, k]\!], [P_k(T)]\ell_j = 0$$

Alors pour $j \in [\![1, k]\!]$,

$$[P_{k+1}(T)]\ell_j = [T - t_{k+1,k+1}I_n][P_k(T)]\ell_j = 0$$

par hypothèse de récurrence. Pour $j = k+1$,

$$[P_{k+1}(T)]\ell_{k+1} = [P_k(T)] \left[\sum_{j=1}^{k} t_{j,k+1}\ell_j \right]$$

Par linéarité,

$$[P_{k+1}(T)]\ell_{k+1} = \sum_{j=1}^{k} t_{j,k+1}[P_k(T)]\ell_j$$

On conclut enfin par hypothèse de récurrence :

$$[P_{k+1}(T)]\ell_{k+1} = 0$$

Cette petite récurrence permet alors conclure à

$$\forall j \in [\![1, n]\!], P_n(T)\ell_j = 0$$

Donc $P_n(T) = 0$. Or, $P_n = (-1)^n \chi_M$. On a alors montré que

$$\chi_f(f) = 0$$

□

La deuxième démonstration n'utilise pas le corps de décomposition, et s'appuie sur les matrices compagnons.

Lemme 24.1 : Polynôme caractéristique d'une matrice compagnon

Soient $P \in \mathbb{K}[X]$ et $n \geqslant 1$:

$$P(X) \overset{\text{déf.}}{=} X^n + \sum_{k=0}^{n-1} a_k X^k$$

On note

$$C_P \overset{\text{déf.}}{=} \begin{pmatrix} 0 & 0 & & & -a_0 \\ 1 & 0 & & & -a_1 \\ 0 & 1 & & & -a_2 \\ \vdots & & \ddots & & \vdots \\ 0 & \cdots & & 1 & -a_{n-1} \end{pmatrix} \in \mathcal{M}_n(\mathbb{K})$$

Alors le polynôme caractéristique de C_P est $(-1)^n P$ (ce signe vient de la convention choisie pour le polynôme caractéristique).

Démonstration du lemme : On fait une démonstration par récurrence sur $n \geqslant 1$.

- Pour $n = 1$, $P(X) = X - a$ avec $a \in \mathbb{K}$ et

$$C_P = (a)$$

qui est effectivement de polynôme caractéristique $a - X$. Le résultat est alors vrai pour $n = 1$.

- Soit $n \geqslant 1$. Supposons le résultat vrai pour tout polynôme de degré n unitaire. On va faire des combinaisons linéaires pour se ramener à ce cas-là. Soit

$$P \overset{\text{déf.}}{=} X^{n+1} + \sum_{k=0}^{n} a_k X^k$$

On note dans cette démonstration Δ le polynôme caractéristique de C_P. Par définition,

$$\Delta = \begin{vmatrix} -X & 0 & & \cdots & 0 & -a_0 \\ 1 & -X & & & & -a_1 \\ 0 & 1 & -X & & & -a_2 \\ \vdots & & \ddots & \ddots & & \vdots \\ \vdots & & & \ddots & \ddots & \vdots \\ 0 & & \cdots & & 1 & -X - a_n \end{vmatrix}_{[n+1]}$$

On fait un petit développement par rapport à la première ligne :

$$\Delta = -X \begin{vmatrix} -X & 0 & & \cdots & 0 & -a_1 \\ 1 & -X & & & & -a_2 \\ 0 & 1 & -X & & & -a_3 \\ \vdots & & \ddots & \ddots & & \vdots \\ \vdots & & & \ddots & \ddots & \vdots \\ 0 & & \cdots & & 1 & -X - a_n \end{vmatrix}_{[n]} + (-1)^n(-a_0) \begin{vmatrix} 1 & -X & & & \\ 0 & 1 & -X & & \\ \vdots & & \ddots & \ddots & \\ \vdots & & & \ddots & \ddots \\ 0 & & \cdots & & 1 \end{vmatrix}_{[n]}$$

Le premier déterminant est la matrice compagnon du polynôme

$$Q = X^n + \sum_{k=0}^{n-1} a_k X^k$$

Par hypothèse de récurrence, ce déterminant vaut donc $(-1)^n Q$. Le deuxième déterminant vaut 1. Ainsi,

$$\Delta = (-1)^{n+1} X Q + (-1)^{n+1} a_0$$

D'où $\Delta = (-1)^{n+1} P$, donc on a prouvé le résultat. □

Remarque : Une autre option pour montrer ce lemme est de faire directement l'opération élémentaire

$$L_1 \leftarrow L_1 + X L_2 + \cdots + X^{n-1} L_n$$

On obtient :

$$\chi_{C_P}(X) = \begin{vmatrix} 0 & 0 & & & -P(X) \\ 1 & 0 & & & -a_1 \\ 0 & 1 & & & -a_2 \\ \vdots & & \ddots & & \vdots \\ 0 & \cdots & & 1 & -a_{n-1} \end{vmatrix}_{[n]}$$

Un développement sur la première ligne conclut.

Démonstration : Soit $x \in E$ non nul. Il existe $p \leqslant n = \dim E$ pour lequel la famille

$$\mathcal{F}_x \overset{\text{déf.}}{=} (x, f(x), \cdots, f^p(x))$$

est libre, et

$$(x, f(x), \cdots, f^p(x), f^{p+1}(x))$$

est liée. Il existe $a_0, \cdots, a_p \in \mathbb{K}$ tels que

$$f^{p+1}(x) + \sum_{j=0}^{p} a_j f^j(x) = 0$$

On note alors

$$P(X) \overset{\text{déf.}}{=} X^{p+1} + \sum_{j=0}^{p} a_j X^j$$

Complétons $\mathcal{F}_x$ et $\mathcal{B}_x$ base de E. Dans cette base, f s'écrit matriciellement par blocs

$$\text{Mat}_{\mathcal{B}_x}(f) = \begin{pmatrix} A & B \\ 0 & C \end{pmatrix}$$

avec

$$A = \begin{pmatrix} 0 & 0 & & & -a_0 \\ 1 & 0 & & & -a_1 \\ 0 & 1 & & & -a_2 \\ \vdots & & \ddots & & \vdots \\ 0 & \cdots & & 1 & -a_p \end{pmatrix} = C_P$$

Il suit (déterminant triangulaire par blocs) :

$$\chi_f = \chi_{C_P} \chi_B$$

Ainsi, par le lemme, il suit

$$[\chi_f(f)](x) = \chi_B \circ P(f)(x)$$

et $[P(f)](x) = 0$ par construction de P. Ainsi,

$$\chi_f(f)(x) = 0$$

Par conséquent, puisque cela est vrai pour tout $x \in E$, il suit que $\chi_f(f) = 0$. □

25 Décomposition de Dunford

Référence : Les maths en tête, Algèbre, Xavier Gourdon

Théorème 25.1 : Théorème de décomposition de Dunford

Soit E un $\mathbb{K}$-espace vectoriel de dimension finie. Soit $f \in \mathcal{L}(E)$ tel que son polynôme caractéristique soit scindé dans $\mathbb{K}$. Alors il existe deux uniques éléments $d, n \in \mathcal{L}(E)$ tels que :

(i) $f = d + n$

(ii) d est diagonalisable sur $\mathbb{K}$

(iii) n est nilpotente

(iv) $d \circ n = n \circ d$

De plus, $u, n \in \mathbb{K}[f]$ s'expriment comme polynômes en f.

Lemme 25.1 : Lemme de décomposition des noyaux

Soit E un $\mathbb{K}$-espace vectoriel de dimension quelconque. Soit $f \in \mathcal{L}(E)$. Enfin, soient $P_1, \cdots, P_s \in \mathbb{K}[X]$ mutuellement premiers entre eux, on note P leur produit :

$$P \overset{\text{déf.}}{=} \prod_{i=1}^{s} P_i$$

Alors :

(i) On dispose de la décomposition suivante :

$$\ker(P(f)) = \bigoplus_{i=1}^{s} \ker\left[P_i(f)\right]$$

(ii) Si $p_i : \ker(P(f)) \longrightarrow \ker\left[P_i(f)\right]$ est la projection sur $\ker\left[P_i(f)^{\alpha_i}\right]$ parallèlement à $\bigoplus_{j\neq i} \ker\left[P_j(f)\right]$, alors p_i est un polynôme en f.

Démonstration du lemme : (i) Le résultat est vrai pour $s = 1$, la décomposition est une décomposition triviale. Par souci de compréhension, on montre par récurrence sur $s \geqslant 2$ la décomposition en somme directe.

• Pour $s = 2$, on se donne $A, B \in \mathbb{K}[X]$ deux polynômes premiers entre eux, et on considère $P = AB$. Montrons que

$$\ker(P(f)) = \ker(A(f)) \oplus \ker(B(f))$$

Puisque A, B sont premiers entre eux, on dispose d'une relation de Bézout entre eux : il existe $U, V \in \mathbb{K}[X]$ tels que :

$$AU + BV \overset{(*)}{=} 1$$

Montrons alors que :

* $\ker(A(f)) \cap \ker(B(f)) = \{0\}$;
* $\ker(P(f)) = \ker(A(f)) + \ker(B(f))$

Soit $x \in \ker(A(f)) \cap \ker(B(f))$. D'après $(*)$, qui n'est rien d'autre que l'égalité de Bézout, on a en évaluant en f :

$$A(f)U(f) + B(f)V(f) = \mathrm{id}_E$$

D'où

$$x = [A(f) \circ U(f)](x) + [B(f) \circ V(f)](x)$$

Or, tous les $\Xi(f)$ considérés sont des polynômes en f, donc commutent :

$$x = [U(f) \circ A(f)](x) + [V(f) \circ B(f)](x)$$

Puisque $A(f)(x) = B(f)(x) = 0$, il suit que $x = 0$, donc l'intersection des deux espaces est bien réduit à $\{0\}$.

Soit $x \in \ker(P(f))$. On dispose toujours de la décomposition :

$$x = [U(f) \circ A(f)](x) + [V(f) \circ B(f)](x)$$

Alors,

$$A(f) \circ \left[V(f) \circ B(f)\right](x) = [AVB(f)](x)$$

Soit encore, puisque $P = AB$:

$$A(f) \circ \left[V(f) \circ B(f)\right](x) = VP(f)(x)$$

Puisque $x \in \ker(P(f))$, il suit

$$A(f) \circ \left[V(f) \circ B(f)\right](x) = 0$$

donc $VB(f)(x) \in \ker(A(f))$. De même, $UA(f)(x) \in \ker(B(f))$. Finalement, le théorème de Bézout nous a directement amené à l'écriture souhaitée :

$$\ker(P(f)) \subset \ker(A(f)) + \ker(B(f))$$

L'inclusion réciproque est vraie : si $x = x_a + x_b$ avec $A(f)(x_a) = 0$ et $B(f)(x_b) = 0$ alors

$$P(f)(x) = P(f)(x_a) + P(f)(x_b)$$

Donc

$$P(f)(x) = B(f) \circ A(f)(x_a) + A(f) \circ B(f)(x_b) = 0$$

On a donc prouvé le lemme des noyaux pour $s = 2$. Notons que les projecteurs p_i de l'énoncé sont donnés ici par

$$p_A = B(f)V(f) \text{ et } p_B = A(f)U(f)$$

donc les projecteurs sont effectivement polynômiaux en f !

• Soit $s \geqslant 2$, et supposons la décomposition en somme directe vérifiée pour tout produit de polynômes premiers entre eux avec s facteurs. On se donne $P_1, \cdots, P_{s+1} \in \mathbb{K}[X]$ mutuellement premiers entre eux, et on note

$$P = \prod_{i=1}^{s+1} P_i$$

Par hypothèse de récurrence :

$$\ker\left(\prod_{i=1}^{s} P_i(f)\right) = \bigoplus_{i=1}^{s} \ker(P_i(f))$$

Et d'après l'initialisation, puisque P_{s+1} et $\prod_{i=1}^{s} P_i$ sont premiers entre eux :

$$\ker(P(f)) = \ker\left(\prod_{i=1}^{s} P_i(f)\right) \oplus \ker(P_{s+1}(f))$$

D'où

$$\ker(P(f)) = \bigoplus_{i=1}^{s+1} \ker(P_i(f))$$

Ce qui conclut l'hérédité pour la décomposition.

(ii) A l'instar de la petite remarque sur l'initialisation de la récurrence, c'est l'égalité de Bézout qui donnera l'expression des projecteurs considérés. On note $P = P_1 \cdots P_s$, avec les P_i mutuellement premiers entre eux. On considère les polynômes

$$Q_i \overset{\text{déf.}}{=} \prod_{j \neq i} P_j$$

Les Q_i sont mutuellement premiers entre eux (pas de facteurs commun), et c'est à eux que nous allons appliquer le théorème de Bézout :

$$1 = \sum_{i=1}^{s} U_i Q_i$$

On définit alors sur $\ker(P(f))$:

$$p_i \overset{\text{déf.}}{=} U_i(f) Q_i(f)$$

Constatons que $p_i \in \mathbb{K}[f]$. Montrons que les p_i sont les projecteurs concernés par ce lemme. Constatons que

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

Fixons $i \in [\![1, s]\!]$. Montrons que :

* $p_i \circ p_i = p_i$;
* Pour $j \neq i$, $p_i \circ p_j = 0$;
* $\mathrm{Im}(p_i) = \ker(P_i(f))$;
* $\ker(p_i) = \bigoplus_{j \neq i} \ker(P_j(f))$.

• Soit $x \in \ker(P(f))$. D'après la relation de Bézout, on a :

$$p_i(x) = \sum_{j=1}^{s} p_i \circ p_j(x)$$

avec pour tout $j \in [\![1, s]\!]$:

$$p_i \circ p_j(x) = Q_i Q_j(f) \circ U_i \circ U_j(f)$$

Or, $Q_i Q_j$ comporte P parmi ses facteurs lorsque $i \neq j$, donc, puisque $P(f)(x) = 0$:

$$p_i \circ p_j(x) = 0$$

lorsque $i \neq j$. De plus, on en déduit

$$p_i(x) = p_i \circ p_i(x)$$

• Soit $x \in \ker(P_i(f))$. D'après l'écriture de Bézout :

$$x = \sum_{j=1}^{s} \left[U_j(f) \prod_{k \neq j} P_k(f)\right](x)$$

Puisque $P_i(f)(x) = 0$, tous les termes concernant P_i sont nuls, d'où

$$x = \left[U_i(f) \prod_{k \neq i} P_k(f)\right](x) = U_i Q_i(f)(x)$$

d'où $x = p_i(x)$. Donc $\ker(P_i(f)) \subset \mathrm{Im}(p_i)$. Réciproquement, si $y = p_i(x)$ avec $x \in \ker(P(f))$, alors

$$[P_i(f)](y) = \left[\left(U_i \prod_{j \neq i} P_j P_i\right)\right](f)(x)$$

Donc,

$$[P_i(f)](y) = U_i P(f)(x) = 0$$

d'où $\mathrm{Im}(p_i) = \ker(P_i(f))$.

• Soit $x \in \ker(P_j(f))$, avec $j \neq i$. Alors :

$$p_i(x) = U_i(x) Q_i(x)$$

avec $i \neq j$, $Q_i(x) = 0$, car il contient le terme P_j. D'où $p_i(x) = 0$. Par linéarité,

$$\bigoplus_{j \neq i} \ker(P_j(f)) \subset \ker(p_i)$$

Réciproquement, si $p_i(x) = 0$, avec $x \in \ker(P(f))$ alors d'après la décomposition de Bézout,

$$x = \sum_{j=1}^{s} U_j Q_j(f)(x)$$

Puisque $p_i(x) = U_i Q_i(f)(x) = 0$, il suit

$$x = \sum_{\substack{j=1 \\ j \neq i}}^{s} U_j Q_j(f)(x)$$

donc $x \in \bigoplus_{j \neq i} \ker(P_j(f))$. □

Passons à la démonstration la plus classique de ce théorème de décomposition.

Démonstration : Montrons l'existence d'un etelle décomposition. D'après le théorème de Cayley-Hamilton,

$$\chi_f(f) = 0$$

On décompose χ_f en ses facteurs premiers dans $\mathbb{K}$, scindé par hypothèse :

$$\chi_f = \prod_{i=1}^{s} (X - \lambda_i)^{m_i}$$

Puisque $\ker(\chi_f(f)) = E$, le lemme des noyaux nous fournit la décomposition suivante :

$$E = \bigoplus_{i=1}^{s} \ker\left((f - \lambda_i \mathrm{id}_E)^{m_i}\right)$$

dont chaque projecteur provenant de ce système de projecteurs est polynomial en f, on les note p_i. On note Γ_i les sous-espaces caractéristiques, qui apparaissent dans cette décomposition. On pose alors

$$d \overset{\text{déf.}}{=} \sum_{i=1}^{s} \lambda_i p_i \text{ et } n \overset{\text{déf.}}{=} f - d$$

Montrons les différentes propriétés.

(i) Par contruction, $f = d + n$.

(ii) d est diagonalisable dans $\mathbb{K}$: dans une base adaptée à la décomposition du lemme des noyaux, d admet en effet une matrice diagonale.

(iii) Montrons que n est nilpotente. On note $m = \max_{1 \leqslant i \leqslant s} m_i$. Alors pour tout $i \in [\![1, s]\!]$, pour $x \in \Gamma_i$:

$$n^m(x) = (f - \lambda_i p_i)^m (x) = 0$$

Donc sur E, $n^m = 0$, n est donc nilpotente.

(iv) Par construction, p_i est un polynôme en f, donc commute avec f. Ainsi, d est un polynôme en f, tout comme n, et donc commutent entre eux.

▷ Montrons l'unicité de la décomposition de Dunford. Soient d' diagonalisable, n' nilpotente deux autres endomorphismes tels que $d + n = d' + n' = f$ et $d'n' = n'd'$. Alors d' commute avec f, donc avec tout polynôme de f, donc avec d : $dd' = d'd$. Ce sont donc deux endomorphismes diagonalisables qui commutent : il existe une base commune de diagonalisation dans laquelle $d - d'$ est diagonalisable. Quand à n et n', ils commutent pour la même raison que d', et par le binôme :

$$(n - n')^{q+q'} = 0$$

avec q et q' les indices de nilpotence. Ainsi, on a

$$d - d' = n' - n$$

qui est un endomorphisme diagonalisable nilpotent. Il est alors nul. D'où l'unicité. □

26 Centre d'un p-groupe - Orbites et stabilisateurs - Théorème de Cayley

Référence : Cours d'algèbre, Daniel Perrin

Cette section est un prétexte pour réviser certaines notions d'action de groupes : orbite, stabilisateur, équations aux classes, etc

Définition 26.1

Soient G un groupe et X un ensemble. On dit que G *agit* sur X par une action ρ si $\rho : G \times X \longrightarrow X$ est une application qui vérifie :

(i) Pour tout $g, h \in G$ et $x \in X$:

$$\rho(gh, x) = \rho(g, \rho(h, x))$$

(ii) Pour tout $x \in X$, $\rho(1_G, x) = x$.

On note lorsque l'action ρ est explicitée :

$$\forall g \in G, \forall x \in X, \rho(g, x) = g \cdot x$$

On définit aussi une action *via* le morphisme de groupe

$$\varphi : \left(\begin{array}{ccc} G & \longrightarrow & \mathfrak{S}(X) \\ g & \longmapsto & \varphi(g) : \left(\begin{array}{ccc} X & \longrightarrow & X \\ x & \longmapsto & \rho(g,x) = g \cdot x \end{array} \right) \end{array} \right)$$

Définition 26.2

Soient X un ensemble et G un groupe qui agit sur X. On définit le *stabilisateur* de $x \in X$ pour cette action :

$$\mathsf{Stab}(x) = G_x \overset{\text{déf.}}{=} \{g \in G \,|\, g \cdot x = x\} < G$$

Pour tout $x \in X$, on définit son *orbite* :

$$\mathsf{Orb}(x) = \omega(x) = G \cdot x \overset{\text{déf.}}{=} \{y \in X \,|\, \exists g \in G, y = g \cdot x\} \subset X$$

Proposition 26.1 : Relation orbite-stabilisateur

Soit X un ensemble, et G un groupe fini qui agit sur X. Alors

$$\forall x \in X, \sharp\mathsf{Orb}(x) \cdot |\mathsf{Stab}(x)| = |G|$$

Démonstration : Soit $x \in X$. On définit l'application

$$\Phi_x : \left(\begin{array}{ccc} {}^{G}/_{\mathsf{Stab}(x)} & \longrightarrow & X \\ g \cdot \mathsf{Stab}(x) & \longmapsto & g \cdot x \end{array} \right)$$

Montrons que :

* Cette application est bien définie ;
* c'est une bijection entre l'ensemble ${}^{G}/_{\mathsf{Stab}(x)}$ et $\omega(x)$.

• Montrer que Φ_x est bien définie signifie qu'on cherche à montrer que si $\bar{g} = \bar{g'}$ alors $\Phi_x(\bar{g}) = \Phi_x(\bar{g'})$. Soient $g, g' \in G$ tels que

$$g\mathsf{Stab}(x) = g'\mathsf{Stab}(x) < G$$

Puisque $g \in g'\mathsf{Stab}(x)$, il existe alors $h \in \mathsf{Stab}(x)$ tel que $g = g'h$. Il suit alors que

$$g \cdot x = (g'h) \cdot x$$

Par propriété d'action de groupe :

$$g \cdot x = g'(h \cdot x)$$

Enfin, puisque $h \in \mathsf{Stab}(x)$, il vient que

$$g \cdot x = g' \cdot x$$

donc $\Phi_x(\bar{g}) = \Phi_x(\bar{g'})$, Φ_x est bien définie.

• Constatons d'abord que

$$\Phi_x\left({}^G/_{\mathsf{Stab}(x)}\right) \subset \omega(x)$$

Puis, pour l'inclusion inverse, si $y \in \omega(x)$, il existe $g \in G$ tel que $y = g \cdot x$. Ainsi, $y = \Phi_x(g\mathsf{Stab}(x))$, Φ_x est donc bien surjective dans $\omega(x)$.

Montrons que Φ_x est injective. Soient $g, g' \in G$ tels que $\Phi_x(\bar{g}) = \Phi_x(\bar{g'})$. Alors

$$g \cdot x = g' \cdot x$$

donc

$$g^{-1}g' \cdot x = g^{-1}g \cdot x$$

donc par propriété d'action de groupes

$$g^{-1}g' \cdot x = x$$

donc $g^{-1}g' \in \mathsf{Stab}(x)$, soit encore $g' \in g\mathsf{Stab}(x)$. Par conséquent,

$$g\mathsf{Stab}(x) = g'\mathsf{Stab}(x)$$

(on dispose de l'inclusion indirecte, et la symétrie du problème donne l'autre). Φ_x est donc bien injective. Ainsi, Φ_x est bien bijective.

G étant fini, on peut conclure en l'égalité des cardinaux :

$$\left|{}^G/_{\mathsf{Stab}(x)}\right| = \sharp\omega(x)$$

□

Remarque : Si G est d'ordre infini, l'application proposée dans la démonstration reste une bijection entre ${}^G/_{\mathrm{Stab}(x)}$ et $\omega(x)$

Théorème 26.1 : Théorème de CAYLEY

Soit G un groupe fini de cardinal n. Alors G est isomorphe à un sous-groupe de $\mathfrak{S}_n$.

Démonstration : On fait agir G sur lui-même *via* la translation à gauche :

$$\left(\begin{array}{ccc} G\times G & \longrightarrow & G \\ (g,x) & \longmapsto & gx \end{array}\right)$$

On définit alors le morphisme de groupes :

$$\varphi : \left(\begin{array}{ccc} G & \longrightarrow & \mathfrak{S}(G) \\ g & \longmapsto & \varphi(g) : \left(\begin{array}{ccc} G & \longrightarrow & G \\ x & \longmapsto & gx \end{array}\right) \end{array}\right)$$

Montrons que φ est injective. Puisque φ est un morphisme de groupes, il suffit de montrer que $\ker(\varphi) = \{1_G\}$. Soit $g \in \ker(\varphi)$. Alors pour tout $x \in G$, $gx = 1_G$. C'est en particulier ke cas pour $x = 1_G$, où on obtient donc $g = 1_G$.

Puisque φ est injective, l'application

$$\varphi : \left(\begin{array}{ccc} G & \longrightarrow & \mathsf{Im}(\varphi) \\ g & \longmapsto & \varphi(g) \end{array}\right)$$

est un isomorphisme de groupes. De plus, $\mathsf{Im}(\varphi)$ est un sous-groupe de $\mathfrak{S}(G)$, isomorphe à $\mathfrak{S}_n$ en notant $G = \{g_1, \cdots, g_n\}$. Il suit alors que G est isomorphe à un sous-groupe de $\mathfrak{S}_n$. □

Proposition 26.2 : Équation aux classes

Soient p un nombre premier et G un p-groupe (*id est* un groupe fini de cardinal p^α avec $\alpha \geqslant 1$). Soit X un ensemble sur lequel G agit. On note

$$X^G \overset{\text{déf.}}{=} \{x \in X \,|\forall g \in G, g \cdot x = x\} \subset X$$

les points fixes de l'action. Alors

$$\sharp X \equiv \sharp X^G \ [p]$$

Démonstration : On introduit la relation d'équivalence sur X suivante :

$$x \mathrel{\mathcal{R}} y \iff \exists g \in G, y = g \cdot x$$

Les deux définitions d'action de groupe donnent directement la réflexivité et la transitivité. Quand à la symétrie, si $y = g \cdot x$ alors

$$g^{-1} \cdot y = g^{-1} \cdot (g \cdot x) = g^{-1}g \cdot x$$

d'où $g^{-1} \cdot y = x$, donc cette relation est symétrique. Il s'agit donc d'une relation d'équivalence qui nous fournit alors une partition de X en ses différentes classes :

$$X = \bigsqcup_{i \in I} \omega(x_i)$$

où x_i est un représentant de $\omega(x_i)$. De plus, pour $x \in X$:

$$x \in X^G \iff \omega(x) = \{x\}$$

Par conséquent, si $x \notin X^G$, $\sharp\omega(x) > 1$. D'après la relation orbite-stabilisateur, $\sharp\omega(x)$ divise $|G|$. Ainsi,

$$\forall x \notin X^G, \sharp\omega(x) \equiv 0 \ [p]$$

Puisque

$$X = \left(\bigsqcup_{\substack{i \in I \\ x_i \in X^G}} \omega(x_i) \right) \sqcup \left(\bigsqcup_{\substack{i \in I \\ x_i \notin X^G}} \omega(x_i) \right)$$

Alors

$$\sharp X = \sum_{\substack{i \in I \\ x_i \in X^G}} \sharp\omega(x_i) + \sum_{\substack{i \in I \\ x_i \notin X^G}} \sharp\omega(x_i)$$

Pour $x_i \in X^G$, $\sharp\omega(x_i) = 1$, et

$$\left\{ i \in I, x_i \in X^G \right\} \simeq X^G$$

D'où finalement,

$$\sharp X \equiv \sharp X^G \ [p]$$

□

Proposition 26.3 : Centre d'un p-groupe

Soient p un nombre premier et G un p-groupe. Alors le centre de G défini par :

$$\mathfrak{Z}(G) \overset{\text{déf.}}{=} \{x \in G \,|\forall g \in G, xg = gx\}$$

n'est pas le sous-groupe trivial.

Démonstration : On fait agir G sur lui-même *via* un automorphisme intérieur :

$$\varphi(g) : \left(\begin{array}{ccc} G & \longrightarrow & G \\ x & \longmapsto & gxg^{-1} \end{array} \right)$$

Alors, pour cette action, l'ensemble des points fixe de l'action (X^G de la proposition **26.2**) n'est rien d'autre que $\mathfrak{Z}(G)$. D'après l'équation aux classes,

$$|\mathfrak{Z}(G)| \equiv |G| \ [p]$$

Et puisque G est un p-groupe,

$$|\mathfrak{Z}(G)| \equiv 0 \ [p]$$

Or, $1_G \in \mathfrak{Z}(G)$, donc $\mathfrak{Z}(G) \neq \varnothing$, donc $|\mathfrak{Z}(G)| \geqslant 1$. C'est donc un multiple de p non nul, donc $\mathfrak{Z}(G)$ n'est pas le sous-groupe trivial ! □

27 Théorèmes de Sylow

Référence : Cours d'algèbre, Daniel PERRIN

Définition 27.1

Soit G un groupe fini d'ordre $|G| = p^\alpha m$ avec m premier avec p. $P < G$ est un *sous-groupe de SYLOW* de G si $|P| = p^\alpha$.

Théorème 27.1 : Premier théorème de SYLOW

Soit G un groupe fini de cardinal $p^\alpha m$ avec p premier avec m. Alors G admet un p-sous-groupe de SYLOW.

Lemme 27.1 : Sous-groupes et Sylow

Soit G un groupe fini de cardinal $p^\alpha m$ avec p ne divisant pas m. Soit H un sous-groupe de G. Soit S un p-sous-groupe de Sylow. Alors il existe $a \in G$ tel que

$$aSa^{-1} \cap H$$

soit un p-Sylow de H.

Démonstration du lemme : G agit sur ${}^G/_S$ par translation à gauche :

$$\begin{pmatrix} G & \longrightarrow & {}^G/_S \\ g & \longmapsto & gS \end{pmatrix}$$

Alors, pour tout $a \in G$, le stabilisateur de aS pour cette action est

$$\mathsf{Stab}(aS) = \{g \in G, gaS = aS\} = aSa^{-1}$$

(La deuxième égalité se montre par double inclusion, étape que l'on peut sauter pour le développement). Par restriction, H agit sur ${}^G/_S$ aussi par translation à gauche. Le stabilisateur de aS pour cette action n'est autre que $aSa^{-1} \cap H$, qui est donc bien un sous-groupe de H, et est un p-sous-groupe de G.

Montrons qu'il existe $a \in G$ tel que $aSa^{-1} \cap H$ soit un p-sous-groupe de Sylow de H. Or, d'après la relation orbite-stabilisateur :

$$\left|{}^H/_{aSa^{-1}\cap H}\right| = \sharp\omega_H(aS)$$

où $\omega_H(aS)$ désigne l'orbite de aS pour l'action de translation dans H. L'orbite pour la translation de G, $\omega(aS)$ est lui multiple de $\sharp\omega_H(aS)$, et vérifie la partition en orbites :

$${}^G/_S = \bigsqcup_{i\in I} \omega(a_i S)$$

Supposons que pour tout $a \in S$, $aSa^{-1} \cap H$ ne soit pas un p-sous-groupe de Sylow de H. Alors $\sharp\omega_H(aS)$ est multiple de p, donc $\sharp\omega(aS)$ aussi. Or, par la mise en partition,

$$\left|{}^G/_S\right| = \sum_{i\in I} \sharp\omega(a_i S)$$

donc est multiple de p. Or, puisque S est un p-Sylow de G,

$$\left|{}^G/_S\right| = m$$

avec p ne divisant pas m. D'où la contradiction recherchée, et donc H admet bien un p-sous-groupe de Sylow.

□

Démonstration : **1.** Dans le cas où $G = \mathsf{GL}_n(\mathbb{F}_p)$, où p est premier, on a

$$|\mathsf{GL}_n(\mathbb{F}_p)| = \prod_{j=0}^{n-1} \left[p^n - p^j\right]$$

donc

$$|\mathsf{GL}_n(\mathbb{F}_p)| = p^{\frac{n(n-1)}{2}} m$$

avec p ne divisant pas m. Le sous-groupe des matrices triangulaires supérieures strictes est alors un p-sous-groupe de Sylow de $\mathsf{GL}_n(\mathbb{F}_p)$:

$$P \overset{\text{déf.}}{=} \left\{ \begin{pmatrix} 1 & & (a_{i,j}) & \\ 0 & 1 & & \\ \vdots & & \ddots & \\ 0 & \cdots & 0 & 1 \end{pmatrix} \middle| a_{i,j} \in \mathbb{F}_p \right\}$$

En effet, les coefficients $a_{i,j}$ quelconques sont au nombre de $p \times p^2 \cdot \dots \cdot p^{n-1}$, soit $p^{\frac{n(n-1)}{2}}$.

2. Revennons au cas général. Si G est un groupe de carnial $p^\alpha m$ avec p ne divisant pas m, le théorème de Cayley permet d'affirmer qu'il existe un morphisme injectif de G dans $\mathfrak{S}_n$, où $n = |G|$. G est donc isomorphe à un sous-groupe de $\mathfrak{S}_n$. Or, l'application

$$\begin{pmatrix} \mathfrak{S}_n & \longrightarrow & \mathsf{GL}_n(\mathbb{F}_p) \\ \sigma & \longmapsto & u_\sigma : \begin{pmatrix} \mathbb{F}_p & \longrightarrow & \mathbb{F}_p \\ e_i & \longmapsto & e_{\sigma(i)} \end{pmatrix} \end{pmatrix}$$

avec $(e_1, \cdots, e_n)$ base de $\mathsf{GL}_n(\mathbb{F}_q)$ est un morphisme injectif, donc permet de voir $\mathfrak{S}_n$ comme un groupe isomorphe à un sous-groupe de $\mathsf{GL}_n(\mathbb{F}_p)$. Ainsi, G est isomorphe à un sous-groupe de $\mathsf{GL}_n(\mathbb{F}_p)$. Or, d'après **1.**, $\mathsf{GL}_n(\mathbb{F}_p)$ admet un p-sous-groupe de Sylow. D'après le lemme, il suit que G admet alors un p-sous-groupe de Sylow.

□

Théorème 27.2 : Deuxième théorème de SYLOW

Soit G un groupe fini de cardinal $p^\alpha m$ où p ne divise pas m.

1. Si H est un p-sous-groupe de G, il existe S un p-sous-groupe de Sylow tel que
$$H \subset S$$
2. Les p-sous-groupes de SYLOW sont tous conjugués entre eux. En conséquence, l'ensemble des p-sous-groupes de SYLOW forment une orbite sous G pour la translation à gauche, et le nombre $s_p(G)$ de p-sous-groupes de SYLOW de G divise $|G|$.
3. On a :
$$s_p(G) \equiv 1\ [p]$$
Par conséquent, $s_p(G)$ divise m.

Un corollaire du point **2.** est le suivant :

Corollaire 27.1

Soit G un groupe de cardinal $p^\alpha m$ avec p ne divisant pas m. Soit S un p-sous-groupe de SYLOW. Les assertions suivantes sont équivalentes :

1. S est un sous-groupe distingué dans G.
2. S est l'unique p-sous-groupe de SYLOW de G : $s_p(G) = 1$.

Démontrons le théorème.

Démonstration : [1.] Soit H un p-sous-groupe de G. Soit S un p-Sylow de G. Par le lemme **27.1**, il existe $a \in G$ tel que
$$aSa^{-1} \cap H < H$$
soit un p-Sylow de H. Pusique H est p-sous-groupe de G, il suit que
$$aSa^{-1} \cap H = H$$
d'où $H \subset aSa^{-1}$, donc H est bien inclus dans p-sous-groupe de Sylow.

[2.] De plus, si H est un p-Sylow, l'inclusion est une égalité par égalité des cardinaux, donc H est conjugué à S, donc tous les p-Sylow sont conjugués. On en déduit alors le corollaire : si S est un p-Sylow distingué, alors toute conjugaison de S n'est rien d'autre que S lui-même. Réciproquement, si S est l'unique p-Sylow de G, il est en particulier stable par conjugaison.

[3.] Puisque **2.** est montré, c'est aussi le cas du corollaire.

On note $\mathcal{X}$ l'ensemble des p-Sylow de G. Soit S un p-Sylow de G. On fait agir G sur $\mathcal{X}$ par conjugaison :
$$\begin{pmatrix} G & \longrightarrow & \mathcal{X} \\ g & \longmapsto & gSg^{-1} \end{pmatrix}$$

S agit lui aussi sur $\mathcal{X}$, par restriction. Par l'équation aux classes,
$$\sharp\mathcal{X} = s_p(G) \equiv \sharp\mathcal{X}^S\ [p]$$
avec $\mathcal{X}^S$ l'ensemble des points fixes de l'action :
$$\mathcal{X}^S \overset{\text{déf.}}{=} \{P \in \mathcal{X}, \forall s \in S, sPs^{-1} = P\}$$
Montrons que $\mathcal{X}^S = \{S\}$. $S \in \mathcal{X}^S$, puisque $sSs^{-1} = S$ pour tout $s \in S$. Soit $P \in \mathcal{X}^S$. Il existe $s \in S$ tel que
$$P = sPs^{-1}$$
Ainsi, P est stable par conjugaison dans S. On note
$$N \overset{\text{déf.}}{=} \langle S, P\rangle$$
alors P et S sont des p-Sylow de N, car p-Sylow de G, et P est distingué dans N. D'après le corollaire, P est alors l'unique p-sous-groupe de Sylow de N. Il suit que $P = S$, donc $\sharp\mathcal{X}^S = 1$. Il suit finalement que
$$s_p(G) \equiv 1\ [p]$$
□

28 Triplets pythagoriciens

Référence : Exercices d'algèbre, Aviva Szpirglas

Théorème 28.1 : Triplets pythagoriciens

On considère l'équation (E) d'inconnues $x, y, z \in \mathbb{Z}$:

$$(E) : \quad x^2 + y^2 = z^2$$

avec $x, y, z > 0$ et $x \wedge y = 1$. Alors $(x, y, z) \in \mathbb{Z}^3$ est solution de (E) si et seulement s'il existe $a > b$ entiers naturels premiers entre eux de parité différente tels que

$$\begin{cases} x &= 2ab \\ y &= a^2 - b^2 \\ z &= a^2 + b^2 \end{cases}$$

Démonstration : On raisonne en quatre étapes.

1. Soient $x, y, z \in \mathbb{Z}$ solution de (E) : ils vérifient $x^2 + y^2 = z^2$ avec $x, y, z > 0$ et $x \wedge y = 1$. Montrons que x, y sont de parité différente.

Si $x, y \in 2\mathbb{N}$ alors $x \wedge y \geqslant 2$, ce qui en contradiction directe avec $x \wedge y = 1$. Si $x, y \in 2\mathbb{N}+1$ alors x^2 et y^2 sont congrus à 1 modulo 4. Ainsi,

$$z^2 \equiv x^2 + y^2 \ [4] \equiv 2 \ [4]$$

Or, modulo 4, on a :

X	0	1	2	3
X^2	0	1	0	1

ce qui contredit $z^2 \equiv 2 \ [4]$. Ainsi, x, y sont bien de parités différentes. Par conséquent, $z^2 = x^2 + y^2$ donne que z est impair.

2. On suppose x pair. Montrons l'existence de $a, b \in \mathbb{N}$, avec $a > b$ tels que

$$\begin{cases} \frac{z-y}{2} &= a^2 \\ \frac{z+y}{2} &= b^2 \end{cases}$$

Puisque x est pair, y est impair, tout comme z. Il existe $x' \in \mathbb{N}$ tel que $x = 2x'$, et donc

$$4x'^2 = z^2 - y^2 = (z-y)(z+y)$$

Il existe $\alpha, \beta \in \mathbb{N}$ tels que $\alpha = \frac{z-y}{2}$ et $\beta = \frac{z+y}{2}$. Ainsi,

$$x'^2 = \alpha\beta$$

Montrons que α et β sont premiers entre eux. On note $d = \alpha \wedge \beta$. Puisque d divise α et β, alors d divise $\beta - \alpha$, donc d divise y. Puis, l'égalité $x'^2 = \alpha\beta$ donne d^2 divise x'^2, donc d divise x', donc d divise x. d divise x et y, qui sont premiers entre eux, donc $d = 1$.

Puisque α et β sont premiers entre eux, et que leur produit est un carré, il suit que $\alpha = a^2$ et $\beta = b^2$ avec $a, b \in \mathbb{N}^*$ (non nuls sans quoi x le serait). $a > b$ car $y^2 < z^2$, sans quoi $x = 0$) et a et b sont premiers entre eux.

3. On en déduit que

$$\begin{cases} z &= a^2 + b^2 \\ y &= a^2 - b^2 \\ x &= 2ab \end{cases}$$

Puis, si a, b sont pairs, alors z le serait, ce qui n'est pas le cas. De même, si a, b sont impairs, z serait aussi pair. Ainsi, a et b sont de parité différente.

4. Reste à faire la synthèse. Soient $a, b \in \mathbb{N}^*$ de parité différente, premiers entre eux, avec $a > b$, et on pose :

$$\begin{cases} z &= a^2 + b^2 \\ y &= a^2 - b^2 \\ x &= 2ab \end{cases}$$

Alors

$$x^2 + y^2 = z^2$$

avec $x, y, z > 0$. Soit $d = x \wedge y$. Puisque y est impair, d ne divise pas 2, d est donc impair. d divise x, donc d^2 divise x^2. d^2 divise aussi y^2. Ainsi, d^2 divise $x^2 + y^2 = z^2$ et d divise z. d divise z, d divise y, donc divise a^2 et b^2, qui sont premier entre eux. D'où $d = 1$, donc x et y sont bien premiers entre eux. □

29 Théorème de Fermat, cas particulier

Référence : Exercices d'algèbre, Aviva Szpirglas

Théorème 29.1 : Théorème de Fermat pour $n = 4$

L'équation d'inconnues $x, y, z \in \mathbb{Z}$

$$x^4 + y^4 = z^4$$

n'admet pas de solutions autres que les solutions triviales.

Remarque : Une solution est dite *triviale* si au moins l'une des inconnues est nulle.

Proposition 29.1 : Équation diophantienne auxiliaire

L'équation diophantienne d'inconnues $x, y, u \in \mathbb{Z}$:

$$x^4 + y^4 = u^2$$

n'admet pas de solution autre que les solutions triviales.

Si la proposition est démontrée, le théorème l'est aussi : si (x, y, z) est solution non trivial du problème de Fermat pour $n = 4$, alors (x, y, z^2) est solution non triviale de l'équation auxiliaire, ce qui est impossible. Le théorème de Fermat est donc montré pour $n = 4$.

Démonstration : 1. Constatons que si $x \wedge y = d$, alors il existe $x' \in \mathbb{Z}$ et $y' \in \mathbb{Z}$ tels que $x = dx'$ et $y = dy'$, auquel cas on aurait

$$d^4 \left(x'^4 + y'^4\right) = u^2$$

donc d^2 divise u, et $\left(\frac{x}{d}, \frac{y}{d}, \frac{u}{d^2}\right)$ est aussi solution, avec x' et y' premier entre eux. Dans la suite, on note $\mathcal{S}_{**}$ l'ensemble des solutions (x, y, u) de l'équation auxiliaire avec $x \wedge y = 1$.

On se donne pour le reste de la démonstration $(x, y, u) \in S_{**}$.

2. Montrons que x et y n'ont pas la même parité. S'ils étaient tous deux pairs, ils ne seraient pas premiers entre eux. Supposons qu'ils soient tous deux impairs. On dispose du tableau de congruence modulo 4 :

x	0	1	2	3
x^2	0	1	0	1

En conséquence,

$$x^4 + y^4 \equiv 2 \ [4]$$

Soit encore $u^2 \equiv 2 \ [4]$, ce qui n'est pas possible. Ainsi, on suppose dans la suite que x et y n'ont pas même parité : par exemple x pair.

3. Montrons qu'il existe alors une autre solution de l'équation auxiliaire, que l'on déduit de (x, y, u).

- Constatons d'abord que (x^2, y^2, u) est un triplet pythagoricien avec x^2 pair, donc d'après la section précédente, il existe $a, b \in \mathbb{Z}$ premier entre eux tels que

$$\begin{cases} x^2 &= 2ab \\ y^2 &= a^2 - b^2 \\ u &= a^2 + b^2 \end{cases}$$

De plus, si a est pair, alors b est impair, et on aurait

$$a^2 - b^2 \equiv -1 \ [4]$$

soit $y^2 \equiv -1 \ [4] \equiv 3 \ [4]$, ce qui est impossible. On suppose donc que a est impair dans la suite, et que b est pair : $b = 2c$, avec $a \wedge c = 1$, sans quoi a, b ne seraient pas premiers entre eux.

- Ainsi, $x^2 = 4ac$, et donc

$$\left(\frac{x}{2}\right)^2 = ac$$

avec a, c premiers entre eux, donc a, c sont des carrés : il existe $\alpha, \gamma \in \mathbb{Z}$ tels que $a = \alpha^2$ et $c = \gamma^2$, avec $\alpha \wedge \gamma = 1$ et α impair. Par conséquent,

$$y^2 = a^2 - b^2 = \alpha^4 - 4\gamma^4$$

donc

$$\left(2\gamma^2\right)^2 + y^2 = \left(\alpha^2\right)^2$$

donc on est encore en présence d'un triplet pythagoricien, avec $2\gamma^2$ pair. Il existe $l, m \in \mathbb{Z}$ premiers entre eux tels que

$$\begin{cases} 2\gamma^2 &= 2lm \\ y &= l^2 - m^2 \\ \alpha^2 &= l^2 + m^2 \end{cases}$$

Ainsi, $\gamma^2 = lm$ avec $l \wedge m = 1$, donc l et m sont des carrés : il existe $r, s \in \mathbb{Z}$ tels que $l = r^2$ et $m = s^2$, avec r, s premiers entre eux. Ainsi,

$$\alpha^2 = r^4 + s^4$$

donc on a $(r, s, \alpha) \in S_{**}$. Cette solution vérifie de plus que $\alpha < u$: en effet,

$$\alpha \leqslant \alpha^2 = a \leqslant a^2 < a^2 + b^2 = u$$

4. On peut alors conclure. Pour tout triplet $(x_1, y_1, u_1) \in S_{**}$, il existe $(x_2, y_2, u_2) \in S_{**}$ tel que $0 < u_2 < u_1$. Ainsi, par récurrence, pour tout triplet $(x_1, y_1, u_1) \in S_{**}$, il existe une suite de solutions $\{(x_n, y_n, u_n)\}_{n\in\mathbb{N}} \in S_{**}^{\mathbb{N}}$ qui vérifie de plus

$$\forall n \in \mathbb{N}^*, 0 < u_{n+1} < u_n$$

ce qui est impossible, puisque les u_n sont des entiers. Par conséquent,

$$S_{**} = \varnothing$$

L'équation auxiliaire n'admet pas de solution non triviale. □

30 Méthode de Newton

Référence : Analyse numérique et équation différentielle, Jean-Pierre DEMAILLY

Cette section est un préliminaire por la section suivante à propos de la décomposition de DUNFORD.

Théorème 30.1 : Méthode de NEWTON

Soit $f \in \mathcal{C}^2([a,b], \mathbb{R})$ telle que $f(a)f(b) < 0$, et $f' \neq 0$ sur $[a,b]$. On définit par récurrence la suite $(x_n)_{n\in\mathbb{N}}$: $x_0 \in [a,b]$ et

$$\forall n \in \mathbb{N}, x_{n+1} \overset{\text{déf.}}{=} x_n - \frac{f(x_n)}{f'(x_n)}$$

On note $x^* \in [a,b]$ est telle que $f(x^*) = 0$ et

$$M \overset{\text{déf.}}{=} \sup_{x\in[a,b]} \left|\frac{f''(x)}{f'(x)}\right|$$

Alors, pour $h \leqslant \frac{1}{M}$ avec $[x^* - h, x^* + h] \subset [a,b]$, on dispose de la majoration : pour tout $x_0 \in [x^* - h, x^* - h]$:

$$\forall n \in \mathbb{N}, |x_n - x^*| \leqslant \frac{1}{M}\left(M|x_0 - x^*|\right)^{2^n}$$

Démonstration : Commençons par dire que x^* existe par le théorème des valeurs intermédiaires. On note pour $x \in [a,b]$

$$\phi(x) \overset{\text{déf.}}{=} x - \frac{f(x)}{f'(x)}$$

La relation de récurrence s'écrit alors $x_{n+1} = \phi(x_n)$. Observons par ailleurs que

$$\phi'(x) = \frac{f(x)f''(x)}{f'(x)^2}$$

- Notons

$$u(x) = \frac{f(x)}{f'(x)}$$

alors

$$u'(x) = 1 - \frac{f''(x)}{f'(x)}u(x)$$

D'où

$$|u'(x)| \leqslant 1 + M|u(x)|$$

De plus, u a le même signe que $x \longmapsto x - x^*$. En effet, f est monotone sur $[a,b]$ car $f' \neq 0$. Deux cas : si f est croissante, $f' > 0$, et par croissance, $u(x) < 0$ sur $[a, x^*[$ et $u(x) > 0$ sur $]x^*, b]$. Si f est décroissante, $f' < 0$, et on a les mêmes conclusions. Par conséquent, pour résoudre cette inégalité différentielle, on peut d'abord se placer sur $[a, x^*[$ où $u > 0$. On a alors

$$u'(x) \leqslant |u'(x)| \leqslant 1 + Mu(x)$$

Ainsi, sur $[a, x^*[$

$$u(x) \leqslant \frac{1}{M}\left(e^{M(x-x^*)}-1\right)$$

Puis, sur $]x^*, b]$, on a

$$-u(x) \leqslant \frac{1}{M}\left(e^{M(x^*-x)}-1\right)$$

Finalement, sur $[a, b]$

$$|u(x)| \leqslant \frac{1}{M}\left(e^{M|x-x^*|}-1\right)$$

- Observons aussi que l'on a

$$\phi'(x) = u(x)\frac{f''(x)}{f'(x)}$$

Ainsi,

$$|\phi'(x)| \leqslant M|u(x)|$$

donc

$$|\phi'(x)| \leqslant e^{M|x-x^*|}-1$$

Par convexité de la fonction exponentielle, qui se situe donc en dessous de sa tangente sur $[x^*-h, x^*+h]$:

$$|\phi'(x)| \leqslant 2M|x-x^*|$$

- Ainsi,

$$|\phi(x)-\phi(x^*)| \leqslant \left|\int_x^{x^*} |\phi'(t)|\,\mathrm{d}t\right|$$

donc

$$|\phi(x)-x^*| \leqslant M|x-x^*|^2$$

Reste à conclure par récurrence. □

31 Théorème de Dunford-Jordan-Chevalley *via* Newton

Références :

— Matthieu Romagny, *Algèbre de dimension finie*

— *Mathématiques Tout-en-un pour la licence, niveau L2*, Jean-Pierre Ramis, 2007, page 268

— *Algèbre pour la license 3*, Jean-Jacques Risler, problème 3.1

Cette section présente un algorithme pour déterminer les matrices D et N de la décomposition de Dunford (-Jordan-Chevalley) sans avoir à déterminer les valeurs propres de A, qui sont quasi-inaccessibles dès lors que la dimension est conséquente. Cet algorithme a le bon goût de terminer en un nombre fini d'étapes, ce qui fournit une nouvelle preuve de ladite décomposition.

L'algorithme utilise le principe de la méthode de Newton que l'on a traité dans la section précédente ! Le principe est le suivant : si on note

$$\chi_A(X) = \prod_{i=1}^{r}(X-\lambda_i)^{m_i}$$

On considère le polynôme

$$T(X) \overset{\text{déf.}}{=} \prod_{i=1}^{r}(X-\lambda_i)$$

avec $T(D)=0$. Inspiré par la méthode de Newton, on construit la suite de matrices suivantes : on se donne $M_0 = A \in \mathcal{M}_n(\mathbb{K})$ ($\mathbb{K}$ est un corps dont on précisera les hypothèses), puis pour $k \in \mathbb{N}$:

$$M_{k+1} \overset{\text{déf.}}{=} M_k - T(M_k)T'(M_k)^{-1}$$

On va montrer que cette suite est bien définie, et converge vers D. Mieux, cela se passe en un nombre d'étapes fini ! Plus précisément, l'algorithme termine pour $\lceil \log_2(n) \rceil$.

Théorème 31.1 : Théorème de décomposition de Jordan-Chevalley

Soit $\mathbb{K}$ un corps de caractéristique 0. Soit $A \in \mathcal{M}_n(\mathbb{K})$, et on suppose que son polynôme caractéristique χ_A est scindé dans $\mathbb{K}$. Alors il existe un unique couple $(D, N) \in \mathcal{M}_n(\mathbb{K})^2$ tel que :

* D est diagonalisable sur $\mathbb{K}$;
* N est nilpotente ;
* $DN = ND$;
* $A = D + N$

De plus, $D, N \in \mathbb{K}[A]$.
Enfin, si $\mathbb{L} \subset \mathbb{K}$ est un sous-corps de $\mathbb{K}$ et que $A \in \mathcal{M}_n(\mathbb{L})$, alors $D, N \in \mathbb{L}[A] \subset \mathcal{M}_n(\mathbb{L})$.

Remarque : L'hypothèse émise pour le corps $\mathbb{K}$ dans le cours de Matthieu Romagny est que $\mathbb{K}$ est *parfait* : $\mathbb{K}$ est de caractéristique 0 ou si $p = \mathrm{Car}(\mathbb{K})$, l'application

$$\mathrm{Frob} : \left(\begin{array}{ccc} \mathbb{K} & \longrightarrow & \mathbb{K} \\ x & \longmapsto & x^p \end{array} \right)$$

est un morphisme de corps surjectif.

Lemme 31.1 : Somme d'un inversible et d'un nilpotent

Soient $A, N \in \mathcal{M}_n(\mathbb{K})$, avec A inversible et N nilpotente, et $AN = NA$. Alors

$$M \overset{\text{déf.}}{=} A + N$$

est inversible, avec $M^{-1} \in A \cdot \mathbb{K}\left[A^{-1}N\right]$.

Démonstration du lemme : Puisque A est inversible, M se factorise :

$$M = A\left(I_n + A^{-1}N\right)$$

Or, en s'inspirant de la somme d'une série géométrique :

$$\frac{1}{1-u} = \sum_{k=0}^{+\infty} u^k$$

On calcule

$$\Xi \overset{\text{déf.}}{=} \left(I_n + A^{-1}N\right) \sum_{k=0}^{n} (-1)^k \left(A^{-1}N\right)^k$$

En développant, on a simplement

$$\Xi = \sum_{k=0}^{n} (-1)^k \left(A^{-1}N\right)^k + \sum_{k=0}^{n} (-1)^k \left(A^{-1}N\right)^{k+1}$$

On fait un changement d'indices :

$$\Xi = \sum_{k=0}^{n} (-1)^k \left(A^{-1}N\right)^k - \sum_{k=1}^{n+1} (-1)^k \left(A^{-1}N\right)^k$$

On regroupe :

$$\begin{aligned} \Xi = \quad & I_n + \sum_{k=1}^{n} (-1)^k \left(A^{-1}N\right)^k \\ & - \sum_{k=1}^{n} (-1)^k \left(A^{-1}N\right)^k + (-1)^n \left(A^{-1}N\right)^{n+1} \end{aligned}$$

Le deux sommes se compensent, et par commutativité de A et N :

$$\Xi = I_n + (-1)^n \left(A^{-1}\right)^{n+1} N^{n+1}$$

Enfin, N est nilpotente, donc $\Xi = I_n$. Autrement dit, $I_n + A^{-1}N$ est inversible, et M est donc produit de deux inversibles, ce qui en fait aussi une matrice inversible. □

Ce lemme permet alors d'affirmer que dans une algèbre commutative, les nilpotents sont les *infinitésimaux* de l'algèbre.

Démonstration : ▷ Intéressons-nous à l'unicité d'un tel couple. On la démontre de la même manière que la section présentant la manière "classique" de prouver ce théorème. On se donne deux décompositions

$$A = D + N = D' + N'$$

Puisque $D, D' \in \mathbb{K}[A]$, ces endomorphismes commutent, et puisqu'ils sont diagonalisables, ces deux matrices admetent une base commune de diagonalisation, donc $D - D'$ est diagonalisable. Puisque $N, N' \in \mathbb{K}[A]$, ces matrices commutent, et la formule du binôme appliqué à la puissance correspondant au maximum des indices de nilpotence montre que $N - N'$ est nilpotente. Ainsi,

$$D - D' = N' - N$$

est une matrice diagonalisable nilpotente. Il suit que cette matrice est nulle ; d'où l'unicité.

▷ Montrons l'existence de cette décomposition. On décompose le polynôme χ_A scindé sur $\mathbb{K}$:

$$\chi_A(X) = \prod_{i=1}^{r}(X - \lambda_i)^{m_i}$$

avec $r \in \mathbb{N}^*$, et pour tout $i \in [\![1, r]\!]$, $\lambda_i \in \mathbb{K}$ distincts, $m_i \in \mathbb{N}^*$. On définit le polynôme $T \in \mathbb{K}[X]$ par :

$$T(X) \overset{\text{déf.}}{=} \prod_{i=1}^{r}(X - \lambda_i)$$

L'objectif est de déterminer $D \in \mathcal{M}_n(\mathbb{K})$ tel que $T(D) = 0$. On obtiendra une matrice qui possède un polynôme annulateur scindé à racines simples, donc D sera bien diagonalisable.

1. T est un polynôme de $\mathbb{K}[X]$ scindé à racines simples. T est alors premier avec T', son polynôme dérivé. Notons

$$m \overset{\text{déf.}}{=} \max\{m_i, 1 \leqslant i \leqslant r\}$$

Alors T^m divise χ_M. Puisque T' est premier avec T^m, est donc premier avec χ_A. Il existe donc une relation de Bézout entre ces deux polynômes : il existe $U, V \in \mathbb{K}[X]$ tels que

$$\chi_A U + T'V = 1$$

Par conséquent, en évaluant en A, le théorème de Cayley-Hamilton nous donne $\chi_A(A) = 0$. Ainsi, $T(A)^m = 0$, donc $T(A)$ est nilpotente. Notons

$$\varepsilon \overset{\text{déf.}}{=} T(A)$$

On a aussi par la relation de Bézout prise en A :

$$T'(A)V(A) = 1$$

donc $T'(A)$ est inversible, avec un inverse dans $\mathbb{K}[A]$. Dans la suite, on note :

$$B = O(\varepsilon) \iff B \in \langle\varepsilon\rangle = \varepsilon\mathbb{K}[A]$$

2. On construit une suite dans $\mathbb{K}[A]$, donnée par récurrence : on note $M_0 \overset{\text{déf.}}{=} A$ et pour tout $k \in \mathbb{N}$:

$$M_{k+1} \overset{\text{déf.}}{=} M_k - T(M_k)T'(M_k)^{-1}$$

Montrons par récurrence sur k que :

- **(i)** La suite $(M_k)_{k\in\mathbb{N}}$ est bien définie, et est une suite de $\mathbb{K}[A]$, qui est une algèbre commutative ;
- **(ii)** Pour tout $k \in \mathbb{N}$,

$$T(M_k) = O\left(\varepsilon^{2^k}\right)$$

- **(iii)** Pour tout $k \in \mathbb{N}$,

$$M_k - A = O(\varepsilon)$$

Ces trois propriétés sont vraies pour $k = 0$, où $M_0 = A$. Pour $k \geqslant 0$, supposons les résultats montrés jusqu'au rang k. Montrons ces résultats au rang $k + 1$.

(i) On a deux choses à montrer : le fait que $T'(M_k)$ soit inversible et que $M_{k+1} \in \mathbb{K}[A]$. Puisque $M_k - A = O(\varepsilon)$, on a en développant :

$$T'(M_k) = T'(A) + O(\varepsilon)$$

En **1.**, on a montré que $T'(A)$ est inversible. De plus, $O(\varepsilon)$ est nilpotente, par commutativité dans $\mathbb{K}[A]$. Ainsi, par le lemme, $T'(M_k)$ est somme d'un nilpotent et d'un inversible dans $\mathbb{K}[A]$ commutatif, donc $T'(M_k)$ est inversible. De plus, toujours d'après le lemme, cet inverse vérifie

$$T'(M_k)^{-1} \in T'(A) \cdot \mathbb{K}\left[T'(A)^{-1}O(\varepsilon)\right]$$

Puisque $T'(A)$ est d'inverse dans $\mathbb{K}[A]$, il suit que $T'(M_k)^{-1} \in \mathbb{K}[A]$, donc

$$M_{k+1} = M_k - T(M_k)T'(M_k)^{-1} \in \mathbb{K}[A]$$

(ii) $\mathbb{K}$ est un corps de caractéristique 0 et $\mathbb{K}[A]$ est une algèbre commutative : on peut appliquer la formule de Taylor à l'ordre 1 en M_k :

$$T(M_{k+1}) = T(M_k) + (M_{k+1} - M_k)T'(M_k) + (M_{k+1} - M_k)^2 Z$$

avec $Z \in \mathbb{K}[A]$. Le miracle de la méthode de Newton a ici lieu :

$$T(M_k) + (M_{k+1} - M_k)T'(M_k) = 0$$

D'où finalement

$$T(M_{k+1}) = T(M_k)^2 Z'$$

avec $Z' \in \mathbb{K}[A]$. La propriété **(ii)** est alors montrée, grâce à l'hypothèse de récurrence.

(iii) Par théorème belge,

$$M_{k+1} - A = M_{k+1} - M_k + M_k - A$$

Par **(ii)**, $M_{k+1} - M_k = O\left(\varepsilon^{2^k}\right)$, et par hypothèse de récurrence, $M_k - A = O(\varepsilon)$, donc

$$M_{k+1} - A = O(\varepsilon)$$

Cela conclut la récurrence.

3. ε est nilpotent, donc $\varepsilon^n = 0$. En particulier, lorsque $2^k \geqslant n$, $T(M_k) = 0$, donc la suite $(M_k)_k$ est constante à partir de ce rang. On note D la matrice des M_k pour ce rang. Alors $T(D) = 0$, avec $D \in \mathbb{K}[A]$. Reste à poser $N = A - D$. D'après **(iii)**, $N \in O(\varepsilon)$, donc est nilpotente. De plus, $N \in \mathbb{K}[A]$. On a alors réussi à conclure sur l'existence de $D, N \in \mathbb{K}[A]$ tels que

$$A = D + N$$

▷ Montrons la dernier fait à propos du sous-corps $\mathbb{L}$. Puisque $A \in \mathcal{M}_n(\mathbb{L})$, $\chi_A \in \mathbb{L}[X]$. *A priori*, les λ_i ne sont pas dans $\mathbb{L}$. Montrons que $T \in \mathbb{L}[X]$ pour conclure sur $D, N \in \mathbb{L}[A]$. On a :

$$T = \frac{\chi_A}{\chi_A \wedge \chi'_A}$$

Or,

$$\chi'_A(X) = \sum_{j=1}^{r} n_j (X - \lambda_j)^{m_j - 1} \prod_{i \neq j} (X - \lambda_i)^{m_i}$$

Si bien que puisque $m_j 1_{\mathbb{K}} \neq 0$ ($\mathbb{K}$ est de caractéristique 0), le PGCD de χ_A par χ'_A n'est rien d'autre que

$$\chi_A \wedge \chi'_A = \prod_{j=1}^{n} (X - \lambda_j)^{m_j - 1}$$

Et donc T est le reste par la division euclidienne de χ_A par $\chi_A \wedge \chi'_A$. χ_A et χ'_A étant deux éléments de $\mathbb{L}[X]$, il suit que $\chi_A \wedge \chi'_A$ aussi, et par algorithme d'Euclide sur les polynômes, c'est aussi le cas pour T. On a donc bien montré que $D, N \in \mathbb{L}[A]$. □

32 Dualité et bidualité pour un groupe fini abélien

Références :

- Matthieu Romagny, *Représentation linéaire des groupes finis* ;
- Gabriel Peyré, *Algèbre discrète de la transformée de* Fourier

Définition 32.1

Soit G un groupe. On appelle *caractère de* G un morphisme de groupes χ de G dans le groupe multiplicatif $(\mathbb{C}^*, \times)$. On note $\hat{G}$ l'ensemble des caractères de G.

$\hat{G}$ est un groupe pour le produit. Le but de cette section est de montrer les théorèmes **32.1** et **32.2** : pour tout groupe fini G, G est isomorphe à $\hat{G}$, et à $\hat{\hat{G}}$. L'intérêt du deuxième isomorphisme est qu'il a le mérite d'être canonique : il ne dépend pas de choix de base.

Proposition 32.1 : Dualité pour un groupe cyclique

Soit G un groupe fini cyclique d'ordre n. Alors $\hat{G} \simeq G$. Plus particulièrement, si ω est une racine n-ème de l'unité, et g un générateur de G, alors

$$\widehat{G} = \left\{ \left(\begin{array}{ccc} G & \longrightarrow & \mathbb{C} \\ g^k & \longmapsto & \omega^{kj} \end{array} \right) \right\}_{j \in [\![0, n-1]\!]}$$

Démonstration : On note pour $\omega \in \mathbb{U}_n$ l'application

$$\chi_\omega : \begin{pmatrix} G & \longrightarrow & \mathbb{C} \\ g^k & \longmapsto & \omega^k \end{pmatrix}$$

On a

$$\mathbb{U}_n = \{\omega^j, j \in [\![0, n-1]\!]\}$$

et on constate que

$$\{\chi_{\omega^j}, j \in [\![0, n-1]\!]\} \subset \hat{G}$$

C'est en effet une famille de morphismes de groupes de G dans $\mathbb{C}$, bien définie : si $g^k = g^{k'}$ alors $g^{k-k'} = 1_G$, donc k et k' sont congrus modulo n. Ainsi,

$$\chi_\omega(g^k) = \omega^k$$

Puisque ω est racine n-ème de l'unité, il suit

$$\chi_\omega(g^k) = \omega^{k'} = \chi_\omega(g^{k'})$$

Soit $\chi \in \hat{G}$. Montrons qu'il existe $j \in [\![0, n-1]\!]$ tel que $\chi = \chi_{\omega^j}$. Pour cela, puisque $g^n = 1$ pour tout $g \in G$, on observe donc que

$$\chi(g)^n = 1$$

Donc, pour tout $g \in G$, il existe $j(g) \in [\![0, n-1]\!]$ tel que $\chi(g) = \omega^{j(g)}$. Si g est générateur de G, alors tout élément de G est de la forme g^k, et donc par propriété de morphisme,

$$\chi(g^k) = \omega^{kj(g)}$$

Donc, on a $\chi = \chi_{\omega^{j(g)}}$ avec g générateur de G. D'où l'égalité souhaitée entre $\hat{G}$ et l'ensemble des χ_{ω_j}. Cela permet alors de conclure sur l'isomorphisme entre G et $\hat{G}$. En effet, puisque G est cyclique, G est isomorphe à $\mathbb{Z}/n\mathbb{Z}$. $\hat{G}$ est quand à lui isomorphe à $\mathbb{U}_n$, *via* $\omega \longmapsto \chi_\omega$. Par l'application exponentielle, et propriété universelle du quotient, $\mathbb{U}_n$ est lui-même isomorphe à $\mathbb{Z}/n\mathbb{Z}$. Ainsi, $\hat{G}$ est isomorphe à G. □

Théorème 32.1 : Dualité pour les groupes finis abélien

Soit G un groupe fini abélien. Alors $G \simeq \hat{G}$. L'isomorphisme n'est pas canonique : il dépend d'un choix d'une racine de l'unité.

Démonstration : On utilise pour cela le théorème de structure des groupes abéliens de type fini : il existe $n_1 | \cdots | n_r > 0$ tels que

$$G \simeq \mathbb{Z}/n_1\mathbb{Z} \times \cdots \times \mathbb{Z}/n_r\mathbb{Z}$$

Puisqu'on a montré en proposition **32.1** que

$$\widehat{\mathbb{Z}/n\mathbb{Z}} \simeq \mathbb{Z}/n\mathbb{Z}$$

il nous suffit de montrer que si G et H sont deux groupes finis cycliques, alors

$$\widehat{G \times H} \simeq \hat{G} \times \hat{H}$$

On note Φ l'application

$$\begin{pmatrix} \hat{G} \times \hat{H} & \longrightarrow & \widehat{G \times H} \\ (\chi, \psi) & \longmapsto & \Phi_{g,h} : \begin{pmatrix} G \times H & \longrightarrow & \mathbb{C}^* \\ (g,h) & \longmapsto & \chi(g)\psi(h) \end{pmatrix} \end{pmatrix}$$

qui permet d'obtenir un isomorphisme qui conclut. Montrons déjà que Φ est bien à valeurs dans $\widehat{G \times H}$. Soit $(\chi, \psi) \in \hat{G} \times \hat{H}$. Montrons que $\Phi_{\chi,\psi}$ est un morphisme de groupes, où on munit $G \times H$ de la multiplication directe :

$$(g_1, h_1) \cdot (g_2, h_2)^{-1} = (g_1 g_2^{-1}, h_1 h_2^{-1})$$

Il s'agit en fait de voir qu'on dispose de la relation

$$\Phi_{\chi,\psi}(g_1 g_2^{-1}, h_1 h_2^{-1}) = \frac{\Phi_{\chi,\psi}(g_1, h_1)}{\Phi_{\chi,\psi}(g_2, h_2)}$$

qui se traduit en fait par :

$$\chi(g_1 g_2^{-1})\psi(h_1 h_2^{-1}) = \frac{\chi(g_1)\psi(h_1)}{\chi(g_2)\psi(h_2)}$$

ce qui est vrai car χ et ψ sont des caractères. Montrons que Φ est un morphisme de groupes. Soient $(\chi_1, \psi_1), (\chi_2, \psi_2) \in \hat{G} \times \hat{H}$, qui est un groupe, avec la multiplication :

$$(\chi_1, \psi_1) \cdot (\chi_2, \psi_2)^{-1} \overset{\text{déf.}}{=} \left(\frac{\chi_1}{\chi_2}, \frac{\psi_1}{\psi_2}\right)$$

Pour tout $(g,h) \in G \times H$,

$$\Phi_{\frac{\chi_1}{\chi_2}, \frac{\psi_1}{\psi_2}}(g,h) = \frac{\chi_1(g)}{\chi_2(g)} \cdot \frac{\psi_1(h)}{\psi_2(h)}$$

Donc, on a finalement

$$\Phi_{\frac{\chi_1}{\chi_2}, \frac{\psi_1}{\psi_2}}(g,h) = \frac{\Phi_{\chi_1,\psi_1}(g,h)}{\Phi_{\chi_2,\psi_2}(g,h)}$$

Ce qui montre que Φ est un morphisme de groupes. Montrons enfin qu'il s'agit d'un isomorphisme. Pour l'injectivité, si $(\chi, \psi) \in \ker \Phi$ alors

$$\forall g \in G, \forall h \in H, \chi(g)\psi(h) = 1$$

donc, pour $g = 1_G$, il suit que ψ est le morphisme trivial sur H dans $\mathbb{C}^*$, donc que χ est lui le morphisme trivial sur G dans $\mathbb{C}^*$. Donc Φ est injective. Pour conclure, puisque G et H sont cycliques, $G \times H$ l'est. Ainsi, par la proposition précédente,

$$\left|\widehat{G \times H}\right| = |G \times H|$$

Par propriété du produit cartésien,

$$\left|\widehat{G \times H}\right| = |G| \cdot |H|$$

Puisque G et H sont cycliques, par la propriété précédente :

$$\left|\widehat{G \times H}\right| = |\hat{G}| \cdot |\hat{H}|$$

Puis, par produit cartésien,

$$\left|\widehat{G \times H}\right| = \left|\hat{G} \times \hat{H}\right|$$

Donc Φ est un morphisme de groups injectif entre deux groupes de même cardinal, donc Φ est un isomorphisme, donc on a bien montré que lorsque G et H sont deux groupes cycliques finis :

$$\widehat{G \times H} \simeq \hat{G} \times \hat{H}$$

Le théorème de structure des groupes abéliens de type fini permet de conclure que pour tout groupe fini abélien G, ce groupe est bien isomorphe à $\hat{G}$. □

Remarque : On peut aussi un autre isomorphisme que Φ, exploité dans le livre de Peyré. On note i_G et i_H les injections :

$$i_G : \begin{pmatrix} G & \longrightarrow & G \times H \\ g & \longmapsto & (g, 1_H) \end{pmatrix}$$

alors

$$\begin{pmatrix} \widehat{G \times H} & \longrightarrow & \hat{G} \times \hat{H} \\ \chi & \longmapsto & (\chi \circ i_G, \chi \circ i_H) \end{pmatrix}$$

est un isomorphisme. Il s'agit en effet d'un morphisme de groupes, injectif : si pour tout $g \in G$ et $h \in H$,

$$(\chi(g, 1_H), \chi(1_G, h)) = (1, 1)$$

Alors

$$\chi(g, h) = \chi(g \cdot 1_G, h \cdot 1_H)$$

donc

$$\chi(g, h) = \chi(g, 1_H)\chi(1_G, h) = 1$$

Ce morphisme est surjectif : si on se donne $(\chi_1, \chi_2) \in \hat{G} \times \hat{H}$, alors

$$\chi : \begin{pmatrix} G \times H & \longrightarrow & \mathbb{C}^* \\ (g, h) & \longmapsto & \chi_1(g)\chi_2(h) \end{pmatrix}$$

en est un antécédent (c'est notre morphisme Φ).

Lemme 32.1 : Prolongement de caractères

Soit G un groupe fini commutatif, et soit H un sous-groupe de G. Soit $\psi_0 \in \hat{H}$. Alors il existe $\chi \in \hat{G}$ tel que

$$\chi_{|H} = \psi_0$$

Démonstration du lemme : On effectue cette démonstration par récurrence (forte) sur l'indice de H dans G : $[G:H]$.

Lorsque $[G:H]=1$, $G=H$, donc il n'y a pas de prolongement à effectuer, le résultat est vrai *via* l'égalité triviale $\chi=\chi$.

Supposons que $[G:H]>1$, et que pour tout tout sous-groupe $K<G$ tel que $[G:K]<[G:H]$, et pour tout $\xi\in\hat{K}$, il existe $\chi\in\hat{G}$ qui prolonge ξ. Puisque $[G:H]>1$, $H\subsetneq G$, et donc on peut considérer $x\in G\backslash H$. Considérons

$$K\overset{\text{déf.}}{=}\langle x,H\rangle$$

D'après le théorème de Lagrange, $x^{|G|}=1_G\in H$. Ainsi, on peut considérer

$$n\overset{\text{déf.}}{=}\min\left\{k\in\mathbb{N}^*,x^k\in H\right\}$$

Alors, tout élément de K s'écrit de manière unique hx^k avec $h\in H$ et $k\in[\![0,n-1]\!]$. En effet, si

$$hx^k=h'x^{k'}$$

avec $h,h'\in H$ et $k,k'\in[\![0,n-1]\!]$ alors

$$x^{k-k'}=h'h^{-1}\in H$$

Puisque $|k-k'|\in[\![0,n-1]\!]$, que $x^{|k-k'|}\in H$, et par la définition de n, il suit que $k=k'$. Il suit que $h=h'$, d'où l'unicité de l'écriture :

$$K=\left\{hx^k\mid h\in H,k\in[\![0,n-1]\!]\right\}$$

Soit $\psi\in\hat{H}$. Montrons qu'il existe $\chi\in\hat{K}$ qui prolonge ψ.

- Supposons que l'on dispose d'un prolongement χ. Alors pour tout $z=hx^k\in K$,

$$\chi(z)=\chi(hx^k)=\chi(h)\chi(x)^k$$

Et puisque χ est un prolongement de ψ,

$$\chi(z)=\psi(h)\chi(x)^k$$

On note

$$\zeta\overset{\text{déf.}}{=}\chi(x)$$

Par définition de n, $x^n\in H$. Alors

$$\zeta^n=\chi(x^n)=\psi(x^n)$$

Supposons alors que l'on ait $\psi(x^n)=1$. Alors ζ est une racine de l'unité. Avec cette hypothèse, il suit que

$$\chi(z)=\psi(h)\zeta^k$$

avec ζ une racine n-ème de l'unité.

- Faisons la synthèse. Soit $\psi\in\hat{H}$. Soit $\zeta\in\mathbb{U}_n$. On définit pour $z=hx^k\in K$:

$$\chi(z)\overset{\text{déf.}}{=}\psi(h)\zeta^k$$

Montrons que $\chi\in\hat{K}$ prolonge ψ. χ est bien définie par unicité de l'écriture de z. Soient $z,z'\in K$. Alors

$$zz'=hh'x^{k+k'}$$

avec $k+k'\in[\![0,2n-2]\!]$. Si $k+k'\in[\![0,n-1]\!]$, alors

$$\chi(zz')=\psi(hh')\zeta^{k+k'}=\chi(z)\chi(z')$$

Sinon, on peut aussi écrire

$$zz'=hh'x^nx^{k+k'-n}$$

avec cette fois $k+k'-n\in[\![0,n-1]\!]$. Puisque $hh'x^n\in H$, on a alors ici l'écriture sous forme unique de zz' dans K décrite précédemment. Ainsi,

$$\chi(zz')=\psi(hh'x^n)\zeta^{k+k'-n}$$

Soit, puisque ψ est un morphisme de groupes et que $\zeta^n=1$,

$$\chi(zz')=\psi(hh')\zeta^{k+k'}$$

d'où $\chi(zz')=\chi(z)\chi(z')$ encore ici.

Ainsi, on a construit un caractère χ qui prolonge le morphisme ψ sur K. Or,

$$[G:H]=[G:K][K:H]$$

avec $[K:H]>1$. Ainsi, $[G:K]<[G:H]$. Par hypothèse de récurrence, il existe alors $\tilde{\chi}\in\hat{G}$ qui prolonge $\chi\in\hat{K}$, qui lui-même prolonge $\psi\in\hat{H}$. Cela conclut donc la récurrence. □

Remarque : Ce prolongement n'est pas unique, et dépend de ζ, tout comme l'isomorphisme entre G et $\hat{G}$.

Théorème 32.2 : Bidualité pour les groupes finis

Soit G un groupe fini abélien. Alors l'application

$$\Phi : \left(\begin{array}{ccc} G & \longrightarrow & \hat{\hat{G}} \\ g & \longmapsto & \Phi_g : \left(\begin{array}{ccc} \hat{G} & \longrightarrow & \mathbb{C} \\ \chi & \longmapsto & \chi(g) \end{array} \right) \end{array} \right)$$

est un isomorphisme entre G et $\hat{\hat{G}}$, qui lui est par conséquent canonique.

Démonstration : D'abord Φ est bien définie : pour tout $g \in G$, Φ_g est un morphisme de groupes, c'est l'évaluation en un point. Montrons que Φ est un morphisme. Si $g, g' \in G$, alors pour tout $\chi \in \hat{G}$,

$$\Phi_{gg'}(\chi) = \chi(gg')$$

Puisque χ est un morphisme de groupes,

$$\Phi_{gg'}(\chi) = \chi(g)\chi(g')$$

D'où

$$\Phi_{gg'}(\chi) = \Phi_g(\chi)\Phi_{g'}(\chi)$$

d'où $\Phi_{gg'} = \Phi_g \Phi_{g'}$. Reste à montrer que Φ est un isomorphisme. D'une part, par le théorème de dualité pour les groupes finis abéliens,

$$|G| = \left|\hat{\hat{G}}\right|$$

Reste à montrer l'injectivité de Φ. Si $g \in \ker(\Phi)$ alors

$$\forall \chi \in \hat{G}, \chi(g) = 1$$

Supposons par l'absurde que $g \neq 1_G$. Construisons un caractère $\chi \in \hat{G}$ tel que $\chi(g) \neq 1$ pour aboutir à une contradiction. On considère $H \overset{\text{déf.}}{=} \langle g \rangle < G$. H est cyclique donc, par la description faite en proposition **32.1**, il existe un caractère $\chi_0 \in \hat{H}$ tel que $\chi_0(g) \neq 1$. Par le lemme **32.1**, on peut prolonger χ_0 en un caractère $\chi \in \hat{G}$. Alors

$$\chi(g) = \chi_0(g) \neq 1$$

ce qui nous permet d'aboutir à notre contradiction, donc Φ est injective, et G est donc isomorphe à son bi-dual, de manière canonique. □

33 Dual du groupe symétrique

Référence : Gabriel PEYRÉ, *Algèbre discrète de la transformée de* FOURIER.

Contrairement à la section précédente, on travaille cette fois-ci un groupe fini, non supposé abélien.

Théorème 33.1 : Dualité pour les groupes finis

Soit G un groupe fini. On a alors :

$$\hat{G} \simeq {}^{G}/_{D(G)}$$

où $D(G)$ désigne le groupe dérivé de G, *id est* le sous-groupe de G engendré par les commutateurs.

Démonstration : Soit $\chi \in \hat{G}$. Puisque $D(G) \subset \ker \chi$, la propriété universelle du quotient nous fournit un diagramme commutatif :

$$\begin{tikzcd} G \arrow[rr, "\chi"] \arrow[dr, "\pi"'] & & \mathbb{C}^* \\ & {}^{G}/_{D(G)} \arrow[ur, dashed, "\bar{\chi}"'] & \end{tikzcd}$$

Alors l'application

$$\Phi : \left(\begin{array}{ccc} \hat{G} & \longrightarrow & \widehat{{}^{G}/_{D(G)}} \\ \chi & \longmapsto & \bar{\chi} \end{array} \right)$$

est un isomorphisme. On note dans cette démonstration $[x]_\partial$ la classe de x dans ${}^{G}/_{D(G)}$:

$$\forall x \in G, [x]_\partial \overset{\text{déf.}}{=} xD(G)$$

Montrons que Φ est un morphisme de groupes. Soient $\chi, \psi \in \hat{G}$. Alors pour tout $x \in G$,

$$\overline{\chi \cdot \psi}\,([x]_\partial) = (\chi \cdot \psi)(x)$$

Par définition du produit de $\hat{G}$, on a alors

$$\overline{\chi \cdot \psi}\,([x]_\partial) = \chi(x)\psi(x)$$

Soit

$$\overline{\chi \cdot \psi}\,([x]_\partial) = \bar{\chi}\,([x]_\partial)\,\bar{\psi}\,([x]_\partial)$$

Donc Φ est bien un morphisme de groupes. Montrons que Φ est injective. Soit $\chi \in \hat{G}$ tel que

$$\forall [x]_\partial \in {}^{G}/_{D(G)}, \bar{\chi}([x]_\partial) = 1$$

Alors, on a simplement par définition de $\bar{\chi}$,

$$\forall x \in G, \chi(x) = 1$$

donc χ est le caractère trivial, don Φ est injective. Montrons que Φ est surjective. Soit $\xi \in \widehat{{}^{G}/_{D(G)}}$. On définit le caractère $\chi \in \hat{G}$ par :

$$\forall x \in G, \chi(x) \overset{\text{déf.}}{=} \xi \circ \pi(x)$$

χ ainsi défini est bien un caractère. Alors pour tout $[x]_\partial \in {}^{G}/_{D(G)}$,

$$\bar{\chi}([x]_\partial) = \chi(x) = \xi([x]_\partial)$$

donc $\bar{\chi} = \xi$, donc Φ est bien surjective. Φ est donc un isomorphisme. Puisque ${}^{G}/_{D(G)}$ est abélien, il suit que $\hat{G} \simeq {}^{G}/_{D(G)}$. □

Pour déterminer la structure de $\widehat{\mathfrak{S}_n}$, rappelons quelques propriétés sur le groupe symétrique.

Proposition 33.1 : Noyau de la signature

On note $\mathfrak{A}_n$ le noyau de $\varepsilon \in \mathrm{Hom}_{\mathrm{gr}}(\mathfrak{S}_n, \{\pm 1\})$. Alors $\mathfrak{A}_n$ est engendré par les cycles de longueur 3.

Démonstration : • $\mathfrak{S}_n$ est engendré par les transpositions de type $(1, i)$. En effet, toute permutation σ se décompose en un produit unique à l'ordre des facteurs près de cycles disjoints. Tout cycle s'exprime comme produit de transpositions. Enfin, si $(i \quad j)$ est une transposition alors

$$(i \quad j) = (1 \quad i)(1 \quad j)(1 \quad i)$$

• La signature d'une permutation vaut -1. Si $\sigma \in \mathfrak{A}_n$, alors $\varepsilon(\sigma) = 1$, donc σ est engendré par un nombre pair de transpositions de type $(1 \quad i)$. Concluons alors en constatant que

$$(1 \quad i)(1 \quad j) = (1 \quad i \quad j)$$

qui est un 3-cycle. □

Proposition 33.2 : Groupe dérivé du groupe symétrique

On a

$$D(\mathfrak{S}_n) = \mathfrak{A}_n$$

Démonstration : Puisque ε est un caractère, tout commutateur est envoyé dans le noyau de ε, donc $D(\mathfrak{S}_n) \subset \mathfrak{A}_n$. Puisque $\mathfrak{A}_n$ est engendré par les 3-cycles, montrons que tout 3-cycle appartient au groupe dérivé. Si $\sigma = (a \quad b \quad c)$ est un 3-cycle, alors

$$\sigma^2 = (a \quad c \quad b)$$

Puisque deux permutations de même longueur sont conjuguées dans $\mathfrak{S}_n$, il existe $\tau \in \mathfrak{S}_n$ tel que

$$\sigma^2 = \tau\sigma\tau^{-1}$$

Autrement dit,

$$\sigma = \tau\sigma\tau^{-1}\sigma^{-1} = [\tau, \sigma] \in D(\mathfrak{S}_n)$$

D'où l'autre inclusion. □

Théorème 33.2 : Dual du groupe symétrique

Le dual du groupe symétrique est donné par :

$$\widehat{\mathfrak{S}_n} \simeq {}^{\mathbb{Z}}/_{2\mathbb{Z}}$$

Démonstration : D'après le théorème **33.1**, et par la proposition **33.2**,

$$\widehat{\mathfrak{S}_n} \simeq {}^{\mathfrak{S}_n}/_{\mathfrak{A}_n}$$

Or, d'après le théorème d'isomorphisme,

$$^{\mathfrak{S}_n}/_{\mathfrak{A}_n} \simeq \mathrm{Im}(\varepsilon) = \{\pm 1\}$$

donc, on a bien

$$\widehat{\mathfrak{S}_n} \simeq {}^{\mathbb{Z}}/_{2\mathbb{Z}}$$

□

34 Répartition de probabilités dans un groupe abélien fini

Référence : Gabriel PEYRÉ, *Algèbre discrète de la transformée de* FOURIER.

Cette section utilise la transformée de FOURIER sur un groupe abélien fini, et de quelques propriétés.

Définition 34.1

Soit G un groupe fini. On note $\mathbb{C}[G]$ l'ensemble des fonctions de G à valeurs dans $\mathbb{C}$. Pour $f \in \mathbb{C}[G]$, on définit sa *transformée de FOURIER* par :

$$\forall \chi \in \hat{G}, \hat{f}(\chi) \overset{\text{déf.}}{=} \sum_{g \in G} f(g)\chi(g)$$

Proposition 34.1 : Formule de PLANCHEREL

Soit G un groupe fini abélien. Soient $f, g \in \mathbb{C}[G]$. Alors

$$\sum_{x \in G} f(x)\overline{g(x)} = \frac{1}{|G|} \sum_{\chi \in \hat{G}} \hat{f}(\chi)\overline{\hat{g}(\chi)}$$

Démonstration : • Par définition de la transformée de Fourier,

$$\sum_{\chi \in \hat{G}} \hat{f}(\chi)\overline{\hat{g}(\chi)} = \sum_{\chi \in \hat{G}} \left[\sum_{x \in G} f(x)\chi(x)\right]\left[\sum_{y \in G} \overline{g(y)\chi(y)}\right]$$

On réorganise pour avoir

$$\sum_{\chi \in \hat{G}} \hat{f}(\chi)\overline{\hat{g}(\chi)} = \sum_{x,y \in G} f(x)\overline{g(y)} \left[\sum_{\chi \in \hat{G}} \chi(x)\overline{\chi(y)}\right]$$

• On a le résultat suivant : si J est l'application

$$\left(\begin{array}{ccc} G & \longrightarrow & \hat{\hat{G}} \\ x & \longmapsto & J_x : \left(\begin{array}{ccc} \hat{G} & \longrightarrow & \mathbb{C}^* \\ \chi & \longmapsto & \chi(x) \end{array}\right) \end{array}\right)$$

alors J est un isomorphisme. Ainsi,

$$\sum_{\chi \in \hat{G}} \chi(x)\overline{\chi(y)} = \sum_{\chi \in \hat{G}} J_x(\chi)\overline{J_y(\chi)}$$

Puisque G est un groupe fini, χ est à valeurs dans $\mathbb{U}_n$, donc

$$\overline{\chi(y)} = \frac{1}{\chi(y)}$$

Et donc,

$$\sum_{\chi\in\hat{G}} \chi(x)\overline{\chi(y)} = \sum_{\chi\in\hat{G}} \frac{J_x(\chi)}{J_y(\chi)}$$

- On note $\phi = \frac{J_x}{J_y} \in \hat{\hat{G}}$. Soit $\psi \in \hat{G}$. Alors

$$\phi(\psi)\sum_{\chi\in\hat{G}} \phi(\chi) = \sum_{\chi\in\hat{G}} \phi(\psi\chi)$$

Puisque $\chi \longmapsto \psi\chi$ est une bijection,

$$\phi(\psi)\sum_{\chi\in\hat{G}} \phi(\chi) = \sum_{\chi\in\hat{G}} \phi(\chi)$$

D'où

$$(1-\phi(\psi))\sum_{\chi\in\hat{G}} \phi(\chi) = 0$$

S'il existe $\psi \in \hat{G}$ tel que $\phi(\psi) \neq 1$, alors $\sum_{\chi\in\hat{G}} \phi(\chi) = 0$. Sinon, ϕ est le caractère trivial. Or, ϕ est ce caractère si et seulement si $J_x = J_y$, et pusique J est un isomorphisme, si et seulement si $x = y$. Par conséquent,

$$\sum_{\chi\in\hat{G}} \chi(x)\overline{\chi(y)} = \delta_{x,y}\left|\hat{G}\right|$$

- Reprenons notre calcul :

$$\sum_{\chi\in\hat{G}} \hat{f}(\chi)\overline{\hat{g}(\chi)} = \sum_{x,y\in G} f(x)g(y)\left[\sum_{\chi\in\hat{G}} \chi(x)\overline{\chi(y)}\right]$$

donne alors

$$\sum_{\chi\in\hat{G}} \hat{f}(\chi)\overline{\hat{g}(\chi)} = \left|\hat{G}\right|\sum_{x\in G} f(x)\overline{g(x)}$$

C'est donc seulement ici qu'on utilise l'hypothèse de commutativité de G : $|G| = \left|\hat{G}\right|$. On a ainsi obtenu la formule de Plancherel. □

Proposition 34.2 : Estimation d'écart avec la probabilité uniforme

Soit G un groupe fini abélien. On considère $P : G \longrightarrow \mathbb{R}_+$ une fonction de répartition d'une loi de probabilité sur G *id est* $\sum_{g\in G} P(g) = 1$. Si $\chi_0 \in \hat{G}$ est la caractère trivial, alors

$$\forall g \in G, \left|P(g) - \frac{1}{|G|}\right|^2 \leqslant \frac{1}{|G|}\sum_{\substack{\chi\in\hat{G}\\ \chi\neq\chi_0}} \left|\hat{P}(\chi)\right|^2$$

Démonstration : • On note U la fonction de répartition uniforme sur G :

$$\forall g \in G, U(g) \overset{\text{déf.}}{=} \frac{1}{|G|}$$

La quantité que l'on cherche à estimer est alors $|P(g) - U(g)|^2$. Soit $g \in G$. On fait alors la majoration brutale suivante :

$$|P(g) - U(g)|^2 \leqslant \sum_{x\in G} |P(x) - U(x)|^2$$

D'après la formule de Plancherel, avec $f = g$, on obtient :

$$\sum_{x\in G} |P(x) - U(x)|^2 = \frac{1}{|G|}\sum_{\chi\in\hat{G}} \left|\hat{P}(\chi) - \hat{U}(\chi)\right|^2$$

- Déterminons $\hat{P}$ et $\hat{U}$. Pour tout $\chi \in \hat{G}$,

$$\hat{U}(\chi) = \frac{1}{|G|}\sum_{g\in G} \chi(g)$$

Or, pour $\chi \neq \chi_0$, $\sum_{g\in G} \chi(g) = 0$ (même calcul que dans la proposition précédente), d'où

$$\hat{U}(\chi) = \delta_{\chi,\chi_0}$$

De plus, en χ_0, on a

$$\hat{P}(\chi_0) = \sum_{g\in G} P(g) = 1$$

Ainsi,

$$\frac{1}{|G|}\sum_{\chi\in\hat{G}} \left|\hat{P}(\chi) - \hat{U}(\chi)\right|^2 = \frac{1}{|G|}\sum_{\substack{\chi\in\hat{G}\\ \chi\neq\chi_0}} \left|\hat{P}(\chi)\right|^2$$

D'où la conclusion :

$$\forall g \in G, \left|P(g) - \frac{1}{|G|}\right|^2 \leqslant \frac{1}{|G|}\sum_{\substack{\chi\in\hat{G}\\ \chi\neq\chi_0}} \left|\hat{P}(\chi)\right|^2$$

□

35 Transformée de Fourier d'un groupe cyclique et multiplication rapide des polynômes

Références :

- Matthieu ROMAGNY, *Représentation linéaire des groupes finis* ;
- Gabriel PEYRÉ, *Algèbre discrète de la transformée de* FOURIER ;
- Thomas CORMEN, Charles LEISERSON, Ronald RIVEST, Clifford STEIN, *Introduction à l'algorithmique* ;
- Claude GASQUET, Patrick WITOMSKI, *Analyse de* FOURIER *et applications*

Proposition 35.1 : Orthogonalité des caractères

Soit G un groupe fini abélien. Pour $\chi, \psi \in \hat{G}$, on note

$$\langle \chi, \psi \rangle \overset{\text{déf.}}{=} \frac{1}{|G|} \sum_{x \in G} \chi(x)\overline{\psi(x)}$$

Si $\hat{G} \overset{\text{déf.}}{=} \{\chi_1, \cdots, \chi_n\}$, alors

$$\forall i, j \in [\![1, n]\!], \langle \chi_i, \chi_j \rangle = \delta_{i,j}$$

En conséquent, $\hat{G}$ est une base orthonormale de $\mathbb{C}[G]$.

Démonstration : Soient $\chi, \psi \in \hat{G}$.

• Dans le cas où G est cyclique (mettons $G = {}^{\mathbb{Z}}/_{n\mathbb{Z}}$), les caractères sont donnés par :

$$\forall j \in [\![0, n-1]\!], \forall [x]_n \in G, \chi_j(x) = \omega^{jx}$$

Ainsi, dans ce cas,

$$\langle \chi_i, \chi_j \rangle = \frac{1}{n} \sum_{[x]_n \in G} \omega^{(i-j)x}$$

Autrement dit,

$$\langle \chi_i, \chi_j \rangle = \frac{1}{n} \sum_{\gamma=0}^{n-1} \omega^{(i-j)\gamma}$$

Cette somme est une somme géométrique nulle lorsque $i \neq j$, et est une somme constante qui vaut n si $i = j$, donc on a bien

$$\langle \chi_i, \chi_j \rangle = \delta_{i,j}$$

• Dans le cas général d'un groupe fini abélien, on se donne $\chi, \psi \in \hat{G}$. Par définition,

$$\langle \chi, \psi \rangle = \frac{1}{|G|} \sum_{x \in G} \chi(x)\overline{\psi(x)}$$

Puisque G est fini, $\psi(x) \in \mathbb{U}_{|G|}$, donc son inverse est son conjugué, donc

$$\langle \chi, \psi \rangle = \frac{1}{|G|} \sum_{x \in G} \frac{\chi(x)}{\psi(x)}$$

On note $\xi \overset{\text{déf.}}{=} \frac{\chi}{\psi} \in \hat{G}$. Pour tout $g \in G$,

$$\xi(g) \sum_{x \in G} \xi(x) = \sum_{x \in G} \xi(gx)$$

et puisque $x \longmapsto gx$ est une bijection, il suit

$$\xi(g) \sum_{x \in G} \xi(x) = \sum_{x \in G} \xi(x)$$

Soit, pour tout $g \in G$,

$$(1 - \xi(g)) \sum_{x \in G} \xi(x) = 0$$

Si ξ n'est pas le morphisme trivial, il existe $g \in G$ tel que $\xi(g) \neq 1$. Pour ce g, on a alors

$$\sum_{x \in G} \xi(x) = 0$$

Or, ξ est le caractère trivial si et seulement si $\chi = \psi$, et donc

$$\langle \chi, \psi \rangle = \frac{1}{|G|} \sum_{x \in G} \delta_{\chi,\psi} = \delta_{\chi,\psi}$$

ce qui montre bien l'orthogonalité des caractères dans un groupe fini abélien. Cela en fait une famille libre, composé de n éléments, dans un espace vectoriel de dimension n (car de base $\{\delta_g\}_{g\in G}$), donc il s'agit bien d'une base de $\mathbb{C}[G]$. $\square$

Proposition 35.2 : Formule d'inversion de Fourier

Soit G un groupe fini abélien. Soit $f \in \mathbb{C}[G]$. Alors :

$$f = \sum_{\chi\in\hat{G}} c_f(\chi)\chi = \frac{1}{|G|}\sum_{\chi\in\hat{G}} \frac{\hat{f}(\chi)}{\chi}$$

où

$$\forall \chi \in \hat{G}, c_f(\chi) \overset{\text{déf.}}{=} \frac{1}{|G|}\sum_{x\in G} f(x)\overline{\chi(x)}$$

Démonstration : Puisque $\hat{G}$ est une base de $\mathbb{C}[G]$, le théorème dé composition dans une base orthonormale donne alors directement

$$f = \sum_{\chi\in\hat{G}} \langle f, \chi\rangle \chi$$

d'où

$$f = \sum_{\chi\in\hat{G}} c_f(\chi)\chi$$

Pour la deuxième somme, on a :

$$\sum_{\chi\in\hat{G}} c_f(\chi)\chi = \frac{1}{|G|}\sum_{x\in G} f(x)\sum_{\chi\in\hat{G}} \chi\overline{\chi(x)}$$

Or,

$$\left(\begin{array}{ccc} \hat{G} & \longrightarrow & \hat{G} \\ \chi\chi(x) & \longmapsto & \bar{\chi}\chi(x) \end{array}\right)$$

est une bijection, d'où

$$\sum_{\chi\in\hat{G}} c_f(\chi)\chi = \frac{1}{|G|}\sum_{x\in G} f(x)\sum_{\chi\in\hat{G}} \bar{\chi}\chi(x)$$

Soit,

$$\sum_{\chi\in\hat{G}} c_f(\chi)\chi = \frac{1}{|G|}\sum_{\chi\in\hat{G}} \frac{\hat{f}(\chi)}{\chi}$$

$\square$

La propriété suivante exprime le constat suivant : la transformation de Fourier d'un groupe cyclique s'identifie avec l'évaluation des polynômes sur les racines de l'unité. La transformation de Fourier inverse s'identifie avec l'interpolation de Lagrange sur les racines de l'unité.

Proposition 35.3 : Transformée de Fourier et interpolation de Lagrange

Soit G un groupe cyclique fini d'ordre N. Il existe deux isomorphismes α et β qui font commuter le diagramme commutatif suivant :

$$\begin{tikzcd}
(\mathbb{C}[G], *) \arrow[r, "\mathcal{F}", "\sim"'] \arrow[d, "\alpha"'] & \left(\mathbb{C}\left[\hat{G}\right], \cdot\right) \arrow[d, "\beta"] \\
\mathbb{C}[X]/\langle X^N - 1\rangle \arrow[r, "\text{TRC}", "\sim"'] & \prod_{k=0}^{N-1} \mathbb{C}
\end{tikzcd}$$

où TRC est le morphisme issu du théorème des restes chinois, et le produit de convolution $*$

sur $\mathbb{C}[G]$ est défini par

$$\forall x \in G, f * g(x) \overset{\text{déf.}}{=} \sum_{y \in G} f(y) g(xy^{-1})$$

Démonstration : • L'application

$$\mathcal{F} : \left(\begin{array}{ccc} (\mathbb{C}[G], *) & \longrightarrow & \left(\mathbb{C}\left[\hat{G}\right], \cdot\right) \\ f & \longmapsto & \hat{f} \end{array} \right)$$

est un isomorphisme de groupes. Il s'agit d'un morphisme : si $f, g \in \mathbb{C}[G]$, pour tout $\chi \in \hat{G}$, on a par définition de la transformée de Fourier,

$$\widehat{f * g}(\chi) = \sum_{x \in G} f * g(x) \chi(x)$$

Puis, par définition de la convolution,

$$\widehat{f * g}(\chi) = \sum_{x \in G} \left[\sum_{y \in G} f(y) g(xy^{-1}) \right] \chi(x)$$

On intervertit les sommes :

$$\widehat{f * g}(\chi) = \sum_{y \in G} f(y) \left[\sum_{x \in G} g(xy^{-1}) \chi(x) \right]$$

On fait le changement de variables $z = xy^{-1}$ dans le somme en x :

$$\widehat{f * g}(\chi) = \sum_{y \in G} f(y) \left[\sum_{z \in G} g(z) \chi(yz) \right]$$

Enfin, χ est un caractère, donc

$$\widehat{f * g}(\chi) = \left[\sum_{y \in G} f(y) \chi(y) \right] \left[\sum_{z \in G} g(z) \chi(z) \right]$$

Soit encore

$$\widehat{f * g}(\chi) = \hat{f}(\chi) \cdot \hat{g}(\chi)$$

d'où $\widehat{f * g} = \hat{f} \cdot \hat{g}$, donc $\mathcal{F}$ est un morphisme de groupes. Montrons qu'il s'agit d'un isomorphisme d'espaces vectoriels. Il suffit pour cela de montrer l'injectivité de $\mathcal{F}$, et de conclure par égalité des dimensions. Soit $f \in \mathbb{C}[G]$ tel que $\hat{f} = 0$. D'après la formule d'inversion de Fourier,

$$f = \frac{1}{|G|} \sum_{\chi \in \hat{G}} \hat{f}(\chi) \chi^{-1}$$

d'où $f = 0$, donc $\mathcal{F}$ est injective. Il s'agit donc bien d'un isomorphisme d'algèbres entre $(\mathbb{C}[G], *)$ et $\left(\mathbb{C}\left[\hat{G}\right], \cdot\right)$.

• Intéressons-nous au morphisme TRC. On note ζ une racine N-ème de l'unité, avec $N \in \mathbb{N}^*$. On dispose du morphisme d'évaluation suivant :

$$\varepsilon_\zeta : \left(\begin{array}{ccc} \mathbb{C}[X] & \longrightarrow & \mathbb{C}^N \\ P & \longmapsto & (P(1), P(\zeta), \cdots, P(\zeta^{N-1})) \end{array} \right)$$

Le noyau de ε_ζ est $\langle X^N - 1 \rangle$, l'idéal engendré par $X^N - 1$, et est surjectif. Par le théorème d'isomorphisme, il suit que

$$\mathbb{C}[X] / \langle X^N - 1 \rangle \simeq \mathbb{C}^N$$

Le morphisme est donné par :

$$\text{TRC}_\zeta : \left(\begin{array}{ccc} \mathbb{C}[X] / \langle X^N - 1 \rangle & \longrightarrow & \mathbb{C}^N \\ [P]_{\langle X^N - 1 \rangle} & \longmapsto & (P(1), P(\zeta) \cdots, P(\zeta^{n-1})) \end{array} \right)$$

On l'appelle aussi *transformation de Fourier discrète*, et on la note DFT_ζ. Son application inverse est l'interpolation de Lagrange. Si $(a_0, \cdots, a_{N-1}) \in \mathbb{C}^N$, alors on définit

$$P(X) = \sum_{k=0}^{n-1} a_k L_k(X)$$

avec

$$L_k(X) \overset{\text{déf.}}{=} \prod_{j \neq k} \frac{X - \zeta^j}{\zeta^k - \zeta^j}$$

La classe de P modulo $X^N - 1$ permet de déterminer l'antécédent de $(a_0, \cdots, a_{N-1})$ par TRC_ζ (qui est bien unique par égalité des dimensions).

• Reste à définir α et β. C'est ici que le fait que G soit cyclique est crucial : on se donne γ un générateur de G. On définit alors pour tout $k \in [\![0, N-1]\!]$,

$$\alpha(\gamma^k) \overset{\text{déf.}}{=} x^k$$

où x est la classe de X modulo $X^N - 1$.

Pour β, puisque G est cyclique, c'est aussi le cas de $\hat{G}$. Soit $\chi \in \hat{G}$ un générateur de $\hat{G}$. On définit alors ici pour tout $\varphi \in \mathbb{C}[\hat{G}]$:

$$\beta(\varphi) = (\varphi(\chi^0), \varphi(\chi), \cdots, \varphi(\chi^{n-1}))$$

avec χ^0 le morphisme trivial. β est un isomorphisme.

• Montrons que le diagramme commute effectivement. Soit $f \in \mathbb{C}[G]$. D'une part, f se décompose sur la base $\{\delta_{\gamma^k}\}_{k \in [\![0, N-1]\!]}$:

$$\alpha(f) = \sum_{k=0}^{N-1} f(\gamma^k)\alpha(\delta_{\gamma^k})$$

Par définition de α, il suit que

$$\alpha(f) = \sum_{k=0}^{N-1} f(\gamma^k)x^k$$

Ainsi, avec la description faite de TRC_ζ, il suit que

$$\mathsf{TRC}_\zeta(\alpha(f)) = \left(\sum_{k=0}^{N-1} f(\gamma^k)\zeta^{kj}\right)_{0\leqslant j\leqslant N-1}$$

Dans l'autre sens, pour tout $k \in [\![0, N-1]\!]$,

$$\hat{f}(\chi^k) = \sum_{j=0}^{N-1} f(\gamma^j)\chi^k(\gamma^j)$$

donc,

$$\hat{f}(\chi^k) = \sum_{j=0}^{N-1} f(\gamma^j)\chi(\gamma)^{kj}$$

Par la description de $\hat{G}$ lorsque G est cyclique, on a $\chi(\gamma) = \zeta$. Ainsi,

$$\hat{f}(\chi^k) = \sum_{j=0}^{N-1} f(\gamma^j)\zeta^{kj}$$

et donc

$$\beta \circ \mathcal{F}(f) = \left(\sum_{j=0}^{N-1} f(\gamma^j)\zeta^{kj}\right)_{0\leqslant k\leqslant N-1}$$

D'où $\beta \circ \mathcal{F} = \mathsf{TRC}_\zeta \circ \alpha$. □

On peut alors calculer la multiplication de deux polynômes par la méthode FFT, la transformée de Fourier rapide. Si $P, Q \in \mathbb{K}[X]$ sont de degré au plus $N-1$, de coefficients $P_0, \cdots, P_{N-1}$ et $Q_0, \cdots, Q_{N-1}$, le coût de calcul est en $O(N\ln(N))$, *via* le schéma suivant :

$$\begin{tikzcd}
\begin{array}{c}(P_0,\cdots,P_{N-1})\\(Q_0,\cdots,Q_{N-1})\end{array} \arrow[rr, "\begin{array}{c}\text{multiplication de}\\\text{polynômes}\end{array}", "O(N^2)"'] \arrow[d, "\text{évaluation}"', "O(N\ln(N))"] & & (R_0, \cdots, R_{N-1}) \\
\begin{array}{c}(P(\xi_0),\cdots,P(\xi_{N-1}))\\(Q(\xi_0),\cdots,Q(\xi_{N-1}))\end{array} \arrow[rr, "O(N)", "\begin{array}{c}\text{multiplication comp.}\\\text{par composantes}\end{array}"'] & & (R(\xi_0), \cdots, R(\xi_{N-1})) \arrow[u, "O(N\ln(N))", "\text{interpolation}"']
\end{tikzcd}$$

L'algorithme se base sur la principe suivant. Soient P et Q de degré au plus $N-1$. On calcule leur transformée de Fourier rapide qui donne l'évaluation aux points ζ^k, les racines $2N$-ème de l'unité. (Le reste de P, Q et PQ sont les mêmes dans la division euclidienne par $X^{2N}-1$.) On fait la multiplication composante par composante. On a alors à ce stade la donnée des $P(\zeta^k)Q(\zeta^k)$ avec $0 \leqslant k \leqslant 2N-1$. Par la transformée discrète inverse de Fourier, on reconstruit alors le polynôme PQ, *via* la formule d'inversion de Fourier : si c_k et le k-ème coefficient du polynôme PQ :

$$PQ(\zeta^k) = \sum_{j=0}^{2N-2} c_j\zeta^{jk}$$

Alors

$$c_k = \frac{1}{2N}\sum_{j=0}^{2N-1} P(\zeta^j)Q(\zeta^j)\zeta^{-jk}$$

36 Théorème de structure des groupes finis abéliens par la dualité

Références :

— Matthieu ROMAGNY, *Représentation linéaire des groupes finis* ;

— Pierre COLMEZ, *Éléments d'analyse et d'algèbre*

Dans cette section, on ne suppose plus connu le théorème de structures des groupes abéliens, donc la démonstration de $\hat{G} \simeq G$ faite dans les sections précédentes ne tient plus. On cherche donc ici : :

* Montrer que si G est un groupe abélien fini, G et $\hat{G}$ ont même *exposant* ;
* en déduire le théorème de structure des groupes abéliens finis.

On repart du théorème de bidualité, que l'on doit aussi redémontrer, car on avait utilisé la propriété de dualité des groupes abéliens finis.

Théorème 36.1 : Bidualité dans les groupes abéliens finis

Soit G un groupe abélien fini. L'application

$$J : \begin{pmatrix} G & \longrightarrow & \hat{\hat{G}} \\ x & \longmapsto & J_x : \begin{pmatrix} \hat{G} & \longrightarrow & \mathbb{C}^* \\ \chi & \longmapsto & \chi(x) \end{pmatrix} \end{pmatrix}$$

est un isomorphisme.

Démonstration : J est un morphisme de groupes. Reste à montrer que J est bijectif. D'une part, on a $|G| = |\hat{\hat{G}}|$, en utilisant les résultats suivants :

— Si $\mathrm{Irr}(G)$ représente l'ensemble des caractères des représentations irréductibles de G, et $\mathrm{Conj}(G)$ toutes les classes de conjugaison de G, alors

$$|\mathrm{Irr}(G)| = |\mathrm{Conj}(G)|$$

— Si G est abélien, toute représentation de G est de dimension 1 : autrement dit $\mathrm{Irr}(G) = \hat{G}$.

Dans le cadre abélien, $G = \mathrm{Conj}(G)$, donc finalement,

$$|G| = |\mathrm{Conj}(G)| = |\mathrm{Irr}(G)| = |\hat{G}|$$

Puisque $\hat{G}$ est lui-même un groupe abélien fini, il suit donc que $|G| = |\hat{\hat{G}}|$.

Montrons que J est injective. Soit $x \in G$. Supposons que pour tout $\chi \in \hat{G}$, $\chi(x) = 1$. Constatons alors que pour tout $a \in G$,

$$\hat{\delta}_a(\chi) = \frac{1}{|G|}\chi(a)$$

donc si pour tout $\chi \in \hat{G}$, $\chi(x) = 1$, alors $\hat{\delta}_x = \frac{1}{|G|}$. Par formule d'inversion de Fourier, pour tout $a \in G$,

$$\frac{1}{|G|}\sum_{\chi \in \hat{G}} \chi(x)\overline{\chi(a)} = \delta_x(a)$$

Or, pour $a = 1_G$, cela donne $\delta_x(1_G) = 1_G$. D'où $x = 1_G$. D'où l'injectivité de J. On a donc bien montré que J est un isomorphisme. □

Définition 36.1

Soit G un groupe fini abélien. On appelle *exposant* du groupe G l'entier $N(G)$ défini par :

$$N(G) \overset{\text{déf.}}{=} \max\left\{n \in \mathbb{N}^* \,|\forall x \in G, x^n = 1_G\right\}$$

Par le théorème de Lagrange, tout ordre existant dans G divise son exposant.

Proposition 36.1 : Exposant du dual d'un groupe fini abélien

Soit G un groupe fini abélien. Alors G et $\hat{G}$ ont même exposant :

$$N(G) = N\left(\hat{G}\right)$$

Démonstration : Soit $\chi \in \hat{G}$. Pour tout $x \in G$,

$$\chi^{N(G)}(x) = \chi\left(x^{N(G)}\right) = 1$$

Donc $\chi^{N(G)} = \chi_0$, le caractère trivial. Ainsi, $N(G) \leqslant N(\hat{G})$. De plus, puisque $\hat{G}$ est un groupe fini cyclique, on a aussi

$$N(\hat{G}) \leqslant N\left(\hat{\hat{G}}\right)$$

Puis, par le théorème **36.1**, $\hat{\hat{G}} \simeq G$, donc

$$N\left(\hat{\hat{G}}\right) = N(G)$$

Il suit que

$$N(G) \leqslant N(\hat{G}) \leqslant N(G)$$

soit $N(G) = N(\hat{G})$. □

Théorème 36.2 : Théorème de structure des groupes abéliens finis

Soit G un groupe fini abélien. Il existe un unique entier $r \in \mathbb{N}$ et des entiers $N_1, \cdots, N_r > 0$ tels que $N_1 | \cdots | N_r$ et

$$G \simeq {}^{\mathbb{Z}}/_{N_1\mathbb{Z}} \oplus \cdots \oplus {}^{\mathbb{Z}}/_{N_r\mathbb{Z}}$$

En particulier, N_1 est l'exposant de G.

Si G est un groupe, et H, K deux sous-groupes de G, on dit que $G = H \oplus K$ (somme directe) si $H \cap K = \{1_G\}$ et pour tout $g \in G$, il existe $h \in H$, $k \in K$ tels que $g = hk$.

Démonstration : On démontre ce résultat par récurrence(forte) sur n, l'ordre de G.

Pour $n = 1$, G est le groupe trivial, il est d'exposant 1, et est isomorphe à ${}^{\mathbb{Z}}/_{\mathbb{Z}}$, ce qui montre le résultat pour $n = 1$.

Supposons que $|G| > 1$. On note $N_1 \overset{\text{déf.}}{=} N(G)$ l'exposant de G. Alors N_1 est aussi l'exposant de $\hat{G}$. Soit $\chi_1 \in \hat{G}$ d'exposant N_1. Pour tout $x \in G$, $\chi_1(x)$ est une racine N_1-ème de l'unité, donc $\chi_1(G)$ est un sous-groupe de $\mathbb{U}_{N_1}$. Puisque $\mathbb{U}_{N_1}$ est cyclique, il suit que

$$\chi(G) = \mathbb{U}_{N_1}$$

Par conséquent, en notant $\zeta = e^{\frac{2\imath\pi}{N_1}}$, il existe $x_1 \in G$ tel que

$$\chi_1(x_1) = \zeta$$

L'ordre de x_1 divise N_1, par définition de l'exposant, et par cette égalité, il suit que N_1 est lui inférieur ou égal à l'ordre de x_1. Donc $H_1 \overset{\text{déf.}}{=} \langle x_1 \rangle$ est d'ordre N_1. Ainsi,

$$H_1 \simeq {}^{\mathbb{Z}}/_{N_1\mathbb{Z}}$$

Montrons que

$$G = H_1 \oplus \ker \chi_1$$

D'une part, ${\chi_1}_{|_H}$ est un morphisme surjectif qui envoie un ensemble de cardinal N_1 sur un ensemble de cardinal N_1, donc χ_1 est un isomorphisme. Si $\alpha : \mathbb{U}_n \longrightarrow H_1$ est l'inverse de χ_1, et $x \in G$, alors $a \overset{\text{déf.}}{=} \alpha(\chi_1(x)) \in H_1$. Ainsi, $b \overset{\text{déf.}}{=} a^{-1}x$ vérifie

$$\chi_1(b) = \chi_1(a)^{-1}\chi_1(x) = 1$$

donc $b \in \ker(\chi_1)$. Ainsi, $x = ab$, donc $G = H_1 \cdot \ker(\chi_1)$. Montrons queleur intersection est réduite à $\{1_G\}$. Pour cela, on considère $x \in H_1 \cap \ker(\chi_1)$. Alors $\chi_1(x) = 1$. Or, χ_1 est injectif sur H_1, donc $x = 1_G$. On a donc bien montré que

$$G = H_1 \oplus \ker \chi_1 \simeq {}^{\mathbb{Z}}/_{N_1\mathbb{Z}} \oplus \ker \chi_1$$

L'hypothèse de récurrence appliqué à $\ker(\chi_1)$ permet alors de conclure cette récurrence, et la démonstration du théorème de structure des groupes abéliens finis. □

37 Construction de variables aléatoires iid de loi uniforme

Référence : Analyse pour l'agrégation, Hervé Queffélec, Claude Zuily

Théorème 37.1

Soit $(\Omega, \mathcal{F}, \mathbb{P}) = ([0,1[, \mathcal{B}([0,1[), \lambda)$ l'espace probabilisé sur $[0,1[$. On fixe $r \geqslant 2$ un entier, et on note $A = [\![0, r-1]\!]$. Il existe une suite $(\varepsilon_n)_{n\in\mathbb{N}} \in A^{\mathbb{N}}$ de variables aléatoires telle que cette suite soit λ-indépendante, et identiquement distribuée de loi uniforme sur A :

$$\forall a \in A, \lambda(\varepsilon_1 = a) = \frac{1}{|A|}$$

Démonstration : La construction des variables aléatoires $(\varepsilon_n)_{n\in\mathbb{N}}$ repose sur la décomposition en base r de tout élément de Ω. Pour $x \in \Omega$, on note donc

$$x \overset{\text{déf.}}{=} \sum_{n=0}^{+\infty} \frac{\varepsilon_n(x)}{r^n}$$

avec $\varepsilon_n(x) \in [\![0, r-1]\!] = A$.

1. Constatons que dans cette écriture, il existe une infinité de n pour lesquels $\varepsilon_n(x) < r-1$, sans quoi le développement de x ne serait pas propre. Autrement dit,

$$\liminf_{n\to+\infty} \varepsilon_n(x) < r-1$$

On note

$$B \overset{\text{déf.}}{=} \left\{ \omega = (\varepsilon_n)_{n\in\mathbb{N}} \in A^{\mathbb{N}} \,\middle|\, \liminf_{n\to+\infty} \varepsilon_n < r-1 \right\}$$

B est non vide, la suite nulle en est un élément. Montrons l'unicité d'un tel développement en base r.

2. On note pour cela pour $\omega \in B$:

$$X_p(\omega) \overset{\text{déf.}}{=} \sum_{n=1}^{p} \frac{\varepsilon_n}{r^n}$$

et

$$X(\omega) \overset{\text{déf.}}{=} \sum_{n=1}^{+\infty} \frac{\varepsilon_n}{r^n}$$

Constatons que $(X_p(\omega))_p$ converge de façon croissante vers $X(\omega)$. Montrons que la suite définie par

$$X_p(\omega) + \frac{1}{r^p}$$

converge en décroissant vers $X(\omega)$ de façon non stationnaire ; autrement dit, on a :

$$\forall p \in \mathbb{N}, X_p(\omega) \leqslant X(\omega) < X_p(\omega) + \frac{1}{r^p}$$

Pour cela, observons que

$$X_{p+1}(\omega) - X_p(\omega) = \frac{\varepsilon_{p+1}}{r^{p+1}}$$

Donc

$$X_{p+1}(\omega) - X_p(\omega) \leqslant \frac{r-1}{r^{p+1}}$$

avec inégalité stricte pour une infinité de p puisque $\omega \in B$. Par conséquent, on obtient

$$X_{p+1}(\omega) + \frac{1}{r^{p+1}} \leqslant X_p(\omega) + \frac{1}{r^p}$$

avec inégalité stricte pour une infinité de p. Ce qui montre bien la convergence non stationnaire et décroissante vers $X(\omega)$.

3. On conclut donc ici en l'injectivité de $\omega \longmapsto X(\omega)$ sur B. Si $\omega \neq \omega'$, avec $\omega = (\varepsilon_n)_n$ et $\omega' = (\delta_n)_n$, on considère

$$p \overset{\text{déf.}}{=} \min \{n \in \mathbb{N}, \delta_n \neq \varepsilon_n\}$$

Alors, en supposant que $\varepsilon_p < \delta_p$:

$$X_p(\omega) = X_{p-1}(\omega) + \frac{\varepsilon_p}{r^p}$$

Par définition de p :

$$X_p(\omega) = X_{p-1}(\omega') + \frac{\varepsilon_p}{r^p}$$

Puis, toujours par définition de p :

$$X_p(\omega) \leqslant X_{p-1}(\omega') + \frac{\delta_p - 1}{r^p}$$

Autrement dit,

$$X_p(\omega) \leqslant X_p(\omega') - \frac{1}{r^p}$$

Or, d'après **2.**,

$$X(\omega) < X_p(\omega) + \frac{1}{r^p}$$

Par le calcul que l'on vient de faire,

$$X_p(\omega) + \frac{1}{r^p} \leqslant X_p(\omega')$$

et par convergence croissante de $X_p(\omega')$:

$$X_p(\omega') \leqslant X(\omega')$$

d'où lorsque $\omega \neq \omega'$, en supposant $\varepsilon_p < \delta_p$:

$$X(\omega) < X(\omega')$$

d'où l'unicité de la décomposition, qui sera cruciale dans les prochains points.

4. Pour $a_1, \cdots, a_N \in A$, on note

$$E \overset{\text{déf.}}{=} \{\varepsilon_1 = a_1, \cdots, \varepsilon_N = a_N\}$$

et

$$F \overset{\text{déf.}}{=} \left[\sum_{n=1}^{N} \frac{a_n}{r^n}, \sum_{n=1}^{N} \frac{a_n}{r^n} + \frac{1}{r^N}\right[$$

Montrons que $E = F$.

▷ Si $x \in E$, et $\omega \overset{\text{déf.}}{=} (\varepsilon_n(x))_{n\in\mathbb{N}}$ alors $x = X(\omega)$. En particulier,

$$X_N(\omega) \leqslant x < X_N(\omega) + \frac{1}{r^N}$$

par tout ce qui a pu être développé en **2.**. Donc, on a bien

$$\sum_{n=1}^{N} \frac{a_n}{r^n} \leqslant x < \sum_{n=1}^{N} \frac{a_n}{r^n} + \frac{1}{r^N}$$

soit $x \in F$.

▷ Pour $x \in F$, il existe $y \in \Omega$ tel que

$$x = \sum_{n=1}^{N} \frac{a_n}{r^n} + \frac{y}{r^N}$$

Puisque $y \in \Omega$, il admet un unique développement propre en base r :

$$y = \sum_{p=1}^{+\infty} \frac{\delta_p}{r^p}$$

donc

$$x = \sum_{n=1}^{N} \frac{a_n}{r^n} + \sum_{p=1}^{+\infty} \frac{\delta_p}{r^{p+N}}$$

par unicité de la décomposition, montrée en **3.**, il suit que

$$\varepsilon_n(x) = \begin{cases} a_n & \text{si} \quad 1 \leqslant n \leqslant N \\ \delta_{n-N} & \text{si} \quad n > N \end{cases}$$

d'où $x \in E$. D'où $E = F$.

5. Concluons alors la démonstration de ce théorème. D'après **4.**, $E = F$, donc

$$\mathbb{P}\left(\bigcap_{k=1}^{N} \varepsilon_k = a_k\right) = \lambda(F)$$

Donc,

$$\mathbb{P}\left(\bigcap_{k=1}^{N} \varepsilon_k = a_k\right) = \frac{1}{r^N}$$

De plus, pour $a_1 \in A$ fixé,

$$\mathbb{P}(\varepsilon_1 = a_1) = \sum_{a_2, \cdots, a_N \in A} \mathbb{P}\left(\bigcap_{k=1}^{N} \varepsilon_k = a_k\right)$$

donc,

$$\mathbb{P}(\varepsilon_1 = a_1) = \frac{|A|^{N-1}}{r^N}$$

soit

$$\mathbb{P}(\varepsilon_1 = a_1) = \frac{1}{r} = \frac{1}{|A|}$$

de même pour tous les ε_k. Ainsi, $(\varepsilon_n)_n$ est une suite de variables aléatoires identiquement distribuées de loi uniforme sur A. De plus,

$$\mathbb{P}\left(\bigcap_{k=1}^{N} \varepsilon_k = a_k\right) = \frac{1}{r^N} = \prod_{k=1}^{N} \mathbb{P}(\varepsilon_k = a_k)$$

Ces variables aléatoires sont donc mutuellement indépendantes. □

Remarque : Quand à l'existence de développement en base r, on définit pour $x \in [0, 1[$:

$$\forall n \in \mathbb{N}, a_n \overset{\text{déf.}}{=} \lfloor r^n x \rfloor - r \left\lfloor r^{n-1} x \right\rfloor$$

Alors

$$\sum_{n=1}^{N} \frac{a_n}{r^n} = \sum_{n=1}^{N} \frac{\lfloor r^n x \rfloor}{r^n} - \frac{\lfloor r^{n-1} x \rfloor}{r^{n-1}}$$

C'est une somme télescopique, donc

$$\sum_{n=1}^{N} \frac{a_n}{r^n} = \frac{\lfloor r^N x \rfloor}{r^N} - \lfloor x \rfloor$$

avec $\lfloor x \rfloor = 0$. Montrons que $\frac{\lfloor r^N x \rfloor}{r^N}$ converge vers x. On note u_n le terme général de cette suite, et on introduit

$$v_N \stackrel{\text{déf.}}{=} \frac{\lfloor r^N x \rfloor + 1}{r^N}$$

Alors on dispose de l'encadrement

$$u_N \leqslant x \leqslant v_N$$

via la définition de la partie entière. De plus,

$$v_N - u_N = \frac{1}{r^N} \xrightarrow[N \to +\infty]{} 0$$

et

$$\forall N \in \mathbb{N}^*, v_N - u_N \geqslant 0$$

Par le théorème des suites adjacentes, il suit que les suites $(u_N)_N$ et $(v_n)_N$ convergent vers une limite commune, qui n'est autre que x par l'encadrement. D'où finalement

$$x = \sum_{n=1}^{+\infty} \frac{a_n}{r^n}$$

38 Théorème de Bernstein

Référence : Analyse pour l'agrégation, Hervé Queffélec, Claude Zuily

Théorème 38.1 : Théorème de Bernstein

Soit $f : [0,1] \longrightarrow \mathbb{C}$ une fonction continue. On note ω son module de continuité :

$$\forall h > 0, \omega(h) \stackrel{\text{déf.}}{=} \sup \{|f(u) - f(v)|, |u - v| \leqslant h\}$$

Pour $n \geqslant 1$, on considère le n-ème polynôme de Bernstein associé à f :

$$\forall x \in \mathbb{R}, B_n(f, x) = B_n(x) \stackrel{\text{déf.}}{=} \sum_{k=0}^{n} \binom{n}{k} x^k (1-x)^{n-k} f\left(\frac{k}{n}\right)$$

Alors :

(1) B_n converge uniformément vers f sur $[0,1]$.

(2) Plus précisément, il existe $C > 0$ tel que

$$\|f - B_n\|_\infty \leqslant C \cdot \omega\left(\frac{1}{\sqrt{n}}\right)$$

(3) La vitesse de convergence est optimale : il existe $f \in \mathcal{C}^0([0,1], \mathbb{R})$ lipschitzienne et $\delta > 0$ tel que

$$\|f - B_n\|_\infty \geqslant \frac{\delta}{\sqrt{n}}$$

Démonstration : La remarque centrale est la suivante : si $X_1, \cdots, X_n$ sont indépendantes et identiquement distribuées de loi $b(x)$, avec $x \in [0,1]$, et S_n est la somme des X_i, alors S_n suit une loi binomiale de paramètres n, x, et par la formule de transfert :

$$\mathbb{E}\left[f\left(\frac{S_n}{n}\right)\right] = \sum_{k=0}^{n} \binom{n}{k} x^k (1-x)^{n-k} f\left(\frac{k}{n}\right) = B_n(x)$$

La loi des grands nombres nous met sur la voie.

[1.] Pour tout $x \in [0,1]$,

$$f(x) - B_n(x) = \mathbb{E}\left[f(x) - \frac{S_n}{n}\right]$$

Donc, par inégalité triangulaire,

$$|f(x) - B_n(x)| \leqslant \mathbb{E}\left[\left|f(x) - \frac{S_n}{n}\right|\right]$$

Soit $\delta > 0$. Alors

$$\begin{aligned}\mathbb{E}\left[\left|f(x) - \tfrac{S_n}{n}\right|\right] = &\quad \mathbb{E}\left[\left|f(x) - \tfrac{S_n}{n}\right| \mathbf{1}_{\{|x - \frac{S_n}{n}| \leqslant \delta\}}\right] \\ &+ \quad \mathbb{E}\left[\left|f(x) - \tfrac{S_n}{n}\right| \mathbf{1}_{\{|x - \frac{S_n}{n}| > \delta\}}\right]\end{aligned}$$

Par définition de ω, on obtient alors

$$\mathbb{E}\left[\left|f(x) - \frac{S_n}{n}\right|\right] \leqslant \omega(\delta) + 2\|f\|_\infty \mathbb{P}\left(\left|\frac{S_n}{n} - x\right| > \delta\right)$$

Or, d'après l'inégalité de Bienaymé-Tchebychev,

$$\mathbb{P}\left(\left|\frac{S_n}{n} - x\right| > \delta\right) \leqslant \frac{1}{\delta^2} \mathrm{Var}\left(\frac{S_n}{n}\right)$$

Par indépendance des X_i, et parce que les X_i sont identiquement distribuées, il suit

$$\mathbb{P}\left(\left|\frac{S_n}{n} - x\right| > \delta\right) \leqslant \frac{\mathrm{Var}(X_1)}{n\delta^2}$$

X_1 suit une loi de Bernouilli, donc sa variance est inférieure ou égale à $\frac{1}{4}$. D'où

$$\mathbb{E}\left[\left|f(x) - \frac{S_n}{n}\right|\right] \leqslant \omega(\delta) + \frac{\|f\|_\infty}{2n\delta^2}$$

On a alors montré que

$$\|f - B_n\|_\infty \leqslant \omega(\delta) + \frac{\|f\|_\infty}{2n\delta^2}$$

Donc, pour tout $\delta > 0$

$$\liminf_{n \to +\infty} \|f - B_n\|_\infty \leqslant \omega(\delta)$$

Or, par continuité uniforme de f sur $[0,1]$,

$$\omega(\delta) \xrightarrow[\delta \to 0^+]{} 0$$

d'où

$$\liminf_{n \to +\infty} \|f - B_n\|_\infty \leqslant 0$$

d'où

$$\|f - B_n\|_\infty \xrightarrow[n \to +\infty]{} 0$$

□

Pour le point **2.**, on étudie un peut plus en détail la fonction ω.

Lemme 38.1 : Sur le module de continuité

ω est une fonction croissante et sous-additive : pour tout $t, s \in [0,1]$ avec $t + s \in [0,1]$,

$$\omega(t+s) \leqslant \omega(t) + \omega(s)$$

En outre,

$$\forall h \in [0,1], \forall \lambda \in \left[0, \frac{1}{h}\right], \omega(\lambda h) \leqslant (\lambda + 1)\omega(h)$$

Démonstration du lemme : • Pour la croissance de ω : il suffit d'observer que pour $h_1 \leqslant h_2$:

$$\{|f(u)-f(v)|\,|\,|u-v| \leqslant h_1\} \subset \{|f(u)-f(v)|\,|\,|u-v| \leqslant h_2\}$$

donc

$$\omega(h_1) \leqslant \omega(h_2)$$

• Pour la sous-additivité, on se donne $t, s \in [0,1]$ tels que $t+s \in [0,1]$. On considère $u, v \in [0,1]$ tels que $|u-v| \leqslant t+s$. On définit

$$w \overset{\text{déf.}}{=} \frac{t}{t+s}u + \frac{s}{t+s}v$$

Alors $|u-w| \leqslant s$ et $|w-v| \leqslant t$. Ainsi,

$$|f(u)-f(v)| \leqslant |f(u)-f(w)| + |f(w)-f(v)|$$

Donc,

$$|f(u)-f(v)| \leqslant \omega(s) + \omega(t)$$

donc

$$\omega(t+s) \leqslant \omega(t) + \omega(s)$$

• Il suit par récurrence que pour tout $h \in [0,1]$ et $r \in \mathbb{N}^*$ tel que $rh \in [0,1]$,

$$\omega(rh) \leqslant r\omega(h)$$

Soit $\lambda > 0$ tel que $\lambda h \in [0,1]$. On note q la partie entière de λ :

$$q \leqslant \lambda < q+1$$

Par croissance de ω,

$$\omega(\lambda h) \leqslant \omega((q+1)h)$$

Par sous-additivité de ω :

$$\omega(\lambda h) \leqslant (q+1)\omega(h)$$

Puis, par définition de q, on a réussi à montrer que

$$\omega(\lambda h) \leqslant (\lambda+1)\omega(h)$$

□

On peut alors montrer le point **2.**

Démonstration : [2.] Par l'inégalité du lemme, avec $\lambda = \sqrt{n}$, on obtient pour $x \in [0,1]$:

$$\omega\left(\left|x - \frac{S_n}{n}\right|\right) \leqslant \left(1 + \sqrt{n}\left|x - \frac{S_n}{n}\right|\right) \cdot \omega\left(\frac{1}{\sqrt{n}}\right)$$

Or, on a :

$$|f(x) - B_n(x)| \leqslant \mathbb{E}\left[\left|f(x) - f\left(\frac{S_n}{n}\right)\right|\right]$$

Donc, par définition du module de continuité,

$$|f(x) - B_n(x)| \leqslant \mathbb{E}\left[\omega\left(\left|x - \frac{S_n}{n}\right|\right)\right]$$

Soit,

$$|f(x) - B_n(x)| \leqslant \left(1 + \sqrt{n}\,\mathbb{E}\left[\left|x - \frac{S_n}{n}\right|\right]\right) \cdot \omega\left(\frac{1}{\sqrt{n}}\right)$$

Par inégalité de Cauchy-Schwarz, on a :

$$|f(x) - B_n(x)| \leqslant \left(1 + \sqrt{n}\sqrt{\mathbb{E}\left[\left|x - \frac{S_n}{n}\right|^2\right]}\right) \cdot \omega\left(\frac{1}{\sqrt{n}}\right)$$

Or, cette quantité dans la racine n'est autre que la variance de X_1, par indépendance et parce ces variables sont identiquement distribuées. D'où

$$|f(x) - B_n(x)| \leqslant \left(1 + \sqrt{n}\sqrt{\frac{x(1-x)}{n}}\right) \cdot \omega\left(\frac{1}{\sqrt{n}}\right)$$

Enfin, $x(1-x) \leqslant \frac{1}{4}$, donc finalement,

$$|f(x) - B_n(x)| \leqslant \frac{3}{2} \cdot \omega\left(\frac{1}{\sqrt{n}}\right)$$

d'où l'estimation

$$\|f - B_n\|_\infty \leqslant \frac{3}{2} \cdot \omega\left(\frac{1}{\sqrt{n}}\right)$$

□

Pour le point **3.**, on utilise un autre lemme : l'inégalité de Khintchine.

Lemme 38.2 : Inégalité de Khintchine

Soient $\varepsilon_1, \cdots, \varepsilon_n$ des variables aléatoires indépendantes de loi de Rademacher. Pour $a_1, \cdots, a_n \in \mathbb{R}$, on note

$$X \overset{\text{déf.}}{=} \sum_{k=1}^{n} a_k \varepsilon_k$$

Alors

$$\|X\|_1 \geqslant \frac{1}{\sqrt{e}} \|X\|_2$$

Remarque : des variables de Rademacher peuvent se construire explicitement sur $\Omega = [0,1]$, muni de sa tribu borélienne et de la mesure de Lebesgue, avec

$$\varepsilon_j(t) \overset{\text{déf.}}{=} \operatorname{sgn}[\sin(2^j \pi t)]$$

Démonstration du lemme : Supposons que $\|X\|_2 = 1$:

$$\sum_{k=1}^{n} a_k^2 = 1$$

On cherche à appliquer l'inégalité suivante pour $Y \in L^\infty$:

$$|\mathbb{E}[XY]| \leqslant \|X\|_1 \|Y\|_\infty$$

On parachute alors ici une variable aléatoire Y qui va nous donner ce qu'on souhaite :

$$Y \overset{\text{déf.}}{=} \prod_{j=1}^{n} (1 + \imath \varepsilon_j a_j)$$

- Déterminons un majorant de $\|Y\|_\infty$. Pour presque tout $\omega \in \Omega$, on a

$$|Y(\omega)| = \prod_{k=1}^{n} \sqrt{1 + a_j^2}$$

Par convexité de l'exponentielle, on a :

$$|Y(\omega)| \leqslant \prod_{k=1}^{n} \sqrt{\exp\left(a_j^2\right)}$$

Donc, par propriété de l'exponentielle,

$$|Y(\omega)| \leqslant \exp\left(\frac{1}{2} \sum_{j=1}^{n} a_j^2\right) = \sqrt{e}$$

d'où $\|Y\|_\infty \leqslant \sqrt{e}$.

- Pour $j \in [\![1, n]\!]$,

$$\mathbb{E}\left[\varepsilon_j Y\right] = \mathbb{E}\left[\varepsilon_j (1 + \imath a_j \varepsilon_j) \prod_{k \neq j} (1 + \imath a_k \varepsilon_k)\right]$$

Par indépendance des variables en jeu :

$$\mathbb{E}\left[\varepsilon_j Y\right] = \mathbb{E}\left[\varepsilon_j (1 + \imath a_j \varepsilon_j)\right] \prod_{k \neq j} \mathbb{E}\left[1 + \imath a_k \varepsilon_k\right]$$

Puis,

$$\varepsilon_j (1 + \imath a_j \varepsilon_j) = \varepsilon_j + \imath a_j$$

Il suit alors que

$$\mathbb{E}\left[\varepsilon_j Y\right] = \imath a_j$$

Par linéarité, on obtient :

$$\mathbb{E}[XY] = \sum_{j=1}^{n} a_j \mathbb{E}[\varepsilon_j Y] = \imath \sum_{j=1}^{n} a_j^2 = \imath$$

- On peut alors conclure,

$$1 = |\mathbb{E}[XY]| \leqslant \|X\|_1 \|Y\|_{L^\infty} \leqslant \sqrt{e}$$

d'où

$$\|X\|_1 \leqslant \frac{1}{\sqrt{e}}$$

Dans le cas général, pour $X \neq 0$,

$$\left\| \frac{X}{\|X\|_2} \right\| \leqslant \frac{1}{\sqrt{e}}$$

d'où l'inégalité de Khintchine, qui reste vraie pour $X = 0$. □

On peut alors conclure sur le point **3**.

Démonstration : 3. La fonction que l'on va considérer est

$$\forall x \in [0,1], f(x) \overset{\text{déf.}}{=} \left| x - \frac{1}{2} \right|$$

Alors, en $\frac{1}{2}$, on a :

$$\|f - B_n\|_\infty \geqslant \left| f\left(\frac{1}{2}\right) - B_n\left(\frac{1}{2}\right) \right| = \left| B_n\left(\frac{1}{2}\right) \right|$$

Or, pour ce choix de f, on a alors

$$B_n\left(\frac{1}{2}\right) = \mathbb{E}\left[\left|\frac{S_n}{n} - \frac{1}{2}\right|\right]$$

avec S_n la somme des variables aléatoires δ_k indépendantes et identiquement distribuées de loi $b\left(\frac{1}{2}\right)$. On réduit au même dénominateur :

$$\|f - B_n\|_\infty \geqslant \frac{1}{2n}\mathbb{E}\left[|2S_n - 1|\right]$$

Autrement dit,

$$\|f - B_n\|_\infty \geqslant \frac{1}{2n}\mathbb{E}\left[\left|\sum_{k=1}^n \varepsilon_k\right|\right]$$

où les ε_k sont indépendantes et identiquement distribuées de loi de Rademacher. D'après l'inégalité de Khintchine :

$$\|f - B_n\|_\infty \geqslant \frac{1}{2\sqrt{e}n}\sqrt{\mathbb{E}\left[\left|\sum_{k=1}^n \varepsilon_k\right|^2\right]}$$

Les variables ε_k sont centrées, donc la quantité sous la racine est la variance, et par indépendance, on a :

$$\|f - B_n\|_\infty \geqslant \frac{1}{2\sqrt{e}\sqrt{n}}$$

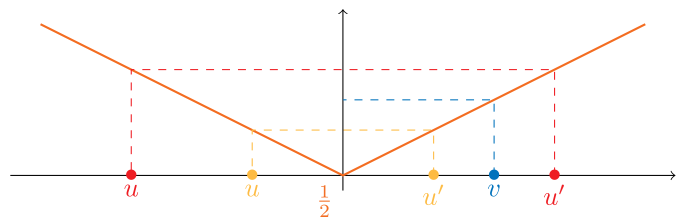
Le choix de la fonction f montre qu'ici :

$$\omega(h) = \sup_{|u-v|\leqslant h} \left\{ \left| \left| u - \frac{1}{2} \right| - \left| v - \frac{1}{2} \right| \right| \right\}$$

donc $\omega(h) \leqslant h$. Détaillons ce fait. Pour $u, v \in [0,1]$,

$$\left| \left| u - \frac{1}{2} \right| - \left| v - \frac{1}{2} \right| \right| = \begin{cases} |u - v| & \text{si} \quad \left(u - \frac{1}{2}\right)\left(v - \frac{1}{2}\right) \geqslant 0 \\ |u' - v| & \text{sinon} \end{cases}$$

avec $u' = 1 - u$ le symétrisé de u par rapport à $\frac{1}{2}$. Dans le premier cas, si $|u - v| \leqslant h$, on a bien $|f(u) - f(v)| \leqslant h$. Dans le deuxième cas, distinguons deux sous-cas.



Supposons $v > \frac{1}{2}$. Si $u < u' \leqslant v$ (cas jaune), on a

$$|v - u'| = v - u' = (v - u) - (u' - u) = |v - u| - |u' - u|$$

d'où

$$|v - u'| \leqslant |u - v| \leqslant h$$

Si $u < v \leqslant u'$, alors par définition de u' :

$$(v - u) - (u' - v) = 2v - 1 = 2\left(v - \frac{1}{2}\right) \geqslant 0$$

d'où

$$|u' - v| = u' - v \leqslant v - u = |u - v| \leqslant h$$

d'où $\omega(h) \leqslant h$. D'où l'inégalité

$$\|f - B_n\|_\infty \geqslant \frac{1}{2\sqrt{e}} \cdot \omega\left(\frac{1}{\sqrt{n}}\right)$$

□

39 Théorème de Paul Lévy

Référence : Analyse pour l'agrégation, Hervé Queffélec, Claude Zuily

Définition 39.1

Soient $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires réelles, et X une variable aléatoire réelle. On dit que $(X_n)_{n\in\mathbb{N}}$ *converge en loi* vers X si et seulement si

$$\forall f \in \mathcal{C}_b^0(\mathbb{R}), \mathbb{E}\left[f(X_n)\right] \xrightarrow[n\to+\infty]{} \mathbb{E}[f(X)]$$

Proposition 39.1 : Définitions équivalentes de la convergence en loi

Soient $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires réelles, et X une variable aléatoire réelle. Les assertions suivantes sont équivalentes :

(i) $(X_n)_n$ converge en loi vers X ;

(ii) Pour tout $f \in \mathcal{C}_0(\mathbb{R})$, l'ensemble des fonctions continues qui tendent vers 0 en $\pm\infty$,

$$\mathbb{E}\left[f(X_n)\right] \xrightarrow[n\to+\infty]{} \mathbb{E}[f(X)]$$

Autrement dit, si $(\mu_n)_{n\in\mathbb{N}}$ est une suite de probabilités sur $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ et μ un probabilité sur ce même espace, alors les propositions sont équivalentes :

(i') Pour tout $f \in \mathcal{C}_b^0(\mathbb{R})$:

$$\int_{\mathbb{R}} f(x) \text{ d}\mu_n(x) \xrightarrow[n\to+\infty]{} \int_{\mathbb{R}} f(x) \text{ d}\mu(x)$$

(ii') Pour tout $f \in \mathcal{C}_0(\mathbb{R})$:

$$\int_{\mathbb{R}} f(x) \text{ d}\mu_n(x) \xrightarrow[n\to+\infty]{} \int_{\mathbb{R}} f(x) \text{ d}\mu(x)$$

Théorème 39.1 : Théorème faible de Paul LÉVY

Soient X une variable aléatoire et $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires. Les deux propositions suivantes sont équivalentes :

(i) $(X_n)_{n\in\mathbb{N}}$ converge en loi vers X ;

(ii) Pour tout $t \in \mathbb{R}$,

$$\varphi_{X_n}(t) \xrightarrow[n\to+\infty]{} \varphi_X(t)$$

Démonstration : [(i) $\implies$ (ii)] Par hypothèse, pour toute fonction f continue bornée sur $\mathbb{R}$,

$$\mathbb{E}[f(X_n)] \xrightarrow[n\to+\infty]{} \mathbb{E}[f(X)]$$

En particulier, pour $f_t(x) = e^{\imath tx}$, on a alors

$$\varphi_{X_n}(t) = \mathbb{E}\left[e^{\imath t X_n}\right] \xrightarrow[n\to+\infty]{} \mathbb{E}\left[e^{\imath t X}\right] = \varphi_X(t)$$

d'où **(ii)**.

[(ii) $\implies$ (i)] On suppose connu le résultat suivant : l'image de $L^1(\mathbb{R})$ de la transformée de Fourier est dense dans $C^0(\mathbb{R})$ pour $\|\cdot\|_\infty$. On note $A(\mathbb{R}) \overset{\text{déf.}}{=} \mathcal{F}\left(L^1(\mathbb{R})\right)$.

- Soit $f \in A(\mathbb{R})$. Il existe $h \in L^1(\mathbb{R})$ telle que

$$\forall x \in \mathbb{R}, f(x) = \int_{\mathbb{R}} e^{\imath tx} h(t) \text{ d}t$$

Alors,

$$\mathbb{E}[f(X_n)] = \mathbb{E}\left[\int_{\mathbb{R}} e^{\imath t X_n} h(t)\ \mathrm{d}t\right]$$

Par le théorème de Fubini, puisque

$$\mathbb{E}\left[\int_{\mathbb{R}} \left|e^{\imath t X_n} h(t)\right|\ \mathrm{d}t\right] \leqslant \|h\|_{L^1} < +\infty$$

On intervertit espérance et intégrale :

$$\mathbb{E}[f(X_n)] = \int_{\mathbb{R}} \mathbb{E}\left[e^{\imath t X_n}\right] h(t)\ \mathrm{d}t$$

Par hypothèse,

$$\mathbb{E}\left[e^{\imath t X_n}\right] \xrightarrow[n\to+\infty]{} \mathbb{E}\left[e^{\imath t X}\right]$$

Par le théorème de convergence dominée, on obtient

$$\mathbb{E}[f(X_n)] \xrightarrow[n\to+\infty]{} \int_{\mathbb{R}} \mathbb{E}\left[e^{\imath t X}\right] h(t)\ \mathrm{d}t$$

Donc, en réutilisant le théorème de Fubini, il suit finalement que

$$\mathbb{E}[f(X_n)] \xrightarrow[n\to+\infty]{} \mathbb{E}[f(X)]$$

ce qui montre le résultat souhaité pour $f \in A(\mathbb{R})$.

- Soit $\varepsilon > 0$. Pour $f \in \mathcal{C}_0(\mathbb{R})$, il existe $f^\varepsilon \in A(\mathbb{R})$ telle que

$$\|f - f^\varepsilon\|_\infty \leqslant \varepsilon$$

Alors pour tout $n \in \mathbb{N}$,

$$\begin{aligned} |\mathbb{E}[f(X_n)] - \mathbb{E}[f(X)]| \leqslant & \quad |\mathbb{E}[f(X_n)] - \mathbb{E}[f^\varepsilon(X_n)]| \\ & + \ |\mathbb{E}[f^\varepsilon(X_n)] - \mathbb{E}[f^\varepsilon(X)]| \\ & + \ |\mathbb{E}[f^\varepsilon(X)] - \mathbb{E}[f(X)]| \end{aligned}$$

Par inégalité triangulaire,

$$|\mathbb{E}[f(X_n)] - \mathbb{E}[f(X)] \leqslant| \leqslant 2\|f - f^\varepsilon\|_\infty + |\mathbb{E}[f^\varepsilon(X_n)] - \mathbb{E}[f^\varepsilon(X)]|$$

Pour n assez grand, le premier terme est plus petit que 2ε et le deuxième est plus petit que ε par le point précédent. D'où pour n assez grand

$$|\mathbb{E}[f(X_n)] - \mathbb{E}[f(X)] \leqslant| \leqslant 3\varepsilon$$

Ce qui conclut sur la convergence en loi de X_n vers X. □

Remarque : On n'a pas totalement tout montré dans ce théorème, à cause de ce résultat de densité, qui se montrerait par le théorème de Stone-Weierstrass d'après Zuily-Queffelec...

Théorème 39.2 : Théorème fort de Paul Lévy

Soit $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires. On suppose que $(\varphi_{X_n})_{n\in\mathbb{N}}$ converge simplement vers une fonction ψ continue en 0. Alors il existe une variable aléatoire X telle que $\psi = \varphi_X$. De plus, $(X_n)_{n\in\mathbb{N}}$ converge en loi vers X.

Remarque : On utilisera encore ici la densité de $A(\mathbb{R})$ dans $\mathcal{C}_0(\mathbb{R})$ dans cette démonstration, et cette version du théorème de représentation de Riesz : si ℓ est une forme linéaire sur $\mathcal{C}_0(\mathbb{R})$ continue positive, et de norme 1 alors il existe une mesure μ positive de masse inférieure à 1 telle que

$$\forall f \in \mathcal{C}_0(\mathbb{R}), \ell(f) = \int_{\mathbb{R}} f(x)\ \mathrm{d}\mu(x)$$

Ceci est démontré dans *Real and complex analysis* de Walter Rudin (il en existe une version traduite *Analyse réelle et complexe*), qui fait intervenir des mesures complexes, et le théorème de Radon-Nykodim, dans une version complexe. La démonstration se fait en 10 étapes...

Démonstration : Pour $f \in \mathcal{C}_0(\mathbb{R})$, on note

$$\forall n \in \mathbb{N}, \ell_n(f) \overset{\text{déf.}}{=} \int_{\mathbb{R}} f(x)\ \mathrm{d}\mu_n(x)$$

où $\mu_n \overset{\text{déf.}}{=} \mathbb{P}_{X_n}$. On observe ici que

$$\forall n \in \mathbb{N}, \forall f \in \mathcal{C}_0(\mathbb{R}), |\ell_n(f)| \leqslant \|f\|_\infty$$

On veut définir une forme linéaire limite ℓ.

1. Supposons que $f \in A(\mathbb{R})$. Il existe $h \in L^1(\mathbb{R})$ telle que

$$\forall x \in \mathbb{R}, f(x) = \hat{h}(x)$$

Alors, pour tout $n \in \mathbb{N}$:

$$\int_{\mathbb{R}} f \, \mathrm{d}\mu_n = \int_{\mathbb{R}} \hat{h} \, \mathrm{d}\mu_n$$

Par le théorème de Fubini,

$$\int_{\mathbb{R}} f \, \mathrm{d}\mu_n = \int_{\mathbb{R}} h(t)\varphi_{X_n}(t) \, \mathrm{d}t$$

Par le théorème de convergence dominée,

$$\int_{\mathbb{R}} f \, \mathrm{d}\mu_n \xrightarrow[n \to +\infty]{} \int_{\mathbb{R}} h(t)\psi(t) \, \mathrm{d}t$$

Ainsi, pour tout $f \in A(\mathbb{R})$, $(\ell_n(f))_{n\in\mathbb{N}}$ converge vers une limite que l'on note $\ell(f)$.

2. Par densité, on va définir ℓ sur $\mathcal{C}_0(\mathbb{R})$. Pour cela, montrons que $(\ell_n(f))_n$ est une suite de Cauchy. Soit $\varepsilon > 0$. Il existe $f^\varepsilon \in A(\mathbb{R})$ telle que

$$\|f - f^\varepsilon\|_\infty \leqslant \varepsilon$$

Soient $n, p \in \mathbb{N}$. On fait le découpage en 3 :

$$\begin{aligned} |\ell_{n+p}(f) - \ell_n(f)| \leqslant \quad & |\ell_{n+p}(f) - \ell_{n+p}(f^\varepsilon)| \\ + \quad & |\ell_{n+p}(f^\varepsilon) - \ell_n(f^\varepsilon)| \\ + \quad & |\ell_n(f^\varepsilon) - \ell_n(f)| \end{aligned}$$

Donc

$$|\ell_{n+p}(f) - \ell_n(f)| \leqslant 2\|f - f^\varepsilon\|_\infty + |\ell_{n+p}(f^\varepsilon) - \ell_n(f^\varepsilon)|$$

Donc, pour n assez grand,

$$|\ell_{n+p}(f) - \ell_n(f)| \leqslant 3\varepsilon$$

La suite $(\ell_n(f))_n$ est donc de Cauchy dans $\mathbb{C}$, donc converge dans $\mathbb{C}$ vers un élément que l'on note $\ell(f)$.

3. ℓ est linéaire, par passage à la limite, et on dispose de l'inégalité

$$|\ell(f)| \leqslant \|f\|_\infty$$

avec $\ell(f) \geqslant 0$ lorsque $f \geqslant 0$. ℓ est donc une forme linéaire sur $\mathcal{C}_0(\mathbb{R})$ positive de norme 1. Par le théorème de Riesz, il exite une mesure μ positive telle que $\mu(\mathbb{R}) \leqslant 1$, et

$$\forall f \in \mathcal{C}_0(\mathbb{R}), \ell(f) = \int_{\mathbb{R}} f \, \mathrm{d}\mu$$

4. Montrons que $\mu(\mathbb{R}) = 1$, *id est* que μ est une probabilité. On note pour tout $t \in \mathbb{R}$,

$$K(t) \overset{\text{déf.}}{=} \frac{1}{2} e^{-|t|}$$

qui vérifie $\int_{\mathbb{R}} K = 1$. Pour $\varepsilon > 0$, on définit

$$K_\varepsilon(t) \overset{\text{déf.}}{=} \frac{1}{\varepsilon} K\left(\frac{t}{\varepsilon}\right)$$

Alors

$$\forall \xi \in \mathbb{R}, \widehat{K_\varepsilon}(\xi) = \hat{K}(\varepsilon\xi) = \frac{1}{1 + \varepsilon^2\xi^2}$$

Et on a les égalités :

$$\ell\left(\widehat{K_\varepsilon}\right) = \int_{\mathbb{R}} K_\varepsilon \psi = \int_{\mathbb{R}} \widehat{K_\varepsilon} \, \mathrm{d}\mu$$

Ce qui se réécrit

$$\frac{1}{2\varepsilon} \int_{\mathbb{R}} \psi(t) e^{\frac{-|t|}{\varepsilon}} \, \mathrm{d}t = \int_{\mathbb{R}} \frac{\mathrm{d}\mu(\xi)}{1 + \varepsilon^2\xi^2}$$

Par convergence dominée, le terme de droite tend vers $\mu(\mathbb{R})$ lorsque $[\varepsilon \to 0^+]$. Quant au terme de droite converge vers $\psi(0) = 1$, par convergence simple des φ_{X_n}, et par continuité de ψ en 0, en utilisant un changement de variables, et par convergence dominée. D'où finalement $\mu(\mathbb{R}) = 1$, donc μ est une probabilité.

5. On dipose alors de l'égalité suivante :

$$\forall f \in \mathcal{C}_0(\mathbb{R}), \int_{\mathbb{R}} f \, \mathrm{d}\mu = \lim_{n \to +\infty} \int_{\mathbb{R}} f \, \mathrm{d}\mu_n$$

Par la proposition, cela est équivalent à

$$\forall f \in \mathcal{C}_b^0(\mathbb{R}), \int_{\mathbb{R}} f \, \mathrm{d}\mu = \lim_{n \to +\infty} \int_{\mathbb{R}} f \, \mathrm{d}\mu_n$$

Appliquons ce résultat à $f_t(x) = e^{\imath tx}$:

$$\forall t \in \mathbb{R}, \int_{\mathbb{R}} e^{\imath tx} \, \mathrm{d}\mu(x) = \lim_{n \to +\infty} \int_{\mathbb{R}} e^{\imath tx} \, \mathrm{d}\mu_n(x)$$

Ce qui se réécrit :

$$\forall t \in \mathbb{R}, \int_{\mathbb{R}} e^{\imath tx} \, \mathrm{d}\mu(x) = \lim_{n \to +\infty} \varphi_{X_n}(t)$$

Soit encore

$$\forall t \in \mathbb{R}, \psi(t) = \int_{\mathbb{R}} e^{\imath tx} \, \mathrm{d}\mu(x)$$

Pour $X \overset{\text{déf.}}{=} \mathrm{id}_{\mathbb{R}}$ sur $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mu)$, il suit finalement que

$$\forall t \in \mathbb{R}, \psi(t) = \varphi_X(t)$$

La convergence en loi de $(X_n)_n$ vers X est alors donné par la version faible du théorème de Lévy. □

40 Gradient à pas optimal

Références :

1. *Optimisation et analyse convexe*, Jean-Baptiste HIRRIAT-URRUTY ;
2. *Méthodes numériques : analyse et application*, Alfio QUARTERONI, Riccardo SACCO, Fausto SALERI

Proposition 40.1

Soient $A \in \mathcal{S}_n^{++}(\mathbb{R})$ et $\mathbf{b} \in \mathbb{R}^n$. On considère l'application :

$$f : \left(\begin{array}{ccc} \mathbb{R}^n & \longrightarrow & \mathbb{R} \\ \mathbf{x} & \longmapsto & \frac{1}{2}\langle A\mathbf{x}, \mathbf{x}\rangle - \langle \mathbf{b}, \mathbf{x}\rangle \end{array} \right)$$

Soit $\mathbf{x}^* \in \mathbb{R}^n$. Les propositions suivantes sont équivalentes :

(i) $\mathbf{x}^*$ est l'unique solution de

$$A\mathbf{x}^* = \mathbf{b}$$

(ii) $\mathbf{x}^*$ est l'unique solution du problème de minimisation :

$$f(\mathbf{x}^*) = \min_{\mathbf{x} \in \mathbb{R}^n} f(\mathbf{x})$$

(iii) $\nabla f(\mathbf{x}^*) = 0$.

Démonstration : f est une fonction convexe (car de matrice hessienne égale à A, symétrique définie positive), donc admet un unique minimum global, caractérisé par

$$\nabla f(\mathbf{x}^*) = 0$$

avec

$$\forall \mathbf{x} \in \mathbb{R}^n, \nabla f(\mathbf{x}) = A\mathbf{x} - b$$

Une manière plus pédestre est de calculer $f(\mathbf{x}+\mathbf{h})$, afin de reconnaître la différentielle, et de montrer que

$$\forall \mathbf{h} \in \mathbb{R}^n, f(\mathbf{x}^*) \leqslant f(\mathbf{x}^* + \mathbf{h})$$

seulement pour $\mathbf{x}^* = A^{-1}\mathbf{b}$.

□

Lemme 40.1 : Inégalité de KANTAROVITCH

Soit $A \in \mathcal{S}_n^{++}(\mathbb{R})$. On note $\lambda_1 \geqslant \cdots \geqslant \lambda_n > 0$ son spectre, par ordre décroissant. Alors pour tout $x \in \mathbb{R}^n$,

$$\langle Ax, x\rangle\langle A^{-1}x, x\rangle \leqslant \frac{1}{4}\left[\sqrt{\frac{\lambda_1}{\lambda_n}} + \sqrt{\frac{\lambda_n}{\lambda_1}}\right]^2 \cdot \|x\|^4$$

On donne deux preuves de ce lemme : la première est plus agréable, mais provient de Wikipédia, donc n'est pas utilisable pour les oraux de l'agrégation. La deuxième provient de la référence, et provient d'un dessin : j'ai renforcé la preuve pour que la démonstration en soit une, au prix de quelques calculs.

Démonstration du lemme : • Pour $t > 0$, la concavité de la fonction logarithme donne l'inégalité suivante :

$$\sqrt{\langle Ax, x\rangle\langle A^{-1}x, x\rangle} \leqslant \frac{1}{2}\left[\frac{\langle Ax, x\rangle}{t} + t\langle A^{-1}x, x\rangle\right]$$

Soit encore

$$\sqrt{\langle Ax, x\rangle\langle A^{-1}x, x\rangle} \leqslant \frac{1}{2}\left\langle \left(\frac{1}{t}A + tA^{-1}\right)x, x\right\rangle$$

Enfin, puisque $B_t \overset{\text{déf.}}{=} \frac{1}{t}A + tA^{-1}$ est symétrique,

$$\sqrt{\langle Ax, x\rangle \langle A^{-1}x, x\rangle} \leqslant \frac{1}{2}\left(\max_{\lambda \in \sigma(B_t)} \lambda\right) \|x\|^2$$

• On cherche donc à minimiser cette quantité par rapport à t. Constatons que puisque A est symétrique définie positive, elle se diagonalise en base orthonormée, et on a alors

$$\sigma(B_t) = \left\{ \frac{\lambda}{t} + \frac{t}{\lambda} \middle| \lambda \in \sigma(A) \right\}$$

On note

$$\forall \lambda > 0, \theta_t(\lambda) \overset{\text{déf.}}{=} \frac{\lambda}{t} + \frac{t}{\lambda}$$

Alors θ_t est une fonction convexe, qui admet son minimum en $\lambda = t$, où elle vaut 2. Ainsi, si $t \in [\lambda_n, \lambda_1]$, il suit que pour tout $\lambda \in [\lambda_n, \lambda_1]$

$$\theta_t(\lambda) \leqslant \max\{\theta_t(\lambda_1), \theta_t(\lambda_n)\}$$

On impose alors que

$$\theta_t(\lambda_1) = \theta_t(\lambda_n)$$

Ce qui donne la valeur de t que l'on va adopter :

$$t = \sqrt{\lambda_1 \lambda_n}$$

et donc pour tout $\lambda \in [\lambda_n, \lambda_1]$:

$$\theta(\lambda) \leqslant \sqrt{\frac{\lambda_1}{\lambda_n}} + \sqrt{\frac{\lambda_n}{\lambda_1}}$$

• On conclut alors de la sorte :

$$\sqrt{\langle Ax, x\rangle \langle A^{-1}x, x\rangle} \leqslant \frac{1}{2}\left(\sqrt{\frac{\lambda_1}{\lambda_n}} + \sqrt{\frac{\lambda_n}{\lambda_1}}\right) \|x\|^2$$

d'où l'inégalité souhaitée. □

Passons à la démonstration plus "géométrique", dans la référence.

Démonstration du lemme : Puisque le résultat est vrai pour $x = 0$ (c'est l'égalité triviale $0 = 0$), on suppose $x \neq 0$.

1. A est symétrique réelle, donc est diagonalisable en base orthonormée : il existe $P \in O_n(\mathbb{R})$ et Δ diagonale à coefficients strictement positifs, puisque A est symétrique définie positive, telles que

$$A = {}^t P \Delta P$$

Par ailleurs, cela donne

$$A^{-1} = {}^t P \Delta^{-1} P$$

Ainsi,

$$\langle Ax, x\rangle = \Delta(Px), Px\rangle$$

et

$$\langle A^{-1}x, x\rangle = \langle \Delta^{-1}(Px), Px\rangle$$

On note donc $y \overset{\text{déf.}}{=} Px$. Puisque $P \in O_n(\mathbb{R})$, $\|y\| = \|x\|$. La quantité dont on cherche un majorant vaut alors ici

$$\langle Ax, x\rangle \langle A^{-1}x, x\rangle = \left(\sum_{k=1}^{n} \lambda_k y_k^2\right)\left(\sum_{j=1}^{n} \frac{y_j^2}{\lambda_j}\right)$$

On note pour tout $j \in [\![1, n]\!]$:

$$\alpha_j \overset{\text{déf.}}{=} \frac{y_i^2}{\|y\|^2} \geqslant 0$$

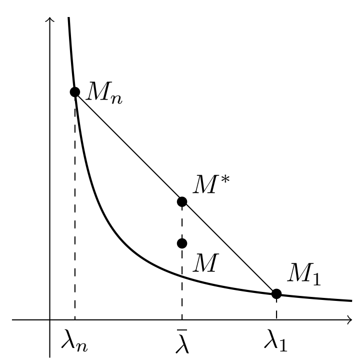
Si bien que $\sum_{j=1}^{n} \alpha_j = 1$. Ainsi, on cherche à estimer

$$\langle Ax, x\rangle \langle A^{-1}x, x\rangle = \|y\|^4 \left(\sum_{k=1}^{n} \alpha_k \lambda_k\right)\left(\sum_{j=1}^{n} \frac{\alpha_j}{\lambda_j}\right)$$

2. Donnons l'intuition géométrique ici. On dispose d'un produit de deux barycentres. On note

$$\bar{\lambda} \overset{\text{déf.}}{=} \sum_{k=1}^{n} \alpha_k \lambda_k$$

et on cherche à exprimer un majorant de la deuxième somme en fonction de $\bar{\lambda}, \lambda_1$ et λ_n. Pour cela, on note M_k les points $\left(\lambda_k, \frac{1}{\lambda_k}\right)$, M le point $\left(\bar{\lambda}, \sum_{k=1}^{n} \frac{\alpha_k}{\lambda_k}\right)$.



D'après le dessin, M^* est au-dessus de M, donc on voit facilement que

$$\sum_{k=1}^{n} \frac{\alpha_k}{\lambda_k} \leqslant \frac{\bar{t}}{\lambda_n} + \frac{1 - \bar{t}}{\lambda_1}$$

avec $\bar{t}$ tel que $\bar{\lambda} = \bar{t}\lambda_n + (1-\bar{t})\lambda_1$, autrement dit

$$\bar{t} \overset{\text{déf.}}{=} \frac{\bar{\lambda} - \lambda_1}{\lambda_n - \lambda_1}$$

3. Naturellement, un tel raisonnement n'est pas suffisant pour assurer qu'il s'agit d'une démonstration. Montrons alors que :

$$\sum_{k=1}^{n} \frac{\alpha_k}{\lambda_k} \leqslant \frac{\bar{t}}{\lambda_n} + \frac{1-\bar{t}}{\lambda_1}$$

Pour cela, on cherche $s \in [0,1]$ tel que

$$\sum_{k=1}^{n} \frac{\alpha_k}{\lambda_k} = \frac{s}{\lambda_n} + \frac{1-s}{\lambda_1}$$

et on montre que $s \leqslant \bar{t}$. La croissance de la fonction $t \longmapsto \frac{t}{\lambda_n} + \frac{1-t}{\lambda_1}$ permettra de conclure. On isole s pour avoir

$$s \overset{\text{déf.}}{=} \sum_{i=2}^{n} \alpha_i \left[\frac{\lambda_n(\lambda_1 - \lambda_i)}{\lambda_i(\lambda_1 - \lambda_n)}\right]$$

à comparer avec

$$\bar{t} = \sum_{j=2}^{n} \alpha_j \left[\frac{\lambda_1 - \lambda_j}{\lambda_1 - \lambda_n}\right]$$

Il suffit de voir que $\frac{\lambda_n}{\lambda_i} \geqslant 1$ pour tout $i \in [\![2,n]\!]$ pour conclure que $s \leqslant \bar{t}$, et donc que

$$\frac{s}{\lambda_n} + \frac{1-s}{\lambda_1} \leqslant \frac{\bar{t}}{\lambda_n} + \frac{1-\bar{t}}{\lambda_1}$$

Soit encore, par définition de s,

$$\sum_{k=1}^{n} \frac{\alpha_k}{\lambda_k} \leqslant \frac{\bar{t}}{\lambda_n} + \frac{1-\bar{t}}{\lambda_1}$$

4. De plus,

$$\frac{\bar{t}}{\lambda_n} + \frac{1-\bar{t}}{\lambda_1} = \frac{1}{\lambda_1} + \frac{1}{\lambda_n} - \frac{\bar{\lambda}}{\lambda_1\lambda_n}$$

Et donc

$$\langle Ax, x\rangle\langle A^{-1}x, x\rangle \leqslant \frac{\|x\|^4}{\lambda_1\lambda_n}\bar{\lambda}(\lambda_1 + \lambda_n - \bar{\lambda})$$

C'est un polynôme de degré 2 en $\bar{\lambda}$, qui se minimise en $\bar{\lambda} = \frac{\lambda_1+\lambda_n}{2}$, et qui donne finalement

$$\langle Ax, x\rangle\langle A^{-1}x, x\rangle \leqslant \frac{\|y\|^4}{4} \cdot \frac{(\lambda_1+\lambda_n)^2}{\lambda_1\lambda_n}$$

d'où l'inégalité de Kantarovitch. □

Théorème 40.1 : Méthode du gradient optimal

On définit $x_0 \in \mathbb{R}^n$, et pour tout $k \in \mathbb{N}$:

$$x_{k+1} \overset{\text{déf.}}{=} x_k - t_k \nabla f(x_k)$$

où $t_k \geqslant 0$ est l'unique réel tel que

$$f\left(x_k - t_k\nabla f(x_k)\right) = \min_{t\in\mathbb{R}} f\left(x_k - t\nabla f(x_k)\right)$$

Alors $(x_k)_{k\in\mathbb{N}}$ converge vers $\mathbf{x}^*$.

Démonstration : **1.** Commençons d'abord par justifier le fait que pour tout $k \in \mathbb{N}$,

$$\xi_k : t \longmapsto f\left(x_k - t\nabla f(x_k)\right)$$

admette un unique minimum global. Il suffit en fait de développer :

$$\begin{aligned}\xi_k(t) = \quad & \tfrac{1}{2}\langle A(x_k - t\nabla f(x_k)), x_k - t\nabla f(x_k)\rangle \\ - \quad & \langle b, x_k - t\nabla f(x_k)\rangle\end{aligned}$$

On obtient alors

$$\begin{aligned}\xi_k(t) = \quad & f(x_k) \\ - \quad & \tfrac{t}{2}\left(\langle A\nabla f(x_k), x_k\rangle + \langle Ax_k, \nabla f(x_k)\rangle\right) \\ + \quad & t\langle b, \nabla f(x_k)\rangle + \tfrac{t^2}{2}\langle A\nabla f(x_k), \nabla f(x_k)\rangle\end{aligned}$$

On réarrange : A est symétrique, et $\nabla f(x_k) = Ax_k - b$, d'où

$$\xi_k(t) = f(x_k) - t\|\nabla f(x_k)\|^2 + \frac{t^2}{2}\langle A\nabla f(x_k), \nabla f(x_k)\rangle$$

Pour $\nabla f(x_k) \neq 0$, on se retrouve donc en présence d'un trinôme du second degré, avec un coefficient dominant positif : par conséquent, ξ_k admet bien un unique minimum, pris en t_k, donné par :

$$t_k \overset{\text{déf.}}{=} \frac{\|\nabla f(x_k)\|^2}{\langle A\nabla f(x_k), \nabla f(x_k)\rangle} > 0$$

2. Puisque A est définie symétrique positive, un calcul dans une base de vecteurs propres de A donne :

$$\forall \mathbf{u} \in \mathbb{R}^n, \langle A\mathbf{u}, \mathbf{u}\rangle = \sum_{i=1}^{n} \lambda_i u_i^2$$

où $\lambda_1 \geqslant \cdots \geqslant \lambda_n > 0$ sont les valeurs propres de A par ordre décroissant. Ainsi,

$$\forall \mathbf{u} \in \mathbb{R}^n, \langle A\mathbf{u}, \mathbf{u}\rangle \geqslant \lambda_n \|\mathbf{u}\|^2$$

On peut donc commencer à estimer $\|x_k - \mathbf{x}^*\|$ grâce à cette remarque :

$$\|x_k - \mathbf{x}^*\|^2 \leqslant \frac{1}{\lambda_n}\langle A(x_k - \mathbf{x}^*), x_k - \mathbf{x}^*\rangle$$

Développons, en utilisant le fait que $A\mathbf{x}^* = \mathbf{b}$:

$$\|x_k - \mathbf{x}^*\|^2 \leqslant \frac{1}{\lambda_n}\left[\langle Ax_k, x_k\rangle - 2\langle \mathbf{b}, \mathbf{x}^*\rangle + \langle \mathbf{b}, \mathbf{x}^*\rangle\right]$$

Puis, en faisant apparaître f, en constatant que :

$$f(\mathbf{x}^*) = \frac{1}{2}\langle A\mathbf{x}^*, \mathbf{x}^*\rangle - \langle \mathbf{b}, \mathbf{x}^*\rangle$$

soit

$$f(\mathbf{x}^*) = \frac{-1}{2}\langle \mathbf{b}, \mathbf{x}^*\rangle$$

On obtient alors

$$\begin{aligned}\|x_k - \mathbf{x}^*\|^2 \leqslant & \quad \tfrac{1}{\lambda_n}[2f(x_k) + 2\langle b, x_k\rangle \\ - & \quad 2\langle b, x_k\rangle - 2f(\mathbf{x}^*)]\end{aligned}$$

D'où

$$\|x_k - \mathbf{x}^*\|^2 \leqslant \frac{2}{\lambda_n}\left[f(x_k) - f(\mathbf{x}^*)\right]$$

3. On note

$$\forall k \in \mathbb{N}, \delta_k \overset{\text{déf.}}{=} f(x_k) - f(\mathbf{x}^*)$$

Déterminons une relation de récurrence sur cette suite. D'après le développement de f fait en **1.**, on a :

$$\delta_{k+1} = f(x_k) - t_k\|\nabla f(x_k)\|^2 + \frac{t_k^2}{2}\langle A\nabla f(x_k), \nabla f(x_k)\rangle - f(\mathbf{x}^*)$$

Puis, par définition de t_k,

$$\delta_{k+1} = f(x_k) - \frac{\|\nabla f(x_k)\|^4}{2\langle A\nabla f(x_k), \nabla f(x_k)\rangle} - f(\mathbf{x}^*)$$

Enfin, si $x_k \neq \mathbf{x}^*$,

$$\delta_{k+1} = \delta_k\left[1 - \frac{\|\nabla f(x_k)\|^4}{2\delta_k\langle A\nabla f(x_k), \nabla f(x_k)\rangle}\right]$$

On peut exprimer $2\delta_k$ au dénominateur en fonction de $\nabla f(x_k)$ et A^{-1}, en sachant que $\mathbf{x}^* = A^{-1}\mathbf{b}$. En effet, par définition de f :

$$2\delta_k = \langle Ax_k, x_k\rangle - 2\langle \mathbf{b}, x_k\rangle + \langle \mathbf{b}, \mathbf{x}^*\rangle$$

On fait ici un petit jeu d'écriture (attention, astuce), avec le fait que A soit symétrique :

$$\begin{aligned}2\delta_k = & \quad \langle Ax_k, A^{-1}Ax_k\rangle - \langle A^{-1}\mathbf{b}, Ax_k\rangle \\ - & \quad \langle \mathbf{b}, x_k\rangle + \langle \mathbf{b}, A^{-1}\mathbf{b}\rangle\end{aligned}$$

Il est alors possible de réorganiser le tout :

$$2\delta_k = \langle A^{-1}(Ax_k - \mathbf{b}), Ax_k - \mathbf{b}\rangle$$

Soit encore

$$2\delta_k = \langle A^{-1}\nabla f(x_k), \nabla f(x_k)\rangle$$

D'où

$$\delta_{k+1} = \delta_k\left[1 - \frac{\|\nabla f(x_k)\|^4}{\langle A^{-1}\nabla f(x_k), \nabla f(x_k)\rangle\langle A\nabla f(x_k), \nabla f(x_k)\rangle}\right]$$

4. D'après l'inégalité de Kantorovitch,

$$\frac{\|\nabla f(x_k)\|^4}{\langle A^{-1}\nabla f(x_k), \nabla f(x_k)\rangle\langle A\nabla f(x_k), \nabla f(x_k)\rangle} \geqslant 4\left[\sqrt{\frac{\lambda_1}{\lambda_n}} + \sqrt{\frac{\lambda_n}{\lambda_1}}\right]^{-2}$$

On note

$$c_2(A) \overset{\text{déf.}}{=} \frac{\lambda_1}{\lambda_n}$$

le conditionnement de A (pour la norme 2). On a alors

$$\frac{\|\nabla f(x_k)\|^4}{\langle A^{-1}\nabla f(x_k), \nabla f(x_k)\rangle\langle A\nabla f(x_k), \nabla f(x_k)\rangle} \geqslant 4\frac{c_2(A)}{[1 + c_2(A)]^2}$$

d'où

$$\delta_{k+1} \leqslant \delta_k\left[1 - \frac{4c_2(A)}{(1 + c_2(A))^2}\right]$$

En réduisant au même dénominateur,

$$\delta_{k+1} \leqslant \delta_k\left[\frac{1 - c_2(A)}{1 + c_2(A)}\right]^2$$

D'où pour tout $k \in \mathbb{N}$,

$$\delta_k \leqslant \delta_0\left[\frac{1 - c_2(A)}{1 + c_2(A)}\right]^{2k}$$

5. Il suit alors que

$$\|x_k - \mathbf{x}^*\|^2 \leqslant \frac{2\delta_0}{\lambda_n}\left[\frac{1 - c_2(A)}{1 + c_2(A)}\right]^{2k}$$

La méthode converge alors vers $\mathbf{x}^*$. $\square$