Préparation à l'Agrégation

Guillaume Kineider

Table des matières

0.1 Théorème taubérien fort de Hardy et Littlewood

Leçons : 209 ; 230 ; 243

Références :

— [CQ09] Denis Choimet, Hervé Queffélec, *Analyse mathématiques : Grands thèmes du vingtième siècle*

— [Kor10] Jacob Korevaar, *Tauberian theory : a century of developments*

— Benjamin Havret, `http://www.normalesup.org/~havret/pdf/developpements_maths%20bhavret.pdf`

Théorème 0.1.1 (taubérien fort de Hardy et Littlewood)**.**
Soit $f(x) = \sum_{n=1}^{+\infty} a_n x^n$ définie pour $|x| < 1$. On suppose :

(i) $f(x) \underset{x \nearrow 1}{\longrightarrow} l$;

(ii) $na_n \geqslant -C$ *pour une constante* $C > 0$.

Alors, $S_N = a_0 + a_1 + ... + a_N \underset{N \to +\infty}{\longrightarrow} l$.

Lemme 0.1.2.
On note $\mathscr{P} = \{P \in \mathbf{R}[X] \mid P(0) = 0 \text{ et } P(1) = 1\}$. *Alors, pour* $x \in]-1, 1[$ *et* $P \in \mathbf{R}[X]$ *tel que* $P(0) = 0$, $\sum a_n P(x)$ *converge et*

$$\text{(i)} \iff \text{(i')} : \forall P \in \mathscr{P}, \quad \sum_{n=0}^{+\infty} a_n P(x^n) \underset{x \nearrow 1}{\longrightarrow} l.$$

Démonstration.
Il s'agit d'un auto-renforcement de la condition (i) puisque le choix de $P(X) = X \in \mathscr{P}$ montre l'implication (i') $\Longrightarrow$ (i). Pour montrer (i) $\Longrightarrow$ (i'), on note que pour $r \in \mathbf{N}^*$ et $x \in]0, 1[$, $0 < x^r < x < 1$ donc :

$$\text{la série } \sum a_n x^{rn} \text{ converge et } \sum_{n=0}^{+\infty} a_n x^{rn} = f(x^r) \underset{x \nearrow 1}{\longrightarrow} l.$$

Puis, par linéarité, pour tout polynôme P fixant 0,

$$\sum_{n=0}^{+\infty} a_n P(x^n) \underset{x \nearrow 1}{\longrightarrow} lP(1)$$

et en divisant par $P(1)$ on obtient le résultat.

□

Démonstration du théorème.

Étape 1 : Réécriture de S_N.

Fixons $N \in \mathbf{N}$. On souhaite écrire S_N comme une somme de la forme qui apparaît dans hypothèse (i'). La première étape est de rendre la somme infinie par l'ajout classique d'une fonction indicatrice :

$$S_N = \sum_{n=0}^{N} a_n = \sum_{n=0}^{+\infty} a_n \mathbb{1}_{[0,N]}(n).$$

On obtient une somme de $a_n \times$ *une fonction de* n. La grande idée de Littlewood est de transformer cette fonction de n en une expression de la forme $g(x^n)$. Une façon simple de faire est d'écrire :

$$S_N = \sum_{n=0}^{+\infty} a_n g(x_N^n)$$

où $g = \mathbb{1}_{[e^{-1},1]}$ et $x_N = e^{-\frac{1}{N}} \underset{N \to +\infty}{\nearrow} 1$.

Étape 2 : Réécriture du problème.

On fixe pour toute la suite $x \in]0, 1[$. On a la paramétrisation des polynômes fixant 0 et 1 suivante :

$$\mathscr{P} = \{X + X(1 - X)Q(X) \mid Q \in \mathbf{R}[X]\}.$$

En effet, les polynômes de cette forme fixent bien 0 et 1. Et réciproquement, si $P \in \mathscr{P}$ P a 0 pour racine donc se factorise en $P = XR$ et $P(1) = R(1)$ donc $P = X(1 + R - 1) = X + X(R - 1)$ avec $R - 1$ ayant 1 pour racine donc qui se factorise par $(1 - X)$: $P = X + X(1 - X)Q(X)$. Et alors,

$$\sum_{n=0}^{+\infty} a_n P(x^n) = \underbrace{\sum_{n=0}^{+\infty} a_n x^n}_{f(x)} + \underbrace{\sum_{n=0}^{+\infty} a_n x^n (1 - x^n) Q(x^n)}_{S_Q(x)}$$

la convergence des sommes étant assurée par le lemme ce qui justifie la séparation des sommes. Et, puisque par hypothèse $f(x)$ converge vers l lorsque x tend vers 1 par valeurs croissantes,

$$\text{(i')} \iff \text{(i")} : \forall Q \in \mathbf{R}[X], \qquad S_Q(x) \underset{x \nearrow 1}{\longrightarrow} 0$$

On s'est ramené à un prédicat sur l'ensemble des polynômes plutôt que sur le sous-ensemble $\mathscr{P}$ ce qui permettra d'utiliser plus tard un résultat de densité. On va donc décomposer de façon analogue g. On écrit

$$g(x) = x + x(1 - x)h(x)$$

où $h(x) = -\frac{1}{1-x}\mathbb{1}_{[0,e^{-1}[}(x) + \frac{1}{x}\mathbb{1}_{[e^{-1},1]}(x)$ est continue par morceaux. Pour achever la démonstration, il suffit de montrer que :

$$S_h(x_N) \underset{N \to +\infty}{\longrightarrow} 0$$

Nous allons montrer que sa limsup est inférieure ou égale à 0, le cas de la liminf se traitant de la même façon. On suppose $Q > h$. On a par linéarité de la somme et donc $Q \mapsto S_Q$:

$$S_h(x_N) = S_Q(x_N) + S_{h-Q}(x_N)$$

et on cherche à majorer le second terme.

Étape 3 : Majoration de l'écart.

Partant de l'hypothèse (ii) $-C/n \leqslant a_n$, on multiplie par $x^n(1 - x^n)[h - Q](x^n) < 0$:

$$a_n x^n (1 - x^n)[h - Q](x^n) \leqslant \frac{C}{n} x^n (1 - x^n)[Q - h](x^n)$$

le membre de droite étant positif. Ensuite, $0 \leqslant 1 - x^n = (1 - x)(1 + \cdots + x^{n-1}) \leqslant n(1 - x)$ donc

$$a_n x^n (1 - x^n)[h - Q](x^n) \leqslant C(x^n - x^{n+1})[Q - h](x^n).$$

On somme sur $n \in \mathbf{N}$ cette relation et on évalue en x_N pour obtenir :

$$S_{h-Q}(x_N) \leqslant C \sum_{n=0}^{+\infty} (x_N^n - x_N^{n+1})[Q - h](x_N^n) \underset{N \to +\infty}{\longrightarrow} C \int_0^1 (Q - h)$$

par sommes de Riemann.

Explications : On utilise la méthode des rectangles à droite sur $[0, 1]$ avec les subdivisions indicées par $N \in \mathbf{N}^*$ suivantes :

$$0 < x_N^{\varphi(N)} < \ldots < x_N^2 < x_N < 1$$

où $\varphi : \mathbf{N} \to \mathbf{N}$ est strictement croissante et telle que $x_N^{\varphi(N)} \underset{N \to +\infty}{\longrightarrow} 0$ (possible à construire car à $N \in \mathbf{N}$ fixé, $x_N^n \underset{n \to +\infty}{\longrightarrow} 0$). Le pas tend vers 0 car :

$$x_N^{\varphi(N)} \underset{N \to +\infty}{\longrightarrow} 0, \quad x_N^n - x_N^{n+1} = x_N^n(1 - x_N) \underset{N \to +\infty}{\longrightarrow} 0 \quad \text{et} \quad x_N - 1 \underset{N \to +\infty}{\longrightarrow} 0$$

Donc

$$\limsup_{N \to +\infty} S_h(x_N) \leqslant C \int_0^1 (Q - h) \qquad \text{pour tout polynôme } Q > h$$

On cherche à rendre l'intégrale de droite arbitrairement petite donc une approximation polynomiale L^1 constamment supérieure à h.

Étape 4 : Approximation polynomiale.

On voudrait utiliser le théorème de Weierstrass mais h n'est pas continue. On fixe $0 < \varepsilon \leqslant e^{-1}/2$ et on considère donc d'abord une approximation continue $s \geqslant h$ telle que :

$$\int_0^1 s - h \leqslant C'\varepsilon, \quad C > 0.$$

Une façon de faire est par exemple de prendre s égale à h en dehors de $[e^{-1} - \varepsilon, e^{-1}]$ et linéaire sur cette intervalle. On vérifie alors la condition sur l'intégrale :

$$\int_0^1 (s-h) = \int_{e^{-1}-\varepsilon}^{e^{-1}} (s(t) - h(t))\mathrm{d}t \leqslant \left(e + \frac{e}{e-1}\right)\varepsilon$$

donc $C' = e + \frac{e}{e-1}$. On considère alors par théorème de Weierstrass un polynôme Q vérifiant :

$$\|Q - (s+\varepsilon)\|_\infty < \varepsilon.$$

Alors, $h \leqslant s < Q < s + 2\varepsilon$ et de plus :

$$\int_0^1 (Q-h) = \int_0^1 (Q-s) + \int_0^1 (s-h) \leqslant 2\varepsilon + C'\varepsilon$$

Étape 5 : Conclusion.

On a donc montré que pour tout $\varepsilon > 0$, $\limsup_{N\to+\infty} S_h(x_N) \leqslant C(C'+2)\varepsilon$. Pour la limite inférieure, on peut faire la même chose en prenant $Q < h$. Cela renverse toutes les inégalités, ce qui explique d'ailleurs que la condition asymétrique $na_n \leqslant -C$ est suffisante, et on obtient pour tout $\varepsilon > 0$

$$C(C'+2)\varepsilon \leqslant \liminf_{N\to+\infty} S_h(x_N) \leqslant \limsup_{N\to+\infty} S_h(x_N) \leqslant C(C'+2)\varepsilon$$

d'où le résultat. □

Remarque :

Historiquement, il s'agit d'un résultat important, conjecture de Hardy prouvé par Littlewood qui marque le début de leur collaboration (plus que fructueuse !) et qui constitue l'un des premiers d'une longue série de théorèmes taubériens qui s'intéressent au comportement de séries entières sur leur disque de convergence. La preuve proposée ici n'est pas celle de Littlewood mais de Wielandt qui est une variation de la preuve de Karamata.

0.2 Marche aléatoire sur Z

Leçons : 190 ; 230 ; 241 ; 264 ; 266

Références :

— [App13] Walter Appel, *Probabilités pour les non-probabilistes*

Théorème 0.2.1.
Soit $(X_i)_{i\geqslant 1}$ une suite de variables aléatoires indépendantes telles que pour tout $i \in \mathbf{N}^*$, $\mathbb{P}\{X_i = 1\} = p$ et $\mathbb{P}\{X_i = -1\} = q = 1-p$. On note $S_n = X_1 + \cdots + X_n$ et on considère la marche aléatoire sur $\mathbf{Z}$ $(S_n)_{n\in\mathbf{N}^*}$.

1. *Si $p \neq \frac{1}{2}$, alors presque sûrement, il n'y aura qu'un nombre fini de retours à 0.*
2. *Si $p = \frac{1}{2}$, il y aura presque sûrement une infinité de retours à 0.*

Démonstration.
On s'intéresse pour $n \in \mathbf{N}^*$ aux évènements $A_n = \{S_n = 0\}$ de probabilité $a_n := \mathbb{P}\{A_n\}$. On va commencer par démontrer le premier point, cas où les pas à gauche et à droite ne sont pas équiprobables.

Cas $p \neq q$.

Montrons que $\limsup_{n\in\mathbf{N}^*} A_n = 0$. L'outil principal pour faire cela est le lemme de Borel-Cantelli "sens facile". Soit $k \in \mathbf{N}^*$. Si on effectue un nombre impair de pas, on ne peut pas se retrouver à l'origine donc on a $a_{2k+1} = 0$ et si on effectue un nombre pair de pas, on est à l'origine si et seulement si on a fait autant de pas d'un côté que de l'autre d'où $a_{2k} = \binom{2k}{k} p^k (1-p)^k$. Par la formule de Stirling :

$$\binom{2k}{k} = \frac{(2k)!}{(k!)^2} \sim \frac{(2k)^{2k}\sqrt{4k\pi}}{e^{2k}} \times \frac{e^{2k}}{k^{2k}(\sqrt{2k\pi})^2} = \frac{4^k}{\sqrt{k\pi}}$$

Donc $a_{2k} \sim \frac{\alpha^k}{\sqrt{k\pi}}$ avec $\alpha = 4p(1-p) < 1$ puisque $p \neq q$. Ainsi $\sum_n a_n$ converge, donc par théorème de Borel-Cantelli $\mathbb{P}\left\{\limsup_{n\in\mathbf{N}^*} A_n\right\} = 0$ ce qui achève la preuve du premier point.

Cas $p = q$.

On pourrait chercher à prouver le deuxième point par l'autre sens de Borel-Cantelli. Mais les évènements A_n ne sont pas indépendants. Pour remédier à ce problème, on va utiliser le caractère "regénératif" des retours à l'équilibre : les temps d'occurrences successives sont indépendants et de même loi. En effet :

Si $n \in \mathbf{N}^*$ et $k \geqslant 1$,

$$\mathbb{P}\{S_{n+k} = 0 \mid S_n = 0\} = \mathbb{P}\{X_{n+1} + \cdots + X_{n+k} = 0\} = \mathbb{P}\{S_k = 0\}$$

car les X_i sont indépendantes et identiquement distribuées. Donc les retours à 0 successifs sont sans mémoire.
On note :

— T l'instant de premier retour à 0 , il s'agit d'une variable aléatoire discrète à valeurs dans $\mathbf{N}^* \cup \{+\infty\}$;
— G_T sa fonction génératrice ;
— $\forall n \geqslant 1$, $b_n := \mathbb{P}\{T = n\}$;
— $b_0 = 0$, $a_0 = 1$
— $A(x) = \sum_{n\geqslant 0} a_n x^n$ la série entière de terme général a_n. Ce n'est a priori pas une fonction génératrice mais les a_n étant des probabilités, son rayon de convergence est au moins 1.

Ainsi,

$$\forall x \in [0,1], \quad G_T(x) = \sum_{n=0}^{+\infty} b_n x^n \qquad \text{et} \qquad \forall x \in [0,1[, \quad A(x) = \sum_{n=0}^{+\infty} a_n x^n$$

Étape 1 : $\forall n \geqslant 1$, $a_n = \sum_{k=0}^{n} b_k a_{n-k}$.

Soit $n \geqslant 1$. On a $\{S_n = 0\} = \{S_n = 0 \cap T \leqslant n\}$ puis par formule des probabilités totales :

$$\begin{aligned}
\mathbb{P}\{S_n = 0\} &= \sum_{k=1}^{n} \mathbb{P}\{S_n = 0 \cap T = k\} \\
&= \mathbb{P}\{T = n\} + \sum_{k=1}^{n-1} \mathbb{P}\{X_{k+1} + \cdots + X_n = 0 \cap T = k\} \\
&= b_n a_0 + \sum_{k=1}^{n-1} b_k a_{n-k} \\
&= \sum_{k=0}^{n} b_k a_{n-k}
\end{aligned}$$

où on a utilisé successivement que l'évènement $\{T = k\}$ est dans la tribu engendrée par les k premiers X_i et est donc indépendant de l'évènement $\{X_{k+1} + \cdots + X_n = 0\}$, que les X_i sont indépendantes et identiquement distribuées et enfin, que puisque $b_0 = 0$, on peut faire commencer la somme à 0.

Étape 2 : $\forall x \in [0,1[,\ G_T(x) = 1 - \frac{1}{A(x)}$.

Soit $x \in [0,1[$. On multiplie la relation de convolution précédente par x^n puis on somme sur tout $n \geqslant 1$. Au membre de gauche, on a $A(x)$ privé de son premier terme $a_0 = 1$ et à droite on reconnait le produit de CAUCHY des deux séries absolument convergentes $A(x)$ et $G_T(x)$:

$$A(x) - 1 = A(x)G_T(x) \tag{1}$$

Puisque $A(x) \geqslant a_0 > 0$, on peut diviser par $A(x)$ et on obtient :

$$G_T(x) = 1 - \frac{1}{A(x)}$$

Les deux premières étapes sont vraies pour tout phénomène regénératif et constituent une méthode assez générale pour étudier ces phénomènes. Il est souvent plus facile de calculer A pour en déduire G_T.

Étape 3 : Calcul de G_T.

Dans notre situation,

$$A(x) = \sum_{n=0}^{+\infty} \frac{1}{4^n} \binom{2n}{n} x^{2n} = \sum_{n=0}^{+\infty} \frac{1}{2^n n!} \frac{(2n)!}{2^n n!} x^{2n}$$

or

$$(2n)! = 1 \times 3 \times 5 \times \cdots \times (2n-1) \times \underbrace{2 \times 4 \times \cdots \times (2n)}_{2^n \times 1 \times 2 \times \cdots \times n = 2^n n!}$$

donc

$$A(x) = \sum_{n=0}^{+\infty} \frac{1 \times 3 \times 5 \times \cdots \times (2n-1)}{2^n n!} (x^2)^n = \frac{1}{\sqrt{1-x^2}}$$

en reconnaissant le développement en série entière de $\frac{1}{\sqrt{1-x}}$. Par l'étape 2 :

$$\forall x \in [0,1[, \qquad G_T(x) = 1 - \sqrt{1-x^2}$$

Étape 4 : Conclusion.

La fonction G_T est continue sur $[0,1]$ par exemple comme série uniformément convergente de fonctions continues donc :

$$G_T(1) = \lim_{x \nearrow 1} 1 - \sqrt{1-x^2} = 1 = \mathbb{P}\{T < +\infty\}.$$

On obtient le point 2 grâce au caractère regénératif du phénomène : si Ω_1 est l'évènement "il y a un retour à 0", $\mathbb{P}\{\Omega_1\} = 1$. Puis si Ω_2 est "il y a un second retour à 0", alors $\mathbb{P}\{\Omega_2 \mid \Omega_1\} = 1$ car les instants de retour à 0 sont sans mémoire, ce qui entraine $\mathbb{P}\{\Omega_2\} = \mathbb{P}\{\Omega_2 \mid \Omega_1\}\mathbb{P}\{\Omega_1\} = 1$ et ainsi, par récurrence on obtient $\mathbb{P}\left\{\limsup_{n \in \mathbf{N}} A_n\right\} = 1$. □

0.3 Méthode de Newton

Leçons : 226 ; 228 ; 229

Références :

— [Rou09] François Rouvière, *Petit guide de calcul différentiel à l'usage de la licence et de l'agrégation*

Théorème 0.3.1.
Soient $c, d \in \mathbf{R}$ avec $c < d$ et $f : [c, d] \to \mathbf{R}$ une fonction de classe $\mathscr{C}^2$. On suppose

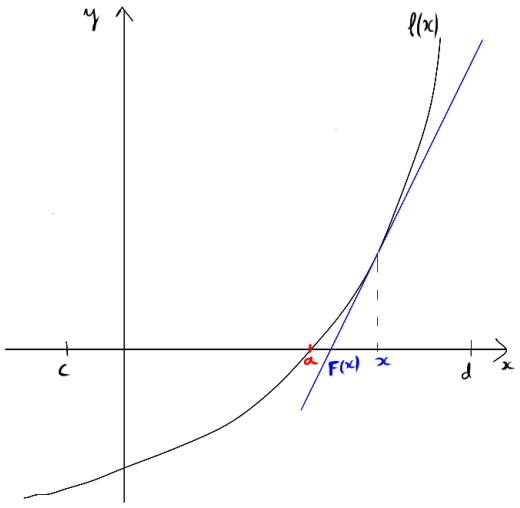
- $f(c) < 0 < f(d)$;
- $f' > 0$.

Posant $F : x \mapsto x - \frac{f(x)}{f'(x)}$, il existe alors $I \subset [c, d]$ stable par F tel que pour tout $x_0 \in I$, la suite définie par :

$$\forall n \in \mathbf{N}, \quad x_{n+1} = F(x_n)$$

converge avec une vitesse quadratique vers l'unique zéro de f sur $[c, d]$, noté a.

On se donne une fonction f de classe $\mathscr{C}^2$ sur un segment $[c, d]$. Par hypothèse, elle est strictement négative en c, strictement positive en d donc par théorème des valeurs intermédiaires, f s'annule au moins une fois. Exactement une fois en fait, car, sa dérivée étant strictement positive, f est strictement croissante. Ce zéro est noté a. Parfois il n'est pas possible d'expliciter a. Ce théorème donne une méthode pour l'approcher, appelé méthode de Newton. L'idée est de partir d'un point x et de remplacer f par sa tangente en x. Cette tangente admet un zéro grâce à l'hypothèse $f' > 0$, et il est bien plus simple à calculer explicitement que celui de f. On a même une formule en fonction de x, donnée par $F(x)$. Le premier théorème que je vais démontrer nous dit que si on part suffisamment proche de a, on peut répéter ce processus et qu'on obtiendra une suite convergeant en vitesse quadratique vers a. Enfin, on peut aussi remarquer a est l'unique solution de $F(x) = a$. Donc la méthode de Newton transforme en fait la résolution d'une équation en un problème de point fixe.



Démonstration.
Soit $x \in [c, d]$. Puisque $f(a) = 0$,

$$F(x) - a = x - a - \frac{f(x) - f(a)}{f'(x)} = \frac{f(a) - f(x) - f'(x)(a - x)}{f'(x)}$$

puis, par théorème de Taylor-Lagrange, il existe $z \in]a, x[$ tel que :

$$F(x) - a = \frac{1}{2}\frac{f''(z)}{f'(x)}(x - a)^2 \tag{2}$$

On aimerait rendre la quantité devant $(x - a)^2$ indépendante de x. Les fonctions $|f'|$ et $|f''|$ sont continues sur le segment $[c, d]$ donc atteignent leurs extremums. Posons $C := \frac{1}{2}\frac{\max |f''|}{\min |f'|} > 0$ bien définie puisque $f' > 0$. Alors,

$$\forall x \in [c, d], \qquad |F(x) - a| \leqslant C|x - a|^2$$

Soit $\alpha > 0$ assez petit pour que $C\alpha < 1$. On pose $I := [a - \alpha, a + \alpha] \subset [c, d]$. Alors, $|F(x) - \alpha| \leqslant C\alpha^2 < \alpha$ d'où $F(I) \subset I$.
Donc si $x_0 \in I$, $x_n \in I$ pour tout $n \in \mathbf{N}$ et

$$C|x_{n+1} - a| = C|F(x) - a| \leqslant C^2|x_n - a|^2 \leqslant C^4|x_{n-1} - a|^4 \leqslant \cdots \leqslant (C|x_0 - \alpha|)^{2^{n+1}} \leqslant (C\alpha)^{2^{n+1}}$$

avec $C\alpha < 1$ d'où la convergence quadratique. □

Théorème 0.3.2.
On suppose de plus $f'' > 0$ sur $[c, d]$. Alors $J :=]a, d]$ est stable par F, et pour tout $x_0 \in J$, (x_n) est strictement décroissante et converge vers a.

Démonstration.
La nouvelle hypothèse implique que f est strictement convexe. On augmente les conditions sur f mais on relâche les conditions sur la valeur initiale. Pour $a < x \leqslant d$, on a $f'(x) > 0$ et $f(x) > 0$ donc

$$F(x) = x - \frac{f(x)}{f'(x)} < x$$

Puis par 2, on a :

$$F(x) - a = \frac{1}{2}\frac{f''(z)}{f'(x)}(x - a)^2 > 0 \tag{3}$$

grâce à la nouvelle hypothèse de convexité, ce qu'on retrouve graphiquement : La tangente en x est en dessous du graphe de f donc rencontre l'axe des abscisses "après" f. Donc J est stable par F. Si $x_0 \in J$, (x_n) est strictement décroissante, minorée par a donc converge vers un point fixe de F par continuité. Or le seul point fixe de F sur $\bar{J}$ est a donc $x_n \to a$. Enfin, on montre de la même manière que dans le premier théorème que la convergence est quadratique à partir de 3. □

Applications :

- Calcul de racines carrées.
- Calcul d'inverses de racines carrées : très utile pour renormaliser des vecteurs, notamment en affichage $3D$ et calcul des ombres. Algorithme Black Magic : `https://fr.wikipedia.org/wiki/Racine_carr%C3%A9e_inverse_rapide`

0.4 Théorème d'inversion de Fourier

Leçons : 209 ; 234 ; 235 ; 236 ; 239 ; 250

Références :

— [Pey12] Jacques Peyrière, *Convolution, séries et intégrales de Fourier*

— [IP19] Lucas Isenmann, Timothée Pecatte, *L'oral à l'agrégation de mathématiques*

Définition 0.4.1.
Pour $f \in L^1(\mathbf{R})$ on définit sa transformée de Fourier par :

$$\forall \xi \in \mathbf{R}, \qquad \hat{f}(\xi) = \int_{\mathbf{R}} e^{-2i\pi\xi x} f(x) \mathrm{d}x$$

Théorème 0.4.2 (Inversion de Fourier)**.**
Si $f \in \mathrm{L}^1(\mathbf{R})$ et $\hat{f} \in \mathrm{L}^1(\mathbf{R})$, alors p.p.$x \in \mathbf{R}$,

$$f(x) = \int_{\mathbf{R}} e^{2i\pi\xi x} \hat{f}(\xi) \mathrm{d}\xi$$

Démonstration.
On aimerait calculer l'intégrale de droite. Puisque la transformée de Fourier de f est elle-même définie par une intégrale, on a tout d'abord l'idée d'écrire l'intégrale double :

$$\int_{\mathbf{R}} e^{2i\pi\xi x} \hat{f}(\xi) \mathrm{d}\xi = \int_{\mathbf{R}} \int_{\mathbf{R}} e^{2i\pi\xi(x-y)} f(y) \mathrm{d}y \mathrm{d}\xi$$

et de chercher à d'appliquer le théorème de Fubini... mais $|e^{2i\pi\xi(x-y)} f(y)| = |f(y)|$ n'est pas intégrable sur $\mathbf{R}^2$! On va résoudre ce problème en ajoutant un poids bien choisi.

Lemme 0.4.3.
Pour tout réel $\delta > 0$, la fonction $\varphi_\delta : x \mapsto e^{-\pi\delta^2 x^2} \in \mathrm{L}^1(\mathbf{R})$ vérifie $\hat{\varphi}_\delta : \xi \mapsto \frac{1}{\delta} e^{-\pi \frac{\xi^2}{\delta^2}}$.

Démonstration.
Soient $\xi \in \mathbf{R}$ et $\delta > 0$. On doit calculer :

$$\hat{\varphi}_\delta(\xi) = \int_{\mathbf{R}} e^{-2i\pi\xi x} e^{-\pi\delta^2 x^2} \mathrm{d}x$$

Heuristique : on sait intégrer les expressions du type $e^{\alpha x}$, $\alpha \in \mathbf{C}$ ou $e^{\beta x^2}$, $\beta \in \mathbf{R}$ mais pas les deux ensemble. En revanche, si on remplace $\mathrm{i}\xi \in \mathrm{i}\mathbf{R}$ par $u \in \mathbf{R}$, on obtient :

$$\int_{\mathbf{R}} e^{-2\pi u x} e^{-\pi\delta^2 x^2} \mathrm{d}x = \int_{\mathbf{R}} e^{-\pi(\delta^2 x^2 + 2xu)} \mathrm{d}x = \int_{\mathbf{R}} e^{-\pi\left(\delta x + \frac{u}{\delta}\right)^2} e^{\pi \frac{u^2}{\delta^2}} \mathrm{d}x = e^{\pi \frac{u^2}{\delta^2}} \int_{\mathbf{R}} e^{-\pi\delta^2 x^2} \mathrm{d}x$$

par invariance par translation. Or, utilisant la valeur de l'intégrale de Gauss, par un changement de variable :

$$\int_{\mathbf{R}} e^{-\pi\delta^2 x^2} \mathrm{d}x = \frac{1}{\delta\sqrt{\pi}} \int_{\mathbf{R}} e^{-y^2} \mathrm{d}y = \frac{1}{\delta}$$

Remplaçant u par $\mathrm{i}\xi$, on obtient $\hat{\varphi}_\delta(\xi) = \frac{1}{\delta} e^{-\pi \frac{\xi^2}{\delta^2}}$. Évidemment il s'agit d'une heuristique, mais nous allons utiliser le théorème de prolongement analytique pour rendre tout cela rigoureux.

On pose

$$f : \left| \begin{array}{ccc} \mathbf{C} \times \mathbf{R} & \to & \mathbf{C} \\ (z, x) & \mapsto & e^{-2\pi z x} e^{-\pi\delta^2 x^2} \end{array} \right.$$

Si $R > 0$,

- Pour $x \in \mathbf{R}$, $f(\cdot, x) \in H(D(0, R))$.
- Pour tout $(z, x) \in D(0, R) \times \mathbf{R}$:

$$|f(z, x)| \leqslant e^{-\pi\delta^2 x^2} e^{2\pi R|x|} \in \mathrm{L}^1(\mathbf{R}).$$

Ainsi, par théorème d'holomorphie sous l'intégrale, la fonction

$$F : \left| \begin{array}{ccc} \mathbf{C} & \to & \mathbf{C} \\ z & \mapsto & \int_{\mathbf{R}} e^{-2\pi z x} e^{-\pi\delta^2 x^2} \mathrm{d}x \end{array} \right.$$

est bien définie et entière car elle l'est sur tout disque du plan complexe. Or le calcul de l'heuristique montre que F coïncide avec $z \mapsto \frac{1}{\delta} e^{\pi \frac{z^2}{\delta^2}} \in H(\mathbf{C})$ sur $\mathbf{R}$ qui est d'accumulation dans $\mathbf{C}$ donc par théorème de prolongement analytique :

$$\forall z \in \mathbf{C}, \qquad F(z) = \frac{1}{\delta} e^{\pi \frac{z^2}{\delta^2}}$$

En particulier, $\hat{\varphi_\delta}(\xi) = F(\mathrm{i}\xi) = \frac{1}{\delta} e^{-\pi \frac{\xi^2}{\delta^2}}$.

□

On a :

$$\varphi_\delta \underset{\delta \to 0}{\longrightarrow} 1 \text{ simplement}$$

et j'admets que

$$(\hat{\varphi_\delta})_{\delta > 0} \text{ est une unité approchée.}$$

Il s'agit de vérifications faciles avec des calculs proches de ceux déjà faits dans l'heuristique.
Soient maintenant $x \in \mathbf{R}$ et $\delta > 0$.On rajoute le poids gaussien φ_δ dans l'intégrale à calculer. On remarque que $\xi \mapsto e^{2i\pi\xi x} \varphi_\delta(\xi)$ est intégrable sur $\mathbf{R}$. On va maintenant pouvoir appliquer le théorème de Fubini sous la forme suivante.

Lemme 0.4.4.
Soit $f, g \in \mathrm{L}^1(\mathbf{R})$. On a $\int_{\mathbf{R}} f\hat{g}\mathrm{d}\lambda = \int_{\mathbf{R}} \hat{f} g \mathrm{d}\lambda$.

Par le lemme :

$$\int_{\mathbf{R}} e^{2\mathrm{i}\pi\xi x} \varphi_\delta(\xi) \hat{f}(\xi) \mathrm{d}\xi = \int_{\mathbf{R}} \hat{\varphi}_\delta(\xi - x) f(\xi) \mathrm{d}\xi = f \star \hat{\varphi}_\delta(x)$$

où la convolution finale est bien définie car $f, \hat{\varphi}_\delta \in \mathrm{L}^1(\mathbf{R})$. Utilisant que $(\hat{\varphi}_\delta)_{\delta>0}$ est une identité approchée, par le théorème d'approximation, $f \star \hat{\varphi}_\delta \underset{\delta \to 0}{\longrightarrow} f$ dans $\mathrm{L}^1(\mathbf{R})$ et on peut donc trouver une suite $(\delta_n)_{n \in \mathbf{N}}$ de réels strictement positifs tels que $\delta_n \underset{n \to +\infty}{\longrightarrow} 0$ et $f \star \hat{\varphi}_{\delta_n} \underset{n \to +\infty}{\longrightarrow} f$ presque partout sur $\mathbf{R}$.

Or, on a d'autre part :

- $\varphi_\delta \underset{\delta \to 0}{\longrightarrow} 1$ simplement ;
- pour tout $\delta > 0$, $\xi \mapsto e^{2\mathrm{i}\pi\xi x} \varphi_\delta(\xi) \hat{f}(\xi)$ est mesurable ;
- pour tout $(\delta, \xi) \in \mathbf{R}_+^* \times \mathbf{R}$, $|e^{2\mathrm{i}\pi\xi x} \varphi_\delta(\xi) \hat{f}(\xi)| \leqslant |\hat{f}(\xi)|$ et $\hat{f} \in \mathrm{L}^1(\mathbf{R})$.

Donc par théorème de convergence dominée, on a :

$$\int_{\mathbf{R}} e^{2\mathrm{i}\pi\xi x} \varphi_\delta(\xi) \hat{f}(\xi) \mathrm{d}\xi \underset{\delta \to 0}{\longrightarrow} \int_{\mathbf{R}} e^{2\mathrm{i}\pi\xi x} \hat{f}(\xi) \mathrm{d}\xi$$

ce qui conclut.

□

0.5 Théorème de Paley-Wiener

Leçons : 207 ; 235 ; 236 ; 239 ; 245 ; 250

Références :

— Clarence Kineider, `http://perso.eleves.ens-rennes.fr/people/clarence.kineider/agreg.html`

Théorème 0.5.1 (de Paley-Wiener)**.**
Pour $h \in \mathrm{L}^1(\mathbf{R})$, on note $\hat{h}$ sa transformée de Fourier définie par $\forall x \in \mathbf{R}, \quad \hat{h}(x) = \int_{\mathbf{R}} e^{2\mathrm{i}\pi xt} h(t)\mathrm{d}t$.

1. *Soit $\varphi \in \mathscr{C}_c^\infty(\mathbf{R})$ avec $\operatorname{supp}\varphi \subset [-r, r]$. Alors φ admet un prolongement holomorphe F sur $\mathbf{C}$ vérifiant :*

$$\forall N \in \mathbf{N},\ \exists C_N > 0,\ \forall z \in \mathbf{C}, \qquad |F(z)| \leqslant \frac{C_N}{(1+|z|)^N} e^{2\pi r|\mathfrak{Im}(z)|}. \tag{4}$$

2. *Réciproquement, soit $F : \mathbf{C} \to \mathbf{C}$ holomorphe vérifiant 4 pour un $r > 0$, alors il existe une unique fonction $\varphi \in \mathscr{C}_c^\infty(\mathbf{R})$ telle que $\operatorname{supp}\varphi \subset [-r, r]$ et $\hat{\varphi} = F_{|\mathbf{R}}$.*

Ce théorème montre que la transformée de Fourier échange la plus grande décroissance à l'infini avec la plus grande régularité. La condition 4 dit que F doit avoir une croissance au plus polynomiale lors de déplacements horizontaux et au plus exponentielle de paramètre r lors de déplacements verticaux.

Démonstration.
1. La fonction F doit correspondre avec $\hat{\varphi}$ sur $\mathbf{R}$. Or on connaît son expression, on va simplement vérifier qu'on peut remplacer ξ dans $\mathbf{R}$ par z dans $\mathbf{C}$ et qu'on obtient un prolongement holomorphe sur $\mathbf{C}$.
Si $z \in \mathbf{C}$, on pose

$$F(z) = \int_{\mathbf{R}} e^{-2\mathrm{i}\pi zt}\varphi(t)\mathrm{d}t.$$

Cette intégrale est bien définie puisque φ est à support compact et évidemment on a $F_{|\mathbf{R}} = \hat{\varphi}$.
Il reste à montrer que :

1. F est entière ;
2. F vérifie 4.

Pour le premier point, on utilise un théorème d'holomorphie sous l'intégrale. Soit

$$f : \left| \begin{array}{ccc} \mathbf{C}\times\mathbf{R} & \to & \mathbf{C} \\ (z,t) & \mapsto & e^{-2\mathrm{i}\pi zt}\varphi(t) \end{array} \right.$$

— Pour tout $t \in \mathbf{R}$, $f(\cdot, t) \in H(\mathbf{C})$.

— Pour $R > 0$, $z \in D(0,R)$ et pour tout $t \in \mathbf{R}$, $|f(z,t)| \leqslant e^{2\pi|t||\mathfrak{Im}(z)|}|\varphi(t)| \leqslant e^{2\pi rR}|\varphi(t)|$ car φ est à support compact dans $[-r, r]$. Et $t \mapsto e^{2\pi rR}|\varphi(t)|$ est intégrable sur $\mathbf{R}$.

Ainsi, par théorème d'holomorphie sous l'intégrale, F est holomorphe sur le disque $D(0,R)$ pour tout rayon R donc $F \in H(\mathbf{C})$.

Montrons maintenant le second point. Soient $N \in \mathbf{N}$ et $z \in \mathbf{C}$. On calcule :

$$\begin{aligned} |z|^N|F(z)| &= \left| \int_{\mathbf{R}} |z|^N e^{-2\mathrm{i}\pi zt}\varphi(t)\mathrm{d}t \right| \\ &= \left| \int_{\mathbf{R}} \frac{1}{(-2\mathrm{i}\pi)^N} \frac{\mathrm{d}^N}{\mathrm{d}t^N}(e^{-2\mathrm{i}\pi zt})\varphi(t)\mathrm{d}t \right| \\ &\overset{N \text{ IPP}}{=} \frac{1}{(2\pi)^N} \left| \int_R e^{-2\mathrm{i}\pi zt}\varphi^{(N)}(t)\mathrm{d}t \right| \\ &\leqslant e^{2\pi r|\mathfrak{Im}(z)|} \underbrace{\frac{1}{(2\pi)^N} \int_{\mathbf{R}} \varphi^{(N)}(t)\mathrm{d}t}_{\tilde{C}_N < +\infty} \end{aligned}$$

On obtient donc

$$\forall N \in \mathbf{N}, \qquad |z|^N|F(z)| \leqslant \tilde{C}_N e^{2\pi r|\mathfrak{Im}(z)|}$$

et F vérifie alors 4 car par binôme de Newton, si $N \in \mathbf{N}$:

$$\begin{aligned}(1+|z|)^N|F(z)| &= \sum_{k=0}^{N}\binom{N}{k}|z|^k|F(z)| \\ &\leqslant \underbrace{\left(\sum_{k=0}^{N}\binom{N}{k}\tilde{C}_N\right)}_{C_N<+\infty} e^{2\pi r|\mathfrak{Im}(z)|}\end{aligned}$$

2. On doit montrer qu'il existe une unique fonction $\mathscr{C}_c^\infty$ solution. On va raisonner par analyse-synthèse.

Analyse : Soit $\varphi \in \mathscr{C}_c^\infty$ solution. On a alors $\varphi \in \mathrm{L}^1(\mathbf{R})$ et par 4, avec par exemple $N=2$, on a $\hat{\varphi} = F_{|\mathbf{R}} \in \mathrm{L}^1(\mathbf{R})$. Donc le théorème d'inversion de Fourier nous donne l'expression de φ au moins presque partout. Mais puisqu'elle est continue et que l'expression trouvée l'est aussi aussi par théorème de continuité sous l'intégrale, on obtient : pour tout $t \in \mathbf{R}$, $\varphi(t) = \int_{\mathbf{R}} e^{2\mathrm{i}\pi tx}F(x)\mathrm{d}x$.

Synthèse : On définit pour tout $t \in \mathbf{R}$,

$$\varphi(t) = \int_{\mathbf{R}} e^{2\mathrm{i}\pi tx}F(x)\mathrm{d}x.$$

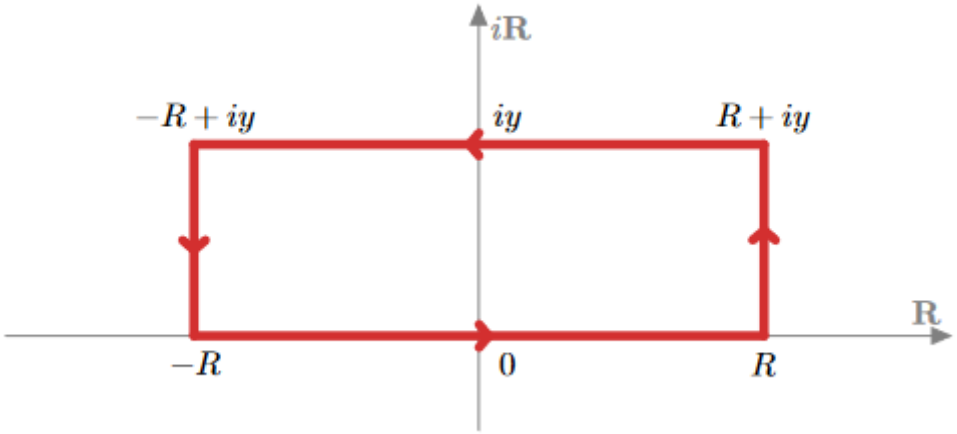
L'intégrande à t fixé est intégrable sur $\mathbf{R}$ puisque F l'est donc φ est bien définie. On montre sans difficulté par théorème de dérivation sous l'intégrale, utilisant 4 pour les dominations, que φ est $\mathscr{C}^\infty$.
Prouvons que supp $\varphi \subset [-r, r]$.

Étape 1 : $\forall t \in \mathbf{R},\ \forall y \in \mathbf{R},\quad \varphi(t) = \int_{\mathbf{R}} e^{2\mathrm{i}\pi t(x+\mathrm{i}y)}F(x+\mathrm{i}y)\mathrm{d}x.$

Autrement dit que φ ne dépend pas de la parallèle horizontale sur laquelle on intègre. On fixe $t \in \mathbf{R}$ et on pose

$$g : \left|\begin{array}{ccc}\mathbf{C} & \to & \mathbf{C} \\ z & \mapsto & e^{2\mathrm{i}\pi tz}F(z)\end{array}\right. \in H(\mathbf{C})$$

On fixe $y \in \mathbf{R}$ et on intègre sur le contour suivant [1] :



On intègre une fonction entière le long d'une courbe fermée donc l'intégrale est nulle.

$$\underbrace{\int_{-R}^{R} g(x)\mathrm{d}x}_{I_1} + \underbrace{\int_0^y g(R+\mathrm{i}s)\mathrm{d}s}_{I_2} - \underbrace{\int_{-R}^{R} g(x+\mathrm{i}y)\mathrm{d}x}_{I_3} - \underbrace{\int_0^y g(-R+\mathrm{i}s)\mathrm{d}s}_{I_4} = 0 \tag{5}$$

Montrons que $I_2,\ I_4 \underset{R\to+\infty}{\longrightarrow} 0$. D'après 4 avec $N=1$, on a, si $s \in [0, y]$:

$$|g(\pm R+\mathrm{i}s)| = \left|e^{2\mathrm{i}\pi t(\pm R+\mathrm{i}s)}F(\pm R+\mathrm{i}s)\right| \leqslant e^{-2\pi ts}\frac{C_1}{1+R}e^{2\pi r|s|} \leqslant \frac{C_1}{1+R}e^{2\pi(|t||s|+r|s|)} \leqslant \frac{C_1}{1+R}e^{2\pi|y|(|t|+r)} \underset{R\to+\infty}{\longrightarrow} 0$$

1. Schéma de Clarence Kineider

d'où une convergence uniforme en $s \in [0, y]$ vers 0 qui permet d'intervertir limite et intégrale donc $I_2, I_4 \underset{R\to+\infty}{\longrightarrow} 0$. Et :

$$I_1 = \int_{-R}^{R} g(x)\mathrm{d}x = \int_{\mathbf{R}} \mathbb{1}_{[-R,R]} g(x)\mathrm{d}x$$

avec l'intégrande dominée par $|F(x)|$ indépendant de R, et de même pour I_3 donc, par théorème de convergence dominée, on peut prendre la limite $R \to +\infty$ dans 5 pour obtenir ce qu'on cherchait à démontrer.

Étape 2 : supp $\varphi \subset [-r, r]$.

Soit maintenant $t \in \mathbf{R}^*$ et $\lambda > 0$. On pose $y = \lambda \operatorname{sign} t$. Par 4 avec $N = 2$: pour tout $x \in \mathbf{R}$,

$$|g(x + \mathrm{i}y)| \leqslant e^{-2\pi t y} \frac{C_2}{(1+|x|)^2} e^{2\pi r|y|} = e^{2\pi\lambda(r-|t|)} \frac{C_2}{(1+|x|)^2}$$

d'où :

$$|\varphi(t)| \leqslant C_2 e^{2\pi\lambda(r-|t|)} \int_{\mathbf{R}} \frac{\mathrm{d}x}{(1+|x|)^2}$$

et prenant $\lambda \to +\infty$, si $|t| > r$ alors $\varphi(t) = 0$ donc supp $\varphi \subset [-r, r]$. □

Remarques :

- En pratique, je ne fais que la démo du point 2. Le premier point se résume à un peu de calcul avec des IPP immédiates et un théorème d'holomorphie sous l'intégrale. En revanche, il donne une application au théorème d'holomorphie sous l'intégrale "originale" et intéressante.
- Idée se cachant derrière le dernier point : pour fixer les idées supposons $t > 0$. Dans le produit $\left|e^{2\mathrm{i}\pi t(x+\mathrm{i}y)} F(x + \mathrm{i}y)\right|$, on a $e^{-2\pi t y} \underset{y\to+\infty}{\searrow} 0$ et $|F(x + iy)| \leqslant C(x) e^{ry} \underset{y\to+\infty}{\nearrow} +\infty$. Mais la croissance du second terme est contrôlé par le paramètre fixe r tandis que le premier terme est contrôlé par le paramètre y que l'on peut prendre aussi grand que l'on veut ! C'est tout l'intérêt de l'introduction de ce nouveau paramètre.

0.6 Théorème de stabilité de Lyapounov

Leçons : 215 ; 220 ; 221

Références :

— [Rou09] François Rouvière, *Petit guide de calcul différentiel à l'usage de la licence et de l'agrégation*

Lemme 0.6.1.
Soit $A \in \mathscr{M}_n(\mathbf{C})$. Pour tout $\varepsilon > 0$ et toute norme d'algèbre $\|\cdot\|$ sur $\mathscr{M}_n(\mathbf{C})$, il existe $C_{\varepsilon,\|\cdot\|} > 0$ telle que :

$$\forall t \in \mathbf{R}, \qquad \|e^{tA}\| \leqslant C_{\varepsilon,\|\cdot\|} e^{t\left(\max\limits_{\lambda\in\sigma(A)} \mathfrak{Re}(\lambda)+\varepsilon\right)}$$

En particulier, si $\max\limits_{\lambda\in\sigma(A)} \mathfrak{Re}(\lambda) < 0$, alors il existe $K > 0$ et $a > 0$ tels que $\|e^{tA}\| \leqslant Ke^{-at}$.

On peut le démontrer avec un peu de calcul en utilisant la décomposition en espaces caractéristiques de A. On peut aussi utiliser le résultat suivant :

Proposition 0.6.2.
Soit $A \in \mathscr{M}_n(\mathbf{C})$. Si γ est un lacet entourant une fois positivement $\sigma(A)$, alors pour toute fonction entière f :

$$f(A) = \frac{1}{2\mathrm{i}\pi} \int_\gamma (zI_n - A)^{-1} f(z)\mathrm{d}z.$$

Idée de démonstration : on montre que les deux membres commutent avec la conjugaison par une matrice inversible. Ce sont également deux fonctions continues de A et la relation est vraie pour les matrices diagonales. Par densité des matrices diagonalisables dans $\mathscr{M}_n(\mathbf{C})$, on obtient le résultat. [Laf10]

Démonstration du lemme.
Il suffit de choisir un lacet simple entourant positivement $\sigma(A)$ et inclus dans $\left\{z \in \mathbf{C} \mid \mathfrak{Re}(z) \leqslant \max\limits_{\lambda\in\sigma(A)} \mathfrak{Re}(\lambda) + \varepsilon\right\}$.
On obtient avec $f : z \mapsto e^{tz}$:

$$e^{tA} = \frac{1}{2\mathrm{i}\pi} \int_\gamma (zI_n - A)^{-1} e^{tz}\mathrm{d}z = \frac{1}{2\mathrm{i}\pi} \int_0^1 (\gamma(s)I_n - A)^{-1} e^{t\gamma(s)}\gamma'(s)\mathrm{d}s$$

et donc

$$\|e^{tA}\| \leqslant \underbrace{\frac{1}{2\pi} \int_0^1 \|(\gamma(s)I_n - A)^{-1}\| \, |\gamma'(s)|\mathrm{d}s}_{C_{\varepsilon,\|\cdot\|}>0} \, e^{t\left(\max\limits_{\lambda\in\sigma(A)} \mathfrak{Re}(\lambda)+\varepsilon\right)}$$

□

Théorème 0.6.3 (de stabilité de Lyapounov)**.**
Soit $f : \mathbf{R}^n \to \mathbf{R}^n$ de classe $\mathscr{C}^1$ telle que $f(0) = 0$. On considère le problème de Cauchy :

$$\left\{ \begin{array}{rcl} y' & = & f(y) \\ y(0) & = & y_0 \end{array} \right. \tag{6}$$

On suppose que la matrice $A := \mathrm{D}f_0$ a toutes ses valeurs propres de parties réelles strictement négatives. Alors 0 est un point asymptotiquement stable, c'est-à-dire :

1. *Pour tout $\varepsilon > 0$, il existe $\delta > 0$ tel que toute solution maximale y dont la valeur initiale est à une distance de l'équilibre $\|y_0\| \leqslant \delta$ est globale et vérifie $\|y(t)\| \leqslant \varepsilon$ pour tout $t \geqslant 0$.*
2. *De plus, il existe $\delta' > 0$ tel que s'il existe $t_0 \in \mathbf{R}^+$ vérifiant $\|y(t_0)\| \leqslant \delta$, alors $y(t) \underset{t\to+\infty}{\longrightarrow} 0$.*

Démonstration.
On définit le produit scalaire B sur $\mathbf{R}^n$ par :

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

On montre qu'il est bien défini grâce au lemme 1 et à l'inégalité de Cauchy-Schwarz. On vérifie très facilement qu'il s'agit bien d'un produit scalaire, mais je vais passer ces vérifications.

Détails : on voit immédiatement qu'il s'agit d'une forme linéaire symétrique. Soient K, a donnés par le lemme 1, on a pour $x, y \in \mathbf{R}^n$:

$$\left|\left\langle e^{tA}x\,,e^{tA}y\right\rangle\right| \leqslant \|e^{tA}x\|\|e^{tA}y\| \leqslant K^2e^{-2at}\|x\|\|y\|$$

donc l'intégrale est absolument convergente. Soit $x \in \mathbf{R}^n$,

$$B(x,x) = \int_0^{+\infty} \|e^{tA}x\|^2\mathrm{d}t \leqslant 0$$

d'où la positivité et si $B(x,x) = 0$, puisque $t \mapsto \|e^{tA}x\|^2$ est continue, positive et d'intégrale nulle, elle doit être nulle et pour $t = 0$, on obtient $x = 0$.

On note q la forme quadratique associée à B et $\|\cdot\|_q$ la norme induite :

$$\forall x \in \mathbf{R}^n, \qquad \|x\|_q = \sqrt{q(x)} = \sqrt{B(x,x)}$$

Puisque f est de classe $\mathscr{C}^1$, elle est localement Lipschitz donc, par théorème de Cauchy-Lipschitz, la problème 6 admet une unique solution maximale définie sur un intervalle de temps ouvert dans $[0, +\infty[$.
Notons y la solution maximale de 6 définie sur $[0, T[$.

Étape 1 : $(q \circ y)'(t) \leqslant -\beta\|y(t)\|_q^2$ dès que $y(t)$ est assez petit.

On définit $r(x) := f(x) - Ax = \underset{x\to 0}{o}(x)$ pour $x \in \mathbf{R}^n$. Il s'agit de l'écart avec l'approximation linéaire de f.
Alors, si $t \in [0, T[$,

$$(q \circ y)'(t) = \mathrm{D}q_{y(t)}.y'(t) = 2B(y(t), y'(t)) = 2B\big(y(t), Ay(t)\big) + 2B\big(y(t), r(y(t))\big)$$

On va appliquer la Grande Idée du Calcul Différentiel selon François Rouvière, c'est-à-dire réduire le problème non-linéaire en un problème d'algèbre linéaire, essentiellement traité par le lemme, et en majorations de normes pour le reste.
Les normes en dimension finie sont toutes équivalentes. On souhaite travailler avec la norme $\|\cdot\|_q$.
On note $c > 0$ telle que $c\|\cdot\|_q \leqslant \|\cdot\|$. Évaluons la contribution de la partie linéaire :

$$\forall x \in \mathbf{R}^n, \quad 2B(x, Ax) = 2\int_0^{+\infty} \left\langle e^{tA}x\,, e^{tA}Ax\right\rangle \mathrm{d}t = \left[\|e^{tA}x\|^2\right]_0^{+\infty} = -\|x\|^2 \leqslant -c^2\|x\|_q^2$$

car $e^{tA} \underset{t\to+\infty}{\longrightarrow} 0$ par le lemme 1.

Évaluons maintenant la contribution du reste. Soient alors $\varepsilon > 0$ suffisamment petit pour que $\beta := c^2 - 2\varepsilon > 0$ et $\alpha > 0$ tel que :

$$\|x\|_q \leqslant \alpha \implies \|r(x)\|_q \leqslant \varepsilon\|x\|_q$$

Supposons que $\|y(t)\|_q \leqslant \alpha$. Par inégalité de Cauchy-Schwarz,

$$|B(y(t), r(y(t)))| \leqslant \|y(t)\|_q\,\|r(y(t))\|_q \leqslant \varepsilon\|y(t)\|_q^2$$

et donc

$$(q \circ y)'(t) \leqslant \underbrace{(-c^2 + 2\varepsilon)}_{-\beta}\|y(t)\|_q^2$$

Étape 2 : y est alors globale.

Supposons $\|y_0\|_q < \alpha$ et montrons que pour tout $t \in [0, T[$, $\|y(t)\|_q < \alpha$. Supposons par l'absurde qu'il existe $t \in]0, T[$ tel que $\|y(t)\|_q \geqslant \alpha$. Alors, on peut considérer par théorème des valeurs intermédiaires :

$$t_0 = \min\left\{t \in]0, T[\mid \|y(t)\|_q = \alpha\right\}.$$

Mais alors par l'étape 1 :

$$(q \circ y)'(t_0) \leqslant -\beta\alpha^2 < 0$$

et $t \mapsto \|y(t)\|$ est strictement décroissante au voisinage de t_0, absurde.
Par conséquent, la trajectoire reste confinée dans la boule de rayon α et, $\mathbf{R}^n$ étant de dimension finie, par théorème

de sorti de tout compact, y est globale.

Remarque : on a montré que si le départ de la trajectoire était à une distance au plus α, alors elle restait à une distance au plus α. Puisque α peut être choisi arbitrairement petit, on a ainsi montré la stabilité.

Étape 3 : Stabilité asymptotique

On déduit de ce qui précède que si $\|y_0\| < \alpha$, $(q \circ y)' \leq -\beta(q \circ y)$ donc

$$e^{\beta t}[(q \circ y)' + \beta(q \circ y)] = \frac{\mathrm{d}}{\mathrm{d}t}\left[e^{\beta t}(q \circ y)\right] \leqslant 0$$

et $t \mapsto e^{\beta t} q(y(t))$ est décroissante sur $\mathbf{R}^+$ d'où $e^{\beta t} q(y(t)) \leqslant q(y_0)$ soit $\|y(t)\|_q \leqslant e^{-\frac{\beta}{2}t}\|y_0\|_q \underset{t\to+\infty}{\longrightarrow} 0$ avec convergence exponentielle. □

0.7 Théorème de Hadamard-Lévy

Leçons : 203 ; 204 ; 214 ; 215 ; 220 ; 267

Références :

— Clarence Kineider, `http://perso.eleves.ens-rennes.fr/people/clarence.kineider/agreg.html`

Théorème 0.7.1 (de Hadamard-Lévy)**.**
Soit $f : \mathbf{R}^n \to \mathbf{R}^n$ de classe $\mathscr{C}^2$. Supposons que :

(i) pour tout $x \in \mathbf{R}^n$ $\mathrm{d}f_x$ est inversible ;

(ii) f est propre, i.e. l'image réciproque d'un compact par f est compacte.

Alors f est un $\mathscr{C}^2$-difféomorphisme global.

Le résultat est vrai si on remplace $\mathscr{C}^2$ par $\mathscr{C}^1$ mais est beaucoup plut technique à démontrer car on ne peut plus utiliser le théorème de Cauchy-Lipschitz pour l'unicité de la solution et la continuité vis-à-vis de la condition initiale. En fait, ce théorème est un cas particulier du suivant en utilisant la simple connexité de $\mathbf{R}^n$: si la variété but est connexe, tout difféomorphisme local propre entre variétés est un revêtement.

Démonstration.
Il suffit pour prouver le théorème de montrer que f est bijective. En effet, on aura alors f injective avec une différentielle inversible en tout point donc le théorème d'inversion global assurera que f est un $\mathscr{C}^2$-difféomorphisme sur son image égale à $\mathbf{R}^n$ par surjectivité. Soit $y \in \mathbf{R}^n$. On cherche à résoudre l'équation $f(x) = y$ et à montrer qu'elle admet une unique solution. Posons $g : x \mapsto f(x) - y$ le translaté de f par y et $S = g^{-1}(\{0\})$. Remarquons que f et g ne diffèrent que d'une constante, donc g vérifient les deux mêmes hypothèses que f. On transforme le problème en une recherche de 0 de fonction. On cherche maintenant à montrer que S est un singleton. Pour trouver les zéros d'une fonction, on dispose de la méthode de Newton dont on va ici utiliser une version continue.

Étape 1 : Schéma de Newton continu.

Soit $F : x \mapsto \mathrm{d}g_x^{-1}.g(x)$. Pour $q \in \mathbf{R}^n$, on pose :

$$\begin{cases} x'(t) &= -F(x(t)) \\ x(0) &= q \end{cases} \tag{7}$$

Puisque g est $\mathscr{C}^2$, F est localement lipschitzienne. Par le théorème de Cauchy-Lipschitz, pour tout $q \in \mathbf{R}^n$, il existe une unique solution maximale à 7 définie sur un intervalle ouvert $[0, T[$, on la note x_q.

Étape 2 : Toute trajectoire est globale.

Pour $t \in [0, T[$, on a :

$$(g \circ x_q)'(t) = \mathrm{d}g_{x_q(t)}.x_q'(t) = -g(x_q(t))$$

donc $(g \circ x_q)$ vérifie l'équation $y' = -y$ avec la condition initiale $g(x_q(0)) = g(q)$ d'où :

$$\forall t \in [0, T[, \qquad g(x_q(t)) = g(q)e^{-t} \tag{8}$$

et si $t \in [0, T[$, $\|g(x_q(t))\| \leqslant \|g(q)\|$ donc $x_q(t) \in g^{-1}\left(\bar{B}(0, \|g(q)\|)\right)$ qui est compacte car g est propre. Puisque $\mathbf{R}^n$ est de dimension finie, par théorème de sorti de tout compact, la trajectoire est globale, soit $T = +\infty$.

Étape 3 : Tout $x \in S$ est un équilibre asymptotiquement stable.

On remarque que pour $x \in \mathbf{R}^n$, $F(x) = 0 \iff g(x) = 0 \iff x \in S$ donc S est l'ensemble des points d'équilibre. Pour montrer la stabilité asymptotique, on veut utiliser le théorème de stabilité de Lyappounov, on a donc besoin de trouver les valeurs propres de la différentielle de $-F$ en x. Si $x \in S$,

$$\mathrm{d}F_x = \underbrace{\mathrm{d}\left(y \mapsto \mathrm{d}g_y^{-1}\right)_x}_{\text{app. linéaire}}.\underbrace{g(x)}_{\text{vecteur}} + \mathrm{d}g_x^{-1} \circ \mathrm{d}g_x = \mathrm{id}$$

Explications : On différencie un produit *matrice* × *vecteur* (bilinéaire) et on ne calcule surtout pas le premier terme, on remarque simplement que $g(x) = 0$ donc ce terme se résume en une application linéaire appliquée au vecteur nulle donc est nul.
Donc $\sigma(-\mathrm{d}F_x) = \{-1\}$, toutes ses valeurs propres sont de partie réelles strictement négatives, donc par théorème de

stabilité de Lyapounov, x est asymptotiquement stable.
On note $W_x = \left\{ q \in \mathbf{R}^n \mid x_q(t) \underset{t\to+\infty}{\longrightarrow} x \right\}$ son bassin d'attraction.

Étape 4 : $\mathbf{R}^n = \bigsqcup_{x\in S} W_x$.

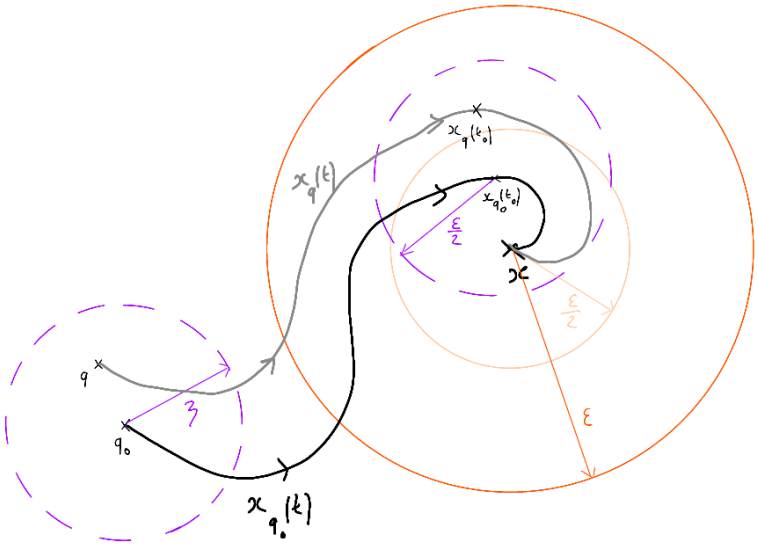
Remarquons déjà qu'une trajectoire ne peut pas converger vers deux points distincts, donc les bassins d'attraction sont disjoints. On cherche donc à montrer que toutes les trajectoires convergent vers un point d'équilibre. Soit $q \in \mathbf{R}^n$ le départ d'une trajectoire. On a montré que toutes les trajectoires étaient bornées donc il existe une suite de temps croissants $(t_k)_{k\in\mathbf{N}}$ donnant une suite de points de cette trajectoire qui converge, soit vérifiant :

$$t_k \underset{k\to+\infty}{\longrightarrow} +\infty \qquad \text{et} \qquad x_q(t_k) \underset{k\to+\infty}{\longrightarrow} l \in \mathbf{R}^n$$

Mais par 8, $g(x_q(t_k)) \underset{k\to+\infty}{\longrightarrow} 0$ donc $g(l) = 0$ par continuité de g, i.e. $l \in S$. Puisque on a montré que l était asymptotiquement stable, dès que la suite se rapproche suffisamment de l, toute la trajectoire converge vers l, donc $q \in W_l$ et $\mathbf{R}^n = \bigsqcup_{x\in S} W_x$.

Étape 5 : Les W_x sont ouverts.

Soit $x \in S$. Puisque x est asymptotiquement stable, il existe $\varepsilon > 0$ tel que $B(x,\varepsilon) \subset W_x$. Toute trajectoire rentrant dans cette boule convergera vers x. Soit $q_0 \in W_x$.
On souhaite montrer que toute trajectoire partant suffisamment proche de q convergera également vers x. Soit $t_0 \in \mathbf{R}^+$ tel que $x_q(t_0) \in B(x,\varepsilon/2)$. La fonction $-F$ étant localement lipschitz, par continuité de la solution par rapport à la condition initiale, il existe $\eta > 0$ tel que pour tout $q \in B(q_0,\eta)$, $x_q(t_0) \in B(x_{q_0}(t_0),\varepsilon/2)$. Par inégalité triangulaire, on a alors $x_q(t_0) \in B(x,\varepsilon)$ donc $q \in W_x$. Ainsi, $B(q_0,\eta) \subset W_x$ et W_x est ouvert.



Étape 6 : S est un singleton.

On a $\mathbf{R}^n = \bigsqcup_{x\in S} W_x$ avec les W_x ouverts. Puisque $\mathbf{R}^n \neq \emptyset$, S est non-vide et puisque $\mathbf{R}^n$ est connexe, S ne peut posséder au plus qu'un élément d'où le résultat. □

Application :

— Perturbation difféomorphe de l'identité `https://www.grenoble-sciences.fr/pap-ebook/lafontaine/sites/lafontaine/files/pdf/cp3_1.pdf` :
La restriction à la boule unité d'un isomorphisme affine assez proche de l'identité se prolonge en un difféomorphisme de $\mathbf{R}^n$ égal à l'identité hors de la boule de rayon 2.

0.8 Réduction des opérateurs compacts symétriques

Leçons : 203 ; 205 ; 208 ; 213

Références :

— Salim Rostam, `https://minerve.ens-rennes.fr/images/Reduction_operateurs_compacts_symetriques.pdf`

Soient $(\mathscr{H}, \langle\cdot\,,\cdot\rangle)$ un espace de Hilbert et $T \in \mathscr{L}(E)$ un opérateur compact symétrique non nul.

Remarque 0.8.1. Un opérateur compact est continu car l'image de la boule unité est bornée.

Proposition 0.8.2.
Si $\lambda \neq 0$, $\ker(\lambda \operatorname{Id} - T)$ est de dimension finie.

Démonstration.
Notons $F = \ker(\lambda \operatorname{Id} - T)$. C'est un sous-espace fermé de $\mathscr{H}$ car T est continue et :

$$\bar{B}(F) = \frac{1}{\lambda} T\left(\bar{B}(F)\right) \subset \frac{1}{\lambda}\overline{T\left(\bar{B}(E)\right)} \cap F$$

où $\bar{B}(F)$ est la boule unité fermé de F. Or T étant compact, le dernier ensemble est compact. Donc, par théorème de Riesz, F est de dimension finie. □

Théorème 0.8.3.
Soient $(\mathscr{H}, \langle\cdot\,,\cdot\rangle)$ un $\mathbf{R}$-espace de Hilbert et $T \in \mathscr{L}(E)$ un opérateur compact symétrique non nul.
Il existe une base hilbertienne de $\mathscr{H}$ formée exclusivement de vecteurs propres de T, l'ensemble de ces vecteurs associés à une valeur propre non-nulle étant au plus dénombrable.
De plus, si (e_n) désigne cette famille et (λ_n) les valeurs propres correspondantes, alors soit elle est finie, soit on peut la réordonner de sorte que $|\lambda_n| \searrow 0$. En outre,

$$\forall x \in \mathscr{H}, \qquad Tx = \sum_{n\in\mathbf{N}} \lambda_n \langle x\,, e_n\rangle\, e_n$$

Démonstration.
Ce théorème généralise le théorème de réduction des endomorphismes auto-adjoints à la dimension infinie. On va d'ailleurs le prouver en suivant la même idée : commencer par trouver une valeur propre puis appliquer une hypothèse de récurrence pour diagonaliser T sur un supplémentaire de l'espace propre associée à la valeur propre trouvée.

Étape 1 : $\|T\|^2 \in \operatorname{vp}(T^2)$.

Soit $x \in \mathscr{H}$ non nul. Utilisant la symétrie de T,

$$\begin{aligned}\left\|T^2x - \|T\|^2 x\right\|^2 &= \|T^2x\|^2 + \|T\|^4\|x\|^2 - 2\|T\|^2\underbrace{\langle T^2x\,, x\rangle}_{\|Tx\|^2} \\ &\leqslant 2\|T\|^2\left(\|T\|^2\|x\|^2 - \|Tx\|^2\right)\end{aligned}$$

donc

$$\left\|\frac{T^2x}{\|T\|^2} - x\right\|^2 \leqslant 2\left(\|T\|^2\|x\|^2 - \|Tx\|^2\right).$$

On cherche un x qui annule le terme de gauche. Par définition de la norme d'opérateur il existe une suite $(x_n)_{n\in\mathbf{N}}$ de vecteurs unitaires tels que $\|Tx_n\| \to \|T\|$. Mais T est compact et (x_n) est bornée donc, quitte à extraire, on peut supposer $Tx_n \underset{n\to+\infty}{\longrightarrow} y \neq 0$, car $\|y\| = \|T\|$ par définition des x_n et $T^2x_n \underset{n\to+\infty}{\longrightarrow} Ty$.
De plus, T étant non-nul, on peut poser $z := \frac{1}{\|T\|^2}Ty$ et alors évaluant l'inégalité en x_n, on obtient $x_n \to z$ donc $z \neq 0$ et $T^2z = \|T\|^2 z$.

Étape 2 : $\|T\|$ ou $-\|T\|$ est dans $\operatorname{vp}(T)$.

On a donc

$$(T^2 - \|T\|^2\operatorname{Id})z = (T + \|T\|\operatorname{Id})\underbrace{(T - \|T\|\operatorname{Id})z}_{z'} = 0$$

et soit $z' = 0$ et alors $Tx = \|T\|x$, soit $z' \neq 0$ et $Tz' = -\|T\|z'$. Dans les deux cas, on a le résultat.

Étape 3 : Construction de (λ_n).

On note $\lambda_1 \in \mathrm{vp}(T)$ avec $|\lambda_1| = \|T\|$ donné par l'étape précédente.
Soit $n \in \mathbf{N}^*$. Si $(\lambda_k)_{1\leqslant k\leqslant n}$ est construite, on pose :

$$T_{n+1} = T_{\left|\left[\bigoplus_{k=1}^{n} \ker(\lambda_k \operatorname{Id} - T)\right]^\perp\right.}$$

et

- si $T_{n+1} = 0$, alors on s'arrête, la suite (λ_n) est alors finie ;
- sinon, T_{n+1} est un opérateur compact symétrique non-nul donc par l'étape 2, il existe $\lambda_{n+1} \in \mathrm{vp}(T_{n+1}) \subset \mathrm{vp}(T)$ et $\lambda_{n+1} \neq \lambda_k$ pour $1 \leqslant k \leqslant n$ par définition de T_{n+1}. De plus,

  $$|\lambda_{n+1}| = \|T_{n+1}\| \leqslant \|T_n\| = |\lambda_n|$$

  car $T_{n+1} = (T_n)_{|[\ker(\lambda_n \operatorname{Id} - T)]^\perp}$.

Donc $(|\lambda_n|)$ est décroissante.

Étape 4 : Si (λ_n) est infinie, alors $|\lambda_n| \searrow 0$.

La suite $(|\lambda_n|)$ est décroissante et minorée par 0 donc $|\lambda_n| \to \lambda \geqslant 0$. Supposons par l'absurde que $\lambda > 0$. Soit, pour tout $n \in \mathbf{N}^*$, e_n vecteur propre unitaire associé à λ_n. Alors pour tout $n \in \mathbf{N}^*$,

$$\left\|\frac{1}{\lambda_n} e_n\right\| = \frac{1}{|\lambda_n|} \leqslant \frac{1}{\lambda}$$

donc $\left(\frac{1}{\lambda_n} e_n\right)$ est bornée. Son image par T admet une sous-suite convergente car T est compact. Mais $\left(T\left(\frac{1}{\lambda_n} e_n\right)\right) = (e_n)$ qui est orthonormée comme vecteurs propres unitaires associée à des valeurs propres distinctes d'un opérateur symétrique donc n'admet pas de valeur d'adhérence, ce qui est absurde.

Étape 5 : Diagonalisation de T.

Posons $F = \bigoplus_{n\geqslant 1} \ker(\lambda_n \operatorname{Id} - T)$ et montrons que $F^\perp \subset \ker T$.

- Si (λ_n) est finie, par construction, cela signifie que $T_{|F^\perp} = 0$ soit $F^\perp \subset \ker T$.
- Sinon, si $x \in F^\perp$, utilisant les notations de l'étape 3, on peut appliquer T_n à x pour tout $n \in \mathbf{N}$ et

  $$\|Tx\| = \|T_n x\| \leqslant \|T_n\|\|x\| = |\lambda_n|\|x\| \underset{n\to+\infty}{\longrightarrow} 0$$

  donc $Tx = 0$ et $x \in \ker T$.

Puisque $\bar{F} = F^{\perp\perp}$, on obtient $(\ker T)^\perp \subset \bar{F}$. De plus, $\ker T$ est fermé car T est continu donc par théorème de projection orthogonale :

$$\mathscr{H} = \ker T \oplus (\ker T)^\perp = \ker T + \bar{F}.$$

On a a priori perdu le caractère direct de la somme. Mais T étant symétrique, ses sous-espaces propres sont orthogonaux deux à deux, donc F, qui est somme orthogonale d'espaces propres distincts du noyau, est orthogonal à $\ker T$. Finalement,

$$\mathscr{H} = \ker T \oplus \bar{F} = \overline{\bigoplus_{n\geqslant 1} \ker(\lambda_n \operatorname{Id} - T)} \oplus \ker T.$$

En concaténant des bases algébriques finies des espaces propres associés aux valeurs propres non nulles, on obtient une base hilbertienne de $\bar{F}$. On concatène alors cette base avec une base hilbertienne du noyau pour avoir le résultat. □

0.9 Optimisation dans un Hilbert

Leçons : 213 ; 219 ; 223 ; 229 ; 253

Références :

— [IP19] Lucas Isenmann, Timothée Pecatte, *L'oral à l'agrégation de mathématiques*

Théorème 0.9.1.
Soient $\mathscr{H}$ un espace de Hilbert sur $\mathbf{R}$ et $J : \mathscr{H} \to \mathbf{R}$ une fonction continue, convexe et coercive : pour toute suite (x_n) de $\mathscr{H}$, si $\|x_n\| \to +\infty$, alors $|J(x_n)| \to +\infty$. Alors, il existe $a \in \mathscr{H}$ tel que :

$$J(a) = \inf_{\mathscr{H}} J$$

Démonstration.
Soit $(x_k)_{k\in\mathbf{N}}$ une suite minimisante : $J(x_k) \to \inf_{\mathscr{H}} J =: m$.

Étape 1 : Pas d'infimum à l'infini.

Supposons (x_k) non bornée. Il existe alors une sous-suite $(x_{\varphi(k)})$ telle que $\|x_{\varphi(k)}\| \to +\infty$ et , J étant coercive, $|J(x_k)| \to +\infty$, ce qui est absurde pour une suite minimisante. Donc (x_k) est bornée par $C > 0$.

On note $F := \mathrm{Vect}(x_k \ ; \ k \in \mathbf{N})$.

Étape 2 : Construction de (y_k) telle que : $\forall v \in F$, $(\langle v \, , y_k\rangle)_k$ converge.

L'idée va être de construire y_k par extraction diagonale des x_k. Puisque la suite des x_k est bornée, la suite de réels $(u_k) = (\langle x_0 \, , x_k\rangle)_{k\in\mathbf{N}}$ est bornée par Cauchy-Schwarz. Par théorème de Bolzano-Weierstrass, elle admet une sous-suite convergente $(u_{\varphi_0(k)})$.

Par récurrence, supposons avoir construit $\varphi_0, \dots, \varphi_i$ des extractions telles que $\left(\left\langle x_i \, , x_{\varphi_0\circ\cdots\circ\varphi_i(k)}\right\rangle\right)_k$ converge. Comme précédemment, $\left(\left\langle x_{i+1} \, , x_{\varphi_0\circ\cdots\circ\varphi_i(k)}\right\rangle\right)_k$ est bornée donc admet une sous-suite convergente

$$\left(\left\langle x_{i+1} \, , x_{\varphi_0\circ\cdots\circ\varphi_i\circ\varphi_{i+1}(k)}\right\rangle\right)_k .$$

On pose alors

$$\Psi : \left| \begin{array}{ccc} \mathbf{N} & \to & \mathbf{N} \\ k & \mapsto & \varphi_0 \circ \cdots \circ \varphi_k(k) \end{array} \right.$$

qui est strictement croissante car les φ_i le sont. On pose $y_k := x_{\Psi(k)}$ pour tout $k \in \mathbf{N}$. Pour tout $i \in \mathbf{N}$, $(\langle x_i \, , y_k\rangle)_{k\geqslant i}$ est extraite de $\left(\left\langle x_i \, , x_{\varphi_0\circ\cdots\circ\varphi_i(k)}\right\rangle\right)_k$ donc converge. Par bilinéarité du produit scalaire, on obtient ce qu'on cherchait à démontrer.

Étape 3 : $\forall u \in \mathscr{H}$, $(\langle u \, , y_k\rangle)_k$ converge.

Par théorème de projection orthogonale, on a $\mathscr{H} = \bar{F} \oplus F^\perp$.
Soient $u \in \mathscr{H}$ ainsi que $v \in \bar{F}$ et $w \in F^\perp$ tels que $u = v + w$. Soient $\varepsilon > 0$ et $\tilde{v} \in F$ tel que $\|v - \tilde{v}\| \leqslant \varepsilon$. Pour tous $k, l \in \mathbf{N}$,

$$|\langle u \, , y_k - y_l\rangle| = |\langle v \, , y_k - y_l\rangle| \leqslant \|v - \tilde{v}\| \, \|y_k - y_l\| + |\langle \tilde{v} \, , y_k - y_l\rangle| .$$

La suite $(\langle \tilde{v} \, , y_k\rangle)_{k\in\mathbf{N}}$ est convergente donc a fortiori de Cauchy. Il existe $N \in \mathbf{N}$ tel que : $\forall k, l \geqslant N$, $|\langle \tilde{v} \, , y_k - y_l\rangle| \leqslant \varepsilon$ et alors

$$|\langle u \, , y_k - y_l\rangle| \leqslant \varepsilon \left(\|y_k\| + \|y_l\|\right) + \varepsilon \leqslant (2C + 1)\varepsilon$$

ce qui montre que $(\langle u \, , y_k\rangle)_{k\in\mathbf{N}}$ est de Cauchy dans $\mathbf{R}$ qui est complet donc converge vers $l_u \in \mathbf{R}$.

Étape 4 : Trouver $a \in \mathscr{H}$ tel que : $\forall u \in \mathscr{H}$, $\langle u \, , y_k\rangle \to \langle u \, , a\rangle$.

On note

$$\Phi : \left| \begin{array}{ccc} \mathscr{H} & \to & \mathbf{R} \\ u & \mapsto & l_u \end{array} \right.$$

On va montrer qu'elle est linéaire et continue.
Linéaire : soit $u_1, u_2 \in \mathscr{H}$, $\lambda \in \mathbf{R}$. On a :

$$l_{u_1+\lambda u_2} = \lim_{k\to+\infty} \langle u_1 + \lambda u_2 \,, y_k \rangle = \lim_{k\to+\infty} \langle u_1 \,, y_k \rangle + \lambda \lim_{k\to+\infty} \langle u_2 \,, y_k \rangle = l_{u_1} + \lambda l_{u_2}$$

Continuité : pour tout $u \in \mathscr{H}$, $|\Phi(u)| \leqslant C\|u\|$ par Cauchy-Schwarz avec (y_k) bornée par C.

Par théorème de représentation de Riesz, il existe (un unique) $a \in \mathscr{H}$ tel que $\forall u \in \mathscr{H}$, $\Phi(u) = \langle u \,, a \rangle$ ce qui conclut cette étape.

Étape 5 : $J(a) = m$.

Pour $\beta > m$, on définit $C_\beta := \{x \in \mathscr{H} \mid J(x) \leqslant \beta\}$ qui est convexe par convexité de J, fermé car J est continue et non vide par définition de l'infimum. On note $p : \mathscr{H} \to C_\beta$ la projection orthogonale sur C_β.

La suite y est extraite de x donc $J(y_k) \to m$ et il existe un rang $N \in \mathbf{N}$ tel que $\forall k \geqslant N$, $y_k \in C_\beta$. On utilise alors la caractérisation du projeté orthogonal :

$$\forall k \geqslant N, \qquad \langle y_k - p(a) \,, a - p(a) \rangle \leqslant 0$$

Or $\langle y_k \,, a - p(a) \rangle \to \langle a \,, a - p(a) \rangle$ donc $\|a - p(a)\|^2 \leqslant 0$ soit $a = p(a)$ et $J(a) \leqslant \beta$.
Ceci étant vrai pour tout $\beta > m$, on en déduit que $J(a) \leqslant m$ et donc $J(a) = m$. □

0.10 Théorème de Riesz-Fischer

Leçons : 201 ; 205 ; 234 ; 241

Références :

— [BP18] Marc Briane, Gilles Pagès, *Théorie de l'intégration : analyse, convolution et transformée de Fourier*

Théorème 0.10.1 (de Riesz-Fischer)**.**
Soient (X, μ) un espace mesuré et $p \in [1, +\infty]$. Alors $\mathrm{L}^p(X, \mu)$ est complet pour la norme $\|\cdot\|_p$. De plus, toute suite convergente dans L^p admet une sous-suite qui converge presque partout.

Démonstration.
On va séparer les cas $p = +\infty$ et p fini. Dans les deux cas, on utilisera la complétude de $\mathbf{C}$ pour exhiber une limite pour la convergence simple presque partout, puis on montrera qu'il y a convergence pour la norme p ce qui donnera en même temps que cette limite est elle-même dans L^p.
Soit $(f_n)_{n\in\mathbf{N}}$ une suite de Cauchy dans L^p.

1) Cas $p = +\infty$.

On pose :

$$\forall m, n \in \mathbf{N}, \qquad B_{m,n} = \{x \in X \mid |f_m(x) - f_n(x)| > \|f_m - f_n\|_\infty\}.$$

Ces ensembles sont de mesures nulles puisque les $f_m - f_n$ sont dans L^∞ donc $E = \bigcup_{m,n\in\mathbf{N}} B_{m,n}$ est de mesure nulle comme union dénombrable d'ensembles négligeables.

Étape 1 : Trouver une limite simple sur $X \setminus E$.

Soit $x \in X \setminus E$. Pour tout $m, n \in \mathbf{N}$, $|f_m(x) - f_n(x)| \leqslant \|f_m - f_n\|_\infty$ et (f_n) étant de Cauchy dans L^∞, $(f_n(x))$ est de Cauchy dans $\mathbf{C}$ qui est complet donc il existe $f(x) \in \mathbf{C}$ tel que

$$f_n(x) \underset{n\to+\infty}{\longrightarrow} f(x)$$

La fonction $x \mapsto f(x)$ vérifie $f_n \overset{CS}{\to} f$ presque partout. Cela permet déjà de montrer le second point du théorème dans le cas $p = +\infty$: une suite convergente dans L^∞ est de Cauchy, et on vient de montrer que toute suite de Cauchy admettait une limite pour la convergence simple presque partout. Donc dans le cas $p = +\infty$, on a même pas besoin de prendre une sous-suite.

Étape 2 : $f_n \to f$ dans $(\mathrm{L}^\infty, \|\cdot\|_\infty)$.

Soit $\varepsilon > 0$. Il existe $N \in \mathbf{N}$ tel que $\forall m, n \geqslant N$, $\|f_m - f_n\| < \varepsilon$ et alors

$$\forall m, n \geqslant N, \qquad |f_m(x) - f_n(x)| < \varepsilon$$

Prenant la limite $[m \to +\infty]$, on obtient

$$\forall n \geqslant N, \qquad |f(x) - f_n(x)| < \varepsilon$$

soit $\|f - f_n\|_\infty < \varepsilon$ pour $n \geqslant N$ avec $f = f_N + (f - f_N) \in \mathrm{L}^\infty$.

2) Cas $p < +\infty$.

On ne peut pas trouver une limite pour la convergence simple presque partout aussi facilement. La raison est que contrairement à la norme infinie, la norme p ne donne pas d'information ponctuelle.
Comme f_n est de Cauchy, il existe une sous-suite $(f_{\varphi(n)})_{n\in\mathbf{N}}$ telle que :

$$\forall n \in \mathbf{N}, \qquad \|f_{\varphi(n+1)} - f_{\varphi(n)}\|_p \leqslant \frac{1}{2^n}$$

Détails : l'extractrice φ est définie par

- $\varphi(0) = 0$;
- $\forall n \geqslant 1$, $\varphi(n) = \min\{k > \varphi(n-1) \mid \forall q, r \geqslant k, \quad \|f_q - f_r\|_p \leqslant 2^{-n}\}$, le minimum étant bien défini car il s'agit d'une partie non-vide de $\mathbf{N}$ par propriété de Cauchy de (f_n).

Étape 1 : $(f_{\varphi(n)}) = \left(f_{\varphi(0)} + \sum_{k=0}^{n-1}(f_{\varphi(k+1)} - f_{\varphi(k)})\right)$ converge simplement presque partout.

Cela revient à montrer la convergence de la série télescopique presque partout. Il suffit donc d'en montrer la convergence absolue.

Soit pour $n \in \mathbf{N}^*$,

$$g_n := \sum_{k=0}^{n-1} |f_{\varphi(k+1)} - f_{\varphi(k)}| : X \to \bar{\mathbf{R}}.$$

La suite (g_n) est une suite croissante de fonctions positives qui converge donc simplement vers une fonction à valeurs dans $\bar{\mathbf{R}}$: $g_n \overset{CS}{\to} g : X \to \bar{\mathbf{R}}$. On souhaiterais prouver que g est finie presque partout. Pour cela, il est suffisant de montrer que $\|g\|_p$ est fini.

Soit $n \in \mathbf{N}^*$, par inégalité triangulaire

$$\|g_n\|_p \leqslant \sum_{k=0}^{n-1} \|f_{\varphi(k+1)} - f_{\varphi(k)}\|_p \leqslant \sum_{k=0}^{n-1} \frac{1}{2^k} < \sum_{k=0}^{+\infty} \frac{1}{2^k} = 2$$

La suite croissante et positive de fonctions mesurables (g_n) converge simplement vers g donc par théorème de convergence monotone, g est mesurable et $\|g_n\|_p \to \|g\|_p$. En particulier, $\|g\|_p \leqslant 2$ donc il existe E négligeable tel que :

$$\forall x \in X \setminus E,\ \exists f(x) \in \mathbf{C}, \qquad f_{\varphi_n}(x) \to f(x)$$

Cela termine cette étape et, comme dans le premier cas, on vient au passage de démontrer le second point du théorème en utilisant le fait qu'une suite convergente est en particulier de Cauchy.

Étape 2 : $f_n \to f$ dans $(\mathrm{L}^p, \|\cdot\|_p)$.

Remarquons qu'on peut se contenter en fait de montrer le résultat pour la sous-suite $(f_{\varphi(n)})$ car une suite de Cauchy admettant une valeur d'adhérence converge vers cette valeur.

Soit $n \in \mathbf{N}$. Par lemme de Fatou,

$$\int_X \liminf_{m\to+\infty} |f_{\varphi(m)} - f_{\varphi(n)}|^p \mathrm{d}\mu = \int_X |f - f_{\varphi(n)}|^p \mathrm{d}\mu \leqslant \liminf_{m\in\mathbf{N}} \int_X |f_{\varphi(m)} - f_{\varphi(n)}|^p \mathrm{d}\mu$$

et

$$\forall m \geqslant n, \qquad \int_X |f_{\varphi(m)} - f_{\varphi(n)}|^p \mathrm{d}\mu = \|f_{\varphi(m)} - f_{\varphi(n)}\|_p^p \leqslant \left(\sum_{k=n}^{m-1} \|f_{\varphi(k+1)} - f_{\varphi(k)}\|_p\right)^p \leqslant \frac{1}{2^{(n-1)p}}$$

donc $\|f - f_{\varphi(n)}\|_p \leqslant \frac{1}{2^{n-1}}$ soit $f_{\varphi(n)} \overset{\|\cdot\|_p}{\to} f$. Enfin, $f = f_{\varphi(1)} + (f - f_{\varphi(1)})$ avec $\|f - f_{\varphi(1)}\|_p \leqslant 1$ donc $f \in \mathrm{L}^p$. □

0.11 Couronnes biholomorphes

Leçons : 207 ; 223 ; 245 ; 267

Références :

— Salim ROSTAM, `https://minerve.ens-rennes.fr/images/Couronnes_biholomorphes.pdf`

Un théorème de RIEMANN stipule que tout ouvert non-vide, simplement connexe est soit égale à $\mathbf{C}$ soit biholomorphe au disque unité. Ce développement s'intéresse à ce qu'il se passe dans un cas particulier d'ouvert avec un trou.

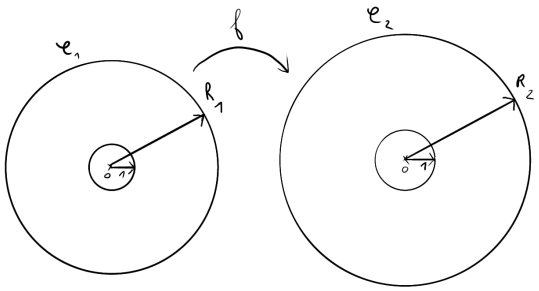
Définition 0.11.1.
Pour $0 < r < R < +\infty$, on définit la couronne $\mathcal{C}(r, R) = \{z \in \mathbf{C} \mid r < |z| < R\}$.

Proposition 0.11.2.
Si $\lambda > 0$, l'homothétie de rapport λ définit un biholomorphisme de $\mathcal{C}(r, R)$ sur $\mathcal{C}(\lambda r, \lambda R)$.

Donc deux couronnes proportionnelles sont biholomorphes. Le théorème qu'on souhaite démontrer constitue la réciproque : si deux couronnes sont biholomorphes, alors elles sont proportionnelles. Toutefois, le cas échéant, l'homothétie correspondante n'est pas le seul biholomorphisme entre les deux couronnes comme on le verra dans la preuve. Il faut ajouter les rotations composées par cette homothétie et aussi un éventuel "retournement".

Théorème 0.11.3.
Si $\mathcal{C}(r_1, R_1)$ et $\mathcal{C}(r_2, R_2)$ sont biholomorphes, alors il existe $\lambda > 0$ tel que $r_2 = \lambda r_1$ et $R_2 = \lambda R_1$.

Démonstration.
On se donne deux couronnes, notées $\mathcal{C}_1$ et $\mathcal{C}_2$ et on suppose qu'il existe un biholomorphisme $f : \mathcal{C}_1 \to \mathcal{C}_2$. Quitte à composer par des homothéties - qui sont, comme on l'a déjà remarqué plus tôt, des biholomorphismes - on peut supposer que $r_1 = r_2 = 1$. On suppose aussi $R_1 \leqslant R_2$, notre but étant alors de prouver que $R_1 = R_2$.



Il suffit de montrer que $f(z) = Cz^\alpha$ où $|C| = 1$ et $\alpha = \frac{\ln R_2}{\ln R_1}$. En effet, par injectivité de f, on aura alors $\alpha = 1$ soit $R_1 = R_2$ car sinon $1 = e^{\frac{2i\pi}{\alpha}}$ mais $f(1) = f\left(e^{\frac{2i\pi}{\alpha}}\right) = C$.

On commencera par montrer la validité de cette expression en terme de module. On note

$$u : \left| \begin{array}{ccc} \mathcal{C}_1 & \to & \mathbf{R} \\ z & \mapsto & \ln|f(z)| - \alpha \ln|z| \end{array} \right.$$

bien définie car les deux couronnes ne contiennent pas 0. Prouvons par principe du maximum que $u = 0$.

Étape 1 : Harmonicité de u sur $\mathcal{C}_1$.

Dans cette partie, on voit u comme une fonction de $\mathbf{R}^2$ dans $\mathbf{R}$. On note $\partial_z := \frac{1}{2}(\partial_x - \mathrm{i}\partial_y)$ et $\partial_{\bar{z}} := \frac{1}{2}(\partial_x + \partial_y)$. On a alors : $\partial_{\bar{z}}\partial_z = \frac{1}{4}(\partial_{xx} + \partial_{yy})$ donc il suffit de vérifier que $\partial_{\bar{z}}\partial_z u = 0$. On va utiliser les deux propriétés suivantes pour g holomorphe :

- $\partial_{\bar{z}} g = 0$ et $\partial_z g = g'$;
- $\partial_z \bar{g} = \frac{1}{2}(\partial_x \bar{g} - \mathrm{i}\partial_y \bar{g}) = \frac{1}{2}(\overline{\partial_x g} - \mathrm{i}\overline{\partial_y g}) = \overline{\partial_{\bar{z}} g} = 0$.

Remarquons que $u(z) = \frac{1}{2}(\ln|f(z)|^2 - \alpha \ln|z|^2) = \frac{1}{2}(\ln(f\bar{f}) - \alpha \ln(z\bar{z}))$. Ceci montre que u est de classe $\mathscr{C}^2$ et justifie la commodité des opérateurs ∂_z et $\partial_{\bar{z}}$. On a donc :

$$2\partial_z u = \frac{(\partial_z f)\bar{f} + f\partial_z \bar{f}}{f\bar{f}} - \alpha \frac{(\partial_z z)\bar{z} + z\partial_z \bar{z}}{z\bar{z}} = \frac{f'}{f} - \frac{\alpha}{z}$$

Puis, les fonctions du membre de droite étant holomorphes, $2\partial_{\bar{z}}\partial_z u = 0$.

Étape 2 : Prolongement de u à $\overline{\mathcal{C}_1}$.

Pour appliquer un principe du maximum, on doit prolonger u de façon continue sur $\overline{\mathcal{C}_1} = \mathcal{C}_1 \cup \mathcal{S}(1) \cup \mathcal{S}(R_1)$. Compte tenu de d'expression de u, il suffit de trouver les valeurs à attribuer à $\ln|f(z)|$ sur les cercles $\mathcal{S}(1)$ et $\mathcal{S}(R_1)$ celles de $\alpha \ln|z|$ y sont évidentes. On s'intéresse donc au comportement de $|f(z)|$ lorsque $|z| \to 1$ et $|z| \to R_1$.
Soit $a \in \mathcal{S}(1)$ et $(z_n)_{n\in\mathbf{N}}$ telle que $z_n \to a$.

1) $\mathrm{Va}(|f(z_n)|) \subset \{1, R_2\}$.

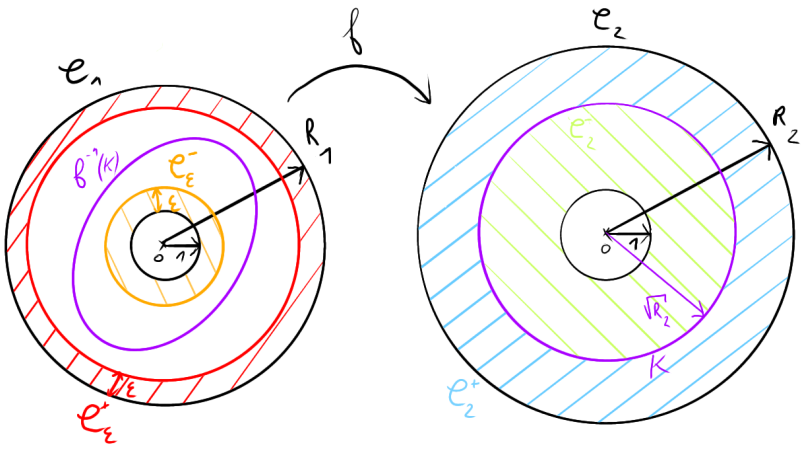
La notation Va désigne l'ensemble des valeurs d'adhérence. Puisque $(f(z_n))$ est dans le compact $\overline{\mathcal{C}_2}$, elle admet une valeur d'adhérence $l \in \overline{\mathcal{C}_2}$: $f(z_{\varphi(n)}) \to l$. Si $l \in \mathcal{C}_2$, f étant bijective il existe $z \in \mathcal{C}_1$ tel que $f^{-1}(l) = z$. Mais f^{-1} étant continue car holomorphe :

$$f^{-1}(f(z_{\varphi(n)})) = z_{\varphi(n)} \to z, \text{ or } |z_{\varphi(n)}| \to 1 \text{ et } |z| > 1$$

ce qui est impossible donc $l \in \overline{\mathcal{C}_2} \setminus \mathcal{C}_2 = \mathcal{S}(1) \cup \mathcal{S}(R_n)$.

2) $\sharp\mathrm{Va}(|f(z_n)|) = 1$.

Montrons par des arguments topologiques que $(|f(z_n)|)$ ne peut admettre les deux valeurs d'adhérence. L'idée derrière est la continuité de f qui ne peut "sauter" d'un voisinage de $\mathcal{S}(1)$ à un voisinage de $\mathcal{S}(R_2)$.
Soit $K := \mathcal{S}(\sqrt{R_2})$ compact, qui jouera le rôle de séparateur entre les cercles de rayons 1 et R_2, alors $\mathcal{C}_2 \setminus K$ a deux composantes connexes $\mathcal{C}_2^- := \mathcal{C}(1, \sqrt{R_2})$ et $\mathcal{C}_2^+ := \mathcal{C}(\sqrt{R_2}, R_2)$. La fonction f étant en particulier un homéomorphisme, $f^{-1}(K)$ est compact et $\mathcal{C}_1 \setminus f^{-1}(K)$ a deux composantes connexes.
Soit $\varepsilon > 0$ tel que $\mathrm{d}(f^{-1}(K), \partial\mathcal{C}_1) > \varepsilon$ dont l'existence est assuré par le fait qu'une distance entre deux compacts est strictement positive. On note $\mathcal{C}_\varepsilon^- :=_C C(1, 1+\varepsilon)$ et $\mathcal{C}_\varepsilon^+ := \mathcal{C}(R_1 - \varepsilon, R_1)$. À partir d'un certain rang, (z_n) est dans le connexe $\mathcal{C}_\varepsilon^-$ dont l'image est donc contenue soit dans $\mathcal{C}_2^-$ soit dans $\mathcal{C}_2^+$.



Ceci montre que $(|f(z_n)|)$ ne peut admettre simultanément les deux valeurs d'adhérence 1 et R_2 car la distance entre, par exemple, $\mathcal{C}_2^-$ et $\mathcal{S}(R_2)$, est strictement positive.
Quitte à considérer $\frac{R_2}{f}$ qui est bien biholomorphe, on peut supposer que $f(\mathcal{C}_\varepsilon^-) \subset \mathcal{C}_2^-$ et donc $\mathrm{Va}(|f(z_n)|) = \{1\}$.

3) u se prolonge par continuité par 0 sur $\mathcal{S}(1)$.

Quels que soit $a \in \mathcal{S}(1)$, toute suite (z_n) convergeant vers a donnera d'après les deux premiers points une suite $(|f(z_n)|)$ ayant une unique valeur d'adhérence parmi 1 et R_2. Mais puisque (z_n) sera contenue à partir d'un certain rang dans la couronne $\mathcal{C}_\varepsilon^-$ et donc, par le choix que l'on a fait, envoyée par f dans $\mathcal{C}_2^-$, cette unique valeur d'adhérence dans ne peut être que 1.
Ainsi, si $a \in \mathcal{S}(1)$, et si $(z_n) \in \mathcal{C}_1^{\mathbf{N}}$ vérifie $z_n \to a$, alors $\mathrm{Va}(|f(z_n)|) = \{1\}$. Or $(|f(z_n)|)$ est à valeurs dans $[1, R_2]$ qui est compact et une suite ayant une unique valeur d'adhérence dans un compact converge vers cette valeur donc $|f(z_n)| \to 1$. On vérifie donc bien que u est prolongeable par continuité sur $\mathcal{S}(1)$ avec la valeur 0.

4) u se prolonge par continuité par 0 sur $\mathcal{S}(R_1)$.

Le même raisonnement s'applique pour le prolongement sur $\mathcal{S}(R_1)$, l'image d'une suite convergeant vers un point du bord va converger en module soit vers 1, soit vers R_2. Il reste à voir que le choix qu'on a fait force en fait cette

valeur à être R_2. Il suffit pour cela de montrer que la couronne $\mathcal{C}_\varepsilon^+$ est envoyée dans $\mathcal{C}_2^+$. Supposons par l'absurde que $f(\mathcal{C}_\varepsilon^+) \subset \mathcal{C}_2^-$. Alors,

$$\mathcal{C}_2 = f(\mathcal{C}_1) = \underbrace{f(\mathcal{C}_\varepsilon^-) \cup f(\mathcal{C}_\varepsilon^+)}_{\subset \mathcal{C}_2^-} \cup \underbrace{f(\overline{\mathcal{C}(1+\varepsilon, R_1 - \varepsilon)})}_{\text{compact}}$$

et $\mathrm{d}(\mathcal{C}_2, \mathcal{S}(R_2)) > 0$ ce qui est absurde, donc $f(\mathcal{C}_\varepsilon^+) \subset \mathcal{C}_2^+$ et u se prolonge par continuité sur $\mathcal{S}(R_1)$ par

$$u(z) = \ln|R_2| - \alpha \ln|R_1| = 0.$$

Étape 3 : Conclusion.

Par principe du maximum, u atteint son maximum sur $\partial\mathcal{C}_1$ mais $u_{|\partial\mathcal{C}_1} = 0$ donc $u = 0$.
En particulier, $\partial_z u = 0$, soit par un calcul déjà fait

$$\frac{f'}{f} = \frac{\alpha}{z} \tag{9}$$

Et en intégrant cette relation sur $\mathcal{S}(\sqrt{R_1})$ paramétré sur $[0,1]$ par $\gamma(t) = \sqrt{R_1} e^{2\mathrm{i}\pi t}$, on obtient :

$$\int_\gamma \frac{f'(z)}{f(z)} \mathrm{d}z = \int_0^1 \frac{f'(\gamma(t))}{f(\gamma(t))} \gamma'(t) \mathrm{d}t = \int_{f\circ\gamma} \frac{\mathrm{d}z}{z} = \alpha \int_\gamma \frac{\mathrm{d}z}{z}$$

soit en divisant par $2\mathrm{i}\pi$

$$\mathrm{ind}_{f\circ\gamma}(0) = \alpha \,\mathrm{ind}_\gamma(0) = \alpha$$

et donc $\alpha \in \mathbf{Z} \cap [1, +\infty[= \mathbf{N}^*$. Et maintenant, on sait résoudre 9, par exemple en remarquant que :

$$\forall z \in \mathcal{C}_1, \qquad \left(f(z) z^{-\alpha}\right)' = f'(z) z^{-\alpha} - \alpha f(z) z^{-\alpha-1} = 0$$

et $\mathcal{C}_1$ étant connexe, il existe $C \in \mathbf{C}^*$ telle que $\forall z \in \mathcal{C}_1,\ f(z) = C z^\alpha$ ce qui conclut.

On peut remarquer qu'on aboutit à $f(z) = Cz$ avec nécessairement $|C| = 1$ ($|f(1+1/n)| \to |C| \geqslant 1$ et $|f(R_1 - 1/n)| \to |C_1| R_1 \leqslant R_1$). On a aussi pu composer par des dilatations et par le retournement $z \mapsto R_2/z$. On obtient ainsi tout les automorphismes possibles entre deux couronnes. □

0.12 Théorème central limite

Leçons : 261 ; 262

Références :

— [App13] Walter Appel, *Probabilités pour les non-probabilistes*

Lemme 0.12.1.
Si $z \in \mathbf{C}$ et $z_n \to z$,

$$\lim_{n\to+\infty} \left(1+\frac{z_n}{n}\right)^n = e^z$$

Démonstration.
La propriété est bien connue lorsque $z = x$ et $z_n = x_n$ sont réels et se démontre par un simple développement limité : si $x \in \mathbf{R}$ et $(x_n) \in \mathbf{R}^{\mathbf{N}}$ vérifient $x_n \to x$:

$$\left(1+\frac{x_n}{n}\right)^n = \exp\left[n\ln\left(1+\frac{x_n}{n}\right)\right] = \exp\left[n\left(\frac{x_n}{n}+o\left(\frac{1}{n}\right)\right)\right] = \exp[x_n+o(1)] \to e^x.$$

Pour k fixé et $n \geqslant k$,

$$a_{n,k} := \binom{n}{k}\frac{1}{n^k} = \frac{n(n-1)\dots(n-k+1)}{k!n^k} \underset{n\to+\infty}{\nearrow} \frac{1}{k!}.$$

Or, pour tout $n \in \mathbf{N}$,

$$\begin{aligned}
\left|e^{z_n} - \left(1+\frac{z_n}{n}\right)^n\right| &= \left|\sum_{k=0}^{+\infty}\frac{z_n^k}{k!} - \sum_{k=0}^{n} a_{n,k}z_n^k\right| \\
&\leqslant \sum_{k=0}^{n}\left|\frac{1}{k!}-a_{n,k}\right||z_n^k| + \sum_{k=n+1}^{+\infty}\frac{|z_n|^k}{k!} \\
&= \sum_{k=0}^{n}\left(\frac{1}{k!}-a_{n,k}\right)|z_n^k| + \sum_{k=n+1}^{+\infty}\frac{|z_n|^k}{k!} \\
&= e^{|z_n|} - \left(1+\frac{|z_n|}{n}\right)^n \\
&\underset{n\to+\infty}{\longrightarrow} 0
\end{aligned}$$

d'après la propriété établie précédemment sur $\mathbf{R}$. D'où finalement

$$\left|e^z - \left(1+\frac{z_n}{n}\right)^n\right| = |e^z - e^{z_n}| + \left|e^{z_n} - \left(1+\frac{z_n}{n}\right)^n\right| \to 0$$

□

Lemme 0.12.2.
Soit $Z \sim \mathcal{N}(0;1)$. Alors on a

$$\forall t \in \mathbf{R}, \qquad \varphi_Z(t) = e^{-\frac{t^2}{2}}$$

Démonstration.
Soit $t \in \mathbf{R}$. On doit calculer :

$$\varphi_Z(t) = \mathbb{E}\left[e^{\mathrm{i}tZ}\right] = \frac{1}{\sqrt{2\pi}}\int_{\mathbf{R}} e^{\mathrm{i}tx}e^{-\frac{x^2}{2}}\mathrm{d}x$$

Heuristique : on sait intégrer les expressions du type $e^{\alpha x}$, $\alpha \in \mathbf{C}$ ou $e^{\beta x^2}$, $\beta \in \mathbf{R}$ mais pas les deux ensemble. En revanche, si on remplace $it \in \mathrm{i}\mathbf{R}$ par $u \in \mathbf{R}$, on obtient :

$$\frac{1}{\sqrt{2\pi}}\int_{\mathbf{R}} e^{ux}e^{-\frac{x^2}{2}}\mathrm{d}x = \frac{1}{\sqrt{2\pi}}\int_{\mathbf{R}} e^{-\frac{(x-u)^2}{2}}e^{\frac{u^2}{2}}\mathrm{d}x = \frac{1}{\sqrt{2\pi}}e^{\frac{u^2}{2}}\int_{\mathbf{R}} e^{-\frac{x^2}{2}}\mathrm{d}x = e^{\frac{u^2}{2}}$$

par invariance par translation et en reconnaissant l'intégrale sur $\mathbf{R}$ de la densité d'une gaussienne. Remplaçant u par it, on a

$$\frac{1}{\sqrt{2\pi}}\int_{\mathbf{R}} e^{\mathrm{i}tx}e^{-\frac{x^2}{2}}\mathrm{d}x = e^{-\frac{t^2}{2}}.$$

Évidemment il s'agit d'une heuristique, mais nous allons utiliser le théorème de prolongement analytique pour rendre tout cela rigoureux.

On pose

$$f : \left| \begin{array}{ccc} \mathbf{C} \times \mathbf{R} & \to & \mathbf{C} \\ (z,x) & \mapsto & e^{zx} e^{-\frac{x^2}{2}} \end{array} \right.$$

- Pour $x \in \mathbf{R}$, $f(\cdot, x) \in H(\mathbf{C})$.
- Si $R > 0$, alors pour tout $(z,x) \in \bar{D}(0,R) \times \mathbf{R}$:

$$|f(z,x)| \leqslant e^{R|x|} e^{-\frac{x^2}{2}} \in \mathrm{L}^1(\mathbf{R}).$$

Ainsi, par théorème d'holomorphie sous l'intégrale, la fonction

$$F : \left| \begin{array}{ccc} \mathbf{C} & \to & \mathbf{C} \\ z & \mapsto & \frac{1}{\sqrt{2\pi}} \int_{\mathbf{R}} e^{zx} e^{-\frac{x^2}{2}} \mathrm{d}x \end{array} \right.$$

est bien définie et holomorphe sur tout disque fermé donc entière. Or le calcul de l'heuristique montre que F coïncide avec la fonction entière $z \mapsto e^{\frac{z^2}{2}}$ sur $\mathbf{R}$ qui est d'accumulation dans $\mathbf{C}$ donc par théorème de prolongement analytique :

$$\forall z \in \mathbf{C}, \qquad F(z) = e^{\frac{z^2}{2}}$$

En particulier, $\varphi_Z(t) = F(\mathrm{i}t) = e^{-\frac{t^2}{2}}$.

□

Théorème 0.12.3 (limite central)**.**
Soit $(X_n)_{n \geqslant 1}$ une suite de variables aléatoires, indépendantes, de même loi et admettant une variance $\sigma^2 > 0$. Notons $m = \mathbb{E}[X_1]$ leur espérance. Posons enfin $S_n = X_1 + \cdots + X_n$. Alors,

$$S_n^* := \frac{S_n - nm}{\sigma\sqrt{n}} \overset{\mathrm{L}}{\Longrightarrow} \mathcal{N}(0;1).$$

Démonstration.
Notons φ la fonction caractéristique de $X_1 - m$ qui est commune à toutes les variables centrées $X_n - m$. Par indépendance des X_i,

$$\varphi_n(t) := \mathbb{E}\left[\exp\left(\mathrm{i}t \frac{1}{\sigma\sqrt{n}} \sum_{k=1}^{n} (X_k - m)\right)\right] = \prod_{k=1}^{n} \mathbb{E}\left[\exp\left(\mathrm{i}t \frac{X_k - m}{\sigma\sqrt{n}}\right)\right] = \varphi\left(\frac{t}{\sigma\sqrt{n}}\right)^n.$$

On cherche maintenant la limite simple de (φ_n). La fonction caractéristique φ est $\mathscr{C}^2$ car X_1 est L^2, d'où par formule de Taylor-Young :

$$\varphi(t) = \varphi(0) + \varphi'(0)t + \varphi''(0)\frac{t^2}{2} + \underset{t \to 0}{o}(t^2)$$

avec $\varphi(0) = \mathbb{E}\left[e^0\right] = 1$, $\varphi'(0) = \mathrm{i}\mathbb{E}[X_n - m] = 0$ et $\varphi''(0) = \mathrm{i}^2 \mathbb{E}\left[(X_n - m)^2\right] = -\sigma^2$, donc :

$$\varphi(t) = 1 - \frac{\sigma^2 t^2}{2} + \underset{t \to 0}{o}(t^2).$$

Ainsi,

$$\varphi_n(t) = \left(1 - \frac{t^2}{2n} + \underset{t \to 0}{o}\left(\frac{1}{n}\right)\right)^n = e^{-\frac{t^2}{2}}$$

grâce au lemme 1. On reconnaît à droite la fonction caractéristique de la loi normale centrée réduite. Par théorème de Lévy, on obtient $S_n^* \overset{\mathrm{L}}{\Longrightarrow} \mathcal{N}(0;1)$. □

0.13 Inégalité de Hoeffding

Leçons : 261 ; 262 ; 266

Références :

— Adrien Laurent, `https://agreg-maths.fr/uploads/versions/874/Inegalite_Hoeffding.pdf`

— Benjamin Havret, `http://www.normalesup.org/~havret/pdf/developpements_maths%20bhavret.pdf`

Lemme 0.13.1.
Soit X une variable aléatoire réelle centrée telle que $|X| \leqslant 1$ p.s., alors

$$\forall t \in \mathbf{R}, \mathbb{E}\left[e^{tX}\right] \leqslant e^{\frac{t^2}{2}}.$$

Démonstration.
Soient $x \in [-1,1]$ et $t \in \mathbf{R}$, utilisant la convexité de l'exponentielle :

$$e^{tx} = \exp\left(\frac{1-x}{2}(-t) + \frac{1+x}{2}t\right) \leqslant \frac{1-x}{2}e^{-t} + \frac{1+x}{2}e^{t}.$$

où les poids $\frac{1-x}{2}$ et $\frac{1+x}{2}$ sont bien dans $[0,1]$ et de somme 1. Puisque $|X| \leqslant 1$ *p.s.*, On peut composer l'inégalité précédente par X pour avoir :

$$e^{tX} \leqslant \frac{1-X}{2}e^{-t} + \frac{1+X}{2}e^{t} \text{ p.s.}$$

et prenant l'espérance avec X centrée

$$\mathbb{E}\left[e^{tX}\right] \leqslant \frac{1}{2}(e^{-t}+e^{t}) = \cosh t = \sum_{n=0}^{+\infty}\frac{t^{2n}}{(2n)!} \leqslant \sum_{n=0}^{+\infty}\frac{t^{2n}}{n!\,2^n} = e^{\frac{t^2}{2}}$$

car si $n \in \mathbf{N}^*$, $\frac{(2n)!}{n!} = (2n)(2n-1)\dots(n+1) \geqslant 2^n$. □

Théorème 0.13.2 (de Hoeffding)**.**
Soit $(X_n)_{n\in\mathbf{N}^*}$ une suite de variables aléatoires réelles indépendantes et centrées. On suppose de plus $|X_n| \leqslant c_n$ p.s. avec $c_n > 0$. Alors, notant $S_n = \sum_{i=1}^n X_i$ et $s_n = \sum_{i=1}^n c_i^2$, pour tout $n \in \mathbf{N}^*$ et tout $\varepsilon > 0$,

$$\mathbb{P}\left\{|S_n| > \varepsilon\right\} \leqslant 2\exp\left(-\frac{\varepsilon^2}{2s_n}\right)$$

Démonstration.
Soient $t > 0$ (il s'agit d'un paramètre que l'on choisira à la fin de sorte à optimiser l'inégalité) et $\varepsilon > 0$. On a $\{S_n > \varepsilon\} = \left\{e^{tS_n} > e^{t\varepsilon}\right\}$ donc ces deux évènements sont de même probabilité, mais la variable e^{tS_n} a le bon goût, contrairement à S_n, d'être positive, d'où par l'inégalité de Markov :

$$\mathbb{P}\left\{S_n > \varepsilon\right\} \leqslant \frac{\mathbb{E}\left[e^{tS_n}\right]}{e^{t\varepsilon}}.$$

Or, par indépendance des X_i,

$$\mathbb{E}\left[e^{tS_n}\right] = \prod_{i=1}^n \mathbb{E}\left[e^{tX_i}\right] = \prod_{i=1}^n \mathbb{E}\left[e^{c_i t\frac{X_i}{c_i}}\right] \leqslant \prod_{i=1}^n e^{\frac{c_i^2 t^2}{2}} \leqslant \exp\left(\frac{t^2}{2}\sum_{i=1}^n c_i^2\right)$$

par le lemme avec X_i/c_i bornée par 1 presque sûrement. Finalement,

$$\mathbb{P}\left\{S_n > \varepsilon\right\} \leqslant \exp\left(\frac{t^2 s_n}{2} - t\varepsilon\right)$$

On va maintenant choisir t de sorte à minimiser le terme de droite, c'est-à-dire minimiser l'expression quadratique $\frac{t^2 s_n}{2} - t\varepsilon$ qui dessine donc une parabole orientée vers le haut, de racines 0 et $\frac{2\varepsilon}{s_n}$. Son minimum est atteint au milieu de ces deux racines, soit en $t_0 = \frac{\varepsilon}{s_n}$ et vaut $-\frac{\varepsilon^2}{2s_n}$. On a alors,

$$\mathbb{P}\left\{S_n > \varepsilon\right\} \leqslant \exp\left(-\frac{\varepsilon^2}{2s_n}\right)$$

Finalement, on fait le même raisonnement avec $(-X_i)_{i\in\mathbf{N}^*}$ pour avoir

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

d'où

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

□

Remarques :

- On peut comparer avec l'inégalité de Bienaymé-Tchebychev pour une somme de lois de Bernoulli ($c_n = 1$) :
  $\mathbb{P}\{|S_n - np| > \varepsilon\} \leqslant \frac{np(1-p)}{n\varepsilon^2} = \frac{p(1-p)}{\varepsilon^2}$ avec Bienaymé-Tchebychev ;
  $\mathbb{P}\{|S_n - np| > \varepsilon\} \leqslant 2\exp\left(-\frac{\varepsilon^2}{2}\right)$ avec Hoeffding.
  On voit l'efficacité de Hoeffding...
- Par ailleurs, l'inégalité de Bienaymé-Tchebychev permet de démontrer une loi faible des grands nombres (convergence en probabilité) en majorant $\mathbb{P}\{|S_n - \mathbb{E}[S_n]| > 1/n\}$ par un terme de limite nulle. Dans le cas de v.a. bornées mais **pas nécessairement de même loi**, Hoeffding permet, non seulement de faire comme avec Bienaymé-Tchebychev, mais la majoration est même sommable sur n, ce qui donne via le théorème de Borel-Cantelli une convergence presque sûre. On obtient le résultat suivant, plus précis qu'une loi forte des grands nombres pour des v.a. bornées :

Théorème 0.13.3 (Loi forte des grands nombres pour des v.a. bornées)**.**
*Si $(X_n)_{n\in\mathbf{N}^*}$ est une suite suite de v.a. indépendantes, centrées **mais pas nécessairement de mêmes lois** telles que $a \leqslant X_i \leqslant b$ p.s., alors presque sûrement :*

$$\limsup_{n\in\mathbf{N}^*} \frac{1}{\sqrt{n\ln n}}\left|\sum_{i=1}^n X_i\right| \leqslant \frac{b-a}{\sqrt{2}}$$

Démonstration.
Il s'agit d'utiliser le théorème de Borel-Cantelli. Si $t > 0$, l'inégalité de Hoeffding donne :

$$\mathbb{P}\left\{\left|\sum_{i=1}^n X_i\right| > t\right\} \leqslant 2\exp\left(-\frac{2t^2}{n(b-a)^2}\right).$$

Soit $\varepsilon > 0$, on évalue en $t = t_n = \sqrt{n\ln n}\left(\frac{b-a}{\sqrt{2}} + \varepsilon\right)$:

$$\mathbb{P}\left\{\left|\sum_{i=1}^n X_i\right| > t_n\right\} \leqslant 2\exp\left(-\ln n\left(1 + \frac{\sqrt{2}\varepsilon}{(b-a)}\right)^2\right) = n^{-\left(1+\frac{\sqrt{2}\varepsilon}{b-a}\right)^2}$$

qui est le terme général d'une série sommable donc le lemme de Borel-Cantelli assure que :

$$\mathbb{P}\left\{\limsup_{n\in\mathbf{N}^*}\left\{\left|\sum_{i=1}^n X_i\right| > t_n\right\}\right\} = 0$$

donc presque sûrement,

$$\limsup_{n\in\mathbf{N}^*} \frac{1}{\sqrt{n\ln n}}\left|\sum_{i=1}^n X_i\right| < \frac{b-a}{\sqrt{2}} + \varepsilon$$

et ce, pour tout $\varepsilon > 0$ donc, par continuité décroissante, presque sûrement

$$\limsup_{n\in\mathbf{N}^*} \frac{1}{\sqrt{n\ln n}}\left|\sum_{i=1}^n X_i\right| < \frac{b-a}{\sqrt{2}}$$

□

Enfin, l'inégalité de Hoeffding admet une généralisation en terme de martingales avec une démonstration assez similaire. En voici un énoncé :

Théorème 0.13.4 (Inégalité d'Azuma)**.**
Soit $(X_n)_n$ une martingale issue de 0 dont les accroissements sont contrôlés par une suite déterministe (c_n), i.e. $|X_n - X_{n-1}| \leqslant c_n$ p.s. pour tout $n \geqslant 1$. Alors, pour tout $\varepsilon > 0$, on a

$$\mathbb{P}\{|X_n| \geqslant \varepsilon\} \leqslant 2e^{-\frac{\varepsilon^2}{2\sigma_n^2}}$$

où $\sigma_n^2 = \sum_{i=1}^n c_i^2$.

0.14 Processus de Galton-Watson

Leçons : 226 ; 264

Références :

— [App13] Walter Appel, *Probabilités pour les non-probabilistes*

— [BK06] Michel Benaïm, Nicole El Karoui, *Promenade aléatoire*

— Salim Rostam `https://minerve.ens-rennes.fr/images/Galton-Watson.pdf`

Théorème 0.14.1 (Processus de Galton-Watson)**.**
Soit X une variable aléatoire à valeurs dans $\mathbf{N}$ admettant une variance ; on note $p_n := \mathbb{P}\{X = n\}$ et $m := \mathbb{E}[X]$. On suppose que $\mathbb{P}\{X = 1\} < 1$. Soit $(X_{i,j})_{i,j\in\mathbf{N}^*}$ une famille de variables aléatoires indépendantes et de même loi que X. On considère alors $(Z_n)_{n\in\mathbf{N}}$ définie par :

$$Z_0 = 1$$
$$\forall n \in \mathbf{N}, \quad Z_{n+1} = \sum_{i=1}^{Z_n} X_{i,n}.$$

et la probabilité $P_{ext} := \mathbb{P}\{\exists n \in \mathbf{N},\ Z_n = 0\}$. Alors,

— *si $m > 1$, $P_{ext} < 1$;*

— *si $m \leqslant 1$, $P_{ext} = 1$.*

Démonstration.
L'idée est de modéliser avec (Z_n) la taille d'une population : à l'instant n, il y a Z_n individus et chaque individu i de la n-ième génération a un nombre $X_{i,n}$ de descendants. Ce développement étudie la suite (Z_n), en particulier s'il existe un n tel que $Z_n = 0$, *i.e.* la population s'éteint. Remarquons une première chose qui va nous servir dans la suite.

Lemme 0.14.2.
Pour $n \in \mathbf{N}^*$, la variable Z_n est indépendante de $X_{i,j}$ pour tout $i \in \mathbf{N}$ et pour tout $j \geqslant n$.

Démonstration.
En effet, on peut remarquer que Z_n ne dépend que de Z_{n-1} et des $(X_{i,n-1})_i$ donc par récurrence, Z_n ne dépend que des $(X_{i,j})_{i\in\mathbf{N},j<n}$ et l'on conclut par le caractère i.i.d. de la suite des $X_{i,j}$. □

Étape 1 : Réécriture.

On remarque maintenant que si $Z_n = 0$, alors $Z_{n+1} = 0$, autrement dit la suite $(\{Z_n = 0\})_n$ est croissante et :

$$P_{ext} = \mathbb{P}\left\{\bigcup_{n\in\mathbf{N}} \{Z_n = 0\}\right\} = \lim_{n\to+\infty} \uparrow \underbrace{\mathbb{P}\{Z_n = 0\}}_{x_n}$$

On aimerait donc calculer les $\mathbb{P}\{Z_n = 0\}$ pour en déduire P_{ext}. Pour cela, on va déterminer la loi de Z_n à travers sa fonction génératrice. Notons G la fonction génératrice de X et G_n celle de Z_n.

Étape 2 : $G_n = \overbrace{G \circ \cdots \circ G}^{n \text{ fois}}$.

Si $t \in [-1, 1]$ et $n \in \mathbf{N}$,

$$\begin{aligned}
G_{n+1}(t) &= \mathbb{E}\left[t^{Z_{n+1}}\right] \\
&= \mathbb{E}\left[\mathbb{E}\left[t^{Z_{n+1}}|Z_n\right]\right] \\
&= \sum_{k=0}^{+\infty} \mathbb{P}\{Z_n = k\}\, \mathbb{E}\left[t^{Z_{n+1}}|Z_n = k\right] \\
&= \sum_{k=0}^{+\infty} \mathbb{P}\{Z_n = k\}\, \mathbb{E}\left[t^{X_{1,n}+X_{2,n}+\cdots+X_{k,n}}\right] \\
&= \sum_{k=0}^{+\infty} \mathbb{P}\{Z_n = k\}\, \mathbb{E}\left[t^{X_{1,n}}\right] \cdot \mathbb{E}\left[t^{X_{2,n}}\right] \ldots \mathbb{E}\left[t^{X_{k,n}}\right] \\
&= G_n(G(t))
\end{aligned}$$

Ainsi, comme G_0=1, on en déduit le résultat par récurrence, et en particulier $\forall n \in \mathbf{N},\ G_{n+1} = G \circ G_n$ d'où :

$$x_{n+1} = G_{n+1}(0) = G(G_n(0)) = G(x_n)$$

et la suite qui nous intéresse est donc la suite des itérées de G issue de 0.

Étape 3 : Convergence.

G est continue sur $[0,1] \ni P_{ext}$ donc $G(P_{ext}) = P_{ext}$. On va mettre montrer plus précisément que P_{ext} est le plus petit point fixe de G.
Soit $l \in [0,1]$ tel que $G(l) = l$. Comme $Z_0 = 1$, $x_0 = 0 \leqslant l$. Une fonction génératrice étant croissante sur $[0,1]$, on peut composer par G dans l'inégalité précédente et utiliser le fait que l est un point fixe pour avoir :

$$x_1 \leqslant g(l) = l$$

et par récurrence

$$\forall n \in \mathbf{N}, \quad x_n \leqslant l.$$

Passant alors à la limite :

$$P_{ext} \leqslant l.$$

P_{ext} est donc le plus petit point fixe de G sur $[0,1]$.

Pour démontrer le théorème, on doit donc s'intéresser au plus petit point d'intersection de la courbe représentative de G avec la courbe $y = x$. On utilisera notamment qu'étant une fonction génératrice :

- G est convexe sur $[0,1]$;
- $G(1) = 1$;
- $G'(1) = m$.

1) Cas $m > 1$. [2]

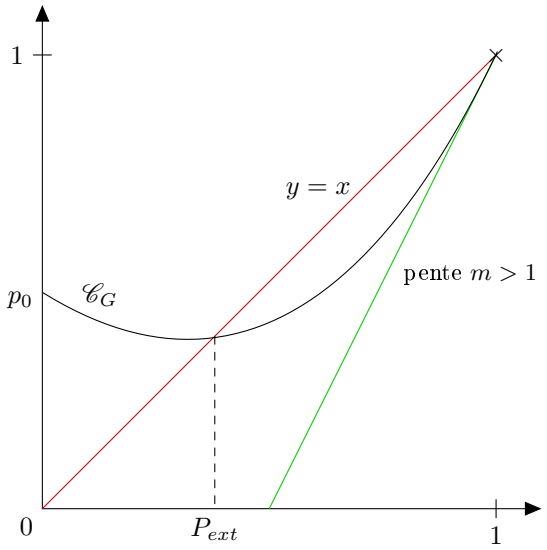


FIGURE 1 – Cas $m > 1$

La courbe se confondant localement avec la tangente, $G - \mathrm{id} < 0$ au voisinage de 1^-, or $G(0) - 0 \geqslant 0$ donc soit $G(0) = p_0 = 0 = P_{ext}$, soit $G(0) > 0$ et par théorème des valeurs intermédiaires $G - \mathrm{id}$ s'annule sur $]0,1[$ et $0 < P_{ext} < 1$.

2) Cas $m < 1$.

Par convexité, G est au-dessus de sa tangente en 1, elle-même au dessus de la droite $y = x$, ne se touchant qu'en 1 donc le seul point fixe possible de G est 1 et $P_{ext} = 1$.

2. Très beaux schémas de Salim Rostam !

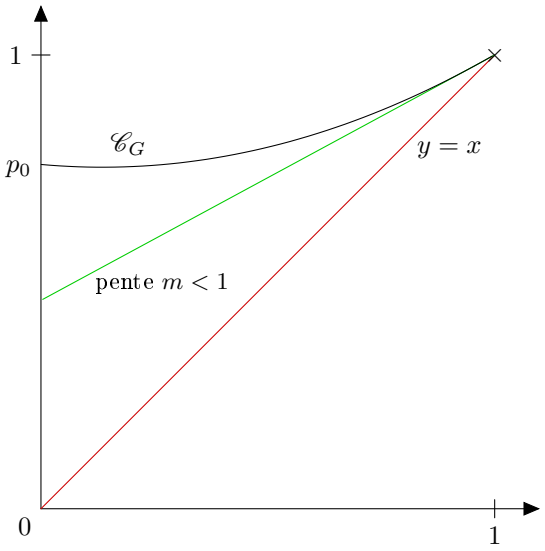


FIGURE 2 – Cas $m < 1$

3) Cas $m = 1$.

Cette fois la tangente en 1 et la droite $y = x$ coïncident. On a donc besoin de connaitre plus précisément la position de G par rapport à sa tangente en 1. X admet une variance donc $G \in \mathscr{C}^2([0,1])$ et si $t \in]0,1]$,

$$G''(t) = \sum_{n=2}^{+\infty} n(n-1)p_n t^{n-2} > 0$$

si et seulement si $p_0 + p_1 < 1$.
Mais si $p_0 + p_1 = 1$, $X \sim \mathscr{B}(m)$ mais $m = 1$ et $X = 1$ *p.s.* ce qui est exclu. Cela correspond au cas inintéressant où tout individu à presque sûrement un unique descendant et donc évidemment $P_{ext} = 0$.

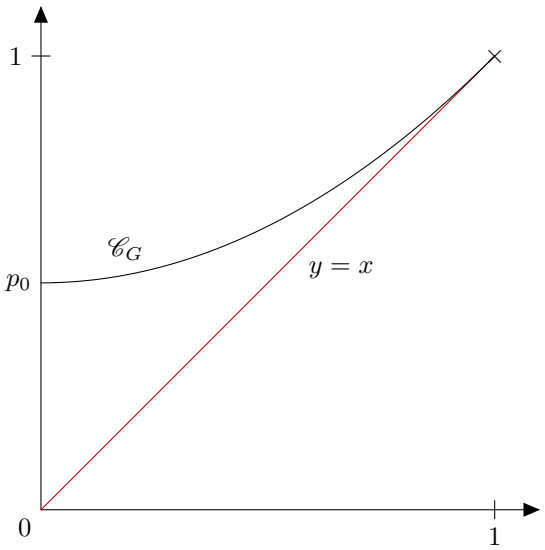


FIGURE 3 – Cas $m = 1$ avec $p_0 + p_1 < 1$

Donc $p_0 + p_1 < 1$ et $G'' > 0$ sur $]0,1]$ d'où G strictement convexe sur $]0,1]$. La courbe représentative de G n'intersecte sa tangente en 1, la droite $y = x$, qu'en 1. Ainsi 1 est l'unique point fixe de G sur $]0,1]$. Enfin $p_0 \neq 0$ car

$$m = \sum_{n=1}^{+\infty} np_n = 1$$

implique $p_1 = 1$, exclu à nouveau. Donc $P_{ext} = 1$. □

Remarques :

- Le seul cas où la probabilité d'extinction est nulle est lorsque $p_0 = 0$; s'il y a une chance, aussi petite soit-elle qu'un individu n'ait aucune descendance, alors la probabilité d'extinction est non nulle.
- On a $\mathbb{E}[Z_{n+1}] = G'_{n+1}(1) = (G \circ G_n)'(1) = G'(1)G_n(1) = mG_n(1)$ et par récurrence :

$$\forall n \in \mathbf{N}, \qquad \mathbb{E}[Z_n] = m^n$$

  Ainsi
  - Dans le cas sous-critique $m < 1$, la population moyenne tend géométriquement vers 0.
  - Dans le cas critique $m = 1$, la population moyenne reste constante.
  - Dans le cas sur-critique $m > 1$, la population moyenne tend géométriquement vers $+\infty$.
- On peut montrer que $\mathbb{E}[Z_{n+1}|Z_n] = mZ_n$ donc Z_n/m^n est une martingale, et par des théorèmes de martingales, elle converge presque sûrement.

0.15 Espace des translatés de dimension finie

Leçons : ; 151 ; 154 ; 159 ; 162 ; 201 ; 221 ; 228

Théorème 0.15.1.
Soit $f : \mathbf{R} \to \mathbf{R}$ dérivable. Il y a équivalence entre :

(i) $F := \mathrm{Vect}(f_a : x \mapsto f(x+a)\ ;\ a \in \mathbf{R})$ *est de dimension finie ;*

(ii) f *est solution d'une équation différentielle linéaire homogène à coefficients constants.*

Pour démontrer ce résultat, on aura besoin de ce lemme :

Lemme 0.15.2.
Soit $f_1, \dots, f_n : \mathbf{R} \to \mathbf{R}$. Il y a équivalence entre :

(1) $(f_1, \dots, f_n)$ *est libre ;*

(2) il existe $\bar{x} = (x_1, \dots, x_n) \in \mathbf{R}^n$ *tels que* $M_{\bar{x}} := (f_i(x_j))_{1 \leqslant i,j \leqslant n}$ *soit inversible.*

C'est sous cette forme qu'on l'utilisera pour démontrer le théorème. Mais écrit comme ceci, sa signification ne saute pas aux yeux... Pour tester la liberté de la famille $(f_1, \dots, f_n)$, il faudrait, pour toute combinaison linéaire non triviale, vérifier qu'elle ne s'annule pas sur tous les éléments de $\mathbf{R}$. Autrement dit, a priori la liberté des f_i devrait dépendre de leur comportement sur chaque réel. Ce théorème montre qu'en fait il existe un n-uplet $\bar{x}$ de $\mathbf{R}$ dépendant uniquement des f_i qui contient l'information sur leurs éventuelles relations. Il suffit donc de vérifier la liberté de la famille des lignes de la matrice $M_{\bar{x}}$.

Démonstration du lemme.
(2) $\implies$ (1) Simple par contraposée car une relation entre les f_i définit aussi une relation entre les lignes de la matrice $M_{\bar{x}}$ quel que soit $\bar{x} \in \mathbf{R}$

(1) $\implies$ (2) Supposons $(f_1, \dots, f_n)$ libre et notons $F := \mathrm{Vect}(f_1, \dots, f_n)$ qui est de dimension n.

Étape 1 : $F^* = \mathrm{Vect}(\mathrm{ev}_a : f \mapsto f(a)\ ;\ a \in \mathbf{R})$.

Ce résultat est naturel : l'information véhiculée par une forme linéaire sur F est contenue dans ses valeurs en chaque élément. Montrons que l'orthogonal de l'espace engendré par les évaluations est nul. Si $f \in F$ vérifie $\mathrm{ev}_a(f) = f(a) = 0$ pour tout $a \in \mathbf{R}$, alors $f = 0$. Or F^* est de dimension n en particulier il est de dimension fini ce qui montre l'égalité recherchée.

Étape 2 : Construction de $\bar{x}$ solution.

Les évaluations contiennent beaucoup trop d'information pour F^* qui est de dimension n. On extrait une base $(\mathrm{ev}_{x_1}, \dots, \mathrm{ev}_{x_n})$ de F^* on on montre que $\bar{x} = (x_1, \dots, x_n)$ convient. Si $L_1, \dots, L_n$ sont les lignes de $M_{\bar{x}}$ et si $(\lambda_1, \dots, \lambda_n) \in \mathbf{R}^n$ vérifie :

$$\sum_{i=1}^{n} \lambda_i L_i = 0,$$

alors

$$\forall j \in \{1, \dots, n\}, \qquad \sum_{i=1}^{n} \lambda_i f_i(x_j) = \mathrm{ev}_{x_j}\left(\sum_{i=1}^{n} \lambda_i f_i\right) = 0$$

donc $\sum_{i=1}^{n} \lambda_i f_i = 0$ puis $\lambda_1 = \cdots = \lambda_n = 0$ par liberté de $(f_1, \dots, f_n)$. □

Remarque : Ce lemme est vrai si on remplace $\mathbf{R}$ par n'importe quel corps $\mathbf{K}$.

Démonstration du théorème.
(ii) $\implies$ (i) Simple car les translatés de f vérifient la même équation différentielle que f et par le théorème de Cauchy-Lipschitz linéaire, l'ensemble des solutions est un espace vectoriel de dimension finie.

(i) $\implies$ (ii) On note $n := \dim(F)$ et $a_1, \dots, a_n$ tels que $(f_{a_1}, \dots, f_{a_n})$ soit une base de F. Puisque f est dérivable, ses translatés aussi et $F \subset \mathscr{D}(\mathbf{R}, \mathbf{R})$. Il suffit de montrer que F est stable par dérivation. On aura alors f de classe $\mathscr{C}^\infty$ et Considérant la famille infinie des dérivées de f, elle admettra une combinaison linéaire nulle non triviale, ce qui est précisément l'équation différentielle recherchée.

Soit $g \in F$. Il existe alors $\lambda_1, \dots, \lambda_n : \mathbf{R} \to \mathbf{R}$ telles que :

$$\forall h \in \mathbf{R}, \qquad g_h = \sum_{i=1}^{n} \lambda_i(h) f_{a_i} \tag{10}$$

et si $h \neq 0$,

$$\forall x \in \mathbf{R}, \qquad \frac{g(x+h) - g(x)}{h} = \sum_{i=1}^{n} \frac{\lambda_i(h) - \lambda_i(0)}{h} f_{a_i}(x)$$

Ainsi, si on montre que les λ_i sont dérivables, notamment en 0, on aura montré que g' est dans F. On va en fait prouver que les λ_i sont dans F. Pour cela, utilisons le lemme : il existe $\bar{x} = (x_1, \dots, x_n) \in \mathbf{R}^n$ tel que $M_{\bar{x}} = (f_{a_i}(x_j))_{1 \leqslant i,j \leqslant n}$ soit inversible. Or, pour tout $a \in \mathbf{R}$ et pour tout $1 \leqslant j \leqslant n$:

$$g(a + x_j) = g_a(x_j) = \sum_{i=1}^{n} \lambda_i(a) f_{a_i}(x_j)$$

soit

$$\begin{pmatrix} g(a+x_1) \\ \vdots \\ g(a+x_n) \end{pmatrix} = \begin{pmatrix} f_{a_1}(x_1) & \cdots & f_{a_n}(x_1) \\ \vdots & & \vdots \\ f_{a_1}(x_n) & \cdots & f_{a_n}(x_n) \end{pmatrix} \begin{pmatrix} \lambda_1(a) \\ \vdots \\ \lambda_n(a) \end{pmatrix} = {}^t M_{\bar{x}} \begin{pmatrix} \lambda_1(a) \\ \vdots \\ \lambda_n(a) \end{pmatrix}$$

donc

$$\begin{pmatrix} \lambda_1(a) \\ \vdots \\ \lambda_n(a) \end{pmatrix} = {}^t M_{\bar{x}}{}^{-1} \begin{pmatrix} g(a+x_1) \\ \vdots \\ g(a+x_n) \end{pmatrix}$$

et puisque les coefficients de ${}^t M_{\bar{x}}{}^{-1}$ ne dépendent pas de a, cette écriture donne les λ_i comme combinaisons linéaires des translatés de g qui sont dans F. Ainsi, $\lambda_1, \dots, \lambda_n \in F$ et $g' = \sum_{i=1}^{n} \lambda_i'(0) f_{a_i} \in F$ ce qui conclut. □

Remarques :

- L'ordre de l'équation différentielle et la dimension de F sont égaux car $p = \dim \mathrm{Vect}(f, f', f'', \dots) \leqslant n = \dim F$ et F est un sev de l'espace des solutions donc $n \leqslant p$.
- On peut supposer f uniquement continue (voire L^1_{loc}) en travaillant avec $F(x) = \int_0^x f(t)\mathrm{d}t$ et en remarquant que si les f_{a_i} sont libres, alors les F_{a_i} le sont aussi. On a alors en intégrant dans 10, $G(x+a) = \sum_{i=1}^{n} \lambda_i(a) \int_0^x f_{a_i}(t)\mathrm{d}t = \sum_{i=1}^{n} \lambda_i(a) F_{a_i}(x)$.

0.16 Sous-groupes compacts de $\mathrm{GL}_n(\mathbf{R})$

Leçons : 101 ; 106 ; 150 ; 154 ; 158 ; 160 ; 170 ; 171 ; 181 ; 191 ; 208 ; 219 ; 229 ; 253

Références :

— Clarence Kineider, `http://perso.eleves.ens-rennes.fr/people/clarence.kineider/agreg.html`

Lemme 0.16.1.
Si K est une partie compacte d'un espace vectoriel E de dimension finie, alors $\mathrm{Conv}(K)$ est encore compacte.

Démonstration.
Notons $n := \dim E$ et posons

$$L := \left\{ (\lambda_1, \dots, \lambda_{n+1}) \in [0,1]^{n+1} \mid \sum_{i=1}^{n} \lambda_i = 1 \right\}$$

partie compacte de $\mathbf{R}^{n+1}$ (fermée et bornée en dimension finie). Alors, par théorème de Carathéodory,

$$\varphi : \left| \begin{array}{ccc} K^{n+1} \times L & \to & \mathrm{Conv}(X) \\ ((x_i),(\lambda_i)) & \mapsto & \sum_{i=1}^{n} \lambda_i x_i \end{array} \right.$$

est surjective, avec $K^{n+1} \times L$ compact d'où $\mathrm{Conv}(K)$ compacte. □

Théorème 0.16.2 (point fixe de Kakutani)**.**
Soient E un $\mathbf{R}$-espace vectoriel de dimension finie et G un sous-groupe compact de $\mathrm{GL}(E)$.
Si $K \subset E$ est compact, convexe, non-vide et stable par l'action de G, alors il existe un unique $x \in E$ tel que pour tout $g \in G$, $g(x) = x$.

Démonstration.
Soit $\|\cdot\|$ une norme euclidienne sur E. À partir de cette norme quelconque, je vais construire une nouvelle norme, invariante par l'action du groupe G. Par compacité de K, il va exister un élément de norme minimale et par convexité il sera unique. Puisque les éléments du groupe G ne modifient pas cette norme, ils fixeront nécessairement tous ce point.
Soit $N : E \to \mathbf{R}$ définie par $N(x) = \max_{g \in G} \|g(x)\|$. Elle est bien définie car l'application d'évaluation en x et $\|\cdot\|$ sont continues et G est compact. Montrons que N est une norme invariante par l'action de G.

- Invariance par l'action de G : si $g \in G$, $x \in E$,
$$N(g(x)) = \max_{h \in G} \|(gh)(x)\| = \max_{h \in G} \|h(x)\| = N(x)$$
la fonction de multiplication par g étant bijective de G sur lui-même.
- Homogénéité : immédiat.
- Positivité : immédiat.
- Définition : si $N(x) = 0$, alors $\forall x \in G$, $g(x) = 0$ donc $x = 0$ avec $g = \mathrm{Id}$ par exemple.
- Inégalité triangulaire : Soient $x, y \in E$ et $g_0 \in G$ tel que $N(x+y) = \|g_0(x+y)\|$. Alors,
$$N(x+y) = \|g_0(x) + g_0(y)\| \leqslant \|g_0(x)\| + \|g_0(y)\| \leqslant N(x) + N(y)$$
avec égalité si et seulement s'il existe $\lambda \in \mathbf{R}^+$ tel que $g_0(x) = \lambda g_0(y)$ soit $g_0(x - \lambda y) = 0$ donc, g_0 étant inversible, si et seulement si $\exists \lambda \in \mathbf{R}^+$, $x = \lambda y$.

Donc N est une norme invariante par l'action de G, continue car une norme est lipschitzienne donc admet un minimum sur K compact : il existe $z \in K$ tel que
$$N(z) = \min_{x \in K} N(x) =: a$$
Montrons que ce minimum est unique. Soit $y \in K$ tel que $N(y) = a$. Alors $\frac{z+y}{2} \in K$ et
$$a \leqslant N\left(\frac{z+y}{2}\right) \leqslant \frac{1}{2}(N(z) + N(y)) = a$$
il y a donc égalité dans l'inégalité triangulaire précédente, z et y sont positivement liés et $N(z) = N(y)$ donc $z = y$.
Puisque N est invariante par l'action de G, pour tout $g \in G$, $g(z) = z$. □

Remarque :

Si G était fini, on aurait pu considérer à la place le produit scalaire :

$$(x,y) \mapsto \sum_{g\in G} \langle g(x)\,, g(y)\rangle$$

et la norme associée serait invariante par G.

Théorème 0.16.3.

Tout sous-groupe compact de $\mathrm{GL}_n(\mathbf{R})$ *est inclus dans le groupe orthogonal d'une norme.*

Démonstration.

Soit $H \leqslant \mathrm{GL}_n(\mathbf{R})$ compact. On considère l'action par congruence $H \curvearrowright \mathscr{S}_n(\mathbf{R})$:

$$\begin{array}{rccc} \rho : & H & \to & \mathrm{GL}(\mathscr{S}_n(\mathbf{R})) \\ & M & \mapsto & (S \mapsto MSM^T) \end{array}$$

Cette action est bien définie puisque $\mathscr{S}_n(\mathbf{R})$ est stable par congruence. Le morphisme ρ est continu car pour tout $M \in H$, on peut donner la matrice de l'application linéaire $\rho(H)$ et ses coefficients sont des polynômes en les coefficients de M.
On note $G := \rho(H) \leqslant \mathrm{GL}(\mathscr{S}_n(\mathbf{R}))$ compact. Ensuite $G \cdot I_n = \{MM^T \mid M \in H\} \subset \mathscr{S}_n^{++}(\mathbf{R})$ comme sous-ensemble de l'orbite de I_n par $\mathrm{GL}_n(\mathbf{R})$-congruence et est compact car G l'est et non-vide. On souhaite appliquer le théorème de point fixe précédent mais $G \cdot I_n$ n'est à priori pas convexe. En fait, on va forcer sa convexité en considérant son enveloppe convexe. Celle-ci reste incluse dans $\mathscr{S}_n^{++}(\mathbf{R})$ car celui-ci est convexe.
Soit $K := \mathrm{Conv}(G \cdot I_n) \subset \mathscr{S}_n^{++}(\mathbf{R})$. Alors K est convexe, non-vide, stable par G et compacte par le lemme 1. Par théorème de Kakutani, $G \curvearrowright K$ admet un point fixe $S \in K \subset \mathscr{S}_n^{++}(\mathbf{R})$. Donc

$$H \subset \{M \in \mathrm{GL}_n(\mathbf{R}) \mid MSM^T = S\} = O(q)$$

où q est la norme associée à S. □

Remarque :

À nouveau, si G était fini, on aurait pu considérer directement le produit scalaire :

$$(x,y) \mapsto \sum_{g\in G} \langle g(x)\,, g(y)\rangle$$

donc le théorème généralise l'existence de produit scalaire moyenné pour les groupes compacts.

Corollaire 0.16.4.

Tout sous-groupe de $\mathrm{GL}_n(\mathbf{R})$ *est inclus dans un conjugué de* $O_n(\mathbf{R})$.
Autrement dit, $O_n(\mathbf{R})$ *est l'unique sous-groupe compact maximal de* $\mathrm{GL}_n(\mathbf{R})$*, à isomorphisme près.*

Démonstration.

On continue la preuve précédente. Il existe (une unique) $R \in \mathscr{S}_n^{++}(\mathbf{R})$ tel que $S = R^2$. Soit $M \in H$.

$$\begin{aligned} MR^2M^T = R^2 &\iff R^{-1}MR^2M^TR^{-1} = I_n \\ &\iff (R^{-1}MR)(R^{-1}MR)^T = I_n \end{aligned}$$

Donc $R^{-1}HR \subset O_n(\mathbf{R})$ soit $H \subset RO_n(\mathbf{R})R^{-1}$. □

0.17 Simplicité de $\mathrm{SO}_n(\mathbf{R})$

Leçons : 103 ; 106 ; 154 ; 204 ; 214

Références :

— Benjamin Havret, `http://www.normalesup.org/~havret/pdf/developpements_maths%20bhavret.pdf`

Lemme 0.17.1.
Soit G un groupe topologique connexe. Si $H \leqslant G$ contient un voisinage du neutre e_g, alors $H = G$.

Démonstration.
Par hypothèse, $e_g \in \overset{\circ}{H}$. On a alors les égalités :

$$H = \bigcup_{h \in H} h\overset{\circ}{H} \qquad \text{et} \qquad G \setminus H = \bigcup_{g \in G \setminus H} g\overset{\circ}{H}$$

Explications : Les deux inclusions $\supset$ sont immédiates car $e_g \in \overset{\circ}{H}$. L'inclusion $\subset$ dans la première égalité est évidente. Dans la seconde, il suffit de montrer que $g\overset{\circ}{H}$ ne rencontre pas H, mais si $h \in H$, $h' \in \overset{\circ}{H}$ et $g \in G \setminus H$ vérifient $h = gh'$, alors $g = h(h')^{-1} \in H$ ce qui est absurde.
La première égalité montre que H est ouvert et la seconde que son complémentaire est ouvert, donc que H est fermé. Le groupe G étant connexe, on obtient $H = \emptyset$, impossible car $e_g \in H$ ou $H = G$ d'où le résultat. □

Proposition 0.17.2.
Pour tout $n \in \mathbf{N}^*$, $\mathrm{SO}_n(\mathbf{R})$ est une sous-variété de dimension $\frac{n(n-1)}{2}$ et son espace tangent en I_n est l'espace vectoriel des matrices antisymétriques $\mathscr{A}_n(\mathbf{R})$.

Démonstration.
On montre le résultat pour $O_n(\mathbf{R})$, ensuite on utilise que $\mathrm{SO}_n(\mathbf{R})$ est la composante neutre de $O_n(\mathbf{R})$. L'ensemble $O_n(\mathbf{R})$ est donné implicitement par $g(X) = 0$ avec $g(X) = {}^tXX - I_n$ de classe $\mathscr{C}^1$ ($\mathscr{C}^\infty$ même) de $\mathscr{M}_n(\mathbf{R})$ dans $\mathscr{S}_n(\mathbf{R})$ l'ensemble des matrices symétriques. Montrons que $\mathrm{d}g_X$ est une surjection pour tout $X \in O_n(\mathbf{R})$.
Soit $X \in O_n(\mathbf{R})$ et $H \in \mathscr{M}_n(\mathbf{R})$.

$$\mathrm{d}g_X.H = {}^tXH + {}^tHX = X^{-1}H + {}^t(X^{-1}H)$$

donc si $S \in \mathscr{S}_n(\mathbf{R})$, $\mathrm{d}g_X.H = S$ admet notamment la solution $H = \frac{1}{2}XY$ ce qui établit la surjectivité. Ainsi, G est une sous-variété de $\mathscr{M}_n(\mathbf{R}) \simeq \mathbf{R}^{n^2}$, de dimension

$$\dim O_n(\mathbf{R}) = \dim \mathscr{M}_n(\mathbf{R}) - \dim \mathscr{S}_n(\mathbf{R}) = \frac{n(n-1)}{2}$$

et son espace tangent en X est $\ker \mathrm{d}g_{I_n} = \left\{ H \in \mathscr{M}_n(\mathbf{R}) \mid {}^t(X^{-1}H) = -X^{-1}H \right\}$.
Lorsque $X = I_n$, il s'agit de $\mathscr{A}_n(\mathbf{R})$. □

Théorème 0.17.3.
Si n est un entier impair différent de 1, alors $\mathrm{SO}_n(\mathbf{R})$ est simple.

En fait, dans le cas $n \geqslant 4$ pair, on a $Z(\mathrm{SO}_n(\mathbf{R})) = \{I_n, -I_n\}$ donc $\mathrm{SO}_n(\mathbf{R})$ ne peut être simple, mais c'est le seul obstacle : $\forall n \in \mathbf{N} \setminus \{1, 2, 4\}$, $\mathbb{P}\mathrm{SO}_n(\mathbf{R})$ est simple. Le cas $n = 4$ est particulier... $\mathbb{P}\mathrm{SO}_4(\mathbf{R})$ n'est pas simple.

Démonstration.
Soit $n \neq 1$ un entier impair et $H \triangleleft \mathrm{SO}_n(\mathbf{R})$ non réduit à $\{I_n\}$. On cherche alors à montrer que $H = \mathrm{SO}_n(\mathbf{R})$. Par le lemme, il est suffisant de prouver que H contient un voisinage de l'identité. On voudrait donc construire un homéomorphisme local au voisinage de l'identité de $\mathrm{SO}_n(\mathbf{R})$ sur H. En fait, il est plus facile de construire un difféomorphisme locale grâce au théorème d'inversion locale. Mais $\mathrm{SO}_n(\mathbf{R})$ n'est pas un ouvert d'un espace vectoriel, donc pour faire du calcul différentiel, on a besoin d'utiliser sa structure de sous-variété.

Étape 1 : Candidat de difféomorphisme

Si $w_1, \dots, w_r \in H$, on pose :

$$\begin{array}{rccc} \varphi : & (\mathrm{SO}_n(\mathbf{R}), I_n) & \to & (H, I_n) \\ & u & \mapsto & [u, w_1] \dots [u, w_r] \end{array}$$

où la notation $(\mathrm{SO}_n(\mathbf{R}), I_n) \to (H, I_n)$ signifie que I_n est envoyé sur lui-même (je vois φ comme une application entre espaces topologiques pointés). On remarque que l'ensemble d'arrivé est bien H puisque il s'agit d'un sous-groupe distingué donc tout commutateur entre un élément de H et un élément de $\mathrm{SO}_n(\mathbf{R})$ reste dans H.
L'application φ dépend du choix des w_i, je décide de ne pas noter la dépendance pour ne pas alourdir les notations. Pour le moment, on va étudier φ en fonction des paramètres w_i, le but étant de trouver un choix de w_i qui fait de φ un difféomorphisme solution.
L'application φ est $\mathscr{C}^1$ car polynomiale en les coefficients de la matrice. Et par un calcul sans difficulté, on obtient :

$$\begin{array}{rccc} \mathrm{d}\varphi_{I_n} : & \mathscr{A}_n(\mathbf{R}) & \to & \mathscr{A}_n(\mathbf{R}) \\ & a & \mapsto & a - \sum_{i=1}^r w_i a w_i^{-1} \end{array}$$

Détails : On calcule en fait la différentielle de φ vue comme fonction $\mathrm{GL}_n(\mathbf{R}) \to \mathrm{GL}_n(\mathbf{R})$ pour se ramener à un ouvert, sinon on ne sait pas calculer... Pour $a \in \mathscr{M}_n(\mathbf{R})$ assez petit, $I_n + a \in \mathrm{GL}_n(\mathbf{R})$ et :

$$\varphi(I_n + a) - \varphi(I_n) = \prod_{i=1}^r (I_n + a) w_i (I_n + a)^{-1} w_i^{-1} \ - I_n = ra - \sum_{i=1}^r w_i a w_i^{-1} + \underset{a \to 0}{o}(a)$$

Donc l'application $\varphi : \mathrm{GL}_n(\mathbf{R}) \to \mathrm{GL}_n(\mathbf{R})$ a pour différentielle en I_n :

$$\begin{array}{rccc} \mathrm{d}\varphi_{I_n} : & \mathscr{M}_n(\mathbf{R}) & \to & \mathscr{M}_n(\mathbf{R}) \\ & a & \mapsto & a - \sum_{i=1}^r w_i a w_i^{-1} \end{array}$$

Cette expression est elle valable pour la différentielle de $\varphi : \mathrm{SO}_n(\mathbf{R}) \to H$?
Soit $a \in \mathscr{A}_n(\mathbf{R})$ et soit γ un arc tracé vérifiant $\gamma(0) = I_n$ et $\gamma'(0) = a$ (définition de l'espace tangent). Alors,

$$(\varphi \circ \gamma)'(0) = \mathrm{d}\varphi_{\gamma(0)}.\gamma'(0) = \mathrm{d}\varphi_{I_n}.a$$

ce calcul étant valable pour $\varphi : \mathrm{GL}_n(\mathbf{R}) \to \mathrm{GL}_n(\mathbf{R})$ ou $\varphi : \mathrm{SO}_n(\mathbf{R}) \to H$ la réponse est oui.

Remarques :

1. Plutôt que d'utiliser directement la structure de variété de $\mathrm{SO}_n(\mathbf{R})$ et son plan tangent, on peut introduire la paramétrisation $\exp : \mathscr{A}_n(\mathbf{R}) \to \mathrm{SO}_n(\mathbf{R})$ au voisinage de l'identité (principe de base en algèbres de Lie).
2. On peut se demander pourquoi s'embêter à regarder $\varphi : \mathrm{SO}_n(\mathbf{R}) \to H$ plutôt que $\varphi : \mathrm{GL}_n(\mathbf{R}) \to \mathrm{GL}_n(\mathbf{R})$... En fait, on a besoin d'avoir $\mathrm{d}\varphi$ inversible pour appliquer un théorème d'inversion locale, mais si on considère la différentielle sur tout $\mathscr{M}_n(\mathbf{R})$, alors le noyau contient toujours les homothéties quels que soient les w_i. Celles-ci disparaissent lorsqu'on se restreint aux matrices antisymétriques.

On aimerait que φ soit un difféomorphisme local, donc que sa différentielle en I_n soit inversible. Puisqu'on est en dimension finie, il suffit de montrer qu'elle est injective.

Étape 2 : Calcul de $\ker \mathrm{d}\varphi_{I_n}$

Soit $a \in \ker \mathrm{d}\varphi_{I_n}$. On a utilisant que les w_i sont orthogonales :

$$\|a\| = \frac{1}{r} \left\| \sum_{i=1} r w_i a w_i^{-1} \right\| \leqslant \frac{1}{r} \sum_{i=1}^r \|w_i a w_i^{-1}\| = \|a\|$$

et il y a ainsi égalité dans l'inégalité triangulaire, donc a et les $w_i a w_i^{-1}$ sont sur une même demi-droite. Étant de même norme, ils sont donc égaux. Finalement, $\forall i \in \{1, \dots, r\}$, $w_i a w_i^{-1} = a$ donc a commute avec tous les w_i. La réciproque étant évidente, on obtient :

$$\ker \mathrm{d}\varphi_{I_n} = \{a \in \mathscr{A}_n(\mathbf{R}) \mid \forall i \in \{1, \dots, r\},\ [a, w_i] = I_n\}$$

On recherche donc des éléments w_i dans H tels que toute matrice antisymétrique commutant avec tous ces éléments est nulle. On sait que H contient un élément distinct de l'identité.

Étape 3 : Tout élément $w \neq I_n$ de H fixe un sous-espace strict de $\mathbf{R}^n$

Soit $w \in H$ avec $w \neq I_n$. Le théorème de structure des matrices orthogonales affirme que w est semblable à :

$$\begin{pmatrix} R_1 & & & & & \\ & \ddots & & & 0 & \\ & & R_p & & & \\ & & & \varepsilon_1 & & \\ & 0 & & & \ddots & \\ & & & & & \varepsilon_q \end{pmatrix}$$

où les R_i sont des blocs de rotation d'ordre 2 et les ε_i sont égaux à 1 ou -1. Puisqu'on est en dimension impaire, il ne peut y avoir que des blocs de rotation, il y a au moins une droite stable. Mais si tous les ε valent -1, le déterminant de w serait $(-1)^n = -1$ ce qui est absurde. Donc :

$$\{0\} \neq \underbrace{\mathrm{Inv}(w)}_{V} \neq \mathbf{R}^n$$

et $d := \dim V \in \{2, \dots, n-1\}$.

Étape 4 : Construction de w_i solutions

Par principe de conjugaison, si $u \in \mathrm{SO}_n(\mathbf{R})$, $\mathrm{Inv}(uwu^{-1}) = u(V)$ et $uwu^{-1} \in H$ car $H \triangleleft \mathrm{SO}_n(\mathbf{R})$. On note $\mathscr{I} = \mathscr{P}_d(\{1, \dots, n\})$ l'ensemble des parties à d éléments de $\{1, \dots, n\}$ et $(e_1, \dots, e_n)$ la base canonique.
Puisque $\mathrm{SO}_n(\mathbf{R})$ agit transitivement sur les grassmanniennes d'ordre d, c'est-à-dire l'ensemble des sous-espaces de dimension d, on peut choisir $u_I \in \mathrm{SO}_n(\mathbf{R})$ telle que $u_I(V) = \mathrm{Vect}(e_i \ ; \ i \in I)$ pour tout $I \in \mathscr{I}$. On pose alors $w_I = u_I w u_I^{-1} \in H$ vérifiant $\mathrm{Inv}(w_I) = \mathrm{Vect}(e_i \ ; \ i \in I)$.
On choisit alors la famille $(w_I)_{I \in \mathscr{I}}$. Si une matrice antisymétrique commute avec tous les éléments de cette famille, elle laisse stable tous les espaces de dimension d engendrés par les vecteurs de la base, donc par intersections toutes les droites engendrées par les e_i. Elle est donc diagonale, mais étant antisymétrique, elle est nulle et on en déduit que cette famille convient ! □

0.18 Exponentielle d'une somme et trigonalisation simultanée

Leçons : 156 ; 157

Références :

— Benjamin Havret `http://www.normalesup.org/~havret/pdf/developpements_maths%20bhavret.pdf`

Si $A, B \in \mathscr{M}_n(\mathbf{C})$ commutent, elles sont cotrigonalisables. On va montrer que cette hypothèse peut être légèrement affaiblie.
Pour cela, on utilisera notamment que si $A, B \in \mathscr{M}_n(\mathbf{C})$ commutent alors tout polynôme en A commute avec B et cela s'étend même à $\exp(A)$ sans difficulté.

Lemme 0.18.1.
Si $A, B \in \mathscr{M}_n(\mathbf{C})$ et si $[A,[A,B]] = 0$ où $[A,B] = AB - BA$ est le commutateur de A et B, alors

$$e^A B e^{-A} = B + [A,B]$$

Démonstration.
Si $t \in \mathbf{R}$,

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}[e^{tA}Be^{-tA}] &= e^{tA}ABe^{-tA} - e^{tA}BAe^{-tA} \\ &= [A,B]e^{tA}e^{-tA} \qquad \text{car } A \text{ commute avec } [A,B] \\ &= [A,B]\end{aligned}$$

En intégrant, on obtient : pour tout $t \in \mathbf{R}$, $e^{tA}Be^{-tA} = B + t[A,B]$ puis on évalue en $t = 1$. □

Proposition 0.18.2.
Si $A, B \in \mathscr{M}_n(\mathbf{C})$ commutent toutes deux avec $N := [A,B]$, alors pour tout $z \in \mathbf{C}$,

$$e^{z(A+B)} = e^{zA}e^{-\frac{z^2}{2}N}e^{zB}$$

Démonstration.
On commence par montrer le résultat sur les réels par le même procédé que le lemme.
Si $t \in \mathbf{R}$,

$$\frac{\mathrm{d}}{\mathrm{d}t}[\underbrace{e^{tA}e^{-\frac{t^2}{2}N}e^{tB}}_{\phi(t)}] = e^{tA}(A + B - tN)e^{-\frac{t^2}{2}N}e^{tB}$$

car on dérive un produit de trois termes qui font "sortir" A, $-tN$ et B respectivement et on peut regrouper ces trois termes d'un côté ou de l'autre de $e^{-\frac{t^2}{2}N}$ par l'hypothèse de commutation. Maintenant, on aimerait regrouper les exponentielles toutes d'un même côté, par exemple à droite, pour obtenir une équation différentielle. On peut échanger e^{tA} avec A ou $-tN$ car A commute avec N. Mais B ne commute pas avec A... C'est là qu'on utilise le lemme 1 qui quantifie le défaut de commutation entre B et e^{tA}.
Par le lemme 1, $e^{tA} = (B + tN)e^{tA}$ donc, le terme correctif tN venant compenser le $-tN$ dans la formule précédente :

$$\phi'(t) = (A + B)\phi(t) \tag{11}$$

d'où $\phi(t) = e^{t(A+B)}$. Mais ϕ que l'on étend sur $\mathbf{C}$ avec la même expression et $z \mapsto e^{z(A+B)}$ sont entières, au sens où chaque composante est une fonction holomorphe sur tout $\mathbf{C}$, et coïncident sur $\mathbf{R}$ qui est d'accumulation dans le connexe $\mathbf{C}$ donc par théorème de prolongement analytique :

$$\forall z \in \mathbf{C}, \qquad \phi(z) = e^{z(A+B)}$$

d'où le résultat. □

Théorème 0.18.3.
Si $A, B \in \mathscr{M}_n(\mathbf{C})$ commutent toutes deux avec $N := [A,B]$, alors N est nilpotente et A et B sont cotrigonalisables.

Démonstration.

Étape 1 : N est nilpotente.

Le résultat précédent se réécrit : $\forall z \in \mathbf{C}$, $e^{\frac{z^2}{2}N} = e^{zA}e^{-z(A+B)}e^{zB}$.
On remarque qu'à gauche on a un terme quadratique en z et à droite des termes linéaires, on va exploiter ce point pour montrer que N est nilpotente.
Soit $\lambda \in \sigma(N)$ et v un vecteur propre associé. Supposons $\lambda \neq 0$ et notons $\mu \in \mathbf{C}$ tel que $\mu^2 = \frac{1}{\lambda}$ et $z(x) = \sqrt{2}\mu x$ pour $x \in \mathbf{R}$. Alors, d'une part

$$e^{\frac{z(x)^2}{2}N}v = e^{\frac{z(x)^2}{2}\lambda}v = e^{x^2}v$$

tandis que

$$\begin{aligned}\left\|e^{z(x)A}e^{-z(x)(A+B)}e^{z(x)B}v\right\| &\leqslant \|e^{z(x)A}\|\,\|e^{-z(x)(A+B)}\|\,\|e^{z(x)B}\|\,\|v\| \\ &\leqslant e^{\sqrt{2}|\mu|(\|A\|+\|B\|+\|A+B\|)x}\|v\|\end{aligned}$$

et on obtient une contradiction prenant $[x \to \infty]$. Donc $\lambda = 0$ et N est nilpotente.

Étape 2 : Cotrigonalisation.

Si $N = 0$, alors A et B commutent donc on a le résultat.
Sinon, N étant nilpotente, elle est non inversible et $\ker N \neq \{0\}$ et est stable par A et B par commutation. Dans une base adaptée à l'inclusion $\ker N \subset \mathbf{R}^n$:

$$A = \begin{pmatrix} A' & * \\ 0 & A'' \end{pmatrix}, \qquad B = \begin{pmatrix} B' & * \\ 0 & B'' \end{pmatrix} \quad \text{et} \quad \begin{pmatrix} 0 & * \\ 0 & N'' \end{pmatrix}$$

puis $AB - BA = N$ donne

$$\begin{cases} A'B' - B'A' = 0 \\ A''B'' - B''A'' = N'' \end{cases}$$

donc A', B' sont cotrigonalisables et $AN = NA$ donne $A''N'' = N''A''$ (idem avec B). Donc A'' et B'' sont deux matrices commutant avec leur commutateur, de taille strictement plus petite que A et B. On en déduit par récurrence, initialisée à $n = 1$ trivialement que A et B sont cotrigonalisables. □

0.19 Exponentielle homéomorphe de $\mathscr{S}_n(\mathbf{R})$ sur $\mathscr{S}_n^{++}(\mathbf{R})$

Leçons : 101 ; 150 ; 153 ; 155 ; 156 ; 158 ; 170 ; 171

Références :

—

Lemme 0.19.1.
Sur $\mathscr{S}_n(\mathbf{R})$, la norme subordonnée à la norme $\|\cdot\|_2$ coïncide avec le rayon spectral ρ.

Démonstration.
Par théorème spectral, soit $(v_1, \dots, v_n)$ une base orthonormée de vecteurs propres de S. On note $\lambda_1, \dots, \lambda_n$ les valeurs propres associées. Quitte à réordonner les v_1, on peut supposer $|\lambda_1| \leqslant \cdots \leqslant |\lambda_n|$.
Si $x = \sum_{i=1}^n x_i v_i$ est de norme 1, on a :

$$\|Sx\|_2 = \left\| \sum_{i=1}^n \lambda_i x_i v_i \right\| \leqslant |\lambda_n| \, \|x\| = \rho(S)$$

avec égalité pour $x = v_n$. □

Lemme 0.19.2.
La conjugaison par une matrice inversible commute avec l'exponentielle :

$$\forall P \in \mathrm{GL}_n(\mathbf{R}), \forall M \in \mathscr{M}_n(\mathbf{R}), \qquad \exp(PMP^{-1}) = P \exp(M) P^{-1}$$

Démonstration.
Cela découle du fait que la conjugaison est un morphisme d'algèbre continu.
Si $M \in \mathscr{M}_n(\mathbf{R})$, $P \in \mathrm{GL}_n(\mathbf{R})$ et $N \in \mathbf{N}$:

$$\sum_{k=0}^N \frac{1}{k!} (PMP^{-1})^k = P \left(\sum_{k=0}^N \frac{1}{k!} M^k \right) P^{-1}$$

puis par continuité de la conjugaison $M \mapsto PMP^{-1}$, $\exp(PMP^{-1}) = P \exp(M) P^{-1}$. □

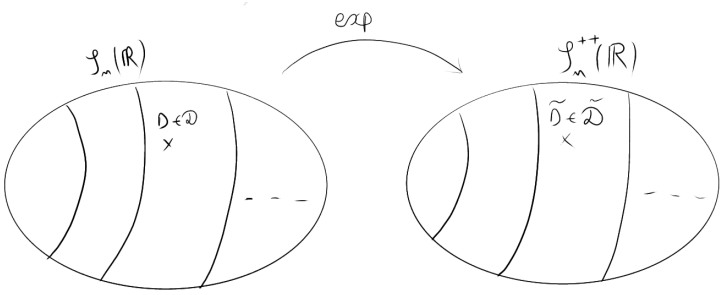
Lemme 0.19.3.
Si $D = \begin{pmatrix} \lambda_1 I_{n_1} & & 0 \\ & \ddots & \\ 0 & & \lambda_k I_{n_k} \end{pmatrix}$, *alors son commutant est*

$$\mathrm{Com}(D) = \left\{ \begin{pmatrix} \lambda_1 A_{n_1} & & 0 \\ & \ddots & \\ 0 & & \lambda_k A_{n_k} \end{pmatrix} \ ; \ A_1 \in \mathscr{M}_{n_1}(\mathbf{R}), \dots, A_k \in \mathscr{M}_{n_k}(\mathbf{R}) \right\}.$$

Démonstration.
Une matrice du commutant laisse stable les sous-espaces propres de D d'où la structure par blocs. Réciproquement, le calcul par blocs montre que toute matrice de cette forme commute avec D. □

Théorème 0.19.4.
L'application $\exp : \mathscr{S}_n(\mathbf{R}) \to \mathscr{S}_n^{++}(\mathbf{R})$ *est un homéomorphisme.*

Démonstration.
Je vais démontrer ce théorème en utilisant le vocabulaire des actions de groupe pour bien dégager le rôle du théorème spectral. On note :
$\mathscr{D} = \{\mathrm{diag}(\lambda_1, \dots, \lambda_n) \ ; \ \lambda_1 \leqslant \cdots \leqslant \lambda_n\}$,
$\widetilde{\mathscr{D}} = \{\mathrm{diag}(\mu_1, \dots, \mu_n) \ ; \ 0 < \mu_1 \leqslant \cdots \leqslant \mu_n\}$.
Par le théorème spectral, les ensembles $\mathscr{S}_n(\mathbf{R})$ et $\mathscr{S}_n^{++}(\mathbf{R})$ sont réunions d'orbites sous l'action de $O_n(\mathbf{R})$ par conjugaison. Et chacune de ces orbites contient une unique matrice de $\mathscr{D}$ et de $\widetilde{\mathscr{D}}$ respectivement. Ces matrices sont des formes normales, qui ont l'intérêt d'avoir une image très simple par l'exponentielle.



Étape 1 : exp : $\mathscr{D} \hookrightarrow\!\!\!\rightarrow \widetilde{\mathscr{D}}$.

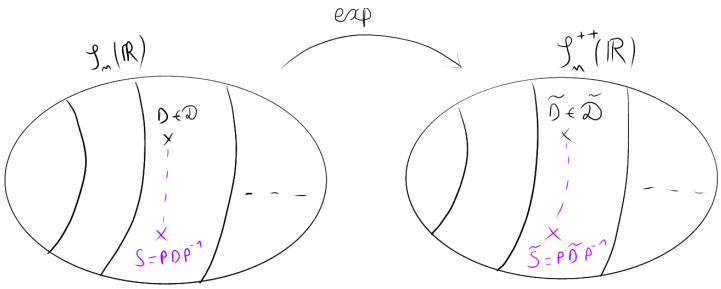
Soient $\widetilde{D} = \mathrm{diag}(\mu_1, \dots, \mu_n) \in \widetilde{\mathscr{D}}$ et $D = \mathrm{diag}(\lambda_1, \dots, \lambda_n)$ une matrice diagonale mais dont les coefficients ne sont pas forcément ordonnés, donc pas nécessairement dans $\mathscr{D}$. Alors, si on résout l'équation en D :

$$\begin{aligned} \exp(D) = \widetilde{D} &\iff \forall i \in [\![1\,,\,n]\!], \quad e^{\lambda_i} = \mu_i \\ &\iff \forall i \in [\![1\,,\,n]\!], \quad \lambda_i = \ln \mu_i \end{aligned}$$

on trouve donc une unique solution par bijection de ln de $\mathbf{R}_+^*$ sur $\mathbf{R}$. De plus, la croissance de la fonction logarithme montre que l'unique solution trouvée est en fait dans $\mathscr{D}$ d'où la bijection recherchée.

Étape 2 : Surjectivité.

Elle découle directement du point précédent et du lemme 2 : si $\widetilde{S} \in \mathscr{S}_n^{++}(\mathbf{R})$ est congruente à $\widetilde{D}$ par la matrice orthogonale P, alors le conjugué par P de l'image réciproque de $\widetilde{D}$ est d'image $\widetilde{S}$.



On remarque aussi que le spectre avec multiplicité que je note Sp est le même pour toute matrice de la même orbite :

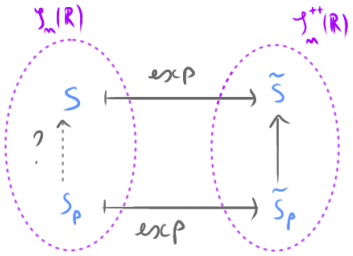
$$\forall S \in \mathscr{S}_n(\mathbf{R}), \qquad \mathrm{Sp}(\exp(S)) = \exp(\mathrm{Sp}(S)). \tag{12}$$

Étape 3 : Injectivité.

Ce qu'on a dit précédemment montre que si deux matrices symétriques sont $O_n(\mathbf{R})$-conjuguées, alors leur images par l'exponentielle sont dans la même orbite. Autrement dit, l'exponentielle induit une bijection entre les orbites de $\mathscr{S}_n(\mathbf{R})$ et de $\mathscr{S}_n^{++}(\mathbf{R})$. Ainsi, il suffit de montrer l'injectivité sur chaque orbite.
Soient $P, Q \in O_n(\mathbf{R})$, $D \in \mathscr{D}$ et $\widetilde{D} := \exp(D)$.
Si $\exp(PDP^{-1}) = \exp(QDQ^{-1})$, alors $P\widetilde{D}P^{-1} = Q\widetilde{D}Q^{-1}$, soit $Q^{-1}P\widetilde{D}(Q^{-1}P)^{-1} = \widetilde{D}$ puis $Q^{-1}P \in \mathrm{Stab}(\widetilde{D}) \subset \mathrm{Com}(\widetilde{D}) = \mathrm{Com}(D)$ par le lemme 3 car D et $\widetilde{D}$ ont la même structure par blocs. Donc $PDP^{-1} = QDQ^{-1}$ d'où l'injectivité.

Étape 4 : Bicontinuité.

L'application exp est continue comme série normalement convergente de fonctions continues.
Pour la continuité de l'inverse, on va montrer la continuité séquentielle. On fixe $\widetilde{S}$ dans $\mathscr{S}_n^{++}(\mathbf{R})$, et on considère une suite $(\widetilde{S}_p)$ dans $\mathscr{S}_n^{++}(\mathbf{R})$ qui tend vers $\widetilde{S}$. On note ensuite S et S_p les images réciproques de $\widetilde{S}$ et $\widetilde{S}_p$. On souhaite montrer qu'alors la suite (S_p) converge vers S.



L'ensemble $\mathscr{S}_n^{++}(\mathbf{R})$ est stable par passage à l'inverse et l'application inverse étant continue :

$$\begin{cases} \widetilde{S}_p \to \widetilde{S} \\ \widetilde{S}_p^{-1} \to \widetilde{S}^{-1} \end{cases}$$

donc ces deux suites sont bornées et par le lemme 1, cette propriété métrique se traduit par une propriété algébrique : l'ensemble des valeurs propres des $\widetilde{S}_p$ est borné ainsi que l'ensemble de leurs inverses. Puisque ces valeurs propres sont strictement positive, on en déduit qu'il existe $c, C > 0$ telles que

$$\forall p \in \mathbf{N}, \qquad \mathrm{Sp}(\widetilde{S}_p) \subset [c, C]^n$$

et donc utilisant 12

$$\forall p \in \mathbf{N}, \qquad \mathrm{Sp}(S_p) \subset [\ln c, \ln C]^n$$

et par le lemme 1 dans l'autre sens, de cette propriété spectrale on déduit la propriété métrique : (S_p) est bornée.

On va montrer que cette suite n'a qu'une unique valeur d'adhérence. Supposons $S_{\varphi(p)} \to \bar{S}$. L'existence de cette valeur d'adhérence est assurée par le caractère bornée en dimension finie. De plus, $\mathscr{S}_n(\mathbf{R})$ est est un espace vectoriel de dimension finie donc est fermé et $\bar{S}$ est symétrique. Par continuité de exp :

$$\begin{array}{ccc} S_{\varphi(p)} & \xrightarrow{p\to+\infty} & \bar{S} \\ \exp\downarrow & & \downarrow\exp \\ \widetilde{S}_{\varphi(p)} & \xrightarrow[p\to+\infty]{} & \widetilde{S} \end{array}$$

et par injectivité $\bar{S} = S$ et (S_p) est bornée avec une unique valeur d'adhérence d'où $S_p \to S$ et la continuité cherchée. □

0.20 Décomposition polaire

Leçons : 101 ; 106 ; 150 ; 153 ; 155 ; 158 ; 170 ; 171

Références :

—

Lemme 0.20.1.
Le sous-ensemble $\mathscr{S}_n^{++}(\mathbf{R})$ est ouvert dans $\mathscr{S}_n(\mathbf{R})$.

Démonstration.
Montrons que son complémentaire $\mathcal{C}$ est fermé. On a $S \in \mathcal{C}$ si et seulement s'il existe $X \in \mathbf{R}^n$ tel que ${}^tXSX \leqslant 0$. Quitte à remplacer X par λX, donc ${}^tXSX$ par $\lambda^2 {}^tXSX$, on peut supposer $X \in \mathbb{S}^{n-1}$.
Soit $S \in \mathscr{S}_n(\mathbf{R})$ et (S_k) une suite de $\mathcal{C}$ qui converge vers S. On a donc une suite (X_k) de $\mathbb{S}^{n-1}$ telle que ${}^tX_kS_kX_k \leqslant 0$. Comme $\mathbb{S}^{n-1}$ est compacte, cette suite a une valeur d'adhérence X et l'on obtient en prenant la limite :

$$ {}^tXSX \leqslant 0$$

donc $S \in \mathcal{C}$. □

Lemme 0.20.2.
Si $S \in \mathscr{S}_n^{++}(\mathbf{R})$, il existe une unique matrice $R \in \mathscr{S}_n^{++}(\mathbf{R})$ telle que $S = R^2$.

Théorème 0.20.3.
Soit $p, q \in \mathbf{N}$ tels que $p + q \leqslant n$. Alors,

$$\overline{\mathrm{Orb}(I_{p,q})} = \bigcup_{\substack{0\leqslant h\leqslant p\\ 0\leqslant k\leqslant q}} \mathrm{Orb}(I_{p-h,q-k}) =: \mathscr{R}.$$

En particulier, $\overline{\mathscr{S}_n^{++}(\mathbf{R})} = \mathscr{S}_n^{+}(\mathbf{R})$.

Démonstration.

Étape 1 : $\mathscr{R} \subset \overline{\mathrm{Orb}(I_{p,q})}$.

Soit $0 \leqslant h \leqslant p$, $0 \leqslant k \leqslant q$ et $\varepsilon > 0$. On pose

$$A_{h,k}(\varepsilon) := \mathrm{Diag}(I_{p-h}, -I_{q-k}, \varepsilon I_h, -\varepsilon I_k, 0_{n-p-q}) \in \mathrm{Orb}(I_{p,q})$$

qui vérifie $A_{h,k}(\varepsilon) \underset{\varepsilon\to 0}{\longrightarrow} I_{p-h,q-k}$. Ainsi, les formes normales des orbites composant $\mathscr{R}$ sont dans $\overline{\mathrm{Orb}(I_{p,q})}$.

Si $S \in \mathrm{Orb}(I_{p-h,q-k})$, il existe $P \in \mathrm{GL}_n(\mathbf{R})$ tel que $S = PI_{p-h,q-k}P^{-1}$ et pour $A \in \mathscr{S}_n(\mathbf{R})$:

$$\underbrace{\left(PA_{h,k}(\varepsilon)P^{-1}\right)}_{\in \mathrm{Orb}(I_{p,q})} \underset{\varepsilon\to 0}{\longrightarrow} S$$

d'où $\mathscr{R} \subset \overline{\mathrm{Orb}(I_{p,q})}$.

Étape 2 : $\mathscr{R}$ est fermé.

On écrit $\mathscr{R}^c = \mathscr{S}_{+\geqslant p+1} \cup \mathscr{S}_{-\geqslant q+1}$ ensembles des matrices symétriques dont les formes quadratiques associées sont définies positives (resp. définies négatives) sur un sous-espace de dimension $p+1$ (resp. $q+1$).
Soit $S \in \mathscr{S}_{+\geqslant p+1}$ et F un sev de $\mathbf{R}^n$ tel que $\dim F = p+1$ et $q_S > 0$ sur F. Soit $\mathscr{B} = (e_1, \ldots, e_{p+1})$ une base de F. Alors, l'application

$$\varphi : \left| \begin{array}{ccc} \mathscr{S}_n(\mathbf{R}) & \to & \mathscr{S}_{p+1}(\mathbf{R}) \\ A & \mapsto & ({}^te_iAe_j)_{1\leqslant i,j\leqslant p+1} \end{array} \right.$$

qui à A associe la matrice de sa forme quadratique associée restreinte à F et dans la base $\mathscr{B}$ est continue et

$$\varphi(A) \in \mathscr{S}_{p+1}^{++}(\mathbf{R}) \implies A \in \mathscr{S}_{+\geqslant p+1}.$$

Donc si U est un voisinage de $\varphi(S)$ inclus dans $\mathscr{S}_{p+1}^{++}(\mathbf{R})$ qui est ouvert dans $S_{p+1}(\mathbf{R})$ par le lemme 1, alors $S \in \varphi^{-1}(U) \subset \mathscr{S}_{+\geqslant p+1}$ et $\mathscr{S}_{+\geqslant p+1}$ est ouvert. On montre de la même façon que $\mathscr{S}_{-\geqslant p+1}$ est ouvert aussi et donc $\mathscr{R}$ est fermé. □

Théorème 0.20.4.
L'application

$$\mu : \left| \begin{array}{ccc} O_n(\mathbf{R}) \times \mathscr{S}_n^{++}(\mathbf{R}) & \to & \mathrm{GL}_n(\mathbf{R}) \\ (O, S) & \mapsto & OS \end{array} \right.$$

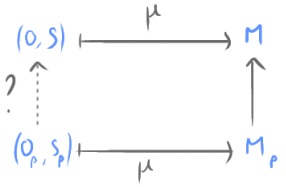
est un homéomorphisme.
En particulier, les orbites pour l'action $O_n(\mathbf{R}) \curvearrowright \mathrm{GL}_n(\mathbf{R})$ par translation contiennent toutes une unique matrice de $\mathscr{S}_n^{++}(\mathbf{R})$.

Démonstration.
On commence par remarquer que μ est bien définie notamment qu'elle est bien à valeurs dans $\mathrm{GL}_n(\mathbf{R})$ et continue.

Surjectivité : soit $M \in \mathrm{GL}(\mathbf{R})$, ${}^tMM \in \mathscr{S}_n^{++}(\mathbf{R})$ car elle est dans l'orbite de I_n par GL_n-congruence. Il existe alors $S \in \mathscr{S}_n^{++}(\mathbf{R})$ telle que $S^2 = {}^tMM$ et donc $M = ({}^tM^{-1}S)S$ avec $O := {}^tM^{-1}S \in O_n(\mathbf{R})$ car $OO^T = {}^tM^{-1}S^2M^{-1} = I_n$.

Injectivité : si $M = OS = O'S'$, alors $S^2 = {}^tS{}^tOOS = {}^tMM = S'^2$ et $S = S'$ par unicité de la racine carrée de ${}^tMM$, puis $O = O'$.

Continuité de μ^{-1} : On montre la continuité séquentielle. On considère $M \in \mathrm{GL}_n(\mathbf{R})$ et une suite (M_p) de $\mathrm{GL}_n(\mathbf{R})$ qui converge vers M. On note $(O, S) = \mu^{-1}(M)$ et $(O_p, S_p) = \mu^{-1}(M_p)$. On cherche à montrer que $O_p \to O$ et $S_p \to S$.



Puisque $O_n(\mathbf{R})$ est compact, il existe $\overline{O} \in O_n(\mathbf{R})$ telle que $O_{\varphi(p)} \to \overline{O}$ et alors

$$S_{\varphi(p)} \to \overline{O}^{-1}M \in \mathrm{GL}_n(\mathbf{R}) \cap \overline{\mathscr{S}_n^{++}(\mathbf{R})} = \mathrm{GL}_n(\mathbf{R}) \cap \mathscr{S}_n^{+}(\mathbf{R}) = \mathscr{S}_n^{++}(\mathbf{R}).$$

Donc notant $\overline{S} = \lim_{p \to +\infty} S_{\varphi(p)}$, $M = \overline{O}\,\overline{S}$ et finalement $O = \overline{O}$, $S = \overline{S}$ par injectivité. □

0.21 Décomposition effective de Dunford

Leçons : 153 ; 155 ; 157

Références :

—

E : k-ev de dimension m.

Lemme 0.21.1.
Si $u \in \mathrm{GL}(E)$, alors $u^{-1} \in k[u]$.

Lemme 0.21.2.
Si $u \in \mathrm{GL}(E)$ et $\nu \in \mathscr{L}(E)$ est nilpotente avec $u\nu = \nu u$, alors $u + \nu \in \mathrm{GL}(E)$.

Démonstration.
Soit $n_0 \in \mathbf{N}$ tel que $\nu^{n_0} = 0$. Alors,

$$u + \nu = u(\mathrm{id} + u^{-1}\nu)$$

or grâce au lemme 1, u^{-1} et ν commutent donc

$$(\mathrm{id} - u^{-1}\nu) \sum_{i=0}^{n_0-1} (u^{-1}\nu)^i = \mathrm{id} - u^{n_0}\nu^{n_0} = \mathrm{id}$$

donc

$$(u + \nu)\left(u^{-1} \sum_{i=0}^{n_0-1} (u^{-1}\nu)^i\right) = \mathrm{id}\,.$$

□

Le but de ce développement est de donner une méthode effective pour calculer la décomposition de Dunford sans avoir à calculer les valeurs propres. On sait qu'un endomorphisme diagonalisable est annulé par un polynôme scindé simple. Le but ici est de choisir un bon polynôme et de chercher un de ses "endomorphismes racines" à l'aide d'un schéma de Newton.

Théorème 0.21.3.
Soit $u \in \mathscr{L}(E)$ dont le polynôme minimal est scindé.
Alors, il existe un unique couple $(d, \nu) \in \mathscr{L}(E)^2$ tel que $d\nu = \nu d$ et $u = d + \nu$ avec d diagonalisable et ν nilpotente.
De plus, $d, \nu \in k[u]$ et si $Q \in k[X]$ est scindé à racines simples tel que $Q(u)$ est nilpotente, alors la suite définie par

$$u_0 = u \qquad \text{et} \qquad \forall n \in \mathbf{N}, \quad u_{n+1} = u_n - Q(u_n)Q'(u_n)^{-1}$$

est bien définie et stationne en d.

Démonstration.
On commence par démontrer un lemme qui contient l'essentiel de la démonstration : il nous dit que notre méthode de Newton est bien définie et converge.

Lemme 0.21.4. *Pour tout $n \in \mathbf{N}$:*

(1) u_n est bien défini ;
(2) $u_n \in k[u]$;
(3) $Q(u_n) \in \left(Q(u)^{2^n}\right)$; (suite d'idéaux décroissants et stationnaires en $\{0\}$)
(4) $Q'(u_n)$ est inversible. (Condition de tangente qui s'annule)

Démonstration. Par récurrence sur $n \in \mathbf{N}$.

Initialisation : $u_0 = u \in k[u]$ d'où (1) et (2) ; $Q(u_0) = Q(u)^{2^0}$ d'où (3).
Ensuite, $\mathrm{pgcd}(Q, Q') = 1$ car Q est scindé à racines simples donc sans facteur carré, donc par théorème de Bézout dans l'anneau principal (euclidien en fait) $k[X]$: il existe $A, B \in k[X]$ tels que

$$BQ'(u_0) = \mathrm{id} - AQ(u_0)$$

et $AQ(u_0)$ est nilpotente donc $BQ'(u_0)$ est inversible par le lemme 2 puis $Q'(u_0)$ inversible d'où (4).

Hérédité : Soit $n \in \mathbf{N}$. On suppose (1), (2), (3) et (4) vraies pour n.

- $Q'(u_n) \in \mathrm{GL}(E)$ donc on a (1).
- $Q'(u_n)^{-1} \in k[Q'(u_n)] \subset k[u]$ en utilisant le lemme 1 et (2) pour n d'où (2) pour $n+1$.
- Par formule de TAYLOR appliquée à Q au point u_n et évaluée en u_{n+1} :

$$\begin{aligned} Q(u_{n+1}) &= \underbrace{Q(u_n) + (u_{n+1} - u_n)Q'(u_n)}_{=0} + (u_{n+1} - u_n)^2 T(u_{n+1}, u_n) \\ &= \left[Q(u_n)Q'(u_n)^{-1}\right]^2 T(u_{n+1} - u_n) \\ &\in \left(Q(u)^{2^{n+1}}\right) \end{aligned}$$

  d'où (3).
- Par une autre formule de TAYLOR appliquée à Q' au point u_n et évaluée en u_{n+1} :

$$\begin{aligned} Q'(u_{n+1}) &= Q'(u_n) + (u_{n+1} - u_n)S(u_{n+1}, u_n) \\ &= \underbrace{Q'(u_n)}_{\text{inversible}} + \underbrace{Q(u_n)}_{\text{nilpotent}} Q'(u_n)^{-1} S(u_{n+1}, u_n) \end{aligned}$$

  donc $Q(u_n)Q'(u_n)^{-1}S(u_{n+1}, u_n)$ est nilpotent car tout commute (que des polynômes en u !) et $Q'(u_{n+1}) \in \mathrm{GL}(E)$ par le lemme 2, soit (4).

□

— L'endomorphisme $Q(u)$ est nilpotent donc il existe $n_0 \in \mathbf{N}$ tel que $Q(u)^{2^{n_0}} = 0$ et alors $Q(u_{n_0}) = 0$ par le point (3) du lemme. Et Q étant scindé à racines simples, $d := u_{n_0}$ est diagonalisable avec $d \in k[u]$.

— On pose

$$\nu := u - d = \sum_{n=0}^{n_0-1} (u_n - u_{n+1}) = \sum_{n=0}^{n_0-1} Q(u_n)Q'(u_n)^{-1}$$

qui est une somme d'endomorphismes nilpotents qui commutent deux-à-deux donc ν est nilpotente et $\nu \in k[u]$.

— $u = d + \nu$ et $\nu, d \in k[u]$ donc ils commutent.

— Si $u = d' + \nu'$ est une décomposition de DUNFORD, alors d',ν' commutent avec u et donc d',ν' commutent avec d et ν, d'où $d' - d = \nu - \nu' = 0$ car à la fois diagonalisable et nilpotente.

□

0.22 Sous-groupes finis de SO(3)

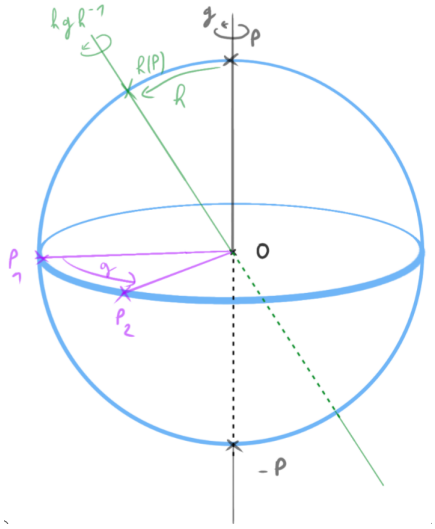
Leçons : 101 ; 104 ; 105 ; 106 ; 154 ; 160 ; 161 ; 190 ; 191

Références :

—

Théorème 0.22.1.
Tout sous-groupe fini de $\mathrm{SO}(3)$ *est soit cyclique, soit un groupe diédral, soit isomorphe à* $\mathfrak{A}_4$, $\mathfrak{S}_4$ *ou* $\mathfrak{A}_5$.

Démonstration.
Soit $G \leqslant \mathrm{SO}(3)$ et $n := |G| \geqslant 2$. La classification des isométries vectorielles montre que $\mathrm{SO}(3)$ ne contient que l'identité et des rotations. Donc tout élément distinct de l'identité a une droite de points fixes qui est l'axe de rotation. Cette axe intersecte la sphère unité en deux pôles opposés.
On note $\mathscr{P}$ l'ensemble des pôles d'éléments de $G \setminus \{\mathrm{id}\}$. On remarque que si P est un pôle de g et $h \in G \setminus \{\mathrm{id}\}$, alors $h(P)$ est un pôle de hgh^{-1}. Donc l'image d'un pôle reste un pôle et $G \curvearrowright \mathscr{P}$ par restriction de l'action naturelle du groupe linéaire. On va décrire G à travers son action sur $\mathscr{P}$.



Étape 1 : Nombre d'orbite.

Soit k le nombre d'orbite. Par formule de Burnside :

$$k = \frac{1}{|G|} \sum_{g \in G} |\operatorname{Fix}(g)| = \frac{1}{n} \left(|\mathscr{P}| + 2(n-1) \right)$$

en séparant l'identité qui fixe tous les pôles et les autres éléments qui ont chacun 2 pôles. De plus,

$$2 \leqslant |\mathscr{P}| \leqslant 2(n-1)$$

car il y a au moins une rotation dans G et la seconde inégalité correspond au cas où toutes les rotations ont des axes distincts. Reportant les inégalités sur $|\mathscr{P}|$ dans la formule de Burnside, on en déduit $2 \leqslant k < 4$ soit $k \in \{2, 3\}$.

Cas $k = 2$: Remplaçant k par 2 dans la formule de Burnside, $|\mathscr{P}| = 2$, autrement dit toutes les rotations ont le même axe de rotation. Donc $\mathscr{P} = (P; -P)$ et G stabilise le plan $(P(-P))^{\perp}$ donc s'identifie à un sous-groupe fini de $\mathrm{SO}(2) \simeq \mathbb{U}$ donc est cyclique.
Cas $k = 3$: Cette fois la formule de Burnside donne $|\mathscr{P}| = 3n - 2n + 2 = n + 2$.

Étape 2 : cardinaux possibles des orbites.

On note $\mathscr{P}_1$, $\mathscr{P}_2$ et $\mathscr{P}_3$ les orbites et les stabilisateurs d'éléments d'une même orbite étant conjugués, ils ont même cardinaux, on note alors m_1, m_2, m_3 les cardinaux des stabilisateurs. Un pôle étant fixé par définition par une rotation de G et par id, on peut supposer, quitte à réordonner les orbite, $2 \leqslant m_1 \leqslant m_2 \leqslant m_3$. Une formule des classes donne :

$$\frac{1}{m_1} + \frac{1}{m_2} + \frac{1}{m_3} = 1 + \frac{2}{n}$$

À partir de cette équation, on peut restreindre les possibilités pour m_1, m_2 et m_3.

- En minorant m_2 et m_3 par m_1 : $\frac{3}{m_1} \geqslant 1 + \frac{2}{n} > 1$ donc $m_1 = 2$ et n est pair.
- On remplace m_1 par 2 et on minore m_3 par m_2 : $\frac{2}{m_2} \geqslant \frac{1}{2} + \frac{2}{n} > \frac{1}{2}$ donc $m_2 \in \{2, 3\}$.
- Si $m_2 = 2$, $\frac{1}{m_3} = \frac{2}{n}$ soit $m_3 = \frac{n}{2}$. Si $m_2 = 3$, $\frac{1}{m_3} = \frac{1}{6} + \frac{2}{n} > \frac{1}{6}$ donc $3 \leqslant m_3 < 6$ soit $m_3 \in \{3, 4, 5\}$.

On résume les possibilités dans le tableau suivant :

m_1	m_2	m_3	n	$\lvert\mathscr{P}_1\rvert$	$\lvert\mathscr{P}_2\rvert$	$\lvert\mathscr{P}_3\rvert$
2	2	$\frac{n}{2}$	–	$\frac{n}{2}$	$\frac{n}{2}$	2
2	3	3	12	6	4	4
2	3	4	24	12	8	6
2	3	5	60	30	20	12

On va traiter ces différents cas dans l'ordre.

Étape 3 : $G \simeq \mathbb{D}_{\frac{n}{2}}$.

Ici n n'est pas entièrement déterminé, on sait juste qu'il est pair. On va avoir besoin de supposer $n \neq 4$ mais puisque les deux seuls groupes d'ordre 4 sont le groupe cyclique et le groupe diédral à 4 éléments, on peut supposer $n \neq 4$.
Alors, si $P \in \mathscr{P}_3$, puisque P est le pôle d'une rotation g si et seulement si $-P$ l'est aussi, $G_P = G_{-P}$ donc les orbites contenant P et $-P$ on le même cardinal. Mais $\mathscr{P}_3$ est la seule orbite a 2 éléments donc $\mathscr{P}_3 = \{-P, P\}$. Ainsi, comme dans le cas $k = 2$, le sous-groupe G_P stabilise le plan $(P(-P))^\perp$ donc s'identifie à un sous-groupe fini de $\mathrm{SO}(2)$ donc est cyclique.
Soit g un générateur de G_p et $x \in \mathscr{P}_1$. La famille

$$\mathrm{Pol} = (g^i(x))_{0 \leqslant i \leqslant \frac{n}{2}-1} \subset \mathscr{P}_1$$

forme un polygone dans le plan $(P(-P))^\perp$ à $\frac{n}{2} = |\mathscr{P}_1|$ sommets donc $|\mathscr{P}_1| = \mathrm{Pol}$. Enfin, puisque $G \curvearrowright \mathscr{P}_1$, G s'identifie à un sous-groupe de $\mathbb{D}_{\frac{n}{2}}$ et par cardinal $G = \mathbb{D}_{\frac{n}{2}}$.

Étape 4 : $G \simeq \mathfrak{A}_4$.

Puisque $|\mathscr{P}_2| = 4$, $G \curvearrowright \mathscr{P}_2$ fidèlement. En effet, une rotation ne fixe que deux pôles, donc un élément fixant les 4 pôles de $\mathscr{P}_2$ est nécessairement l'identité. Donc $G \hookrightarrow \mathfrak{S}_4$ et le seul sous-groupe de $\mathfrak{S}_4$ d'ordre 12 étant $\mathfrak{A}_4$, $G \simeq \mathfrak{A}_4$.

Étape 5 : $G \simeq \mathfrak{S}_4$.

Même raisonnement que le cas précédent, mais on doit raffiner un peu, sinon on obtient simplement une injection de G dans $\mathfrak{S}_6$. Comme dans l'étape 3, l'orbite $\mathscr{P}_2$ est la seule de cardinal 8 donc contient 4 points et leurs opposés. $\mathscr{P}_2 = \{\pm P_i \ ; \ i \in [\![1\,,\,4]\!]\}$ et par linéarité $G \overset{\alpha}{\curvearrowright} \{(P_i, -P_i) \ ; \ i \in [\![1\,,\,4]\!]\}$.
Montrons que cette action est fidèle. Soit $g \in \ker \alpha$. Supposons par l'absurde $g \neq \mathrm{id}$. Cette rotation envoie chaque P_i sur lui-même ou sur son opposé. Dans tous les cas, $\forall i$, $g^2(P_i) = P_i$ donc $g^2 = \mathrm{id}$ car sinon c'est une rotation qui fixe 8 points distincts. Donc g est d'ordre 2. Or $|G_{P_i}| = m_2 = 3$ donc une rotation qui fixe l'un des P_i est d'ordre 3 et g ne peut fixer aucun P_i :

$$\forall i \in [\![1\,,\,4]\!], \qquad g(P_i) = -P_i$$

De plus, pour tout $h \in G$, $hgh^{-1} \neq \mathrm{id}$ car $g \neq \mathrm{id}$ et $hgh^{-1} \in \ker \alpha \triangleleft G$ donc les mêmes arguments montrent que :

$$\forall i \in [\![1\,,\,4]\!], \qquad hgh^{-1}(P_i) = -P_i$$

puis $hgh^{-1}g^{-1} = \mathrm{id}$ car fixe 8 points distincts et $g \in \mathbf{Z}(G) \subset \mathbf{Z}(\mathrm{SO}(3)) = \{\mathrm{id}\}$ d'où l'absurdité.
Ainsi α est injective et par égalité des cardinaux $G \simeq \mathfrak{S}_4$.

Étape 6 : $G \simeq \mathfrak{A}_5$.

La même démonstration qu'à l'étape précédente donne une injection $G \hookrightarrow \mathfrak{S}_6$. Mais $|\mathfrak{S}_6| = 720$ et il resterait à trouver ses sous-groupes d'ordre 60... On peut montrer qu'ils sont tous isomorphes à $\mathfrak{A}_5$ mais ce n'est pas facile. À la place, je vais utiliser le lemme suivant :

Lemme 0.22.2.
Le seul groupe simple d'ordre 60 *est* $\mathfrak{A}_5$ *à isomorphisme près.*

Notre but va donc être de montrer que G est simple à travers son action sur sur les paires de pôles opposés. En effet, les orbites étant de cardinaux distincts, un pôle et son opposé sont dans la même orbite et par linéarité G agit sur les paires de points opposés : $G \curvearrowright \{(P,-P) \mid P \in \mathscr{P}\}$.
Il y a toujours 3 orbites de tailles divisées par 2 car on a rassemblé les points par paires. Mais g stabilise $(P,-P)$ si et seulement si $g(P) = P$ donc $G_{(P,-P)} = G_P$. L'avantage de regarder l'action sur des couples plutôt que sur les pôles individuellement ou les droites est que les stabilisateurs s'intersectent trivialement : une rotation ne peut pas fixer 4 points non alignés.
Tout $g \in G \setminus \{\mathrm{id}\}$ a un axe donc est dans le stabilisateur d'un couple dont les ordres sont 2, 3 et 5 qui sont premiers donc g est d'ordre 2, 3 ou 5.
Ainsi,

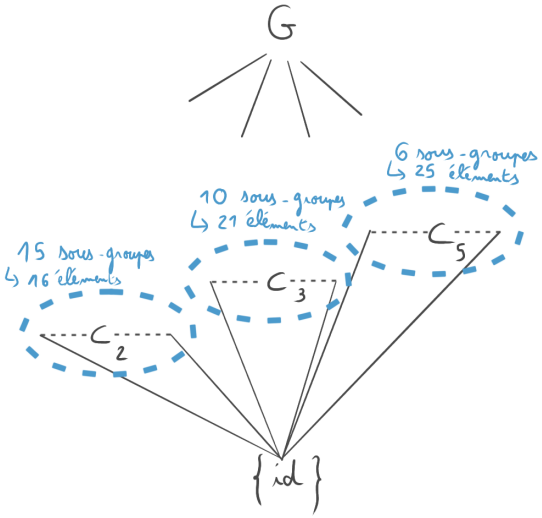
— $\mathscr{P}_1$ donne 15 sous-groupes d'ordre 2 conjugués ;

— $\mathscr{P}_2$ donne 10 sous-groupes d'ordre 3 conjugués ;

— $\mathscr{P}_3$ donne 6 sous-groupes d'ordre 5 conjugués.

Si $H \triangleleft G$, $H \neq \{\mathrm{id}\}$ et si $x \in H$, tous les conjugués de $< g >$ sont dans H car il est distingué donc

- si x est d'ordre 2, H contient 16 au moins éléments (id et un autre élément par sous-groupe d'ordre 2) ;
- si x est d'ordre 3, H contient 21 au moins éléments ;
- si x est d'ordre 5, H contient 25 au moins éléments.



Or $|H| \mid |G| = 60$ donc H contient au moins deux éléments d'ordres distincts et $|H| > 30$ donc $|H| = 60$ ce qui conclut la preuve.

□

0.23 Classification des isométries planes

Leçons : 106 ; 108 ; 151 ; 154 ; 160 ; 161 ; 191

Références :

— [SS88] Yvonne Sortais, René Sortais, *Géométrie del'espace et du plan* pour l'application.

Définition 0.23.1.

1. Rotation plane : on appelle rotation plane de centre A et d'angle θ l'isométrie affine ayant pour unique point fixe A et pour partie linéaire la rotation vectorielle d'angle θ.
2. Symétrie glissée : on appelle symétrie glissée toute composée d'une réflexion d'axe D et d'une translation de vecteur non-nul dirigeant D. Cette composition est alors commutative.

Théorème 0.23.2.
Toute isométrie affine plane peut s'écrire comme produit d'au plus 3 réflexions.

Théorème 0.23.3.

1. *Les déplacements du plan sont l'identité, les translations et les rotations.*
2. *Les anti-déplacements du plan sont les réflexions et les symétries glissées.*
3. *Soit f une isométrie du plan. Alors,*
   - $\dim \mathrm{Fix}(f) = 2 \iff f = \mathrm{id}$;
   - $\dim \mathrm{Fix}(f) = 1 \iff f$ *est une réflexion* ;
   - $\dim \mathrm{Fix}(f) = 0 \iff f$ *est une rotation affine* ;
   - $\mathrm{Fix}(f) = \emptyset \iff f$ *est une translation ou une symétrie glissée.*

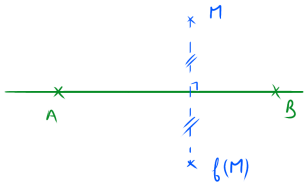
Démonstration.
Soit $f \in \mathrm{Isom}(\mathbf{R}^2)$.

$\dim \mathrm{Fix}(f) = 2$: f préserve tous le plan donc $f = \mathrm{id}$ qui est produit de 0 réflexion.

$\dim \mathrm{Fix}(f) = 1$: $\mathrm{Fix}(f) = (AB)$ et si $M \notin (AB)$,

$$AM = Af(M)$$
$$BM = Bf(M)$$

donc $(AB) = (M : f(M))$ qui dénote la médiatrice du segment $[M, f(M)]$.

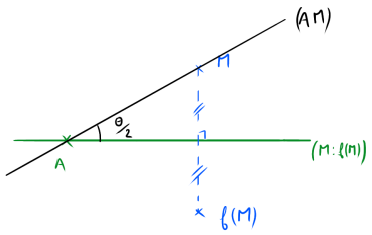


Ainsi, composant par la réflexion d'axe (AB), $\sigma_{(AB)} \circ f$ fixe A, B et M donc son espace de points fixes est de dimension 2 et par le cas précédent $\sigma_{(AB)} \circ f = \mathrm{id}$ soit $f = \sigma_{(AB)}$ donc f est une réflexion.

$\dim \mathrm{Fix}(f) = 0$: $\mathrm{Fix}(f) = \{A\}$ et si $M \neq A$, comme précédemment, puisque f est une isométrie, $A \in (M : f(M))$ et $\sigma_{(M:f(M))} \circ f$ fixe A et M. Donc $d := \dim \mathrm{Fix}(\sigma_{(M:f(M))} \circ f) \geqslant 1$.

— $d = 2$: on a alors $\sigma_{(M:f(M))} \circ f = \mathrm{id}$ soit $f = \sigma_{(M:f(M))}$, absurde car f a un unique point fixe.

— $d = 1$: par le cas précédent $\sigma_{(M:f(M))} \circ f = \sigma_{(AM)}$ donc $f = \sigma_{(M:f(M))} \circ \sigma_{(AM)}$ et f est produit de 2 réflexions.



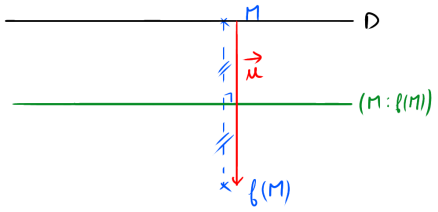
En particulier, $\det \overrightarrow{f} = 1$ et a un unique point fixe donc f est une rotation.

On en déduit aussi qu'une rotation est le produit de deux réflexions d'axes sécants, l'une d'entre elle pouvant être choisie arbitrairement (car M est arbitraire).

Fix$(f) = \emptyset$: soit $M \in \mathbf{R}^2$. L'isométrie $g := \sigma_{(M:f(M))} \circ f$ fixe M donc vérifie $\mathrm{Fix}(f) \neq \emptyset$. On note $d := \dim \mathrm{Fix}(f)$.

— $d = 2$: g est l'identité donc $f = \sigma_{(M:f(M))}$, absurde car f n'a pas de point fixe.

— $d = 1$: g est une réflexion et il existe une droite D contenant M telle que $f = \sigma_{(M:f(M))} \circ \sigma_D$.
Si $D \cap (M : f(M)) \neq \emptyset$, f a un point fixe, impossible. Donc $D /\!\!/ (M : f(M))$ et $\overrightarrow{f} = \mathrm{id}$ car les parties linéaires des deux réflexions sont les mêmes et d'ordre 2 d'où $f = \tau_{\overrightarrow{u}}$ est une translation.

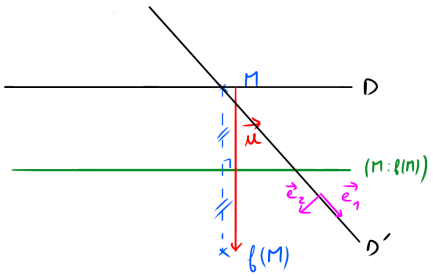


Réciproquement, toute translation ne fixe aucun point et vérifie $d = 1$ donc on en déduit que toute translation est le produit de deux réflexions d'axes parallèles, l'une d'entre elle pouvant être choisie arbitrairement.

— $d = 0$: $\mathrm{Fix}(f) = M$ et par le cas précédent g est une rotation de centre M que l'on peut écrire comme le produit de deux réflexions d'axes sécants, l'un des deux choisi arbitrairement. On choisit $M \in D /\!\!/ (M : f(M))$ et D' telles que $g = \sigma_D \circ \sigma_{D'}$. Donc

$$f = \sigma_{(M:f(M))} \circ \sigma_D \circ \sigma_{D'} = \tau_{\overrightarrow{u}}.$$

On a donc décomposé f en produit de 3 réflexions, ce qui achève la démonstration du premier théorème. On a pas encore une symétrie glissée car $\overrightarrow{u}$ n'est pas de même direction que D' a priori.



Soit $\overrightarrow{e_1}$ un vecteur directeur de D' et $\overrightarrow{e_2}$ orthogonal et non-nul à $\overrightarrow{e_1}$. Alors, $(\overrightarrow{e_1}, \overrightarrow{e_2})$ est une base de $\overrightarrow{E}$. On décompose $\overrightarrow{u}$ dans cette base : $\overrightarrow{u} = \lambda\overrightarrow{e_1} + \mu\overrightarrow{e_2}$. Alors,

$$f = \tau_{\lambda\overrightarrow{e_1}} \circ \tau_{\mu\overrightarrow{e_2}} \circ \sigma_{D'} = \tau_{\lambda\overrightarrow{e_1}} \circ \sigma_{D' + \frac{\mu}{2}\overrightarrow{e_2}} \tag{13}$$

et f est une symétrie glissée.

□

Détail de la dernière égalité de 13 :

$$\tau_{\mu\overrightarrow{e_2}} \circ \sigma_{D'}(N) = N \iff \sigma_{D'}(N) = \tau_{-\mu\overrightarrow{e_2}}(N) \iff D' = (N : N - \mu\overrightarrow{e_2}) \iff N \in D' + \frac{1}{2}\mu\overrightarrow{e_2}$$

Théorème 0.23.4.
Soit Δ_1, Δ_2 et Δ_3 trois droites concourantes. Il existe un unique triangle ABC pour lequel $\Delta_1 = (A : B)$, $\Delta_2 = (B : C)$ et $\Delta_3 = (C : A)$ à homothétie près.

Démonstration.
Dans tout triangle, les médiatrices des côtés sont concourantes en un point. On cherche une réciproque qui dit que les médiatrices caractérisent le triangle à homothétie près. On note O le point d'intersection de Δ_1, Δ_2 et Δ_3.

Analyse

Soit ABC un triangle solution. Les points A, B, C sont différents de O. Et

$$\begin{array}{ccccccc} A & \overset{\sigma_{\Delta_1}}{\mapsto} & B & \overset{\sigma_{\Delta_2}}{\mapsto} & C & \overset{\sigma_{\Delta_3}}{\mapsto} & A \\ O & \mapsto & O & \mapsto & O & \mapsto & O \end{array}$$

donc $\sigma_{\Delta_3} \circ \sigma_{\Delta_2} \circ \sigma_{\Delta_1}$ est un antidéplacement car produit d'un nombre impair de réflexions et fixe au moins une droite donc, par la classification précédente, c' est la réflexion $\sigma_{(OA)}$. Donc A est sur l'axe de la réflexion $\sigma_{\Delta_3} \circ \sigma_{\Delta_2} \circ \sigma_{\Delta_1}$. Une fois A fixé, B et C sont entièrement déterminés comme image de A par σ_{Δ_1} et $\sigma_{\Delta_2} \circ \sigma_{\Delta_1}$. S'il y a une solution, elle est unique à homothétie près.

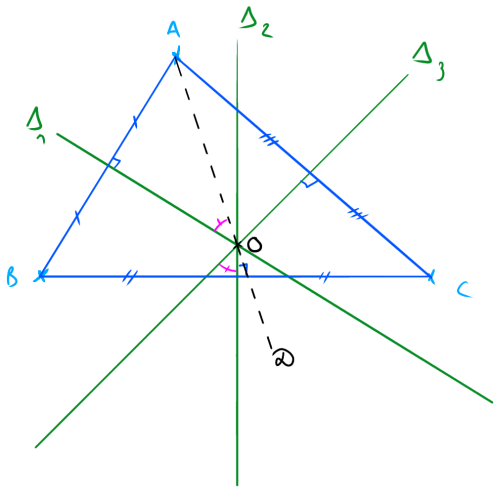
Synthèse

Étape 1 : Construction de l'axe $\mathscr{D}$ de $\sigma_{\mathscr{D}} = \sigma_{\Delta_3} \circ \sigma_{\Delta_2} \circ \sigma_{\Delta_1}$.

On sait que $\mathscr{D}$ contient O donc il suffit d'en trouver un vecteur directeur. Soit $\overrightarrow{u_1}, \overrightarrow{u_2}, \overrightarrow{u_3}, \overrightarrow{u}$ des vecteurs directeurs respectivement de Δ_1, Δ_2, Δ_3 et $\mathscr{D}$. On a :

$$\sigma_{\Delta_3} \circ \sigma_{\Delta_2} = \sigma_{\mathscr{D}} \circ \sigma_{\Delta_1}$$

Or $\begin{array}{l} \sigma_{\Delta_3} \circ \sigma_{\Delta_2} \text{ est la rotation de centre } O, \text{ d'angle } 2\overline{(\overrightarrow{u_2}, \overrightarrow{u_3})} \\ \sigma_{\mathscr{D}} \circ \sigma_{\Delta_1} \text{ est la rotation de centre } O, \text{ d'angle } 2\overline{(\overrightarrow{u_1}, \overrightarrow{u})} \end{array}$ donc $\overline{(\overrightarrow{u_1}, \overrightarrow{u})} = \overline{(\overrightarrow{u_2}, \overrightarrow{u_3})} + k\pi$ ce qui détermine la direction de $\mathscr{D}$.



Étape 2 : Construction d'un triangle solution.

Soit $A \in \mathscr{D}$, $A \neq O$ quelconque, puis $B = \sigma_{\Delta_1}(A)$ et $C = \sigma_{\Delta_2}(B)$. Il reste à vérifier que $\Delta_3 = (A : C)$, soit $\sigma_{\Delta_3}(A) = C$.
Or,

$$\underbrace{\sigma_{\Delta_2} \circ \sigma_{\Delta_1}(A)}_{=C} = \sigma_{\Delta_3} \circ \underbrace{\sigma_{\mathscr{D}}(A)}_{=A} = \sigma_{\Delta_3}(A)$$

car $A \in \mathscr{D}$. Donc ABC construit est solution. □

0.24 Théorème de Carathéodory et équations diophantiennes

Leçons : 126 ; 151 ; 162 ; 181

Références :

— Clarence Kineider, `http://perso.eleves.ens-rennes.fr/people/clarence.kineider/agreg.html`

Théorème 0.24.1.
Soit A une partie de $\mathbf{R}^n$, on a :

$$\mathrm{Conv}(A) = \left\{ \sum_{i=1}^{n+1} \lambda_i a_i \mid a_i \in A,\ \lambda_i \geqslant 0,\ \sum_{i=1}^{n+1} \lambda_i = 1 \right\}$$

Démonstration.
L'inclusion de doite à gauche est claire.

Si $x \in \mathrm{Conv}(A)$, on peut donc écrire x comme combinaison convexe d'éléments de A :

$$x = \sum_{i=1}^{p} \lambda_i a_i \qquad \text{où } a_i \in A,\ \lambda_i \geqslant 0,\ \sum_{i=1}^{p} \lambda_i = 1,$$

avec p minimal. Supposons par l'absurde $p > n+1$. Notre but est de modifier légèrement les λ_i de sorte à garder une combinaison convexe qui vaut x mais dans laquelle l'un des coefficients s'annule.
On pose

$$\varphi : \left| \begin{array}{ccc} \mathbf{R}^p & \to & \mathbf{R}^n \times \mathbf{R} \\ (\nu_i)_{1\leqslant i\leqslant n} & \mapsto & (\sum_{i=1}^p \nu_i a_i, \sum_{i=1}^p \nu_i) \end{array} \right.$$

qui est linéaire donc par théorème du rang :

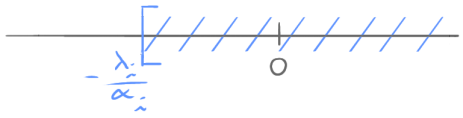
$$\dim \ker \varphi = p - \mathrm{rg}\, \varphi \geqslant p - (n+1) \geqslant 1$$

Cette argument de dimension nous dit que φ qui va d'un espace vectoriel dans un espace de dimension plus petite ne peut être injective. Soit $\alpha = (\alpha_i)_{1\leqslant i\leqslant p} \in \ker \varphi \setminus \{0\}$.
On pose $\mu_i(t) = \lambda_i + t\alpha_i$ pour $i \in [\![1\,,\,p]\!]$ et $t \in \mathbf{R}$. On a :

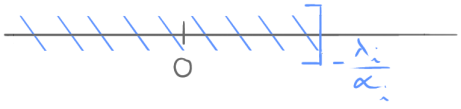
$$\sum_{i=1}^{p} \mu_i(t) a_i = x + t\sum_{i=1}^{p} \alpha_i a_i = x \qquad \text{et} \qquad \sum_{i=1}^{p} \mu_i(t) = \sum_{i=1}^{p} \lambda_i + t\sum_{i=1}^{p} \alpha_i = 1$$

On doit maintenant choisir le paramètre t pour que les $\mu_i(t)$ soient positifs et annuler un des coefficients.

- Si $\alpha_i = 0$: $\mu_i(t) = \lambda_i \geqslant 0$ pour tout $t \in \mathbf{R}$.
- Si $\alpha_i > 0$: il faut que $t \geqslant -\frac{\lambda_i}{\alpha_i}$.



- Si $\alpha_i < 0$: il faut que $t \leqslant -\frac{\lambda_i}{\alpha_i}$.



Puisque $\alpha \neq 0$ et $\sum_{i=1}^p \alpha_i = 0$, il existe j, k tels que $\alpha_j < 0$ et $\alpha_k > 0$. On pose alors :

$$\tau = \max \left\{ -\frac{\lambda_i}{\alpha_i} \mid 1 \leqslant i \leqslant p,\ \alpha_i > 0 \right\} < 0$$

qui vérifie $\mu_i(\tau) \geqslant 0$ pour tout $i \in [\![1\,,\,p]\!]$ et $\mu_{i_0} = 0$ pour l'indice i_0 en lequel le maximum est atteint. D'où

$$x = \sum_{i=1}^{p} \mu_i(\tau) a_i = \sum_{i \neq i_0} \mu_i(\tau) a_i$$

ce qui contredit la minimalité de p et conclut. □

Proposition 0.24.2.
Soit $m, n \geqslant 1$ et $A \in \mathscr{M}_{m,n}(\mathbf{Z})$. Il y a équivalence entre :

(i) *$Ax = 0$ admet une solution non-nulle $x \in \mathbf{N}^n$;*

(ii) *$0_{\mathbf{R}^m} \in \mathrm{Conv}(A_1, \dots, A_n)$ où $A_1, \dots, A_n$ sont les colonnes de A.*

Démonstration.
(i) $\Longrightarrow$ (ii) : Soit $x \in \mathbf{N}^n$ non-nul tel que $Ax = 0 = \sum_{i=1}^n x_i A_i$. Notant $\lambda = \sum_{i=1}^n x_i$, on a

$$\sum_{i=1}^n \frac{x_i}{\lambda} A_i = 0_{\mathbf{R}^n} \qquad \text{avec} \sum_{i=1}^n \frac{x_i}{\lambda} = 1.$$

(ii) $\Longrightarrow$ (i) : Soit l minimal tel que $0_{\mathbf{R}^m} \in \mathrm{Conv}(A_{i_1}, \dots, A_{i_l})$ et soit $r := \mathrm{rg}(A_{i_1}, \dots, A_{i_l})$.
On a $r < l$ car la famille est liée (il existe une combinaison convexe donc a fortiori linéaire non triviale donnant 0). Et par théorème de Carathéodory : $l \leqslant r + 1$ donc finalement $l = r + 1$.
Par algorithme du pivot de Gauss, il existe $P \in \mathrm{GL}_n(\mathbf{Z})$ telle que :

$$P \begin{pmatrix} A_{i_1} \vdots \cdots \vdots A_{i_{r+1}} \end{pmatrix} = \begin{pmatrix} M \\ \dots \\ 0_{1,r+1} \end{pmatrix}$$

avec $M \in \mathscr{M}_{r,r+1}(\mathbf{Z})$ de rang r sur $\mathbf{R}$ mais aussi sur $\mathbf{Q}$ puisque le rang est invariant par extension de corps. Par théorème du rang, son noyau est de dimension 1 sur $\mathbf{Q}$ et sur $\mathbf{R}$. Or il existe $x \in (\mathbf{R}_+^*)^{r+1}$ tel que

$$\begin{pmatrix} A_{i_1} \vdots \cdots \vdots A_{i_{r+1}} \end{pmatrix} x = 0$$

Comme $\ker_{\mathbf{Q}} M \subset \ker_{\mathbf{R}} M$, si $y \in \ker_{\mathbf{Q}} M \setminus \{0\}$, il existe $\lambda \in \mathbf{R}^*$ tel que $y = \lambda x$ et on peut considérer que $y_i \in \mathbf{N}$ pour tout $i \in [\![1\,,\, r+1]\!]$.
On pose alors $\tilde{y} = (\tilde{y}_i) \in \mathbf{R}^n$ tel que $\tilde{y}_{i_k} = y_k$ et $\tilde{y}_i = 0$ sinon. Puisque $My = 0$, $A\tilde{y} = 0$ avec $\tilde{y} \in (\mathbf{N}^n)^*$. □

Autre application qui sert pour montrer que $O_n(\mathbf{R})$ est l'unique sous-groupe maximal de $\mathrm{GL}_n(\mathbf{R})$ à conjugaison près :

Proposition 0.24.3.
Si K est une partie compacte d'un espace vectoriel E de dimension finie, alors $\mathrm{Conv}(K)$ est encore compacte.

Démonstration.
Notons $n := \dim E$ et posons

$$L := \left\{ (\lambda_1, \dots, \lambda_{n+1}) \in [0,1]^{n+1} \mid \sum_{i=1}^{n+1} \lambda_i = 1 \right\}$$

partie compacte de $\mathbf{R}^{n+1}$ (fermée et bornée en dimension finie). Alors, par théorème de Carathéodory,

$$\varphi : \left| \begin{array}{ccc} K^{n+1} \times L & \to & \mathrm{Conv}(X) \\ ((x_i), (\lambda_i)) & \mapsto & \sum_{i=1}^n \lambda_i x_i \end{array} \right.$$

est surjective, avec $K^{n+1} \times L$ compact d'où $\mathrm{Conv}(K)$ compacte. □

0.25 Disques de Gerschgorin

Leçons : 144 ; 149 ; 191 ; 204 ; 228 ; 267

Références :

— [Ser02] Denis Serre, *Matrices : Theory and Applications*

Théorème 0.25.1.
Soit $A \in \mathscr{M}_n(\mathbf{C})$. Pour $i \in [\![1\,,\,n]\!]$, on note

- $D_i := D\left(a_{ii}, \sum_{j\neq i} |a_{ij}|\right)$ *le i-ème disque de* Gerschgorin ;
- $\mathscr{D} := \bigcup_{i=1}^n D_i$ *le domaine de* Gerschgorin.

Alors $\sigma(A) \subset \mathscr{D}$.

Démonstration.
Soient $\lambda \in \sigma(A)$ et x un vecteur propre associé. Soit $i \in [\![1\,,\,n]\!]$ tel que $|x_i| = \|x\|_\infty$. Alors $x_i \neq 0$ car x est non-nul en tant que vecteur propre et la i-ème ligne de $(A - \lambda I_n)x = 0$ donne :

$$(a_{ii} - \lambda)x_i + \sum_{j\neq i} a_{ij}x_j = 0$$

puis divisant par x_i et prenant la valeur absolue

$$|a_{ii} - \lambda| = \left|\sum_{j\neq i} a_{ij}\frac{x_j}{x_i}\right| \leqslant \sum_{j\neq i} |a_{ij}|$$

ce qui conclut. □

Théorème 0.25.2.
Soit C une composante connexe de $\mathscr{D}$. Il existe alors $I \subset [\![1\,,\,n]\!]$ tel que $C = \bigcup_{i\in I} D_i$ On note alors $p := |I|$ et $m(B)$ le nombre de valeurs propres de $B \in \mathscr{M}_n(\mathbf{C})$ dans C comptées avec multiplicité.
Alors $m(A) = p$.

Démonstration.
Soit $t \in [0,1]$. On va utiliser un argument d'homotopie en exploitant le fait qu'il est très facile de localiser les valeurs propres d'une matrice diagonale... On note :

- $A(t) = (1-t)\mathrm{diag}(a_{11}, \cdots, a_{nn}) + tA$;
- $\mathscr{D}_t$ le domaine de Gerschgorin de $A(t)$;
- χ_t son polynôme caractéristique ;
- $m(t)$ pour $m(A(t))$.

Alors on remarque que

- $A(0) = \mathrm{diag}(a_{11}, \cdots, a_{nn})$;
- $A(1) = A$;
- $\mathscr{D}_t \subset \mathscr{D}$;
- $m(0) = p$ car $a_{jj} \in C \iff D_j \subset C$.

Montrons que l'application m est continue. Comme C et $\mathscr{D}\backslash C$ sont compacts, on peut trouver une courbe de Jordan Γ les séparant et entourant positivement C. Puis, comme $\mathscr{D}_t \subset \mathscr{D}$, χ_t ne s'annule par sur Γ et est entière car polynomiale donc par principe de l'argument :

$$m(t) = \frac{1}{2\mathrm{i}\pi}\int_\Gamma \frac{\chi_t'(z)}{\chi_t(z)}\mathrm{d}z.$$

Or $t \mapsto A(t)$ est continue, puis $t \mapsto \chi_t$ et $t \mapsto \chi_t'$ aussi d'où m continue par continuité sous le signe intégral. Et $[0,1]$ étant connexe, m qui est continue sur un connexe et à valeurs dans $\mathbf{N}$, est constante et $m(A) = m(1) = m(0) = p$. □

0.26 Automorphismes de $\mathfrak{S}_n$

Leçons : 101 ; 103 ; 104 ; 105 ; 108 ; 190

Références :

—

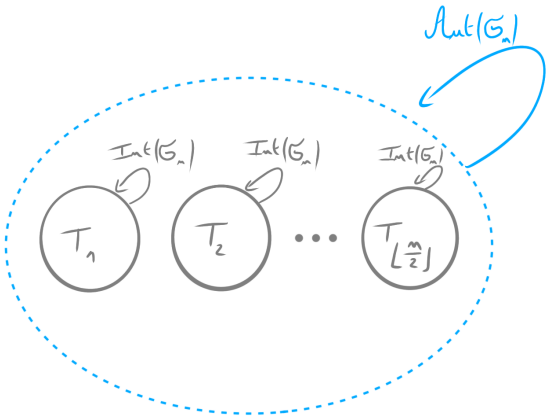
Théorème 0.26.1.
Si $n \neq 6$, $\mathrm{Aut}(\mathfrak{S}_n) = \mathrm{Int}(\mathfrak{S}_n)$.

Démonstration.
Puisqu'un automorphisme envoie une classe de conjugaison sur une autre, on a $\mathrm{Aut}(\mathfrak{S}_n) \curvearrowright \mathrm{Conj}(\mathfrak{S}_n)$. Pour $k \in [\![1\,,\, \lfloor \frac{n}{2} \rfloor]\!]$, on note $T_k = \{\sigma \in \mathfrak{S}_n \mid a(\sigma) = (k, 0, \dots, 0)\}$, c'est-à-dire l'ensemble des k-transpositions. Ce sont des éléments d'un même type donc des classes de conjugaison. De plus,

— un automorphisme préserve les ordres ;

— $\sigma \in \mathfrak{S}$ est d'ordre 2 $\iff$ $\sigma \in \bigcup_k T_k$.

Donc l'action se restreint à $\mathrm{Aut}(\mathfrak{S}_n) \curvearrowright \left\{T_1, \dots, T_{\lfloor \frac{n}{2} \rfloor}\right\}$.



On remarque que par définition de classe de conjugaison, les automorphismes intérieurs fixent tous les T_k. Réciproquement, nous allons démontrer et utiliser le lemme suivant :

Lemme 0.26.2.
Soit $\varphi \in \mathrm{Aut}(\mathfrak{S}_n)$. *Si* $\varphi(T_1) = T_1$, *alors* $\varphi \in \mathrm{Int}(\mathfrak{S}_n)$.

Démonstration.
Supposons $\varphi(T_1) = T_1$. On note $T = \{\tau_i := (1 \quad i) \mid i \in [\![2\,,\, n]\!]\} \subset T_1$. Alors $\mathfrak{S}_n =< T >$. Donc si φ coïncide avec un morphisme intérieur sur T, le lemme est démontré.
Le résultat est évident pour $n = 2$, on suppose $n \geqslant 3$.
Si $\mathrm{supp}\,(\varphi(\tau_2)) \cap \mathrm{supp}\,(\varphi(\tau_3)) = \emptyset$, $\varphi(\tau_2\tau_3) = \varphi(\tau_2)\varphi(\tau_3) \in T_2$ est d'ordre 2 car l'ordre est le ppcm des longueurs des cycles dans la décomposition en cycles à supports disjoints mais $\tau_2\tau_3 = (1 \quad 3 \quad 2)$ est d'ordre 3, ce qui est absurde.
Donc, $\begin{matrix}\varphi(\tau_2) = (a_1 \quad a_2) \\ \varphi(\tau_3) = (a_1 \quad a_3)\end{matrix}$ avec $a_2 \neq a_3$ par injectivité de φ. On va montrer que les images des τ_i permutent toutes a_1 avec un autre élément. Si $n = 3$, c'est bon. Supposons $n \geqslant 4$ et soit $i \in [\![4\,,\, n]\!]$.
On a $\begin{matrix}\tau_2\tau_i = (1 \quad i \quad 2) \\ \tau_3\tau_i = (1 \quad i \quad 3)\end{matrix}$ donc par le même argument que précédemment, $\varphi(\tau_i)$ a un support qui intersecte ceux de $\varphi(\tau_2)$ et de $\varphi(\tau_3)$ et soit $\varphi(\tau_i) = (a_1 \quad a_i)$ soit $\varphi(\tau_i) = (a_2 \quad a_3)$ mais alors

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

et $\tau_2\tau_i\tau_3 = (1 \quad 3 \quad i \quad 2)$, impossible par préservation des ordres par un automorphisme. Donc

$$\forall i \in [\![2\,,\, n]\!], \qquad \varphi(\tau_i) = (a_1 \quad a_i) \text{ où les } a_i \text{ sont distincts.}$$

On pose alors $\alpha \in \mathfrak{S}_n$ tel que $\alpha(i) = a_i$, et on obtient :

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

d'où $\varphi \in \mathrm{Int}(\mathfrak{S}_n)$. □

Soit $\varphi \in \mathrm{Aut}(\mathfrak{S}_n)$. Soit k tel que $\varphi(T_1) = T_k$. Par le lemme, il suffit de montrer que $n = 1$. En particulier, $|T_k| = |T_1|$ mais

$$|T_k| = \frac{\binom{n}{2}\binom{n-2}{2}\dots\binom{n-2k+2}{2}}{k!} = \frac{\frac{n!}{2(n-2)!}\frac{(n-2)!}{2(n-4)!}\cdots\frac{(n-2k+2)!}{2(n-2k)!}}{k!} = \frac{n(n-1)\dots(n-2k+1)}{2^k k!}$$

car le choix d'une k-transposition correspond au choix de k couples disjoints et le produit étant commutatif, il faut diviser par le nombre de façons d'arranger k transpositions. Donc on doit avoir :

$$\frac{n(n-1)\dots(n-2k+1)}{2^k k!} = \frac{n(n-1)}{2} \quad \textit{i.e.} \quad (n-2)(n-3)\dots(n-2k+1) = 2^{k-1}k!$$

C'est évidemment vrai pour $k = 1$...

- Si $k = 2$: $(n-2)(n-3) = 4$ or $n-2$ ou $n-3$ est impair donc impossible.
- Si $k > 3$: $(n-2)(n-3)\dots(n-k+1)\binom{n-k}{k} = 2^{k-1}$ et l'un des facteurs à gauche est impair donc impossible.
- Si $k = 3$: $(n-2)(n-3)(n-4) = 2^3 \times 3$, vérifiée si $n = 6$ uniquement car le membre de gauche est une fonction strictement croissante de n.

Donc si $n \neq 6$, $\varphi(T_1) = T_1$ et $\varphi \in \mathrm{Int}(\mathfrak{S}_n)$ par le lemme. □

0.27 Transformée de Fourier discrète et FFT

Leçons : 102 ; 120 ; 162

Références :

— [Pey12] Jacques Peyrière, *Convolution, séries et intégrales de Fourier*

Soit K un corps tel que $\mu_n(K) = V(X^n - 1)$ soit de cardinal n.

Remarques :

- On a $|\mu_n(K)| \leqslant n$ car $X^n - 1 \in K[X]$ a au plus n racines.
- Si $\text{car}(K) = p > 0$ et $p \mid n$, alors $n = pm$ et $X^n - 1 = (X^m - 1)^p$ donc $\mu_n(K) \subset \mu_m(K)$ et le cardinal est au plus m.
- En pratique, on choisi pour n une puissance de 2 donc seule la caractéristique 2 pose problème.

Le but de ce développement est de proposer un algorithme efficace de multiplication de polynômes de $K[X]$ qui repose sur la transformée de Fourier discrète et l'algorithme FFT du type "diviser pour régner".

Théorème 0.27.1.
Soit $n \in \mathbf{N}^*$ une puissance de 2. Si $P = \sum_{i=0}^{n-1} a_i X^i$ et $Q = \sum_{i=0}^{n-1} b_i X^i$ sont deux polynômes de $K[X]$ de degrés $< n/2$, alors, notant $PQ = \sum_{i=0}^{n} c_i X^i$, on peut calculer les c_i en $O(n \ln n)$ opérations dans K.

Autrement dit, on cherche à calculer la décomposition dans la base canonique de PQ. Mais la base canonique n'est pas pratique pour la multiplication : un algorithme naïf demande $O(n^2)$ opérations pour chaque c_i. On va ici utiliser la transformée de Fourier discrète pour écrire les polynômes P et Q dans une base qui permet d'exprimer facilement le produit.

L'interpolation de Lagrange nous dit que P,Q et PQ sont uniquement déterminés par leurs valeurs sur un n-uplet de points distincts. L'avantage de la base de Lagrange c'est que les coefficients du produit PQ sont très simples à calculer : ce sont juste les produits de ceux de P et de Q.

Il reste à choisir le n-uplet de points d'évaluation. L'idée de l'algorithme FFT est de prendre les racines n-ièmes de l'unité pour exploiter certaines symétries.

Démonstration.
Soit ω une racine primitive de l'unité , c'est-à-dire un générateur de $\mu_n(K)$.
Si $R \in K[X]$ est multiple de $X^n - 1$, alors $R(\omega^i) = 0$ pour $i \in [\![0, n-1]\!]$, on peut donc poser le morphisme de K-algèbres :

$$\begin{array}{rccc} DFT_\omega : & K[X]/(X^n-1) & \to & K^n \\ & \bar{P} & \mapsto & \big(P(1), \dots, P(\omega^{n-1})\big) \end{array}$$

Remarques :

- La multiplication sur K^n est le produit terme à terme
- C'est un morphisme de K-algèbres car l'évaluation d'un polynôme l'est.
- Morphisme de K-algèbre = qui respecte les polynômes.
- L'application DFT_ω est la traduction de la transformée de Fourier sur le groupe localement compact et abélien $\mathbf{Z}/n\mathbf{Z}$ dont le dual peut être identifier à $\mu_n(K)$ par le choix d'une racine primitive de l'unité (ici ω) :

$$\begin{array}{ccc} (\mathscr{A}(\mathbf{Z}/n\mathbf{Z}, K), \star) & \xrightarrow{\sim} & K[X]/(X^n-1) \\ \delta_a & \mapsto & X^a \end{array} \quad \left| \quad \begin{array}{ccc} (\mathscr{A}(\mathbf{Z}/n\mathbf{Z}, K), \cdot) = K^{\mu_n(K)} & \xrightarrow{\sim} & K^n \\ f & \mapsto & \big(f(1), \dots, f(\omega^{n-1})\big) \end{array} \right.$$

Lemme 0.27.2.
L'application DFT_ω est un isomorphisme de K-algèbres. Avec pour base $(\bar{1}, \bar{X}, \dots, \bar{X}^{n-1})$ au départ et la base canonique de $\mathbf{C}^n$ à l'arrivée, sa matrice est la matrice de Vandermonde :

$$V(\omega) = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ 1 & \omega & \cdots & \omega^{n-1} \\ \vdots & \vdots & & \vdots \\ 1 & \omega^{n-1} & \cdots & \omega^{(n-1)^2} \end{pmatrix}$$

De plus, $DFT_\omega^{-1} = \frac{1}{n} DFT_{\omega^{-1}}$.

Démonstration du lemme.

Si $i \in [\![0\,,\,n-1]\!]$, $DFT_\omega(\bar{X}^i) = (1, \omega^i, \dots, \omega^{(n-1)i})$ qui est la i-ième colonne de $V(\omega)$ d'où la matrice de DFT_ω. Montrons que $A := V(\omega)V(\omega^{-1}) = nI_n$. Pour $0 \leqslant i, j \leqslant n-1$, on calcule le coefficient (i,j) de A :

$$a_{i,j} = \sum_{k=0}^{n-1} \omega^{ik}\omega^{-kj} = \sum_{k=0}^{n-1} \omega^{(i-j)} = n\delta_{i,j}$$

En effet, si $i = j$, on somme n fois 1, sinon $\omega^{(i-j)} \neq 1$ donc $a_{i,j} = \frac{\omega^{(i-j)n}-1}{\omega^{i-j}-1} = 0$. □

$DFT_\omega^{-1} = \frac{1}{n} DFT_{\omega^{-1}}$ est la formule d'inversion de Fourier dans notre cas. Elle nous dit que les problèmes d'évaluation et d'interpolation sont en fait le même problème. Si on a un algorithme efficace pour écrire un polynôme dans la base de Lagrange (évaluation), on dispose également d'un algorithme efficace pour revenir à la représentation dans la base canonique (interpolation).

Algorithme FFT

On veut calculer

$$DFT_\omega(P) = V(\omega) \begin{pmatrix} a_0 \\ \vdots \\ a_{n-1} \end{pmatrix}$$

avec n une puissance de 2. Puisque $1 = \omega^n - 1 = (\omega^{\frac{n}{2}} + 1)(\omega^{\frac{n}{2}} - 1)$ et que $\omega^{\frac{n}{2}} \neq 1$, $\omega^{\frac{n}{2}} = -1$. On remarque alors que :

$$V(\omega) = \begin{pmatrix} \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \\ \vdots & 1 & \omega^i & \omega^{2i} & \omega^{3i} & \omega^{4i} & \cdots & \omega^{(n-1)i} & \vdots \\ \vdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & & \vdots \\ \vdots & 1 & -\omega^i & \omega^{2i} & -\omega^{3i} & \omega^{4i} & \cdots & -\omega^{(n-1)i} & \vdots \\ & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \end{pmatrix} \begin{matrix} \\ \text{ligne } i \in [\![1\,,\,n/2]\!] \\ \\ \\ \text{ligne } i + \frac{n}{2} \\ \\ \end{matrix}$$

Donc la deuxième moitié des lignes se déduit de la première : les termes d'ordre pair sont les mêmes et ceux d'ordre impair diffèrent sont opposés. On pose alors $Q_1 = \sum_{i=0}^{n/2} a_{2i}X^i$ et $Q_2 = \sum_{i=0}^{n/2} a_{2i+1}X^i$.

On remarque que ω^2 est une racine primitive $\frac{n}{2}$-ème de l'unité et que et on calcule :

$$\hat{Q}_1 = DFT_{\omega^2}(Q_1) = \left(\sum_{j=0}^{n/2-1} a_{2j}\omega^{2ij} \right)_{0 \leqslant i \leqslant \frac{n}{2}-1} \quad \text{et} \quad \hat{Q}_2 = DFT_{\omega^2}(Q_2) = \left(\sum_{j=0}^{n/2-1} a_{2j+1}\omega^{2ij} \right)_{0 \leqslant i \leqslant \frac{n}{2}-1}$$

Alors, $\hat{Q}_1$ donnera les termes de puissances pairs et $\hat{Q}_2$ ceux de puissance impairs et pour $i \in [\![0\,,\,\frac{n}{2}-1]\!]$,

$$DFT_\omega(P)_i = (\hat{Q}_1)_i + \omega^i(\hat{Q}_2)_i$$
$$DFT_\omega(P)_{i+\frac{n}{2}} = (\hat{Q}_1)_i - \omega^i(\hat{Q}_2)_i$$

d'où l'algorithme suivant :

Algorithm 1 Algorithme FFT

Require: $P = \sum_{i=0}^{n-1} a_i X^i$
1: **if** $n = 1$ **then**
2: **return** a_0
3: **end if**
4: $Q_1 \leftarrow \sum_{i=0}^{n/2} a_{2i}X^i$; $Q_2 \leftarrow \sum_{i=0}^{n/2} a_{2i+1}X^i$
5: $u \leftarrow DFT_{\omega^2}(Q_1)$; $v \leftarrow DFT_{\omega^2}(Q_2)$
6: **return** $\left(u_1 + v_1, u_2 + \omega v_2, \dots, u_{\frac{n}{2}} + \omega^{\frac{n}{2}} v_{\frac{n}{2}}, u_1 - v_1, u_2 - \omega v_2, \dots, u_{\frac{n}{2}} - \omega^{\frac{n}{2}} v_{\frac{n}{2}}\right)$

Complexité

On note $C(m)$ la complexité (c'est-à-dire le nombre d'opérations dans K) pour $n = 2^m$. Alors, si $m \neq 0$,

$$C(m) = 2C(m-1) + 2 \times 2^m$$

donc en posant $T(m) = \frac{C(m)}{2^m}$, on a $T(m) = T(m-1) + 2$ pour tout $n \in \mathbf{N}^*$ soit $T(m) = 2m + T(0) = 2m$. Finalement, $C(m) = 2^m \times 2m = O(n \ln n)$. □

0.28 Théorème de Kronecker

Leçons : 102 ; 105 ; 141 ; 144 ; 152

Références :

—

Théorème 0.28.1.
Tout polynôme $P \in \mathbf{Z}[X]$ unitaire et dont les racines complexes sont toutes dans $\overline{D} := \{z \in \mathbf{C} \mid |z| \leqslant 1\}$ est produit d'une puissance de X et de polynômes cyclotomiques.

Démonstration.
On note $A_n := \{P \in \mathbf{Z}[X]$ unitaire, de degré n tels que $V(P) \subset \overline{D}\}$.

Étape 1 : A_n est fini.

Pour démontrer ce point, il suffit de borner chaque coefficient de P car ceux-ci sont entiers. On a une majoration sur les racines et on souhaite obtenir une majoration des coefficients d'où l'idée d'utiliser les relations coefficients-racines. Soit $P = \prod_{i=1}^{n}(X - \xi_i) = X^n + p_{n-1}X^{n-1} + \cdots + p_0 \in A_n$. Alors, si $k \in [\![1\,,\, n]\!]$,

$$p_{n-k} = (-1)^k \sum_{1 \leqslant i_1 < \cdots < i_k \leqslant n} \xi_{i_1} \cdots \xi_{i_k}$$

donc $|p_{n-k}| \leqslant \sum_{i_1 < \cdots < i_k} 1 = \binom{n}{k}$ et les coefficients de P ne peuvent prendre qu'un nombre fini de valeurs.

Étape 2 : $\forall k \in \mathbf{N}, \quad P_k := \prod_{i=1}^{n}(X - \xi_i^k) \in A$.

Soit $k \in \mathbf{N}$. On a déjà P_k unitaire, de degré n, et à racines dans $\overline{D}$. Il reste à voir qu'il est à coefficients entiers.

Version 1 : par le théorème de structure des polynômes symétriques.
Réutilisant les relations coefficients-racines, on peut développer P_k :

$$P_k = X^n + \sum_{j=1}^{n}(-1)^j \Sigma_j(\xi_1^k, \ldots, \xi_n^k) X^{n-j}$$

où Σ_j est le j-ème polynôme symétrique élémentaire. Puisque $S_j = \Sigma_j(X_1^k, \ldots, X_n^k) \in \mathbf{Z}[X_1, \ldots, X_n]^{\mathfrak{S}_n}$, le théorème de structure des polynômes symétriques nous assure qu'il existe $Q_j \in \mathbf{Z}[X_1, \ldots, X_n]$ tel que $S_j = Q_j(\Sigma_1, \ldots, \Sigma_n)$. Ainsi,

$$\Sigma_j(\xi_1^k, \ldots, \xi_n^k) = S_j(\xi_1, \ldots, \xi_n) = Q_j\big(\underbrace{\Sigma_1(\xi_1, \ldots, \xi_n)}_{-p_{n-1} \in \mathbf{Z}}, \ldots, \underbrace{\Sigma_n(\xi_1, \ldots, \xi_n)}_{(-1)^n p_0 \in \mathbf{Z}}\big)$$

d'où $P_k \in \mathbf{Z}[X]$ et finalement $P_k \in A_n$.

Version 2 : par un résultant.
On pose $Q_k = X^k - Y \in \mathbf{Z}[X, Y]$ et :

$$R_k(Y) := \operatorname{Res}_X(P, Q_k) = \begin{vmatrix} 1 & & & & 1 & & \\ p_{n-1} & \ddots & & & 0 & \ddots & \\ \vdots & & \ddots & & \vdots & & 1 \\ \vdots & & & 1 & \vdots & & 0 \\ \vdots & & & p_{n-1} & \vdots & & \vdots \\ p_0 & & & \vdots & 0 & & \vdots \\ & \ddots & & \vdots & -Y & & \vdots \\ & & \ddots & \vdots & & \ddots & 0 \\ & & & p_0 & & & -Y \end{vmatrix} \in \mathbf{Z}[Y].$$

L'idée derrière ce résultant est que ses racines doivent être des puissances k-ème de racines de P. Or,

$$R_k(Y) = \prod_{i=1}^{n} Q_k(\xi_i, Y) = \prod_{i=1}^{n} (\xi_i^k - Y) = (-1)^n P_k(Y)$$

d'où $P_k \in \mathbf{Z}[X]$ et finalement $P_k \in A_n$.

Étape 3 : Décomposition de P.

Les deux points précédents donnent que $\{P_k \ ; \ k \in \mathbf{N}\}$ est fini, donc l'ensemble de leurs racines

$$\{\xi_i^k \ ; \ k \in \mathbf{N}, \ i \in [\![1\,,\,n]\!]\}$$

aussi. Donc si $i \in [\![1\,,\,n]\!]$, l'ensemble des puissances de ξ_i est fini d'où $\xi_i = 0$ ou ξ_i est d'ordre fini dans $\mathbf{C}^*$ et ξ_i est une racine de l'unité.
On écrit donc, quitte à renuméroter les ξ_i, $P = X^{n-r} \prod_{j=1}^{r} (X - \xi_j)$ avec ξ_j racine primitive n_j-ème de l'unité. Ainsi $P \mid X^{n-r} \prod_{j=1}^{r} \phi_{n_j}$ dans $\mathbf{C}[X]$ mais, puisque ce sont tous des polynômes de $\mathbf{Z}[X]$, aussi dans $\mathbf{Q}[X]$. Finalement, puisque X et les ϕ_{n_j} sont irréductibles dans $\mathbf{Q}[X]$ qui est principal donc a fortiori factoriel, on obtient le résultat. □

Voici une application de ce théorème :

Proposition 0.28.2.
Soient $B \in \mathscr{M}_n(\mathbf{Z})$ et $k \geqslant 2$ un entier tel que $I_n + kB$ soit d'ordre fini m.
Si $k \geqslant 3$, alors $B = 0$ et si $k = 2$, alors soit $B = 0$, soit $m = 2$.

Démonstration.
La matrice $I_n + kB$ est diagonalisable et $\sigma(I_n + kB) \subset \mathbb{U}_m$ donc $B = \frac{1}{k}((I_n + kB) - I_n)$ est diagonalisable, avec pour valeurs propres :

$$\left\{ \frac{\xi - 1}{k} \ ; \ \xi \in \sigma(I_n + kB) \right\} \subset \overline{D}$$

Or ce sont les racines de $\chi_B \in \mathscr{M}_n(\mathbf{Z})$ donc, par théorème de Kronecker, ses racines sont soit nulles, soit des racines de l'unité.

Si $k \geqslant 3$: $|\frac{\xi-1}{k}| < 1$ donc $\sigma(B) = \{0\}$ et, étant diagonalisable, $B = 0$.

Si $k = 2$: Pour tout $\xi \in \sigma(I_n + kB)$, d'une part $|\frac{\xi-1}{2}| \in \{0, 1\}$ et d'autre part $\xi \in \mathbb{U}_m$ car $I_n + kB$ est d'ordre m. Par égalité dans l'inégalité triangulaire, nécessairement $\xi \in \{-1, 1\}$ et donc $I_n + kB$ est d'ordre $m = 1$, soit $B = 0$, ou d'ordre $m = 2$. □

0.29 Automorphismes de $k(X)$

Leçons : 122 ; 125 ; 141 ; 142

Références :

—

Soit A un anneau principal et K son corps des fractions.

Définition 0.29.1.
Le contenu d'un polynôme $P \in A[X]$, noté $c(P)$, est le pgcd de ses coefficients (défini à un inversible près).

Lemme 0.29.2 (Gauss)**.**
Soit $P, Q \in A[X]$. On a $c(PQ) = c(P)c(Q)$.

Démonstration.
Puisque $PQ = c(Q)c(P)\frac{P}{c(P)}\frac{Q}{c(Q)}$, on peut supposer $c(P) = c(Q) = 1$.
Supposons par l'absurde que $c(PQ) \neq 1$. Alors les coefficients de PQ ont un diviseur irréductible commun d. Comme A est principal, l'idéal (d) est maximal donc $L := {}^{A}\!/_{(d)}$ est un corps. Projeté dans $L[X]$:

$$\overline{PQ} = \overline{P}\,\overline{Q} = 0$$

mais $L[X]$ est intègre car L l'est. Donc $\overline{P} = 0$ ou $\overline{Q} = 0$, impossible car $c(P) = c(Q) = 1$. □

Lemme 0.29.3.
Soit $P \in A[X]$ tel que $c(1)$. Alors,

$$P \text{ est irréductible dans } A[X] \iff P \text{ est irréductible dans } K[X].$$

Démonstration.
$\Longleftarrow$: Supposons P irréductible dans $K[X]$. Si $P = QR$ dans $A[X]$, cette décomposition est valable aussi dans $K[X]$ donc Q ou R est inversible dans K, disons Q.
Alors, $Q \in K^\times \cap A = A \setminus \{0\}$ et par lemme de Gauss

$$c(P) = 1 = Qc(R)$$

donc $Q \in A^\times$ d'où P irréductible dans $A[X]$.

$\Longrightarrow$: Supposons P irréductible dans $A[X]$. Si $P = QR$ dans $K[X]$, l'idée est de faire remonter cette égalité dans $A[X]$, il existe $a, b \in A \setminus \{0\}$ tels que $aQ, bR \in A[X]$ et $c(aQ) = c(bR) = 1$. Or, à nouveau par lemme de Gauss,

$$c(abP) = ab = c(aQ)c(bR) = 1$$

donc $a, b \in A^\times$ et $Q, R \in A[X]$ d'où Q ou R inversible dans K. □

Théorème 0.29.4.
Soit k un corps. Pour $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{GL}_2(k)$ et $F \in k(X)$, on note $F_A = F\left(\frac{aX+b}{cX+d}\right)$. Alors,

$$\mathrm{Aut}_{k-alg}(k(X)) = \{\varphi_A : F \mapsto F_A \mid A \in \mathrm{GL}_2(k)\}$$

Démonstration.
$\supset$: L'évaluation étant un morphisme d'anneaux, si $A \in \mathrm{GL}_2(A)$, φ_A est un morphisme de corps. De plus, si $\lambda \in k$, $(\lambda F)_A = \lambda F_A$ donc φ_A est un morphisme d'algèbre.
Or $\varphi_{I_2} = \mathrm{id}$ et un simple calcul montre que si $A, B \in \mathrm{GL}_2(k)$, $\varphi_A \circ \varphi_B = \varphi_{AB}$ donc φ_A est inversible, d'inverse $\varphi_{A^{-1}}$.

$\subset$: Soit $\sigma \in \mathrm{Aut}(k(X))$. Notons $F := \sigma(X) = \frac{P}{Q}$ avec $P, Q \in k[X]$ premiers entre eux et $n = \max\{\deg P, \deg Q\}$.
Par propriétés de morphisme, σ est entièrement déterminée par sa valeur en X, c'est-à-dire par F. on commence par montrer que $n = 1$.
L'image de $k(X)$ par σ est $k(F)$ d'une part, mais étant un automorphisme $k(F) = k(X)$ et en particulier $[k(X) : k(F)] = 1$. Si on montre que n coïncide avec l'indice de $k(X)$ dans $k(F)$, on aura le résultat. Pour cela, il suffit de montrer que X est algébrique sur $k(F)$ et que son polynôme minimal est de degré n. On considère :

$$W(T) = P(T) - FQ(T) \in k[F][T]$$

qui s'annule en X et est de degré n. Cela montre déjà que X est bien algébrique, il reste à voir que W est irréductible sur $k(F)$.
Comme $k(X) = k(F)$, $[k(F) : k]$ est infini donc $k[F]$ est un anneau de polynômes.

Puisque $\deg_F W = 1$, W est irréductible dans $k(X)[F]$, et comme il est primitif, par le lemme implication de droite à gauche, il est irréductible dans $k[X][F] = k[F][X]$. Et utilisant le lemme dans l'autre sens, W est irréductible dans $k(F)[X]$. Finalement, $n = [k(X) : k(F)] = 1$.

Donc $P = aX + b$, $Q = cX + d \neq 0$ avec P ou Q de degré 1. Et, étant premiers entre eux, ils ne sont pas associés donc $\left(\binom{a}{b}, \binom{c}{d}\right)$ est libre et

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{GL}_2(k).$$

□

0.30 Théorèmes de Chevalley-Warning et Erdös-Ginzburg-Ziv

Leçons : 121 ; 123 ; 126 ; 144 ; 190

Références :

— Benjamin Havret `http://www.normalesup.org/~havret/pdf/developpements_maths%20bhavret.pdf`

Lemme 0.30.1.
Soient q une puissance d'un nombre premier et $0 \leqslant k < q$. Alors,

$$\sum_{x \in \boldsymbol{F}_q} x^k = 0$$

Démonstration.

— Si $k = 0$: $\sum_{x \in \mathbf{F}_q} x^k = q = 0$.

— Sinon $1 \leqslant k < q$: soit g tel que $\mathbf{F}_q^\times =< g >$. On a alors $g^k \neq 1$ et

$$\sum_{h \in \mathbf{F}_q^\times} h^k = \sum_{h \in \mathbf{F}_q^\times} (gh)^k = g^k \sum_{h \in \mathbf{F}_q^\times} h^k$$

donc

$$\sum_{h \in \mathbf{F}_q^\times} h^k = \sum_{x \in \mathbf{F}_q} x^k = 0.$$

□

Théorème 0.30.2 (Chevalley-Warning)**.**
Soient $n \in \mathbf{N}^*$, p un nombre premier et q une puissance de p.
Soit $P_1, \dots, P_r \in \mathbf{F}_q[X_1, \dots, X_n]$ tels que $\deg P_1 + \cdots + \deg P_r < n$. On note $V \subset \boldsymbol{F}_q^n$ leurs zéros communs. Alors,

$$|V| = 0 \mod p.$$

Ce théorème dit qu'un système d'équations algébriques dans un corps fini avec suffisamment d'inconnues possède un nombre de solutions divisible par p. En fait, il est utile pour garantir l'existence de solutions non triviales comme on le verra avec le second théorème.

Démonstration.
Le point clé (et élégant...) de cette démonstration est de transformer l'objet combinatoire $|V|$ en un objet algébrique. On considère

$$S = \prod_{i=1}^{r} (1 - P_i^{q-1}) \in \mathbf{F}_q[X_1, \dots, X_n].$$

On remarque que sa fonction associée est la fonction indicatrice de V. En effet,

- Si $\underline{x} = (x_1, \dots, x_n) \in V$: $P_i(\underline{x}) = 0$ pour tout $i \in [\![1\,,\, r]\!]$ et $S(\underline{x}) = 1$.
- Si $\underline{x} \notin V$: il existe $i \in [\![1\,,\, r]\!]$ tel que $P_i(\underline{x}) \in \mathbf{F}_p^\times$ qui est un groupe d'ordre $p-1$ donc par théorème de Lagrange $P_i(\underline{x})^{q-1} = 1$ d'où $S(\underline{x}) = 0$.

Ainsi, $|V| = \sum_{\underline{x} \in \mathbf{F}_q^n} S(\underline{x})$ et comme $\mathbf{F}_q$ est de caractéristique p :

$$|V| = 0 \mod p \iff \sum_{\underline{x} \in \mathbf{F}_q^n} S(\underline{x}) = 0.$$

On va donc démontrer cette égalité. Si $S = 0$, on a le résultat. Sinon, on va en fait voir que chaque monôme de S donne une somme nulle : soit $S_\alpha = c_\alpha X^{\alpha_1} \dots X_n^{\alpha_n}$ un monôme de S. Par la condition sur la somme des degrés des P_i, on a $S_\alpha \leqslant \deg S < (q-1)n$. Ainsi, tous les α_i ne peuvent pas dépasser $q-1$, il existe i_0 tel que $\alpha_{i_0} < q-1$, or :

$$\sum_{\underline{x} \in \mathbf{F}_q^n} S_\alpha(\underline{x}) = c_\alpha \sum_{x_1 \in \mathbf{F}_q} x_1^{\alpha_1} \left(\sum_{x_2 \in \mathbf{F}_q} x_2^{\alpha_2} \dots \left(\sum_{x_n \in \mathbf{F}_q} x_n^{\alpha_n} \right) \dots \right) = c_\alpha \prod_{i=1}^{n} \left(\sum_{x \in \mathbf{F}_q} x^{\alpha_i} \right) = 0$$

par le lemme. □

Théorème 0.30.3 (Erdös-Ginzburg-Ziv).
Soient $n \in \mathbf{N}^*$ et $a_1, \dots, a_{2n-1} \in \mathbf{Z}$.
Il existe une partie $I \subset [\![1\,,\,2n-1]\!]$ à exactement n éléments telle que

$$\sum_{i \in I} a_i = 0 \mod n.$$

Démonstration.
La démonstration se fera en deux étapes, d'abord le cas où $n = p$ est premier puis on montrera que le résultat est stable par produits : si il est vrai pour n et m, alors il l'est pour nm.

Étape 1 : Cas $n = p$ premier.

Soient p un nombre premier et $a_1, \dots, a_{2p+1} \in \mathbf{Z}$. On va à nouveau traduire les assertions combinatoires en assertions algébriques. On cherche des polynômes dont les zéros communs ont p coordonnées non nulles de somme divisible par p.
On pose

$$P_1 = \sum_{k=1}^{2p-1} \overline{a_k} X_k^{p-1} \in \mathbf{F}_p[X_1, \dots, X_{2p-1}].$$

Si $\underline{x} \in \mathbf{F}_p^{2p-1}$ vérifie $P_1(\underline{x}) = 0$, et si $I_{\underline{x}} = \{i \in [\![1\,,\,2p-1]\!] \mid x_i \neq 0\}$:

$$\sum_{i \in I_{\underline{x}}} \overline{a_i} = 0$$

Dit autrement, la fonction associée à P_1 vérifie

$$\tilde{P_1}(\underline{x}) = \sum_{i=1}^{2p-1} a_i \mathbb{1}_{\{x_i \neq 0\}}(x_i).$$

Les zéros de cette fonction donneront donc des entiers de somme divisible par p car on est en caractéristique p. Il faut ajouter une condition pour s'assurer d'avoir exactement p termes dans la somme. On pose aussi

$$P_2 = \sum_{k=1}^{2p-1} X_k^{p-1}$$

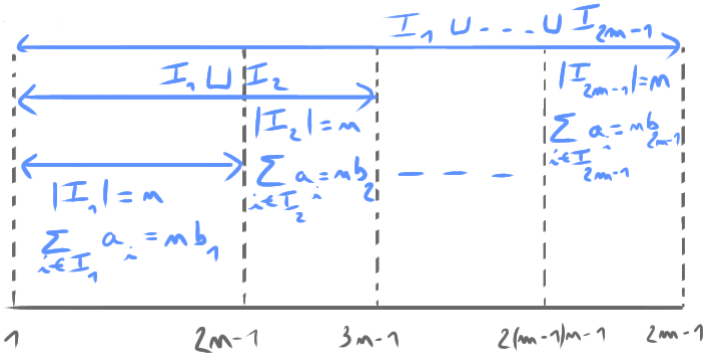
de sorte que si $\underline{x} \in \mathbf{F}_p^{2p-1}$ est tel que $P_1(\underline{x}) = 0$,

$$P_2(\underline{x}) = 0 \iff p \mid |I_{\underline{x}}| \iff \begin{cases} |I_{\underline{x}}| = 0 & \iff \underline{x} = 0 \\ & \text{ou} \\ |I_{\underline{x}}| = p \end{cases}$$

car $|I_{\underline{x}}| \leqslant 2p-1$. On cherche donc un zéro non-nul commun à ces deux polynômes.
Mais $\deg P_1 + \deg P_2 = 2p-2 < 2p-1$ et $(0, \dots, 0)$ annule P_1 et P_2 donc par théorème de Chevalley-Warning, il existe au moins $p-1$ autres solutions, ce qui conclut la première étape.

Étape 2 : Stabilité par produits.

On suppose le résultat vrai pour $m, n \in \mathbf{N}^*$, on cherche à montrer qu'il est encore vrai pour mn.
Soit $a_1, \dots, a_{2mn-1} \in \mathbf{Z}$. On soit sélectionner mn indices parmi les $2mn-1$. L'idée est de former tout d'abord $2m-1$ paquets de n indices.



On considère les $2n-1$ premiers a_i. Par le cas précédent, on peut en prendre n dont la somme est un multiple de n. Ensuite on ajoute aux $n-1$ a_i non sélectionnés les n a_i suivants pour avoir à nouveau $2n-1$ entiers, parmi lesquels on choisit un nouvel ensemble de n éléments de somme multiple de n. On répète l'opération pour former $2m-1$ ensembles distincts de n entiers de somme multiple de n.

Cela donne $b_1, \ldots, b_{2m-1} \in \mathbf{Z}$ et, par le premier point avec l'entier m, on peut trouver $J \subset [\![1\,,\, 2m-1]\!]$ ayant m éléments et tel que :

$$\sum_{j \in J} b_j = 0 \mod m.$$

Finalement, l'ensemble $K = \bigsqcup_{j \in J} I_j$ est solution car $|K| = mn$ et :

$$\sum_{k \in K} a_k = \sum_{j \in J} \sum_{i \in I_j} a_i = \sum_{j \in J} n b_j = 0 \mod mn.$$

□

Remarque sur l'optimalité : considérons un ensemble de $2n-2$ entiers composé de $n-1$ zéros et de $n-1$ fois 1. Un sous-ensemble de n entiers parmi ceux-ci sera toujours dans $[\![1\,,\, n-1]\!]$ donc non divisible par n.

0.31 Algorithme de Berlekamp

Leçons : 123 ; 125 ; 141 ; 142 ; 151

Références :

— [SR02] Philippe Saux Picart, Éric Rannou, *Cours de calcul formel : corps finis, systèmes polynomiaux, applications*

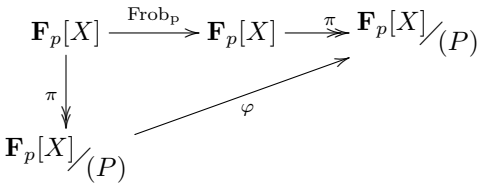
— [Dem08] Michel Demazure, *Cours d'algèbre*

Théorème 0.31.1 (Berlekamp)**.**
Soit p un nombre premier et $P \in \mathbf{F}_p[X]$ sans facteur carré. Notons $P_1, \dots, P_r$ ses facteurs irréductibles. L'application

$$\varphi : \left| \begin{array}{ccc} \boldsymbol{F}_p[X]/(P) & \to & \boldsymbol{F}_p[X]/(P) \\ \overline{Q} & \mapsto & \overline{Q^p} \end{array} \right.$$

est un endomorphisme de $\mathbf{F}_p$-algèbre et $V := \ker(\varphi - \mathrm{id})$ est de dimension r.
Si $r = 1$, alors P est irréductible. Sinon, si $T \in \boldsymbol{F}_p[X]$ est non constant vérifiant $\deg T < \deg P$ et $\overline{T} \in V$, alors il existe $s \in \boldsymbol{F}_p$ tel que $T - s$ soit un facteur non trivial de P.

Démonstration.
On commence par remarquer que si $Q \in \mathbf{F}_p[X]$, $(PQ)^p = P^p Q^p = 0 \mod P$ donc φ est bien définie. Donc on a la factorisation suivante, qui n'est autre que la fonction φ du théorème.

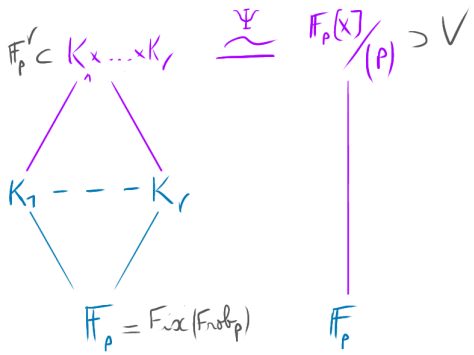


De plus, puisque le Frobenius et la surjection canonique de $\mathbf{F}_p[X]$ sur son quotient par (P) sont des morphismes de $\mathbf{F}_p$-algèbres, φ en est un morphisme de $\mathbf{F}_p$-algèbres.

Pour $i \in [\![1\,,\,r]\!]$, on note $K_i := \mathbf{F}_p[X]/(P_i)$ qui est un corps car P_i est irréductible. Le théorème des restes chinois nous dit que :

$$\Psi : \left| \begin{array}{ccc} \mathbf{F}_p[X]/(P) & \to & K_1 \times \cdots \times K_r \\ Q \mod P & \mapsto & (Q \mod P_1, \dots, Q \mod P_r) \end{array} \right.$$

est un isomorphisme d'anneaux.



Si $Q \in \mathbf{F}_p$,

$$\begin{aligned} \overline{Q} \in V &\iff \overline{Q}^p = \overline{Q} \\ &\iff \Psi(\overline{Q^p}) = \Psi(\overline{Q}) \\ &\iff \forall i \in [\![1\,,\,r]\!], \quad Q^p = Q \mod P_i \\ &\iff \forall i \in [\![1\,,\,r]\!], \quad (Q \mod P_i) \in \mathbf{F}_p \\ &\iff \Psi(\overline{Q}) \in \mathbf{F}_p^r \subset K_1 \times \cdots \times K_r \end{aligned}$$

ainsi Ψ induit une bijection entre $\mathbf{F}_p^n$ et V d'où $|V| = p^r$ soit $\dim V = r$.

Supposons $r \geqslant 2$. Les polynômes constants modulo P sont de dimension 1 donc il existe $T \in \mathbf{F}_p[X]$ tel que $\overline{T} \in V$ et $\overline{T}$ non constant donc T non constant.
Alors, puisque $X^p - X = \prod_{\alpha \in \mathbf{F}_p}(X - \alpha)$, on a aussi :

$$T^p - T = \prod_{\alpha \in \mathbf{F}_p} (T - \alpha)$$

où les $T - \alpha$ sont premiers deux à deux comme le montrent par exemple les relations de Bézout :

$$\frac{T - \beta}{\alpha - \beta} - \frac{T - \alpha}{\alpha - \beta} = 1$$

pour $\alpha, \beta \in \mathbf{F}_p$, $\alpha \neq \beta$. Or $T^p - T = 0 \mod P$ donc

$$P = \operatorname{pgcd}(P, T^p - T) = \operatorname{pgcd}\left(P, \prod_{\alpha \in \mathbf{F}_p} (T - \alpha)\right) = \prod_{\alpha \in \mathbf{F}_p} \operatorname{pgcd}(P, T - \alpha)$$

car les facteurs $T - \alpha$ sont premiers entre eux. Les facteurs $\operatorname{pgcd}(P, T - \alpha)$ ne peuvent tous être constants car leur produit vaut P qui est de degré $\geqslant 2$ et ne peuvent être associés à P puisque leur degré est inférieur ou égal à $\deg P - 1$. □

0.32 Primalité des nombres de Mersenne

Leçons : 120 ; 121 ; 122 ; 123 ; 125 ; 126 ; 141

Références :

— Clarence Kineider, `http://perso.eleves.ens-rennes.fr/people/clarence.kineider/agreg.html`

Définition 0.32.1.
Soit $q \in \mathbf{N}^*$. Le q-ième nombre de Mersenne est $M_q = 2^q - 1$.

Proposition 0.32.2.
Si $n \in \mathbf{N}^*$ n'est pas premier, alors M_n n'est pas premier non plus.

Démonstration.
Si $n = ab$ avec $a, b \geqslant 2$, alors

$$M_n = 2^{ab} - 1 = (2^a - 1)(1 + 2^a + \cdots + 2^{(b-1)a})$$

par sommation télescopique, ces deux termes étant > 1. □

Remarque 0.32.3. $M_2 = 3$ est premier. Il reste donc à traiter la primalité de M_q pour q premier impair.

Théorème 0.32.4.
Soit q premier impair. L'entier M_q est premier si et seulement si $(2+\sqrt{3})^{2^{q-1}} \equiv -1 \mod M_q$.

Démonstration.
Le $\sqrt{3} \mod M_q$ n'est pas bien défini dans l'énoncé. On va commencer par montrer que 3 n'est jamais un carré dans $\mathbf{Z}/M_q\mathbf{Z}$. Il faudra en fait se placer dans un anneau plus grand contenant une racine carrée de 3.

$\left(\frac{3}{M_q}\right) = -1.$

Puisque $2 \equiv -1 \mod 3$, $M_q \equiv (-1)^q - 1 \equiv 1 \mod 3$ donc par la loi de réciprocité quadratique :

$$\left(\frac{3}{M_q}\right) = (-1)^{\frac{2(2^q-2)}{4}} \left(\frac{M_q}{3}\right) = -1.$$

Cette relation est valable même si M_q n'est pas premier, en remplaçant les symboles de Legendre par ceux de Jacobi car M_q est impair. Puisque le symbole de Jacobi $\left(\frac{3}{M_q}\right)$ vaut -1, on peut affirmer que 3 n'est pas un carré modulo M_q.

On note $\mathscr{A} := (\mathbf{Z}/M_q\mathbf{Z})[X]/(X^2 - 3)$ et $\sqrt{3}$ la classe de X dans $\mathscr{A}$.

$\boxed{\Longrightarrow}$ On suppose M_q premier donc $\mathbf{F}_{M_q} = \mathbf{Z}/M_q\mathbf{Z}$ est un corps et $X^2 - 3$ y étant irréductible car de degré 2 et sans racine, $K = \mathscr{A}$ est un corps de caractéristique M_q.
On cherche à monter que $(2+\sqrt{3})^{2^{q-1}} = -1$ dans K. Pour cela, on va trouver une racine de $2+\sqrt{3}$ et montrer qu'à la puissance $2^q = M_q + 1$, elle vaut -1. Mais pour exprimer cette racine, on a besoin de trouver une racine de 2 dans K. En fait elle sera même dans le corps premier $\mathbf{F}_{M_q}$.
On a $2(2^q - 1) \equiv 2M_q \equiv 0 \mod M_q$ donc $2^{q+1} \equiv 2 \mod M_q$. Comme $q+1$ est pair, on peut poser $\sqrt{2} := 2^{\frac{q+1}{2}} \in \mathbf{F}_{M_q}$.
Posons alors $\rho := \frac{1+\sqrt{3}}{\sqrt{2}}$ et $\bar{\rho} := \frac{1-\sqrt{3}}{\sqrt{2}}$. Alors $\rho^2 = 2 + \sqrt{3}$ et

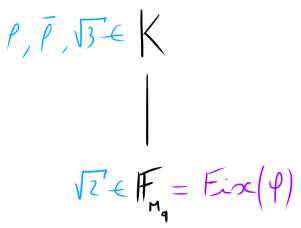
$$(2+\sqrt{3})^{2^{q-1}} = (\rho^2)^{2^{q-1}} = \rho^{2^q} = \rho \cdot \rho^{M_q}$$

Or l'application

$$\varphi : \left| \begin{array}{ccc} K & \to & K \\ x & \mapsto & x^{M_q} \end{array} \right.$$

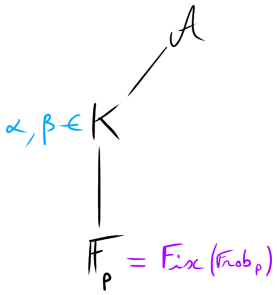
est le morphisme de Frobenius car K est de caractéristique M_q. On va utiliser les deux points suivants :

- C'est un morphisme d'anneaux.
- $\mathrm{Fix}(\varphi) = \mathbf{F}_{M_q}$ car K est un corps.



Ces deux points donnent en particulier, $\forall Q \in \mathbf{F}_{M_q}[X], \quad \varphi(Q(X)) = Q(\varphi(X))$. Ainsi, ρ et $\bar{\rho}$ étant les deux racines de $(\sqrt{2}X-1)^2 - 3 \in \mathbf{F}_{M_q}[X]$, ρ^{M_q} est soit ρ, soit $\bar{\rho}$ mais ρ n'est pas dans $\mathbf{F}_{M_q}$ donc n'est pas fixé par le FROBENIUS et finalement on a $\rho^{M_p} = \bar{\rho}$. Mais $\rho\bar{\rho} = -1$ dans K ce qui conclut pour cette implication.

$\boxed{\Longleftarrow}$ On suppose $(2+\sqrt{3})^{2^{q-1}} = -1$ dans l'anneau $\mathscr{A}$. Remarquons que M_q est impair. Soit $p \geqslant 3$ premier tel que $p \mid M_q$.
On cherche à montrer que $p = M_q$. On commence par remarquer que p est un diviseur de zéro car $p \cdot \frac{M_q}{p} = M_q = 0$ dans $\mathscr{A}$, et donc $p \notin \mathscr{A}^\times$. On peut ainsi considérer $\mathcal{M}$ un idéal maximal de $\mathscr{A}$ contenant p car $\mathscr{A}$ est fini.
Alors $K := \mathscr{A}/\mathcal{M}$ est un corps de caractéristique p et on pose $\alpha := \overline{2+\sqrt{3}}^K$, $\beta := \overline{2-\sqrt{3}}^K$.



L'hypothèse donne $\alpha^{2^{q-1}} = \overline{-1}^K$ dans K et la classe de -1 est différente de celle de 1 car sinon K serait de caractéristique 2. Donc α est d'ordre 2^q. On considère :

$$Q(X) = (X-\alpha)(X-\beta) = X^2 - 4X + 1 \in \mathbf{F}_p[X].$$

Par les mêmes propriétés du FROBENIUS qu'au point précédent, α^p est racine de Q car α l'est et que Q est à coefficient dans le corps premier $\mathbf{F}_p$ de K donc on a $\alpha^p = \alpha$ ou $\alpha^p = \beta$.

Cas $\alpha^p = \alpha$: cela équivaut à $\alpha \in \mathbf{F}_p$ donc $\alpha^{p-1} = 1$ mais α est d'ordre 2^q d'où $2^q \mid p-1$ or $p \leqslant 2^q - 1$, absurde.

Cas $\alpha^p = \beta$: on remarque que $\beta = \alpha^{-1}$ donc $\alpha^{p+1} = 1$ et $2^q \mid p+1$, ce qui implique $p = M_q$. □

Remarque : le K dans les deux implications sont *a posteriori* les mêmes, ce qui justifie cette notation.

0.33 Théorème des deux carrés

Leçons : 120 ; 121 ; 122 ; 126

Références :

—

Proposition 0.33.1.
L'anneau $\mathbf{Z}[\mathrm{i}] = \{a + \mathrm{i}b \mid a, b \in \mathbf{Z}\}$ *est euclidien muni du stathme* N *défini par* $N(a + \mathrm{i}b) = a^2 + b^2$.

Proposition 0.33.2.
Le stathme N *est multiplicatif :* $\forall z, z' \in \mathbf{Z}[\mathrm{i}],\ N(zz') = N(z)N(z')$.
De plus, les inversibles de $\mathbf{Z}[\mathrm{i}]$ *sont ses éléments de norme* 1.

Démonstration.
Le premier point est évident. Pour le second :
— Si $z \in \mathbf{Z}[\mathrm{i}]^\times$, $zz^{-1} = 1$ donc $N(z)N(z^{-1}) = N(1) = 1$ avec les normes dans $\mathbf{N}$ donc $N(z) = 1$.
— Si $N(z) = 1$, alors $z\bar{z} = 1$, or $\bar{z} \in \mathbf{Z}[\mathrm{i}]$ donc $z \in \mathbf{Z}[\mathrm{i}]^\times$.

□

Théorème 0.33.3.
On note $\mathscr{P}$ *l'ensemble des nombres premiers.*
Soit $S = \{n \in \mathbf{N} \mid \exists a, b \in \mathbf{N},\ n = a^2 + b^2\}$. *Alors, pour* $n \in \mathbf{N}$,

$$n \in S \iff \forall p \in \mathscr{P} \text{ tel que } p \equiv 3 \mod 4, \quad \nu_p(n) \text{ est pair.}$$

Démonstration.
Soit $n \in \mathbf{N}$. On cherche un critère pour affirmer que n peut s'écrire comme la somme de deux carrés, c'est-à-dire qu'il appartient à l'ensemble S défini dans l'énoncé du théorème. Le point clé de la démonstration est de remarquer que si $a, b \in \mathbf{N}, a^2 + b^2 = (a + \mathrm{i}b)(a - \mathrm{i}b)$ d'où l'idée de travailler dans l'anneau des entiers de Gauss On commence avec la caractérisation suivante :

$$n \in S \iff \exists z \in \mathbf{Z}[\mathrm{i}], \quad n = N(z)$$

La norme N étant multiplicative, on en déduit déjà un résultat important : S est stable par produit. La décomposition en facteurs premiers nous dit qu'il suffit d'étudier les puissances de nombres premiers. Soit $p \in \mathscr{P}$. On va commencer par les facteurs sans puissance.

Étape 1 : $p \in S \iff p$ est réductible dans $\mathbf{Z}[\mathrm{i}]$.

— Si $p \in S$, $p = N(z) = z\bar{z}$. De plus, $p = N(z) = N(\bar{z}) > 1$ donc z et $\bar{z}$ ne sont pas inversibles d'où p réductible.
— Si $p = zz'$, alors $N(p) = p^2 = N(z)N(\bar{z})$, égalité dans $\mathbf{N}$, et $N(z), N(z') \neq 1$ car z, z' ne sont pas inversibles donc $N(z) = N(z') = p \in S$.

Notons $(\mathbf{F}_p^\times)^2$ les carrés de $\mathbf{F}_p^\times$.

Étape 2 : p est réductible dans $\mathbf{Z}[\mathrm{i}] \iff -1 \in (\mathbf{F}_p^\times)^2$.

Commençons par remarquer que :

$$\mathbf{Z}[\mathrm{i}]/(p) \simeq \frac{\mathbf{Z}[X]/(X^2+1)}{(p)} \simeq \frac{\mathbf{Z}[X]/(p)}{(X^2+1)} \simeq \mathbf{F}_p[X]/(X^2+1)$$

d'où les équivalences :

$$\begin{aligned} p \text{ irréductible dans } \mathbf{Z}[\mathrm{i}] &\iff (p) \text{ idéal premier de } \mathbf{Z}[\mathrm{i}] \\ &\iff \mathbf{Z}[\mathrm{i}]/(p) \text{ intègre} \\ &\iff \mathbf{F}_p[X]/(X^2+1) \text{ intègre} \\ &\iff X^2 + 1 \text{ irréductible sur } \mathbf{F}_p \\ &\iff -1 \notin (\mathbf{F}_p^\times)^2 \end{aligned}$$

On a donc besoin de savoir quand -1 est un carré dans $\mathbf{F}_p$. Dans la suite, on aura besoin de supposer p impair donc on commence par traiter le cas $p = 2$. Comme $2 = 1^2 + 1^2$, $2 \in S$. Supposons maintenant $p > 2$.

Étape 3 : $x \in (\mathbf{F}_p^\times)^2 \iff x^{\frac{p-1}{2}} = 1$.

Notons $A = \left\{x \in \mathbf{F}_p^\times \mid x^{\frac{p-1}{2}} = 1\right\}$. Comme $\mathbf{F}_p$ est un corps, le polynôme $X^{\frac{p-1}{2}} - 1$ possède au plus $\frac{p-1}{2}$ racines, donc $|A| \leqslant \frac{p-1}{2}$. De plus,

$$\varphi : \left| \begin{array}{ccc} \mathbf{F}_p^\times & \to & (\mathbf{F}_p^\times)^2 \\ x & \mapsto & x^2 \end{array} \right.$$

est un morphisme surjectif qui vérifie $\ker\varphi = \{1, -1\}$ son noyau étant les racines de $X^2 - 1$ dans $\mathbf{F}_p$ avec $p \neq 2$ d'où $|(\mathbf{F}_p^\times)^2| = \frac{p-1}{2}$. Enfin, si $x \in (\mathbf{F}_p^\times)^2$, il existe $y \in \mathbf{F}_p^\times$ tel que $x = y^2$ et $x^{\frac{p-1}{2}} = y^{p-1} = 1$. Ainsi $(\mathbf{F}_p^\times)^2 \subset A$ et $|A| \leqslant |(\mathbf{F}_p^\times)^2|$ donc $A = (\mathbf{F}_p^\times)^2$.

On en déduit en particulier que :

$$-1 \in (\mathbf{F}_p^\times)^2 \iff \frac{p-1}{2} \text{ est pair} \iff p \equiv 1 \mod 4$$

où la deuxième équivalence se montre facilement en utilisant le fait que p est premier impair donc $p \equiv 1 \mod 4$ ou $p \equiv 3 \mod 4$.

Donc finalement, mettant bout-à-bout toutes les équivalences

$$p \in S \iff p \equiv 1 \mod 4$$

Étape 4 : Conclusion.

Notons $n = \prod p_i^{\nu_{p_i}}$.

— Si $p_i = 2$, alors $p_i^{\nu_{p_i}} \in S$.

— Si $p_i \equiv 1 \mod 4$, idem.

— Si $p_i \equiv 3 \mod 4$, $p_i \notin S$ mais $p_i^2 = p_i^2 + 0^2 \in S$.

Cela suffit pour montrer $\Longleftarrow$ grâce à la stabilité de S par produit.

Supposons $n = a^2 + b^2$ et $p \equiv 3 \mod 4$. Si $\nu_p(n) = 0$, alors c'est bon, sinon $p \mid n = (a + \mathrm{i}b)(a - \mathrm{i}b)$ donc par l'étape 1, il est irréductible dans $\mathbf{Z}[\mathrm{i}]$ qui est euclidien donc factoriel, et il divise l'un des facteurs. Étant réel, $p \mid a, b$ soit $a = pa'$, $b = pb'$ et $n = p^2(a'^2 + b'^2)$.
Ainsi, $\frac{n}{p^2} \in S$ et $\nu_p(\frac{n}{p^2}) = \nu_p(n) - 2$ d'où le résultat par récurrence. □

Références

[App13] Walter Appel. *Probabilités pour les non-probabilistes*. Paris : H & K, 2013. ISBN : 978-2-35141-298-5.

[BK06] Michel Benaïm et Nicole el Karoui. *Promenade aléatoire: chaînes de Markov et simulations ; martingales et stratégies*. 3. impr. Mathématiques appliquées. Palaiseau : Les éd. de l'École polytechnique, 2006. 312 p. ISBN : 978-2-7302-1168-0.

[BP18] Marc Briane et Gilles Pagès. *Théorie de l'intégration: analyse convolution et transformée de Fourier*. 7e éd. Louvain-la-Neuve Paris : De Boeck supérieur, 2018. ISBN : 978-2-8073-1788-8.

[CQ09] Denis Choimet et Hervé Queffélec. *Analyse mathématique: grands théorèmes du vingtième siècle*. Tableau noir 104. Paris : Calvage & Mounet, 2009. ISBN : 978-2-916352-10-7.

[Dem08] Michel Demazure. *Cours d'algèbre*. Nouvelle bibliothèque mathématique 1. Paris : Cassini, 2008. ISBN : 978-2-84225-127-7.

[IP19] Lucas Isenmann et Timothée Pecatte. *L'oral à l'agrégation de mathématiques: une sélection de développements*. Références sciences. Paris : Ellipses, 2019. ISBN : 978-2-340-03527-0.

[Kor10] Jacob Korevaar. *Tauberian Theory: A Century of Developments*. 2010. ISBN : 978-3-642-05919-3.

[Laf10] Jacques Lafontaine. *Introduction aux variétés différentielles*. Nouvelle éd. Collection Grenoble sciences. Les Ulis : EDP sciences, 2010. ISBN : 978-2-7598-0572-3.

[Pey12] Jacques Peyrière. *Convolution, séries et intégrales de Fourier*. Références sciences. Paris : Ellipses, 2012. ISBN : 978-2-7298-7205-2.

[Rou09] François Rouvière. *Petit guide de calcul différentiel à l'usage de la licence et de l'agrégation*. 3e éd. revue et corrigée. Enseignement des mathématiques 4. Paris : Cassini, 2009. ISBN : 978-2-84225-109-3.

[Ser02] D. Serre. *Matrices: Theory and Applications*. Graduate Texts in Mathematics 216. New York : Springer, 2002. 202 p. ISBN : 978-0-387-95460-8.

[SR02] Philippe Saux Picart et Eric Rannou. *Cours de calcul formel: corps finis, syst??mes polynomiaux : applications*. Paris : Ellipses, 2002. ISBN : 978-2-7298-1025-2.

[SS88] Yvonne Sortais et René Sortais. *Géométrie de l'espace et Du Plan: Synthèse de Cours, Exercices Résolus*. Actualités Scientifiques et Industrielles 1424. Paris : Hermann, 1988. 394 p. ISBN : 978-2-7056-1424-9.