♯ Développements

Antoine Diez

École Normale Supérieure de Rennes - Université de Rennes 1

2016 - 2017

Cher lecteur,

J'ai rassemblé ici l'ensemble des développements que j'ai préparés et consciencieusement rédigés pendant un an. À l'instar des plans, ceux d'analyse et probabilités sont peut être un peu plus ambitieux et j'espère, de meilleure facture. Beaucoup de développements m'ont servi d'aide-mémoire pour les plans et contiennent donc parfois un peu plus que ce que je prévoyais de présenter. La dernière partie contient des développements que j'ai abandonné en cours de route, parfois avec soulagement, souvent avec regret. Il s'agit donc principalement de développements que je juge intéressants mais dont la constitution heurte les exigences cadenassées d'un concours. Ou alors ils étaient trop difficiles. Ou nuls.

La plupart de ces développements ont été élaborés en collaboration [1] (ou au moins relus) avec mes camarades que je remercie ici : Grégoire Clarté, Gabriel Lepetit, David Michel, Alexandre Eimer entre autres et par pdf interposés les agrégatifs des années précédentes, Adrien Laurent et Florian Lemonnier en tête.

Ce document fait partie d'un triptyque : un autre document contient quelques conseils bibliographiques et un troisième la plupart des plans rédigés cette année, numérisés et commentés.

1. C'est à dire que j'ai honteusement tout recopié sur eux.

Table des matières

1 Algèbre

1.1 Autour des endomorphismes semi-simples

On se place dans E, un $\mathbf{K}$-espace vectoriel de dimension finie n.

Définition. On dit que $f \in \mathcal{L}(E)$ est semi-simple lorsque tout sous-espace vectoriel de E stable par f admet un supplémentaire stable par f.

Lemme. *Soit* $\mathbf{L}/\mathbf{K}$ *une extension de corps. Alors* $\Pi_{f,\mathbf{K}} = \Pi_{f,\mathbf{L}}$.

Preuve. C'est une conséquence de l'indépendance du rang vis à vis du corps de base (qui provient de l'indépendance du résultat du calcul des mineurs). Maintenant, on a déjà :

$$\Pi_{f,\mathbf{L}} | \Pi_{f,\mathbf{K}}$$

et comme ces polynômes sont unitaires, il suffit de montrer qu'ils sont de même degré pour conclure. Or, le degré du polynôme minimal de f sur $\mathbf{L}$ est égal au rang de la famille $(id, f, \dots, f^{n-1})$ dans $\mathcal{L}(E)$ qui est un espace vectoriel de dimension finie n^2. Comme le rang ne dépend pas du corps de base et quitte à tout mettre dans une grosse matrice, on en déduit l'égalité annoncée. □

Lemme. *Soit* F *un sous-espace stable par* f. *On note* $\Pi_f = P_1^{\alpha_1} \dots P_r^{\alpha_r}$. *On a :*

$$F = \bigoplus_{i=1}^{r} \Big[\operatorname{Ker} P_i^{\alpha_i}(f) \cap F \Big].$$

Preuve. Par le lemme des noyaux, on sait que :

$$F = \bigoplus_{i=1}^{r} \operatorname{Ker} P_i^{\alpha_i}(f|_F) = \bigoplus_{i=1}^{r} \Big[\operatorname{Ker} P_i^{\alpha_i}(f) \cap F \Big]$$

□

Théorème. *Un endomorphisme* f *est semi-simple si et seulement si son polynôme minimal* Π_f *est un produit de polynômes irréductibles unitaires distincts deux à deux.*

Preuve. Progressivement :

Étape 1. Lorsque Π_f *est irréductible.*

On va montrer que f est semi-simple, considérons donc F un sous-espace stable par f. Si $F = E$, il n'y a rien à faire. Sinon, soit $x \in E \setminus F$ et

$$E_x = \{P(f)(x),\ P \in \mathbf{K}[X]\}.$$

Clairement E_x est stable par f. Pour conclure et quitte à itérer le processus, il suffit de montrer que

$$F \text{ et } E_x \text{ sont en somme directe.}$$

L'idéal $I_x = \{P \in \mathbf{K}[X],\ P(f)(x) = 0\}$ est non réduit à 0 (il y a Π_f) et principal donc il est engendré par un unique polynôme unitaire Π_x. Comme $\Pi_x | \Pi_f$, ce polynôme est irréductible.

Soit $y = P(f)(x) \in E_x \cap F$ que l'on suppose non nul. Alors $P \notin I_x$, c'est à dire que Π_x ne divise pas P et comme il est irréductible, P et Π_x sont premiers entre eux. Par le théorème de Bézout, on peut écrire :

$$UP + V\Pi_x = 1.$$

On trouve :

$$x = U(f) \circ P(f)(x) = U(f)(y) \in F \text{ car } y \in F.$$

C'est absurde !

Avant de continuer, voici une seconde preuve plus ocnceptuelle de cette première étape : comme Π_f est irréductible, on peut considérer le corps $\mathbf{L} = \mathbf{K}[X]/(\Pi_f)$ et munir E d'une structure de $\mathbf{L}$-espace vectoriel en posant $\bar{P} \cdot x = P(f)(x)$. On note $E_{\mathbf{L}}$ cette nouvelle structure et $E_{\mathbf{K}}$ l'ancienne. On vérifie immédiatement que $F \subset E_{\mathbf{K}}$ est un sous-espace vectoriel stable par f si et seulement si $F \subset E_{\mathbf{L}}$ est un sous-espace vectoriel. Il suffit de considérer G un supplémentaire de F dans $E_{\mathbf{L}}$.

Étape 2. Cas général, condition nécéssaire.

Soit $f \in \mathcal{L}(E)$ un endomorphisme semi-simple de polynôme minimal $\Pi_f = P_1^{\alpha_1} \dots P_r^{\alpha_r}$. Supposons qu'il existe $\alpha_i \geq 2$. On écrit alors $\Pi_f = P^2 Q$.

$F = \operatorname{Ker} P(f)$ est un sous-espace stable par f qui admet un supplémentaire stable noté S. Si $x \in S$, alors

- $\Pi_f(f)(x) = P(f)P(f)Q(f)(x) = 0$ donc $P(f)Q(f)(x) \in F$.
- S est stable par f donc $P(f)Q(f)(x) \in S$.

Finalement, $P(f)Q(f)(x) \in F \cap S = \{0\}$ et $P(f)Q(f)$ s'annule sur S.

Mais $P(f)Q(f) = Q(f)P(f)$ donc par définition de F, $P(f)Q(f)$ s'annule aussi sur F. Puisque F et S sont supplémentaires, le polynôme PQ annule f ce qui contredit la minimalité de Π_f.

Étape 3. Cas général, condition suffisante.

Soit $f \in \mathcal{L}(E)$ dont le polynôme minimal est de la forme $\Pi_f = P_1 \dots P_r$ où les P_i sont des polynômes irréductibles distincts. Soit F un sous-espace stable par f. Pour tout $i \in \{1, \dots r\}$, $F \cap \operatorname{Ker} P_i(f)$ est stable par $f|_{\operatorname{Ker} P_i(f)}$. Puisque P_i est un polynôme irréductible qui annule $f|_{\operatorname{Ker} P_i(f)}$, c'est le polynôme minimal de $f|_{\operatorname{Ker} P_i(f)}$. La première étape fournit l'existence d'un sous-espace S_i stable par $f|_{\operatorname{Ker} P_i(f)}$ (donc par f) tel que :

$$\operatorname{Ker} P_i(f) = (F \cap \operatorname{Ker} P_i(f)) \cap S_i.$$

Il suffit d'écrire :

$$E = \bigoplus_{i=1}^{r} \Big[F \cap \operatorname{Ker} P_i(f) \oplus S_i \Big] = \left[\bigoplus_{i=1}^{r} \Big(F \cap \operatorname{Ker} P_i(f) \Big) \right] \oplus \bigoplus_{i=1}^{r} S_i = F \oplus S$$

et S est stable par f qui est donc semi-simple. □

Lorsque $\mathbf{K}$ est algébriquement clos, les polynômes irréductibles sont de degré 1 donc f est semi-simple si et seulement si f est diagonalisable. On note maintenant M la matrice de f dans une base et on dit qu'elle est semi-simple lorsque f l'est.

Théorème. *Si le corps* $\mathbf{K}$ *est de caractéristique nulle, alors* M *est semi-simple si et seulement s'il existe une extension* $\mathbf{L}/\mathbf{K}$ *dans laquelle* M *est diagonalisable.*

Preuve. Soit $\mathbf{K}$ de caractéristique nulle et $\mathbf{L}/\mathbf{K}$ une extension de corps. On commence par montrer que M est semi-simple sur $\mathbf{K}$ si et seulement si M l'est sur $\mathbf{L}$ (ici, M est à coefficients dans $\mathbf{K}$). Le polynôme minimal de M sur $\mathbf{K}$ est le même que celui de M sur $\mathbf{L}$. Il suffit donc de montrer que Π_M est sans facteur carré dans $\mathbf{K}[X]$ si et seulement s'il est sans facteur carré dans $\mathbf{L}[X]$.

Dans un corps de caractéristique nulle, P est sans facteur carré équivaut à $P \wedge P' = 1$. Mais comme le calcul du pgcd s'effectue dans $\mathbf{K}$, le fait que P et P' soient premiers entre eux ne dépend pas du corps considéré.

Prouvons le théorème : supposons que M est semi-simple dans $\mathbf{K}$. Alors soit $\mathbf{L}$ un corps de décomposition de $\Pi_M \in \mathbf{K}[X]$. Dans $\mathbf{L}[X]$, le polynôme Π_M est scindé à racines simples donc M est diagonalisable. Réciproquement, si M est diagonalisable dans $\mathbf{L}$ alors M est semi-simple dans $\mathbf{L}$ et on vient de montrer que ce fait était équivalent à la semi-simplicité de M sur $\mathbf{K}$. □

Références.
X. Gourdon, *Algèbre*
V. Beck, J. Malick, G. Peyré, *Objectif Agrégation*

122 Anneaux principaux. Applications. *(allez...)*
141 Polynômes irréductibles à une indéterminée. Corps de rupture. Exemples et applications.
153 Polynômes d'endomorphisme en dimension finie. Réduction d'un endomorphisme en dimension finie. Applications.
154 Sous-espaces stables par un endomorphisme ou une famille d'endomorphismes d'un espace vectoriel de dimension finie. Applications.
155 Endomorphismes diagonalisables en dimension finie.

1.2 Autour du déterminant de Cayley-Menger

C'est très joli quand on a compris. La preuve de la première proposition est à la fois astucieuse et calculatoire et méritent donc d'être admise. On doit pouvoir mentionner ce résultat dans les leçons de géométrie affine. Je n'ai pas compris le truc de Zavidovique avec la récurrence d'ordre 2, pour moi l'ordre 1 suffit mais ça n'est pas très important je crois.

On se place dans $\mathbf{R}^n$ muni de structure affine standard.

Définition. Soient $x_0, \dots, x_n$ des points de $\mathbf{R}^n$ et $d_{i,j} = \|x_i - x_j\|$ les distances associées. Le déterminant de Cayley-Menger est défini par :

$$\Gamma(x_0, \dots, x_n) = \begin{vmatrix} 0 & 1 & \dots & 1 \\ 1 & \ddots & d_{i,j}^2 & \\ \vdots & d_{i,j}^2 & \ddots & \\ 1 & & & 0 \end{vmatrix}.$$

C'est un déterminant de taille $n+2$. On convient que $\Gamma(x_0) = \begin{vmatrix} 0 & 1 \\ 1 & 0 \end{vmatrix} = -1$.

Notons bien qu'on a seulement besoin des distances $(d_{i,j})$ pour définir un déterminant de Cayley-Meger et pas des points. C'est précisément le but du théorème de montrer le lien entre les deux.

Proposition. *La volume du parallélogramme*[1] *défini par les points $x_0, \dots, x_n$ vérifie :*

$$\det(x_1 - x_0, \dots, x_n - x_0)^2 = \frac{(-1)^{n+1}}{2^n} \Gamma(x_0, \dots, x_n).$$

Théorème. *Soient $(d_{i,j})_{0 \le i,j \le n}$ des réels positifs vérifiant :*

$$\forall i \neq j, \;\; d_{i,j} = d_{j,i} > 0 \quad \textit{et} \quad d_{i,i} = 0.$$

Il y a équivalence entre :

(1) Il existe des points $x_0, \dots, x_n \in \mathbf{R}^n$ qui sont les sommets d'un simplexe non dégénéré avec $d_{i,j} = \|x_i - x_j\|$.

(2) Pour toute sous-famille $i_1, \dots, i_k$, le déterminant de Cayley-Menger associé aux distances $(d_{i_p,i_q})_{1 \le p,q \le k}$ est de signe $(-1)^k$.

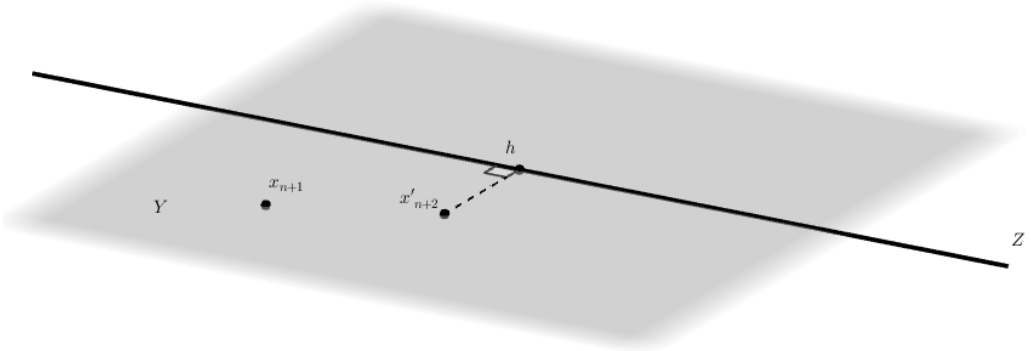
L'implication (1) $\Rightarrow$ (2) est donnée par la proposition précédente. On ne prouvera donc que la réciproque.

Preuve. On va procéder par récurrence sur la dimension n de l'espace ambiant. En dimension 1, on place un point x_0 n'importe où puis un autre point à la distance voulue, le déterminant de Cayley-Menger associé vaut $2d^2 > 0$ et si on ne prend qu'un point on se rappelle qu'on a convenu qu'il valait -1. Place à l'hérédité.

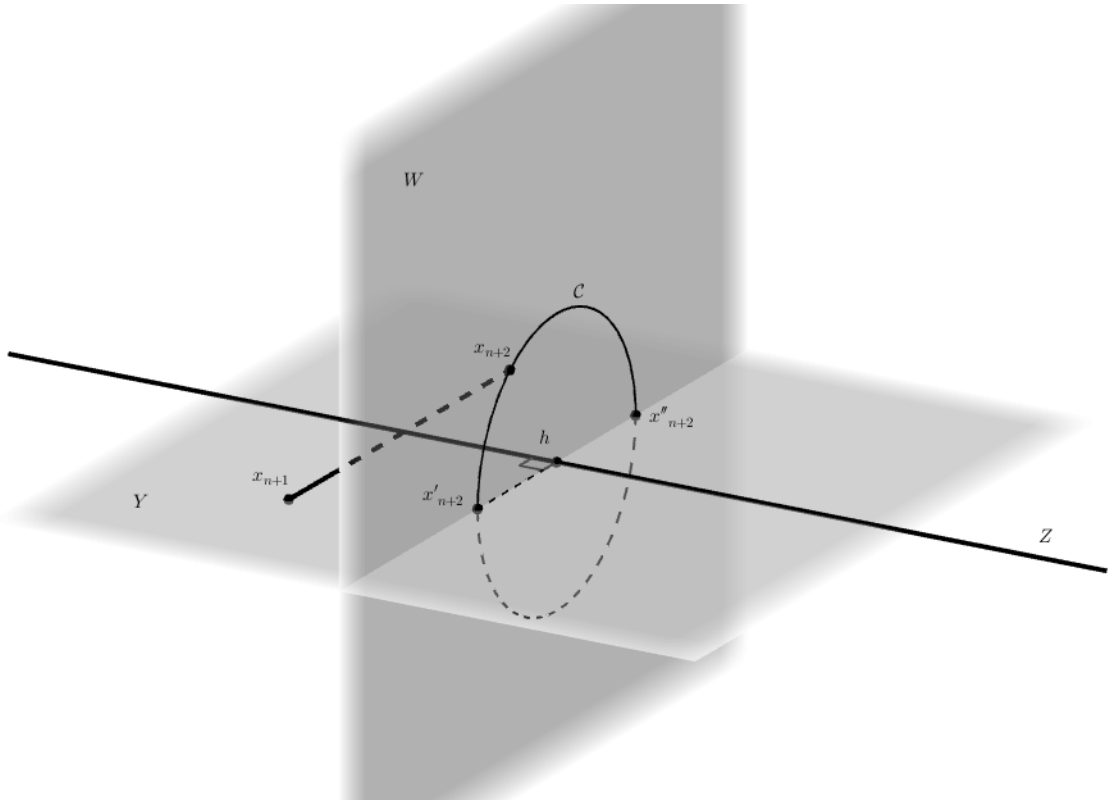
On se place dans $\mathbf{R}^{n+2}$, $n \ge 0$ et on se donne $(d_{i,j})_{0 \le i,j \le n+2}$ des distances comme dans l'énoncé. La récurrence est d'ordre 1 mais on aura quand même besoin de se donner de la marge, d'où le $n+2$ et pas $n+1$.

1. Un simplexe est l'enveloppe convexe d'une base et un parallélogramme l'ensemble des combinaisons linéaires avec des coefficients dans $[0,1]$ des vecteurs de la base. Le théorème de Fubini montre facilement que Volume(simplexe)=Volume(parallélogramme)/$n!$.

- On commence par appliquer l'hypothèse de récurrence avec les $(d_{i,j})_{0\leq i,j\leq n+1}$, cela donne naissance à $(n+2)$ points $x_0,\dots,x_{n+1}$ qui engendrent un hyperplan (affine) dans $\mathbf{R}^{n+2}$, noté Y.
- Dans l'hyperplan Y qui est de dimension $(n+1)$, on note $Z\subset Y$ l'hyperplan affine de dimension n engendré par $x_0,\dots,x_n$. On applique à nouveau l'hypothèse de récurrence avec les $(d_{i,j})_{i.j\in\{0,\dots,n,n+2\}}$ en faisant coïncider les $(n+1)$ premiers points avec $x_0,\dots,x_n$ (c'est possible à une isométrie affine près). Un nouveau point est né, on l'appelle $x'_{n+2}\in Y$. Comme le simplexe engendré par $x_0,\dots,x_n,x'_{n+2}$ est non dégénéré, on est sûr que $x'_{n+2}\notin Z$.
- On projette orthogonalement x'_{n+2} sur Z et on appelle h son projeté.



- On appelle W le supplémentaire orthogonal de Z passant par h. Il est de dimension 2. On trace $\mathcal{C}$ le cercle inclus dans W de centre h et passant par x'_{n+2}. On appelle x''_{n+2} l'autre point d'intersection de $\mathcal{C}$ et Y. Quitte à changer de nom, on suppose que x_{n+1} et x'_{n+2} sont du même côté comme sur le dessin.



On peut commencer à réfléchir. Par construction, pour tout $x \in \mathcal{C}$ et tout $i \in \{0, \ldots, n\}$, on a :

$$\|x - x_i\| = \|x'_{n+2} - x_i\| = d_{i,n+2}$$

et il ne reste plus qu'à trouver un point $x_{n+2} \in \mathcal{C}$ vérifiant :

$$\|x_{n+2} - x_{n+1}\| = d_{n+1,n+2}.$$

Pour $\xi \in \mathbf{R}$, on définit

$$G(\xi) = \begin{vmatrix} & & & & 1 & 1 \\ & & \Gamma(x_0, \ldots, x_n) & & d^2_{0,n+1} & d^2_{n+2,0} \\ & & & & d^2_{1,n+1} & d^2_{1,n+2} \\ & & & & \vdots & \vdots \\ 1 & d^2_{n+1,0} & d^2_{n+1,1} & \cdots & 0 & \xi \\ 1 & d^2_{n+2,0} & d^2_{n+2,1} & \cdots & \xi & 0 \end{vmatrix}.$$

C'est le déterminant de Cayley-Menger d'un système de $(n+3)$ points dans lequel on a remplacé la distance $d_{n+1,n+2}$ par ξ. C'est un polynôme de degré 2 en ξ et en développant par rapport aux deux dernières colonnes, il est de la forme :

$$G(\xi) = -\Gamma(x_0, \ldots, x_n)\xi^2 + a\xi + b, \quad a, b \in \mathbf{R}.$$

Par hypothèse, $\Gamma(x_0, \ldots, x_n)$ est de signe $(-1)^{n+1}$ et $G(d^2_{n+1,n+2})$ est de même signe. De plus,

$$G(\|x_{n+1} - x'_{n+2}\|^2) = G(\|x_{n+1} - x''_{n+2}\|^2) = 0$$

car les simplexes correspondants sont dégénérés : tous les points sont dans l'hyperplan Y.

Comme G est un polynôme du second degré, $G(\xi)$ est de signe opposé au coefficient dominant lorsque :

$$\xi \in \left[\|x_{n+1} - x'_{n+2}\|^2, \|x_{n+1} - x''_{n+2}\|^2\right] = I.$$

Mais $d^2_{n+1,n+2} \in I$ car son image par G est du même signe. Or, quand x parcourt $\mathcal{C}$, $\|x - x_{n+1}\|^2$ parcourt I donc par le théorème des valeurs intermédiaires :

$$\exists\, x_{n+2} \in \mathcal{C},\ \|x_{n+2} - x_{n+1}\| = d_{n+1,n+2}$$

□

Adapté d'un travail TROP WAOOUH de l'inénarrable Grégoire Clarté.

Référence. M. Zavidovique, *Un max de maths*

151 Dimension d'un espace vectoriel (on se limitera au cas de la dimension finie). Rang. Exemples et applications.
152 Déterminant. Exemples et applications.

1.3 Classification des groupes d'ordre 12

Quelques prérequis.

- Il y a à isomorphisme près exactement deux groupes d'ordre 4 : le groupe cyclique $\mathbf{Z}/4\mathbf{Z}$ et le groupe de Klein $V_4 \simeq (\mathbf{Z}/2\mathbf{Z})^2$. On notera $V_4 = \{1, a_1, a_2, a_3\}$ avec $a_1 a_2 = a_3$.
- Le groupe des automorphismes de V_4 noté $\mathrm{Aut}(V_4)$ est isomorphe à $\mathfrak{S}_3$, le groupe des permutations de l'ensemble $\{a_1, a_2, a_3\}$ qui a trois éléments.
- Si G est un groupe et $N \triangleleft G$ alors on a une suite exacte :

  $$1 \longrightarrow N \longrightarrow G \underset{p}{\longrightarrow} G/N \longrightarrow 1.$$

  Une *section* est un morphisme $s : G/N \to G$ tel que $p \circ s = id_{G/N}$. L'existence d'une section est équivalente à l'existence d'un sous-groupe $H \subset G$ tel que

  $$p|_H : H \xrightarrow{\sim} G/N$$

  (prendre $H = s(G/N)$). Dans ce cas, il existe un produit semi-direct tel que

  $$G \simeq N \rtimes H.$$

  On dit alors que H relève G/N ou que G/N se relève en H.

- Si N est d'indice fini, les sous-groupes H qui relèvent G/N sont exactement ceux qui vérifient :
$$|H| = [G : N] \quad \text{et} \quad N \cap H = \{1\}.$$
Pour le voir, il suffit de noter que $\operatorname{Ker} p|_H = \{1\}$ et de conclure par égalité des cardinaux.

Ce qu'on va montrer.

Théorème. *Il existe à isomorphisme près exactement cinq groupes d'ordre 12 :*
- *les abéliens :* $\mathbf{Z}/12\mathbf{Z}$, $\mathbf{Z}/6\mathbf{Z} \times \mathbf{Z}/2\mathbf{Z}$.
- *les non abéliens :* $\mathbf{Z}/3\mathbf{Z} \rtimes \mathbf{Z}/4\mathbf{Z}$, $\mathbf{Z}/3\mathbf{Z} \rtimes V_4$, $V_4 \rtimes \mathbf{Z}/3\mathbf{Z}$.

Preuve. Soit G un groupe d'ordre $12 = 2^2 \times 3$. Le théorème de Sylow dit que le nombre de 3-Sylow que l'on note r vérifie
$$r|4 \quad \text{et} \quad r \equiv 1\ [3].$$
On en déduit que $r = 1$ ou 4 et on va distinguer ces deux cas.

Premier cas : $r = 1$.

On appelle N l'unique 3-Sylow de G. C'est un sous-groupe distingué isomorphe à $\mathbf{Z}/3\mathbf{Z}$ et si H est un 2-Sylow (ça existe), alors
$$|H| = 4 = [G : N] \quad \text{et} \quad N \cap H = \{1\} \quad \text{car} \quad 4 \wedge 3 = 1.$$
On en déduit que n'importe quel 2-Sylow H relève G/N et comme il n'y a que deux groupes d'ordre 4 :
$$G \simeq \mathbf{Z}/3\mathbf{Z} \rtimes \mathbf{Z}/4\mathbf{Z} \quad \text{ou} \quad G \simeq \mathbf{Z}/3\mathbf{Z} \rtimes V_4$$
selon la forme de H. Pour savoir si plusieurs produits semi-directs non isomorphes peuvent exister, on étudie les morphismes $H \to \operatorname{Aut}(\mathbf{Z}/3\mathbf{Z})$ selon la forme de H. Le groupe des automorphismes $\operatorname{Aut}(\mathbf{Z}/3\mathbf{Z})$ est isomorphe à $(\mathbf{Z}/3\mathbf{Z})^\times \simeq \mathbf{Z}/2\mathbf{Z}$ et on notera donc :
$$\operatorname{Aut}(\mathbf{Z}/3\mathbf{Z}) = \{id, u\}.$$
- Il y a deux morphismes de $\mathbf{Z}/4\mathbf{Z} \to \operatorname{Aut}(H)$ définis par l'image de 1 :
$$1 \mapsto id \quad \text{et dans ce cas} \qquad G \simeq \mathbf{Z}/3\mathbf{Z} \times \mathbf{Z}/4\mathbf{Z} \simeq \mathbf{Z}/12\mathbf{Z}.$$
$$1 \mapsto u \quad \text{et dans ce cas} \qquad G \simeq \mathbf{Z}/3\mathbf{Z} \rtimes \mathbf{Z}/4\mathbf{Z}$$
et ce dernier produit semi-direct est le seul qui n'est pas direct.
- Il y a quatre morphismes $V_4 \to \{id, u\}$ définis par les images de a_1 et a_2. Le morphisme trivial donne naissance au groupe
$$G \simeq \mathbf{Z}/3\mathbf{Z} \times V_4 \simeq \mathbf{Z}/6\mathbf{Z} \times \mathbf{Z}/2\mathbf{Z}.$$
Les trois autres morphismes donnent naissance à potentiellement trois produits semi-directs non directs mais en fait il sont tous isomorphes. On a donc sans ambiguïté :
$$G \simeq \mathbf{Z}/3\mathbf{Z} \rtimes V_4.$$

Deuxième cas : $r = 4$.

Le groupe G contient quatre 3-Sylow, il n'est pas abélien et comme les intersections entre deux 3-Sylow sont triviales, il y a exactement $4 \times (3-1) = 8$ éléments d'ordre 3. Il y a au moins un 2-Sylow qui est d'ordre 4 et qui intersecte donc les 3-Sylow de manière triviale. Ce 2-Sylow est donc unique, il est distingué dans G et on le note N.

Comme avant, si H est un 3-Sylow, on a $|H| = [G : N] = 3$ et $N \cap H = \{1\}$ donc les 3-Sylow sont autant de relèvements de G/N. On en déduit :

$$G \simeq \mathbf{Z}/4\mathbf{Z} \rtimes \mathbf{Z}/3\mathbf{Z} \quad \text{ou} \quad G \simeq V_4 \rtimes \mathbf{Z}/3\mathbf{Z}$$

selon la forme de N. Comme tout à l'heure, il s'agit de regarder les morphismes de $\mathbf{Z}/3\mathbf{Z} \to \mathrm{Aut}(N)$ selon la forme de N. Comme $\mathbf{Z}/3Z$ est cyclique d'ordre 3, on va regarder les éléments d'ordre 3 dans $\mathrm{Aut}(N)$.

- D'abord $\mathrm{Aut}(\mathbf{Z}/4\mathbf{Z}) \simeq (\mathbf{Z}/4\mathbf{Z})^\times \simeq \mathbf{Z}/2\mathbf{Z}$ et comme il n'y a pas d'élément d'ordre 3 dans $\mathbf{Z}/2\mathbf{Z}$ le seul morphisme $\mathbf{Z}/3\mathbf{Z} \to \mathrm{Aut}(\mathbf{Z}/4\mathbf{Z})$ est le morphisme trivial donc $G \simeq \mathbf{Z}/4\mathbf{Z} \times \mathbf{Z}/3\mathbf{Z}$. C'est un groupe abélien, ce qui est exclu.
- D'après les remarques préliminaires, $\mathrm{Aut}(V_4) \simeq \mathfrak{S}_3$ et il y a deux éléments d'ordre 3 (les 3 cycles (123) et (132)). Comme précédemment, les produits semi-directs qui en découlent sont isomorphes et alors :
$$G \simeq V_4 \rtimes \mathbf{Z}/3\mathbf{Z}.$$

La conclusion.

On a donc exhibé les cinq groupes annoncés et il reste à voir qu'ils sont deux à deux non isomorphes. C'est facile pour les groupes abéliens et pour ceux non abéliens, il suffit d'appliquer le théorème de Sylow pour exhiber leur structure en remarquant que si $G = N \rtimes H$ alors N et H se plongent dans G. On obtient :

$V_4 \rtimes \mathbf{Z}/3\mathbf{Z}$	possède un unique 2-Sylow(V_4) et quatre 3-Sylow	$\simeq \mathbf{Z}/3Z$
$\mathbf{Z}/3\mathbf{Z} \rtimes V_4$	possède un unique 3-Sylow($\mathbf{Z}/3\mathbf{Z}$) et trois 2-Sylow	$\simeq V_4$
$\mathbf{Z}/3\mathbf{Z} \rtimes \mathbf{Z}/4\mathbf{Z}$	possède un unique 3-Sylow($\mathbf{Z}/3\mathbf{Z}$) et trois 2-Sylow	$\simeq \mathbf{Z}/4\mathbf{Z}$

□

Après il peut être intelligent de remarquer qu'en fait :

$$\mathfrak{A}_4 \simeq V_4 \rtimes \mathbf{Z}/3\mathbf{Z} \quad \text{et} \quad \mathcal{D}_6 \simeq \mathbf{Z}/3\mathbf{Z} \rtimes V_4.$$

Référence. M. Alessandri, *Thèmes de Géométrie. Groupes en situation géométrique.*

103 Exemples de sous-groupes distingués et de groupes quotients. Applications.
104 Groupes finis. Exemples et applications.

1.4 Des noyaux itérés à la réduction de Jordan en pasant par les tableaux de Young

On note $\mathcal{N}_n(\mathbf{C})$ l'ensemble des matrices nilpotentes de $\mathcal{M}_n(\mathbf{C})$. Le but est l'étude de l'action par conjugaison de $GL_n(\mathbf{C})$ sur $\mathcal{N}_n(\mathbf{C})$.

Soit $A \in \mathcal{N}_n(\mathbf{C})$ d'indice de nilpotence $m \leq n$. On note $K_i = \operatorname{Ker} A^i$ les noyaux emboîtés :

$$\{0\} = K_0 \subsetneq K_1 \subset K_2 \subset \ldots \subset K_n = \mathbf{C}^n.$$

Pour tout entier i, on pose :

$$k_i = \dim \operatorname{Ker} A^i \quad \text{et} \quad \lambda_i = k_i - k_{i-1}.$$

Notons que $(k_i)_i$ et $(\lambda_i)_i$ sont constants sur chaque orbite de nilpotence. C'est la réciproque qui est intéressante.

Proposition. *La suite des dimensions d'essouffle au sens où :*

$$\forall i \in \{1, \ldots, n-1\},\ 0 \leq \lambda_{i+1} \leq \lambda_i.$$

De plus, $(k_i)_i$ est strictement croissante avant de devenir stationnaire au rang m.

Preuve. Soit $i \in \{1, \ldots, n-1\}$, la première inégalité est triviale puisque $K_i \subset K_{i+1}$. Et comme $AK_{i+1} \subset K_i$, on peut composer les morphismes :

$$\nu : K_{i+1} \to K_i,\ X \mapsto AX \quad \text{et} \quad \pi_i : K_i \to K_i/K_{i-1},\ Y \mapsto \overline{Y}.$$

On a :

$$\operatorname{Ker}(\pi \circ \nu) = \nu^{-1}\big(\pi_i^{-1}(\{0\})\big) = \nu^{-1}(K_{i-1}) = K_i$$

et on peut passer au quotient pour en déduire une injection $K_{i+1}/K_i \hookrightarrow K_i/K_{i-1}$.

Enfin si $\operatorname{Ker} A^j = \operatorname{Ker} A^{j+1}$ alors si $k \geq j+1$ et $x \in \operatorname{Ker} A^k$ alors $A^{k-j-1}x \in \operatorname{Ker} A^{j+1} = \operatorname{Ker} A^j$ donc $x \in \operatorname{Ker} A^{k-1}$ et on conclut facilement que $\operatorname{Ker} A^j = \operatorname{Ker} A^k$. □

Remarquons que $\sum_{j=1}^{n} \lambda_j = n$ est une partition de l'entier n que l'on représente sous la forme d'un tableau de Young que l'on note $Y(A)$ et que l'on aurait pu tout aussi bien noter $Y(\mathcal{O})$ où $\mathcal{O}$ est l'orbite de A. Remplissons le :

(1) Soit G_m un supplémentaire de K_{m-1} dans $K_m = \mathbf{C}^n$:

$$K_{m-1} \oplus G_m = K_m \quad \text{et} \quad \dim G_m = \lambda_m.$$

On décide que $(v_m^1, \ldots, v_m^{\lambda_m})$ est une base de G_m.

(2) On remarque que $(Av_m^1, \ldots, Av_m^{\lambda_m})$ est une famille libre qui engendre un sous-espace qui intersecte K_{m-2} trivialement puisque :

$$A^{m-2}Av_m^k = A^{m-1}v_m^k \neq 0.$$

(3) On a le droit de compléter :

$$(Av_m^1, \ldots, Av_m^{\lambda_m}, \ldots, v_{m-1}^{\lambda_m+1}, \ldots, v_{m-1}^{\lambda_{m-1}}) \text{ est une base de } G_{m-1}$$

avec

$$K_{m-2} \oplus G_{m-1} = K_{m-1}.$$

(4) Par récurrence descendante, on construit comme cela des supplémentaires G_r de K_{r-1} dans K_r.

On stocke tout cela dans le tableau de Young :

v_m^1	v_m^2	$\dots$	$v_m^{\lambda_m}$						
Av_m^1	Av_m^2	$\dots$	$Av_m^{\lambda_m}$	$v_{m-1}^{\lambda_m+1}$	$\dots$	$v_{m-1}^{\lambda_{m-1}}$			
$\dots$	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$		
$A^{n-r}v_m^1$	$A^{n-r}v_m^2$	$\dots$	$A^{n-r}v_m^{\lambda_m}$	$A^{n-r-1}v_{m-1}^{\lambda_m+1}$	$\dots$	$A^{n-r-1}v_{m-1}^{\lambda_{m-1}}$	$\dots$		
$\dots$	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$		
$A^{m-1}v_m^1$	$A^{m-1}v_m^2$	$\dots$	$A^{m-1}v_m^{\lambda_m}$	$A^{m-2}v_{m-1}^{\lambda_m+1}$	$\dots$	$A^{m-2}v_{m-1}^{\lambda_{m-1}}$	$\dots$	$\dots$	$v_1^{\lambda_1}$

On vient de construire une base de $\mathbf{C}^n$ et si on lit le tableau de bas en haut et de gauche à droite, on voit que la matrice de A dans cette base est de la forme :

$$\begin{pmatrix} J_1 & & 0 \\ & \ddots & \\ 0 & & J_r \end{pmatrix}$$

où les J_k sont des blocs de Jordan. On peut aussi compter le nombre de blocs de Jordan : c'est le nombre de colonnes de même taille. Plus précisément, si $k \in \{1, \dots, m\}$, il y a exactement $\lambda_k - \lambda_{k+1}$ blocs de Jordan de taille k.

Théorème (Jordan)**.** *Le classe de similitude d'une matrice nilpotente est caractérisée par son diagramme de Young. Autrement dit, si A et B sont deux matrices nilpotentes, alors*

$$\mathscr{O}_A = \mathscr{O}_B \iff Y(A) = Y(B).$$

Preuve. Le sens direct est immédiat. Pour la réciproque, on vient de montrer que les matrices A et B était semblables à la même réduite de Jordan. □

Références.
H2G2

P. Lax, *Linear and Its Applications, Second Edition*
R. Mansuy, *Algèbre linéaire, Réduction des endomorphismes*

101 Groupe opérant sur un ensemble. Exemples et applications. *(euh...)*
150 Exemples d'actions de groupes sur les espaces de matrices.
157 Endomorphismes trigonalisables. Endomorphismes nilpotents.

1.5 Sur les invariants de similitude

On se placera toujours dans E, un $\mathbf{K}$-espace vectoriel de dimension sur un corps quelconque. Génériquement, u désignera un endomorphisme dans $\mathscr{L}(E)$ dont le polynôme minimal est noté Π_u et le polynôme caractéristique χ_u.

Quelques pré-requis.

Définition. Soit $u \in \mathscr{L}(E)$ et soit $x \in E$. On appelle *polynôme minimal de u en x* l'unique générateur unitaire de l'idéal

$$\{P \in \mathbf{K}[X],\ P(u)(x) = 0\}.$$

On le note $\Pi_{u,x}$. On a $\Pi_{u,x} | \Pi_u$.

Proposition. *Il existe $x \in E$ tel que $\Pi_u = \Pi_{u,x}$.*

Preuve. On écrit $\Pi_u = \prod_{i=1}^r P_i^{m_i}$ où P_i sont des irréductibles distincts. On note $K_i = \operatorname{Ker} P_i^{m_i}(u)$ et $u_i = u|_{K_i}$. Par le lemme des noyaux :

$$E = \oplus_i K_i.$$

Montrons le résultat sur chaque sous-espace K_i. Par l'absurde, si le résultat ne tenait pas, alors pour tout $x_i \in K_i$, Π_{u_i,x_i} diviserait strictement $\Pi_{u_i} = P_i^{m_i}$ donc diviserait $P_i^{m_i-1}$ par irréductibilité. Mais alors $P_i^{m_i-1}(u_i)$ serait nul sur tout K_i, ce qui est impossible par minimalité de Π_{u_i}. On dispose donc d'éléments x_i comme dans l'énoncé sur chaque sous-espace K_i. Montrons que $x = x_1 + \ldots + x_r$ convient. On a :

$$0 = \Pi_{u,x}(u)(x) = \sum_i \Pi_{u,x}(u)(x_i)$$

donc $\Pi_{u,x}(u)(x_i) = 0$ puisque les K_i sont en somme directe. Ainsi, $P_i^{m_i} = \Pi_{u_i,x_i} | \Pi_{u,x}$ pour tout i. Puisque les $P_i^{m_i}$ sont premiers entre eux, leur produit qui est égal à Π_u divise aussi $\Pi_{u,x}$, ce qui conclut. □

Ce qu'on va montrer.

Théorème. *Soit $u \in \mathscr{L}(E)$. Il existe une unique famille $P_1, \ldots, P_r$ de polynômes unitaires et une famille $E_1, \ldots, E_r$ de sous-espaces de E vérifiant :*

(i) $P_r | \ldots | P_1$

(ii) $E = E_1 \oplus \ldots \oplus E_r$

(iii) Pour tout $i \in \{1, \ldots, r\}$, E_i est stable par u et $u_{|E_i}$ est cyclique de polynôme P_i.

Les polynômes $P_1, \ldots . P_r$ sont appelés les invariants de similitudes de u.

Avant toute chose, remarquons que nécessairement $P_1 = \Pi_u$ car $P_1(u) = 0$ et $\Pi_u(u|_{E_1}) = 0$.

Preuve. Comme d'habitude, on procède par récurrence mais on ne l'écrit pas.

Existence. Soit $d = \deg(\Pi_u)$ et soit $x \in E$ tel que $\Pi_{u,x} = \Pi_u$. On note :

$$F = \mathrm{Vect}(x, u(x), \ldots, u^{d-1}(x)).$$

Bien sûr, F est stable par u et $u|_F$ est cyclique. On va montrer par dualité que F admet un supplémentaire stable par u. Soit $\varphi \in E^*$ tel que :

$$\varphi(x) = \varphi\big(u(x)\big) = \ldots = \varphi\big(u^{d-2}(x)\big) = 0 \quad \text{et} \quad \varphi\big(u^{d-1}(x)\big) = 1.$$

La famille $(\varphi, \varphi \circ u, \ldots, \varphi \circ u^{d-1})$ est une famille libre de E^* et on note Φ le sous-espace vectoriel de E^* engendré par cette famille. On pose alors :

$$G := \Phi^\circ = \{y \in E,\ \forall \psi \in \Phi,\ \psi(y) = 0\}$$

et on montre que c'est un supplémentaire de F stable par u. Il y a trois choses à voir :

- *G est u-stable.* Soit $y \in G$, alors par construction on a déjà :

  $$\forall k \in \{0, \ldots, d-2\},\ \varphi \circ u^k\big(u(y)\big) = 0.$$

  Comme le polynôme minimal de u est de degré d, on a :

  $$u^d(y) \in \mathrm{Vect}\left(y, u(y), \ldots, u^{d-1}(y)\right)$$

  et donc $\varphi \circ u^{d-1}\big(u(y)\big) = \varphi\big(u^d(y)\big) = 0$ par ce qui précède.
- $F \cap G = \{0\}$. Soit $y \in F \cap G$, alors on peut écrire :

  $$y = a_0 x + \ldots a_{d-1} u^{d-1}(x)$$

  et en appliquant $\varphi \circ u^i$ pour i allant de 0 à $d-1$, on trouve que tous les a_k sont nuls.
- $\dim F + \dim G = n$. C'est une propriété générale de l'orthogonal au sens de la dualité :

  $$\dim \Phi + \dim \Phi^\circ = n.$$

Et bien sûr, $\Pi_{u_{|G}} | \Pi_u$ puisque Π_u annule $u_{|G}$. À une récurrence près, on a achevé la preuve de l'existence.

Unicité. On suppose l'existence d'une autre famille de polynôme $Q_1, \ldots, Q_s$ donnant lieu à une autre décomposition $F_1 \oplus \ldots \oplus F_s$ comme dans l'énoncé. On a déjà $P_1 = Q_1 = \Pi_u$. Soit $j > 1$ l'indice minimal tel que $P_j \neq Q_j$ (il existe car les sommes des degrés des P_i et des Q_j sont égales). Alors, on a d'une part :

$$P_j(u)(E) = P_j(u)(E_1) \oplus \ldots \oplus P_j(u)(E_{j-1})$$

et d'autre part :

$$P_j(u)(E) = P_j(u)(F_1) \oplus \ldots \oplus P_j(u)(F_{j-1}) \oplus P_j(u)(F_j) \oplus \ldots \oplus P_j(u)(F_s).$$

Mais pour $i < j$, on a :

$$\dim P_j(u)(E_i) = \dim P_j(u)(F_i)$$

donc

$$0 = \dim P_j(u)(F_j) = \ldots = \dim P_j(u)(G_s)$$

ce qui prouve $Q_j | P_j$ et par symétrie $P_j | Q_j$. C'est absurde car $P_j \neq Q_j$. Finalement $r = s$ et $P_i = Q_i$ pour tout i. □

Corollaire (Décomposition de Frobénius)**.** *Soit $u \in \mathscr{L}(E)$. Il existe une base dans laquelle la matrice de u est de la forme*

$$\begin{pmatrix} C_{P_1} & & \\ & \ddots & \\ & & C_{P_r} \end{pmatrix}$$

où C_{P_i} est la matrice compagnon associée au polynôme P_i avec $P_r | \ldots | P_1$. De plus, on a

$$\chi_u = P_1 \ldots P_r.$$

Corollaire. *u et v sont semblables si et seulement s'ils ont les mêmes invariants de similitude.*

Preuve. Si u et v sont semblables, considérer $F_i = \varphi(E_i)$ où E_i sont les sous-espaces associés à u et φ tel que $\varphi \circ u = v \circ \varphi$. Ou alors, reprendre la preuve de l'unicité. □

Corollaire. *Soit $u \in \mathscr{L}(E)$. Alors u est semblable à sa transposée.*

Preuve. Il suffit de le montrer pour les endomorphismes cycliques. Le changement de base

$$e'_i = a_1 e_1 + \ldots a_{n-i} e_{n-i} + e_{n-i+1}$$

conduit au résultat. □

Corollaire (Décomposition de Jordan des endomorphismes nilpotents)**.** *Tout est dans le titre.*

Preuve. Puisque $\chi_u = X^n$, les invariants de similitudes sont de la forme X^{n_i}. □

Trucs à savoir.

- Les invariants de similitude ne dépendent pas du corps de base.
- La théorie des $\mathbf{K}[X]$-modules donne une façon simple pour calculer les invariants de similitude :

  Théorème. *Si U est la matrice de $u \in \mathscr{L}(E)$ dans une certaine base, alors les invariants de similitude de u sont les facteurs invariants non inversibles de la matrice $U - XI_n \in \mathcal{M}_n(\mathbf{K}[X])$.*

  Preuve. On montre par des opérations élémentaires sur les lignes et les colonnes qu'une matrice de la forme $C_P - XI$ est équivalente à

$$\begin{pmatrix} 1 & & & 0 \\ & \ddots & & \\ & & 1 & \\ 0 & & & P \end{pmatrix}$$

  et on utilise la décomposition de Frobenius pour conclure. □

Références.
H2G2
X. Gourdon, *Algèbre*
V. Beck, J, Malick, G. Peyré, *Objectif agrégation*

151 Dimension d'un espace vectoriel (on se limitera au cas de la dimension finie). Rang. Exemples et applications.
153 Polynômes d'endomorphisme en dimension finie. Réduction d'un endomorphisme en dimension finie. Applications.
154 Sous-espaces stables par un endomorphisme ou une famille d'endomorphisme d'un espace vectoriel de dimension finie. Applications.

1.6 L'anneau $\mathbf{Z}[i]$ et le théorème des deux carrés

On définit

$$\mathbf{Z}[i] := \{a + ib \in \mathbf{C},\ a, b \in \mathbf{Z}\}$$

l'anneau des *entiers de Gauss* muni de l'automorphsime de conjugaison et de la « norme » hérités de $\mathbf{C}$:

$$\begin{array}{ccc} \sigma : \mathbf{Z}[i] & \longrightarrow & \mathbf{Z}[i] \\ z = a + ib & \longmapsto & \overline{z} = a - ib \end{array} \quad \text{et} \quad \begin{array}{ccc} N : \mathbf{Z}[i] & \longrightarrow & \mathbf{N} \\ z = a + ib & \longmapsto & z\overline{z} = a^2 + b^2 \end{array}$$

De l'étude de $\mathbf{Z}[i]$, on va déduire le théorème des deux carrés dont le but est de préciser l'ensemble :

$$\Sigma := \{n \in \mathbf{N},\ \ n = a^2 + b^2,\ a, b \in \mathbf{N}\}.$$

Propriétés. *On liste ici les propriétés structurelles de* $\mathbf{Z}[i]$ *:*

(i) L'anneau $\mathbf{Z}[i]$ *est un anneau intègre.*

(ii) Les inversibles de $\mathbf{Z}[i]$ *sont connus :* $\mathbf{Z}[i]^\times = \{\pm 1, \pm i\} = \{z \in \mathbf{Z}[i],\ \ N(z) = 1\}$.

(iii) L'anneau $\mathbf{Z}[i]$ *est euclidien, donc principal, pour le stathme* N.

Preuve. Dans l'ordre :

(i) C'est un sous-anneau de $\mathbf{C}$ qui est intègre.

(ii) Si $z = a + ib \in \mathbf{Z}[i]^\times$, alors on note son inverse z' et puisque la norme est à valeurs dans $\mathbf{N}$:

$$N(z)N(z') = 1 \implies N(z) = N(z') = 1$$

de sorte que $a^2 + b^2 = 1$ et $z \in \{\pm 1, \pm i\}$. Réciproquement ces éléments sont bien inversibles.

(iii) C'est presque de l'analyse : si $z, t \in \mathbf{Z}[i] \setminus \{0\}$, alors on commence par écrire dans $\mathbf{C}$:

$$\frac{z}{t} = x + iy \in \mathbf{C} \text{ et } q = a + ib \in \mathbf{Z}[i] \text{ avec } |x - a| \leq \frac{1}{2} \text{ et } |y - b| \leq \frac{1}{2}.$$

Par construction, $|z/t - q| \leq \sqrt{1/4 + 1/4} = \sqrt{2}/2 < 1$ et le reste de la *division euclidienne* de z par t est :

$$r = z - qt \in \mathbf{Z}[i] \text{ et } |r| = |t||z/t - q| < |t|.$$

On a trouvé $q, r \in \mathbf{Z}[i]$ tels que $z = qt + r$ et $N(r) < N(t)$.

□

Le théorème des deux carrés est en fait peu ou prou une reformulation arithmétique de ce que sont les irréductibles de $\mathbf{Z}[i]$. Comme ça n'est pas l'objectif ici, on renvoie à la fin pour un théorème qui les décrit précisément (mais la preuve n'utilise rien de plus que ce qui va suivre). Notons qu'il ne faut pas confondre *nombre premier* et *entier vu dans dans* $\mathbf{Z}[i]$ *qui s'avère être premier dans cet anneau.* Comme l'anneau est euclidien donc factoriel, on ne parlera pas d'éléments premiers mais seulement d'irréductibles.

Lemme. *Soit* $p \in \mathbf{N}$ *un nombre premier. On a :*

$$p \in \Sigma \Longleftrightarrow p \text{ n'est pas irréductible dans } \mathbf{Z}[i].$$

Preuve. Si $p = a^2 + b^2$, alors $p = (a + ib)(a - ib)$ et $a, b \neq 0$ donc $p \in \mathbf{Z}[i]$ n'est pas irréductible. Réciproquement, si $p = zz'$ avec z, z' non inversibles, alors $N(p) = N(z)N(z') = p^2$ et nécessairement $N(z) = N(z') = p$ donc $p \in \Sigma$. □

Théorème. *Soit* $p \in \mathbf{N}$ *un nombre premier. On a :*

$$p \in \Sigma \iff p = 2 \text{ ou } p \equiv 1 \ (mod.\ 4).$$

Preuve. La condition est nécéssaire car un carré modulo 4 vaut 0 ou 1 et $2 = 1+1$. Il reste à montrer le sens direct. Compte-tenu du lemme précédent, il suffit de supposer que p n'est pas irréductible, c'est à dire, dans un anneau factoriel, que l'idéal (p) n'est pas premier, ou encore que le quotient $\mathbf{Z}[i]/(p)$ n'est pas intègre. D'abord, on se convainc que :

$$\mathbf{Z}[i] \simeq \mathbf{Z}[X]/(X^2 + 1)$$

Ensuite, on utilise un théorème d'isomorphisme pour montrer :

$$\mathbf{Z}[i]/(p) \simeq \frac{\mathbf{Z}[X]}{(X^2+1, p)} \simeq \frac{\mathbf{Z}[X]/(p)}{(X^2+1)} \simeq \frac{\mathbf{Z}/p\mathbf{Z}[X]}{(X^2+1)}.$$

Et dire que ce dernier anneau n'est pas intègre signifie que $X^2 + 1$ n'est pas irréductible dans $\mathbf{F}_p[X]$, ce qui équivalent (c'est un polynôme de degré 2) à dire que que $X^2 + 1$ a une racine dans $\mathbf{F}_p$. Finalement,

$$p \in \Sigma \iff -1 \in \mathbf{F}_p^{\times 2}.$$

Mais on connait une caractérisation des carrés dans les corps finis : si $p > 2$

$$-1 \in \mathbf{F}_p^{\times 2} \iff (-1)^{\frac{p-1}{2}} = 1 \iff \frac{p-1}{2} \text{ pair} \iff p \equiv 1 \text{ (mod. 4)}.$$

et bien sûr -1 est un carré dans $\mathbf{F}_2$. □

On vient de terminer la description de l'ensemble des des nombres premiers appartenant à Σ. Le théorème des deux carrés dans toute sa généralité en découle, modulo la factorialité de $\mathbf{Z}$.

Théorème (des deux carrés)**.** *Soit* $n \in \mathbf{N}$, $n \geq 2$. *Alors*

$$n \in \Sigma \iff v_p(n) \text{ pair pour } p \equiv 3 \ (mod.\ 4).$$

Preuve. Comme Σ est stable par multiplication et qu'un carré est toujours dans Σ, le théorème précédent donne le sens réciproque. Pour le sens direct, on fixe $p \equiv 3 \pmod 4$ et on procède par récurrence sur $v_p(n)$. En voici les arguments :

- Si $v_p(n) = 0$ alors c'est fini.
- Si $v_p(n) \geq 1$, alors $p|a^2 + b^2 = (a + ib)(a - ib)$. Mais par le théorème précédent, $p \notin \Sigma$ et le lemme indique que p est irréductible dans $\mathbf{Z}[i]$. Disons par exemple que p divise $a + ib$ dans $\mathbf{Z}[i]$.
- Comme p est entier, $p|a$ et $p|b$ donc $p^2|n$ et on peut même écrire :

$$\frac{n}{p^2} = a'^2 + b'^2 \in \Sigma \quad \text{et} \quad v_p\left(\frac{n}{p^2}\right) = v_p(n) - 2 \text{ est pair par hypothèse de récurrence.}$$

Finalement $v_p(n)$ est aussi pair et on a fini.

□

Quelques compléments et précisions.

- D'abord, on a par les mêmes arguments une description précise des irréductibles de $\mathbf{Z}[i]$:

  Théorème (Irréductibles de $\mathbf{Z}[i]$)**.** *Les irréductibles de $\mathbf{Z}[i]$ sont aux inversibles près :*

  (i) Les nombres premiers $p \in \mathbf{N}$ avec $p \equiv 3$ (mod. 4).

  (ii) Les entiers de Gauss $a + ib$ dont la norme $a^2 + b^2$ est un nombre premier p. Dans ce cas, on a $p = 2$ ou $p \equiv 1$ (mod. 4).

  Notons qu'il existe une preuve un peu plus pédestre dans [Jeanneret, Lines].

- Le théorème d'isomorphisme en question dans la preuve est le suivant :

  Théorème (d'isomorphisme)**.** *Soient A un anneau commutatif unitaire et I un idéal de A. Alors, les idéaux de A/I sont de la forme J/I où J est un idéal de A tel que $I \subset J$ et dans ce cas on a un isomorphisme :*

$$\frac{A/I}{J/I} \simeq \frac{A}{J}.$$

  La preuve repose sans doute sur le théorème de correspondance des idéaux qui donne une bijection entre l'ensemble des idéaux de A contenant I et l'ensemble des idéaux de A/I. Ici, on applique le théorème avec $A = \mathbf{Z}[X]$ et $I = (p) \subset (X^2 + 1, p) = J$ et $I = (X^2 + 1) \subset (X^2 + 1, p) = J$.

- On a aussi passé sous silence le fait que $Z[X]/(p) \simeq \mathbf{F}_p[X]$, qui repose sur le théorème d'isomorphisme « standard ».
- La caractérisation des carrés dans les corps fini repose sur l'étude du nombre de racines du polynôme $X^{\frac{q-1}{2}} - 1$.

Références.

D. Perrin, *Cours d'Algèbre*
A. Jeanneret, D. Lines, *Invitation à l'Algèbre*

120 Anneaux $\mathbf{Z}/n\mathbf{Z}$. Applications.
121 Nombres premiers. Applications.
122 Anneaux principaux. Applications.
126 Exemples d'équations diophantiennes.

1.7 La décomposition de Dunford-Newton

On se place dans E un $\mathbf{K}$-espace vectoriel de dimension finie d où $\mathbf{K}$ est un corps de caractéristique nulle.

Théorème (Dunford)**.** *Soit $u \in \mathscr{L}(E)$ dont le polynôme caractéristique est scindé. Alors il existe un unique couple (d, n) d'endomorphsimes de E tels que d est diagonalisable, n est nilpotent, d et n commutent et $u = d + n$. De plus d et n sont des polynômes en u.*

La construction de d et n est effective et ne nécessite pas le calcul des valeurs propres de u. Plus précisément, on va montrer :

Théorème (Dunford-Newton)**.** *La suite d'endomorphismes définie par :*

$$\left\{ \begin{array}{rcl} u_{n+1} & = & u_n - P(u_n)P'(u_n)^{-1} \\ u_0 & = & u \end{array} \right. \qquad où \quad P = \frac{\chi_u}{\chi_u \wedge \chi_u'}$$

est bien définie et stationne vers d qui vérifie l'énoncé du théorème précédent.

Preuve. On montre par récurrence les propriétés suivantes pour tout $n \in \mathbf{N}$

$$(i)\ P'(u_n) \in GL_d(\mathbf{K}) \qquad (ii)\ P(u_n) \in P(u)^{2^n}\mathbf{K}[u] \qquad (iii)\ u_n \in \mathbf{K}[u]$$

- Pour $n = 0$, il n'y a que le (i) à montrer et cela vient de la définition de P. Par construction, c'est un polynôme scindé (car χ_u l'est) à racines simples. En fait, en caractéristique nulle :
$$\chi_u = \prod (X - \lambda_i)^{m_i} \implies P = \prod (X - \lambda_i).$$
Bien sûr, $P \wedge P' = 1$ donc $P^m \wedge P' = 1$ pour tout $m \in \mathbf{N}$. Mais compte-tenu de la forme de P, il existe $m \in \mathbf{N}$ tel que $\chi_u | P^m$. Le théorème de Bézout donne pour un certain $A \in \mathbf{K}[X]$:
$$A(u)P'(u) = I_d \quad i.e \quad P'(u) \in GL_d(\mathbf{K}).$$
- Supposons le résultat vrai jusqu'au rang n, $n \geq 0$. La première chose à voir est que u_{n+1} est bien défini et est un polynôme en u. Puis on commence par écrire la formule de Taylor pour les polynômes : il existe $Q \in \mathbf{K}[X, Y]$ tel que :
$$P(X + Y) = P(X) + YP'(X) + Y^2 Q(X, Y).$$
On trouve :
$$\begin{array}{rcl} P(u_{n+1}) & = & P\big(u_n + (u_{n+1} - u_n)\big) \\ & = & P(u_n) + (u_{n+1} - u_n)P'(u_n) + (u_{n+1} - u_n)^2 Q(u_n, u_{n+1} - u_n) \\ & = & P(u_n) - P(u_n) + P(u_n)^2 P'(u_n)^{-2} Q(u_n, u_{n+1} - u_n) \\ & = & P(u)^{2^{k+1}} \Big[P'(u_n)^{-2} Q(u_n, u_{n+1} - u_n)] \in P(u)^{2^{n+1}}\mathbf{K}[u] \end{array}$$

car l'inverse est un polynôme. On vient de montrer (ii). Pour montrer (i), il suffit d'écrire, toujours avec la même formule mais à l'ordre 1 :

$$P'(u_{n+1}) = P'(u_n) + (u_{n+1} - u_n)R(u_n)$$

puis de constater que $u_{n+1} - u_n = P(u_n)P'(u_n)^{-1}$ est nilpotent car $P(u_n)$ l'est car $u_n \in P(u)^{2^n}\mathbf{K}[u]$ et $P(u)$ est nilpotent. Finalement,

$$P'(u_{n+1}) = \text{ inversible} + \text{nilpotent}, \quad \text{les deux commutant.}$$

Il ne reste plus qu'à voir la stationnarité de la suite : c'est la même argument, à partir d'un certain rang pour lequel $\chi_u | P^{2^n}$:

$$P(u)^{2^n} = P^{2^n}(u) = 0 \implies P(u_n) = 0 \implies u_{n+1} = u_n.$$

Notons $d = u_N$ la limite comme dans l'énoncé. C'est un polynôme en u et $P(d) = 0$ donc d est diagonalisable car annulé par un polynôme scindé à racines simples. De plus,

$$n = u - d = \sum_{n=0}^{N-1} u_n - u_{n+1} = -\sum_{n=0}^{N-1} P(u_n)P'(u_n)^{-1}$$

est une somme d'endomorphismes nilpotents qui commutent donc est nilpotent. Bien sûr c'est aussi un polynôme en u donc commute avec d. □

Preuve (unicité). S'il existait un autre couple (d', n') comme dans l'énoncé alors comme tout le monde est un polynôme en u, tout le monde commute avec tout le monde et en particulier d et d' sont co-diagonalisbales donc $d - d' = n - n'$ est diagonalisable. Mais $n - n'$ est un endomorphisme nilpotent comme somme d'endomorphismes nilpotents qui commutent. Le seul endomorphisme nilpotent diagonalisable étant 0, on obtient $n = n'$ puis $d = d'$. □

Référence. ?

153 Polynômes d'endomorphisme en dimension finie. Réduction d'un endomorphisme en dimension finie. Applications.
154 Sous-espaces stables par un endomorphisme ou une famille d'endomorphismes d'un espace vectoriel de dimension finie. Applications.
155 Endomorphismes diagonalisables en dimension finie.
157 Endomorphismes trigonalisables. Endomorphismes nilpotents.

1.8 La transformée de Fourier rapide

Dans $\mathbf{C}^N$, on considère

$$\omega \equiv \omega_N = \exp\left(-\frac{2i\pi}{N}\right)$$

et si $f = (f(0), \dots, f(N-1))^T \in \mathbf{C}^N$, on appelle $\hat{f} \in \mathbf{C}^N$ sa transformée de Fourier discrète définie par :

$$\forall k \in \{0, \dots, N-1\}, \quad \hat{f}(k) = \sum_{n=0}^{N-1} f(n)\omega^{kn}.$$

On appelle $F_N \in \mathcal{M}_N(\mathbf{C})$ la matrice :

$$\begin{pmatrix} 1 & 1 & 1 & \dots & 1 \\ 1 & \omega & \omega^2 & \dots & \omega^{N-1} \\ 1 & \omega^2 & \omega^4 & \dots & \omega^{2(N-1)} \\ \vdots & & & & \end{pmatrix}$$

ce qui permet d'écrire de façon plus concise :

$$\hat{f} = F_N f.$$

Théorème. *Il existe un algorithme permettant de calculer le vecteur $\hat{f}$ à partir du vecteur f en $\mathcal{O}(N \log N)$ opérations.*

Preuve. Quitte à rajouter quelques zéros, on peut supposer que N est pair et même que c'est une puissance de deux. La grande idée est de réordonner les colonnes de F_N : d'abord celles d'indice pair et ensuite les autres. Autrement dit, en notant en colonnes $F_N = (e_0, \dots, e_{N-1})$, on regarde

$$F_N^r = (e_0, \dots, e_{N-2}, e_1, \dots, e_{N-1}).$$

Comme $\omega^N = 1$ et $\omega^{N/2} = -1$, la matrice F_N^r s'écrit :

$$F_N^r = \left(\begin{array}{cccc|cccc} 1 & 1 & \dots & 1 & 1 & 1 & \dots & 1 \\ 1 & \omega^2 & \dots & \omega^{N-2} & \omega & \omega^3 & \dots & \omega^{N-1} \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 1 & \omega^{N-2} & \dots & & \omega^{\frac{N}{2}-1} & & & \\ \hline 1 & 1 & \dots & 1 & -1 & \dots & & -1 \\ 1 & \omega^2 & \dots & \omega^{N-2} & -\omega & \dots & & -\omega^{N-1} \\ \vdots & \vdots & & \vdots & \vdots & & & \vdots \\ 1 & \omega^{N-2} & \dots & & -\omega^{\frac{N}{2}-1} & \dots & & \end{array}\right)$$

Et comme $\omega_N^2 = \omega_{N/2}$ on peut finalement écrire :

$$F_N^r = \left(\begin{array}{c|c} F_{N/2} & B \\ \hline F_{N/2} & -B \end{array}\right)$$

En plus, en factorisant la ligne $k \in \{1, \dots, N/2\}$ de B par ω^{k-1}, on trouve :

$$B = DF_{N/2} \quad \text{avec} \quad D = \operatorname{diag}(1, \dots, \omega^{\frac{N}{2}-1}).$$

Finalement, avec les vecteurs réordonnés :

$$f^r = \begin{pmatrix} f_{\text{pair}} \\ f_{\text{impair}} \end{pmatrix} := \begin{pmatrix} f(0) \\ \vdots \\ f(N-2) \\ f(1) \\ \vdots \\ f(N-1) \end{pmatrix} \quad \text{et} \quad \hat{f} = \begin{pmatrix} \hat{f}^1 \\ \hat{f}^2 \end{pmatrix} := \begin{pmatrix} \hat{f}(0) \\ \vdots \\ \hat{f}(N/2-1) \\ \hat{f}(N/2) \\ \vdots \\ \hat{f}(N-1) \end{pmatrix}$$

on doit calculer :

$$\hat{f} = F_N^r f^r$$

ce qui s'écrit pr blocs :

$$\begin{pmatrix} \hat{f}^1 \\ \hat{f}^2 \end{pmatrix} = \left(\begin{array}{c|c} F_{N/2} & DF_{N/2} \\ \hline F_{N/2} & -DF_{N/2} \end{array}\right) \begin{pmatrix} f_{\text{pair}} \\ f_{\text{impair}} \end{pmatrix}$$

C'est fini :

$$\begin{aligned} \hat{f}^1 &= F_{N/2} f_{\text{pair}} + DF_{N/2} f_{\text{impair}} \\ \hat{f}^2 &= F_{N/2} f_{\text{pair}} - DF_{N/2} f_{\text{impair}}. \end{aligned}$$

On voit donc qu'il suffit de calculer deux transformée de Fourier de taille $N/2$ et il ne reste plus qu'à compter : si $C(N)$ est le nombre de multiplications nécessaires au calcul du vecteur $\hat{f}$ à partir de f, on voit :

$$C(N) = 2C(N/2) + \mathcal{O}(N)$$

et on en déduit en écrivant $N = 2^p$ et $C'(N) = C(N)/N$:

$$C(N) = \mathcal{O}(N \log_2 N).$$

□

Une application rapide et élégante de ce résultat est le calcul et l'étude de la stabilité du θ-schéma pour l'équation de la chaleur.

Références.

P. D. Lax, *Linear Algebra and its Applications, Second Edition*
G. Peyré, *L'Algèbre Discrète de la Transformée de Fourier*

102 Groupe des nombres complexes de module 1. Sous-groupes des racines de l'unité. Applications.
110 Caractères d'un groupe abélien fini et transformée de Fourier discrète. Applications.

1.9 Le lemme d'Artin

Lemme (Dedekind)**.** *Soient $n \geq 1$ et $\varphi_1, \dots \varphi_n : \mathbf{L} \to \mathbf{K}$ des homorphismes de corps distincts entre les corps $\mathbf{K}$ et $\mathbf{L}$. Alors le système engendré par les $\varphi_1, \dots, \varphi_n$ est libre :*

$$\left[\forall x \in \mathbf{K},\ \sum_{i=1}^{n} \alpha_i \varphi_i(x) = 0\ \right] \implies \alpha_1 = \ldots = \alpha_n = 0.$$

Preuve. Raisonnons par l'absurde. Comme un homomorphisme de corps n'est jamais nul, on considère $r \geq 2$ le nombre minimal tel que, quitte à re-numéroter les φ_i, il existe une relation de dépendance linéaire :

$$\forall x \in \mathbf{K},\ \sum_{i=1}^{r} \alpha_i \varphi_i(x) = 0. \tag{1.1}$$

Puisque $\varphi_1 \neq \varphi_r$, il existe $y \in \mathbf{K}$ tel que $\varphi_1(y) \neq \varphi_r(y)$. On a :

$$\forall x \in \mathbf{K},\ \sum_{i=1}^{r} \alpha_i \varphi_i(yx) = \sum_{i=1}^{r} \alpha_i \varphi_i(y)\varphi_i(x) = 0. \tag{1.2}$$

Il suffit d'opérer : (1.2)$\leftarrow$(1.2)$-\varphi_1(y)\times$(1.1) et on conclut :

$$\sum_{i=2}^{r} \alpha_i (\varphi_1(y) - \varphi_i(y))\varphi_i = 0$$

ce qui contredit la minimalité de r. □

Lemme (Artin)**.** *Soient $\mathbf{L}$ est un corps et H un sous-groupe fini de $\mathrm{Aut}(L)$. Alors l'extension $\mathbf{L}/\mathbf{L}^H$ est un extension finie de degré $[\mathbf{L} : \mathbf{L}^H] = |H|$.*

Preuve. On note une bonne fois pour toutes :

$$m = [\mathbf{L} : \mathbf{L}^H] \in \mathbf{N} \cup \{\infty\}, \quad n = |H| < \infty \quad \text{et} \quad H = \{\sigma_1, \dots, \sigma_n\}.$$

Étape 1. On montre que $n \leq m$.

Par l'absurde, on suppose que $m < n < \infty$ et on considère $x_1, \dots, x_m$ une $\mathbf{L}^H$-base de $\mathbf{L}$. La matrice des $(\sigma_j(x_i))_{i,j}$ est de taille $m \times n$ avec $m < n$ donc il existe une solution $(y_1, \dots, y_n)$ non nulle au système :

$$\forall i \in \{1, \dots, m\},\ \sum_{j=1}^{n} \sigma_j(x_i) y_j = 0.$$

Comme les $x_1, \dots, x_m$ forment une $\mathbf{L}^H$-base de $\mathbf{L}$ on a obtenu une relation de dépendance linéaire entre les σ_j car si $x = \sum_{i=1}^{m} \alpha_i x_i \in \mathbf{L}$ avec $\alpha_i \in \mathbf{L}^H$, on a :

$$\sum_{j=1}^{n} y_j \sigma_j(x) = \sum_{j=1}^{n} \sum_{i=1}^{m} y_j \alpha_i \sigma_j(x_i) = \sum_{i=1}^{m} \alpha_i \sum_{j=1}^{m} \sigma_j(x_i) y_j = 0.$$

Ce qui contredit le lemme de Dedekind. Donc $n \leq m$.

Étape 2. On montre que $m \leq n$.

Toujours par l'absurde, on suppose que $n < m \in \mathbf{N} \cup \{\infty\}$. C'est dire qu'il existe $n+1$ éléments $x_1, \ldots, x_{n+1} \in \mathbf{L}$ linéairement indépendants sur $\mathbf{L}^H$. Comme tout à l'heure, il existe une solution $(y_1, \ldots, y_{n+1})$ non nulle au système :

$$\forall i \in \{1, \ldots, n\},\ \sum_{j=1}^{n+1} \sigma_i(x_j) y_j = 0.$$

Quitte à re-numéroter les y_j, on peut supposer que le système est de rang r et se ré-écrit :

$$\forall i \in \{1, \ldots, n\},\ \sum_{j=1}^{r} \sigma_i(x_j) y_j = 0. \tag{1.3}$$

En faisant opérer $\sigma \in H$ sur ce système, ce système est équivalent à :

$$\forall i \in \{1, \ldots, n\},\ \sum_{j=1}^{r} \sigma_i(x_j) \sigma(y_j) = 0. \tag{1.4}$$

Maintenant on écrit $\sigma(y_1) \times (1.3) - y_1 \times (1.4)$ et on obtient :

$$\forall i \in \{1, \ldots, r\},\ \sum_{j=2}^{r} \sigma_i(x_j)(y_j \sigma(y_1) - \sigma(y_j) y_1) = 0.$$

Par minimalité de r, on en déduit :

$$\forall j \in \{2, \ldots, r\},\ y_j \sigma(y_1) - \sigma(y_j) y_1 = 0 \quad \textit{i.e.}\ \ y_j y_1^{-1} \in \mathbf{L}^H.$$

On peut conclure : en notant $y_j = y_1 z_j$ avec $z_j \in \mathbf{L}^H$, le système (1.3) donne pour l'indice i tel que $\sigma_i = id_{\mathbf{L}}$:

$$\sum_{j=1}^{r} x_j y_1 z_j = 0 \ \text{ donc } \ \sum_{j=1}^{r} x_j z_j = 0$$

car $y_1 \neq 0$. Mais c'est impossible car les x_j sont linéairement indépendants. □

Corollaire. *Soient* $\mathbf{L}$ *un corps et* H *un sous-groupe fini de* $\mathrm{Aut}(\mathbf{L})$. *Alors* $\mathbf{L}/\mathbf{L}^H$ *est finie et*

$$H = \mathrm{Aut}(\mathbf{L}/\mathbf{L}^H).$$

Preuve. On note $G = \mathrm{Aut}(\mathbf{L}/\mathbf{L}^H)$. On a déjà $H \subset G$.

(1) Comme $\mathbf{L}/\mathbf{L}^H$ est une extension finie, G est un groupe fini. En effet, si $a_1, \ldots, a_n$ est une $\mathbf{L}^H$-base de $\mathbf{L}$, on peut considérer

$$P = P_1 \ldots P_r$$

où les P_i sont les polynômes minimaux des a_i sur $\mathbf{L}^H$. On appelle R l'ensemble des racines de P contenue dans $\mathbf{L}$. Alors l'application :

$$\Psi : G \longrightarrow \mathrm{Bij}(R),\ \ \sigma \longmapsto \sigma|_R$$

est un homomorphisme injectif et comme R est fini, il en est de même pour $\mathrm{Bij}(R)$ et pour G.

(2) Regardons les inclusions :

$$\mathbf{L}^G \subset \mathbf{L}^H \subset \mathbf{L} \quad \text{et} \quad \mathbf{L}^H \subset \mathbf{L}^G \subset \mathbf{L}$$

où le deuxième triptyque d'inclusions provient du fait que $\mathbf{L}^{\mathrm{Aut}(\mathbf{L}/\mathbf{L}^H)}$ est un corps intermédiaire (c'est presque la définition).

(3) Finalement, $\mathbf{L}^G = \mathbf{L}^H$ et par le théorème précédent :

$$|G| = [\mathbf{L} : \mathbf{L}^G] = [\mathbf{L} : \mathbf{L}^H] = |H|$$

et comme $H \subset G$, on conclut $G = H$.

□

Sur la théorie de Galois.

Le lemme d'Artin est une étape cruciale du théorème de correspondance de Galois :

Théorème (Galois)**.** *Soit* $\mathbf{L}/\mathbf{K}$ *une extension finie galoisienne*[2]*. Alors, entre autres choses, il y a une bijection*

$$\{\textit{corps intermédiaire de } \mathbf{L}/\mathbf{K}\} \longrightarrow \{\textit{sous-groupe de } \mathrm{Gal}(\mathbf{L}/\mathbf{K})\}$$

donnée par :

$$\mathit{Gal} : \mathbf{M} \mapsto \mathrm{Gal}(\mathbf{L}/\mathbf{M}) \quad \textit{et} \quad \mathit{Fix} : H \mapsto \mathbf{L}^H.$$

Le corollaire du lemme d'Artin dit que $\mathit{Gal} \circ \mathit{Fix}(H) = H$. Dans l'autre sens c'est plus un peu plus conceptuel mais c'est aussi plus simple : il suffit de voir que si $\mathbf{K} \subset \mathbf{M} \subset \mathbf{L}$ alors $\mathbf{L}/\mathbf{M}$ est une extension galoisienne[3] donc $\mathbf{L}^{\mathrm{Gal}(\mathbf{L}/\mathbf{M})} = \mathbf{M}$, *i.e.* $\mathit{Fix} \circ \mathit{Gal}(\mathbf{M}) = \mathbf{M}$.

Après, il s'agit aussi de montrer le lien entre les sous-groupes normaux de $\mathrm{Gal}(\mathbf{L}/\mathbf{K})$ et les corps intermédiaires $\mathbf{M}$ tels que $\mathbf{M}/\mathbf{K}$ est galoisienne.

Référence. A. Jeanneret, D. Lines, *Invitation à l'Algèbre*

125 Extensions de corps. Exemples et applications.
151 Dimension d'un espace vectoriel (on se limitera au cas de la dimension finie). Rang. Exemples et applications.
162 Systèmes d'équations linéaires ; opérations élémentaires, aspects algorithmiques et conséquences théoriques.

2. C'est à dire une extension normale (tout polynôme irréductible dans $\mathbf{K}[X]$ qui a une racine dans $\mathbf{L}$ a toutes ses racines dans $\mathbf{L}$) et séparable (tout polynôme irréductible dans $\mathbf{K}$ ou dans $\mathbf{L}$ n'a que des racines simples dans son corps des racines). C'est équivalent à dire que $\mathbf{L}/\mathbf{K}$ est algébrique et

$$\mathbf{L}^{\mathrm{Gal}(\mathbf{L}/\mathbf{K})} = \mathbf{K}$$

où on note plus volontier $\mathrm{Aut}(\mathbf{L}/\mathbf{K}) \equiv \mathrm{Gal}(\mathbf{L}/\mathbf{K})$. C'est aussi équivalent à dire que $\mathbf{L}$ est le corps des racines d'un polynôme de $\mathbf{K}[X]$ ou encore que $|\mathrm{Gal}(\mathbf{L}/\mathbf{K})| = [\mathbf{L} : \mathbf{K}]$.

3. Si $\mathbf{M}$ est un corps intermédiaire de $\mathbf{L}/\mathbf{K}$, alors, puisque $\mathbf{L}/\mathbf{K}$ est galoisienne, $\mathbf{L}$ est le corps des racines d'un certain $P \in \mathbf{K}[X]$. C'est aussi le corps des racines de ce même polynôme vu dans $\mathbf{M}[X]$ donc $\mathbf{L}/\mathbf{M}$ est galoisienne.

1.10 Le théorème de structure des groupes abéliens finis

On considère un groupe abélien fini G et on rappelle que son dual est le groupe des caractères $\widehat{G} = \mathrm{Hom}(V, \mathbf{C}^\times)$ muni de la multiplication sur les valeurs. On rappelle aussi que la terminologie *caractère* est un abus dans ce contexte.

Quelques prérequis.

Lemme. *Un groupe est fini si et seulement si toutes ses représentations irréductibles sont de dimension 1, c'est à dire que son groupe des caractères est égal à l'ensemble de ses caractères irréductibles.*

Lemme. *On a l'égalité des cardinaux* $|G| = |\widehat{G}|$.

Preuve. Il suffit d'écrire la suite d'égalités :

$$|G| = |\operatorname{Conj} G| = |\operatorname{Irr} G| = |\widehat{G}|.$$

On peut aussi utiliser le lemme de prolongement des caractères de G. Peyré. □

Lemme. *L'application de bidualité* $\mathrm{ev} : G \to \widehat{\widehat{G}}$ *défini par* $\mathrm{ev}(x)(\chi) = \chi(x)$ *est un isomorphisme de groupes.*

Preuve. On vérifie que ev est un morphisme et on montre sa bijectivité. On déduit du lemme précédent que $|G| = |\widehat{\widehat{G}}|$ et on montre l'injectivité de ev. Il suffit de prendre $g \in \operatorname{Ker} \mathrm{ev}$ et d'écrire :

$$\delta_g = \sum_{\chi \in \widehat{G}} \langle \delta_g, \chi \rangle \chi = \ldots = \frac{1}{|G|} \sum_{\chi \in \widehat{G}} \chi$$

de sorte que $\delta_g(e) = 1$ et $g = e$. □

Lemme. *Les groupes* G *et* $\widehat{G}$ *ont même exposant.*

Preuve. On note $N(G)$ l'exposant de G. Soit $\chi \in \widehat{G}$. On a pour tout $x \in G$:

$$\chi^{N(G)}(x) = \chi(x)^{N(G)} = \chi(x^{N(G)}) = \chi(1) = 1$$

donc $\chi^{N(G)} = 1$ et l'exposant de $\widehat{H}$ divise celui de H. En faisant pareil avec $\widehat{G}$ et $\widehat{\widehat{G}}$ et puisque $G \simeq \widehat{\widehat{G}}$, on a le résultat. □

Ce qu'on va montrer.

Théorème. *Soit* G *un groupe abélien fini. Il existe* $r \in \mathbf{N}$ *et des entiers* $n_r | \ldots | n_1$ *tels que*

$$G \simeq (\mathbf{Z}/n_1\mathbf{Z}) \times \ldots (\mathbf{Z}/n_r\mathbf{Z}).$$

Preuve. On procède par récurrence sur $|G|$. C'est bon avec $r = 0$ pour $|G| = 1$ et si $|G| > 1$ alors notons $n = n_1$ l'exposant de G.

(1) Pour tout $\chi \in \widehat{G}$ et tout $x \in G$, $\chi(x)$ est une racine n-ème de l'unité. De plus, comme n est aussi l'ordre de $\widehat{G}$, il existe [4] $\chi_1 \in \widehat{G}$ d'ordre n. Finalement, $\chi_1(G) \subset \mathbf{U}_n$ et il existe $x_1 \in G$ tel que
$$\chi_1(x_1) = e^{2i\pi/n}.$$
L'ordre de x_1 est aussi n et le sous-groupe de $H_1 \subset G$ engendré par x_1 est isomorphe à $\mathbf{Z}/n\mathbf{Z}$.

(2) On va montrer que $G \simeq H_1 \times G_1$ où $G_1 = \operatorname{Ker} \chi_1$. On commence par voir que χ_1 induit un isomorphisme $H_1 \to \mathbf{U}_n$. En effet, c'est un morphisme surjectif entre deux groupes de même cardinal . Son inverse sera noté
$$\alpha : \mathbf{U}_n \to H_1.$$

(3) Soit $x \in G$. On définit :
$$a = \alpha(\chi_1(x)) \quad \text{et} \quad b = a^{-1}x.$$
En particulier
$$\chi_1(b) = \chi_1(a)^{-1}\chi_1(x) = 1$$
donc $b \in G_1$ et tout élément de $x \in G$ peut s'écrire $x = ab$ où $a \in H_1$ et $b = G_1$.

(4) Par injectivité de χ_1, il est clair que $H_1 \cap G_1 = \{1\}$ et on peut conclure
$$G \simeq H_1 \times G_1.$$

(5) Puisque l'exposant d'un sous-groupe divise celui du groupe et que G_1 est isomorphe à un sous-groupe de G, la relation de divisibilité à lieu et on peut terminer la récurrence.

□

Références.
P. Colmez, *Éléments d'analyse et d'algèbre (et de théorie des nombres)*
G. Peyré, *L'Algèbre discrète de la Transformée de Fourier*

104 Groupes finis. Exemples et applications.
107 Représentations et caractères d'un groupe fini sur un C-espace vectoriel. Exemples.
110 Caractères d'un groupe abélien fini et transformée de Fourier discrète. Applications.

1.11 Le théorème de structure des polynômes symétriques

On se place dans un anneau intègre A et l'on définit d'abord plusieurs notions :

(1) Le **poids** du monôme $aX_1^{i_1} \dots X_m^{i_m} \in A[X_1, \dots, X_m]$ est $i_1 + 2i_2 + \dots mi_m$. Le poids d'un polynôme est le maximum des poids des monômes.

4. Dans un groupe **abélien** fini, il existe un élément d'ordre l'exposant du groupe. Pour le voir il suffit de montrer que si a est d'ordre m et b est d'ordre n alors il existe un élément d'ordre $m \vee n$. Lorsque $m \wedge n = 1$, ab convient. Dans le cas contraire, on considère $m' \wedge n' = 1$ tels que $m'|m$, $n'|n$ et $m'n' = m \vee n$. Ensuite on voit que $a^{\frac{m}{m'}} b^{\frac{n}{n'}}$ est d'ordre $m'n'$.

(2) Dans $A[t_1, \ldots, t_n]$, la **k-ème fonction symétrique élémentaire** est la fonction :

$$s_k(t_1, \ldots, t_n) = \sum_{1 \leq i_1 < i_2 < \ldots < i_k \leq i_n} t_{i_1} t_{i_2} \ldots t_{i_k}.$$

(3) Pour $j = 0, \ldots, n-1$, s_{n-j} est le coefficient devant le monôme d'ordre j du **polynôme général de degré** n :

$$F(X) = (X - t_1)(X - t_2) \ldots (X - t_n) \in A[t_1, \ldots, t_n][X].$$

(4) On appelle **polynôme symétrique** de $A[t_1, \ldots, t_n]$ tout polynôme invariant par l'action de $\mathfrak{S}_n$ par permutation des indéterminés. L'ensemble des polynômes symétriques est noté $\mathcal{S}$ et forme un sous-anneau de $A[t_1, \ldots, t_n]$.

Comme les fonctions symétriques élémentaires sont des polynômes symétrique, $\mathcal{S}$ contient le sous-anneau qu'elles engendrent. En fait, la réciproque est vraie.

Théorème. *Tout polynôme symétrique à coefficients dans A est un polynôme en les fonctions symétriques élémentaires. Autrement dit,*

$$\mathcal{S} = A[s_1, \ldots, s_n].$$

Preuve. Soit P un polynôme symétrique en les indéterminés $t_1, \ldots, t_n$. On va montrer par double récurrence sur n et sur $d = \deg P$ le fait suivant :

$\mathscr{P}(n, d)$: Il existe un polynôme $Q \in A[X_1, \ldots, X_n]$ de poids $\leq d$ tel que :

$$P(t_1, \ldots, t_n) = Q(s_1(t_1, \ldots, t_n), \ldots, s_n(t_1, \ldots, t_n)).$$

On va montrer par récurrence sur n la proposition « $\forall \tilde{d} \in \mathbf{N},\ \mathscr{P}(n, \tilde{d})$ ».

Pour $n = 1$, il n'y a rien à montrer puisque $s_1 = t_1$.

Soit $n \geq 2$. On suppose « $\forall \tilde{d},\ \mathscr{P}(n-1, \tilde{d})$ ». On montre par récurrence sur $d \in \mathbf{N}$ la propriété $\mathscr{P}(n, d)$. Pour $d = 0$, $\mathscr{P}(n, 0)$ est toujours vraie. Soit $d \geq 1$, on suppose $\mathscr{P}(n, d-1)$.

Soit $P \in A[t_1, \ldots, t_n]$ un polynôme symétrique de degré d. Alors, en regardant le polynôme :

$$\tilde{P}(t_1, \ldots, t_{n-1}) = P(t_1, \ldots, t_{n-1}, 0)$$

on déduit par hypothèse de récurrence « $\forall \tilde{d},\ \mathscr{P}(n-1, \tilde{d})$ », l'existence d'un polynôme $Q \in A[X_1, \ldots, X_{n-1}]$ de poids $\leq d$ qui vérifie :

$$P(t_1, \ldots, t_{n-1}, 0) = Q(s_1(t_1, \ldots, t_{n-1}), \ldots, s_{n-1}(t_1, \ldots, t_{n-1})).$$

Comme $s_i(t_1, \ldots, t_{n-1}, 0) = s_i(t_1, \ldots, t_{n-1})$ pour $i = 1, \ldots, n-1$ on peut même se permettre d'écrire :

$$P(t_1, \ldots, t_{n-1}, 0) = Q(s_1(t_1, \ldots, t_{n-1}, 0), \ldots, s_{n-1}(t_1, \ldots, t_{n-1}, 0)).$$

On pose :

$$P_1(t_1, \ldots, t_n) = P(t_1, \ldots, t_n) - Q(s_1(t_1, \ldots, t_n), \ldots, s_{n-1}(t_1, \ldots, t_n)).$$

On sait déjà que Q est de poids $\leq d$ donc $Q(s_1(t_1,\dots,t_n),\dots,s_{n-1}(t_1,\dots,t_n))$ est de degré $\leq d$. Ainsi, P_1 est un polynôme symétrique de degré $\leq d$.

Par construction, $P_1(t_1,\dots,t_{n-1},0)=0$ donc t_n divise P_1 et par symétrie, il en est de même pour tous les t_i. En fait, on peut même dire que $t_1\dots t_n = s_n(t_1,\dots,t_n)$ divise P_1. Autrement dit, il existe $P_2 \in A[t_1,\dots,t_n]$ tel que :

$$P(t_1,\dots,t_n) = P_2(t_1,\dots,t_n)s_n(t_1,\dots,t_n) + Q(s_1(t_1,\dots,t_n),\dots,s_{n-1}(t_1,\dots,t_n)).$$

On voit que P_2 est symétrique, de degré $\leq d-n < d$. Par hypothèse de récurrence $\mathscr{P}(n,\tilde{d})$ pour $\tilde{d} < d$, il existe $Q_2 \in A[X_1,\dots,X_n]$ de poids $\leq d-n$ tel que :

$$P_2(t_1,\dots,t_n) = Q_2(s_1(t_1,\dots,t_n),\dots,s_n(t_1,\dots,t_n)).$$

On peut conclure :

$$P(t_1,\dots,t_n) = Q_2(s_1,\dots,s_n)s_n + Q(s_1,\dots,s_{n-1})$$

et le polynôme

$$Q(X_1,\dots,X_n)X_n + Q(X_1,\dots,X_{n-1})$$

est bien de poids $\leq d$.

On a montré $\mathscr{P}(n,d)$ donc par récurrence $\ll \forall d,\ \ \mathscr{P}(n,d) \gg$. Puis par récurrence $\ll \forall n,\ \forall d,\ \ \mathscr{P}(n,d) \gg$. □

En ajoutant le concept d'ordre dans la preuve, on peut ensuite montrer que le polynôme Q est unique.

Référence.
A. Jeanneret, D. Lines, *Invitation à l'Algèbre*
E. Ramis, C. Deschamps, J. Odoux, *Cours de mathématiques spéciales, tome 1*

105 Groupe des permutations d'un ensemble fini. Applications.
142 Algèbre des polynômes à plusieurs indéterminées. Applications.
144 Racines d'un polynôme. Fonctions symétriques élémentaires. Exemples et applications

1.12 Les théorèmes de Chevalley-Warning et d'Erdös-Ginzburg-Ziv

K désigne un corps fini de cardinal $q = p^\alpha$.

Théorème (Chevalley-Warning)**.** *Soit $(f_a)_{a\in A}$ une famille de polynômes de $\mathbf{K}[X_1,\dots,X_m]$, indexée par un ensemble A. On suppose que les degrés de ces polynômes vérifient :*

$$\sum_{a\in A} \deg(f_a) < m.$$

On pose $V \subset K^m$ l'ensemble des points où tous les f_a s'annulent simultanément. Le cardinal de V vérifie :

$$\mathrm{Card}(V) \equiv 0 \ \ [p].$$

Preuve. On considère le polynôme :

$$P = \prod_{a \in A} (1 - f_a^{q-1})$$

et on va faire ça progressivement :

(1) Il est clair que $P(x) = 1_{\mathbf{K}}$ si $x \in V$ et s'il existe $a \in A$ tel que $f_a(x) \neq 0$ alors, puisque que $\mathbf{K}^\times$ est cyclique d'ordre $q-1$, on a $f_a^{q-1}(x) = 1_{\mathbf{K}}$ et donc $P(x) = 0_{\mathbf{K}}$. Finalement, en définissant $S(f) := \sum_{x \in K^m} f(x) \in \mathbf{K}$ pour tout polynôme $f \in \mathbf{K}[X_1, \ldots, X_m]$, on a :

$$S(P) = \sum_{x \in \mathbf{K}^m} P(x) = \mathrm{Card}(V) 1_{\mathbf{K}}.$$

(2) On va montrer que $S(P) = 0_{\mathbf{K}}$. D'abord, la condition $\sum_{a \in A} \deg(f_a) < m$ entraîne $\deg(P) < m(q-1)$. Il suffit donc de montrer $S(X^u) = 0_{\mathbf{K}}$ pour tout multi-indice $u = (u_1, \ldots, u_m)$ tel que $\sum_{i=1}^m u_i < m(q-1)$. Par le principe des tiroirs, il existe un indice i tel que $u_i < q-1$. On calcule :

$$S(X^u) = S(X_i^{u_i}) S(X_1^{u_1} \ldots \hat{X}_i^{u_i} \ldots X_m^{u_m}).$$

(3) Ah oui, mais si $y \in \mathbf{K}^\times$ est un générateur de $\mathbf{K}^\times$ et avec la convention $0^0 = 0$:

$$S(X_i^{u_i}) = \sum_{x \in \mathbf{K}} x^{u_i} = \sum_{x \in \mathbf{K}^\times} (yx)^{u_i} = y^{u_i} S(X_i^{u_i}).$$

Comme $u_i < q-1$, on est sûr que $y^{u_i} \neq 1_{\mathbf{K}}$ et donc $S(X_i^{u_i}) = 0 = S(X^u)$.

(4) En conclusion, on a montré :

$$\mathrm{Card}(V) 1_{\mathbf{K}} = 0_{\mathbf{K}}$$

et comme $\mathbf{K}$ est de caractéristique p, on a bien :

$$\mathrm{Card}(V) \equiv 0 \ [p].$$

□

Théorème (Erdös-Ginzburg-Ziv)**.** *Soit $n \in \mathbf{N}$. Parmi $2n-1$ entiers $a_1, \ldots, a_{2n-1}$, on peut toujours en trouver n dont la somme est divisible par n. En plus, c'est optimal.*

Preuve. Il faut commencer par :

Étape 1. Le cas où l'entier n est premier auquel cas il sera noté p.

On sa place dans le corps $K = \mathbf{F}_p$ et notera $\overline{a}$ la classe modulo p de $a \in \mathbf{N}$. On va applique le théorème de Chevalley-Warning avec les polynômes :

$$P_1(X_1, \ldots, X_{2p-1}) = \sum_{i=1}^{2p-1} X_i^{p-1} \quad \text{et} \quad P_2(X_1, \ldots, X_{2p-1}) = \sum_{i=1}^{2p-1} \overline{a}_i X_i^{p-1}.$$

Puisque $0 \in \mathbf{K}^{2p-1}$ est une racine commune à ces polynômes, on est assuré de l'existence d'une racine on triviale notée $(x_1, \ldots x_{2p-1})$. Il y a deux choses à voir :

(1) D'abord comme $x_i^{p-1} = 1_{\mathbf{K}}$ si $x_i \neq 0_{\mathbf{K}}$ et $x_i^{p-1} = 0_{\mathbf{K}}$ sinon, on en déduit (toujours car $\mathbf{K}$ est de caractéristique p) que la relation $P_1(x_1, \ldots, x_{2p-1}) = 0_{\mathbf{K}}$ implique qu'il existe très exactement p éléments $x_{n_1}, \ldots, x_{n_p}$ non nuls.

(2) C'est fini en considérant la deuxième relation $P_2(x_1, \ldots, x_{2p-1}) = 0_{\mathbf{K}}$ puisque :

$$0_{\mathbf{K}} = P_2(x_1, \ldots, x_{2p-1}) = \sum_{i=1}^{p} \overline{a}_{n_i} x_{n_i}^{p-1} = \sum_{i=1}^{p} \overline{a}_{n_i}.$$

Étape 2. Le cas général où l'entier p redevient $n \in \mathbf{N}$.

On va procéder par récurrence (forte) sur $n \in \mathbf{N}$. L'initialisation pour $n = 1$ n'est pas difficile. Supposons donc le résultat montré jusqu'au rang $n-1$, $n \geq 1$. Si n est premier, c'est l'étape 1, sinon on écrit $n = pn'$ avec p premier et $n' < n$.

(1) On écrit $2n - 1 = (2n' - 1)p + p - 1$ et par hypothèse de récurrence on peut construire $(2n' - 1)$ sous-ensembles disjoints de $E := \{a_1, \ldots, a_{2n-1}\}$ de la façon suivante : pour $i \in \{1, \ldots, 2n' - 1\}$, $E_i \subset E \setminus (E_1 \cup \ldots \cup E_{i-1})$ est de cardinal p et la somme de ses éléments est divisible par p. À la fin, $E \setminus (E_1 \cup \ldots \cup E_{2n'-1})$ est de cardinal $p - 1$ et on ne peut plus continuer.

(2) Pour $i \in \{1, \ldots, 2n' - 1\}$, on note s_i la somme des éléments de E_i et $s_i = ps_i'$. On applique encore l'hypothèse de récurrence avec les s_i' : il existe $k_1, \ldots, k_{n'}$ tel que n' divise $s_1' + \ldots + s_{k_{n'}}'$.

(3) Pour conclure, il suffit de considérer le sous-ensemble

$$\bigcup_{j=1}^{n'} E_{k_j} \subset \{a_1, \ldots, a_{2n-1}\}$$

qui est de cardinal $pn' = n$ et dont la somme de ses éléments vaut

$$\sum_{j=1}^{n'} s_{k_j} = p \sum_{j=1}^{n'} s_{k_j}'$$

qui est divisible par $pn' = n$.

Étape 3. Le résultat est optimal.

On considère $(2n - 2)$ entiers parmi lesquels $(n - 1)$ valent 0 et $(n - 1)$ valent 1. On ne peut pas trouver n éléments dont la somme soit divisible par n, puisqu'elle est toujours inférieure à n et strictement positive. □

Références. M. Zavidovique, *Un Max de Maths*

120 Anneaux $\mathbf{Z}/n\mathbf{Z}$. Applications.
121 Nombres premiers. Applications.
123 Corps finis. Applications.
142 Algèbre des polynômes à plusieurs indéterminées. Applications.
144 Racines d'un polynôme. Fonctions symétriques élémentaires. Exemples et applications

1.13 Le nombre de polynômes irréductibles unitaires dans $\mathbf{F}_q$

Théorème (Formule d'inversion de Möbius)**.** *Soient $f : \mathbf{N}^* \to \mathbf{R}$ et $g : n \in \mathbf{N}^* \mapsto \sum_{d|n} f(d)$. Alors pour tout $n \geq 1$:*

$$f(n) = \sum_{d|n} \mu\left(\frac{n}{d}\right) g(d) = \sum_{d|n} \mu(d) g\left(\frac{n}{d}\right).$$

Preuve. On commence par voir que si $n = 1$, $\sum_{d|n} \mu(d) = 1$ et si $n = p_1^{\alpha_1} \dots p_r^{\alpha_r}$, alors :

$$\begin{aligned} \sum_{d|n} \mu(d) &= \mu(1) + \sum_{1 \leq i \leq r} \mu(p_i) + \dots + \mu(p_1 \dots p + r) \\ &= 1 - \binom{r}{1} + \binom{r}{2} - \dots + (-1)^r = (1-1)^r = 0 \end{aligned}$$

de sorte que l'on peut calculer :

$$\begin{aligned} \sum_{d|n} \mu(d) g\left(\frac{n}{d}\right) &= \sum_{d|n} \mu(d) \sum_{d' | \frac{n}{d}} f(d') = \sum_{dd'|n} f(d') \\ &= \sum_{d'|n} f(d') \sum_{d | \frac{n}{d'}} \mu(d) = f(n) \end{aligned}$$

Et on passe d'une somme à l'autre en posant $d' = n/d$. □

Fort de ce résultat, on va montrer l'estimation suivante.

Théorème. *Pour $n \in \mathbf{N}^*$, $A(n,q)$ désigne le nombre de polynômes irréductibles unitaires sur $\mathbf{F}_q$. On note $I(n,q) = \operatorname{Card} A(n,q)$. Il existe des polynômes irréductibles de tout degré sur $\mathbf{F}_q$ et*

$$I(n,q) \sim \frac{q^n}{n}.$$

Preuve. La première partie consiste à étudier les diviseurs des polynômes $X^{q^n} - X$.

(1) Si P est un facteur irréductible unitaire de $X^{q^n} - X$ dans $\mathbf{F}_q$, alors notons d son degré. Comme $X^{q^n} - X$ est scindé sur $\mathbf{F}_{q^n}$, P est aussi scindé dans $\mathbf{F}_{q^n}$ et si $x \in \mathbf{F}_{q^n}$ est une racine de P, $\mathbf{F}_q(x)$ est un corps intermédiaire entre $\mathbf{F}_q$ et $\mathbf{F}_{q^n}$ de degré d. Alors, la théorème de la base télescopique donne :

$$[\mathbf{F}_{q^n} : \mathbf{F}_q(x)][\mathbf{F}_q(x) : \mathbf{F}_q] = [\mathbf{F}_{q^n} : \mathbf{F}_q] = n$$

et on peut déjà affirmer que $d|n$. De plus, comme $X^{q_n} - X$ est à racines simples dans $\mathbf{F}^{q^n}$ (c'est la définition), ses facteurs irréductibles dans $\mathbf{F}_q$ sont de multiplicité 1 et on a :

$$X^{q^n} - X \Big| \prod_{d|n} \prod_{P \in A(d,q)} P.$$

(2) Maintenant, si $d|n$ et $P \in A(d,q)$, alors soit $\overline{\mathbf{F}}_q$ une clôture algébrique de $\mathbf{F}_q$.

- Si $x \in \overline{F}_q$ est une racine de P alors par unicité des corps finis, $\mathbf{F}_q(x) \simeq \mathbf{F}_{q^d}$ donc x est aussi une racine de $X^{q^d} - X$.

- Comme $d|n$ un savant bidouillage avec les puissances montre que $X^{q^d}-X\Big|X^{q^n}-X$.
- Comme P Est irréductible, il est à racines simples dans $\overline{F}_q$, en effet sinon il existerait $\alpha\in\overline{F}_q$ tel que $X-\alpha|P$ et $X-\alpha|P'$. Alors $X-\alpha$ divise $P\wedge P'$ qui est donc de degré supérieur ou égal à 1. Mais comme $P\wedge P'|P$ et que ce dernier est irréductible, $P\wedge P'=P$, c'est à dire $P'=0$. Le morphisme de Frobenius dit alors que P est de la forme $P(X)=R(X^p)=R(X)^p$, c'est absurde.

En conclusion, on vient de montrer en mettant tout ensemble que :

$$\prod_{d|n}\prod_{P\in A(d,q)}P\Big|X^{q^n}-X.$$

Finalement on peut écrire :

$$X^{q^n}-X=\prod_{d|n}\prod_{P\in A(d,q)}P$$

et la suite n'est que dénombrement : en regardant les degrés, on trouve

$$q^n=\sum_{d|n}dI(d,q).$$

La formule d'inversion de Möbius donne :

$$nI(n,q)=\sum_{d|n}\mu\left(\frac{n}{d}\right)q^d\implies I(n,q)=\frac{q^n+r_n}{n}$$

où

$$r_n=\sum_{d|n,\,d<n}\mu\left(\frac{n}{d}\right)q^d.$$

Cela montre déjà que $I(q,n)\neq 0$ pour tout n et en majorant :

$$|r_n|\leq\sum_{d=1}^{\lfloor\frac{n}{2}\rfloor}q^d=q\frac{q^{\lfloor\frac{n}{2}\rfloor}-1}{q-1}\leq\frac{q^{\lfloor\frac{n}{2}\rfloor+1}}{q-1}=o(q^n)$$

et on en déduit l'équivalent annoncé. □

Référence. S. Francinou, H. Gianella, *Exercices de Mathématiques pour l'Agrégation, Algèbre 1*

123 Corps finis. Applications.
125 Extensions de corps. Exemples et applications.
141 Polynômes irréductibles à une indéterminée. Corps de rupture. Exemples et applications.
190 Méthodes combinatoires, problèmes de dénombrement.

1.14 Le groupe $\mathcal{O}(p,q)$

Soient deux entiers p, q dont la somme vaut n. $\mathcal{O}(p,q) \subset GL_n(\mathbf{R})$ désigne le groupe des isométries de la forme quadratique standard sur $\mathbf{R}^n$ de signature (p,q) :

$$\mathcal{O}(p,q) = \{M \in GL_{p+q}(\mathbf{R}),\ \ MI_{p,q}M^* = I_{p,q}\}$$

où $I_{p,q}$ désigne la matrice :

$$I_{p,q} := \begin{pmatrix} I_p & 0 \\ 0 & -I_q \end{pmatrix}.$$

Théorème. *Il existe un homéomorphisme :*

$$\mathcal{O}(p,q) \cong \mathcal{O}(p) \times \mathcal{O}(q) \times \mathbf{R}^{pq}.$$

Preuve. La preuve peut être vue comme une application de la décomposition polaire sur $GL_n(\mathbf{R})$.

Étape 1. $\mathcal{O}(p,q)$ est stable par transposition et la décomposition polaire y est interne.

Soit $M \in \mathcal{O}(p,q)$. Puisqu'une matrice commute avec son inverse, on a :

$${}^tMI_{p,q}M = I_{p,q} \Longleftrightarrow ({}^tMI_{p,q})(MI_{p,q}) = I_n \Longleftrightarrow (MI_{p,q})({}^tMI_{p,q}) = I_n \Longleftrightarrow MI_{p,q}{}^tM = I_{p+q}$$

donc ${}^tM \in \mathcal{O}(p,q)$. Écrivons maintenant $M = OS$ où $O \in \mathcal{O}(n)$ et $S \in \mathscr{S}_n^{++}(\mathbf{R})$. Il s'agit de montrer que O et S sont dans $\mathcal{O}(p,q)$. On va montrer que S l'est, ce qui est suffisant. Par le théorème spectral et l'unicité de la racine carrée, on peut écrire :

$$S = P\,\mathrm{diag}(\sqrt{\lambda_1}, \ldots, \sqrt{\lambda_n})P^{-1}$$

où $P \in GL_n(\mathbf{R})$ et les $\lambda_i > 0$. En outre, comme $S^2 = {}^tMM \in \mathcal{O}(p,q)$, il est facile de voir que :

$$S^2 I_{p,q} = I_{p,q}S^{-2} \Longrightarrow L(S^2)I_{p,q} = I_{p,q}L(S^{-2})$$

pour tout polynôme $L \in \mathbf{R}[X]$. En particulier, si L est le polynôme interpolateur de Lagrange vérifiant :

$$\forall i \in \{1, \ldots, n\}, \quad L(\lambda_i) = \sqrt{\lambda_i} \ \text{ et } \ L(\lambda_i^{-1}) = \left(\sqrt{\lambda_i}\right)^{-1}$$

on a $S = L(S^2)$ et $S^{-1} = L(S^{-2})$ et la conclusion suit.

Conclusion provisoire. Comme la décomposition polaire induit un homéomorphisme de $GL_n(\mathbf{R})$ dans $\mathcal{O}(n) \times \mathscr{S}_n^{++}(\mathbf{R})$, on en déduit par restriction que

$$\mathcal{O}(p,q) \cong \big(\mathcal{O}(p,q) \cap \mathcal{O}(n)\big) \times \big(\mathcal{O}(p,q) \cap \mathscr{S}_n^{++}(\mathbf{R})\big). \tag{1.5}$$

Étape 2. Étude de l'intersection $\mathcal{O}(p,q) \cap \mathcal{O}(n)$.

Soit $M \in \mathcal{O}(p,q) \cap \mathcal{O}(n)$. Puisque $I_{p,q}M = MI_{p,q}$, les sous-espaces propres de $I_{p,q}$ sont stables par M qui s'écrit alors par blocs sous la forme :

$$M = \begin{pmatrix} M_p & 0 \\ 0 & M_q \end{pmatrix}.$$

Et comme ${}^tMM = I_n$, on trouve que $M_p \in \mathcal{O}(p)$ et $M_q \in \mathcal{O}(q)$. D'où l'homéomorphisme :

$$\mathcal{O}(p,q) \cap \mathcal{O}(n) \cong \mathcal{O}(p) \times \mathcal{O}(q). \tag{1.6}$$

Étape 3. Étude de l'intersection $\mathcal{O}(p,q) \cap \mathscr{S}_n^{++}(\mathbf{R})$.

Introduisons l'ensemble

$$\mathcal{U}(p,q) := \big\{N \in \mathcal{M}_n(\mathbf{R}),\ \ NI_{p,q} + I_{p,q}N = 0\big\}.$$

On va montrer que exp réalise un homéomorphisme [5] :

$$\exp : \mathcal{U}(p,q) \cap \mathscr{S}_n(\mathbf{R}) \xrightarrow{\sim} \mathcal{O}(p,q) \cap \mathscr{S}_n^{++}(\mathbf{R}). \tag{1.7}$$

- D'abord $\exp : \mathscr{S}_n(\mathbf{R}) \to \mathscr{S}_n^{++}(\mathbf{R})$ est continue et injective car si $\exp(A) = \exp(A')$, en écrivant :

  $$A = P\,\mathrm{diag}(\lambda_1, \ldots, \lambda_n)P^{-1} \Longrightarrow \exp(A) = P\,\mathrm{diag}(e^{\lambda_1}, \ldots, e^{\lambda_n})P^{-1} \in \mathscr{S}_n^{++}(\mathbf{R})$$

  on obtient en considérant le polynôme interpolateur de Lagrange L tel que $L(e^{\lambda_i}) = \lambda_i$ pour tout $i \in \{1, \ldots, n\}$:

  $$L(\exp(A')) = L(\exp(A)) = A$$

  et comme A' commute avec $L(\exp(A'))$, on vient de prouver que A et A' commutent. Ces deux matrices sont donc co-diagonalisables : il existe P_0 inversible telle que :

  $$A = P_0\,\mathrm{diag}(\lambda_1, \ldots, \lambda_n)P_0^{-1} \ \text{ et } \ A' = P_0\,\mathrm{diag}(\lambda'_1, \ldots, \lambda'_n)P_0^{-1}.$$

  En passant à l'exponentielle, on trouve que pour tout $i \in \{1, \ldots, n\}$, $e^{\lambda_i} = e^{\lambda'_i}$, d'où $\lambda_i = \lambda'_i$ et $A = A'$.
- Si $N \in \mathcal{U}(p,q) \cap \mathscr{S}_n(\mathbf{R})$, alors $\exp(N) \in \mathcal{O}(p,q)$ car

  $$I_{p,q}N = -NI_{p,q} \Longrightarrow I_{p,q}\exp(N) = \exp(-N)I_{p,q} \Longrightarrow {}^t\exp(N)I_{p,q}\exp(N) = I_{p,q}.$$

- Enfin, si $M \in \mathcal{O}(p,q) \cap \mathscr{S}_n^{++}(\mathbf{R})$ on peut écrire pour une certaine matrice $P \in \mathcal{O}(n)$ et certains $\lambda_i > 0$:

  $$M = P\,\mathrm{diag}(\lambda_1, \ldots, \lambda_n)P^{-1}.$$

  Alors la matrice

  $$N = P\,\mathrm{diag}(\ln\lambda_1, \ldots, \ln\lambda_n)P^{-1}$$

  est bien définie, appartient à $\mathscr{S}_n(\mathbf{R})$, dépend continûment de M et vérifie $\exp(N) = M$. Il reste à voir que $N \in \mathcal{U}(p,q)$. Une fois encore, comme :

  $$MI_{p,q} = I_{p,q}M^{-1} \Longrightarrow L(M)I_{p,q} = I_{p,q}L(M^{-1})$$

  pour tout polynôme $L \in \mathbf{R}[X]$, il suffit de trouver L tel que $N = L(M)$ et $-N = L(M^{-1})$. Le polynôme d'interpolation de Lagrange vérifiant :

  $$\forall i \in \{1, \ldots, n\},\ \ L(\lambda_i) = \ln\lambda_i \ \text{ et } \ L(\lambda_i^{-1}) = -\ln\lambda_i$$

  convient.

5. On montre en fait incidemment que $\exp : \mathscr{S}_n(\mathbf{R}) \to \mathscr{S}_n^{++}(\mathbf{R})$ est un homéomorphisme.

Finalement, un calcul par blocs montre que

$$U = \begin{pmatrix} A & B \\ {}^tB & C \end{pmatrix} \in \mathcal{U}(p,q) \cap \mathscr{S}_n(\mathbf{R}) \Longrightarrow UI_{p,q} + I_{p,q}U = 2\begin{pmatrix} A & 0 \\ 0 & -D \end{pmatrix}.$$

Et donc

$$\mathcal{U}(p,q) \cap \mathscr{S}_n(\mathbf{R}) = \left\{ \begin{pmatrix} 0 & B \\ {}^tB & 0 \end{pmatrix},\ B \in \mathcal{M}_{p,q}(\mathbf{R}) \right\}$$

D'où l'homéomorphisme :

$$\mathcal{U}(p,q) \cap \mathscr{S}_n(R) \cong \mathbf{R}^{pq} \tag{1.8}$$

Conclusion. Le théorème découle de (1.5), (1.6), (1.7) et (1.8). □

Références. Zavidovique et H2G2

170 Formes quadratiques sur un espace vectoriel de dimension finie. Orthogonalité, isotropie. Applications.

1.15 Les polygônes réguliers constructibles

Quelques pré-requis

Théorème (Wanzel)**.** *Un nombre est constructible si et seulement s'il appartient à une tout d'extensions quadratiques de* $\mathbf{Q}$.

Le polygone régulier à n côtés est constructible lorsque l'angle $2\pi/n$ est constructible, ce qui revient à dire $\cos(2\pi/n)$ est constructible.

Les nombres de Fermat sont les nombres de la forme $2^{(2^m)} + 1$, $m \in \mathbf{N}$.

Ce qu'on va montrer.

Théorème (Gauß-Wanzel)**.** *Les polygones réguliers constructibles sont ceux dont le nombre de côtés n sont de la forme 2^α avec $\alpha \geq 2$ ou de la forme $2^\alpha p_1 p_2 \dots p_r$ où les p_i sont des nombres premiers de Fermat distincts.*

La preuve de ce théorème résulte de la concaténation des quatre résultats suivants, classés par ordre de difficulté.

Proposition 1. *Les angles de la forme $\widehat{\frac{2\pi}{2^\alpha}}$ sont constructibles.*

Preuve. Il suffit de savoir construire des bissectrices. □

Proposition 2. *Si m et n sont premiers entre eux, l'angle $\widehat{\frac{2\pi}{mn}}$ est constructible si et seulement si $\widehat{\frac{2\pi}{n}}$ et $\widehat{\frac{2\pi}{m}}$ sont constructibles. En conséquence, si n a pour décomposition en facteurs premiers $n = p_1^{\alpha_1} \dots p_k^{\alpha_k}$, le polygône régulier à n côtés est constructible si et seulement si les angles $\widehat{\frac{2\pi}{p_1^{\alpha_1}}}, \dots, \widehat{\frac{2\pi}{p_k^{\alpha_k}}}$ le sont.*

Preuve. Prouvons la première partie de l'énoncé. Le sens direct est facile et ne nécessite pas que m et n soient premiers entre eux : les angles $\widehat{\frac{2\pi}{n}}$ et $\widehat{\frac{2\pi}{m}}$ sont des multiples entiers de l'angle $\widehat{\frac{2\pi}{mn}}$. Pour la réciproque, on écrit une relation de Bézout entre m et n et il s'agit ensuite de construire la somme de deux angles constructibles, ce qui est possible. □

Proposition 3. *Soit $p \geq 3$ un nombre premier. Si $\widehat{\frac{2\pi}{p^\alpha}}$ est constructible alors $\alpha = 1$ et p est un nombre de Fermat.*

Preuve. Notons $q = p^\alpha$ et $\omega = \exp(2i\pi/q)$. Par le théorème de Wanzel, on a :

$$[\mathbf{Q}(\omega) : \mathbf{Q}] = 2^m, \quad \text{pour un certain } m \in \mathbf{N}.$$

Mais comme le q-ème polynôme cyclotomique Φ_q est le polynôme annulateur de ω sur $\mathbf{Q}$, on a :

$$[\mathbf{Q}(\omega) : \mathbf{Q}] = 2^m = \varphi(q) = p^{\alpha-1}(p-1).$$

Comme p est impair, on a déjà $\alpha = 1$ et $p = 2^m + 1$. Reste à montrer que m est une puissance de 2. En écrivant $m = 2^\beta\lambda$ avec λ impair, on a : $p = 1 + (2^{(2^\beta)})^\lambda$ et comme $1 + X | 1 + X^\lambda$ (car (-1) est racine lorsque λ est impair), on a $1 + 2^{(2^\beta)} | p$ et le résultat suit puisque p est premier. □

Proposition 4. *La réciproque de la proposition 3 est vraie.*

Preuve. Notons $p = 1 + 2^n$, ω une racine primitive p-ème de l'unité et $K = \mathbf{Q}(\omega)$. On a $[K : \mathbf{Q}] = p - 1$ et une base de K sur $\mathbf{Q}$ est $\{1, \omega, \ldots, \omega^{p-2}\}$.

Étape 1. Le groupe des automorphismes de K, sa vie, son oeuvre.

On note G le groupe des automorphismes de $K/\mathbf{Q}$. En appliquant $g \in G$ à $\Phi_p(\omega) = 0$, on voit que g est entièrement déterminé par l'image de ω qui ne peut être qu'un ω^k, $1 \leq k \leq p-1$. On note désormais $g_k \in G$ tel que

$$g_k(\omega) = \omega^k.$$

On a donc $G = \{g_1, \ldots, g_{p-1}\}$ mais on peut en dire plus : en considérant l'application :

$$\psi : G \to (\mathbf{Z}/p\mathbf{Z})^\times, \;\; g_k \mapsto \bar{k}$$

dont on peut montrer que c'est un isomorphisme de groupes, on sait aussi que G est un groupe cyclique d'ordre $p-1 = 2^n$, disons engendré par $g \in G$. Intéressons-nous maintenant aux sous-groupes : $G_i := \langle g^{2^i} \rangle$ pour $i \in \{1, \ldots, n\}$. Ils sont d'ordre 2^{n-i} et on a :

$$\{1\} = G_n \subset G_{n-1} \subset \ldots \subset G_1 \subset G_0 = G.$$

Étape 2. De la théorie de Galois sans le dire.

L'idée principale de la preuve consiste à associer aux sous-groupes G_i les sous corps :

$$K_i = \{z \in K, \; g^{2^i}(z) = z\} \subset K.$$

De ces sous-corps, proviendra la tour d'extensions quadratiques recherchée. Comme $g^{2^{i+1}} = (g^{2^i})^2$, on a déjà $K_i \subseteq K_{i+1}$ et comme $g^n = Id$, on a $K_n = K$. Reste à montrer :

$$(i)\ K_0 = \mathbf{Q} \qquad (ii)\ \forall i \in \{0, \ldots, n-1\},\ [K_{i+1} : K_i] = 2.$$

Étape 3. Faisons le.

(i) On a clairement $\mathbf{Q} \subset K_0$. De plus, si $z \in K$, z s'écrit de façon unique :

$$z = \lambda_0\omega + \lambda_1 g(\omega) + \ldots + \lambda_{p-2}g^{p-2}(\omega).$$

Ainsi, si $z \in K_0$, on trouve en appliquant g à cette égalité : $\lambda_0 = \lambda_1 = \ldots = \lambda_{p-2}$ et alors :

$$z = \lambda_0(\omega + \omega^2 + \ldots + \omega^{p-1}) = -\lambda_0 \in \mathbf{Q}.$$

(ii) On montre d'abord que les inclusions $K_i \subset K_{i+1}$ sont strictes. Il suffit pour cela de voir que :

$$z = \sum_{h=0}^{2^{n-i-1}-1} g^{h2^{i+1}} \in K_{i+1} \setminus K_i.$$

Ensuite, il n'y a qu'à écrire :

$$2^n = [K_n : K_0] = [K_n : K_{n-1}][K_{n-1} : K_{n-2}] \ldots [K_1 : K_0].$$

Il y a n facteurs non égaux à 1 (car les inclusions sont strictes) d'où le résultat.

Étape 4. La conclusion.

Pour revenir à $\cos(2\pi/p)$, il suffit d'écrire :

$$\cos\frac{2\pi}{p} = \frac{1}{2}(\omega + \omega^{-1}) \in \mathbf{Q}(\omega).$$

En fait, on a même :

$$K_{n-1} = \mathbf{Q}\left(\cos\frac{2\pi}{p}\right)$$

mais c'est inutile. □

Référence. J-C. Carrega, *Théorie des corps, la règle et le compas*

102 Groupe des nombres complexes de module 1. Sous-groupe des racines de l'unité. Applications.
121 Nombres premiers. Applications.
125 Extensions de corps. Exemples et applications.
183 Utilisation des groupes en géométrie.

1.16 Quaternions et rotations

Quelques pré-requis

Définition (Réalisation matricielle des quaternions)**.** L'ensemble des quaternions $\mathbf{H}$ est l'ensemble des matrices de $\mathcal{M}_2(\mathbf{C})$ de la forme :

$$q = \begin{pmatrix} z & -\overline{w} \\ w & \overline{z} \end{pmatrix}, \quad z, w \in \mathbf{C}$$

C'est une algèbre à division non commutative où l'inverse d'un élément q est donné par :

$$q^{-1} = \frac{1}{\det(q)} \begin{pmatrix} \overline{z} & \overline{w} \\ -w & z \end{pmatrix}.$$

- Le corps $\mathbf{C}$ est le sous-corps de $\mathbf{H}$ constitué des matrices $\mathrm{diag}(z, \overline{z})$.
- La conjugaison quaternionique d'un élément q, notée $\overline{q}$, est définie comme la trans-conjuguée de la matrice qui le représente.
- La norme d'un élément $q \in \mathbf{H}$ est $N(q) = q\overline{q}$.

Proposition. *En tant qu'espace vectoriel,* $\mathbf{H}$ *est de dimension 4 et admet pour base :*

$$1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad j = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}, \quad j = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \quad k = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}.$$

On a $i^2 = j^2 = k^2 = -1$ *et la multiplication entre les éléments de* $\mathbf{H}$ *est donnée par le diagramme suivant :*

$$\begin{array}{ccccc} & & i & & \\ & \swarrow & & \nwarrow & \\ j & & \longrightarrow & & k \end{array}$$

- *Le sous-espace des quaternions imaginaires purs est* $\mathbf{I} = \mathbf{R}i \oplus \mathbf{R}j \oplus \mathbf{R}k$.
- *Le centre de* $\mathbf{H}$ *est réduit aux quaternions réels* $\mathbf{R}$.
- *En écrivant* $q = x + yi + zj + tk$*, on a :*

  $$\overline{q} = x - yi - zj - tk \quad \text{et} \quad N(q) = x^2 + y^2 + z^2 + t^2.$$

  et $N(qq') = N(q)N(q')$.
- *L'application* N *est une forme quadratique sur* $\mathbf{H}$ *associée à la forme bilinéaire* $(q, q') \mapsto \frac{1}{2}(q\overline{q}' + q'\overline{q})$.
- *La base* $(1, i, j, k)$ *est orthonormée. En particulier, le sous-espace des quaternions imaginaires purs* $\mathbf{I}$ *est l'orthogonal de* $\mathbf{R} = \mathbf{R}1$.

Ce qu'on va monter.

On note G le groupe des quaternions de norme 1, c'est à dire le noyau de l'homomorphisme de groupes $N : \mathbf{H}^* \to \mathbf{R}_+^*$.

Théorème. *Il existe un isomorphisme explicite :*

$$G/\{\pm 1\} \simeq SO_3(\mathbf{R}).$$

Preuve. Il s'agit de remarquer que G agit par automorphsimes intérieurs sur $\mathbf{H}$:

$$\begin{array}{rccccccc} S: & G & \longrightarrow & \mathrm{Aut}(\mathbf{H}) & & & & \\ & h & \longmapsto & S_h: & \mathbf{H} & \longrightarrow & \mathbf{H} \\ & & & & q & \longmapsto & hqh^{-1} = hq\overline{h} \end{array}$$

L'application linéaire S_h est bien un automorphisme, son inverse est donné par $S_{\overline{h}} = (S_h)^{-1}$. L'application S est bien un homomorphisme car $S_{h_1h_2}(q) = h_1h_2q\overline{h}_2h_1 = S_{h_1}S_{h_2}(q)$.

(1) Pour tout $h \in G$, l'application S_h respecte la norme. Comme 1 est central dans $\mathbf{H}$, l'application S_h préserve $\mathbf{R}$ et son orthogonal $\mathbf{I}$. On a donc le droit de restreindre $S : G \to \mathcal{O}(\mathbf{I})$. Via le choix d'une base qui donne un isomorphisme entre $\mathcal{O}(\mathbf{I})$ et le groupe orthogonal $\mathcal{O}(3, \mathbf{R})$, on en déduit un morphisme :

$$S : G \to \mathcal{O}(3, \mathbf{R}).$$

(2) En munissant $\mathcal{O}(3, \mathbf{R}) \subset \mathcal{M}_3(\mathbf{R})$ de sa topologie usuelle, on voit que l'application S est continue (on peut le voir en écrivant la matrice de S_h dans la base (i, j, k) dont les coefficients sont polynômiaux en les coordonnées de h). On peut écrire :

$$G = \left\{(x, y, z, t) \in \mathbf{R}^4,\ \ x^2 + y^2 + z^2 + t^2 = 1\right\}$$

de sorte que G est homéomorphe à la sphère $\mathbf{S}^3$ et en particulier connexe. Ainsi, l'image $S(G)$ est aussi connexe et puisqu'elle contient l'identité, on dispose en fait d'un morphisme :

$$S : G \to SO_3(\mathbf{R}).$$

(3) Le noyau de S est $\mathrm{Ker}\, S = Z(\mathbf{H}) \cap G = \mathbf{R} \cap G = \{\pm 1\}$. Il ne reste plus qu'à montrer la surjectivité de S et on pourra conclure par théorème d'isomorphisme.

(4) Pour la surjectivité, il suffit de montrer que l'image $S(G)$ contient tous les retournements (car ils engendre $SO_3(\mathbf{R})$). Soit $h \in \mathbf{I} \cap \mathbf{S}^3$ et r_h le retournement d'axe $\mathbf{R}h$. On va montrer que $S_h = r_h$. Il suffit de voir que S_h laisse stable h (ce qui est évident car $S_h(h) = hhh^{-1} = h$) et de montrer que S_h est une involution :

$$(S_h)^2 = S_{h^2} = S_{-1} = Id$$

car $h \in \mathbf{I} \cap G$ donc $\overline{h} = -h$ et $h^2 = -h\overline{h} = -1$.

□

Références.
H2G2
D. Perrin, *Cours d'algèbre*
A. Jeanneret, D. Lines, *Invitation à l'algèbre*

101 Groupe opérant sur un ensemble. Exemples et applications.
161 Isométries d'un espace affine euclidien de dimension finie. Applications en dimensions 2 et 3.
182 Applications des nombres complexes à la géométrie.
183 Utilisation des groupes en géométrie.

1.17 Sous-groupes distingués et noyaux de caractères

Lemme (Noyau des caractères)**.** *Soit G un groupe fini et $\rho : G \to GL(V)$ une représentation de caractère χ_V où V est un espace de dimension d. Alors*

$$K_{\chi_V} = \{g \in G,\ \chi_V(g) = \chi_V(1)\}$$

est un sous-groupe distingué de G appelé noyau de la représentation.

Preuve. Soit $g \in G$ d'ordre k. Comme $g^k = 1$, on a $\rho(g)^k = Id$ donc $\rho(g)$ est annulé par un polynôme scindé à racines simples. C'est donc un endomorphisme diagonalisable et ses valeurs propres $\omega_1, \ldots, \omega_d$ sont des racines de l'unité. En particulier :

$$|\chi_V(g)| = |\omega_1 + \ldots + \omega_d| \leq d = |\chi_V(1)|.$$

Si $g \in K_{\chi_V}$, alors l'inégalité précédente est une égalité et les ω_i sont positivement liés. Nécessairement,

$$\omega_1 = \ldots = \omega_d = 1$$

et $\rho(g) = Id$. Réciproquement c'est aussi vrai donc

$$K_{\chi_V} = \mathrm{Ker}(\rho)$$

est un sous-groupe distingué. □

Soit G un groupe fini dont on note $\widehat{G} = \{\rho_1, \ldots, \rho_r\}$ le dual formé de représentants des représentations irréductibles non isomorphes.

Théorème. *Les sous-groupes distingués de G sont exactement du type*

$$\bigcap_{i \in I} K_{\chi_i} \quad où \quad I \subset \{1, \ldots, r\}.$$

Preuve. Soit $N \lhd G$ un sous-groupe distingué. On appelle U la représentation régulière de G/N, on sait qu'elle est fidèle, c'est à dire que ρ_U est injective.
(1) On note $\pi : G \to G/N$ la projection canonique et on pose :

$$\forall g \in G,\ \tilde{\rho}_U(g) := \rho_U \circ \pi(g).$$

Alors $\tilde{\rho}_U : G \to GL(U)$ est une représentation de G de noyau

$$\mathrm{Ker}(\tilde{\rho}) = \mathrm{Ker}(\rho \circ \pi) = N.$$

(2) Notons χ le caractère de $\tilde{\rho}_U$. D'après le lemme préliminaire, $N = K_\chi$.

(3) On décompose χ en fonction des caractères irréductibles :

$$\chi = a_1\chi_1 + \ldots + a_r\chi_r.$$

Comme dans la preuve du lemme, on a :

$$\forall g \in G,\ |\chi(g)| \leq \sum_{i=1}^{r} a_i|\chi(g)| \leq \sum_{i=1}^{r} a_i|\chi(1)| = \chi(1)$$

et $g \in K_\chi$ si et seulement si $\chi(g) = \chi(1)$ si et seulement si l'inégalité est une égalité si et et seulement si :

$$\forall i \in \{1, \ldots, r\},\ a_i\chi_i(g) = a_i\chi_i(1).$$

On obtient finalement

$$g \in K_\chi \iff \Big[a_i > 0 \ \Rightarrow\ g \in K_{\chi_i}\Big].$$

et le résultat avec $I = \{i\ a_i > 0\}$.

Le réciproque est évidente puisqu'une intersection de sous-groupes distingués est un sous-groupe distingué. □

Corollaire. *G est simple si et seulement si*

$$\forall i \neq 1,\ \forall g \in G,\ \chi_i(g) \neq \chi_i(1).$$

Preuve. S'il existe $g \neq 1$ dans G tel que $\chi_1(g) = \chi_i(1)$ alors $K_i \subset G$ est un sous-groupe distingué non trivial donc G n'est pas simple. Réciproquement, si G n'est pas simple, il existe $g \in N$, $g \neq 1$ où $N \lhd G$ est un sous-groupe distingué. Le théorème précédent dit que :

$$N = \bigcap_{i \in I} K_i$$

donc $g \in K_i$ pour un certain i ce qui signifie encore que $\chi_i(g) = \chi_i(1)$. □

Référence. G. Peyré, *L'Algèbre discrète de la Transformée de Fourier*

103 Exemples de sous-groupes distingués et de groupes quotients. Applications.
107 Représentations et caractères d'un groupe fini sur un C-espace vectoriel. Exemples.

1.18 Sur les groupes paveurs du plan

Soit $\mathcal{E}$ un plan affine euclidien de direction E.

Définition (Groupe paveur)**.** On dit qu'un sous-groupe G de $Is^+(\mathcal{E})$ est un groupe paveur (direct) lorsqu'il existe un compact connexe P d'intérieur non vide tel que G vérifie les axiomes suivants :

(i) $\bigcup_{g\in G} g(P) = \mathcal{E}$.

(ii) $g(\mathring{P}) \cap h(\mathring{P}) \Rightarrow g = h$.

On notera $T \subset Is^+(\mathcal{E})$ l'ensemble des translations. Soit G un groupe paveur.

Lemme. *Soit Γ un sous-groupe de E. On suppose que :*

(i) Γ est discret au sens où pour tout $u \in \Gamma$, il existe $r > 0$ tel que $B(u,r) \cap \Gamma = \{u\}$.

(ii) Γ engendre E en tant qu'espace vectoriel.

Alors Γ est un réseau au sens où il existe une famille $\mathbf{R}$-libre $(u,v) \in E \times E$ (appelée base de Γ) telle que

$$\Gamma = \mathbf{Z}u \oplus \mathbf{Z}v.$$

De plus, on peut prendre pour base de Γ tout couple $(u,v) \in \Gamma$ vérifiant :

$$\|u\| = \min_{x\in\Gamma\setminus\{0\}} \|x\|, \quad et \quad \|v\| = \min_{x\in\Gamma\setminus\mathbf{Z}u} \|x\|.$$

Preuve. Puisque Γ est discret, on peut choisir $u \in \Gamma \setminus \{0\}$ tel que $\|u\|$ soit minimal, puis $v \in \Gamma \setminus \mathbf{R}u$ tel que $\|v\|$ soit minimal. On va montrer que $\Gamma \subset \mathbf{Z}u \oplus \mathbf{Z}v$. Soit $w = \lambda u + \mu v \in \Gamma \setminus \{0\}$ avec $\lambda, \mu \in [0,1[$ (quitte à translater). Si λ ou μ est nul, cela contredit la minimalité de u ou v. Si λ et μ sont non nuls, alors un calcul montre que :

$$\|w\|^2 < (a+b)\|u\|^2 \quad \text{donc} \quad a+b > 1.$$

Mais en refaisant pareil avec $w' = u + v - w$ on trouve $a + b < 1$. C'est contradictoire. Finalement $w = 0$ et la conclusion suit. □

Théorème. *Il y a à conjugaison près dans $GL(\mathbf{R}^2)$ exactement cinq groupes paveurs.*

Preuve. Il y a trois étapes.

Étape 1. Étude des translations de G.

On note $T(G)$ l'ensemble des translations de G et $\Gamma(G) \subset E$ le sous-espaces des $\vec{u} \in E$ tels que $t_{\vec{u}} \in T(G)$. **Il s'agit de montrer que $\Gamma(G)$ est un réseau de E.**

- D'abord, $\Gamma(G)$ est discret car si on choisit $\varepsilon > 0$ tel que P contienne une boule de rayon ε alors, puisque pour tout $g \in G \setminus \{id\}$, $g(\mathring{P}) \cap \mathring{P} = \emptyset$, on a :

  $$\vec{u} \in \Gamma(G) \Rightarrow \|\vec{u}\| \geq 2\varepsilon.$$

  et la structure de groupe de $\Gamma(G)$ permet de conclure.
- Supposons que $\Gamma(G) = \{0\}$ alors G ne contient que des rotations. Si deux rotations avaient des centres distincts, leur commutateur serait une translation non triviale donc toutes les rotations ont même centre et comme P est compact, $G(P)$ aussi et le premier axiome n'est pas vérifié. Supposons ensuite que $\Gamma(G) \subset \mathbf{R}\vec{u}$. Alors si $r \in G \setminus T(G)$, on a $r \circ t_{\vec{u}} \circ r^{-1} = t_{\vec{r}(\vec{u})} \in T(G)$ donc $\vec{r}(\vec{u})$ est colinéaire à $\vec{u}$ et r est une symétrie centrale. La composée de deux symétries centrales autour de deux centres distincts A et B est la translation de vecteur $2\overrightarrow{AB}$. Finalement, les centres des symétries sont sur une même droite dirigée par $\vec{u}$ et le premier axiome ne peut pas être vérifié.

Finalement, $\Gamma(G)$ est discret et contient une base de E donc c'est un réseau par le lemme préliminaire.

Étape 2. Étude des rotations de G.

On note $\vec{G} = \{\vec{f},\ f \in G\} \subset GL(E)$. **Il s'agit de montrer que $\vec{G}$ est isomorphe à un groupe cyclique d'ordre 1, 2, 3, 4 ou 6**.

Soit $g \in G$. On considère une base $(\vec{u}, \vec{v})$ du réseau $\Gamma(G)$. Comme

$$g \circ t_{\vec{u}} \circ g^{-1} = t_{\vec{g}(\vec{u})} \in G$$

alors $\vec{g}(\vec{u}) \in \mathbf{Z}\vec{u} + \mathbf{Z}\vec{v}$ et de même pour $\vec{v}$. Ainsi, la matrice de $\vec{g}$ dans la base $(\vec{u}, \vec{v})$ est à coefficients entiers. En notant θ l'angle de la rotation $\vec{g}$, on trouve : $\operatorname{Tr} \vec{g} = 2\cos\theta \in \mathbf{Z}$, de sorte que $\vec{g} \in \{id, -id, R_{\pi/2}, R_{\pi/3}, R_{2\pi/3}\}$. Puisque la composée de deux telles rotations distinctes et distinctes de $\pm id$ n'est pas une rotation listée, on est assuré que $\vec{G}$ est cyclique avec un ordre parmi ceux annoncés.

Étape 3. Conclusion.

D'abord, il existe des groupes paveurs dont l'ordre de la partie linéaire est un des cinq listés : il suffit d'en exhiber. Ensuite :

- Soient $n \in \{1, 2\}$ et G_1 et G_2 deux groupes paveurs de partie linéaire d'ordre n. Ainsi, pour $i \in \{1, 2\}$, $\vec{G}_i$ est engendré par les translations $t_{\vec{u}_i}$ et $t_{\vec{v}_i}$ et éventuellement la symétrie centrale s_i de centre A_i lorsque $n = 2$. On considère l'application affine $f \in GA(\mathcal{E})$ définie par
$$f(A_1) = A_2 \text{ et } \vec{f}(\vec{u}_1) = \vec{u}_2, \quad \vec{f}(\vec{v}_1) = \vec{v}_2.$$
On a bien $G_2 = fG_1f^{-1}$.
- Soient $n \in \{3, 4, 6\}$ et G_1 et G_2 deux groupes paveurs de partie linéaire d'ordre n. Pour $i \in \{1, 2\}$ on considère r_i une rotation de centre A_i qui engendre $\vec{G}_i$ et $\vec{u}_i \in \Gamma(G_i) \setminus \{0\}$ de norme minimale. Soit $\vec{v}_i = \vec{r}_i(\vec{u}_i)$. Alors $(\vec{u}_i, \vec{v}_i)$ est une famille libre et comme $\|\vec{u}_i\| = \|\vec{v}_i\|$ c'est une base de $\Gamma(G_i)$ (lemme préliminaire). On considère la similitude affine $f \in GA(\mathcal{E})$ définie par
$$f(A_1) = A_2 \text{ et } \vec{f}(\vec{u}_1) = \vec{u}_2, \quad \vec{f}(\vec{v}_1) = \vec{v}_2.$$
On a bien $G_2 = fG_1f^{-1}$.

□

Références.

M. Berger, *Géométrie, Tome 1*
R. Krust, `https://www.lycee-champollion.fr/IMG/PavagesDirects_b.pdf`

161 Isométries d'un espace affine euclidien de dimension finie. Applications en dimensions 2 et 3.
183 Utilisation des groupes en géométrie.

2 Analyse et Probabilités

2.1 Au bord du disque de convergence - théorèmes abéliens et taubériens

Théorème (Abel angulaire)**.** *Soit $f(z) = \sum a_n z^n$ une série entière de rayon de convergence 1. Soit $\theta_0 \in [0, \pi/2[$. On pose ;*

$$\Delta_{\theta_0} = \{z \in \mathbf{C},\ |z| < 1 \text{ et } \exists \rho > 0,\ \exists \theta \in [-\theta_0, \theta_0],\ \ z = 1 - \rho e^{i\theta}\}.$$

Si la série $\sum a_n$ converge alors :

$$\lim_{z \to 1,\, z \in \Delta_{\theta_0}} f(z) = \sum_{n=0}^{+\infty} a_n.$$

Preuve. On note $R_n = \displaystyle\sum_{k=n+1}^{+\infty} a_k$. Une transformation d'Abel donne pour $(n,p) \in \mathbf{N} \times \mathbf{N}^*$ et $z \in \Delta_{\theta_0}$:

$$\sum_{k=n+1}^{n+p} a_k z^k = \sum_{k=n+1}^{n+p} (R_{k-1} - R_k) z^k = R_n z^{n+1} + \sum_{k=n+1}^{n+p-1} R_k (z^{k+1} - z_k) - R_{n+p} z^{n+p}.$$

Soient $\varepsilon > 0$ et $N_0 \in \mathbf{N}$ tel que pour tout $n \geq N_0$, $|R_n| \leq \varepsilon$. Alors, pour $n \geq N_0$:

$$\left| \sum_{k=n+1}^{n+p} a_k z^n \right| \leq \varepsilon |z^{n+1}| + \varepsilon \sum_{k=n+1}^{n+p-1} |z^{k+1} - z^k| + \varepsilon |z^{n+p}| \leq 2\varepsilon + \varepsilon \frac{|z-1|}{1-|z|}.$$

Maintenant, si $z = 1 - \rho e^{i\varphi} \in \Delta_{\theta_0} \cap D(1, \cos\theta_0)$, on a :

$$\frac{|z-1|}{1-|z|} = \frac{\rho}{2\rho\cos\varphi - \rho^2}(1 + |z|) \leq \frac{2}{2\cos\varphi - \rho} \leq \frac{2}{\cos\theta_0}$$

de sorte que

$$\forall z \in \Delta_{\theta_0} \cap D(1, \cos\theta_0),\ \ \forall n \geq N_0,\ \ \forall p \in \mathbf{N}^*,\ \ \left| \sum_{k=n+1}^{n+p} a_k z^n \right| \leq \left(2 + \frac{2}{\cos\theta_0}\right) \varepsilon.$$

Cette dernière majoration est aussi valable pour $z = 1$. Finalement, un critère de Cauchy uniforme justifie la continuité de f en 1 et la limite annoncée. □

Théorème (Taubérien faible)**.** *Soit $f(z) = \sum s_n z^n$ une série entière de rayon de convergence 1 avec $a_n = o(1/n)$. Si :*

$$\exists S \in \mathbf{C}, \quad \lim_{x \to 1,\, x<1} f(x) = S$$

alors $\sum a_n$ converge et $\displaystyle\sum_{n=0}^{+\infty} a_n = S$.

Preuve. Pour $n \in \mathbf{N}$, on note $S_n = \sum_{k=0}^{n} a_k$. On écrit :

$$\forall n \in \mathbf{N}^*,\ \forall x \in]0,1[, \quad S_n - f(x) = \sum_{k=1}^{n} a_k(1-x^k) - \sum_{k=n+1}^{+\infty} a_k x^k$$

et comme $(1-x^k) = (1-x)(1+x+\ldots+x^{k-1}) \leq k(1-x)$ pour $0 < x < 1$ on en déduit :

$$|S_n - f(x)| \leq (1-x)\sum_{k=1}^{n} k|a_k| + \sum_{k=n+1}^{+\infty} \frac{k|a_k|}{n} x^k \leq (1-x)Mn + \frac{\sup_{k>n} k|a_k|}{n(1-x)}$$

où M est un majorant de la suite $(k|a_k|)_k$. Soit $0 < \varepsilon < 1$. On a en particulier :

$$\forall n \in \mathbf{N}^*, \quad \left|S_n - f\left(1 - \frac{\varepsilon}{n}\right)\right| \leq M\varepsilon + \frac{\sup_{k>n} k|a_k|}{\varepsilon}.$$

On choisit N_0 tel que $\sup_{k>N_0} k|a_k| \leq \varepsilon^2$, et on trouve :

$$\forall n \geq N_0, \quad \left|S_n - f\left(1 - \frac{\varepsilon}{n}\right)\right| \leq (M+1)\varepsilon.$$

On choisit $N_1 \geq N_0$ tel que pour $n \geq N_1$, $|f(1-\varepsilon/n) - S| < \varepsilon$ et on trouve :

$$\forall n \geq N_1, \quad |S_n - S| \leq \left|S_n - f\left(1 - \frac{\varepsilon}{n}\right)\right| + \left|f\left(1 - \frac{\varepsilon}{n}\right) - S\right| \leq (M+2)\varepsilon$$

ce qui conclut. □

Théorème (Taubérien fort, Hardy-Littlewood)**.** *Soit $f(z) = \sum s_n z^n$ une série entière de rayon de convergence 1 avec $a_n = \mathcal{O}(1/n)$. Si :*

$$\exists S \in \mathbf{C}, \quad \lim_{x \to 1,\, x<1} f(x) = S$$

alors $\sum a_n$ converge et $\displaystyle\sum_{n=0}^{+\infty} a_n = S$.

Preuve. On se restreint sans mal au cas $S = 0$ (quitte à considérer $a_0 - S$ comme premier terme). On note Φ l'ensemble des fonctions $\varphi : [0,1] \to \mathbf{R}$ telles que

(i) Pour tout $x \in [0,1[$, la série $\sum a_n \varphi(x^n)$ converge

(ii) $\lim_{x \to 1^-} \sum_{n=0}^{+\infty} a_n \varphi(x^n) = 0$.

Clairement, Φ contient les fonctions polynômes qui s'annulent en zéro. De plus, pour $q : x \mapsto x^k$, on a :

$$\forall x \in [0,1[,\ \ (1-x)\sum_{n=0}^{+\infty} x^n q(x^n) = (1-x)\sum_{n=0}^{+\infty}(x^{k+1})^n = \frac{1-x}{1-x^{k+1}} \underset{x\to 1^-}{\longrightarrow} \frac{1}{k+1} = \int_0^1 q(t)dt$$

et le résultat reste vrai par linéarité pour tout polynôme. On va maintenant montrer que $g := \mathbf{1}_{[1/2,1]} \in \Phi$ ce qui conclura puisque :

$$\forall x \in [0,1[,\ \sum_{n=0}^{+\infty} a_n g(x^n) = \sum_{n=0}^{\lfloor -\log 2/\log x \rfloor} a_n.$$

La condition (i) est claire puisque $g(x^n) = 0$ dès que $x^n < 1/2$. Pour le reste, on va approcher g par des polynômes p_1, p_2 tels que :

1. $p_1(0) = p_2(0) = 0$ et $p_1(1) = p_2(1) = 0$
2. $p_1 \le g \le p_2$ sur $[0,1]$
3. $\displaystyle\int_0^1 q(x)dx < \varepsilon$ avec $q(x) = \frac{p_2(x) - p_1(x)}{x(1-x)}$.

On pose $h(x) := \dfrac{g(x)-x}{x(1-x)} = -\dfrac{1}{1-x}\mathbf{1}_{[0,1/2[}(x) + \dfrac{1}{x}\mathbf{1}_{[1/2,1]}(x)$. Soit $\varepsilon > 0$. Un dessin intelligent montre qu'il existe deux fonctions continues s_1 et s_2 telles que

$$s_1 \le h \le s_2 \quad \text{et} \quad \int_0^1 (s_2 - s_1)(t)dt \le \varepsilon.$$

Puis, par le théorème de Weierstrass, on exhibe deux polynômes $\tilde{s}_1$ et $\tilde{s}_2$ tels que $|s_1 - \tilde{s}_1| < \varepsilon$ et $|s_2 - \tilde{s}_2| < \varepsilon$. On définit alors $u_1 = \tilde{s}_1 - \varepsilon$ et $u_2 = \tilde{s}_2 + \varepsilon$ qui sont des polynômes vérifiant :

$$u_1 \le h \le u_2 \ \text{ et } \ , \int_0^1 (u_2 - u_1)(t)dt \le \int_0^2 (s_2 - s_1)(t)dt + 4\varepsilon \le 5\varepsilon.$$

Comme $g(x) = x + x(1-x)h(x)$ sur $[0,1]$, les polynômes

$$p_1(x) = x + x(1-x)u_1(x) \ \text{ et } \ p_2(x) = x + x(1-x)u_2(x)$$

conviennent. Finalement :

$$\forall x \in [0,1[,\ \left|\sum_{n=0}^{+\infty} a_n g(x^n)\right| \le \sum_{n=0}^{+\infty} |a_n||p_1(x^n)| + \sum_{n=0}^{+\infty} |a_n||(g-p_1)(x^n)|.$$

Le premier terme est $< \varepsilon$ pour $x > \lambda$ pour un certain $0 < \lambda < 1$ puisque $p_1 \in \Phi$. Pour le second terme, on écrit :

$$\begin{aligned}\sum_{n=0}^{+\infty} |a_n||(g-p_1)(x)| &\le \sum_{n=1}^{+\infty} |a_n|(p_2 - p_1)(x^n) \\ &\le M\sum_{n=1}^{+\infty} \frac{x^n(1-x^n)}{n} q(x^n) \le M(1-x)\sum_{n=1}^{+\infty} x^n q(x^n)\end{aligned}$$

où $|a_n| \leq M/n$ pour tout $n \in \mathbf{N}^*$ et puisque $(1-x^n) = (1-x)(1+x+\ldots+x^{n-1}) \leq n(1-x)$. Or, on a :

$$\lim_{x\to 1^-} (1-x)\sum_{n=1}^{+\infty} x^n q(x) = \int_0^1 q(t)dt < \varepsilon$$

d'où la conclusion.

□

Références.
V. Beck, J. Malick, G. Peyré, *Objectif agrégation*
X. Gourdon, *Analyse*

207 Prolongement de fonctions. Exemples et applications.
209 Approximation d'une fonction par des polynômes et des polynômes trigonométriques. Exemples et applications. *(le deuxième)*
223 Suites numériques. Convergence, valeurs d'adhérence. Exemples et applications.
230 Séries de nombres réels ou complexes. Comportement des restes ou des sommes partielles des séries numériques. Exemples.
235 Problèmes d'interversion de limites et d'intégrales.
243 Convergence des séries entières, propriétés de la somme. Exemples et applications.

2.2 L'équation de la chaleur dans un anneau

On peut tout faire en justifiant au fur et à mesure avec les théorèmes de Lebesgue, c'est un peu plus long mais on montre l'unicité en même temps et le développement rentre dans des leçons pas drôles. Cependant la preuve de l'unicité à partir de l'énergie transcrit une propriété importante de l'équation de la chaleur. On pourrait aussi parler du principe du maximum.

On cherche à résoudre l'équation de la chaleur dans un anneau, c'est à dire le problème de Cauchy :

$$\begin{cases} \partial_t u(t,x) = \partial^2_{xx} u(t,x), & \text{pour tout } x \in [0,1] \text{ et tout } t > 0 \\ \lim_{t\to 0} u(t,x) = f(x), & \text{pour tout } x \in [0,1] \end{cases} \tag{2.1}$$

où f est une fonction C^2 périodique (de période 1). On cherche u périodique de période 1 sous la forme :

$$u(t,x) = \sum_{n\in\mathbf{Z}} u_n(t) e^{2i\pi nx}$$

où la sommation est pour l'instant purement formelle.

Étape 1 : analyse.

On effectue dans un premier temps des calculs purement formels pour déterminer les coefficients u_n. Le système (2.1) s'écrit alors : pour tout $n \in \mathbf{Z}$,

$$\begin{cases} u_n'(t) = -4\pi^2 n^2 u_n(t) \\ \lim\limits_{t\to 0} u_n(t) = \displaystyle\int_{[0,1]} f(y) e^{-2i\pi ny} dy \end{cases}$$

On trouve pour tout $n \in Z$:

$$u_n(t) = \left(\int_{[0,1]} f(y) e^{-2i\pi n y} dy \right) e^{-4\pi^2 n^2 t}$$

et au prix d'une interversion $\sum$ et $\int$, on a :

$$u(t,x) = \int_{[0,1]} \left(\sum_{n \in Z} e^{2i\pi n(x-y)} e^{-4\pi^2 n^2 t} \right) f(y) dy = K_t * f(x) \tag{2.2}$$

où on définit le *noyau de la chaleur* :

$$K_t(x) := \sum_{n \in \mathbf{Z}} e^{-4\pi^2 n^2 t} e^{2i\pi n x}$$

qui est bien défini comme somme d'une série de fonction normalement convergente sur tout compact de $]0, +\infty[\times[0,1]$.

Étape 2 : synthèse.

On vérifie que la fonction u définie en (2.2) est solution du problème de Cauchy (2.1). Comme pour tout $k \geq 0$ la série de fonctions

$$\sum_{n \in \mathbf{Z}} n^k e^{-4\pi^2 n^2 t} e^{2i\pi n x}$$

est normalement convergente sur tout compact de $]0, +\infty[\times[0,1]$, on peut dériver sous l'intégrale en t et en x. De plus, les fonctions $(t,x) \mapsto e^{-4\pi^2 n^2 t} e^{2i\pi n x}$ sont solutions de l'équation de la chaleur donc il en est de même pour u. De plus, au prix d'une interversion $\sum$ et $\int$ qui est licite, on reconnaît lorsque $t = 0$ dans (2.2), la somme de la série de Fourier de f qui converge bien ponctuellement compte tenu de sa régularité.

Étape 3 : unicité de la solution

Par linéarité de l'équation, si u_1 et u_2 sont solutions de (2.1), alors $u = u_1 - u_2$ est solution du problème :

$$\begin{cases} \partial_t u(t,x) = \partial^2_{xx} u(t,x), \text{ pour tout } x \in [0,1] \text{ et tout } t > 0 \\ u(0,x) = 0, \text{ pour tout } x \in [0,1] \end{cases}$$

Posons pour tout $t \in [0, +\infty[$,

$$E(t) = \int_0^1 (u(t,x))^2 dx.$$

On peut dériver sans crainte sous l'intégrale compte tenu du caractère compact de $[0,1]$ et de la régularité de u :

$$E'(t) = \int_0^1 2u(t,x) \partial_t u(t,x) dx = \int_0^1 2u(t,x) \partial^2_{xx} u(t,x) dx = -2 \int_0^1 (\partial_x u(t,x))^2 dx.$$

où la dernière égalité provient d'une intégration par partie (le caractère 1-périodique de u permet d'annuler le terme de bord). Ainsi, E est positive et décroissante sur $[0, +\infty[$. Comme $E(0) = 0$, E est identiquement nulle, ce qui prouve que $u = 0$.

Le cas non périodique

Théorème. *Si $f \in \mathscr{S}(\mathbf{R})$, alors il existe un et un seul élément $u \in C^\infty(\mathbf{R}_+, \mathscr{S}(\mathbf{R}))$ tel que*

$$\begin{cases} \partial_t u(t,x) = \partial^2_{xx} u(t,x), & \text{pour tout } x \in \mathbf{R} \text{ et tout } t > 0 \\ \lim_{t\to 0} u(t,x) = f(x), & \text{pour tout } x \in \mathbf{R} \end{cases}$$

Preuve. La régularité de u nous autorise à prendre la transformée de Fourier de l'équation, qui devient alors :

$$\partial_t \hat{u}(t,\xi) = -4\pi^2\xi^2 \hat{u}(t,\xi) \quad \text{et} \quad \hat{u}(0,\xi) = \hat{f}(\xi).$$

Comme dans le cas périodique, c'est une équation différentielle simple à résoudre :

$$\hat{u}(t,\xi) = e^{-4\pi^2\xi^2 t}\hat{f}.$$

Et nécessairement,

$$u(t,x) = \mathscr{F}^{-1}(e^{-4\pi^2\xi^2 t}\hat{f})(x).$$

Comme on sait calculer la transformée d'une gaussienne, on peut écrire plus simplement,

$$u(t,x) = K_t * f(x)$$

où le noyau de la chaleur K_t est défini par :

$$K_t(x) = \frac{1}{\sqrt{4\pi t}} e^{-\frac{x^2}{4t}}.$$

Compte-tenu de la régularité, on peut vérifier facilement que u ainsi défini est solution de l'équation de la chaleur. □

Remarque. On ne dit absolument rien sur l'existence ou l'unicité de la solution dans un autre cadre fonctionnel que celui-là. En particulier, on pourra vérifier que la fonction :

$$v(t,x) = \begin{cases} \frac{x}{t^3/2} e^{-\frac{x^2}{4t}} & \text{si } t > 0 \\ 0 & \text{si } t = 0 \end{cases}$$

est solution de l'équation de la chaleur avec condition initiale nulle. Pour tout $t \in \mathbf{R}_+$, $v(t) \in \mathscr{S}(\mathbf{R})$ mais v n'est bornée sur aucun voisinage de $(0,0)$ (donc même pas continue).

Références.

B. Candelpergher, *Calcul Intégral*
J. Rauch, *Partial differential equations*

209 Approximation d'une fonction par des polynômes et des polynômes trigonométriques. Exemples et applications.
213 Espaces de Hilbert. Bases hilbertiennes. Exemples et applications. *(euh...)*
222 Exemples d'équations aux dérivées partielles linéaires.
235 Problèmes d'interversion de limites et d'intégrales.*(mais si)*
239 Fonctions définies par une intégrale dépendant d'un paramètre. Exemples et applications. (lol)
246 Séries de Fourier. Exemples et applications.

2.3 Des bases presque orthogonales et des opérateurs compacts

Théorème (Paley, Wiener, Birkhoff, Rota, Nagy, Lax)**.** *Soit H un espace de Hilbert muni d'une base orthonormée $\{x_n\}$. Soit $\{y_n\}$ une famille d'éléments de H « proches »des $\{x_n\}$ au sens où :*

$$\sum \|x_n - y_n\|^2 < +\infty.$$

Si aucun y_i n'est dans l'adhérence du sous-espace engendré par les autres y_n, alors $\{y_n\}$ est une famille totale. En particulier, si les y_n sont orthonormés, alors $\{y_n\}$ est une base hilbertienne.

Preuve. On va commencer par le résultat de base sur les opérateurs compacts.

Théorème. *Soit $C : X \to X$ un opérateur compact d'un espace de Banach. On note $T = I - C$.*

(i) Si $(C_n)_n$ est une suite d'opérateurs compacts qui converge en norme d'opérateur vers un opérateur C. Alors C est compact.

(ii) $\mathrm{Im}(T)$ *est fermée.*

(iii) (Fredholm) Si T est injectif alors T est surjectif.

Preuve. (i) Soit $\varepsilon > 0$ et $n \in \mathbf{N}$ tel que $\|C_n - C\| < \varepsilon$. Puisque l'image de la boule unité par C_n peut être recouverte par un nombre fini de boules de rayon ε, l'image de la boule unité par C peut être recouverte par un nombre fini de boules de rayons 2ε.

(ii) Soit $(y_k)_k$ une suite de points de $\mathrm{Im}\, T$ qui converge vers $y \in X$:

$$\lim_{k\to+\infty} y_k = y, \quad y_k = Tx_k.$$

On note $d_k = \mathrm{dist}(x_k, \mathrm{Ker}\, T)$ et on montre que la suite des $(d_k)_k$ est bornée. Pour cela, on considère une suite $(z_k)_k$ de points de $\mathrm{Ker}\, T$ telle que $|w_k| := |x_k - z_k| < 2d_k$. Clairement,

$$Tw_k = Tx_k - Tz_k = y_k.$$

Si la suite $(d_k)_k$ n'était pas bornée, alors $u_k := w_k/d_k$ vérifie $|u_k| < 2$ et $Tu_k = y_k/d_k \to 0$ puisque $(y_k)_k$ est bornée. Par définition de T compacité de C et continuité de T on a quitte à extraire :

$$u_k - Cu_k = Tu_k \to 0 \implies u_k \to u \in \mathrm{Ker}\, T.$$

C'est contradictoire avec $|u_k - u| \geq 1$ qui découle de la définition de u_k. Donc $(d_k)_k$ est bornée, toute comme $(w_k)_k$. Mais alors, on peut extraire une sous-suite de $(Cw_k)_k$ convergente et quitte à extraire :

$$w_k - Cw_k = y_k \to y \implies w_k \to w.$$

Par continuité de T, on a :

$$w - Cw = Tw = y.$$

(*iii*) On suppose que T est injectif et on considère la suite :

$$X_0 := X \quad \text{et} \quad \forall n \in \mathbf{N},\ X_{n+1} := TX_n = T^{n+1}X \subset X_n.$$

On veut montrer que $X_1 = X$ donc on suppose le contraire. Par injectivité de T, les inclusions sont strictes. De plus comme :

$$T^n = (I - C)^n = I + \sum_{k=1}^{n} \binom{n}{k} (-1)^k C^k$$

le premier point montre que l'image de T^n est fermée pour tout $n \in \mathbf{N}$. Ainsi par le lemme super important ci-dessous, pour tout $k \in \mathbf{N}$, il existe $x_k \in X_k$ tel que :

$$|x_k| = 1 \quad \text{et} \quad \operatorname{dist}(x_k, X_{k+1}) \geq \frac{1}{2}.$$

Alors, si $m < n$:

$$Cx_m - Cx_n = x_m - (Tx_m + x_n - Tx_n) \in x_m + X_{m+1}$$

et $|Cx_m - Cx_n| \geq 1/2$. C'est contradictoire avec la compacité de C car alors $(Cx_n)_n$ ne contient aucune sous-suite convergente.

□

On peut maintenant prouver le théorème à proprement parler. Prenons $u \in H$ et écrivons :

$$u = \sum a_n x_n, \quad a_n = \langle x_n, u \rangle.$$

On définit ensuite l'application linéaire $B : H \to H$ définit par :

$$Bu = \sum a_n (y_n - x_n) + \sum a_n x_n = \lim_{N \to +\infty} \sum_{n=0}^{N} a_n y_n =: \sum a_n y_n.$$

Il n'y a pas de problème de définition puisque :

$$\sum |a_n| \|y_n - x_n\| \leq \left(\sum |a_n|^2 \right)^{1/2} \left(\sum \|y_n - x_n\|^2 \right)^{1/2} \leq C\|u\| < +\infty.$$

On vient même de montrer que $B - I$ est une application linéaire continue. On prétend que $B - I$ est un opérateur compact. En effet, on peut écrire $B - I$ comme la limite (en norme d'opérateur) des opérateurs :

$$G_N : u \mapsto G_N u := \sum_{n=0}^{N} a_n (y_n - x_n)$$

qui sont compacts puisque leur image est de dimension finie. Comme $B = I + (B - I)$, il suffit de montrer que B est injectif pour conclure que B est surjectif. C'est évident car

$$Bu = 0 \;\Rightarrow\; \sum a_n y_n = 0 \;\Rightarrow\; \forall n \in \mathbf{N},\ a_n = 0 \;\; i.e \;\; u = 0$$

par l'hypothèse d'indépendance des y_n. □

Lemme (Super important)**.** *Soit V un espace vectoriel normé et $W \subset V$ un sous-espace vectoriel strict fermé de V. Alors il existe $v \in V$ tel que $\|v\| = 1$ et $d(v, W) \geq 1/2$.*

Preuve. On choisit $y \in V \setminus W$ et $w \in W$ tel que $\|y - w\| \leq 2d(y, W)$. On vérifie que $v := \dfrac{y - w}{\|y - w\|}$ convient. □

Référence. P. D. Lax, *Functional Analysis*

203 Utilisation de la notion de compacité.
208 Espaces vectoriels normés, applications linéaire continues. Exemples.
213 Espaces de Hilbert. Bases hilbertiennes. Exemples et applications.

2.4 À propos de l'équation de Schrödinger linéaire

Il faut un peu adapter pour pouvoir présenter ce développement : je pense qu'une bonne solution est de tout montrer sauf l'unicité au sens faible et d'éventuellement rajouter la transformée de Fourier de la gaussienne. Le détail de l'unicité pour les solutions faibles est probablement dangereux à présenter. À noter que PLEIN de leçons d'analyse peuvent contenir ce développement au moins en exemple : 201, 202, 205, 207, 208, 213, 222, 234, 235, 239, 241, 250...

On cherche à résoudre le problème suivant :

$$\begin{cases} i\partial_t u(t,x) + \Delta u(t,x) = 0 \quad \text{dans } \mathcal{D}'(\mathbf{R} \times \mathbf{R}^d) \\ u(0,\cdot) = u_0 \in H^s(\mathbf{R}^d), \quad s \in \mathbf{R} \end{cases} \tag{S}$$

On va d'abord s'intéresser aux solutions *tempérées* de (S).

Théorème 5 (Solutions tempérées)**.** *Pour tout $u_0 \in \mathscr{S}(\mathbf{R}^d)$, il existe une unique solution de* (S) *appartenant à $C^\infty(\mathbf{R}, \mathscr{S}(\mathbf{R}^d))$.*

Preuve. Soit u une telle solution. Alors en prenant la transformée de Fourier selon x de (S), on trouve [1] :

$$\partial_t \hat{u} = \widehat{i\Delta u}(t,\xi) = -4i\pi^2 |\xi|^2 \hat{u}(t,\xi).$$

1. La régularité est importante : le théorème de convergence dominée permet de montrer que si $u \in C^\infty(\mathbf{R}, \mathscr{S}(\mathbf{R}^d))$, alors $t \mapsto \widehat{u(t)}$ est aussi dans $C^\infty(\mathbf{R}, \mathscr{S}(\mathbf{R}^d))$, ce qui justifie la notation $\hat{u}(t,\xi)$. Mais ça ne suffirait pas de supposer que pour tout $t \in \mathbf{R}$, $u(t) \in \mathscr{S}(\mathbf{R}^d)$ (voir un contre-exemple à la fin du développement sur l'équation de la chaleur).

C'est une équation différentielle linéaire en t avec un paramètre ξ. Son unique solution s'écrit :

$$\hat{u}(t,\xi) = e^{-4\pi^2 it|\xi|^2}\widehat{u_0}(\xi).$$

On en déduit l'unique solution tempérée de (S) en prenant la transformée de Fourier inverse dans $\mathscr{S}(\mathbf{R}^d)$. Et il n'y a pas de problème pour dériver indéfiniment en t puisque tout est dans $\mathscr{S}(\mathbf{R}^d)$. □

Il convient de remarquer que l'on peut écrire cette solution :

$$u(t,x) = e^{it\Delta}u_0(x)$$

où on a défini pour $t \in \mathbf{R}$, l'opérateur :

$$\begin{array}{rccc} e^{it\Delta}: & \mathscr{S}(\mathbf{R}^d) & \longrightarrow & \mathscr{S}(\mathbf{R}^d) \\ & f & \longmapsto & e^{it\Delta}f := \mathcal{F}^{-1}\big(e^{-4\pi^2 it|\xi|^2}\widehat{f}\big). \end{array}$$

Lemme 6. *Listons quelques propriétés importantes :*

- *L'opérateur $e^{it\Delta}$ se prolonge de manière unique en une isométrie de $H^s(\mathbf{R}^d)$:*

$$\forall f \in H^s(\mathbf{R}^d), \quad \|e^{it\Delta}f\|_{H^s} = \|f\|_{H^s} \tag{2.3}$$

- *Pour tout $t_1, t_2 \in \mathbf{R}$, $e^{i(t_1+t_2)\Delta} = e^{it_1\Delta}e^{it_2\Delta}$.*
- *Pour tout $f \in H^s(\mathbf{R}^d)$, $e^{it\Delta}f \in C^0(\mathbf{R}, H^s)$*

Preuve. • Comme $\mathscr{S}(\mathbf{R}^d)$ est dense dans $H^s(\mathbf{R}^d)$, il suffit de vérifier (2.3) pour tout $f \in \mathscr{S}(\mathbf{R}^d)$ pour conclure grâce au théorème de prolongement des applications linéaires continues. Par définition de la norme H^s :

$$\|e^{it\Delta}f\|_{H^s} = \big\|e^{-4\pi^2 it|\xi|^2}(1+|\xi|^2)^{s/2}\hat{f}\big\|_{L^2(\mathbf{R}^d)} = \|f\|_{H^s}.$$

- La propriété de semi-groupe se lit dans la définition pour $f \in \mathscr{S}(\mathbf{R}^d)$ et demeure vraie dans H^s par densité et continuité.
- Vérifions la continuité en $t = 0$: pour tout $g \in H^s(\mathbf{R}^d)$,

$$\|e^{it\Delta}g - g\|_{H^s}^2 = \int_{\mathbf{R}^d}(1+|\xi|^2)^s\big|(e^{-4\pi^2 it|\xi|^2}-1)\hat{g}(\xi)\big|^2 d\xi \longrightarrow 0$$

par convergence dominée.

□

Théorème 7. *Pour tout $u_0 \in H^s(\mathbf{R}^d)$, le problème est bien posé dans $C^0(\mathbf{R}, H^s) \subset \mathcal{D}'(\mathbf{R}\times\mathbf{R}^d)$: $u = e^{it\Delta}u_0 \in C^0(\mathbf{R}, H^s)$ est l'unique solution de* (S).

Preuve. **Existence.** Avant toute chose, il faut préciser le sens que l'on donne à la distribution $u = e^{it\Delta}u_0$ (c'est à dire au sens de l'injection $C^0(\mathbf{R}, H^s) \hookrightarrow \mathcal{D}'(\mathbf{R}\times\mathbf{R}^d)$) . Pour tout $\varphi \in \mathcal{D}(\mathbf{R}\times\mathbf{R}^d)$, on définit :

$$\langle u, \varphi\rangle = \int_{\mathbf{R}}\big\langle u(t,\cdot), \varphi(t,\cdot)\big\rangle dt$$

et on vérifie que c'est bien une distribution : supposons que φ soit à support dans $K_t \times K_x$.

$$|\langle u, \varphi\rangle| \leq \int_{\mathbf{R}^d} \|u(t)\|_{H^s} \|\varphi(t)\|_{H^{-s}} dt \leq \|u_0\|_{H^s} \int_{K_t} \|\varphi(t)\|_{H^{-s}} dt.$$

Et on voit que :

$$\|\varphi(t)\|^2_{H^{-s}} \leq \|\varphi(t)\|^2_{H^{\lfloor |s| \rfloor + 1}} \leq C(K_x) \sum_{|\alpha| \leq \lfloor |s| \rfloor + 1} \|\partial_x^\alpha \varphi(t)\|^2_{L^\infty(\mathbf{R}^d)}.$$

Montrons maintenant que u est solution de (S) au sens des distributions. Considérons une suite $(u_0^n)_n$ d'éléments de $\mathscr{S}(\mathbf{R}^d)$ qui converge vers u_0 dans $H^s(\mathbf{R}^d)$ et posons $u^n = e^{it\Delta} u_0^n$. Alors pour tout $\varphi \in \mathcal{D}(\mathbf{R} \times \mathbf{R}^d)$:

$$|\langle u^n - u, \varphi\rangle| \leq \|u_0^n - u_0\|_{H^s} \cdot C \sum_{|s|+1} \|\partial_x^\alpha \varphi\|_{L^\infty(\mathbf{R}^d)} \to 0$$

donc $u^n \to u$ au sens des distributions. Pareil pour leurs dérivées à tout ordre et donc :

$$i\partial_t u(t,x) + \Delta u(t,x) = 0 \quad \text{dans } \mathcal{D}'(\mathbf{R} \times \mathbf{R}^d).$$

Et par continuité de $e^{it\Delta}$ par rapport à t, on a bien sûr $u \in C^0(\mathbf{R}, H^s)$ et $u(0) = u_0$.

Unicité. Pour une équation homogène linéaire, il suffit de montrer que la solution nulle est l'unique solution lorsque $u_0 = 0$. Soit donc u une solution de (**??**) avec $u_0 = 0$. L'idée générale pour montrer l'unicité de la solution des problèmes linéaires est de se ramener à montrer l'existence pour le problème dual inhomogène. Plus précisément il s'agit d'écrire :

$$\langle \partial_t u - i\Delta u, \varphi\rangle_{\mathcal{D}',\mathcal{D}} = 0 \iff \langle u, \partial_t \varphi + i\Delta\varphi\rangle_{\mathcal{D}',\mathcal{D}} = 0.$$

Dans $C^0(\mathbf{R}, H^s)$ cela équivaut à :

$$\int_{\mathbf{R}} \big\langle u(t,\cdot), \big(\partial_t\varphi + i\Delta\varphi\big)(t,\cdot)\big\rangle_{\mathscr{S}',\mathscr{S}} \, dt = 0 \tag{2.4}$$

Et on veut montrer :

$$\forall T \in \mathbf{R}, \ \ u(T,\cdot) = 0 \text{ dans } H^s(\mathbf{R}^d) \ \ i.e \ \ \forall T \in \mathbf{R}, \ \forall h \in \mathscr{S}(\mathbf{R}^d), \ \ \langle u(T,\cdot), h\rangle = 0.$$

De deux choses l'une : d'abord, on remarque que $\mathscr{S}(\mathbf{R} \times \mathbf{R}^d) \subset C^\infty(\mathbf{R}; \mathscr{S}(\mathbf{R}^d)$ et ensuite on sait résoudre le problème dual :

$$\left\{ \begin{array}{rcl} \partial_t\varphi + i\Delta\varphi & = & 0 \text{ dans } \mathscr{S}(\mathbf{R} \times \mathbf{R}^d) \\ \varphi(T,\cdot) & = & h \in \mathscr{S}(\mathbf{R}^d) \end{array} \right.$$

il suffit de prendre $\varphi(t,x) = e^{-i(t-T)\Delta} h$. La conclusion suit alors directement de l'identité *IPP-like* :

$$\forall T > 0, \ \int_0^T \big\langle u(t,\cdot), \big(\partial_t\varphi + i\Delta\varphi\big)(t,\cdot)\big\rangle_{\mathscr{S}',\mathscr{S}} \, dt = \langle u(T,\cdot), \varphi(T,\cdot)\rangle_{\mathscr{S}',\mathscr{S}}$$

laquelle résulte [2] d'un jeu de découpages-collages.

2. Pour B. Perthame dans *Transport Equations in Biology* page 153, c'est évident mais j'ai pas compris pourquoi.

Étape 1. On montre que pour tous $t_1 < t_2$ et tout $\varphi \in \mathcal{D}(\mathbf{R} \times \mathbf{R}^d)$:

$$\int_{t_1}^{t_2} \big\langle u(t,\cdot), \partial_t \varphi(t,\cdot) + i\Delta\varphi(t,\cdot)\big\rangle dt = \big\langle u(t_2), \varphi(t_2)\big\rangle - \big\langle u(t_1), \varphi(t_1)\big\rangle.$$

Prenons pour tout n suffisamment grand, $\chi_n \in \mathcal{D}(\mathbf{R})$ dont le support est contenu dans $[t_1, t_2]$ et valant 1 sur $[t_1 + 1/n, t_2 - 1/n]$. Alors (2.4) appliqué à $\chi_n\varphi$ s'écrit :

$$\begin{aligned} 0 &= \int_{t^1}^{t_2} \big\langle u(t), (\partial_t + i\Delta)(\chi_n\varphi)(t)\big\rangle dt \\ &= \int_{t_1}^{t_2} \big\langle u(t), \chi_n(\partial_t + i\Delta)(\varphi(t))\big\rangle dt + \int_{t_1}^{t_2} \langle u(t), \chi_n'(t)\varphi(t)\big\rangle dt =: I_1^n + I_2^n \end{aligned}$$

La quantité $|I_1^n|$ est majorée indépendamment de n par les mêmes calculs que précédemment et par le théorème de convergence dominée :

$$I_1^n \underset{n\to+\infty}{\longrightarrow} \int_{t_1}^{t_2} \big\langle u(t), (\partial_t + i\Delta)(\varphi(t))\big\rangle dt.$$

Quant à I_2^n, on écrit :

$$I_2^n = \int_{t_1}^{t_1+1/n} \langle u(t), \chi_n'(t)\varphi(t)\big\rangle dt + \int_{t_2-1/n}^{t_2} \langle u(t), \chi_n'(t)\varphi(t)\big\rangle dt$$

et pour la première intégrale :

$$\int_{t_1}^{t_1+1/n} \langle u(t), \chi_n'(t)\varphi(t)\big\rangle dt = \int_{t_1}^{t_1+1/n} \chi_n'(t) \underbrace{\Big(\big\langle u(t), \varphi(t)\big\rangle - \big\langle u(t_1), \varphi(t_1)\big\rangle\Big)}_{\to 0} dt + \langle u(t_1), \varphi(t_1)\rangle.$$

Idem pour la seconde intégrale. Le résultat annoncé s'ensuit. Tout cela est vrai parce que $\chi_n\varphi(t)$ converge vers $\varphi(t)$ dans $\mathscr{S}(\mathbf{R}^d)$

Étape 2. Le résultat tient toujours lorsque $\varphi \in C^\infty(\mathbf{R}, \mathscr{S}(\mathbf{R}^d))$.

Il suffit d'appliquer l'étape 1 à $\chi_1(t)\chi_2(\varepsilon x)\varphi(t,x)$ où $\chi_1 \in \mathcal{D}(\mathbf{R})$ vaut 1 sur $[t_1, t_2]$ et $\chi_2 \in \mathcal{D}(\mathbf{R}^d)$ vaut 1 en 0. Le théorème de convergence dominée permet de conclure.

□

Quelques remarques complémentaires

- On dispose d'une formule explicite pour calculer $e^{it\Delta}f$ lorsque $f \in L^2(\mathbf{R}^d)$:

$$(e^{it\Delta}f)(x) = \left(\frac{e^{i|\cdot|^2/4t}}{(4\pi it)^{d/2}} * f\right)(x)$$

ce que l'on peut voir en calculant la transformée de Fourier d'une gaussienne de paramètre complexe. La distribution :

$$E(t,x) = \frac{e^{i|x|^2/4t}}{(4\pi it)^{d/2}} \quad \text{si } t > 0 \quad \text{et} \quad E(0,\cdot) = \delta \in H^{-d/2-\varepsilon}$$

vérifie $E(t) = e^{it\Delta}\delta$ et est appelée solution fondamentale puisqu'on obtient toutes les autres solutions en fonction de cette solution.

- De l'intérêt de chercher un solution dans H^s : pour $s = 1$, H^1 est l'espace d'énergie. Et comme $\delta \in H^{d/2-\varepsilon}$ c'est comme ça qu'on justifie la définition de la solution fondamentale.
- C'est une EDP dispersive qui ne rentre pas dans la classification standard des EDP linéaires d'ordre 2.
- On n'a pas utilisé la forme de $e^{it\Delta}$: c'est une méthode assez générale pour les problèmes du type $P(\partial_t, \partial_x)u = 0$. En particulier, on peut faire pareil pour l'équation de la chaleur, des ondes.
- Lorsque l'équation n'est plus homogène, on dispose de la formule de Duhamel :

$$u(t) = e^{it\Delta}u_0 - i\int_0^t e^{i(t-s)\Delta}F(s)ds.$$

Référence. J. Rauch, *Partial differential equations*

222 Exemples d'équations aux dérivées partielles linéaires.
250 Transformation de Fourier. Applications.
+ toutes celles citées au début.

2.5 Hypercyclicité et critère de Kitai

Soit (E, d) un $\mathbf{C}$-espace vectoriel métrique complet et séparable dont S est une partie dénombrable dense. Soit A un endomorphisme continu de E. On dit qu'un point $x \in E$ est *hypercyclique* pour A lorsque son orbite $\{A^n(x),\ n \in \mathbf{N}\}$ est dense dans E. On note $HC(A)$ l'ensemble des points hypercycliques pour A.

Théorème. *$HC(A)$ est soit l'ensemble vide, soit un G_δ dense dans E.*

Preuve. Il est facile de voir que :

$$HC(A) = \{x \in E/\ \forall (s,k) \in S \times \mathbf{N}^*,\ \exists n \in \mathbf{N}^*,\ d(A^n(x), s) < 1/k\}. \tag{2.5}$$

En termes ensemblistes :

$$HC(A) = \bigcap_{(s,k)\in S\times\mathbf{N}^* n\in\mathbf{N}^*} \bigcup (A^n)^{-1}\big(B(s, 1/k)\big)$$

Puisque S est dénombrable et A est continu, $HC(A)$ est bien une intersection dénombrable d'ouverts. Supposons que $HC(A) \neq \emptyset$ et prenons $x \in HC(A)$. Pour tout $n \in \mathbf{N}$, $A^n(x) \in HC(A)$ (car une partie dénombrable dense privée d'un nombre fini de points reste dense) donc l'orbite de x, qui est dense, est contenue dans $HC(A)$ qui l'est tout autant. □

Lorsqu'on est dans le second cas, l'opérateur A est dit *hypercyclique*.

Théorème (Critère de Kitaï)**.** *Supposons qu'existent X et Y deux parties denses de E et $B : Y \to Y$ un opérateur vérifiant les trois conditions suivantes :*

$$(i)\ x \in X \Rightarrow A^n(x) \underset{n\to+\infty}{\longrightarrow} 0; \quad (ii)\ y \in Y \Rightarrow B^n(x) \underset{n\to+\infty}{\longrightarrow} 0; \quad (iii)\ y \in Y \Rightarrow AB(y) = y$$

alors A est hypercyclique.

Preuve. Reprenons l'égalité (2.5) que l'on écrit ici :

$$HC(A) = \bigcap_{(s,k)\in S\times\mathbf{N}^*} \Omega_{s,k}, \quad \text{où} \quad \Omega_{s,k} = \bigcup_{n\in\mathbf{N}^*} (A^n)^{-1}\big(B(s,1/k)\big)$$

Comme E est un espace métrique complet, il suffit de montrer que chacun des $\Omega_{s,k}$ est dense pour conclure que A est hypercyclique (c'est le théorème de Baire). Soient donc $b \in E$ que l'on cherche à approcher à $\varepsilon > 0$ près par un élément de $\Omega_{s,k}$. Puisque X et Y sont denses, posons pour $n \in \mathbf{N}$

$$\rho_n = x + B^n(y)$$

où $d(x,b) < \varepsilon/2$ et $d(y,s) < 1/(2k)$. Grâce à la condition (ii), on sait que pour $n \in \mathbf{N}$ suffisamment grand,

$$d(\rho_n, b) \leq d(\rho_n, x) + d(x,b) < \varepsilon.$$

De plus, comme pour tout $n \in \mathbf{N}$, la condition (iii) implique :

$$A^n(\rho_n) = A^n(x) + y$$

on a pour n suffisamment grand et grâce à (i) :

$$d(A^n(\rho_n), s) \leq d(A^n(x), s) + d(y,s) < 1/k \quad i.e \quad \rho_n \in (A^n)^{-1}\big(B(s,1/k)\big) \subset \Omega_{s,k}$$

et la conclusion suit. □

Quelques applications

- Lorsque E est de dimension finie, A n'est jamais hypercyclique. En effet, quitte à décomposer E en somme directe des sous-espaces caractéristiques de A, on peut considérer que A est de la forme $A = \lambda Id + N$ où N est un endomorphisme nilpotent. Ainsi, notant d la dimension ambiante, on a pour tout $n \geq d$:

$$A^n = \sum_{k=0}^{d} \binom{n}{k} \lambda^{n-k} N^k.$$

  Soit alors $x \in E$. En notant $y_k = N^k(x)$ pour $k \in \{0, \ldots, d-1\}$, on a donc :

$$A^n(x) = \sum_{k=0}^{d-1} \alpha_k^{(n)} y_k.$$

  Pour que cette orbite soit dense il faut d'une part que les y_k forment une base de l'espace et d'autre part que les suites $(\alpha_k^{(n)})_{n\in\mathbf{N}}$ soient denses dans $\mathbf{C}$. Cette dernière condition n'est jamais réalisée puisqu'on peut voir que ces suites tendent en module ou bien vers 0 ou bien vers $+\infty$ ou bien sont de module constant.

- Soient $E = H(\mathbf{C})$ l'espace des fonctions entières (qui est métrisable, complet et séparable pour la topologie de la convergence uniforme sur tout compact) et T l'opérateur de translation qui à $f \in H(\mathbf{C})$ associe l'application $z \mapsto f(z+1)$. On montre que T est hypercyclique en appliquant le critère de Kitaï avec l'opérateur

$$B : f \in H(\mathbf{C}) \mapsto T^{-1}f \in H(\mathbf{C}), \quad \text{où} \quad T^{-1}f(z) := f(z-1) \quad \text{pour tout } z \in \mathbf{C}$$

et avec les parties

$$X = \left\{ z \mapsto e^{-z} P(z),\ \ P \in \mathbf{C}[X] \right\} \ \text{ et } \ Y = \left\{ z \mapsto e^{z} Q(z),\ \ Q \in \mathbf{C}[X] \right\}$$

qui sont denses dans $H(\mathbf{C})$ puisque l'espace des polynômes complexes l'est pour la convergence uniforme sur tout compact (c'est le développement en série de Taylor).

- Dans le même esprit, on montre que l'opérateur de dérivation sur $H(\mathbf{C})$ est hypercyclique en considérant $X = Y = \mathbf{C}[X]$ et B l'opérateur qui à un polynôme associe son unique primitive nulle en zéro.

Référence. S. Gonnord, N. Tosel, *Topologie et analyse fonctionnelle*

202 Exemples de parties denses et applications.
205 Espaces complets. Exemples et applications.
226 Suites vectorielles et réelles définies par une relation de récurrence $u_{n+1} = f(u_n)$. Exemples. Applications à la résolution approchée d'équations.

2.6 La construction du mouvement Brownien par Paul Lévy

S'expose avec brio en vingt minutes en anglais. En français et en quinze minutes il faut parler un peu plus vite et passer rapidement sur le début : ne pas réécrire la définition et commencer tout de suite par un dessin pour expliquer la construction (voir l'annexe de la leçon 262).

Définition 8 (Brownian Motion)**.** A standard Brownian Motion on $[0,1]$ is a stochastic process $\{B_t,\ 0 \leq t \leq 1\}$ with $B_0 = 0$ and which satisfies the following properties :

(*i*) For all $t, s \in [0,1]$ such that $t + s \in [0,1]$, $B_{t+s} - B_s$ is $\mathcal{N}(0,t)$

(*ii*) If $0 \leq t_1 < t_2 < \ldots < t_n \leq 1$, then the increments B_{t_1}, $B_{t_2} - B_{t_1}$, $\ldots$, $B_{t_n} - B_{t_{n-1}}$ are independent.

Théorème 9 (Lévy)**.** *It exists.*

Preuve. The idea is to construct the Brownian Motion on the set of dyadic numbers :

$$\mathcal{D} = \bigcup_{n \in \mathbf{N}} \mathcal{D}_n \ \text{ where } \ \mathcal{D}_n = \left\{ \frac{k}{2^n},\ k = 0, 1, \ldots, 2^n \right\}$$

and to interpolate in between.

For $n \in \mathbf{N}$, let us consider the Schauder tent functions f_{nk} defined by :

$$f_{nk}(t) = \begin{cases} 2^{-(n+2)/2} & \text{if } t = (k + \frac{1}{2})2^{-n} \\ 0 & \text{if } t \notin (k2^{-n}, (k+1)2^{-n}) \\ \text{linear} & \text{in between} \end{cases}$$

Suppose that we can construct a sequence $(X, X_{nk})_{n,k}$ of i.i.d $\mathcal{N}(0,1)$ random variables. We define for $n \in \mathbf{N}$:

$$F_n(t) = \sum_{k=0}^{2^n - 1} X_{nk} f_{nk}(t).$$

Step 1. *Almost surely the series*

$$B_t = tX + \sum_{n=0}^{+\infty} F_n(t)$$

converges uniformly for $0 \leq t \leq 1$.

Let us define for $m \leq 0$ the partial sums

$$B_t^{(m)} = tX + \sum_{n=0}^{m} F_n(t),\ m \geq 0$$

it is sufficient to show that with probability one $(B_t^{(m)})_m$ converges uniformly on $[0,1]$. Note that since the $f_{n,k}$ have disjoint supports :

$$\sup_{t\in[0,1]} |B_t^{(m)} - B_t^{(m-1)}| = \|F_m\|_\infty \leq 2^{-(m+2)/2} \max\{|X_{mk}|,\ k = 0, \ldots, 2^m - 1\}.$$

The normal distribution has the following property which is an easy consequence of a change of variables [3].

Lemme 10. *Let* Y *be a* $\mathcal{N}(0,\sigma^2)$ *random variable. Then for* $\lambda > 0$ *:*

$$\mathbf{P}(|Y| \geq \lambda) \leq \sqrt{\frac{2}{\pi}} \frac{\sigma}{\lambda} e^{-\frac{\lambda^2}{2\sigma^2}}.$$

As a consequence,

$$\begin{aligned} \mathbf{P}\left(2^{-\frac{m+2}{2}} \max\{|X_{mk}|,\ k = 0, \ldots, 2^m - 1\} > \frac{1}{m^2}\right) &\leq 2^m \mathbf{P}\left(|X_{m1}| \geq \frac{2^{-(m+2)/2}}{m^2}\right) \\ &\leq \frac{1}{\sqrt{2\pi}} 2^{\frac{m}{2}-1} m^2 e^{-\frac{2^{m+1}}{m^4}}. \end{aligned}$$

Since the right-hand side is summable, by the Borel-Cantelli lemma :

$$\sup_{t\in[0,1]} |B_t^{(m+1)} - B_t^{(m)}| \leq \frac{1}{m^2} \text{ for large enough } m \text{ a.s.}$$

and the uniform convergence follows. Note that since the functions F_n are continuous, so is $t \mapsto B_t$ almost surely.

Step 2. $\{B_t,\ 0 \leq t \leq 1\}$ *is a standard Brownian motion on* $[0,1]$.

We will prove by induction on $m \geq 0$:

$(i)_m$ For all $t, s \in \mathcal{D}_m$ such that $t + s \in [0,1]$, $B_{t+s} - B_t$ is $\mathcal{N}(0,t)$

$(ii)_m$ If $0 \leq t_1 < t_2 < \ldots < t_n \leq 1$ with $t_i \in \mathcal{D}_m$, then the increments B_{t_1}, $B_{t_2} - B_{t_1}$, $\ldots$, $B_{t_n} - B_{t_{n-1}}$ are independent.

3. Write that $\mathbf{P}(|Y| \geq \lambda) \leq \frac{2}{\sigma\sqrt{2\pi}} \int_{y\geq\lambda} e^{-\frac{y^2}{2\sigma^2}} \frac{y}{\lambda} dy$.

We will then deduce (i) and (ii) on the set $\mathcal{D}$ and then on $[0,1]$ by density of the dyadic numbers.

Clearly $(i)_m$ and $(ii)_m$ are true for $m = 0$ since X is a $\mathcal{N}(0,1)$ random variable and $\mathcal{D}_0 = \{0,1\}$. Let us define for $m \geq 1$ and $k \in \{0, \dots 2^m - 1\}$, the increment :

$$\Delta_{mk} = B_{(k+1)2^{-m}} - B_{k2^{-m}} = B^{(m-1)}_{(k+1)2^{-m}} - B^{(m-1)}_{k2^{-m}}.$$

Fix $m \geq 1$ and suppose that $(i)_m$ and $(ii)_m$ are true. We are going to prove that the increments $\Delta_{m+1,k}$ are gaussian $\mathcal{N}(0, 2^{-(m+1)})$ and independent. In fact, we are going to prove a little bit more : for a given $m \in \mathbf{N}$, the vector $(\Delta_{mk})_{k \in \{0,\dots,2^m-1\}}$ is gaussian with mean zero and covariance matrix $2^{-(m+1)} I_{2^m}$.

- First, we can suppose without loss of generality (we'll see why later) that $d := k2^{-(m+1)} \notin \mathcal{D}_m$. Then :

  $$\Delta_{m+1,k} = \frac{1}{2}\Delta - 2^{-(m+2)/2} X_{mk'}$$

  where (figure 1)

  $$\frac{k}{2^{m+1}} = \left(k' + \frac{1}{2}\right) 2^{-m} \text{ and } \Delta = B\left(d + 2^{-(m+1)}\right) - B\left(d - 2^{-(m+1)}\right) \sim \mathcal{N}(0, 2^{-m}).$$

  Since $d \pm 2^{-(m+1)} \in \mathcal{D}_{m-1}$, Δ and $X_{mk'}$ are independent and from the induction hypothesis we conclude that Δ_{mk} is gaussian with

  $$\mathbf{E}(\Delta_{m+1,k}) = 0 \text{ and } \operatorname{Var}(\Delta_{m+1,k}) = \frac{1}{4}\operatorname{Var}(\Delta) + \frac{1}{2^{m+2}} = 2^{-(m+1)}.$$

  The same is true if $d \in \mathcal{D}_m$ and it is easy to see that the vector (Δ_{mk}) is gaussian.

- Two successive increments $\Delta_{m+1,k}$ and $\Delta_{m+1,k-1}$ are independent since

  $$\begin{aligned} \operatorname{cov}(\Delta_{m+1,k}, \Delta_{m+1,k-1}) &= \mathbf{E}\left((\frac{1}{2}\Delta - 2^{-(m+2)/2} X_{mk'})(\frac{1}{2}\Delta + 2^{-(m+2)/2} X_{mk'})\right) \\ &= \frac{1}{4}2^{-m} - 2^{-(m+2)} = 0 \end{aligned}$$

  since Δ and $X_{mk'}$ are independent. The same is true if $d \in \mathcal{D}_m$. Two any increments over disjoint intervals are thus independent by summing independent increments over intervals of the form $(k2^{-(m+1)}, (k+1)2^{-(m+1)})$.

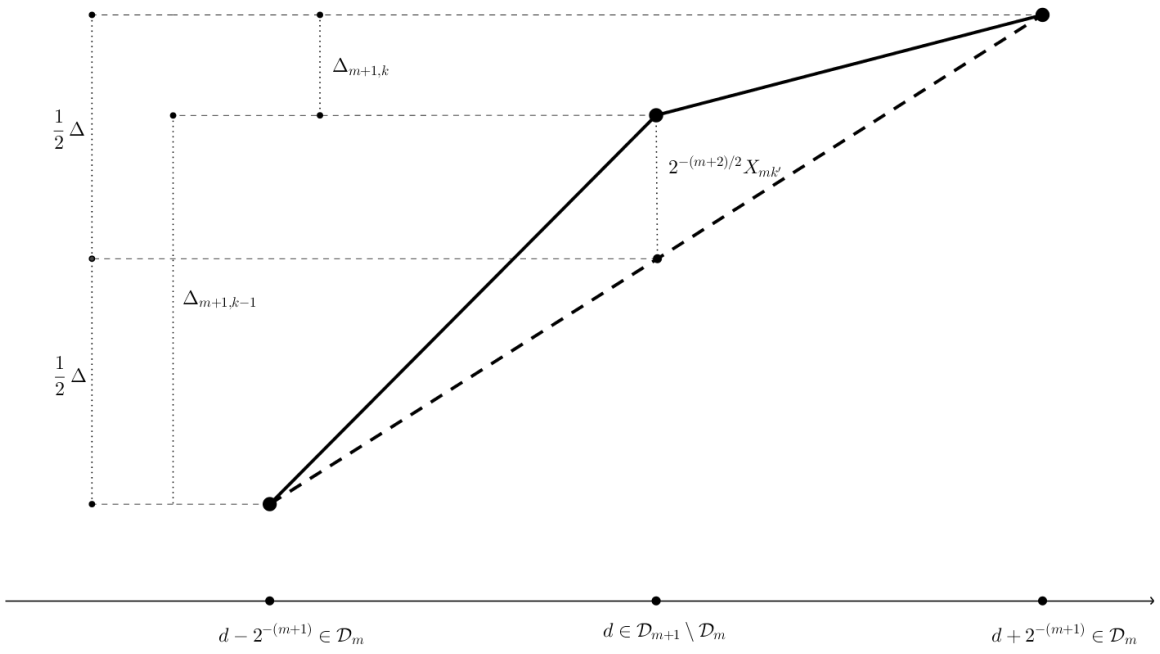


Figure 2.1: Detail around d

Since the set of dyadic numbers is dense in $[0,1]$ the conclusion follows by writing any vector of increments as the limit of gaussian vectors with mean zero and convergent covariance matrices (use the characteristic functions). □

Quelques dessins

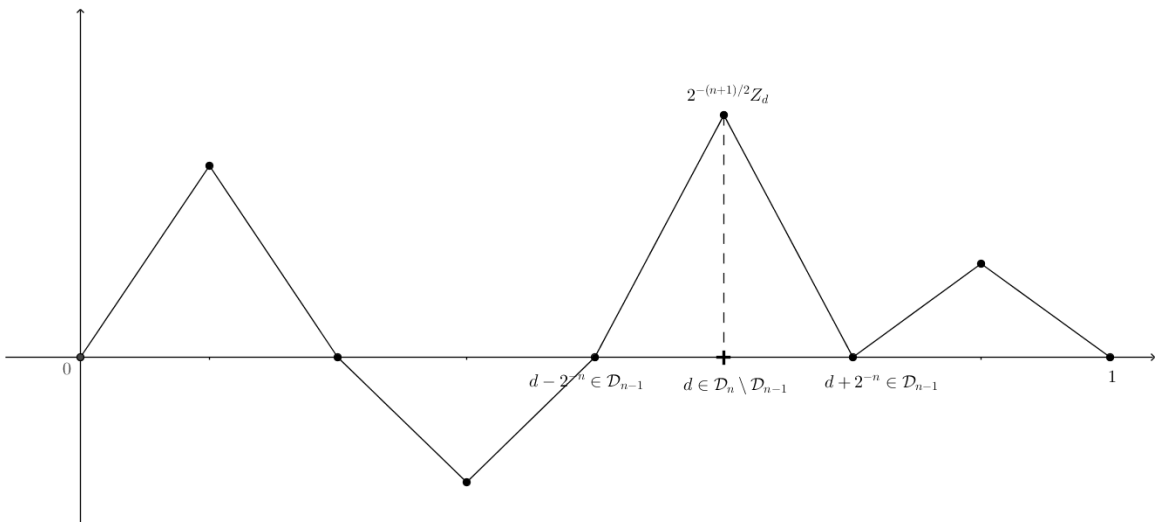


Figure 2.2: Des tentes

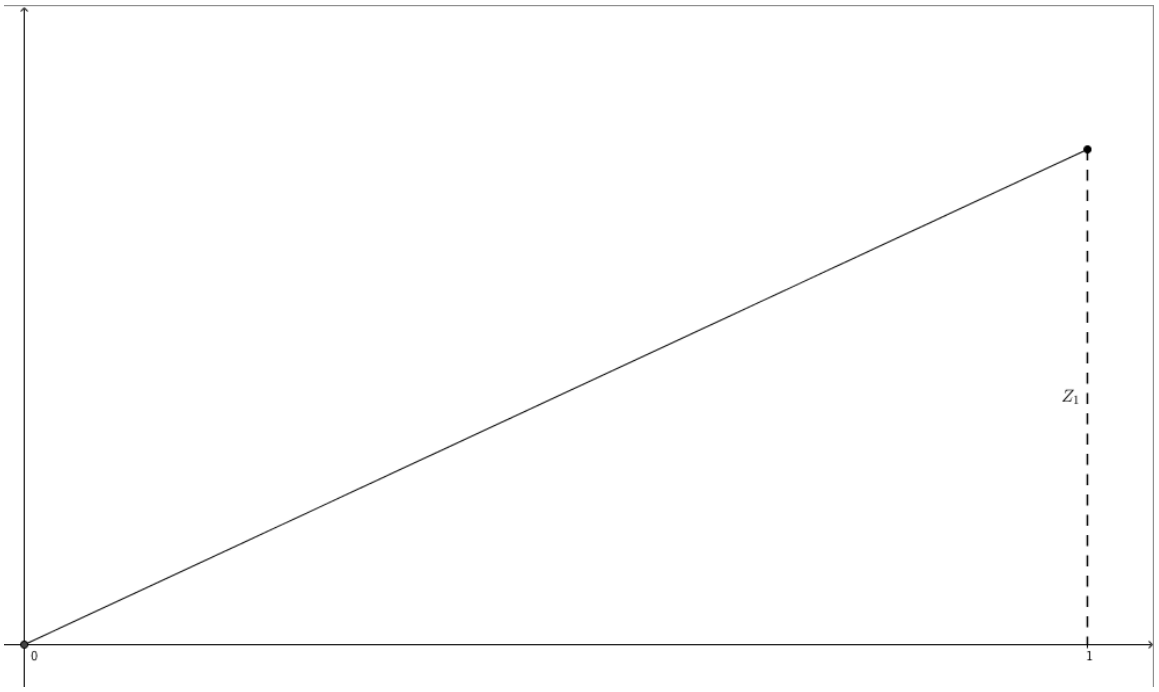


Figure 2.3: Au commencement

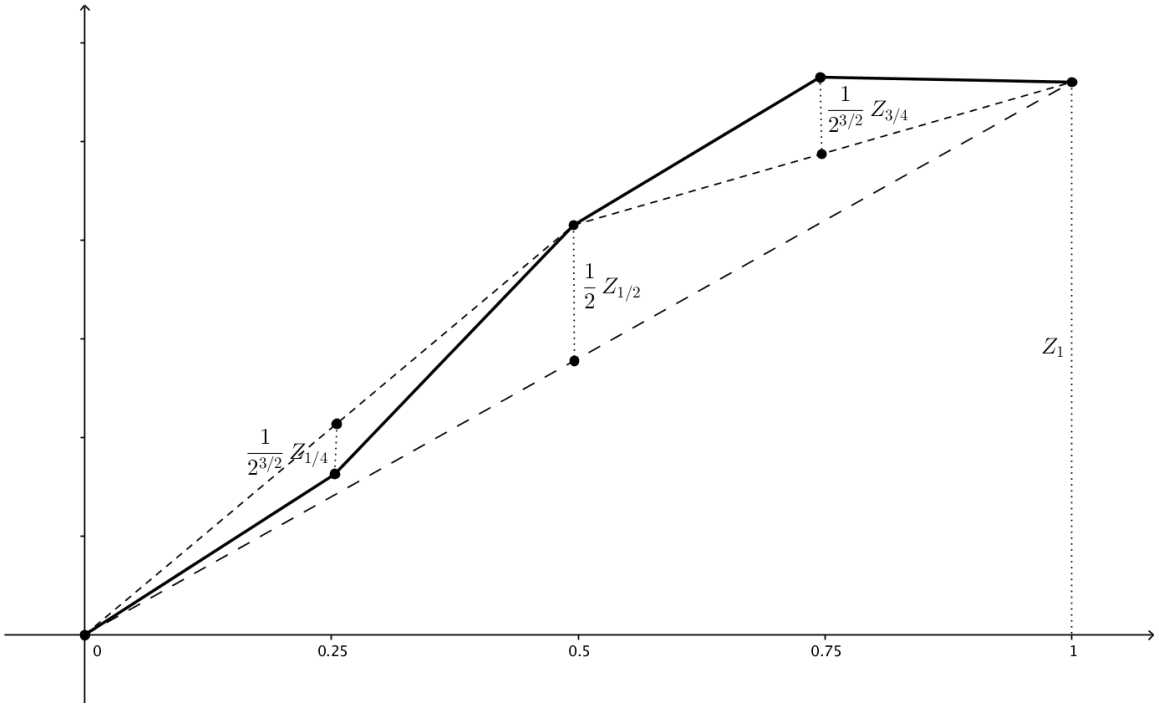


Figure 2.4: Après quelques itérations

Reference. J. B. Walsh, *Knowing the Odds : an Introduction to Probability*

241 Suites et séries de fonctions. Exemples et contre-exemples.
260 Espérance, variance et moments d'une variable aléatoire.
262 Modes de convergence d'une suite de variables aléatoires. Exemples et applications.
263 Variables aléatoires à densité. Exemples et applications.

2.7 La transformée de Bargmann ou comment voir $L^2(\mathbf{R})$ comme un espace de fonctions anaytiques

Quelques pré-requis

Les *polynômes d'Hermite* sont définis sur $\mathbf{R}$ par :

$$H_0 = 1 \quad \text{et} \quad \forall n \in \mathbf{N}^*,\ H_n(x) = \frac{(-1)^n}{2^{n-1/4}\pi^{n/2}\sqrt{n!}} e^{2\pi x^2} \partial^n (e^{-2\pi x^2}).$$

Les *fonctions d'Hermite* sont définies sur $\mathbf{R}$ par :

$$\forall n \in \mathbf{N},\ \ h_n(x) = e^{-\pi x^2} H_n(x).$$

Proposition. *Les fonctions d'Hermite forment une base hilbertienne de $L^2(\mathbf{R})$.*

La transformée de Bargmann et l'espace du même nom

Pour tout $z \in \mathbf{C}$, et $f \in L^2(\mathbf{C})$, on définit :

$$Bf(z) := \sum_{n\geq 0} \frac{z^n}{\sqrt{n!}} \langle h_n, f\rangle.$$

Comme $|\langle h_n, f\rangle| \leq \|f\|$, cela définit une fonction analytique dans tout $\mathbf{C}$ appelée *transformée de Bargmann de f*. Notons que par le théorème de Fubini, on peut écrire plus simplement :

$$\begin{aligned} Bf(z) = \sum_{n\geq 0} \frac{z^n}{\sqrt{n!}} \langle h_n, f\rangle &= \sum_{n\geq 0} \frac{z^n}{\sqrt{n!}} \int_{\mathbf{R}} h_n(x) f(x) dx \\ &= \int_{\mathbf{R}} \left(\sum_{n\geq 0} \frac{z^n}{\sqrt{n!}} h_n(x) \right) f(x) dx \end{aligned}$$

Entre parenthèses, c'est une série entière (avec un paramètre $x \in \mathbf{R}$) dont le terme général vaut :

$$2^{1/4} e^{\pi x^2} \frac{\left(\frac{-z}{2\sqrt{\pi}}\right)^n}{n!} \partial^n (e^{-2\pi x^2}).$$

On reconnait la série de Taylor au point x de la fonction $y \mapsto e^{-2\pi y^2}$ qui est analytique sur tout $\mathbf{C}$. Ainsi :

$$Bf(z) = \int_{\mathbf{R}} 2^{1/4} e^{\pi x^2} e^{-2\pi(x - \frac{z}{2\sqrt{\pi}})^2} f(x) dx = 2^{1/4} \int_{\mathbf{R}} e^{-\pi x^2 + 2xz\sqrt{\pi} - \frac{1}{2}z^2} f(x) dx.$$

Notons que la transformée de Bargmann B est injective car :

$$Bf = 0 \Rightarrow \forall n \in \mathbf{N},\ \langle h_n, f\rangle \Rightarrow f = 0.$$

L'image de $L^2(\mathbf{R})$ par la transformée de Bargmann B est *l'espace de Bargmann*, noté $\mathfrak{Barg}(\mathbf{C})$ dans la suite. C'est un sous-espace de l'espace des fonctions analytiques sur tout $\mathbf{C}$. Bien évidemment :

$$B : L^2(\mathbf{R}) \to \mathfrak{Barg}(\mathbf{C})$$

est un isomorphisme qui induit une structure d'espace de Hilbert sur $\mathfrak{Barg}(\mathbf{C})$ pour laquellle B est continue. On pose pour cela :

$$\langle Bf, Bg\rangle_{\mathfrak{Barg}(\mathbf{C})} := \langle f, g\rangle_{L^2(\mathbf{C})}.$$

Proposition. *L'espace* $\mathfrak{Barg}(\mathbf{C})$ *est l'espace des fonctions analytiques* $g(z) = \sum a_n z^n$ *sur tout* $\mathbf{C}$ *telles que* $\sum_{n\geq 0} n!|a_n|^2 < +\infty$.

Preuve. Si $f \in L^2(\mathbf{R})$ et $Bf(z) = \sum_{n\geq 0} a_n z^n$, alors pour tout $n \in \mathbf{N}$:

$$\langle h_n, f\rangle = \sqrt{n!}\, a_n$$

et par l'égalité de Parseval, on a bien $\sum_{n\geq 0} n!|a_n|^2 < +\infty$.
Réciproquement, si $(a_n)_n$ vérifie cette condition de sommabilité, alors en posant :

$$f = \sum_{n\geq 0} \sqrt{n!}\, a_n h_n \in L^2(\mathbf{R})$$

on a bien $Bf(z) = \sum_{n\geq 0} a_n z^n \in \mathfrak{Barg}(\mathbf{C})$. □

Une application : la transformée de Fourier dans l'espace de Bargmann

On note $\mathscr{F}$ la transformée de Fourier dans $L^2(\mathbf{R})$ que l'on peut transporter par l'isomorphisme B en un opérateur de $\mathfrak{Barg}(\mathbf{C})$ appelé *transformée de Fourier dans l'espace de Bargmann* et noté simplement $\mathfrak{F}$. Il s'agit de faire commuter le diagramme :

$$\begin{array}{ccc} L^2(\mathbf{R}) & \xrightarrow{B} & \mathfrak{Barg}(\mathbf{C}) \\ \mathscr{F}\downarrow & & \downarrow\mathfrak{F} \\ L^2(\mathbf{R}) & \xrightarrow[B]{} & \mathfrak{Barg}(\mathbf{C}) \end{array}$$

Le calcul de $\mathfrak{F}(B\varphi) = B(\mathscr{F}\varphi)$ est facile lorsque $\varphi \in \mathcal{D}(\mathbf{R})$: il suffit d'appliquer le théorème de Fubini et de se souvenir du calcul de

$$\int_{\mathbf{R}} e^{-\alpha x^2 + 2\beta x} dx = \sqrt{\frac{\pi}{\alpha}} e^{\frac{\beta^2}{\alpha}},\ \ (\alpha > 0,\ \beta \in \mathbf{C}).$$

On trouve :

$$\begin{aligned} B(\mathscr{F}\varphi)(z) &= 2^{1/4} \int_{\mathbf{R}} e^{-\pi\xi^2 + 2\xi z\sqrt{\pi} - \frac{1}{2}z^2} \left(\int_{\mathbf{R}} e^{-2i\pi x\xi} \varphi(x) dx \right) d\xi \\ &= 2^{1/4} e^{-\frac{1}{2}z^2} \int_{\mathbf{R}} \left(\int_{\mathbf{R}} e^{-\pi\xi^2 + 2\xi(z\sqrt{\pi} - i\pi x)} d\xi \right) \varphi(x) dx \\ &= 2^{1/4} \int_{\mathbf{R}} e^{-\pi x^2 + 2x(-iz)\sqrt{\pi} - \frac{1}{2}(-iz)^2} \varphi(x) dx = B\varphi(-iz) \end{aligned}$$

Par densité de $\mathcal{D}(\mathbf{R})$ dans $L^2(\mathbf{R})$ et continuité de tous les opérateurs, le résultat subsiste pour tout $f \in L^2(\mathbf{R})$ et on trouve :

$$\forall \mathfrak{g} \in \mathfrak{Barg}(\mathbf{C}),\ \ \mathfrak{F}(\mathfrak{g})(z) = \mathfrak{g}(-iz).$$

En particulier, il devient facile de calculer $\mathscr{F}(h_n)$: dans l'espace de Bargmann, $B(h_n) = \frac{z^n}{\sqrt{n!}}$, d'où :

$$\mathfrak{F}(Bh_n)(z) = (-i)^n \frac{z^n}{\sqrt{n!}} = (-i)^n B(h_n)$$

et donc :

$$B^{-1}\mathfrak{F}B(h_n) = \mathscr{F}(h_n) = (-i)^n h_n$$

ce qui signifie encore que $\mathscr{F}$ est diagonalisée sur la base d'Hermite.

Référence. B. Candelpergher, *Calcul intégral*

201 Espaces de fonctions ; exemples et applications.
213 Espace de Hilbert. Bases hilbertiennes. Exemples et applications.
236 Illustrer par des exemples quelques méthodes de calcul d'intégrales de fonctions d'une ou plusieurs variables. *(mais si)*
243 Convergence des séries entières, propriétés de la somme. Exemples et applications. *(ou pas)*
245 Fonctions holomorphes sur un ouvert de $\mathbf{C}$. Exemples et applications. *(ou pas)*
250 Transformation de Fourier. Applications.

2.8 Un exemple de perturbation d'une système linéaire : le pendule de Van der Pol

Pour $\varepsilon > 0$, l'équation du pendule de van der Pol est :

$$x'' + \varepsilon(x^2 - 1)x' + x = 0$$

que l'on s'empresse d'écrire comme un système du premier ordre :

$$\begin{array}{rcl} x' & = & -y \\ y' & = & x - \varepsilon(x^2-1)y \end{array} \tag{VdP}$$

On le notera plus volontiers :

$$X' = AX + \varepsilon f(X) = F_\varepsilon(X), \quad X = \begin{pmatrix} x \\ y \end{pmatrix}$$

avec

$$A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \quad \text{et} \quad f(X) = \begin{pmatrix} 0 \\ (x^2-1)y \end{pmatrix}.$$

C'est une perturbation de l'équation $x'' + x = 0$ dont la solution $t \mapsto x(t)$ est 2π-périodique et de la forme :

$$t \mapsto x_0 \cos t + \dot{x_0} \sin t.$$

On va s'intéresser aux solution périodique du système perturbé (VdP).

Proposition. *Pour tout $\varepsilon > 0$ et toute condition initiale $X_0 = (x_0, y_0) \in \mathbf{R}^2$, le système non-linéaire* (VdP) *possède une unique solution qui est globale.*

Preuve. Le champ de vecteur est C^1 donc il existe une unique solution maximale, définie sur un intervalle ouvert, on la note $X(t) = (x(t), y(t))$. On définit :

$$E(t) = \frac{1}{2}y^2 + \frac{1}{2}x^2$$

Comme

$$E'(t) = -\varepsilon(x^2 - 1)y^2$$

on peut séparer l'étude en deux cas : d'abord si $|x(t)| \leq 1$ alors $y(t)$ est bornée par le lemme de Gronwall et si $|x(t)| > 1$ alors $E' < 0$ donc E est décroissante et $\|X(t)\|_2$ est bornée. Dans les deux cas, le critère d'explosion en temps fini permet d'affirmer que la solution est globale. □

À partir de maintenant on fixe $\xi > 0$. On notera [4] $\phi_t(\xi; \varepsilon)$ le flot de (VdP) associé à la condition initiale

$$x(0; \varepsilon) := \xi \quad \text{et} \quad y(0; \varepsilon) = 0.$$

Proposition (Poincaré)**.** *Pour $\varepsilon > 0$ suffisamment petit, il existe une fonction C^1, $(\xi, \varepsilon) \mapsto T(\xi, \varepsilon)$ telle que*

$$\phi_{T(\xi,\varepsilon)}(\xi; \varepsilon) \in \mathbf{R}_+ \times \{0\}, \quad T(\xi; 0) = 2\pi \quad \textit{et} \quad \phi_{T(\xi;0)}(\xi; 0) = (\xi, 0).$$

On définit alors la fonction

$$(\xi, \varepsilon) \mapsto P(\xi, \varepsilon) := x(T(\xi, \varepsilon); \varepsilon).$$

Preuve. Soit $\xi > 0$. On va appliquer le théorème des fonctions implicites à la fonction :

$$(t, \varepsilon) \in \mathbf{R} \times \mathbf{R}_+^* \mapsto g(t, \xi, \varepsilon) = (\phi_t(\xi; \varepsilon) - (\xi, 0)) \cdot F_\varepsilon(\xi, 0).$$

En effet, $\phi_t(\xi; \varepsilon) \in \mathbf{R}_+ \times \{0\}$ si et seulement si $\phi_t(\xi; \varepsilon) - (\xi, 0)$ est orthogonal à $F_\varepsilon(\xi, 0)$.

Puisque la solution du système non-perturbé est 2π-périodique :

$$g(2\pi, \xi, 0) = 0 \quad \text{et} \quad \frac{\partial g}{\partial t}(t, \xi, \varepsilon) = F_\varepsilon(\xi, 0) \cdot F_\varepsilon(\xi, 0).$$

De sorte que $\dfrac{\partial g}{\partial t}(2\pi, \xi, 0) = \|F_0(\xi, 0)\|^2 > 0.$

□

4. Noter la subtile utilisation du point-virgule.

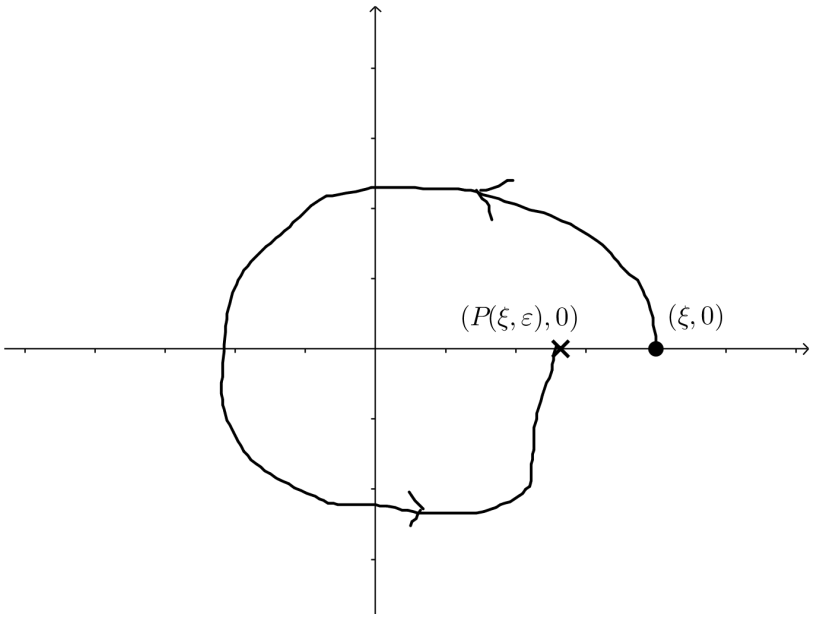


Figure 2.5: Cicéron c'est Poincaré

S'intéresser aux solutions périodiques revient à étudier les points fixes de $\xi \mapsto P(\xi,\varepsilon)$, c'est à dire aux zéros de la fonction

$$\delta(\xi,\varepsilon) := P(\xi,\varepsilon) - \xi \quad \text{qui vérifie} \quad \delta(\xi,0) = 0 \quad \text{pour tout } \xi > 0.$$

Plus précisément, on cherche une courbe $\varepsilon \mapsto \beta(\varepsilon)$ telle que

$$\delta(\beta(\varepsilon),\varepsilon) \equiv 0 \quad \text{pour } \varepsilon > 0 \text{ suffisamment petit.}$$

On ne peut pas appliquer le théorème des fonctions implicites à δ mais comme :

$$\delta(\xi,\varepsilon) = \varepsilon\partial_\varepsilon\delta(\xi,0) + \mathcal{O}(\varepsilon^2)$$

on peut poser

$$\Delta(\xi,\varepsilon) = \frac{1}{\varepsilon}\delta(\xi,\varepsilon)$$

qui se prolonge par continuité en $\varepsilon = 0$ et qui vérifie :

$$\Delta(\beta(\varepsilon),\varepsilon) \equiv 0 \iff \delta(\beta(\varepsilon),\varepsilon) \equiv 0.$$

Pour appliquer le théorème des fonctions implicites à Δ, il faut trouver $\xi > 0$ tel que :

$$\Delta(\xi,0) = 0 \quad \text{et} \quad \partial_\xi\Delta(\xi,0) \neq 0 \quad \text{i.e} \quad \partial^2_{\xi\varepsilon}\delta(\xi,0) \neq 0.$$

On calcule avec la définition de δ :

$$\partial_\varepsilon\delta(\xi,0) = x'(T(\xi,0),0)\partial_\varepsilon T(\xi,0) + \partial_\varepsilon x(T(\xi,0),\varepsilon).$$

Or, $T(\xi,0) = 2\pi$ et $x'(T(\xi,0),0) = -y(0,0) = 0$. Il n'y a donc que le second terme à calculer. Pour ce faire, on dérive par rapport à ε le système (VdP) et on trouve (lemme de Schwarz) :

$$\begin{aligned}(\partial_\varepsilon x)'(t;0) &= \partial_\varepsilon y(t;0)\\ (\partial_\varepsilon y)'(t;0) &= \partial_\varepsilon x(t;0) - (x^2(t;0)-1)y(t;0)\end{aligned}$$

avec les conditions initiales :

$$\partial_\varepsilon x(0;0) = 0 \quad \text{et} \quad \partial_\varepsilon y(0;0) = 0.$$

C'est un système linéaire inhomogène que l'on résout par variation de la constante :

$$\begin{pmatrix} \partial_\varepsilon x(t;0) \\ \partial_\varepsilon y(t;0) \end{pmatrix} = \int_0^t e^{(t-s)A} \begin{pmatrix} 0 \\ (1-x(s;0)^2)y(s;0) \end{pmatrix} ds.$$

En particulier, comme on sait résoudre explicitement le système non perturbé, un calcul avec des sinus et des cosinus montre que :

$$\partial_\varepsilon x(2\pi,0) = \partial_\varepsilon \delta(\xi,0) = \Delta(\xi,0) = \int_0^{2\pi} \sin s[(1-\xi^2\cos^2 s)\xi \sin s]ds = \frac{\pi}{4}\xi(4-\xi^2).$$

En conclusion $\xi = 2$ est le seul zéro de $\Delta(\xi,0)$ et il est simple. Par le théorème des fonctions implicites, il existe une fonction $\varepsilon \mapsto \beta(\varepsilon)$ définie dans un voisinage de $\varepsilon = 0$ telle que $\beta(0) = 2$ et pour tout $\varepsilon > 0$ dans le domaine de définition de β, le système (VdP) a une orbite périodique avec condition initiale $(x(0;\varepsilon), y(0;\varepsilon)) = (\beta(\varepsilon), 0)$.

Il parait que l'on peut aussi calculer un développement asymptotique quand $\varepsilon \to 0$ de la période des orbites périodiques du système (VdP) en développant en série de Taylor :

$$\varepsilon \mapsto T(\beta(\varepsilon), \varepsilon).$$

à partir de l'identité :

$$y(T(\beta(\varepsilon),\varepsilon),\varepsilon) \equiv 0.$$

La suite consisterait à montrer que l'unique trajectoire périodique est un cycle limite pour les autres trajectoires, ce qui résulte directement du théorème de Poincaré-Bendixson (difficile, il y a une preuve dans le Gonnord-Tosel). En voici une illustration :

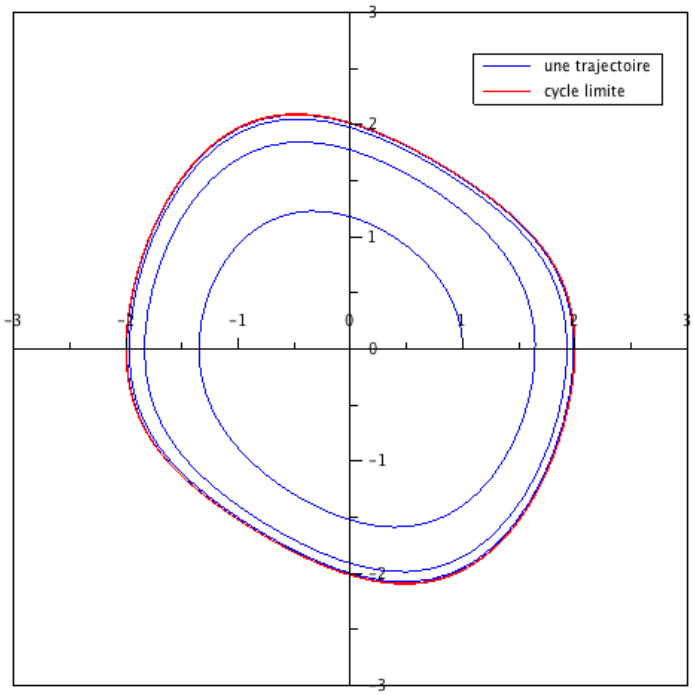


Figure 2.6: Existence d'un cycle limite pour le pendule de Van Der Pol

Réalisé sous l'oeil sévère de Grégoire Clarté.

Référence. C. Chicone, *Ordinary Differential Equations with Applications, 2nd Edition.*

214 Théorème d'inversion locale, théorème des fonctions implicites. Exemples et applications en analyse et en géométrie.
220 Équations différentielles $X' = f(t, X)$. Exemples d'étude des solutions en dimension 1 et 2.
221 Équations différentielles linéaires. Système d'équations différentielles linéaires. Exemples et applications.

2.9 Le problème des moments de Hamburger

Théorème. *Soit $(\alpha_k)_k$ une suite de réels telle que la série entière $\sum_{k\geq 0} \frac{\alpha_k}{k!} r^k$ ait un rayon de convergence strictement positif. Alors il existe au plus une mesure de probabilité μ sur $\mathbf{R}$ dont les moments sont les α_k :*

$$\forall k \in \mathbf{N},\ \ \alpha_k = \int x^k d\mu(x).$$

Preuve. Supposons l'existence d'une telle mesure μ et notons φ sa fonction caractéristique. La fonction caractéristique a au moins deux vertus : elle caractérise la loi et les moments sont donnés par ses dérivées en 0. La preuve repose sur ces deux idées : on va montrer que sous la condition énoncée, la fonction caractéristique de μ est analytique, ainsi le théorème de prolongement analytique montrera que deux mesures ayant les mêmes moments auront la même fonction caractéristique, celles-ci coïncidant autour de zéro.

Étape 1. Précisons le développement de Taylor de φ.

Pour $x \in \mathbf{R}$, $t, h \in \mathbf{R}$ la formule de Taylor avec reste intégral pour $s \mapsto e^{isx}$ entre 0 et h donne :

$$e^{i(t+h)x} = e^{itx}\left(\sum_{m=0}^{n} \frac{(ix)^m}{m!} + \int_0^h \frac{(h-s)^n}{n!}(ix)^n e^{isx} ds\right)$$

Ce qui en intégrant sur $x \in \mathbf{R}$ selon μ donne :

$$\varphi(t+h) - \sum_{m=0}^{n} \frac{h^m}{m!}\varphi^{(m)}(t) = \int_{\mathbf{R}} \int_0^h \frac{(h-s)^n}{n!}(ix)^n e^{isx} ds\, d\mu(x)$$

En prenant les modules de chaque côté, on trouve finalement :

$$\left|\varphi(t+h) - \sum_{m=0}^{n} \frac{h^m}{m!}\varphi^{(m)}(t)\right| \leq \frac{|h|^{n+1}}{(n+1)!}\beta_{n+1} \tag{2.6}$$

où on a noté :

$$\forall k \in \mathbf{N},\ \ \beta_k = \int |x|^k d\mu(x).$$

Étape 2. Quid des β_k ?

La majoration (2.6) montre que φ est analytique en $t \in \mathbf{R}$ pour peu que l'on trouve $r > 0$ tel que

$$\forall |h| < r,\ \ \frac{|h|^n}{n!}\beta_n \underset{n\to+\infty}{\longrightarrow} 0.$$

Bien sûr, on va utiliser l'existence de $0 < s < 1$ tel que $\alpha_n s^n/n! \to 0$. On remarque d'abord que $\beta_{2n} = \alpha_{2n}$ de sorte que seul les termes avec un indice impair sont à contrôler. On a par l'inégalité de Cauchy-Schwarz et l'inégalité arithmético-géométrique :

$$\beta_{2n+1} \leq \sqrt{\alpha_{2n}\alpha_{2n+2}} \leq \frac{1}{2}\left(\alpha_{2n} + \alpha_{2n+2}\right).$$

Ainsi il suffit de prendre $r > 0$ tel que $(2n+2)r^{2n+1} < s^{2n+2}$ pour tout $n \in \mathbf{N}$ pour avoir :

$$\frac{r^{2n+1}}{(2n+1)!}\beta_{2n+1} \underset{n\to+\infty}{\longrightarrow} 0.$$

Étape 3. Un prolongement analytique et c'est fini.

Supposons que ν soit une mesure de probabilité sur $\mathbf{R}$ admettant les α_k pour moments. Alors, on vient de montrer que les fonctions caractéristiques de μ et de ν coïncident sur l'ouvert $B(0,r)$ donc partout puisqu'elle sont analytiques [5]. Finalement, $\mu = \nu$.

Remarquons pour terminer qu'on a utilisé que les moments d'ordre pair.

□

Les moments caractérisent parfois la loi.

- La loi normale $\mathcal{N}(0,1)$ est caractérisée par ses moments :

  $$\alpha_{2k} = \frac{(2k)!}{2^k k!}$$

  et

  $$\frac{\alpha_{2k}^{1/2k}}{2k} \sim \frac{C}{2k} \times \frac{k}{\sqrt{k}} \sim \frac{C}{\sqrt{k}}.$$

- La loi log-normale n'est pas caractérisée par ses moments. Elle est définie par la densité :

  $$f_0(x) = (2\pi)^{-1/2} x^{-1} \exp(-(\log x)^2/2), \;\; x \geq 0.$$

  Et pour $-1 \leq a \leq 1$, la densité :

  $$f_a(x) = f_0(x)\big(1 + a\sin(2\pi \log x)\big)$$

  a les mêmes moments que f_0. Pour le voir, on va montrer que

  $$\forall r \in \mathbf{N}, \;\; \int_0^{+\infty} x^r f_0(x) \sin(2\pi \log x) dx = 0.$$

  Il suffit d'écrire le changement de variable $x = \exp(s+r)$:

  $$(2\pi)^{-1/2} \int_{-\infty}^{+\infty} \exp(rs + r^2) \exp(-(s+r)^2/2) \sin(2\pi(s+r)) ds$$
  $$= (2\pi)^{-1/2} \exp(r^2/2) \int_{-\infty}^{+\infty} \exp(-s^2/2) \sin(2\pi s) ds = 0$$

  par imparité de la dernière intégrande.

- En fait, les mesures caractérisées par leurs moments sont celles qui décroissent vite à l'infini (voir « tension de mesures »et théorème de Prokhorov dans l'annexe). On peut déjà voir que si le support de la mesure est contenu dans dans un segment $[-M, M]$, en approchant F uniformément par des polynômes, le résultat est vrai. (?)

5. C'est hyper simple de l'écrire ici parce que le $r > 0$ ne dépend pas du t : on peut arguer que $\mathbf{R}$ est archimédien...

La méthode des moments.

Pour montrer qu'une suite de variables aléatoires à valeurs réelles convergent en loi on peut utiliser le critère suivant quand on sait que la limite est caractérisée par ses moments :

Théorème (Méthode des moments)**.** *Soit X une variable aléatoire réelle caractérisée par ses moments et $(X_n)_n$ une suite de variables aléatoires telles que*

$$\forall r \in \mathbf{N},\ \mathbf{E}(X_n^r) \underset{n\to+\infty}{\longrightarrow} \mathbf{E}(X^r).$$

Alors $X_n \Rightarrow X$.

Preuve. On note classiquement μ_n la loi de X_n et μ la loi de X. Alors comme la suite $\left(\mathbf{E}(X_n^2)\right)_n$ converge, elle est majorée, disons par $K > 0$ et l'inégalité de Markov montre que :

$$\mathbf{P}(|X_n| \geq x) \leq \frac{K}{x^2}$$

de sorte que la suite $(\mu_n)_n$ est tendue. Pour montrer qu'elle converge, il suffit donc de montrer qu'elle a une unique valeur d'adhérence. Supposons donc que $\mu_{n_k} \Rightarrow \nu$ et considérons Y une variable aléatoire réelle de loi ν. Il s'agit d'un problème d'uniforme intégrabilité : pour tout $r > 0$ on a comme tout à l'heure pour un certain $K > 0$:

$$K \geq \int X_n^{2r} d\mathbf{P} \geq \alpha \int_{|X_n|^r \geq \alpha} |X_n|^r d\mathbf{P}$$

de sorte que la suite $(X_n^r)_n$ est uniformément intégrable et comme elle converge en loi vers Y^r (facile à voir avec la caractérisation par l'intégrale sur une fonction continue bornée) le théorème de Vitali (mais s'appelle-t-il bien Vitali ?) montre que

$$\mathbf{E}(X_n^r) \to \mathbf{E}(Y^r).$$

Ainsi X et Y ont les mêmes moments et ont donc la même loi : $\nu = \mu$ □

Quelques remarques complémentaires.

- La condition que l'on donne n'est pas vraiment optimale. Une condition légèrement plus forte est connue sous le nom de condition de Carleman (peut être qu'on peut raffiner la preuve pour la retrouver) :
$$\sum_{k=1}^{+\infty} \frac{1}{\alpha_{2k}^{1/2k}} = +\infty.$$
Cela étant, on peut montrer que cette condition est suffisante mais pas nécéssaire. Il existe aussi des conditions nécessaires mais pas suffisantes.
- En fait, le problème des moments n'est pas inhérent aux probabilités. Dans le cas général (pour une mesure de Borel), une condition nécéssaire et suffisante est donnée par l'analyse fonctionnelle (voir [Lax, *Functional Analysis*]) :
$$\sum_{n,k} \alpha_{n+k}\xi_n\xi_k \geq 0$$
pour toute famille finie de de réels $(\xi_n)_{0\leq n\leq N}$. Il semblerait que des applications existent en physique quantique.
- Il existe d'autres problèmes de moments et en particulier le problème de Stieltjes lorsqu'on se restreint à des mesures dont le support est contenu dans $[0, +\infty)$. Il y a des liens entre les différents problèmes.
- P. Billingsley donne de jolies applications de la méthode des moments, pour prouver le théorème central limite ou en théorie des nombres notamment.
- Quand la mesure μ est à densité, il y a plus simple pour vérifier l'analyticité de φ : c'est le théorème de Paley et Wiener (coucou cher jury).

Annexe : Sur la converge en loi.

Théorème (Skorohod)**.** *Si F_n converge en loi vers F, alors il existe des variables aléatoires Y, Y_n associées aux fonctions de répartition respectives F et F_n telles que $Y_n \to Y$ presque sûrement.*

Preuve. On considère $\Omega = (0,1)$, $\mathcal{F}$ les boréliens de Ω et $\mathbf{P}$ la mesure de Lebesgue sur Ω. On définit :

$$Y_n(x) = \sup\{y, F_n(y) < x\} \quad \text{et} \quad Y(x) = \sup\{y,\ F(y) < x\}.$$

Ces variables aléatoires Y, Y_n admettent F et F_n pour fonctions de répartition respectives (c'est presque une tautologie mais il faudrait y réfléchir). Pour conclure, on se convainc que :

$$\liminf_{n\to+\infty} Y_n(x) \geq Y(x) \quad \text{et} \quad \limsup_{n\to+\infty} Y_n(x) \leq Y(x).$$

□

Théorème (Helly)**.** *De toute suite de fonctions de répartition $(F_n)_n$, on peut extraire une sous-suite qui converge simplement vers une fonction F, continue à droite et croissante, en tout point de continuité de F. La limite F n'est pas nécessairement une fonction de répartition.*

Preuve. L'argument principal est la séparabilité de $\mathbf{R}$ couplé à une extraction diagonale. Plus précisément, puisque $\mathbf{Q}$ est dénombrable, par le théorème de Bolzano-Weierstrass et un argument d'extraction diagonale, il existe une extractrice $n(k)$ telle que :

$$\forall q \in \mathbf{Q},\ \ F_{n(k)}(q) \underset{k\to+\infty}{\longrightarrow} G(q)$$

où G est une fonction définie a priori sur $\mathbf{Q}$. On pose :

$$F(x) := \inf\{G(q),\ q \in \mathbf{Q},\, q > x\}.$$

La fonction F ainsi définie et continue à droite et croissante. Pour terminer, soit x un point de continuité de F, on considère des rationnels $r_1 < r_2 < s$ tels que

$$F(x) - \varepsilon < F(r_1) \leq F(r_2) \leq F(x) \leq F(s) < F(x) + \varepsilon.$$

Pour k assez grand, on vérifie alors :

$$F(x) - \varepsilon < F_{n(k)}(r_2) \leq F_{n(k)}(x) \leq F_{n(k)}(s) < F(x) + \varepsilon.$$

□

Théorème (Prokhorov)**.** *Si les mesures μ_n associées aux F_n vérifient la condition*

$$\forall \varepsilon > 0,\ \ \exists M_\varepsilon > 0,\ \ \limsup_{n\to+\infty} \big(1 - \mu_n([-M_\varepsilon, M_\varepsilon])\big) \leq \varepsilon \tag{2.7}$$

alors toute limite simple en tout point de continuité d'une sous-suite de $(F_n)_n$ est une fonction de répartition. Remarquons aussi que :

$$\mu_n([-M_\varepsilon, M_\varepsilon]) = F_n(M_\varepsilon) - F(-M_\varepsilon).$$

Cela montre que de toute suite de mesure tendue on peut extraire une sous-suite convergente (pour la convergence en loi).

Preuve. Supposons (2.7) et considérons une sous-suite $(F_{n(k)})_k$ qui converge simplement vers une fonction F continue à droite et croissante en tout point de continuité de F. Pour que F soit une fonction de répartition, il suffit de montrer que $\mu(\mathbf{R}) = 1$. Il suffit de voir que pour $r < -M_\varepsilon$ et $s > M_\varepsilon$ des points de continuité de F :

$$\begin{aligned} 1 - F(s) + F(r) &= \lim_{k\to+\infty} 1 - F_{n(k)}(s) + F_{n(k)}(r) \\ &\leq \limsup_{k\to+\infty} 1 - F_n(M_\varepsilon) + F(M_\varepsilon) \leq \varepsilon \end{aligned}$$

et donc pour tout $\varepsilon > 0$, $\limsup_{x\to+\infty} 1 - F(x) + F(-x) \leq \varepsilon$. □

Corollaire. *Soit $(\mu_n)_n$ une suite tendue de mesures de probabilité sur $\mathbf{R}$ qui admet une unique valeur d'adhérence. Alors la suite converge vers cette valeur d'adhérence.*

Théorème (Vitali)**.** *Si $X_n \Rightarrow X$ et si les $(X_n)_n$ sont uniformément intégrables au sens où*

$$\lim_{\alpha\to+\infty} \sup_n \int_{|X_n|\geq\alpha} |X_n| d\mathbf{P} = 0$$

alors

$$\mathbf{E}(X_n) \to \mathbf{E}(X).$$

Preuve. Par le théorème de Skorohod, on peut considérer $\tilde{X}_n \to \tilde{X}$ presque sûrement et il suffit de considérer

$$\tilde{X}_n^{(\alpha)} = \tilde{X}_n \mathbf{1}_{|\tilde{X}_n|\leq\alpha} \quad \text{et} \quad \tilde{X}^{(\alpha)} = \tilde{X}\mathbf{1}_{|\tilde{X}|\leq\alpha}$$

pour conclure en utilisant notamment le théorème de convergence dominée [6]. □

Références.
P. Billingsley, *Probability and Measure*
R. Durrett, *Probability : Theory and Examples*
B. Candelpergher, *Théorie des probabilités, un introduction élémentaire*

241 Suites et séries de fonctions. Exemples et contre-exemples.
260 Espérance, variance et moments d'une variable aléatoire.
263 Variables aléatoires à densité. Exemples et applications.

2.10 Le processus de Galton-Watson

Version périmée même pas terminée. En fait c'est mieux dans le livre de J. Walsh, Knowing the Odds et ENCORE MIEUX chez Grégoire Clarté.

Soient ξ_i^n, $i, n \geq 0$ des variables aléatoires i.i.d. à valeurs dans $\mathbf{N}$. On note :

$$p_k = \mathbf{P}(\xi_i^k = k) \quad \text{et} \quad \mu = \mathbf{E}(\xi_i^m) \in (0, +\infty).$$

On définit :

$$Z_0 = 1 \quad \text{et} \quad \forall n \in \mathbf{N}, \quad Z_{n+1} = \begin{cases} \xi_1^{n+1} + \ldots + \xi_{Z_n}^{n+1} & \text{si } Z_n > 0 \\ 0 & \text{si } Z_n = 0 \end{cases}$$

6. le théorème de Vitali en est une généralisation, il dit en fait un peu plus que ça

Théorème. *Si $\mu > 1$, alors $\mathbf{P}(Z_n > 0, \text{ pour tout } n) > 0$.*

Preuve. Pour $s \in [0,1]$, on définit la fonction génératrice :

$$\varphi(s) = \sum_{k\geq 0} p_k s^k.$$

La preuve du théorème (et même un peu plus) découle de l'étude de φ et des trois lemmes suivants :

Lemme 11. *Si $\theta_m = \mathbf{P}(Z_m = 0)$, alors $\theta_m = \sum_{k=0}^{+\infty} p_k \theta_{m-1}^k = \varphi(\theta_{m-1})$.*

Preuve (lemme 15). □

Lemme 12. *La fonction génératrice est C^1 sur $[0,1]$, croissante, convexe et $\varphi'(1) = \mu$.*

Preuve (lemme 12). La série entière $\sum_{k\geq 0} p_k s^k$ a un rayon de convergence égal à 1, on peut donc différentier à l'intérieur du disque ouvert de convergence. Pour tout $0 \leq s < 1$:

$$\varphi'(s) = \sum_{k=1}^{+\infty} k p_k s^{k-1} \geq 0.$$

De plus, la série de fonctions $\sum_{k\geq 1} k p_k s^{k-1}$ converge normalement donc uniformément sur $[0,1]$ (en norme infinie et parce que $\mu < +\infty$). Par un théorème de dérivation sous l'intégrale, on a la régularité et la croissance de demandée avec en prime $\varphi'(1) = \lim_{s\uparrow 1} \varphi'(s) = \mu$. Pour la convexité, on dérive une deuxième fois pour $0 \leq s < 1$:

$$\varphi''(s) = \sum_{k=2}^{+\infty} k(k-1) p_k s^{k-2} \geq 0.$$

Dans l'intérieur, c'est gagné et par continuité et croissance, c'est aussi bon aux bords. □

Lemme 13. *Si $\mu > 1$, il existe un unique $\rho < 1$ tel que $\varphi(\rho) = \rho$.*

Preuve (lemme 13). D'abord, $\varphi(0) \geq 0$, $\varphi(1) = 1$ et $\varphi'(1) = \mu > 1$ donc $\varphi(1-\varepsilon) < 1-\varepsilon$ pour $\varepsilon > 0$ assez petit (écrire le développement de Taylor). En conséquence, $\varphi - id$ est continue, positive en 0 et négative en $1-\varepsilon$: le théorème de Rolle assure l'existence d'un point fixe de φ dans $(0,1)$.
Maintenant, si on suppose $\mu > 1$, il existe $k > 1$ tel que $p_k > 0$ et $\varphi'' > 0$ donc φ est strictement convexe dans $(0,1)$. Donc si $\rho < 1$ est un point fixe de de φ, $\varphi(x) < x$ pour $x \in (\rho, 1)$. D'où l'unicité de ρ. □

Lemme 14. *Lorsque $m \uparrow \infty$ et en supposant $\mu > 1$, $\theta_m \uparrow \rho$.*

Preuve (lemme 14). Comme $\theta_0 = 0$, $\varphi(\rho) = \rho$ et φ est croissante donc $(\theta_m)_m$ est croissante et $\theta_m \leq \rho$. La suite converge vers $\theta_\infty \leq \rho$ et donc $\theta_\infty = \rho$. □

Remarquons que si $\mu \leq 1$, alors 1 est l'unique point fixe de φ et $\theta_m \to 1$. □

Référence. R. Durrett, *Probability : Theory and Examples.*

223 Suites numériques. Convergence, valeurs d'adhérence. Exemples et applications.
229 Fonctions monotones. Fonctions convexes. Exemples et applications.
243 Convergence des séries entières, propriétés de la somme. Exemples et applications.
260 Espérance, variance et moments d'une variable aléatoire.
264 Variables aléatoire discrètes. Exemples et applications.

2.11 Le théorème de Morgenstern et les fonctions lisses analytiques nulle part

Théorème (Morgenstern)**.** *Il existe un G_δ dense de fonctions de $C^\infty([0,1],\mathbf{R})$ analytiques nulle part pour la topologie de la distance :*

$$\forall f,g\in C^\infty([0,1],\mathbf{R}),\ \ d(f,g)=\sum_{n\in\mathbf{N}}\min\{2^{-n},\ \|(f-g)^{(n)}\|_\infty\}$$

qui fait de $C^\infty([0,1],\mathbf{R})$ un espace métrique complet.

Preuve. Pour $a\in[0,1]\cap\mathbf{Q}$ et $n\in\mathbf{N}$, on définit :

$$T(a,n)=\{f\in C^\infty([0,1],\mathbf{R}),\ \forall k\in\mathbf{N},\ |f^{(k)}(a)|\leq k!n^k\}$$

et on montre que c'est un fermé d'intérieur vide.

- Soit $(f_k)_k$ une suite de fonctions de $T(a,n)$ qui converge vers f pour la topologie de d. En particulier, pour tout $i\in\mathbf{N}$:

  $$\|f_k^{(i)}-f^{(i)}\|_\infty \underset{k\to+\infty}{\longrightarrow} 0$$

  ce qui entraîne la convergence simple de toutes les dérivées en a, de sorte que :

  $$\forall i\in\mathbf{N},\ \ |f^{(i)}(a)|\leq i!n^i \quad\text{et}\quad f\in T(a,n).$$

- Soient $f\in T(a,n)$ et $\varepsilon>0$. Pour $A>0$ et $b>0$m=, on définit :

  $$f_{A,b}(x)=A\cos(b(x-a)).$$

  On va montrer que pour des valeurs intelligentes de A et b, la fonction

  $$s(x)=f(x)+\frac{\varepsilon}{2}f_{A,b}(x)$$

  est à distance $<\varepsilon$ de f mais n'appartient pas à $T(a,n)$. D'abord, on remarque que pour tout $k\in\mathbf{N}$:

  $$\|f_{A,b}^{(k)}\|_\infty=Ab^k \quad\text{et}\quad f_{A,b}^{(2k)}(a)=(-1)^kAb^{2k}.$$

  Puis on choisit :

  $$k\in\mathbf{N}\ \text{ tel que }\ \sum_{i\geq k}2^{-i}\leq\frac{\varepsilon}{2}\quad\text{et}\quad b>2\ \text{ tel que }\ \varepsilon b^k\geq 4(2k)!n^{2k}\quad\text{et}\quad A=b^{-k}.$$

  De sorte que :

  $$d(s,f)\leq\frac{\varepsilon}{2}\sum_{i=0}^{k-1}2^{i-k}+\sum_{i\geq k}2^{-i}<\varepsilon$$

et
$$|f^{(2k)}(a) - s^{(2k)}(a)| = \frac{\varepsilon}{2}|f^{(2k)}_{b^{-k},b}(a)| = \frac{\varepsilon}{2}b^{-k}b^{2k} > 2(2k)!n^{2k}$$
ce qui prouve que $s \notin T(a,n)$ puisque $|f^{(2k)}(a)| < (2k)!n^{2k}$. Donc $T(a,n)$ est d'intérieur vide.

Maintenant, le théorème de Baire affirme que :
$$\bigcup_{a\in[0,1]\cap\mathbf{Q}} \bigcup_{n\in\mathbf{N}^*} T(a,n)$$
est d'intérieur vide. Son complémentaire est donc un G_δ dense et les fonctions qu'il contient ne sont analytiques nulle part. En effet, si f est analytique quelque part, disons en $a \in]0,1[$, alors f est analytique dans un voisinage de a donc par densité de $\mathbf{Q}$ dans $\mathbf{R}$ on peut déjà supposer sans peine que $a \in]0,1[\cap\mathbf{Q}$. De plus le critère d'Hadamard impose :
$$\sup_{k\in\mathbf{N}^*}\left\{\left(|f^{(k)}(a)|/k!\right)^{1/k}\right\} < +\infty$$
de sorte que $f \in T(a,n)$ pour n assez grand. □

Exemple. La fonction :
$$f(x) = \sum_{n=1}^{+\infty} \frac{\cos(n!x)}{(n!)^n}$$
est $C^\infty(\mathbf{R})$, 2π-périodique mais analytique nulle part.

- f n'est pas analytique en $x = 0$, car pour $k \in \mathbf{N}$:
$$|f^{(2k)}(0)| = \sum_{n\geq 0}(n!)^{2k-n} \geq (k!)^k$$
En particulier,
$$\left(|f^{(2k)}(0)|/(2k)!\right)^{1/2k} \geq \frac{\sqrt{k!}}{(2k)!^{1/2k}} \sim c\frac{\sqrt{k!}}{(2k)^{3/2}} \to +\infty.$$
- Comme pour tout $n, m \in \mathbf{Z}$ avec $m \neq 0$, on a :
$$x \mapsto f\left(x + 2\pi\frac{n}{m}\right) - f(x) = \sum_{k=0}^{m} \frac{\cos(k!x)}{k!^k}$$
est analytique sur $\mathbf{R}$ et en particulier en $x = 0$, f n'est analytique en aucun point de la forme $x = 2\pi n/m$. Par densité de $\mathbf{Q}$ dans $\mathbf{R}$, f n'est analytique nulle part.

Références.
M. Zavidovique, *Un Max de Maths.*
F. John, *Partial Differential Equations, 4th Edition.*

201 Espaces de fonctions ; exemples et applications.
205 Espaces complets. Exemples et applications.
228 Continuité et dérivabilité des fonctions réelles d'une variable réelles. Exemples et applications.
243 Convergence des séries entières, propriétés de la somme. Exemples et applications.

2.12 Le théorème de Müntz-Szász

Théorème (Müntz-Szász)**.** *Soit $(\lambda_j)_{j\geq 1}$ une suite de réels strictement positifs qui tend vers $+\infty$. Il y a équivalence entre :*

(i) L'espace vectoriel engendré par la famille $X = \{1, t^{\lambda_1}, t^{\lambda_2}, \ldots\}$ est dense dans $\mathscr{C} := C([0,1], \mathbf{C})$ pour la norme uniforme.

(ii) $\displaystyle\sum \frac{1}{\lambda_j} = +\infty$

où on note un peu abusivement t^{λ_j} la fonction $t \mapsto t^{\lambda_j}$.

Preuve. La preuve consiste à relier l'étude des zéros de certaines fonctions holomorphes au critère de densité suivant : « $\operatorname{Vect} X$ *est dense dans $\mathscr{C}$ si et seulement si toute forme linéaire continue sur $\mathscr{C}$ qui s'annule sur X est nulle sur $\mathscr{C}$* », lequel est une conséquence du théorème de Hahn-Banach. On notera $\lambda_0 = 1$

(**ii**) $\Rightarrow$ (**i**). Soit ℓ une forme linéaire qui s'annule sur X.

(1) Pour $\zeta \in \mathbf{C}$, $\operatorname{Re}\zeta > 0$, on considère la fonction :

$$f(\zeta) = \ell(t^\zeta)$$

qui est holomorphe sur $\{\operatorname{Re} z > 0\}$ car pour $h \in \mathbf{C}$:

$$\frac{f(\zeta+h) - f(\zeta)}{h} - \ell\big(\log(t)t^\zeta\big) = \ell\left(\frac{t^{\zeta+h} - t^\zeta}{h} - \log(t)t^\zeta\right) \to 0.$$

(Écrire le développement de Taylor de $\zeta \to t^\zeta \in \mathbf{C}$ et prendre le sup en t.) De plus, comme $\|t^\zeta\|_\infty \leq 1$ et que ℓ est continue, f est aussi bornée sur $\{\operatorname{Re}\zeta > 0\}$. Par hypothèse,

$$\forall j \geq 1,\ \ f(\lambda_j) = 0.$$

(2) Pour $N \geq 1$, on définit la *produit de Blaschke* $B_N(\zeta)$:

$$B_N(\zeta) := \prod_{j=1}^{N} \frac{\zeta - \lambda_j}{\zeta + \lambda_j}.$$

Le produit B_N vérifie les propriétés suivantes :

(a) $B_N(\lambda_j) = 0$ pour $j = 1, \ldots, N$ et $B_N(\zeta) \neq 0$ si $\zeta \neq \lambda_j$.

(b) $|B_N(\zeta)| \to 1$ lorsque $\operatorname{Re}\zeta \to 0$.

(c) $|B_N(\zeta)| \to 1$ lorsque $|\zeta| \to +\infty$.

De plus, puisque les zéros de B_N sont simples, la fonction :

$$g_N(\zeta) := \frac{f(\zeta)}{B_N(\zeta)}$$

est bien définie et holomorphe sur $\{\operatorname{Re}\zeta > 0\}$.

(3) On prétend que g_N est bornée sur $\{\mathrm{Re}\,\zeta > 0\}$. D'abord f est bornée et on peut supposer sans perte de généralité que $|f(\zeta)| \leq 1$. En utilisant les propriétés (b) et (c), pour tout $\varepsilon > 0$, il existe $\delta > 0$ tel que

$$\mathrm{Re}\,\zeta = \delta \Rightarrow |g_N(\zeta)| \leq 1 + \varepsilon \quad \text{et} \quad |\zeta| = \delta^{-1} \Rightarrow |g_N(\zeta)| \leq 1 + \varepsilon.$$

Par le principe du maximum, la borne $|g_N(\zeta)| \leq 1 + \varepsilon$ est valable sur tout le domaine $\{\mathrm{Re}\,\zeta \geq \delta\} \cap \{|\zeta| \leq \delta^{-1}\}$. En laissant tendre $\varepsilon, \delta \to 0$, on obtient $|g_N(\zeta)| \leq 1$ sur tout $\{\mathrm{Re}\,\zeta > 0\}$.

(4) On a finalement prouvé que :

$$|f(\zeta)| \prod_{j=1}^{N} \left| \frac{\lambda_j + \zeta}{\lambda_j - \zeta} \right| = |f(\zeta)| \prod_{j=1}^{N} \left| 1 + \frac{2\zeta}{\lambda_j - \zeta} \right| \leq 1.$$

Et tout est en place pour conclure la première étape de la preuve : le théorème de sommation des relations de comparaison indique que si la série $\sum 1/\lambda_j$ diverge, alors le produit infini diverge et ce, pour tout $\zeta \in \mathbf{R}$ distinct des λ_j. La borne uniforme impose alors $\ell(t^\zeta) = f(\zeta) = 0$ pour tout $\zeta \in \mathbf{R}$. En particulier $\ell(t^k) = 0$ pour tout $k \in \mathbf{N}$ et on déduit du théorème de Weierstrass que $\ell \equiv 0$.

(**i**) $\Rightarrow$ (**ii**). Par contraposée, on suppose la convergence de la série $\sum 1/\lambda_j < +\infty$ et on construit une forme linéaire continue non nulle sur $\mathscr{C}$ mais nulle sur X.

(1) En s'inspirant de la construction précédente mais en bricolant un peu plus, on considère la fonction :

$$f(z) = \frac{z}{(2+z)^3} \prod_{j=1}^{+\infty} \frac{\lambda_j - z}{2 + \lambda_j + z} = \frac{z}{(2+z)^3} \prod_{j=1}^{+\infty} \left(1 - \frac{2z+2}{2 + \lambda_j + z} \right)$$

qui est holomorphe sur $\{\mathrm{Re}\,z > -2\}$ car le produit converge normalement sur tout compact par hypothèse. Les zéros de f sont exactement les λ_j auxquels on adjoint 0. Comme les facteurs du produit sont de module ≤ 1 pour $\mathrm{Re}\,z \geq -1$ (dessin), le produit est majoré par 1 et

$$|f(z)| \leq \frac{\text{const.}}{|z|^2}, \quad \mathrm{Re}\,z \geq -1.$$

En particulier f est intégrable sur le droite $\mathrm{Re}\,z = -1$.

(2) Écrivons la formule de Cauchy pour $f(z)$ sur le contour $\Gamma_R := \{|\zeta + 1| = R, \mathrm{Re}\,\zeta \geq -1\} \cup [-1 - iR, -1 + iR]$:

$$f(z) = \frac{1}{2i\pi} \int_{\Gamma_R} \frac{f(\zeta)}{\zeta - z} d\zeta, \quad \mathrm{Re}\,z \geq -1.$$

Sur le demi-cercle $C_R := \{-1 + Re^{i\theta},\ \theta \in [-\pi/2, \pi/2]\}$, l'intégrale est inférieure à :

$$\frac{\text{const.}}{R^2} \int_{-\pi/2}^{\pi/2} \frac{R}{|Re^{i\theta} - z - 1|} dt \underset{R \to +\infty}{\longrightarrow} 0$$

de sorte qu'il ne reste que l'intégrale sur le droite $\{\mathrm{Re}\,z = -1\}$:

$$f(z) = -\frac{1}{2\pi} \int_{-\infty}^{+\infty} \frac{f(-1+is)}{-1 + is - z} ds, \quad \mathrm{Re}\,z \geq -1$$

(3) Très astucieusement, on remarque que :

$$\frac{1}{1+z-is} = \int_0^1 t^{z-is} dt, \quad \operatorname{Re} z \geq -1.$$

De sorte que :

$$f(z) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} f(-1+is) \left(\int_0^1 t^{z-is} dt \right) ds.$$

L'interversion est licite car tout est intégrable :

$$f(z) = \int_0^1 t^z \left(\frac{1}{2\pi} \int_{-\infty}^{+\infty} f(-1+is) t^{-is} ds \right) dt =: \int_0^1 t^z w(t) dt.$$

(4) On définit la forme linéaire sur $\mathscr{C}$:

$$\ell(g) = \int_0^1 g(t) w(t) dt.$$

qui est continue puisque w est intégrable (car $|f(-1+is)| \leq \text{const.}/|1+s^2|$). De plus, pour tout ζ de partie réelle > 0 :

$$\ell(t^\zeta) = f(\zeta) \neq 0, \text{ si } \zeta \neq \lambda_j.$$

On a même montré un peu mieux : si $\lambda \neq \lambda_j$, alors t^λ n'appartient pas à $\overline{\operatorname{Vect} X}$.

□

Le détail de la preuve de l'holomorphie de f pour $(ii) \Rightarrow (i)$ et plein d'autres trucs sont chez Grégoire Clarté. Grâce lui en soit rendue.

Références.
P. Lax, *Functional Analysis*
W. Rudin, *Analyse réelle et complexe.*

201 Espaces de fonctions ; exemples et applications.
202 Exemples de parties denses et applications.
209 Approximation d'une fonction par des polynômes et des polynômes trigonométriques. Exemples et applications.
245 Fonctions holomorphes sur un ouvert de **C**. Exemples et applications.

2.13 Le théorème de relèvement continu

Version semi-périmée.

Théorème (Relèvement continu)**.** *Soient $[a,b]$ un intervalle de $\mathbf{R}$ et $\theta_0 \in \mathbf{R}$. On considère $\gamma : [a,b] \to \mathbf{U}$ une application continue telle que*

$$\gamma(a) = e^{i\theta_0}.$$

Alors il existe une unique application continue $\theta : [a,b] \to \mathbf{R}$ envoyant a sur θ_0 telle que

$$\forall t \in [a,b], \ \gamma(t) = e^{i\theta(t)}. \tag{2.8}$$

Une application continue $[a,b] \to \mathbf{U}$ vérifiant (2.8) s'appelle un relèvement continu de γ. Le lemme suivant sera important :

Lemme 15. *Soient $\theta_1, \theta_2 \in \mathcal{R}(\gamma)$. Alors la fonction $\theta_1 - \theta_2$ est une constante appartenant à $2\pi\mathbf{Z}$.*

Preuve. On a pour tout $t \in I$, $\theta_1(t) - \theta_2(t) \in 2\pi\mathbf{Z}$ par définition d'un relèvement. Puisque θ_1 et θ_2 sont des relèvements continus et que l'image d'un connexe par arcs par une fonction continue est connexe par arcs, la conclusion est immédiate. □

En particulier, si on relève γ sur deux sous-intervalles de $[a,b]$ d'intersection non vide, il est possible de prolonger de façon unique chacun de ces relèvements à la réunion des deux sous-intervalles.

Preuve (du théorème de relèvement continu). L'unicité d'un tel relèvement découle du lemme 15 [7]. Bien qu'il n'existe pas de fonction argument qui soit continue sur tout $\mathbf{C}^*$, il est facile d'en construire une sur n'importe quel plan fendu $\mathbf{C} \setminus \mathbf{R}_- e^{i\alpha}$. Cette remarque permet de construire pour tout $t \in I$ un relèvement local de γ restreint à un voisinage de t. La preuve consiste ensuite à globaliser la construction par connexité.

Étape 1. Soit $t \in I$. Par continuité de γ, il existe un intervalle ouvert $I_t \subset [a,b]$ contenant t et tel que $\gamma(I_t) \subset \mathbf{U} \setminus \{-\gamma(t)\}$. Sur le plan fendu $\mathbf{C} \setminus \mathbf{R}_- \gamma(t)$, on définit une fonction argument continue $\Theta_{\gamma(t)} : \mathbf{C} \setminus \mathbf{R}_- \gamma(t) \to \mathbf{R}$. $\Theta_{\gamma(t)} \circ \gamma$ est un relèvement continu de $\gamma|_{I_t}$.

Étape 2. On définit sur $[a,b]$ la relation $t \sim t'$ si et seulement s'il existe un sous intervalle de I contenant t et t' sur lequel γ se relève continûment [8]. Cette relation est clairement symétrique, elle est réflexive grâce à l'étape 1 et comme on peut prolonger les relèvements elle est aussi transitive. C'est donc une relation d'équivalence. De plus, toujours grâce à l'étape 1, il est facile de voir que les classes de $\sim$ sont ouvertes. Par connexité de $[a,b]$, il n'y a qu'une seule classe qui est $[a,b]$ tout entier. En particulier $a \sim b$ et le théorème est prouvé. □

Application : l'Antipodensatz de Borsuk

Définition (Degré d'un lacet, d'une application)**.** Soit $\gamma : [a,b] \to \mathbf{C}^*$ un lacet. On appelle *degré* de γ l'entier

$$\frac{\theta(b) - \theta(a)}{2\pi}$$

où θ est un relèvement quelconque de $\gamma/|\gamma|$. On définit le degré d'une application continue $f : \mathbf{U} \to \mathbf{C}^*$ comme le degré du lacet :

$$\tilde{f} : t \in [-\pi, \pi] \mapsto f(e^{it}).$$

Lemme (Le degré est localement constant)**.** *Soient γ_1 et γ_2 deux lacets tracés sur $\mathbf{C}^*$. Si $\|\gamma_1 - \gamma_2\|_\infty < \|\gamma_1\|_\infty$ alors γ_1 et γ_2 ont même degré.*

7. Puisque $\mathbf{R}$ est un revêtement de $\mathbf{U}$ via $\theta \mapsto e^{i\theta}$, c'est aussi un cas particulier du théorème d'unicité des relèvements en topologie algébrique, qui se prouve aussi par un argument de connexité !

8. *i.e.* la restriction de γ à ce sous-intervalle admet un relèvement continu

Preuve. La condition implique que $\|\frac{\gamma_2}{\gamma_1} - 1\|_\infty < 1$. Donc le lacet γ_2/γ_1 est tracé dans le plan fendu $\mathbf{C} \setminus \mathbf{R}_-$ et est donc de degré nul. Et il est facile de voir que

$$\deg(\gamma_2/\gamma_1) = \deg(\gamma_2) - \deg(\gamma_1).$$

□

Théorème (Antipodensatz de Borsuk)**.** *Soit $g : \mathbf{S}^2 \to \mathbf{R}^2$ une application continue. Alors il existe un point $\omega \in \mathbf{S}^2$ tel que $g(\omega) = g(-\omega)$.*

Preuve. On définit les applications continues $p : \overline{\mathbf{D}} \to \mathbf{S}^2$ par :

$$\forall z \in \overline{\mathbf{D}},\ \ p(z) = (\mathrm{Re}(z), \mathrm{Im}(z), \sqrt{1 - \mathrm{Re}(z)^2 - \mathrm{Im}(z)^2})$$

et $f : \overline{\mathbf{D}} \to \mathbf{R}^2$ par :

$$\forall z \in \overline{\mathbf{D}},\ \ f(z) = g(p(z)) - g(-p(z)).$$

La restriction de p à $\mathbf{U}$ est impaire, tout comme celle de f. Avec les notations précédentes, cela signifie que pour tout $t \in [-\pi, \pi]$, $\tilde{f}(t + \pi) = -\tilde{f}(t)$. Cela entraîne facilement que le degré de f est impair.
Si f ne s'annulait pas sur $\overline{\mathbf{D}}$, alors considérons l'homotopie

$$H(s, t) := f(se^{it})/|f(se^{it})|$$

qui est une fonction continue en s et en t. Ainsi, $s \mapsto \deg(H(s, \cdot))$ est continue. Comme elle est à valeurs dans $\mathbf{Z}$ elle est constante. Or, $H(0, \cdot)$ est de degré nul et il en est de même pour $H(1, \cdot) = \tilde{f}$. Mais 0 n'est pas impair, d'où contradiction. □

Remarque. Les mêmes arguments conduisent au théorème de Brouwer pour le disque fermé $\overline{\mathbf{D}}$: il suffit de voir qu'il n'existe pas de retraction de $\overline{\mathbf{D}}$ sur $\mathbf{U}$ (ou de façon équivalente que l'identité sur $\mathbf{U}$, qui est impaire, ne peut pas se prolonger en une fonction continue qui envoie $\overline{\mathbf{D}}$ sur $\mathbf{U}$).

Je ne connais pas de référence précise pour ce développement : il s'agit d'une adaptation d'un polycopié de cours de N. Tosel. Cependant, on trouve un théorème un peu différent mais dans un cadre peut être plus adapté à l'agrégation dans *Analyse fonctionnelle* de S. Gonnord et N. Tosel.

204 Connexité. Exemples et applications.
207 Prolongement de fonctions. Exemples et applications. *(mouais)*

2.14 Le théorème de Riesz-Fischer

Théorème (Riesz-Fischer)**.** *L^p est un espace de Banach pour tout $1 \leq p \leq \infty$*

Preuve. On traite d'abord le cas $p = \infty$. Soit (f_n) une suite de Cauchy dans L^∞. Pour tout $k \in \mathbf{N}^*$, il existe $N_k \in \mathbf{N}$ tel que

$$\forall m, n \geq N_k,\ \ \|f_m - f_n\|_{L^\infty} \leq \frac{1}{k}.$$

C'est à dire qu'il existe E_k négligeable tel que

$$\forall x \in \Omega \setminus E_k,\ \forall m, n \geq N_k,\ \ |f_m(x) - f_n(x)| \leq \frac{1}{k}.$$

Soit $E = \cup_k E_k$ qui est négligeable. Pour tout $x \in \Omega \setminus E$, $(f_n(x))$ est une suite de Cauchy dans $\mathbf{R}$. Il existe donc $f(x)$ tel que $f_n(x) \to f(x)$. En passant à la limite $m \to +\infty$ dans la précédente inégalité, on obtient :

$$\forall x \in \Omega \setminus E,\ \forall n \geq N_k,\ \ |f(x) - f_n(x)| \leq \frac{1}{k}.$$

Donc $f \in L^\infty$ et $\|f - f_n\|_{L^\infty} \leq 1/k$ pour tout $n \geq N_k$. Par conséquent $\|f - f_n\|_{L^\infty} \to 0$ ce qui conclut le premier cas.

Supposons maintenant que $1 \leq p < \infty$. Soit (f_n) une suite de Cauchy dans L^p. Pour conclure il suffit de montrer qu'une sous-suite extraite converge dans L^p. On extrait une sous-suite (f_{n_k}) telle que :

$$\forall k \geq 1,\ \ \|f_{n_{k+1}} - f_{n_k}\|_{L^p} \leq \frac{1}{2^k}.$$

On va montrer que la suite (f_{n_k}) converge dans L^p. On pose :

$$g_m(x) = \sum_{k=1}^{m} |f_{n_{k+1}}(x) - f_{n_k}(x)|$$

de sorte que

$$\|g_m\|_{L^p} \leq 1.$$

Le théorème de convergence monotone assure la convergence presque partout de $(g_m)_m$ vers $g \in L^p$. D'autre part : pour $k \geq j \geq 2$:

$$|f_{n_k}(x) - f_{n_j}(x)| \leq |f_{n_k}(x) - f_{n_{k-1}}(x)| + \ldots + |f_{n_{j+1}}(x) - f_{n_j}(x)| \leq g(x) - g_{j-1}(x).$$

Il en résulte que presque partout, $(f_{n_k}(x))$ est une suite de Cauchy et converge vers une limite notée $f(x)$. Presque partout, on a :

$$|f(x) - f_{n_k}(x)| \leq g(x)$$

donc $f \in L^p$. Enfin, on a $|f_{n_k}(x) - f(x)| \to 0$ presque partout et $|f_{n_k}(x) - f(x)|^p \leq |g(x)|^p \in L^1$ donc par le théorème de convergence dominée, on a bien :

$$\|f_{n_k} - f\|_{L^p} \to 0.$$

□

Proposition. *Soient (f_n) une suite de L^p et $f \in L^p$ telle que $\|f_n - f\|_{L^p} \to 0$ avec $1 \leq p \leq \infty$. Alors il existe une sous-suite de (f_n) qui converge presque partout vers f et qui est uniformément majorée par une fonction de L^p.*

Preuve. Le résultat est évident pour $p = \infty$. Supposons donc que $1 \leq p < \infty$. Comme la suite est de Cauchy, on peut extraire une sous-suite (f_{n_k}) comme tout à l'heure. suivant la démonstration précédente, cette sous-suite converge presque partout vers une fonction $\tilde{f}$ avec presque partout :

$$\forall k \in \mathbf{N}^*, \quad |\tilde{f}(x) - f_{n_k}(x)| \leq g(x).$$

Donc $\tilde{f} \in L^p$ et comme tout à l'heure $f_{n_k} \to \tilde{f}$. Par unicité de la limite, on a $\tilde{f} = f$ presque partout. □

Référence. H. Brézis, *Analyse fonctionnelle*

205 Espaces complets. Exemples et applications.
234 Espaces L^p, $1 \leq p \leq +\infty$.

2.15 Le théorème ergodique de Von Neumann

Théorème (Von Neumann)**.** *Soient H un espace de Hilbert. Soit $U(t)$, $t \in \mathbf{R}$, un semi-groupe fortement continu d'opérateurs unitaires*[9] *de H.*

(i) Le sous-espace $F := \bigcap_{t \geq 0} \mathrm{Ker}(U(t) - I)$ est un sous-espace fermé de H et on note P la projection orthogonale sur ce sous-espace.

(ii) Pour $t \geq 0$, on définit l'opérateur « moyenne temporelle » :

$$M(t)g := \frac{1}{t} \int_0^t U(s)g \, ds.$$

C'est un opérateur[10] *continu de norme ≤ 1.*

Lorsque $t \to +\infty$, on a :

$$\forall g \in H, \quad \lim_{t \to +\infty} M(t)g = Pg.$$

Preuve (Hopf). D'abord, $\mathrm{Ker}(U(t) - I)$ est fermé pour tout $t \geq 0$ car c'est l'image réciproque de $\{0\}$ qui est fermé par une application continue, donc F est fermé. De plus, il est facile de voir que si U est unitaire, alors :

$$\mathrm{Ker}(U - I) = \mathrm{Im}(U - I)^{\perp}.$$

En effet,

$$\forall g, h \in H, \quad \langle (U - I)g, h \rangle = \langle g, (I - U)U^* h \rangle.$$

À partir de maintenant, on considère $r \geq 0$ et l'opérateur $U := U(r)$. D'après ce qui précède,

$$H = \mathrm{Ker}(U - I) \oplus_{\perp} \overline{\mathrm{Im}(U - I)}.$$

Si $g \in H$, on note alors $g = e + z$ selon la décomposition précédente.

9. C'est à dire : $U(0) = Id$, $U(t + s) = U(t)U(s)$, $U(t)x \underset{t \to 0}{\longrightarrow} x$ et $U(t)U(t)^* = Id$

10. Pas d'inquiétude, on peut définir l'intégrale de Riemann d'une fonction continue d'un intervalle de $\mathbf{R}$ à valeurs dans un Banach : c'est la limite des sommes de Riemann dont ont peut montrer la cauchyssitude. Bien sûr, toutes les propriétés de l'intégrale sont valables.

(1) On montre que $M(t)z \to 0$. Pour cela, on approche z à ε près par :

$$\|z - z_\varepsilon\| < \varepsilon, \quad z_\varepsilon \in \mathrm{Im}(U - I).$$

On a $\|M(t)z - M(t)z_\varepsilon\| \leq \varepsilon$ et pour un certain $h \in H$:

$$\begin{aligned} M(t)z_\varepsilon &= M(t)(U-I)h = M(t)U(r)h - M(t)h \\ &= \frac{1}{t}\int_0^t U(s+r)h\,ds - \frac{1}{t}\int_0^t U(s)h\,ds \\ &= \frac{1}{t}\int_r^{t+r} U(s)h\,ds - \frac{1}{t}\int_0^t U(s)h\,ds \\ &= \frac{1}{t}\int_t^{t+r} U(s)h\,ds - \frac{1}{t}\int_0^r U(s)h\,ds \end{aligned}$$

et chacun des deux termes est de norme inférieure à $r\|h\|/t \to 0$.

(2) On montre que $M(t)e$ converge vers un élément de $\mathrm{Ker}(U - I)$. D'abord, on remarque que $U(t + r)e = U(t)U(r)e = U(t)e$ donc par r-périodicité en écrivant $t = nr + q$:

$$M(t)e = \frac{1}{t}\int_0^t U(s)e\,ds = \frac{n}{t}\int_0^r U(s)e\,ds + \frac{1}{t}\int_0^q U(s)e\,ds.$$

Lorsque $t \to +\infty$, on obtient :

$$M(t)e \to \frac{1}{r}\int_0^r U(s)e\,ds \in \mathrm{Ker}(U(r) - I).$$

Puisque $r \geq 0$ était choisi arbitrairement, on en déduit que pour tout $g \in H$, la limite de $M(t)g$ existe et appartient à F. Montrons pour conclure que cette limite est bien la projection annoncée. D'abord, on voit que pour tout $g \in F$, et tout $t \geq 0$, $U(t)g = g$ *i.e* $U(t)^*g = g$. Cela implique que $F^\perp$ est stable par tous les $U(s)$ car

$$\forall w \in F^\perp,\ \forall f \in F,\ \ \langle U(s)w, f\rangle = \langle w, U(s)^* f\rangle = \langle w, f\rangle = 0.$$

Du coup, pour tout $t \geq 0$, $M(t)$ stabilise aussi $F^\perp$. Mais alors :

$$\forall g \in F^\perp,\ \ \lim_{t\to+\infty} M(t)g \in F^\perp \cap F = \{0\}$$

et

$$\forall g \in F,\ \ [\forall t \geq 0,\ M(t)g = g] \ \Rightarrow\ \lim_{t\to+\infty} M(t)g = g.$$

Ce qui conclut. □

De la mécanique statistique.

On considère $H = L^2(\mu)$ où μ est une mesure de probabilité sur $\mathbf{R}^{6N}$. On considère le système différentiel :

$$\frac{dZ}{dt}(t) = F(Z(t))$$

qui est de divergence nulle donc le flot préserve la mesure au sens où pour tout $\Omega \subset \mathbf{R}^{6N}$ et tout $t \geq 0$:

$$\mu(X(t,0,\Omega)) = \mu(\Omega).$$

De plus, puisque le système est homogène, le flot définit un semi-groupe fortement continu :

$$U(t)g(z) := g(X(t,0,z)).$$

Puisque le flot préserve la mesure, les opérateurs $U(t)$ sont unitaires. On peut alors montrer (regarder $f^{-1}(]\infty, c[)$ pour $f \in F$ et $c \in \mathbf{R}$) que le sous-espace F est réduit aux constantes si et seulement s'il n'existe pas de sous-ensemble (mesurable) non trivial invariant par le flot (on dit alors que la transformation est **ERGODIQUE**). Dans ce cas, le théorème ergodique stipule que

$$\forall g \in H,\ \frac{1}{t}\int_0^t g(X(s,0,\cdot))ds \overset{L^2(\mu)}{\longrightarrow} Pg = \int_{\mathbf{R}^{6N}} g(z)d\mu(z).$$

Au fondement de la mécanique statistique se tient *l'hypothèse ergodique* qui affirme l'égalité entre les moyennes spatiales et temporelles. Au vu des échelles de temps dans la réalité véritable lorsque $N \sim 10^{23}$, ça ne parait pas délirant.

Enfin, il existe une version discrète de ce théorème sont la preuve est en tout point semblable, on la trouvera au choix dans [Gonnord, Tosel, *Analyse Fonctionnelle*], [Beck, Malick, Peyré, *Objectif Agrégation*] ou [Petersen, *Ergodic Theory*].

Référence. P. D. Lax, *Functional Analysis*

208 Espaces vectoriels normés, applications linéaire continues. Exemples.
213 Espaces de Hilbert. Bases hilbertiennes. Exemples et applications.

2.16 Les théorèmes de Lebesgue et de Rademacher

Il y a deux développements ici.

Le cas unidimensionnel des fonctions à variation bornée.

Théorème (Lebesgue)**.** *Une fonction à variation bornée est presque partout différentiable (au sens de la mesure de Lebesgue μ).*

Preuve. Soit $f : [a,b] \to \mathbf{R}$ à variation bornée. On note D l'ensemble au plus dénombrable de ses points de discontinuité. On notera :

$$f^+(x) = \limsup_{y\to x} \frac{f(y)-f(x)}{y-x} \quad \text{et} \quad \liminf_{y\to x} \frac{f(y)-f(x)}{y-x}$$

et on définit les ensembles :

$$A^+ = \{x \in]a,b[\backslash D,\ f^+(x) = +\infty\},\ \ A^- = \{x \in]a,b[\backslash D,\ f^-(x) = -\infty\}$$

et

$$B = \{x \in]a, b[\setminus D,\ f^+(x) > f^-(x)\}.$$

Étape 1. Un lemme de recouvrement de Vitali.

Lemme (Vitali)**.** *Soient* $(I_n =]x_n - r_n,\ x_n + r_n[)_{1 \leq n \leq N}$ N *intervalle de* $\mathbf{R}$*. Il existe une partie* $J \subset \{1, \ldots, N\}$ *telle que les* $(I_j)_{j \in J}$ *soient deux à deux disjoints et :*

$$\bigcup_{n=1}^{N} I_n \subset \bigcup_{j \in J}]x_j - 3r_j,\ x_j + 3r_j[.$$

En particulier :

$$\mu\left(\bigcup_{i=1}^{N} I_n\right) \leq 3 \sum_{j \in J} \mu(I_j).$$

Preuve. On ordonne les rayons par ordre décroissant : $r_1 \geq r_2 \geq \ldots \geq r_N$. On commence par considérer $J_1 \subset \{1, \ldots, N\}$ maximal telle que les $(I_j)_{j \in J_1}$ ne s'intersectent pas et soient de rayon r_1. Ensuite, soit $J_2 \subset \{1, \ldots, N\}$ maximal tel que les $(I_j)_{j \in J_2}$ soient de rayon r_2 et que les $(I_j)_{J_1 \cup J_2}$ ne s'intersectent pas. On définit de la même façon J_1, $J_2, \ldots, J_k$ et on pose $J = J_1 \cup \ldots J_k$. Cette partie convient puisque si $B(x', r') \cap B(x, r) \neq \emptyset$ entraîne $B(x', r') \subset B(x, 3r)$ donc toute boule éliminée est incluse dans une boule $B(x_j, 3r_j)$ pour un certain $j \in J$. □

On utilisera le corollaire suivant : pour toute famille $(I_x)_{x \in X}$ d'intervalles ouverts de $]a, b[$, on peut d'abord extraire un recouvrement dénombrable $(I_{x_n})_{n \in \mathbf{N}}$ (séparabilité [11] de $\mathbf{R}$) puis en appliquant le lemme de Vitali à la famille $(I_{x_n})_{0 \leq n \leq N}$ telle que

$$\mu\left(\bigcup_{n=0}^{N} I_{x_n}\right) \geq \frac{1}{2}\mu\left(\bigcup_{n \in \mathbf{N}} I_{x_n}\right)$$

on peut affirmer [12] l'existence d'une partie finie F telle que les $(I_x)_{x \in F}$ sont deux à deux disjoints et

$$\sum_{x \in F} \mu(I_x) \geq \frac{1}{6}\mu\left(\bigcup_{x \in X} I_x\right).$$

Étape 2. A^+ et A^- sont de mesure nulle.

Par l'absurde, si $\mu(A^+) > 0$ alors, pour tout $x \in A^+$ et tout $K > 0$, il existe un intervalle $I_x =]a_x, b_x[$ contenant x tel que

$$f(b_x) - f(a_x) > K(b_x - a_x)$$

11. Dans un contexte plus général, il s'agit d'un théorème de Lindelöf que l'on peut trouver dans *Analyse fondamentale* de S. Dolecki. Voici l'idée adaptée au contexte présent : on considère $\{B(a_n, 1/k)\}_{n,k} =: \{U_m\}_{m \in \mathbf{N}}$ où a_n est une suite dense dans $\mathbf{R}$. On note $\mathcal{N} = \{n \in \mathbf{N},\ \exists x \in X,\ U_n \subset I_x\}$. Si $n \in \mathcal{N}$, on note $x_n \in X$ tel que $U_n \subset I_{x_n}$ et on voit que $\bigcup_{n \in \mathcal{N}} I_{x_n}$ convient.

12. Il est raisonnable d'admettre ça en lemme préliminaire avant de développer la preuve.

(on trouve a_x ou b_x et on utilise la continuité de f en x pour trouver l'autre). L'étape 1 donne alors l'existence d'une partie finie F telle que

$$\frac{1}{K}\sum_{x\in F}(f(b_x)-f(a_x)) \geq \sum_{x\in F}\mu(I_x) \geq \frac{1}{6}\mu(A^+)$$

ce qui contredit le fait que f est à variation bornée puisque K est arbitraire.

Étape 3. On suppose par l'absurde que $\mu(B) > 0$.

Pour $(\alpha,\beta) \in \mathbf{Q}_+^* \times \mathbf{Q}$, on définit :

$$C_{\alpha,\beta} = \{x \in B,\ f^+(x) > \beta+\alpha \ \text{ et } \ f^-(x) < \beta-\alpha\}$$

et puisque par densité de $\mathbf{Q}$: $B = \bigcup C_{\alpha,\beta}$, il existe par additivité dénombrable de la mesure un ensemble $C_{\alpha,\beta}$ de mesure strictement positive. Quitte à retrancher $x \mapsto \beta x$ à f, on peut supposer $\beta = 0$.

On va nier la rectifiabilité du graphe de f : prenons F une partie finie de $[a,b]$ contenant a et b. On note a_F la fonction affine par morceau interpolant f aux points de F. Pour tout $x \in C_{\alpha,0} \setminus F$, on choisit $I_x =]a_x, b_x[$ contenu dans $]a,b[$, contenant x tel que $I_x \cap F = \emptyset$ et :

$$\begin{cases} f(b_x)-f(a_x) & \leq & -\alpha(b_x-a_x) & \text{si } a_F \text{ est croissante sur } I_x \\ f(b_x)-f(a_x) & \geq & \alpha(b_x-a_x) & \text{sinon.} \end{cases}$$

Par l'étape 1, il existe une partie finie $E \subset C_{\alpha,0} \setminus F$ telle que les $(I_x)_{x\in F}$ soient deux à deux disjoints et vérifient

$$\sum_{x\in E}\mu(I_x) \geq \frac{\mu(C_{\alpha,0})}{6}.$$

On pose $G = E \cup F$.

Étape 4. La longueur des graphes des fonctions affines par morceaux.

On notera $\ell(a_G)$ et $\ell(a_F)$ la longueur des graphes respectifs de a_F et a_G. Un dessin intelligent montre que :

$$\ell(a_G) \geq \ell(a_F) + (\sqrt{1+\alpha^2}-1)\sum_{x\in E}\mu(I_x) \geq \ell(a_F) + \frac{\mu(C_{\alpha,0})}{6}(\sqrt{1+\alpha^2}-1)$$

ce qui contredit la rectifiabilité du graphe de f, F étant arbitraire.

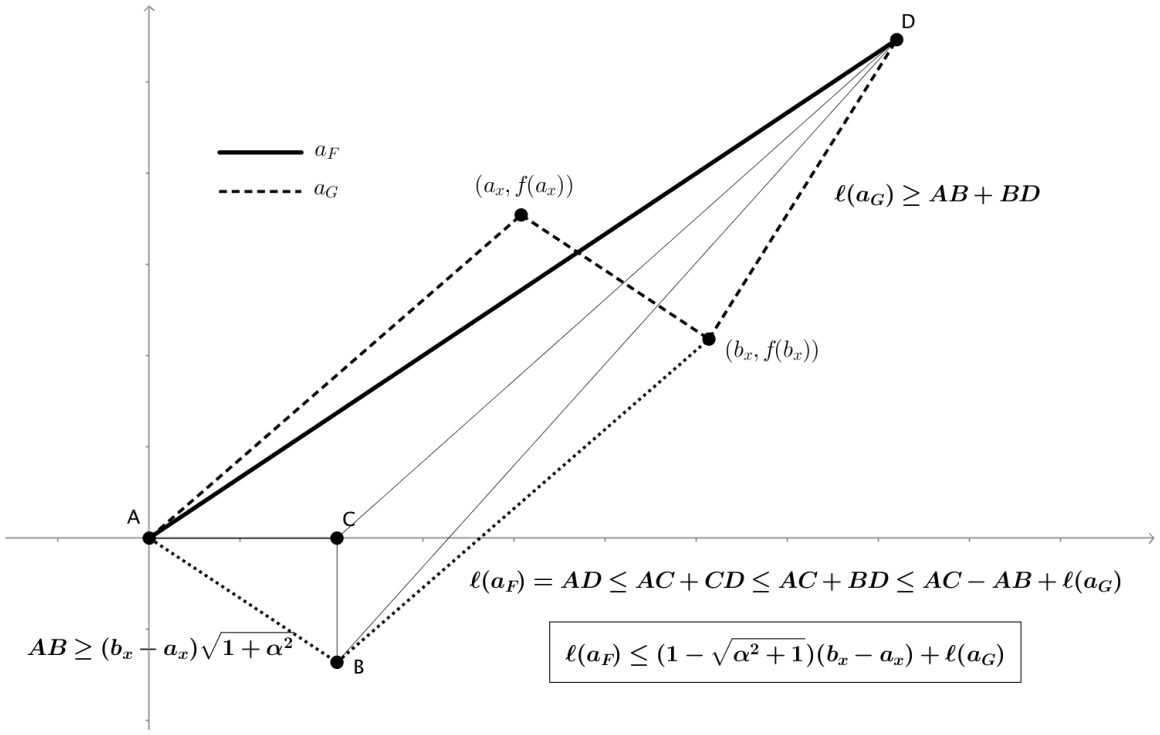


Figure 2.7: Un dessin intelligent. *"Autour de I_x"*.

□

Le cas multidimensionnel des fonctions lipschitziennes.

Théorème (Rademacher)**.** *Toute fonction $f : \mathbf{R}^n \to \mathbf{R}^m$ lipschitzienne est différentiable presque partout (au sens de la mesure de Lebesgue).*

Preuve. On se ramène sans peine à des fonctions $f : \mathbf{R}^n \to \mathbf{R}$ et on admet le cas $n = 1$ qui relève du paragraphe précédent puisque les fonctions lipschitziennes sont à variation bornée. On notera C une constante de Lipschitz de f. En particulier on sait que $\nabla f(x)$ existe pour presque tout $x \in \mathbf{R}^n$ et si $e \in \mathbf{S}^{n-1}$, la dérivée directionnelle

$$L_x(e) = \lim_{t \to 0} \frac{f(x+te) - f(x)}{t}$$

existe pour presque tout $x \in \mathbf{R}^n$. Ces fonctions sont mesurables et dans L^∞.

Étape 1. $L_x(e) = \langle \nabla(x), e \rangle$ au sens faible.

Soit $\varphi \in C_c^\infty(\mathbf{R}^n)$. D'une part, comme $\left| \frac{f(x+te) - f(x)}{t} \right| \leq C$, le théorème de convergence dominée assure :

$$\int_{\mathbf{R}^n} \frac{f(x+te) - f(x)}{t} \varphi(x) dx \underset{t \to 0}{\longrightarrow} \int_{\mathbf{R}^n} L_x(e) \varphi(x) dx.$$

D'autre part, on a aussi toujours par convergence dominée :

$$\int_{\mathbf{R}^n} \frac{f(x+te) - (x)}{t} \varphi(x) dx = \int_{\mathbf{R}^n} \frac{\varphi(x-te) - \varphi(x)}{t} f(x) dx \underset{t \to 0}{\longrightarrow} - \int_{\mathbf{R}^n} \langle \nabla \varphi(x), e \rangle f(x) dx.$$

Maintenant, en écrivant $e = x_1e_1 + \dots x_ne_n$, on développe le gradient et on utilise une dernière fois le théorème de convergence dominée pour justifier :

$$\begin{aligned}\int_{\mathbf{R}^n} \langle \nabla\varphi(x), e\rangle f(x)dx &= \sum_{i=1}^n x_i \lim_{t\to 0} \int_{\mathbf{R}^n} \frac{\varphi(x+te_i)-\varphi(x)}{t} f(x)dx \\ &= \sum_{i=1}^n x_i \lim_{t\to 0} \int_{\mathbf{R}^n} \frac{f(x-te_i)-f(x)}{t} \varphi(x)dx = -\int_{\mathbf{R}^n} \langle \nabla f(x), e\rangle \varphi(x)dx\end{aligned}$$

de sorte que :

$$\forall \varphi \in C_c^\infty(\mathbf{R}^n),\ \int_{\mathbf{R}^n} (L_x(e) - \langle \nabla f(x), e\rangle)\varphi(x)dx = 0$$

Étape 2. Pour presque tout $x \in \mathbf{R}^n$, $L_x(e) = \langle \nabla(x), e\rangle$.

En creux, il s'agit de prouver l'injection $L^\infty(\mathbf{R}^n) \subset L^1_{\text{loc}}(\mathbf{R}^n) \hookrightarrow \mathcal{D}'(\mathbf{R}^n)$. Soit donc $h \in L^1_{\text{loc}}(\mathbf{R}^n)$ tel que

$$\forall \varphi \in C_c^\infty(\mathbf{R}^n),\ \int_{\mathbf{R}^n} h(x)\varphi(x)dx = 0.$$

On montre que $h = 0$ presque partout sur toute boule $B(0,r)$, $r > 0$. On pose $g = h\mathbf{1}_{B(0,2r)}$ et on considère ρ_ε une suite régularisante telle que $\operatorname{Supp}\rho_\varepsilon \subset B(0, r_\varepsilon)$. On sait alors que $\rho_\varepsilon * g \to g$ dans $L^1(\mathbf{R}^n)$. Soit $x \in B(0,r)$, on a :

$$\rho_\varepsilon * g(x) = \int_{B(0,2r)} \rho_\varepsilon(x-y)g(y)dy = \int_{|x-y|\le r_\varepsilon} \rho_\varepsilon(x-y)h(y)dy = \int_{\mathbf{R}^n} \rho_\varepsilon(x-y)h(y)dy$$

car $|x-y| \le r_\varepsilon \Rightarrow |y| \le |y-x| + |x| \le 2r$ pour ε assez petit. Comme $y \mapsto \rho_\varepsilon(x-y)$ est C^∞ à support compact, on a par hypothèse :

$$\rho_\varepsilon * g(x) = 0.$$

En faisant tendre ε vers 0, on obtient $g = 0$ presque partout dans $B(0,r)$ et la conclusion.

Étape 3. Dérivées directionnelles et différentiabilité.

On note A l'ensemble des points où $\nabla f(x)$ existe. Pour conclure quant à la différentiabilité de f, il suffit de montrer que la convergence :

$$\frac{f(x+te)-f(x)}{t} \underset{t\to 0}{\longrightarrow} \langle \nabla f(x), e\rangle$$

qui est vraie à $e \in \mathbf{S}^{n-1}$ fixé pour presque tout $x \in A$ est vraie **uniformément** en e.

Soit $(e_i)_{i\ge 1}$ une suite dense de $\mathbf{S}^{n-1}$. Par additivité dénombrable de la mesure,

$$B = \left\{ x \in A,\ \forall i \ge 1,\ \frac{f(x+te_i)-f(x)}{t} \underset{t\to 0}{\longrightarrow} \langle \nabla f(x), e_i\rangle \right\}.$$

vérifie $\mu(\mathbf{R}^n \setminus B) = 0$. Soit $x \in B$, on définit pour $t \in \mathbf{R}$

$$\phi_t : e \in \mathbf{S}^{n-1} \longmapsto \frac{f(x+te)-f(x)}{t}.$$

Toutes ces fonctions sont C-lipschitziennes. Soit $\varepsilon > 0$. Par compacité de la sphère unité, il existe $i_1, \ldots, i_N$ tels que :

$$\mathbf{S}^{n-1} \subset \bigcup_{k=1}^{N} B(e_{i_k}, \varepsilon/2\widetilde{C})$$

où $\widetilde{C} = C + \|\nabla f(x)\|$. On découpe :

$$\begin{aligned} |\phi_t(e) - \langle \nabla f(x), e\rangle| &\leq |\phi_t(e) - \phi_t(e_{i_k})| + |\phi_t(e_{i_k}) - \langle \nabla f(x), e_{i_k}\rangle| + |\langle \nabla f(x), e_{i_k} - e\rangle| \\ &\leq \widetilde{C}\|e - e_{i_k}\| + |\phi_t(e_{i_k}) - \langle \nabla f(x), e_{i_k}\rangle| \\ &\leq \varepsilon/2 + \varepsilon/2 = \varepsilon \end{aligned}$$

pour t assez grand, dépendant seulement des $e_{i_1}, \ldots, e_{i_N}$ qui sont en nombre fini. Ce qui conclut. □

Appendice : sur les fonctions à variation bornée.

Définition. Une fonction $f : [a, b] \to \mathbf{R}$ est dite à variation bornée lorsqu'elle vérifie l'une des trois conditions équivalentes suivantes :

(i) La variation totale de f définie par :

$$V_{a,b}(f) = \sup \left\{ \sum_{i=1}^{n-1} |f(x_{i+1}) - f(x_i)|,\ a \leq x_1 < x_2 < \ldots < x_n \leq b \right\}$$

est finie.

(ii) Le graphe de f est rectifiable au sens où si F est une partie finie de $[a, b]$ contenant a et b et si a_F désigne la fonction affine par morceaux a_F interpolant f aux points de F,

$$\sup_F \ell(a_F) < \infty$$

où $\ell(a_F)$ est la longueur du graphe de a_F.

(iii) f est la différence de deux fonctions croissantes.

Preuve. L'équivalence (i) $\Leftrightarrow$ (ii) est immédiate [13] en considérant la norme 1 dans $\mathbf{R}^2$. L'implication (iii) $\Rightarrow$ (i) est aussi claire puisqu'une fonction croissante est à variation bornée. Pour la réciproque, il suffit de voir que $x \mapsto V_{a,x}(f)$ et $x \mapsto V_{a,x}(f) - f(x)$ sont croissantes. □

La condition (iii) donne immédiatement : une fonction à variation bornée est continue sauf peut être sur un ensemble au plus dénombrable et une fonction à variation bornée est réglée.

Contre-exemple. La fonction $f : [0, 1] \to \mathbf{R}$ définie par

$$f(x) = x \cos\left(\frac{1}{x}\right) \text{ si } x \neq 0 \text{ et } f(0) = 0$$

est continue mais pas à variation bornée. Pour le voir, il suffit de considérer la subdivision :

$$\sigma_n : 0 < \frac{1}{n\pi} < \frac{1}{(n-1)\pi} < \ldots < \frac{1}{2\pi} < \frac{1}{\pi} < 1.$$

Référence. S. Gonnord, N. Tosel, *Thèmes d'Analyse pour l'agrégation : Calcul Différentiel.*

201 Espaces de fonctions ; exemples et applications.
215 Applications différentiables définies sur un ouvert de $\mathbf{R}^n$. Exemples et applications.
228 Continuité et dérivabilité des fonctions réelles d'une variable réelles. Exemples et applications.
229 Fonctions monotones. Fonctions convexes. Exemples et applications.

13. Ou pas : $\ell(a_F) = \sum \|(x_{i+1} - x_i, f(x_{i+1}) - f(x_i))\|_2 \simeq \sum \|(x_{i+1} - x_i, f(x_{i+1}) - f(x_i))\|_1 \simeq b - a + V_{a,b}(f)$.

2.17 La méthode de Laplace

Théorème (Méthode de Laplace)**.** *Soient $[a,b]$ un intervalle de $\mathbf{R}$ (borné ou non), $\varphi : [a,b] \to \mathbf{R}$ une fonction C^2 et $f : [a,b] \to \mathbf{C}$ L^1 telle que la fonction*

$$F(x) = \int_a^b e^{x\varphi(t)} f(t)dt$$

soit bien définie continue sur $x \in \mathbf{R}_+$.

1. *Si φ atteint un maximum ordinaire en un unique point $t_0 \in]a,b[$, c'est à dire :*

$$\varphi'(t_0) = 0, \quad \varphi''(t_0) < 0, \quad \text{et} \quad \varphi' \text{ ne s'annule qu'en } t_0$$

alors en supposant $f(t_0) \neq 0$:

$$F(x) \underset{x\to+\infty}{\sim} f(t_0) \left(\frac{2\pi}{-\varphi''(t_0)}\right)^{1/2} \frac{e^{x\varphi(t_0)}}{\sqrt{x}}.$$

2. *Si φ atteint son maximum en a et φ' ne s'annule pas sur $]a,b[$, alors en supposant $f(a) \neq 0$:*

$$F(x) \underset{x\to+\infty}{\sim} \frac{f(a)}{-\varphi'(a)} \frac{e^{x\varphi(a)}}{x}.$$

Preuve. On commence par le premier cas et on regarde l'intégrale sur $[t_0,b[$. Le comportement de φ est quadratique au voisinage de t_0 :

$$\varphi(t) - \varphi(t_0) \underset{t\to t_0}{\sim} \frac{1}{2}\varphi''(t_0)(t-t_0)^2.$$

Comme φ est un C^1-difféormorphsime de $[t_0,b[$ sur $[\varphi(t_0),\varphi(b)[$, on peut écrire le changement de variable

$$s = (\varphi(t_0) - \varphi(t))^{1/2} \iff t = \psi(s) := \varphi^{-1}(\varphi(t_0) - s^2).$$

On obtient :

$$\int_{t_0}^b e^{x\varphi(t)} f(t)dt = e^{x\varphi(t_0)} \int_0^{(\varphi(t_0)-\varphi(b))^{1/2}} e^{-xs^2} g(s)ds$$

où $g(s) = f(\psi(s))\psi'(s)$.

Or, le théorème de convergence dominée et la continuité de g en 0 assurent pour α petit :

$$\int_0^\alpha e^{-xs^2} g(s)ds = \frac{1}{\sqrt{x}} \int_0^{\alpha\sqrt{x}} e^{-u^2} g\left(\frac{u}{\sqrt{x}}\right) du \underset{x\to+\infty}{\sim} \frac{g(0)\sqrt{\pi}}{2} \frac{1}{\sqrt{x}}.$$

Le résultat reste valable pour tout $\alpha > 0$ puisque :

$$g \in L^1([a,b[) \implies \int_\alpha^\beta |e^{-xs^2} g(s)|ds \leq e^{-x\alpha^2} \int_\alpha^\beta |g(s)|ds \to 0.$$

Ici, on trouve finalement :

$$e^{x\varphi(t_0)} \int_0^{(\varphi(t_0)-\varphi(b))^{1/2}} e^{-xs^2} g(s)ds \underset{x\to+\infty}{\sim} \frac{\sqrt{\pi}}{2} g(0) \frac{e^{x\varphi(t_0)}}{\sqrt{x}}$$

Il ne reste plus qu'à calculer $g(0) = f(\psi(0))\psi'(0) = f(t_0)\psi'(0)$. Or :

$$\psi^{-1} \circ \psi(s) = s \ \Rightarrow \ \psi'(0)(\psi^{-1})'(t_0) = 1$$

et

$$\psi^{-1}(t) = \sqrt{\varphi(t_0) - \varphi(t)} \ \Rightarrow \ (\psi^{-1})'(t) = \frac{-\varphi'(t)}{2\sqrt{\varphi(t_0) - \varphi(t)}} \underset{t \to t_0}{\sim} \frac{-(t - t_0)\varphi''(t_0)}{\sqrt{2}\sqrt{-\varphi''(t_0)}(t_0 - t)} = \sqrt{\frac{\varphi''(t_0)}{2}}.$$

D'où :

$$g(0) = f(t_0)\sqrt{\frac{2}{-\varphi''(t_0)}}.$$

On trouve le résultat en écrivant les mêmes arguments sur $[a, t_0]$ et en sommant.

Dans le second cas, c'est pareil avec le changement de variable

$$s = \varphi(a) - \varphi(t)$$

du coup, c'est plus simple. □

Théorème (Phase stationnaire)**.** *Soient $\varphi : [a, b] \to \mathbf{R}$ C^∞, $f : [a, b] \to \mathbf{C}$ C_c^∞. Alors la fonction*

$$F(x) = \int_a^b e^{ix\varphi(t)} f(t) dt$$

est bien définie et continue sur $\mathbf{R}_+$. Si φ' s'annule en un unique point t_0, si ce point est intérieur à $[a, b]$ et tel que :

$$\varphi''(t_0) \neq 0 \quad et \quad f(t_0) \neq 0$$

alors :

$$F(x) \underset{x \to +\infty}{\sim} f(t_0) \frac{\sqrt{2\pi}}{\sqrt{|\varphi''(t_0)|}} e^{\mathrm{sgn}(\varphi''(t_0))i\pi/4} \frac{e^{ix\varphi(t_0)}}{\sqrt{x}}.$$

Références.
V. Beck, J. Malick, G. Peyré, *Objectif agrégation*
X. Gourdon, *Analyse*

207 Prolongement de fonctions. Exemples et applications.
223 Suites numériques. Convergence, valeurs d'adhérence. Exemples et applications.
230 Séries de nombres réels ou complexes. Comportement des restes ou des sommes partielles des séries numériques. Exemples.
235 Problèmes d'interversion de limites et d'intégrales.
243 Convergence des séries entières, propriétés de la somme. Exemples et applications.

2.18 Stabilité et instabilité en première approximation

Un lemme d'algèbre linéaire

Fait. *Soient A un endomorphisme de $\mathbf{R}^d$ en $\varepsilon > 0$. Il existe une base dans laquelle la matrice (du complexifié) de A soit triangulaire supérieure avec tous les termes surdiagonaux de module inférieur à ε.*

Preuve. On considère une base (e_i) de trigonalisation de A et on pose $f_i = \alpha^{i-1}e_i$. Le résultat suit pour α suffisamment petit. □

Lemme. *Soit u un endomorphisme de $\mathbf{R}^d$ dont on note $P = P_1P_2$ le polynôme caractéristique où P_1 (resp. P_2) n'à que des racines de partie réelle dans $\mathbf{R}_+^*$ (resp. $\mathbf{R}_-$). On note $E_1 = \operatorname{Ker} P_1(u)$ et $E_2 = \operatorname{Ker} P_2(u)$. Le lemme des noyaux donne $\mathbf{R}^d = E_1 \oplus E_2$. Alors il existe un produit scalaire euclidien $\langle \cdot, \cdot \rangle$ sur $\mathbf{R}^d$ et un réel $\alpha > 0$ tels que si $\| \cdot \|$ désigne la norme euclidienne associée, on a :*

(i) E_1 *et* E_2 *sont orthogonaux pour* $\langle \cdot, \cdot \rangle$

(ii) $\forall h \in E_1$, $\langle df(0)h, h \rangle \geq 2\alpha \|h\|^2$

(iii) $\forall h \in E_2$, $\langle df(0)h, h \rangle \leq \alpha \|h\|^2$

Preuve. Soit $\varepsilon > 0$. Dans une base adaptée à $\mathbf{R}^d = E_1 \oplus E_2$ telle que la matrice $A = (a_{i,j})_{i,j}$ de u soit triangulaire supérieure (par blocs) avec tous les termes surdiagonaux de module inférieur à ε, on pose pour $x = (x_i)_i$ et $y = (y_i)_i$:

$$\langle x, y \rangle = \operatorname{Re} \sum_{i=1}^{d} \overline{x_i} y_i.$$

De sorte que :

$$\langle u(x), x \rangle = \operatorname{Re} \sum_{i=1}^{d} a_{i,i} |x_i|^2 + \operatorname{Re} \sum_{i<j} a_{i,j} \overline{x_j} x_i.$$

Notons que :

$$\left| \sum_{i<j} a_{i,j} \overline{x_j} x_i \right| \leq \varepsilon \sum_{i,j} |x_i||x_j| = \varepsilon \left(\sum_i |x_i|^2 \right)^2 \leq \varepsilon d \|x\|^2.$$

- Si $x \in E_1$, soit $\rho > 0$ tel que toutes les valeurs propres de A soit de partie réelle $> \rho$. On a
$$\langle u(x), x \rangle \geq \rho \|x\|^2 - d\varepsilon \|x\|^2.$$
- Si $x \in E_2$ alors
$$\langle u(x), x \rangle \leq d\varepsilon \|x\|^2.$$

D'où le résultat avec $\alpha = \dfrac{\rho}{3}$ et $\varepsilon = \dfrac{\rho}{3d}$. □

Deux critères de stabilité et d'instabilité

Théorème (Lyapunov)**.** *Soit $f : \Omega \subset \mathbf{R}^d \to \mathbf{R}^d$ un champ localement lipschitzien dont $x_0 \in U \subset \Omega$ est un équilibre. On suppose qu'il existe une fonction $V : U \subset \Omega \to \mathbf{R}$ C^1, dite de Lyapunov, telle que :*

(i) $V(x_0) = 0$

(ii) $\forall x \in U \setminus \{x_0\}$, $V(x) > 0$

(iii) $\forall x \in U, \quad dV(x)f(x) \leq 0.$

Alors x_0 est un équilibre stable et si de plus $dV(x)f(x) < 0$ pour tout $x \in \Omega \setminus \{x_0\}$ alors x_0 est asymptotiquement stable.

Preuve. On note $B_\varepsilon(x_0)$ la boule ouverte de centre x_0 et de rayon ε et $S_\varepsilon(x_0)$ la sphère correspondante.

1. Soit $\varepsilon > 0$ tel que $\overline{B}_\varepsilon \subset U$. Puisque V ne s'annule pas sur le compact $S_\varepsilon(x_0)$, il existe $m > 0$ tel que $V(x) \geq m$ pour tout $x \in S_\varepsilon(x_0)$. En outre, par continuité de V, il existe $0 < \delta < \varepsilon$ tel que $V(x) \leq m/2$ pour tout $x \in \overline{B}_\delta(x_0)$. Soit $x \in B_\delta(x_0)$. Par décroissance de la fonction $t \mapsto V(\phi_t(x))$, on a $V(\phi_t(x)) < m$ pour tout temps où la solution $\phi_t(x)$ est définie. Par continuité des trajectoires, $\phi_t(x)$ ne peut pas traverser $S_\varepsilon(x_0)$ donc pour tout t où la solution est définie, $\phi_t(x_0) \in B_\varepsilon(x_0)$. Par le théorème de sortie de tout compact, la solution est définie pour tout $t > 0$ et reste dans $B_\varepsilon(x_0)$.
2. Si la fonction de Lyapunov est stricte, soit $x \in B_\delta(x_0)$. Par compacité de $\overline{B}_\varepsilon(x_0)$, ou bien $\phi_t(x) \to x_0$ lorsque $t \to +\infty$, ou bien il existe une sous-suite croissante $(t_k)_k$ telle que $\phi_{t_k}(x) \to x_*$ avec $x_* \neq x_0$. Si x_0 n'est pas asymptotiquement stable, on est dans la deuxième situation pour au moins un point $x \in B_\delta(x_0)$. Par continuité et stricte décroissance sur les orbites de V, on a déjà $V(\phi_{t_k}(x)) \to V(x_*)$ et $V(\phi_{t_k}(x)) > V(x_*)$. De plus, toujours par stricte décroissance de V sur les orbites :

$$\lim_{k\to+\infty} V(\phi_{1+t_k}(x)) = \lim_{k\to+\infty} V(\phi_1(\phi_{t_k}(x))) = V(\phi_1(x_*)) < V(x_*).$$

Soient ℓ tel que $V(\phi_{1+t_\ell}(x)) < V(x_*)$ et $j > \ell$ tel que $t_j > 1 + t_\ell$. On a :

$$V(\phi_{t_j}(x)) < V(\phi_{1+t_\ell}(x)) < V(x_*).$$

C'est une contradiction.

□

Théorème (Cetaev)**.** *Soit $f : \Omega \subset \mathbf{R}^d \to \mathbf{R}^d$ un champ localement lipschitzien dont x_0 est un équilibre. On suppose qu'il existe un ouvert $U \subset \Omega$ et une fonction $V : \Omega \to \mathbf{R}$ C^1 telle que :*

(i) $\forall x \in U$, $V(x) > 0$

(ii) $\forall x \in U$, $dV(x)f(x) > 0$

(iii) $\forall x \in \partial U$, $V(x) = 0$

(iv) $x_0 \in \partial U$

Alors l'équilibre x_0 est instable.

Preuve. On peut supposer que $x_0 = 0$. Soit $\varepsilon > 0$ tel que $B := B(0, \varepsilon) \subset \Omega$. Grâce à (iv), on peut choisir $0 < \varepsilon' < \varepsilon$ et $x \in U$ de norme majorée par ε'.

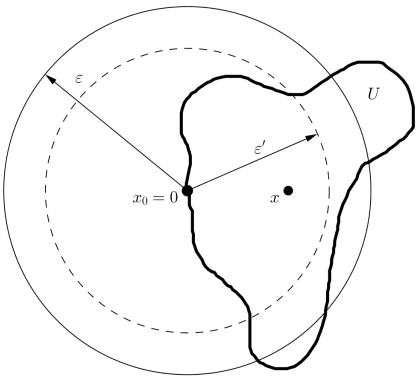


Figure 2.8: Un dessin

Supposons par l'absurde que $\phi_t(x)$ soit définie pour tout $t \in \mathbf{R}_+$, à valeurs dans B. On définit :

$$g : \mathbf{R}_+ \to \mathbf{R}, \;\; t \mapsto V(\phi_t(x))$$

qui est une fonction C^1 sur $\mathbf{R}_+$ avec :

$$g'(t) = dV(\phi_t(x))f(\phi_t(x)).$$

Il est pas trop difficile de voir que $\phi_t(x) \in U$ pour tout $t \in \mathbf{R}_+$. On considère le compact :

$$K = \{y \in \overline{U} \cap B, \; V(y) \geq V(x)\}.$$

Par croissance de g, $\phi_t(x) \in K$ pour tout $t \in \mathbf{R}_+$. De plus $K \subset U$ donc il existe $\alpha > 0$ tel que :

$$\forall y \in K, \;\; dV(y)f(y) \geq \alpha$$

de sorte que $g'(t) \geq \alpha$ pour tout $t \in \mathbf{R}_+$ et $g(t) \underset{t \to +\infty}{\longrightarrow} +\infty$. C'est absurde puisque V est bornée sur K. □

Deux théorèmes de stabilité et d'instabilité en première approximation

Théorème (Stabilité)**.** *Soit $f : \Omega \subset \mathbf{R}^d \to \mathbf{R}^d$ un champ de vecteur C^1 tel que x_0 soit un équilibre et dont la matrice $df(0)$ a toutes ses valeurs propre de partie réelle strictement négative. Alors x_0 est un équilibre asymptotiquement stable.*

Preuve. On suppose que $x_0 = 0$. Par le lemme précédent, il existe $\alpha > 0$, un produit scalaire $\langle \cdot, \cdot \rangle$ et sa norme associée $\| \cdot \|$ tels que :

$$\forall h \in \mathbf{R}^d, \;\; \langle df(0)h, h \rangle \leq -\alpha \|h\|^2.$$

Soit $r > 0$ tel que $B := B(0, r) \subset \Omega$ (boule ouverte pour $\| \cdot \|$) et

$$\forall x \in B, \;\; \|f(x) - df(0)x\| \leq \frac{\alpha}{2}\|x\|.$$

La fonction $V|_B : x \mapsto \|x\|^2$ est une fonction de Lyapunov pour $f|_B$ puisque pour tout $x \in B$:

$$dV(x)f(x) = 2\langle x, f(x)\rangle = 2\Big(\langle x, df(0)x\rangle + \langle x, f(x) - df(0)x\rangle\Big) \leq 2\left(-\alpha\|x\|^2 + \frac{\alpha}{2}\|x\|^2\right) = -\alpha\|x\|^2.$$

Le théorème de Lyapunov permet de conclure. □

Théorème (Instabilité)**.** *Soit $f : \Omega \subset \mathbf{R}^d \to \mathbf{R}^d$ un champ de vecteur C^1 tel que x_0 soit un équilibre et dont la matrice $df(0)$ a au moins une valeur propre de partie réelle strictement positive. Alors x_0 est un équilibre instable.*

Preuve. On suppose que $x_0 = 0$. Soient E_1, E_2, $\alpha > 0$ et un produit scalaire $\langle \cdot, \cdot \rangle$ comme dans le lemme préliminaire. Pour $v = v_1 + v_2 \in E_1 \oplus E_2$, on pose $V(v) = \|v_1\|^2 - \|v_2\|^2$ et on va tenter d'appliquer théorème de Cetaev.

On pose $U_1 = \{x \in \mathbf{R}^d,\ V(x) > 0\}$ et on note :

$$A = df(0), \quad f(x) = Ax + g(x) \ \text{ où } \ g(x) = o(\|x\|).$$

Pour $x = x_1 + x_2 \in U_1$ et $h = h_1 + h_2$, on a :

$$dV(x)f(x) = 2\Big(\langle x_1, Ax_1\rangle - \langle x_2, Ax_2\rangle + \langle x_1 - x_2, g(x)\rangle\Big).$$

Soit $\varepsilon > 0$ tel que pour $\|x\| \leq \varepsilon$:

$$\|g(x)\| \leq \frac{\alpha}{4}\|x\|.$$

Comme $\|x_1 - x_2\| = \|x\|$, on a $|\langle x_1 - x_2, g(x)\rangle| \leq \alpha\|x\|^2/4 \leq \alpha\|x_1\|^2/2$ et

$$dV(x)f(x) \geq 2\alpha(2\|x_1\|^2 - \|x_2\|^2) + 2\langle x_1 - x_2, g(x)\rangle \geq \alpha\|x_1\|^2 > 0.$$

D'où le résultat en appliquant le théorème de Cetaev sur l'ouvert $U_1 \cap B(0, \varepsilon)$. □

Si on ne veut pas parler du résultat d'instabilité, on peut écrire une preuve plus simple du lemme préliminaire (cas où les valeurs propres sont toutes de parties réelles strictement positives), laquelle se trouve dans *Analyse fonctionnelle* de S.Gonnord et N. Tosel ou bien chez Grégoire Clarté.

Références.

S. Gonnord, N. Tosel, *Thèmes d'analyse pour l'agrégation : calcul différentiel*
C. Chicone, *Ordinary Differential Equations with Applications, 2nd Edition.*

220 Équations différentielles $X' = f(t, X)$. Exemples d'étude des solutions en dimension 1 et 2.
221 Équations différentielles linéaires. Système d'équations différentielles linéaires. Exemples et applications.

2.19 Une façon de prolonger la fonction zêta

Quelques pré-requis.

La fonction thêta définie pour $t > 0$ par $\theta(t) = \sum_{n\in\mathbf{Z}} e^{-\pi n^2 t}$ vérifie l'équation fonctionnelle :

$$\theta(t) = \frac{1}{\sqrt{t}}\theta\left(\frac{1}{t}\right).$$

C'est une conséquence immédiate de la formule de Poisson :

$$\sum_{n\in\mathbf{Z}} f(n) = \sum_{n\in\mathbf{Z}} \hat{f}(n)$$

avec la bonne convention pour la transformée de Fourier, de telle sorte que :

$$\mathcal{F}\left(e^{-\alpha|\cdot|^2}\right)(\xi) = \left(\frac{\pi}{\alpha}\right)^{d/2} e^{-\frac{\pi^2}{\alpha}\xi^2}.$$

La fonction $\tilde{\theta}(t) = \sum_{n\geq 1} e^{-\pi n^2 t}$ vérifie alors l'équation fonctionnelle :

$$\tilde{\theta}(t) = \frac{1}{2\sqrt{t}}\left(2\tilde{\theta}\left(\frac{1}{t}\right) + 1\right) - \frac{1}{2}. \tag{2.9}$$

La fonction Γ vérifie :

$$\frac{1}{\Gamma(z)} = ze^{\gamma z}\prod_{n=1}^{+\infty}\left(1 + \frac{z}{n}\right)e^{-\frac{z}{n}} \tag{2.10}$$

expression à partir de laquelle on voit que $1/\Gamma$ se prolonge en une fonction entière (et même $1/z\Gamma$).

Ce qu'on va montrer.

Théorème. *La fonction ζ définie sur le demi-plan $\operatorname{Re} s > 1$ par :*

$$\zeta(s) = \sum_{n=1}^{+\infty}\frac{1}{n^s}$$

se prolonge en une fonction holomorphe sur $\mathbf{C}\setminus\{1\}$ et vérifie l'identité :

$$\pi^{-s/2}\zeta(s)\Gamma\left(\frac{s}{2}\right) = \pi^{-(1-s)/2}\zeta(1-s)\Gamma\left(\frac{1-s}{2}\right). \tag{2.11}$$

Preuve. Il s'agit essentiellement de montrer que ζ satisfait l'équation fonctionnelle annoncée. Montrons la d'abord pour $\operatorname{Re} s > 1$.

Étape 1. Au commencement, une astuce.

Grâce à l'astucieux changement de variable $x = \pi n^2 y$, on va relier Γ et ζ :

$$\Gamma\left(\frac{s}{2}\right) = \int_0^{+\infty} e^{-x}x^{\frac{s}{2}}\frac{dx}{x} = \pi^{s/2}n^s\int_0^{+\infty} e^{-\pi n^2 y}y^{s/2}\frac{dy}{y}$$

de sorte, qu'au moins formellement,

$$\pi^{-s/2}\zeta(s)\Gamma\left(\frac{s}{2}\right) = \sum_{n\geq 1}\int_0^{+\infty} e^{-\pi n^2 y} y^{s/2}\frac{dy}{y}.$$

Bien sûr, on s'empresse d'intervertir la somme et l'intégrale, puisque

$$\sum_{n=1}^{+\infty}\int_0^{+\infty} |e^{-\pi n^2 y} y^{s/2}|\frac{dy}{y} = \sum_{n=1}^{+\infty}\int_0^{+\infty} e^{-\pi n^2 y} y^{\mathrm{Re}\, s/2}\frac{dy}{y} = \pi^{\mathrm{Re}\, s/2}\zeta(\mathrm{Re}\, s)\Gamma(\mathrm{Re}\, s/2) < +\infty.$$

On trouve :

$$\pi^{-s/2}\zeta(s)\Gamma\left(\frac{s}{2}\right) = \int_0^{+\infty}\tilde{\theta}(y) y^{s/2}\frac{dy}{y}.$$

Étape 2. Un peu de calcul.

En utilisant (2.9), on va d'abord écrire :

$$\begin{aligned}\int_0^1 \tilde{\theta}(y) y^{s/2}\frac{dy}{y} &= \int_0^1 \tilde{\theta}\left(\frac{1}{y}\right) y^{(s-1)/2}\frac{dy}{y} + \frac{1}{2}\int_0^1 y^{\frac{s}{2}-\frac{3}{2}}dy - \frac{1}{2}\int_0^1 y^{\frac{s}{2}-\frac{1}{2}}dy \\ &= \int_1^{+\infty}\tilde{\theta}(y) y^{-(s-1)/2}\frac{dy}{y} - \frac{1}{s(1-s)}.\end{aligned}$$

Et finalement,

$$\pi^{-s/2}\zeta(s)\Gamma\left(\frac{s}{2}\right) = \int_1^{+\infty}\left(y^{\frac{s}{2}} + y^{\frac{1-s}{2}}\right)\tilde{\theta}(y)\frac{dy}{y} - \frac{1}{s(1-s)}.$$

Étape 3. Reste à vérifier que tout va bien.

Le terme de droite est clairement invariant par $s \mapsto 1-s$, ce qui montre (2.11) pour $\mathrm{Re}\, s > 1$. Il faut maintenant s'attarder sur des questions de régularité :

(i) On peut sans mal diviser par $\Gamma(s/2)$, on regarde d'abord :

$$\frac{1}{\Gamma(s/2)}\left(\frac{1}{s-1} - \frac{1}{s}\right).$$

Il y a un pôle en $s=1$ mais la singularité en $s=0$ est effaçable, comme le montre (2.10).

(ii) On va montrer que l'intégrale définit une fonction holomorphe sur tout $\mathbf{C}$ et diviser par $\Gamma(s/2)$ ne changera rien. Comme l'intégrande est clairement holomorphe sur $\mathbf{C}$ pour (presque) tout $y \in (1,+\infty)$, on peut appliquer le théorème d'holomorphie sous l'intégrale car :

$$\left|\left(y^{\frac{s}{2}} + y^{\frac{1-s}{2}}\right)\tilde{\theta}(y)\frac{1}{y}\right| \leq \left(\sum_{n=1}^{+\infty} e^{-\pi n y}\right)\left(y^{\frac{\mathrm{Re}\, s}{2}} + y^{\frac{1-\mathrm{Re}\, s}{2}}\right)\frac{1}{y} \leq \frac{y^{\frac{b}{2}} + y^{\frac{1-a}{2}}}{y(e^{\pi n y}-1)} \in L^1_y\big((1,+\infty)\big)$$

si $a < \mathrm{Re}\, s < b$.

En conclusion, la formule :

$$\zeta(s) = \frac{\pi^{s/2}}{\Gamma(s/2)} \int_1^{+\infty} \left(y^{\frac{s}{2}} + y^{\frac{1-s}{2}}\right)\tilde{\theta}(y)\frac{dy}{y} - \frac{\pi^{s/2}}{\Gamma(s/2)s(1-s)}$$

définit bien une fonction holomorphe dans $\mathbf{C} \setminus \{1\}$. □

Remarques complémentaires.

1. L'équation (2.10) s'obtient en fait assez facilement, il suffit d'écrire, par convergence dominée :

$$\Gamma(z) = \lim_{N\to+\infty} \int_0^N \left(1 - \frac{t}{N}\right)^N t^{z-1}dt$$

et d'intégrer N fois par parties pour obtenir :

$$\Gamma(z) = \lim_{N\to+\infty} \frac{N!N^z}{z(z+1)\dots(z+N)}.$$

En inversant cette relation on trouve :

$$\frac{1}{\Gamma(z)} = \lim_{N\to+\infty} N^{-z} \prod_{n=1}^{N} \left(1 + \frac{z}{n}\right).$$

Ensuite, on remarque que :

$$N^{-z} = e^{-z\log N} = e^{z(1+\frac{1}{2}+\dots+\frac{1}{N}-\log N)} \prod_{n=1}^{N} e^{-\frac{z}{n}}$$

ce qui fait apparaitre la constante γ.

2. D'autres formulations de l'équation fonctionnelle (2.11) existent. La plupart s'obtiennent grâce à d'astucieuses formules vérfiées par Γ :

$$\Gamma(1+s)\Gamma(1-s) = \frac{\pi s}{\sin(\pi s)} \quad \text{et} \quad \Gamma(s+1) = 2^s\Gamma(s/2+1)\Gamma(s/2+1/2)\pi^{-1/2}.$$

3. Enfin, la formule de Poisson pour une fonction continue f vérifiant :

$$\exists M > 0, \alpha > 1,\ \forall x \in \mathbf{R}, |f(x)| \leq M(1+|x|)^{-\alpha} \quad \text{et} \quad \sum_{n\in\mathbf{Z}} |\hat{f}(n)| < +\infty$$

n'a rien à voir avec le sujet. Elle s'obtient en écrivant la décomposition en série de Fourier de la fonction 1-périodique :

$$g(x) = \sum_{n\in\mathbf{Z}} f(x+n).$$

Références.

B. Candelpergher, *Calcul intégral*
C. Zuily, H. Queffélec, *Analyse pour l'agrégation*

207 Prolongement de fonctions. Exemples et applications.
235 Problèmes d'interversion de limites et d'intégrales.
239 Fonctions définies par une intégrale dépendant d'un paramètre. Exemples et applications.
241 Suites et séries de fonctions. Exemples et contre-exemples.
245 Fonctions holomorphes sur un ouvert de **C**. Exemples et applications.

2.20 Une formule d'inversion de Fourier

Si μ est une mesure de probabilité sur $\mathbf{R}$ on définit $\varphi(t) = \int e^{itx}\mu(dx)$ sa fonction caractéristique.

Théorème (Inversion)**.** *Si $a < b$, alors la limite suivante existe et vaut :*

$$\lim_{T\to+\infty} \frac{1}{2\pi}\int_{-T}^{T} \frac{e^{-ita}-e^{-itb}}{it}\varphi(t)dt = \mu(a,b) + \frac{1}{2}\mu(\{a,b\}).$$

Preuve. On note :

$$I_T = \int_{-T}^{T} \frac{e^{-ita}-e^{-itb}}{it}\varphi(t)dt = \int_{-T}^{T}\int \frac{e^{-ita}-e^{-itb}}{it}e^{itx}\mu(dx)dt.$$

Ensuite, on calcule bêtement :

$$\begin{aligned}
\frac{e^{-ita}-e^{-itb}}{it}e^{itx} &= \frac{e^{it(b-a)/2}-e^{-it(b-a)/2}}{it}e^{it(x-(a+b)/2)} \\
&= 2\frac{\sin\left(\frac{t(b-a)}{2}\right)}{t}\Big(\cos\big(t(x-(a+b)/2)\big) + i\sin\big(t(x-(a+b)/2)\big)\Big) \\
&= \frac{\sin\big(t(x-a)\big)}{t} - \frac{\sin\big(t(x-b)\big)}{t} + 2i\frac{\sin\big(t(b-a)/2\big)}{t}\sin\big(t(x-(a+b)/2)\big)
\end{aligned}$$

Pour tout $x \in \mathbf{R}$, c'est la somme de trois fonctions continues en t donc intégrables sur $(-T,T)$. Par imparité, l'intégrale de la troisième fonction sur cet ouvert est nulle. Finalement, par le théorème de Fubini (qui est justifié car tout est intégrable) :

$$I_T = \int\left(\int_{-T}^{T}\frac{\sin\big(t(x-a)\big)}{t}dt - \int_{-T}^{T}\frac{\sin\big(t(x-b)\big)}{t}dt\right)\mu(dx).$$

On définit $R(\theta,T) := \int_{-T}^{T}\sin(\theta t)/t\ dt$. Un changement de variable donne :

$$R(\theta,T) = 2\int_0^{T\theta}\frac{\sin x}{x}dx = 2S(T\theta)$$

avec $S(T) := \int_0^T \sin x/x\ dx$. Une rapide étude de signe montre que :

$$\forall\theta\in\mathbf{R},\quad R(\theta,T) = 2\,\mathrm{sgn}(\theta)S(T|\theta|)$$

Et comme $S(T)\to\pi/2$ lorsque $T\to+\infty$, on a finalement :

$$R(x-a,T)-R(x-b,T) \underset{T\to+\infty}{\longrightarrow} \begin{cases} 2\pi & \text{si } a<x<b \\ \pi & \text{si } x=a \text{ ou } x=b \\ 0 & \text{si } x<a \text{ ou } x>b \end{cases}$$

Comme $|R(\theta,T)| \leq 2\sup_y S(y) < +\infty$, on peut appliquer le théorème de convergence dominée et on obtient :

$$I_T \underset{T\to+\infty}{\longrightarrow} 2\pi\mu(a,b) + \pi\mu(\{a\}) + \pi\mu(\{b\}).$$

□

Théorème. *Si φ est intégrable sur $\mathbf{R}$ par rapport à la mesure de Lebesgue, alors μ admet une densité continue et bornée :*

$$f(y) := \frac{1}{2\pi}\int e^{-ity}\varphi(t)dt.$$

Preuve. Puisque φ est intégrable, il en est de même de l'intégrande dans le théorème précédent et on peut écrire :

$$\mu(a,b) + \frac{1}{2}\mu(\{a,b\}) \leq \frac{1}{2\pi}\int_{\mathbf{R}} \left|\frac{e^{-ita} - e^{-itb}}{it}\varphi(t)\right| dt \leq \frac{b-a}{2\pi}\int_{\mathbf{R}} |\varphi(t)|dt.$$

Ce qui montre que μ n'a pas d'atome (prendre $b_n \downarrow a$) et on peut écrire :

$$\begin{aligned}
\mu(x,x+h) &= \frac{1}{2\pi}\int \frac{e^{-itx} - e^{-it(x+h)}}{it}\varphi(t)dt \\
&= \frac{1}{2\pi}\int \left(\int_x^{x+h} e^{-ity}dy\right)\varphi(t)dt \\
&= \int_x^{x+h}\left(\frac{1}{2\pi}\int e^{-ity}\varphi(t)dt\right)dy
\end{aligned}$$

où le théorème de Fubini justifie l'interversion. La continuité et la bornitude de f découlent du théorème de convergence dominée. □

Fait. $\lim_{T\to+\infty}\int_0^T \frac{\sin x}{x}dx = \frac{\pi}{2}$

Preuve. Pour $T > 0$:

$$\int_0^T dx\int_0^{+\infty} dy|e^{-xy}\sin x| = \int_0^T \frac{|\sin x|}{x}dx < \infty$$

donc on peut appliquer le théorème de Fubini :

$$\int_0^T \frac{\sin x}{x}dx = \int_0^{+\infty} dy\int_0^T dx\, e^{-xy}\sin x$$

et

$$\int_0^T e^{x(i-y)}dx = \frac{1}{1+y^2}(y+i)(1-e^{T(i-y)})$$

de sorte que :

$$\int_0^T e^{-xy}\sin x\, dx = \frac{1}{1+y^2}\big(1 - e^{-yT}(y\sin T + \cos T)\big)$$

et on conclut en appliquant le théorème de convergence dominée. □

Référence. R. Durrett, *Probability : Theory and Examples*

236 Illustrer par des exemples quelques méthodes de calcul d'intégrales de fonctions d'une ou plusieurs variables.
261 Fonction caractéristique d'une variable aléatoire. Exemples et applications.
263 Variables aléatoires à densité. Exemples et applications.

3 Non exclusifs

3.1 Analyse du θ-schéma pour l'équation de la chaleur

C'est long mais on peut choisir d'en présenter seulement quelques parties. Par exemple, je prévoyais de développer seulement la stabilité et la convergence et d'admettre la consistance qui est un calcul fort peu intéressant.

Introduction

On cherche une méthode numérique de résolution de l'équation de la chaleur (dans un cercle) :

$$\left\{ \begin{array}{rcl} \partial_t u & = & \nu \partial^2_{xx} u \quad (x,t) \in (0,1) \times \mathbf{R}_+^* \\ u(0,x) & = & u_0(x) \end{array} \right.$$

On admet que le problème a une unique solution, notée u qui est suffisamment régulière. On discrétise le temps et l'espace :

$$\Delta t > 0, \quad \Delta x = \frac{1}{N+1}, \quad (t_n, x_j) = (n\Delta t, j\Delta x).$$

On pose pour $j \in \{0, \dots, N+1\}$, $u_j^0 = u_0(x_j)$. On définit le θ-schéma pour $\theta \in [0,1]$:

$$F\Big(\{u_j^n\}\Big) := \frac{u_j^{n+1} - u_j^n}{\Delta t} + \theta\nu \frac{-u_{j-1}^{n+1} + 2u_j^{n+1} - u_{j+1}^{n+1}}{(\Delta x)^2} + (1-\theta)\nu \frac{-u_{j-1}^n + 2u_j^n - u_{j+1}^n}{(\Delta x)^2} = 0$$

où u_j^n est une approximation de $u(t_n, x_j)$ pour $n \geq 0$ et $j \in \{0, \dots, N+1\}$. Les conditions au bord sont périodiques : $u_{N+1}^n = u_0^N = 0$ de sorte qu'on s'intéresse à $(u_j^n)_{1 \leq j \leq N} \in \mathbf{C}^N$ pour tout $n \in \mathbf{N}$. De façon plus concise on peut écrire :

$$(Id + s\theta A)u^{n+1} = (Id - s(1-\theta)A)u^n$$

où

$$s = \nu \frac{\Delta t}{(\Delta x)^2} \quad \text{et} \quad A = \begin{pmatrix} 2 & -1 & & & \\ -1 & 2 & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & \ddots & \ddots & -1 \\ & & & -1 & 2 \end{pmatrix} \in \mathcal{M}_N(\mathbf{C}).$$

Consistance

On dit que le schéma donné par F est consistant lorsque $F\Big(\big\{u(t_n,x_j)\big\}\Big) \to 0$ lorsque $\Delta t, \Delta x \to 0$. On dit qu'il est d'ordre (p,q) lorsque $F\Big(\big\{u(t_n,x_j)\big\}\Big) = \mathcal{O}\Big((\Delta t)^p + (\Delta x)^q\Big)$.

Appliquons les formules de Taylor :

$$\frac{u(t_{n+1},x_j)-u(t_n,x_j)}{\Delta t} = \partial_t u(t_n,x_j) + \frac{\Delta t}{2}\partial^2_{tt}(t_n,x_j) + \mathcal{O}\big((\Delta t)^2\big).$$

$$\frac{u(t_n,x_{j-1})-2u(t_n,x_j)+u(t_n,x_{j+1})}{(\Delta x)^2} = \partial^2_{xx} u(t_n,x_j) + \mathcal{O}\big((\Delta x)^2\big).$$

$$\begin{aligned}\frac{u(t_n,x_{j-1})-2u(t_n,x_j)+u(t_n,x_{j+1})}{(\Delta x)^2} &= \partial^2_{xx} u(t_n,x_j) + \mathcal{O}\big((\Delta x)^2\big)\\ &= \frac{1}{\nu}\partial_t u(t_{n+1},x_j) + \mathcal{O}\big((\Delta x)^2\big)\\ &= \frac{1}{\nu}\partial_t u(t_n,x_j) + \frac{\Delta t}{\nu}\partial^2_{tt} u(t_n,x_j) + \mathcal{O}\big((\Delta x)^2+(\Delta t)^2\big)\end{aligned}$$

De sorte que :

$$F\Big(\big\{u(t_n,x_j)\big\}\Big) = \Delta t\left(\frac{1}{2}-\theta\right)\partial^2_{tt} u(t_n,x_j) + \mathcal{O}\big((\Delta x)^2+(\Delta t)^2\big).$$

Stabilité L^2

On dit que le schéma est stable en norme $\|\cdot\|$ lorsqu'il existe une constante $K>0$ indépendante de Δt et Δx telle que pour tout $n>0$, $\big\|\{u_j^n\}_j\big\| \le K\big\|\{u_j^0\}_j\big\|$. Dans la suite, on étudie la stabilité du θ-schéma en norme L^2.

On note u^n la fonction constante par morceaux égale à u_j^n sur chaque intervalle $[(j-1/2)\Delta x,(j+1/2)\Delta x]$, 1-périodique. Alors,

$$\|u^n\|_{L^2} = \left(\int_0^1 |u^n(x)|^2 dx\right)^{1/2} = \left(\sum_{j=0}\Delta x|u_j^n|^2\right)^{1/2} = \big\|\{u_j^n\}_j\big\|_{\ell^2}.$$

En tant que fonction périodique de $L^2(0,1)$, u^n se décompose en série de Fourier :

$$u^n(x) = \sum_{k\in\mathbf{Z}} \hat{u}^n(k)\exp(2i\pi kx) \quad \text{où} \quad \hat{u}^n(k) = \int_0^1 u^n(x)\exp(-2i\pi kx)dx.$$

Notons que si $v^n(x) = u^n(x+\Delta x)$ alors $\hat{v}^n(k) = \hat{u}^n(k)\exp(2i\pi k\Delta x)$ et que la norme L^2 de u^n est reliée à la norme ℓ^2 des coefficients par la formule de Plancherel :

$$\|u^n\|_{L^2} = \left(\sum_{k\in\mathbf{Z}}|\hat{u}^n(k)|^2\right)^{1/2}.$$

En Fourier-transformant la définition du schéma, on trouve :

$$\frac{\hat{u}^{n+1}(k)-\hat{u}^n(k)}{\Delta t}+\theta\nu\frac{-\hat{u}^{n+1}(k)e^{-2i\pi k\Delta x}+2\hat{u}^{n+1}(k)-\hat{u}^{n+1}(k)e^{2i\pi k\Delta x}}{(\Delta x)^2}$$
$$+(1-\theta)\nu\frac{-\hat{u}^n(k)e^{-2i\pi k\Delta x}+2\hat{u}^n(k)-\hat{u}^n(k)e^{2i\pi k\Delta x}}{(\Delta x)^2}=0$$

c'est à dire :

$$\hat{u}^{n+1}(k)=\frac{1-2\frac{(1-\theta)\nu\Delta t}{(\Delta x)^2}(1-\cos(2\pi k\Delta x))}{1+2\frac{\theta\nu\Delta t}{(\Delta x)^2}(1-\cos(2\pi k\Delta x))}\hat{u}^n(k)=:A(k)\hat{u}^n(k).$$

La formule de Plancherel donne :

$$\|u^n\|_2^2=\sum_{k\in\mathbf{Z}}|\hat{u}^n(k)|^2=\sum_{k\in\mathbf{Z}}|A(k)^n\hat{u}^0(k)|^2$$

d'où l'on voit que la stabilité du schéma est équivalente à la condition de von Neumann :

$$\forall k\in\mathbf{Z},\ \ |A(k)|\leq 1.$$

(prendre u^0 avec un seul mode de Fourier non nul tel que $|A(k)|>1$). Dans le cas présent, on résume la situation par :

Proposition. *Le θ-schéma est stable en norme L^2 inconditionnellement si $1/2\leq\theta\leq 1$ et sous le condition CFL (Courant-Friedrichs-Lewy)*

$$2(1-2\theta)\nu\Delta t\leq(\Delta x)^2$$

si $0\leq\theta<1/2$.

Preuve. La condition s'écrit :

$$-1-2s\theta(1-\cos(2\pi k\Delta x))\leq 1-2(1-\theta)s(1-\cos(2\pi k\Delta x))$$

C'est à dire :

$$s(1-2\theta)(1-\cos(2\pi k\Delta x))\leq 1.$$

Et une formule de trigonométrie donne :

$$2s(1-2\theta)\sin^2(\pi k\Delta x)\leq 1.$$

Ceci devant être vrai pour tout $k\in\mathbf{Z}$ et tout $\Delta x>0$. □

Convergence

En notant $e_j^n=u_j^n-u(t_n,x_j)$, on montre (c'est un cas local du théorème de Lax) que sous hypothèses de consistance et de stabilité L^2, le θ-schéma est convergent au sens où :

$$\forall T>0,\quad\lim_{\Delta t,\Delta x\to 0}\left(\sup_{t_n\leq T}\|e^n\|_{\ell^2}\right)=0.$$

Comme le schéma est linéaire, on peut écrire :

$$u^{n+1} = Au^n$$

pour une certaine matrice A. En notant $\tilde{u}_j^n = u(t_n, x_j)$, on a par consistance du schéma

$$\tilde{u}^{n+1} = A\tilde{u}^n + \Delta t \varepsilon^n \quad \text{où} \quad \lim_{\Delta t, \Delta x \to 0} \|\varepsilon^n\| = 0.$$

Puis :

$$e^{n+1} = Ae^n - \Delta t \varepsilon^n.$$

et par stabilité ℓ^2 du schéma, $\|A^n\| \leq K$ donc :

$$\|e^n\|_{\ell^2} \leq \Delta t \sum_{k=1}^{n} \|A^{n-k}\| \|\varepsilon^{k-1}\| \leq \Delta t n K C \Big((\Delta x)^p + (\Delta t)^q \Big).$$

Calcul numérique et FFT.

On définit la convolution circulaire de deux vecteurs $f, g \in \mathbf{C}^N$ par :

$$f * g(j) = \sum_{k=1}^{N} f(k) g(j-k)$$

où les coefficients sont vus modulo N avec la convention $f(N+1) = g(N+1) = f(0) = g(0) = 0$. Plus précisément :

$$f * g(j) = \sum_{k=1}^{N} f(k) g(j-k) = \sum_{k=1}^{j-1} f(k) g(j-k) + \sum_{k=j+1}^{N} f(k) g(j-k+N).$$

Avec cette écriture et en posant $g = (s,\, 0, \ldots,\, s,\, -2s) \in \mathbf{C}^N$, le θ-schéma s'écrit :

$$u^{n+1} - u^n = g * (\theta u^{n+1} + (1-\theta) u^n).$$

Maintenant, on définit la transformée de Fourier d'un vecteur $v \in \mathbf{C}^N$ par :

$$\forall k \in \{1, \ldots, N\}, \quad \hat{v}(k) := \sum_{n=1}^{N} v(n) e^{2i\pi kn/N}.$$

Proposition. *La transformée de Fourier discrète jouit des propriétés suivantes :*

- *Pour $f, g \in \mathbf{C}^N$, on a :*

$$\forall k \in \{1, \ldots, N\}, \quad \widehat{f * g}(k) = \hat{f}(k) \hat{g}(k).$$

- *On a la formule d'inversion suivante :*

$$\forall n \in \{1, \ldots, N\}, \quad f(n) = \frac{1}{N} \sum_{k=1}^{N} \hat{f}(k) e^{2i\pi nk/N}.$$

- *Il existe un algorithme permettant de calculer la transformée de Fourier discrète d'un vecteur de $\mathbf{C}^N$ (ou son inverse) en $\mathcal{O}(N \log N)$ opérations.*

On dispose dès lors d'un moyen efficace pour le calcul numérique du θ-schéma puisque :

$$\forall k \in \{1, \dots, N\}, \ \hat{u}^{n+1}(k) = \frac{1 + (1-\theta)\hat{g}(k)}{1 - \theta\hat{g}(k)} \hat{u^n}(k)$$

avec

$$\forall k \in \{1, \dots, N\}, \ \hat{g}(k) = 2s\left(\cos\left(\frac{2k\pi}{N}\right) - 1\right).$$

Il suffit d'inverser la relation pour trouver $\hat{u}^n$ à l'itération voulue. Remarquons qu'on retrouve l'étude de stabilité de von Neumann, la condition de stabilité étant ici :

$$\forall k \in \{1, \dots, N\}, \quad \left|\frac{1 + (1-\theta)\hat{g}(k)}{1 - \theta\hat{g}(k)}\right| \leq 1.$$

Références.
G. Allaire, *Analyse numérique et à optimisation*
G. Peyré, *L'algèbre discrète de la transformée de Fourier*

110 Caractères d'un groupe abélien fini et transformée de Fourier discrète. Applications. *(après FFT)*
218 Applications des formules de Taylor. *(ou pas, du coup)*
222 Exemples d'équations aux dérivées partielles linéaires.
233 Méthodes itératives en analyse numérique matricielle.
246 Séries de Fourier. Exemples et applications.

3.2 De la manière de battre les cartes en Amérique

C'est trop super (quand on a compris) !

Un paquet de cartes est un ensemble fini $\mathscr{C} = \{1, \dots, n\}$ muni d'une relation d'ordre total $\prec$ (où $i \prec j$ signifie que la carte i est au dessus de la carte j). L'ensemble des relations d'ordre totales sur $\mathscr{C}$ est en bijection [1] avec l'ensemble $\mathfrak{S}_n$ des permutations de $\mathscr{C}$: par exemple si $X \in \mathfrak{S}_n$, on définit l'ordre $\prec$ par : $X(1) \prec \dots \prec X(n)$. Un paquet de cartes est donc aussi une permutation de $\mathscr{C}$. Un mélange du paquet est l'action par composition d'une permutation de $\mathscr{C}$ sur le paquet : par exemple, si $\sigma \in \mathfrak{S}_n$, le paquet $X(1) \prec \dots \prec X(n)$ devient le paquet $\sigma \circ X(1) \prec \dots \prec \sigma \circ X(n)$. [2]

1. Il y a plein de bijections, on prendra celle explicitée ci-après.
2. Il n'est pas utile de noter différemment la relation d'ordre pour au moins deux raisons : d'abord parce que sa définition est contenue dans l'écriture ci-dessus et aussi parce que $\prec$ pourra garder de fait sa signification concrète tout au long du texte. En cas de réel besoin, on notera $\prec_\sigma$ l'ordre induit par la permutation σ mais génériquement, un paquet sera toujours noté $X(1) \prec \dots \prec X(n)$.

3. Non exclusifs

Dans la suite, un paquet X est une variable aléatoire à valeurs dans $\mathfrak{S}_n$. Il est dit *bien mélangé* lorsque X suit la loi uniforme sur $\mathfrak{S}_n$, notée π. On s'intéresse à des mélanges successifs du paquet, modélisés par la chaîne de Markov $(X_k)_{k\in\mathbf{N}}$:

$$X_{k+1} = f(X_k, A_k), \quad X_0 = id$$

où les $(A_k)_{k\in\mathbf{N}^*}$ sont des variables aléatoires indépendantes à valeurs dans un certaine espace et qui suivent une loi donnée Q appelée *loi du mélange* (c'est la manière choisie de battre les cartes). On note μ_n la loi de X_n.

Problème. *La loi de $(X_k)_k$ converge-t-elle vers la loi uniforme π ? Dans quel sens ? Et à quelle vitesse ?*

On quantifie la convergence à l'aide de la :

Définition. On appelle *distance de variation totale* entre deux lois Q_1 et Q_2 sur $\mathfrak{S}_n$ la quantité :

$$d_v(Q_1, Q_2) := \max_{S\subset\mathfrak{S}_n} |Q_1(S) - Q_2(S)| = \frac{1}{2}\sum_{\sigma\in\mathfrak{S}_n} \big|Q_1\big(\{\sigma\}\big) - Q_2\big(\{\sigma\}\big)\big|$$

On peut montrer que cette distance métrise la convergence en loi.

Le *mélange américain* consiste à couper aléatoirement le paquet en deux et à insérer les cartes du premier paquet aléatoirement au sein du second en gardant l'ordre relatif des paquets. Une façon équivalente voir les choses est de considérer le *mélange inverse* : on choisit aléatoirement k cartes que l'on place au dessus du paquet en gardant l'ordre relatif. Dit autrement : on se donne une partie $\mathcal{A} = \{i_1, \dots, i_k\} \subset \{1, \dots, n\}$ et on place les cartes en positions $i_1, \dots, i_k$ au dessus du paquet en gardant l'ordre relatif. Par exemple si $\mathcal{A} = \{1, 3\}$ et $n = 4$, le paquet $2 \prec 4 \prec 1 \prec 3$ devient le paquet $2 \prec 1 \prec 4 \prec 3$ et la permutation associé par le mélange inverse est :

$$\pi_0 = \begin{pmatrix} 2 & 4 & 1 & 3 \\ 2 & 1 & 4 & 3 \end{pmatrix}$$

elle dépend du paquet :

$$X = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 1 & 3 \end{pmatrix} \quad \text{qui devient} \quad \pi_0 \circ X = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 4 & 3 \end{pmatrix}$$

C'est complètement tordu mais c'est très pratique parce qu'il suffit d'écrire que le paquet $X(1) \prec \dots \prec X(n)$ devient le paquet $\pi_0 \circ X(1) \prec \dots \prec \pi_0 \circ X(n)$ et en fait on n'a pas du tout besoin de savoir comment s'écrit π_0 : étant donné un paquet X et une partie $\mathcal{A}$ on peut en déduire le nouveau paquet [3] : on vient de définir le f de la chaîne de Markov.

Remarquons que tirer une variable aléatoire $\mathscr{A}$ à valeurs dans $\mathcal{P}(\{1\dots, n\})$ suivant la loi uniforme revient à tirer une variable aléatoire π_0 à valeurs dans $\mathfrak{S}_n$ (associée par le mélange inverse) qui suit la loi :

$$Q(\{\sigma\}) := \begin{cases} \frac{n+1}{2^n} & \text{si } \sigma = id \\ \frac{1}{2^n} & \text{si } \sigma \in H \setminus \{id\} \\ 0 & \text{sinon.} \end{cases}$$

3. On pourrait l'écrire mais bon...

où H est l'ensemble des permutations σ comportant au plus une décroissance ($X(C_i) \prec X(C_j)$ et $X(C_j) \prec_\sigma X(C_i)$). En effet, le choix des parties de la forme $\{1, \dots, j\}$ conduisent à l'identité. C'est néanmoins très inutile d'écrire cela.

Théorème. *Après avoir effectués k mélanges à l'américaine sur un jeu de n cartes, la distance de variation totale vérifie :*

$$d_v(\mu_k, \pi) \leq 1 - \prod_{i=1}^{n-1} \left(1 - \frac{i}{2^k}\right).$$

Preuve. Avec les notations précédentes, notons $(\mathscr{A}_k)_{k\in\mathbf{N}^*}$ une suite de variables aléatoires indépendantes à valeurs dans $\mathcal{P}(\{1, \dots, n\})$ et suivant la loi uniforme. Soit $(\pi_k)_k$ la suite des permutations associée par le mélange inverse et $X_k = \pi_k \circ X_{k-1}$ le paquet à l'instant k. Soit $i \in \{1, \dots, n\}$. Pour tout $k \in \mathbf{N}^*$, on pose $b_k^{(i)} = \mathbf{1}_{i\in\mathcal{A}_k^c}$: puisque les $\mathscr{A}_k$ suivent la loi uniforme et sont indépendantes, les $(b_k^{(i)})_k$ sont des variables aléatoires de Bernoulli indépendantes de paramètre 1/2. On note enfin $b^{(k)}(i) = b_k^{(i)} \dots b_1^{(i)}$ le nombre binaire aléatoire construit par concaténation [4] Soit T le plus petit instant où tous les nombres $b^{(k)}(1), \dots, b^{(k)}(n)$ sont distincts. À l'instant T le paquet est bien mélangé au sens défini précédemment. En effet, à cet instant (mais pas avant), les $b^{(T)}(i)$ définissent un ordre total sur $\mathscr{C}$ et comme toutes les variables aléatoires sont indépendantes et suivent une loi uniforme, l'ordre construit suit aussi une loi uniforme. Or, on a pour tout $\mathfrak{P} \subset \mathfrak{S}_n$:

$$\begin{aligned}
\mu_k(\mathfrak{P}) &= \mathbf{P}(X_k \in \mathfrak{P}) = \mathbf{P}(X_k \in \mathfrak{P}, T \leq k) + \mathbf{P}(X_k \in \mathfrak{P}, T > k) \\
&\leq \sum_{j=0}^{k} \mathbf{P}(X_k \in \mathfrak{P}, T = j) + \mathbf{P}(T > k) \\
&\leq \sum_{j=0}^{k} \pi(\mathfrak{P})\mathbf{P}(T = j) + \mathbf{P}(T > k) \\
&\leq \pi(\mathfrak{P}) + \mathbf{P}(T > k)
\end{aligned}$$

où π est on le rappelle la loi uniforme sur $\mathfrak{S}_n$. En faisant pareil avec $\mathfrak{P}^c$, on trouve :

$$\forall \mathfrak{P} \subset \mathfrak{S}_n, \ \ |\mu_k(\mathfrak{P}) - \pi(\mathfrak{P})| \leq \mathbf{P}(T > k) \Longrightarrow d_v(\mu_n, \pi) \leq \mathbf{P}(T > k).$$

Il y a 2^k possibilités pour le nombre associé à une carte. Comme les $b^{(k)}(i)$ sont indépendants et uniformément distribués, la probabilité que tous les $b^{(k)}(i)$ soient différents au k-ème mélange vaut :

$$\frac{2^k(2^k-1)\dots(2^k-(n-1))}{(2^k)^n} = \prod_{i=1}^{n-1}\left(1 - \frac{i}{2^k}\right)$$

(c'est le "paradoxe des anniversaires"). Alors :

$$\mathbf{P}(T > k) = 1 - \prod_{i=1}^{n-1}\left(1 - \frac{i}{2^k}\right)$$

□

4. Interprétation : on attribue 0 aux cartes que l'on place au dessus, 1 aux autres. On garde en mémoire l'histoire de chaque carte. Un schéma s'imposerait, si ça ne demandait pas tant de prouesses techniques.

Remarquons, qu'à n fixé,

$$d_v(\mu_k, \pi) \underset{k\to+\infty}{\sim} \frac{n-1}{2^k}.$$

Mais comme ça va très très vite, il convient de s'interroger sur le sens de l'infini.

Quelques remarques complémentaires

Toutes les remarques intéressantes (et il y en a) sont à retrouver chez Grégoire Clarté (et dans un texte de modélisation option A sur le sujet).

Référence. M. Aigner, G. Ziegler, *Raisonnements divins*

Adapté d'un travail remarquable de l'inénarrable Grégoire Clarté.

105 Groupe des permutations d'un ensemble fini. Applications. *(heureusement que c'était une impasse)*
190 Méthodes combinatoires, problèmes de dénombrement.
262 Modes de convergence d'une suite de variables aléatoires. Exemples et applications.
264 Variables aléatoires discrètes. Exemples et applications.

3.3 Le fameux ellipsoïde de John et Loewner, assorti de quatre preuve de la log-concavité du déterminant (la quatrième va vous surprendre)

Le théorème qui suit, comme tout bon résultat d'analyse convexe, va servir à montrer l'unicité d'un extremum.

Théorème. *Soient $A, B \in \mathscr{S}_n^{++}(\mathbf{R})$ et $t \in [0,1]$. Alors :*

$$\det(tA + (1-t)B) \geq (\det A)^t (\det B)^{1-t}.$$

Avec égalité si et seulement si $A = B$.

On va donner quatre preuves (plus ou moins détaillées) de ce résultat.

La preuve algébriste. On invoque le théorème de peudo-réduction simultanée et on écrit :

$$A = {}^tPP \quad \text{et} \quad B = {}^tPDP \quad \text{où} \quad P \in \mathcal{O}_n(\mathbf{R}) \quad \text{et} \quad D = \operatorname{diag}(\lambda_1, \dots, \lambda_n).$$

Ensuite, on remplace :

$$(\det A)^t (\det B)^{1-t} = \det P^2 (\det D)^{1-t} \quad \text{et} \quad \det(tA + (1-t)B) = \det P^2 \det(tI_n + (1-t)D).$$

En prenant le logarithme, on doit montrer :

$$\sum_{i=1}^{n} \ln(t + (1-t)\lambda_i) \geq (1-t) \sum_{i=1}^{n} \ln \lambda_i.$$

Mais comme le logarithme est concave, c'est évident. □

La preuve astucieuse. On pose $C = B^{-1}A \in \mathscr{S}_n^{++}$ et on remplace :

$$\det(tA + (1-t)B) = \det B \det(tC + (1-t)I_n) \quad \text{et} \quad (\det A)^t (\det B)^{1-t} = \det B (\det C)^t$$

de sorte qu'il suffit de montrer :

$$\det(tC + (1-t)I_n) \geq (\det C)^t$$

ou encore :

$$\prod_{i=1}^{n} (tc_j + 1 - t) \geq \prod_{i=1}^{N} c_j^t$$

où les c_j sont les valeurs propres de C. Mais là encore, c'est évident car en prenant le logarithme, on voit que pour tout $c > 0$:

$$c^t \leq tc + (1-t)$$

avec égalité si et seulement si $t = 0$ ou $t = 1$. □

La preuve par le calcul différentiel. On étudie la convexité de la fonction $\log\det$ définie sur le convexe $\mathscr{S}_n^{++}$. Une caractérisation usuelle est la suivante :

$$\log\det \text{ est convexe } \Leftrightarrow \forall A, B \in \mathscr{S}_n^{++},\ f(t) = \log\det(tA + (1-t)B) \text{ est convexe.}$$

Et il n'y a plus qu'à différentier deux fois cette dernière fonction pour conclure. On trouve :

$$\frac{d^2 f}{dt^2}(t) = -\operatorname{Tr}\left((Y^{-1}\dot{Y})^2\right) \quad \text{où} \quad Y(t) = B + t(A-B).$$

Maintenant, on écrit que la trace du carré d'une matrice est la somme de ses valeurs propres au carré et on se débrouille pour justifier que $Y^{-1}\dot{Y}$ a des valeurs propres réelles. □

La preuve par le calcul intégral (by P. D. Lax). On commence par remarquer que si $H \in \mathscr{S}_n^{++}(\mathbf{R})$, alors en diagonalisant H en base othonormée et en effectuant le même changement de variable que dans la preuve du théorème de l'ellipsoïde de John et Loewner, on a :

$$\int_{\mathbf{R}^n} e^{-\langle x, Hx\rangle} dx = \frac{\pi^{n/2}}{\sqrt{\det H}}.$$

Il suffit d'appliquer cette jolie formule et l'inégalité de Hölder pour conclure :

$$\begin{aligned}
\frac{\pi^{n/2}}{\sqrt{\det(tA + (1-t)B)}} &= \int_{\mathbf{R}^n} e^{-t\langle x, Ax\rangle} e^{-(1-t)\langle x, Bx\rangle} dx \\
&\leq \left(\int_{\mathbf{R}^n} e^{-\langle x, Ax\rangle} dx\right)^t \left(\int_{\mathbf{R}^n} e^{-\langle x, Bx\rangle} dx\right)^{1-t} \\
&= \frac{\pi^{n/2}}{\sqrt{(\det A)^t (\det B)^{1-t}}}
\end{aligned}$$

Pour $t \in (0,1)$, il y a égalité dans l'inégalité de Hölder si et seulement si pour tout $x \in \mathbf{R}^n$, $e^{-\langle x,Ax\rangle}$ et $e^{-\langle x,Bx\rangle}$ sont colinéaires. En spécifiant en $x = 0$ on a égalité si et seulement si $A = B$. □

Théorème (John-Loewner)**.** *Soit K un compact d'intérieur non vide de $\mathbf{R}^n$. Il existe un unique ellipsoïde centré en 0 de volume minimal contenant K.*

Preuve. Lorsque $\mathbf{R}^n$ est muni de sa structure euclidienne usuelle, un ellipsoïde (plein) centré en 0 a une équation du type $q(x) \leq 1$ où q appartient à $\mathcal{Q}_{++}$ l'ensemble des formes quadratiques définies positives. On note dans ce cas :

$$\mathscr{E}_q = \{x \in \mathbf{R}^n,\ q(x) \leq 1\}.$$

Étape 1. Calculons sans complexe le volume d'un ellipsoïde et théorisons.

Soit $q \in \mathcal{Q}_{++}$ de matrice $S \in \mathscr{S}_n^{++}$ qui dans une base orthonormale s'écrit $q(x) = \sum_{i=1}^n a_i x_i^2$. N'ayons pas peur :

$$\mathscr{V}(\mathscr{E}_q) \ = \ \iint \dots \int_{a_1x_1^2+\dots a_nx_n^2\leq 1} dx_1 \dots dx_n = \iint \dots \int_{x_1^2+\dots x_n^2\leq 1} \frac{dx_1 \dots dx_n}{\sqrt{a_1 \dots a_n}} = \frac{V_0}{\sqrt{a_1 \dots a_n}}$$

où V_0 désigne le volume de la boule unité en dimension n. Voici une reformulation du problème :

Maximiser $D(q) := \det S = a_1 \dots a_n$ **sur le domaine** $\mathscr{A} = \{q \in \mathcal{Q}_+,\ \forall x \in K,\ q(x) \leq 1\}$.

Avant toute chose et pour traiter ce problème d'optimisation, il faut une norme sur $\mathcal{Q} \supset \mathscr{A}$ l'espace des formes quadratiques. On pose :

$$N(q) = \sup_{\|x\|\leq 1} |q(x)|.$$

Étape 2. De l'optimisation.

Il s'agit essentiellement de montrer que $\mathscr{A}$ est un compact non vide et on pourra conclure par continuité de D.

- $\mathscr{A}$ **est non vide.** Soit $M > 0$ tel que $K \subset B(0,M)$. Alors, on pose :

  $$\forall x \in \mathbf{R}^n,\ q(x) = \frac{\|x\|^2}{M^2}$$

  et on a bien $q \in \mathscr{A}$.
- $\mathscr{A}$ **est fermé.** Adoptons un point de vue séquentiel : soit $\mathscr{A} \ni q_n \to q$ pour la topologie de la norme N. Alors :

  $$\forall x \in \mathbf{R}^n,\ |q_n(x) - q(x)| \leq N(q - q_n)\|x\| \implies \forall x \in \mathbf{R}^n \lim_{n\to+\infty} q_n(x) = q(x).$$

- $\mathscr{A}$ **est borné.** Soient $a \in K$ et $r > 0$ tel que $B(a,r) \subset K$. Maintenant, pour $q \in \mathscr{A}$ on a pour tout $x \in \mathbf{R}^n$ tel que $\|x\| \leq r$:

  $$\sqrt{q(x)} = \sqrt{(q(a+x-a)} \leq \sqrt{q(a+x)} + \sqrt{q(-a)} = \sqrt{q(a+x)} + \sqrt{q(a)} \leq 2.$$

  Et par homogénéité, si $\|x\| \leq 1$:

  $$|q(x)| = q(x) = \frac{1}{r^2} q(rx) \leq \frac{4}{r^2}.$$

La fonction D est continue sur le compact $\mathscr{A}$ donc atteint son maximum en q_0 qui est définie positive.

Étape 3. Ne pas oublier l'unicité.

Pour l'unicité, on va montrer que $\mathscr{A}$ est convexe et utiliser la log-concavité du déterminant. Allons-y : soient $q, q' \in \mathscr{A}$ et $t \in (0,1)$. Pour tout $x \in \mathbf{R}^n$:

$$0 \leq (tq + (1-t)q')(x) \leq t + (1-t) = 1.$$

donc $tq+(1-t)q' \in \mathscr{A}$. Maintenant, s'il existait $q \in \mathscr{A}$ distincte de q_0 telle que $D(q) = D(q_0)$, on écrirait :

$$D\left(\frac{1}{2}(q+q_0)\right) = \det\left(\frac{1}{2}(S+S_0)\right) > (\det S)^{1/2}(\det(S_0))^{1/2} \geq \det S_0 = D(q_0)$$

et ce serait bien sûr contradictoire avec la maximalité de $D(q_0)$. □

Une remarque complémentaire.

Il existe au moins une application de ce résultat. Remarquons d'abord qu'il se généralise sans mal à un espace euclidien E quelconque quitte à le munir d'une structure qui le rend isomorphe à $\mathbf{R}^n$ pour un certain n. Lorsqu'on a vu cela, on est prêt à montrer sans utiliser le théorème de pont fixe de Kakutani qu'un sous groupe compact de $GL(E)$ est toujours contenu dans le groupe des isométries d'une certaine forme quadratique (d'après M. Alessandri).

Le livre dédié aux Équations aux Dérivées Partielles de F. John est très bien.

Références.
S. Francinou, H. Gianella, S. Nicolas, *Oraux X-ENS. Algèbre 3.*
P. D. Lax, *Linear Algebra and Its Applications. Second Edition.*
M. Alessandri, *Thèmes de Géométrie.*

152 Déterminant. Exemples et applications.
170 Formes quadratiques sur un espace vectoriel de dimension finie. Orthogonalité, isotropie. Applications.
171 Formes quadratiques réelles. Coniques. Exemples et applications.
203 Utilisation de la notion de compacité.
219 Extrema : existence, caractérisation, recherche. Exemples et applications.
253 Utilisation de la notion de convexité en analyse.

3.4 L'exponentielle matricielle est surjective

Théorème. *Soit $A \in \mathcal{M}_n(\mathbf{C})$. Alors,* $\exp(\mathbf{C}[A]) = \mathbf{C}[A] \cap GL_n(\mathbf{C})$. *En particulier,* $\exp : \mathcal{M}_n(\mathbf{C}) \to GL_n(\mathbf{C})$ *est surjective et un antécédent de $A \in GL_n(\mathbf{C})$ est un polynôme (complexe) en A.*

Preuve. *Étape 1 : quelques résultats préliminaires.*

- On commence par observer l'égalité $\mathbf{C}[A]^\times = \mathbf{C}[A] \cap GL_n(\mathbf{C})$ où $\mathbf{C}[A]^\times$ est le groupe des inversibles de $\mathbf{C}[A]$. Seule l'inclusion $\supset$ pose question : il s'agit de voir que l'inverse d'une matrice M est un polynôme en M (en effet le coefficient constant de son polynôme minimal est non nul : $\mu_M = \alpha + XP$ et $M^{-1} = -P(M)/\alpha$). Ainsi, l'inverse d'un élément de $\mathbf{C}[A] \cap GL_n(\mathbf{C})$ reste dans $\mathbf{C}[A]$ (c'est un polynôme de polynôme en A)
- Pour tout $M \in \mathcal{M}_n(\mathbf{C})$, $\exp(M) \in \mathbf{C}[M]$: en effet, c'est une limite dans $\mathcal{M}_n(\mathbf{C})$ (pour la norme d'algèbre) d'éléments de $\mathbf{C}[M]$ qui est un sous-espace vectoriel de dimension finie donc fermé. En conséquence,

  $$\exp : \mathbf{C}[A] \to \mathbf{C}[A]^\times$$

  est un morphisme de groupes.
- $\mathbf{C}[A]^\times = \mathbf{C}[A] \cap \det^{-1}(\mathbf{R}^*)$ est un ouvert de $\mathbf{C}[A]$. Il est aussi connexe par arcs (donc connexe) car si $M, N \in \mathbf{C}[A]^\times$, la fonction

  $$z \in \mathbf{C} \mapsto \det(zM + (1-z)N)$$

  est polynomiale en z et non nulle donc admet un nombre fini de zéros. 0 et 1 ne sont pas des zéros de ce polynôme donc on peut construire une courbe $z(t) \in \mathbf{C}$ reliant 0 et 1 en évitant ces zéros [5]. Ainsi

  $$t \in [0,1] \mapsto z(t)M + (1 - z(t))N$$

  est une courbe tracée dans $\mathbf{C}[A]^\times$ reliant continûment N et M.

Étape 2 : exp *est localement un difféomorphisme*

Comme la différentielle de $\exp : \mathcal{M}_n(\mathbf{C}) \to GL_n(\mathbf{C})$ en 0 est l'identité de $\mathcal{M}_n(\mathbf{C})$, on a aussi en restreignant $\exp : \mathbf{C}[A] \to \mathbf{C}[A]^\times$:

$$d\exp(0) = id_{\mathbf{C}[A]}.$$

En particulier cette différentielle est bijective et le théorème d'inversion locale assure l'existence de deux ouverts $\mathcal{U} \subset \mathbf{C}[A]$ et $\mathcal{V} \subset \mathbf{C}[A]^\times$ contenant respectivement 0 et Id tel que $\exp : \mathcal{U} \to \mathcal{V}$ soit un difféomorphisme. Comme exp est un morphisme de groupes, le résultat demeure au voisinage de chaque point $M \in \mathbf{C}[A]$:

$$\exp : M + \mathcal{U} \to \exp(M)\mathcal{V}$$

est un difféomorphisme.

Étape 3 : un argument de connexité pour conclure.

L'étape 2 implique en fait que $\exp(\mathbf{C}[A])$ est un ouvert de $\mathbf{C}[A]^\times$. Mais c'est aussi un fermé en remarquant que

$$\mathbf{C}[A]^\times \setminus \exp(\mathbf{C}[A]) = \bigcup_{M \in \mathbf{C}[A]^\times \setminus \exp(\mathbf{C}[A])} M\exp(\mathbf{C}[A])$$

(l'inclusion $\supset$ se prouve par contraposée). En vertu de la connexité de $\mathbf{C}[A]^\times$, on conclut que

$$\exp(\mathbf{C}[A]) = \mathbf{C}[A]^\times = \mathbf{C}[A] \cap GL_n(\mathbf{C}).$$

□

5. On montre même que $\mathbf{R}^2 \setminus D$ où D est dénombrable est connexe par arcs

Une application

Proposition. *L'image par l'application exponentielle de $\mathcal{M}_n(\mathbf{R})$ est l'ensemble*

$$\exp(\mathcal{M}_n(\mathbf{R})) = \{A^2,\ \ A \in GL_n(\mathbf{R})\}.$$

Preuve.
$\subset$: Il suffit de remarquer que $\exp(M) = \exp(\frac{1}{2}M)^2$
$\supset$: Soit $M = A^2$ où $A \in GL_n(\mathbf{R})$. Il existe un polynôme $P \in \mathbf{C}[X]$ tel que $A = \exp(P(A))$. Comme A est réelle, on a aussi $\exp(\overline{P}(A)) = \overline{A} = A$ et donc

$$\exp((P + \overline{P})(A)) = A^2 = M.$$

□

Référence. M. Zavidovique, *Un max de maths*

156 Exponentielle de matrices. Applications.
204 Connexité. Exemples et applications.

3.5 Le théorème de Cartan - Von Neumann

Théorème (Cartan-von Neumann)**.** *Tout sous-groupe fermé de $GL_n(\mathbf{R})$ est une sous-variété de $\mathcal{M}_n(\mathbf{R})$.*

Preuve. Soit G un sous-groupe fermé de $GL_n(\mathbf{R})$. On montre que G est localement difféomorphe à un ouvert de $\mathcal{M}_n(\mathbf{R}) \simeq \mathbf{R}^{n^2}$. À cause de la structure de groupe et du caractère C^∞ de la translation, il suffit de la prouver pour un voisinage de I_n dans G.

Étape 1. Une sous-algèbre digne d'intérêt

On pose :

$$\mathcal{L}_G := \{m \in \mathcal{M}_n(\mathbf{R}),\ \forall t \in \mathbf{R},\ e^{tm} \in G\}$$

et on montre que c'est une sous-algèbre de $\mathcal{M}_n(\mathbf{R})$. La seule chose à vérifier est la stabilité par la somme. Pour cela, remarquons d'abord que la différentielle de exp en 0 est inversible donc, exp est localement inversible et on note L son inverse. Pour m au voisinage de 0 :

$$e^m = I_n + m + o(\|m\|) \ \text{ et } \ L(I_n + m) = m + o(\|m\|).$$

Soient $a, b \in \mathcal{L}_G$ et $t \in \mathbf{R}$, on a pour $k \in \mathbf{N}$ suffisamment grand :

$$\begin{array}{rcl} (e^{ta/k}e^{tb/k})^k & = & e^{kL(e^{ta/k}e^{tb/k})} \\ & = & e^{kL(I_n + t\frac{a+b}{k} + o(1/k))} \\ & = & e^{t(a+b)+o(1)} \\ & \underset{k\to+\infty}{\longrightarrow} & e^{t(a+b)} \end{array}$$

et la fermeture de G montre que $a + b \in \mathcal{L}_G$.

Étape 2. Là où le théorème d'inversion locale intervient de façon cruciale.

Soient M un supplémentaire de $\mathcal{L}_G$ dans $\mathcal{M}_n(\mathbf{R})$ et $\varphi : \mathcal{M}_n(\mathbf{R}) \to GL_n(\mathbf{R})$ l'application qui a $x = l+m$, $(l,m) \in \mathcal{L}_G \times M$, associe $\varphi(x) = e^l e^m$. φ est C^∞ et sa différentielle en 0 vaut l'identité. Par le théorème d'inversion locale, il existe un voisinage U de 0 dans $\mathcal{M}_n(\mathbf{R})$ tel que φ soit un C^1-difféomorphisme de U sur $\varphi(U)$. On va montrer que, quitte à restreindre U,

$$\varphi : U \cap \mathcal{L}_G \to \varphi(U) \cap G$$

est un C^1-difféomorphisme, ce qui conclura (en on aura aussi $\dim G = \dim \mathcal{L}_G$). Il suffit pour cela, de montrer que $\varphi(U \cap \mathcal{L}_G) = \varphi(U) \cap G$. Remarquons que l'inclusion $\subset$ est évidente : c'est la définition de $\mathcal{L}_G$.

Étape 3. L'autre inclusion et c'est fini.

On veut montrer qu'il existe un voisinage U de 0 dans $\mathcal{M}_n(\mathbf{R})$ tel que $\varphi(U) \cap G \subset \varphi(U \cap \mathcal{L}_G)$. On va montrer :

$$\exists k \in \mathbf{N}^*, \forall x_k \in B(0, 1/k), \varphi(x_k) \in G \Longrightarrow x_k \in \mathcal{L}_G.$$

Si ça n'était pas le cas, alors comme $\mathcal{M}_n(\mathbf{R}) = \mathcal{L}_G \oplus M$, il existerait deux suites $(l_k)_k$ et $(m_k)_k$ respectivement de $\mathcal{L}_G$ et $M \setminus \{0\}$, de limites nulles, et telles que pour tout $k \in \mathbf{N}$, $\varphi(l_k + m_k)$ soit dans G. Alors, par définition de $\mathcal{L}_G$, on aurait $e^{m_k} \in G$ pour tout k. Puisque $m_k \neq 0$ posons, pour tout $k \in \mathbf{N}^*$:

$$\varepsilon_k = \frac{m_k}{\|m_k\|} \in M.$$

Quitte à extraire, on peut supposer que (ε_k) converge vers $\varepsilon \in M$ de norme 1. Soit $t \in \mathbf{R}$ et notons :

$$\frac{t}{\|m_k\|} = \lambda_k + \mu_k, \quad \text{où } \lambda_k \in \mathbf{Z} \text{ et } |\mu_k| < \frac{1}{2}.$$

Alors :

$$e^{t\varepsilon} = \lim_{k \to +\infty} e^{t \frac{m_k}{\|m_k\|}} = \lim_{k \to +\infty} e^{\lambda_k m_k}$$

car $e^{\mu_k m_k} \to I_n$. Comme $\lambda_k \in \mathbf{Z}$, $e^{t\varepsilon} \in G$ comme limite d'une suite de points de G qui est fermé. Ce qui prouve que $\varepsilon \in \mathcal{L}_G \cap M = \{0\}$. C'est absurde puisque ε est de norme 1. □

Quelques remarques complémentaires

- En fait, $\mathcal{L}_G$ est une sous-algèbre de Lie de $\mathcal{M}_n(\mathbf{R})$, c'est à dire un sous-espace de $\mathcal{M}_n(\mathbf{R})$ stable par $(a,b) \mapsto [a,b] = ab - ba$. Pour le voir, il suffit de montrer en développant à l'ordre 2 :
$$(e^{a/k} e^{b/k} e^{-a/k} e^{-b/k})^{k^2} \to e^{ab-ba}.$$
- Il vaut mieux définir le "logarithme matriciel" :
$$L(I_n + m) = \sum_{k=0}^{+\infty} \frac{(-1)^k m^{k+1}}{k+1}, \quad \|m\| < 1$$
qui vérifie $e^{L(1+m)} = 1 + m$.

- On peut voir que G est discret si et seulement si $\mathcal{L}_G = \{0\}$. Pour le sens réciproque, on montre en réutilisant le résultat sur les suites de l'étape 3 que I_n est isolé, ce qui est suffisant.
- Soit $m \in \mathcal{L}_G$. Alors $t \mapsto e^{tm}$ est une courbe tracée dans G passant par I_n en $t = 0$ et le vecteur tangent à cette courbe en I_n est m. Par suite, $\mathcal{L}_G$ est inclus dans l'espace tangent à G en Id. Mais on a prouvé qu'ils ont même dimension, ils sont donc égaux.
- Le sous-groupe de G engendré par $\exp \mathcal{L}_G$ est la composante connexe de I_n dans G.
- En pratique, ça ne sert pas à grand chose.

Références.
S. Gonnord, N. Tosel, *Calcul différentiel*
R. Mneimné, F. Testard, *Introduction aux groupes de Lie classiques*

106 Groupe linéaire d'un espace vectoriel de dimension fini E, sous-groupe de $GL(E)$. Applications.
156 Exponentielle de matrices. Applications.
214 Théorème d'inversion locale, théorème des fonctions implicites. Exemples et applications en analyse et en géométrie
215 Applications différentiables sur un ouvert de $\mathbf{R}^n$. Exemples et applications.

3.6 Le théorème de Krein-Milmann

Ne faire que le dimension finie (sans blague) en admettant tout ce qu'il y a jusqu'à l'existence d'un hyperplan d'appui. La partie sur le théorème de Birkhoff est mieux rédigée dans le livre de P. D. Lax, en plus il y a un dessin.

Bien que les arguments soient les mêmes, il convient de distinguer le cas de la dimension finie et celui de la dimension infinie. Notons cependant que tout repose sur la propriété de séparation stricte des convexes compacts qui est facile en dimension finie mais qui, en dimension infinie, porte le nom de théorème de Hahn-Banach géométrique et la marque du lemme de Zorn. Commençons par énoncer quelques définitions et propriétés indépendantes de ces considérations dimensionnelles.

Deux définitions.

En toute généralité, E est ici un espace de Banach.

Définition 16 (Hyperplan affine)**.** On appelle hyperplan affine associé à $\Lambda \in E' \setminus \{0\}$ et $\alpha \in \mathbf{R}$ l'ensemble :

$$H_{\Lambda,\alpha} := \{x \in E,\ \langle \Lambda, x\rangle = \alpha\}.$$

$H_{\Lambda,\alpha}$ peut évidemment s'écrire comme l'image de l'hyperplan vectoriel $H_{\Lambda,0}$ par une translation.

En dimension finie, $E' \simeq E$ et Λ s'appellera $s \in E$.

3. Non exclusifs

Définition 17 (Point extrémal)**.** Soit C un convexe non vide de E. Un point $x \in C$ est dit extrémal lorsque pour tout triplet $x_1, x_2 \in C$ et $t \in (0,1)$,

$$x = (1-t)x_1 + tx_2 \implies x = x_1 = x_2.$$

On note $\mathcal{E}(C)$ l'ensemble des points extrémaux de C.

Proposition 18. *Si C est compact, alors $\mathcal{E}(C) \neq \emptyset$.*

Preuve. Il suffit de considérer $\bar{x} \in C$ qui réalise le maximum de $x \mapsto \|x\|$ sur C. □

En dimension finie.

Ici, E est un espace affine réel de dimension finie d.

Théorème 19 (Séparation)**.** *Soient C un convexe fermé non vide de E et $x \notin E$. Alors il existe un hyperplan affine qui sépare strictement x et C, c'est à dire :*

$$\exists s \in E \setminus \{0\},\ \exists \alpha \in \mathbf{R},\ \forall c \in C,\ \langle s, c\rangle < \alpha < \langle s, x\rangle$$

ou de façon équivalente

$$\exists s \in E \setminus \{0\},\ \ \langle s, x\rangle > \sup\big\{\langle s, c\rangle,\ c \in C\big\}.$$

Preuve. Soit $s = x - p_C(x) \neq 0$. Alors la condition :

$$\forall c \in C,\ \langle s, c - p_C(x)\rangle \leq 0$$

et équivalente à :

$$\forall c \in C,\ \ \langle s, c\rangle \leq \langle s, x - s\rangle = \langle s, x\rangle - \|s\|^2.$$

□

Raffinons un tout petit peu la propriété de séparation.

Définition 20 (Hyperplan d'appui)**.** Un hyperplan affine $H_{s,\alpha}$ est appelé hyperplan d'appui à un convexe C lorsque C est contenu dans un demi-plan affine de frontière $H_{s,\alpha}$, c'est à dire :

$$[\forall c \in C,\ \langle s, c\rangle \geq \alpha] \ \text{ ou } \ [\forall c \in C,\ \langle s, c\rangle \leq \alpha].$$

S'il existe un point $x \in H_{s,\alpha} \cap \partial C$, on dit que $H_{s,\alpha}$ est un hyperplan d'appui en x.

Proposition 21 (Existence d'un hyperplan d'appui)**.** *Soit C un convexe non vide de E. Alors pour tout $x \in \partial C$, il existe un hyperplan d'appui à C en x.*

Preuve. Soient $x \in \partial C$ et $(x_k)_k$ une suite de $E \setminus \overline{C}$ qui converge vers x. Le théorème de séparation assure l'existence d'une suite $(s_k)_k$ de vecteurs de $E \setminus \{0\}$ tels que :

$$\forall k \in \mathbf{N},\ \forall c \in C,\ \ \langle s_k, x_k - c\rangle > 0.$$

Quitte à supposer $\|s_k\| = 1$ pour tout $k \in N$ et à extraire une sous-suite convergente, on peut supposer $s_k \to s \in E \setminus \{0\}$. Alors :

$$\forall c \in C,\ \ \langle s_k, x_k - x\rangle \to \langle s, x - c\rangle \geq 0.$$

Le résultat suit avec $\alpha := \langle s, x\rangle$. □

3. Non exclusifs

La preuve du théorème de Krein-Milman en dimension finie consiste à récurrer sur la dimension, l'argument principal résidera alors dans le :

Lemme 22. *Soient C un compact convexe de E et $H_{s,\alpha}$ un hyperplan d'appui à C en un point $x \in C$. Alors x est un point extrémal de C si et seulement si x est un point extrémal de $H_{s,\alpha} \cap C$.*

Preuve. Si x est un point extremal de $H_{s,\alpha} \cap C$, alors soient $x_1, x_2 \in C$ et $t \in (0,1)$ tels que $x = (1-t)x_1 + tx_2$. On a par définition de $H_{s,\alpha}$ (quitte à changer α en $-\alpha$) :

$$\langle s, x_i \rangle \leq \alpha, \quad i = 1, 2.$$

Finalement,

$$\alpha = \langle s, x \rangle = (1-t)\langle s, x_1 \rangle + t\langle s, x_2 \rangle \leq (1-t)\alpha + t\alpha = \alpha.$$

Toutes les inégalités sont des égalités donc $x_1, x_2 \in H_{s,\alpha}$ et le résultat suit. □

On peut enfin énoncer (et prouver) le :

Théorème 23 (Krein-Milman)**.** *Soit C un compact convexe non vide de E. Alors C est l'enveloppe convexe de ses points extrémaux : $C = \mathrm{co}\left(\mathcal{E}(C)\right)$.*

Preuve. On raisonne par récurrence sur la dimension de l'espace. En dimension 0, C est réduit à un point et c'est évident. Supposons donc le résultat vrai en dimension $d-1$ avec $d \geq 1$. Soit $x \in C$, deux cas se présentent :

- Si $x \in \partial C$, alors en considérant H un hyperplan d'appui à C en x, on sait par le lemme précédent que x est un point extrémal de $H \cap C$ qui est un convexe compact d'un espace affine de dimension $d-1$. Par hypothèse de récurrence, x est dans l'enveloppe convexe des points extrémaux de $H \cap C$ qui sont aussi des points extrémaux de C, toujours par le lemme.
- Si $c \in C \setminus \partial C$, alors soit $x' \neq x$ dans C (c'est possible en dimension ≥ 1). La droite affine passant par x et x' coupe la frontière de C en (au moins) deux points. Ces deux points sont dans $\mathrm{co}\left(\mathcal{E}(C)\right)$ d'après le premier cas. C'est fini.

□

En voici une application :

Théorème 24 (Birkhoff)**.** *L'ensemble des matrices bistochastiques $\mathcal{B}_n$ est l'enveloppe convexe dans $\mathcal{M}_n(\mathbf{R})$ de l'ensemble des matrices de permutation.*

Preuve. En vertu du théorème de Krein-Milman, il suffit de montrer que les matrices de permutation sont les points extrémaux de l'ensemble des matrices bistochastiques. D'abord, les matrices de permutation sont des points extrémaux de $\mathcal{B}_n$. En effet, si $P = (1-\alpha)M + \alpha N$ avec $\alpha \in (0,1)$ et $M, N \in \mathcal{B}_n$, alors

- Si $p_{i,j} = 0$, comme les coefficients de M et N sont positifs, on a $m_{i,j} = n_{i,j} = 0$.
- Si $p_{i,j} = 1$, comme les coefficients de M et N sont ≤ 1, on a $m_{i,j} = n_{i,j} = 1$.

Montrons donc qu'il n'y a pas d'autres points extrémaux. Soit $M \in \mathcal{B}_n$ qui n'est pas une matrice de permutation. Alors, il existe $m_{i_1,j_1} \in (0,1)$ et comme M est une matrice stochastique, il existe aussi $m_{i_1,j_2} \in (0,1)$ avec $j_2 \neq j_1$. Mais comme M^T est aussi stochastique, il

existe aussi $m_{i_2,j_2} \in (0,1)$ avec $i_2 \neq i_1$. On construit comme cela une suite d'indices (i_1, j_1), (i_2, j_2), ... avec la propriété :

$$m_{i_k,j_k} \in (0,1) \quad \text{et} \quad m_{i_{k-1},j_k} \in (0,1).$$

Puisqu'il n'y a que n indices possibles, il existe $r \in \mathbf{N}^*$ tel que $j_{r+1} = j_1$. On définit la matrice $B \in \mathcal{M}_n(\mathbf{R})$ par :

$$b_{i_k,j_k} = 1, \quad b_{i_k,j_{k+1}} = -1 \quad \text{et} \quad b_{i,j} = 0 \quad \text{sinon.}$$

En notant $e = (1 \dots 1)^T$, on a par construction $Be = 0$ et $B^T e = 0$. Ainsi, pour tout $\alpha > 0$ suffisamment petit, les matrices $M \pm \alpha B$ sont bistochastiques. Comme M est le milieu du segment $[M - \alpha B, M + \alpha B]$, M n'est pas un point extrémal. □

En dimension infinie.

Ici, E redevient un espace de Banach dans toute sa généralité. Toutes les propriétés de séparation sont remplacée par le :

Théorème 25 (Hahn-Banach géométrique)**.** *Soient C_1 et C_2 deux convexes non vides de E. On suppose que C_1 est fermé et que C_2 est compact. Alors il existe un hyperplan fermé qui sépare C_1 et C_2 au sens strict :*

$$\exists \mu \in E' \setminus \{0\}, \quad \sup_{x \in C_1} \langle \Lambda, x \rangle < \inf_{y \in C_2} \langle \Lambda, y \rangle.$$

Corollaire 26. *Le dual topologique de E sépare les points : si $x \neq y$ sont deux points de E, il existe $\mu \in E'$ tel que $\langle \Lambda, x \rangle \neq \langle \Lambda, y \rangle$.*

Et le théorème prend la forme suivante :

Théorème 27 (Krein-Milman)**.** *Soit C un compact convexe de E. Alors C est l'enveloppe convexe fermée de ses points extrémaux :*

$$C = \overline{\mathrm{co}}\big(\mathcal{E}(C)\big).$$

On appelle *ensemble extrémal de C* tout sous-ensemble convexe $S \subset C$ dont aucun point n'est le barycentre de deux points dans $K \setminus C$.

Preuve. Le lemme de Zorn est crucial, non seulement car on utilise le théorème de Hahn-Banach mais aussi dans la première étape.

Étape 1. Tout ensemble extrémal de C contient un point extrémal.

Notons $\mathscr{P}$ l'ensemble de tous les ensembles extrémaux compacts de C. Comme $C \in \mathscr{P}$, cet ensemble est non vide. On utilisera les propriétés suivantes :

(*i*) Toute intersection non vide d'éléments de $\mathscr{P}$ est encore dans $\mathscr{P}$.

(*ii*) Si $S \in \mathscr{P}$, $\Lambda \in E'$ et $\mu \in \mathbf{R}$ le maximum de Λ sur S, alors :

$$S_\Lambda := \{x \in S, \ \langle \Lambda, x \rangle = \mu\}$$

est un élément de $\mathscr{P}$ (pour le voir, il suffit s'appliquer Λ à $(1-t)x + ty = z \in S_\Lambda$).

3. Non exclusifs

Soient $S \in \mathscr{P}$ et $\mathscr{P}' \neq \emptyset$ l'ensemble des éléments de $\mathscr{P}$ qui sont des sous-ensembles de S. On va appliquer le lemme de Zorn pour justifier l'existence d'un élément minimal de $\mathscr{P}'$. On montrera ensuite que cet élément est un singleton.

D'abord, $\mathscr{P}'$ est partiellement ordonné par l'inclusion et si Ω est une famille totalement ordonnée de $\mathscr{P}'$, l'intersection de tous les éléments de Ω est non vide comme intersection de compacts emboîtés et est dans $\mathscr{P}$ par la propriété (i). Bien sûr on a aussi $M \in \mathscr{P}'$ et c'est un minorant de Ω. Par suite, $\mathscr{P}'$ est un ensemble inductif et par le lemme de Zorn, il possède un élément minimal M.

Puisque M n'a pas de sous-ensemble propre, la propriété (ii) entraîne que pour tout $\Lambda \in E'$, $M_\Lambda = M$. Autrement dit, toute forme linéaire continue est constante sur M. Mais comme E' sépare les points, M est nécessairement un singleton : $M = \{x\} \subset S$ et $x \in S$ est un point extrémal. On vient de montrer que :

$$\forall S \in \mathscr{P}, \quad \mathcal{E}(C) \cap S \neq \emptyset. \tag{3.1}$$

Étape 2. Conclusion.

Puisque C est compact et convexe, on a déjà :

$$\overline{\mathrm{co}}\big(\mathcal{E}(C)\big) \subset C.$$

Par l'absurde, supposons qu'il existe $x_0 \in C \setminus \overline{\mathrm{co}}\big(\mathcal{E}(C)\big)$. Le théorème de Hahn-Banach assure l'existence de $\Lambda \in E'$ tel que :

$$\forall x \in \overline{\mathrm{co}}\big(\mathcal{E}(C)\big), \quad \langle \Lambda, x\rangle < \langle \Lambda, x_0\rangle.$$

Avec les notations précédentes, on a $C_\Lambda \in \mathscr{P}$ et

$$\overline{\mathrm{co}}\big(\mathcal{E}(C)\big) \cap C_\Lambda = \emptyset.$$

C'est contradictoire avec (3.1). □

Références.

J.-B. Hiriart-Urruty, C. Lemaréchal, *Fundamentals of Convex Analysis*
P. D. Lax, *Linear Algebra*
P. D. Lax, *Functional Analysis*
H. Brézis, *Functional Analysis*
W. Rudin, *Functional Analysis*

159 Formes linéaires et dualité en dimension finie. Exemples et applications.
181 Barycentre dans un espace affine réel de dimension finie, convexité. Applications
253 Utilisation de la notion de convexité en analyse.

3.7 Méthodes de gradient

Introduction générale.

Il s'agit de résoudre un système linéaire du type :

$$Ax = b \tag{3.2}$$

où, *a priori*, $A \in GL_n(\mathbf{R})$ et $b \in \mathbf{R}^n$. En fait, on prendra $A \in S_n^{++}(\mathbf{R})$. Un problème équivalent consiste à trouver le point qui minimise la fonctionnelle :

$$\Phi(y) = \frac{1}{2} y^T A y - y^T b.$$

En effet, il est facile de voir que

$$\nabla\Phi(y) = \frac{1}{2}(A^T + A)y - b = Ay - b. \tag{3.3}$$

Et si x est solution du système linéaire, alors

$$\Phi(y) = \Phi(x + (y - x)) = \Phi(x) + \frac{1}{2}(y - x)^T A(y - x), \quad \text{i.e} \quad \frac{1}{2}\|y - x\|_A^2 = \Phi(y) - \Phi(x)$$

où $\|z\|_A^2 = z^T A z$ est la norme d'énergie que l'on utilisera toujours par la suite. Une *méthode de gradient* consiste à partir d'un point $x_0 \in \mathbf{R}^n$ et à construire la suite

$$x_{k+1} = x_k + \alpha_k d_k \tag{3.4}$$

où $d_k \in \mathbf{R}^n$ est une direction à choisir et $\alpha_k \in \mathbf{R}$. Une idée naturelle est de choisir α_k de sorte à optimiser $\Phi(x_{k+1})$ dans la direction d_k, c'est à dire tel que :

$$\frac{d}{d\alpha_k}\Phi(x_k + \alpha_k d_k) = -d_k^T r_k + \alpha_k d_k^T A d_k = 0$$

où $-r_k := \nabla\Phi(x_k) = Ax_k - b$. On trouve :

$$\alpha_k = \frac{\langle d_k, r_k\rangle}{\|d_k\|_A^2} \tag{3.5}$$

(c'est bien défini lorsque $d_k \neq 0$ car $A \in S_n^{++}(\mathbf{R})$).

Théorème. *Soit x la solution du système* (3.2) *ou de façon équivalente, la solution du problème de minimisation* (3.3). *Si α_k est choisi comme dans* (3.5), *alors la suite* (3.4) *vérifie :*

$$\|x_{k+1} - x\|_A^2 = (1 - \sigma_k)\|x_k - x\|_A^2$$

où

$$\sigma_k = \frac{\langle d_k, r_k\rangle^2}{\|d_k\|_A^2 \|r_k\|_{A^{-1}}^2} \in (0, 1].$$

Preuve. Il suffit de calculer :

$$\begin{aligned} \|x_{k+1} - x\|_A^2 &= \|x_k - x + \alpha_k d_k\|_A^2 \\ &= \|x_k - x\|_A^2 + \alpha_k^2\|d_k\|_A^2 + 2\alpha_k\langle d_k, A(x_k - x)\rangle \\ &= \|x_k - x\|_A^2 + \alpha_k^2\|d_k\|_A^2 - 2\alpha_k\langle d_k, r_k\rangle \end{aligned}$$

car $A(x_k - x) = Ax_k - b = -r_k$ et $\|x_k - x\|_A^2 = \|r_k\|_{A^{-1}}^2$. Et en remplaçant α_k par son expression :

$$\|x_{k+1} - x\|_A^2 = \left(1 - \frac{\langle d_k, r_k\rangle^2}{\|d_k\|_A^2 \|r_k\|_{A^{-1}}^2}\right) \|x_k - x\|_A^2.$$

□

Méthode de gradient à pas optimal.

On choisit pour direction la "plus grande pente", autrement dit :

$$d_k = -\nabla\Phi(x_k) = -Ax_k + b = r_k.$$

Dans ce cas, $d_k \neq 0$ tant qu'on a pas atteint la solution et la convergence découle du théorème et de inégalité de Kantorovich [6] :

Lemme (Inégalité de Kantorovich)**.** *En notant $0 < \lambda_1 \leq \ldots \leq \lambda_n$ les valeurs propres de A, on a pour tout $y \in \mathbf{R}^n$,*

$$\frac{\|y\|^4}{\|y\|_A^2 \|y\|_{A^{-1}}^2} \geq \frac{4\lambda_n\lambda_1}{(\lambda_n + \lambda_1)^2}.$$

Preuve. On va montrer l'inégalité équivalente :

$$\forall y \in \mathbf{R}^n, \ \ \|y\|^4 \leq \frac{1}{4}\left(\sqrt{\frac{\lambda_n}{\lambda_1}} + \sqrt{\frac{\lambda_1}{\lambda_n}}\right)^2.$$

On va même supposer que $\|y\| = 1$ et commencer par remarquer :

$$1 = \|y\|^2 = \langle y, AA^{-1}y\rangle \leq \|y\|_A \|A^{-1}y\|_A = \|y\|_A \|y\|_{A^{-1}}.$$

Et dans une base orthonormale de vecteurs propres :

$$\begin{aligned}
\|y\|_A\|y\|_{A^{-1}} &= \sqrt{\left(\sum_{i=1}^n \lambda_i y_i^2\right)\left(\sum_{i=1}^n \frac{1}{\lambda_i} y_i^2\right)} = \sqrt{\frac{\lambda_1}{\lambda_n}\left(\sum_{i=1}^n \frac{\lambda_i}{\lambda_1} y_i^2\right)\left(\sum_{i=1}^n \frac{\lambda_n}{\lambda_i} y_i^2\right)} \\
&\leq \frac{1}{2}\sqrt{\frac{\lambda_1}{\lambda_n}}\left(\left(\sum_{i=1}^n \frac{\lambda_i}{\lambda_1} y_i^2\right) + \left(\sum_{i=1}^n \frac{\lambda_n}{\lambda_i} y_i^2\right)\right) \\
&\leq \frac{1}{2}\sqrt{\frac{\lambda_1}{\lambda_n}}\left(\sum_{i=1}^n \left(\frac{\lambda_i}{\lambda_1} + \frac{\lambda_n}{\lambda_i}\right) y_i^2\right)
\end{aligned}$$

La fonction $x \mapsto \frac{x}{\lambda_1} + \frac{\lambda_n}{x}$ admet un maximum en λ_1 ou en λ_n et il vaut dans les deux cas : $1 + \frac{\lambda_n}{\lambda_1}$. Ainsi,

$$\|y\|_A\|y\|_{A^{-1}} \leq \frac{1}{2}\sqrt{\frac{\lambda_1}{\lambda_n}}\left(\sum_{i=1}^n \left(1 + \frac{\lambda_n}{\lambda_1}\right) y_i^2\right) \leq \frac{1}{2}\left(\sqrt{\frac{\lambda_n}{\lambda_1}} + \sqrt{\frac{\lambda_1}{\lambda_n}}\right)$$

et le résultat suit en élevant au carré. □

6. Comme c'est une inégalité de convexité, on peut la développer dans les leçons qui leur sont dévolues mais en fait, on n'en a pas besoin pour conclure : on peut obtenir une majoration de l'erreur (un peu différente mais pas pire) beaucoup plus rapidement et beaucoup plus simplement comme le fait P. D. Lax dans son *Linear Algebra* (dans la deuxième édition) !

Et sachant que $\mathrm{cond}(A) = \lambda_n/\lambda_1$, on obtient le :

Théorème. *Avec les choix précédents et $d_k = r_k$, la suite* (3.4) *converge vers x avec :*

$$\|x_k - x\|_A \leq \frac{\lambda_n - \lambda_1}{\lambda_n + \lambda_1}\|x_k - x\|_A.$$

Un calcul supplémentaire donne :

$$\|x_k - x\| \leq \sqrt{\mathrm{cond}(A)}\left(\frac{\mathrm{cond}(A) - 1}{\mathrm{cond}(A) + 1}\right)^k \|x_0 - x\|.$$

Preuve. La première inégalité découle directement de l'inégalité de Kantorovich. Pour la seconde, il s'agit de voir que pour tout $y \in \mathbf{R}^n$,

$$\lambda_1\|y\|^2 \leq \|y\|_A^2 \leq \lambda_n\|y\|^2.$$

□

De la dernière inégalité, on voit que la convergente peut être lente lorsque la matrice est mal conditionnée.

Méthode de gradient conjugué.

Remarquons que pour tout $k \in \mathbf{N}$:

$$r_{k+1} = r_k - \alpha_k A d_k \tag{3.6}$$

et α_k est choisi de sorte à ce que

$$\langle r_{k+1}, d_k\rangle = 0. \tag{3.7}$$

Idée. Construire des directions (d_k) deux à deux A-orthogonales, comme ça r_{k+1} sera orthogonal à $\mathrm{Vect}(d_0, \ldots, d_k)$.

Notations. Pour $x, y \in \mathbf{R}^n$, on note $x \perp y$ lorsque x et y sont orthogonaux pour le produit scalaire euclidien et $x \perp_A y$ lorsque x et y sont orthogonaux pour le produit scalaire donné par A. On étend naturellement cette notation à des sous-espaces de $\mathbf{R}^n$.

On pose $d_0 = r_0$ et pour $k \in \mathbf{N}$, on construit d_{k+1} comme l'orthogonalisé de Gram-Schmidt pour le produit scalaire donné par A de r_{k+1} relativement à $\mathrm{Vect}(d_k)$:

$$d_{k+1} = r_{k+1} - \beta_k d_k \tag{3.8}$$

où

$$\beta_k = \frac{\langle r_{k+1}, Ad_k\rangle}{\|d_k\|_A^2} \quad \text{si } d_k \neq 0, \quad \beta_k = 0 \text{ sinon.} \tag{3.9}$$

Remarquons que si $d_k = 0$ alors r_k et d_{k-1} sont colinéaires et comme ils sont aussi orthogonaux par (3.7), $r_k = 0$.

Lemme. *Avec le choix* (3.9)*, les directions* (3.8) *vérifient pour tout $k \in \mathbf{N}$ la propriété suivante : si $r_0, \ldots, r_k$ ne sont pas nuls alors,*

(i) $\mathrm{Vect}(r_0, \ldots, r_k) = \mathrm{Vect}(d_0, \ldots, d_k)$

(ii) $r_{k+1} \perp \text{Vect}(d_0, \ldots, d_k)$

(iii) $d_{k+1} \perp_A \text{Vect}(d_0, \ldots, d_k)$

Preuve. On procède par récurrence sur $k \in \mathbf{N}$. Lorsque $k = 0$, (i), (ii) et (iii) sont vrais grâce aux relations $r_0 = d_0$, (3.7) et (3.8) et bien sûr $r_0 \neq 0$ sinon il n'y a rien à faire. Supposons donc le résultat vrai au rang $k-1$, $k \in \mathbf{N}^*$.

(i) Par (3.8), on a : $d_k = r_k - \beta_{k-1} d_{k-1}$.

(ii) Par (3.7), on a déjà $r_{k+1} \perp d_k$ et si $j \in \{0, \ldots, k-1\}$, la relation (3.6) couplée à l'hypothèse de récurrence (ii) et (iii) donne $r_{k+1} \perp d_j$.

(iii) Par (3.8), on a déjà $d_{k+1} \perp_A d_k$ (c'est la définition) et si $j \in \{0, \ldots, k-1\}$, la relation (3.8) couplée à l'hypothèse de récurrence (iii) donne :

$$\langle d_{k+1}, Ad_j \rangle = \langle r_{k+1}, Ad_j \rangle.$$

Montrons que $Ad_j \in \text{Vect}(r_0, \ldots, r_k)$, ce qui conclura grâce aux relations (i) et (ii) que l'on vient de prouver. Grâce à la relation (3.6) avec $k = j$, il suffit de montrer que $\alpha_j \neq 0$, ce qui est la cas car :

$$\alpha_j = 0 \overset{(3.5)}{\iff} \langle r_j, d_j \rangle = 0 \overset{(3.8)}{\iff} r_j = 0$$

et on a justement supposé le contraire.

□

Théorème. *La méthode de gradient associée aux directions* (3.8) *avec le choix* (3.9) *converge vers la solution x du problème* (3.2) *en au plus n itérations.*

Preuve. Les conditions (i) et (ii) du lemme précédent assurent que la famille $(r_k)_k$ est une famille orthogonale donc libre. On est en dimension n. □

Référence. A. Quarteroni, R. Sacco, F. Saleri, *Numerical Mathematics*

158 Matrices symétriques réelles, matrices hermitiennes.
162 Systèmes d'équations linéaires ; opérations élémentaires, aspects algorithmiques et conséquences théoriques.
219 Extrema : existence, caractérisation, recherche. Exemples et applications.
226 Suites vectorielles et réelles définies par une relation de récurrence $u_{n+1} = f(u_n)$. Exemples. Applications à la résolution approchée d'équations.
233 Méthodes itératives en analyse numérique matricielle.

3.8 Les sous-groupes compacts de $GL(E)$

Soit E un $\mathbf{R}$-espace vectoriel euclidien de dimension finie. On commence par un théorème de point fixe.

Théorème (Kakutani)**.** *Soient G un sous-groupe compact de $GL(E)$ et K un convexe compact non vide tel que pour tout $g \in G$, $g(K) \subset K$. Alors il existe $x \in K$ tel que pour tout $g \in G$, $g(x) = x$.*

Preuve. (1) On commence par montrer le résultat où G est remplacé par un singleton. Soit donc $v \in GL(E)$ qui stabilise K. Alors à partir de $x_0 \in K$, on définit la suite :

$$x_k = \frac{1}{k+1} \sum_{j=0}^{k} v^j(x_0).$$

Par compacité, on peur extraire une sous-suite convergente $x_{\varphi(k)} \to x \in K$. Alors :

$$v(x_k) = x_k + \frac{1}{k+1}(v^{k+1}(x_0) - x_0)$$

et en particulier

$$v(x_{\varphi(k)}) = x_{\varphi(k)} + \varepsilon_k$$

où $(\varepsilon_k)_k$ est une suite qui tend vers 0 (car K est v-stable). Par continuité, on a $v(x) = x$.

(2) On note $\|\cdot\|$ la norme euclidienne sur E et on définit :

$$\|x\|_G = \max_{g \in G} \|g(x)\|.$$

On a écrit max car G est compact. On a défini une norme sur E qui est invariante par G. Étudions le cas d'égalité dans l'inégalité triangulaire : si $x, y \in E$, on a pour un certain $u_0 \in G$:

$$\|x+y\|_G = \|u_0(x) + u_0(y)\| \leq \|u_0(x)\| + \|u_0(y)\| \leq \|x\|_G + \|y\|_G.$$

S'il y a égalité alors $u_0(x)$ et $u_0(y)$ sont positivement liés. Puisque $u_0 \in GL(E)$, c'est aussi le cas de x et y. Réciproquement, c'est aussi vrai.

(3) Par la première étape, on sait que pour tout $u \in G$,

$$K^u = \{x \in K,\ u(x) = x\} \neq \emptyset$$

mais on veut montrer que

$$\bigcap_{u \in G} K^u \neq \emptyset.$$

Comme les K^u sont des fermés du compact K, il suffit de considérer les intersections finies :

$$\bigcap_{1 \leq k \leq p} K^{u_k} \neq \emptyset, \quad p \in \mathbf{N}.$$

On pose avec ces notations :

$$v = \frac{1}{p} \sum_{k=1}^{p} u_k \quad \text{et} \quad v(x) = x \in K$$

L'existence de x est encore assurée par la première étape. Avec la norme $\|\cdot\|_G$:

$$\|x\|_G = \|v(x)\|_G \leq \frac{1}{p} \sum_{k=1}^{p} \|u_k(x)\|_G = \|x\|_G$$

et le cas d'égalité étudié en (2) donne la positive-liaison des $u_k(x)$ et donc leur égalité puisqu'ils sont de même norme $\|\cdot\|_G$. Par suite, x est dans l'intersection des K^{u_k}.

□

3. Non exclusifs

Signalons avant de poursuivre une preuve un rien plus concise du théorème de Kakutani : avec le produit scalaire de la deuxième étape, on considère $x \in K$ qui minimise la norme $\|\cdot\|_G$ sur K (c'est un compact) et on montre qu'il convient. Pour cela, on regarde :

$$\frac{1}{2}(x + g(x)) \in K$$

pour $g \in G$. On a :

$$\|x\|_G \leq \|1/2(x + g(x))\|_G \leq \|x\|_G.$$

donc il y a égalité dans l'inégalité triangulaire et x et $g(x)$ sont positivement liés : comme ils ont même norme $\|\cdot\|_G$ ils sont égaux. Ceci est vrai pour tout $g \in G$.

En revanche, la première étape de la preuve est intéressante en elle-même, elle permet par exemple de prouver très simplement le théorème de Massera sur les équations différentielles linéaires pérodiques (voir Gonnord, Tosel, *Calcul Différentiel*).

Théorème. *Soit G un sous-groupe compact de $GL_n(\mathbf{R})$. Alors il existe un produit scalaire euclidien $\langle\ ,\ \rangle^G$ sur $\mathbf{R}^n$ de forme quadratique associée q^G telle que $G \subset \mathcal{O}(q^G)$*

Preuve. L'idée est d'utiliser le formalisme des actions de groupes pour voir le problème $G \subset \mathrm{Stab}(S)$ pour un certain $S \in \mathscr{S}_n^{++}$ comme un problème de point fixe.

Pour $A \in G$, on pose pour $S \in \mathscr{S}_n$:

$$\rho(A)(S) = ASA^T.$$

On note que $\rho : G \to GL(\mathscr{S}_n)$ est continue car polynômiale. On a défini une action de groupe sur $\mathscr{S}_n$. On regarde maintenant $\mathscr{G} = \rho(G)$. C'est un sous-groupe compact de $GL(\mathscr{S}_n)$.

(1) On regarde $\mathrm{Orb}_\rho(I_n) = \{MM^T,\ M \in G\}$: c'est un compact non vide du convexe $\mathscr{S}_n^{++}$ donc, par le théorème de Carathéodory, c'est aussi le cas de son enveloppe convexe, que l'on note K.

(2) Par construction, K est $\mathscr{G}$-stable car pour tout $A \in G$:

$$\rho(A)(MM^T) = (AM)(AM)^T \in \mathrm{Orb}_\rho(I_n)$$

et on conclut par linéarité.

(3) On utilise le théorème de point fixe de Kakutani : il existe $S \in K \subset \mathscr{S}_n^{++}$ fixé par tous les éléments de $\mathscr{G}$. En termes d'action de groupes,

$$G \subset \mathrm{Stab}_\rho(S)$$

et en termes de formes quadratiques :

$$\forall A \in G,\ \ ASA^T = S.$$

On a fini. En prenant la racine carrée de S, on peut aussi écrire :

$$\forall A \in G,\ \ (\sqrt{S}^{-1}A\sqrt{S})(\sqrt{S}^{-1}A\sqrt{S})^T = I_n \ \ \textit{i.e}\ \ \sqrt{S}^{-1}G\sqrt{S} \subset \mathcal{O}(E).$$

□

Quelques remarques complémentaires.

- Pas de problème sur $\mathbf{C}$ en remplaçant les transposées par des transconjuguées. Du coup, on est dans un espace hermitien.
- Le théorème de Carathéodory dit que dans un espace affine de dimension n, l'enveloppe convexe d'une partie A est l'ensemble des barycentres à coefficients positifs ou nul de $n+1$ points de A. La conséquence évoquée plus haut, à savoir que l'enveloppe convexe d'un compact est compacte vient du fait suivant :

$$\Lambda_{n+1} \times K^{n+1} \to \mathrm{Conv}(K),\ (\lambda, x) \mapsto \sum_{i=1}^{n+1} \lambda_i x_i$$

  est surjective continue et l'espace de départ est compact.
- En dimension infinie, l'enveloppe convexe d'un compact est précompacte.
- Il y a aussi une preuve qui utilise l'ellipsoïde de John et Loewner !
- Mesure de Haar.

Référence. M. Alessandri, *Thèmes de Géométrie. Groupes en situation géométrique.*

106 Groupe linéaire d'un espace vectoriel de dimension finie E, sous-groupes de $GL(E)$. Applications.
150 Exemples d'actions de groupes sur les espaces de matrices.
170 Formes quadratiques sur un espace vectoriel de dimension finie. Orthogonalité, isotropie. Applications.
171 Formes quadratiques réelles. Coniques. Exemples et applications.
181 Barycentres dans un espace affine réel de dimension finie, convexité. Applications.

3.9 Sur la proportion des couples d'entiers premiers entre eux

Pour $n \in \mathbf{N}^*$, on note r_n la proportion des couples d'entiers de $\{1, \dots, n\}$ formés d'entiers premiers entre eux.

Proposition. *Si μ désigne la fonction de Möbius, on a :*

$$r_n = \frac{1}{n^2} \sum_{d=1}^{n} \mu(d) \left\lfloor \frac{n}{d} \right\rfloor^2.$$

En particulier :

$$\lim_{n \to +\infty} r_n = \frac{6}{\pi^2}.$$

Preuve. On note A_n l'ensemble des couples $(a, b) \in \{1, \dots, n\}^2$ tels que $a \wedge b = 1$. On a donc :

$$r_n = \frac{\mathrm{Card}\, A_n}{n^2}.$$

3. Non exclusifs

Soient $p_1, \ldots, p_k$ les nombres premiers inférieurs à n et U_i l'ensemble des couples $(a, b) \in \{1, \ldots, n\}^2$ tels que $p + i | a$ et $p_i | b$, On a donc :

$$A_n = \left(\bigcup_{i=1}^{k} U_i \right)^c.$$

On utilise la formule du crible pour calculer le cardinal de A_n :

$$\operatorname{Card} \left(\bigcup_{i=1}^{k} U_i \right) = \sum_{\emptyset \neq I \subset \{1,\ldots,k\}} (-1)^{1+\operatorname{Card} I} \operatorname{Card} \left(\bigcap_{i \in I} U_i \right).$$

Maintenant, si $I \subset \{1, \ldots, k\}$ est non vide, l'intersection des U_i pour $i \in I$ est exactement l'ensemble des couples de multiples strictement positifs de $\prod_{i \in I} p_i$ inférieurs ou égaux à n. On a donc :

$$\operatorname{Card} \left(\bigcap_{i \in I} U_i \right) = \left\lfloor \frac{n}{\prod_{i \in I} p_i} \right\rfloor^2.$$

Ainsi :

$$\begin{aligned} \operatorname{Card} A_n &= n^2 - \sum_{\emptyset \neq I \subset \{1,\ldots,k\}} (-1)^{1+\operatorname{Card} I} \operatorname{Card} \left(\bigcap_{i \in I} U_i \right) \\ &= n^2 - \sum_{\emptyset \neq I \subset \{1,\ldots,k\}} (-1)^{1+\operatorname{Card} I} \left\lfloor \frac{n}{\prod_{i \in I} p_i} \right\rfloor^2 = \sum_{d=1}^{n} \mu(d) \left\lfloor \frac{n}{d} \right\rfloor^2. \end{aligned}$$

On en déduit la formule pour r_n. Pour le calcul de la limite, on aura besoin du :

Lemme. *Pour tout entier $n \in \mathbf{N}^*$,* $\displaystyle\sum_{d|n} \mu(d) = \begin{cases} 1 & si \quad n = 1 \\ 0 & si \quad n \geq 2 \end{cases}$

Preuve. Il n'y a rien à prouver pour $n = 1$ et si $n \geq 2$, on écrit sa décomposition en facteurs premiers :

$$n = \prod_{i=1}^{k} p_i^{\alpha_i}$$

de sorte que

$$\sum_{d|n} \mu(d) = \sum_{i=0}^{k} \binom{k}{i} (-1)^i = (1-1)^k = 0.$$

□

Maintenant, on écrit :

$$\left| r_n - \sum_{d=1}^{n} \frac{\mu(d)}{d^2} \right| = \left| \sum_{d=1}^{n} \mu(d) \left(\frac{1}{n^2} \left\lfloor \frac{n}{d} \right\rfloor^2 - \frac{1}{d^2} \right) \right|.$$

Puis on utilise la majoration :

$$\left\lfloor \frac{n}{d} \right\rfloor > \frac{n}{d} - 1 \quad \Rightarrow \quad \frac{1}{n^2} - \frac{2}{dn} < \frac{1}{n^2} \left\lfloor \frac{n}{d} \right\rfloor^2 - \frac{1}{d^2} \geq 0$$

pour continuer :

$$\left| r_n - \sum_{d=1}^{n} \frac{\mu(d)}{d^2} \right| \leq \sum_{d=1}^{n} \left(\frac{2}{dn} - \frac{1}{n^2} \right) \leq \frac{2}{n} \sum_{d=1}^{n} \frac{1}{d} - \frac{1}{n} = \mathcal{O}\left(\frac{\ln n}{n} \right).$$

Puisqu'il y a absolue convergence de la série des $\mu(d)/d$:

$$\lim_{n \to +\infty} r_n = \sum_{d=1}^{+\infty} \frac{\mu(d)}{d^2}.$$

Enfin, par sommabilité des séries en question, on applique le théorème sur le produit des séries :

$$\begin{aligned} \left(\sum_{d=1}^{+\infty} \frac{\mu(d)}{d^2} \right) \left(\sum_{n=1}^{+\infty} \frac{1}{n^2} \right) &= \sum_{d,n \geq 1} \frac{\mu(d)}{(dn)^2} = \sum_{d \geq 1,\, d|m} \frac{\mu(d)}{m^2} \\ &= \sum_{m \geq 1} \sum_{d|m} \frac{\mu(d)}{m^2} = \sum_{m \geq 1} \frac{1}{m^2} \sum_{d|m} \mu(d) = 1. \end{aligned}$$

D'où la conclusion. □

Référence. S. Francinou, H. Gianella, S. Nicolas *Oraux X-ENS : Algèbre 1*

190 Méthodes combinatoires, problèmes de dénombrement.
230 Séries de nombres réels ou complexes. Comportement des restes ou des sommes partielles des séries numériques. Exemples.

4 Rebut

4.1 La formule de Taylor généralisée

C'est très calculatoire et assez inutile mais a une esthétique certaine. Ça vaut peut être le coup comme exemple dans la leçon 218 notamment.

Rappelons la formule de Taylor avec reste intégral pour une fonction C^∞ sur $\mathbf{R}$:

$$f(x) = \sum_{n=0}^{N} f^{(n)}(0)\frac{x^n}{n!} + \int_0^x \frac{(x-t)^N}{N!} f^{(n+1)}(t)dt. \tag{4.1}$$

Théorème. *Soient X une variable aléatoire réelle et f une fonction indéfiniment dérivable sur $\mathbf{R}$ et telle que $f^{(n)}(X)$ soit intégrable pour tout n. Alors, la suite de polynômes $(q_n)_{n\geq 0}$ définie par :*

$$q_0 = 1, \qquad \frac{q_n'}{n!} = \frac{q_{n-1}}{(n-1)!}, \qquad \int_{\mathbf{R}} q_n d\mathbf{P}_X = 0 \ \textit{pour tout}\ n \geq 1 \tag{4.2}$$

vérifie pour tout $x \in \mathbf{R}$:

$$f(x) = \sum_{n=0}^{N} \mathbf{E}\big(f^{(n)}(X)\big)\frac{q_n(x)}{n!} + \mathbf{E}\left(\int_0^{x-X} \frac{q_N(x-s)}{N!} f^{(N+1)}(X+s)ds\right).$$

Preuve. Soient $x, y \in \mathbf{R}$. On écrit en intégrant par parties :

$$\begin{aligned}
&\int_0^{x-y} \frac{q_N(x-s)}{N!} f^{(N+1)}(y+s)ds \\
&\quad = \frac{q_N(y)}{N!}f^{(N)}(x) - \frac{q_N(x)}{N!}f^{(N)}(y) + \int_0^{x-y} \frac{q_{N-1}(x-s)}{(N-1)!} f^{(N)}(y+s)ds \\
&\quad = \dots \\
&\quad = \sum_{n=1}^{N}\left(\frac{q_n(y)}{n!}f^{(n)}(x) - \frac{q_n(x)}{n!}f^{(n)}(y)\right) + \int_0^{x-y} q_0(x-s)f'(y+s)ds \\
&\quad = \sum_{n=0}^{N}\left(\frac{q_n(y)}{n!}f^{(n)}(x) - \frac{q_n(x)}{n!}f^{(n)}(y)\right)
\end{aligned}$$

En prenant l'espérance de la variable aléatoire :

$$\sum_{n=0}^{N}\left(\frac{q_n(X)}{n!}f^{(n)}(x) - \frac{q_n(x)}{n!}f^{(n)}(X)\right)$$

on trouve :

$$\mathbf{E}\left[\sum_{n=0}^{N}\left(\frac{q_n(X)}{n!}f^{(n)}(x)-\frac{q_n(x)}{n!}f^{(n)}(X)\right)\right]=f(x)-\sum_{n=0}^{N}\frac{q_n(x)}{n!}\mathbf{E}\big(f^{(n)}(X)\big)$$

qui est bien la formule annoncée. □

Remarque. Les conditions (4.2) découlent de la relation :

$$\frac{e^{xz}}{\mathbf{E}(e^{Xz})}=\sum_{n\geq 0}\frac{q_n(x)}{n!}z^n \tag{4.3}$$

pour tout $x, z \in \mathbf{R}$ et dont on montre qu'elle est nécessaire en considérant le cas $f(x)=e^{xz}$.

Quelques exemples et applications

Pour des lois de probabilités bien choisies, on retrouve à peu de frais des développements asymptotiques classiques. Soit à partir de maintenant f une fonction vérifiant les hypothèses du théorème pour les lois considérées.

- **La formule de Taylor usuelle.** Pour $\mathbf{P}_X=\delta_0$ on retrouve la formule de Taylor (4.1).
- **La formule d'Euler-MacLaurin.** Lorsque X suit la loi uniforme sur $[0,1]$, les polynômes q_n sont les polynômes de Bernoulli B_n et :

$$g(x)=\sum_{n=0}^{N}\frac{B_n(x)}{n!}\int_0^1 g^{(n)}(y)dy+\int_0^1 dy\int_0^{x-y}ds\frac{B_N(x-s)}{N!}g^{(N+1)}(y+s).$$

pour toute fonction g vérifiant les hypothèses du théorème. En particulier, si $g(x)=f(x+k)$ pour $k\in\mathbf{N}$, on trouve en $x=0$:

$$f(k)=\int_k^{k+1}f(y)dy+\sum_{n=1}^{N}\Big(f^{(n-1)}(k+1)-f^{(n-1)}(k)\Big)\frac{B_n(0)}{n!}$$
$$-\int_0^1 dy\int_0^y ds\frac{B_N(s)}{N!}f^{(N+1)}(y-s+k)$$

En intervertissant les intégrales dans le reste, il est égal à :

$$R_N^{(k)}(f):=-\frac{1}{N!}\int_0^1 dt\,B_N(t)\int_1^t dy\,f^{(N+1)}(y-t+k)=\frac{-1}{N!}\int_0^1 B_N(t)\Big(f^{(N)}(k+1-t)-f^{(N)}(k)\Big)dt.$$

Comme les polynômes de Bernoulli sont d'intégrale nulle, on trouve après changement de variable :

$$R_N(f)=\frac{-1}{N!}\int_k^{k+1}B_N(k+1-t)f^{(N)}(t)dt.$$

Pour $t\in[k,k+1]$, on peut écrire $k+1-t=1-(t-\lfloor t\rfloor)$ et puisque $B_n(1-t)=(-1)^nB_n(t)$, on a finalement :

$$R_N^{(k)}(f)=\frac{(-1)^{N+1}}{N!}\int_k^{k+1}B_N(t-\lfloor t\rfloor)f^{(N)}(t)dt.$$

Et il suffit de sommer k entre 1 et $n-1$ pour retrouver la formule d'Euler-MacLaurin :

$$\sum_{k=1}^{n-1} f(k) = \int_1^n f(y)dy + \sum_{m=1}^{N} \left(f^{(m-1)}(n) - f^{(m-1)}(1)\right)\frac{B_m(0)}{m!} + \frac{(-1)^{N+1}}{N!}\int_1^n B_N(t-\lfloor t\rfloor)f^{(N)}(t)dt.$$

- **Le développement de Tchebychev-Hermite.** On considère cette fois $X \sim \mathcal{N}(0,1)$. La relation (4.3) définit la suite des polynômes q_n, ce sont les polynômes de Hermite

$$q_n(x) = H_n(x) = (-1)^n e^{x^2/2}\partial^n(e^{-x^2/2}).$$

Finalement,

$$f(x) = \sum_{n=0}^{N} \frac{1}{\sqrt{2\pi}}\left(\int_{\mathbf{R}} e^{-t^2/2}f^{(n)}(t)dt\right)\frac{H_n(x)}{n!} + \frac{1}{\sqrt{2\pi}}\int_{\mathbf{R}} dy\ e^{-y^2/2}\int_0^{x-y} dt\frac{H_N(x-t)}{N!}f^{(N+1)}(y+t)$$

En intégrant n fois par parties, on voit que

$$\frac{1}{\sqrt{2\pi}}\int_{\mathbf{R}} e^{-t^2/2}\partial^n f(t)dt = (-1)^n\int_{\mathbf{R}}\partial^n(e^{-t^2/2})f(t)dt = \frac{1}{\sqrt{2\pi}}\int_{\mathbf{R}} f(t)H_n(t)e^{-t^2/2}dt.$$

En posant $\widetilde{H_n} = H_n/\sqrt{n!}$, on trouve :

$$f(x) = \sum_{n=0}^{N}\langle f, \widetilde{H_n}\rangle\widetilde{H_n} + R_N(f).$$

où $\langle\cdot,\cdot\rangle$ est le produit scalaire défini par

$$\langle f, g\rangle = \frac{1}{\sqrt{2\pi}}\int_{\mathbf{R}} f(t)g(t)e^{-t^2/2}dt$$

et où $R_N(f)$ est un reste explicitement calculable dont on montre qu'il tend vers 0 en norme 2 pour le produit scalaire considéré. En particulier, on retrouve (en partie) le fait que les polynômes de Hermite modulo une constante forment une base hilbertienne pour le produit scalaire considéré.

Référence. Bernard Candelpergher, *Théorie des probabilités*

4.2 La décomposition de Bruhat et les drapeaux

Pourquoi pas.

4. Rebut

On note T_s l'ensemble des matrices triangulaires supérieures inversibles sur $\mathbf{K}^n$.

Théorème (Décomposition de Bruhat)**.** *Le groupe produit $T_s \times T_s$ agit transitivement sur $GL_n(\mathbf{K})$ et un système de représentants des orbites est donné par les matrices de permutation :*

$$GL_n(\mathbf{K}) = \bigsqcup_{\sigma \in \mathfrak{S}_n} T_s P_\sigma T_s.$$

Preuve. Il s'agit d'adapter l'algorithme du pivot de Gauss. Prenons $A = (a_{i,j})_{i,j} \in GL_n(\mathbf{K})$ et procédons comme suit :

(1) La première colonne de A n'est pas nulle et i_1 est le plus grand indice tel que $a_{i_1,1} \neq 0$. On opère :

- À gauche pour $k < i_1$, $T_{i_1,k}(-a_{k,1}/a_{i_1,1}) : L_k \leftarrow L_k - \frac{a_{k,1}}{a_{i_1,1}} L_{i_1}$
- À droite pour $k > 1$, $T_{1,k}(-a_{i_1,k}/a_{i_1,1}) : C_k \leftarrow C_k - \frac{a_{i_1,k}}{a_{i_1,1}} C_1$
- On dilate $D_{i_1}(1/a_{i_1,1}) : L_1 \leftarrow \frac{1}{a_{i_1,1}} L_1$

La matrice ressemble maintenant à :

$$\begin{pmatrix} 0 & & & \\ \vdots & & * & \\ 0 & & & \\ 1 & 0 & \dots & 0 \\ 0 & & & \\ \vdots & & * & \\ 0 & & & \end{pmatrix}$$

(2) La matrice est toujours inversible donc la deuxième colonne n'est pas nulle et i_2 est le plus grand indice tel que le coefficient $(i_2, 2)$ n'est pas nul. En procédant comme en (1), on annule la deuxième colonne, la i_2-ème ligne et le coefficients $(i_2, 2)$ vaut 1. Bien sûr $i_2 \neq i_1$.

(3) On construit une suite injective $i_1, \dots, i_n$. Elle est aussi bijective et en notant σ la permutation associée, on a :

$$T_1 A T_2 = P_\sigma$$

où $T_1, T_2 \in T_s$ car les matrices de transvections et de dilations sont dans T_s. Finalement,

$$A = T_1^{-1} P_\sigma T_2^{-1} \in \bigsqcup_{\sigma \in \mathfrak{S}_n} T_s P_\sigma T_s.$$

Il reste à montrer l'unicité d'une telle décomposition : supposons qu'il existe $T, T' \in T_s$ et des permutations σ et σ' telles que $TP_\sigma = P_{\sigma'} T'_{\sigma'}$. *Multiplier par une matrice de permutation revient à permuter les lignes/colonnes*[1] et plus précisément, soit $j \in \{1, \dots, n\}$:

- Le coefficient $(\sigma'(j), j)$ de $P_{\sigma'} T' = TP_\sigma$ est le coefficient (j, j) de T'. Il est non nul car T' est inversible.
- Le coefficient $(\sigma'(j), j)$ de TP_σ est le coefficient $(\sigma'(j), \sigma(j))$ de T. Il est nul si $\sigma'(j) > \sigma(j)$.

1. La j-ème ligne de $P_\sigma T$ est la $\sigma(j)$-ème ligne de T et ATTENTION, la k-ème colonne de TP_σ est la $\sigma^{-1}(k)$-ème colonne de T.

• On a donc montré $\sigma'(j) \leq \sigma(j)$. Par symétrie, $\sigma(j) = \sigma'(j)$ et $\sigma = \sigma'$.

□

Un *drapeau* de $\mathbf{K}^n$ est une suite croissante de sous-espaces $\{0\} = F_0 \subset \ldots \subset F_n = \mathbf{K}^n$ telle que F_k est de dimension k. On note $\mathcal{D}$ l'ensemble des drapeaux de $\mathbf{K}^n$.

Théorème. *Le groupe $GL_n(\mathbf{K})$ agit transitivement sur $\mathcal{D}$. De plus, l'action de $GL_n(\mathbf{K})$ sur $\mathcal{D} \times \mathcal{D}$ a $n!$ orbites.*

Preuve. (1) Si $d = (F_0, \ldots, F_n)$ est un drapeau et $A \in GL_n(\mathbf{K})$, alors

$$A \cdot d = (A(F_0), A(F_1), \ldots, A(F_n))$$

définit une action qui est clairement transitive (par le théorème de la base incomplète, il existe une base $(f_1, \ldots, f_n)$ de $\mathbf{K}^n$ telle que pour tout $k \in \{1, \ldots, n\}$, $F_k = \mathrm{Vect}(f_1, \ldots, f_k)$ et d est alors dans l'orbite du drapeau canonique).

(2) Puisqu'il n'y a qu'une seule classe, la relation orbite/stabilisateur donne une bijection entre $\mathcal{D}$ et les classes à gauche :

$$\mathcal{D} \simeq GL_n(\mathbf{K})/\operatorname{Stab}_{GL_n(\mathbf{K})}(\delta)$$

où δ est le drapeau canonique formé des sous-espaces $\mathrm{Vect}(e_1, \ldots, e_k)$ avec $(e_1, \ldots, e_n)$ la base canonique de $\mathbf{K}^n$. Ah oui mais il est facile de voir que $\operatorname{Stab}_{GL_n(\mathbf{K})}(\delta) = T_s$ et on peut donc identifier (c'est une bijection)

$$\mathcal{D} \simeq GL_n(\mathbf{K})/T_s.$$

(3) On considère l'action de $GL_n(\mathbf{K})$ sur $\mathcal{D} \times \mathcal{D}$:

$$A \cdot (d, d') = (A \cdot d, A \cdot d').$$

Compte-tenu de (2), en identifiant $d = X \cdot \delta$ avec $\overline{X} \in GL_n(\mathbf{K})/T_s$ et $d' = Y \cdot \delta$ avec $\overline{Y} \in GL_n(\mathbf{K})/T_s$, on obtient une action de $GL_n(\mathbf{K})$ sur $GL_n(\mathbf{K})/T_s \times GL_n(\mathbf{K})/T_s$ définie par :

$$A \cdot (\overline{X}, \overline{Y}) = (\overline{AX}, \overline{AY}).$$

(4) Comptons maintenant les orbites de cette dernière action. Soit $(\overline{X}, \overline{Y}) \in (GL_n(\mathbf{K})/T_s)^2$. On utilise la décomposition de Bruhat de $X^{-1}Y$ pour écrire :

$$(\overline{X}, \overline{Y}) = X \cdot (\overline{I_n}, \overline{T_1 P_\sigma T_2}) = XT_1 \cdot (\overline{T_1^{-1}}, \overline{P_\sigma T_2}) = XT_1 \cdot (\overline{I_n}, \overline{P_\sigma}).$$

Ainsi : $(\overline{X}, \overline{Y}) \in \operatorname{Orb}_{GL_n(\mathbf{K})}(\overline{I_n}, \overline{P_\sigma})$. Un système de représentants des orbites est donc donné par les $(\overline{I_n}, \overline{P_\sigma})$ et toutes les orbites sont distinctes par unicité de σ dans la décomposition de Bruhat. Il y a donc $n!$ orbites.

□

Références.

S. Francinou, H. Gianella, S. Nicolas, *Oraux X-ENS, Algèbre 1*
R. Mneimné, F. Testard *Introduction aux groupes de Lie classiques*
H2G2, tome 2.

101 Groupe opérant sur un ensemble. Exemples et applications.
105 Groupe des permutations d'un ensemble fini. Applications.
150 Exemples d'actions de groupes sur les espaces de matrices.
157 Endomorphismes trigonalisables. Endomorphismes nilpotents.
162 Systèmes d'équations linéaires ; opérations élémentaires, aspects algorithmiques et conséquences théoriques.

4.3 Le théorème de Riesz-Thorin

Il me semble qu'il y a des subtilités pas toujours détaillées dans les livres. En plus, je ne connais pas d'application simple (mais il y a des applications importantes : cf. Linares, Ponce, Introduction to Nonlinear Dispersive Equations).

Théorème (Riesz-Thorin)**.** *Soient (X,μ) et (Y,ν) des espaces mesurés et $p_0 \neq p_1$, $q_0 \neq q_1$ dans $[1,\infty]$. On se donne une application linéaire :*

$$T : L^{p_i}(X) \to L^{q_i}(Y)$$

continue pour $i \in \{0,1\}$ avec pour normes respectives M_0 et M_1. Pour $a \in (0,1)$, on considère p_a et q_a dans $[1,\infty]$ tels que :

$$\frac{1}{p(a)} = \frac{1-a}{p_0} + \frac{a}{p_1} \quad et \quad \frac{1}{q(a)} = \frac{1-a}{q_0} + \frac{a}{q_1}.$$

Alors T est une application linéaire continue de $L^{p(a)}(X) \to L^{q(a)}(Y)$ de norme M_a telle que :

$$M_a \leq M_0^{1-a} M_1^a.$$

On montre d'abord le théorème dit des *trois droites*.

Théorème (Hadamard)**.** *Soit ϕ une fonction holomorphe bornée sur la bande $\{0 \leq \operatorname{Re} z \leq 1\}$. On note pour $a \in [0,1]$:*

$$N(a) := \sup_{\eta} |\phi(a+i\eta)|$$

Alors

$$N(a) \leq N(0)^{1-a} N(1)^a.$$

Preuve. On suppose que $N(0)$ et $N(1)$ sont non nuls. Soit $c = \log N(0)/N(1)$. La fonction $z \mapsto \phi(z)e^{cz}$ bornée dans la bande $\{0 \leq \operatorname{Re} z \leq 1\}$ et son module est inférieur à $N(0)$ lorsque $\operatorname{Re} z = 0$ ou $\operatorname{Re} z = 1$. Par le principe du maximum :

$$|\phi(a+i\eta)| e^{ca} \leq N(0)$$

et le résultat suit de la définition de c. □

Preuve. (Riesz-Thorin). Pour $p \in [1,\infty]$ on notera $p' \in [1,\infty]$ tel que

$$\frac{1}{p} + \frac{1}{p'} = 1.$$

4. Rebut

Prenons $p(a)$ et $q(a)$ comme dans l'énoncé et $f \in L^{p(a)}$. Notons que T est bien défini sur $L^{p_0}(X) \cap L^{p_1}(Y) \to L^{q_0}(Y) \cap L^{q_1}(Y)$ donc $T : L^{p(a)}(X) \to L^{q(a)}(Y)$ l'est aussi.

On veut majorer la norme $M_a \in \overline{\mathbf{R}_+}$ de $T : L^{p(a)}(X) \to L^{q(a)}(Y)$. Par dualité :

$$M_a = \sup_{\|f\|_{L^p}=1,\ \|h\|_{L^q}=1} |\langle h, Tf\rangle|.$$

Soient donc $f = |f|e^{i\alpha} \in L^p(X)$ et $h = |h|e^{i\beta} \in L^{q'}(Y)$ de norme 1. On pose pour $0 \leq \operatorname{Re} z \leq 1$:

$$f_z = |f|^{p(a)/p(z)}e^{i\alpha} \quad \text{et} \quad h_z = |h|^{q'(a)/q'(a)}e^{i\beta}.$$

La fonction :

$$\phi(z) = \langle h_z, Tf_z\rangle = \int_Y h_z(y)\, Tf_z(y)dy$$

est bien définie et holomorphe sur $\{0 < \operatorname{Re} z < 1\}$ comme un calcul direct le montrerait. Notons

$$N(a) = \sup_{\operatorname{Re} z=a} |\phi(z)|.$$

On va majorer $N(0)$ et $N(1)$: soit $z = i\eta$, $\eta \in \mathbf{R}$, alors par définition :

$$\frac{p(a)}{p(z)} = \frac{p(a)}{p_0} + \text{imag.} \quad \text{et} \quad \frac{q'(a)}{q'(z)} = \frac{q'(a)}{q'_0} + \text{imag.}.$$

Ainsi :

$$|f_z|^{p_0} = |f|^{\operatorname{Re}\left(p_0 \times p(a)/p(z)\right)} = |f|^{p(a)}$$

de sorte que :

$$\|f_z\|_{p_0}^{p_0} = \|f\|_{L^{p(a)}}^{p(a)} = 1 \quad \text{et} \quad \|h_z\|_{L^{q'_0}}^{q'_0} = \|h\|_{L^{q'(a)}}^{q'(a)} = 1.$$

L'inégalité de Hölder permet de conclure :

$$|\phi(z)| \leq \|h_z\|_{L^{q'_0}} \|Tf_z\|_{L^{q_0}} \leq M_0.$$

Finalement $N(0) \leq M_0$ et par un raisonnement analogue, $N(1) \leq M_1$. La théorème résulte alors du lemme des trois droites de Hadamard appliqué à ϕ :

$$|\phi(a)| \leq N(a) \leq M_0^{1-a} M_1^a$$

et comme $f_a = f$ et $h_a = h$, on a le résultat. □

TOUT EST FAUX

Référence. P. D. Lax, *Functional Analysis*

4.4 Un anneau vraiment passionnant

Théorème. *L'anneau*

$$\mathbf{Z}\left[\frac{1+i\sqrt{19}}{2}\right] = \left\{z = a + b\frac{1+i\sqrt{19}}{2} \in \mathbf{C},\ a, b \in \mathbf{Z}\right\}$$

est principal mais non euclidien.

Proposition. *Soit A un anneau euclidien. Il existe $x \in A \setminus A^\times$ tel que la restriction à $A^\times \cup \{0\}$ de la projection canonique de A sur $A/(x)$ soit surjective.*

Preuve. Si A est un corps, $x = 0$ convient. Sinon, parmi les éléments de A non nuls et non inversibles, on choisit x de stathme minimal. Alors, si $a \in A$, on écrit $a = xq + r$ avec $r = 0$ ou $\nu(r) < \nu(x)$. Si $r \neq 0$, alors r est inversible par définition de x. Finalement, modulo (x), a est égal à 0 ou à un élément inversible. □

Appliquons ce résultat. On va noter $\alpha = \mathbf{Z}\left[\frac{1+i\sqrt{19}}{2}\right]$ et $\overline{\alpha}$ son conjugué. Remarquons que :

$$\alpha + \overline{\alpha} = 1 \quad \text{et} \quad \alpha\overline{\alpha} = 5$$

donc α vérifie l'équation $\alpha^2 - \alpha + 5 = 0$. En tant que sous-anneau de $\mathbf{C}$, $\mathbf{Z}[\alpha]$ est intègre et stable par conjugaison puisque $\overline{\alpha} = 1 - \alpha$. On définit pour $z \in \mathbf{Z}[\alpha]$ ce qu'on appellera *norme* :

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

L'objectif est le calcul des inversibles. Soit $z \in \mathbf{Z}[\alpha]^\times$, alors :

$$1 = N(1) = N(zz^{-1}) = N(z)N(z^{-1})$$

donc $N(z) = 1$ (c'est un inversible de $\mathbf{N} \subset \mathbf{Z}$) et on a la relation :

$$a^2 + ab + 5b^2 = 1 \ \text{ où } \ z = a + b\alpha.$$

Or, on voit que :

$$b^2 + a^2 + ab \geq b^2 + a^2 - |ab| \geq (|b| - |a|)^2 \geq 0$$

donc

$$1 = a^2 + ab + 5b^2 \geq 4b^2$$

de sorte que $b = 0$ et $a = \pm 1$. Finalement, $\mathbf{Z}[\alpha]^\times = \{\pm 1\}$.

Si cet anneau était euclidien, il existerait $x \in \mathbf{Z}[\alpha]$ tel que $\mathbf{Z}[\alpha]/(x)$ soit un corps à 2 ou 3 éléments. D'où un homomorphisme :

$$\varphi : \mathbf{Z}[\alpha] \longrightarrow \mathbf{K} \ \text{ avec } \ \mathbf{K} = \mathbf{F}_2 \ \text{ ou } \ \mathbf{K} = \mathbf{F}_3$$

dont la restriction à $\mathbf{Z}$ est la projection canonique. Alors $\beta = \varphi(\alpha)$ vérifie $\beta^2 - \beta + 5 = 0$ dans $\mathbf{K}$. Que ce soit dans $\mathbf{F}_2$ ou dans $\mathbf{F}_3$, il n'y a pas de solution. C'est absurde.

Proposition. *Soient $a, b \in \mathbf{Z}[\alpha] \notin \{0\}$. Il existe $q, r \in \mathbf{Z}[\alpha]$ avec :*

(i) $r = 0$ ou $N(r) < N(b)$

(ii) $a = bq + r$ *ou* $2a = bq + r$.

Preuve. Soit $x = \frac{a}{b} = \frac{a\bar{b}}{b\bar{b}} \in \mathbf{C}$ que l'on écrit $x = u + v\alpha$ avec $u, v \in \mathbf{Q}$. On note $n = \lfloor v \rfloor$. On distingue :

1. Si $v \notin]n + 1/3, n + 2/3[$, alors, soient s et t les entiers les plus proches de u et v respectivement, on a en posant $q = s + t\alpha$:

$$N(x - q) = (s - u)^2 + (s - u)(t - v) + 5(t - v)^2 \leq \frac{1}{4} + \frac{1}{2} \times \frac{1}{3} + \frac{5}{9} = \frac{35}{36} < 1.$$

   et on a le résultat voulu en posant $r = a - bq = b(x - q)$.

2. Dans l'autre cas, on note que $2v \in]2n + 2/3, 2n + 1 + 1/3[$. Ainsi, en posant $m = \lfloor 2v \rfloor$, on a $2v \notin]m + 1/3, m + 2/3[$ et on est ramené au cas précédent : on a $2a = bq + r$ avec $N(r) < N(b)$.

□

On montre maintenant que $\mathbf{Z}[\alpha]$ est principal. On commence par voir que l'idéal (2) est maximal car

$$\mathbf{Z}[\alpha] \simeq \mathbf{Z}[T]/(T^2 - T + 5)$$

donc en vertu du théorème d'isomorphisme :

$$\mathbf{Z}[\alpha]/(2) \simeq (\mathbf{Z}/2\mathbf{Z})[T]/(T^2 + T + 1)$$

et la conclusion puisque $T^2 + T + 1$ est irréductible sur $\mathbf{F}_2$.

Ensuite, on prend $I \neq \{0\}$ un idéal de $\mathbf{Z}[\alpha]$ et $a \in I$ non nul de norme minimale. Si $I = (a)$ c'est fini, sinon prenons $x \in I \notin (a)$. On écrit la pseudo division euclidienne :

(i) Si $x = aq + r$ avec $N(r) < N(a)$ on a $r = 0$ et $x \in (a)$ c'est absurde.

(ii) Si $2x = aq + r$ avec $N(r) < N(a)$ on a $r = 0$ et $2x = aq$. Comme (2) est maximal donc premier, on a $a \in (2)$ ou $q \in (2)$. Le deuxième cas est impossible car il entrainerait $x \in (a)$. Donc $a = 2a'$ et $x = a'q \in (a')$. Mais (2) est maximal et ne contient pas q donc $(2, q) = \mathbf{Z}[\alpha]$ de sorte qu'on peut écrire

$$\lambda 2 + \mu q = 1 \quad \lambda, \mu \in \mathbf{Z}[\alpha].$$

On en déduit :

$$a' = \lambda 2a' + \mu q a' = \lambda a + \mu x$$

donc $a' \in I$ et on a une contradiction car $N(a') < N(2a') = N(a)$.

Référence. D. Perrin, *Cours d'Algèbre*

122 Anneaux principaux. Applications. Youhouu

4.5 Une équation aux dérivées partielles linéaire sans solution

On va dire que c'est culturel.
C'est le même John que dans John-Loewner je crois.

Théorème (Lewy)**.** *Il existe une fonction $F(x,y,t) \in C^\infty(\mathbf{R}^3)$ telle que l'équation :*

$$Lu := \partial_x u + \partial u - 2i(x+iy)\partial_t u = F \tag{4.4}$$

n'admette pas de solution C^1 de dérivée Hölderienne, sur aucun ouvert $\Omega \in \mathbf{R}^3$.

Preuve. On commence par un lemme :

Lemme. *Soit $\psi(t) \in C^\infty(\mathbf{R})$ à valeurs réelles. Si u est une solution C^1 de l'équation $Lu = \psi'(t)$ sur un voisinage ouvert Ω de l'origine, alors ψ est analytique en 0.*

Preuve. Si u est une telle solution, on note $x+iy = z$ et on pose pour $r < R$ et $|t| < R$, où R est petit :

$$V(r,t) = \int_{|z|=r} u(x,y,t)\,dz = i\int_0^{2\pi} u(r\cos\theta, r\sin\theta, t) r e^{i\theta} d\theta$$

Par la formule de Green, on a :

$$V(r,t) = i\int_{|z|\le r} (\partial_x u + i\partial_y u)(x,y,t)dx\,dy = i\int_0^r\int_0^{2\pi} (\partial_x u + i\partial_y u)(\rho\cos\theta, \rho\sin\theta, t)\rho d\rho d\theta.$$

De sorte qu'en dérivant :

$$\frac{\partial V}{\partial r} = i\int_0^{2\pi} (\partial_x u + i\partial_y u)(r\cos\theta, r\sin\theta, t) r d\theta = \int_{|z|=r} (\partial_x u + i\partial_y u)(x,y,t) r\frac{dz}{z}.$$

En posant $s = r^2$, on trouve finalement :

$$\begin{aligned}\frac{\partial V}{\partial s} &= \frac{1}{2r}\frac{\partial V}{\partial r} = \int_{|z|=r} (\partial_x u + i\partial_y u)(x,y,t)\frac{dz}{2z} \\ &= i\int_{|z|=r} \partial_t u(x,y,t)dz + \int_{|z|=r} \psi'(t)\frac{dz}{2z} = i\frac{\partial V}{\partial t} + \pi i\psi'(t).\end{aligned}$$

De sorte que la fonction $U(t,s) = V(s,t) + \pi\psi(t)$ vérifie l'équation de Cauchy-Riemann :

$$\frac{\partial U}{\partial t} + i\frac{\partial U}{\partial s} = 0.$$

C'est donc une fonction holomorphe sur l'ouvert $\{0 < s < R^2,\ |t| < R\}$, continue sur le bord $s = 0$ et telle que $U(t,0) = \pi\psi(t) \in \mathbf{R}$. En posant $U(t,-s) = -\overline{U(t,s)}$, on vient de construire une fonction holomorphe sur un voisinage de 0. En particulier, $U(t,0) = \pi\psi(t)$ est analytique. □

Le même argument montre que pour tout $(x_0, y_0, t_0) \in \mathbf{R}^3$, l'existence d'une solution C^1 sur un voisinage de (x_0, y_0, t_0) pour

$$Lu(x, y, t) = \psi'(t + 2y_0x - 2x_0y)$$

implique l'analyticité de ψ en t_0.

À partir de maintenant, on considère une fonction périodique $\psi \in C^\infty$ qui n'est analytique nulle part et on note $\{Q_j = (x_j, y_j, t_j)\}_{j\in\mathbf{N}}$ une famille dénombrable dense dans $\mathbf{R}^3$. Comme ψ est périodique (donc toutes ses dérivées sont majorées), en prenant $c_j = 2^{-j}\exp(-|x_j| - |y_j|)$, on voit que la série :

$$F_\varepsilon(x, y, t) = \sum_{j\in\mathbf{N}} \varepsilon_j c_j \psi'(t + 2y_jx - 2x_jy)$$

définit une fonction C^∞ sur $\mathbf{R}^3$ pour toute suite $(\varepsilon_j)_j \in \ell^\infty$ avec

$$|F_\varepsilon| \leq M\|\varepsilon\|.$$

On va montrer qu'il existe un $\varepsilon \in \ell^\infty$ pour lequel le problème

$$Lu = F_\varepsilon \tag{4.5}$$

n'a pas de solution C^1 de dérivées Hölderiennes.

On commence par définir pour $j \in \mathbf{N}$:

$$\begin{aligned} E_j &= \big\{\varepsilon \in \ell^\infty, \text{ le problème (4.5) a une solution locale } u \text{ sur un voisinage de } Q_j \text{ avec } u(Q_j) = 0\big\} \\ &= \bigcup_{n\in\mathbf{N}} E_{j,n} \end{aligned}$$

où $E_{j,n}$ est l'ensemble des $\varepsilon \in \ell^\infty$ tels qu'il existe $u \in C^1\big(B(Q_j, n^{-1/2})\big)$ vérifiant :

$$u(Q_j) = 0$$

$$|\partial^\alpha u| \leq n, \ \ |\alpha| = 1$$

$$\forall P, Q \in \mathbf{R}^3, \ |\partial^\alpha u(P) - \partial^\alpha u(Q)| \leq n|P - Q|^{1/n}, \ \ |\alpha| = 1$$

Lemme. *Les $E_{j,n}$ sont des fermés d'intérieur vide.*

Preuve. Soit $(\varepsilon^k)_k$ une suite d'éléments de $E_{j,n}$ qui converge vers $\varepsilon \in \ell^\infty$. Alors F_{ε^k} converge uniformément vers F_ε. On note (u_k) la suite des solutions associées aux (ε^k). Comme ces fonctions et leurs premières dérivées sont équi-continues, on peut extraire une sous suite qui converge vers u ainsi que ses premières dérivées (Ascoli). Cette solution vérifie $Lu = F_\varepsilon$ et toutes les propriétés voulues. Donc $E_{j,n}$ est fermé. Montrons par l'absurde qu'il est d'intérieur vide : si ε est dans l'intérieur de $E_{j,n}$, alors, en notant :

$$\delta = (0, \ldots, 0, 1/c_j, 0, \ldots) \in \ell^\infty$$

la suite

$$\varepsilon' = \varepsilon + \theta\delta$$

est aussi dans $E_{j,n}$ pour θ assez petit. Si u et u' sont les solutions associées à ε et ε', alors $u'' = (u' - u)/\theta$ vérifie :

$$Lu'' = F_\delta = \psi'(t_j + 2y_jx - 2x_jy)$$

et le premier lemme contredit la non analyticité de ψ en t_j. □

On peut conclure. Par l'absurde, s'il existait une solution $u \in C^1(\Omega)$ au problème (4.4) avec une donnée F_ε sur un ouvert $\Omega \in \mathbf{R}^3$, alors par densité et quitte à poser $u - u(Q_j)$, il existe $j \in \mathbf{N}$ et $n \in \mathbf{N}$ tels que $\varepsilon \in E_{j,n}$. Autrement dit :

$$\ell^\infty \subset \bigcup_{j,n \in \mathbf{N}} E_{j,n}.$$

C'est impossible par le théorème de Baire puisque ℓ^∞ est complet et ne peut pas s'écrire comme une réunion dénombrable de fermés d'intérieur vide. □

Quelques remarques complémentaires.

- Il y a plein de fonctions (périodiques) qui sont de classe $C^\infty(\mathbf{R})$ mais analytiques nulle part. En voici deux exemples (resp. [John] et Wikipédia) :

$$f(x) = \sum_{n=1}^{+\infty} \frac{\cos(n!x)}{(n!)^n} \quad \text{et} \quad g(x) = \sum_{j \in \mathbf{N}} e^{-\sqrt{2^j}} \cos(2^j x).$$

  Voir une condition nécéssaire et suffisante d'analyticité dans [John].
- La condition d'Hölderiennité est artificielle : Hartmann a montré qu'elle était inutile.
- Le défaut d'analyticité est essentiel : en fait, si les données sont analytiques, le théorème de Cauchy-Kowaleski assure l'existence d'une solution locale.

Références.

F. John, *Partial Differential Equations*
G. Folland, *Introduction to Partial Differential Equations, Second Edition*

205 Espaces complets. Exemples et applications.
215 Applications différentiables définies sur un ouvert de $\mathbf{R}^n$. Exemples et applications.
222 Exemples d'équations aux dérivées partielles linéaires.
241 Suites et séries de fonctions. Exemples et contre-exemples.
245 Fonctions holomorphes sur un ouvert de $\mathbf{C}$. Exemples et applications.

5 Leçon ? Développements.

101 Groupe opérant sur un ensemble. Exemples et applications.

- Quaternions et rotations
- Sous-groupes compacts de $GL_n(\mathbf{R})$
- Réduction de Jordan

102 Groupe des nombres complexes de module 1. Sous-groupes des racines de l'unité. Applications.

- Polygones réguliers constructibles
- Transformée de Fourier Rapide

103 Exemples de sous-groupes distingués et de groupes quotients. Applications.

- Classification des groupes d'ordre 12
- Sous-groupes distingués et noyaux de caractères

104 Groupes finis. Exemples et applications.

- Structure des groupes abéliens finis
- Classification des groupes d'ordre 12

105 Groupe des permutations d'un ensemble fini. Applications.

- Cartes
- Bruhat
- Structure des polynômes symétriques

106 Groupe linéaire d'un espace vectoriel de dimension finie E, sous-groupes de GL(E). Applications.

- Cartan-Von Neumann
- Sous-groupes compacts de $GL_n(\mathbf{R})$

107 Représentations et caractères d'un groupe fini sur un C-espace vectoriel. Exemples.

- Structure des groupes abéliens finis
- Sous-groupes distingués et noyaux de caractères

108 Exemples de parties génératrices d'un groupe. Applications.

- Groupes paveurs
- Quaternions en rotations
- $SO_3(\mathbf{R})$ est simple

110 Caractères d'un groupe abélien fini et transformée de Fourier discrète. Applications.

- Transformée de Fourier Rapide
- Structure des groupes abéliens finis

120 Anneaux $\mathbf{Z}/n\mathbf{Z}$. Applications.

- Chevalley-Warning et Erdös-Ginzburg-Ziv
- Théorème des deux carrés

121 Nombres premiers. Applications.

- Polygones réguliers constructibles
- Chevalley-Warning et Erdös-Ginzburg-Ziv
- Théorème des deux carrés

122 Anneaux principaux. Applications.

- Théorème des deux carrés
- Endomorphismes semi-simples
- L'anneau moche

123 Corps finis. Applications.

- Chevalley-Warning et Erdös-Ginzburg-Ziv
- Irréductibles de $\mathbf{F}_q$

125 Extensions de corps. Exemples et applications.

- Polygones réguliers constructibles
- Lemme d'Artin
- Irréductibles de $\mathbf{F}_q$

126 Exemples d'équations diophantiennes.

RIEN DU TOUT HAHAHA

141 Polynômes irréductibles à une indéterminée. Corps de rupture. Exemples et applications.

- Irréductibles de $\mathbf{F}_q$

- Endomorphismes semi-simples

142 Algèbre des polynômes à plusieurs indéterminées. Applications.

- Chevalley-Warning et Erdös-Ginzburg-Ziv
- Structure des polynômes symétriques

144 Racines d'un polynôme. Fonctions symétriques élémentaires. Exemples et applications

- Chevalley-Warning et Erdös-Ginzburg-Ziv
- Structure des polynômes symétriques

150 Exemples d'actions de groupes sur les espaces de matrices.

- Réduction de Jordan
- Sous-groupes compacts de $GL_n(\mathbf{R})$

151 Dimension d'un espace vectoriel (on se limitera au cas de la dimension finie). Rang. Exemples et applications.

- Invariants de similitude
- Cayley-Meger
- Lemme d'Artin
- Polygones réguliers constructibles

152 Déterminant. Exemples et applications.

- Cayley-Meger
- John-Loewner

153 Polynômes d'endomorphisme en dimension finie. Réduction d'un endomorphisme en dimension finie. Applications.

- Invariants de similitude
- Décomposition de Dunford-Newton
- Endomorphismes semi-simples
- Réduction de Jordan

154 Sous-espaces stables par un endomorphisme ou une famille d'endomorphismes d'un espace vectoriel de dimension finie. Applications.

- Invariants de similitude
- Décomposition de Dunford-Newton
- Endomorphismes semi-simples

155 Endomorphismes diagonalisables en dimension finie.

- Décomposition de Dunford-Newton

- Endomorphismes semi-simples

156 Exponentielle de matrices. Applications.

1. Exponentielle est surjective
2. Cartan-Von Neumann

157 Endomorphismes trigonalisables. Endomorphismes nilpotents.

- Réduction de Jordan
- Décomposition de Dunford-Newton

158 Matrices symétriques réelles, matrices hermitiennes.

- Méthodes de gradient
- $\mathcal{O}(p,q)$

159 Formes linéaires et dualité en dimension finie. Exemples et applications.

- Krein-Milman
- Invariants de similitude

160 Endomorphismes remarquables d'un espace vectoriel euclidien (de dimension finie).

- Sous-groupes compacts de $GL_n(\mathbf{R})$
- John-Loewner

161 Isométries d'un espace affine euclidien de dimension finie. Applications en dimensions 2 et 3.

- Quaternions et rotations
- Groupes paveurs

162 Systèmes d'équations linéaires ; opérations élémentaires, aspects algorithmiques et conséquences théoriques.

- Méthodes de gradient Lemme d'Artin

170 Formes quadratiques sur un espace vectoriel de dimension finie. Orthogonalité, isotropie. Applications.

- $\mathcal{O}(p,q)$
- John-Loewner

171 Formes quadratiques réelles. Coniques. Exemples et applications.

- $\mathcal{O}(p,q)$

- John-Loewner

181 Barycentres dans un espace affine réel de dimension finie, convexité. Applications.

- Krein-Milman
- Sous-groupes compacts de $GL_n(\mathbf{R})$

182 Applications des nombres complexes à la géométrie.

- Quaternions et rotations
- Polygones réguliers constructibles

183 Utilisation des groupes en géométrie.

- Polygones réguliers constructibles
- Quaternions et rotations

190 Méthodes combinatoires, problèmes de dénombrement.

- Cartes
- Irréductibles de $\mathbf{F}_q$
- Couples de nombres premiers entre eux

201 Espaces de fonctions ; exemples et applications.

- Bargmann
- Théorème de Müntz
- Théorème de Morgenstern
- Théorème de Lebesgue et Rademacher

202 Exemples de parties denses et applications.

- Critère de Kitai
- Théorème de Müntz

203 Utilisation de la notion de compacité.

- Des bases presque orthogonales
- John-Loewner
- Sous-groupes compacts de $GL_n(\mathbf{R})$

204 Connexité. Exemples et applications.

- Exponentielle est surjective
- Théorème de relèvement

205 Espaces complets. Exemples et applications.

- Critère de Kitai
- Riesz-Fischer
- Théorème de Morgenstern

207 Prolongement de fonctions. Exemples et applications.

- Théorème de relèvement
- Prolongement de la fonction ζ
- Théorèmes abéliens et taubériens

208 Espaces vectoriels normés, applications linéaire continues. Exemples.

- Des bases presque orthogonales
- Théorème ergodique de Von Neumann

209 Approximation d'une fonction par des polynômes et des polynômes trigonométriques. Exemples et applications.

- Équation de la chaleur
- Théorème de Müntz

213 Espaces de Hilbert. Bases hilbertiennes. Exemples et applications.

- Bargmann
- Des bases presque orthogonales
- Théorème ergodique de Von Neumann

214 Théorème d'inversion locale, théorème des fonctions implicites. Exemples et applications en analyse et en géométrie.

- Cartan-Von Neumann
- Le pendule de Van Der Pol
- Exponentielle est surjective

215 Applications différentiables définies sur un ouvert de $\mathbf{R}^n$. Exemples et applications.

- Cartan-Von Neumann
- Théorème de Lebesgue et Rademacher

218 Applications des formules de Taylor.

- Problème et méthode des moments
- Analyse du θ-schéma
- Méthode de Laplace

219 Extrema : existence, caractérisation, recherche. Exemples et applications.

- Méthodes de gradient
- John-Loewner

220 Équations différentielles $X' = f(t, X)$. Exemples d'étude des solutions en dimension 1 et 2.

- Le pendule de Van Der Pol
- Stabilité et instabilité

221 Équations différentielles linéaires. Système d'équations différentielles linéaires. Exemples et applications.

- Le pendule de Van Der Pol
- Stabilité et instabilité

222 Exemples d'équations aux dérivées partielles linéaires.

- Équation de Schrödinger linéaire
- Analyse du θ-schéma
- Équation de la chaleur

223 Suites numériques. Convergence, valeurs d'adhérence. Exemples et applications.

- Processus de Galton-Watson
- Théorèmes abéliens et taubériens

224 Exemples de développements asymptotiques de suites et de fonctions.

- Problème et méthode des moments
- Méthode de Laplace

226 Suites vectorielles et réelles définies par une relation de récurrence $u_{n+1} = f(u_u)$. Exemples. Applications à la résolution approchée d'équations.

- Critère de Kitai
- Méthodes de gradient

228 Continuité et dérivabilité des fonctions réelles d'une variable réelles. Exemples et applications.

- Théorème de Morgenstern
- Théorème de Lebesgue et Rademacher

229 Fonctions monotones. Fonctions convexes. Exemples et applications.

- Processus de Galton-Watson
- John-Loewner

- Théorème de Lebesgue et Rademacher

230 Séries de nombres réels ou complexes. Comportement des restes ou des sommes partielles des séries numériques. Exemples.

- Couples de nombres premiers entre eux
- Théorèmes abéliens et taubériens

233 Méthodes itératives en analyse numérique matricielle.

- Méthodes de gradient
- Analyse du θ-schéma

234 Espaces L^p, $1 \leq p \leq +\infty$.

- Riesz-Fischer
- Bargmann

235 Problèmes d'interversion de limites et d'intégrales.

- Prolongement de la fonction ζ
- Théorèmes abéliens et taubériens
- Équation de la chaleur

236 Illustrer par des exemples quelques méthodes de calcul d'intégrales de fonctions d'une ou plusieurs variables.

- Bargmann
- Inversion de la fonction caractéristique

239 Fonctions définies par une intégrale dépendant d'un paramètre. Exemples et applications.

- Prolongement de la fonction ζ
- Méthode de Laplace

241 Suites et séries de fonctions. Exemples et contre-exemples.

- Prolongement de la fonction ζ
- Une construction du mouvement Brownien

243 Convergence des séries entières, propriétés de la somme. Exemples et applications.

- Processus de Galton-Watson
- Théorème de Morgenstern
- Théorèmes abéliens et taubériens
- Bargmann

245 Fonctions holomorphes sur un ouvert de **C**. Exemples et applications.

- Théorème de Müntz
- Prolongement de la fonction ζ

246 Séries de Fourier. Exemples et applications.

- Équation de la chaleur
- Analyse du θ-schéma

250 Transformation de Fourier. Applications.

- Équation de Schrödinger linéaire
- Bargmann

253 Utilisation de la notion de convexité en analyse.

- Krein-Milman
- John-Loewner

260 Espérance, variance et moments d'une variable aléatoire.

- Processus de Galton-Watson
- Une construction du mouvement Brownien

261 Fonction caractéristique d'une variable aléatoire. Exemples et applications.

- Inversion de la fonction caractéristique
- Problème et méthode des moments

262 Modes de convergence d'une suite de variables aléatoires. Exemples et applications.

- Cartes
- Problème et méthode des moments

263 Variables aléatoires à densité. Exemples et applications.

- Une construction du mouvement Brownien
- Inversion de la fonction caractéristique

264 Variables aléatoire discrètes. Exemples et applications.

- Cartes
- Processus de Galton-Watson