

[accueil](#)

Fichier Pdf

L'axiomatisation et l'algébrisation de la mathématique qui, prétend-on, se poursuivent depuis déjà plus de 50 ans ont conduit à rendre illisibles un si grand nombre de textes mathématiques que la menace (qui a toujours pesé) de voir les mathématiques perdre contact avec la physique et les sciences naturelles est devenue réelle.

Vladimir Arnold, 1996

Analyse de Hilbert

1. Produit scalaire
2. Polynômes orthogonaux
 - 2-a : Relation de récurrence
 - 2-b : Formule de Christoffel-Darboux
 - 2-c : Noyau
3. Le théorème de Weierstrass
4. Les zéros des polynômes orthogonaux (1)
5. Approximation
 - 5-a : Théorème
 - 5-b : Formule de quadrature de Gauss
 - 5-c : Nombres de Christoffel
 - 5-d : Evaluation de l'erreur
 - 5-e : Formule d'interpolation d'Hermite
6. Les zéros des polynômes orthogonaux (2)
7. Espace de Hilbert
8. Représentations conformes
9. Un exemple : le produit hermitien et les inégalités de Heisenberg
 - 9-a : Produit scalaire hermitien
 - 9-b : Inégalités de Heisenberg
10. Equations intégrales
 - 10-a : Aperçu de la question
 - 10-b : La méthode de Fredholm
11. Noyaux de Fourier
 - 11-a : Transformée de Mellin
 - 11-b : Exemples de noyaux de Fourier
 - 11-c : Formules d'inversion non symétriques
12. Convolution



1. Introduction et bibliographie

De très nombreux résultats et méthodes ont vu le jour au fil des siècles dans les domaines de l'analyse numérique, de la résolution des équations différentielles, de l'approximation, etc. Quand on lit les divers ouvrages proposés dans la bibliographie, dont les dates de parution s'échelonnent sur à peine un siècle, on voit la diversité des approches et des motivations : untel sera plus intéressé par le côté purement mathématique, tel autre par la résolution d'équations aux dérivées partielles, tel autre par l'approximation, etc.

Arriver à se faire une idée un peu claire de la chose n'est pas très facile, et c'est ce que nous allons essayer de faire malgré tout. Dans un premier temps nous faisons quelques aperçus théoriques en évitant de développer trop le côté purement mathématique (le lecteur intéressé peut lire Schwartz par exemple), puis nous nous intéresserons au côté pratique de la chose avec divers types de polynômes orthogonaux (Legendre, Laguerre, Hermite), enfin nous terminerons par un bref aperçu sur la théorie de l'approximation avec les splines et les nurbs.

Comme pour la plupart des textes présents sur le site la plus grande place est faite aux applications, mais ce ne sera malheureusement pas le cas dans cette première partie.

Je ne résiste pas néanmoins à citer l'Universalis

*« L'enseignement de ces questions en France est conçu selon un plan rigide : étude des modes de convergence dans un cadre abstrait, validation des opérations sur les séries et les intégrales, représentation des fonctions et, enfin, résolution de problèmes. Cette démarche est **contraire** à la pratique scientifique où l'approfondissement théorique des modes de représentation et d'approximation va de pair avec l'étude des problèmes visés.*

*Les méthodes de représentation et d'approximation jouent un **rôle central** dans l'analyse mathématique. Elles sont présentées de façon synthétique dans les cinq premiers chapitres, qui renvoient pour plus de détails sur chacune des méthodes décrites aux divers articles d'analyse de l'Encyclopédie. Dans les trois derniers chapitres, nous approfondissons les problèmes d'approximation en abordant notamment les questions de stabilité et de vitesse de convergence, spécialement utiles en analyse numérique. »*

Jean-Louis Ovaert, Jean-Luc Verley

© Encyclopædia Universalis 2004, tous droits réservés

Bibliographie

N. Boccara, *Distributions*, Ellipses, 1996

Un petit livre pas très épais mais très clair et facile à lire. Recommandé pour débiter.

A. Angot, *Compléments de Mathématiques*, Masson et Cie, 1970

Très clair et succinct, amplement suffisant jusqu'en Spé. On le trouve parfois d'occasion. Bien complet et des applications intéressantes à l'Electricité.

H. Hochstadt, *Les fonctions de la physique mathématique*, Masson et Cie, 1973

Epuisé, mais très intéressant ; les calculs sont parfois rudes mais méritent de s'y attarder.

W. Appel, *Mathématiques pour la physique et les physiciens*, H&K éd., 2002

Passe un peu vite sur pas mal de trucs, mais reste compréhensible.

- G. Demengel, *Transformations de Laplace*, Ellipses, 2002
Malgré une présentation un peu fouillis on y trouve l'essentiel à connaître sur ce genre de questions.
- E. T. Whittaker & G. N. Watson, *A modern course of analysis*, Cambridge University Press, 1927 (2003)
Bien que ce soit en anglais c'est assez facile à lire. La référence est une réimpression de la 4^{ème} édition datée de 1927. Le livre par lui-même contient beaucoup de choses autres.
- I. N. Sneddon, *Fourier Transforms*, Dover, New-York, 1951 (1995)
Beaucoup d'applications à des domaines variés. Explications assez claires même s'il faut un bon niveau parfois.
- L. Schwartz, *Analyse hilbertienne*, Hermann, Paris, 1979
Assez théorique, mais reste lisible avec un niveau Bac+3.
- G. Valiron, *Equations fonctionnelles*, Masson et Cie, Paris, 1950 (rééd. J. Gabay, 1989)
Ca a pas mal vieilli, mais on trouve quand même l'essentiel sur la résolution des équations différentielles et les principales méthodes.
- J. E. Rombaldi, *Interpolation et Approximation*, Vuibert, 2005
Bon petit ouvrage sur l'approximation et les noyaux.
- G. Demengel & J.P. Pouget, *Modèles de Bézier, des B-splines et des NURBS*, Ellipses, 1998
Très utilisable et très concret.
- M. Attéia & J. Gaches, *Approximation hilbertienne*, EDP sciences, 1999
La première partie sur les splines est illisible. La partie ondelettes est plus sympa.
- B. Burke Hubbard, *Ondes et ondelettes*, Belin, 1995
Vulgarisation sur les ondelettes, mais touche à pas mal de choses. Intéressant et très lisible.
- Encyclopédie Universalis, article *Fonctions (représentation et approximation)* et ceux associés, ed. électronique, 2005.
L'ensemble est assez bon. Peut servir de base.

Sur Internet, pas mal de choses...voir par exemple :

Quelques applications : <http://www.unice.fr/DeptPhys/pilot/node1.html>

A propos de la théorie de la chaleur :

<http://www.sciences.univ-nantes.fr/physique/perso/blanquet/conducti/cddex.htm>

Une autre page sur la Chaleur, mais en plus complet :

<http://gershwin.ens.fr/vdaniel/Doc-Locale/Cours-Mirrored/Operateurs-Differentiels/www.chez.com/touslescours/math/cours/opdiff/node1.html>

Pour une tripotée de formules et de liens on consultera évidemment :

<http://mathworld.wolfram.com/BesselFunctionoftheFirstKind.html>

ainsi que tous les liens du site.

Beaucoup de choses intéressantes sur <http://www.sciences.ch/htmlfr/introduction.php>, on profitera également des liens vers de nombreux textes disponibles en pdf.

Cours d'analyse numérique : <http://cel.ccsd.cnrs.fr/cours/cel-19/numeri.pdf>

Equations aux dérivées partielles :

<http://w3-phystheo.ups-tlse.fr/~robert/licadmin/ul1b/www/1er.03-04/ced2/ced2.html>

Un peu de Maple : <http://www.lptl.jussieu.fr/users/viot/td.html>

Transformée de Hankel : <http://www.journals.cms.math.ca/cgi-bin/vault/public/view/betancor8407.prepub/body/PDF/betancor8407.prepub.pdf?file=betancor8407.prepub>

Beaucoup de choses ici, mais en fouillant : <http://www.mathpages.com/home/index.htm>

2. Produit scalaire

On définit en général le produit scalaire (*produit scalaire hermitien* ou *p.s. intérieur*) $\langle f, g \rangle$ dans les fonctions à valeurs dans \mathbb{C} par les propriétés suivantes :

1. $\langle f, g \rangle = \overline{\langle g, f \rangle}$
2. $\langle \alpha f + \beta g, h \rangle = \alpha \langle f, h \rangle + \beta \langle g, h \rangle$
3. $\langle f, f \rangle = 0 \Leftrightarrow f = 0$;

pour le p.s. ainsi défini on définit la norme de f par $\|f\| = \langle f, f \rangle^{1/2}$ qui possède alors les propriétés suivantes :

1. $\|f\| = 0 \Leftrightarrow f = 0$, $\|f\| \geq 0$
2. $\|\alpha f\| = |\alpha| \|f\|$
3. $\|f + g\| \leq \|f\| + \|g\|$

(inégalité triangulaire : l'égalité a lieu si et seulement si les vecteurs f et g sont indépendants).

Cette dernière inégalité résulte de la propriété suivante : $|\langle f, g \rangle| \leq \|f\| \|g\|$; posons $\langle f, g \rangle = |\langle f, g \rangle| e^{i\theta}$

et $\alpha = \frac{|\langle f, g \rangle|}{\|f\|^2} e^{-i\theta}$; effectuons l'opération suivante :

$$\langle \alpha f - g, \alpha f - g \rangle \|f\|^2 = \|f\|^2 \|g\|^2 - |\langle f, g \rangle|^2$$

qui est évidemment positif. L'inégalité précédente est une égalité si $\alpha f - g = 0$ donc si f et g sont linéairement dépendants. La réciproque est immédiate avec les propriétés du p.s.

Deux fonctions f et g seront orthogonales si et seulement si leur p.s. est nul, de même une famille de fonctions $(f_n)_{n \in \mathbb{N}}$ sera orthogonale si toutes les fonctions f_n sont orthogonales deux à deux. La famille sera orthonormale si les normes de tous les vecteurs valent 1. L'ensemble des combinaisons linéaires de la famille $(f_n)_{n \in \mathbb{N}}$ constitue évidemment un espace vectoriel de dimension infinie (ou finie si le nombre de fonctions est fini...) dont les f_n constituent une base.

Si nous prenons une fonction f quelconque la question qui se pose est alors de savoir si elle peut s'exprimer dans la base $(f_n)_{n \in \mathbb{N}}$; pour ce faire il faut déterminer les coordonnées de f dans la base, coordonnées obtenues en faisant le p.s. de f avec tous les vecteurs de la base : en effet supposons que f s'écrive $f = \sum_{n \in \mathbb{N}} \alpha_n f_n$ alors $\langle f, f_i \rangle = \sum_{j \in \mathbb{N}} \alpha_j \langle f_j, f_i \rangle = \alpha_i \langle f_i, f_i \rangle = \alpha_i \|f_i\|^2$. Si la base est orthonormée

on récupère directement la coordonnée α_i .

Si ce n'est pas le cas, ce n'est pas très grave car nous pouvons construire une base orthonormée équivalente à la base $(f_n)_{n \in \mathbb{N}}$: considérons donc que nos vecteurs $(f_n)_{n \in \mathbb{N}}$ sont simplement

linéairement indépendants et posons $e_0 = \frac{f_0}{\|f_0\|}$ qui est évidemment normé ; prenons maintenant

$e_1 = \frac{f_1 - \langle f_1, e_0 \rangle e_0}{\|f_1 - \langle f_1, e_0 \rangle e_0\|}$ qui est normé et orthogonal à e_1 ; d'une manière générale si on prend

$e_k = \frac{f_k - \sum_{i=0}^{k-1} \langle f_k, e_i \rangle e_i}{\|f_k - \sum_{i=0}^{k-1} \langle f_k, e_i \rangle e_i\|}$ on obtient une famille orthonormée de vecteurs (la démonstration est

laissée aux bons soins du lecteur). C'est le procédé d'orthogonalisation de Gram-Schmidt. Ce

procédé de calcul est récurrent, mais on peut obtenir une expression directe des e_k en considérant le déterminant

$$d_k = \begin{vmatrix} \langle f_1, f_1 \rangle & \langle f_1, f_2 \rangle & \dots & \dots & \langle f_1, f_k \rangle \\ \langle f_2, f_1 \rangle & \langle f_2, f_2 \rangle & \dots & \dots & \langle f_2, f_k \rangle \\ \dots & \dots & \dots & \dots & \dots \\ \langle f_k, f_1 \rangle & \langle f_k, f_2 \rangle & \dots & \dots & \langle f_k, f_k \rangle \\ f_1 & f_2 & \dots & \dots & f_k \end{vmatrix}$$

dont le p.s. avec chaque $e_j, j < k$ donne 0 et $\langle e_k, d_k \rangle = \sum_{j=0}^k \beta_{j,k} \langle f_j, d_k \rangle$; comme les lignes k et j sont

identiques, on a $e_k = \frac{d_k}{\|d_k\|}$.

3. Polynômes orthogonaux

Ce que nous venons de faire ne fait pas intervenir le type des fonctions f , aussi intéressons nous à une famille de fonctions particulières : les polynômes réels ; nous les prenons définis sur un intervalle $[a, b]$ à priori quelconque que nous allons munir d'un p.s. : choisissons un polynôme quelconque $p(x)$ et une fonction **poïds** $w(x)$ pour laquelle l'intégrale $\int_a^b p(x)w(x)dx$ existe. Pour la suite des événements il vaut mieux que w soit strictement positive et si nous parlons d'intégrale au sens de Lebesgue, **presque partout** strictement positive. Nous introduisons maintenant notre produit scalaire de la manière suivante :

$$\langle f, g \rangle = \int_a^b f(x)g(x)w(x)dx \text{ pour des fonctions réelles.}$$

Si les fonctions sont complexes il faut alors prendre le produit hermitien

$$\langle f, g \rangle = \int_a^b f(x)\overline{g(x)}w(x)dx .$$

Vérifions que nous avons bien un p.s. : du fait que nous prenons des polynômes sur \mathbb{R} nous avons immédiatement $\langle f, g \rangle = \overline{\langle g, f \rangle} = \langle g, f \rangle$; de même la condition de linéarité est vérifiée. Pour la 3^{ème} condition $\langle f, f \rangle = \int_a^b f^2(x)w(x)dx$ qui est donc positif puisque w l'est ; de même comme $f^2(x)w(x) \geq 0$ l'intégrale ne peut être nulle que si f^2 est nulle ; la réciproque est immédiate.

Cherchons donc une famille de polynômes orthogonaux pour ce p.s. : nous prenons la famille linéairement indépendante $(1, x, x^2, \dots, x^n)$ que nous pouvons orthogonaliser avec la méthode de Gram-Schmidt ; par exemple avec $w(x) = 1$ et $[a, b] = [-1, 1]$ nous avons :

$$f_0 = 1 \text{ d'où } \|f_0\|^2 = \int_{-1}^1 1dx = 2 \text{ et } e_0 = \frac{1}{\sqrt{2}} ;$$

$$f_1 = x \text{ d'où } d_1 = \begin{vmatrix} \langle 1, 1 \rangle & \langle 1, x \rangle \\ 1 & x \end{vmatrix} = \begin{vmatrix} 2 & 0 \\ 1 & x \end{vmatrix} = 2x \text{ et } e_1 = \frac{2x}{\|2x\|} = \sqrt{\frac{3}{2}}x ;$$

$$f_2 = x^2 \text{ d'où } d_2 = \begin{vmatrix} \langle 1, 1 \rangle & \langle 1, x \rangle & \langle 1, x^2 \rangle \\ \langle x, 1 \rangle & \langle x, x \rangle & \langle x, x^2 \rangle \\ 1 & x & x^2 \end{vmatrix} = \begin{vmatrix} 2 & 0 & \frac{2}{3} \\ 0 & \frac{2}{3} & 0 \\ 1 & x & x^2 \end{vmatrix} = \frac{4}{3}x^2 - \frac{4}{9} \text{ et } e_2 = \sqrt{\frac{5}{8}}(3x^2 - 1) \dots$$

3-a : Relation de récurrence

La méthode constructive précédente est malheureusement un peu lourde et difficile à mettre en œuvre dès que n devient grand, aussi allons nous chercher une formule de récurrence pour les polynômes dans le cas général.

Il est immédiat que les polynômes e_n sont de degré n (à cause de d_n), aussi posons $e_n = k_n x^n + \dots + k_0$ et calculons

$$(1) \quad \langle x e_n, e_j \rangle = \int_a^b x e_n(x) e_j(x) w(x) dx = \langle e_n, x e_j \rangle = 0 \text{ pour } j \leq n-2$$

puisque e_j est au plus de degré $n-1$; calculons maintenant

$$x e_n = k_n x^{n+1} + \dots = \frac{k_n}{k_{n+1}} (k_{n+1} x^{n+1} + \dots) = \frac{k_n}{k_{n+1}} \left[e_{n+1} + \sum_{j=0}^n a_j e_j \right]$$

(les a_j sont les coordonnées du polynôme restant dans la base) ; posons $A_n = \frac{k_{n+1}}{k_n}$ et écrivons

$e_{n+1} - A_n x e_n = \sum_{j=0}^n a_j e_j$; effectuons le p.s. sur cette égalité :

$$(2) \quad \langle e_{n+1} - A_n x e_n, e_j \rangle = -A_n \langle x e_n, e_j \rangle = a_j \text{ pour } j \leq n.$$

Grâce à (1) nous avons alors $a_j = 0$ pour $j \leq n-2$ et $e_{n+1} - A_n x e_n = a_n e_n + a_{n-1} e_{n-1}$. Pouvons nous trouver une expression agréable de a_{n-1} ? (2) donne également

$$\begin{aligned} a_{n-1} &= -A_n \langle x e_n, e_{n-1} \rangle = -A_n \langle e_n, x e_{n-1} \rangle = \\ &= -A_n \left\langle e_n, \frac{1}{A_{n-1}} \left(e_n + \sum_{j=0}^{n-1} \alpha_j e_j \right) \right\rangle = -\frac{A_n}{A_{n-1}} \langle e_n, e_n \rangle = -\frac{A_n}{A_{n-1}}. \end{aligned}$$

Finalement nous pouvons écrire la relation de récurrence entre les polynômes orthogonaux :

$$e_{n+1} - (A_n x + a_n) e_n + B_n e_{n-1} = 0 \text{ avec } A_n = \frac{k_{n+1}}{k_n} \text{ et } B_n = \frac{A_n}{A_{n-1}}, B_0 = 0.$$

3-b : Formule de Christoffel-Darboux

Au même titre qu'en géométrie la distance liée au p.s. peut s'exprimer dans une base orthonormée par une expression de la forme $x^2 + y^2 + \dots$ on peut se demander ce que donnerait une expression de la forme $\sum_{j=0}^n e_j^2$, mais comme les e_j sont des polynômes on peut généraliser à des expressions de

la forme $K_n(x, y) = \sum_{j=0}^n e_j(x) e_j(y)$.

Nous avons pour l'instant $e_0 = k_0$ et $e_1 = k_1 x - k_1 k_0^2 \int_a^b x w(x) dx$ (revenir à la définition).

Utilisons la relation de récurrence sur les e_j : $e_{n+1}(x) - (A_n x + a_n) e_n(x) + B_n e_{n-1}(x) = 0$ où nous allons nous débarrasser des a_n ; $A_n x + a_n = \frac{e_{n+1}(x) + B_n e_{n-1}(x)}{e_n(x)}$; en réécrivant la même chose pour y et en soustrayant, nous tirons : $A_n(x-y) = \frac{e_{n+1}(x) + B_n e_{n-1}(x)}{e_n(x)} - \frac{e_{n+1}(y) + B_n e_{n-1}(y)}{e_n(y)}$, soit

$$e_n(x) e_n(y) = \frac{1}{A_n} \left(\frac{e_{n+1}(x) e_n(y) - e_{n+1}(y) e_n(x)}{x-y} \right) - \frac{B_n}{A_n} \left(\frac{e_n(x) e_{n-1}(y) - e_n(y) e_{n-1}(x)}{x-y} \right)$$

avec $B_n = \frac{A_n}{A_{n-1}}$ d'où

$$e_n(x)e_n(y) = \frac{1}{A_n} \left(\frac{e_{n+1}(x)e_n(y) - e_{n+1}(y)e_n(x)}{x-y} \right) - \frac{1}{A_{n-1}} \left(\frac{e_n(x)e_{n-1}(y) - e_n(y)e_{n-1}(x)}{x-y} \right).$$

Or nous sommes observateurs et nous remarquons que

$$e_n(x)e_n(y) = K_n(x, y) - K_{n-1}(x, y) ;$$

en comparant avec la relation précédente, nous concluons immédiatement que

$$K_n(x, y) = \frac{k_n}{k_{n+1}} \left(\frac{e_{n+1}(x)e_n(y) - e_{n+1}(y)e_n(x)}{x-y} \right).$$

Vous pouvez vérifier que cette relation est valable pour $n = 0$, la récurrence est alors établie. Si nous faisons tendre y vers x , nous obtenons donc en faisant une légère manipulation :

$$K_n(x, x) = \frac{k_n}{k_{n+1}} \lim_{y \rightarrow x} \frac{e_{n+1}(x)[e_n(y) - e_n(x)] - e_n(x)[e_{n+1}(y) - e_{n+1}(x)]}{x-y},$$

soit

$$K_n(x, x) = \frac{k_n}{k_{n+1}} [e_{n+1}(x)e'_n(x) - e_n(x)e'_{n+1}(x)] = \sum_{k=0}^n e_k^2(x) \geq 0.$$

3-c : **Noyau**

Fixons maintenant la valeur de y dans l'intervalle $[a, b]$ et étudions le polynôme

$$q(x) = \sum_{k=0}^n e_k(x)e_k(y).$$

* q est de degré n et s'écrit donc $q(x) = \sum_{k=0}^n \alpha_k e_k(x)$;

* d'autre part nous souhaitons avoir un polynôme normé, soit $\|q(y)\| = 1$ donc $\sum_{k=0}^n \alpha_k^2 = 1$;

* par ailleurs nous pouvons considérer que $q(y)$ est le p.s. de deux vecteurs de composantes (α_k) et $(e_k(y))$.

Utilisons la relation de Cauchy-Schwarz sur $q^2(y)$:

$$q^2(y) \leq \sum_{k=0}^n \alpha_k^2 \sum_{k=0}^n e_k^2(y) = \sum_{k=0}^n e_k^2(y),$$

l'égalité ayant lieu lorsque les vecteurs sont linéairement indépendants, auquel cas $q^2(y)$ est maximum.

Si nos deux vecteurs sont linéairement dépendants, il existe λ tel que $\alpha_k = \lambda e_k(y)$ pour tout k ,

d'où $q(x) = \lambda \sum_{k=0}^n e_k(x)e_k(y) = \lambda K_n(x, y)$; toujours pour avoir un polynôme normé nous écrivons :

$$\|q\| = 1 \Leftrightarrow \lambda^2 \sum_{k=0}^n e_k^2(y) = \lambda^2 K_n(y, y) = 1 \Leftrightarrow \lambda = \frac{\pm 1}{\sqrt{K_n(y, y)}} \text{ et finalement}$$

$$q(x) = \frac{\pm K_n(x, y)}{\sqrt{K_n(y, y)}}.$$

Que peut-on dire de $|q(y)|$ dans ce cas ?

Comme de toutes manières $q(y) = \sum_{k=0}^n e_k^2(y)$, c'est sa valeur maximale et les polynômes $q(x)$ précédents sont ceux qui rendent $|q(y)|$ maximal pour une valeur fixée de y .

Terminons ce paragraphe en remarquant que si p est un polynôme de degré inférieur ou égal à n , alors $p(x) = \sum_{k=0}^n \langle p, e_k \rangle e_k(x)$ (les coordonnées de p sont obtenues en faisant le p.s. de p et de chaque e_k puisque le p.s. consiste à projeter p sur chacun des e_k) ; de plus

$$\langle p(x), K_n(x, y) \rangle = \sum_{k=0}^n \langle p(x), e_k(x) e_k(y) \rangle = \sum_{k=0}^n \langle p(x), e_k(x) \rangle e_k(y) = \sum_{k=0}^n \langle p, e_k \rangle e_k(y) = p(y).$$

Le terme $K_n(x, y)$ est appelé *noyau régénérateur* ou *noyau hilbertien* ; en particulier si $p=1$ ($p(x)=p(y)=1$), nous obtenons $\langle 1, K_n(x, y) \rangle = \int_a^b K_n(x, y) w(x) dx = 1$.

4. Le théorème de Weierstrass

Nous avons vu dans le livre une approche du théorème à partir des pôlynômes de Bernstein sous une forme probabiliste. Améliorons ceci en faisant la démonstration...

Prenons donc lesdits polynômes

$$B_n(x, f) = \sum_{k=0}^n \binom{n}{k} f\left(\frac{k}{n}\right) x^k (1-x)^{n-k}$$

avec x dans $[0, 1]$ (on peut toujours se ramener à l'intervalle $[a, b]$ moyennant une transformation affine : $t \in [a, b] \rightarrow \frac{b-t}{b-a} \in [0, 1]$ et réciproquement). Il nous faut donc montrer que ces polynômes convergent simplement vers la fonction f (continue), soit que pour tout $\varepsilon > 0$, il existe N tel que pour tout n supérieur à N , $|f(x) - B_n(x, f)| < \varepsilon$.

Prenons tout d'abord $f=1$:

$$B_n(x, 1) = \sum_{k=0}^n \binom{n}{k} x^k (1-x)^{n-k} = (x+1-x)^n = 1$$

(développement du binôme) et tout va bien. Un peu plus loin avec $f(x)=x$:

$$B_n(x, x) = \sum_{k=0}^n \binom{n}{k} \frac{k}{n} x^k (1-x)^{n-k} = \sum_{k=0}^n \binom{n-1}{k-1} x^k (1-x)^{n-k} = x \sum_{k=0}^n \binom{n-1}{k-1} x^{k-1} (1-x)^{n-1-(k-1)} = x.$$

Plus complexe avec $f(x) = x^2$:

$$\left| x^2 - B_n(x, x^2) \right| = \left| x^2 - \sum_{k=0}^n \binom{n}{k} \frac{k^2}{n^2} x^k (1-x)^{n-k} \right| = x \left| x - \sum_{k=0}^n \binom{n-1}{k-1} \frac{k}{n} x^{k-1} (1-x)^{n-k} \right|, \text{ bof...}$$

Dérivons $(x+y)^n$ par rapport à x puis multiplions par x/n : on a

$$x(x+y)^{n-1} = \sum_{k=0}^n \binom{n}{k} \frac{k}{n} x^k y^{n-k},$$

redérivons cette relation et remultiplions par x/n , on a alors :

$$x^2(x+y)^{n-2} + \frac{xy}{n}(x+y)^{n-2} = \sum_{k=0}^n \binom{n}{k} \frac{k^2}{n^2} x^k y^{n-k}.$$

Remplaçons y par $1-x$: $x^2 + \frac{x(1-x)}{n} = B_n(x, x^2)$; par conséquent

$$\left| x^2 - B_n(x, x^2) \right| = \frac{x(1-x)}{n} \leq \frac{1}{4n}$$

(la fonction $x(1-x)$ est inférieure à $1/4$ sur $[0, 1]$).

Passons au cas général : f est uniformément continue, donc bornée ; il existe donc M tel que

$|f(x)| \leq M$, par conséquent $\left|f(x) - f\left(\frac{k}{n}\right)\right| \leq 2M$; par ailleurs pour tout $\varepsilon > 0$, il existe δ tel que si

$\left|x - \frac{k}{n}\right| < \delta$ alors $\left|f(x) - f\left(\frac{k}{n}\right)\right| < \varepsilon$. Prenons un $\varepsilon > 0$; pour notre polynôme B_n , il existe des points

k/n pour lesquels les valeurs de B_n sont à une distance inférieure à ε de f et d'autres pour lesquels cette distance est supérieure, nous pouvons donc écrire

$$|f(x) - B_n(x, f)| \leq \sum_{\left|x - \frac{k}{n}\right| < \delta} \binom{n}{k} \left|f(x) - f\left(\frac{k}{n}\right)\right| x^k (1-x)^{n-k} + \sum_{\left|x - \frac{k}{n}\right| \geq \delta} \binom{n}{k} \left|f(x) - f\left(\frac{k}{n}\right)\right| x^k (1-x)^{n-k};$$

La première somme se majore facilement par $\frac{\varepsilon}{2}$; pour la deuxième, nous avons

$$S_2 \leq 2M \sum_{\left|x - \frac{k}{n}\right| \geq \delta} \binom{n}{k} x^k (1-x)^{n-k} \leq \frac{2M}{\delta^2} \sum_{\left|x - \frac{k}{n}\right| \geq \delta} \binom{n}{k} \left(x - \frac{k}{n}\right)^2 x^k (1-x)^{n-k}$$

d'où en développant le terme carré :

$$S_2 \leq \frac{2M}{\delta^2} [x^2 B_n(x, 1) - 2x B_n(x, x) + B_n(x, x^2)] = \frac{2M}{\delta^2} \left[x^2 - 2x^2 + \frac{x(1-x)}{n} \right] \leq \frac{2M}{4n\delta^2}.$$

En choisissant $n > \frac{M}{\delta^2 \varepsilon}$, on obtient $|f(x) - B_n(x, f)| \leq \frac{\varepsilon}{2} + \frac{M}{2n\delta^2}$. Cette démonstration est

constructive puisqu'on sait à peu près quel polynôme de Bernstein on devra utiliser pour approcher f . Par contre, comme on le voit ci-dessous la vitesse de convergence est particulièrement lente.

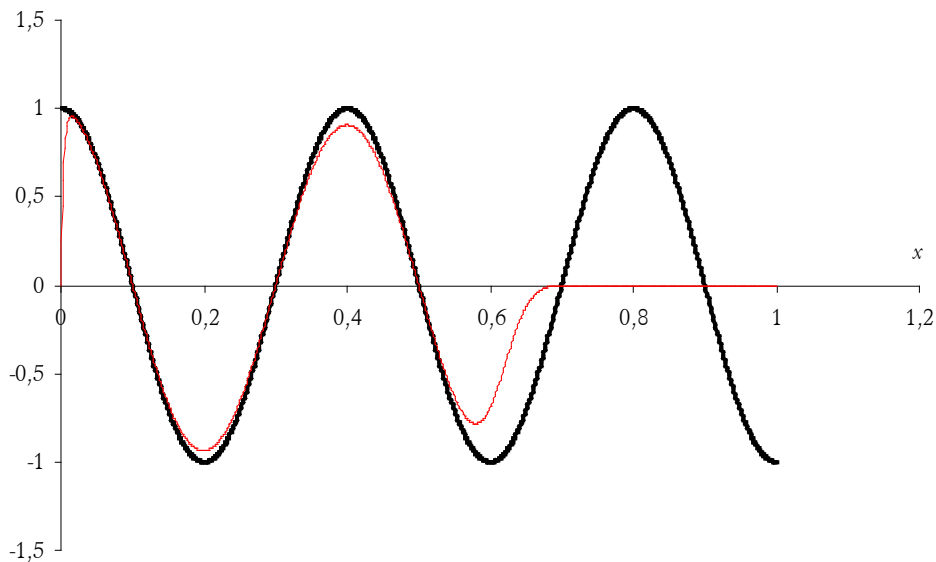


fig. 1 : Approximation de $\cos x$ par les polynômes de Bernstein, $n = 180$

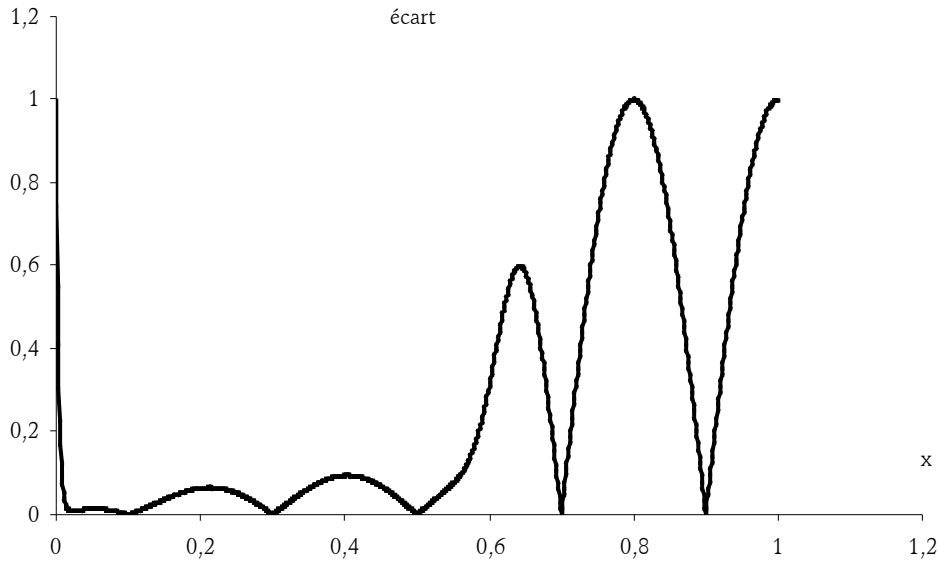


fig. 2 : écart entre $\cos x$ et B_n

5. Les zéros des polynômes orthogonaux (1)

Les polynômes $e_n(x)$ étant de degré n , ils ont n zéros complexes ou réels (aux ordres de multiplicité près), mais nous allons montrer que ces zéros sont tous réels, simples et dans $[a, b]$!

Supposons que nous ayons $k < n$ zéros réels pour $e_n(x)$, alors le polynôme

$$p_k(x) = \prod_{j=1}^k (x - x_j) = \sum_{j=0}^k a_j e_j(x)$$

sera tel que $\langle e_n, p_k \rangle = 0$.

Le produit $p_k(x)e_n(x)$ ne contiendra que des termes au carré multipliés par une expression de signe constant ; supposons que ce produit est positif, nous aurons alors

$$\langle e_n, p_k \rangle = \int_a^b e_n(x) p_k(x) w(x) dx > 0 ;$$

il y a donc une contradiction qui ne peut être levée qu'en prenant $k = n$. Il y a donc n zéros simples dans $[a, b]$.

Un autre résultat intéressant est que e_n et e_{n+1} ne s'annulent pas simultanément et que les zéros de e_{n+1} séparent ceux de e_n : reprenons la relation de récurrence du 3-a : soient x_m et x_{m+1} deux zéros successifs de e_n , on a

$$\sum_{j=0}^n e_j^2(x_m) = \frac{k_n}{k_{n+1}} [-e'_n(x_m)e_{n+1}(x_m)] \geq 0$$

donc

$$e'_n(x_m)e_{n+1}(x_m) \leq 0 ,$$

de même que

$$e'_n(x_{m+1})e_{n+1}(x_{m+1}) \leq 0 .$$

Comme on a deux zéros consécutifs, $e'_n(x_m)$ et $e'_n(x_{m+1})$ sont de signes opposés ; il en est alors de même pour $e_{n+1}(x_m)$ et $e_{n+1}(x_{m+1})$, donc e_{n+1} doit s'annuler au moins une fois entre les deux zéros.

De la même manière e_n doit s'annuler entre deux zéros de e_{n+1} .

Montrons que ces zéros ne sont pas identiques : si c'était le cas on aurait également $e_{n-1}(x_m) = 0$ et ainsi de suite en descendant les valeurs de n ; mais comme $e_0 > 0$ c'est impossible.

Nous verrons d'autres propriétés de ces zéros ultérieurement.

6. Approximation

6-a : Théorème fondamental

Le théorème fondamental de l'approximation des fonctions est le suivant :

Parmi tous les polynômes de degré n , il en existe un rendant minimum la distance

$$\|f(x) - p_n(x)\|$$

pour la norme du p.s. défini précédemment.

Le polynôme en question est donné par $p_n(x) = \sum_{j=0}^n \alpha_j e_j(x)$ avec $\alpha_j = \int_a^b f(x) e_j(x) w(x) dx$.

Cette approximation n'a de sens que pour la fonction poids w et sur l'intervalle $[a, b]$.

Prenons q_n un polynôme quelconque et p_n celui du théorème ; nous avons $q_n(x) = \sum_{j=0}^n \beta_j e_j(x)$ et

nous posons $g(x) = f(x) - p_n(x)$: pour $k \leq n$,

$$\langle g, e_k \rangle = \left\langle f - \sum_{j=0}^n \alpha_j e_j, e_k \right\rangle = \langle f, e_k \rangle - \sum_{j=0}^n \alpha_j \langle e_j, e_k \rangle = \langle f, e_k \rangle - \alpha_k \langle e_k, e_k \rangle = 0$$

Par conséquent $\langle g, q_n \rangle = 0$; de plus

$$\|g - q_n\|^2 = \langle g - q_n, g - q_n \rangle = \|g\|^2 - 2\langle g, q_n \rangle + \|q_n\|^2 = \|g\|^2 + \|q_n\|^2$$

et la plus petite valeur de cette norme est $\|g\|^2$, minimum atteint lorsque $q_n(x) = 0$ pour tout x .

Comme g est orthogonal avec tous les e_j , on peut appliquer ce qu'on a dit sur les zéros en considérant g comme e_n , donc g s'annule n fois et change de signe $n+1$ fois (ou bien g est nul).

6-b : Formule de quadrature de Gauss

Nous avons parlé brièvement dans le livre de la formule d'interpolation de Lagrange qui permet d'obtenir un polynôme passant exactement par n points donnés de f . Précisons ce point :

prenons la suite de points x_j , $j=1\dots n$, zéros du polynôme orthogonal e_n , d'images $f(x_j)$ par f

continue et considérons $F(x) = \sum_{j=1}^n f(x_j) \frac{e_n(x)}{(x-x_j)e_n'(x_j)}$: F est un polynôme de degré $n-1$ et en plus

$F(x_j) = \lim_{x \rightarrow x_j} F(x) = f(x_j)$; si f est un polynôme de degré $n-1$, alors $f = F$.

Prenons f un polynôme de degré $2n-1$, $F(x) - f(x)$ est également de degré $2n-1$ et s'annule pour les

x_j , la fonction $r(x) = \frac{F(x) - f(x)}{e_n(x)}$ est alors un polynôme de degré $n-1$ et nous pouvons écrire

$f(x) = F(x) - (F(x) - f(x))$, soit

$$(1) \quad f(x) = \sum_{j=1}^n f(x_j) \frac{e_n(x)}{(x-x_j)e_n'(x_j)} - r(x)e_n(x).$$

Intégrons cette relation, histoire de faire réapparaître nos p.s. :

$$\int_a^b f(x)w(x)dx = \sum_{j=1}^n f(x_j) \int_a^b \frac{e_n(x)}{(x-x_j)e_n'(x_j)} w(x)dx - \int_a^b r(x)e_n(x)w(x)dx,$$

posons $\int_a^b \frac{e_n(x)}{(x-x_j)e'_n(x_j)} w(x) dx = \lambda_{j,n}$ et remarquons que $\int_a^b r(x)e_n(x)w(x) dx = 0$ (la remarque sur le degré de r sert ici) ; les nombres $\lambda_{j,n}$ sont appelés les *nombre de Christoffel*, sont indépendants de f et donnent la *formule de quadrature de Gauss* :

$$\int_a^b f(x)w(x) dx = \sum_{j=1}^n f(x_j)\lambda_{j,n}.$$

On pourrait se dire que sachant à priori calculer l'intégrale d'un polynôme sans trop de mal la formule précédente ne sert pas à grand chose, mais heureusement elle reste à peu près valable même si f n'est pas un polynôme.

Par exemple avec $n = 2$ et le poids $w(x)=1$ nous avons $e_2(x) = \sqrt{\frac{5}{8}}(3x^2 - 1)$, d'où sur $[-1 ; 1]$ les nombres de Christoffel $\lambda_{1,2} = \lambda_{2,2} = 1$; calculons par exemple

$$\int_{-1}^1 \frac{1}{x+3} dx \approx \frac{1}{3-\frac{1}{\sqrt{3}}} \cdot 1 + \frac{1}{3+\frac{1}{\sqrt{3}}} \cdot 1 = \frac{9}{13} \approx 0,6923,$$

proche du $\ln 2$ obtenu par calcul direct.

6-c : **Nombres de Christoffel**

On trouve d'autres définitions des nombres de Christoffel n'utilisant pas d'intégrale : reprenons le noyau hilbertien et appliquons le à un zéro de e_n :

$$(2) \quad K_n(x, x_k) = \sum_{j=0}^{n-1} e_j(x)e_j(x_k) = \frac{k_n}{k_{n+1}} \left[\frac{-e_n(x)e_{n+1}(x_k)}{x-x_k} \right]$$

d'où

$$\frac{e_n(x)}{x-x_k} = \frac{-k_{n+1}K_n(x, x_k)}{k_n e_{n+1}(x_k)}$$

et en remplaçant dans la définition des nombres de Christoffel :

$$\lambda_{k,n} = \int_a^b \frac{e_n(x)}{(x-x_k)e'_n(x_k)} w(x) dx = \frac{-k_{n+1}}{k_n e_{n+1}(x_k) e'_n(x_k)} \int_a^b K_n(x, x_k) w(x) dx = \frac{-k_{n+1}}{k_n e_{n+1}(x_k) e'_n(x_k)}$$

(la dernière intégrale vaut 1).

De même si on fait tendre x vers x_k dans (2) on a $K_n(x_k, x_k) = \sum_{j=0}^{n-1} e_j^2(x_k) = \frac{-k_n}{k_{n+1}} e_{n+1}(x_k) e'_n(x_k)$ d'où

$$\lambda_{k,n} = \frac{1}{\sum_{j=0}^{n-1} e_j^2(x_k)}.$$

Si nous appliquons la formule de quadrature de Gauss à la fonction $\left[\frac{e_n(x)}{(x-x_k)e'_n(x_k)} \right]^2$ nous

obtenons

$$\begin{aligned} \int_a^b \left[\frac{e_n(x)}{(x-x_k)e'_n(x_k)} \right]^2 w(x) dx \\ = \sum_{j=1}^n \lambda_{j,n} \lim_{x \rightarrow x_j} \left[\frac{e_n(x)}{(x-x_k)e'_n(x_k)} \right]^2 = \sum_{j=1}^{n-1} \lambda_{j,n} \left[\frac{e_n(x_j)}{(x_j-x_k)e'_n(x_k)} \right]^2 + \lambda_{k,n} = \lambda_{k,n} \end{aligned}$$

(tous les termes dans le carré sont nuls).

6-d : Evaluation de l'erreur

Dans tout ce qui précède nous n'avons utilisé que la continuité et la dérivée première des e_j , mais ces derniers sont davantage dérivables, aussi peut-on voir si on n'obtiendrait pas une évaluation de l'erreur commise (en remplaçant f par p_n) à l'aide des dérivées d'ordre supérieur, un peu à la manière de Taylor. Reprenons (1) et considérons

$$g(x) = f(x) - \sum_{j=1}^n f(x_j) \frac{e_n(x)}{(x-x_j)e'_n(x_j)} = f(x) - \sum_{j=1}^n f(x_j) \psi_j(x) ;$$

la fonction g a au moins les mêmes zéros que e_n et pour une valeur X fixée, nous pouvons trouver une constante K telle que $\theta(x) = g(x) - Ke_n(x)$ s'annule en X : $K = \frac{g(X)}{e_n(X)}$ convient d'ailleurs très

bien. Dérivons n fois cette relation : $\theta^{(n)}(x) = g^{(n)}(x) - Kn!k_n$, mais comme θ s'annule au moins $n+1$ fois, sa dérivée n -ième doit s'annuler au moins une fois, pour une valeur u par exemple :

$g^{(n)}(u) = Kn!k_n$; on a donc $K = \frac{1}{n!k_n} g^{(n)}(u) = \frac{1}{n!k_n} f^{(n)}(u)$ (la fonction ψ_j est un polynôme de degré $n-1$ et sa dérivée n -ième est nulle).

Finalement nous avons

$$f(x) = \sum_{j=1}^n f(x_j) \psi_j(x) + \frac{f^{(n)}(u)}{n!k_n} e_n(x), \quad \psi_j(x) = \frac{e_n(x)}{(x-x_j)e'_n(x_j)},$$

soit la *formule de Lagrange avec reste* (le x de la formule correspond en fait au X précédent, ce qui donne la formule).

6-e : Formule d'interpolation d'Hermite

Dans la formule précédente on obtient un polynôme qui suit les points de f , mais on ne s'est pas occupé de la forme du polynôme : il serait (très) intéressant d'obtenir en plus des tangentes identiques pour f et pour le polynôme... grosso-modo dans la formule de Lagrange on remplace la fonction f par des segments de droite puisqu'on utilise la dérivée première des e_n ; pour obtenir des courbes nous allons utiliser des arcs de parabole et donc utiliser les dérivées secondes.

Nous cherchons donc un polynôme de degré $2n-1$ (on va réutiliser Gauss) tel que ses valeurs et celles de sa dérivée coïncident avec celles de f en x_1, x_2, \dots, x_n (le lecteur peut penser aux B-splines pour lesquelles on trace des arcs de courbe successifs avec une fonction de degré $3 = 2 \cdot 2 - 1$).

Il faut ici que nous arrivions à écrire f sous la forme

$$f(x) = \sum_{j=1}^n f(x_j) \sigma_j(x) + f'(x_j) \tau_j(x) + R_n(x) ;$$

supposons que nous ayons trouvé nos fonctions σ_j et τ_j , posons

$$F(x) = \sum_{j=1}^n f(x_j) \sigma_j(x) + f'(x_j) \tau_j(x),$$

on doit alors avoir $f(x_j) = F(x_j)$ et $R_n(x_j) = 0$ ainsi que $f'(x_j) = F'(x_j)$ et $R'_n(x_j) = 0$; comme R_n est lié aux e_n , il faut faire intervenir le carré des e_n afin d'assurer la nullité, aussi écrivons

$R_n(x) = Ke_n^2(x)$; le même raisonnement que précédemment donne la valeur de K : $K = \frac{f^{(2n)}(u)}{k_n^2(2n)!}$ et

$$R_n(x) = \frac{f^{(2n)}(u)}{k_n^2(2n)!} e_n^2(x) .$$

Pour les fonctions σ_j et τ_j nous devons avoir

$$\sigma_j(x_j) = 1, \tau_j(x_j) = 0, \sigma'_j(x_j) = 0, \tau'_j(x_j) = 1$$

et qu'ils s'annulent pour tous les autres zéros ; il faut également que ce soient des polynômes

donc nous devons retrouver dans chacun le terme $\left[\frac{e_n(x)}{(x-x_j)e'_n(x_j)} \right]^2 = \psi_j^2(x)$.

Posons $\sigma_j(x) = a(x)\psi_j^2(x)$, comme $\psi_j(x_j) = 1$ nous devons avoir $a(x_j) = 1$; dérivons :

$$\sigma'_j(x) = a'(x)\psi_j^2(x) + 2a(x)\psi'_j(x)\psi_j(x)$$

d'où

$$0 = a'(x_j) + 2\psi'_j(x_j) \Leftrightarrow a'(x_j) = -2\psi'_j(x_j) ;$$

calculons la dérivée de ψ :

$$\psi_j(x) = \frac{1}{e'_n(x_j)} \frac{e'_n(x)(x-x_j) - e_n(x_j)}{(x-x_j)^2} \xrightarrow{x \rightarrow x_j} \frac{e''_n(x_j)}{e'_n(x_j)}$$

qui ne s'annule évidemment pas en x_j , donc en multipliant par $\frac{(x-x_j)}{2}$ nous obtenons

$$a(x) = 1 - (x-x_j) \frac{e''_n(x)}{e'_n(x)}.$$

Le même type de raisonnement permet d'obtenir $\tau_j(x) = (x-x_j)\psi_j^2(x)$. Nous avons finalement obtenu la forme cherchée, nous calculons alors :

$$\int_a^b \tau_j(x)w(x)dx = \int_a^b \frac{1}{[e'_n(x_j)]^2} \frac{e_n(x)}{(x-x_k)} e_n(x)w(x)dx = A \left\langle \frac{e_n}{(x-x_k)}, e_n \right\rangle = 0$$

puisque le terme de gauche du p.s. est de degré $n-1$. De même nous obtenons

$$\begin{aligned} \int_a^b \sigma_j(x)w(x)dx &= \int_a^b \left[1 - (x-x_j) \frac{e''_n(x_j)}{e'_n(x_j)} \right] \left[\frac{e_n(x)}{(x-x_j)e'_n(x_j)} \right]^2 w(x)dx \\ &= \int_a^b \left[\frac{e_n(x_j)}{(x-x_j)e'_n(x_j)} \right]^2 - \frac{e''_n(x_j)e_n(x)}{(x-x_j)[e'_n(x_j)]^2} e_n(x)w(x)dx = \int_a^b \left[\frac{e_n(x_j)}{(x-x_j)e'_n(x_j)} \right]^2 w(x)dx = \lambda_{j,n}. \end{aligned}$$

Finalement nous récupérons la formule de Gauss avec un reste d'évaluation de l'erreur commise :

$$\int_a^b f(x)w(x)dx = \sum_{j=1}^n f(x_j)\lambda_{j,n} + \int_a^b \frac{f^{(2n)}(u)e_n^2(x)}{b_n^2(2n)!} w(x)dx.$$

7. Les zéros des polynômes orthogonaux (2)

En reprenant la démonstration du § 5 avec $n+m$ au lieu de $n+1$, on montre de la même manière que e_{n+m} s'annule au moins une fois entre deux zéros de e_n . Par contre on dispose de quelques résultats sur la distribution des zéros : considérons la fonction f définie par

$$\begin{cases} f(x) = (x-\alpha)(\beta-x) - \varepsilon \text{ si } x \in [\alpha, \beta] \subset [a, b], \\ f(x) = -\varepsilon \text{ ailleurs.} \end{cases}$$

Grâce au théorème de Weierstrass nous savons qu'il existe un polynôme p tel que

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

pour x dans $]a, b[$; pour ε suffisamment petit f est positive et $\int_a^b f(x)w(x)dx > 0$ ainsi que $\int_a^b p(x)w(x)dx > 0$ (alors que p est négatif par construction en dehors de $[\alpha, \beta]$); prenons maintenant e_n avec n assez grand, tel que $\text{degré}(p) \leq 2n - 1$ et que e_n ne s'annule pas dans $[\alpha, \beta]$. La formule de quadrature de Gauss donne alors

$$\int_a^b p(x)w(x)dx = \sum_{j=1}^n p(x_j)\lambda_{j,n}$$

qui est strictement positif, mais si tous les x_j sont en dehors de $[\alpha, \beta]$, les $p(x_j)$ sont négatifs et comme les nombres de Christoffel sont positifs, la somme précédente doit être négative. On a donc une contradiction et e_n s'annule au moins une fois dans $[\alpha, \beta]$, ce que nous traduisons en disant que e_n s'annule au moins une fois dans tout intervalle inclus dans $[a, b]$ pour n suffisamment grand.

Théorème de séparation : prenons maintenant la suite strictement croissante u_n avec $u_0 = a$ et $u_n = b$ telle que u_k soit l'unique valeur pour laquelle $I(u_k) = \int_a^{u_k} w(x)dx = \sum_{j=1}^k \lambda_{j,n}$; alors les zéros de e_n s'insèrent exactement entre les termes de la suite :

$$a = u_0 < x_1 < u_1 < x_2 < \dots < u_{n-1} < x_n < u_n = b.$$

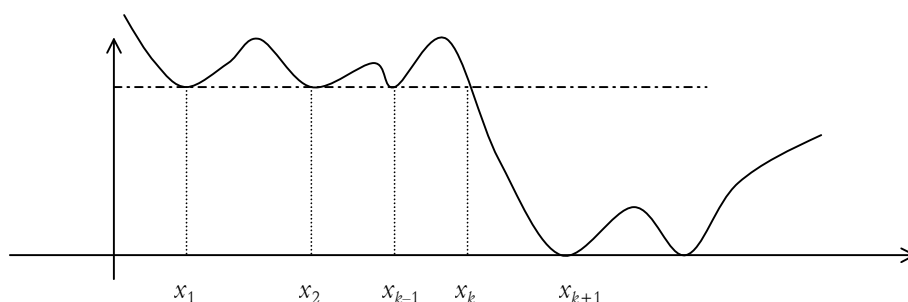
Prenons un polynôme $q(x)$ de degré $2n - 2$ défini par les $2n - 1$ conditions suivantes :

$$\begin{cases} q(x_j) = 1 \text{ si } 1 \leq j \leq k, \\ q(x_j) = 0 \text{ si } k+1 \leq j \leq n, \\ q'(x_j) = 0 \text{ si } j \neq k. \end{cases}$$

Comme il y a $2n-1$ conditions, notre polynôme q est défini de manière unique et est de degré $2n-2$, sa dérivée est de degré $2n-3$ et s'annule pour tous les x_j sauf un donc $n-1$ fois ; de plus entre deux x_j avec j inférieur à k elle doit s'annuler au moins une fois (q passe de 1 à 1, q' s'annule donc au moins $k-1$ fois), enfin quand q est nul sa dérivée est nulle (elle s'annule au moins $n-k-1$ fois) ; conclusion q' s'annule au moins $n-1+n-2$ fois, soit $2n-3$ fois.

Mais comme q' est de degré $2n-3$, elle s'annule exactement $2n-3$ fois.

Supposons maintenant que $q(x) \leq 1$ pour $x \leq x_k$, comme $q(x_k) = 1$, q est croissante donc $q'(x_k) > 0$.



Comme $q(x_{k+1}) = 0$, pour qu'il n'y ait pas d'incohérence dans le comportement de q , q' devrait s'annuler entre x_k et x_{k+1} mais comme on a compté tous les zéros de q' c'est impossible.

Conclusion, $q(x) > 1$ pour $x \leq x_k$.

On obtient donc que $\int_a^b q(x)w(x)dx > \int_a^{x_k} q(x)w(x)dx > \int_a^{x_k} w(x)dx$ et que

$$\int_a^b q(x)w(x)dx = \sum_{j=1}^k \lambda_{j,n} q(x_j) = \sum_{j=1}^k \lambda_{j,n} = \int_a^{y_k} w(x)dx ;$$

finalement

$$\int_a^{y_k} w(x)dx > \int_a^{x_k} w(x)dx \Rightarrow y_k > x_k .$$

De la même manière en utilisant le polynôme \tilde{q} défini par

$$\begin{cases} \tilde{q}(x_j) = 0 \text{ si } 1 \leq j \leq k, \\ \tilde{q}(x_j) = 1 \text{ si } k+1 \leq j \leq n, \\ \tilde{q}'(x_j) = 0 \text{ si } j \neq k. \end{cases}$$

nous obtiendrons que $x_{k+1} > y_k$.

8. Espace de Hilbert

Si on considère l'ensemble des fonctions de carré sommable pour la densité w , noté $L_2(w)$, et constitué des fonctions f telles que $\int_a^b f^2(x)w(x)dx < +\infty$ nous obtenons un espace vectoriel pour les opérations habituelles ; en introduisant un p.s. comme $\langle f, g \rangle = \int_a^b f(x)g(x)w(x)dx$, on définit la norme de f par

$$\|f\| = \left[\int_a^b f^2(x)w(x)dx \right]^{1/2} = \sqrt{\langle f, f \rangle} .$$

Si en plus cet espace est *complet* pour la norme considérée, on aura obtenu un *espace de Hilbert*.

Rappelons qu'un ensemble E muni d'une norme $\|\cdot\|$ est *complet* si toute suite de Cauchy converge dans E au sens de la norme $\|\cdot\|$. Rappelons également qu'une suite (u_n) est de Cauchy lorsque pour tout $\varepsilon > 0$ il existe N tel que pour tout $n, m > N$ on ait $\|u_n - u_m\| < \varepsilon$. Si f_n est une suite convergente de fonctions de E , alors sa limite f est dans E .

L'intérêt ici est que $L_2(w)$ est complet et est donc un espace de Hilbert (voir ci-dessous).

Reprenons nos polynômes orthogonaux, nous avons vu que le polynôme p pour lequel $\|f - p\|$ est minimale est donné par $p(x) = \sum_{j=0}^n \alpha_j e_j(x)$ où $\alpha_j = \langle f, e_j \rangle$; comme les e_j sont orthonormés nous avons

$$0 \leq \|f - p\|^2 = \|f\|^2 - 2 \left\langle f, \sum_{j=0}^n \alpha_j e_j \right\rangle + \left\| \sum_{j=0}^n \alpha_j e_j \right\|^2 = \|f\|^2 - 2 \sum_{j=0}^n \langle f, e_j \rangle \alpha_j + \sum_{j=0}^n \alpha_j^2 = \|f\|^2 - \sum_{j=0}^n \alpha_j^2 .$$

Par conséquent $\|f\|^2 \geq \sum_{j=0}^n \alpha_j^2$ pour tout n , c'est l'*inégalité de Bessel*. En faisant tendre n vers l'infini,

la série de droite est bornée et converge ; si par bonheur on pouvait avoir l'égalité, la distance entre f et p pourrait être rendue aussi petite qu'on le souhaite (avec un choix judicieux de p). Le fait est que c'est effectivement possible et ceci grâce à la complétude de $L_2(w)$.

Démonstration de la complétude de $L_2(w)$

Quelques définitions pour commencer.

1. Une suite orthonormée sera dite *complète* dans $L_2(w)$ si pour tout f de $L_2(w)$ on a

$$\|f\|^2 = \sum_{j=0}^{\infty} \alpha_j^2 \quad (\text{égalité de Parseval});$$

2. Une famille de fonctions $\{\varphi_n\}$ dans un espace de Hilbert sera fermée si lorsque pour toute fonction f telle que $\langle f, \varphi_n \rangle = 0$ alors $f=0$ (ceci se comprend intuitivement : il n'y a pas d'autre fonction que la fonction nulle qui soit orthogonale aux φ_n).

Premier théorème

Un ensemble orthonormé de fonctions $\{\varphi\}$ dans $L_2(w)$ est dénombrable.

Plaçons nous sur $[0, 1]$ pour simplifier ; l'ensemble $\{r_n\}$ des nombres rationnels de $[0, 1]$ est dénombrable (voir en Annexe), par conséquent nous pouvons définir une famille dénombrable de

fonctions S_r telles que $\begin{cases} S_{r_i}(x) = 1 \text{ si } 0 \leq x \leq r_i \\ S_{r_i}(x) = 0 \text{ si } r_i < x \leq 1 \end{cases}$; soit g une fonction telle que

$$\int_0^1 g(x)S_{r_i}(x)w(x)dx = \int_0^{r_i} g(x)w(x)dx = 0$$

pour tout entier i , alors g est nulle presque partout : en effet la fonction $G(x) = \int_0^x g(t)w(t)dt$ est continue et s'annule pour les valeurs rationnelles de x , donc G est nulle et $G'(x) = g(x) = 0$ presque partout. Prenons un sous-ensemble dénombrable $\{\varphi_n\}$ de $\{\varphi\}$ de fonctions pour

lesquelles $\int_0^1 \varphi_n(x)S_{r_i}(x)w(x)dx \neq 0$; si $\{\varphi\}$ était constitué uniquement de tels sous-ensembles il

serait dénombrable ; choisissons alors un sous-ensemble $\{\varphi\}_0$ de $\{\varphi\}$ non-dénombrable de fonctions possédant la propriété précédente et partitionnons l'ensemble des réels positifs par les

intervalles semi-ouverts $A_k = \left[\frac{1}{k+1}, \frac{1}{k} \right]$, k entier. Nous prenons maintenant dans $\{\varphi\}_0$ les

fonctions φ telles que le nombre $C^2 = \left[\int_0^1 \varphi(x)S_{r_i}(x)w(x)dx \right]^2$ soit dans A_k . Nous pouvons ainsi

diviser $\{\varphi\}_0$ en sous-ensembles dont au moins un n'est pas dénombrable, appelons le $\{\varphi\}_k$.

Ecrivons maintenant l'inégalité de Bessel pour S_{r_i} : on a $\alpha_j = \langle S_{r_i}, \varphi_j \rangle = \int_0^1 S_{r_i}(x)\varphi_j(x)w(x)dx$ avec

$\varphi_j \in \{\varphi\}_k$ d'où $\alpha_j^2 = C^2 \geq \frac{1}{k+1}$; la somme des α_j^2 ne peut donc converger, contredisant l'inégalité de Bessel.

Conclusion : chaque $\{\varphi\}_k$ est fini, et $\{\varphi\}_0$ est au plus dénombrable (il pourrait même être fini !).

Deuxième théorème

Dans un espace hilbertien une suite orthonormé $\{\varphi_n\}$ est complète si, et seulement si, elle est fermée.

Supposons $\{\varphi_n\}$ fermée et considérons une fonction f de L_2 ainsi que $g_n(x) = f(x) - \sum_{j=0}^n \alpha_j \varphi_j(x)$;

la suite g_n est de Cauchy puisque

$$\|g_n - g_m\| = \left\| \sum_{j=m+1}^n \alpha_j \varphi_j \right\| = \sqrt{\sum_{j=m+1}^n \alpha_j^2}$$

et donc converge vers une fonction g .

On a alors pour $k < n$, $|\langle g, \varphi_k \rangle| = |\langle g_n - g, \varphi_k \rangle| \leq \|g_n - g\|$ grâce à l'inégalité de Cauchy-Schwarz. Mais la norme de droite tend vers 0 à l'infini et $\langle g, \varphi_k \rangle = 0$; comme $\{\varphi_n\}$ est fermée, g est nulle

et $\lim_{n \rightarrow \infty} \left(\|f\|^2 - \sum_{j=1}^n \alpha_j^2 \right) = \lim_{n \rightarrow \infty} (\|g_n\|^2) = 0$ donc $\{\varphi_n\}$ est complète.

Dans l'autre sens maintenant... $\{\varphi_n\}$ est complète et prenons f telle que $\alpha_j = \langle f, \varphi_j \rangle = 0$ alors

$\lim_{n \rightarrow \infty} \left(\|f\|^2 - \sum_{j=1}^n \alpha_j^2 \right) = \lim_{n \rightarrow \infty} (\|f\|^2) = \|f\|^2 = 0$ donc f est nulle et $\{\varphi_n\}$ est fermée.

Troisième (et dernier) théorème

La suite des polynômes orthonormés $\{\varphi_n\}$ est complète dans L_2 .

Premier point : soit f continue et telle que $\int_0^1 f(x)x^n w(x) dx = 0$ pour n entier positif ou nul ; nous écrivons f sous la forme $f = p + r$ où p est un polynôme, alors

$$0 = \int_0^1 f(x)p(x)w(x) dx = \int_0^1 f^2(x)w(x) dx - \int_0^1 f(x)r(x)w(x) dx ;$$

on a alors pour tout $\varepsilon > 0$,

$$\int_0^1 f^2(x)w(x) dx = \left| \int_0^1 f(x)r(x)w(x) dx \right| \leq \varepsilon \int_0^1 |f(x)|w(x) dx ,$$

cette dernière intégrale peut être rendue aussi petite que l'on veut et on conclut que $\int_0^1 f^2(x)w(x) dx = 0$; comme f est continue, f est nulle.

Posons maintenant $F(x) = \int_0^x f(t)w(t) dt$, on a $F(0) = F(1) = 0$ (remplacer n par 0...) et $F'(x) = f(x)$

d'où $\int_0^1 x^n F'(x) dx = 0$; intégrons par parties :

$$\left[x^n F(x) \right]_0^1 - n \int_0^1 x^{n-1} F(x) dx = -n \int_0^1 x^{n-1} F(x) dx = 0$$

et F est nulle d'où $\int_0^1 f(x)w(x) dx = 0$; comme $w > 0$ presque partout, f est nulle presque partout ce qui démontre le théorème.

Nous nous sommes placés sur $[0, 1]$ et rien ne nous empêche de passer sur $[a, b]$ par une simple transformation, par contre le passage à des intervalles non bornés ne se fait pas aussi simplement. Pour certains polynômes (Laguerre, Hermite) correspondant à des densités différentes, il faudra redémontrer le caractère complet de L_2 .

Passons dans le domaine complexe...

9. Représentations conformes

Nous prenons au lieu d'un segment une courbe C rectifiable sans points multiple du plan complexe et nous appelons s la longueur d'arc prise pour paramètre ; la fonction densité w à valeurs réelles et strictement positive sera une fonction de s .

Nous pouvons alors reprendre les notions précédentes et les appliquer à des polynômes orthogonaux dans \mathbb{C} , aussi nous reprenons la définition du p.s. hermitien donnée au début du chapitre.

Intéressons nous au cas où C est fermée et où $w(s)=1$ et reprenons le noyau hilbertien

$$K_n(a, z) = \sum_{j=0}^n \overline{e_j(a)} e_j(z).$$

Comme en 3-c : nous avons $\langle p(a), K_n(a, z) \rangle = \int_C p(z) \overline{K_n(a, z)} ds = p(z)$ et le polynôme

$$G_n(z) = \frac{\pm K_n(a, z)}{\sqrt{K_n(a, a)}} \text{ est tel que } \|G_n\|^2 = \langle G_n, G_n \rangle = \int_C |G_n(z)|^2 dz = 1 \text{ et } |G_n(a)| \geq |p(a)| \text{ où } p(z) \text{ est}$$

n'importe quel polynôme de norme unité et de degré n .

Rappelons qu'il existe toujours au moins une transformation conforme $\gamma(z)$ (ou une fonction analytique) de \mathbb{C} qui transforme un domaine D inclus dans C (D simplement connexe ayant au moins deux points sur C) en le cercle unité, et ceci de manière univoque ; si en plus on impose $\gamma(a) = 0$ et $\gamma'(a) > 0$ alors γ est unique.

La démonstration de ce théorème de Riemann n'est pas constructive, mais l'utilisation de G_n va nous permettre d'approcher γ de manière efficace.

Passons tout d'abord à la limite sur $K_n(a, z)$ quand n tend vers l'infini : dans tout domaine intérieur à C , K_n est uniformément bornée ; montrons le en appliquant la formule intégrale de Cauchy :

$$\begin{aligned} K_n(a, a) &= \frac{1}{2\pi i} \int_C \frac{K_n(a, a)}{z-a} dz = \frac{1}{2\pi i} \int_C \frac{K_n^2(a, a)}{(z-a)K_n(a, a)} dz \\ &\leq \frac{1}{2\pi} \int_C \frac{|G_n(z)|^2}{|z-a|} dz \leq \frac{1}{2\pi\delta} \int_C |G_n(z)|^2 dz = \frac{1}{2\pi\delta}. \end{aligned}$$

où δ est la distance de a à C . Si on appelle d la plus petite des distances de a à C et de z à C , nous avons grâce à Cauchy-Schwarz :

$$|K_n(a, z)|^2 \leq K_n(a, a)K_n(z, z) = \frac{1}{4\pi d^2}.$$

La limite K de K_n existe donc et par là même la limite G de G_n : $G(z) = \lim_{n \rightarrow \infty} \frac{K_n(a, z)}{\sqrt{K_n(a, a)}}$ et sa norme est également 1.

G est également solution du problème du maximum : soit γ transformant l'intérieur de C en le cercle unité et g sa réciproque alors $ds = \frac{d\theta}{|\gamma'(z)|} = |g'(\gamma(\theta))| d\theta$ où $d\theta$ est la longueur d'arc du cercle unité. En intégrant G on a alors avec $\gamma = e^{i\theta}$:

$$\int_C |G(z)|^2 dz = \int_0^{2\pi} |G(g(\gamma))|^2 |g'(\gamma)| d\theta = \int_0^{2\pi} |G(g(\gamma))\sqrt{g'(\gamma)}|^2 d\theta = 1.$$

Posons $F(\gamma) = G(g(\gamma))\sqrt{g'(\gamma)} = \sum_{n=0}^{\infty} a_n \gamma^n$, nous avons alors $2\pi \sum_{n=0}^{\infty} |a_n|^2 = 1$ grâce à l'intégrale précédente.

Pour rendre $|G(a)|^2$ maximum, il faut rendre $\left| \frac{F(\gamma)}{\sqrt{g'(\gamma)}} \right|^2$ maximum au point $g(\gamma) = a$, soit pour $\gamma(a) = 0$. Il nous faut donc $F(\gamma(a)) = F(0) = a_0$ avec $2\pi |a_0|^2 = 1$, soit

$$\max |G(a)|^2 = \frac{1}{2\pi |g'(0)|^2} = \frac{\gamma'(a)}{2\pi}.$$

On cherche maintenant à voir le lien entre K_n et γ , aussi considérons la norme de leur différence, soit

$$J_n = \int_C \left| K_n(a, z) - \frac{1}{2\pi} \sqrt{\gamma'(a)\gamma'(z)} \right|^2 ds.$$

Nous avons

$$J_n = \int_C |K_n(a, z)|^2 ds + \frac{\gamma'(a)}{(2\pi)^2} \int_C |\gamma'(z)|^2 ds - \frac{1}{\pi} \operatorname{Re} \int_C K_n(a, z) \sqrt{\gamma'(a)\overline{\gamma'(z)}} ds;$$

la première intégrale vaut $K_n(a, a) \int_C |G_n(z)|^2 ds = K_n(a, a)$; pour la seconde la courbe C s'envoie sur le cercle unité et $ds = |g'(\gamma(\theta))| d\theta$ d'où $\int_C |\gamma'(z)| ds = \int_0^{2\pi} d\theta = 2\pi$ et notre intégrale vaut $\frac{\gamma'(a)}{2\pi}$.

Pour la dernière :

$$\begin{aligned} \int_C K_n(a, z) \sqrt{\gamma'(a)\overline{\gamma'(z)}} ds &= \int_0^{2\pi} K_n(a, g(\gamma)) \sqrt{\frac{\gamma'(a)}{g'(\gamma)}} |g'(\gamma)| d\theta \\ &= \sqrt{\gamma'(a)} \int_0^{2\pi} K_n(a, g(\gamma)) \sqrt{g'(\gamma)} d\theta = \sqrt{\gamma'(a)} 2\pi K_n(a, g(0)) \sqrt{g'(0)} = 2\pi K_n(a, a). \end{aligned}$$

en utilisant la formule de la moyenne sur les fonctions harmoniques.

Finalement nous obtenons :

$$J_n = K_n(a, a) + \frac{\gamma'(a)}{2\pi} - 2K_n(a, a) = \frac{\gamma'(a)}{2\pi} - K_n(a, a).$$

Comme $\frac{\gamma'(a)}{2\pi} = \lim_{n \rightarrow \infty} K_n(a, a)$, la limite de J_n est 0. On a donc $K_n(a, z) = \frac{1}{2\pi} \sqrt{\gamma'(a)\overline{\gamma'(z)}}$ presque partout d'où

$$\gamma'(z) = \frac{(2\pi)^2}{\gamma'(a)} K_n^2(a, z) = \lim_{n \rightarrow \infty} 2\pi \frac{K_n^2(a, z)}{K_n(a, a)},$$

et comme $\gamma(a) = 0$ on a finalement :

$$\gamma(z) = \lim_{n \rightarrow \infty} \frac{2\pi}{K_n(a, a)} \int_a^z K_n^2(a, u) du.$$

Prenons par exemple la représentation conforme du cercle unité sur lui-même dont on sait qu'elle est donnée par

$$\gamma(z) = \frac{z-a}{1-\bar{a}z}$$

pour tout point a à l'intérieur du cercle unité ; on a bien $\gamma(a) = 0$ et

$$\gamma'(z) = \frac{1-\bar{a}z + \bar{a}(z-a)}{(1-\bar{a}z)^2} \Rightarrow \gamma'(a) = \frac{1}{1-\bar{a}a} = \frac{1}{1-|a|^2} > 0.$$

Nous allons retrouver ce résultat avec ce qui vient d'être fait.

Nous choisissons donc C le cercle unité et a un point à l'intérieur ; les polynômes orthogonaux e_n sont alors tels que

$$\langle e_j, e_k \rangle = \int_C e_j(z) \overline{e_k(z)} dz = \int_0^{2\pi} e^{i(j-k)\theta} d\theta = 2\pi \delta_{j,k}$$

(puisque les polynômes e_n sont sur le cercle, ils s'écrivent sous la forme $e_n(z) = e^{if_n(\theta)}$), comme on doit avoir $a_n = \langle e_n, ze_{n-1} \rangle = 1$ il nous faut

$$\int_C \overline{z} e^{i[f_n(\theta) - f_{n-1}(\theta)]} dS = \int_C e^{i[f_n(\theta) - f_{n-1}(\theta) - \theta]} dS$$

et par conséquent $f_n(\theta) = f_{n-1}(\theta) + \theta$, soit $f_n(\theta) = n\theta$; on normalise en divisant par $\sqrt{2\pi}$ d'où la famille orthonormée

$$e_n(z) = \frac{1}{\sqrt{2\pi}} z^n = \frac{1}{\sqrt{2\pi}} e^{in\theta}.$$

Cherchons le noyau hilbertien :

$$K(a, z) = \sum_{j=0}^{\infty} \overline{e_j(a)} e_j(z) = \frac{1}{2\pi} \sum_{j=0}^{\infty} \overline{a}^j z^j = \frac{1}{2\pi} \sum_{j=0}^{\infty} (\overline{a}z)^j = \frac{1}{2\pi(1-\overline{a}z)}$$

(somme d'une suite géométrique) et

$$\gamma(z) = 2\pi(2\pi(1-\overline{a}a)) \int_a^z \frac{du}{4\pi^2(1-\overline{a}u)^2} = \frac{1-\overline{a}a}{-\overline{a}} \int_a^z \frac{-\overline{a}du}{(1-\overline{a}u)^2} = \frac{1-\overline{a}a}{-\overline{a}} \left[\frac{1}{1-\overline{a}z} - \frac{1}{1-\overline{a}a} \right] = \frac{z-a}{1-\overline{a}z}.$$

On reconnaît bien évidemment dans ce cas les polynômes complexes servant à décomposer une fonction en série de Fourier, ceux-ci n'étant finalement qu'un cas particulier d'une situation beaucoup plus générale.

10. Un exemple : le produit hermitien et les inégalités de Heisenberg

10-a : Produit scalaire hermitien

Prenons l'exemple de la base orthonormée sur $[0 ; 1]$ des fonctions $(e^{2i\pi nx})_{n \in \mathbb{Z}}$:

$$\langle e^{2i\pi nx}, e^{2i\pi mx} \rangle = \int_0^1 e^{2i\pi nx} e^{-2i\pi mx} dx = \int_0^1 e^{2i\pi(n-m)x} dx = \delta_{nm}$$

δ_{nm} vaut 1 si $n = m$ et 0 sinon (symbole de Kronecker).

Posons $k = n - m$:

$$\int_0^1 e^{2i\pi kx} dx = \frac{1}{2i\pi k} \left[e^{2i\pi kx} \right]_0^1 = \frac{e^{2i\pi k} - 1}{2i\pi k} = \frac{e^{i\pi k} - e^{-i\pi k}}{2i} = \frac{e^{i\pi k}}{\pi k} \sin(\pi k) = 0.$$

Ces fonctions sont linéairement indépendantes, chose que nous avons obtenue en étudiant les séries de Fourier. Faisons le produit scalaire d'une fonction f et d'un élément de $(e^{2i\pi nx})_{n \in \mathbb{Z}}$:

$$\langle e^{2i\pi n}, f \rangle = \int_0^1 e^{-2i\pi nx} f(x) dx$$

soit le coefficient c_n de la série de Fourier de f qui correspond en fait à la projection orthogonale de f sur le vecteur de base $e^{2i\pi n}$.

Lemme de Lebesgue : un résultat important est le suivant :

$$\lim_{t \rightarrow +\infty} \int_a^b f(x) e^{itx} dx = 0 ;$$

intuitivement le fait de multiplier par e^{itx} revient à envoyer f sur le cercle trigonométrique, aussi lorsque t tend vers l'infini, on fait une infinité de fois le tour du cercle et la somme de toutes les valeurs s'annule en moyenne.

10-b : Inégalités de Heisenberg

La démonstration de ces inégalités que nous avons simplement approchées dans le chapitre 13 utilise la TF ainsi que quelques notions de probabilités. Considérons une fonction f telle que

$$\int_{-\infty}^{+\infty} |f(x)|^2 dx = 1,$$

donc telle que $|f|^2$ soit une densité de probabilité ; la TF de f , \hat{f} est également une distribution de probabilité : considérons deux variables aléatoires x et y de distributions respectives f et \hat{f} , de moyennes x_m et y_m , alors les variances respectives de x et y sont

$$\text{var } x = \int_{-\infty}^{+\infty} (x - x_m)^2 |f(x)|^2 dx \quad \text{et} \quad \text{var } y = \int_{-\infty}^{+\infty} (y - y_m)^2 |\hat{f}(y)|^2 dy ;$$

posons $X = x - x_m$ et $Y = y - y_m$, on a alors

$$\text{var } x = \text{var } X = \int_{-\infty}^{+\infty} X^2 |f(x)|^2 dX = \int_{-\infty}^{+\infty} X^2 |f(X + x_m)|^2 dX ;$$

posons $g(X) = e^{2i\pi X y_m} f(X + x_m)$ qui est également normalisée :

$$\text{var } X = \int_{-\infty}^{+\infty} X^2 |g(X)|^2 dX \quad \text{et} \quad \hat{g}(Y) = e^{-2i\pi x_m(Y + y_m)} \hat{f}(X + x_m),$$

soit $\text{var } Y = \text{var } y = \int_{-\infty}^{+\infty} Y^2 |\hat{g}(Y)|^2 dY$. Calculons la TF de f en faisant une intégration par parties :

$$\hat{f}'(k) = \int_{-\infty}^{+\infty} f'(x) e^{-2i\pi kx} dx = \left[f(x) e^{-2i\pi kx} \right]_{-\infty}^{+\infty} + 2i\pi k \int_{-\infty}^{+\infty} f(x) e^{-2i\pi kx} dx = 2i\pi k \hat{f}(k) ;$$

le crochet est nul si f est bornée (ce qui est le cas ici) : la somme $M e^{-2i\pi kx}$ est nulle car on fait une infinité de fois le tour du cercle de centre O et de rayon M.

Conséquence de ceci, $k\hat{f}(k) = \frac{1}{2\pi i} \hat{f}'(k)$ d'où :

$$\int_{-\infty}^{+\infty} Y^2 |\hat{g}(Y)|^2 dY = \int_{-\infty}^{+\infty} |Y \hat{g}(Y)|^2 dY = \int_{-\infty}^{+\infty} \left| \frac{1}{2\pi i} \hat{g}'(Y) \right|^2 dY = \frac{1}{4\pi^2} \int_{-\infty}^{+\infty} |\hat{g}'(Y)|^2 dY .$$

La TF est une *isométrie*, soit $\int_{-\infty}^{+\infty} |\hat{g}'(Y)|^2 dY = \int_{-\infty}^{+\infty} |g'(X)|^2 dX$. Grâce à l'inégalité de Schwarz nous pouvons écrire

$$(1) \quad \left(\int_{-\infty}^{+\infty} X^2 |g(X)|^2 dX \right) \left(\int_{-\infty}^{+\infty} |g'(X)|^2 dX \right) \geq \left(\int_{-\infty}^{+\infty} X |g(X)| |g'(X)| dX \right)^2 \\ = \left(\int_{-\infty}^{+\infty} |Xg(X)g'(X)| dX \right)^2$$

en utilisant les «vecteurs» $Xg(X)$ et $g'(X)$.

La fonction f initiale n'est pas forcément réelle, et même en Mécanique Quantique elle est souvent complexe, de même pour g ; or pour deux complexes a et b on a l'inégalité $|ab| \geq \frac{1}{2} (a\bar{b} + \bar{a}b)$ qui donne appliquée à nos deux vecteurs :

$$|Xg(X)g'(X)| \geq \frac{1}{2} \left(Xg(X)\overline{g'(X)} + \overline{Xg(X)}g'(X) \right).$$

Par ailleurs si on a la curiosité de dériver $|g(X)|^2$:

$$\frac{d}{dX}(|g(X)|^2) = g(X)\overline{g'(X)} + \overline{g(X)}g'(X)$$

d'où enfin :

$$|Xg(X)g'(X)| \geq \frac{1}{2} X \frac{d}{dX}(|g(X)|^2).$$

On peut donc encore minorer dans (1) par

$$\frac{1}{4} \left(\int_{-\infty}^{+\infty} X \frac{d}{dX}(|g(X)|^2) dX \right)^2 = \frac{1}{4} \left(\left[X|g(X)|^2 \right]_{-\infty}^{+\infty} - \int_{-\infty}^{+\infty} |g(X)|^2 dX \right)^2 = \frac{1}{4} (-1)^2 = \frac{1}{4} ;$$

on a encore intégré par parties, le crochet est nul (g est bornée et sur \mathbb{R} X vaut 0 en moyenne), quand à l'intégrale restante elle vaut 1.

Récrivons donc (1) avec les variances :

$$\text{var } x \text{ var } y \geq \frac{1}{4\pi^2} \cdot \frac{1}{4} = \frac{1}{16\pi^2} \text{ (inégalité de Heisenberg).}$$

Il y aura égalité lorsque la TF de f est égale à f , c'est à dire lorsque la distribution de probabilité est une loi Normale !

11. Equations intégrales

11-a : Aperçu de la question

L'introduction des équations intégrales s'est faite en Analyse avec Laplace (1782) qui considère les équations d'inconnue y : $f(x) = \int e^{xu} y(u) du$ et $f(x) = \int u^{x-1} y(u) du$ en relation avec la résolution d'équations différentielles (méthode de la résolvante de Laplace). La première équation de ce type dont on a trouvé la solution est celle de Fourier : $f(x) = \int_{-\infty}^{+\infty} \cos(xu) y(u) du$ pour laquelle une

solution est sous certaines conditions $y(u) = \frac{2}{\pi} \int_0^{+\infty} \cos(ux) f(x) dx$.

Abel s'attaquera à la question (Œuvres, I, pp 11, 97) et obtiendra quelques solutions mais le bond en avant sera dû à Liouville qui trouvera une méthode importante pour la résolution de ces équations. Par la suite de nombreux développements seront apportés... Voyons ce qu'en dit l'Encyclopédie Universalis :

La forme usuelle d'une équation intégrale (inhomogène) est :

$$y(x) - \lambda \int_A K(x, u) y(u) du = f(x)$$

où A est une partie de \mathbb{R}^m décrite par chacune des variables x et u , K une fonction donnée sur A^2 appelée *noyau de l'équation*, f une fonction donnée sur A , qui est la constante 0 dans l'équation *homogène* :

$$y(x) = \lambda \int_A K(x, u) y(u) du,$$

enfin la fonction y est l'inconnue de l'équation et λ un paramètre ; toutes ces quantités sont de préférence complexes.

En prenant deux séries de valeurs $x_i, i = 1 \dots N$ et $u_j, j = 1 \dots M$ de sorte que (x_i, u_j) soit dense dans A^2 , on peut écrire alors en première approximation

$$f(x_i) = y(x_i) - \lambda \sum_{j=1}^M \int_A K(x_i, u_j) y(u_j) du$$

avec $\delta_i = (u_j - u_i)$

$$\Leftrightarrow f(x_i) = y(x_i) - \lambda \sum_{j=1}^M K(x_i, u_j) y(u_j) \delta_i \Leftrightarrow \mathbf{F} = (\mathbf{Id} - \lambda \delta \mathbf{K}) \mathbf{Y}$$

où l'écriture en gras représente soit des vecteurs de dimension N , soit des matrices $N \times M$.

Si la fonction F est nulle, le problème se ramène à la recherche des valeurs propres de \mathbf{K} , sinon le problème va se ramener à savoir si la matrice $(\mathbf{Id} - \lambda \mathbf{K})$ est inversible (*théorie spectrale*) : si tel est le cas on a une solution unique $\mathbf{Y} = (\mathbf{Id} - \lambda \mathbf{K})^{-1} \mathbf{F}$. Les méthodes de l'algèbre linéaire permettent normalement de trouver la réponse à la question dans les cas simples et en tout cas de trouver des approximations satisfaisantes. Malheureusement le déterminant de $(\mathbf{Id} - \lambda \mathbf{K})$ est souvent presque nul (*matrice presque non-inversible*) et le calcul est délicat.

Le *problème de Sturm-Liouville* concerne les valeurs du paramètre réel λ pour lesquelles l'équation différentielle linéaire homogène :

$$Ly - \lambda ry = 0$$

(où L est un opérateur différentiel d'ordre n à coefficients continus sur un intervalle compact $[a ; b]$ de \mathbb{R} et r une fonction continue strictement positive sur cet intervalle) a des solutions non nulles vérifiant n conditions aux limites données.

Si 0 n'est pas l'une de ces valeurs de λ , on définit une *fonction de Green* G du problème, continue sur $[a ; b]^2$; si l'on connaît G , la formule intégrale :

$$y(x) = \lambda \int_a^b G(x, u) f(u) du$$

donne la solution de l'équation non homogène : $Ly = f(x)$ vérifiant les conditions aux limites données, de sorte que le problème de Sturm-Liouville est transformé en l'équation intégrale homogène :

$$y(x) = \lambda \int_a^b G(x, u) r(u) y(u) du.$$

Le *problème de Dirichlet*, dans un ouvert borné D de \mathbb{R}^m , pour une fonction continue f donnée sur la frontière Γ de D , consiste à trouver la fonction, unique d'après le principe du maximum, continue sur : $\bar{D} = D \cup \Gamma$, harmonique sur D , qui coïncide avec f sur Γ . En 1877, C. G. Neumann proposait la méthode suivante pour la solution de ce problème, en supposant $m = 2$ et Γ pourvue d'une tangente continue ; on désignera par $L(\Gamma)$ la longueur de la courbe Γ :

Soit (x, y) le point courant de Γ , d'abscisse curviligne s , et (a, b) les cosinus directeurs de la normale en ce point orientée vers D .

On appelle potentiel de double couche d'une densité continue μ sur Γ la limite, quand δ tend vers 0, du quotient par 2δ de la différence entre le potentiel de la densité μ au point $(x + a\delta, y + b\delta)$ et celui de la densité $-\mu$ au point (x, y) . Le potentiel de double couche est la fonction :

$$h(X, Y) = \int_0^{L(\Gamma)} \mu(s) \frac{a(X-x) + b(Y-y)}{(X-x)^2 + (Y-y)^2} ds = \int_{\Gamma} \mu d\omega$$

où $d\omega$ est l'angle orienté sous lequel, du point $(X, Y) \in D$, on voit l'arc ds de Γ . Cette fonction est harmonique sur D et, en un point $(X, Y) \in \Gamma$, d'abscisse curviligne S , elle a pour limite :

$$\pi\mu(S) + \int_0^{L(\Gamma)} \mu(s) \frac{a(X-x) + b(Y-y)}{(X-x)^2 + (Y-y)^2} ds ;$$

en égalant cette limite à $f(X, Y)$, on obtient une équation intégrale non homogène où les variables sont S et s , mais sans paramètre λ .

Henri Poincaré pressentit, dès 1896, le rôle que ce paramètre jouerait dans les résultats ; cette intuition fut confirmée en 1903 par les remarquables travaux du Suédois Ivar Fredholm.

11-b : La méthode de Fredholm

On est donc en présence de l'équation $y(x) - \lambda \int_A K(x, u)y(u)du = f(x)$ où f est continue, λ un paramètre, K une fonction de deux variables, appelée le noyau ; on prendra A un intervalle compact de la forme $(a ; b)$ inclus dans $\mathbb{R} = \mathbb{R} \cup \{-\infty, +\infty\}$ avec éventuellement des bornes infinies ; le cas où A est compact mais de forme plus compliquée débouche sur la théorie des distributions, ce qui nous compliquerait un peu trop la vie. Nous choisissons de découper $(a ; b)$ en n parties égales, de sorte que l'on a à résoudre le système :

$$f(x_i) = y(x_i) - \lambda \sum_{j=1}^n K(x_i, x_j)y(x_j)\delta, \delta = (x_p - x_{p-1}), x_0 = a, x_n = b.$$

On applique la méthode vue ci-dessus, il faut donc trouver notre déterminant :

$$D_n(\lambda) = \det(\mathbf{Id} - \lambda\delta\mathbf{K}) = \begin{vmatrix} 1 - \lambda\delta K(x_1, x_1) & -\lambda\delta K(x_1, x_2) & \dots & -\lambda\delta K(x_1, x_n) \\ -\lambda\delta K(x_2, x_1) & 1 - \lambda\delta K(x_2, x_2) & \dots & -\lambda\delta K(x_2, x_n) \\ \dots & \dots & \dots & \dots \\ -\lambda\delta K(x_n, x_1) & -\lambda\delta K(x_n, x_2) & \dots & 1 - \lambda\delta K(x_n, x_n) \end{vmatrix} =$$

$$1 - \lambda\delta \sum_{p=1}^n K(x_p, x_p) + \frac{\lambda^2}{2!} \delta^2 \sum_{p,q=1}^n \begin{vmatrix} K(x_p, x_p) & K(x_p, x_q) \\ K(x_q, x_p) & K(x_q, x_q) \end{vmatrix} - \frac{\lambda^3}{3!} \delta^3 \sum_{p,q,r=1}^n \begin{vmatrix} K(x_p, x_p) & K(x_p, x_q) & K(x_p, x_r) \\ K(x_q, x_p) & K(x_q, x_q) & K(x_q, x_r) \\ K(x_r, x_p) & K(x_r, x_q) & K(x_r, x_r) \end{vmatrix} + \dots$$

que nous pouvons réécrire avec des intégrales ($\delta \rightarrow 0, n \rightarrow \infty$) :

$$D(\lambda) = 1 - \lambda \int_a^b K(u_1, u_1) du_1 + \frac{\lambda^2}{2!} \int_a^b \int_a^b \begin{vmatrix} K(u_1, u_1) & K(u_1, u_2) \\ K(u_2, u_1) & K(u_2, u_2) \end{vmatrix} du_1 du_2$$

$$- \frac{\lambda^3}{3!} \int_a^b \int_a^b \int_a^b \begin{vmatrix} K(u_1, u_1) & K(u_1, u_2) & K(u_1, u_3) \\ K(u_2, u_1) & K(u_2, u_2) & K(u_2, u_3) \\ K(u_3, u_1) & K(u_3, u_2) & K(u_3, u_3) \end{vmatrix} du_1 du_2 du_3 + \dots$$

Par ailleurs si $D_n(x_\mu, x_\nu ; \lambda)$ est le cofacteur de $K(x_\mu, x_\nu)$ dans $D_n(\lambda)$, la solution correspondante du système est

$$y(x_\mu) = \frac{1}{D_n(\lambda)} \left[\sum_{k=1}^n f(x_k) D_n(x_\mu, x_k) \right].$$

Cherchons alors la forme limite de $D_n(x_\mu, x_\nu ; \lambda)$: si $\mu \neq \nu$, on a

$$D_n(x_\mu, x_\nu; \lambda) =$$

$$\lambda \delta K(x_\mu, x_\nu) - \lambda^2 \delta^2 \sum_{p=1}^n \begin{vmatrix} K(x_\mu, x_\nu) & K(x_\mu, x_p) \\ K(x_p, x_\nu) & K(x_p, x_p) \end{vmatrix} + \frac{\lambda^3}{2!} \delta^3 \sum_{p,q=1}^n \begin{vmatrix} K(x_\mu, x_\nu) & K(x_\mu, x_p) & K(x_\mu, x_q) \\ K(x_p, x_\nu) & K(x_p, x_p) & K(x_p, x_q) \\ K(x_q, x_\nu) & K(x_q, x_p) & K(x_q, x_q) \end{vmatrix} + \dots$$

Passons à la limite :

$$D(x_\mu, x_\nu; \lambda) = \lambda K(x_\mu, x_\nu) - \lambda^2 \int_a^b \begin{vmatrix} K(x_\mu, x_\nu) & K(x_\mu, u_1) \\ K(u_1, x_\nu) & K(u_1, u_1) \end{vmatrix} du_1 \\ - \frac{\lambda^3}{3!} \int_a^b \int_a^b \begin{vmatrix} K(x_\mu, x_\nu) & K(x_\mu, u_1) & K(x_\mu, u_2) \\ K(u_1, x_\nu) & K(u_1, u_1) & K(u_1, u_2) \\ K(u_2, x_\nu) & K(u_2, u_1) & K(u_2, u_2) \end{vmatrix} du_1 du_2 + \dots$$

Ce qui donne finalement la solution

$$y(x) = \frac{1}{D(\lambda)} \left[\int_a^b D(x, u; \lambda) f(u) du \right].$$

Un exemple : prenons l'équation $y(x) = x + \lambda \int_0^1 xuy(u)du$.

Le noyau est $K(x, u) = xu$, d'où $\begin{vmatrix} K(u_1, u_1) & K(u_1, u_2) \\ K(u_2, u_1) & K(u_2, u_2) \end{vmatrix} = \begin{vmatrix} u_1^2 & u_1 u_2 \\ u_2 u_1 & u_2^2 \end{vmatrix} = 0$ ainsi que tous les autres

déterminants qui suivent. On a alors $D(\lambda) = 1 - \lambda \int_0^1 u^2 du = 1 - \frac{1}{3}\lambda$ et $D(x, u; \lambda) = \lambda xu$, soit la

$$\text{solution } y(x) = \frac{x}{1 - \frac{1}{3}\lambda}.$$

Il y a quand même un certain nombre de choses à préciser : regardons la convergence de la série :

$$D(\lambda) = 1 + \sum_{n=1}^{\infty} \frac{c_n}{n!} \lambda^n \quad \text{où } c_n = (-1)^n \int_a^b \dots \int_a^b \begin{vmatrix} K(u_1, u_1) & K(u_1, u_2) & \dots & K(u_1, u_n) \\ K(u_2, u_1) & K(u_2, u_2) & \dots & K(u_2, u_n) \\ \dots & \dots & \dots & \dots \\ K(u_n, u_1) & K(u_n, u_2) & \dots & K(u_n, u_n) \end{vmatrix} du_1 \dots du_n.$$

Un calcul de Hadamard (*Bulletin des Sciences Mathématiques*, 1893) montre que la valeur maximale du déterminant $D = |a_{ij}|_{i,j=1\dots n}$ où a_{mj} est réel et où $\sum_{j=1}^n a_{mj}^2 = 1, m = 1\dots n$ est 1, aussi, si dans

$D(\lambda)$ on divise tout par le maximum M de K , qui est continue et par conséquent bornée, on a

$$|K(u_i, u_j)| < M \quad \text{d'où } \sum_{j=1}^n (K(u_i, u_j))^2 \leq nM^2 \Rightarrow \sum_{j=1}^n \left(\frac{K^2(u_i, u_j)}{nM^2} \right) \leq 1 \Rightarrow \sum_{j=1}^n \left(\frac{K(u_i, u_j)}{\sqrt{n}M} \right)^2 \leq 1 \quad \text{et par}$$

chaque déterminant est majoré par $(\sqrt{n}M)^n$, soit $|c_n| \leq n^{\frac{n}{2}} M^n (b-a)^n$. Posons

$$\frac{n}{n^2} M^n (b-a)^n = n! d_n, \quad \text{alors } \lim_{n \rightarrow \infty} \frac{d_{n+1}}{d_n} = \lim_{n \rightarrow \infty} \frac{(b-a)M}{\sqrt{n+1}} \sqrt{\left(1 + \frac{1}{n}\right)^n} = 0 \quad \text{puisque } \lim_{n \rightarrow \infty} \left(1 + \frac{1}{n}\right)^n = e.$$

D est donc une fonction entière de λ .

Le même type de calcul sur $D(x, u; \lambda)$ montre que l'on a également une fonction entière de λ et en plus que la série $D(x, u; \lambda) - \lambda K(x, u)$ est une série uniformément convergente de fonctions continues en x et u sur (a, b) .

On peut alors écrire $D(x, u; \lambda) = \lambda D(\lambda)K(x, u) + \sum_{n=1}^{\infty} \frac{(-1)^n Q_n(x, u)}{n!} \lambda^{n+1}$ où

$$Q_n(x, u) = \int_a^b \dots \int_a^b \begin{vmatrix} 0 & K(x, u_1) & \dots & K(x, u_n) \\ K(u_1, u) & K(u_1, u_1) & \dots & K(u_1, u_n) \\ \dots & \dots & \dots & \dots \\ K(u_n, u) & K(u_n, u_1) & \dots & K(u_n, u_n) \end{vmatrix} du_1 \dots du_n .$$

Si on développe par rapport à la première colonne, on aura n termes de la forme $K(u_k, u)P_k$ où P_k

est de la forme $\begin{vmatrix} K(x, u_1) & K(x, u_2) & \dots & K(x, u_n) \\ K(u_1, u_1) & K(u_1, u_2) & \dots & K(u_1, u_n) \\ \dots & \dots & \dots & \dots \\ K(u_n, u_1) & K(u_n, u_2) & \dots & K(u_n, u_n) \end{vmatrix}$ avec un signe quelconque dû aux

interversions de lignes. De toutes manières tous ces déterminants sont égaux puisqu'ils contiennent les mêmes termes et

$$Q_n(x, u) = -n \int_a^b \dots \int_a^b K(u_1, u) P_k du_1 \dots du_n .$$

On a donc

$$D(x, u; \lambda) = \lambda D(\lambda)K(x, u) + \lambda \int_a^b D(x, u_1; \lambda)K(u_1, u) du_1 .$$

Considérons maintenant l'équation

$$y(u) = f(u) + \lambda \int_a^b K(u, z)y(z) dz ,$$

multiplions par $D(x, u; \lambda)$, intégrons et inversons les intégrales :

$$\int_a^b D(x, u; \lambda)y(u) du = \int_a^b D(x, u; \lambda)f(u) du + \lambda \int_a^b \left[\int_a^b D(x, u; \lambda)K(u, z) du \right] y(z) dz ,$$

on a alors

$$\begin{aligned} \int_a^b D(x, u; \lambda)f(u) du &= \int_a^b D(x, u; \lambda)y(u) du - \lambda \int_a^b [D(x, z; \lambda) - \lambda D(\lambda)K(x, z)] y(z) dz \\ &= \lambda D(\lambda) \int_a^b K(x, z)y(z) dz = D(\lambda)[y(x) - f(x)] . \end{aligned}$$

On peut donc conclure que si $D(\lambda) \neq 0$ et si l'équation de Fredholm a une solution, alors elle ne peut être autre que

$$y(x) = f(x) + \int_a^b f(u) \frac{D(x, u; \lambda)}{D(\lambda)} du .$$

De plus, si on remplace y par cette écriture dans l'équation, c'en est une solution ; il existe donc bien une unique solution continue à l'équation.

12. Noyaux de Fourier

12-a : Transformée de Mellin

La transformation qui à $f(x)$ associe $\hat{f}(u) = \int_a^b f(x)K(u, x)dx$ est évidemment linéaire et on peut la considérer comme un opérateur linéaire L associé à une matrice de dimension infinie \mathbf{M} : on aura une relation du type $L(f) = \hat{f}(u)$.

Si pour une fonction f , il existe une seule fonction \hat{f} telle que $L(f)(u) = \hat{f}(u)$ et réciproquement, alors cet opérateur est inversible et il existe un opérateur inverse L^{-1} tel que $f(x) = L^{-1}(\hat{f})(x)$: on doit alors avoir les deux relations :

$$\begin{cases} \hat{f}(u) = \int_a^b f(x)K(u, x)dx \\ f(x) = \int_{a'}^{b'} \hat{f}(u)H(u, x)du \end{cases}$$

Si on peut écrire la deuxième relation, on a un théorème d'inversion et on dira que K est un noyau de Fourier si $H = K$.

Le cas présentant le plus d'intérêt est lorsque $K(u, x) = K(x, u) = K(ux)$, le principal résultat concernant cette situation est dans le théorème suivant :

On appelle transformée de Mellin \mathbf{F} l'opérateur L dont le noyau est $K(u, x) = x^{u-1}$:

$$\mathbf{F}(u) = \int_0^{+\infty} f(x)x^{u-1}dx.$$

Une condition nécessaire pour que la fonction $K(ux)$ soit un noyau de Fourier est que la transformée de Mellin $\mathbf{K}(s)$ de la fonction $K(x)$ satisfasse l'équation fonctionnelle

$$\mathbf{K}(s)\mathbf{K}(1-s) = 1.$$

Multiplications des deux côtés $\hat{f}(u) = \int_a^b f(x)K(u, x)dx = \int_0^{+\infty} f(x)K(ux)dx$ par u^{s-1} et intégrons par rapport à u entre 0 et $+\infty$:

$$\int_0^{+\infty} \hat{f}(u)u^{s-1}du = \int_0^{+\infty} u^{s-1}du \int_0^{+\infty} f(x)K(ux)dx = \int_0^{+\infty} K(ux)u^{s-1}du \int_0^{+\infty} f(x)dx,$$

le passage de K d'une intégrale à l'autre peut-être justifié par le théorème de Fubini sous des conditions très larges. Posons $y = ux$:

$$\int_0^{+\infty} K(ux)u^{s-1}du = \int_0^{+\infty} K(y)\left(\frac{y}{x}\right)^{s-1}\left(\frac{1}{x}\right)dy = x^{-s} \int_0^{+\infty} K(y)y^{s-1}dy = x^{-s}\mathbf{K}(s)$$

et

$$\int_0^{+\infty} \hat{f}(u)u^{s-1}du = \mathbf{K}(s) \int_0^{+\infty} f(x)x^{-s}dx = \mathbf{K}(s)\mathbf{F}(1-s), \text{ soit } \hat{\mathbf{F}}(s) = \mathbf{K}(s)\mathbf{F}(1-s).$$

Mais l'intégrale de gauche est la transformée de Mellin de \hat{f} , aussi si K est bien un noyau de Fourier, on a $f(x) = \int_0^{+\infty} \hat{f}(u)K(ux)du$. Si on refait la même séquence d'opérations que précédemment, on a alors

$$\mathbf{F}(s) = \hat{\mathbf{F}}(1-s)\mathbf{K}(s) \Rightarrow \mathbf{F}(1-s) = \hat{\mathbf{F}}(1-1+s)\mathbf{K}(1-s) = \hat{\mathbf{F}}(s)\mathbf{K}(1-s) = \mathbf{K}(s)\mathbf{F}(1-s)\mathbf{K}(1-s)$$

et finalement $\mathbf{K}(s)\mathbf{K}(1-s) = 1$.

En fait cette condition peut être suffisante, mais nous en reparlerons plus loin.

12-b : Exemples de noyaux de Fourier

Les noyaux les plus connus sont évidemment ceux permettant de faire les transformées classiques : Fourier-cosinus, Fourier-sinus, Laplace, Mellin...

A titre d'exemple regardons le noyau $K(x) = a \cos x$ où a est une constante :

$$\mathbf{K}(s) = a \int_0^{+\infty} x^{s-1} \cos x dx = \frac{1}{2} \int_0^{+\infty} x^{s-1} e^{ix} dx + \frac{1}{2} \int_0^{+\infty} x^{s-1} e^{-ix} dx,$$

mais nous savons que $\int_0^{+\infty} x^{s-1} e^{-px} dx = \frac{\Gamma(s)}{p^s}$ d'où $\int_0^{+\infty} x^{s-1} e^{\pm ix} dx = \frac{\Gamma(s)}{\pm i^s} = e^{\pm \frac{\pi}{2} is} \Gamma(s)$; ceci donne alors

$$\mathbf{K}(s) = a \left[\frac{1}{2} e^{\frac{\pi}{2} is} \Gamma(s) + \frac{1}{2} e^{-\frac{\pi}{2} is} \Gamma(s) \right] = a \cos\left(\frac{\pi}{2} s\right) \Gamma(s),$$

$$\mathbf{K}(1-s) = a \cos\left(\frac{\pi}{2} - \frac{\pi}{2} s\right) \Gamma(s) = a \sin\left(\frac{\pi}{2} s\right) \Gamma(s)$$

$$\mathbf{K}(s)\mathbf{K}(1-s) = a^2 \cos\left(\frac{\pi}{2} s\right) \sin\left(\frac{\pi}{2} s\right) \Gamma(s)\Gamma(1-s) = \frac{1}{2} a^2 \sin(\pi s) \frac{\pi}{\sin(\pi s)} = \frac{\pi}{2} a^2.$$

Pour que K soit un noyau de Fourier, il faut donc que $a = \sqrt{\frac{2}{\pi}}$ et on aura alors

$$\begin{cases} \hat{f}(u) = \sqrt{\frac{\pi}{2}} \int_0^{+\infty} f(x) \cos(ux) dx \\ f(x) = \sqrt{\frac{\pi}{2}} \int_0^{+\infty} \hat{f}(u) \cos(ux) du \end{cases}.$$

Les calculs précédents sont assez formels dans l'ensemble et nécessitent des démonstrations un peu plus élaborées, mais dans l'ensemble ça marche comme ça. De la même manière on obtiendrait le même type de résultat avec $K(x) = a \sin x$.

Un exemple traité au chapitre Fonctions de Bessel est lorsque $K(x) = \sqrt{x} J_\nu(x)$ pour lequel on a

$$\mathbf{K}(s) = 2^{s-\frac{1}{2}} \frac{\Gamma\left(\frac{1}{2}\nu + \frac{1}{2}s + \frac{1}{4}\right)}{\Gamma\left(\frac{1}{2}\nu + \frac{1}{2}s + \frac{3}{4}\right)}; \text{ dans ce cas on est en présence de la } \underline{\text{transformée de Hankel}}.$$

12-c : Formules d'inversion non symétriques

Dans les équations
$$\begin{cases} \hat{f}(u) = \int_a^b f(x) K(u, x) dx \\ f(x) = \int_{a'}^{b'} \hat{f}(u) H(u, x) du \end{cases}$$
, rien n'oblige à ce que $K = H$. Par contre le même

calcul que précédemment amène au théorème suivant :

Une condition nécessaire pour que l'équation intégrale $\hat{f}(u) = \int_0^{+\infty} f(x) K(u, x) dx$ ait une solution de la forme $f(x) = \int_0^{+\infty} \hat{f}(u) H(u, x) du$ est que les transformées de Mellin $\mathbf{K}(s)$ et $\mathbf{H}(s)$ des fonctions $K(x)$ et $H(x)$ satisfassent l'équation fonctionnelle

$$\mathbf{K}(s)\mathbf{H}(1-s) = 1.$$

Par exemple le calcul de \mathbf{K} pour $K(x) = e^{-px}$ donne $\mathbf{K}(s) = \int_0^{+\infty} x^{s-1} e^{-px} dx = \frac{\Gamma(s)}{p^s}$, il faut donc trouver H pour que $\mathbf{K}(s)\mathbf{H}(1-s) = 1 \Leftrightarrow \mathbf{H}(1-s) = \frac{p^s}{\Gamma(s)} \Leftrightarrow \mathbf{H}(s) = \frac{p^{1-s}}{\Gamma(1-s)} = \frac{1}{\pi} p \sin(\pi s) \frac{\Gamma(s)}{p^s}$. Les choses ne sont donc visiblement pas très simples...

13. Convolution

(même appellation en anglais, *Faltung* en allemand).

Lorsque nous multiplions deux nombres que faisons-nous ? Nos deux nombres s'écrivent :

$$a = \overline{a_n \dots a_0} = a_n 10^n + \dots + a_1 10 + a_0 \text{ et } b = \overline{b_m \dots b_n \dots b_0} = b_m 10^m + \dots + b_n 10^n + \dots + b_1 10 + b_0 ;$$

effectuons la multiplication :

$$a.b = a_0 b_0 + (a_0 b_1 + a_1 b_0) 10 + (a_0 b_2 + a_1 b_1 + a_2 b_0) 10^2 + \dots$$

Le terme générique de cette série est

$$(a.b)_k = \sum_{i=\max(0, k-m)}^{i=\min(k, n)} a_i b_{k-i} \text{ pour } k = 0, \dots, m+n.$$

Evidemment dans le cas d'une multiplication on doit tenir compte des retenues, mais on peut comprendre que seul le résultat nous intéresse et que la question de l'écriture du nombre est sans importance : on effectue simplement la multiplication.

Notons maintenant $f(k) = a_k$ les termes de la suite (a_k) , $g(k)$ ceux de la suite (b_k) et x la base de

numération : on écrira alors (en mettant des zéros partout où c'est nécessaire) $a = \sum_{n=-\infty}^{+\infty} f(n)x^n$,

$b = \sum_{n=-\infty}^{+\infty} g(n)x^n$ et le produit des deux sera une série de terme général

$$(a * b)_k = \sum_{n=-\infty}^{+\infty} f(n)g(k-n)x^k ;$$

la notation $(a * b)$ représentant le *produit de convolution* des séries a et b . Si on passe alors aux intégrales, le produit en question sera représenté par l'intégrale :

$$(1) (f * g)(z) = \int_{-\infty}^{+\infty} g(u)f(z-u)du = \int_{-\infty}^{+\infty} g(z-u)f(u)du.$$

Supposons que f et g soient telles que l'on puisse leur faire une transformée de Fourier et que l'on puisse inverser l'ordre d'intégration dans l'intégrale (1) :

$$f(z-u) = \int_{-\infty}^{+\infty} \hat{f}(t)e^{-2\pi it(z-u)} dt \text{ ainsi que } g(u) = \int_{-\infty}^{+\infty} \hat{g}(t)e^{-2\pi itu} dt$$

d'où

$$\begin{aligned} \int_{-\infty}^{+\infty} g(u)f(z-u)du &= \int_{-\infty}^{+\infty} g(u) \left[\int_{-\infty}^{+\infty} \hat{f}(t)e^{-2\pi it(z-u)} dt \right] du \\ &= \int_{-\infty}^{+\infty} \hat{f}(t)e^{-2\pi itz} \left[\int_{-\infty}^{+\infty} g(u)e^{2\pi itu} du \right] dt = \int_{-\infty}^{+\infty} \hat{f}(t)\hat{g}(t)e^{-2\pi itz} dt. \end{aligned}$$

On a donc le *théorème de convolution* où $f * g$ est la *convoluée* de f et g :

$$(f * g)(z) = \int_{-\infty}^{+\infty} g(u)f(z-u)du = \int_{-\infty}^{+\infty} \hat{f}(t)\hat{g}(t)e^{-2\pi iz} dt .$$

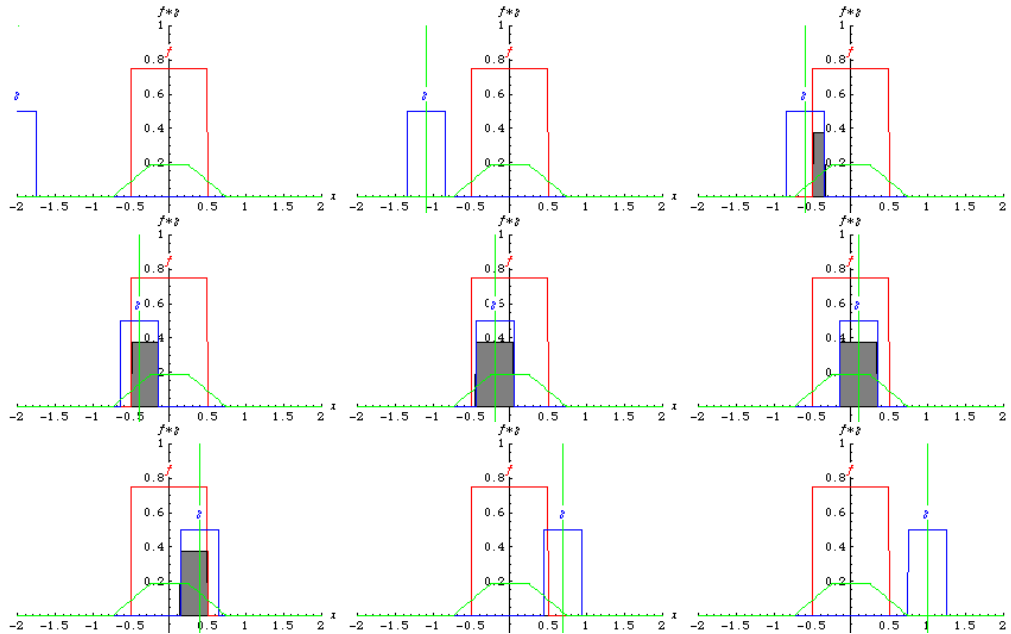


fig. 3 : Effets de la convolution de deux fonctions « rectangle », emprunté à E. Weisstein

<http://mathworld.wolfram.com/Convolution.html>

Si on fait $x = 0$, on a alors $\int_{-\infty}^{+\infty} g(u)f(-u)du = \int_{-\infty}^{+\infty} \hat{f}(t)\hat{g}(t)dt$

Si en plus on fait $f = g$, on a $\int_{-\infty}^{+\infty} f(u)f(-u)du = \int_{-\infty}^{+\infty} \hat{f}(t)\hat{f}(t)dt$.

En fait comme f est complexe, on a

$$f(-u) = \int_{-\infty}^{+\infty} \hat{f}(t)e^{2\pi i t u} dt \text{ et } \bar{f}(u) = \int_{-\infty}^{+\infty} \hat{\bar{f}}(t)e^{2\pi i t u} dt = \int_{-\infty}^{+\infty} \overline{\hat{f}(t)}e^{2\pi i t u} dt .$$

Considérons alors la fonction d'autocorrélation C définie par $C(z) = \int_{-\infty}^{+\infty} \bar{f}(u)f(u+z)du$ et remplaçons f et \bar{f} par leurs transformées de Fourier respectives :

$$\begin{aligned} C(z) &= \int_{-\infty}^{+\infty} \bar{f}(u)f(u+z)du = \int_{-\infty}^{+\infty} \left[\int_{-\infty}^{+\infty} \overline{\hat{f}(z')}e^{2\pi i z' u} dz' \right] \left[\int_{-\infty}^{+\infty} \hat{f}(z'')e^{-2\pi i z''(u+z)} dz'' \right] du \\ &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \overline{\hat{f}(z')}e^{2\pi i z' u} \hat{f}(z'')e^{-2\pi i z''(u+z)} dz' dz'' du = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \overline{\hat{f}(z')} \hat{f}(z'') e^{-2\pi i (z''-z')u} e^{-2\pi i z'' z} dz' dz'' du \\ &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \overline{\hat{f}(z')} \hat{f}(z'') \left[\int_{-\infty}^{+\infty} e^{-2\pi i (z''-z')u} du \right] e^{-2\pi i z'' z} dz' dz'' = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \overline{\hat{f}(z')} \hat{f}(z'') \delta(z''-z') e^{-2\pi i z'' z} dz' dz'' \\ &= \int_{-\infty}^{+\infty} \overline{\hat{f}(z')} \hat{f}(z) e^{-2\pi i z z} dz = \int_{-\infty}^{+\infty} |\hat{f}(z)|^2 e^{-2\pi i z z} dz \\ &= \widehat{|\hat{f}(z)|^2} \end{aligned}$$

C'est le théorème de Wiener-Khinchin.

Un résultat également important est :

$$\int_{-\infty}^{+\infty} (f * g)(u) du = \int_{-\infty}^{+\infty} \left[\int_{-\infty}^{+\infty} g(v) f(u-v) dv \right] du = \int_{-\infty}^{+\infty} g(v) dv \int_{-\infty}^{+\infty} f(w) dw .$$

La convolution est un outil fondamental en traitement du signal et en électronique. La multiplication rapide sur les ordinateurs utilise la *transformée de Fourier rapide* en lien avec la convolution de deux nombres, etc.