



# Polynômes orthogonaux

**Mathieu Lavoie**

**Maîtrise en mathématiques**

Québec, Canada

© Mathieu Lavoie, 2015



# Résumé

Les polynômes orthogonaux sont introduits par la théorie de Sturm-Liouville, puis les équivalences existantes entre leurs définitions classiques sont montrées. Certains résultats de base de la théorie sont ensuite décortiqués. On termine en introduisant des résultats préliminaires de la théorie analytique des polynômes, qui étudie les liens entre les coefficients d'un polynôme, ses zéros et ses points critiques.



# Table des matières

<b>Résumé</b>	<b>iii</b>
<b>Table des matières</b>	<b>v</b>
<b>Introduction</b>	<b>1</b>
<b>1 Théorie de Sturm-Liouville</b>	<b>3</b>
1.1 Problèmes issus de la physique . . . . .	3
1.2 Contexte et outils utilisés . . . . .	7
1.3 Opérateur de Sturm-Liouville . . . . .	9
1.4 Quelques exemples classiques . . . . .	12
<b>2 Définitions équivalentes des polynômes orthogonaux</b>	<b>15</b>
2.1 Une première définition . . . . .	15
2.2 Cas des solutions d'un problème de Sturm-Liouville . . . . .	18
2.3 Formule de Rodrigues . . . . .	20
2.4 Relation de récurrence . . . . .	24
2.5 Séries génératrices . . . . .	25
2.6 Polynômes orthogonaux moins connus . . . . .	31
<b>3 Propriétés des polynômes orthogonaux</b>	<b>37</b>
3.1 Formule de Christoffel . . . . .	37
3.2 Séries de Fourier généralisées, propriétés d'extrema et de fermeture . . . . .	40
3.3 Noyau de Christoffel-Darboux . . . . .	43
3.4 Quadratures de Gauss . . . . .	46
<b>4 Zéros et points critiques</b>	<b>51</b>
4.1 Zéros des polynômes orthogonaux . . . . .	51
4.2 Polynômes généraux . . . . .	54
4.3 Classification des points critiques et interprétation physique de ceux de type 2	61
4.4 Théorème de Gauss-Lucas et une réciproque . . . . .	62
<b>Conclusion</b>	<b>71</b>
<b>A Table de polynômes orthogonaux classiques</b>	<b>73</b>
A.1 Polynômes de Legendre . . . . .	73
A.2 Polynômes de Gegenbauer (ou ultrasphériques) . . . . .	74
A.3 Polynômes de Jacobi . . . . .	74

A.4 Polynômes de Tchebychev (première espèce) . . . . .	75
A.5 Polynômes de Tchebychev (seconde espèce) . . . . .	76
A.6 Polynômes de Hermite (forme physique) . . . . .	76
A.7 Polynômes de Laguerre . . . . .	77
A.8 Polynômes de Laguerre associés . . . . .	78
<b>B Schéma d'Askey</b>	<b>79</b>
<b>C Rappel sur l'interpolation de Lagrange</b>	<b>81</b>
<b>D Zéros et points critiques des polynômes de Legendre</b>	<b>83</b>
<b>Pages Wikipédia utiles</b>	<b>85</b>
<b>Bibliographie</b>	<b>87</b>

*À Rachel*



# Introduction

Les polynômes orthogonaux sont un sujet d'étude pour les mathématiciens depuis longtemps. À titre d'exemple, Adrien-Marie Legendre en était arrivé dès le début du XIX<sup>e</sup> siècle à considérer la suite de polynômes auxquels son nom est maintenant associé, les polynômes de Legendre, dans le cadre de ses calculs concernant la mécanique céleste. Depuis cette époque jusqu'à aujourd'hui, la théorie concernant ces polynômes n'a cessé de croître en précision et aussi en importance, puisque d'autres applications se sont développées. En effet, avec l'avènement des ordinateurs, les polynômes orthogonaux sont devenus des outils d'approximation et d'encodage-décodage très utiles. C'est pourquoi leur étude continue encore aujourd'hui.

Ce mémoire se veut une (brève) introduction à l'étude des polynômes orthogonaux, qui soit accessible et rigoureuse. Aussi, les définitions puis les théorèmes seront développés dans le cadre le plus général possible.

Le premier chapitre introduit le sujet sensiblement comme il semble être apparu historiquement. En effet, certaines suites de polynômes orthogonaux peuvent être vues comme des solutions d'équations différentielles. Des problèmes qui mènent à de telles équations différentielles sont donc d'abord proposés. Ensuite, le contexte utilisé pour notre étude est introduit. Après cela, nous étudions de manière systématique les opérateurs de Sturm-Liouville (qui mènent à des suites de polynômes orthogonaux), et enfin nous en donnons quelques exemples.

Le second chapitre relie ensemble les définitions usuelles qu'une suite de polynômes orthogonaux peut avoir. Nous commençons par donner une première définition, simplement via la notion de produit scalaire, et quelques conséquences de celle-ci. Ceci fait, nous situons les solutions d'un problème de Sturm-Liouville dans ce nouveau cadre. Nous continuons en montrant les équivalences existantes entre ce qui précède et la formule de Rodrigues, puis la relation de récurrence permettant de calculer un polynôme dans une suite à partir des deux précédents. Il suit une discussion sur les séries génératrices des polynômes orthogonaux. Nous terminons ce chapitre par l'étude d'une classe de polynômes orthogonaux moins connue, à savoir ceux issus d'un processus discret, qui sont néanmoins tout aussi utiles que leurs congénères.

Nous décrivons dans le troisième chapitre quelques propriétés classiques des polynômes orthogonaux. Une formule de Christoffel est établie ; les séries de Fourier généralisées sont introduites

avec certaines de leurs conséquences ; le noyau de Christoffel-Darboux est étudié et finalement, nous présentons les quadratures de Gauss. Elles forment en effet un outil important, autant dans les applications que dans la suite de la théorie.

Dans le quatrième et dernier chapitre, nous nous aventurons dans un domaine de recherche très actif : celui de la recherche de lien entre les coefficients, les zéros et les points critiques des polynômes. Des résultats existants pour les polynômes orthogonaux sont d'abord présentés, puis nous élargissons notre perspective pour étudier des résultats de base de la théorie analytique des polynômes généraux, à coefficients dans  $\mathbb{C}$ . Ces résultats nous mèneront au théorème de Gauss-Lucas, puis à une de ses réciproques, récemment démontrée, qui sera le point culminant du chapitre.

Pour référence, nous avons placé en Annexe A une table des suites de polynômes rencontrées tout au long du mémoire contenant diverses informations, notamment pour l'utilisation des définitions présentées, ainsi que quelques commentaires sur les particularités des polynômes de chaque sorte. Le schéma d'Askey, qui est mentionné en 2.6, est représenté dans l'annexe B. L'Annexe C est un rappel sur l'interpolation de Lagrange, qui sera utilisée en 3.4. Enfin, j'ai rassemblé dans l'annexe D quelques graphiques des polynômes de Legendre illustrant des théorèmes du chapitre 4. Nous y référerons en temps voulu.

# Chapitre 1

## Théorie de Sturm-Liouville

Dans cette première section, nous allons étudier un type de problème d'où provient une partie importante des polynômes orthogonaux classiques. Nous verrons d'abord deux exemples d'équations tirées de la physique dont la résolution mène à des problèmes dits de Sturm-Liouville. Dans les sections suivantes, nous généraliserons notre approche. Nous regarderons d'abord avec quels outils nous travaillerons, puis nous étudierons les propriétés des opérateurs de Sturm-Liouville, et enfin leurs solutions.

### 1.1 Problèmes issus de la physique

Le texte de cette section est inspiré de (5) et de (1).

#### 1.1.1 Équation de Laplace

Considérons l'opérateur de Laplace en trois dimensions

$$\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}.$$

Cet opérateur est utilisé par exemple pour modéliser le comportement d'une distribution de température, de charge électrique, ou de potentiel gravitationnel sur un domaine donné (voir (5, p. 196)). Plus précisément, considérons par exemple l'équation suivante

$$\nabla^2 f = c \frac{\partial f}{\partial t},$$

où  $c$  est une constante positive et  $f = f(x, y, z, t)$  est une fonction de quatre variables. Sous certaines conditions,  $f$  représente la température d'un matériau homogène isotrope. Avec la constante  $c$  appropriée, il s'agit en effet de l'« équation de la chaleur », telle que donnée dans (7) par Jean-Baptiste-Joseph Fourier lui-même. Sans analyser en profondeur le problème comme l'a fait Fourier, nous pouvons vérifier que le modèle se comporte avec une certaine cohérence par l'argument suivant. L'opérateur peut être vu comme donnant la différence entre la valeur

moyenne de la température dans le voisinage d'un point et la température en ce point. Ainsi, si  $\nabla^2 f(x) > 0$ , la température en  $x$  est inférieure à celle autour, et le point aura tendance à se réchauffer, et donc  $\frac{\partial f}{\partial t} > 0$ .

Maintenant que nous sommes confiants de l'utilité physique de notre discussion, résolvons l'équation

$$\nabla^2 f = 0.$$

Comme, par nature, les problèmes de ce type s'appliquent sur une sphère autour d'un point, nous passons en coordonnées polaires. L'équation devient

$$\frac{\partial}{\partial r} \left( r^2 \frac{\partial f}{\partial r} \right) + \frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left( \sin \theta \frac{\partial f}{\partial \theta} \right) + \frac{1}{\sin^2 \theta} \frac{\partial^2 f}{\partial \phi^2} = 0.$$

Utilisons la séparation de variables en posant

$$f(r, \theta, \phi) = R(r)\Theta(\theta)\Phi(\phi).$$

Il suit les trois équations suivantes

$$\begin{aligned} \frac{-1}{\Phi} \frac{d^2 \Phi}{d\phi^2} &= m_l^2, \\ \frac{1}{R} \frac{d}{dr} \left( r^2 \frac{dR}{dr} \right) &= l(l+1), \\ -\frac{1}{\sin \theta} \frac{d}{d\theta} \left( \sin \theta \frac{d\Theta}{d\theta} \right) + \frac{m_l^2 \Theta}{\sin^2 \theta} &= l(l+1)\Theta. \end{aligned}$$

En résolvant la première équation, nous trouvons que  $\Phi_{m_l}$  est bien définie pour  $\phi \in \mathbb{R}$  si et seulement si  $m_l \in \mathbb{Z}$ . Intéressons-nous maintenant à la troisième équation. Nous posons  $w = \cos \theta$  et il vient

$$\frac{d}{dw} \left[ (1-w^2) \frac{d\Theta}{dw} \right] + \left[ l(l+1) - \frac{m_l^2}{1-w^2} \right] \Theta = 0. \quad (1.1)$$

Les solutions de cette équation sont appelées « fonctions de Legendre associées ». Pour les trouver, posons  $\Theta_{lm_l}(w) = (1-w^2)^{|m_l|/2} \frac{d^{|m_l|} P_l(w)}{dw^{|m_l|}}$  et calculons, en utilisant les notations suivantes pour alléger les équations.

$$\begin{aligned} \Gamma &:= \frac{d^{|m_l|} P_l(w)}{dw^{|m_l|}}; \\ \lambda &:= |m_l|; \\ D^n &:= \frac{d^n}{dw^n}. \end{aligned}$$

Allons-y étape par étape. Nous avons

$$\begin{aligned}
\Theta &= (1 - w^2)^{\lambda/2} \Gamma; \\
D\Theta &= -\lambda w(1 - w^2)^{\lambda/2-1} \Gamma + (1 - w^2)^{\lambda/2} D\Gamma; \\
(1 - w^2)D\Theta &= -\lambda w(1 - w^2)^{\lambda/2} \Gamma + (1 - w^2)^{\lambda/2+1} D\Gamma; \\
D[(1 - w^2)D\Theta] &= -\lambda(1 - w^2)^{\lambda/2} \Gamma + \lambda^2 w^2(1 - w^2)^{\lambda/2-1} \Gamma - \lambda w(1 - w^2)^{\lambda/2} D\Gamma \\
&\quad - 2w\left(\frac{\lambda}{2} + 1\right)(1 - w^2)^{\lambda/2} D\Gamma + (1 - w^2)^{\lambda/2+1} D^2\Gamma.
\end{aligned}$$

En faisant le changement de variable, l'équation (1.1) devient donc, après quelques manipulations simples

$$(1 - w^2)^{\lambda/2} [(1 - w^2)D^2\Gamma - 2w(\lambda + 1)D\Gamma + (l(l + 1) - \lambda(\lambda + 1))\Gamma] = 0,$$

ou encore

$$(1 - w^2)D^2\Gamma - 2w(\lambda + 1)D\Gamma + (l(l + 1) - \lambda(\lambda + 1))\Gamma = 0,$$

Nous remarquons maintenant qu'il s'agit de la dérivée d'ordre  $\lambda$  de l'équation suivante, appelée l'équation de Legendre,

$$(1 - w^2) \frac{d^2 P_l}{dw^2} - 2w \frac{dP_l}{dw} + l(l + 1)P_l = 0,$$

ou, sous une autre forme,

$$\frac{d}{dw} \left[ (1 - w^2) \frac{dP_l}{dw} \right] + l(l + 1)P_l = 0.$$

Nous reconnaitrons cette dernière comme un problème de Sturm-Liouville.

### 1.1.2 L'oscillateur harmonique

Un système physique est appelé un *oscillateur harmonique* si le potentiel associé est quadratique, de la forme  $V(x) = cx^2$  (pour un système en une dimension), où  $c$  est une constante. L'exemple le plus connu d'un tel système est sans doute le montage masse-ressort. Ce potentiel est particulièrement important en physique car un grand nombre d'autres systèmes peuvent être décrits par ce potentiel, au moins dans une certaine limite d'application. Il y a une raison fort simple à cela, qu'amène l'argument que voici. Considérons un potentiel  $U(x)$  arbitraire, mais possédant un minimum en  $x = x_0$  (c'est un *point d'équilibre*). En développant  $U(x)$  en série entière, ainsi que les physiciens le font souvent, il vient

$$U(x) = U(x_0) + U'(x_0)(x - x_0) + \frac{1}{2}U''(x_0)(x - x_0)^2 + O((x - x_0)^3).$$

Maintenant, comme il y a un minimum en  $x_0$ , il suit  $U'(x_0) = 0$  et  $U''(x_0) \geq 0$ , et comme de plus le potentiel peut toujours être redéfini pour avoir  $U(x_0) = 0$ , cela peut en fait s'écrire

$$U(x) = U(x_0) \frac{1}{2}U''(x_0)(x - x_0)^2 + O((x - x_0)^3) \quad (x \leftarrow x_0).$$

Si les mouvements de la particule autour de  $x_0$  ont une amplitude suffisamment petite pour que les termes en  $O((x-x_0)^3)$  soient négligeables devant celui en  $(x-x_0)^2$ , nous pouvons considérer qu'à toute fin pratique nous avons affaire à un oscillateur harmonique. C'est comme cela que sont étudiées les vibrations des atomes d'une molécule autour de leur position d'équilibre, ou encore les oscillations des atomes ou des ions d'un réseau cristallin ( voir (4, p. 481)). Ce potentiel sert aussi à donner une première approximation du comportement d'un potentiel au voisinage d'un point d'équilibre.

Ceci étant dit, introduisons l'équation de Schrödinger qui, en mécanique quantique, régit le comportement des particules ayant une certaine masse  $m$ , placées dans un certain potentiel  $V$ . Rappelons également que  $\hbar$  est la constante de Planck.

$$i\hbar \frac{\partial}{\partial t} \Psi(\vec{x}, t) = -\frac{\hbar^2}{2m} \nabla^2 \Psi(\vec{x}, t) + V(\vec{x}, t) \Psi(\vec{x}, t).$$

Pour les fins du présent exemple, il nous suffira de travailler avec la version en une dimension, avec le potentiel  $V(x) := \frac{\hbar^2}{2m} x^2$ , c'est-à-dire un oscillateur harmonique tel que présenté ci-haut, indépendant du temps, mais avec une constante de rappel bien choisie. L'équation précédente devient

$$i\hbar \frac{\partial}{\partial t} \Psi(x, t) = -\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} \Psi(x, t) + V(x) \Psi(x, t).$$

Pour extraire de cette équation la dépendance au temps, utilisons la séparation de variables. Posons donc que

$$\Psi(x, t) = \psi(x) \chi(t).$$

L'équation de Schrödinger devient

$$\begin{aligned} i\hbar \psi(x) \frac{d\chi(t)}{dt} &= -\frac{\hbar^2}{2m} \chi(t) \frac{d^2\psi(x)}{dx^2} + \chi(t) V(x) \psi(x) \\ \iff \frac{i\hbar}{\chi(t)} \frac{d\chi(t)}{dt} &= \frac{1}{\psi(x)} \left[ -\frac{\hbar^2}{2m} \frac{d^2\psi(x)}{dx^2} \right] + V(x). \end{aligned}$$

En utilisant  $E := \hbar\omega$  comme constante de séparation, nous obtenons

$$\begin{cases} i\hbar \frac{d\chi(t)}{dt} = \hbar\omega \chi(t) \implies \chi(t) = A e^{-i\omega t} \\ -\frac{\hbar^2}{2m} \frac{d^2\psi(x)}{dx^2} + V(x) \psi(x) = \hbar\omega \psi(x). \end{cases}$$

Nous n'avons fait tout ce chemin que pour aboutir à la seconde équation, qui est pertinente pour la présente étude. Réécrivons-la, en insérant le potentiel choisi précédemment. Cela donne

$$\frac{d^2\psi(x)}{dx^2} + \left( \frac{2m\omega}{\hbar} - x^2 \right) \psi(x) = 0.$$

Posons  $\frac{2m\omega}{\hbar} = 2n + 1$ . Nous obtenons

$$\frac{d^2\psi}{dx^2} + ((2n + 1) - x^2) \psi = 0,$$

En posant maintenant  $u := \exp(\frac{x^2}{2})\psi(x)$ , l'équation précédente devient

$$\frac{d^2u}{dx^2} - 2x\frac{du}{dx} + 2nu = 0,$$

à savoir l'équation de Hermite (dont les solutions utiles ici sont les polynômes de Hermite).

Elle peut être réécrite

$$\frac{d}{dx} \left[ e^{-x^2} \frac{du}{dx} \right] + 2ne^{-x^2}u = 0.$$

Encore une fois, la dernière forme est celle d'un problème de Sturm-Liouville.

## 1.2 Contexte et outils utilisés

Le texte de cette section est inspiré de (2, chap. 7) et de (15, section G5). Soit  $X$  un espace mesurable et  $\mu$  une mesure positive sur  $X$ . Nous travaillerons sur l'espace

$$L^2(\mu) := \{f : X \rightarrow \mathbb{R} : f \text{ mesurable, } \|f\| < \infty\},$$

où la norme est celle induite par le produit scalaire

$$\langle f|g \rangle := \int_X fg d\mu,$$

Naturellement, nous travaillons modulo la relation d'équivalence « être égal presque partout ».

Aussi, l'inégalité de Cauchy-Schwarz

$$|\langle f|g \rangle| \leq \|f\| \|g\|$$

assure que ce produit scalaire est bien défini pour toutes fonctions  $f, g \in L^2(\mu)$ . Nous appellerons  $m$  la mesure de Lebesgue.

Nous utiliserons pour notre analyse les mesures de *Lebesgue-Stieltjes*, et l'espace  $X$  consistera le plus souvent en un intervalle appelé *segment fondamental*. Nous supposons dans la suite le lecteur familier avec la théorie de l'intégration de Lebesgue. Ceci dit, révisons maintenant un peu la théorie concernant lesdites mesure de Lebesgue-Stieltjes.

**Théorème 1.1.** *Soit  $F : \mathbb{R} \rightarrow \mathbb{R}$  croissante et continue à droite. Alors il existe une unique mesure  $m_F$  sur  $\mathcal{B}(\mathbb{R})$  telle que  $m_F((a, b]) = F(b) - F(a)$  pour tout  $a < b$ .*

La définition suivante vient alors.

**Définition 1.2.**

- (i) La mesure de Lebesgue-Stieltjes associée à  $F$  est appelée  $m_F$ .
- (ii) On note  $\int f dF := \int f dm_F$ .

**Remarque 1.3.** Dans le point (ii) de la définition qui précède, nous voyons vraiment le membre de gauche de l'égalité comme une notation. Dans le théorème suivant, nous verrons par contre que cette notation est heureuse et correspond bien, dans les cas où  $F \in C^1$ , à la notation différentielle que nous connaissons.

**Théorème 1.4** (de calcul pour les intégrales de Lebesgue-Stieltjes). *Soit  $g$  une fonction mesurable sur un espace mesurable  $X$ , et soit  $F \in C^1$  telle que  $F' \geq 0$ . Alors*

$$(i) \quad m_F(A) = \int_A F' dm,$$

$$(ii) \quad \int_A g dm_F = \int_A g F' dm.$$

*Démonstration.*

- (i) Pour chaque ensemble mesurable  $A$ , posons  $\mu(A) = \int_A F' dm$ . Alors  $\mu$  est une mesure borélienne et

$$\mu((a, b]) = \int_{(a, b]} F' dm = \int_a^b F'(x) dx = F(b) - F(a),$$

la dernière égalité provenant du théorème fondamental du calcul. La mesure  $m_F$  étant l'unique mesure avec cette propriété, nous en déduisons  $m_F = \mu$ .

- (ii) C'est un type de preuve courant en théorie de la mesure. Le résultat est montré pour les fonctions (a) étagées, (b) positives, (c) intégrables.

(a) Supposons que  $g = \sum_{i=1}^n c_i \mathbb{1}_{E_i}$  est une fonction étagée sous forme canonique, pour des  $E_i \subseteq X$  formant une partition de  $X$ . Alors

$$\begin{aligned} \int_X g dm_F &\stackrel{\text{def}}{=} \sum_{i=1}^n c_i m_F(E_i) \stackrel{(i)}{=} \sum_{i=1}^n c_i \int_{E_i} F' dm = \sum_{i=1}^n c_i \int_X F' \mathbb{1}_{E_i} dm \\ &= \int_X \sum_{i=1}^n c_i \mathbb{1}_{E_i} F' dm = \int_X g F' dm. \end{aligned}$$

(b) Supposons maintenant  $g$  positive. Alors par le théorème d'approximation, il existe  $(\phi_k)$ , une suite de fonctions étagées telles que  $\phi_k \uparrow g$ , et alors par le théorème de convergence monotone

$$\int_X \phi_k dm_F = \int_X \phi_k F' dm \uparrow \int_X g F' dm.$$

En effet,  $F' \geq 0 \implies \phi_k F' \uparrow g F'$ .

(c) Pour  $g$  intégrable,

$$\begin{aligned} \int_X g dm_F &= \int_X g^+ dm_F - \int_X g^- dm_F \\ &= \int_X g^+ F' dm - \int_X g^- F' dm \\ &= \int_X g F' dm. \end{aligned}$$

□

**Définition 1.5.** Nous dirons qu'une suite de fonctions  $(f_n)$  converge vers  $f$  dans  $L^2$  si

$$\lim_{n \rightarrow \infty} \|f_n - f\| = 0.$$

Nous écrirons  $f_n \xrightarrow{L^2} f$ .

Finalement, nous tenons pour acquis les résultats suivants :

- $C[a, b]$  est dense dans  $L^2[a, b]$ .
- $L^2$  est complet.

### 1.3 Opérateur de Sturm-Liouville

Dans cette section, nous cherchons sous quelles conditions un opérateur différentiel du second ordre est auto-adjoint. La motivation de cette quête se trouve dans le théorème selon lequel les valeurs propres d'un tel opérateur sont réelles et ses vecteurs propres sont orthogonaux deux à deux dans  $L^2[a, b]$ .

Nous considérons donc une EDO du second ordre de la forme

$$p(x)y'' + q(x)y' + r(x)y = f(x), \quad (1.2)$$

où  $p, q, r$  sont des fonctions définies sur un intervalle (potentiellement infini)  $I := (a, b)$ , et sur lesquelles nous ne supposons rien en terme de continuité ou de dérivabilité à ce stade-ci, de manière à garder la théorie la plus générale possible. Notons  $L$  l'opérateur linéaire du second ordre

$$L := p(x) \frac{d^2}{dx^2} + q(x) \frac{d}{dx} + r(x).$$

L'EDO précédente s'écrit donc simplement

$$Ly = f(x).$$

Nous en rencontrerons par ailleurs souvent dans la suite sous la forme

$$\frac{1}{p(x)} Ly = \frac{f(x)}{p(x)}.$$

Cela motive la définition suivante. Si  $p(x) \neq 0$  pour tout  $x \in I$ , alors l'EDO est dite *régulière*. Sinon, elle est dite *singulière* et alors les points  $x$  où  $p(x) = 0$  sont appelés *singularités*. Nous pourrions remarquer dans la suite que les équations menant à plusieurs types de polynômes orthogonaux classiques sont singulières.

**Définition 1.6.** Soit  $L$  un opérateur et  $m_F$  une mesure de Lebesgue-Stieltjes. Posons  $w := F'$  et  $L^2(w) := L^2(m_F)$ .

- (i) Nous appelons l'*adjoint* de  $L$  dans  $L^2(w)$  l'opérateur  $L^\dagger$  tel que  $\langle Lf|g \rangle = \langle f|L^\dagger g \rangle$   $\forall f, g \in L^2(w)$ .
- (ii)  $L$  est *auto-adjoint* dans  $L^2(w)$  si  $L = L^\dagger$ .
- (iii) Nous appelons l'*adjoint formel* de  $L := \sum_{k=0}^n a_k(x) \frac{d^k}{dx^k}$  l'opérateur  $L^*u := \sum_{k=0}^n (-1)^k \frac{d^k}{dx^k} [\overline{a_k(x)}u]$ , où les fonctions  $a_k(x)$  sont à valeurs dans  $\mathbb{C}$ .
- (iv)  $L$  est *formellement auto-adjoint* dans  $L^2(w)$  si  $wL(f) = L^*(wf) \forall f \in L^2(w)$ .

Comme prévu, nous voulons connaître les conditions sous lesquelles  $L$  est auto-adjoint. Cela nous mène au théorème suivant, inspiré de (1).

**Théorème 1.7.** Si dans (1.2), nous avons que  $p \in C^2(I)$ ,  $q \in C^1(I)$  et  $r \in C(I)$  et que  $r$  et  $q$  sont à valeurs réelles, alors

- (i)  $L$  est formellement auto-adjoint dans  $L^2(w)$  si et seulement si  $p$  est réelle et  $wq \equiv (wp)'$  ;
- (ii) si  $L$  est formellement auto-adjoint et

$$wp(u'\bar{v} - u\bar{v}')|_a^b = 0, \quad \forall u, v \in C^2(I) \cap L^2(I),$$

alors  $L$  est auto-adjoint ;

- (iii) si  $L$  est auto-adjoint dans  $L^2(w)$ , alors les valeurs propres de  $-L$ , c'est-à-dire les valeurs  $\lambda$  qui solutionnent

$$Lu + \lambda u = 0$$

sont réelles et les fonctions propres associées à des valeurs propres distinctes sont orthogonales deux à deux.

*Démonstration.* Dans l'ordre :

- (i) Par définition,  $L$  est formellement auto-adjoint dans  $L^2(w)$  si

$$\begin{aligned} wLu &= L^*wu \Leftrightarrow \\ wpu'' + wqu' + wru &= (\bar{p}wu)'' - (\bar{q}wu)' + (\bar{r}wu) \\ &= \bar{p}wu'' + (2(\bar{p}w)' - \bar{q}w)u' + ((\bar{p}w)'' - (\bar{q}w)' + \bar{r}w)u \end{aligned}$$

L'égalité précédente doit être vérifiée si et seulement si

$$\begin{aligned} wp &= \bar{p}w \\ wq &= 2(\bar{p}w)' - w\bar{q} \\ wr &= (\bar{p}w)'' - (\bar{q}w)' + \bar{r}w, \end{aligned}$$

Il est clair qu'il est suffisant que  $\bar{p} = p$  et que  $wq = (wp)'$  pour que ces équations soient satisfaites. Montrons maintenant que ces conditions sont en fait nécessaires. Comme  $w$  est à valeurs réelles, la première des trois équations donne que  $p = \bar{p}$ . En se rappelant de plus l'hypothèse selon laquelle  $r$  et  $q$  sont à valeurs réelles, les deux autres peuvent se réécrire

$$\begin{aligned} 2wq &= 2(pw)' \\ (wp)'' - (wq)' &= wr - wr = 0, \end{aligned}$$

d'où il suit évidemment que  $(wp)' = wq$ .

- (ii) Faisons d'abord le calcul suivant, pour  $u, v \in L^2(I)$ , dont le but est de faire passer  $L$  de l'autre côté du produit scalaire. En utilisant l'intégration par parties,

$$\begin{aligned} \langle Lu|v \rangle &= \int_I (pu'' + qu' + ru)\bar{v}w dm \\ &= p\bar{v}wu'|_a^b - \int_I u'(p\bar{v}w)' dm + q\bar{v}wu|_a^b - \int_I u(q\bar{v}w)' dm + \int_I ru\bar{v}w dm \\ &= [p\bar{v}wu' - (p\bar{v}w)'u]|_a^b + \int_I u(p\bar{v}w)'' dm + q\bar{v}wu|_a^b - \int_I u(q\bar{v}w)' dm + \int_I ru\bar{v}w dm \\ &= \left\langle u \left| \frac{1}{w} [(\bar{p}vw)'' - (\bar{q}vw)' + \bar{r}vw] \right. \right\rangle + [pw((u)'\bar{v} - (u)\bar{v}') + (wq - (pw)')w\bar{v}]|_a^b \\ &= \left\langle u \left| \frac{1}{w} L^*(wv) \right. \right\rangle + [wp(u'\bar{v} - u\bar{v}') + (wq - (pw)')u\bar{v}]|_a^b. \end{aligned}$$

Comme  $L$  est supposé formellement auto-adjoint,  $(wp)' = wq$  et nous pouvons écrire

$$\langle Lu|v \rangle = \langle u|Lv \rangle + [wp(u'\bar{v} - u\bar{v}')]|_a^b,$$

et le résultat suit de l'hypothèse

$$wp(u'\bar{v} - u\bar{v}')|_a^b = 0, \quad \forall u, v \in C^2(I) \cap L^2(I).$$

- (iii) Supposons que  $\lambda \in \mathbb{C}$  est une valeur propre de  $-L$ . Alors il existe une fonction propre  $u \neq 0$  telle que  $Lu + \lambda u = 0$ . Alors, comme  $L$  est auto-adjoint,

$$\begin{aligned} \lambda \|u\|^2 &= \langle \lambda u|u \rangle = -\langle Lu|u \rangle \\ &= -\langle u|Lu \rangle = -\langle u|\lambda u \rangle \\ &= \bar{\lambda} \|u\|^2. \end{aligned}$$

Comme  $u \neq 0 \implies \|u\| \neq 0$ , nous pouvons diviser par  $\|u\|^2$  et nous trouvons alors  $\lambda = \bar{\lambda}$ , c'est-à-dire que  $\lambda \in \mathbb{R}$ . Maintenant, si  $\mu$  est une autre valeur propre de  $-L$  associée à la fonction propre  $v$ ,

$$\begin{aligned}\lambda \langle u|v \rangle &= -\langle Lu|v \rangle = -\langle u|Lv \rangle = \mu \langle u|v \rangle \\ (\lambda - \mu) \langle u|v \rangle &= 0 \implies \langle u|v \rangle = 0,\end{aligned}$$

car en effet nous avons supposé  $\mu \neq \lambda$ .

□

Il suit directement du théorème le corollaire suivant, et c'en était la motivation.

**Corollaire 1.8.** *Soient  $p$  et  $r$  des fonctions définies sur un intervalle  $I$ . Alors l'opérateur*

$$L := \frac{d}{dx} \left( wp \frac{d}{dx} \right) + r$$

*est formellement auto-adjoint. Il suit qu'il est auto-adjoint si*

$$wp(u'\bar{v} - u\bar{v}')|_a^b = 0, \quad \forall u, v \in C^2(I) \cap L^2(I).$$

**Remarque 1.9.** Étant donné un problème de Sturm-Liouville sous la forme

$$p(x)y'' + q(x)y' + \lambda = 0,$$

où maintenant  $\lambda \in \mathbb{R}$ , nous pouvons retrouver  $w(x)$  par la formule

$$w(x) = \frac{1}{p(x)} \exp \left( \int_{x_0}^x \frac{q(t)}{p(t)} dt \right).$$

## 1.4 Quelques exemples classiques

Dans cette dernière section, nous mentionnons quelques exemples classiques de ces opérateurs, auxquels nous appliquons le corollaire précédent pour fin d'illustration. Rappelons que les caractéristiques principales des polynômes classiques cités en exemple dans ce mémoire sont regroupées dans l'annexe A.

### 1.4.1 L'équation de Legendre

Souvenons nous de l'équation

$$(1 - x^2)u'' - 2xu' + n(n + 1)u = 0,$$

que nous voulons résoudre sur  $[-1, 1]$ . Elle peut être mise sous la forme équivalente

$$\frac{d}{dx} \left( (1 - x^2) \frac{du}{dx} \right) + n(n + 1)u = 0,$$

et alors le corollaire précédent nous indique que les paires  $u, v$  de fonctions propres de cette équation seront orthogonales si

$$(1 - x^2)(u'\bar{v} - u\bar{v}')|_{-1}^1 = 0, \quad \forall u, v \in C^2(I) \cap L^2(I),$$

ce qui est évidemment vrai.

### 1.4.2 L'équation de Hermite

$$u'' - 2xu' + 2nu = 0$$

à résoudre sur  $\mathbb{R}$ . Cette équation peut être mise sous la forme équivalente

$$\frac{d}{dx} \left[ e^{-x^2} \frac{du}{dx} \right] + e^{-x^2} u = 0,$$

tel que vu précédemment. Ainsi écrite, le corollaire précédent nous indique que les paires  $u, v$  de fonctions propres de cette équation seront orthogonales si

$$e^{-x^2}(u'\bar{v} - u\bar{v}')|_{-\infty}^{\infty} = 0, \quad \forall u, v \in C^2(I) \cap L^2(I),$$

ce qui est encore clairement vrai.

### 1.4.3 L'équation de Laguerre

$$xu'' + (1 - x)u' + nu = 0$$

à résoudre sur  $(0, \infty)$ . Elle peut être mise sous la forme équivalente

$$\frac{d}{dx} \left[ xe^{-x} \frac{du}{dx} \right] + ne^{-x} u = 0,$$

et donc les paires  $u, v$  de fonctions propres de cette équation seront orthogonales si

$$xe^{-x}(u'\bar{v} - u\bar{v}')|_0^{\infty} = 0, \quad \forall u, v \in C^2(I) \cap L^2(I),$$

ce qui est toujours vérifié.

Nous savons donc maintenant générer des suites de polynômes orthogonaux à partir d'opérateurs différentiels. Il apparaît qu'historiquement, ces opérateurs sont apparus avant ce qui suivra dans le prochain chapitre, et nous avons choisi de suivre cet ordre pour introduire le sujet. Nous entrerons cependant dans la suite plus sérieusement dans le vif du sujet, et nous étudierons des formules pour lesdites suites de polynômes qui sont certes plus aisées à manipuler.



## Chapitre 2

# Définitions équivalentes des polynômes orthogonaux

Dans ce chapitre, nous donnerons d'abord une première définition d'une suite de polynômes orthogonaux et ferons quelques remarques. Comme nous avons annoncé dans le premier chapitre que les solutions à des problèmes de Sturm-Liouville menaient à des polynômes orthogonaux, on montrera l'équivalence entre les deux concepts. Puis nous donnerons la formule de Rodrigues et la relation de récurrence générale que suivent les polynômes orthogonaux. Finalement, nous verrons qu'il existe des fonctions génératrices de ces suites de polynômes et en discuterons via quelques exemples.

### 2.1 Une première définition

Rappelons d'abord que nous travaillons sur un espace

$$L^2(m_F) := \{f : X \rightarrow \mathbb{R} : f \text{ mesurable, } \|f\| < \infty\}$$

de norme induite par le produit scalaire

$$\langle f|g \rangle := \int_X fg dm_F,$$

où  $m_F$  est une mesure de Lebesgue-Stieltjes et  $X$  un espace mesurable.

**Définition 2.1.** Une suite de polynômes  $(P_n)_{n \geq 0}$ , où  $P_n$  est de degré  $n$ , est dite *orthogonale* dans  $L^2(m_F)$  si  $\langle P_n|P_m \rangle = 0$  lorsque  $n \neq m$ .

Une telle suite forme évidemment une base orthogonale de l'espace des polynômes et il est bien connu qu'étant donné un produit scalaire fixé, nous pouvons en construire une en appliquant le procédé de *Gram-Schmidt* à la base canonique des polynômes, à savoir  $\{1, x, x^2, x^3, \dots\}$ . Il vaut la peine d'en rappeler les détails ici, pour l'importance du résultat et pour fixer les notations, voire les idées. Commençons par une définition.

**Définition 2.2.** La *projection d'un vecteur  $f$  sur un vecteur  $g$*  est l'opération

$$\text{proj}(f : g) = \frac{\langle f | g \rangle}{\|g\|^2} g.$$

Cela généralise la notion habituelle de projection en trois dimensions : nous recherchons la «composante de  $f$  dans la direction de  $g$ ».

**Définition 2.3.** Soit  $B := \{u_i : i = 1, 2, 3, \dots, n\}$ , une base d'un sous-espace vectoriel  $\mathbb{E} \subseteq \mathbb{F}$  (ou encore  $n$  vecteurs linéairement indépendants d'un espace  $\mathbb{F}$  de dimension au moins  $n$ ), et soit  $f \in \mathbb{F}$ . La *projection de  $f$  sur  $\mathbb{E}$*  s'écrit

$$\text{proj}(f : u_1, u_2, \dots, u_n) := \sum_{i=1}^n \text{proj}(f : u_i).$$

Autorisons-nous ici une petite digression. Cette dernière définition peut se généraliser aux espaces et sous-espaces vectoriels de taille infinie. Cela amène, outre beaucoup de questions légitimes sur la convergence, les concepts suivants.

**Définition 2.4.** Nous dirons que  $B := \{u_1, u_2, u_3, \dots\}$  est une *base d'approximation* de  $\mathbb{E}$  si chaque élément de  $\mathbb{E}$  peut être approché par une combinaison linéaire des éléments de  $B$ , c'est-à-dire si pour tout  $f$  dans  $\mathbb{E}$  et tout  $\epsilon > 0$ , il existe des scalaires  $a_1, a_2, \dots, a_n$  et des  $u_1, u_2, \dots, u_n \in B$  ( $n \in \mathbb{N}$ ) tels que

$$\left| f - \sum_{i=1}^n a_i u_i \right| < \epsilon.$$

**Définition 2.5.** Soit  $B := \{u_1, u_2, u_3, \dots\}$  une base d'approximation d'un espace  $\mathbb{E}$  et soit  $f \in \mathbb{E}$ . La série que donne

$$\text{proj}(f : u_1, u_2, \dots) := \sum_{i=1}^{\infty} \text{proj}(f : u_i)$$

se nomme la *série de Fourier généralisée de  $f$  dans la base  $B$* .

La série ainsi définie ne converge pas toujours vers  $f$ . Par contre, si nous demandons en plus que la base d'approximation considérée soit orthogonale, nous aurons la convergence au sens des moindres carrés, c'est-à-dire que

$$\|f - \text{proj}(f : u_1, u_2, \dots, u_n)\| \rightarrow 0 \quad (n \rightarrow \infty).$$

Nous étudierons cela plus en détail au chapitre suivant. Pour l'heure, rappelons un résultat important mais simple sur la projection. (Un développement complet de la méthode est fait dans (5, section 2.3 page 46).)

**Lemme 2.6.** *Étant donnés  $f$  et  $\phi$  deux vecteurs d'un espace  $\mathbb{E}$  quelconque, le vecteur  $f - \text{proj}(f : \phi)$  est orthogonal à  $\phi$ . Cette opération peut être vue comme « enlever à  $f$  sa partie linéairement dépendante de  $\phi$  ». Plus généralement, étant donnés  $f$  et  $\{u_1, u_2, \dots, u_n\}$ , des vecteurs d'un espace  $\mathbb{E}$  quelconque, le vecteur  $f - \text{proj}(f : u_1, u_2, \dots, u_n)$  est orthogonal à  $u_1, u_2, \dots, u_n$ .*

*Démonstration.* Ce n'est qu'un petit calcul qui utilise la linéarité du produit scalaire.

$$\begin{aligned} \langle f - \text{proj}(f : \phi) | \phi \rangle &= \left\langle f - \frac{\langle f | \phi \rangle}{\langle \phi | \phi \rangle} \phi | \phi \right\rangle \\ &= \langle f | \phi \rangle - \frac{\langle f | \phi \rangle}{\langle \phi | \phi \rangle} \langle \phi | \phi \rangle \\ &= \langle f | \phi \rangle - \langle f | \phi \rangle \\ &= 0. \end{aligned}$$

□

Voici maintenant le théorème où nous voulions aboutir. La méthode de Gram-Schmidt en est la démonstration.

**Théorème 2.7.** *Soient  $U := \{u_1, u_2, u_3, \dots\}$ , des vecteurs linéairement indépendants d'un espace  $\mathbb{E}$ . Il existe un ensemble de vecteurs orthogonaux  $\Phi := \{\phi_1, \phi_2, \phi_3, \dots\}$  tels que les  $k$  premiers éléments de  $\Phi$  forment une base du sous-espace engendré par les  $k$  premiers éléments de  $U$ .*

*Démonstration.* La preuve de ce résultat suit de l'application du lemme précédent pour une construction explicite de l'ensemble  $\Phi$ . Fixons d'abord  $\phi_1 := u_1$ . Posons ensuite  $\phi_2 := u_2 - \text{proj}(u_2 : \phi_1)$ . Le lemme donne donc que  $\phi_2$  est orthogonal à  $\phi_1$ . De plus,  $\phi_1$  et  $\phi_2$  sont indépendants et engendrent le même sous-espace que  $u_1$  et  $u_2$ . De même, pour chaque  $k > 2$ , nous poserons  $\phi_k := u_k - \text{proj}(u_k : \phi_1, \phi_2, \dots, \phi_{k-1})$  et, encore une fois, le lemme nous assurera que  $\phi_k$  est orthogonal à tous les vecteurs qui le précèdent dans la liste. □

Maintenant, sachant que dans chaque espace  $L^2(m_F)$ , une suite de polynômes orthogonaux existe, nous pouvons constater qu'elle est unique à un facteur scalaire près. Pour ce faire, énonçons d'abord le lemme suivant.

**Lemme 2.8.** *Soit  $P_0, P_1, P_2, \dots$  une suite de polynômes orthogonaux dans  $L^2(w)$ . Alors pour chaque  $n \geq 1$ , le polynôme  $P_n$  est orthogonal à  $1, x, x^2, \dots, x^{n-1}$ .*

*Démonstration.* Ce résultat est clair dans la mesure où chaque  $x^i$  ( $i < n$ ) peut être exprimé comme une combinaison linéaire des  $P_0, P_1, \dots, P_i$ , tous orthogonaux à  $P_n$  par hypothèse. La conclusion suit par la linéarité du produit scalaire. □

Nous en sommes au théorème (dont la démonstration présentée ici provient de (5)), presque aussi simple, mais néanmoins important. Il stipule qu'à un facteur de normalisation près, une suite de polynômes orthogonaux est unique. Pour la suite, appelons  $\text{cd}(p)$  le coefficient directeur de  $p$ .

**Théorème 2.9** (Unicité). *Soit  $P_0, P_1, P_2, \dots$  une suite de polynômes orthogonaux dans  $L^2(w)$ , et soit  $p_n$  un polynôme de degré  $n$  et orthogonal à  $1, x, x^2, \dots, x^{n-1}$  dans  $L^2(w)$ . Alors  $p_n$  est un multiple scalaire de  $P_n$ .*

*Démonstration.* Soit  $C$ , un scalaire tel que  $\text{cd}(CP_n) = \text{cd}(p_n)$ . Alors  $f_n := CP_n - p_n$  est un polynôme de degré au plus  $n - 1$ . Or, par le lemme pour  $P_n$  et par hypothèse pour  $p_n$ , il suit par linéarité que  $f_n$  est orthogonal à  $1, x, x^2, \dots, x^{n-1}$ . Le polynôme  $f_n$  est donc orthogonal à toute combinaison linéaire de ces monômes, en particulier à lui même, i.e.  $\langle f_n | f_n \rangle = 0 \implies f_n = 0 \implies p_n = CP_n$ .  $\square$

## 2.2 Cas des solutions d'un problème de Sturm-Liouville

### 2.2.1 Théorie générale des équations différentielles

Vu le titre du présent écrit, il est entendu que nous allons nous intéresser aux problèmes menant à des polynômes orthogonaux. Énonçons d'abord le théorème suivant, qui provient de (3, p. 129). Il s'y trouve aussi la démonstration du cas qui nous intéresse, soit celui du degré deux, avec  $x_0 = 0$  (3, p. 138).

**Théorème 2.10.** *Soit une équation différentielle  $L(y) = 0$ , où  $L$  est de la forme*

$$L(y) = y^{(n)} + a_1(x)y^{(n-1)} + \dots + a_n(x)y,$$

*où les  $a_i(x)$  sont des fonctions dont les développements en série convergent au voisinage d'un point  $X_0$  d'un intervalle  $|x - x_0| < r$ .*

*Alors il existe une solution générale  $\phi$  au problème  $L(y) = 0$  qui s'exprime comme une série entière  $\phi(x) = \sum_{n=0}^{\infty} a_n(x - x_0)^n$ , convergente pour  $|x - x_0| < r$ .*

Nous verrons que dans le cas qui nous concerne, la résolution des équations par la méthode des séries entières donne deux solutions linéairement indépendantes. Or, cela est suffisant, dans le sens du théorème suivant. Rappelons que le wronskien de deux solutions  $y_1$  et  $y_2$  est le déterminant donné par

$$W(t) := \begin{vmatrix} y_1(t) & y_2(t) \\ y_1'(t) & y_2'(t) \end{vmatrix}.$$

**Théorème 2.11.** *Si  $y_1$  et  $y_2$  sont deux solutions de*

$$y'' + p(t)y' + q(t)y = 0 \quad t \in I$$

et s'il existe  $t_0 \in I$  tel que le wronskien  $W(t_0) \neq 0$ , alors l'ensemble de solutions

$$\{y = c_1 y_1(t) + c_2 y_2(t) : c_1, c_2 \in \mathbb{R}\}$$

contient toutes les solutions de l'équation.

Maintenant, pour résoudre des problèmes de Sturm-Liouville dans les applications, il faut habituellement que la solution converge non seulement sur l'intervalle ouvert  $I$ , mais sur sa fermeture. Pour cela, la série devra se terminer (c'est-à-dire que les termes devront être tous nuls à partir d'un certain rang), et nous aurons bien affaire alors à des polynômes. Voyons quelques exemples.

### 2.2.2 Exemples de solutions

Le cas le plus simple est bien sûr celui de l'équation de Legendre

$$y'' - \frac{2x}{1-x^2}y' + \frac{n(n+1)}{(1-x^2)}y = 0.$$

En faisant les calculs (voir (11)), nous trouvons la solution générale

$$\phi(x) = \sum_{n=0}^{\infty} a_n x^n = \sum_{k=0}^{\infty} a_{2k} x^{2k} + \sum_{k=0}^{\infty} a_{2k+1} x^{2k+1},$$

où, pour chaque  $k \geq 1$  entier,

$$\begin{cases} a_{2k} = (-1)^k \frac{(n-0)(n-2)\dots(n-(2k-2))(n+1)(n+3)\dots(n+(2k-1))}{(2k)!} a_0 \\ a_{2k+1} = (-1)^k \frac{(n-1)(n-3)\dots(n-(2k-1))(n+2)(n+4)\dots(n+(2k-2))}{(2k+1)!} a_1. \end{cases}$$

Tous les coefficients sont alors déterminés en fonction des deux premiers, soit  $a_0$  et  $a_1$ , qui sont donc des constantes arbitraires. Par ailleurs, en développant les coefficients de l'équation de départ en séries entières, nous obtenons

$$\frac{-2x}{1-x^2} = \sum_{k=0}^{\infty} (-2)x^{2k+1}$$

et

$$\frac{n(n+1)}{1-x^2} = n(n+1) \sum_{k=0}^{\infty} x^{2k}.$$

Ces deux séries ont un rayon de convergence de 1, donc en vertu du théorème ci-haut, notre solution générale converge pour  $|x| < 1$ . Maintenant nous voulons avoir convergence sur l'intervalle fermé. Or, en prenant un paramètre  $n$  entier, une des deux séries est finie. En mettant

le coefficient de celle-ci (i.e.  $a_0$  ou  $a_1$  selon le cas) égal à 1 et celui de l'autre série égal à zéro, la solution est donc polynômiale, ce qui satisfait notre désir.

Tel que vu dans le premier chapitre, les paires de solutions d'une telle équation sont orthogonales. La suite de polynômes ainsi obtenue est donc orthogonale pour la mesure de Lebesgue. En voici les premiers spécimens, normalisés pour que  $P_n(1) = 1$  pour chaque  $n$ .

$$\begin{aligned} P_0(x) &= 1, \\ P_1(x) &= x, \\ P_2(x) &= \frac{3}{2}x^2 - \frac{1}{2}, \\ P_3(x) &= \frac{5}{2}x^3 - \frac{3}{2}x. \end{aligned}$$

Un développement semblable peut être fait pour l'équation de Hermite,

$$y'' - 2xy' + 2ny = 0.$$

En faisant les calculs, nous trouvons les coefficients

$$\begin{cases} a_{2k} = (-1)^k 2^k \frac{n(n-2)\dots(n-(2k-2))}{(2k)!} a_0, \\ a_{2k+1} = (-1)^k 2^k \frac{(n-1)(n-3)\dots(n-(2k-1))}{(2k+1)!} a_1. \end{cases}$$

Ici, toutes les séries convergent en principe dans  $\mathbb{R}$ , mais le même manège que pour l'équation précédente peut quand-même être répété. Nous obtenons une autre suite de polynômes orthogonaux, cette fois avec  $dm_F = e^{-x^2} dm$  dont les premiers membres, normalisés pour avoir des coefficients entiers (ce qui donne un coefficient directeur valant  $2^n$  pour chaque  $H_n$ ), sont

$$\begin{aligned} H_0(x) &= 1, \\ H_1(x) &= 2x, \\ H_2(x) &= 4x^2 - 2, \\ H_3(x) &= 8x^3 - 12x. \end{aligned}$$

## 2.3 Formule de Rodrigus

### 2.3.1 Cas général

Pour les polynômes orthogonaux provenant des problèmes de Sturm-Liouville, la suite peut en fait être calculée sans résoudre le problème au long comme dans la section précédente. Il suffit d'utiliser le théorème suivant.

**Théorème 2.12.** *Considérons un problème de Sturm-Liouville*

$$p(x)y'' + q(x)y' + \lambda = 0,$$

où il est supposé que  $\deg(p(x)) \leq 2$ ,  $\deg(q(x)) \leq 1$ , que  $\lambda$  est une constante et que l'opérateur  $L := p \frac{d^2}{dx^2} + q \frac{d}{dx} + \lambda$  est auto-adjoint. Appelons  $w(x) = \frac{1}{p(x)} \exp\left(\int_{x_0}^x \frac{q(t)}{p(t)} dt\right)$  (voir remarque 1.9). Alors la suite des polynômes orthogonaux qui solutionnent ce problème est donnée par la formule de Rodrigues

$$P_n(x) = \frac{1}{w(x)} \frac{d^n}{dx^n} \left( w(x) [p(x)]^n \right),$$

à un facteur de normalisation près.

*Démonstration.*

(i) Posons  $v(x) := w(x)[p(x)]^n$  et montrons que  $\frac{1}{w(x)} \frac{d^n}{dx^n} v(x)$  solutionne le problème de Sturm-Liouville de départ. Calculons d'abord

$$\begin{aligned} w' &= \left( -\frac{p'}{p^2} + \frac{1}{p} \cdot \frac{q}{p} \right) \exp\left(\int_{x_0}^x \frac{q(t)}{p(t)} dt\right) \\ &= \frac{q - p'}{p} w. \end{aligned}$$

Alors d'une part

$$\begin{aligned} \left( \frac{v^{(n)}}{w} \right)' &= \frac{v^{(n+1)} w - v^{(n)} w'}{w^2} = \frac{v^{(n+1)} - v^{(n)} \left( \frac{q-p'}{p} \right)}{w} \\ &= \frac{v^{(n+1)}}{w} + \frac{(p' - q)v^{(n)}}{wp}, \end{aligned}$$

et puis

$$\begin{aligned} \left( \frac{v^{(n)}}{w} \right)'' &= \left( \frac{v^{(n+1)}}{w} \right)' + \left( \frac{(p' - q)v^{(n)}}{wp} \right)' \\ &= \frac{v^{(n+2)}}{w} + \frac{(p' - q)v^{(n+1)}}{wp} + \frac{[(p' - q)v^{(n)}]' wp - (p' - q)v^{(n)} wq}{w^2 p^2} \\ &= \frac{v^{(n+2)}}{w} + \frac{(p' - q)v^{(n+1)}}{wp} + \frac{(p'' - q')v^{(n)} + (p' - q)v^{(n+1)}}{wp} - \frac{(p' - q)v^{(n)} q}{wp^2} \\ &= \frac{v^{(n+2)}}{w} + \frac{2(p' - q)v^{(n+1)}}{wp} + \left[ \frac{(p'' - q')}{wp} - \frac{(p' - q)q}{wp^2} \right] v^{(n)}. \end{aligned}$$

Donc

$$\begin{aligned} p \left[ \frac{v^{(n)}}{w} \right]'' + q \left[ \frac{v^{(n)}}{w} \right]' + \lambda \left[ \frac{v^{(n)}}{w} \right] &= \\ \frac{p}{w} v^{(n+2)} + \frac{2(p' - q)}{w} v^{(n+1)} + \left[ \frac{(p'' - q')}{w} - \frac{(p' - q)q}{wp} \right] v^{(n)} &+ \\ + \frac{q}{w} v^{(n+1)} + \frac{q(p' - q)}{wp} v^{(n)} + \frac{\lambda}{w} v^{(n)}. & \end{aligned}$$

En simplifiant, cela donne

$$p \left[ \frac{v^{(n)}}{w} \right]'' + q \left[ \frac{v^{(n)}}{w} \right]' + \lambda \left[ \frac{v^{(n)}}{w} \right] = \frac{p}{w} v^{(n+2)} + \frac{2p' - q}{w} v^{(n+1)} + \left[ \frac{(p'' - q')}{w} + \frac{\lambda}{w} \right] v^{(n)}. \quad (2.1)$$

D'autre part,

$$\begin{aligned} pv' &= p(wp^n)' \\ &= p(w'p^n + wnp^{n-1}p') \\ &= w'p^{n+1} + nwp^n p' \\ &= \left( \frac{q - p'}{p} \right) wp^{n+1} + nwp^n p' \\ &= (q - p')wp^n + nwp^n p' \\ &= (q - p' + np')wp^n \\ &= (q - p' + np')v \\ &= (q + (n - 1)p')v. \end{aligned}$$

En dérivant  $n + 1$  fois l'égalité précédente avec la règle du produit, nous obtenons

$$pv^{(n+2)} + (n+1)p'v^{(n+1)} + \frac{n(n+1)}{2}p''v^{(n)} = [q + (n-1)p']v^{(n+1)} + (n+1)[q' + (n-1)p'']v^{(n)}.$$

En ramenant tout du même côté et en réorganisant, cela donne

$$\begin{aligned} 0 &= pv^{(n+2)} + [(n+1)p' - (q + (n-1)p')]v^{(n+1)} + \left[ \frac{n(n+1)}{2}p'' - (n+1)(q' + (n-1)p'') \right] v^{(n)} \\ &= pv^{(n+2)} + [2p' - q]v^{(n+1)} + \left[ \left( \frac{n(n+1)}{2} - (n+1)(n-1) \right) p'' - (n+1)q' \right] v^{(n)} \\ &= pv^{(n+2)} + [2p' - q]v^{(n+1)} + \left[ p'' - q' + \left( \frac{n - n^2}{2} \right) p'' - nq' \right] v^{(n)}. \end{aligned}$$

En comparant cela avec 2.1, nous trouvons que la fonction  $\frac{1}{w(x)} \frac{d^n}{dx^n} v(x)$  est bien une solution au problème de départ, les valeurs propres de l'équation étant données par

$$\lambda = \left( \frac{n - n^2}{2} \right) p'' - nq'.$$

- (ii) Nous avons montré que  $P_n = \frac{1}{w(x)} \frac{d^n}{dx^n} (w(x)[p(x)]^n)$  solutionne le problème de Sturm-Liouville associé pour chaque  $n$  fixé, et donc qu'il s'agit bien d'une suite de fonctions propres orthogonales dans  $L^2(w)$ . Il reste à montrer que ce sont bien des polynômes.

Cela se déduit du résultat un peu plus général suivant. Soit  $r = r(x)$  un polynôme quelconque. Alors

$$\frac{d^n}{dx^n} rwp^n = ws,$$

où  $s = s(x)$  est un polynôme. En effet, si  $n = 1$ ,

$$\begin{aligned} \frac{d}{dx} rwp &= r'wp + r(wp)' \\ &= r'wp + rwq \\ &= w[r'p + rq] \\ &= ws, \end{aligned}$$

où  $s(x) = r'p + rq$ . Supposons maintenant, pour une récurrence, que le résultat soit vrai pour un certain  $n$  fixé. Alors

$$\begin{aligned} \frac{d^{n+1}}{dx^{n+1}} rwp^{n+1} &= \frac{d^n}{dx^n} [r'wp^{n+1} + r(wp)'p^n + rwp(p^n)'] \\ &= \frac{d^n}{dx^n} [r'wp^{n+1} + rwqp^n + rwpnp^{n-1}p'] \\ &= \frac{d^n}{dx^n} wp^n [r'p + rq + rnp'] \\ &\stackrel{H.R.}{=} ws \end{aligned}$$

pour un certain polynôme  $s = s(x)$ , puisque  $r_0(x) := r'p + rq + rnp'$  est un polynôme.

□

### 2.3.2 Deux cas particuliers

Considérons l'équation de Legendre que voici

$$(1 - x^2)u'' - 2xu' + n(n + 1)u = 0.$$

La formule générale de Rodrigues nous donne que les polynômes de Legendre peuvent être mis sous la forme

$$P_n(x) = \frac{d^n}{dx^n} [(1 - x^2)^n].$$

Cependant, pour avoir la normalisation  $P_n(1) = 1$ , écrivons plutôt

$$P_n(x) = \frac{1}{2^n n!} \frac{d^n}{dx^n} [(x^2 - 1)^n].$$

Introduisons maintenant l'équation de Gegenbauer, une généralisation de celle de Legendre.

$$(1 - x^2)y'' - (2\alpha + 1)xy' + n(n + 2\alpha)y = 0.$$

Nous pouvons calculer que dans ce cas-ci,  $w(x) = (1 - x^2)^{\alpha - \frac{1}{2}}$ , et alors nous obtenons la formule de Rodrigues des polynômes de Gegenbauer (non normalisée)

$$C_n^{(\alpha)} = (1 - x^2)^{\frac{1}{2} - \alpha} \frac{d^n}{dx^n} [(1 - x^2)^{n + \alpha - \frac{1}{2}}].$$

## 2.4 Relation de récurrence

### 2.4.1 Résultat général

Malgré que les méthodes précédentes pour obtenir une suite de polynômes orthogonaux semblent laborieuses, il est intéressant de savoir que toute suite de polynômes orthogonaux peut s'obtenir à l'aide d'une relation de récurrence, particulièrement lorsqu'on veut faire du calcul numérique. C'est le contenu du théorème suivant.

**Théorème 2.13.** *Soit  $\mathbb{O} := \{P_n \mid n \geq 0\}$  une suite de polynômes orthogonaux. Alors il existe une relation de récurrence selon laquelle nous pouvons calculer  $P_{n+1}$  en fonction de  $P_n$  et  $P_{n-1}$ , donnée par la formule*

$$P_{n+1}(x) = (A_n x + B_n)P_n(x) - C_n P_{n-1}(x),$$

où les constantes  $A_n, B_n, C_n$  sont

$$A_n := \frac{cd(P_{n+1})}{cd(P_n)}, \quad B_n := -A_n \frac{\langle xP_n | P_n \rangle}{\langle P_n | P_n \rangle}, \quad C_n := A_n \frac{\langle xP_n | P_{n-1} \rangle}{\langle P_{n-1} | P_{n-1} \rangle}$$

et où  $cd(P_n)$  note le coefficient directeur de  $P_n$ .

*Démonstration.* Considérons le polynôme  $P_{n+1}(x) - A_n x P_n(x)$ . De la définition de la constante  $A_n$ , nous voyons assez bien que les termes en  $x^{n+1}$  s'annulent, et donc que ce polynôme est de degré au plus  $n$ . Il peut alors être exprimé comme une combinaison linéaire des  $n$  premiers éléments de  $\mathbb{O}$ . Cela s'écrit

$$P_{n+1}(x) - A_n x P_n(x) = a_n P_n(x) + a_{n-1} P_{n-1}(x) + \dots + a_0 P_0(x).$$

La valeur de chaque  $a_j$  peut être trouvée en prenant le produit scalaire de cette expression de chaque côté avec le  $P_j$  correspondant. À gauche, nous obtenons, pour chaque  $P_j$  avec  $0 \leq j < n-1$ ,

$$\begin{aligned} \langle P_{n+1}(x) - A_n x P_n(x) | P_j(x) \rangle &= \langle P_{n+1}(x) | P_j(x) \rangle - A_n \langle x P_n(x) | P_j(x) \rangle \\ &= \langle P_{n+1}(x) | P_j(x) \rangle - A_n \langle P_n(x) | x P_j(x) \rangle \\ &= 0, \end{aligned}$$

résultat obtenu grâce au fait que les degrés de  $P_j$  et de  $xP_j$  sont strictement inférieurs au degré de  $P_n$  et aussi, à plus forte raison, au degré de  $P_{n+1}$ . Pour les  $j$  correspondants, nous obtenons à droite :  $a_j \langle P_j(x) | P_j(x) \rangle$ , avec bien sûr,  $\langle P_j(x) | P_j(x) \rangle \neq 0$ . En égalisant les deux côtés, il vient donc que chacun de ces  $a_j$  doit être nul.

Pour  $j = n-1$ , nous obtenons de même, à gauche :

$$\begin{aligned} \langle P_{n+1}(x) - A_n x P_n(x) | P_{n-1}(x) \rangle &= \langle P_{n+1}(x) | P_{n-1}(x) \rangle - A_n \langle x P_n(x) | P_{n-1}(x) \rangle \\ &= -A_n \langle x P_n(x) | P_{n-1}(x) \rangle \end{aligned}$$

et à droite :  $a_{n-1}\langle P_{n-1}|P_{n-1}\rangle$ , d'où, en égalisant

$$a_{n-1} = -A_n \frac{\langle xP_n|P_{n-1}\rangle}{\langle P_{n-1}|P_{n-1}\rangle} =: -C_n.$$

Pour  $j = n$ , nous obtiendrons, par le même procédé

$$a_n = -A_n \frac{\langle xP_n|P_n\rangle}{\langle P_n|P_n\rangle} =: B_n.$$

La formule est donc bien démontrée. □

Fait à remarquer, il est sous-entendu dans ce théorème que le coefficient directeur de  $P_{n+1}$  est connu *avant* de connaître  $P_{n+1}$ . Il s'agit en fait d'un faux problème, car  $A_n$  ne dépend en réalité que de la normalisation imposée à la suite de polynômes orthogonaux. Ainsi, en utilisant  $A_n = 1$  dans la récurrence, la suite obtenue est tout à fait valable, mais on gère habituellement l'apparence des coefficients des polynômes en modifiant cette valeur.

## 2.4.2 Exemples

**Tchebychev :** Les polynômes de Tchebychev (de première espèce), sont ceux satisfaisant l'équation

$$(1 - x^2)T_n''(x) - xT_n'(x) + n^2T_n(x) = 0.$$

Ils satisfont la relation de récurrence

$$\begin{aligned} P_0 &= 1, \\ P_1 &= x, \\ P_{n+1}(x) &= 2xP_n(x) - P_{n-1}(x). \end{aligned}$$

Ici, en choisissant  $A_n = 2$ , nous avons obtenu  $B_n = 0$ ,  $C_n = 1$ .

**Legendre :** Pour les polynômes de Legendre, nous choisissons plutôt  $A_n = \frac{2n+1}{n+1}$ , ce qui donne  $B_n = 0$  et  $C_n = \frac{n}{n+1}$ , d'où la formule de récurrence

$$\begin{aligned} P_0 &= 1, \\ P_1 &= x, \\ P_{n+1}(x) &= \frac{2n+1}{n+1}xP_n(x) - \frac{n}{n+1}P_{n-1}(x). \end{aligned}$$

Cette dernière équation se rencontre aussi sous l'appellation *Formule de récurrence de Bonnet*.

## 2.5 Séries génératrices

### 2.5.1 Généralités

Le domaine des séries génératrices en général est vaste et complexe. Certains utilisent encore la terminologie « fonctions génératrices » pour les décrire, mais l'appellation « séries » nous a

semblé préférable, car il s'agit de séries formelles pour lesquelles on ne se soucie pas toujours de convergence. Il en existe plusieurs types. En théorie, toute suite a une série génératrice de chaque type. Cependant, on rencontre en pratique seulement celles qui ont une forme utilisable pour des problèmes spécifiques ou encore celles dont la forme fermée est relativement simple. Pour les suites de polynômes orthogonaux  $P_n(x)$ , on rencontre surtout des séries génératrices ordinaires, exponentielles et de Dirichlet. Par exemple, pour les polynômes de Tchebychev (toujours de première espèce), voici les séries (et leur forme fermée correspondante)

(i) Ordinaire

$$\sum_{n=0}^{\infty} T_n(x)t^n = \frac{1-tx}{1-2tx+t^2};$$

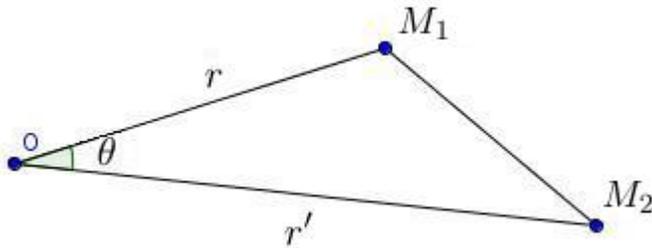
(ii) Exponentielle

$$\sum_{n=0}^{\infty} T_n(x) \frac{t^n}{n!} = \frac{1}{2} e^x \cosh(t\sqrt{x^2-1});$$

(iii) De Dirichlet

$$\sum_{n=0}^{\infty} T_n(x) \frac{t^n}{n} = -\frac{1}{2} \ln(1-2tx+t^2).$$

Il serait facile de croire que ce ne sont que des artifices mathématiques servant à démontrer plus joliment certains théorèmes, mais en fait, pour certains polynômes, la série génératrice est apparue *en premier*. Considérons le problème suivant à titre d'exemple (voir (10, section 4)). Étant données deux masses  $M_1$  et  $M_2$  situées à des distances  $r$  et  $r'$  respectivement de l'origine  $O$ , avec  $r < r'$ , nous voulons calculer le potentiel newtonien entre les deux masses.



La loi des cosinus donne d'abord, pour la distance entre  $M_1$  et  $M_2$ ,

$$\overline{M_1M_2} = \sqrt{r^2 - 2rr' \cos \theta + r'^2} = r'(h^2 - 2h \cos \theta + 1)^{1/2},$$

où  $h := \frac{r}{r'} < 1$ .

Le potentiel est donc donné par

$$U(h) = \frac{G M_1 M_2}{r'(h^2 - 2h \cos \theta + 1)^{1/2}},$$

où  $G$  est la constante universelle de la gravitation. Posons  $x = \cos \theta$  et considérons l'expression

$$\Phi(x, h) = \frac{1}{\sqrt{1 - 2xh + h^2}} \quad |h| < 1.$$

Il s'agit d'une série génératrice des polynômes de Legendre. En effet, en la développant en série de puissance en  $h$ , nous obtenons lesdits polynômes comme coefficients. Voici une première manière de le montrer.

**Théorème 2.14.** *Développons l'expression de  $\Phi(x, h)$  comme une série entière par rapport à  $h$ . Cela peut s'écrire*

$$\Phi(x, h) = P_0(x) + hP_1(x) + h^2P_2(x) + \dots = \sum_{l=0}^{\infty} h^l P_l(x),$$

et les  $P_l(x)$  sont les polynômes de Legendre.

*Démonstration.* Observons d'abord que cette série entière en  $h$  (où  $x \leq 1$  est considéré comme un paramètre) converge pour  $|h| < 1$  (Ce qui était d'ailleurs l'hypothèse du problème précédent). Maintenant, vérifions le cas où  $x = 1$ .

$$\begin{aligned} \Phi(1, h) &= (1 - 2h + h^2)^{-1/2} = \frac{1}{1 - h} \\ &= 1 + h + h^2 + \dots \\ &= P_0(1) + P_1(1)h + P_2(1)h^2 + \dots \end{aligned}$$

Il vient donc bien que  $P_l(1) = 1$  pour chaque  $l \in \mathbb{N}$ . En différentiant la définition de  $\Phi(x, h)$  et en remplaçant, nous obtenons facilement que

$$(1 - x^2) \frac{\partial^2 \Phi}{\partial x^2} - 2x \frac{\partial \Phi}{\partial x} + h \frac{\partial^2}{\partial h^2} (h\Phi) = 0.$$

En substituant par la série entière en  $h$  de tout à l'heure, cette équation devient

$$(1 - x^2) \sum_{l=0}^{\infty} h^l P_l''(x) - \sum_{l=0}^{\infty} h^l P_l'(x) + \sum_{l=0}^{\infty} l(l+1)h^l P_l(x) = 0.$$

Cette égalité forçant tous les coefficients des  $h^l$  à être nuls, il vient pour chaque  $l$

$$(1 - x^2)P_l''(x) - 2xP_l'(x) + l(l+1)P_l(x) = 0,$$

c'est-à-dire que tous les  $P_l$  ainsi définis solutionnent l'équation de Legendre. Par l'unicité d'un tel phénomène traitée dans la section 2.2, ces polynômes sont donc bien ceux de Legendre.  $\square$

### 2.5.2 Calcul de séries génératrices

En général, il n'y a pas de manière simple de trouver une série génératrice à partir d'une suite donnée de polynômes. Cependant, étant donné le résultat, certaines démonstrations sont relativement simples (voir à ce sujet (17)). Une des difficultés provient de la forme des  $A_n$ ,  $B_n$  et  $C_n$  de la relation de récurrence. Par contre, lorsqu'ils sont constants, comme dans le cas des polynômes de Tchebychev dont il était question tout à l'heure, cela devient très simple. En effet, considérons la relation de récurrence que suivent lesdits polynômes, à savoir

$$\begin{aligned} T_0 &= 1, \\ T_1 &= x, \\ T_n(x) &= 2xT_{n-1}(x) - T_{n-2}(x). \end{aligned}$$

Multiplions par  $t^n$ , sommons sur les  $n$  et brassons un peu. Nous obtenons

$$\begin{aligned} \sum_{n=2}^{\infty} T_n(x)t^n &= 2x \sum_{n=2}^{\infty} T_{n-1}(x)t^n - \sum_{n=2}^{\infty} T_{n-2}(x)t^n \\ &= 2xt \sum_{n=2}^{\infty} T_{n-1}(x)t^{n-1} - t^2 \sum_{n=2}^{\infty} T_{n-2}(x)t^{n-2}. \end{aligned}$$

Alors

$$\begin{aligned} \sum_{n=0}^{\infty} T_n(x)t^n - xt - 1 &= 2xt \sum_{n=1}^{\infty} T_n(x)t^n - t^2 \sum_{n=0}^{\infty} T_n(x)t^n \\ &= 2xt \sum_{n=0}^{\infty} T_n(x)t^n - 2xt - t^2 \sum_{n=0}^{\infty} T_n(x)t^n, \end{aligned}$$

donc

$$1 + xt - 2xt = (1 - 2xt + t^2) \sum_{n=0}^{\infty} T_n(x)t^n,$$

c'est-à-dire

$$\sum_{n=0}^{\infty} T_n(x)t^n = \frac{1 - tx}{1 - 2tx + t^2}.$$

La série génératrice ordinaire donnée précédemment a donc été facilement retrouvée. Voyons maintenant un cas un peu plus complexe. Partons de la relation de récurrence décrivant les polynômes de Legendre, à savoir

$$nP_n(x) = (2n - 1)xP_{n-1}(x) - (n - 1)P_{n-2}(x).$$

Multiplions par  $t^{n-1}$  et sommons sur les  $n$ .

$$\sum_{n=1}^{\infty} nP_n(x)t^{n-1} = \sum_{n=1}^{\infty} (2n - 1)xP_{n-1}(x)t^{n-1} - \sum_{n=1}^{\infty} (n - 1)P_{n-2}(x)t^{n-1}.$$

Posons

$$h(t) := \sum_{n=0}^{\infty} P_n t^n \implies h'(t) := \sum_{n=0}^{\infty} n P_n t^{n-1}.$$

Alors

$$\begin{aligned} h'(t) &= \sum_{n=0}^{\infty} (2n+1)xP_n(x)t^n - \sum_{n=0}^{\infty} (n+1)P_n(x)t^{n+1} \\ &= 2xt \sum_{n=0}^{\infty} nP_n(x)t^{n-1} + x \sum_{n=0}^{\infty} P_n(x)t^n - \sum_{n=0}^{\infty} nP_n(x)t^{n+1} - \sum_{n=0}^{\infty} P_n(x)t^{n+1} \\ &= 2xth'(t) + xh(t) - t^2h'(t) - th(t) \end{aligned}$$

et donc

$$\begin{aligned} (x-t)h(t) &= (1-2xt+t^2)h'(t) \\ \frac{h'(t)}{h(t)} &= \frac{x-t}{1-2xt+t^2}. \end{aligned}$$

En intégrant, nous obtenons

$$\ln(h(t)) = -\frac{1}{2} \ln(1-2xt+t^2),$$

d'où

$$h(t) = (1-2xt+t^2)^{-\frac{1}{2}},$$

formule que nous connaissons maintenant très bien. La méthode précédente se généralise entre autres aux polynômes de Gegenbauer (voir (16, p.82)).

### 2.5.3 Utilisation des séries génératrices

Outre le fait qu'elles contiennent de manière concise beaucoup d'information et qu'elles adviennent telles quelles dans des applications, une autre raison d'utiliser des séries génératrices est d'obtenir des représentations dites « explicites » de polynômes orthogonaux. En voici deux exemples.

Dans le cas des polynômes de Hermite, nous pouvons écrire

$$\begin{aligned}
\sum_{n=0}^{\infty} \frac{H_n(x)t^n}{n!} &= e^{2xt-t^2} \\
&= \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} \frac{(2xt)^k}{k!} \cdot \frac{(-t^2)^j}{j!} \\
&= \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} \frac{(-1)^j (2x)^k t^{2j+k}}{j!k!} \\
&= \sum_{n=0}^{\infty} \sum_{2j+k=n} \frac{(-1)^j (2x)^k t^n}{j!k!} \\
&= \sum_{n=0}^{\infty} \left( \sum_{j=0}^{\lfloor n/2 \rfloor} \frac{(-1)^j (2x)^{n-2j}}{j!(n-2j)!} t^n \right).
\end{aligned}$$

Nous avons trouvé

$$H_n = n! \sum_{j=0}^{\lfloor n/2 \rfloor} \frac{(-1)^j (2x)^{n-2j}}{j!(n-2j)!}.$$

Pour ceux de Laguerre (généralisés), nous avons de même

$$\begin{aligned}
\sum_{n=0}^{\infty} L_n^{(\alpha)}(x)w^n &= \frac{1}{(1-w)^{\alpha+1}} \exp\left(-\frac{xw}{1-w}\right) \\
&= \frac{1}{(1-w)^{\alpha+1}} \sum_{k=0}^{\infty} \frac{1}{k!} \left(-\frac{xw}{1-w}\right)^k \\
&= \sum_{k=0}^{\infty} \frac{1}{k!} \left(\frac{(-xw)^k}{(1-w)^{k+\alpha+1}}\right) \\
&= \sum_{k=0}^{\infty} \left(\frac{(-xw)^k}{k!} \sum_{j=0}^{\infty} \frac{(k+\alpha+j)!}{(k+\alpha)!} \cdot \frac{w^j}{j!}\right) \\
&= \sum_{k=0}^{\infty} \sum_{j=0}^{\infty} \frac{(-xw)^k}{k!} \binom{n+\alpha}{j} w^j \\
&= \sum_{n=0}^{\infty} \left(\sum_{j+k=n} \frac{(-x)^k}{k!} \binom{n+\alpha}{j}\right) w^n \\
&= \sum_{n=0}^{\infty} \left(\sum_{k=0}^n \frac{(-x)^k}{k!} \binom{n+\alpha}{n-k}\right) w^n.
\end{aligned}$$

Nous avons maintenant trouvé

$$L_n^{(\alpha)}(x) = \sum_{k=0}^n \frac{(-x)^k}{k!} \binom{n+\alpha}{n-k}.$$

## 2.6 Polynômes orthogonaux moins connus

Au début de ce chapitre, nous avons introduit une définition très générale d'une suite de polynômes orthogonaux. Pourtant, nous ne nous sommes concentrés jusqu'à maintenant que sur celles s'obtenant à partir d'un problème de Sturm-Liouville. C'est pour justifier cette vision plus générale que nous jetterons un coup d'oeil ici à l'exemple des polynômes de Krawtchouk. Ces polynômes ont été introduits en 1929 par le mathématicien ukrainien Mikail Krawtchouk comme une généralisation de ceux de Hermite (voir (9)). Les applications ont suivi. À propos de ce qui précède, remarquons que non seulement les polynômes de Krawtchouk sont une généralisation de ceux de Hermite, mais qu'en fait il ne s'agit que d'une classe particulière de polynômes dans le *schéma d'Askey* (voir l'annexe C et (8)). Ce schéma donne les liens entre les polynômes dits, dans le cadre de cette classification, *classiques*. Toujours dans cette classification, les polynômes issus d'un problème de Sturm-Liouville sont quant à eux dits *très classiques*. C'est pour éviter d'entrer dans ce débat de nomenclature que nous avons titré cette section comme tel.

### 2.6.1 Origine et théorie

Dans l'article de M. Krawtchouk de 1929 (9), le produit scalaire suivant est d'abord considéré.

$$\langle f|g \rangle = \sum_{i=0}^N p_i f(x_i)g(x_i),$$

où nous convenons que  $x_{i+1} - x_i = 1$ ,  $p_i > 0$  et  $\sum_{i=0}^N p_i = 1$ . Cela revient simplement à utiliser notre produit scalaire habituel, mais avec une mesure de Lebesgues-Stieltjes engendrée par une fonction escalier. Pour écrire la suite, rappelons le symbolisme de la fonction de Heaviside

$$H(x) = \begin{cases} 1 & \text{si } x \geq 0, \\ 0 & \text{si } x < 0. \end{cases}$$

En utilisant la fonction escalier

$$F_N(x) = \sum_{i=0}^N p_i H(x - x_i),$$

nous obtenons le produit scalaire

$$\langle f|g \rangle = \int_{\mathbb{R}} fg dF = \sum_{i=0}^N p_i f(x_i)g(x_i). \quad (2.2)$$

L'égalité (2.2) est naturelle si  $F$  est vue comme une distribution. Cependant, pour plus de cohérence avec notre propos, donnons-en une preuve en théorie de la mesure, à travers le théorème suivant.

**Théorème 2.15.** Soit  $(p_i)_{i \in \mathbb{N}}$  une suite de réels positifs. Soit  $x_0 \in \mathbb{R}$  et posons  $x_{i+1} = x_i + 1$ . Posons enfin

$$F_N(x) := \sum_{i=0}^N p_i H(x - x_i).$$

Alors étant donnée une fonction  $g \in L^2(m_{F_N})$ , nous avons l'égalité

$$\int g dF_N = \sum_{i=0}^N p_i g(x_i).$$

*Démonstration.* Considérons une fonction étagée, sous forme canonique

$$g := \sum_{k=0}^m c_k \mathbb{1}_{A_k}.$$

Nous montrons le résultat par induction sur  $N$ .

- Si  $N = 0$ , nous pouvons écrire  $F_0 = p_0 H(x - x_0)$ , et alors

$$\int g dF_0 := \sum_{k=0}^m c_k m_F(A_k)$$

S'il existe  $M$  et  $k_0$  tels que  $A_{k_0} \subseteq (-M, x_0)$ , alors  $0 \leq m_F(A_{k_0}) \leq m_F((-\infty, x_0)) = F(x_0^-) - F(-M) = 0$ .

S'il existe  $M$  tel que  $A_{k_0} \subseteq (x_0, M)$ , alors  $0 \leq m_F(A_{k_0}) \leq m_F((x_0, M)) = F(M) - F(x_0) = 0$ . Il y a donc un seul élément non nul dans la somme, à savoir celui pour lequel  $x_0 \in A_{k_1}$ . Nous avons donc

$$\int g dF_0 = c_{k_1} m_F(A_{k_1}) = g(x_0)(F(x_0) - F(x_0^-)) = g(x_0)p_0.$$

- Supposons le résultat vrai jusqu'à un certain  $N$  fixé. Considérons

$$F_{N+1} := \sum_{i=0}^{N+1} p_i H(x - x_i).$$

Posons

$$g = \sum_{k=0}^m c_k \mathbb{1}_{A_k \cap (-\infty, x_{N+1})} + \sum_{k=0}^m c_k \mathbb{1}_{A_k \cap [x_{N+1}, \infty)}.$$

Alors

$$\int g dF_{N+1} = \sum_{k=0}^m c_k m_{F_{N+1}}(A_k \cap (-\infty, x_{N+1})) + \sum_{k=0}^m c_k m_{F_{N+1}}(A_k \cap [x_{N+1}, \infty)).$$

Pour la somme de gauche, comme  $F_N = F_{N+1}$  sur  $(-\infty, x_{N+1})$  et que  $m_{F_N}([x_{N+1}, \infty)) = 0$ , il vient  $m_{F_{N+1}}(A_k \cap (-\infty, x_{N+1})) = m_{F_N}(A_k \cap (-\infty, x_{N+1})) = m_{F_N}(A_k)$  et nous pourrions appliquer l'hypothèse de récurrence. Pour la somme de droite, un raisonnement semblable à celui de l'étape de base s'applique, et nous trouvons que le seul  $A_k \cap [x_{N+1}, \infty)$  de mesure non nulle est celui contenant  $x_{N+1}$ , disons  $A_{k_0}$ . Nous écrivons donc

$$\begin{aligned} \int g dF_{N+1} &= \sum_{k=0}^m c_k m_{F_N}(A_k) + c_{k_0} m_{F_{N+1}}(A_{k_0} \cap [x_{N+1}, \infty)) \\ &= \int g dF_N + c_{k_0} (F_{N+1}(x_{N+1}) - F_{N+1}(x_{N+1}^-)) \\ &\stackrel{H.R.}{=} \sum_{i=0}^N p_i g(x_i) + p_{N+1} g(x_{N+1}) \\ &= \sum_{i=0}^{N+1} p_i g(x_i). \end{aligned}$$

Maintenant, ce résultat peut sans peine être appliqué aux fonctions positives avec le théorème de convergence monotone, puis aux fonctions intégrables en séparant leurs parties positive et négative.  $\square$

Le produit scalaire en question est donc bien du type auquel nous sommes habitués. Maintenant, rappelons que pour un  $N$  fixé, nous considérons des  $p_i > 0$ ,  $i = 0, 1, \dots, N$  tels que  $\sum p_i = 1$ . Tchebychev a considéré le cas  $p_i = p_j$ ,  $\forall i, j$ . Cela a donné une généralisation des polynômes de Legendre. Nous considérons pour l'heure le cas où  $F$  est la fonction de répartition d'une loi binomiale, ou autrement dit où  $p_i = p(i)$ ,  $i = 0, 1, \dots, N$ , avec

$$p(i) = \binom{N}{i} p^i (1-p)^{N-i}, \quad 0 < p < 1. \quad (2.3)$$

Cela nous conduira tel qu'annoncé aux polynômes de Krawtchouk. Pour  $N$  quelconque, calculons avec Gram-Schmidt et  $K_0 = 1$ ,

$$\begin{aligned} K_1^{(p)} &= x - \text{proj}(x : 1) \\ &= x - \frac{\langle x | 1 \rangle}{\langle 1 | 1 \rangle} \\ &= x - \left( \sum_{i=0}^N p_i x_i \right) \underbrace{\left( \sum_{i=0}^N p_i \right)^{-1}}_{=1}. \end{aligned}$$

Nous remplaçons maintenant  $p_i$  et  $x_i$  dans la somme qui reste par leur valeur respective, puis nous manipulons la somme obtenue pour en arriver à utiliser une formule que nous connaissons

bien. Cela donne

$$\begin{aligned}
K_1^{(p)} &= x - \sum_{i=1}^N i \binom{N}{i} p(1-p)^{N-i} \\
&= x - \sum_{i=1}^N \frac{iN!}{i!(N-i)!} p^i (1-p)^{N-i} \\
&= x - \sum_{i=1}^N \frac{N!}{(i-1)!(N-i)!} p^i (1-p)^{N-i} \\
&= x - \sum_{i=0}^{N-1} \frac{N!}{i!(N-i-1)!} p^{i+1} (1-p)^{N-i-1} \\
&= x - Np \underbrace{\sum_{i=0}^{N-1} \binom{N-1}{i} p^i (1-p)^{N-1-i}}_{=1} \\
&= x - Np.
\end{aligned}$$

Nous pourrions de manière analogue en arriver à trouver

$$K_2^{(p)} = \frac{1}{2} [N^2 p^2 + x(2p + x - 1) - Np(p - 2x)].$$

Ainsi, jusqu'à maintenant, les polynômes de Krawtchouk se comportent de manière exactement analogue aux polynômes classiques que nous traitons en exemple depuis le début, si ce n'est que pour chaque  $N$  fixé, il se trouve exactement  $N + 1$  polynômes formant une base des polynômes de degré  $N$ , soit un ensemble fini. C'est que, comme l'indique le produit scalaire, ils sont issus d'un processus discret plutôt que continu. Ils ne sont donc pas solutions d'un problème de Sturm-Liouville, mais bien d'une *équation aux différences*, à savoir dans ce cas-ci

$$p(N-x)y(x+1) - [p(N-x) + x(1-p) - n]y(x) + x(1-p)y(x-1) = 0.$$

Ils ont aussi une formule de type Rodrigues

$$K_n^{(p)}(x, N) = (-1)^n \binom{N}{n} \left[ \binom{N}{x} \left( \frac{p}{1-p} \right)^x \right]^{-1} \nabla^N \left[ \binom{N-n}{x} \left( \frac{p}{1-p} \right)^x \right],$$

où  $\nabla f(x) := f(x) - f(x-1)$ . Dans le cas de la relation de récurrence, notre traitement était tout à fait général et il peut donc encore être appliqué. Cela donne ici

$$K_{n+1}^{(p)}(x, N) = \frac{1}{n+1} \left[ (x-n-N+2)K_n^{(p)}(x, N) - p(1-p)(N-n+1)K_{n-1}^{(p)}(x, N) \right].$$

Enfin, il en existe des « sommes génératrices », dont voici deux spécimens.

$$\sum_{n=0}^N (-1)^n p^{-n} K_n^{(p)}(x, N) t^n = \left( 1 - \frac{1-p}{p} t \right)^x (1+t)^{N-x};$$

$$\sum_{n=0}^N K_n^{(p)}(x, N) w^n = (1 + (1-p)w)^x (1-pw)^{N-x}.$$

Ayant vu tout cela, nous pouvons nous douter en quoi les polynômes de Krawtchouk sont une généralisation de ceux de Hermite. C'est le lien que donne la formule

$$\lim_{N \rightarrow \infty} \left( \frac{2}{p(1-p)} \right)^{n/2} \binom{N}{n}^{-1/2} \sqrt{n!} K_n^{(p)}(x) = H_n(z),$$

où  $x = Np + z\sqrt{2Np(1-p)}$ .

Nous n'en ferons pas la démonstration, puisqu'elle utilise d'autres formules qui précèdent, que nous n'avons pas non plus démontrées<sup>1</sup>. Le tout serait donc soit très lourd, soit peu convaincant. De toute façon, tel qu'annoncé d'entrée de jeu, cette section se voulait une ouverture pour une vue plus large de la théorie des polynômes orthogonaux, tout en lui conservant une certaine cohérence.

### 2.6.2 Application

Avec ce que nous avons vu jusqu'à maintenant, nous pourrions croire que les polynômes de Krawtchouk et leurs congénères (car ils en ont) ne sont qu'une curiosité théorique. En fait, c'est tout le contraire. Il existe par exemple une méthode de traitement d'images binaires qui utilise les moments d'ensembles de polynômes orthogonaux. Parmi les polynômes utilisés, il y a ceux de Legendre et de Zernike. Cependant, l'utilisation des moments desdits polynômes comporte certains inconvénients, entre autres celui de devoir approximer numériquement des intégrales continues. Il est raisonnable de concevoir que ce problème puisse être évité en utilisant des polynômes orthogonaux issus d'un processus discret, comme ceux de Krawtchouk. Donnons donc maintenant un aperçu de cette *méthode des moments de Krawtchouk*.

Pour un certain  $N$  fixé, définissons pour la cause les polynômes de Krawtchouk  $k_n$  par leur produit scalaire (avec la fonction  $p(x)$  définie à l'équation (2.3))

$$\sum_{x=0}^N p(x) k_n(x) k_m(x) = \rho \delta_{mn},$$

où  $\delta_{mn}$  est bien entendu le delta de Kroenecker et  $\rho = \rho(N, n, p)$  est une certaine fonction de pondération appropriée (comme  $N$  et  $p$  sont habituellement fixés, nous écrivons  $\rho(N, n, p) = \rho(n)$ ).

Les moments de Krawtchouk d'ordre  $m+n$  sont alors définis par

$$\begin{aligned} K_{mn} &= \frac{1}{\rho(n)\rho(m)} \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} p(x) k_n(x) p(y) k_m(y) f(x, y) \\ &= \frac{1}{\rho(n)\rho(m)} \left\langle k_n(x) \middle| \langle k_m(y) | f(x, y) \rangle \right\rangle \quad m, n \in \mathbb{N}, \end{aligned}$$

---

1. On peut cependant voir (16), pp. 35 à 37 pour plus de détails.

où  $f(x, y)$  est la fonction de l'image, qui pourra être reconstruite à l'aide de la transformation inverse,

$$f(x, y) = \sum_{m=0}^N \sum_{n=0}^N K_{mn} k_n(x) k_m(y).$$

L'article de Yap, Raveendran et Ong (18) montre expérimentalement que la méthode des moments est significativement meilleure lorsque des polynômes issus d'un processus discret sont utilisés, en particulier ceux de Krawtchouk. En effet, pour un développement du même ordre, un résultat plus précis est obtenu en moins de temps lors de la reconstruction d'une image avec ou sans bruit.

Forts maintenant de toutes ces définitions et toujours plus confiants de l'applicabilité de notre sujet d'étude, nous sommes prêts à établir quelques propriétés remarquables des polynômes orthogonaux, ce qui est l'objet du chapitre suivant.

## Chapitre 3

# Propriétés des polynômes orthogonaux

Dans ce chapitre, nous continuerons notre étude des propriétés des polynômes orthogonaux. Nous commencerons par établir une formule de Christoffel, qui permet de générer toute une famille de suites de polynômes orthogonaux à partir d'une seule de ces suites. Ensuite, la notion importante des séries de Fourier généralisées sera vue plus en détail, ainsi que certains résultats d'analyse fonctionnelle qui y sont liés. Nous en viendrons alors à étudier le noyau de Christoffel-Darboux, qui a quelques propriétés remarquables. Nous finirons ce chapitre par l'étude des méthodes de quadrature de Gauss, un outil important de l'intégration numérique qui utilise les racines des polynômes orthogonaux. Nous étudierons ces dernières plus en profondeur au chapitre suivant.

### 3.1 Formule de Christoffel

Étant donnée une suite de polynômes orthogonaux, le théorème suivant (provenant de (16)) permet d'en construire une infinité d'autres.

**Théorème 3.1.** *Soit  $(p_n)_{n \geq 1}$  une suite de polynômes, orthogonaux par rapport à une mesure  $dF(x)$  de segment fondamental  $[a, b]$ . Soit de plus  $\rho(x) := c(x - x_1)(x - x_2)(x - x_1) \cdots (x - x_k)$ ,  $c \neq 0$  un polynôme de degré  $k$ , positif sur  $[a, b]$ , avec  $x_i \neq x_j$ ,  $\forall i \neq j$ . Alors l'ensemble  $(q_n)_{n \geq 1}$ , où*

$$q_n := \frac{1}{\rho(x)} \begin{vmatrix} p_n(x) & p_{n+1}(x) & \cdots & p_{n+k}(x) \\ p_n(x_1) & p_{n+1}(x_1) & \cdots & p_{n+k}(x_1) \\ \vdots & \vdots & \ddots & \vdots \\ p_n(x_k) & p_{n+1}(x_k) & \cdots & p_{n+k}(x_k) \end{vmatrix},$$

*est un ensemble de polynômes, orthogonaux par rapport à la mesure  $\rho(x)dF(x)$ .*

Dans le cas où un des  $x_i$  serait de multiplicité  $\mu > 1$ , le théorème fonctionne encore, mais en remplaçant les  $\mu - 1$  lignes supplémentaires en  $x_i$  par leur dérivée première, seconde, ...,  $\mu^e$ .

*Démonstration.* Montrons d'abord que le déterminant précédent (appelons-le  $D$ ) n'est pas identiquement nul. Développons-le par rapport à la première ligne, pour obtenir

$$D = p_n(x)D_0 - p_{n+1}(x)D_1 + p_{n+2}(x)D_2 + \dots + (-1)^k p_{n+k}(x)D_k,$$

où nous avons appelé  $D_i$  la sous-matrice de  $D$  obtenue en lui retranchant la première ligne et la  $i + 1^e$  colonne. Comme lesdits  $D_i$  sont tout simplement des nombres, le membre de droite de l'équation précédente est une combinaison linéaire de polynômes. Aussi sont-ils de degrés tous différents, donc linéairement indépendants. La combinaison linéaire ne peut alors être nulle que si tous les  $D_i$  sont nuls. Montrons donc que

$$D_k := \begin{vmatrix} p_n(x_1) & p_{n+1}(x_1) & \cdots & p_{n+k-1}(x_1) \\ p_n(x_2) & p_{n+1}(x_2) & \cdots & p_{n+k-1}(x_2) \\ \vdots & \vdots & \ddots & \vdots \\ p_n(x_k) & p_{n+1}(x_k) & \cdots & p_{n+k-1}(x_k) \end{vmatrix} \neq 0.$$

Supposons le contraire. Alors il existe une combinaison linéaire des colonnes de la matrice qui soit nulle. Autrement dit, il existe des constantes  $a_i$  ( $i = 0, 1, 2, \dots, k - 1$ ) non toutes nulles telles que

$$H(x) := a_0 p_n + a_1 p_{n+1} + \dots + a_{k-1} p_{n+k-1}$$

s'annule pour  $x = x_1, x_2, \dots, x_k$ .  $H(x)$  est donc de la forme  $\rho(x)G(x)$  pour un certain polynôme  $G(x)$  de degré inférieur ou égal à  $n - 1$ , qui est donc orthogonal à  $H(x)$ . Cela s'écrit

$$\int H(x)G(x)dF(x) = \int G(x)^2 \rho(x)dF(x) = 0.$$

Puisque  $\rho$  est positif sur  $[a, b]$ , ceci implique que

$$G(x) \equiv 0 \implies H(x) \equiv 0,$$

une contradiction.

Remarquons maintenant que par le calcul précédent, ledit déterminant doit être un polynôme de degré exactement  $n + k$ . De plus, comme chaque  $x_i$ ,  $i = 1, 2, \dots, k$  l'annule évidemment, nous voyons que ledit déterminant est divisible par  $\rho(x)$ . Ainsi,  $q_n(x)$  est bien un polynôme de degré  $n$ . Soit maintenant  $q(x)$ , un polynôme arbitraire de degré (inférieur ou égal à)  $n - 1$ . Alors

$$\begin{aligned} \int q(x)q_n(x)(\rho(x)dF(x)) &= \int q(x)(q_n(x)\rho(x))dF(x) \\ &= \int q(x) \begin{vmatrix} p_n(x) & p_{n+1}(x) & \cdots & p_{n+k}(x) \\ p_n(x_1) & p_{n+1}(x_1) & \cdots & p_{n+k}(x_1) \\ \vdots & \vdots & \ddots & \vdots \\ p_n(x_k) & p_{n+1}(x_k) & \cdots & p_{n+k}(x_k) \end{vmatrix} dF(x). \end{aligned}$$

Maintenant, le déterminant dans l'intégrale est en fait une combinaison linéaire de  $p_n, p_{n+1}, \dots, p_{n+k}$ , orthogonaux par hypothèse à tout polynôme de degré au plus  $n - 1$  (par rapport à la mesure  $dF(x)$ ). Nous avons trouvé

$$\int q(x)q_n(x)(\rho(x)dF(x)) = \int q(x)(q_n(x)\rho(x))dF(x) = 0.$$

Ainsi, comme  $q_n$  est orthogonal à tout polynôme de degré strictement inférieur pour chaque  $n$ , nous avons bien obtenu que la suite des  $(q_n)_{n \geq 1}$  forme un ensemble de polynômes orthogonaux. □

À titre d'exemple, pour illustrer l'intérêt de ce résultat somme toute fort simple, remarquons que les polynômes de Gegenbauer  $C_n^{(\alpha)}$ , avec  $\alpha + 1/2$  un entier, peuvent être obtenus ainsi à partir des polynômes de Legendre  $P_n$ . Rappelons que pour ces derniers polynômes,  $dF(x) = dx$ ,  $[a, b] = [-1, 1]$ , et prenons  $\rho(x) = (1 - x^2)^{\alpha - \frac{1}{2}}$ . Alors

- avec  $\alpha = \frac{1}{2}$ , la formule nous donne directement le résultat bien connu  $C_n^{(\frac{1}{2})} = P_n$ . D'où le fait que les polynômes de Gegenbauer sont une généralisation des polynômes de Legendre.
- avec  $\alpha = \frac{3}{2}$ , le résultat donne

$$C_n^{(\frac{3}{2})} = \frac{1}{1 - x^2} \begin{vmatrix} p_n(x) & p_{n+1}(x) & p_{n+2}(x) \\ p_n(1) & p_{n+1}(1) & p_{n+2}(1) \\ p_n(-1) & p_{n+1}(-1) & p_{n+2}(-1) \end{vmatrix}.$$

En faisant quelques calculs, nous trouvons

$$C_0^{(\frac{3}{2})} = \frac{1}{1 - x^2} \begin{vmatrix} 1 & x & \frac{3x^2-1}{2} \\ 1 & 1 & 1 \\ 1 & -1 & 1 \end{vmatrix} = 3,$$

$$C_1^{(\frac{3}{2})} = \frac{1}{1 - x^2} \begin{vmatrix} x & \frac{3x^2-1}{2} & \frac{5x^3-3x}{2} \\ 1 & 1 & 1 \\ -1 & 1 & -1 \end{vmatrix} = -5x,$$

$$C_2^{(\frac{3}{2})} = \frac{1}{1 - x^2} \begin{vmatrix} \frac{3x^2-1}{2} & \frac{5x^3-3x}{2} & \frac{35x^4-30x^2+3}{8} \\ 1 & 1 & 1 \\ -1 & 1 & -1 \end{vmatrix} = \frac{35x^2 - 7}{4},$$

qui sont bien proportionnels aux polynômes de Gegenbauer classiques.

## 3.2 Séries de Fourier généralisées, propriétés d'extrema et de fermeture

### 3.2.1 Contexte

Les séries de Fourier classiques sont utilisées pour représenter une fonction périodique ou encore toute autre fonction sur un intervalle fini avec des « monômes trigonométriques », c'est-à-dire des fonctions de la forme  $\sin nx$  ou  $\cos nx$ ,  $n \in \mathbb{N}$ . Fait à remarquer, la suite de ces fonctions est en fait orthogonale par rapport à la mesure de Lebesgues sur  $[0, 2\pi]$ . De plus, en étudiant les théorèmes de convergence de telles séries, on peut constater que cette dernière propriété est fondamentale, pour ne pas dire à la base de tout. Une question légitime apparaît dès lors : une fonction peut-elle être développée en série suivant une suite quelconque de fonctions orthogonales ? De manière générale, cela pourra être fait si la suite de fonctions en question est une base d'approximation de l'espace où vit la fonction à développer. C'est ce qui est appelé les séries de Fourier généralisées, telles qu'abordées à la définition 2.5, que voici.

**Définition 3.2.** Soit  $B := \{u_1, u_2, u_3, \dots\}$  une base d'approximation d'un espace séparable  $\mathbb{E}$  et soit  $f \in \mathbb{E}$ . La série que donne «  $\text{proj}(f : u_1, u_2, \dots)$  » se nomme la « série de Fourier généralisée de  $f$  dans la base  $B$  ».

### 3.2.2 Convergence

L'objet de ce paragraphe est d'étudier dans quel contexte la série de Fourier généralisée d'une fonction converge effectivement vers cette fonction. Que l'on veuille bien d'abord se rappeler le théorème d'approximation de Weierstrass, qui nous assure que pour toute fonction  $f$  continue dans un intervalle fermé et à valeurs réelles, il existe une suite de polynômes qui convergent uniformément vers  $f$ . Aussi, il est admis que toute suite de polynômes à degrés étagés (c'est-à-dire une suite contenant un polynôme de degré  $n$  pour chaque  $n = 0, 1, 2, \dots$ ) forme une base de l'espace des polynômes. Nous pouvons donc déjà conclure que toute suite de polynômes orthogonaux par rapport à une mesure de segment fondamental fermé  $[a, b]$  est une base d'approximation pour  $C[a, b]$ . C'était notre premier théorème de fermeture. Concrétisons cela d'abord avec le théorème d'extrema suivant.

**Théorème 3.3.** Soit  $\{u_1, u_2, \dots, u_n\}$  une base orthogonale d'un sous-espace  $H_n$  d'un  $\mathbb{R}$ -espace vectoriel  $V$  et soit  $f \in V$ . L'élément  $h \in H_n$  qui minimise la norme  $\|f - h\|$  s'écrit

$$h = \text{proj}(f : u_1, u_2, \dots, u_n) = \sum_{k=1}^n \frac{a_k}{\|u_k\|^2} u_k,$$

où  $a_k := \langle f | u_k \rangle$ .

*Démonstration.* Soit  $g := \sum_{k=1}^n b_k u_k$ , un élément arbitraire de  $H_n$ , alors

$$\begin{aligned} \|f - g\|^2 &= \langle f - g | f - g \rangle = \langle f | f \rangle - 2\langle f | g \rangle + \langle g | g \rangle \\ &= \|f\|^2 - 2 \sum_{k=1}^n b_k \langle f | u_k \rangle + \sum_{k=1}^n b_k^2 \|u_k\|^2 \\ &= \|f\|^2 - 2 \sum_{k=1}^n b_k a_k + \sum_{k=1}^n b_k^2 \|u_k\|^2 \\ &= \|f\|^2 + \sum_{k=1}^n \left( b_k \|u_k\| - \frac{a_k}{\|u_k\|} \right)^2 - \sum_{k=1}^n \frac{a_k^2}{\|u_k\|^2}. \end{aligned}$$

Ce calcul donne clairement que les valeurs de  $b_k$  qui minimisent  $\|f - g\|^2$  sont bien les valeurs de  $\frac{a_k}{\|u_k\|^2}$ , c'est-à-dire que la norme est minimale lorsque  $g = h$ .  $\square$

Avec les mêmes hypothèses et notations que dans le théorème, nous obtenons comme corollaire immédiat de ce résultat l'égalité

$$\|f - \text{proj}(f : u_1, u_2, \dots, u_n)\|^2 = \|f\|^2 - \sum_{k=1}^n \left( \frac{a_k}{\|u_k\|} \right)^2,$$

qui mène à l'inégalité, dite *inégalité de Bessel*,

$$\sum_{k=1}^n \left( \frac{a_k}{\|u_k\|} \right)^2 \leq \|f\|^2.$$

Lorsque  $n \rightarrow \infty$ , le terme de gauche de cette inégalité ne peut qu'augmenter. Or, il est borné par  $\|f\|^2$ . Cela nous permet de conclure que la série  $\sum_{k=1}^{\infty} \left( \frac{a_k}{\|u_k\|} \right)^2$  converge.

Un autre corollaire immédiat du théorème précédent va comme suit.

**Corollaire 3.4.** *Soit  $(p_n)_{n \geq 1}$ , une suite de polynômes orthogonaux. En prenant  $f := p_{n+1}$  et  $H_n = \{p_1, p_2, \dots, p_n\}$ , nous obtenons que l'élément qui minimise la norme  $\|f - h\|$  est  $h = 0$ , c'est-à-dire que  $p_{n+1}$  (normalisé) est le polynôme de degré  $n + 1$  de norme minimale.*

Pour la suite, il nous faut maintenant introduire un certain type de convergence.

**Définition 3.5.** Une série  $\sum_{k=1}^{\infty} a_k u_k$  (où les  $a_k$  sont des scalaires et les  $u_k$  des éléments d'un espace vectoriel  $V$ ) converge en moyenne s'il existe  $g \in V$  tel que

$$\left\| g - \sum_{k=1}^n a_k u_k \right\| \rightarrow 0 \quad (n \rightarrow \infty).$$

Il se dit aussi que « la série converge vers  $g$  au sens des moindres carrés ».

**Remarque 3.6.** En appliquant ce qui précède, nous voyons que l'égalité de Parseval suit si la série  $\sum_{k=0}^{\infty} \text{proj}(f : p_k)$  converge en moyenne vers  $f$ , c'est-à-dire

$$\|f\|^2 = \sum_{k=0}^{\infty} \|\text{proj}(f : p_k)\|^2.$$

### 3.2.3 Polynômes et bases d'approximation

Pour que le résultat précédent ait un intérêt quelconque pour notre étude, il faut évidemment que l'espace que nous cherchons à étudier ait comme base d'approximation un ensemble de polynômes orthogonaux. Mais est-ce toujours le cas? Nous avons vu en début de section que la réponse est affirmative pour  $C[a, b]$  avec  $a$  et  $b$  finis. Nous pourrions aussi montrer le résultat plus général suivant (voir (16), p. 10).

**Théorème 3.7.** *Soient  $p \geq 1$ , une mesure  $d\alpha(x)$  de Lebesgues-Stieltjes et  $f(x) \in L^p_\alpha[a, b]$  où  $a$  et  $b$  sont finis. Alors pour chaque  $\epsilon > 0$ , il existe une fonction  $F(x) \in C[a, b]$  telle que  $\|f(x) - F(x)\|_p^p < \epsilon$ .*

En appliquant le théorème de Weierstrass à ce résultat, nous obtenons bien qu'une suite de polynômes orthogonaux est une base d'approximation pour  $L^2_\alpha[a, b]$ . Qu'en est-il maintenant si le segment fondamental n'est pas fini, comme pour les polynômes de Laguerre ou de Hermite? Le résultat précédent ne dit rien dans ce cas. Pour le cas particulier de ces deux ensembles de polynômes, une autre démonstration fonctionne. Cependant, il n'est pas nécessairement vrai en général que les polynômes sont denses dans les fonctions continues vivant dans un espace de Lebesgues-Stieltjes. Le contre-exemple suivant illustre cela.

Prenons la mesure  $d\alpha(x) = \exp(-x^\mu \cos(\mu\pi))dx$ , et la fonction  $f(x) = \sin(x^\mu \sin(\mu\pi))$ , où  $0 < \mu < 1/2$ . Alors il est possible de montrer que (voir (16), p. 40),

$$\int_0^\infty f(x)x^n d\alpha(x) = 0 \quad \forall n \geq 0.$$

Donc, étant donné un polynôme arbitraire  $\rho(x)$ ,

$$\begin{aligned} A &= \int_0^\infty (f(x))^2 d\alpha(x) \\ &= \int_0^\infty f(x)(f(x) - \rho(x)) d\alpha(x) \\ &\leq \int_0^\infty |f(x)||f(x) - \rho(x)| d\alpha(x) \\ &\leq \int_0^\infty |f(x) - \rho(x)| d\alpha(x) \\ &\stackrel{\text{C.-S.}}{\leq} \left( \int_0^\infty |f(x) - \rho(x)|^2 d\alpha(x) \right)^{1/2} \left( \int_0^\infty d\alpha(x) \right)^{1/2}, \end{aligned}$$

Cela montre que les intégrales

$$\int_0^\infty |f(x) - \rho(x)| d\alpha(x) \text{ et } \int_0^\infty |f(x) - \rho(x)|^2 d\alpha(x)$$

ne peuvent être rendues arbitrairement petites, ou autrement dit que  $f(x)$  ne peut être approchée par des polynômes dans cet espace  $L^k_\alpha[0, \infty]$ , avec  $k = 1, 2$ .

### 3.2.4 Un dernier théorème d'extrema

Voici maintenant un dernier théorème d'extrema, qui servira à introduire la prochaine section.

**Théorème 3.8.** *Soit  $x_0 \in \mathbb{C}$  et soit l'ensemble*

$$S := \{p, \text{ polynômes de degré } n \text{ tels que } \|p(x)\|^2 = 1 \text{ dans } L_\alpha^2[a, b]\}.$$

*Soit de plus  $\{p_n\}$  la suite de polynômes orthogonaux dudit espace. Alors le polynôme  $\rho \in S$  maximisant la quantité  $|\rho(x_0)|^2$  s'écrit*

$$\rho(x) = K_n(x_0, x_0)^{-1/2} K_n(x_0, x),$$

où

$$K_n(x_0, x) = \sum_{k=0}^n \frac{\overline{p_k(x_0)} p_k(x)}{\langle p_k | p_k \rangle} = \sum_{k=0}^n \frac{p_k(\overline{x_0}) p_k(x)}{\langle p_k | p_k \rangle}.$$

*Le maximum est donné par  $K_n(x_0, x_0)$ .*

**Remarque 3.9.**  $K_n(x, y)$  est appelé le « noyau de Christoffel-Darboux » et sera étudié plus en détail à la section suivante. Aussi, par définition même,  $K_n(x, x) \geq 0$ .

*Démonstration.* Écrivons

$$\rho = \sum_{k=0}^n \frac{\langle \rho | p_k \rangle}{\langle p_k | p_k \rangle} p_k = \sum_{k=0}^n \lambda_k \frac{p_k}{\|p_k\|}.$$

Alors la condition  $\|\rho\|^2 = 1$  peut s'écrire  $\sum_{k=0}^n |\lambda_k|^2 = 1$ , et l'inégalité de Cauchy donne alors

$$|\rho(x_0)|^2 \leq \sum_{k=0}^n |\lambda_k|^2 \sum_{k=0}^n \frac{|p_k(x_0)|^2}{\langle p_k | p_k \rangle} = K_n(x_0, x_0).$$

Cette borne est évidemment atteinte si  $\lambda_k = \lambda p_k(x_0)$ , où  $\lambda$  est déterminé par la condition

$$|\lambda|^2 \sum_{k=0}^n |p_k(x_0)|^2 = 1.$$

□

## 3.3 Noyau de Christoffel-Darboux

Le noyau de Christoffel-Darboux, tel que défini à la section précédente, comporte assez de propriétés intéressantes pour se mériter une section du présent chapitre. Une première de ces propriétés est qu'il peut être utilisé pour représenter les sommes partielles du développement en séries de Fourier généralisées d'une fonction. En effet,

$$\sum_{k=0}^n \text{proj}(f : p_k) = \sum_{k=0}^n \frac{p_k(x)}{\langle p_k | p_k \rangle} \int_a^b f(t) p_k(t) d\alpha(t) = \int_a^b f(t) K_n(t, x) d\alpha(t).$$

De ce qui précède, il vient aussi que pour un polynôme arbitraire  $\rho(x)$  de degré au plus  $n$

$$\int_a^b K_n(t, x)\rho(t)d\alpha(t) = \rho(x). \quad (3.1)$$

Par ailleurs, remarquons que cette équation définit  $K_n(t, x)$  de manière unique. En effet, supposons qu'il existe un autre polynôme  $k(x, t)$  de degré au plus  $n$  en  $t$  tel que

$$\int_a^b k_n(t, x)\rho(t)d\alpha(t) = \rho(x)$$

pour chaque polynôme  $\rho$  de degré au plus  $n$ . Alors

$$\int_a^b (K_n(t, x) - k_n(t, x))\rho(t)d\alpha(t) \equiv 0$$

pour chaque polynôme  $\rho$  de degré au plus  $n$ , c'est-à-dire que  $K_n(t, x) - k_n(t, x)$  est un polynôme de degré au plus  $n$  en  $t$  qui est orthogonal à tout polynôme de degré au plus  $n$  en  $t$ . La seule possibilité est que

$$K_n(t, x) - k_n(t, x) \equiv 0 \implies K_n(t, x) = k_n(t, x),$$

une contradiction.

Le résultat suivant se déduit de l'équation (3.1).

**Théorème 3.10.** *Supposons que  $a$  et  $x_0$  soient finis, avec  $x_0 \leq a$ . Alors les polynômes  $\{K_n(x_0, x)\}$  sont orthogonaux par rapport à la mesure  $(x - x_0)d\alpha(x)$ .*

*Démonstration.* Posons  $x = x_0$ ,  $\rho(t) = (t - x_0)r(t)$  dans l'équation précédente, avec  $r(t)$  un polynôme arbitraire de degré  $n - 1$ . Nous obtenons

$$\int_a^b K_n(t, x_0)(t - x_0)r(t)d\alpha(t) = (x_0 - x_0)r(x_0) = 0,$$

c'est-à-dire que  $K_n(x, x_0)$  est orthogonal à tout polynôme de degré  $n - 1$  par rapport au produit scalaire mentionné dans l'énoncé, ce qui est suffisant pour conclure.  $\square$

Voyons maintenant le dernier résultat de cette section, mais non le moindre.

**Théorème 3.11** (Formule de Christoffel-Darboux).

$$K_n(x, y) = \frac{cd(P_n)}{cd(P_{n+1})\|P_n\|^2} \frac{P_{n+1}(x)P_n(y) - P_n(x)P_{n+1}(y)}{(x - y)}.$$

*Démonstration.* Il suffit de montrer que le membre de droite de l'égalité précédente satisfait l'égalité (3.1) plus haut. L'intégrale s'écrit, pour un polynôme arbitraire  $\rho(x)$  de degré  $n$ ,

$$I := \frac{cd(P_n)}{\|P_n\|^2 cd(P_{n+1})} \int_a^b \frac{P_{n+1}(x)P_n(t) - P_n(x)P_{n+1}(t)}{x - t} \rho(t) d\alpha(t).$$

En réorganisant, cela peut se réécrire

$$I = I_1 + I_2 + I_3,$$

où

$$\begin{aligned} I_1 &= \frac{\text{cd}(P_n)}{\|P_n\|^2 \text{cd}(P_{n+1})} \int_a^b [P_{n+1}(x)P_n(t) - P_n(x)P_{n+1}(t)] \frac{\rho(t) - \rho(x)}{x-t} d\alpha(t), \\ I_2 &= \frac{\text{cd}(P_n)}{\|P_n\|^2 \text{cd}(P_{n+1})} \rho(x) \int_a^b P_n(t) \frac{P_{n+1}(x) - P_{n+1}(t)}{x-t} d\alpha(t), \\ I_3 &= \frac{\text{cd}(P_n)}{\|P_n\|^2 \text{cd}(P_{n+1})} \rho(x) \int_a^b P_{n+1}(t) \frac{P_n(t) - P_n(x)}{x-t} d\alpha(t). \end{aligned}$$

Maintenant, remarquons que  $\frac{\rho(t)-\rho(x)}{x-t}$  et  $\frac{P_n(t)-P_n(x)}{x-t}$  sont en fait des polynômes en  $t$  de degré  $n-1$ . Ils sont donc orthogonaux à  $P_n$  et  $P_{n+1}$ . Vues comme des produits scalaires, il est donc clair que la première et la troisième intégrale sont nulles. Aussi, dans la seconde, remarquons que  $\frac{P_{n+1}(x)-P_{n+1}(t)}{x-t}$  est en fait un polynôme de degré  $n$  en  $t$ . En traitant encore l'intégrale comme un produit scalaire, nous voyons que seul le coefficient de  $t^n$  compte. Puisqu'évidemment

$$\frac{\text{cd}(P_n)}{\text{cd}(P_{n+1})} \frac{P_{n+1}(x) - P_{n+1}(t)}{x-t} = \text{cd}(P_n)t^n + \dots,$$

il vient

$$\begin{aligned} I = I_2 &= \frac{\text{cd}(P_n)}{\|P_n\|^2 \text{cd}(P_{n+1})} \rho(x) \int_a^b P_n(t) \frac{P_{n+1}(x) - P_{n+1}(t)}{x-t} d\alpha(t) \\ &= \frac{1}{\|P_n\|^2} \rho(x) \langle P_n(t) | \text{cd}(P_n)t^n + \dots \rangle \\ &= \rho(x). \end{aligned}$$

□

Le résultat aurait aussi pu être déduit de la relation de récurrence entre les polynômes orthogonaux.

**Corollaire 3.12.** *Si  $x = y$ , alors*

$$K_n(x, x) = \frac{\text{cd}(P_n)}{\|P_n\|^2 \text{cd}(P_{n+1})} [P'_{n+1}(x)P_n(x) - P'_n(x)P_{n+1}(x)].$$

*Démonstration.* Laissons tendre  $y$  vers  $x$  dans la formule précédente et appliquons la règle de l'Hospital.

$$\begin{aligned} &\lim_{y \rightarrow x} \frac{P_{n+1}(x)P_n(y) - P_n(x)P_{n+1}(y)}{x-y} \\ &\stackrel{H}{=} \lim_{y \rightarrow x} -[P_{n+1}(x)P'_n(y) - P_n(x)P'_{n+1}(y)] \\ &= P'_{n+1}(x)P_n(x) - P'_n(x)P_{n+1}(x). \end{aligned}$$

□

### 3.4 Quadratures de Gauss

Il est question de méthode de quadrature lorsqu'on veut approximer numériquement la valeur d'une intégrale. Une méthode très simple et bien connue s'appelle la méthode du trapèze. Il s'agit de faire la moyenne de la valeur de la fonction aux deux extrémités de l'intervalle d'intégration et de la multiplier par la longueur de cet intervalle. Cela donne

$$\int_a^b f(x)dx \approx (b-a) \frac{f(a) + f(b)}{2}.$$

Cette méthode requiert donc deux évaluations de  $f(x)$  et est de degré de précision 1, c'est-à-dire que la quadrature est exacte dans le cas des polynômes de degré 1. Cherchons maintenant à atteindre un degré de précision supérieur avec une méthode analogue

$$\int_a^b f(x)dx \approx w_1 f(t_1) + w_2 f(t_2).$$

Les  $w_k$  sont appelés *poinds d'intégration* et les  $t_k$  *points d'intégration*. En prenant  $w_1 = w_2 = \frac{b-a}{2}$ ,  $t_1 = a$ ,  $t_2 = b$ , nous retrouvons la méthode du trapèze. Mais il y a mieux. Ce sont les quadratures de Gauss.

#### 3.4.1 Quadratures de Gauss-Legendre et de Gauss-Laguerre à deux points

Le type de quadrature à utiliser dépend de la forme de l'intervalle d'intégration de l'intégrale à approximer. En effet, cet intervalle doit être de la même forme que l'intervalle fondamental du produit scalaire des polynômes orthogonaux utilisés. (Nous verrons tout à l'heure pourquoi il est question de polynômes ici.) Par exemple, pour un intervalle fini  $[a, b]$ , nous pouvons utiliser Gauss-Legendre (voir (6, chap. 6), pour une demi-droite  $[0, \infty)$ , nous pouvons utiliser Gauss-Laguerre et pour tous les réels  $\mathbb{R} = (-\infty, \infty)$ , nous utilisons Gauss-Hermite. Clarifions tout cela par deux exemples.

Pour le cas d'un intervalle fermé, il est toujours possible de se ramener à une intégrale sur  $[-1, 1]$  par le changement de variable

$$x = \frac{(b-a)t + (a+b)}{2} \implies dx = \frac{b-a}{2} dt.$$

En effet, cela donne

$$\int_a^b f(x)dx = \int_{-1}^1 f\left(\frac{(b-a)t + (a+b)}{2}\right) \frac{b-a}{2} dt = \frac{b-a}{2} \int_{-1}^1 g(t)dt,$$

où nous avons posé

$$g(t) := f\left(\frac{(b-a)t + (a+b)}{2}\right).$$

Nous nous contenterons donc pour l'instant de regarder ce cas particulier. Cherchons une méthode de la forme

$$\int_{-1}^1 f(x)dx \approx w_1 f(t_1) + w_2 f(t_2),$$

avec un degré de précision qui soit le plus élevé possible. Comme il y a quatre inconnues dans l'équation précédente, quatre contraintes peuvent être formulées. Nous voulons donc pouvoir évaluer précisément des intégrales impliquant des polynômes de degrés 0, 1, 2 et 3. Cela peut s'écrire

$$\begin{cases} \int_{-1}^1 dx = w_1 + w_2 \\ \int_{-1}^1 x dx = w_1 t_1 + w_2 t_2 \\ \int_{-1}^1 x^2 dx = w_1 t_1^2 + w_2 t_2^2 \\ \int_{-1}^1 x^3 dx = w_1 t_1^3 + w_2 t_2^3, \end{cases}$$

d'où le système non linéaire

$$\begin{cases} 2 = w_1 + w_2 \\ 0 = w_1 t_1 + w_2 t_2 \\ \frac{2}{3} = w_1 t_1^2 + w_2 t_2^2 \\ 0 = w_1 t_1^3 + w_2 t_2^3. \end{cases}$$

En résolvant ce système, il vient

$$\begin{cases} w_1 = w_2 = 1 \\ t_1 = -\sqrt{\frac{1}{3}} \\ t_2 = \sqrt{\frac{1}{3}}. \end{cases}$$

**Remarque 3.13.**

- (i) Ici, les  $t_i$  sont les racines du polynôme de Legendre  $P_2$ .
- (ii) Pour un même nombre de points d'intégration, nous avons gagné en degré de précision par rapport à la méthode du trapèze (3 vs 1).

Nous avons donc trouvé la méthode de Gauss-Legendre à deux points

$$\int_{-1}^1 f(x) dx \approx f\left(-\sqrt{\frac{1}{3}}\right) + f\left(\sqrt{\frac{1}{3}}\right).$$

Considérons maintenant des intégrales sur des demi-droites  $[a, \infty)$  ou  $(-\infty, b]$ . Pour celles-ci, nous utilisons le contexte des polynômes de Laguerre. Nous nous ramenons donc sur l'intervalle  $[0, \infty)$  avec un changement de variable approprié, puis nous considérons

$$\int_0^\infty f(x) dx = \int_0^\infty (f(x)e^x) e^{-x} dx = \int_0^\infty g(x) e^{-x} dx,$$

où  $g(x) := f(x)e^x$ . En appliquant le même principe que précédemment, nous obtenons le système

$$\begin{cases} 1 = \int_0^\infty e^{-x} dx = w_1 + w_2 \\ 1 = \int_0^\infty x e^{-x} dx = w_1 t_1 + w_2 t_2 \\ 2 = \int_0^\infty x^2 e^{-x} dx = w_1 t_1^2 + w_2 t_2^2 \\ 6 = \int_0^\infty x^3 e^{-x} dx = w_1 t_1^3 + w_2 t_2^3, \end{cases}$$

nettement plus difficile à résoudre que le premier. (D'ailleurs, le système que donnent les quadratures de Gauss-Laguerre à trois points ou plus n'a pas de solution analytique.) La solution en est néanmoins connue. On peut en effet trouver

$$\begin{cases} w_1 = \frac{2+\sqrt{2}}{4} \\ w_2 = \frac{2-\sqrt{2}}{4} \\ t_1 = 2 - \sqrt{2} \\ t_2 = 2 + \sqrt{2}. \end{cases}$$

**Remarque 3.14.** De manière analogue à l'exemple précédent, les  $t_i$  sont les racines du polynôme de Laguerre  $L_2$ . Nous verrons dans la suite que ce comportement est tout à fait général.

### 3.4.2 Cas général

Nous verrons maintenant quelques théorèmes auxquels obéissaient les exemples précédents. Ceux-ci feront apparaître le lien entre les quadratures de Gauss et les polynômes orthogonaux.

**Théorème 3.15.** Soit  $(p_n)_{n \geq 1}$ , une suite de polynômes, orthogonaux par rapport à la mesure  $d\alpha(x)$ . Alors il existe des constantes  $\lambda_{kn}$  ( $k = 1, 2, \dots, n$ ) et  $x_{1n} < x_{2n} < \dots < x_{nn}$  telles que

$$\int_I \rho(x) d\alpha(x) = \sum_{k=1}^n \lambda_{kn} \rho(x_{kn})$$

pour tout polynôme  $\rho$  de degré  $2n - 1$  (ou moins), déterminées uniquement par  $n$  et  $d\alpha(x)$ .

**Remarque 3.16.** Nous justifierons au chapitre suivant le fait que les zéros de  $p_n$  soient réels et distincts.

*Démonstration.* Contruisons par interpolation de Lagrange<sup>1</sup> le polynôme de degré  $n - 1$  qui coïncide avec  $\rho(x)$  aux points  $x_{kn}$ , que nous choisirons comme étant les zéros de  $p_n$ . Cela s'écrit

$$L(x) = \sum_{k=1}^n \rho(x_{kn}) \frac{p_n(x)}{p'_n(x_{kn})(x - x_{kn})} = \sum_{k=1}^n \rho(x_{kn}) l_{kn}(x),$$

où il a évidemment été posé

$$l_{kn}(x) := \frac{p_n(x)}{p'_n(x_{kn})(x - x_{kn})}.$$

Maintenant,  $p_n(x)$  divise  $\rho(x) - L(x)$ , c'est-à-dire qu'il existe  $r(x)$ , un polynôme de degré au plus  $n - 1$ , tel que

$$\rho(x) - L(x) = p_n(x)r(x).$$

---

1. Voir l'annexe B pour un peu de théorie concernant l'interpolation de Lagrange.

Ainsi

$$\begin{aligned}\int_I \rho(x) d\alpha(x) &= \int_I L(x) d\alpha(x) + \underbrace{\int_I p_n(x)r(x) d\alpha(x)}_{=0} \\ &= \int_I L(x) d\alpha(x) \\ &= \sum_{k=1}^n \rho(x_{kn}) \int_I l_{kn}(x) d\alpha(x).\end{aligned}$$

Nous avons donc bien le résultat, moyennant que

$$\lambda_{kn} = \int_I l_{kn}(x) d\alpha(x).$$

Concernant l'unicité, supposons qu'il existe des nombres  $\lambda_{kn}$  ( $k = 1, 2, \dots, n$ ) et  $x_1 < x_2 < \dots < x_n$  tels que l'équation

$$\int_I \rho(x) d\alpha(x) = \sum_{k=1}^n \lambda_{kn} \rho(x_k)$$

tienne pour un polynôme  $\rho(x)$  de degré  $2n - 1$  arbitraire. Prenons  $\rho(x) = q(x)r(x)$ , avec  $q(x) = (x - x_1)(x - x_2) \cdots (x - x_n)$  et  $r(x)$  un polynôme de degré  $n - 1$  arbitraire. Alors l'équation précédente donne

$$\int_I q(x)r(x) d\alpha(x) = 0,$$

ce qui signifie qu'en fait  $q(x)$  est un multiple de  $p_n$ , et donc que les  $x_i$  doivent bien en être les racines. L'unicité des  $\lambda_{kn}$  suit.  $\square$

Les  $\lambda_{kn}$  que nous avons considérés dans le théorème sont parfois appelés *nombres de Christoffel* et ont certaines propriétés que voici.

**Propriétés 3.17.**

(i) Nous avons la formule

$$\sum_{k=1}^n \lambda_{kn} = \int_I d\alpha(x) = \alpha(b) - \alpha(a)$$

(ii) et les représentations

$$\lambda_{kn} = \int_I \left( \frac{p_n(x)}{p'_n(x_{kn})(x - x_{kn})} \right)^2 d\alpha(x),$$

(iii)

$$\lambda_{kn} = \frac{\text{cd}(p_{n+1})}{\text{cd}(P_n)} \frac{-\|p_n\|^2}{p_{n+1}(x_{kn})p'_n(x_{kn})} = \frac{\text{cd}(p_n)}{\text{cd}(P_{n-1})} \frac{\|p_n\|^2}{p_{n-1}(x_{kn})p'_n(x_{kn})},$$

(iv)

$$\lambda_{kn}^{-1} = \sum_{i=0}^n (p_i(x_{kn}))^2 = K_n(x_{kn}, x_{kn}),$$

où  $K_n(x, y)$  est le noyau de Christoffel-Darboux.

(v) Les  $\lambda_{kn}$  sont positifs.

*Démonstration.* Dans l'ordre,

(i) Il suffit de prendre  $\rho(x) = 1$  dans l'équation du théorème 3.15.

(ii) De même, il suffit de prendre  $\rho(x) = (l_{kn}(x))^2 = \left( \frac{p_n(x)}{p'_n(x_{kn})(x-x_{kn})} \right)^2$  dans ladite équation.

(iii) C'est une application de la formule de Christoffel-Darboux

$$K_n(x, y) = \frac{\text{cd}(p_n)}{\text{cd}(p_{n+1})} \frac{p_{n+1}(x)p_n(y) - p_n(x)p_{n+1}(y)}{\|p_n\|^2(x-y)},$$

ainsi que de la formule (3.1) de la section 3.3. En effet, posons  $y = x_{kn}$  et intégrons par rapport à  $d\alpha(x)$ . Cela donne

$$\begin{aligned} 1 &= \int_I K_n(x, x_{kn}) d\alpha(x) \\ &= \int_I \frac{\text{cd}(p_n)}{\text{cd}(p_{n+1})} \frac{p_{n+1}(x)p_n(x_{kn}) - p_n(x)p_{n+1}(x_{kn})}{\|p_n\|^2(x-x_{kn})} d\alpha(x) \\ &= \frac{\text{cd}(p_n)}{\text{cd}(p_{n+1})} \int_I \frac{-p_n(x)p_{n+1}(x_{kn})}{\|p_n\|^2(x-x_{kn})} d\alpha(x) \\ &= -\frac{\text{cd}(p_n)}{\text{cd}(p_{n+1})} \frac{p_{n+1}(x_{kn})p'_n(x_{kn})\lambda_{kn}}{\|p_n\|^2}. \end{aligned}$$

(iv) C'est une application du corollaire de la formule de Christoffel-Darboux et de la propriété (iii).

$$\begin{aligned} K_n(x_{kn}, x_{kn}) &= \frac{\text{cd}(p_n)}{\|p_n\|^2 \text{cd}(p_{n+1})} [p'_{n+1}(x_{kn})p_n(x_{kn}) - p'_n(x_{kn})p_{n+1}(x_{kn})] \\ &= \frac{\text{cd}(p_n)}{\|p_n\|^2 \text{cd}(p_{n+1})} [-p'_n(x_{kn})p_{n+1}(x_{kn})] \\ &= \frac{1}{\lambda_{kn}}. \end{aligned}$$

(v) C'est une conséquence évidente de la représentation (ii).

□

## Chapitre 4

# Zéros et points critiques

Un des domaines de l'analyse mathématique toujours actifs de nos jours est la recherche de liens entre les coefficients d'un polynôme, ses zéros et ses points critiques (c'est-à-dire les zéros de sa dérivée). Pour les polynômes orthogonaux en particulier, on trouve très peu de résultats généraux de cette sorte, et la première section de ce chapitre sera consacrée à trois d'entre eux. Dans la seconde section, nous donnerons une classification des points critiques ainsi qu'une interprétation physique de ceux de type 2. Il existe par ailleurs une foule de théorèmes sur les zéros de suites particulières de polynômes orthogonaux, que nous n'aborderons par ici. Nous préférons exposer dans les troisième et quatrième sections des résultats concernant des polynômes d'une variable en général. Dans la section 2, des résultats de base tirés de (13) seront donnés, puis dans la section 3, nous donnerons le théorème de Gauss-Lucas ainsi qu'une réciproque, récemment démontrée dans (12) (en 2014).

### 4.1 Zéros des polynômes orthogonaux

Les deux théorèmes suivants, bien que démontrés avec des outils de base, ont une grande importance dans la théorie des polynômes orthogonaux. Le troisième est un théorème de densité des zéros intéressant par lui-même.

**Théorème 4.1.** *Le  $n^e$  polynôme d'une suite de polynômes orthogonaux,  $P_n$  ( $n \geq 0$ ) admet  $n$  racines réelles distinctes, toutes situées à l'intérieur du segment fondamental.*

*Démonstration.* Soit  $m \geq 0$ , le nombre de zéros de  $P_n$  qui soient réels et situés sur le segment fondamental, que nous noterons  $[a, b]$ . Le théorème fondamental de l'algèbre assure d'abord que  $m \leq n$ . Soit donc  $\Phi := \{x_1, x_2, \dots, x_m\}$  l'ensemble (potentiellement vide) desdits zéros de  $P_n$ . Soit de plus le polynôme  $S(x) := (x - x_1)(x - x_2) \dots (x - x_m)$  où nous nous rappelons la convention selon laquelle un produit vide vaut 1. Ce polynôme possède les mêmes zéros que  $P_n$  dans  $[a, b]$ , et donc le produit  $S(x)P_n(x)$  doit être du même signe sur tout l'intervalle  $[a, b]$ , et ne s'annuler qu'en chaque  $x_i$ . Selon les conditions énoncées sur le poids d'intégration  $w(x)$ ,

il doit en être de même pour le produit  $S(x)P_n(x)w(x)$ .

Ainsi, le produit scalaire  $\langle S|P_n \rangle$  doit être non nul, puisque nous intégrons une fonction non identiquement nulle qui ne change pas de signe sur l'intervalle d'intégration. Or,  $P_n$  est orthogonal à tous les polynômes de degré strictement inférieur (puisqu'ils sont combinaisons linéaires des  $P_k$  pour  $k = 0, 1, \dots, n-1$ ). Il faut donc avoir  $m \geq n$ . Mais nous avons déjà  $m \leq n$ . Nous avons donc montré que  $m = n$ , d'où le résultat.  $\square$

Remarquons que si toutes les racines d'un polynôme sont sur le segment fondamental, il en est de même pour ses points critiques. En effet, par le théorème de Rolle (dont Gauss-Lucas qui suivra se veut une généralisation), il doit s'en retrouver au moins un entre chaque paire de zéros voisins du polynôme, donc en fait exactement un. Maintenant, en regard du théorème précédent, le résultat suivant paraît assez naturel.

**Théorème 4.2.** *Soit  $\{P_n : n = 0, 1, 2, \dots\}$  une suite de polynômes orthogonaux. Alors les racines de  $P_n$  se trouvent strictement entre les racines de  $P_{n+1}$ .*

*Démonstration.* Nous utilisons le noyau de Christoffel-Darboux, développé au chapitre précédent. Nous avons vu que d'une part,  $K_n(x, x) \geq 0$ , et d'autre part que

$$K_n(x, x) = P'_{n+1}(x)P_n(x) - P'_n(x)P_{n+1}(x). \quad (4.1)$$

Soit donc deux zéros consécutifs  $\alpha$  et  $\beta$  de  $P_{n+1}$ . Comme ils sont consécutifs,  $P'_{n+1}(\alpha)$  et  $P'_{n+1}(\beta)$  doivent être de signes opposés. Or (4.1) donne, avec ce qui la précédait

$$\begin{cases} P'_{n+1}(\alpha)P_n(\alpha) \geq 0 \\ P'_{n+1}(\beta)P_n(\beta) \geq 0. \end{cases}$$

Il suit que  $P_n(\alpha)$  et  $P_n(\beta)$  sont de signes opposés, et donc que  $P_n$  a un zéro dans  $[\alpha, \beta]$  par le théorème des valeurs intermédiaires. Comme finalement il y a  $n$  tels intervalles  $[\alpha, \beta]$  et que  $P_n$  a au plus  $n$  zéros réels, le théorème suit.  $\square$

Dans l'annexe D sont consignés quelques exemples de graphiques qui aident à voir que le théorème est tout naturel.

Passons maintenant au troisième et dernier théorème de la section.

**Théorème 4.3** (densité des zéros). *Soit  $(p_n)_{n \geq 1}$  une suite de polynômes, orthogonaux par rapport à une mesure de Lebesgues-Stieltjes  $m_\alpha(x)$ . Pour tout intervalle  $[a', b'] \subseteq [a, b]$  (avec naturellement  $a' < b'$ ), il existe un  $N \in \mathbb{N}$  tel que pour chaque  $n \geq N$ , le polynôme  $p_n$  a au moins un zéro dans  $[a', b']$ .*

*Démonstration.* (Par contradiction.) Supposons le contraire, c'est-à-dire supposons que pour chaque rang  $m$ , il existe un rang  $n$  tel que  $2n-1 > m$  et tel que le polynôme  $p_n$  n'a aucun de ses zéros dans  $[a', b']$ . Appelons  $(x_{\nu n})$  ( $\nu = 1, 2, \dots, n$ ) les zéros de  $p_n$  pour chaque  $n$ .

Considérons la fonction

$$f(x) = \begin{cases} 0 & \text{dans } [a, b] \setminus [a', b'], \\ (x - a')(b' - x) & \text{dans } [a', b']. \end{cases}$$

Comme  $f$  est continue sur un intervalle fermé borné, nous savons par Weierstrass que pour tout  $\epsilon > 0$ , il existe un polynôme  $r(x)$  d'un certain degré  $M$  tel que  $|f(x) - r(x)| < \epsilon$  pour tout  $x \in [a, b]$ . Par hypothèse, il existe un rang  $N$  tel que  $2N - 1 > M$  et tel que  $p_N$  n'a aucun de ses zéros  $x_{\nu N}$  dans  $[a', b']$ . Alors, en utilisant une quadrature de Gauss, tel que présenté à la section 3.4., nous pouvons obtenir

$$\int_a^b r(x) d\alpha(x) = \sum_{\nu=1}^N \lambda_{\nu N} r(x_{\nu N}).$$

Comme tous les  $x_{\nu N}$  sont dans  $[a, b] \setminus [a', b']$ , que  $f \equiv 0$  sur cet ensemble et que  $|f(x) - r(x)| < \epsilon$ , il vient que  $|r(x_{\nu N})| < \epsilon$  pour chaque  $x_{\nu N}$ . De plus, les  $\lambda_{\nu N}$  sont tous positifs (voir la discussion à ce propos en 3.4). Ainsi

$$\sum_{\nu=1}^N \lambda_{\nu N} r(x_{\nu N}) \leq \sum_{\nu=1}^N \lambda_{\nu N} \epsilon.$$

En mettant tout cela ensemble, nous pouvons faire le calcul suivant

$$\begin{aligned} \int_a^b f(x) d\alpha(x) &\leq \int_a^b (r(x) + \epsilon) d\alpha(x) \\ &= \sum_{\nu=1}^N \lambda_{\nu N} r(x_{\nu N}) + \epsilon(\alpha(b) - \alpha(a)) \\ &\leq \epsilon \sum_{\nu=1}^N \lambda_{\nu N} + \epsilon(\alpha(b) - \alpha(a)) \\ &= \epsilon \left( \sum_{\nu=1}^N \lambda_{\nu N} + (\alpha(b) - \alpha(a)) \right) \\ &= 2\epsilon(\alpha(b) - \alpha(a)). \end{aligned}$$

En laissant  $\epsilon \rightarrow 0$ , l'inégalité reste vraie et donne

$$\int_a^b f(x) d\alpha(x) \leq 0,$$

ce qui est évidemment une contradiction. □

Dans ce théorème, nous avons supposé dès le départ que le segment fondamental était fini. Or, sous certaines conditions, notamment si  $\max_{\nu} |x_{\nu n}| = o(n)$  ( $n \rightarrow \infty$ ), le théorème est aussi valide pour les segments fondamentaux infinis (voir (16, p.111)).

## 4.2 Polynômes généraux

Dans cette section et dans la suivante, nous travaillons avec des polynômes généraux à coefficients dans  $\mathbb{C}$ . Un tel polynôme est dit *unitaire* si son coefficient directeur vaut 1. Commençons par une proposition très simple qui donne les coefficients d'un polynôme en fonction de ses zéros.

**Proposition 4.4** (Formules de Viète). *Considérons un polynôme unitaire de degré  $n$*

$$f(z) := \sum_{n=0}^n a_n z^n = \prod_{\nu=1}^n (z - z_\nu). \quad (4.2)$$

Alors les formules suivantes sont vérifiées.

$$\begin{aligned} a_{n-1} &= - \sum_j z_j; \\ a_{n-2} &= \sum_{1 \leq j < k \leq n} z_j z_k; \\ &\dots \\ a_0 &= (-1)^n \prod_i z_i. \end{aligned}$$

*Démonstration.* Le résultat suit en développant le produit de droite dans (4.2) et en égalisant les coefficients des monômes de même puissance de chaque côté.  $\square$

Le prochain théorème est une réponse partielle à la question générale : « étant donné un certain sous-ensemble de  $\mathbb{C}$ , combien un polynôme donné a-t-il de zéros dans cet ensemble ? »

**Théorème 4.5.** *Soit  $f(z) := \sum_{\nu=0}^n a_\nu z^\nu$ , un polynôme unitaire de degré  $n$ , et soit  $m \leq n$ . Alors  $f$  a au moins  $m$  zéros de module ne dépassant pas*

$$2 \max_{0 \leq \nu \leq m-1} |a_\nu|^{\frac{1}{n-\nu}}.$$

Pour faire la preuve du théorème, montrons-en d'abord un cas particulier.

**Lemme 4.6.** *Soit  $f(z) := \sum_{\nu=0}^n a_\nu z^\nu$  un polynôme unitaire de degré  $n$  tel que  $|a_\nu| \leq 1$  pour  $\nu = 0, \dots, m-1$ , ( $m \leq n$ ). Nommons les zéros  $z_j$  de  $f$  pour avoir  $|z_1| \leq \dots \leq |z_n|$ . Alors  $|z_m| \leq 2$ .*

*Démonstration du lemme.*

Si  $m = n$ , il n'y a rien à montrer, le polynôme étant supposé unitaire. Nous supposons donc

$m < n$ . Si  $|z_{m+1}| \leq 2$ , il n'y a rien à montrer non plus. Supposons donc  $|z_{m+1}| > 2$ . Appelons  $\zeta$  un zéro générique parmi  $z_{m+1}, z_{m+2}, \dots, z_n$  et posons

$$g(z) := \frac{f(z)}{z - \zeta} =: \sum_{\nu=0}^{n-1} b_\nu z^\nu,$$

ainsi que  $b_{-1} := b_n := 0$ . Développer le membre de droite de cette définition et égaliser les coefficients donne

$$a_\nu = -\zeta b_\nu + b_{\nu-1}, \quad \nu = 0, \dots, n.$$

Ainsi les  $b_j$  peuvent être écrits

$$b_j = -\sum_{\nu=0}^j \frac{a_\nu}{\zeta^{j-\nu+1}} \quad (j = 0, \dots, n-2),$$

et alors il vient, pour  $\mu = 0, \dots, m-1$ ,

$$|b_\mu| \leq \sum_{\nu=0}^{\mu} \frac{|a_\nu|}{2^{\mu-\nu+1}} \leq \sum_{l=1}^{\mu+1} 2^{-l} < 1.$$

En répétant le processus pour chacun des zéros concernés, nous aboutissons au polynôme de degré  $m$

$$h(z) := \sum_{\nu=0}^m c_\nu z^\nu := \frac{f(z)}{(z - z_{m+1}) \cdots (z - z_n)},$$

où  $c_m = 1$ , et  $|c_\mu| < 1$  pour  $\mu = 0, \dots, m-1$ .

Maintenant, considérons un zéro  $w$  de  $h$ . Alors

$$w^m = -\sum_{\nu=0}^{m-1} c_\nu w^\nu. \quad (4.3)$$

Ainsi, si nous supposons  $|w| > 1$ , il vient, en divisant d'abord (4.3) par  $w^{m-1}$  puis en utilisant l'inégalité triangulaire

$$|w| \leq \sum_{\nu=0}^{m-1} |c_\nu| |w|^{\nu-m+1} \leq \sum_{l=0}^{\infty} |w|^{-l} = \frac{1}{1 - |w|^{-1}}.$$

Puisque  $0 < \frac{1}{|w|} < 1$  nous pouvons alors faire le calcul suivant,

$$\begin{aligned} |w| &\leq \frac{1}{1 - |w|^{-1}} \\ &\iff |w|(1 - |w|^{-1}) \leq 1 \\ &\iff |w| - 1 \leq 1 \\ &\iff |w| \leq 2, \end{aligned}$$

c'est-à-dire que dans tous les cas,  $|z_m| \leq 2$ , et c'était le résultat souhaité.  $\square$

La preuve du théorème se fait maintenant en quelques lignes.

*Démonstration du théorème.*

Posons

$$\lambda := \max_{0 \leq \mu \leq m-1} |a_\mu|^{\frac{1}{n-\mu}}.$$

Si  $\lambda = 0$ , le théorème est trivial. Supposons donc  $\lambda > 0$ . Alors le polynôme unitaire

$$\phi(z) := \lambda^{-n} f(\lambda z)$$

est bien tel que dans le lemme, c'est-à-dire que les modules de ses  $m$  premiers coefficients ne dépassent pas 1. Ledit lemme s'applique donc, et alors  $\phi$  a au moins  $m$  zéros de modules ne dépassant pas 2. Il suit le résultat, à savoir que  $f$  doit avoir au moins  $m$  zéros de module ne dépassant pas  $2\lambda$ .

□

Le résultat suivant dit que les coefficients d'un polynôme dépendent continûment de ses zéros. Cela semble assez naturel, mais maintenant nous précisons exactement ce que ça veut dire.

**Théorème 4.7** (de continuité). *Soit  $f(z) := \sum_{\nu=0}^n a_\nu z^\nu = \prod_{j=1}^k (z - z_j)^{m_j}$  ( $\sum m_j = n$ ), un polynôme unitaire de degré  $n$ , avec  $k$  zéros distincts  $z_1, z_2, \dots, z_k$ , de multiplicités respectives  $m_1, \dots, m_k$ . Alors étant donné*

$$0 < \epsilon < \frac{1}{2} \min_{1 \leq i < j \leq k} |z_i - z_j|,$$

*il existe un  $\delta > 0$  tel que tout polynôme  $g(z) = \sum_{\nu=0}^n b_\nu z^\nu$ , avec  $|b_\nu - a_\nu| < \delta$  pour  $\nu = 1, \dots, n-1$ , a exactement  $m_j$  zéros dans le disque  $D_j$  de centre  $z_j$  et de rayon  $\epsilon$ , pour chaque  $j$ .*

*Démonstration.* Écrivons

$$f(z + z_j) = \sum_{\nu=0}^n a_\nu^{[j]} z^\nu \quad \text{et} \quad g(z + z_j) =: g_j(z) = \sum_{\nu=0}^n b_\nu^{[j]} z^\nu,$$

où

$$a_\nu^{[j]} = \sum_{k=\nu}^n \binom{k}{\nu} a_k z_j^{k-\nu} \quad \text{et} \quad b_\nu^{[j]} = \sum_{k=\nu}^n \binom{k}{\nu} b_k z_j^{k-\nu}$$

(incluant  $a_n^{[j]} = a_n = b_n = b_n^{[j]} = 1$ ). Comme  $z_j$  est un zéro de multiplicité  $m_j$  de  $f$ ,  $a_\mu^{[j]} = 0$  pour  $\mu = 0, 1, \dots, m_j - 1$ . Nous pouvons donc écrire, pour ces mêmes  $\mu$ ,

$$b_\mu^{[j]} = b_\mu^{[j]} - a_\mu^{[j]} = \sum_{k=\mu}^{n-1} \binom{k}{\mu} (b_k - a_k) z_j^{k-\mu}.$$

En appliquant le théorème précédent à  $g_j$  avec  $m = m_j$  et en posant

$$\rho_j := 2 \max_{0 \leq \mu \leq m_j - 1} \left( \sum_{k=\mu}^{n-1} \binom{k}{\mu} |b_k - a_k| |z_j|^{k-\mu} \right)^{\frac{1}{n-\mu}},$$

nous trouvons que  $D_j := \{z \in \mathbb{C} : |z - z_j| \leq \rho_j\}$  contient au moins  $m_j$  zéros de  $g$ . En prenant  $\delta$  assez petit, les  $\rho_j$  ( $j = 1, \dots, k$ ), peuvent être tous plus petits qu' $\epsilon$ , et alors les  $D_j$  sont disjoints. Ainsi, nous pouvons conclure que chaque  $D_j$  contient exactement  $m_j$  zéros de  $g$ .  $\square$

Des théorèmes classiques d'analyse complexe peuvent être déduits de ce dernier théorème de continuité dans le cas particulier, bien sûr, des polynômes. Ainsi nous verrons, pour finir cette section, d'abord une variante du principe du maximum (dont la démonstration se fera principalement dans le prochain lemme), ensuite le théorème de Rouché et finalement celui de Hurwitz.

**Lemme 4.8.** *Soient  $f$  et  $g$ , deux polynômes n'ayant aucun zéro en commun et tels que  $|f(x)| \leq |g(x)|$  sur  $\mathbb{R}$ . Supposons de plus que chaque zéro  $z_i$  de  $g$  soit dans le demi-plan supérieur  $P_S$  (c'est-à-dire  $\text{Im}(z_i) \geq 0$ ). Alors (i)  $|f(z)| \leq |g(z)|$  sur le demi-plan inférieur  $P_I$  (c'est-à-dire  $\text{Im}(z_i) \leq 0$ ) et (ii) l'égalité est atteinte si et seulement si  $f(z) \equiv e^{i\theta} g(z)$  pour un  $\theta \in [0, 2\pi)$ .*

*Démonstration.* Il se déduit immédiatement des hypothèses que  $|g(x)| > 0$  sur  $\mathbb{R}$ . Il s'en déduit de plus que  $\deg(f) \leq \deg(g) =: n$ . En effet, dans le cas contraire, lorsque  $z$  tendrait vers l'infini sur l'axe réel, nous obtiendrions pour sûr qu'à partir d'un certain point  $|f(x)| \geq |g(x)|$ , ce qui est évidemment une contradiction d'une des hypothèses. Par la même logique, en écrivant

$$f(z) = \sum_{\nu=0}^n a_\nu z^\nu \quad g(z) = \sum_{\nu=0}^n b_\nu z^\nu,$$

il suit que  $|a_n| \leq |b_n|$  (avec potentiellement  $a_n = 0$ ).

Ceci étant dit, il suffit de montrer le résultat suivant. Sous les mêmes hypothèses, si  $\zeta$  est tel que  $\text{Im}(\zeta) < 0$ , alors

$$|f(\zeta)| \geq |g(\zeta)| \implies f(z) \equiv e^{i\theta} g(z)$$

pour un  $\theta \in [0, 2\pi)$

Ce résultat implique en effet le lemme car il montre d'une part que l'inégalité stricte ( $|f| > |g|$ ) est impossible, donc que (i) tient, et d'autre part que si l'égalité est atteinte en un point, alors elle est atteinte partout, ce qui donne une des implications de (ii), l'autre étant triviale.

Maintenant, pour montrer l'implication précédente, il nous suffira d'obtenir que le polynôme

$$h(z) := g(z)f(\zeta) - f(z)g(\zeta)$$

est identiquement nul (nous verrons pourquoi tout à l'heure, montrons-le d'abord dans la boîte suivante).

---

**Boîte 1** Supposons le contraire, c'est-à-dire  $h \neq 0$ , et considérons donc un point  $\lambda \in \mathbb{R}$  tel que  $h(\lambda) \neq 0$  (qui existe par le principe d'identité, que nous tenons pour acquis ici). Considérons aussi la famille de polynômes en  $z$

$$F(t, z) := \frac{z^n [g(z^{-1} + \lambda)f(\zeta) - tf(z^{-1} + \lambda)g(\zeta)]}{g(\lambda)f(\zeta) - tf(\lambda)g(\zeta)}, \quad (4.4)$$

où  $t \in [0, 1]$ .

Lorsque  $t$  est fixé, observons que  $F$  est un polynôme unitaire de degré  $n$ , ce qui se voit relativement bien, et que le dénominateur dans l'équation est non nul. (En effet, nous savons que  $|f(\zeta)| \geq |g(\zeta)|$  et que  $|g(\lambda)| \geq |f(\lambda)|$  (par hypothèse, puisque  $\lambda \in \mathbb{R}$ ) et donc  $|f(\zeta)g(\lambda)| \geq |g(\zeta)f(\lambda)| > t|g(\zeta)f(\lambda)|$  pour  $t \in [0, 1)$ . Si  $t = 1$ , le dénominateur est simplement  $h(\lambda)$ , supposé non nul.) De plus, nous affirmons que

$$\phi : z \mapsto \frac{1}{z - \lambda} \quad \text{et} \quad \phi^{-1} : z \mapsto z^{-1} + \lambda$$

sont des bijections de  $\mathbb{C} \cup \{\infty\}$ , qui envoient l'ensemble  $\mathbb{R} \cup \{\infty\}$  sur lui-même, le demi-plan supérieur  $P_S$  sur le demi-plan inférieur  $P_I$  et vice-versa.

Maintenant, les zéros du polynôme  $F(0, z)$  étant ceux de  $g$  passés dans  $\phi$  (c'est-à-dire que si nous notons  $z_i$  les zéros de  $g$ , alors les  $\phi(z_i)$  seront zéros de  $F$ ), ils sont tous dans  $P_I$ . Par ailleurs  $F(1, z)$  s'annule évidemment en  $z = (\zeta - \lambda)^{-1}$ , un point de  $P_S \setminus \mathbb{R}$ . Par le théorème de continuité, cela signifie que lorsque  $t$  passe de 0 à 1, un des zéros de  $F$  traverse l'axe réel. Il existe donc un  $t' \in [0, 1)$  et un  $x' \in \mathbb{R}$  tels que  $F(t', x') = 0$ . Il se distingue alors deux cas.

(i) Si  $x' = 0$ , (4.4) devient

$$0 = \frac{b_n f(\zeta) - t' a_n g(\zeta)}{g(\lambda) f(\zeta) - t' f(\lambda) g(\zeta)},$$

où  $a_n$  et  $b_n$  sont tels que définis plus haut. En brassant, nous obtenons

$$t' a_n g(\zeta) = b_n f(\zeta),$$

d'où

$$|a_n g(\zeta)| > |b_n f(\zeta)| \geq |b_n g(\zeta)|,$$

d'où finalement  $|a_n| > |b_n|$ , une contradiction.

(ii) Si  $x' \neq 0$ , (4.4) donne de manière analogue

$$0 = g((x')^{-1} + \lambda) f(\zeta) - t' f((x')^{-1} + \lambda) g(\zeta),$$

ce qui implique

$$|g((x')^{-1} + \lambda)| < |f((x')^{-1} + \lambda)|,$$

encore une contradiction.

---

Rappelons que nous supposons  $|f(\zeta)| \geq |g(\zeta)|$  (A). Nous avons maintenant  $h(z) := g(z)f(\zeta) - f(z)g(\zeta) \equiv 0$  (B). Montrons tel que promis que tout cela implique que  $g(z) = e^{i\theta}f(z)$ . D'abord, de (B) nous déduisons

$$g(z)f(\zeta) \equiv f(z)g(\zeta) \quad (\text{B}')$$

$$|g(z)||f(\zeta)| \equiv |f(z)||g(\zeta)| \quad (\text{B}'')$$

De (A) et (B''), il suit  $|g(z)| < |f(z)|$ , pour chaque  $z \in \mathbb{C}$ , donc en particulier pour  $\zeta$ . Nous obtenons  $|g(\zeta)| = |f(\zeta)|$ , ou encore  $g(\zeta) = e^{i\theta}f(\zeta)$  pour un  $\theta \in [0, 2\pi)$  (A2). Remplaçons (A2) dans (B'). Cela donne (rappelons que  $f(\zeta) \neq 0$ ),

$$\begin{aligned} g(z)f(\zeta) &\equiv f(z)e^{i\theta}f(\zeta) \\ g(z) &\equiv f(z)e^{i\theta}, \end{aligned}$$

ce qui était le résultat recherché. □

Dans cette démonstration, l'axe réel aurait aussi bien pu être remplacé par une courbe de Jordan, comme dans le théorème de Rouché qui suivra.

**Théorème 4.9** (variante du principe du maximum). *Considérons deux polynômes  $f$  et  $g$  tels que  $|f(x)| \leq |g(x)|$  sur  $\mathbb{R}$ , et supposons que chaque zéro  $z_j$  de  $g$  soit tel que  $\text{Im}(z_j) \geq 0$ . Alors  $|f(z)| \leq |g(z)|$  pour l'ensemble des  $z$  du demi-plan inférieur  $\text{Im}(z) \leq 0$ . De plus, l'égalité est atteinte dans ce demi-plan si et seulement si  $f(z) = e^{i\theta}g(z)$  pour un  $\theta \in [0, 2\pi)$ .*

*Démonstration.* Appelons  $d = d(x)$  le plus grand commun diviseur unitaire de  $f$  et  $g$  et écrivons  $f = df_1$  et  $g = dg_1$ . Nous avons maintenant  $|f_1(x)| \leq |g_1(x)|$  sur  $\mathbb{R}$ , sauf possiblement aux zéros réels de  $d$ .

Advenant qu'ils existent, appelons-en un  $y$  et supposons que  $f_1(y) > g_1(y)$ . Comme  $d$  est un polynôme donc une fonction continue, il doit exister un intervalle  $I := \{x : |x - y| < \epsilon\}$  tel que  $|f_1(x)| > |g_1(x)|$  sur  $I$ . Or,  $d$  étant un polynôme, ses zéros sont isolés il doit donc exister un  $x_0 \in I$  qui ne soit pas un zéro de  $d$ ; c'est une contradiction. Nous avons donc bien  $|f_1(x)| \leq |g_1(x)|$  sur  $\mathbb{R}$  tout entier.

Comme de plus les polynômes  $f_1$  et  $g_1$  n'ont plus de zéros communs, ils respectent les hypothèses du lemme; celui-ci nous donne alors que (i)  $|f_1(z)| \leq |g_1(z)|$  sur  $P_I$  et (ii) l'égalité

est atteinte si et seulement si  $f_1(z) \equiv e^{i\theta}g_1$  pour un  $\theta \in \mathbb{R}$ . Les conclusions précédentes ne sont pas affectées par la multiplication par  $|d|$  et  $d$ . Nous pouvons en effet déduire que (i)  $|d||f_1(z)| \leq |d||g_1(z)|$  sur  $P_I$  et (ii) l'égalité est atteinte si et seulement si  $df_1(z) \equiv e^{i\theta}dg_1$  pour un  $\theta \in \mathbb{R}$ , ce qui est une réécriture de la conclusion du théorème. □

**Théorème 4.10** (de Rouché). *Soit  $\gamma$ , une courbe de Jordan dans  $\mathbb{C}$ . Considérons  $f$  et  $g$ , deux polynômes, définis sur  $\gamma \cup \text{int}(\gamma)$ , et tels que  $|g(z)| < |f(z)|$  sur  $\gamma$ . Alors  $f$  et  $f + g$  ont le même nombre de zéros dans  $\text{int}(\gamma)$ , en comptant leur multiplicité.*

*Démonstration.* Le raisonnement utilisé ici est similaire à celui qui est utilisé dans la démonstration du lemme 4.8.

Appelons  $\deg(f) =: n$ . Soit  $\lambda \in \gamma$  et posons  $\phi : z \mapsto \frac{1}{z-\lambda}$ . Alors, de manière analogue à ce qui a déjà été vu,

$$F(t, z) := \frac{z^n [f(z^{-1} + \lambda) + tg(z^{-1} + \lambda)]}{f(\lambda) + tg(\lambda)}$$

est un polynôme unitaire en  $z$ , non nul pour toute paire  $(t, z) \in [0, 1] \times \phi(\gamma)$ . Conséquemment, quand  $t$  passe de 0 à 1, aucun des zéros de  $F(t, z)$  (vu comme un polynôme en  $z$  pour chaque  $t$ ) ne peut traverser  $\phi(\gamma)$ . Il suit que  $F(0, z)$  et  $F(1, z)$  doivent avoir le même nombre de zéros dans  $\phi(\text{int}(\gamma))$ , ce qui donne précisément que  $f$  et  $f + g$  ont le même nombre de zéros dans  $\text{int}(\gamma)$ . □

**Remarque 4.11.** Dans le théorème suivant, la notation  $f_n \xrightarrow{u} f$  est utilisée pour signifier qu'une suite de fonctions  $f_n$  converge uniformément vers une fonction  $f$  lorsque  $n \rightarrow \infty$ .

**Théorème 4.12** (de Hurwitz). *Soit  $(f_n)_{n \in \mathbb{N}}$  une suite de polynômes définis dans une région  $\Omega \subseteq \mathbb{C}$ . Supposons que  $f_n \xrightarrow{u} f \not\equiv 0$  sur tout sous-ensemble compact de  $\Omega$ . Alors  $\zeta \in \Omega$  est un zéro de  $f$  de multiplicité  $m$  si et seulement s'il existe un voisinage  $V \subset \Omega$  de  $\zeta$  tel que dans tous les disques  $D(\zeta, \epsilon) \subseteq V$  avec  $\epsilon > 0$ , il existe un  $N(\epsilon)$  tel que chaque  $f_n$  avec  $n \geq N(\epsilon)$  a exactement  $m$  zéros, en comptant la multiplicité.*

*Démonstration.* Considérons un point quelconque  $\zeta \in \Omega$ . Comme  $f \not\equiv 0$  est holomorphe, ses zéros sont isolés, et donc il existe un disque fermé  $V \subseteq \Omega$  tel que  $f(z) \neq 0$  sur  $V \setminus \{\zeta\}$ . Maintenant, si pour  $\epsilon > 0$  nous avons  $D(\zeta, \epsilon) \subseteq V$ , alors  $\min_{|z-\zeta|=\epsilon} |f(z)| =: c > 0$ .

Par ailleurs,  $\max_{|z-\zeta|=\epsilon} |f(z) - f_n(z)| < c$  pour  $n$  assez grand. En posant  $g(z) := f_n(z) - f(z)$ , nous pouvons appliquer le théorème précédent avec les fonctions  $f$  et  $g$ , et  $\gamma$  le cercle  $C := |z - \zeta| = \epsilon$ . Il suit que  $f_n$  et  $f$  ont le même nombre de zéros dans  $\text{int}(C)$ , ce qui était la conclusion recherchée. □

### 4.3 Classification des points critiques et interprétation physique de ceux de type 2

Cette section se veut une introduction à la section suivante. Nous verrons que les points critiques de polynômes ne sont pas toujours que des objets mathématiques abstraits.

Considérons un polynôme unitaire  $f$ . Celui-ci peut être représenté par

$$f(z) = \prod_{j=1}^k (z - z_j)^{m_j},$$

où  $\sum_j m_j = n$  et où les  $z_j$  sont distincts. Les points critiques de ce polynôme peuvent être classés en deux catégories. Si  $m_j > 1$ , alors nous voyons facilement que  $z_j$  sera un point critique d'ordre  $m_j - 1$  de  $f$  en plus d'en être un zéro. Un tel point critique est dit de type 1. Nous pouvons maintenant constater que  $f$  aura exactement  $n - k$  points critiques de ce type. Comme  $f'$  doit être de degré  $n - 1$ , il en reste  $k - 1$ , qui sont dits de type 2. (Remarquons que nous avons montré précédemment que tous les points critiques des polynômes orthogonaux réels sont de type 2.) Comme ces derniers ne sont pas des zéros de  $f$ , ce seront des zéros de la fonction rationnelle  $f'/f$ , définie sur  $\mathbb{C} \setminus \{z_i : i = 1, \dots, k\}$ , qui peut être exprimée

$$\frac{f'(z)}{f(z)} = \sum_{j=1}^k \frac{m_j}{z - z_j}.$$

Pour des raisons évidentes, le quotient  $f'/f$  est appelé la *dérivée logarithmique* de  $f$ .

Gauss a donné l'interprétation suivante des points critiques de type 2. Nous plaçons  $k$  objets  $O_j$  de « masses »  $m_j$  dans le plan euclidien  $\mathbb{R}^2$ , aux points respectifs  $\vec{x}_j = (x_j, y_j)$ . Supposons que chaque objet  $O_j$  génère une force attractive  $\vec{F}_j$  qui soit proportionnelle à la masse  $m_j$  et à la réciproque de la distance euclidienne. Un nouvel objet  $O$ , quelconque, est placé en  $\vec{x} := (x, y)$  dans le même plan. La force exercée par chacun des autres objets sur celui-ci s'écrit alors

$$\vec{F}_j(\vec{x}) = -cm_j \frac{\vec{x} - \vec{x}_j}{|\vec{x} - \vec{x}_j|^2}.$$

En identifiant  $(x, y) \equiv z := x + iy$  et, de manière analogue,  $\vec{x}_j$  à  $z_j$ , nous obtenons que la force totale exercée en  $\vec{x}$  s'écrit

$$\vec{F}(\vec{x}) := \sum_{j=1}^k \vec{F}_j(\vec{x}) = -c \sum_{j=1}^k \frac{m_j}{\bar{z} - \bar{z}_j} = -c \overline{\left( \frac{f'(z)}{f(z)} \right)}.$$

Les points critiques de type 2 étant précisément ceux qui annulent le dernier membre de cette équation, nous en déduisons qu'ils représentent les points d'équilibre du champ de force que nos zéros engendraient au départ.

## 4.4 Théorème de Gauss-Lucas et une réciproque

Le théorème de Gauss-Lucas est central dans l'étude de la théorie analytique des polynômes. Nous en présenterons ici une démonstration relativement simple. Tel qu'annoncé, une réciproque de ce théorème, récemment démontrée, sera ensuite donnée. Cela terminera ce chapitre, qui se voulait une légère introduction à l'étude de certaines propriétés des polynômes analytiques.

Introduisons maintenant le concept central pour la suite.

**Définition 4.13.** L'enveloppe convexe d'un ensemble  $E$  est le plus petit ensemble convexe contenant  $E$  (au sens de l'inclusion). Nous le notons  $\text{conv}(E)$ .

Rappelons qu'un ensemble  $E$  est convexe si toute combinaison linéaire convexe d'éléments de  $E$  est encore dans  $E$ , c'est-à-dire qu'étant donnés  $\{z_\nu\}_{\nu=1,\dots,n} \subseteq E$  et  $\{\alpha_\nu\}_{\nu=1,\dots,n} \subseteq [0, 1]$  tels que  $\sum_{\nu=1}^n \alpha_\nu = 1$ , nous avons que

$$\sum_{\nu=1}^n \alpha_\nu z_\nu \in E.$$

La notation suivante sera aussi utile. Si nous appelons  $Z(f)$  l'ensemble des zéros de  $f$ ,  $\mathcal{K}(f) := \text{conv}(Z(f))$ , sera l'enveloppe convexe des zéros de  $f$ . Nous voici donc prêts pour le théorème principal de cette section.

**Théorème 4.14** (Gauss-Lucas). *Soit  $f$  un polynôme. Alors  $\mathcal{K}(f') \subseteq \mathcal{K}(f)$ .*

*Démonstration.* Montrons qu'un point critique donné  $\zeta$  d'un polynôme  $f$  de degré  $n$  peut s'exprimer comme une combinaison linéaire convexe des zéros  $z_i$  ( $i = 1, \dots, n$ ) de  $f$ , c'est-à-dire qu'il existe des  $\alpha_\nu$  tels que

$$\zeta = \sum_{\nu=1}^n \alpha_\nu z_\nu,$$

où  $\alpha_\nu \geq 0$  pour chaque  $\nu$  et  $\sum_{\nu=1}^n \alpha_\nu = 1$ . Il y a deux cas possibles.

Si  $\zeta$  est de type 1,  $\zeta = z_\mu$  pour un  $\mu$ , et il suffit donc de prendre  $\alpha_\mu = 1$  et  $\alpha_\nu = 0$  pour  $\nu \neq \mu$ .

Si  $\zeta$  est de type 2, nous avons vu qu'il doit être un zéro de la dérivée logarithmique de  $f$ . Alors

$$0 = \sum_{\nu=1}^n \frac{1}{\zeta - z_\nu} = \sum_{\nu=1}^n \frac{1}{\bar{\zeta} - \bar{z}_\nu},$$

donc

$$\sum_{\nu=1}^n \frac{\zeta - z_\nu}{|\zeta - z_\nu|^2} = 0.$$

Prendre

$$\alpha_\nu := \frac{|\zeta - z_\nu|^{-2}}{\sum_{\mu=1}^n |\zeta - z_\mu|^{-2}} \quad \nu = 1, \dots, n$$

mène donc à la conclusion recherchée. En effet

$$\begin{aligned} 0 &= \sum_{\nu=1}^n \frac{\zeta - z_\nu}{|\zeta - z_\nu|^2} = \sum_{\nu=1}^n \frac{\zeta}{|\zeta - z_\nu|^2} - \sum_{\nu=1}^n \frac{z_\nu}{|\zeta - z_\nu|^2} \\ \implies \zeta \sum_{\nu=1}^n \frac{1}{|\zeta - z_\nu|^2} &= \sum_{\nu=1}^n \frac{z_\nu}{|\zeta - z_\nu|^2}, \end{aligned}$$

et alors

$$\begin{aligned} \zeta &= \frac{\sum_{\nu=1}^n |\zeta - z_\nu|^{-2} z_\nu}{\sum_{\nu=1}^n |\zeta - z_\nu|^{-2}} \\ &= \sum_{\nu=1}^n \frac{|\zeta - z_\nu|^{-2} z_\nu}{\sum_{\mu=1}^n |\zeta - z_\mu|^{-2}} \\ &= \sum_{\nu=1}^n \alpha_\nu z_\nu. \end{aligned}$$

De plus

$$\sum_{\nu=1}^n \alpha_\nu = \frac{\sum_{\nu=1}^n |\zeta - z_\nu|^{-2}}{\sum_{\nu=1}^n |\zeta - z_\nu|^{-2}} = 1.$$

□

Le théorème de Gauss-Lucas énonce une propriété de l'enveloppe convexe des polynômes. Nous avons mentionné que le résultat suivant en était une réciproque. En effet, il caractérise tout les opérateurs ayant ladite propriété sur l'enveloppe convexe des polynômes. Nous notons  $\mathcal{P}_n$  les polynômes à coefficients complexes de degré au plus  $n$ . Nous aurons aussi besoin des définitions suivantes.

**Définition 4.15.**

- (i) Soit un ensemble  $E \subseteq \mathbb{C}$ . Le diamètre de  $E$  est le nombre

$$\text{diam}(E) := \sup_{a,b \in E} |a - b|.$$

Nous posons de plus  $\text{diam}(\emptyset) := 0$ . Avec cette convention, le diamètre d'un ensemble est toujours positif ou nul, c'est-à-dire  $\text{diam}(E) \geq 0$  quel que soit  $E$ . Pour fin de notation, étant donnée une fonction complexe  $f$  quelconque, posons  $\mathcal{D}(f) := \text{diam}(Z(f))$ .

- (ii) Soit  $P \in \mathcal{P}_n$ , plus  $n$  et  $0 \leq k \leq n - 2$ . Posons

$$d_{n,k}(P) := \min\{d \in \mathbb{R}^+ : \mathcal{D}(P^{(k)}) \leq d\mathcal{D}(P)\}$$

et

$$d_{n,k} := \max_{P: k < \deg(P) \leq n} (d_{n,k}(P)).$$

L'existence de telles constantes est bien entendu garantie par le théorème de Gauss-Lucas.

**Théorème 4.16.** Soit  $L : \mathcal{P}_n \rightarrow \mathbb{C}[z]$  (avec  $n \geq 1$ ), un opérateur linéaire tel que  $L[P] \neq 0$  pour chaque polynôme  $P$ . Alors  $L$  a la propriété que  $\mathcal{D}(L[P]) \leq \mathcal{D}(P)$  si et seulement s'il a une des formes suivantes :

(i)  $L = zl_1 + l_2$ , où  $l_1$  et  $l_2$  sont des fonctionnelles linéaires ;

(ii)  $L = (z - c)^m l_3$ , pour un  $m \geq 2$ , un  $c \in \mathbb{C}$  et une fonctionnelle linéaire non nulle  $l_3$  ;

(iii)  $L[P] = c(P \circ p)^{(k)}$ , pour un  $c \in \mathbb{C} \setminus \{0\}$ , un  $k$  tel que  $0 \leq k \leq n - 2$  et  $p := az + b$ , où  $a, b \in \mathbb{C}$  et  $|a| \geq d_{n,k}$ .

*Démonstration.*

( $\Leftarrow$ ) Il faut montrer que chacune des trois formes données de  $L$  implique que  $\mathcal{D}(L[P]) \leq \mathcal{D}(P)$ .

(i)  $L = zl_1 + l_2$ .

D'abord, il est évident que dans ce cas  $\mathcal{D}(L[P]) = 0$ , puisque  $L[P]$  a au plus un zéro et que le diamètre d'un singleton (et de l'ensemble vide) est toujours nul. Or,  $\mathcal{D}(P) \geq 0$  trivialement.

(ii)  $L = (z - c)^m l_3$ .

Comme en (i),  $\mathcal{D}(L[P]) = 0$  clairement, et le résultat suit.

(iii)  $L[P] = c(P \circ p)^{(k)}$ .

Calculons

$$L[P] = c(P(az + b))^{(k)} = ca^k P^{(k)}(az + b).$$

Ainsi

$$Z(L[P]) = Z(ca^k P^{(k)}(az + b)) = Z(P^{(k)}(az + b)),$$

Donc

$$\begin{aligned} \mathcal{D}(L[P]) &= \mathcal{D}(P^{(k)}(az + b)) \\ &= \frac{1}{|a|} \mathcal{D}(P^{(k)}) \\ &\leq \frac{d_{n,k}}{|a|} \mathcal{D}(P) \\ &\leq \mathcal{D}(P). \end{aligned}$$

( $\Rightarrow$ ) Au cours de la preuve, des « sous-résultats » seront mis en boîte pour une lecture plus aisée. Commençons par quelques préliminaires qui nous serviront pour la suite.

Si l'opérateur  $L$  est une fonctionnelle linéaire, alors il est de la première forme. Supposons donc que  $L$  ne soit pas une fonctionnelle (mais reste un opérateur linéaire), dans le sens où

il existe  $r$  tel que  $0 \leq r \leq n$  et  $\deg(L[z^r]) \geq 1$ . Supposons  $r$  minimal par rapport à cette propriété. Soient de plus  $s$  tel que  $r \leq s \leq n$  et  $\alpha \in \mathbb{C}$ . Alors il existe des nombres  $a_s, b_s \in \mathbb{R}$  et  $q_s \in \mathbb{N}$  tels que

$$(a_s(\alpha)z + b_s(\alpha))^{q_s(\alpha)} = L[(z + \alpha)^s] = \sum_{j=0}^s \binom{s}{j} \alpha^j L[z^{s-j}]. \quad (4.5)$$

La première égalité provient de l'hypothèse  $\mathcal{D}(L[P]) \leq \mathcal{D}(P)$ . En effet,  $(z + \alpha)^s$  n'a qu'un seul zéro, donc  $\mathcal{D}((z + \alpha)^s) = 0$ , et alors  $L[(z + \alpha)^s]$  est forcé d'avoir au plus un zéro, ce qui s'écrit bien en toute généralité comme le membre de gauche de (4.5). La seconde égalité provient quant à elle du développement du binôme et de l'utilisation de la linéarité de  $L$ .

Considérons la somme dans (4.5), et regardons-la comme un polynôme en  $z$ . Le degré de ce polynôme doit être supérieur à 1. De plus, il ne dépend pas de la valeur de  $\alpha$ , sauf possible-ment pour un certain ensemble fini de ces  $\alpha$ , notons-le  $A_s$  pour chaque  $s$  fixé. Cet ensemble contiendrait en fait les zéros du polynôme en  $\alpha$  formant le coefficient directeur du polynôme en  $z$  que définit la somme. Appelons  $U := \cup_{j=r}^n A_j$ . La preuve qui suivra sera valable si  $0 \notin U$ . Si  $0 \in U$ , choisissons  $\epsilon \notin U$ . En posant  $z' := z - \epsilon$  dans ce qui précède et refaisant la construction faite dans ce paragraphe avec  $z$  remplacé par  $z'$ , nous trouverons  $0 \notin U'$ , et la preuve qui suivra s'appliquera à la variable  $z'$ . La preuve suivante est donc générale.

Pour finir les préliminaires, remarquons que  $s_1 > s_0 \implies q_{s_1} > q_{s_0}$ . En effet, avec  $\alpha \notin U$ , la seconde égalité dans (4.5) montre cela clairement.

Nous en venons maintenant au coeur de la démonstration. Exécutons d'abord le calcul suivant. Soit  $\alpha \notin U$ . Alors  $q_s(\alpha) = q_s(0) =: q_s \geq 1$  et  $a_s(0) \neq 0$  et

$$L[(z + \alpha)^r] = (a_r(\alpha)z + b_r(\alpha))^{q_r} = L[z^r] + \sum_{j=1}^r \binom{r}{j} \alpha^j L[z^{r-j}].$$

Or, pour  $t < r$ ,  $L$  agit par hypothèse comme une fonctionnelle linéaire sur  $z^t$ . Il vient

$$(a_r(\alpha)z + b_r(\alpha))^{q_r} = L[z^r] + Q_r(\alpha) = (a_r(0)z + b_r(0))^{q_r} + Q_r(\alpha),$$

où  $Q_r(\alpha)$  est une constante par rapport à  $z$ . En particulier, en développant les deux membres de cette égalité et en égalisant les coefficients de mêmes puissances de  $z$ , il vient

$$\begin{aligned} a_r(\alpha) &= a_r(0) =: a_r, \\ b_r(\alpha) &= b_r(0) =: b_r \\ \text{et } Q_r(\alpha) &= 0. \end{aligned}$$

Traitons maintenant 2 cas.

(I)  $q_r \geq 2$ . Nous avons vu que

$$Q_r(\alpha) = \sum_{j=1}^r \binom{r}{j} \alpha^j L[z^{r-j}] = 0.$$

En regardant maintenant cette somme comme un polynôme en  $\alpha$ , il vient  $L[1] = L[z] = L[z^2] = \dots = L[z^{r-1}] = 0$ .

- Si  $r = n$ , alors pour chaque polynôme  $P_n \in \mathcal{P}_n$

$$L[P_n] = \text{cd}(P_n)L[z^r] = \text{cd}(P_n)a_r^{q_r} \left( z + \frac{b_r}{a_r} \right)^{q_r},$$

et  $L$  est de la forme (ii).

- Pour traiter le cas  $r < n$ , calculons

$$L[(z + \alpha)^{r+1}] = (a_{r+1}(\alpha)z + b_{r+1}(\alpha))^{q_{r+1}} \quad (4.6)$$

$$= L[z^{r+1}] + \alpha(r+1)L[z^r] \quad (4.7)$$

$$= (a_{r+1}(0)z + b_{r+1}(0))^{q_{r+1}} + \alpha(r+1)(a_r z + b_r)^{q_r}. \quad (4.8)$$

Remarquons qu'ici,  $q_r \leq q_{r+1}$ .

---

Boîte 1 Posons  $\frac{-b_r}{a_r}$ , et montrons que

$$(a_{r+1}(0)z + b_{r+1}(0)) = 0$$

par contradiction, c'est-à-dire supposons que  $(a_{r+1}(0)z + b_{r+1}(0))^{q_{r+1}} \neq 0$ . Alors pour  $\alpha \neq 0$ , les équations 4.6 à 4.8 nous donnent l'égalité

$$-\frac{\alpha(r+1)(a_r z + b_r)^{q_r}}{(a_{r+1}(0)z + b_{r+1}(0))^{q_{r+1}}} = 1 - \left( \frac{a_{r+1}(\alpha)z + b_{r+1}(\alpha)}{a_{r+1}(0)z + b_{r+1}(0)} \right)^{q_{r+1}}.$$

Comme  $\frac{-b_r}{a_r}$  est un zéro de multiplicité  $q_r$  du membre de gauche, il en est de même pour le membre de droite. Or, il est relativement facile de se convaincre que ledit membre de droite doit soit s'annuler identiquement, soit avoir des zéros distincts, ce qui contredit le fait que  $q_r \geq 2$ .

---

Ce résultat en boîte donne directement

$$L[z^{r+1}] = (a_{r+1}(0)z + b_{r+1}(0))^{q_{r+1}} = 0$$

Cet argument peut être répété successivement pour chaque  $k$  tel que  $r+1 < k < n$ , et il suit qu'encre dans le cas présent,  $L$  est de la forme (ii) avec  $c = \frac{-b_r}{a_r}$  et  $m = q_r$ .

(II)  $q_r = 1$ .

Supposons que  $L$  ne soit pas de la forme (ii). Soit  $t$  tel que  $q_r = \dots = q_t = 1$  et  $q_{t+1} \geq 2$ .

Alors

$$(a_{t+1}(\alpha)z + b_{t+1}(\alpha))^{q_{t+1}} = (a_{t+1}(0)z + b_{t+1}(0))^{q_{t+1}} + \sum_{j=1}^{t+1} \binom{t+1}{j} \alpha^j L[z^{t+1-j}].$$

---

**Boîte 2** Montrons que nous ne pouvons avoir  $q_{t+1} \geq 3$ , par contradiction. En effet, si  $q_{t+1} \geq 3$ , en comparant les coefficients de  $z^{q_{t+1}}$  puis de  $z^{q_{t+1}-1}$  des deux membres de l'équation précédente, il vient  $a_{t+1}(\alpha) = a_{t+1}(0)$  et  $b_{t+1}(\alpha) = b_{t+1}(0)$ . Cela est déjà une contradiction, puisque pour  $\alpha \notin U$ , le degré de la somme n'est pas nul, donc à plus forte raison la somme n'est pas identiquement nulle.

---

Donc  $q_{t+1} = 2$ .

---

**Boîte 3** Montrons maintenant, toujours par contradiction, que nous ne pouvons avoir  $t > r$ . Supposons-le donc. Pour  $M \in \mathbb{N}$ , appelons

$$P_M(z) := z^{t+1} - Mz^{t-1}.$$

Alors un petit calcul montre que  $\mathcal{D}(P_M) = 2\sqrt{M}$ , et alors

$$\lim_{M \rightarrow \infty} \frac{\mathcal{D}(P_M)}{M} = \lim_{M \rightarrow \infty} \frac{2}{\sqrt{M}} = 0.$$

Par ailleurs,

$$L[P_M] = (a_{t+1}(0)z + b_{t+1}(0))^2 - M(a_{t-1}(0)z + b_{t-1}(0)).$$

Les zéros de ce polynôme sont naturellement

$$\frac{-B \pm \sqrt{B^2 - 4AC}}{2A},$$

avec

$$\begin{cases} A = a_{t+1}^2(0) \\ B = 2a_{t+1}(0)b_{t+1}(0) - Ma_{t-1}(0) \\ C = b_{t+1}^2(0) - Mb_{t-1}(0). \end{cases}$$

Ainsi,  $\mathcal{D}(L[P_M]) = \frac{\sqrt{B^2 - 4AC}}{A}$  et alors

$$\begin{aligned} \lim_{M \rightarrow \infty} \frac{\mathcal{D}(L[P_M])}{M} &= \lim_{M \rightarrow \infty} \frac{\sqrt{B^2 - 4AC}}{AM} \\ &= \lim_{M \rightarrow \infty} \frac{\left(\frac{B^2}{M} - \frac{4AC}{M^2}\right)^{1/2}}{A} \\ &= \frac{a_{t-1}(0)}{a_{t+1}^2(0)} \neq 0. \end{aligned}$$

Cela contredit l'hypothèse de départ du théorème, à savoir  $\mathcal{D}(L[P_M]) \leq \mathcal{D}(P_M)$ .

---

Donc  $t = r$  et il vient  $q_{r+1} = 2$ . Ces conditions plus précises en main, nous avons

$$(a_{r+1}(\alpha)z + b_{r+1}(\alpha))^2 = (a_{r+1}(0)z + b_{r+1}(0))^2 + \alpha(r+1)(a_r z + b_r) + Q_{r+1}(\alpha),$$

où  $Q_{r+1}(\alpha) = \sum_{j=2}^{r+1} \binom{r+1}{j} \alpha^j L[z^{r+1-j}]$  est une constante par rapport à  $z$ . Il suit, en appelant  $b_{r+1} := b_{r+1}(0)$ , que

$$\begin{cases} a_{r+1}(\alpha) = a_{r+1}(0) =: a_{r+1}; \\ b_{r+1}(\alpha) = b_{r+1} + \frac{\alpha(r+1)a_r}{2a_{r+1}}; \\ b_{r+1}^2(\alpha) = b_{r+1}^2 + \alpha(r+1)b_r + Q_{r+1}(\alpha). \end{cases}$$

Les deux dernières égalités donnent

$$\frac{\alpha(r+1)a_r b_{r+1}}{a_{r+1}} + \frac{\alpha^2(r+1)^2 a_r^2}{4a_{r+1}^2} = \alpha(r+1)b_r + Q_{r+1}(\alpha),$$

en retranchant un  $b_{r+1}^2$  de chaque côté. Maintenant, en égalisant dans cette dernière égalité les coefficients des mêmes puissances de  $\alpha$ , nous trouvons

$$\begin{cases} \frac{b_{r+1}}{a_{r+1}} = \frac{b_r}{a_r}, \text{ pour } r \geq 1 \\ L[z^{r-1}] = \frac{(r+1)a_r^2}{2a_{r+1}^2} \text{ et} \\ L[z^{r-2}] = \dots = L[1] = 0. \end{cases}$$

Comme  $L$  est un opérateur linéaire, il peut être normalisé de telle sorte que  $\tilde{L}[z^{r-1}] = (r-1)!$ , tout en conservant  $\tilde{L}[z^{r-2}] = \dots = \tilde{L}[1] = 0$ . Si nous appelons maintenant  $a := \frac{a_r}{r!}$ ,  $b := \frac{b_r}{r!}$ , et  $p(z) := az + b$ , ce qui précède mène à  $a_{r+1}^2 = \frac{(r+1)!a^2}{2}$  et  $b_{r+1}^2 = \frac{(r+1)!b^2}{2}$ , et par suite à

$$\tilde{L} = a^{1-r}(P \circ p)^{(r-1)},$$

pour  $P(z) = z^j$  ( $0 \leq j \leq r+1$ ). Nous obtiendrions donc bien la forme (iii) si : (a) le résultat tenait pour  $0 \leq j \leq n$  et (b) si  $|a| \geq d_{n,k}$ . Voici donc les preuves de ces deux chaînons manquants. La première est une récurrence.

(a) Supposons que le résultat tienne pour  $j = 0, \dots, k$ , où  $r+1 \leq k \leq n-1$ . Alors la même preuve que dans ce qui précède peut-être refaite :

- En comparant les coefficients comme dans la boîte 2, nous pouvons montrer que soit  $q_{k+1} = q_k + 1$ , soit  $q_{k+1} = q_k$ .
- Ensuite, nous pouvons montrer comme dans la boîte 3 que le cas  $q_{k+1} = q_k + 1$  est impossible en considérant le polynôme  $z^{k+1} - Mz^{k-1}$ .

Après un calcul élémentaire, il vient  $q_{k+1} = k + 2 - r \geq 3$ , et alors

$$\begin{aligned}\tilde{L}[(z + \alpha)^{k+1}] - \tilde{L}[z^{k+1}] &= (a_{k+1}(\alpha) + b_{k+1}(\alpha))^{q_{k+1}} - (a_{k+1}(0) + b_{k+1}(0))^{q_{k+1}} \\ &= \sum_{j=1}^{k+1} \binom{k+1}{j} \alpha^j \tilde{L}[z^{k+1-j}] \\ &= \tilde{P}_{k+1}(z + \alpha) - \tilde{P}_{k+1}(z),\end{aligned}$$

où  $\tilde{P}_{k+1}(z) := \frac{(k+1)!}{q_{k+1}!} (az + b)^{q_{k+1}}$ . En particulier,  $a_{k+1}(\alpha) = a_{k+1}(0)$ .

---

**Boîte 4** Remarquons maintenant que si  $l \geq 3$  et  $\beta \neq 0$  sont tels que, pour tout  $w \in \mathbb{C}$ , il existe des nombres  $\delta, \gamma$  tels que

$$(w + \beta)^l - w^l = d[(w + \gamma)^l - (w + \delta)^l],$$

alors de deux choses l'une : soit  $(d, \beta, \delta) = (1, \gamma, 0)$ , soit  $(d, \beta, \gamma) = (-1, \delta, 0)$ . Cela provient du fait que si  $\lambda_0, \dots, \lambda_l$  sont des nombres complexes tous différents les uns des autres, alors les  $(w + \lambda_0)^l, \dots, (w + \lambda_l)^l$  forment une base des polynômes de degré  $l$ , donc sont tous linéairement indépendants.

---

De ce résultat en boîte, il vient facilement que

$$\tilde{L}[z^{k+1}] = \tilde{P}_{k+1}(z),$$

ce qui complète l'induction.

(b) L'hypothèse  $\mathcal{D}(L[P]) \leq \mathcal{D}(P)$  force  $|a| \geq d_{n,k}$ . En effet

$$\begin{aligned}Z(L[P]) &= Z(a^{-k}(P \circ p)^{(k)}) \\ &= Z(P \circ p)^{(k)} \\ &= Z(a^k P^{(k)} \circ p) \\ &= Z(P^{(k)} \circ p),\end{aligned}$$

où  $k = r - 1$  et  $p = az + b$ . Donc

$$\begin{aligned}\mathcal{D}((L[P])) &= \mathcal{D}((P^{(k)} \circ p)) \\ &= \frac{1}{|a|} \mathcal{D}(P^{(k)}) \\ &\stackrel{\text{Hyp.}}{\leq} \mathcal{D}(P) \\ \iff \mathcal{D}(P^{(k)}) &\leq |a| \mathcal{D}(P).\end{aligned}$$

Comme  $d_{n,k}$  est le plus petit nombre tel que cette inégalité soit satisfaite, il vient  $|a| \geq d_{n,k}$ .



Ce dernier résultat a été inclus ici simplement pour illustrer que notre sujet, tout ancien qu'il soit, est toujours d'actualité.

# Conclusion

Rappelons que la motivation de ce mémoire était de donner une introduction suffisamment exhaustive et rigoureuse à la théorie des polynômes orthogonaux. Dans ce but, nous avons introduit le sujet par l'intermédiaire des problèmes de Sturm-Liouville, puis nous avons établi les équivalences existantes entre les principales définitions que l'on rencontre pour les polynômes orthogonaux. Nous avons ensuite décortiqué certaines propriétés de base desdits polynômes, et finalement, nous avons abordé la théorie analytique des polynômes.

En cours de route, nous avons ouvert des portes sur d'autres aspects de la recherche sur les polynômes orthogonaux. Nous avons entrevu que le schéma d'Askey classifiait ceux-ci avec les différentes relations qui les relient. Nous aurions aussi pu continuer à étudier la théorie analytique des polynômes, et, qui sait, peut-être que lorsqu'appliqués aux polynômes orthogonaux en particulier, certains résultats généraux gagnent-ils en puissance. Ceci est sans parler des sujets liés que nous n'avons pas abordés du tout, mentionnons la théorie des mesures secondaires, ou encore l'utilisation d'ensembles de polynômes dits « biorthogonaux » en codage et décodage de signaux.

Bref, les objets en apparence simples mais pleins de potentiel que sont les polynômes sont encore étudiés sous plusieurs angles. Il est donc justifié de croire que la théorie continuera encore de se développer et de donner des applications intéressantes.



# Annexe A

## Table de polynômes orthogonaux classiques

### A.1 Polynômes de Legendre

Intervalle fondamental :  $[-1, 1]$ .

Mesure :  $d\mu(x) = dx$ .

Normalisation standard :  $P_n(1) = 1$ .

Coefficient  $A_n$  de la relation de récurrence avec cette normalisation :

$$\frac{2n+1}{n+1}.$$

Carré de la norme du polynôme normalisé :

$$\frac{2}{2n+1}.$$

Équation différentielle :

$$(1-x^2)y'' - 2xy' + n(n+1)y = 0.$$

Les premiers de ces polynômes sont :

$$\begin{aligned}P_0 &= 1, \\P_1 &= x, \\P_2 &= \frac{3}{2}x^2 - \frac{1}{2}, \\P_3 &= \frac{5}{2}x^3 - \frac{3}{2}x.\end{aligned}$$

Les polynômes de Legendre sont un cas particulier ( $\alpha = 1/2$ ) des polynômes de Gegenbauer qui suivent.

## A.2 Polynômes de Gegenbauer (ou ultrasphériques)

Intervalle fondamental :  $[-1, 1]$ .

Mesure :  $d\mu(x) = (1 - x^2)^{\alpha - \frac{1}{2}} dx$ ,  $\alpha > -\frac{1}{2}$ .

Normalisation standard (si  $\alpha \neq 0$ ) :

$$C_n^{(\alpha)}(1) = \frac{\Gamma(n + 2\alpha)}{n! \Gamma(2\alpha)}.$$

Coefficient  $A_n$  de la relation de récurrence avec cette normalisation :

$$\frac{2(n + 2\alpha)}{n + 1}.$$

Carré de la norme du polynôme normalisé :

$$\frac{\pi 2^{1-2\alpha} \Gamma(n + 2\alpha)}{n! (n + \alpha) (\Gamma(\alpha))^2}.$$

Équation différentielle :

$$(1 - x^2)y'' - (2\alpha + 1)xy' + n(n + 2\alpha)y = 0.$$

Les premiers de ces polynômes sont :

$$\begin{aligned} C_0^{(\alpha)} &= 1, \\ C_1^{(\alpha)} &= 2\alpha x, \\ C_2^{(\alpha)} &= -\alpha + 2\alpha(1 + \alpha)x^2, \\ C_3^{(\alpha)} &= -2\alpha(1 + \alpha)x + \frac{4}{3}\alpha(1 + \alpha)(2 + \alpha)x^3. \end{aligned}$$

Les polynômes de Gegenbauer sont un cas particulier ( $\alpha = \beta$ ) des polynômes de Jacobi qui suivent.

## A.3 Polynômes de Jacobi

Intervalle fondamental :  $[-1, 1]$ .

Mesure :  $d\mu(x) = (1 - x)^\alpha (1 + x)^\beta dx$ ,  $\alpha, \beta > -1$ .

Normalisation standard (si  $\alpha \neq 0$ ) :

$$P_n^{(\alpha, \beta)}(1) = \frac{\Gamma(n + 1 + \alpha)}{n! \Gamma(1 + \alpha)}.$$

Coefficient  $A_n$  de la relation de récurrence avec cette normalisation :

$$\frac{(2n+1+\alpha+\beta)(2n+2+\alpha+\beta)}{2(n+1)(n+1+\alpha+\beta)}.$$

Carré de la norme du polynôme normalisé :

$$\frac{\pi 2^{\alpha+\beta+1} \Gamma(n+\alpha+1) \Gamma(n+\beta+1)}{n!(2n+\alpha+\beta+1) \Gamma(n+\alpha+\beta+1)}.$$

Équation différentielle :

$$(1-x^2)y'' - [\beta - \alpha - (\alpha + \beta + 2)x]y' + n(n+1+\alpha+\beta)y = 0.$$

Les premiers de ces polynômes sont :

$$P_0^{(\alpha,\beta)} = 1,$$

$$P_1^{(\alpha,\beta)} = \frac{1}{2} [2(\alpha+1) + (\alpha+\beta+2)(x-1)],$$

$$P_2^{(\alpha,\beta)} = \frac{1}{8} [4(\alpha+1)(\alpha+2) + 4(\alpha+\beta+3)(\alpha+2)(z-1) + (\alpha+\beta+3)(\alpha+\beta+4)(z-1)^2].$$

## A.4 Polynômes de Tchebychev (première espèce)

Intervalle fondamental :  $[-1, 1]$ .

Mesure :  $d\mu(x) = (1-x^2)^{-\frac{1}{2}} dx$ .

Normalisation standard :  $T_n(1) = 1$ .

Coefficient  $A_n$  de la relation de récurrence avec cette normalisation : 2.

Carré de la norme du polynôme normalisé :

$$\begin{cases} \pi & \text{si } n = 0, \\ \pi/2 & \text{si } n \neq 0. \end{cases}.$$

Équation différentielle :

$$(1-x^2)y'' - xy' + n^2y = 0.$$

En résolvant les intégrales appropriées à la main, on peut remarquer que l'orthogonalité des  $T_n$  provient en fait directement de celle des  $\cos nx$ . Cela se concrétise avec la formule suivante, qui sert le plus souvent aux polynômes de Tchebychev,

$$T_n(x) = \cos(n \arccos x).$$

Les premiers de ces polynômes sont :

$$\begin{aligned}T_0 &= 1, \\T_1 &= x, \\T_2 &= 2x^2 - 1, \\T_3 &= 4x^3 - 3x.\end{aligned}$$

## A.5 Polynômes de Tchebychev (seconde espèce)

Intervalle fondamental :  $[-1, 1]$ .

Mesure :  $d\mu(x) = (1 - x^2)^{\frac{1}{2}} dx$ .

Normalisation standard :  $U_n(1) = n + 1$ .

Coefficient  $A_n$  de la relation de récurrence avec cette normalisation : 2.

Carré de la norme du polynôme normalisé :

$$\frac{2}{2n + 1}.$$

Équation différentielle :

$$(1 - x^2)y'' - 3xy' + n(n + 2)y = 0.$$

De manière analogue aux polynômes de première espèce, ceux de deuxième espèce suivent une équation trigonométrique, à savoir

$$U_n(\cos \theta) = \frac{\sin((n + 1)\theta)}{\sin \theta}.$$

Les premiers de ces polynômes sont :

$$\begin{aligned}U_0 &= 1, \\U_1 &= 2x, \\U_2 &= 4x^2 - 1, \\U_3 &= 8x^3 - 4x.\end{aligned}$$

## A.6 Polynômes de Hermite (forme physique)

Intervalle fondamental :  $(-\infty, \infty)$ .

Mesure :  $d\mu(x) = e^{-x^2} dx$ .

Normalisation standard :  $cd(H_n) = 2^n$ .

Coefficient  $A_n$  de la relation de récurrence avec cette normalisation : 2.

Carré de la norme du polynôme normalisé :

$$2^n n! \sqrt{\pi}.$$

Équation différentielle :

$$y'' - 2xy' + 2ny = 0.$$

Les premiers de ces polynômes sont :

$$H_0 = 1,$$

$$H_1 = 2x,$$

$$H_2 = 4x^2 - 2,$$

$$H_3 = 8x^3 - 12x.$$

On dit que les polynômes précédents sont sous forme « physique » par opposition à la forme dite « probabiliste »  $\hat{H}_n$ . La relation entre ces deux formes s'écrit

$$\hat{H}_n(x) = (-1)^n (2^{\frac{n}{2}} n!)^{-1} H_n(2^{-\frac{1}{2}} x).$$

## A.7 Polynômes de Laguerre

Intervalle fondamental :  $(0, \infty)$ .

Mesure :  $d\mu(x) = e^{-x} dx$ .

Normalisation standard :  $cd(L_n) = \frac{(-1)^n}{n!}$ .

Coefficient  $A_n$  de la relation de récurrence avec cette normalisation :  $\frac{-1}{n+1}$ .

Carré de la norme du polynôme normalisé : 1.

Équation différentielle :

$$xy'' + (1-x)y' + ny = 0.$$

Les premiers de ces polynômes sont :

$$L_0 = 1,$$

$$L_1 = -x + 1,$$

$$L_2 = \frac{1}{2}(x^2 - 4x + 2),$$

$$L_3 = \frac{1}{6}(-x^3 + 9x^2 - 18x + 6).$$

Les polynômes de Laguerre se généralisent aux polynômes de Laguerre associés dont la description suit.

## A.8 Polynômes de Laguerre associés

Intervalle fondamental :  $(0, \infty)$ .

Mesure :  $d\mu(x) = x^\alpha e^{-x} dx$ ,  $\alpha > -1$ .

normalisation standard :  $cd(L_n) = \frac{(-1)^n}{n!}$ .

Coefficient  $A_n$  de la relation de récurrence avec cette normalisation :  $\frac{-1}{n+1}$ .

Carré de la norme du polynôme normalisé : 1.

Équation différentielle :

$$xy'' + (\alpha + 1 - x)y' + ny = 0.$$

Les premiers de ces polynômes sont :

$$L_0^\alpha = 1,$$

$$L_1^\alpha = -x + \alpha + 1,$$

$$L_2^\alpha = \frac{1}{2}[x^2 - 2(\alpha + 2)x + (\alpha + 2)(\alpha + 1)],$$

$$L_3^\alpha = \frac{1}{6}[-x^3 + 3(\alpha + 3)x^2 - 3(\alpha + 2)(\alpha + 3)x + (\alpha + 1)(\alpha + 2)(\alpha + 3)].$$

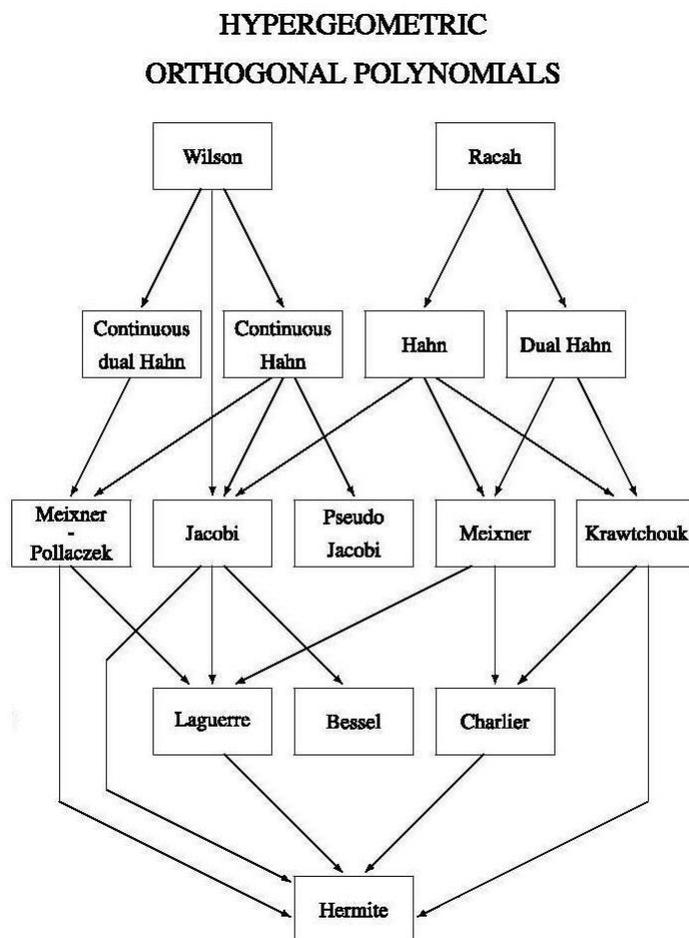
Les polynômes de Hermite peuvent se réduire aux polynômes de Laguerre associés avec  $\alpha = \pm \frac{1}{2}$  à l'aide des formules suivantes (voir (16, p. 106)).

$$H_{2m} = (-1)^m 2^{2m} m! L_m^{(-\frac{1}{2})}(x^2)$$

$$H_{2m+1} = (-1)^m 2^{2m+1} m! x L_m^{(\frac{1}{2})}(x^2).$$

## Annexe B

# Schéma d'Askey



Cette image provient de (8). Les classes de polynômes au bout d'une flèche sont des cas particuliers ou des cas limites de ceux à l'origine de la même flèche.



## Annexe C

# Rappel sur l'interpolation de Lagrange

Étant donnés  $n + 1$  points (distincts)  $(x_0, y_0), (x_1, y_1), \dots, (x_n, y_n)$  du graphe d'une fonction inconnue, nous voulons construire un polynôme de degré  $n$  interpolant ces points (c'est-à-dire passant par tous ces points). L'idée de base de l'interpolation de Lagrange est la suivante. On construit  $n + 1$  polynômes  $L_0, L_1, \dots, L_n$  de degré  $n$ , tels que

$$L_i(x_j) = \begin{cases} 1 & \text{si } i = j, \\ 0 & \text{si } i \neq j. \end{cases}$$

Ensuite, le polynôme d'interpolation recherché est simplement

$$L(x) = \sum_{i=0}^n y_i L_i(x).$$

La méthode la plus simple pour construire de tels  $L_i$  est évidemment de prendre

$$L_i(x) = \prod_{\substack{j=0 \\ j \neq i}}^n \frac{x - x_j}{x_i - x_j}.$$

Maintenant, il est d'un certain intérêt de remarquer que le polynôme satisfaisant les conditions de départ est unique et alors celui que nous avons construit est le seul qui satisfasse lesdites conditions. En effet, si  $P$  et  $Q$  sont deux polynômes de degré  $n$  passant par les  $n + 1$  points  $(x_i, y_i)$ , alors  $P - Q$  est aussi un polynôme de degré  $n$ , mais qui a au moins les  $n + 1$  valeurs  $x_i$  comme racines. Or, on sait bien qu'un polynôme de degré  $n$  ayant  $n + 1$  racines doit être identiquement nul, et donc  $P = Q$ .

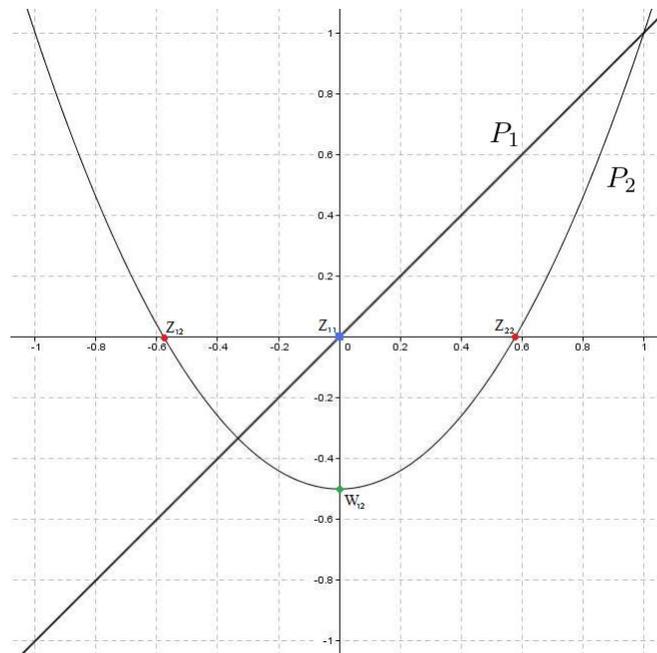


## Annexe D

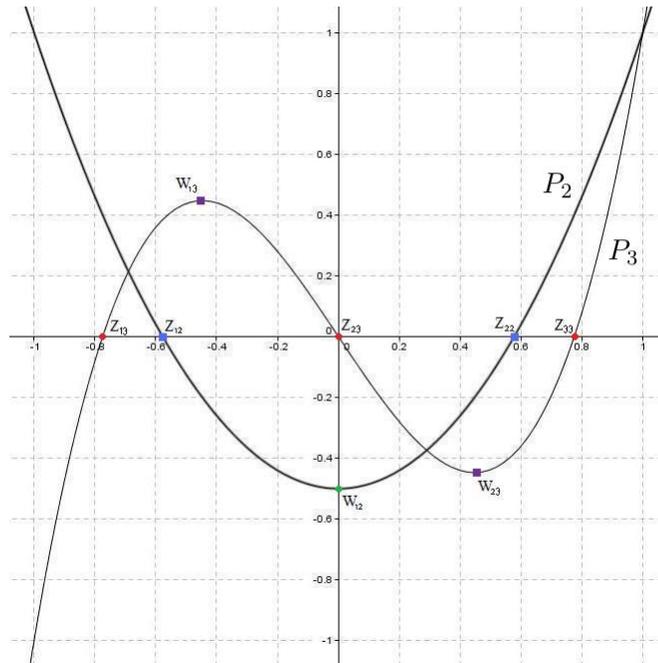
# Zéros et points critiques des polynômes de Legendre

Nous avons consigné ici quelques représentations graphiques de polynômes de Legendre, comparés deux à deux. Le but de faire cela est de donner l'intuition que les théorèmes présentés sur les zéros et les points critiques des polynômes orthogonaux sont naturels, étant données les formes que prennent ces polynômes. Sur ces graphiques, les  $Z_{ij}$  sont des zéros et les  $W_{ij}$  des points critiques.

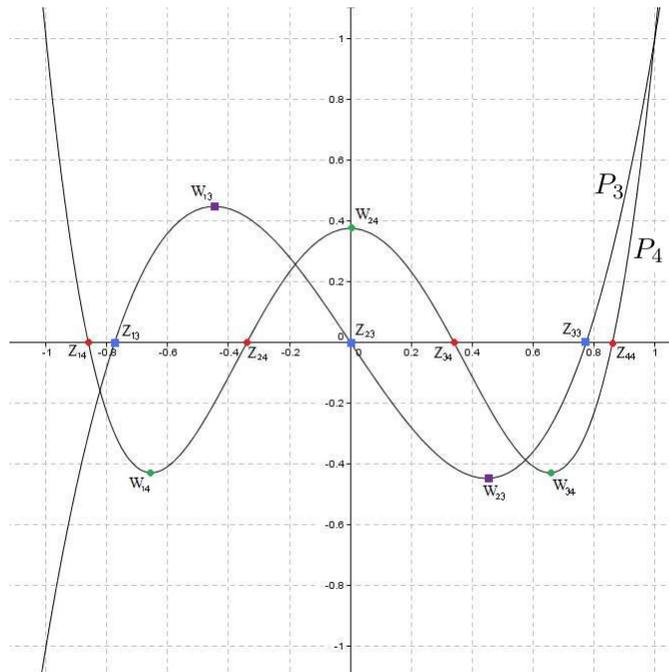
$$P_1 = x, P_2 = \frac{3x^2-1}{2}$$



$$P_2 \text{ et } P_3 = \frac{5x^3 - 3x}{2}$$



$$P_3 \text{ et } P_4 = \frac{35x^4 - 30x^2 + 3}{8}$$



Dans chacun de ces graphiques, nous voyons bien que les zéros de  $P_n$  sont situés strictement entre ceux de  $P_{n+1}$ . De plus, nous voyons que les points critiques de chaque polynôme sont situés strictement entre les zéros du même polynôme.

# Pages Wikipédia utiles

Polynômes orthogonaux,

[http://fr.wikipedia.org/wiki/Polynomes\\_orthogonaux](http://fr.wikipedia.org/wiki/Polynomes_orthogonaux)

Polynômes de Legendre,

[http://fr.wikipedia.org/wiki/Polynome\\_de\\_Legendre](http://fr.wikipedia.org/wiki/Polynome_de_Legendre)

Legendre Polynomials,

[http://en.wikipedia.org/wiki/Legendre\\_polynomials](http://en.wikipedia.org/wiki/Legendre_polynomials)

Generalized Fourier Series,

[http://en.wikipedia.org/wiki/Generalized\\_Fourier\\_series](http://en.wikipedia.org/wiki/Generalized_Fourier_series)

Série Génératrice,

[http://fr.wikipedia.org/wiki/Serie\\_generatrice](http://fr.wikipedia.org/wiki/Serie_generatrice)

Polynôme de Tchebyshev,

[http://fr.wikipedia.org/wiki/Polynome\\_de\\_Tchebychev](http://fr.wikipedia.org/wiki/Polynome_de_Tchebychev)

Méthodes de quadratures de Gauss

[http://fr.wikipedia.org/wiki/Methodes\\_de\\_quadrature\\_de\\_Gauss](http://fr.wikipedia.org/wiki/Methodes_de_quadrature_de_Gauss)



# Bibliographie

- [1] M. AL-GWAIZ : *Sturm-Liouville Theory and its Applications*. Springer-Verlag, 2008.
- [2] F. E. BURK : *A Garden of Integrals*. The mathematical Association of America, 2007.
- [3] E. A. CODDINGTON : *An Introduction to Ordinary Differential Equations*. Prentice-Hall Inc., 1961.
- [4] C. COHEN-TANNOUJDI, B. DIU et F. LALOE : *Mécanique quantique I*. Hermann, éditeur des sciences et des arts, 1973.
- [5] H. F. DAVIS : *Fourier Series and orthogonal Functions*. Allyn and Bacon inc., 1963.
- [6] A. FORTIN : *Analyse numérique pour ingénieur*. Clarendon press-Oxford, 2002.
- [7] J. FOURIER : *Théorie analytique de la chaleur dans les corps solides*. Firmin Didot, père et fils, 1822.
- [8] R. KOEKOEK, P. A. LESKY et R. F. SWARTTOUW : *Hypergeometric Polynomials and Their  $q$ -Analogues*. Springer-Verlag, 2010.
- [9] M. KRAWTCHOUK : Sur une généralisation des polynômes d'Hermite. *Comptes Rendus Mathématiques*, 189(17):620–622, 1929.
- [10] F. LAROCHE : Polynômes et fonctions de Legendre. *Promenades mathématiques*, <http://promenadesmaths.free.fr/>, Consulté le 20 juillet 2014.
- [11] J. MASHREGHI : *Équations différentielles*. (notes de cours), Hiver 2012.
- [12] N. NICOLOV et B. SENDOV : A converse of the Gauss-Lucas theorem. *American Mathematical Monthly*, 121(6):541–544, 2014.
- [13] Q. RAHMAN et G. SCHMEISSER : *Analytic Theory of Polynomials*. Presses internationales Polytechnique, 2011.
- [14] R. RESNICK et R. EISBERG : *Quantum Physics of Atoms, Molecules, Solids, Nuclei and Particles*. John Wiley and sons inc., 1985.

- [15] J. ROSTAND : *Mesure et intégration*. (notes de cours), Automne 2012.
- [16] G. SZEGÖ : *Orthogonal Polynomials*. American mathematical society, 1939.
- [17] N. THAKARE et H. MADEHERA : Use of Hermite's Method to Obtain Generating Functions for Classical Orthogonal Polynomials. *Indian J. pure appl. Math*, 13(2):183–189, 1982.
- [18] P. YAP, P. RAVEENDRAN et S. ONG : Krawtchouk Moments as a New Set of Discrete Orthogonal Moments for Image Reconstruction. *IEEE*, 1:908–912, 2002.