

# Approximation numérique

Patrick Fischer <sup>1</sup>, Lisl Weynans, Charles Dossal

<sup>1</sup>IMB, Université Bordeaux 1, 351 Cours de la Libération, 33405 Talence, France  
(Patrick.Fischer@math.u-bordeaux1.fr)

# Contents

<b>1</b>	<b>La représentation des fonctions</b>	<b>1</b>
1.1	Introduction . . . . .	1
1.2	Interpolation polynomiale . . . . .	1
1.2.1	Polynômes de Lagrange . . . . .	1
1.2.2	Polynômes de Newton . . . . .	3
1.2.3	Choix des points de discrétisation . . . . .	4
1.2.4	Polynômes d’Hermite . . . . .	7
1.2.5	Interpolation locale: lissage par fonctions splines . . . . .	8
1.3	C.A.O.: les courbes de Bézier . . . . .	10
1.3.1	Polynômes de Bernstein . . . . .	10
1.3.2	Les courbes de Bézier, définition . . . . .	11
1.3.3	Propriétés . . . . .	12
1.3.4	Élévation de la longueur d’une courbe de Bézier . . . . .	13
1.3.5	Algorithme de De Casteljaou . . . . .	13
1.4	Approximation d’une fonction . . . . .	14
1.4.1	Qu’est-ce qu’une approximation ? . . . . .	14
1.4.2	Caractérisation de la meilleure approximation . . . . .	15
1.4.3	Polynômes orthogonaux . . . . .	15
1.5	Approximation au sens des moindres carrés discrets . . . . .	17
<b>2</b>	<b>Dérivation et intégration numériques</b>	<b>20</b>
2.1	Dérivation numérique . . . . .	20
2.2	Intégration numérique . . . . .	22
2.2.1	Formule de quadrature du type interpolation . . . . .	22
2.2.2	Accélération de la convergence: méthode de Romberg . . . . .	26
2.2.3	Formule d’Euler - Mac Laurin . . . . .	27
2.2.4	Formules de Gauss . . . . .	29
2.3	Méthodes de Monte-Carlo . . . . .	33
<b>3</b>	<b>Introduction aux méthodes numériques de résolution d’équations différentielles</b>	<b>35</b>
3.1	Méthodes à 1 pas . . . . .	35
3.1.1	Méthode d’Euler . . . . .	35
3.1.2	Schéma général . . . . .	36
3.1.3	Erreur de discrétisation . . . . .	38
3.2	Méthodes numériques plus précises . . . . .	39

3.2.1	Méthode du point milieu . . . . .	39
3.2.2	Méthode de Runge-Kutta . . . . .	39
3.2.3	Méthodes implicites . . . . .	40
3.2.4	Méthodes à pas variable . . . . .	40
3.3	Méthodes à pas multiples . . . . .	40
<b>4</b>	<b>Traitement du signal: Analyse de Fourier et ondelettes</b>	<b>42</b>
4.1	Généralités . . . . .	42
4.2	Analyse de Fourier . . . . .	43
4.2.1	Transformée de Fourier discrète . . . . .	45
4.2.2	Fast Fourier Transform . . . . .	45
4.2.3	Phénomène de Gibbs . . . . .	46
4.3	Filtre et échantillonnage . . . . .	47
4.3.1	Notion de filtre . . . . .	47
4.3.2	Transformée de Fourier, théorème de Shannon . . . . .	48
4.4	Analyse temps-fréquence . . . . .	49
4.5	La transformée de Fourier à fenêtre . . . . .	50
4.6	La transformée en ondelettes . . . . .	51
4.6.1	La transformée en ondelettes continues . . . . .	51
4.6.2	La transformée en ondelettes discrètes . . . . .	53
4.6.3	La transformée en ondelettes orthogonales . . . . .	53
4.7	Paquets d'ondelettes . . . . .	55
4.7.1	Cas 1D . . . . .	55
4.7.2	Cas 2D . . . . .	56
4.8	Le critère entropique et l'algorithme du choix de la meilleure base . . . . .	56
4.8.1	L'entropie . . . . .	57
4.8.2	Algorithme du choix de la meilleure base . . . . .	58
<b>5</b>	<b>Exercices</b>	<b>60</b>
5.1	. . . . .	60
5.2	. . . . .	61
5.3	. . . . .	62
5.4	. . . . .	63
5.5	. . . . .	64
5.6	. . . . .	65

Ce résumé de cours est une introduction aux méthodes d'approximation numérique. Il s'adresse à des étudiants de troisième année de licence en ingénierie mathématique. Il est basé sur les notes de cours de Charles-Henri Bruneau, l'ouvrage de R. Théodor *Initiation à l'analyse numérique*, Masson et l'ouvrage de C. Gasquet et P. Witomski *Analyse de Fourier et Applications*, Masson.

# Chapter 1

## La représentation des fonctions

### 1.1 Introduction

On se donne une fonction  $f$  définie sur un intervalle réel  $I$  et à valeurs réelles. Cette fonction est caractérisée par une propriété particulière: une équation, un jeu de données numériques ... mais on ne la connaît pas explicitement. On cherche à déterminer une fonction  $f_h$  appartenant à un ensemble de fonctions bien connues qui sera une approximation de la fonction  $f$  initiale. Deux approches sont possibles pour le calcul de cette approximation:

- On impose que  $f$  et  $f_h$  coïncident (et éventuellement leurs dérivées) en des points choisis. Cette approche conduit aux méthodes d'interpolation polynomiale. Elle permet également d'approcher la fonction en dehors de l'intervalle initial.
- On cherche à minimiser  $\|f - f_h\|$  dans un espace fonctionnel normé à préciser. Ce type de méthode conduit à un problème d'optimisation dont le résultat dépend de la norme choisie.

### 1.2 Interpolation polynomiale

On se donne une fonction  $f$  continue dont on connaît les valeurs  $f(x_i)$  en  $n + 1$  points différents  $x_0, x_1, \dots, x_n$ . On cherche à reconstituer  $f$  sur un intervalle contenant les  $x_i$  (on peut considérer les  $x_i$  comme étant ordonnés:  $x_i < x_{i+1}$ ). Le problème de l'interpolation (c'est-à-dire sur  $[x_0, x_n]$ ) ou de l'extrapolation (c'est-à-dire hors de  $[x_0, x_n]$ ) polynomiale consiste à construire un polynôme  $p$  de degré minimal tel que

$$\forall i \in \{0, \dots, n\}, \quad p(x_i) = f(x_i). \quad (1.1)$$

#### 1.2.1 Polynômes de Lagrange

**Théorème 1.1** *Il existe un unique polynôme  $p$  de degré  $n$  vérifiant (1.1).*

**Preuve:** Existence:

On considère les polynômes de Lagrange:

$$L_j(x) = \prod_{k \neq j} \frac{x - x_k}{x_j - x_k} \quad (1.2)$$

Ils sont de degré  $n$  et vérifient:

$$L_j(x) = \begin{cases} 1 & \text{si } x = x_j \\ 0 & \text{si } x = x_i, i \neq j \end{cases} \quad (1.3)$$

On pose ensuite

$$p(x) = \sum_{j=0}^n L_j(x) f(x_j) \quad (1.4)$$

qui répond bien à la question.

Unicité:

Si  $p_1$  et  $p_2$  sont deux polynômes qui vérifient les hypothèses alors  $q = p_1 - p_2$  est un polynôme de degré  $n$  qui admet  $n + 1$  racines distinctes ( $x_0, x_1, \dots, x_n$ ); il est donc nul et donc  $p_1 = p_2$ . ■

**Cas particulier:** On suppose les points  $x_i$  équidistants. On note  $h = x_{i+1} - x_i$ . On a alors  $x_j - x_k = (j - k)h$ , et

$$\prod_{k \neq j} (x_j - x_k) = h^n \prod_{k \neq j} (j - k) = (-1)^{n-j} h^n j! (n - j)!. \quad (1.5)$$

On obtient alors:

$$L_j(x) = \frac{(-1)^{n-j}}{h^n j! (n - j)!} \prod_{k \neq j} (x - x_k) \quad (1.6)$$

**Théorème 1.2** Si  $f$  est de classe  $C^{n+1}$  sur  $I$  contenant les points  $x_i$  ordonnés, alors pour tout  $x \in I$ , il existe  $c \in ]x_0, x_n[$  tel que

$$f(x) - p(x) = \frac{1}{(n + 1)!} L(x) f^{(n+1)}(c) \quad (1.7)$$

où  $L(x) = \prod_{k=0}^n (x - x_k)$ .

**Preuve:** Si  $x = x_k$  alors on a  $0 = 0$ .

Sinon si  $x \neq x_k$ , on définit  $\varphi(t) = f(t) - p(t) - \frac{f(x) - p(x)}{L(x)} L(t)$  qui est une fonction de classe  $C^{n+1}$ . La fonction  $\varphi$  vérifie alors  $\varphi(x_k) = 0$ ,  $0 \leq k \leq n$  et  $\varphi(x) = 0$ . Cette fonction a donc au moins  $n + 2$  zéros, sa dérivée première  $n + 1$  zéros, et donc  $\varphi^{(n+1)}$  a au moins un zéro noté  $c$  (th. de Rolle).

D'où

$$\varphi^{(n+1)}(c) = f^{(n+1)}(c) - \frac{f(x) - p(x)}{L(x)} (n + 1)! = 0 \quad (1.8)$$

■

**Corollaire 1.1**

$$|f(x) - p(x)| \leq |L(x)| \frac{M_{n+1}}{(n+1)!} \quad (1.9)$$

où  $M_{n+1} = \max_{x \in I} |f^{(n+1)}(x)|$ .

**Remarque:** On a donc intérêt à choisir les points  $x_i$  de façon à minimiser  $|L(x)|$ . Mais l'approximation ne sera alors bonne que pour le  $x$  en question. Les polynômes de Lagrange ne sont pas pratiques puisque d'un point de vue numérique, il est difficile de déduire  $L_{j+1}$  à partir de  $L_j$ . Pour cela on introduit le polynôme d'interpolation de Newton.

### 1.2.2 Polynômes de Newton

- **Points quelconques**

Les polynômes  $e_k$  de la base de Newton sont définis comme suit :

$$e_k(x) = \prod_{i=0}^{k-1} (x - x_i) = (x - x_0)(x - x_1) \cdots (x - x_{k-1}), \quad k = 1, \dots, n, \quad (1.10)$$

avec pour convention  $e_0 = 1$ . On obtient alors:

$$\begin{aligned} e_1 &= (x - x_0) \\ e_2 &= (x - x_0)(x - x_1) \\ e_3 &= (x - x_0)(x - x_1)(x - x_2) \\ &\vdots \\ e_n &= (x - x_0)(x - x_1) \cdots (x - x_{n-1}) \end{aligned} \quad (1.11)$$

L'ensemble des polynômes  $(e_k)_{0 \leq k \leq n}$  forment une base de l'espace  $\mathcal{P}_n$  des polynômes de degré au plus  $n$ , puisqu'il s'agit d'une famille de  $(n+1)$  polynômes de degré zéro à  $n$ .

Le polynôme d'interpolation de Newton de degré  $n$  relatif aux données  $\{(x_0, f(x_0)), (x_1, f(x_1)), \dots, (x_n, f(x_n))\}$  s'écrit :

$$\begin{aligned} P_n(x) &= \sum_{k=0}^n \alpha_k e_k(x) \\ &= \alpha_0 + \alpha_1(x - x_0) + \alpha_2(x - x_0)(x - x_1) \\ &\quad + \dots + \alpha_n(x - x_0)(x - x_1) \cdots (x - x_{n-1}) \end{aligned} \quad (1.12)$$

avec  $P_n(x_i) = f(x_i)$ ,  $\forall i = 0, \dots, n$ . Il faut alors déterminer les coefficients  $(\alpha_k)_{0 \leq k \leq n}$ .

**Définition 1.1** Soit  $f$  dont on connaît les valeurs en  $n+1$  points. On appelle différences divisées d'ordre  $0, 1, \dots, n$  les quantités:

$$\begin{aligned} \delta^0 f(x_k) &= f(x_k), \quad \forall 0 \leq k \leq n \\ \delta^1 f(x_0, x_1) &= \frac{\delta^0 f(x_1) - \delta^0 f(x_0)}{x_1 - x_0} \\ \delta^2 f(x_0, x_1, x_2) &= \frac{\delta^1 f(x_1, x_2) - \delta^1 f(x_0, x_1)}{x_2 - x_0} \\ &\vdots \\ \delta^n f(x_0, \dots, x_n) &= \frac{\delta^{n-1} f(x_1, \dots, x_n) - \delta^{n-1} f(x_0, \dots, x_{n-1})}{x_n - x_0}. \end{aligned} \quad (1.13)$$

Le polynôme d'interpolation de Newton de degré  $n$  s'écrit alors à l'aide des différences divisées successives :

$$P_n(x) = \delta^0 f(x_0) + \sum_{k=1}^n \delta^k f(x_0, \dots, x_k) e_k(x). \quad (1.14)$$

**Remarques:** Le polynôme d'interpolation de Newton conduit à la même estimation d'erreur que le polynôme de Lagrange.

- **Points équidistants**

Quand les points sont équidistants ( $x_{i+1} - x_i = h$ ), nous pouvons faire le calcul en utilisant les différences finies. On pose:

$$\begin{aligned} \Delta^1 f(x) &= f(x+h) - f(x) \\ \Delta^2 f(x) &= \Delta^1(\Delta^1 f(x)) = \Delta^1 f(x+h) - \Delta^1 f(x) \\ &\vdots \\ \Delta^n f(x) &= \Delta^1(\Delta^{n-1} f(x)), \end{aligned} \quad (1.15)$$

et le polynôme d'interpolation de Newton de degré  $n$  s'écrit alors:

$$P_n(x) = f(x_0) + \frac{y}{1!} \Delta^1 f(x_0) + \dots + \frac{y(y-1) \cdots (y-n+1)}{n!} \Delta^n f(x_0) \quad (1.16)$$

où  $y = \frac{x - x_0}{h}$ .

C'est le polynôme progressif car les différences sont à droites. Nous pouvons aussi définir le polynôme régressif avec des différences à gauche en partant de  $x_n$  au lieu de  $x_0$ . Le polynôme progressif est adapté au cas où l'on cherche à déterminer la valeur en un point  $x$  situé au début de l'intervalle de calcul, et le polynôme régressif adapté au cas où l'on cherche à déterminer la valeur en fin d'intervalle.

- **Attention !** On n'obtient pas de meilleurs résultats en augmentant le degré du polynôme d'interpolation. Supposons que nous cherchions à approcher la fonction  $f(x) = \frac{1}{1+25x^2}$  sur l'intervalle  $[-1, 1]$ . Un polynôme d'interpolation de degré trop élevé sur des points équidistants conduit à de fortes oscillations aux bords de l'intervalle (cf. Figure 1.1). Ce phénomène est connu sous le nom de phénomène de Runge.

En pratique, on effectue des interpolations avec des polynômes de degrés faibles sur des petits intervalles plutôt que des polynômes de degrés élevés sur de grands intervalles.

### 1.2.3 Choix des points de discrétisation

Soit une fonction  $f : [a, b] \rightarrow \mathbb{R}$  et soient  $x_0, x_1, \dots, x_n \in [a, b]$ . Nous avons vu (1.9) que si  $f \in C^{n+1}([a, b])$  alors

$$|f(x) - P_n(x)| \leq \max_{x \in [a, b]} \frac{|f^{(n+1)}(x)|}{(n+1)!} \max_{x \in [a, b]} \|(x-x_0)(x-x_1) \cdots (x-x_n)\| \quad (1.17)$$



Figure 1.1: Phénomène de Runge: En rouge  $f(x)$ , en bleu interpolation avec un polynôme de degré 5, et en vert avec un polynôme de degré 9

où  $P_n(x)$  est un polynôme d'interpolation de type Lagrange ou Newton. L'erreur est majorée par le produit de deux termes:

$$\begin{aligned} \max_{x \in [a,b]} \frac{|f^{(n+1)}(x)|}{(n+1)!} \\ \max_{x \in [a,b]} |v(x)| &= \max_{x \in [a,b]} |(x-x_0)(x-x_1) \cdots (x-x_n)| \end{aligned} \tag{1.18}$$

Le premier terme dépend de la fonction à interpoler mais le polynôme  $v$  de degré  $n+1$  en est indépendant. On peut donc chercher les points  $x_i$  qui minimiseront ce terme dans la majoration de l'erreur. On note  $E_n([a,b])$  l'ensemble des polynômes de degré  $n$  dont le coefficient principal (devant  $x^n$ ) est 1. Le meilleur choix des  $x_i$  est alors donné par les racines du polynôme  $q \in E_{n+1}([a,b])$  vérifiant:

$$\forall v \in E_{n+1}([a,b]), \max_{x \in [a,b]} |q(x)| \leq \max_{x \in [a,b]} |v(x)| \tag{1.19}$$

On peut déterminer  $q$  dans le cas particulier où  $a = -1$  et  $b = 1$  et étendre ensuite le résultat au cas général par un changement de variable.

**Définition 1.2** On appelle fonction polynomiale de Tchébicheff de degré  $n$  la fonction  $T_n : [-1, 1] \rightarrow \mathbb{R}$  définie par:

$$T_n(x) = \cos(n \arccos(x)). \tag{1.20}$$

**Théorème 1.3** Les polynômes de Tchébicheff vérifient la relation de récurrence  $T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x)$ .

Le coefficient du terme de plus haut degré  $x^n$  de  $T_n$  est  $2^{n-1}$ .

**Preuve:** On utilise la formule de trigonométrie habituelle:

$$\cos((n+1)\theta) = \cos(n\theta)\cos(\theta) - \sin(n\theta)\sin(\theta) \quad (1.21)$$

qui nous permet d'écrire:

$$\cos((n+1)\theta) + \cos((n-1)\theta) = 2\cos(n\theta)\cos(\theta). \quad (1.22)$$

On obtient alors la formule du théorème. ■

**Théorème 1.4**  $T_n$  possèdent  $n$  racines simples:

$$x_k = \cos\left(\frac{2k-1}{2n}\pi\right), \quad k = 1, 2, \dots, n. \quad (1.23)$$

$T_n$  atteint ses extrema sur l'intervalle  $[-1, 1]$  aux  $n+1$  points:

$$x'_k = \cos\left(\frac{k}{n}\pi\right), \quad k = 0, 1, \dots, n \quad (1.24)$$

pour lesquels il prend alternativement les valeurs  $-1$  et  $1$ .

**Preuve:** On vérifie facilement que les  $\{x_k\}_{k=1, \dots, n}$  sont bien racines de  $T_n$ , et comme  $T_n$  est de degré  $n$ , ce sont les seules racines. Ensuite, on calcule la dérivée de  $T_n(x)$ :

$$T'_n(x) = \sin(n \arccos(x)) \cdot \frac{n}{\sqrt{1-x^2}}. \quad (1.25)$$

On vérifie que  $T'_n(x'_k) = 0$  pour  $k = 1, \dots, n$  et  $T_n(x'_k) = (-1)^k$ . On a aussi  $T_n(x'_0) = 1$  et  $T_n(x'_n) = (-1)^n$ . ■

**Définition 1.3** On appelle fonction polynôme de Tchébicheff normalisée  $\bar{T}_n = \frac{1}{2^{n-1}}T_n$  pour  $n \geq 1$ .

**Théorème 1.5**

$$\forall p \in E_n, \quad \frac{1}{2^{n-1}} = \max_{x \in [-1, 1]} |\bar{T}_n(x)| \leq \max_{x \in [-1, 1]} |p(x)| \quad (1.26)$$

**Preuve:** Nous cherchons donc à montrer que le polynôme  $\bar{T}_n$  minimise  $\max_{x \in [-1, 1]} |p(x)|$   $\forall p \in E_n$ . Supposons qu'il existe  $p \in E_n$  tel que,

$$\max_{x \in [-1, 1]} |p(x)| < \max_{x \in [-1, 1]} |\bar{T}_n(x)| = \frac{1}{2^{n-1}}. \quad (1.27)$$

On pose  $r = \bar{T}_n - p$  qui est une polynôme de degré au plus  $n - 1$ . On remarque que

$$r(x'_k) = \bar{T}_n(x'_k) - p(x'_k) = \frac{(-1)^k}{2^{n-1}} - p(x'_k), \quad k = 0, \dots, n. \quad (1.28)$$

prend alternativement le signe positif ou négatif (car  $|p(x'_k)| < \frac{1}{2^{n-1}}$ ). Comme le polynôme  $r$  est continue, il possède donc  $n$  racines. Mais comme il s'agit d'un polynôme de degré  $n - 1$ , on a donc forcément  $r = 0$ . ■

Sur l'intervalle  $[-1, 1]$  et en choisissant les abscisses d'interpolation  $x_k = \cos\left(\frac{2k+1}{2(n+1)}\pi\right)$ ,  $k = 0, 1, \dots, n$  on obtient alors la majoration de l'erreur d'interpolation:

$$|f(x) - P_n(x)| \leq \frac{1}{2^n} \frac{1}{(n+1)!} \max_{x \in [-1, 1]} |f^{(n+1)}(x)| \quad (1.29)$$

## 1.2.4 Polynômes d'Hermite

De la même manière que dans l'interpolation de Lagrange, on impose que le polynôme prenne un certain nombre de valeurs, mais de plus, on fixe les valeurs des dérivées en ces points. Soient les  $n + 1$  triplets  $(x_i, f(x_i), f'(x_i))$  pour  $i = 0, \dots, n$  où les  $x_i$  sont tous distincts.

On cherche un polynôme  $p$  tel que:

$$\begin{cases} p(x_i) &= f(x_i) \\ p'(x_i) &= f'(x_i) \end{cases} \quad (1.30)$$

**Théorème 1.6** *Il existe un polynôme  $p$  et un seul de degré au plus  $2n + 1$  tel que (1.30) est vérifié.*

**Preuve:** Unicité: Supposons qu'il existe des polynômes  $p$  et  $q$  de degré au plus  $2n + 1$  vérifiant (1.30). Le polynôme  $r = p - q$ , qui est aussi de degré au plus  $2n + 1$ , admet chaque  $x_i$  comme racine double ( $r(x_i) = r'(x_i) = 0$ ). Il possède donc  $2n + 2$  racines au moins,  $r$  est donc le polynôme nul.

Existence: On montre que le polynôme  $p$  suivant satisfait les conditions précédentes:

$$p(x) = \sum_{i=0}^n h_i(x) f(x_i) + \sum_{i=0}^n k_i(x) f'(x_i). \quad (1.31)$$

où:

$$h_i(x) = (1 - 2(x - x_i) l'_i(x_i)) l_i^2(x) \quad (1.32)$$

$$k_i(x) = (x - x_i) l_i^2(x) \quad (1.33)$$

avec

$$l_i(x) = \prod_{j=0, j \neq i}^n \frac{x - x_j}{x_i - x_j} \quad (1.34)$$

■

Evaluation de l'erreur d'interpolation:

**Théorème 1.7** Soit  $f \in C^{2n+2}[a, b]$ . Si  $a \leq x_0 < x_1 < x_2 < \dots < x_n \leq b$ , alors pour tout  $x \in [a, b]$ , il existe  $\xi \in [a, b]$  t.q.:

$$f(x) - p(x) = \frac{(x - x_0)^2(x - x_1)^2 \cdots (x - x_n)^2}{(2n + 2)!} f^{(2n+2)}(\xi) \quad (1.35)$$

où  $a \leq \min(x_i, x) < \xi < \max(x_i, x) \leq b$ .

**Preuve:** Démonstration analogue à celle de l'interpolation de Lagrange. ■

**Remarque:** Nous pouvons définir des formules d'interpolation mixte Lagrange-Hermite où les valeurs des dérivées ne sont utilisées que pour certains points. Nous pouvons également faire intervenir des dérivées d'ordre plus élevée.

## 1.2.5 Interpolation locale: lissage par fonctions splines

L'interpolation polynomiale possède deux défauts majeurs inévitables:

- Le coût des calculs devient élevé lorsque le degré du polynôme est grand.
- Les effets de bord sont importants si l'intervalle d'interpolation est grand.

Une solution consiste à travailler localement: on fait alors de l'interpolation par morceaux, c'est-à-dire que sur un sous intervalle donné on fait de l'interpolation polynomiale de degré faible, puis on impose des conditions de régularité (continuité ou plus) pour relier les morceaux.

Les splines vont réaliser une interpolation polynomiale par morceaux mais on imposera de plus un degré de régularité aux points de discrétisation appelés aussi "noeuds". Nous ne présenterons qu'un exemple élémentaire de fonctions splines: les splines d'interpolation cubiques.

Soient  $x_0, x_1, \dots, x_n$  les points d'interpolation. Sur chaque segment  $[x_i, x_{i+1}]$ , on cherche un polynôme  $s_i$  tel que les conditions d'interpolation  $s_i(x_i) = f(x_i)$  pour  $i = 0, \dots, n$  soient vérifiées. De plus, on impose des conditions de raccord de façon à obtenir une fonction de classe  $C^2$ . Pour  $i = 1, \dots, n - 1$ :

$$s'_{i-1}(x_i) = s'_i(x_i) \quad (1.36)$$

$$s''_{i-1}(x_i) = s''_i(x_i) \quad (1.37)$$

Ces deux conditions entraînent la continuité des dérivées premières et secondes aux noeuds de discrétisation. Elles nous incitent à chercher  $s_i$  sous la forme d'un polynôme de degré trois. Par interpolation linéaire, nous pouvons calculer  $s''_i$  en posant  $h_i = x_{i+1} - x_i$ :

$$s''_i(x) = M_i \frac{x_{i+1} - x}{h_i} + M_{i+1} \frac{x - x_i}{h_i}, \quad (1.38)$$

où les coefficients  $M_i$  sont à déterminer.

En intégrant deux fois, on obtient:

$$s_i(x) = M_i \frac{(x_{i+1} - x)^3}{6h_i} + M_{i+1} \frac{(x - x_i)^3}{6h_i} + a_i(x_{i+1} - x) + b_i(x - x_i), \quad (1.39)$$

où  $a_i$  et  $b_i$  dénotent les constantes d'intégration.

Ces constantes peuvent se déterminer en écrivant les relations de continuité aux noeuds  $s_i(x_i) = f(x_i)$  et  $s_i(x_{i+1}) = f(x_{i+1})$ . C'est-à-dire:

$$s_i(x_i) = M_i \frac{h_i^2}{6} + a_i h_i = f(x_i). \quad (1.40)$$

D'où

$$a_i = \frac{f(x_i)}{h_i} - M_i \frac{h_i}{6} \quad (1.41)$$

et

$$s_i(x_{i+1}) = M_{i+1} \frac{h_i^2}{6} + b_i h_i = f(x_{i+1}) \quad (1.42)$$

implique

$$b_i = \frac{f(x_{i+1})}{h_i} - M_{i+1} \frac{h_i}{6}. \quad (1.43)$$

En remplaçant alors  $a_i$  et  $b_i$  par leurs expressions, on obtient:

$$\begin{aligned} s_i(x) = & M_i \left( \frac{(x_{i+1} - x)^3}{6h_i} - \frac{h_i}{6}(x_{i+1} - x) \right) + M_{i+1} \left( \frac{(x - x_i)^3}{6h_i} - \frac{h_i}{6}(x - x_i) \right) \\ & + \frac{f(x_i)}{h_i}(x_{i+1} - x) + \frac{f(x_{i+1})}{h_i}(x - x_i). \end{aligned} \quad (1.44)$$

Il reste alors à écrire les conditions de continuité pour les dérivées premières,

$$s'_i(x) = M_i \left( \frac{-(x_{i+1} - x)^2}{2h_i} + \frac{h_i}{6} \right) + M_{i+1} \left( \frac{(x - x_i)^2}{2h_i} - \frac{h_i}{6} \right) - \frac{f(x_i)}{h_i} + \frac{f(x_{i+1})}{h_i}. \quad (1.45)$$

La condition de continuité  $s'_i(x_i) = s'_{i-1}(x_i)$  en  $x_i$  s'écrit alors:

$$-M_i \frac{2h_i}{6} - M_{i+1} \frac{h_i}{6} - \frac{f(x_i)}{h_i} + \frac{f(x_{i+1})}{h_i} = M_{i-1} \frac{h_{i-1}}{6} + 2M_i \frac{h_{i-1}}{6} - \frac{f(x_{i-1})}{h_{i-1}} + \frac{f(x_i)}{h_{i-1}}. \quad (1.46)$$

Cette dernière relation peut se réécrire sous la forme d'un système linéaire à résoudre:

$$h_i M_{i+1} + 2(h_i + h_{i-1})M_i + h_{i-1}M_{i-1} = 6 \left( \frac{f(x_{i+1}) - f(x_i)}{h_i} - \frac{f(x_i) - f(x_{i-1})}{h_{i-1}} \right). \quad (1.47)$$

Nous obtenons alors un système linéaire de  $n - 1$  équations à  $n + 1$  inconnues. A moins de connaître explicitement  $M_0$  et  $M_n$ , on prend souvent  $M_0 = M_n = 0$ . D'autres choix sont évidemment possibles.

La théorie de l'interpolation locale peut s'étendre de la même façon à des dimensions supérieures. Il existe d'autres méthodes d'interpolation, comme l'interpolation trigonométrique utilisée en traitement du signal.

## 1.3 C.A.O.: les courbes de Bézier

Dans l'industrie, on utilise des courbes pour la conception d'objets: ailes d'avions, fuselage, profile de voiture, fabrication de prothèses, etc. Les courbes et les surfaces polynomiales sont les plus habituellement utilisées depuis les années 60, i.e. depuis que Bézier et De Casteljau leur ont donné une détermination mathématique exploitable sur le plan informatique à l'aide de points de contrôle via les polynômes de Bernstein. Dans cette partie, nous nous limitons à la description des courbes dites Bézier polynomiales. Une telle courbe permet le stockage et le tracé d'une courbe polynomiale à l'aide d'un ensemble fini de points appelé polygone de contrôle.

### 1.3.1 Polynômes de Bernstein

**Définition 1.4** Soit  $\mathbb{P}_n$  l'espace des polynômes à une variable réelle de degré inférieur ou égal à  $n$ . On appelle  $i^{\text{ème}}$  polynôme de Bernstein de degré  $n$  sur l'intervalle  $[0, 1]$ :

$$B_i^n(t) = C_n^i (1-t)^{n-i} t^i, \quad \forall t \in [0, 1], \quad (1.48)$$

où  $C_n^i = \frac{n!}{(n-i)! i!}$ .

**Proposition 1.1** L'ensemble des  $B_i^n(t)$ ,  $i = 0, \dots, n$  est une base de  $\mathbb{P}_n$ . Les polynômes de Bernstein vérifient les propriétés:

1. positivité:  $\forall t \in [0, 1], \forall i = 0, \dots, n, B_i^n(t) \geq 0$ ,
2. symétrie:  $\forall t \in [0, 1], \forall i = 0, \dots, n, B_i^n(t) = B_{n-i}^n(1-t)$ ,
3. le polynôme  $B_i^n(t)$  atteint son maximum sur  $[0, 1]$  en  $\frac{i}{n}$ ,
4. partition de l'unité:  $\forall t \in [0, 1], \sum_{i=0}^n B_i^n(t) = 1$ ,
5. relation de récurrence:

$$\begin{aligned} B_0^n(t) &= (1-t)B_0^{n-1}(t) \\ B_i^n(t) &= (1-t)B_i^{n-1}(t) + tB_{i-1}^{n-1}(t), \quad \forall i = 1, \dots, n-1, \\ B_n^n(t) &= tB_{n-1}^{n-1}(t). \end{aligned} \quad (1.49)$$

6. en notant  $DB_i^n(t)$  la dérivée première de  $B_i^n(t)$ , alors,

$$DB_i^n(t) = n(B_{i-1}^{n-1}(t) - B_i^{n-1}(t)), \quad \forall i = 0, \dots, n. \quad (1.50)$$

**Preuve:** On montre que c'est une base en montrant que la famille des polynômes de Bernstein est libre, et comme il y en a  $n + 1$  dans un espace de dimension  $n + 1$ , c'est forcément une base. Pour montrer que la famille est libre, on prend une combinaison linéaire  $\sum_i a_i B_i^n(t) = 0$ . On prend  $t = 0$ . On obtient alors  $a_0 = 0$ . On dérive ensuite la combinaison linéaire et on prend à nouveau  $t = 0$ . On trouve alors  $a_1 = 0$ . Etc. Les différents points de la proposition se vérifient ensuite directement. ■

### 1.3.2 Les courbes de Bézier, définition

**Définition 1.5** *On se donne un ensemble  $\mathcal{P}$  de  $n + 1$  points  $\mathcal{P} = \{P_0, \dots, P_n\}$  dans un espace de dimension  $d \geq 2$ . La courbe de Bézier polynomiale associée est la courbe paramétrée définie pour  $t \in [0, 1]$  par:*

$$B_{\mathcal{P}}(t) = \sum_{i=0}^n B_i^n(t) P_i. \quad (1.51)$$

L'ensemble  $\mathcal{P}$  est appelé polygone de contrôle ou polygone de Bézier. L'entier  $n$ , nombre de cotés de  $\mathcal{P}$ , est appelé la longueur de la courbe. Cette longueur est supérieure ou égale au degré de la courbe. La courbe de Bézier est utilisée pour  $t \in [0, 1]$  bien qu'elle soit définie pour  $t \in \mathbb{R}$  en tant que courbe polynomiale.

**Exemples:**

Pour  $n = 1, 2, 3, 4$ :

$$\begin{aligned} B_{\mathcal{P}}(t) &= (1-t)P_0 + tP_1, & (1.52) \\ B_{\mathcal{P}}(t) &= (1-t)^2P_0 + 2(1-t)tP_1 + t^2P_2, \\ B_{\mathcal{P}}(t) &= (1-t)^3P_0 + 3(1-t)^2tP_1 + 3(1-t)t^2P_2 + t^3P_3, \\ B_{\mathcal{P}}(t) &= (1-t)^4P_0 + 4(1-t)^3tP_1 + 6(1-t)^2t^2P_2 + 4(1-t)t^3P_3 + t^4P_4, \end{aligned}$$

Pour  $n = 1$ , il n'y a que deux points, et la courbe de Bézier correspondante est le segment

Figure 1.2: Courbe de Bézier avec  $n = 3$

reliant  $P_0$  à  $P_1$ . Pour  $n = 2$ , on est au point  $P_0$  en  $t = 0$  et en  $P_2$  en  $t = 1$ . La position du

point  $P_1$  influe sur la convexité de la courbe, sans que la courbe ne passe par  $P_1$ . Dans le cas général pour  $n \geq 3$ , les situations sont plus complexes. La courbe a pour extrémités  $P_0$  et  $P_n$ . Les points de contrôle  $P_1, \dots, P_{n-1}$  agissent comme des aimants vis à vis de la courbe sans être des points d'interpolation.

Une courbe polynomiale est déterminée par un polygone de contrôle. Toute déformation peut être obtenue par déplacement des points de contrôle. Lorsque le paramètre  $t$  est proche de  $\frac{i}{n}$ , le poids  $B_i^n(t)$  est prépondérant et la courbe  $B_{\mathcal{P}}$  est "proche" de  $P_i$ . Un déplacement du point  $P_i$  étire, avec une certaine inertie, la courbe vers ce point.

### 1.3.3 Propriétés

**Proposition 1.2** 1. Pour  $t$  fixé,  $B_{\mathcal{P}}(t)$  est le barycentre des points  $P_i$  affectés des poids  $B_i^n(t)$ . Ainsi le point  $B_{\mathcal{P}}(t)$  est toujours situé dans l'enveloppe convexe du polygone de contrôle.

2. Symétrie: on note  $\mathcal{Q} = \{P_n, P_{n-1}, \dots, P_1, P_0\}$ . On a alors:  $B_{\mathcal{P}}(1-t) = B_{\mathcal{Q}}(t)$ .

3. La dérivée d'ordre  $k$  de  $B_{\mathcal{P}}$ , notée  $D^k B_{\mathcal{P}}$  a pour expression:

$$D^k B_{\mathcal{P}}(t) = \frac{n!}{(n-k)!} \sum_{i=0}^{n-k} B_i^{n-k}(t) \Delta^k P_i, \quad (1.53)$$

où  $\Delta^k P_i$  est la différence progressive d'ordre  $k$  calculée à partir de  $P_i$ :

$$\begin{aligned} \Delta^0 P_i &= P_i \\ \Delta P_i &= P_{i+1} - P_i \\ \Delta^2 P_i &= \Delta(\Delta P_i) \\ &= \Delta P_{i+1} - \Delta P_i \\ &= P_{i+2} - 2P_{i+1} + P_i \\ \Delta^k P_i &= \Delta(\Delta^{k-1} P_i) \\ &= \Delta^{k-1} P_{i+1} - \Delta^{k-1} P_i \\ &= \sum_{j=0}^k (-1)^{k-j} C_k^j P_{i+j} \end{aligned} \quad (1.54)$$

4. Obtention de courbes de Bézier fermées globalement  $C^k$ . Il suffit que:

$$\begin{aligned} P_0 &= P_n \\ \Delta P_0 &= \Delta P_{n-1} \\ &\vdots \\ \Delta^k P_0 &= \Delta^k P_{n-k} \end{aligned} \quad (1.55)$$

5. Transformation affine. Une transformation affine  $A$  respecte les barycentres, donc  $A(B_{\mathcal{P}}(t)) = B_{A\mathcal{P}}(t)$  où  $A\mathcal{P} = \{AP_0, \dots, AP_n\}$ .



6. N'importe quelle courbe paramétrée polynomiale peut se réécrire comme une courbe de Bézier en effectuant un changement de base de la base  $\{1, t, \dots, t^n\}$  vers la base de Bernstein.

### 1.3.4 Elévation de la longueur d'une courbe de Bézier

La courbe de Bézier  $B_{\mathcal{P}}(t)$  de longueur  $n$  peut être décrite dans la base des polynômes de Bernstein de degré  $n + 1$  :  $B_{\mathcal{Q}}(t) = \sum_{i=0}^{n+1} B_i^{n+1}(t)Q_i$  où les sommets  $Q_i$  sont définis par :

$$\begin{aligned} Q_0 &= P_0 \\ \vdots &= \vdots \\ Q_i &= \frac{iP_{i-1} + (n - i + 1)P_i}{n + 1} \\ \vdots &= \vdots \\ Q_{n+1} &= P_n \end{aligned} \tag{1.56}$$

Par exemple pour  $n = 3$  :  $Q_0 = P_0$ ,  $Q_1 = \frac{1}{4}(P_0 + 3P_1)$ ,  $Q_2 = \frac{1}{2}(P_1 + P_2)$ ,  $Q_3 = \frac{1}{4}(3P_2 + P_3)$  et  $Q_4 = P_3$ . Le procédé peut se poursuivre, et on peut montrer que la suite des polygones converge uniformément vers le support de la courbe.

### 1.3.5 Algorithme de De Casteljau

Il permet de calculer tout point d'une courbe de Bézier et repose sur la proposition suivante.

**Proposition 1.3** Soit  $B_{\mathcal{P}}(t) = \sum_{i=0}^n B_i^n(t)P_i$  une courbe de Bézier; cette courbe peut aussi s'écrire pour  $j = 0, \dots, n$ :

$$B_{\mathcal{P}}(t) = \sum_{i=0}^{n-j} B_i^{n-j}(t)P_i^j(t), \tag{1.57}$$

où  $P_i^j(t)$  est défini par récurrence:

$$\begin{aligned} P_i^0(t) &= P_i, \quad i = 0, \dots, n \\ P_i^j(t) &= (1 - t)P_i^{j-1}(t) + tP_{i+1}^{j-1}(t), \quad j = 1, \dots, n, \quad i = 0, \dots, n - j. \end{aligned} \tag{1.58}$$

**Algorithme de De Casteljau:**

On applique la proposition précédente avec  $j = n$  et  $B_{\mathcal{P}}(t) = P_0^n(t)$ :

*Initialisation:*

Pour  $i = 0$  à  $n$ , on pose  $P_i^0(t) = P_i$ .

*Itérations:*

Pour  $j = 1$  à  $n$  faire

Pour  $i = 0$  à  $n - j$  faire

$$P_i^j(t) = (1 - t)P_i^{j-1}(t) + tP_{i+1}^{j-1}(t),$$

*Résultat:*

$$B_{\mathcal{P}}(t) = P_0^n(t)$$

## 1.4 Approximation d'une fonction

### 1.4.1 Qu'est-ce qu'une approximation ?

On veut approcher  $x^4$  sur  $[0, 1]$  par un polynôme de degré 1  $p(x) = ax + b$ . Comment choisir  $a$  et  $b$  ? On peut poser le problème de différentes manières: trouver  $p$  tel que:

1.  $\int_0^1 (x^4 - p(x))^2 dx$  soit minimum
2.  $\int_0^1 (x^4 - p(x))^2 dx + \int_0^1 \frac{d}{dx}(x^4 - p(x))^2 dx$  soit minimum
3.  $\max_{x \in [0, 1]} |x^4 - p(x)|$  soit minimum.

Chacun de ces critères conduira à des réponses différentes:

1.  $p(x) = \frac{4}{5}x - \frac{1}{5}$
2.  $p(x) = \frac{4}{5}x - \frac{6}{25}$
3.  $p(x) = x - \frac{3}{16}2^{1/3}$

Nous avons besoin d'un outil mathématique pour mesurer la distance de deux fonctions. Pour cela, on se place dans le cadre des espace vectoriels normés. Une fois que la norme est choisie le problème consiste à répondre aux trois questions suivantes:

- Comment caractériser la meilleure approximation ?
- Si elle existe, est-elle unique ?
- Comment la calculer ?

On se donne une fonction  $f$  définie sur un intervalle réel et à valeurs dans  $\mathbb{R}$ ; Cette fonction est caractérisée par une propriété particulière: une équation, un jeu de données... mais on ne la connaît pas explicitement. On cherche à en réaliser une meilleure approximation, dans un sens à déterminer. Soit  $f \in H$  un espace vectoriel normé. On cherche  $u \in V \subset H$  réalisant:

$$\|u - f\| = \inf_{v \in V} \|v - f\|. \quad (1.59)$$

L'espace  $V$  est en général de dimension finie; on le note  $V_n$  où  $n$  est sa dimension, et  $u$  sera alors noté  $u_n$ . Bien entendu, si  $n$  croît alors la quantité  $\|u_n - f\|$  va décroître. On parlera d'approximation convergente lorsque

$$\forall \epsilon > 0, \exists n_0 \in \mathbb{N}, n > n_0 \implies \|u_n - f\| \leq \epsilon. \quad (1.60)$$

**Proposition 1.4** Soit  $H$  un espace vectoriel normé tel que  $V_n \subset H$ . Pour toute fonction  $f \in H$ , il existe  $p \in V_n$  tel que  $\|f - p\| = \inf_{q \in V_n} \|f - q\|$ .

**Preuve:** On considère une suite minimisante  $p_m$  ( $\lim_{m \rightarrow +\infty} \|f - p_m\| = \inf_{q \in V_n} \|f - q\|$ ). On montre qu'elle est bornée:  $\|p_m\| \leq \|f - p_m\| + \|f\| \leq C, \forall m$ . Comme  $V_n$  est de dimension finie, on peut en extraire une sous-suite convergente qui a pour limite  $p$  (Th. de Bolzano-Weierstrass). Comme  $V_n$  est fermé alors la limite appartient à  $V_n$ . et donc  $\|f - p_m\| \rightarrow \|f - p\|$  car la norme est continue. ■

## 1.4.2 Caractérisation de la meilleure approximation

On suppose que  $H$  est muni d'un produit scalaire noté  $(f, g)$  qui induit la norme  $(f, f) = \|f\|^2$ . Si l'espace  $H$  est complet alors il s'agit d'un espace de Hilbert. On suppose le sous espace  $V_n$  de dimension finie. On peut alors construire une base  $\{p_1, \dots, p_n\}$  et caractériser la solution de (1.59) par sa décomposition sur cette base  $u_n = \sum_{j=1}^n a_j p_j$ . Le problème est maintenant ramené à un problème sur  $\mathbb{R}^n$  c'est-à-dire à la détermination des coefficients  $a_j$ . En fonction du choix de la norme, et du choix de la base, nous obtenons alors des résultats différents.

On peut considérer des bases locales, des bases orthonormées (avec des polynômes orthogonaux) ou des bases quelconques (comme la base canonique  $\{1, x, x^2, \dots, x^n\}$ ).

## 1.4.3 Polynômes orthogonaux

**Proposition 1.5** Soit  $H$  un espace de Hilbert tel que  $V_n \subset H, \forall n \in \mathbb{N}$ . Il existe une unique suite de polynômes  $p_n$  de degré  $n$  de coefficient  $x_n$  égal à 1 tel que  $(p_n, q) = 0, \forall q \in V_{n-1}$ .

**Preuve:** La suite  $p_n$  est obtenue par le procédé d'orthogonalisation de Gram-Schmidt à partir de la suite de monômes  $1, x, \dots, x^n, \dots$ . La projection sur le sous espace engendré par les  $k + 1$  premiers polynômes  $p_j$  est définie par  $P_{V_k}(x) = \sum_{j=0}^k \frac{(x, p_j)}{(p_j, p_j)} p_j(x)$ . Ce qui donne, appliqué aux monômes:

$$\begin{aligned} p_0(x) &= 1 & (1.61) \\ p_1(x) &= x - \frac{(x, 1)}{(1, 1)} p_0(x) \\ p_n(x) &= x^n - \sum_{j=0}^{n-1} \frac{(x, p_j)}{(p_j, p_j)} p_j(x). \end{aligned}$$

■

**Proposition 1.6** Les polynômes orthogonaux  $p_n$  vérifient la relation de récurrence:

$$p_n(x) = (x - \lambda_n) p_{n-1}(x) - \mu_n p_{n-2}(x), \quad n \geq 1 \quad (1.62)$$

où  $p_{-1}(x) = 0, p_0(x) = 1, \lambda_n = \frac{(x p_{n-1}, p_{n-1})}{\|p_{n-1}\|^2}, \mu_n = \frac{\|p_{n-1}\|^2}{\|p_{n-2}\|^2}, \mu_1 = 0$ .

**Preuve:** On va calculer  $p_n(x) - xp_{n-1}(x)$  pour montrer que cela peut s'écrire comme indiqué dans la proposition.

Par construction  $p_n - xp_{n-1} \in V_{n-1}$ , et peut donc s'écrire comme  $\sum_{i=0}^{n-1} \alpha_i p_i$ . On fait alors le produit scalaire par  $p_j$  avec  $j \leq n-1$ . On obtient alors  $\alpha_j(p_j, p_j) = -(xp_{n-1}, p_j) = -(p_{n-1}, xp_j)$  par orthogonalité. Ce dernier terme est nul sauf pour  $j = n-2$  ou  $j = n-1$ . Pour  $j = n-2$ , on obtient  $\alpha_{n-2} \|p_{n-2}\|^2 = -(p_{n-1}, xp_{n-2}) = -(p_{n-1}, p_{n-1}) + (p_{n-1}, p_{n-1} - xp_{n-2}) = -\|p_{n-1}\|^2$ .

Pour  $j = n-1$ , on obtient  $\alpha_{n-1} \|p_{n-1}\|^2 = -(p_{n-1}, xp_{n-1})$ .

D'où le résultat de la proposition avec  $\lambda_n = -\alpha_{n-1}$  et  $\mu_n = -\alpha_{n-2}$ . ■

**Théorème 1.8** Soit  $H$  un espace de Hilbert tel que  $V_n \subset H, \forall n \in \mathbb{N}$ . Il existe un unique polynôme de meilleure approximation  $u_n \in V_n$  caractérisé par:

$$(u_n - f, q) = 0, \quad \forall q \in V_n, \quad (1.63)$$

$$\text{avec } u_n = \sum_{k=0}^n \frac{(f, p_k)}{\|p_k\|^2} p_k \text{ et } \|f - u_n\|^2 = \|f\|^2 - \sum_{k=0}^n \frac{(f, p_k)^2}{\|p_k\|^2}.$$

Les  $p_k$  sont des polynômes orthogonaux.

**Preuve:** Commençons par montrer que l'équation (1.63) permet effectivement de caractériser la meilleure approximation. Soit  $q$  un polynôme quelconque de  $V_n$ . On a alors:

$$\|f - q\|^2 = \|f - u_n + u_n - q\|^2 = \|f - u_n\|^2 + 2(f - u_n, u_n - q) + \|u_n - q\|^2. \quad (1.64)$$

On en déduit:

$$\|f - u_n\|^2 = \|f - q\|^2 - \|u_n - q\|^2 - 2(f - u_n, u_n - q). \quad (1.65)$$

Et la caractérisation entraîne donc:

$$\|f - u_n\| \leq \|f - q\|, \quad \forall q \in V_n \quad (1.66)$$

Montrons l'unicité. Supposons qu'il existe un autre polynôme  $v_n$  vérifiant la même caractérisation:

$$(u_n - f, q) = (v_n - f, q) = 0, \quad \forall q \in V_n. \quad (1.67)$$

Si on choisit en particulier  $q = u_n - v_n$ , cela entraîne  $\|u_n - v_n\| = 0$  et donc  $u_n = v_n$ .

Enfin, on obtient la décomposition de  $u_n$  sur la base des polynômes orthogonaux  $p_j$  en considérant pour  $q$  chacun des  $p_j$

:

$$(f - \sum_{k=0}^n a_k p_k, p_j) = 0, \quad \forall j \leq n. \quad (1.68)$$

■

**Exemples de quelques familles de polynômes orthogonaux classiques:**

On prend  $H = L_w^2(a, b)$  muni du produit scalaire  $(f, g) = \int_a^b f(x)g(x)w(x)dx$  où les bornes  $a$  et  $b$  peuvent être finies ou non et le poids  $w(x)$  est une fonction continue strictement positive sur  $]a, b[$ .

- Legendre:

$$a = -1, \quad b = 1, \quad w(x) = 1, \quad p_n(x) = \frac{d^n}{dx^n} \left( \frac{(x^2 - 1)^n}{2^n n!} \right) \quad (1.69)$$

- Tchebycheff 1:

$$a = -1, \quad b = 1, \quad w(x) = \frac{1}{\sqrt{1-x^2}}, \quad p_n(x) = \cos(n\theta), \quad \text{où } x = \cos \theta \quad (1.70)$$

- Tchebycheff 2:

$$a = -1, \quad b = 1, \quad w(x) = \sqrt{1-x^2}, \quad p_n(x) = \frac{\sin((n+1)\theta)}{\sqrt{1-x^2}}, \quad \text{où } x = \cos \theta \quad (1.71)$$

- Hermite:

$$a = -\infty, \quad b = +\infty, \quad w(x) = e^{-x^2}, \quad p_n(x) = (-1)^n e^{x^2} \frac{d^n}{dx^n} (e^{-x^2}) \quad (1.72)$$

- Laguerre:

$$a = 0, \quad b = +\infty, \quad w(x) = e^{-x}, \quad p_n(x) = \frac{e^x}{n!} \frac{d^n}{dx^n} (x^n e^{-x}) \quad (1.73)$$

**Remarque:**

Il existe des polynômes orthogonaux discrets correspondant à un produit scalaire discret  $(p, q) = \sum_k p(x_k)q(x_k)w_k$  calculé sur une discrétisation de l'intervalle de définition. Ces polynômes ne sont en général pas calculés par le procédé d'orthogonalisation de Gram-Schmidt mais par la formule de récurrence (1.62).

## 1.5 Approximation au sens des moindres carrés discrets

On dispose d'une suite de  $n$  données (ou mesures) expérimentales  $(x_k, y_k)$ , et on se propose de déterminer une relation de la forme  $y = f(x)$ , à partir de ces informations. On ne cherche pas ici le polynôme d'interpolation, mais à approcher ces points par un polynôme de faible degré, ou un élément d'un espace vectoriel de faible dimension.

Soit  $V$  un espace vectoriel de dimension finie. Le problème consiste à chercher  $f \in V$  vérifiant:

$$\sum_{k=1}^n |y_k - f(x_k)|^2 w_k = \inf_{p \in V} \sum_{k=1}^n |y_k - p(x_k)|^2 w_k \quad (1.74)$$

où  $w_k$  sont des poids permettant de donner plus d'importance à certaines régions. On dira alors que  $f$  réalise la meilleure approximation des données dans  $V$  au sens des moindres carrés discrets. Le cas continu où la somme est remplacée par une intégrale rentre dans le cadre décrit dans le paragraphe précédent. Le choix de l'espace d'approximation  $V$  est très important car il détermine la "nature" de l'approximation obtenue.

Soit  $V_m$  un espace vectoriel de dimension finie  $m$  et de base  $\{\varphi_j\}_{j=1,\dots,m}$ . On cherche alors à minimiser la fonctionnelle

$$J(\alpha) = \sum_{k=1}^n |y_k - \sum_{j=1}^m \alpha_j \varphi_j(x_k)|^2 w_k. \quad (1.75)$$

Par souci de simplification, nous ne traiterons dans la suite que le cas  $w_k = 1, \forall k$ . Le cas général se déduit facilement du cas particulier. Soit  $A$  la matrice de taille  $(n, m)$  de coefficients  $a_{ij} = \varphi_j(x_i)$ . La fonctionnelle  $J(\alpha)$  à minimiser peut se réécrire:

$$J(\alpha) = \|A\alpha - y\|_2^2 \quad (1.76)$$

où  $\alpha = {}^t(\alpha_1, \dots, \alpha_m)$  et  $y = {}^t(y_1, \dots, y_n)$  sont des vecteurs colonnes.

**Théorème 1.9** *On suppose que  $m \leq n$ . Alors  $\alpha$  réalise le minimum de  $J$  si et seulement si  ${}^tAA\alpha = {}^tAy$ . Ce système admet une unique solution si  $\text{rang}(A) = \text{rang}({}^tAA) = m$ .*

**Preuve:** La fonctionnelle  $J(\alpha) = (A\alpha - y, A\alpha - y)$  est quadratique convexe de gradient  $\nabla J(\alpha) = 2({}^tAA\alpha - {}^tAy)$ . Le vecteur  $\alpha$  minimise la fonctionnelle  $J$  si et seulement si  $\nabla J(\alpha) = 0$ , c'est-à-dire si et seulement si  ${}^tAA\alpha = {}^tAy$ .

Soit  $y \in \mathbb{R}^n$ . Alors  $y$  peut s'écrire  $y = y_1 + y_2$  avec  $y_1 \in \text{Im}(A)$  et  $y_2 \in \text{Ker}({}^tA)$ . Il existe alors  $\alpha$  tel que  $y_1 = A\alpha$ . Ce qui implique  ${}^tAA\alpha = {}^tAy_1 = {}^tAy$ .

Si  $\text{rang}({}^tAA) = \dim \text{Im}({}^tAA) = m$  alors  ${}^tAA$  est inversible et  $\alpha = ({}^tAA)^{-1}{}^tAy$  est la solution unique.

De plus,  $\text{rang}(A) = \dim \text{Im}(A) = m - \dim \text{Ker}(A) = \text{rang}({}^tAA) = m - \dim \text{Ker}({}^tAA)$  car les noyaux sont égaux. ■

Si  $\text{rang}(A) < m$ , il y a plusieurs solutions mais il en existe une de norme minimale. La résolution se fait ensuite avec n'importe quelle méthode matricielle d'inversion de matrice.

**Exemple: la droite de régression linéaire:**

**Proposition 1.7** *Soit  $n \geq 2$ , et  $n$  points  $(x_k, y_k), k = 1, \dots, n$  distincts. Alors il existe une droite unique  $f(x) = a^*x + b^*$  qui réalise le minimum de la fonctionnelle  $J(a, b) = \sum_{k=1}^n |y_k - ax_k - b|^2$ .*

**Preuve:** Le minimum est atteint en  $(a^*, b^*)$  solution de:

$$\frac{\partial J}{\partial a}(a^*, b^*) = \frac{\partial J}{\partial b}(a^*, b^*) = 0, \quad (1.77)$$

ce qui conduit aux équations:

$$\begin{aligned} a^* \left( \sum_{i=1}^n x_i^2 \right) + b^* \left( \sum_{i=1}^n x_i \right) &= \sum_{i=1}^n x_i y_i \\ a^* \left( \sum_{i=1}^n x_i \right) + n b^* &= \sum_{i=1}^n y_i \end{aligned} \tag{1.78}$$

Il s'agit d'un système linéaire dont le déterminant  $n \left( \sum_{i=1}^n x_i^2 \right) - \left( \sum_{i=1}^n x_i \right)^2$  est strictement positif (utiliser l'inégalité de Cauchy-Schwartz). Ce qui assure l'unicité de la solution. ■

Le même résultat peut s'écrire en considérant des points avec pondérations.

# Chapter 2

## Dérivation et intégration numériques

### 2.1 Dérivation numérique

Soit  $f$  une fonction de  $\mathbb{R}$  dans  $\mathbb{R}$  que l'on suppose dérivable. On veut calculer une approximation du nombre:

$$f'(x_0) = \lim_{h \rightarrow 0} \frac{f(x_0 + h) - f(x_0)}{h}. \quad (2.1)$$

On peut obtenir une approximation grossière de  $f'(x_0)$  en prenant pour  $h$  fixé:  $\frac{f(x_0 + h) - f(x_0)}{h}$ . Cette approximation nécessite la connaissance de  $f$  en  $x_0$  et  $x_0 + h$ . L'idée est alors d'approcher  $f$  par un polynôme d'interpolation et de considérer sa dérivée comme une approximation de la dérivée de  $f$ . Une autre solution consiste à utiliser un développement de Taylor de  $f$ . En combinant les termes de développements de Taylor calculés en différents points, il sera alors possible de calculer la dérivée de  $f$  selon différentes formules.

#### **Théorème 2.1 Formule de Taylor**

Soient  $E$  et  $F$  deux espaces vectoriels normés,  $\Omega$  un ouvert de  $E$ , et  $(x, y) \in \Omega^2$  tels que  $x + t(y - x) \in \Omega, \forall t \in [0, 1]$ . Si  $f \in C^{p+1}(\Omega, F)$ ,  $p \in \mathbb{N}$  alors:

$$f(y) = f(x) + f'(x)(y - x) + \dots + \frac{1}{p!} f^{(p)}(x)(y - x, \dots, y - x) \quad (2.2)$$

$$+ \int_0^1 \frac{(1-t)^p}{p!} f^{(p+1)}(x + t(y-x))(y-x, \dots, y-x) dt. \quad (2.3)$$

**Proposition 2.1** *Le reste intégral dans la formule précédente est en  $O(\|y - x\|^{p+1})$ :*

$$f(y) = f(x) + f'(x)(y - x) + \dots + \frac{1}{p!} f^{(p)}(x)(y - x, \dots, y - x) + O(\|y - x\|^{p+1}). \quad (2.4)$$



La méthode consiste à approcher les dérivées aux points  $x_k$ ,  $0 \leq k \leq n$  espacés du pas  $h_k = x_k - x_{k-1}$ ,  $0 \leq k \leq n$ . Si les points sont équidistants, on note  $h$  le pas, et on a :

$$\begin{aligned} f(x_k + 2h) &= f(x_k) + f'(x_k) \frac{2h}{1!} + f''(x_k) \frac{4h^2}{2!} + f'''(x_k) \frac{8h^3}{3!} + \cdots + O((2h)^{p+1}) \\ f(x_k + h) &= f(x_k) + f'(x_k) \frac{h}{1!} + f''(x_k) \frac{h^2}{2!} + f'''(x_k) \frac{h^3}{3!} + \cdots + O((h)^{p+1}) \\ f(x_k - h) &= f(x_k) - f'(x_k) \frac{h}{1!} + f''(x_k) \frac{h^2}{2!} - f'''(x_k) \frac{h^3}{3!} + \cdots + O((h)^{p+1}) \\ f(x_k - 2h) &= f(x_k) - f'(x_k) \frac{2h}{1!} + f''(x_k) \frac{4h^2}{2!} - f'''(x_k) \frac{8h^3}{3!} + \cdots + O((2h)^{p+1}) \\ &\vdots = \vdots \end{aligned}$$

dont les combinaisons linéaires permettent de donner des approximations centrées ou décentrées des dérivées successives au point  $x_k$  :

$$f'(x_k) = \frac{f(x_{k+1}) - f(x_k)}{h} + O(h) \quad (2.5)$$

$$f'(x_k) = \frac{f(x_k) - f(x_{k-1})}{h} + O(h) \quad (2.6)$$

$$f'(x_k) = \frac{f(x_{k+1}) - f(x_{k-1})}{2h} + O(h^2) \quad (2.7)$$

$$f'(x_k) = \frac{-f(x_{k+2}) + 8f(x_{k+1}) - 8f(x_{k-1}) + f(x_{k-2})}{12h} + O(h^4) \quad (2.8)$$

$$f'(x_k) = \frac{-f(x_{k+2}) + 4f(x_{k+1}) - 3f(x_{k-1})}{2h} + O(h^2) \quad (2.9)$$

$$f'(x_k) = \frac{3f(x_k) - 4f(x_{k-1}) + f(x_{k-2})}{2h} + O(h^2) \quad (2.10)$$

$$f''(x_k) = \frac{f(x_{k+1}) - 2f(x_k) + f(x_{k-1}))}{h^2} + O(h^2) \quad (2.11)$$

$$f''(x_k) = \frac{-f(x_{k+2}) + 16f(x_{k+1}) - 30f(x_k) + 16f(x_{k-1}) - f(x_{k-2}))}{12h^2} + O(h^4) \quad (2.12)$$

$\vdots$

Ces formules peuvent aussi être obtenues en utilisant la méthode basée sur l'interpolation polynomiale  $p_n(x) = \sum_{k=0}^n f(x_k)L_k(x)$  dont les dérivées sont des approximations des dérivées de  $f$ .

**Proposition 2.2** *Si  $f$  est de classe  $C^{n+1}$  sur l'intervalle  $I$  contenant les points  $x_k$  or-*

donnés, alors il existe  $c$  et  $d$  dans  $]x_0, x_n[$  tels que,  $\forall x \in I$

$$\text{Si } x = x_k, \quad f'(x_k) - p'(x_k) = \frac{1}{(n+1)!} \prod_{j=0, j \neq k}^n (x_k - x_j) f^{(n+1)}(c) \quad (2.13)$$

$$\text{Si } x \neq x_k, \quad f'(x) - p'(x) = \frac{1}{(n+1)!} L'(x) f^{(n+1)}(c) + \frac{1}{(n+2)!} L(x) f^{(n+2)}(d) \quad (2.14)$$

où  $L(x) = \prod_{k=0}^n (x - x_k)$  et  $p(x)$  désigne le polynôme d'interpolation de Lagrange.

**Preuve:** On a vu (1.7) que  $f(x) - p(x) = \frac{1}{(n+1)!} L(x) f^{(n+1)}(c)$  où  $c$  dépend de  $x$ . On note  $g(x) = \frac{1}{(n+1)!} f^{(n+1)}(c)$ . On calcule la dérivée et on obtient:  $f'(x) - p'(x) = L'(x)g(x) +$

$L(x)g'(x)$ . Il suffit alors de remarquer que  $L(x_k) = 0$  et que  $L'(x_k) = \prod_{j=0, j \neq k}^n (x_k - x_j)$ . ■

Ces résultats peuvent s'étendre aux calculs des dérivées supérieures.

## 2.2 Intégration numérique

Soit l'intégrale  $I(f) = \int_a^b f(x)dx$  avec  $b > a$ . Nous souhaitons trouver une valeur approchée de cette intégrale au moyen d'une somme finie. Pour cela, nous allons chercher à construire des formules de quadrature.

**Définition 2.1** On appelle formule de quadrature à  $n+1$  points une formule du type:

$$I_{n+1}(f) = \sum_{k=0}^n A_k^n f(x_k) \quad (2.15)$$

où les coefficients  $A_k^n$  sont indépendants de la fonction  $f$ , et les points  $x_k$  sont des points de l'intervalle  $[a, b]$ .

On note  $E_{n+1}(f) = I(f) - I_{n+1}(f)$  l'erreur de l'approximation de  $I(f)$  par  $I_{n+1}(f)$ .

**Définition 2.2** On dit que la formule de quadrature est de degré  $m$  si  $E_{n+1}(p) = 0$  pour tout polynôme  $p$  de degré inférieur ou égal à  $m$ .

### 2.2.1 Formule de quadrature du type interpolation

#### Formules globales:

Nous avons vu au chapitre précédent que nous pouvons approcher une fonction  $f$  par un polynôme d'interpolation  $p_n(x) = \sum_{k=0}^n f(x_k)L_k(x)$ . L'idée va donc être de remplacer

le calcul de  $I(f)$  par  $\int_a^b p_n(x)dx$ . Cela se justifie pour les deux raisons suivantes:

- l'intégration de polynômes est très simple, et ne nécessite que les quatre opérations élémentaires,
- dans la pratique, assez souvent, on ne connaît pas la forme analytique de  $f$ , mais seulement sa valeur en certains points  $x_k$ .

**Définition 2.3** On appelle formule de quadrature de type interpolation la formule obtenue par le calcul de  $\int_a^b p_n(x)dx$  pour approcher  $\int_a^b f(x)dx$ ,  $p_n(x)$  étant un polynôme d'interpolation de  $f$ .

**Proposition 2.3** Une formule de quadrature à  $n + 1$  points est de degré au moins  $n$  si et seulement si elle est de type interpolation à  $n + 1$  points.

**Preuve:** Si la formule est du type interpolation à  $n+1$  points, alors  $\forall p \in \mathbb{P}_n$  (espace des polynômes de degré  $n$ ),  $p$  est égal à son polynôme d'interpolation à  $n + 1$  points. Donc  $E_{n+1}(p) = 0$ ,  $\forall p \in \mathbb{P}_n$ . Réciproquement, si la formule est de degré au moins  $n$ , on a :

$$\begin{aligned} \sum_{k=0}^n A_k^n &= b - a \\ \sum_{k=0}^n A_k^n x_k &= \int_a^b x dx \\ &\vdots = \vdots \\ \sum_{k=0}^n A_k^n x_k^n &= \int_a^b x^n dx \end{aligned}$$

C'est un système à  $n + 1$  équations et  $n + 1$  inconnues qui admet une solution unique (matrice de Vandermonde de déterminant non nul), et les coefficients  $A_k^n = \int_a^b L_k(x)dx$ ,  $k = 0, \dots, n$  constituent une solution. ■

**Exemple:** Les formules de Newton-Cotes.

On construit  $n+1$  points équidistants en posant  $h = \frac{b-a}{n}$ :  $x_0 = a$ ,  $x_1 = a+h, \dots, x_n = b$ . On obtient alors les formules de Newton-Cotes fermées. (Si on exclut les extrémités  $a$  et  $b$  et on pose  $h = \frac{b-a}{n+2}$ , on obtient alors les formules dites ouvertes)

Formule de Newton-Cotes fermée de degré 1 et 2:

$$\int_a^b f(x)dx = (b-a) \left[ \frac{1}{2}f(a) + \frac{1}{2}f(b) \right] \quad (\text{formule des trapèzes}) \quad (2.16)$$

$$\int_a^b f(x)dx = (b-a) \left[ \frac{1}{6}f(a) + \frac{4}{6}f\left(\frac{a+b}{2}\right) + \frac{1}{6}f(b) \right] \quad (\text{formule de Simpson}) \quad (2.17)$$

**Théorème 2.2** *Estimation de l'erreur pour les formules de Newton-Cotes.*

Si  $f \in C^l([a, b])$  avec  $l = n + 2$  si  $n$  est pair et  $l = n + 1$  si  $n$  est impair, alors il existe  $c \in [a, b]$  t.q.:

$$E_{n+1} = \frac{1}{(n+2)!} \int_0^n t \prod_{j=0}^n (t-j) dt \quad h^{n+3} f^{(n+2)}(c), \quad \text{si } n \text{ est pair} \quad (2.18)$$

$$E_{n+1} = \frac{1}{(n+1)!} \int_0^n \prod_{j=0}^n (t-j) dt \quad h^{n+2} f^{(n+1)}(c), \quad \text{si } n \text{ est impair} \quad (2.19)$$

Une formule de quadrature basée sur l'interpolation peut présenter les mêmes problèmes lorsque l'intervalle est grand (phénomène de Runge). On peut éviter cela en utilisant des formules d'interpolation locales. On choisit un degré d'interpolation faible qu'on utilise sur chaque segment  $[x_k, x_{k+1}]$ .

### Formules locales:

#### **Polynômes de degré 1 (formule des trapèzes):**

Sur chaque segment  $[x_k, x_{k+1}]$  on écrit une formule d'interpolation de degré 1, puis on fait la somme sur tous les segments. Soit  $p_1(x)$  le polynôme d'interpolation de Lagrange de degré 1 sur l'intervalle  $[x_k, x_{k+1}]$ :

$$p_1(x) = f(x_k) \frac{x - x_{k+1}}{x_k - x_{k+1}} + f(x_{k+1}) \frac{x - x_k}{x_{k+1} - x_k} \quad (2.20)$$

On calcule l'intégrale de  $p_1(x)$  sur l'intervalle  $[x_k, x_{k+1}]$ :

$$\begin{aligned} I_{1,k} &= \int_{x_k}^{x_{k+1}} p_1(x) dx \\ &= (x_{k+1} - x_k) \left( \frac{f(x_k) + f(x_{k+1})}{2} \right) \end{aligned}$$

puis on effectue la somme de tous les segments:

$$I_1 = \sum_{k=0}^{n-1} I_{1,k} = h \sum_{k=0}^{n-1} \left( \frac{f(x_k) + f(x_{k+1})}{2} \right) \quad (2.21)$$

si  $h = x_{k+1} - x_k, \forall k$ .

**Calcul de l'erreur:** Il suffit d'intégrer l'erreur d'interpolation correspondante et de faire la somme sur tous les segments de l'intervalle. Sur  $[x_k, x_{k+1}]$ , on a  $f(x) - p_1(x) = \frac{1}{2}(x - x_k)(x - x_{k+1})f''(c_k)$ . Cela entraîne:

$$I - I_1 = \frac{1}{2} \sum_{k=0}^{n-1} f''(c_k) \int_{x_k}^{x_{k+1}} (x - x_k)(x - x_{k+1}) dx. \quad (2.22)$$

On effectue alors un changement de variable  $x = x_k + t(x_{k+1} - x_k)$ :

$$\begin{aligned} I - I_1 &= \frac{1}{2} \sum_{k=0}^{n-1} f''(c_k) \int_0^1 t(t-1)h^3 dt \\ &= -\frac{h^3}{12} \sum_{k=0}^{n-1} f''(c_k). \end{aligned}$$

On obtient donc une majoration de l'erreur:

$$|I - I_1| \leq (b-a) \frac{h^2}{12} \max_{x \in [a,b]} |f''(x)|. \quad (2.23)$$

### Polynômes de degré 2 (formule de Simpson):

On note  $x_{k+1/2}$  le point milieu de l'intervalle  $[x_k, x_{k+1}]$ . Soit  $p_2(x)$  le polynôme d'interpolation de degré 2 sur l'intervalle  $[x_k, x_{k+1}]$  de longueur  $h$ :

$$\begin{aligned} p_2(x) &= 2f(x_k) \frac{(x - x_{k+1/2})(x - x_{k+1})}{h^2} - 4f(x_{k+1/2}) \frac{(x - x_k)(x - x_{k+1})}{h^2} \\ &\quad + 2f(x_{k+1}) \frac{(x - x_k)(x - x_{k+1/2})}{h^2}. \end{aligned}$$

On calcule alors l'intégrale de  $p_2(x)$ :

$$\begin{aligned} I_{2,k} &= \int_{x_k}^{x_{k+1}} p_2(x) dx \\ &= 2f(x_k) \int_0^1 (t - 1/2)(t - 1) h dt - 4f(x_{k+1/2}) \int_0^1 t(t - 1) h dt + 2f(x_{k+1}) \int_0^1 t(t - 1/2) h dt \\ &= \frac{h}{6} (f(x_k) + 4f(x_{k+1/2}) + f(x_{k+1})) \end{aligned}$$

On déduit ensuite l'intégrale totale:

$$\begin{aligned} I_2 &= \sum_{k=0}^{n-1} I_{2,k} \\ &= h \sum_{k=0}^{n-1} \frac{f(x_k) + 4f(x_{k+1/2}) + f(x_{k+1})}{6} \end{aligned}$$

On effectue alors l'évaluation de l'erreur de la même manière que pour la formule des trapèzes, et on trouve la majoration de l'erreur suivante:

$$|I - I_2| \leq (b-a) \frac{h^4}{2880} \max_{x \in [a,b]} |f^{(4)}(x)|. \quad (2.24)$$

On gagne donc alors un ordre de grandeur dans l'erreur en utilisant  $n = 2$  au lieu de  $n = 1$ .

**Théorème 2.3** *Si le nombre de points d'interpolation est  $n + 1$ , dans les formules de Newton-Cotes fermées composites avec  $n$  pair, alors l'erreur est en  $h^{n+2}$ . Par contre, si  $n$  est impair, l'erreur est en  $h^{n+1}$ .*

## 2.2.2 Accélération de la convergence: méthode de Romberg

On veut calculer

$$I = \int_a^b f(x) dx \quad (2.25)$$

Nous avons étudié jusqu'à présent des méthodes qui donnent une approximation  $I(h)$  de  $I$  où  $h = \frac{b-a}{n}$  est le pas de discrétisation, et avec l'hypothèse que pour des fonctions suffisamment régulières, on a  $\lim_{h \rightarrow 0} I(h) = I(0) = I$ .

Nous allons voir maintenant un algorithme permettant d'améliorer les résultats des méthodes précédentes. Supposons que l'on puisse faire un développement limité de  $I(h)$  en 0, c'est-à-dire que l'on puisse écrire:

$$I(h) = I(0) + a_1 h + \frac{a_2}{2!} h^2 + \dots + h^p \epsilon(p). \quad (2.26)$$

On a alors:

$$I\left(\frac{h}{2}\right) = I(0) + a_1 \frac{h}{2} + \frac{a_2}{4 \cdot 2!} h^2 + \dots + h^p \epsilon(p). \quad (2.27)$$

En prenant une combinaison linéaire de  $I(h)$  et  $I\left(\frac{h}{2}\right)$ , on pourra obtenir une approximation numérique de  $I(0)$  meilleure que celle de  $I(h)$ . Cette constatation est à la base de la méthode de Romberg.

### Exemple:

On considère la formule du trapèze (interpolation sur 2 points avec un polynôme de degré 1):

$$\begin{aligned} I(h) &= \frac{h}{2} (f(a) + f(b)) \\ I\left(\frac{h}{2}\right) &= \frac{h}{4} \left( f(a) + 2f\left(\frac{a+b}{2}\right) + f(b) \right) \\ \frac{4I\left(\frac{h}{2}\right) - I(h)}{3} &= \frac{h}{6} \left( f(a) + 4f\left(\frac{a+b}{2}\right) + f(b) \right) \end{aligned}$$

On retrouve alors la formule de Simpson. On peut alors à nouveau recommencer le procédé à partir de cette dernière formule. D'une manière générale, on définit le processus itératif:

$$\begin{aligned} R(0, 0) &= \frac{1}{2} (b-a)(f(a) + f(b)) \\ R(n, 0) &= \frac{1}{2} R(n-1, 0) + h_n \sum_{k=1}^{2^{n-1}} f(a + (2k-1)h_n) \\ R(n, m) &= R(n, m-1) + \frac{1}{4^m - 1} (R(n, m-1) - R(n-1, m-1)) \end{aligned}$$

où  $n \geq 1$ ,  $m \geq 1$ ,  $h_n = \frac{b-a}{2^n}$ . L'erreur d'approximation de l'intégrale pour  $R(n, m)$  est alors en  $O(h_n^{2^{m+1}})$ .

La méthode de Romberg converge très vite et il suffit en général de calculer quelques itérations pour obtenir une bonne approximation de l'intégrale.

### 2.2.3 Formule d'Euler - Mac Laurin

C'est une formule d'approximation de l'intégrale construite par intégration par parties et qui s'exprime à l'aide des polynômes de Bernoulli.

**Définition 2.4** Les polynômes de Bernoulli sont définis par:

$$B_n(x) = \sum_{k=0}^n \binom{n}{k} b_k x^{n-k}, \quad (2.28)$$

pour  $n \geq 0$ , et où les coefficients  $b_k$  désignent les nombres de Bernoulli. Ces derniers peuvent être définis par le développement:

$$\frac{x}{e^x - 1} = \sum_{n=0}^{+\infty} b_n \frac{x^n}{n!}. \quad (2.29)$$

On a donc  $b_n = B_n(0)$ , et on peut montrer que  $b_{2k+1} = 0$  pour  $k \geq 1$ .

De même, les polynômes de Bernoulli peuvent s'obtenir aussi à partir d'un développement:

$$\frac{te^{xt}}{e^t - 1} = \sum_{n=0}^{\infty} B_n(x) \frac{t^n}{n!}. \quad (2.30)$$

Les polynômes de Bernoulli peuvent aussi se définir à partir de la relation de récurrence:

$$\begin{aligned} B_0(x) &= 1, \\ B'_n(x) &= nB_{n-1}(x) \text{ et } \int_0^1 B_n(x) dx = 0 \text{ pour } n \geq 1. \end{aligned} \quad (2.31)$$

Les premiers polynômes sont:

$$\begin{aligned} B_1(x) &= x - 1/2, & B_2(x) &= x^2 - x + 1/6, \\ B_3(x) &= x^3 - \frac{3}{2}x^2 + \frac{1}{2}x, & B_4(x) &= x^4 - 2x^3 + x^2 - \frac{1}{30}, \dots \end{aligned}$$

**Idée de base de la formule d'Euler - Mac Laurin:** Supposons que nous souhaitons calculer l'intégrale d'une fonction  $f(x)$  sur l'intervalle  $[0, 1]$ . Nous allons effectuer une intégration par parties qui va permettre de réécrire l'intégrale à calculer:

$$\begin{aligned} \int_0^1 f(x) dx &= \int_0^1 1 \cdot f(x) dx \\ &= [xf(x)]_0^1 - \int_0^1 xf'(x) dx \\ &= f(1) - \int_0^1 \left(x - \frac{1}{2}\right) f'(x) dx - \frac{1}{2} \int_0^1 f'(x) dx \\ &= \frac{1}{2} (f(0) + f(1)) - \int_0^1 \left(x - \frac{1}{2}\right) f'(x) dx \end{aligned}$$

où on voit apparaître  $B_1(x) = x - \frac{1}{2}$  le premier polynôme de Bernoulli. On peut alors effectuer une nouvelle intégration par parties de la dernière intégrale (en utilisant la formule de récurrence):

$$\begin{aligned} \int_0^1 B_1(x) f'(x) dx &= \int_0^1 \frac{1}{2} B_2'(x) f'(x) dx \\ &= \left[ \frac{1}{2} B_2(x) f'(x) \right]_0^1 - \int_0^1 \frac{1}{2} B_2(x) f''(x) dx \end{aligned}$$

On obtient alors la formule d'intégration:

$$\int_0^1 f(x) dx = \frac{1}{2} (f(0) + f(1)) - \frac{B_2(0)}{2} (f'(1) - f'(0)) + \int_0^1 \frac{1}{3!} B_3'(x) f''(x) dx. \quad (2.32)$$

On peut répéter ces développements par parties, et on obtient alors la formule générale:

$$\begin{aligned} \int_0^1 f(x) dx &= \frac{1}{2} (f(0) + f(1)) + \sum_{k=1}^{n-1} (-1)^k \frac{B_{k+1}(0)}{(k+1)!} (f^{(k)}(1) - f^{(k)}(0)) \\ &\quad + (-1)^n \int_0^1 \frac{1}{(n+1)!} B_{n+1}'(x) f^{(n)}(x) dx. \end{aligned} \quad (2.33)$$

**Proposition 2.4** *Soit  $f$  de classe  $C^{2m+2}$  sur  $[0, 1]$ , alors il existe  $c \in ]0, 1[$  tel que:*

$$\int_0^1 f(x) dx = \frac{1}{2} (f(0) + f(1)) - \sum_{l=1}^m \frac{b_{2l}}{(2l)!} (f^{(2l-1)}(1) - f^{(2l-1)}(0)) + \frac{b_{2m+2}}{(2m+2)!} f^{(2m+2)}(c). \quad (2.34)$$

**Preuve:** Les deux premiers termes de cette formule sont obtenus par construction en tenant compte du fait que  $b_{2k+1} = 0$  pour  $k \geq 1$ . Ne figurent donc ici que les termes d'indices pairs. Nous allons maintenant déterminer le troisième terme, le reste  $r$ , qui s'écrit:

$$\begin{aligned} r &= \int_0^1 \frac{1}{(2m+2)!} B_{2m+2}'(x) f^{(2m+1)}(x) dx \\ &= \frac{1}{(2m+2)!} [B_{2m+2}(x) f^{(2m+1)}(x)]_0^1 - \frac{1}{(2m+2)!} \int_0^1 B_{2m+2}(x) f^{(2m+2)}(x) dx \\ &= -\frac{1}{(2m+2)!} \int_0^1 (B_{2m+2}(x) - B_{2m+2}(0)) f^{(2m+2)}(x) dx \end{aligned}$$

Or  $B_{2m+2}(x) - B_{2m+2}(0)$  est un polynôme de signe constant sur  $[0, 1]$ . On peut donc appliquer le théorème de la valeur moyenne:  $\exists c \in ]0, 1[$  t.q.

$$r = -\frac{1}{(2m+2)!} f^{(2m+2)}(c) \int_0^1 (B_{2m+2}(x) - B_{2m+2}(0)) dx \quad (2.35)$$



On obtient alors le résultat attendu car  $\int_0^1 B_{2m+2}(x) dx = 0$ . ■

La formule générale pour un intervalle  $[a, b]$  quelconque s'obtient facilement par un changement de variable  $x = a + t(b - a)$ ,  $t \in [0, 1]$ . En introduisant une subdivision de  $[a, b]$  en  $n$  sous-intervalles égaux et en appliquant la formule générale à chaque segment  $[x_k, x_{k+1}]$ , on peut alors obtenir une nouvelle formule composite en calculant la somme de tous les segments.

## 2.2.4 Formules de Gauss

Dans les formules étudiées jusqu'à présent, on choisissait  $x_k = a + kh$ . On souhaite améliorer les résultats en déterminant au mieux les points de discrétisation  $x_k$  (comme dans le cas de l'interpolation polynomiale). Dans la formule de quadrature,

$$\int_a^b f(x) dx = \sum_{k=0}^n c_k f(x_k), \quad (2.36)$$

nous allons considérer les coefficients  $c_k$  et les points  $x_k$  comme inconnus, et nous souhaitons ainsi améliorer les résultats précédents en imposant à la formule d'être exacte sur  $\mathbb{P}_{n'}$  avec  $n' > n$ . Par exemple, nous pouvons remarquer que  $x_0 = \frac{a+b}{2}$  et  $c_0 = b - a$  permet d'intégrer de façon exacte les polynômes  $p \in \mathbb{P}_0$  et  $p \in \mathbb{P}_1$  (degré 1 et non 0 seulement). Dans le cas général, on a  $2(n+1)$  degrés de liberté ( $c_k$  et  $x_k$ ), et donc le choix optimal permet d'intégrer de façon exacte un polynôme de degré  $2n+1$ , et non pas  $n$  comme pour les formules de quadrature précédentes.

**Théorème 2.4** *La formule de quadrature à  $n+1$  points est de degré  $2n+1$  si et seulement si:*

1. elle est du type interpolation à  $n+1$  points,
2. les abscisses d'interpolation sont telles que,

$$L(x) = \prod_{k=0}^n (x - x_k) \quad (2.37)$$

vérifie

$$\int_a^b x^q L(x) dx = 0 \quad (2.38)$$

$$\forall 0 \leq q \leq n.$$

**Preuve:**

- Supposons la formule de degré  $2n+1$ . Alors elle est en particulier exacte pour les polynômes de degré  $n$ , et d'après un résultat précédent elle est donc du type interpolation (cela démontre (1)). Soit  $q \leq n$  un entier positif. Comme  $L(x)$  est un polynôme de degré  $n+1$ , alors

$x^q L(x)$  est un polynôme de degré inférieur ou égal à  $2n + 1$ . On applique alors la formule de quadrature à  $x^q L(x)$ :

$$\int_a^b x^q L(x) dx = \sum_{k=0}^n c_k x_k^q L(x_k) \quad (2.39)$$

Or  $L(x_k) = 0$ . Donc  $\int_a^b x^q L(x) dx = 0, \quad \forall q : 0 \leq q \leq n$ .

- Réciproquement, supposons (1) et (2).  
D'après (1), la formule de quadrature est exacte pour les polynômes de degré  $n$ . Soit  $P \in \mathbb{P}_{2n+1}$  que l'on divise par  $L$ :

$$P = LQ + R \quad (2.40)$$

avec  $d^\circ Q \leq n$  et  $d^\circ R \leq n$ .

Donc

$$\int_a^b P(x) dx = \int_a^b Q(x)L(x) dx + \int_a^b R(x) dx \quad (2.41)$$

D'après (2),  $\int_a^b Q(x)L(x) dx = 0$ , et pour  $R$  de degré inférieur à  $n$ , la formule est exacte:

$$\int_a^b R(x) dx = \sum_{k=0}^n c_k R(x_k). \quad (2.42)$$

Or  $R(x_k) = P(x_k) - L(x_k)Q(x_k) = P(x_k)$ .

Donc:

$$\int_a^b P(x) dx = \sum_{k=0}^n c_k P(x_k), \quad \forall P \in \mathbb{P}_{2n+1} \quad (2.43)$$

■

Nous venons de voir les conditions imposées aux points  $x_k$  (2.38). Nous nous posons alors les questions suivantes: Existe-t-il des points  $x_k$  vérifiant ces conditions ? Cette solution est-elle unique ? Comment la calculer ?

**Théorème 2.5** *Il existe un unique polynôme  $L(x) = \prod_{k=0}^n x - x_k$  dont les  $x_k \in [a, b]$  sont distincts et tel que:*

$$\int_a^b x^q L(x) dx = 0, \quad \forall q : 0 \leq q \leq n. \quad (2.44)$$

**Preuve:**

- Soient  $L_1$  et  $L_2$  deux polynômes vérifiant les conditions précédentes. On pose  $P(x) = L_1(x) - L_2(x)$ .  $P$  est un polynôme de degré inférieur ou égal à  $n$  car le terme en  $x^{n+1}$  s'élimine. Ce polynôme vérifie:

$$\int_a^b x^q P(x) dx = 0, \quad \forall q : 0 \leq q \leq n. \quad (2.45)$$

Or  $P$  peut s'écrire:  $P(x) = a_0 + a_1x + a_2x^2 + \dots + a_nx^n$ . On calcule alors l'intégrale de  $P^2$ :

$$\int_a^b P^2(x) dx = a_0 \int_a^b P(x) dx + a_1 \int_a^b xP(x) dx + \dots + a_n \int_a^b x^n P(x) dx = 0 \quad (2.46)$$

Or, comme l'application  $x \rightarrow P^2(x)$  est continue et positive sur  $[a, b]$ , alors  $P^2(x) = 0, \forall x \in [a, b]$ , et donc  $P = 0$ .

- On pose  $L_{n+1}(x) = L(x) = (x - x_0)(x - x_1) \dots (x - x_n)$ . Nous allons construire par récurrence la suite des polynômes  $L_n(x)$  vérifiant les conditions du théorème (orthogonalisation de Gram-Schmidt). Ces polynômes doivent donc vérifier:

$$\int_a^b L_k(x)L_n(x) dx = 0, \quad \forall k < n \quad (2.47)$$

avec le coefficient de plus haut degré égal à 1.

On a donc  $L_0(x) = 1$ .  $L_1(x)$  sera de la forme  $L_1(x) = x + a_0L_0(x)$ .

Or,

$$\int_a^b L_0(x)L_1(x) dx = \int_a^b x dx + a_0 \int_a^b dx = 0. \quad (2.48)$$

On peut alors calculer  $a_0$ :

$$a_0 = -\frac{1}{b-a} \int_a^b x dx \quad (2.49)$$

Supposons  $L_0, L_1, \dots, L_n$  construits, et posons,

$$L_{n+1}(x) = x^{n+1} + a_nL_n(x) + a_{n-1}L_{n-1}(x) + \dots + a_0L_0(x). \quad (2.50)$$

On doit avoir pour  $j \leq n$ :

$$\int_a^b L_{n+1}(x) \cdot L_j(x) dx = 0. \quad (2.51)$$

On remplace  $L_{n+1}(x)$  par son expression, on multiplie par  $L_j(x)$  et on intègre:

$$\int_a^b x^{n+1}L_j(x) dx + a_j \int_a^b L_j^2(x) dx = 0. \quad (2.52)$$

On obtient alors  $a_j$ :

$$a_j = \frac{-\int_a^b x^{n+1}L_j(x) dx}{\int_a^b L_j^2(x) dx}. \quad (2.53)$$

- Nous allons montrer que les racines de  $L_n(x) = 0$  sont réelles, simples, et comprises entre  $a$  et  $b$ .

Supposons que  $L_n(x)$  possède  $k$  racines réelles de multiplicité impaire entre  $a$  et  $b$  ( $L_n(x)$  peut aussi posséder des racines de multiplicité paire ou des racines en dehors

de  $[a, b]$ ). Nous allons alors montrer que  $k = n$  et que les racines sont simples. Soient  $x_1, x_2, \dots, x_k$  ces racines. On construit le polynôme  $g_k(x) = (x - x_1)(x - x_2) \cdots (x - x_k)$ . Le produit  $L_n \cdot g_k$  ne change pas de signe sur  $[a, b]$  car n'ayant que des racines de multiplicité paire. On a donc:

$$\int_a^b L_n(x) g_k(x) dx \neq 0. \quad (2.54)$$

Le cas  $k < n$  est impossible par construction de  $L_n$ . Donc  $k = n$ , et les  $n$  racines sont des racines de multiplicité 1.  $L_n(x)$  a toute ses racines réelles entre  $a$  et  $b$ . Elles sont toutes simples. ■

**Proposition 2.5** *Les formules de Gauss à  $n + 1$  points ne sont pas exactes pour les polynômes de degré supérieur à  $2n + 1$ .*

**Preuve:** Il suffit de remarquer que  $L^2(x)$  est de degré  $2n + 2$ . Si la formule était exacte pour  $L^2(x)$  on aurait:

$$\int_a^b L^2(x) dx = \sum_{k=0}^n A_k^n L^2(x_k) = 0. \quad (2.55)$$

Ce qui n'est pas possible car l'intégrale d'une fonction positive est positive. ■

**Théorème 2.6** *Si  $f$  est de classe  $C^{2n+2}$  sur  $[a, b]$ , alors il existe  $c \in ]a, b[$  tel que:*

$$E_{n+1}(f) = \frac{1}{(2n+2)!} f^{(2n+2)}(c) \int_a^b L^2(x) dx \quad (2.56)$$

**Preuve:** On considère que  $f$  est interpolée par un polynôme d'Hermite de degré  $2n + 1$ :

$$\begin{aligned} p_{2n+1}(x_k) &= f(x_k) \\ p'_{2n+1}(x_k) &= f'(x_k). \end{aligned}$$

On a alors (erreur d'interpolation d'Hermite):

$$f(x) - p_{2n+1}(x) = L^2(x) \frac{1}{(2n+2)!} f^{(2n+2)}(\xi), \quad (2.57)$$

avec  $\xi$  qui dépend de  $x$ . On intègre alors cette expression:

$$\begin{aligned} \int_a^b f(x) dx - \int_a^b p_{2n+1}(x) dx &= \int_a^b L^2(x) \frac{1}{(2n+2)!} f^{(2n+2)}(\xi) dx \\ \int_a^b f(x) dx - \sum_{k=0}^n A_k^n f(x_k) &= \int_a^b L^2(x) \frac{1}{(2n+2)!} f^{(2n+2)}(\xi) dx \end{aligned} \quad (2.58)$$

Or  $L^2(x)$  est positif et on peut donc appliquer la formule de la moyenne et on obtient alors le résultat annoncé. ■

**Remarque:** Il peut être utile de faire intervenir dans la formule une ou les deux extrémités de l'intervalle d'intégration; les autres points sont alors choisis au mieux selon le critère de Gauss. Nous obtenons alors les formules de Gauss-Radau lorsqu'une des deux extrémités est assignée, et Gauss-Lobato lorsque les deux extrémités sont assignées.

## 2.3 Méthodes de Monte-Carlo

On souhaite étendre les formules d'intégration précédentes aux dimensions supérieures. Par exemple, pour intégrer une fonction de  $\mathbb{R}^2$  dans  $\mathbb{R}$  sur le carré unité, on peut appliquer les formules de quadrature unidimensionnelles pour chacune des variables:

$$\begin{aligned} I = \int_0^1 \int_0^1 f(x, y) dx dy &\approx \int_0^1 \left( \sum_{k=0}^n A_k^n f(x_k, y) \right) dy \\ &\approx \sum_{l=0}^n A_l^n \left( \sum_{k=0}^n A_k^n f(x_k, y_l) \right) \\ &\approx \sum_{l=0}^n \sum_{k=0}^n A_l^n A_k^n f(x_k, y_l). \end{aligned}$$

La formule s'étend facilement pour les dimensions supérieures.

Cette méthode ne peut cependant plus être utilisée si l'intégration ne se fait pas sur un rectangle. On peut par contre faire des approximations sur des triangles en se ramenant au triangle unité:

$$I = \int_0^1 \int_0^{1-x} f(x, y) dy dx \tag{2.59}$$

En effet, on peut trouver des points qui permettent d'intégrer de façon exacte des polynômes de degré croissant.

Dans le cas du calcul d'une intégrale sur un domaine quelconque, on a alors la possibilité de découper le domaine en sous-domaines élémentaires sur lesquels on approche la fonction par des polynômes de degré faible.

Ces méthodes d'intégration dites *déterministes* peuvent se révéler très coûteuses dans le cas d'intégration de fonctions de  $\mathbb{R}^n$  dans  $\mathbb{R}^p$  sur des domaines quelconques. En effet, ces méthodes nécessitent de calculer les valeurs de  $f$  en chacun des points de discrétisation. Les méthodes de Monte Carlo, quant à elles, choisissent de façon aléatoire les points où la fonction doit être évaluée.

**Définition 2.5** *On appelle méthode de Monte Carlo toute méthode visant à calculer une valeur numérique, et utilisant des procédés aléatoires, c'est-à-dire des techniques probabilistes. Le nom de ces méthodes fait référence aux jeux de hasard pratiqués dans les casinos de Monte-Carlo.*

### Exemple 1: Détermination de la superficie d'un lac

Supposons que nous souhaitions connaître la superficie d'un lac de forme quelconque.

Nous dessinons alors un rectangle fictif (dont on connaît la surface) plus grand que le lac et qui contiendra celui-ci. Pour déterminer la surface du lac, on va choisir  $N$  points du rectangle de façon aléatoire. On compte ensuite le nombre de points  $n$  qui sont dans le lac. La probabilité qu'un point soit choisi dans le lac est lié au rapport entre la surface du lac et la surface du rectangle. On note  $S_l$  la surface du lac,  $S_r$  la surface du rectangle. On a alors:

$$\frac{n}{N} \approx \frac{S_l}{S_r} \quad (2.60)$$

qui nous donne une estimation de la surface du lac:  $S_l \approx \frac{n}{N} S_r$ .

Cet exemple se traduit mathématiquement par le calcul de l'intégrale de la fonction indicatrice du lac qui vaut 1 dans le lac et 0 ailleurs sur le rectangle.

### Exemple 2: Détermination de la valeur de $\pi$

Dessinons un carré de coté de longueur 2. On peut alors inscrire dans ce carré un cercle de rayon 1. On choisit alors des points au hasard dans le carré. La proportion de points situés dans le cercle par rapport au nombre de points total donnera une approximation de  $\pi/4$ .

### Algorithme:

Il n'existe pas une méthode unique dite de Monte Carlo. Le terme désigne plutôt une large classe d'approches basées sur la génération de points aléatoires. On peut cependant définir l'algorithme de base classique. On souhaite calculer l'intégrale suivante:

$$I = \int_{x_1^a}^{x_1^b} \int_{x_2^a}^{x_2^b} \cdots \int_{x_n^a}^{x_n^b} f(x_1, \dots, x_n) dx_1 \cdots dx_n \quad (2.61)$$

sur un hypercube délimité par  $[x_1^a, x_1^b] \times [x_2^a, x_2^b] \times \cdots \times [x_n^a, x_n^b]$ . On génère alors de façon aléatoire  $N$  points  $X_1, \dots, X_N$  situés dans l'hypercube, et l'intégrale peut être alors approchée par:

$$I \approx E(f; N) = V \cdot \langle f \rangle = V \frac{1}{N} \sum_{i=1}^N f(X_i). \quad (2.62)$$

où  $V$  désigne le volume de l'hypercube et  $\langle f \rangle$  la moyenne discrète de  $f$ . L'estimation de l'erreur décroît comme  $\sigma(f)/\sqrt{N}$  où  $\sigma(f)$  désigne la variance discrète de  $f$ . Autrement dit, pour réduire l'erreur d'un facteur 10, il faut multiplier par 100 le nombre de points. Cependant, des variantes existent pour optimiser cet algorithme en privilégiant des zones pour la génération des points.

# Chapter 3

## Introduction aux méthodes numériques de résolution d'équations différentielles

Pour la plupart des équations différentielles, les solutions ne s'expriment pas à partir de fonctions élémentaires ou de primitives de fonctions élémentaires. Il est alors nécessaire d'avoir recours à des méthodes numériques pour obtenir des approximations des solutions. Pour une équation différentielle  $x'(t) = f(t, x)$  où  $f$  est une application suffisamment régulière, nous allons donc chercher une approximation d'une fonction dont on connaît la dérivée en chaque point de son domaine de définition et la valeur en un point donné  $x(t_0) = x_0$ .

### 3.1 Méthodes à 1 pas

#### 3.1.1 Méthode d'Euler

Si  $t$  augmente d'un petit pas  $h$ , le graphe de la solution reste proche de sa tangente. Si l'on suit cette tangente, on aboutit au point:

$$(t_1, x_1) = (t_0 + h, x_0 + hf(t_0, x_0)) \quad (3.1)$$

Cette approximation peut se justifier en écrivant un développement à l'ordre 2 d'une solution  $u(t)$  autour du point  $t_0$ :

$$\begin{aligned} u(t_0 + h) &= u(t_0) + hu'(t_0) + \frac{1}{2}h^2u''(t_0 + \theta h) \\ &= x_0 + hf(t_0, x_0) + \frac{1}{2}h^2u''(t_0 + \theta h) \end{aligned}$$

où  $0 < \theta < 1$ . Si  $h$  est petit et  $u''$  borné, alors l'approximation de  $u(t_0 + h)$  par  $x_0 + hf(t_0, x_0)$  est de l'ordre de  $h^2$ . On recommence ensuite avec  $(t_1, x_1)$  comme condition initiale.

**Définition 3.1** Soient le rectangle  $R = [a, b] \times [c, d]$ ,  $(t_0, x_0) \in R$  une condition initiale, et  $h$  un pas de discrétisation.

On définit la suite de points  $(t_n, x_n)$  par:

$$\begin{cases} t_{n+1} &= t_n + h \\ x_{n+1} &= x_n + hf(t_n, x_n) \end{cases}$$

avec  $a = t_0 < t_1 < \dots < t_N = b$ .

On obtient alors une approximation  $x_n$  de la suite de points  $x(t_n)$ .

**Exemple:**  $x'(t) = 2t$

### 3.1.2 Schéma général

Dans les méthodes à 1 pas, le calcul de  $x_{n+1}$  se fait à partir de  $t_n, x_n, h$ .

**Définition 3.2** Une méthode à 1 pas est une méthode de la forme:

$$\begin{cases} x_{n+1} = x_n + h\phi(t_n, x_n, h) \\ x_0 = x(t_0) \end{cases}$$

Le choix d'une méthode se fait en choisissant  $\phi$ .

**Exemple:**  $\phi(t, x, h) = f(t, x)$  Euler.

Nous allons étudier les conditions qu'il faut imposer à  $\phi$  et le lien entre  $\phi$  et  $f$  pour que la méthode soit jugée bonne.

### Définition 3.3 Convergence

La méthode est convergente si:

$$\lim_{h \rightarrow 0} \max_n |x_n - x(t_n)| = 0, \quad (3.2)$$

quelle que soit la condition initiale.

### Définition 3.4 Stabilité

Soient  $\{y_n; 0 \leq n \leq N\}$  et  $\{z_n; 0 \leq n \leq N\}$  les solutions respectives des systèmes:

$$\begin{cases} y_{n+1} = y_n + h\phi(t_n, y_n, h) \\ y_0 \end{cases}$$

et

$$\begin{cases} z_{n+1} = z_n + h(\phi(t_n, z_n, h) + \epsilon_n) \\ z_0 \end{cases}$$

La méthode est dite stable s'il existe deux constantes  $M_1$  et  $M_2$  indépendantes de  $h$  telles que:

$$\max_n |y_n - z_n| \leq M_1 |y_0 - z_0| + M_2 \max_n |\epsilon_n|. \quad (3.3)$$

Cette notion de stabilité signifie qu'une petite perturbation sur les données  $(y_0, \phi)$  n'entraîne qu'une petite perturbation sur la solution et ceci indépendamment de  $h$ . Cette condition de stabilité est indispensable pour traiter numériquement une équation différentielle. Il faut aussi que la méthode approche l'équation différentielle.



**Définition 3.5 Consistance**

La méthode est dite consistante avec l'équation différentielle si:

$$\lim_{h \rightarrow 0} \left( \max_n \left| \frac{x(t_{n+1}) - x(t_n)}{h} - \phi(t_n, x(t_n), h) \right| \right) = 0, \quad (3.4)$$

pour toute solution  $x$  de  $x' = f(t, x)$ .

**Théorème 3.1** Si une méthode à 1 pas est consistante et stable, alors elle est convergente.

**Preuve:** Si la méthode est consistante, elle vérifie alors la relation:

$$y(t_{n+1}) = y(t_n) + h(\phi(t_n, y(t_n), h) + \epsilon_n), \quad (3.5)$$

où  $\lim_{h \rightarrow 0} \max_n |\epsilon_n| = 0$ . De plus, comme elle est stable, il existe  $M_2$  telle que:

$$\max_n |y_n - y(x_n)| \leq M_2 \max_n |\epsilon_n|. \quad (3.6)$$

Elle est donc convergente. ■

**Théorème 3.2** Une condition nécessaire et suffisante pour que la méthode soit consistante est que:

$$\phi(t, x, 0) = f(t, x) \quad (3.7)$$

**Preuve:** Supposons la méthode consistante. On a alors:

$$\lim_{h \rightarrow 0} \left( \max_n \left| \frac{x(t_{n+1}) - x(t_n)}{h} - \phi(t_n, x(t_n), h) \right| \right) = 0, \quad (3.8)$$

pour toute solution  $x$  de  $x' = f(t, x)$ . Cela implique que  $\forall n = 0, \dots, N$ ,

$$\lim_{h \rightarrow 0} \left| \frac{x(t_{n+1}) - x(t_n)}{h} - \phi(t_n, x(t_n), h) \right| = 0. \quad (3.9)$$

Or

$$\lim_{h \rightarrow 0} \frac{x(t_{n+1}) - x(t_n)}{h} = x'(t_n) = f(t_n, x(t_n)). \quad (3.10)$$

et

$$\lim_{h \rightarrow 0} \phi(t_n, x(t_n), h) = \phi(t_n, x(t_n), 0). \quad (3.11)$$

Donc:  $\forall n = 0, \dots, N$ ,  $\phi(t_n, x(t_n), 0) = f(t_n, x(t_n))$ . Et donc:  $\phi(t, x, 0) = f(t, x)$ .

La réciproque est évidente. ■

**Théorème 3.3** Si la fonction  $\phi$  vérifie une condition de Lipschitz par rapport à la deuxième variable, c'est-à-dire:

$$\exists M > 0, \forall t \in [a, b], \forall x, y \in \mathbb{R} |\phi(t, x, h) - \phi(t, y, h)| \leq M|x - y| \quad (3.12)$$

pour  $h$  suffisamment petit,  $M$  ne dépendant pas de  $h$ , alors la méthode est stable.

**Preuve:** Soient  $\{y_n; 0 \leq n \leq N\}$  et  $\{z_n; 0 \leq n \leq N\}$  vérifiant:

$$y_{n+1} = y_n + h\phi(t_n, y_n, h)$$

et

$$z_{n+1} = z_n + h(\phi(t_n, z_n, h) + \epsilon_n).$$

On a alors:

$$\begin{aligned} |y_{n+1} - z_{n+1}| &\leq |y_n - z_n| + h|\phi(t_n, y_n, h) - \phi(t_n, z_n, h)| + h|\epsilon_n| \\ &\leq |y_n - z_n| + hM|y_n - z_n| + h|\epsilon_n| \\ &\leq (1 + hM)|y_n - z_n| + h|\epsilon_n| \end{aligned}$$

On peut alors déduire (dem. par récurrence) que:

$$|y_n - z_n| \leq (1 + hM)^n |y_0 - z_0| + \frac{(1 + hM)^n - 1}{M} \max_{k \leq n} |\epsilon_k|. \quad (3.13)$$

On utilise ensuite que  $1 + hM \leq e^{hM}$  pour  $hM \geq 0$ . On obtient alors:

$$|y_n - z_n| \leq e^{nhM} |y_0 - z_0| + \frac{e^{nhM} - 1}{M} \max_{k \leq n} |\epsilon_k|. \quad (3.14)$$

On remarque que  $nh \leq b - a$ . On arrive enfin à:

$$|y_n - z_n| \leq e^{(b-a)M} |y_0 - z_0| + \frac{e^{(b-a)M} - 1}{M} \max_{k \leq n} |\epsilon_k|. \quad (3.15)$$

On pose  $M_1 = e^{(b-a)M}$  et  $M_2 = \frac{e^{(b-a)M} - 1}{M}$ , et on obtient la définition de la stabilité pour la méthode. ■

### 3.1.3 Erreur de discrétisation

**Définition 3.6** On définit l'erreur de discrétisation par

$$e_n = x_n - x(t_n) \quad (3.16)$$

**Définition 3.7** Ordre d'une méthode

La méthode est d'ordre  $\geq p$  si pour toute solution  $x(t)$  de  $x'(t) = f(t, x)$

$$\max_n \left| \frac{x(t_{n+1}) - x(t_n)}{h} - \phi(t_n, x(t_n), h) \right| = h^{p-1} \epsilon(h) = O(h^p) \quad (3.17)$$

**Théorème 3.4** Si  $\phi$  vérifie une condition de Lipschitz par rapport à la deuxième variable et si la méthode est d'ordre  $\geq p$  alors:

$$\max_n |e_n| = \max_n |x_n - x(t_n)| = h^{p-1} \epsilon(h) \quad (3.18)$$

**Preuve:** On a :

$$x_{n+1} = x_n + h\phi(t_n, x_n, h), \quad (3.19)$$

et

$$x(t_{n+1}) = x(t_n) + h(\phi(t_n, x(t_n), h) + \epsilon_n) \quad (3.20)$$

avec  $\max_n |\epsilon_n| = O(h^p)$  (car la méthode est d'ordre  $\geq p$ ). Comme la méthode est stable (condition de Lipschitz sur la deuxième variable de  $\phi$ ), il existe deux constantes  $M_1$  et  $M_2$  telles que:

$$\max_n |x_n - x(t_n)| \leq M_1 |x(a) - x_0| + M_2 O(h^p). \quad (3.21)$$

D'où le résultat. ■

**Exemple:** Méthode d'Euler.

## 3.2 Méthodes numériques plus précises

Deux méthodes numériques très classiques utilisent une technique calquée sur la méthode d'Euler: ayant fixé un pas  $h$ , on construit une suite:

$$\begin{cases} t_{n+1} = t_n + h \\ x_{n+1} = x_n + h.m \end{cases}$$

Le nombre  $m$  peut être assimilé à une pente moyenne sur l'intervalle  $[t_n, t_{n+1}]$ . Dans la méthode d'Euler,  $m$  est simplement la pente  $f(t_n, x_n)$  au point  $(t_n, x_n)$ .

### 3.2.1 Méthode du point milieu

On prend pour  $m$  la valeur de  $f$  au milieu du segment  $[(t_n, x_n), (t_n + h, x_n + hf(t_n, x_n))]$ , soit  $m_M = f(t_n + h/2, x_n + (h/2)f(t_n, x_n))$ .

Dans le cas d'une équation  $x'(t) = g(t)$ , on retrouve la méthode du point milieu pour l'évaluation d'une intégrale.

La méthode du point milieu est une méthode d'ordre 2.

### 3.2.2 Méthode de Runge-Kutta

On prend pour  $m$  un nombre  $m_{RK}$ , moyenne pondérée des valeurs de  $f$  en quatre points. On pose:

- $m_1 = f(t_n, x_n)$  : pente au point de départ
- $m_2 = f(t_n + h/2, x_n + h/2.m_1)$ : pente au milieu d'un segment de pente  $m_1$
- $m_3 = f(t_n + h/2, x_n + h/2.m_2)$ : pente au milieu d'un segment de pente  $m_2$
- $m_4 = f(t_n + h, x_n + h.m_3)$ : pente au bout d'un segment de pente  $m_3$
- $m_{RK} = \frac{1}{6}(m_1 + 2m_2 + 2m_3 + m_4)$

Dans le cas particulier où l'équation s'écrit  $x'(t) = g(t)$ , on a  $m_2 = m_3$  et on retrouve exactement la méthode de Simpson pour le calcul de l'intégrale.

Cette méthode RK est une méthode d'ordre 4.

### 3.2.3 Méthodes implicites

La valeur  $x_{n+1}$  est donnée implicitement à partir de  $t_n, x_n$  et  $h$ . La méthode d'Euler implicite est définie par:

$$\begin{cases} t_{n+1} = t_n + h \\ x_{n+1} = x_n + hf(t_{n+1}, x_{n+1}) \end{cases}$$

Il faut alors résoudre une équation pour calculer  $x_{n+1}$ .

### 3.2.4 Méthodes à pas variable

Quand les solutions cherchées ont des comportements très différents selon les régions du plan, on a intérêt à adapter le pas  $h$  en fonction de la région. Des tests sont effectués périodiquement, et lorsque la valeur obtenue après un pas  $h$  est trop différente de celle obtenue après deux pas de  $h/2$ , c'est qu'il est préférable de diminuer le pas.

## 3.3 Méthodes à pas multiples

Nous avons étudié deux types de méthodes: les unes basées sur la formule de Taylor et les autres du type Runge-Kutta. Les premières sont numériquement mauvaises (et nécessitent le calcul et le stockage des dérivées successives de  $f$  si l'on souhaite augmenter la précision). Les secondes sont bonnes, mais font intervenir le calcul de  $f(x, y)$  en des points ne servant pas directement au calcul. Dans les méthodes à pas multiples, on calcule  $x_{n+k}$  en fonction de  $x_{n+k-1}, x_{n+k-2}, \dots, x_n$ , sans faire intervenir de valeurs intermédiaires. Soit l'équation différentielle:

$$\begin{cases} x'(t) = f(t, x), & \forall t \in [a, b] \\ x(a) = \alpha \end{cases}$$

où  $f$  est une fonction de classe  $C^1$  et donc l'unique solution est notée  $x(t)$ .

Pour  $h = \frac{b-a}{N}$ , posons  $t_n = a + nh$ ,  $x_n$  une approximation de  $x(t_n)$  et  $f_n = f(t_n, x_n)$  pour  $n = 0, \dots, N$ .

**Définition 3.8** On appelle méthode à  $k$  pas, une méthode qui a pour forme générale:

$$\alpha_k x_{n+k} + \alpha_{k-1} x_{n+k-1} + \dots + \alpha_0 x_n = h(\beta_k f_{n+k} + \dots + \beta_0 f_n) \quad (3.22)$$

où les  $(\alpha_i)$  et les  $\beta_i$  sont indépendants de  $n$  et  $\alpha_k \neq 0$ .

Lorsque l'on connaît les quantités  $x_n, x_{n+1}, \dots, x_{n+k-1}$  donc  $f_n, f_{n+1}, \dots, f_{n+k-1}$ , on peut alors calculer  $x_{n+k}$ :

- directement si  $\beta_k = 0$ , c'est une méthode à  $k$  pas explicite
- en résolvant l'équation  $\alpha_k x_{n+k} - h\beta_k f(t_{n+k}, x_{n+k}) = d$ ,  $d$  étant connu. C'est une méthode implicite, lorsque  $\beta_k \neq 0$ .

On distingue deux grandes catégories de méthodes à pas multiples:

- On remplace  $f(t, x(t))$  par un polynôme d'interpolation et on intègre. Par exemple:

$$\int_{t_n}^{t_{n+2}} x'(t) dt = \int_{t_n}^{t_{n+2}} f(t, x(t)) dt. \quad (3.23)$$

Si on utilise la méthode de Simpson pour approcher la valeur de l'intégrale, on obtient:

$$x_{n+2} - x_n = \frac{h}{3}(f_{n+2} + 4f_{n+1} + f_n). \quad (3.24)$$

- On approche la dérivée en un point. Par exemple, à partir de  $x'(t_{n+2}) = f(t_{n+2}, x(t_{n+2}))$ , on approche  $x'(t_{n+2})$  par  $\frac{3}{2h} \left( x_{n+2} - \frac{4}{3}x_{n+1} + \frac{1}{3}x_n \right)$ . On obtient alors la formule:

$$x_{n+2} - \frac{4}{3}x_{n+1} + \frac{1}{3}x_n = \frac{2}{3}hf(t_{n+2}, x_{n+2}). \quad (3.25)$$

# Chapter 4

## Traitement du signal: Analyse de Fourier et ondelettes

### 4.1 Généralités

La théorie du signal a pour objet l'étude des signaux et des systèmes qui les transmettent. Mais qu'est-ce qu'un signal ? La notion de signal est très variée. Cela peut être l'évolution d'une quantité qui dépend du temps, de l'espace, d'une fréquence ou de toute autre variable. Ces quantités mesurables seront appelées signaux; elles correspondent en mathématiques à la notion de fonction, ou plus exactement à celle de distribution. Il y a différentes façons de considérer un signal:

- déterministe ou aléatoire
- la variable peut être continue (signal analogique)  $x = x(t)$  par exemple, ou discrète (signal discret)  $x = (x_n)_{n \in \mathbb{Z}}$ . Un signal discret sera souvent obtenu par échantillonnage d'un signal analogique.

Nous n'étudierons ici que les signaux déterministes.

On appelle système toute entité (appareil physique, opérateur mathématique, etc) qui reçoit un signal d'entrée, et fournit un signal de sortie. S'il y a plusieurs signaux d'entrée ou de sortie, on dira alors que le signal est vectoriel. En théorie du signal, on ne s'intéresse pas nécessairement aux composantes du système, mais surtout à la façon dont il transforme un signal d'entrée en un signal de sortie. C'est une *boite noire* modélisée par un opérateur agissant sur des fonctions:

$$y = Ax \tag{4.1}$$

avec  $x \in X$  ensemble des signaux d'entrée, et  $y \in Y$  ensemble des signaux de sortie.

#### Exemples de signaux:

- fonction de Heaviside:  $s(t) = 0$  si  $t < 0$  et  $s(t) = 1$  si  $t \geq 0$ .
- signal sinusoïdal (ou monochromatique)  $s(t) = \alpha \cos(\omega t + \varphi)$  où  $|\alpha| = \max |s(t)|$  est l'*amplitude* du signal  
 $\omega$  est la *pulsation*

$a = \frac{2\pi}{\omega}$  est la plus petite période  
 $\lambda = \frac{1}{a}$  est la fréquence  
 $\varphi$  est la phase initiale.

Les valeurs d'un signal sont, en principe, des nombres réels, et la fréquence un nombre positif. Cependant, on utilise souvent une fonction à valeurs complexes pour représenter des signaux  $z(t) = \alpha e^{i(\omega t + \varphi)}$ . Cette écriture conduit à l'emploi de fréquences négatives, ce qui physiquement n'a aucun sens. Le signal  $z(t)$  peut aussi s'écrire  $z(t) = \beta e^{2i\pi\lambda t}$  où  $\beta$  est complexe et inclut la phase initiale  $\varphi$ .

## 4.2 Analyse de Fourier

**Définition 4.1** On appelle polynôme trigonométrique de degré inférieur ou égal à  $N$  toute fonction s'écrivant:

$$p(t) = \sum_{k=-N}^N c_k e_k(t) \quad (4.2)$$

où  $e_k(t) = e^{\frac{2i\pi kt}{a}}$  (de période  $a$ ).

Le polynôme  $p(t)$  peut se réécrire en fonction des fonctions sinus et cosinus.

**Proposition 4.1** On note  $\mathcal{T}_N$  l'ensemble des polynômes trigonométriques  $p$  de degré inférieur ou égal à  $N$ , muni du produit scalaire

$$(p, q) = \int_0^a p(t)q^*(t) dt. \quad (4.3)$$

Les fonctions  $e_n(t)_{n \in \mathbb{Z}}$  forment une famille orthogonale

$$(e_n, e_m) = \begin{cases} 0 & \text{si } n \neq m \\ a & \text{si } n = m. \end{cases} \quad (4.4)$$

On note  $L_p^2(0, a)$  l'espace des fonctions d'énergie finie périodiques de période  $a$  sur l'intervalle  $[0, a]$ :

$$L_p^2(0, a) = \left\{ f : \mathbb{R} \longrightarrow \mathbb{C}, \text{ de période } a \text{ t.q. } \int_0^a |f(t)|^2 dt < +\infty \right\} \quad (4.5)$$

**Théorème 4.1** Si  $f \in L_p^2(0, a)$  alors il existe un unique polynôme trigonométrique  $f_N$  dans  $\mathcal{T}_N$  tel que

$$\|f - f_N\|_2 = \min_{p \in \mathcal{T}_N} \{\|f - p\|_2\}. \quad (4.6)$$

Ce polynôme est donné par

$$f_N(t) = \sum_{n=-N}^N c_n e_n(t) \quad (4.7)$$

où les coefficients  $c_n$  sont définis par

$$c_n = \frac{1}{a} \int_0^a f(t) e_n^*(t) dt. \quad (4.8)$$

De plus, on a:

$$\lim_{N \rightarrow +\infty} \int_0^a |f(t) - f_N(t)|^2 dt = 0 \quad (4.9)$$

La famille de fonctions  $(e_n)_{n \in \mathbb{Z}}$  forme donc une base de l'espace  $L_p^2(0, a)$ . Les coefficients  $c_n$  sont appelés *coefficients de Fourier* de la fonction périodique  $f$ .

**Théorème 4.2** *Egalité de Parseval:*

$$\frac{1}{a} \int_0^a |f(t)|^2 dt = \sum_{n=-\infty}^{+\infty} |c_n|^2 \quad (4.10)$$

Les résultats précédents ne concernent que la convergence en moyenne quadratique de la série de Fourier de  $f$  vers  $f$ , mais ne donnent aucune information sur sa convergence ponctuelle. Pour cela, nous avons besoin d'hypothèses plus fines sur la fonction  $f$ .

**Théorème 4.3** *Convergence locale*

Soit  $f \in L_p^1(0, a)$ . Si en un point  $t_0$  les limites  $f(t_0^+)$  et  $f(t_0^-)$  existent, de même que les dérivées à droite et à gauche en  $t_0$ , alors

$$f_N(t_0) \longrightarrow \frac{1}{2} (f(t_0^+) + f(t_0^-)) \quad (4.11)$$

**Théorème 4.4** *Convergence uniforme*

Si la fonction  $f$  de période  $a$  est continue sur  $\mathbb{R}$  et si elle admet sur  $[0, a]$  une dérivée; si de plus  $f'$  est continue par morceaux sur  $[0, a]$  alors:

- La série de Fourier de  $f'$  s'obtient en dérivant terme à terme celle de  $f$ .
- Les coefficients de Fourier de  $f$  vérifient:

$$\sum_{n=-\infty}^{+\infty} |c_n| < +\infty \quad (4.12)$$

- La série de Fourier de  $f$  converge normalement sur  $\mathbb{R}$  vers  $f$ .

**Proposition 4.2** Soit  $f \in L_p^2(0, a)$ . Alors les deux propriétés suivantes sont équivalentes:

1. Les coefficients de Fourier de  $f$  sont à décroissance rapide.
2. La fonction  $f$  est indéfiniment dérivable.

**Définition 4.2** *Spectre d'un signal périodique*

Soit un signal  $f$ , de période  $a$ , et développé en série de Fourier sous la forme:

$$f(t) = \sum_{n=-\infty}^{+\infty} c_n e^{2i\pi n t/a}. \quad (4.13)$$

On appelle *spectre* de  $f$  l'ensemble des couples  $(n/a, c_n)_{n \in \mathbb{Z}}$ .



### 4.2.1 Transformée de Fourier discrète

Nous avons vu dans la section précédente la définition théorique des séries de Fourier pour les fonctions périodiques. Nous allons maintenant proposer une méthode pour calculer une approximation des coefficients  $c_n$ . Nous supposons que nous connaissons la période  $a$  de la fonction  $f$ , ainsi qu'un nombre fini  $N$  de ses valeurs, régulièrement espacées sur une période,  $f(k\frac{a}{N}) = y_k$  pour  $k = 0, \dots, N-1$ . Nous voulons déduire de ces informations une approximation des coefficients de Fourier  $c_n$  de  $f$ . Nous supposons de plus qu'il y a convergence ponctuelle de la série de Fourier de  $f$  vers  $f$ . Ayant  $N$  données, nous allons chercher à approcher les  $N$  principaux coefficients de Fourier  $c_n$  avec  $n = -N/2, \dots, N/2-1$ . Nous devons donc effectuer un calcul approché de l'intégrale:

$$c_n = \frac{1}{a} \int_0^a f(t) e^{-2i\pi n \frac{t}{a}} dt. \quad (4.14)$$

Pour cela, nous allons approcher cette intégrale par la formule des trapèzes:

$$c'_n = \frac{1}{N} \sum_{k=0}^{N-1} y_k e^{-2i\pi n \frac{k}{N}} \quad (4.15)$$

qui peut se réécrire:

$$c'_n = \frac{1}{N} \sum_{k=0}^{N-1} y_k \omega_N^{-nk} \quad (4.16)$$

où  $\omega_N = e^{2i\frac{\pi}{N}}$ . En tenant compte de la périodicité, on calcule dans la pratique:

$$Y_n = \frac{1}{N} \sum_{k=0}^{N-1} y_k \omega_N^{-nk}, \quad n = 0, \dots, N-1. \quad (4.17)$$

On peut alors reconstruire le signal initial par la formule:

$$y_k = \sum_{n=0}^{N-1} Y_n \omega_N^{nk}, \quad k = 0, \dots, N-1. \quad (4.18)$$

### 4.2.2 Fast Fourier Transform

Le calcul des valeurs  $\{Y_n\}_{n=0, \dots, N-1}$  par la formule des trapèzes nécessite  $N^2$  opérations. Un algorithme développé par Cooley et Tuckey exploite la forme particulière de la matrice de transformation construite à partir des racines de l'unité.

Supposons que  $N$  soit pair,  $N = 2m$ , et regroupons dans (4.17) les termes d'indices pairs et les termes d'indices impairs:

$$Y_n = \frac{1}{2} (P_n + \omega_N^{-n} I_n) \quad (4.19)$$

avec

$$P_n = \frac{1}{m} \left( y_0 + y_2 \omega_N^{-2n} + \dots + y_{N-2} \omega_N^{-(N-2)n} \right)$$

$$I_n = \frac{1}{m} \left( y_1 + y_3 \omega_N^{-2n} + \dots + y_{N-1} \omega_N^{-(N-2)n} \right)$$

Nous pouvons remarquer que:

$$\begin{aligned} P_{n+m} &= P_n \\ I_{n+m} &= I_n \\ \omega_N^{-(n+m)} &= -\omega_N^{-n} \end{aligned}$$

On peut ainsi exploiter ces propriétés pour diviser par deux le nombre d'opérations. On commence par calculer  $P_n$  et  $\omega_N^{-n} I_n$ . On en déduit ensuite  $Y_n = \frac{1}{2} (P_n + \omega_N^{-n} I_n)$  et  $Y_{n+m} = \frac{1}{2} (P_n - \omega_N^{-n} I_n)$  pour  $n = 0, 1, \dots, m-1$ .

Cette première astuce permet de diviser par deux le nombre d'opérations, mais on peut aussi remarquer que  $P_n$  et  $I_n$  sont aussi des transformées de Fourier discrètes d'ordre  $m = N/2$  indépendantes l'une de l'autre. Il suffit alors de répéter la décomposition précédente à condition que  $m$  soit encore pair. En choisissant donc  $N$  comme une puissance de 2, on peut poursuivre ces décompositions jusqu'à obtenir des transformées de Fourier discrètes d'ordre 2 qui se résument à des calculs de la forme:

$$\begin{aligned} Y &= \frac{y+z}{2} \\ Z &= \frac{y-z}{2}. \end{aligned}$$

On obtient à l'issue de ces calculs les coefficients de Fourier dans le désordre. Il suffit alors de les réordonner, ou de changer l'ordre des valeurs dès le début pour obtenir le bon résultat à la fin. Il existe aussi des astuces pour permuter les coefficients de façon efficace. Le coût de l'algorithme est alors en  $N \log_2 N$ .

### 4.2.3 Phénomène de Gibbs

Nous étudierons ce problème sous la forme d'un exercice. On utilisera la formule suivante:

$$\frac{1}{2} + \sum_{k=1}^n \cos(kx) - \frac{\cos(nx)}{2} = \frac{\sin(nx)}{2 \tan(x/2)}$$

1. On considère la fonction "en dents de scie"  $s(x)$ ,  $2\pi$ -périodique et définie par  $s(x) = \frac{\pi-x}{2}$  sur  $[0, 2\pi[$ .

- (a) Représenter  $s(x)$ .
- (b) Montrer que la formule d'inversion appliquée à  $s$  donne formellement:

$$s(x) = \sum_{k=1}^{\infty} \frac{\sin(kx)}{k}$$

Dans la suite, on considère les sommes partielles de cette série:

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

2. Montrer la relation:

$$s_n(x) - \int_0^x \frac{\sin(nt)}{t} dt = \int_0^x \sin(nt) \left( \frac{1}{2 \tan(t/2)} - \frac{1}{t} \right) + \frac{\cos(nt)}{2} - \frac{1}{2} dt$$

En déduire que pour  $|x| \leq 1$  on a:

$$s_n(x) = \int_0^x \frac{\sin(nt)}{t} dt - \frac{x}{2} + O(1/n)$$

3. Etudier les variations de la fonction

$$G(a) = \int_0^a \frac{\sin(t)}{t} dt$$

et en déduire que  $G(\pi) > G(+\infty)$  (on commencera par montrer que l'intégrale est convergente en  $+\infty$ ).

4. Comment se comporte la suite  $s_n(\pi/n)$  quand  $n$  tend vers  $+\infty$ ? (On admettra que  $G(+\infty) = \frac{\pi}{2}$ ).

En déduire que  $\limsup s_n(x) > s(0^+)$  et  $\liminf s_n(x) < s(0^+)$  quand  $n \rightarrow +\infty$  et  $x \rightarrow 0^+$ .

5. Tracer  $s_n(x)$  pour  $n$  grand sur le même graphe que  $s(x)$ .

## 4.3 Filtre et échantillonnage

### 4.3.1 Notion de filtre

Le terme de *filtre* désigne à la fois un système physique ayant des propriétés particulières, et sa modélisation mathématique définie comme étant un système linéaire, continu et invariant.

**Proposition 4.3** *Tout signal monochromatique  $e_\lambda(t) = e^{2i\pi\lambda t}$  est une fonction propre de tout filtre  $A$ . Il existe donc une fonction scalaire  $H(\lambda)$  telle que*

$$\forall \lambda \in \mathbb{R}, \quad A(e_\lambda) = H(\lambda)e_\lambda. \quad (4.20)$$

La fonction  $H : \mathbb{R} \rightarrow \mathbb{C}$  est appelée *fonction de transfert du filtre  $A$* .

**Exemple (exercice):**

On considère un circuit électrique de type  $RC$ , composé d'une alimentation, d'une résistance et d'un condensateur.

On note  $x(t)$  la tension en entrée et  $v(t)$  la tension aux bornes du condensateur. La différence de potentiel aux bornes d'un condensateur de charge  $Q$  étant  $v = Q/C$ , la loi d'Ohm fournit l'équation  $Ri(t) + v(t) = x(t)$ . Sachant que  $i(t) = Q'(t)$ , on obtient donc une équation différentielle linéaire du premier ordre:  $RCv'(t) + v(t) = x(t)$ .

1. Déterminer la fonction de transfert de ce filtre. Quelle est la nature qualitative de ce filtre ?

2. En effectuant le changement de variable  $v(t) = w(t) \exp(-\frac{t}{RC})$ , et en supposant que la réponse à l'entrée nulle est nulle, montrer que ce circuit est linéaire par rapport à la tension d'entrée  $x(t)$ .

**Exemple:** filtres de butterworth

### 4.3.2 Transformée de Fourier, théorème de Shannon

Dans la formule de reconstruction (4.13) valable pour les signaux périodiques, le coefficient  $\frac{n}{a}$  est homogène à une fréquence, et lorsque  $n$  décrit  $\mathbb{Z}$ ,  $\frac{n}{a}$  décrit un ensemble de fréquences discrètes dépendantes de  $a$ .

Si le signal  $f$  n'est pas périodique, cette relation n'est plus valable. Dans ce cas, on peut considérer un signal non périodique comme étant périodique mais de période  $a$  infinie. Ainsi, lorsque  $a$  est très "grand", alors  $\frac{n}{a}$  décrit presque toutes les fréquences possibles. En passant à la limite  $a \rightarrow +\infty$ , on passe alors d'un ensemble de fréquences discrètes à un ensemble de fréquences continues. Il en découle alors, que dans la formule 4.13, on passe alors d'une somme discrète à une somme continue, c'est-à-dire une intégrale. On parle alors de transformée de Fourier et non plus de série de Fourier.

**Définition 4.3** Si  $f$  est une fonction de  $L^1(\mathbb{R})$ , sa transformée de Fourier est définie pour tout  $\omega \in \mathbb{R}$  par :

$$\hat{f}(\omega) = \int_{t \in \mathbb{R}} f(t) e^{-2i\pi\omega t} dt \quad (4.21)$$

La transformée de Fourier d'une fonction de  $L^1(\mathbb{R})$  est continue et tend vers 0 à l'infini. Nous pouvons alors reconstruire  $f$  à partir de sa transformée de Fourier à l'aide d'une intégrale. Si  $\hat{f}$  appartient à  $L^1(\mathbb{R})$  on a la formule d'inversion suivante :

$$f(t) = \int_{\omega \in \mathbb{R}} \hat{f}(\omega) e^{2i\pi\omega t} d\omega \quad (4.22)$$

**Définition 4.4** Si  $f \in L^2(\mathbb{R})$ , on définit la transformée de Fourier par

$$\hat{f}(\omega) = \lim_{N \rightarrow \infty} \int_{t=-N}^N f(t) e^{-2i\pi\omega t} dt \quad (4.23)$$

Cette limite existe pour presque tout  $\omega \in \mathbb{R}$  et la fonction  $\hat{f}$  ainsi définie appartient à  $L^2(\mathbb{R})$ . On a par ailleurs  $\|f\|_2 = \|\hat{f}\|_2$ .

La formule de Shannon est une formule d'interpolation exprimant la valeur  $f(t)$  d'un signal à un instant  $t$  quelconque, en fonction de valeurs discrètes équiréparties sur  $\mathbb{R}$ ,  $(f(na))_{n \in \mathbb{Z}}$ . Le signal sera donc totalement déterminé par ses valeurs discrétisées. Cette propriété est évidemment fautive pour un signal quelconque, et on devra imposer des hypothèses restrictives sur le signal  $f$ .

**Définition 4.5** On appelle sinus cardinal la fonction

$$s(t) = \frac{\sin(\frac{\pi}{a}t)}{\frac{\pi}{a}t}. \quad (4.24)$$

**Théorème 4.5 Formule de Shannon. Cas général.** Soit  $f$  un signal de  $E_{\lambda_c}$ :

$$\hat{f} \in L^2(\mathbb{R}), \quad \text{et} \quad \text{Supp}(\hat{f}) \subset [-\lambda_c, \lambda_c]. \quad (4.25)$$

Alors on a:

$$\forall a > 0, \quad \sum_{n=-\infty}^{+\infty} |f(na)|^2 < +\infty \quad (4.26)$$

et

$$\forall a \leq \frac{1}{2\lambda_c}, \quad f(t) = \sum_{n=-\infty}^{+\infty} f(na) \frac{\sin\left(\frac{\pi}{a}(t - na)\right)}{\frac{\pi}{a}(t - na)}. \quad (4.27)$$

Cette égalité a lieu au sens de la norme  $L^2(\mathbb{R})$ . Elle a lieu au sens de la convergence uniforme si en outre:

$$\sum_{n=-\infty}^{+\infty} |f(na)| < +\infty. \quad (4.28)$$

**Théorème 4.6 Formule de Shannon. Cas particulier pour un signal trigonométrique.** Pour un signal trigonométrique,

$$f(t) = \sum_{n=-N}^N c_n e^{2i\pi\lambda_n t}, \quad \lambda_n \in \mathbb{R} \quad (4.29)$$

on a la formule de Shannon suivante:

$$\forall a \in ]0, \frac{1}{2\lambda_c}], \quad \forall t \in \mathbb{R}, \quad f(t) = \sum_{n=-\infty}^{+\infty} f(na) \frac{\sin\left(\frac{\pi}{a}(t - na)\right)}{\frac{\pi}{a}(t - na)}, \quad (4.30)$$

avec  $\lambda_c = \max\{|\lambda_n|, |n| \leq N\}$ , et la convergence est ponctuelle sur  $\mathbb{R}$ .

Le théorème de Shannon apporte une réponse au problème de l'échantillonnage lorsque l'on cherche à calculer le spectre d'une fonction. Il donne une borne pour la fréquence d'échantillonnage en fonction de la fréquence maximale présente dans le signal.

## 4.4 Analyse temps-fréquence

L'analyse de Fourier est bien adaptée à l'analyse de signaux périodiques et réguliers. Dans ce cas, les coefficients de Fourier  $c_n$  deviennent vite négligeables quand  $|n|$  augmente, et on peut alors caractériser le signal correctement dans le domaine des fréquences à l'aide de quelques coefficients seulement. Mais dès que le signal devient irrégulier, ou lorsqu'il comporte des composantes transitoires, l'analyse de Fourier n'est plus adaptée à ce type de signal. En effet, dans le spectre  $\hat{f}(\lambda)$  tous les aspects temporels du signal disparaissent, comme par exemple le début et la fin d'une composante transitoire ou l'apparition d'une singularité.

L'intérêt de l'analyse de Fourier, largement utilisée pour étudier les composantes spectrales d'un phénomène, est de pouvoir décrire un système par ses caractéristiques fréquentielles plutôt que temporelles, venant ainsi compléter la représentation classique. Cependant, une mauvaise description dans une des deux représentations peut entraîner une

mauvaise compréhension générale d'un phénomène. La représentation idéale est celle qui donne la meilleure description possible pour un système donné. Pour cela, l'utilisation d'une méthode qui puisse réaliser simultanément une analyse avec des paramètres de temps et de fréquence, comme une partition de musique où sont indiquées la fréquence et la durée des notes, est donc nécessaire. Nous présentons dans cette partie quelques exemples d'analyses *temps-fréquence* qui ont toutes en commun le principe de projection sur une famille de fonctions dépendant de deux variables respectivement reliées à la fréquence et au temps:

$$C_f(a, b) = \int_{\mathbb{R}} f(t) \psi_{a,b}^*(t) dt, \quad (4.31)$$

où  $\psi_{a,b}(t)$  joue le même rôle que l'exponentielle complexe dans la transformée de Fourier.

## 4.5 La transformée de Fourier à fenêtre

Une possibilité pour le choix de  $\psi_{a,b}(t)$  est d'utiliser une fonction  $\omega(t)$  qui soit translatée et modulée:

$$\psi_{a,b}(t) = \omega(t - b) e^{2i\pi at}, \quad a, b \in \mathbb{R} \quad (4.32)$$

où  $\omega(t)$  est une fonction fenêtre, gaussienne en général. Avec cette construction, on dispose de fonctions analysantes bien localisées en temps et en fréquence.

**Théorème 4.7** Soit  $\omega \in L^1 \cap L^2$  une fenêtre telle que  $|\hat{\omega}|$  soit une fonction paire et  $\|\omega\|_2 = 1$ . On pose:

$$\psi_{a,b}(t) = \omega(t - b) e^{2i\pi at}, \quad a, b \in \mathbb{R} \quad (4.33)$$

Pour tout signal  $f \in L^2(\mathbb{R})$ , on considère les coefficients

$$C_f(a, b) = \int_{\mathbb{R}} f(t) \psi_{a,b}^*(t) dt. \quad (4.34)$$

Alors on a:

- Conservation de l'énergie.

$$\int_{\mathbb{R}} \int_{\mathbb{R}} |C_f(a, b)|^2 da db = \int_{\mathbb{R}} |f(t)|^2 dt. \quad (4.35)$$

- Formule de reconstruction

$$f(t) = \int_{\mathbb{R}} \int_{\mathbb{R}} C_f(a, b) \psi_{a,b}(t) da db \quad (4.36)$$

au sens suivant: si

$$g_A(t) = \int_{b \in \mathbb{R}} \int_{|a| \leq A} C_f(a, b) \psi_{a,b}(t) da db \quad (4.37)$$

alors  $g_A \rightarrow f$  dans  $L^2$  quand  $A \rightarrow +\infty$ .

Pour la transformée de Fourier à fenêtre dans  $L^2$ , nous avons donc des formules analogues à celles de la transformée de Fourier classique. Dans la pratique, on choisit une fonction  $\omega$  localisée autour de l'origine comme une fonction gaussienne par exemple. La fonction  $\psi_{a,b}$  est alors localisée autour du point  $t = b$  tandis que  $\hat{\psi}_{a,b}$  est localisée autour du point  $\xi = a$ . Les coefficients  $C_f(a, b)$  contiennent alors une information à la fois en temps et en fréquence autour du point  $(a, b)$ . Dans la pratique, les calculs se font numériquement à partir d'une discrétisation du plan temps-fréquence.

Malgré les avantages apportés par cette représentation, elle n'est cependant pas parfaite et en particulier n'est pas adaptée pour décrire précisément des fonctions très irrégulières. En effet, la description de hautes fréquences très localisées nécessite une fenêtre avec un petit support alors que la description de basses fréquences diffuses est meilleure avec une fenêtre plus large. Pour pallier à ce problème, des fonctions analysantes dont la taille du support est reliée à leur fréquence peuvent être définies. Cette constatation conduit au développement de la transformée en ondelettes.

## 4.6 La transformée en ondelettes

L'idée de base dans la théorie des ondelettes est d'appliquer des dilatations en plus des translations précédemment introduites. Partant d'une fonction bien localisée en temps et en fréquence, une famille de fonctions analysantes se construit par la formule suivante,

$$\psi_{a,b}(t) = |a|^{-1/2} \psi\left(\frac{t-b}{a}\right), \quad a \in \mathbb{R}^*, b \in \mathbb{R}, \quad (4.38)$$

à partir d'une fonction initiale  $\psi(t)$  appelée l'ondelette *mère*. Dans cette écriture  $t$  est le paramètre de temps et  $1/a$  est homogène à une fréquence. On peut alors définir la transformée en ondelettes correspondante à l'aide de la relation (4.31). Selon le type d'application considérée, différentes familles d'ondelettes peuvent être choisies: on peut choisir de laisser les paramètres  $a$  et  $b$  varier de façon continue dans leur domaine  $\mathbb{R}^* \times \mathbb{R}$ , ou de restreindre leurs valeurs à une grille discrète. Ce choix génère deux sortes d'ondelettes, respectivement appelées ondelettes continues et ondelettes discrètes.

### 4.6.1 La transformée en ondelettes continues

Les caractéristiques de  $\psi$  appelée aussi, *ondelette mère* sont différentes de celles d'une fenêtre comme nous l'avons vu dans la section précédente. Ici, la fonction  $\psi$  sera d'intégrale nulle et donc oscillante. On imposera également que  $\psi$  et  $\hat{\psi}$  soient bien localisées entraînant donc une décroissance rapide vers 0 à l'infini. Les fonctions de base ressemblent donc à des vagues amorties (d'où le nom de la théorie).

**Exemples:**

- Ondelette de Morlet:  $\psi(t) = e^{-t^2/2} \cos(5t)$
- DOGs (derivative of Gaussian). La dérivée seconde est aussi appelée le *chapeau Méricain*:  $\psi(t) = \frac{1}{\sqrt{2\pi}\sigma^3} \left(1 - \frac{t^2}{\sigma^2}\right) \exp\left(\frac{-t^2}{2\sigma^2}\right)$

**Théorème 4.8** Soit  $\psi \in L^1 \cap L^2$ , une ondelette mère vérifiant les conditions suivantes:

1.  $\int_{-\infty}^{+\infty} \frac{|\hat{\psi}(\lambda)|^2}{|\lambda|} d\lambda = K < +\infty$
2.  $\|\psi\|_2 = 1$

On construit alors les ondelettes de base:

$$\psi_{a,b}(t) = |a|^{-1/2} \psi\left(\frac{t-b}{a}\right), \quad a \in \mathbb{R}^*, b \in \mathbb{R}, \quad (4.39)$$

et pour tout signal  $f \in L^2$  on considère ses coefficients d'ondelettes:

$$C_f(a, b) = \int_{\mathbb{R}} f(t) \psi_{a,b}^*(t) dt. \quad (4.40)$$

Alors on a:

1. Conservation de l'énergie:

$$\frac{1}{K} \int \int_{\mathbb{R}^2} |C_f(a, b)|^2 \frac{da db}{a^2} = \|f\|_2^2. \quad (4.41)$$

2. Formule de reconstruction:

$$f(t) = \frac{1}{K} \int \int_{\mathbb{R}^2} C_f(a, b) \psi_{a,b}(t) \frac{da db}{a^2}, \quad (4.42)$$

au sens suivant: si,

$$f_\epsilon(t) = \frac{1}{K} \int_{b \in \mathbb{R}} \int_{|a| \geq \epsilon} C_f(a, b) \psi_{a,b}(t) \frac{da db}{a^2}, \quad (4.43)$$

alors  $f_\epsilon \rightarrow f$  dans  $L^2$  quand  $\epsilon \rightarrow 0^+$ .

**Remarques:**

- La première hypothèse sur  $\psi$  implique que  $\hat{\psi}(0) = \int_{\mathbb{R}} \psi(t) dt = 0$ . L'ondelette mère doit donc être une fonction oscillante.
- Si le signal  $f \in L^1 \cap L^2$ , alors la formule de reconstruction est plus simple car les intégrales existent au sens usuel en prenant  $\epsilon = 0$ .

Contrairement à l'analyse de Fourier standard, l'analyse en ondelettes n'est pas adaptée pour des calculs explicites, même dans le cas de signaux et d'ondelettes mères simples. Toutes les intégrales sont donc calculées à l'aide d'approximations numériques. On restreint alors les paramètres  $a$  et  $b$  à un sous-ensemble discret de  $\mathbb{R}$ .



## 4.6.2 La transformée en ondelettes discrètes

Les ondelettes discrètes, comme leur nom l'indique, correspondent à un choix particulier de discrétisation de  $a$  et  $b$ ,  $a = a_0^m$ ,  $b = nb_0a_0^m$ , indiquant que le paramètre de translation  $b$  dépend du coefficient de dilatation choisi. La famille d'ondelettes devient alors, pour  $n \in \mathbb{Z}$ ,

$$\psi_{m,n}(t) = a_0^{-m/2} \psi(a_0^{-m}t - nb_0). \quad (4.44)$$

Cet ensemble définit une base d'ondelettes discrète, a priori non orthogonale, donc redondante; ce qui entraîne que la reconstruction d'une fonction  $f(t)$  de façon numérique à partir de ses coefficients  $\langle f, \psi_{m,n} \rangle$  n'est pas immédiate. Un tel algorithme de reconstruction est possible pour les  $\psi_{m,n}$  s'il existe deux constantes  $A > 0$  et  $B < +\infty$  telles que,

$$A\|f\|^2 \leq \sum_{m,n} |\langle f, \psi_{m,n} \rangle|^2 \leq B\|f\|^2, \quad \forall f \in L^2(\mathbb{R}). \quad (4.45)$$

Si  $A$  et  $B$  sont égaux à 1 alors la famille  $\{\psi_{m,n}\}_{m,n \in \mathbb{Z}}$  forme une base orthonormale.

Le système vu jusqu'à présent n'est pas orthogonal. Dans la pratique, cela signifie que c'est une description redondante de l'information qui est donnée par les coefficients  $C_f(a, b)$ . On peut cependant obtenir une base orthogonale en utilisant une discrétisation particulière.

**Exemple:** Système de Haar.

## 4.6.3 La transformée en ondelettes orthogonales

La construction théorique de familles d'ondelettes orthogonales est liée à la notion d'analyse multirésolution introduite par Mallat en 1989.

**Définition 4.6** Une analyse multirésolution est une décomposition de l'espace  $L^2(\mathbb{R})$  en une suite de sous-espaces emboîtés,

$$\dots \subset V_2 \subset V_1 \subset V_0 \subset V_{-1} \subset V_{-2} \dots$$

telle que

- $\bigcap_{j \in \mathbb{Z}} V_j = \{0\}$  et  $\bigcup_{j \in \mathbb{Z}} V_j$  soit dense dans  $L^2(\mathbb{R})$
- $f(x) \in V_j \Leftrightarrow f(2x) \in V_{j-1}$
- $f(x) \in V_0 \Leftrightarrow f(x - k) \in V_0$
- Il existe une fonction  $\varphi \in V_0$ , appelée l'ondelette père, telle que  $\{\varphi(x - k)\}_{k \in \mathbb{Z}}$  forme une base orthonormale de  $V_0$

Nous allons maintenant construire une base orthogonale de  $L^2(\mathbb{R})$  à partir d'une analyse multirésolution.

**Théorème 4.9** Soit  $W_j$  le sous-espace orthogonal complémentaire de  $V_j$  dans  $V_{j-1}$ :

$$V_j \oplus W_j = V_{j-1} \quad (4.46)$$

Cet espace contient la différence d'information entre  $V_j$  et  $V_{j-1}$ , et permet la décomposition de  $L^2(\mathbb{R})$  en une somme directe:

$$L^2(\mathbb{R}) = \bigoplus_{j \in \mathbb{Z}} W_j \quad (4.47)$$

Il existe alors une fonction  $\psi \in W_0$ , l'ondelette mère, telle que  $\{\psi(x - k)\}_{k \in \mathbb{Z}}$  soit une base orthonormale de  $W_0$ .

Les bases d'ondelettes correspondantes sont alors définies par:

$$\varphi_{j,k}(x) = 2^{-j/2} \varphi(2^{-j}x - k), \quad k, j \in \mathbb{Z}, \quad (4.48)$$

$$\psi_{j,k}(x) = 2^{-j/2} \psi(2^{-j}x - k), \quad k, j \in \mathbb{Z}. \quad (4.49)$$

**Définition 4.7** On dira que l'ondelette mère  $\psi$  possède  $M$  moments nuls si elle vérifie:

$$\int_{\mathbb{R}} \psi(x) x^m dx = 0, \quad m = 0 \dots M - 1, \quad (4.50)$$

### Algorithme de transformée en ondelettes rapide

Puisque l'ondelette père, aussi appelée fonction d'échelle,  $\varphi(x)$ , et l'ondelette mère  $\psi(x)$  appartiennent à  $V_{-1}$ , nous pouvons écrire leurs développements respectifs sur la base de  $V_{-1}$ :

$$\varphi(x) = \sqrt{2} \sum_{k=0}^{L-1} h_k \varphi(2x - k), \quad h_k = \langle \varphi, \varphi_{-1,k} \rangle, \quad (4.51)$$

$$\psi(x) = \sqrt{2} \sum_{k=0}^{L-1} g_k \varphi(2x - k), \quad g_k = (-1)^k h_{L-k-1}, \quad (4.52)$$

où le nombre  $L$  de coefficients est relié au nombre  $M$  de moments nuls et à certaines propriétés (régularité, symétrie, décroissance à l'infini...) qui peuvent être imposées à  $\varphi(x)$ . Les familles  $\{h_k\}$  et  $\{g_k\}$  forment en fait une paire de filtres miroirs en quadrature  $H$  et  $G$ .

**Définition 4.8** Soit un couple de filtres discrets  $(H, G) = (h_k, g_k)$  dont leurs transformées de Fourier  $h(\omega)$  et  $g(\omega)$  sont des fonctions  $2\pi$ -périodiques. Si  $h(\omega)$  et  $g(\omega)$  vérifient:

$$\begin{aligned} |h(\omega)|^2 + |g(\omega)|^2 &= 1, \quad \forall \omega \\ h(\omega) \overline{g(\omega)} + h(\omega + \pi) \overline{g(\omega + \pi)} &= 1, \quad \forall \omega \end{aligned}$$

Alors les filtres  $H$  et  $G$  sont appelés filtres miroirs en quadrature.

Les fonctions vérifiant (4.51) ou (4.52) ont leur support inclus dans  $[0 \dots L - 1]$ . De plus, s'il existe une échelle la plus "grossière",  $j = n$ , et une échelle la plus "fine",  $j = 0$ , les fonctions de base peuvent se réécrire:

$$\varphi_{j,k}(x) = \sum_{l=0}^{L-1} h_l \varphi_{j-1,2k+l}(x), \quad j = 1 \dots n, \quad (4.53)$$

et

$$\psi_{j,k}(x) = \sum_{l=0}^{L-1} g_l \varphi_{j-1,2k+l}(x), \quad j = 1 \dots n. \quad (4.54)$$

Le calcul de la transformée en ondelettes d'une fonction  $f(x)$  s'effectue alors par la détermination de deux ensembles de coefficients  $d_k^j$  et  $r_k^j$  définis par

$$d_k^j = \int_{\mathbb{R}} dx f(x) \psi_{j,k}(x), \quad (4.55)$$

et

$$r_k^j = \int_{\mathbb{R}} dx f(x) \varphi_{j,k}(x). \quad (4.56)$$

## Calcul pratique

A partir d'un ensemble initial de coefficients  $r_k^0$ , et en utilisant (4.53) et (4.54), il est alors possible de déterminer les coefficients  $d_k^j$  et  $r_k^j$  par les formules de récurrence suivantes:

$$d_k^j = \sum_{l=0}^{L-1} g_l r_{2k+l}^{j-1}, \quad (4.57)$$

et

$$r_k^j = \sum_{l=0}^{L-1} h_l r_{2k+l}^{j-1}. \quad (4.58)$$

Les coefficients  $d_k^j$ , et  $r_k^j$  sont considérés dans (4.57) et (4.58) comme des suites périodiques de période  $2^{n-j}$ . L'ensemble  $d_k^j$  est composé par les coefficients correspondant à la décomposition de  $f(x)$  sur la base  $\psi_{j,k}$  et  $r_k^j$  peut être interprété comme l'ensemble des approximations entre les différentes échelles.

## 4.7 Paquets d'ondelettes

### 4.7.1 Cas 1D

**Définition 4.9** Soient  $H$  et  $G$  une paire de filtres miroirs en quadrature dont les coefficients sont respectivement notés  $h_j$  et  $g_j$ . On note par  $\psi_0$  et  $\psi_1$  l'ondelette père et l'ondelette mère correspondantes. En utilisant les filtres  $H$  et  $G$ , on peut alors définir la suite de fonctions suivante:

$$\begin{aligned} \psi_{2n}(x) &= \sqrt{2} \sum_{j \in \mathbb{Z}} h_j \psi_n(2x - j), \\ \psi_{2n+1}(x) &= \sqrt{2} \sum_{j \in \mathbb{Z}} g_j \psi_n(2x - j). \end{aligned} \quad (4.59)$$

L'ensemble des fonctions  $\{\psi_n\}_n$  définit alors les paquets d'ondelettes associés aux filtres  $H$  et  $G$ .

Une base de paquets d'ondelettes orthonormale de  $L^2(\mathbb{R})$  pourra alors être constituée en choisissant n'importe quel système libre et générateur parmi les fonctions  $2^{s/2}\psi_n(2^s x - j)$ . La procédure de sélection des fonctions, appelée aussi algorithme du choix de la meilleure base, sera décrite dans la suite. Chaque élément de la base est alors défini par trois paramètres: l'échelle  $s$ , le nombre d'onde (lié à la fréquence)  $n$  et le temps  $j$ . Les coefficients de paquets d'ondelettes peuvent être représentés de façon concise à l'aide de rectangles dyadiques. Par exemple, la décomposition en paquets d'ondelettes d'un signal composé de 8 points  $\{x_1, \dots, x_8\}$  pourra avoir la représentation en arbre donnée dans le Tableau 4.1.

$x_1$	$x_2$	$x_3$	$x_4$	$x_5$	$x_6$	$x_7$	$x_8$
$r_1$	$r_2$	$r_3$	$r_4$	$d_1$	$d_2$	$d_3$	$d_4$
$rr_1$	$rr_2$	$dr_1$	$dr_2$	$rd_1$	$rd_2$	$dd_1$	$dd_2$
$rrr_1$	$drr_1$	$rdr_1$	$ddr_1$	$rrd_1$	$drd_1$	$rdd_1$	$ddd_1$

Table 4.1: Représentation dyadique des coefficients de paquets d'ondelettes

Chaque ligne est obtenue à partir de la ligne précédente par application des filtres  $H$  ou  $G$ . L'application du filtre  $H$  est notée  $r$  comme "résumé" et l'application du filtre  $G$  est notée  $d$  comme "différence". Par exemple, l'ensemble  $\{rd_1 \quad rd_2\}$  est obtenu par application du filtre  $H$  à  $\{d_1 \quad d_2 \quad d_3 \quad d_4\}$ , et  $\{dd_1 \quad dd_2\}$  est obtenu par application du filtre  $G$ .

### 4.7.2 Cas 2D

Les paquets d'ondelettes bidimensionnelles peuvent être obtenus par produits tensoriels  $\psi_{snk}(x) \cdot \psi_{s'n'k'}(y)$  des paquets à une dimension. Le support de ces fonctions est alors exactement le produit cartésien des supports de  $\psi_{snk}(x)$  et  $\psi_{s'n'k'}(y)$ . La même échelle  $s = s'$  sera utilisée dans la suite pour les deux directions d'espace. Ces ensembles de fonctions peuvent être représentés par des carrés dyadiques correspondant à l'application de l'un des filtres suivants  $H \otimes H = H_x H_y$ ,  $H \otimes G = H_x G_y$ ,  $G \otimes H = G_x H_y$ , ou  $G \otimes G = G_x G_y$ . Une décomposition en paquets d'ondelettes bidimensionnelles peut être représentée graphiquement par le schéma de la Figure 4.1.

## 4.8 Le critère entropique et l'algorithme du choix de la meilleure base

Les paquets d'ondelettes constituent un large ensemble de fonctions à partir duquel une base peut être sélectionnée. Le principal critère de choix consiste à chercher une base sur laquelle les coefficients, rangés par ordre décroissant, décroissent le plus rapidement

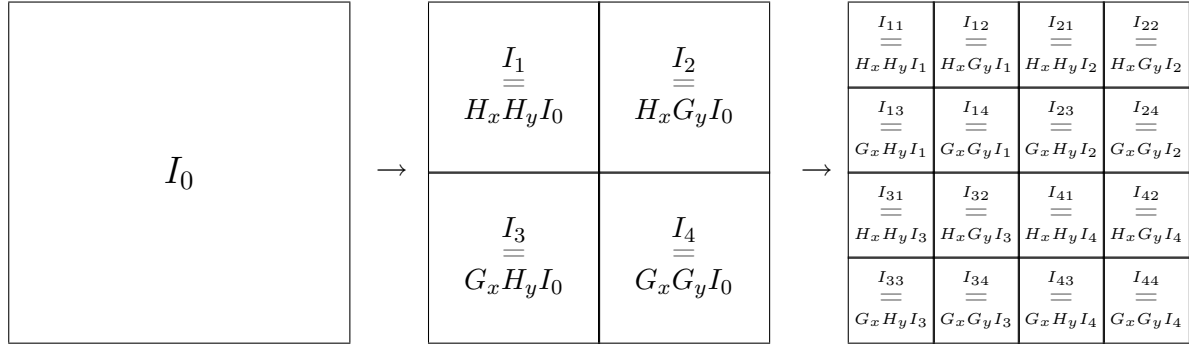


Figure 4.1: Deux niveaux d'une décomposition en paquets d'ondelettes bidimensionnels

possible.

Différents critères peuvent être définis à partir de fonctions *coûts* préalablement choisies.

### 4.8.1 L'entropie

La fonction coût la plus largement utilisée est l'*entropie*.

**Définition 4.10** Pour un vecteur  $u = \{u_k\}$  donné, l'entropie correspondante est définie par:

$$E(u) = \sum_k p(k) \log\left(\frac{1}{p(k)}\right), \quad (4.60)$$

où  $p(k) = \frac{|u_k|^2}{\|u\|_2^2}$  est l'énergie normalisée du  $k^{\text{ième}}$  élément du vecteur considéré. Si  $p(k) = 0$ , alors on pose  $p(k) \log\left(\frac{1}{p(k)}\right) = 0$ .

Tous les termes de la somme sont positifs. En fait, l'entropie mesure le logarithme du nombre de coefficients significatifs dans le signal original. Le vecteur  $p = \{p(k)\}_k$  peut être interprété comme étant une fonction de distribution de probabilité discrète car  $0 \leq p(k) \leq 1, \forall k$  et  $\sum_k p(k) = 1$ . On peut facilement montrer que si le nombre de valeurs  $p(k)$  non nulles est égal  $N$ , alors  $E(u) \leq \log N$ . Une telle fonction de distribution de probabilité est dite concentrée en au plus  $N$  valeurs. Une faible valeur pour  $E(u)$  indique une forte concentration de l'énergie de  $u$  sur peu de composantes  $u_k$ . D'autres choix sont possibles pour la fonction coût, dont par exemple:

- Concentration en norme  $l^p$ :

$$E(u) = \sum_k \frac{|u_k|^p}{\|u\|_p^p}. \quad (4.61)$$

- Nombre de coefficients au-dessus d'un seuil donné  $\epsilon$ :

$$E(u) = \#\{|u_k| \text{ t.q. } |u_k| > \epsilon\}. \quad (4.62)$$

### 4.8.2 Algorithme du choix de la meilleure base

Une fois la fonction coût choisie, il reste à définir un algorithme permettant d'extraire la meilleure base (au sens de la minimisation de la fonction coût) de l'ensemble des paquets d'ondelettes.

1. Calcul des fonctions coûts de chacun des éléments de l'arbre. Si l'on considère par exemple un vecteur à 8 composantes, et si l'on choisit le *nombre de coefficients au-dessus d'un seuil donné* pour la fonction coût, la première étape conduit à un tableau du type du Tableau 4.2.

8							
4				3			
2		1		1		2	
1	0	1	1	1	0	1	0

Table 4.2: Arbre d'entropie, première étape

2. Parcours de l'arbre en partant du bas, et comparaison de la somme de l'entropie obtenue par les enfants à celle obtenue par le parent. Par exemple, si l'on considère les deux premiers enfants de la dernière ligne, ils obtiennent respectivement 1 et 0. Alors que le parent obtient 2. Il est donc plus avantageux de représenter la même information à l'aide des deux enfants (qui conduisent à une entropie totale de  $1 + 0$ ) qu'avec le parent. Les enfants seront donc choisis à la place du parent.
3. Le parent se voit alors attribué l'entropie induite par les enfants pour continuer l'algorithme en itérant la même comparaison avec les parents de la ligne supérieure suivante.

4 (8)							
2 (4)				2 (3)			
1 (2)		1		1		1 (2)	
1	0	1	1	1	0	1	0

Table 4.3: Arbre d'entropie, deuxième étape

A la fin du processus, la base possédant l'entropie la plus faible aura été sélectionnée parmi toutes les bases possibles. Dans notre exemple, on s'aperçoit qu'un signal avec une entropie initiale de 8 peut être représenté dans une base avec une entropie de 4.

Les éléments qui ont été sélectionnés sont indiqués en gras dans le Tableau 4.3. Avec cet algorithme, la base est choisie automatiquement de façon à fournir la meilleure représentation du vecteur initial au sens de la minimisation de la fonction coût choisie.

**Remarques:**

Nous n'avons présenté ici qu'un bref aperçu de la théorie des ondelettes. Toutes les notions s'étendent au cas de plusieurs variables. Cette théorie relativement récente (début des années 80) connaît un grand développement tant d'un point de vue théorique qu'appliqué. Elle permet également l'analyse d'objets fractales, ces objets géométriques dont les structures demeurent inchangées lorsqu'ils sont observés à différentes échelles.

# Chapter 5

## Exercices

### 5.1

#### Exercice 1. Comparaison entre interpolation de Lagrange et interpolation composite linéaire

On considère la fonction  $f(x) = e^{2x}$ , avec  $x \in [0, 1]$ . Soit  $N$  un nombre entier, on pose  $H = \frac{1}{N}$ , et on note  $\Pi_1^H f$  le polynôme composite linéaire par morceaux qui interpole  $f$  aux noeuds  $x_i = iH$ , pour  $i = 1, \dots, N$ .

1. Calculer le nombre minimal  $N$  de sous-intervalles pour que l'erreur d'interpolation  $E_1^H(f) = \max_{x \in [0,1]} |f(x) - \Pi_1^H f(x)|$  soit inférieure à  $10^{-4}$ .
2. Soit  $\Pi_n f$  le polynôme d'interpolation de Lagrange de degré  $n$  qui interpole  $f$  aux noeuds  $x_i = iH$ , pour  $i = 1, \dots, n$ . Est-ce que l'erreur d'interpolation  $E_n(f) = \max_{x \in [0,1]} |f(x) - \Pi_n^H f(x)|$  tend vers zéro quand  $n \rightarrow +\infty$  ?
3. Est-ce que le nombre de noeuds nécessaires pour que  $E_n(f)$  soit inférieure à  $10^{-4}$  est du même ordre de grandeur que celui de la question 1) ?

#### Exercice 2. Formule de Newton

Soient  $(x_i, y_i)_{i=0, \dots, n}$  un ensemble de points donnés. On définit les différences finies divisées par les notations suivantes:

$$\begin{aligned} y[x_i] &= y_i \\ \partial_y[x_i, x_j] &= \frac{y[x_j] - y[x_i]}{x_j - x_i} \\ \partial_y^2[x_i, x_j, x_k] &= \frac{\partial_y[x_j, x_k] - \partial_y[x_i, x_j]}{x_k - x_i} \\ &\cdot \quad \cdot \\ &\cdot \quad \cdot \\ &\cdot \quad \cdot \\ \partial_y^n[x_{i_0}, x_{i_1}, \dots, x_{i_n}] &= \frac{\partial_y^{n-1}[x_{i_1}, \dots, x_{i_n}] - \partial_y^{n-1}[x_{i_0}, \dots, x_{i_{n-1}}]}{x_{i_n} - x_{i_0}} \end{aligned}$$



Montrer par récurrence que le polynôme d'interpolation de degré  $n$  qui passe par les  $n + 1$  points  $(x_0, y_0), \dots, (x_n, y_n)$ , avec les  $x_i$  distincts, est unique et donné par:

$$p(x) = y[x_0] + (x-x_0)\partial_y[x_0, x_1] + (x-x_0)(x-x_1)\partial_y^2[x_0, x_1, x_2] + \dots + (x-x_0)\dots(x-x_{n-1})\partial_y^n[x_0, x_1, \dots, x_n]$$

Quel est l'avantage de cette formule par rapport à la définition habituelle du polynôme d'interpolation de Lagrange?

### Exercice 3. Convergence uniforme

Soit  $n \in \mathbb{N}$ . Soit  $(P_k)_{k \in \mathbb{N}}$  une suite de polynômes de degré  $n$  convergeant simplement vers  $f$  sur l'intervalle  $[a, b]$ . En utilisant l'interpolation de Lagrange sur  $(n + 1)$  points distincts de  $[a, b]$ , montrer que  $f$  est un polynôme de degré  $\leq n$  et que la convergence est uniforme.

### Exercice 4. Une autre interpolation

Soient des points  $(x_i, y_i)$ , avec les  $x_i$  ordonnés:  $x_0 < x_1 < \dots < x_n$ . Soit  $s$  une fonction  $[a, b] \rightarrow \mathbb{R}$ , avec  $a = x_0, b = x_n$ , polynômiale de degré 3, vérifiant

- $s(x_i) = y_i$ , pour  $i = 0, \dots, n$
- $s$  est de classe  $C^2$

1. A quel type d'interpolation correspond cette fonction  $s$ ?
2. Montrer que  $\forall f : [a, b] \rightarrow \mathbb{R}$  vérifiant les deux propriétés précédentes, et telle que

$$s''(b)(f'(b) - s'(b)) = s''(a)(f'(a) - s'(a))$$

on a:

$$\int_a^b s''(x)^2 dx \leq \int_a^b f''(x)^2 dx$$

## 5.2

### Exercice 5. Courbes de Béziérs 1

Soit  $C_1$  le cercle unité dans  $\mathbb{R}$ .

1. Ecrire l'équation de la courbe polynomiale de degré  $\leq 4$  qui approche  $C_1$  au voisinage de  $t = 0$ ,  $t$  étant le paramètre de la courbe polynomiale.
2. En utilisant la formule des dérivées successives à l'origine, déterminer les coordonnées des points de contrôle de la courbe.
3. On souhaite que la courbe soit fermée. En supposant les points  $P_0$  et  $P_1$  fixés, à quelle condition sur les autres points peut-on obtenir une courbe fermée de classe  $C^1$ ? De classe  $C^2$ ?

- On suppose maintenant que la courbe est fermée et de classe  $C^1$ .  $P_0, P_1, P_3$  et  $P_4$  sont fixés conformément au résultat de la question précédente, et  $P_2$  est libre. En utilisant l'algorithme de Casteljau, calculer  $BP[P](1/2)$ . A quelle condition ce point se trouve-t-il sur l'axe des abscisses?

### Exercice 6. Courbes de Béziérs 2

- On se donne quatre points  $P_0, P_1, P_2, P_3$ . On place  $P_0$  à l'origine. Déterminer une condition sur la disposition de  $P_1, P_2$  et  $P_3$  pour que la courbe de Béziérs associée à ces points admette une représentation univoque de la forme  $y = f(x)$ .
- On suppose maintenant que  $P_0 = P_3 = (0, 0)$ ,  $P_1 = (1, 1)$ . Expliciter la courbe quand  $P_2 = P_1$ , puis lorsque  $P_2$  est sur la droite d'équation  $y = x$ .
- Les points  $P_0, P_1$  et  $P_3$  sont inchangés. Déterminer une condition sur  $P_2$  pour que la courbe traverse l'axe des  $x$ .

## 5.3

**Exercice 7.** Soit  $\alpha \in ]0, 1]$ .

- Calculer  $\omega_0$  et  $\omega_1$  pour que la formule de quadrature  $\int_0^1 f(x)dx \approx \omega_0 f(0) + \omega_1 f(\alpha)$  soit exacte pour toute fonction polynomiale de degré 1.
- Déterminer  $\alpha$  pour que la formule soit exacte pour toute fonction polynomiale de degré 2. Quelles sont alors les valeurs de  $\omega_0$  et  $\omega_1$ ?

**Exercice 8.** On se donne pour  $n \in \mathbb{N}$ ,  $n + 1$  réels distincts  $(x_i)_{0 \leq i \leq n}$  d'un intervalle  $[a, b]$  de  $\mathbb{R}$ ., et  $f$  une fonction de  $[a, b]$  dans  $\mathbb{R}$ .

On s'intéresse à une méthode d'intégration du type:

$$\int_a^b f(x)dx \approx \sum_{i=0}^n W_i f(x_i)$$

où les  $W_i$  sont  $n + 1$  réels quelconques.

- Montrer que la formule de quadrature ci-dessus est d'ordre  $n$ , (c'est-à-dire exacte pour les polynômes de degré  $\leq n$ ) ssi

$$\forall j \in \{0, \dots, n\} \int_a^b x^j dx = \sum_{i=0}^n W_i x_i^j \quad (*)$$

- Ecrire le système  $(*)$  sous forme matricielle

3. Pourquoi la matrice suivante est-elle inversible?

$$M = \begin{pmatrix} 1 & 1 & \dots & \dots & 1 \\ x_0 & x_1 & \dots & \dots & x_n \\ x_0^2 & x_1^2 & \dots & \dots & x_n^2 \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ x_0^n & x_1^n & \dots & \dots & x_n^n \end{pmatrix}$$

4. En déduire que si la méthode de quadrature est d'ordre  $n$  alors les coefficients  $(W_i)_{0 \leq i \leq n}$  sont uniques et donner leur expression théorique.

5. Que peut-on dire sur les  $(W_i)_{0 \leq i \leq n}$  si la méthode est d'ordre  $< n$ .

## 5.4

**Exercice 9.** Le  $k$ -ème polynôme de Tchebychev est défini par:

$$\begin{aligned} C_k(x) &= \cos(k \arccos(x)) \text{ pour } |x| \leq 1 \\ &= 0.5[(x + \sqrt{x^2 - 1})^k + (x - \sqrt{x^2 - 1})^k] \text{ pour } |x| > 1 \end{aligned}$$

1. Etablir la relation de récurrence suivante:

$$\begin{aligned} C_{k+2}(x) &= 2xC_{k+1}(x) - C_k(x) \\ C_1(x) &= x \\ C_0(x) &= 1 \end{aligned}$$

En déduire que  $C_k$  est effectivement un polynôme de degré  $k$ .

2. En quels points s'annule  $C_k$ ?

3. Montrer que sur l'intervalle  $[-1, 1]$ ,  $C_k$  admet  $k + 1$  extrema situés aux points  $x_i = \cos(\frac{i\pi}{k})$ , avec  $i \in \{0, \dots, k\}$  sur lesquels il prend les valeurs  $+1$  et  $-1$ , et que  $C_k(x_i)C_k(x_{i+1}) = -1$ , pour  $i \in \{0, \dots, k - 1\}$ .

4. Montrer que  $C_k$  est solution du problème de minimisation suivant:  
Soit  $a \in \mathbb{R}$ ,  $|a| > 1$ , trouver un polynôme  $p$  de degré  $\leq k$  tel que:

$$\max_{-1 \leq x \leq 1} |p(x)| = \inf_{q \in P_k, q(a) = C_k(a)} \max_{-1 \leq x \leq 1} |q(x)|$$

5. Montrer que la solution du problème de minimisation précédent est unique.

6. Montrer que les polynômes de Tchebychev sont orthogonaux sur  $[-1, 1]$  pour le poids  $w(x) = \frac{1}{\sqrt{1-x^2}}$ , c'est-à-dire:

$$\int_{-1}^1 C_n(x)C_m(x) \frac{dx}{\sqrt{1-x^2}} = c_n \delta_n^m \frac{\pi}{2}$$

avec  $c_n = 2$  si  $n = 0$  et  $c_n = 1$  sinon.

## 5.5

### Exercice 10. Méthodes de quadrature de Gauss à poids

En cours ont été abordées les formules de Gauss, permettant de calculer une approximation à l'ordre  $2n + 1$  de l'intégrale  $\int_a^b f(x)dx$  avec la formule de quadrature  $\sum_{i=0}^n c_i f(x_i)$  avec les points  $x_i$  et les coefficients  $c_i$  choisis de manière appropriée.

Ici nous nous intéressons à des formules de quadrature du type

$$\int_a^b f(x)w(x)dx \simeq \sum_{i=0}^n c_i f(x_i) \quad (*)$$

où le poids  $w(x)$  est une fonction strictement positive donnée.

Dans ce cas, les  $c_i$  sont définis par:

$$c_i = \int_a^b \prod_{p \neq i} \left( \frac{x - x_p}{x_i - x_p} \right) w(x) dx.$$

1. On pose  $L(x) = \prod_{k=0}^n (x - x_k)$ . Montrer que la formule (\*) est de degré  $2n+1$  ssi  $\int_a^b x^p L(x)w(x)dx = 0, \forall p \in 0, \dots, n$ .
2. Montrer que la formule (\*) ne peut pas être d'ordre plus élevé que  $2n + 1$ .

On appelle polynômes orthogonaux relativement à un poids  $w$  positif et intégrable sur un intervalle  $[a, b]$  la suite des polynômes  $P_0, P_1, \dots, P_n, \dots$  ayant les propriétés suivantes:

- $\forall n, P_n$  est de degré  $n$
  - $\forall n, \forall P$  de degré  $k \leq n - 1, \int_a^b P_n(x)P(x)w(x)dx = 0$ .
3. Montrer que quel que soit le poids  $w > 0$  intégrable sur  $[a, b]$ , il existe une unique suite de polynômes orthogonaux dont le coefficient de plus haut degré vaut 1, qui satisfont à:

$$P_n = x^n - \sum_{i=0}^{n-1} c_{in} P_i$$

avec

$$c_{in} = \frac{\int_a^b x^n P_i(x)w(x)dx}{\int_a^b P_i(x)P_i(x)w(x)dx}.$$

4. Soit  $w$  un poids intégrable et strictement positif sur  $[a, b]$ . Montrer que, quel que soit  $n$ , toutes les racines du  $n$ ème polynôme orthogonal  $P_n$  sont simples et sont dans  $]a, b[$ .

**Exercice 11. Cas particulier: formule de Gauss-Tchebychev**  
 Les polynômes de Tchebychev sont définis sur l'intervalle  $[-1, 1]$  par:

$$T_n(x) = \cos(n \arccos(x))$$

1. Pour  $n$  entier  $\geq 0$ , et  $\alpha \in ]0, \pi[$ , on pose:

$$I_n(\alpha) = \int_0^\pi \frac{\cos(n\theta) - \cos(n\alpha)}{\cos(\theta) - \cos(\alpha)} d\theta.$$

Montrer que  $\forall n \geq 0, I_{n+2}(\alpha) + I_n(\alpha) = 2\cos(\alpha)I_{n+1}(\alpha)$ .

2. En déduire que  $\forall n \geq 0, I_n(\alpha) = \pi \frac{\sin(n\alpha)}{\sin(\alpha)}$ .

3. On note  $s_0, s_1, \dots, s_n$  les  $(n+1)$  racines du polynôme de Tchebychev  $T_{n+1}$ , et  $\forall 0 \leq k \leq n, Q_k(x) = \frac{T_{n+1}(x)}{x-s_k}$ . Montrer que pour tout  $k$  compris entre 0 et  $n$ :

$$\int_{-1}^1 \frac{Q_k(x)}{\sqrt{1-x^2}} = \pi \frac{\sin((n+1)\theta_k)}{\sin(\theta_k)}$$

avec  $\theta_k = \arccos(s_k)$ .

4. En déduire que

$$\int_{-1}^1 \frac{Q_k(x)}{\sqrt{1-x^2}} = \frac{\pi}{n+1} Q_k(s_k)$$

5. Soit  $l_k$  la suite des polynômes élémentaires de Lagrange, et soit  $L(x) = \prod_{k=0}^n (x-s_k)$ . Montrer que:

$$l_k(x) = \frac{L(x)}{(x-s_k)L'(s_k)} = \frac{Q_k(x)}{T'_{n+1}(s_k)}$$

6. Montrer que la formule de Gauss-Tchebychev est vraie pour des polynômes de degré  $\leq 2n+1$ :

$$\int_{-1}^1 \frac{f(x)}{\sqrt{1-x^2}} dx \simeq \frac{\pi}{n+1} \sum_{i=0}^n f\left(\cos\left(\frac{(2i+1)\pi}{2(n+1)}\right)\right)$$

## 5.6

### Exercice 12. Convolution

On définit la convolution de 2 suites  $N$ -périodiques  $y$  et  $z$  par:  $(y * z)_k = \sum_{j=0}^{N-1} y_{k-j} z_j$   
 Montrer que  $y * z$  est également  $N$ -périodique, et que:

$$F_N(y * z)_i = NF_n(y)_i \times F_N(z)_i$$

où  $F_N(y)$  désigne la transformée de Fourier discrète de  $y$ .

### Exercice 13. Egalité de Parseval

Démontrer l'égalité de Parseval pour la transformée de Fourier discrète:

$$N \sum_{k=0}^{N-1} |z_k|^2 = \sum_{k=0}^{N-1} |y_k|^2 \quad \text{avec } z_k = \frac{1}{N} \sum_{j=0}^{N-1} y_j \omega_N^{-jk}$$

### Exercice 14. Théorème d'échantillonnage

1. Pour une suite  $y_0, y_1, y_2, \dots, y_{N-1}$  donnée, soit  $\{z_k\}$  sa transformée de Fourier discrète. Si  $N$  est pair, montrer que le polynôme trigonométrique défini par:

$$p_N(x) = \sum_{k=-N/2}^{N/2-1} z_k e^{ikx}$$

vérifie:  $p_N(x_j) = y_j$  avec  $x_j = \frac{2\pi j}{N}$

2. On suppose que  $f(x) = \sum_{k=-\infty}^{+\infty} c_k e^{ikx}$ , et  $y_j = f(x_j)$ . Montrer que

$$z_k - c_k = \sum_{l=-\infty, l \neq 0}^{+\infty} c_{k+lN}$$

3. En déduire que  $|f(x) - p_N(x)| \leq 2 \sum_{|k| \geq N/2} |c_k|$  Quelle conclusion peut-on en tirer?

### Exercice 15. Décroissance rapide

Montrer que si  $f$  est une fonction périodique de classe  $C^k$  alors  $c_n \sim O(|n|^{-k})$  quand  $n \rightarrow +\infty$ .

### Exercice 16. Calcul

Calculer, par les formules usuelles et par l'algorithme de FFT, la transformée de Fourier discrète de  $\{0, 1, 2, 3, 0, -3, -2, -1\}$ .