

République Algérienne Démocratique et Populaire  
Ministère de l'Enseignement Supérieur et de la Recherche  
Scientifique

11/5/10.06

Université 8 Mai 1945 Guelma

Faculté des Mathématiques et de l'Informatique  
et des Sciences de la Matière  
Département de Mathématiques



### Mémoire

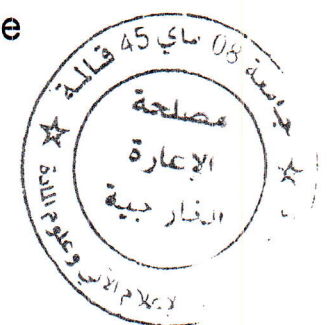
Présenté en vue de l'obtention du diplôme de  
**Master Académique en Mathématiques**

Option : **Analyse**

Par :

**Ounaissia Nadia**

**Houamria Radia**



### **Intitulé**

*Une méthode spectrale pour la résolution d'une  
équation différentielle*

Dirigé par : Dr Benrabah.R

Devant le jury

**PRESIDENT** Hitta. A  
**RAPPORTEUR** Benrabah.R  
**EXAMINATEUR** Lakhal. F

**Univ-Guelma**  
**Univ-Guelma**  
**Univ-Guelma**

Session Juin 2013

# Table des Matières

3

---

3.1 Méthode de différences finies : Principe de la méthode.....	36
3.1.1 Questions d'analyse numérique.....	38
3.2 Etude de la méthode différences finies pour un problème elliptique unidimen- -sionnel.....	39
3.3 Convergence de la méthode des différences finies.....	40
3.3.1 Programme de comparaison.....	45
3.4 Conclusion.....	48
<b>Bibliographie</b>	<b>50</b>

# Chapitre 1

## INTRODUCTION

Durant les dix dernières années beaucoup d'efforts ont été effectués dans le développement de méthodes numériques capables de caractériser des structures guide d'ondes composites.

Un grand nombre de méthodes telles la méthode d'adaptation modale (Mode-Matching Method), la technique de la résonance transverse, la méthode des lignes, et la méthode spectrale, ont été proposées et raffinées selon les applications.

En générale le choix de la méthode numérique est basé sur un compromis entre la précision, l'efficacité, l'adaptation à la structure étudiée, etc... et ce choix n'est pas forcément unique. Pour des lignes de transmission planaires, simple couches ou multicouche de diélectrique, la méthode approuvée et utilisée est la méthode spectrale.

L'utilisation de la méthode des moments et plus particulièrement de la méthode de Galerkin, permet d'obtenir un résultat précis bien que le déterminant associé à l'équation caractéristique, soit de dimension faible. La méthode est habituellement employée en choisissant un jeu de fonctions de base qui satisfait les conditions de singularité sur le bord des conducteurs, pour l'approche du courant sur les rubans ou des champs dans les fentes.

D'un point de vue mathématique, la méthode spectrale transforme un problème d'équation intégrale dans le domaine spatial en une équation algébrique dans le domaine spectral associé à la théorie des transformations de Fourier. Son principal avantage est la simplification du traitement analytique et numérique.

## 1.1 Polynômes orthogonaux et les méthodes spectrales

Les méthodes spectrales ont commencé à être utilisées lors de l'apparition des premiers algorithmes de transformée de Fourier rapide (Cooley & Turkey 1965). On notera cependant que l'algorithme avait déjà été décrit au XIX<sup>e</sup> siècle (Gauss 1866). Depuis, elles ont souvent été utilisées dans les problèmes turbulence homogène isotrope (Orszag & Patterson 1972) car elles permettent d'obtenir une vitesse de convergence très élevée et donc excellente résolution des petites échelles.

### 1.1.1 Polynômes orthogonaux

**Définition 1.1.1.** Soit  $(p_n)_{n \in \mathbb{N}}$  une famille de polynômes, cette famille sera orthogonale relativement à une fonction-poids  $\omega$  donnée, sur un intervalle  $[a, b]$  si

$$\int_a^b p_n(x)p_m(x)\omega(x)dx = 0, \quad (1.1)$$

pour tout  $n, m$  ;  $n \neq m$ , et

$$\int_a^b (p_n(x))^2 \omega(x)dx = c(n), \quad (1.2)$$

où  $c$  est une constante ne dépend que de  $n$ .

Toute famille de polynômes  $p_0, p_1, \dots$ , où chaque  $p_n$  est de degré  $n$ , est une base de l'espace vectoriel  $\mathbb{k}[x]$  (de dimension infinie) de tous les polynômes, « adaptée au drapeau  $(\mathbb{k}_n[x])_{n \in \mathbb{N}}$  ».

**Propriétés des polynômes orthogonaux** Toute suite  $(p_n)_{n \in \mathbb{N}}$  de polynômes orthogonaux possède un grand nombre de propriétés remarquables. Pour commencer :

**Lemme 1.1.1.**  $(p_0, \dots, p_n)$  est une base de  $\mathbb{k}_n[x]$ .

**Lemme 1.1.2.**  $p_n$  est orthogonal à  $\mathbb{k}_{n-1}[x]$ .

**Lemme 1.1.3.** Pour toute suite de polynômes orthogonaux, il existe une relation de récurrence relativement à trois polynômes consécutifs

$$p_{n+1} = (a_n x + b_n) p_n - c_n p_{n-1}$$

Les coefficients  $a_n, b_n, c_n$  sont donnés par

$$a_n = \frac{k_{n+1}}{k_n}, \quad b_n = a_n \left( \frac{k'_{n+1}}{k_{n+1}} - \frac{k'_n}{k_n} \right), \quad c_n = a_n \left( \frac{k_{n-1} h_n}{k_n h_{n-1}} \right),$$

où  $k_j$  et  $k'_j$  désignent les deux premiers coefficients de  $p_j$  :

$$p_j(x) = k_j x^j + k'_j x^{j-1} + \dots$$

et  $h_j$  le produit scalaire de  $p_j$  par lui-même :  $h_j = \langle p_j, p_j \rangle$ .

**Proposition 1.1.1.** Pour tout entier positif  $n$ , les zéros du polynôme  $p_n$  sont réels, distincts, strictement compris entre  $-1$  et  $1$ .

*Démonstration.* Soit  $l$  le nombre de zéros distincts de  $p_n$ , qui sont réels, d'ordre impair et strictement compris entre  $-1$  et  $1$ . Notons  $x_1, x_2, \dots, x_l$  les zéros correspondants. Si  $l$  est inférieur à  $n$ , le polynôme  $p_n$  sera orthogonal au polynôme  $(x - x_1) \cdot \dots \cdot (x - x_l)$ , d'où  $\int_{-1}^1 p_n(x)(x - x_1) \cdot \dots \cdot (x - x_l) \omega(x) dx = 0$ , ce qui est impossible puisque la fonction intégrée ne change pas de signe. Donc  $l = n$ . □

### 1.1.2 Polynôme de Legendre

On appelle *polynôme de Legendre* de degré  $n$ , le polynôme défini par la relation de récurrence

$$(n + 1) L_{n+1}(x) = (2n + 1)xL_n(x) - nL_{n-1}(x) \tag{1.3}$$

initialisée par

$$L_0(x) = 1, \quad L_1(x) = x$$

(1) Les polynômes de Legendre sont des polynômes orthogonaux relativement à la fonction de poids  $\omega(x) = 1$  sur l'intervalle  $\Omega = [-1, 1]$ .

$$\int_{-1}^{+1} L_n(x) L_m(x) dx = \frac{2}{2n + 1} \delta_{n,m}$$

En particulier,  $L_1(x) = 1$  et

$$\|L_n\|_2 = \left( \int_{-1}^{+1} L_n^2(x) dx \right)^{1/2} = \sqrt{\frac{2}{2n + 1}}$$

(2) Pour tout entier  $n \geq 0$ , le polynôme  $L_n$  vérifie l'équation différentielle

$$\frac{d}{dx} \left( (1-x^2) L'_n(x) \right) + n(n+1)L_n(x) = 0 \quad (1.4)$$

*Démonstration.* On remarque que le polynôme  $\frac{d}{dx}((1-x^2)L'_n(x))$  est de degré  $\leq n$  et qu'il vérifie, pour tout polynôme  $\varphi$  de degré  $\leq n-1$

$$\begin{aligned} \int_{-1}^1 \frac{d}{dx} \left( (1-x^2) L'_n(x) \right) \varphi(x) dx &= - \int_{-1}^1 (1-x^2) L'_n(x) \varphi'(x) dx \\ &= \int_{-1}^1 L_n(x) \frac{d}{dx} \left( (1-x^2) \varphi'(x) \right) dx \end{aligned}$$

Comme  $\frac{d}{dx}((1-x^2)\varphi'(x))$  est alors un polynôme de degré  $\leq n-1$ , ceci entraîne

$$\int_{-1}^1 \frac{d}{dx} \left( (1-x^2) L'_n(x) \right) \varphi(x) dx = 0$$

On en déduit qu'il existe un nombre réel  $\lambda_n$  tel que

$$\frac{d}{dx} \left( (1-x^2) L'_n(x) \right) + \lambda_n L_n(x) = 0$$

En comparant les coefficients dominants, on trouve bien  $\lambda_n = n(n+1)$  □

L'opérateur  $A$  défini par

$$A\varphi = -\frac{d}{dx} \left( (1-x^2) \varphi'(x) \right) \quad (1.5)$$

est clairement auto-adjoint et positif dans  $L^2(\Omega)$ . Il est de type Sturm-Liouville. L'équation (1.4) se traduit par le fait que les polynômes de Legendre en sont des *fonctions propres*, ceci est à l'origine du qualificatif "spectral" des méthodes exposées ici.

Une conséquence de l'équation (1.4) est que l'on obtient par intégration par parties, pour tous entiers  $m$  et  $n$  positifs ou nuls

$$\int_{-1}^1 L'_m(x) L'_n(x) (1-x^2) dx = n(n+1) \int_{-1}^1 L_m(x) L_n(x) dx \quad (1.6)$$

Ceci signifie que les  $L'_n$ ,  $n \geq 1$  forment une famille de polynômes deux à deux orthogonaux dans l'espace  $L^2_{1-x^2}(\Omega)$ .

(3) Les polynômes de Legendre satisfont la relation de récurrence

$$(1-x^2) L'_n(x) = -nxL_n(x) + nL_{n-1}(x)$$

(4) Formule de Rodrigues

$$L_n(x) = \frac{1}{2^n n!} \frac{d^n}{dx^n} ((x^2 - 1)^n)$$

(5) Majorations

$$\forall x \in [-1, 1], \quad |L_n(x)| \leq 1$$

$$\forall x \in [-1, 1], \quad |L'_n(x)| \leq \frac{n(n+1)}{2}$$

$$\forall x \in [-1, 1], \quad |L_n(x)| \leq \frac{1}{\sqrt{8\pi n(1-x^2)}}$$

$$\frac{1 - L_n^2(x)}{(2n-1)(n+1)} \leq L_n^2(x) - L_{n-1}(x)L_{n+1}(x) \leq \frac{2n+1}{3n(n+1)}$$

(6) Les polynômes de Legendre vérifient les relations

$$L_n(\pm 1) = (\pm 1)^n.$$

(7) Les polynômes de Legendre vérifient la formule

$$\int_{-1}^{+1} L_n(x) \frac{dx}{\sqrt{1-x^2}} = \frac{2^{3/2}}{2n+1}$$

(8) On peut aussi définir cette suite de polynômes par sa fonction génératrice

$$\varphi(z) = \frac{1}{\sqrt{1-2xz+z^2}} = \sum_{n=0}^{\infty} L_n(x)z^n.$$

### Relation de Récurrence et équation différentielle.

Rappelons la relation de récurrence entre les polynômes orthogonaux

$$p_{n+1} = (a_n x + b_n) p_n - c_n p_{n-1}, \text{ avec } a_n = \frac{k_{n+1}}{k_n} \text{ et } c_n = \frac{a_n}{a_{n-1}}, c_0 = 0.$$

Dans le cas des polynômes de Legendre, on a

$$a_n = \frac{2n+1}{n+1}, \quad b_n = 0 \text{ et } c_n = \frac{n}{n+1},$$

soit la relation

$$L_{n+1}(x) - \frac{2n+1}{n+1} x L_n(x) + \frac{n}{n+1} L_{n-1}(x) = 0$$

$$\iff (n+1)L_{n+1}(x) - (2n+1)xL_n(x) + nL_{n-1}(x) = 0 \tag{1.7}$$

Si on dérive  $\varphi$  par rapport à  $x$ , on a

$$\frac{\partial \varphi(z, x)}{\partial x} = z(1 - 2zx + z^2)^{-\frac{3}{2}} = \frac{z}{1 - 2zx + z^2} \varphi(z, x)$$

$$\iff (1 - 2zx + z^2) \frac{\partial \varphi(z, x)}{\partial x} - z\varphi(z, x) = 0;$$

remplaçons  $\varphi$  par son expression en fonction des  $L_n(x)$  et identifions les termes

$$L'_{n+1}(x) - 2xL'_n(x) + L'_{n-1}(x) - L_n(x) = 0, \quad n = 0, 1, \dots \quad (1.8)$$

Dérivons (1.7)

$$(n+1)L'_{n+1}(x) - (2n+1)L_n(x) - (2n+1)xL'_n(x) + nL'_{n+1}(x) = 0$$

et éliminons  $L'_{n-1}(x)$  avec (1.8)

$$L'_{n-1}(x) - xL'_n(x) = (n+1)L_n(x), \quad n = 0, 1, 2, \dots \quad (1.9)$$

De même éliminons  $L'_{n+1}(x)$ , toujours avec (1.8)

$$xL'_n(x) - L'_{n-1}(x) = nL_n(x), \quad n = 1, 2, \dots \quad (1.10)$$

Ajoutons (1.9) + (1.10)

$$L'_{n+1}(x) - L'_{n-1}(x) = (2n+1)L_n(x), \quad n = 1, 2, \dots \quad (1.11)$$

enfin remplaçons  $n$  par  $n-1$  dans (1.10) et éliminons  $L'_{n-1}(x)$ , ce qui donne

$$(1-x^2)L'_n(x) = nL_{n-1}(x) - nxL_n(x), \quad n = 1, 2, \dots \quad (1.12)$$

On obtient ainsi une expression de  $L'_n(x)$  en fonction des polynômes de Legendre ; si on dérive (1.12) et que l'on réutilise (1.9) pour éliminer  $L'_{n-1}(x)$ , on obtient

$$[(1-x^2)L'_n(x)]' + n(n+1)L_n(x) = 0, \quad n = 0, 1, 2, \dots \quad (1.13)$$

Ceci montre que les  $L_n$  sont solution de l'équation

$$[(1-x^2)y']' + n(n+1)y = 0 \iff (1-x^2)y'' - 2xy' + n(n+1)y = 0$$



**Exemple 1.** (*Evaluation d'une combinaison linéaire de polynômes de Legendre*)

1. **Première étape** : Écriture d'une fonction  $y = \text{CombLinLeg}(x, c)$  calculant les valeurs d'une combinaison linéaire donnée de polynômes de Legendre (1.14) en des points constituant les abscisses d'un vecteur  $x$

$$y = \sum_{k=1}^p c_k L_{k-1}(x) \quad (1.14)$$

2. **Deuxième étape** : Écriture d'un script appelant la fonction  $\text{CombLinLeg}$ , avec  $x$  qui discrétise l'intervalle  $[-1, 1]$  suffisamment finement pour une bonne représentation graphique et  $c$  contenant les coefficients correspondants à la combinaison linéaire  $L_0 - 2L_1 + 3L_5$ .

Le programme suivant réalise ces deux étapes :

---

```

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% fonction y=CombLinLeg(x,c)
%Evaluation d'une combinaison linéaire en des points x1,...xn
% Arguments d'entrée :
% x : un vecteur de composantes (x1<x2<...<xn)
% c : un vecteur de composantes (c1,c2,...,cp)
% Argument de sortie :
% y : un vecteur de composantes (y1,y2,...yn)
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
n=length(x); % nombre de points où on calcule la valeur de la combinaison linéaire
p=length(c); % degré maximum des polynômes de la combinaison linéaire
if p>0,
    pol1=ones(1,n);
    y=c(1)*pol1;
    else
    y=[];
end
if p>1,
    pol2=x;
```

```

y=c(1)*pol1+c(2)*pol2;
  for k=3 :p
    pol=((2*k-3)*x.*pol2-(k-2)*pol1)/(k-1);
    y=y+c(k)*pol;
    pol1=pol2;
    pol2=pol;
  end
end

```

---

La représentation graphique de  $L_0 - 2L_1 + 3L_5$  sur l'intervalle  $[-1, 1]$  est faite dans le script :

```

PlotPolLeg
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% PlotPolLeg : Comparaison du temps de calcul avec la fonction matlab legendre
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
close all
x=linspace(-1,1,500); % abscisses de représentation graphique
deg=5;
c=[1;-2;0;0;0;3]
%deg=50;
%c=rand(deg+1,1);
  tic;
y=CombLinLeg(x,c);
Rec=toc;
tic;
L=legendre(0,x);
z=c(1)*L(1, :);
  for i=1 :deg
    L=legendre(i,x);
    z=z+c(i+1)*L(1, :);
  end

```

```
end
    Mat=toc;
plot(x,y,x,z)
legend('récurrence','matlab')
title(['Comb. lin. de Pol. de Legendre de degré ',num2str(deg)])
fprintf('temps de calcul par récurrence= %f et par matlab= %f',Rec,Mat);
```

---

**Remarque 1.1.1.** *Il existe une fonction Matlab , de syntaxe d'appel  $L=legendre(n,x)$  qui renvoie un tableau  $L$  de  $(n + 1)$  lignes , dont la  $(m + 1)$ -ième ligne contient les valeurs de la fonction de Legendre aux points spécifiés dans le vecteur  $x$ . Donc on peut calculer les valeurs du polynôme de Legendre avec cette fonction, en n'utilisant que la première ligne du tableau renvoyé en sortie.*

**Remarque 1.1.2.** *Le script PlotPolLeg compare aussi les temps de calcul des deux méthodes, en appelant la fonction **tic** avant l'appel de **CombLinLeg** et la fonction **toc** just après. La valeur renvoyée par **toc** contient le temps d'exécution. On refait la même chose avant et après le groupe de commandes pour la méthode utilisant Legendre. Pour que les temps de calculs soit significatifs, il vaut mieux augmenter le nombre de points de calcul à 500 points, ainsi que le degré de la combinaison linéaire à 50. Le rapport entre les temps de calcul, alors supérieur à 100, est incontestablement en faveur du script **CombLinLeg**, l'utilisation de la fonction **legendre** dans ce contexte impliquant une grande quantité de calculs unitils ou bien redondants.*

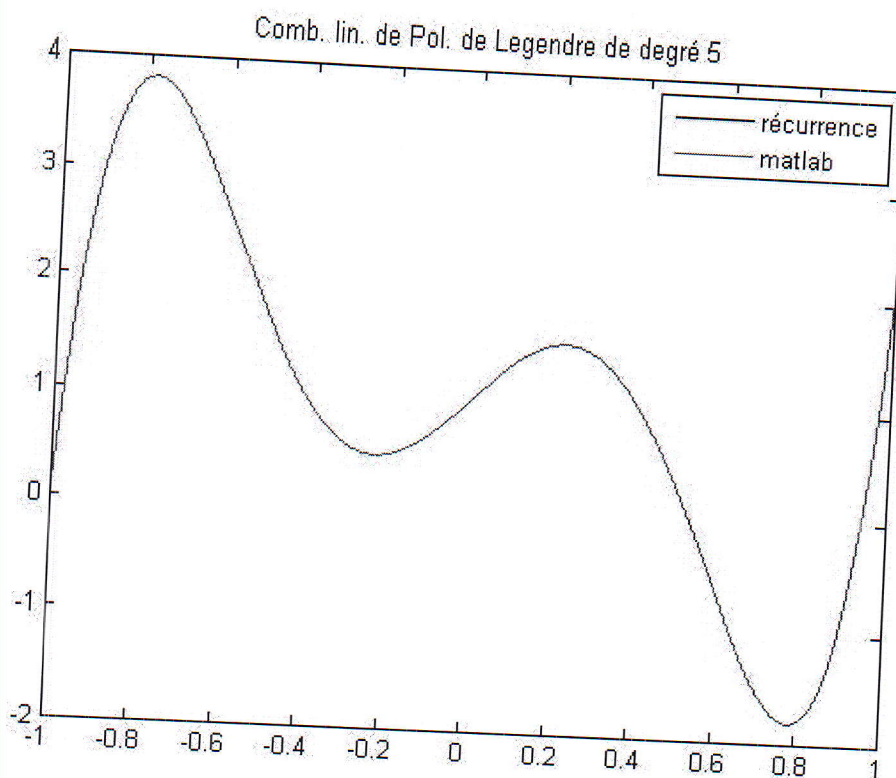


Figure 1.1 Combinaison linéaire de  $L_0 - 2L_1 + 3L_5$

## 1.2 Approximation polynômiale

### 1.2.1 Approximation polynomiale par morceaux

Soit  $f$  une fonction continue, définie sur l'intervalle  $[0,1]$ , on cherche à approcher  $f$  par une autre fonction  $S$  qui soit polynomiale par morceaux, une telle fonction est appelée *spline*. L'intérêt de l'approximation polynomiale par morceaux réside dans la possibilité de mieux contrôler les problèmes liés au manque de régularité globale de la fonction, un autre intérêt pratique est la stabilité dans les calculs : on commet moins d'erreur en calculant plusieurs polynômes de bas degré qu'en calculant un seul polynôme de degré élevé, à précision de calcul égale.

On découpe l'intervalle  $I = [0,1]$  en sous-intervalle  $I_i$  de même longueur,  $I_i = [x_i, x_{i+1}]$ ,  $x_i = ih$ ,  $h = 1/n$ . Sur chaque sous-intervalle  $I_i$  on approche  $f$  par un polynôme  $p_{n,i}$  de degré  $n$ . On note  $S_n$  la fonction polynomiale par morceaux coïncidant avec  $p_{n,i}$  sur l'intervalle  $I_i$ .

**Approximation constante par morceaux**

Soit  $S_0$  une fonction constante sur chaque intervalle  $I_i$  et coïncidant avec  $f$  aux points  $x_{i+1/2} = (x_i + x_{i+1})/2$

$$S_{0 \setminus I_i}(x) = f(x_{i+1/2}).$$

Supposons la fonction  $f$  de classe  $C^1$  sur  $I$ , pour tout  $x \in I_i = [x_i, x_{i+1}]$ ; il existe  $\xi_{x,i} \in I_i$  tel que

$$f(x) - S_0(x) = (x - x_{i+1/2})f'(\xi_{x,i}).$$

On en déduit, avec  $x_{i+1} - x_i = h = 1/n$ , que

$$\|f - S_0\|_\infty \leq \frac{h}{2} M_1,$$

où  $M_1$  est une borne supérieure de  $f'$  sur  $I$ . Donc, lorsque  $h \rightarrow 0$ ,  $S_0$  converge uniformément vers  $f$ .

**Approximation affine (linéaire)**

Cette fois, l'approximation  $S_1$  est affine sur chaque intervalle  $I_i$  et coïncide avec  $f$  aux points  $x_i$  et  $x_{i+1}$

$$S_{1, I_i}(x) = \frac{(f(x_{i+1}) - f(x_i))}{h} (x - x_i) + f(x_i)$$

Supposons d'abord que la fonction  $f$  soit de classe  $C^2$  sur  $I$ , pour tout  $x \in I_i = [x_i, x_{i+1}]$ , il existe  $\xi_{x,i} \in I_i$  tel que

$$f(x) - S_1(x) = \frac{(x - x_i)(x - x_{i+1})}{2} f''(\xi_{x,i})$$

On en déduit, avec  $x_{i+1} - x_i = h = \frac{1}{n}$ , que

$$\|f - S_1\|_\infty \leq \frac{h^2}{8} M_2$$

où  $M_2$  est une borne supérieure de  $f''$  sur  $I$ . Donc, lorsque  $h \rightarrow 0$ ,  $S_1$  converge uniformément vers  $f$ .

### Approximation cubique par morceaux (spline cubique)

Cette fois, l'approximation  $S_3$  est de classe  $C^2$  sur  $I$ , cubique sur chaque  $I_i$  et qui coïncide avec  $f$  aux points  $x_i$  et  $x_{i+1}$ . Soit  $p_i$  la restriction de  $S_3$  à l'intervalle  $I_i$ . On peut chercher  $p_i$  sous la forme

$$p_i(x) = a_i(x - x_i)^3 + b_i(x - x_i)^2 + c_i(x - x_i) + d_i$$

Les fonctions  $f$  et  $p_i$  coïncidant aux points  $x_i$  et  $x_{i+1}$ , on a

$$d_i = f(x_i) \tag{1.15}$$

et

$$f(x_{i+1}) = a_i h^3 + b_i h^2 + c_i h + d_i \tag{1.16}$$

Nous allons exprimer les inconnues  $a_i$ ,  $b_i$  et  $c_i$  en fonction des dérivées secondes de la fonction.

On note  $\alpha_i = p_i''(x_i)$ . On écrivant la continuité de la dérivée seconde aux points  $x_i$  (i.e.  $p_{i-1}''(x_i) = p_i''(x_i)$ ), on obtient

$$b_i = \frac{1}{2}\alpha_i \tag{1.17}$$

$$a_i = \frac{\alpha_{i+1} - \alpha_i}{6h} \tag{1.18}$$

Substituant ces deux relations dans (1.16), on obtient

$$f(x_{i+1}) = \frac{\alpha_{i+1} - \alpha_i}{6} h^2 + \frac{1}{2}\alpha_i h^2 + c_i h + f_i$$

d'où l'expression de  $c_i$

$$c_i = \frac{f_{i+1} - f_i}{h} - \frac{2\alpha_i + \alpha_{i+1}}{6} h. \tag{1.19}$$

En écrivant maintenant la continuité de la dérivée première au point  $x_i$ , on obtient

$$3a_{i-1}h^2 + 2b_{i-1}h + c_{i-1} = c_i \tag{1.20}$$

Remplaçant les  $a_i$ ,  $b_i$  et  $c_i$  figurant dans (1.20) par leurs expressions données par (1.17), (1.18), (1.19), on obtient  $n - 1$  équations qui expriment une relation de récurrence entre  $\alpha_{i-1}$ ,  $\alpha_i$  et  $\alpha_{i+1}$

$$h(\alpha_{i-1} + 4\alpha_i + \alpha_{i+1}) = \frac{6}{h}(f_{i-1} - 2f_i + f_{i+1}).$$

À ces  $n - 1$  équations, il faut en ajouter deux autres pour fermer le système et déterminer les  $n + 1$  valeurs  $\alpha_i$ . On peut par exemple, poser  $\alpha_0 = \alpha_n = 0$ , on parle alors de spline naturelle.

On peut aussi fixer les pentes de la spline aux extrémités de l'intervalle. Dans le premier cas ( $\alpha_0 = \alpha_n = 0$ ), le vecteur  $\alpha = (\alpha_1, \dots, \alpha_{n-1})^t$  est solution du système linéaire tridiagonale

$$Ax = b$$

avec

$$A = h \begin{pmatrix} 4 & 1 & 0 & \dots & 0 \\ 1 & 4 & 1 & & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & & & 1 & 4 & 1 \\ 0 & \dots & 0 & 1 & 4 \end{pmatrix}, b = \frac{6}{h} \begin{pmatrix} f_0 - 2f_1 + f_2 \\ \vdots \\ f_{i-1} - 2f_i + f_{i+1} \\ \vdots \\ f_{n-2} - 2f_{n-1} + f_n \end{pmatrix}.$$

La matrice  $A$  est inversible, car à diagonale strictement dominante. Connaissant les  $\alpha_i$ , on détermine les  $p_i$  par les relations (1.15), (1.17), (1.18), (1.19) et la spline est parfaitement déterminée sur chaque intervalle.

## 1.3 Intégration numérique

D'une façon générale, les méthodes d'intégration numérique permettent de calculer une valeur approchée de l'intégrale d'une fonction quand on ne connaît pas sa primitive mais qu'on peut évaluer les valeurs de la fonction elle-même en un certain nombre de points. Ces méthodes se répartissent en deux grandes catégories : les méthodes composées dans lesquelles la fonction  $f$  est remplacée par un polynôme d'interpolation sur chaque intervalle élémentaire  $[x_i, x_{i+1}]$  de la subdivision et les méthodes de Gauss fondées sur les polynômes orthogonaux pour lesquelles les points de la subdivision sont imposés.

### 1.3.1 Intégration numérique par la méthode des trapèzes

Soit  $f$  une fonction continue sur  $[a, b]$ , dérivable sur  $]a, b[$  et  $a = x_0 < x_1 \dots < x_{n-1} < x_n = b$  une subdivision régulière de l'intervalle  $[a, b]$ . On note  $h$  le pas de cette subdivision. Dans la méthode des trapèzes, la fonction  $f$  est remplacée sur chaque intervalle  $[x_i, x_{i+1}]$  par la droite joignant les points  $(x_i, f(x_i))$  et  $(x_{i+1}, f(x_{i+1}))$ , soit

$$h(x) = \frac{(x - x_i)f(x_{i+1}) - (x - x_{i+1})f(x_i)}{x_{i+1} - x_i}; \quad x \in [x_i, x_{i+1}]$$

La méthode s'écrit

$$\int_a^b f(x) dx \simeq \sum_{i=0}^{n-1} (x_{i+1} - x_i) \frac{f(x_i) + f(x_{i+1})}{2}$$

Lorsque la subdivision se réduit à sa plus simple expression,  $x_0 = a$ ,  $x_1 = b$  on a

$$\int_a^b f(x) dx \simeq \frac{1}{2}(b-a)(f(a) + f(b))$$

La méthode des trapèzes est une méthode d'ordre 1. L'erreur dans la méthode des trapèzes est donnée par l'expression

$$\left| \int_a^b f(x) dx - S \right| \leq \frac{1}{12} \frac{(b-a)^3}{n^2} \sup_{x \in [a,b]} |f''(x)|$$

La somme  $S$  s'exprime par

$$S = \frac{h}{2} \left( f(a) + f(b) + \sum_{i=1}^{n-1} f(x_i) \right)$$

Pour améliorer la précision, on considère parfois la formule des trapèzes corrigée suivante

$$\int_a^b f(x) dx \simeq \frac{h}{2} \left( f(a) + f(b) + \sum_{i=1}^{n-1} f(x_i) \right) - \frac{h^2}{12} (f'(b) - f'(a)).$$

### 1.3.2 Intégration numérique par la méthode de Simpson

Dans la méthode de Thomas Simpson (1710-1761), la fonction  $f$  est remplacée par un polynôme du second degré définissant un arc de parabole passant par les points d'ordonnées  $f(x_i)$ ,  $f(x_{i+1})$  et  $f(x_{i+2})$ . La méthode s'écrit

$$\int_a^b f(x) dx \simeq \sum_{i=0}^{n-1} \frac{1}{6} (x_{i+1} - x_i) \left( f(x_{i+1}) + f(x_i) + 4f\left(\frac{x_{i+1} + x_i}{2}\right) \right)$$

Lorsque la subdivision se réduit à sa plus simple expression,  $x_0 = a$ ,  $x_1 = (a+b)/2$ ,  $x_2 = b$  la formule précédente devient

$$\int_a^b f(x) dx \simeq \frac{1}{3}(b-a) \left( f(a) + 4f\left(\frac{a+b}{2}\right) + f(b) \right)$$

La méthode de Simpson est une méthode d'ordre 4. L'erreur dans la méthode de Simpson est donnée par

$$\left| \int_a^b f(x) dx - S \right| \leq \frac{1}{2880} \frac{(b-a)^5}{n^4} \sup_{x \in [a,b]} |f^{(5)}(x)|$$

La somme  $S$  qui approche l'intégrale s'exprime par

$$S = \frac{h}{2} \sum_{i=0}^{n-1} \left( f(a + ih) + f(a + (i+1)h) + 4f\left(a + ih + \frac{h}{2}\right) \right).$$



### 1.3.3 Intégration numérique par quadrature de Gauss-Legendre

Lorsque la famille de polynôme orthogonaux est la famille de polynôme de Legendre relativement à la fonction de pondération  $\omega(x) = 1$  sur l'intervalle  $[-1, 1]$ , l'intégrale est approchée par la formule

$$\int_{-1}^1 f(x)dx = \sum_{i=1}^s f(x_i)\omega_i + R_s(f)$$

où les  $x_i$  sont les zéros du polynôme de Legendre  $L_s$  et où les poids  $\omega_i$  sont donnés par la relation

$$\omega_i = \frac{2}{(1-x_i^2)[L'_s(x_i)]^2} \cdot \tag{1.21}$$

Le reste

$$R_s(f) = \frac{2^{2s+1}(s!)^4}{(2s+1)[(2s)!]^3} f^{(2s)}(\xi), \quad \text{avec } \xi \in ]-1, 1[$$

est nul pour les polynômes de  $\mathbb{k}_{2s-1}$ . Pour ces fonctions, la formule de quadrature

$$\int_{-1}^1 f(x)dx \simeq \sum_{i=1}^s f(x_i)\omega_i, \tag{1.22}$$

est exacte.

Pour utiliser numériquement cette formule, il faut tout d'abord calculer les poids  $\omega_i$  et les points  $x_i$  d'intégration. On propose pour cela la méthode ci-dessous.

Le relation (1.3) sur les polynômes  $L_j$  pour  $j = 0, \dots, s$  peuvent s'écrire sous forme matricielle  $Mu = xu + v$ , en posant  $a_j = j/(2j + 1)$ ,  $b_j = (j + 1)/(2j + 1)$ , et

$$M = \begin{pmatrix} 0 & 1 & & & & & & & & & \\ a_1 & 0 & b_1 & & & & & & & & \\ & \ddots & \ddots & \ddots & & & & & & & \\ & & & a_{s-2} & 0 & b_{s-2} & & & & & \\ & & & & a_{s-1} & 0 & & & & & \end{pmatrix}, v = b_{s-1}L_s(x) \begin{pmatrix} 0 \\ \vdots \\ \vdots \\ 0 \\ 1 \end{pmatrix}, u = \begin{pmatrix} L_0(x) \\ \vdots \\ \vdots \\ \vdots \\ L_{s-1}(x) \end{pmatrix}$$

Pour  $x$  racine de  $L_s$ , on a  $v = 0$  et le système linéaire devient  $Mu = xu$ . Donc les valeurs propres de la matrice tridiagonale  $M$  sont les racines de  $L_s$ .

Pour calculer les poids  $\omega_i$ , donnés par la formule (1.21) on combine la relation de récurrence (1.3) avec la relation suivante

$$(1-x^2)L'_s(x) = -s x L_s(x) + s L_{s-1}(x), \quad s \geq 1.$$

Comme  $L_s(x_i) = 0$ , on obtient finalement

$$\omega_i = \frac{2(1 - x_i^2)}{(sL_{s-1}(x_i))^2}$$

et on calcule la valeur  $L_{s-1}(x_i)$  en utilisant la relation de récurrence (1.3).

**Calcul les poids et les Abscisses :**

Le calcul des abscisses et poids d'intégration de Gauss se fait dans la fonction **xwGauss** :

---

```

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% fonction [x,w]=xwGauss(s)
% Fonction pour calculer les abscisses et poids de Gauss de degré s
% Argument d'entrée :
% s : degré de la méthode d'intégration
% Arguments de sortie
% x : abscisses de Gauss x1...xs
% w : poids de Gauss w1...ws
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% Calcul des zéros du polynôme  $L_s$ 
A=zeros(s,s);
for i=1 :s-2
    A(i+1,i+2)=(i+1)/(2*i+1);
    A(i+1,i)=i/(2*i+1);
end
if s>1
    A(s,s-1)=(s-1)/(2*s-1);
    A(1,2)=1;
    x=eig(A);
    x=sort(real(x));
else
    x=[0];
end

```

```

% Calcul des poids
p0=ones(s,1);
p1=x;
for j=2 :s-1
    p2=(2*j-1)*x.*p1/j-(j-1)*p0/j;
    p0=p1;
    p1=p2;
end
if s==1
    w=[1];
elseif s==2;
    w=[0.5;0.5];
else
    w=2*(1-x.^2)./(s*p2).^2;
end

```

---

**Remarque 1.3.1.** Les abscisses sont les valeurs propres de la matrice  $M$ , on construit donc cette dernière et on utilise la fonction Matlab **eig** pour obtenir son spectre. Le script **TestIntGauss** teste la formule de quadrature sur une fonction régulière, ici le fonction **exp(x)**, et compare également cette méthode d'intégration avec la méthode proposée par Matlab, programmée dans la fonction **quad**.

---

Ce programme mesure le temps de calculs respectifs des deux méthodes : (Simpson et Gauss)

**Algorithme : TestIntGauss**

```

% Intégration d'une fonction test avec la quadrature de Gauss Legendre

```

```

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

```

```

[x,w]=xwGauss(20);

```

```

fprintf('fonction exp : int. exacte=%f quad. Gauss=%f\n',exp(1.)-exp(-1.),exp(x)'*w)

```

```

n=10
IPmat=quad(@exp,-1,1);
for i=2 :n
    [x,w]=xwGauss(i);
    fprintf('%d points pour exp : int. exacte=%f quad. Gauss=
%f quad. matlab=%f\n',i,exp(1.)-exp(-1.),exp(x)*w,IPmat)
end
nbeval=1000;          % nombre de calculs d'intégrales
tic;
n=4;                 % 4 points suffisent pour avoir
[x,w]=xwGauss(n); % 6 chiffres significatifs dans l'approximation
for i=1 :nbeval
    exp(x)*w;
end
cpu_qg=toc;
tic;
for i=1 :nbeval
    IPmat=quad(@exp,-1,1);
end
cpu_mat=toc;
fprintf(' %d évaluations de Quad Gauss en %f secondes, de quad
en %f secondes',nbeval,cpu_qg,cpu_mat);

```

---

**Remarque 1.3.2.** Pour comparer les performances de la méthode d'intégration de **quad** avec la méthode de **Gauss** du point vue temps de calcul, on choisit un nombre de points suffisant pour que la quadrature de Gauss donne la valeur de l'intégrale exacte avec six chiffres significatifs : quatre points suffisent pour la fonction  $\exp(x)$  par exemple et on mesure le temps de calculs respectifs des deux méthodes pour évaluer un grand nombre de fois la même intégrale, à l'aide des fonction *tic* et *toc*. La méthode de Simpson étant moins précise que la

*méthode de Gauss de degré 4, elle nécessite plus d'évaluation de l'intégrale pour obtenir la même précision et est plus lente. Donc il est avantageux d'utiliser la quadrature de Gauss.*

## Chapitre 2

# Méthode spectrale sur un problème unidimensionnel

### 2.1 Position du problème

Les méthodes spectrales sont des techniques d'approximation des solutions d'équations aux dérivées partielles. Elles consistent à rechercher la solution approchée dans un espace de polynômes. La précision de ces méthodes n'est limitée que par la régularité de la fonction à approcher, à l'opposé des approximations de type différences finies ou éléments finis basées sur une discrétisation du domaine de définition de la solution. La technique pour calculer la solution approchée repose essentiellement sur la formulation variationnelle du problème continue. L'approximation consiste à remplacer l'espace de fonctions test par un espace de polynômes et à évaluer les intégrales au moyen de formules de quadrature appropriées.

Cet mémoire propose de mettre en oeuvre une méthode spectrale pour résoudre numériquement le problème aux limites posé sur l'intervalle  $\Omega = [-1, 1]$

$$\begin{cases} -u'' + cu = f, \\ u(-1) = 0, \\ u(1) = 0, \end{cases} \quad (2.1)$$

avec  $f \in L^2(\Omega)$  et  $c$  un réel positif ou nul.

## 2.2 Présentation générale des méthodes spectrales

Le principe de base des méthodes spectrales est de décomposer la solution  $u$  d'une équation différentielle sur une base (infinie) de fonction orthogonale

$$u(x) = \sum_{k=0}^{\infty} \tilde{u}_k \phi_k(x) \quad (2.2)$$

où les fonctions  $\phi_k$  sont une base orthogonale de l'espace des fonctions sur  $(x_0, x_1)$  muni d'un produit scalaire, c'est-à-dire

$$(\phi_k, \phi_l) = \int_{x_0}^{x_1} \phi_k(x) \phi_l(x) \omega(x) dx = \delta_{kl}$$

où  $\omega$  la fonction de poids associée au produit scalaire. On peut alors vérifier aisément que les  $\tilde{u}_k$  s'obtient par projection de  $u$  sur la base

$$\tilde{u}_k = \int_{x_0}^{x_1} u(x) \phi_k(x) \omega(x) dx$$

La discrétisation numérique consiste alors à tronquer la base des  $\phi_k$  de sorte que l'on écrive la solution sous la forme

$$u_d(x) = \sum_{k=0}^N \hat{u}_k \phi_k(x)$$

où on notera que les coefficients  $\hat{u}_k$  sont *a priori* différents de ceux de la décomposition exacte (2.2).

On peut alors obtenir

**Méthode de Galerkin :** Dans cette méthode, on cherche à obtenir la projection exacte de  $u$  sur la base tronquée, c'est à dire  $\hat{u}_k = \tilde{u}_k$  pour  $k \in [0, N]$ . Cette méthode présente l'avantage de fournir comme solution une fonction de l'espace réel, qui peut donc être calculée sur une grille spatiale aussi fine que voulue.

## 2.3 Formulation variationnelle

La formulation variationnelle du problème (2.1) est la suivante

$$\left\{ \begin{array}{l} \text{Trouver } u \in H_0^1(-1,1) \\ a(u, v) = L(v) \quad \forall v \in H_0^1(-1,1) \\ u(1) = u(-1) = 0 \end{array} \right. \quad (2.3)$$

où

$$a(u, v) = \int_{-1}^1 u'(x)v'(x)dx + c \int_{-1}^1 u(x)v(x)dx \text{ et } L(v) = \int_{-1}^1 f(x)v(x)dx.$$

On cherche à montrer l'existence et l'unicité de la solution de (2.3). Pour cela, on utilise le théorème de Lax-Milgram, qu'on rappelle ici :

**Théorème 2.3.1. (Lax-Milgram)** Soient  $H$  un espace de Hilbert réel ou complexe muni de son produit scalaire noté  $\langle \cdot, \cdot \rangle$  de norme associée notée  $\|\cdot\|$ ,  $a(\cdot, \cdot)$  une forme bilinéaire (ou une forme sesquilinéaire si  $H$  est complexe) qui est

- continue sur  $H \times H$

$$\exists c > 0, \forall (u, v) \in H^2, |a(u, v)| \leq c \|u\| \|v\|$$

- coercive sur  $H$

$$\exists \alpha > 0, \forall u \in H, a(u, u) \geq \alpha \|u\|^2$$

et  $L(\cdot)$  une forme linéaire continue sur  $H$ . Sous ces hypothèses il existe un unique  $u$  de  $H$  tel que l'équation  $a(u, v) = L(v)$  soit vérifiée pour tout  $v$  de  $H$

$$\exists! u \in H, \forall v \in H, a(u, v) = L(v)$$

Si de plus la forme bilinéaire  $a$  est symétrique, alors  $u$  est l'unique élément de  $H$  qui minimise la fonctionnelle  $J : H \rightarrow \mathbb{R}$  définie par  $J(v) = \frac{1}{2}a(v, v) - L(v)$  pour tout  $v$  de  $H$ , c'est-à-dire

$$\exists! u \in H, J(u) = \min_{v \in H} J(v)$$

À noter qu'une solution régulière  $u$  (i.e dans  $H_0^2(-1,1)$ ) du problème aux limites (2.1) est aussi solution du problème (2.3) et réciproquement : toute solution régulière de (2.3) est une solution de (2.1).

## 2.4 Développement d'une fonction en série de Legendre

Pour une fonction  $f \in L^2(-1,1)$ , on définit la série de Legendre

$$\mathcal{L}(f) = \sum_{j=0}^{\infty} \hat{f}_j L_j$$



où les coefficients  $\widehat{f}_j$  sont donnés par

$$\widehat{f}_j = \frac{2j+1}{2} \int_{-1}^1 f(x) L_j(x) dx \quad (2.4)$$

On définit une approximation de la fonction  $f$  par sa série tronquée à l'ordre  $p$

$$\mathcal{L}_p(f) = \sum_{j=0}^p \widehat{f}_j L_j.$$

Cette approximation est exacte pour les polynômes de degré inférieurs à  $p$  puisque les  $(L_j)_{j=0}^p$  forment une base orthogonale de  $\mathbb{k}_p$ . Le calcul des coefficients  $\widehat{f}_j$  doit se faire par une formule de quadrature approchée. L'erreur induite par cette approximation des termes  $\widehat{f}_j$  doit être du même ordre ou négligeable devant l'erreur globale de la méthode, il faut donc utiliser une quadrature d'ordre élevée, et pour cela la méthode d'intégration de Gauss-Legendre étudiée au chapitre précédent est toute indiquée.

**Exemple 2.** (*Approximation de la fonction  $f$  avec  $\mathcal{L}_p(f)$* )

**Étape 1 :** Écriture d'un script qui calcule la série de Legendre  $\mathcal{L}_p(f)$  d'une fonction  $f$  à l'ordre  $p$  qui contient :

- Calcul des abscisses et poids d'intégration pour la formule (1.22) pour un  $s$  fixé.
- Calcul approché des coefficients  $(\widehat{f}_k)_{k=0}^p$  définie par (2.4)
- Représentation sur le même graphique de la fonction  $f$  et de sa série  $\mathcal{L}_p(f)$  sur l'intervalle  $[-1, 1]$ .

**Étape 2 :** Comparaisons avec l'exemple représenter sur la figure 2.1 pour

$$f(x) = \sin(6x) \exp(-x)$$

**Étape 3 :** Exécution le script avec des fonctions moins régulières. (On pourra utiliser par exemple les fonctions  $abs(x)$  et  $sign(x)$ .)

**Étape 4 :** Trace la courbe d'erreur en norme infinie entre la fonction et sa série en fonction du degré de la série.

**Remarque 2.4.1.** On commence par calculer les (absc et poids) par un programme qu'on appelle `xwGauss` :

Et un autre programme qui calcul une combinaison linéaire entre les polynômes de Legendre

### CombLinLeg

#### Algorithme : CalSerLeg

```

%%%%%%%%%%
% Function [x,y,err]=CalSerLeg(s,p,npt,Test)
%
% Calcul et représentation graphique de la série de Legendre d'une fonction
%
% Arguments d'entrée :
% s : degré de la quadrature de Gauss utilisée pour calculer les coefficients de la série
% p : degré de la série
% npt : nombre de points d'évaluation de la série, dans l'intervalle [-1,1]
% Test : fonction dont on calcule la série
% arguments de sortie :
% x : points d'évaluation de la série
% y : valeurs de la série aux points x
% err : erreur en norme infinie entre la fonction et sa série
%%%%%%%%%%
%
% Calcul des poids et abscisses d'intégration de Gauss de degré s
    [absc,poids]=xwGauss(s);
    t=feval(Test,absc);
% Calcul des polynomes de Legendre aux points absc, dans les tableaux LX0, LX1 et LX2.
% Calcul des coefficients de la série dans le tableau c
    u=t.*poids;
    LX0=ones(s,1);
    LX1=absc;
    c=zeros(p+1,1);
    c(1)=t'*poids/2;
    c(2)=3*u'*LX1/2;

```

```

    for k=3 :p+1
% Calcul des valeurs de  $L_{k-1}$  aux points d'intégration  $nL_n(\text{absc})=(2n-1).\text{absc}.L_{n-1}(\text{absc})-(n-1)L_{n-2}(\text{absc})$ 
        LX2=((2*k-3)*absc.*LX1-(k-2)*LX0)/(k-1);
% Calcul de  $\hat{f}_{k-1}$  dans  $c(k)$ 
        c(k)=(2*k-1)*u'*LX2/2;
% On vérifie que les polynômes sont orthogonaux entre eux
        fprintf('(LX2.*LX1)*poids = %f\n', (LX2.*LX1)*poids);
        fprintf('(LX2.*LX0)*poids = %f\n', (LX2.*LX0)*poids);
        LX0=LX1;
        LX1=LX2;
    end
% Evaluation de la série en npt points de l'intervalle [-1,1]
    x=linspace(-1,1,npt);
    y=CombLinLeg(x,c);
% Erreur entre la série de Legendre et la fonction
    err=norm(Test(x)-y,inf)

```

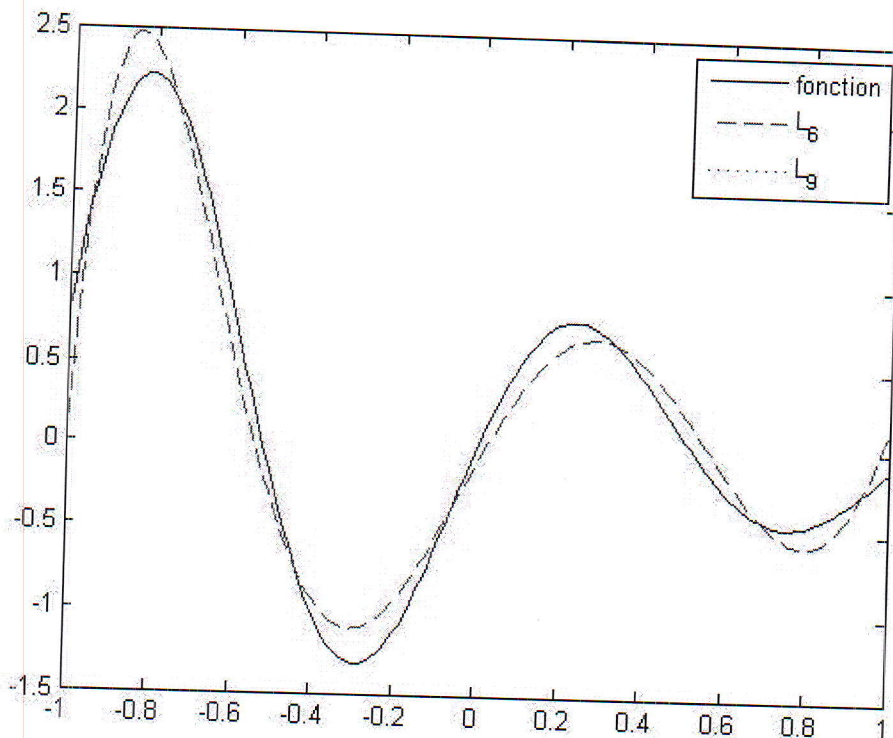


Figure 2.1 Approximation de la fonction  $f(x)$  avec  $\mathcal{L}_6$  et  $\mathcal{L}_9$

**Remarque 2.4.2.** Le choix du nombre de points d'intégration  $s$  dépend du degré de la série qu'on veut calculer. En effet, on sait que la formule de quadrature (1.22) sur  $s$  points est exacte pour les polynômes de  $\mathbb{k}_{2s-1}$ . Par ailleurs, la décomposition en série de Legendre tronquée à l'ordre  $p$  d'un polynôme de degré  $p$  est en fait une décomposition dans la base des  $(L_i)_{0 \leq i \leq p}$ . On aimerait donc que le calcul des coefficients  $\hat{f}_k$  soit exact si  $f$  est un polynôme de degré  $p$ . Dans ce cas le terme de plus haut degré,  $\hat{f}_p$  consiste à intégrer un polynôme de degré  $2p$ , il faut donc que  $s$  soit au moins égal à  $p+1$  pour assurer l'exactitude de la formule de quadrature. On peut vérifier numériquement que si la fonction est de classe  $C^\infty$ , l'erreur en norme  $L^\infty$  décroît plus vite que n'importe quel polynôme en fonction de l'inverse du degré de la série.

**Remarque 2.4.3.** Quand on fait tourner le script pour une fonction moins régulière, c'est à dire qui n'est pas de classe  $C^\infty$ , les résultats se détériorent : pour la fonction  $|x|$ , l'approximation converge lentement quand on augmente le degré de la série et il reste toujours une

erreur visible à l'oeil nu près de la singularité. La figure 2.2 montre la dépendance linéaire de l'erreur en fonction de  $1/p$ .

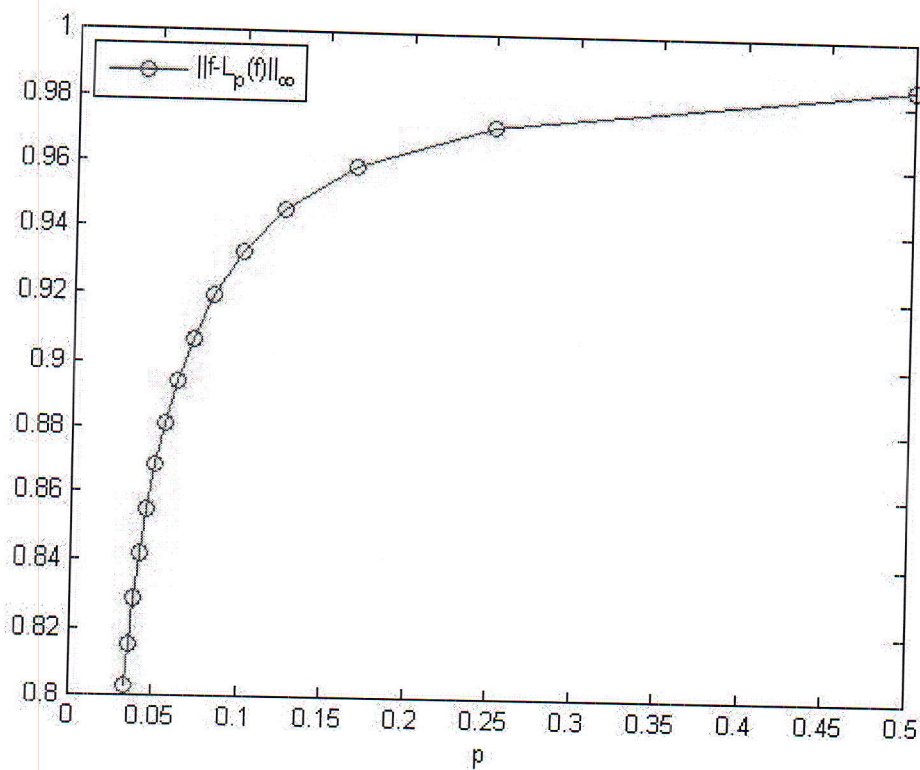
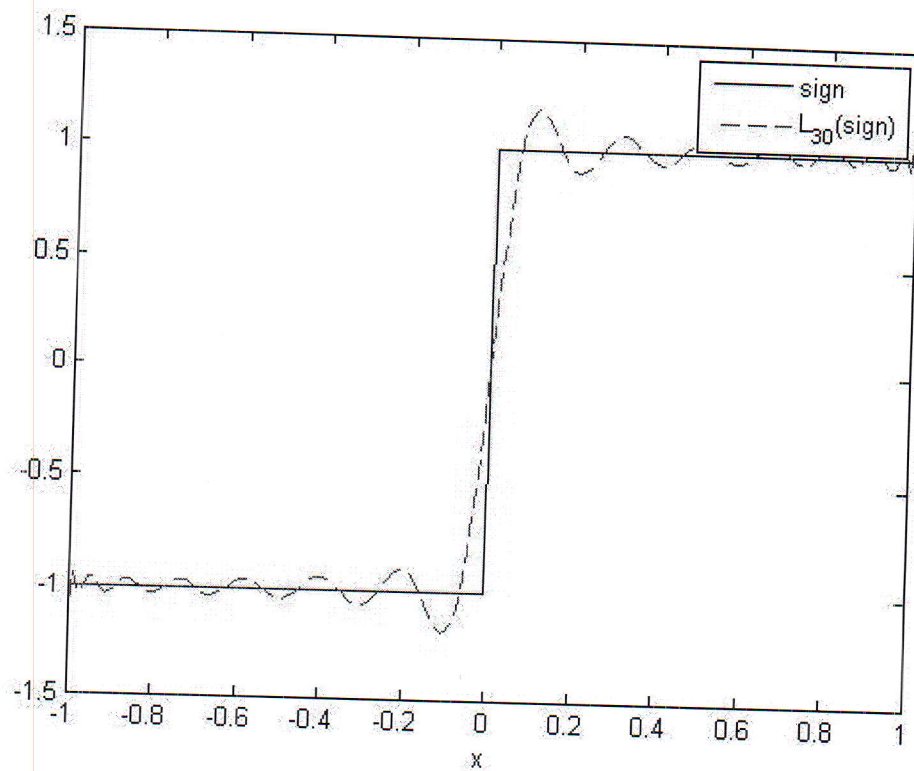


Figure 2.2 Erreur en norme  $L^\infty$  entre  $f(x)=|x|$  et sa série tronquée à l'ordre  $p$

**Remarque 2.4.4.** Pour la fonction discontinue  $\text{sign}$ , la série ne converge pas en norme  $L^\infty$ . Elle présente des oscillations d'amplitude bornée mais dans la fréquence avec le degré, c'est le phénomène de Gibbs. La figure 2.3 et la figure 2.4 représentent les fonctions  $|x|$  et  $\text{sign}(x)$  et leur série tronquées à l'ordre 30.



*Figure 2.3* Comparaison de la fonction  $\text{sign}(x)$  avec sa série tronquée à l'ordre 30

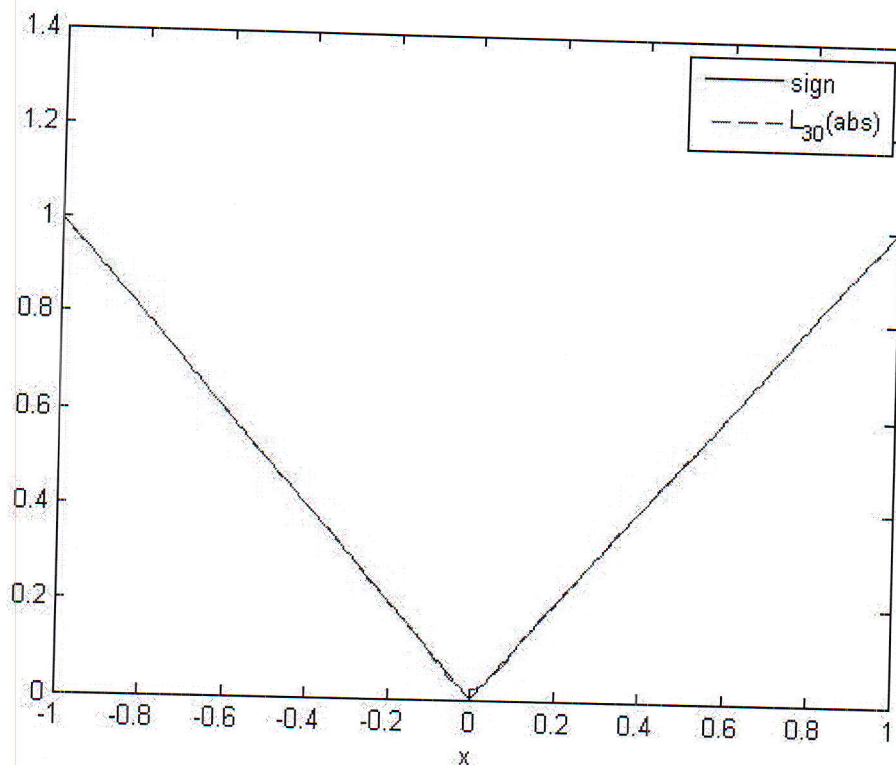


Figure 2.4 Comparaison de la fonction  $|x|$  avec sa série tronquée à l'ordre 30

## 2.5 Résolution par la méthode spectrale du problème (2.3)

On considère le sous-espace vectoriel de  $\mathbb{k}_m$

$$\mathbb{k}_m^0(\Omega) = \{p \in \mathbb{k}_m, p(-1) = p(1) = 0\}.$$

Cet espace est de dimension  $m - 1$  et s'écrit aussi sous la forme

$$\mathbb{k}_m^0(\Omega) = \{p = (1 - x^2)q, q \in \mathbb{k}_{m-2}\}.$$

Il est clair que  $\mathbb{k}_m^0(\Omega) \subset H_0^1(\Omega)$  et l'on peut donc sans difficulté définir une méthode d'approximation variationnelle, dite méthode spectrale variationnelle de Galerkin.

$$\left\{ \begin{array}{l} \text{Trouver } u_m \in \mathbb{k}_m^0(\Omega) \text{ tel que, pour tout } v_m \in \mathbb{k}_m^0(\Omega), \text{ on ait} \\ \int_{-1}^1 u_m'(x)v_m'(x)dx + c \int_{-1}^1 u_m(x)v_m(x)dx = \int_{-1}^1 f(x)v_m(x)dx. \end{array} \right. \quad (2.5)$$

On choisit comme base de  $\mathbb{K}_m^0(\Omega)$  la famille de polynômes  $\mathcal{F}_m = (F_i)_{1 \leq i \leq m-1}$

$$F_i = (1 - x^2)L'_i$$

où les  $L_i$  sont les polynômes de Legendre étudiés précédemment. La forme matricielle de la formulation variationnelle s'obtient en écrivant la fonction inconnue  $u_m$  dans cette base

$$u_m = \sum_{i=1}^{m-1} u_{m,i} F_i \quad (2.6)$$

et en faisant varier  $v_m$  dans  $\mathcal{F}_m$  dans la formulation variationnelle (2.5).

### 2.5.1 Construction du système linéaire

En injection la combinaison linéaire (2.6) dans (2.5) et en choisissant comme fonction test successivement  $F_i$ , pour  $i = 1, \dots, m-1$  on obtient un système linéaire de  $m-1$  équations dont la solution est le vecteur  $\bar{u}_m = (u_{m,i})_{1 \leq i \leq m-1}$  des coefficients de  $u_m$  dans la base  $\mathcal{F}_m$

$$A_m \bar{u}_m = b_m,$$

où

$$(A_m)_{i,j} = \frac{[i(i+1)]^2}{i+1/2} \delta_{i,j} + c \frac{i(i+1)j(j+1)}{(2i+1)(2j+1)} \left( \frac{4(2i+1)\delta_{i,j}}{(2i-1)(2i+3)} - \frac{2\delta_{i,j-2}}{2i+3} - \frac{2\delta_{i,j+2}}{2i-1} \right),$$

$$(b_m)_i = \int_{-1}^1 f(x) F_i(x) dx = \frac{2i(i+1)}{2i+1} \left( \frac{1}{2i-1} \hat{f}_{i-1} - \frac{1}{2i+3} \hat{f}_{i+1} \right).$$

Une fois résolu le système linéaire, on connaît les coefficients spectraux de la solution  $\bar{u}_m$  dans la base  $\mathcal{F}_m$ . On obtient facilement l'expression du  $\bar{u}_m$  dans la base  $(L_j)_{j=0}^m$ , en utilisant les relations

$$(1-x^2)L'_i = \frac{i(i+1)}{2i+1} (L_{i-1} - L_{i+1}). \quad (2.7)$$

### 2.5.2 Etude la convergence de la méthode

Nous allons montrer quelques résultats de majoration de la distance de fonctions à régularité donnée à un espace de polynômes, toujours dans le cas des polynômes de Legendre. Cette distance sera calculée au moyen d'opérateurs de projection orthogonale.

On fixe  $N$  un entier positif, on note  $\mathbb{K}_N(\Omega)$  l'espace des polynômes de degrés inférieurs ou égaux à  $N$ , et on note  $\pi_N$  l'opérateur de projection orthogonale de  $L^2(\Omega)$  sur  $\mathbb{K}_N(\Omega)$ .



**Théorème 2.5.1.** *Pour tout entier  $m \geq 0$ , il existe une constante  $c$  positive ne dépendant que de  $m$  telle que, pour toute fonction  $\varphi$  de  $H^m(\Omega)$ , on ait*

$$\|\varphi - \pi_N \varphi\|_{L^2(\Omega)} \leq cN^{-m} \|\varphi\|_{H^m(\Omega)}. \quad (2.8)$$

On commence par prouver un résultat de continuité sur l'opérateur auto-adjoint  $A$  défini par

$$A\varphi = -\frac{d}{dx} \left( (1-x^2)\varphi'(x) \right)$$

**Lemme 2.5.1.** *Pour tout entier  $l \geq 0$ , l'opérateur  $A$  est continu de  $H^{l+2}(\Omega)$  dans  $H^l(\Omega)$ . Pour tous entiers  $l \geq 0$  et  $m \geq 0$ , l'opérateur  $A^m$  est continu de  $H^{l+2m}(\Omega)$  dans  $H^l(\Omega)$ .*

*Démonstration.* On vérifie grâce à Leibniz que, pour tout entier  $k \geq 0$ ,

$$\frac{d^k(A\varphi)}{dx^k} = -(1-x^2)\frac{d^{k+2}\varphi}{dx^{k+2}} + 2(k+1)x\frac{d^{k+1}\varphi}{dx^{k+1}} + k(k+1)\frac{d^k\varphi}{dx^k}$$

En appliquant cette formule, on voit que, pour tout  $k$ ,  $0 \leq k \leq l$ ,

$$\left\| \frac{d^k(A\varphi)}{dx^k} \right\|_{L^2(\Omega)} \leq c \left( \left\| \frac{d^{k+2}\varphi}{dx^{k+2}} \right\|_{L^2(\Omega)} + \left\| \frac{d^{k+1}\varphi}{dx^{k+1}} \right\|_{L^2(\Omega)} + \left\| \frac{d^k\varphi}{dx^k} \right\|_{L^2(\Omega)} \right)$$

d'où la première affirmation du lemme. On en déduit la seconde en itérant  $m$  fois le résultat.  $\square$

**Démonstration du théorème.** Etant donné une fonction  $\varphi$  de  $H^m(\Omega)$ , que l'on décompose selon la base des polynômes de Legendre, il faut estimer

$$\|\varphi - \pi_N \varphi\|_{L^2(\Omega)}^2 = \sum_{n=N+1}^{\infty} (\varphi^n)^2 \|L_n\|_{L^2(\Omega)}^2$$

On va distinguer deux cas, suivant que  $m$  est pair ou impair :

1) Lorsque  $m$  est pair égal à  $2k$ , d'après l'équation différentielle (1.4) vérifiée par les polynômes  $L_n$ ,  $n \geq 0$ , on a

$$\varphi^n = \frac{1}{\|L_n\|_{L^2(\Omega)}^2} \int_{-1}^1 \varphi(x)L_n(x)dx = \frac{1}{\|L_n\|_{L^2(\Omega)}^2} \frac{1}{n(n+1)} \int_{-1}^1 \varphi(x)(AL_n)(x)dx$$

Comme l'opérateur  $A$  est auto-adjoint dans  $L^2(\Omega)$ , on obtient

$$\varphi^n = \frac{1}{\|L_n\|_{L^2(\Omega)}^2} \frac{1}{n(n+1)} \int_{-1}^1 (A\varphi)(x)L_n(x)dx$$

En itérant  $k$  fois ce résultat, on en déduit

$$\varphi^n = \frac{1}{\|L_n\|_{L^2(\Omega)}^2} \frac{1}{(n(n+1))^k} \int_{-1}^1 (A^k \varphi)(x) L_n(x) dx$$

On constate donc que

$$\|\varphi - \pi_N \varphi\|_{L^2(\Omega)}^2 = \sum_{n=N+1}^{\infty} \frac{1}{(n(n+1))^{2k}} \left( \frac{\int_{-1}^1 (A^k \varphi)(x) L_n(x) dx}{\|L_n\|_{L^2(\Omega)}^2} \right)^2 \|L_n\|_{L^2(\Omega)}^2$$

On minore alors les  $n(n+1)$  par  $N^2$ , ce qui donne

$$\|\varphi - \pi_N \varphi\|_{L^2(\Omega)}^2 \leq N^{-4k} \sum_{n=N+1}^{\infty} \left( \frac{\int_{-1}^1 (A^k \varphi)(x) L_n(x) dx}{\|L_n\|_{L^2(\Omega)}^2} \right)^2 \|L_n\|_{L^2(\Omega)}^2$$

Comme les  $\frac{\int_{-1}^1 (A^k \varphi)(x) L_n(x) dx}{\|L_n\|_{L^2(\Omega)}^2}$  sont les coefficients de  $A^k \varphi$  dans la base des polynômes de Legendre, on a

$$\|\varphi - \pi_N \varphi\|_{L^2(\Omega)}^2 \leq N^{-4k} \|A^k \varphi\|_{L^2(\Omega)}^2.$$

En utilisant le Lemme (2.5.1), on conclut

$$\|\varphi - \pi_N \varphi\|_{L^2(\Omega)}^2 \leq cN^{-2m} \|\varphi\|_{H^m(\Omega)}^2.$$

2) Lorsque  $m$  est impair égal à  $2k+1$ , on obtient comme précédemment

$$\varphi^n = \frac{1}{\|L_n\|_{L^2(\Omega)}^2} \frac{1}{(n(n+1))^k} \int_{-1}^1 (A^k \varphi)(x) L_n(x) dx$$

puis on utilise une fois de plus l'équation différentielle (1.4) et on intègre par parties une seule fois. On en déduit

$$\varphi^n = \frac{1}{\|L_n\|_{L^2(\Omega)}^2} \frac{1}{(n(n+1))^{k+1}} \int_{-1}^1 (A^k \varphi)'(x) L_n'(x) (1-x^2) dx$$

On voit alors que

$$\|\varphi - \pi_N \varphi\|_{L^2(\Omega)}^2 \leq \sum_{n=N+1}^{\infty} \frac{1}{(n(n+1))^{2(k+1)}} \frac{\left( \int_{-1}^1 (A^k \varphi)'(x) L_n'(x) (1-x^2) dx \right)^2}{\|L_n\|_{L^2(\Omega)}^2} \|L_n\|_{L^2(\Omega)}^2$$

On note que les polynômes  $L_n'$ ,  $n \geq 1$ , sont deux à deux orthogonaux pour la mesure  $(1-x^2)dx$ , toute fonction  $\psi$  de  $H^1(\Omega)$  admet le développement

$$\psi = \sum_{n=0}^{\infty} \psi^n L_n, \text{ avec } \psi^n = \frac{\int_{-1}^1 \psi'(x) L_n'(x) (1-x^2) dx}{\int_{-1}^1 L_n'^2(x) (1-x^2) dx} \quad \text{pour } n \geq 1$$

de la formule (1.6), on déduit alors

$$\int_{-1}^1 \psi'^2(x)(1-x^2)dx = \sum_{n=0}^{\infty} \frac{1}{n(n+1)} \frac{\left(\int_{-1}^1 \psi'(x)L'_n(x)(1-x^2)dx\right)^2}{\|L_n\|_{L^2(\Omega)}^2}$$

En appliquant cette formule pour la fonction  $\psi = A^k\varphi$  et en minorant  $(n(n+1))^{2k+1}$  par  $N^{2(2k+1)}$ , on voit que

$$\|\varphi - \pi_N\varphi\|_{L^2(\Omega)}^2 \leq N^{-2(2k+1)} \int_{-1}^1 (A^k\varphi)'^2(1-x^2)dx$$

Et on conclut

$$\|\varphi - \pi_N\varphi\|_{L^2(\Omega)}^2 \leq cN^{-2m} \left\| (A^k\varphi)' \right\|_{L^2(\Omega)}^2 \leq cN^{-2m} \|A^k\varphi\|_{H^1(\Omega)}^2,$$

d'où, d'après le Lemme (2.5.1),

$$\|\varphi - \pi_N\varphi\|_{L^2(\Omega)}^2 \leq cN^{-2m} \|\varphi\|_{H^m(\Omega)}^2.$$

## Chapitre 3

# Etude de comparaison (Méthode spectrale et différence finie)

### 3.1 Méthode de différences finies : Principe de la méthode

On considère le problème unidimensionnel

$$-u''(x) = f(x), \quad \forall x \in [-1, 1] \quad (3.1)$$

$$u(-1) = u(1) = 0, \quad (3.2)$$

où  $f \in C([-1, 1])$ . Les conditions aux limites (3.2) considérées ici sont dites de type Dirichlet homogène (le terme homogène désigne les conditions nulles). Cette équation modélise par exemple la diffusion de la chaleur dans un barreau conducteur chauffé (terme source  $f$ ) dont les deux extrémités sont plongées dans de la glace.

Soit  $(x_k)_{k=0, \dots, N+1}$  une subdivision de  $[-1, 1]$ , avec

$$x_0 = -1 < x_1 < x_2 < \dots < x_N < x_{N+1} = 1.$$

Pour  $i = 0, \dots, N$ , on note  $h_{i+1/2} = x_{i+1} - x_i$  et on définit le "pas" du maillage par

$$h = \max_{i=0, \dots, N} h_{i+1/2}. \quad (3.3)$$

Pour simplifier l'exposé, on se limitera dans un premier temps à un pas constant

$$h_{i+1/2} = h \quad \forall i \in [0, N].$$

On écrit l'équation aux dérivées partielles (3.1) aux points  $x_i$

$$-u''(x_i) = f(x_i), \forall i = 1, \dots, N,$$

Effectuons un développement de Taylor en  $x_i$

$$\begin{aligned} u(x_{i+1}) &= u(x_i) + hu'(x_i) + \frac{h^2}{2}u''(x_i) + \frac{h^3}{6}u'''(x_i) + \frac{h^4}{24}u^{(4)}(\zeta_i), \\ u(x_{i-1}) &= u(x_i) - hu'(x_i) + \frac{h^2}{2}u''(x_i) - \frac{h^3}{6}u'''(x_i) + \frac{h^4}{24}u^{(4)}(\eta_i), \end{aligned}$$

avec  $\zeta_i \in [x_i, x_{i+1}]$ ,  $\eta_i \in [x_{i-1}, x_i]$ . En additionnant, on obtient

$$u(x_{i+1}) + u(x_{i-1}) = 2u(x_i) + h^2u''(x_i) + O(h^2)$$

Il semble donc raisonnable d'approcher la dérivée seconde  $-u''(x_i)$  par le "quotient différentiel"

$$\frac{2u(x_i) - u(x_{i-1}) - u(x_{i+1}))}{h^2}.$$

Sous des hypothèses de régularité sur  $u$ , on peut montrer que cette approximation est d'ordre 2 au sens

$$R_i = u''(x_i) + \frac{2u(x_i) - u(x_{i-1}) - u(x_{i+1}))}{h^2} = O(h^2)$$

On appelle erreur de consistance au point  $x_i$  la quantité  $R_i$ .

### 3.1.1 Questions d'analyse numérique

Voici un certain nombre de questions, qui sont typiquement du domaine de l'analyse numérique, auxquelles nous tenterons de répondre dans la suite :

1. Le problème qu'on a obtenu en dimension finie, (avec des inconnues localisées aux noeuds du maillage dans le cas de la méthode des différences finies) admet-il une (unique) solution ? On montrera que oui.
2. La solution du problème discret converge-t-elle vers la solution du problème continu lorsque le pas du maillage  $h$  tend vers 0 ? Dans le cas des différences finies en une dimension d'espace, le pas du maillage est défini par

$$h = \sup_{i=1 \dots N} |x_{i+1} - x_i|.$$

### 3.2 Etude de la méthode différences finies pour un problème elliptique unidimensionnel

On cherche à discrétiser le problème aux limites, suivant

$$\begin{cases} -u''(x) + c(x)u(x) = f(x), & -1 < x < 1, \\ u(-1) = u(1) = 0, \end{cases} \quad (3.4)$$

où  $c \in C([-1, 1], \mathbb{R}_+)$ , et  $f \in C([-1, 1], \mathbb{R})$ , qui peut modéliser par exemple un phénomène de diffusion-réaction d'une espèce chimique. On se donne un pas du maillage constant  $h = \frac{1}{N+1}$  et une subdivision de  $] -1, 1[$ , notée  $(x_k)_{k=0, \dots, N+1}$  avec :  $x_0 = -1 < x_1 < x_2 < \dots < x_N < x_{N+1} = 1$ . Soit  $u_i$  l'inconnue discrète associée au noeud  $i$  ( $i = 1, \dots, N$ ). On pose  $u_0 = u_{N+1} = 0$ . On obtient les équations discrètes en approchant  $u''(x_i)$  par quotient différentiel par développement de Taylor,

$$\begin{cases} \frac{1}{h^2} (2u_i - u_{i-1} - u_{i+1}) + c_i u_i = f_i & i = 1, \dots, N \\ u_0 = u_{N+1} = 0. \end{cases} \quad (3.5)$$

avec  $c_i = c(x_i)$  et  $f_i = f(x_i)$ . On peut écrire ces équations sous forme matricielle

$$A_h U_h = b_h, \text{ avec } U_h = \begin{pmatrix} u_1 \\ \vdots \\ u_N \end{pmatrix} \text{ et } b_h = \begin{pmatrix} f_1 \\ \vdots \\ f_N \end{pmatrix} \quad (3.6)$$

$$\text{et } A_h = \frac{1}{h^2} \begin{pmatrix} 2 + c_1 h^2 & -1 & 0 & \dots & 0 \\ -1 & 2 + c_2 h^2 & -1 & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & -1 & 2 + c_{N-1} h^2 & -1 \\ 0 & \dots & 0 & -1 & 2 + c_N h^2 \end{pmatrix} \quad (3.7)$$

Les questions suivantes surgissent alors naturellement :

1. Le système (3.6) admet-il un unique solution ?
2. A-t-on convergence de  $U_h$  vers  $u$  ?

Nous allons répondre par l'affirmative à ces deux questions. Commençons par la première.

**Proposition 3.2.1.** *Soit  $c = (c_1, \dots, c_N)^t \in \mathbb{R}^N$  tel que  $c_i \geq 0$  pour  $i = 1, \dots, N$ ; alors la matrice  $A_h$  définie par (3.7) est symétrique définie positive, et donc inversible.*

*Démonstration.* La matrice  $A_h$  est évidemment symétrique. Montrons qu'elle est définie positive. Soit  $v = (v_1 \dots v_N)^t$ , on pose  $v_0 = v_{N+1} = 0$ . Calculons le produit scalaire  $A_h v \cdot v = v^t A_h v$ . On a

$$A_h v \cdot v = \frac{1}{h^2} (v_1 \dots v_N) \begin{pmatrix} 2 + c_1 h^2 & -1 & & 0 \\ -1 & \ddots & \ddots & \\ & \ddots & & -1 \\ 0 & & -1 & 2 + c_N h^2 \end{pmatrix} \begin{pmatrix} v_1 \\ \vdots \\ \vdots \\ v_N \end{pmatrix},$$

c'est-à-dire

$$A_h v \cdot v = \frac{1}{h^2} \sum_{i=1}^N v_i (-v_{i-1} + (2 + c_i h^2) v_i - v_{i+1}).$$

On a donc, par changement d'indice

$$A_h v \cdot v = \frac{1}{h^2} \left[ \sum_{i=1}^N (-v_{i-1} v_i) + \sum_{i=1}^N (2 + c_i h^2) v_i^2 - \sum_{j=2}^{N+1} v_{j-1} v_j \right].$$

Et comme on a posé  $v_0 = 0$  et  $v_{N+1} = 0$ , on peut écrire

$$A_h v \cdot v = \frac{1}{h^2} \sum_{i=1}^N (2 + c_i h^2) v_i^2 + \frac{1}{h^2} \sum_{i=1}^N (-2v_i v_{i-1})$$

soit encore

$$A_h v \cdot v = \sum_{i=1}^N c_i v_i^2 + \frac{1}{h^2} \sum_{i=1}^N (-2v_i v_{i-1} + v_i^2 + v_{i-1}^2) + v_N^2.$$

On a donc finalement

$$A_h v \cdot v = \sum_{i=1}^N c_i v_i^2 + \frac{1}{h^2} \sum_{i=1}^N (v_i - v_{i-1})^2 + v_N^2 \geq 0, \quad \forall v = (v_1, \dots, v_N) \in \mathbb{R}^N.$$

Si on suppose  $A_h v \cdot v = 0$ , on a alors

$$\sum_{i=1}^N c_i h^2 v_i^2 = 0 \text{ et } v_i - v_{i-1} = 0, \quad \forall i = 1, \dots, N.$$

On a donc  $v_1 = v_2 = \dots = v_N = v_0 = v_{N+1} = 0$ . Remarquons que ces égalités sont vérifiées même si les  $c_i$  sont nuls. Ceci démontre que la matrice  $A_h$  est bien définie.  $\square$

### 3.3 Convergence de la méthode des différences finies

Nous allons maintenant nous préoccuper de la question de la convergence.

**Définition 3.3.1. (Matrices monotones)** Soit  $A \in M_N(\mathbb{R})$ , de coefficients  $a_{i,j}$ ,  $i = 1, \dots, N$  et  $j = 1, \dots, N$ . On dit que  $A$  est positive (ou  $A \geq 0$ ) si  $a_{i,j} \geq 0$ ,  $\forall i, j = 1, \dots, N$ . On dit que  $A$  est monotone si  $A$  est inversible et  $A^{-1} \geq 0$ .

L'avantage des schémas à matrices monotones est de satisfaire la propriété de conservation de la positivité, qui peut être cruciale dans les applications physiques.

**Définition 3.3.2. (Conservation de la positivité)** Soit  $A \in M_N(\mathbb{R})$  de coefficients  $a_{i,j}$ ,  $i = 1, \dots, N$  et  $j = 1, \dots, N$ ; on dit que  $A$  conserve la positivité si  $Av \geq 0$  entraîne  $v \geq 0$  (les inégalités s'entendent composante par composante).

On a en effet la proposition suivante :

**Proposition 3.3.1. (Monotonie et positivité)** Soit  $A \in M_N(\mathbb{R})$ . Alors  $A$  conserve la positivité si et seulement si  $A$  est monotone.

*Démonstration.* Supposons d'abord que  $A$  conserve la positivité, et montrons que  $A$  inversible et que  $A^{-1}$  a des coefficients  $\geq 0$ . Si  $x$  est tel que  $Ax = 0$ , alors  $Ax \geq 0$  et donc, par hypothèse,  $x \geq 0$ . Mais on a aussi  $Ax \leq 0$ , soit  $A(-x) \geq 0$  et donc par hypothèse,  $x \leq 0$ . On en déduit  $x = 0$ , ce qui prouve que  $A$  est inversible. La conservation de la positivité donne alors que  $y \geq 0 \Rightarrow A^{-1}y \geq 0$ . En prenant  $y = e_1$  on obtient que la première colonne de  $A^{-1}$  est positive, puis en prenant  $y = e_i$  on obtient que la  $i$ -ème colonne de  $A^{-1}$  est positive, pour  $i = 2, \dots, N$ . Donc  $A^{-1}$  a tous ses coefficients positifs.

Réciproquement, supposons maintenant que  $A$  est inversible et que  $A^{-1}$  a des coefficients positifs. Soit  $x \in \mathbb{R}^N$  tel que  $Ax = y \geq 0$ , alors  $x = A^{-1}y \geq 0$ . Donc  $A$  conserve la positivité.  $\square$

**Remarque 3.3.1. (Principe du maximum)** On appelle principe du maximum continu le fait que si  $f \geq 0$  alors le minimum de la fonction  $u$  solution du problème (3.4) est atteint sur les bords. Cette propriété mathématique correspond à l'intuition physique qu'on peut avoir du phénomène : si on chauffe un barreau tout en maintenant ses deux extrémités à une température fixe, la température aux points intérieurs du barreau sera supérieure à celle des extrémités. Il est donc souhaitable que la solution approchée satisfasse la même propriété.



**Lemme 3.3.1.** Soit  $c = (c_1, \dots, c_N)^t \in \mathbb{R}^N$ , et  $A_h \in M_N(\mathbb{R})$  définie par (3.7). Si  $c_i \geq 0$  pour tout  $i = 1, \dots, N$ , alors  $A_h$  est monotone.

*Démonstration.* On va montrer que si  $v \in \mathbb{R}^N$   $A_h v \geq 0$  alors  $v \geq 0$ . On peut alors utiliser la proposition (3.3.1) pour conclure. Soit  $v = (v_1, \dots, v_N)^t \in \mathbb{R}^N$ . Posons  $v_0 = v_{N+1} = 0$ . Supposons que  $A_h v \geq 0$ . On a donc

$$-\frac{1}{h^2}v_{i-1} + \left(\frac{2}{h^2} + c_i\right)v_i - \frac{1}{h^2}v_{i+1} \geq 0, \quad i = 1, \dots, N \quad (3.8)$$

Soit

$$p = \min\{i \in \{1, \dots, N\}; v_p = \min_{j=1, \dots, N} v_j\}.$$

Supposons que  $\min_{j=1, \dots, N} v_j < 0$ . On a alors  $p \geq 1$  et

$$\frac{1}{h^2}(v_p - v_{p-1}) + c_p v_p + \frac{1}{h^2}(v_p - v_{p+1}) \geq 0.$$

On en déduit que

$$\frac{2}{h^2}c_p v_p \geq \frac{1}{h^2}(v_{p-1} - v_p) + \frac{1}{h^2}(v_{p+1} - v_p) \geq 0$$

Si  $c_p > 0$ , on a donc  $v_p \geq 0$ , et donc  $v_i \geq 0, \forall i = 1, \dots, N$ . Si  $c_p = 0$ , on doit alors avoir  $v_{p-1} = v_p = v_{p+1}$  ce qui est impossible car  $p$  est le plus petit indice  $j$  tel que  $v_j = \min_{i=1, \dots, N} v_i$ . Donc dans ce cas le minimum ne peut pas être atteint pour  $j = p > 1$ . On a ainsi finalement montré que  $\min_{i \in \{1, \dots, N\}} v_i \geq 0$ , on a donc  $v \geq 0$ .  $\square$

**Définition 3.3.3. (Erreur de consistance)** On appelle erreur de consistance la quantité obtenue en remplaçant l'inconnue par la solution exacte dans le schéma numérique. Dans le cas du schéma (3.5), l'erreur de consistance au point  $x_i$  est donc définie par

$$R_i = \frac{1}{h^2}(2u(x_i) - u(x_{i-1}) - u(x_{i+1})) + c(x_i)u(x_i) - f(x_i). \quad (3.9)$$

L'erreur de consistance  $R_i$  est donc l'erreur qu'on commet en remplaçant l'opérateur  $-u''$  par le quotient différentiel

$$\frac{1}{h^2}(2u(x_i) - u(x_{i-1}) - u(x_{i+1})).$$

Cette erreur peut être évaluée si  $u$  est suffisamment régulière, en effectuant des développements de Taylor.

**Définition 3.3.4. (Ordre du schéma)** On dit qu'un schéma de discrétisation à  $N$  points est d'ordre  $p$  s'il existe  $C \in \mathbb{R}$ , ne dépendant que de la solution exacte, tel que l'erreur de consistance satisfasse

$$\max_{i=1,\dots,N} (R_i) < Ch^p,$$

où  $h$  est le pas du maillage défini par (3.3) (c.à.d. le maximum des écarts  $x_{i+1} - x_i$ ). On dit qu'un schéma de discrétisation est consistant si

$$\max_{i=1,\dots,N} (R_i) \longrightarrow 0 \text{ lorsque } h \longrightarrow 0,$$

où  $N$  est le nombre de points de discrétisation.

**Lemme 3.3.2.** Si la solution de (3.4) vérifie  $u \in C^4([-1, 1])$ , alors le schéma (3.5) est consistant d'ordre 2, et on a plus précisément

$$|R_i| \leq \frac{h^2}{12} \sup_{[-1,1]} |u^{(4)}|, \quad \forall i = 1, \dots, N. \quad (3.10)$$

*Démonstration.* (voir : [7] ) □

**Remarque 3.3.2. (Sur l'erreur de consistance)**

1. Si on note  $\bar{U}_h = (u(x_i))_{i=1\dots N}$  le vecteur dont les composantes sont les valeurs exactes de la solution de (3.4), et  $U_h = (u_1 \dots u_N)^t$  la solution de (3.5), on a

$$R_i = A_h (U_h - \bar{U}_h) \quad (3.11)$$

2. On peut remarquer que si  $u^{(4)} = 0$  les développements de Taylor effectués ci-dessus se résument à :

$$-u''(x_i) = \frac{2u(x_i) - u(x_{i-1}) - u(x_{i+1}))}{h^2},$$

et on a donc  $R_i = 0$ , pour tout  $i = 1, \dots, N$ , et donc  $u_i = u(x_i)$ , pour tout  $i = 1 \dots N$ . Dans ce cas, le schéma de discrétisation donne la valeur exacte de la solution en  $x_i$ , pour tout  $i = 1, \dots, N$ . Cette remarque est bien utile lors de la phase de validation de méthodes numériques et/ou programmes informatiques pour la résolution de l'équation (3.4). En effet, si on choisit  $f$  telle que la solution soit un polynôme de degré inférieur ou égal à 3, alors on doit avoir une erreur entre solution exacte et approchée inférieure à l'erreur machine.

La preuve de convergence du schéma utilise la notion de consistance, ainsi qu'une notion de stabilité, que nous introduisons maintenant :

**Proposition 3.3.2.** *On dit que le schéma (3.5) est stable, au sens où la matrice de discrétisation  $A_h$  satisfait*

$$\|A_h^{-1}\|_{\infty} \leq \frac{1}{8}. \quad (3.12)$$

*Démonstration.* (voir : [7]) □

**Définition 3.3.5. (Erreur de discrétisation)** *On appelle erreur de discrétisation en  $x_i$ , la différence entre la solution exacte en  $x_i$  et la  $i$ -ème composante de la solution donnée par le schéma numérique*

$$e_i = u(x_i) - u_i, \quad \forall 1, \dots, N. \quad (3.15)$$

**Théorème 3.3.1.** *Soit  $u$  la solution exacte de*

$$\begin{cases} -u'' + cu = f, \\ u(-1) = u(1) = 0. \end{cases}$$

*On suppose  $u \in C^4([-1, 1])$ . Soit  $u_h$  la solution de (3.5). Alors l'erreur de discrétisation définie par (3.15) satisfait*

$$\max_{i=1, \dots, N} |e_i| \leq \frac{h^2}{96} \|u^{(4)}\|_{\infty}.$$

*Le schéma est donc convergente d'ordre 2.*

*Démonstration.* Soit  $U_h = (U_1, \dots, U_n)^t$  et  $\bar{U} = (u(x_1), \dots, u(x_N))^t$ , on cherche à majorer  $\|\bar{U}_h - U_h\|_{\infty}$ . On a  $A(\bar{U}_h - U_h) = R$  où  $R$  est l'erreur de consistance (voir remarque 3.3.2). On a donc

$$\|\bar{U}_h - U_h\|_{\infty} \leq \|A_h^{-1}\|_{\infty} \|R\|_{\infty} \leq \frac{1}{8} \times \frac{h^2}{12} \|u^{(4)}\|_{\infty}$$

□

**Remarque 3.3.3. (Sur la convergence)** *On peut remarquer que la preuve de la convergence s'appuie sur la stabilité (elle-même déduite de la conservation de la positivité) et sur la consistance. Dans certains livres d'analyse numérique, vous trouverez la "formule" : stabilité + consistance  $\implies$  convergence. Il faut toutefois prendre garde au fait que ces notions de stabilité et convergence peuvent être variables d'un type de méthode à un autre.*

**Exemple 3.3.1.** *Pour tester la méthode spectrale, la programmation Matlab, exige les étapes suivantes :*

- Étape 1 :** Construction de la matrice  $A_m$
- Étape 2 :** Construction du second membre  $b_m$ . Cette étape comprend le calcul des coefficients  $\hat{f}_k$ .
- Étape 3 :** Résolution du système linéaire.
- Étape 4 :** Calcul des coefficients  $\hat{u}_{mk}$  de la solution numérique, en utilisant la relation (2.7)
- Étape 5 :** Calcul de l'erreur en norme infinie entre la solution exacte et la spectrale.
- Étape 6 :** Calcul de la solution différences finies.
- Étape 7 :** Représentation sur le même graphique de la fonction  $u$  et de la solution numérique : spectrale et différences finies.

**Remarque 3.3.4.** On choisit comme cas test la fonction  $u(x) = \sin(\pi x) \cos(10x)$  qui vérifie les conditions aux limites  $u(-1) = u(1) = 0$  et prenant comme second membre

$$f(x) = (\pi^2 + 130) \sin(\pi x) \cos(10x) + 20\pi \cos(\pi x) \sin(10x)$$

et la constante  $c = 30$ , la fonction  $u$  est solution du problème (2.1). Le programme ci-dessous permet d'enchaîner les différentes étapes du calcul et finalement de comparer la solution calculée par la méthode spectrale avec la solution par différence finies.

### 3.3.1 Programme de comparaison

```

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%Résolution de -u'' + cu = f par la méthode spectrale de Galerkin
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
m=16;           % degré de l'approximation par Legendre
s=m+1;         % degré de la quadrature de Gauss pour le second membre.
global c
c=30;
% Construction de la matrice
A=zeros(m,m);
for i=1 :m

```

```

    A(i,i)=(i*(i+1))^2*(1./(0.5+i)+ 4.*c/((2.*i+1.)*(2*i-1)*(2*i+3)));
end
for i=1 :m-2
    A(i,i+2)=-2*c*i*(i+1)*(i+2)*(i+3)/((2*i+1)*(2*i+3)*(2*i+5));
end
for i=3 :m
    A(i,i-2)=-2*c*i*(i+1)*(i-2)*(i-1)/((2*i-1)*(2*i-3)*(2*i+1));
end
% Construction du second membre
[absc,poids]=xwGauss(s);
t=fbe(absc); u=t.*poids;
LX0=ones(s,1);
LX1=absc;
C=zeros(m+2,1);
C(1)=t'*poids/2; C(2)=3*u'*LX1/2;
for k=2 :m+1
    % calcul de  $f_k$  dans  $c(k+1)$ 
    % calcul des valeurs de  $L_k$  aux points d'intégration
    %  $kL_k = (2k - 1)xL_{k-1} - (k - 1)L_{k-2}$ 
    LX2=((2*k-1)*absc.*LX1-(k-1)*LX0)/k;
    C(k+1)=(2*k+1)*u'*LX2/2;
    LX0=LX1;
    LX1=LX2;
end
B=zeros(m,1);
for i=1 :m
    B(i)=2*i*(i+1)*(C(i)/(2*i-1)-C(i+2)/(2*i+3))/(2*i+1);
end
% Résolution du système linéaire
U=A\B;

```

```

%
% Changement de base
%(1 - x^2)L'_i = (i(i + 1)/(2i + 1)).(L_{i-1} - L_{i+1})
UN=zeros(1,m+2);
for k=1 :m
    CC=(k+1)*k*U(k)/(2*k+1);
    UN(k)=UN(k)+CC;
    UN(k+2)=UN(k+2)-CC;
end
%
% Calcul de la solution approchée et de l'erreur
%
n=100;
xa=linspace(-1,1,n);
y=CombLinLeg(xa,UN);
es=norm(y-speciale(xa),inf);
%
% Calcul de la solution différences finies
%
mdf=50; h=2/mdf;
xdf=linspace(-1+h,1-h,mdf-1)';
A=toeplitz([2,-1,zeros(1,mdf-3)])/h^2+c*eye(mdf-1,mdf-1);
B=fbe(xdf); ydf=A\B;
%
% Représentation graphique
%
plot (xa,speciale(xa),xa,y,'-',xdf,ydf,'x')
legend('exacte','spectrale','diff. finies')
fprintf('erreur spectrale= %e diff. finies= %e\n ',...es,norm(ydf-speciale(xdf),inf));

```

---

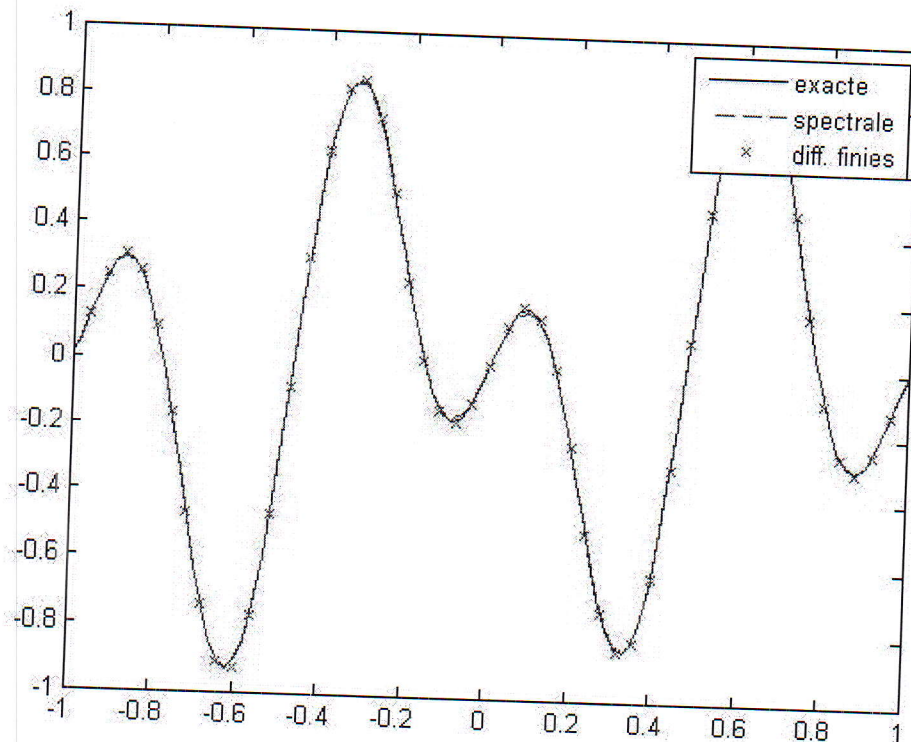


Figure 3.1 Comparaison de la solution exacte, spectrale et la solution différences finies avec 50 inconnues

### 3.4 Conclusion

1-Il est clair sur la figure 3.1 que la **méthode spectrale** est beaucoup plus précise que la **méthode des différences finies** puisque la solution approchée dans  $\mathbb{K}_{21}^0$  se confond avec la solution exacte.

2-L'erreur en norme sup est  $5 \cdot 10^{-5}$  alors qu'elle est égale à  $6 \cdot 10^{-2}$  pour la solution différences finies. Il faut faire le calcul aux différences finies sur environ 800 points pour arriver à une erreur en norme infinie aussi petite qu'avec la méthode spectrale.

3-En revanche, dès que la solution du problème continu n'est pas  $C^\infty$ , les performances de la méthode spectrale en terme de précision chutent et deviennent comparable-voire inférieure, à celle des différences finies.

4-Le lecteur pourra vérifier ce point en calculant le second membre  $f$  du problème (2.1) de manière à ce que la dérivée seconde de la solution exacte soit constante par morceaux.



# Bibliographie

- [1] Danaila, P. Joly, Kaber et M. Postel, *Introduction au calcul scientifique par la pratique*, Dunod, 2005.
- [2] Delabrière, M. Postel, *Méthodes d'approximation. Equations différentielles. Applications scilab*, Ellipses, Paris 2004
- [3] Bernardy, Y. Maday, F. Rapetti, *Discrétisations variationnelles de problèmes aux limites elliptiques*, Mathématiques et Application, vol 45, Springer-Verlage Mai 2004.
- [4] Frank Jedrzejewski, *Introduction aux méthodes numériques*, Springer-Verlage France, Paris 2005-2006.
- [5] Crouzeix, A.L. Mignot, *Analyse numérique des équations différentielles*, Masson, Paris 1984.
- [6] Jean-louis Marrien, *Analyse numérique avec Matlab*, Dunod, Paris, 2007.
- [7] Raphaèle Herbin, *Analyse numérique des équations aux dérivées partielles*, Master de mathématiques Université Aix Marseille 1, 18 octobre 2007.
- [8] Martin Campos Pinto, *Approximation polynômiale des fonctions et des intégrales*, 21 décembre 2006.
- [9] B. Mercier, *Stabilité et convergence des méthodes spectrales polynômiales. Application à l'équation d'advection*, AFCET, 1982.