Résumé du cours d'Analyse Numérique du Professeur Marco Picasso

jean-eloi.lombard@epfl.ch

19 juillet 2007

Table des matières

1 Problèmes d'interpolation

Base de Lagrange La base de Lagrange est donnée par les polynômes :

$$\phi_j(t) = \prod_{\substack{k=0 \\ k \neq j}}^{n} \frac{t - t_k}{t_j - t_k}$$

avec t_j des valeurs distinctes données.

Polynôme interpollant Le polynôme interpollant de la fonction f est donné par :

$$p(t) = \sum_j f(t_j)\phi_j(t)$$

Theoreme 1.1 (Erreur maximale) *Soit f une fonction $(n+1)$ fois dérivable sur $[a; b]$.*
Si p_n (le polynôme interpollant de f) est défini, alors :

$$\max_{t\in[a;b]} |f(t) - p_n(t)| \leq \frac{1}{2(n+1)} \left(\frac{b-a}{n}\right)^{n+1} \max_{t\in[a;b]} |f^{(n+1)}(t)|$$

2 Dérivation Numérique

2.1 Dérivées numériques d'ordre 1

Soit $f \in C^1 : \mathbb{R} \to \mathbb{R}$, $x_0 \in \mathbb{R}$ et $h > 0$. On approxime $f'(x_0)$ par les trois opérateurs suivants :

1. $f'(x_0) \approx \frac{f(x_0+h)-f(x_0)}{h} = \frac{\Delta_h f(x_0)}{h}$
2. $f'(x_0) \approx \frac{f(x_0)-f(x_0-h)}{h} = \frac{\nabla_h f(x_0)}{h}$
3. $f'(x_0) \approx \frac{f(x_0+\frac{h}{2})-f(x_0-\frac{h}{2})}{h} = \frac{\delta_h f(x_0)}{h}$

Définition 2.1 (Opérateurs de différence première finie) *Lorsque $h > 0$ est donné, les opérateurs Δ_h, ∇_h et δ_h sont appelés opérateurs de différence première respectivement progressive, rétrograde et centrée.*

Theoreme 2.1 *Les opérateurs de différence première Δ_h, ∇_h et δ_h sont linéaires.*

Theoreme 2.2 *Si $f \in C^2 : \mathbb{R} \to \mathbb{R}$ et si $x_0 \in \mathbb{R}$ est fixé et si $h_0 \in \mathbb{R}^+$ alors :*

$$\left| f'(x_0) - \frac{f(x_0+h) - f(x_0)}{h} \right| < Ch \qquad \forall h \leq h_0$$

Theoreme 2.3 *Si $f \in C^3 : \mathbb{R} \to \mathbb{R}$, $x_0 \in \mathbb{R}$ fixé et $h_0 \in \mathbb{R}^+$ donné alors il existe $C \in \mathbb{R}^+$ tel que :*

$$\left| f'(x_0) - \frac{f(x_0+\frac{h}{2}) - f(x_0-\frac{h}{2})}{h} \right| \leq Ch^2 \qquad \forall h \leq h_0$$

Démonstration Supposons $f \in C^3$. Le développement limité de f autour de x_0 est donc donné par :

$$f\left(x_0 + \frac{h}{2}\right) = f(x_0) + \frac{h}{2} f'(x_0) + \frac{f''(x_0)}{2!}\frac{h^2}{4} + \frac{f'''(\xi)}{3!}\frac{h^3}{8} \tag{1}$$

$$f\left(x_0 - \frac{h}{2}\right) = f(x_0) - \frac{h}{2} f'(x_0) + \frac{f''(x_0)}{2!}\frac{h^2}{4} - \frac{f'''(\eta)}{3!}\frac{h^3}{8} \tag{2}$$

avec $\xi \in]x_0, x_0+\frac{h}{2}[$ et $\eta \in]x_0-\frac{h}{2}, x_0[$. En soustrayant l'équation 1 à l'équation 2 et en divisant le résultat par h, il vient :

$$\left| f'(x_0) - \frac{f(x_0+\frac{h}{2}) - f(x_0-\frac{h}{2})}{h} \right| = \left| \frac{f'''(\xi) + f'''(\eta)}{2} \right| \frac{h^2}{24}$$

et pour finir, il suffit de poser :

$$C = \frac{1}{24} \max_{x \in]x_0-\frac{h}{2}, x_0+\frac{h}{2}[} |f'''(x)|$$

Définition 2.2 (Erreurs de troncatures) $\frac{\Delta_h f(x_0)}{h}$, $\frac{\nabla_h f(x_0)}{h}$ *et* $\frac{\delta_h f(x_0)}{h}$ *sont les formules de différences finies progressives, rétrogrades et centrées pour l'approximation de $f'(x_0)$. Les différences*

$$\left| f'(x_0) - \frac{f(x_0+h) - f(x_0)}{h} \right|$$

$$\left| f'(x_0) - \frac{f(x_0) - f(x_0-h)}{h} \right|$$

$$\left| f'(x_0) - \frac{f(x_0+\frac{h}{2}) - f(x_0-\frac{h}{2})}{h} \right|$$

sont les erreurs de troncatures. *Les deux premières sont d'ordre h et on dit qu'elles sont consistantes à l'ordre 1 en h, alors que la troisième est de l'ordre h^2 et consistante à l'odre 2 en h. Ainsi la formule de différences finies centrées est la plus précise.*

2.2 Dérivées numériques d'ordre supérieur

Définition 2.3 (Opérateurs de différences m finies) *Les opérateurs aux différences finies sont généralisés par les définitions récursives suivantes :*

$$\begin{aligned} \Delta_h^m f &= \Delta_h(\Delta_h^{m-1} f) \\ \nabla_h^m f &= \nabla_h(\nabla_h^{m-1} f) \\ \delta_h^m f &= \delta_h(\delta_h^{m-1} f) \end{aligned}$$

Ainsi l'opérateur de différence seconde centrée finie est donné par :

$$\delta_h^2 f(x) = f(x+h) - 2f(x) + f(x-h)$$

et l'approximation de la dérivée seconde est donnée par :

$$f''(x_0) \simeq \frac{\delta_h^2 f(x_0)}{h^2} = \frac{f(x_0+h) - 2f(x_0) + f(x_0-h)}{h^2}$$

Theoreme 2.4 *Soit $m \in \mathbb{N}$, $f \in C^{m+1} : \mathbb{R} \to \mathbb{R}$, $x_0 \in \mathbb{R}$ et $h_0 \in \mathbb{R}^+$ alors il existe $C \in \mathbb{R}^+$ tel que*

$$\left| f^m x_0 - \frac{\Delta_h^m f(x_0)}{h^m} \right| < Ch \qquad \forall h \leq h_0$$

$$\left| f^m x_0 - \frac{\nabla_h^m f(x_0)}{h^m} \right| < Ch \qquad \forall h \leq h_0$$

Theoreme 2.5 *Soit $m \in \mathbb{N}$, $f \in C^{m+2} : \mathbb{R} \to \mathbb{R}$, $x_0 \in \mathbb{R}$ et $h_0 \in \mathbb{R}^+$ alors il existe $C \in \mathbb{R}^+$ tel que :*

$$\left| f^m x_0 - \frac{\delta_h^m f(x_0)}{h^m} \right| < Ch^2 \qquad \forall h \leq h_0$$

3 Intégration Numérique

3.1 Généralités

Définition 3.1 (Formule du Trapèze) *Si g est une fonction continue sur [-1 ;1], la formule de quadrature*

$$J(g) = \sum_{j=1}^{M} \omega_j g(t_j)$$

est définie par la donnée de M points d'intégrations $t_1, \dots, t_M$ et M réels $\omega_1, \dots, \omega_M$ appelés poids de la formule de quadrature.

Définition 3.2 *La formule de quadrature*

$$J(g) = \sum_{j=1}^{M} \omega_j g(t_j)$$

pour calculer numériquement $\int_{-1}^{+1} p(t)dt$ est exacte pour les polynomes de degrés $r \geq 0$ si

$$J(p) = \int_{t_1}^{t_M} p(t)dt$$

pour tout polynôme p de degré $\leq r$

Theoreme 3.1 *Supposons :*

1. *$J(g) = \sum_{j=1}^{M} \omega_j g(t_j)$ exacte pour des polynômes de degré r*
2. *f une fonction donnée sur $[a; b]$*
3. *$L_h(f)$ la formule composite définie par*

$$L_h(f) = \sum_{i=0}^{N-1} \frac{x_{i+1} - x_i}{2} \sum_{j=1}^{M} \omega_j f\left(x_i + \frac{(x_{i+1} - x_i)(t_j + 1)}{2}\right)$$

4. *h le pas d'intégration.*
5. *f $(r+1)$ fois continuement dérivable sur $[a; b]$*

alors il existe une constante C indépendante de x_i telle que

$$\left| \int_a^b f(x)dx - L_h(f) \right| \leq Ch^{r+1} \tag{3}$$

Theoreme 3.2 *Soit $t_1 \leq \ldots \leq t_M$, M points distincts de $[-1; 1]$ et soit $\phi_1, \dots, \phi_M$ la base de Lagrange de $\mathbb{P}_{M-1}$ associée à ces M points, alors*

$$J(g) = \sum_{j=1}^{M} \omega_j g(t_j)$$

est exacte pour les polynômes de degré $M - 1$ si et seulement si

$$\omega_j = \int_{-1}^{1} \phi_j(t)dt \qquad \forall j = 1, \dots, M \tag{4}$$

3.1.1 Formule du rectangle

Définition 3.3 (Formule du rectangle) *Formule à 1 point* ($t_1 = 0$)

$$J(g) = 2g(0)$$

3.1.2 Formule de Gauss-Legendre

Définition 3.4 (Polynôme de Legendre)

$$L_M(t) = \frac{1}{2^M M!}\frac{d^M}{dt^M}(t^2-1)^M$$

Proposition 3.1 *Les M polynômes de Legendre forment une base orthogononale de $\mathbb{P}_M$.*

Proposition 3.2 *L_M a exactement M racines distinctes dans $[-1;1]$. Ces zéros sont appelés* points de Gauss.

Définition 3.5 (Formule de Gauss-Legendre) *La formule*

$$J(g) = \sum_{j=1}^{M} \omega_j g(t_j)$$

est la formule de Gauss-Legendre à M points si

1. *les points d'intégration $t_1 \leq \ldots \leq t_M$ sont les M solutions du polynôme de Legendre L_M.*
2. *les poids ω_j sont défini par :*

$$\omega_j = \int_{-1}^{1} \phi_j(t)dt \qquad j = 1, \ldots, M$$

   ou $\phi_1, \ldots, \phi_M$ est la base de Lagrange de $\mathbb{P}_{M-1}$ associé aux M points de Gauss.

Proposition 3.3 *La formule de Gauss-Legendre à M points est exacte pour les polynomes de degrés $r = 2M - 1$.*

3.1.3 Formule de Simpson

Définition 3.6 (Formule de Simpson) *Formule à trois points $t_1 = -1$, $t_2 = 0$, $t_3 = 1$. La base de Lagrange associée à ces trois points est donnée par :*

$$\phi_1(t) = \frac{t^2 - t}{2} \qquad \phi_2(t) = 1 - t^2 \qquad \phi_3(t) = \frac{t^2 + t}{2}$$

d'où, d'après l'équation 4 :

$$\omega_1 = \int_{-1}^{1} \phi_1(t)dt = \frac{1}{3} \qquad \omega_2 = \int_{-1}^{1} \phi_2(t)dt = \frac{4}{3} \qquad \omega_3 = \int_{-1}^{1} \phi_3(t)dt = \frac{1}{3}$$

la formule de Simpson s'écrit donc :

$$J(g) = \frac{1}{3}g(-1) + \frac{4}{3}g(0) + \frac{1}{3}g(1)$$

Proposition 3.4 *La formule est exacte pour des polynômes de degré 3.*

Proposition 3.5 *L'erreur est donnée par l'équation 3 qui devient :*

$$\left| \int_a^b f(x)dx - L_h(f) \right| \leq Ch^4$$

4 Décomposition de Cholesky et LU

Définition 4.1 (Matrice régulière) *Une matrice est dite* régulière *si et seulement si elle est inversible. Les deux termes sont équivalents.*

4.1 Décomposition LU

Theoreme 4.1 *Si A est une $N \times N$ matrice dont toutes les sous-matrices principales (les matrices carrées formée sur la diagonale) sont régulières, alors il existe une décomposition unique*

$$A = LU$$

*avec L un matrice "**L**ower Triangular" et U une matrice "**U**pper Triangular" avec que des 1 sur la diagonale. La matrice U est obtenue par l'élimination de Gauss.*

4.1.1 Exemple

Soit $A \in Mat(3,3,\mathbb{R})$ et cherchons sa décomposition LU, on a donc :

$$L = \begin{bmatrix} l_{11} & & \\ l_{21} & l_{22} & \\ l_{31} & l_{32} & l_{33} \end{bmatrix} \text{ et } U = \begin{bmatrix} 1 & u_{12} & u_{13} \\ & 1 & u_{23} \\ & & 1 \end{bmatrix}$$

en effectuant le produit LU il vient :

$$LU = \begin{bmatrix} l_{11} & l_{11} + u_{12} & l_{11}u_{13} \\ l_{21} & l_{21}u_{12} + l_{22} & l_{21}u_{13} + l_{22}u_{23} \\ l_{31} & l_{31}u_{21} + l_{32} & l_{31}u_{13} + l_{32}u_{23} + l_{33} \end{bmatrix}$$

puis il suffit d'identifier avec la matrice A :

$$A = LU$$

$$\Leftrightarrow \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = \begin{bmatrix} l_{11} & l_{11} + u_{12} & l_{11}u_{13} \\ l_{21} & l_{21}u_{12} + l_{22} & l_{21}u_{13} + l_{22}u_{23} \\ l_{31} & l_{31}u_{21} + l_{32} & l_{31}u_{13} + l_{32}u_{23} + l_{33} \end{bmatrix}$$

4.2 Décomposition LL^T

Définition 4.2 (Matrice symétrique définie positive) $A \in Mat(n,n,\mathbb{R})$ *est dite* symétrique *définie positive si :*

1. $A = A^T$ *(A est symétrique)*
2. $y^T Ay \geq 0$ *pour tout* $y \in \mathbb{R}^n$
3. $y^T Ay = 0$ *si et seulement si* $y = 0$ *pour tout* $y \in \mathbb{R}^n$

Soit $A \in Mat(n,n,\mathbb{R})$ avec $n \in \mathbb{N}$.

Theoreme 4.2 *Si* $A \in Mat(n,n,\mathbb{R})$ *est symétrique définie positive alors toutes ses sous-matrices principales sont symétriques définies positives et sont donc régulière.*

Theoreme 4.3 (de Cholesky) *Si* A *est une matrice symétrique définie positive alors il existe une unique matrice triangulaire inférieure à valeurs diagonales positives, notée L, telle que* $LL^T = A$.

4.2.1 Exemple

Soit $A \in Mat(3, 3, \mathbb{R})$, cherchons sa décomposition de Cholesky. La matrice L à la forme :

$$L = \begin{bmatrix} l_{11} & & \\ l_{21} & l_{22} & \\ l_{31} & l_{32} & l_{33} \end{bmatrix}$$

donc L^T est donée par :

$$L^T = \begin{bmatrix} l_{11} & l_{21} & l_{31} \\ & l_{22} & l_{32} \\ & & l_{33} \end{bmatrix}$$

et le produit LL^T est donné par :

$$LL^T = \begin{bmatrix} l_{11}^2 & l_{11}l_{21} & l_{11}l_{31} \\ l_{21}l_{11} & l_{21}^2 + l_{22}^2 & l_{21}l_{31} + l_{22}l_{32} \\ l_{11}l_{31} & l_{31}l_{21} + l_{32}l_{22} & l_{31}^2 + l_{32}^2 + l_{33}^2 \end{bmatrix}$$

il suffit ensuite d'identifier termes à termes avec les éléments de A.

5 Résolution de systèmes linéaires par des méthodes itératives

5.1 But

Lorsqu'il s'agit de résoudre un système de N équations à N inconnues de la forme :

$$A\vec{x} = \vec{b} \tag{5}$$

il faut utiliser des algorithmes (élimination de Gauss, décomposition LU, ou décomposition de Cholesky si la matrice est symétrique définie positive) de complexité $o(N^3)$. Il est donc intéressant d'approximer les solutions de tels systèmes par des méthodes itératives.

5.2 Méthode

Soit $A, K, M \in Mat(N, N, \mathbb{R})$ telles que :

$$A = K - M \tag{6}$$

avec K une matrice régulière, alors l'équation 5 devient :

$$\vec{x} = K^{-1}M\vec{x} + K^{-1}\vec{b}$$

De cette égalité nous construisons le schéma suivant :

1. choix arbitraire de $\vec{x}_0$
2. ittération sur $\vec{x}^{n+1} = K^{-1}M\vec{x}^n + K^{-1}\vec{b}$

5.3 Méthodes de Jacobi et de Gauss

Soit D la diagonale de la matrice A, $-F$ les élèments au dessus de la diagonale et $-E$ ceux en dessous, alors :

$$A = D - E - F$$

5.3.1 Méthode de Jacobi

Posons $K = D$ et $M = E + F$, alors la matrice :

$$J = K^{-1}M = D^{-1}(E + F)$$

est appelée la matrice de Jacobi. Ainsi le schéma devient :

1. choix arbitraire de $\vec{x}_0$
2. ittération sur $\vec{x}^{n+1} = J\vec{x}^n + D^{-1}\vec{b}$

Theoreme 5.1 (Condition de Convergence) *Si A et $2D - A$ sont des matrices symétriques définies positives alors la méthode de Jacobi est convergente.*

5.3.2 Méthode de Gauss-Seidel

Posons $K = D - E$ et $M = F$, ce qui donne la matrice de Gauss-Seidel définie par :

$$G = K^{-1}M = (D - E)^{-1}F$$

et pour le schéma :

1. choix arbitraire de $\vec{x}_0$
2. ittération sur $(D - E)\vec{x}^{n+1} = F\vec{x}^n + \vec{b}$

Theoreme 5.2 (Condition de convergence) *Si A est une matrice symétrique définie positive, alors la méthode de Gauss-Seidel est convergente.*

6 Équations non-linéaires

Pour chercher un zéro d'une fonction $f \in C^1 : \mathbb{R} \to \mathbb{R}$ il faut procéder au schéma numérique suivant :

1. localisation de x_0 solution grossière par une méthode graphique
2. construction d'une suite $x_n : \mathbb{N} \to \mathbb{R}$ telle que $\lim_{n\to\infty} x_n = \bar{x}$ avec $\bar{x}$ solution de $f(x) = 0$.

6.1 Méthode de Newton-Raphson

Soit $f \in C^1 : \mathbb{R} \to \mathbb{R}$ admettant $\bar{x}$ comme solution vérifiant $f'(\bar{x}) \neq 0$. Supposons connu x_0 une approximation grossière de la solution $\bar{x}$. La méthode de Newton-Raphson est définie par la suite :

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} \tag{7}$$

Theoreme 6.1 *Soit :*

1. $f \in C^2 : \mathbb{R} \to \mathbb{R}$
2. $\bar{x}$ *vérifiant* $f(\bar{x}) = 0$ *et* $f'(\bar{x}) \neq 0$.

alors il existe $\epsilon \in \mathbb{R}^+$ *et* $x_0 \in \mathbb{R}$ *tel que si* $|x_0 - \bar{x}| \leq \epsilon$ *la suite donnée par la méthode de Newton-Raphson 7 converge quadratiquement vers* $\bar{x}$.

Theoreme 6.2 (Condition de convergence d'une méthode de point fixe) *Soit* $g \in C^1 : \mathbb{R} \to \mathbb{R}$ *et* $\bar{x}$ *un point fixe de* g *(soit* $g(\bar{x}) = \bar{x}$*).*
Si $|g'(\bar{x})| < 1$
Alors il existe $\epsilon > 0$ *tel que si* $|\bar{x} - x_0| \leq \epsilon$ *alors la suite :*

$$x_{n+1} = g(x_n)$$

converge linéairement lorsque $n \to \infty$.

6.2 Systèmes non-linéaires

Soit $N \in \mathbb{N}$ et $F : \mathbb{R}^N \to \mathbb{R}^N$ une fonction dont on cherche une solution, c'est à dire un vecteur $\vec{x}_{solution}$ vérifiant $f(\vec{x}_{solution}) = \vec{0}$. L'équation $f(\vec{x}) = \vec{0}$ est donc un système non-linéaire (a priori) de N équations à N inconnues. En posant $\vec{x} = (x_1, \ldots, x_N)$ l'équation $f(\vec{x}) = \vec{0}$ est équivalent à :

$$f(\vec{x}) = \vec{0} \Leftrightarrow \begin{pmatrix} f_1(\vec{x}) \\ \vdots \\ f_N(\vec{x}) \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$$

en admettant $f_1, \ldots, f_N \in C^1$ il est possible de définir la matrice Jacobienne de f associée au point $\vec{x} \in \mathbb{R}^N$ par :

$$Df(x) = \begin{bmatrix} \frac{\partial f_1(x)}{\partial x_1} & \cdots & \frac{\partial f_1(x)}{\partial x_N} \\ \vdots & \vdots & \vdots \\ \frac{\partial f_N(x)}{\partial x_1} & \cdots & \frac{\partial f_N(x)}{\partial x_N} \end{bmatrix}$$

La méthode de Newton-Raphson est ainsi généarlisée aux systèmes non-linéaires par :

$$\begin{aligned} & \vec{x}^{n+1} = \vec{x}^n - Df(\vec{x}^n)^{-1} f(\vec{x}^n) \\ \Leftrightarrow \quad & Df(\vec{x}_n)(\vec{x}^n - \vec{x}^{n+1}) = f(\vec{x}^n) \end{aligned}$$

Proposition 6.1 *Si la solution grossière $\vec{x}_0$ est "suffisament proche" de la solution $\vec{x}_{solution}$ alors la méthode de Newton-Raphson "généralisée au systèmes non-linéaire" converge quadratiquement.*

Le schéma numérique pour la résolution d'un système non-linéaire est donc donnée par :

1. construction du vecteur $\vec{b} = f(\vec{x}^n)$
2. construction de la matrice $A = Df(\vec{x}^n)$
3. résolution du système $A\vec{y} = \vec{b}$ (avec $\vec{y} = \vec{x}^n - \vec{x}^{n+1}$) par :
   (a) élimination de Gauss
   (b) décomposition (LU ou LL^T)
   (c) méthode de Gauss-Seidel ou Jacobi
4. on pose $\vec{x}^{n+1} = \vec{x}^n - \vec{y}$

7 Équations différentielles

Soit $f \in C^1 : (x,t) \in \mathbb{R} \times \mathbb{R}^+ \to \mathbb{R}$ et $u_0 \in \mathbb{R}$ la valeur initiale. Nous cherchons une fonction $f \in C^1 : \mathbb{R}^+ \to \mathbb{R}$ qui satisfait :

$$\begin{aligned} \dot{u}(t) &= f(u(t),t) \qquad \forall t \geq 0 \\ u(0) &= u_0 \end{aligned}$$

la recherche d'une telle fonction est dite "problème de Cauchy".

Theoreme 7.1 *Soit :*

1. $f \in C^1 : \mathbb{R} \to \mathbb{R}$
2. $l : \mathbb{R} \to \mathbb{R}$ *telle que pour tout* $x, y \in \mathbb{R}$ *et* $t \in \mathbb{R}^+$

$$(f(x,t) - f(y,t))\,(x-y) \leq l(t)|x-y|^2$$

alors le probème de Cauchy admet une solution globale unique.

Theoreme 7.2 (Cauchy-Lipschitz) *Soit :*

1. $f \in C^1 : \mathbb{R} \to \mathbb{R}$
2. $L \in \mathbb{R}$, $x, y \in \mathbb{R}$ *et* $t \in \mathbb{R}^+$ *tel que*

$$|f(x,t) - f(y,t)| \leq L|x-y|$$

alors le problème de Cauchy admet une solution globale unique.

7.1 Schémas d'Euler

Dans les deux cas, on approxime :

$$\dot{u}(t) = \frac{u(t_{n+1}) - u(t_n)}{t_{n+1} - t_n} = \frac{u(t_{n+1}) - u(t_n)}{h_n}$$

7.1.1 Schéma d'Euler progressif

Le schéma est défini par l'approximation suivante du problème de Cauchy :

$$\begin{aligned} \frac{u^{n+1} - u^n}{h_n} &= f(u^n, t_n) \qquad \forall n = 0, \ldots, n \\ u^0 &= u_0 \end{aligned}$$

ce schéma est explicite, c'est-à dire qu'il permet de calculer directement u^{n+1} en fonction de u^n :

$$u^{n+1} = u^n + h_n f(u^n, t_n)$$

7.1.2 Schéma d'Euler rétrograde

Le schéma est défini par :

$$\begin{aligned} \frac{u^{n+1} - u^n}{h_n} &= f(u^{n+1}, t_{n+1}) \qquad \forall n = 0, \ldots, n \\ u^0 &= u_0 \end{aligned}$$

Ce schéma est implicite car il n'est pas possible de calculer directement u^{n+1} en fonction de u_n car :

$$u^{n+1} - h_n f(u^{n+1}, t_{n+1}) = u^n$$

8 Problèmes aux limites unidimensionnels

Étant donné deux fonction $f, c \in C^1 : [0;1] \to \mathbb{R}$, on cherche une fonction $u \in C^2 : [0;1] \to \mathbb{R}$ vérifiant :

$$\begin{aligned} &-u''(x) + c(x)u(x) = f(x) \qquad 0 < x < 1 \\ &u(0) = u(1) = 0 \end{aligned}$$

8.1 Méthode des différences finies

Soit $N \in \mathbb{N}^*$ le nombre de points de discrétisation et $h = \frac{1}{N+1}$ le pas d'espace. On pose $x_j = jh \ \forall j = 1, \ldots, N+1$ et on approxime $u''(x)$ par :

$$\begin{aligned} u''(x) &= \frac{\delta_h^2 u(x)}{h^2} + O(h^2) \\ &= \frac{u(x+h) - 2u(x) + u(x-h)}{h^2} + O(h^2) \end{aligned}$$

Le problème est donc approximé par la méthode de différence finie centrée :

$$\begin{aligned} &\frac{-u_{j-1} + 2u_j - u_{j+1}}{h^2} + c(x_j)u_j = f(x_j) \qquad 1 \le j \le N \\ &u_0 = u_{N+1} = 0 \end{aligned} \tag{8}$$

Si $\vec{u}$ est un vecteur de dimension N de composantes $u_1, \ldots, u_n$ et si $\vec{f}$ est le vecteur de dimension N de composantes $f(x_1), \ldots, f(x_n)$ et si A est la matrice tri-diagonale définie par :

$$A = \frac{1}{h^2}\begin{bmatrix} 2+c_1h^2 & -1 & & & \\ -1 & 2+c_2h^2 & -1 & & \\ & -1 & \ddots & \ddots & \\ & & \ddots & \ddots & -1 \\ & & & -1 & 2+c_nh^2 \end{bmatrix}$$

avec $c_i = c(x_i)$ alors le problème 8 est équivalent à $A\vec{u} = \vec{f}$.

Theoreme 8.1 *On suppose :*

1. $c(x) \geq 0, \forall x \in [0;1]$
2. $u \in C^4$

alors il existe une constante C telle que :

$$\max_{1\leq j\leq N} |u(x_j) - u_j| \leq Ch^2$$

L'erreur est donc 4 fois plus petite chaque fois que l'on double le nombre de points de discrétisation.

9 Problèmes paraboliques

9.1 Présentation du problème

Soit :

1. un barreau métallique sur l'intervalle $[0;L]$
2. soit $w(x)$ la température au temps t_0 au point x.
3. soit $f(x,t)$ la puissance par unité de longueur fournie au point $x \in [0;L]$ au temps $t > t_0$.
4. $\rho \in \mathbb{R}^+$ la densité volumique, $c_p \in \mathbb{R}^+$ la chaleur spécifique et $k \in \mathbb{R}^+$ la conductivité thermique.
5. alors la température $u(x,t)$ du barreau au point x à l'instant t est donné par :

$$\begin{aligned} \rho c_p \frac{\partial u(x,t)}{\partial t} - k\frac{\partial^2 u(x,t)}{\partial x^2} &= f(x,t) && \forall x \in]0;L[, \forall t > 0 \qquad (9) \\ u(0,t) = u(L,t) &= 0 && \forall t > 0 \\ u(x,0) &= w(x) && \forall x \in]0;L[\end{aligned}$$

9.2 Discrétisation

Discrétisons l'intervalle représentant le barreau en posant $N \in \mathbb{N}$ le nombre de points de discrétisation, $h = \frac{1}{N+1}$ le pas d'espace et $x_i = ih$ avec $i = 0, \ldots, N+1$. On a donc $\vec{u}(x,t) \approx \vec{u}_i(t)$. En posant $\tau \in \mathbb{R}^+$ le pas de temps et $n \in \mathbb{N}$ on a $t_n = n\tau$ donc $\vec{u}(t_n) \approx \vec{u}^n$.

9.3 Intégration numérique

9.3.1 Schéma d'Euler progressif

En utilisant le schéma d'Euler progressif on a donc :

$$\begin{aligned} \frac{\vec{u}^{n+1} - \vec{u}^n}{\tau} &= -A\vec{u}^n + \vec{f}(n\tau) \qquad n \in \mathbb{N} \\ \vec{u}^0 &= \vec{w} \end{aligned}$$

ce qui, en explicitant $\vec{u}^{n+1}$ donne :

$$\begin{aligned} \vec{u}^{n+1} &= (I - \tau A)\vec{u}^n + \tau \vec{f}(n\tau) \qquad n \in \mathbb{N} \\ \vec{u}^0 &= \vec{w} \end{aligned}$$

avec

$$A = \frac{k}{h^2} \begin{bmatrix} 2 & -1 & & \\ -1 & \ddots & \ddots & \\ & \ddots & \ddots & -1 \\ & & -1 & 2 \end{bmatrix}$$

et I la matrice identité.

Proposition 9.1 (Condition de stabilité) *Le pas temporel τ est limité par la condition de stabilité :*

$$\tau \leq \frac{h^2}{2k}$$

9.3.2 Schéma d'Euler rétrograde

En utilisant le schéma d'Euler rétrograde, le problème parabolique devient :

$$\begin{aligned} \vec{u}^{n+1} &= (I - \tau A)^{-1} \left(\vec{u}^n + \tau \vec{f}((n+1)\tau)\right) \qquad n \in \mathbb{N} \\ \vec{u}^0 &= \vec{w} \end{aligned}$$

Contrairement au schéma progressif, le schéma rétrograde est stable pour tout pas de temps $\tau \in \mathbb{R}^+$.

9.4 Convergence

Les deux schémas sont d'ordre 1 en temps et 2 en espace. L'erreur commise est donc d'ordre $\tau + h^2$.

10 Problèmes hyperboliques

10.1 Équation de transport 1D et différence finie

10.1.1 Présentation du problème

Soit $f, c \in C^1 : \mathbb{R} \times \mathbb{R}^+ \to \mathbb{R}$. Nous cherchons $u : \mathbb{R} \times \mathbb{R}^+$ vérifiant l'équation :

$$\begin{aligned} \frac{\partial u(x,t)}{\partial t} + c(x,t)\frac{\partial u(x,t)}{\partial x} &= f(x,t) & \forall x \in \mathbb{R}, \forall t > 0 \\ u(x,0) &= w(x) & \forall x \in \mathbb{R} \end{aligned}$$

10.1.2 Discrétisation

Pour approximer le problème utilisons un pas spatial h tel que $x_j = jh$ $\forall j = \pm 1, \pm 2, \ldots$ et un pas de temps τ, ce qui donne $t_n = n\tau$ avec $n = 0, 1, \ldots$. On a donc $u(x,t) \approx u_j^n$.

10.1.3 Approximation par la méthode de différence finie décentrée

Le schéma numérique de la méthode de différence finie décentrée est donnée par :

$$\begin{aligned} &\frac{u_j^{n+1} - u_j^n}{\tau} + \frac{\left((c_j^n)^+\right)(u_j^n - u_{j-1}^n)}{h} + \frac{\left((c_j^n)^-\right)(u_{j+1}^n - u_j^n)}{h} = f(jh, n\tau) \\ &j = 0, \pm 1, \pm 2, \ldots \quad \text{et } n = 0, 1, 2, \ldots \end{aligned}$$

avec :

$$(c_j^n)^+ = \begin{cases} c(jh, n\tau) & \text{si } c(jh, n\tau) > 0 \\ 0 & \text{sinon} \end{cases}$$

et

$$(c_j^n)^- = \begin{cases} c(jh, n\tau) & \text{si } c(jh, n\tau) < 0 \\ 0 & \text{sinon} \end{cases}$$

10.1.4 Condition de stabilité

La condition de stabilité sur τ et h est donnée par :

$$\frac{\tau}{h} \leq \frac{1}{\sup_{x\in\mathbb{R},t>0}|c(x,t)|}$$

10.2 Équation des ondes 1D

Soit $f \in C^0 : [0;1]\times\mathbb{R}^+ \to \mathbb{R}$, $v, w : [0;1] \to \mathbb{R}$ et $c \in \mathbb{R}^+$. Nous cherchons une fonction $u : [0;1] \times \mathbb{R}^+ \to \mathbb{R}$ telle que :

$$\begin{aligned}
\frac{\partial^2 u(x,t)}{\partial t^2} - c^2\frac{\partial^2 u(x,t)}{\partial x^2} &= f(x,t) && \forall x \in]0;1[, \quad \forall t > 0 \\
u(0,t) = u(1,t) &= 0 && \forall t \\
u(x,0) &= w(x) && \forall x \in]0;1[\\
\frac{\partial u(x,0)}{\partial t} &= v(x) && \forall x \in]0;1[
\end{aligned}$$

10.2.1 Discrétisation

Pour approximer le problème utilisons un pas spatial h tel que $x_j = jh$ $\forall j = \pm 1, \pm 2, \ldots$ et un pas de temps τ, ce qui donne $t_n = n\tau$ avec $n = 0, 1, \ldots$. On a donc $u(x,t) \approx u_j^n$.

10.2.2 Résolution par la méthode de différence finie centrée

Suite à une discrétisation en temps et en espace, l'équation d'onde devient :

$$\begin{aligned}
\vec{u}^{n+1} &= (2I - \tau^2 c^2 A)\vec{u}^n - u^{n-1} + \tau^2 \vec{f}(n\tau) \qquad (10) \\
u_0^n &= u_{N+1}^n = 0 \\
\vec{u}^0 &= \vec{w}(x) \\
\frac{u_j^{n+1} - u_j^n}{\tau} &= v_j(x)
\end{aligned}$$

10.2.3 Stabilité

Theoreme 10.1 *Le schéma 10 est stable si la condition suivante est satisfaite :*

$$\tau \leq \frac{h}{c}$$

11 Problèmes de convection-diffusion

Soit $f, c \in C^0 : [0;1] \to \mathbb{R}$ et $\epsilon \in \mathbb{R}^+$. Nous cherchons une fonction $u : [0;1] \to \mathbb{R}$ vérifiant :

$$\begin{aligned} &-\epsilon u''(x) + c(x)u'(x) = f(x) \qquad \forall x \in]0;1[\\ &u(0) = u(1) = 0 \end{aligned} \tag{11}$$

11.1 Discrétisation

Soit $N \in \mathbb{N}$ et $h = \frac{1}{N+1}$, on a alors $x_j = jh$ avec $j = 0, 1, \ldots, N+1$ et $u_j \approx u(x_j)$.

11.2 Schéma

En posant $\alpha_j \in [0;1]$ il est possible d'approcher :

$$u'(x_j) \approx \alpha_j \frac{u_j - u_{j-1}}{h} + (1 - \alpha_j) \frac{u_{j+1} - u_j}{h}$$

qui est une moyenne pondérée de la différence finie rétrograde et progressive. Utilisons donc le schéma numérique décentré suivant pour discrétiser 11 :

$$\begin{aligned} &\epsilon \frac{2u_j - u_{j+1} - u_{j-1}}{h^2} + \frac{c(x_j)}{h} \Big(\alpha_j (u_j - u_{j-1}) + (1 - \alpha_j)(u_{j+1} - u_j) \Big) = f(x_j) \\ &\forall j = 1, \ldots, N \\ &u_0 = u_{N+1} = 0 \end{aligned}$$

De manière à obtenir la meilleur approximation il faut poser :

$$\alpha_j = \frac{1}{2} + \frac{1}{2} \coth \frac{hc(x_j)}{2\epsilon} - \frac{\epsilon}{hc(x_j)}$$

Références

[1] Picasso : *Analyse Numérique pour Ingénieurs* Presses Polytecniques et Universitaires Romandes (2004)