Ecole Polytechnique, Promotion 2005
Analyse numérique et optimisation (MAP 431)
Contrôle Hors Classement du mardi 30 avril 2007

Corrigé proposé par G. Allaire

1 Différences finies

1. On remplace u_j^n par $u(t^n, x_j)$, où u est une fonction régulière, et on définit l'erreur de troncature du schéma par

$$E = (\Delta t)^{-1} \Big(u(t^{n+1}, x_j) - c_- u(t^n, x_{j-1}) - c_0 u(t^n, x_j) - c_+ u(t^n, x_{j+1}) \Big) .$$

où on n'a pas oublié de diviser par $1/\Delta t$ pour que la formule du schéma soit sous une forme "standard" (cf. la Remarque 2.2.5 du polycopié). On fait un développement de Taylor en t autour de t^n et en x autour du point x_j

$$u(t^{n+1}, x_j) = u(t^n, x_j) + \Delta t \frac{\partial u}{\partial t}(t_n, x_j) + \frac{1}{2}(\Delta t)^2 \frac{\partial^2 u}{\partial t^2}(t_n, x_j) + \mathcal{O}\Big((\Delta t)^3\Big)$$

et

$$u(t^n, x_{j\pm1}) = u(t^n, x_j) \pm \Delta x \frac{\partial u}{\partial x}(t_n, x_j) + \frac{1}{2}(\Delta x)^2 \frac{\partial^2 u}{\partial x^2}(t_n, x_j) + \frac{1}{6}(\Delta x)^3 \frac{\partial^3 u}{\partial x^3}(t_n, x_j) + \mathcal{O}\Big((\Delta x)^4\Big)$$

qui conduit à

$$E = \frac{1 - c_- - c_0 - c_+}{\Delta t} u(t^n, x_j) + \frac{\partial u}{\partial t}(t_n, x_j) + \frac{\Delta t}{2} \frac{\partial^2 u}{\partial t^2}(t_n, x_j) + \frac{(c_- - c_+)\Delta x}{\Delta t} \frac{\partial u}{\partial x}(t_n, x_j)$$

$$- \frac{(c_- + c_+)(\Delta x)^2}{2\Delta t} \frac{\partial^2 u}{\partial x^2}(t_n, x_j) + \frac{(c_- - c_+)(\Delta x)^3}{6\Delta t} \frac{\partial^3 u}{\partial x^3}(t_n, x_j) + \mathcal{O}\Big((\Delta t)^2 + c_\pm (\Delta x)^4/\Delta t\Big).$$

Pour que E ait une limite quand Δt et Δx tendent vers 0 (indépendemment), et puisque c_-, c_0, c_+ ne dépendent de Δt et Δx qu'à travers le rapport $\Delta t/(\Delta x)^2$, il faut que

$$c_- + c_0 + c_+ = 1 \quad \text{et} \quad c_- = c_+.$$

On note $a = c_-(\Delta x)^2/(\nu \Delta t)$ et on en déduit une simplification de l'erreur de troncature

$$E = \frac{\partial u}{\partial t}(t_n, x_j) + \frac{\Delta t}{2} \frac{\partial^2 u}{\partial t^2}(t_n, x_j) - a\nu \frac{\partial^2 u}{\partial x^2}(t_n, x_j) + \mathcal{O}\Big((\Delta t)^2 + a(\Delta x)^2\Big),$$

où on a précisé le reste en fonction de a. Clairement, le schéma ne peut être consistant que si $a = 1$: dans ce cas, si u est une solution de l'équation de la chaleur, alors $E = \mathcal{O}\Big((\Delta t) + (\Delta x)^2\Big)$. Réciproquement, si u n'est pas solution de l'équation de la chaleur, alors l'erreur de troncature ne tend pas vers zéro. Autrement dit, le seul schéma consistant est

$$u_j^{n+1} = u_j^n \frac{\nu \Delta t}{(\Delta x)^2} \left(u_{j-1}^n - 2u_j^n + u_{j+1}^n \right) .$$

Il est d'ordre 2 en espace mais seulement d'ordre 1 en temps car on ne peut pas "éliminer" le terme de dérivée seconde en temps. Il s'agit, bien sûr, du schéma explicite centré en espace. Le calcul de l'erreur de troncature nous donne automatiquement l'équation équivalente

$$\frac{\partial u}{\partial t} + \frac{\Delta t}{2}\frac{\partial^2 u}{\partial t^2} - \nu\frac{\partial^2 u}{\partial x^2} = 0$$

ou bien, si on remplace la dérivée seconde en temps par une dérivée en espace,

$$\frac{\partial u}{\partial t} - \nu\frac{\partial^2 u}{\partial x^2} + \frac{\nu^2\Delta t}{2}\frac{\partial^4 u}{\partial x^4} = 0.$$

2. Pour obtenir un schéma de type Lax-Wendroff on considère le développement de Taylor

$$u(t_{n+1}, x_j) = u(t_n, x_j) + (\Delta t)\frac{\partial u}{\partial t}(t_n, x_j) + \frac{(\Delta t)^2}{2}\frac{\partial^2 u}{\partial t^2}(t_n, x_j) + \mathcal{O}\Big((\Delta t)^3\Big),$$

qui, pour la solution u de l'équation de la chaleur devient

$$u(t_{n+1}, x_j) = u(t_n, x_j) + \nu(\Delta t)\frac{\partial^2 u}{\partial x^2}(t_n, x_j) + \frac{(\nu\Delta t)^2}{2}\frac{\partial^4 u}{\partial x^2}(t_n, x_j) + \mathcal{O}\Big((\Delta t)^3\Big).$$

Comme d'habitude on note u_j^n une approximation de $u(t_n, x_j)$. Une approximation à l'ordre 2 de $\frac{\partial^2 u}{\partial x^2}(t_n, x_j)$ est

$$\frac{u_{j-1}^n - 2u_j^n + u_{j+1}^n}{(\Delta x)^2}.$$

Sur ce modèle une approximation à l'ordre 2 de $\frac{\partial^4 u}{\partial x^4}(t_n, x_j)$ est

$$\frac{(u_{j-2}^n - 2u_{j-1}^n + u_j^n) - 2(u_{j-1}^n - 2u_j^n + u_{j+1}^n) + (u_j^n - 2u_{j+1}^n + u_{j+2}^n)}{(\Delta x)^4}$$

$$= \frac{u_{j-2}^n - 4u_{j-1}^n + 6u_j^n - 4u_{j+1}^n + u_{j+2}^n}{(\Delta x)^4}.$$

Le schéma de type Lax-Wendroff est donc

$$u_j^{n+1} = c^2/2u_{j-2}^n + c(1-2c)u_{j-1}^n + (1-2c+3c^2)u_j^n + c(1-2c)u_{j+1}^n + c^2/2u_{j+2}^n,$$

où $c = \nu\Delta t/(\Delta x)^2$. Ce schéma est à 5 points en espace. Par construction, il est au moins d'ordre 2 en temps et en espace puisque le développement de Taylor en temps est fait jusqu'à l'ordre 2 inclus et que les approximations des dérivées spatiales sont d'ordre 2. Tous les coefficients de la formule du schéma sont positifs si $0 \leq c \leq 1/2$. Par conséquent, le schéma vérifie le principe du maximum discret et est donc stable L^∞ sous la condition CFL $2\nu\Delta t \leq (\Delta x)^2$. Comme il est linéaire, consistant et stable, il converge d'après le théorème de Lax.

2 Diffusion non locale

1. Soit le problème aux limites

$$\begin{cases} -\Delta u(x) + \int_\Omega c(x,y)\Big(u(x)-u(y)\Big)dy = f(x) & \text{dans } \Omega, \\ u = 0 \quad \text{sur } \partial\Omega. \end{cases} \tag{1}$$

Pour obtenir une formulation variationnelle de (1) on multiplie l'équation par une fonction test v, avec v s'annulant sur le bord $\partial\Omega$, et on intègre par parties. On obtient

$$\int_\Omega \nabla u \cdot \nabla v \, dx + \int_\Omega \int_\Omega c(x,y)\Big(u(x)-u(y)\Big)v(x)\, dy\, dx = \int_\Omega f v \, dx. \tag{2}$$

On peut réécrire le terme non-local de manière symétrique. En effet, en échangeant x et y

$$\begin{aligned} \int_\Omega \int_\Omega c(x,y)\Big(u(x)-u(y)\Big)v(x)\, dy\, dx &= \int_\Omega \int_\Omega c(y,x)\Big(u(y)-u(x)\Big)v(y)\, dx\, dy \\ &= -\int_\Omega \int_\Omega c(x,y)\Big(u(x)-u(y)\Big)v(y)\, dx\, dy \end{aligned}$$

par symétrie de $c(x,y)$. Par conséquent, en gardant la moitié du terme non-local identique et en changeant l'autre moitié avec la formule ci-dessus, (2) est équivalent à

$$\int_\Omega \nabla u \cdot \nabla v \, dx + \frac{1}{2}\int_\Omega \int_\Omega c(x,y)\Big(u(x)-u(y)\Big)\Big(v(x)-v(y)\Big)dx\, dy = \int_\Omega f v \, dx.$$

Pour que tous les termes dans (2) aient un sens et pour vérifier la conditions aux limites de Dirichlet on choisit $H_0^1(\Omega)$ comme espace de Hilbert. Finalement, la formulation variationnelle proposée est :

trouver $u \in H_0^1(\Omega)$ tel que (2) ait lieu pour tout $v \in H_0^1(\Omega)$.

On vérifie maintenant les hypothèses du théorème de Lax-Milgram 3.3.1. On définit une forme linéaire

$$L(v) = \int_\Omega f v \, dx$$

et une forme bilinéaire

$$a(u,v) = \int_\Omega \nabla u \cdot \nabla v \, dx + \frac{1}{2}\int_\Omega \int_\Omega c(x,y)\Big(u(x)-u(y)\Big)\Big(v(x)-v(y)\Big)dx\, dy$$

qui sont clairement continues par application de l'inégalité de Cauchy-Schwarz pour majorer les intégrales. Vérifions maintenant la coercivité de la forme bilinéaire. On a

$$\begin{aligned} a(v,v) &= \int_\Omega |\nabla v|^2 dx + \frac{1}{2}\int_\Omega \int_\Omega c(x,y)\Big(v(x)-v(y)\Big)^2 dx\, dy \\ &\geq \int_\Omega |\nabla v|^2 dx \geq C\|v\|^2_{H^1(\Omega)} \end{aligned}$$

car $c \geq 0$ et grâce à l'inégalité de Poincaré (Ω est borné). Par conséquent il existe une unique solution $u \in H_0^1(\Omega)$ de la formulation variationnelle (2).

2. Vérifions maintenant que cette solution de la formulation variationnelle est bien une solution du problème (1). Pour cela on admet que $u \in H^2(\Omega)$. On utilise la formule de Green (4.22) qui implique que (2) est équivalent à

$$\int_\Omega \left(-\Delta u + \int_\Omega c(x,y)\Big(u(x) - u(y)\Big)dy - f \right) v \, dx = 0.$$

Comme v est une fonction quelconque dans $H_0^1(\Omega)$, on en déduit par application du Corollaire 4.2.2 que

$$-\Delta u(x) + \int_\Omega c(x,y)\Big(u(x) - u(y)\Big)dy = f(x)$$

comme fonctions de $L^2(\Omega)$, ce qui implique que cette égalité a lieu presque partout dans Ω. Par ailleurs, comme $u \in H_0^1(\Omega)$, le théorème de trace 4.3.13 (ou son corollaire 4.3.16) affirme que, sur le bord $\partial\Omega$, $u = 0$ en tant que fonctions de $L^2(\partial\Omega)$, ce qui implique que les conditions aux limites de Dirichlet ont lieu presque partout dans $\partial\Omega$. On a donc bien retrouvé le système (2) au sens "presque partout".

3. D'après la proposition 3.3.4, puisque la forme bilinéaire $a(u, v)$ est symétrique, la solution u minimise aussi l'énergie

$$E(v) = \frac{1}{2} a(v, v) - L(v),$$

c'est-à-dire

$$E(v) = \frac{1}{2} \int_\Omega |\nabla v|^2 dx + \frac{1}{4} \int_\Omega \int_\Omega c(x,y)\Big(v(x) - v(y)\Big)^2 dx \, dy - \int_\Omega f v \, dx.$$

4. La matrice de rigidité $\mathcal{K}$ est la somme de deux contributions

$$\mathcal{K} = \mathcal{K}_0 + \mathcal{K}_1,$$

où $\mathcal{K}_0$ est la matrice de rigidité du Laplacien et $\mathcal{K}_1$ est celle due au terme non-local. D'après la section 6.2.1 du cours on a

$$\mathcal{K}_0 = h^{-1} \begin{pmatrix} 2 & -1 & & & 0 \\ -1 & 2 & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & -1 & 2 & -1 \\ 0 & & & -1 & 2 \end{pmatrix}.$$

On note $\phi_j(x) = \phi\left(\frac{x - x_j}{h}\right)$ la j-ème fonction de base des éléments finis de Lagrange $\mathbb{P}_1$, où $\phi(x) = \max(0, 1 - |x|)$ est la fonction "chapeau" usuelle. Pour le terme non-local on a

$$\Big(\mathcal{K}_1\Big)_{ij} = \int_0^1 \Big(\phi_i(x) - \phi_i(y)\Big)\phi_j(x) \, dy \, dx = \int_0^1 \phi_i(x)\phi_j(x) \, dx - \left(\int_0^1 \phi_i(x) \, dx\right)\left(\int_0^1 \phi_j(x) \, dx\right)$$

et, grâce à la formule de quadrature des trapèzes, on en déduit

$$\left(\mathcal{K}_1\right)_{ij} = h\delta_{ij} - h^2.$$

Avec la notation $\mathbb{1}$ pour le vecteur dont toutes les composantes sont égales à 1, on a donc obtenu

$$\mathcal{K}_1 = h\mathrm{Id} - h^2 \mathbb{1} \otimes \mathbb{1}.$$

5. La matrice $\mathcal{K}_1$, et donc $\mathcal{K}$, est pleine et il est couteux de résoudre, par une méthode directe (de type factorisation de Cholesky ou élimination de Gauss), le système linéaire qui découle de la méthode des éléments finis de Lagrange $\mathbb{P}_1$. Une meilleur idée est de réécrire $\mathcal{K}$ sous la forme

$$\mathcal{K} = \mathcal{T} - h^2 \mathbb{1} \otimes \mathbb{1},$$

où $\mathcal{T}$ est tridiagonale et donc facile à factoriser. On utilise alors la formule suivante qui permet de calculer l'inverse de $\mathcal{K}$ en fonction de celui de $\mathcal{T}$

$$\mathcal{K}^{-1} = \mathcal{T}^{-1} + \frac{h^2}{1 - h^2 \mathcal{T}^{-1}\mathbb{1} \cdot \mathbb{1}} \mathcal{T}^{-1}\mathbb{1} \otimes \mathcal{T}^{-1}\mathbb{1},$$

qui n'a un sens que si $h^2\mathcal{T}^{-1}\mathbb{1} \cdot \mathbb{1} \neq 1$, ce qui est vrai dès que h est petit. Le système linéaire issu de la méthode des éléments finis de Lagrange $\mathbb{P}_1$

$$\mathcal{K}u = b \tag{3}$$

admet donc la solution

$$u = \mathcal{T}^{-1}b + \frac{h^2\, b \cdot \mathcal{T}^{-1}\mathbb{1}}{1 - h^2 \mathcal{T}^{-1}\mathbb{1} \cdot \mathbb{1}} \mathcal{T}^{-1}\mathbb{1}.$$

Bien sûr, pour résoudre (3) on ne calcule pas la matrice inverse $\mathcal{T}^{-1}$! En pratique, on factorise la matrice tridiagonale $\mathcal{T}$ puis on résout un premier système linéaire pour calculer $\mathcal{T}^{-1}\mathbb{1}$ et un deuxième système linéaire pour calculer u

$$\mathcal{T}u = b + \frac{h^2\, b \cdot \mathcal{T}^{-1}\mathbb{1}}{1 - h^2 \mathcal{T}^{-1}\mathbb{1} \cdot \mathbb{1}} \mathbb{1}.$$

L'ordre des matrices étant n, cette manière de faire demande de l'ordre de $\mathcal{O}(n^2)$ opérations (puisque la factorisation LU ou de Cholesky préserve la structure bande de la matrice) alors que l'approche naive consistant à factoriser $\mathcal{K}$ demande de l'ordre de $\mathcal{O}(n^3)$ opérations (même chose pour la place mémoire).