Calcul Scientifique
M2 Mapi^3 – 2020-2021

Christophe Besse

First printing, September 2020

Table des matières

1 Rappels et compléments 7

1.1 Les différences finies 7

1.1.1 Un problème modèle 7

1.1.2 Rappels sur les différences finies 8

1.2 Les éléments finis 10

1.3 Transformée de Fourier discrète 13

1.4 La Transformée en Sinus Discrète 18

1.5 Volumes finis 20

2 Équations aux dérivées partielles paraboliques 25

2.1 Problème modèle 25

2.2 Solutions classiques dans $\mathbb{R}^N$ 26

2.3 Équation de la chaleur dans un domaine borné 28

2.3.1 Unicité de la solution - Stabilité 29

2.3.2 Principe du maximum 31

2.3.3 Résolution de l'équation de la chaleur par séparation des variables 33

2.4 Résolution par transformée de Fourier discrète 35

2.5 Résolution par différences finies 36

2.6 Conditions aux limites transparentes 42

3 Équations de Schrödinger 45

3.1 Problème modèle 45

3.2 Équations aux dérivées partielles dispersives 46

3.3 Solutions classiques dans $\mathbb{R}^n$ 47

3.4 Schémas numériques **48**

3.5 Équation de Schrödinger non linéaire **51**

3.5.1 Schéma de splitting - pas fractionnaire 51

3.5.2 Le schéma de Crank-Nicolson 54

3.5.3 Schéma de relaxation 56

Bibliographie 61

Notes de Cours **61**

Articles **61**

Livres **62**

1. Rappels et compléments

Le but de ce chapitre est de rappeler les techniques de discrétisation spatiale vue les années précédentes et de les compléter.

1.1 Les différences finies

1.1.1 Un problème modèle

Nous considérons pour débuter le problème modèle suivant

$$\begin{cases} -u''(x) = f(x), & x \in]0,1[, \\ u(0) = u(1) = 0. \end{cases} \tag{1.1}$$

Il s'agit en fait dans ce cas simple mono dimensionnel d'une équation différentielle ordinaire mais où la condition de Cauchy est remplacée par des conditions de Dirichlet homogènes.

Dans ce cas très simple, on peut résoudre l'équation par intégrations successives

$$u'(s) = -\int_0^s f(t)\,dt + c_1$$

où c_1 est une constante à déterminer. Puis,

$$\begin{aligned} u(x) &= \int_0^x u'(s)\,ds \\ &= -\int_0^x \left(\int_0^s f(t)\,dt\right) ds + c_1 x + c_2, \end{aligned}$$

où c_2 est une constante à déterminer. Or, on a $u(0) = u(1) = 0$ et donc $c_2 = 0$ et

$$0 = u(1) = -\int_0^1 \left(\int_0^s f(t)\,dt\right) ds + c_1$$

d'où la valeur de c_1 et on a alors

$$u(x) = \int_0^x \left(\int_0^s f(t)\,dt\right) ds + x\int_0^1 \left(\int_0^s f(t)\,dt\right) ds. \tag{1.2}$$

R Il faut évidemment demander $f \in L^1(0,1)$.

Dans le cas présent, il serait facile de déterminer une approximation de (1.2) mais la connaissance de solutions explicites de type (1.2) est extrêmement rare et il vaut mieux construire une approximation qui donne une solution approchée dans tous les cas.

Le principe de base est de construire un maillage du domaine $\Omega = [0,1]$. En dimension 1, un maillage est simplement constitué d'une collection de points $(x_j)_{0\leqslant j\leqslant n+1}$ c'est à dire une subdivision de $[0,1]$

$$x_0 = 0 < x_1 < \cdots < x_n < x_{n+1} = 1.$$

Le maillage sera dit uniforme si les points x_j sont équidistants, c'est à dire $x_j = jh$ avec $h = 1/(n+1)$, $0 \leqslant j \leqslant n+1$.

Il n'est pas nécessaire d'avoir un maillage uniforme.

Les points x_j sont appelés sommets du maillage. On produit des intervalles (mailles)

$$I_k =]x_{k-1}, x_k[, \qquad k = 1, 2, \cdots, n+1.$$

Si le maillage est uniforme, $h = x_k - x_{k-1}$ pour tout k. Sinon, on définit le pas du maillage par

$$h = \max_{1\leqslant k\leqslant n+1}(x_k - x_{k-1}).$$

Le pas h est un indicateur de la finesse du maillage.

1.1.2 Rappels sur les différences finies

Le principe est basé sur le développement de Taylor des fonctions :

Théorème 1.1.1 Soient $I \subset \mathbb{R}$, $a \in I$, et $k \in \mathbb{N}$, $k \geqslant 1$ et $f : I \to \mathbb{R}$ une fonction k fois dérivable en a. Alors, il existe une fonction $h_k : I \to \mathbb{R}$ telle que

$$f(x) = f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \cdots + \frac{f^{(k)}(a)}{k!}(x-a)^k + h_k(x)(x-a)^k$$

et $\lim_{x\to a} h_k(x) = 0$. On définit le terme de reste $R_k(x) = h_k(x)(x-a)^k$ et ainsi $R_k(x) = o(|x-a|^k)$ quand $x \to a$.

Nous pouvons obtenir des formules pour R_k en supposant que f est $(k+1)$ fois dérivable.

Proposition 1.1.2

- si $f^{(k)} \in C^0(I)$, alors $R_k(x) = \dfrac{f^{(k+1)}(\xi)}{(k+1)!}(x-a)^{k+1}$.
- si $f^{(k+1)} \in C^0(I)$, alors $R_k(x) = \displaystyle\int_0^x \frac{f^{(k+1)}(t)}{k!}(x-t)^k dt$.
- si $f^{(k+1)} \in C^0(I)$, et I un ensemble fermé (ou $|f^{(k+1)}| \leqslant M$), alors $|R_k(x)| \leqslant M\dfrac{|x-a|^{k+1}}{(k+1)!}$ et $R_k(x) = O(|x-a|^{k+1}$.

Grâce à ces relations, on peut construire une version approchée des opérateurs différentiels (pour des fonctions f qui sont régulières). Soit $h \in \mathbb{R}$ (h doit être suffisamment petit $|h| \ll 1$), $f : I \to \mathbb{R}$ régulière, alors on a

(I) $f(x+h) = f(x) + hf'(x) + \frac{h^2}{2}f''(x) + \frac{h^3}{6}f'''(x) + \frac{h^4}{24}f^{\text{iv}}(x) + O(h^5)$.

(II) $f(x-h) = f(x) - hf'(x) + \frac{h^2}{2}f''(x) - \frac{h^3}{6}f'''(x) + \frac{h^4}{24}f^{\text{iv}}(x) + O(h^5)$.

De (I), on a $hf'(x) = f(x+h) - f(x) + O(h^2)$ ce qui donne $f'(x) = \dfrac{f(x+h)-f(x)}{h} + O(h)$. Comme $h \to 0$, on a

$$f'(x) - \frac{f(x+h)-f(x)}{h} = O(h) \to 0.$$

Nous obtenons donc une manière d'approcher $f'(x)$ par différence finie décentrée aval

$$f'(x) \approx \frac{f(x+h)-f(x)}{h}.$$

Puisque le terme de reste est $O(h)$, on dit qu'on a une approximation du premier ordre.
De (II), en suivant le même calcul, nous obtenons une approximation par différence finie décentrée amont

$$f'(x) \approx \frac{f(x)-f(x-h)}{h}.$$

En soustrayant (II) à (I), nous avons

$$f(x+h) - f(x-h) = 2hf'(x) + O(h^3)$$

et nous obtenons une nouvelle approximation centrée du second ordre

$$f'(x) \approx \frac{f(x+h)-f(x-h)}{2h}.$$

En ajoutant (II) à (I) conduit à

$$f(x+h) + f(x-h) = 2f(x) + h^2 f''(x) + O(h^4)$$

ainsi nous avons

$$f''(x) = \frac{f(x+h) - 2f(x) + f(x-h)}{h^2} + O(h^2)$$

ce qui donne une approximation du second ordre de $f''(x)$

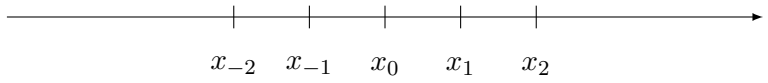
$$f''(x) \approx \frac{f(x+h) - 2f(x) + f(x-h)}{h^2}.$$

Ces constructions peuvent être étendues en plusieurs dimensions. Par exemple, soit $f : \mathbb{R}^2 \to \mathbb{R}$ une fonction régulière, alors on a l'approximation du second ordre suivante

$$\begin{aligned}
\partial_x f(x,y) &\approx \frac{f(x+h,y) - f(x-h,y)}{2h}, \\
\partial_y^2 f(x,y) &\approx \frac{f(x,y+h) - f(x,y) + f(x,y-h)}{h^2}, \\
\partial_{xy} f(x,y) &\approx \frac{f(x+h,y+k) + f(x+h,y-k) - f(x-h,y+k) + f(x-h,y-k)}{4hk}.
\end{aligned}$$

Ces formules sont utiles pour l'approximation numérique des EDPs. Nous échantillonnons $\mathbb{R}$

avec un taux d'échantillonnage h et définissons $x_j = jh$, $j \in \mathbb{Z}$,

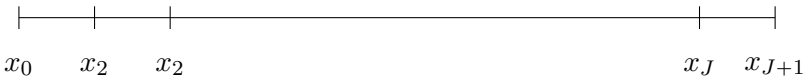


Nous évaluons la fonction régulière $f : \mathbb{R} \to \mathbb{R}$ à chaque $(x_j)_{j\in\mathbb{Z}}$ et nous obtenons la suite $(f(x_j) := f_j)_{j\in\mathbb{Z}}$. Afin d'obtenir une approximation de f' en chaque nœud x_j, nous utilisons les relations précédentes et

$$\begin{aligned}
f'(x_j) \text{ est approchée par } & \frac{f_{j+1} - f_j}{h} && 1^{\text{er}} \text{ ordre,}\\
& \frac{f_j - f_{j-1}}{h} && 1^{\text{er}} \text{ ordre,}\\
& \frac{f_{j+1} - f_{j-1}}{2h} && 2^{\text{nd}} \text{ ordre,}\\
f''(x_j) \text{ est approchée par } & \frac{f_{j+1} - 2f_j + f_{j-1}}{h^2} && 2^{\text{nd}} \text{ ordre.}
\end{aligned}$$

Évidemment, ici $j \in \mathbb{Z}$ et nous pouvons évaluer ces approximations pour chaque j. Dans des expériences numériques réelles, nous nous intéressons à l'approximation de solutions d'EDPs dans des domaines bornés Ω. Ces EDPs sont donc complétées avec des conditions aux limites sur la frontière $\Gamma = \partial\Omega$ (Dirichlet, Neumann, ...).

La question est de savoir comment prendre en compte ces conditions avec les relations liées aux différences finies. Pour présenter l'idée, nous considérons un cas mono-dimensionnel et définissons l'intervalle $I = [x_\ell, x_r]$ de longueur $x_r - x_\ell = \mu$. Nous maillons I avec $J + 2$ nœuds régulièrement espacés avec des sous-intervalles de longueur $h = (x_r - x_\ell)/(J + 1)$.



- Pour les conditions de Dirichlet homogènes, nous demandons $u(x_0) = u_0 = 0$ et $u(x_{J+1}) = u_{J+1} = 0$.
- Pour les conditions de Neumann homogènes, plusieurs possibilités existent dépendant du problème. Nous présentons ici deux méthodes :

  (CLN$_1$) Pour le nœud x_0, nous utilisons une approximation décentré aval d'ordre 1 $(u_1 - u_0)/h = 0$. Pour le nœud x_{J+1}, nous utilisons une approximation décentrée amont d'ordre 1 $(u_{J+1} - u_J)/h = 0$.

  (CLN$_2$) Pour cette seconde approximation, on introduit deux nœuds fantômes $x_{-1} = x_0 - h$ et $x_{J+2} = x_{J+1} + h$. Pour le nœud x_0, on utilise une différence finie centrée $(u_1 - u_{-1})/2h = 0$. Et pour le nœud x_{J+1}, $(u_{J+2} - u_J)/2h = 0$.

1.2 Les éléments finis

Nous rappelons la méthode sur un exemple mono-dimensionnel standard

$$\begin{cases} -\partial_x^2 u + \alpha u = 1, & x \in]0,1[,\ \alpha > 0,\\ \partial_x u(0) = u(1) = 0. \end{cases} \tag{1.3}$$

La méthode des éléments finis repose sur la formulation variationnelle associée au problème (1.3). On multiplie l'équation en volume par une fonction v suffisamment régulière et on intègre. Par intégration par partie, il vient

$$\int_0^1 u'(x)v'(x)\,dx - [u'v]_0^1 + \alpha \int_0^1 u(x)v(x)\,dx = \int_0^1 v(x)\,dx,$$

soit encore

$$\int_0^1 u'v' + \alpha uv\,dx - (u'(1)v(1) - \underbrace{u'(0)}_{=0}\,v(0)) = \int_0^1 v(x)\,dx.$$

Afin de donner un sens aux différentes intégrales, il faut imposer $(u, u') \in (L^2(0,1))^2$, et une régularité identique pour v et v'. On manque d'information sur le terme résiduel $u'(1)v(1)$. Sans hypothèse supplémentaire sur u' en $x = 1$, on est donc amener à considérer $v(1) = 0$. On construit l'espace de Hilbert

$$V = \{f \in H^1(0,1) \,|\, \gamma_0 f(1) = 0\},$$

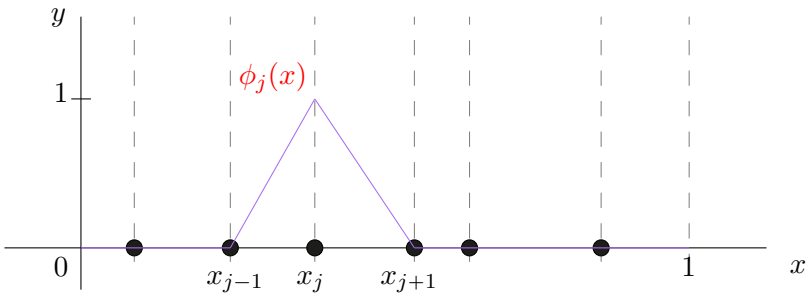
où γ_0 est l'application trace. On a donc la formulation variationnelle

$$\text{Trouver } u \in V \text{ telle que } a(u,v) := \int_0^1 u'v' + \alpha uv\,dx = \int_0^1 v(x)\,dx := \ell(v), \quad \forall v \in V. \tag{1.4}$$

Par application du théorème de Lax-Milgram (dont on s'assure que les hypothèses sont vérifiées), il existe une unique solution à (1.3). On montre que V est un espace de Hilbert séparable. Il existe donc une suite $(\varphi_j)_{j\in\mathbb{N}}$ libre et totale dans V. On construit le sous espace de dimension finie V_m de V par $V_m = \text{Vect}\{\varphi_1, \cdots, \varphi_m\}$ et on construit le problème approché

$$\text{Trouver } u_m \in V_m \text{ telle quea}(u_m, v) = \ell(v), \quad \forall v \in V_m. \tag{1.5}$$

Fabriquer de manière explicite la base $(\varphi_j)_{j\in\mathbb{N}}$ n'est pas utile pour la méthode des éléments finis. On lui préfère une base de V_m dont les supports des vecteurs sont à support compact et formant une base "quasi" orthogonale. Les fonctions de base qu'on utilise pour V_m sont des fonctions chapeaux



La fonction chapeau ϕ_j associée au sommet x_j est définie pour $j = 1, \cdots, n$ par

$$\phi_j(x) = \begin{cases} 0, & \text{si } x \notin [x_{j-1}, x_{j+1}] = I_j \cup I_{j+1}, \\ \dfrac{x - x_{j-1}}{x_j - x_{j-1}}, & \text{si } x \in [x_{j-1}, x_j] = I_j, \\ \dfrac{x_{j+1} - x}{x_{j+1} - x_j}, & \text{si } x \in [x_j, x_{j+1}] = I_{j+1}. \end{cases}$$

On a donc

$$\phi_j(x_j) = 1 \qquad \text{et} \qquad \phi_j(x_k) = 0, \ \forall k \neq j,$$

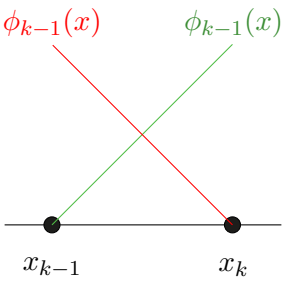
et $\phi_j \in V_m$. Lorsque le maillage est uniforme,

$$\phi_j(x) = \begin{cases} 0, & \text{si } x \notin I_j \cup I_{j+1}, \\ \dfrac{x - x_{j-1}}{h}, & \text{si } x \in I_j, \\ \dfrac{x_{j+1} - x}{h}, & \text{si } x \in I_{j+1}, \end{cases}$$

et on remarque alors que les ϕ_j se définissent à partir d'une unique fonction ϕ par

$$\phi_j(x) = \phi\left(\frac{x - x_j}{h}\right), \qquad \phi(x) = \begin{cases} 1 - |x|, & \text{si } |x| \leqslant 1, \\ 0 & \text{si } |x| > 1. \end{cases}$$

Dans le cas général, on a sur une maille I_k deux fonctions de base qui coexistent



et on peut faire correspondre $\phi_k(x)|_{I_k}$ à la fonction $\widehat{\phi_1}(x) = x$, $0 \leqslant x \leqslant 1$ et $\phi_{k-1}(x)|_{I_k}$ à la fonction $\widehat{\phi_2}(x) = 1 - x$, $0 \leqslant x \leqslant 1$. On a ainsi

$$\phi_{k-1}(x)|_{I_k} = \widehat{\phi_2}\left(\frac{x - x_{k-1}}{x_k - x_{k-1}}\right) \quad \text{et} \quad \phi_k(x)|_{I_k} = \widehat{\phi_1}\left(\frac{x - x_{k-1}}{x_k - x_{k-1}}\right).$$

L'espace V_m est un sous espace de $C^0([0,1])$ de dimension finie n, et toute fonction $v \in V_m$ est définie de manière unique par ses valeurs aux sommets $(x_j)_{1 \leqslant j \leqslant n}$

$$v(x) = \sum_{j=1}^{m} v(x_j)\phi_j(x), \quad \forall\, x \in [0,1].$$

Cette base permet donc de caractériser une fonction de V_m par ses valeurs aux sommets du maillage : on parle dans ce cas d'éléments finis de Lagrange.

Afin d'approcher (1.5), on décompose alors u_m sur la base de $\{\phi_j\}_{1 \leqslant j \leqslant m}$ et on prend $v = \phi_i$, $1 \leqslant i \leqslant m$ ce qui donne

$$\sum_{j=1}^{m} a(\phi_j, \phi_i) u_m(x_j) = l(\phi_i).$$

En notant $U_m = (u_m(x_j))_{1 \leqslant j \leqslant n} \in \mathbb{R}^m$, $b_m = (l(\phi_i))_{1 \leqslant i \leqslant n} \in \mathbb{R}^m$, et en introduisant la matrice

$$A_m = (a(\phi_j, \phi_i))_{1 \leqslant i,j \leqslant m} \in \mathbb{R}^{m \times m},$$

la formulation variationnelle revient à résoudre dans $\mathbb{R}^m$ le système linéaire

$$A_m U_m = b_m.$$

Il faut calculer la matrice A_m. En utilisant le fait que les vecteurs de base de V_m soient à support compact, il est facile de voir que la matrice A_m est creuse. Pour des nœuds du maillage loin des bords, on a

$$(-\frac{1}{h} + \alpha\frac{h}{6})u_{i-1} + (\frac{2}{h} + \alpha\frac{h2}{3})u_i + (-\frac{1}{h} + \alpha\frac{h}{6})u_{i+1} = 2\frac{h}{2}, \quad 2 \leqslant i \leqslant m.$$

Après prise en compte des conditions aux limites, la matrice A_m est la somme d'une matrice de rigidité S_m et d'une matrice de masse M_m dont les valeurs sont

$$S_m = \frac{1}{h}\begin{pmatrix} 1 & -1 & & & \\ -1 & 2 & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & -1 & 2 & -1 \\ & & & -1 & 2 \end{pmatrix}, \qquad M_m = \frac{\alpha h}{6}\begin{pmatrix} 2 & 1 & & & \\ 1 & 4 & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & 1 & 4 & 1 \\ & & & 1 & 2 \end{pmatrix}.$$

1.3 Transformée de Fourier discrète

On rappelle que si $f \in L^1(\mathbb{R})$, on peut définie la transformée de Fourier de f par

$$\mathscr{F}(f)(\xi) = \hat{f}(\xi) = \int_{\mathbb{R}} e^{-ix\xi} f(x)\, dx, \quad \xi \in \mathbb{R}, \tag{1.6}$$

où ξ est identifiée comme la fréquence. En dimension supérieure, on a

$$\hat{f}(\boldsymbol{\xi}) = \int_{\mathbb{R}^d} e^{-i\mathbf{x}\cdot\boldsymbol{\xi}} f(\mathbf{x})\, d\mathbf{x}, \quad \boldsymbol{\xi} \in \mathbb{R}^d.$$

Si $\hat{f} \in L^1(\mathbb{R})$, alors on définit la transformée de Fourier inverse

$$\mathscr{F}^{-1}(\hat{f})(x) = \frac{1}{2\pi} \int_{\mathbb{R}} e^{ix\xi} \hat{f}(\xi)\, d\xi, \quad x \in \mathbb{R}. \tag{1.7}$$

Ces définitions peuvent être étendues à des espaces fonctionnels plus généraux (L^2, espace de Schwartz, ...).

Si f est localement absolument continue (ce qui pourrait être traduit par dérivable et égale à « l'intégrale de sa dérivée »), et si $f \in L^1(\mathbb{R})$, $f' \in L^1(\mathbb{R})$, alors

$$\mathscr{F}(f')(\xi) = \hat{f}(\xi) = \int_{\mathbb{R}} e^{-ix\xi} f'(x)\, dx = i\xi \int_{\mathbb{R}} e^{-ix\xi} f(x)\, dx,$$

et nous avons donc la relation donnant la transformée de Fourier d'une dérivée en termes de la transformée de Fourier de la fonction

$$\widehat{f'}(\xi) = i\xi \hat{f}(\xi).$$

En supposant plus de régularité sur f, nous avons aussi

$$\widehat{f''}(\xi) = -\xi^2 \hat{f}(\xi).$$

R Il existe d'autres conventions. On peut par exemple définir la transformée de Fourier par

$$\mathscr{F}(f)(\nu) = \hat{f}(\nu) = \int_{\mathbb{R}} e^{-i2\pi\nu t} f(t)\, dt, \quad \nu \in \mathbb{R}, \tag{1.8}$$

et

$$f(t) = \mathscr{F}^{-1}(\hat{f})(t) = \int_{\mathbb{R}} e^{i2\pi\nu t} \hat{f}(\nu)\, d\nu, \quad t \in \mathbb{R}. \tag{1.9}$$

Dans ce cas, t désigne le temps (en seconde) et ν la fréquence en Hertz. On a maintenant

$$\mathscr{F}(f')(\nu) = i(2\pi\nu)\hat{f}(\nu) \tag{1.10}$$

et

$$\mathscr{F}(f'')(\nu) = -(2\pi\nu)^2 \hat{f}(\nu). \tag{1.11}$$

Nous souhaiterions trouver un outil similaire pour les fonctions discrètes, soit les suites. On rappelle que pour des fonctions μ-périodiques, on peut définit leurs séries de Fourier. Soit

$$L^2_p(0,\mu) = \left\{ f : \mathbb{R} \to \mathbb{C},\ \mu\text{-périodique}\ ,\ \text{telle que}\ \int_0^\mu |f(t)|^2\, dt < \infty \right\}.$$

Considérons le polynôme trigonométrique

$$p(t) = \sum_{n=-N}^{N} c_n e^{\frac{2i\pi nt}{\mu}}, \quad c_n \in \mathbb{C}.$$

Définissons $e_n(t) = e^{\frac{2i\pi nt}{\mu}}$, alors

$$\langle e_n, e_m \rangle_{L^2_p} = \int_0^\mu e_n(t)\overline{e_m}(t)dt = \begin{cases} 0 & \text{if } n \neq m, \\ \mu & \text{if } n = m. \end{cases}$$

Donc, la famille $\{e_n\}_{-N \leqslant n \leqslant N}$ est une base orthogonale de l'ensemble des polynômes trigonométriques de degré $\leqslant N$. Nous pouvons donc calculer les coefficients c_n par produit scalaire de p avec chaque e_n puisque

$$\langle p, e_n \rangle_{L^2_p} = c_n \|e_n\|^2_{L^2_p} = \mu c_n$$

et nous obtenons

$$c_n = \frac{1}{\mu}\langle p, e_n \rangle_{L^2_p} = \frac{1}{\mu}\int_0^\mu p(t)e^{-\frac{2i\pi}{\mu}nt}\,dt = \frac{1}{\mu}\int_{-\mu/2}^{\mu/2} p(t)e^{-\frac{2i\pi}{\mu}nt}\,dt.$$

On se pose maintenant la question de savoir si $f \in L^2_p(0,\mu)$, il est possible de trouver un polynôme trigonométrique optimal p tel que $\|f-p\|_{L^2_p}$ soit minimal ? Soit $p(t) = \sum_{n=-N}^{N} x_n e_n(t)$ et $c_n(f) = \langle f, e_n \rangle / \mu$. Premièrement, rappelons que pour tout $(x,y) \in \mathbb{C}^2$, nous avons

$$|x-y|^2 = |x|^2 + |y|^2 - 2\mathrm{Re}(x\overline{y}).$$

De manière similaire, nous avons

$$\|f-p\|^2_{L^2_p} = \|f\|^2_{L^2_p} + \|p\|^2_{L^2_p} - 2\mathrm{Re}\langle f, p \rangle_{L^2_p}.$$

Le dernier produit scalaire est

$$\langle f, p \rangle_{L^2_p} = \langle f, \sum_{n=-N}^{N} x_n e_n \rangle_{L^2_p} = \sum_{n=-N}^{N} \langle f, x_n e_n \rangle_{L^2_p} = \sum_{n=-N}^{N} \mu c_n(f)\overline{x_n}.$$

Donc

$$\begin{aligned} \|f-p\|^2_{L^2_p} &= \|f\|^2_{L^2_p} + \mu \sum_{n=-N}^{N} |x_n|^2 - 2\mu\mathrm{Re}\left(\sum_{n=-N}^{N} c_n\overline{x_n}\right) \\ &= \|f\|^2_{L^2_p} + \mu \sum_{n=-N}^{N} \left(|x_n|^2 - 2\mathrm{Re}(c_n\overline{x_n})\right) \\ &= \|f\|^2_{L^2_p} + \mu \sum_{n=-N}^{N} \left(|c_n - x_n|^2 - |c_n|^2\right). \end{aligned}$$

La somme $\sum_{n=-N}^{N} \left(|c_n - x_n|^2 - |c_n|^2\right)$ est minimale si $x_n = c_n$. Ainsi,

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

est la meilleure approximation polynomiale trigonométrique de f. On peut prouver que

$$\lim_{N\to\infty} \|f - f_N\|^2_{L^2_p} = 0.$$

Théorème 1.3.1 Soit $f \in C^0(\mathbb{R})$, μ-périodique, f dérivable sur $[0,\mu]$ (à l'exception éventuelle d'un nombre fini de points) et f' continue par morceaux sur $[0,\mu]$ ($f' \in L^2_p(0,\mu)$), alors, la série de Fourier de f

$$S(f) = \sum_{n\in\mathbb{Z}} c_n(f) e_n$$

satisfait

1. La série de Fourier converge et on a $\sum_{n\in\mathbb{Z}} |c_n(f)|^2 < \infty$.
2. La série de Fourier de f' est calculée par dérivation de chaque terme de la série de Fourier de f,
$$S(f') = \sum_{n\in\mathbb{Z}} c_n(f) e'_n = \sum_{n\in\mathbb{Z}} \frac{2i\pi}{\mu} c_n(f) e_n.$$

R

- Si f est paire, alors $c_n(f) = c_{-n}(f)$, $\forall n \in \mathbb{Z}$.
- Si f est impaire, alors $c_n(f) = -c_{-n}(f)$, $\forall n \in \mathbb{Z}$.
- Il est possible d'établir plusieurs formules pour un polynôme trigonométrique p. En effet,

$$\begin{aligned} p(t) &= \sum_{n=-N}^{N} c_n e_n \\ &= c_0 + \sum_{n=-N}^{-1} c_n e_n + \sum_{n=1}^{N} c_n e_n \\ &= c_0 + \sum_{n=1}^{N} c_n e_n + c_{-n} e_{-n} \\ &= c_0 + \sum_{n=1}^{N} c_n \left(\cos\left(\frac{2\pi n}{\mu} t\right) + i \sin\left(\frac{2\pi n}{\mu} t\right)\right) + c_{-n}\left(\cos\left(\frac{2\pi n}{\mu} t\right) - i\sin\left(\frac{2\pi n}{\mu} t\right)\right). \end{aligned}$$

Ainsi, nous avons

$$p(t) = \frac{a_0}{2} + \sum_{n=1}^{N} a_n \cos\left(\frac{2\pi n}{\mu} t\right) + b_n \sin\left(\frac{2\pi n}{\mu} t\right)$$

où $a_0 = 2c_0$, $a_n = c_n + c_{-n}$ et $b_n = i(c_n - c_{-n})$, ou encore

$$a_n = \frac{2}{\mu}\int_0^{\mu} p(t)\cos\left(\frac{2\pi n}{\mu} t\right) dt = \frac{2}{\mu}\int_{-\mu/2}^{\mu/2} p(t)\cos\left(\frac{2\pi n}{\mu} t\right) dt$$

et

$$b_n = \frac{2}{\mu}\int_0^{\mu} p(t)\sin\left(\frac{2\pi n}{\mu} t\right) dt = \frac{2}{\mu}\int_{-\mu/2}^{\mu/2} p(t)\sin\left(\frac{2\pi n}{\mu} t\right) dt.$$

Si p est paire, on peut en conclure que $b_n = 0$, et si p est impaire, $a_n = 0$.

Afin de trouver la version discrète de la transformée de Fourier, nous avons à déterminer une manière de calculer les coefficients $c_n(f)$. Pour faire cela, nous échantillonnons la fonction f sur $[0,\mu)$ avec N points et on définit les valeurs $f_k = f(k\mu/N)$, $k = 0, \cdots, N-1$. Nous voulons calculer numériquement c_n, $n = -N/2, \cdots, N/2 - 1$ où

$$c_n = \frac{1}{\mu}\int_0^{\mu} f(t) e^{-\frac{2i\pi}{\mu} tn} dt.$$

On utilise pour cela la formule des rectangles à gauche et on obtient

$$\begin{aligned} c_n = \frac{1}{\mu}\sum_{p=0}^{N-1}\int_{x_p}^{x_{p+1}} f(t)e^{-\frac{2i\pi}{\mu}tn}dt &\approx \frac{1}{\mu}\sum_{p=0}^{N-1}\frac{\mu}{N}f_p e^{-\frac{2i\pi}{\mu}n\frac{p\mu}{N}} \\ &\approx \frac{1}{N}\sum_{p=0}^{N-1} f_p e^{-\frac{2i\pi}{N}np}. \end{aligned}$$

Nous définissons donc l'approximation de c_n par

$$c_n^N = \frac{1}{N}\sum_{p=0}^{N-1} f_p e^{-\frac{2i\pi}{N}np} = \frac{1}{N}\sum_{p=0}^{N-1} f_p \omega_N^{-np},$$

où $\omega_N = e^{-\frac{2i\pi}{N}}$.

Grâce à la périodicité, nous avons $c_{n+N}^N = c_n^N$, et donc considérer $\left(c_n^N\right)_{n=-N/2}^{N/2-1}$ est équivalent à $\left(c_n^N\right)_{n=0}^{N-1}$.

Définition 1.3.1 On appelle **Transformée de Fourier Discrète** l'opération linéaire

$$\begin{array}{rccc} \mathscr{F}_N : & \mathbb{C}^N & \longrightarrow & \mathbb{C}^N \\ & f = (f_j)_{j=0}^{N-1} & \longmapsto & F = (F_k)_{k=0}^{N-1} \end{array}$$

où

$$F_k = \sum_{j=0}^{N-1} f_j \omega_N^{-jk}.$$

R D'après la construction ci-dessus, la définition devrait être

$$F_k = \frac{1}{N}\sum_{j=0}^{N-1} f_j \omega_N^{-jk}.$$

La convention veut que l'on fasse porter le pré-facteur $1/N$ dans la transformation inverse.

Soit Ω_N la matrice inversible $\Omega_N = \left(\omega_N^{jk}\right)_{0\leqslant j,k\leqslant N-1}$, qui est en fait

$$\Omega_N = \begin{pmatrix} 1 & 1 & 1 & 1 & \cdots & 1 \\ 1 & \omega_N & \omega_N^2 & \omega_N^3 & \cdots & \omega_N^{N-1} \\ 1 & \omega_N^2 & \omega_N^4 & & & \vdots \\ \vdots & \vdots & & & & \vdots \\ 1 & \omega_N^{N-1} & \cdots & \cdots & \cdots & \omega_N^{(N-1)^2} \end{pmatrix}.$$

Proposition 1.3.2 $\mathscr{F}_N$ est un opérateur linéaire bijectif et le vecteur F est donné par

$$F = \overline{\Omega_N} f. \tag{1.12}$$

Définition 1.3.2 On appelle Transformée de Fourier Discrète Inverse l'opération linéaire

$$\begin{array}{rccc} \mathscr{F}_N^{-1}: & \mathbb{C}^N & \longrightarrow & \mathbb{C}^N \\ & F=(F_k)_{k=0}^{N-1} & \longmapsto & f=(f_j)_{j=0}^{N-1} \end{array}$$

où

$$f_j = \frac{1}{N}\sum_{k=0}^{N-1} F_k \omega_N^{jk}, \quad 0 \leqslant j \leqslant N-1,$$

d'où $f = \frac{1}{N}\Omega_N F$.

La complexité du calcul de la Transformée de Fourier discrète par (1.12) est $O(N^2)$ ce qui est très coûteux. On préfère utiliser la fameuse transformée de Fourier rapide - Fast Fourier Transform (FFT) - et sa transformation inverse (IFFT) qui ont chacune une complexité de $O(N \log N)$. Nous avertissons le lecteur que traditionnellement, en utilisant la périodicité, l'algorithme de la (FFT) stocke les coefficients dans l'ordre suivant

$$(F_k)_{k=0}^{N-1} = \left(F_0, \cdots, F_{N-1}, F_{-N/2}, \cdots, F_{-1}\right),$$

ce qui peut être une source de confusion.

Il est intéressant de noter les différents parallèles respectivement entre la transformée de Fourier (continue), les séries de Fourier (périodiques) et la transformée de Fourier discrète

continu	continue, périodique	discret, périodique
$\mathscr{F}(f)=\hat{f}$	$c_n(f)$	$\mathscr{F}_N(f) = (F_k)_{0\leqslant k\leqslant N-1}$
$f \in L^2(\mathbb{R})$	f μ-périodique	$f=(f_j)_{0\leqslant j\leqslant N-1}$

Réciproquement, on a les transformées inverses

continu	continue, périodique	discret, périodique
$f=\mathscr{F}^{-1}(f)$	$S(f)=\sum_n c_n(f)e^{i2\pi nt/\mu}$	$f=\mathscr{F}_N^{-1}\left((F_k)_{0\leqslant k\leqslant N-1}\right)$

Concernant la dérivation, on peut faire un parallèle identique et on a en utilisant la deuxième formulation de la transformée de Fourier continue (1.8)

continu	continue, périodique	discret, périodique
$f'=\mathscr{F}^{-1}(i2\pi\nu\hat{f})$	$S(f')=\displaystyle\sum_n \frac{i2\pi n}{\mu}c_n(f)e^{i2\pi nt/\mu}$	"f'" $=\mathscr{F}_N^{-1}\left(\frac{i2\pi n}{\mu}(F_k)_{0\leqslant k\leqslant N-1}\right)$

Afin d'utiliser l'efficacité de l'algorithme de la (FFT) pour l'approximation des opérateurs différentiels, nous nous reposons sur les résultats connus pour les séries de Fourier.

■ **Exemple 1.1** On veut calculer l'approximation de la dérivée première de la fonction f. On rappelle que

$$S(f') = \sum_{n\in\mathbb{Z}} \frac{2i\pi}{\mu} c_n(f) e_n.$$

L'approximation c_n^N de c_n est calculée par `fft(f)`. On construit le multiplicateur fréquentiel $(2i\pi n)/\mu$, puis on multiplie terme à terme, définissant $d_n = (2i\pi n)/\mu\ c_n$ et on forme la somme $\sum d_n e_n$ qui est donnée par la transformée de Fourier discrète. Avec Matlab, l'algorithme est

```
k=(2*pi/mu)*[0:N/2-1,-N/2:-1]';
c=fft(f);
Df=ifft(c.*k);
```

Pour la dérivée seconde f'', l'opération est aussi extrêmement simple et l'implémentation Matlab est

```
D2f=ifft((-k.^2).*c);
```

■

Il y a deux intérêts majeurs à l'utilisation de la transformée de Fourier discrète.

1. Nous disposons de la transformée de Fourier rapide.
2. Soit $f \in C^0$, 2π-périodique, et $I_N f$ le polynôme trigonométrique avec N modes de Fourier qui interpole f aux N points équidistants $x_j = j2\pi/N$ défini par

$$I_N f(x) = \sum_{j=-N/2}^{N/2-1} c_j e^{ijx}, \quad (c_j) = F_N(f(x_j)).$$

Alors, on peut prouver le théorème suivant (voir [9], page 77)

Théorème 1.3.3 Supposons que la dérivée d'ordre s $\partial_x^s f \in L^2$, pour $s \geqslant 1$. Alors, l'erreur d'interpolation est bornée dans L^2 par

$$\|f - I_N f\| \leqslant CN^{-s}\|\partial_x^s f\|,$$

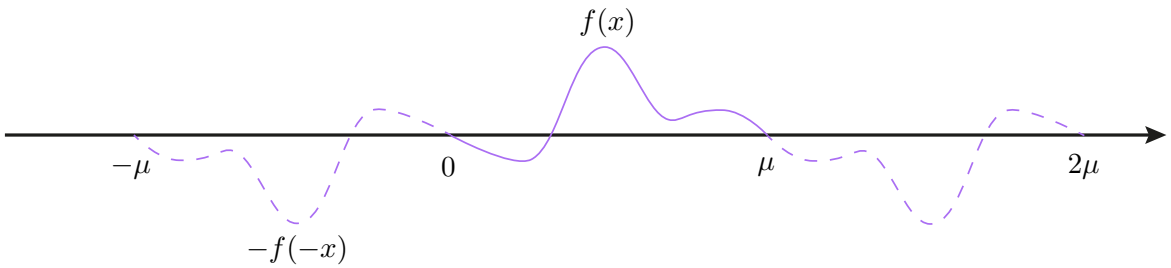
où $C = C(s)$ seulement. On peut aussi montrer que sous des hypothèses plus fortes de régularité

$$\|\partial_x^m (f - I_N f)\| \leqslant CN^{-s}\|\partial_x^{s+m} f\|.$$

Donc, comparée à la méthode des différences finies, la méthode spectrale basée sur Fourier est bien meilleure.

1.4 La Transformée en Sinus Discrète

La Transformée de Fourier Discrète précédente nécessite que f soit périodique ou que l'on considère des conditions aux limites périodiques. Si on s'intéresse à des conditions aux limites de Dirichlet homogènes, on doit travailler avec des fonctions qui sont toujours périodiques mais nulles au bord. On choisit d'étendre la fonction f comme une fonction impaire, ce qui permet à la dérivée de f d'être continue (voir la figure ci-dessous).



La nouvelle fonction g est 2μ-périodique. On peut donc construire sa série de Fourier

$$g_N(t) = \frac{a_0}{2} + \sum_{n=1}^{N} a_n \cos\left(\frac{2\pi n}{2\mu}t\right) + b_n \sin\left(\frac{2\pi n}{2\mu}t\right).$$

Puisque g est impaire, $a_n = 0$, $n = 0, \cdots, N-1$, d'où

$$g_N(t) = \sum_{n=1}^{N} b_n \sin\left(\frac{\pi n}{\mu}t\right)$$

et la série de Fourier de g est

$$S(g) = \sum_{n=1}^{\infty} b_n \sin\left(\frac{\pi n}{\mu} t\right),$$

avec

$$b_n = \frac{2}{2\mu}\int_0^{2\mu} g(t)\sin\left(\frac{\pi n}{\mu}t\right)\,dt = \frac{1}{\mu}\int_0^{2\mu} g(t)\sin\left(\frac{\pi n}{\mu}t\right)\,dt = \frac{1}{\mu}\int_{-\mu}^{\mu} g(t)\sin\left(\frac{\pi n}{\mu}t\right)\,dt.$$

Donc,

$$\begin{aligned} b_n &= \frac{1}{\mu}\left[\int_0^{\mu} g(t)\sin\left(\frac{\pi n}{\mu}t\right)\,dt + \int_{-\mu}^{0} g(t)\sin\left(\frac{\pi n}{\mu}t\right)\,dt\right] \\ &= \frac{1}{\mu}\left[\int_0^{\mu} g(t)\sin\left(\frac{\pi n}{\mu}t\right)\,dt + \int_0^{\mu} g(-s)\sin\left(\frac{\pi n}{\mu}(-s)\right)\,ds\right]. \end{aligned}$$

Puisque g et sin sont deux fonctions impaires, on obtient

$$b_n = \frac{2}{\mu}\int_0^{\mu} g(t)\sin\left(\frac{\pi n}{\mu}t\right)\,dt.$$

Cette transformation s'appelle transformée en sinus discrète . Nous avons

$$S(f') = \sum_{n=1}^{\infty} \frac{\pi n}{\mu} b_n \cos\left(\frac{\pi n}{\mu}t\right),$$

et

$$S(f'') = -\sum_{n=1}^{\infty} \left(\frac{\pi n}{\mu}\right)^2 b_n \sin\left(\frac{\pi n}{\mu}t\right).$$

Définition 1.4.1 On appelle **Transformée en Sinus Discrète** l'opération linéaire

$$\begin{array}{rccc} \mathscr{F}_N : & \mathbb{C}^N & \longrightarrow & \mathbb{C}^N \\ & f = (f_j)_{j=0}^{N-1} & \longmapsto & F = (F_k)_{k=0}^{N-1} \end{array}$$

où

$$F_k = \frac{2}{N}\sum_{j=1}^{N-1} f_j \sin\left(\frac{\pi jk}{N}\right), \quad 1 \leqslant k \leqslant N-1$$

et la transformée inverse est donnée par

$$f_j = \sum_{k=1}^{N-1} F_k \sin\left(\frac{\pi jk}{N}\right), \quad 1 \leqslant j \leqslant N-1.$$

R Comme pour la transformée de Fourier discrète, la convention est usuellement

$$F_k = \sum_{j=1}^{N-1} f_j \sin\left(\frac{\pi jk}{N}\right) \text{ et } f_j = \frac{2}{N}\sum_{k=1}^{N-1} F_k \sin\left(\frac{\pi jk}{N}\right).$$

On peut calculer de manière efficace la transformée en sinus discrète à partir de la FFT. On suit la construction de la fonction g au niveau discret. Nous construisons la suite impaire à $(2N-1)$-points

$$g_j = \begin{cases} f_j, & 0 < j < N, \\ 0, & j = 0, N \\ -f_{2N-j}, & N < j \leqslant 2N-1. \end{cases}$$

Calculons la Transformée de Fourier Discrète de $(g_j)_{0\leqslant j\leqslant 2N-1}$. Soit $0 \leqslant k \leqslant 2N-1$, alors

$$G_k = \sum_{j=0}^{2N-1} g_j e^{-\frac{2i\pi}{2N}jk} = \sum_{j=1}^{N-1} g_j e^{-\frac{i\pi}{N}jk} + \sum_{j=N+1}^{2N-1} g_j e^{-\frac{i\pi}{N}jk},$$

soit encore

$$G_k = \sum_{j=1}^{N-1} f_j e^{-\frac{i\pi}{N}jk} - \sum_{j=N+1}^{2N-1} f_{2N-j} e^{-\frac{i\pi}{N}jk}.$$

Nous appliquons le changement de variable $\ell = 2N - j$ dans la deuxième somme ce qui donne

$$G_k = \sum_{j=1}^{N-1} f_j e^{-\frac{i\pi}{N}jk} - \sum_{\ell=1}^{N-1} f_\ell e^{-i2\pi k} e^{\frac{i\pi}{N}\ell k},$$

ou encore

$$G_k = \sum_{j=1}^{N-1} f_j (e^{-\frac{i\pi}{N}jk} - e^{-i2\pi k} e^{\frac{i\pi}{N}jk}).$$

Nous obtenons donc

$$G_k = -2i \sum_{j=1}^{N-1} f_j \sin\left(\frac{\pi jk}{N}\right),$$

qui est la Transformée en Sinus Discrète de f, notée par $\mathrm{DST}(f_j)$. D'où,

$$\left(\mathrm{DST}(f_j)\right)_k = \frac{\left(\mathrm{FFT}(g_j)\right)_k}{-2i}, \quad 1 \leqslant k \leqslant N-1.$$

Pour des conditions aux limites de Neumann homogènes, on utiliserait la transformée en cosinus discrète.

1.5 Volumes finis

Considérons l'équation mono dimensionnelle à coefficients variables

$$\begin{cases} -\partial_x(k(x)\partial_x u) + u = s(x) \quad , \quad x \in]0,1[, \\ \partial_x u(0) = \partial_x u(1) = 0, \end{cases} \tag{1.13}$$

avec $k(x) > 1$ pour tout $x \in]0,1[$. On discrétise le domaine spatial en N cellules (mailles ou encore volumes de contrôle) de taille identique $h = 1/N$ et on définit $x_j = h/2 + jh$, de telle sorte que les points x_j sont au centre des mailles. Les côtés de la maille j sont $x_{j-1/2}$ et $x_{j+1/2}$.

Dans la méthode des volumes finis, les inconnues approchent la moyenne de la solution sur une cellule. Plus précisément, on définit q_j comme étant l'approximation suivante sur la maille $I_j = [x_{j-1/2}, x_{j+1/2}]$

$$q_j \approx u_j := \frac{1}{h} \int_{x_{j-1/2}}^{x_{j+1/2}} u(x)\, dx.$$

On intègre (1.13) sur la cellule I_j

$$-\int_{x_{j-1/2}}^{x_{j+1/2}} \partial_x(k(x)\partial_x u(x))\,dx + \int_{x_{j-1/2}}^{x_{j+1/2}} u(x)\,dx = \int_{x_{j-1/2}}^{x_{j+1/2}} S(x)\,dx$$

soit encore

$$-\Big[k(x)\partial_x\Big]_{x_{j-1/2}}^{x_{j+1/2}} + hu_j = \int_{x_{j-1/2}}^{x_{j+1/2}} S(x)\,dx := jS_j.$$

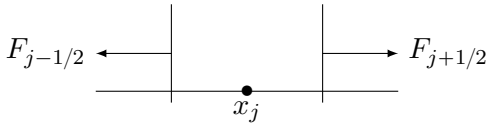
En développant, on obtient

$$hu_j - \big(k(x_{j+1/2})\partial_x u x_{j+1/2} - k(x_{j-1/2})\partial_x x_{j-1/2}\big) = hS_j.$$

On définit le flux $F_j = F(x_j) = -k(x_j)\partial_x u(x_j)$. On obtient ainsi

$$F_{j+1/2} - F_{j-1/2} + hu_j = hS_j. \tag{1.14}$$

Bien que dire que les données $F_{j\pm1/2}$ sont des flux est un peu abusif en dimension 1, on interpréte $F_{j+1/2}$ et $F_{j-1/2}$ comme étant les flux de chaleur à travers les frontières droite et gauche de la cellule I_j



Il nous reste à formuler une approximation des flux pour obtenir une méthode numérique utilisable. On sait par la formule de quadrature du point milieu que

$$\int_a^b f(t)\,dt = (b-a)f\left(\frac{a+b}{2}\right) + \mathcal{O}(|b-a|^3),$$

ainsi

$$f\left(\frac{a+b}{2}\right) = \frac{1}{b-a}\int_a^b f(t)\,dt + \mathcal{O}(|b-a|^2). \tag{1.15}$$

Ainsi, la valeur au milieu de la cellule I_j est une approximation à l'ordre 2 de la moyenne. Nous avons donc pour u régulière

$$F_{j-1/2} = -k(x_{j-1/2})\partial_x u(x_{j-1/2}) = -k(x_{j-1/2})\frac{u(x_j) - u(x_{j-1})}{h} + \mathcal{O}(h^2)$$

et

$$u(x_j) = \frac{1}{h}\int_{x_{j-1/2}}^{x_{j+1/2}} u(x)\,dx + \mathcal{O}(h^2) := u_j + \mathcal{O}(h^2)$$

d'où

$$F_{j-1/2} = -k(x_{j-1/2})\frac{u_j - u_{j-1}}{h} + \mathcal{O}(h) \approx \tilde{F}_{j-1/2} = -k(x_{j-1/2})\frac{q_j - q_{j-1}}{h}.$$

On obtient ainsi un schéma numérique pour les points intérieurs $1 \leqslant j \leqslant N-2$ en approchant (1.14) qui devient

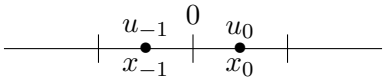
$$-\left(k(x_{j+1/2})\frac{q_{j+1} - q_j}{h} - k(x_{j-1/2})\frac{q_j - q_{j-1}}{h}\right) + hq_j = hS_j$$

soit

$$-\frac{k_{j+1/2}q_{j+1} - (k_{j+1/2} + k_{j-1/2})q_j + k_{j-1/2}q_{j-1}}{h^2} + q_j = S_j, \quad 1 \leqslant j \leqslant N-2. \tag{1.16}$$

L'équation précédente ressemble à une approximation différence finies. Ce comportement serait différent pour un pas de maillage non uniforme.

Il reste à prendre en compte les conditions aux limites. En effet, si on considère $j = 0$ ou $j = N - 1$ dans (1.16), les termes q_{-1} et q_N sont inconnus. On utilise donc les conditions aux limites. Pour cela, on introduit des cellules fictives (ou fantômes) I_{-1} et I_N. La condition aux limites en $x = 0$ est $\partial_x u = 0$ (traduisant un flux nul).



On étend formellement la définition de la solution u pour $x < 0$:

$$0 = \partial_x u(0) = \frac{u(x_0) - u(x_{-1})}{h} + \mathcal{O}(h^2).$$

Par (1.15), on a

$$u(x_j) = u_j + \mathcal{O}(h^2).$$

Alors,

$$0 = \partial_x u(0) = \frac{u_0 - u_{-1}}{h} + \mathcal{O}(h)$$

et donc

$$u_{-1} = u_0 + \mathcal{O}(h^2).$$

En terme d'approximation de u_j, cela donne $q_{-1} = q_0$. En remplaçant pour $j = 0$ dans (1.16), on obtient

$$q_0 - k_{1/2}\frac{q_1 - q_0}{h^2} = S_0.$$

On a de même pour la maille extrême à droite du domaine

$$q_{N-1} - k_{N-3/2}\frac{q_{N-2} - q_{N-1}}{h^2} = S_{N-1}.$$

La première version du schéma volume fini pour résoudre l'équation (1.13) est donné par :

Trouver $\mathbf{q} = (q_j)_{0 \leqslant j \leqslant N-1} \in \mathbb{R}^N$ tel que

$$\begin{aligned} &hq_j - k_{j+1/2}\tfrac{q_{j+1}-q_j}{h} + k_{j-1/2}\tfrac{q_j - q_{j-1}}{h} = hS_j, \quad 1 \leqslant j \leqslant N-1, \\ &hq_0 - k_{1/2}\tfrac{q_1 - q_0}{h} = hS_0, \\ &hq_{N-1} + k_{N-3/2}\tfrac{q_{N-1}-q_{N-2}}{h} = hS_{N-1}. \end{aligned}$$

On peut écrire une deuxième version en terme de flux :

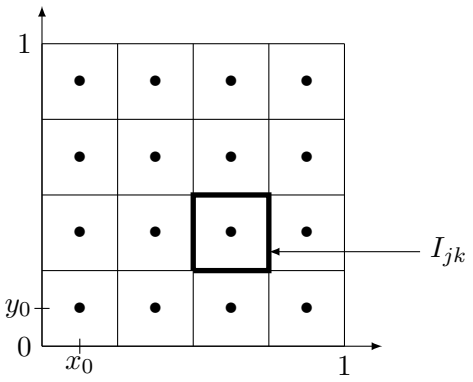
Trouver $\mathbf{q} = (q_j)_{0 \leqslant j \leqslant N-1} \in \mathbb{R}^N$ tel que

$$\begin{aligned} &hq_j - F_{j+1/2}(q) - F_{j-1/2}(q) = hS_j, \quad 1 \leqslant j \leqslant N-1, \\ &hq_0 + F_{1/2}(q) = hS_0, \\ &hq_{N-1} - F_{N-3/2}(q) = hS_{N-1}. \end{aligned}$$

L'extension à la dimension 2 est relativement aisée. On s'intéresse au problème

$$\begin{cases} -\Delta u + u = S(x,y), & (x,y) =]0,1[^2, \\ \partial_n u = 0, & (x,y) \in \Gamma. \end{cases} \tag{1.17}$$

On maille le carré $[0,1]^2$ avec $N \times N$ cellules de taille $h \times h$, $h = 1/N$. Les points milieu (x_j, y_k) des cellules I_{jk} sont définis par $x_j = h/2 + jh$ et $y_k = h/2 + kh$. La surface d'une cellule est donc h^2.



L'inconnue est

$$q_{jk} \approx u_{jk} := \frac{1}{h^2} \int_{I_{jk}} u(x,y)\, dxdy.$$

L'intégrale de (1.17) sur la cellule de I_{jk} donne

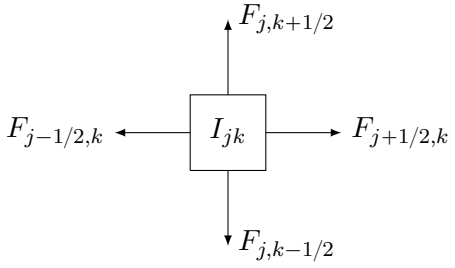
$$\int_{I_{jk}} u - \Delta u dx = \int_{I_{jk}} S\, dxdy.$$

Après formule de Green, on obtient

$$\int_{\partial I_{jk}} -\nabla u \cdot \mathbf{n}\, d\sigma + \int_{I_{jk}} u\, dxdy \int_{I_{jk}} S\, dxdy = h^2 f_{jk}. \tag{1.18}$$

Notons $-\nabla u = (f_1, f_2)^T$. Soit

$$F_{j,k\pm1/2} = \int_{x_{j-1/2}}^{x_{j+1/2}} f_2(x, y_{k\pm1/2})\, dx, \qquad F_{j\pm1/2,k} = \int_{y_{k-1/2}}^{y_{k+1/2}} f_2(x_{j\pm1/2}, y)\, dy.$$



L'équation (1.18) devient

$$hu_{jk} - \left(F_{j+1/2,k} - F_{j-1/2,k} + F_{j,k+1/2} - F_{j,k-1/2}\right) = h^2 S_{jk}.$$

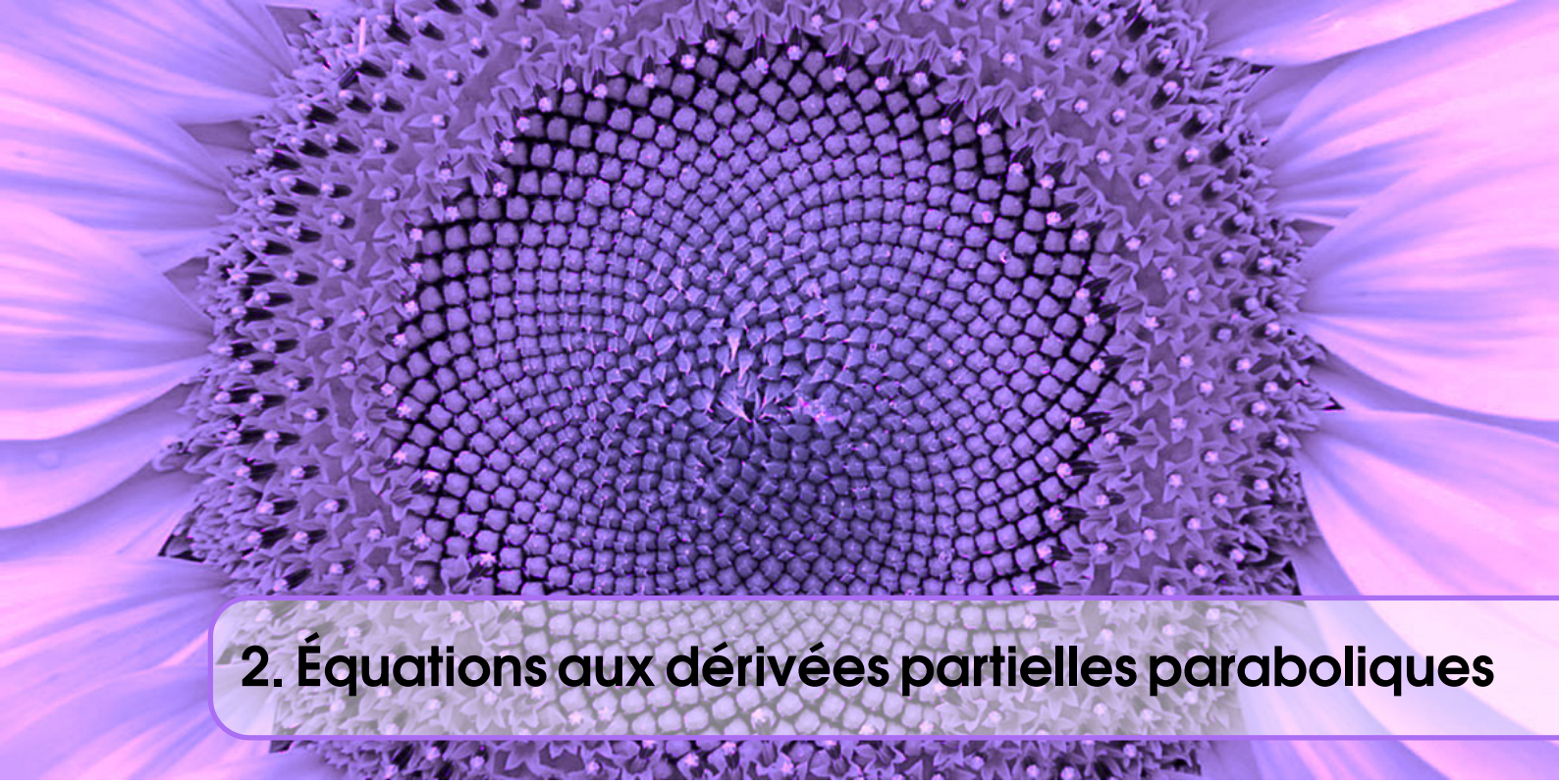
Il reste comme en dimension 1 à procéder à une approximation des flux. Comme en 1D, on a par exemple

$$\begin{aligned}
F_{j,k-1/2} &= -\int_{x_{j-1/2}}^{x_{j+1/2}} \partial_y u(x, y_{k-1/2})\, dx = -\int_{x_{j-1/2}}^{x_{j+1/2}} \frac{u(x,y_k) - u(x,y_{k-1})}{h}\, dx + \mathcal{O}(h^2) \\
&= -\frac{1}{h}\int_{x_{j-1/2}}^{x_{j+1/2}} u(x,y_k)\, dx + \frac{1}{h}\int_{x_{j-1/2}}^{x_{j+1/2}} u(x,y_{k-1})\, dx + \mathcal{O}(h^2) \\
&= -\frac{1}{h^2}\int_{I_{jk}} u(x,y)\, dy + \frac{1}{h^2}\int_{I_{j,k-1}} u(x,y)\, dy + \mathcal{O}(h) \\
&= -u_{j,k} + u_{j,k-1} + \mathcal{O}(h) \\
&\approx -q_{j,k} + q_{j,k-1}.
\end{aligned}$$

Le schéma volume fini 2D pour approcher (1.17) devient pour les mailles internes

$$q_{j+1,k} + q_{j-1,k} + q_{j,k+1} + q_{j,k-1} - (4-h)q_{j,k} = h^2 S_{j,k}.$$

Il reste à prendre en compte les conditions aux limites et on fait des opérations similaires à la dimension 1.



2. Équations aux dérivées partielles paraboliques

Avant de proposer la construction de schémas numériques adaptés à une équation aux dérivées partielles particulière, il est important de s'assurer de son caractère bien posé. On dit qu'un problème est bien posé si

1. le problème a une solution,
2. la solution est unique,
3. la solution dépend continûment des données du problème,

C'est un préalable nécessaire sans quoi un schéma numérique serait inutile, conduisant à des solutions sans signification.

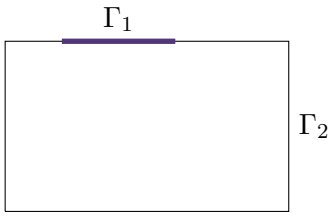
2.1 Problème modèle

L'équation aux dérivées partielles linéaires parabolique standard est l'équation de la chaleur

$$\partial_t u(t,x) - \Delta u(t,x) = f,$$

à laquelle il convient d'ajouter des conditions initiales et des conditions aux limites appropriées si nécessaire. Ici, t est une quantité positive désignant le temps, $x = (x_1, \cdots, x_N)$ est la variable d'espace avec $x \in \Omega$, $\Omega \subset \mathbb{R}^N$ ou bien $x \in \mathbb{R}^N$, et f est un terme source à valeurs réelles.

Un exemple simple d'utilisation de l'équation de la chaleur intervient dans la simulation de l'évolution de la température dans une pièce chauffée par un radiateur. Prenons une pièce idéalisée



La pièce est chauffée par un radiateur situé à la frontière sur Γ_1. La partie restante de la frontière est notée Γ_2, de sorte que $\partial\Omega = \Gamma_1 \cup \Gamma_2$.

Deux situations peuvent être envisagées. La pièce est mal isolée et on impose sur Γ_2 la température extérieure. On a donc les conditions aux limites suivantes

$$u(x) = \begin{cases} u_{\text{rad}} \text{ sur } \Gamma_1, \\ u_{\text{ext}} \text{ sur } \Gamma_2. \end{cases}$$

Une autre situation serait de considérer une pièce isolée, ce qui se traduit par un flux de chaleur nul sur Γ_2. Si on définit $\mathbf{n}$ comme étant la normale unitaire sortante à Ω, alors les conditions aux limites sont

$$u(x) = \begin{cases} u_{\text{rad}} \text{ sur } \Gamma_1, \\ \nabla u \cdot \mathbf{n} = 0 \text{ sur } \Gamma_2. \end{cases}$$

Il faut compléter les conditions aux limites par une donnée initiale telle que $u(x,0) = u^i(x)$.

2.2 Solutions classiques dans $\mathbb{R}^N$

On appelle solutions classiques des solutions $u \in C^2(\mathbb{R}_+^* \times \mathbb{R}^N, \mathbb{R}) \cap C(\mathbb{R}_+ \times \mathbb{R}^N, \mathbb{R})$ qui sont à distinguer des solutions faibles (avec moins de régularité). On suppose que la donnée initiale $u^i \in C(\mathbb{R}^N, \mathbb{R})$ et on cherche des solutions classiques de

$$\begin{cases} \partial_t u - \Delta u = 0, \quad x \in \mathbb{R}^N,\ t > 0, \\ u(0,x) = u^i(x), \quad x \in \mathbb{R}^N. \end{cases} \tag{2.1}$$

Commençons par un calcul formel, qu'il convient de justifier par la suite. On suppose que l'on peut appliquer la transformée de Fourier à u et u^i. On prend pour cela la convention suivante. Pour $f \in L^1(\mathbb{R}^N)$, la transformée de Fourier de f est définie par

$$\hat{f}(\xi) = \frac{1}{(2\pi)^{N/2}} \int_{\mathbb{R}^N} e^{-i\langle x,\xi\rangle} f(x)\, dx.$$

On rappelle que si $f \in L^1(\mathbb{R}^N)$, alors $\hat{f}$ est continue sur $\mathbb{R}^N$ et bornée. La transformée de Fourier ci-dessus s'étend à $L^2(\mathbb{R}^N)$ par densité et elle laisse invariant l'espace de Schwartz $\mathscr{S}(\mathbb{R}^N) = \{u \in C^\infty,\ \forall \alpha,\ \forall \beta \in \mathbb{N}^N,\ \lim_{|x|\to\infty} x^\alpha D^\beta u = 0\}$. Supposons que u soit solution de (2.1) et ait la régularité suffisante pour lui appliquer la transformée de Fourier. On a alors

$$\begin{cases} \partial_t \hat{u}(t,\xi) + |\xi|^2 \hat{u}(t,\xi) = 0, \quad \xi \in \mathbb{R}^N,\ t > 0, \\ \hat{u}(0,\xi) = \widehat{u^i}(\xi), \quad \xi \in \mathbb{R}^N. \end{cases}$$

La solution de cette équation différentielle ordinaire est

$$\hat{u}(t,\xi) = e^{-|\xi|^2 t} \widehat{u^i}(\xi), \quad \xi \in \mathbb{R}^N,\ t \geqslant 0.$$

On constate donc que les hautes fréquences éventuellement présentes dans u^i sont amorties. Si on change le signe devant l'opérateur Laplacien, elles seraient augmentées et on conclut que le système serait mal posé.

Cherchons maintenant à construire $u(t,x)$ à partir de cette relation. Soit $g(t) \in L^1(\mathbb{R}^N)$ telle que

$$\widehat{g(t)}(\xi) = e^{-|\xi|^2 t}.$$

On a donc

$$\hat{h}(t,\cdot) = \widehat{g(t)}\widehat{u^i} = \widehat{g(t) * u^i}, \quad t \geqslant 0.$$

Ainsi,

$$u(t,\cdot) = (2\pi)^{N/2} g(t) * u^i, \quad t \geqslant 0.$$

Il nous reste donc à calculer $g(t)$. Comme $\widehat{g(t)} \in L^1(\mathbb{R}^N)$, on peut appliquer la transformée de Fourier inverse

$$g(t)(x) = \frac{1}{(2\pi)^{N/2}} \int_{\mathbb{R}^N} e^{i\langle x,\xi\rangle} e^{-|\xi|^2 t}\, d\xi.$$

On fait le changement de variable $\xi = \eta/\sqrt{2t}$. Ainsi

$$g(t)(x) = \frac{1}{(2\pi)^{N/2}} \frac{1}{(2t)^{N/2}} \int_{\mathbb{R}^N} e^{i\langle x,\eta/\sqrt{2t}\rangle} e^{-|\eta|^2/2}\, d\eta.$$

Or, on sait que $\widehat{e^{-|y|^2/2}}(\xi) = e^{-|\xi|^2/2}$. On a de ce fait

$$g(t)(x) = \frac{1}{(2t)^{N/2}} \frac{1}{(2\pi)^{N/2}} \int_{\mathbb{R}^N} e^{i\langle x/\sqrt{2t},\eta\rangle} e^{-|\eta|^2/2}\, d\eta = \frac{1}{(2t)^{N/2}} \mathscr{F}^{-1}\left(e^{-|\eta|^2/2}\right)\left(\frac{x}{\sqrt{2t}}\right).$$

On a donc

$$g(t)(x) = \frac{1}{(2t)^{N/2}} e^{-|x|^2/4t}.$$

Au final,

$$u(t,x) = \frac{1}{(4\pi t)^{N/2}} \int_{\mathbb{R}^N} e^{-|x-y|^2/4t} u^i(y)\, dy, \quad x \in \mathbb{R}^N,\ t > 0.$$

Donnons maintenant des conditions sur u^i pour que u soit solution classique. Les deux cas de figures suivants sont favorables

- $u^i \in \left(L^1(\mathbb{R}^N) + L^\infty(\mathbb{R}^N)\right) \cap C(\mathbb{R}^N, \mathbb{R})$.
- $u^i \in C(\mathbb{R}^N, \mathbb{R})$ et il existe une constante $C \in \mathbb{R}$ et un entier $p \in \mathbb{N}$ tels que $|u^i(x)| \leqslant C(1 + |x|^p)$.

Il reste à s'assurer de l'unicité de la solution construite. Sans hypothèse supplémentaire que $u^i \in C(\mathbb{R}^N, \mathbb{R})$, on peut montrer qu'il existe plusieurs solution au problème

$$\begin{cases} \partial_t u - \Delta u = 0, & x \in \mathbb{R}^N,\ t > 0, \\ u(0,x) = 0, & x \in \mathbb{R}^N. \end{cases}$$

En dimension un on a par exemple $u(t,x) = \displaystyle\sum_{k=0}^{\infty} \frac{1}{2k!} x^{2k} \frac{d^k}{dt^k} e^{-1/t^2}$. On est donc amener à rajouter des contraintes de croissance sur la solution.

Théorème 2.2.1 Si $u^i \in C(\mathbb{R}^N, \mathbb{R})$ et si il existe une constante $C \in \mathbb{R}$ et un entier $p \in \mathbb{N}$ tels que $|u^i(x)| \leqslant C(1 + |x|^p)$, alors il existe un unique u vérifiant

1. u est une solution classique de (2.1)
2. Pour tout $T > 0$, il existe une constante $C = C(T) \in \mathbb{R}$ et $P = P(T) \in \mathbb{N}$ telles que

$$|u(t,x)| \leqslant C(1 + |x|^P), \quad x \in \mathbb{R}^N,\ t \in [0,T].$$

R

1. On peut montrer que $u(t,\cdot) \in C(\mathbb{R}^N)$.
2. On peut associer à la solution une structure de semi-groupe. On a vu que

$$u(t,x) = \frac{1}{(2\pi)^{N/2}} g(t) * u^i := G(t)u^i.$$

On a $G(0) = Id$ et on peut vérifier que $G(t+s) = G(t)G(s)$ sous la condition que $s \geqslant 0$ et $t \geqslant 0$. La suite d'opérations suivantes est autorisée

$$u^i \xrightarrow{G(t)} u(t) \xrightarrow{G(t')} u(t+t'), \quad t \geqslant 0,\ t' \geqslant 0.$$

ce qui caractérise la structure de semi-groupe. En revanche, $G(t)$ n'agit pas continûment sur les espaces fonctionnels pour $t < 0$. On ne peut pas inverser le sens du temps, ce qui traduit le caractère mal posé pour les temps négatifs.

3. Si on travaille dans l'espace des distributions $\mathscr{D}'$, et si $u^i = \delta$, alors, pour tout temps $t > 0$, aussi petit que l'on veut, $u(t, \cdot) > 0$ en des points aussi éloignés de l'origine que l'on veut. Ceci est caractéristique d'une propagation à vitesse infinie.

Nous n'avons pour l'instant fourni de solutions que pour l'équation de la chaleur (2.1) homogène. Intéressons nous à l'équation inhomogène

$$\begin{cases} \partial_t u - \Delta u = f(t,x), & x \in \mathbb{R}^N,\ t > 0, \\ u(0,x) = u^i(x), & x \in \mathbb{R}^N. \end{cases} \tag{2.2}$$

On sait que pour les équations différentielles ordinaires $u' + au = f(t)$, avec a constante et $u(0) = u^i$, on a par formule de variation de la constante

$$u(t) = e^{-at}u^i + \int_0^t e^{-a(t-s)}\,ds.$$

Si on écrit $G(t) = e^{-at}$, on obtient la formule de Duhamel

$$u(t) = G(t)u^i + \int_0^t G(t-s)f(s)\,ds.$$

La propriété de semi-groupe pour l'équation de la chaleur permet de mimer ce comportement et on a

$$u(t,x)G(t)u^i + \int_0^t G(t-s)f(s)\,ds, \quad G(t) = \frac{1}{(2\pi)^{N/2}}g(t) * \cdot$$

2.3 Équation de la chaleur dans un domaine borné

Le matériel présenté ici est très fortement inspiré des notes de cours de F. Lagoutière [8]. On considère l'équation de la chaleur sur $[0,1] \in \mathbb{R}$ avec des conditions aux limites de Dirichlet homogènes

$$\begin{cases} \partial_t u - \kappa\partial_x^2 u = f(t,x), & x \in]0,1[,\ t \in]0,T[, \\ u(t,0) = u(1,t) = 0, & t \in]0,T[, \\ u(0,x) = u^i(x), & x \in]0,1[. \end{cases} \tag{2.3}$$

On suppose que $\kappa > 0$ et que $f \in C^0([0,1] \times [0,T])$. On souhaite reproduire la preuve vue pour $\mathbb{R}^N$ tout entier qui utilisait la transformée de Fourier. On remplace la transformée de Fourier par des séries de Fourier. Plus précisément, comme on considère des conditions aux limites de Dirichlet homogènes, il faut utiliser des séries de sinus. On se rend en outre compte que pour tout $k \in \mathbb{N}$, les fonctions $e^{-\kappa k^2\pi^2 t}\sin(k\pi x)$ sont solutions de (2.3), en oubliant pour un moment la condition initiale.

Proposition 2.3.1 La famille $\{\sqrt{2}\sin(k\pi\cdot)\}_{k\in\mathbb{N}*}$ est une base hilbertienne de $L^2(]0,1[)$

$$f(x) = \sqrt{2}\sum_{k\in\mathbb{N}*}\hat{f}(k)\sin(k\pi x), \quad \text{dans } L^2(]0,1[),$$

où

$$\hat{f}(k) = \sqrt{2}\int_0^1 f(x)\sin(k\pi x)\,dx, \quad k \in \mathbb{N}^*.$$

Comme on l'a vu dans le chapitre précédent, si $g \in C^1([0,1]) \cap C^2(]0,1[)$, avec $g'' \in L^2(]0,1[)$ telle que $g(0) = g(1) = 0$, on a

$$\widehat{g''}(k) = -k^2\pi^2\hat{g}(k), \quad k \in \mathbb{N}^*.$$

2.3.1 Unicité de la solution - Stabilité

Supposons que u soit solution de (2.3) avec $u \in C^1(]0,T[; C^2(]0,1[)$ et $u(t,0) = u(t,1) = 1$. On suppose que $f(t,\cdot)$, $u(t,\cdot)$, $\partial_t u(t,\cdot)$ et $\partial_x^2 u(t,\cdot)$ sont développables sur la base Hilbertienne pour tout $t \in]0,T]$. D'après le théorème de Lebesgue de dérivation sous le signe intégral, $\widehat{u(t,\cdot)}(k)$ est dérivable par rapport à t, $t > 0$ pour tout $k \in \mathbb{N}^*$ et

$$\partial_t(\widehat{u(t,\cdot)})(k) = \sqrt{2}\int_0^1 \partial_t u(t,x)\sin(k\pi x)\,dx = \widehat{\partial_t u(t,\cdot)}(k).$$

Donc,

$$\partial_t u(t,x) = \sqrt{2}\sum_{k\in\mathbb{N}^*} \widehat{\partial_t u(t,\cdot)}(k) sin(k\pi x) = \sqrt{2}\sum_{k\in\mathbb{N}^*} \partial_t \hat{u}(t,\cdot)(k) sin(k\pi x).$$

Tout ceci est rendu possible car

- $u(\cdot,x)\sin(k\pi x)$ est dérivable par rapport à t
- $u(\cdot,x)\sin(k\pi x)$ est intégrable par rapport à x
- Pour tout $\varepsilon \in]0,T[$, pour tout $t \in [\varepsilon,T]$, on a pour tout x

$$|\partial_t u(t,x)\sin(k\pi x)| \leqslant \|\partial_t u\|_{L^\infty([\varepsilon,T]\times[0,1])} \in L^1(]0,1[).$$

On fait le choix de noter $\hat{u}(k)(t)$ plutôt que $\widehat{u(t,\cdot)}(k)$ et on a donc

$$\partial_t \hat{u}(k)(t) = \partial_t\widehat{u(t,\cdot)}(k) = \widehat{\partial_t u(t,\cdot)}(k).$$

On a aussi $\widehat{\partial_x^2 u}(k)(t) = -k^2\pi^2\hat{u}(k)(t)$. Comme u est solution de (2.3), on a

$$\sum_{k\in\mathbb{N}^*}\left[\hat{u}(k)'(t) - \kappa\widehat{\partial_x^2 u}(k)(t)\right]\sin(k\pi x) = \sum_{k\in\mathbb{N}^*}\hat{f}(k)(t)\sin(k\pi x)$$

et donc

$$\sum_{k\in\mathbb{N}^*}\left[\hat{u}(k)'(t) + \kappa k^2\pi^2\hat{u}(k)(t) - \hat{f}(k)(t)\right]\sin(k\pi x) = 0.$$

Comme $\{\sqrt{2}\sin(k\pi\cdot)\}_{k\in\mathbb{N}^*}$ est une base hilbertienne, cela signifie que chacun des coefficients dans la somme précédente est nul

$$\hat{u}(k)'(t) + \kappa k^2\pi^2\hat{u}(k)(t) = \hat{f}(k)(t), \quad k \in \mathbb{N}^*,\ t \in]0,T[.$$

On a donc transformé l'EDP (2.3) en un famille d'EDO. Si on fournit la donnée initiale, on a

$$\hat{u}(k)(t) = \left[\hat{u}(k)(0) + \int_0^t \hat{f}(k)(s)e^{\kappa k^2\pi^2 s}\,ds\right]e^{-\kappa k^2\pi^2 t}.$$

Ainsi,

$$u(t,x) = \sqrt{2}\sum_{k\in\mathbb{N}^*}\hat{u}(k)(t)\sin(k\pi x) \quad \text{dans } L^2(]0,1[).$$

Il reste à définie $\hat{u}(k)(0)$. On sait que

$$\lim_{t\to 0^+} u(t,\cdot) = u^i \quad \text{dans } L^2(]0,1[).$$

Ceci est équivalent à

$$\lim_{t\to 0^+} \sqrt{2}\sum_{k\in\mathbb{N}^*} \hat{u}(k)(t)\sin(k\pi x) = \sqrt{2}\sum_{k\in\mathbb{N}^*} \widehat{u^i}(k)\sin(k\pi x) \quad \text{dans } L^2(]0,1[).$$

Comme $\{\sqrt{2}\sin(k\pi\cdot)\}_{k\in\mathbb{N}^*}$ est une base hilbertienne

$$\lim_{t\to 0^+} \hat{u}(k)(t) = \widehat{u^i}(k)$$

et donc $\widehat{u^i}(k) = \hat{u}(k)(0)$. On a ainsi construit la transformée en sinus de u solution de (2.3)

$$\hat{u}(k)(t) = \left[\hat{u}(k)(0) + \int_0^t \hat{f}(k)(s)e^{\kappa k^2\pi^2 s}\,ds\right] e^{-\kappa k^2\pi^2 t}. \tag{2.4}$$

Au final, si u^1 et u^2 sont deux solutions de (2.3) avec la même donnée initiale, on a $u = u^1 - u^2 = 0$ d'où l'unicité.

Nous pouvons maintenant fournir un résultat de stabilité ou encore une estimation d'énergie. On utilise pour cela la relation de Bessel-Parseval qui lie les normes L^2 dans l'espace physique et l'espace de fréquence (espace de Fourier)

$$E(t) := \|u(t,\cdot)\|_{L^2(]0,1[)} = \|\{\hat{u}(k)(t)\}_{k\in\mathbb{N}^*}\|_{\ell^2} := \left(\sum_{k\in\mathbb{N}^*} |\hat{u}(k)(t)|^2\right)^{1/2}.$$

$E(t)$ désigne l'énergie thermique contenue dans un fil au temps t. D'après la construction de u (2.4)

$$\begin{aligned}\|\{\hat{u}(k)(t)\}_{k\in\mathbb{N}^*}\|_{\ell^2} &\leqslant \left\|\left\{\widehat{u^i}(k)e^{-\kappa k^2\pi^2 t}\right\}_{k\in\mathbb{N}^*}\right\|_{\ell^2} + \left\|\left\{\int_0^t \hat{f}(k)(s)e^{\kappa k^2\pi^2 s}\,ds e^{-\kappa k^2\pi^2 t}\right\}_{k\in\mathbb{N}^*}\right\|_{\ell^2} \\ &\leqslant \left\|\left\{\widehat{u^i}(k)\right\}_{k\in\mathbb{N}^*}\right\|_{\ell^2} + \sqrt{\sum_{k\in\mathbb{N}^*}\int_0^t |\hat{f}(k)(s)|^2\,ds \int_0^t e^{2\kappa k^2\pi^2(s-t)}\,ds}.\end{aligned}$$

On a

$$\int_0^t e^{2\kappa k^2\pi^2(s-t)}\,ds = \frac{1-e^{-2\kappa k^2\pi^2 t}}{2\kappa k^2\pi^2} \leqslant \frac{1}{2\kappa k^2\pi^2} \leqslant \frac{1}{2\kappa\pi^2}, \quad k\in\mathbb{N}^*$$

et donc

$$\int_0^t |\hat{f}(k)(s)|^2\,ds \int_0^t e^{2\kappa k^2\pi^2(s-t)}\,ds \leqslant \frac{1}{2\kappa\pi^2}\int_0^t |\hat{f}(k)(s)|^2\,ds.$$

Finalement,

$$\begin{aligned}\|\{\hat{u}(k)(t)\}_{k\in\mathbb{N}^*}\|_{\ell^2} &\leqslant \left\|\left\{\widehat{u^i}(k)\right\}_{k\in\mathbb{N}^*}\right\|_{\ell^2} + \frac{1}{\sqrt{2\kappa}\pi}\sqrt{\sum_{k\in\mathbb{N}^*}\int_0^t |\hat{f}(k)(s)|^2\,ds} \\ &\leqslant \left\|\left\{\widehat{u^i}(k)\right\}_{k\in\mathbb{N}^*}\right\|_{\ell^2} + \frac{1}{\sqrt{2\kappa}\pi}\sqrt{\int_0^t \sum_{k\in\mathbb{N}^*} |\hat{f}(k)(s)|^2\,ds},\end{aligned}$$

où on utilise le théorème de Beppo-Levi pour permuter les signes somme et intégral.

Finalement, par Parseval, on a

$$\|\{\hat{u}(k)(t)\}_{k\in\mathbb{N}^*}\|_{\ell^2} \leqslant \left\|\left\{\widehat{u^i}(k)\right\}_{k\in\mathbb{N}^*}\right\|_{\ell^2} + \frac{1}{\sqrt{2\kappa}\pi}\underbrace{\left(\int_0^t \|f(s,\cdot)\|^2_{L^2(]0,1[)}\,ds\right)^{1/2}}_{=\|f\|_{L^2(]0,T[\times]0,1[)}}.$$

Ainsi, pour tout t, on a

$$\sup_{t\in[0,T]} \|u(t,\cdot)\|_{L^2(]0,1[)} \leqslant \|u^i\|_{L^2(]0,1[)} + \frac{1}{\sqrt{2\kappa\pi}}\|f\|_{L^2(]0,T[\times]]0,1[)}.$$

Physiquement, cette inégalité traduit que l'énergie contenue dans un fil au temps t est inférieure à la somme de l'énergie initialement contenue dans le fil et de l'énergie fournie par chauffage du fil entre $t = 0$ et l'instant t.

Mathématiquement, cette inégalité traduit la continuité de la solution u par rapport aux données u^i et f. En effet, soit $\bar{u^i}$ et $\bar{f}$ proches de u^i et f au sens

$$\|u^i - \bar{u^i}\|_{L^2(]0,1[)} \leqslant \varepsilon, \qquad \|f - \bar{f}\|_{L^2(]0,T[\times]]0,1[)} \leqslant \varepsilon.$$

La quantité $u - \bar{u}$ est solution de (2.3) avec $f - \bar{f}$ comme terme source, $u^i - \bar{u^i}$ comme donnée initiale et avec des conditions aux limite inchangées. D'après l'inégalité, on a

$$\begin{aligned}\|u(t,\cdot) - \bar{u}(t,\cdot)\|_{L^2(]0,1[)} &\leqslant \|u^i - \bar{u^i}\|_{L^2(]0,1[)} + \tfrac{1}{\sqrt{2\kappa\pi}}\|f - \bar{f}\|_{L^2(]0,T[\times]]0,1[)} \\ &\leqslant \varepsilon(1 + \tfrac{1}{\sqrt{2\kappa\pi}}), \quad t \in]0,T[.\end{aligned}$$

Donc $\bar{u}(t,\cdot)$ reste proche de $u(t,\cdot)$ en norme $L^2(]0,1[)$ ce qui traduit la stabilité L^2.

On a montré l'unicité et la stabilité de la solution. On a construit une solution $u \in L^2(]0,T[\times]]0,1[)$. Il faudrait montrer qu'en fait, $u \in C^1(]0,T[;C^2(]0,1[))$. C'est une preuve un peu longue disponible dans [8]. On peut donc conclure en l'existence et l'unicité d'une solution régulière.

R

1. Conditions aux limites de Dirichlet inhomogènes

$$\begin{cases} \partial_t u - \kappa\partial_x^2 u = f(t,x), \quad x \in]0,1[,\ t \in]0,T[, \\ u(t,0) = u_0,\ u(1,t) = u_1, \quad t \in]0,T[, \\ u(0,x) = u^i(x), \quad x \in]0,1[. \end{cases} \tag{2.5}$$

   On remarque que $g(t,x) = u_0 + (u_1 - u_0)x$ vérifie cette équation sans terme source. Soit maintenant v la solution de

$$\begin{cases} \partial_t v - \kappa\partial_x^2 v = f(t,x), \quad x \in]0,1[,\ t \in]0,T[, \\ v(t,0) = 0,\ v(1,t) = 0, \quad t \in]0,T[, \\ v(0,x) = u^i(x) - g(0,x), \quad x \in]0,1[. \end{cases}$$

   On sait que la solution de cette équation existe. Alors, $u(t,x) = v(t,x) + g(t,x)$ est solution de (2.5).
2. Conditions aux limites de Neumann homogènes. Au lieu d'utiliser une base de sinus, on prendrait la famille $\{1, \{\cos(k\pi\cdot)\}_{k\in\mathbb{N}*}\}$ comme base hilbertienne de $L^2(]0,1[)$

2.3.2 Principe du maximum

On a montré la stabilité de la solution u en norme L^2. On souhaite faire une étude similaire pour la norme L^∞. Cette stabilité L^∞ porte le nom de principe du maximum. Il existe plusieurs preuves. On choisit celle utilisant la notion d'entropie. On appelle entropie pour (2.1) toute fonction $C^2(\mathbb{R})$ telle que si u est solution de (2.1),

$$\partial_t S(u)(t,x) - \kappa\partial_x^2 S(u)(t,x) \leqslant S'(u)(t,x)f(t,x).$$

Proposition 2.3.2 Si $u^i \in L^2(]0,1[)$ et $f \in C^2([0,T]\times[0,1])$ avec $f(t,0) = f(t,1) = 0$, alors toute fonction S de classe $C^2(\mathbb{R})$ et convexe sur $\mathbb{R}$ est une entropie.

Démonstration. On a les relations suivantes

$$\partial_t S(u) = S'(u)\partial_t u, \quad \partial_x S(u) = S'(u)\partial_x u, \quad \partial_x^2 S(u) = S''(u)(\partial_x u)^2 + S'(u)\partial_x^2 u.$$

Donc,

$$\begin{aligned}\partial_t S(u)(t,x) - \kappa\partial_x^2 S(u)(t,x) &= S'(u)\partial_t u - \kappa S''(u)(\partial_x u)^2 - \kappa S'(u)\partial_x^2 \\ &\leqslant S'(u)(\partial_t u - \kappa\partial_x^2 u) = S'(u)f(t,x)\end{aligned}$$

■
■

On suppose maintenant pour simplifier que $f(t,x) = 0$.

Théorème 2.3.3 $u^i \in L^2(]0,1[) \cap L^\infty(]0,1[)$ et $u(t,x)$ l'unique solution de l'équation de la chaleur avec conditions aux limites de Dirichlet homogènes et $f = 0$. Alors,

$$\|u(t,\cdots)\|_{L^\infty(]0,1[)} \leqslant \|u^i\|_{L^\infty(]0,1[)}, \quad \forall t \geqslant 0.$$

Démonstration. Soit $S \in C^2(\mathbb{R})$ et $S'' > 0$. Alors, S est une entropie pour (2.1) et vérifie

$$\partial_t S(u)(t,x) - \kappa\partial_x^2 S(u)(t,x) \leqslant 0.$$

Notons $M = \max(0, \operatorname{sup\,ess}\, u^i)$. On rappelle que $\operatorname{sup\,ess}\, u^i$ est le plus petit des majorants de u^i et que $\operatorname{inf\,ess}\, u^i$ est le plus grand des minorants de u^i. On a le choix de S. On construit donc S telle que

$$S(u) = \begin{cases} 0, \text{ si } u < M, \\ (u-M)^3, \text{ si } u \geqslant M. \end{cases}$$

La fonction S est bien de classe C^2 et $S'' > 0$. On intègre l'inégalité d'entropie

$$\int_0^1 \left[\partial_t S(u)(t,x) - \kappa\partial_x^2 S(u)(t,x) \leqslant 0\right] dx \Longleftrightarrow \int_0^1 \partial_t S(u)(t,x)\,dx - \kappa\left[\partial_x S(u)(t,x)\right]_0^1 \leqslant 0.$$

Or, pour tout $t > 0$, $u(t,0) = u(t,1) = 0 \leqslant M$. Donc, en utilisant la définition de S, on a

$$\partial_x S(u)(t,0) = S'(u(t,0))\partial_x u(t,0) = S'(0)\partial_x u(t,0) = 0$$

et

$$\partial_x S(u)(t,1) = S'(u(t,1))\partial_x u(t,1) = S'(0)\partial_x u(t,1) = 0.$$

Ainsi,

$$\int_0^1 \partial_t S(u)(t,x)\,dx = \partial_t \int_0^1 S(u)(t,x)\,dx \leqslant 0$$

On intègre enfin par rapport à t de sorte que

$$\int_0^1 S(u)(t,x)\,dx \leqslant \int_0^1 S(u)(0,x)\,dx$$

Or, $M = \max(0, \operatorname{sup\,ess}\, u^i)$. D'où, pour presque tout $x \in [0,1]$, $S(u)(0,x) = 0$ d'où

$$\int_0^1 S(u)(t,x)\,dx \leqslant 0.$$

Comme $S(u) \geqslant 0$ pour tout $x \in \mathbb{R}$, cela implique que $S(u)(t,x) = 0$ pour presque tout x et tout t. Ceci signifie que

$$u(t,x) \leqslant M, \quad \text{pp } x, \forall t.$$

On refait les mêmes estimations pour $m = \min(0, \inf \text{ess } u^i)$ et

$$\tilde{S}(u) = \begin{cases} -(u-m)^3, \text{ si } u \leqslant m, \\ 0, \text{ si } u > m. \end{cases}$$

On démontre de la même manière que $u(t,x) \geqslant m$ pour presque tout x et tout t. ■

Physiquement, ceci indique que la plus forte chaleur est atteinte à $t = 0$ et que la chaleur ne peut que décroître, mais aussi qu'elle ne peut pas être inférieure à la plus petite valeur de u^i.

Si on considère des conditions aux limites de Dirichlet inhomogènes $u(t,0) = u_0$ et $u(t,1) = u_1$, on montrerait

$$\min(u_0, u_1, \inf \text{ess } u^i) \leqslant u(t,x) \leqslant \max(u_0, u_1, \sup \text{ess } u^i).$$

La chaleur maximale et minimale sont atteintes soit à $t = 0$, soit sur les bords.

2.3.3 Résolution de l'équation de la chaleur par séparation des variables

La méthode utilisée pour montrer l'existence et l'unicité de solution de (2.1) avec $f = 0$ est en fait un cas particulier de la méthode de séparation des variables. Celle technique revient à chercher u sous la forme

$$u(t,x) = \psi(t)\varphi(x).$$

Nous présentons ici le cas des conditions aux limites de Dirichlet homogènes. Le lecteur intéressé pourra consulter le livre [5] où d'autres cas sont traités.

Étape 1

On injecte cette relation dans l'équation de la chaleur. On obtient

$$\psi'(t)\varphi(x) = \kappa\psi(t)\varphi''(x)$$

ou encore

$$\frac{1}{\kappa}\frac{\psi'(t)}{\psi(t)} = \frac{\varphi''(x)}{\varphi(x)}.$$

Comme le membre de gauche de cette égalité ne dépend que de t et que celui de droite ne dépend que de x, on en déduit qu'ils sont constants. C'est à dire qu'il existe une constante $\lambda \in \mathbb{R}$ telle que

$$\psi'(t) = \lambda\kappa\psi(t), \qquad \varphi''(x) = \lambda\varphi(x).$$

On a donc deux équations différentielles à variables séparées.

Étape 2

On cherche les solutions non nulles $\psi(x)$ vérifiant les conditions aux limites

$$\begin{cases} \varphi''(x) = \lambda\varphi(x), \\ \varphi(0) = \varphi(1) = 0. \end{cases}$$

Il faut traiter trois cas différents.

$\underline{\lambda > 0}$ Les solutions sont données par $\varphi(x) = Ae^{\sqrt{\lambda}x} + Be^{-\sqrt{\lambda}x}$. Les conditions aux limites conduisent à

$$A + B = 0, \quad \text{et} \quad Ae^{\sqrt{\lambda}} + Be^{-\sqrt{\lambda}} = 0.$$

L'unique solution est $(A,B) = (0,0)$ et donc il n'y a pas de solutions non nulles.

$\underline{\lambda = 0}$ Les solutions sont données par $\varphi(x) = Ax + B$. Les conditions aux limites conduisent à $(A, B) = (0, 0)$. Il n'y a donc pas de solutions non nulles.

$\underline{\lambda < 0}$ Les solutions sont données par $\varphi(x) = A\sin(\sqrt{-\lambda}x) + B\cos(\sqrt{-\lambda}x)$. Les conditions aux limites conduisent à

$$B = 0, \quad \text{et} \quad A\sin(\sqrt{\lambda}) = 0.$$

On est face à deux alternatives. Soit A est nulle et on retrouve les deux cas précédents, soit $A \neq 0$ et $\sqrt{-\lambda} = n\pi$, $n > 0$. On obtient donc une famille de solutions non nulles

$$\varphi_n(x) = \sin(n\pi x), \quad n > 0,$$

associées aux $\lambda_n = -n^2\pi^2$. Nous avons donc une infinité de solutions et φ_n et λ_n peuvent respectivement être associée à la notion de fonction propre et de valeur propre.

Étape 3

On remarque que $\langle \varphi_n, \varphi_m \rangle_{L^2} = 0$ pour $n \neq m$. Donc la famille $\{\varphi_n\}_{n\in\mathbb{N}}$ est une base sur laquelle on peut développer la solution

Étape 4

On doit maintenant résoudre $\psi_n'(t) = \lambda_n \kappa \psi_n(t)$. Les solutions sont

$$\psi_n(t) = c_n e^{\kappa\lambda_n t} = c_n e^{-\kappa n^2\pi^2 t},$$

et les c_n sont déterminés par la condition initiale.

Étape 5

Comme l'équation (2.1) est linéaire, la somme de plusieurs solutions à l'équation est toujours solution de l'équation. Ainsi, u est la somme de toutes les solutions élémentaires

$$u(t, x) = \sum_{n\in\mathbb{N}^*} \psi_n(y)\varphi_n(x) = \sum_{n\in\mathbb{N}^*} c_n e^{-\kappa n^2\pi^2 t}\sin(n\pi x).$$

Or $u(0, x) = u^i(x)$. D'où,

$$\sum_{n\in\mathbb{N}^*} \psi_n(0)\varphi_n(x) = \sum_{n\in\mathbb{N}^*} c_n\varphi_n(x) = u^i(x).$$

Ainsi, les c_n sont les coordonnées de la décomposition de u^i dans la base $\{\varphi_n\}_{n\in\mathbb{N}}$ et on a

$$c_n = \frac{\langle u^i, \varphi_n\rangle_{L^2}}{\langle \varphi_n, \varphi_m\rangle_{L^2}} = 2\int_0^1 u^i(x)\sin(n\pi x)\,dx.$$

Cette méthode s'applique pour d'autres EDP linéaires et d'autres conditions aux limites. Elle s'étend également aux dimensions supérieures. Si on considère par exemple le cas bidimensionnel avec l'équation de la chaleur

$$\begin{cases} \partial_t u - \Delta u = 0,\ x \in \Omega = [0, a] \times [0, b],\ t > 0, \\ u(t, 0) = u^i(x), \\ u(0, x) = 0,\ x \in \partial\Omega. \end{cases}$$

On recherche u sous la forme $u(t, x, y) = \psi(t)f(x)g(y)$. L'équation vérifiée est

$$\psi' fg - \psi(f''g + fg'') = 0$$

soit encore

$$\frac{\psi'(t)}{\psi(t)} = \frac{f''(x)g(y) + f(x)g''(y)}{f(x)g(y)} = \lambda.$$

On est donc amené à résoudre respectivement

$$\psi'(t) = \lambda\psi(t)$$

et

$$f''(x)g(y) + f(x)g''(y) = \lambda f(x)g(y).$$

Cette dernière équation est équivalente à

$$\frac{f''(x)}{f(x)} + \frac{g''(y)}{g(y)} = \lambda$$

soit encore

$$\frac{f''(x)}{f(x)} = \lambda - \frac{g''(y)}{g(y)}.$$

Chaque côté de l'égalité est constant pour l'autre et on est donc amener à résoudre

$$f''(x) = \mu f(x) \quad \text{et} \quad g''(y) = \nu g(y), \text{ où } \nu = \lambda - \mu$$

avec $f(0 = f(a) = 0$ et $g(0) = g(b) = 0$. On est donc ramené au cas de la dimension 1.

2.4 Résolution par transformée de Fourier discrète

Nous présentons dans cette section comment utiliser la transformée de Fourier discrète vue dans le chapitre précédent comme méthode numérique de résolution. On présente la méthode sur l'équation de la chaleur mono-dimensionnelle avec conditions aux limites périodiques

$$\begin{cases} \partial_t u = \partial_x^2 u = 0, & 0 \leqslant x < a,\ t > 0, \\ u(0,x) = u^i(x), & 0 \leqslant x < a, \\ u(t,0) = u(t,a), & t > 0, \end{cases}$$

où u^i est une fonction a-périodique régulière.

Nous avons vu qu'il y a un lien fort entre transformée de Fourier continue, discrète et série de Fourier. On sait par exemple que si f est une fonction a-périodique, on a

$$S(f'') = \sum_{n\in\mathbb{Z}} -\frac{4\pi^2 n^2}{a^2} C_n(f) e_n(t),$$

où $C_n(f) = \int_0^a f(t)e_{-n}(t)\,dt/a$.

Discrétisons l'inconnue u sur un maillage uniforme de sorte que $u_k(t)$ désigne une approximation de $u(t, ka/N)$, $0 \leqslant k < N$. Soit u^0 le vecteur de composantes $u_k^0 = u^i(ka/N)$, $0 \leqslant k < N$. En remplaçant $u(t,x)$ par $u_k(t)$ dans l'équation de la chaleur, on a

$$u_k'(t) = \mathscr{F}_N^{-1}\left(-\frac{4\pi^2 j^2}{a^2}\mathscr{F}_N\left(u_k(t)\right)_j\right)_k, \quad 0 \leqslant k < N.$$

Cette équation est une équation différentielle ordinaire pour chaque mode de Fourier k. En effet, on a

$$\mathscr{F}_N\left(u_k'(t)\right)_j = -\frac{4\pi^2 j^2}{a^2}\mathscr{F}_N\left(u_k(t)\right)_j, \quad 0 \leqslant k < N.$$

Or, par définition de la transformée de Fourier discrète, on a

$$\left(\mathscr{F}_N(u_k)\right)_j = \sum_{k=0}^{N-1} u_j(t) e^{-2i\pi jk/N}.$$

Ainsi,

$$\left(\mathscr{F}_N(u_k')\right)_j = \sum_{k=0}^{N-1} u_j'(t) e^{-2i\pi jk/N} = \frac{d}{dt}\left(\mathscr{F}_N(u_k)\right)_j .$$

Nous pouvons en déduire

$$\frac{d}{dt}\mathscr{F}_N\left(u_k(t)\right)_j = -\frac{4\pi^2 j^2}{a^2}\mathscr{F}_N\left(u_k(t)\right)_j .$$

et donc

$$\mathscr{F}_N\left(u_k(t)\right)_j = e^{-\frac{4\pi^2 j^2}{a^2}t}\mathscr{F}_N\left(u^0\right)_j .$$

Ainsi, on a simplement

$$u_k(t) = \mathscr{F}_N^{-1}\left(e^{-\frac{4\pi^2 j^2}{a^2}t}\mathscr{F}_N\left(u^0\right)_j\right)_k ,$$

soit encore

$$U(t) = \mathscr{F}_N^{-1}\left(e^{-\frac{4\pi^2 j^2}{a^2}t}\mathscr{F}_N\left(u^0\right)\right), \tag{2.6}$$

où $U(t) = [u_0(t), u_1(t), \cdots, u_{N-1}(t)]^T$.

Aucune approximation temporelle n'est nécessaire ce qui simplifie grandement la mise en place du schéma numérique. En outre, on sait que les calculs des transformées de Fourier discrètes (direct ou inverse) sont rendus très rapides par l'utilisation de la FFT (Fast Fourier Transform). Pour d'autres conditions aux limites standards (Dirichlet homogène ou Neumann homogène), on utiliserait les transformées en sinus ou en cosinus.

2.5 Résolution par différences finies

Nous présentons dans cette section une discrétisation complète espace-temps. Les pas d'espace et de temps sont respectivement notés δx et δt. Les temps discrets sont $t^n = n\delta t$, $n \in \{0, \cdots, N\}$, de sorte que $\delta t = T/N$. La variable d'espace $x \in [0,1]$ est discrétisée avec $J+2$ points et on a $x_j = j\delta x$, $j \in \{0, \cdots, J+1\}$, $x_0 = 0$ et $x_{J+1} = 1$. Ainsi, $\delta x = 1/J+1$. On cherche des approximations de $u(t^n, x_j)$ et on définit $u_j^0 = u^i(x_j)$. On a pour la discrétisation spatiale

$$\partial_x^2 u(t^n, x_j) = \frac{u(t^n, x_{j+1}) - 2u(t^n, x_j) + u(t^n, x_{j-1})}{\delta x^2} + \mathcal{O}(\delta x^2).$$

Concernant la discrétisation temporelle, on peut considérer des discrétisations amont (explicite) ou aval (implicite) avec

$$\begin{aligned}
\partial_t u(t^n, x_j) &= \frac{u(t^n, x_j) - u(t^{n-1}, x_j)}{\delta t} + \mathcal{O}(\delta t) \quad \text{explicite,} \\
\partial_t u(t^n, x_j) &= \frac{u(t^{n+1}, x_j) - u(t^n, x_j)}{\delta t} + \mathcal{O}(\delta t) \quad \text{implicite.}
\end{aligned}$$

Étudions le cas du schéma explicite

$$\frac{u_j^{n+1} - u_j^n}{\delta t} - \kappa \frac{u_{j+1}^2 - 2u_j^n + u_{j-1}^n}{\delta x^2} = f(t^n, x_j), \tag{2.7}$$

où u_j^n désigne une approximation de $u(t^n, x_j)$.

Définition 2.5.1 Erreur de consistance

On appelle erreur de consistance l'erreur commise en remplaçant l'équation exacte par l'équation aux différences finies

Ici,

$$\varepsilon_j^n(u) = \frac{u(t^{n+1},x_j)-u(t^n,x_j)}{\delta t} - \kappa\frac{u(t^n,x_{j+1})-2u(t^n,x_j)+u(t^n,x_{j-1})}{\delta x^2} - f(t^n,x_j),$$

et on pose $\varepsilon^n(u) = \left(\varepsilon_j^n(u)\right)_{1\leqslant j\leqslant J} \in \mathbb{R}^J$.

Proposition 2.5.1 Si $u \in C_t^2C_x^4$ est solution de (2.1), alors il existe $C \in \mathbb{R}$ telle que

$$\max_{0\leqslant n\leqslant N-1} \|\varepsilon^n(u)\|_\infty \leqslant c(\delta t + \delta x^2).$$

On dit que le schéma est d'ordre 1 en temps et 2 en espace.

R Il en résulte $\lim_{\delta x,\delta t\to 0} \max_{0\leqslant n\leqslant N-1} \|\varepsilon^n(u)\|_\infty = 0$. Le schéma est donc consistant avec (2.1).

Démonstration. La preuve repose sur la formule de Taylor.

$$\varepsilon_j^n(u) = \frac{\delta t}{2}\partial_t^2 u(\tau,x_j) - \kappa\frac{\delta x^2}{24}(\partial_x^4 u(t^n,y) + \partial_x^4 u(t^n,z)),$$

où

$$\tau \in [t^n,t^{n+1}], \quad y\in[x_j,x_{j+1}], \quad z\in[x_{j-1},x_j].$$

La constante C est donc

$$C = \frac{1}{2}\max\left(\max_{[0,T]\times[0,1]}|\partial_t^2 u|, \frac{\kappa}{6}\max_{[0,T]\times[0,1]}|\partial_x^4 u|\right).$$

■

Théorème 2.5.2 Soit $e^n(u) = \left(e_j^n(u)\right)_{1\leqslant j\leqslant J}$ le vecteur de l'erreur globale au temps t^n défini par

$$e_j^n(u) = u_j^n - u(t^n,x_j), \quad 0\leqslant n\leqslant N,\ 0\leqslant j\leqslant J+1.$$

Supposons $\kappa\dfrac{\delta t}{\delta x^2} \leqslant \dfrac{1}{2}$, alors il existe $C\in\mathbb{R}$ telle que

$$\|e^n(u)\|_\infty \leqslant C(\delta t+\delta x^2), \quad 0\leqslant n\leqslant N.$$

Démonstration. On a $e_0^n(u) = e_{J+1}^n(u) = 0$ pour tout n. D'où, pour $1\leqslant j\leqslant J$, on a

$$\begin{aligned}
e_j^{n+1}(u) &= u_j^{n+1} - u(t^{n+1},x_j)\\
&= u_j^n + \kappa\frac{\delta t}{\delta x^2}(u_{j+1}^2 - 2u_j^n + u_{j-1}^n) + \delta t f(t^n,x_j)\\
&\quad -u(t^n,x_j) - (u(t^n+1,x_j) - u(t^n,x_j))\\
&= e_j^n + \kappa\frac{\delta t}{\delta x^2}(u_{j+1}^2 - 2u_j^n + u_{j-1}^n) + \delta t f(t^n,x_j)\\
&\quad - \left(\delta t\varepsilon_j^n(u) + \kappa\frac{\delta t}{\delta x^2}\left(u(t^n,x_{j+1} - 2u(t^n,x_j) + u(t^n,x_{j-1}))\right) + \delta t f(t^n,x_j)\right)\\
&= e_j^n + \kappa\frac{\delta t}{\delta x^2}(e_{j+1}^2 - 2e_j^n + e_{j-1}^n) - \delta t\varepsilon_j^n(u).
\end{aligned}$$

Ainsi, e_j^n, $1\leqslant j\leqslant J$, est solution de

$$\frac{e_j^{n+1}(u)-e_j^n(u)}{\delta t} - \kappa\frac{e_{j+1}^n(u) - 2e_j^n(u) + e_{j-1}^n(u)}{\delta x^2} = -\varepsilon_j^n(u).$$

On a donc

$$e_j^{n+1}(u) = \left(1 - 2\kappa\frac{\delta t}{\delta x^2}\right) e_j^n(u) + \kappa\frac{\delta t}{\delta x^2} e_{j+1}^n(u) + \kappa\frac{\delta t}{\delta x^2} e_j^n(u) - \delta t\varepsilon_j^n(u).$$

On a supposé $0 \leqslant 2\kappa\delta t/\delta x^2 \leqslant 1$, d'où l'estimation pour $0 \leqslant n \leqslant N-1$, $0 \leqslant j \leqslant J+1$,

$$\begin{aligned} |e_j^{n+1}(u)| &\leqslant \left(1 - 2\kappa\frac{\delta t}{\delta x^2}\right)\|e^n(u)\|_\infty + \frac{2\kappa\delta t}{\delta x^2}\|e^n(u)\|_\infty + \delta\|\varepsilon^n(u)\|_\infty, \\ &\leqslant \|e^n(u)\|_\infty + \delta t\|\varepsilon^n(u)\|_\infty. \end{aligned}$$

On a donc

$$\begin{aligned} \|e^{n+1}(u)\|_\infty &\leqslant \|e^n(u)\|_\infty + C\delta t(\delta t + \delta x^2) \\ &\leqslant \underbrace{\|e^0(u)\|_\infty}_{=0} + C(n+1)\delta t(\delta t + \delta x^2) \underbrace{\leqslant}_{n\delta t \leqslant T} CT(\delta t + \delta x^2) \end{aligned}$$

■

Donc, le schéma explicite converge si δx et δt tendent vers 0, et si on choisit δt suffisamment petit pour vérifier la **condition CFL**

$$\kappa\frac{\delta t}{\delta x^2} \leqslant \frac{1}{2}.$$

On dit que le schéma explicite est conditionnellement convergent dans L^∞.

Il est intéressant d'écrire le schéma sous forme matricielle

$$\frac{U^{n+1} - U^n}{\delta t} + \kappa\frac{\delta t}{\delta x^2} AU^n = F^n,$$

où

$$A = \begin{pmatrix} 2 & -1 & & & \\ -1 & 2 & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & -1 & 2 & -1 \\ & & & -1 & 2 \end{pmatrix} \in M_J(\mathbb{R}), \quad U = \begin{pmatrix} u_1^n \\ \vdots \\ u_J^n \end{pmatrix} \in \mathbb{R}^J, \quad F^n = \begin{pmatrix} f(t^n, x_1) \\ \vdots \\ f(t^n, x_J) \end{pmatrix} \in \mathbb{R}^J.$$

Pour le schéma implicite, on aurait

$$\left(\frac{I}{\delta t} + \frac{\kappa}{\delta x^2} A\right) U^{n+1} + \left(-\frac{I}{\delta t}\right) U^n = F^n.$$

On regroupe ces deux schémas sous une même formulation

$$\left(\frac{I}{\delta t} + \frac{\theta\kappa}{\delta x^2} A\right) U^{n+1} + \left(-\frac{I}{\delta t} + \frac{(1-\theta)\kappa}{\delta x^2} A\right) U^n = F^{n+1/2}, \quad 0 \leqslant \theta \leqslant 1, \tag{2.8}$$

où $F^{n+1/2} = (f(t^{n+1/2}, x_1), \cdots, f(t^{n+1/2}, x_J))^T$. Cette modification permet pour une valeur de θ particulière d'obtenir un schéma d'ordre 2 en temps. On a appelle la méthode (2.8) le ***θ*-schéma**. Pour $\theta = 1$, on retrouve le schéma implicite et pour $\theta = 0$ le schéma explicite (à la modification sur le second membre près).

R On pourrait être intéressé d'utiliser une méthode d'éléments finis pour l'approximation spatiale. On partirait pour cela de la formulation variationnelle. On multiplie l'équation par une fonction test ϕ et on intègre en espace

$$\int_0^1 \Big[\partial_t u - \partial_x^2 u = f\Big]\phi(x)\,dx.$$

On définit

$$u_h^n(x) = \sum_{j=1}^{J} i_{h,j}^n \phi_j(x) \in V_h,$$

où $\{\phi_j\}_{1\leqslant j\leqslant J}$ est une base d'éléments finis de V_h. On remplace $\phi(x)$ dans la formulation variationnelle par tous les ϕ_i. Simultanément, on remplace les dérivées temporelles par leur approximation différences finies et on adapte pour obtenir un θ-schéma

$$\sum_{j=1}^{J} \frac{u_{h,j}^{n+1} - u_{h,j}^n}{\delta t} \int_0^1 \phi_j\phi_i\,dx + \theta\kappa u_{h,j}^{n+1} \int_0^1 \partial_x\phi_j\partial_x\phi_i\,dx + (1-\theta)\kappa u_{h,j}^n \int_0^1 \partial_x\phi_j\partial_x\phi_i\,dx = \int_0^1 f(t^{n+1/2}, x_j)\phi_i(x)\,dx,\ 1\leqslant i\leqslant J.$$

On a donc

$$\left(\frac{M}{\delta t} + \theta\kappa S\right) U^{n+1} + \left(-\frac{M}{\delta t} + (1-\theta)\kappa S\right) U^n = F^{n+1/2},$$

où M et S sont respectivement les matrices de masse et de rigidité et $F_j^{n+1/2} = \int_0^1 f(t^{n+1/2}, x)\phi_j(x)\,dx$, pour $1\leqslant j\leqslant J$.

On étudie maintenant le θ-schéma différences finies. L'erreur de consistance est

$$\varepsilon_j^n(u) = \frac{u(t^{n+1},x_j) - u(t^n,x_j)}{\delta t} - \kappa\left((1-\theta)\frac{u(t^n,x_{j+1}) - 2u(t^n,x_j) + u(t^n,x_{j-1})}{\delta x^2} + \theta\frac{u(t^{n+1},x_{j+1}) - 2u(t^{n+1},x_j) + u(t^{n+1},x_{j-1})}{\delta x^2}\right) - f(t^{n+1/2},x_j).$$

Proposition 2.5.3 Soit $u \in C_t^3 C_x^4$ solution de (2.1). Alors, il existe $C \in \mathbb{R}$ et $D \in \mathbb{R}$ telles que

$$\max_{0\leqslant n\leqslant N-1} \|\varepsilon^n(u)\|_\infty \leqslant C|1-2\theta|\delta t + D(\delta t^2 + \delta x^2).$$

Le θ-schéma est consistant pour tout θ, d'ordre 1 en temps et 2 en espace sauf pour $\theta = 1/2$ où il est d'ordre 2 en temps.

Étudions maintenant l'erreur $e_j^n(u) = u_j^n - u(t^n, x_j)$ pour $0\leqslant n\leqslant N$ et $1\leqslant j\leqslant J$. Par linéarité, on a

$$\left(\frac{I}{\delta t} + \frac{\theta\kappa}{\delta x^2}A\right) e^{n+1}(u) + \left(-\frac{I}{\delta t} + \frac{(1-\theta)\kappa}{\delta x^2}A\right) e^n(u) = -\varepsilon^n(u).$$

La matrice $\tilde{A}_\theta = \left(\frac{I}{\delta t} + \frac{\theta\kappa}{\delta x^2}A\right)$ n'est pas diagonale si $\theta \neq 0$. L'étude de l'erreur ℓ^∞ est donc difficile. On préfère mener une étude ℓ^2. On note

$$|||e^n|||_2 = \sqrt{\frac{1}{J}\sum_{j=1}^{J} |e_j^n|^2}.$$

Étudions les propriétés de la matrice $\tilde{A}_\theta$. On sait que les valeurs propres de A sont

$$\alpha_j = 4\sin^2\left(\frac{j\pi}{2(J+1)}\right) > 0, \quad 1 \leqslant j \leqslant J,$$

avec les vecteurs propres associés

$$V_j = \left(\sin\left(\frac{j\pi}{J+1}\right), \sin\left(\frac{2j\pi}{J+1}\right), \cdots, \sin\left(\frac{Jj\pi}{J+1}\right)\right)^T.$$

La matrice $\tilde{A}_\theta$ est symétrique réelle. Elle est donc diagonalisable sur $\mathbb{R}$. Ses valeurs propres sont

$$\mu_j = \frac{1}{\delta t} + \frac{\theta\kappa}{\delta x^2}\alpha_j > 0, \quad 0 \leqslant \theta \leqslant 1$$

et les mêmes vecteurs propres que ceux de A. La matrice $\tilde{A}_\theta$ est donc définie positive et par la même inversible.

On note

$$B_\theta = I + \theta\frac{\kappa\delta t}{\delta x^2}A = \delta t\tilde{A}_\theta.$$

On a

$$e^{n+1}(u) = B_\theta^{-1}\Big[B_{\theta-1}e^n(u) - \delta t\varepsilon^n(u)\Big].$$

On définit la matrice d'amplification du schéma

$$L_\theta = B_\theta^{-1}B_{\theta-1}.$$

Alors,

$$e^{n+1}(u) = L_\theta e^n(u) - \delta B_\theta^{-1}\varepsilon^n(u).$$

On peut donc estimer

$$\begin{aligned}|||e^{n+1}|||_2 &\leqslant |||L_\theta e^n|||_2 + \delta t|||B_\theta^{-1}\varepsilon^n|||_2 \\ &\leqslant |||L_\theta|||_2|||e^n|||_2 + \delta t|||B_\theta^{-1}|||_2|||\varepsilon^n|||_2.\end{aligned}$$

Appliquer à L_θ, la norme $|||\cdot|||_2$ est une norme matricielle

$$|||M|||_2 = \sup_{v\in\mathbb{R}^J}\frac{|||Mv|||_2}{|||v|||_2} = \sqrt{\rho(M^*M)},$$

où $\rho(C)$ désigne le rayon spectrale d'une matrice C. On sait que les valeurs propres de B_θ sont pour $1 \leqslant j \leqslant J$

$$\beta_j = 1 + \theta\frac{\kappa\delta t}{\delta x^2}\alpha_j = \delta t\mu_j \geqslant 1, \quad \forall\theta \geqslant 0.$$

Ainsi, B_θ est inversible et les valeurs propres de B_θ^{-1} sont toutes dans l'intervalle $]0,1]$. Or, B_θ est symétrique réelle et donc B_θ^{-1} aussi. D'où $(B_\theta^{-1})^* = B_\theta^{-1}$ et

$$|||B_\theta^{-1}|||_2 = \sqrt{\rho((B_\theta^{-1})^*B_\theta^{-1})} = \rho(B_\theta^{-1}) \leqslant 1.$$

Ainsi,

$$|||e^{n+1}|||_2 \leqslant |||L_\theta|||_2|||e^n|||_2 + \delta t|||\varepsilon^n|||_2.$$

Par récurrence,

$$|||e^{n+1}|||_2 \leqslant |||L_\theta|||_2^{n+1}\underbrace{|||e^0|||_2}_{=0} + \delta t\sum_{p=0}^{n}|||L_\theta|||_2^{n-p}|||\varepsilon^p|||_2.$$

Or, $\sqrt{\sum_{j=1}^{J}|\varepsilon_j^p(u)|^2} \leqslant \sqrt{J\|\varepsilon^p(u)\|_\infty^2}$ et donc $\|\!|\varepsilon^p|\!\|_2 \leqslant \|\!|\varepsilon^p|\!\|_\infty$. On obtient donc

$$\begin{aligned}\|\!|e^{n+1}|\!\|_2 &\leqslant \delta t \sum_{p=0}^{n} \|\!|L_\theta|\!\|_2^{n-p} \|\!|\varepsilon^p|\!\|_\infty \\ &\leqslant \delta t \sum_{p=0}^{n} \|\!|L_\theta|\!\|_2^{n-p} \Big[C|1-2\theta|\delta t + D(\delta t^2 + \delta x^2)\Big].\end{aligned}$$

Si $\|\!|L_\theta|\!\|_2 \leqslant 1$, on a

$$\|\!|e^{n+1}|\!\|_2 \leqslant T\Big[C|1-2\theta|\delta t + D(\delta t^2 + \delta x^2)\Big]$$

ce qui nous permet de conclure à la convergence du schéma numérique.

Il reste à déterminer les conditions pour avoir $\|\!|L_\theta|\!\|_2 \leqslant 1$. Rappelons que

$$L_\theta = B_\theta^{-1} B_{\theta-1} = \left(I + \theta\frac{\kappa\delta t}{\delta x^2}A\right)^{-1}\left(I - (1-\theta)\frac{\kappa\delta t}{\delta x^2}A\right).$$

On sait que B_θ et B_θ^{-1} ont les mêmes vecteurs propres et des valeurs propres inverses. De même, les vecteurs propres de $B_{\theta-1}$ sont les mêmes que ceux de A, à savoir les vecteurs V_j, et ses valeurs propres réelles sont données par

$$\sigma_j = 1 - (1-\theta)\kappa\frac{\delta t}{\delta x^2}\alpha_j, \quad 1 \leqslant j \leqslant J.$$

Ainsi, les vecteurs propres de L_θ sont donc les vecteurs V_j et ses valeurs propres sont les réels donnés par

$$\lambda_j = \frac{\sigma_j}{\beta_j} = \frac{1 - 4(1-\theta)\kappa\frac{\delta t}{\delta x^2}\sin^2\left(\frac{j\pi}{2(J+1)}\right)}{1 + 4\theta\kappa\frac{\delta t}{\delta x^2}\sin^2\left(\frac{j\pi}{2(J+1)}\right)}, \quad 1 \leqslant j \leqslant J.$$

Comme L_θ est une matrice symétrique réelle, on peut estimer sa 2-norme par son rayon spectral

$$\|\!|L_\theta|\!\|_2 \leqslant 1 \Longleftrightarrow \rho(L_\theta) \leqslant 1 \Longleftrightarrow |\lambda_j| \leqslant 1, \quad 1 \leqslant j \leqslant J.$$

La condition $\rho(L_\theta) \leqslant 1$ est la condition de stabilité de Von Neumann.

Proposition 2.5.4

- Si $\theta \geqslant 1/2$, $\lambda(L_\theta) \leqslant 1$
- Si $\theta \in [0, 1/2[$, $\lambda(L_\theta) \leqslant 1$ si et seulement si $\kappa\dfrac{\delta t}{\delta x^2} \leqslant \dfrac{1}{2(1-2\theta)}$.

Démonstration. Il faut étudier la fonction g définie par

$$g(s) = \frac{1 - 4(1-\theta)\kappa\frac{\delta t}{\delta x^2}s}{1 + 4\theta\kappa\frac{\delta t}{\delta x^2}s}, \quad 0 \leqslant s \leqslant 1.$$

La fonction g est décroissante. Elle atteint sa valeur maximale en $s = 0$ et minimale en $s = 1$. Réciproquement, son module $|g|$ est maximal en $s = 0$ et minimal en $s = 1$. Or, $g(0) = 1$. Étudions

$$g(1) = \frac{1 - 4(1-\theta)\kappa\frac{\delta t}{\delta x^2}}{1 + 4\theta\kappa\frac{\delta t}{\delta x^2}}, \quad 0 \leqslant \theta \leqslant 1.$$

Avoir $|g(1)| \leqslant 1$ est équivalent à

$$-(1 + 4\theta\kappa\frac{\delta t}{\delta x^2}) \leqslant 1 - 4(1-\theta)\kappa\frac{\delta t}{\delta x^2} \leqslant 1 + 4\theta\kappa\frac{\delta t}{\delta x^2}$$

soit encore

$$\kappa\frac{\delta t}{\delta x^2}(4-8\theta)\leqslant 2 \quad \text{et} \quad \kappa\frac{\delta t}{\delta x^2}\geqslant 0$$

ce qui conduit à

$$\theta\geqslant 1/2 \quad \text{ou} \quad \kappa\frac{\delta t}{\delta x^2}\leqslant\frac{1}{2(1-2\theta)}.$$

■

En conséquence, on a le théorème suivant

Théorème 2.5.5 Sous les mêmes hypothèses que précédemment, si $\theta\leqslant 1/2$ ou $\kappa\delta t/\delta x^2\leqslant 1/2(1-2\theta)$, alors il existe C et D deux réels tels que

$$|||e^n(u)|||_2\leqslant T\Big[C|1-2\theta|\delta t+D(\delta t^2+\delta x^2)\Big],\quad 0\leqslant n\leqslant N$$

R Pour $\theta=1/2$, le schéma est inconditionnellement convergent et d'ordre 2.

2.6 Conditions aux limites transparentes

Nous avons considéré jusque là des conditions aux limites standards (Dirichlet ou Neumann) dans le cas où $x\in\Omega\subset\mathbb{R}^d$. Dans le cas où $x\in\mathbb{R}^d$ et en supposant que le terme source f soit égal à 0, nous avons construit une solution qui s'écrit

$$u(t,x)=\frac{1}{(2\pi)^{d/2}}\int_{\mathbb{R}^d}e^{i\langle x,\xi\rangle}e^{-|\xi|^2t}\widehat{u^i}(\xi)\,d\xi=\frac{1}{(4\pi t)^{d/2}}\int_{\mathbb{R}^N}e^{-|x-y|^2/4t}u^i(y)\,dy,\quad x\in\mathbb{R}^d,\ t>0.$$

Dans le cas où $x\in]0,1[$ et v satisfait des conditions aux limites de Dirichlet homogènes, nous avons pu représenter la solution dans la base hilbertienne $\{\sqrt{2}\sin(k\pi\cdot)\}_{k\in\mathbb{N}*}$ de $L^2(]0,1[)$. On a trouvé

$$v(t,x)=\sqrt{2}\sum_{k\in\mathbb{N}*}\hat{v}(k)(t)\sin(k\pi x)\quad\text{dans } L^2(]0,1[).$$

avec

$$\hat{v}(k)(t)=\widehat{u^i}(k)e^{-\kappa k^2\pi^2t}.$$

On pourrait établir un résultat similaire pour $x\in]-L/2,L/2[$, $L>0$. La question est de savoir si on peut espérer avoir v proche de la restriction de u à $]-L/2,L/2[$. A moins d'avoir $L=+\infty$, la réponse est évidemment négative et il n'est donc pas possible de penser que v est une approximation de u. On est donc amené à rechercher des conditions aux limites telles que la solution de

$$\begin{cases}\partial_t v=\partial_x^2 v & (t,x)\in[0,T]\times]-L/2,L/2[,\\ v(0,x)=u^i(x), & x\in]-L/2,L/2[,\\ Bv(t,x)=0, & t\in[0,T],\ x\in\{-L/2,L/2\},\end{cases}$$

soit telle que

$$v(t,x)=u(t,x),\quad (t,x)\in[0,T]\times]-L/2,L/2[.$$

Ici, B est un opérateur encodant les conditions aux limites. Il est possible de construire B en factorisant l'opérateur $\partial_t-\partial_x^2$ qui s'applique à la fois à u et v. Le calcul suivant peut être rendu rigoureux, mais dépasse le cadre de ce cours. On a

$$\partial_t-\partial_x^2=(\sqrt{\partial_t}-\partial_x)(\sqrt{\partial_t}+\partial_x).$$

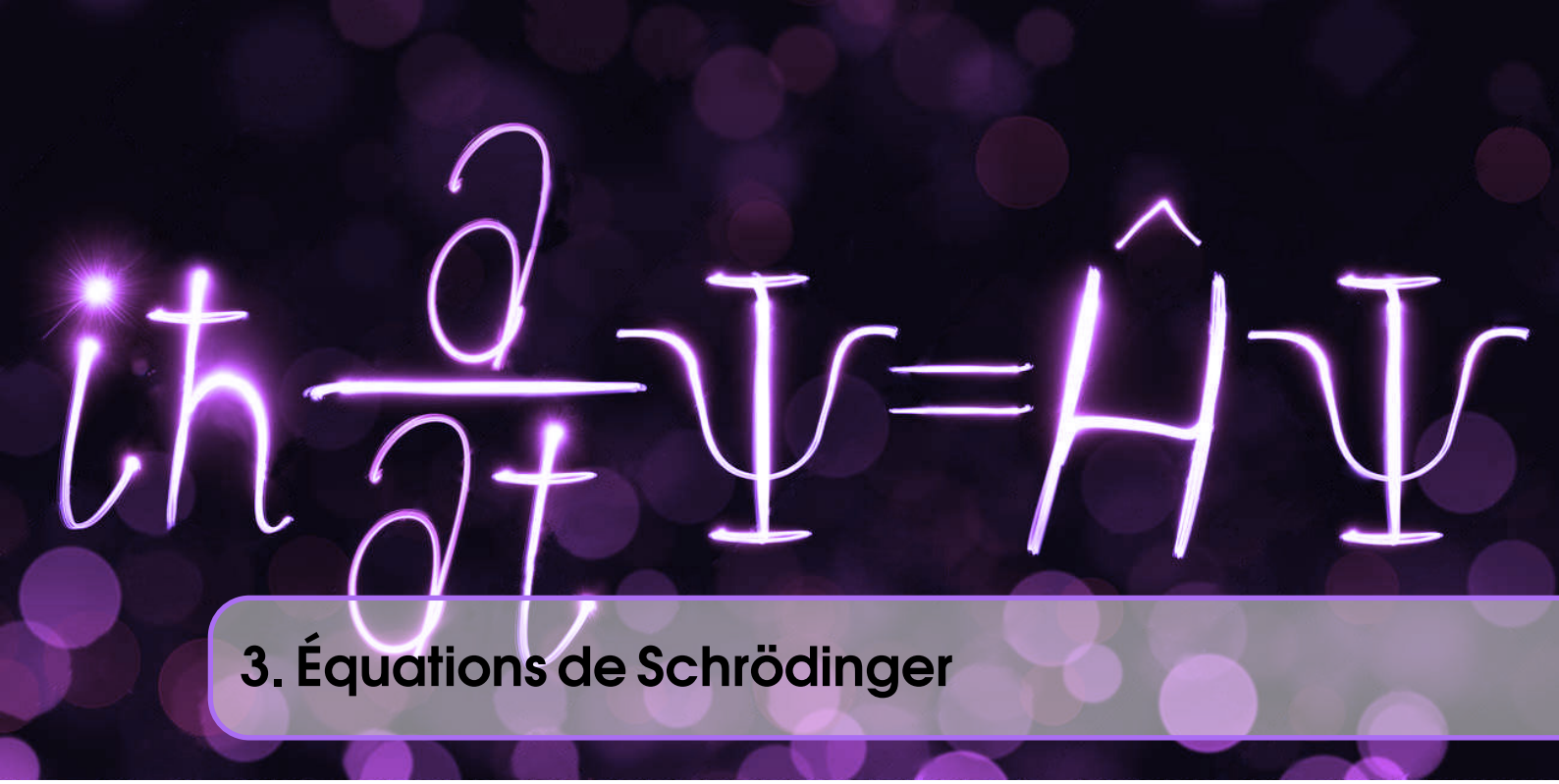
On en déduit la condition aux limites transparentes

$$Bv(t,x) = \partial_n v + \partial_t^{1/2} v = 0, \quad (t,x) \in [0,T] \times \{-L/2, L/2\},$$

où $\partial_n = \partial_x$ si $x = L/2$ et $\partial_n = -\partial_x$ si $x = -L/2$. On définit l'opérateur

$$\sqrt{\partial_t} f = \partial_t^{1/2} f = \frac{1}{\sqrt{\pi}} \partial_t \int_0^t \frac{f(s)}{\sqrt{t-s}}\, ds.$$

C'est un opérateur non local, dit pseudo-différentiel. Il faut donc pour calculer la solution procéder à une convolution en temps avec la trace de v le long des droites $(t,x) \in [0,T] \times \{\pm L/2\}$. Il faut donc garder trace de la solution à la frontière à tous les temps pour calculer la solution. On peut montrer que cette condition aux limites conduit bien à la propriété recherchée si la donnée initiale est à support compact dans $]-L/2, L/2[$.



3. Équations de Schrödinger

3.1 Problème modèle

On sait depuis le début du siècle dernier qu'il existe un principe de dualité onde-corpuscule. C'est Louis de Broglie qui le premier affirma que toute matière (et pas seulement la lumière) a une nature ondulatoire. Par la suite, Erwin Schrödinger proposa une équation pour décrire l'évolution spatio-temporelle d'une fonction d'onde associée à la particule. Il rechercha une équation pour une onde harmonique

$$\psi(t,\mathbf{x}) = \psi_0 e^{i(\mathbf{k}\cdot\mathbf{x}-\omega t)}, \qquad \psi_0 \in \mathbb{R},\ i \in \mathbb{C},\ i^2 = -1, \quad \mathbf{x} \in \mathbb{R}^3.$$

L'équation porte aujourd'hui son nom

Équation de Schrödinger linéaire pour une particule décrivant l'évolution de la fonction d'onde ψ

$$i\hbar\partial_t\psi = -\frac{\hbar^2}{2m}\Delta\psi + V(\mathbf{x})\psi, \tag{3.1}$$

où $\Delta\psi = \partial^2_{x_1} + \partial^2_{x_2} + \partial^2_{x_3}$, $h = 2\pi\hbar = 6.62 \cdot 10^{-34} J.s$ est la constante de Planck, m la masse de la particule et V un potentiel dans lequel évolue la particule.

Les physiciens doutaient de l'équation établie par Schrödinger car ils ne pouvaient pas trouver de moyens de récupérer les variables physiques. La question fondamentale était « comment trouver la particule à partir de sa fonction d'onde ». La réponse a été donnée par Born en 1926 : la fonction d'onde elle-même n'a pas de réalité physique, mais le carré de son module $|\psi|^2 = \overline{\psi}\psi$ est une densité de probabilité. La quantité $|\psi(t,\mathbf{x})|^2\, d\mathbf{x}$ représente la probabilité de trouver la particule dans une boule de centre $\mathbf{x}$ et de rayon $d\mathbf{x}$. Par conséquent, la probabilité de trouver la particule dans tout l'espace est de 100%, ce qui en mathématiques se lit

$$\int_{\mathbb{R}^3} |\psi(t,\mathbf{x})|^2\, d\mathbf{x} = 1. \tag{3.2}$$

Cette relation est connue comme la "conservation de la masse". On peut obtenir la conservation de la masse directement depuis l'équation. On multiplie l'équation de Schrödinger (3.1) par $\overline{\psi}$, on intègre en espace puis on prend ensuite la partie réelle de l'équation résultante.

Une autre quantité conservée importante est l'énergie obtenue à partir de la relation

$$\mathscr{E}(\psi)(t) = \int_{\mathbb{R}^3} \frac{\hbar^2}{2m} |\nabla\psi|^2 + V(\mathbf{x})|\psi|^2 \, d\mathbf{x}. \tag{3.3}$$

Pour obtenir cette relation, on multiplie l'équation de Schrödinger (3.1) par $\overline{\partial_t \psi}$, on intègre en espace puis on prend ensuite la partie imaginaire de l'équation résultante.

Les deux relations de conservation de la masse et de l'énergie ont un sens mathématique si $\psi \in H^1(\mathbb{R}^3)$.

Après adimensionnement, on convient d'étudier l'équation

$$\begin{cases} i\partial_t \psi + \Delta\psi = V(\mathbf{x})\psi, \quad x \in \mathbb{R},\, t > 0, \\ \psi(0,x) = \psi^i(x), \quad x \in \mathbb{R}. \end{cases} \tag{3.4}$$

3.2 Équations aux dérivées partielles dispersives

Soit L un opérateur différentiel à coefficients constants

$$Lu(x) = \sum_{|\alpha| \leqslant k} c_\alpha \partial_x^\alpha u(x), \quad k \geqslant 1,\, k \in \mathbb{N},$$

où $\alpha \in \mathbb{N}^d$ est un multi-entier avec $|\alpha| = \alpha_1 + \cdots + \alpha_d$ et

$$\partial_x^\alpha = \partial_x^{\alpha_1} \cdots \partial_x^{\alpha_d}.$$

On pourrait définir les équations aux dérivées partielles paraboliques par

$$\begin{cases} \partial_t u(t,x) = Lu(t,x), \\ u(0,x) = u^i(x), \end{cases}$$

où on demande que l'opérateur L soit auto-adjoint ($L = L^*$). Sur le même modèle, on définit les EDP dispersives.

Définition 3.2.1 Une équation aux dérivées partielles dispersives à coefficients constants

$$\begin{cases} \partial_t u(t,x) = Lu(t,x), \\ u(0,x) = u^i(x), \end{cases}$$

est dispersive si

$$u : \mathbb{R} \times \mathbb{R}^d \to V, \quad V \text{ un espace de Hilbert}$$

et si L est un opérateur anti-adjoint $L^* = -L$ ou encore iL est auto-adjoint.

■ **Exemple 3.1** Considérons l'équation de Airy $\partial_t u + \partial_x^3 u = 0$. L'opérateur est $L = -\partial_x^3$. Considérons que $V = H^3(R)$. Alors

$$\langle Lu, v\rangle = \int_{\mathbb{R}} -\partial_x^3 u \, v \, dx = \int_{\mathbb{R}} u \partial_x^3 v \, dx = -\langle u, Lv\rangle = \langle u, -Lv\rangle = \langle u, L^* v\rangle$$

d'où $L^* = -L$ qui est bien anti-adjoint.

Considérons maintenant l'équation de Schrödinger linéaire $\partial_t u - i\partial_x^2 u = 0$. On a $L = i\partial_x^2$ et

$$\langle Lu, v\rangle = \int_{\mathbb{R}} i\partial_x^2 u \, \bar{v} \, dx = \int_{\mathbb{R}} u \left(i\partial_x^2 \bar{v}\right) \, dx.$$

Mais, pour deux nombres complexes z_1 et z_2, on a $\overline{z_1 z_2} = \bar{z}_1 \bar{z}_2$. Ainsi

$$\langle Lu, v\rangle = \int_{\mathbb{R}} u\overline{-i\partial_x^2 \bar{v}}\, dx = \langle u, L^* v\rangle.$$

Ainsi, $L^* = -i\partial_x^2 = -L$ et L est anti-adjoint.

■

3.3 Solutions classiques dans $\mathbb{R}^n$

On suppose $V = 0$ et on procède comme pour l'équation de la chaleur en utilisant la transformée de Fourier. L'équation (3.4) devient

$$\begin{cases} i\partial_t \hat{\psi} - |\xi|^2 \hat{\psi} = 0, & \xi \in \mathbb{R},\ t > 0, \\ \hat{\psi}(0,\xi) = \widehat{\psi^i}(\xi), & \xi \in \mathbb{R}. \end{cases}$$

Ainsi,

$$\hat{\psi}(t,\xi) = e^{-i|\xi|^2 t}\widehat{\psi^i}(\xi), \quad \xi \in \mathbb{R}^N,\ t \geqslant 0.$$

On définit ainsi pour t fixé

$$\hat{S}(t,\cdot) : \xi \mapsto e^{-i|\xi|^2 t}.$$

Soit $\hat{S}(t)$ l'opérateur de multiplication dans $\mathscr{S}'(\mathbb{R}^N)$ par $\hat{S}(t,\cdot)$. Alors, la famille $\{\hat{S}(t)\}_{t\in\mathbb{R}}$ est un groupe dans $\mathscr{S}'(\mathbb{R}^N)$. En effet, contrairement à l'équation de la chaleur où $\widehat{g(t)}(\xi) = e^{-|\xi|^2 t}$ était non borné pour $t < 0$, l'opérateur $e^{-i|\xi|^2 t}$ est borné pour tout $t \in \mathbb{R}$. On a donc un opérateur réversible ce qui permet à $\{\hat{S}(t)\}_{t\in\mathbb{R}}$ d'être un groupe

$$\psi^i \xrightarrow{S(t)} \psi(t) \xrightarrow{S(-t)} \psi^i, \quad t \geqslant 0.$$

Si $\widehat{\psi^i} \in L^2(\mathbb{R}^N)$ au lieu de $\mathscr{S}'(\mathbb{R}^N)$, alors $\hat{\psi}(t,\xi) = e^{-i|\xi|^2 t}\widehat{\psi^i}(\xi) \in L^2$ et on a

$$\|\hat{\psi}\|_{L^2} = \|\widehat{\psi^i}\|_{L^2}$$

ce qui par Parseval traduit la conservation de la masse. Ainsi, $\{\hat{S}(t)\}_{t\in\mathbb{R}}$ est un groupe unitaire sur L^2.

En revenant dans $\mathscr{S}'(\mathbb{R}^N)$, on a

$$\psi(t,x) = \mathscr{F}^{-1}\left(\hat{S}(t,\cdot)\widehat{\psi^i}\right) = \mathscr{F}^{-1}\left(\hat{S}(t,\cdot)\right) * \psi^i = S(t,\cdot) * \psi^i.$$

On montre, en faisant les mêmes manipulations que pour l'équation de la chaleur, que

$$S(t,\cdot) = \frac{1}{(4\pi i t)^{N/2}} e^{-\frac{|x|^2}{4it}}.$$

Si on pose $G(t) = S(t,\cdot)*$, alors les opérations de convolution $\{G(t)\}_{t\in\mathbb{R}}$ constituent un groupe dans $\mathscr{S}$ et $\mathscr{S}'$, et un groupe unitaire sur L^2.

R Contrairement à l'équation de la chaleur, $G(t)$ appliqué à ψ^i n'a pas d'effet régularisant.

On a donc

$$\psi(t,x) = \frac{1}{(4\pi i t)^{N/2}} \int_{\mathbb{R}^N} e^{-\frac{|x-y|^2}{4it}} \psi^i(y)\, dy.$$

Si $\psi^i \in L^1(\mathbb{R}^N)$, on a immédiatement par cette formule que $\psi(t,\cdot) \in L^\infty(\mathbb{R}^N)$ et on a

$$\|\psi(t,\cdot)\|_{L^\infty} \leqslant Ct^{-N/2}\|\psi^i\|_{L^1}.$$

La norme L^∞ de ψ décroît donc avec le temps alors que la norme L^2 est conservée. De manière générale, si on considère deux entiers p et q, $2 \leqslant p \leqslant \infty$, et $1/p + 1/q = 1$, alors il existe un constante $c > 0$ tel que pour tout $\psi^i \in L^q(\mathbb{R}^N)$, $\psi \in L^p(\mathbb{R}^N)$ et

$$\|\psi(t,\cdot)\|_{L^p} \leqslant Ct^{-N(1/2-1/p)}\|\psi^i\|_{L^q}.$$

3.4 Schémas numériques

On considère à nouveau le cas où $V = 0$.

La méthode basée sur la transformée de Fourier est comme pour l'équation de la chaleur très efficace pour l'équation de Schrödinger non linéaire et par Parseval, on a automatiquement la conservation de la norme L^2 (masse). La modification dans la formule (2.6) définissant le schéma est mineure puisqu'il suffit de rajouter un i dans la formule qui devient

$$U(t) = \mathscr{F}_N^{-1}\left(e^{-i\frac{4\pi^2 j^2}{a^2}t}\mathscr{F}_N\left(u^0\right)\right).$$

Le θ-schéma est également une méthode très populaire et on a les mêmes résultats de convergence et de stabilité. Il convient en plus de s'assurer de la conservation de la masse et de l'énergie. Prenons le cas $\theta = 1/2$ et démontrons ces résultats. Nous considérons l'équation linéaire de Schrödinger sur un domaine régulier borné

Équation de Schrödinger linéaire (LS)

$$\begin{cases} i\partial_t\psi = -\dfrac{1}{2}\Delta\psi, & t > 0,\ \mathbf{x} \in \Omega \in \mathbb{R}^d, \\ \psi(0,\mathbf{x}) = \psi_0(\mathbf{x}), & \mathbf{x} \in \Omega, \\ \text{CL sur } \partial\Omega, & t > 0. \end{cases} \tag{3.5}$$

Nous supposons dans ce paragraphe que $u(t,\mathbf{x}) = 0$ pour $\mathbf{x} \in \partial\Omega$. Soit $k = \delta t$ le pas de temps constant et $t_n = nk$. Désignons par $\psi^n(x)$ une approximation de la solution exacte $\psi(t_n, x)$.

Le schéma **semi-discret de Crank-Nicolson** s'écrit

$$\begin{cases} i\dfrac{\psi^{n+1}-\psi^n}{\delta t} = -\dfrac{1}{2}\Delta\dfrac{\psi^{n+1}+\psi^n}{2}, & \mathbf{x} \in \Omega, \\ \psi^0(\mathbf{x}) = \psi_0(\mathbf{x}), & \mathbf{x} \in \Omega. \end{cases} \tag{3.6}$$

Proposition 3.4.1 Le schéma Crank-Nicolson du second ordre (3.6) préserve la norme L^2 (alias la masse).

Démonstration. Nous devons imiter la preuve de l'équation continue. Multiplions le schéma par $(\overline{\psi^{n+1}+\psi^n})/2$. On a

$$i\frac{\psi^{n+1}-\psi^n}{\delta t}\frac{\overline{\psi^{n+1}+\psi^n}}{2} = -\frac{1}{2}\Delta\frac{\psi^{n+1}+\psi^n}{2}\frac{\overline{\psi^{n+1}+\psi^n}}{2},$$

ce qui s'écrit aussi

$$\frac{i}{2\delta t}(|\psi^{n+1}|^2 - |\psi^n|^2 + 2i\mathrm{Im}(\psi^n\overline{\psi^{n+1}})) = -\frac{1}{2}\Delta\frac{\psi^{n+1}+\psi^n}{2}\frac{\overline{\psi^{n+1}+\psi^n}}{2}.$$

En intégrant cette équation par rapport à $\mathbf{x}$ donne

$$\begin{aligned}\frac{i}{2\delta t}\int_\Omega(|\psi^{n+1}|^2 - |\psi^n|^2 + 2i\mathrm{Im}(\psi^n\overline{\psi^{n+1}}))d\mathbf{x} &= \frac{1}{2}\int_\Omega|\nabla\frac{\psi^{n+1}+\psi^n}{2}|^2 d\mathbf{x} \\ &\quad -\frac{1}{2}\int_{\partial\Omega}\partial_{\mathbf{n}}\frac{\psi^{n+1}+\psi^n}{2}\frac{\overline{\psi^{n+1}+\psi^n}}{2}d\sigma.\end{aligned}$$

Grâce aux conditions aux limites de Dirichlet, le dernier terme s'annule, et en prenant la partie imaginaire de l'équation résultante, on obtient

$$\int_\Omega |\psi^{n+1}|^2 d\mathbf{x} = \int_\Omega |\psi^n|^2 d\mathbf{x}.$$

■

Pour la conservation de l'énergie, on répète les opérations mais en multipliant par $(\overline{\psi^{n+1} - \psi^n})/\delta t$.

Nous pouvons maintenant faire une approximation de la variable spatiale. Nous simplifions la présentation en considérant le cas 1D et $\Omega = (x_\ell, x_r)$. Soit $h = \delta x = (x_r - x_\ell)/(J+1)$ la taille du maillage et définissons les nœuds $x_j = jh + x_\ell$, $j = 0, \cdots, J+1$. On écrit ψ_j^n l'approximation de $\psi(t_n, x_j)$. Nous approximons l'opérateur ∂_x^2 par la différence finie standard du second ordre

$$\partial_x^2 \psi(t_n, x_j) \approx \frac{\psi_{j+1}^n - 2\psi_j^n + \psi_{j-1}^n}{\delta x^2}, \quad 1 \leqslant j \leqslant J.$$

Le schéma de Crank-Nicolson discret complet est

$$\begin{cases} i\dfrac{\psi_j^{n+1} - \psi_j^n}{\delta t} = -\dfrac{1}{4}\left(\dfrac{\psi_{j+1}^{n+1} - 2\psi_j^{n+1} + \psi_{j-1}^{n+1} + \psi_{j+1}^n - 2\psi_j^n + \psi_{j-1}^n}{\delta x^2}\right), \quad n \geqslant 0,\ 1 \leqslant j \leqslant J, \\ \psi_0^{n+1} = \psi_{J+1}^{n+1} = 0. \end{cases} \tag{3.7}$$

On peut simplifier le schéma en définissant l'opérateur Laplacien discret

$$\Delta_h v_j = \frac{v_{j+1} - 2v_j + v_{j-1}}{\delta x^2}.$$

Le schéma (3.7) devient

$$i\frac{\psi_j^{n+1} - \psi_j^n}{\delta t} = -\frac{1}{4}\Delta_h \frac{\psi_j^{n+1} + \psi_j^n}{2}, \quad 1 \leqslant j \leqslant J.$$

En utilisant la matrice D_2

$$D_2 = \frac{1}{h^2}\begin{pmatrix} 2 & -1 & & & & \\ -1 & 2 & -1 & & & \\ & -1 & 2 & -1 & & \\ & & \ddots & \ddots & \ddots & \\ & & & -1 & 2 & -1 \\ & & & & -1 & 2 \end{pmatrix}, \tag{3.8}$$

le système linéaire complet à résoudre est

$$i\frac{\Psi^{n+1} - \Psi^n}{\delta t} = -\frac{1}{4} D_2 \frac{\Psi^{n+1} + \Psi^n}{2}, \quad \Psi^n = (\psi_j^n)_{1 \leqslant j \leqslant J} \in \mathbb{C}^J.$$

Ceci est équivalent à

$$\left[I - \frac{i\delta t}{4} D_2\right] \Psi^{n+1} = \left[I + \frac{i\delta t}{4} D_2\right] \Psi^n,$$

où I désigne la matrice identité.

Comme pour l'équation de la chaleur, nous avons le résultat de convergence suivant

Théorème 3.4.2 Soit ψ une solution régulière à (3.5) sur $(0,T)$ avec des conditions aux limites de Dirichlet homogènes et Ψ^n, $n = 0, \cdots, N$, $T = N\delta t$, la solution au schéma Crank-Nicolson totalement discret (3.7) avec $\Psi^0 = (\psi_j^0)_{j=1}^J$, $\psi_j^0 = \psi_0(x_j)$. Alors, pour δt assez petit,

$$\max_{1\leqslant n\leqslant N} \|\psi^n - \Psi^n\|_h \leqslant c(h^2 + k^2)$$

où $k = \delta t$ et $h = \delta x$, c étant une constante indépendante de k et h.

Ici, $\psi^n = (\psi(t_n, x_j))_{j=1}^J$ et les normes discrètes sont données par

$$\|v\|_h = \langle v, v\rangle_h^{1/2}, \quad \langle v, w\rangle_h = h\sum_{j=1}^J v_j\overline{w}_j, \quad |v|_{1,h} = \left(h\sum_{j=0}^J \left|\frac{v_{j+1} - v_j}{h}\right|^2\right)^{1/2}.$$

Démonstration. Soit $r^n \in \mathbb{C}_0^{J+2} = \{v = (v_0, \cdots, v_{J+1}) \in \mathbb{C}^{J+2},\ v_0 = v_{J+1} = 0\}$ l'erreur de consistance du schéma Crank-Nicolson. Avec les développements de Taylor pour k et h petit, nous obtenons facilement

$$r_j^n = \frac{\psi_j^{n+1} - \psi_j^n}{k} - i\Delta_h\frac{\psi_j^{n+1} + \psi_j^n}{2}, \quad j = 1, \cdots, J,$$

et

$$\max_{j,n} |r_j^n| \leqslant C(k^2 + h^2).$$

Soit $e^n = \psi^n - \Psi^n$, $n = 0, \cdots, N$. Nous avons, puisque tout est linéaire,

$$\frac{e_j^{n+1} - e_j^2}{k} = i\Delta_h\frac{e_j^{n+1} + e_j^n}{2} + r_j^n, \quad j = 1, \cdots, J$$

Il est facile de montrer que $-\langle\Delta_h v, v\rangle_h = |v|_{1,h}^2$. En prenant le produit scalaire avec $(e_j^{n+1} + e_j^n)/2$, on obtient

$$\langle\frac{e^{n+1} - e^n}{k}, \frac{e^{n+1} + e^n}{2}\rangle_h = -i\left|\frac{e^{n+1} + e^n}{2}\right|_{1,h} + \langle r^n, \frac{e^{n+1} + e^n}{2}\rangle_h,$$

où $e^n = (e_j^n)_j$ et $r^n = (r_j^n)_j$. Prenant la partie réelle,

$$\|e^{n+1}\|_h^2 - \|e^n\|_h^2 = 2k\,\mathrm{Re}\langle r^n, \frac{e^{n+1} + e^n}{2}\rangle_h.$$

Par l'inégalité de Schwartz,

$$\|e^{n+1}\|_h^2 - \|e^n\|_h^2 \leqslant Ck\|r^n\|_h(\|e^{n+1}\|_h + \|e^n\|_h),$$

qui se lit également comme suit

$$(\|e^{n+1}\|_h + \|e^n\|_h)(\|e^{n+1}\|_h - \|e^n\|_h) \leqslant Ck(h^2 + k^2)(\|e^{n+1}\|_h + \|e^n\|_h).$$

Cela donne immédiatement

$$\|e^{n+1}\|_h \leqslant \|e^n\|_h + Ck(h^2 + k^2)$$

ou encore

$$\|e^n\|_h \leqslant \underbrace{\|e^0\|_h}_{=0} + \underbrace{\sum_{\ell=0}^{n-1} Ck(h^2 + k^2)}_{\leqslant CT(h^2+k^2)}.$$

■

3.5 Équation de Schrödinger non linéaire

Lorsque l'on considère plusieurs particules bosoniques en interaction, il est possible de montrer qu'à température proche de zéro, l'ensemble des particules se comporte comme une particule géante vérifiant une équation de Schrödinger non linéaire cubique. Cet ensemble de particule s'appelle un condensat de Bose-Einstein (voir [2]).

Équation de Schrödinger non linéaire cubique

$$\begin{aligned} i\partial_t\psi(t,\mathbf{x}) &= \left[-\frac{1}{2}\Delta + V(\mathbf{x}) + \beta|\psi(t,\mathbf{x})|^2\right]\psi(t,\mathbf{x}), \\ E(\psi) &= \int_{\mathbb{R}^3} \frac{1}{2}|\nabla\psi|^2 + V(\mathbf{x})|\psi|^2 + \frac{\beta}{2}|\psi|^4\, d\mathbf{x}, \end{aligned} \tag{3.9}$$

Comme pour l'équation de Schrödinger linéaire, la masse est également conservée $N(t) = N(\psi(t,\cdot)) = \int_{\mathbb{R}^d} |\psi(t,\mathbf{x})|^2 d\mathbf{x} = \int_{\mathbb{R}^d} |\psi_0(x)|^2 d\mathbf{x} = 1,\ t \geqslant 0.$

R L'équation de Schrödinger non linéaire cubique intervient également dans la modélisation en optique non linéaire pour décrire la propagation de la lumière. Elle apparaît également en mécanique des fluides.

3.5.1 Schéma de splitting - pas fractionnaire

Afin de présenter l'idée, nous allons considérer le cas simple des équations différentielles ordinaires linéaires.

L'idée est simple : diviser pour mieux régner. La méthode du splitting est probablement la technique la plus systématique pour le développement d'algorithmes en général pour les systèmes **autonomes**.

Supposons que nous souhaitions résoudre $z'(t) = f(z) = f_1(z) + f_2(z)$ où chacun des deux ODE est complètement intégrable $z'(t) = f_1(z)$ et $z'(t) = f_2(z)$. La méthode consiste à

- suivre le flot associé à f_1 à partir de z_n. Résoudre $z'(t) = f_1(z(t))$, $z(0) = z_n$ et le faire évoluer exactement pour l'étape h : $z_* = z_1(h, z_n)$.
- suivre le flot associé à f_2 à partir de z_*. Résoudre $z'(t) = f_2(z(t))$, $z(0) = z_*$ et le faire évoluer exactement pour l'étape h : $z_{n+1} = z_2(h, z_*)$.

En désignant les applications flot des champs vectoriels f_1 et f_2 par φ_{t,f_1} et φ_{t,f_2} , nous venons de calculer

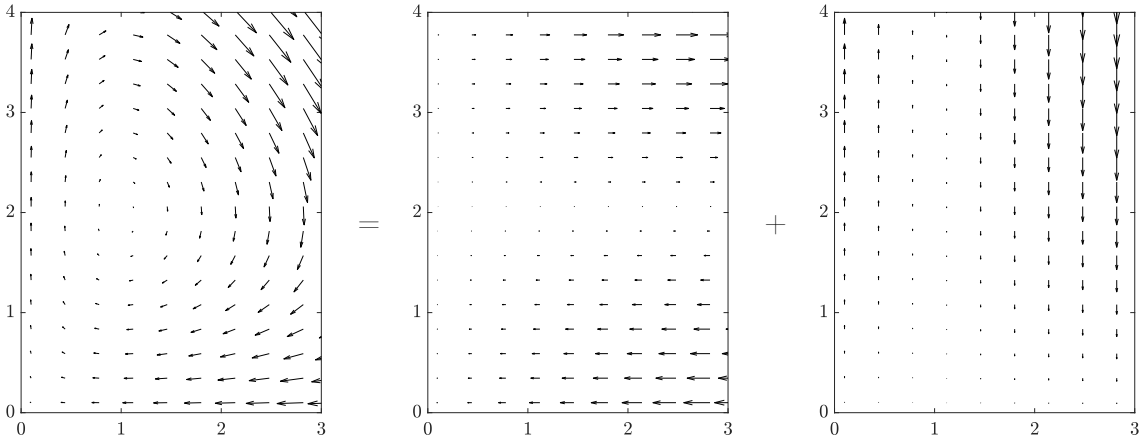
$$z_{n+1} = \varphi_{h,f_2}(\varphi_{h,f_1}(z_n)) = (\varphi_{h,f_2} \circ \varphi_{h,f_1})(z_n).$$

Une telle méthode est appelée méthode de pas fractionnaire (splitting), et celle que nous venons de présenter est connue sous le nom de schéma de splitting de *Lie-Trotter*.

■ **Exemple 3.2** Si nous considérons le modèle de Lotka-Volterra (3.10)

$$\begin{aligned} u' &= u(v-2), \\ v' &= v(1-u). \end{aligned} \tag{3.10}$$

alors $f(u,v) = (u(v-2), v(1-u))$ et une décomposition pourrait être $f_1(u,v) = (u(v-2), 0)$ et $f_2(u,v) = (0, v(1-u))$. L'image des champs vectoriels associés est donnée ci-dessous



Champ vectoriel f = Champ vectoriel f_1 + Champ vectoriel f_2.

■

Il est plus simple de comprendre la méthode sur les EDO linéaires. Considérons $y' = Ay$, $y(0) = y_0$ avec $y \in \mathbb{R}^d$, $A \in M_{d,d}(\mathbb{R})$. La solution est donc donnée par $y(h) = e^{Ah}y_0 := \varphi_h y_0$, où l'exponentielle de la matrice Ah est donnée par

$$e^{Ah} = \sum_{j \geqslant 0} \frac{(hA)^j}{j!}.$$

Supposons que $A = B + C$, donc $y' = (B + C)y$ et le schéma de splitting de Lie s'écrit

$$\begin{cases} y_1' = By_1, & y_1(0) = y_0, \\ y_2' = Cy_2, & y_2(0) = y_1(h), \end{cases}$$

et nous obtenons donc

$$y_2(h) = e^{Ch}e^{Bh}y_0 := \phi_y y_0.$$

Calculons l'erreur locale.

$$\begin{aligned} \phi_h(y_0) - \varphi_h(y_0) &= e^{Ch}e^{Bh}y_0 - e^{Ah}y_0 \\ &= (I + Ch + \frac{(Ch)^2}{2} + O(h^3))(I + Bh + \frac{(Bh)^2}{2} + O(h^3))y_0 \\ &\quad -(I + Ah + \frac{(Ah)^2}{2} + O(h^3))y_0 \\ &= (I + (B + C)h + \frac{h^2}{2}(B^2 + 2CB + C^2) + O(h^3))y_0 \\ &\quad -(I + Ah + \frac{(Ah)^2}{2} + O(h^3))y_0 \\ &= \frac{h^2}{2}(B^2 + 2CB + C^2 - (B + C)^2)y_0 + O(h^3) \\ &= \frac{h^2}{2}(CB - BC)y_0 + O(h^3) \\ &= \frac{h^2}{2}[C, B]y_0 + O(h^3), \end{aligned}$$

où $[C, B]$ est le commutateur des deux matrices B et C. Si le commutateur est nul, on dit que les deux matrices commutent. Sinon, l'erreur locale est $O(h^2)$ et l'erreur globale est $O(h)$.

On peut construire aisément un splitting du second ordre, connu comme le *splitting de Strang*, par

$$\phi_h = e^{Ch/2}e^{Bh}e^{Ch/2}.$$

Pour plus d'informations concernant les méthodes de splitting pour les équations différentielles ordinaires, on pourra consulter [7].

Considérons maintenant l'équation de Schrödinger en vue d'appliquer le splitting de Strang

$$i\partial_t\psi + \frac{1}{2}\Delta\psi = \beta|\psi|^{p-1}\psi + V(\mathbf{x})\psi,$$

Nous découpons l'équation en deux parties

$$i\partial_t v + \frac{1}{2}\Delta v = 0$$

et

$$i\partial_t w = \beta|w|^{p-1}w + V(\mathbf{x})w.$$

Nous avons maintenant à considérer une EDP et une EDO. La solution de l'EDO est en fait explicite. En effet, si nous multiplions par $\bar{w}$, nous obtenons

$$\partial_t w\bar{w} = -i\beta|w|^{p+1} - iV(\mathbf{x})|w|^2.$$

On sait que

$$\frac{d}{dt}|w|^2 = 2\mathrm{Re}(\partial_t w\bar{w}).$$

Ainsi, en prenant la partie réelle des l'équation précédente, nous avons

$$\frac{1}{2}\frac{d}{dt}|w|^2 = 0$$

et le module de $|w|$ est ainsi conservé. Nous obtenons donc une solution explicite

$$w(t,\mathbf{x}) = e^{-i[\beta|w(0,\mathbf{x})|^{p-1}+V(\mathbf{x})]t}w(0,\mathbf{x}).$$

Le splitting de Strang du temps t_n au temps t_{n+1} est donné par les trois étapes

Une étape du splitting de Strang
1. Calculer $u_1 = e^{-i[\beta|\psi^n|^{p-1}+V(\mathbf{x})]/\delta t/2}\psi^n$.
2. Résoudre $i\partial_t u_2 + \Delta u_2/2 = 0$ avec $u_2(t=0) = u_1$ sur un intervalle de longueur δt soit par DST ou FFT ou une méthode de différences finies, éléments finis ou volumes finis.
3. Calculer $\psi^{n+1} = e^{-i[\beta|u_2|^{p-1}+V(\mathbf{x})]/\delta t/2}u_2$.

Le splitting de Strang a une structure intéressante. En effet, pour chaque pas, nous avons
1. $\|u_1\|_{L^2} = \|\psi^n\|_{L^2}$ puisque l'opérateur en jeu est unitaire.
2. $\|u_2\|_{L^2} = \|u_1\|_{L^2}$ par la propriété de l'équation de Schrödinger linéaire.
3. $\|\psi^{n+1}\|_{L^2} = \|u_2\|_{L^2}$.

Ainsi, la masse est automatiquement conservée. De plus, si nous partons de ψ^{n+1} et appliquons le schéma avec un pas de temps inversé $(-\delta t)$ et en prenant le conjugué de ψ^n, le schéma est réversible dans le temps. Une autre propriété très intéressante est que le schéma avec DST ou FFT est explicite (implicite avec la méthode des différences finies et le schéma de Crank-Nicolson, voir le paragraphe 3.5.2). L'énergie n'est cependant pas préservée. Nous pouvons obtenir un schéma d'ordre supérieur assez facilement (en temps) en prenant un splitting temporel d'ordre supérieur.

La preuve de convergence pour les splittings de Lie et de Strang est disponible dans [4] et [10].

3.5.2 Le schéma de Crank-Nicolson

Le schéma que nous présentons ici a été introduit pour la première fois par Delfour, Fortin et Payre [6] suite à une idée de Strauss et Vazquez [11] pour l'équation de Schrödinger non linéaire (NLS) suivante

$$i\partial_t\psi = -\frac{1}{2}\Delta\psi + V(\mathbf{x})\psi + \beta|\psi|^{p-1}\psi.$$

Il s'appuie sur deux remarques différentes.

R Supposons que $G'(v) = |v|^{p-1}p$, alors si $v \in \mathbb{R}$,

$$G(v) = \frac{|v|^{p+1}}{p+1}.$$

Si $v \in \mathbb{C}$, $G(v) = |v|^{2\sigma+2} = (|v|^2)^{\sigma+1}$, $d_v(|v|^2)(h) = d_v(v\bar{v})(h) = 2\mathrm{Re}(\bar{v}h)$ et

$$d_v(G(v))(h) = 2(\sigma+1)|v|^{2\sigma}\mathrm{Re}(\bar{v}h).$$

R Les opérations effectuées pour montrer la conservation de la masse et de l'énergie sont les suivantes

- masse : multiplier (NLS) par $\bar{\psi}$, intégrer par rapport à x et prendre la partie imaginaire. La contribution du terme non linéaire est $\int|\psi|^{p-1}|\psi|^2 d\mathbf{x}$.
- énergie : multiplier (NLS) par $\overline{\partial_t\psi}$, intégrer par rapport à $\mathbf{x}$ et prendre la partie réelle. La contribution du terme non linéaire est

$$\mathrm{Re}\int|\psi|^{p-1}\psi\partial_t\bar{\psi}d\mathbf{x} = \frac{1}{p+1}\frac{d}{dt}\int|\psi|^{p+1}d\mathbf{x}.$$

Nous voulons construire un schéma numérique qui a les mêmes comportements.

La discrétisation de $|\psi|^{p-1}\psi = G'(\psi)$ est choisie telle que

$$\frac{G(\psi^{n+1}) - G(\psi^n)}{|\psi^{n+1}|^2 - |\psi^n|^2}(\psi^{n+1} + \psi^n).$$

Si $(a,b) \in \mathbb{R}^2$, alors

$$\frac{G(b) - G(a)}{|b|^2 - |a|^2}(b+a) = \frac{G(b) - G(a)}{b-a} = G'(\frac{a+b}{2}) + O((b-a)^2).$$

C'est une approximation de second ordre de G' au temps $t_{n+1/2} = (t_n + t_{n+1})/2$.

Si nous imitons la preuve des conservations de la masse et de l'énergie, nous avons

- Pour la conservation de la masse : $|\psi|^{p-1}\psi\bar{\psi}$ devient

$$\frac{G(\psi^{n+1}) - G(\psi^n)}{|\psi^{n+1}|^2 - |\psi^n|^2}(\psi^{n+1} + \psi^n)(\overline{\psi^{n+1} + \psi^n}) = \frac{G(\psi^{n+1}) - G(\psi^n)}{|\psi^{n+1}|^2 - |\psi^n|^2}|\psi^{n+1} + \psi^n|^2 \in \mathbb{R}$$

  et la partie imaginaire est nulle.
- Pour la conservation de l'énergie $|\psi|^{p-1}\psi\partial_t\bar{\psi}$ devient

$$\frac{G(\psi^{n+1}) - G(\psi^n)}{|\psi^{n+1}|^2 - |\psi^n|^2}(\psi^{n+1} + \psi^n)(\overline{\frac{\psi^{n+1} - \psi^n}{\delta t}})$$

  ce qui est

$$\frac{1}{\delta t}\frac{G(\psi^{n+1}) - G(\psi^n)}{|\psi^{n+1}|^2 - |\psi^n|^2}(|\psi^{n+1}|^2 - |\psi^n|^2 + 2i\mathrm{Im}(\psi^n\overline{\psi^{n+1}})).$$

En prenant la partie réelle, on a

$$\int_{\mathbb{R}^d} \frac{G(\psi^{n+1}) - G(\psi^n)}{\delta t} d\mathbf{x},$$

ce qui correspond à $\dfrac{1}{p+1}\dfrac{d}{dt}\displaystyle\int |\psi|^{p+1} d\mathbf{x}$.
La discrétisation du terme non linéaire est donc le bon choix pour garantir une approximation de second ordre avec conservation de la masse et de l'énergie.

Le schéma semi-discret de Crank-Nicolson semi-discret est

$$\begin{aligned} i\frac{\psi^{n+1}-\psi^n}{\delta t} + \frac{1}{2}\Delta\frac{\psi^{n+1}+\psi^n}{2} &= \beta\frac{G(\psi^{n+1})-G(\psi^n)}{|\psi^{n+1}|^2-|\psi^n|^2}(\psi^{n+1}+\psi^n) + V(\mathbf{x})\frac{\psi^{n+1}+\psi^n}{2} \\ &= \frac{\beta}{p+1}\frac{|\psi^{n+1}|^{p+1}-|\psi^n|^{p+1}}{|\psi^{n+1}|^2-|\psi^n|^2}(\psi^{n+1}+\psi^n) + V(\mathbf{x})\frac{\psi^{n+1}+\psi^n}{2}. \end{aligned}$$

Si $p-1 = 2\sigma$, le terme non linéaire dans l'équation (NLS) devient $|\psi|^{2\sigma}\psi$ ce qui est approximé par

$$\frac{\beta}{\sigma+1}\frac{|\psi^{n+1}|^{2\sigma+2}-|\psi^n|^{2\sigma+2}}{|\psi^{n+1}|^2-|\psi^n|^2}\frac{\psi^{n+1}+\psi^n}{2}.$$

■ **Exemple 3.3**

- $\sigma = 1$ (non linéarité cubique). Le terme non linéaire est approché par

$$\frac{1}{2}\frac{|\psi^{n+1}|^4-|\psi^n|^4}{|\psi^{n+1}|^2-|\psi^n|^2}\frac{\psi^{n+1}+\psi^n}{2} = \frac{|\psi^{n+1}|^2+|\psi^n|^2}{2}\frac{\psi^{n+1}+\psi^n}{2}.$$

- $\sigma = 2$ (non linéarité quintique). Puisque $(a^6-b^6)/(a^2-b^2) = a^4+a^2b^2+b^4$, le terme non linéaire est approché par

$$|\psi|^4\psi \approx \frac{|\psi^{n+1}|^4+|\psi^{n+1}|^2|\psi^n|^2+|\psi^n|^4}{3}\frac{\psi^{n+1}+\psi^n}{2}.$$

- Le cas général est

$$\frac{|\psi^{n+1}|^{2\sigma+2}-|\psi^n|^{2\sigma+2}}{|\psi^{n+1}|^2-|\psi^n|^2} = \sum_{p=0}^{\sigma}|\psi^{n+1}|^{2p}|\psi^n|^{2(\sigma-p)}.$$

■

On peut montrer que ce schéma est :

- réversible en temps,
- préserve la masse et l'énergie,
- est inconditionnellement stable (semi-implicite),

mais il est

- non explicite (il y a un système non linéaire à résoudre, par exemple par une méthode de point fixe),

En dimension 1, le système complètement discret s'écrit

$$i\frac{\psi_j^{n+1}-\psi_j^n}{\delta t}+\frac{1}{2}\Delta_h\frac{u_j^{n+1}+\psi_j^n}{2} = \beta\frac{G(\psi_j^{n+1})-G(\psi_j^n)}{|\psi^{n+1}|^2-|\psi_j^n|^2}(u_j^{n+1}+\psi_j^n)+V(x_j)\frac{\psi_j^{n+1}+\psi_j^n}{2}, \quad 1\leqslant j\leqslant J.$$

En définissant

$$\varphi(u_j^{n+1},\psi_j^n) = \frac{G(\psi_j^{n+1})-G(\psi_j^n)}{|\psi^{n+1}|^2-|\psi_j^n|^2}(u_j^{n+1}+\psi_j^n), \tag{3.11}$$

et en utilisant les notations vectorielles et matricielles, nous avons

$$iI\frac{\Psi^{n+1}-\Psi^n}{\delta t}+\frac{1}{4}D_2(\Psi^{n+1}+\Psi^n)=\beta\varphi(\Psi^{n+1},\Psi^n)+V\frac{\Psi^{n+1}+\Psi^n}{2},$$

où $\Psi^n=(\psi_j^n)_{j=1}^J$ et D_2 est définie par (3.8).

La méthode de point fixe appliquée au système non linéaire est donnée par l'algorithme suivant

Algorithm 1: Méthode de point fixe pour le schéma de Crank Nicolson

Data: $W_0=\Psi^n$, $s=0$, test=TRUE
while *test=TRUE* **do**
  $\left[i\frac{I}{\delta t}+\frac{1}{4}D_2-\frac{V}{2}\right]W_{s+1}=\left[i\frac{I}{\delta t}-\frac{1}{4}D_2+\frac{V}{2}\right]\Psi^n+\beta\varphi(W_s,\Psi^n)$;
  test=$\|W_{s+1}-W_s\|>\varepsilon$;
  $W_s=W_{s+1}$;
end
$\Psi^{n+1}=W_s$

Pour $v,w\in\mathbb{C}_0^{J+2}$, on définit les normes discrètes suivantes

$$\langle v,w\rangle_h=h\sum_{j=1}^J v_j\overline{w_j},$$
$$\|v\|_h=\langle v,w\rangle_h^{1/2}=\|v\|_{h,2}.$$

Il est possible de démontrer le résultat de convergence suivant (voir [1])

Théorème 3.5.1 Soit ψ une solution (assez régulière) de l'équation de Schrödinger non linéaire cubique 1d et Ψ^n solution du schéma de Crank-Nicolson et $k=o(h^{1/4})$. Alors, pour k suffisamment petit,

$$\max_{1\leqslant n\leqslant N}\|\psi(t_n)-\Psi^n\|_h\leqslant c(k^2+h^2). \tag{3.12}$$

3.5.3 Schéma de relaxation

Ce schéma, construit dans [3], est consacré à l'équation de Schrödinger non linéaire cubique

$$i\partial_t\psi+\Delta\psi=\beta|\psi|^2\psi,\quad \psi(t=0)=\psi_0. \tag{3.13}$$

L'idée du schéma de relaxation est d'écrire cette équation comme

$$\begin{cases}\gamma=|\psi|^2,\\ i\partial_t\psi+\Delta\psi=\beta\gamma\psi,\end{cases}$$

et de discrétiser ces deux équations en deux instants différents (comme pour une grille décalée), en restant du second ordre. Pour la première équations, discrétisée au temps t_n, nous avons

$$\frac{\gamma^{n+1/2}+\gamma^{n-1/2}}{2}=|\psi^n|^2$$

et l'équation de Schrödinger est discrétisée autour de $t_{n+1/2}$ avec le schéma de Crank-Nicolson

$$i\frac{\psi^{n+1}-\psi^n}{\delta t}+\frac{1}{2}\Delta\frac{\psi^{n+1}+\psi^n}{2}=\beta\gamma^{n+1/2}\frac{\psi^{n+1}+\psi^n}{2}.$$

Pour assurer une précision de deuxième ordre du schéma, il est nécessaire d'avoir une approximation de $\gamma^{-1/2}(\mathbf{x})=|\psi(-\delta t/2,\mathbf{x})|^2$, au moins d'ordre 2. Cela peut être fait avec le précédent schéma de Crank-Nicolson au pas de temps $-\delta t/2$. Finalement,

Le schéma de relaxation s'écrit

$$\begin{cases} \gamma^{n+1/2} = 2|\psi^n|^2 - \gamma^{n-1/2}, & (3.14) \\ i\dfrac{\psi^{n+1}-\psi^n}{\delta t} + \dfrac{1}{2}\Delta\dfrac{\psi^{n+1}+\psi^n}{2} = \beta\gamma^{n+1/2}\dfrac{\psi^{n+1}+\psi^n}{2}. & (3.15) \end{cases}$$

Le schéma est maintenant linéairement implicite. Aucune méthode de Newton ou de point fixe n'est nécessaire pour résoudre un système non linéaire. Il peut être prouvé que ce schéma converge vers la solution de (3.13) avec une précision du second ordre [3]. Cependant, la preuve est réellement compliquée et ne peut être reproduite ici. Le principal problème est que nous n'avons pas d'équation d'évolution pour γ. Nous pouvons en construire une en prenant la dérivée temporelle de $\gamma = |\psi|^2$. Cependant, à un niveau discret, nous obtiendrions

$$\gamma^{n+1/2} = \sum_k \delta t \frac{\psi^{k+1}|^2 - |\psi^k|^2}{\delta t}$$

qui est une approximation de $\gamma = \int \partial_s \gamma(s) ds$.

Ce schéma présente deux caractéristiques importantes : il est linéaire, mais il préserve également la masse et l'énergie.

Proposition 3.5.2 Le schéma de relaxation (3.14)-(3.15) conserve la masse

$$\int_{\mathbb{R}^d} |\psi^n| \, d\mathbf{x} = \int_{\mathbb{R}^d} |\psi^0| \, d\mathbf{x}, \quad n \geqslant 0, \tag{3.16}$$

et l'énergie

$$\int_{\mathbb{R}^d} \frac{|\nabla\psi^n|^2}{2} + \beta\gamma^{n-1/2}|\psi^n|^2 - \beta\frac{(\gamma^{n-1/2})^2}{2} \, d\mathbf{x} = \int_{\mathbb{R}^d} \frac{|\nabla\psi^0|^2}{2} + \beta\gamma^{-1/2}|\psi^0|^2 - \beta\frac{(\gamma^{-1/2})^2}{2} \, d\mathbf{x}, \tag{3.17}$$

ce qui s'écrit également

$$\int_{\mathbb{R}^d} \frac{|\nabla\psi^n|^2}{2} + \beta\frac{\gamma^{n+1/2}\gamma^{n-1/2}}{2} \, d\mathbf{x} = \int_{\mathbb{R}^d} \frac{|\nabla\psi^0|^2}{2} + \beta\frac{\gamma^{1/2}\gamma^{-1/2}}{2} \, d\mathbf{x}. \tag{3.18}$$

Démonstration. La preuve de la conservation de la masse est similaire à celle du schéma de Crank-Nicolson. Concernant l'énergie, nous multiplions l'équation (3.14) par $\overline{\psi^{n+1}-\psi^n}$, prenons la partie réelle et intégrons plus de $\mathbb{R}^d$. D'abord,

$$\frac{\psi^{n+1}+\psi^n}{2}\overline{\psi^{n+1}-\psi^n} = \frac{1}{2}\left(|\psi^{n+1}|^2 - |\psi^n|^2 + 2i\mathrm{Im}(\psi^n\overline{\psi^{n+1}})\right).$$

Donc,

$$\mathrm{Re}\left[\gamma^{n+1/2}\frac{\psi^{n+1}+\psi^n}{2}\overline{\psi^{n+1}-\psi^n}\right] = \frac{1}{2}\left(\gamma^{n+1/2}|\psi^{n+1}|^2 - \gamma^{n+1/2}|\psi^n|^2\right).$$

En ajoutant $0 = \gamma^{n-1/2}|\psi^n|^2 - \gamma^{n-1/2}|\psi^n|^2$ à cette égalité, nous obtenons

$$\mathrm{Re}\left[\gamma^{n+1/2}\frac{\psi^{n+1}+\psi^n}{2}\overline{\psi^{n+1}-\psi^n}\right] = \frac{1}{2}\left(\gamma^{n+1/2}|\psi^{n+1}|^2 - \gamma^{n+1/2}|\psi^n|^2 + \gamma^{n-1/2}|\psi^n|^2 - \gamma^{n-1/2}|\psi^n|^2\right).$$

En multipliant (3.14) par $\gamma^{n+1/2} - \gamma^{n-1/2}$ conduit à

$$\frac{(\gamma^{n+1/2})^2 - (\gamma^{n-1/2})^2}{2} = |\psi^n|^2(\gamma^{n+1/2} - \gamma^{n-1/2}).$$

Ainsi,

$$\mathrm{Re}\left[\gamma^{n+1/2}\frac{\psi^{n+1}+\psi^n}{2}\overline{\psi^{n+1}-\psi^n}\right] = \frac{1}{2}(\gamma^{n+1/2}|\psi^{n+1}|^2-\gamma^{n-1/2}|\psi^n|^2)-\frac{1}{2}|\psi^n|^2(\gamma^{n+1/2}-\gamma^{n-1/2})$$
$$= \frac{1}{2}(\gamma^{n+1/2}|\psi^{n+1}|^2-\gamma^{n-1/2}|\psi^n|^2)-\frac{(\gamma^{n+1/2})^2-(\gamma^{n-1/2})^2}{4}.$$

De là, en multipliant (3.15) par $\overline{\psi^{n+1}-\psi^n}$, en intégrant et en prenant la partie réelle, nous obtenons

$$-\frac{1}{4}\int_{\mathbb{R}^d}|\nabla\psi^{n+1}|^2-|\nabla\psi^n|^2\}\,d\mathbf{x} = \frac{\beta}{2}\int_{\mathbb{R}^d}\gamma^{n+1/2}|\psi^{n+1}|^2-\gamma^{n-1/2}|\psi^n|^2-\frac{(\gamma^{n+1/2})^2-(\gamma^{n-1/2})^2}{2}\,d\mathbf{x}$$

et finalement

$$\int_{\mathbb{R}^d}\frac{|\nabla\psi^{n+1}|^2}{2}+\beta\gamma^{n+1/2}|\psi^{n+1}|^2-\beta\frac{(\gamma^{n+1/2})^2}{2}\,d\mathbf{x} = \int_{\mathbb{R}^d}\frac{|\nabla\psi^n|^2}{2}+\beta\gamma^{n-1/2}|\psi^n|^2-\beta\frac{(\gamma^{n-1/2})^2}{2}\,d\mathbf{x}.$$

Pour obtenir (3.18), nous multiplions (3.14) par $\gamma^{n-1/2}$ ce qui donne

$$\gamma^{n-1/2}|\psi^n|^2-\frac{(\gamma^{n-1/2})^2}{2} = \frac{\gamma^{n+1/2}\gamma^{n-1/2}}{2}$$

puis nous remplaçons dans (3.17). ■

Mise en œuvre

- Conditions aux limites homogènes de Dirichlet. Nous discrétisons la variable spatiale par la méthode des différences finies et introduisons comme auparavant le vecteur inconnu $\Psi^n=(\psi_j^n)_{j=1}^J$) et la matrice D_2 définie par (3.8) qui se approxime l'opérateur de Laplace. Nous introduisons également le vecteur $\Gamma^{n+1/2}=(\gamma_j^{n+1/2})_{j=1}^J$. Supposons que nous connaissions $\Gamma^{n+1/2}$, alors (3.15) devient

$$i\frac{\Psi^{n+1}-\Psi^n}{\delta t}+\frac{1}{2}D_2\frac{\Psi^{n+1}+\Psi^n}{2} = \beta\Gamma^{n+1/2}\frac{\Psi^{n+1}+\Psi^n}{2}.$$

En définissant $\Psi^{n+1/2}=(\Psi^{n+1}+\Psi^n)/2$, l'algorithme est

Algorithm 2: Schéma de relaxation

Data: Ψ^n, $\Gamma^{n+1/2}$
while $n<N$ **do**
  $\Psi^{n+1/2} = \frac{2i}{\delta t}\left[\frac{2i}{\delta t}I+\frac{D_2}{2}-\beta\Gamma^{n+1/2}\right]^{-1}\Psi^n$;
  $\Psi^{n+1} = 2\Psi^{n+1/2}-\Psi^n$;
  $\Gamma^{n+3/2} = 2|\Psi^{n+1}|^2-\Gamma^{n-1/2}$;
end

- Conditions aux limites périodiques. Comme pour le schéma numérique précédent, nous voulons utiliser la transformée de Fourier discrète. Pour dériver l'algorithme, il suffit de le présenter en version continue avec une transformée de Fourier. L'équation (3.15) se réduit à

$$\left(\frac{2i}{\delta t}I+\frac{\Delta}{2}-\beta\gamma^{n+1/2}\right)\psi^{n+1/2} = \frac{2i}{\delta t}\psi^n$$

ou encore

$$\left(I+\left(I-i\delta t\frac{\Delta}{4}\right)^{-1}i\beta\gamma^{n+1/2}\right)\psi^{n+1/2} = \left(I-i\delta t\frac{\Delta}{4}\right)^{-1}\psi^n.$$

Soit M l'opérateur

$$M = \left(I + \left(I - i\delta t \frac{\Delta}{4} \right)^{-1} i\beta\gamma^{n+1/2} \right).$$

Alors, (3.15) est équivalent à

$$M(\psi^{n+1/2}) = \left(I - i\delta t \frac{\Delta}{4} \right)^{-1} \psi^n. \tag{3.19}$$

Avec la transformée de Fourier, le membre de droite devient

$$\frac{\widehat{\psi^n}}{1 + i\delta t \|\boldsymbol{\xi}\|^2/4}.$$

Le membre de gauche est

$$M(\psi^{n+1/2}) = \psi^{n+1/2} + \mathscr{F}^{-1} \left(\frac{i\beta \mathscr{F}(\gamma^{n+1/2}\psi^{n+1/2})}{1 + i\delta t \|\boldsymbol{\xi}\|^2/4} \right). \tag{3.20}$$

On peut utiliser un solveur de Krylov pour résoudre (3.19) en définissant l'opérateur $M(\varphi)$ par (3.20).

Bibliographie

Notes de cours

[8] Frédéric Lagoutière. *Équations aux dérivées partielles et leurs approximations*. Notes de cours. Université Lyon 1. URL : http://math.univ-lyon1.fr/homes-www/lagoutiere/poly.pdf (cf. pages 28, 31).

Articles

[1] Georgios D. Akrivis. "Finite difference discretization of the cubic Schrödinger equation". In : *IMA J. Numer. Anal.* 13.1 (1993), pages 115-124. ISSN : 0272-4979. DOI : 10.1093/imanum/13.1.115. URL : https://doi.org/10.1093/imanum/13.1.115 (cf. page 56).

[2] Weizhu Bao et Yongyong Cai. "Mathematical theory and numerical methods for BEC". In : *Kinetic and Related Models* 6.1 (2013) (cf. page 51).

[3] Christophe Besse. "A relaxation scheme for the nonlinear Schrödinger equation". In : *SIAM J. Numer. Anal.* 42.3 (2004), pages 934-952. ISSN : 0036-1429. DOI : 10.1137/S0036142901396521. URL : https://doi.org/10.1137/S0036142901396521 (cf. pages 56, 57).

[4] Christophe Besse, Brigitte Bidégaray et Stéphane Descombes. "Order estimates in time of splitting methods for the nonlinear Schrödinger equation". In : *SIAM J. Numer. Anal.* 40.1 (2002), pages 26-40. ISSN : 0036-1429. DOI : 10.1137/S0036142900381497. URL : https://doi.org/10.1137/S0036142900381497 (cf. page 53).

[6] M. Delfour, M. Fortin et G. Payre. "Finite-Difference Solutions of a Non-linear Schrödinger Equation". In : *Journal of Computation Physics* 44 (1981), pages 277-288 (cf. page 54).

[10] Christian Lubich. "On splitting methods for Schrödinger-Poisson and cubic nonlinear Schrödinger equations". In : *Math. Comp.* 77.264 (2008), pages 2141-2153. ISSN : 0025-5718. DOI : 10.1090/S0025-5718-08-02101-7. URL : https://doi.org/10.1090/S0025-5718-08-02101-7 (cf. page 53).

[11] W. Strauss et Vazquez. “Numerical Solution of a Nonlinear Klein-Gordon Equation”. In : *Journal of Computation Physics* 28 (1978), pages 271-278 (cf. page 54).

Livres

[5] Richard Courant et David Hilbert. *Methods of Mathematical Physics*. John Wiley et Sons, 1966 (cf. page 33).

[7] Ernst Hairer, Christian Lubich et Gerhard Wanner. *Geometric Numerical Integration*. 2^{e} édition. Springer, 2006 (cf. page 53).

[9] Christian Lubich. *From Quantum to Classical Molecular Dynamics : Reduced Models and Numerical Analysis*. European Mathematical Society, 2008 (cf. page 18).