Chapitre 2

MÉTHODE DES DIFFÉRENCES FINIES

2.1 Introduction

Dans ce chapitre nous analysons les schémas numériques de différences finies. Nous définissons la **stabilité** et la **consistance** d'un schéma et nous montrons que, pour les équations aux dérivées partielles linéaires à coefficients constants, la stabilité combinée à la consistance d'un schéma impliquent sa **convergence**.

Le plan de ce chapitre est le suivant. La Section 2.2 traite le cas de l'équation de la chaleur introduite au Chapitre 1. La Section 2.3 généralise les résultats précédents aux cas de l'équation des ondes ou de l'équation d'advection. Un des buts de ce chapitre est de fournir un cadre de conception et d'analyse des schémas de différences finies pour des modèles beaucoup plus généraux. Le lecteur ne devrait pas avoir de mal à étendre les concepts présentés ici à son modèle préféré et à concevoir ainsi des schémas numériques originaux.

Terminons cette introduction en disant que la méthode des différences finies est une des plus anciennes méthodes de simulation numérique qui est encore utilisée pour certaines applications, comme la propagation d'ondes (sismiques ou électromagnétiques) ou la mécanique des fluides compressibles. Pour d'autres applications, comme la mécanique du solide ou celle des fluides incompressibles, on lui préfère souvent la méthode des éléments finis. Néanmoins, de nombreux concepts en différences finies se retrouvent dans toutes les autres méthodes numériques. Ainsi, les schémas numériques du Chapitre 8 combineront des éléments finis pour la discrétisation spatiale et des différences finies pour la discrétisation temporelle. La généralité et la simplicité de la méthode des différences finies motive donc son exposition détaillée en début de cet ouvrage.

2.2 Différences finies pour l'équation de la chaleur

2.2.1 Divers exemples de schémas

Nous nous limitons à la dimension un d'espace et nous renvoyons à la Sous-section 2.2.6 pour le cas de plusieurs dimensions d'espace. Nous considérons l'équation de la chaleur dans le domaine borné $(0,1)$

$$\begin{cases} \dfrac{\partial u}{\partial t} - \nu \dfrac{\partial^2 u}{\partial x^2} = 0 \text{ pour } (x,t) \in (0,1) \times \mathbb{R}_*^+ \\ u(0,x) = u_0(x) \text{ pour } x \in (0,1). \end{cases} \tag{2.1}$$

Pour discrétiser le domaine $(0,1)\times\mathbb{R}^+$, on introduit un pas d'espace $\Delta x = 1/(N+1) > 0$ (avec N un entier positif) et un pas de temps $\Delta t > 0$, et on définit les noeuds d'un maillage régulier

$$(t_n, x_j) = (n\Delta t, j\Delta x) \text{ pour } n \geq 0, j \in \{0, 1, ..., N+1\}.$$

On note u_j^n la valeur d'une solution discrète approchée au point (t_n, x_j), et $u(t,x)$ la solution exacte de (2.1). La donnée initiale est discrétisée par

$$u_j^0 = u_0(x_j) \text{ pour } j \in \{0, 1, ..., N+1\}.$$

Les conditions aux limites de (2.1) peuvent être de plusieurs types, mais leur choix n'intervient pas dans la définition des schémas. Ici, nous utilisons des conditions aux limites de Dirichlet

$$u(t,0) = u(t,1) = 0 \text{ pour tout } t \in \mathbb{R}_*^+$$

qui se traduisent en

$$u_0^n = u_{N+1}^n = 0 \text{ pour tout } n > 0.$$

Par conséquent, à chaque pas de temps nous avons à calculer les valeurs $(u_j^n)_{1\leq j\leq N}$ qui forment un vecteur de $\mathbb{R}^N$. Nous donnons maintenant plusieurs schémas possibles pour l'équation de la chaleur (2.1). Tous ces schémas sont définis par N équations (en chaque point x_j, $1 \leq j \leq N$) qui permettent de calculer les N valeurs u_j^n. Au Chapitre 1 nous avons déjà parlé du **schéma explicite**

$$\frac{u_j^{n+1} - u_j^n}{\Delta t} + \nu \frac{-u_{j-1}^n + 2u_j^n - u_{j+1}^n}{(\Delta x)^2} = 0 \tag{2.2}$$

pour $n \geq 0$ et $j \in \{1, ..., N\}$, ainsi que du **schéma implicite**

$$\frac{u_j^{n+1} - u_j^n}{\Delta t} + \nu \frac{-u_{j-1}^{n+1} + 2u_j^{n+1} - u_{j+1}^{n+1}}{(\Delta x)^2} = 0. \tag{2.3}$$

Il est facile de vérifier que le schéma implicite (2.3) est effectivement bien défini, c'est-à-dire qu'on peut calculer les valeurs u_j^{n+1} en fonction des u_j^n : en effet, il faut inverser la matrice tridiagonale carrée de taille N

$$\begin{pmatrix} 1+2c & -c & & & 0 \\ -c & 1+2c & -c & & \\ & \ddots & \ddots & \ddots & \\ & & -c & 1+2c & -c \\ 0 & & & -c & 1+2c \end{pmatrix} \quad \text{avec } c = \frac{\nu \Delta t}{(\Delta x)^2}, \tag{2.4}$$

dont il est aisé de vérifier le caractère défini positif, donc inversible. En faisant une combinaison convexe de (2.2) et (2.3), pour $0 \leq \theta \leq 1$, on obtient le **θ-schéma**

$$\frac{u_j^{n+1} - u_j^n}{\Delta t} + \theta\nu \frac{-u_{j-1}^{n+1} + 2u_j^{n+1} - u_{j+1}^{n+1}}{(\Delta x)^2} + (1-\theta)\nu \frac{-u_{j-1}^n + 2u_j^n - u_{j+1}^n}{(\Delta x)^2} = 0. \tag{2.5}$$

Bien sûr, on retrouve le schéma explicite (2.2) si $\theta = 0$, et le schéma implicite (2.3) si $\theta = 1$. Le θ-schéma (2.5) est implicite dès que $\theta \neq 0$. Pour la valeur $\theta = 1/2$, on obtient le **schéma de Crank-Nicholson**. Un autre schéma implicite, dit à six points, est donné par

$$\begin{aligned} &\frac{u_{j+1}^{n+1} - u_{j+1}^n}{12\Delta t} + \frac{5(u_j^{n+1} - u_j^n)}{6\Delta t} + \frac{u_{j-1}^{n+1} - u_{j-1}^n}{12\Delta t} \\ &+\nu \frac{-u_{j-1}^{n+1} + 2u_j^{n+1} - u_{j+1}^{n+1}}{2(\Delta x)^2} + \nu \frac{-u_{j-1}^n + 2u_j^n - u_{j+1}^n}{2(\Delta x)^2} = 0. \end{aligned} \tag{2.6}$$

Exercice 2.2.1 Montrer que le schéma (2.6) n'est rien d'autre que le θ-schéma avec $\theta = 1/2 - (\Delta x)^2/12\nu\Delta t$.

Tous les schémas qui précèdent sont dits **à deux niveaux** car ils ne font intervenir que deux indices de temps. On peut bien sûr construire des schémas multiniveaux : les plus populaires sont à trois niveaux. En plus du schéma (instable) de Richardson vu au Chapitre 1, on cite le **schéma de DuFort-Frankel**

$$\frac{u_j^{n+1} - u_j^{n-1}}{2\Delta t} + \nu \frac{-u_{j-1}^n + u_j^{n+1} + u_j^{n-1} - u_{j+1}^n}{(\Delta x)^2} = 0, \tag{2.7}$$

le **schéma de Gear**

$$\frac{3u_j^{n+1} - 4u_j^n + u_j^{n-1}}{2\Delta t} + \nu \frac{-u_{j-1}^{n+1} + 2u_j^{n+1} - u_{j+1}^{n+1}}{(\Delta x)^2} = 0. \tag{2.8}$$

Nous voilà en face de beaucoup trop de schémas ! Et la liste ci-dessus n'est pas exhaustive ! Un des buts de l'analyse numérique va être de comparer et de sélectionner les meilleurs schémas suivant des critères de précision, de coût, ou de robustesse.

Remarque 2.2.1 S'il y a un second membre $f(t,x)$ dans l'équation de la chaleur (2.1), alors les schémas se modifient en remplaçant zéro au second membre par une approximation consistante de $f(t,x)$ au point (t_n, x_j). Par exemple, si on choisit l'approximation $f(t_n, x_j)$, le schéma explicite (2.2) devient

$$\frac{u_j^{n+1} - u_j^n}{\Delta t} + \nu \frac{-u_{j-1}^n + 2u_j^n - u_{j+1}^n}{(\Delta x)^2} = f(t_n, x_j).$$

Remarque 2.2.2 Les schémas ci-dessus ont une écriture plus ou moins compacte, c'est-à-dire qu'il font intervenir un nombre fini, plus ou moins restreint, de valeurs u_j^n. La collection des couples (n', j') qui interviennent dans l'équation discrète au point (n, j) est appelé **stencil** du schéma (terme anglais qu'on peut essayer de traduire par **support**). En général, plus le stencil est large, plus le schéma est coûteux et difficile à programmer (en partie à cause des "effets de bord", c'est-à-dire des cas où certains des couples (n', j') sortent du domaine de calcul).

Remarque 2.2.3 On peut remplacer les conditions aux limites de Dirichlet dans (2.1) par des conditions aux limites de Neumann, ou bien des conditions aux limites de périodicité (entres autres). Commençons par décrire deux manières différentes de discrétiser les conditions de Neumann

$$\frac{\partial u}{\partial x}(t, 0) = 0 \text{ et } \frac{\partial u}{\partial x}(t, 1) = 0.$$

Tout d'abord, on peut écrire

$$\frac{u_1^n - u_0^n}{\Delta x} = 0 \text{ et } \frac{u_{N+1}^n - u_N^n}{\Delta x} = 0$$

qui permet d'éliminer les valeurs u_0^n et u_{N+1}^n et de ne calculer que les N valeurs $(u_j^n)_{1 \le j \le N}$. Cette discrétisation de la condition de Neumann n'est que du premier ordre. Si le schéma est du deuxième ordre, cela engendre une perte de précision près du bord. C'est pourquoi on propose une autre discrétisation (du deuxième ordre)

$$\frac{u_1^n - u_{-1}^n}{2\Delta x} = 0 \text{ et } \frac{u_{N+2}^n - u_N^n}{2\Delta x} = 0$$

qui est plus précise, mais nécessite l'ajout de 2 "points fictifs" x_{-1} et x_{N+2}. On élimine les valeurs u_{-1}^n et u_{N+2}^n, correspondant à ces points fictifs, et il reste maintenant $N+2$ valeurs à calculer, à savoir $(u_j^n)_{0 \le j \le N+1}$.

D'autre part, les conditions aux limites de périodicité s'écrivent

$$u(t, x+1) = u(t, x) \text{ pour tout } x \in [0, 1], t \ge 0.$$

Elles se discrétisent par les égalités $u_0^n = u_{N+1}^n$ pour tout $n \ge 0$, et plus généralement $u_j^n = u_{N+1+j}^n$.

2.2.2 Consistance et précision

Bien sûr, les formules des schémas ci-dessus ne sont pas choisies au hasard : elles résultent d'une approximation de l'équation par développement de Taylor comme nous l'avons expliqué au Chapitre 1. Pour formaliser cette approximation de l'équation aux dérivées partielles par des différences finies, on introduit la notion de **consistance** et de **précision**. Bien que pour l'instant nous ne considérons que l'équation de la chaleur (2.1), nous allons donner une définition de la consistance valable pour n'importe quelle équation aux dérivées partielles que nous notons $F(u) = 0$. Remarquons que $F(u)$ est une notation pour une fonction de u et de ses dérivées partielles en tout point (t, x). De manière générale un schéma aux différence finies est défini, pour tous les indices possibles n, j, par la formule

$$F_{\Delta t,\Delta x}\left(\{u_{j+k}^{n+m}\}_{m^-\leq m\leq m^+,\ k^-\leq k\leq k^+}\right) = 0 \tag{2.9}$$

où les entiers m^-, m^+, k^-, k^+ définissent la largeur du stencil du schéma (voir la Remarque 2.2.2).

Définition 2.2.4 *Le schéma aux différence finies (2.9) est dit consistant avec l'équation aux dérivées partielles $F(u) = 0$, si, pour toute solution $u(t, x)$ suffisamment régulière de cette équation,* **l'erreur de troncature** *du schéma, définie par*

$$F_{\Delta t,\Delta x}\left(\{u(t + m\Delta t, x + k\Delta x)\}_{m^-\leq m\leq m^+,\ k^-\leq k\leq k^+}\right), \tag{2.10}$$

tend vers zéro, uniformément par rapport à (t, x), lorsque Δt et Δx tendent vers zéro indépendemment.

De plus, on dit que le schéma est précis à l'ordre p en espace et à l'ordre q en temps si l'erreur de troncature (2.10) tend vers zéro comme $\mathcal{O}\Big((\Delta x)^p + (\Delta t)^q\Big)$ lorsque Δt et Δx tendent vers zéro.

Remarque 2.2.5 Il faut prendre garde dans la formule (2.9) à une petite ambiguïté quant à la définition du schéma. En effet, on peut toujours multiplier n'importe quelle formule par une puissance suffisamment élevée de Δt et Δx de manière à ce que l'erreur de troncature tende vers zéro. Cela rendrait consistant n'importe quel schéma ! Pour éviter cet inconvénient, on supposera toujours que la formule $F_{\Delta t,\Delta x}(\{u_{j+k}^{n+m}\}) = 0$ a été écrite de telle manière que, pour une fonction régulière $u(t, x)$ qui n'est pas solution de l'équation de la chaleur, la limite de l'erreur de troncature n'est pas nulle. •

Concrètement on calcule l'erreur de troncature d'un schéma en remplaçant u_{j+k}^{n+m} dans la formule (2.9) par $u(t + m\Delta t, x + k\Delta x)$. Comme application de la Définition 2.2.4, nous allons montrer le lemme suivant.

Lemme 2.2.6 *Le schéma explicite (2.2) est consistant, précis à l'ordre 1 en temps et 2 en espace. De plus, si on choisit de garder constant le rapport $\nu\Delta t/(\Delta x)^2 = 1/6$, alors ce schéma est précis à l'ordre 2 en temps et 4 en espace.*

Remarque 2.2.7 Dans la deuxième phrase de l'énoncé du Lemme 2.2.6 on a légèrement modifié la définition de la consistance en spécifiant le rapport entre Δt et Δx lorsqu'ils tendent vers zéro. Ceci permet de tenir compte d'éventuelles compensations entre termes apparaissant dans l'erreur de troncature. En pratique, on observe effectivement de telles améliorations de la précision si on adopte le bon rapport entre les pas Δt et Δx. •

Démonstration. Soit $v(t,x)$ une fonction de classe $\mathcal{C}^6$. Par développement de Taylor autour du point (t,x), on calcule l'erreur de troncature du schéma (2.2)

$$\begin{aligned}\frac{v(t+\Delta t,x)-v(t,x)}{\Delta t} &+\nu\frac{-v(t,x-\Delta x)+2v(t,x)-v(t,x+\Delta x)}{(\Delta x)^2}\\ &= \Big(v_t-\nu v_{xx}\Big)+\frac{\Delta t}{2}v_{tt}-\frac{\nu(\Delta x)^2}{12}v_{xxxx}+\mathcal{O}\Big((\Delta t)^2+(\Delta x)^4\Big),\end{aligned}$$

où v_t, v_x désignent les dérivées partielles de v. Si v est une solution de l'équation de la chaleur (2.1), on obtient ainsi aisément la consistance ainsi que la précision à l'ordre 1 en temps et 2 en espace. Si on suppose en plus que $\nu\Delta t/(\Delta x)^2 = 1/6$, alors les termes en Δt et en $(\Delta x)^2$ se simplifient car $v_{tt} = \nu v_{txx} = \nu^2 v_{xxxx}$. □

Schéma	Erreur de troncature	Stabilité
Explicite (2.2)	$\mathcal{O}\Big(\Delta t+(\Delta x)^2\Big)$	stable L^2 et L^∞ si condition CFL $2\nu\Delta t\le(\Delta x)^2$
Implicite (2.3)	$\mathcal{O}\Big(\Delta t+(\Delta x)^2\Big)$	stable L^2 et L^∞
Crank-Nicholson (2.5) (avec $\theta=1/2$)	$\mathcal{O}\Big((\Delta t)^2+(\Delta x)^2\Big)$	stable L^2
θ-schéma (2.5) (avec $\theta\neq 1/2$)	$\mathcal{O}\Big(\Delta t+(\Delta x)^2\Big)$	stable L^2 si condition CFL $2(1-2\theta)\nu\Delta t\le(\Delta x)^2$
Schéma à 6 points (2.6)	$\mathcal{O}\Big((\Delta t)^2+(\Delta x)^4\Big)$	stable L^2
DuFort-Frankel (2.7)	$\mathcal{O}\Big((\frac{\Delta t}{\Delta x})^2+(\Delta x)^2\Big)$	stable L^2
Gear (2.8)	$\mathcal{O}\Big((\Delta t)^2+(\Delta x)^2\Big)$	stable L^2

TAB. 2.1 – Erreurs de troncature et stabilité de divers schémas pour l'équation de la chaleur

Exercice 2.2.2 Pour chacun des schémas de la Sous-section 2.2.1, vérifier que l'erreur de troncature est bien du type annoncé dans le Tableau 2.1. (On remarquera que tous ces schémas sont consistants sauf celui de DuFort-Frankel.)

2.2.3 Stabilité et analyse de Fourier

Dans le Chapitre 1 nous avons évoqué la stabilité des schémas de différences finies sans en donner une définition précise. Tout au plus avons nous expliqué que, numériquement, l'instabilité se manifeste par des oscillations non bornées de la solution numérique. Il est donc temps de donner une définition mathématique de la stabilité. Pour cela nous avons besoin de définir une norme pour la solution numérique $u^n = (u_j^n)_{1\le j\le N}$. Nous reprenons les normes classiques sur $\mathbb{R}^N$ que nos pondérons simplement par le pas d'espace Δx :

$$\|u^n\|_p = \left(\sum_{j=1}^{N} \Delta x |u_j^n|^p\right)^{1/p} \quad \text{pour } 1 \le p \le +\infty, \tag{2.11}$$

où le cas limite $p = +\infty$ doit être compris dans le sens $\|u^n\|_\infty = \max_{1\le j\le N} |u_j^n|$. Remarquons que la norme ainsi définie dépend de Δx à travers la pondération mais aussi à travers l'entier N car $\Delta x = 1/(N+1)$. Grâce à la pondération par Δx, la norme $\|u^n\|_p$ est identique à la norme $L^p(0,1)$ pour les fonctions constantes par morceaux sur les sous-intervalles $[x_j, x_{j+1}[$ de $[0,1]$. Souvent, on l'appellera donc "norme L^p". En pratique on utilise surtout les normes correspondant aux valeurs $p = 2, +\infty$.

Définition 2.2.8 *Un schéma aux différence finies est dit* **stable** *pour la norme $\|\ \|$, définie par (2.11), s'il existe une constante $K > 0$ indépendante de Δt et Δx (lorsque ces valeurs tendent vers zéro) telle que*

$$\|u^n\| \le K\|u^0\| \text{ pour tout } n \ge 0, \tag{2.12}$$

quelle que soit la donnée initiale u^0.

Si (2.12) n'a lieu que pour des pas Δt et Δx astreints à certaines inégalités, on dit que le schéma est **conditionnellement stable***.*

Remarque 2.2.9 Puisque toutes les normes sont équivalentes dans $\mathbb{R}^N$, le lecteur trop rapide pourrait croire que la stabilité par rapport à une norme implique la stabilité par rapport à toutes les normes. Malheureusement il n'en est rien et il existe des schémas qui sont stables par rapport à une norme mais pas par rapport à une autre (voir plus loin l'exemple du schéma de Lax-Wendroff avec les Exercices 2.3.2 et 2.3.3). En effet, le point crucial dans la Définition 2.2.8 est que la majoration est uniforme par rapport à Δx alors même que les normes définies par (2.11) dépendent de Δx. •

Définition 2.2.10 *Un schéma aux différences finies est dit* **linéaire** *si la formule $F_{\Delta t,\Delta x}(\{u_{j+k}^{n+m}\}) = 0$ qui le définit est linéaire par rapport à ses arguments u_{j+k}^{n+m}.*

La stabilité d'un schéma linéaire à deux niveaux est très facile à interpréter. En effet, par linéarité tout schéma linéaire à deux niveaux peut s'écrire sous la forme condensée

$$u^{n+1} = Au^n, \tag{2.13}$$

où A est un opérateur linéaire (une matrice, dite d'itération) de $\mathbb{R}^N$ dans $\mathbb{R}^N$. Par exemple, pour le schéma explicite (2.2) la matrice A vaut

$$\begin{pmatrix} 1-2c & c & & & 0 \\ c & 1-2c & c & & \\ & \ddots & \ddots & \ddots & \\ & & c & 1-2c & c \\ 0 & & & c & 1-2c \end{pmatrix} \text{ avec } c = \frac{\nu\Delta t}{(\Delta x)^2}, \qquad (2.14)$$

tandis que pour le schéma implicite (2.3) la matrice A est l'inverse de la matrice (2.4). A l'aide de cette matrice d'itération, on a $u^n = A^n u^0$ (attention, la notation A^n désigne ici la puissance n-ème de A), et par conséquent la stabilité du schéma est équivalente à

$$\|A^n u^0\| \leq K\|u^0\| \quad \forall\, n \geq 0, \forall\, u^0 \in \mathbb{R}^N.$$

Introduisant la norme matricielle subordonnée (voir la Définition 13.1.1)

$$\|M\| = \sup_{u\in\mathbb{R}^N, u\neq 0} \frac{\|Mu\|}{\|u\|},$$

la stabilité du schéma est équivalente à

$$\|A^n\| \leq K \quad \forall\, n \geq 0, \qquad (2.15)$$

qui veut dire que la suite des puissances de A est bornée.

Stabilité en norme L^∞.
La stabilité en norme L^∞ est très liée avec le principe du maximum discret que nous avons déjà vu au Chapitre 1. Rappelons la définition de ce principe.

Définition 2.2.11 *Un schéma aux différences finies vérifie le* **principe du maximum discret** *si pour tout $n \geq 0$ et tout $1 \leq j \leq N$ on a*

$$\min\left(0, \min_{0\leq j\leq N+1} u_j^0\right) \leq u_j^n \leq \max\left(0, \max_{0\leq j\leq N+1} u_j^0\right)$$

quelle que soit la donnée initiale u^0.

Remarque 2.2.12 Dans la Définition 2.2.11 les inégalités tiennent compte non seulement du minimum et du maximum de u^0 mais aussi de zéro qui est la valeur imposée au bord par les conditions aux limites de Dirichlet. Cela est nécessaire si la donnée initiale u^0 ne vérifie pas les conditions aux limites de Dirichlet (ce qui n'est pas exigé), et inutile dans le cas contraire. •

Comme nous l'avons vu au Chapitre 1 (voir (1.33) et l'Exercice 1.4.1), la vérification du principe du maximum discret permet de démontrer le lemme suivant.

Lemme 2.2.13 *Le schéma explicite (2.2) est stable en norme L^∞ si et seulement si la condition CFL $2\nu\Delta t \leq (\Delta x)^2$ est satisfaite. Le schéma implicite (2.3) est stable en norme L^∞ quelque soit les pas de temps Δt et d'espace Δx (on dit qu'il est inconditionnellement stable).*

Exercice 2.2.3 Montrer que le schéma de Crank-Nicholson (2.5) (avec $\theta = 1/2$) est stable en norme L^∞ si $\nu\Delta t \leq (\Delta x)^2$, et que le schéma de DuFort-Frankel (2.7) est stable en norme L^∞ si $2\nu\Delta t \leq (\Delta x)^2$

Stabilité en norme L^2.

De nombreux schémas ne vérifient pas le principe du maximum discret mais sont néanmoins de "bons" schémas. Pour ceux-là, il faut vérifier la stabilité dans une autre norme que la norme L^∞. La norme L^2 se prête très bien à l'étude de la stabilité grâce à l'outil très puissant de l'analyse de Fourier que nous présentons maintenant. Pour ce faire, nous supposons désormais que les conditions aux limites pour l'équation de la chaleur sont des **conditions aux limites de périodicité**, qui s'écrivent $u(t, x+1) = u(t, x)$ pour tout $x \in [0, 1]$ et tout $t \geq 0$. Pour les schémas numériques, elles conduisent aux égalités $u_0^n = u_{N+1}^n$ pour tout $n \geq 0$, et plus généralement $u_j^n = u_{N+1+j}^n$. Il reste donc à calculer $N + 1$ valeurs u_j^n.

A chaque vecteur $u^n = (u_j^n)_{0 \leq j \leq N}$ on associe une fonction $u^n(x)$, constante par morceaux, périodique de période 1, définie sur $[0, 1]$ par

$$u^n(x) = u_j^n \text{ si } x_{j-1/2} < x < x_{j+1/2}$$

avec $x_{j+1/2} = (j + 1/2)\Delta x$ pour $0 \leq j \leq N$, $x_{-1/2} = 0$, et $x_{N+1+1/2} = 1$. Ainsi définie, la fonction $u^n(x)$ appartient à $L^2(0, 1)$. Or, d'après l'analyse de Fourier, toute fonction de $L^2(0, 1)$ peut se décomposer en une somme de Fourier (voir le théorème 4.5.13 de [6]). Plus précisément on a

$$u^n(x) = \sum_{k \in \mathbb{Z}} \hat{u}^n(k) \exp(2i\pi kx), \tag{2.16}$$

avec $\hat{u}^n(k) = \int_0^1 u^n(x) \exp(-2i\pi kx)\, dx$ et la formule de Plancherel

$$\int_0^1 |u^n(x)|^2\, dx = \sum_{k \in \mathbb{Z}} |\hat{u}^n(k)|^2. \tag{2.17}$$

Remarquons que même si u^n est une fonction réelle, les coefficients $\hat{u}^n(k)$ de la série de Fourier sont complexes. Une propriété importante pour la suite de la transformée de Fourier des fonctions périodiques est la suivante : si on note $v^n(x) = u^n(x + \Delta x)$, alors $\hat{v}^n(k) = \hat{u}^n(k) \exp(2i\pi k\Delta x)$.

Expliquons maintenant la méthode sur l'exemple du schéma explicite (2.2). Avec nos notations, on peut réécrire ce schéma, pour $0 \leq x \leq 1$,

$$\frac{u^{n+1}(x) - u^n(x)}{\Delta t} + \nu \frac{-u^n(x - \Delta x) + 2u^n(x) - u^n(x + \Delta x)}{(\Delta x)^2} = 0.$$

Par application de la transformée de Fourier, il vient

$$\hat{u}^{n+1}(k) = \left(1 - \frac{\nu\Delta t}{(\Delta x)^2}\left(-\exp(-2i\pi k\Delta x) + 2 - \exp(2i\pi k\Delta x)\right)\right)\hat{u}^n(k).$$

Autrement dit

$$\hat{u}^{n+1}(k) = A(k)\hat{u}^n(k) = A(k)^{n+1}\hat{u}^0(k) \text{ avec } A(k) = 1 - \frac{4\nu\Delta t}{(\Delta x)^2}(\sin(\pi k\Delta x))^2.$$

Pour $k \in \mathbb{Z}$, le coefficient de Fourier $\hat{u}^n(k)$ est borné lorsque n tend vers l'infini si et seulement si le facteur d'amplification vérifie $|A(k)| \leq 1$, c'est-à-dire

$$2\nu\Delta t(\sin(\pi k\Delta x))^2 \leq (\Delta x)^2. \tag{2.18}$$

Si la condition CFL (1.31), i.e. $2\nu\Delta t \leq (\Delta x)^2$, est satisfaite, alors l'inégalité (2.18) est vraie quelque soit le mode de Fourier $k \in \mathbb{Z}$, et par la formule de Plancherel on en déduit

$$\|u^n\|_2^2 = \int_0^1 |u^n(x)|^2 dx = \sum_{k\in\mathbb{Z}} |\hat{u}^n(k)|^2 \leq \sum_{k\in\mathbb{Z}} |\hat{u}^0(k)|^2 = \int_0^1 |u^0(x)|^2 dx = \|u^0\|_2^2,$$

ce qui n'est rien d'autre que la stabilité L^2 du schéma explicite. Si la condition CFL n'est pas satisfaite, le schéma est instable. En effet, il suffit de choisir Δx (éventuellement suffisamment petit) et k_0 (suffisamment grand) et une donnée initiale ayant une seule composante de Fourier non nulle $\hat{u}^0(k_0) \neq 0$ avec $\pi k_0 \Delta x \approx \pi/2$ (modulo π) de telle manière que $|A(k_0)| > 1$. On a donc démontré le lemme suivant.

Lemme 2.2.14 *Le schéma explicite (2.2) est stable en norme L^2 si et seulement si la condition CFL $2\nu\Delta t \leq (\Delta x)^2$ est satisfaite.*

De la même façon on va démontrer la stabilité du schéma implicite.

Lemme 2.2.15 *Le schéma implicite (2.3) est stable en norme L^2.*

Remarque 2.2.16 Pour les schémas explicite (2.2) et implicite (2.3) la condition de stabilité L^2 est la même que celle de stabilité L^∞. Cela n'est pas toujours le cas pour d'autres schémas. •

Démonstration. Un raisonnement analogue à celui utilisé pour le schéma explicite conduit, pour $0 \leq x \leq 1$, à

$$\frac{u^{n+1}(x) - u^n(x)}{\Delta t} + \nu\frac{-u^{n+1}(x-\Delta x) + 2u^{n+1}(x) - u^{n+1}(x+\Delta x)}{(\Delta x)^2} = 0,$$

et par application de la transformée de Fourier

$$\hat{u}^{n+1}(k)\left(1 + \frac{\nu\Delta t}{(\Delta x)^2}\left(-\exp(-2i\pi k\Delta x) + 2 - \exp(2i\pi k\Delta x)\right)\right) = \hat{u}^n(k).$$

Autrement dit

$$\hat{u}^{n+1}(k) = A(k)\hat{u}^n(k) = A(k)^{n+1}\hat{u}^0(k) \text{ avec } A(k) = \left(1 + \frac{4\nu\Delta t}{(\Delta x)^2}(\sin(\pi k \Delta x))^2\right)^{-1}.$$

Comme $|A(k)| \leq 1$ pour tout mode de Fourier k, la formule de Plancherel permet de conclure à la stabilité L^2 du schéma. □

Remarque 2.2.17 L'analyse de Fourier repose sur le choix des conditions aux limites de périodicité. On peut aussi l'effectuer si l'équation aux dérivées partielles a lieu sur tout $\mathbb{R}$ au lieu de $[0,1]$ (on a alors affaire à une intégrale de Fourier plutôt qu'à une série de Fourier). Néanmoins, il n'est pas très réaliste de parler de schéma numérique sur tout $\mathbb{R}$ puisque cela implique un nombre infini de valeurs u_j^n à chaque pas de temps n alors qu'un ordinateur ne peut que traiter un nombre fini de valeurs.

La stabilité L^2 peut aussi se démontrer dans le cas de conditions aux limites de Dirichlet. Il faut alors adapter les idées de l'analyse de Fourier. Par exemple, ce qui remplace la transformée de Fourier dans ce cas est la décomposition sur une base de vecteurs propres de la matrice d'itération (2.13) qui permet de passer du vecteur u^n au vecteur u^{n+1}. •

Remarque 2.2.18 (Essentielle d'un point de vue pratique) Traduisons sous forme de "recette" la méthode de l'analyse de Fourier pour prouver la stabilité L^2 d'un schéma. On injecte dans le schéma un mode de Fourier

$$u_j^n = A(k)^n \exp(2i\pi k x_j) \quad \text{avec} \quad x_j = j\Delta x,$$

et on en déduit la valeur du facteur d'amplification $A(k)$. Rappelons que, pour l'instant, nous nous sommes limité au cas scalaire, c'est-à-dire que $A(k)$ est un nombre complexe dans $\mathbb{C}$. On appelle **condition de stabilité de Von Neumann** l'inégalité

$$|A(k)| \leq 1 \text{ pour tout mode } k \in \mathbb{Z}. \tag{2.19}$$

Si la condition de stabilité de Von Neumann est satisfaite (avec éventuellement des restrictions sur Δt et Δx), alors le schéma est stable pour la norme L^2, si non il est instable.

En général, un schéma stable (et consistant) est convergent (voir la Sous-section 2.2.4). En pratique, un schéma instable est totalement "inutilisable". En effet, même si on part d'une donnée initiale spécialement préparée de manière à ce qu'aucun des modes de Fourier instables ne soit excité par elle, les inévitables erreurs d'arrondi vont créer des composantes non nulles (bien que très petites) de la solution sur ces modes instables. La croissance exponentielle des modes instables entraîne qu'après seulement quelques pas en temps ces "petits" modes deviennent "énormes" et polluent complètement le reste de la solution numérique. •

Exercice 2.2.4 Montrer que le θ-schéma (2.5) est stable en norme L^2 inconditionnellement si $1/2 \leq \theta \leq 1$, et sous la condition CFL $2(1-2\theta)\nu\Delta t \leq (\Delta x)^2$ si $0 \leq \theta < 1/2$.

Exercice 2.2.5 Montrer que le schéma à 6 points (2.6) est inconditionnellement stable en norme L^2.

Remarque 2.2.19 Certains auteurs utilisent une autre définition de la stabilité, moins restrictive que la Définition 2.2.8 mais plus complexe. Dans cette définition le schéma est dit stable pour la norme $\| \ \|$ si pour tout temps $T > 0$ il existe une constante $K(T) > 0$ indépendante de Δt et Δx telle que

$$\|u^n\| \leq K(T)\|u^0\| \text{ pour tout } 0 \leq n \leq T/\Delta t,$$

quelle que soit la donnée initiale u^0. Cette nouvelle définition permet à la solution de croître avec le temps comme c'est le cas, par exemple, pour la solution de l'équation

$$\frac{\partial u}{\partial t} - \nu \frac{\partial^2 u}{\partial x^2} = cu \text{ pour } (t,x) \in \mathbb{R}^+ \times \mathbb{R},$$

qui, par le changement d'inconnue $v(t,x) = e^{-ct}u(t,x)$, se ramène à l'équation de la chaleur (donc pour $c > 0$ suffisamment grand, la solution u croît exponentiellement en temps). Avec une telle définition de la stabilité, la condition de stabilité de Von Neumann devient l'inégalité

$$|A(k)| \leq 1 + C\Delta t \text{ pour tout mode } k \in \mathbb{Z}.$$

Par souci de simplicité nous préférons nous en tenir à la Définition 2.2.8 de la stabilité. •

2.2.4 Convergence des schémas

Nous avons maintenant tous les outils pour démontrer la convergence des schémas de différences finies. Le résultat principal de cette sous-section est le Théorème de Lax qui affirme que, pour un schéma linéaire, **consistance et stabilité implique convergence**. La portée de ce résultat dépasse en fait de beaucoup la méthode des différences finies. Pour toute méthode numérique (différences finies, éléments finis, etc.) la convergence se démontre en conjuguant deux arguments : stabilité et consistance (leurs définitions précises varient d'une méthode à l'autre). D'un point de vue pratique, le Théorème de Lax est très rassurant : si l'on utilise un schéma consistant (ils sont construits pour cela en général) et que l'on n'observe pas d'oscillations numériques (c'est-à-dire qu'il est stable), alors la solution numérique est proche de la solution exacte (le schéma converge).

Théorème 2.2.20 (Lax) *Soit $u(t,x)$ la solution suffisamment régulière de l'équation de la chaleur (2.1) (avec des conditions aux limites appropriées). Soit u_j^n la solution numérique discrète obtenue par un schéma de différences finies avec la donnée initiale $u_j^0 = u_0(x_j)$. On suppose que le schéma est linéaire, à deux niveaux, consistant, et stable pour une norme $\| \ \|$. Alors le schéma est convergent au sens où*

$$\forall T > 0, \quad \lim_{\Delta t, \Delta x \to 0} \left(\sup_{t_n \leq T} \|e^n\| \right) = 0, \tag{2.20}$$

avec e^n le vecteur "erreur" défini par ses composantes $e_j^n = u_j^n - u(t_n, x_j)$.

De plus, si le schéma est précis à l'ordre p en espace et à l'ordre q en temps, alors pour tout temps $T > 0$ il existe une constante $C_T > 0$ telle que

$$\sup_{t_n \leq T} \|e^n\| \leq C_T \Big((\Delta x)^p + (\Delta t)^q \Big). \tag{2.21}$$

Remarque 2.2.21 Nous n'avons pas encore démontré l'existence et l'unicité de la solution de l'équation de la chaleur (2.1) (avec des conditions aux limites de Dirichlet ou périodiques). Pour l'instant nous faisons donc l'hypothèse de l'existence et l'unicité d'une telle solution (ainsi que de sa régularité), mais nous verrons au Chapitre 8 que ce résultat est vrai de manière générale. •

Démonstration. Pour simplifier, on suppose que les conditions aux limites sont de Dirichlet. La même démonstration est aussi valable pour des conditions aux limites de périodicité ou des conditions aux limites de Neumann (en supposant ces dernières discrétisées avec le même ordre de précision que le schéma). Un schéma linéaire à deux niveaux peut s'écrire sous la forme condensée (2.13), i.e.

$$u^{n+1} = A u^n,$$

où A est la matrice d'itération (carrée de taille N). Soit u la solution (supposée suffisamment régulière) de l'équation de la chaleur (2.1). On note $\tilde{u}^n = (\tilde{u}^n_j)_{1 \leq j \leq N}$ avec $\tilde{u}^n_j = u(t_n, x_j)$. Comme le schéma est consistant, il existe un vecteur ϵ^n tel que

$$\tilde{u}^{n+1} = A\tilde{u}^n + \Delta t \epsilon^n \text{ avec } \lim_{\Delta t, \Delta x \to 0} \|\epsilon^n\| = 0, \tag{2.22}$$

et la convergence de ϵ^n est uniforme pour tout les temps $0 \leq t_n \leq T$. Si le schéma est précis à l'ordre p en espace et à l'ordre q en temps, alors $\|\epsilon^n\| \leq C((\Delta x)^p + (\Delta t)^q)$. En posant $e^n_j = u^n_j - u(t_n, x_j)$ on obtient par soustraction de (2.22) à (2.13)

$$e^{n+1} = Ae^n - \Delta t \epsilon^n$$

d'où par récurrence

$$e^n = A^n e^0 - \Delta t \sum_{k=1}^{n} A^{n-k} \epsilon^{k-1}. \tag{2.23}$$

Or, la stabilité du schéma veut dire que $\|u^n\| = \|A^n u^0\| \leq K \|u^0\|$ pour toute donnée initiale, c'est-à-dire que $\|A^n\| \leq K$ où la constante K ne dépend pas de n. D'autre part, $e^0 = 0$, donc (2.23) donne

$$\|e^n\| \leq \Delta t \sum_{k=1}^{n} \|A^{n-k}\| \|\epsilon^{k-1}\| \leq \Delta t n K C \Big((\Delta x)^p + (\Delta t)^q \Big),$$

ce qui donne l'inégalité (2.21) avec la constante $C_T = TKC$. La démonstration de (2.20) est similaire. □