Université de Rennes 2023-2024
Agrégation externe

Compléments sur l'approximation numérique de l'équation de chaleur

Miguel Rodrigues

Ces notes de cours sont consacrées à l'approximation numérique des solutions de l'équation de la chaleur. Une partie du matériel discuté ici peut être trouvé dans les livres classiques d'analyse numérique, notamment dans [Fil13, Chapitre 7].

1 Premiers commentaires

On cherche à calculer des approximations de solutions de

$$\partial_t \theta - \partial_x^2 \theta = Q\,, \tag{1.1}$$

où l'inconnue est θ. L'équation doit être complétée par une donnée initiale et, selon le domaine spatial, par des conditions de bord adaptées.
Dans les problèmes les plus simples le terme source est donné, mais le cas où la source est une fonction de l'inconnue peut encore être analysé à un niveau élémentaire par variation de la constante et point fixe.
En l'absence de conditions de bord et de terme source, l'équation préserve l'intégrale et la positivité. De même, elle fait décroître toute fonctionnelle du type $\int \eta(\theta)$ avec η convexe, en particulier elle fait décroître strictement [1] toute norme L^p, $1 \leq p < \infty$. Elle fait également décroître strictement [2] la norme L^∞.
On pourrait utiliser un schéma numérique lié au mouvement brownien, qui serait alors utlisé comme les caractéristiques dans les méthodes particulaires pour les équations de transport. Cependant, comme pour la méthode de Monte Carlo pour le calcul approché d'intégrale cela n'aurait d'intérêt qu'en dimension grande.
Nous allons nous focaliser sur deux types de méthodes,

- la méthode spectrale ;
- la méthode des différences finies.

2 Méthode spectrale

La méthode spectrale (dans la version que nous utilisons) consiste à discrétiser directement la formule de résolution en Fourier. A priori, cela devrait le restreindre au cas avec des conditions de bord périodiques, mais en dimension 1 les cas Dirichlet, Neumann ou Dirichlet d'un côté Neumann de l'autre s'y ramènent.

1. Sauf pour la solution nulle.
2. Sauf pour les solutions constantes.

Explicitement, si θ résout (1.1) avec, pour tout $t \geq 0$, $u(t,\cdot)$ $\mathbf{Z}$-périodique, alors pour tout $t' > t \geq 0$, les coefficients de Fourier [3] de u vérifient pour $k \in \mathbf{Z}$

$$c_k(u(t',\cdot)) \;=\; \mathrm{e}^{-(t'-t)\,(2\pi\,k)^2}\, c_k(u(t,\cdot)) + \int_t^{t'} \mathrm{e}^{-(t'-\tau)\,(2\pi\,k)^2}\, c_k(Q(\tau,\cdot))\,\mathrm{d}\tau\,.$$

Étant donnés des pas d'espace et de temps Δx et Δt tels que $N_x := 1/\Delta x \in \mathbf{N}^*$, on note N_t la partie entière de $T/\Delta t$ et on introduit les nœuds de subdivisions $x_j = j\Delta x$, $j \in [\![1, N_x]\!]$, et $t_n = n\Delta t$, $n \in [\![0, N_t]\!]$. On souhaite calculer u_j^n, des valeurs en (t_n, x_j) censées fournir une approximation des valeurs $\theta(t_n, x_j)$ de la solution. Insistons sur le fait que bien nous ne le marquions pas par des indices et des exposants, (u_j^n, t_n, x_j) ne dépendent pas que de (j,n) mais aussi de $(\Delta t, \Delta x)$ ou de manière équivalente de $(\Delta t, N_x)$. Introduisons

$$\mathcal{F}_{N_x} : \mathbf{C}^{N_x} \to \mathbf{C}^{N_x}, \quad (f_1, \cdots, f_{N_x}) \longmapsto \left(\frac{1}{N_x} \sum_{j=1}^{N_x} f_j\, \mathrm{e}^{-\mathrm{i}k\,2\pi\,\frac{j}{N_x}} \right)_{k \in [\![-\lfloor \frac{N_x}{2} \rfloor, N_x - \lfloor \frac{N_x}{2} \rfloor - 1]\!]}$$

et posons $\Theta^n := (\theta_1^n, \cdots, \theta_{N_x}^n)$, $C^n := \mathcal{F}_{N_x}(\Theta^n)$ et $G^n := \mathcal{F}_{N_x}((Q(t^n, x_j))_{j[\![1,N_x]\!]})$. Le choix de schéma ne dépend que d'un choix de méthode de quadrature pour la partie liée au terme source. Le plus simple est pour $k \in\in [\![-\lfloor \frac{N_x}{2} \rfloor, N_x - \lfloor \frac{N_x}{2} \rfloor - 1]\!]$ et $n \in [\![0, N_t - 1]\!]$

$$C_k^{n+1} \;=\; \mathrm{e}^{-(\Delta t)\,(2\pi\,k)^2}\, C_k^n + \mathrm{e}^{-(\Delta t)\,(2\pi\,k)^2}\, G_k^n\, \Delta t\,.$$

Le schéma est directement écrit avec les coefficients de Fourier discrets et diagonal dans ces coordonnées. L'analyse de stabilité est immédiate : le schéma est inconditionnellement [4] L^2 stable de constante inférieure à 1 et la formule de Duhamel est diagonale

$$C_k^n \;=\; \mathrm{e}^{-t^n\,(2\pi\,k)^2}\, C_k^0 + \sum_{\ell=1}^{n} \mathrm{e}^{-(t^n - t^{\ell-1})\,(2\pi\,k)^2}\, G_k^{\ell-1}\, \Delta t\,.$$

L'analyse de consistance est assez directe quand on dispose déjà des estimations d'erreur pour la transformée de Fourier discrète. Sans terme source, le schéma est d'ordre infini en temps et en espace. Avec un terme source non nul, le schéma est d'ordre 1 en temps en temps et d'ordre infini en espace. La limitation d'ordre en temps est due dans ce cas au fait que nous avons choisi une méthode de quadrature d'ordre 1 pour l'intégrale de la formulation de Duhamel entre t^n et t^{n+1}.

Quand le terme source est vraiment une donnée on peut facilement monter en ordre en choisissant une autre méthode de quadrature. Quand le terme source est une fonction de l'inconnue, alors comme dans les méthodes de résolution d'équations différentielles ordinaires, il faut faire attention à écrire une méthode qui n'utilise que des valeurs déjà calculées. On peut par exemple utiliser un schéma comme ci-dessus dans des étages de type Runge-Kutta. Notons par ailleurs que quand le terme source est une fonction de l'inconnue (exprimée en fonctions des valeurs pas des coefficients de Fourier), pour le calculer on est de fait obligé de faire des allers-retours réguliers entre les coefficients de Fourier et les valeurs. On rappelle qu'il faut faire attention au problème de recentrage lorsqu'on utilise les boîtes noires de calcul.

De manière générale, quand on calcule une approximation d'évolution il vaut mieux n'afficher qu'une fraction des pas de temps calculés pour éviter l'accumulation des données, le ralentissement de l'affichage ou l'effet de scintillement, sans perdre en précision. Pour les schémas spectraux c'est d'autant plus

3. On utilise sans rappel les normalisations des notes de rappel.
4. Pas de condition CFL.

avantageux que quand le terme source est donné l'on ne calculera alors le retour en espace que pour les pas de temps que l'on souhaite afficher.
Comme toujours l'ordre maximal du schéma peut se trouver réduit par le défaut de régularité des termes sources ou de la donnée initiale. On rappelle qu'ici la régularité est mesurée par la décroissance des coefficients de Fourier, ce qui correspond à de la régularité en tant que fonction périodique, ce qui se traduit par de la régularité sur $[0,1]$ plus des raccordements entre les valeurs de la fonction et de ses dérivées en 0 et en 1.
Quand on veut appliquer ce qui précède à la résolution des problèmes avec conditions de bord de Dirichlet ou de Neumann on a intérêt à choisir une discrétisation qui préserve la parité/imparité. Ici il faut seulement faire attention à ce que le domaine soit symétrique par rapport à 0. Cela correspond à choisir N_x impair de sorte qu'avec $n_x := \lfloor \frac{N_x}{2} \rfloor$, l'on ait

$$[\![-\lfloor \tfrac{N_x}{2} \rfloor, N_x - \lfloor \tfrac{N_x}{2} \rfloor - 1]\!] = [\![-n_x, n_x]\!] .$$

L'imparité au temps t^n (liée aux conditions de Dirichlet sur $[0,1/2]$) se lit alors sur $C_k^n = -C_{-k}^n$ alors que la parité au temps t^n (liée aux conditions de Neumann sur $[0,1/2]$) se lit alors sur $C_k^n = C_{-k}^n$. Il est immédiat de vérifier que le schéma préserve ces conditions. La régularité se traduit par de la régularité sur $[0,1/2]$ plus des annulations des dérivées (les paires pour Dirichlet, les impaires pour Neumann) en 0 et en $1/2$.
On montre aussi sans difficulté que dans le cas sans terme source, sans condition sur les pas, le schéma respecte le comportement en temps long : convergence exponentiellement rapide vers la moyenne pour les cas périodique et Neumann, vers zéro pour le cas Dirichlet.
Les méthodes spectrales ont beaucoup d'avantages. Elles ont cependant des inconvénients évidents.

1. Pour ce qui est de l'équation de la chaleur, pour des coefficients de diffusion non constants les méthodes spectrales ne sont plus vraiment applicables.
2. Elles perdent un peu de leur efficacité quand le terme source dépend de l'inconnu, à cause des allers-retours espace-fréquence.
3. On ne peut pas les étendre facilement aux équation transport-diffusion avec un champ de vitesse non constant. On peut cependant utiliser une technique de type *splitting* basée sur un développement du type $\mathrm{e}^{\Delta t(A+B)} = \mathrm{e}^{\Delta t\, A}\,\mathrm{e}^{\Delta t\, B} + \mathcal{O}((\Delta t)^2)$ ou $\mathrm{e}^{\Delta t(A+B)} = \mathrm{e}^{\frac{1}{2}\Delta t\, A}\,\mathrm{e}^{\Delta t\, B}\,\mathrm{e}^{\frac{1}{2}\Delta t\, A} + \mathcal{O}((\Delta t)^3)$. Cela permet d'alterner des schémas pour la partie transport et des schémas pour la partie diffusion. On notera cependant que dans le cas des équations de transport-diffusion, les astuces de réduction des conditions de Dirichlet ou Neumann à celles périodiques ne fonctionnent plus, ce qui réduit les méthodes spectrales au cas périodique.
4. Elles requièrent des points équidistants, donc ne permettent pas des techniques adaptatives en espace.

Quand elles fonctionnent, elles sont cependant très rapides et très faciles à programmer (avec les boîtes noires Fourier discret).

3 Méthode des différences finies

La méthode des différences finies est moins efficace mais plus robuste. En particulier, bien que nous n'en fassions l'analyse qu'avec des pas équidistants et des conditions de bord périodiques, le schéma fonctionne bien plus largement.

On se contentera d'analyser le cas sans source.

3.1 Conditions de bord périodiques

Comme ci-dessus, étant donnés des pas d'espace et de temps Δx et Δt tels que $N_x := 1/\Delta x \in \mathbf{N}^*$, on note N_t la partie entière de $T/\Delta t$ et on introduit les nœuds de subdivisions $x_j = j\Delta x$, $j \in [\![1, N_x]\!]$, et $t_n = n\Delta t$, $n \in [\![0, N_t]\!]$. On souhaite calculer θ_j^n, des valeurs en (t_n, x_j) censées fournir une approximation des valeurs $\theta(t_n, x_j)$ de la solution. Insistons sur le fait que bien nous ne le marquions pas par des indices et des exposants, (θ_j^n, t_n, x_j) ne dépendent pas que de (j, n) mais aussi de $(\Delta t, \Delta x)$ ou de manière équivalente de $(\Delta t, N_x)$.
Le schéma le plus simple est

$$\theta_j^{n+1} = \theta_j^n + \frac{\Delta t}{(\Delta x)^2}\left(\theta_{j+1}^n - 2\theta_j^n + \theta_{j-1}^n\right), \qquad j \in [\![1, N_x]\!], \qquad n \in [\![0, N_t - 1]\!], \tag{DFc}$$

où les valeurs en dehors de $j \in [\![1, N_x]\!]$ sont obtenues par périodicité, $\theta_0^n := \theta_{N_x}^n$, $\theta_{N_x+1}^n := \theta_1^n$. En Fourier, il devient

$$C_k^{n+1} = \left(1 + 2\frac{\Delta t}{(\Delta x)^2}\,(\cos(k\,2\pi\,\Delta x) - 1)\right) C_k^n, \qquad k \in [\![-\lfloor \tfrac{N_x}{2} \rfloor, N_x - \lfloor \tfrac{N_x}{2} \rfloor - 1]\!], \qquad n \in [\![0, N_t - 1]\!].$$

Si on fixe $\lambda := \Delta t/(\Delta x)^2$, l'analyse de stabilité L^2 montre que le schéma est instable si $\lambda > 1/2$ (et Δx est assez petit) et stable si $\lambda \le 1/2$.
L'analyse de consistance donne de l'ordre 1 en temps et 2 en espace. On peut la mener en espace ou en fréquence. En fréquence, cela se lit sur

$$\frac{\mathrm{e}^{-(2\pi k)^2\,\Delta t} - \left(1 + 2\,\frac{\Delta t}{(\Delta x)^2}\,(\cos(2\pi\,k\,\Delta x) - 1)\right)}{\Delta t} \overset{(\Delta t, \Delta x) \to (0,0)}{=} \mathcal{O}(\Delta t + (\Delta x)^2).$$

Sous la condition CFL $\lambda \le 1/2$ on peut aussi montrer la stabilité L^∞ (avec constante 1).
Sous la condition CFL $\lambda < 1/2$ le schéma respecte le comportement en temps long.

3.2 Conditions de Dirichlet ou de Neumann

Quand le problème périodique a été obtenu à partir d'un problème de Dirichlet ou de Neumann sur $[0, 1/2]$, le schéma ci-dessus peut être projeté comme un schéma sur $[0, 1/2]$.
Supposons N_x impair et posons $n_x := \lfloor \frac{N_x}{2} \rfloor$.
Pour les conditions de Dirichlet sur $[0, 1/2]$, on obtient $\theta_0^n = 0$, $\theta_{n_x}^n = 0$ et l'on calcule récursivement θ_j^n, $j \in [\![1, n_x - 1]\!]$, grâce à

$$\theta_j^{n+1} = \theta_j^n + \frac{\Delta t}{(\Delta x)^2}\left(\theta_{j+1}^n - 2\theta_j^n + \theta_{j-1}^n\right), \qquad j \in [\![1, n_x - 1]\!], \qquad n \in [\![0, N_t - 1]\!].$$

Pour les conditions de Neumann sur $[0, 1/2]$, on calcule récursivement θ_j^n, $j \in [\![0, n_x]\!]$, grâce à

$$\theta_j^{n+1} = \theta_j^n + \frac{\Delta t}{(\Delta x)^2}\left(\theta_{j+1}^n - 2\theta_j^n + \theta_{j-1}^n\right), \qquad j \in [\![0, n_x]\!], \qquad n \in [\![0, N_t - 1]\!],$$

avec $\theta_{-1}^n = \theta_1^n$ et $\theta_{n_x+1}^n = \theta_{n_x-1}^n$.

3.3 Version implicite

La contrainte CFL est plus forte que dans le cas des équations de transport, le pas de temps devant être quadratiquement petit en le pas d'espace.
Pour lever cette contrainte, on peut préférer une version implicite

$$\theta_j^{n+1} = \theta_j^n + \frac{\Delta t}{(\Delta x)^2}\left(\theta_{j+1}^{n+1} - 2\theta_j^{n+1} + \theta_{j-1}^{n+1}\right), \qquad j \in [\![1, N_x]\!], \qquad n \in [\![0, N_t - 1]\!],$$

avec $\theta_0^{n+1} := \theta_{N_x}^{n+1}$, $\theta_{N_x+1}^{n+1} := \theta_1^{n+1}$. L'analyse L^2 de ce schéma peut être obtenue de la manière que pour (DFc). On obtient les mêmes conclusions (convergence, ordre, temps long) sans condition CFL mais au prix de la résolution d'un système linéaire à chaque pas de temps.
On peut penser cette discrétisation comme une discrétisation en espace de la discrétisation en temps

$$\theta^{n+1} - (\Delta t)\,\partial_x^2\theta^{n+1} \;=\; \theta^n$$

qui a une forme familière. En particulier on a reproduit ci-dessus au niveau discret la structure de Riesz/Lax-Milgram

$$\Theta^{n+1} + (\Delta t)\,D_{\Delta x}^*\,D_{\Delta x}\Theta^{n+1} \;=\; \Theta^n$$

avec au choix $D_{\Delta x}(v) = (v_{j+1} - v_j)/(\Delta x)$ d'adjoint [5] $D_{\Delta x}^*(v) = -(v_j - v_{j-1})/(\Delta x)$, ou l'inverse avec $D_{\Delta x}(v) = (v_j - v_{j-1})/(\Delta x)$ et $D_{\Delta x}^*(v) = -(v_{j+1} - v_j)/(\Delta x)$. Elle est associée à la norme hilbertienne de carré $\Theta \mapsto \|\Theta\|_{\ell^2}^2 + (\Delta t)\,\|D_{\Delta x}\Theta\|_{\ell^2}^2$.

Références

[Fil13] F. Filbet. *Analyse numérique - Algorithme et étude mathématique.* Dunod, 2013.

5. Pour rendre le calcul d'adjoint rigoureux il faut inclure les conditions de bord.