



## L'EQUATION DE LA CHALEUR ET L'ENSEIGNEMENT DES METHODES DE SIMULATION NUMERIQUE DE LA PHYSIQUE

Eric Deleersnijder

Institut d'Astronomie et de Géophysique G. Lemaître  
Université Catholique de Louvain, 2 Chemin du cyclotron, 1348 Louvain-la-Neuve

*Although [...] more engineers now use "off-the-shelf" software packages, it is still important to understand what is in them and how to use them properly. It is the author's experience that one can obtain this kind of understanding only by actually writing at least some simple programs and "playing" with the methods; one does not really understand a method until one has programmed it. I have met many people who use commercial packages and complain about their experience with them. It often turns out that they are either asking a program to do something it was not intended to do or that they are trying to solve a problem that does not have a solution (an ill-posed problem). A little background in the fundamentals of how the method works, what it is good for, and how it should be used would have helped these people enormously.*

Joel H. Ferziger (1998) (à propos des méthodes numériques de l'ingénierie)

**Summary.** Numerically solving partial differential problems to simulate the behaviour of natural or artificial systems is now part of the scientific method — along with the experimental and theoretical approaches — mainly because the power of the computers has enormously increased over the last decades. This is the reason why it is argued that physics students should be given an initiation to methods for solving partial differential problems — preferably during their third year at the university. It is believed that basic properties of finite-difference numerical schemes — *i.e.* consistency, stability, convergence, conservativity, accuracy, efficiency — may be introduced and illustrated while building methods for numerically solving simple partial differential equations supposedly known to all physicists, such as the heat equation. Accordingly, several heat conduction problems are set up in such a way that the analytical solutions of the continuous problems and their discrete counterparts are available, rendering it easy to understand the properties of the corresponding solution methods. A particular emphasis is put on one-dimensional problems, but a few two-dimensional methods are also considered. Finally, the importance of advective heat transport processes is mentioned, as well as the specific difficulties relevant to the numerical treatment of these phenomena. Hopefully, these examples will be useful to those willing to introduce physics students to the fundamentals of numerical simulation in physics.

## Introduction

Depuis le XVII<sup>ème</sup> siècle environ, on a construit des modèles mathématiques<sup>1</sup>, c'est-à-dire des relations mathématiques susceptibles de décrire l'état de systèmes naturels ou artificiels, ainsi que l'évolution de cet état. En d'autres termes, ces systèmes ont été "mis en équations". Ces dernières sont souvent des équations différentielles ordinaires ou aux dérivées partielles. On a établi de nombreuses propriétés qualitatives utiles des solutions de ces équations, mais il a rarement été possible d'en découvrir les solutions analytiques générales. Dans la plupart des cas, obtenir des solutions analytiques n'a été possible que pour des problèmes idéalisés, découplant d'hypothèses fortes sur la symétrie des solutions, la forme du domaine d'intérêt, la proximité d'un état d'équilibre, etc. Ces difficultés mathématiques ont sans doute freiné l'interaction entre les développements théoriques et l'expérimentation ou l'observation.

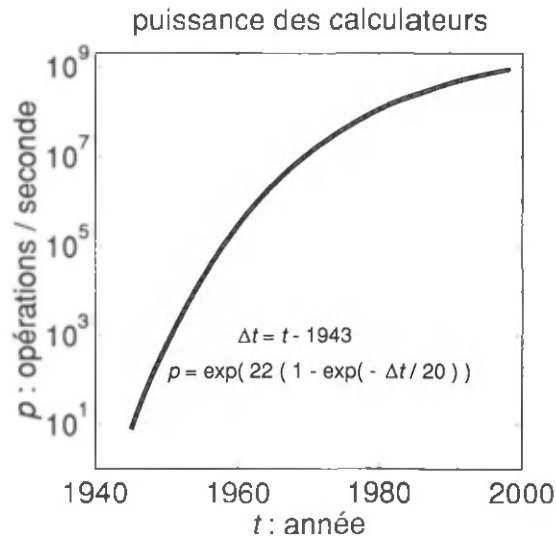
Rechercher une solution approximative d'un problème mathématique constitue fréquemment une alternative fructueuse à la détermination de la solution exacte. Ainsi, les solutions asymptotiques sont connues depuis plusieurs siècles, mais on ne peut construire de telles approximations que dans des conditions bien précises, qui limitent drastiquement la construction et l'emploi de ces expressions. D'un autre côté, il est quasiment toujours possible de bâtir des algorithmes numériques susceptibles de fournir une approximation de la solution exacte d'un problème. Cette approche n'a cependant pris un réel essor qu'avec l'apparition de calculateurs suffisamment puissants (Figure 1), c'est-à-dire vers la fin des années septante. Les premiers ordinateurs, qui datent des années quarante, étaient capables d'effectuer quelques opérations arithmétiques par seconde. Aujourd'hui, une station de travail commune peut réaliser, chaque seconde, de l'ordre de 100 millions de multiplications de nombres comportant environ 16 chiffres significatifs. Une telle puissance de calcul permet de produire des approximations de la solution de problèmes mathématiques relativement complexes de sorte qu'on peut maintenant simuler numériquement avec un haut niveau de réalisme le comportement de nombreux systèmes naturels ou artificiels. On peut affirmer que la simulation numérique est maintenant devenue un acteur à part entière de la démarche scientifique<sup>2</sup>, au même titre que les développements théoriques et l'expérimentation en laboratoire, ou encore l'observation sur le terrain.

Aujourd'hui, la simulation numérique tend à se substituer aux expériences sur des systèmes en grandeur réelle ou réduite dès que ces dernières sont plus chères que la simulation — à condition, bien entendu, que la simulation du système considéré soit possible. Ainsi, par exemple, on remplace de plus en plus les essais en soufflerie par des programmes informatiques visant à résoudre les équations de l'aérodynamique — que l'on qualifie parfois de "soufflerie numérique". D'autre part, il existe des systèmes inaccessibles à l'expérimentation au sens classique du terme. Le système climatique de la Terre appartient clairement à cette catégorie: on

<sup>1</sup> Les *Principia* de Newton, publiés en 1687, furent sans doute le premier grand succès de la modélisation mathématique, en prouvant clairement la capacité des modèles mathématiques à représenter et, dans une certaine mesure, expliquer des phénomènes naturels.

<sup>2</sup> A propos de la simulation numérique, on appréciera ces mots de Robinson (1987): "The use of such simulation in scientific research (numerical experimentation, sensitivity and process studies, etc.) is thought by many to represent the first major step forward in the basic *scientific method* since the seventeenth century. Science is now a tripartite endeavor with simulation added to the two classical components, Experiment and Theory".

ne peut “mettre un tel système en éprouvette” pour étudier et comprendre son fonctionnement, mais on peut simuler son comportement à l’aide de moyens informatiques et pratiquer alors des “expériences numériques”.



**Figure 1.** Courbe représentant l’évolution depuis la fin des années quarante de la vitesse approximative du processeur le plus puissant de l’époque, exprimée en nombre d’opérations arithmétiques par seconde (selon Hack, 1992).

Nombre de simulations numériques reposent sur la production d’approximations discrètes de la solution de problèmes différentiels. Les algorithmes correspondants sont basés sur des schémas numériques dont la mise au point n’est généralement pas triviale. S’il existe actuellement des logiciels commerciaux — comme *Mathematica* ou *Matlab* — permettant de fournir des solutions numériques de problèmes différentiels aux dérivées ordinaires — sans que leur utilisateur doive s’impliquer dans les détails des schémas numériques —, la résolution numérique de problèmes différentiels aux dérivées partielles exige souvent la mise au point d’un algorithme spécialement conçu pour le problème considéré. Le scientifique ne peut donc s’en remettre à un logiciel qui prendrait en charge les “détails techniques” de la résolution numérique du problème. Et même quand un tel logiciel existe, l’utiliser valablement demande des compétences numériques significatives. En d’autres termes, pratiquer la simulation numérique — et, en particulier, la résolution numérique de problèmes différentiels aux dérivées partielles — ne s’improvise pas.

Pour simuler numériquement un système, il faut d’abord établir les équations qui régissent l’état de ce système. Ensuite, il faut bâtir une ou plusieurs approximations discrètes de ces équations et, finalement, implémenter sur un ordinateur les programmes correspondants. On peut penser que le programme de la licence en physique des universités francophones belges

couvre actuellement très bien la première étape, à savoir la mise en équations du système, ainsi que l'établissement des propriétés qualitatives des solutions, la découverte de solutions analytiques dans des cas idéalisés et, dans une moindre mesure, la construction d'approximations asymptotiques. Par contre, il est regrettable que, dans la plupart des cas, les étudiants ne connaissent quasiment rien des méthodes numériques visant à obtenir des approximations discrètes des problèmes différentiels — surtout aux dérivées partielles. Il serait peut-être opportun de combler cette lacune, afin que nos licenciés en physique soient mieux armés pour obtenir des emplois dans les domaines — toujours plus nombreux — de la recherche fondamentale ou appliquée où la simulation numérique joue un rôle important.

Initier les étudiants aux méthodes de simulation numérique de la physique — et, en particulier, à la résolution numérique d'équations différentielles aux dérivées partielles — n'implique pas nécessairement une lourde charge d'enseignement. En effet, en première licence<sup>3</sup>, on pourrait se fixer l'objectif d'introduire et illustrer la plupart des notions de base de la technique des différences finies<sup>4</sup> en traitant quelques problèmes différentiels élémentaires qui concernent tous les physiciens. Ci-dessous, on va montrer que la résolution de problèmes de conduction de chaleur peut aisément aider à atteindre ce but.

### Equation de la chaleur uni-dimensionnelle

On considère le problème classique de conduction de chaleur dans une barre isolée — sauf en ses extrémités. La longueur de la barre vaut  $2L$ , et la position de chaque point de cette barre est repérée par la coordonnée  $x$ , telle que  $-L \leq x \leq L$ . On divise la barre en  $2K$  segments de longueur égale  $\Delta x = (2L)/(2K) = L/K$ . On identifie chaque segment à l'aide de l'indice  $k$ , qui peut prendre les valeurs suivantes (Figure 2):

$$k = -K+1, -K+2, \dots, K-1, K. \quad (1)$$

Le centre du segment associé à l'indice  $k$  est situé à la coordonnée

$$x_k = (k - 1/2) \Delta x. \quad (2)$$

Les extrémités du segment  $k$  sont donc situées en  $x_{k \pm 1/2} = (k - 1/2 \pm 1/2) \Delta x$ . La température évolue dans le temps, que l'on note  $t$ . On identifie des instants successifs, séparés par la durée

<sup>3</sup> La première licence pourrait être le moment idéal pour initier les étudiants à la résolution numérique d'équations différentielles aux dérivées partielles. En candidature, la maturité scientifique — notamment les bases mathématiques et la capacité à établir des relations fructueuses entre les matières de différents cours — est insuffisante. En seconde licence, l'initiation aux méthodes dont il est question ici viendrait trop tard pour qu'un mémoire y faisant largement appel puisse être entrepris. Et, comme le mémoire conditionne souvent le choix des études et recherches de troisième cycle — ou, parfois, l'obtention d'un emploi dans l'industrie —, ne pas être en mesure de pratiquer valablement la simulation numérique en seconde licence pourrait revenir à y renoncer *de facto* pendant une période significative de la carrière.

<sup>4</sup> Les "différences finies" sont une famille de méthodes numériques couramment employées pour résoudre des problèmes différentiels. Schématiquement, on peut dire que les autres familles sont les "éléments finis" et les "méthodes spectrales". Si les frontières entre ces trois classes de méthodes sont plutôt floues, il est par contre très clair que les différences finies constituent la catégorie la plus simple à mettre en oeuvre — et donc à enseigner. Cependant, les concepts de base, comme la précision, stabilité, la convergence, etc., des schémas numériques sont valables pour toute méthode. Donc, les enseigner dans le cadre de la méthode des différences finies ne peut que faciliter l'apprentissage ultérieur d'autres familles de méthodes numériques.

$\Delta t$ , à l'aide de l'indice

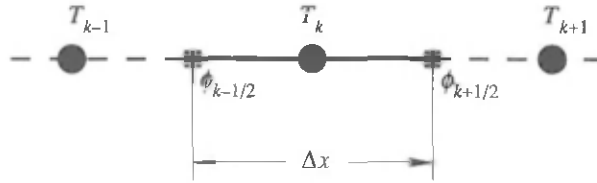
$$l = 0, 1, 2, \dots, \quad (3)$$

de telle sorte que

$$t_l = l \Delta t. \quad (4)$$

Pour toute variable  $a$ , on utilise les indices  $k$  et  $l$  comme suit:

$$a_{l,k} \equiv a(t_l, x_k). \quad (5)$$



**Figure 2.** Représentation schématique du  $k$ -ème segment de barre, avec la position des températures et flux de chaleur.

L'énergie interne comprise dans le  $k$ -ème segment est définie comme l'intégrale sur le segment considéré du produit de la masse volumique  $\rho$ , de la chaleur spécifique  $c$  et de la température  $T$ . La barre étant isolée, le premier principe de la thermodynamique indique que seuls les flux de chaleur aux extrémités du segment considéré sont susceptibles de faire varier son énergie interne. Ainsi, si  $\phi$  désigne le flux de chaleur transitant dans la barre — associé au phénomène de “conduction” — le bilan d'énergie du  $k$ -ème segment s'écrit

$$\int_{x_{k-1/2}}^{x_{k+1/2}} \rho c T(t_{l+1}, x) dx = \int_{x_{k-1/2}}^{x_{k+1/2}} \rho c T(t_l, x) dx + \int_{t_l}^{t_{l+1}} [\phi(t, x_{k-1/2}) - \phi(t, x_{k+1/2})] dt. \quad (6)$$

Selon la loi de Fourier, le flux de chaleur tend à transporter de la chaleur de la région la plus chaude à la région la plus froide et peut être paramétrisé comme suit:

$$\phi = -\kappa \frac{\partial T}{\partial x}, \quad (7)$$

où  $\kappa$  représente la conductivité thermique.

Les équations (6) et (7) sont les ingrédients essentiels à l'établissement de l'équation différentielle aux dérivées partielles régissant l'évolution de la température — communément appelée “équation de la chaleur” — et d'une approximation discrète de cette même équation.

A partir de ce point, pour simplifier les développements mathématiques, on supposera que la masse volumique, la chaleur spécifique et la conductivité thermique sont constantes. Par conséquent, la diffusivité de température,

$$\lambda = \frac{\kappa}{\rho c}, \quad (8)$$

est également constante.

En introduisant les développements de Taylor

$$T(t, x) = T(t, x_k) + \sum_{m=1}^{\infty} \left( \frac{\partial^m T}{\partial x^m} \right)_{x_k} \frac{(x - x_k)^m}{m!} \quad (9)$$

et

$$\phi(t, x) = \phi(t_l, x) + \sum_{m=1}^{\infty} \left( \frac{\partial^m \phi}{\partial t^m} \right)_{t_l} \frac{(t - t_l)^m}{m!} \quad (10)$$

dans (6), on obtient, après quelques manipulations,

$$\rho c \frac{T_{l+1,k} - T_{l,k}}{\Delta t} = - \frac{\phi_{l,k+1/2} - \phi_{l,k-1/2}}{\Delta x} + O(\Delta t, \Delta x^2) . \quad (11)$$

Si l'on augmente indéfiniment le nombre de segments, l'expression (11) s'applique alors "en chaque point" de la barre. Ainsi, à la limite  $\Delta x/L \rightarrow 0$  ou  $K \rightarrow \infty$ , (11) conduit à l'équation continue

$$\rho c \frac{\partial T}{\partial t} = - \frac{\partial \phi}{\partial x} . \quad (12)$$

En intégrant (12) sur le domaine d'intérêt, on obtient la relation régissant l'évolution de l'énergie interne de la barre,

$$E(t) = \int_{-L}^{+L} \rho c T \, dx , \quad (13)$$

à savoir

$$E(t_b) = E(t_a) + \int_{t_a}^{t_b} [\phi(t, x = -L) - \phi(t, x = +L)] \, dt , \quad (14)$$

où  $t_a$  et  $t_b$  sont des instants arbitrairement choisis. Ceci montre que l'énergie interne de la barre ne varie qu'en fonction du bilan des flux thermiques en ses extrémités. C'est "physiquement" facile à comprendre: la barre étant isolée, la chaleur ne peut en "sortir" ou y "entrer" que par ses extrémités! Par conséquent, si les deux extrémités de la barre sont également isolées — aucun flux de chaleur ne les traverse —, alors l'énergie interne totale de la barre reste constante.

Substituant la loi de Fourier (7) dans l'équation de bilan de chaleur (12), utilisant (8), on obtient l'équation de la chaleur uni-dimensionnelle,

$$\frac{\partial T}{\partial t} = \lambda \frac{\partial^2 T}{\partial x^2} , \quad (15)$$

qui est le prototype des *équations paraboliques*<sup>5</sup>. Si l'on adjoint à cette équation la distribution initiale de température et des conditions aux limites adéquates, on obtient un *problème*

<sup>5</sup> Hirsch (1988) (section 3.2.3) fait partie des auteurs qui exposent les critères permettant de ranger une équation différentielle aux dérivées partielles quasi-linéaire donnée dans la catégorie elliptique, parabolique, hyperbolique, ou dans une classe hybride entre deux des trois formes citées précédemment.

différentiel aux dérivées partielles dont la solution est la distribution de la température dans la barre à tout instant postérieur à l'instant initial. On peut donc prévoir l'évolution du champ thermique dans la barre — si l'on est en mesure de résoudre le présent problème différentiel.

### Equation discrète

La plupart des problèmes différentiels aux dérivées partielles n'ont pas de solution analytique générale connue. Par contre, on peut chercher des approximations numériques de leur solution par la méthode des différences finies<sup>6</sup>. Il faut alors renoncer à déterminer la valeur de la solution en chaque point et se contenter de l'estimer en un nombre limité de points. Pour ce faire, la relation (11) constitue un point de départ utile. En effet, en négligeant les termes d'ordre supérieur en les incréments temporel et spatial —  $O(\Delta t, \Delta x^2)$  —, on obtient l'algorithme discret

$$T_{l+1,k}^* = T_{l,k}^* - \frac{\Delta t}{\rho c} \frac{\phi_{l,k+1/2}^* - \phi_{l,k-1/2}^*}{\Delta x}, \quad (16)$$

où chaque variable munie d'un astérisque symbolise l'équivalent discret de la variable continue correspondante — qui est dépourvue d'astérisque. L'expression (16) n'est évidemment pas la seule expression discrète susceptible de fournir une approximation de la solution de (12), mais elle est sans doute la plus simple.

L'énergie interne de la solution discrète est

$$E_l^* = \sum_{k=-K+1}^K \rho c T_{l,k}^* \Delta x. \quad (17)$$

En combinant (16) et (17), on obtient l'équivalent discret de la relation (14):

$$E_{l_b}^* = E_{l_a}^* + \sum_{l=l_a}^{l_b-1} (\phi_{l,-K}^* - \phi_{l,+K}^*) \Delta t. \quad (18)$$

où  $l_a$  et  $l_b$  sont des nombres entiers dont la valeur peut être choisie arbitrairement. Donc, pour la solution discrète, comme pour la solution continue, seuls les flux de chaleur traversant les extrémités de la barre sont susceptibles de modifier l'énergie interne: on dit que le schéma numérique (16) est *conservatif*. Ce résultat ne dépend en aucune manière de la façon dont on estime les flux de chaleur à l'intérieur de la barre. Ainsi, par exemple, cette propriété resterait valable si on évaluait les flux de chaleur à l'intérieur de la barre à l'aide d'un générateur de nombres aléatoires — une option de modélisation qui ne fournirait évidemment pas la meilleure approximation numérique de la solution...

En manipulant le développement de Taylor de la température

$$T(t, x) = T(t, x_{k\pm 1/2}) + \sum_{m=1}^{\infty} \left( \frac{\partial^m T}{\partial x^m} \right)_{x_{k\pm 1/2}} \frac{(x - x_{k\pm 1/2})^m}{m!}, \quad (19)$$

on peut obtenir une approximation discrète simple de la loi de Fourier,

<sup>6</sup> Ici, on ne fait pas de distinction entre "différences finies" et "volumes finis".

$$\phi_{l,k\pm 1/2}^* = -\kappa \frac{\pm T_{l,k\pm 1}^* \mp T_{l,k}^*}{\Delta x} . \quad (20)$$

En introduisant cette dernière dans l'algorithme (16), il vient

$$T_{l+1,k}^* = T_{l,k}^* + \frac{\lambda \Delta t}{\Delta x^2} (T_{l,k+1}^* - 2T_{l,k}^* + T_{l,k-1}^*) \quad (21a)$$

ou, une forme équivalente,

$$\frac{T_{l+1,k}^* - T_{l,k}^*}{\Delta t} = \lambda \frac{T_{l,k+1}^* - 2T_{l,k}^* + T_{l,k-1}^*}{\Delta x^2} . \quad (21b)$$

Cet algorithme constitue une discrétisation élémentaire de l'équation de la chaleur (15), qui permet d'estimer la température à l'instant  $t_{l+1}$  aux points  $x_k$  si la température est connue à l'instant  $t_l$  aux points appropriés. Ainsi, si l'on dispose de conditions initiales et aux limites adéquates, on peut actualiser progressivement la température aux points  $x_k$  du domaine d'intérêt.

En manipulant les séries de Taylor adéquates, on peut ré-écrire (21) sous la forme

$$\left( \frac{T_{l+1,k}^* - T_{l,k}^*}{\Delta t} - \lambda \frac{T_{l,k+1}^* - 2T_{l,k}^* + T_{l,k-1}^*}{\Delta x^2} \right) - \left[ \left( \frac{\partial T^*}{\partial t} \right)_{l,k} - \lambda \left( \frac{\partial^2 T^*}{\partial x^2} \right)_{l,k} \right] = e_{l,k} , \quad (22)$$

où  $e_{l,k}$  désigne l'erreur de troncature, c'est-à-dire la différence, au point de discrétisation de la température  $(t_l, x_k)$ , entre l'équation aux différences (21b) et l'équation de la chaleur,

$$e_{l,k} = \frac{1}{2} \left( \frac{\partial^2 T^*}{\partial t^2} \right)_{l,k} \Delta t - \frac{1}{12} \left( \frac{\partial^4 T^*}{\partial x^4} \right)_{l,k} \Delta x^2 + O(\Delta t^2, \Delta x^4) . \quad (23)$$

Cette dernière tend vers zéro à mesure que l'on diminue le pas de temps  $\Delta t$  et l'incrément spatial  $\Delta x$ . On dit que le schéma (21) est compatible avec l'équation de la chaleur (15), et est du premier ordre de précision selon la dimension temporelle et du second ordre selon l'espace. Si l'erreur de troncature n'était pas au moins du premier ordre en chaque incrément des variables indépendantes, alors cette erreur ne tendrait pas vers zéro à mesure que l'on diminue ces incréments, indiquant que l'expression discrète considérée, représenterait, en fait, une discrétisation d'une équation continue différente.

En examinant uniquement l'erreur de troncature, on pourrait croire que, pour augmenter la précision de l'approximation discrète de la solution, c'est-à-dire faire qu'elle soit plus proche de la solution exacte, il suffit d'augmenter l'ordre de précision du schéma et/ou de diminuer les incréments des variables indépendantes. Bien qu'elles semblent en accord avec le bon sens le plus élémentaire, ces suggestions ne sont que partiellement correctes: l'erreur de troncature ne représente pas une mesure très fiable de la précision du schéma numérique considéré. Deux raisons complémentaires expliquent cela:

— L'erreur de troncature comprend des dérivées de l'inconnue dont l'ordre est le plus souvent supérieur à celles que l'on trouve dans l'équation continue. Comme on n'a généralement que peu d'informations quantitatives ou simplement qualitatives sur ces dérivées, la seule information pertinente que l'on tire de l'analyse de l'erreur de troncature est la puissance la plus



petite à laquelle est élevé chaque incrément des variables indépendantes — ici, il s'agit de la première puissance pour l'incrément temporel et de la seconde pour le pas spatial. Or, une telle information se révèle en définitive peu utile car, dans un processus de résolution numérique, les incréments des variables indépendantes ne deviennent jamais arbitrairement petits<sup>7</sup>.

— L'erreur de troncature est une quantité à caractère local, car on l'évalue à l'aide des valeurs de la solution prises uniquement dans le voisinage du point considéré de l'espace des variables indépendantes. Par conséquent, cette quantité ne donne que peu de renseignements, voire aucun, sur la propagation et l'amplification éventuelle de la différence entre la solution exacte et son approximation numérique.

C'est parce qu'elle peut être évaluée assez facilement et, surtout, sans connaître la solution exacte de l'équation continue que l'erreur de troncature est très attrayante. Mais, comme on vient de le démontrer, cette quantité est peu pertinente et souvent trompeuse. Il faudrait plutôt analyser la *convergence* du schéma, c'est-à-dire le comportement de l'écart entre la solution exacte et son approximation numérique en fonction des valeurs des incréments des variables indépendantes. Pour ce faire, il faudrait disposer de la solution exacte. Mais, si celle-ci était connue, pourquoi alors en calculer une approximation discrète? On touche ici du doigt la difficulté majeure inhérente à tout processus de résolution numérique d'une équation différentielle aux dérivées partielles: *estimer la qualité de la solution numérique est très difficile*. Et, puisque l'évaluation de la solution discrète est malaisée, déterminer la façon de l'améliorer l'est encore moins...

La résolution numérique d'équations différentielles aux dérivées partielles n'est cependant pas un jeu de hasard. Différentes méthodes existent pour inférer certaines propriétés de l'approximation numérique choisie. Parmi celles-ci, l'analyse de la stabilité du schéma — pour les équations d'évolution — joue un rôle essentiel. Il s'agit d'empêcher que la solution numérique, pour des raisons indépendantes de la nature même du processus représenté par l'équation continue, présente des modes susceptibles de croître indéfiniment. Ainsi, un schéma d'un ordre de précision élevé au sens de l'erreur de troncature mais instable n'a aucune chance de fournir une approximation valable de la solution exacte, alors qu'un schéma d'un ordre de précision plus faible mais stable sera très certainement plus utile. Pour les problèmes différentiels aux dérivées partielles vérifiant des conditions malheureusement très restrictives, le théorème de Lax (Richtmeyer and Morton, 1957) stipule que tout schéma compatible et stable est également convergent. Pour la plupart des problèmes à caractère réaliste, ce théorème ne s'applique pas. On s'y réfère cependant, au moins implicitement, de façon heuristique: pour tout schéma numérique, on se doit de démontrer la compatibilité avec l'équation continue et analyser, dans la mesure du possible, la stabilité — ce qui exige souvent une ou plusieurs simplifications du schéma, qui comprennent la linéarisation si celui-ci est non-linéaire.

---

<sup>7</sup> La limitation de la mémoire et de la puissance de calcul des ordinateurs oblige souvent le modélisateur à employer des pas de temps et d'espace qui sont loin d'être beaucoup plus petits que les temps et longueurs caractéristiques des phénomènes à simuler. Ainsi, la "grossièreté" relative du maillage employé constitue l'une des excuses les plus couramment invoquées pour justifier la piètre qualité d'une solution numérique.

### Résolution d'un problème de conduction de chaleur

On va maintenant illustrer la discussion de la qualité d'une approximation numérique en poursuivant l'étude du problème de conduction de chaleur entamée plus haut. Pour cela, on doit prescrire des conditions initiales et aux limites. On considère que les deux extrémités de la barre sont isolées. Par conséquent, à tout instant, les conditions aux limites

$$\left( -\lambda \frac{\partial T}{\partial x} \right)_{x=\pm L} = 0 \quad (24)$$

s'appliquent. Le profil initial de température est en "marche d'escalier":

$$T(t=0, x) = T_+, \quad \text{pour } 0 < x \leq L, \quad (25a)$$

$$T(t=0, x) = T_-, \quad \text{pour } -L \leq x < 0, \quad (25b)$$

où  $T_+$  et  $T_-$  sont des constantes. Le problème à résoudre se compose donc de l'équation (15) et des conditions (24)-(25).

**Tableau I.** Définition des problèmes adimensionnels continu et discret de conduction de chaleur à résoudre.

| problème continu                                                           | problème discret                                                                                       |
|----------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------|
| domaine d'intérêt                                                          | domaine d'intérêt                                                                                      |
| $t \geq 0, -1 \leq x \leq +1$                                              | $n = 0, 1, 2, \dots; -K+1 \leq k \leq K$                                                               |
| équation d'évolution                                                       | équation d'évolution                                                                                   |
| $\frac{\partial T}{\partial t} = \frac{\partial^2 T}{\partial x^2}$        | $\frac{T_{l+1,k}^* - T_{l,k}^*}{\Delta t} = \frac{T_{l,k+1}^* - 2T_{l,k}^* + T_{l,k-1}^*}{\Delta x^2}$ |
| condition initiale                                                         | condition initiale                                                                                     |
| $T(t=0, x) = -1$ si $-1 \leq x < 0$<br>$T(t=0, x) = +1$ si $0 < x \leq +1$ | $T_{0,k} = -1$ si $-K+1 \leq k \leq 0$<br>$T_{0,k} = +1$ si $1 \leq k \leq +K$                         |
| conditions aux limites                                                     | conditions aux limites                                                                                 |
| $\left( \frac{\partial T}{\partial x} \right)_{x=\pm 1} = 0$               | $\frac{T_{l,\pm K+1} - T_{l,\pm K}}{\Delta x} = 0$                                                     |

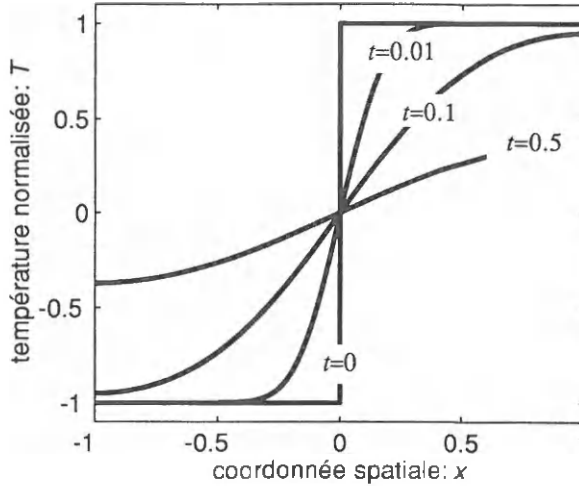
Le caractère général de la solution augmente et les expressions mathématiques deviennent plus simples si l'on rend toutes les variables adimensionnelles. On définit ces dernières comme suit:

$$(t', x', T') = \left( \frac{\lambda t}{L^2}, \frac{x}{L}, \frac{T - (T_+ + T_-)/2}{(T_+ - T_-)/2} \right), \quad (26)$$

On répète la même opération pour les variables de l'approximation discrète de la solution. En particulier, le pas de temps et le pas spatial adimensionnels s'écrivent

$$(\Delta t', \Delta x') = \left( \frac{\lambda \Delta t}{L^2}, \frac{\Delta x}{L} = \frac{1}{K} \right). \quad (27)$$

A partir de ce point et jusqu'à nouvel ordre, on considèrera uniquement les variables adimensionnelles et, pour simplifier les notations, on omettra les primes qui caractérisent ces variables. On peut alors reformuler le problème continu et son pendant discret comme indiqué dans le Tableau I.



**Figure 3.** Représentation de la solution exacte (28)-(29) du problème de conduction de chaleur à différents instants, à savoir  $t = 0, 0.01, 0.1$  et  $0.5$ .

La solution du problème continu posé dans le Tableau I s'exprime comme une somme infinie de "modes de Fourier" (Figure 3):

$$T(t, x) = \sum_{n=1}^{\infty} \alpha_n e^{-\varepsilon_n t} \sin(\mu_n x), \quad (28)$$

où  $\alpha_n$ ,  $\varepsilon_n$  et  $\mu_n$  représentent, pour le  $n$ -ième mode, l'amplitude, le coefficient d'amortissement et le nombre d'onde. Ces paramètres valent

$$(\alpha_n, \varepsilon_n, \mu_n) = \left( \frac{2}{(n-1/2)\pi}, (n-1/2)^2 \pi^2, (n-1/2)\pi \right). \quad (29)$$

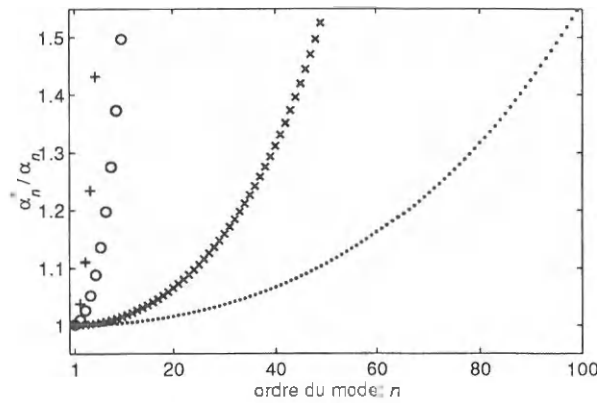
En faisant appel à la théorie des équations aux différences (voir, par exemple, Bender and Orszag, 1978) on obtient<sup>8</sup> la solution analytique du problème discret du Tableau I,

$$T_{l,k}^* = \sum_{n=1}^K \alpha_n^* e^{-\varepsilon_n^* t} \sin(\mu_n^* x) . \quad (30)$$

Cette série, au contraire de celle constituant la solution continue, ne comprend qu'un nombre fini de modes,  $K$  en l'occurrence. Si le nombre d'onde de la solution discrète équivaut à celui du mode correspondant de la solution continue,

$$\mu_n^* = (n - 1/2)\pi = \mu_n , \quad \text{pour } n \leq K , \quad (31)$$

l'amplitude et le coefficient d'amortissement de la solution discrète ne sont pas égaux aux paramètres correspondants de la solution continue. La longueur d'onde la plus courte vaut donc  $2\Delta x/(1 - \Delta x/2)$  — qui tend vers  $2\Delta x$  si le nombre de mailles tend vers l'infini.



**Figure 4.** Rapport de l'amplitude des modes,  $\alpha_n^* / \alpha_n$ , pour  $K=5$  (+),  $K=10$  (o),  $K=50$  (x) et  $K=100$  (\*).

L'amplitude d'un mode de la solution discrète,

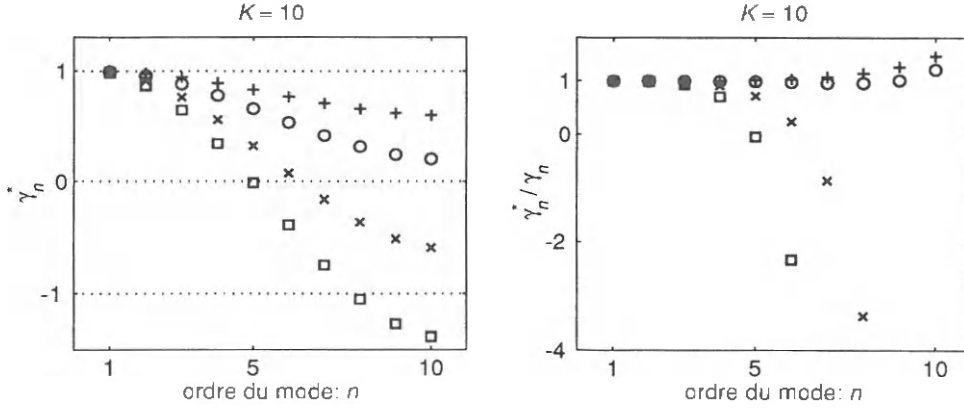
$$\alpha_n^* = \frac{1}{K \sin \frac{(n-1/2)\pi}{2K}} , \quad (32)$$

est toutefois asymptotique à  $\alpha_n$  si le nombre de mailles,  $K = 1/\Delta x$ , est élevé et si l'ordre relatif,  $n/K$ , du mode considéré reste petit. En effet, on a

<sup>8</sup> S'il est rare que l'on puisse trouver la solution analytique d'un problème différentiel aux dérivées partielles, il est vraiment exceptionnel que la solution analytique d'un problème discret associé soit disponible. C'est évidemment parce que ces deux solutions analytiques sont connues — et relativement simples — que l'on a choisi de traiter le présent problème de conduction de chaleur.

$$\alpha_n^* \equiv \alpha_n + \frac{(n-1/2)\pi}{12} \frac{1}{K^2}, \quad \text{pour } n/K \rightarrow 0 \text{ et } K \rightarrow \infty. \quad (33)$$

Ainsi, raffiner la discrétisation spatiale — en augmentant  $K$  — améliore la valeur de l'amplitude des modes d'ordre relativement peu élevé. Par contre, quoi qu'on fasse, l'amplitude des modes d'ordre proche de  $K$  reste relativement éloignée de celle des modes correspondants de la solution continue (Figure 4).



**Figure 5.** Représentation du facteur d'amplification des modes numériques,  $\gamma_n^*$  (panneau de gauche), et du rapport du facteur d'amplification des modes numériques au facteur d'amplification correspondant de la solution exacte,  $\gamma_n^*/\gamma_n$ , pour  $K=10$ , dans le cas où  $\Delta t/\Delta x^2=0.1$  (+),  $\Delta t/\Delta x^2=0.2$  (o),  $\Delta t/\Delta x^2=0.4$  (x) et  $\Delta t/\Delta x^2=0.6$  (carré). Les deux premiers cas correspondent à des modes numériques stables, le troisième représente des modes stables mais dont certains oscillent, tandis que le dernier est un exemple d'instabilité.

Pour étudier la stabilité des solutions, il est utile d'examiner le “facteur d'amplification”, qui est le facteur par lequel le  $n$ -ième mode de la solution est amplifié au cours d'un pas de temps. Ainsi, pour le  $n$ -ième mode de la solution continue et de la solution discrète, le facteur d'amplification est défini par

$$(\gamma_n, \gamma_n^*) = (e^{-\varepsilon_n \Delta t}, e^{-\varepsilon_n^* \Delta t}). \quad (34)$$

On voit sans peine que

$$0 < \gamma_n = e^{-(n-1/2)^2 \pi^2 \Delta t} < 1. \quad (35)$$

Donc, aucun mode de la solution exacte ne croît à mesure que le temps passe. Pour que le schéma numérique soit *stable*, il est indispensable que cette propriété soit vérifiée par tous les modes numériques. Par conséquent, on exige que

$$\gamma_n^* = 1 - \frac{2\Delta t}{\Delta x^2} [1 - \cos(\mu_n^* \Delta x)] \quad (36)$$

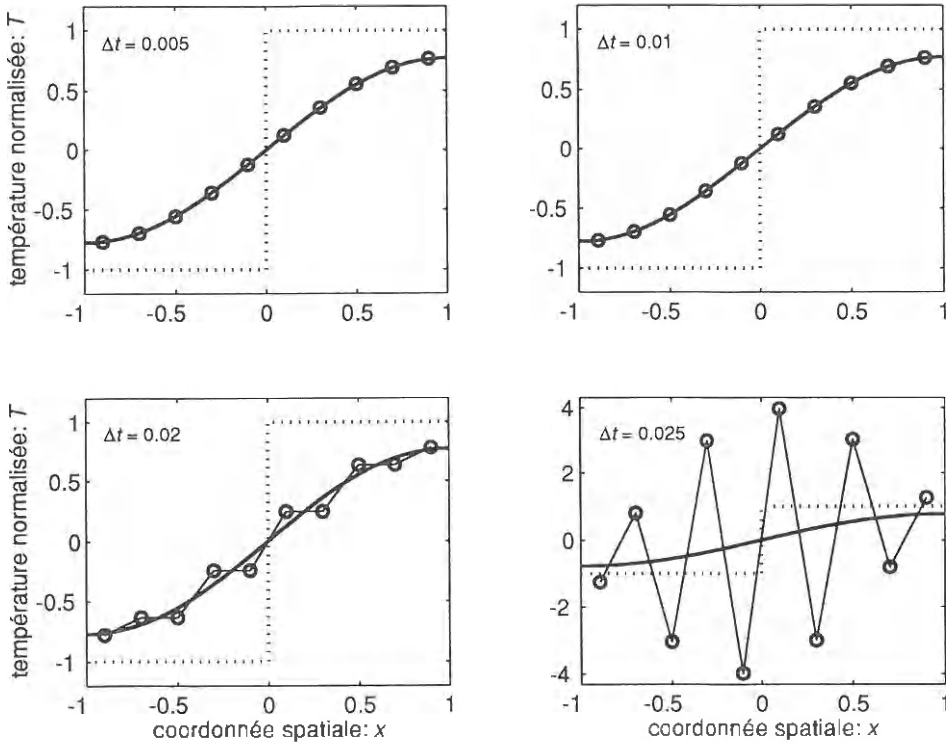
vérifie l'inégalité

$$|\gamma_n^*| \leq 1, \quad (37)$$

ce qui implique que le pas de temps et l'incrément spatial doivent être tels que

$$\frac{\Delta t}{\Delta x^2} \leq \frac{1}{1 - \cos \frac{(K-1/2)\pi}{K}}. \quad (38)$$

Ce résultat est essentiel. Il indique en effet que réduire le pas spatial  $\Delta x$  — dans l'espoir d'augmenter la précision de l'approximation numérique — sans réduire de façon adéquate le pas de temps peut très bien conduire à l'instabilité du schéma, c'est-à-dire détruire totalement sa précision. *On ne pouvait pas déduire ceci de la seule analyse de l'erreur de troncature!*



**Figure 6.** Comparaison de la température initiale (trait pointillé), de la température exacte (trait plein) en  $t=0.2$  et de son approximation numérique (o) en  $t=0.2$ , pour  $K=5$  et différentes valeurs du pas de temps  $\Delta t$ . Tous les schémas sont stables, sauf celui correspondant au panneau inférieur droit.

Le facteur d'amplification  $\gamma_n$  est strictement positif pour tous les modes, ce qui implique qu'aucun mode n'oscille dans le temps. Pour qu'il en soit également ainsi des modes numériques, il faut leur imposer une condition plus restrictive que la condition de stabilité (38), à savoir

$$\frac{\Delta t}{\Delta x^2} \leq \frac{1/2}{1 - \cos \frac{(K-1/2)\pi}{K}} . \quad (39)$$

La Figure 5 illustre le fait que le facteur d'amplification des modes numériques d'ordre élevé, c'est-à-dire proche de  $K$ , est, largement erroné quelle que soit la finesse relative de la discrétisation spatio-temporelle.

Le coefficient d'amortissement des modes discrets vaut

$$\varepsilon_n^* = -\frac{1}{\Delta t} \log \left\{ 1 - \frac{2\Delta t}{\Delta x^2} [1 - \cos(\mu_n^* \Delta x)] \right\} , \quad (40)$$

une expression qui est complexe si l'argument du logarithme est négatif. Sauf si le schéma est instable, tous les modes discrets sont amortis. Par conséquent, la différence entre la solution exacte et son approximation par un algorithme stable n'est jamais très grande en valeur absolue pour le problème qui nous occupe (Figure 6). De toute façon, on peut montrer que le schéma considéré ici est convergent dès lors que ses paramètres le rendent stable. La démonstration formelle de cette propriété sort du cadre de cette note, mais l'argumentation peut être ébauchée facilement: l'amplitude et le coefficient d'amortissement des modes discrets d'ordre relativement peu élevé sont d'autant plus corrects que les incréments spatiaux et temporels tendent vers zéro — en préservant la stabilité du schéma —, tandis que l'erreur sur les modes d'ordre élevé importe peu car ces modes sont très amortis.

### Autres schémas numériques

C'est l'existence des solutions analytiques continues et discrètes qui a rendu possible la discussion approfondie ci-dessus. La plupart des problèmes différentiels aux dérivées partielles sont plus difficiles de telle sorte que, généralement, aucune solution exacte n'est disponible. Il est cependant souhaitable de disposer d'un outil pour déterminer, au moins approximativement, certaines propriétés des schémas numériques — en particulier leur stabilité — avant même leur implémentation sur ordinateur. Cet outil est l'analyse de von Neumann, dont on peut comprendre les bases en se référant à la solution du problème résolu ci-dessus. Cette approche consiste à examiner le comportement de solutions de la forme "modes de Fourier" d'équations discrètes linéaires à coefficients constants<sup>9</sup>. Les longueurs d'onde des modes peuvent prendre toutes les valeurs supérieures à  $2\Delta x$ . Ainsi, en utilisant de nouveau des variables dimensionnelles, on introduit dans l'équation discrète (21) des solutions de la forme

$$T_{l,k}^* = (\gamma^*)^l e^{ikv} , \quad (41)$$

<sup>9</sup> Si le schéma à analyser n'est pas linéaire à coefficients constants, on le rend tel en appliquant les simplifications que l'on juge raisonnable.

où  $v = \mu \Delta x$ . Comme on considère que  $\mu \leq (2\pi)/(2\Delta x)$ , on a

$$0 \leq v \leq \pi . \quad (42)$$

Le facteur d'amplification des solutions de ce type vaut

$$\gamma^* = 1 - \frac{2\lambda \Delta t}{\Delta x^2} (1 - \cos v) . \quad (43)$$

Pour que le schéma soit stable, il faut  $|\gamma^*| < 1$ , pour toute valeur admissible de  $v$ . Donc, les incréments des variables indépendantes doivent vérifier l'inégalité

$$\frac{\lambda \Delta t}{\Delta x^2} \leq \frac{1}{2} , \quad (44)$$

qui est équivalente — en variables dimensionnelles — à (38) quand  $K \equiv \Delta x^{-1} \rightarrow \infty$ .

On va maintenant utiliser largement l'analyse de von Neumann pour appréhender les propriétés essentielles d'autres schémas de discrétisation de l'équation de la chaleur (15).

Dans (21b), la dérivée temporelle de la température — le second membre de cette relation — est évaluée à l'instant  $l \Delta t$ , c'est-à-dire l'instant le plus précoce concerné par l'équation discrète. Un tel schéma est dit *explicite*. Dans un schéma *implicite*, par contre, on évalue la dérivée temporelle à l'instant le plus tardif concerné par l'équation discrète: pour transformer (21) en un schéma implicite, il suffit d'évaluer la dérivée temporelle en  $(l+1) \Delta t$ . Cette discrétisation s'écrit

$$\frac{T_{l+1,k}^* - T_{l,k}^*}{\Delta t} = \lambda \frac{T_{l+1,k+1}^* - 2T_{l+1,k}^* + T_{l+1,k-1}^*}{\Delta x^2} . \quad (45)$$

Le facteur d'amplification de ce schéma,

$$\gamma^* = \frac{1}{1 + \frac{2\lambda \Delta t}{\Delta x^2} (1 - \cos v)} , \quad (46)$$

est toujours positif et inférieur ou égal à l'unité, quelles que soient les valeurs de la diffusivité thermique, de l'incrément temporel et du pas spatial. En d'autres termes, ce schéma est toujours stable. On dit que ce schéma est *inconditionnellement stable* — alors que le premier schéma était *conditionnellement stable*.

Le prix à payer pour bénéficier d'une propriété aussi avantageuse que la stabilité inconditionnelle est la nécessité de résoudre un système d'équations algébriques linéaires pour progresser d'un pas de temps. En ré-écrivant l'équation (45) sous la forme

$$-\frac{\lambda \Delta t}{\Delta x^2} T_{l+1,k-1}^* + \left(1 + 2 \frac{\lambda \Delta t}{\Delta x^2}\right) T_{l+1,k}^* - \frac{\lambda \Delta t}{\Delta x^2} T_{l+1,k+1}^* = T_{l,k}^* , \quad (47)$$

on voit immédiatement que la matrice du système est tridiagonale. C'est pourquoi il n'est pas nécessaire d'employer l'algorithme d'élimination de Gauss pour résoudre ce système. On peut se contenter de l'algorithme de Thomas, qui exige d'effectuer un nombre d'opérations arithmétiques proportionnel à  $K$ , le nombre de points de discrétisation, alors que l'élimination de Gauss demande un nombre d'opérations proportionnel à  $K^3$ . Le nombre d'opérations exigées par le schéma implicite étant du même ordre de grandeur que pour le schéma explicite,



on pourrait conclure qu'il est toujours souhaitable d'employer le schéma implicite si sa précision n'était pas assez modeste. En effet, cet algorithme est du premier ordre de précision selon le temps et du second ordre selon l'espace — c'est-à-dire qu'il n'est pas meilleur, de ce point de vue, que le schéma explicite.

Le schéma *leapfrog* — saute-mouton —,

$$\frac{T_{l+1,k}^* - T_{l-1,k}^*}{2\Delta t} = \lambda \frac{T_{l,k+1}^* - 2T_{l,k}^* + T_{l,k-1}^*}{\Delta x^2}, \quad (48)$$

estimant la dérivée temporelle de la température par une différence centrée, est du second ordre de précision dans le temps et l'espace. On peut donc penser qu'il est supérieur aux deux schémas envisagés plus haut. Malheureusement, il est *inconditionnellement instable*, et donc inutilisable! Voici un excellent exemple du peu de pertinence des informations fournies par l'erreur de troncature...

La solution la plus pratique pour augmenter la précision du schéma réside dans la construction de schémas *semi-implicites*. En utilisant l'interpolation linéaire

$$T_{l+\chi,k}^* \approx (1-\chi)T_{l,k}^* + \chi T_{l+1,k}^*, \quad \text{avec } 0 \leq \chi \leq 1, \quad (49)$$

le schéma semi-implicite résultant de la combinaison du schéma explicite (21) et du schéma implicite (45) s'écrit

$$\frac{T_{l+1,k}^* - T_{l,k}^*}{\Delta t} = \lambda \frac{T_{l+\chi,k+1}^* - 2T_{l+\chi,k}^* + T_{l+\chi,k-1}^*}{\Delta x^2}, \quad (50)$$

où  $\chi$  représente le niveau d'"implicité" du schéma. Bien entendu, si  $\chi$  vaut zéro ou l'unité, on retrouve le schéma explicite (21) ou le schéma implicite (45). L'erreur de troncature de l'algorithme (49) est du second ordre dans l'espace et du premier ordre dans le temps, sauf quand  $\chi = 1/2$ . Dans ce cas, l'erreur de troncature est du second ordre dans le temps. Le facteur d'amplification est

$$\gamma^* = \frac{1 - (1-\chi) \frac{2\lambda\Delta t}{\Delta x^2} (1 - \cos v)}{1 + \chi \frac{2\lambda\Delta t}{\Delta x^2} (1 - \cos v)}, \quad (51)$$

conduisant à la condition de stabilité

$$\frac{\lambda\Delta t}{\Delta x^2} \leq \frac{1}{2 \max(0, 1 - 2\chi)}. \quad (52)$$

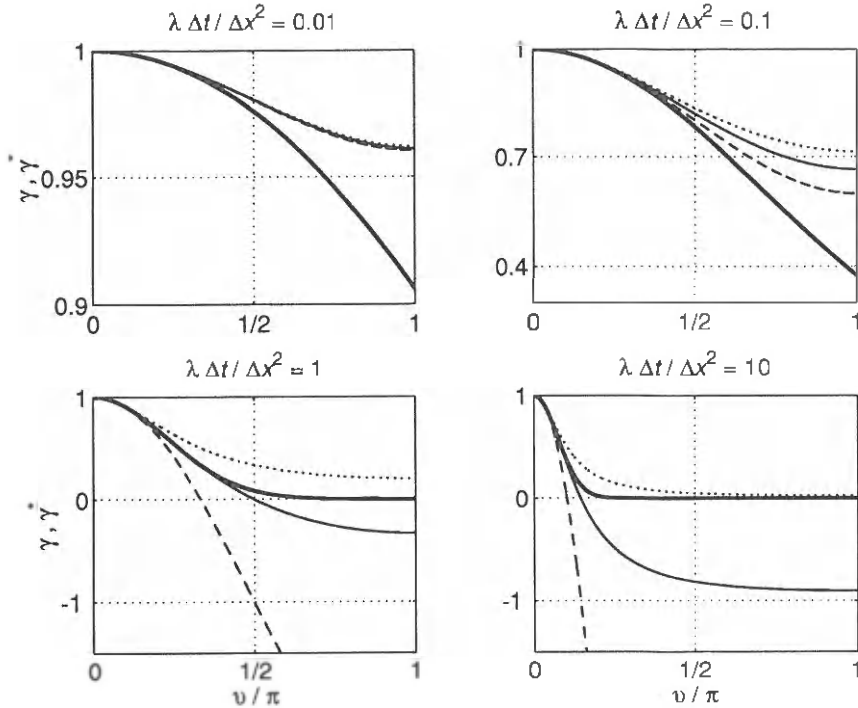
Ainsi, dès que  $\chi > 0$ , le schéma est "plus stable" que le schéma explicite — c'est-à-dire que le plus grand pas de temps admissible est plus grand que celui associé au schéma explicite. Le schéma est conditionnellement stable pour  $\chi < 1/2$ , et inconditionnellement stable sinon — c'est-à-dire quand la partie implicite "domine" la partie explicite.

Au delà de la stabilité, l'analyse de von Neumann permet de comparer le comportement du facteur d'amplification du schéma numérique,  $\gamma^*$ , avec celui de la solution exacte,  $\gamma$ . On obtient ce dernier en introduisant dans l'équation de la chaleur continue (15) une solution du type "mode de Fourier", c'est-à-dire

$$T(t, x) = e^{-\varepsilon t + i\mu x}, \quad (53)$$

ce qui fournit

$$\gamma \equiv \frac{T(t + \Delta t, x)}{T(t, x)} = e^{-\frac{\lambda \Delta t}{\Delta x^2} v^2}. \quad (54)$$



**Figure 7.** Comparaison du facteur d'amplification de la solution continue (trait plein épais),  $\gamma$ , avec celui de la solution discrète,  $\gamma^*$ , en fonction du nombre d'onde normalisé,  $v$ . On considère les schémas explicite (trait interrompu), implicite (pointillés) et semi-implicite (trait plein) pour  $\chi=1/2$ .

Dans la Figure 7, on voit que le coefficient d'amplification de la solution discrète est généralement très éloigné de celui de la solution continue pour les modes dont la longueur d'onde est courte, c'est-à-dire proche de  $2 \Delta x$ . Cette constatation est valable pour quasiment tous les schémas numériques, qu'ils visent ou non à résoudre l'équation de la chaleur. En d'autres termes, les phénomènes dont la longueur caractéristique est proche de la taille de la maille sont quasiment toujours mal représentés par un schéma numérique. Par contre, tous les algorithmes examinés dans la Figure 7 — explicite, implicite et semi-implicite — représentent très bien l'évolution des modes de grande longueur d'onde. Ceci n'est guère étonnant, car le

temps caractéristique d'amortissement de ces modes est nettement plus grand que l'incrément temporel,  $\Delta t$ , de sorte que les détails de la discrétisation temporelle importent assez peu pour ces modes.

Une autre approche de la résolution numérique de l'équation de la chaleur est celle des méthodes de *marches aléatoires*, selon laquelle on considère un grand nombre de particules ponctuelles transportant chacune une quantité d'énergie interne appropriée avec une vitesse telle que les déplacements de ces particules simulent la diffusion moléculaire de la chaleur. Si  $X(t)$  est la coordonnée spatiale d'une de ces particules à l'instant  $t$ , alors à l'instant  $t + \Delta t$  se trouvera à la position

$$X(t + \Delta t) = X(t) + (6\lambda\Delta t)^{1/2} r , \quad (55)$$

où  $r$  est un nombre aléatoire, uniformément distribué entre  $-1$  et  $+1$ . L'analyse de cette technique sort du cadre de la présente note.

### Equation de la chaleur multi-dimensionnelle

De nombreux processus diffusifs ne peuvent être valablement modélisés par une équation de diffusion à une dimension spatiale. On ne peut donc se limiter à l'étude d'équations et d'algorithmes uni-dimensionnels; il faut considérer des problèmes où l'inconnue dépend d'au moins deux coordonnées spatiales. Ainsi, il est souhaitable d'examiner une équation du type

$$\frac{\partial T}{\partial t} = \lambda_x \frac{\partial^2 T}{\partial x^2} + \lambda_y \frac{\partial^2 T}{\partial y^2} , \quad (56)$$

où  $\lambda_x$  et  $\lambda_y$  représentent les diffusivités associées aux directions des axes des coordonnées spatiales  $x$  et  $y$ . Si  $\Delta x$  et  $\Delta y$  symbolisent les pas spatiaux selon les axes  $x$  et  $y$ , alors le schéma numérique *explicite* le plus simple visant à la résolution de (56) est sans doute

$$\begin{aligned} \frac{T_{l+1,k_x,k_y}^* - T_{l,k_x,k_y}^*}{\Delta t} = \lambda_x \frac{T_{l,k_x+1,k_y}^* - 2T_{l,k_x,k_y}^* + T_{l,k_x-1,k_y}^*}{\Delta x^2} \\ + \lambda_y \frac{T_{l,k_x,k_y+1}^* - 2T_{l,k_x,k_y}^* + T_{l,k_x,k_y-1}^*}{\Delta y^2} , \end{aligned} \quad (57)$$

où les indices entiers  $k_x$  et  $k_y$  servent à repérer la position des points de discrétisation selon les directions des axes  $x$  et  $y$ . Par une méthode analogue à celle utilisée pour les problèmes uni-dimensionnels, on peut montrer que le facteur d'amplification de l'algorithme (57) vaut

$$\gamma^* = 1 - \frac{2\lambda_x \Delta t}{\Delta x^2} (1 - \cos v_x) - \frac{2\lambda_y \Delta t}{\Delta y^2} (1 - \cos v_y) , \quad (58)$$

de sorte que la condition de stabilité de ce schéma s'écrit

$$\frac{\lambda_x \Delta t}{\Delta x^2} + \frac{\lambda_y \Delta t}{\Delta y^2} \leq \frac{1}{2} . \quad (59)$$

Si les diffusivités et les pas spatiaux sont identiques dans les deux directions, le plus grand

incrément temporel admissible est égal à la moitié de celui découlant du critère (44), qui prévaut pour le schéma équivalent à une dimension. Cette relative perte de stabilité s'accroît encore si l'on considère trois dimensions spatiales. C'est pourquoi il est tentant d'implémenter un schéma semi-implicite ou implicite. Cependant, à deux dimensions, un tel algorithme exige la résolution d'un système d'équations algébriques linéaires dont la matrice est pentadiagonale, et dont deux bandes sont largement distantes de la diagonale principale. Comme une telle opération peut demander une quantité prohibitive d'opérations arithmétiques, il peut être approprié d'envisager un algorithme d'un autre type, à savoir la technique dite *implicite aux directions alternées*. Cette dernière consiste à progresser selon deux "demi-pas de temps" avec un algorithme implicite dans une direction et explicite dans l'autre, la direction implicite changeant d'un demi-pas de temps à l'autre:

$$\frac{T_{l+1/2,k_x,k_y}^* - T_{l,k_x,k_y}^*}{\Delta t/2} = \lambda_x \frac{T_{l+1/2,k_x+1,k_y}^* - 2T_{l+1/2,k_x,k_y}^* + T_{l+1/2,k_x-1,k_y}^*}{\Delta x^2} + \lambda_y \frac{T_{l,k_x,k_y+1}^* - 2T_{l,k_x,k_y}^* + T_{l,k_x,k_y-1}^*}{\Delta y^2}, \quad (60a)$$

$$\frac{T_{l+1,k_x,k_y}^* - T_{l+1/2,k_x,k_y}^*}{\Delta t/2} = \lambda_x \frac{T_{l+1/2,k_x+1,k_y}^* - 2T_{l+1/2,k_x,k_y}^* + T_{l+1/2,k_x-1,k_y}^*}{\Delta x^2} + \lambda_y \frac{T_{l+1,k_x,k_y+1}^* - 2T_{l+1,k_x,k_y}^* + T_{l+1,k_x,k_y-1}^*}{\Delta y^2}. \quad (60b)$$

Si le pas de temps est relativement grand, un des demi-pas de temps peut être instable. Mais, en tout cas, le schéma complet reste inconditionnellement stable. En effet, son facteur d'amplification est

$$\gamma^* = \frac{1 - \frac{\lambda_x \Delta t}{\Delta x^2} (1 - \cos v_x)}{1 + \frac{\lambda_x \Delta t}{\Delta x^2} (1 - \cos v_x)} \frac{1 - \frac{\lambda_y \Delta t}{\Delta y^2} (1 - \cos v_y)}{1 + \frac{\lambda_y \Delta t}{\Delta y^2} (1 - \cos v_y)}, \quad (61)$$

une quantité dont la valeur absolue est toujours inférieure ou égale à l'unité.

### Equation d'advection-diffusion

Comme l'illustre la Figure 6, la précision des schémas numériques visant à fournir une solution approchée du problème de conduction considéré plus haut est généralement bonne, sauf si le schéma est instable. Grossièrement parlant, ici, il suffit quasiment d'assurer la stabilité du schéma pour obtenir une approximation numérique relativement précise. Ceci est dû principalement au fait que l'amortissement des modes de Fourier de la solution est très fort. Donc, il importe finalement assez peu que l'amortissement des modes discrets soit ou ne soit pas correct — pour autant qu'ils restent stables. En fait, ce problème différentiel aux dérivées

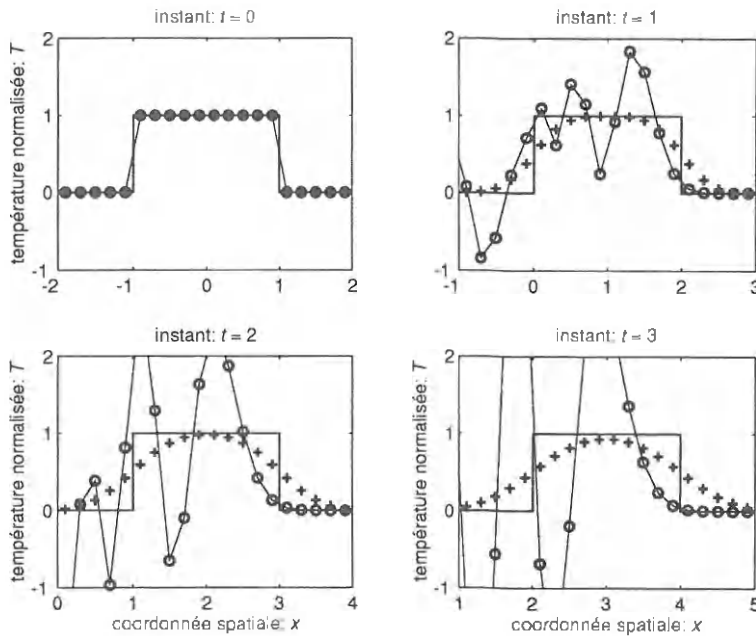
partielles est sans doute un des plus simples à résoudre numériquement. Cette simplicité cesse d'être de mise dès que l'on considère des processus physiques qui ne conduisent pas à des modes amortis. Ainsi, si l'on passe d'un milieu solide à un milieu fluide, il convient d'ajouter à l'équation de la chaleur (15) le transport convectif — ou "advection". On obtient alors l'équation

$$\frac{\partial T}{\partial t} = \frac{\partial}{\partial x} \left( -uT + \lambda \frac{\partial T}{\partial x} \right), \quad (62)$$

où  $u$  symbolise la vitesse du fluide. Si cette dernière est une constante positive, on peut adimensionnaliser (62) de la façon suivante:

$$\frac{\partial T}{\partial t} = -\frac{\partial T}{\partial x} + \frac{1}{Pe} \frac{\partial^2 T}{\partial x^2}, \quad (63)$$

où  $Pe$  est un coefficient adimensionnel généralement appelé "nombre de Péclet". Ce dernier mesure l'importance du transport convectif par rapport au transport diffusif — et a une signification équivalente à celle du nombre de Reynolds dans les équations de bilan de quantité de mouvement des fluides.



**Figure 8.** Solution exacte (trait plein) et solutions discrètes fournies par l'algorithme (68) (o) et (69) (+) du problème d'advection uni-dimensionnelle (64)-(67) à différents instants. L'instabilité du schéma centré (68) apparaît clairement, tout comme la dissipation erronée de (69).

Si le transport advectif domine largement le transport diffusif, obtenir une approximation numérique précise de l'équation (63) est très difficile, surtout si l'on traite une version multidimensionnelle de cette équation — c'est-à-dire si plus d'une variable indépendante spatiale doit être retenue. Des milliers d'articles ont été consacrés à ce problème! On est alors confronté à un problème où il ne suffit plus — loin s'en faut! — d'assurer la stabilité de l'algorithme. On va l'illustrer en examinant brièvement deux cas limites de l'équation d'advection-diffusion (63).

Quand le nombre de Péclet est très grand, il est tentant de négliger le terme diffusif. Dans un domaine infini, effectuer cette simplification ne provoque généralement pas de grandes erreurs. Ce faisant, on obtient l'équation hyperbolique

$$\frac{\partial T}{\partial t} = - \frac{\partial T}{\partial x}, \quad (64)$$

dont on cherche la solution dans le domaine  $-\infty < x < +\infty$ . Cette dernière est bien connue:

$$T(t, x) = T(t=0, x-t). \quad (65)$$

Le profil initial de température est donc transporté, sans déformation, avec une vitesse adimensionnelle unitaire. Ici, on suppose que la température initiale est le "trait Morse" défini par

$$T(t=0, x) = 1, \quad \text{pour } |x| < 1, \quad (66a)$$

$$T(t=0, x) = 0, \quad \text{pour } 1 < |x|. \quad (66b)$$

La solution de ce problème est alors

$$T(t, x) = 1, \quad \text{pour } |x-t| < 1, \quad (67a)$$

$$T(t=0, x) = 0, \quad \text{pour } 1 < |x-t|. \quad (67b)$$

On va maintenant bâtir deux approximations discrètes de la solution de (64), valables aux instants  $t_l = l\Delta t$  ( $l = 0, 1, 2, 3, \dots$ ) et aux points  $x_k = (k - 1/2)\Delta x$  ( $k = 0, \pm 1, \pm 2, \pm 3, \dots$ ). En utilisant des techniques semblables à celles employées pour le problème de diffusion de chaleur, on peut établir les schémas

$$\frac{T_{l+1,k}^* - T_{l,k}^*}{\Delta t} = - \frac{T_{l,k+1}^* - T_{l,k-1}^*}{2\Delta x}, \quad (68)$$

$$\frac{T_{l+1,k}^* - T_{l,k}^*}{\Delta t} = - \frac{T_{l,k}^* - T_{l,k-1}^*}{\Delta x}, \quad (69)$$

qui diffèrent par la façon dont est discrétisée la dérivée spatiale. Le premier schéma contient une approximation centrée de la dérivée spatiale, qui est précise au second ordre en  $\Delta x$ . Malheureusement, ce schéma est inconditionnellement instable, c'est-à-dire qu'il est instable quelles que soient les valeurs de l'incrément temporel et du pas spatial (Figure 8). Le second schéma, basé sur une approximation décentrée — ou *upwind* — de la dérivée spatiale n'est précis qu'au premier ordre en  $\Delta x$ , mais est stable à condition que le nombre de Courant,  $\Delta t/\Delta x$ , soit inférieur à l'unité. L'inconvénient majeur de ce schéma réside dans le fort taux de

dissipation numérique qu'il contient, ce qui constitue une erreur significative<sup>10</sup> car la solution exacte n'est nullement dissipée ou amortie (Figure 8). Pour obtenir une approximation discrète acceptable du problème d'advection ci-dessus, il faut avoir recours à des méthodes relativement sophistiquées — dont la discussion sort du cadre de ces notes.

On aborde maintenant le second cas limite de l'équation d'advection-diffusion (63), à savoir la situation *stationnaire* — c'est-à-dire indépendante du temps. L'équation à résoudre est alors

$$-\frac{\partial T}{\partial x} + \frac{1}{Pe} \frac{\partial^2 T}{\partial x^2} = 0. \quad (70)$$

Le domaine d'intérêt est  $-1 \leq x \leq 1$  et les conditions aux limites sont

$$T(x=0) = 1, \quad (71a)$$

$$T(x=1) = 0. \quad (71b)$$

La solution du problème (70)-(71) vaut

$$T(x) = \frac{1 - e^{-Pe(1-x)}}{1 - e^{-Pe}}. \quad (72)$$

Si le nombre de Péclet  $Pe$  est petit, alors le second terme de (70) tend à dominer l'équation, qui tend à se réduire à  $\partial^2 T / \partial x^2 = 0$ , dont la solution est  $T(x) = 1 - x$ . Par contre, si  $Pe$  est grand, le premier terme de (70) tend à dominer, de telle sorte que l'équation est à peu près  $\partial T / \partial x = 0$ . La solution de cette dernière est  $T(x) = 1$ . A proximité de la frontière  $x = 1$ , cette solution n'est évidemment pas compatible avec la condition (71b) imposée à cette extrémité du domaine d'intérêt. La solution s'ajuste alors vers la valeur prescrite à la frontière dans une région de fort gradient, appelée *couche limite*<sup>11</sup>, dont l'épaisseur est de l'ordre de  $Pe^{-1}$ . La Figure 9 illustre clairement cette discussion.

Le problème discret composé de l'équation

$$-\frac{T_{k+1} - T_{k-1}}{2\Delta x} + \frac{1}{Pe} \frac{T_{k+1} - 2T_k + T_{k-1}}{\Delta x^2} = 0 \quad (73)$$

où  $T_k \equiv T(x=k\Delta x)$  — avec  $k = 0, 1, 2, \dots, K = \Delta x^{-1}$  — et des conditions aux limites

$$T_0 = 1, \quad (74a)$$

$$T_K = 0, \quad (74b)$$

est compatible avec le problème différentiel (70)-(71). Il se fait que ce problème discret admet, pour  $Pe \neq 2$ , une solution analytique simple:

<sup>10</sup> Le schéma upwind du premier ordre fournit la solution exacte si  $\Delta t / \Delta x = 1$ . Ceci est cependant sans grand intérêt, car le nombre de Courant doit être inférieur à l'unité dès que la diffusion est prise en compte. En outre, à plusieurs dimensions spatiales — c'est-à-dire pour les problèmes réellement intéressants ou utiles —, il n'existe aucune combinaison des incréments des variables indépendantes qui soit susceptible de conduire à la solution exacte — sauf dans des cas très particuliers.

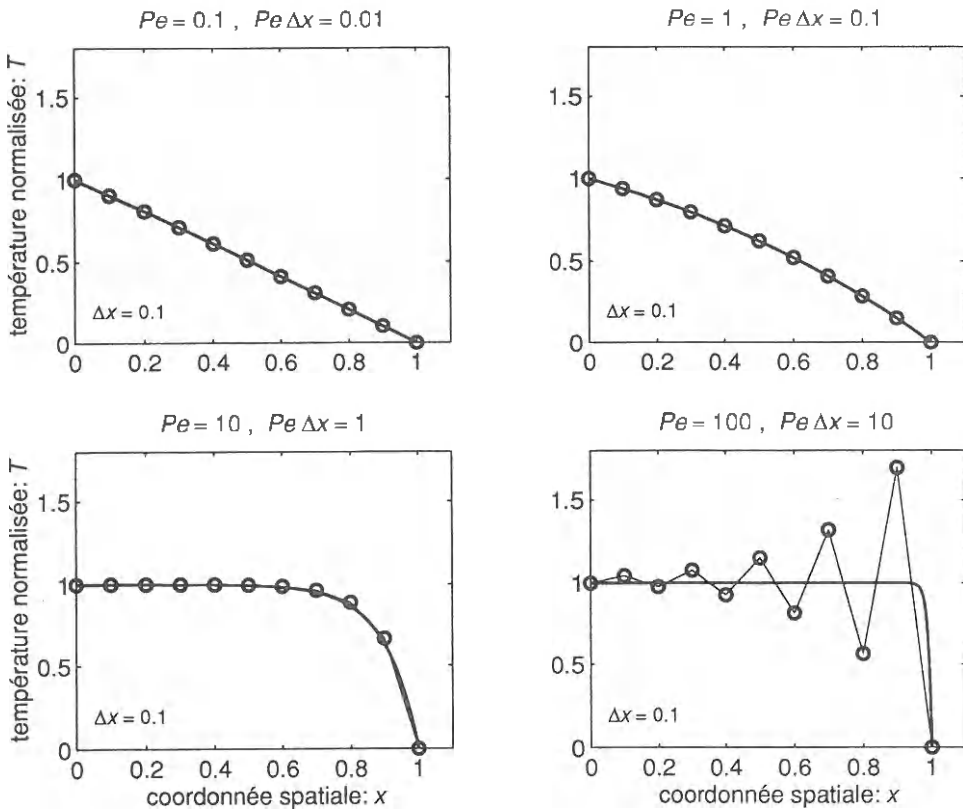
<sup>11</sup> La notion de "couche limite" fut sans doute développée pour la première fois en mécanique des fluides, mais peut maintenant être considérée comme un concept mathématique associé aux problèmes de perturbation singulière — que l'on rencontre dans de nombreux modèles mathématiques totalement indépendants de la mécanique des fluides.

$$T_k = \frac{1 - \left( \frac{2 - Pe \Delta x}{2 + Pe \Delta x} \right)^{K-k}}{1 - \left( \frac{2 - Pe \Delta x}{2 + Pe \Delta x} \right)^K}, \quad k = 0, 1, 2, \dots, K. \quad (75)$$

Cette expression présente un comportement oscillatoire si  $Pe \Delta x > 2$ . Comme la solution exacte n'oscille pas, un tel comportement est difficilement admissible. Il est donc souhaitable d'imposer que l'incrément spatial  $\Delta x$  satisfasse la condition

$$\Delta x < \frac{2}{Pe}. \quad (76)$$

Comme la limite supérieure admissible — selon (76) — de l'incrément spatial est du même ordre de grandeur que l'épaisseur de la couche limite, la condition (76) signifie que le pas spatial ne peut être significativement plus grand que l'épaisseur de la couche limite. En d'autres termes, la couche limite doit être *résolue* par la discrétisation choisie.



**Figure 9.** Solution exacte (trait plein) et solution discrète approchée (o) du problème (70)-(71) pour différentes valeurs du nombre de Péclet  $Pe$ .



## Conclusion

Les exemples ci-dessus illustrent les idées suivantes:

— La conception d'une méthode de résolution numérique d'équations différentielles aux dérivées partielles ne s'improvise pas. Il faut maîtriser un certain nombre de concepts de base, comme la compatibilité, la stabilité, la convergence. Il faut également être en mesure de choisir à bon escient la famille du schéma à mettre en oeuvre, à savoir le caractère explicite, implicite, semi-implicite, etc.

— D'autre part, pour traiter des problèmes réellement intéressants, c'est-à-dire des problèmes beaucoup plus complexes que les exemples académiques abordés ici, il faut s'être forgé une expérience suffisante. En effet, pour ces problèmes, aucun résultat théorique rigoureux n'est généralement disponible quant aux propriétés des schémas de résolution concevables. Par conséquent, seule l'expérience — assistée d'un bagage théorique approprié — permet de faire les "bons choix" de modélisation et de discrétisation.

— La simulation numérique d'un système naturel ou artificiel demande la maîtrise simultanée de la "physique" du problème, des méthodes numériques appropriées et de l'informatique nécessaire à la mise en oeuvre sur ordinateur du schéma numérique choisi. En outre, il faut être capable de réaliser la symbiose entre ces trois domaines,

*Il est donc illusoire de penser que la mise en oeuvre d'une méthode numérique est une activité de nature "technique", que l'on pourra apprendre au pied levé, le jour où le besoin s'en fera sentir. C'est pourquoi il est indispensable que le programme de la licence en physique de nos universités comporte une initiation aux méthodes de simulation numérique de la physique, avec une priorité sensible pour la résolution de problèmes différentiels aux dérivées partielles. Il faut enseigner les notions théoriques de base — les exemples ci-dessus montrent d'ailleurs qu'il s'agit d'une charge d'enseignement plutôt modérée — et, surtout, exiger que les étudiants se forment par la pratique, c'est-à-dire en résolvant numériquement, et de façon complète, des problèmes aux dérivées partielles.*

**Remerciements.** Michel Crucifix a accepté de relire et commenter la version préliminaire de ce texte. L'auteur a eu des discussions fructueuses à propos du sujet abordé ici avec Jean-Marc Gérard, Jan Govaerts et Bernard Piraux. L'auteur est Chercheur qualifié au Fonds National de la Recherche Scientifique de Belgique.

## Références

- Bender C. M. and S. A. Orszag, 1978, *Advanced Mathematical Methods for Scientists and Engineers*, McGraw-Hill.
- Ferziger J. H., 1998 (2nd ed.), *Numerical Methods for Engineering Application*, Wiley.
- Hack J. J., 1992, Climate system simulation: basic numerical & computational concepts, in: *Climate System Modeling*, K. E. Trenberth (Ed.), Cambridge University Press, pp. 283-318.

- Hirsch C., 1988, *Numerical Computation of Internal and External Flows - Volume 1: Fundamentals of Numerical Discretization*, Wiley.
- Richtmeyer R. D. and K. W. Morton, 1957, *Difference Methods for Initial-Value Problems*, Wiley.
- Robinson A. R., 1987, Predicting open ocean currents, fronts and eddies, in: *Three-Dimensional Models of Marine and Estuarine Dynamics*, J. C. J. Nihoul and B. M. Jamart (Eds.), Elsevier, pp. 89-111.