

## 1.4 Calcul numérique par différences finies

### 1.4.1 Principes de la méthode

A part dans quelques cas très particuliers, il est impossible de calculer explicitement des solutions des différents modèles présentés ci-dessus. Il est donc nécessaire d'avoir recours au calcul numérique sur ordinateur pour estimer qualitativement et quantitativement ces solutions. Le principe de toutes les méthodes de résolution numérique des équations aux dérivées partielles est d'obtenir des valeurs numériques discrètes (c'est-à-dire en nombre fini) qui **“approchent”** (en un sens convenable à préciser) la solution exacte. Dans ce procédé il faut bien être conscient de deux points fondamentaux : premièrement, on ne calcule pas des solutions exactes mais approchées ; deuxièmement, on **discrétise** le problème en représentant des fonctions par un nombre fini de valeurs, c'est-à-dire que **l'on passe du “continu” au “discret”**.

Il existe de nombreuses méthodes d'approximation numérique des solutions d'équations aux dérivées partielles. Nous présentons maintenant une des plus anciennes et des plus simples, appelée méthode des différences finies (nous verrons plus loin une autre méthode, dite des éléments finis). Pour simplifier la présentation, nous nous limitons à la dimension un d'espace (voir la Sous-section 2.2.6 pour les dimensions supérieures). Nous n'abordons pour l'instant que les principes pratiques de cette méthode, c'est-à-dire la construction de ce qu'on appelle des **schémas numériques**. Nous réservons pour le Chapitre 2 la justification théorique de ces schémas, c'est-à-dire l'étude de leur convergence (en quel sens les solutions approchées discrètes sont proches des solutions exactes continues).

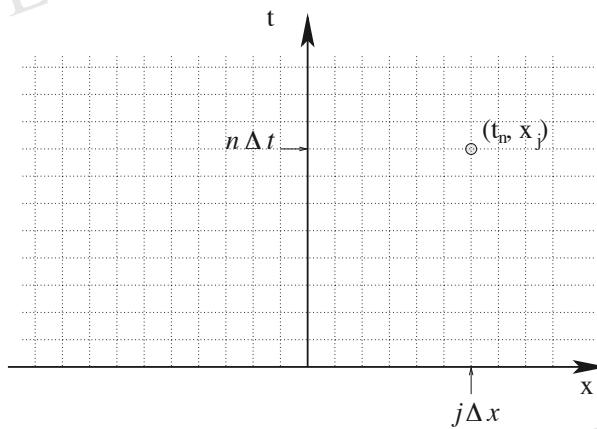


FIG. 1.4 – Maillage en différences finies.

Pour discrétiser le continuum spatio-temporel, on introduit un **pas d'espace**  $\Delta x > 0$  et un **pas de temps**  $\Delta t > 0$  qui seront les plus petites échelles représentées par la méthode numérique. On définit un maillage ou des coordonnées discrètes de

l'espace et du temps (voir la Figure 1.4)

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

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 (inconnue). Le principe de la méthode des différences finies est de remplacer les dérivées par des différences finies en utilisant des formules de Taylor dans lesquelles on néglige les restes. Par exemple, on approche la dérivée seconde en espace (le Laplacien en dimension un) par

$$-\frac{\partial^2 u}{\partial x^2}(t_n, x_j) \approx \frac{-u_{j-1}^n + 2u_j^n - u_{j+1}^n}{(\Delta x)^2} \quad (1.25)$$

où l'on reconnaît la formule de Taylor

$$\begin{aligned} -u(t, x - \Delta x) + 2u(t, x) - u(t, x + \Delta x) &= -(\Delta x)^2 \frac{\partial^2 u}{\partial x^2}(t, x) \\ &\quad - \frac{(\Delta x)^4}{12} \frac{\partial^4 u}{\partial x^4}(t, x) + \mathcal{O}((\Delta x)^6) \end{aligned} \quad (1.26)$$

Si  $\Delta x$  est “petit”, la formule (1.25) est une “bonne” approximation (elle est naturelle mais pas unique). La formule (1.25) est dite **centrée** car elle est symétrique en  $j$ .

Pour discréteriser l'équation de convection-diffusion

$$\frac{\partial u}{\partial t} + V \frac{\partial u}{\partial x} - \nu \frac{\partial^2 u}{\partial x^2} = 0 \quad (1.27)$$

il faut aussi discréteriser le terme de convection. Une formule centrée donne

$$V \frac{\partial u}{\partial x}(t_n, x_j) \approx V \frac{u_{j+1}^n - u_{j-1}^n}{2\Delta x}$$

Il ne reste plus qu'à faire la même chose pour la dérivée en temps. On a encore le choix dans la formule de différences finies : soit centrée, soit décentrée. Examinons trois formules “naturelles”.

1. En premier lieu, la différence finie centrée

$$\frac{\partial u}{\partial t}(t_n, x_j) \approx \frac{u_j^{n+1} - u_j^{n-1}}{2\Delta t}$$

conduit au schéma complètement symétrique par rapport à  $n$  et  $j$  (appelé schéma centré ou **schéma de Richardson**)

$$\frac{u_j^{n+1} - u_j^{n-1}}{2\Delta t} + V \frac{u_{j+1}^n - u_{j-1}^n}{2\Delta x} + \nu \frac{-u_{j-1}^n + 2u_j^n - u_{j+1}^n}{(\Delta x)^2} = 0. \quad (1.28)$$

Aussi “naturel” et évident soit-il, ce schéma est incapable de calculer des solutions approchées de l'équation de convection-diffusion (1.27) (voir

l'exemple numérique de la Figure 1.5)! Nous justifierons cette incapacité du schéma à approcher la solution exacte dans le Lemme 2.2.23. Pour l'instant, indiquons simplement que la difficulté provient du caractère centré de la différence finie qui approche la dérivée en temps.

2. Un deuxième choix est la la différence finie décentrée amont (on remonte le temps ; on parle aussi de **schéma d'Euler rétrograde**)

$$\frac{\partial u}{\partial t}(t_n, x_j) \approx \frac{u_j^n - u_j^{n-1}}{\Delta t}$$

qui conduit au schéma

$$\frac{u_j^n - u_j^{n-1}}{\Delta t} + V \frac{u_{j+1}^n - u_{j-1}^n}{2\Delta x} + \nu \frac{-u_{j-1}^n + 2u_j^n - u_{j+1}^n}{(\Delta x)^2} = 0. \quad (1.29)$$

3. Le troisième choix est le symétrique du précédent : la différence finie décentrée aval (on avance dans le temps ; on parle aussi de **schéma d'Euler progressif**)

$$\frac{\partial u}{\partial t}(t_n, x_j) \approx \frac{u_j^{n+1} - u_j^n}{\Delta t}$$

conduit au schéma

$$\frac{u_j^{n+1} - u_j^n}{\Delta t} + V \frac{u_{j+1}^n - u_{j-1}^n}{2\Delta x} + \nu \frac{-u_{j-1}^n + 2u_j^n - u_{j+1}^n}{(\Delta x)^2} = 0. \quad (1.30)$$

La différence principale entre ces deux derniers schémas est que (1.29) est dit **implicite** car il faut résoudre un système d'équations linéaires pour calculer les valeurs  $(u_j^n)_{j \in \mathbb{Z}}$  en fonctions des valeurs précédentes  $(u_j^{n-1})_{j \in \mathbb{Z}}$ , tandis que (1.30) est dit **explicite** puisqu'il donne immédiatement les valeurs  $(u_j^{n+1})_{j \in \mathbb{Z}}$  en fonction des  $(u_j^n)_{j \in \mathbb{Z}}$ . Le décalage de 1 sur l'indice  $n$  entre les schémas (1.29) et (1.30) n'est évidemment qu'apparent puisqu'on peut réécrire de manière équivalente (1.30) sous la forme

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

Dans les trois schémas que nous venons de définir, il y a bien sûr une donnée initiale pour démarrer les itérations en  $n$  : les valeurs initiales  $(u_j^0)_{j \in \mathbb{Z}}$  sont définies, par exemple, par  $u_j^0 = u_0(j\Delta x)$  où  $u_0$  est la donnée initiale de l'équation de convection-diffusion (1.27). Remarquons que pour le “mauvais” schéma centré (1.28) il y a une difficulté supplémentaire au démarrage : pour  $n = 1$  on a aussi besoin de connaître les valeurs  $(u_j^1)_{j \in \mathbb{Z}}$  qu'il faut donc calculer autrement (par exemple, par application d'un des deux autres schémas).



FIG. 1.5 – Schéma centré instable avec  $\nu\Delta t = 0.1(\Delta x)^2$ .

#### 1.4.2 Résultats numériques pour l'équation de la chaleur

Commençons par faire quelques tests numériques très simples dans le cas où  $V = 0$  et  $\nu = 1$ , c'est-à-dire que **l'on résout numériquement l'équation de la chaleur**. On choisit comme condition initiale la fonction

$$u_0(x) = \max(1 - x^2, 0).$$

Pour pouvoir comparer les solutions numériques approchées avec la solution exacte (1.14), nous voudrions travailler sur le domaine infini  $\Omega = \mathbb{R}$ , c'est-à-dire calculer, pour chaque  $n \geq 0$ , une infinité de valeurs  $(u_j^n)_{j \in \mathbb{Z}}$ , mais l'ordinateur ne le permet pas car sa mémoire est finie ! En première approximation, nous remplaçons donc  $\mathbb{R}$  par le “grand” domaine  $\Omega = (-10, +10)$  muni de conditions aux limites de Dirichlet. Nous admettrons la validité de cette approximation (qui est confirmée par les comparaisons numériques ci-dessous). Nous fixons le pas d'espace à  $\Delta x = 0.05$  : il y a donc 401 valeurs  $(u_j^n)_{-200 \leq j \leq +200}$  à calculer. Rappelons pour mémoire que les valeurs  $u_j^n$  calculées par l'ordinateur sont entachées d'erreurs d'arrondi et ne sont donc pas les valeurs exactes des schémas discrets : néanmoins, dans les calculs présentés ici, ces erreurs d'arrondi sont totalement négligeables et ne sont en aucune manière la cause des différent phénomènes que nous allons observer. Sur toutes les figures nous représentons la solution exacte, calculée avec la formule explicite (1.14), et la solution approchée numérique considérée.

Réglons tout de suite le sort du schéma centré (1.28) : comme nous l'avions annoncé, ce schéma est incapable de calculer des solutions approchées de l'équation de

la chaleur. Quel que soit le choix du pas de temps  $\Delta t$ , ce schéma est **instable**, c'est-à-dire que la solution numérique oscille de manière non bornée si l'on diminue les valeurs des pas  $\Delta x$  et  $\Delta t$ . Ce phénomène très caractéristique (et d'apparition très rapide) est illustré par la Figure 1.5. Insistons sur le fait que **quel que soit le choix** des pas  $\Delta t$  et  $\Delta x$ , on observe ces oscillations (non physiques, bien sûr). On dit que le schéma est inconditionnellement instable. Une justification rigoureuse en sera donnée au chapitre suivant (voir le Lemme 2.2.23).



FIG. 1.6 – Schéma implicite avec  $\nu\Delta t = 2(\Delta x)^2$ .

A l'opposé du précédent schéma, le schéma implicite (1.29) calcule de “bonnes” solutions approchées de l’équation de la chaleur **quel que soit** le pas de temps  $\Delta t$  (voir la Figure 1.6). En particulier, on n’observe jamais d’oscillations numériques quel que soit le choix des pas  $\Delta t$  et  $\Delta x$ . On dit que le schéma implicite est inconditionnellement stable.

Considérons maintenant le schéma explicite (1.30) : des expériences numériques montrent facilement que selon les valeurs du pas de temps  $\Delta t$  des oscillations numériques apparaissent ou non (voir la Figure 1.7). La limite de stabilité est facile à trouver expérimentalement : quel que soit le choix des pas  $\Delta t$  et  $\Delta x$  qui vérifient la condition

$$2\nu\Delta t \leq (\Delta x)^2 \quad (1.31)$$

le schéma est stable, tandis que si (1.31) n’est pas vérifiée, alors le schéma est instable. On dit que le schéma explicite est conditionnellement stable. La condition de stabilité (1.31) est **une des remarques les plus simples et les plus profondes de l’analyse numérique**. Elle fut découverte en 1928 (avant l’apparition des pre-

miers ordinateurs !) par Courant, Friedrichs, et Lewy. Elle porte depuis le nom de **condition CFL ou condition de Courant, Friedrichs, et Lewy**.

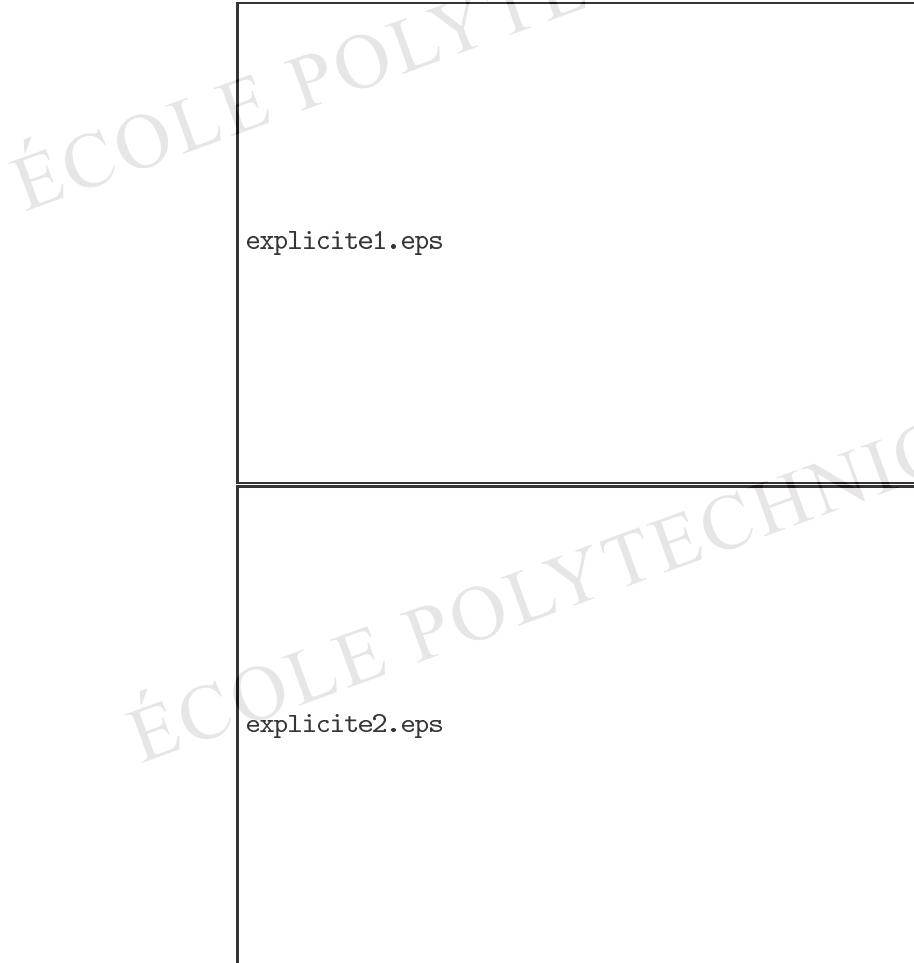


FIG. 1.7 – Schéma explicite avec  $\nu\Delta t = 0.4(\Delta x)^2$  (haut) et  $\nu\Delta t = 0.51(\Delta x)^2$  (bas).

Nous allons justifier brièvement cette condition de stabilité (une analyse plus poussée sera effectuée au prochain chapitre). Réécrivons le schéma explicite sous la forme

$$u_j^{n+1} = \frac{\nu\Delta t}{(\Delta x)^2} u_{j-1}^n + \left(1 - 2\frac{\nu\Delta t}{(\Delta x)^2}\right) u_j^n + \frac{\nu\Delta t}{(\Delta x)^2} u_{j+1}^n. \quad (1.32)$$

Si la condition CFL est vérifiée, alors (1.32) montre que  $u_j^{n+1}$  est une combinaison convexe des valeurs au temps précédent  $u_{j-1}^n, u_j^n, u_{j+1}^n$  (tous les coefficients dans le

membre de droite de (1.32) sont positifs et leur somme vaut 1). En particulier, si la donnée initiale  $u_0$  est bornée par deux constantes  $m$  et  $M$  telles que

$$m \leq u_j^0 \leq M \text{ pour tout } j \in \mathbb{Z},$$

alors une récurrence facile montre que les mêmes inégalités restent vraies pour tous les temps ultérieurs

$$m \leq u_j^n \leq M \text{ pour tout } j \in \mathbb{Z} \text{ et pour tout } n \geq 0. \quad (1.33)$$

La propriété (1.33) empêche le schéma d'osciller de manière non bornée : il est donc stable sous la condition CFL. La propriété (1.33) est appelée principe du maximum discret : il s'agit de l'équivalent **discret** du principe du maximum **continu** pour les solutions exactes que nous avons vu à la Remarque 1.2.10.

Supposons au contraire que la condition CFL ne soit pas vérifiée, c'est-à-dire que

$$2\nu\Delta t > (\Delta x)^2.$$

Alors, pour certaines données initiales le schéma est instable (il peut être stable pour certaines données initiales “exceptionnelles” : par exemple, si  $u_0 \equiv 0$ !). Prenons la donnée initiale définie par

$$u_j^0 = (-1)^j$$

qui est bien uniformément bornée. Un calcul simple montre que

$$u_j^n = (-1)^j \left(1 - 4 \frac{\nu\Delta t}{(\Delta x)^2}\right)^n$$

qui croît en module vers l'infini lorsque  $n$  tend vers l'infini car  $1 - 4 \frac{\nu\Delta t}{(\Delta x)^2} < -1$ . Le schéma explicite est donc instable si la condition CFL n'est pas satisfaite.

**Exercice 1.4.1** Le but de cet exercice est de montrer que le schéma implicite (1.29), avec  $V = 0$ , vérifie aussi le principe du maximum discret. On impose des conditions aux limites de Dirichlet, c'est-à-dire que la formule (1.29) est valable pour  $1 \leq j \leq J$  et on fixe  $u_0^n = u_{J+1}^n = 0$  pour tout  $n \in \mathbb{N}$ . Soit deux constantes  $m \leq 0 \leq M$  telles que  $m \leq u_j^0 \leq M$  pour  $1 \leq j \leq J$ . Vérifier que l'on peut bien calculer de manière unique les  $u_j^{n+1}$  en fonction des  $u_j^n$ . Montrer que pour tous les temps  $n \geq 0$  on a encore les inégalités  $m \leq u_j^n \leq M$  pour  $1 \leq j \leq J$  (et ceci sans condition sur  $\Delta t$  et  $\Delta x$ ).

Si nous avons à peu près élucidé la question de la stabilité du schéma explicite, nous n'avons rien dit sur sa convergence, c'est-à-dire sur sa capacité à approcher correctement la solution exacte. Nous répondrons rigoureusement à cette question au prochain chapitre. Remarquons que la stabilité est, bien sûr, une condition nécessaire de convergence, mais pas suffisante. Contentons nous pour l'instant de vérifier expérimentalement la convergence du schéma, c'est-à-dire que lorsque les pas d'espace et de temps deviennent de plus en plus petits les solutions numériques correspondantes

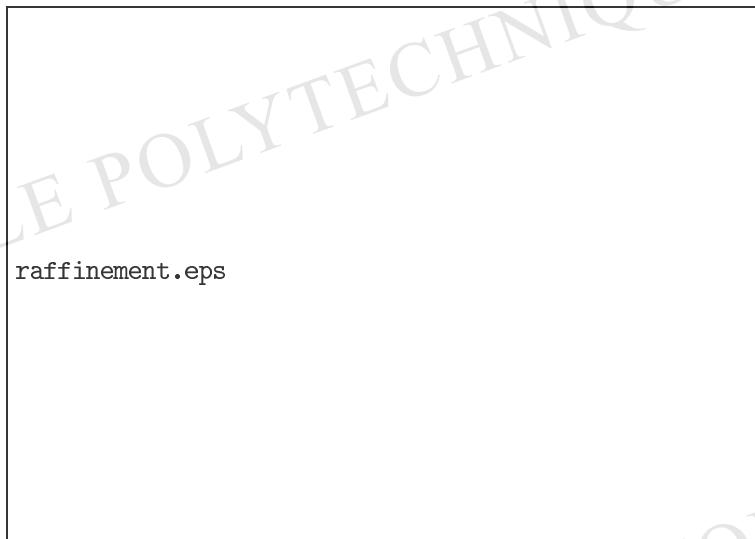


FIG. 1.8 – Schéma explicite avec  $\nu\Delta t = 0.4(\Delta x)^2$  pour diverses valeurs de  $\Delta x$ .

convergent et que leur limite est bien la solution exacte (nous pouvons vérifier ce dernier point puisqu’ici la solution exacte est disponible). Sur la Figure 1.8 nous vérifions numériquement que, si l’on raffine le pas d’espace  $\Delta x$  (qui prends les valeurs 0.5, 0.1, et 0.05) ainsi que le pas de temps  $\Delta t$  en gardant constant le rapport  $\nu\Delta t/(\Delta x)^2$  (le nombre CFL), alors la solution numérique est de plus en plus proche de la solution exacte. (La comparaison s’effectue au même temps final  $t = 1$ , donc le nombre de pas de temps augmente lorsque le pas de temps  $\Delta t$  diminue.) Ce procédé de “**vérification numérique de la convergence**” est très simple et on ne doit jamais hésiter à l’utiliser faute de mieux (c’est-à-dire si l’analyse théorique de la convergence est impossible ou trop difficile).

#### 1.4.3 Résultats numériques pour l’équation d’advection

Effectuons une deuxième série d’expériences numériques sur **l’équation de convection-diffusion** (1.27) avec une vitesse  $V = 1$  non nulle. Nous reprenons les mêmes données que précédemment et nous choisissons le schéma explicite avec  $\nu\Delta t = 0.4(\Delta x)^2$ . Nous regardons l’influence de la valeur de la constante de diffusion  $\nu$  (ou inverse du nombre de Péclet) sur la stabilité du schéma. La Figure 1.9 montre que le schéma est stable pour  $\nu = 1$ , instable pour  $\nu = 0.01$ , et que pour la valeur intermédiaire  $\nu = 0.1$ , le schéma semble stable mais la solution approchée est légèrement différente de la solution exacte. On comprend bien que plus l’inverse du nombre de Péclet  $\nu$  est petit, plus le terme convectif est prédominant sur le terme diffusif. Par conséquent, la condition CFL (1.31), obtenue lorsque la vitesse  $V$  est

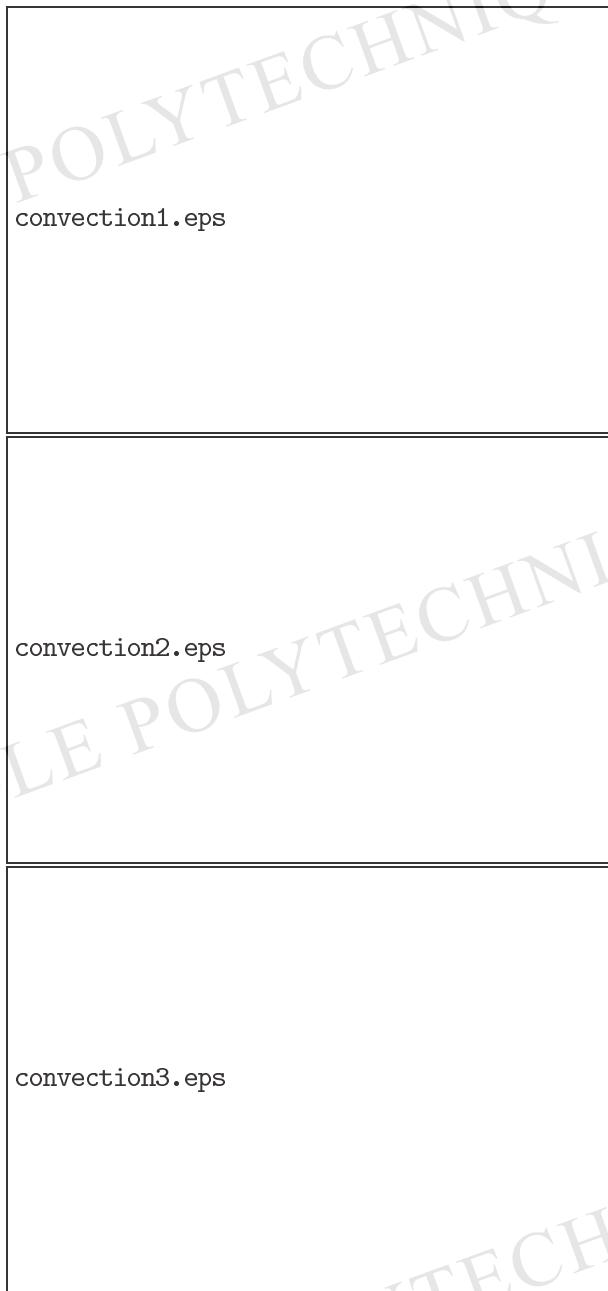


FIG. 1.9 – Schéma explicite pour l'équation de convection-diffusion avec  $\nu\Delta t = 0.4(\Delta x)^2$  et  $V = 1$ . En haut,  $\nu = 1$ , au milieu  $\nu = 0.1$ , et en bas  $\nu = 0.01$ .

nulle, est de moins moins en valable au fur et à mesure que  $\nu$  diminue.

Pour comprendre ce phénomène, examinons l'**équation d'advection** qui s'obtient à la limite  $\nu = 0$ . Remarquons tout d'abord que la condition CFL (1.31) est automatiquement satisfaite si  $\nu = 0$  (quel que soit  $\Delta t$  et  $\Delta x$ ), ce qui semble contradictoire avec le résultat expérimental du bas de la Figure 1.9.

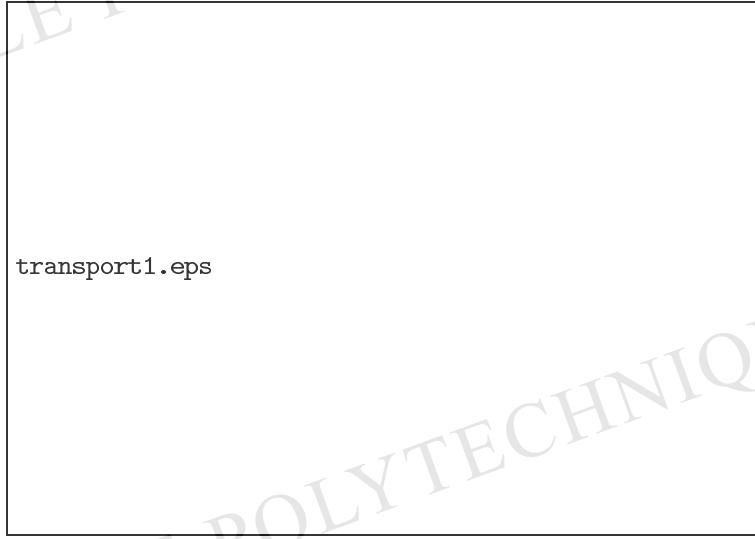


FIG. 1.10 – Schéma explicite centré pour l'équation d'advection avec  $\Delta t = 0.9\Delta x$ ,  $V = 1$ ,  $\nu = 0$ .

Pour l'équation d'advection (c'est-à-dire (1.27) avec  $\nu = 0$ ), le schéma explicite (1.30) peut se réécrire

$$u_j^{n+1} = \frac{V\Delta t}{2\Delta x} u_{j-1}^n + u_j^n - \frac{V\Delta t}{2\Delta x} u_{j+1}^n. \quad (1.34)$$

Ce schéma conduit aux oscillations de la Figure 1.10 dans les mêmes conditions expérimentales que le bas de la Figure 1.9. On voit bien que  $u_j^{n+1}$  n'est jamais (quel que soit  $\Delta t$ ) une combinaison convexe de  $u_{j-1}^n$ ,  $u_j^n$ , et  $u_{j+1}^n$ . Il ne peut donc y avoir de principe du maximum discret pour ce schéma, ce qui est une indication supplémentaire de son instabilité (une preuve rigoureuse en sera donnée au Lemme 2.3.1). L'origine de cette instabilité est que, dans le schéma explicite (1.34), nous avons choisi de traiter le terme convectif de manière centré. Nous pouvons cependant décentrer ce terme comme nous l'avons fait pour la dérivée en temps. Deux choix sont possibles : décentrer vers la droite ou vers la gauche. Le signe de la vitesse  $V$  est bien sûr crucial : ici nous supposons que  $V > 0$  (un argument symétrique est valable si  $V < 0$ ). Pour  $V > 0$ , le décentrement à droite est dit **décentrement aval** : on obtient

$$V \frac{\partial u}{\partial x}(t_n, x_j) \approx V \frac{u_{j+1}^n - u_j^n}{\Delta x}$$

en allant chercher “l’information” en suivant le courant. Ce choix conduit à un schéma décentré aval “désastreux”

$$\frac{u_j^{n+1} - u_j^n}{\Delta t} + V \frac{u_{j+1}^n - u_j^n}{\Delta x} = 0 \quad (1.35)$$

qui est tout aussi instable que le schéma centré. Au contraire le **décentrement amont** (c’est-à-dire à gauche si  $V > 0$ ), qui va chercher “l’information” en remontant le courant

$$V \frac{\partial u}{\partial x}(t_n, x_j) \approx V \frac{u_j^n - u_{j-1}^n}{\Delta x}$$

conduit au schéma explicite décentré amont

$$\frac{u_j^{n+1} - u_j^n}{\Delta t} + V \frac{u_j^n - u_{j-1}^n}{\Delta x} = 0 \quad (1.36)$$

qui donne les résultats de la Figure 1.11. On vérifie aisément que le schéma (1.36) est

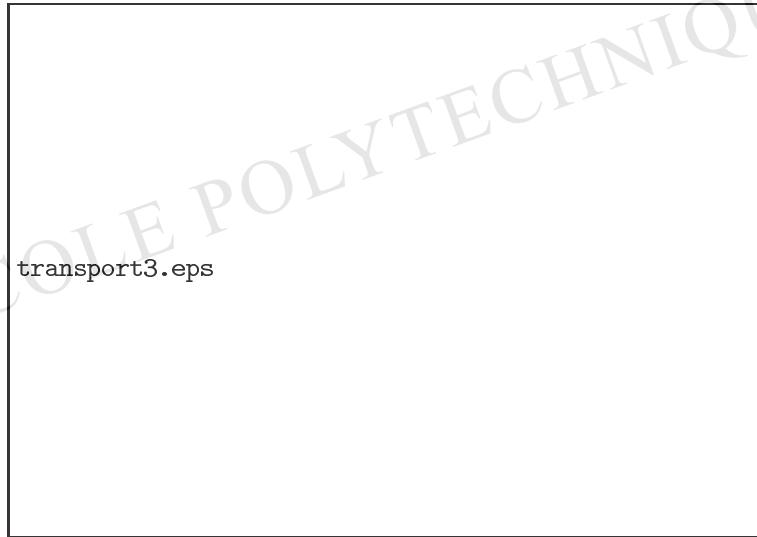


FIG. 1.11 – Schéma explicite décentré amont pour l’équation d’advection avec  $\Delta t = 0.9\Delta x$ ,  $V = 1$ .

stable sous une nouvelle condition CFL (différente de la précédente condition CFL (1.31))

$$|V|\Delta t \leq \Delta x. \quad (1.37)$$

En effet, on peut réécrire (1.36) sous la forme

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

qui montre que, si la condition (1.37) est satisfaite,  $u_j^{n+1}$  est une combinaison convexe de  $u_{j-1}^n$  et  $u_j^n$ . Par conséquent, le schéma décentré amont (1.36) vérifie un principe du maximum discret, ce qui entraîne sa stabilité conditionnelle. L'idée du **décentrement amont est une autre idée majeure de l'analyse numérique**. Elle est particulièrement cruciale dans tous les problèmes de mécanique des fluides où elle fut d'abord découverte (en anglais on parle de **upwinding**, c'est-à-dire de remonter le vent ou le courant), mais elle apparaît dans bien d'autres modèles.

La conclusion de cette étude sur l'équation d'advection est que pour le modèle de convection-diffusion avec faible valeur de la constante de diffusion  $\nu$ , il faut absolument décenter vers l'amont le terme convectif et suivre la condition CFL (1.37) plutôt que celle (1.31). A ce prix on peut améliorer les résultats de la Figure 1.9.

**Exercice 1.4.2** Montrer que, si la condition CFL (1.37) n'est pas satisfaite, le schéma décentré amont (1.36) pour l'équation d'advection est instable pour la donnée initiale  $u_j^0 = (-1)^j$ .

**Exercice 1.4.3** Écrire un schéma explicite centré en espace pour l'équation des ondes (1.18) en une dimension d'espace et sans terme source. Préciser comment démarrer les itérations en temps. Vérifier l'existence d'un cône de dépendance discret analogue à celui continu illustré par la Figure 1.3. En déduire que, si ce schéma converge, les pas de temps et d'espace doivent nécessairement satisfaire la condition (de type CFL)  $\Delta t \leq \Delta x$ .

Les conclusions de cette section sont nombreuses et vont nourrir les réflexions du prochain chapitre. Tout d'abord, tous les schémas numériques "raisonnables" ne fonctionnent pas, loin s'en faut. On rencontre des problèmes de stabilité (sans parler de convergence) qui nécessitent d'analyser en détails ces schémas : c'est la raison d'être de l'analyse numérique qui concilie objectifs pratiques et études théoriques. Enfin, les "bons" schémas numériques doivent respecter un certain nombre de propriétés (comme par exemple, le principe du maximum discret, ou le décentrement amont) qui ne sont que la traduction (au niveau discret) de propriétés physiques ou mathématiques de l'équation aux dérivées partielles. **On ne peut donc pas faire l'économie d'une bonne compréhension de la modélisation physique et des propriétés mathématiques des modèles si l'on veut réaliser de bonnes simulations numériques.**

## 1.5 Remarques sur les modèles mathématiques

Nous terminons ce chapitre par un certain nombre de définitions qui permettront au lecteur de s'y retrouver dans le vocabulaire employé ici comme dans les ouvrages classiques sur l'analyse numérique.

### 1.5.1 Notion de problème bien posé

**Définition 1.5.1** *On appelle problème aux limites une équation aux dérivées partielles munie de conditions aux limites sur la totalité de la frontière du domaine sur lequel elle est posée.*

Par exemple, le Laplacien (1.20) est un problème aux limites. A contrario, l'équation différentielle ordinaire

$$\begin{cases} \frac{dy}{dt} = f(t, y) \text{ pour } 0 < t < T \\ y(t = 0) = y_0 \end{cases} \quad (1.38)$$

n'est pas un problème aux limites puisqu'êtant posée sur un segment  $(0, T)$ , avec  $0 < T \leq +\infty$ , elle n'a de conditions "au bord" qu'en  $t = 0$  (et pas en  $t = T$ ).

**Définition 1.5.2** *On appelle problème de Cauchy une équation aux dérivées partielles où, pour au moins une variable (généralement le temps  $t$ ), les conditions "au bord" sont des conditions initiales (c'est-à-dire ne portent que sur un bord  $t = 0$ , et pas en  $t = T$ ).*

Par exemple, l'équation différentielle ordinaire (1.38) est un problème de Cauchy, mais pas le Laplacien (1.20) (quel que soit le choix de la composante de la variable d'espace  $x$  à qui on ferait jouer le rôle du temps).

De nombreux modèles sont à la fois des problèmes aux limites et des problèmes de Cauchy. Ainsi, l'équation de la chaleur (1.8) est un problème de Cauchy par rapport à la variable de temps  $t$  et un problème aux limites par rapport à la variable d'espace  $x$ . Tous les modèles que nous allons étudier dans ce cours rentrent dans une de ces deux catégories de problème.

Le fait qu'un modèle mathématique soit un problème de Cauchy ou un problème aux limites n'implique pas automatiquement qu'il s'agisse d'un "bon" modèle. L'expression **bon modèle** n'est pas employée ici au sens de la pertinence physique du modèle et de ses résultats, mais au sens de sa cohérence mathématique. Comme nous allons le voir cette cohérence mathématique est une condition nécessaire avant de pouvoir même envisager des simulations numériques et des interprétations physiques. Le mathématicien Jacques Hadamard a donné une définition de ce qu'est un "bon" modèle, en parlant de **problème bien posé** (un problème mal posé est le contraire d'un problème bien posé). On décide de noter  $f$  les données (le second membre, les données initiales, le domaine, etc.),  $u$  la solution recherchée, et  $\mathcal{A}$  "l'opérateur" qui agit sur  $u$ . Il s'agit ici de notations abstraites,  $\mathcal{A}$  désignant à la fois l'équation aux dérivées partielles et le type de conditions initiales ou aux limites. Le problème est donc de trouver  $u$  solution de

$$\mathcal{A}(u) = f \quad (1.39)$$

**Définition 1.5.3** *On dit que le problème (1.39) est bien posé si pour toute donnée  $f$  il admet une solution unique  $u$ , et si cette solution  $u$  dépend continûment de la donnée  $f$ .*

Examinons en détail cette définition de Hadamard : elle contient en fait trois conditions pour qu'un problème soit bien posé. Premièrement, il faut qu'il existe au moins une solution : c'est bien la moindre des choses à demander à un modèle sensé représenter la réalité ! Deuxièmement, il faut que la solution soit unique : c'est un point plus délicat car, s'il est clair que, lorsque en météorologie on prévoit le temps qu'il va faire demain, il vaut mieux pouvoir prédire "soleil" ou "pluie" (avec un "ou" exclusif) mais pas les deux avec des chances égales, il existe d'autres problèmes qui admettent "raisonnablement" plusieurs ou une infinité de solutions. Par exemple, les problèmes de plus court chemin admettent souvent plusieurs solutions : pour aller du pôle sud au pôle nord tout méridien convient, et de même, pour aller en avion de Paris à New York, votre agence de voyage vous fait passer tantôt par Bruxelles ou Londres, plutôt qu'un trajet direct, car plus court chemin veut dire ici plus économique. Hadamard exclut de sa définition ce type de problème car la multiplicité des solutions cache une indétermination du modèle : pour choisir finalement un chemin parmi tous ceux qui sont les plus courts, on utilise un autre critère (qu'on avait "oublié" jusque là) comme par exemple, le trajet le plus pratique ou confortable. C'est une situation courante en mathématiques appliquées : quand un modèle admet trop de solutions, il faut lui ajouter un critère de sélection de la "bonne" solution (voir l'exemple typique de la dynamique des gaz [18]). Troisièmement, et c'est la condition la moins évidente a priori, il faut que la solution dépende continûment des données. Au premier abord, cela semble une fantaisie de mathématicien, mais c'est pourtant crucial dans une perspective **d'approximation numérique**. En effet, faire un calcul numérique d'une solution approchée de (1.39) revient à perturber les données (qui de continues deviennent discrètes) et à résoudre (1.39) pour ces données perturbées. Si de petites perturbations des données conduisent à de grandes perturbations de la solution, il n'y a aucune chance pour que la simulation numérique soit proche de la réalité (ou du moins de la solution exacte). Par conséquent, cette dépendance continue de la solution par rapport aux données est une condition absolument nécessaire pour envisager des simulations numériques précises. Remarquons que cette condition est aussi très importante d'un point de vue physique car les appareils de mesure physique des données n'ont qu'une précision relative : si l'on est incapable de distinguer deux données très proches mais conduisant à des phénomènes très différents, le modèle représenté par (1.39) n'a aucune valeur prédictive, et donc un intérêt pratique à peu près nul.

Terminons en avouant qu'à ce niveau de généralité la Définition (1.5.3) est bien floue, et que pour lui donner un sens mathématique précis il faut bien sûr dire dans quels espaces de fonctions on place les données et on cherche la solution, et quelles normes ou topologies on utilise pour la continuité. Il n'est pas rare en effet qu'un changement d'espace (bien anodin en apparence) entraîne des propriétés d'existence ou d'unicité fort différentes !

**Exercice 1.5.1** Le but de cet exercice est de montrer que le problème de Cauchy pour le Laplacien est mal posé. Soit le domaine bidimensionnel  $\Omega = (0, 1) \times (0, 2\pi)$ . On considère

le problème de Cauchy en  $x$  et le problème aux limites en  $y$  suivant

$$\begin{cases} -\frac{\partial^2 u}{\partial x^2} - \frac{\partial^2 u}{\partial y^2} = 0 & \text{dans } \Omega \\ u(x, 0) = u(x, 2\pi) = 0 & \text{pour } 0 < x < 1 \\ u(0, y) = 0, \frac{\partial u}{\partial x}(0, y) = -e^{-\sqrt{n}} \sin(ny) & \text{pour } 0 < y < 2\pi \end{cases}$$

Vérifier que  $u(x, y) = \frac{e^{-\sqrt{n}}}{n} \sin(ny) \operatorname{sh}(nx)$  est une solution. Montrer que la condition initiale et toutes ses dérivées en  $x = 0$  convergent uniformément vers 0, tandis que, pour tout  $x > 0$ , la solution trouvée  $u(x, y)$  et toutes ses dérivées ne sont pas bornées quand  $n$  tend vers l'infini. Conclure.

### 1.5.2 Classification des équations aux dérivées partielles

**Définition 1.5.4** On appelle **ordre** d'une équation aux dérivées partielles l'ordre de la plus grande dérivée présente dans l'équation.

Par exemple, le Laplacien (1.20) est une équation du deuxième ordre, tandis que l'équation des plaques (1.24) est une équation du quatrième ordre. On distingue souvent l'ordre par rapport à la variable de temps  $t$  et par rapport à la variable d'espace  $x$ . Ainsi, on dira que l'équation de l'équation de la chaleur (1.8) est du premier ordre en temps et du deuxième ordre en espace, alors que l'équation des ondes (1.18) est du deuxième ordre en espace-temps.

Pour comprendre le vocabulaire souvent employé d'équation aux dérivées partielles **soit elliptique, soit parabolique, soit hyperbolique**, nous allons brièvement classifier les équations aux dérivées partielles linéaires du deuxième ordre portant sur des fonctions réelles de deux variables réelles  $u(x, y)$  (nous ne cherchons absolument pas à effectuer une classification systématique de toutes les e.d.p.). Une telle équation s'écrit

$$a \frac{\partial^2 u}{\partial x^2} + b \frac{\partial^2 u}{\partial x \partial y} + c \frac{\partial^2 u}{\partial y^2} + d \frac{\partial u}{\partial x} + e \frac{\partial u}{\partial y} + fu = g. \quad (1.40)$$

Pour simplifier nous supposons que les coefficients  $a, b, c, d, e, f$  sont constants.

**Définition 1.5.5** On dit que l'équation (1.40) est **elliptique** si  $b^2 - 4ac < 0$ , **parabolique** si  $b^2 - 4ac = 0$ , et **hyperbolique** si  $b^2 - 4ac > 0$ .

L'origine de ce vocabulaire est bien sûr la classification des coniques du plan, sur laquelle la Définition 1.5.5 est calquée. En effet, il est bien connu que l'équation du deuxième degré

$$ax^2 + bxy + cy^2 + dx + ey + f = 0$$

définit une courbe plane qui est (sauf certains cas dégénérés) une ellipse si  $b^2 - 4ac < 0$ , une parabole si  $b^2 - 4ac = 0$ , et une hyperbole si  $b^2 - 4ac > 0$ .

Si on applique la Définition 1.5.5 aux divers modèles du deuxième ordre que nous avons évoqué dans ce chapitre (en remplaçant le couple  $(x, y)$  par les variables  $(t, x)$  en une dimension d'espace), nous concluons que l'**équation de la chaleur est parabolique** (de même que l'équation de convection-diffusion), que le **Laplaciens est elliptique**, et que l'**équation des ondes est hyperbolique**. Une généralisation adéquate de cette définition permet

d'affirmer que l'équation d'advection est hyperbolique, et que les équations de Stokes, de l'élasticité, ou des plaques sont elliptiques. En règle générale les problèmes stationnaires (indépendants du temps) sont modélisés par des e.d.p. elliptiques, tandis que les problèmes d'évolution sont modélisés par des e.d.p. paraboliques ou hyperboliques.

Nous verrons plus loin que les problèmes aux limites sont bien posés pour les équations aux dérivées partielles elliptiques, tandis que les problèmes de Cauchy en temps et aux limites en espace sont bien posés pour les équations aux dérivées partielles paraboliques ou hyperboliques. Il y a donc une différence importante de comportement entre ces types d'équations.

**Remarque 1.5.6** Le caractère elliptique, parabolique, ou hyperbolique de l'équation (1.40) n'est pas modifié par un changement de variables. Soit  $(x, y) \rightarrow (X, Y)$  un tel changement de variables non singulier, c'est-à-dire tel que son Jacobien  $J = X_x Y_y - X_y Y_x$  ne s'annule pas (on note  $Z_z$  la dérivée de  $Z$  par rapport à  $z$ ). Un calcul, simple dans le principe mais long dans le détail, montre que (1.40) devient

$$A \frac{\partial^2 u}{\partial X^2} + B \frac{\partial^2 u}{\partial X \partial Y} + C \frac{\partial^2 u}{\partial Y^2} + D \frac{\partial u}{\partial X} + E \frac{\partial u}{\partial Y} + Fu = G,$$

avec notamment  $A = aX_x^2 + bX_x X_y + cX_y^2$ ,  $B = 2aX_x Y_x + b(X_x Y_y + X_y Y_x) + 2cX_y Y_y$ ,  $C = aY_x^2 + bY_x Y_y + cY_y^2$ , et on vérifie que  $B^2 - 4AC = J^2(b^2 - 4ac)$ . En particulier, un changement de variables adéquat permet de simplifier l'équation aux dérivées partielles (1.40) pour la ramener sous sa forme "canonique". Ainsi, toute équation elliptique peut se ramener au Laplacien  $\frac{\partial^2}{\partial X^2} + \frac{\partial^2}{\partial Y^2}$ , toute équation parabolique à l'équation de la chaleur  $\frac{\partial}{\partial X} - \frac{\partial^2}{\partial Y^2}$ , et toute équation hyperbolique à l'équation des ondes  $\frac{\partial^2}{\partial X^2} - \frac{\partial^2}{\partial Y^2}$ . •

**Remarque 1.5.7** On sait bien que l'équation générale des coniques du plan admet un certain nombre de cas dégénérés où elle ne décrit plus une conique mais un ensemble de droites, voire un seul point. La même situation de dégénérescence peut avoir lieu avec l'équation aux dérivées partielles (1.40). Par exemple, l'équation  $\frac{\partial^2 u}{\partial x^2} = 1$  avec  $a = 1$  et  $b = c = d = e = f = 0$  n'est pas parabolique en dimension deux (bien que  $b^2 - 4ac = 0$ ) mais elliptique en dimension un (la variable  $y$  ne joue ici aucun rôle). Il faut donc faire un peu attention avant de conclure sur le type de ces équations "dégénérées". •