

---

# Approximation numérique des équations différentielles et aux dérivées partielles

---

*Grégory Vial*



Cours d'approfondissement S7

31 août 2015



# Table des matières

<b>1</b>	<b>Introduction : modèles mathématiques, numériques, informatiques</b>	<b>5</b>
<b>2</b>	<b>Équations différentielles ordinaires</b>	<b>7</b>
2.1	Repères théoriques . . . . .	7
2.2	La méthode d'Euler . . . . .	8
2.3	Notions générales de consistance et de stabilité . . . . .	9
2.4	Quelques méthodes classiques à un pas . . . . .	10
2.5	Propriétés qualitatives des schémas . . . . .	11
2.5.1	Stabilité asymptotique . . . . .	11
2.5.2	Conservativité . . . . .	12
<b>3</b>	<b>Équations aux dérivées partielles</b>	<b>15</b>
3.1	Introduction . . . . .	15
3.2	Repères théoriques . . . . .	16
3.2.1	L'équation de Laplace . . . . .	16
3.2.2	L'équation de la chaleur . . . . .	16
3.2.3	L'équation des ondes et le transport linéaire . . . . .	17
3.3	Méthodes numériques pour les problèmes elliptiques . . . . .	19
3.3.1	En dimension 1 . . . . .	19
3.3.2	En dimension supérieure . . . . .	21
3.3.3	Un cas non-linéaire . . . . .	21
3.4	Problèmes d'évolution paraboliques . . . . .	22
3.4.1	L'équation de la chaleur . . . . .	22
3.4.2	Les problèmes de réaction-diffusion . . . . .	23
3.5	Phénomènes de propagation . . . . .	24
3.5.1	L'équation de transport linéaire en une dimension . . . . .	24
3.5.2	L'équation des ondes . . . . .	24
	<b>Références</b>	<b>25</b>



# Introduction : modèles mathématiques, numériques, informatiques

La modélisation mathématique consiste à décrire à l'aide du langage mathématique un phénomène « concret ». Selon qu'il provienne la physique, la biologie, l'économie, etc, les techniques de modélisation diffèrent. En tout état de cause, le modèle est une *approximation* de la situation réelle, dont on n'est en général pas capable de décrire tous les mécanismes (mais qu'on cherche parfois aussi à simplifier davantage).

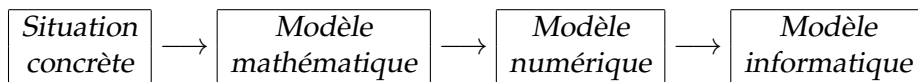
Cette étape de modélisation « passage du concret aux équations », n'est pas *stricto sensu* un travail de mathématicien, bien qu'elle fasse usage de notations et concepts mathématiques. En effet, il ne s'agit pas ici de démontrer, de prouver les différents développements conduisant au modèle, mais bien *d'écrire* ce modèle à l'aide de considérations physiques, chimiques, biologiques, économiques, etc.

Une fois le modèle écrit, on peut alors l'aborder en tant qu'objet mathématique. Le mathématicien intervient alors pour prouver, par exemple, qu'il admet une unique solution, et qu'elle satisfait certaines propriétés. Insistons sur le fait que l'intuition physique (au sens large, elle peut aussi bien être économique) guide la formulation des questions mathématiques à considérer, mais en aucun cas n'aide à les résoudre. Un argument du type : « la solution du modèle est une fonction positive car elle représente une concentration chimique » n'est pas valable. La validation mathématique de la positivité de la solution constitue précisément une validation du modèle *a posteriori*.

La plupart du temps, les modèles mathématiques sont trop complexes pour pouvoir être résolus explicitement (« à la main »). On a alors recours à une approximation numérique qui – elle – sera calculable. On parle alors parfois de modèle numérique. Encore une fois, il convient de construire et surtout valider le modèle numérique en se basant exclusivement sur le modèle mathématique, et non pas en le couplant avec des ingrédients physiques (même si l'intuition demeure un guide).

L'étape suivante consiste à programmer le modèle numérique pour le résoudre sur machine. On aboutit à un modèle informatique.

Le schéma ci-dessous résume les différentes étapes dans la résolution du problème initial.



Il est important d'avoir en tête les différents types de modèles intervenant, afin de pouvoir mieux valider chaque processus, et déterminer où se situe une erreur, le cas échéant.

Dans ce cours d'approfondissement, on donnera un aperçu de la modélisation mathématique, à l'aide d'exemples issus d'applications diverses. Le cœur du cours consiste à détailler la modélisation numérique pour les modèles de type équations différentielles ordinaires (EDO) et équations aux dérivées partielles (EDP). La modélisation informatique sera abordée avec le logiciel `matlab`, mais sa maîtrise n'est pas un objectif ici.

Le présent polycopié n'a pas vocation à être un cours complet, mais plutôt un document de référence reprenant les points principaux qui auront été vus en séances.

# Équations différentielles ordinaires

Dans ce chapitre, on considère une équation (ou un système d'équations) différentielle(s), écrit sous la forme

$$\begin{cases} u'(t) = f(t, u(t)) & \text{pour } t \in [0, T], \\ u(0) = u_0. \end{cases} \quad (2.1)$$

Les données du problème sont

- ◇ un temps final  $T$ ,
- ◇ un domaine  $\Omega \subset \mathbb{R}^d$ ,
- ◇ une fonction continue  $f : [0, T] \times \Omega \rightarrow \mathbb{R}^d$ ,
- ◇ un vecteur initial  $u_0 \in \Omega$ .

## Remarque 1

*Si la dimension  $d$  vaut 1, on parle d'équation, sinon de système d'équations. Lorsque la fonction  $f$  ne dépend pas explicitement du temps, i.e.  $f(t, v) = \tilde{f}(v)$ , on dit que le système est autonome.*

La présentation est volontairement très synthétique ; pour plus de détails on pourra utilement consulter les livres [2, 3, 4, 5]

## 2.1 Repères théoriques

Dans la suite, on se placera sous l'hypothèse suivante :

$$\exists L > 0, \quad \forall t \in [0, T], \quad \forall v, w \in \Omega, \quad \|f(t, v) - f(t, w)\| \leq C\|v - w\|, \quad (2.2)$$

qui assure, d'après le théorème de Cauchy-Lipschitz, l'existence et l'unicité d'une solution au problème (2.1).

Bien que le cadre précédent soit agréable d'un point de vue mathématique, peu de modèles de type EDO en relèvent. Citons toutefois l'équation du pendule simple

$$\begin{cases} \theta''(t) = -\frac{g}{l} \sin(\theta) - \alpha \theta' & \text{pour } t \in [0, T], \\ \theta(0) = \theta_0, \\ \theta'(0) = \zeta_0. \end{cases}$$

En effet, en posant  $u(t) = (\theta(t), \theta'(t)) \in \mathbb{R}^2$ , le problème s'écrit sous la forme (2.1) avec

$$f : (t, v) \mapsto \left(-\frac{g}{\ell} \sin(v_1) - \alpha v_2, v_2\right) \quad \text{et} \quad u_0 = (\theta_0, \zeta_0).$$

## 2.2 La méthode d'Euler

On se place sous l'hypothèse (2.2) et on note  $\varphi$  la solution du problème (2.1). Afin de construire une approximation numérique de la fonction  $\varphi$ , on introduit une subdivision  $t_0 < t_1 < \dots < t_N$  de l'intervalle  $[0, T]$  dont le pas maximal est noté  $h$  :

$$h = \max_{0 \leq n \leq N-1} h_n \quad \text{avec} \quad h_n = t_{n+1} - t_n.$$

On écrit alors l'équation différentielle sous forme intégrale :

$$\varphi(t_{n+1}) = \varphi(t_n) + \int_{t_n}^{t_{n+1}} f(s, \varphi(s)) ds$$

Les approximations  $u_0, u_1, \dots, u_N$  sont fournies par l'approximation de l'intégrale dans l'égalité précédente à l'aide de la méthode du rectangle, à savoir

$$u_{n+1} = u_n + h_n f(t_n, u_n) \quad (n = 0, 1, \dots, N-1).$$

L'initialisation a lieu à partir de la condition initiale (déjà notée  $u_0$ ) et la détermination de  $u_{n+1}$  connaissant  $u_n$  est un simple calcul. La figure 2.1 montre la construction géométrique des itérés successifs de la méthode d'Euler. La courbe rouge correspond à la solution issue du point  $(t_0, u_0)$  (celle qu'on cherche à approcher), la courbe bleue celle issue du point  $(t_1, u_1)$ .

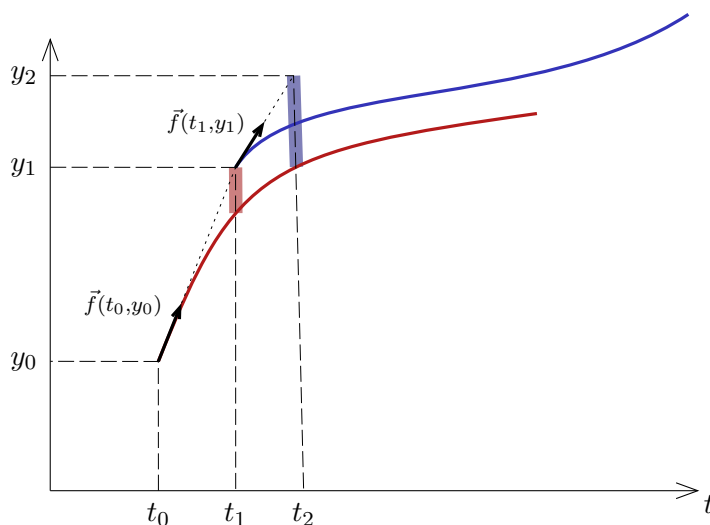


FIGURE 2.1 – La méthode d'Euler en dimension 1.

Les bandes rouge et bleue matérialisent l'erreur commise à chaque étape, où il apparaît clairement un phénomène de propagation des erreurs déjà présentes, cumulé avec une erreur intrinsèque à chaque étape. Ce constat conduit à séparer en deux parties les sources d'erreur pour étudier la convergence de la méthode (pour cette étude, on suppose pour simplifier les notations que le pas de la subdivision est constant :  $t_{n+1} - t_n = h$  pour tout  $n$ ) :



- ◇ *Consistance.* on définit l'erreur de consistance<sup>1</sup> comme l'écart à la courbe exacte auquel conduit le schéma en une itération, sans erreur initiale : à l'étape  $n$  elle est définie par

$$\varepsilon_n = \varphi(t_{n+1}) - \varphi(t_n) - h_n f(t_n, \varphi(t_n)).$$

Si la solution  $y$  est suffisamment régulière (de classe  $\mathcal{C}^2$ , ce qui est assuré par l'hypothèse  $f \in \mathcal{C}^1$ ), un développement de Taylor fournit immédiatement l'estimation

$$|\varepsilon_n| \leq \frac{h^2}{2} \sup_{t \in [0, T]} |\varphi''(t)|. \quad (2.3)$$

- ◇ *Stabilité.* cette notion mesure la manière dont le schéma propage les erreurs. On considère un schéma perturbé à chaque étape :

$$v_{n+1} = v_n + h_n f(t_n, v_n) + \mu_n \quad (n = 0, 1, \dots, N-1),$$

(avec  $v_0 = u_0$ ). On cherche à savoir si  $v_n$  est proche de  $u_n$ . Pour cela, on écrit la majoration

$$|v_{n+1} - u_{n+1}| \leq (1 + hL)|v_n - u_n| + |\mu_n|,$$

où  $L$  est une constante de Lipschitz pour la fonction  $f$ . Une récurrence immédiate<sup>2</sup> montre que

$$|u_n - v_n| \leq \sum_{k=0}^{n-1} (1 + hL)^k |\mu_{n-k-1}|$$

d'où

$$\max_{0 \leq n \leq N} |v_n - u_n| \leq e^{LT} \sum_{n=0}^{N-1} |\mu_n|. \quad (2.4)$$

- ◇ *Convergence.* Si l'on choisit  $v_n = \varphi(t_n)$ , alors la perturbation  $\mu_n$  n'est autre que l'erreur de consistance  $\varepsilon_n$ . Les inégalités (2.3) et (2.4) se combinent pour fournir

$$\max_{0 \leq n \leq N} |\varphi(t_n) - u_n| \leq e^{LT} N \frac{h^2}{2} \sup_{t \in [0, T]} |\varphi''(t)| = Te^{LT} \frac{h}{2} \sup_{t \in [0, T]} |\varphi''(t)|.$$

Il s'agit bien d'un résultat de convergence des itérés de la méthode vers les valeurs de la solution exacte lorsque le pas  $h$  tend vers 0.

## 2.3 Notions générales de consistance et de stabilité

Considérons un schéma qui s'écrit sous la forme

$$u_{n+1} = G(\mathbf{h}_n; u_n, u_{n-1}, \dots, u_{n-p}),$$

où  $\mathbf{h}_n$  désigne le vecteur des pas  $(h_i)_{n-p \leq i \leq n}$ . On dit que le schéma est à  $p+1$  pas (la méthode d'Euler est à un pas). On note toujours  $\varphi$  la solution continue du problème de Cauchy.

### Définition 1 (consistance)

L'erreur de consistance à l'étape  $n$  (avec  $p \leq n \leq N-1$ ) est la quantité

$$\varepsilon_n = \varphi(t_{n+1}) - G(\mathbf{h}_n, \varphi(t_n), \varphi(t_{n-1}), \dots, \varphi(t_{n-p})).$$

1. Il s'agit d'un anglicisme ; « cohérence » serait plus juste, mais contraire à l'usage. ...  
2. Ce résultat est parfois appelé *lemme de Grönwall discret*

Le schéma (2.3) est consistant si  $\varepsilon = o(h)$  lorsque  $h \rightarrow 0$ . Il est dit consistant d'ordre  $k$  si  $\varepsilon_n = \mathcal{O}(h^{k+1})$ .

### Définition 2 (stabilité)

Soit  $(\mu_n)_{p \leq n \leq N-1}$  donnée. On suppose  $v_i = u_i$  pour  $i \leq p$  et on pose

$$v_{n+1} = G(\mathbf{h}_n; v_n, v_{n-1}, \dots, v_{n-p}) + \mu_n.$$

On dit que le schéma (2.3) est stable s'il existe une constante  $C > 0$  indépendante de  $h$  telle que

$$\max_{n=0}^{N-1} \|v_n - u_n\| \leq C \sum_{n=p}^{N-1} \|\mu_n\|.$$

Le résultat suivant, connu sous le nom de *théorème de Lax*, est une conséquence directe des définitions.

### Théorème 1 (Lax)

Un schéma consistant et stable est convergent, dans le sens où

$$\max_{n=0}^{N-1} \|\varphi(t_n) - u_n\| \rightarrow 0 \quad \text{lorsque } h \rightarrow 0.$$

Si le schéma est, de plus, d'ordre  $k$  alors la convergence a lieu en  $\mathcal{O}(h^k)$ .

## 2.4 Quelques méthodes classiques à un pas

Citons ici quelques schémas couramment utilisés. La mention *explicite* signifie que la détermination de  $u_{n+1}$  à partir de  $u_n$  est un simple calcul. Pour les méthodes dites *implicites*, la résolution d'une équation est nécessaire pour cela.

**La méthode d'Euler rétrograde** (implicite, ordre 1)

$$u_{n+1} = u_n + h_n f(t_{n+1}, u_{n+1}).$$

**La méthode de Heun** (explicite, ordre 2)

$$u_{n+1} = u_n + \frac{h_n}{2} \left[ f(t_n, u_n) + f(t_{n+1}, u_n + h_n f(t_n, u_n)) \right].$$

**La méthode de Crank-Nicolson** (implicite, ordre 2)

$$u_{n+1} = u_n + \frac{h_n}{2} \left[ f(t_n, u_n) + f(t_{n+1}, u_{n+1}) \right].$$

**La méthode de Runge-Kutta RK4** (explicite, ordre 4)

$$\begin{cases} k_1^n &= f(t_n, u_n) \\ k_2^n &= f(t_n + \frac{h_n}{2}, u_n + \frac{h_n}{2}k_1^n) \\ k_3^n &= f(t_n + \frac{h_n}{2}, u_n + \frac{h_n}{2}k_2^n) \\ k_4^n &= f(t_{n+1}, u_n + h_n k_3^n) \\ u_{n+1} &= u_n + \frac{h_n}{6} (k_1^n + 2k_2^n + 2k_3^n + k_4^n) \end{cases}$$

## 2.5 Propriétés qualitatives des schémas

L'ordre de consistance d'une méthode n'est pas le seul indicateur d'efficacité. En effet, il n'est pas toujours possible de choisir le pas  $h$  aussi petit que nécessaire (par exemple, lorsqu'on est intéressé par la comportement en temps long de la solution et que le temps de calcul est compté). Il est alors crucial d'employer des méthodes qui respectent certaines propriétés de la solution continue du problème de Cauchy.

### 2.5.1 Stabilité asymptotique

Il s'agit ici de passer un schéma au crible de la résolution du problème linéaire scalaire (donc très simple !)

$$u' = -\lambda u \quad \text{avec} \quad \lambda > 0.$$

La solution est donnée par  $\varphi(t) = u_0 e^{-\lambda t}$ , si bien qu'elle reste positive en tout temps (on a supposé  $u_0 > 0$ ). Il est légitime d'imposer à un schéma numérique de faire de même. Une récurrence immédiate conduit aux expressions suivantes (pour un pas fixe, i.e.  $\forall n, h_n = h$ ) :

- ◇ Euler explicite :  $u_n = (1 - \lambda h)^n u_0$ .
- ◇ Euler rétrograde :  $u_n = (1 + \lambda h)^{-n} u_0$ .

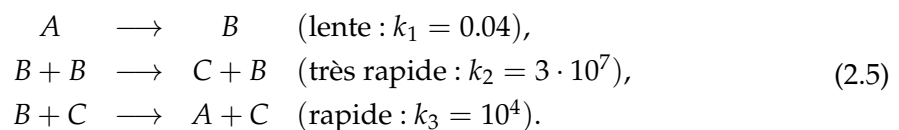
Ainsi, une condition apparaît pour que la méthode d'Euler explicite respecte la positivité :

$$h \leq \frac{1}{\lambda},$$

qui devient très restrictive pour des grandes valeurs de  $\lambda$ . En revanche, la méthode implicite conserve la positivité quel que soit le pas  $h$ .

On dit que la méthode d'Euler explicite est *conditionnellement* asymptotiquement stable (ou A-stable), et la méthode d'Euler implicite est *inconditionnellement* asymptotiquement stable. On peut retenir que les méthodes implicites, en général, sont préférables aux méthodes explicites lorsque la solution continue présente des variations fortes.

On parle alors de problème *raides*, qu'on rencontre fréquemment en cinétique chimique lorsque des réaction avec des vitesses de grandeurs différentes sont mises en jeu. Le prototype de ce genre d'équations est donné par le problème de Robertson (1966) : les quantités chimiques en jeu sont appelées  $A, B, C$  pour plus de simplicité ; le bilan global de la réaction est  $A \rightarrow C$ . On décompose le mécanisme chimique complexe en trois réactions élémentaires dont les lois cinétiques sont supposées d'ordre 1 :



La première réaction est nommée *amorçage*, il s'agit de la formation lente du catalyseur  $B$ . Le deuxième réaction produit le composé  $C$  très rapidement, et la troisième traduit la recombinaison du catalyseur  $B$  (*rupture*); cette dernière justifie le terme d'autocatalyse.

Les vitesses des trois réactions en jeu sont très différentes comme le montrent les constantes de réaction situées à droite des équations chimiques. Les lois de la cinétique chimique permettent d'obtenir le système différentiel suivant, où  $y_a(t), y_b(t), y_c(t)$  désignent les quantités respectives des composés  $A, B$  et  $C$  à l'instant  $t$ ,

$$\begin{cases} y'_a(t) &= -k_1 y_a(t) + k_3 y_b(t) y_c(t) \\ y'_b(t) &= k_1 y_a(t) - k_3 y_b(t) y_c(t) - k_2 [y_b(t)]^2 \\ y'_c(t) &= k_2 [y_b(t)]^2. \end{cases} \quad (2.6)$$

Ces équations sont assorties de conditions initiales, traduisant la seule présence du réactif  $A$  :

$$\begin{aligned} y_a(0) &= 1, \\ y_b(0) &= 0, \\ y_c(0) &= 0. \end{aligned} \quad (2.7)$$

## 2.5.2 Conservativité

Certains problèmes préservent naturellement une quantité au cours du temps (énergie dans les problèmes conservatifs en mécanique, par exemple). Il est très naturel d'imposer au niveau discret (i.e. au niveau du schéma numérique) cette conservation.

L'exemple typique d'un tel comportement est donné par le système linéaire

$$x' = -y \quad \text{et} \quad y' = x.$$

La solution associée à la condition initiale  $x(0) = 1, y(0) = 0$  s'écrit

$$\varphi(t) = (\cos t, \sin t),$$

si bien que la quantité

$$E(t) = x(t)^2 + y(t)^2$$

est constante au cours du temps (ce qui se voit aussi bien en calculant la dérivée  $E'(t)$  et en utilisant les équations satisfaites par les fonctions  $x$  et  $y$ ).

On peut facilement déterminer l'évolution de l'équivalent discret  $E_n = x_n^2 + y_n^2$  pour les différentes méthodes décrites plus haut.

Euler explicite	$x_{n+1}^2 + y_{n+1}^2 = (1 + h^2)(x_n^2 + y_n^2) ;$
Euler implicite	$x_{n+1}^2 + y_{n+1}^2 = (1 + h^2)^{-1} (x_n^2 + y_n^2) ;$
Heun	$x_{n+1}^2 + y_{n+1}^2 = \left(1 + \frac{h^4}{4}\right) (x_n^2 + y_n^2) ;$
Crank-Nicolson	$x_{n+1}^2 + y_{n+1}^2 = x_n^2 + y_n^2 ;$
RK4	$x_{n+1}^2 + y_{n+1}^2 = \left(1 - \frac{h^6}{72} + \frac{h^8}{576}\right) (x_n^2 + y_n^2) ;$

Il apparaît que la méthode de Cank-Nicolson préserve exactement la quantité  $E_n$  et est, en cela, meilleure que la méthode de RK4 pourtant d'ordre plus élevé. Notons que,

toutefois, seules les méthodes d'Euler sont à proscrire car les méthodes de Heun et RK4 présentent des déviations tout-à-fait acceptables de l'énergie.

La situation est plus complexe à étudier dans le cas de systèmes non-linéaires (cf. l'étude des méthodes symplectiques pour les systèmes hamiltoniens), comme le célèbre système proie-prédateur de Lotka-Volterra ( $x$  représente la densité de population des proies,  $y$  celle des prédateurs) :

$$\begin{cases} x' = ax - bxy, \\ y' = -cy + dxy. \end{cases}$$

L'énergie associée s'écrit

$$E(t) = dx - c \ln x + by - a \ln y.$$



# Équations aux dérivées partielles

**Notation.** Dans tout le chapitre, on notera  $\frac{\partial u}{\partial x_i}$  ou  $\partial_{x_i} u$  ou encore  $\partial_i u$  la dérivée partielle de la fonction  $u$  selon la variable  $x_i$ .

## 3.1 Introduction

Une équation aux dérivées partielles (EDP) est une égalité qui fait intervenir des dérivées d'une fonction par rapport à ses différentes variables. Nous allons nous concentrer sur les équations d'ordre 2 (c'est-à-dire que les dérivées impliquées sont d'ordre au plus 2) qui sont de loin les plus fréquentes. Enfin, nous nous restreignons ici au cas des équations linéaires car le cas non linéaire est d'un tout autre niveau de difficulté, tant d'un point de vue théorique que numérique.

Plaçons-nous, pour fixer les idées, en dimension 2 et considérons donc une équation d'inconnue  $u$ , sous la forme

$$a \frac{\partial^2 u}{\partial x^2} + b \frac{\partial^2 u}{\partial y^2} + c \frac{\partial^2 u}{\partial x \partial y} + \alpha \frac{\partial u}{\partial x} + \beta \frac{\partial u}{\partial y} + \gamma u = f,$$

où les coefficients  $a, b, c, d, e, f$  sont des constantes fixées. On peut montrer qu'à l'aide d'un changement de variable (et, éventuellement de fonction) adéquat, il est possible de ramener cette EDP à une forme canonique qui relève d'une catégorie parmi les trois possibles ci-dessous

◇ *Cas elliptique.*

$$\frac{\partial^2 U}{\partial X^2} + \frac{\partial^2 U}{\partial Y^2} = F.$$

◇ *Cas hyperbolique.*

$$\frac{\partial^2 U}{\partial X^2} - \frac{\partial^2 U}{\partial Y^2} = F.$$

◇ *Cas parabolique.*

$$\frac{\partial U}{\partial X} \pm \frac{\partial^2 U}{\partial Y^2} = F.$$

Il est donc crucial d'étudier ces trois équations-modèles, tant q'un point de vue théorique que des méthodes numériques. Dans les modèles physiques – et aussi largement d'un point de vue mathématique – la variable temporelle  $t$  joue un rôle particulier par

rapport aux variables d'espace  $\mathbf{x} = (x_1, x_2, \dots, x_d)$ . Dans les trois équations modèles ci-dessus,  $X$  désignera souvent la variable temporelle. Dans le cas parabolique, le choix du sens de variation du temps (vers le futur ou vers le passé) et le choix du signe  $\pm$  dans l'équation ne peuvent être quelconques, comme on le verra au paragraphe 3.2.

## 3.2 Repères théoriques

### 3.2.1 L'équation de Laplace

On considère l'équation modèle elliptique, qui s'écrit en dimension  $d \geq 1$

$$\Delta u(\mathbf{x}) = f(\mathbf{x}) \quad \text{avec} \quad \Delta = \frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2} + \dots + \frac{\partial^2}{\partial x_d^2}.$$

Bien sûr, le problème n'est pas bien défini tel quel. Tout d'abord, précisons que la relation précédente doit avoir lieu pour  $\mathbf{x} \in \Omega$ , où  $\Omega$  est un sous-ensemble (ouvert) de  $\mathbb{R}^d$ . Il n'y a manifestement pas unicité de la solution (ajouter une constante à  $u$  ne modifie pas l'équation, par exemple). On ajoute donc à l'équation une condition aux limites (ou au bord), par exemple

$$u(\mathbf{x}) = 0 \quad \text{sur} \quad \partial\Omega.$$

Le cas de la dimension 1, avec  $\Omega = ]0, 1[$ , s'écrit

$$u''(x) = f(x) \quad \text{avec} \quad u(0) = u(1) = 0.$$

Il est alors élémentaire de montrer qu'il existe une unique solution donnée, après double intégration, par

$$\forall x \in ]0, 1[, \quad u(x) = \int_0^x \int_0^y f(s) \, ds - x \int_0^1 \int_0^y f(s) \, ds.$$

La fonction  $u$  ainsi définie est clairement de classe  $\mathcal{C}^2$  dès que  $f$  est continue, et le sens à donner à l'EDP est clair. La situation est plus complexe en dimension  $d \geq 2$  car, bien sûr, on n'est plus capables d'effectuer des calculs explicites. Néanmoins, on peut montrer que le problème admet une unique solution, de classe  $\mathcal{C}^2$  si le bord de  $\Omega$  est une courbe (lorsque  $d = 2$ ) ou une surface (pour  $d = 3$ ) suffisamment *régulière* (ou *lisse*). Dans le cas où  $\Omega$  est moins régulier, la notion de solution doit être affaiblie pour donner un sens à l'EDP. Toutefois, nous ne rentrerons pas dans ces considérations ici. Retenons que, si  $f$  et  $\Omega$  sont de classe  $\mathcal{C}^\infty$ , il en sera de même de la solution  $u$ .

### 3.2.2 L'équation de la chaleur

Considérons l'équation de la chaleur en dimension 1 (le paramètre  $\sigma > 0$  mesure la conductivité thermique) :

$$\frac{\partial u}{\partial t}(x, t) - \sigma \frac{\partial^2 u}{\partial x^2}(x, t) = 0, \tag{3.1}$$

pour  $x \in ]0, 1[$  et  $t > 0$ , à laquelle on adjoint une condition initiale

$$u(x, 0) = u_0(x), \tag{3.2}$$

et des conditions aux limites

$$u(0, t) = u(1, t) = 0. \tag{3.3}$$



On peut, ici encore, donner une solution explicite (sous forme d'une série, toutefois) à l'aide de l'analyse de Fourier. En effet, notons  $\tilde{u}_0$  la fonction définie sur  $\mathbb{R}$ , 2-périodique et impaire, qui coïncide avec  $u_0$  sur  $]0, 1[$ . Si  $u_0$  est dans  $L^2(]0, 1[$ , on a le développement en série de Fourier

$$\tilde{u}_0(x) = \sum_{n \geq 1} b_n \sin(n\pi x) \quad \text{dans } L^2.$$

### **Théorème 2**

Le problème (3.1)–(3.2)–(3.3) admet une unique solution, donnée par

$$u(x, t) = \sum_{n \geq 1} b_n e^{-n^2 \pi^2 \sigma t} \sin(n\pi x). \quad (3.4)$$

L'analyse de la convergence normale de la série et des séries dérivées permet de montrer que la fonction  $x \mapsto u(x, t)$  ainsi obtenue est de classe  $\mathcal{C}^\infty$  dès que  $t > 0$ , même si  $u_0$  n'est pas régulière (c'est l'effet *régularisant* de l'équation).

Indiquons que le même genre d'analyse est possible en dimension supérieure, les fonctions sinus étant remplacées par des fonctions plus complexes (fonctions propres du Laplacien dans  $\Omega$ , dont on ne connaît pas de formule explicite).

### **Remarque 2**

Enfin, insistons sur le choix du signe «  $-$  » dans l'équation (3.1). S'il est modifié en «  $+$  » (ou, de manière équivalente, si l'on considère  $t < 0$ ), le problème change de nature puisque l'effet régularisant joue dans l'autre sens. En effet, la série (3.4) ne converge en aucun sens pour  $t > 0$  en raison du terme exponentiellement croissant. On ne considère donc jamais l'équation  $\partial_t u + \partial_x^2 u = 0$  pour  $t > 0$ , ou l'équation  $\partial_t u - \partial_x^2 u = 0$  pour  $t < 0$ .

## **3.2.3 L'équation des ondes et le transport linéaire**

### **L'équation de transport**

On considère l'équation

$$\frac{\partial u}{\partial t}(x, t) - c \frac{\partial u}{\partial x}(x, t) = 0, \quad (3.5)$$

munie de la condition initiale  $v(x, 0) = u_0(x)$ . On ne se préoccupe par ici des conditions aux limites, on peut faire l'hypothèse que l'équation est écrite pour  $x \in \Omega = \mathbb{R}$ . La résolution est immédiate :  $u(x, t) = u_0(x + ct)$ . La donnée initiale  $u_0$  est donc *transportée* à vitesse  $c$  au cours du temps, par simple translation.

Lorsque la vitesse n'est plus constante, mais est une fonction  $(x, t) \mapsto c(x, t)$ , la résolution est encore possible, mais pas aussi explicite. Introduisons en effet les *caractéristiques*  $t \mapsto X_{x_0, t_0}(t)$  définies par

$$X'_{x_0, t_0}(t) = c(X_{x_0, t_0}(t), t) \quad \text{avec} \quad X_{x_0, t_0}(t_0) = x_0. \quad (3.6)$$

Un simple calcul montre alors que, pour toute solution  $u$  de (3.5) satisfait

$$\forall t, \quad \frac{d}{dt} [u(X_{x_0, t_0}(t), t)] = 0.$$

Ainsi,  $u$  est constante le long des caractéristiques.

### **Théorème 3**

Soient  $x_0$  et  $t_0$  fixés. La solution de l'équation de transport est donnée par

$$u(x_0, t_0) = u_0(X_{x_0, t_0}(0)).$$

### **L'équation des ondes**

L'équation hyperbolique modèle est l'équation des ondes (où  $c$  désigne une vitesse)

$$\frac{\partial^2 u}{\partial t^2}(x, t) - c^2 \frac{\partial^2 u}{\partial x^2}(x, t) = 0, \quad (3.7)$$

pour  $x \in \mathbb{R}$  et  $t > 0$ , à laquelle on adjoint des conditions initiales (deux car le problème est d'ordre 2 en temps)

$$u(x, 0) = u_0(x), \quad \frac{\partial u}{\partial t}(x, 0) = v_0(x). \quad (3.8)$$

On n'impose pas de condition aux limites car le problème est ici posé sur  $\mathbb{R}$ . En posant  $w(x, t) = \partial_t u(x, t) + c \partial_x u(x, t)$ , on obtient le système

$$\begin{cases} \frac{\partial w}{\partial t}(x, t) - c \frac{\partial w}{\partial x}(x, t) = 0, \\ \frac{\partial u}{\partial t}(x, t) + c \frac{\partial u}{\partial x}(x, t) = w. \end{cases}$$

On a ainsi découpé l'opérateur des ondes, d'ordre 2, en deux opérateurs de transport (à vitesse constante) d'ordre 1.

### **Théorème 4 (Formule de d'Alembert)**

Si  $u_0 \in \mathcal{C}^2(\mathbb{R})$  et  $u_1 \in \mathcal{C}^1(\mathbb{R})$ , alors le problème (3.7)–(3.8) posé sur  $\Omega = \mathbb{R}$  admet une unique solution donnée par

$$u(x, t) = \frac{1}{2} (u_0(x + ct) + u_0(x - ct)) + \frac{1}{2c} \int_{x-ct}^{x+ct} u_1(s) ds.$$

En l'absence de vitesse initiale (i.e.  $v_0 = 0$ ), la solution  $u$  est la somme de deux ondes qui se propagent en sens inverse à vitesses  $\pm c$ .

### **Remarque 3**

Si l'on considère  $\Omega = ]0, 1[$ , en ajoutant les conditions aux limites  $u(0, t) = u(1, t) = 0$ , il est possible de résoudre le problème à l'aide des séries de Fourier. La solution est donnée par

$$u(x, t) = \sum_{n \geq 1} b_n \cos(n\pi ct) \sin(n\pi x) + \frac{b'_n}{n\pi c} \sin(n\pi ct) \sin(n\pi x),$$

où  $b_n$  et  $b'_n$  désignent les coefficients dans les développements en sinus de  $u_0$  et  $v_0$ , respectivement. À l'aide de formules de trigonométrie élémentaires, on peut retrouver une expression faisant intervenir les variables  $x \pm ct$ , comme dans la formule de d'Alembert.

## 3.3 Méthodes numériques pour les problèmes elliptiques

### 3.3.1 En dimension 1

On considère le problème suivant :

$$\begin{cases} -u''(x) + k(x)u(x) = f(x) & \text{pour } x \in ]0,1[, \\ u(0) = u(1) = 0. \end{cases} \quad (3.9)$$

On a ajouté le terme d'ordre 0, à coefficient non constant, afin que la résolution explicite ne soit pas possible. Pour assurer l'existence et l'unicité d'une solution, on suppose la fonction  $k$  positive.

#### Méthode de tirs

Pour  $\alpha \in \mathbb{R}$  fixé, on introduit le problème

$$\begin{cases} -u''(x) + k(x)u(x) = f(x) & \text{pour } x \in ]0,1[, \\ u(0) = 0, \quad u'(0) = \alpha. \end{cases} \quad (3.10)$$

Il s'agit d'un problème *de Cauchy*, qui diffère du problème (3.9) en ceci qu'il possède des conditions *initiales* et non *aux limites*. Il est donc redevable de la théorie des équations différentielles ordinaires. Si les fonctions  $f$  et  $k$  sont supposées continues, le théorème de Cauchy-Lipschitz assure l'existence et l'unicité d'une solution à ce problème linéaire, notée  $u_\alpha$ .

Il reste à voir s'il est possible de trouver  $\alpha$  tel que  $u_\alpha(1) = 0$ . Pour cela, on remarque que l'application  $\alpha \mapsto u_\alpha(1)$  est affine (en vertu de la linéarité de l'équation). Par ailleurs, elle n'est pas constante : en effet, si  $u_\alpha(1) = u_\beta(1)$ , alors la différence  $v = u_\alpha - u_\beta$  satisfait  $-v'' + kv = 0$  et  $v(0) = 0, v(1) = 0$ . On peut alors écrire à l'aide d'une intégration par parties

$$0 = \int_0^1 (-v'' + kv)v = \int_0^1 v'^2 + kv^2,$$

d'où  $v' = 0$  et donc  $v = 0$  car  $v(0) = 0$ .

En conclusion, il existe un unique  $\alpha^*$  tel que  $u_{\alpha^*}$  satisfasse le problème (3.9). Les arguments théoriques précédents permettent de mettre en œuvre une méthode numérique, dite *méthode de tirs* :

- ◇ on calcule  $u_\alpha(1)$  pour deux valeurs distinctes de  $\alpha$ .
- ◇ on détermine  $\alpha^*$ .
- ◇ on calcule  $u_{\alpha^*}$ .

Pour la résolution du problème (3.10) (nécessaire 3 fois), on utilise une méthode numérique pour les EDO (Euler, RK4, ...).

#### Remarque 4

Cette méthode est spécifique de la dimension 1 car on fait jouer à la variable d'espace le rôle d'une variable de temps. Elle ne se généralise pas en dimension supérieure, contrairement à la méthode de différences finies décrite dans le paragraphe suivant.

## Méthode de différences finies

On discrétise l'intervalle  $]0, 1[$  :  $J$  étant donné, on note

$$\Delta x = \frac{1}{J+1} \quad \text{et} \quad x_j = j\Delta x \quad (0, 1, \dots, J+1).$$

On va construire un vecteur  $(u_0, u_1, \dots, u_{J+1})$  dont on espère qu'il approchera la solution  $u$  de l'équation aux points  $x_0, x_1, \dots, x_{J+1}$ .

Au vu des conditions aux limites, le choix  $u_0 = u_{J+1} = 0$  s'impose. Pour  $1 \leq j \leq J$ , on écrit

$$-u''(x_j) \simeq \frac{-u(x_{j-1}) + 2u(x_j) - u(x_{j+1}))}{\Delta x^2},$$

si bien qu'il est naturel d'écrire

$$\frac{-u_{j-1} + 2u_j - u_{j+1}}{\Delta x^2} + k(x_j)u_j = f(x_j).$$

Avec les notations  $k_j = k(x_j)$ ,  $f_j = f(x_j)$  et

$$A = \begin{bmatrix} 2 & -1 & 0 & \cdots & 0 \\ -1 & 2 & -1 & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & -1 & 2 & -1 \\ 0 & \cdots & 0 & -1 & 2 \end{bmatrix} \in \mathbb{R}^{J \times J}, \quad K = \begin{bmatrix} k_1 & & & & \\ & k_2 & & & \\ & & \ddots & & \\ & & & \ddots & \\ & & & & k_{J+1} \end{bmatrix} \in \mathbb{R}^{J \times J},$$

$$U = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_{J+1} \end{bmatrix} \in \mathbb{R}^J, \quad F = \begin{bmatrix} f_1 \\ f_2 \\ \vdots \\ f_{J+1} \end{bmatrix} \in \mathbb{R}^J$$

le problème discret s'écrit sous la forme du système linéaire  $\left(\frac{1}{\Delta x^2}A + K\right)U = F$ .

### Remarque 5

*Insistons sur le fait que les conditions aux limites  $u_0 = u_{J+1} = 0$  sont incorporées dans l'écriture matricielle précédente, au travers des première et dernière ligne de la matrice  $A$ . Si les conditions sont remplacées par  $u(0) = \alpha$  et  $u(1) = \beta$ , la matrice est inchangée mais le second membre doit être modifié en remplaçant le premier et le dernier coefficients par*

$$f_1 + \frac{\alpha}{\Delta x^2} \quad \text{et} \quad f_{J+1} + \frac{\beta}{\Delta x^2}.$$

*La prise en compte de conditions de Neumann  $u'(0) = u'(1) = 0$  est plus délicate. Une manière simple (mais qui se révèle peu précise) consiste à poser  $u_0 = u_1$  et  $u_J = u_{J+1}$ ; dans la matrice les coefficients 2 dans les première et dernière ligne sont remplacés par des 1.*

Il est possible de montrer l'estimation d'erreur suivante (on renvoie à [1] pour la preuve de ce résultat).

### **Théorème 5**

On suppose que la solution satisfait  $u \in \mathcal{C}^4([0, 1])$ , ce qui est réalisé lorsque les fonctions  $f$  et  $k$  sont de classe  $\mathcal{C}^2[0, 1]$ . Alors

$$\max_{0 \leq j \leq J+2} |u_j - u(x_j)| \leq \frac{\Delta x^2}{96} \sup_{x \in [0, 1]} |u^{(4)}(x)|$$

### **3.3.2 En dimension supérieure**

On considère le problème de Laplace :

$$-\Delta u = f, \quad \text{pour } \mathbf{x} \in \Omega, \quad (3.11)$$

avec la condition aux limites de Dirichlet  $u = 0$  sur  $\partial\Omega$ .

La méthode utilisée précédemment s'applique en dimension  $d \geq 2$ , mais reste bien adapté à des géométries cartésiennes. Si  $\Omega = ]0, a[ \times ]0, b[ \in \mathbb{R}^2$  est un carré, on introduit  $K$  et  $L$  et on les pas d'espace

$$\Delta x = \frac{a}{K+1} \quad \text{et} \quad \Delta y = \frac{b}{L+1}$$

et la subdivision pour  $0 \leq k \leq K+1$  et  $0 \leq \ell \leq L+1$

$$x_k = k\Delta x, \quad y_\ell = \ell\Delta y.$$

Le Laplacien est alors discrétisé sur la grille comme ceci pour  $1 \leq k \leq K$  et  $1 \leq \ell \leq L$

$$-\Delta u(x_k, y_\ell) \simeq \frac{-u(x_{k-1}, y_\ell) + 2u(x_k, y_\ell) - u(x_{k+1}, y_\ell)}{\Delta x^2} + \frac{-u(x_k, y_{\ell-1}) + 2u(x_k, y_\ell) - u(x_k, y_{\ell+1})}{\Delta y^2}.$$

Il faut ensuite numéroté les points de la grille  $\mathbf{x}_j = (x_k, y_\ell)$  pour écrire un système linéaire comme plus haut. Si l'on choisit l'ordre lexicographique, la matrice est pentadiagonale.

### **3.3.3 Un cas non-linéaire**

Soit le problème

$$-\Delta u + u^3 = f, \quad \text{dans } \Omega.$$

avec conditions de Dirichlet homogènes  $u = 0$  sur  $\partial\Omega$ . Pour l'approximation numérique, note  $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_d$  la subdivision de  $\Omega$ ,  $A_2$  la matrice du Laplacien-Dirichlet en dimension 2 (cf. paragraphe précédent), et on introduit la fonction  $\Phi : \mathbb{R}^J \rightarrow \mathbb{R}^J$  définie par

$$\Phi : U \mapsto A_2 U + U^3 - F,$$

où  $U^3$  représente le vecteur dont les composantes sont les cubes de celles de  $U$ . Il s'agit donc de trouver le zéro de la fonction  $\Phi$ . Pour cela, on peut utiliser la méthode de Newton, qui s'écrit

$$U^{n+1} = U^n - [J\Phi(U^n)]^{-1} \Phi(U^n),$$

où la matrice jacobienne de  $\Phi$  est donnée par

$$J\Phi(U) = A_2 + 3D(U^2),$$

avec  $D(U^2)$  la matrice diagonale dont les entrées diagonales sont les carrés des composantes de  $U$ .

### 3.4 Problèmes d'évolution paraboliques

#### 3.4.1 L'équation de la chaleur

On considère le problème mono-dimensionnel

$$\frac{\partial u}{\partial t} - \sigma \frac{\partial^2 u}{\partial x^2} = 0 \quad \text{pour } x \in ]0, 1[ \quad \text{et } t > 0,$$

avec la condition initiale  $u(x, 0) = u_0(x)$  et les conditions aux limites  $u(0, t) = u(1, t) = 0$ . Comme plus haut, on introduit le pas d'espace  $\Delta x = 1/(J+2)$  et la subdivision  $x_j = j\Delta x$ . En notant  $U(t)$  le vecteur des approximations construites par remplacement de l'opérateur  $-\partial_x^2$  par la multiplication par la matrice  $A$  (cf. § 3.3.1), il vient

$$U'(t) + \frac{\sigma}{\Delta x^2} AU(t) = 0.$$

Il s'agit d'un système d'équations différentielles en dimension  $J$ , pour lequel les méthodes vues dans le chapitre précédent sont disponibles. Concentrons-nous sur les plus simples : en notant  $\Delta t$  le pas de temps, et  $U^n$  l'approximation construite pour  $U(n\Delta t)$ ,

◇ *Euler explicite :*

$$U^{n+1} = \left( Id - \sigma \frac{\Delta t}{\Delta x^2} A \right) U^n,$$

◇ *Euler implicite :*

$$U^{n+1} = \left( Id + \sigma \frac{\Delta t}{\Delta x^2} A \right)^{-1} U^n,$$

Dans la suite, on pose  $\beta = \sigma \frac{\Delta t}{\Delta x^2} > 0$ . D'après l'expression sous forme de série de Fourier (cf. Théorème 2), la solution  $u$  est bornée (et tend même vers 0 lorsque  $t \rightarrow +\infty$ ). Il est raisonnable de souhaiter qu'il en soit de même pour la solution discrète. Il s'agit donc de savoir sous quelle condition la matrice  $(B^n)$  est bornée pour  $n \in \mathbb{N}$ , lorsque la matrice  $B$  vaut  $Id - \beta A$  ou  $(Id + \beta A)^{-1}$ . Cette condition s'écrit en fonction du rayon spectral :  $\rho(B) < 1$ .

↪ *Estimation du rayon spectral par une norme subordonnée.*

Partant de l'inégalité  $\rho(B) \leq \|B\|$  pour toute norme subordonnée, on peut calculer la norme  $\|B\|_\infty$  :

$$\|B\|_\infty = \max_i \sum_j |B_{ij}|$$

Pour  $B = Id - \beta A$ , on obtient  $\|B\|_\infty = |1 - 2\beta| + 2\beta$  qui est inférieur strictement à 1 si et seulement si  $2\beta \leq 1$ .

Pour  $B = (Id + \beta A)^{-1}$ , il n'est pas facile de calculer cette norme.

↪ *Estimation du rayon spectral à l'aide des valeurs propres.*

Il est facile de montrer que la matrice  $A$  est définie positive. Ainsi, les valeurs propres de  $Id + \beta A$  sont strictement supérieures à 1 et celles de  $B = (Id + \beta A)^{-1}$  sont strictement inférieures à 1. On en déduit immédiatement que  $\rho(B) < 1$ .

Pour  $B = Id - \beta A$ , si la plus grande valeur propre est strictement inférieure à 1 (car  $A$  est définie positive), ce n'est pas évident pour la plus petite. Un calcul explicite montre que les valeurs propres de  $A$  sont données par

$$\lambda_k = 4 \sin^2 \left( \frac{k\pi}{2(J+1)} \right) \quad (1, \leq k \leq J),$$

si bien que celles de  $B$  valent  $1 - \beta\lambda_k$ . La plus petite est strictement supérieure à  $-1$  sous la condition  $\beta \leq \frac{1}{2}$ .

En conclusion, il apparaît que la suite  $(U^n)$  construite à l'aide de la méthode d'Euler implicite est bornée sans condition, alors que la suite  $(U^n)$  correspondant à la méthode d'Euler explicite est bornée sous la condition  $2\beta \leq 1$ , soit

$$\sigma \frac{\Delta t}{\Delta x^2} \leq \frac{1}{2}.$$

Cette condition, dite CFL (Courant-Friedrichs-Lévy), contraint fortement le choix du pas de temps.

### 3.4.2 Les problèmes de réaction-diffusion

Une extension naturelle des méthodes exposées précédemment concerne les équations non-linéaires, dites de *réaction-diffusion*, qui s'écrivent

$$\frac{\partial u}{\partial t} - \sigma \Delta u = f(u),$$

où la non-linéarité  $f : \mathbb{R} \rightarrow \mathbb{R}$  correspond aux aspects *réactifs*. Comme on l'a vu, il est préférable d'utiliser une méthode implicite pour la partie diffusie. Le terme de réaction, quant à lui, peut-être approché de manière explicite ou implicite :

◇ *Schéma semi-implicite (Diffusion-implicite, réaction-explicite)*

$$U^{n+1} = U^n - \sigma \frac{\Delta t}{\Delta x^2} A U^{n+1} + \Delta t f[U^n].$$

◇ *Schéma totalement implicite*

$$U^{n+1} = U^n - \sigma \frac{\Delta t}{\Delta x^2} A U^{n+1} + \Delta t f[U^{n+1}]$$

Dans les formules précédentes, la notation  $f[U]$  désigne le vecteur de composantes  $f(u_j)$ . Le schéma semi-implicite nécessite la résolution d'un système linéaire à chaque itération, le schéma totalement implicite, quant à lui, requiert la résolution d'un système d'équations non-linéaires. Le surcoût lié à la non-linéarité fait qu'on réserve la méthode totalement implicite aux problèmes raides, cf. §2.5.1. Mentionnons enfin les méthodes de *splitting*, qui proposent une résolution alternée de l'équation de la chaleur, d'une part, et de l'équation différentielle ordinaire  $u' = f(u)$ , d'autre part (permettant l'utilisation de méthodes performantes pour chacune d'entre-elles, qui ne seraient pas opérantes sur l'équation globale).

## 3.5 Phénomènes de propagation

### 3.5.1 L'équation de transport linéaire en une dimension

On considère le problème suivant, posé pour  $x$  dans  $\mathbb{R}$  (ce qui explique l'absence de conditions aux limites) :

$$\frac{\partial u}{\partial t}(x, t) - c \frac{\partial u}{\partial x}(x, t) = 0,$$

avec la condition initiale  $u(x, 0) = u_0(x)$ . On a vu au paragraphe 3.2.3 que la solution est constante le long des caractéristiques qui sont ici des droites :

$$u(x, t) = u_0(x + ct).$$

Pour l'approximation numérique, on introduit deux pas de discrétisation  $\Delta x$  et  $\Delta t$ . On propose les schémas suivants, où l'on souhaite que  $u_j^n$  approche  $u(j\Delta x, n\Delta t)$  – on a noté dans la suite  $\beta = c \frac{\Delta t}{\Delta x}$ .

◇ Schéma centré

$$\frac{u_j^{n+1} - u_j^n}{\Delta t} - c \frac{u_{j+1}^n - u_{j-1}^n}{2\Delta x} = 0 \quad \text{soit} \quad u_j^{n+1} = u_j^n + \frac{\beta}{2} u_{j+1}^n - \frac{\beta}{2} u_{j-1}^n.$$

◇ Schéma décentré à gauche

$$\frac{u_j^{n+1} - u_j^n}{\Delta t} - c \frac{u_j^n - u_{j-1}^n}{\Delta x} = 0 \quad \text{soit} \quad u_j^{n+1} = (1 + \beta)u_j^n - \beta u_{j-1}^n.$$

◇ Schéma décentré à droite

$$\frac{u_j^{n+1} - u_j^n}{\Delta t} - c \frac{u_{j+1}^n - u_j^n}{\Delta x} = 0 \quad \text{soit} \quad u_j^{n+1} = (1 - \beta)u_j^n + \beta u_{j+1}^n.$$

Dans les trois cas,  $u_j^{n+1}$  apparaît comme barycentre des  $(u_\ell^n)$  car la somme des coefficients vaut 1. Si l'on suppose que la condition initiale  $u_0$  appartient à un intervalle  $[a, b]$ , imposer au schéma de respecter cette hypothèse à chaque étape  $n$  revient à exiger aux coefficients d'être positifs (au sens large). Dans le cas où  $c > 0$ , on a  $\beta > 0$  et seul le schéma décentré à droite respecte ce cahier des charges sous la condition

$$c \frac{\Delta t}{\Delta x} \leq 1. \tag{3.12}$$

Toujours sous cette condition, le schéma décentré à gauche est adapté à des vitesses  $c$  négatives. Quant au schéma centré (fournissant une erreur de consistance pourtant plus faible), il est instable au sens de ce critère pour tous choix des pas de temps et d'espace.

La condition (3.12) peut être interprétée à l'aide du schéma de la figure 3.1. La stabilité est reliée au fait que la caractéristique exacte doit être située à l'intérieur du domaine de dépendance numérique.

### 3.5.2 L'équation des ondes

Soit le problème

$$\frac{\partial^2 u}{\partial t^2}(x, t) - c^2 \frac{\partial^2 u}{\partial x^2}(x, t) = 0, \tag{3.13}$$



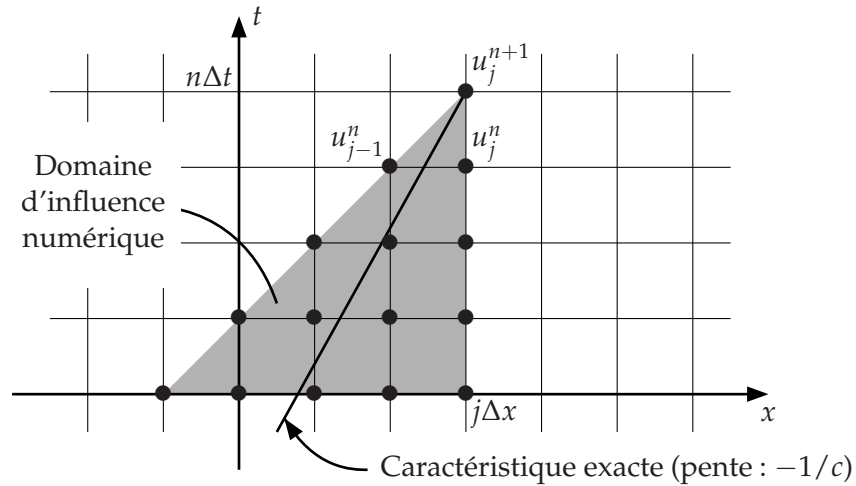


FIGURE 3.1 – Interprétation de la condition CFL (3.12) pour  $c < 0$ .

pour  $x \in \mathbb{R}$  et  $t > 0$ , avec les conditions initiales

$$u(x, 0) = u_0(x), \quad \frac{\partial u}{\partial t}(x, 0) = v_0(x). \quad (3.14)$$

Au paragraphe 3.2.3, on a vu que l'équation des ondes pouvait être ramenée à un système d'équations de transport. Cette écriture peut être utilisée à des fins de calcul numérique. On ne donne pas de détails ici – voir [3] par exemple – mais on mentionne simplement le schéma de Newmark, stable et consistant d'ordre 2 en temps et en espace :

$$\frac{u_j^{n+1} - 2u_j^n + u_j^{n-1}}{\Delta t^2} - c^2 \frac{u_{j-1}^n - 2u_j^n + u_{j+1}^n}{\Delta x^2} = 0,$$

soit

$$u_j^{n+1} = 2(1 - \beta^2)u_j^n - u_j^{n-1} + \beta^2 u_{j-1}^n + \beta^2 u_{j+1}^n.$$

Il s'agit d'un schéma à deux pas, pour lequel l'initialisation n'est pas fournie directement par les conditions initiales. En effet, si la condition  $u(x, 0) = u_0(x)$  donne naturellement

$$u_j^0 = u_0(j\Delta x),$$

la détermination de  $(u_j^1)_j$ , on peut écrire simplement, suivant un développement de Taylor à l'ordre 1 :

$$u_j^1 = u_j^0 + \Delta t v_0(j\Delta x),$$

mais l'erreur commise détériore l'ordre global de la méthode. En utilisant un développement à l'ordre 2, on peut proposer l'initialisation suivante :

$$u_j^1 = u_j^0 + \Delta t v_0(j\Delta x) + \frac{c^2 \Delta t^2}{2\Delta x^2} \left( -u_{j-1}^0 + 2u_j^0 - u_{j+1}^0 \right).$$

### Remarque 6

*Si le problème est posé sur un intervalle borné (ou semi-borné), la prise en compte des conditions aux limites s'effectue comme dans le cas elliptique.*

# Références

- [1] Philippe G. Ciarlet. *Introduction à l'analyse numérique matricielle et à l'optimisation*. Collection Mathématiques Appliquées pour la Maîtrise. [Collection of Applied Mathematics for the Master's Degree]. Masson, Paris, 1982.
- [2] Michel Crouzeix and Alain L. Mignot. *Analyse numérique des équations différentielles*. Collection mathématiques appliquées pour la maîtrise. Masson, Paris, 1984.
- [3] Florence Hubert and John Hubbard. *Calcul scientifique*. Vuibert, 2006.
- [4] Alfio Quarteroni, Riccardo Sacco, and Fausto Saleri. *Numerical mathematics*, volume 37 of *Texts in Applied Mathematics*. Springer-Verlag, New York, 2000.
- [5] J. Stoer and R. Bulirsch. *Introduction to numerical analysis*, volume 12 of *Texts in Applied Mathematics*. Springer-Verlag, New York, third edition, 2002. Translated from the German by R. Bartels, W. Gautschi and C. Witzgall.