

Chapitre 2

Equations Hyperboliques

Le but de ce chapitre est d'étudier aussi bien du point de vue théorique que numérique l'équation de transport :

$$(2.1) \quad \frac{\partial u}{\partial t} + c \frac{\partial u}{\partial x} = 0 \quad \text{dans } \mathbb{R} \times (0, T),$$

où $T > 0$ et $c \in \mathbb{R}$ sont des constantes données et $u : \mathbb{R} \times (0, T) \rightarrow \mathbb{R}$ est la fonction inconnue. Cette équation est associée à la donnée initiale :

$$(2.2) \quad u(x, 0) = u_0(x) \quad \text{dans } \mathbb{R},$$

où u_0 est une fonction donnée qui est au moins continue.

2.1 L'étude théorique

Le premier résultat montre que l'on a une solution explicite si u_0 est de classe \mathcal{C}^1 .

Proposition 2.1. *Si $u_0 \in \mathcal{C}^1(\mathbb{R})$, la fonction u définie par :*

$$(2.3) \quad u(x, t) = u_0(x - ct), \quad \text{pour } x \in \mathbb{R}, t \in [0, T],$$

est solution de (2.1)-(2.2).

Preuve : On utilise la méthode des caractéristiques qui consiste à supposer qu'il existe une solution u de classe \mathcal{C}^1 de (2.1)-(2.2) et à chercher des courbes $t \mapsto (x(t), t)$ le long desquelles u est constant, *i.e.*

$$u(x(t), t) = c^{\text{te}}.$$

On dérive cette propriété par rapport à t :

$$\left(\frac{\partial u}{\partial t} + \dot{x}(t) \frac{\partial u}{\partial x} \right) (x(t), t) = 0$$

et en identifiant avec l'équation, on trouve $\dot{x}(t) = c$. Donc, si $y \in \mathbb{R}$ et $t \in (0, T)$,

$$u(y + ct, t) = c^{\text{te}} = u_0(y),$$

où la seconde égalité est obtenue en prenant $t = 0$. En inversant cette formule, c'est-à-dire en prenant $x = y + ct$, on trouve

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

Réciproquement, on vérifie trivialement que u donné par cette formule est solution de l'équation. \square

Plus généralement, la méthode dite "des caractéristiques" s'applique aux équations où c dépend de x :

$$(2.4) \quad \frac{\partial u}{\partial t} + c(x) \frac{\partial u}{\partial x} = 0 \quad \text{dans } \mathbb{R} \times (0, T).$$

On cherche ici encore des courbes $(x(t), t)$ le long desquelles u est constant, *i.e.*

$$u(x(t), t) = c^{\text{te}},$$

en dérivant par rapport à t on obtient

$$\left(\frac{\partial u}{\partial t} + \dot{x}(t) \frac{\partial u}{\partial x} \right) (x(t), t) = 0,$$

et en identifiant on arrive cette fois à :

$$(2.5) \quad \dot{x}(t) = c(x(t)).$$

Les trajectoires $x(\cdot)$ sont donc solutions d'une équation différentielle ordinaire (EDO). Pour que cette EDO ait une solution définie pour tout $t \in (0, T)$, on va supposer que la fonction c est localement lipschitzien (ce qui entraîne l'existence locale), et que c est borné (ce qui entraîne l'existence globale puisqu'il n'y a pas d'explosion en temps fini). Dans ce cas :

$$u(x(t), t) = u_0(x(0)).$$

Pour inverser cette formule, il est clair qu'il faut résoudre "backward" ⁽¹⁾ l'EDO, c'est-à-dire introduire la solution avec donnée finale en $s = t$ de :

$$\begin{cases} \dot{x}(s) = c(x(s)), & s \in [0, t] \\ x(t) = x \in \mathbb{R} \end{cases}$$

et trouver la valeur initiale $x(0)$. On aura alors :

$$u(x, t) = u_0(x(0)).$$

⁽¹⁾de manière rétrograde in french

Cette formule est évidemment moins explicite que dans le cas où c est constant (puisque dans ce cas on a $x(0) = x - ct$), mais elle permet tout de même de calculer u dans certains cas.

Notre but est maintenant double : d'une part donner un sens à la solution $u(x, t) = u_0(x - ct)$ si u_0 est seulement continue, car alors u n'est plus nécessairement de classe \mathcal{C}^1 ; d'autre part étudier l'unicité d'une telle solution.

Pour cela, on introduit l'espace $\mathcal{C}_0^1(\mathbb{R} \times [0, T])$ des fonctions de classe \mathcal{C}^1 à support compact dans $\mathbb{R} \times [0, T]$.

Définition 2.1. Une fonction $u \in L^\infty(\mathbb{R} \times [0, T])$ est dite solution faible de (2.1)-(2.2) si et seulement si

$$(2.6) \quad \int_0^T \int_{\mathbb{R}} u \left(\frac{\partial \phi}{\partial t} + c \frac{\partial \phi}{\partial x} \right) dx dt + \int_{\mathbb{R}} u_0(x) \phi(x, 0) dx = 0,$$

pour toute fonction $\phi \in \mathcal{C}_0^1(\mathbb{R} \times [0, T])$.

Remarque : Il s'agit essentiellement d'une solution au sens des distributions, mis à part le deuxième terme qui prend en compte la donnée initiale.

On a le résultat suivant :

Théorème 2.1. (i) Si $u_0 \in \mathcal{C}^1(\mathbb{R})$, alors la fonction u définie par

$$(2.7) \quad u(x, t) = u_0(x - ct) \quad \text{pour } x \in \mathbb{R}, t \in [0, T[$$

est solution faible de (2.1)-(2.2).

(ii) Si $u_0 \in \mathcal{C}(\mathbb{R})$, la fonction définie par (2.7) est solution faible de (2.1)-(2.2).

(iii) Si $u \in L^\infty(\mathbb{R} \times [0, T])$ est solution faible de (2.1) associée à la donnée initiale u_0 et v est solution faible de (2.1) associée à la donnée initiale v_0 , alors, pour tout $R > 0$ et $ct \leq R$:

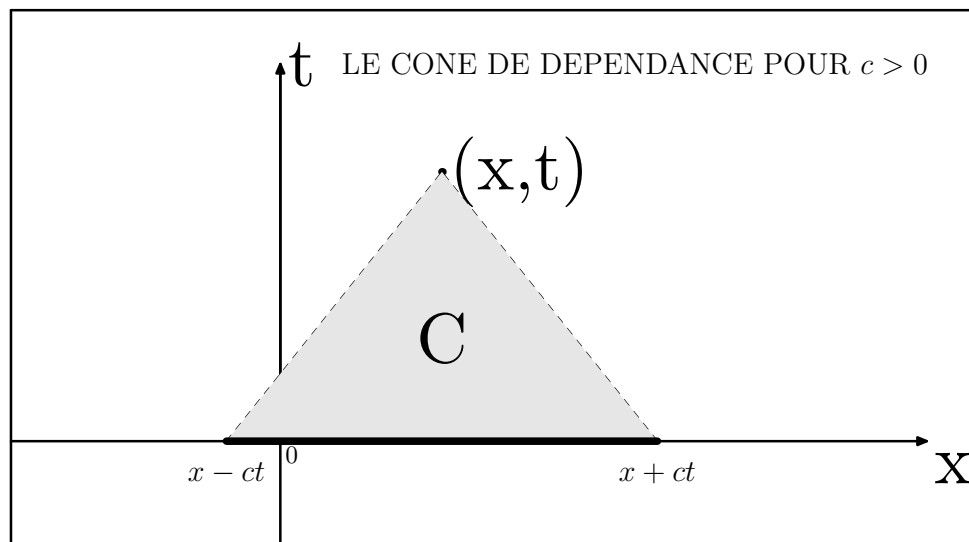
$$u_0 = v_0 \text{ dans } [-R, R] \Rightarrow u(x, t) = v(x, t) \text{ dès que } |x| \leq R - ct.$$

Avant de passer à la preuve de ce théorème, quelques remarques : (i) signifie qu'une solution "forte" est toujours solution "faible", ce qui semble naturel. (ii) sera une conséquence du fait qu'une limite de solutions faibles - limite localement uniforme dans le cas présent - sera encore une solution faible; enfin (iii) est un résultat de type "vitesse finie de propagation" : l'information disant que u_0 et v_0 sont différents (éventuellement) pour $|x| > R$ ne parvient au point $x = 0$ qu'au temps $t = R/c$.

Le cône suivant :

$$(2.8) \quad C_{(x,t)} = \{(y, s) \in \mathbb{R} \times [0, t] : |y - x| \leq c(t - s)\}$$

est appelé “cône de dépendance théorique” de la solution puisque $u(x, t)$ ne dépend que des valeurs de u dans ce cône, et en particulier sur l'intervalle $[x - ct, x + ct]$.



Preuve du théorème 2.1 : La preuve de (i) consiste simplement en une intégration par parties, ce qui ne pose aucun problème puisque u est de classe \mathcal{C}^1 .

Maintenant, si u_0 est seulement continue, il existe une suite $(u_0^\varepsilon)_\varepsilon$ de fonctions de classe \mathcal{C}^1 qui converge localement uniformément vers u_0 dans \mathbb{R} . Si on note u^ε la solution associée à u_0^ε , il est clair que :

$$u^\varepsilon(x, t) = u_0^\varepsilon(x - ct) \rightarrow u(x, t) = u_0(x - ct)$$

uniformément sur tout compact de $\mathbb{R} \times [0, T[$. Pour démontrer que u est solution faible, il suffit de passer à la limite dans l'égalité :

$$(2.9) \quad \int_0^T \int_{\mathbb{R}} u^\varepsilon \left(\frac{\partial \phi}{\partial t} + c \frac{\partial \phi}{\partial x} \right) dx dt + \int_{\mathbb{R}} u_0^\varepsilon(x) \phi(x, 0) dx = 0,$$

ce qui montre (ii).

Pour prouver (iii), on pose $w = u - v$. Si $\phi(x, 0)$ est à support compact dans l'intervalle $] - R, R[$, on a :

$$\int_0^T \int_{\mathbb{R}} w \left(\frac{\partial \phi}{\partial t} + c \frac{\partial \phi}{\partial x} \right) dx dt = 0,$$

puisque $u_0 = v_0$ dans $] - R, R[$. Comme $w \in L^\infty(\mathbb{R} \times [0, T[)$, $w \in L^2_{\text{loc}}(\mathbb{R} \times [0, T[)$; or les fonctions de \mathcal{C}^1_0 sont denses dans L^1_{loc} , ainsi il existe une suite $(w^\varepsilon)_\varepsilon$ de fonctions de classe \mathcal{C}^1 telles que

$$w^\varepsilon \rightarrow w \quad \text{dans} \quad L^2(C),$$

où $C = \{(x, t) \in \mathbb{R} \times [0, T] : |x| < R - ct\}$. On a alors le lemme suivant :

Lemme 2.1. *Pour toute fonction χ de classe C^1 , à support compact dans C , il existe $\phi^\varepsilon \in C_0^1(\mathbb{R} \times [0, T])$ telle que $\phi^\varepsilon(x, 0)$ est à support compact dans $] -R, R[$ et telle que :*

$$\frac{\partial \phi^\varepsilon}{\partial t} + c \frac{\partial \phi^\varepsilon}{\partial x} = \chi w^\varepsilon \quad \text{dans } \mathbb{R} \times [0, T].$$

Poursuivons la démonstration du théorème en supposant le lemme acquis (nous reviendrons sur le lemme juste après). On injecte alors ϕ^ε comme fonction-test pour w et on passe à la limite dans :

$$\int_0^T \int_{\mathbb{R}} w w^\varepsilon \chi \, dx dt = 0,$$

ce qui montre que pour toute fonction χ à support compact dans C ,

$$\int_0^T \int_{\mathbb{R}} w^2 \chi \, dx dt = 0.$$

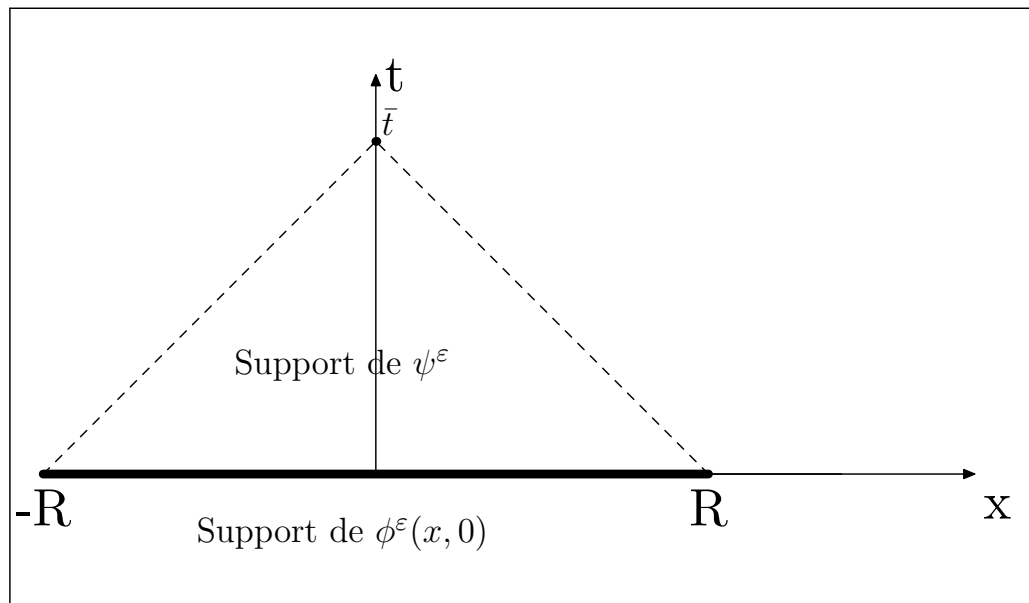
Cela implique clairement que $w^2 = 0$ dans C (prendre $\chi \geq 0$ et $\chi > 0$ dans $|x| < R - \varepsilon - ct$), d'où le résultat (iii). \square

Preuve du Lemme 2.1 : Posons $\psi^\varepsilon = \chi w^\varepsilon$ et utilisons la méthode des caractéristiques mais de manière différente : utilisant $\dot{x}(t) = c$, on calcule :

$$\frac{d}{dt} [\phi^\varepsilon(x(t), t)] = \left(\frac{\partial \phi^\varepsilon}{\partial t} + c \frac{\partial \phi^\varepsilon}{\partial x} \right) (x(t), t) \, dx dt = \psi^\varepsilon(x(t), t)$$

d'où :

$$\phi^\varepsilon(x(t), t) = \phi^\varepsilon(x(0), 0) + \int_0^t \psi^\varepsilon(x(s), s) \, ds$$



et ainsi :

$$\phi^\varepsilon(x, t) = \phi^\varepsilon(x - ct, 0) + \int_0^t \psi^\varepsilon(x - c(t - s), s) ds.$$

Comme ψ^ε est à support compact dans C et $\phi^\varepsilon(\cdot, 0)$ dans $] -R, R[$, le fait que ϕ^ε est à support compact en x est clair. C'est moins évident en temps. On pose alors $\bar{t} = R/c$; il suffit de choisir la donnée initiale ϕ^ε telle que

$$\phi^\varepsilon(x, 0) = - \int_0^{\bar{t}} \psi^\varepsilon(x + cs, s) ds,$$

et il est alors facile de vérifier que $\phi^\varepsilon \in \mathcal{C}_0^1(\mathbb{R} \times [0, T])$ convient. \square

2.2 L'approche numérique

Pour calculer numériquement la solution de (2.1)-(2.2), on introduit une grille qui, cette fois, est en x et en t , en posant :

$$\begin{aligned} x_j &= j\Delta x, & j &\in \mathbb{Z}, \\ t_n &= n\Delta t, & n &\in [0, n] \cap \mathbb{N}, \quad N\Delta t = T, \end{aligned}$$

et on notera u_j^n une approximation de $u(x_j, t_n)$. Il est à remarquer cette fois que $j \in \mathbb{Z}$, ce qui n'est évidemment pas très compatible avec un calcul sur ordinateur! Pour simplifier la présentation, on va donc d'abord présenter des schémas numériques définis pour $j \in \mathbb{Z}$ puis on expliquera comment se ramener à un nombre fini de valeurs de j (cf. section 2.2.6).

2.2.1 Les schémas explicites les plus standard

Dans notre cadre, puisque l'on a affaire à une équation d'évolution, un schéma numérique est un algorithme qui permet de calculer $u^{n+1} = (u_j^{n+1})_j$ en fonction de $u^n = (u_j^n)_j$ puisque la donnée initiale $u^0 = (u_j^0)_j$ est connue.

On distingue alors deux types de schémas : les schémas "explicites" où les u_j^{n+1} s'expriment directement en fonction des u_j^n et les schémas "implicites" où l'on obtient les u_j^{n+1} en résolvant une équation (qui dépend, bien entendu, des u_j^n).

EXEMPLES - Le schéma suivant :

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

est un schéma explicite "décentré avant". Il est explicite car on peut donner explicitement u_j^{n+1} en fonction de u^n :

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

et on dit qu'il est "décentré avant" car l'approximation de $\frac{\partial u}{\partial x}(x_j, t)$ par $\frac{1}{\Delta x}(u_{j+1}^n - u_j^n)$ utilise le point x_{j+1} qui est "en avant" de x_j . De même on a les schémas explicites "décentré arrière" et "centré" suivants :

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

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

On démontre facilement que ces trois schémas sont consistants en faisant des développements de Taylor ; par exemple pour (SEDAv)⁽²⁾ :

$$\begin{aligned} \frac{1}{\Delta t}(u_j^{n+1} - u_j^n) + \frac{c}{\Delta x}(u_{j+1}^n - u_j^n) &= \frac{1}{\Delta t} \int_{t_n}^{t_n+\Delta t} \frac{\partial u}{\partial t}(x_j, s) ds \\ &\quad + \frac{c}{\Delta x} \int_{x_j}^{x_j+\Delta x} \frac{\partial u}{\partial x}(x, t_n) dx. \end{aligned}$$

En intégrant par parties, il vient :

$$\begin{aligned} \frac{1}{\Delta t}(u_j^{n+1} - u_j^n) + \frac{c}{\Delta x}(u_{j+1}^n - u_j^n) &= \left[\frac{1}{\Delta t}(t - t_n - \Delta t) \frac{\partial u}{\partial t}(x_j, s) \right]_{t_n}^{t_n+\Delta t} \\ &\quad - \int_{t_n}^{t_n+\Delta t} \frac{1}{\Delta t}(t - t_n - \Delta t) \frac{\partial^2 u}{\partial t^2}(x_j, s) ds \\ &\quad + \left[\frac{c}{\Delta x}(x - x_j - \Delta x) \frac{\partial u}{\partial x}(x, t_n) \right]_{x_j}^{x_j+\Delta x} \\ &\quad - \int_{x_j}^{x_j+\Delta x} \frac{c}{\Delta x}(x - x_j - \Delta x) \frac{\partial^2 u}{\partial x^2}(x, t_n) dx \\ &= \frac{\partial u}{\partial t}(x_j, t_n) + c \frac{\partial u}{\partial x}(x_j, t_n) + \mathcal{O}(\Delta t) + \mathcal{O}(\Delta x). \end{aligned}$$

Le schéma est donc d'ordre 1 en t et en x , puisque l'erreur théorique commise est en $(\Delta t)^1$ et en $(\Delta x)^1$ quand on remplace l'équation par le schéma.

2.2.2 Étude de la stabilité des principaux schémas explicites

L'étude de la stabilité de ces schémas est faite via une analyse de Fourier ; on teste le schéma sur des données initiales de la forme $u_j^0 = \exp(ikx_j)$ pour $j \in \mathbb{Z}$, c'est-à-dire que l'on prend comme données initiales des ondes de fréquence k . Dans le cas des schémas ci-dessus, il est clair que l'on a :

$$u^{n+1} = g\left(\frac{c\Delta t}{\Delta x}, k\Delta x\right)u^n,$$

⁽²⁾ le lecteur sera peut-être surpris de voir des intégrales et de l'intégration par parties dans un calcul de développement de Taylor ... mais la notion de "formule de Taylor avec reste intégral" ne devrait pas lui être étrangère ! L'autre option, plus classique consiste évidemment à faire des développements limités (exercice : le faire!).

où $g\left(\frac{c\Delta t}{\Delta x}, k\Delta x\right) \in \mathbb{C}$. On posera désormais $\lambda = c\Delta t/\Delta x$; c'est le coefficient de CFL (Courant-Friedrichs-Lévy).

Définition 2.2. *On dira que le schéma est stable si et seulement si :*

$$|g(\lambda, k\Delta x)| \leq 1 \quad \text{pour tout } k.$$

On remarque tout d'abord que si $|g(\lambda, k\Delta x)| \leq 1$, alors $|u_j^{n+1}| \leq |u_j^n| \leq |u_j^0|$ pour tout $j \in \mathbb{Z}$, et qu'ainsi les u_j^n restent bornés indépendamment de n (et évidemment de j , si j reste lui-même borné).

SCHÉMA (SEDAV) - si $u_j^n = \gamma \exp(ik\Delta x)$, alors :

$$u_j^{n+1} = u_j^n \left[1 - \lambda(\exp(ik\Delta x) - 1) \right],$$

donc :

$$g(\lambda, k\Delta x) = 1 - \lambda(\exp(ik\Delta x) - 1).$$

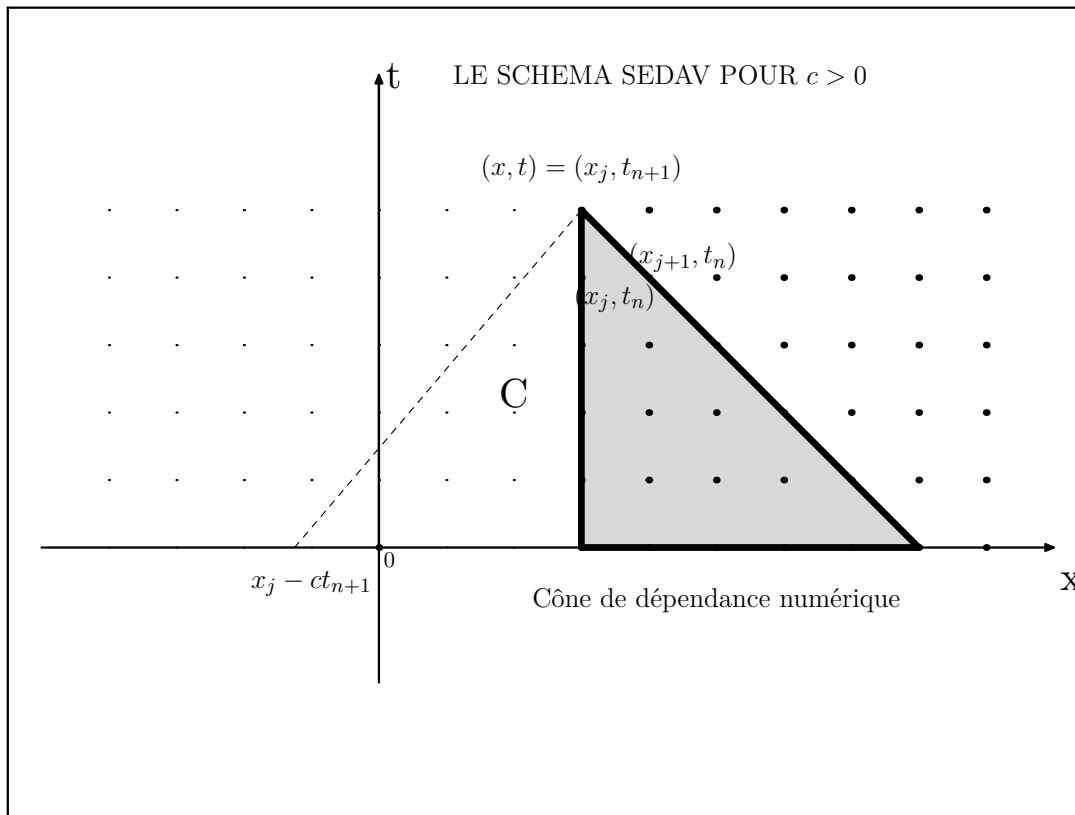
Calculons maintenant :

$$\begin{aligned} |g(\lambda, k\Delta x)|^2 &= \left[1 - \lambda(\cos(k\Delta x) - 1) \right]^2 + \lambda^2 \sin^2(k\Delta x) \\ &= 1 - 2\lambda(\cos(k\Delta x) - 1) + \lambda^2(\cos(k\Delta x) - 1)^2 + \lambda^2 \sin^2(k\Delta x) \\ &= 1 + 4\lambda \sin^2\left(\frac{k\Delta x}{2}\right) + 4\lambda^2 \sin^2\left(\frac{k\Delta x}{2}\right) \\ &= 1 + 4\lambda(1 + \lambda) \sin^2\left(\frac{k\Delta x}{2}\right). \end{aligned}$$

Il est clair en premier lieu que si $\lambda > 0$, le schéma est instable.

Donnons une explication de ce phénomène en considérant l'action du schéma : la valeur de u_j^{n+1} ne dépend que des valeurs de u_ℓ^0 pour $\ell \in (j, j + n + 1)$, alors que la valeur de la vraie solution est $u_0(x_j - ct_{n+1})$. Donc si $c > 0$, ce qui correspond à $\lambda > 0$, le point $x_j - ct_{n+1}$ n'est pas dans le cône de dépendance numérique, d'où l'instabilité (voir la figure ci-dessous). Une règle empirique est la suivante :

“Le cône de dépendance numérique doit contenir le cône de dépendance théorique”



Si maintenant $\lambda < 0$ (ce qui correspond à $c < 0$), alors $|g(\lambda, k\Delta x)| \leq 1$ pour tout k est équivalent à :

$$\lambda = \frac{c\Delta t}{\Delta x} \geq -1,$$

c'est la condition de Courant-Friedrichs-Lévy (CFL)⁽³⁾; il faut que Δt et Δx satisfassent cette propriété pour que le schéma soit stable. Dans ce cas, on vérifie bien (et c'est même équivalent) que le point $x_j - ct_{n+1}$ est situé dans le cône de dépendance numérique.

On a un schéma "diffusif" puisque $|g(\lambda, k\Delta x)| < 1$ pour la plupart des valeurs de λ et k . Ce schéma introduit en général une perte d'amplitude de la solution puisque quand n devient grand, $|g|^n$ tend vers zéro. Alors que pour la solution continue, on n'a pas de perte d'amplitude :

$$\begin{aligned} u(x_j, t_{n+1}) &= u(x_j - c\Delta t, t_n) \\ &= \gamma \exp(ik(x_j - c\Delta t)) \\ &= \gamma \exp(ikx_j) \cdot \exp(-ic\Delta t) \\ &= u(x_j, t_n) \cdot \exp(-ic\Delta t), \end{aligned}$$

⁽³⁾ de façon générale, on appelle "condition de (CFL)" la condition sur λ qui assure la stabilité du schéma - voir plus loin d'autres exemples de condition (CFL).

ce qui donne un coefficient d'amplification théorique :

$$\begin{aligned} G(\lambda, k\Delta x) &= \exp(-ikc\Delta t) \\ &= \exp(-i\lambda k\Delta x), \end{aligned}$$

dont le module est exactement 1. On peut aussi définir l'erreur de phase :

$$E_\phi(\lambda, k\Delta x) = -\left(\arg g(\lambda, k\Delta x) - \arg G(\lambda, k\Delta x) \right),$$

qui décrit, en fait, dans le cas de l'équation continue, l'erreur que l'on commet sur le déplacement des ondes se propageant à vitesse c . L'erreur d'amplitude étant :

$$E_a(\lambda, k\Delta x) = |g(\lambda, k\Delta x)| - |G(\lambda, k\Delta x)|.$$

L'erreur de phase concerne les problèmes dits de "dispersion" ; il est assez souvent difficile de calculer explicitement les erreurs de phase et d'amplitude mais on peut en donner des équivalents lorsque $K = k\Delta x$ est petit par exemple. Dans le cas du schéma (SEDAv), pour λ fixé :

$$E_a(\lambda, K) \underset{K \rightarrow 0}{\sim} \frac{K^2}{2} \lambda(1 + \lambda),$$

et :

$$\begin{aligned} E_\phi(\lambda, K) &= -\left[\arg(1 - \lambda\{\exp(iK) - 1\} - 1) + \lambda K \right] \\ &\underset{K \rightarrow 0}{\sim} -K^3 \left(\frac{\lambda}{6} + \frac{\lambda^2}{2} + \frac{\lambda^3}{3} \right). \end{aligned}$$

En effet, par des formules classiques de trigonométrie :

$$g(\lambda, K) = 1 + \sin\left(\frac{K}{2}\right) \left(\sin\left(\frac{K}{2}\right) - i \cos\left(\frac{K}{2}\right) \right)$$

donc :

$$\begin{aligned} \tan [\arg g(\lambda, K)] &= -\frac{2\lambda \sin(K/2) \cos(K/2)}{1 + 2\lambda \sin^2(K/2)} \\ &= -\lambda \sin(K) (1 - 2\lambda \sin^2(K/2) + \mathcal{O}(K^4)) \\ &= -\lambda \sin(K) + 2\lambda^2 \sin(K) \sin^2(K/2) + \mathcal{O}(K^5) \\ &= -\lambda K + K^3 \left(\frac{\lambda}{6} + \frac{\lambda^2}{2} \right) + o(K^3). \end{aligned}$$

Finalement, on voit que :

$$\arg g(\lambda, K) = -\lambda K + \beta K^3 + o(K^3), \quad \text{avec} \quad \beta = \frac{\lambda}{6} + \frac{\lambda^2}{2} + \frac{\lambda^3}{3}.$$

SCHÉMA (SEC) - On calcule d'abord $g(\lambda, k\Delta x)$: si $u_j^n = \gamma \exp(ik\Delta x)$, alors :

$$u_j^{n+1} = u_j^n \left[1 - \frac{\lambda}{2} (\exp(ik\Delta x) - \exp(-ik\Delta x)) \right],$$

d'où :

$$g(\lambda, k\Delta x) = 1 - i\lambda \sin(k\Delta x).$$

On a donc :

$$|g(\lambda, k\Delta x)|^2 = 1 + \lambda^2 \sin^2(k\Delta x),$$

ce qui montre que le schéma est instable puisque $|g| > 1$, au moins pour la plupart des valeurs de k .

SCHÉMA (SEDAR) - Ce type de schéma se traite exactement comme le schéma (SEDAv) ; il est stable si et seulement si $0 \leq \lambda \leq 1$.

Exercice 11. Pour l'équation de transport :

$$\frac{\partial u}{\partial t} + c \frac{\partial u}{\partial x} = 0 \quad \text{dans } \mathbb{R} \times (0, T),$$

on considère la famille de schémas d'approximation numérique :

$$u_j^{n+1} = u_j^n + \frac{\lambda}{2} (u_{j+1}^n - u_{j-1}^n) + \mu(\Delta t, \Delta x) (u_{j+1}^n + u_{j-1}^n - 2u_j^n),$$

où $\lambda = c \frac{\Delta t}{\Delta x}$ et $\mu(\Delta t, \Delta x)$ est un paramètre dépendant du schéma choisi.

1. Donner des conditions sur $\mu(\Delta t, \Delta x)$ pour que ce schéma soit consistant et préciser son ordre.

2. Déterminer les conditions de stabilité que doivent satisfaire λ et $\mu(\Delta t, \Delta x)$.
(Indication : on pourra tout exprimer en fonction de $\theta = \sin^2(\frac{k\Delta x}{2})$)

2.2.3 Étude de la stabilité des principaux schémas implicites

Le cas des schémas implicites est un peu plus complexe car nous avons à résoudre une équation à chaque itération du temps ; il convient donc de se poser la question de l'existence et de l'unicité de cette solution. Nous allons étudier deux exemples :

$$\text{(SIDAv)} \quad \frac{u_j^{n+1} - u_j^n}{\Delta t} + c \left(\frac{u_{j+1}^{n+1} - u_j^{n+1}}{\Delta x} \right) = 0$$

$$\text{(SIC)} \quad \frac{u_j^{n+1} - u_j^n}{\Delta t} + c \left(\frac{u_{j+1}^{n+1} - u_{j-1}^{n+1}}{2\Delta x} \right) = 0$$

SCHÉMA (SIDAV) - Avec les notations habituelles, on peut réécrire le schéma (SIDAV) sous la forme :

$$(1 - \lambda)u_j^{n+1} + \lambda u_{j+1}^{n+1} = u_j^n, \quad j \in \mathbb{Z}.$$

On est donc en présence d'une équation de la forme $Au^{n+1} = u^n$, où A est un opérateur linéaire qui agit sur les suites indéxées par \mathbb{Z} (*i.e.* c'est un endomorphisme de $\ell(\mathbb{Z})$). Comme en dimension finie, il est naturel de chercher le noyau de A puisque cela est directement relié à l'unicité des solutions. Soit donc $(x_j)_{j \in \mathbb{Z}}$ un élément du noyau de A . Alors :

$$(1 - \lambda)x_j + \lambda x_{j+1} = 0 \quad \text{pour tout } j \in \mathbb{Z},$$

soit :

$$x_{j+1} = \left(\frac{\lambda - 1}{\lambda}\right)x_j \quad (j \in \mathbb{Z}).$$

Si on suppose $x_0 = 1$, on a donc :

$$x_j = \left(\frac{\lambda - 1}{\lambda}\right)^j.$$

Comme $j \in \mathbb{Z}$, $|x_j| \rightarrow +\infty$ quand $j \rightarrow +\infty$ ou quand $j \rightarrow -\infty$ (seul le cas $\lambda = 1/2$ échappe à cette règle). Si on veut avoir une solution unique (qui doit être la suite identiquement nulle), on doit restreindre l'espace des suites sur lequel on travaille pour éviter ces suites qui sont dans le noyau. Un choix classique est de prendre les suites de $\ell^2(\mathbb{Z})$, *i.e.* les suites $(u_j)_{j \in \mathbb{Z}}$ telles que

$$\sum_{j=-\infty}^{\infty} |u_j|^2 < +\infty,$$

où celles de $\ell^\infty(\mathbb{Z})$ pour lesquelles $\max\{|u_j|, j \in \mathbb{Z}\} < +\infty$. On voit que dans $\ell^\infty(\mathbb{Z})$ par exemple, s'il existe une solution, elle est unique sauf dans le cas $\lambda = 1/2$ car le choix de $\ell^\infty(\mathbb{Z})$ n'élimine pas la suite $((-1)^j)_{j \in \mathbb{Z}}$, alors que, si on choisit $\ell^2(\mathbb{Z})$, la solution est, cette fois, toujours unique. On constate donc que le choix de l'espace de suites considéré n'est pas tout à fait anodin...

Dans les cas où l'on a unicité, pour $u_j^n = \exp(ikj\Delta x)$, une solution (donc LA solution) est donnée par $u_j^{n+1} = g(\lambda, k\Delta x)u_j^n$ avec :

$$g(\lambda, k\Delta x) = \left[1 + \lambda(\exp(ik\Delta x) - 1)\right]^{-1}.$$

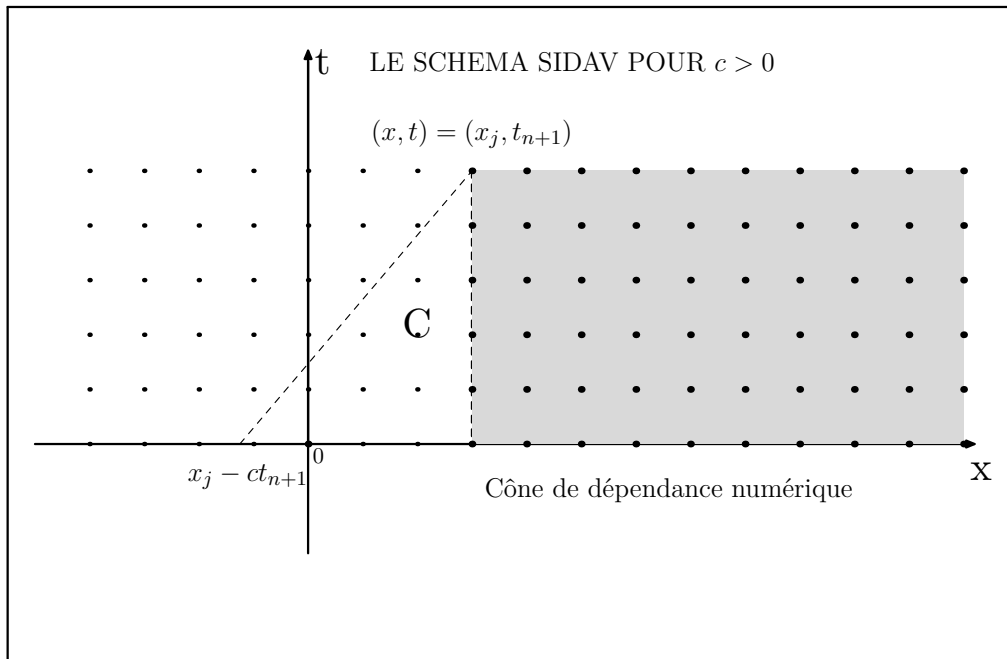
Or on a vu que :

$$\left|1 + \lambda(\exp(ik\Delta x) - 1)\right|^2 = 1 + 4(-\lambda + \lambda^2) \sin^2\left(\frac{k\Delta x}{2}\right),$$

et il est donc clair que le schéma (SIDAV) est stable si et seulement si $-\lambda + \lambda^2 \geq 0$, c'est-à-dire si et seulement si :

$$\lambda \in]-\infty, 0[\cup]1, \infty[.$$

L'interprétation est claire si l'on repense au cône de dépendance numérique : u_j^{n+1} est calculé à partir de u_j^n et de u_{j+1}^n , ce qui donne un cône de dépendance numérique de la forme indiquée sur la figure ci-dessous :



On voit donc que si $c > 0$, ce qui est équivalent à $\lambda > 0$, alors le point $j\Delta x - c(n+1)\Delta t$ n'est pas dans le cône ; alors que si $c < 0$, il l'est pour tout choix de Δt et Δx ; on dira que le schéma est inconditionnellement stable si $c < 0$. Ce qui donne beaucoup de souplesse du point de vue numérique.

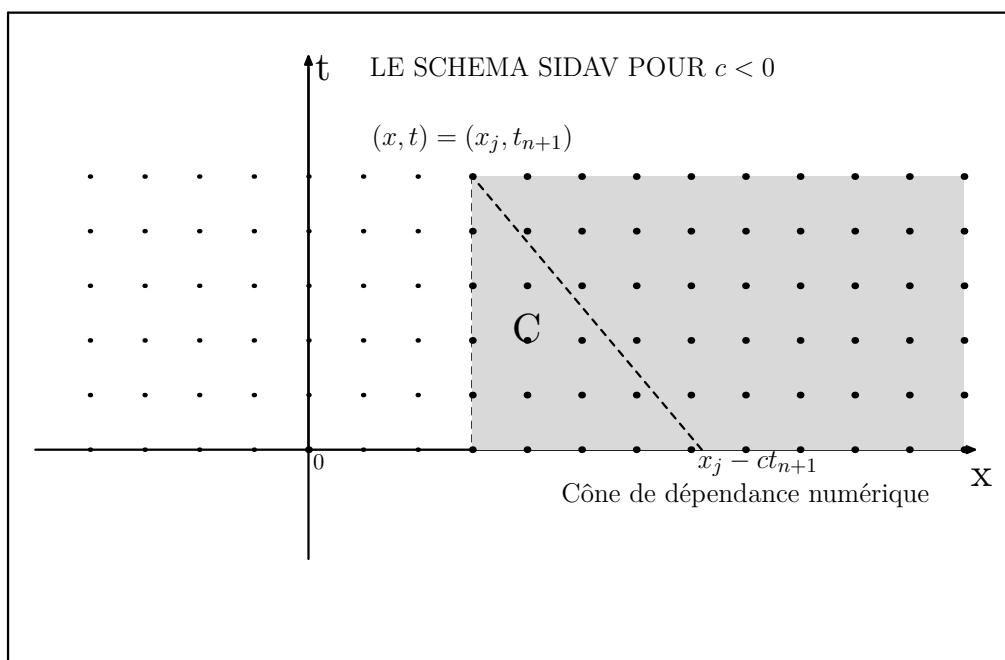


SCHÉMA (SIC) - Ce schéma est inconditionnellement stable et ceci, pour toutes les valeurs de c ; on peut le comprendre intuitivement car le cône de dépendance numérique contient tous les points $(\ell\Delta x, k\Delta t)$ pour $\ell \in \mathbb{Z}$ et $k \leq n + 1$. Voici deux preuves plus formelles de ce résultat :

(i) le coefficient d'amplification est donné par :

$$g(\lambda, k\Delta x) = \left[1 - i\lambda \sin(k\Delta x) \right]^{-1},$$

et donc :

$$\left| g(\lambda, k\Delta x) \right|^2 = \left[1 + \lambda^2 \sin^2(k\Delta x) \right]^{-1} \leq 1.$$

(ii) La deuxième preuve consiste à travailler dans $\ell^2(\mathbb{Z})$ On multiplie le schéma par u_j^{n+1} et on fait la somme de $-J$ à $+J$ ($J > 0$); il vient :

$$\begin{aligned} \sum_{j=-J}^J (u_j^{n+1})^2 &= \sum_{j=-J}^J u_j^{n+1} u_j^n - \frac{\lambda}{2} \sum_{j=-J}^J (u_{j+1}^{n+1} - u_{j-1}^{n+1}) u_j^{n+1} \\ &= \sum_{j=-J}^J u_j^{n+1} u_j^n - \frac{\lambda}{2} (u_{J+1}^{n+1} u_J^{n+1} - u_{J-1}^{n+1} u_{-J}^{n+1}). \end{aligned}$$

On utilise alors l'inégalité de Cauchy-Schwarz pour le premier terme du membre de droite et on fait tendre J vers $+\infty$: comme $(u_j^{n+1})_j \in \ell^2(\mathbb{Z})$, le deuxième terme tend vers zéro et finalement :

$$\left(\sum_{j=-\infty}^{\infty} (u_j^{n+1})^2 \right)^{1/2} \leq \left(\sum_{j=-\infty}^{\infty} (u_j^n)^2 \right)^{1/2},$$

ce qui donne la stabilité dans $\ell^2(\mathbb{Z})$.

Remarque : cette preuve est l'analogie de celle qui consiste à multiplier l'équation par u et intégrer sur \mathbb{R} ; en effet ,

$$\left(\frac{\partial u}{\partial t} + c\frac{\partial u}{\partial x}\right)u = \frac{\partial}{\partial t}\left(\frac{u^2}{2}\right) + \frac{\partial}{\partial x}\left(\frac{u^2}{2}\right)$$

et en intégrant :

$$\frac{d}{dt}\left(\int_{\mathbb{R}} \frac{u^2}{2} dx\right) + c \int_{\mathbb{R}} \frac{\partial}{\partial x}\left(\frac{u^2}{2}\right) dx = 0$$

mais $\int_{\mathbb{R}} \partial_x(u^2/2) = 0$ car $u \in L^2(\mathbb{R})$, et donc finalement :

$$\int_{\mathbb{R}} u^2(t, x) dx = \int_{\mathbb{R}} u_0^2(x) dx .$$

2.2.4 Un résultat de convergence

Pour obtenir la convergence des schémas numériques que nous venons d'étudier, une première méthode consiste à adapter la méthode présentée dans le chapitre 1 sur les équations elliptiques. Nous en présentons ici une deuxième sur l'exemple du schéma (SIC) où l'on travaille dans L^2 avec des solutions faibles.

Théorème 2.2. *Pour tout choix de Δx et Δt , la solution $(u_j^n)_{n,j}$ du schéma (SIC) converge dans L^2 faible vers l'unique solution u de (2.1)-(2.2) si $u_0 \in \mathcal{C}(\mathbb{R}) \cap L^2(\mathbb{R})$.*

Preuve : Remarquons tout d'abord que le schéma (SIC) admet une solution par des arguments de suites récurrentes d'ordre 2 ou en utilisant le fait que le schéma s'écrit sous la forme :

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

i.e. $Au^{n+1} = u^n$ avec A satisfaisant :

$$(P) \quad \langle Au, u \rangle = \|u\|^2 ,$$

où $\langle \cdot, \cdot \rangle$ désigne le produit scalaire dans $\ell^2(\mathbb{Z})$, *i.e.* ,

$$\langle u, v \rangle = \sum_{j=-\infty}^{\infty} u_j v_j \quad \text{si } u = (u_j)_j, v = (v_j)_j$$

et $\|\cdot\|$ est la norme associée. On voit alors que :

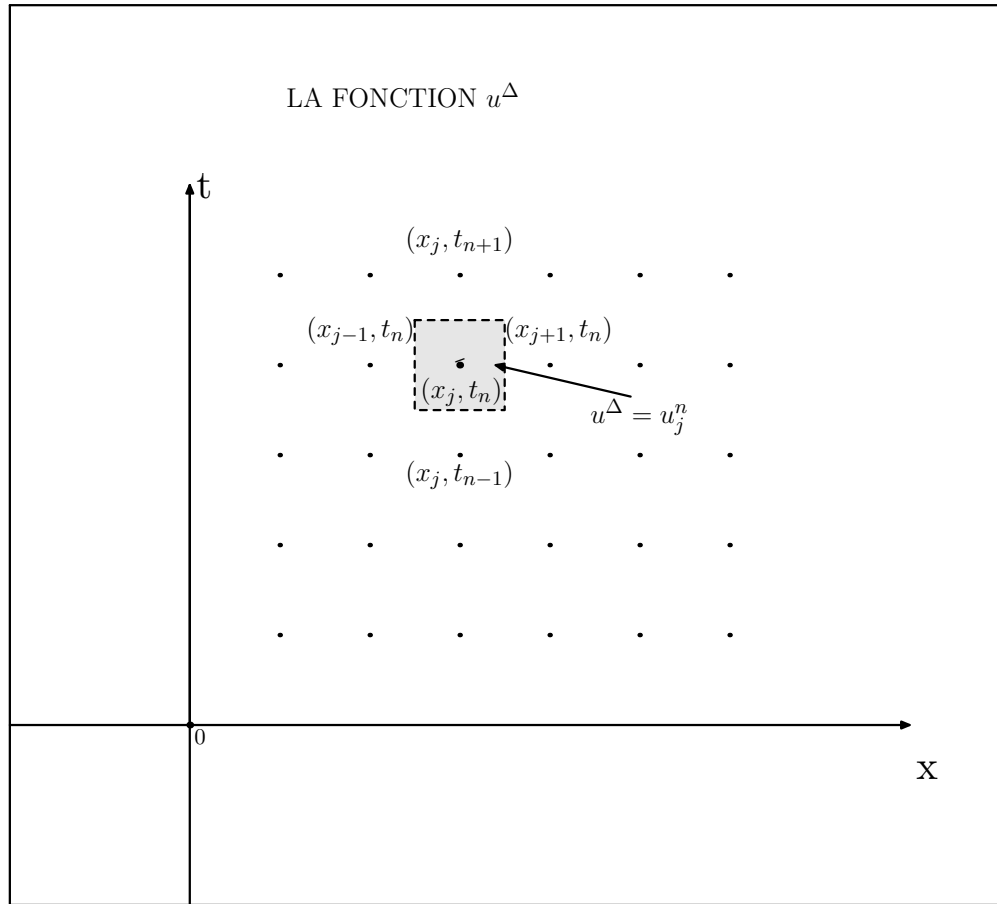
(i) $\text{Im}(A)$ est dense car si $u \in (\text{Im } A)^\perp$, alors :

$$\|u\|^2 = \langle Au, u \rangle = 0 .$$

(ii) $\text{Im}(A) = \ell^2(\mathbb{Z})$ car, si $f \in \ell^2(\mathbb{Z})$ il existe une suite $f_\varepsilon \rightarrow f$ telle que $f_\varepsilon \in \text{Im}(A)$ pour tout $\varepsilon > 0$. Mais si u_ε satisfait $Au_\varepsilon = f_\varepsilon$, on a, d'après (P) :

$$\begin{aligned} \|u_\varepsilon - u_{\varepsilon'}\|^2 &= \langle Au_\varepsilon - Au_{\varepsilon'}, u_\varepsilon - u_{\varepsilon'} \rangle \\ &= \langle f_\varepsilon - f_{\varepsilon'}, u_\varepsilon - u_{\varepsilon'} \rangle \\ &\leq \|f_\varepsilon - f_{\varepsilon'}\| \cdot \|u_\varepsilon - u_{\varepsilon'}\| \end{aligned}$$

en divisant par $\|u_\varepsilon - u_{\varepsilon'}\|$, il résulte que $(u_\varepsilon)_\varepsilon$ est une suite de Cauchy dans l'espace de Hilbert $\ell^2(\mathbb{Z})$ et donc $u_\varepsilon \rightarrow u$ qui est solution de $Au = f$, puisque l'opérateur A est borné, donc continu.



On note maintenant u^Δ la fonction définie sur $\mathbb{R} \times [0, T]$ par :

$$u^\Delta(x, t) = u_j^n \quad \text{si :} \quad \begin{cases} |x - x_j| \leq \frac{\Delta x}{2}, & |t - t_n| \leq \frac{\Delta t}{2} & \text{pour } n > 0 \\ |x - x_j| \leq \frac{\Delta x}{2}, & 0 \leq t \leq \frac{\Delta t}{2} & \text{pour } n = 0 \\ |x - x_j| \leq \frac{\Delta x}{2}, & T - \frac{\Delta t}{2} \leq t \leq T & \text{pour } n = N \end{cases}$$

D'après le calcul de stabilité $\ell^2(\mathbb{Z})$, on voit que :

$$\begin{aligned} \|u^\Delta\|_{L^2(\mathbb{R} \times (0, T))} &= \sum_{n=1}^{N-1} \Delta x \cdot \Delta t \sum_{j \in \mathbb{Z}} (u_j^n)^2 + \Delta x \frac{\Delta t}{2} \sum_{j \in \mathbb{Z}} (u_j^0)^2 + \Delta x \frac{\Delta t}{2} \sum_{j \in \mathbb{Z}} (u_j^N)^2 \\ &\leq \Delta x \cdot N \cdot \Delta t \sum_{j \in \mathbb{Z}} (u_j^0)^2. \end{aligned}$$

Si on pose :

$$u_j^0 = \frac{1}{\Delta x} \int_{x_j - \Delta x/2}^{x_j + \Delta x/2} u_0(x) dx$$

qui satisfait :

$$|u_j^0| \leq \frac{1}{(\Delta x)^{1/2}} \left(\int_{x_j - \Delta x/2}^{x_j + \Delta x/2} |u_0(x)|^2 dx \right)^{1/2},$$

on a alors (en utilisant $N\Delta t = 1$) :

$$\|u^\Delta\|_{L^2(\mathbb{R} \times (0, T))} \leq \int_{\mathbb{R}} |u_0(x)|^2 dx.$$

Les fonctions $(u^\Delta)_\Delta$ sont donc bornées dans L^2 et on peut ainsi en extraire une sous-suite (que nous noterons encore $(u^\Delta)_\Delta$) qui converge dans L^2 -faible. Il s'agit maintenant de déterminer les propriétés de la limite faible \bar{u} .

Soit $\phi \in C_0^1(\mathbb{R} \times [0, T])$, et notons $\phi_j^n = \phi(n\Delta t, j\Delta x)$; on construit ϕ^Δ comme ci-dessus pour u^Δ et on calcule le produit de ϕ_j^{n+1} par le schéma en sommant sur n et sur j :

$$\sum_{n=0}^{N-1} \sum_{j \in \mathbb{Z}} \left(\frac{u_j^{n+1} - u_j^n}{\Delta t} \right) \phi_j^{n+1} + \frac{c}{2\Delta x} \sum_{n=0}^{N-1} \sum_{j \in \mathbb{Z}} (u_{j+1}^{n+1} - u_{j-1}^{n+1}) \phi_j^{n+1} = 0,$$

ces sommes ayant un sens puisque ϕ est à support compact en x . On réécrit ces sommes sous la forme :

$$\sum_{n=0}^{N-1} \sum_{j \in \mathbb{Z}} u_j^n \left(\frac{\phi_j^{n-1} - \phi_j^n}{\Delta t} \right) - \sum_{j \in \mathbb{Z}} \frac{u_j^0 \phi_j^1}{\Delta t} + \frac{c}{2\Delta x} \sum_{n=0}^{N-1} \sum_{j \in \mathbb{Z}} u_j^{n+1} (\phi_{j-1}^{n+1} - \phi_{j+1}^{n+1}) = 0,$$

et on multiplie ces sommes par $\Delta t \cdot \Delta x$; puis on exprime le tout en fonction de u^Δ et des fonctions $\Delta_t \phi$ et $\Delta_x \phi$ définies pour $0 < n < N$ par :

$$\Delta_t \phi(x, t) = \frac{\phi_j^n - \phi_j^{n+1}}{\Delta t} \text{ si } |x - x_j| \leq \frac{\Delta x}{2}, \quad |t - t_n| \leq \frac{\Delta t}{2},$$

et

$$\Delta_x \phi(x, t) = \frac{\phi_{j+1}^n - \phi_{j-1}^n}{2\Delta x} \text{ si } |x - x_j| \leq \frac{\Delta x}{2}, \quad |t - t_n| \leq \frac{\Delta t}{2}.$$

On obtient alors :

$$\int_0^T \int_{\mathbb{R}} u^\Delta \Delta_t \phi \, dx \, dt + \int_{\mathbb{R}} u^\Delta(x, 0) \phi^\Delta(x, 0) \, dx + c \int_0^T \int_{\mathbb{R}} u^\Delta \Delta_x \phi \, dx \, dt = 0.$$

On passe à la limite en utilisant d'une part la convergence de u^Δ vers \bar{u} dans L^2 -faible, et d'autre part le fait que

$$\begin{aligned} \Delta_t \phi &\rightarrow \frac{\partial \phi}{\partial t} \text{ dans } L^2\text{-fort} \\ \Delta_x \phi &\rightarrow \frac{\partial \phi}{\partial x} \text{ dans } L^2\text{-fort} \end{aligned}$$

ce qui entraîne que \bar{u} satisfait le problème limite suivant :

$$\int_0^T \int_{\mathbb{R}} \bar{u} \frac{\partial \phi}{\partial t} \, dx \, dt + \int_{\mathbb{R}} \bar{u}(x, 0) \phi(x, 0) \, dx + c \int_0^T \int_{\mathbb{R}} \bar{u} \frac{\partial \phi}{\partial x} \, dx \, dt$$

En remarquant que $\bar{u}(x, 0) = u_0(x)$, on voit donc que \bar{u} est solution faible de (2.1)-(2.2), d'où finalement $\bar{u} = u$, de par l'unicité de la solution faible de ce problème. \square

2.2.5 Schémas d'ordre supérieur

Pour obtenir des schémas d'ordre supérieur, on utilise en général une méthode que nous décrivons maintenant sur le schéma (SEC). On commence par un calcul de type consistance pour ce schéma :

$$\begin{aligned} \frac{1}{\Delta t} (u_j^{n+1} - u_j^n) + \frac{c}{2\Delta x} (u_{j+1}^n - u_{j-1}^n) &= \left(\frac{\partial u}{\partial t} + c \frac{\partial u}{\partial x} \right) (x_j, t_n) \\ &+ \frac{\Delta t}{2} \frac{\partial^2 u}{\partial t^2} (x_j, t_n) + \mathcal{O}((\Delta t)^2) + \mathcal{O}((\Delta x)^2). \end{aligned}$$

Pour obtenir un schéma d'ordre supérieur en temps (et même sans les deux variables puisqu'en général, Δx et Δt sont du même ordre), il suffit d'éliminer le terme en $\partial^2 u / \partial t^2$. Pour cela on utilise l'équation qui donne :

$$\frac{\partial^2 u}{\partial t^2} = -\frac{\partial}{\partial t} \left(c \frac{\partial u}{\partial x} \right) = c \frac{\partial}{\partial x} \left(\frac{\partial u}{\partial t} \right) = c^2 \frac{\partial^2 u}{\partial x^2},$$

et on discrétise cette dérivée seconde en x , ce qui conduit au schéma explicite suivant :

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

Le coefficient d'amplification de ce schéma est :

$$g(\lambda, k\Delta x) = 1 - 2\lambda^2 \sin^2\left(\frac{k\Delta x}{2}\right) - i\lambda \sin(k\Delta x),$$

d'où :

$$\begin{aligned} |g(\lambda, k\Delta x)|^2 &= 1 - 4\lambda^2 \sin^2\left(\frac{k\Delta x}{2}\right) + 4\lambda^4 \sin^4\left(\frac{k\Delta x}{2}\right) + \lambda^2 \sin^2(k\Delta x) \\ &= 1 - 4\lambda^2 \sin^2\left(\frac{k\Delta x}{2}\right) [1 - \cos^2\left(\frac{k\Delta x}{2}\right)] + 4\lambda^4 \sin^4\left(\frac{k\Delta x}{2}\right) \\ &= 1 - 4\lambda^2 \sin^4\left(\frac{k\Delta x}{2}\right) (1 - \lambda^2). \end{aligned}$$

Le schéma est donc stable si et seulement si $|\lambda| \leq 1$ (condition de CFL).

Exercice 12. (Autres approches pour construire des schémas)

1. (Méthode semi-lagrangienne) On souhaite résoudre :

$$\frac{\partial u}{\partial t} + c(x) \frac{\partial u}{\partial x} = 0 \quad \text{dans } \mathbb{R} \times (0, T),$$

$$u(x, t) = u_0(x), \quad \text{dans } \mathbb{R},$$

pour $u_0 \in \mathcal{C}(\mathbb{R})$ et c est une fonction lipschitzienne (ou \mathcal{C}^1) bornée. Si $X(\cdot)$ est la solution de :

$$\dot{X}(s) = c(X(s)) \quad , \quad X(0) = x ,$$

justifier l'égalité $u(x, t) = u(X(-h), t - h)$ pour $h > 0$ et utiliser cette propriété pour construire un (et même toute une famille de) schéma(s) numérique(s). (On pourra ressortir de leur boîte les cours d'Analyse Numérique des équations différentielles ordinaires pour approcher $X(-\Delta t)$ car $h = \Delta t$ et remettre à jour vos connaissances en interpolation pour en déduire une approximation de $u(X(-h), t - h)$.)

Discuter la stabilité de ces schémas.

2. (Semi-discrétisation en espace) On suppose pour simplifier que u_0 est 1-périodique et que c est constant. Montrer qu'en discrétisant seulement en espace (disons par (SEDAv) ou (SEDAr)) et avec un choix $\Delta x = N^{-1}$, $N \in \mathbb{N}$, on se ramène à une équation différentielle ordinaire du type :

$$\frac{du}{dt} + Au = 0 .$$

Examiner quelques méthodes à un pas pour la résolution de cette équation différentielle ordinaire.

2.2.6 Approximation par un domaine de calcul borné

La méthode la plus naturelle consiste à remplacer (2.1)-(2.2) par :

$$(2.10) \quad \frac{\partial u^R}{\partial t} + c \frac{\partial u^R}{\partial x} = 0 \text{ dans }]-R, R[\times (0, T),$$

$$(2.11) \quad u^R(x, 0) = u_0(x) \text{ dans }]-R, R[,$$

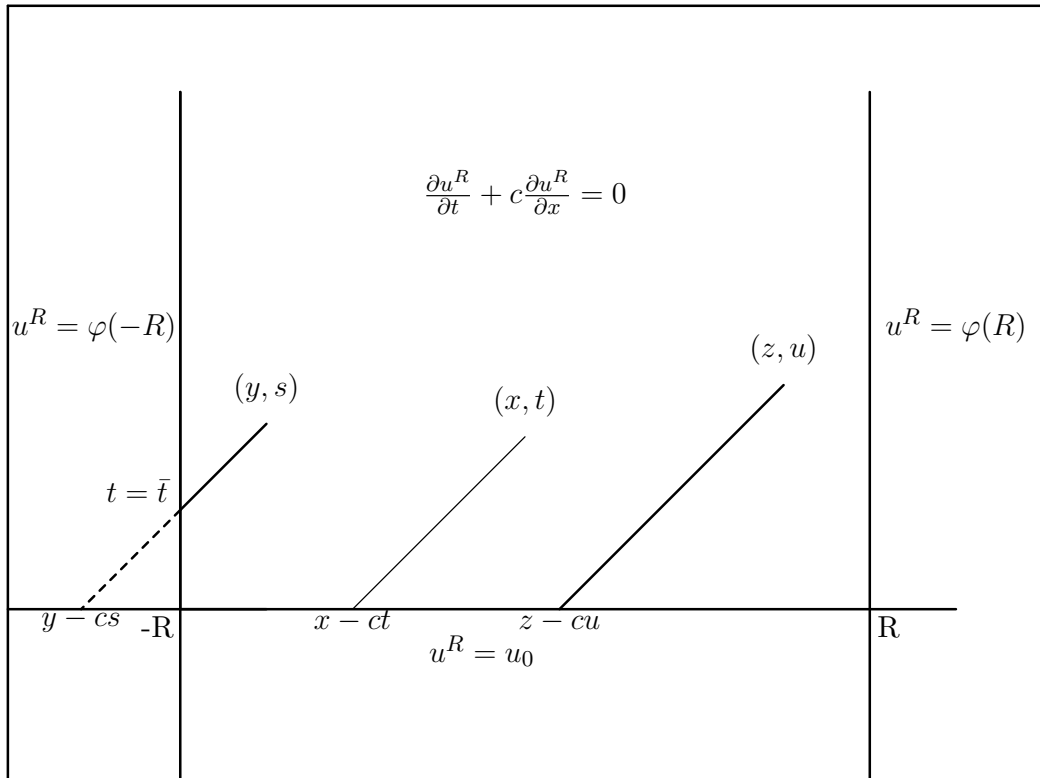
avec

$$(CLA) \quad u^R(\pm R, t) = \varphi(\pm R) \text{ dans } t \in (0, T),$$

où φ est une fonction continue. On a donc une équation posée dans un domaine borné et on a besoin d'ajouter des conditions aux limites artificielles en R et $-R$ matérialisées par φ .

Par la méthode des caractéristiques, on calcule facilement u^R . Comme le montre la figure ci-dessous, quand $c > 0$,

$$u^R(x, t) = \begin{cases} u_0(x - ct) & \text{si } x - ct \geq -R \\ \varphi(-R) & \text{sinon} \end{cases}$$



Plusieurs remarques :

- (i) Le choix $\varphi(-R) = u_0(-R)$ paraît naturel pour assurer la continuité de u^R ; ce n'est pas une nécessité, mais c'est plutôt agréable.
- (ii) La condition aux limites en $+R$ n'est pas vue puisque les caractéristiques sont dirigées vers l'intérieur du domaine quand on se place sur le bord $x = +R$: elles sont "rentrantes". Ce n'est pas le cas en $-R$ où elles sont sortantes.
- (iii) Si on calcule u^R , on aura une bonne approximation de u seulement pour $x - ct \geq -R$ et pas dans le domaine tout entier. Une idée naturelle consisterait à "pencher" le domaine de calcul, mais cette idée que l'on applique parfois, est délicate à mettre en place si, par exemple, c dépend de x .