

Introduction aux équations aux dérivées partielles et à leur approximation numérique

Anne-Sophie Bonnet-BenDhia, Sonia Fliss,
Patrick Joly et Philippe Moireau

April 25, 2012

Table des matières

1	Introduction aux problèmes hyperboliques linéaires	9
1.1	Exemples de modèles hyperboliques	9
1.1.1	Equation de transport linéaire	9
1.1.2	L'équation des ondes et ses variantes	11
1.2	La méthode des caractéristiques pour l'équation de transport	14
1.2.1	Le problème de Cauchy pour l'équation de transport linéaire	14
1.2.2	Résolution de l'équation de transport à coefficients constants	15
1.2.3	Résolution de l'équation de transport à coefficients variables	19
1.3	Systèmes hyperboliques linéaires	25
1.3.1	Systèmes hyperboliques linéaires	25
1.3.2	L'équation des ondes en dimension 1	27
2	Approximation par différences finies	31
2.1	Présentation de la méthode de différences finies	31
2.1.1	Approximation de quelques opérateurs différentiels	31
2.1.2	La démarche : Etude de la consistance et la stabilité	35
2.1.3	Application aux équations hyperboliques linéaires	36
2.2	Le schéma explicite centré pour l'équation d'advection	38
2.2.1	Consistance et ordre du schéma	40
2.2.2	Propagation numérique et condition nécessaire de convergence	42
2.2.3	Analyse de la stabilité d'un schéma : principes généraux	43
2.2.4	Analyse de stabilité L^2 du schéma explicite centré : la méthode de Fourier-Von Neumann	46
2.3	Le schéma de Lax-Friedrichs pour l'équation d'advection	51
2.3.1	Consistance et ordre du schéma	52
2.3.2	Vitesse de propagation numérique	53
2.3.3	Analyse de stabilité	54
2.3.4	Convergence du schéma	56
2.4	Le schéma saute-mouton pour l'équation des ondes	61
2.4.1	Présentation du schéma	61
2.4.2	Le schéma de démarrage : choix des conditions initiales approchées	61
2.4.3	Ordre du schéma : consistance et erreur de troncature	63

2.4.4	Analyse de stabilité L^2 par techniques énergétiques	63
3	Problèmes hyperboliques non linéaires :	67
3.1	Lois de conservation	67
3.1.1	Introduction	67
3.1.2	Construction d'une équation de conservation	68
3.1.3	Exemples	69
3.1.4	Ajout d'un terme de viscosité	71
3.2	Solutions classiques : méthode des caractéristiques	72
3.2.1	Solutions classiques au problème de Cauchy	72
3.2.2	Solutions classiques locales et naissance d'un choc	77
3.3	Solutions faibles - Solutions entropiques	80
3.3.1	Solutions faibles	81
3.3.2	Solutions C^1 par morceaux. Condition de choc	82
3.3.3	Solutions entropiques	87
3.3.4	Existence et unicité de la solution entropique	95
3.4	Le problème de Riemann	96
3.4.1	Présentation du problème	96
3.4.2	Une solution auto-semblable	97
3.4.3	Le cas d'une onde de détente	98
3.4.4	Le cas d'une onde de choc	98
3.4.5	Solution du problème de Riemann à 2 états	99
3.4.6	Le problème de Riemann à trois états	99
4	Discrétisation des équations hyperboliques non linéaires 1D	105
4.1	Présentation du problème	105
4.2	Propriétés des schémas	106
4.2.1	Schémas sous forme conservative	106
4.2.2	Consistance	108
4.2.3	Linéarisation du schéma - Stabilité L^2	109
4.2.4	Vers la convergence : le théorème de Lax-Wendroff	111
4.2.5	Un exemple de schéma conservatif : le schéma de Murman Roe	113
4.2.6	Schémas entropiques	114
4.2.7	Schémas monotones	116
4.2.8	Deux exemples de schémas monotones	121
4.3	Un exemple fondamental - Le schéma de Godounov	123

Préambule

Les **équations aux dérivées partielles** sont aux fonctions de **plusieurs variables** ce que les équations différentielles sont aux fonctions d'une variable réelle. Elles expriment, sous forme d'égalités, des relations que doivent satisfaire les dérivées partielles d'une certaine fonction inconnue u de plusieurs variables afin de décrire un phénomène physique, satisfaire une propriété prescrite, ...etc ... Quand on s'intéresse à une application donnée, on est très vite amené à faire la distinction entre les **EDPs stationnaires** (où toutes les variables jouent un rôle équivalent) et les **EDPs d'évolution** (où une des variables joue un rôle privilégié : le temps). On rencontre de telles équations dès qu'on s'intéresse à des questions de modélisation : en physique, en électromagnétisme, en mécanique du solide et des fluides bien sûr, mais aussi en biologie, en chimie, en économie, en finance... On y est naturellement confronté quand on s'intéresse à des problèmes de calcul des variations ou plus généralement d'optimisation. Savoir manipuler des équations aux dérivées partielles fait de nos jours partie du quotidien de l'ingénieur.

L'objectif de ce cours est d'introduire à l'étude et à la résolution approchée sur ordinateur des équations aux dérivées partielles les plus simples qu'on rencontre dans les applications. Ce cours est un cours de mathématiques appliquées qui prétend à un certain niveau de rigueur et de généralité. Toutefois, compte tenu de la complexité technique du sujet, nous serons parfois amenés à sacrifier quelque peu cette rigueur et cette généralité : la théorie des équations aux dérivées partielles et de leur approximation numérique est une science difficile, au coeur de la recherche moderne en mathématiques, dont seuls les rudiments sont accessibles en première année d'Ecole d'ingénieurs. C'est pourquoi nous avons fait le choix pédagogique d'aborder le sujet par le « petit bout de la lorgnette » en choisissant de présenter, plutôt qu'une grande théorie, des exemples et des problèmes modèles simples pour lesquels on peut aller assez loin dans la description des propriétés des solutions, souvent par le biais de calculs explicites. Qu'on ne s'y trompe pas pourtant : les propriétés que nous mettrons en évidence sont le plus souvent représentatives de ce qui se passe dans le cas général (cela sera d'ailleurs mis en lumière de temps à autre). De plus, la compréhension de ces exemples simples est très utile pour aborder des théories plus élaborées.

Nous avons l'habitude de classer les équations aux dérivées partielles en trois grandes classes fondamentales d'équations : les **équations elliptiques** (qui servent typique-

ment à décrire des phénomènes d'équilibre en physique) pour les problèmes stationnaires, les **équations paraboliques** (qui permettent de décrire des phénomènes de diffusion) et les **équations hyperboliques** (qui permettent de décrire les phénomènes de propagation) pour les problèmes d'évolution. C'est cette dernière classe d'équations qui fera l'objet de ce cours. Les deux autres catégories seront abordés dans deux cours de deuxième année sur la méthode des éléments finis : les équations elliptiques dans le cours MA201 et les équations paraboliques dans le cours MA206.

Dans ce cours, nous abordons les problèmes linéaires et non linéaires que nous traiterons par le biais d'une équation modèle. Pour les équations hyperboliques linéaires, ce sera l'**équation de transport** (Chapitre 1). Pour les équations hyperboliques non linéaires, enfin, ce seront les lois de conservation scalaires dont le prototype est l'**équation de Burgers** (Chapitre 3). Par souci de simplicité, nous nous limiterons la plupart du temps à la dimension 1 d'espace.

Rares sont les situations réalistes pour lesquelles on sait calculer une solution analytique à la main. Pour aller plus loin dans la connaissance fine et « quantitative » des solutions d'une EDPs (dont l'analyse mathématique ne fournira en général que l'existence d'une solution, voire l'unicité et quelques propriétés qualitatives), l'ingénieur sera amené à se tourner vers l'utilisation de méthodes numériques pour calculer une solution approchée. Le second objectif du cours est d'introduire le lecteur à la méthode de discrétisation la plus classique (et la plus simple) : la **méthode des différences finies** qui est la « mère » des méthodes de discrétisation plus élaborées telles que les méthodes d'éléments finis, de volumes finis ou de Galerkin discontinu (qui seront abordées plus tard dans le cursus). La mise en oeuvre de ces méthodes est en général simple. Leur théorie ne l'est pas forcément et un des buts du cours est d'aborder - dans le cas des différences finies - les aspects théoriques des méthodes numériques, ce que l'on appelle l'**analyse numérique**. Nous introduirons en particulier les concepts de **consistance**, de **stabilité** et de **convergence** et les liens étroits entre ces trois notions. Le lecteur exigeant pourra s'offusquer de voir que nous avons choisi la plupart du temps de présenter et analyser les méthodes numériques sur les cas simples pour lesquels on saurait par exemple calculer la solution à la main. A nouveau, le but est pédagogique : il est bien clair qu'une méthode numérique doit bien marcher dans ces cas simples si on veut avoir une chance qu'elle marche dans les cas compliqués. Par ailleurs, sur ces cas simples, on arrive assez facilement à obtenir des résultats d'analyse numérique précis.

Signalons qu'il y a des connections fortes entre ce cours et d'autres cours de Mathématiques de 1ère année, en particulier le cours MA102 sur les outils fondamentaux d'analyse et le cours AO102 sur la théorie des équations différentielles ordinaires, dans lesquels le lecteur trouvera les prérequis nécessaires pour appréhender aisément le contenu de ce cours.

Ce polycopié est le résultat de l'évolution de ce cours au fil des dernières années. Il a bénéficié d'un grand nombre de contributions dont celle de Véronique Duwig et Jerónimo Rodríguez et de précieuses relectures dont celles de Eliane Bécache, Stéphanie Chaillat, Julien Coatléven, François Févotte, Sébastien Impériale et de quelques élèves. Il est cependant encore largement perfectible : n'hésitez donc pas à nous faire part de vos remarques, corrections et critiques.

Chapitre 1

Introduction à la théorie et l'approximation des problèmes hyperboliques linéaires

1.1 Exemples de modèles hyperboliques

Commençons par donner quelques exemples d'équations intervenant dans des applications classiques issues de la physique, de la biologie ou des sciences de l'ingénieur. Elles sont toutes de nature hyperbolique, notion qui sera définie plus loin, ce qui signifie concrètement que la vitesse de propagation est finie, contrairement à l'équation de la chaleur.

1.1.1 Equation de transport linéaire

L'équation hyperbolique la plus simple est l'équation de transport linéaire. Elle consiste à trouver la solution $u(x, t) \in \mathbb{R}$ de l'équation aux dérivées partielles suivante :

$$\begin{cases} \frac{\partial u}{\partial t} + c(x, t) \frac{\partial u}{\partial x} = 0, & t \geq 0, x \in \mathbb{R}, \\ u(x, t = 0) = u^0(x) & x \in \mathbb{R}. \end{cases} \quad (1.1)$$

Ici le scalaire $c(x, t)$ est donné : c'est par définition la vitesse de propagation associée à l'équation, au point x à l'instant t . On comprendra plus loin cette dénomination.

Cette équation qui peut sembler très simple pose des difficultés numériques considérables, elle est toujours l'objet de recherches actuellement (il s'agit notamment de savoir calculer la solution sur des temps très grands). Par ailleurs, couplée à d'autres équations, elle pose des difficultés théoriques également.

On rencontre cette équation dans un grand nombre d'applications. Citons en quelques unes.

EXEMPLE 1.1.1 (LA CIRCULATION AUTOMOBILE)

On étudie la circulation automobile sur une route. La variable $u(x, t)$ représente la quantité de voitures présentes entre les bornes x et $x + \Delta x$ à l'instant t et on appelle $F(x, t)$ le flux de voitures par minute qui passent à l'instant t devant la borne x . On suppose que chaque conducteur ajuste la vitesse de sa voiture en fonction uniquement de la vitesse de la voiture qui le précède. Alors la **conservation de la quantité** de voitures (il n'y a ni station-service, ni itinéraires de délestage) se traduit par :

$$\frac{\partial u}{\partial t}(x, t) + \frac{\partial F}{\partial x}(x, t) = 0.$$

Si toutes les voitures roulaient à vitesse constante donnée c , on aurait :

$$F(x, t) = c u(x, t)$$

et donc l'équation de transport linéaire (1.1).

EXEMPLE 1.1.2 (LES ÉQUATIONS CINÉTIQUES)

La physique cinétique décrit les plasma ou gaz dilués et fournit un ensemble important d'équations de transport dont les variables sont : le temps t , la position $x \in \Omega$ des particules avec Ω le domaine d'étude et leur vitesse $v \in V$ avec V l'ensemble des vitesses admissibles. L'exemple le plus classique est l'équation de scattering décrivant l'évolution d'une densité $f(x, t, v)$ de particules (neutrons, amibes ou bactéries en ce qui concerne les applications à la biologie)

$$\begin{cases} \frac{\partial}{\partial t} f(x, t, v) + v \frac{\partial}{\partial x} f(x, t, v) = K[f], \\ f(x, t = 0, v) = f_0(x, v) \quad \text{donnée,} \end{cases}$$

où K est donnée par :

$$K[f] = \int_V k(v, v') f(x, t, v') dv' - \int_V k(v', v) f(x, t, v') dv',$$

où $k(v, v')$ représente une probabilité de « tourner » d'une vitesse v' à une vitesse v .

Cette équation découle encore d'une relation de **conservation** de la quantité de particules.

EXEMPLE 1.1.3 (MODÈLE DE DÉMOGRAPHIE/RENOUVELLEMENT CELLULAIRE)

Ici $u(x, t)$ représente la densité d'individus d'âge $x \geq 0$ à l'instant t . Le taux de mortalité est noté $d(x)$ et les individus peuvent donner naissance à des nouveaux nés d'âge

$x = 0$ avec un taux de fécondité $b(x)$:

$$\begin{cases} \frac{\partial}{\partial t} u(x, t) + \frac{\partial}{\partial x} u(x, t) + d(x)u(x, t) = 0, \\ u(0, t) = \int_0^{+\infty} b(x') u(x', t) dx' \end{cases}$$

Dans le cas de la mitose cellulaire, il est naturel de choisir $d(x) = \chi_{\{x > x_0\}}$ ¹ (disparition des cellules qui se divisent) et $b(x) = 2\chi_{\{x > x_0\}}$ (la mitose donne naissance à deux cellules identiques).

1.1.2 L'équation des ondes et ses variantes

L'équation des ondes fournit un des exemples les plus simples de système hyperbolique. Voyons cet aspect en détail. Les ondes scalaires se propagent dans un milieu monodimensionnel selon l'équation :

$$\begin{cases} \frac{\partial^2 u}{\partial t^2} - c^2 \frac{\partial^2 u}{\partial x^2} = 0, & t \in \mathbb{R}, x \in \mathbb{R} \\ u(x, t = 0) = u^0(x), \\ \frac{\partial u}{\partial t}(x, t = 0) = u^1(x). \end{cases} \quad (1.2)$$

Ici la constante c désigne la vitesse de l'onde. Cette équation du deuxième ordre peut aussi être écrite comme un système d'équations hyperboliques du premier ordre. En effet, posons :

$$v = \frac{\partial u}{\partial t}, \quad w = c \frac{\partial u}{\partial x}.$$

On a alors :

$$\begin{cases} \frac{\partial w}{\partial t} - c \frac{\partial v}{\partial x} = 0, \\ \frac{\partial v}{\partial t} - c \frac{\partial w}{\partial x} = 0. \end{cases}$$

Si on introduit le vecteur U et la matrice C définis par :

$$U = \begin{bmatrix} w \\ v \end{bmatrix} \quad \text{et} \quad C = \begin{bmatrix} 0 & -c \\ -c & 0 \end{bmatrix}$$

l'équation (1.2) se met sous la forme :

$$\frac{\partial U}{\partial t} + C \frac{\partial U}{\partial x} = 0$$

¹On désigne par χ_A la fonction caractéristique de A

qu'on peut voir comme une équation de transport linéaire vectorielle où la solution est à valeurs dans \mathbb{R}^2 . C'est en fait un premier exemple de système hyperbolique linéaire tel que nous les définirons plus loin.

Nous donnons dans la suite des exemples de modèles de propagation d'ondes en physique. La liste est évidemment non exhaustive : nous aurions pu citer également la propagation des ondes élastiques issue des équations de l'élastodynamique, ...

EXEMPLE 1.1.4 (CORDE VIBRANTE)

Le problème est de calculer le mouvement d'une corde, de longueur L , fixée en ses extrémités et qui est, soit écartée de sa position d'équilibre et lâchée (corde de guitare), soit frappée (corde de piano) de façon à lui imprimer, en ses différents points, des vitesses de déplacement vertical.

Ainsi au repos, la corde occupe un intervalle $[a, b]$. Sous l'action d'une force normale d'intensité f (dans le cas de la corde de piano) ou d'un écart de sa position d'équilibre u^0 (écart de position) et u^1 (écart de vitesse), la corde se déforme. Si on appelle $u(x, t)$ le déplacement latéral de la corde au point x à l'instant t , les variations de u sont décrites par les équations :

$$\begin{cases} \frac{\partial^2 u}{\partial t^2} - \frac{\partial^2 u}{\partial x^2} = f, & \text{pour } t \in \mathbb{R}^+, x \in [a, b], \\ u = 0 & \text{en } x = a \text{ et } x = b \\ u(x, 0) = u^0(x), & \text{pour } x \in [a, b], \\ \frac{\partial}{\partial t} u(x, 0) = u^1(x), & \text{pour } x \in [a, b]. \end{cases}$$

On montre que dans le cas d'une peau de tambour par exemple, le déplacement

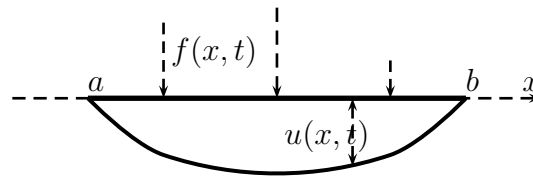


FIGURE 1.1 – Déplacement d'une corde vibrante

normal de la peau vérifie également une équation des ondes, la seule différence est que la position x est dans un domaine de \mathbb{R}^2 .

EXEMPLE 1.1.5 (ACOUSTIQUE LINÉAIRE)

Le son est en fait une conséquence d'un mouvement matériel d'oscillation : une corde qui vibre ou la membrane d'un haut-parleur par exemple. Cette vibration provoque un mouvement des atomes l'avoisinant qui va se déplacer de proche en proche

sous forme d'onde de pression. La vitesse du son varie suivant le milieu dans lequel il se propage. Le principal facteur est la densité de ce milieu : dans un gaz, sa vitesse est plus faible que dans un liquide. Par exemple, le son se propage approximativement à 340 m/s dans l'air à 15°C, et à 1500 m/s (5400 km/h) dans l'eau.

On appelle $p(x, t)$ la pression du fluide au point x à l'instant t et c la vitesse du son dans le milieu qu'on suppose constante ici (le milieu est homogène). La fonction p est alors solution de l'équation des ondes dans \mathbb{R}^3 :

$$\frac{\partial^2 p}{\partial t^2} - c^2 \left(\frac{\partial^2 p}{\partial x^2} + \frac{\partial^2 p}{\partial y^2} + \frac{\partial^2 p}{\partial z^2} \right) = f(x, y, z, t)$$

où f est la source à l'origine du son.

EXEMPLE 1.1.6 (LES ÉQUATIONS DE MAXWELL)

Dans cet exemple, la solution recherchée n'est plus un scalaire mais un vecteur. Les inconnues de ce problème sont le champ électrique $E(x, t) \in \mathbb{R}^3$, le champ magnétique $H(x, t) \in \mathbb{R}^3$, l'induction électrique $D(x, t) \in \mathbb{R}^3$ et l'induction magnétique $B(x, t) \in \mathbb{R}^3$.

Ces champs obéissent, en l'absence de charges et de courants, aux équations de Maxwell :

$$\begin{cases} \frac{\partial B}{\partial t} = -\text{rot } E, \\ \frac{\partial D}{\partial t} = \text{rot } H, \end{cases}$$

et sont par ailleurs reliés par les lois de comportement :

$$\begin{cases} B(x, t) = \mu H(x, t), \\ D(x, t) = \varepsilon E(x, t), \end{cases}$$

où μ est la perméabilité magnétique et ε la permittivité diélectrique du milieu (constantes strictement positives dans le cas où le milieu est homogène). Ils caractérisent le comportement électromagnétique du matériau dans lequel l'onde se propage.

On montre facilement qu'on peut éliminer B , D et H des équations pour aboutir à une équation où seul le champ électrique E apparaît :

$$\varepsilon \frac{\partial^2 E}{\partial t^2} + \frac{1}{\mu} \text{rot} (\text{rot } E) = 0$$

ce qui correspond à une équation des ondes vectorielle dans \mathbb{R}^3 .

Si on suppose que E est de direction constante, on montre facilement qu'on retombe sur une équation d'onde scalaire.

1.2 La méthode des caractéristiques pour l'équation de transport

1.2.1 Le problème de Cauchy pour l'équation de transport linéaire

On considère ici le problème suivant : trouver $u(x, t) : \mathbb{R} \times \mathbb{R}^+ \rightarrow \mathbb{R}$ tel que

$$\begin{cases} \frac{\partial u}{\partial t} + c(x, t) \frac{\partial u}{\partial x} = 0, & x \in \mathbb{R}, \quad t > 0, \quad (i) \\ u(x, 0) = u^0(x), & x \in \mathbb{R}, \quad (ii) \end{cases} \quad (1.3)$$

où, au moins pour définir une solution classique (voir plus loin), on suppose que la donnée initiale u^0 est une fonction de classe C^1 (on peut admettre une régularité moindre, par exemple $u^0 \in L^\infty$ ou $u^0 \in L^2$, avec le concept de solution faible - voir remarque 1.2.3).

L'hypothèse sur la régularité du coefficient $c(x, t)$ est plus fondamentale. Nous supposons que $c(\cdot, \cdot)$ est une fonction continue de $x, t \in \mathbb{R} \times \mathbb{R}$ i. e. $c \in C^0(\mathbb{R} \times \mathbb{R})$ (nous autorisons le temps à être négatif) lipschitzienne en x , uniformément par rapport à t :

$$\exists L > 0 \quad / \quad \forall t \in \mathbb{R}, \quad \forall (x_1, x_2) \in \mathbb{R} \times \mathbb{R}, \quad |c(x_1, t) - c(x_2, t)| \leq L |x_1 - x_2|. \quad (1.4)$$

Pour simplifier un peu les aspects techniques des démonstrations, nous travaillerons avec l'hypothèse légèrement plus forte suivante :

$$\begin{aligned} &c(\cdot, \cdot) \text{ est continûment dérivable en } x \\ \text{et } &\exists L > 0 \quad / \quad \forall t \in \mathbb{R}, \quad \forall x \in \mathbb{R}, \quad \left| \frac{\partial c}{\partial x}(x, t) \right| \leq L. \end{aligned} \quad (1.5)$$

REMARQUE 1.2.1

Contrairement à ce qui se passait pour le coefficient de diffusion dans l'équation de la chaleur, le signe du coefficient c n'a aucune importance fondamentale, en tous cas aucune incidence sur l'existence ou non d'une solution. Il en est du même du signe ou du sens du temps. L'équation de transport peut s'intégrer indifféremment pour $t < 0$ ou $t > 0$. C'est une équation réversible.

REMARQUE 1.2.2

Rappelons que, le fait que $x \rightarrow c(x, t)$ soit lipschitzienne entraîne (c'est un résultat non trivial de théorie de la mesure) que la dérivée au sens des distributions

$$\frac{\partial c}{\partial x},$$

est une fonction L^∞ . En ce sens les hypothèses (1.4) et (1.5) sont très proches.

Solutions classiques du problème de Cauchy

DÉFINITION 1.2.1

On dit que u est une solution classique de l'équation (1.3) si c est une fonction \mathcal{C}^1 de x et de t et si elle satisfait (1.3) point par point.

Dans ce qui suit, nous allons montrer que le problème (1.3) admet une unique solution classique et que celle-ci peut être construite par la technique des caractéristiques.

REMARQUE 1.2.3 (SUR LA RÉGULARITÉ DES DONNÉES)

Le lecteur notera que, pour espérer l'existence d'une solution classique, il est nécessaire que la donnée initiale u^0 soit \mathcal{C}^1 et que la fonction $c(\cdot, \cdot)$ soit continue en ses deux variables. L'utilité de la condition de Lipschitz (1.4) (et même de (1.5)) (qui est toujours essentielle) n'apparaîtra qu'au moment où nous exposerons la méthode des caractéristiques. Lorsque u^0 est moins régulière ($u^0 \in L^p$ par exemple) on peut toujours donner un sens à grâce à la notion de solution faible. Ceci sera abordé plus en détail dans l'étude des équations hyperboliques non linéaires.

1.2.2 Résolution de l'équation de transport à coefficients constants

Dans le cas où $c(\cdot, \cdot)$ est une fonction constante, $c(x, t) = c$, l'opérateur de transport peut s'interpréter comme un opérateur de dérivation dans une direction oblique du plan (x, t) :

$$\frac{\partial}{\partial t} + c \frac{\partial}{\partial x} = \nabla_{x,t} \cdot \begin{pmatrix} c \\ 1 \end{pmatrix}, \quad \nabla_{x,t} = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial t} \right)^t.$$

Par conséquent, les solutions classiques de (1.3) sont constantes le long de la famille de droites parallèles au vecteur $(c, 1)^t$ du plan (x, t)

$$\mathcal{D}_\lambda = \left\{ (y, s) \in \mathbb{R} \times \mathbb{R} / y - cs = \lambda \right\}. \quad (1.6)$$

Par définition, les droites \mathcal{D}_λ sont la famille des droites caractéristiques de (1.3). Par tout point (x, t) , passe une droite caractéristique et une seule

$$\mathcal{D}(x, t) = \left\{ (y, s) \in \mathbb{R} \times \mathbb{R} / (y - x) - c(s - t) = 0 \right\}.$$

Bien entendu, on observe que

$$\mathcal{D}(x, t) = \mathcal{D}(x', t') \iff x - x' = c(t - t'), \quad (1.7)$$

alors que si on fixe t par exemple, l'ensemble des droites caractéristiques $\mathcal{D}(x, t)$ forme, quand x décrit \mathbb{R} , un ensemble de droites parallèles disjointes qui remplissent le plan.

$$\mathcal{D}(x, t) \cap \mathcal{D}(x', t) = \emptyset \quad \text{pour } x \neq x', \quad \bigcup_{x \in \mathbb{R}} \mathcal{D}(x, t) = \mathbb{R}^2. \quad (1.8)$$

Voyons maintenant comment construire u la solution de l'équation de transport (1.3) à l'aide des droites caractéristiques. Il nous faut bien évidemment utiliser la condition initiale. Fixons $t > 0$, pour calculer la valeur de la solution en (x, t) nous remarquons que $\mathcal{D}(x, t)$ coupe l'axe $t = 0$ en un point unique, appelé pied de la caractéristique issue de (x, t) à savoir :

$$(x_0, 0), \quad x_0 = x - ct.$$

La solution u de (1.3) étant, si elle existe, constante le long de $\mathcal{D}(x, t)$, cette solution est nécessairement donnée par :

$$u(x, t) = u(x - ct, 0) = u^0(x - ct). \quad (1.9)$$

A posteriori, on vérifie que la solution donnée par (1.9) est bien solution de (1.3).

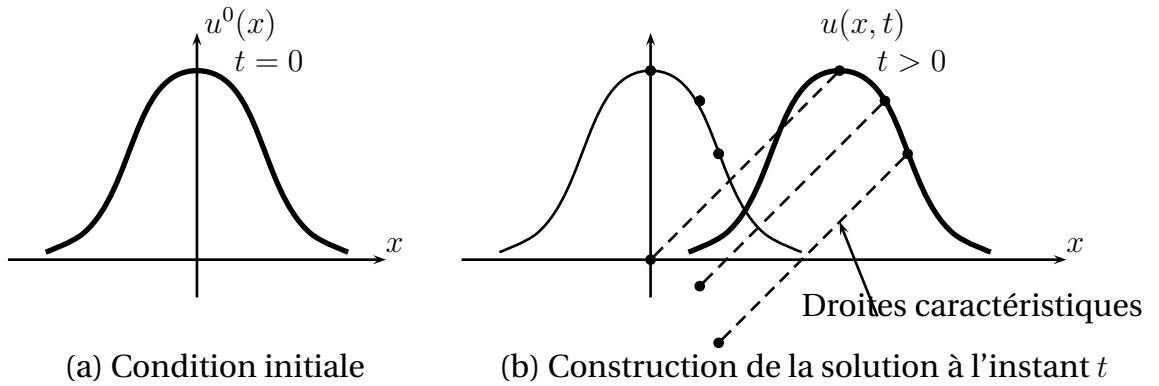


FIGURE 1.2 – Méthode des caractéristiques pour l'équation du transport

Nous avons établi le

THÉORÈME 1.2.1

Le problème admet une unique solution classique $u \in C^1(\mathbb{R} \times \mathbb{R}^+)$ donnée par (1.9).

Le graphe de la solution à l'instant t se déduit du graphe de la donnée initiale u^0 par une translation de longueur ct selon l'axe des abscisses : la formule (1.9) représente une fonction qui se propage (ou se transporte) sans déformation à la vitesse

constante c (voir figure (1.2)). Ceci justifie le fait que l'on appelle le coefficient c vitesse de l'équation de transport.

Il est immédiat à partir de (1.9) de vérifier les propriétés suivantes.

- **Propagation à vitesse finie.** De façon évidente

$$\text{supp } u^0 \subset [\alpha, \beta] \implies \text{supp } u(\cdot, t) \subset [\alpha + ct, \beta + ct].$$

- **Préservation de la régularité.** De façon tout à fait évidente, la régularité en la variable x de la solution $u(\cdot, t)$ est pour tout t la même que celle de u^0 .
- **Conservation de la norme L^p .** Du fait de l'invariance de l'intégrale par translation, on a évidemment

$$\forall p \in [1, +\infty], \quad \|u(\cdot, t)\|_{L^p} = \|u^0\|_{L^p}.$$

ce qui signifie que l'application linéaire $S(t) : u^0 \mapsto u(\cdot, t)$ est une isométrie dans tous les espaces L^p . On a donc aucun effet de dissipation.

Quand on évoque la conservation de la norme L^2 , on parle souvent de conservation de l'énergie. Ce résultat peut s'obtenir directement. En effet

$$\int_{-\infty}^{+\infty} \frac{\partial u}{\partial t} u \, dx + c \int_{-\infty}^{+\infty} \frac{\partial u}{\partial x} u \, dx = 0.$$

$$\text{Or } \int_{-\infty}^{+\infty} \frac{\partial u}{\partial x} u \, dx = \int_{-\infty}^{+\infty} \frac{\partial}{\partial x} \left[\frac{u^2}{2} \right] = 0 \quad \text{et par conséquent}$$

$$\frac{d}{dt} \int_{-\infty}^{+\infty} u^2 \, dx = 0.$$

On va utiliser la transformation de Fourier en espace, dont nous rappelons les principales propriétés.

Rappels sur la transformation de Fourier

Nous considérons la transformation de Fourier définie par :

$$f \in L^1(\mathbb{R}) \xrightarrow{\mathcal{F}} \hat{f} \in C^0(\mathbb{R}) \cap L^\infty(\mathbb{R}),$$

où \hat{f} est donnée par

$$\hat{f}(k) := \int_{-\infty}^{\infty} f(x) e^{-ikx} \, dx.$$

La transformation \mathcal{F} est injective et si $\hat{f} \in L^1(\mathbb{R})$, on a la formule d'inversion :

$$(\mathcal{F}^{-1}\hat{f})(x) = f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{f}(k) e^{ikx} dk.$$

On rappelle enfin que \mathcal{F} s'étend aux distributions en un isomorphisme de $\mathcal{S}'(\mathbb{R})$ et que

$$\mathcal{F}(\delta) = 1, \quad \mathcal{F}\left(\frac{\partial}{\partial x}\right) = i k \mathcal{F}.$$

Par ailleurs, il est bien connu que \mathcal{F} envoie $L^2(\mathbb{R})$ dans lui même et qu'on a le théorème de Plancherel :

$$\forall f \in L^2(\mathbb{R}), \quad \int_{-\infty}^{+\infty} |\hat{f}(k)|^2 dk = 2\pi \int_{-\infty}^{+\infty} |f(x)|^2 dx,$$

On rappelle enfin que, si on définit le produit de convolution de deux fonctions dans $L^2(\mathbb{R})$ par

$$(f * g)(x) := \int_{-\infty}^{\infty} f(x-y) g(y) dy = \int_{-\infty}^{\infty} f(y) g(x-y) dy, \quad \in L^\infty(\mathbb{R}),$$

alors

$$\mathcal{F}(f * g) = \hat{f} \hat{g}.$$

Enfin, nous rappelons les propriétés de la transformation de Fourier vis à vis des gaussiennes et d'un changement d'échelle

$$e^{-\frac{k^2}{2}} \xrightarrow{\mathcal{F}^{-1}} e^{-\frac{x^2}{2}} \quad \text{et} \quad \hat{u}(ak) \xrightarrow{\mathcal{F}^{-1}} \frac{1}{a} u\left(\frac{x}{a}\right) \quad (a \in \mathbb{R}),$$

et finalement des translations ($a \in \mathbb{R}$)

$$\forall f \in L^2(\mathbb{R}), \quad \tau_a f(x) = f(x+a) \implies \mathcal{F}(\tau_a f)(k) = e^{ika} \mathcal{F}f(k). \quad (1.10)$$

Résolution de l'équation de transport à coefficients constants par Fourier

Il est instructif de chercher à résoudre l'équation de transport (1.3) en utilisant la transformation de Fourier en espace :

$$u(x, t) \rightarrow \hat{u}(k, t).$$

En appliquant la transformation de Fourier spatiale à (1.3) on obtient l'équation différentielle ordinaire

$$\frac{d\hat{u}}{dt} + ick \hat{u} = 0,$$

ce qui donne, compte tenu de la condition initiale

$$\hat{u}(k, t) = e^{-ickt} \hat{u}^0(k).$$

Le lecteur notera alors que

- L'expression (1.9) de la solution exacte se retrouve en utilisant les propriétés de la transformation de Fourier vis à vis des translations.
- On a en particulier

$$|\widehat{u}(k, t)| = |\widehat{u}^0(k)|.$$

Grâce au théorème de Plancherel, on retrouve la conservation de la norme L^2 .

- On a également

$$\widehat{u}(k, t + \Delta t) = \widehat{S}(k, \Delta t) \widehat{u}(k, t), \quad \widehat{S}(k, \Delta t) = e^{-ick\Delta t}. \quad (1.11)$$

Le coefficient d'amplification complexe $\widehat{S}(k, \Delta t)$

- est de module 1 : c'est la manifestation du caractère non dissipatif de l'équation de transport à coefficients constants,
- a une phase proportionnelle à k : c'est la manifestation du caractère non dispersif de l'équation de transport à coefficients constants.

1.2.3 Résolution de l'équation de transport à coefficients variables

Cette fois, l'opérateur de transport à coefficients variables

$$\frac{\partial}{\partial t} + c(x, t) \frac{\partial}{\partial x}$$

n'est plus vraiment un opérateur de dérivation oblique. Par contre, il peut s'interpréter, comme nous allons le détailler, comme un opérateur de dérivation tangente le long d'une famille de courbes, appelées courbes caractéristiques, qui vont jouer le rôle des droites \mathcal{D}_λ dans le cas où c était constant. Géométriquement, ces courbes sont les courbes enveloppes du champ de vecteur $2D$

$$(x, t) \implies \begin{pmatrix} c(x, t) \\ 1 \end{pmatrix}.$$

Pour les introduire de façon plus analytique, cherchons à priori les courbes du plan (x, t) paramétrées par le temps, soit $\{ (X(s), s) \mid s \in \mathbb{R} \}$ où $X(s)$ est une fonction - supposée dérivable - à déterminer, telles que toute solution classique u de

$$\frac{\partial u}{\partial t} + c(x, t) \frac{\partial u}{\partial x} = 0 \quad (1.12)$$

soit constante le long de ces courbes. Nous étudions les variations de $U(s) = u(X(s), s)$ le long d'une telle courbe. Il vient

$$U'(s) = \frac{d}{ds} u(X(s), s) = \frac{\partial u}{\partial t} (X(s), s) + \frac{\partial u}{\partial x} (X(s), s) X'(s).$$

On voit, grâce à (1.12), que si on impose à X d'être solution de l'équation différentielle ordinaire :

$$X'(s) = c(X(s), s), \quad (1.13)$$

alors on aura

$$U'(s) = 0 \implies u(X(s), s) = \text{constante}.$$

Les courbes cherchées sont donc les courbes intégrales (ou trajectoires) de l'équation différentielle (1.13), ce qui établit un lien intime entre la résolution des équations de transport linéaires et la théorie des équations différentielles. Comme on a affaire à une équation différentielle ordinaire du premier ordre, on s'attend à ce qu'il existe une trajectoire et une seule passant par un point (x, t) : cela revient à ajouter à (1.13) une condition « initiale » en $s = t$. Cela nous amène naturellement à introduire une fonction X non plus de la seule variable s mais des trois variables $(s; x, t)$ de la façon suivante : (x, t) étant fixés et jouant le rôle de paramètres, $s \rightarrow X(s; x, t)$ est la solution du problème de Cauchy :

$$\begin{cases} \frac{\partial X}{\partial s}(s; x, t) = c(X(s; x, t), s), & s \in \mathbb{R}, & (i) \\ X(t; x, t) = x. & & (ii) \end{cases} \quad (1.14)$$

Notons que nous cherchons à intégrer (1.13) aussi bien pour $s < t$ que $s > t$.

L'hypothèse (1.4) est précisément l'hypothèse qui nous permet d'appliquer le célèbre théorème de Cauchy-Lipschitz relatif à la résolution des équations différentielles ordinaires :

LEMME 1.2.2

Pour tout $(x_0, t_0) \in \mathbb{R} \times \mathbb{R}$, le problème (1.14) admet une unique solution

$$s \mapsto X(s; x, t) \in C^1(\mathbb{R}).$$

L'application $X(\cdot; \cdot, \cdot)$ ainsi définie est appelée flot de l'équation différentielle (1.13). Par ailleurs, sous l'hypothèse (1.5), on peut démontrer (mais ce n'est pas si évident que cela) que X est également une fonction de classe C^1 par rapport aux variables x et t et que les fonctions

$$\begin{cases} s \rightarrow \dot{X}_x(s; x, t) := \frac{\partial X}{\partial x}(s; x, t), \\ s \rightarrow \dot{X}_t(s; x, t) := \frac{\partial X}{\partial t}(s; x, t), \end{cases}$$

sont respectivement les solutions des équations différentielles linéaires

$$\begin{cases} \frac{\partial \dot{X}_x}{\partial s}(s; x, t) - \frac{\partial c}{\partial x}(X(s; x, t), s) \dot{X}_x(s; x, t) = 0, & s \in \mathbb{R}, \\ \dot{X}_x(t; x, t) = 1, \end{cases} \quad (1.15)$$

et

$$\begin{cases} \frac{\partial \dot{X}_t}{\partial s}(s; x, t) - \frac{\partial c}{\partial x}(X(s; x, t), s) \dot{X}_t(s; x, t) = 0, & s \in \mathbb{R}, \\ \dot{X}_t(t; x, t) = -c(x, t). \end{cases} \quad (1.16)$$

A titre d'exercice instructif, nous invitons le lecteur à établir, au moins formellement, les équations (1.15) et (1.16). Ces équations s'intègrent « à la main » et on obtient :

$$\begin{cases} \frac{\partial X}{\partial x}(s; x, t) = \exp \left(\int_t^s \frac{\partial c}{\partial x}(X(\tau; x, t), \tau) d\tau \right), & (i) \\ \frac{\partial X}{\partial t}(s; x, t) = -c(x, t) \exp \left(\int_t^s \frac{\partial c}{\partial x}(X(\tau; x, t), \tau) d\tau \right). & (ii) \end{cases} \quad (1.17)$$

En particulier on voit que que s étant fixé la fonction $(x, t) \rightarrow X(s; x, t)$ vérifie une équation aux dérivées partielles qui n'est autre que l'équation de transport,

$$\forall x \in \mathbb{R}, \quad \forall (s, t) \in \mathbb{R} \times \mathbb{R}, \quad \frac{\partial X}{\partial t}(s; x, t) + c(x, t) \frac{\partial X}{\partial x}(s; x, t) = 0. \quad (1.18)$$

REMARQUE 1.2.4

L'expression (1.17) montre que

$$\frac{\partial X}{\partial x}(s; x, t) > 0,$$

et par conséquent que, pour tout $(s, t) \in \mathbb{R} \times \mathbb{R}$, la fonction $x \rightarrow X(s; x, t)$ est strictement croissante.

REMARQUE 1.2.5

Si c ne vérifie que la condition (1.4), on peut montrer que X admet presque partout des dérivées partielles par rapport aux variables x et t qui sont toujours caractérisées par les équations (1.15) et (1.16), la dérivée en x de c étant prise au sens des distributions.

Nous appellerons courbes caractéristiques de (1.12) les courbes intégrales de (1.14), c'est à dire les courbes du plan (y, s) (paramétrées par s) définies par :

$$\mathcal{C}(x, t) = \left\{ (y, s) \in \mathbb{R} \times \mathbb{R} / y = X(s; x, t), s \in \mathbb{R} \right\}.$$

La courbe $\mathcal{C}(x, t)$ est la caractéristique de (1.12) qui passe par le point (x, t) . C'est l'équivalent de la droite $\mathcal{D}(x, t)$ du paragraphe précédent (cf. (1.6)).

Les propriétés (1.7) et (1.8) sont donc les équivalents pour les courbes $\mathcal{C}(x, t)$ de celles que nous avons rencontrés pour les des droites $\mathcal{D}(x, t)$. Elles se traduisent au travers

de celles de la fonction X et leur démonstration est simplement un exercice d'application (répétée) du théorème de Cauchy-Lipschitz et principalement du résultat d'unicité associé.

La première propriété précise ce qui est presque une tautologie, à savoir que deux courbes caractéristiques $\mathcal{C}(x, t)$ et $\mathcal{C}(x', t')$ coïncident si et seulement si (x, t) et (x', t') appartiennent à une même trajectoire de l'équation différentielle (1.13) :

$$\mathcal{C}(x, t) = \mathcal{C}(x', t') \iff x' = X(t'; x, t) \iff x = X(t; x', t'). \quad (1.19)$$

En termes du flot X , cela traduit la relation fondamentale :

$$\forall (t, t') \in \mathbb{R}^2, \quad \forall s \in \mathbb{R}, \quad X(s; x, t) = X(s; X(t'; x, t), t'). \quad (1.20)$$

Cette propriété exprime une propriété (assez intuitive d'ailleurs) dite de semi-groupe de l'équation différentielle (1.13) qui exprime le fait que intégrer (1.13) de s à t en partant de la position x équivaut à intégrer (1.13) d'abord entre t et t' pour calculer $X(t'; x, t)$, puis repartir de t' pour aller jusqu'au temps s à partir de la nouvelle position « initiale » $X(t'; x, t)$ au temps t' .

Par ailleurs, à un instant t donné, l'ensemble des courbes caractéristiques $\mathcal{C}(x, t)$ forme, quand x décrit \mathbb{R} un ensemble de courbes disjointes qui remplissent le plan (on dit qu'elles forment un fibrage du plan).

$$\mathcal{C}(x, t) \cap \mathcal{C}(x', t) = \emptyset \quad \text{pour } x \neq x', \quad \bigcup_{x \in \mathbb{R}} \mathcal{C}(x, t) = \mathbb{R}^2. \quad (1.21)$$

Cela se traduit en termes de l'application X par la propriété que

$$\text{Pour tout } (s, t), \text{ l'application } X(s; \cdot, t) \text{ est une bijection de } \mathbb{R} \text{ dans } \mathbb{R}. \quad (1.22)$$

On en connaît en outre l'inverse :

$$\forall (s, t) \in \mathbb{R} \times \mathbb{R}, \quad X(s; \cdot, t)^{-1} = X(t; \cdot, s), \quad (1.23)$$

ce qui exprime que pour revenir de x à x , on peut tout d'abord aller au point $X(s; x, t)$ en intégrant de t à s en partant du point x à l'instant t , puis revenir en x en intégrant en marche arrière de s à t , en partant du point $X(s; x, t)$ à l'instant s .

PREUVE DES PROPRIÉTÉS (1.19) À (1.23):

Démontrer (1.20) est essentiellement un jeu d'écriture. Fixons t, t' et x et introduisons les deux fonctions de la variable réelle s

$$Y_1(s) = X(s; x, t), \quad Y_2(s) = X(s; X(t'; x, t), t').$$

Par définition de $X(\cdot; \cdot, \cdot)$ (voir 1.14), Y_1 et Y_2 satisfont la même équation différentielle du premier ordre :

$$\frac{dY_1}{ds}(s) = c(Y_1(s), s), \quad \frac{dY_2}{ds}(s) = c(Y_2(s), s).$$

De plus si on regarde en $s = t'$, il vient grâce à (1.14-(ii))

$$Y_2(t') = X(t'; X(t'; x, t), t') = X(t'; x, t) = Y_1(t').$$

Grâce au théorème de Cauchy-Lipschitz (et plus précisément le résultat d'unicité), nous savons que Y_1 et Y_2 coïncident pour tout s , ce que nous voulions démontrer.

Démontrons la première équivalence de (1.19). La seconde se déduit de (1.23).

L'égalité $C(x, t) = C(x', t')$ entraîne en particulier que (x', t') appartient à $C(x, t)$ ce qui veut dire que $x' = X(t'; x, t)$.

La réciproque repose sur (1.20). En effet, si $x' = X(t'; x, t)$

$$\left| \begin{aligned} C(x', t') &= \{(s, X(s; x', t')), s \in \mathbb{R}\}, && \text{(par définition)} \\ &= \{(s, X(s; X(t'; x, t), t')), s \in \mathbb{R}\}, && \text{(par hypothèse)} \\ &= \{(s, X(s; x, t)), s \in \mathbb{R}\}, && \text{(par (1.20))} \\ &= C(x, t). && \text{(par définition)} \end{aligned} \right.$$

L'injectivité de $X(s; \cdot, t)$ résulte du fait que cette fonction est strictement croissante (cf. remarque 1.2.4). On peut aussi la voir comme une conséquence de Cauchy-Lipschitz : si $X(s; x, t) = X(s; x', t)$, les fonctions $\tau \rightarrow X(\tau; x, t)$ et $\tau \rightarrow X(\tau; x', t)$ seraient deux solutions de (1.13) qui coïncideraient en $\tau = s$. D'après le théorème d'unicité on en déduit qu'elles sont égales pour tout τ et donc pour $\tau = t$ ce qui donne $x = x'$ compte tenu de (1.14-(ii)).

Si nous choisissons $t = s$ et $t' = t$ dans (1.20) nous obtenons, compte tenu de (1.14-(ii)),

$$x = X(s; X(t; x, s), t),$$

ce qui démontre à la fois la surjectivité de $X(s; \cdot, t)$ et la formule (1.23).

Finalement, la propriété (1.21) ne fait que traduire géométriquement la bijectivité des applications $X(s; \cdot, t)$. ■

On peut retrouver (1.18) à partir de (1.20). En effet, dérivons (1.20) par rapport à t' , (x, s, t) étant fixés. En utilisant le théorème de dérivée composée, il vient :

$$0 = \frac{\partial X}{\partial t}(s; X(t'; x, t), t') + \frac{\partial X}{\partial x}(s; X(t'; x, t), t') \frac{\partial X}{\partial t}(t'; x, t),$$

c'est dire, compte tenu de (1.14-(i))

$$\frac{\partial X}{\partial t}(s; X(t'; x, t), t') + c(X(t'; x, t), t') \frac{\partial X}{\partial x}(s; X(t'; x, t), t') = 0.$$

En choisissant $t' = t$ et en utilisant (1.14-(ii)), nous aboutissons à (1.18).

Nous pouvons maintenant mettre en place la méthode des caractéristiques. Supposons que (1.3) ait une solution classique. Pour calculer sa valeur en (x, t) , nous utilisons le fait que u est constante le long de la caractéristique $\mathcal{C}(x, t)$. Autrement dit :

$$u(x, t) = u(X(s; x, t), s) \quad \forall s \in \mathbb{R}.$$

En particulier, en prenant $s = 0$, on introduit naturellement le pied de la caractéristique $\mathcal{C}(x, t)$, à savoir $(X(0; x, t), 0)$ et on obtient en utilisant la condition initiale :

$$u(x, t) = u^0(X(0; x, t)) . \quad (1.24)$$

Il reste pour conclure à vérifier que la fonction u donnée par (1.24) est bien solution. C'est bien une fonction de classe C^1 en tant que composée de deux fonctions de classe C^1 . Par ailleurs,

$$\left[\frac{\partial u}{\partial t} + c \frac{\partial u}{\partial x} \right] (x, t) = u^0(X(0; x, t)) \left[\frac{\partial X}{\partial t}(0; x, t) + c(X(0; x, t)) \frac{\partial X}{\partial x}(0; x, t) \right] = 0,$$

grâce à (1.18). Nous avons donc démontré le

THÉORÈME 1.2.3

Sous les hypothèses (1.4), le problème admet une unique solution classique donnée par (1.24).

On peut démontrer aisément à partir de la formule que certaines propriétés de l'équation de transport à coefficients constants sont préservées.

COROLLAIRE 1.2.4

La solution u se propage à vitesse finie

$$\text{supp } u^0 \subset [\alpha, \beta] \implies \text{supp } u(\cdot, t) \subset [X(t; \alpha, 0), X(t; \beta, 0)] \quad (1.25)$$

Elle vérifie la propriété de conservation

$$\forall t \geq 0, \quad \sup_{x \in \mathbb{R}} u(x, t) = \sup_{x \in \mathbb{R}} u^0(x), \quad \inf_{x \in \mathbb{R}} u(x, t) = \inf_{x \in \mathbb{R}} u^0(x), \quad \|u(\cdot, t)\|_{L^\infty} = \|u^0\|_{L^\infty} \quad (1.26)$$

En revanche certaines propriétés sont perdues. Ainsi, comme les caractéristiques peuvent s'écarter ou se rapprocher au cours du temps, la solution ne se propage pas à vitesse constante (elle peut « ralentir » ou « accélérer ») et le profil (ou graphe) de la solution $u(\cdot, t)$ va se déformer au cours du temps. De ce fait, on n'a plus en général conservation de la norme L^p de la solution pour $p < +\infty$.

1.3 Systèmes hyperboliques linéaires

1.3.1 Systèmes hyperboliques linéaires

Pour simplifier l'exposé, nous nous limiterons dans ce qui suit aux systèmes hyperboliques linéaires à coefficients constants.

Considérons un système $n \times n$ du premier ordre sous la forme suivante :

$$\frac{\partial}{\partial t} u_i + \sum_{k=1}^d \sum_{j=1}^n a_{ij}^{(k)} \frac{\partial}{\partial x_k} u_j = 0, \quad i \in \{1, 2, \dots, n\}. \quad (1.27)$$

On cherche ici la solution $(u_1(\mathbf{x}, t), u_2(\mathbf{x}, t), \dots, u_n(\mathbf{x}, t))$, avec \mathbf{x} qui, dans les cas les plus simples, est un scalaire ($\mathbf{x} \in \mathbb{R}, d = 1$), un vecteur du plan ($\mathbf{x} \in \mathbb{R}^2, d = 2$) ou de l'espace ($\mathbf{x} \in \mathbb{R}^3, d = 3$). La donnée est une famille de matrices :

$$A^{(k)} = (a_{ij}^{(k)}) \in M_{n \times n}, \quad \forall k \in \{1, \dots, d\}.$$

Par exemple, pour le cas $n = 2$ et $d = 1$, on retombe sur l'équation des ondes 1D (1.2) avec :

$$A = \begin{bmatrix} 0 & -c \\ -c & 0 \end{bmatrix}$$

La complexité de cette structure à trois indices peut s'éclaircir en introduisant la notion d'onde plane.

DÉFINITION 1.3.1

Une onde plane $v(y, t)$ de direction $\nu = (\nu_1, \dots, \nu_d) \in \mathbb{S}^{d-1}$ (ensemble des vecteurs de \mathbb{R}^d de norme 1) est une solution ne dépendant que d'une variable d'espace, de direction ν :

$$u(x, t) = v(x \cdot \nu, t) \in \mathbb{R}^d$$

LEMME 1.3.1

Une onde plane solution du système (1.27) satisfait, pour $1 \leq i \leq n$, l'équation monodimensionnelle

$$\frac{\partial}{\partial t} v_i(y, t) + \sum_{j=1}^n a_{ij}(\nu) \frac{\partial}{\partial y} v_j(y, t) = 0,$$

avec la notation

$$a_{ij}(\nu) = \sum_{k=1}^d \nu_k a_{ij}^{(k)}.$$

Par la suite, on notera également

$$A(\nu) = \sum_{k=1}^d \nu_k A^{(k)} \in M_{n \times n}.$$

DÉFINITION 1.3.2

Considérons le système linéaire du premier ordre (1.27),

1. ce système est dit *hyperbolique* si $\forall \nu \in \mathbb{S}^{d-1}$, la matrice $A(\nu)$ est diagonalisable dans \mathbb{R} . On note $\lambda_l(\nu)$, $1 \leq l \leq n$ ses valeurs propres ordonnées,
2. ce système est dit *strictement hyperbolique* si les n valeurs propres sont distinctes, on a alors $\lambda_1(\nu) < \dots < \lambda_n(\nu)$.

REMARQUE 1.3.1

Pour un système hyperbolique $n \times n$ en dimension un ($d = 1$), il existe une matrice Λ diagonale et une matrice de passage P telle que $A = P\Lambda P^{-1}$, avec :

$$\Lambda = \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix}.$$

Posons $v(x, t) = P^{-1}u(x, t)$ et $v^0(x) = P^{-1}u^0(x)$, on peut alors écrire n équations de transport découplées dont les inconnues $v_j(x, t)$ sont les composantes de $v(x, t)$:

$$\begin{cases} \frac{\partial}{\partial t} v_j + \lambda_j \frac{\partial}{\partial x} v_j = 0, & j = 1, 2, \dots, n, \\ v_j(x, 0) = v_j^0(x). \end{cases}$$

Pour tout $u^0 \in C^1(\mathbb{R})^n$, l'unique solution classique de ce problème de Cauchy sera donc donnée par :

$$\begin{cases} u(x, t) = Pv(x, t), \\ v_j(x, t) = v_j^0(x - \lambda_j t), \\ v^0(x) = P^{-1}u^0(x). \end{cases}$$

Ainsi, l'étude des systèmes hyperboliques linéaires en dimension un se ramène facilement à l'étude de plusieurs équations de transport découplées.

REMARQUE 1.3.2

En dimension d d'espace, cette base dépend de la direction ν de l'onde plane et il n'est que très rarement vrai que l'on puisse trouver une base commune. Considérer par exemple l'équation des ondes en dimension $d > 1$.

1.3.2 L'équation des ondes en dimension 1

Nous allons donner ici des résultats spécifiques à l'équation des ondes 1D :

$$\begin{cases} \frac{\partial^2 u}{\partial t^2} - c^2 \frac{\partial^2 u}{\partial x^2} = 0, & t \in \mathbb{R}, x \in \mathbb{R}, \\ u(x, 0) = u^0(x), \\ \frac{\partial u}{\partial t}(x, 0) = u^1(x). \end{cases} \quad (1.28)$$

La formule de D'Alembert**THÉORÈME 1.3.2**

Toute solution de l'équation des ondes homogène (1.28) se décompose sous la forme

$$u(x, t) = u^+(x, t) + u^-(x, t).$$

- $u^+(x, t) = f(x - ct)$ est une onde progressive se propageant à la vitesse c dans la direction $x > 0$:

$$u^+(x + L, t) = u^+(x, t - \frac{L}{c})$$

(le graphe de $x \rightarrow u^+(x, t + T)$ se déduit de celui de $x \rightarrow u^+(x, t)$ par une simple translation de $L = cT$ vers la droite). Une telle solution est constante sur les droites $x - ct = Cte$. Cette famille de droites constitue ce que l'on appelle la première famille de courbes caractéristiques associée à l'équation des ondes (1.28).

- $u^-(x, t) = g(x + ct)$ est une onde progressive se propageant à la vitesse c dans la direction $x < 0$:

$$u^-(x - L, t) = u^-(x, t - \frac{L}{c})$$

(le graphe de $x \rightarrow u^-(x, t + T)$ se déduit de celui de $x \rightarrow u^-(x, t)$ par une simple translation de $L = cT$ vers la gauche). Une telle solution est constante sur les droites $x + ct = Cte$ qui constitue la seconde famille de courbes caractéristiques associée à l'équation des ondes.

PREUVE: Nous avons vu au début de ce chapitre que l'équation des ondes (1.28) se met sous forme hyperbolique :

$$\frac{\partial U}{\partial t} + A \frac{\partial U}{\partial x} = 0$$

où :

$$U = \begin{pmatrix} \frac{\partial u}{\partial t} \\ \frac{\partial u}{\partial x} \end{pmatrix} \quad \text{et} \quad A = \begin{bmatrix} 0 & -c \\ -c & 0 \end{bmatrix}$$

On diagonalise aisément sous la forme $A = P\Lambda P^{-1}$, où :

$$\Lambda = \begin{bmatrix} c & 0 \\ 0 & -c \end{bmatrix} \quad P = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \quad P^{-1} = \begin{bmatrix} 1/2 & -1/2 \\ 1/2 & 1/2 \end{bmatrix}$$

Avec $V(x, t) = P^{-1}U(x, t)$ et $V_0(x) = P^{-1}U_0(x)$, on a de manière évidente :

$$V_0(x) = \frac{1}{2} \begin{pmatrix} u^1(x) - c \frac{\partial u^0}{\partial x}(x) \\ u^1(x) + c \frac{\partial u^0}{\partial x}(x) \end{pmatrix}$$

et par application de la méthode des caractéristiques :

$$V(x, t) = \frac{1}{2} \begin{pmatrix} u^1(x - ct) - c \frac{\partial u^0}{\partial x}(x - ct) \\ u^1(x + ct) + c \frac{\partial u^0}{\partial x}(x + ct) \end{pmatrix}$$

d'où l'on déduit immédiatement que :

$$U(x, t) = \frac{1}{2} \begin{pmatrix} u^1(x - ct) + u^1(x + ct) - c \frac{\partial u^0}{\partial x}(x - ct) + c \frac{\partial u^0}{\partial x}(x + ct) \\ -u^1(x - ct) + u^1(x + ct) + c \frac{\partial u^0}{\partial x}(x - ct) + c \frac{\partial u^0}{\partial x}(x + ct) \end{pmatrix}$$

Alors, en intégrant par rapport au temps la première composante du vecteur U ci-dessus, on obtient :

$$u(x, t) - u^0(x) = \frac{1}{2} \int_0^t u^1(x + cs) ds + \frac{1}{2} \int_0^t u^1(x - cs) ds + \frac{c}{2} \int_0^t \frac{\partial u^0}{\partial x}(x + cs) ds - \frac{c}{2} \int_0^t \frac{\partial u^0}{\partial x}(x - cs) ds,$$

soit, après deux changements de variable évidents :

$$u(x, t) = \frac{1}{2c} \int_x^{x+ct} u^1(s) ds - \frac{1}{2c} \int_x^{x-ct} u^1(s) ds + \frac{1}{2} u^0(x + ct) + \frac{1}{2} u^0(x - ct),$$

et donc la décomposition demandée avec :

$$u^+(x, t) = \frac{1}{2} u^0(x - ct) + \frac{1}{2c} \int_{x-ct}^x u^1(s) ds \quad \text{et} \quad u^-(x, t) = \frac{1}{2} u^0(x + ct) + \frac{1}{2c} \int_x^{x+ct} u^1(s) ds,$$

ce qui donne la célèbre formule de d'Alembert :

$$u(x, t) = \frac{1}{2} (u^0(x + ct) + u^0(x - ct)) + \frac{1}{2c} \int_{x-ct}^{x+ct} u^1(s) ds. \quad (1.29)$$

■

Propagation à vitesse finie

La formule (1.29) montre que la valeur de la solution u au point x à l'instant t ne dépend que des valeurs des données initiales dans l'intervalle $[x - ct; x + ct]$ qui est aussi la base du cône caractéristique (ou cône de dépendance) $D(x, t)$ issu du point (x, t) c'est à dire le cône délimité par les deux droites caractéristiques de l'équation des ondes qui passent par le point (x, t) (voir figure 1.3) :

$$D(x, t) = \left\{ (y, s) , 0 \leq s \leq t, |y - x| \leq c(t - s) \right\}.$$

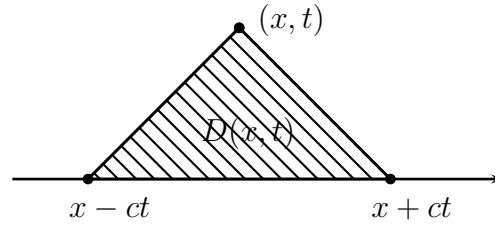


FIGURE 1.3 – Le cône de dépendance

On en déduit que la solution se propage à vitesse finie au sens où, si on part des données à support compact, la solution reste à support compact à tout instant. Plus précisément (voir la figure 1.4) :

$$\text{Supp } u^0 \cup \text{Supp } u^1 \subset [a, b] \Rightarrow \text{Supp } u(\cdot, t) \subset [a - ct; b + ct]$$

Conservation de l'énergie

Il est naturel d'associer à toute solution u de (1.28) la densité d'énergie :

$$e(x, t) = \frac{1}{2} \left(\left| \frac{\partial u}{\partial t} \right|^2(x, t) + c^2 \left| \frac{\partial u}{\partial x} \right|^2(x, t) \right).$$

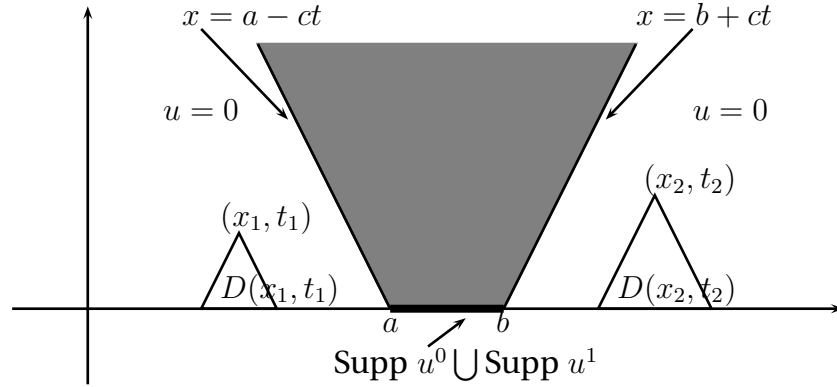


FIGURE 1.4 – Propagation à vitesse finie

On dit que la solution u est d'énergie finie (les solutions qui ont un sens physique sont généralement d'énergie finie) dès que :

$$E(t) = \int_{\mathbb{R}} e(x, t) dx < +\infty,$$

où $E(t)$ est par définition l'énergie totale de la solution.

THÉORÈME 1.3.3

Il y a conservation d'énergie au cours du temps :

$$\frac{\partial}{\partial t} E(t) = 0.$$

PREUVE: En multipliant l'équation des ondes par $\frac{\partial u}{\partial t}$, on obtient :

$$\frac{\partial^2 u}{\partial t^2} \frac{\partial u}{\partial t} - c^2 \frac{\partial^2 u}{\partial x^2} \frac{\partial u}{\partial t} = 0.$$

Nous observons ensuite que :

$$\int_{\mathbb{R}} \frac{\partial^2 u}{\partial t^2} \frac{\partial u}{\partial t} dx = \frac{1}{2} \frac{d}{dt} \left(\int_{\mathbb{R}} \left| \frac{\partial u}{\partial t} \right|^2 dx \right),$$

alors que, moyennant une intégration par parties :

$$-c^2 \int_{\mathbb{R}} \frac{\partial^2 u}{\partial x^2} \frac{\partial u}{\partial t} dx = \frac{c^2}{2} \frac{d}{dt} \left(\int_{\mathbb{R}} \left| \frac{\partial u}{\partial x} \right|^2 dx \right),$$

Ainsi en ajoutant ces deux dernières égalités, on obtient :

$$\frac{\partial}{\partial t} E(t) = 0.$$

■

Chapitre 2

Approximation par différences finies - Application aux problèmes hyperboliques linéaires

2.1 Présentation de la méthode de différences finies

2.1.1 Approximation de quelques opérateurs différentiels

La méthode de différences finies permet de calculer une approximation de la solution d'une EDP en des points qui sont distribués sur une grille. L'objectif est alors de construire des approximations des dérivées des fonctions intervenant dans l'EDP à l'aide de valeurs discrètes de celle-ci, par le biais de formules de Taylor.

Nous nous plaçons, sauf mention contraire, dans le cadre d'une discrétisation par différences finies en une dimension d'espace. Ceci implique que nous discrétisons le continuum spatio-temporel par une grille régulière de pas Δt en temps et h en espace tel que les coordonnées discrètes soient :

$$(x_j, t^n) = (jh, n\Delta t).$$

La solution discrète sera calculée en ces points (voir la figure 2.1). Le principe des différences finies est donc de remplacer les dérivées en ces points par des différences finies en utilisant des formules de Taylor dont on néglige les restes.

Pour commencer, reprenons la définition usuelle de la dérivée

$$u'(x) := \lim_{h \rightarrow 0} \frac{u(x+h) - u(x)}{h},$$

afin d'introduire *l'approximation décentrée à droite*

$$u'(x) \simeq D_h^+ u(x) := \frac{u(x+h) - u(x)}{h}.$$

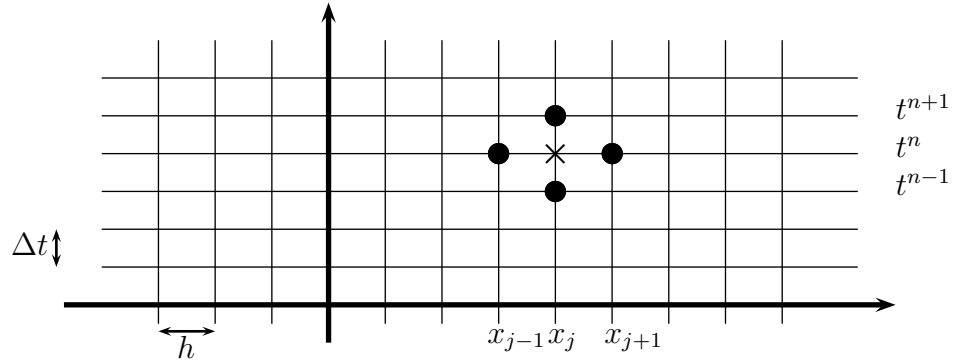


FIGURE 2.1 – Grille espace temps

Evidemment, il est, *a priori*, tout aussi légitime de définir une *approximation décentrée à gauche*

$$u'(x) \simeq D_h^- u(x) := \frac{u(x) - u(x-h)}{h},$$

ou encore une *approximation centrée*

$$u'(x) \simeq D_h^0 u(x) := \frac{u(x+h) - u(x-h)}{2h}.$$

Nous allons voir que ce choix n'est en fait pas anodin et que certains choix *a priori* raisonnables ne permettent pas une bonne approximation (dans un sens qu'il nous faudra définir) de la solution de l'EDP initiale.

La première obligation pour notre approximation est *la consistance*. Elle exprime le fait que l'opérateur aux différences finies converge vers l'opérateur différentiel lorsque le pas de maillage tend vers 0. Ainsi, pour un maillage de pas h , on note D_h un opérateur d'approximation aux différences finies de $\frac{d}{dx}$ et on appelle *erreur de troncature* pour cet opérateur la quantité

$$\varepsilon(x, h) := D_h u(x) - u'(x).$$

L'approximation est *consistante* pour la norme $\|\cdot\|$, si pour u suffisamment régulière

$$\lim_{h \rightarrow 0} \|\varepsilon(x, h)\| = 0.$$

On peut même aller plus loin en disant que l'approximation est *d'ordre* $k \in \mathbb{N}$ si

$$\varepsilon(x, h) = O(h^k).$$

En pratique, l'analyse de consistance s'effectue en considérant les développements de Taylor de la fonction u . Par exemple pour l'approximation décentrée à droite on sait pour $u \in C^2(I)$, qu'il existe $\theta \in]0, 1[$ tel que

$$u(x+h) = u(x) + h u'(x) + \frac{h^2}{2} u''(x + \theta h).$$

On obtient ainsi

$$\frac{u(x+h) - u(x)}{h} - u'(x) = \frac{h}{2} u''(x + \theta h),$$

d'où l'estimation

$$\varepsilon(x, h) \leq \frac{h}{2} \sup |u''|. \quad (2.1)$$

On aurait obtenu le même type d'inégalité pour D_h^- donc les approximations décentrées à gauche et à droite sont d'ordre 1. En pratique, on voit que l'inégalité (2.1) nous indique même que le choix de h doit être déterminé par la valeur de la dérivée seconde.

Pour l'approximation centrée, les calculs sont un peu plus longs mais pour un bénéfice sur l'ordre de l'approximation. On considère $u \in C^3(I)$. Il existe $\theta^+ \in]0, 1[$ tel que

$$u(x+h) = u(x) + h u'(x) + \frac{h^2}{2} u''(x + \theta h) + \frac{h^3}{6} u^{(3)}(x + \theta^+ h)$$

et $\theta^- \in]0, 1[$ tel que

$$u(x-h) = u(x) - h u'(x) + \frac{h^2}{2} u''(x + \theta h) - \frac{h^3}{6} u^{(3)}(x + \theta^- h)$$

ce qui donne par soustraction

$$\frac{u(x+h) - u(x-h)}{2h} - u'(x) = \frac{h^2}{3} \left(u^{(3)}(x + \theta^+ h) + u^{(3)}(x + \theta^- h) \right).$$

D'après le théorème des valeurs intermédiaires, il existe $\theta \in]-1, 1[$ tel que

$$\frac{u(x+h) - u(x-h)}{2h} - u'(x) = \frac{h^2}{6} u^{(3)}(x + \theta h),$$

et l'approximation centrée est donc d'ordre 2.

Ces trois approximations sont loin d'être les seules et il existe en fait une infinité de possibilités d'approximation dont on jugera la pertinence notamment en fonction de l'erreur de consistance. Il est en particulier possible de faire intervenir plus de points dans D_h afin d'obtenir une approximation plus précise. Par exemple si nous combinons les approximations suivantes (obtenus comme précédemment mais en choisissant désormais les formules de Taylor-Young au lieu de celles de Taylor-Lagrange), nous obtenons

$$\frac{u(x+h) - u(x-h)}{2h} = u'(x) + \frac{h^2}{6} u^{(3)}(x) + O(h^4),$$

et

$$\frac{u(x+2h) - u(x-2h)}{4h} = u'(x) + 4 \frac{h^2}{6} u^{(3)}(x) + O(h^4),$$

ce qui donne pour tout coefficient de pondération $\alpha \in [0, 1]$

$$\begin{aligned} D_h^{(4)} u &= \alpha \frac{u(x+h) - u(x-h)}{2h} + (1-\alpha) \frac{u(x+2h) - u(x-2h)}{4h} \\ &= u'(x) + (\alpha + 4(1-\alpha)) \frac{h^2}{6} u^{(3)}(x) + O(h^4). \end{aligned}$$

Ainsi pour un choix adéquate $\alpha = \frac{4}{3}$ de pondération entre les deux approximations initiales, on obtient une approximation d'ordre 4.

Afin de présenter les différents choix d'opérateurs d'approximation D_h , on adoptera une représentation symbolique beaucoup plus visuelle comme utilisée Figure 2.2 pour les exemples d'approximation que nous venons de présenter.

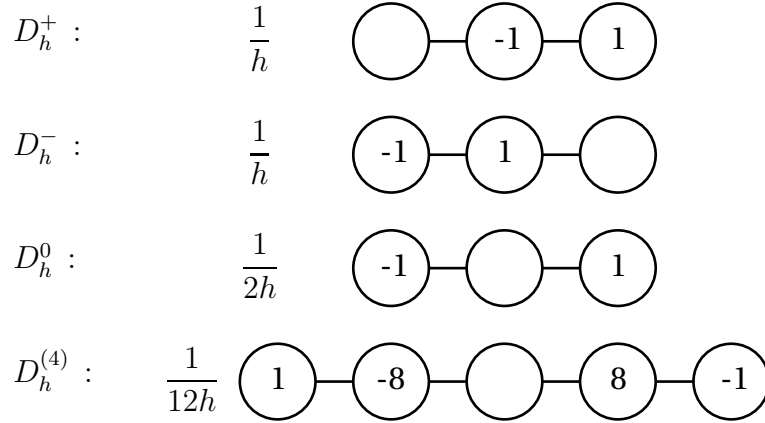


FIGURE 2.2 – Représentation symbolique des opérateurs aux différences finies

De tout évidence la construction d'opérateurs d'approximation ne se limite pas aux dérivées d'ordre 1. On utilise le même principe pour les dérivées d'ordre supérieur. Notamment, on a

$$u''(x) \simeq \frac{u'(x+h/2) - u'(x-h/2)}{h},$$

avec par exemple

$$u'(x+h/2) \simeq \frac{u'(x+h) - u'(x-h)}{h},$$

et

$$u'(x-h/2) \simeq \frac{u'(x+h) - u'(x-h)}{h}.$$

On peut ainsi définir un opérateur D_h^2 d'approximation aux différences finies de la dérivée seconde sous la forme

$$u''(x) \simeq D_h^2 u(x) := \frac{u(x+h) - 2u(x) + u(x-h)}{h^2}.$$

La consistance de cet opérateur se détermine toujours à partir de formule de Taylor (ici Taylor-Lagrange). Il existe $\theta^+ \in]0, 1[$ tel que

$$u(x + h) = u(x) + h u'(x) + \frac{h^2}{2} u''(x) + \frac{h^3}{6} u^{(3)}(x) + \frac{h^4}{24} u^{(4)}(x + \theta^+ h),$$

et $\theta^+ \in]0, 1[$ tel que

$$u(x - h) = u(x) - h u'(x) + \frac{h^2}{2} u''(x) - \frac{h^3}{6} u^{(3)}(x) + \frac{h^4}{24} u^{(4)}(x + \theta^+ h).$$

On a donc l'existence de $\theta \in]-1, 1[$ tel que

$$D_h^2 u(x) = \frac{u(x + h) - 2u(x) + u(x - h))}{h^2} = u''(x) + \frac{h^2}{12} u^{(4)}(x + \theta h), \quad (2.2)$$

et l'approximation est donc d'ordre 2.

2.1.2 La démarche : Etude de la consistance et la stabilité

Pour une équation aux dérivées partielles donnée, l'idée est donc de construire une équation discrète où les opérateurs différentiels sont remplacés par des opérateurs d'approximation. La solution continue n'a *a priori* pas de raison d'être solution de cette équation discrète.

La question qui se pose donc naturellement est de savoir si, pour une équation discrète particulière (on parlera également de schéma) obtenue en utilisant la méthode de différences finies décrite ci-dessus, la solution de l'équation discrète obtenue approche de la solution du problème continu et d'autre part quand il y a convergence, de quantifier la vitesse de convergence.

Cette analyse se fait toujours en deux étapes fondamentales qui seront donc le dénominateur commun de chaque étude de schéma de cet ouvrage :

- **La consistance** qui correspond à l'erreur d'approximation liées lorsqu'on remplace l'EDP par des différences finies. Il est fondamental de comprendre que, bien que très intuitive en tant que condition nécessaire, la consistance d'une méthode ne suffit pas à obtenir la convergence.
- **La stabilité** qui assure que l'opérateur discret est bien inversible et la norme de son inverse est bornée indépendamment du pas de discrétisation.

On montrera alors que dès qu'un schéma est consistant et stable alors il est convergent, c'est à dire que la solution numérique calculée à partir de ce schéma tend, quand les paramètres de discrétisation tendent vers 0, vers la solution de l'équation aux dérivées partielles initiale.

L'objectif de ce chapitre étant de présenter et d'analyser quelques schémas classiques pour les équations hyperboliques linéaires, nous étudierons donc par différentes approches la consistance et la stabilité des schémas avant d'obtenir le résultat final de convergence.

2.1.3 Application aux équations hyperboliques linéaires

Avant de passer à la construction de schémas de discrétisation particuliers, indiquons dès à présent une particularité de la résolution des problèmes d'évolution : avec la plupart des méthodes numériques couramment utilisées, la résolution ne s'effectue pas de la même manière en temps et en espace. En effet il existe classiquement deux situations bien distinctes suivant que l'EDP fait intervenir ou non le temps :

- Si on résout par exemple l'équation de Laplace en dimension 2

$$-\frac{\partial^2 u}{\partial x^2} - \frac{\partial^2 u}{\partial y^2} = f \quad \text{dans } \mathbb{R}^2,$$

on peut montrer que la méthode des différences finies conduit à inverser un système linéaire dans lequel les deux variables x et y jouent le même rôle : d'une certaine façon, toutes les valeurs de la solution approchée sont calculées simultanément.

- Avec un problème d'évolution en dimension 1 d'espace, bien que la solution soit à nouveau une fonction de deux variables, le temps et l'espace sont traités différemment. Cela se comprend intuitivement puisque le temps est naturellement orienté et que nous disposons donc de conditions initiales et pas de conditions finales alors que les directions x et $-x$ jouent le même rôle. De ce fait, la solution approchée se calculera pas à pas en temps : la solution à l'instant t^n est déterminée à partir de la solution aux instants précédents. Les conditions initiales sont alors utilisées pour déterminer la solution aux premiers pas de temps. En pratique, pour les problèmes du premier ordre en temps, comme c'est le cas de l'équation d'advection par exemple, on pourra se contenter d'un schéma à un pas de temps ; la solution à l'instant t^n est calculée à partir de la solution à l'instant précédent t^{n-1} . A un instant donné t^n :
 - soit les valeurs u_j^n , $j \in \mathbb{Z}$ sont calculées indépendamment les unes des autres : on parle alors de schéma *explicite*,
 - soit les valeurs u_j^n , $j \in \mathbb{Z}$ sont calculées en résolvant un système d'équations couplées. On parle alors de schéma *implicite*.

Afin de présenter les différents schémas numériques possibles pour les équations hyperboliques, nous allons considérer les deux exemples principaux que sont :

- l'équation d'advection sur \mathbb{R} avec une vitesse constante c

$$\begin{cases} \frac{\partial u}{\partial t} + c \frac{\partial u}{\partial x} = 0, & (x, t) \in \mathbb{R} \times]0, T[, \\ u(0, x) = u^0(x), & x \in \mathbb{R}. \end{cases} \quad (2.3)$$

- l'équation des ondes sur \mathbb{R} avec une vitesse constante c

$$\begin{cases} \frac{\partial^2 u}{\partial t^2} - c^2 \frac{\partial^2 u}{\partial x^2} = 0, & \text{pour } t \geq 0, x \in \mathbb{R} \\ u(t = 0, x) = u^0(x), & \text{pour } x \in \mathbb{R} \\ \frac{\partial u}{\partial t}(t = 0, x) = u^1(x), & \text{pour } x \in \mathbb{R} \end{cases}$$

Dans toute la suite et sauf mention contraire, nous supposons la vitesse de propagation c positive.

Comme annoncé précédemment, nous approchons la solution u de ces EDPs par une solution discrète u_j^n . On cherche alors à calculer pour chaque n ($n \in \mathbb{N}$, $n\Delta t \leq T$) un vecteur $u^n = (u_j^n)_{j \in \mathbb{Z}}$ tel que u_j^n soit une approximation de la valeur de la solution u du problème précédent en x_j à l'instant t^n .

REMARQUE 2.1.1

Il peut sembler surprenant de recourir à une résolution numérique approchée pour ces problèmes où, la vitesse étant constant, on connaît analytiquement les solutions exactes. L'intérêt est de calculer dans des cas simples les propriétés des différents schémas qui seront ensuite généralisés au cas non linéaire. Un schéma qui ne permet pas de résoudre le cas linéaire sera automatiquement à proscrire dans le cas non linéaire.

Nous commencerons par l'équation d'advection et présenterons le *schéma explicite centré* qui semble le plus naturel. À partir de celui ci nous introduirons les notions fondamentales pour l'étude des schémas aux différences finies (schéma explicite/implicite, consistance et erreur de troncature et stabilité). Nous verrons alors que ce schéma est inconditionnellement instable et donc qu'*a priori* la solution discrète ne peut converger vers la solution du problème continu. Cet exemple aura pour but de sensibiliser le lecteur sur le fait qu'un schéma consistant n'est pas nécessairement convergent. Nous étudierons ensuite le *schéma de Lax-Friedrichs* pour lequel

nous introduirons la notion de convergence. Le schéma saute-mouton sera enfin étudié comme approximation de l'équation des ondes.

2.2 Le schéma explicite centré pour l'équation d'advection

Nous allons écrire pour tout instant t^n , $n > 0$ et tout point x_j , une première équation discrète qui va remplacer (et approcher) l'équation (2.3) en (x_j, t^n) . On utilise

- une approximation décentrée à droite (vers le futur), d'ordre 1, pour $\frac{\partial u}{\partial t}$

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

- une approximation centrée, d'ordre 2, pour $\frac{\partial u}{\partial x}$, écrite à l'instant t^n (voir Section 2.1) :

$$\frac{\partial u}{\partial x}(x_j, t_n) \simeq D_h^0 u(x_j, t_n) = \frac{u_{j+1}^n - u_{j-1}^n}{2h}.$$

Le schéma numérique au point (x_j, t^n) s'écrit donc

$$\frac{u_j^{n+1} - u_j^n}{\Delta t} + c \frac{u_{j+1}^n - u_{j-1}^n}{2h} = 0, \quad j \in \mathbb{Z}, \quad n \geq 0. \quad (2.4)$$

Il s'agit d'un schéma à un pas (ou deux points) en temps et à trois points en espace. Ce schéma est bien explicite puisqu'on calcule explicitement la solution à l'instant $n+1$ à partir de la solution à l'instant n sans recourir à l'inversion d'un système linéaire. On a en effet

$$u_j^{n+1} = u_j^n + \frac{\alpha}{2}(u_{j+1}^n - u_{j-1}^n), \quad j \in \mathbb{Z},$$

où

$$\alpha = \frac{c\Delta t}{h}.$$

Evidemment, pour que (2.4) définisse entièrement la solution approchée, il suffit de connaître les u_j^0 , $j \in \mathbb{Z}$, ce pour quoi on va utiliser la condition initiale et on choisira typiquement (noter que le choix (i) nécessite que u^0 soit une fonction continue)

$$(i) \quad u_j^0 = u^0(x_j) \quad \text{ou} \quad (ii) \quad u_j^0 = \frac{1}{h} \int_{x_j - \frac{h}{2}}^{x_j + \frac{h}{2}} u^0(x) dx. \quad (2.5)$$

Pour l'analyse, il est commode d'associer à une suite $(u_j)_{j \in \mathbb{Z}}$ à valeurs réelles, une fonction de la variable x . Nous choisissons dans la suite l'interpolation linéaire par

morceaux (\mathbb{P}_1 désigne ci-après l'espace des polynômes d'une variable de degré inférieur ou égal à 1). Ainsi pour une suite $(u_j)_{j \in \mathbb{Z}}$ on associe

$$u_h \in V_h = \{ v_h \in C^0(\mathbb{R}) / \forall j \in \mathbb{Z}, v_h|_{[x_j, x_{j+1}]} \in \mathbb{P}_1 \},$$

telle que u_h est définie par

$$\forall j \in \mathbb{Z}, \quad u_h(x_j) = u_j.$$

On peut démontrer que, pour tout réel $p > 1$,

$$(u_j)_{j \in \mathbb{Z}} \in \ell^p = \left\{ (u_j)_{j \in \mathbb{Z}} / \sum_{j \in \mathbb{Z}} |u_j|^p < +\infty \right\} \iff u_h \in L_h^p = V_h \cap L^p(\mathbb{R}).$$

et qu'il existe des constantes $C_*(p)$ et $C^*(p)$, indépendantes de h , telles que

$$\forall (u_j) \in \ell^p, \quad C_*(p) \left(\sum_{j \in \mathbb{Z}} |u_j|^p h \right)^{\frac{1}{p}} \leq \|u_h\|_{L^p} \leq C^*(p) \left(\sum_{j \in \mathbb{Z}} |u_j|^p h \right)^{\frac{1}{p}}.$$

De même

$$(u_j)_{j \in \mathbb{Z}} \in \ell^\infty = \left\{ (u_j) / \sup_{j \in \mathbb{Z}} |u_j| < +\infty \right\} \iff u_h \in L_h^\infty = V_h \cap L^\infty(\mathbb{R}).$$

et on a l'identité

$$\|u_h\|_{L^\infty} = \sup_{j \in \mathbb{Z}} |u_j|.$$

Nous allons pouvoir réécrire (2.4) de façon abstraite en introduisant l'opérateur linéaire A_h tel que

$$v(x) : \mathbb{R} \rightarrow \mathbb{R}, \quad \mapsto \quad A_h v(x) = c \frac{v(x+h) - v(x-h)}{2h}.$$

Puisque A_h commute avec les opérateurs de translation par un multiple de h , il envoie linéairement V_h dans lui-même donc

$$A_h \in \mathcal{L}(V_h).$$

Mieux, on démontre facilement que

$$\forall p \in [1, +\infty], \quad A_h \in \mathcal{L}(L^p) \quad \text{et} \quad \|A_h\|_{\mathcal{L}(L^p)} = \frac{\|A_1\|_{\mathcal{L}(L^p)}}{h^2}.$$

On voit alors que, avec des notations évidentes, le schéma (2.4, 2.5-(i)) équivaut à

$$\text{Trouver } u_h^n \in V_h, \quad n = 0, 1, 2, \dots \quad \text{tels que}$$

$$\frac{u_h^{n+1} - u_h^n}{\Delta t} + A_h u_h^n = 0, \quad n = 0, 1, 2, \dots, \quad (2.6)$$

et satisfaisant la condition initiale

$$u_h^0 = \Pi_h u^0, \quad (2.7)$$

où Π_h désigne l'opérateur d'interpolation linéaire de $C^0(\mathbb{R})$ dans V_h .

Il est bien évident que, compte tenu des propriétés de A_h ,

$$u_h^0 \in L^p \implies u_h^n \in L^p, \quad n = 1, 2, \dots$$

La solution de (2.6) s'écrit donc

$$u_h^n = (S_h(\Delta t))^n u_h^0 \quad \text{avec} \quad S_h(\Delta t) = I - \Delta t A_h. \quad (2.8)$$

2.2.1 Consistance et ordre du schéma

Rappelons que l'étude de la consistance d'un schéma vise à justifier (et quantifier) la façon dont les équations discrètes approchent les équations continues. Elle ne dit rien sur la façon dont la solution discrète approche la solution continue. Pour cela il faut aussi avoir la stabilité du schéma, concept qui sera abordé dans la prochaine section. Il est néanmoins bien évidemment nécessaire de vérifier la consistance d'un schéma pour espérer sa convergence.

Pour ce faire, il faut définir l'erreur de troncature du schéma numérique. Le principe consiste à partir de la solution exacte $u(x, t)$ de l'équation d'advection (2.3), supposée régulière, d'introduire la suite des valeurs discrètes

$$U_j^n = u(x_j, t^n). \quad (2.9)$$

Si on remplace u_j^n par U_j^n dans le membre de gauche de (2.4), on n'a *a priori* aucune raison de trouver 0 comme résultat (sinon le schéma serait « exact »), mais on est en droit d'exiger que cela soit « petit ». Ce raisonnement justifie la

DÉFINITION 2.2.1

On appelle *erreur de troncature du schéma (2.4)* la suite

$$\varepsilon_j^n = \frac{U_j^{n+1} - U_j^n}{\Delta t} + c \frac{U_{j+1}^n - U_{j-1}^n}{2h},$$

où U_j^n est définie par (2.9), u étant une solution régulière de (2.3).

On dira que le schéma est consistant si, pour tout segment borné I et tout temps $T > 0$

$$\lim_{(h, \Delta t) \rightarrow 0} \varepsilon_j^n = 0.$$

Étant donnés deux entiers strictement positifs k et p , on dit que le schéma est d'ordre p en temps et k en espace (ce qui entraîne *a fortiori* la consistance) si

$$\varepsilon_j^n = \mathcal{O}(\Delta t^p + h^k).$$

THÉORÈME 2.2.1

Le schéma (2.4) est d'ordre 1 en temps et d'ordre 2 en espace.

DÉMONSTRATION: C'est une application des développements de Taylor. Il existe $\theta_t \in]0, 1[$ tel que :

$$U_j^{n+1} = U_j^n + \Delta t \left(\frac{\partial u}{\partial t} \right)_j^n + \frac{\Delta t^2}{2} \frac{\partial^2 u}{\partial t^2}(x_j, t + \theta_t \Delta t),$$

où nous avons adopté la notation

$$\left(\frac{\partial u}{\partial t} \right)_j^n = \frac{\partial u}{\partial t}(x_j, t^n), \quad \left(\frac{\partial u}{\partial x} \right)_j^n = \frac{\partial u}{\partial x}(x_j, t^n), \quad \dots \quad (2.10)$$

Par conséquent

$$\frac{U_j^{n+1} - U_j^n}{\Delta t} = \left(\frac{\partial u}{\partial t} \right)_j^n + \frac{\Delta t}{2} \frac{\partial^2 u}{\partial t^2}(x_j, t + \theta_t \Delta t). \quad (2.11)$$

De même, en développant par rapport à la variable x et avec la notation (2.10), il existe $\theta^+ \in]0, 1[$ et $\theta^- \in]0, 1[$ tels que

$$\begin{cases} U_{j+1}^n = U_j^n + h \left(\frac{\partial u}{\partial x} \right)_j^n + \frac{h^2}{2} \left(\frac{\partial^2 u}{\partial x^2} \right)_j^n + \frac{h^3}{6} \frac{\partial^3 u}{\partial x^3}(x_j + \theta^+ h, t^n), \\ U_{j-1}^n = U_j^n - h \left(\frac{\partial u}{\partial x} \right)_j^n + \frac{h^2}{2} \left(\frac{\partial^2 u}{\partial x^2} \right)_j^n - \frac{h^3}{6} \frac{\partial^3 u}{\partial x^3}(x_j - \theta^- h, t^n). \end{cases}$$

Après soustraction nous obtenons

$$\frac{U_{j+1}^n - U_{j-1}^n}{2h} = \left(\frac{\partial u}{\partial x} \right)_j^n + \frac{h^2}{12} \left[\frac{\partial^3 u}{\partial x^3}(x_j + \theta^+ h, t^n) + \frac{\partial^3 u}{\partial x^3}(x_j - \theta^- h, t^n) \right].$$

D'après le théorème des valeurs intermédiaires, il existe donc $\theta \in]-1, 1[$ tel que

$$\frac{U_{j+1}^n - U_{j-1}^n}{2h} = \left(\frac{\partial u}{\partial x} \right)_j^n + \frac{h^2}{6} \frac{\partial^3 u}{\partial x^3}(x_j + \theta h, t^n) \quad (2.12)$$

De (2.11) et (2.12), nous déduisons, par définition de la solution u , que

$$\begin{cases} \varepsilon_j^n = \left(\frac{\partial u}{\partial t} \right)_j^n + c \left(\frac{\partial u}{\partial x} \right)_j^n + \frac{\Delta t}{2} \frac{\partial^2 u}{\partial t^2}(x_j, t + \theta_t \Delta t) + c \frac{h^2}{6} \frac{\partial^3 u}{\partial x^3}(x_j + \theta h, t^n) \\ \quad = \frac{\Delta t}{2} \frac{\partial^2 u}{\partial t^2}(x_j, t + \theta_t \Delta t) + c \frac{h^2}{6} \frac{\partial^3 u}{\partial x^3}(x_j + \theta h, t^n). \end{cases} \quad (2.13)$$

Ainsi nous obtenons

$$\varepsilon_j^n = \mathcal{O}(\Delta t + h^2).$$

■

Pour l'analyse de convergence, il sera utile d'introduire une version « continue en espace » de l'erreur de troncature, à savoir, en reprenant les notations de la section précédente,

$$\varepsilon_h^n = \frac{U_h^{n+1} - U_h^n}{\Delta t} + A_h U_h^n.$$

où $U_h^n = \Pi_h u(\cdot, t^n)$. Reprenant l'expression (2.13), on déduit une estimation L^∞ de ε_h^n (avec une constante numérique C convenable)

$$\|\varepsilon_h^n\|_{L^\infty} \leq C \left(\Delta t \sup_{t \in [t^n, t^{n+1}]} \left\| \frac{\partial^2 u}{\partial t^2}(\cdot, t) \right\|_{L^\infty} + h^2 \left\| \frac{\partial^3 u}{\partial x^3}(\cdot, t^n) \right\|_{L^\infty} \right). \quad (2.14)$$

Nous pouvons également obtenir une estimation équivalente de ε_h^n pour la norme L^2 , c'est à dire qu'il existe une constante C telle que

$$\|\varepsilon_h^n\|_{L^2} \leq C \left(\Delta t \sup_{t \in [t^n, t^{n+1}]} \left\| \frac{\partial^2 u}{\partial t^2}(\cdot, t) \right\|_{L^2} + h^2 \left\| \frac{\partial^3 u}{\partial x^3}(\cdot, t^n) \right\|_{L^2} \right). \quad (2.15)$$

Le lecteur se convaincra aisément que les estimations (2.14) et (2.15) se généralisent à toute norme L^p (la constante C ci-dessous dépendant de p)

$$\|\varepsilon_h^n\|_{L^p} \leq C \left(\Delta t \sup_{t \in [t^n, t^{n+1}]} \left\| \frac{\partial^2 u}{\partial t^2}(\cdot, t) \right\|_{L^p} + h^2 \left\| \frac{\partial^3 u}{\partial x^3}(\cdot, t^n) \right\|_{L^p} \right).$$

2.2.2 Propagation numérique et condition nécessaire de convergence

La formule (2.4) (ou (2.8)) démontre que le support de la solution numérique se propage à vitesse finie au sens où :

$$\text{supp } u_h^0 \subset [a, b] \implies \text{supp } u_h^n \subset [a - nh, b + nh].$$

On peut même dire que la propagation du support de la solution numérique est donc donnée par

$$V_{num}(\Delta, h) = \frac{h}{\Delta t}.$$

Si le support de la solution numérique se propage moins vite que le support de la solution exacte, la solution numérique n'aura aucune chance de s'approcher de la solution exacte puisqu'il va exister une partie du support de la solution exacte sur laquelle elle sera toujours nulle (voir figure 2.3-(a)).

Une condition nécessaire de convergence du schéma est donc que la vitesse numérique doit être supérieure à la vitesse de propagation de la solution exacte (voir figure 2.3-(b)) :

$$\frac{h}{\Delta t} \geq c \iff \alpha = \frac{c\Delta t}{h} \leq 1. \quad (2.16)$$

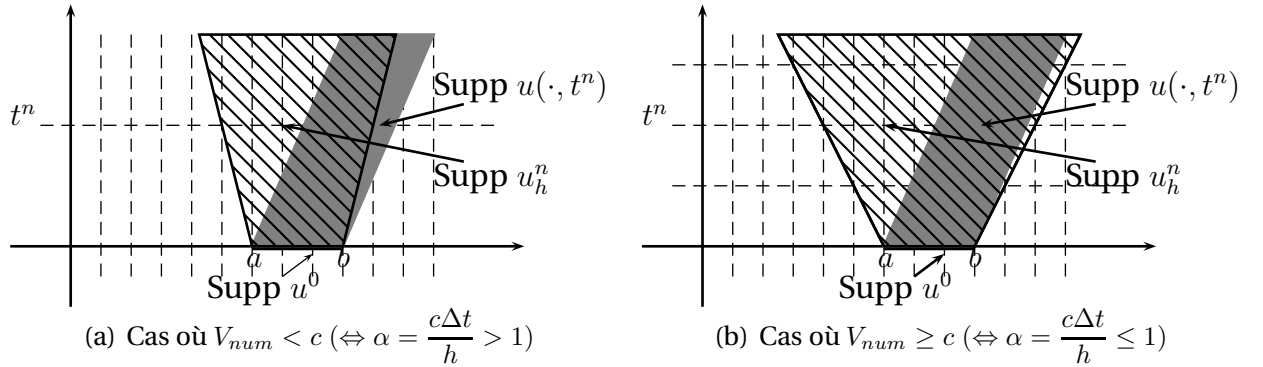


FIGURE 2.3 – Condition nécessaire de convergence : comparaison du support de $u(\cdot, t)$ (zone grise) et du support de u_h^n (zone hachurée).

REMARQUE 2.2.1 (REMARQUE SUR LA CONDITION NÉCESSAIRE DE CONVERGENCE)

La condition (2.16) a une interprétation géométrique simple. En effet, pour l'équation de transport, on sait que la connaissance de u à l'instant t^n dans l'intervalle $[x_{j-1}, x_{j+1}]$ permet de calculer la solution dans l'intervalle

$$x_{j-1} - ct^n \leq x - ct \leq x_{j+1} - ct^n.$$

La condition (2.16) signifie donc que le point (x_j, t_{n+1}) doit se trouver dans la bande grisée de la Figure 2.4.

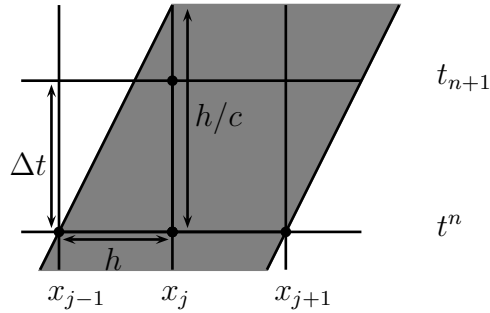


FIGURE 2.4 – Condition (2.16) : $\Delta t \leq \frac{h}{c}$

2.2.3 Analyse de la stabilité d'un schéma : principes généraux

Avant d'aborder la notion de stabilité du schéma numérique, une notion intuitivement simple mais délicate à définir mathématiquement, il est utile de préciser ce que nous entendons par convergence du schéma. Intuitivement, la convergence signifie que, dans une certaine norme, l'erreur entre la solution exacte et la solution approchée tend vers 0 quand Δt et h tendent vers 0. C'est le sens plus précis du « tend vers 0 » qu'il nous faut définir.

On supposera dans toute la suite que nous nous intéressons à l'approximation de la solution dans un intervalle de temps fixe $[0, T]$, $T > 0$. Nous nous limiterons à la convergence L^∞ en temps à valeur L^p en espace.

DÉFINITION 2.2.2

On dit qu'un schéma d'approximation de (2.3) produisant une suite $u_h^n \in L^p$ converge dans $L^\infty(0, T; L^p)$ si et seulement si, pour toute donnée initiale u^0 dans L^p ,

$$\lim_{(h, \Delta t) \rightarrow 0} \sup_{t^n \leq T} \|u(\cdot, t^n) - u_h^n\|_{L^p} = 0.$$

Selon le principe qu'une condition nécessaire pour qu'une suite converge est qu'elle soit bornée, nous introduisons le principe de stabilité L^p (qui devrait s'appeler plutôt stabilité $L^\infty(0, T; L^p)$).

DÉFINITION 2.2.3

On dit qu'un schéma numérique est L^p -stable si et seulement si, pour tout temps $T > 0$, il existe une constante $C(T)$ indépendante de h et Δt , telle que pour toute donnée initiale u^0 dans L^p ,

$$\forall t^n \leq T, \quad \|u_h^n\|_{L^p} \leq C(T) \|u^0\|_{L^p}. \quad (2.17)$$

REMARQUE 2.2.2

Comme on le verra, le résultat de stabilité (2.17) (ou encore le résultat de convergence) peut n'être satisfait que sous une condition de stabilité de la forme

$$(h, \Delta t) \in \mathbf{S} \subset \mathbb{R}_+^* \times \mathbb{R}_+^*,$$

où \mathbf{S} définit le domaine de stabilité de la méthode. Évidemment tout ceci n'a de sens que si l'ensemble \mathbf{S} admet $(0, 0)$ comme point d'accumulation.

Dans ce qui suit, nous allons nous intéresser aux schémas numériques à un pas produisant une suite u_h^n satisfaisant une relation de récurrence du type

$$u_h^{n+1} = S_h(\Delta t) u_h^n, \quad S_h(\Delta t) \in \mathcal{L}(L^p). \quad (2.18)$$

Le schéma explicite (2.6) entre donc dans cette catégorie.

REMARQUE 2.2.3

Un schéma consistant de la forme (2.18) vérifie, au moins formellement

$$\lim_{\Delta t \rightarrow 0} S_h(\Delta t) = I.$$

REMARQUE 2.2.4

Démontrer un résultat de stabilité L^p pour un schéma à un pas du type (2.18) correspond à l'obtention d'une estimation uniforme en $(h, \Delta t)$ de

$$\|S_h(\Delta t)^n\|_{\mathcal{L}(L^p)} = \sup_{v \in L^p} \frac{\|S_h(\Delta t)^n v\|_{L^p}}{\|v\|_{L^p}}.$$

En effet, la définition équivaut à

$$\forall t^n \leq T, \quad \|S_h(\Delta t)^n\|_{\mathcal{L}(L^p)} \leq C(T). \quad (2.19)$$

Une condition suffisante de stabilité, très utile dans la pratique, est donnée par le

THÉORÈME 2.2.2

Une condition suffisante de stabilité L^p du schéma (2.18) est qu'il existe une constante $\nu \geq 0$ telle que

$$\forall (\Delta t, h) \in \mathbf{S}, \quad \|S_h(\Delta t)\|_{\mathcal{L}(L^p)} \leq 1 + \nu \Delta t. \quad (2.20)$$

Auquel cas on a

$$\|S_h(\Delta t)^n\|_{\mathcal{L}(L^p)} \leq e^{\nu t^n},$$

ce qui entraîne en particulier (2.17) avec $C(T) = e^{\nu T}$.

DÉMONSTRATION: Par définition de $\|S_h(\Delta t)\|_{\mathcal{L}(L^p)}$, il est clair que

$$\|S_h(\Delta t)^n\|_{\mathcal{L}(L^p)} \leq (1 + \nu \Delta t)^n.$$

Pour conclure, il suffit d'utiliser le résultat suivant : pour tout $x > 0$ et tout entier n

$$\left(1 + \frac{x}{n}\right)^n \leq e^x.$$

En effet, pour $n = 1$, l'inégalité $1 + x \leq e^x$ traduit simplement le fait que le graphe de la fonction $x \mapsto e^x$ est au dessus de sa tangente au point d'abscisse $x = 0$. Plus généralement, la fonction $x \mapsto e^x$ étant croissante, il vient

$$1 + \frac{x}{n} \leq e^{\frac{x}{n}} \implies \left(1 + \frac{x}{n}\right)^n \leq e^x.$$

■

Notons que si la condition (2.20) est réalisée avec $\nu = 0$, c'est à dire si

$$\forall (\Delta t, h) \in \mathbf{S}, \quad \|S_h(\Delta t)\|_{\mathcal{L}(L^p)} \leq 1,$$

alors on a

$$\|u_h^n\|_{L^p} \leq \|u_h^0\|_{L^p},$$

qui est l'équivalent discret de

$$\|u(\cdot, t)\|_{L^p} \leq \|u^0\|_{L^p}.$$

On parlera alors de *stabilité uniforme*.

Signalons ainsi que le théorème 2.2.2 admet une réciproque, dont la démonstration, qui nécessite un peu de connaissances sur la théorie spectrale, dépasse le cadre de ce cours.

THÉORÈME 2.2.3

Pour qu'un schéma numérique du type (2.18) soit L^p stable, il faut qu'il existe une constante $\nu \geq 0$ telle que

$$\forall (\Delta t, h) \in \mathbf{S}, \quad \|S_h(\Delta t)\|_{\mathcal{L}(L^p)} \leq 1 + \nu \Delta t. \quad (2.21)$$

2.2.4 Analyse de stabilité L^2 du schéma explicite centré : la méthode de Fourier-Von Neumann

Le schéma explicite (2.6) entre dans la classe (2.18) avec

$$S_h(\Delta t) = I - \Delta t A_h. \quad (2.22)$$

Pour étudier $\|S_h(\Delta t)\|_{\mathcal{L}(L^2)}$, nous utilisons la transformation de Fourier et introduisons l'opérateur

$$\widehat{S}_h(\Delta t) = \mathcal{F} S_h(\Delta t) \mathcal{F}^{-1} \in \mathcal{L}(L^2). \quad (2.23)$$

On montre alors qu'il existe $\widehat{S}_h(k, \Delta t) \in L^\infty(\mathbb{R}, \mathbb{C})$ tel que

$$\forall k \in \mathbb{R}, \quad \widehat{u}_h^{n+1}(k) = (\widehat{S}_h(\Delta t) \widehat{u}_h^n)(k) = \widehat{S}_h(k, \Delta t) \widehat{u}_h^n(k), \quad (2.24)$$

où on a posé $\widehat{u}_h^n = \mathcal{F} u_h^n$.

Grâce au théorème de Plancherel, nous avons

$$\|\widehat{S}_h(\Delta t)\|_{\mathcal{L}(L^2)} = \|S_h(\Delta t)\|_{\mathcal{L}(L^2)}.$$

Le théorème 2.2.4 va nous donner une autre expression pour $\|S_h(\Delta t)\|_{\mathcal{L}(L^2)}$ qui sera beaucoup utilisée en pratique. Pour cela, il est utile d'introduire quelques définitions et énoncer (sans preuve) quelques résultats de base (que nous invitons le lecteur à démontrer à titre d'exercice) :

- Étant donné $\widehat{b}(k) \in L^\infty(\mathbb{R}, \mathbb{C})$, on appelle opérateur de multiplication par \widehat{b} , l'opérateur \widehat{B} , linéaire et continu de $L^p(\mathbb{R}; \mathbb{C})$ dans lui-même défini par :

$$\forall v \in L^p(\mathbb{R}, \mathbb{C}), \quad (\widehat{B}v)(k) = \widehat{b}(k) v(k), \quad \text{p.p. } k \in \mathbb{R}.$$

C'est un exercice très simple de montrer que :

$$\forall p \in [1, +\infty], \quad \|\widehat{B}\|_{\mathcal{L}(L^p)} = \|\widehat{b}\|_{L^\infty} \quad (2.25)$$

- Dans le cas $p = 2$, on appelle opérateur de symbole \widehat{b} , l'opérateur

$$B = \mathcal{F} \widehat{B} \mathcal{F}^{-1} \in \mathcal{L}(L^2).$$

Du théorème de Plancherel et de (2.25), nous déduisons aisément que

$$\|B\|_{\mathcal{L}(L^2)} = \|\widehat{b}\|_{L^\infty} \quad (2.26)$$

D'après les relations (2.23-2.24), $S_h(\Delta t)$ est un opérateur à symbole, on en déduit donc d'après (2.26)

THÉORÈME 2.2.4

$$\forall \Delta t, \quad \|S_h(\Delta t)\|_{\mathcal{L}(L^2)} = \sup_{k \in \mathbb{R}} |\widehat{S}_h(k, \Delta t)|$$

Il nous reste donc à étudier la norme L^∞ de $k \mapsto \widehat{S}_h(k, \Delta t)$. Pour ce faire, il nous faut caractériser la transformée de Fourier de A_h . Celui ci est également un cas particulier d'opérateur à symbole :

LEMME 2.2.5

On a l'identité

$$\forall u \in L^2, \quad \widehat{A_h u}(k) = \widehat{A}_h(k) \widehat{u}(k), \quad \forall k \in \mathbb{R}, \quad (2.27)$$

et l'opérateur A_h a pour symbole $\widehat{A}_h(k)$ donné par

$$\widehat{A}_h(k) = i c \frac{\sin kh}{h}.$$

$$\begin{array}{ccc} u & \xrightarrow{A_h} & A_h u \\ \mathcal{F} \downarrow & & \downarrow \mathcal{F} \\ \widehat{u} & \xrightarrow{\widehat{A}_h} & \widehat{A}_h \widehat{u} \end{array}$$

DÉMONSTRATION: La propriété (1.10) signifie que l'opérateur τ_a (défini par $\tau_a u(x) = u(x+a)$) n'est autre que l'opérateur de symbole e^{iak} . Comme

$$A_h = c \frac{\tau_h - \tau_{-h}}{2h},$$

on en déduit que A_h est l'opérateur de symbole

$$\widehat{A}_h(k) = c \frac{e^{ikh} - e^{-ikh}}{2h} = i c \frac{\sin kh}{h}$$

ce qui achève la démonstration. ■

REMARQUE 2.2.5

On peut remarquer que

$$\widehat{A}_h(k) = i c |k| \left(1 + O(|k|^2 h^2) \right).$$

On retrouve alors sous une autre forme, le fait que l'opérateur A_h dont le symbole est $\widehat{A}_h(k)$ approche l'opérateur $c \frac{\partial}{\partial x}$ dont le symbole est $i c k$, et ce au second ordre en h .

De (2.22), nous déduisons facilement le

COROLLAIRE 2.2.6

L'opérateur $S_h(\Delta t)$ est l'opérateur de symbole

$$\widehat{S}_h(k, \Delta t) = 1 - i \alpha \sin kh, \quad \text{avec} \quad \alpha = \frac{c \Delta t}{h}.$$

Par conséquent, d'après le théorème 2.2.4

$$\|S_h(\Delta t)\|_{\mathcal{L}(L^2)} = \sup_{k \in \mathbb{R}} |\widehat{S}_h(k, \Delta t)| = (1 + \alpha^2)^{1/2}.$$

Finalement, nous avons tous les éléments pour établir le théorème suivant

THÉORÈME 2.2.7

Le schéma explicite centré (2.4) est inconditionnellement instable dans L^2 .

Au final, le schéma explicite centré qui semblait à première vue naturel est donc, malgré sa consistance, instable. La solution numérique calculée grâce à ce schéma ne peut *a priori* pas converger vers la solution du problème continu (2.3).

La figure 2.5 illustre l'instabilité du schéma explicite centré. Pour plusieurs pas de

temps, nous représentons en rouge la solution exacte (de l'équation d'advection (2.3) pour une vitesse c particulière) et en bleu la solution numérique calculée en utilisant le schéma explicite centré (2.4).

REMARQUE 2.2.6 (REMARQUE SUR LA MÉTHODE DE FOURIER-VON NEUMANN)

Les calculs précédents apparaissent tout naturellement quand on cherche la solution explicite de (2.6) en appliquant la transformation de Fourier. Nous procédons donc ici de la même façon que pour l'équation continue (2.3) à la section 1.2.2.

Appliquons la transformation de Fourier à (2.6). Si \widehat{u}_h^n désigne la transformée de Fourier de u_h^n , en utilisant (2.27), nous obtenons :

$$\frac{\widehat{u}_h^{n+1}(k) - \widehat{u}_h^n(k)}{\Delta t} + i c \frac{\sin kh}{h} \widehat{u}_h^n(k) = 0,$$

c'est à dire

$$\widehat{u}_h^{n+1}(k) = \widehat{S}_h(k, \Delta t) \widehat{u}_h^n(k) \quad (2.28)$$

Cette formule est à comparer à la formule (1.11) pour la solution exacte : le symbole $\widehat{S}_h(k, \Delta t)$ apparaît donc comme le coefficient d'amplification du schéma.

Par suite

$$\widehat{u}_h^n(k) = \widehat{S}_h(k, \Delta t)^n \widehat{u}_h^0(k).$$

Par le théorème de Plancherel

$$\|u_h^n\|_{L^2}^2 = 2\pi \int_{-\infty}^{+\infty} \left| \widehat{S}_h(k, \Delta t) \right|^n |\widehat{u}_h^0(k)|^2 dk$$

et la recherche de l'estimation (2.17) (pour $p = 2$) mène ainsi naturellement au résultat du théorème 2.2.7.

Retenons ici une recette de calcul pour la mise en oeuvre de la méthode de Fourier-Von Neumann lors de l'analyse de stabilité L^2 d'un schéma numérique approchant une équation aux dérivées partielles d'évolution en une dimension d'espace :

(i) On cherche une solution de la forme

$$u_j^n = u_h^n(k) e^{ikx_j}, \quad j \in \mathbb{Z}, \quad n \geq 0, \quad (u_h^n(k) \in \mathbb{C}).$$

(ii) On détermine une relation de récurrence satisfaite par la suite $u_h^n(k)$,

(iii) On étudie le comportement de la famille de suites $(u_h^n(k), k \in \mathbb{R})$.

Remarquons alors que :

- cette méthode n'est applicable que si l'on s'intéresse à l'approximation d'une équation linéaire à coefficients constants posée sur \mathbb{R} , sur un maillage régulier,
- dans le cas des schémas numériques à un pas de temps, la relation de récurrence est nécessairement de la forme (2.28) et l'analyse de stabilité se réduit à l'étude du coefficient d'amplification numérique.

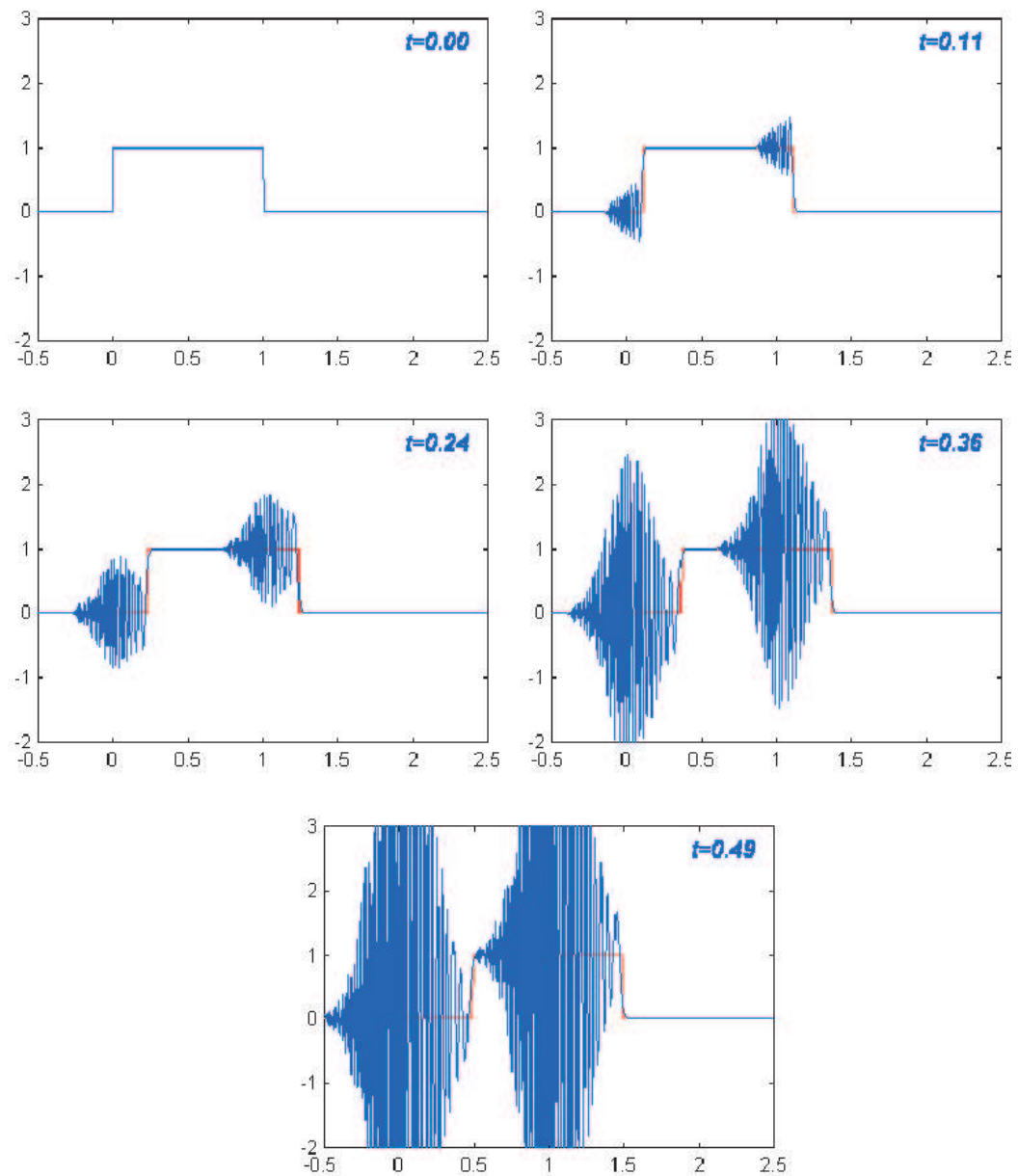


FIGURE 2.5 – Schéma explicite centré (inconditionnellement instable).

2.3 Le schéma de Lax-Friedrichs pour l'équation d'advection

À la différence du schéma explicite centré, le schéma de Lax-Friedrichs cherche à symétriser le traitement sur la variable temporelle tout en conservant le caractère explicite du schéma, le caractère centré. On utilise donc :

- l'approximation de $\frac{\partial u}{\partial t}$ suivante

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

- une approximation centrée, d'ordre 2, pour $\frac{\partial u}{\partial x}$, écrite à l'instant t^n (voir Section 2.1)

$$\frac{\partial u}{\partial x}(x_j, t^n) \simeq D_h^0 u(x_j, t^n) = \frac{u_{j+1}^n - u_{j-1}^n}{2h}.$$

L'équation du schéma de Lax-Friedrichs au point (x_j, t^n) s'écrit donc

$$\frac{2u_j^{n+1} - u_{j+1}^n - u_{j-1}^n}{2\Delta t} + c \frac{u_{j+1}^n - u_{j-1}^n}{2h} = 0. \quad (2.29)$$

Il s'agit d'un schéma à un pas (ou deux points) en temps et à deux points en espace. Ce schéma est bien explicite. Il permet de calculer explicitement la solution à l'instant $n+1$ à partir de la solution à l'instant n

$$u_j^{n+1} = \frac{1-\alpha}{2} u_{j+1}^n - \frac{1+\alpha}{2} u_{j-1}^n, \quad j \in \mathbb{Z},$$

où nous rappelons que

$$\alpha = \frac{c\Delta t}{h}.$$

Evidemment, pour que (2.29) définisse entièrement la solution approchée, il suffit de connaître les u_j^0 , $j \in \mathbb{Z}$, ce pour quoi on va, comme indiqué dans (2.5), utiliser la condition initiale.

REMARQUE 2.3.1 (INTERPRÉTATION GÉOMÉTRIQUE DU SCHÉMA DE LAX FRIEDRICHS)

On rappelle que la solution exacte est constante sur les droites caractéristiques qui sont ici plus précisément les droites de pente c . On a donc

$$u(x_j, t^{n+1}) = u(x^*, t^n), \quad \text{avec } x^* = x_j - c\Delta t.$$

On remarque ainsi aisément que (voir Figure 2.6) :

$$x^* = \frac{1+\alpha}{2} x_{j-1} + \frac{1-\alpha}{2} x_{j+1}.$$

Le schéma de Lax-Friedrichs consiste donc à considérer une interpolation linéaire sur le maillage de la solution continue.

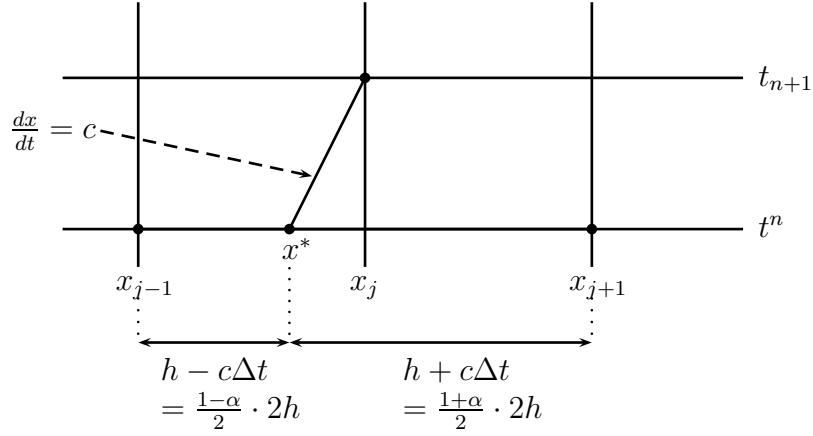


FIGURE 2.6 – Interprétation géométrique du schéma de Lax Friedrichs

2.3.1 Consistance et ordre du schéma

Nous commençons par l'étude de l'erreur de troncature qui nous conduit à l'analyse de la consistance. Contrairement au schéma explicite centré, l'analyse est ici un peu plus complexe.

Respectant la définition 2.2.1, l'erreur de troncature du schéma (2.29) est donnée par

$$\varepsilon_j^n = \frac{U_j^{n+1} - \frac{U_{j+1}^n + U_{j-1}^n}{2}}{\Delta t} + c \frac{U_{j+1}^n - U_{j-1}^n}{2h}, \quad (2.30)$$

où U_j^n est définie par $U_j^n = u(x_j, t^n)$, u étant une solution régulière de (2.3).

On peut alors démontrer

THÉORÈME 2.3.1

Le schéma (2.29) est d'ordre 1 en temps et en espace.

DÉMONSTRATION: On commence par remarquer que

$$\frac{U_{j+1}^n + U_{j-1}^n}{2} = U_j^n + \frac{U_{j+1}^n - 2U_j^n + U_{j-1}^n}{2},$$

en utilisant la définition 2.2.1 de l'erreur de troncature pour le schéma explicite centré que nous notons $(\varepsilon_j^n)_{EC}$, il vient facilement que l'erreur de troncature pour le schéma de Lax-Friedrichs ε_j^n est donnée par :

$$\varepsilon_j^n = (\varepsilon_j^n)_{EC} + \frac{h^2}{2\Delta t} \frac{U_{j+1}^n - 2U_j^n + U_{j-1}^n}{h^2}.$$

Si on fixe $\alpha = \frac{c\Delta t}{h}$ et en utilisant l'opérateur d'approximation de la dérivée seconde donnée dans (2.2), on obtient

$$\varepsilon_j^n = (\varepsilon_j^n)_{EC} + \frac{ch}{2\alpha} \left[\left(\frac{\partial^2 u}{\partial x^2} \right)_j^n + \mathcal{O}(h^2) \right]. \quad (2.31)$$

En utilisant le calcul de l'erreur de consistance pour le schéma explicite centré (voir la preuve du Théorème 2.2.1)

$$(\varepsilon_j^n)_{EC} = \mathcal{O}(\Delta t + h^2),$$

on obtient finalement

$$\varepsilon_j^n = \mathcal{O}(\Delta t + h).$$

■

Pour l'analyse de convergence, nous utilisons la version « continue en espace » de l'erreur de troncature, à savoir, en reprenant les notations de la section 2.2,

$$\varepsilon_h^n(x) = \frac{u(x, t^{n+1}) - \frac{u(x+h, t^n) + u(x-h, t^n)}{2}}{\Delta t} + c \frac{u(x+h, t^n) - u(x-h, t^n)}{2h}.$$

À partir des expressions (2.13) et (2.31), on déduit une estimation L^∞ de ε_h^n (avec une constante numérique C dépendant de α)

$$\|\varepsilon_h^n\|_{L^\infty} \leq C(\alpha) \left(\Delta t \sup_{t \in [t^n, t^{n+1}]} \left\| \frac{\partial^2 u}{\partial t^2}(\cdot, t) \right\|_{L^\infty} + h \sup_{t \in [t^n, t^{n+1}]} \left\| \frac{\partial^2 u}{\partial x^2}(\cdot, t) \right\|_{L^\infty} \right).$$

Nous pouvons également obtenir une estimation équivalente de ε_h^n pour la norme L^p , c'est à dire qu'il existe une constante $C(\alpha)$ telle que

$$\|\varepsilon_h^n\|_{L^p} \leq C(\alpha) \left(\Delta t \sup_{t \in [t^n, t^{n+1}]} \left\| \frac{\partial^2 u}{\partial t^2}(\cdot, t) \right\|_{L^p} + h \sup_{t \in [t^n, t^{n+1}]} \left\| \frac{\partial^2 u}{\partial x^2}(\cdot, t) \right\|_{L^p} \right),$$

ce qui pour le problème de Cauchy se ramène au

THÉORÈME 2.3.2

Pour le schéma de Lax-Friedrichs (2.29), l'erreur de troncature est telle que pour tout p , il existe une constante $C(\alpha)$ dépendant de α telle que

$$\|\varepsilon_h^n\|_{L^p} \leq C(\alpha) h \left\| \frac{\partial^2 u}{\partial x^2}(\cdot, t) \right\|_{L^p}. \quad (2.32)$$

2.3.2 Vitesse de propagation numérique

La formule (2.29) démontre que le support de la solution numérique se propage à vitesse finie au sens où

$$\text{supp } u_h^0 \subset [a, b] \implies \text{supp } u_h^n \subset [a - nh, b + nh]$$

La vitesse de propagation de la solution numérique est donc donnée par

$$V_{num}(\Delta, h) = \frac{h}{\Delta t}.$$

Comme pour le schéma explicite centré, une condition nécessaire de convergence du schéma est donc que la vitesse numérique doit être supérieure à la vitesse de propagation de la solution exacte (voir Figure 2.3) :

$$\frac{h}{\Delta t} \geq c \quad \Leftrightarrow \quad \alpha = \frac{c\Delta t}{h} \leq 1. \quad (2.33)$$

2.3.3 Analyse de stabilité

Le schéma de Lax Friedrichs (2.29) produit une suite u_h^n satisfaisant la relation de récurrence

$$u_h^{n+1} = S_h(\Delta t) u_h^n \quad (2.34)$$

avec

$$[S_h(\Delta t) u](x) = \frac{1-\alpha}{2} u(x+h) + \frac{1+\alpha}{2} u(x-h).$$

En appliquant la méthode de Fourier-Von Neumann développée à la section 2.2.3 ou l'algorithme proposé en Remarque 2.2.6 sur la stabilité L^2 , on montre que la relation précédente mène au coefficient d'amplification

$$\widehat{S}_h(k, \Delta t) = \frac{1-\alpha}{2} e^{ikh} + \frac{1+\alpha}{2} e^{-ikh},$$

soit (voir Figure 2.7 pour la représentation dans le plan complexe de $k \mapsto \widehat{S}_h(k, \Delta t)$)

$$\widehat{S}_h(k, \Delta t) = \cos(kh) - i\alpha \sin(kh).$$

On rappelle qu'un schéma donné par une relation de récurrence d'ordre 1 est L^2 -stable si et seulement si il existe une constante $\nu \geq 0$ telle que

$$\forall (\Delta t, h) \in \mathbf{S}, \quad \|S_h(\Delta t)\|_{\mathcal{L}(L^2)} = \sup_{k \in \mathbb{R}} |\widehat{S}_h(k, \Delta t)| \leq 1 + \nu \Delta t.$$

On en déduit alors le résultat de stabilité L^2 pour le schéma de Lax-Friedrichs.

THÉORÈME 2.3.3

Le schéma de Lax-Friedrichs (2.29) est stable (et même uniformément stable) pour la norme L^2 si et seulement si

$$\alpha = \frac{c\Delta t}{h} \leq 1. \quad (2.35)$$

Cette condition est appelée condition CFL (Courant-Friedrichs-Lewy).

DÉMONSTRATION: Il suffit de remarquer que

$$\sup_{k \in \mathbb{R}} |\widehat{S}_h(k, \Delta t)| = \max(1, \alpha).$$

■

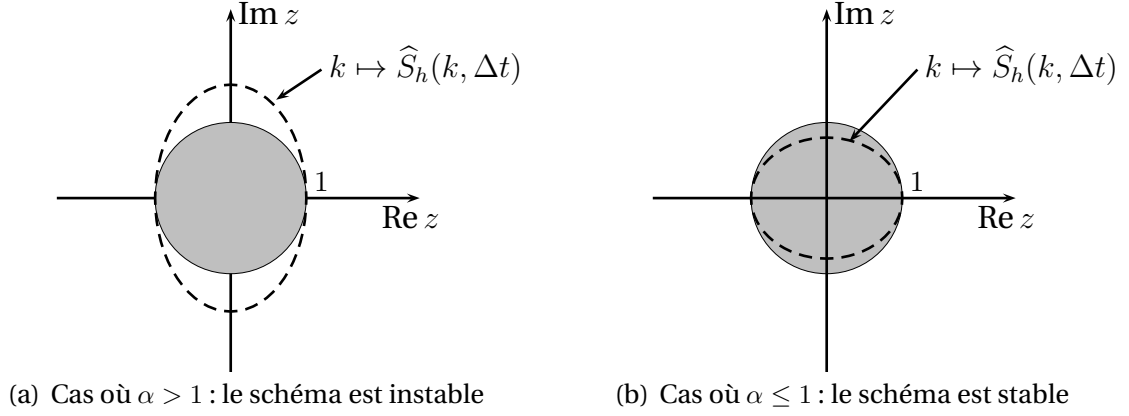


FIGURE 2.7 – Coefficient d'amplification dans le plan complexe en fonction de $\alpha = \frac{c\Delta t}{h}$

La stabilité L^2 est donc obtenue par le calcul du coefficient d'amplification en utilisant la méthode de Fourier-Von Neumann. Plus généralement, pour la stabilité L^p , il faut utiliser des arguments différents. Pour le schéma de Lax-Friedrichs, on utilise un argument de convexité.

THÉORÈME 2.3.4

Le schéma de Lax-Friedrichs (2.29) est stable (et même uniformément stable) pour la norme L^p si et seulement si

$$\alpha = \frac{c\Delta t}{h} \leq 1$$

DÉMONSTRATION: Supposons que $\alpha \leq 1$. On rappelle que pour tout u

$$[S_h(\Delta t)u](x) = \frac{1-\alpha}{2} u(x+h) + \frac{1+\alpha}{2} u(x-h),$$

avec

$$\frac{1-\alpha}{2} \geq 0, \quad \frac{1+\alpha}{2} \geq 0, \quad \frac{1-\alpha}{2} + \frac{1+\alpha}{2} = 1.$$

Par convexité de $x \mapsto |x|^p$ pour $p \geq 1$, on obtient pour toute fonction $u \in L^p$, d'une part que $S_h(\Delta t)u \in L^p$ et d'autre part que

$$\int_R |[S_h(\Delta t)u](x)|^p dx \leq \frac{1-\alpha}{2} \int_R |u(x+h)|^p dx + \frac{1+\alpha}{2} \int_R |u(x-h)|^p dx$$

ce qui entraîne

$$\|S_h(\Delta t)u\|_{L^p} \leq \|u\|_{L^p}.$$

En utilisant, le théorème 2.2.2, on en déduit que le schéma est stable pour la norme L^p . ■

La condition CFL (2.35) correspond exactement à la condition nécessaire de convergence (2.33) obtenue en comparant la vitesse numérique et la vitesse de propagation de la solution continue. Elle dit que le pas de temps doit être inférieur à h/c . Cette condition est quelque peu contraignante : même si on se place à la limite de la condition de stabilité, lorsque qu'on divise le pas d'espace par 2, le pas de temps doit être divisé par 2. Si on raffine le maillage en espace par 10, il faut raffiner le maillage en temps par 10 !

L'intérêt du schéma explicite est que les calculs à effectuer à chaque pas de temps sont très simples et peu coûteux. La contre-partie est que la condition de stabilité contraint à choisir un pas de temps aussi petit que le pas en espace ce qui peut engendrer un gros surcoût de calcul.

2.3.4 Convergence du schéma

Passons maintenant à la convergence du schéma (2.29) dans $L^\infty(0, T; L^p)$ en montrant d'après la définition 2.2.3 que pour toute donnée initiale dans L^p

$$\lim_{(h, \delta t) \rightarrow 0} \sup_{t^n \leq T} \|u(\cdot, t^n) - u_h^n\|_{L^p} = 0,$$

où u est la solution continue de (2.3) et u_h^n est la suite produite par le schéma (2.29).

Nous introduisons donc l'erreur commise sur la solution

$$e_h^n = u(\cdot, t^n) - u_h^n,$$

et obtenons la convergence du schéma comme une conséquence de sa consistance et de sa stabilité.

La première étape consiste donc à relier l'erreur e_h^n à l'erreur de troncature ε_h^n . C'est particulièrement simple ici grâce à la linéarité des équations. En effet, des égalités (2.34) et (2.30), nous déduisons que

$$\begin{cases} u(\cdot, t^{n+1}) = S_h(\Delta t) u(\cdot, t^n) + \Delta t \varepsilon_h^n, \\ u_h^{n+1} = S_h(\Delta t) u_h^n, \end{cases}$$

soit par différence

$$e_h^{n+1} = S_h(\Delta t) e_h^n + \Delta t \varepsilon_h^n. \quad (2.36)$$

Ainsi, l'erreur e_h^n apparaît comme la solution d'un problème d'évolution discret entretenu par un terme source qui n'est autre que l'erreur de troncature.

Plus généralement, nous allons en fait montrer que si une suite e_h^n vérifie la relation (2.36) où l'opérateur $S_h(\Delta t)$ satisfait l'estimation de stabilité L^p (2.19), alors on peut

estimer e_h^n en fonction de e_h^0 et des ε_h^k uniformément en Δt , tant que $t^n \leq T$.

En effet, nous observons que

$$\begin{cases} e_h^1 = S_h(\Delta t) e_h^0 + \Delta t \varepsilon_h^0, \\ e_h^2 = S_h(\Delta t) e_h^1 + \Delta t \varepsilon_h^1 = S_h(\Delta t)^2 e_h^0 + \Delta t \left(\varepsilon_h^1 + S_h(\Delta t) \varepsilon_h^0 \right), \end{cases}$$

et, de proche en proche,

$$e_h^n = S_h(\Delta t)^n e_h^0 + \Delta t \left(\sum_{k=0}^{n-1} S_h(\Delta t)^k \varepsilon_h^{n-1-k} \right).$$

En passant aux normes, on obtient alors

$$\|e_h^n\|_{L^p} = \|S_h(\Delta t)^n\|_{\mathcal{L}(L^p)} \|e_h^0\|_{L^p} + \Delta t \left(\sum_{k=0}^{n-1} \|S_h(\Delta t)^k\|_{\mathcal{L}(L^p)} \|\varepsilon_h^{n-1-k}\|_{L^p} \right).$$

Si on suppose l'hypothèse de stabilité (2.21) satisfaite, on a

$$\|S_h(\Delta t)\|_{\mathcal{L}(L^p)} \leq 1 + \nu \Delta t.$$

Or on sait que

$$\forall t^q \leq T, \quad \|S_h(\Delta t)^q\|_{\mathcal{L}(L^p)} \leq e^{\nu t^q} \leq e^{\nu T}.$$

Nous avons donc,

$$\forall t^n \leq T, \quad \|e_h^n\|_{L^p} \leq e^{\nu T} \left(\|e_h^0\|_{L^p} + t^n \sup_{k \leq n-1} \|\varepsilon_h^k\|_{L^p} \right),$$

ce qui mène à l'estimation d'erreur

$$\sup_{t^n \leq T} \|e_h^n\|_{L^p} \leq e^{\nu T} \left(\|e_h^0\|_{L^p} + T \sup_{t^n \leq T} \|\varepsilon_h^n\|_{L^p} \right).$$

Bien entendu, dans le cas d'un schéma uniformément stable ($\nu = 0$), l'estimation devient

$$\sup_{t^n \leq T} \|e_h^n\|_{L^p} \leq \left(\|e_h^0\|_{L^p} + T \sup_{t^n \leq T} \|\varepsilon_h^n\|_{L^p} \right). \quad (2.37)$$

Nous pouvons appliquer ce résultat général pour établir des estimations d'erreur L^p pour le schéma de Lax Friedrichs (2.29).

THÉORÈME 2.3.5

Soient u la solution de (2.3) et u_h^n la solution de (2.29, 2.7). Si $\frac{c\Delta t}{h} \leq 1$ et si, pour $p \in [1, +\infty]$, on a l'estimation d'erreur $L^\infty(0, T; L^p)$:

$$\sup_{t^n \leq T} \|u(\cdot, t^n) - u_h^n\|_{L^p} \leq h C(\alpha) T \left\| \frac{d^2 u^0}{dx^2} \right\|_{L^p}.$$

DÉMONSTRATION: D'après le théorème 2.3.4, sous la condition CFL (2.35), nous savons que le schéma est uniformément stable dans L^p . Nous pouvons donc appliquer la formule (2.37) dans laquelle nous injectons l'estimation (2.32) de l'erreur de troncature.

Le terme de (2.37) concernant e_0^h est nul si la condition initiale imposée pour la solution numérique est (choix (i) de (2.7))

$$\forall j, \quad u_j^0 = u(x_j, t = 0) = u^0(x_j).$$

Si la condition initiale imposée est autre (par exemple le choix (ii) de (2.7)), pour estimer le terme en e_0^h nous utilisons le résultat classique suivant (sur l'interpolation linéaire - exercice)

$$\|e_h^0\|_{L^p} = \|\Pi_h u^0 - u^0\|_{L^p} \leq Ch \left\| \frac{d^2 u^0}{dx^2} \right\|_{L^p}.$$

■

D'après la Section 2.3.2 et le théorème, la condition CFL (2.35) est donc une condition nécessaire et suffisante de convergence du schéma de Lax-Friedrichs. Les figures 2.8 et 2.9 illustrent l'importance de cette condition dans la stabilité et donc la convergence du schéma : pour plusieurs pas de temps, nous représentons en rouge la solution exacte (de l'équation d'advection (2.3) pour une vitesse c particulière) et en bleu la solution numérique calculée en utilisant le schéma de Lax Friedrichs (2.29).

De manière générale, nous venons d'établir dans cette section, le théorème de Lax.

THÉORÈME 2.3.6 (LAX)

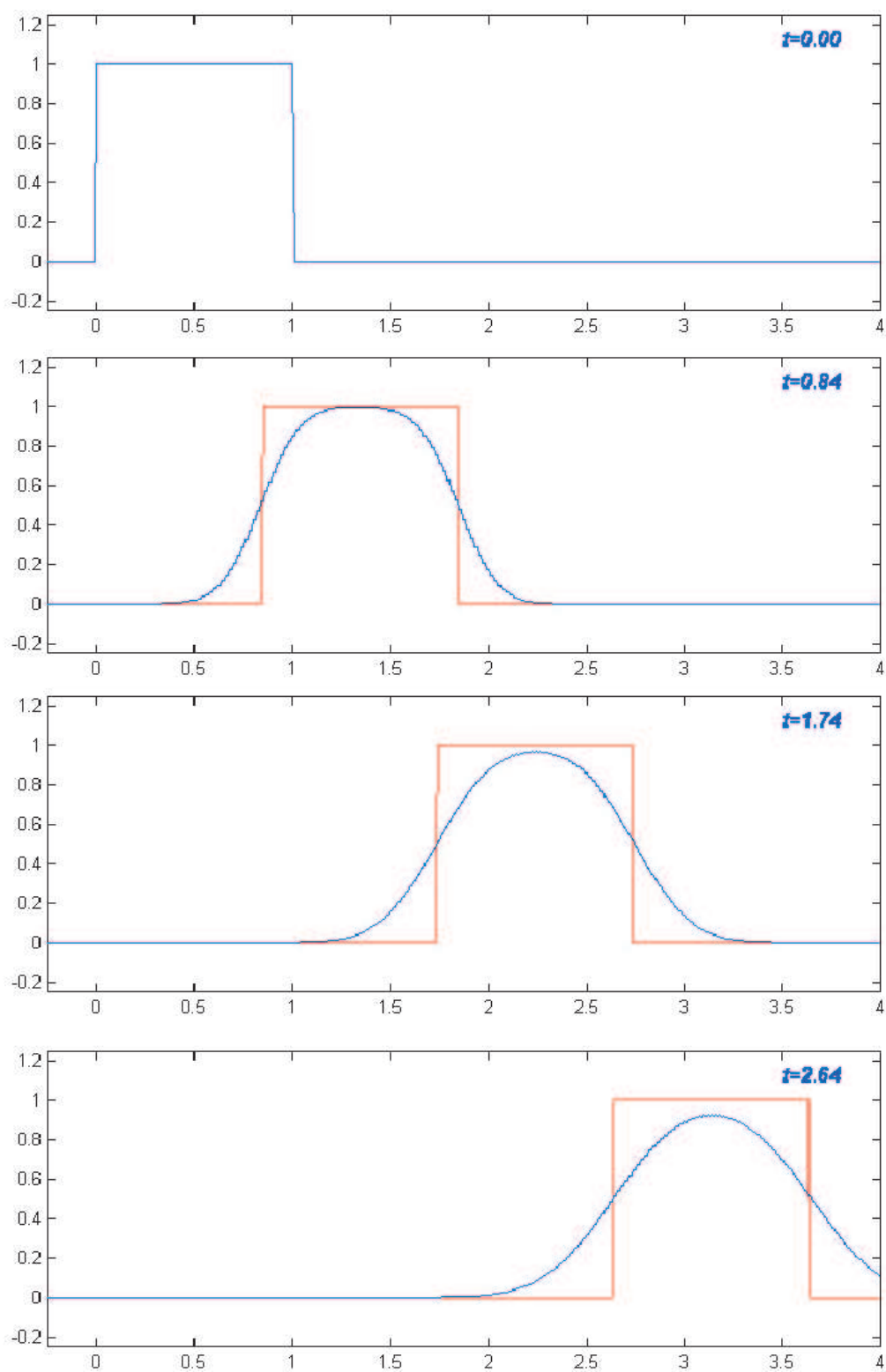
Soit u une solution suffisamment régulière d'un système différentiel et u_j^n la solution numérique discrète obtenue par un schéma aux différences finies avec la donnée initiale $u_j^0 = u^0(x_j)$. On suppose le schéma linéaire, consistant et, pour un certain domaine de stabilité \mathbf{S} admettant $(0, 0)$ comme point d'accumulation, stable pour une norme $\|\cdot\|$, alors le schéma est convergent au sens où

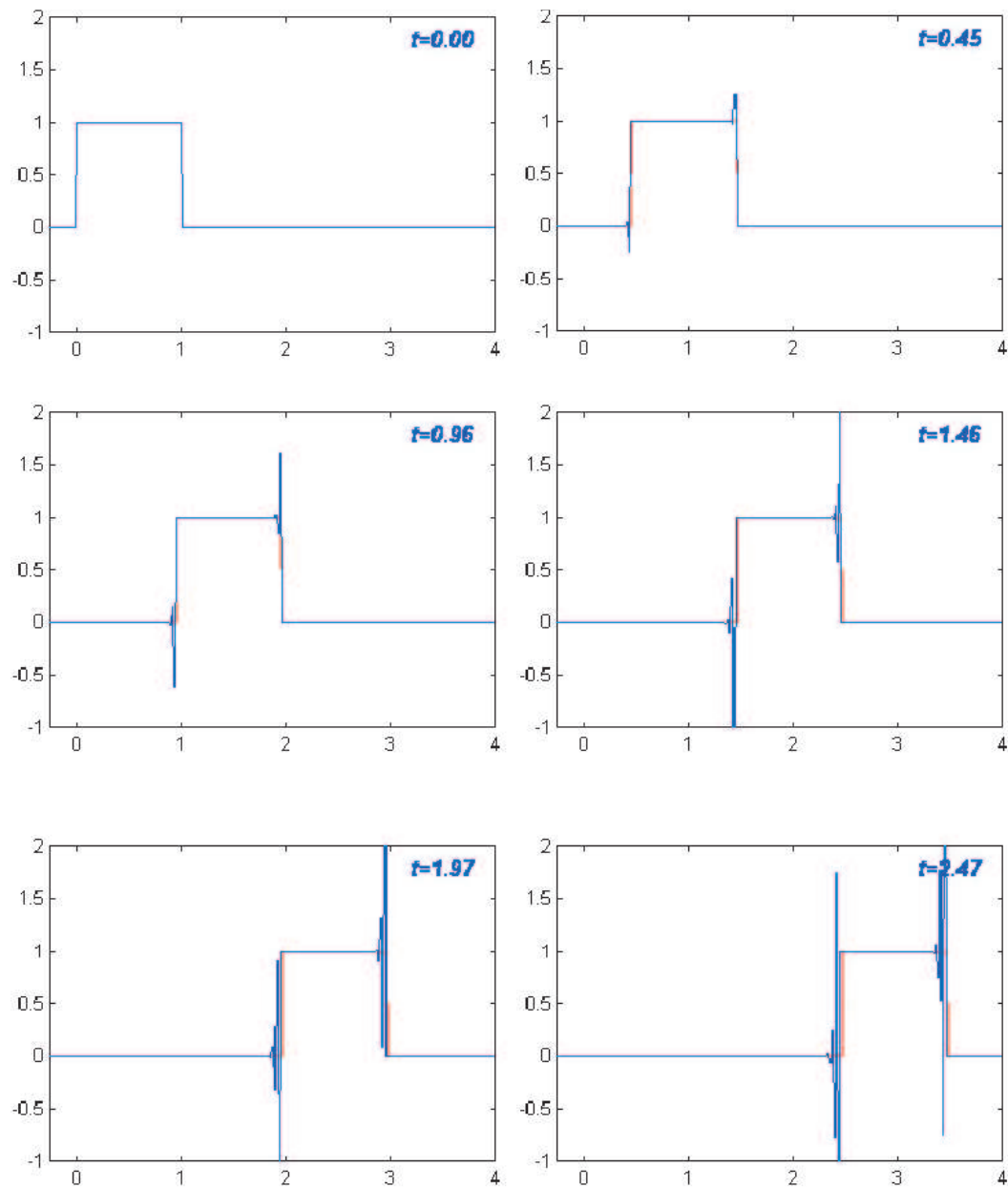
$$\forall (h, \Delta t) \in \mathbf{S}, \quad \forall T > 0, \quad \lim_{\Delta t, h \rightarrow 0} \sup_{t^n \leq T} \|e_h^n\| = 0,$$

avec $e_h^n = u_h^n - u(t^n, \cdot)$ l'erreur commise sur la solution.

De plus, l'ordre de convergence est donné par l'erreur de troncature : c'est à dire que si le schéma est d'ordre p en temps et d'ordre q en espace alors

$$\forall n, \quad e_h^n = \mathcal{O}(\Delta t^p + h^q).$$

FIGURE 2.8 – Schéma de Lax-Friedrichs avec $c\Delta t = 0.3h$

FIGURE 2.9 – Schéma de Lax-Friedrichs avec $c\Delta t = 1.01h$

2.4 Le schéma saute-mouton pour l'équation des ondes

Nous considérons ici l'équation des ondes sur \mathbb{R} à vitesse constante c

$$\begin{cases} \frac{\partial^2 u}{\partial t^2} - c^2 \frac{\partial^2 u}{\partial x^2} = 0, & \text{pour } t \geq 0, x \in \mathbb{R} \\ u(t = 0, x) = u^0(x), & \text{pour } x \in \mathbb{R} \\ \frac{\partial u}{\partial t}(t = 0, x) = u^1(x), & \text{pour } x \in \mathbb{R} \end{cases} \quad (2.38)$$

Nous discrétisons toujours l'espace avec un pas h et le temps avec $\Delta t > 0$ et cherchons toujours à calculer l'inconnue discrète $u^n = (u_j^n)_{j \in \mathbb{Z}}$ à chaque pas de temps.

2.4.1 Présentation du schéma

Le schéma saute-mouton est un schéma explicite centré :

$$\forall n, \forall j, \quad \frac{u_j^{n+1} - 2u_j^n + u_j^{n-1}}{(\Delta t)^2} - c^2 \frac{u_{j+1}^n - 2u_j^n + u_{j-1}^n}{h^2} = 0. \quad (2.39)$$

La première remarque fondamentale sur le plan pratique, est que, son caractère explicite permet une mise en oeuvre peut couteuse en résolvant simplement :

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

où le nombre sans dimension :

$$\alpha = \frac{c\Delta t}{h},$$

mesure le rapport entre la distance parcourue par l'onde pendant le temps Δt (soit $c\Delta t$) et la taille de la maille en espace h . Nous verrons un peu plus loin que, comme pour l'équation d'advection, cette quantité est liée à la condition CFL des schémas introduits.

Bien entendu pour pouvoir déterminer complètement la solution discrète, il faut la connaître aux deux premiers instants. C'est pour la détermination de la solution aux deux premiers instants que vont intervenir les conditions initiales. Les difficultés qui en découlent sont caractéristiques de l'équation des ondes qui est d'ordre 2 en espace et en temps.

2.4.2 Le schéma de démarrage : choix des conditions initiales approchées

Pour le problème continu, les données initiales sont définies par les valeurs de la solution à $t = 0$, ainsi que celles de sa dérivée en temps au même instant.

Pour le problème discrétisé, pour déterminer u_h^0 et u_h^1 on doit approcher $u(x_j, 0)$ et $u(x_j, \Delta t)$.

Pour u_h^0 , il est naturel de choisir

$$\forall j, \quad u_j^0 = u^0(x_j). \quad (2.40)$$

Pour approcher $u(x_j, \Delta t)$, on utilise simplement un développement de Taylor au premier ordre

$$u(x_j, \Delta t) = u(x_j, 0) + \Delta t \frac{\partial u}{\partial t}(x_j, 0) + \mathcal{O}(\Delta t^2),$$

ce qui mène à l'approximation

$$\forall j, \quad u_j^1 = u^0(x_j) + \Delta t u^1(x_j).$$

Nous verrons que la précision d'un tel schéma de démarrage est insuffisante. Pour en avoir l'intuition, il suffit de remarquer que ce choix correspond à une approximation d'ordre 1 en temps de la condition initiale

$$\frac{\partial u}{\partial t}(x, 0) = u^1(x).$$

Ainsi cette dernière équation se réécrit

$$\forall j, \quad \frac{u_j^1 - u_j^0}{\Delta t} = u^1(x_j),$$

où le terme de gauche peut être vue comme une approximation décentrée (donc d'ordre 1) de la dérivée en temps de la solution à l'instant $t = 0$.

Pour avoir un schéma de démarrage d'ordre 2 en temps (plus cohérent avec l'ordre du schéma (2.39) qui, comme on le verra plus loin, est égal à 2), il suffit de pousser le développement de Taylor un cran plus loin

$$u(x_j, \Delta t) = u^0(x_j) + \Delta t u^1(x_j) + \frac{\Delta t^2}{2} \frac{\partial^2 u}{\partial t^2}(x_j, 0) + \mathcal{O}(\Delta t^3),$$

et, comme on ne dispose pas de la dérivée seconde de la solution en tant que donnée initiale, d'utiliser l'équation des ondes écrite à l'instant $t = 0$ pour en déduire que

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

ce qui mène au schéma de démarrage

$$\forall j, \quad u_j^1 = u_j^0 + \Delta t u^1(x_j) + \frac{c^2 \Delta t^2}{2} \frac{u_{j+1}^0 - 2u_j^0 + u_{j-1}^0}{h^2}. \quad (2.41)$$

Le schéma numérique est donc complètement défini par (2.39), (2.40) et (2.41).

2.4.3 Ordre du schéma : consistance et erreur de troncature

De la même façon que dans la définition 2.2.1, l'erreur de troncature du schéma (2.39) est donnée par

$$\varepsilon_j^n = \frac{U_j^{n+1} - 2U_j^n + U_j^{n-1}}{\Delta t^2} - c^2 \frac{U_{j+1}^n - 2U_j^n + U_{j-1}^n}{h^2}.$$

où U_j^n est définie par $U_j^n = u(x_j, t^n)$, u étant une solution régulière de (2.38).

On peut alors montrer que

THÉORÈME 2.4.1

Le schéma (2.39) est consistant à l'ordre 2 en temps et en espace :

$$\varepsilon_j^n = \mathcal{O}(\Delta t^2 + h^2).$$

DÉMONSTRATION: Il suffit d'utiliser l'opérateur d'approximation de la dérivée seconde donnée dans (2.2). ■

2.4.4 Analyse de stabilité L^2 par techniques énergétiques

Comme nous l'avons vu pour d'autres types d'équation, la stabilité d'un schéma (dans une certaine norme) est un préalable incontournable à sa convergence (dans cette même norme). Nous allons mener cette analyse dans le cadre le plus adapté à l'équation des ondes qui est le cadre L^2 du fait de la périodicité des solutions de l'équation des ondes. Ayant déjà étudié la méthode de Fourier-Von Neumann pour l'équation d'advection, nous allons ici utiliser une autre méthode fondée sur une propriété fondamentale de l'équation des ondes : la conservation de l'énergie.

Un résultat de conservation d'énergie discrète

Il est assez naturel de se demander si le schéma utilisé satisfait une relation de conservation équivalente à la conservation d'énergie pour la solution du problème continu. En effet, nous avons vu au théorème 1.3.3 que l'équation des ondes vérifie pour tout $t > 0$,

$$E(t) = E(0) \quad \text{avec} \quad E(t) = \frac{1}{2} \int_{\mathbb{R}} \left| \frac{\partial u}{\partial t}(t, x) \right|^2 dx + \frac{c^2}{2} \int_{\mathbb{R}} \left| \frac{\partial u}{\partial x}(t, x) \right|^2 dx.$$

Introduisons alors l'énergie discrète (à l'instant $t^{n+\frac{1}{2}} = n + \frac{1}{2}\Delta t$) :

$$E^{n+\frac{1}{2}} = \frac{1}{2} \sum_{j \in \mathbb{Z}} \left(\frac{u_j^{n+1} - u_j^n}{\Delta t} \right)^2 + \frac{1}{2} a_h(u_h^{n+1}, u_h^n), \quad n \geq 1, \quad (2.42)$$

avec

$$a_h(u, v) = c^2 \sum_{j \in \mathbb{Z}} \left(\frac{u_{j+1} - u_j}{h} \right) \left(\frac{v_{j+1} - v_j}{h} \right).$$

et où u_h^n , définie sur $(0, 1)$ est la version « continue en espace » de la solution discrète, en reprenant les notations de la section 2.2.

En remarquant que $a_h(u, v)$ est un équivalent discret de

$$c^2 \int_{\mathbb{R}} \frac{\partial u}{\partial x} \frac{\partial v}{\partial x} dx,$$

on voit que la quantité (2.42) peut être vu comme une approximation de l'énergie $E(t^{n+\frac{1}{2}})$ où l'on a remplacé le produit scalaire $L^2(\mathbb{R})$ sur la variable d'espace par le produit scalaire ℓ^2 sur le vecteur $(u_j^n)_j$:

$$\|u_h^n\|_{\ell^2}^2 = \langle u_h^n, u_h^n \rangle_{\ell^2} = \sum_{j \in \mathbb{Z}} (u_j^n)^2.$$

Autrement dit

$$E_h^{n+\frac{1}{2}} = \frac{1}{2} \left\| \frac{u_h^{n+1} - u_h^n}{\Delta t} \right\|_{\ell^2}^2 + \frac{1}{2} a_h(u_h^{n+1}, u_h^n).$$

THÉORÈME 2.4.2

Si la suite $(u_j^n)_j$ est solution du schéma saute-mouton (2.39), l'énergie discrète $E_h^{n+\frac{1}{2}}$ se conserve au cours du temps :

$$E_h^{n+\frac{1}{2}} = E_h^{n-\frac{1}{2}}, \quad \forall n \geq 1.$$

PREUVE: Nous laissons les détails de la démonstration au lecteur pour nous concentrer sur les principes de la démonstration.

Rappelons que la conservation de l'énergie continue $E(t)$ s'obtient en multipliant l'équation des ondes par $\frac{\partial u}{\partial t}$ et en intégrant le résultat en espace. Nous procédons de manière équivalente au niveau discret en multipliant par

$$\frac{u_h^{n+1} - u_h^n}{2\Delta t}.$$

Or on observe l'apparition d'identités remarquables

$$\begin{aligned} \left\langle \frac{u_h^{n+1} - 2u_h^n + u_h^{n-1}}{\Delta t^2}, \frac{u_h^{n+1} - u_h^n}{2\Delta t} \right\rangle_{\ell^2} &= \frac{1}{2\Delta t} \left(\left\| \frac{u_h^{n+1} - u_h^n}{\Delta t} \right\|_{\ell^2}^2 - \left\| \frac{u_h^n - u_h^{n-1}}{\Delta t} \right\|_{\ell^2}^2 \right), \\ \left\langle \frac{u_h^n - 2u_h^{n-1} + u_h^{n-2}}{h^2}, \frac{u_h^{n+1} - u_h^n}{2\Delta t} \right\rangle_{\ell^2} &= \frac{1}{\Delta t} (a_h(u_h^{n+1}, u_h^n) - a_h(u_h^n, u_h^{n-1})). \end{aligned}$$

Ceci conduit bien en additionnant ces deux relations à

$$\frac{1}{2\Delta t}(E_h^{n+1/2} - E_h^{n-1/2}) = 0,$$

c'est à dire le résultat annoncé. ■

Principe de la stabilité à partir de la conservation d'énergie

Une fois cette conservation établie il nous faut maintenant l'utiliser pour assurer un résultat de stabilité de type

$$\|u_h^n\|_{L^2(\mathbb{R})} \leq \|u_h^0\|_{L^2(\mathbb{R})}.$$

Le problème majeur dans l'étude discrète est que l'énergie discrète n'est pas nécessairement une quantité positive dû au terme

$$a_h(u_h^{n+1}, u_h^n).$$

Ainsi même si l'énergie se conserve on ne peut pas être sûr qu'on conserve une norme sur u_h^n . Toutefois si le pas de temps est suffisamment petit (nous voyons apparaître ici une condition de type CFL), u_h^{n+1} sera proche de u_h^n et $a_h(u_h^{n+1}, u_h^n)$ sera donc presque un carré. Dans ce cas, on pourra obtenir une forme « d'équivalence de norme » qui permettra de relier la stabilité de la norme $L^2(\mathbb{R})$ à celle de l'énergie.

On obtient alors le théorème suivant que nous admettrons

THÉORÈME 2.4.3 (*Admis*)

Soit u_h^n la solution discrète. Si la solution du problème est assez régulière alors

$$\alpha = \frac{c\Delta t}{h} < 1 \implies \forall T > 0 \lim_{\Delta t, h \rightarrow 0} \sup_{t^n \leq T} \|e_h^n\|_{L^2(\mathbb{R})} = 0$$

avec $e_h^n = u_h^n - u(t^n, \cdot)$ l'erreur commise sur la solution. Plus précisément, on montre que si on pose l'énergie discrète de l'erreur

$$E_h^{n+1/2}(e_h^n) = \frac{1}{2} \left\| \frac{e_h^{n+1} - e_h^n}{\Delta t} \right\|_{\ell^2}^2 + \frac{1}{2} a_h(e_h^{n+1}, e_h^n)$$

alors

$$E_h^{n+1/2}(e_h^n) > 0$$

et

$$\sqrt{E_h^{n+1/2}(e_h^n)} \leq \sqrt{E^{1/2}(e^0)} + \frac{\sqrt{2}}{2\sqrt{1-\alpha^2}} \Delta t \sum_{k=1}^n \|\varepsilon_h^k\|_{\ell^2}$$

où $\Delta t \sum_{k=1}^n \|\varepsilon_h^k\|_{\ell^2}$ est l'erreur de troncature du schéma sommée à chaque pas de temps.

Nous n'entrons pas dans les détails de la démonstration car notre objectif était simplement de vous montrer que l'écriture d'un schéma numérique, où les identités remarquables permettent d'écrire une énergie discrète, sont un outil extrêmement puissant pour l'obtention d'un schéma stable.

Chapitre 3

Problèmes hyperboliques non linéaires : étude des lois de conservation scalaires

3.1 Lois de conservation

3.1.1 Introduction

Dans ce chapitre, nous nous intéressons aux équations scalaires de la forme :

$$\frac{\partial u}{\partial t} + \frac{\partial}{\partial x} (f(u)) = 0, \quad (3.1)$$

où f est une fonction régulière de u . Comme nous l'expliquons au paragraphe 3.1.2, cette équation traduit la conservation de la quantité u . On dit qu'il s'agit d'une loi de conservation.

Plus particulièrement, nous nous intéresserons à l'équation aux dérivées partielles de Burgers :

$$\frac{\partial u}{\partial t} + \frac{\partial}{\partial x} \left(\frac{u^2}{2} \right) = 0.$$

Notre objectif est de résoudre le problème de Cauchy, qui consiste à trouver une fonction $u(x, t)$ telle que l'on ait :

$$\begin{cases} \frac{\partial u}{\partial t} + \frac{\partial}{\partial x} (f(u)) = 0, & \forall x \in \mathbb{R}, \quad \forall t > 0, \\ u(x, 0) = u^0(x), & \forall x \in \mathbb{R}, \end{cases}$$

où la condition initiale u^0 est une fonction donnée.

Notre but dans ce chapitre est d'ébaucher la théorie de ce problème. Nous montrons tout d'abord comment déterminer la solution classique, quand elle existe, par la méthode des caractéristiques (section 3.2.1). Cependant, de par le caractère non-linéaire de l'équation, de telles solutions classiques ne peuvent exister globalement en temps et, typiquement, des discontinuités peuvent apparaître et disparaître au cours du temps : ce sont les chocs. La fonction u est alors une solution faible du problème, au sens des distributions (section 3.3.1). Nous montrons que la vitesse de propagation du choc doit vérifier une condition, appelée condition de Rankine-Hugoniot (section 3.3.2).

Mais on vérifie aisément qu'il n'y a pas unicité de la solution faible du problème de Cauchy. L'unique solution physique est celle qui vérifie une condition supplémentaire dite **condition d'entropie** (section 3.3.3). Pour conclure (section 3.4), nous calculons explicitement la solution dans le cas du problème de Riemann où la condition initiale u^0 est de la forme :

$$u^0(x) = \begin{cases} u_g & \text{si } x < 0, \\ u_d & \text{si } x > 0. \end{cases}.$$

Cette étude théorique du problème continu est un préalable indispensable à l'étude des schémas d'approximation par différences finies que nous présenterons dans le cours suivant. En effet, une méthode numérique de discrétisation devra tenir compte au mieux des propriétés de la solution du problème continu.

Nous considérerons la plupart du temps une fonction f régulière mais **quelconque**. Toutefois, on verra que, lorsque la fonction f est **strictement convexe** (dont le prototype $f(u) = u^2/2$ donne l'**équation de Burgers**), un certain nombre de résultats prennent une forme plus simple. En particulier, nous nous limiterons à ce cas pour l'étude du problème de Riemann.

3.1.2 Construction d'une équation de conservation

L'équation (3.1) résulte en fait de l'écriture d'une simple loi de conservation.

En effet, considérons par exemple une répartition linéique de matière. Notons $u(x, t)$ la densité de matière en x ($x \in \mathbb{R}$) à l'instant t ($t \geq 0$) et soit $F(x, t)$ le flux de matière passant par le point x à l'instant t . Alors, la variation de la quantité de matière dans l'intervalle $[x_1, x_2]$ entre les temps t_1 et t_2 est égale à la différence entre le flux de matière entrant en x_1 et le flux sortant en x_2 pendant cette période. Autrement dit :

$$\int_{x_1}^{x_2} \{u(\xi, t_2) - u(\xi, t_1)\} d\xi = \int_{t_1}^{t_2} \{F(x_1, s) - F(x_2, s)\} ds. \quad (3.2)$$

Ceci s'écrit également :

$$\frac{1}{x_2 - x_1} \int_{x_1}^{x_2} \frac{u(\xi, t_2) - u(\xi, t_1)}{t_2 - t_1} d\xi = \frac{-1}{t_2 - t_1} \int_{t_1}^{t_2} \frac{F(x_1, s) - F(x_2, s)}{x_1 - x_2} ds.$$

Lorsque $(x_2 - x_1)$ et $(t_2 - t_1)$ tendent vers 0, si les fonctions u et F sont assez régulières, cette identité intégrale devient l'équation aux dérivées partielles suivante :

$$\frac{\partial u}{\partial t}(x, t) + \frac{\partial F}{\partial x}(x, t) = 0. \quad (3.3)$$

Supposons de plus que le flux F soit lui-même une fonction de la densité u :

$$F(x, t) = f(u(x, t)). \quad (3.4)$$

Alors l'équation (3.3) prend la forme suivante :

$$\frac{\partial u}{\partial t}(x, t) + \frac{\partial}{\partial x} f(u(x, t)) = 0. \quad (3.5)$$

On dit que l'équation (3.5) est sous **forme conservative**, par opposition à sa forme non conservative :

$$\frac{\partial u}{\partial t}(x, t) + c(u(x, t)) \frac{\partial u}{\partial x}(x, t) = 0 \quad (3.6)$$

où $c(u) = f'(u)$.

Dans le cas particulier où $f(u) = \frac{u^2}{2}$, les équations (3.5) et (3.6) s'écrivent :

$$\frac{\partial u}{\partial t} + \frac{\partial}{\partial x} \left(\frac{u^2}{2} \right) = 0, \quad (3.7)$$

$$\frac{\partial u}{\partial t} + u \frac{\partial u}{\partial x} = 0.$$

Il s'agit alors de l'équation de Burgers.

3.1.3 Exemples

La circulation automobile

L'équation (3.5) fournit un modèle simple pour la circulation automobile sur une route à une seule voie : u représente la quantité de voitures par centaine de mètres et F le flux de voitures par minute qui passent à l'heure t devant la borne x . On suppose que chaque conducteur ajuste la vitesse de sa voiture en fonction uniquement de la vitesse de la voiture qui le précède. Alors la conservation de la quantité de voitures (il n'y a ni station-service, ni itinéraire de délestage) se traduit par l'équation (3.5).

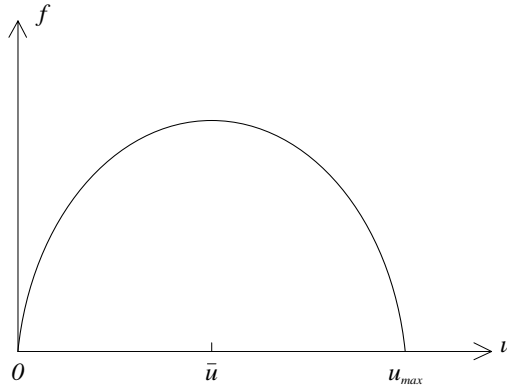
Si toutes les voitures roulaient à une vitesse constante donnée c , on aurait :

$$\begin{cases} f(u) = c u \\ \frac{\partial u}{\partial t} + c \frac{\partial u}{\partial x} = 0 \end{cases}$$

et on retrouve l'équation de transport qui est linéaire.

En réalité, le flux de voitures F est une fonction f non-linéaire de la densité de voitures u . Essayons d'imaginer à quoi peut ressembler cette fonction...

Bien sûr, le flux de voitures est nul si la densité est elle-même nulle, mais il est également nul si la densité est trop importante (u voisin de la densité de saturation u_{max}) : on a alors un embouteillage. Enfin, on observe un flux maximum au voisinage d'une densité intermédiaire idéale \bar{u} .



Supposons que la fonction f est strictement concave sur l'intervalle des densités admissibles $[0, u_{max}]$. Notons v l'écart entre la densité de voitures idéale \bar{u} et la densité effective u :

$$v(x, t) = \bar{u} - u(x, t).$$

On a alors, sur l'intervalle $[\bar{u} - u_{max}, \bar{u}]$:

$$\frac{\partial v}{\partial t} + \frac{\partial}{\partial x}(g(v)) = 0 \quad (3.8)$$

où g est la fonction strictement convexe donnée par $g(v) = -f(\bar{u} - v)$.

La dynamique des gaz

En réalité, la loi de conservation scalaire (3.5) a surtout été étudiée parce qu'elle constitue un modèle simple pour l'étude des **systèmes de lois de conservation** qui

modélisent de nombreux phénomènes physiques. L'exemple fondamental est celui de la dynamique des gaz.

En effet, considérons un fluide parfait compressible. Alors un écoulement unidimensionnel est modélisé, en coordonnées Eulériennes, par :

- L'équation de conservation de la masse

$$\frac{\partial \rho}{\partial t} + \frac{\partial}{\partial x}(\rho u) = 0 ;$$

- L'équation de conservation de la quantité de mouvement

$$\frac{\partial}{\partial t}(\rho u) + \frac{\partial}{\partial x}(\rho u^2 + p) = 0 ;$$

- L'équation de conservation de l'énergie

$$\frac{\partial}{\partial t}(\rho e) + \frac{\partial}{\partial x}((\rho e + p)u) = 0 ;$$

où ρ représente la densité du fluide, u la vitesse et e l'énergie totale spécifique. La pression p est donnée par une équation d'état en fonction des quantités ρ , u et e .

L'étude de ce système de trois équations et trois inconnues présente de nombreuses difficultés théoriques dont certaines n'ont pas encore été résolues à ce jour. C'est pourquoi l'on s'intéresse à l'équation scalaire (3.5) qui, bien que beaucoup plus simple à étudier, présente les mêmes caractères fondamentaux que le système précédent.

3.1.4 Ajout d'un terme de viscosité

Jusque là, nous avons fait l'hypothèse que le flux $F(x, t)$ avait la forme (3.4). Pour beaucoup de problèmes issus de la physique, il est plus réaliste de supposer que l'on a :

$$F(x, t) = f(u(x, t)) - \varepsilon \frac{\partial u}{\partial x}(x, t) \quad (3.9)$$

où ε est un petit paramètre strictement positif. L'équation (3.5) devient alors :

$$\frac{\partial u}{\partial t} + \frac{\partial}{\partial x}(f(u)) - \varepsilon \frac{\partial^2 u}{\partial x^2} = 0. \quad (3.10)$$

Le terme

$$-\varepsilon \frac{\partial^2 u}{\partial x^2}$$

permet de tenir compte des effets de **dissipation de l'énergie** et de **dispersion**. Par analogie avec la dynamique des gaz, on dit qu'il s'agit d'un **terme de viscosité**.

Pour l'exemple de la circulation automobile, l'expression du flux (3.9) traduit le fait que le conducteur ne tient pas seulement compte de la densité de voitures u là où il se trouve mais qu'il anticipe en fonction de ce qu'il aperçoit devant lui.

L'équation (3.5) est en fait la limite de l'équation visqueuse (3.10) lorsque le paramètre de viscosité ε tend vers 0. Or, contrairement à l'équation (3.5), l'équation (3.10) admet une solution unique et très régulière. Cette remarque, très importante, nous permettra de repérer la solution physique parmi toutes les solutions faibles de l'équation (3.5).

3.2 Solutions classiques : méthode des caractéristiques

On considère le problème suivant :

$$\frac{\partial u}{\partial t} + \frac{\partial}{\partial x}(f(u)) = 0 \quad t > 0, \quad x \in \mathbb{R} \quad (3.11)$$

$$u(x, 0) = u^0(x) \quad x \in \mathbb{R} \quad (3.12)$$

où la condition initiale u^0 est une fonction \mathcal{C}^1 de x . Rappelons que la fonction f est supposée être très régulière (disons au moins de classe \mathcal{C}^2) et posons :

$$c(u) = f'(u), \quad (3.13)$$

de sorte que c est une fonction **croissante** dès que f est **convexe**.

REMARQUE 3.2.1

Dans le cas linéaire, $f(u) = cu$, $c(u) = f'(u) = c$ n'est autre que la vitesse de propagation (constante) des solutions de l'équation de transport. En quelque sorte, dans le cas non linéaire, on peut dire que la vitesse de propagation dépend de la solution !

3.2.1 Solutions classiques au problème de Cauchy

DÉFINITION 3.2.1 (SOLUTION CLASSIQUE)

On dit que u est une solution classique de l'équation (3.11) dans un domaine ouvert \mathcal{Q} de $\mathbb{R} \times \mathbb{R}^+$ si c est une fonction \mathcal{C}^1 de x et de t , dans \mathcal{Q} et si elle satisfait (3.11) point par point dans \mathcal{Q} .

Nous allons montrer que le problème (3.11-3.12) admet, au moins pour $t < T \leq +\infty$, une solution classique (on parle de solution classique locale si $T < +\infty$) et que celle-ci peut être construite par la technique des droites caractéristiques.

REMARQUE 3.2.2 (CAS D'UNE DONNÉE INITIALE CONTINUE ET C^1 PAR MORCEAUX)

Dans tout ce paragraphe, nous supposons que la donnée initiale u^0 est C^1 et nous nous intéresserons aux solutions classiques du problème (3.11-3.12). Cependant, tous les résultats que nous allons établir peuvent se généraliser à une donnée initiale u^0 continue qui est seulement C^1 par morceaux. La méthode des caractéristiques permet alors de construire une fonction u continue et C^1 par morceaux qui vérifie l'équation (3.11) point par point dans tout ouvert où elle est C^1 , et telle que (3.12) soit satisfait. Nous verrons au paragraphe suivant qu'une telle fonction est une solution faible du problème.

Droites caractéristiques

Soit u une solution classique de l'équation (3.11). Alors u satisfait l'équation sous forme non conservative :

$$\frac{\partial u}{\partial t} + c(u) \frac{\partial u}{\partial x} = 0 \quad (3.14)$$

où $c(u)$ est donné par (3.13).

Considérons maintenant la courbe $(x(t), t)$ du plan (x, t) définie par :

$$\begin{cases} \frac{dx}{dt}(t) = c(u(x(t), t)) \\ x(\bar{t}) = \bar{x} \end{cases} \quad (3.15)$$

où (\bar{x}, \bar{t}) est un point quelconque du plan.

Alors on peut dériver u par rapport au temps le long de cette courbe et l'on obtient :

$$\frac{d}{dt}u(x(t), t) = \frac{\partial u}{\partial x}(x(t), t) \frac{dx}{dt}(t) + \frac{\partial u}{\partial t}(x(t), t).$$

D'après (3.14) et (3.15), cette quantité est nulle. Autrement dit, u est constant le long de la courbe (3.15) :

$$u(x(t), t) = u(\bar{x}, \bar{t}) \quad \forall t.$$

On en déduit que la courbe en question est une droite du plan (x, t) d'équation :

$$x = \bar{x} + c(u(\bar{x}, \bar{t}))(t - \bar{t}). \quad (3.16)$$

DÉFINITION 3.2.2 (DROITES CARACTÉRISTIQUES)

Soit une solution classique u de l'équation (3.11). Alors les droites de la forme (3.15) ou (3.16) sont appelées droites caractéristiques pour u de l'équation (3.11). De plus, u est constante le long de chacune de ces droites.

Construction d'une solution classique du problème de Cauchy

Supposons que u est, au moins pour $t < T$, une solution classique du problème (3.11-3.12). Considérons alors un point (\bar{x}, \bar{t}) de la forme $(\xi, 0)$. D'après (2.6), la droite caractéristique passant par $(\xi, 0)$ a pour équation :

$$x = \xi + c(u^0(\xi))t \quad (3.17)$$

et le long de cette droite, on a :

$$u(x, t) = u^0(\xi). \quad (3.18)$$

Ceci permet donc de calculer la valeur de u en tout point (x, t) par lequel passe une et une seule droite caractéristique.

Considérons tout d'abord le cas de l'équation de transport (la vitesse c est constante)

$$\frac{\partial u}{\partial t} + c \frac{\partial u}{\partial x} = 0.$$

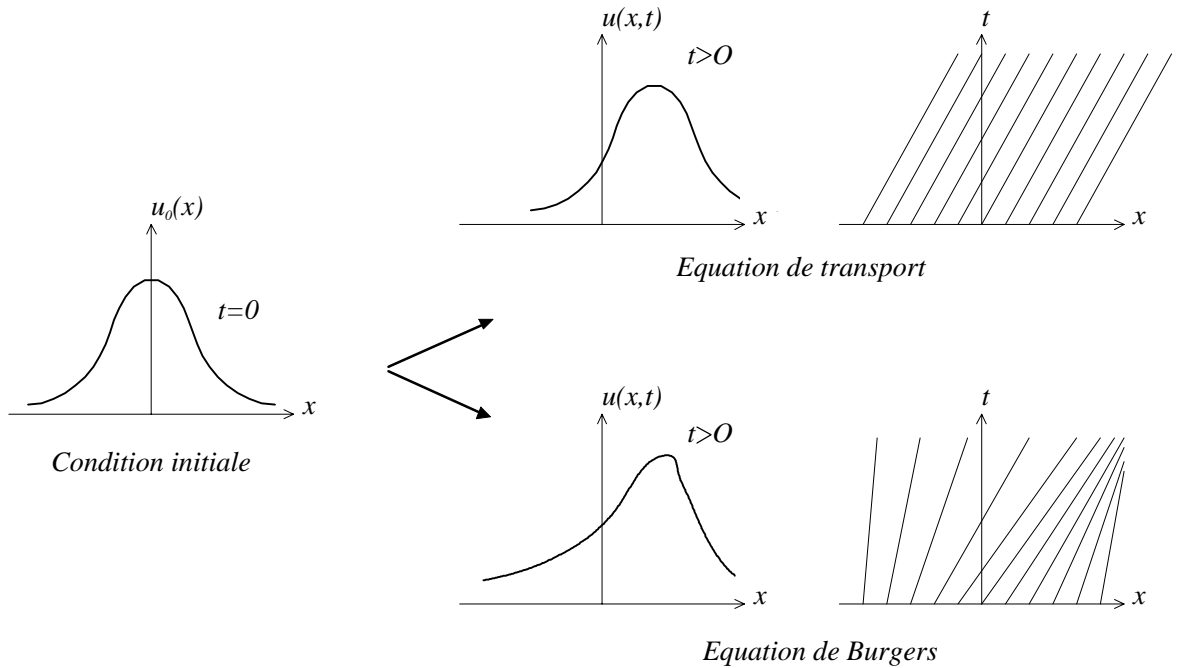
Les droites caractéristiques ont pour équation

$$x = \xi + ct$$

et sont donc parallèles. D'après (3.18), la solution u est alors donnée, $\forall t \geq 0$, par :

$$u(x, t) = u^0(x - ct).$$

Il s'agit d'une onde qui se propage « vers l'avant » à la vitesse c sans se déformer.



En revanche, dans le cas non-linéaire de l'équation (3.11), les droites caractéristiques ne sont pas parallèles. Toutes les valeurs de u^0 ne se propagent pas à la même vitesse. La solution se déforme au cours du temps.

REMARQUE 3.2.3 (ATTENTION !)

On représente toujours les droites caractéristiques dans le plan (x, t) et non dans le plan (t, x) . La pente de la droite d'équation (3.17) est donc $1/c(u^0(\xi))$. En particulier, si $c(u^0(\xi)) = 0$, la droite caractéristique issue de ξ est verticale.

Pour que la méthode des caractéristiques définisse sans ambiguïté la valeur de la solution au point (x, t) , il faut qu'il n'y ait qu'une droite caractéristique qui passe par ce point. Ceci nous amène à introduire la fonction

$$c_0(x) = c(u^0(x)). \quad (3.19)$$

Notons que $c_0(x)$ peut s'interpréter comme la vitesse de propagation initiale de la solution au point x : nous l'appellerons vitesse initiale. La caractéristique associée passant par le point $(\xi, 0)$ a pour équation :

$$x = \xi + c_0(\xi) t$$

Le cas d'une vitesse initiale croissante : existence d'une unique solution globale

THÉORÈME 3.2.1 (EXISTENCE D'UNE SOLUTION CLASSIQUE GLOBALE)

Si $c_0 = c(u^0)$ est croissante sur \mathbb{R} , le problème de Cauchy (3.11-3.12) admet une unique solution classique globale, construite par la méthode des caractéristiques.

DÉMONSTRATION: On fixe $t \in \mathbb{R}^+$. Si c_0 est croissante, la fonction continue :

$$\xi \longrightarrow F(\xi) = \xi + c_0(\xi) t$$

définit une bijection de \mathbb{R} dans \mathbb{R} . En effet :

$$\frac{dF}{d\xi}(\xi) = 1 + c'_0(\xi) t. \quad (3.20)$$

De plus, comme $c_0(\xi) \geq c_0(0)$ pour $\xi > 0$ et $c_0(\xi) \leq c_0(0)$ pour $\xi < 0$:

$$\lim_{\xi \rightarrow \pm\infty} F(\xi) = \pm\infty.$$

Pour tout couple (x, t) , il existe donc un unique $\xi(x, t)$ tel que :

$$x = \xi(x, t) + c_0(\xi(x, t)) t. \quad (3.21)$$

Si la solution classique existe, elle est donc donnée par :

$$u(x, t) = u^0(\xi(x, t)).$$

Vérifions que u est effectivement solution de l'équation (3.14). On a :

$$\begin{cases} \frac{\partial u}{\partial x}(x, t) = u'_0(\xi(x, t)) \frac{\partial \xi}{\partial x}(x, t) \\ \frac{\partial u}{\partial t}(x, t) = u'_0(\xi(x, t)) \frac{\partial \xi}{\partial t}(x, t) \end{cases}$$

et donc

$$\frac{\partial u}{\partial t}(x, t) + c(u(x, t)) \frac{\partial u}{\partial x}(x, t) = u'_0(\xi(x, t)) \left[\frac{\partial \xi}{\partial t}(x, t) + c_0(\xi(x, t)) \frac{\partial \xi}{\partial x}(x, t) \right] \quad (3.22)$$

Or, en dérivant (3.21) respectivement par rapport à x et t , il vient :

$$\begin{cases} \frac{\partial \xi}{\partial x}(x, t) [1 + c'_0(\xi) t] = 1 \\ \frac{\partial \xi}{\partial t}(x, t) [1 + c'_0(\xi) t] = -c_0(\xi(x, t)) \end{cases} \quad (3.23)$$

Mais par hypothèse :

$$c'_0(\xi(x, t)) \geq 0$$

donc la quantité entre crochets dans (3.23) reste strictement positive pour tout $t \geq 0$ et on déduit de (3.22) que l'équation (3.14) est donc vérifiée pour tout temps positif. ■

REMARQUE 3.2.4 (CAS CONVEXE)

Lorsque f est convexe, c est croissante et la condition « c_0 est croissante » devient simplement « u^0 est croissante »

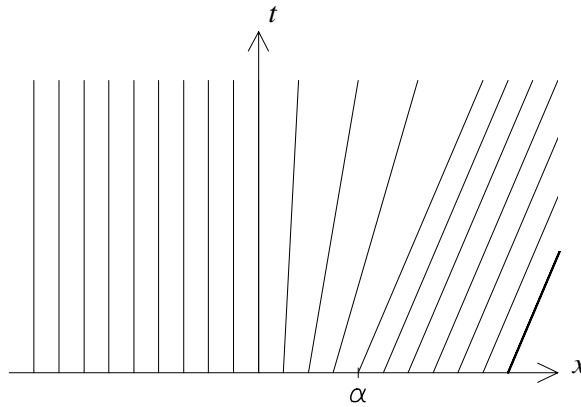
Exemple

Considérons par exemple la condition initiale suivante :

$$u^0(x) = \begin{cases} 0 & \text{si } x \leq 0 \\ \frac{x}{\alpha} & \text{si } 0 \leq x \leq \alpha \\ 1 & \text{si } x \geq \alpha \end{cases}$$

où α est un réel strictement positif. Nous nous permettons de considérer une donnée initiale qui n'est que \mathcal{C}^1 par morceaux en vertu de la remarque 3.2.2. On veut résoudre le problème de Cauchy (3.11-3.12) pour l'équation de Burgers ($f(u) = u^2/2$ et $c(u) = u$). D'après (3.17), la droite caractéristique issue de $(\xi, 0)$ a pour équation :

$$\begin{cases} x = \xi & \text{si } \xi \leq 0 \\ x = \xi + \frac{\xi}{\alpha} t & \text{si } 0 \leq \xi \leq \alpha \\ x = \xi + t & \text{si } \xi \geq \alpha \end{cases} .$$

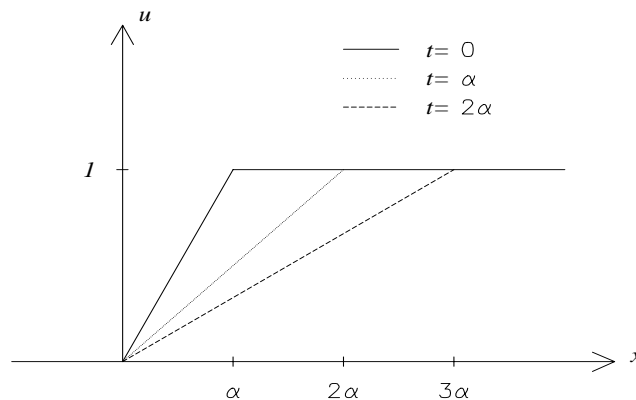


La solution u est alors donnée par :

$$u(x, t) = u^0(\xi) = \begin{cases} 0 & \text{si } x \leq 0 \\ \frac{x}{\alpha + t} & \text{si } 0 \leq x \leq \alpha + t \\ 1 & \text{si } x \geq \alpha + t \end{cases} .$$

On remarque que pour tout t fixé, la fonction u reste croissante en x .

On a représenté ci-dessous la solution u en fonction de x à différents temps :



3.2.2 Solutions classiques locales et naissance d'un choc

Nous abandonnons maintenant l'hypothèse que vitesse initiale c_0 définie par (3.19) est une fonction croissante. Dans ce cas, les solutions classiques peuvent ne pas exister pour tout temps mais seulement localement.

Dans cette section, pour des raisons purement techniques, nous ferons l'hypothèse

supplémentaire, peu restrictive en pratique :

$$\text{La fonction } x \mapsto u^0(x) \text{ est bornée sur } \mathbb{R}. \quad (3.24)$$

On en déduit bien sur que la vitesse initiale c_0 est elle même bornée.

Existence d'une solution classique locale

THÉORÈME 3.2.2 (TEMPS D'APPARITION D'UN CHOC)

Si la dérivée de c_0 prend des valeurs strictement négatives, alors il n'existe pas de solution classique pour tout temps $t > 0$. Le temps maximal t^ tel qu'il existe une solution classique pour $0 \leq t < t^*$ est donné par :*

$$t^* = \frac{1}{\sup_{\xi \in \mathbb{R}} \{-c'_0(\xi)\}} = \inf_{\xi \in \mathbb{R}} \{-c'_0(\xi)^{-1}\}.$$

REMARQUE 3.2.5

L'hypothèse assure que $0 < t^ < +\infty$.*

DÉMONSTRATION: Nous commençons par montrer l'existence d'une unique solution classique locale pour $0 < t < t^*$. On reprend en fait la démonstration du théorème 3.2.1. Pour tout $t < t^*$, la fonction $\xi \rightarrow F(\xi)$ définie dans la preuve du lemme 3.2.1 est strictement croissante (sa dérivée - cf. (3.20) - est supérieure à $1 - t/t^*$) et tend vers $\pm\infty$ quand $\xi \rightarrow \pm\infty$ (le lecteur notera que c'est l'hypothèse technique (3.24) qui permet de le garantir). Elle définit donc à nouveau une bijection de \mathbb{R} dans \mathbb{R} . L'équation (3.21) admet alors une solution $\xi(x, t)$ unique, ce qui permet d'établir l'existence d'une solution classique u pour $t \in [0, t^*[$.

Notons que, d'après les formules (3.23), les dérivées de la solution classique tendent vers l'infini lorsque t tend vers t^* . Démontrons directement que la solution classique ne peut exister au delà du temps t^* .

Soit t_{max} , le temps maximum t_{max} d'existence d'une solution classique, défini par

$$t_{max} = \sup \{ t > 0 / \text{il existe une solution classique sur } \mathbb{R}^+ \times [0, t[\}$$

Nous savons déjà que $t_{max} \geq t^*$. Supposons qu'il existe $\xi_2 > \xi_1$ tels que $c_0(\xi_1) > c_0(\xi_2)$. Alors, les droites caractéristiques issues des points ξ_1 et ξ_2 se croisent à l'instant \bar{t} et au point \bar{x} donnés par :

$$\begin{cases} \bar{t} = \frac{\xi_2 - \xi_1}{c_0(\xi_1) - c_0(\xi_2)}, \\ \bar{x} = \xi_1 + c_0(\xi_1) \bar{t} = \xi_2 + c_0(\xi_2) \bar{t} \end{cases}$$

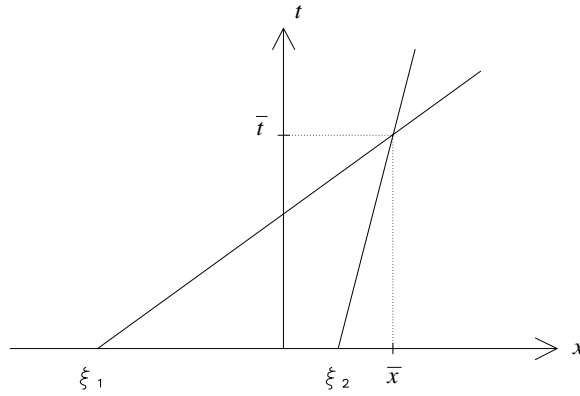
S'il existait une solution classique u jusqu'à l'instant \bar{t} , elle devrait donc vérifier simultanément $u(\bar{x}, \bar{t}) = u^0(\xi_1)$ et $u(\bar{x}, \bar{t}) = u^0(\xi_2)$, ce qui est impossible (comme $c_0(\xi_1) > c_0(\xi_2)$). Donc

$$\bar{t} \geq t_{max}.$$

Soit ξ tel que $c'_0(\xi) < 0$. Pour $h > 0$ assez petit on a nécessairement $c_0(\xi + h) < c_0(\xi)$ et donc d'après ce qui précède avec $\xi_1 = \xi$ et $\xi_2 = \xi + h$

$$t_{max} \leq \frac{h}{c_0(\xi) - c_0(\xi + h)} \implies t_{max} \leq -\frac{1}{c'_0(\xi)} \quad (\text{en faisant tendre } h \text{ vers } 0).$$

Ceci étant vrai pour tout ξ , on en déduit que $t_{max} \leq t^*$ ■



REMARQUE 3.2.6

La naissance d'une discontinuité est un phénomène purement non-linéaire, qui peut se produire même si u^0 est très régulière.

Exemple

Considérons cette fois la condition initiale suivante :

$$u^0(x) = \begin{cases} 1 & \text{si } x \leq 0, \\ 1 - \frac{x}{\alpha} & \text{si } 0 \leq x \leq \alpha, \\ 0 & \text{si } x \geq \alpha. \end{cases}$$

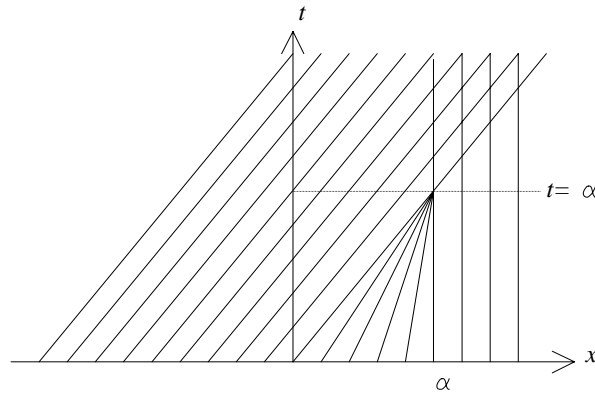
où α est un réel strictement positif.

Pour l'équation de Burgers, la droite caractéristique issue de ξ a pour équation :

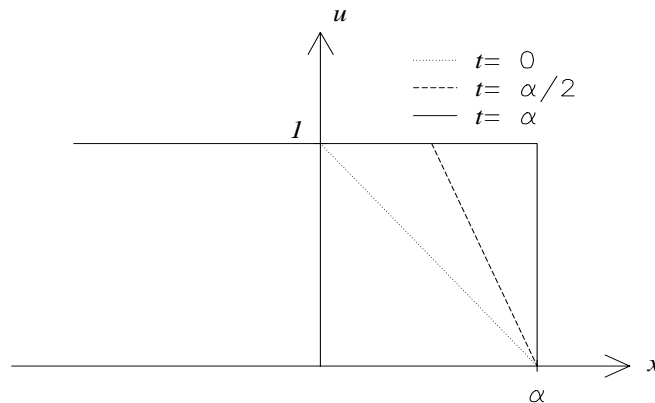
$$\begin{cases} x = \xi + t & \text{si } \xi \leq 0, \\ x = \xi + \left(1 - \frac{\xi}{\alpha}\right) t & \text{si } 0 \leq \xi \leq \alpha, \\ x = \xi & \text{si } \xi \geq \alpha. \end{cases}$$

Le temps critique t^* est égal à α . La solution u , pour $0 \leq t < \alpha$, vaut :

$$u(x, t) = \begin{cases} 1 & \text{si } x \leq t, \\ \frac{x - \alpha}{t - \alpha} & \text{si } t \leq x \leq \alpha, \\ 0 & \text{si } x \geq \alpha. \end{cases}$$



On peut représenter la solution pour différents temps :



On remarque que la solution reste décroissante en x pour tout t , $0 \leq t < \alpha$. Lorsque t augmente de 0 à α , la pente de la solution devient de plus en plus raide. A l'instant $t = \alpha$, la dérivée en x de u est infinie en $x = \alpha$. La solution u est discontinue en ce point.

3.3 Solutions faibles - Solutions entropiques

On a vu au paragraphe précédent que, même si la donnée initiale u^0 est très régulière, il peut ne pas exister de solution classique au problème de Cauchy (3.11-3.12) pour

tout temps $t > 0$. A partir d'un temps critique t^* , la solution u devient discontinue. En quel sens, peut-on prolonger cette solution au delà du temps t^* ? C'est la question à laquelle nous allons chercher à répondre.

3.3.1 Solutions faibles

Rappelons-nous que l'équation (3.11) a été obtenue comme conséquence de la loi de conservation intégrale (3.2). Or cette dernière a un sens pour les fonctions discontinues. Nous allons donc repartir de cette idée en reformulant cette notion à l'aide de fonctions tests.

Soit u une solution classique de (3.11-3.12) et φ une fonction $\mathcal{C}^1(\mathbb{R} \times [0, +\infty[)$ à support compact (c'est notre fonction test). Alors, on a, par intégration par parties :

$$\begin{aligned} 0 &= \int_{x=-\infty}^{+\infty} \int_{t=0}^{+\infty} \left(\frac{\partial u}{\partial t} + \frac{\partial}{\partial x}(f(u)) \right) \varphi \, dt \, dx \\ &= \int_{x=-\infty}^{+\infty} \int_{t=0}^{+\infty} \left(-u \frac{\partial \varphi}{\partial t} - f(u) \frac{\partial \varphi}{\partial x} \right) dt \, dx - \int_{x=-\infty}^{+\infty} u^0(x) \varphi(x, 0) \, dx. \end{aligned}$$

Cette égalité ne fait plus appel explicitement aux dérivées de u et a donc un sens même si u n'est pas a priori régulière : il suffit par exemple que u soit localement intégrable ! Cette observation nous conduit à introduire la notion de solution faible qui va seulement demander à la solution d'être localement bornée (donc en particulier localement intégrable) sur $\mathbb{R} \times [0, +\infty[$.

DÉFINITION 3.3.1 (SOLUTION FAIBLE)

Soit $u^0 \in L_{loc}^\infty(\mathbb{R})$. Une fonction $u \in L_{loc}^\infty(\mathbb{R} \times [0, +\infty[)$, est appelée une solution faible du problème de Cauchy (3.11-3.12) si elle satisfait

$$\int_{x=-\infty}^{+\infty} \int_{t=0}^{+\infty} \left(u \frac{\partial \varphi}{\partial t} + f(u) \frac{\partial \varphi}{\partial x} \right) dx \, dt + \int_{x=-\infty}^{+\infty} u^0(x) \varphi(x, 0) dx = 0 \quad (3.25)$$

pour tout φ de classe \mathcal{C}^1 à support compact dans $\mathbb{R} \times [0, +\infty[$.

Le lemme suivant assure la cohérence des définitions 3.2.1 et 3.3.1.

LEMME 3.3.1 (LIEN SOLUTION FAIBLE - SOLUTION CLASSIQUE)

Si u est une solution classique du problème de Cauchy (3.11-3.12), alors c'est une solution faible. Si $u \in \mathcal{C}^1(\mathbb{R} \times [0, +\infty[)$ est une solution faible du problème de Cauchy (3.11-3.12), alors c'est une solution classique.

DÉMONSTRATION:

La première partie a été démontrée. Réciproquement, soit $u \in \mathcal{C}^1(\mathbb{R} \times [0, +\infty[)$ une solution

faible de (3.11-3.12). Alors, d'après (3.25), on a :

$$\int_{x=-\infty}^{+\infty} \int_{t=0}^{+\infty} \left(u \frac{\partial \varphi}{\partial t} + f(u) \frac{\partial \varphi}{\partial x} \right) dt dx = 0$$

pour tout $\varphi \in \mathcal{D}(\mathbb{R} \times]0, +\infty[)$. Autrement dit, on a :

$$\frac{\partial u}{\partial t} + \frac{\partial}{\partial x}(f(u)) = 0$$

au sens des distributions. Mais comme u est régulière, ceci est en fait vérifié au sens classique, point par point.

Soit alors une fonction $\varphi \in \mathcal{C}^1(\mathbb{R} \times [0, +\infty[)$ à support compact dans $\mathbb{R} \times [0, +\infty[$. En intégrant (3.25) par parties, on obtient :

$$\begin{aligned} 0 &= - \int_{x=-\infty}^{+\infty} \int_{t=0}^{+\infty} \left(\frac{\partial u}{\partial t} + \frac{\partial}{\partial x}(f(u)) \right) \varphi dt dx + \int_{x=-\infty}^{+\infty} (u^0(x) - u(x, 0)) \varphi(x, 0) dx \\ &= \int_{x=-\infty}^{+\infty} (u^0(x) - u(x, 0)) \varphi(x, 0) dx. \end{aligned}$$

Ceci étant vrai pour tout φ , on en déduit (3.12). ■

3.3.2 Solutions \mathcal{C}^1 par morceaux. Condition de choc

Dans cette section, nous allons nous intéresser à une classe particulière de solutions faibles, les solutions \mathcal{C}^1 par morceaux.

DÉFINITION 3.3.2 (FONCTION \mathcal{C}^1 PAR MORCEAUX)

Soit \mathcal{O} un ouvert borné de $\mathbb{R} \times [0, +\infty[$. Une fonction $v(x, t)$ est dite \mathcal{C}^1 par morceaux sur \mathcal{O} s'il existe un nombre fini de courbes $\Sigma_1, \dots, \Sigma_p$ de la forme :

$$\Sigma_i : \begin{cases} x = \xi_i(t), t \in [t_i^{(1)}, t_i^{(2)}], \\ \text{où } \xi_i \text{ est une fonction } \mathcal{C}^1 \text{ de } t \end{cases}$$

telles que v soit égale à la restriction d'une fonction de $\mathcal{C}^1(\mathbb{R}^2)$ dans chaque composante connexe de $\mathcal{O} \setminus \Sigma_1 \cup \Sigma_2 \cdots \cup \Sigma_p$.

Une fonction $v(x, t)$ est dite \mathcal{C}^1 par morceaux sur $\mathbb{R} \times [0, +\infty[$ si elle est \mathcal{C}^1 par morceaux sur tout ouvert borné de $\mathbb{R} \times [0, +\infty[$.

Notons que si une fonction v , \mathcal{C}^1 par morceaux, peut admettre, à travers un nombre

fini de courbes Σ_j , des discontinuités en valeur ou en dérivée cette fonction, ainsi que ses dérivées en temps et en espace admettent des limites de part et d'autre de ces courbes.

Nous pouvons maintenant considérer des solutions du problème qui soient seulement \mathcal{C}^1 par morceaux. Nous allons voir cependant que les discontinuités d'une telle solution ne sont pas arbitraires.

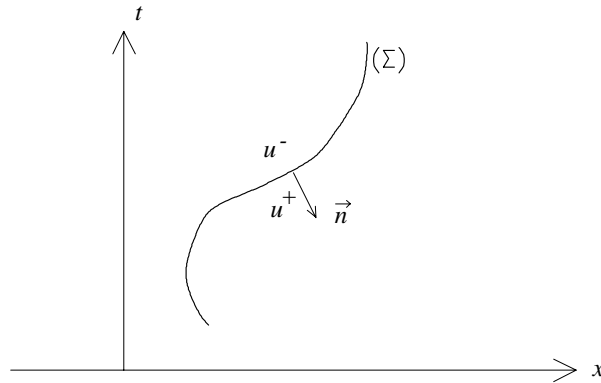
Pour fixer les notations, si u est une fonction \mathcal{C}^1 par morceaux dans le plan (x, t) et Σ une courbe de discontinuité de u de la forme $(\xi(t), t)$, on note :

- n la normale unitaire à Σ en $(\xi(t), t)$ orientée vers les x positifs, de sorte que :

$$n = \begin{pmatrix} n_x \\ n_t \end{pmatrix} = \frac{1}{\sqrt{1+s^2}} \begin{pmatrix} 1 \\ -s \end{pmatrix} \text{ où } s = \xi'(t). \quad (3.26)$$

- u^- et u^+ les valeurs de u « à gauche » et « à droite » de Σ :

$$\begin{cases} u^- = \lim_{\varepsilon \rightarrow 0} u((x, t) - \varepsilon n), \\ u^+ = \lim_{\varepsilon \rightarrow 0} u((x, t) + \varepsilon n). \end{cases}$$



Nous pouvons maintenant démontrer le théorème suivant

THÉORÈME 3.3.2 (EQUATIONS DES DISCONTINUITÉS)

Soit u une fonction \mathcal{C}^1 par morceaux dans $\mathbb{R} \times [0, +\infty[$. Alors u est une solution faible du problème de Cauchy (3.11-3.12) si et seulement si :

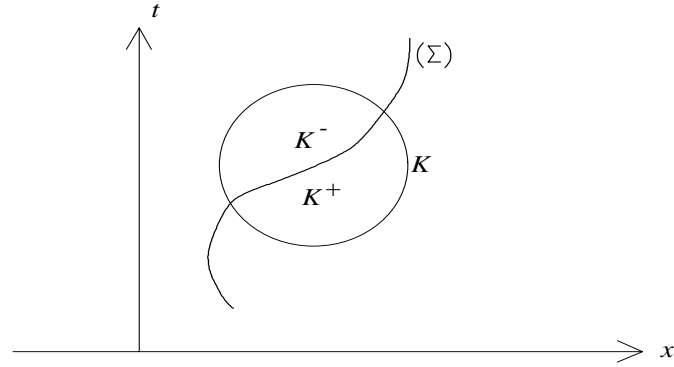
(i) u est une solution classique de (3.11-3.12) dans tout domaine où elle est \mathcal{C}^1

(ii) u satisfait la condition :

$$\xi'(t) (u^+ - u^-) = f(u^+) - f(u^-) \quad (3.27)$$

le long de toute courbe de discontinuité Σ .

DÉMONSTRATION: Soit u une fonction \mathcal{C}^1 par morceaux dans le plan (x, t) et Σ une courbe de discontinuité de u . Soit alors K un ouvert borné que la courbe Σ partage en deux composantes connexes K^- et K^+ . Pour simplifier, on supposera que K est contenu dans le demi-plan $t > 0$ mais ceci n'est pas restrictif. Enfin, on suppose que u est \mathcal{C}^1 dans l'adhérence de chacun des ouverts K^- et K^+ .



Soit $\varphi \in \mathcal{D}(K)$. Comme u est régulière sur K^- et K^+ :

$$\begin{aligned} \int_{K^-} \left(u \frac{\partial \varphi}{\partial t} + f(u) \frac{\partial \varphi}{\partial x} \right) dx dt &= - \int_{K^-} \left(\frac{\partial u}{\partial t} + \frac{\partial}{\partial x} (f(u)) \right) \varphi dx dt \\ &\quad + \int_{\Sigma} (u^- n_t + f(u^-) n_x) \varphi d\gamma, \end{aligned}$$

et de même :

$$\begin{aligned} \int_{K^+} \left(u \frac{\partial \varphi}{\partial t} + f(u) \frac{\partial \varphi}{\partial x} \right) dx dt &= - \int_{K^+} \left(\frac{\partial u}{\partial t} + \frac{\partial}{\partial x} (f(u)) \right) \varphi dx dt \\ &\quad - \int_{\Sigma} (u^+ n_t + f(u^+) n_x) \varphi d\gamma. \end{aligned}$$

D'où finalement :

$$\begin{aligned} \int_K \left(u \frac{\partial \varphi}{\partial x} + f(u) \frac{\partial \varphi}{\partial x} \right) dx dt &= - \int_{K^- \cup K^+} \left(\frac{\partial u}{\partial t} + \frac{\partial}{\partial x} (f(u)) \right) \varphi dx dt \\ &\quad - \int_{\Sigma} \{ (u^+ - u^-) n_t + (f(u^+) - f(u^-)) n_x \} \varphi d\gamma. \end{aligned} \quad (3.28)$$

Le théorème s'en déduit. En effet, si u est une solution faible du problème (3.11-3.12), elle satisfait l'équation (3.11) au sens des distributions dans K et au sens classique dans K^- et K^+ . Autrement dit :

$$0 = \int_K \left(u \frac{\partial \varphi}{\partial t} + f(u) \frac{\partial \varphi}{\partial x} \right) dx dt = \int_{K^- \cup K^+} \left(\frac{\partial u}{\partial t} + \frac{\partial}{\partial x} (f(u)) \right) dx dt.$$

D'après (3.28), on a donc :

$$\int_{\Sigma} \{ (u^+ - u^-) n_t + (f(u^+) - f(u^-)) n_x \} \varphi d\gamma = 0.$$

Comme ceci est vrai pour tout $\varphi \in \mathcal{D}(K)$, on en déduit en utilisant l'expression exacte de la normale (3.26) :

$$(u^+ - u^-)(-\xi'(t)) + f(u^+) - f(u^-) = 0,$$

le long de $\Sigma \cap K$.

Réciproquement, si u satisfait (i) et (ii), on a d'après (3.28) :

$$\int_K \left(u \frac{\partial \varphi}{\partial t} + f(u) \frac{\partial \varphi}{\partial x} \right) dx dt = 0$$

pour tout $\varphi \in \mathcal{D}(K)$, donc u est solution faible du problème. ■

On écrira souvent la relation (3.27) sous la forme synthétique :

$$s[u] = [f(u)] \text{ où } s = \xi'(t).$$

Par analogie à la dynamique des gaz, cette condition de choc est généralement appelée **condition de Rankine Hugoniot**.

REMARQUE 3.3.1 (EQUATION DE TRANSPORT - ÉQUATION DE BÜRGERS)

Considérons deux cas particuliers.

- Dans le cas particulier de l'équation de transport, on a : $f(u) = cu$. La condition de choc s'écrit donc : $s = c$. Autrement dit, dans le cas linéaire, les discontinuités se propagent nécessairement le long des caractéristiques.

- Dans le cas de l'équation de Burgers, la condition de Rankine-Hugoniot s'écrit :

$$s = \frac{u^+ + u^-}{2}.$$

Autrement dit, la vitesse du choc est la moyenne de la vitesse des caractéristiques de part et d'autre.

Nous sommes maintenant en mesure de justifier rigoureusement la remarque 3.2.2.

COROLLAIRE 3.3.3 (CAS DES FONCTIONS \mathcal{C}^1 PAR MORCEAUX ET CONTINUES)

Soit u une fonction \mathcal{C}^1 par morceaux et continue dans $\mathbb{R} \times [0, +\infty[$. Si u est une solution classique du problème (3.11-3.12) dans tout domaine où elle est \mathcal{C}^1 , alors c'est une solution faible du problème (3.11-3.12).

DÉMONSTRATION: En effet, la condition de choc (3.27) est automatiquement satisfaite car on a toujours $u^+ \equiv u^-$. ■

Exemple

Nous allons maintenant construire une solution faible globale pour la condition initiale de l'exemple section 3.2.2. On rappelle qu'il existait une solution continue jusqu'au temps $t = \alpha$. La fonction u vaut alors :

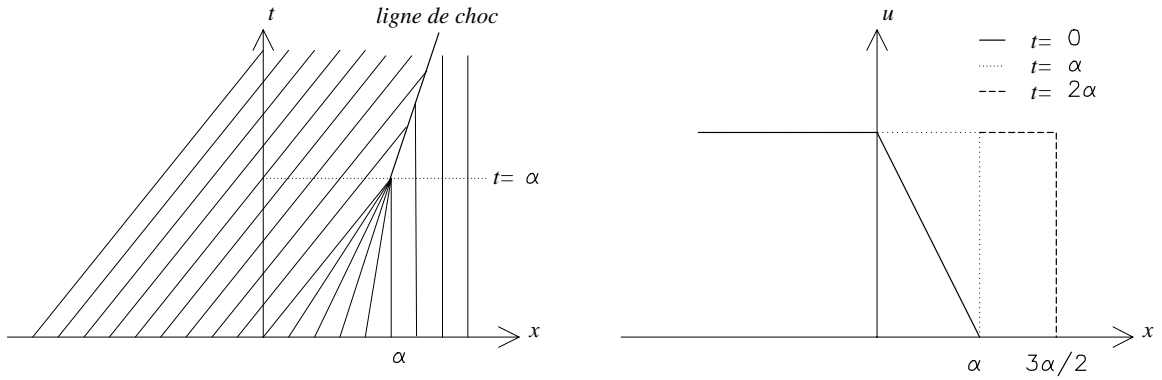
$$u(x, \alpha) = \begin{cases} 1 & \text{si } x < \alpha, \\ 0 & \text{si } x > \alpha. \end{cases}$$

Elle est discontinue en $x = \alpha$. Il nous faut maintenant calculer la vitesse de propagation de la discontinuité. D'après (3.27), on doit avoir : $\xi'(t) = \frac{1}{2}$. La fonction u donnée par :

$$u(x, t) = \begin{cases} 1 & \text{si } x < \alpha + \frac{t - \alpha}{2}, \\ 0 & \text{si } x > \alpha + \frac{t - \alpha}{2}. \end{cases}$$

est bien solution du problème pour $t > \alpha$.

On a représenté ci-dessous les caractéristiques, et la solution à différents temps.



3.3.3 Solutions entropiques

Non unicité de la solution faible - un exemple

Grâce à la notion de solution faible, nous avons calculé une solution globale au problème de Cauchy (3.11-3.12) dans des situations où il n'existait pas de solution classique. La question qui se pose alors est de savoir si cette solution faible globale est unique.

L'exemple suivant montre que ce n'est malheureusement pas le cas.

Considérons l'équation de Burgers pour la donnée initiale suivante :

$$u^0(x) = \begin{cases} 0 & \text{si } x < 0, \\ 1 & \text{si } x > 0. \end{cases}$$

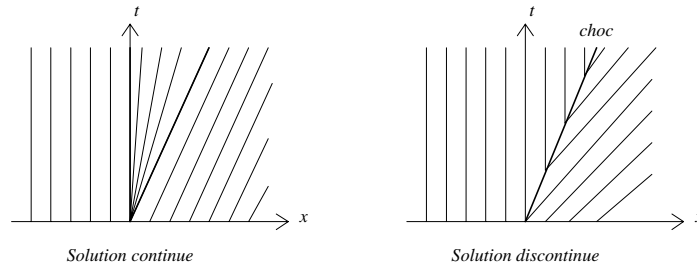
Nous pouvons alors calculer une solution continue au problème de Cauchy : pour cela, il suffit de considérer l'exemple section 3.2.1 et de faire tendre α vers 0. On obtient :

$$u(x, t) = \begin{cases} 0 & \text{si } x \leq 0, \\ \frac{x}{t} & \text{si } 0 \leq x \leq t, \\ 1 & \text{si } x \geq t, \end{cases}$$

On vérifie aisément a posteriori que u est solution du problème, car c'est une solution classique de l'équation de Burgers là où elle est \mathcal{C}^1 .

Mais en suivant la démarche de l'exemple section 3.3.2, nous pouvons aussi calculer une solution discontinue. La vitesse de propagation du choc est donnée par la relation $\xi'(t) = \frac{1}{2}$. La solution qui en résulte est donc :

$$u(x, t) = \begin{cases} 0 & \text{si } x < \frac{t}{2}, \\ 1 & \text{si } x > \frac{t}{2}. \end{cases}$$



REMARQUE 3.3.2

Il est intéressant de noter qu'une condition initiale discontinue peut donner naissance à une solution continue. C'est encore un phénomène purement non-linéaire.

Nous avons pu construire une deuxième solution au problème en introduisant une discontinuité. En fait, on peut même construire une infinité de solutions si l'on admet trois courbes de discontinuité.

En effet, pour tout $\gamma < 1$, la fonction suivante est solution du problème :

$$u(x, t) = \begin{cases} 0 & \text{si } x < \frac{\gamma}{2} t, \\ \gamma & \text{si } \frac{\gamma}{2} t < x < \frac{t}{2}, \\ 1 - \gamma & \text{si } \frac{t}{2} < x < (1 - \frac{\gamma}{2}) t, \\ 1 & \text{si } x > (1 - \frac{\gamma}{2}) t. \end{cases}$$

En fait, la solution physique est la solution continue.

Passage à la limite dans l'équation dissipative

Nous allons maintenant présenter un critère qui permettra de trier les solutions faibles pour extraire l'unique solution physique du problème.

Comme nous l'avons signalé, l'équation

$$\frac{\partial u}{\partial t} + \frac{\partial}{\partial x}(f(u)) = 0 \quad (3.29)$$

est généralement obtenue comme simplification de l'équation **visqueuse** ou **dissipative** :

$$\frac{\partial u_\varepsilon}{\partial t} + \frac{\partial}{\partial x}(f(u_\varepsilon)) - \varepsilon \frac{\partial^2 u_\varepsilon}{\partial x^2} = 0, \quad (3.30)$$

où ε est un petit paramètre strictement positif.

Nous admettons que le problème de Cauchy constitué de l'équation (3.30) et de la condition initiale

$$u(x, 0) = u^0(x) \quad \text{p.p. } x \in \mathbb{R} \quad (3.31)$$

où $u^0 \in L^\infty(\mathbb{R})$, possède une solution unique u_ε qui appartient à $L^\infty(\mathbb{R} \times [0, +\infty[)$. De plus, u_ε est très régulière sur $\mathbb{R} \times]0, +\infty[$.

Nous souhaitons donc caractériser la solution physique de (3.29-3.31), comme limite de u_ε lorsque ε tend vers 0.

LEMME 3.3.4

Soit $(u_\varepsilon)_{\varepsilon>0}$ une suite de solutions de (3.30) telles que :

$$\|u_\varepsilon\|_{L^\infty(\mathbb{R} \times [0, +\infty[)} \leq C \quad \forall \varepsilon > 0 \quad (3.32)$$

et

$$u_\varepsilon \longrightarrow u \text{ quand } \varepsilon \longrightarrow 0 \text{ presque partout dans } \mathbb{R} \times [0, +\infty[, \quad (3.33)$$

alors u est solution au sens des distributions de l'équation (3.29).

DÉMONSTRATION: En utilisant (3.32), (3.33) et le théorème de convergence dominée de Lebesgue, on vérifie aisément que $u_\varepsilon \rightarrow u$ et que $f(u_\varepsilon) \rightarrow f(u)$, au sens des distributions, c'est-à-dire dans $\mathcal{D}'(\mathbb{R} \times [0, +\infty[)$.

Comme la dérivation est une opération continue dans l'espace des distributions, on en déduit que :

$$\frac{\partial u_\varepsilon}{\partial t} \rightarrow \frac{\partial u}{\partial t}, \quad \frac{\partial^2 u_\varepsilon}{\partial x^2} \rightarrow \frac{\partial^2 u}{\partial x^2} \quad \text{et} \quad \frac{\partial}{\partial x}(f(u_\varepsilon)) \rightarrow \frac{\partial}{\partial x}(f(u))$$

dans $\mathcal{D}'(\mathbb{R} \times [0, +\infty[)$. Le lemme s'en déduit. ■

Pour caractériser la limite u des solutions de l'équation visqueuse, on utilise un critère qui repose sur la notion mathématique d'**entropie**. Précisons cela.

DÉFINITION 3.3.3 (ENTROPIE)

On appelle entropie pour l'équation (3.29) tout couple (U, F) de fonctions \mathcal{C}^1 de \mathbb{R} dans \mathbb{R} telles que :

- (i) U est une fonction strictement convexe,
- (ii) $F'(u) = U'(u) f'(u) \quad \forall u \in \mathbb{R}$.

Dès que $U(\cdot)$ et $F(\cdot)$ sont deux fonctions régulières reliées par la relation (ii) ci-dessus, si u est une solution classique de l'équation (3.29) on a :

$$U'(u) \frac{\partial u}{\partial t} + U'(u) f'(u) \frac{\partial u}{\partial x} = 0$$

d'où :

$$\frac{\partial}{\partial t} U(u) + \frac{\partial}{\partial x} F(u) = 0. \quad (3.34)$$

En revanche, si u est une solution faible de (3.29), par exemple \mathcal{C}^1 par morceaux, elle ne satisfait pas en général l'équation (3.34) au sens des distributions. En effet, il faudrait pour cela que la relation $s[U(u)] = [F(u)]$ soit satisfaite le long de toute courbe de discontinuité de u . Cela est généralement incompatible avec la relation de Rankine-Hugoniot.

Par contre, lorsque U est **convexe**, on peut montrer que, si la solution faible u est construite par passage à la limite sur l'équation avec viscosité, l'égalité (3.34) devient une inégalité au sens des distributions. Plus précisément, on peut démontrer le théorème suivant.

THÉORÈME 3.3.5

Soit $(u_\varepsilon)_{\varepsilon>0}$ une suite de solutions régulières de (3.30) vérifiant (3.32) et (3.33), et soit (U, F) une entropie pour l'équation (3.29). Alors u vérifie

$$\frac{\partial}{\partial t} U(u) + \frac{\partial}{\partial x} F(u) \leq 0 \quad (3.35)$$

au sens des distributions.

DÉMONSTRATION:

Comme u_ε est une solution régulière de (3.30), on a :

$$U'(u_\varepsilon) \frac{\partial u_\varepsilon}{\partial t} + U'(u_\varepsilon) f'(u_\varepsilon) \frac{\partial u_\varepsilon}{\partial x} - \varepsilon U'(u_\varepsilon) \frac{\partial^2 u_\varepsilon}{\partial x^2} = 0,$$

qui s'écrit aussi :

$$\frac{\partial}{\partial t} U(u_\varepsilon) + \frac{\partial}{\partial x} F(u_\varepsilon) - \varepsilon \frac{\partial^2}{\partial x^2} U(u_\varepsilon) = -\varepsilon U''(u_\varepsilon) \left(\frac{\partial u_\varepsilon}{\partial x} \right)^2.$$

Par des arguments similaires à ceux de la démonstration du lemme 3.3.4, on a :

$$\frac{\partial}{\partial t}U(u_\varepsilon) \rightarrow \frac{\partial}{\partial t}U(u), \frac{\partial}{\partial x}F(u_\varepsilon) \rightarrow \frac{\partial}{\partial x}F(u) \quad \text{et} \quad \frac{\partial^2}{\partial x^2}U(u_\varepsilon) \rightarrow \frac{\partial^2}{\partial x^2}U(u) \quad \text{dans} \quad \mathcal{D}'(\mathbb{R} \times [0, +\infty[).$$

En revanche, le terme $\varepsilon U''(u_\varepsilon) \left(\frac{\partial u_\varepsilon}{\partial x} \right)^2$ ne peut être traité de façon similaire.

Malgré la présence du facteur ε , ce terme ne tend pas nécessairement vers 0 au sens des distributions. Cela est dû au fait que les hypothèses n'empêchent pas $\frac{\partial u_\varepsilon}{\partial x}$ de tendre vers $+\infty$. Cependant, comme U est convexe, on a :

$$\left\langle -\varepsilon U''(u_\varepsilon) \left(\frac{\partial u_\varepsilon}{\partial x} \right)^2, \varphi \right\rangle \leq 0$$

pour toute fonction φ positive de $\mathcal{D}(\mathbb{R} \times [0, +\infty[)$. A la limite, on obtient donc :

$$\left\langle \frac{\partial}{\partial t}U(u) + \frac{\partial}{\partial x}F(u), \varphi \right\rangle \leq 0$$

pour toute fonction φ positive de

$$\mathcal{D}(\mathbb{R} \times [0, +\infty[).$$

■

La condition (3.35) est appelée **condition d'entropie**.

On peut maintenant définir la notion de solution entropique.

DÉFINITION 3.3.4 (SOLUTION ENTROPIQUE)

Une solution faible u du problème de Cauchy (3.29-3.31) est appelée solution entropique si elle satisfait la condition d'entropie (3.35) pour toute entropie (U, F) de l'équation (3.29).

Solutions entropiques C^1 par morceaux. Chocs entropiques

Pour les solutions C^1 par morceaux étudiées au paragraphe 3.3.2, la recherche de solutions entropiques va limiter les chocs admissibles : ce sont les chocs entropiques.

THÉORÈME 3.3.6 (ENTROPIE POUR LES FONCTIONS \mathcal{C}^1 PAR MORCEAUX)

Soit u une fonction \mathcal{C}^1 par morceaux dans $\mathbb{R} \times [0, +\infty[$ solution faible de (3.29-3.31), c'est à dire qui vérifie :

- u est une solution classique du problème de Cauchy (3.29-3.31) là où elle est \mathcal{C}^1 .
- u vérifie le long de toute courbe de discontinuité Σ :

$$\xi'(t)(u^+ - u^-) = f(u^+) - f(u^-).$$

Alors u est une solution entropique si et seulement si, pour toute entropie (U, F) , le long de toute courbe de discontinuité, Σ on a

$$\xi'(t) [U(u^+(t)) - U(u^-(t))] \geq F(u^+(t)) - F(u^-(t)). \quad (3.36)$$

DÉMONSTRATION: Soit u une fonction \mathcal{C}^1 par morceaux et (U, F) une entropie pour l'équation (3.29). En reprenant les notations et la démarche du théorème 3.3.2, on vérifie aisément que, pour tout $\varphi \in \mathcal{D}(K)$:

$$\begin{aligned} \int_K \left(U(u) \frac{\partial \varphi}{\partial t} + F(u) \frac{\partial \varphi}{\partial x} \right) dx dt &= - \int_{K^- \cup K^+} \left(\frac{\partial}{\partial t} U(u) + \frac{\partial}{\partial x} F(u) \right) \varphi dx dt \\ &\quad - \int_{\Sigma} \{ (U(u^+) - U(u^-)) n_t + (F(u^+) - F(u^-)) n_x \} \varphi d\gamma. \end{aligned} \quad (3.37)$$

Soit u une solution faible du problème satisfaisant la condition d'entropie (3.35). On a :

$$\int_K \left(U(u) \frac{\partial \varphi}{\partial t} + F(u) \frac{\partial \varphi}{\partial x} \right) dx dt \geq 0 \quad \text{et} \quad \int_{K^- \cup K^+} \left(\frac{\partial}{\partial t} U(u) + \frac{\partial}{\partial x} F(u) \right) \varphi dx dt = 0$$

pour toute fonction φ positive de $\mathcal{D}(K)$. D'après (3.37), on a, en utilisant l'expression exacte de la normale,

$$(U(u^+) - U(u^-)) (-\xi'(t)) + (F(u^+) - F(u^-)) \leq 0$$

au sens des distributions le long de Σ . Autrement dit, on a :

$$\xi'(t) (U(u^+) - U(u^-)) \geq (F(u^+) - F(u^-)) \quad \text{dans } \mathcal{D}'.$$

■

Le critère (3.36) est relativement peu pratique pour vérifier qu'un choc est entropique ou non. Heureusement, on peut en donner une forme équivalente plus simple, à commencer par le cas où f est convexe.

THÉORÈME 3.3.7 (CHOC ENTROPIQUE. CAS CONVEXE)

On suppose que f est strictement convexe. Alors la condition de choc entropique est équivalente à

$$u^- > u^+ \quad \text{le long de toute courbe de discontinuité } \Sigma. \quad (3.38)$$

Avant de démontrer ce théorème, nous allons montrer que la condition (3.38) signifie que lorsqu'il y a choc, les caractéristiques doivent **converger vers la ligne de choc** et non s'en écarter. On rappelle que c désigne la dérivée de f .

LEMME 3.3.8

Soit u une solution entropique du problème de Cauchy (3.29-3.31), qui soit C^1 par morceaux dans $\mathbb{R} \times [0, +\infty[$. Alors, u vérifie le long de toute courbe de discontinuité Σ :

$$c(u^+) < \xi'(t) < c(u^-).$$

DÉMONSTRATION: Comme f est strictement convexe, la fonction :

$$u \mapsto \frac{f(u) - f(u^+)}{u - u^+}$$

est continue et strictement croissante sur $[u^+, u^-]$. Il en résulte que :

$$c(u^+) < \frac{f(u^-) - f(u^+)}{u^- - u^+}.$$

Mais, par la relation de Rankine-Hugoniot, ceci s'écrit : $c(u^+) < \xi'(t)$. On démontre de même la seconde inégalité. ■

REMARQUE 3.3.3 (CAS PARTICULIER DE L'ÉQUATION DE BÜRGERS)

Dans le cas de l'équation de Burgers, on a tout simplement :

$$u^+ < s = \frac{u^- + u^+}{2} < u^-.$$

La condition d'entropie permet donc d'éliminer les solutions discontinues non physiques de l'exemple section 3.3.3.

DÉMONSTRATION DU THÉORÈME 3.3.7 : Nous partons de la fin de la démonstration du théorème 3.3.6. Considérons l'application G suivante :

$$u \mapsto G(u) = \frac{f(u) - f^-}{u - u^-} (U(u) - U(u^-)) - (F(u) - F(u^-)).$$

Nous allons montrer que G est une fonction décroissante de u . Admettons-le pour l'instant. Comme on a $G(u^-) = 0$ et que (3.36) s'écrit $G(u^+) \geq 0$, il résulte de la décroissance de G que les conditions (3.36) et (3.38) sont équivalentes, ce qui démontre le théorème.

Prouvons maintenant que G est décroissante. On a :

$$G'(u) = \frac{f'(u)(u - u^-) - (f(u) - f^-)}{(u - u^-)} (U(u) - U(u^-)) + \frac{f(u) - f^-}{u - u^-} U'(u) - F'(u).$$

Mais comme $F'(u) = U'(u)f'(u)$, ceci s'écrit aussi

$$\begin{aligned} G'(u) &= \frac{f'(u)(u - u^-) - (f(u) - f^-)}{(u - u^-)^2} (U(u) - U(u^-)) + \frac{f(u) - f^- - f'(u)(u - u^-)}{u - u^-} U'(u) \\ &= \frac{(f'(u)(u - u^-) - f(u) + f^-) (U(u) - U(u^-) - U'(u)(u - u^-))}{(u - u^-)^2}. \end{aligned}$$

Or f et U sont convexes, d'où :

$$f'(u)(u - u^-) - (f(u) - f^-) < 0 \quad \text{et} \quad U'(u)(u - u^-) - (U(u) - U(u^-)) < 0, \quad \forall u \in \mathbb{R},$$

ce qui achève la démonstration. ■

REMARQUE 3.3.4 (ENTROPIE ET DYNAMIQUE DES GAZ)

En dynamique des gaz, la condition d'entropie se déduit du second principe de la thermodynamique. En effet, ce principe implique que l'entropie augmente à travers un choc, de l'amont vers l'aval de l'écoulement. Il en résulte que la vitesse normale relative du fluide diminue de l'amont vers l'aval. C'est ce qui justifie l'appellation « condition d'entropie ».

Nous donnons maintenant la version générale du théorème 3.3.7. Sa démonstration est toutefois plus délicate que celle du théorème 3.3.7 et sera admise ici.

THÉORÈME 3.3.9 (CHOC ENTROPIQUE. CAS GÉNÉRAL)

La condition de choc entropique (3.36) est équivalente à $\forall \alpha \in [0, 1]$

$$\begin{cases} f(\alpha u^- + (1 - \alpha)u^+) \geq \alpha f(u^-) + (1 - \alpha)f(u^+) & \text{si } u^+ > u^- , \\ f(\alpha u^- + (1 - \alpha)u^+) \leq \alpha f(u^-) + (1 - \alpha)f(u^+) & \text{si } u^+ < u^- . \end{cases} \quad (3.39)$$

REMARQUE 3.3.5 (INTERPRÉTATION GÉOMÉTRIQUE)

La condition s'interprète facilement d'un point de vue géométrique. Un choc est entropique si et seulement si une des deux conditions suivantes est réalisée :

- (i) $u^+ > u^-$ et le graphe de $u \mapsto f(u)$ est au dessus de sa corde sur le segment $[u^-, u^+]$,
- (ii) $u^+ < u^-$ et le graphe de $u \mapsto f(u)$ est en dessous de sa corde sur le segment $[u^-, u^+]$.

Le graphe d'une fonction convexe étant toujours en dessous de sa corde et on retombe bien sur la condition $u^+ < u^-$.

3.3.4 Existence et unicité de la solution entropique

D'un point de vue mathématique, le gros avantage de la notion de solution entropique est qu'elle permet d'obtenir un résultat d'unicité. Pour conclure, nous allons énoncer, sans le démontrer, le résultat fondamental de la théorie des lois de conservation scalaire qui assure l'existence et l'unicité d'une solution entropique au problème de Cauchy.

THÉORÈME 3.3.10 (EXISTENCE ET UNICITÉ DE LA SOLUTION ENTROPIQUE)

Soit $u^0 \in L^\infty(\mathbb{R})$. Alors le problème de Cauchy (3.11-3.12) admet une solution entropique unique $u \in L^\infty(\mathbb{R} \times [0, +\infty[)$ qui vérifie l'estimation

$$\|u(\cdot, t)\|_{L^\infty(\mathbb{R})} \leq \|u^0\|_{L^\infty(\mathbb{R})} \text{ p.p. } t \geq 0.$$

Si u et v sont les solutions entropiques associées aux conditions initiales u^0 et v^0 , on a

$$\int_{x_1}^{x_2} |u(x, t) - v(x, t)| dx \leq \int_{x_1 - At}^{x_2 + At} |u^0(x) - v^0(x)| dx \quad (3.40)$$

pour presque tout $t \geq 0$, $\forall x_1 < x_2$, où

$$A = \max \left\{ |f'(\xi)| / |\xi| \leq \max \left(\|u^0\|_{L^\infty(\mathbb{R})}, \|v^0\|_{L^\infty(\mathbb{R})} \right) \right\}.$$

REMARQUE 3.3.6 (VITESSE DE PROPAGATION)

L'inégalité (3.40) signifie que la valeur de u au point (x, t) ne dépend que des valeurs de u^0 dans l'intervalle $[x - At, x + At]$. Autrement dit, la solution entropique a une vitesse de propagation finie.

REMARQUE 3.3.7 (A PROPOS DE LA DÉMONSTRATION DU THÉORÈME)

La démonstration de ce théorème est longue et difficile et fait appel à des outils et des arguments mathématiques qui dépassent largement le cadre de ce cours. Nous renvoyons à GODLEWSKI-RAVIART¹ pour la démonstration détaillée de ce théorème. Signalons toutefois que

- La démonstration du résultat d'existence est constructive et fait appel à la technique de viscosité. Les principales étapes sont les suivantes (pour une donnée u^0 assez régulière) :
 - On démontre tout d'abord, pour tout $\epsilon > 0$, l'existence et l'unicité de la solution (régulière) de (3.11-3.12).
 - A l'aide d'estimation a priori et de techniques de compacité, on montre que cette solution u_ϵ satisfait les hypothèses (3.32) et (3.33) du lemme 3.3.4.

¹E. GODLEWSKI, P.A. RAVIART : Hyperbolic systems of conservation laws, Mathématiques et Applications n°3 – 4, Publication de la SMAI, Ed. Ellipses, 1991.

- On utilise le lemme 3.3.4 et le théorème pour passer à la limite et montrer que la limite u est bien solution entropique de (3.30-3.12).
- Le résultat d'unicité est bien entendu une conséquence immédiate du résultat de « stabilité L^1 » (3.40). Il n'est pas très difficile de voir que (3.40) découle de l'inégalité au sens des distributions (pour tout couple (u, v) de solutions entropiques, $\text{sgn}(\cdot)$ désigne la fonction « signe »)

$$\frac{\partial}{\partial t} |u - v| + \frac{\partial}{\partial x} \left(\text{sgn}(u - v) (f(u) - f(v)) \right) \leq 0, \quad \text{dans } \mathcal{D}'(\mathbb{R} \times]0, +\infty[)$$

dont la démonstration, qui est elle beaucoup plus délicate, fait appel aux entropies de Kruzkov $U_k(x) = |u - k|$, $k \in \mathbb{R}$.

THÉORÈME 3.3.11 (MONOTONIE)

Si u et v sont les solutions entropiques de (3.11-3.12) associées aux conditions initiales u^0 et v^0 , on a

$$u^0 \geq v^0 \Rightarrow u(\cdot, t) \geq v(\cdot, t) \text{ p.p. } t \geq 0. \quad (3.41)$$

On en déduit en particulier que

$$\begin{cases} x \rightarrow u^0(x) \text{ croissante} & \Leftrightarrow x \rightarrow u(x, t) \text{ croissante, p.p. } t > 0, \\ a \leq u^0(x) \leq b \text{ p.p. } x \in \mathbb{R} & \Leftrightarrow a \leq u(x, t) \leq b \text{ p.p. } (x, t) \in \mathbb{R} \times \mathbb{R}^+. \end{cases} \quad (3.42)$$

Nous invitons le lecteur à démontrer (3.42) à partir de (3.41).

3.4 Le problème de Riemann

Dans cette section, nous nous limitons pour simplifier au cas où f est strictement convexe. Dans les cinq premières parties de cette section, nous considérons le problème de Riemann à 2 états. Dans la dernière partie, nous nous intéresserons au problème de Riemann à 3 états et on utilisera les résultats établis dans les sections précédentes.

3.4.1 Présentation du problème

Pour illustrer les résultats obtenus dans ce chapitre, nous allons calculer la solution entropique du problème de Cauchy suivant :

$$\begin{cases} \frac{\partial u}{\partial t} + \frac{\partial}{\partial x} f(u) = 0 & x \in \mathbb{R}, t > 0 \\ u(x, 0) = \begin{cases} u_g \text{ si } x < 0, \\ u_d \text{ si } x > 0, \end{cases} \end{cases} \quad (3.43)$$

où u_g et u_d sont deux constantes données.

REMARQUE 3.4.1 (LIEN AVEC LA DYNAMIQUE DES GAZ)

On dit qu'il s'agit d'un problème de Riemann par analogie à un problème étudié par ce dernier en dynamique des gaz : on considère un gaz dans un tube cylindrique très fin et très long. On suppose que le tube est séparé en deux par une membrane et que le gaz qui est au repos n'a pas les mêmes pression et densité des deux côtés de la membrane. A l'instant $t = 0$, la membrane est déchirée et l'on s'intéresse à l'écoulement qui s'en suit.

Nous avons déjà calculé la solution du problème (3.43) pour l'équation de Burgers dans les cas suivants :

- $u_g = 0$ et $u_d = 1$: c'est l'exemple de la section 3.2.1,
- $u_g = 1$ et $u_d = 0$: c'est l'exemple de la section 3.3.2,

et nous avons pu remarquer que les solutions obtenues étaient de natures très différentes.

Pour l'étude du problème général également, il nous faudra distinguer deux cas suivant que $u_g < u_d$ ou $u_g > u_d$.

3.4.2 Une solution auto-semblable

Avant tout, nous allons établir le lemme suivant.

LEMME 3.4.1 (SOLUTION AUTO-SEMBLABLE)

La solution u du problème de Riemann (3.43) est auto-semblable, c'est-à-dire qu'elle est de la forme

$$u(x, t) = v(x/t).$$

De plus, la fonction v doit vérifier, là où elle est \mathcal{C}^1 ,

$$v'(\xi) = 0 \quad \text{ou} \quad c(v(\xi)) = \xi. \quad (3.44)$$

DÉMONSTRATION: Soit $\alpha > 0$. Posons : $\bar{x} = \alpha x$, $\bar{t} = \alpha t$ et $\bar{u}(\bar{x}, \bar{t}) = u(x, t)$. Il est clair que \bar{u} est solution tout comme u du problème (3.43). Autrement dit, pour tout $\alpha > 0$:

$$u(x, t) = u(\alpha x, \alpha t).$$

Ceci signifie exactement que u est auto-semblable. De plus, si u est \mathcal{C}^1 , on doit avoir :

$$\frac{\partial u}{\partial t} + c(u) \frac{\partial u}{\partial x} = -\frac{x}{t^2} v'(x/t) + \frac{1}{t} c(v(x/t)) v'(x/t) = 0$$

D'où :

$$v'(x/t) (c(v(x/t)) - x/t) = 0.$$

■

Nous pouvons maintenant résoudre explicitement le problème.

3.4.3 Le cas d'une onde de détente

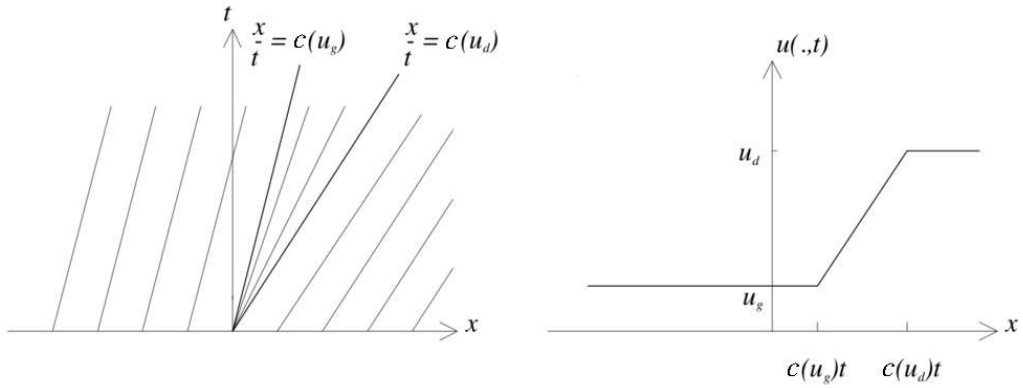
On se place dans le cas où $u_g < u_d$. Dans ce cas, tout comme pour l'exemple de la section 3.2.1, la solution entropique est continue. La méthode des caractéristiques permet de construire la solution dans les domaines $x \leq c(u_g)t$ et $x \geq c(u_d)t$. On a :

$$u(x, t) = \begin{cases} u_g & \text{si } \frac{x}{t} \leq c(u_g) \\ u_d & \text{si } \frac{x}{t} \geq c(u_d) \end{cases}.$$

Pour calculer u dans le domaine $c(u_g)t \leq x \leq c(u_d)t$, on résout l'équation (3.44). C'est toujours possible car f est strictement croissante, donc c est inversible. On obtient :

$$u(x, t) = c^{-1}\left(\frac{x}{t}\right) \text{ si } c(u_g) \leq \frac{x}{t} \leq c(u_d).$$

La fonction u ainsi construite est effectivement continue.



On dit qu'il s'agit d'une onde de détente.

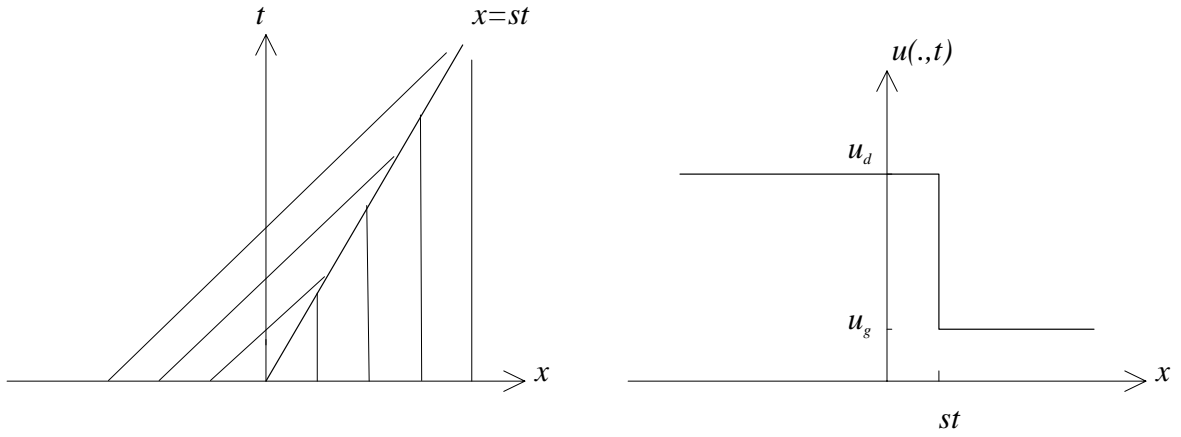
3.4.4 Le cas d'une onde de choc

On se place dans le cas où $u_g > u_d$. Dans ce cas, tout comme pour l'exemple de la section 3.3.2, la solution est constante de part et d'autre d'un choc. La vitesse de propagation du choc est donnée par la relation de Rankine-Hugoniot (elle est constante) :

$$s = \frac{f(u_g) - f(u_d)}{u_g - u_d}.$$

Le choc est entropique puisque $u_g > u_d$. La solution est

$$u(x, t) = \begin{cases} u_g & \text{si } x < st, \\ u_d & \text{si } x > st. \end{cases}$$



3.4.5 Solution du problème de Riemann à 2 états

En résumé, la solution entropique du problème de Riemann (3.43) peut être notée :

$$w_R\left(\frac{x}{t}; u_g, u_d\right)$$

- Si $u < v$:

$$w_R(\xi; u, v) = \begin{cases} u & \text{si } \xi \leq c(u) \\ c^{-1}(\xi) & \text{si } c(u) \leq \xi \leq c(v) \\ v & \text{si } \xi \geq c(v) \end{cases}$$

- Si $u > v$:

$$w_R(\xi; u, v) = \begin{cases} u & \text{si } \xi < s, \\ v & \text{si } \xi > s, \end{cases} \quad \text{où } s = \frac{f(u) - f(v)}{u - v}.$$

3.4.6 Le problème de Riemann à trois états

Nous nous proposons de résoudre explicitement le problème de Cauchy suivant (contrairement aux paragraphes précédents, nous nous restreignons ici au cas de l'équation

de Burgers)

$$\begin{cases} \frac{\partial u}{\partial t} + \frac{\partial}{\partial x} \left(\frac{u^2}{2} \right) = 0 & x \in \mathbb{R}, t > 0, \\ u(x, 0) = \begin{cases} u_1 & \text{si } x < 0, \\ u_2 & \text{si } 0 < x < 1, \\ u_3 & \text{si } x > 1, \end{cases} \end{cases}$$

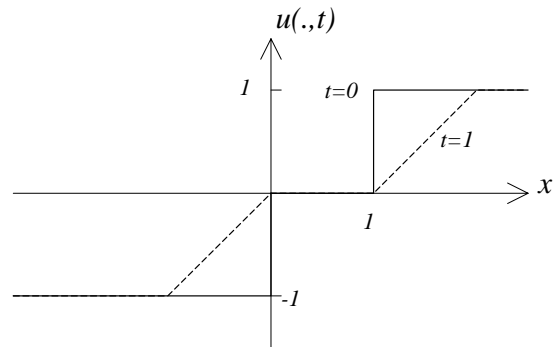
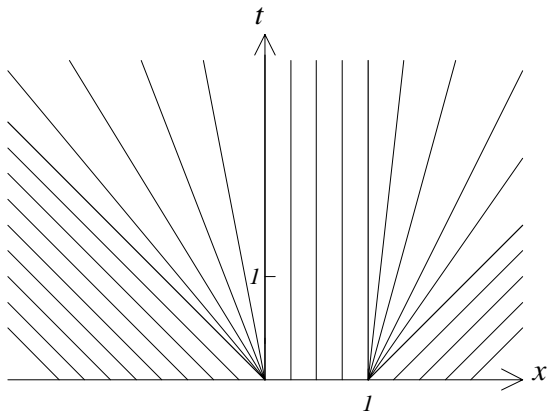
où u_1, u_2 et u_3 sont trois constantes données.

Cas d'une condition initiale croissante

On suppose $u_1 < u_2 < u_3$. Dans ce cas, la donnée initiale étant croissante, il existe une solution continue et croissante pour tout temps $t > 0$. Les deux détentes issues des points $x = 0$ et $x = 1$ n'interféreront jamais. La solution est donnée par :

$$u(x, t) = \begin{cases} u_1 & \text{si } x < u_1 t, \\ x/t & \text{si } u_1 t < x < u_2 t, \\ u_2 & \text{si } u_2 t < x < u_2 t + 1, \\ (x - 1)/t & \text{si } u_2 t + 1 < x < u_3 t + 1, \\ u_3 & \text{si } u_3 t + 1 < x. \end{cases}$$

Application numérique : $u_1 = -1, u_2 = 0, u_3 = 1$.



Cas d'une condition initiale décroissante

On suppose $u_1 > u_2 > u_3$. A l'instant $t = 0$ naissent deux chocs entropiques issus des points $x = 0$ et $x = 1$. Ils se rencontreront à un temps ultérieur t^* qu'il faudra déterminer. La trajectoire du premier choc $x = \xi_1(t)$ est donnée, pour t proche de 0, par la relation de Rankine-Hugoniot :

$$\xi_1(t) = \frac{1}{2} (u_1 + u_2) t$$

et celle du second choc $x = \xi_2(t)$ par :

$$\xi_2(t) = \frac{1}{2} (u_2 + u_3) t + 1.$$

Le temps t^* satisfait donc l'équation suivante :

$$\frac{1}{2} (u_1 + u_2) t^* = \frac{1}{2} (u_2 + u_3) t^* + 1$$

d'où :

$$t^* = \frac{2}{u_1 - u_3}.$$

A cet instant, on a :

$$\xi_1(t) = \xi_2(t) = \frac{u_1 + u_2}{u_1 - u_3}.$$

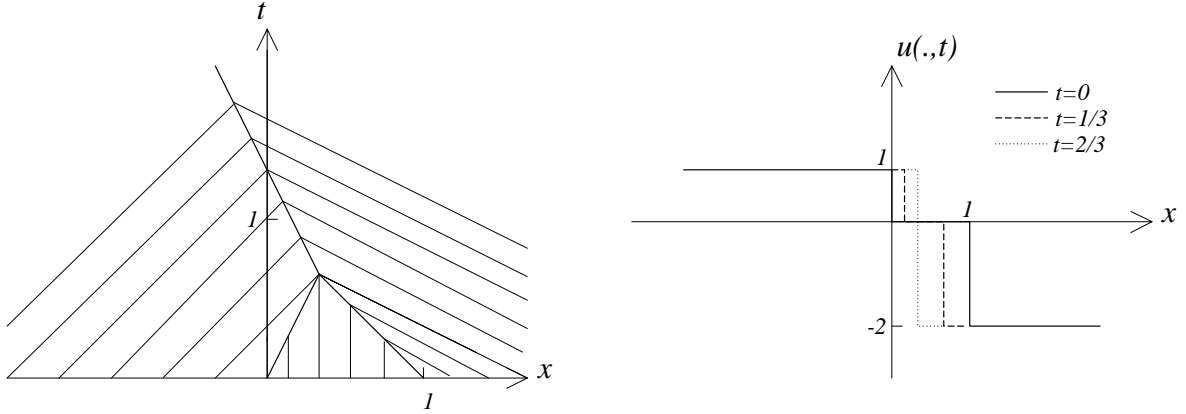
Au delà du temps t^* , il ne reste qu'un seul choc dont la trajectoire est donnée par :

$$x = \xi(t) = \frac{1}{2} (u_1 + u_3) t + \frac{u_2 - u_3}{u_1 - u_3}.$$

La solution est finalement donnée par

$$\begin{aligned} \text{si } t < t^*, \quad u(x, t) &= \begin{cases} u_1 & \text{si } x < \xi_1(t), \\ u_2 & \text{si } \xi_1(t) < x < \xi_2(t), \\ u_3 & \text{si } \xi_2(t) < x, \end{cases} \\ \text{si } t > t^*, \quad u(x, t) &= \begin{cases} u_1 & \text{si } x < \xi(t), \\ u_3 & \text{si } \xi(t) < x, \end{cases} \end{aligned}$$

Application numérique : $u_1 = 1$, $u_2 = 0$, $u_3 = -2$.



Cas d'une condition initiale non monotone

On suppose $u_1 < u_2$ et $u_2 > u_3$. Le dernier cas ($u_1 > u_2$ et $u_2 < u_3$) est laissé en exercice au lecteur !!! On observe donc une détente issue du point $x = 0$ et un choc entropique issu du point $x = 1$. Ce choc et cette détente vont interférer à partir d'un temps t^* que l'on déterminera. Pour $t < t^*$, la trajectoire du choc est donnée par :

$$x = \xi(t) \text{ où } \xi(t) = \frac{1}{2} (u_2 + u_3) t + 1.$$

Par conséquent, on a, pour $t < t^*$:

$$u(x, t) = \begin{cases} u_1 & \text{si } x < u_1 t, \\ x/t & \text{si } u_1 t < x < u_2 t, \\ u_2 & \text{si } u_2 t < x < \xi(t), \\ u_3 & \text{si } x > \xi(t). \end{cases}$$

Le temps t^* satisfait donc l'équation :

$$u_2 t^* = \xi(t^*)$$

d'où :

$$t^* = \frac{2}{u_2 - u_3}.$$

Pour $t > t^*$, la trajectoire du choc $x = \xi(t)$ est influencée par la détente. En effet, la valeur de la solution à gauche du choc est :

$$u(\xi(t), t) = \frac{\xi(t)}{t},$$

et par conséquent, la relation de Rankine-Hugoniot s'écrit :

$$\xi'(t) = \frac{1}{2} \left(\frac{\xi(t)}{t} + u_3 \right).$$

La solution générale de cette équation différentielle est donnée par :

$$\xi(t) = \alpha \sqrt{t} + u_3 t.$$

Comme de plus, on sait que :

$$\xi(t^*) = \frac{2u_2}{u_2 - u_3}$$

on trouve :

$$\alpha = \sqrt{2(u_2 - u_3)}.$$

Pour $t > t^*$, la solution se prolonge donc par :

$$u(x, t) = \begin{cases} u_1 & \text{si } x < u_1 t, \\ x/t & \text{si } u_1 t < x < \xi(t), \\ u_3 & \text{si } x > \xi(t). \end{cases}$$

Par la suite, de deux choses l'une. Soit $u_1 < u_3$ et les formules ci-dessus restent valables pour tout temps $t > t^*$. Soit $u_1 > u_3$ et ces formules restent valables tant que

$$u_1 t < \xi(t),$$

c'est-à-dire pour $t < t^{**}$ où t^{**} est solution de

$$u_1 t^{**} = \sqrt{2(u_2 - u_3)t^{**}} + u_3 t^{**}.$$

On a donc :

$$t^{**} = \frac{2(u_2 - u_3)}{(u_1 - u_3)^2}.$$

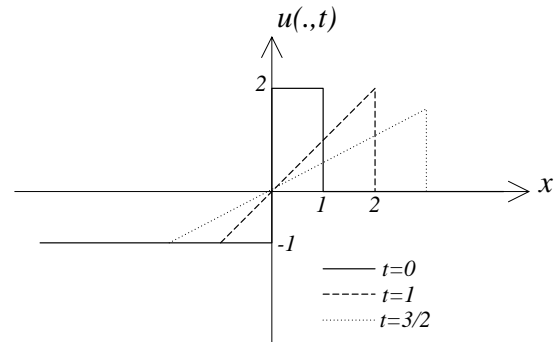
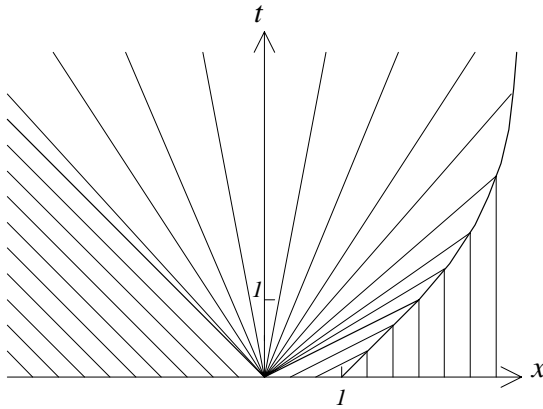
Pour $t > t^{**}$, la trajectoire du choc est donnée par :

$$\xi(t) = \frac{1}{2} (u_1 + u_3) t + \frac{u_2 - u_3}{u_1 - u_3}$$

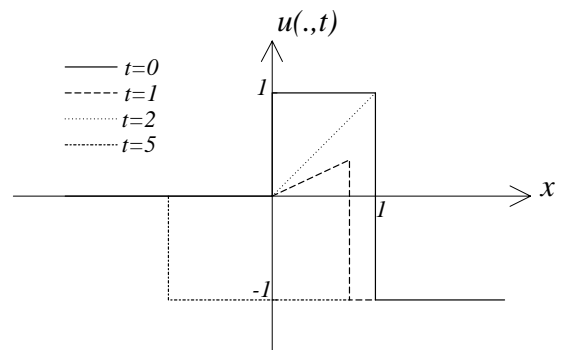
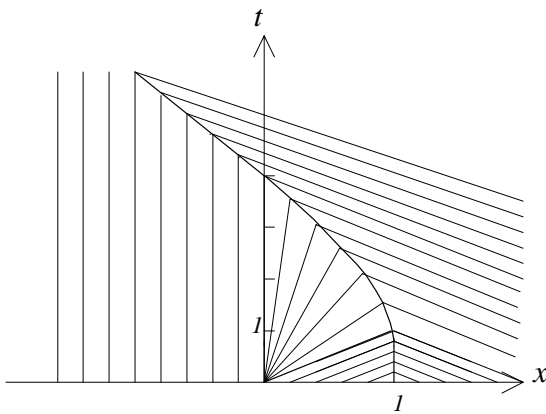
et la solution par :

$$u(x, t) = \begin{cases} u_1 & \text{si } x < \xi(t), \\ u_3 & \text{si } x > \xi(t). \end{cases}$$

Application numérique : $u_1 = 0$, $u_2 = 1$, $u_3 = -1$.



Le cas $u_1 = -1$, $u_2 = 2$, $u_3 = 0$



Le cas $u_1 = 0$, $u_2 = 1$, $u_3 = -1$

Chapitre 4

Discrétisation des équations hyperboliques non linéaires 1D

4.1 Présentation du problème

Ce chapitre est consacré à l'approximation numérique par différences finies du problème de Cauchy :

$$\left\{ \begin{array}{l} \text{Trouver } u(x, t) : \mathbb{R} \times \mathbb{R}^+ \rightarrow \mathbb{R} \text{ tel que} \\ \frac{\partial u}{\partial t} + \frac{\partial}{\partial x} f(u) = 0, \quad x \in \mathbb{R}, t > 0, \\ u(x, 0) = u^0(x), \quad x \in \mathbb{R}, \end{array} \right. \quad (4.1)$$

où f est une fonction strictement convexe (par exemple, $f(u) = \frac{u^2}{2}$ pour l'équation de Burgers).

On se donne un pas de discrétisation en espace $h > 0$ et un pas de discrétisation en temps $\Delta t > 0$. On pose :

$$\left\{ \begin{array}{l} x_j = jh, \quad j \in \mathbb{Z}, \\ t^n = n\Delta t, \quad n \in \mathbb{N}, \end{array} \right.$$

et on se propose de calculer une approximation u_j^n de $u(x_j, t^n)$ ¹ où u est la solution de (4.1).

¹ou plutôt une approximation de $\frac{1}{h} \int_{x_j - \frac{h}{2}}^{x_j + \frac{h}{2}} u(x, t^n) dx$ si l'on considère une approche par volumes finis (cf. section 4.2.1)

Dans toute la suite, nous nous intéressons à des schémas explicites à un pas de temps et à trois pas d'espace qui s'écrivent

$$u_j^{n+1} = H(u_{j-1}^n, u_j^n, u_{j+1}^n), \quad (4.2)$$

où H est une fonction régulière. Les schémas implicites sont difficiles à mettre en oeuvre pour les problèmes non-linéaires et nous n'en parlerons pas ici.

4.2 Propriétés des schémas

4.2.1 Schémas sous forme conservative

Nous avons vu que la forme conservative :

$$\frac{\partial u}{\partial t} + \frac{\partial}{\partial x} f(u) = 0 \quad (4.3)$$

et la forme non conservative :

$$\frac{\partial u}{\partial t} + f'(u) \frac{\partial u}{\partial x} = 0$$

des lois de conservation ne sont équivalentes que pour les solutions u régulières du problème et que la notion de solution faible repose sur la forme conservative de l'équation. On suit le même principe pour le calcul numérique d'une solution faible ; on utilisera un schéma « sous forme conservative ».

Le principe de base des schémas conservatifs est la méthode des volumes finis qui vise à trouver une approximation de $u_j(t) = \frac{1}{h} \int_{x_{j-\frac{1}{2}}}^{x_{j+\frac{1}{2}}} u(x, t) dx$ à tous les temps $t = t^n$

où $x_{j\pm\frac{1}{2}} = x_j \pm \frac{h}{2}$. Pour cela, on intègre tout d'abord l'équation (4.3) entre $x_{j-\frac{1}{2}}$ et $x_{j+\frac{1}{2}}$ (et on divise par h),

$$\frac{d}{dt} u_j(t) + \frac{1}{h} \int_{x_{j-\frac{1}{2}}}^{x_{j+\frac{1}{2}}} \frac{\partial}{\partial x} f(u) dx = 0$$

ou de manière équivalente

$$\frac{d}{dt} u_j(t) + \frac{f(u(x_{j+\frac{1}{2}}, t)) - f(u(x_{j-\frac{1}{2}}, t))}{h} = 0. \quad (4.4)$$

Le schéma numérique conservatif est une approximation de l'équation (4.4) : on approche la quantité $f(u(x_{j+\frac{1}{2}}, t))$ (resp. $f(u(x_{j-\frac{1}{2}}, t))$) en utilisant les $u_l(t)$. Un choix naturel consiste à prendre $f(u(x_{j+\frac{1}{2}}, t)) \approx g(u_j(t), u_{j+1}(t))$ et $f(u(x_{j-\frac{1}{2}}, t)) \approx g(u_{j-1}(t), u_j(t))$ où g est le flux numérique qu'il faut alors définir. Après une discrétisation en temps

explicite, on obtient le schéma (4.5).

DÉFINITION 4.2.1 (SCHÉMA CONSERVATIF)

On dit que le schéma (4.2) peut se mettre sous forme conservative s'il existe une fonction g régulière telle que l'on ait :

$$u_j^{n+1} = u_j^n - \frac{\Delta t}{h} \{g(u_j^n, u_{j+1}^n) - g(u_{j-1}^n, u_j^n)\} \quad (4.5)$$

La fonction g est appelée le flux numérique du schéma (elle est définie à une constante additive près).

On utilisera souvent la notation abrégée suivante

$$u_j^{n+1} = u_j^n - \alpha \left(g_{j+\frac{1}{2}}^n - g_{j-\frac{1}{2}}^n \right).$$

où :

$$\begin{cases} \alpha &= \frac{\Delta t}{h}, \\ g_{j+\frac{1}{2}}^n &= g(u_j^n, u_{j+1}^n). \end{cases}$$

REMARQUE 4.2.1

Remarque sur le schéma décentré Le schéma suivant

$$u_j^n = \begin{cases} u_j^n - \alpha c(u_j^n) (u_j^n - u_{j-1}^n) & \text{si } c(u_j^n) \geq 0 \\ u_j^n - \alpha c(u_j^n) (u_{j+1}^n - u_j^n) & \text{si } c(u_j^n) \leq 0 \end{cases}$$

est une généralisation naturelle au cas non-linéaire du schéma décentré. Pourtant, ce schéma est à proscrire car (on le vérifie aisément) il ne peut pas être mis sous forme conservative.

Si on applique ce schéma à la donnée initiale suivante :

$$u_j^0 = \begin{cases} u_g & \text{si } j \leq 0 \\ u_d & \text{si } j > 0 \end{cases}$$

où u_g et u_d sont tels que $c(u_g) \geq 0$ et $c(u_d) \leq 0$, on obtient une solution stationnaire :

$$u_j^n = u_j^0 \quad \forall j, \forall n$$

qui n'est pas, en général, une solution faible de l'équation.

L'importance de la forme conservative pour un schéma sera illustrée par le théorème de Lax-Wendroff (théorème 4.2.2) énoncé plus loin. Mais avant tout, il nous faut préciser sous quelle condition, portant sur le flux numérique, un schéma sous forme conservative est consistant avec l'équation (4.3).

4.2.2 Consistance

Les notions d'erreur de troncature, de consistance et d'ordre sont définies comme dans le cas linéaire. L'erreur de troncature du schéma (4.5) s'écrit

$$\varepsilon_j^n = \frac{1}{\Delta t} (\bar{u}_j^{n+1} - \bar{u}_j^n) + \frac{1}{h} (\bar{g}_{j+\frac{1}{2}}^n - \bar{g}_{j-\frac{1}{2}}^n),$$

où

$$\begin{cases} \bar{u}_j^n &= \bar{u}(x_j, t^n). \\ \bar{g}_{j+\frac{1}{2}}^n &= \bar{g}(\bar{u}_j^n, \bar{u}_{j+1}^n), \end{cases}$$

et \bar{u} est la solution régulière de l'équation (4.3).

Remarquons en particulier que l'erreur de troncature se calcule à l'aide de solutions **régulières** de l'équation (4.3). Ces notions n'auront donc pas de sens au voisinage des chocs.

LEMME 4.2.1 (CONDITION DE CONSISTANCE)

Un schéma mis sous la forme conservative (4.5) est consistant avec l'équation (4.3) si et seulement si :

$$g(u, u) = f(u) + cte \quad \forall u \in \mathbb{R}. \quad (4.6)$$

Il est alors d'ordre 1 au moins.

DÉMONSTRATION: Soit un schéma de la forme :

$$u_j^{n+1} = u_j^n - \alpha \left(g_{j+\frac{1}{2}}^n - g_{j-\frac{1}{2}}^n \right).$$

Calculons son erreur de troncature

$$\varepsilon_j^n = \frac{1}{\Delta t} (\bar{u}_j^{n+1} - \bar{u}_j^n) + \frac{1}{h} (\bar{g}_{j+\frac{1}{2}}^n - \bar{g}_{j-\frac{1}{2}}^n),$$

où l'on a utilisé les notations définies précédemment. Or on a

$$\begin{cases} \bar{u}_j^{n+1} - \bar{u}_j^n &= \Delta t \frac{\partial \bar{u}}{\partial t}(x_j, t^n) + \mathcal{O}(\Delta t^2), \\ \bar{u}_{j+m}^n - \bar{u}_j^n &= mh \frac{\partial \bar{u}}{\partial x}(x_j, t^n) + \mathcal{O}(h^2) \text{ avec } m = \pm 1. \end{cases}$$

D'où, si $\frac{\partial g}{\partial u}$ et $\frac{\partial g}{\partial v}$ désignent les 2 dérivées partielles de g :

$$\begin{cases} g(\bar{u}_j^n, \bar{u}_{j+1}^n) &= g(\bar{u}_j^n, \bar{u}_j^n) + \frac{\partial g}{\partial v}(\bar{u}_j^n, \bar{u}_j^n) h \frac{\partial \bar{u}}{\partial x}(x_j, t^n) + \mathcal{O}(h^2), \\ g(\bar{u}_{j-1}^n, \bar{u}_j^n) &= g(\bar{u}_j^n, \bar{u}_j^n) - \frac{\partial g}{\partial u}(\bar{u}_j^n, \bar{u}_j^n) h \frac{\partial \bar{u}}{\partial x}(x_j, t^n) + \mathcal{O}(h^2). \end{cases}$$

On en déduit enfin, pour $\alpha = \frac{\Delta t}{h}$ fixé :

$$\varepsilon_j^n = \frac{\partial \bar{u}}{\partial t}(x_j, t^n) + \left\{ \frac{\partial g}{\partial v}(\bar{u}_j^n, \bar{u}_j^n) + \frac{\partial g}{\partial u}(\bar{u}_j^n, \bar{u}_j^n) \right\} \frac{\partial \bar{u}}{\partial x}(x_j, t^n) + \mathcal{O}(h).$$

Le schéma est donc consistant (et au moins d'ordre 1) si et seulement si

$$\frac{\partial g}{\partial u}(u, u) + \frac{\partial g}{\partial v}(u, u) = f'(u) \quad \forall u \in \mathbb{R},$$

d'où l'égalité (4.6). ■

Insistons sur le fait qu'il existe **des schémas consistants et non conservatifs**.

4.2.3 Linéarisation du schéma - Stabilité L^2

La stabilité asymptotique d'un schéma non-linéaire se définit de la manière suivante :

DÉFINITION 4.2.2 (STABILITÉ ASYMPTOTIQUE)

Le schéma (4.2) est asymptotiquement stable pour la norme $\|\cdot\|$ si pour tout couple de données initiales (u^0, v^0) , le couple de solutions (u^n, v^n) vérifie

$$\forall n \in \mathbb{N}, \quad \|u^n - v^n\| \leq C \|u^0 - v^0\|$$

où la constante C est indépendante de $u^0, v^0, n, \Delta t$ et h .

Contrairement au cas linéaire, la stabilité L^2 d'un schéma non-linéaire ne peut plus s'étudier de manière simple à l'aide de la transformée de Fourier. Cependant, on peut obtenir un critère heuristique de stabilité par linéarisation.

Considérons un schéma sous la forme conservative (4.5) et soient u^n et v^n les solutions correspondant aux données initiales u^0 et v^0 .

$$\begin{cases} u_j^{n+1} = u_j^n - \alpha (g(u_j^n, u_{j+1}^n) - g(u_{j-1}^n, u_j^n)), \\ v_j^{n+1} = v_j^n - \alpha (g(v_j^n, v_{j+1}^n) - g(v_{j-1}^n, v_j^n)). \end{cases} \quad (4.7)$$

On pose alors

$$w_j^n = u_j^n - v_j^n.$$

Si w_j^n est supposé très petit devant u_j^n , on peut calculer la différence entre les deux égalités (4.7) et développer le résultat par rapport aux puissances de w . Si on ne conserve que les termes du premier ordre, on obtient :

$$\begin{aligned} w_j^{n+1} = w_j^n - \alpha \left\{ \left(\frac{\partial g}{\partial u}(u_j^n, u_{j+1}^n) - \frac{\partial g}{\partial v}(u_{j-1}^n, u_j^n) \right) w_j^n \right. \\ \left. + \frac{\partial g}{\partial v}(u_j^n, u_{j+1}^n) w_{j+1}^n - \frac{\partial g}{\partial u}(u_{j-1}^n, u_j^n) w_{j-1}^n \right\}. \end{aligned} \quad (4.8)$$

Relativement à w , il s'agit d'un schéma linéaire à coefficients variables. On dit qu'il s'agit du **schéma linéarisé associé au schéma (4.5)**.

L'étude de la stabilité du schéma (4.8) ne peut pas se faire comme au chapitre précédent par analyse de Fourier, puisque les coefficients du schéma sont ici variables. Pour pallier cette difficulté, il nous faut faire une nouvelle hypothèse : nous supposons que u_j^n varie lentement en fonction de j . On peut alors considérer que, au voisinage d'un point x_k , le schéma (4.8) « ressemble » beaucoup au schéma suivant :

$$w_j^{n+1} = w_j^n - \alpha \left\{ \left(\frac{\partial g}{\partial u} - \frac{\partial g}{\partial v} \right) w_j^n + \frac{\partial g}{\partial v} w_{j+1}^n - \frac{\partial g}{\partial u} w_{j-1}^n \right\} \quad (4.9)$$

où toutes les dérivées $\frac{\partial g}{\partial u}$ et $\frac{\partial g}{\partial v}$ sont calculées en (u_k^n, u_k^n) . Ce schéma étant à coefficients constants, sa stabilité s'analyse de manière classique.

DÉFINITION 4.2.3 (STABILITÉ LINÉAIRE)

On dit que le schéma (4.5) est linéairement stable L^2 si le schéma (4.9) est stable L^2 pour tout $k \in \mathbb{Z}$.

REMARQUE 4.2.2 (UNE STABILITÉ NON GARANTIE)

Le critère de stabilité linéaire que nous venons d'explicitier n'a qu'une valeur indicative. En effet, les approximations que nous avons faites ne sont pas justifiées au voisinage des chocs ni même dans les zones où la dérivée devient trop importante. En résumé, les schémas linéairement instables sont toujours à proscrire. En revanche, il est seulement plausible, mais nullement garanti, qu'un schéma linéairement stable ait quelques propriétés de stabilité.

REMARQUE 4.2.3 (UNE STABILITÉ LINÉAIRE QUI PEUT ÊTRE CONDITIONNELLE)

Certains schémas sont linéairement stables sous une condition portant sur le rapport des pas Δt et h . Tout comme dans le cas linéaire, cette condition CFL peut s'interpréter géométriquement. En effet, nous avons vu au cours précédent que, pour l'équation

$$\frac{\partial u}{\partial t} + \frac{\partial}{\partial x} f(u) = 0,$$

la connaissance de u à l'instant t^n dans l'intervalle $[x_{j-1}, x_{j+1}]$ détermine la solution dans l'ensemble du cône :

$$x_{j-1} + A(t - t^n) \leq x \leq x_{j+1} - A(t - t^n) \quad (4.10)$$

où $A = \sup \left\{ |c(\xi)|; |\xi| \leq \sup_{x \in \mathbb{R}} |u(x, t^n)| \right\}$. La condition (4.10) signifie que le point (x_j, t_{n+1}) doit se trouver dans ce cône.

4.2.4 Vers la convergence : le théorème de Lax-Wendroff

Nous allons maintenant démontrer un résultat important concernant les schémas sous forme conservative. Il exprime le fait que si un schéma sous forme conservative consistant avec l'équation (4.3) converge, c'est nécessairement vers une solution faible du problème.

On suppose que la donnée initiale u^0 du problème (4.1) a été discrétisée comme suit :

$$u_j^0 = \frac{1}{h} \int_{x_{j-\frac{1}{2}}}^{x_{j+\frac{1}{2}}} u^0(x) dx \quad \forall j \in \mathbb{Z},$$

où :

$$x_{j+\frac{1}{2}} = \left(j + \frac{1}{2}\right) h \quad \forall j \in \mathbb{Z}.$$

On construit alors, pour Δt et h donnés, une suite $u^n = (u_j^n)_{j \in \mathbb{Z}}$ par un schéma de la forme

$$u_j^{n+1} = u_j^n - \frac{\Delta t}{h} \left(g_{j+\frac{1}{2}}^n - g_{j-\frac{1}{2}}^n \right), \quad (4.11)$$

consistant avec l'équation (4.3), c'est-à-dire

$$g(u, u) = f(u), \quad \forall u \in \mathbb{R}.$$

Notons qu'on a omis la constante pour simplifier la démonstration. On note u_Δ la fonction définie par :

$$u_\Delta(x, t) = u_j^n \text{ pour } x_{j-\frac{1}{2}} \leq x < x_{j+\frac{1}{2}} \text{ et } t^n \leq t < t_{n+1}.$$

A chaque couple $(\Delta t, h)$ correspond une fonction u_Δ . On a alors le théorème suivant.

THÉORÈME 4.2.2 (LAX-WENDROFF)

Supposons qu'il existe une suite $(\Delta t_k, h_k)_{k \in \mathbb{N}}$, une suite de fonctions $(u_{\Delta k})_{k \in \mathbb{N}}$ associées, et une fonction u telles que, pour tout domaine K borné de $\mathbb{R} \times [0, +\infty[$:

$$(i) \quad u \in L^\infty(K),$$

$$(ii) \quad u_{\Delta k} \longrightarrow u \text{ dans } L^\infty(K) \text{ quand } k \longrightarrow +\infty.$$

Alors u est une solution faible du problème (4.1).

DÉMONSTRATION: Soit $\varphi \in \mathcal{D}(\mathbb{R}^2)$. On doit montrer que

$$\int_{x=-\infty}^{+\infty} \int_{t=0}^{+\infty} \left(u \frac{\partial \varphi}{\partial t} + f(u) \frac{\partial \varphi}{\partial x} \right) dx dt + \int_{x=-\infty}^{+\infty} u^0(x) \varphi(x, 0) dx = 0. \quad (4.12)$$

Posons $\varphi_j^n = \varphi(x_j, t^n)$, pour $j \in \mathbb{Z}$, $n \in \mathbb{N}$. En multipliant (4.11) par φ_j^n et en sommant, on obtient (on omet l'indice k pour alléger les notations) :

$$h \sum_{j,n} \left(u_j^{n+1} - u_j^n \right) \varphi_j^n + \Delta t \sum_{j,n} \left(g_{j+\frac{1}{2}}^n - g_{j-\frac{1}{2}}^n \right) \varphi_j^n = 0.$$

Par « sommation par parties », on obtient :

$$h \sum_{j,n} u_j^{n+1} \left(\varphi_j^n - \varphi_j^{n+1} \right) - h \sum_j u_j^0 \varphi_j^0 + \Delta t \sum g_{j+\frac{1}{2}}^n \left(\varphi_j^n - \varphi_{j+1}^n \right) = 0. \quad (4.13)$$

Si on note φ_Δ et g_Δ les fonctions définies par :

$$\begin{cases} \varphi_\Delta(x, t) = \varphi_j^n & \text{avec } x_{j-\frac{1}{2}} \leq x < x_{j+\frac{1}{2}} \quad \text{et } t^n \leq t < t_{n+1}, \\ g_\Delta(x, t) = g_{j+\frac{1}{2}}^n & \text{avec } x_j \leq x < x_{j+1} \quad \text{et } t^n \leq t < t_{n+1}, \end{cases}$$

l'égalité (4.13) s'écrit aussi :

$$\begin{aligned} & \int_{x=-\infty}^{+\infty} \int_{t=\Delta t}^{+\infty} u_\Delta(x, t) \frac{\varphi_\Delta(x, t - \Delta t) - \varphi_\Delta(x, t)}{\Delta t} dx dt - \int_{x=-\infty}^{+\infty} u_\Delta(x, 0) \varphi_\Delta(x, 0) dx \\ & + \int_{x=-\infty}^{+\infty} \int_{t=0}^{+\infty} g_\Delta(x, t) \frac{\varphi_\Delta(x - \frac{h}{2}, t) - \varphi_\Delta(x + \frac{h}{2}, t)}{h} dx dt = 0. \end{aligned}$$

On vérifie alors facilement que la première intégrale converge vers :

$$- \int_{x=-\infty}^{+\infty} \int_{t=0}^{+\infty} u(x, t) \frac{\partial \varphi}{\partial t}(x, t) dx dt$$

et la seconde vers :

$$\int_{x=-\infty}^{+\infty} u^0(x) \varphi(x, 0) dx.$$

On remarque ensuite que g_Δ vérifie par définition :

$$g_\Delta(x, t) = g \left(u_\Delta \left(x - \frac{h}{2}, t \right), u_\Delta \left(x + \frac{h}{2}, t \right) \right) \quad \forall x, \forall t.$$

Donc g_Δ converge uniformément sur tout compact vers la fonction $g(u(x, t), u(x, t))$ qui, d'après la consistance du schéma est identiquement égale à $f(u(x, t))$. La troisième intégrale converge donc vers :

$$- \int_{x=-\infty}^{+\infty} \int_{t=0}^{+\infty} f(u(x, t)) \frac{\partial \varphi}{\partial x}(x, t) dx dt,$$

d'où le résultat (4.12). ■

Ce théorème nous assure que les schémas sous forme conservative sont adaptés au calcul de solutions faibles. Notons qu'un schéma qui ne peut pas s'écrire sous forme conservative peut très bien converger vers une fonction qui n'est pas solution faible du problème.

4.2.5 Un exemple de schéma conservatif : le schéma de Murman Roe

Il est naturel de vouloir étendre au cas non-linéaire le schéma décentré. Mais il y a une difficulté puisque nous avons vu que le schéma décentré à gauche est stable sous une condition CFL si la vitesse c est positive. Si celle-ci est négative, il est impératif de décentrer à droite. Autrement dit, le schéma non-linéaire suivant :

$$u_j^{n+1} = u_j^n - \alpha(f(u_j^n) - f(u_{j-1}^n))$$

est linéairement instable dans toute zone où $c(u)$ est négatif.

Pour remédier à cette difficulté, on peut par exemple considérer le schéma suivant, appelé schéma de Murman-Roe

$$u_j^{n+1} = u_j^n - \alpha \{g(u_j^n, u_{j+1}^n) - g(u_{j-1}^n, u_j^n)\}$$

où

$$g(u, v) = \frac{1}{2}(f(u) + f(v) + |c(u, v)|(u - v)),$$

avec

$$c(u, v) = \begin{cases} \frac{f(u) - f(v)}{u - v} & \text{si } u \neq v \\ f'(u) & \text{si } u = v. \end{cases}$$

Ce schéma coïncide avec le schéma décentré lorsque f est monotone.

Le schéma de Murman-Roe possède les propriétés suivantes.

- **Conservatif.** On remarque que le schéma de Murman-Roe est sous forme conservative en prenant pour le flux numérique la fonction g définie précédemment.
- **Consistant.** Il est consistant car $g(u, u) = f(u)$ pour tout $u \in \mathbb{R}$.
- **Linéairement stable dans certains cas.** Il n'est pas différentiable aux points (u, v) pour lesquels $c(u, v)$ s'annule. On ne peut donc pas linéariser le schéma au voisinage de ces points. On peut seulement affirmer que, sous la condition CFL (4.23), le schéma de Murman-Roe est linéairement stable L^2 dans les zones où f est monotone.

D'après le théorème de Lax-Wendroff, le schéma de Murman Roe va tendre vers une solution faible du problème qui n'est pas nécessairement la solution entropique du problème.

En effet, supposons que l'on veuille résoudre numériquement à l'aide de ce schéma le problème (4.1) avec la donnée initiale :

$$u^0(x) = \begin{cases} u_g & \text{si } x \leq 0, \\ u_d & \text{si } x > 0, \end{cases}$$

où u_g et u_d sont tels que : $f(u_g) = f(u_d)$. Pour la donnée initiale discrète suivante

$$u_j^0 = \begin{cases} u_g & \text{si } j \leq 0, \\ u_d & \text{si } j \geq 1, \end{cases}$$

le schéma de Murman-Roe fournit une solution discrète stationnaire, i.e. :

$$u_j^n = u_j^0 \quad \forall n \in \mathbb{N}.$$

En effet, on a :

$$g(u_g, u_g) = g(u_d, u_d) = g(u_g, u_d) = g(u_d, u_g) = f(u_g) = f(u_d).$$

Ainsi, le schéma de Murman-Roe permet de calculer une solution du problème donnée par

$$u(x, t) = u^0(x) \quad \forall t \geq 0.$$

C'est une solution faible car la relation de Rankine-Hugoniot est satisfaite.

Si $u_g \geq u_d$, il s'agit bien de la solution entropique. En revanche, si $u_g < u_d$, on sait que la solution entropique correspond à une onde de détente et non à ce choc stationnaire. Le schéma de Murman-Roe ne permet pas de « sélectionner » la solution entropique.

On verra plus tard que ce schéma est **non entropique, non monotone**.

4.2.6 Schémas entropiques

Un schéma sous forme conservative convergent ne converge pas toujours vers la solution entropique du problème (voir précédemment le schéma de Murman-Roe).

Nous allons maintenant énoncer un critère (malheureusement difficile à vérifier en pratique) qui assure la convergence (si cette dernière a lieu) vers la solution entropique.

On rappelle tout d'abord que la solution entropique est caractérisée par le fait qu'elle satisfait la condition d'entropie

$$\frac{\partial}{\partial t} U(u) + \frac{\partial}{\partial x} F(u) \leq 0 \quad (4.14)$$

pour toute entropie (U, F) , c'est-à-dire pour tout couple de fonctions \mathcal{C}^1 , U étant strictement convexe et F étant définie par :

$$F'(v) = U'(v)f'(v) \quad \forall v \in \mathbb{R}.$$

Nous sommes en mesure de définir la notion de consistance d'un schéma avec une condition entropique.

DÉFINITION 4.2.4 (SCHÉMA CONSISTANT AVEC UNE CONDITION D'ENTROPIE)

Le schéma (4.5) est dit consistant avec la condition d'entropie (4.14) si et seulement si il existe une fonction $G(u, v)$ telle que :

(i) G est consistante avec le flux d'entropie F ,

$$G(u, u) = F(u), \quad \forall u \in \mathbb{R};$$

(ii) Pour toute solution (u_j^n) de (4.5), on a

$$U_j^{n+1} \leq U_j^n - \alpha \left(G_{j+\frac{1}{2}}^n - G_{j-\frac{1}{2}}^n \right)$$

où l'on a posé :

$$\begin{cases} U_j^n &= U(u_j^n), \\ G_{j+\frac{1}{2}}^n &= G(u_j^n, u_{j+1}^n), \end{cases} \quad \forall j \in \mathbb{Z}, \forall n \in \mathbb{N}.$$

La fonction G est appelée le flux d'entropie numérique.

On peut alors établir, par une démarche tout à fait analogue à celle que nous avons suivie pour démontrer le théorème de Lax-Wendroff, le théorème suivant.

THÉORÈME 4.2.3

Sous les hypothèses du théorème 4.2.2, si de plus le schéma considéré est consistant avec toute condition d'entropie, alors la limite u est l'unique solution entropique du problème (4.1).

Ceci conduit finalement à définir les schémas entropiques.

DÉFINITION 4.2.5 (SCHÉMA ENTROPIQUE)

On dit que le schéma (4.5) est entropique s'il est consistant avec toute condition d'entropie.

Nous allons étudier une classe particulière de schémas entropiques : les schémas monotones.

4.2.7 Schémas monotones

Condition de monotonie

On a vu dans le chapitre précédent que l'opérateur qui à u^0 associe la solution entropique du problème de Cauchy (4.1) est monotone. Autrement dit, si u et v sont les solutions pour les données initiales u^0 et v^0 , on a :

$$u^0(x) \geq v^0(x) \text{ p.p. } x \in \mathbb{R} \Rightarrow u(x, t) \geq v(x, t) \text{ p.p. } x \in \mathbb{R}, t > 0.$$

Il est naturel de s'intéresser aux schémas possédant la même propriété.

DÉFINITION 4.2.6 (SCHÉMA MONOTONE)

Un schéma est dit monotone si

$$\{u_j^n \geq v_j^n, \forall j \in \mathbb{Z}\} \Rightarrow \{u_j^{n+1} \geq v_j^{n+1}, \forall j \in \mathbb{Z}\}.$$

Nous disposons d'un critère simple pour voir si un schéma est monotone ou pas.

LEMME 4.2.4 (CRITÈRE DE MONOTONIE)

Le schéma

$$u_j^{n+1} = H(u_{j-1}^n, u_j^n, u_{j+1}^n) \quad (4.15)$$

est monotone si et seulement si H est une fonction croissante de chacun de ses arguments.

DÉMONSTRATION: La démonstration de ce lemme est immédiate. ■

En pratique, la fonction H dépend des pas de discrétisation Δt et h et la monotonie n'est réalisée que sous une condition sur $\alpha = \frac{\Delta t}{h}$ (condition de type CFL).

Propriétés des schémas monotones

Nous pouvons déjà énoncer quelques propriétés immédiates des schémas monotones.

LEMME 4.2.5 (PROPRIÉTÉS DES SCHÉMAS MONOTONES)

Soit un schéma monotone de la forme (4.15) et tel que $H(u, u, u) = u$ pour tout $u \in \mathbb{R}$. Alors :

(i) il préserve la monotonie i.e.

$$\begin{cases} \forall j, & u_j^n \leq u_{j+1}^n & \Rightarrow & \forall j, & u_j^{n+1} \leq u_{j+1}^{n+1}, \\ \forall j, & u_j^n \geq u_{j+1}^n, \forall j & \Rightarrow & \forall j, & u_j^{n+1} \geq u_{j+1}^{n+1}, \end{cases}$$

(ii) il vérifie le principe du maximum suivant :

$$\forall k \in \mathbb{Z} \quad \inf_{j \in \mathbb{Z}} u_j^n \leq u_k^{n+1} \leq \sup_{j \in \mathbb{Z}} u_j^n,$$

(iii) il est borné en norme l^∞ :

$$\|u^{n+1}\|_{\ell^\infty} \leq \|u^n\|_{\ell^\infty}$$

$$\text{où } \|u^n\|_{\ell^\infty} = \sup_{j \in \mathbb{Z}} |u_j^n|.$$

DÉMONSTRATION:

(i) il suffit de poser $v_j^n = u_{j+1}^n$ (puis $v_j^n = u_{j-1}^n$) et d'appliquer la définition.

(ii) De même, il suffit de poser : $v_j^n = \inf_{k \in \mathbb{Z}} u_k^n$ (puis $v_j^n = \sup_{k \in \mathbb{Z}} u_k^n$) $\forall j$ et d'appliquer la définition en utilisant l'identité $H(u, u, u) = u$.

Enfin, (iii) est une conséquence directe de (ii). ■

En particulier, si la condition initiale est monotone, la solution approchée calculée à l'aide d'un schéma monotone ne comportera jamais d'oscillations.

Une propriété essentielle des schémas monotones est donnée par le théorème suivant que nous admettrons :

THÉORÈME 4.2.6 (UN RÉSULTAT DE CONVERGENCE)

Soit un schéma sous forme **conservative consistant avec (4.3) et monotone**. Alors, de toute suite h_k tendant vers 0, on peut extraire une sous-suite telle que la suite de fonctions associées $(u_{\Delta k})$ converge presque partout vers une solution faible u du problème (4.1) (les notations sont celles du théorème 4.2.2).

Autrement dit, les schémas monotones sont, en un certain sens, convergents. De plus, la limite est bien la solution entropique du problème.

On a en effet le théorème suivant.

THÉORÈME 4.2.7 (LIEN SCHÉMA ENTROPIQUE - SCHÉMA MONOTONE)

Un schéma sous forme conservative consistant avec (4.3) et monotone est entropique.

DÉMONSTRATION: On considère un schéma de la forme suivante :

$$u_j^{n+1} = H(u_{j-1}^n, u_j^n, u_{j+1}^n)$$

où $H(u, v, w) = v - \alpha\{g(v, w) - g(u, v)\}$. On doit montrer qu'il est consistant avec toute condition d'entropie, mais on admettra qu'il suffit de considérer les entropies de Kruzkov qui sont de la forme :

$$\begin{cases} U(u) &= |u - k|, \\ F(u) &= \text{signe}(u - k)(f(u) - f(k)), \end{cases} \quad k \in \mathbb{R}.$$

D'après la définition 4.2.4, il nous faut trouver une fonction $G(u, v)$ telle que l'on ait :

$$G(u, u) = \text{signe}(u - k)(f(u) - f(k)) \quad \forall u \in \mathbb{R} \quad (4.16)$$

et

$$|u_j^{n+1} - k| \leq |u_j^n - k| - \alpha(G(u_j^n, u_{j+1}^n) - G(u_{j-1}^n, u_j^n)) \quad (4.17)$$

pour toute solution (u_j^n) du schéma. Posons alors :

$$G(u, v) = g(u \vee k, v \vee k) - g(u \wedge k, v \wedge k)$$

où l'on a noté :

$$a \vee b = \max(a, b) \text{ et } a \wedge b = \min(a, b).$$

On a alors, puisque le schéma est consistant avec (4.3) :

$$\begin{aligned} G(u, u) &= g(u \vee k, u \vee k) - g(u \wedge k, u \wedge k) \\ &= f(u \vee k) - f(u \wedge k) \\ &= \text{signe}(u - k)(f(u) - f(k)). \end{aligned}$$

L'identité (4.16) est donc satisfaite.

Posons alors : $v_j^n = u_j^n \vee k$. Comme le schéma est monotone :

$$\begin{cases} v_j^n \geq u_j^n \quad \forall j & \Rightarrow H(v_{j-1}^n, v_j^n, v_{j+1}^n) \geq u_j^{n+1} \\ v_j^n \geq k \quad \forall j & \Rightarrow H(v_{j-1}^n, v_j^n, v_{j+1}^n) \geq H(k, k, k) = k \end{cases}$$

d'où finalement :

$$H(u_{j-1}^n \vee k, u_j^n \vee k, u_{j+1}^n \vee k) \geq u_j^{n+1} \vee k. \quad (4.18)$$

On démontrerait de même l'inégalité suivante :

$$H(u_{j-1}^n \wedge k, u_j^n \wedge k, u_{j+1}^n \wedge k) \leq u_j^{n+1} \wedge k. \quad (4.19)$$

Mais comme $|u - k| = u \vee k - u \wedge k$, on déduit de (4.18) et (4.19) que :

$$|u_j^{n+1} - k| \leq H(u_{j-1}^n \vee k, u_j^n \vee k, u_{j+1}^n \vee k) - H(u_{j-1}^n \wedge k, u_j^n \wedge k, u_{j+1}^n \wedge k),$$

et on vérifie aisément que cette dernière inégalité coïncide avec (4.17). ■

Malheureusement, la portée des schémas monotones est limitée par leur ordre :

THÉORÈME 4.2.8 (ORDRE DES SCHÉMAS MONOTONES)

Un schéma sous forme conservative consistant avec (4.3) et monotone est d'ordre 1 (exactement).

DÉMONSTRATION: On considère comme précédemment un schéma de la forme :

$$u_j^{n+1} = H(u_{j-1}^n, u_j^n, u_{j+1}^n),$$

où $H(u, v, w) = v - \alpha\{g(v, w) - g(u, v)\}$.

1) Soit \bar{u} une solution régulière de (4.3). En poursuivant le calcul amorcé lors de la démonstration du lemme 4.2.1, on vérifie aisément que l'erreur de troncature s'écrit

$$\varepsilon_j^n = \frac{\Delta t}{2} \left[\frac{\partial^2 \bar{u}}{\partial t^2} \right]_j^n + \frac{h^2}{2\Delta t} \alpha \left[\frac{\partial}{\partial x} \left(\left(\frac{\partial g}{\partial v}(\bar{u}, \bar{u}) - \frac{\partial g}{\partial u}(\bar{u}, \bar{u}) \right) \frac{\partial \bar{u}}{\partial x} \right) \right]_j^n + \mathcal{O}(h^2).$$

En dérivant (4.3) nous avons

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

Ainsi, l'erreur de troncature s'écrit finalement :

$$\varepsilon_j^n = \frac{\Delta t}{2} \left[\frac{\partial}{\partial x} \left(\beta(\bar{u}, \alpha) \frac{\partial \bar{u}}{\partial x} \right) \right]_j^n + \mathcal{O}(h^2)$$

où

$$\beta(\bar{u}, \alpha) = c(\bar{u})^2 - \frac{1}{\alpha} \left(\frac{\partial g}{\partial u}(\bar{u}, \bar{u}) - \frac{\partial g}{\partial v}(\bar{u}, \bar{u}) \right).$$

2) Le schéma étant monotone, on a :

$$\begin{aligned} \forall u \in \mathbb{R}, \quad \frac{\partial H}{\partial u}(u, u) &= \alpha \frac{\partial g}{\partial u}(u, u) \geq 0 \\ \forall u \in \mathbb{R}, \quad \frac{\partial H}{\partial v}(u, u) &= 1 - \alpha \left(\frac{\partial g}{\partial u}(u, u) - \frac{\partial g}{\partial v}(u, u) \right) \geq 0 \\ \forall u \in \mathbb{R}, \quad \frac{\partial H}{\partial w}(u, u) &= -\alpha \frac{\partial g}{\partial v}(u, u) \geq 0. \end{aligned}$$

On a donc finalement

$$\forall u \in \mathbb{R}, \quad \begin{cases} \frac{\partial g}{\partial u}(u, u) \geq 0 \\ \frac{\partial g}{\partial v}(u, u) \leq 0 \\ \frac{\partial g}{\partial u}(u, u) - \frac{\partial g}{\partial v}(u, u) \leq \frac{1}{\alpha} \end{cases} \quad (4.20)$$

3) Le schéma étant consistant, on a :

$$\forall u \in \mathbb{R}, \quad g(u, u) = f(u) + cste,$$

d'où

$$\forall u \in \mathbb{R}, \quad \frac{\partial g}{\partial u}(u, u) + \frac{\partial g}{\partial v}(u, u) = c(u). \quad (4.21)$$

De (4.20) et (4.21), on déduit finalement :

$$c(u)^2 \leq \left(\frac{\partial g}{\partial u}(u, u) - \frac{\partial g}{\partial v}(u, u) \right)^2 \leq \frac{1}{\alpha} \left(\frac{\partial g}{\partial u}(u, u) - \frac{\partial g}{\partial v}(u, u) \right).$$

Donc $\beta(\bar{u}, \alpha)$ est toujours négatif ou nul.

Pour que le schéma soit d'ordre supérieur ou égal à 2, il faut que $\beta(u, \alpha)$ soit nul pour tout u .

D'après ce qui précède, il faut pour cela que l'on ait :

$$\forall u \in \mathbb{R}, \quad \left(\frac{\partial g}{\partial u}(u, u) - \frac{\partial g}{\partial v}(u, u) \right)^2 = \left(\frac{\partial g}{\partial u}(u, u) + \frac{\partial g}{\partial v}(u, u) \right)^2,$$

et

$$\frac{\partial g}{\partial u}(u, u) - \frac{\partial g}{\partial v}(u, u) = \frac{1}{\alpha}.$$

Ceci n'est vrai que si :

$$\frac{\partial g}{\partial u}(u, u) = \frac{1}{\alpha} \text{ et } \frac{\partial g}{\partial v}(u, u) = 0,$$

où

$$\frac{\partial g}{\partial v}(u, u) = -\frac{1}{\alpha} \text{ et } \frac{\partial g}{\partial u}(u, u) = 0.$$

Plaçons nous par exemple dans le premier cas. D'après (4.21), on a alors :

$$\forall u \in \mathbb{R}, \quad a(u) = \frac{1}{\alpha}.$$

Il s'agit donc du cas linéaire de l'équation de transport avec la vitesse $\frac{1}{\alpha}$. Le schéma s'écrit :

$$u_j^{n+1} = u_{j-1}^n,$$

et il est exact car il coïncide avec la méthode des caractéristiques.

Ainsi nous avons démontré qu'un schéma monotone est d'ordre 1 exactement, sauf si $f(u) =$

cu et $c \frac{\Delta t}{h} = 1$ auquel cas il est d'ordre ∞ .

■

4.2.8 Deux exemples de schémas monotones

Le schéma de Lax-Friedrichs

Le schéma de Lax dans le cas linéaire s'étend de manière naturelle au cas non-linéaire sous la forme :

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

Le schéma de Lax-Friedrichs possède les propriétés suivantes.

- **Conservatif.** On vérifie aisément qu'il peut se mettre sous la forme conservative avec pour flux numérique :

$$g(u, v) = \frac{1}{2} \left\{ f(u) + f(v) + \frac{1}{\alpha}(u - v) \right\}.$$

- **Consistant.** Le schéma de Lax-Friedrichs est consistant d'ordre 1 au moins. En effet,

$$g(u, u) = \frac{1}{2} \left\{ f(u) + f(u) - \frac{1}{\alpha}(u - u) \right\} = f(u).$$

On verra ensuite qu'il est monotone donc qu'il est exactement d'ordre 1.

- **Linéairement stable.** Pour le schéma de Lax-Friedrichs, (4.8) s'écrit :

$$w_j^{n+1} = \frac{1}{2} (w_{j-1}^n + w_{j+1}^n) - \frac{\alpha}{2} c(u_k^n) (w_{j+1}^n - w_{j-1}^n)$$

où c désigne la dérivée de f . C'est exactement le schéma de Lax-Friedrichs linéaire avec une vitesse égale à $c(u_k^n)$. On sait que ce schéma est stable L^2 sous la condition CFL :

$$\frac{|c(u_k^n)| \Delta t}{h} \leq 1. \quad (4.22)$$

Le schéma non-linéaire de Lax-Friedrichs est donc linéairement stable L^2 si (4.22) est vérifiée pour tout k . Comme cette condition dépend de n , le pas de temps doit être ajusté à chaque itération. Autrement dit, on doit avoir :

$$\sup_{k \in \mathbb{Z}} |c(u_k^n)| \frac{\Delta t^{(n)}}{h} \leq 1 \quad \forall n \in \mathbb{N}. \quad (4.23)$$

- **Monotone.** On a

$$H(u, v, w) = \frac{1}{2} (u + w) - \frac{\alpha}{2} (f(w) - f(u)).$$

D'où :

$$\begin{aligned}\frac{\partial H}{\partial u}(u, v, w) &= \frac{1}{2} + \frac{\alpha}{2}c(u), \\ \frac{\partial H}{\partial v}(u, v, w) &= 0, \\ \frac{\partial H}{\partial w}(u, v, w) &= \frac{1}{2} - \frac{\alpha}{2}c(w).\end{aligned}$$

Par conséquent, le schéma de Lax-Friedrichs est monotone sous la condition CFL (4.23).

Le schéma d'Engquist-Osher

Il est défini comme suit :

$$u_j^{n+1} = u_j^n - \frac{\alpha}{2} \left\{ f(u_{j+1}^n) - f(u_{j-1}^n) + \int_{u_{j-1}^n}^{u_j^n} |c(\xi)| d\xi - \int_{u_j^n}^{u_{j+1}^n} |c(\xi)| d\xi \right\}$$

où l'on a noté a la dérivée de f .

On retrouve le schéma décentré à gauche dans les zones où f est croissante et à droite lorsque f est décroissante.

Le schéma d'Engquist-Osher possède les propriétés suivantes.

- **Conservatif.** Il peut s'écrire sous forme conservative à condition de poser :

$$g(u, v) = \frac{1}{2} \left\{ f(u) + f(v) - \int_u^v |c(\xi)| d\xi \right\}.$$

- **Consistant.** On vérifie aisément que $f(u, u) = g(u)$. Le schéma d'Engquist-Osher est donc consistant et d'ordre 1 au moins. On va voir qu'il est monotone donc qu'il est exactement d'ordre 1.
- **Linéairement stable.** Avec une démarche similaire à celle effectuée pour le schéma de Lax-Friedrichs, on montre qu'il est linéairement stable L^2 sous la condition CFL (4.23).
- **Monotone.** On a

$$H(u, v, w) = v - \frac{\alpha}{2} \left\{ f(w) - f(u) + \int_u^v |c(\xi)| d\xi - \int_v^w |c(\xi)| d\xi \right\}$$

D'où :

$$\begin{aligned}\frac{\partial H}{\partial u}(u, v, w) &= \frac{\alpha}{2}c(u) + \frac{\alpha}{2}|c(u)| \\ \frac{\partial H}{\partial v}(u, v, w) &= 1 - \alpha|c(v)| \\ \frac{\partial H}{\partial w}(u, v, w) &= -\frac{\alpha}{2}c(w) + \frac{\alpha}{2}|c(w)|.\end{aligned}$$

Par conséquent, le schéma d'Engquist-Osher est monotone sous la condition CFL (4.23).

4.3 Un exemple fondamental - Le schéma de Godounov

Dans son principe, le schéma consiste à générer une suite u_h^n de fonctions constantes par morceaux

$$u_h^n(x) = u_j^n, \quad x \in \left]x_j - \frac{h}{2}, x_j + \frac{h}{2}\right[$$

Le passage de u_h^n à u_h^{n+1} repose sur deux outils

- la résolution de problèmes de Riemann locaux
- une projection orthogonale locale

La construction de u_h^{n+1} à partir de u_h^n se fait ainsi en deux étapes :

- **Première étape** : On note $u_h^n(x)$ la fonction définie par

$$u_h^n(x) = u_j^n \text{ si } x_{j-\frac{1}{2}} \leq x < x_{j+\frac{1}{2}}, \quad (4.24)$$

où $x_{j+\frac{1}{2}} = \left(j + \frac{1}{2}\right)h$, $\forall j \in \mathbb{Z}$. On calcule alors \tilde{u}_h^{n+1} solution entropique du problème suivant :

$$\begin{cases} \frac{\partial \tilde{u}}{\partial t} + \frac{\partial}{\partial x} f(\tilde{u}) = 0, & x \in \mathbb{R}, t \in [t^n, t^{n+1}] \\ \tilde{u}(x, t^n) = u_h^n(x), & x \in \mathbb{R} \end{cases} \quad (4.25)$$

Le lemme suivant montre que ce problème peut être résolu de façon exacte

LEMME 4.3.1*Sous la condition*

$$\sup_{k \in \mathbb{Z}} |c(u_k^n)| \frac{\Delta t^{(n)}}{\Delta x} \leq \frac{1}{2}, \quad (4.26)$$

la solution du problème (4.25) est donnée par :

$$\tilde{u}_h^{n+1}(x, t) = w_R \left(\frac{x - x_{j+\frac{1}{2}}}{t - t^n}; u_j^n, u_{j+1}^n \right) \text{ si } x_j < x < x_{j+1}, \quad (4.27)$$

où $w_R \left(\frac{x}{t}; u_g, u_d \right)$ est la solution du problème de Riemann pour les états (u_g, u_d) .

DÉMONSTRATION: Le problème (4.25) consiste en une succession de problèmes de Riemann de la forme

$$(P_j) \begin{cases} \frac{\partial v}{\partial t} + \frac{\partial}{\partial x} f(v) = 0, & x \in [x_j, x_{j+1}], t \in [t^n, t^{n+1}], \\ v(x, t^n) = \begin{cases} u_j^n & \text{si } x < x_{j+\frac{1}{2}}, \\ u_{j+1}^n & \text{si } x > x_{j+\frac{1}{2}}. \end{cases} \end{cases}$$

La solution de ce problème pour $x \in \mathbb{R}$ est donnée par :

$$v(x, t) = w_R \left(\frac{x - x_{j+\frac{1}{2}}}{t - t^n}; u_j^n, u_{j+1}^n \right). \quad (4.28)$$

Il s'agit d'une onde de détente si $u_j^n \leq u_{j+1}^n$ et d'un choc dans le cas contraire. La « juxtaposition » de solutions de la forme (4.28) définit bien la solution de (4.25) si les différents problèmes n'ont pas « interagi » avant l'instant t^{n+1} . Autrement dit, il faut que la solution v de chacun des problèmes (P_j) vérifie :

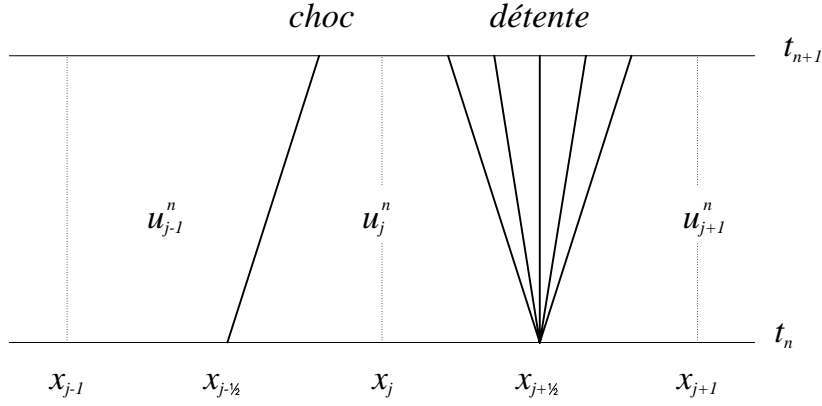
$$\forall t \in [t^n, t^{n+1}[, \quad \begin{cases} v(x_j, t) & = u_j^n, \\ v(x_{j+1}, t) & = u_{j+1}^{n+1}, \end{cases}$$

Ceci s'écrit encore :

$$\forall \tau \in [0, \Delta t[, \quad \begin{cases} w_R \left(-\frac{h}{2\tau}; u_j^n, u_j^{n+1} \right) = u_j^n, \\ w_R \left(\frac{h}{2\tau}; u_j^n, u_j^{n+1} \right) = u_j^{n+1} \end{cases}$$

et ceci est automatiquement vérifié sous la condition :

$$\frac{h}{2\Delta t} \geq \max(|c(u_j^n)|, |c(u_{j+1}^n)|).$$



Autrement dit, la vitesse de propagation de la solution du problème (P_j) étant :

$$V_j = \max \left(|c(u_j^n)|, |c(u_j^{n+1})| \right),$$

l'onde de choc ou l'onde de détente n'auront pas atteint les bornes de l'intervalle x_j et x_{j+1} avant l'instant :

$$T_j = t^n + V_j \frac{h}{2}.$$

Il suffit donc de choisir t^{n+1} tel que : $t^{n+1} \leq T_j, \forall j \in \mathbb{Z}$. ■

- Deuxième étape : On définit u_h^{n+1} comme projection de \tilde{u}_h^{n+1} sur les fonctions constantes par morceaux :

$$u_j^{n+1} = P_h \tilde{u}_h^{n+1}. \quad (4.29)$$

où la projection P_h est définie par

$$\forall v \in L^2(\mathbb{R}), \quad P_h v(x) = \frac{1}{h} \int_{x_{j-1/2}}^{x_{j+1/2}} v(x) dx, \quad \forall x \in [x_{j-1/2}, x_{j+1/2}]$$

Défini ainsi, le schéma de Godunov présente les propriétés suivantes :

- **Il est conservatif.** Nous allons maintenant vérifier que le schéma de Godunov décrit précédemment peut se mettre sous forme conservative.

LEMME 4.3.2

- (i) L'application $\xi \rightarrow f(w_R(\xi; u, v))$ est continue en 0, quels que soient u et v .
- (ii) Le schéma de Godunov peut s'écrire sous forme conservative et son flux numérique est donné par :

$$g(u, v) = f(w_R(0; u, v)). \quad (4.30)$$

DÉMONSTRATION:

(i) On sait que l'application $\xi \rightarrow w_R(\xi; u, v)$ est continue partout sauf en $\xi = \frac{f(u) - f(v)}{u - v}$ si $u > v$. Elle est donc continue en 0 sauf si : $u > v$ et $f(u) = f(v)$. Or, on a dans ce cas : $w_R(0^-; u, v) = u$ et $w_R(0^+; u, v) = v$. D'où : $f(w_R(0^-; u, v)) = f(u) = f(w) = f(w_R(0^+; u, v))$.

(ii) Intégrons l'équation (4.25) sur la maille

$$\left[x_{j-\frac{1}{2}}, x_{j+\frac{1}{2}} \right] \times [t^n, t^{n+1}].$$

On a :

$$0 = \int_{x_{j-\frac{1}{2}}}^{x_{j+\frac{1}{2}}} \int_{t^n}^{t^{n+1}} \left(\frac{\partial v}{\partial t} + \frac{\partial}{\partial x} f(v) \right) dx dt.$$

A priori, cette écriture n'a de sens que si la maille considérée n'est pas traversée par une ligne de choc. Dans le cas contraire, il faudrait intégrer de part et d'autre du choc. Cependant, comme v satisfait la relation de Rankine-Hugoniot, on vérifie aisément que l'intégrale sur la ligne de choc issue de l'intégration par parties est nulle. Tout se passe donc comme si v était régulière !

On a alors

$$0 = \int_{x_{j-\frac{1}{2}}}^{x_{j+\frac{1}{2}}} (v(x, t^{n+1}) - v(x, t^n)) dx + \int_{t^n}^{t^{n+1}} \left(f(v(x_{j+\frac{1}{2}}, t)) - f(v(x_{j-\frac{1}{2}}, t)) \right) dt.$$

D'où, en vertu de (4.29), (4.27) et du point (i) :

$$0 = h \left(u_j^{n+1} - u_j^n \right) + \Delta t \left(f(w_R(0; u_j^n, u_{j+1}^n)) - f(w_R(0; u_{j-1}^n, u_j^n)) \right).$$

C'est ce qu'il fallait démontrer. ■

Remarquons que le schéma de Godounov coïncide avec le schéma décentré si f est monotone. En effet, si par exemple

$$c(u_j^n) \geq 0, \forall j \in \mathbb{Z},$$

alors

$$w_R(0; u_j^n, u_{j+1}^n) = u_j^n, \forall j \in \mathbb{Z}$$

et le schéma de Godounov s'écrit

$$u_j^{n+1} = u_j^n - \alpha \left(f(u_j^n) - f(u_{j-1}^n) \right).$$

- **Il est consistant d'ordre 1.** En utilisant (4.30) on remarque que $g(u, u) = f(u)$ et donc que le schéma de Godounov est consistant d'ordre 1 au moins. Comme il est monotone (cf. ci-après), il est d'ordre 1 exactement.
- **Il est monotone et donc entropique.** Le schéma de Godounov est monotone. En effet, la première étape qui consiste en une résolution exacte d'un problème de Cauchy est monotone (voir cours précédent). La seconde étape l'est trivialement.