



**HAL**  
open science

# Schémas numériques pour la résolution de l'équation des ondes

Eliane Bécache

► **To cite this version:**

Eliane Bécache. Schémas numériques pour la résolution de l'équation des ondes. Master. France. 2009. hal-04208566

**HAL Id: hal-04208566**

**<https://hal.science/hal-04208566v1>**

Submitted on 15 Sep 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# **Schémas numériques pour la résolution de l'équation des ondes**

MASTER Modélisation et simulation

**Eliane Bécache**

ENSTA - Janvier 2009



# 1 Introduction et généralités sur les ondes

L'objectif de ce cours est d'appréhender les problèmes de propagation d'ondes, de les étudier sur le plan mathématique et de proposer et d'analyser des méthodes numériques pour les résoudre. Cette partie du cours est plus spécifiquement consacrée aux méthodes numériques et nous renvoyons au cours de P. Joly pour ce qui concerne l'étude mathématique des équations continues.

Les phénomènes de propagation d'ondes se rencontrent dans de nombreuses applications. On rencontre essentiellement trois types d'ondes: *les ondes acoustiques*, c'est à dire les ondes qui se propagent dans un fluide (eau ou air par exemple) ; *les ondes élastiques*, c'est à dire les ondes qui se propagent dans un solide ; *les ondes électromagnétiques*, par exemple la lumière. Nous étudierons principalement l'équation des ondes acoustiques qui est le modèle le plus simple (modèle scalaire) mais qui est déjà très riche car il permet d'aborder les principales notions communes à tous ces modèles et d'en comprendre les propriétés essentielles. Nous verrons que ces différents modèles conduisent à un même type d'équations: les équations hyperboliques d'ordre 2, c'est à dire les équations de la forme:

$$\frac{d^2u}{dt^2} + Au = f$$

où  $A$  est un opérateur différentiel en espace d'ordre 2.

Nous reviendrons dans cette introduction générale sur différentes applications pour lesquelles on comprend bien que l'outil de modélisation numérique est essentiel : il est bien sûr dans la plupart des cas, en particulier des applications réelles, impossible de résoudre ces équations de façon exacte explicitement. On doit donc avoir recours au calcul numérique d'une solution approchée. Ce qui va nous intéresser ici, ce sont les différentes méthodes pour approcher ces équations et les difficultés qu'on rencontre pour mettre en oeuvre numériquement ces méthodes et pour obtenir une "bonne solution approchée".

Les questions essentielles lorsqu'on approche  $u$ , la solution exacte, par  $u_h$  une solution approchée, sont:

- **la convergence:** c'est à dire savoir si  $u_h$  tend vers  $u$  et en quel sens ?
- **la précision:** si  $\|u - u_h\|$  tend vers 0, à quelle vitesse ?
- **la stabilité:** si on perturbe les données, comment est perturbée la solution ?

En fait, la question de la convergence est une question difficile et nous nous intéresserons dans un premier temps à la question de savoir si nos méthodes permettent de bien approcher un cas particulier d'ondes qui jouent un rôle essentiel, *les ondes planes*, c'est à dire les solutions de la forme  $u(x, t) = e^{i(\omega t - kx)}$ . Cette étude est un outil classique de l'analyse des schémas pour les équations d'ondes et est appelée *analyse de dispersion des ondes planes*.

Nous nous restreindrons dans ce cours à l'étude de schémas pour l'équation des ondes acoustiques essentiellement 1D mais nous illustrerons de temps en temps par des exemples 2D.

Avant d'aborder les schémas, nous indiquons brièvement quelques propriétés des différents modèles.

## 1.1 Les ondes acoustiques

Si  $\Omega$  désigne un ouvert de  $\mathbb{R}^3$  (borné ou non) rempli de fluide, la propagation d'ondes acoustiques dans ce fluide dépend de  $\rho(x)$  la densité de fluide au point  $x$ ,  $c(x)$  la vitesse locale des ondes acoustiques.

L'équation qui régit les variations de la pression  $p(x, t)$  du fluide en un point  $x$  et à l'instant  $t$  est l'équation des ondes :

$$(1) \quad \frac{1}{\mu(x)} \frac{\partial^2 p}{\partial t^2} - \operatorname{div} \left( \frac{1}{\rho(x)} \nabla p \right) = g, \quad x \in \Omega, t > 0$$

où  $\mu(x) = \rho(x)c^2(x)$ . En milieu homogène, cette équation devient simplement :

$$(2) \quad \frac{\partial^2 p}{\partial t^2} - c^2 \Delta p = g, \quad x \in \Omega, t > 0.$$

La fonction  $g(x, t)$  est la source.

**Remarque 1** Cette équation représente aussi les petites vibrations transversales d'une membrane homogène tendue. Elle s'appelle aussi l'équation des membranes vibrantes et la quantité que l'on calcule dans ce cas est le déplacement normal.

On doit adjoindre à cette équation, pour qu'elle soit bien posée :

- des conditions initiales (2 car l'équation est d'ordre 2):

$$p(x, 0) = p_0(x), \quad \text{et} \quad \frac{\partial p}{\partial t}(x, 0) = p_1(x)$$

- des conditions aux limites posées sur le bord du domaine (par exemple homogènes):

- Condition de Dirichlet:  $p = 0$  sur  $\partial\Omega$ .
- Condition de Neumann:  $\frac{\partial p}{\partial n} = 0$  sur  $\partial\Omega$ .
- Condition d'impédance:  $\frac{\partial p}{\partial n} + \frac{\partial p}{\partial t} = 0$  sur  $\partial\Omega$ .

En acoustique, l'équation (1) décrit la propagation d'une onde de pression qui se propage à une vitesse  $c$ . En 2D, on peut penser à la propagation d'une onde dans l'eau lorsqu'on jette un caillou ... L'onde due à l'excitation ponctuelle (caillou) se propage suivant des cercles et s'atténue au cours du temps mais ne disparaît pas. Un exemple d'onde acoustique en 3D, c'est bien sûr le son qui se propage dans l'air: cette fois-ci l'onde passe et disparaît (heureusement!). De façon plus générale, il y a une différence de comportement de l'onde entre les dimensions paires ou impaires: en dimension paire, la solution fondamentale (c'est à dire l'onde créée par une source ponctuelle en espace et en temps) "remplit" le cône de lumière alors qu'en dimension impaire elle a pour support le bord du cône.

**Exemple important d'ondes: les ondes planes.** Les ondes planes sont des solutions particulières de l'équation des ondes, qui jouent un rôle important car on peut montrer que toute solution peut s'écrire comme superposition d'ondes planes (Fourier). En 1D, par exemple, l'équation des ondes dans tout l'espace, sans terme source, s'écrit:

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

Les ondes planes sont alors de la forme:

$$u(x, t) = e^{i(\omega t - kx)}$$

et sont solutions sous la condition que

$$\omega^2 = c^2 k^2$$

appelée *relation de dispersion*.

**Applications.** Les ondes acoustiques interviennent par exemple en acoustique musicale, en acoustique de salle, en acoustique sous-marine, ou pour étudier les phénomènes de houle.

## 1.2 Les ondes élastiques

Les équations du mouvement sont:

$$(3) \quad \rho \frac{\partial^2 \vec{u}}{\partial t^2} - \text{div}(\sigma) = \rho \vec{g}$$

où  $\vec{u}(x, t)$  est un vecteur représentant le déplacement d'une particule du point  $x$  à l'instant  $t$ ,  $\sigma$  est le tenseur des contraintes, relié au tenseur des déformations  $\varepsilon(\vec{u})$  par la loi de comportement du matériau. Ici on considère la loi linéaire de Hooke:

$$(4) \quad \sigma = C\varepsilon(\vec{u}) \iff \sigma_{ij} = C_{ijkl}\varepsilon_{kl}$$

avec  $\varepsilon_{ij}(\vec{u}) = \frac{1}{2}(\partial_i u_j + \partial_j u_i)$  et où  $C$  est un tenseur d'ordre 4, vérifiant des conditions de symétrie

$$C_{ijkl} = C_{jikl} = C_{klij}$$

et de positivité:

$$C_{ijkl}\xi_{ij}\xi_{kl} \geq \alpha |\xi|^2, \quad \forall \xi \text{ tenseur symétrique}$$

(avec la convention d'Einstein des indices répétés). Un cas particulier important de matériaux sont les matériaux dits isotropes, c'est à dire dont les propriétés sont identiques dans toutes les directions. Dans ce cas, le milieu ne dépend plus que de deux paramètres,  $\lambda$  et  $\mu$  appelés coefficients de Lamé,

$$C_{ijkl} = \lambda \delta_{ij}\delta_{kl} + \mu(\delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk})$$

et la relation de Hooke devient simplement:

$$\sigma_{ij} = \lambda \text{div}(\mathbf{u})\delta_{ij} + 2\mu\varepsilon_{ij}(\vec{u}) \iff \sigma = \lambda \text{Tr}(\varepsilon)\mathbf{I} + 2\mu\varepsilon$$

L'équation en milieu isotrope homogène ( $\lambda = cte$  et  $\mu = cte$ ) peut alors se réécrire:

$$(5) \quad \begin{aligned} \rho \frac{\partial^2 \vec{u}}{\partial t^2} - (\lambda + \mu)\vec{\text{grad}}(\text{div}\vec{u}) - \mu\vec{\Delta}\vec{u} &= \rho \vec{g} \\ \iff \\ \rho \frac{\partial^2 \vec{u}}{\partial t^2} - (\lambda + 2\mu)\vec{\text{grad}}(\text{div}\vec{u}) + \mu\vec{\text{rot}}(\text{rot}\vec{u}) &= \rho \vec{g} \end{aligned}$$

La dernière équation provient de l'identité  $\vec{\Delta} = \vec{\text{grad}}(\text{div}) - \vec{\text{rot}}(\text{rot})$  ou le deuxième rotationnel noté  $\vec{\text{rot}}$  sans flèche est soit scalaire (en 2D) soit vectoriel (et égal dans ce cas à  $\vec{\text{rot}}$ , en 3D). Cette équation décrit la propagation de 2 types d'ondes se propageant à des vitesses différentes : les ondes de Pression ou ondes P (Pressure) et les ondes de cisaillement ou ondes S (Shear). Ces deux ondes correspondent aux opérateurs divergence et rotationnel. La divergence représente une variation de volume, un mouvement de compression ou de traction, alors que le rotationnel représente une déformation sans variation de volume, donc un cisaillement. Posons  $d = \text{div} u$  et  $r = \text{rot} u$ . Alors il est facile de voir que ces deux quantités obéissent à des équations d'ondes :

- Prenons la divergence de (5), alors en utilisant que  $\text{div}(\vec{\text{grad}}) = \Delta$  et  $\text{div}(\vec{\text{rot}}) = 0$ , on voit que  $d$  vérifie une équation des ondes scalaire:

$$\frac{\partial^2 d}{\partial t^2} - C_p^2 \Delta d = \text{div } \vec{g}$$

où  $C_p = \sqrt{(\lambda + 2\mu)/\rho}$  est la vitesse des ondes P.

- Prenons le rotationnel de (5), alors en utilisant que  $\text{rot}(\vec{\text{grad}}) = 0$  et que  $\vec{\text{rot}}(\text{rot } r) = -\Delta r$  (car  $\text{div } r = \text{div}(\text{rot } u) = 0$ ), on voit que  $r$  vérifie une équation des ondes scalaire en 2D et vectorielle en 3D:

$$\frac{\partial^2 r}{\partial t^2} - C_s^2 \Delta r = \text{rot } \vec{g}$$

où  $C_s = \sqrt{\mu/\rho}$  est la vitesse des ondes S.

**Exercice 1** En cherchant des solutions particulières de (5) sous la forme d'ondes planes:

$$(6) \quad \vec{u}(\vec{x}, t) = e^{i\omega t - \vec{k} \cdot \vec{x}} \vec{d} = e^{i\omega(t - \frac{\vec{n} \cdot \vec{x}}{c})} \vec{d}$$

où  $\vec{k}$  est le vecteur d'onde,  $\vec{n} = \vec{k}/|\vec{k}|$ ,  $\vec{d}$  la direction de propagation et  $\omega$  la pulsation, retrouver les deux types d'ondes planes qui se propagent dans un milieu isotrope. Indiquer la direction de propagation de chacune de ces deux ondes. Que se passe-t'il en 3D?

**Applications.** Les ondes élastiques se rencontrent notamment lorsqu'on étudie des problèmes de sismique ou de géophysique, ou encore dans les problèmes de contrôle non destructif.

### 1.3 Les ondes électromagnétiques

Dans un milieu que nous supposons diélectrique linéaire isotrope, un champ électromagnétique dans un domaine  $\Omega$  est décrit par le champ électrique  $\vec{E}(x, t)$  et le champ magnétique  $\vec{H}(x, t)$  qui vérifient les équations de Maxwell

$$(7) \quad \begin{cases} \mu \frac{\partial \vec{H}}{\partial t} + \text{rot } \vec{E} = 0 \\ \varepsilon \frac{\partial \vec{E}}{\partial t} + \sigma \vec{E} - \text{rot } \vec{H} + \vec{j}_s = 0 \end{cases}$$

où  $\varepsilon(x) > 0$  est la permittivité diélectrique du milieu,  $\mu(x) > 0$  est la perméabilité magnétique du milieu,  $\sigma(x) \geq 0$  est la conductivité du matériau,  $\vec{j}_s$  est un courant source. La vitesse des ondes du milieu est donnée par la relation:

$$c^2 = \frac{1}{\varepsilon\mu}$$

Le vide est un cas particulier de milieu homogène, il est non conducteur donc  $\sigma_0 = 0$  et on a

$$(8) \quad \begin{cases} \varepsilon(x) = \varepsilon_0 = (36\pi 10^9)^{-1} \\ \mu(x) = \mu_0 = 4\pi \cdot 10^{-7} \end{cases}$$

ce qui donne la valeur bien connue pour la vitesse de la lumière :

$$(9) \quad c_0 = 3 \cdot 10^8 \text{ m/s}$$

Lorsque la conductivité est positive, une onde qui se propage dans un milieu conducteur est atténuée au cours de sa propagation c'est à dire qu'il y a un phénomène d'absorption. Nous supposons par la suite que le milieu est non conducteur. Le champ électromagnétique vérifie alors les équations de Maxwell:

$$(10) \quad \begin{cases} \mu \frac{\partial \vec{H}}{\partial t} + \text{rot } \vec{E} = 0 & (i) \\ \varepsilon \frac{\partial \vec{E}}{\partial t} - \text{rot } \vec{H} + \vec{j}_s = 0 & (ii) \end{cases}$$

On peut remarquer en prenant la divergence de (i) que  $(\text{div rot} = 0)$

$$\frac{\partial \text{div}(\mu \vec{H})}{\partial t} = 0$$

Par conséquent si on suppose qu'à l'instant initial on a  $\text{div}(\mu \vec{H}_0) = 0$ , cette propriété reste vraie pour tout instant:

$$\text{div}(\mu \vec{H}) = 0.$$

De même si on prend la divergence de (ii), en supposant que  $\text{div} \vec{j}_s = 0$  et que  $\text{div}(\varepsilon \vec{E}_0) = 0$  alors

$$\text{div}(\varepsilon \vec{E}) = 0$$

Nous ferons ces hypothèses par la suite.

### Formulation en $\vec{E}$

On peut éliminer  $\vec{H}$ , en utilisant la première équation et en dérivant la seconde équation par rapport à  $t$ ,

$$(11) \quad \varepsilon \frac{\partial^2 \vec{E}}{\partial t^2} + \text{rot}(\mu^{-1} \text{rot } \vec{E}) + \frac{\partial \vec{j}_s}{\partial t} = 0$$

### Formulation en $\vec{H}$

De même on peut éliminer  $\vec{E}$

$$(12) \quad \mu \frac{\partial^2 \vec{H}}{\partial t^2} + \text{rot}(\varepsilon^{-1} \text{rot } \vec{H}) + \text{rot}(\varepsilon^{-1} \vec{j}_s) = 0$$

On peut noter une parenté entre les systèmes (11) ou (12) et l'équation des ondes scalaires et sous certaines hypothèses supplémentaires, on retombe exactement sur une équation d'ondes. Plaçons nous par exemple dans le vide. En utilisant la relation  $\Delta = \text{grad div} - \text{rot rot}$ , (11) se réécrit

$$\frac{\partial^2 \vec{E}}{\partial t^2} + c_0^2 (\text{grad div} \vec{E} - \Delta \vec{E}) = -\frac{1}{\varepsilon_0} \frac{\partial \vec{j}_s}{\partial t}$$

ce qui, compte tenu de la relation  $\text{div} \vec{E} = 0$  devient une équation d'ondes:

$$\frac{\partial^2 \vec{E}}{\partial t^2} - c_0^2 \Delta \vec{E} = -\frac{1}{\varepsilon_0} \frac{\partial \vec{j}_s}{\partial t}$$

à laquelle on pourra adjoindre une condition aux limites de conducteur parfait (si  $\Omega$  est en contact avec un conducteur parfait (un métal par exemple)):

$$(13) \quad \vec{E} \wedge n|_{\Gamma} = 0$$



**Applications.** Les ondes électromagnétiques sont beaucoup utilisées pour la détection d'objets volants (furtivité radar). Elles interviennent bien sûr aussi dans tout ce qui concerne les télécommunications, en particulier dans l'étude des antennes ou en optique (fibres optiques, optique intégrée).

### 1.4 Lien entre problème stationnaire et transitoire

Il est commun de distinguer deux types de problèmes associés à ces modèles: les problèmes en temps (ou problèmes transitoires) et les problèmes en fréquence (ou problèmes harmoniques). Dans le premier cas, le temps fait explicitement partie des variables de l'inconnue du problème et on s'intéresse alors à un problème d'évolution. Dans le second cas, la dépendance en temps est imposée a priori, par l'intermédiaire de la source par exemple. Elle est supposée périodique en temps, et même harmonique : on cherche une solution proportionnelle à  $e^{i\omega t}$ , où  $\omega > 0$  désigne la pulsation (on parle aussi de fréquence). Le temps n'intervient alors plus que par l'intermédiaire de cette fréquence, qui joue le rôle de paramètre, et l'inconnue recherchée est une fonction des seules variables spatiales. Ainsi, l'équation des ondes donne naissance à l'équation de Helmholtz

$$-c^2 \Delta u - \omega^2 u = 0.$$

Il convient de souligner que ces deux types de problèmes, bien qu'issus d'une même modélisation physique, ont des propriétés mathématiques radicalement différentes et nécessitent le développement de méthodes numériques adaptées. En outre, la justification mathématique précise du passage de l'un à l'autre s'avère souvent délicate.

Toutefois, il est un point commun entre ces problèmes : l'existence d'une dimension caractéristique qu'on appelle la longueur d'onde. Très intuitivement, cette dimension est la longueur sur laquelle la solution recherchée varie substantiellement (typiquement la période d'une solution périodique). Dans le cas d'une propagation en milieu hétérogène, il conviendra de parler de plusieurs longueurs d'onde, celle-ci pouvant varier d'une région de l'espace à l'autre. Sa connaissance a bien entendu une influence fondamentale sur les propriétés de la solution et donc sur le choix d'une méthode numérique. Un problème sera considéré comme de grande taille si le rapport entre la taille de ce problème (à savoir la taille du domaine de calcul dans le cas d'un problème harmonique ou la distance parcourue par l'onde pendant le temps d'intégration dans le cas d'un problème transitoire) et la longueur d'onde est grand.

Nous montrons sur les figures qui suivent deux exemples. Dans le premier exemple (Figure1) la source est harmonique et donne naissance à une onde qui devient elle-même harmonique au bout d'un certain temps. Dans le deuxième exemple (Figure2), la source est une dérivée de Gaussienne.

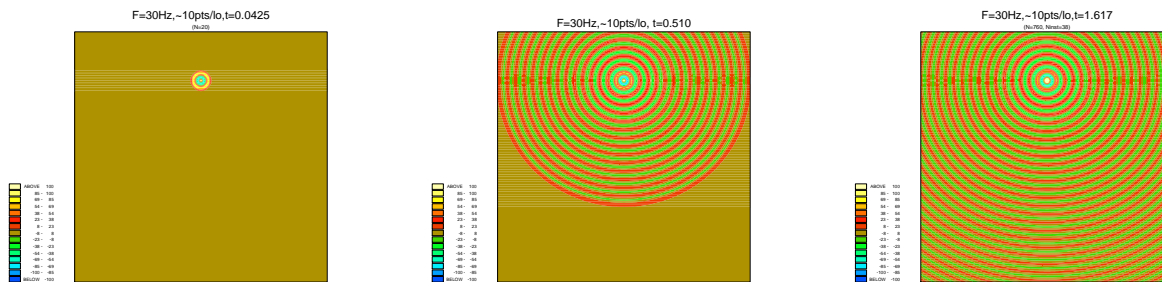


Figure 1: Source harmonique. Régime harmonique en train de s'établir

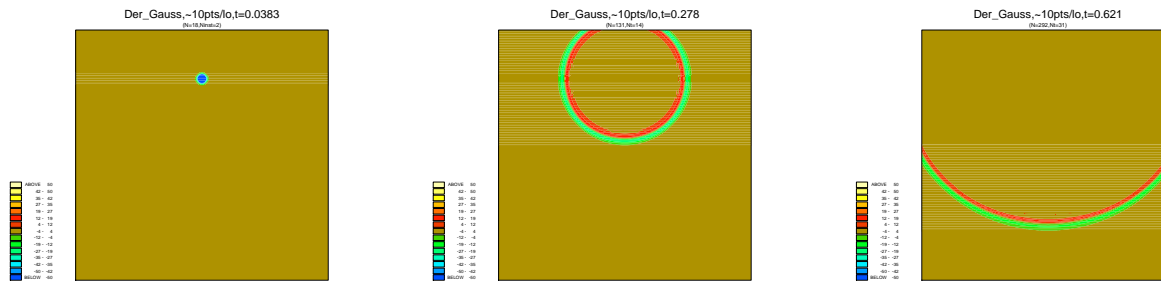


Figure 2: Source=Dérivée de Gaussienne.

## 1.5 Techniques de résolution numériques de ces problèmes

Les problèmes auxquels on est souvent confronté sont de grande taille au sens où le rapport entre la taille du domaine de calcul (ou la distance parcourue par l'onde pendant la durée de la simulation en régime transitoire) et la longueur d'onde peut être grande. La grande taille des problèmes conduit à des calculs sur ordinateur à la fois longs et gourmands en place mémoire. Par ailleurs, des phénomènes parasites, tels que la dispersion numérique par exemple, peuvent venir entacher la fiabilité des résultats. Pour bien comprendre ces phénomènes, et améliorer les résultats en termes de précision et d'efficacité, il est nécessaire d'étudier les méthodes numériques utilisées.

Dans ce cours, nous allons étudier deux types de schémas numériques, basés sur l'utilisation des différences finies et des éléments finis. Nous verrons que les différences finies utilisent une grille de calcul régulière ce qui a pour avantages la facilité d'implémentation et la rapidité de calculs, mais qui présente l'inconvénient de ne pas pouvoir prendre en compte de manière précise des géométries complexes (approximation en marches d'escalier). La méthode des éléments finis repose sur un maillage du domaine (triangulation) qui permet donc de suivre la géométrie du domaine de façon précise. Nous verrons par contre qu'elle ne conduit pas directement à un schéma explicite car elle fait apparaître une matrice de masse. Cette difficulté peut être résolue en approchant cette matrice par une matrice diagonale, ce qui est possible en utilisant des formules de quadrature adéquates, c'est à dire en faisant de la condensation de masse.

Signalons l'existence d'une troisième méthode que nous n'étudierons pas dans ce cours: la méthode des potentiels retardés (ou des équations intégrales en temps). Cette méthode est basée sur une représentation intégrale de la solution en fonction de ses traces sur la frontière du domaine. L'équation intégrale est alors une équation posée sur la frontière et l'inconnue est une des traces de la solution. Cette méthode présente donc l'avantage de réduire considérablement la taille du domaine de calcul et de ne pas nécessiter l'introduction de frontières artificielles pour borner le domaine (cf. §milieux non bornés, plus loin). Cependant, outre le fait qu'elle est beaucoup plus difficile à mettre en oeuvre, elle requiert la connaissance de la solution fondamentale, ce qui limite son application à des milieux particuliers (homogènes, stratifiés...).

**Un exemple de schéma aux différences finies.** Donnons un exemple des difficultés qui peuvent apparaître lorsqu'on résout numériquement l'équation des ondes. Si on veut résoudre le problème

$$\begin{cases} \frac{\partial^2 u}{\partial t^2} - \Delta u = g(t)f(x), & \text{dans } \Omega \\ u(t=0) = 0, \quad \partial_t u(t=0) = 0, & \text{dans } \Omega \\ u = 0, & \text{sur } \Gamma = \partial\Omega \end{cases}$$

On discrétise le domaine :  $h$  est le pas en espace,  $\Delta t$  le pas de temps. On approche la solution  $u(ih, jh, n\Delta t)$  par  $u_{ij}^n = u_h(x_i, y_j, t_n)$  et les dérivées par une expression dite aux différences et dont on peut vérifier la validité (et l'ordre) à l'aide d'un développement de Taylor

$$\frac{\partial^2 u}{\partial t^2} = \frac{u(t + \Delta t) - 2u(t) + u(t - \Delta t)}{\Delta t^2} + O(\Delta t^2)$$

ce qui donne le schéma:

$$\left\{ \begin{array}{l} \frac{u_{ij}^{n+1} - 2u_{ij}^n + u_{ij}^{n-1}}{\Delta t^2} - \frac{u_{i+1j}^n - 2u_{ij}^n + u_{i-1j}^n}{\Delta x^2} \\ \quad - \frac{u_{ij+1}^n - 2u_{ij}^n + u_{ij-1}^n}{\Delta y^2} = g(n\Delta t)f(x_i, y_j) \\ u_{ij}^0 = u_{ij}^1 = 0 \end{array} \right.$$

Le pas  $h$  doit être choisi en fonction des fréquences de la source  $g$ . Si  $h$  est trop grand, on a un problème de *dispersion numérique* et également d'anisotropie numérique, voir Figures 3 et 4. Le pas  $\Delta t$  doit être choisi de telle sorte à satisfaire une condition de stabilité  $c\Delta t/h < \sqrt{2}/2$ . Si  $h$  diminue trop, on doit aussi diminuer  $\Delta t$  ce qui donne un coût de calcul élevé.

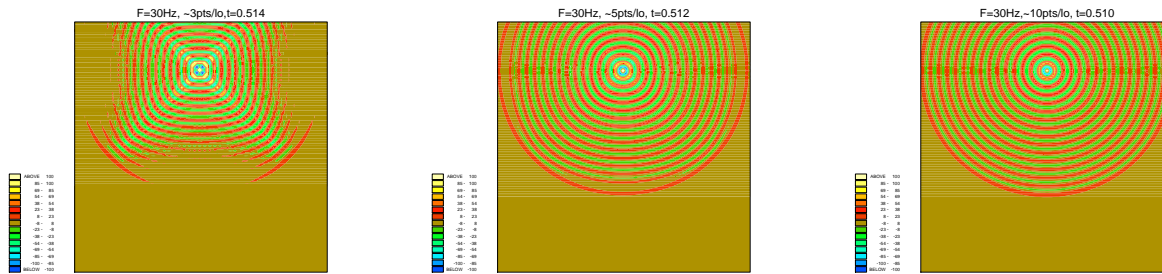


Figure 3: Anisotropie et dispersion numériques en fonction du nombre de points par longueur d'onde. Source harmonique

**Résolution de Problèmes en milieux non bornés.** De nombreux problèmes de propagation d'onde se posent en milieu non borné ou du moins très grand par rapport à la zone d'intérêt: nous pensons par exemple aux problèmes de la diffraction d'une onde électromagnétique par un avion ou à la propagation d'une onde élastique dans le sous-sol. Pour des raisons pratiques évidentes, on est amené à réduire les calculs effectifs à un domaine borné en espace. Se pose alors le problème du traitement de la frontière artificielle ainsi introduite afin de simuler le fait que le milieu de propagation réel est infini. C'est ce qui amène à introduire les notions de conditions aux limites transparentes (i.e. qui n'ont pas d'influence sur la solution), de conditions aux limites absorbantes (conditions aux limites qui sont censées "laisser sortir" les ondes du domaine de calcul en minimisant les réflexions parasites) ou de couches absorbantes (des petites bandes qui sont rajoutées autour du domaine de calcul dans lesquelles on travaille avec un modèle mathématique qui permet de laisser rentrer les ondes dans la couche puis les absorbe). Nous illustrons sur la figure 5 la différence entre une simulation dans un domaine borné avec des conditions aux limites absorbantes d'un côté et des conditions aux limites de Neumann de l'autre.

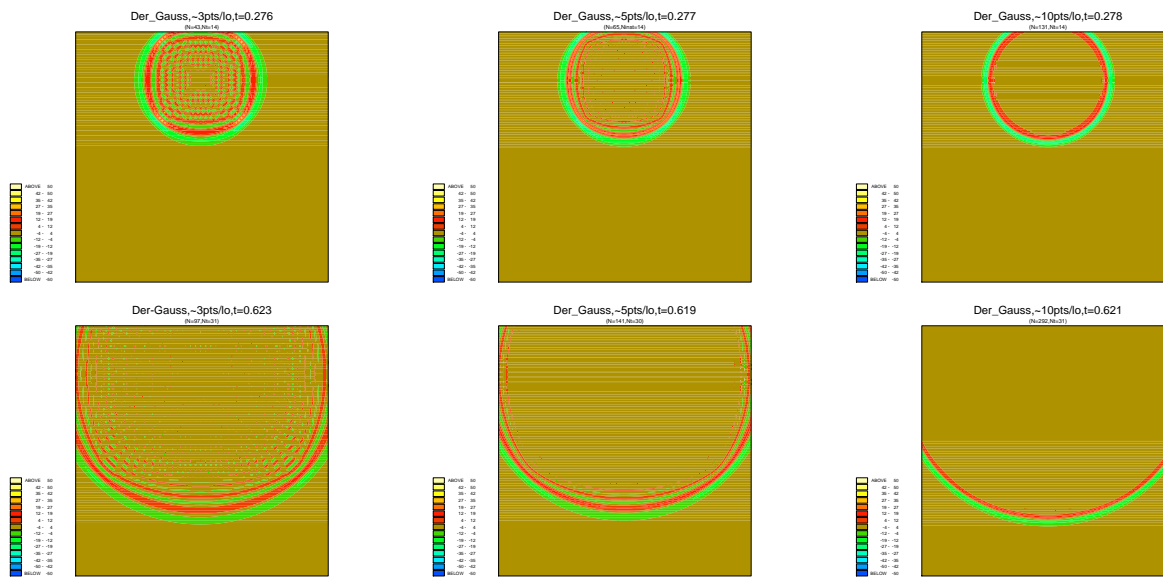


Figure 4: Anisotropie et dispersion numériques en fonction du nombre de points par longueur d'onde. Source transitoire (dérivée de Gaussienne)

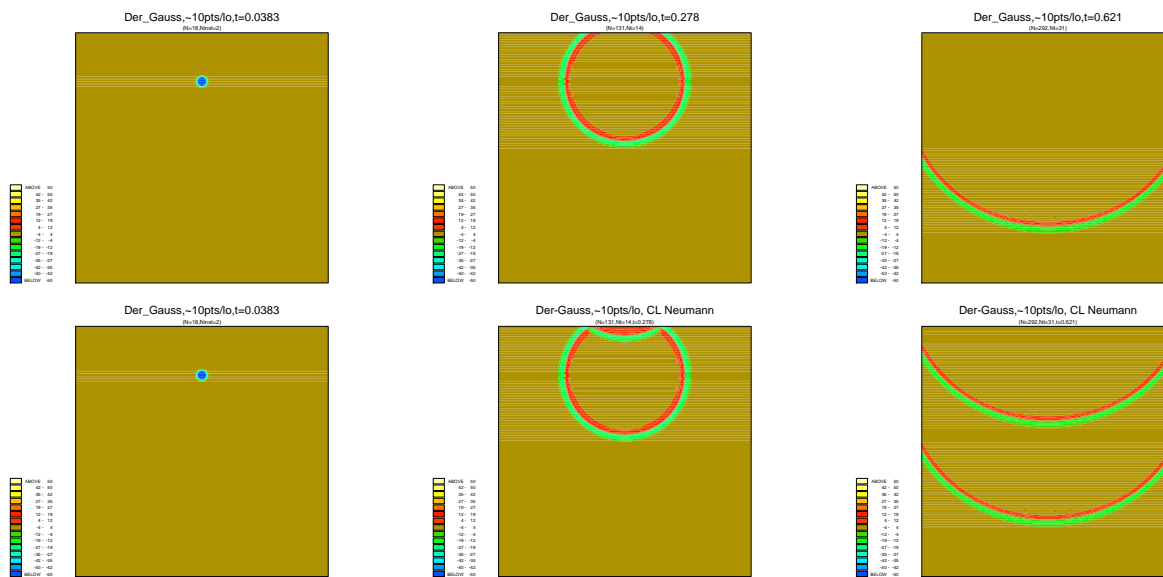


Figure 5: Conditions aux limites absorbantes et conditions aux limites de Neumann

## 2 Quelques propriétés de l'équation des ondes 1D

Pour les propriétés de l'équation des ondes continue, nous renvoyons au cours de P. Joly. Nous faisons ici seulement un bref survol des propriétés essentielles qui nous serviront également pour l'analyse des schémas.

### 2.1 Le problème modèle

Nous considérons le problème de propagation dans tout l'espace suivant :

$$(14) \quad \left\{ \begin{array}{ll} \text{Trouver } u(x, t) : \mathbf{R}^N \times \mathbf{R}^+ \rightarrow \mathbf{R} / & \\ \rho \frac{\partial^2 u}{\partial t^2} - \operatorname{div}(\mu \nabla u) = f & \mathbf{R}^N \times \mathbf{R}^+ \\ u(x, 0) = u_0(x) & \mathbf{R}^N \\ \frac{\partial u}{\partial t}(x, 0) = u_1(x) & \mathbf{R}^N \end{array} \right.$$

Les coefficients  $\rho(x)$  et  $\mu(x)$  caractérisant le milieu de propagation sont supposés satisfaire les hypothèses suivantes :

$$\left\| \begin{array}{l} \bullet \rho(x), \mu(x) \text{ mesurables} \\ \bullet 0 < \rho_- \leq \rho(x) \leq \rho_+ < +\infty \quad p.p. x \in \mathbf{R}^N \\ \bullet 0 < \mu_- \leq \mu(x) \leq \mu_+ < +\infty \quad p.p. x \in \mathbf{R}^N \end{array} \right.$$

Les données du problème sont donc, outre les coefficients  $\rho(x)$  et  $\mu(x)$  :

- les données initiales  $(u_0(x), u_1(x))$
- le second membre  $f(x, t)$ .

L'inconnue du problème est la fonction  $u(x, t)$ .

### 2.2 La formule de D'Alembert et ses conséquences

On considère l'équation des ondes 1D avec une vitesse constante  $c$  et on s'intéresse au problème:

$$(15) \quad \left\{ \begin{array}{l} \frac{\partial^2 u}{\partial t^2} - c^2 \frac{\partial^2 u}{\partial x^2} = f \\ u(x, 0) = u_0(x) \\ \frac{\partial u}{\partial t}(x, 0) = u_1(x) \end{array} \right.$$

**Théorème 1** *La solution du problème (15) est donnée par la formule de d'Alembert :*

$$(16) \quad \left\{ \begin{array}{l} u(x, t) = \frac{u_0(x+ct) + u_0(x-ct)}{2} + \frac{1}{2c} \int_{x-ct}^{x+ct} u_1(s) ds \\ + \frac{1}{2c} \int_0^t ds \int_{|y-x| < c(t-s)} f(y, s) dy \end{array} \right.$$

**Remarque 2** *Cette formule montre en particulier que la solution du problème de Cauchy (ie pour  $f = 0$ ) se décompose sous la forme*

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

où

$$\begin{aligned} u^+(x, t) &= f(x - ct) && \text{correspond à une onde qui se propage dans la direction des } x > 0 \\ u^-(x, t) &= g(x + ct) && \text{correspond à une onde qui se propage dans la direction des } x < 0 \end{aligned}$$

### 2.2.1 Démonstration de la formule de D'Alembert

On se restreint au cas où  $f = 0$ , le cas  $f \neq 0$  est un peu plus technique mais le principe de la démonstration est le même. L'idée est d'utiliser l'identité

$$\partial_t^2 - c^2 \partial_x^2 = (\partial_t - c \partial_x)(\partial_t + c \partial_x)$$

c'est à dire de voir l'équation des ondes comme deux équations de transport se propageant en sens inverse, et d'introduire des variables qui permettent d'intégrer directement:

$$\partial_\xi \longleftrightarrow \partial_t + c \partial_x, \quad \partial_\eta \longleftrightarrow \partial_t - c \partial_x$$

c'est à dire

$$\begin{cases} \xi = x + ct \\ \eta = x - ct \end{cases} \iff \begin{cases} x = \frac{\xi + \eta}{2} \\ t = \frac{\xi - \eta}{2c} \end{cases}$$

et  $U(\xi, \eta) = u(x, t)$ . On a donc

$$\begin{cases} \partial_\xi U = \partial_x u \partial_\xi x + \partial_t u \partial_\xi t = \frac{1}{2c} (\partial_t + c \partial_x) u \\ \partial_\eta U = \partial_x u \partial_\eta x + \partial_t u \partial_\eta t = -\frac{1}{2c} (\partial_t - c \partial_x) u \end{cases}$$

et l'équation d'ondes se réécrit simplement:

$$(17) \quad \partial_{\xi\eta}^2 U(\xi, \eta) = 0$$

Les conditions initiales en  $t = 0$  deviennent des conditions en  $\xi = \eta$ :

$$(18) \quad \begin{aligned} U(\xi, \xi) &= u_0(\xi) \\ \partial_\eta U(\xi, \xi) - \partial_\xi U(\xi, \xi) &= -\frac{u_1(\xi)}{c} \end{aligned}$$

En intégrant (17) resp. par rapport à  $\xi$  et  $\eta$  on obtient:

$$\begin{aligned} (i) \quad \partial_\eta U(\xi, \eta) &= G(\eta), \\ (ii) \quad \partial_\xi U(\xi, \eta) &= F(\xi) \end{aligned}$$

et la deuxième condition initiale devient:

$$G(\xi) - F(\xi) = -\frac{u_1(\xi)}{c}$$

Intégrons (i) par rapport à  $\eta$ , entre  $\eta_0 = \xi$  et  $\eta$ :

$$U(\xi, \eta) - U(\xi, \xi) = \int_\xi^\eta G(s) ds$$

ce qui se réécrit grâce à la première CI:

$$U(\xi, \eta) - u_0(\xi) = \int_{\xi}^{\eta} G(s) ds$$

De même on intègre (ii) par rapport à  $\xi$ , entre  $\xi_0 = \eta$  et  $\xi$ :

$$U(\xi, \eta) - U(\eta, \eta) = - \int_{\xi}^{\eta} F(s) ds \Rightarrow U(\xi, \eta) - u_0(\eta) = - \int_{\xi}^{\eta} F(s) ds$$

et en sommant:

$$2U(\xi, \eta) = u_0(\xi) + u_0(\eta) + \int_{\eta}^{\xi} \frac{u_1(s)}{c} ds$$

ce qui donne la formule de D'Alembert. ■

**Exercice 2** On peut retrouver la formule de D'Alembert en passant par une transformée de Fourier en espace:

$$(19) \quad \mathcal{F}(f)(k) = \hat{f}(k) = \frac{1}{\sqrt{2\pi}} \int f(x) e^{-ikx} dx$$

On rappelle Fourier inverse:

$$(20) \quad f(x) = \frac{1}{\sqrt{2\pi}} \int \hat{f}(k) e^{ikx} dk$$

1- Montrer que la transformée de Fourier de la solution de (15) a l'expression suivante:

$$(21) \quad \hat{u}(k, t) = \hat{u}_0(k) \cos(ckt) + \hat{u}_1(k) \frac{\sin(ckt)}{ck}$$

2- Quelques rappels sur les transformées de Fourier de distributions... On rappelle que  $\hat{\delta} = 1/\sqrt{2\pi}$ . Notons  $\tau_{x_0}$  l'opérateur de translation:

$$\tau_{x_0} f(x) = f(x - x_0)$$

Montrer les propriétés suivantes:

$$\mathcal{F}(\tau_{x_0} f)(k) = e^{-ikx_0} \hat{f}(k)$$

$$\mathcal{F}(f \star g)(k) = \sqrt{2\pi} \hat{f}(k) \hat{g}(k)$$

3- Montrer que

$$\frac{1}{\sqrt{2\pi}} \mathcal{F}^{-1}(\cos(ckt)) = \frac{1}{2} (\delta(x - ct) + \delta(x + ct))$$

$$\hat{H}(k) = \frac{1}{i\sqrt{2\pi}} vp \frac{1}{k} + \sqrt{\frac{\pi}{2}} \delta$$

où  $H$  est la fonction de Heaviside et  $vp \frac{1}{k}$  représente la valeur principale de  $1/k$  définie par

$$\langle vp \frac{1}{k}, \phi \rangle = \lim_{\varepsilon \rightarrow 0} \int_{|k| > \varepsilon} \frac{\phi(k)}{k} dk$$

En déduire que

$$\mathcal{F}(H(x - x_0) - H(x + x_0))(k) = -\sqrt{\frac{2}{\pi}} \frac{\sin kx_0}{k}$$

4- Conclure.

### 2.2.2 Cône de dépendance et propagation à vitesse finie

D'après la formule de D'Alembert, on voit que pour calculer la solution au point  $M = (x, t)$ , on n'a besoin des conditions initiales et du second membre  $f$  seulement dans le cône de dépendance  $D(M)$  (cf Fig. 6). On en déduit une propriété essentielle des ondes qui est leur *propagation à vitesse finie*.

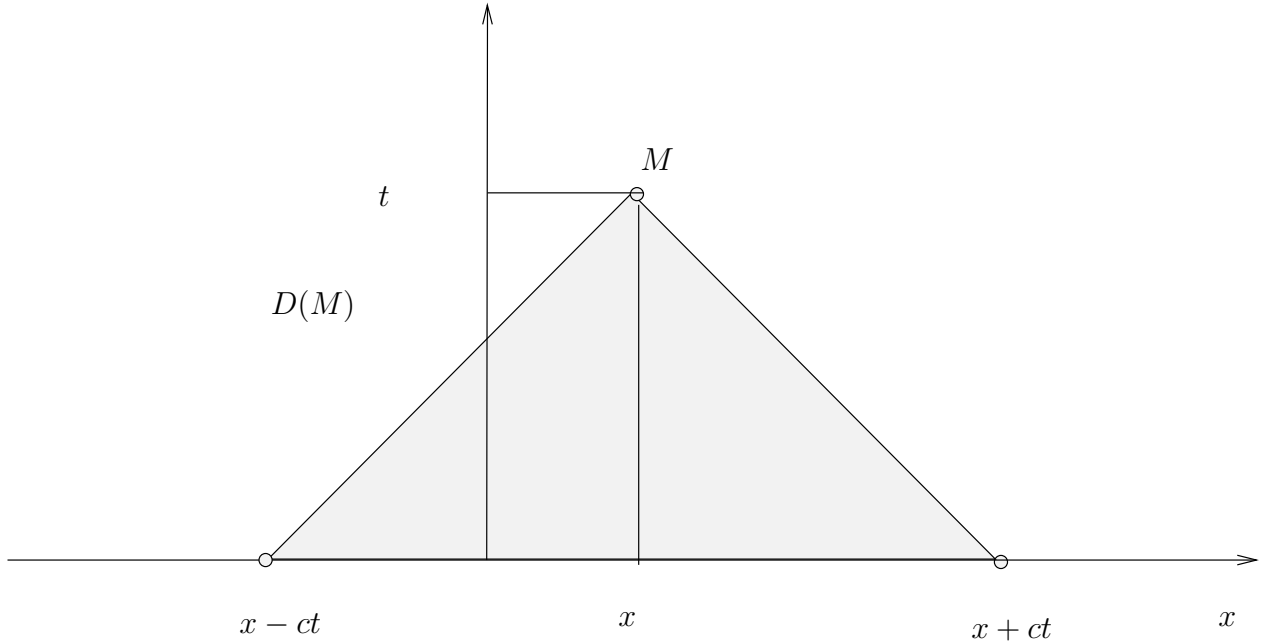


Figure 6: Cône de dépendance

En effet, si on se donne  $f$  à support compact  $K = [a, b]$  (par rapport à  $x$ ) et les conditions initiales  $u_0$  et  $u_1$  dans le même support  $K$ , à un instant  $t$ , la solution  $u(x, t)$  est à support  $K_t = [a - ct, b + ct]$ .

### 2.2.3 Régularité de la solution du problème de Cauchy

D'après la formule de D'Alembert, on voit que si  $u_0 \in C^{k+1}(\mathbb{R})$  et  $u_1 \in C^k(\mathbb{R})$ , alors  $u(., t) \in C^{k+1}(\mathbb{R})$ . Autrement dit, dans l'équation des ondes,  $\partial_t u$  et  $c\partial_x u$  sont de même nature, et si on suppose que  $(\partial_t u(0), c\partial_x u(0)) \in C^k(\mathbb{R}) \times C^k(\mathbb{R})$  alors on a  $(\partial_t u(., t), c\partial_x u(., t)) \in C^k(\mathbb{R}) \times C^k(\mathbb{R})$ . On dit que l'équation des ondes conserve la régularité, et ceci est une caractéristique de la famille des *équations hyperboliques linéaires*. Les équations paraboliques (équation de la chaleur) régularisent : même pour une donnée non régulière, la solution est régulière pour  $t > 0$ . Les équations hyperboliques non linéaires peuvent introduire des singularités, même pour des données très régulières (équations de Burgers et chocs).

### 2.2.4 Régularité de la solution du problème source

Si on compare aux problèmes elliptiques, type Laplacien, il est naturel de se demander si on gagne deux crans de régularité en passant de  $f$  à  $u$ ? La réponse en général est non... On va voir qu'on ne gagne qu'un cran de régularité.

On n'obtient pas facilement d'informations à partir de la formule explicite de la solution, mais plutôt à partir de son expression en Fourier. On peut montrer que la solution du problème source



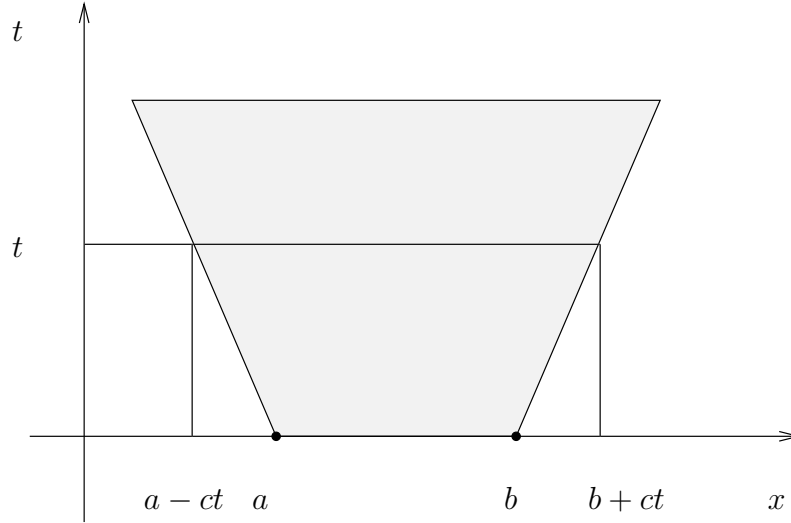


Figure 7: Cône de dépendance

vérifie

$$\hat{u}(k, t) = \int_0^t \frac{\sin(ck(t-s))}{ck} \hat{f}(k, s) ds$$

ce qui implique

$$c \frac{\partial \hat{u}}{\partial x}(k, t) = i \int_0^t \sin(ck(t-s)) \hat{f}(k, s) ds$$

$$\frac{\partial \hat{u}}{\partial t}(k, t) = \int_0^t \cos(ck(t-s)) \hat{f}(k, s) ds$$

Par Cauchy-Schwartz on obtient alors

$$\left| c \frac{\partial \hat{u}}{\partial x}(k, t) \right|^2 \leq t \int_0^t |\hat{f}(k, s)|^2 ds$$

$$\left| \frac{\partial \hat{u}}{\partial t}(k, t) \right|^2 \leq t \int_0^t |\hat{f}(k, s)|^2 ds$$

Finalement Plancherel nous permet de conclure:

$$\left\| c \frac{\partial u}{\partial x} \right\|_{L^2}^2(t) \leq t \int_0^t \int_{\mathbb{R}} |f(x, s)|^2 dx ds$$

$$\left\| \frac{\partial u}{\partial t} \right\|_{L^2}^2(t) \leq t \int_0^t \int_{\mathbb{R}} |f(x, s)|^2 dx ds$$

ce qui montre que

$$f \in L_{loc}^2(\mathbb{R}^+, L^2(\mathbb{R})) \implies u \in L^\infty(\mathbb{R}^+, H^1(\mathbb{R})) \cap W^{1,\infty}(\mathbb{R}^+, L^2(\mathbb{R}))$$

### 2.2.5 Conservation de l'énergie- Unicité

On définit l'énergie par

$$(22) \quad E(t) = \frac{1}{2} \int_{\mathbb{R}} \left( \left| \frac{\partial u}{\partial t} \right|^2 + \left| \frac{\partial u}{\partial x} \right|^2 \right) dx$$

**Exercice 3** 1- En absence de terme source ( $f = 0$ ), montrer qu'il y a conservation de l'énergie, i.e.

$$E(t) = E(0) = \frac{1}{2} \|u_1\|^2 + \frac{1}{2} \left\| \frac{\partial u_0}{\partial x} \right\|^2, \quad \forall t > 0$$

On pourra montrer ce résultat de deux façons. La première façon consiste à partir de l'équation des ondes, à la multiplier par  $\partial_t u$  et à intégrer par parties. La deuxième démonstration se fait directement à partir de la formule de D'Alembert.

2- En déduire l'unicité de la solution.

3- On rajoute maintenant un terme source  $f \in L^2(\mathbb{R})$ . Montrer que, à tout instant, l'énergie reste bornée':

$$(23) \quad E(t)^{1/2} \leq E(0)^{1/2} + C \int_0^t \|f\|(s) ds$$

## 2.2.6 Stabilité par Fourier

### 2.3 Ondes planes harmoniques

Ce sont des solutions particulières de l'équation des ondes, qui s'écrivent sous la forme :

$$(24) \quad u(x, t) = e^{i(kx - \omega t)}$$

où  $k \in \mathbb{R}$  est le nombre d'onde,  $\omega$  la pulsation. On note  $\lambda = \frac{2\pi}{k}$  la longueur d'onde (période en espace) et  $T = \frac{2\pi}{\omega}$  la période. La fonction  $u$  est une solution de l'équation des ondes si la *relation de dispersion* est vérifiée:

$$\omega^2 = c^2 k^2$$

ce qui montre en particulier que  $\lambda = \frac{2\pi c}{\omega} = cT$ .

Ces solutions particulières jouent un rôle important: en utilisant la transformation de Fourier en espace (19) on montre que toute solution d'énergie finie de l'équation des ondes est une superposition d'ondes planes harmoniques. Plus précisément, on peut montrer que si  $u$  est une solution d'énergie finie, on peut l'écrire comme

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

où  $u^+$  est une onde se propageant vers la droite,  $u^-$  est une onde se propageant vers la gauche, avec

$$u^+(x, t) = \int a^+(k) e^{i(kx - \omega^+(k)t)} dk, \quad \omega^+(k) = ck$$

$$u^-(x, t) = \int a^-(k) e^{i(kx - \omega^-(k)t)} dk, \quad \omega^-(k) = -ck$$

les amplitudes  $a^+(k)$  et  $a^-(k)$  dépendant des conditions initiales:

$$a^+(k) = \frac{1}{2} \hat{u}_0(k) + \frac{i}{2ck} \hat{u}_1(k)$$

$$a^-(k) = \frac{1}{2} \hat{u}_0(k) - \frac{i}{2ck} \hat{u}_1(k)$$

Ce résultat s'obtient facilement à partir de l'expression en Fourier de la solution (21) établi à l'exercice 2, qui montre que:

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

### 3 Semi-discrétisation en espace de l'équation des ondes 1D par un schéma aux différences finies

On s'intéresse ici à l'approximation en espace de l'équation des ondes (problème de Cauchy) :

$$(25) \quad \begin{cases} \frac{1}{c^2} \frac{\partial^2 u}{\partial t^2} - \frac{\partial^2 u}{\partial x^2} = 0, & x \in \mathbb{R}, t > 0 \\ u(x, t = 0) = u_0(x), & x \in \mathbb{R} \\ \frac{\partial u}{\partial t}(x, t = 0) = u_1(x), & x \in \mathbb{R} \end{cases}$$

et on se propose dans un premier temps de l'approcher par différences finies. Nous verrons au paragraphe 5 un autre type d'approximations issu de la méthode des éléments finis. Dans tout ce qui suit on supposera que les données initiales  $u_0$  et  $u_1$  sont à support compact

$$u_0(x) = u_1(x) = 0 \quad \text{pour } |x| \geq L$$

#### 3.1 Quelques préliminaires sur les différences finies.

Soit  $f$  une fonction "assez régulière" (au moins  $C^2$ ) définie sur  $\mathbb{R}$  et soit  $h$  le pas de discrétisation. On veut approcher les dérivées de  $f$  au point  $x_j = jh$  en utilisant uniquement les valeurs de  $f$  en ces points. Il est naturel d'approcher la dérivée de  $f$  en un point  $x$  par un taux d'accroissement, par exemple décentré à droite:

$$f'(x) \sim \frac{f(x+h) - f(x)}{h}$$

ou décentré à gauche:

$$f'(x) \sim \frac{f(x) - f(x-h)}{h}$$

On peut vérifier l'ordre de ces approximations grâce à la formule de Taylor, qui nous dit (si  $f \in C^2$ )

$$(26) \quad f(x_{j+1}) = f(x_j) + hf'(x_j) + \frac{h^2}{2} f''(\xi_j), \quad \xi_j \in ]x_j, x_{j+1}[$$

ce qui montre que :

$$f'(x_j) = \frac{f(x_{j+1}) - f(x_j)}{h} + O(h)$$

où  $|O(h)| = \left| \frac{h}{2} f''(\xi_j) \right| \leq \frac{h}{2} \sup_{\xi \in ]x_j, x_{j+1}[} |f''(\xi)|$ . L'approximation à l'aide du taux d'accroissement décentré est donc une approximation décentrée d'ordre 1. On peut, en faisant une moyenne des deux approximations décentrées à gauche et à droite, obtenir une approximation centrée et, toujours grâce à un développement de Taylor, montrer qu'on gagne un ordre de précision, mais on doit pour cela supposer plus de régularité ( $f \in C^3$ ) pour obtenir:

$$f'(x_j) = \frac{f(x_{j+1}) - f(x_{j-1}))}{2h} + \frac{h^2}{6} (f'''(\xi_j^+) + f'''(\xi_j^-))$$

Cette approximation centrée est d'ordre 2.

Pour approcher la dérivée seconde, on procède de la même manière. On a vu que l'approximation centrée est la plus précise, c'est donc celle qu'on va retenir et on écrit que  $f''(x)$  est la dérivée de  $f'(x)$ , qu'on approche par:

$$\begin{aligned} f''(x) &\sim \frac{f'(x+h/2) - f'(x-h/2)}{h/2} \\ &\sim \frac{2}{h} \left( \frac{f(x+h) - f(x)}{h/2} - \frac{f(x) - f(x-h)}{h/2} \right) \\ &\sim \frac{f(x+h) - 2f(x) + f(x-h)}{h^2} \end{aligned}$$

Afin de vérifier l'ordre de cette approximation, on utilise de nouveau un développement de Taylor:

$$\begin{aligned} f(x_{j+1}) &= f(x_j) + hf'(x_j) + \frac{h^2}{2}f''(x_j) + \frac{h^3}{6}f^{(3)}(x_j) + \frac{h^4}{24}f^{(4)}(\xi_j^+), \quad \xi_j^+ \in ]x_j, x_{j+1}[ \\ f(x_{j-1}) &= f(x_j) - hf'(x_j) + \frac{h^2}{2}f''(x_j) - \frac{h^3}{6}f^{(3)}(x_j) + \frac{h^4}{24}f^{(4)}(\xi_j^-), \quad \xi_j^- \in ]x_{j-1}, x_j[ \end{aligned}$$

et en sommant:

$$f''(x_j) = \frac{f(x_{j+1}) - 2f(x_j) + f(x_{j-1}))}{h^2} + O(h^2)$$

ce qui montre que c'est une approximation centrée d'ordre 2 valable si  $f \in C^4$ .

### 3.2 Schéma semi-discrétisé en espace.

On approche  $u(x_j, t)$  par  $u_j(t)$  solution du schéma obtenu en utilisant une approximation centrée d'ordre 2:

$$(27) \quad \begin{cases} \frac{1}{c^2} \frac{d^2 u_j}{dt^2} - \frac{u_{j+1} - 2u_j + u_{j-1}}{h^2} = 0 \\ u_j(0) = u_{0,j} \\ \frac{du_j}{dt}(0) = u_{1,j} \end{cases}$$

où les suites  $u_{0,h} = (u_{0,j})_{j \in Z}$  et  $u_{1,h} = (u_{1,j})_{j \in Z}$  sont des approximations des données initiales  $u_0$  et  $u_1$ . Si les données initiales sont continues, on peut prendre par exemple

$$(28) \quad u_{\alpha,j} = u_{\alpha}(x_j), \quad \alpha = 0, 1,$$

Un autre choix classique consiste à prendre les valeurs moyennes:

$$(29) \quad u_{\alpha,j} = \frac{1}{h} \int_{x_j-h/2}^{x_j+h/2} u_{\alpha}(x) dx, \quad \alpha = 0, 1,$$

on verra plus loin que ce deuxième choix correspond à une projection sur l'espace approché  $L_h^2$ .

**Exercice 4** Écrire un schéma centré d'ordre 2 pour l'équation des ondes en 2D. Représenter sur un graphique quels points sont utilisés pour l'approximation de l'opérateur  $-c^2 \Delta$ .

### 3.3 Analyse par Fourier discret

**Définition de  $L_h^2$ .** Pour analyser ce problème, on le considère comme un problème d'évolution posé dans l'espace de Hilbert:

$$(30) \quad L_h^2 = \left\{ u_h = (u_j)_{j \in \mathbb{Z}}, \sum_j |u_j|^2 < +\infty \right\}$$

muni de la norme  $\|u_h\|^2 = h \sum_j |u_j|^2$  et du produit scalaire associé  $(u_h, v_h) = h \sum_j u_j \bar{v}_j$ . L'espace  $L_h^2$  est isomorphe au sous-espace de  $L^2(\mathbb{R})$  des fonctions constantes par morceaux: en effet  $u_h$  peut être assimilé à la fonction

$$u_h(x) = u_j \quad \text{si } x \in ]x_{j-1/2}, x_{j+1/2}[ \equiv ]x_j - h/2, x_j + h/2[$$

et le produit scalaire sur  $L_h^2$  coïncide ainsi avec le produit scalaire usuel de  $L^2(\mathbb{R})$ :

$$(u_h, v_h) = \int_{\mathbb{R}} u_h(x) v_h(x) dx$$

Si  $\pi_h$  est la projection orthogonale de  $L^2(\mathbb{R})$  sur  $L_h^2$ , pour une fonction  $f \in L^2(\mathbb{R})$  on définit  $f_h = \pi_h f$  par

$$(f_h, v_h) = (f, v_h), \quad \forall v_h \in L_h^2$$

En choisissant  $v_h = \chi_j = 1_{]x_{j-1/2}, x_{j+1/2}[}$  on a donc:

$$h f_j = \int_{x_j - h/2}^{x_j + h/2} f(x) dx$$

On voit par conséquent que le deuxième choix de conditions initiales (29) correspond à

$$(31) \quad u_{0,h} = \pi_h u_0, \quad u_{1,h} = \pi_h u_1$$

**Rappel sur la transformée de Fourier discrète.** On pose  $K_h = \left[ -\frac{\pi}{h}, \frac{\pi}{h} \right]$  et on introduit:

$$\left\{ \begin{array}{l} \mathcal{F}_h : L_h^2 \longrightarrow L^2(K_h) \\ u_h \longrightarrow \hat{u}_h = \mathcal{F}_h u_h \\ \hat{u}_h(k) = \frac{1}{\sqrt{2\pi}} \sum_j u_j e^{-ikx_j h} \end{array} \right.$$

où la série converge au sens  $L^2$ .  $\mathcal{F}_h$  est un isomorphisme et on a la formule d'inversion :

$$u_j = \frac{1}{\sqrt{2\pi}} \int_{K_h} \hat{u}_h(k) e^{ikx_j} dk$$

De plus, l'identité de Plancherel-Parseval montre que  $\mathcal{F}_h$  est une isométrie:

$$(32) \quad \sum_j |u_j|^2 h = \int_{K_h} |\hat{u}_h(k)|^2 dk$$

Enfin,  $\mathcal{F}_h$  diagonalise l'opérateur de translation: ( $q \in \mathbb{Z}$ )

$$\left\{ \begin{array}{l} \tau_q : L_h^2 \longrightarrow L_h^2 \\ u_h \longrightarrow v_h = \tau_q u_h \\ v_j = u_{j+q}, \quad \forall j \in \mathbb{Z} \end{array} \right.$$

On note  $\widehat{\tau}_q$  l'opérateur défini par  $\widehat{\tau}_q \widehat{u}_h = \widehat{\tau_q u_h}$ .

$$\begin{array}{ccc} L_h^2 & \xrightarrow{\tau_q} & L_h^2 \\ \mathcal{F}_h \downarrow & & \downarrow \mathcal{F}_h \\ L^2(K_h) & \xrightarrow{\widehat{\tau}_q} & L^2(K_h) \end{array}$$

On montre facilement que (exercice):

$$\widehat{\tau}_q \widehat{u}_h(k) = \exp(iqhk) \widehat{u}_h(k)$$

ce qui montre que l'opérateur  $\widehat{\tau}_q$  ( $\widehat{\tau}_q(k) = \exp(iqhk)$ ) est un opérateur de multiplication. On l'appelle le *symbole* de  $\tau_q$ .

**Application: réécriture du schéma (27):**

on peut réécrire l'équation du schéma sous la forme:

$$\frac{d^2 u_j}{dt^2} - \frac{c^2}{h^2} ((\tau_1 + \tau_{-1} - 2\tau_0) u_h)_j = 0,$$

Le problème semi-discrétisé consiste donc à chercher  $u_h \in L_h^2$  solution de:

$$(33) \quad \left\{ \begin{array}{l} \frac{d^2 u_h}{dt^2} + A_h u_h = 0 \quad (i) \\ u_h(0) = u_{0,h} \quad (ii) \\ \frac{du_h}{dt}(0) = u_{1,h} \quad (iii) \end{array} \right.$$

où on a posé  $A_h = -\frac{c^2}{h^2}(\tau_1 + \tau_{-1} - 2\tau_0)$  et où on suppose que les données initiales approchées sont dans  $L_h^2$

$$(34) \quad u_{0,h} \in L_h^2, \quad u_{1,h} \in L_h^2$$

L'opérateur  $A_h$  étant une combinaison linéaire de translations, il est facile de calculer son symbole:

$$\begin{aligned} \widehat{A}_h(k) &= -\frac{c^2}{h^2}(\widehat{\tau}_1 + \widehat{\tau}_{-1} - 2\widehat{\tau}_0) = -\frac{c^2}{h^2}(e^{ikh} + e^{-ikh} - 2) \\ &= \frac{c^2}{h^2}(2 \cos kh - 2) = \frac{4c^2}{h^2} \sin^2 \frac{kh}{2} \end{aligned}$$

**Théorème 2** *L'opérateur  $A_h$  est autoadjoint positif.*

**Démonstration : (i) Par Fourier :** on a vu que  $A_h$  était composé uniquement de sommes de translations et Fourier diagonalise les translations d'où

$$\widehat{A_h u_h}(k) = D_h(k) \widehat{u_h}(k)$$

où  $D_h(k) = \widehat{A_h}(k) = \frac{4c^2}{h^2} \sin^2 \frac{kh}{2}$  est le symbole de l'opérateur  $A_h$ . En appliquant Plancherel:

$$(A_h u_h, v_h) = \int_{K_h} D_h(k) \widehat{u_h}(k) \overline{\widehat{v_h}(k)} dk$$

Or  $D_h(k) \in \mathbb{R}$  ce qui implique que  $A_h$  est autoadjoint. De plus  $D_h(k) \geq 0$  ( $D_h(k) = 0$  si  $kh = 2n\pi$ ) ce qui implique que  $A_h$  est positif, et même défini positif car

$$\begin{aligned} (A_h u_h, u_h) = 0 &\implies \int D_h(k) |\widehat{u_h}(k)|^2 dk = 0 \\ &\implies D_h(k) |\widehat{u_h}(k)|^2 = 0, \text{ pp } k \implies |\widehat{u_h}(k)|^2 = 0, \text{ pp } k. \blacksquare \end{aligned}$$

**(i) Par des formules de Green discrètes**

**Lemme 1** *On montre l'identité:*

$$(35) \quad (A_h u_h, v_h) = a_h(u_h, v_h)$$

avec

$$(36) \quad a_h(u_h, v_h) = c^2 h \sum_j \frac{u_{j+1} - u_j}{h} \frac{v_{j+1} - v_j}{h}$$

le résultat est alors immédiat.  $\blacksquare$

**Exercice 5** *Démontrer le lemme 1 (intégrations par parties discrètes).*

**Remarque 3** *On peut faire un parallèle systématique entre les problèmes discrets et continus. Ainsi:*

- $A_h$  est l'équivalent discret de  $A = -c^2 \frac{d^2}{dx^2}$ .
- $D_h(k)$  est l'équivalent de  $\widehat{A}(k) = c^2 k^2$  et on a

$$D_h(k) = c^2 k^2 (1 + O(k^2 h^2))$$

- $a_h$  est l'équivalent de

$$a(u, v) = \int c^2 \frac{du}{dx} \frac{\overline{dv}}{dx} dx$$

- l'égalité  $a_h(u_h, v_h) = (A_h u_h, v_h)$  est l'équivalent discret de l'intégration par parties:

$$-c^2 \int \frac{d^2 u}{dx^2} \overline{v} dx = \int \frac{du}{dx} \frac{\overline{dv}}{dx} dx$$

**Existence et unicité de la solution par Fourier.**

Nous allons montrer de manière analogue au cas continu le résultat suivant:

**Théorème 3** *Le problème (33) admet une solution unique  $u_h \in C^\infty(\mathbb{R}^+, L_h^2)$*

Ce théorème peut être vu comme la conséquence de résultats généraux sur les EDO (Equations Différentielles Ordinaires) dans un espace de Banach []. Toutefois, la démonstration peut se faire comme dans le cas continu de façon constructive. On a vu à l'exercice 2 qu'on peut établir la formule de D'Alembert à partir de la transformée de Fourier

$$\left\{ \begin{array}{l} \frac{\partial^2 u}{\partial t^2} - c^2 \frac{\partial^2 u}{\partial x^2} = 0 \\ u(0) = u_0, \quad \frac{du}{dt}(0) = u_1 \end{array} \right. \implies \left\{ \begin{array}{l} \frac{d^2 \hat{u}}{dt^2} + c^2 k^2 \hat{u}(k) = 0 \\ \hat{u}(0) = \hat{u}_0, \quad \frac{d\hat{u}}{dt}(0) = \hat{u}_1 \end{array} \right.$$

ce qui en intégrant donne:

$$\hat{u}(k, t) = \hat{u}_0 \cos(\omega t) + \frac{\hat{u}_1}{\omega} \sin(\omega t)$$

où  $\omega = ck$  (relation de dispersion continue).

De même, si on prend la transformée de Fourier discrète du schéma, on obtient alors un simple problème d'évolution qu'on peut intégrer explicitement:

$$\left\{ \begin{array}{l} \frac{d^2 u_h}{dt^2} + A_h u_h = 0 \\ u_h(0) = u_{0,h}, \quad \frac{du_h}{dt}(0) = u_{1,h} \end{array} \right. \implies \left\{ \begin{array}{l} \frac{d^2 \hat{u}_h}{dt^2} + D_h(k) \hat{u}_h(k) = 0 \\ \hat{u}_h(0) = \hat{u}_{0,h}, \quad \frac{d\hat{u}_h}{dt}(0) = \hat{u}_{1,h} \end{array} \right.$$

$D_h(k)$  étant positif, on peut l'écrire sous la forme  $D_h(k) = \omega_h^2(k)$ , qui est la *relation de dispersion du schéma*, ce qui donne

$$(37) \quad \hat{u}_h(k, t) = \hat{u}_{0,h} \cos(\omega_h t) + \frac{\hat{u}_{1,h}}{\omega_h} \sin(\omega_h t)$$

avec

$$(38) \quad \omega_h(k) = \frac{2c}{h} \left| \sin \frac{kh}{2} \right| = ck \frac{2}{kh} \left| \sin \frac{kh}{2} \right|$$

( $|\omega_h(k)| \leq ck, \omega_h(k) = ck(1 + O(k^2 h^2))$ ). On vérifie facilement que la solution ainsi construite est bien dans  $C^\infty(\mathbb{R}^+, L_h^2)$ , en remarquant que

$$|\hat{u}_h(k, t)| \leq |\hat{u}_{0,h}| + t|\hat{u}_{1,h}|$$

et on conclut en utilisant l'hypothèse sur les données initiales  $u_{0,h} \in L_h^2$  et  $u_{1,h} \in L_h^2$ .

On voit que ce qui joue le rôle de  $\omega = ck$  dans le cas discret c'est la variable  $\omega_h$  et la relation (38) est la relation de dispersion discrète:  $\omega_h$  n'est plus une fonction linéaire de  $k$ , il y a dispersion numérique des ondes.



**Dispersion numérique.** On a défini une relation de dispersion continue et une relation de dispersion du schéma. Ces relations peuvent s'obtenir simplement en recherchant les ondes planes solutions. Pour le problème continu, une onde plane est de la forme  $u(x, t) = e^{i(\omega t - kx)}$  et est solution si  $\omega^2 = c^2 k^2$ . Pour le problème semi-discrétisé:

$$\frac{\partial^2 u_h}{\partial t^2} + A_h u_h = 0$$

une onde plane est de la forme  $u_j(t) = e^{i(\omega_h t - kx_j)}$  et est solution si  $D_h(k) = \omega_h^2$ . On voit sur ces expressions que la vitesse de l'onde continue est  $\omega/k = c$  et celle de l'onde semi-discrétisée est  $\omega_h/k = c_h$ . La différence  $(c - \omega_h/k)t$  à un instant donné  $t$  représente donc le déphasage entre l'onde continue et l'onde semi-discrétisée. Ce déphasage augmente au cours du temps. Pour quantifier un schéma, on représente les *courbes de dispersion*, obtenues en traçant le rapport entre vitesse continue et semi-discrète:

$$q_h = \frac{\omega_h}{kc} = \frac{2}{kh} \sin \frac{kh}{2}$$

Ce rapport ne dépend en fait que de la quantité  $K = kh$ . On a l'habitude de représenter  $q_h$  en fonction de l'inverse du nombre de points par longueur d'onde  $G = \frac{K}{2\pi} = \frac{h}{\lambda}$ .

### Une première estimation de la solution - Stabilité.

Nous faisons ici les hypothèses suivantes sur l'approximation des données initiales: il existe une constante  $C > 0$  indépendante de  $h$  telle que, pour toute fonction  $u_0$  et  $u_1$ , et pour  $h$  assez petit, on a:

$$(39) \quad \begin{aligned} (i) \quad & \|u_{0h}\|_{L^2} \leq C \|u_0\|_{L^2} \\ (ii) \quad & \|u_{1h}\|_{L^2} \leq C \|u_1\|_{L^2} \end{aligned}$$

**Théorème 4** On suppose que  $u_{0h}$  et  $u_{1h}$  vérifient (39). La solution du schéma (33) vérifie alors l'estimation suivante:

$$(40) \quad \|u_h(t)\|_{L^2} \leq C(\|u_0\|_{L^2} + t \|u_1\|_{L^2})$$

où  $C$  est une constante positive indépendante de  $h$ .

**Démonstration:** la démonstration découle de la formule (37). ■

**Exercice 6** Les hypothèses (39) sont elles vérifiées pour le choix (29) de conditions initiales approchées ? pour le choix (28), en supposant que  $u_0$  et  $u_1$  sont dans  $C^0(\mathbb{R}) \cap H^1(\mathbb{R})$  ?

## 3.4 Analyse de stabilité par des techniques d'énergie

Les techniques d'énergie reposent sur des estimations a priori de la solution (i.e. sans connaître son expression) et ont l'avantage de s'appliquer en milieu hétérogène ainsi que sur des grilles irrégulières (voir paragraphe éléments finis). Par souci de simplicité, nous présentons cependant les démonstrations dans le cas homogène mais l'extension au cas hétérogène ne présente pas de difficulté particulière.

**Théorème 5** Nous supposons que  $u_0 \in H^1$  et  $u_1 \in L^2$ , que  $u_{0,h}$  et  $u_{1,h}$  vérifient (39) et de plus que

$$(41) \quad \exists C > 0 (\text{ indép. de } h), \quad a_h(u_{0,h}, u_{0,h}) \leq Ca(u_0, u_0)$$

1. La solution du problème (33) satisfait la conservation d'énergie

$$(42) \quad E_h(t) = E_h(0), \quad \forall t \geq 0$$

où l'énergie discrète est définie par

$$(43) \quad E_h(t) = \frac{1}{2} \left\| \frac{du_h}{dt} \right\|^2 + \frac{1}{2} a_h(u_h, u_h)$$

2. L'énergie initiale est bornée par

$$(44) \quad E_h(0) \leq \frac{1}{2} \int \left( |u_1|^2 + c^2 \left| \frac{du_0}{dx} \right|^2 \right) dx = E(0)$$

**Démonstration:** 1- Conservation de l'énergie discrète. La conservation de l'énergie discrète s'obtient comme dans le cas continu: on multiplie scalairement l'équation du schéma (33)-(i) par  $\frac{du_h}{dt}$ :

$$\left( \frac{d^2 u_h}{dt^2} + A_h u_h, \frac{du_h}{dt} \right) = 0$$

ce qui peut encore s'écrire,  $A_h$  étant autoadjoint:

$$\frac{1}{2} \frac{d}{dt} \left\| \frac{du_h}{dt} \right\|^2 + \frac{1}{2} \frac{d}{dt} (A_h u_h, u_h) = 0$$

soit encore  $\frac{dE_h}{dt}(t) = 0$ .

2- Majoration de l'énergie discrète indépendamment de  $h$ . Le deuxième point est immédiatement obtenu grâce aux hypothèses faites sur les données initiales (39) et (41). ■

**Exercice 7** Montrer que l'hypothèse (41) est bien satisfaite pour les deux choix de  $u_{0,h}$  donnés par (28) et (29).

**Estimation de la solution à partir de l'estimation d'énergie.** Il est facile d'obtenir une estimation de stabilité sur la solution à partir de l'estimation d'énergie (44) qui peut se réécrire:

$$\left\| \frac{du_h}{dt} \right\|^2 + a_h(u_h, u_h) \leq 2E(0) \iff (*) \quad h \sum_j \left| \frac{du_j}{dt} \right|^2 + c^2 h \sum_j \left| \frac{u_{j+1} - u_j}{h} \right|^2 \leq 2E(0)$$

Si on veut obtenir une estimation sur  $u_h$ , on utilise l'identité:

$$u_j(t) = \underbrace{u_j(0)}_{(u_{0h})_j} + \underbrace{\int_0^t \frac{du_j}{dt}(s) ds}_{\equiv v_j}$$

c'est à dire

$$u_h = u_{0h} + v_h, \quad \text{avec} \quad v_j = \int_0^t \frac{du_j}{dt}(s) ds$$

d'où

$$\|u_h\| \leq \|u_{0h}\| + \|v_h\| \leq \|u_0\| + \|v_h\|$$

et par définition:

$$\|v_h\|^2 = h \sum_j |v_j(t)|^2 = h \sum_j \left| \int_0^t \frac{du_j}{dt}(s) ds \right|^2 \leq h \sum_j \int_0^t \left| \frac{du_j}{dt}(s) \right|^2 ds \times t$$

Or d'après (\*) on a  $h \sum_j \left| \frac{du_j}{dt} \right|^2 \leq 2E(0)$  d'où

$$h \sum_j \int_0^t \left| \frac{du_j}{dt}(s) \right|^2 ds \leq 2tE(0)$$

et par conséquent

$$\|v_h\|^2 \leq 2t^2 E(0) \implies \|v_h\| \leq \sqrt{2t} \sqrt{E(0)}$$

et on a donc

$$\|u_h\| \leq \|u_0\| + \sqrt{2t} \sqrt{E(0)}$$

On retrouve le même type d'estimation que celle qu'on avait obtenue à partir de l'expression de la solution (en Fourier). Mais ici  $E(0)$  est borné non pas par les normes  $L^2$  des conditions initiales mais par  $\|u_1\|_{L^2} + \|u_0\|_{H^1}$ . En Fourier on obtenait une estimation  $L^2$  de la solution en fonction des normes  $L^2$  des conditions initiales, alors qu'ici on obtient une estimation dans la norme de l'énergie en fonction de l'énergie initiale (on obtient donc aussi un contrôle de la dérivée spatiale de la solution). Il est plus naturel de se placer dans la norme de l'énergie que dans la norme  $L^2$ , car c'est elle qui est liée à la physique du pb, c'est bien une énergie qui se propage, et les solutions qui nous intéressent physiquement sont les solutions d'énergie finie.

### 3.5 Convergence du schéma semi-discrétisé

On peut introduire différentes notions de convergence: il faut préciser ce qu'on entend comparer avec la solution exacte et pour quelles normes... Nous allons ici introduire deux notions de convergence. Dans le premier cas, nous démontrerons une convergence par des techniques d'énergie, dans le deuxième par Fourier. L'étude de la convergence qui passe par des techniques d'énergie compare deux suites de  $L^2_h$ : la suite donnée par le schéma  $(u_j(t))_j$  et la suite des valeurs de la solution exacte aux noeuds  $x_j$ ,  $(u(x_j, t))_j$ . Ceci revient bien entendu à comparer les deux fonctions constantes par morceaux associées. L'étude de la convergence par Fourier compare deux fonctions de  $L^2$ : la solution exacte  $u$  et une fonction  $\tilde{u}^h$  qui vérifie une "version continue du schéma". On a illustré ces deux notions différentes sur la figure 8.

#### 3.5.1 Convergence du schéma semi-discrétisé par techniques d'énergie

**Exercice 8** Soit  $u$  la solution (supposée régulière) de (25) avec des données initiales très régulières. On note  $\bar{u}_h$  la fonction de  $L^2_h$  définie à partir des valeurs de  $u$  aux noeuds  $x_j$ :

$$\bar{u}_j(t) = u(x_j, t)$$

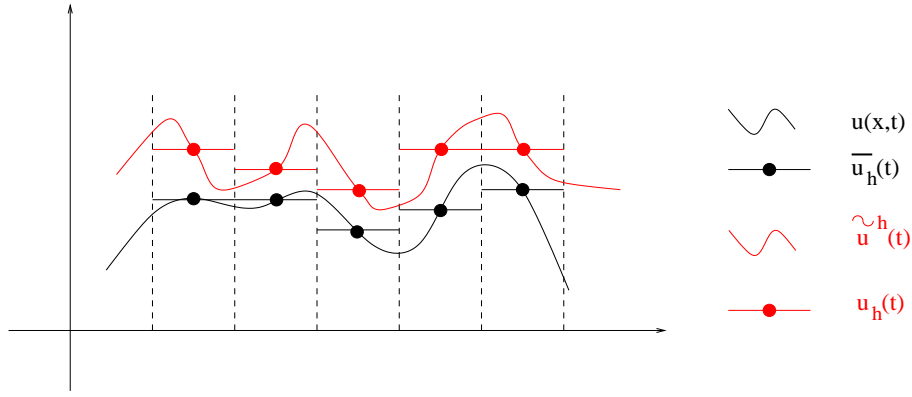


Figure 8: Convergence du schéma semi-discrétisé. Convergence par l'énergie: on compare  $\bar{u}_h$  et  $u_h$  dans  $L_h^2$ . Convergence par Fourier: on compare  $u$  et  $\tilde{u}^h$  dans  $L^2$ .

et  $u_h$  la solution approchée, donnée par le schéma (33).

1- Montrer que  $\bar{u}_h$  vérifie le schéma suivant:

$$(45) \quad \begin{cases} \frac{d^2 \bar{u}_h}{dt^2} + A_h \bar{u}_h = \varepsilon_h & (i) \\ \bar{u}_j(0) = u_0(x_j) & (ii) \\ \frac{d\bar{u}_j}{dt}(0) = u_1(x_j) & (iii) \end{cases}$$

On explicitera  $\varepsilon_h$  appelé erreur de consistance du schéma. De quel ordre est cette erreur? En utilisant un développement de Taylor avec reste intégral, on pourra montrer que

$$|\varepsilon_j(t)| \leq Ch \int_{x_{j-1}}^{x_{j+1}} \left| \frac{\partial^4 u}{\partial x^4} \right|(\xi, t) d\xi$$

et en déduire que

$$(46) \quad \|\varepsilon_h(t)\| \leq Ch^2 \left\| \frac{\partial^4 u}{\partial x^4}(t) \right\|_{L^2}$$

2- Pour évaluer l'erreur d'approximation du schéma sur la solution on introduit l'erreur de convergence:

$$e_h(t) = \bar{u}_h - u_h \in L_h^2$$

Ecrire le système vérifié par  $e_h$ . On note  $E_h$  l'énergie correspondante:

$$E_h(t) = \frac{1}{2} \left\| \frac{de_h}{dt} \right\|^2 + \frac{1}{2} a_h(e_h, e_h)$$

Montrer qu'elle vérifie l'estimation suivante:

$$(47) \quad (E_h(t))^{1/2} \leq (E_h(0))^{1/2} + C \int_0^t \|\varepsilon_h\|(s) ds$$

3- Pour en déduire une estimation de l'erreur, on utilise l'identité:

$$e_h(t) = e_h(0) + \int_0^t \frac{de_h}{dt}(s) ds$$

Montrer que l'erreur est majorée par:

$$(48) \quad \|e_h(t)\| \leq C \left( \|e_h(0)\| + t \left\| \frac{de_h}{dt}(0) \right\| + t \sqrt{a_h(e_h(0), e_h(0))} + t \int_0^t \|\varepsilon_h\|(s) ds \right)$$

En déduire en particulier que si les conditions initiales sont approchées par (28), la convergence est d'ordre 2 et:

$$(49) \quad \|e_h(t)\| \leq C t h^2 \left\| \frac{\partial^4 u}{\partial x^4} \right\|_{L^1(0,T;L^2(\mathbb{R}))}$$

### 3.5.2 Convergence du schéma semi-discrétisé par Fourier

Rappelons que la solution exacte vérifie le problème

$$\begin{cases} \frac{\partial^2 u}{\partial t^2} - c^2 \frac{\partial^2 u}{\partial x^2} = 0 \\ u(x, 0) = u_0(x) \\ \frac{\partial u}{\partial t}(x, 0) = u_1(x) \end{cases}$$

et peut s'exprimer en Fourier :

$$\hat{u}(k, t) = \hat{u}_0(k) \cos(ckt) + \hat{u}_1(k) \frac{\sin(ckt)}{ck}$$

La solution  $u_h$  du schéma semi-discrétisé vérifie:

$$\begin{cases} \frac{d^2 u_h}{dt^2} + A_h u_h = 0 \\ u_h(0) = u_{0,h} \\ \frac{du_h}{dt}(0) = u_{1,h} \end{cases}$$

et peut s'exprimer en Fourier discret:

$$\hat{u}_h(k, t) = \hat{u}_{0,h}(k) \cos(\omega_h t) + \hat{u}_{1,h}(k) \frac{\sin(\omega_h t)}{\omega_h}$$

On introduit  $\tilde{u}^h \in L^2(\mathbb{R})$  qui vérifie

$$\begin{cases} \frac{\partial^2 \tilde{u}^h}{\partial t^2} + \tilde{A}_h \tilde{u}^h = 0 \\ \tilde{u}^h(0) = u_0 \\ \frac{\partial \tilde{u}^h}{\partial t}(0) = u_1 \end{cases}$$

avec

$$\tilde{A}_h \tilde{u}^h(x) = -c^2 \frac{\tilde{u}^h(x+h) - 2\tilde{u}^h(x) + \tilde{u}^h(x-h)}{h^2}$$

Si on approche les conditions initiales par les valeurs ponctuelles:

$$u_{0,j} = u_0(x_j), \quad u_{1,j} = u_1(x_j),$$

c'est à dire qu'on suppose qu'on ne commet pas d'erreur sur les conditions initiales aux points  $x_j$ , alors il est facile de voir que la solution donnée par le schéma coïncide avec les valeurs de  $\tilde{u}^h$  aux noeuds  $x_j$ :

$$u_j(t) \equiv \tilde{u}^h(x_j, t)$$

La fonction constante par morceaux  $u_h(t)$  et la fonction  $\tilde{u}^h(t)$  sont alors deux interpolations différentes de la suite  $u_h(t)$ .

L'opérateur  $\tilde{A}_h$  est une combinaison de translations et il est facile de voir que sa transformée de Fourier est un opérateur de multiplication:

$$\mathcal{F}(\tilde{A}_h)(k) = -\frac{c^2}{h^2}(e^{ikh} - 2 + e^{-ikh}) = \frac{4c^2}{h^2} \sin^2 \frac{kh}{2} \equiv \omega_h(k)^2 = \mathcal{F}_h(A_h)(k)$$

On en déduit aisément l'expression de la transformée de Fourier de  $\tilde{u}^h(t)$ :

$$\mathcal{F}(\tilde{u}^h)(k, t) = \hat{u}_0(k) \cos(\omega_h t) + \hat{u}_1(k) \frac{\sin(\omega_h t)}{\omega_h}$$

Posons

$$\mathcal{F}(e_h)(k, t) = \mathcal{F}(\tilde{u}^h)(k, t) - \mathcal{F}(u)(k, t)$$

D'après ce qui précède, cette erreur est composée de deux termes:

$$\mathcal{F}(e_h)(k, t) = \mathcal{F}(e_h^0)(k, t) + \mathcal{F}(e_h^1)(k, t)$$

avec

$$\mathcal{F}(e_h^0)(k, t) = (\cos(\omega_h(k)t) - \cos(ckt))\hat{u}_0(k)$$

$$\mathcal{F}(e_h^1)(k, t) = \left( \frac{\sin(\omega_h(k)t)}{\omega_h(k)} - \frac{\sin(ckt)}{ck} \right) \hat{u}_1(k)$$

Il nous faut estimer ces deux erreurs en norme  $L^2$  pour obtenir une estimation de

$$\|e_h(t)\|_{L^2}^2 = \|\tilde{u}^h(t) - u(t)\|_{L^2}^2 = \int_{\mathbf{R}} |\mathcal{F}(\tilde{u}^h)(k, t) - \mathcal{F}(u)(k, t)|^2 dk$$

la dernière égalité provenant de Plancherel.

**Estimation de  $e_h^0$ .**

$$(50) \quad \|e_h^0(t)\|_{L^2}^2 = \int_{\mathbf{R}} |\cos(\omega_h(k)t) - \cos(ckt)|^2 |\hat{u}_0(k)|^2 dk$$

En utilisant la formule des accroissements finis:

$$\frac{f(x) - f(y)}{x - y} = f'(y_0), \quad y_0 \in ]x, y[$$

appliquée à  $f(x) = \cos x$ , on obtient:

$$\left| \frac{\cos(\omega_h(k)t) - \cos(ckt)}{\omega_h(k)t - ckt} \right| = |\sin y_0| \quad y_0 \in ]\omega_h(k)t, ckt[$$

d'où

$$(51) \quad |\cos(\omega_h(k)t) - \cos(ckt)|^2 \leq (\omega_h(k)^2 - c^2k^2)t^2$$

Or

$$|\omega_h(k) - ck| = \left| ck \left( \frac{\sin x}{x} - 1 \right) \right|, \quad \text{avec } x = \frac{kh}{2}$$

Pour estimer cette quantité on utilise le développement:

$$\sin x = x - \frac{x^3}{3!} \cos y_0, \quad y_0 \in ]0, x[$$

ce qui montre que

$$\left| \frac{\sin x}{x} - 1 \right| = \left| \frac{x^2}{3!} \cos y_0 \right| \leq \frac{x^2}{6}$$

et par conséquent:

$$(52) \quad |\omega_h(k) - ck| \leq |ck| \frac{k^2 h^2}{24} = \left| \frac{ck^3 h^2}{24} \right|$$

En reportant dans (51), on en déduit que

$$(53) \quad |\cos(\omega_h(k)t) - \cos(ckt)|^2 \leq (\omega_h(k)^2 - c^2k^2)t^2 \leq \frac{c^2 k^6 h^4 t^2}{24^2}$$

et finalement dans (50):

$$\|e_h^0(t)\|_{L^2}^2 \leq \frac{c^2 h^4 t^2}{24^2} \int_{\mathbf{R}} k^6 |\hat{u}_0(k)|^2 dk = \frac{c^2 h^4 t^2}{24^2} \int_{\mathbf{R}} \left| \frac{d^3 u_0}{dx^3} \right|^2 dx = \frac{c^2 h^4 t^2}{24^2} \left\| \frac{d^3 u_0}{dx^3} \right\|_{L^2}^2$$

**Estimation de  $e_h^1$ .**

$$(54) \quad \|e_h^1(t)\|_{L^2}^2 = \int_{\mathbf{R}} \left| \frac{\sin(\omega_h(k)t)}{\omega_h(k)} - \frac{\sin(ckt)}{ck} \right|^2 |\hat{u}_1(k)|^2 dk$$

On décompose:

$$\frac{\sin(\omega_h(k)t)}{\omega_h(k)} - \frac{\sin(ckt)}{ck} = \sin(\omega_h(k)t) \left( \frac{1}{\omega_h(k)} - \frac{1}{ck} \right) + \frac{1}{ck} (\sin(\omega_h(k)t) - \sin(ckt))$$

Le premier terme:

$$\left| \sin(\omega_h(k)t) \left( \frac{1}{\omega_h(k)} - \frac{1}{ck} \right) \right| \leq \underbrace{|\sin(\omega_h(k)t)|}_{\leq |\omega_h(k)t|} \frac{|\omega_h(k) - ck|}{|\omega_h(k)ck|} \leq t \frac{|\omega_h(k) - ck|}{|ck|}$$

est majoré, en utilisant (52), par:

$$\left| \sin(\omega_h(k)t) \left( \frac{1}{\omega_h(k)} - \frac{1}{ck} \right) \right| \leq t \frac{\left| \frac{ck^3 h^2}{24} \right|}{|ck|} \leq t \frac{k^2 h^2}{24}$$

Pour le deuxième terme, on utilise de nouveau la formule des accroissements finis qui nous donne:

$$\frac{\sin x - \sin y}{x - y} = \cos y_0 \quad y_0 \in ]x, y[$$

d'où

$$\left| \frac{1}{ck} \right| \left| \sin(\omega_h(k)t) - \sin(ckt) \right| \leq \frac{1}{|ck|} |\omega_h(k)t - ckt| \leq \frac{1}{|ck|} t \left| \frac{ck^3 h^2}{24} \right| \leq t \frac{k^2 h^2}{24}$$

Finalement on obtient donc:

$$\left| \frac{\sin(\omega_h(k)t)}{\omega_h(k)} - \frac{\sin(ckt)}{ck} \right| \leq t \frac{k^2 h^2}{12}$$

et dans (54):

$$\left\| e_h^1(t) \right\|_{L^2}^2 \leq t^2 \frac{h^4}{12^2} \int_{\mathbf{R}} k^4 |\widehat{u}_1(k)|^2 dk = t^2 \frac{h^4}{12^2} \left\| \frac{d^2 u_1}{dx^2} \right\|_{L^2}^2$$

On obtient finalement le:

**Théorème 6** Sous l'hypothèse  $(u_0, u_1) \in H^3(\mathbf{R}) \times H^2(\mathbf{R})$  on a l'estimation d'erreur:  $\forall t \in [0, T]$

$$(55) \quad \left\| \widetilde{u}^h(t) - u(t) \right\|_{L^2} \leq C t h^2 \left( \left\| \frac{d^3 u_0}{dx^3} \right\|_{L^2} + \left\| \frac{d^2 u_1}{dx^2} \right\|_{L^2} \right)$$



## 4 Discrétisation totale par un schéma aux différences finies en espace et en temps

### 4.1 Schéma explicite centré

On introduit un découpage en temps:  $\Delta t$  est le pas de temps,  $t^n = n\Delta t$ . On note  $u_j^n$  une approximation de  $u(x_j, t^n)$ . Pour obtenir un schéma totalement discrétisé, on approche la dérivée en temps  $\frac{\partial^2 u}{\partial t^2}(x_j, t^n)$  par une différence finie centrée d'ordre 2, ce qui conduit au schéma suivant:

$$(56) \quad \frac{u_j^{n+1} - 2u_j^n + u_j^{n-1}}{\Delta t^2} - c^2 \frac{u_{j+1}^n - 2u_j^n + u_{j-1}^n}{\Delta x^2} = 0$$

ou encore:

$$(57) \quad \boxed{\frac{u_h^{n+1} - 2u_h^n + u_h^{n-1}}{\Delta t^2} + A_h u_h^n = 0}$$

avec  $(A_h u_h)_j^n = -c^2 \frac{u_{j+1}^n - 2u_j^n + u_{j-1}^n}{\Delta x^2}$ . C'est un schéma explicite:

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

Schéma centré au point  $\times$  - Calcul au point  $\circ$

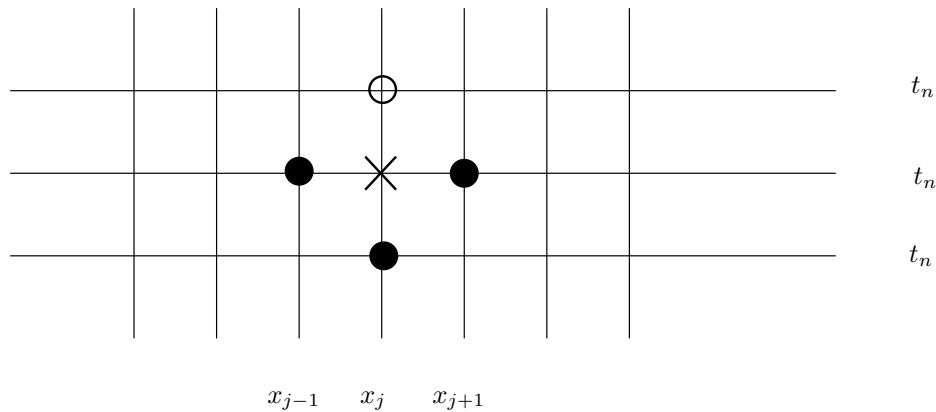


Figure 9: Schéma explicite centré

#### Démarrage: Choix des conditions initiales approchées

Pour le problème continu, comme pour le problème semi-discrétisé, les données initiales étaient définies par les valeurs de la solution à  $t = 0$ , ainsi que celles de sa dérivée en temps à  $t = 0$ :

Problème continu $u(x, 0) = u_0(x)$ $\frac{\partial u}{\partial t} = u_1(x)$	Problème semi-discrétisé en espace $u_h(0) = u_{0,h}$ $\frac{du_h}{dt}(0) = u_{1,h}$
--	--

où on a noté  $u_{\alpha,h}$  des approximations (en espace) de  $u_\alpha$ .

Pour le problème totalement discrétisé,  $u_h^n = (u_j^n)_{j \in \mathbb{Z}}$  est une approximation de  $u(x_j, t^n)$  et on a besoin de  $u_h^0$  et de  $u_h^1$  pour pouvoir démarrer le schéma. On doit donc approcher  $u(x_j, 0) = u_0(x_j)$  et  $u(x_j, \Delta t)$ . Pour le premier instant il est naturel de choisir, comme pour le problème semi-discrétisé:

$$(58) \quad \boxed{u_h^0 = u_{0,h}}$$

Pour approcher  $u(x_j, \Delta t)$ , on peut utiliser la formule de Taylor:

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

ce qui conduit à l'approximation:

$$(59) \quad u_h^1 = u_{0,h} + \Delta t u_{1,h}$$

Ce choix correspond alors à une approximation d'ordre 1 en temps de  $\partial u / \partial t$ .

Pour avoir de l'ordre 2 en temps (plus cohérent avec le schéma à l'intérieur, comme on le verra plus loin) on pousse plus loin le développement:

$$u(x_j, \Delta t) = u_0(x_j) + \Delta t u_1(x_j) + \frac{\Delta t^2}{2} \frac{\partial^2 u}{\partial t^2}(x_j, 0) + O(\Delta t^3)$$

et on utilise l'équation:  $\frac{\partial^2 u}{\partial t^2} = c^2 \frac{\partial^2 u}{\partial x^2}$ , d'où

$$u(x_j, \Delta t) = u_0(x_j) + \Delta t u_1(x_j) + \frac{c^2 \Delta t^2}{2} \frac{\partial^2 u}{\partial x^2}(x_j, 0) + O(\Delta t^3)$$

Ce choix conduit cette fois à l'approximation:

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

qui s'écrit encore:

$$(60) \quad \boxed{u_h^1 = \left(I - \frac{\Delta t^2}{2} A_h\right) u_h^0 + \Delta t u_{1,h}}$$

C'est ce dernier choix que nous adopterons dans la suite du cours (sauf mention du contraire): le schéma est donc défini par (57), (58) et (60).

## 4.2 Ordre du schéma: consistance et erreur de troncature

On se donne une solution  $u$  régulière de l'équation des ondes

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

On notera par  $\bar{u}_j^n = u(x_j, t^n)$ , pour distinguer des valeurs de la suite définie par le schéma  $u_j^n$ , et on définit l'erreur de troncature:

$$(61) \quad \varepsilon_j^n = \frac{\bar{u}_j^{n+1} - 2\bar{u}_j^n + \bar{u}_j^{n-1}}{\Delta t^2} + (A_h \bar{u}_h)_j^n$$

**Définition 1** On dit que le schéma est consistant si l'erreur de troncature tend vers 0 ( $\varepsilon_j^n \rightarrow 0$ ) lorsque  $\Delta t$  et  $\Delta x$  tendent vers 0.

Le schéma est consistant à l'ordre  $m$  en temps et  $k$  en espace si l'erreur de troncature est d'ordre  $m$  en temps et  $k$  en espace, i.e. si

$$\varepsilon_j^n = O(\Delta t^m + \Delta x^k),$$

lorsque  $\Delta t$  et  $\Delta x$  tendent vers 0.

**Remarque 4** L'erreur de troncature est une erreur qui indique comment l'équation est approchée par le schéma. Ce n'est pas une erreur entre la solution exacte et la solution approchée (erreur de convergence), mais c'est une erreur qui quantifie à quel ordre la solution exacte vérifie le schéma.

**Lemme 2** Le schéma (57) est consistant à l'ordre 2 en temps et en espace :

$$\varepsilon_j^n = O(\Delta t^2 + \Delta x^2)$$

**Preuve:** Si  $u$  est régulière, on peut faire des développements de Taylor:

$$\frac{\bar{u}_j^{n+1} - 2\bar{u}_j^n + \bar{u}_j^{n-1}}{\Delta t^2} = \frac{\partial^2 u}{\partial t^2}(x_j, t^n) + \frac{2}{4!} \Delta t^2 \frac{\partial^4 u}{\partial t^4}(x_j, t^n) + O(\Delta t^4)$$

$$\frac{\bar{u}_{j+1}^n - 2\bar{u}_j^n + \bar{u}_{j-1}^n}{\Delta x^2} = \frac{\partial^2 u}{\partial x^2}(x_j, t^n) + \frac{2}{4!} \Delta x^2 \frac{\partial^4 u}{\partial x^4}(x_j, t^n) + O(\Delta x^4)$$

d'où

$$\begin{aligned} \varepsilon_j^n &= \underbrace{\frac{\partial^2 u}{\partial t^2}(x_j, t^n) - c^2 \frac{\partial^2 u}{\partial x^2}(x_j, t^n)}_{=0} \\ &+ \frac{2}{4!} \left( \Delta t^2 \frac{\partial^4 u}{\partial t^4}(x_j, t^n) - c^2 \Delta x^2 \frac{\partial^4 u}{\partial x^4}(x_j, t^n) \right) \\ &+ O(\Delta t^4 + \Delta x^4) \end{aligned}$$

le premier terme est nul car  $u$  vérifie l'équation des ondes. On en déduit aussitôt que  $\varepsilon_j^n = O(\Delta t^2 + \Delta x^2)$ . Si on examine de plus près le second terme, en réutilisant l'équation des ondes, on a :

$$\frac{\partial^4 u}{\partial t^4} = c^2 \frac{\partial^4 u}{\partial x^2 \partial t^2} = c^4 \frac{\partial^4 u}{\partial x^4}$$

donc le second terme est égal à:

$$\frac{2}{4!} (\Delta t^2 c^4 - c^2 \Delta x^2) \frac{\partial^4 u}{\partial x^4}(x_j, t^n) = \frac{2}{4!} c^2 \Delta x^2 (\alpha^2 - 1) \frac{\partial^4 u}{\partial x^4}(x_j, t^n)$$

avec  $\alpha = \frac{c\Delta t}{\Delta x}$ , ce qui montre qu'on gagne un ordre lorsque  $\alpha = 1$ . On a donc un schéma d'ordre 2 exactement lorsque  $\alpha \neq 1$  et un schéma d'ordre au moins 4 lorsque  $\alpha = 1$ . En fait on peut montrer que lorsque  $\alpha = 1$  le schéma est d'ordre infini, c'est à dire que  $\varepsilon_j^n = 0$  et que le schéma donne la solution exacte. On va donner une interprétation de ce résultat à partir du cône de dépendance.

**Exercice 9** 1- Montrer que l'erreur de troncature est nulle (en 1D) si  $\alpha = 1$  (on utilisera l'expression de la solution exacte  $u(x, t) = f(x + ct) + g(x - ct)$ ).

2- Ecrire un schéma de discrétisation totale de l'équation des ondes en 2D, centré et d'ordre 2.

3- Montrer que pour ce schéma, et en supposant que  $\Delta x = \Delta y = h$ , l'erreur de troncature reste d'ordre 2 même si  $c\Delta t = h$ , contrairement au cas 1D.

### 4.3 Cône de dépendance numérique- Condition nécessaire de convergence

Bien sûr, le fait que le schéma approche l'équation ne suffit pas à garantir que la solution discrète  $u_h^n$  converge vers la solution exacte, lorsque les paramètres de discrétisation  $h$  et  $\Delta t$  tendent vers 0. Nous donnons ici une condition nécessaire de convergence. On note toujours  $\alpha = \frac{c\Delta t}{\Delta x}$ , et on définit la vitesse numérique:

$$V_{num} = \frac{\Delta x}{\Delta t}$$

Nous regardons la convergence lorsque  $h$  et  $\Delta t$  tendent vers 0, avec  $\alpha$  fixé.

#### Premier point de vue(cf Fig 10)

La solution exacte  $u$  issue des conditions initiales  $u_0$  et  $u_1$  est comprise dans le cône situé entre les deux droites "limites" de pente  $\pm 1/c$ . La solution discrète  $u_h$  se propage sur les points de la grille, avec une vitesse  $V_{num}$  (en gagnant un point à chaque itération). Si la pente  $\frac{1}{V_{num}} > \frac{1}{c}$ , alors  $u_h = 0$  dans la zone comprise entre les 2 cônes, alors que  $u \neq 0$ , donc il ne peut pas y avoir convergence. Une **condition nécessaire de convergence** est donc que:

$$(62) \quad |\alpha| \leq 1 \iff \left| \frac{c\Delta t}{\Delta x} \right| \leq 1$$

Nous verrons que cette condition, appelée condition CFL (Courant-Friedrichs-Levy), est également une condition nécessaire de stabilité du schéma.

#### Deuxième point de vue(cf Fig 11)

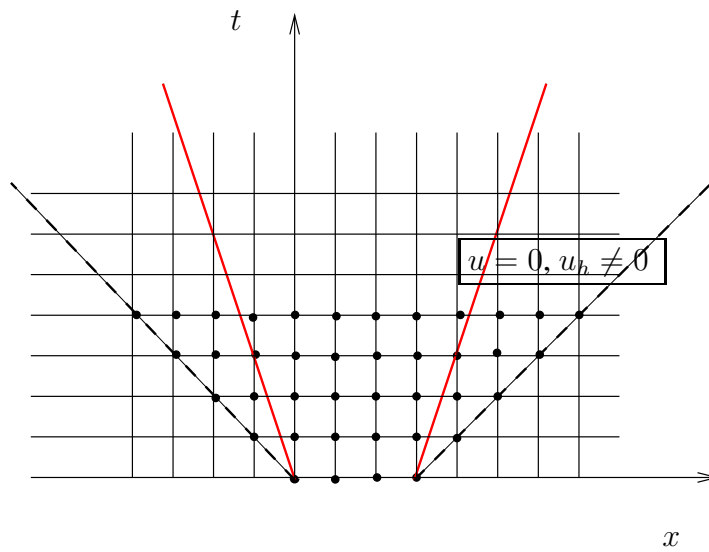
Soit  $M = (x_j, t^n)$  un point de la grille de calcul. Les points qui ont servi à calculer  $u_j^n$  sont contenus dans un cône de sommet  $M$  d'arête les demi droites issues de  $M$  et de pente  $\pm \Delta x / \Delta t$ . On le note  $K_\alpha^-(M)$ . La solution  $u_j^n$  dépend donc uniquement des valeurs de  $u^0$  et  $u^1$  sur le segment  $[M_{0,\alpha}^-, M_{0,\alpha}^+] = [x_{j-n}, x_{j+n}] = [(j-n)h, (j+n)h]$ . La solution exacte, elle dépend des valeurs de  $u^0$  et  $u^1$  entre  $[M^-, M^+]$ , avec  $M^- = (x_j - ct^n, 0)$  et  $M^+ = (x_j + ct^n, 0)$ . On a donc 2 cas:

**Cas 1:**  $[M^-, M^+] \subset [M_{0,\alpha}^-, M_{0,\alpha}^+]$ , i.e., le cône de dépendance numérique contient le cône de dépendance exact et on a donc toutes les infos nécessaires pour construire la solution approchée. Ce cas est le cas où  $(j-n)h \leq x_j - ct^n \leq x_j + ct^n \leq (j+n)h$  c'est à dire  $c\Delta t \leq h$ .

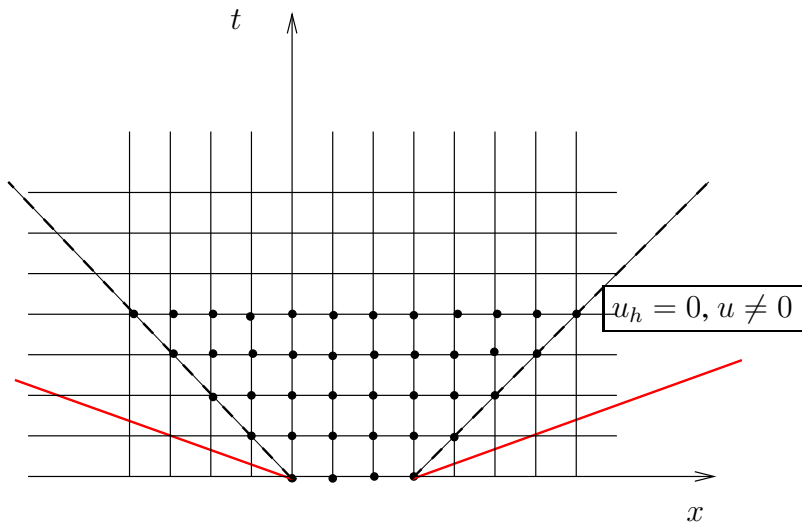
**Cas 2:**  $[M_{0,\alpha}^-, M_{0,\alpha}^+] \subset [M^-, M^+]$ , i.e., le cône de dépendance numérique est strictement contenu dans le cône de dépendance exact et la solution approchée ne tient pas compte des valeurs de la solution (non nulle) à l'extérieur du cône numérique. Il ne peut donc pas y avoir convergence.

### 4.4 Analyse de stabilité par Fourier

Avec la notion de consistance définie précédemment, la deuxième notion fondamentale associée à un schéma numérique est la stabilité. On dira que le schéma est stable si la solution discrète à l'instant  $t^n$ ,  $u_h^n$ , est bornée indépendamment des paramètres de discrétisation  $h$  et  $\Delta t$ , pour toutes conditions initiales. La stabilité est une condition nécessaire de convergence. En fait, on peut montrer que si le schéma est consistant, alors c'est aussi une condition suffisante (si un schéma est stable et consistant alors il est convergent).



Cas 1 :  $1/c > 1/V_{num}$  peut converger



Cas 2 :  $1/c < 1/V_{num}$  : ne peut pas converger

Figure 10: Condition nécessaire de convergence:  $1/V_{num} \leq 1/c \iff \alpha = \frac{c\Delta t}{\Delta x} \leq 1$

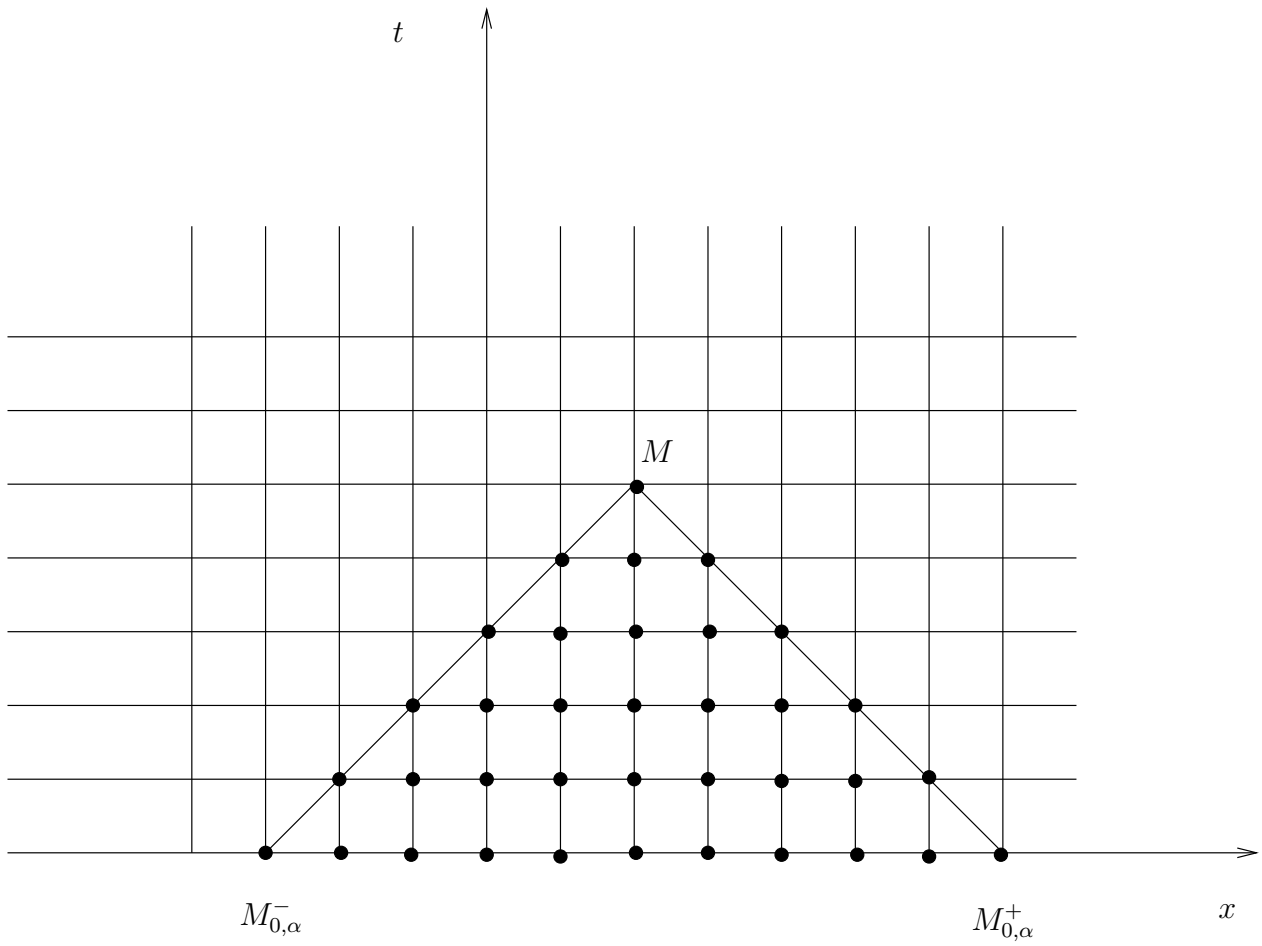


Figure 11: Convergence

On suppose ici que  $u_0$  et  $u_1$  sont données dans  $L^2$  et que leurs approximations  $u_{0h}$  et  $u_{1h}$  vérifient les hypothèses (39). On cherche à obtenir des estimations de la solution discrète, indépendantes de  $h, \Delta t$ , pour toutes conditions initiales. Le schéma peut s'écrire

$$(63) \quad \frac{u_h^{n+1} - 2u_h^n + u_h^{n-1}}{\Delta t^2} + (A_h u_h)^n = 0$$

Comme pour le schéma semi-discretisé en espace, on applique à (63) une transformée de Fourier discrète d'où

$$(64) \quad \frac{\hat{u}_h^{n+1} - 2\hat{u}_h^n + \hat{u}_h^{n-1}}{\Delta t^2} + D_h(k)\hat{u}_h^n = 0$$

avec  $D_h(k) = \hat{A}_h(k) = \frac{4c^2}{h^2} \sin^2 \frac{kh}{2}$  ce qui est équivalent à

$$\hat{u}_h^{n+1} - 2(1 - 2\alpha^2 \sin^2 \frac{kh}{2})\hat{u}_h^n + \hat{u}_h^{n-1} = 0$$

Par conséquent, la solution est de la forme:

$$\hat{u}_h^n = A_h r_1(k)^n + B_h r_2(k)^n$$

où  $r_1$  et  $r_2$  sont les racines du polynôme caractéristique:

$$(65) \quad r^2 - 2(1 - \beta)r + 1 = 0$$

avec  $\beta = 2\alpha^2 \sin^2 \frac{kh}{2} = 2\frac{c^2 \Delta t^2}{h^2} \sin^2 \frac{kh}{2}$ . On a en particulier  $0 \leq \beta \leq 2\alpha^2$  et  $D_h(k)\Delta t^2 = 2\beta$ .

Une condition nécessaire de stabilité  $L^2$  est donc  $\max_k (|r_1(k)|, |r_2(k)|) \leq 1$ . Les deux racines vérifient:

$$(66) \quad \begin{cases} r_1 r_2 = 1 \\ r_1 + r_2 = 2(1 - \beta) \end{cases}$$

Les constantes  $A_h$  et  $B_h$  sont déterminées par les conditions initiales:

$$\begin{aligned} \hat{u}_h^0 &= A_h + B_h \\ \hat{u}_h^1 &= A_h r_1(k) + B_h r_2(k) \end{aligned} \iff \begin{aligned} A_h(r_2 - r_1) &= \hat{u}_h^0 r_2 - \hat{u}_h^1 \\ B_h(r_2 - r_1) &= \hat{u}_h^1 - r_1 \hat{u}_h^0 \end{aligned}$$

ce qui donne si  $r_1 \neq r_2$  l'expression de la solution

$$(67) \quad \hat{u}_h^n = \frac{\hat{u}_h^0 r_2 - \hat{u}_h^1}{r_2 - r_1} r_1^n + \frac{\hat{u}_h^1 - r_1 \hat{u}_h^0}{r_2 - r_1} r_2^n = \hat{u}_h^0 \frac{r_2 r_1^n - r_1 r_2^n}{r_2 - r_1} + \hat{u}_h^1 \frac{r_2^n - r_1^n}{r_2 - r_1}$$

Le discriminant de (65) est :

$$\Delta(k) = (1 - \beta)^2 - 1 = - \underbrace{\beta}_{\geq 0} (2 - \beta)$$

On a donc 2 cas à considérer:

**Cas 1:**  $\exists k, tq. \Delta(k) > 0$

Les 2 racines sont donc réelles distinctes  $(r_1, r_2) \in \mathbb{R}^2$  et d'autre part on a d'après la première relation de (66)

$$|r_1| |r_2| = 1$$

donc dans ce cas on a  $\max(|r_1|, |r_2|) > 1$  ce qui montre que le schéma est instable (il existe des conditions initiales pour lesquelles la solution explose: si par exemple  $R > 1$ ,  $R^n = e^{n \ln R} = e^{\frac{n \Delta t}{\Delta t} \ln R} = e^{\frac{n}{\Delta t} \ln R} \rightarrow +\infty$  lorsque  $\Delta t \rightarrow 0$ ).

**Cas 2:**  $\forall k, \Delta(k) \leq 0$

D'après ce qui précède, une condition nécessaire de stabilité s'écrit donc

$$\Delta(k) \leq 0, \forall k \iff 2 - \beta \geq 0, \forall k \iff |\alpha| \leq 1$$

Dans ce cas, les deux racines sont soit complexes conjuguées soit réelles et identiques. Dans les deux cas:

$$r_2 = \bar{r}_1$$

donc  $|r_1| = |r_2| = 1$  et par conséquent:

$$(68) \quad |\widehat{u}_h^n| \leq |A_h| + |B_h|$$

donc la solution n'explose pas en temps et il n'y a pas non plus d'amortissement ce qui est conforme au phénomène étudié et traduit la conservation d'énergie. Pour montrer qu'on a stabilité  $L^2$ , il reste à vérifier que la solution est bornée indépendamment de  $h$ .

**Estimation de la solution sous l'hypothèse  $\alpha \leq 1$**

Pour obtenir une estimation de la solution, (68) ne permet pas de conclure. On va reprendre l'expression de la solution en fonction des conditions initiales. Les étapes sont les suivantes:

On rappelle le choix des conditions initiales (cohérentes avec de l'ordre 2 en temps) (58)-(60):

$$u_h^0 = u_{0,h}$$

$$u_h^1 = (I - \frac{\Delta t^2}{2} A_h) u_h^0 + \Delta t u_{1,h}$$

ce qui montre que

$$\widehat{u}_h^0 = \widehat{u}_{0,h}$$

$$\widehat{u}_h^1 = (1 - \frac{\Delta t^2}{2} D_h(k)) \widehat{u}_h^0 + \Delta t \widehat{u}_{1,h} = (1 - \beta) \widehat{u}_h^0 + \Delta t \widehat{u}_{1,h} = \frac{r_1 + r_2}{2} \widehat{u}_h^0 + \Delta t \widehat{u}_{1,h}$$

On injecte ces expressions dans l'expression de la solution (67), ce qui donne:

$$\widehat{u}_h^n = \frac{1}{r_2 - r_1} \left( \frac{1}{2} (r_2 r_1^n - r_1 r_2^n + r_2^{n+1} - r_1^{n+1}) \widehat{u}_h^0 + \Delta t (r_2^n - r_1^n) \widehat{u}_{1,h} \right)$$

En utilisant la relation  $r_1 r_2 = 1$  et en posant  $r_2 = r$  et  $r_1 = 1/r$  on obtient:

$$\widehat{u}_h^n = \frac{1}{2} \frac{r^{2n} + 1}{r^n} \widehat{u}_h^0 + \Delta t \frac{r^{2n} - 1}{r^2 - 1} \frac{1}{r^{n-1}} \widehat{u}_{1,h}$$

ce qui implique ( $|r| = 1$ )

$$|\widehat{u}_h^n| \leq \frac{1}{2} \left| r^n + \frac{1}{r^n} \right| |\widehat{u}_h^0| + \Delta t \left| \frac{r^{2n} - 1}{r^2 - 1} \right| |\widehat{u}_{1,h}|$$



**Lemme 3** Si  $|r| = 1$ , alors

$$\frac{1}{2} \left| r^n + \frac{1}{r^n} \right| \leq 1, \text{ et } \left| \frac{r^{2n} - 1}{r^2 - 1} \right| \leq n$$

**Preuve:** La première estimation est immédiate:

$$\frac{1}{2} \left| r^n + \frac{1}{r^n} \right| \leq \frac{1}{2} |r|^n + \left| \frac{1}{r} \right|^n = \frac{1}{2}(1 + 1) = 1$$

La deuxième estimation découle de l'identité:

$$\left| \frac{1 - (r^2)^n}{1 - r^2} \right| = \left| 1 + r^2 + \dots + (r^2)^{n-1} \right| \leq 1 + \dots + 1 = n \blacksquare$$

A partir du lemme, il est immédiat d'obtenir une estimation de la solution :

$$\begin{aligned} |\widehat{u}_h^n(k)| &\leq \left| \widehat{u}_h^0 \right| + \Delta t n |\widehat{u}_{1,h}| \\ \implies |\widehat{u}_h^n(k)|^2 &\leq 2 \left| \widehat{u}_h^0 \right|^2 + 2(t^n)^2 |\widehat{u}_{1,h}|^2 \\ \implies \|\widehat{u}_h^n\|_{L^2}^2 &\leq 2 \|\widehat{u}_h^0\|_{L^2}^2 + 2(t^n)^2 \|\widehat{u}_{1,h}\|_{L^2}^2 \\ \implies \|u_h^n\|_{L^2}^2 &\leq 2 \|u_h^0\|_{L^2}^2 + 2(t^n)^2 \|u_{1,h}\|_{L^2}^2 \end{aligned}$$

ce qui implique encore, en utilisant d'une part  $\sqrt{a^2 + b^2} \leq |a| + |b|$  et d'autre part les hypothèses (39):

$$\|u_h^n\|_{L^2} \leq C(\|u_0\|_{L^2} + t^n \|u_1\|_{L^2})$$

qui est l'analogue de l'estimation (40) du cas semi-discret. On résume ces résultats dans la

**Proposition 1** Soient  $u_0$  et  $u_1$  données dans  $L^2$  et leurs approximations  $u_{0h} \in L_h^2$  et  $u_{1h} \in L_h^2$  vérifiant les hypothèses (39). Alors, sous la condition CFL:  $\left| \frac{c\Delta t}{h} \right| \leq 1$ , le schéma totalement discrétisé admet une solution unique et on a l'estimation:

$$(69) \quad \|u_h^n\|_{L^2} \leq C(\|u_0\|_{L^2} + t^n \|u_1\|_{L^2})$$

où  $C$  est une constante positive indépendante de  $h$  et  $\Delta t$ , ce qui montre la stabilité sur un intervalle de temps borné.

## 4.5 Analyse de dispersion

Comme pour le schéma semi-discrétisé, on peut faire une analyse par ondes planes. Ce paragraphe suit la même démarche que celle présentée pour schéma semi-discrétisé (le lecteur peut essayer de le faire comme exercice).

### Ondes planes continues

On rappelle que l'analyse par ondes planes du problème continu consiste à s'intéresser aux solutions particulières de la forme  $u(x, t) = e^{i(-\omega t + kx)}$ , ce qui conduit à la relation de dispersion  $\omega^2 = c^2 k^2$ . La quantité  $\omega/k = \pm c$  est la vitesse de phase de l'onde.

## Ondes planes numériques

Ce sont les solutions particulières de la forme

$$u_j^n = e^{i(j\Delta x k - \omega_{h,\Delta t} n \Delta t)}$$

En injectant cette expression dans le schéma, on obtient la relation de dispersion du schéma:

$$\begin{aligned} \frac{2 \cos \omega_{h,\Delta t} \Delta t - 2}{\Delta t^2} - c^2 \frac{2 \cos \Delta x - 2}{\Delta x^2} &= 0 \\ \iff \frac{4}{\Delta t^2} \sin^2 \frac{\omega_{h,\Delta t} \Delta t}{2} &= \frac{4c^2}{h^2} \sin^2 \frac{kh}{2} \equiv D_h(k) \end{aligned}$$

Pour le schéma semi-discretisé, rappelons qu'on avait obtenu comme relation (en cherchant des solutions de la forme  $u_j(t) = e^{i(jhk - \omega_h t)}$ )

$$\omega_h^2 = D_h(k) = \frac{4c^2}{h^2} \sin^2 \frac{kh}{2}$$

le deuxième terme  $D_h(k)$  provenant de la discrétisation en espace et le premier terme étant le symbole de l'opérateur  $\partial_t^2$  qui n'a pas été discrétisé, et on avait  $\omega_h^2 = k^2 c^2 + O(h^2)$ . L'erreur de dispersion (rapport des vitesses continue et semi-discretisée) s'exprimait alors:

$$\frac{\omega_h^\pm}{k} \frac{1}{c} = \pm \frac{2}{kh} \sin \frac{kh}{2}$$

Pour le schéma totalement discrétisé, ce premier terme est remplacé par le symbole de l'opérateur discret approchant  $\partial_t^2$ :

$$(70) \quad \frac{4}{\Delta t^2} \sin^2 \frac{\omega_{h,\Delta t} \Delta t}{2} = D_h(k)$$

et on a  $\omega_{h,\Delta t} = k^2 c^2 + O(h^2 + \Delta t^2)$ . On peut réécrire (70) sous la forme

$$(71) \quad \omega_{h,\Delta t} = \pm \frac{2}{\Delta t} \text{Arcsin} \left( \frac{c\Delta t}{h} \sin \frac{kh}{2} \right) \equiv \pm \frac{2}{\Delta t} \text{Arcsin} \left( \alpha \sin \frac{kh}{2} \right)$$

ou encore, en posant  $K = kh$ ,

$$(72) \quad \frac{1}{c} \frac{\omega_{h,\Delta t}}{k} = \pm \frac{2}{\alpha K} \text{Arcsin} \left( \alpha \sin \frac{K}{2} \right)$$

Le nombre  $G = \frac{K}{2\pi} = \frac{h}{\lambda}$  représente l'inverse du nombre de points par longueur d'onde  $\lambda = \frac{2\pi}{k}$ .

La quantité  $\frac{1}{c} \frac{\omega_{h,\Delta t}}{k}$  représente le rapport entre la vitesse de phase numérique et la vitesse de phase continue. Notons  $\omega_{h,\Delta t}^+$  la solution telle que  $0 \leq \omega_{h,\Delta t}^+ \Delta t / 2 \leq \pi/2$ , et

$$q(\alpha, G) = \frac{1}{c} \frac{\omega_{h,\Delta t}^+}{k} = \frac{2}{\alpha K} \text{Arcsin} \left( \alpha \sin \frac{K}{2} \right) = \frac{1}{\alpha \pi G} \text{Arcsin}(\alpha \sin \pi G).$$

**Quelques remarques** (cf Figure 12)

- Si  $\alpha$  est fixé et  $h \rightarrow 0$ , on a

$$q(\alpha, G) = 1 - (1 - \alpha^2)K^2/2 + \dots$$

On retrouve que le schéma est d'ordre 2 (et d'ordre infini si  $\alpha = 1$ ).

- Pour  $\alpha$  fixé, la fonction  $G \rightarrow q(\alpha, G)$  est décroissante (+ on a de points par longueur d'onde, plus  $q$  est proche de 1).
- Pour  $G$  fixé, la fonction  $\alpha \rightarrow q(\alpha, G)$  est une fonction croissante. Le meilleur schéma est donc obtenu pour le plus grand  $\alpha$  possible (limité par la CFL) c'est à dire pour  $\alpha = 1$ . En 1D, on a même  $q(1, G) = 1$  et le schéma est exact (faux en 2D, cf exercice 9).
- Sous la condition  $\alpha \leq 1$  on retrouve la stabilité du schéma puisqu'alors  $\omega_{h,\Delta t}^+$  est réel. Si  $\omega$  était complexe, ça voudrait dire qu'il existe des solutions

$$u_j^n = e^{i(jhk - \omega n \Delta t)} = e^{i(jhk - \omega_R n \Delta t)} e^{\omega_I n \Delta t}$$

et on a 2 solutions  $\omega^\pm$  donc une des deux est telle que  $\omega_I > 0$  donc explose. Cette technique est très pratique pour déterminer la condition de stabilité d'un schéma.

- La limite quand  $\alpha \rightarrow 0$  correspond à

$$q(\alpha \rightarrow 0, G) \sim \frac{1}{\pi G} \sin \pi G = \frac{2}{K} \sin \frac{K}{2}$$

qui correspond à l'erreur de dispersion du schéma semi-discrétisé. On voit donc que sa dispersion est plus grande que celle du schéma totalement discrétisé.

**Exercice 10** On peut établir la condition de stabilité en 2D facilement, par le même argument que en 1D (cf. 4.).

1- Montrer que la relation de dispersion du schéma 2D s'écrit

$$(73) \quad \frac{4}{\Delta t^2} \sin^2 \frac{\omega_{h,\Delta t} \Delta t}{2} = \frac{4c^2}{h^2} \left( \sin^2 \frac{k_1 h}{2} + \sin^2 \frac{k_2 h}{2} \right)$$

2- Montrer que le schéma est stable sous la condition CFL

$$(74) \quad |\alpha| = \left| \frac{c \Delta t}{h} \right| \leq \frac{\sqrt{2}}{2}$$

On peut interpréter cette condition de la façon suivante: en 2D, la solution élémentaire (fonction de Green) est:

$$G(x, t) = \frac{H(t - |x|/c)}{2\pi \sqrt{t^2 - |x|^2/c^2}}$$

dont le support est le disque  $\{|x| \leq ct\}$ . A l'instant  $t^n = n\Delta t$ , elle a donc atteint tous les points  $|x| \leq cn\Delta t$ . La solution numérique, elle, se propage sur un losange de demi-diagonale  $nh$  et donc de côté  $l = nh/\sqrt{2}$ . La condition CFL exprime le fait que le losange doit strictement contenir le disque. S'il existait des points du disque qui sont strictement à l'extérieur du losange, cela signifierait que en ces points la solution exacte est non nulle alors que la solution approchée l'est. Cette condition apparaît donc de nouveau comme une condition nécessaire de convergence du schéma. (Le losange contient strictement le disque si  $l \geq cn\Delta t$  donc si  $nh/\sqrt{2} \geq cn\Delta t \iff c\Delta t/h \leq \sqrt{2}/2$ , cf Figure 13).

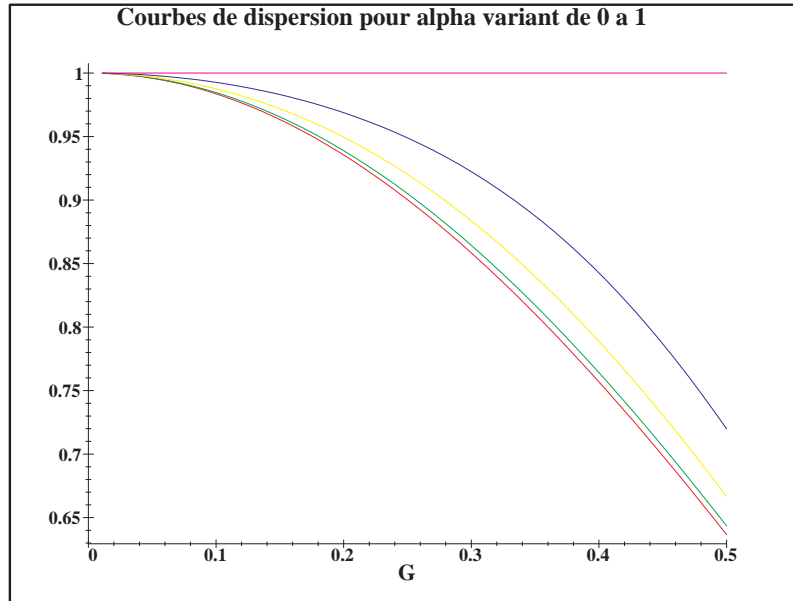


Figure 12: Courbes de dispersion du schéma aux différences finies

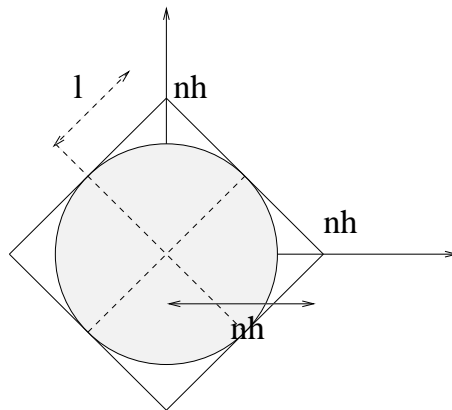


Figure 13: Support de la solution élémentaire exacte (disque) et de la solution élémentaire approchée (losange)

## 4.6 Stabilité par techniques énergétiques

Comme dans le cas continu et dans le cas semi-discret, on peut établir la stabilité en montrant la conservation d'une énergie.

**Théorème 7** Soient  $u_0 \in H^1$ ,  $u_1 \in L^2$  et  $u_h^n$  la solution du schéma (57)-(58)-(60) où  $u_{0h}$  et  $u_{1h}$  vérifient les hypothèses (39) et (41).

• La quantité

$$(75) \quad E_h^{n+1/2} = \frac{1}{2} \left\| \frac{u_h^{n+1} - u_h^n}{\Delta t} \right\|^2 + \frac{1}{2} (A_h u_h^n, u_h^{n+1})$$

se conserve c'est à dire

$$(76) \quad E_h^{n+1/2} = E_h^{n-1/2}, \quad \forall n.$$

• Sous la condition CFL  $\alpha = \frac{c\Delta t}{h} < 1$ ,  $E_h^{n+1/2}$  est une forme quadratique positive et par conséquent le schéma est stable. Plus précisément, on a l'estimation suivante:

$$(77) \quad \frac{1}{2} (1 - \alpha^2) \left\| \frac{u_h^{n+1} - u_h^n}{\Delta t} \right\|^2 \leq E_h^{n+1/2}$$

Nous laissons au lecteur la preuve de ce théorème, au travers de l'exercice suivant.

**Exercice 11** 1- En multipliant scalairement le schéma

$$\frac{u_h^{n+1} - 2u_h^n + u_h^{n-1}}{\Delta t^2} + A_h u_h^n = 0$$

par l'approximation centrée de  $(\partial_t u)^n$ , montrer que la quantité

$$E_h^{n+1/2} = \frac{1}{2} \left\| \frac{u_h^{n+1} - u_h^n}{\Delta t} \right\|^2 + \frac{1}{2} (A_h u_h^n, u_h^{n+1})$$

se conserve (i.e.  $E_h^{n+1/2} = E_h^{n-1/2}$ ,  $\forall n$ ). Comparer aux cas continus et semi-discrets. Commenter.

2- Montrer que pour les conditions initiales (58) et (60) (où  $u_{0h}$  et  $u_{1h}$  vérifient les hypothèses (39) et (41)) on a

$$E_h^{1/2} \leq CE(0) = C \left( \frac{1}{2} \|u_1\|^2 + \frac{1}{2} a(u_0, u_0) \right)$$

3- Montrer que pour tout  $v_h$ , on a

$$(78) \quad (A_h v_h, v_h) \leq \frac{4c^2}{h^2} \|v_h\|^2$$

4- En utilisant l'identité

$$a_h(u, v) = \frac{1}{4} a_h(u + v, u + v) - \frac{1}{4} a_h(u - v, u - v)$$

en déduire que

$$E_h^{n+1/2} \geq \frac{1}{2} \left( 1 - \frac{c^2 h^2}{\Delta t^2} \right) \left\| \frac{u_h^{n+1} - u_h^n}{\Delta t} \right\|^2$$

et conclure.

### Estimation de la solution

On suppose que  $\alpha = \frac{c\Delta t}{h} < 1$ . L'estimation d'énergie obtenue au théorème 7 permet d'obtenir le résultat de stabilité suivant:

**Corollaire 1** *La solution du schéma discrétisé vérifie l'estimation suivante:*

$$(79) \quad \|u_h^n\| \leq C(\alpha)(\|u_0\| + t^n \sqrt{E(0)})$$

Cette estimation est à rapprocher de l'estimation qu'on avait établie pour le schéma semi-discrétisé:

$$\|u_h(t)\| \leq \|u^0\| + \sqrt{2}t\sqrt{E(0)}$$

**Démonstration:** l'estimation (77) peut se réécrire, en utilisant l'identité (76):

$$\frac{1}{2}(1 - \alpha^2) \left\| \frac{u_h^{n+1} - u_h^n}{\Delta t} \right\|^2 \leq E_h^{n+1/2} = E_h^{1/2} \leq CE(0)$$

d'où

$$(80) \quad \|u_h^{n+1} - u_h^n\| \leq M\Delta t$$

avec

$$(81) \quad M = C\sqrt{\frac{2E(0)}{1 - \alpha^2}}$$

On a donc

$$\|u_h^n\| \leq \|u_h^n - u_h^{n-1}\| + \|u_h^{n-1}\| \leq M\Delta t + \|u_h^{n-1}\|$$

soit

$$\|u_h^n\| \leq nM\Delta t + \|u_h^0\|$$

■

### Estimation d'énergie pour le schéma avec terme source

Par la même technique, il est facile de montrer:

**Corollaire 2** *Sous les mêmes hypothèses que celles du théorème 7, si  $u_h^n$  est solution du schéma avec un terme source:*

$$(82) \quad \frac{u_h^{n+1} - 2u_h^n + u_h^{n-1}}{\Delta t^2} + A_h u_h^n = f_h^n$$

avec les conditions initiales (58) et (60), alors l'énergie  $E_h^{n+1/2}$  vérifie l'identité suivante:

$$E_h^{n+1/2} - E_h^{n-1/2} = \frac{1}{2}(f_h^n, u_h^{n+1} - u_h^{n-1})$$

De plus, si  $\alpha < 1$ , nous avons l'estimation d'énergie suivante:

$$(83) \quad \sqrt{E_h^{n+1/2}} \leq \sqrt{E_h^{1/2}} + \frac{\sqrt{2}}{2\sqrt{1 - \alpha^2}} \Delta t \sum_{k=1}^n \|f_h^k\|$$

**Preuve:** l'identité (82) est immédiate. Il est par ailleurs facile de voir qu'elle implique:

$$\underbrace{E_h^{n+1/2} - E_h^{n-1/2}}_{(\sqrt{E_h^{n+1/2}} - \sqrt{E_h^{n-1/2}})(\sqrt{E_h^{n+1/2}} + \sqrt{E_h^{n-1/2}})} \leq \frac{1}{2} \|f_h^n\| (\|u_h^{n+1} - u_h^n\| + \|u_h^n - u_h^{n-1}\|)$$

Si  $\alpha < 1$ , l'estimation (77) montre que

$$\|u_h^{n+1} - u_h^n\| \leq \Delta t \frac{\sqrt{2}}{\sqrt{1-\alpha^2}} \sqrt{E_h^{n+1/2}}$$

ce qui, avec l'inégalité précédente conduit à:

$$\sqrt{E_h^{n+1/2}} - \sqrt{E_h^{n-1/2}} \leq \frac{1}{2} \|f_h^n\| \Delta t \frac{\sqrt{2}}{\sqrt{1-\alpha^2}}$$

et permet d'obtenir (83). ■

**Remarque 5** La constante  $C(\alpha)$  de l'estimation (79) dépend de  $\alpha$ : elle est en  $\frac{1}{\sqrt{1-\alpha^2}}$ . On perd donc l'estimation quand  $\alpha$  se rapproche de 1, or on sait par ailleurs que le schéma s'améliore lorsque  $\alpha$  se rapproche de 1. En fait, il existe une autre méthode, toujours liée à une énergie discrète, qui permet d'obtenir une estimation avec une constante  $C$  indépendante de  $\alpha$  (voir Annexe B).

## 4.7 Convergence du schéma totalement discrétisé par différences finies

Le résultat général sur les schémas disant que si un schéma est consistant et stable alors il est convergent, s'applique ici. Plus précisément on a :

**Théorème 8** Soit  $u_h^n$  la solution du schéma (57)-(58)-(60) où  $u_{0h}$  et  $u_{1h}$  vérifient les hypothèses (39) et (41)) et approchent  $u_0$  et  $u_1$  en norme  $L^2$ . Si la solution du problème continu est assez régulière et si  $\alpha < 1$  alors

$$\lim_{h \rightarrow 0} \sup_n \|u_h^n - \bar{u}_h^n\|_{L^2} = 0$$

où  $\bar{u}_h(t^n)$  la fonction de  $L_h^2$  définie à partir des valeurs de  $u$  aux noeuds ( $u \in C^0$ )

$$\bar{u}_j(t^n) = u(x_j, t^n)$$

Plus précisément, on montre que

$$(84) \quad \sqrt{E_h^{n+1/2}} \leq \sqrt{E_h^{1/2}} + \frac{\sqrt{2}}{2\sqrt{1-\alpha^2}} \Delta t \sum_{k=1}^n \|\varepsilon_h^k\|$$

où on a posé

$$(85) \quad E_h^{n+1/2} = \frac{1}{2} \left\| \frac{e_h^{n+1} - e_h^n}{\Delta t} \right\|^2 + \frac{1}{2} (A_h e_h^n, e_h^{n+1})$$

et où  $e_h^n = \bar{u}_h^n - u_h^n$  est l'erreur de convergence.

**Preuve:** linéarité, il est facile de voir que l'erreur de convergence vérifie :

$$\begin{cases} \frac{e_h^{n+1} - 2e_h^n + e_h^{n-1}}{\Delta t^2} + A_h e_h^n = \varepsilon_h^n \\ e_h^0 = \bar{u}_h^0 - u_h^0, & (e_j^0 = u_0(x_j) - u_{0j}) \\ e_h^1 = \bar{u}_h^1 - u_h^1, & (e_j^1 = u(x_j, \Delta t) - u_j^1) \end{cases}$$

où  $\varepsilon_h^n$  est l'erreur de troncature, définie en (61). D'après l'estimation (83) du corollaire 2, l'énergie associée à ce problème (85) vérifie:

$$\sqrt{E_h^{n+1/2}} \leq \sqrt{E_h^{1/2}} + \frac{\sqrt{2}}{2\sqrt{1-\alpha^2}} \Delta t \sum_{k=1}^n \|\varepsilon_h^k\|$$

Sous la condition CFL  $\alpha < 1$ , c'est à dire **si le schéma est stable**, cette énergie est une quantité quadratique positive, ce qui montre que si le second membre de cette estimation tend vers 0, alors l'énergie tend vers 0 et ceci impliquera que l'erreur tend vers 0. On remarque qu'il y a deux sources d'erreur:

- le terme

$$\sqrt{E_h^{1/2}} = \frac{1}{2} \left\| \frac{e_h^1 - e_h^0}{\Delta t} \right\|^2 + \frac{1}{2} (A_h e_h^0, e_h^1)$$

qui représente l'erreur due à l'approximation des conditions initiales. Ce terme tend vers 0 si les données initiales sont bien approchées.

- le terme  $\Delta t \sum_{k=1}^n \|\varepsilon_h^k\|$  provenant de l'erreur de troncature. **Pour n'importe quel schéma consistant**, on peut montrer que ce terme tend vers 0. En effet, dans ce cas on sait que  $\varepsilon_j^k \rightarrow 0$  lorsque  $\Delta t$  et  $\Delta x$  tendent vers 0, ce qui veut aussi dire qu'on a une estimation du type:

$$|\varepsilon_j^k| \leq C(\Delta t, h) \nu_j^k$$

où  $\nu_h^k$  est une suite de fonctions bornées de  $L_h^2$  et  $C(\Delta t, h) \rightarrow 0$  lorsque  $\Delta t$  et  $\Delta x$  tendent vers 0. Il est facile d'en déduire que

$$\|\varepsilon_h^k\| \leq C(\Delta t, h) \|\nu_h^k\|, \quad \forall k$$

Par conséquent:

$$\Delta t \sum_{k=1}^n \|\varepsilon_h^k\| \leq \Delta t C(\Delta t, h) \sum_{k=1}^n \|\nu_h^k\|$$

Plus particulièrement, pour notre schéma, on peut montrer<sup>1</sup> que si la solution est assez régulière (on voit la régularité nécessaire dans l'estimation), on a

$$\Delta t \sum_{k=1}^n \|\varepsilon_h^k\| \leq C \left( \Delta t^2 \left\| \frac{\partial^4 u}{\partial t^4} \right\|_{C^0(0,T;L^2)} + h^2 \left\| \frac{\partial^4 u}{\partial x^4} \right\|_{C^0(0,T;L^2)} + \Delta t^2 h \left\| \frac{\partial^5 u}{\partial t^4 \partial x} \right\|_{C^0(0,T;L^2)} \right)$$

**Donc l'erreur de consistance est d'ordre 2.**

Ce qui montre que **le schéma est convergent.** ■

---

<sup>1</sup>Nous ne faisons pas la démonstration ici qui est technique sans être difficile: elle repose sur des développements de Taylor avec reste intégral



**Remarque 6 - Sur l'erreur due à l'approximation des conditions initiales**

Le terme  $\sqrt{E_h^{1/2}}$  dans le membre de droite de l'estimation d'erreur (84) représente l'erreur commise sur l'approximation des conditions initiales. Si par exemple, nous faisons le choix d'approcher  $u_0$  et  $u_1$  par  $u_{0h}$  et  $u_{1h}$  définis en (28), c'est à dire les valeurs ponctuelles, on a

$$e_h^0 = 0$$

et

$$e_j^1 = u(x_j, \Delta t) - u_j^1 = u_0(x_j) + \Delta t u_1(x_j) + \frac{\Delta t^2}{2} c^2 \frac{\partial^2 u_0}{\partial x^2}(x_j) + O(\Delta t^3) - u_j^1$$

C'est sur ce terme qu'on va voir l'importance du choix du démarrage.

- Si on avait fait le choix (59), c'est à dire:

$$u_j^1 = u_{0j} + \Delta t u_{1j}$$

on voit que on obtient

$$e_j^1 = \underbrace{u_0(x_j) - u_{0j}}_{=0} + \Delta t \underbrace{(u_1(x_j) - u_{1j})}_{=0} + O(\Delta t^2) = O(\Delta t^2)$$

Or dans le terme d'erreur ce n'est pas  $e_h^1$  qui intervient mais  $e_h^1/\Delta t$ :

$$E_h^{1/2} = \frac{1}{2} \left\| \frac{e_h^1 - e_h^0}{\Delta t} \right\|^2 + \frac{1}{2} (A_h e_h^0, e_h^1) = \frac{1}{2} \left\| \frac{e_h^1}{\Delta t} \right\|^2 = O(\Delta t^2) \Rightarrow \sqrt{E_h^{1/2}} = O(\Delta t)$$

- Par contre, si on démarre avec (60), c'est à dire:

$$u_j^1 = u_{0j} - \frac{\Delta t^2}{2} (A_h u_{0h})_j + \Delta t u_{1j}$$

on obtient:

$$\begin{aligned} e_j^1 &= u_0(x_j) - u_{0j} + \Delta t (u_1(x_j) - u_{1j}) + \frac{\Delta t^2}{2} (c^2 \frac{\partial^2 u_0}{\partial x^2} + (A_h u_{0h})_j) + O(\Delta t^3) \\ &= \frac{\Delta t^2}{2} \underbrace{(c^2 \frac{\partial^2 u_0}{\partial x^2}(x_j) + (A_h u_{0h})_j)}_{O(h^2)} + O(\Delta t^3) \end{aligned}$$

et cette fois ci on peut montrer que

$$\sqrt{E_h^{1/2}} = \frac{1}{\sqrt{2}} \left\| \frac{e_h^1}{\Delta t} \right\| = O(\Delta t^2)$$

## 5 Approximation variationnelle (Galerkin)

Nous allons maintenant nous placer dans un cadre variationnel. Nous ferons ici une présentation très succincte de l'approche variationnelle et de la méthode des éléments finis, et nous renvoyons le lecteur aux nombreux ouvrages ou cours consacrés à cette méthode.

Du point de vue de l'approximation les principaux intérêts d'utiliser un schéma variationnel par rapport à un schéma aux différences finies sont les suivants: ils permettent d'avoir une approche systématique pour construire des schémas (y compris d'ordre élevé), on peut en faire l'analyse en milieu hétérogène, on dispose de résultats de convergence et des estimations d'erreur, on peut utiliser des grilles non régulières (bonne approximation de la géométrie). Une des différences essentielles est qu'ils font apparaître une matrice de masse qui rend le schéma non explicite. Cependant, nous allons voir que pour des éléments finis standards, on peut approcher cette matrice par une matrice diagonale (condensation de masse).

### 5.1 Formulation variationnelle continue

On considère l'équation à coefficients variables posée sur un domaine borné  $\Omega = ]0, L[$ :

$$(86) \quad \begin{cases} \rho \frac{\partial^2 u}{\partial t^2} - \frac{\partial}{\partial x} \left( \mu \frac{\partial u}{\partial x} \right) = 0, & x \in ]0, L[, t > 0 \\ u(0, t) = u(L, t) = 0 \\ u(x, 0) = u_0(x), \quad \frac{\partial u}{\partial t}(x, 0) = u_1(x) \end{cases}$$

et on fait les hypothèses suivantes sur les coefficients et sur les conditions initiales

$$(87) \quad \begin{cases} \rho \in L^\infty(0, L), & 0 < \rho_* \leq \rho(x) \leq \rho^* \\ \mu \in L^\infty(0, L), & 0 < \mu_* \leq \mu(x) \leq \mu^* \\ u_0 \in H_0^1(0, L) = V, & u_1 \in L^2(0, L) = H \end{cases}$$

Sous ces hypothèses, le problème (86) admet une solution unique:

$$u \in C^0(0, T; V); \quad \frac{du}{dt} \in C^0(0, T; H)$$

qui vérifie le problème variationnel suivant:

$$(88) \quad \begin{cases} \frac{d^2}{dt^2} (u(t), v)_\rho + a(u(t), v) = 0, & \forall v \in V \\ u(0) = u_0; \quad \frac{du}{dt}(0) = u_1 \end{cases}$$

où nous avons posé:

$$(89) \quad \begin{cases} (u, v)_\rho = \int_0^L \rho(x) u(x) v(x) dx \\ a(u, v) = \int_0^L \mu(x) \frac{du}{dx} \frac{dv}{dx} dx \end{cases}$$

Grâce aux hypothèses sur  $\rho$  et  $\mu$ , le produit scalaire  $(\cdot, \cdot)_\rho$  est équivalent au produit scalaire usuel sur  $H$  et  $a(\cdot, \cdot)$  est une forme bilinéaire symétrique continue sur  $V \times V$  et coercive sur  $V$  (inégalité de Poincaré sur  $H_0^1$ :  $\|v\|_{L^2} \leq C \|\nabla v\|_{L^2}$ ).

## 5.2 Approximation de Galerkin (semi-discrétisation en espace)

### 5.2.1 La formulation approchée

On introduit un sous-espace vectoriel  $V_h \subset V$ , de dimension finie  $N$  qui approche  $V$  au sens suivant:

$$(H1) \quad \lim_{h \rightarrow 0} \inf_{v_h \in V_h} \|v - v_h\|_V = 0, \quad \forall v \in V.$$

Soit  $(\varphi_I)_{I=1, N}$  une base de  $V_h$ . Le problème approché consiste à se placer dans  $V_h$  au lieu de  $V$ , c'est à dire à chercher  $u_h(t) \in V_h$  tel que

$$(90) \quad \begin{cases} \frac{d^2}{dt^2}(u_h(t), v_h)_\rho + a(u_h(t), v_h) = 0, & \forall v_h \in V_h \\ u_h(0) = u_{0h} \in V_h ; \frac{du_h}{dt}(0) = u_{1h} \in V_h \end{cases}$$

où  $u_{0h}$  approche  $u_0$  dans  $V$  et  $u_{1h}$  approche  $u_1$  dans  $H$ , i.e.

$$(H2) \quad \lim_{h \rightarrow 0} \|u_0 - u_{0h}\|_V = 0, \quad \lim_{h \rightarrow 0} \|u_1 - u_{1h}\|_H = 0.$$

En décomposant  $u_h(t)$  sur la base

$$u_h(t) = \sum_{I=1}^N u_I(t) \varphi_I$$

et en notant  $U(t)$  le vecteur de composantes  $u_I(t)$ , ceci peut s'écrire matriciellement :

$$(91) \quad \begin{cases} M \frac{d^2 U}{dt^2} + K U = 0, \\ U(0) = U_0 ; \frac{dU}{dt}(0) = U_1 \end{cases}$$

où les matrices  $M$  et  $K$  sont respectivement les matrices de masse et de rigidité, définies par

$$M_{IJ} = (\varphi_I, \varphi_J)_\rho ; \quad K_{IJ} = a(\varphi_I, \varphi_J)$$

**Exercice 12** Montrer que les matrices  $M$  et  $K$  sont symétriques et définies positives.

Le système (91) est maintenant un système différentiel ordinaire.

### 5.2.2 Estimation d'énergie

#### Exercice 13 . Estimation a priori sur l'énergie

1- Montrer que la solution  $u_h(t)$  du problème semi-discrétisé vérifie la conservation de l'énergie suivante:

$$E_h(t) = \frac{1}{2} \left\| \frac{du_h}{dt} \right\|_\rho^2 + \frac{1}{2} a(u_h, u_h)$$

2- Montrer que sous les hypothèses faites sur les conditions initiales on peut majorer  $E_h(0)$  indépendamment de  $h$ ,

$$\exists C > 0, \quad C \text{ indépendante de } h, \quad E_h(0) \leq C(\|u_1\|_H^2 + \|u_0\|_V^2)$$

Conclure.

3- Montrer que le choix suivant de conditions initiales convient:

$$\begin{cases} u_{0h} = \text{la projection orthogonale de } u_0 \text{ sur } V_h \text{ pour le produit scalaire } V \\ u_{1h} = \text{la projection orthogonale de } u_1 \text{ sur } V_h \text{ pour le produit scalaire } H \end{cases}$$

On pourra admettre (ou démontrer) le lemme suivant:

**Lemme 4** Si  $V$  et  $H$  sont 2 espaces de Hilbert, tels que  $V \subset H$ ,  $V$  dense dans  $H$  et avec injection continue alors l'hypothèse (H1) implique

$$\lim_{h \rightarrow 0} \inf_{v_h \in V_h} |v - v_h|_H = 0, \quad \forall v \in H.$$

### 5.2.3 Convergence du schéma semi discrétisé

Nous avons vu pour le schéma aux différences finies que la notion de convergence devait être précisée et nous en avons vu deux différentes selon ce qu'on compare à la solution exacte. Dans une approche variationnelle, cette notion est beaucoup plus simple à définir puisque la solution approchée est cherchée dans un sous espace de l'espace auquel appartient la solution exacte. On peut donc comparer ces deux fonctions, en tant qu'éléments de cet espace. Soit  $u$  solution de (88) et  $u_h$  solution de (90). Nous supposons ici que  $u \in C^2(0, T; V)$  et que  $u_h \in C^2(0, T; V_h)$ , ce qui suppose en particulier que  $u_0$  et  $u_1$  sont dans  $V$ . On supposera que  $u_{0h}$  converge vers  $u_0$  dans  $V$  et  $u_{1h}$  converge vers  $u_1$  dans  $H$ . L'erreur  $u - u_h$  vérifie le problème suivant:

$$(92) \quad \begin{cases} (\partial_{tt}^2(u - u_h), v_h)_\rho + a(u - u_h, v_h) = 0, & \forall v_h \in V_h \\ (u - u_h)(0) = u_0 - u_{0h} ; \frac{d(u - u_h)}{dt}(0) = u_1 - u_{1h} \end{cases}$$

La première équation peut encore s'écrire en introduisant  $w_h \in V_h$ :

$$(93) \quad \begin{aligned} & (\partial_{tt}^2(w_h - u_h), v_h)_\rho + a(w_h - u_h, v_h) = \\ & = (\partial_{tt}^2(w_h - u), v_h)_\rho + a(w_h - u, v_h). \end{aligned}$$

En choisissant  $v_h = \partial_t(w_h - u_h)$  on obtient:

$$\begin{aligned} & (\partial_{tt}^2(w_h - u_h), \partial_t(w_h - u_h))_\rho + a(w_h - u_h, \partial_t(w_h - u_h)) = \\ & = (\partial_{tt}^2(w_h - u), \partial_t(w_h - u_h))_\rho + a(w_h - u, \partial_t(w_h - u_h)) \end{aligned}$$

soit encore:

$$\begin{aligned} & \frac{1}{2} \frac{d}{dt} \|\partial_t(w_h - u_h)\|_\rho^2 + \frac{1}{2} \frac{d}{dt} a(w_h - u_h, w_h - u_h) = \\ & = (\partial_{tt}^2(w_h - u), \partial_t(w_h - u_h))_\rho + a(w_h - u, \partial_t(w_h - u_h)) \end{aligned}$$

Posons

$$(94) \quad \mathcal{E}_h(t) = \frac{1}{2} \|\partial_t(w_h - u_h)\|_\rho^2 + \frac{1}{2} a(w_h - u_h, w_h - u_h)$$

L'égalité précédente se réécrit:

$$(95) \quad \frac{d}{dt} \mathcal{E}_h(t) = (\partial_{tt}^2(w_h - u), \partial_t(w_h - u_h))_\rho + a(w_h - u, \partial_t(w_h - u_h))$$

**Projection elliptique** . Pour pouvoir obtenir une estimation de l'énergie on veut majorer le second membre par une "fonction de  $u$  multipliée par une quantité qui fait intervenir l'énergie". Le terme  $a(w_h - u, \partial_t(w_h - u_h))$  est gênant car il est majoré par  $\|a\| \|w_h - u\|_V \|\partial_t(w_h - u_h)\|_V$  or dans l'énergie c'est  $\|\partial_t(w_h - u_h)\|_H$  qui intervient. On s'en débarrasse en introduisant l'opérateur de *projection elliptique*:

$$(96) \quad \begin{aligned} P_h : V &\longrightarrow V_h \\ u &\longrightarrow P_h u \text{ avec} \\ &a(P_h u - u, v_h) = 0, \quad \forall v_h \in V_h \end{aligned}$$

L'opérateur  $P_h$  est bien défini par Lax-Milgram. D'autre part on montre que

**Lemme 5**

$$(97) \quad \boxed{\|P_h u - u\|_V \leq \frac{\|a\|}{\alpha} \inf_{v_h \in V_h} \|v_h - u\|_V}$$

**Démonstration:** Par définition, on a :

$$a(P_h u - u, P_h u - u) = a(P_h u - u, v_h - u), \quad \forall v_h \in V_h$$

d'où (coercivité + continuité de  $a$ )

$$\alpha \|P_h u - u\|_V^2 \leq \|a\| \|P_h u - u\|_V \|v_h - u\|_V$$

ce qui permet de conclure. ■

On peut aussi facilement montrer que

$$P_h(\partial_t u) = \partial_t(P_h u)$$

En effet, en dérivant l'équation définissant la projection elliptique par rapport au temps:

$$\frac{d}{dt} a(P_h u, v_h) = \frac{d}{dt} a(u, v_h), \quad \forall v_h \in V_h$$

Si  $u \in C^1(0, T; V)$  alors on a  $P_h u \in C^1(0, T; V_h)$  et on peut donc écrire :

$$a(\partial_t(P_h u), v_h) = a(\partial_t u, v_h), \quad \forall v_h \in V_h$$

ce qui définit  $\partial_t(P_h u)$  comme la projection elliptique de  $\partial_t u$ . On a supposé que  $u \in C^2(0, T; V)$ , donc le résultat est valable pour les dérivées secondes en temps. En particulier, on a aussi l'estimation (97) sur les dérivées en temps.

L'estimation sur la projection elliptique est en norme  $V$ . Pour montrer l'estimation d'erreur en norme  $L^2$ , il est utile d'avoir également une estimation sur la projection elliptique en norme  $L^2$ . On peut bien sûr la majorer par la norme  $V$ , mais en fait, il est possible de montrer qu'on gagne un ordre en  $h$ , en utilisant le:

**Lemme 6** . (lemme de Aubin - Nitsche). Pour tout  $u \in V$ , on a l'estimation:

$$(98) \quad \boxed{\|P_h u - u\|_H \leq \|a\| \|P_h u - u\|_V \sup_{g \in H} \frac{1}{\|g\|_H} \inf_{v_h \in V_h} \|v_h - \varphi_g\|_V}$$

où  $\varphi_g$  est solution du problème adjoint, défini pour tout  $g \in H$  par

$$(99) \quad \begin{cases} \text{pour tout } g \in H, \text{ trouver } \varphi_g \in V \text{ tel que} \\ a(v, \varphi_g) = (g, v)_H, \quad \forall v \in V \end{cases}$$

**Démonstration:** On utilise la définition de la norme dans  $H$ :

$$\|P_h u - u\|_H = \sup_{g \in H} \frac{|(g, P_h u - u)|}{\|g\|_H}$$

On a d'une part  $a(u - P_h u, v_h) = 0$  pour tout  $v_h$  et d'autre part  $a(u - P_h u, \varphi_g) = (g, u - P_h u)$ , d'où par différence

$$(g, u - P_h u) = a(u - P_h u, \varphi_g - v_h) \leq \|a\| \|P_h u - u\|_V \|\varphi_g - v_h\|_V, \quad \forall v_h \in V_h$$

d'où le résultat. ■

Pour déduire de ce lemme une estimation qui montre l'ordre en  $h$ , nous avons besoin de quelques hypothèses supplémentaires. Nous supposons qu'il existe un opérateur  $\Pi_h$  qui envoie  $V$  dans  $V_h$  et tel que pour toute fonction  $v \in H^{k+1}(\Omega) \cap V$

$$(100) \quad \begin{aligned} (i) \quad & \|\Pi_h v - v\|_{1,\Omega} \leq Ch^k |v|_{k+1,\Omega} \\ (ii) \quad & \|\Pi_h v - v\|_{0,\Omega} \leq Ch^{k+1} |v|_{k+1,\Omega} \end{aligned}$$

Un tel opérateur existe bien lorsqu'on utilise la méthode des éléments finis: il s'agit de l'opérateur d'interpolation. Pour pouvoir utiliser ce résultat, nous avons besoin de régularité:

**Définition 2** *Le problème adjoint (99) est dit régulier si pour tout  $g \in H$ , la solution  $\varphi_g$  est dans  $V \cap H^2(\Omega)$  et s'il existe une constante  $C > 0$  telle que*

$$(101) \quad \|\varphi_g\|_{2,\Omega} \leq C \|g\|_{0,\Omega}$$

**Lemme 7** *Si le problème adjoint (99) est régulier, on a l'estimation suivante, pour tout  $u \in V$ :*

$$(102) \quad \boxed{\|P_h u - u\|_H \leq Ch \|P_h u - u\|_V}$$

**Démonstration:** Si le problème adjoint est régulier, pour tout  $g \in H$ , la solution  $\varphi_g$  est dans  $H^2(\Omega)$  et vérifie (101), donc l'estimation d'interpolation (100)-(ii) nous indique que

$$\|\Pi_h \varphi_g - \varphi_g\|_V \leq Ch |\varphi_g|_{2,\Omega} \leq Ch |g|_{0,\Omega}$$

la deuxième inégalité provenant de (101). On en déduit donc que pour tout  $g \in H$

$$\frac{1}{\|g\|_H} \inf_{v_h \in V_h} \|v_h - \varphi_g\|_V \leq \frac{1}{\|g\|_H} \|\Pi_h \varphi_g - \varphi_g\|_V \leq Ch$$

ce qui montre (102) en utilisant l'estimation (98). ■

**Estimation d'erreur  $L^2$**  .

**Lemme 8** *On a l'estimation d'énergie suivante:*

$$(103) \quad \mathcal{E}_h^{1/2}(t) \leq \mathcal{E}_h^{1/2}(0) + C \int_0^t \|(I - P_h) \partial_{tt}^2 u\|_H(s) ds$$

**Démonstration:** On reprend (95) en choisissant  $w_h = P_h u$ , ce qui donne

$$\underbrace{\frac{d}{dt} \mathcal{E}_h(t)}_{2\mathcal{E}_h^{1/2} \frac{d}{dt} \mathcal{E}_h^{1/2}} = (\partial_{tt}^2(w_h - u), \partial_t(w_h - u_h))_\rho \leq \left\| \partial_{tt}^2(w_h - u) \right\|_H \underbrace{\left\| \partial_t(w_h - u_h) \right\|_H}_{\leq C\mathcal{E}_h^{1/2}}$$

d'où

$$\frac{d}{dt} \mathcal{E}_h^{1/2} \leq C \left\| \partial_{tt}(w_h - u) \right\|_H$$

on en déduit (103) en intégrant en temps. ■

**Théorème 9** Si  $u$  est solution de (88) et  $u_h$  solution de (90), avec  $u \in C^2(0, T; V)$  et  $u_h \in C^2(0, T; V_h)$ , et si le problème adjoint est régulier, alors on a les estimations suivantes:

$$(i) \quad \left\| \partial_t u - \partial_t u_h \right\|_H(t) \leq C \left( \mathcal{E}_h^{1/2}(0) + h \left\| (I - P_h) \partial_t u \right\|_V(t) \right. \\ \left. + h \int_0^t \left\| (I - P_h) \partial_{tt}^2 u \right\|_V(s) ds \right) \\ (ii) \quad \left\| u - u_h \right\|_V(t) \leq \left\| (I - P_h) u \right\|_V + C \left( \mathcal{E}_h^{1/2}(0) + h \int_0^t \left\| (I - P_h) \partial_{tt}^2 u \right\|_V(s) ds \right)$$

**Démonstration:** Par Cauchy-Schwartz:

$$\left\| \partial_t u - \partial_t u_h \right\|_H \leq \left\| \partial_t u - \partial_t P_h u \right\|_H + \underbrace{\left\| \partial_t P_h u - \partial_t u_h \right\|_H}_{\leq C\mathcal{E}_h^{1/2}}$$

et on obtient (104)-(i) en utilisant (103) et (102). De même, on a :

$$\left\| u - u_h \right\|_V \leq \left\| u - P_h u \right\|_V + \underbrace{\left\| P_h u - u_h \right\|_V}_{\leq C\mathcal{E}_h^{1/2}}$$

qui permet d'obtenir (104)-(ii). ■

**Lemme 9** • Si on choisit comme conditions initiales approchées

$$u_{0h} = P_h u_0, \quad \text{et} \quad u_{1h} = P_h u_1$$

alors  $\mathcal{E}_h(0) = 0$ .

• Pour d'autres choix de conditions initiales approchées on a

$$(105) \quad \mathcal{E}_h^{1/2}(0) \leq C \left( \left\| (I - P_h) u_1 \right\|_H + \left\| u_1 - u_{1h} \right\|_H + \left\| u_0 - u_{0h} \right\|_V \right)$$

**Démonstration:** l'énergie discrète initiale s'écrit:

$$\mathcal{E}_h(0) = \frac{1}{2} \left\| P_h u_1 - u_{1h} \right\|_H^2 + \frac{1}{2} a(P_h u_0 - u_{0h}, P_h u_0 - u_{0h})$$

on obtient donc immédiatement le premier point. Sinon, on utilise Cauchy Schwartz pour majorer:

$$\left\| P_h u_1 - u_{1h} \right\|_H \leq \left\| (I - P_h) u_1 \right\|_H + \left\| u_1 - u_{1h} \right\|_H$$

D'autre part on a vu que

$$\|P_h u_0 - u_{0h}\|_V \leq \frac{\|a\|}{\alpha} \|v_h - u_0\|_V, \quad \forall v_h \in V_h$$

donc en particulier pour  $v_h = u_{0h}$  :

$$\|P_h u_0 - u_{0h}\|_V \leq \frac{\|a\|}{\alpha} \|u_{0h} - u_0\|_V.$$

■

**Corollaire 3** *On se place sous les hypothèses du théorème 9 et on suppose qu'il existe un opérateur défini comme en (100). On suppose de plus que la solution a la régularité  $\partial_t^s u(t) \in H^{k+1}(\Omega)$  pour  $s = 0, 1, 2$ , et que  $u_1 \in H^{k+1}(\Omega)$ . Alors on a les estimations d'erreur suivantes:*

$$(106) \quad \begin{aligned} (i) \quad \|\partial_t(u - u_h)\|_H(t) &\leq C(\|u_1 - u_{1h}\|_H + \|u_0 - u_{0h}\|_V + \\ &\quad + h^{k+1} \left( |u_1|_{k+1} + |\partial_t u|_{k+1} + \int_0^t |\partial_{tt}^2 u|_{k+1}(s) ds \right)) \\ (ii) \quad \|u - u_h\|_V(t) &\leq C(\|u_1 - u_{1h}\|_H + \|u_0 - u_{0h}\|_V + \\ &\quad + h^k \left( h|u_1|_{k+1} + |u|_{k+1} + h \int_0^t |\partial_{tt}^2 u|_{k+1}(s) ds \right)) \end{aligned}$$

**Démonstration:** On regroupe le différents résultats précédents: si le problème adjoint est régulier on a l'estimation (102) qui relie les normes  $L^2$  aux normes  $V$ . En utilisant la régularité  $H^{k+1}$  on a :

$$\|(I - P_h)\partial_t^s u\|_V \leq Ch^k |\partial_t^s u|_{k+1}$$

on en déduit facilement le résultat en injectant dans l'estimation (104). ■

## 5.3 La méthode des éléments finis.

### 5.3.1 Quelques rappels

La méthode des éléments finis est une méthode particulière pour construire l'espace d'approximation  $V_h$ . Nous ne considérons ici que l'approximation à l'aide d'éléments finis de Lagrange, pour lesquels les fonctions de base sont définies par leurs valeurs aux noeuds (il existe d'autres types d'éléments finis, par exemple ceux de Hermitte pour lesquels les dérivées des fonctions de base interviennent également). Nous la présentons dans le cas général (avec des illustrations en 2D) avant de voir son application au cas 1D.

On introduit un maillage du domaine  $\Omega$  (ou triangulation  $\mathcal{T}_h$ ), c'est à dire une partition en  $L$  polyèdres (triangles, rectangles...),  $\Omega = \cup_{l=1}^L K_l$  tels que

- $\forall l, \overset{\circ}{K}_l \neq \emptyset$
- $K_l \cap K_{l'} = \emptyset$ , si  $l \neq l'$
- Toute face d'un élément est soit la face d'un autre élément, soit une face portée par le bord



Un *élément fini de Lagrange d'ordre k* est la donnée d'un triplet  $(K, \Sigma, P)$  où

- $K$  est un fermé borné
- $\Sigma$  est un ensemble de  $n_K$  points  $(M_i^K)_{i=1, n_K} \in K$  appelés Degrés de Liberté
- $P$  est un espace vectoriel de polynômes définis sur  $K$  tel que  $P \supset P^k(K)$  et tel que  $\exists n_K$  fonctions de base telles que  $\tau_i^K(M_j^K) = \delta_{ij}, \forall i, j = 1, n_k$

A chaque  $K_l$  on associe un élément fini  $E_l = (K_l, \Sigma_l, P_l)$ , on note  $n_l = \text{card } \Sigma_l$  et  $\tau_i^l$  les fonctions de base locales. On définit l'espace d'approximation par

$$V_h = \{v_h \in C^0(\bar{\Omega}), v_h|_{K_l} \in P_l, \forall K_l \in \mathcal{T}_h\}$$

Les fonctions de base globales sont alors définies par:

$$\varphi_I = \begin{cases} \tau_i^l, & \text{si } I \text{ est un DL de } K_l \\ & \text{numéroté } i \text{ dans la numérotation locale} \\ 0 & \text{sinon} \end{cases}$$

**Exemple 1 : Élément fini  $P^1$  en 1D:**  $\Omega = [-M, +M] = \cup_l K_l$

$$\begin{aligned} K &= [0, 1] \\ \Sigma &= \{0, 1\} \\ P &= P^1 = \{p(x) = ax + b, (a, b) \in \mathbb{R}^2\}, \quad \dim P = 2 \\ \tau_1 &= x ; \quad \tau_2 = 1 - x \end{aligned}$$

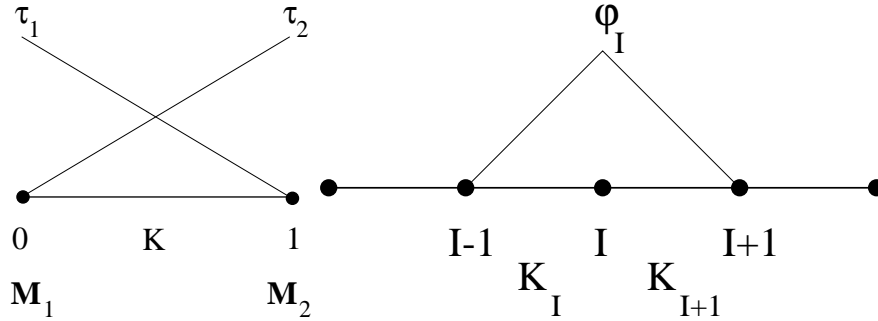


Figure 14: Gauche: Fonctions de base locales sur l'élément de référence. Droite: Fonction de base globale associée au noeud I

**Exemple 2 : Élément fini  $P^2$  en 1D:**  $\Omega = [-M, +M] = \cup_l K_l$

$$\begin{aligned} K &= [0, 1] \\ \Sigma &= \{0, 1/2, 1\} \\ P &= P^2 = \{p(x) = ax^2 + bx + c, (a, b, c) \in \mathbb{R}^3\}, \quad \dim P = 3 \\ \tau_1 &= (1-x)(1-2x) ; \quad \tau_2 = x(1-2x) ; \quad \tau_3 = 4x(1-x) \end{aligned}$$

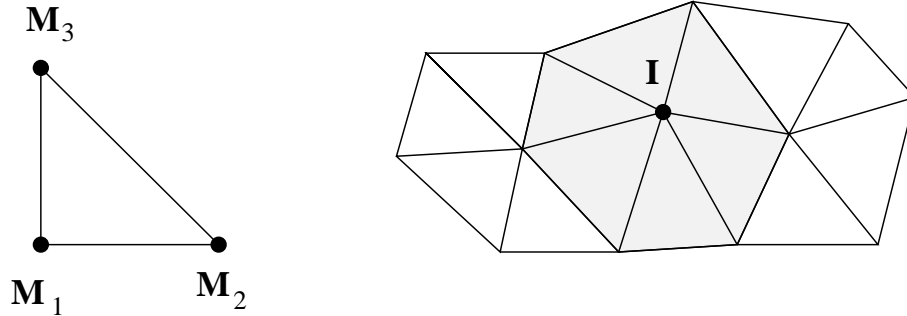
**Exemple 3 : Élément fini  $P^1$  en 2D:**  $\Omega = \cup_l K_l$

$$K = T$$

$$\Sigma = \{M_1, M_2, M_3\}$$

$$P = P^1(T) = \{p(x, y) = ax + by + c, (a, b, c) \in \mathbb{R}^3\}, \dim P = 3$$

$$\tau_1 = 1 - x - y ; \tau_2 = x ; \tau_3 = y$$



**Exercice 14** Décrire l'élément fini  $Q^1$  en 2D.

En pratique pour générer une famille d'éléments finis, on définit un élément fini de Lagrange de référence  $(\widehat{K}, \widehat{P}, \widehat{\Sigma})$  et on construit

$$K = F_K(\widehat{K}), P = \{p : K \rightarrow \mathbb{R}; p \circ F_K \in \widehat{P}\}, \Sigma = F_K(\widehat{\Sigma})$$

où  $F_K$  est une bijection affine, ce qui définit bien un élément fini de Lagrange. La transformation étant affine, tout  $x \in K$  peut s'écrire:

$$x = F_K(\widehat{x}) = B_K \widehat{x} + b_K$$

où  $B_K$  est une matrice inversible et  $b_K$  un vecteur de  $\mathbb{R}^n$  ( $n =$  dimension de l'espace). On supposera que  $\det(B_K) > 0$ . À toute fonction  $v$  définie sur  $K$ , on associe une fonction définie sur  $\widehat{K}$ , notée  $\widehat{v}$ , définie par:

$$\widehat{v}(\widehat{x}) = v(x), \forall \widehat{x} \in \widehat{K} \iff v = \widehat{v} \circ F_K^{-1}$$

On peut en particulier passer d'une intégrale sur  $K$  à une intégrale sur  $\widehat{K}$  par la formule de changement de variable:

$$(107) \quad \int_K v(x) dx = \det(B_K) \int_{\widehat{K}} \widehat{v}(\widehat{x}) d\widehat{x}$$

qui montre (en prenant  $v = 1$ ) que

$$(108) \quad \det(B_K) = \frac{\text{mes}(K)}{\text{mes}(\widehat{K})}$$

On introduit les caractéristiques géométriques suivantes:

$h_K =$  diamètre de  $K = \max$  de la distance euclidienne entre 2 points de  $K$   
 $\rho_K =$  rondeur de  $K =$  diamètre maximum des sphères (cercles) contenues dans  $K$  et on pose  $h = \max_K h_K$ . On supposera que il existe une constante  $\sigma \leq 1$  telle que

$$(109) \quad \forall h, \forall K \in \mathcal{T}_h, \frac{h_K}{\rho_K} \leq \sigma, \forall K$$

ce qui revient à dire que les éléments ne sont pas trop aplatis.

**Opérateur d'interpolation de Lagrange.** Si on utilise des éléments finis de Lagrange, les degrés de liberté sont les noeuds  $(M_j^K)_{j=1,n_K}$  et on définit l'opérateur d'interpolation local par:

$$\Pi_K v = \sum_{j=1}^{n_K} v(M_j^K) \tau_j^K$$

pour toute fonction  $v$  définie sur  $K$ . On suppose que

$$P^k(K) \subset P \subset H^{k+1}(K)$$

On a alors le

**Théorème 10** *Il existe une constante  $C$  qui ne dépend que de l'éléments fini de référence  $(\widehat{K}, \widehat{P}, \widehat{\Sigma})$  telle que pour tout  $m$ ,  $0 \leq m \leq k + 1$*

$$(110) \quad \forall v \in H^{k+1}(K), \quad |v - \Pi_K v|_{m,K} \leq C \frac{h_K^{k+1}}{\rho_K^m} |v|_{k+1,K}$$

**Application.** En utilisant l'hypothèse (109), on voit que l'opérateur d'interpolation vérifie: pour tout  $v \in H^{k+1}(K)$ ,

$$(111) \quad \begin{aligned} (i) \quad & \|\Pi_K v - v\|_{1,K} \leq Ch^k |v|_{k+1,K} \\ (ii) \quad & \|\Pi_K v - v\|_{0,K} \leq Ch^{k+1} |v|_{k+1,K} \end{aligned}$$

On introduit l'opérateur d'interpolation global  $\Pi_h$ :

$$(112) \quad v \in C^0(\overline{\Omega}) \rightarrow \Pi_h v \in L^2(\Omega), \quad \text{tel que } \forall K \in \mathcal{T}_h, \Pi_h v|_K = \Pi_K v$$

donc, pour  $v \in C^0(\overline{\Omega})$  on a

$$(113) \quad \Pi_h v = \sum_{I=1}^N v(M_I) w_I$$

En particulier si  $v \in H^{k+1}(\Omega)$ , avec  $k \geq 1$ , alors (injections de Sobolev, si  $m > n/2$  alors  $H^m(\Omega) \subset C^0(\overline{\Omega})$ ) on a  $v \in C^0(\overline{\Omega})$  et on peut définir l'opérateur d'interpolation.

On déduit du théorème 10 l'erreur globale d'interpolation:

$$|v - \Pi_h v|_{m,\Omega} = \left( \sum_K |v - \Pi_h v|_{m,K}^2 \right)^{1/2}$$

Il existe une constante  $C$  telle que pour tout  $m$ ,  $0 \leq m \leq k + 1$

$$(114) \quad \forall v \in H^{k+1}(\Omega), \quad |v - \Pi_h v|_{m,\Omega} \leq Ch^{k+1-m} |v|_{k+1,\Omega}$$

### 5.3.2 Application à l'approximation de l'équation des ondes 1D.

On reprend notre problème approché et on va considérer une approximation à l'aide d'éléments finis  $P^1$ . On pose  $\Delta x = L/(N + 1)$ ,  $x_j = j\Delta x$  pour  $j = 0, \dots, N + 1$ . L'espace  $V_h$  est donc défini par:

$$V_h = \left\{ v_h \in C^0[0, L], \quad v_h|_{[x_j, x_{j+1}]} \in P^1, \quad \forall j = 0, \dots, N, \quad v_h(0) = v_h(L) = 0 \right\}$$

Une base de  $V_h$  est définie par les fonctions “chapeaux”  $\varphi_j \in V_h$  telles que  $\varphi_j(x_i) = \delta_{ij}$ . En particulier on peut vérifier que toute fonction  $v_h \in V_h$  se décompose sur cette base de la façon suivante:

$$v_h = \sum_{j=1}^N v_h(x_j) \varphi_j$$

**Exercice 15** 1- Montrer (sans calcul) que les matrices de masse et de rigidité obtenues à partir de la méthode des éléments finis  $P^1$  sont tridiagonales.

2- Expliciter ces matrices en milieu homogène et comparer le schéma ainsi obtenu avec le schéma aux différences finies étudié au chapitre précédent. Commenter.

3- On se place en milieu hétérogène et on suppose que  $\rho$  et  $\mu$  sont constantes par morceaux:

$$\begin{aligned} \rho(x) &= \rho_{i+1/2} \quad \text{sur } ]x_i, x_{i+1}[ \\ \mu(x) &= \mu_{i+1/2} \quad \text{sur } ]x_i, x_{i+1}[ \end{aligned}$$

Montrer que le schéma s’écrit cette fois sous la forme

$$\begin{aligned} \frac{\rho_{i+1/2}}{6} \frac{d^2 U_{i+1}}{dt^2} + \frac{1}{3} (\rho_{i-1/2} + \rho_{i+1/2}) \frac{d^2 U_i}{dt^2} + \frac{\rho_{i-1/2}}{6} \frac{d^2 U_{i-1}}{dt^2} \\ + \frac{1}{h^2} (-\mu_{i-1/2} U_{i-1} + (\mu_{i-1/2} + \mu_{i+1/2}) U_i - \mu_{i+1/2} U_{i+1}) = 0 \end{aligned}$$

### 5.3.3 Analyse de dispersion

En homogène le schéma  $P^1$  semi-discrétisé s’écrit:

$$\frac{d^2}{dt^2} \left( \frac{2}{3} U_i + \frac{1}{6} U_{i-1} + \frac{1}{6} U_{i+1} \right) - c^2 \frac{U_{i+1} - 2U_i + U_{i-1}}{h^2} = 0$$

La relation de dispersion s’écrit :

$$\omega^2 \left( \frac{2}{3} + \frac{1}{6} e^{-ikh} + \frac{1}{6} e^{ikh} \right) = \frac{4c^2}{h^2} \sin^2 \frac{kh}{2}$$

soit:

$$\omega^2 = \frac{12c^2}{h^2} \frac{\sin^2 \frac{kh}{2}}{3 - 2 \sin^2 \frac{kh}{2}}$$

ce qui donne comme rapport des vitesses (avec  $K = kh$ )

$$q_{P^1} = \frac{\omega}{kc} = \frac{2\sqrt{3}}{K} \frac{\sin \frac{K}{2}}{\sqrt{3 - 2 \sin^2 \frac{kh}{2}}}$$

à comparer avec celui obtenu pour les différences finies:

$$q_{DF} = \frac{2}{K} \sin \frac{K}{2}$$

On remarque que la vitesse numérique pour le schéma  $P^1$  approche la vitesse par le haut alors que celle du schéma aux différences finies l’approche par le bas:

$$q_{P^1} = 1 + \frac{1}{24} K^2 + O(K^4)$$

$$q_{DF} = 1 - \frac{1}{24} K^2 + O(K^4)$$

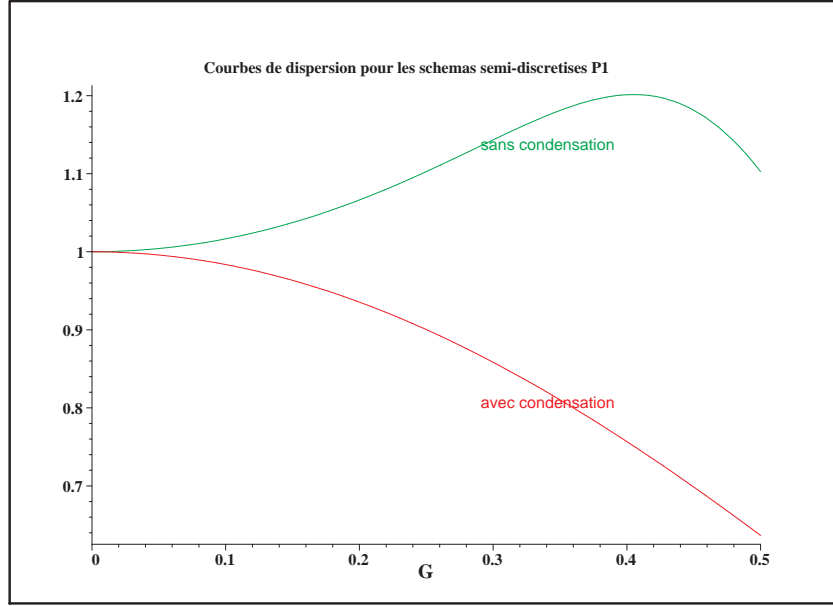


Figure 15: Comparaison des courbes de dispersion avec et sans condensation

#### 5.4 Discrétisation totale - Formules de quadrature, condensation de masse

Pour la discrétisation en temps, on utilise une différence finie centrée d'ordre 2, exactement comme pour le schéma aux différences finies. Le problème approché consiste donc à trouver  $u_h^n \in V_h$  tel que

$$(115) \quad \begin{cases} \left( \frac{u_h^{n+1} - 2u_h^n + u_h^{n-1}}{\Delta t^2}, v_h \right)_\rho + a(u_h^n, v_h) = 0, & \forall v_h \in V_h \\ u_h^0, \text{ et } u_h^1 \text{ donnés dans } V_h \end{cases}$$

On peut par exemple démarrer le schéma avec:

$$(116) \quad \begin{aligned} u_h^0 &= u_{0,h} \\ (u_h^1, v_h)_\rho &= (u_{0,h}, v_h)_\rho - \frac{\Delta t^2}{2} a(u_{0,h}, v_h) + \Delta t (u_{1h}, v_h)_\rho, \quad \forall v_h \in V_h \end{aligned}$$

ce qui peut s'écrire matriciellement:

$$(117) \quad \begin{cases} M \frac{U^{n+1} - 2U^n + U^{n-1}}{\Delta t^2} + KU^n = 0 \\ U^0 = U_0 \\ MU^1 = MU_0 - \frac{\Delta t^2}{2} KU_0 + \Delta t MU_1 \end{cases}$$

Le schéma ainsi obtenu n'est pas explicite: pour calculer  $U^{n+1}$ , on doit inverser la matrice de masse  $M$  tridiagonale symétrique:

$$U^{n+1} = M^{-1}(-\Delta t^2 KU^n + 2U^n - U^{n-1})$$

On peut cependant approcher  $M$  par une matrice diagonale, en utilisant la formule de quadrature

$$(118) \quad \int_a^b f(x)dx \approx (b-a) \frac{f(a) + f(b)}{2}$$

sans perdre de précision sur le schéma.

**Exercice 16** 1- Calculer la matrice de masse approchée obtenue en utilisant la formule de quadrature (118) et écrire le schéma correspondant.

3- Appliquer ce résultat au cas homogène et comparer le nouveau schéma obtenu avec le schéma aux différences finies.

## 5.5 Stabilité par techniques d'énergie (milieux hétérogènes)

En milieu hétérogène la méthode qui passe par Fourier ne peut plus s'appliquer. Par contre, on peut de la même façon que ce qu'on avait fait au §4.6 obtenir une estimation d'énergie. Le cadre variationnel est en fait le plus adapté, puisqu'on travaille dans les espaces liés à l'énergie. On va l'appliquer à n'importe quel schéma de la forme:

$$M \frac{U^{n+1} - 2U^n + U^{n-1}}{\Delta t^2} + KU^n = 0$$

En multipliant scalairement cette équation par  $\frac{U^{n+1} - U^{n-1}}{2\Delta t}$  on obtient l'égalité

$$\left\| \frac{U^{n+1} - U^n}{\Delta t} \right\|_M^2 + (KU^n, U^{n+1}) = \left\| \frac{U^n - U^{n-1}}{\Delta t} \right\|_M^2 + (KU^{n-1}, U^n)$$

où  $\|\cdot\|_M^2$  est la norme associée à la matrice  $M$ , c'est à dire  $\|U\|_M^2 = (MU, U)$ . On a donc une quantité qui se conserve:

$$(119) \quad E^{n+1/2} = \left\| \frac{U^{n+1} - U^n}{\Delta t} \right\|_M^2 + (KU^n, U^{n+1})$$

**Lemme 10** La quantité définie par (119) se conserve:

$$E^{n+1/2} = E^{n-1/2} = E^{1/2}, \quad \forall n$$

En utilisant l'identité

$$(KU, V) = \frac{1}{4}(K(U+V), U+V) - \frac{1}{4}(K(U-V), U-V)$$

on peut réécrire cette quantité sous la forme:

$$E^{n+1/2} = \left( \mathcal{M} \frac{U^{n+1} - U^n}{\Delta t}, \frac{U^{n+1} - U^n}{\Delta t} \right) + \left( K \frac{U^n + U^{n+1}}{2}, \frac{U^n + U^{n+1}}{2} \right)$$

avec

$$\mathcal{M} = M - \frac{\Delta t^2}{4}K$$

Si  $\mathcal{M}$  est définie positive, cette quantité définit bien une énergie et on a stabilité (pour des Conditions Initiales bien choisies). Cette condition est une condition de type CFL et peut encore s'écrire sous la forme:

$$(CFL) \quad \boxed{\frac{\Delta t^2}{4} \sup_{V \neq 0} \frac{(KV, V)}{(MV, V)} < 1}$$

ou encore

$$(120) \quad \frac{\Delta t^2}{4} \sup_{v_h \neq 0} \frac{a(v_h, v_h)}{(v_h, v_h)_\rho} < 1$$

On reconnaît le quotient de Rayleigh:

$$\mathcal{R}(v_h) = \frac{a(v_h, v_h)}{(v_h, v_h)_\rho}$$

Chercher le sup de ce quotient revient à trouver le max des valeurs propres du problème

$$\text{Trouver } u_h \text{ et } \lambda_h \text{ tels que } a(u_h, v_h) = \lambda_h (u_h, v_h)_\rho$$

ou encore du problème matriciel

$$KU = \lambda_h MU$$

**Remarque 7** On sait déjà que ces valeurs propres sont strictement positives puisque les matrices  $K$  et  $M$  sont définies positives.

**Remarque 8** Les matrices  $M$  et  $K$  sont de la forme:

$$K = \frac{1}{h} K_0 \quad \text{où } K_0 = O(1); \quad M = h M_0 \quad \text{où } M_0 = O(1)$$

La condition CFL revient donc à chercher les valeurs propres  $\lambda_0$  de  $M_0^{-1} K_0$  et peut se réécrire sous la forme

$$(121) \quad \frac{\Delta t^2}{4h^2} \|M_0^{-1} K_0\|_2^2 < 1$$

**Lemme 11** Le schéma variationnel (??) est stable si les paramètres de discrétisation vérifient la condition (CFL) (ou encore la condition (121)).

**Condition suffisante pour le schéma avec condensation, en milieu hétérogène.** La condition (CFL) qu'on vient d'écrire paraît un peu abstraite... En pratique, pour le schéma avec condensation, la condition de stabilité est l'extension naturelle de celle obtenue en milieu homogène:

**Lemme 12**

Le schéma  $P^1$  avec condensation est stable sous la condition

$$(122) \quad \frac{c^* \Delta t}{h} < 1$$

où  $c^*$  représente la vitesse maximum des ondes qui se propagent dans le milieu. Plus précisément:

$$c^* = \max_j \sqrt{\frac{\mu_j}{\rho_j}}$$

et où  $\mu_j = (\mu_{j-1/2} + \mu_{j+1/2})/2$  et  $\rho_j = (\rho_{j-1/2} + \rho_{j+1/2})/2$ .

**Démonstration:**

On ne montre que le caractère suffisant de cette condition. Il est facile de montrer que

$$(123) \quad (KU, U) = \frac{1}{h} \sum_j \mu_{j+1/2} |U_{j+1} - U_j|^2$$

En effet, on a

$$\begin{aligned} (KU, U) &= \frac{1}{h} \sum_j \left( -\mu_{j-1/2} U_{j-1} U_j + (\mu_{j-1/2} + \mu_{j+1/2}) U_j^2 - \mu_{j+1/2} U_{j+1} U_j \right) \\ &= \frac{1}{h} \left( \sum_j \mu_{j-1/2} U_j (U_j - U_{j-1}) + \sum_j \mu_{j+1/2} U_j (U_j - U_{j+1}) \right) \end{aligned}$$

d'où l'expression (123). D'autre part, pour le schéma avec condensation, on a

$$(124) \quad (MU, U) = h \sum_j \underbrace{\frac{\rho_{j+1/2} + \rho_{j-1/2}}{2}}_{=\rho_j} |U_j|^2$$

Ce qui montre que

$$C_{cfl} \equiv \frac{\Delta t^2}{4} \sup_{V \neq 0} \frac{(KV, V)}{(MV, V)} = \frac{\Delta t^2}{4h^2} \frac{\sum_j \mu_{j+1/2} |U_{j+1} - U_j|^2}{\sum_j \rho_j |U_j|^2}$$

En utilisant que  $(a - b)^2 \leq 2(a^2 + b^2)$ , on peut majorer cette quantité par

$$C_{cfl} \leq \frac{\Delta t^2}{4h^2} \frac{2 \sum_j \mu_{j+1/2} (|U_{j+1}|^2 + |U_j|^2)}{\sum_j \rho_j |U_j|^2} \leq \frac{\Delta t^2}{4h^2} \frac{4 \sum_j \mu_j |U_j|^2}{\sum_j \rho_j |U_j|^2}$$

où  $\mu_j = (\mu_{j-1/2} + \mu_{j+1/2})/2$ . Maintenant, si on pose

$$c_j^2 = \frac{\mu_j}{\rho_j} \leq (c^*)^2$$

on voit que

$$C_{cfl} \leq \frac{\Delta t^2}{h^2} \frac{\sum_j \rho_j c_j^2 |U_j|^2}{\sum_j \rho_j |U_j|^2} \leq \frac{(c^*)^2 \Delta t^2}{h^2}$$

Par conséquent si on impose que  $c^* \Delta t/h < 1$  la condition (CFL) sera bien satisfaite. ■

**Application en milieu homogène et sur maillage uniforme.** En se plaçant dans un milieu homogène, et en calculant les valeurs propres, on peut ainsi déterminer les conditions CFL des schémas (et retrouver la CFL du schéma aux différences finies):

**Lemme 13**

• Le schéma  $P^1$  sans condensation est stable sous la condition

$$(125) \quad |\alpha| < \frac{\sqrt{3}}{3}$$

• Le schéma  $P^1$  avec condensation est stable sous la condition

$$(126) \quad |\alpha| < 1$$



Pour le schéma avec condensation, qui est identique au schéma aux différences finies en milieu homogène, on retrouve la CFL qu'on avait déjà vue. La condition de stabilité est plus contraignante pour le schéma sans condensation. Notons que ces conditions ne sont que suffisantes, ce sont en fait les CFL du problème posé en domaine infini.

**Démonstration:** on indique uniquement la démonstration du premier point, l'autre étant similaire. Rappelons que en milieu homogène on a

$$(MU)_i = \rho h \left( \frac{2}{3}U_i + \frac{1}{6}U_{i-1} + \frac{1}{6}U_{i+1} \right)$$

$$(KU)_i = -\frac{\mu}{h}(U_{i+1} - 2U_i + U_{i-1})$$

Le problème aux valeurs propres  $KU = \lambda MU$  revient donc à trouver  $\lambda$  et  $U \neq 0$  tels que

$$-c^2 \frac{U_{i+1} - 2U_i + U_{i-1}}{h^2} = \lambda \left( \frac{2}{3}U_i + \frac{1}{6}U_{i-1} + \frac{1}{6}U_{i+1} \right)$$

$$U_0 = U_{N+1} = 0$$

Posons

$$\nu = \frac{\lambda h^2}{6c^2}$$

le schéma devient:

$$(1 + \nu)U_{i+1} + 2(2\nu - 1)U_i + (1 + \nu)U_{i-1} = 0$$

On cherche la solution sous la forme

$$U_j = Ar_1^j + Br_2^j$$

et on est ramené à chercher les racines du polynome caractéristique

$$(1 + \nu)r^2 + 2(2\nu - 1)r + 1 + \nu = 0$$

Ces deux racines vérifient aussi:

$$r_1 r_2 = 1, \quad \text{et } r_1 + r_2 = -\frac{2(2\nu - 1)}{1 + \nu}$$

Les conditions aux limites imposent alors pour avoir une solution non nulle que

$$B = -A \neq 0, \quad \text{et } r_1^{N+1} = r_2^{N+1}$$

Si on cherche  $r_1 = \rho_1 e^{iz_1}$  et  $r_2 = \rho_2 e^{iz_2}$ , on doit donc avoir

$$\rho_1 e^{iz_1(N+1)} = \rho_2 e^{iz_2(N+1)}$$

$$\rho_1 \rho_2 e^{i(z_1+z_2)} = 1$$

$$\rho_1 e^{iz_1} + \rho_2 e^{iz_2} = -\frac{2(2\nu - 1)}{1 + \nu}$$

Les deux premières relations donnent finalement

$$z_1 = \frac{m\pi}{N+1} = \frac{m\pi h}{L}, \quad r_1 = e^{im\pi/(N+1)}$$

et comme fonctions propres:

$$(U_j)^{(m)} = 2iA \sin(m\pi j/(N+1)), \quad m = 1, N$$

Les valeurs propres sont alors données par la troisième relation

$$\rho_1 e^{iz_1} + \rho_2 e^{iz_2} = -\frac{2(2\nu-1)}{1+\nu}$$

soit:

$$(1+\nu)2 \cos z_1 = 2(1-2\nu) \Rightarrow \nu = \frac{1 - \cos z_1}{2 + \cos z_1}$$

c'est à dire

$$\lambda_m = \frac{6c^2}{h^2} \nu_m = \frac{6c^2}{h^2} \frac{1 - \cos \frac{m\pi}{N+1}}{2 + \cos \frac{m\pi}{N+1}}$$

La fonction

$$x \in [-1, 1] \rightarrow \frac{1-x}{2+x}$$

est strictement décroissante: le max est donc obtenu pour la plus petite valeur du cos c'est à dire pour  $m = N$  et cette valeur est strictement inférieure à celle obtenue pour  $x = -1$  c'est à dire 2:

$$\max \lambda_m = \lambda_N = \frac{6c^2}{h^2} \frac{1 - \cos \frac{N\pi}{N+1}}{2 + \cos \frac{N\pi}{N+1}} < \frac{12c^2}{h^2}$$

Donc la condition CFL est vérifiée si

$$\frac{\Delta t^2}{4} \frac{6c^2}{h^2} \frac{1 - \cos \frac{N\pi}{N+1}}{2 + \cos \frac{N\pi}{N+1}} < 1$$

qui est a fortiori vérifiée si on demande

$$\frac{\Delta t^2}{4} \frac{12c^2}{h^2} < 1 \iff \frac{\sqrt{3}c\Delta t}{h} < 1$$

En fait la condition exacte s'écrit:

$$\begin{aligned} \frac{3c^2 \Delta t^2}{2h^2} \frac{(1 - \cos \pi(1 - \frac{h}{L}))}{(2 + \cos \pi(1 - \frac{h}{L}))} &= \frac{3c^2 \Delta t^2}{2h^2} \frac{(1 + \cos \pi \frac{h}{L})}{(2 - \cos \pi \frac{h}{L})} \\ &= \frac{3c^2 \Delta t^2}{2h^2} \frac{(2 - \frac{\pi^2 h^2}{L^2} + O(h^4))}{(1 + \frac{\pi^2 h^2}{L^2} + O(h^4))} \\ &= \frac{3c^2 \Delta t^2}{h^2} (2 - 3 \frac{\pi^2 h^2}{L^2}) + O(h^4) \end{aligned}$$

La condition donnée en (125) est donc une approximation de la CFL exacte à l'ordre 2. C'est la condition CFL lorsqu'on se trouve sur un domaine infini (ou semi-infini, faire  $L \rightarrow +\infty$ ).

### 5.5.1 Analyse de dispersion et de stabilité des deux schémas en milieu homogène et avec maillage uniforme

**Relation de dispersion.** Pour le schéma avec condensation, on va retrouver l'étude de dispersion faite pour le schéma aux DF puisque ces deux schémas coïncident en milieu homogène. La relation de dispersion est donc:

$$\frac{4}{\Delta t^2} \sin^2 \frac{\omega \Delta t}{2} = \frac{4c^2}{h^2} \sin^2 \frac{kh}{2}$$

Le schéma  $P^1$  sans condensation s'écrit en milieu homogène:

$$\begin{aligned} & \frac{2}{3} \frac{u_i^{n+1} - 2u_i^n + u_i^{n-1}}{\Delta t^2} + \frac{1}{6} \frac{u_{i+1}^{n+1} - 2u_{i+1}^n + u_{i+1}^{n-1}}{\Delta t^2} \\ & + \frac{1}{6} \frac{u_{i-1}^{n+1} - 2u_{i-1}^n + u_{i-1}^{n-1}}{\Delta t^2} = c^2 \frac{u_{i+1}^n - 2u_i^n + u_{i-1}^n}{h^2} \end{aligned}$$

d'où la relation de dispersion:

$$\left( \frac{2}{3} + \frac{1}{6} e^{-ikh} + \frac{1}{6} e^{ikh} \right) \frac{e^{i\omega \Delta t} - 2 + e^{-i\omega \Delta t}}{\Delta t^2} = c^2 \frac{e^{ikh} - 2 + e^{-ikh}}{h^2}$$

soit

$$\frac{4}{\Delta t^2} \sin^2 \frac{\omega \Delta t}{2} = \frac{4c^2}{h^2} \sin^2 \frac{kh}{2} \frac{1}{1 - \frac{2}{3} \sin^2 \frac{kh}{2}}$$

**Stabilité.** On retrouve la condition de stabilité du schéma  $P^1$ , en exprimant que pour tout  $k$  on doit avoir  $\omega(k) \in \mathbb{R}$ , i.e.

$$0 \leq \frac{c^2 \Delta t^2}{h^2} \sin^2 \frac{kh}{2} \frac{1}{1 - \frac{2}{3} \sin^2 \frac{kh}{2}} \leq 1, \quad \forall k$$

Soit  $F$  la fonction définie dans  $[0, 1]$  par:

$$F(x) = \frac{\alpha^2 x}{1 - 2/3x}$$

$F$  est une fonction strictement croissante

$$0 \leq F(x) \leq F(1) = 3\alpha^2$$

Par conséquent la CFL s'écrit:

$$3\alpha^2 \leq 1 \iff |\alpha| \leq \frac{\sqrt{3}}{3}$$

**Courbes de dispersion.** On rappelle que pour le schéma aux différences finies, on avait: ( $K = kh$ ,  $G = K/(2\pi)$ ),

$$\begin{aligned} q_{DF} &= \frac{2}{\alpha K} \arcsin\left(\alpha \sin \frac{K}{2}\right) = \frac{1}{\alpha \pi G} \arcsin(\alpha \sin \pi G) \\ &= 1 - \frac{1}{24}(1 - \alpha^2)K^2 + O(K^4) \end{aligned}$$

La condition de stabilité imposait de prendre  $\alpha \leq 1$ , la meilleure courbe de dispersion étant obtenue pour  $\alpha = 1$  pour laquelle le schéma est exact.

Pour le schéma  $P^1$  sans condensation, on a :

$$\begin{aligned} q_{P^1} &= \frac{2}{\alpha K} \arcsin \left( \alpha \frac{\sin \frac{K}{2}}{\sqrt{1 - \frac{2}{3} \sin^2 \frac{K}{2}}} \right) \\ &= 1 + \frac{1}{24}(1 + \alpha^2)K^2 + O(K^4) \end{aligned}$$

Comme pour le schéma semi-discrétisé on remarque que la vitesse est approchée par le haut alors qu'elle l'était par le bas pour le schéma aux différences finies. D'autre part on voit sur le développement de Taylor que plus  $\alpha$  est grand, plus la vitesse numérique s'éloigne de la vitesse continue. La meilleure valeur est donc obtenue pour  $\alpha = 0$ , pour laquelle on a :

$$q_{P^1}(\alpha = 0) = \frac{2}{\alpha K} \frac{\sin \frac{K}{2}}{\sqrt{1 - \frac{2}{3} \sin^2 \frac{K}{2}}}$$

qui est différent de 1. La meilleure courbe obtenue pour  $P^1$  est donc moins bonne que la meilleure obtenue avec les différences finies.

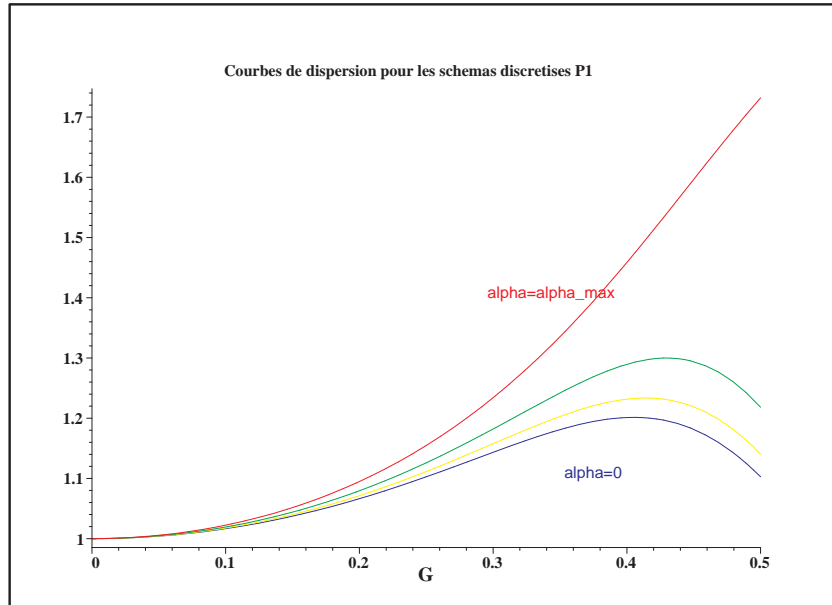


Figure 16: Courbes de dispersion du schéma totalement discrétisé sans condensation, pour  $\alpha$  variant de 0 à  $\sqrt{3}/3$  ( $\alpha = 0, 0.25, 0.4, \sqrt{3}/3$ ).

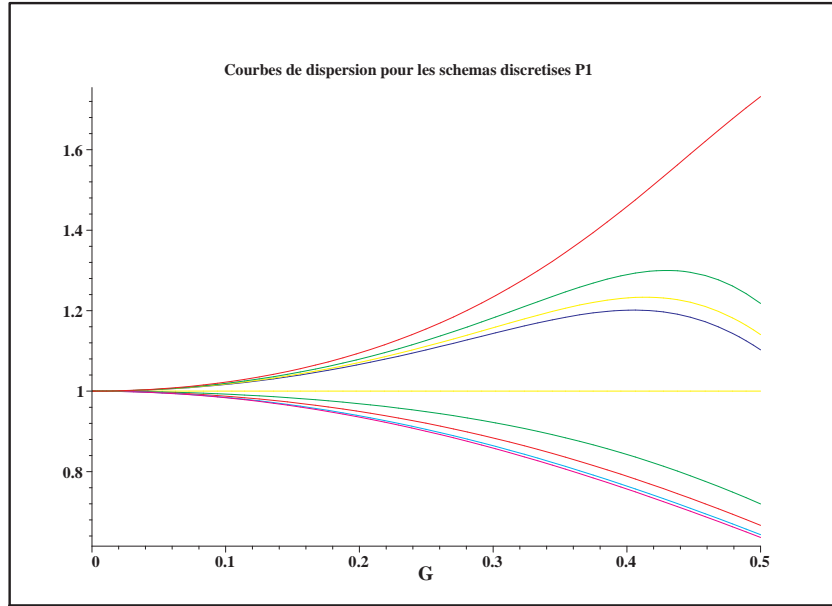


Figure 17: Comparaison des courbes de dispersion du schéma totalement discrétisé avec et sans condensation.

## 5.6 Convergence du schéma totalement discrétisé

On indique les étapes pour un problème assez général :

**Le problème continu:**

$$(P) \begin{cases} \text{Trouver } u \in H^1(0, T; H) \cap L^2(0, T; V) \\ \frac{d^2}{dt^2}(u, v) + a(u, v) = 0, & \forall v \in V \\ u(0) = u_0 \\ \frac{du}{dt}(0) = u_1 \end{cases}$$

On suppose que  $a(\cdot, \cdot)$  est coercive sur  $V$ .

**Application:**  $V = H_0^1$ ,  $H = L^2$

$$(u, v) = \int_{\Omega} \rho u v dx; \quad a(u, v) = \int_{\Omega} \mu \nabla u \cdot \nabla v dx$$

**Approximation par différences finies en temps et Galerkin en espace:**

$$(P_h) \begin{cases} \text{Trouver } u_h^n \in V_h \\ \left( \frac{u_h^{n+1} - 2u_h^n + u_h^{n-1}}{\Delta t^2}, v_h \right) + a(u_h^n, v_h) = 0, & \forall v_h \in V_h \\ u_h^0 = u_{0h} \\ (u_h^1, v_h) = (u_{0h}, v_h) - \frac{\Delta t^2}{2} a(u_{0h}, v_h) + \Delta t (u_{1h}, v_h), & \forall v_h \in V_h \end{cases}$$

où  $u_{0h}$  et  $u_{1h}$  sont des approximations de  $u_0$  et  $u_1$ . On suppose que  $u$  est très régulière (c'est l'estimation d'erreur qui nous indiquera la régularité nécessaire qu'on doit supposer) et que  $u_0 \in V$  et  $u_1 \in V$ . On pose  $\bar{u}^n = u(t^n)$ .

La différence entre le problème vérifié par la solution exacte ( $P$ ) et le problème approché ( $P_h$ ) donne un schéma vérifié par l'erreur

$$(127) \quad \begin{cases} e_h^n = u_h^n - \bar{u}^n \\ \left( \frac{e_h^{n+1} - 2e_h^n + e_h^{n-1}}{\Delta t^2}, v_h \right) + a(e_h^n, v_h) = (\varepsilon_h^n, v_h), \quad \forall v_h \in V_h \\ e_h^0 = u_h^0 - \bar{u}_h^0, \quad e_h^1 = u_h^1 - \bar{u}_h^1 \end{cases}$$

où

$$\varepsilon_h^n = \frac{\partial^2 u}{\partial t^2}(t^n) - \frac{\bar{u}_h^{n+1} - 2\bar{u}_h^n + \bar{u}_h^{n-1}}{\Delta t^2}$$

L'idée c'est d'obtenir une estimation de l'erreur en fonction de  $\varepsilon_h^n$ , qui est "petit" (de l'ordre de  $\Delta t^2$ ), ce qui permettra de montrer la convergence (et l'ordre d'approximation). L'erreur  $\varepsilon_h^n$  représente une erreur de troncature. Si on pouvait choisir  $v_h = \frac{e_h^{n+1} - e_h^{n-1}}{2\Delta t}$  dans (127), on obtiendrait une estimation d'énergie sur l'erreur et ce serait fini... Mais ce n'est pas possible car  $e_h^n \notin V_h$ . Comme pour le problème semi-discrétisé, on introduit la projection elliptique  $w_h^n = P_h \bar{u}^n \in V_h$  (cf (96)), et on décompose l'erreur sous la forme:

$$e_h^n = \underbrace{u_h^n - w_h^n}_{\delta_h^n \in V_h} + \underbrace{w_h^n - \bar{u}^n}_{-r_h^n}$$

On remarque que maintenant  $\delta_h^n = u_h^n - P_h \bar{u}^n$  est bien dans  $V_h$  et  $r_h^n = \bar{u}^n - P_h \bar{u}^n$  représente l'erreur de projection elliptique qui est "petite". On réécrit le problème (127) sous la forme:

$$\left\{ \begin{array}{l} \left( \frac{\delta_h^{n+1} - 2\delta_h^n + \delta_h^{n-1}}{\Delta t^2}, v_h \right) + a(\delta_h^n, v_h) = (\varepsilon_h^n, v_h) + \left( \frac{r_h^{n+1} - 2r_h^n + r_h^{n-1}}{\Delta t^2}, v_h \right) + a(r_h^n, v_h), \quad \forall v_h \in V_h \\ e_h^0 = u_h^0 - \bar{u}_h^0, \quad e_h^1 = u_h^1 - \bar{u}_h^1 \end{array} \right.$$

Par définition de la projection elliptique, on a

$$a(r_h^n, v_h) = 0, \quad \forall v_h \in V_h$$

ce qui permet finalement de réécrire le problème sous la forme:

$$(128) \quad \left\{ \begin{array}{l} \left( \frac{\delta_h^{n+1} - 2\delta_h^n + \delta_h^{n-1}}{\Delta t^2}, v_h \right) + a(\delta_h^n, v_h) = (\varepsilon_h^n, v_h) + \left( \frac{r_h^{n+1} - 2r_h^n + r_h^{n-1}}{\Delta t^2}, v_h \right), \quad \forall v_h \in V_h \\ e_h^0 = u_h^0 - \bar{u}_h^0, \quad e_h^1 = u_h^1 - \bar{u}_h^1 \end{array} \right.$$

En choisissant  $v_h = \frac{\delta_h^{n+1} - \delta_h^{n-1}}{2\Delta t}$ , on obtient:

$$\frac{E_h^{n+1/2} - E_h^{n-1/2}}{\Delta t} = (\varepsilon_h^n, \frac{\delta_h^{n+1} - \delta_h^{n-1}}{2\Delta t}) + \left( \frac{r_h^{n+1} - 2r_h^n + r_h^{n-1}}{\Delta t^2}, \frac{\delta_h^{n+1} - \delta_h^{n-1}}{2\Delta t} \right)$$

avec

$$E_h^{n+1/2} = \frac{1}{2} \left\| \frac{\delta_h^{n+1} - \delta_h^n}{\Delta t} \right\|^2 + \frac{1}{2} a(\delta_h^n, \delta_h^{n+1})$$

À partir de là, les étapes sont les suivantes;

1. On suppose la condition CFL:

$$(129) \quad \alpha^2 = \frac{\Delta t^2}{4} \sup_{v_h \in V_h} \frac{a(v_h, v_h)}{\|v_h\|^2} < 1$$

Sous cette condition on montre que

$$(130) \quad E_h^{n+1/2} \geq \frac{1}{2}(1 - \alpha^2) \left\| \frac{\delta_h^{n+1} - \delta_h^n}{\Delta t} \right\|^2$$

2. A partir de l'identité d'énergie:

$$\frac{E_h^{n+1/2} - E_h^{n-1/2}}{\Delta t} \leq \underbrace{\left( \|\varepsilon_h^n\| + \left\| \frac{r_h^{n+1} - 2r_h^n + r_h^{n-1}}{\Delta t^2} \right\| \right)}_{=\mu_n} \left\| \frac{\delta_h^{n+1} - \delta_h^{n-1}}{2\Delta t} \right\|$$

or en utilisant Cauchy-Schwartz et (130) on a

$$\begin{aligned} \left\| \frac{\delta_h^{n+1} - \delta_h^{n-1}}{2\Delta t} \right\| &\leq \frac{1}{2} \left\| \frac{\delta_h^{n+1} - \delta_h^n}{\Delta t} \right\| + \frac{1}{2} \left\| \frac{\delta_h^n - \delta_h^{n-1}}{\Delta t} \right\| \\ &\leq \frac{1}{2} \frac{\sqrt{2}}{\sqrt{1 - \alpha^2}} \sqrt{E_h^{n+1/2}} + \frac{1}{2} \frac{\sqrt{2}}{\sqrt{1 - \alpha^2}} \sqrt{E_h^{n-1/2}} \end{aligned}$$

ce qui implique que

$$\begin{aligned} E_h^{n+1/2} - E_h^{n-1/2} &\leq \Delta t \mu_n \frac{\sqrt{2}}{2\sqrt{1 - \alpha^2}} (\sqrt{E_h^{n+1/2}} + \sqrt{E_h^{n-1/2}}) \\ \implies (\sqrt{E_h^{n+1/2}} + \sqrt{E_h^{n-1/2}})(\sqrt{E_h^{n+1/2}} - \sqrt{E_h^{n-1/2}}) &\leq \Delta t \mu_n \frac{\sqrt{2}}{2\sqrt{1 - \alpha^2}} (\sqrt{E_h^{n+1/2}} + \sqrt{E_h^{n-1/2}}) \\ \implies \sqrt{E_h^{n+1/2}} &\leq \sqrt{E_h^{n-1/2}} + \Delta t \mu_n \frac{\sqrt{2}}{2\sqrt{1 - \alpha^2}} \end{aligned}$$

d'où finalement:

$$(131) \quad \sqrt{E_h^{n+1/2}} \leq \sqrt{E_h^{1/2}} + \Delta t \frac{\sqrt{2}}{2\sqrt{1 - \alpha^2}} \sum_{m=1}^n \mu_m$$

D'après la définition de  $\mu_m$ , le terme de droite contient deux erreurs: l'erreur de troncature  $\Delta t \sum_{m=1}^n \|\varepsilon_h^m\|$  et l'erreur de projection elliptique  $\Delta t \sum_{m=1}^n \left\| \frac{r_h^{m+1} - 2r_h^m + r_h^{m-1}}{\Delta t^2} \right\|$  que nous estimons dans les deux prochains points.

3. **Erreur de troncature.** Sous l'hypothèse que  $u \in C^4(0, T; L^2)$  on montre que

$$(132) \quad \Delta t \sum_{m=1}^n \|\varepsilon_h^m\| \leq Ct^n \Delta t^2 \sup_{[0, T]} \left\| \frac{\partial^4 u}{\partial t^4}(s) \right\|$$

4. **Erreur de projection elliptique.** On rappelle que  $r_h^n = \bar{u}^n - P_h \bar{u}^n$ . On montre que si  $u \in C^4(0, T; V)$ , on a

$$(133) \quad \Delta t \sum_{m=1}^n \left\| \frac{r_h^{m+1} - 2r_h^m + r_h^{m-1}}{\Delta t^2} \right\| \leq Ct^n \left( \sup_{[0, T]} \left\| (I - P_h) \frac{\partial^2 u}{\partial t^2}(s) \right\| + \Delta t^2 \left\| (I - P_h) \frac{\partial^4 u}{\partial t^4}(s) \right\| \right)$$

5. **Conclusion.** En regroupant tous ces résultats, on obtient finalement:

$$(134) \quad \sqrt{E_h^{n+1/2}} \leq \sqrt{E_h^{1/2}} + \frac{Ct^n}{\sqrt{1-\alpha^2}} \left( \Delta t^2 \sup_{[0,T]} \left\| \frac{\partial^4 u}{\partial t^4}(s) \right\| + \sup_{[0,T]} \left\| (I - P_h) \frac{\partial^2 u}{\partial t^2}(s) \right\| + \Delta t^2 \left\| (I - P_h) \frac{\partial^4 u}{\partial t^4}(s) \right\| \right)$$

Si on utilise par exemple des éléments finis  $P^1$ , on rappelle que si  $v \in H^2$  on a

$$\|(I - P_h)v\|_V = O(h), \quad \|(I - P_h)v\|_{L^2} = O(h^2)$$

donc en supposant que  $\partial_t^2 u \in H^2$  et  $\partial_t^4 u \in H^2$ , on obtient une erreur d'ordre 2 (à condition que  $\sqrt{E_h^{1/2}}$  soit d'ordre 2, c'est à dire l'erreur sur les conditions initiales).



## 6 Analyse de dispersion et de stabilité de schémas en 2D

### 6.1 Formulation variationnelle

On considère l'équation des ondes en milieu hétérogène posée dans un domaine  $\Omega \subset \mathbb{R}^2$

$$(135) \quad \begin{cases} \rho \frac{\partial^2 u}{\partial t^2} - \operatorname{div}(\mu \nabla u) = f, & \text{dans } \Omega \\ u(x, y, 0) = u_0(x, y), \quad \partial_t u(x, y, 0) = u_1(x, y) \\ u(x, y, t) = 0 \quad \text{sur } \partial\Omega \end{cases}$$

On fait les mêmes hypothèses (87) sur les coefficients et sur les conditions initiales que en 1D (en remplaçant  $x$  par  $(x, y)$  et  $(0, L)$  par  $\Omega$ ). On pose

$$V = H_0^1(\Omega), \quad H = L^2(\Omega)$$

La formulation variationnelle de ce problème s'écrit:

$$(136) \quad \begin{cases} \frac{d^2}{dt^2}(u(t), v)_\rho + \int_\Omega \mu \nabla u(t) \cdot \nabla v dx = \int_\Omega f v dx, \quad \forall v \in V \\ u(0) = u_0 ; \frac{du}{dt}(0) = u_1 \end{cases}$$

### 6.2 Approximation par différences finies

**Le schéma.** On peut vérifier que le schéma aux différences finies centrées d'ordre 2 est équivalent à utiliser le schéma  $P^1$  avec condensation, en utilisant comme formule de quadrature:

$$\int_K f(x) dx = \frac{|K|}{3} \sum_{i=1}^3 f(S_i)$$

où  $S_i$  sont les sommets du triangle. Le schéma s'écrit alors, en milieu homogène:

$$\frac{u_{ij}^{n+1} - 2u_{ij}^n + u_{ij}^{n-1}}{\Delta t^2} - \frac{c^2}{h^2} (u_{i+1j}^n - 2u_{ij}^n + u_{i-1j}^n + u_{ij+1}^n - 2u_{ij}^n + u_{ij-1}^n) = 0$$

et sa relation de dispersion s'écrit:

$$\frac{4}{\Delta t^2} \sin^2 \frac{\omega \Delta t}{2} = \frac{4c^2}{h^2} \left( \sin^2 \frac{k_x h}{2} + \sin^2 \frac{k_y h}{2} \right)$$

**Stabilité.** Le schéma est stable si  $\forall (k_x, k_y)$ , la solution  $\omega$  est réelle i.e., si  $\forall (k_x, k_y)$

$$0 \leq \alpha^2 \left( \sin^2 \frac{k_x h}{2} + \sin^2 \frac{k_y h}{2} \right) \leq 1$$

Or

$$\sin^2 \frac{k_x h}{2} + \sin^2 \frac{k_y h}{2} \leq 2, \quad \forall (k_x, k_y)$$

et pour  $k_x = k_y = \pi/h$ ,  $\sin^2 \frac{k_x h}{2} + \sin^2 \frac{k_y h}{2} = 2$  donc

$$\min \sin^2 \frac{k_x h}{2} + \sin^2 \frac{k_y h}{2} = 0, \quad \text{et} \quad \max \sin^2 \frac{k_x h}{2} + \sin^2 \frac{k_y h}{2} = 2$$

La condition de stabilité s'écrit donc:

$$2\alpha^2 \leq 1 \iff |\alpha| \leq \frac{\sqrt{2}}{2}$$

**Vitesse de phase numérique.** La relation de dispersion donne  $\omega(k_x, k_z, \alpha)$ . On introduit  $\theta$  l'angle d'incidence et on exprime la vitesse de phase en fonction de  $K = |k|h$  avec  $|k| = \sqrt{k_x^2 + k_z^2}$  et de  $\theta = \arctan(k_z/k_x)$  (i.e.,  $k = |k|(\cos \theta, \sin \theta)$ ).

$$V(\alpha, K, \theta) = \frac{2c}{\alpha K} \arcsin\left(\alpha \left(\sin^2 \frac{K \cos \theta}{2} + \sin^2 \frac{K \sin \theta}{2}\right)^{1/2}\right)$$

La vitesse numérique dépend de l'angle  $\theta$  ce qui montre que le schéma introduit une anisotropie numérique. On peut faire un développement de Taylor:

$$V(\alpha, K, \theta) = c \left( 1 - \frac{K^2}{24} \left( 1 - \alpha^2 - \frac{1}{2} \sin^2 2\theta \right) + O(K^4) \right)$$

ce qui montre qu'on a une approximation d'ordre 2 de la vitesse. Contrairement au cas 1D, le schéma n'est plus exact pour  $\alpha = 1$  (ou pour toute autre valeur de  $\alpha$ , cf aussi exercice ? sur l'erreur de troncature en 2D). Pour  $K$  assez petit, on voit que la vitesse est approchée par le bas ( $V(\alpha, K, \theta) \leq c$ ).

Pour  $\alpha = \frac{\sqrt{2}}{2}$ , (valeur maximale pour stabilité) et  $\theta = \pi/4$ , on obtient la vitesse exacte

$$V\left(\alpha = \frac{\sqrt{2}}{2}, K, \theta = \pi/4\right) = c$$

ce qui montre que, pour cette valeur de  $\alpha$  le schéma est exact dans les directions diagonales. Pour  $\alpha$  fixé, la dispersion décroît pour  $\theta$  variant de 0 à  $\pi/4$ , ce qui confirme que l'erreur est plus mauvaise dans les directions du maillage et est meilleure dans les directions diagonales. On représente:

- les courbes de dispersion: sur chaque figure, on fixe l'angle de propagation  $\theta$  (de 0 à  $\pi/4$ ), et on représente les courbes de dispersion obtenues pour différentes valeurs du paramètre

$$\gamma = \frac{\alpha}{\alpha_{max}}$$

ici  $\alpha_{max} = \frac{\sqrt{2}}{2}$ . Les courbes sont tracées en fonction de  $G = K/(2\pi)$ , i.e.,

$$q_{DF}^{2D}(\alpha, G, \theta) = \frac{1}{\alpha \pi G} \arcsin\left(\alpha \left(\sin^2 \pi G \cos \theta + \sin^2 \pi G \sin \theta\right)^{1/2}\right)$$

- les courbes d'anisotropie: on fixe  $\gamma = \frac{\alpha}{\alpha_{max}}$ , et on représente la vitesse de phase numérique en fonction de  $\theta$ , pour différentes valeurs de  $N = 1/G$ , le nombre de points par longueur d'ondes. On représente ces courbes pour :  $\gamma = 0, 0.5, 1$  et pour  $N = 2, 3, 4, 5, 10, 20, 100$ .

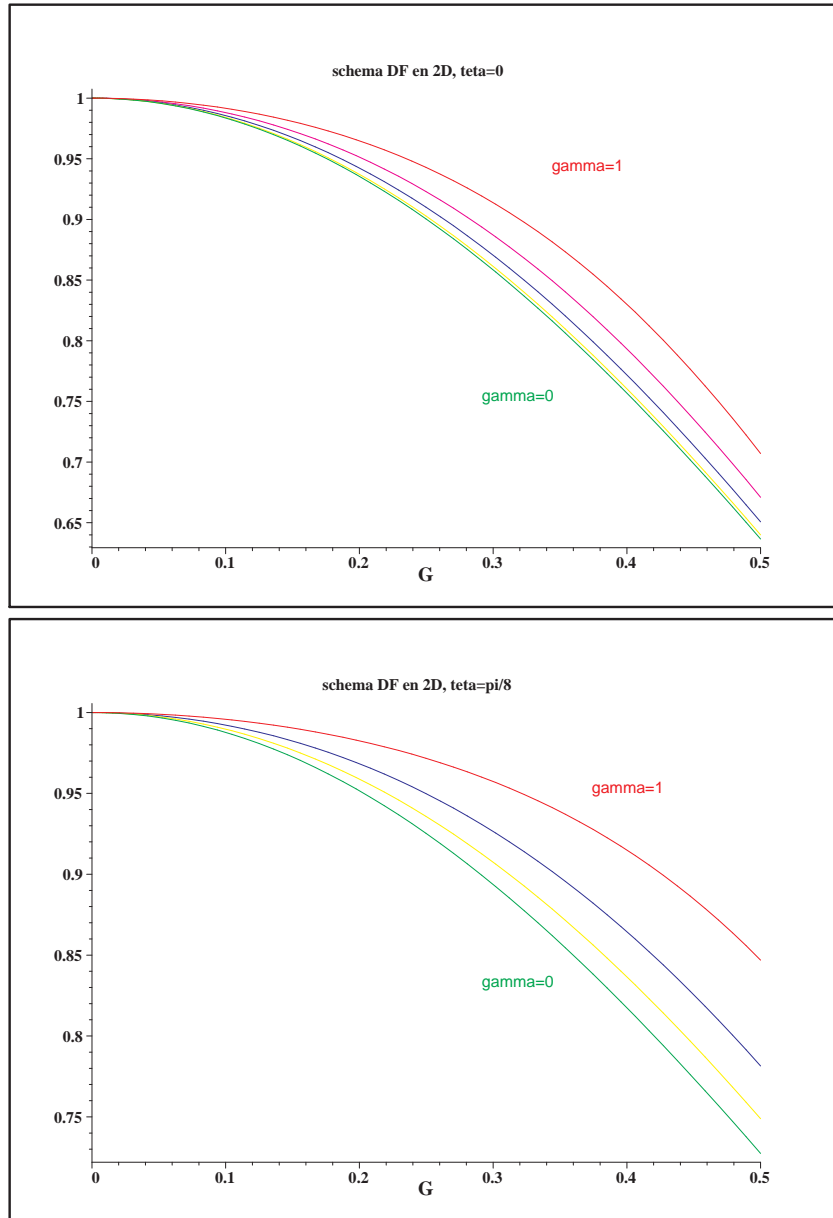


Figure 18: Courbes de dispersion du schéma totalement discrétisé 2D Différences finies,  $\theta = 0$  et  $\theta = \frac{\pi}{8}$ .

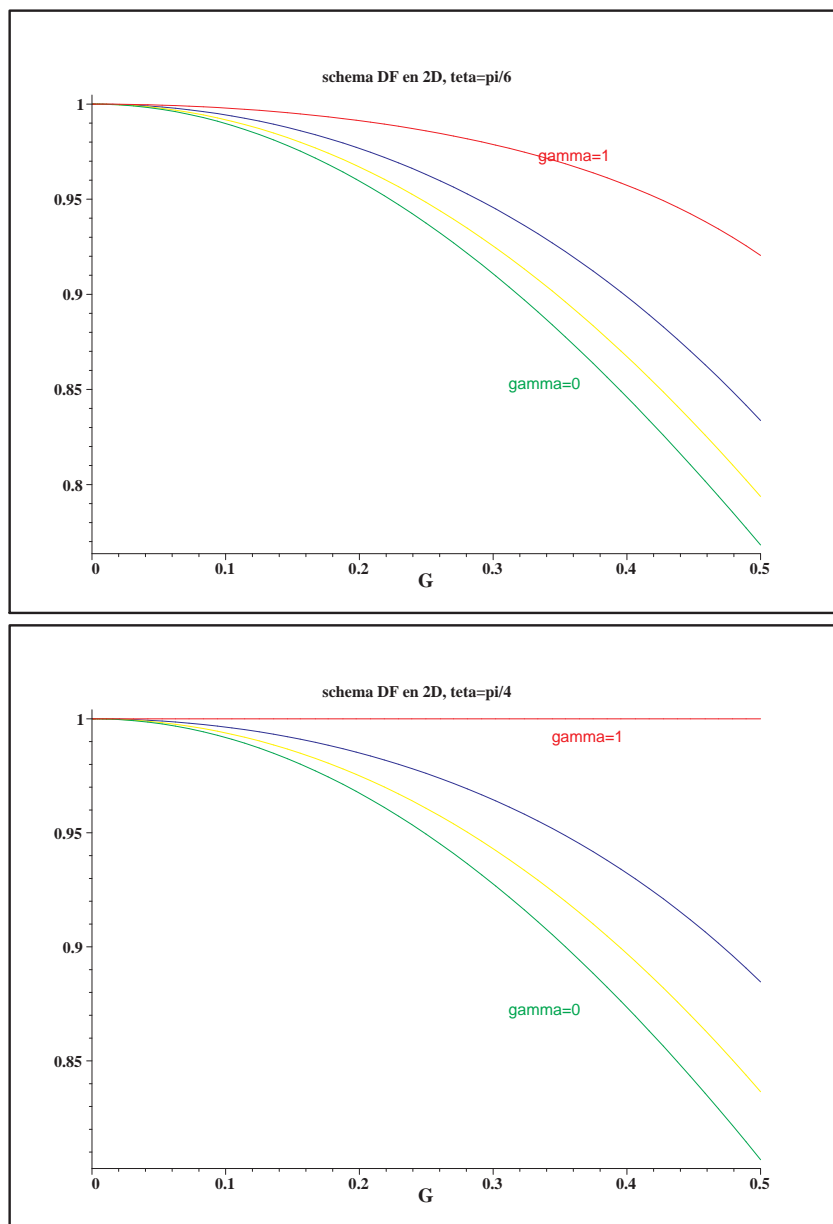
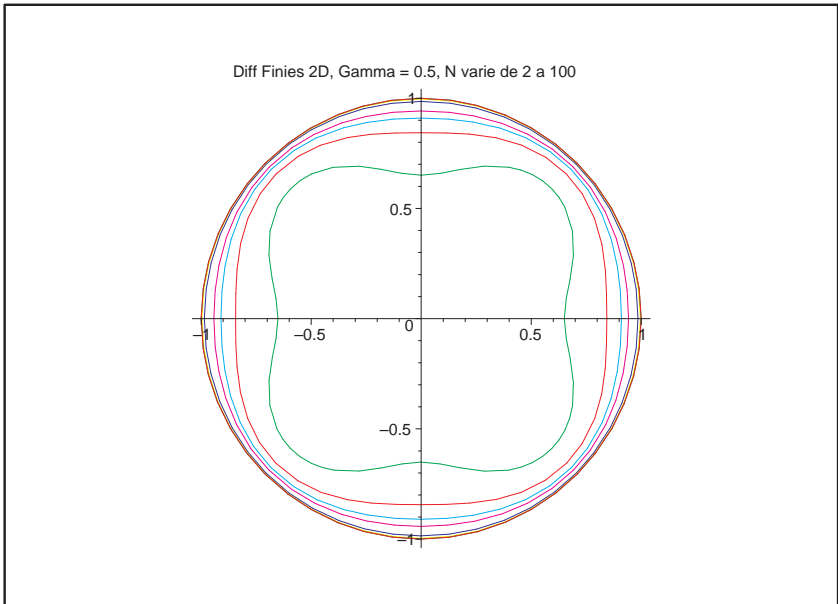
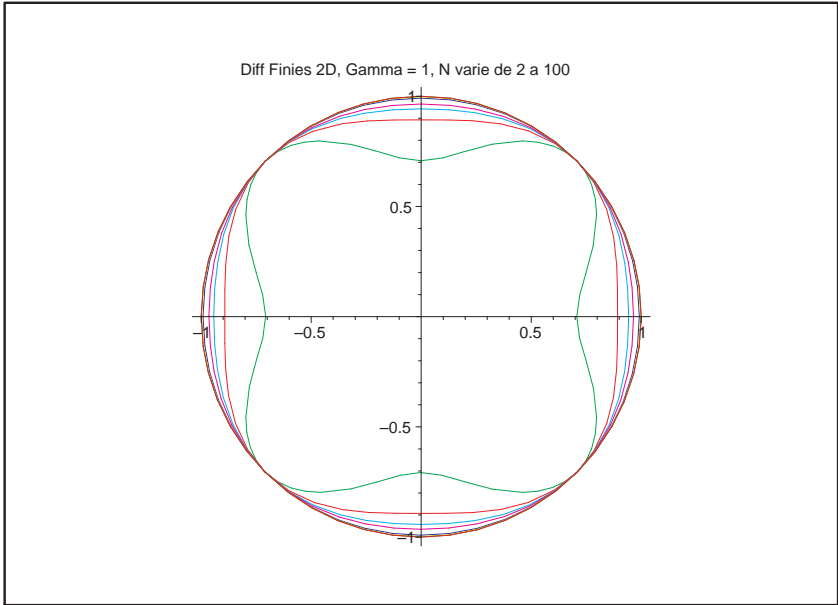


Figure 19: Courbes de dispersion du schéma totalement discrétisé 2D Différences finies,  $\theta = \frac{\pi}{6}$  et  $\theta = \frac{\pi}{4}$ .



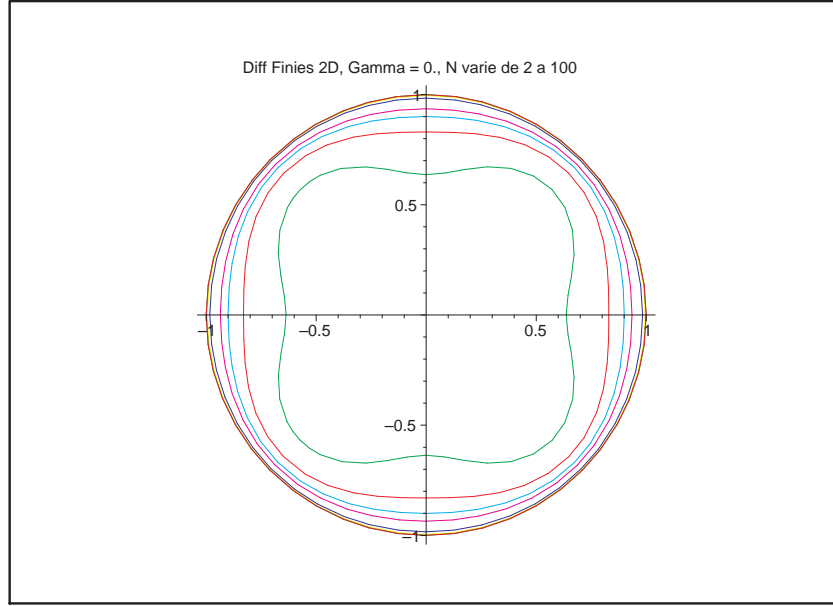


Figure 20: Courbes d'anisotropie du schéma totalement discrétisé 2D Différences finies.

### 6.3 Approximation par éléments finis $Q^1$

**Le schéma.** On suppose de nouveau qu'on fait de la condensation de masse, en utilisant comme formule de quadrature:

$$\int_K f(x) dx = \frac{|K|}{4} \sum_{i=1}^4 f(S_i)$$

où  $S_i$  sont les sommets de l'élément. Le schéma s'écrit alors, en milieu homogène:

$$\frac{u_{ij}^{n+1} - 2u_{ij}^n + u_{ij}^{n-1}}{\Delta t^2} - \frac{c^2}{3h^2} (u_{i+1j+1}^n + u_{i+1j}^n + u_{i+1j-1}^n + u_{i-1j-1}^n + u_{i-1j}^n + u_{i-1j+1}^n + u_{ij+1}^n + u_{ij-1}^n - 8u_{ij}^n) = 0$$

et sa relation de dispersion s'écrit:

$$\frac{4}{\Delta t^2} \sin^2 \frac{\omega \Delta t}{2} = \frac{c^2}{3h^2} (8 - 2 \cos k_x h - 2 \cos k_z h - 2 \cos(k_x + k_z)h - 2 \cos(k_x - k_z)h)$$

**Stabilité.** Le schéma est stable si  $\forall (k_x, k_y)$ , la solution  $\omega$  est réelle i.e., si  $\forall (k_x, k_y)$

$$0 \leq \frac{\alpha^2}{12} (8 - 2 \cos k_x h - 2 \cos k_z h - 2 \cos(k_x + k_z)h - 2 \cos(k_x - k_z)h) \leq 1$$

Posons:

$$F(k_x, k_z) = \frac{\alpha^2}{12} (8 - 2 \cos k_x h - 2 \cos k_z h - 2 \cos(k_x + k_z)h - 2 \cos(k_x - k_z)h)$$

**Lemme 14** On a :

$$\min F(k_x, k_z) = 0, \quad \text{et} \quad \max F(k_x, k_z) = \alpha^2$$

On en déduit aussitôt le

**Théorème 11** *Le schéma  $Q^1$  est stable sous la condition CFL  $|\alpha| \leq 1$ .*

**Dém du lemme:** ça vient de l'identité:

$$\cos(k_x + k_z)h + \cos(k_x - k_z)h = 2 \cos k_x h \cos k_z h$$

d'où

$$F(k_x, k_z) = \frac{\alpha^2}{12} (9 - (1 + 2 \cos k_x h)(1 + 2 \cos k_z h))$$

or on a

$$-1 \leq 1 + 2 \cos k_x h \leq 3$$

d'où

$$-3 \leq (1 + 2 \cos k_x h)(1 + 2 \cos k_z h) \leq 9$$

ce qui montre que

$$0 \leq F(k_x, k_z) \leq \frac{\alpha^2}{12} (9 + 3) = \alpha^2$$

D'autre part  $F(0, 0) = 0$  donc le min est atteint et  $F(0, \pi/h) = \alpha^2$  donc le max est atteint. ■

**Vitesse de phase numérique.** La relation de dispersion donne

$$\omega(k_x, k_z, \alpha) = \frac{2}{\Delta t} \arcsin \sqrt{F(k_x, k_z)}$$

On introduit comme précédemment  $\theta$  l'angle d'incidence et on exprime la vitesse de phase en fonction de  $K = |k|h$  avec  $|k| = \sqrt{k_x^2 + k_z^2}$  et de  $\theta = \arctan(k_z/k_x)$  (i.e.,  $k = |k|(\cos \theta, \sin \theta)$ ).

$$V(\alpha, K, \theta) = \frac{2c}{\alpha K} \arcsin \sqrt{F(\alpha, K, \theta)}$$

avec

$$F(\alpha, K, \theta) = \frac{\alpha^2}{12} (9 - (1 + 2 \cos(K \cos \theta))(1 + 2 \cos(K \sin \theta)))$$

La vitesse numérique dépend de l'angle  $\theta$  ce qui montre que le schéma introduit de nouveau une anisotropie numérique. On peut faire un développement de Taylor:

$$V(\alpha, K, \theta) = c \left( 1 - \frac{K^2}{24} (1 - \alpha^2 + \frac{1}{2} \sin^2 2\theta) + O(K^4) \right)$$

ce qui montre qu'on a une approximation d'ordre 2 de la vitesse. Pour  $K$  assez petit, on voit que la vitesse est approchée par le bas car  $1 - \alpha^2 + \frac{1}{2} \sin^2 2\theta \geq 1 - \alpha^2 \geq 0$  ( $V(\alpha, K, \theta) \leq c$ ). Pour  $\alpha = 1$ , (valeur maximale pour stabilité) et  $\theta = 0$ , on obtient la vitesse exacte

$$V(\alpha = 1, K, \theta = 0) = c$$

ce qui montre que, pour cette valeur de  $\alpha$  le schéma est exact dans les directions du maillage. Pour  $\alpha$  fixé, la fonction  $\theta \in [0, \pi/4] \rightarrow 1 - \alpha^2 + \frac{1}{2} \sin^2 2\theta$  est une fonction croissante ce qui montre que la

dispersion augmente avec l'angle  $\theta$  fait entre la direction de propagation et les directions du maillage. Ceci confirme que l'erreur est plus mauvaise dans les directions diagonales et est meilleure dans les directions du maillage. On représente:

- les courbes de dispersion: sur chaque figure, on fixe l'angle de propagation  $\theta$  (de 0 à  $\pi/4$ ), et on représente les courbes de dispersion obtenues pour différentes valeurs du paramètre

$$\gamma = \frac{\alpha}{\alpha_{max}}$$

ici  $\alpha_{max} = 1$ . Les courbes sont tracées en fonction de  $G = K/(2\pi)$ .

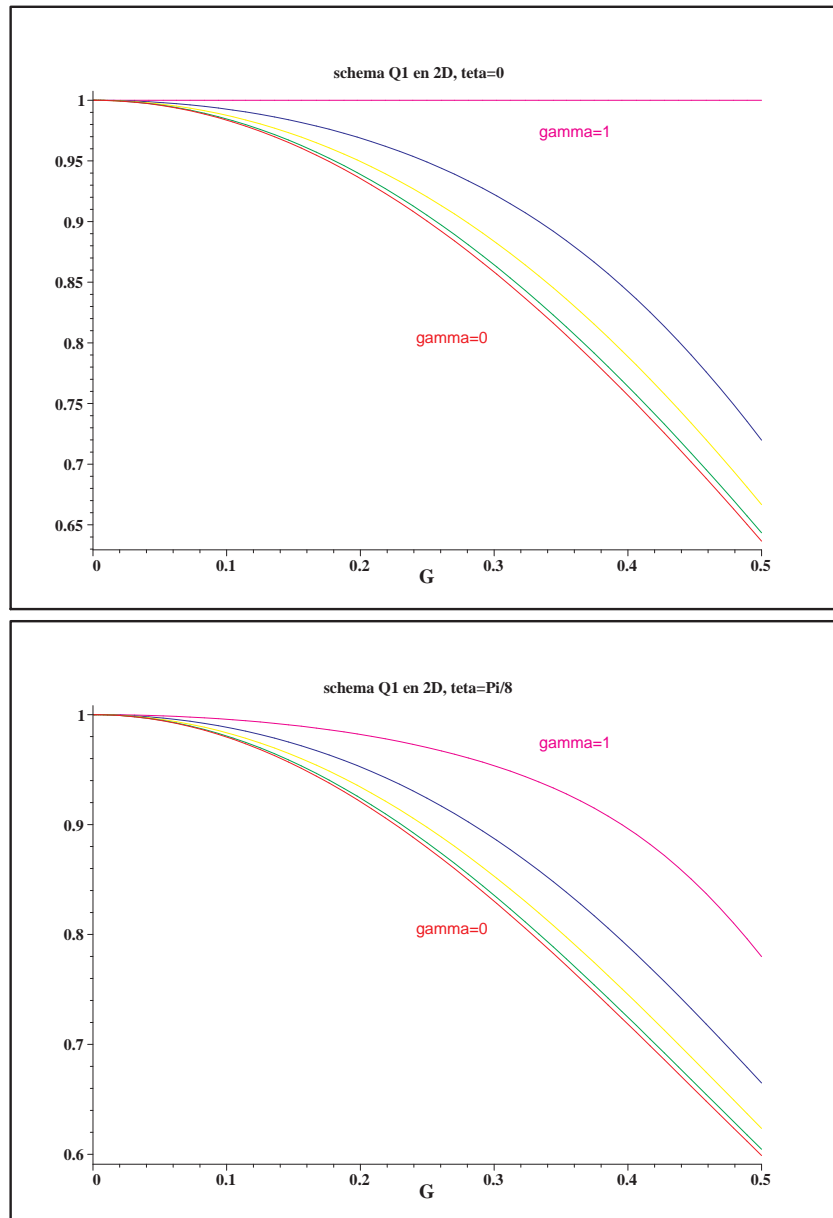


Figure 21: Courbes de dispersion du schéma totalement discrétisé 2D -  $Q^1$ ,  $\theta = 0$  et  $\theta = \frac{\pi}{8}$ .



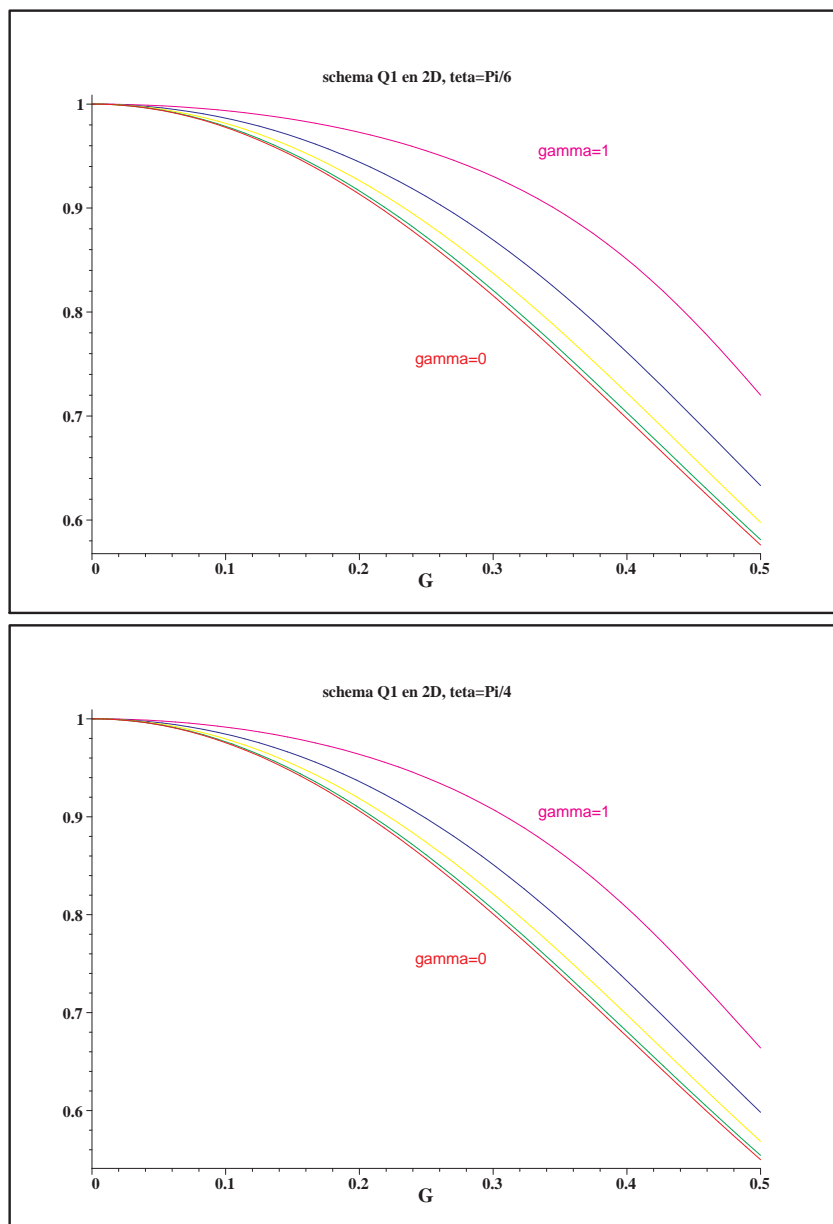
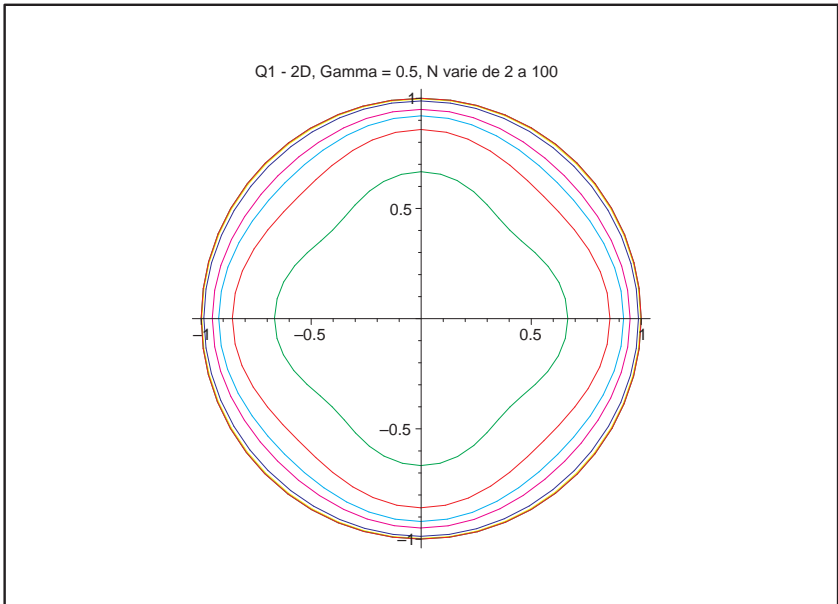
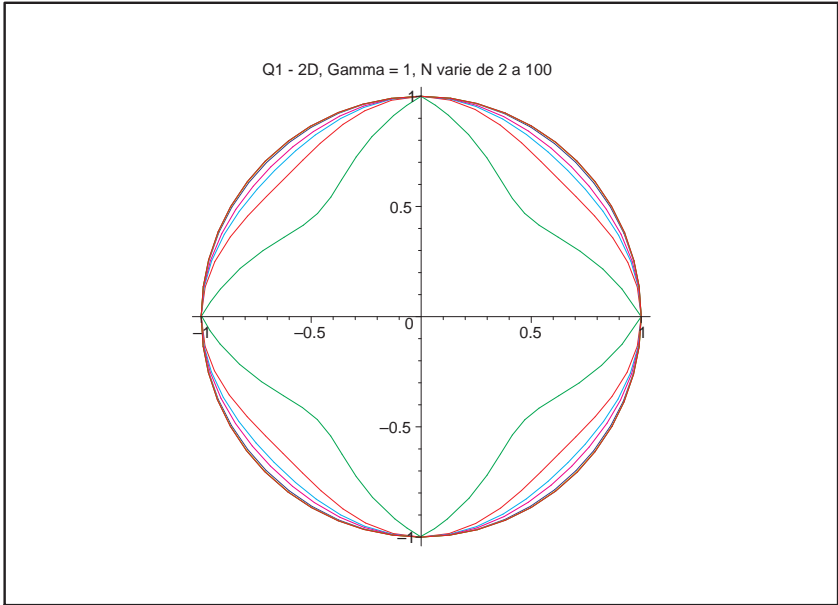


Figure 22: Courbes de dispersion du schéma totalement discrétisé 2D Différences finies,  $\theta = \frac{\pi}{6}$  et  $\theta = \frac{\pi}{4}$ .



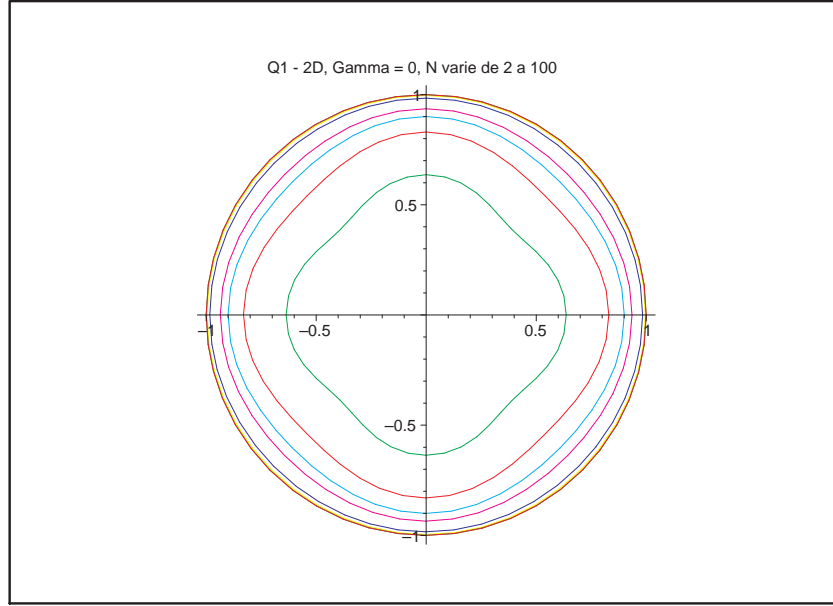


Figure 23: Courbes d'anisotropie du schéma totalement discrétisé 2D  $Q^1$

## Annexe

### A Quelques corrigés

#### Corrigé de l'exercice 5

On démontre le lemme 1 par des intégrations par parties discrètes:

$$\begin{aligned}
 (A_h u_h, v_h) &= h \sum_j \frac{-c^2}{h} (u_{j+1} - 2u_j + u_{j-1}) \bar{v}_j = -hc^2 \sum_j \left( \frac{u_{j+1} - u_j}{h^2} + \frac{u_{j-1} - u_j}{h^2} \right) \bar{v}_j \\
 &= -hc^2 \left( \sum_j \frac{u_{j+1} - u_j}{h^2} \bar{v}_j - \sum_j \frac{u_j - u_{j-1}}{h^2} \bar{v}_j \right) \\
 &= -hc^2 \left( \sum_j \frac{u_{j+1} - u_j}{h^2} \bar{v}_j - \sum_j \frac{u_{j+1} - u_j}{h^2} \bar{v}_{j+1} \right) \\
 &= -hc^2 \sum_j \frac{u_{j+1} - u_j}{h} \frac{\bar{v}_j - \bar{v}_{j+1}}{h}
 \end{aligned}$$

#### Corrigé de l'exercice 6.

- **Les hypothèses (39) sont-elles vérifiées pour le choix (29) de conditions initiales?**

Dans ce cas, les hypothèses sont satisfaites puisque, pour toute projection orthogonale on a

$$\|\pi_h u_\alpha\| \leq \|u_\alpha\|$$

- **Les hypothèses (39) sont-elles vérifiées pour le choix (28) de conditions initiales, en supposant que  $u_0$  et  $u_1$  sont dans  $C^0(\mathbb{R}) \cap H^1(\mathbb{R})$  ?**

Faisons la démonstration pour  $u_0 \in C^0(\mathbb{R}) \cap H^1(\mathbb{R})$ . On montre dans ce cas que  $u_{0h}$  converge vers  $u_0$  dans  $L^2$ . Par conséquent, pour  $h \leq h_0$  assez petit, on peut majorer  $\|u_{0,h} - u_0\|_{L^2}$  par  $\|u_0\|_{L^2}$  et donc

$$\|u_{0,h}\|_{L^2} \leq \|u_{0,h} - u_0\|_{L^2} + \|u_0\|_{L^2} \leq 2 \|u_0\|_{L^2}$$

Pour montrer que  $u_{0h}$  converge vers  $u_0$  dans  $L^2$ , on écrit:

$$\begin{aligned} \|u_{0,h} - u_0\|_{L^2}^2 &= \int |u_{0,h}(x) - u_0(x)|^2 dx \\ &= \sum_j \int_{x_{j-1/2}}^{x_{j+1/2}} |u_0(x_j) - u_0(x)|^2 dx \\ &= \sum_j \int_{x_{j-1/2}}^{x_{j+1/2}} \left| \int_{x_j}^x u_0'(s) ds \right|^2 dx \\ &\leq \sum_j \int_{x_{j-1/2}}^{x_{j+1/2}} (x - x_j) \int_{x_j}^x |u_0'(s)|^2 ds dx \\ &\leq \sum_j \int_{x_{j-1/2}}^{x_{j+1/2}} \frac{h}{2} |u_0'(s)|^2 h ds \leq Ch^2 \|u_0'\|_{L^2}^2 \end{aligned}$$

On a donc

$$\|u_{0,h} - u_0\|_{L^2} \leq Ch \|u_0'\|_{L^2}$$

ce qui montre que  $u_{0,h}$  converge vers  $u_0$  dans  $L^2$ , à l'ordre 1.

On peut remarquer qu'on a été obligé de demander plus de régularité sur les conditions initiales. En effet, les estimations (39) ne sont pas vérifiées lorsqu'on approche les conditions initiales par les valeurs ponctuelles si ces conditions initiales sont seulement dans  $L^2 \cap C^0$ . Ce qui revient à dire que pour une fonction  $f \in L^2 \cap C^0$  on ne peut pas avoir une estimation du type:

$$h \sum_j |f(x_j)|^2 \leq C \int_{\mathbb{R}} |f(x)|^2 dx$$

### Corrigé de l'exercice 7

Montrer que l'hypothèse (41) est bien satisfaite pour les deux choix de  $u_{0,h}$  donnés par (28) et (29).

On suppose que  $c = cte$ , donc montrer que  $a_h(u_{0,h}, u_{0,h}) \leq Ca(u_0, u_0)$  revient à montrer que

$$h \sum_j \left| \frac{u_{0,j+1} - u_{0,j}}{h} \right|^2 \leq C \int_{\mathbb{R}} \left| \frac{du_0}{dx}(s) \right|^2 ds$$

- **Premier choix (28).** On suppose que  $u_0 \in C^0$  et

$$u_{0,j} = u_0(x_j)$$

On a donc

$$u_{0,j+1} - u_{0,j} = \int_{x_{j-1/2}}^{x_{j+1/2}} \frac{du_0}{dx}(x) dx$$

qu'on majore par Cauchy-Schwartz:

$$|u_{0,j+1} - u_{0,j}|^2 \leq h \int_{x_{j-1/2}}^{x_{j+1/2}} \left| \frac{du_0}{dx} \right|^2(x) dx$$

d'où en sommant:

$$\sum_j |u_{0,j+1} - u_{0,j}|^2 \leq h \int_{\mathbb{R}} \left| \frac{du_0}{dx} \right|^2(x) dx$$

ce qui est le résultat attendu.

- **Deuxième choix (29):**

$$u_{0,j} = \frac{1}{h} \int_{x_j-h/2}^{x_j+h/2} u_0(x) dx$$

d'où en faisant la différence:

$$u_{0,j+1} - u_{0,j} = \frac{1}{h} \int_{x_j-h/2}^{x_j+h/2} (u_0(x+h) - u_0(x)) dx$$

↓ (Cauchy-Schwartz)

$$\begin{aligned} |u_{0,j+1} - u_{0,j}|^2 &\leq \frac{1}{h^2} \int_{x_j-h/2}^{x_j+h/2} |u_0(x+h) - u_0(x)|^2 dx \times \int_{x_j-h/2}^{x_j+h/2} 1^2 dx \\ &\leq \frac{1}{h} \int_{x_j-h/2}^{x_j+h/2} |u_0(x+h) - u_0(x)|^2 dx \end{aligned}$$

↓ (on somme / à j)

$$\sum_j \left| \frac{u_{0,j+1} - u_{0,j}}{h} \right|^2 \leq \frac{1}{h} \int_{\mathbb{R}} \left| \frac{u_0(x+h) - u_0(x)}{h} \right|^2 dx$$

D'autre part

$$\begin{aligned} u_0(x+h) - u_0(x) &= \int_0^h \frac{du_0}{dx}(x+t) dt \\ \implies |u_0(x+h) - u_0(x)|^2 &\leq h \int_0^h \left| \frac{du_0}{dx}(x+t) \right|^2 dt \end{aligned}$$

d'où

$$\sum_j \left| \frac{u_{0,j+1} - u_{0,j}}{h} \right|^2 \leq \frac{1}{h} \int_{\mathbb{R}} \frac{1}{h} \int_0^h \left| \frac{du_0}{dx}(x+t) \right|^2 dt dx$$

On fait le changement de variable dans l'intégrale en  $x$ :  $u = x+t$ . Lorsque  $x \in \mathbb{R}$  et  $0 \leq t \leq h$ ,  $u$  parcourt  $\mathbb{R}$

$$\sum_j \left| \frac{u_{0,j+1} - u_{0,j}}{h} \right|^2 \leq \frac{1}{h^2} \int_{\mathbb{R}} \int_0^h \left| \frac{du_0}{dx}(u) \right|^2 dt du = \frac{1}{h} \int_{\mathbb{R}} \left| \frac{du_0}{dx}(u) \right|^2 du$$

donc finalement

$$a_h(u_{0,h}, u_{0,h}) \leq c^2 h \frac{1}{h} \int_{\mathbb{R}} \left| \frac{du_0}{dx}(u) \right|^2 du = c^2 \int_{\mathbb{R}} \left| \frac{du_0}{dx}(u) \right|^2 du = a(u_0, u_0)$$

**Corrigé de l'exercice 8 : convergence du schéma semi-discrétisé.**

1- La solution exacte vérifie

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

ce qui, par définition de  $\bar{u}_j(t)$ , s'écrit aussi:

$$\frac{d^2 \bar{u}_j}{dt^2}(t) = c^2 \frac{\partial^2 u}{\partial x^2}(x_j, t)$$

On en déduit:

$$\begin{aligned} \frac{d^2 \bar{u}_j}{dt^2}(t) + (A_h \bar{u}_h)_j(t) &= c^2 \frac{\partial^2 u}{\partial x^2}(x_j, t) + (A_h \bar{u}_h)_j(t) \\ &= c^2 \frac{\partial^2 u}{\partial x^2}(x_j, t) - c^2 \frac{u(x_{j+1}, t) - 2u(x_j, t) + u(x_{j-1}, t)}{h^2} \\ &\equiv \varepsilon_j(t) \end{aligned}$$

L'erreur de consistance  $\varepsilon_h$  est d'ordre 2. Plus précisément, faisons un développement de Taylor avec reste intégral:

$$\begin{aligned} \bar{u}_{j+1}(t) &= \bar{u}_j(t) + h \frac{\partial u}{\partial x}(x_j, t) + \frac{h^2}{2} \frac{\partial^2 u}{\partial x^2}(x_j, t) + \frac{h^3}{3!} \frac{\partial^3 u}{\partial x^3}(x_j, t) + \frac{1}{3!} \int_{x_j}^{x_{j+1}} \frac{\partial^4 u}{\partial x^4}(\xi, t) (x_{j+1} - \xi)^3 d\xi \\ \bar{u}_{j-1}(t) &= \bar{u}_j(t) - h \frac{\partial u}{\partial x}(x_j, t) + \frac{h^2}{2} \frac{\partial^2 u}{\partial x^2}(x_j, t) - \frac{h^3}{3!} \frac{\partial^3 u}{\partial x^3}(x_j, t) + \frac{1}{3!} \int_{x_j}^{x_{j-1}} \frac{\partial^4 u}{\partial x^4}(\xi, t) (x_{j-1} - \xi)^3 d\xi \end{aligned}$$

ce qui donne:

$$\frac{\bar{u}_{j+1} - 2\bar{u}_j + \bar{u}_{j-1}}{h^2} = \frac{\partial^2 u}{\partial x^2}(x_j, t) + \frac{1}{3!} \int_{x_j}^{x_{j+1}} \frac{\partial^4 u}{\partial x^4}(\xi, t) \frac{(x_{j+1} - \xi)^3}{h^2} d\xi + \frac{1}{3!} \int_{x_j}^{x_{j-1}} \frac{\partial^4 u}{\partial x^4}(\xi, t) \frac{(x_{j-1} - \xi)^3}{h^2} d\xi$$

et on en déduit:

$$\begin{aligned} |\varepsilon_j(t)| &\leq \frac{1}{3!} \int_{x_j}^{x_{j+1}} \left| \frac{\partial^4 u}{\partial x^4}(\xi, t) \right| \frac{|x_{j+1} - \xi|^3}{h^2} d\xi + \frac{1}{3!} \int_{x_{j-1}}^{x_j} \left| \frac{\partial^4 u}{\partial x^4}(\xi, t) \right| \frac{|x_{j-1} - \xi|^3}{h^2} d\xi \\ &\leq Ch \int_{x_{j-1}}^{x_{j+1}} \left| \frac{\partial^4 u}{\partial x^4}(\xi, t) \right| d\xi \end{aligned}$$

L'erreur de consistance est donc majorée en norme  $L^2$  par:

$$\begin{aligned} \|\varepsilon_h(t)\|^2 &= h \sum_j |\varepsilon_j(t)|^2 \leq Ch^3 \sum_j \left( \int_{x_{j-1}}^{x_{j+1}} \left| \frac{\partial^4 u}{\partial x^4}(\xi, t) \right| d\xi \right)^2 \\ &\leq Ch^3 \sum_j h \int_{x_{j-1}}^{x_{j+1}} \left| \frac{\partial^4 u}{\partial x^4}(\xi, t) \right|^2 d\xi \leq Ch^4 \left\| \frac{\partial^4 u}{\partial x^4}(t) \right\|_{L^2}^2 \end{aligned}$$

d'où (46).

2- En faisant la différence entre les systèmes vérifiés par  $\bar{u}_h$  et  $u_h$ , on montre que l'erreur de convergence vérifie le système suivant:

$$(137) \quad \begin{cases} \frac{d^2 e_h}{dt^2} + A_h e_h = \varepsilon_h & (i) \\ e_h(0) = \bar{u}_{0h} - u_{0h} & (ii) \quad (e_j(0) = u_0(x_j) - u_{0j}) \\ \frac{de_h}{dt}(0) = \bar{u}_{1h} - u_{1h} & (iii) \quad (\frac{de_j}{dt}(0) = u_1(x_j) - u_{1j}) \end{cases}$$

Ici l'énergie ne se conserve pas car il y a la présence d'un terme source, mais avec le même principe que sans terme source il est facile de voir, en multipliant scalairement l'équation par  $de_h/dt$  que:

$$\begin{aligned} \left( \frac{d^2 e_h}{dt^2} + A_h e_h, \frac{de_h}{dt} \right) &= \left( \varepsilon_h, \frac{de_h}{dt} \right) \Rightarrow \underbrace{\frac{dE_h}{dt}}_{2E_h^{1/2} \frac{dE_h^{1/2}}{dt}} = \underbrace{\left( \varepsilon_h, \frac{de_h}{dt} \right)}_{\leq \|\varepsilon_h\| \left\| \frac{de_h}{dt} \right\| \leq C \|\varepsilon_h\| E_h^{1/2}} \\ \Rightarrow \frac{dE_h^{1/2}}{dt} &\leq C \|\varepsilon_h\| \\ \Rightarrow E_h^{1/2}(t) - E_h^{1/2}(0) &\leq C \int_0^t \|\varepsilon_h\|(s) ds \end{aligned}$$

ce qui est l'estimation (47) annoncée.

3 - Pour en déduire une estimation de l'erreur, on utilise l'identité:

$$e_h(t) = e_h(0) + \int_0^t \frac{de_h}{dt}(s) ds$$

qui implique:

$$\begin{aligned} \|e_h(t)\| &\leq \|e_h(0)\| + \left\| \int_0^t \frac{de_h}{dt}(s) ds \right\| \\ &\leq \|e_h(0)\| + \left( \int_{\mathbb{R}} \left( \int_0^t \frac{de_h}{dt}(x, s) ds \right)^2 dx \right)^{1/2} \\ &\leq \|e_h(0)\| + \left( \int_{\mathbb{R}} t \int_0^t \left| \frac{de_h}{dt} \right|^2(x, s) ds dx \right)^{1/2} \\ &\leq \|e_h(0)\| + \left( t \int_0^t \int_{\mathbb{R}} \left| \frac{de_h}{dt} \right|^2(x, s) dx ds \right)^{1/2} \\ &\leq \|e_h(0)\| + \left( t \int_0^t \left\| \frac{de_h}{dt} \right\|^2(s) ds \right)^{1/2} \end{aligned}$$

Or on a :

$$\begin{aligned} \left\| \frac{de_h}{dt} \right\|^2(s) &\leq C E_h(s) \leq C \left( (E_h(0))^{1/2} + \int_0^s \|\varepsilon_h\|(v) dv \right)^2 \\ &\leq C \left( E_h(0) + \left( \int_0^s \|\varepsilon_h\|(v) dv \right)^2 \right) \leq C \left( E_h(0) + \left( \int_0^t \|\varepsilon_h\|(v) dv \right)^2 \right), \quad \forall s \leq t \end{aligned}$$

d'où en reportant dans l'inégalité précédente:

$$\begin{aligned}\|e_h(t)\| &\leq \|e_h(0)\| + \left( Ct^2 \left( E_h(0) + \left( \int_0^t \|\varepsilon_h\|(v)dv \right)^2 \right) \right)^{1/2} \\ &\leq \|e_h(0)\| + Ct(E_h(0))^{1/2} + Ct \int_0^t \|\varepsilon_h\|(v)dv\end{aligned}$$

Or on a

$$E_h(0) = \frac{1}{2} \left\| \frac{de_h}{dt}(0) \right\|^2 + \frac{1}{2} a_h(e_h(0), e_h(0))$$

d'où le résultat (48) annoncé.

D'après l'estimation sur l'erreur de consistance (46), on a

$$\int_0^t \|\varepsilon_h(v)\| dv \leq Ch^2 \int_0^t \left\| \frac{\partial^4 u}{\partial x^4}(v) \right\|_{L^2} dv \leq Ch^2 \left\| \frac{\partial^4 u}{\partial x^4} \right\|_{L^1(0,T;L^2(\mathbb{R}))}$$

Si les conditions initiales sont approchées par (28), c'est à dire par les valeurs ponctuelles, l'erreur de convergence démarre de conditions initiales nulles:

$$e_h(0) = \frac{de_h}{dt}(0) = 0$$

et il reste:

$$\|e_h(t)\| \leq Ct \int_0^t \|\varepsilon_h\|(v)dv \leq Cth^2 \left\| \frac{\partial^4 u}{\partial x^4} \right\|_{L^1(0,T;L^2(\mathbb{R}))}$$

c'est à dire (49).

### Corrigé de l'exercice 9.

1- Pour  $\alpha = 1$ , montrons que l'erreur de troncature est nulle. On sait que la solution exacte s'écrit sous la forme:

$$u(x, t) = f(x + ct) + g(x - ct)$$

Pour  $\alpha = 1$  ( $c\Delta t = \Delta x$ ), le schéma se réécrit:

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

$$\iff u_j^{n+1} - 2u_j^n + u_j^{n-1} = u_{j+1}^n - 2u_j^n + u_{j-1}^n$$

$$\iff u_j^{n+1} + u_j^{n-1} = u_{j+1}^n + u_{j-1}^n$$

Pour la solution exacte, notons  $\bar{u}_j^n = u(x_j, t^n)$ , on a:

$$\begin{aligned}\bar{u}_j^{n+1} + \bar{u}_j^{n-1} &= f(x_j + ct^{n+1}) + g(x_j - ct^{n+1}) + f(x_j + ct^{n-1}) + g(x_j - ct^{n-1}) \\ &= f((j+n+1)h) + g((j-n-1)h) + f((j+n-1)h) + g((j-n+1)h) \\ &= f((j+1)h + nc\Delta t) + g((j+1)h - nc\Delta t) \\ &\quad + f((j-1)h + nc\Delta t) + g((j-1)h - nc\Delta t) \\ &= \bar{u}_{j+1}^n + \bar{u}_{j-1}^n\end{aligned}$$



La solution exacte vérifie donc le schéma de façon exacte ce qui montre que l'erreur de troncature est nulle,  $\varepsilon_j^n = 0$ .

2- et 3- En 2D, on ne gagne plus d'ordre si  $\alpha = 1$  (en supposant que  $\Delta x = \Delta y = h$ ). On approche l'équation

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

par le schéma

$$\frac{u_{ij}^{n+1} - 2u_{ij}^n + u_{ij}^{n-1}}{\Delta t^2} - c^2 \left( \frac{u_{i+1j}^n - 2u_{ij}^n + u_{i-1j}^n}{h^2} + \frac{u_{ij+1}^n - 2u_{ij}^n + u_{ij-1}^n}{h^2} \right) = 0$$

Si  $\bar{u}_{ij}^n = u(x_i, y_j, t^n)$  avec  $u$  solution exacte régulière de l'équation continue, l'erreur de troncature est définie par

$$\begin{aligned} \varepsilon_{ij}^n &= \frac{\bar{u}_{ij}^{n+1} - 2\bar{u}_{ij}^n + \bar{u}_{ij}^{n-1}}{\Delta t^2} - c^2 \left( \frac{\bar{u}_{i+1j}^n - 2\bar{u}_{ij}^n + \bar{u}_{i-1j}^n}{h^2} + \frac{\bar{u}_{ij+1}^n - 2\bar{u}_{ij}^n + \bar{u}_{ij-1}^n}{h^2} \right) \\ &= \left( \frac{\partial^2 u}{\partial t^2} \right)_{ij}^n + \frac{2}{4!} \Delta t^2 \left( \frac{\partial^4 u}{\partial t^4} \right)_{ij}^n + O(\Delta t^4) \\ &\quad - c^2 \left( \left( \frac{\partial^2 u}{\partial x^2} \right)_{ij}^n + \frac{2}{4!} \Delta x^2 \left( \frac{\partial^4 u}{\partial x^4} \right)_{ij}^n + O(\Delta x^4) \right) \\ &\quad - c^2 \left( \left( \frac{\partial^2 u}{\partial y^2} \right)_{ij}^n + \frac{2}{4!} \Delta y^2 \left( \frac{\partial^4 u}{\partial y^4} \right)_{ij}^n + O(\Delta y^4) \right) \\ &= \frac{2}{4!} \left( \Delta t^2 \left( \frac{\partial^4 u}{\partial t^4} \right)_{ij}^n \right) - c^2 \left( \Delta x^2 \left( \frac{\partial^4 u}{\partial x^4} \right)_{ij}^n + \Delta y^2 \left( \frac{\partial^4 u}{\partial y^4} \right)_{ij}^n \right) + O(\Delta t^4 + h^4) \end{aligned}$$

On réutilise l'équation, qu'on dérive  $\partial_t^2$ :

$$\frac{\partial^4 u}{\partial t^4} = c^2 \frac{\partial^2}{\partial t^2} \Delta u = c^4 \Delta^2 u = c^4 \left( \frac{\partial^4 u}{\partial x^4} + 2 \frac{\partial^4 u}{\partial x^2 \partial y^2} + \frac{\partial^4 u}{\partial y^4} \right)$$

d'où en supposant  $\Delta x = \Delta y = h$ , (et en oubliant les indices)

$$\varepsilon_{ij}^n = \frac{2}{4!} \left( \Delta t^2 c^4 \left( \frac{\partial^4 u}{\partial x^4} + 2 \frac{\partial^4 u}{\partial x^2 \partial y^2} + \frac{\partial^4 u}{\partial y^4} \right) - c^2 h^2 \left( \frac{\partial^4 u}{\partial x^4} + \frac{\partial^4 u}{\partial y^4} \right) \right) + O(\Delta t^4 + h^4)$$

On voit que si  $c^2 \Delta t^2 = h^2$  les termes en  $\frac{\partial^4 u}{\partial x^4}$  et  $\frac{\partial^4 u}{\partial y^4}$  disparaissent mais il reste les dérivées croisées, donc  $\varepsilon_{ij}^n = O(\Delta t^2)$  et le schéma reste d'ordre 2.

### Corrigé de l'exercice 11

1- On multiplie le schéma par  $\frac{u_h^{n+1} - u_h^{n-1}}{2\Delta t}$ , ce qui donne

$$\begin{aligned}
 & \left( \frac{u_h^{n+1} - 2u_h^n + u_h^{n-1}}{\Delta t^2}, \frac{u_h^{n+1} - u_h^{n-1}}{2\Delta t} \right) + \left( A_h u_h^n, \frac{u_h^{n+1} - u_h^{n-1}}{2\Delta t} \right) = 0 \\
 \Rightarrow & \left( \frac{u_h^{n+1} - u_h^n}{\Delta t^2}, \frac{u_h^{n+1} - u_h^n}{2\Delta t} \right) + \left( \frac{u_h^{n+1} - u_h^n}{\Delta t^2}, \frac{u_h^n - u_h^{n-1}}{2\Delta t} \right) \\
 & - \left( \frac{u_h^n - u_h^{n-1}}{\Delta t^2}, \frac{u_h^{n+1} - u_h^n}{2\Delta t} \right) - \left( \frac{u_h^n - u_h^{n-1}}{\Delta t^2}, \frac{u_h^n - u_h^{n-1}}{2\Delta t} \right) \\
 & + \frac{1}{2\Delta t} (A_h u_h^n, u_h^{n+1}) - \frac{1}{2\Delta t} (A_h u_h^n, u_h^{n-1}) = 0 \\
 \Rightarrow & \left\| \frac{u_h^{n+1} - u_h^n}{\Delta t} \right\|^2 - \left\| \frac{u_h^n - u_h^{n-1}}{\Delta t} \right\|^2 + (A_h u_h^n, u_h^{n+1}) - (A_h u_h^n, u_h^{n-1}) = 0
 \end{aligned}$$

Dans le cas semi-discret on avait

$$E_h(t) = \frac{1}{2} \left\| \frac{du_h}{dt} \right\|^2 + \frac{1}{2} a_h(u_h, u_h)$$

qui définissait une énergie car la forme  $a_h$  est une forme quadratique. Ici, la quantité  $a_h(u_h^n, u_h^{n+1})$  n'est pas quadratique, on ne connaît pas son signe. On va montrer que sous la condition CFL, la quantité  $E_h^{n+1/2}$  définit encore une énergie.

2- Par définition on a

$$\begin{aligned}
 E_h^{1/2} &= \frac{1}{2} \left\| \frac{u_h^1 - u_h^0}{\Delta t} \right\|^2 + \frac{1}{2} (A_h u_h^0, u_h^1) \\
 &= \frac{1}{2} \left\| u_{1h} - \frac{\Delta t}{2} A_h u_h^0 \right\|^2 + \frac{1}{2} (A_h u_h^0, (I - \frac{\Delta t^2}{2} A_h) u_h^0) + \frac{\Delta t}{2} (A_h u_h^0, u_{1h}) \\
 &= \frac{1}{2} \left( \|u_{1h}\|^2 + \frac{\Delta t^2}{4} \|A_h u_h^0\|^2 - \Delta t (u_{1h}, A_h u_h^0) \right) \\
 &\quad + \frac{1}{2} (A_h u_h^0, u_h^0) - \frac{\Delta t^2}{4} \|A_h u_h^0\|^2 + \frac{\Delta t}{2} (A_h u_h^0, u_{1h}) \\
 &= \frac{1}{2} (A_h u_h^0, u_h^0) - \frac{\Delta t^2}{8} \|A_h u_h^0\|^2 + \frac{1}{2} \|u_{1h}\|^2 \\
 &\leq \frac{1}{2} (A_h u_h^0, u_h^0) + \frac{1}{2} \|u_{1h}\|^2
 \end{aligned}$$

et d'après les hypothèses (39) et (41) ça implique donc que

$$E_h^{1/2} \leq C \left( \frac{1}{2} a(u_0, u_0) + \frac{1}{2} \|u_1\|^2 \right) = CE(0)$$

3- On utilise les identités (35) et (36):

$$\begin{aligned} (A_h v_h, v_h) &= a_h(v_h, v_h) = c^2 h \sum_j \left| \frac{v_{j+1} - v_j}{h} \right|^2 \leq c^2 h \sum_j \frac{2(|v_{j+1}|^2 + |v_j|^2)}{h^2} \\ &\leq \frac{4c^2}{h^2} h \sum_j |v_j|^2 = \frac{4c^2}{h^2} \|v_h\|^2 \end{aligned}$$

4- On réécrit l'énergie à l'aide de l'identité rappelée

$$\begin{aligned} E_h^{n+1/2} &= \frac{1}{2} \left\| \frac{u_h^{n+1} - u_h^n}{\Delta t} \right\|^2 \\ &+ \frac{1}{2} \left( \frac{1}{4} a_h \left( \frac{u_h^n + u_h^{n+1}}{2}, \frac{u_h^n + u_h^{n+1}}{2} \right) - \frac{\Delta t^2}{4} a_h \left( \frac{u_h^{n+1} - u_h^n}{\Delta t}, \frac{u_h^{n+1} - u_h^n}{\Delta t} \right) \right) \end{aligned}$$

et l'estimation précédente montre donc que

$$E_h^{n+1/2} \geq \frac{1}{2} \left\| \frac{u_h^{n+1} - u_h^n}{\Delta t} \right\|^2 \left( 1 - \frac{c^2 \Delta t^2}{h^2} \right)$$

On voit donc que sous la condition CFL,  $\left| \frac{c\Delta t}{h} \right| < 1$ , on a  $E_h^{n+1/2} > 0$ . On obtient ici le *caractère suffisant* de cette condition. ■

### Corrigé de l'exercice 13

**1- Conservation d'énergie:** on réécrit (90) sous la forme:

$$\left( \frac{\partial^2 u_h}{\partial t^2}, v_h \right)_\rho + a(u_h(t), v_h) = 0, \forall v_h \in V_h$$

puis, pour  $v_h = \partial u_h / \partial t$  on obtient le résultat.

**2- Majoration de l'énergie.** Si on suppose que

$$u_{0h} \rightarrow u_0, \text{ dans } V \text{ et } u_{1h} \rightarrow u_1, \text{ dans } H$$

on a en particulier pour  $h$  assez petit

$$\|u_0 - u_{0h}\|_V \leq \|u_0\|_V \text{ et } \|u_1 - u_{1h}\|_H \leq \|u_1\|_H$$

d'où

$$\|u_{0h}\|_V \leq \|u_0 - u_{0h}\|_V + \|u_0\|_V \leq 2 \|u_0\|_V$$

$$\|u_{1h}\|_H \leq \|u_1 - u_{1h}\|_H + \|u_1\|_H \leq 2 \|u_1\|_H$$

ce qui montre que

$$E_h(0) = \frac{1}{2} \|u_{1h}\|_H^2 + \frac{1}{2} a(u_{0h}, u_{0h}) \leq C(\|u_1\|_H^2 + \|u_0\|_V^2)$$

**3- Choix des conditions initiales.** On veut montrer que avec ce choix, on a

$$\begin{aligned} u_{0h} &\rightarrow u_0, \text{ dans } V & (i) \\ u_{1h} &\rightarrow u_1, \text{ dans } H & (ii) \end{aligned}$$

(i) On a :

$$\|u_0 - u_{0h}\|_V^2 = (u_0 - u_{0h}, u_0 - u_{0h})_V = (u_0 - u_{0h}, u_0 - v_h)_V, \quad \forall v_h \in V_h$$

d'où

$$\begin{aligned} \|u_0 - u_{0h}\|_V &\leq \|u_0 - v_h\|_V, \quad \forall v_h \in V_h \\ \|u_0 - u_{0h}\|_V &\leq \inf_{v_h \in V_h} \|u_0 - v_h\|_V, \end{aligned}$$

ce qui montre le résultat bien connu sur la projection orthogonale!

$$\|u_0 - u_{0h}\|_V = \inf_{v_h \in V_h} \|u_0 - v_h\|_V$$

et par l'hypothèse d'approximation (H1) on peut conclure que  $\|u_0 - u_{0h}\|_V \rightarrow 0$  lorsque  $h \rightarrow 0$ .

(ii) Même chose mais avec des normes  $H$  au lieu de  $V$ , on a donc:

$$\|u_1 - u_{1h}\|_H = \inf_{v_h \in V_h} \|u_1 - v_h\|_H,$$

et le lemme permet de conclure que  $\|u_1 - u_{1h}\|_H \rightarrow 0$ .

**Démo du lemme:** Soit  $v \in H$ .

- Par densité de  $V$  dans  $H$ : il existe une suite  $v^N \in V$ , telle que  $\|v - v^N\|_H \rightarrow 0$ .
- (H1) implique que pour tout  $v^N$  on a  $\lim_{h \rightarrow 0} \inf_{v_h \in V_h} \|v^N - v_h\|_V = 0$ .
- L'injection continue implique que  $\|\cdot\|_H \leq \|\cdot\|_V$  donc  $\lim_{h \rightarrow 0} \inf_{v_h \in V_h} \|v^N - v_h\|_H = 0$ .

Pour tout  $v_h$  on a

$$\|v - v_h\|_H \leq \|v - v^N\|_H + \|v^N - v_h\|_H, \quad \forall N$$

d'où

$$\inf_{v_h} \|v - v_h\|_H \leq \underbrace{\inf_{v_h} (\|v - v^N\|_H + \|v^N - v_h\|_H)}_{\text{indep de } v_h} = \|v - v^N\|_H + \inf_{v_h} \|v^N - v_h\|_H,$$

On fait tendre  $h$  vers 0:  $\forall N$

$$\lim_{h \rightarrow 0} \inf_{v_h \in V_h} \|v - v_h\|_H \leq \|v - v^N\|_H + \underbrace{\lim_{h \rightarrow 0} \inf_{v_h} \|v^N - v_h\|_H}_{=0}$$

d'où

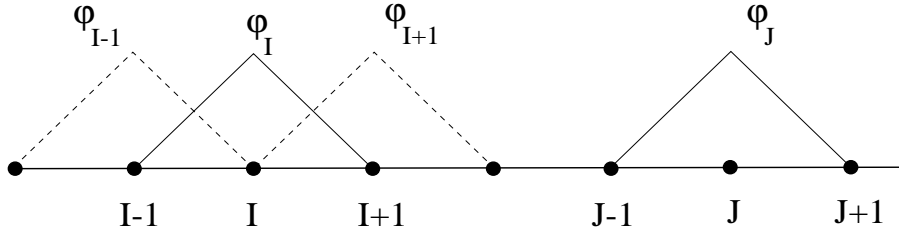
$$\lim_{h \rightarrow 0} \inf_{v_h \in V_h} \|v - v_h\|_H \leq \|v - v^N\|_H, \quad \forall N$$

On fait tendre  $N$  vers  $+\infty$ , donc  $\|v - v^N\|_H$  tend vers 0 et le terme de gauche est indépendant de  $N$  donc il est nul:

$$\lim_{h \rightarrow 0} \inf_{v_h \in V_h} \|v - v_h\|_H = 0. \quad \blacksquare$$

**Corrigé de l'exercice 15:**

**1- Matrices tridiagonales:** les fonctions de base ont pour support 2 intervalles adjacents et donc les supports de 2 fonctions de base ne s'intersectent que si ces fonctions de base sont associées au même noeud ou à un noeud voisin, i.e.



$$(\varphi_I, \varphi_J)_\rho = 0, \quad \text{si } J \neq I-1, I, I+1 \quad \text{car } \text{Supp } \varphi_I \cap \text{Supp } \varphi_J = \emptyset$$

**2- Calcul des matrices en homogène:**

**Calcul de la matrice de masse:** *Matrices élémentaires*

On calcule d'abord les interactions des fonctions de base sur l'élément de référence (cf Fig. 14)

$$\widehat{M}^e = \rho \begin{pmatrix} 1/3 & 1/6 \\ 1/6 & 1/3 \end{pmatrix}$$

En milieu homogène  $\rho(x) = \rho = cte$ , donc les matrices élémentaires associées à tous les éléments sont identiques: sur  $K_{i+1/2} = [x_i, x_{i+1}]$  par exemple:

$$M^{i+1/2} = \rho h \begin{pmatrix} 1/3 & 1/6 \\ 1/6 & 1/3 \end{pmatrix}$$

*Assemblage:*

$$M_{ii} = M_{22}^{i-1/2} + M_{11}^{i+1/2} = \frac{2}{3}\rho h; \quad M_{ii+1} = M_{12}^{i+1/2} = \frac{1}{6}\rho h$$

d'où la matrice de masse dans le cas homogène:

$$(MU)_i = \rho h \left( \frac{2}{3}U_i + \frac{1}{6}U_{i-1} + \frac{1}{6}U_{i+1} \right)$$

**Calcul de la matrice de rigidité:** on peut de nouveau faire le calcul des matrices élémentaires qui ici sont particulièrement simples car les fonctions de base étant  $P^1$ , les dérivées sont des constantes.

$$K_{ij} = a(\varphi_i, \varphi_j) = \int_0^L \mu \frac{d\varphi_i}{dx} \frac{d\varphi_j}{dx} dx$$

Or  $d\varphi_i/dx = 1/h$  sur  $]x_{i-1}, x_i[$  et  $d\varphi_i/dx = -1/h$  sur  $]x_i, x_{i+1}[$  donc

$$K_{ii} = \mu \int_{x_{i-1}}^{x_i} \frac{1}{h^2} dx + \int_{x_i}^{x_{i+1}} \frac{1}{h^2} dx = \frac{2\mu}{h}$$

$$K_{ii+1} = \mu \int_{x_i}^{x_{i+1}} \frac{-1}{h^2} dx = -\frac{\mu}{h}$$

On a donc

$$(KU)_i = -\frac{\mu}{h}(U_{i+1} - 2U_i + U_{i-1})$$

En divisant tout par  $\rho h$  on peut réécrire le schéma sous la forme:

$$\frac{M}{\rho h} \frac{d^2 U}{dt^2} + \frac{K}{\rho h} U = 0$$

et on remarque que

$$\left(\frac{K}{\rho h} U\right)_i = -\frac{\mu}{\rho} \frac{U_{i+1} - 2U_i + U_{i-1}}{h^2} = -c^2 \frac{U_{i+1} - 2U_i + U_{i-1}}{h^2}$$

coïncide avec l'opérateur  $A_h$  aux différences finies ( $c^2 = \frac{\mu}{\rho}$ ). Par contre le schéma aux différences finies ne faisait pas intervenir de matrice de masse, on avait seulement  $d^2 u/dt^2$  alors que maintenant on a  $M d^2 U/dt^2$  et  $M$  est une matrice tridiagonale, donc ce n'est plus un schéma explicite.

**3- Calcul des matrices en hétérogène:** on refait le même calcul que précédemment. Les matrices élémentaires sont cette fois différentes:

$$M^{i+1/2} = \rho_{i+1/2} h \begin{pmatrix} 1/3 & 1/6 \\ 1/6 & 1/3 \end{pmatrix}$$

d'où après assemblage

$$M_{ii} = M_{22}^{i-1/2} + M_{11}^{i+1/2} = \frac{h}{3}(\rho_{i+1/2} + \rho_{i-1/2})$$

$$M_{ii+1} = M_{12}^{i+1/2} = \frac{1}{6}\rho_{i+1/2}h; \quad M_{ii-1} = \frac{1}{6}\rho_{i-1/2}h$$

*Rigidité*

$$K_{ii} = \frac{\mu_{i-1/2}}{h} + \frac{\mu_{i+1/2}}{h}; \quad K_{ii+1} = -\frac{\mu_{i+1/2}}{h}$$

d'où le schéma. ■

## Corrigé de l'exercice 16

**1- Matrice condensée.** Rappelons que les éléments de la matrice de masse s'expriment:

$$M_{ij} = \int_0^L \rho(x) \varphi_i(x) \varphi_j(x) dx$$

On a déjà vu que seuls  $M_{ii+1}$ ,  $M_{ii}$  et  $M_{ii-1}$  sont non nuls. Regardons par exemple le terme:

$$M_{ii+1} = \int_0^L \rho(x) \varphi_i(x) \varphi_{i+1}(x) dx = \int_{x_i}^{x_{i+1}} \rho_{i+1/2} \varphi_i(x) \varphi_{i+1}(x) dx$$

$$\approx \rho_{i+1/2} h \frac{\overbrace{\varphi_i(x_i) \varphi_{i+1}(x_i)}^0 + \overbrace{\varphi_i(x_{i+1}) \varphi_{i+1}(x_{i+1})}^0}{2}$$

donc  $M_{ii+1}^{ap} = 0$ . Pour le terme diagonal, on a

$$\begin{aligned} M_{ii} &= \int_0^L \rho(x)(\varphi_i(x))^2 dx = \int_{x_{i-1}}^{x_i} \rho_{i-1/2}(\varphi_i(x))^2 dx + \int_{x_i}^{x_{i+1}} \rho_{i+1/2}(\varphi_i(x))^2 dx \\ &\approx \rho_{i-1/2} h \frac{\overbrace{(\varphi_i(x_{i-1}))^2}^0 + \overbrace{(\varphi_i(x_i))^2}^1}{2} + \rho_{i+1/2} h \frac{\overbrace{(\varphi_i(x_i))^2}^1 + \overbrace{(\varphi_i(x_{i+1}))^2}^0}{2} \\ &\approx h \frac{\rho_{i-1/2} + \rho_{i+1/2}}{2} \end{aligned}$$

On obtient donc le schéma explicite:

$$\frac{U_i^{n+1} - 2U_i^n + U_i^{n-1}}{\Delta t^2} + \frac{1}{h} \frac{2}{\rho_{i-1/2} + \rho_{i+1/2}} (KU^n)_i = 0$$

**2- En milieu homogène:** on a donc simplement

$$M^{ap} = \rho h I$$

et en utilisant  $((KU)_i = -\frac{\mu}{\rho}(U_{i+1} - 2U_i + U_{i-1}))$ , le schéma s'écrit:

$$\frac{U_i^{n+1} - 2U_i^n + U_i^{n-1}}{\Delta t^2} - \frac{\mu}{\rho} \frac{U_{i+1}^n - 2U_i^n + U_{i-1}^n}{h^2} = 0$$

On retombe donc sur le schéma aux différences finies.

Ce résultat est encore vrai en dimensions supérieures (en 2D ou en 3D): sur une *grille régulière* et avec les éléments finis  $P^1$ , on peut réinterpréter le schéma variationnel comme un schéma aux DF. Si en plus on utilise une formule de quadrature pour condenser la masse, il s'agit du schéma aux DF usuel.

**Remarque 9** *Le schéma avec condensation de masse étant identique au schéma obtenu par différences finies, en milieu homogène, on voit en particulier qu'il est bien d'ordre 2 (la dispersion est d'ordre 2 etc...).*

## B Stabilité du schéma totalement discrétisé

On se propose de montrer dans ce paragraphe qu'on peut obtenir une estimation de la solution, à partir de techniques d'énergie, uniforme par rapport à  $\alpha$  et par conséquent qui n'explose pas quand  $\alpha$  se rapproche de 1. On considère le schéma suivant:

$$M \frac{U^{n+1} - 2U^n + U^{n-1}}{\Delta t^2} + AU^n = 0$$

On suppose que  $A = BB^t$  et on réécrit le schéma en posant:

$$V^{n+1/2} = \frac{U^{n+1} - U^n}{\Delta t}$$

$$W^n = B^t U^n$$

ce qui donne:

$$\begin{cases} M \frac{V^{n+1/2} - V^{n-1/2}}{\Delta t} + BW^n = 0 & (i) \\ \frac{W^{n+1} - W^n}{\Delta t} - B^t V^{n+1/2} = 0 & (ii) \end{cases}$$

On a alors les deux conservations d'énergies discrètes:

$$E_1^{n+1/2} = E_1^{n-1/2} \text{ et } E_2^{n+1} = E_2^n$$

avec

$$\begin{cases} E_1^{n+1/2} = \frac{1}{2} (MV^{n+1/2}, V^{n+1/2}) + \frac{1}{2} (W^n, W^{n+1}) \\ = \frac{1}{2} \left( \left( M - \frac{\Delta t^2}{4} A \right) V^{n+1/2}, V^{n+1/2} \right) + \frac{1}{2} \left\| \frac{W^n + W^{n+1}}{2} \right\|^2 \\ E_2^n = \frac{1}{2} \|W^n\|^2 + \frac{1}{2} (V^{n+1/2}, V^{n-1/2})_M \\ = \frac{1}{2} \left( \left( I - \frac{\Delta t^2}{4} B^t M^{-1} B \right) W^n, W^n \right) + \frac{1}{2} \left\| \frac{V^{n+1/2} + V^{n-1/2}}{2} \right\|_M^2 \end{cases}$$

Posons:

$$\alpha_1 = \sup_v \frac{\Delta t^2 (Av, v)}{4 \|v\|_M^2}, \quad \alpha_2 = \sup_v \frac{\Delta t^2 (B^t M^{-1} Bv, v)}{4 \|v\|^2}$$

Il est alors facile de montrer le

**Lemme 15** • *La quantité  $E_1^{n+1/2}$  est positive pour tout  $n$  si et seulement si la condition  $\alpha_1 < 1$  est satisfaite. De plus, si cette condition est satisfaite, on a :*

$$(138) \quad E_1^{n+1/2} \geq \frac{1}{2} (1 - \alpha_1^2) \|V^{n+1/2}\|_M^2 + \frac{1}{2} \left\| \frac{W^n + W^{n+1}}{2} \right\|^2$$

• *La quantité  $E_2^n$  est positive pour tout  $n$  si et seulement si la condition  $\alpha_2 < 1$  est satisfaite. De plus, si cette condition est satisfaite, on a :*

$$(139) \quad E_2^n \geq \frac{1}{2} (1 - \alpha_2^2) \|W^n\|^2 + \frac{1}{2} \left\| \frac{V^{n+1/2} + V^{n-1/2}}{2} \right\|_M^2$$

• *Les deux conditions sont équivalentes:*

$$\alpha_1 < 1 \iff \alpha_2 < 1$$

Nous laissons la démonstration au lecteur. À partir de ce résultat, on montre que, si la condition CFL est satisfaite alors on obtient une estimation de la solution. En effet, si  $\alpha_2 < 1$ , l'estimation (139) montre qu'on peut contrôler en particulier

$$\left\| \frac{V^{n+1/2} + V^{n-1/2}}{2} \right\|_M^2 \leq 2E_2^0$$



qui est une approximation de l'énergie cinétique  $\left\| \frac{\partial u}{\partial t} \right\|^2$ . Or cette quantité peut aussi s'écrire:

$$\|U^{n+1} - U^{n-1}\|_M \leq 2\sqrt{2}\Delta t \sqrt{E_2^0}$$

Si par exemple  $n$  est pair,  $n = 2p$ , on a

$$\begin{aligned} \|U^{2p+1}\|_M &\leq \|U^{2p+1} - U^{2p-1}\|_M + \|U^{2p-1}\|_M \\ &\leq 2\sqrt{2}\Delta t \sqrt{E_2^0} + \|U^{2p-1}\|_M \\ &\leq 2\sqrt{2}p\Delta t \sqrt{E_2^0} + \|U^1\|_M \\ &\leq 2\sqrt{2}T \sqrt{E_2^0} + \|U^1\|_M \end{aligned}$$

De même si  $n = 2p + 1$ , on a

$$\|U^{2p+2}\|_M \leq 2\sqrt{2}T \sqrt{E_2^0} + \|U^0\|_M$$

Dans tous les cas, on obtient donc une estimation uniforme de la solution approchée par rapport aux données.

# Contents

<b>1</b>	<b>Introduction et généralités sur les ondes</b>	<b>3</b>
1.1	Les ondes acoustiques . . . . .	3
1.2	Les ondes élastiques . . . . .	5
1.3	Les ondes électromagnétiques . . . . .	6
1.4	Lien entre problème stationnaire et transitoire . . . . .	8
1.5	Techniques de résolution numériques de ces problèmes . . . . .	9
<b>2</b>	<b>Quelques propriétés de l'équation des ondes 1D</b>	<b>12</b>
2.1	Le problème modèle . . . . .	12
2.2	La formule de D'Alembert et ses conséquences . . . . .	12
2.2.1	Démonstration de la formule de D'Alembert . . . . .	13
2.2.2	Cône de dépendance et propagation à vitesse finie . . . . .	15
2.2.3	Régularité de la solution du problème de Cauchy . . . . .	15
2.2.4	Régularité de la solution du problème source . . . . .	15
2.2.5	Conservation de l'énergie- Unicité . . . . .	16
2.2.6	Stabilité par Fourier . . . . .	17
2.3	Ondes planes harmoniques . . . . .	17
<b>3</b>	<b>Semi-discrétisation en espace de l'équation des ondes 1D par un schéma aux différences finies</b>	<b>18</b>
3.1	Quelques préliminaires sur les différences finies. . . . .	18
3.2	Schéma semi-discrétisé en espace. . . . .	19
3.3	Analyse par Fourier discret . . . . .	20
3.4	Analyse de stabilité par des techniques d'énergie . . . . .	24
3.5	Convergence du schéma semi-discrétisé . . . . .	26
3.5.1	Convergence du schéma semi-discrétisé par techniques d'énergie . . . . .	26
3.5.2	Convergence du schéma semi-discrétisé par Fourier . . . . .	28
<b>4</b>	<b>Discrétisation totale par un schéma aux différences finies en espace et en temps</b>	<b>32</b>
4.1	Schéma explicite centré . . . . .	32
4.2	Ordre du schéma: consistance et erreur de troncature . . . . .	33
4.3	Cône de dépendance numérique- Condition nécessaire de convergence . . . . .	35
4.4	Analyse de stabilité par Fourier . . . . .	35
4.5	Analyse de dispersion . . . . .	40
4.6	Stabilité par techniques énergétiques . . . . .	44
4.7	Convergence du schéma totalement discrétisé par différences finies . . . . .	46
<b>5</b>	<b>Approximation variationnelle (Galerkin)</b>	<b>49</b>
5.1	Formulation variationnelle continue . . . . .	49
5.2	Approximation de Galerkin (semi-discrétisation en espace) . . . . .	50
5.2.1	La formulation approchée . . . . .	50
5.2.2	Estimation d'énergie . . . . .	50
5.2.3	Convergence du schéma semi discrétisé . . . . .	51
5.3	La méthode des éléments finis. . . . .	55
5.3.1	Quelques rappels . . . . .	55

5.3.2	Application à l'approximation de l'équation des ondes 1D. . . . .	58
5.3.3	Analyse de dispersion . . . . .	59
5.4	Discrétisation totale - Formules de quadrature, condensation de masse . . . . .	60
5.5	Stabilité par techniques d'énergie (milieux hétérogènes) . . . . .	61
5.5.1	Analyse de dispersion et de stabilité des deux schémas en milieu homogène et avec maillage uniforme . . . . .	66
5.6	Convergence du schéma totalement discrétisé . . . . .	68
<b>6</b>	<b>Analyse de dispersion et de stabilité de schémas en 2D</b>	<b>72</b>
6.1	Formulation variationnelle . . . . .	72
6.2	Approximation par différences finies . . . . .	72
6.3	Approximation par éléments finis $Q^1$ . . . . .	77
<b>A</b>	<b>Quelques corrigés</b>	<b>82</b>
<b>B</b>	<b>Stabilité du schéma totalement discrétisé</b>	<b>94</b>