Schémas d'ordre élevé en espace et/ou en temps pour l'équation des ondes acoustiques 1-D

Laurent Anné, Patrick Joly, Quang Huy Tran

▶ **To cite this version:**

Laurent Anné, Patrick Joly, Quang Huy Tran. Schémas d'ordre élevé en espace et/ou en temps pour l'équation des ondes acoustiques 1-D. [Rapport de recherche] RR-3759, INRIA. 1999. inria-00072903

HAL Id: inria-00072903
https://inria.hal.science/inria-00072903

Submitted on 24 May 2006

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.

INRIA

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Schémas d'ordre élevé en espace et/ou en temps pour l'équation des ondes acoustiques 1-D

Laurent ANNÉ , Patrick JOLY, Quang Huy TRAN

N ° 3759

Septembre 1999

THÈME 4

ISSN 0249-6399

Schémas d'ordre élevé en espace et/ou en temps pour l'équation des ondes acoustiques 1-D

Laurent ANNÉ*, Patrick JOLY†, Quang Huy TRAN‡

Thème 4 — Simulation et optimisation
de systèmes complexes
Projet Ondes

Rapport de recherche n° 3759 — Septembre 1999 — 36 pages

Résumé : Dans cet article, nous nous proposons de :

1. Établir toutes les propriétés, déjà signalées en conjecture [2] ou jusque-là méconnues, concernant la discrétisation à un ordre élevé du laplacien 1D.
2. Généraliser aux ordres supérieurs le schéma aux différences finies 4-4 de Cohen [6] pour l'équation des ondes acoustiques en milieu homogène 1-D.

Pour la famille de schémas 2m-2m obtenue, des propriétes variées concernant la stabilité et la dispersion numérique sont démentrées. La preuve de cse résultats se fait à l'aide de mathématiques élémentaires.

Mots-clé : équation des ondes, différences finies, stabilité, dispersion numérique.

(Abstract: pto)

* SIMULOG. lanne@grenoble.simulog.fr
† INRIA-rocquencourt. Patrick.Joly@inria.fr
‡ IFP. Q-Huy.Tran@ifp.fr

Unité de recherche INRIA Rocquencourt
Domaine de Voluceau, Rocquencourt, BP 105, 78153 LE CHESNAY Cedex (France)
Téléphone : 01 39 63 55 11 - International : +33 1 39 63 55 11
Télécopie : (33) 01 39 63 53 30 - International : +33 1 39 63 53 30

Higher-order numerical schemes in space and/or time for solving the 1D acoustic wave equation

Abstract: In this article, our objective is:

1. To establish theoretical properties, some of them mentionned as conjectures in [2], about higher order finite difference discretizations of the 1D Laplace operator.

2. Generalize to higher orders the finite difference scheme of Cohen [6] to the 1-D homogeneous acoustic wave equation.

For the corresponding family of 2m-2m schemes, various properties concerning the stability and the dispersion of the scheme can be proven. The proofs of hese results are based on elementary mathematics.

Key-words: wave equation, finite differences, stability, numerical dispersion.

Introduction

Vingt ans après avoir été introduits dans le monde de la géophysique par Alford et al. [1], les schémas aux différences finies sont devenus aujourd'hui l'outil de référence pour la résolution industrielle de l'équation des ondes acoustiques. Derrière l'aspect scientifique de la question, l'enjeu économique est de taille. Aussi, loin d'évoluer vers une solution bien figée, ce problème continue-t-il à attirer des contributions variées de la part des ingénieurs, physiciens, informaticiens, numériciens et mathématiciens.

Le schéma originel de Alford et al. [1, 16] est du deuxième ordre en temps et en espace (2-2). La recherche d'une meilleure précision sur des longues simulations a débouché sur une multitude de nouveaux schémas. Certains sont des variantes plus ou moins sophistiquées du schéma 2-2 : Bamberger et al. [4], Dablain [9], Mufti [22], Sei [23] ou plus récemment Anné et al. [2]. D'autres sont l'aboutissement d'une démarche radicalement différente : Holberg [14] ou Kosloff et Baysal [17]. Quoi qu'il en soit, ces schémas-là s'attachent essentiellement à améliorer la précision en espace, laissant celle en temps à l'ordre 2. Cette erreur temporelle détériore notablement l'effet du gain obtenu en espace sur l'ensemble du schéma.

La plupart des codes d'onde, parmi lesquels celui de l'IFP, utilisent couramment un schéma dit 2-4 car il est du deuxième ordre en temps et du quatrième ordre en espace. Afin de remédier à l'erreur temporelle du schéma 2-4, Shubin-Bell [24] puis Cohen [6] ont préconisé un schéma 4-4 du quatrième ordre en temps et en espace. Ce schéma a été d'être généralisé par Cohen et Joly [7], [8] aux milieux hétérogènes. En milieu homogène 1-D, cependant, le schéma 4-4 requiert le même nombre d'opérations informatiques que le schéma 2-4. De surcroît, la constante de stabilité du schéma, définie comme la valeur maximale autorisée pour le coefficient $c\tau/h$ - où c désigne la vitesse de propagation des ondes, τ le pas de temps et h le pas d'espace¡égale à 1 pour le schéma 4-4 , et donc meilleure que celle du schéma 2-4 qui est de $\sqrt{3}/2$.

Ces deux avantages non négligeables du schéma 4-4 nous ont incités à l'incorporer au sein d'une famille de schémas dits *2m-2m*, avec $m \geq 1$, lesquels posséderaient des propriétés analogues. Dans le cadre de ce papier, nous nous restreignons à un milieu homogène 1-D. Un tel choix simplificateur est dicté par notre désir d'alléger la démonstration de certains faits spécifiques observés sur le tracé des courbes de dispersion. On verra notamment que :

- la pulsation numérique est toujours une fonction croissante et concave du nombre d'onde ;
- la vitesse de groupe numérique est toujours inférieure à la vitesse de phase numérique, elle-même inférieure à la vitesse du milieu ;
- les vitesses numériques de groupe et de phase sont des fonctions décroissantes du nombre d'onde, croissantes du taux de stabilité et de l'ordre du schéma ;

On démontre que la constante de stabilité (CFL) des schémas 2m-2m vaut toujours 1. Ceci plaide avec force en leur faveur, d'autant plus que la relation de dispersion devient exacte lorsque le CFL choisi pour la simulation tend vers 1.

Bien entendu, on ne manquera pas de comparer ces schémas 2m-2m aux schémas 2-2m de Anné et al. [2], dont on reprend et développe les idées essentielles. Pour construire un laplacien discret d'ordre élevé, on combine plusieurs opérateurs d'ordre 2. Comme on le verra plus en détails, il y a 3 manières, distinctes mais complémentaires, de voir le laplacien discret d'ordre $2m$. Chaque manière de voir fait intervenir un jeu de coefficients, donnés par plusieurs formules simples et une représentation intégrale, à partir desquelles découlent des propriétés d'alternance de signe ou de complète monotonie [27]. L'une d'entre elles fournit une interprétation du laplacien discret en termes d'éléments finis mixtes [13], ainsi que la positivité de l'opérateur laplacien discret.

Notre texte s'organise en trois parties. Dans un premier temps (§1), on passe en revue toutes les propriétés du laplacien discret d'ordre élevé, lequel correspond à la partie spatiale de l'opérateur des ondes. Ce travail conduit directement aux schémas 2-2m, qui font l'objet de la deuxième section (§2). Enfin, on abordera (§3) la construction des schémas 2m-2m proprement dit.

1. Laplacien discret d'ordre 2m

Commençons par convenir de quelques notations qui seront employées tout au long de ce récit. Soit v une fonction de $\mathbb{R}$ dans $\mathbb{C}$; pour tout point $x \in \mathbb{R}$ et tout pas $h > 0$, on définit le laplacien discret élémentaire

$$\Delta_h v(x) = \frac{v(x+h) - 2v(x) + v(x-h)}{h^2}. \tag{1.1}$$

connue pour être l'approximation à trois points du laplacien. Si v est suffisamment régulière, on sait que $\Delta_h v(x)$ est une approximation d'ordre 2 de la dérivée seconde v''. Plus précisément, en supposant que $v \in \mathcal{C}^{2m+2}(\mathbb{R})$ pour un entier $m \geq 1$, on a

$$\Delta_h v(x) = v''(x) + 2\sum_{r=2}^{m} \frac{h^{2r-2}}{(2r)!} v^{(2r)}(x) + O(h^{2m}), \tag{1.2}$$

comme le montre sans difficulté un développement limité de v au voisinage de x. Soit maintenant p un entier ≥ 1. On aura besoin de considérer $\Delta_{ph} v(x)$, i.e. le laplacien d'ordre 2 calculé avec un pas égal à ph, qui vaut

$$\Delta_{ph} v(x) = \frac{v(x+ph) - 2v(x) + v(x-ph)}{p^2 h^2} \tag{1.3}$$

conformément à la définition (1.1). Il faut prendre garde à ne pas confondre $\Delta_{ph} v(x)$ avec le symbole $\Delta_h^p v(x)$, qui représente le p-ème itéré de Δ_h, appliqué à v en x. Pour ce dernier, on démontre [21] que

$$\Delta_h^p v(x) = \frac{1}{h^{2p}} \sum_{r=-p}^{p} (-1)^{p+r} \binom{2p}{p+r} v(x+rh) \tag{1.4}$$

où

$$\binom{2p}{p+r} = \frac{(2p)!}{(p+r)!(p-r)!} \tag{1.5}$$

est la notation anglo-saxonne pour les coefficients du binôme. Un développement limité, plus compliqué que celui de (1.2), montre que Δ_h^p est une approximation d'ordre 2 de la dérivée $v^{(2p)}$, à savoir

$$\Delta_h^p v(x) = v^{(2p)}(x) + O(h^2). \tag{1.6}$$

Les opérateurs ci-dessus s'appliquent aussi à des suites indexées dans $\mathbb{Z}$: il suffit de mettre v_{j+q} à la place de $v(x+qh)$. Ainsi, on aura

$$\Delta_h^p v_j = \frac{1}{h^{2p}} \sum_{r=-p}^{p} (-1)^{p+r} \binom{2p}{p+r} v_{j+r}. \tag{1.7}$$

Enfin, lorsque la droite $\mathbb{R}$ est découpée en un maillage régulier de pas h et de nœuds x_j, on écrira indifféremment v_j ou $v(x_j)$.

Notre objectif est de chercher une approximation à l'ordre $2m$ du laplacien, avec $m \in \mathbb{N}^*$ arbitraire. Désignons par ${}_m\Delta_h$ un tel opérateur discret, l'indice m à gauche du Δ au premier membre désigne la moitié de l'ordre souhaité. On va voir qu'une tel opérateur aux différences n'utilisant que $2m+1$ points existe, et qu'il peut même être vu sous trois angles différents, à savoir :

1. en tant que combinaison de laplaciens discrets d'ordre 2 calculés avec des pas différents, comme l'indique (1.8) ;
2. en tant que développement en série entière appliquée au laplacien d'ordre 2, comme l'indique (1.36) ;
3. en tant que convolution en espace sur des opérateurs laplaciens d'ordre 2 voisins, comme l'indique (1.50).

1.1 Première vision

On se propose de chercher une formule de type

$${}_m\Delta_h = \sum_{p=1}^{m} \alpha_p^m \Delta_{ph}\,. \tag{1.8}$$

de sorte que ${}_m\Delta_h$ constitue une approximation à l'ordre $2m$ du laplacien. Bien sûr, on a ${}_1\Delta_h = \Delta_h$, ce qui entraîne $\alpha_1^1 = 1$. Ce cas trivial étant vu, on va déterminer les coefficients α_p^m pour m quelconque. D'après (1.2), on a

$$\Delta_{ph} v(x) = v''(x) + 2\sum_{r=2}^{m} \frac{p^{2r-2} h^{2r-2}}{(2r)!} v^{(2r)}(x) + O(h^{2m}),$$

pour chaque p entre 1 et m. En combinant ces Δ_{ph} selon (1.8), on obtient

$$ {}_m\Delta_h v(x) = \left(\sum_{p=1}^{m} \alpha_p^m\right) v''(x) + 2\sum_{r=2}^{m} \frac{v^{(2r)}(x)}{(2r)!} \left(\sum_{p=1}^{m} \alpha_p^m p^{2r-2}\right) h^{2r-2} + O(h^{2m}). $$

Une condition nécessaire et suffisante pour atteindre la précision d'ordre $2m$ est que le coefficient de v'' soit égal à 1 et que, en même temps, les facteurs des h^{2r-2} soient nuls. Ainsi, les α_p^m doivent satisfaire le système linéaire

$$ \begin{pmatrix} 1 & 1 & 1 & \cdots & \cdots & 1 \\ 1^2 & 2^2 & 3^2 & \cdots & \cdots & m^2 \\ 1^4 & 2^4 & 3^4 & \cdots & \cdots & m^4 \\ \vdots & \vdots & \vdots & \ddots & & \vdots \\ \vdots & \vdots & \vdots & & \ddots & \vdots \\ 1^{2m'} & 2^{2m'} & 3^{2m'} & \cdots & \cdots & m^{2m'} \end{pmatrix} \begin{pmatrix} \alpha_1^m \\ \alpha_2^m \\ \alpha_3^m \\ \vdots \\ \vdots \\ \alpha_m^m \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ \vdots \\ 0 \end{pmatrix} \tag{1.9} $$

dans lequel on a posé $m' = m - 1$.

1.1.1 Formule explicite des coefficients

En dépit de son apparente complexité, ce système de Van der Monde admet une solution extrêmement simple. Les résultats qui suivent sont nouveaux par rapport à Anné et al. [2].

Théorème 1.1 *Pour tout $m \in \mathbb{N}^*$, les coefficients α_p^m sont donnés par*

$$ \alpha_p^m = \prod_{1\le r\le m}^{r\neq p} \left(1 - \frac{p^2}{r^2}\right)^{-1} \tag{1.10} $$

et vérifient l'identité

$$ \prod_{1\le p\le m} \left(1 - \frac{z^2}{p^2}\right)^{-1} = \sum_{p=1}^{m} \alpha_p^m \left(1 - \frac{z^2}{p^2}\right)^{-1} \tag{1.11} $$

pour tout $z \in \mathbb{C}$. Ce sont donc les coefficients de la décompostion en éléments simples de la fraction rationnelle du premier membre de (1.11).

Démonstration- Appelons W_m la matrice du système (1.9), et α^m le vecteur $(\alpha_1^m, , \alpha_2^m, \ldots, \alpha_m^m)^T$. Soit $\ell = (\ell_1, \ell_2, \ldots, \ell_m)^T$ un vecteur quelconque de $\mathbb{R}^m$. En multipliant les deux membres de (1.9) par ℓ^T, on arrive à

$$ \ell^T W_m \alpha^m = \ell_1 \tag{1.12} $$

Un système de Van der Monde renferme des liens étroits avec l'interpolation polynômiale [10]. C'est pourquoi on introduit

$$ P_\ell(z) = \ell_1 + \ell_2 z^2 + \ell_3 z^4 + \cdots + \ell_m z^{2m-2}, \tag{1.13} $$

polynôme pair de degré $2m-2$ associé au vecteur $\ell \in \mathbb{R}^m$. Alors, il est facile de voir que

$$ \ell^T W_m = (P_\ell(1), P_\ell(2), \ldots, P_\ell(m)), $$

ce qui permet de ré-écrire l'équation (1.12) sous la forme

$$ \sum_{r=1}^{m} P_\ell(r)\alpha_r^m = P_\ell(0). \tag{1.14} $$

Demander que (1.12) soit vrai pour tout $\ell \in \mathbb{R}^m$ revient à demander que (1.14) ait lieu pour tout polynôme $P \in \mathbb{R}[X]$, pair et de degré $2m-2$. Or, pour p fixé entre 1 et m, le polynôme

$$ Q_p(z) = \prod_{1\le r\le m}^{r\neq p} \left(1 - \frac{z^2}{r^2}\right) $$

est de la forme imposée, s'annule en tout point $z \in \{1, \ldots, m\}$ sauf en $z = p$. Dans ces conditions, la relation (1.14) fournit

$$Q_p(p)\alpha_p^m = Q_p(0) = 1,$$

d'où $\alpha_p^m = 1/Q_p(p)$ qui est la formule (1.10). Quant à l'identité (1.11), elle s'obtient en décomposant directement le premier-membre. □

La formule (1.10) peut servir à prouver plusieurs observations bien connues par les praticiens des ondes.

Corollaire 1.1 *Les coefficients α_p^m possèdent les propriétés suivantes :*

1.

$$\alpha_p^m = (-1)^{p-1} \frac{2\, m!^2}{(m-p)!\,(m+p)!}\,,$$

2. Leurs valeurs absolues, données par

$$|\alpha_p^m| = \frac{2\, m!^2}{(m-p)!\,(m+p)!}\,,$$

décroissent avec p et croissent avec m.

3. Leurs sommes partielles sont positives, à savoir

$$\sum_{r=1}^{p} \alpha_r^m > 0\,.$$

DÉMONSTRATION- 1. Le produit (1.10) peut se transformer en

$$\alpha_p^m = \prod_{1\le r\le m}^{r\neq p} \left(1 - \frac{p}{r}\right)^{-1} \times \prod_{1\le r\le m}^{r\neq p} \left(1 + \frac{p}{r}\right)^{-1}$$

En réduisant chaque terme au même dénominateur, on montre aisément que le premier sous-produit est égal à

$$\prod_{1\le r\le m}^{r\neq p} \left(1 - \frac{p}{r}\right)^{-1} = (-1)^{p-1} \frac{m!}{p!\,(m-p)!},$$

et le second sous-produit vaut

$$\prod_{1\le r\le m}^{r\neq p} \left(1 + \frac{p}{r}\right)^{-1} = 2\,\frac{m!\,p!}{(m+p)!}.$$

La valeur de α_p^m s'en déduit aisément.

2. On déduit de ce qui précède que

$$\frac{|\alpha_{p+1}^m|}{|\alpha_p^m|} = \frac{(m-p)!\,(m+p)!}{(m-p-1)!\,(m+p+1)!} = \frac{m-p}{m+p+1}.$$

Ce rapport est manifestement inférieur à 1, d'où la décroissance en p. La croissance vis-à-vis de m se démontre de la même façon.

3. L'alternance des signes et la décroissance en valeur absolue impliquent

$$\alpha_{2q-1}^m + \alpha_{2q}^m = |\alpha_{2q-1}^m| - |\alpha_{2q}^m| > 0$$

pour tout $q \geq \lfloor m/2 \rfloor$ ($\lfloor a \rfloor$ désignant la partie entière du réel a). Ceci garantit la positivité des sommes partielles de rang pair. Celles de rang impair sont aussi positifs, car $\alpha_{2q+1}^m > 0$. □

En spécifiant $m = 2$ dans (1.10), on retrouve la combinaison classique [6] pour le laplacien discret d'ordre 4, i.e.

$$\alpha_1^2 = \tfrac{4}{3} \quad \text{et} \quad \alpha_2^2 = -\tfrac{1}{3}.$$

L'annexe A.1 fournit la liste des α_p^m pour $m \leq 6$.

1.1.2 Ordre infini

À ce stade, il est tentant de faire tendre m vers l'infini en vue d'obtenir, du moins théoriquement, un laplacien d'ordre ∞.

Proposition 1.1 *Lorsque m tend vers l'infini à p fixé, on a*

$$\alpha_p^\infty = \lim_{m\to\infty} \alpha_p^m = 2(-1)^{p-1}.$$

DÉMONSTRATION- Nous donnons deux preuves de ce résultat.

1. La fomule pour α_p^m peut encore s'écrire:

$$\alpha_p^m = (-1)^{p-1}\,\frac{m(m-1)\dots(m-p+1)}{(m+1)(m+2)\dots(m+p)}$$

α_p^m apparaît donc comme le quotient de deux polynômes en m de même degré p. La limite lorsque n tend vers l'infini est donc donné par le rapport des coefficients en m^p.

2. La formule (1.10) du Théorème 1.1 peut encore se mettre sous la forme

$$\alpha_p^m = \lim_{z\to p} \frac{1}{E_m(z)}\left(1-\frac{z^2}{p^2}\right) \tag{1.15}$$

avec par définition

$$E_m(z) = \prod_{1\le r\le m}\left(1-\frac{z^2}{r^2}\right).$$

Quand $m\to\infty$, il est connu [3, 11] que ce produit tend vers

$$\frac{\sin\pi z}{\pi z} = \prod_{r=1}^{\infty}\left(1-\frac{z^2}{r^2}\right). \tag{1.16}$$

La convergence de ce produit infini étant uniforme sur tout disque $|z|<R$, on peut intervertir les passages à la limite en m et en z pour obtenir

$$\alpha_p^\infty = \lim_{z\to p}\frac{\pi z}{\sin\pi z}\left(1-\frac{z^2}{p^2}\right).$$

Le changement de variable $y = 1-z/p$ conduit à

$$\alpha_p^\infty = \lim_{y\to 0}\frac{\pi p(1-y)}{(-1)^{p-1}\sin\pi py}\,y(2-y) = 2(-1)^{p-1}$$

comme annoncé. □

Les polynômes

$$E_p(z) = \prod_{r=1}^{p}\left(1-\frac{z^2}{r^2}\right), \tag{1.17}$$

avec par convention $E_0(z)=1$, réapparaîtront plus tard, dans les schémas 2m-2m pour l'équation des ondes.

Cela étant, on se rend compte rapidement que l'opérateur

$${}_\infty\Delta_h = 2\sum_{p=1}^{\infty}(-1)^{p-1}\Delta_{ph}\,, \tag{1.18}$$

surnommé *opérateur de Fourier* par certains auteurs, n'a aucun intérêt pratique: il est non-local car il fait intervenir une infinité de points de grille.

Remarque 1.1 Lorsque v est une fonction dans $\mathcal{C}_0^\infty(\mathbb{R})$ on peut toujours calculer ${}_\infty\Delta_h v(x)$ puique la somme (1.18) ne comporte qu'un nombre fini de termes. Cela reste vrai si la fonction v décroît suffisamment vite à l'infini, en particulier si v appartient à $\mathcal{S}(\mathbb{R})$ (on a alors une série convergente). Par contre, cette série ne converge pas nécessairement si v est simplement $\mathcal{C}^\infty(\mathbb{R})$. Il suffit par exemple de prendre la fonction quadratique $v(x) = \frac{1}{2}x^2$, laquelle vérifie $\Delta_{ph} v(x) = 1$ quel que soit $x \in \mathbb{R}$, pour tomber sur la série divergente $2\sum(-1)^{p-1}$!

Remarque 1.2 L'ordre infini n'est pas synonyme de *exact* même si $v \in \mathcal{S}(\mathbb{R})$. Il signifie seulement que l'erreur de troncature $v''(x) - {}_\infty\Delta_h v(x)$ décroît plus vite que toute puissance de h lorsque h tend vers 0. On illustrera ce phénomène à la section suivante.

1.1.3 Implémentation

Informatiquement, la mise en œuvre efficace de l'opérateur ${}_m\Delta_h$ passe par un autre jeu de coefficients, définis comme

$$a_0^m = -2\sum_{p=1}^{m} \frac{\alpha_p^m}{p^2} \quad \text{et} \quad a_q^m = \frac{\alpha_q^m}{q^2} \quad \text{pour } 1 \le q \le m. \tag{1.19}$$

On parle alors de *molécule* avec l'expression

$${}_m\Delta_h v_j = \frac{1}{h^2}\left\{ a_0^m v_j + \sum_{q=1}^{m} a_q^m (v_{j+q} + v_{j-q}) \right\}. \tag{1.20}$$

C'est cette molécule qui sera effectivement codée.

Remarque 1.3 On verra à la section 1.3 que $a_0^m = -2\sum_{p=1}^{m} \frac{1}{p^2}$. □

Remarque 1.4 Rien ne s'oppose à ce que l'on considère les limites a_p^∞ lorsque $m \to \infty$. D'après (1.19), on a

$$a_0^\infty = -\frac{\pi^2}{3} \quad \text{et} \quad a_p^\infty = \frac{2(-1)^{p-1}}{p^2} \qquad \text{pour } p \ge 1. \tag{1.21}$$

1.2 Deuxième vision

Même si l'on dispose déjà de beaucoup de résultats sur le laplacien discret ${}_m\Delta_h$, il est possible de jeter un regard entièrement neuf sur celui-ci. À cette fin, le passage vers le domaine fréquentiel et la manipulation des symboles sont riches d'enseignements.

1.2.1 Interprétation de ${}_m\Delta_h$ en tant que symbole

Rappelons que la notion de *symbole* d'un opérateur commutant avec les translations est lié à l'utilisation de la transformation de Fourier. Pour notre usage dans cet article, il nous suffira pour le définir de décrire l'action d'un tel opérateur sur la fonction $v(x) = \exp(ikx)$, avec $i^2 = -1$ et $k \in \mathbb{R}$. Ainsi, puisque $v'' = -k^2 v$, on dit que $-k^2$ est le symbole de d^2/dx^2, et l'on écrit

$$\mathcal{F}(d^2/dx^2) = -k^2 \tag{1.22}$$

L'action de Δ_h sur v est différente, car l'on a [2] sans difficulté

$$\Delta_h v = \frac{2}{h^2}\left[\cos(kh) - 1\right] v = -\frac{4}{h^2}\sin^2\left(\frac{kh}{2}\right) v$$

ce qui entraîne

$$\mathcal{F}(\Delta_h) = -\frac{4}{h^2}\sin^2\left(\frac{kh}{2}\right). \tag{1.23}$$

Dans le but de rendre les formules plus compactes, on introduit

$$K = \tfrac{1}{2}kh, \tag{1.24}$$

variable sans dimension qui représente l'inverse du nombre de pas d'espace par période, à un facteur multiplicatif près. Les correspondances (1.22) et (1.23) deviennent

$$\mathcal{F}(d^2/dx^2) = -\frac{4}{h^2}\,K^2 \qquad \text{et} \qquad \mathcal{F}(\Delta_h) = -\frac{4}{h^2}\,\sin^2 K.$$

Par composition, il est clair qu'on aura

$$\mathcal{F}(d^{2p}/dx^{2p}) = (-1)^p \left(\frac{4}{h^2}\right)^p K^{2p} \tag{1.25}$$

et

$$\mathcal{F}(\Delta_h^p) = (-1)^p \left(\frac{4}{h^2}\right)^p \sin^{2p} K. \tag{1.26}$$

Enfin, n'oublions pas que

$$\mathcal{F}(\Delta_{ph}) = -\frac{2}{p^2h^2}\left[1 - \cos(2pK)\right]. \tag{1.27}$$

Cherchons maintenant le symbole de ${}_m\Delta_h$. En vertu de la définition (1.8) et de la correspondance (1.27), on a

$$\mathcal{F}({}_m\Delta_h) = \frac{2}{h^2}\sum_{p=1}^{m} \frac{\alpha_p^m}{p^2}\left[\cos(2pK) - 1\right], \tag{1.28}$$

ou encore, en utilisant les coefficients a_p^m de (1.19),

$$\mathcal{F}({}_m\Delta_h) = \frac{1}{h^2}\{\, a_0^m + 2\sum_{p=1}^{m} a_p^m \cos(2pK)\,\}. \tag{1.29}$$

Il est possible de réarranger ce symbole sous la forme d'un polynôme de degré m en la variable $\sin^2 K$. Énonçons plus précisément ce résultat fondamental.

Théorème 1.2 *Le symbole de ${}_m\Delta_h$ peut se mettre sous la forme*

$$\mathcal{F}({}_m\Delta_h) = -\frac{4}{h^2}\sum_{p=1}^{m} \beta_p \sin^{2p} K, \tag{1.30}$$

dans laquelle les coeffients β_p sont

1. *indépendants de m ;*
2. *strictement positifs et égaux à*

$$\beta_p = \frac{2^{2p-1}(p-1)!^2}{(2p)!}\,; \tag{1.31}$$

DÉMONSTRATION- 1. Les formules trigonométriques permettent de dire d'emblée que (1.28) peut être transformée en

$$\mathcal{F}({}_m\Delta_h) = -\frac{4}{h^2}\sum_{p=1}^{m} \beta_p^m \sin^{2p} K,$$

où, a priori, les coefficients β_p^m dépendent de m. L'argument clé de la démonstration réside dans l'ordre du schéma. En effet, si ${}_m\Delta_h$ est d'ordre $2m$ en espace, son symbole se doit d'être d'ordre $2m+2$ en fréquence, i.e. l'écart entre $\mathcal{F}(d^2/dx^2)$ et $\mathcal{F}({}_m\Delta_h)$ est en K^{2m+2} :

$$-\frac{4}{h^2}K^2 = -\frac{4}{h^2}\sum_{p=1}^{m} \beta_p^m \sin^{2p} K + O(K^{2m+2})$$

Les β_p^m apparaissent donc comme les m premiers coefficients du développement de K^2 en série de $\sin K$.

2. En posant $z = \sin K$ pour $|K| \le \pi/2$, la relation ci-dessus devient

$$\arcsin^2 z = \sum_{p=1}^{m} \beta_p z^{2p} + O(z^{2m+2}) \tag{1.32}$$

Or, la fonction $\Upsilon(z) = \arcsin^2 z$ vérifie l'équation différentielle

$$(1-z^2)\Upsilon'' - z\Upsilon = 2 \tag{1.33}$$

sur l'intervalle $]-1,1[$. En reportant formellement la série $\Upsilon(z) = \sum \beta_p z^{2p}$ dans (1.33), on obtient

$$2\beta_1 + \sum_{p=1}^{\infty} \left[(2p+1)(2p+2)\beta_{p+1} - 4p^2\beta_p\right] z^{2p} = 2,$$

à partir de quoi il vient, par identification, que

$$\beta_1 = 1 \qquad \text{et} \qquad \beta_{p+1} = \frac{4p^2}{(2p+1)(2p+2)}\beta_p.$$

La formule (1.31) en découle. Tous les β_p sont bien positifs.

En vue de justifier rigoureusement le calcul précédent, on forme a posteriori la série $\Upsilon(z) = \sum \beta_p z^{2p}$, où les β_p sont imposés par (1.31). Puisque

$$\lim_{p\to\infty} \frac{\beta_{p+1}}{\beta_p} = 1,$$

le rayon de convergence de cette série est égal à 1. On sait qu'à l'intérieur du disque de convergence, une fonction définie par une série entière est toujours dérivable terme à terme [20]. En menant à bout cette opération, on voit que Υ obéit à (1.33) sur $]-1,1[$. □

Remarque 1.5 On peut retrouver les α_p^m à partir de la connaissance des β_p. Voici les grandes étapes du calcul. En élevant à la puissance p-ème la relation

$$\sin K = \frac{\exp(iK) - \exp(-iK)}{2i},$$

on obtient

$$\sin^{2p} K = 2^{-2p}\binom{2p}{p} + 2^{1-2p}\sum_{l=1}^{p}(-1)^l\binom{2p}{p-l}\cos(2lK).$$

On reporte cette expression dans (1.30) après avoir remplacé les β_p par leurs valeurs puis on permute l'ordre des sommations (en faisant attention aux indices). On identifie alors terme à terme le résultat obtenu avec (1.28). Tous calculs effectués, on obtient

$$\alpha_p^m = 2(-1)^{p-1}p^2 \sum_{l=p}^{m} \frac{(l-1)(l-2)\dots(l-p+1)}{l(l+1)(l+2)\dots(l+p)}. \tag{1.34}$$

La somme au second-membre se calcule simplement si l'on dispose d'une primitive discrète pour le terme générique, c'est-à-dire d'une fonction $\wp$ telle que

$$p^2\,\frac{(l-1)(l-2)\dots(l-p+1)}{l(l+1)(l+2)\dots(l+p)} = \wp(l) - \wp(l-1). \tag{1.35}$$

On voit qu'alors, en ajoutant ces égalités membre à membre, on se ramène à une somme téléscopique , de sorte que

$$\alpha_p^m = 2(-1)^{p-1}[\wp(m) - \wp(p-1)].$$

Or, on vérifie aisément que la fonction

$$\wp(z) = \frac{z(z-1)(z-2)\dots(z-p+1)}{(z+1)(z+2)\dots(z+p)}$$

remplit la condition (1.35). De plus, $\wp(p-1) = 0$. Ainsi, $\alpha_p^m = 2(-1)^{p-1}\wp(m)$, ce qui cohérent avec le Corollaire 1.1.

Les premiers termes de la suite β sont: $\beta_1 = 1$, $\beta_2 = \frac{1}{3}$ et $\beta_3 = \frac{8}{45}$. Une liste plus exhaustive peut être consultée à l'annexe A.2. Qu'il soit sous la forme (1.28), (1.29) ou (1.30), le symbole $\mathcal{F}(_m\Delta_h)$ est avant tout

une fonction paire et 2π-périodique de la variable K. La Figure 1 représente les variations en fonction de K la fonction

$$\overline{\mathcal{F}}(_m\Delta_h) = -\tfrac{1}{4}h^2\,\mathcal{F}(_m\Delta_h)$$

pour quelques valeurs de m, ainsi que celles de la même fonction pour le symbole exact.

Attardons-nous quelques instants sur une propriété des coefficients β_p qui nous sera utile plus tard.

Corollaire 1.2 *La série $\sum\beta_p$ converge et a pour somme $\pi^2/4$.*

DÉMONSTRATION- En se référant à la preuve du Théorème 1.2, on a

$$\sum_{p=1}^{\infty}\beta_p z^{2p} = \arcsin^2 z \qquad \text{pour tout } z\in\,]-1,\,1\,[.$$

Soit $N\in\mathbb{N}$. Comme tous les β_p sont positifs, on a

$$\sum_{p=1}^{N}\beta_p z^{2p} < \arcsin^2 z.$$

Lorsque $z\to 1$, chacun des deux membres de l'inégalité précédente possède une limite finie. Les deux limites respectent encore l'inégalité

$$\sum_{p=1}^{N}\beta_p \le \arcsin^2 1 = \pi^2/4.$$

Cela prouve que la série à termes positifs $\sum\beta_p$ converge et sa somme est majorée par $\pi^2/4$. Remarquons maintenant qu'en faisant tendre N vers l'infini dans l'inégalité évidente

$$\sum_{p=1}^{N}\beta_p z^{2p} < \sum_{p=1}^{N}\beta_p\,,$$

on obtient

$$\arcsin^2 z = \sum_{p=1}^{\infty}\beta_p z^{2p} \le \sum_{p=1}^{\infty}\beta_p$$

pour tout $z\in\,]-1,\,1\,[$. Il ne reste plus qu'à faire $z\to 1$ pour minorer la somme $\sum\beta_p$ par $\pi^2/4$. □

Remarque 1.6 La convergence de la série de terme général β_p peut se vérifier en remarquemt que, grâce à l'utilisation de la formule de Stirling:

$$n!\sim e^{-n}n^n\sqrt{2\pi n} \qquad n\to+\infty$$

on a:

$$\beta_p \simeq \frac{sqrt\pi}{2p^{\frac{3}{2}}} \qquad p\to+\infty.$$

1.2.2 Positivité de $_m\Delta_h$ en tant qu'opérateur dans ℓ_h^2

Retournons au domaine spatial afin de voir ce qu'y implique le Théorème 1.2. Des correspondances déjà vues entre opérateur et symboles, on tire l'égalité

$$[-_m\Delta_h] = -\Delta_h + \sum_{p=2}^{m} b_p h^{2p-2}(-\Delta_h)^p \tag{1.36}$$

avec $b_p = 4^{1-p}\beta_p$. Par conséquent, le laplacien discret $_m\Delta_h$ s'interprète comme étant le laplacien élémentaire Δ_h auquel ont été apportées $m-1$ corrections successives utilisant les puissances Δ_h^p. Les coefficients de ces termes correctifs ne dépendent pas de m.

Cette manière de voir a des répercussions favorables sur l'analyse de $-[_m\Delta_h]$. Soit ℓ^2_h l'ensemble des suites réelles indexées sur $\mathbb{Z}$, de carrée sommable. Muni du produit scalaire

$$\langle u, v\rangle = \sum_j u_j v_j\, h, \tag{1.37}$$

l'espace ℓ^2_h est de Hilbert. Il est bien connu que l'opérateur $-\Delta_h$ est autoadjoint et positif dans ℓ^2_h en vertu de la formule:

$$\forall\, u,\, v \in \ell^2_h \qquad \langle -\Delta_h u,\, v\rangle = \langle \nabla^+_h u,\, \nabla^+_h v\rangle$$

où le gradient discret ∇^+_h est défini dans par:

$$\nabla^+_h u_j = \frac{u_{j+1} - u_j}{h} \tag{1.38}$$

Il suffit en fait de remarquer que:

$$\Delta_h = \nabla^-_h \nabla^+_h$$

où l'opérateur ∇^-_h est défini dans ℓ^2_h par:

$$\nabla^-_h u_j = \frac{u_j - u_{j-1}}{h} \tag{1.39}$$

n'est autre que l'opposé de l'adjoint, dans ℓ^2_h, de l'opérateur ∇^-_h. On en déduit que tous les itérés $(-\Delta_h)^p$ de $-\Delta_h$ sont autoadjoint positifs dans ℓ^2_h. On a en particulier la formule:

$$\forall\, u,\, v \in \ell^2_h \qquad \langle (-\Delta^p_h)u,\, v\rangle = \langle \heartsuit^p_h u,\, \heartsuit^p_h v\rangle. \tag{1.40}$$

où les opérateurs $\heartsuit^p_h$ sont définis par:

$$\heartsuit^{2q}_h = \Delta^q_h \qquad \text{et} \qquad \heartsuit^{2q+1}_h = \nabla^+_h \Delta^q_h \tag{1.41}$$

Résumons toutes nos observations par la

Proposition 1.2 *L'opérateur $-[_m\Delta_h]$ est symétrique et positif sur ℓ^2_h. Plus précisément, si $b_p = 4^{1-p}\beta_p$, alors*

$$\forall\, u,\, v \in \ell^2_h \qquad \langle -[_m\Delta_h]u,\, v\rangle = \sum_{p=1}^{m} b_p h^{2p-2} \langle \heartsuit^p_h u,\, \heartsuit^p_h v\rangle.$$

La positivité est une propriété essentielle pour la résolution numérique de l'équation des ondes. C'est elle qui assure en effet la stabilité des schémas. Notons que, dans le cas de $_m\Delta_h$, cette propriété ne découle pas immédiatement de la première vision, à cause de l'alternance de signe des coefficients α^m_p. Le recours à la deuxième vision, dans laquelle les coefficients β_p sont positifs, est indispensable.

1.2.3 Liens avec les éléments finis

Supposons qu'on ait à résoudre $-u'' = f$ sur $\mathbb{R}$ par une méthode de différences finies utilisant $_m\Delta_h$. Il n'est un secret pour personne [4] que si $m = 1$, alors Δ_h possède un équivalent variationnel en termes d'éléments finis P_1. En s'inspirant des travaux de Glowinski et Pironeau [13], on peut généraliser ce résultat pour m arbitraire.

Théorème 1.3 *L'utilisation de $_m\Delta_h$ pour le problème $-u'' = f$ sur $\mathbb{R}$ équivaut à l'application, sur un maillage de pas h, de la méthode des éléments finis P_1, avec condensation de masse, à la formulation sous forme de système du second ordre de l'équation modifiée*

$$-u'' + \sum_{p=2}^{m} (-1)^p b_p h^{2p-2} u^{2p} = f. \tag{1.42}$$

DÉMONSTRATION- A] Écriture sous forme de système. Le problème modifié (1.42) équivaut, du moins formellement, à trouver $m+1$ fonctions u, ν_1, ..., ν_m telles que

$$\left\{\begin{array}{rcl} -u'' &=& \nu_1 \\ -\nu_1'' &=& \nu_2 \\ \vdots & \vdots & \vdots \\ -\nu_{m-1}'' &=& \nu_m \\ \nu_1 + \sum_{p=2}^m b_p h^{2p-2} \nu_p &=& f \end{array}\right. \tag{1.43}$$

où les ν_k jouent le rôle de fonctions auxiliaires.

B] Formulation variationnelle. Chercher $m+1$ fonctions u, ν_1, ..., ν_m, tous appartenant à $H^1(\mathbb{R})$, telles que

$$\left\{\begin{array}{rcll} \int u'\mu_1'\,dx &=& \int \nu_1\mu_1\,dx & \forall \mu_1 \in H^1(\mathbb{R}) \\ \int \nu_1'\mu_2'\,dx &=& \int \nu_2\mu_2\,dx & \forall \mu_2 \in H^1(\mathbb{R}) \\ \vdots & \vdots & \vdots & \vdots \in \vdots \\ \int \nu_{m-1}'\mu_m'\,dx &=& \int \nu_m\mu_m\,dx & \forall \mu_m \in H^1(\mathbb{R}) \\ \int [\nu_1 + \sum_{p=2}^m b_p h^{2p-2}\nu_p]\mu\,dx &=& \int f\mu dx & \forall \mu \in H^1(\mathbb{R}) \end{array}\right. \tag{1.44}$$

l'intégration étant faite sur $\mathbb{R}$.

C] Discrétisation P1. On choisit $V_h \subset H^1(\mathbb{R})$, espace des fonctions affines par morceaux sur un maillage de pas h. La traduction de (1.44) dans la base canonique de V_h donne lieu aux systèmes linéaires infinis

$$\left\{\begin{array}{rcl} \mathbf{K}\mathbf{u} &=& \mathbf{M}\boldsymbol{\nu}_1 \\ \mathbf{K}\boldsymbol{\nu}_1 &=& \mathbf{M}\boldsymbol{\nu}_2 \\ \vdots & \vdots & \vdots \\ \mathbf{K}\boldsymbol{\nu}_{m-1} &=& \mathbf{M}\boldsymbol{\nu}_m \\ \mathbf{M}\,[\boldsymbol{\nu}_1 + \sum_{p=2}^m b_p h^{2p-2}\boldsymbol{\nu}_p] &=& \mathbf{f} \end{array}\right. \tag{1.45}$$

avec

$\mathbf{K}$	matrice (infinie) de rigidité
$\mathbf{M}$	matrice (infinie) de masse
$\mathbf{f}$	vecteur (infini) des second-membres
$\mathbf{u}$	vecteur (infini) des valeurs de u
$\boldsymbol{\nu}_k$	vecteur (infini) des valeurs de ν_k

D] Condensation de masse. On approche $\mathbf{M}$ par l'opérateur identité $\mathbf{I}$, ce qui peut se justifier par exemple par l'application d'une formule de quadrature numérique. Les équations (1.45) se simplifient, et en éliminant les $\boldsymbol{\nu}_k$, on obtient

$$[\mathbf{K} + \sum_{p=2}^m b_p h^{2p-2}\mathbf{K}^p]\,\mathbf{u} = \mathbf{f}.$$

Puisque $\mathbf{K}$ représente $-\Delta_h$, tout se passe comme si le vecteur $\mathbf{u}$ vérifie

$$[-\Delta_h + \sum_{p=2}^m (-1)^p b_p h^{2p-2}\Delta_h^p]\,\mathbf{u} = \mathbf{f},$$

c'est-à-dire $-[_m\Delta_h]\mathbf{u} = \mathbf{f}$, à l'ordre 2. □

1.2.4 Ordre infini

On constate que

$$\overline{\mathcal{F}}(_\infty\Delta_h) \equiv \overline{\mathcal{F}}(d^2/dx^2) \quad \text{si et seulement si} \quad |K| \le \pi/2,$$

ce qui est au demeurant normal, car dans la démonstration du Théorème 1.2, on a vu que

$$K^2 = \sum_{p=1}^\infty \beta_p \sin^{2p} K$$

seulement pour $|K| \le \pi/2$, auquel cas on a bien $K = \arcsin z$. Il existe de ce fait des situations dans lesquelles l'ordre infini est exact.

Corollaire 1.3 *Soit $v \in \mathcal{C}^\infty(\mathbb{R})$ admettant une transformée de Fourier et dont le spectre est borné, i.e. inclus dans $[-k_{\max}, k_{\max}]$. Alors, $_\infty\Delta_h v = v''$ si l'on veille à ce que $h \le \pi/k_{\max}$.*

DÉMONSTRATION- Pour assurer $_\infty\Delta_h v = v''$, il suffit que pour chaque nombre d'onde k du spectre, les symboles $\mathcal{F}(_\infty\Delta_h)$ et $\mathcal{F}(d^2/dx^2)$ prennent la même valeur. Ceci se produit lorsque

$$|K| = \frac{|k|h}{2} \le \frac{\pi}{2}$$

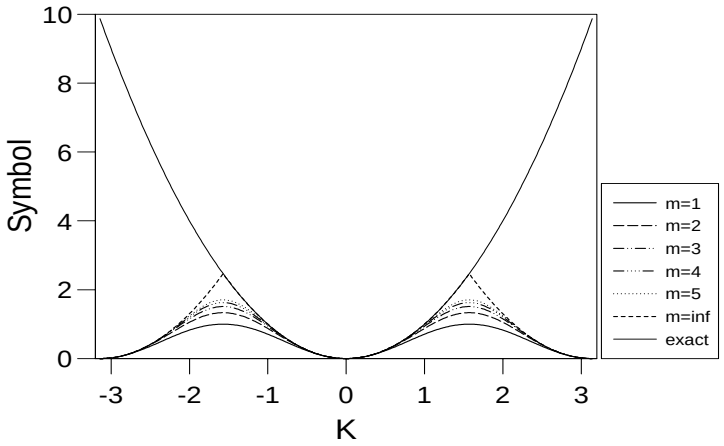

FIG. 1 – *Graphe de* $\overline{\mathcal{F}}({}_m\Delta_h)$ *pour* $m = 1, 2, 3, \infty$ *et celui de* $\overline{\mathcal{F}}(d^2/dx^2)$ *en fonction de la variable normalisée* K

c'est-à-dire si $|k| \leq \pi/h$. □

Remarque 1.7 Ce résultat est à rapprocher du théorème d'échantillonnage de Shannon [12] qui affirme qu'une fonction dont le spectre est borné est entièrement déterminée par ses valeurs sur un maillage suffisamment fin de $\mathbb{R}$. La condition sur le pas s'écrit encore

$$h \leq \tfrac{1}{2}\lambda_{\min}\,, \tag{1.46}$$

où $\lambda_{\min} = 2\pi/k_{\max}$ est la plus petite longueur d'onde : c'est le critère de Nyquist.

Par conséquent, c'est parmi les fonctions u à spectre non-borné qu'il faudra chercher des contre-exemples pour lesquels ${}_\infty\Delta_h u \neq u''$. Nous prétendons que la gaussienne

$$u(x) = \exp(-\tfrac{1}{2}x^2) \tag{1.47}$$

répond à la question. En effet, sa décomposition en harmoniques est

$$u(x) = \int_{\mathbb{R}} \widehat{u}(k)\exp(ikx)\,dk \qquad \text{avec} \quad \widehat{u}(k) = \frac{1}{\sqrt{2\pi}}\exp(-\tfrac{1}{2}k^2).$$

En dérivant sous le signe somme, on obtient

$$\begin{aligned} {}_m\Delta_h u(x) - u''(x) &= -2\int_0^\infty [\mathcal{F}(d^2/dx^2) - \mathcal{F}({}_m\Delta_h)]\,\widehat{u}(k)\cos(kx)\,dk \\ &= \frac{8}{h^2}\int_0^\infty [\overline{\mathcal{F}}(d^2/dx^2) - \overline{\mathcal{F}}({}_m\Delta_h)]\,\widehat{u}(k)\cos(kx)\,dk. \end{aligned}$$

La présence de $\cos(kx)$ est due à la parité des symboles, considérés maintenant comme fonctions de k. L'intégrand est un produit de facteurs positifs. La différence entre crochet s'annule pour $k \in [0, \pi/h]$. Pour $k \geq \pi/h$, elle peut être majorée par le premier terme $\overline{\mathcal{F}}(d^2/dx^2) = \frac{1}{4}k^2h^2$. Au total,

$$0 < {}_m\Delta_h u(x) - u''(x) \leq \sqrt{\frac{2}{\pi}}\int_{\pi/h}^\infty k^2\exp(-\tfrac{1}{2}k^2)\,dk.$$

Une intégration par parties montre que

$$\int_{\pi/h}^\infty k^2\exp(-\tfrac{1}{2}k^2)\,dk = \frac{\pi}{h}\exp\left(-\frac{\pi^2}{2h^2}\right) + \sqrt{2}\,\mathrm{erfc}\left(\frac{\pi}{h\sqrt{2}}\right)$$

où erfc(.) est la fonction d'erreur complémentaire, dont un équivalent [19] pour les grands arguments ($h \to 0$) est

$$\mathrm{erfc}\left(\frac{\pi}{h\sqrt{2}}\right) \approx \sqrt{\frac{2}{\pi}}\,\frac{h}{\pi}\,\exp\left(-\frac{\pi^2}{2h^2}\right).$$

Ce terme tend vers 0. On en déduit que

$$ {}_m\Delta_h u(x) - u''(x) \approx \sqrt{2\pi}\, h^{-1} \exp(-\pi^2/2h^2) \tag{1.48} $$

lorsque $h \to 0$. Cette expression décroît plus vite que tout monôme de h certes, mais elle n'est pas nulle !

Remarque 1.8 On peut retrouver les résultats de la Remarque 1.3. Pour $m \to \infty$, l'égalité (1.29) devient

$$ \mathcal{F}({}_\infty\Delta_h) = \frac{1}{h^2}\{a_0^\infty + 2\sum_{p=1}^{\infty} a_p^\infty \cos(2pK)\}\,. $$

Si $|K| \le \pi/2$, on sait que ce symbole coïncide avec le symbole exact, d'où

$$ -4K^2 = a_0^\infty + 2\sum_{p=1}^{\infty} a_p^\infty \cos(2pK)\,, $$

soit, par le changement de variable $\kappa = 2K$,

$$ -\tfrac{1}{2}\kappa^2 = \tfrac{1}{2}a_0^\infty + \sum_{p=1}^{\infty} a_p^\infty \cos(p\kappa)\,, \tag{1.49} $$

pour $\kappa \in\,]-\pi, \pi[$. Autrement dit, les coefficients a_p^∞ sont ceux du développement en série de Fourier de la fonction 2π-périodique qui vaut $-\frac{1}{2}\kappa^2$ sur l'intervalle $]-\pi, \pi[$. Il s'ensuit que

$$ a_p^\infty = \frac{1}{\pi}\int_{-\pi}^{\pi} [-\tfrac{1}{2}\kappa^2]\cos(p\kappa)\, d\kappa = \frac{2(-1)^{p-1}}{p^2} \quad \text{si } p \ge 1, \quad = -\frac{\pi^2}{3} \quad \text{si } p = 0. $$

On retombe bien sur (1.21). □

1.3 Troisième vision

Malgré l'abondance des résultats déjà établis, le sujet du laplacien discret ${}_m\Delta_h$ est pourtant loin d'être épuisé. Il subsiste un troisième point de vue, plus familier sans doute aux spécialistes de traitement du signal.

Prenons l'exemple de l'opérateur ${}_2\Delta_h$, auquel on fait subir les transformations suivantes :

$$ \begin{aligned} {}_2\Delta_h v_j &= \tfrac{4}{3}\Delta_h v_j - \tfrac{1}{3}\Delta_{2h} v_j \\ &\asymp -\tfrac{1}{12}v_{j+2} + \tfrac{4}{3}v_{j+1} - \tfrac{5}{2}v_j + \tfrac{4}{3}v_{j-1} - \tfrac{1}{12}v_{j-2} \\ &= -\tfrac{1}{12}\Delta_h v_{j-1} + \tfrac{7}{6}\Delta_h v_j - \tfrac{1}{12}\Delta_h v_{j+1}. \end{aligned} $$

Dans la deuxième ligne, le signe $\asymp$ signale qu'il faut diviser le second-membre par h^2, facteur que l'on a omis pour des raisons typographiques. La dernière ligne traduit que ${}_2\Delta_h v_j$ est une combinaison linéaire de $\Delta_h v$ aux points $j-1$, j et $j+1$. Plus généralement, on peut se demander s'il existe une formule de *convolution* symétrique de type

$$ {}_m\Delta_h v_j = \gamma_0^m \Delta_h v_j + \sum_{q=1}^{m-1} \gamma_q^m (\Delta_h v_{j-q} + \Delta_h v_{j+q})\,. \tag{1.50} $$

En anticipation du théorème qui va suivre, on introduit quelques notations supplémentaires. Pour tout entier $r \in \mathbb{N}^*$, on considère

$$ H_r = \sum_{n=1}^{r} \frac{(-1)^{n-1}}{n} \tag{1.51} $$

$$ \zeta_r = \sum_{n=1}^{r} \frac{1}{n^2} \tag{1.52} $$

$$ \Psi_r = \sum_{n=1}^{r} \frac{(-1)^{n-1}}{n^2} \tag{1.53} $$

et on adopte les conventions $H_r = \zeta_r = \Psi_r = 0$ si $r \le 0$. Muni de ces notations, on peut énoncer le

Théorème 1.4 *Les coefficients γ_q^m possèdent les propriétés suivantes :*

1. *Leurs signes alternent avec q de sorte que $(-1)^q \gamma_q^m > 0$.*

2. *Leurs valeurs absolues décroissent avec q et croissent avec m.*

3. *Ils sont reliés à la série harmonique et à celle des carrés inverses par*

$$\gamma_q^m = 2H_{2m} - q\zeta_m + 2q\Psi_{q-1} - 2H_{q-1} + \sum_{p=1}^{q-1}(q-p)[a_p^m - a_p^\infty]$$

DÉMONSTRATION- On essaie de relier les γ_q^m aux a_p^m et β_p déjà connus. Les calculs sont extrêment fastidieux, c'est pourquoi nous les laissons aux lecteurs intéressés. □

Cette formule met bien en évidence le rôle des sommes partielles H_r, ζ_r et Ψ_r. Elle est toutefois assez longue, et pourrait à terme être améliorée. Signalons, toujours sans démonstration, quelques cas particuliers intéressants.

Corollaire 1.4 *Les γ_q^m vérifient les propriétés suivantes:*

– *la somme des coefficients de la molécule de convolution est égale à 1*

$$\gamma_0^m + 2\sum_{q=1}^{m-1}\gamma_q^m = 1$$

– *le coefficient central γ_0^m est le double de la somme partielle de rang $2m$ de la série harmonique alternée*

$$\gamma_0^m = 2\sum_{n=1}^{2m}\frac{(-1)^{n-1}}{n} = 2\ln 2 - \int_0^1 \xi^{2m}\,\frac{2}{1+\xi}\,d\xi.$$

– *le coefficient a_0^m de la section 1.1.3 est relié aux γ_q^m par*

$$\frac{a_0^m}{2} = \gamma_0^m - \gamma_1^m = \sum_{p=1}^{m}\frac{1}{p^2}.$$

Bien entendu, on peut faire tendre m vers l'infini et obtenir une convolution infinie, donnée par

Corollaire 1.5 *Les coefficients γ_q^∞ sont donnés par*

$$\gamma_q^\infty = 2\ln 2 - q\pi^2/6 + 2q\Psi_{q-1} - 2H_{q-1}.$$

Ils forment une suite alternée, décroissante au sens strict. Sa série converge et l'on a $\gamma_0^\infty + 2\sum_{q=1}^\infty \gamma_q^\infty = 1$.

2. Schémas 2-2m pour les ondes acoustiques

Prolongement logique de l'étude du laplacien discret, cette partie a pour objet la résolution de l'équation des ondes acoustiques, en milieu 1-D homogène et infini, par une famille de schémas appelés *2-2m*.

2.1 Construction

Situons d'abord le contexte du problème. On cherche une approximation de la solution $u(t,x)$, avec $t > 0$ et $x \in \mathbb{R}$, de l'équation des ondes

$$\frac{1}{c^2}\frac{\partial^2 u}{\partial t^2} - \frac{\partial^2 u}{\partial x^2} = 0, \tag{2.1}$$

complétée par les conditions initiales de Cauchy

$$u(0,x) = u_0(x) \quad \text{et} \quad \frac{\partial u}{\partial t}(0,x) = v_0(x). \tag{2.2}$$

Pour cela, on se donne deux pas de discrétisation, à savoir

$$\begin{cases} h > 0 & : \quad \text{le pas d'espace} \\ \tau > 0 & : \quad \text{le pas de temps} \end{cases}$$

positifs mais arbitraires pour le moment. Les nœuds du maillage en espace sont notés x_j, alors que les instants multiples de τ sont désignés par t^n. Avec ces notations, le schéma de Alford et al. [1] s'écrit

$$\Delta_\tau u_j^n - c^2 \Delta_h u_j^n = 0,$$

équation à partir de laquelle on tire la formule de *mise à jour*

$$u_j^{n+1} = 2u_j^n - u_j^{n-1} + \left(\frac{c\tau}{h}\right)^2 [u_{j+1}^n - 2u_j^n + u_{j-1}^n].$$

On voit qu'il y intervient le quotient

$$\lambda = \frac{c\tau}{h}, \tag{2.3}$$

appelé coefficient de stabilité ou *le CFL* dans le jargon numéricien. Quant au schéma 2-2m, il s'écrit

$$\Delta_\tau u_j^n - c^2 [{}_m\Delta_h] u_j^n = 0. \tag{2.4}$$

Il est explicite, d'ordre 2 en temps et d'ordre $2m$ en espace au sens de l'erreur de troncature : si $\tilde{u}$ est la solution exacte de (2.1) et (2.2),supposée trèsrégulière, on a

$$[\Delta_\tau \tilde{u} - c^2 {}_m\Delta_h \tilde{u}](t^n, x_j) = O(\tau^2) + O(h^{2m}). \tag{2.5}$$

2.2 Analyse de Fourier

La *relation de dispersion* est définie comme la condition nécessaire et suffisante pour que le schéma, en l'occurrence l'équation (2.4), accepte une solution de la forme

$$u_j^n = \exp i(n\omega\tau - jkh), \tag{2.6}$$

avec

$$\begin{cases} \omega \in \mathbb{R} & : \text{ la pulsation,} \\ k \in \mathbb{R} & : \text{ le nombre d'onde.} \end{cases}$$

Connaissant le symbole de Δ_τ et celui de $[{}_m\Delta_h]$, on voit immédiatement qu'une telle condition est

$$-\frac{4}{\tau^2}\sin^2\left(\frac{\omega\tau}{2}\right) = -\frac{4c^2}{h^2}\sum_{p=1}^{m}\beta_p \sin^{2p}\left(\frac{kh}{2}\right). \tag{2.7}$$

Allégeons les écritures en définissant les nombres sans dimension

$$\begin{cases} \Omega = \frac{1}{2}\omega\tau & : \text{ la pulsation réduite} \\ K = \frac{1}{2}kh & : \text{ le nombre d'onde réduit} \end{cases}$$

lesquels transforment la relation (2.7) en

$$\sin^2\Omega = \lambda^2 \sum_{p=1}^{m}\beta_p \sin^{2p} K, \tag{2.8}$$

le CFL λ étant celui de (2.3) et les β_p ceux de la partie 2. Outre sa simplicité, la relation (2.8) est précieuse à double titre comme nous allons le voir.

2.3 Étude de stabilité

D'une part, elle nous mène à la condition de *stabilité*. On a en effet

Théorème 2.1 *Une condition nécessaire et suffisante de stabilité pour le schéma 2-2m est*

$$\lambda \leq [\sum_{p=1}^{m}\beta_p]^{-1/2} \stackrel{\text{def}}{=} \Lambda_m. \tag{2.9}$$

DÉMONSTRATION- On peut assimiler (2.8) à une équation définissant implicitement Ω en fonction de $K \in \mathbb{R}$. On aura alors stabilité L^2 si et seulement si $\Omega(K)$ appartient à $\mathbb{R}$ pour tout K réel. Pour cela, il faut et il suffit que le second-membre de (2.8) soit inférieur à 1. Ce dernier atteignant son maximum en $K = \pi/2$, on doit avoir

$$\lambda^2 \sum_{p=1}^{m} \beta_p \leq 1,$$

ce qui achève la démonstration. □

Par définition, Λ_m est la constante de stabilité du schéma $2-m$. L'annexe B.1 dresse la liste de ces constantes de stabilité Λ_m. Les deux premiers d'entre eux sont incontestablement les plus populaires : $\Lambda_1 = 1$ et $\Lambda_2 = \sqrt{3}/2$.

Remarque 2.1 En vertu du Corollaire 1.2, Λ_m diminue avec m. La condition de stabilité est donc d'autant plus restrictive que m est grand. D'autre part, le seuil de stabilité pour le schéma 2-∞ est

$$\Lambda_\infty = \frac{2}{\pi}, \tag{2.10}$$

et la condition

$$\lambda \leq \frac{2}{\pi} \tag{2.11}$$

est une condition de stabilité suffisante et uniforme pour tous les schémas 2-2m. □

2.4 Étude de dispersion

D'autre part, la relation (2.8) permet de prévoir l'effet dispersif du schéma sur chaque composante k. Considérant (2.8) comme une équation permettant de définir ω en fonction de k, il est de tradition [5, 25, 26] dans ce genre de problème de définir

$$\begin{cases} v_\varphi = \dfrac{\omega}{k} & : \text{ la vitesse de phase,} \\ v_g = \dfrac{\partial \omega}{\partial k} & : \text{ la vitesse de groupe.} \end{cases}$$

Ces vitesses ne peuvent malheureusement pas s'exprimer en fonction des variables réduites Ω, K et λ que l'on a sélectionnées. Cependant, les grandeurs que l'on souhaite vraiment examiner sont les rapports de ces vitesses devant la célérité c du milieu, i.e.

$$\begin{cases} V_\varphi = \dfrac{v_\varphi}{c} & : \text{ vitesse de phase normalisée} \\ V_g = \dfrac{v_g}{c} & : \text{ vitesse de groupe normalisée} \end{cases}$$

et ceux-ci se calculent parfaitement à partir des variables réduites suivant

$$V_\varphi = \frac{1}{\lambda}\frac{\Omega}{K} \qquad \text{et} \qquad V_g = \frac{1}{\lambda}\frac{\partial \Omega}{\partial K}. \tag{2.12}$$

Le signe ∂ de dérivation partielle s'explique par le fait que l'on considère Ω comme une fonction des trois paramètres m, λ et K. Avant de parler de dérivée, il est de bon ton de mieux cerner le domaine de définition de cette fonction. Fixons $\lambda \in]\,0, \Lambda_m[$ et $K \in \mathbb{R}$. Il y a en principe deux solutions possibles pour Ω, à savoir

$$\Omega(m, \lambda, K) = \pm \arcsin\left\{ \lambda \,[\sum_{p=1}^{m} \beta_p \sin^{2p} K\,]^{-1/2} \right\}. \tag{2.13}$$

On retient la branche positive. La fonction obtenue est paire et π-périodique[1] par rapport à K : on se contentera donc de l'étudier sur $[0, \pi/2]$. Dans toute la suite de cette section, on suppose

$$\lambda \in]\,0, \Lambda_m[\qquad \text{et} \qquad K \in [0, \pi/2]. \tag{2.14}$$

1. les exposants de sin K étant tous pairs

Par ailleurs, gardons présent à l'esprit le fait que $\Omega(m,\lambda,0)=0$ et que par développement limité, on a

$$\begin{array}{rcccl} V_\varphi(m,\lambda,0) & = & \lim_{K\to 0} V_\varphi(m,\lambda,K) & = & 1 \\ V_g(m,\lambda,0) & = & \lim_{K\to 0} V_g(m,\lambda,K) & = & 1 \end{array}$$

pour tout $m\in\mathbb{N}^*$.

L'idéal serait d'avoir $V_\varphi = V_g = 1$ pour tout $K\in\mathbb{R}$. C'est ce que réalise l'équation exacte, dont la relation de dispersion est

$$\Omega^2 = \lambda^2 K^2. \tag{2.15}$$

Le schéma 2-2m s'en éloigne souvent assez, parfois beaucoup. Pour visualiser l'ampleur des dégâts, il faut tracer les courbes représentant V_φ et V_g en fonction de leurs arguments.Commençons par donner quelques résultats sur ces fonctions (Dans ce qui suit, on dit qu'une fonction $F(m,\lambda,K)$ définie pour $0\le\lambda\le\Lambda_m$ et $o\le K\le\pi/2$ est croissante par rapport à m si lorsque $\lambda\le\Lambda_{m+1}$, $F(m+1,\lambda,K)\ge F(m,\lambda,K))$).

Théorème 2.2 *Au sein de la famille 2-2m,*

1. *la pulsation $\Omega(m,\lambda,K)$ est positive et croît avec m, λ et K ;*
2. *les vitesses V_φ et V_g sont positives et croissent avec m et λ.*

DÉMONSTRATION- 1. La croissance de Ω en fonction de ses arguments est évidente d'après la formule (2.13) et du choix de la branche positive.

2_φ. Il est alors clair que $V_\varphi = \Omega/\lambda K$ est positif et croît avec m dès que $\lambda<\inf(\Lambda_m,\Lambda_{m+1})=\Lambda_{m+1}$. Les variables auxiliaires

$$\Upsilon_m(K) = [\sum_{p=1}^{m}\beta_p \sin^{2p} K]^{1/2} \qquad \text{et} \qquad C_m(K) = \frac{\Upsilon_m(K)}{K}$$

indépendantes de λ permettent d'exprimer la vitesse de phase réduite par

$$V_\varphi(K) = C_m(K)\,\frac{\arcsin\left[\lambda\Upsilon_m(K)\right]}{\lambda\Upsilon_m(K)}.$$

Par l'hypothèse (2.14), on a $\lambda\Upsilon_m<1$. Comme la fonction $z\mapsto \arcsin z/z$ est croissante sur $[0,1]$, il en est de même pour V_φ vis-à-vis de λ.

2_g. La pulsation Ω étant une fonction croissante de K, sa dérivée par rapport à K est positive, d'où $V_g\ge 0$. Notons Υ'_m la dérivée de Υ_m par rapport à K. En élevant au carré la relation

$$\frac{\partial\Omega}{\partial K} = \frac{\lambda\Upsilon'_m}{\sqrt{1-\lambda^2\Upsilon_m^2}},$$

puis en divisant par λ^2, on obtient

$$V_g^2 = \frac{\Upsilon'^2_m}{1-\lambda^2\Upsilon_m^2}.$$

Cette fonction est manifestement croissante par rapport à λ.

Il reste à montrer que V_g croît avec m, à λ et K fixés. On peut supposer $K>0$, ayant vu que $V_g(m,\lambda,0)=1$ pour tout m. Le principe de la preuve que nous allons présenter sera réutilisé pour une propriété analogue concernant le schéma 2m-2m.

Posons $F_m=\lambda^2\Upsilon_m^2$. En dérivant les deux membres de la relation de dispersion $\sin^2\Omega = F_m$ par rapport à K, et en appliquant quelques manipulations trigonométriques, on obtient (en ne faisant explicitement apparaître que la dépendance en m de V_g)

$$V_g^2(m) = \frac{F'^2_m}{4\lambda^2 F_m(1-F_m)},$$

où F'_m désigne la dérivée de F_m par rapport à K. De ce fait, montrer que $V_g^2(m+1)\ge V_g^2(m)$ revient à montrer que

$$\frac{F'^2_{m+1}}{4\lambda^2F_{m+1}(1-F_{m+1})} \ge \frac{F'^2_m}{4\lambda^2F_m(1-F_m)}$$

soit encore, comme tous les termes sont positifs, à montrer que

$$\frac{F'^2_{m+1}}{F'^2_m} \ge \frac{F_{m+1}(1-F_{m+1})}{F_m(1-F_m)} \tag{2.16}$$

sous la réserve habituelle $\lambda < \Lambda_{m+1}$. Appelons

$$\delta_m = \lambda^2 \beta_{m+1} \sin^{2m+2} K \qquad (< 1) \tag{2.17}$$

la différence entre F_{m+1} et F_m. Alors, on a

$$\frac{F_{m+1}}{F_m} = 1 + \frac{\delta_m}{F_m} \qquad \text{et} \qquad \frac{F'_{m+1}}{F'_m} = 1 + \frac{\delta'_m}{F'_m},$$

et l'inégalité à démontrer (2.16) devient

$$\left(1 + \frac{\delta'_m}{F'_m}\right)^2 \geq \left(1 + \frac{\delta_m}{F_m}\right)\left(1 - \frac{\delta_m}{1 - F_m}\right). \tag{2.18}$$

La clé essentielle de la démonstration réside dans l'inégalité

$$\frac{\delta'_m}{F'_m} \geq \frac{\delta_m}{F_m}. \tag{2.19}$$

Pareille minoration donnerait lieu en effet à

$$\begin{aligned} \left(1 + \frac{\delta'_m}{F'_m}\right)^2 &\geq \left(1 + \frac{\delta_m}{F_m}\right)\left(1 + \frac{\delta_m}{F_m}\right) \\ &\geq \left(1 + \frac{\delta_m}{F_m}\right)\left(1 - \frac{\delta_m}{1 - F_m}\right) \end{aligned}$$

qui est justement l'objectif (2.18). Cela dit, comment établit-on (2.19)? En dérivant (2.17) par rapport à $K > 0$, on a

$$\delta'_m = (2m + 2)\, \frac{\cos K}{\sin K}\, \delta_m.$$

Par conséquent,

$$\begin{aligned} \frac{\delta'_m}{\delta_m} F_m &= [(2m + 2)\mathrm{cotan} K]\, \lambda^2 \sum_{p=1}^{m} \beta_p \sin^{2p} K \\ &= \lambda^2 \sum_{p=1}^{m} (2m + 2) \beta_p \cos K \sin^{2p-1} K \\ &\geq \lambda^2 \sum_{p=1}^{m} 2p \beta_p \cos K \sin^{2p-1} K \end{aligned}$$

et la dernière somme n'est autre que F'_m. On aboutit ainsi à (2.19). □

Sur la Figure 2, on a tracé en fonction de K les trois grandeurs Ω, V_φ et V_g pour $\lambda = 0.6$ et plusieurs indices d'ordre m. On remarquera qu'il n'y a pas monotonie de V_φ et V_g vis à vis de K. Toutefois, les cas extrêmes $m = 1$ et $m = \infty$ méritent une attention spéciale.

Pour $m = 1$, la pulsation réduite $\Omega(1) = \arcsin(\lambda \sin K)$ est toujours inférieure à λK : on étudiera la fonction $z \mapsto \arcsin(\lambda z) - \lambda \arcsin(z)$ sur $z \in [0, 1]$ pour s'en convaincre. La vitesse de phase normalisée $V_\varphi(1)$ est donc toujours inférieure à 1. Par ailleurs, on a aussi $V_g(1) \leq 1$ car

$$V_g^2(1) = \frac{1 - \sin^2 K}{1 - \lambda^2 \sin^2 K}.$$

Autrement dit, le schéma 2-2 propage les composantes du signal à des vitesses plus lentes que la célérité c du milieu. Comme elles sont des fonctions croissantes de λ et constituent des approximations par défaut, on en déduit que l'erreur de dispersion numérique commise tant sur la vitesse de phase que de groupe diminue losque le coefficient λ augmente de 0 à 1. On peut enfin vérifier que V_φ et V_g sont des fonctions décroissantes de K.

Pour $m = \infty$, on assiste au phénomène inverse. Grâce à l'identité

$$K^2 = \sum_{p=1}^{\infty} \beta_p \sin^{2p} K,$$

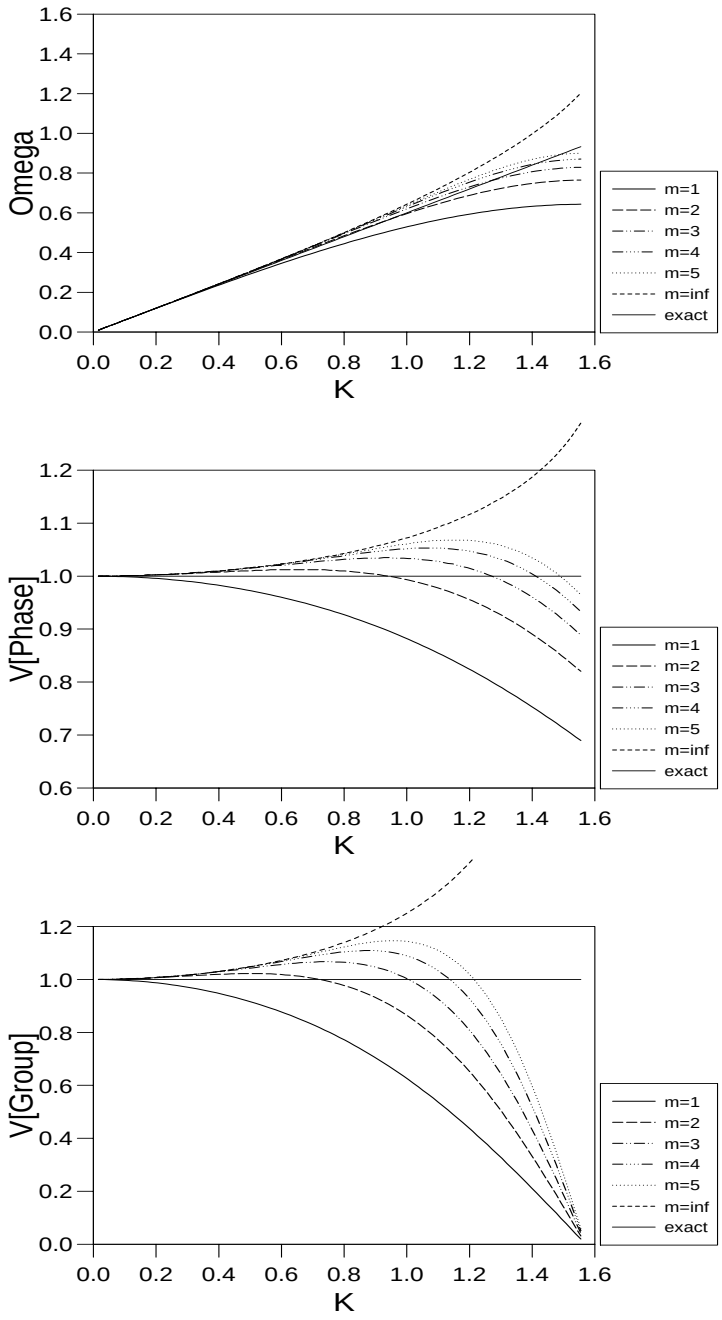

FIG. 2 – *Courbes de dispersion du schéma 2-2m pour* $\lambda = 0.6$.

démontrée pour $|K| \leq \pi/2$ au cours de la preuve du théorème 1.2, on a $\Omega(\infty) = \arcsin(\lambda K)$. Ce dernier est toujours plus supérieur à λK. En le dérivant par rapport à K et en élevant le résultat au carré, on trouve

$$V_g^2(\infty) = \frac{K^2}{1-\lambda^2 K^2}.$$

Le second membre est croissant en K. Comme sa valeur en $K=0$ est de 1, on en conclut que $V_g(\infty) \geq 1$. Le schéma 2-∞ propage les composantes du signal à des vitesses toutes plus rapides que la célérité c du milieu. Comme elles sont croissantes en λ et constituent cette fois des approximations par excès, l'erreur de dispersion augmente avec λ. Ceci est normal, la discrétisation en espace étant "exacte", le seul facteur dérreur est la discrétisation en temps: on réduit donc l'erruer en diminuant le pas de temps. On vérifie enfin facilement que V_φ et V_g sont des fonctions croissantes de K.

Pour $2 \leq m < \infty$, la Figure 2 indique que $V_\varphi(m)$ et $V_g(m)$ peuvent se situer de part et d'autre de 1 suivant la valeur du nombre d'onde réduit K. Plus précisément, le schéma 2-2m accélère les composantes correspondant à des K petits, et ralentit celles correspondant à des K grands.

3. Schémas 2m-2m pour les ondes acoustiques

Les schémas 2-2m sont naturels, mais ils présentent le défaut de n'être que d'ordre 2 en temps. Ils ne sont donc globalement que du second ordre si on choisit τ proportionnel à h, c'est à dire λ constant. Pour voir cela, mettons la relation de dispersion exacte (2.15) sous la forme

$$\sin^2 \Omega = \sin^2(\lambda K), \tag{3.1}$$

puis retranchons le second-membre de (2.8) à celui de (3.1). La quantité obtenue

$$\epsilon(m,\lambda,K) = \sin^2(\lambda K) - \lambda^2 \sum_{p=1}^{m} \beta_p \sin^{2p} K \tag{3.2}$$

est définie comme l'erreur de troncature du schéma 2-2m *dans l'espace des phases* (Ω, K). C'est l'équivalent de l'erreur de troncature (2.5) dans le domaine (t,x). Or, on verra plus loin, Lemme 3.1, que le développement de $\sin^2(\lambda K)$ en puissances de $\sin K$ est

$$\sin^2(\lambda K) = \lambda^2 \sum_{p=1}^{\infty} \beta_p E_{p-1}(\lambda) \sin^{2p} K,$$

où les E_{p-1} sont les polynômes définis en (1.17). Par suite,si on tyravaille, ce qui est le cas en pratique, avec λ constant

$$\epsilon(m,\lambda,K) = \lambda^2 \sum_{p=2}^{m} \beta_p [E_{p-1}(\lambda) - 1] \sin^{2p} K + \lambda^2 \sum_{p=m+1}^{\infty} \beta_p E_{p-1}(\lambda) \sin^{2p} K.$$

Comme $E_1(\lambda) \neq 1$ sauf en $\lambda = 0$, valeur qu'on a exclue pour le CFL, l'erreur de troncature fréquentielle est constamment en K^4 quel que soit m : la limitation due à l'ordre 2 en temps a été ainsi mise en évidence. Encore une fois, ceci est logique, τ et h variant proportionnellement, l'ordre effectif du schéma est imposé par la discrétisation la moins précise, en l'occurrence la discrétisation en temps.

3.1 Idée de la construction

On souhaite améliorer l'approximation de la dérivée seconde en temps. Une première idée est de remplacer Δ_τ par ${}_l\Delta_\tau$ dans (2.4) et d'obtenir ainsi le schéma 2l-2m

$$[{}_l\Delta_\tau] u_j^n - c^2 [{}_m\Delta_h] u_j^n = 0. \tag{3.3}$$

Au niveau de l'implémetation pratique, la discrétisation naïve (3.3) demanderait le stockage du champ u aux $2l$ pas de temps précédant u^{n+l}, ce qui est coûteux en termes de mémoire. Le problème est en fait plus profond. La relation de dispersion

$$\sum_{q=1}^{l} \beta_q \sin^{2q} \Omega = \lambda^2 \sum_{p=1}^{m} \beta_p \sin^{2p} K$$

admet malheureusement des racines Ω complexes pour $K \in \mathbb{R}$, et ce quelle que soit la valeur de λ. En effet, $K \in \mathbb{R}$ etant fixé, cette relation de dispersion apparaît comme une équation polynomiale paire de degré $2l$ en $X = sinK$. La positivité des coefficients β_q entraine que cette équation ná que deux racines rélles (opposées). Dès que $l > 1$, elle admet donc $2l-2$ autres racines non réelles. Les Ω correspondants ont donc une parie imaginaire non nulle, qui a le mauvais signe (si Ω est solution, -ω l'est aussi) pour la moitié d'entre eux. Ce schéma est donc inconditionnellement instable.

Il existe une méthode bien connue pour contourner cet obstacle qui, comme nous allons le voir, aboutit à un schéma stable n'utilisant que trois pas de temps. Considérons une fonction $u(x,t)$ très régulière. L'erreur de troncature commise en approchant $\frac{\partial^2 u}{\partial t^2}$ par $\Delta_\tau u$ vérifie:

$$\frac{\partial^2 u}{\partial t^2} = \Delta_\tau u - 2\sum_{q=2}^{l} \frac{\tau^{2q-2}}{(2q)!}\frac{\partial^{2q} u}{\partial t^{2q}} + O(\tau^{2l})$$

C'est la prèsence des dérivées d'ordre de plus en plus élevé en temps $\frac{\partial^{2q} u}{\partial t^{2q}}$ qui oblige à considérer de plus en plus de pas de temps dans la construction de ${}_l\Delta_\tau u$.

Toutefois, si u est une solution exacte de (2.1), alors, en dérivant deux fois par rapport au temps, on a

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

En répétant cette opération plusieurs fois, on obtient

$$\frac{\partial^{2q} u}{\partial t^{2q}} = c^{2q} \frac{\partial^{2q} u}{\partial x^{2q}}, \tag{3.4}$$

ce qui nous suggère de *remplaçant* dans les dérivées d'ordre supérieur en temps par des dérivées en espace. On obtient ainsi

$$\frac{\partial^2 u}{\partial t^2} = \Delta_\tau u - 2\sum_{q=2}^{l} \frac{\tau^{2q-2}}{(2q)!} c^{2q} \frac{\partial^{2q} u}{\partial x^{2q}} + O(h^{2l})$$

sous réserve que le rapport $\lambda/c = \tau/h$ soit maintenu constant.

Il importe de bien discrétiser les dérivées spatiales dans la somme ci-dessus. Compte-tenu de la présence des puissances τ^{2q-2}, il suffit d'approcher

$$\frac{\partial^{2q} u}{\partial x^{2q}} \quad \text{à l'ordre } 2l-2q+2.$$

On sait par (1.6) que la puissance $\Delta_h^q u$ n'approche $\partial^{2q} u/\partial x^{2q}$ qu'à l'ordre 2. Elle n'est donc pas directement utilisable si $q < l$. Afin de monter en ordre de précision, on peut envisager des corrections de type

$$\frac{\partial^{2q} u}{\partial x^{2q}} = \sum_{r=q}^{l} (-1)^{r-q}\, {}_l b_r^q\, h^{2r-2q} \Delta_h^r u + O(h^{2l-2q+2}), \tag{3.5}$$

où les coefficients ${}_l b_r^q$ sont à déterminer [2]. Par exemple, pour $q=2$ et $l=3$, on a

$$\frac{\partial^4 u}{\partial x^4} = \Delta_h^2 u - \tfrac{1}{6} h^2 \Delta_h^3 u + O(h^4).$$

Supposons tous les ${}_l b_r^q$ connus. À l'ordre $2l$, on aurait

$$\begin{aligned} \frac{\partial^2 u}{\partial t^2} &\approx \Delta_\tau u - 2\sum_{q=2}^{l} \frac{\tau^{2q-2} c^{2q}}{(2q)!} \sum_{r=q}^{l} (-1)^{r-q}\, {}_l b_r^q\, h^{2r-2q} \Delta_h^r u \\ &= \Delta_\tau u - c^2 \sum_{r=2}^{l} (-1)^{r-1} [\sum_{q=2}^{r} (-1)^{q-1} \frac{2\, {}_l b_r^q}{(2q)!} \left(\frac{c\tau}{h}\right)^{2q-2}] h^{2r-2} \Delta_h^r u. \end{aligned}$$

2. ils sont en réalité indépendants de l

Le nouveau schéma 2l-2m s'écrit

$$\Delta_\tau u_j^n = c^2 \sum_{p=1}^{m} (-1)^{p-1}\, b_p\, h^{2p-2} \Delta_h^p u_j^n + c^2 \sum_{r=2}^{l} (-1)^{r-1} Q_r^l\left(\frac{c\tau}{h}\right) h^{2r-2} \Delta_h^r u_j^n,$$

où Q_r^l désigne le polynôme

$$Q_r^l(\lambda) = \sum_{q=2}^{r} (-1)^{q-1} \frac{2\ {}_l b_r^q}{(2q)!} \lambda^{2q-2}.$$

L'ordre global du schéma étant $2 \times \inf(l,m)$, il semble raisonnable de spécifier $l = m$. C'est l'hypothèse que l'on va émettre. Le schéma 2m-2m prend alors la forme

$$\Delta_\tau u_j^n - c^2 \sum_{p=1}^{m} (-1)^{p-1}\, \Xi_p^m\left(\frac{c\tau}{h}\right) h^{2p-2} \Delta_h^p u_j^n = 0, \tag{3.6}$$

dans laquelle le polynôme Ξ_p^m est de degré $2p-2$. Cela étant, le calcul effectif des coefficients ${}_l b_r^q$ vérifiant (3.5) s'annonce fastidieux. Cette direction ne sera pas explorée davantage, d'autant qu'il existe une méthode plus élégante pour obtenir les polynômes Ξ_p^m.

La méthode que nous venons de décrire pour obtenir des schémas d'ordre élevé en temps est parfois référencée dans la littérature comme la méthode de l'équation modifiée [24]. Tout se passe en effet comme si on discrétisait l'équation des ondes modifiée:

$$\frac{\partial^2 u}{\partial t^2} - c^2 \sum_{p=1}^{m} (-1)^{p-1}\, \Xi_p^m\left(\frac{c\tau}{h}\right) h^{2p-2} \frac{\partial^{2p} u}{\partial x^{2p}} u = 0$$

3.2 Forme explicite du schéma

Commençons par donner un résultat préliminaire, déjà évoqué au début de cette section..

Lemme 3.1 *Pour tout $\lambda \in \mathbb{R}$ et tout $K \in [-\pi/2, \pi/2]$, on a*

$$\sin^2(\lambda K) = \lambda^2 \sum_{p=1}^{\infty} \beta_p\, E_{p-1}(\lambda) \sin^{2p} K. \tag{3.7}$$

Si de plus $|\lambda| \leq 1$, alors tous les termes de la série sont positifs.

DÉMONSTRATION- Comme $|K| \leq \pi/2$, il existe $z \in [-1,1]$ tel que $K = \arcsin z$. Vis-à-vis de la variable z, la fonction

$$\Upsilon(\lambda, z) = \sin^2(\lambda \arcsin z)$$

satisfait à l'équation différentielle

$$(1-z^2)\frac{\partial^2 \Upsilon}{\partial z^2} - z\frac{\partial \Upsilon}{\partial z} + 4\lambda^2 \Upsilon = 2\lambda^2. \tag{3.8}$$

En injectant formellement la série $\Upsilon(\lambda, z) = \sum T_p(\lambda) z^{2p}$ dans (3.8), on obtient

$$2T_1(\lambda) + \sum_{p=1}^{\infty} \left[(2p+1)(2p+2)T_{p+1}(\lambda) - 4(p^2-\lambda^2)T_p(\lambda)\right] z^{2p} = 2\lambda^2,$$

à partir de quoi il vient, par identification, que

$$T_1(\lambda) = 1 \qquad \text{et} \qquad T_{p+1}(\lambda) = \frac{4(p^2-\lambda^2)}{(2p+1)(2p+2)}\, T_p(\lambda).$$

Il en découle que

$$T_p(\lambda) = \frac{2^{2p-1}}{(2p)!}\, \lambda^2 (1^2-\lambda^2)(2^2-\lambda^2)\ldots((p-1)^2-\lambda^2) = \lambda^2 \beta_p E_{p-1}(\lambda).$$

En vue de justifier rigoureusement le calcul précédent, on forme a posteriori la série $\Upsilon(\lambda,z)=\sum T_p(\lambda)z^{2p}$, la valeur des $T_p(\lambda)$ étant imposée par la formule ci-dessus. Puisque

$$\lim_{p\to\infty}\frac{T_{p+1}(\lambda)}{T_p(\lambda)}=1\,,$$

le rayon de convergence est égal à 1. En dérivant terme à terme à l'intérieur du disque de convergence, on montre que Υ obéit à (3.8) sur $]-1,1[$. On va même montrer que la convergence a lieu pour $z=\pm1$.

Remarquons auparavant que, par définition, le polynôme

$$E_{p-1}(\lambda)=\prod_{r=1}^{p-1}\left(1-\frac{\lambda^2}{r^2}\right)$$

s'annule dès que $\lambda\in\mathbb{Z}^*\cap[1-p,p-1]$. Si $|\lambda|=1$, tous les $E_{p-1}(1)$ sont nuls sauf $E_0(1)=1$: la série est à termes positifs. Au sens large, on vérifie aisément que $E_{p-1}(\lambda)$ est du même signe que

$$(-1)^{\ell(p,|\lambda|)}\qquad\text{avec}\qquad\ell(p,\rho)=\inf(p-1,\lfloor\rho\rfloor),$$

où $\lfloor\rho\rfloor$ désigne la partie entière du réel $\rho\in\mathbb{R}^+$. Par conséquent, si $|\lambda|<1$, on a $\ell(p,|\lambda|)=0$ pour tout $p\geq1$, et tous les termes de la série (3.7) sont positifs.

Du fait de la parité en λ de Υ et des E_{p-1}, on va supposer $\lambda>0$ mais arbitraire. Les termes de la série (3.7) sont alors de même signe à partir du rang $p=\lfloor\lambda\rfloor+1$. Quitte à travailler avec la série opposée, on décrète que tous ces termes-là sont positifs. Écrivons

$$\Upsilon(\lambda,z)-\sum_{p=1}^{\lfloor\lambda\rfloor}\beta_p\,E_{p-1}(\lambda)z^{2p}=\sum_{p=\lfloor\lambda\rfloor+1}^{\infty}\beta_p\,E_{p-1}(\lambda)z^{2p}\tag{3.9}$$

pour $z\in]-1,1[$, et notons Υ_λ la fonction de z qui est au premier-membre de (3.9). Soit $N>\lfloor\lambda\rfloor+1$. Comme tous les termes du second-membre de (3.9) sont positifs, on a

$$\sum_{p=\lfloor\lambda\rfloor+1}^{N}\beta_p\,E_{p-1}(\lambda)z^{2p}\leq\Upsilon_\lambda(z).$$

Chacun des deux membres de cette inégalité admet une limite finie quand $z\to1$. Les valeurs limites respectent le sens de l'inégalité, i.e.

$$\sum_{p=\lfloor\lambda\rfloor+1}^{N}\beta_p\,E_{p-1}(\lambda)\leq\Upsilon_\lambda(1).$$

Cela prouve que la série à termes positifs $\sum_{\lfloor\lambda\rfloor+1}\beta_p\,E_{p-1}(\lambda)$ converge et sa somme est majorée par $\Upsilon_\lambda(1)$. Par ailleurs, en faisant tendre N vers l'infini dans l'inégalité évidente

$$\sum_{p=\lfloor\lambda\rfloor+1}^{N}\beta_p\,E_{p-1}(\lambda)z^{2p}\leq\sum_{p=\lfloor\lambda\rfloor+1}^{N}\beta_p\,E_{p-1}(\lambda),$$

on obtient

$$\Upsilon_\lambda(z)\leq\sum_{p=\lfloor\lambda\rfloor+1}^{\infty}\beta_p\,E_{p-1}(\lambda).$$

Il ne reste plus qu'à faire $z\to1$ pour minorer la somme au second-membre par $\Upsilon_\lambda(1)$. L'égalité entre $\Upsilon_\lambda(1)$ et $\sum_{\lfloor\lambda\rfloor+1}\beta_p\,E_{p-1}(\lambda)$ est le résultat de convergence désiré pour $z=1$. La convergence pour $z=-1$ (la série est alors alternée à partir d'un certain rang) s'en déduit immédiatement : on a même convergence absolue. □

On peut maintenant énoncer le

Théorème 3.1 *Le schéma 2m-2m est donné par*

$$\Delta_\tau u_j^n-c^2\sum_{p=1}^{m}(-1)^{p-1}\,b_p\,E_{p-1}\Big(\frac{c\tau}{h}\Big)\,h^{2p-2}\Delta_h^p u_j^n=0,\tag{3.10}$$

où les E_{p-1} sont les polynômes définis par (1.17). Sa relation de dispersion est

$$\sin^2 \Omega = \lambda^2 \sum_{p=1}^{m} \beta_p \, E_{p-1}(\lambda) \sin^{2p} K. \tag{3.11}$$

DÉMONSTRATION- On a vu que le schéma 2m-2m est de la forme (3.6) et que les Ξ_p^m peuvent a priori dépendre de m. La relation de dispersion associée à un tel schéma s'écrit

$$\sin^2 \Omega = \sum_{p=1}^{m} \Theta_p^m(\lambda) \sin^{2p} K.$$

Si l'erreur de troncature en (x,t) est d'ordre $2m$, celle en (Ω, K) doit être d'ordre $2m+2$, ce qui oblige

$$\sin^2(\lambda K) = \sum_{p=1}^{m} \Theta_p^m(\lambda) \sin^{2p} K + O(K^{2m+2}).$$

Autrement dit, les $\Theta_p^m(\lambda)$ sont les m premiers coefficients du développement de $\sin^2(\lambda K)$ en série de $\sin K$. Ils sont donc indépendants de m et en vertu du Lemme 3.1, on a

$$\Theta_p^m(\lambda) = \lambda^2 \, \beta_p \, E_{p-1}(\lambda).$$

En revenant dans le domaine (x,t) par l'intermédiaire de (1.26), on obtient

$$\Xi_p^m(\lambda) = b_p \, E_{p-1}(\lambda).$$

ce qui est le résultat désiré. □

Tel qu'il est défini, le schéma 2m-2m généralise le schéma 4-4 de Cohen [6]. La partie temporelle de (3.10) ne fait intervenir que les deux pas de temps précédant u^{n+1}. Sa partie spatiale conduit à une molécule de $2m+1$ points dont on peut évaluer une fois pour toutes[3] les coefficients : c'est ce que décrit l'annexe B.2 d'un point de vue plus pratique. En résumé, le schéma 2m-2m a la même structure informatique et le même nombre d'opérations élémentaires que le schéma 2-2m. Autrement dit, en passant du schéma 2-2m au schéma 2m-2m on gagne en précision à coût nul !

3.3 Étude de stabilité

À partir de la relation de dispersion (3.11), on obtient

Théorème 3.2 *Le schéma 2m-2m est stable sous la condition:*

$$\lambda \leq 1. \tag{3.12}$$

Plus précisément, le plus grand nombre $\lambda_{\max}^m$ pour lequel le schéma 2m-2m est stable lorsque λ appartient à l'intervalle $[0, \lambda_{\max}^m]$ est égal à 1, pour tout $m \geq 1$.

Aux points entiers $\lambda = 1, 2, \ldots, m$, le schéma 2m-2m dégénère en un schéma aux caractéristiques (donc exact) de pas d'espace λh.

DÉMONSTRATION- D'après le Lemme 3.1, si $\lambda \in [0,1]$, tous les termes de la série (3.7) sont positifs. Le second-membre de la relation de dispersion (3.11) se laisse alors majorer par

$$\lambda^2 \sum_{p=1}^{m} \beta_p \, E_{p-1}(\lambda) \sin^{2p} K \leq \sin^2(\lambda K) \leq 1,$$

de sorte que la solution $\Omega(m, \lambda, K)$ est réelle. Cela reflète la stabilité du schéma pour $\lambda \in [0,1]$.

Appelons F et F_m les second-membres respectifs de (3.7) et (3.11), vus comme une fonction de λ et de K. Si $\lambda \in \mathbb{N} \cap [1,m]$, on a $F_m(\lambda, K) = F(\lambda, K)$ car $E_{p-1}(\lambda) = 0$ pour $p \geq m+1$. La relation de dispersion (3.11) devient alors exacte. Quant au schéma, il devient

$$\Delta_\tau u_j^n - c^2 \sum_{p=1}^{m} \Delta_{\lambda h} u_j^n = 0, \tag{3.13}$$

3. avant de démarrer les itérations en temps

avec $\tau = \lambda h$. Autrement dit, c'est un schéma caractéristique.

Au point $\lambda = 1$, on a $F_m(1, \pi/2) = \sin^2(\pi/2) = 1$. Or, on établira à la Proposition 3.3 de la section suivante que

$$\frac{\partial F_m}{\partial \lambda}(\lambda, K) > 0$$

pour tout $\lambda \in]0, 1]$. La fonction F_m continue donc à croître pour $\lambda \in [1, 1+\epsilon[$. Sur cet intervalle, $F_m(\lambda, \pi/2) > 1$. On ne peut plus trouver $\Omega \in \mathbb{R}$, d'où l'instabilité. □

3.4 Étude de dispersion

Le second-membre de la relation de dispersion (3.11) définit Introduisons la fonction de deux variables

$$F_m(\lambda, K) = \lambda^2 \sum_{p=1}^{m} \beta_p \, E_{p-1}(\lambda) \sin^{2p} K \tag{3.14}$$

de telle sorte que la relation de dispersion (3.11) s'ecrit encore:

$$sin^2\Omega = F_m(\lambda, K) \tag{3.15}$$

Par commodité nous appellerons cette fonction *fonction forme* de rang m, pour la simple raison que c'est elle qui dicte la forme de la courbe $\Omega(m, \lambda, K)$. Les variables (λ, K) décrivent $\mathbb{R} \times [-\pi/2, \pi/2]$. En préparation de nombreux résultats concernant le schéma 2m-2m, examinons le comportement de F_m par rapport à ses variables.

3.4.1 Quelques préliminaires

Vis-à-vis de K, la fonction F_m est la somme de rang m de la série

$$F(\lambda, K) = \sum_{p=1}^{\infty} \beta_p \, E_{p-1}(\lambda) \sin^{2p} K, \tag{3.16}$$

laquelle vérifie l'équation différentielle

$$F'' + 4\lambda^2 F - 2\lambda^2 = 0, \tag{3.17}$$

que l'on déduit immédiatement de (3.8) par le changement de variable $z = \sin K$ et le changement de fonction $\Upsilon(\lambda, z) = F(\lambda, K)$. Le signe $'$ indique la dérivée par rapport à K. Si l'on applique l'opérateur différentiel du premier-membre de (3.17) à F_m au lieu de F, on obtient après calculs

$$F_m'' + 4\lambda^2 F_m - 2\lambda^2 = -R_{m+1}(\lambda) \sin^{2m} K, \tag{3.18}$$

avec

$$R_{m+1}(\lambda) = (2m+1)(2m+2)\, \beta_{m+1} \, \lambda^2 E_m(\lambda). \tag{3.19}$$

Usage sera copieusement fait de l'équation différentielle (3.18).

Vis-à-vis de λ, la fonction F_m est un polynôme pair de degré $2m$. Les valeurs entières $\lambda = 0, \pm 1, \pm 2, \ldots, \pm m$ étant toutes racines de $\lambda^2 E_{p-1}(\lambda)$ pour $p \geq m+1$, il est facile de montrer que

$$\forall\, \lambda \in \mathbb{Z} \cap [-m, m], \qquad F_m(\lambda, K) = F(\lambda, K).$$

Le polynôme F_m prend donc exactement les valeurs de F aux $2m+1$ points entiers $\lambda = 0, \pm 1, \pm 2, \ldots, \pm m$. Comme son degré est $2m$, on en conclut que F_m est le polynôme d'interpolation de Lagrange de F aux $2m+1$ points énumérés ci-dessus.

Le schéma 2m-2m étant d'ordre m, l'erreur entre F_m et F doit être en K^{2m+2} pour $K \to 0$. Le fait de voir F_m comme polynôme d'interpolation de F permet d'étendre cette estimation d'erreur à tout $K \in [0, \pi/2]$.

3.4.2 Une estimation d'erreur uniforme

Proposition 3.1 *L'estimation d'erreur*

$$0 \leq F(\lambda, K) - F_m(\lambda, K) \leq 2m(m+1)\beta_{m+1}\lambda E_m(\lambda) K^{2m+2} \tag{3.20}$$

est valable pour tout $K \in [0, \pi/2]$ *et* $\lambda \in]0, 1[$.

DÉMONSTRATION- On sait déjà que $F_m \leq F$ par le Lemme 3.1. La théorie de l'interpolation [10, 21] dit que le reste $F - F_m$ peut s'écrire

$$F(\lambda, K) - F_m(\lambda, K) = \frac{\lambda(\lambda^2 - 1^2)(\lambda^2 - 2^2)\dots(\lambda^2 - m^2)}{(2m+1)!} \frac{\partial^{2m+1} F}{\partial \lambda^{2m+1}}(\xi_\lambda)$$

avec $\xi_\lambda \in [-m, m]$. Dans le cas présent, la dérivée

$$\frac{\partial^{2m+1} F}{\partial \lambda^{2m+1}}(\xi_\lambda) = (-1)^m 2^{2m} K^{2m+1} \sin(2\xi_\lambda K)$$

peut être majorée par

$$\begin{aligned} \left| \frac{\partial^{2m+1} F}{\partial \lambda^{2m+1}}(\xi_\lambda) \right| &\leq 2^{2m} K^{2m+1} |2\,\xi_\lambda K| \\ &\leq m 2^{2m+1} K^{2m+2}, \end{aligned}$$

ce qui donne

$$|F(\lambda, K) - F_m(\lambda, K)| \leq m \frac{2^{2m+1}\, m!^2}{(2m+1)!} \lambda E_m(\lambda)\, K^{2m+2}.$$

Le second-membre n'est autre que $2m(m+1)\,\beta_{m+1}\,\lambda E_m(\lambda)\,K^{2m+2}$. □Notons que l'estimation d'erreur (3.20) est à nouveau une illustration de l'ordre du schéma, elle ne permet de vérifier que $F_m(\lambda, K)$ converge, lorsque m tend vers l'infini, vers $F_m(\lambda, K)$ que lorsque $K < 1$. En ce sens, cette estimation n'est pas optimale.

3.4.3 Propriétés des courbes de dispersion

On est maintenant en mesure de passer en revue les propriétés de dispersion du schéma 2m-2m. On suppose désormais

$$\lambda \in]\,0,\,1\,[\qquad \text{et} \qquad K \in [0, \pi/2].$$

La fonction F_m est alors à valeurs dans $[0, 1]$. Elle est croissante en K, ainsi que

$$\Omega = \arcsin \sqrt{F_m(\lambda, K)},$$

d'où la positivité des vitesses normalisées V_φ et V_g définies par (2.12). On a aussi $V_\varphi(m, \lambda, 0) = V_g(m, \lambda, 0) = 1$ par développement limité. Notre premier résultat concerne la vitesse de groupe.

Lemme 3.2 *Dans le schéma 2m-2m, on a toujours* $V_g \leq 1$.

DÉMONSTRATION- On multiplie les deux membres de l'équation différentielle (3.18) par F'_m. Le premier-membre est reconnu comme étant la dérivée par rapport à K de

$$\tfrac{1}{2}[\,F'^2_m - 4\lambda^2 F_m(1 - F_m)\,], \tag{3.21}$$

tandis que le second-membre est toujours négatif. La fonction (3.21) est par conséquent décroissante. Sa valeur en $K = 0$ étant nulle, elle est négative sur $K \in [0, \pi/2]$. Cela implique

$$\frac{F'^2_m}{4\lambda^2 F_m(1 - F_m)} \leq 1.$$

Or, le premier-membre de cette inégalité n'est autre que V_g^2, comme on peut le voir en dérivant $\sin^2 \Omega = F_m$ par rapport à K. □

Dans la section 3.1, on a vu que la vitesse de groupe normalisée V_g pouvait dépasser 1 pour le schéma 2-2m. La majoration du Lemme 3.2 est donc propre au schéma 2m-2m. Également spécifiques au schéma 2m-2m sont les résultats suivants.

Théorème 3.3 *Dans le schéma 2m-2m,*

1. *la pulsation* Ω *est concave en* K ;
2. *les vitesses* V_φ *et* V_g *décroissent avec* K, *et l'on a* $V_g \leq V_\varphi \leq 1$.

DÉMONSTRATION- 1. En dérivant deux fois la relation $\sin^2\Omega = F_m$, on obtient

$$\frac{\partial^2\Omega}{\partial K^2}\sin(2\Omega) + 2\left(\frac{\partial\Omega}{\partial K}\right)^2\cos(2\Omega) = F_m''.$$

De l'égalité

$$\left(\frac{\partial\Omega}{\partial K}\right)^2 = \frac{F_m'^2}{4F_m(1-F_m)},$$

on tire que

$$\frac{\partial^2\Omega}{\partial K^2}\sin(2\Omega) = F_m'' - \frac{F_m'^2}{2F_m(1-F_m)}[1-2F_m]. \tag{3.22}$$

Puisque $\sin(2\Omega) \geq 0$, il suffit de démontrer que le second-membre de (3.22) est négatif pour conclure à la concavité de Ω. Formons la différence

$$G_m = F_m''F_m(1-F_m) - \tfrac{1}{2}F_m'^2(1-2F_m).$$

La dérivée de G_m par rapport à K est égale, après simplification, à

$$G_m' = F_m'''F_m(1-F_m) + F_m'^3.$$

Comme G_m s'annule en $K=0$, il suffit de démontrer que $G_m' \leq 0$ pour s'assurer de la négativité de G_m. Pour cela, on fait de nouveau appel à l'équation (3.18) que l'on dérive par rapport à K, soit

$$F_m''' + 4\lambda^2F_m' = -2m\,R_{m+1}(\lambda)\,\sin^{2m-1}K\cos K. \tag{3.23}$$

Le second-membre de (3.23) étant toujours négatif, il en est de même pour le premier-membre. En multipliant ce dernier par $F_m(1-F_m)$, on peut donc dire que

$$F_m'''F_m(1-F_m) + 4\lambda^2F_m'F_m(1-F_m) \leq 0. \tag{3.24}$$

Or (cf. preuve du Lemme 3.2), on sait que

$$4\lambda^2F_m(1-F_m) \geq F_m'^2,$$

ce qui permet de minorer le premier-membre de (3.24) par G_m'.

2. Il s'ensuit que

$$\frac{\partial V_g}{\partial K} = \frac{1}{\lambda}\frac{\partial^2\Omega}{\partial K^2} \leq 0,$$

d'où la décroissance de V_g en K. En ce qui concerne V_φ, on a

$$\frac{\partial V_\varphi}{\partial K} = \frac{\partial}{\partial K}\left(\frac{\Omega}{K}\right) = \frac{1}{\lambda}\frac{1}{K^2}\left[K\frac{\partial\Omega}{\partial K} - \Omega\right],$$

et il suffit de démontrer que la fonction entre crochet est négative. Sa dérivée

$$\frac{\partial\Omega}{\partial K} + K\frac{\partial^2\Omega}{\partial K^2} - \frac{\partial\Omega}{\partial K} = K\frac{\partial^2\Omega}{\partial K^2}$$

est négative d'après 1. Elle est donc décroissante et en fait négative car sa valeur en $K=0$ est nulle.

Le crochet reste négatif si on le divise par K, ce qui donne

$$\frac{\partial\Omega}{\partial K} \leq \frac{\Omega}{K},$$

soit $V_g \leq V_\varphi$. Enfin, la décroissance de V_φ entraîne $V_\varphi \leq V_\varphi(K=0) = 1$. □

Il ressort de ce théorème que l'erreur de groupe $1-V_g$ est toujours plus grande que l'erreur de phase $1-V_\varphi$. Au voisinage de $K=0$, on a le résultat classique, valable pour n'importe quel schéma d'ordre m.

$$\lim_{K\to 0}\frac{1-V_g(m,\lambda,K)}{1-V_\varphi(m,\lambda,K)} = 2m+1. \tag{3.25}$$

Le rapport entre les deux erreurs de vitesse tend vers l'infini lorsque $m \to \infty$. Cela signifie juste que vis-à-vis de m, ces deux erreurs ne tendent pas vers 0 à la même cadence. La dépendance des grandeurs Ω, V_φ et V_g par rapport à m s'étudie d'ailleurs très facilement.

Proposition 3.2 *Au sein de la famille 2m-2m, la pulsation Ω et les vitesses V_φ et V_g croissent avec m.*

DÉMONSTRATION- Le fait que Ω et V_φ sont croissants par rapport à m est trivial. Seule la croissance de V_g est quelque peu délicate : on procède comme dans la démonstration du Théorème 2.2. □

Il nous reste à élucider le comportement du schéma 2m-2m vis-à-vis du quotient de stabilité λ. Dans le schéma 2-2m, la fonction forme (2.8) est simplement proportionnelle à λ^2, ce qui facilite énormément la démonstration de la croissance des grandeurs Ω, V_φ et V_g par rapport à λ. Dans le schéma 2m-2m, la fonction forme F_m fait intervenir λ par le biais des fonctions non monotones $\lambda^2 E_{p-1}(\lambda)$, ce qui rend le travail beaucoup plus délicat.

Proposition 3.3 *Dans le schéma 2m-2m, Ω croît avec λ sur $[0,1]$.*

DÉMONSTRATION- Il suffit de prouver que la fonction $F_m = \sin^2\Omega$ est croissante vis-à-vis de λ. La dérivation terme à terme par rapport à λ de (3.14) conduit à une impasse. C'est pourquoi on va emprunter une voie plus détournée via une représentation intégrale de F_m.

La fonction de Green [18], définie par

$$G_\lambda(K,\chi) = \begin{cases} \dfrac{1}{2\lambda}\sin[2\lambda(K-\chi)] & \text{si } \chi \le K \\ 0 & \text{si } \chi \ge K \end{cases} \tag{3.26}$$

vérifie, au sens des distributions, l'équation

$$\frac{\partial^2 G_\lambda}{\partial \chi^2}(K,\chi) + 4\lambda^2 G_\lambda(K,\chi) = \delta_K(\chi).$$

La théorie générale [18] de Green, appliquée à l'équation différentielle (3.18), montre alors que

$$F_m(\lambda,K) = \int_0^{\pi/2} \left\{2\lambda^2 - R_{m+1}(\lambda)\sin^{2m}\chi\right\} G_\lambda(K,\chi)\, d\chi, \tag{3.27}$$

où R_{m+1} a été défini en (3.19). En introduisant

$$U_m(K) = \tfrac{1}{2}(2m+1)(2m+2)\beta_{m+1}\sin^{2m}K \tag{3.28}$$

et en tenant compte de (3.26), on ramène (3.27) à

$$F_m(\lambda,K) = \int_0^K \{1 - E_m(\lambda)U_m(\chi)\}\ \lambda\sin[2\lambda(K-\chi)]\, d\chi. \tag{3.29}$$

Dérivons alors F_m par rapport à chaque variable et remultiplions le résultat par la variable concernée. Par rapport à K, on a

$$K\frac{\partial F_m}{\partial K} = 2\lambda^2 K \int_0^K \{1 - E_m(\lambda)U_m(\chi)\}\cos[2\lambda(K-\chi)]\, d\chi. \tag{3.30}$$

Par rapport à λ, on a

$$\begin{aligned} \lambda\frac{\partial F_m}{\partial \lambda} &= K\frac{\partial F_m}{\partial K} - \lambda^2 E_m'(\lambda)\int_0^K U_m(\chi)\sin[2\lambda(K-\chi)]\, d\chi \\ &+ F_m - 2\lambda^2\int_0^K \{1 - E_m(\lambda)U_m(\chi)\}\,\chi\cos[2\lambda(K-\chi)]\, d\chi, \end{aligned}$$

où l'on s'est servi de (3.30). Le dernier terme du second-membre de l'équation précédente peut être intégré par parties. On obtient

$$2\lambda^2\int_0^K \{1 - E_m(\lambda)U_m(\chi)\}\,\chi\cos[2\lambda(K-\chi)]\, d\chi = F_m - \lambda E_m(\lambda)\int_0^K U_m'(\chi)\chi\sin[2\lambda(K-\chi)]\, d\chi$$

si bien que les F_m s'éliminent, et

$$\begin{aligned} \lambda\frac{\partial F_m}{\partial \lambda} &= K\frac{\partial F_m}{\partial K} - \lambda^2 E_m'(\lambda)\int_0^K U_m(\chi)\sin[2\lambda(K-\chi)]\, d\chi \\ &+ \lambda E_m(\lambda)\int_0^K U_m'(\chi)\chi\sin[2\lambda(K-\chi)]\, d\chi \end{aligned} \tag{3.31}$$

Il est clair que les polynômes $E_m(\lambda)$ sont décroissants sur $[0,1]$ et donc que $E'_m(\lambda) < 0$, tandis que les $U_m(\chi)$ sont croissants sur $[0,\pi/2]$ et donc que $U'_m(\chi) > 0$. D'autre part, on a vu que F_m est croissant par rapport à K. Tous les termes du second-membre de (3.31) étant positifs, il en est de même pour le premier-membre. □Nous n'avons pas réussi à établir des résultats analogues pour les vitesses de phase et de groupe V_φ et V_g. Toutefois, sur la base de résultats numériques, nous proposons la conjecture suivante:

Conjecture 3.1 *Les vitesses V_φ et V_g croissent avec λ sur $]0,1[$.*

Si le résultat de la conjecture 3.1 sont exacts, compte ternu que les vitesses de phase et de groupe V_φ et V_g sont approchées par défaut (cf. théorème 3.3), la dispersion numérique diminue donc lorsque λ augmente de 0 à 1, résultat bien connu pour le schéma d'ordre 2 (cf la fin de la section 2).

Illustrons graphiquement les résultats précédents. La Figure 3 représente les grandeurs Ω, V_φ et V_g en fonction de K avec différents indices d'ordre m et pour $\lambda = 0.6$. Cette figure est à comparer avec la Figure 2. Dans le cas du schéma 2m-2m, le dessin confirme que (i) ni V_φ ni V_g ne peut dépasser 1 et (ii) qu'ils sont décroissants en K. Les courbes avec $m = \infty$ coïncident avec les courbes exactes car l'on est sur $[0,\pi/2]$.

La Figure 4 correspond à un CFL plus grand, à savoir $\lambda = 0.9$. On voit que les vitesses V_φ et V_g se rapprochent de 1, uniformément en K.

La Figure 5 vient en appui de la Conjecture 3.1. Elle représente les grandeurs Ω, à $K = \pi/4$ fixé, et ce pour différents ordres m.

4. Conclusion

À l'aide d'outils mathématiques relativement élémentaires, on a pu établir un grand nombre de résultats concernant le laplacien discret d'ordre élevé et le schéma 2m-2m pour l'équation des ondes acoustiques en milieu 1-D homogène. Doté par construction d'agréables propriétés de convergence, ce schéma l'emporte nettement sur le schéma 2-2m, et ce à coût égal.

La transposition théorique aux cas 2-D et 3-D du schéma 2-2m ne pose pas de problèmes majeurs: il suffit d'appliquer le laplacien discret ${}_m\Delta_h$ dans chaque direction. L'anisotropie s'introduira. La condition de stabilité sera aussi affectée: si d est la dimension, le CFL sera de $\Lambda_m/\sqrt{d}$. Ceci alourdira notablement les calculs.

Par contre, la transposition aux cas 2-D et 3-D du schéma 2m-2m n'est pas immédiate. Elle nécessite en effet le développement en série de la relation de dispersion exacte par rapport à plusieurs variables. Chacun des termes du développement être le symbole d'un certain opérateur local en espace, ce qui n'est guère évident.

Remerciements

Nous tenons à exprimer toute notre gratitude envers Alain Bamberger, dont l'encouragement dans la rédaction de cet article nous a été extrêmement profitable.

Références

[1] R. M. Alford, K. R. Kelly et D. M. Boore, Accuracy of Finite-Difference Modeling of the Acousic Wave Equation, *Geophysics* **39**, 834–842 (1974).

[2] L. Anné, W. W. Symes et Q. H. Tran, Dispersion and Cost Analysis of Some Finite Difference Schemes in One-Parameter Acoustic Wave Modelling, *Computational Geosciences* **1**, 1–33 (1997).

[3] R. Ayoub, Euler and the Zeta Function, *Amer. Math. Monthly* **81**, 1067–1086 (1974).

[4] A. Bamberger, G. Chavent et P. Lailly, *Étude de Schémas Numériques pour les Équations de l'Élastodynamique Linéaire*, Rapport de Recherche **41**, INRIA, Rocquencourt, 1980.

[5] L. Brillouin, *Wave Propagation and Group Velocity*, Academic Press, New-York, 1960.

[6] G. Cohen, Différences Finies d'Ordre 4 pour l'Équation des Ondes, dans *Méthodes Numériques d'Ordre Élevé pour les Ondes en Régime Transitoire*, École des Ondes, G. Cohen éd., Collection Didactique, INRIA Rocquencourt, 1994.

[7] G. Cohen et P. Joly, Fourth Order Schemes for the Heterogeneous Acoustics Equation, *Comp. Meth. in Appl. Mech. Eng.* **80**, 397–407 (1990).

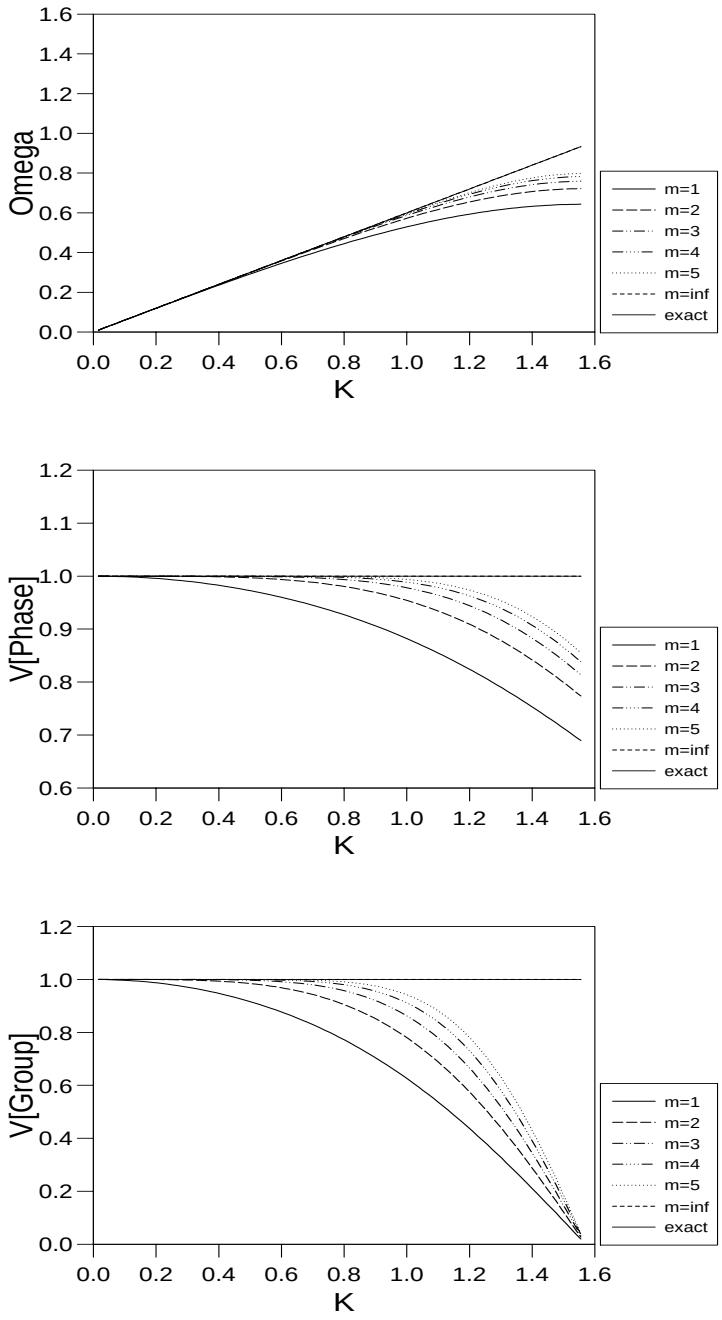

FIG. 3 – *Courbes de dispersion du schéma 2m-2m pour* $\lambda = 0.6$.

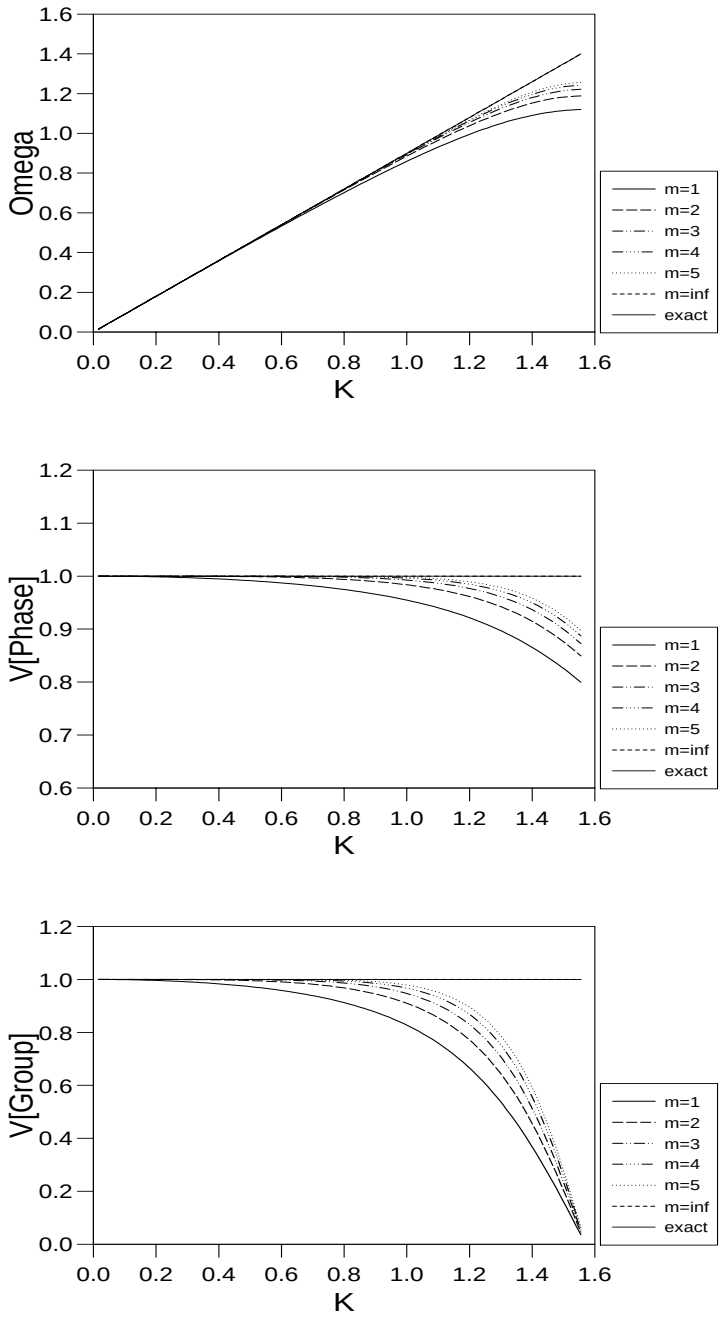

FIG. 4 – *Courbes de dispersion du schéma 2m-2m pour* $\lambda = 0.9$.

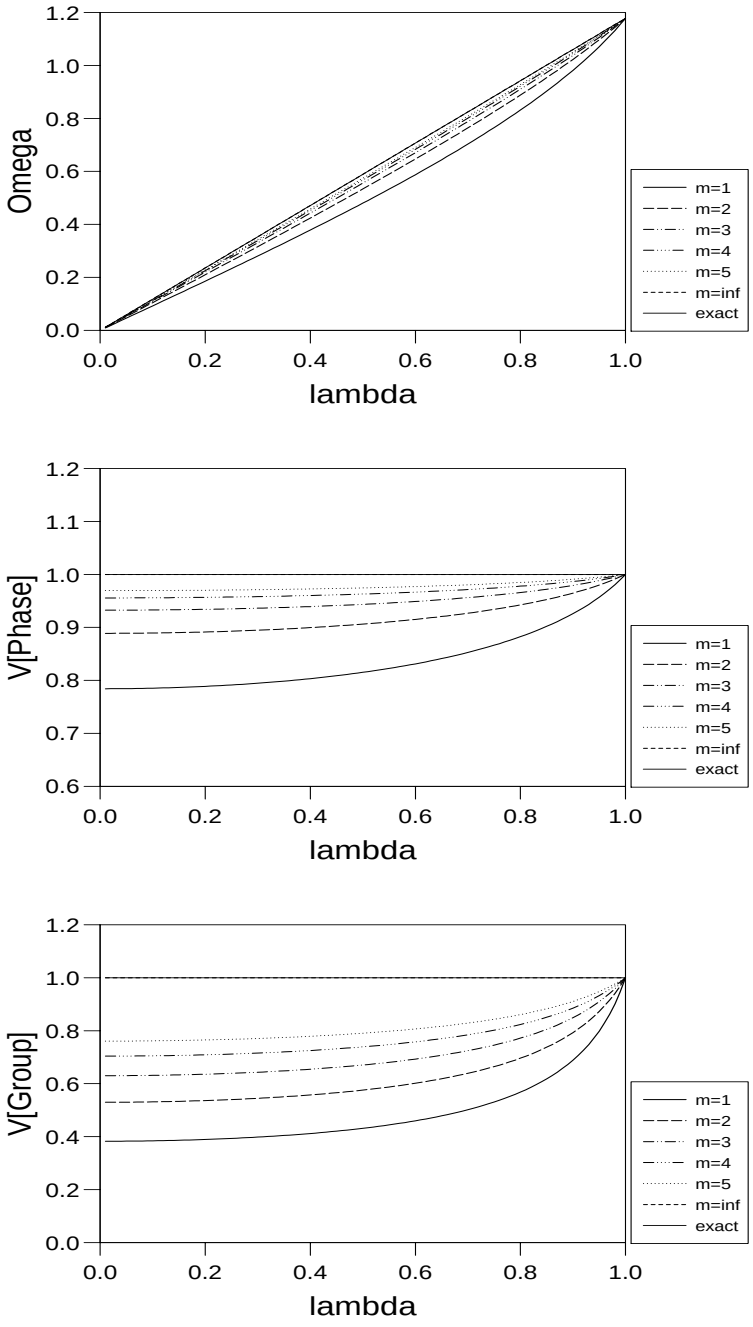

FIG. 5 – *Grandeurs* Ω, V_φ *et* V_g *en fonction de* λ *pour* $K = 3\pi/8$.

[8] G. Cohen et P. Joly, Analysis of fourth order finite difference schemes for the wave equation in heterogeneous media *SIAM J. on Num. Anal.* **33**, no. 4, 1266–1302 (1996).

[9] M. A. Dablain, The Application of High-Order Differencing to the Scalar Wave Equation, *Geophysics* **51**, 54–66 (1986).

[10] P. J. Davis, *Interpolation and Approximation*, Dover Publications Inc., New-York, 1975.

[11] A. Erdélyi, W. Magnus, F. Oberhettinger et F. G. Tricomi, *Higher Transcendental Functions*, Vol. I, McGraw-Hill Book Co., New-York, 1953.

[12] P. Faurre et M. Robin, *Éléments d'Automatique*, Dunod, Paris, 1984.

[13] R. Glowinski et O. Pironneau, Numerical Methods for the First Biharmonic Equation and for the 2-D Stokes Problem, *SIAM Review* **21**, 167–212 (1979).

[14] O. Holberg, Computational Aspects of the Choice of Interval for Numerical Differentiation in Large-Scale Simulation of Wave Phenomena, *Geophys. Prosp.* **35**, 629–655 (1987).

[15] P. Joly, *Une Remarque sur les Schémas aux Différences Finies d'Ordre Arbitraire en Espace pour l'Équation des Ondes*, Communication personnelle, 1995.

[16] K. R. Kelly, R. W. Ward, S. Treitel et R. M. Alford, Synthetic Seismograms: a Finite-Difference Approach, *Geophysics* **41**, 2–27 (1976).

[17] D. Kosloff et E. Baysal, Forward Modeling by a Fourier Method, *Geophysics* **47**, 1402–1412 (1982).

[18] C. Lanczos, *Linear Differential Operators*, D. Van Nostrand Company Ltd., London, 1961.

[19] N. N. Lebedev, *Special Functions and Their Applications*, Prentice-Hall Inc., Englewood Cliffs, 1965.

[20] J. Lelong-Ferrand et J. M. Arnaudiès, *Cours de Mathématique*, Tome 2 : Analyse, Dunod, Paris, 1977.

[21] L. M. Milne-Thomson, *The Calculus of Finite Difference*, MacMillan and Co., London, 1960.

[22] I. Mufti, Large-Scale Three-Dimensional Seismic Models and Their Interpretative Significance, *Geophysics* **55**, 1166–1182 (1990).

[23] A. Sei, *Étude de Schémas Numériques pour les Modèles de Propagation d'Ondes en Milieux Hétérogènes*, Thèse de Doctorat, Université de Paris IX-Dauphine, 1991.

[24] A. Shubin et J.B. Bell, *A modified equation approach to constructing fourth order methods for acoustic wave propagation*, SIAM J. on Sci. Stat. Comp. **8**, (2), 135–151, 1987.

[25] L. Trefethen, Group Velocity in Finite Difference Schemes, *SIAM Review* **24**, 113–136 (1982).

[26] R. Vichnevetsky et J. B. Bowles, *Fourier Analysis of Numerical Approximations of Hyperbolic Equations*, SIAM Studies in Applied Mathematics, SIAM, Philadelphia, 1982.

[27] D. V. Widder, *The Laplace Transform*, Princeton University Press, Princeton, 1946.

A. Laplacien discret

A.1 Première vision

Tableau des α_p^m pour $1 \leq p \leq m \leq 6$. Les lignes sont indexées par m, les colonnes par p. Pour avoir les coefficients a_p^m de la molécule, il suffit de diviser α_p^m par p^2 si $p \geq 1$.

1	1					
2	$\frac{4}{3}$	$-\frac{1}{3}$				
3	$\frac{3}{2}$	$-\frac{3}{5}$	$\frac{1}{10}$			
4	$\frac{8}{5}$	$-\frac{4}{5}$	$\frac{8}{35}$	$-\frac{1}{35}$		
5	$\frac{5}{3}$	$-\frac{20}{21}$	$\frac{5}{14}$	$-\frac{5}{63}$	$\frac{1}{126}$	
6	$\frac{12}{7}$	$-\frac{15}{14}$	$\frac{10}{21}$	$-\frac{1}{7}$	$\frac{2}{77}$	$-\frac{1}{462}$

A.2 Deuxième vision

Liste des β_p pour $1 \leq p \leq 6$. L'indice p repère les colonnes. Pour avoir les coefficients b_p, il suffit de multiplier β_p par 4^{1-p}.

1	*2*	*3*	*4*	*5*	*6*
1	$\frac{1}{3}$	$\frac{8}{45}$	$\frac{4}{35}$	$\frac{128}{1575}$	$\frac{128}{2079}$

A.3 Troisième vision

Tableau des γ_q^m pour $0 \leq q < m \leq 6$. Les lignes sont indexées par m, les colonnes par $q+1$.

1	1					
2	$\frac{7}{6}$	$-\frac{1}{12}$				
3	$\frac{37}{30}$	$-\frac{23}{180}$	$\frac{1}{90}$			
4	$\frac{533}{420}$	$-\frac{779}{5040}$	$\frac{11}{504}$	$-\frac{1}{560}$		
5	$\frac{1627}{1260}$	$-\frac{4343}{25200}$	$\frac{43}{1400}$	$-\frac{109}{25200}$	$\frac{1}{3150}$	
6	$\frac{18107}{13860}$	$-\frac{5697}{30800}$	$\frac{751}{19800}$	$-\frac{1949}{277200}$	$\frac{191}{207900}$	$-\frac{1}{16632}$

B. Schémas numériques

B.1 Schéma 2-2m

Constantes de stabilité Λ_m élevées au carré, pour $1 \leq m \leq 6$. L'indice m est en colonne.

1	*2*	*3*	*4*	*5*	*6*	∞
1	$\frac{3}{4}$	$\frac{45}{68}$	$\frac{315}{512}$	$\frac{75}{128}$	$\frac{51975}{91904}$	$\frac{4}{\pi^2}$

B.2 Schéma 2m-2m

Formule de mise à jour

$$u_j^{n+1} = 2u_j^n - u_j^{n-1} + \lambda^2 \left\{ \widetilde{a}_0^m(\lambda) u_j^n + \sum_{q=1}^{m} \widetilde{a}_q^m(\lambda) \left[u_{j+q}^n + u_{j-q}^n \right] \right\}$$

où

$$\widetilde{a}_q^m(\lambda) = (-1)^{q-1} \sum_{p=q}^{m} \binom{2p}{p-q} b_p \, E_{p-1}(\lambda),$$

avec

$$\left| \begin{array}{rcl} E_{p-1}(\lambda) & = & \displaystyle\prod_{r=1}^{p-1} \left(1 - \frac{\lambda^2}{r^2} \right), \\ b_p & = & \dfrac{2\,(p-1)!^2}{(2p)!}, \end{array} \right.$$

et les conventions

$$b_0 = 0 \qquad \text{et} \qquad E_{-1}(\lambda) = E_0(\lambda) = 1.$$

Molécule 4-4

$$-\tfrac{5-\lambda^2}{2}\, u_j^n + \tfrac{4-\lambda^2}{3} \left[u_{j-1}^n + u_{j+1}^n \right] - \tfrac{1-\lambda^2}{12} \left[u_{j-2}^n + u_{j+2}^n \right].$$

Molécule 6-6

$$\begin{array}{lll} -\frac{83-13\lambda^2+2\lambda^4}{36}\, u_j^n & + & \frac{36-13\lambda^2+\lambda^4}{24} \left[u_{j-1}^n + u_{j+1}^n \right] \\ & - & \frac{9-10\lambda^2+\lambda^4}{60} \left[u_{j-2}^n + u_{j+2}^n \right] \quad + \quad \frac{4-5\lambda^2+\lambda^4}{360} \left[u_{j-3}^n + u_{j+3}^n \right]. \end{array}$$

Unit´e de recherche INRIA Lorraine, Technopôle de Nancy-Brabois, Campus scientifique,
615 rue du Jardin Botanique, BP 101, 54600 VILLERS LÈS NANCY
Unit´e de recherche INRIA Rennes, Irisa, Campus universitaire de Beaulieu, 35042 RENNES Cedex
Unit´e de recherche INRIA Rhône-Alpes, 655, avenue de l'Europe, 38330 MONTBONNOT ST MARTIN
Unit´e de recherche INRIA Rocquencourt, Domaine de Voluceau, Rocquencourt, BP 105, 78153 LE CHESNAY Cedex
Unit´e de recherche INRIA Sophia-Antipolis, 2004 route des Lucioles, BP 93, 06902 SOPHIA-ANTIPOLIS Cedex

Éditeur
INRIA, Domaine de Voluceau, Rocquencourt, BP 105, 78153 LE CHESNAY Cedex (France)
http://www.inria.fr
ISSN 0249-6399