

WONDERBUBBLELAND

Rapport de stage génie mathématique 2020-2021

Averil PROST

sous la direction de

Alexandre IMPÉRIALE

Commissariat à l'énergie atomique
et aux énergies alternatives, Saclay



INSA

INSTITUT NATIONAL
DES SCIENCES
APPLIQUÉES
ROUEN NORMANDIE

cea

Table des matières

Introduction	1
1 Cadre de l'étude	2
1.1 Quelques définitions	2
1.1.1 Matériau	2
1.1.2 Équation d'onde	2
1.2 Problème bien posé	3
1.2.1 Étude statique préliminaire	3
1.2.2 Introduction du temps	5
1.3 Propriétés de la solution	7
1.3.1 Énergie I	7
1.3.2 Expression dans une base spectrale	8
1.3.3 Aspect en dimension 1	9
2 Méthodes de Galerkin	11
2.1 Semi-discrétisation en espace	11
2.1.1 Formulation faible	11
2.1.2 Énergie II	13
2.1.3 Éléments de convergence	13
2.1.3.1 Base spectrale	13
2.1.3.2 Base éléments finis	14
2.2 Semi-discrétisation en temps	20
2.2.1 Schéma	20
2.2.2 Énergie III	20
2.3 Schéma complètement discréteisé	21
2.3.1 Schéma	21
2.3.2 Énergie IV	22
2.3.3 Condition CFL	22
2.3.4 Condensation de masse	25
3 Enrichissement par fonctions bulles	27
3.1 Quelques définitions	27
3.1.1 Espace W_h	27
3.1.2 Représentation dans l'espace enrichi	28
3.1.3 Cas considérés	28

3.2	Schéma IMEX	29
3.2.1	Problème variationnel associé	29
3.2.2	Énergie V	32
3.2.3	Condition CFL	32
3.3	Orthogonalisation	33
3.3.1	Transformation de l'espace bulle	34
3.3.2	Condition CFL	36
3.3.2.1	Matrice de masse	36
3.3.3	Représentation dans l'espace orthogonalisé	37
3.3.4	Un mot sur la quasi-orthogonalisation	39
4	Étude numérique	44
4.1	Obtention d'une solution de référence	44
4.1.1	Première idée : équation non linéaire	44
4.1.2	Deuxième idée : prédiction-correction	46
4.1.3	Troisième idée : résolution numérique	49
4.2	Choix d'implémentation	49
4.2.1	Configurations	49
4.2.2	Maillages	50
4.3	Résultats	51
4.3.1	Comportement de la CFL	51
4.3.2	Influence de la microstructure sur le modèle macroscopique	52
4.3.3	Influence des fonctions bulles	53
4.3.4	Influence de l'orthogonalisation	54
4.3.5	Influence de la condensation de masse	55
4.3.6	Influence du type de bulles	57
4.3.7	Découplage des problèmes	57
Conclusion	59	
Références	60	
Grille de déroulement du stage	61	

Introduction

Ce document présente les travaux réalisés durant le stage de 4ème année d'école d'ingénieur, spécialité Génie Mathématique à l'Institut National des Sciences Appliquées (INSA) Rouen Normandie.

J'ai eu la chance d'être accueillie au Laboratoire de Simulation et Modélisation en Acoustique (LSMA) au CEA, sous la direction d'Alexandre IMPÉRIALE. Ce laboratoire fait partie d'un ensemble axé sur le contrôle non destructif, une famille de méthodes de contrôle s'appuyant largement sur la simulation d'ondes. En particulier, la modélisation de la propagation en milieu non homogène trouve des applications immédiates dans le contrôle des soudures, et représente un enjeu important pour les activités nucléaires. C'est dans ce contexte que se place mon sujet de stage, présenté comme l'étude d'un schéma explicite – implicite pour la discréétisation de l'équation des ondes en régime transitoire en présence d'une micro-structure globalement distribuée.

Nous commençons par l'étude de la solution exacte du problème considéré, en établissant son existence, unicité et quelques propriétés pertinentes. On poursuit par la discréétisation de la solution pour obtenir un schéma numérique, dont on étudie quelques propriétés de stabilité et de convergence.

La troisième partie est dédiée à la méthode multi-échelle proposée comme sujet d'étude. Le schéma implicite-explicite associé est inspiré des récents travaux de Methenni [2021], qui s'intéressent à la propagation d'ondes dans des plaques fines. Nous avons eu pour ambition d'en transposer l'idée directrice à notre problème, pour éventuellement améliorer les méthodes existantes.

La quatrième partie présente les résultats obtenus à l'aide d'un code développé le long du stage.

Ce rapport s'accompagne d'un certain nombre d'illustrations animées, indiquées par les balises vertes.

je suis une balise verte 

On trouvera ces ressources, ainsi que l'intégralité des résultats, sur le dépôt git associé.

Chapitre 1

Cadre de l'étude

1.1 Quelques définitions

1.1.1 Matériaux

Soit un bloc de solide idéal, dont on connaît parfaitement toutes les caractéristiques en chaque point. On se donne deux points A et B distincts de ce solide, et on pose la définition suivante :

Définition 1.1.1 (domaine Ω) On notera $\Omega =]a, b[$ la coupe selon la direction $[A, B]$ du solide considéré. On considère que Ω est non vide et totalement inclus dans le matériau.

Nous connaissons le module de Young E et la masse volumique ρ en chaque point. Nous noterons $\kappa : \Omega \rightarrow \mathbb{R}$ la fonction qui, à chaque point de Ω , associe le rapport E/ρ en ce point.

Définition 1.1.2 (discrétisation) On fait l'hypothèse que la fonction κ est constante par morceaux. Soit $\{x_i\}_{i=0}^N$ la discrétisation de Ω induite par κ , c'est-à-dire telle que $\Omega_i :=]x_i, x_{i+1}[$ soit non vide, la restriction $\kappa|_{\Omega_i}$ soit constante, et $\kappa|_{\Omega_i} \neq \kappa|_{\Omega_{i+1}}$.

On note $\kappa_i := \kappa|_{\Omega_i}$. On fait l'hypothèse (qui suit le sens physique) que κ soit strictement positive et bornée : on notera $\kappa_{\min} > 0$ et $\kappa_{\max} < \infty$ deux constantes telles que $\forall i \in \llbracket 0, N-1 \rrbracket$, $\kappa_{\min} \leq \kappa_i \leq \kappa_{\max}$.

1.1.2 Équation d'onde

Supposons que nous connaissons deux fonctions u_0 et u_1 de Ω dans \mathbb{R} , qui représentent un cliché à l'instant $t = 0$ d'une onde et de sa dérivée dans le matériau. Soit $T \in \mathbb{R}_+^*$ un temps de simulation fini. Nous chercherons à approcher la solution du problème suivant :

Définition 1.1.3 (Équation d'onde)

$$\begin{cases} \partial_t^2 u - \partial_x (\kappa \partial_x u) = 0 & \text{sur } \Omega \times]0, T[\\ u|_{t=0} = u_0, (\partial_t u)|_{t=0} = u_1 & \text{sur } \Omega \\ u|_{\partial\Omega} = 0 & \text{sur } \partial\Omega \times]0, T[\end{cases} \quad (1.1)$$

Nous n'avons pas précisé dans quel espace la solution u se trouve : c'est l'objet de la section suivante. Nous adopterons deux points de vue : selon les propriétés étudiées, u désignera soit une fonction réelle à deux variables, définie de $\Omega \times [0, T] \rightarrow \mathbb{R}$, soit une fonction de la variable temporelle à valeurs dans un espace fonctionnel. Selon cette deuxième définition, $u(t)$ est une fonction définie sur Ω à valeurs dans \mathbb{R} . On se permettra l'utilisation des notations suivantes :

$$u(t) = u(\cdot, t) \quad \text{et} \quad u(t)(x) = u(x, t)$$

Définition 1.1.4 Un **opérateur** est une application linéaire de son domaine $D \subset H$ dans H .

Nous aurons besoin de l'application suivante :

Définition 1.1.5 (opérateur laplacien inhomogène) On définit

$$\mathcal{A}u = -\partial_x(\kappa \partial_x u)$$

Partout où elle est définie, \mathcal{A} est une application linéaire qui correspond à une version inhomogène de l'opérateur laplacien. On désignera par

$$\mathcal{D}(\mathcal{A}) = \left\{ u \in H_0^1(\Omega) \mid \mathcal{A}(u) \in L^2(\Omega) \right\}$$

le sous-espace linéaire de $H_0^1(\Omega)$ pour lequel l'équation (1.1) met en relation des éléments de classe L^2 en espace.

Remarque 1.1.1 Pour κ discontinue, le domaine $\mathcal{D}(\mathcal{A})$ ne contient pas les fonctions familières très régulières : si $\partial_x u(t)$ est continue, $\kappa \partial_x u(t)$ ne l'est plus, et $\mathcal{A}u(t)$ n'est pas dans $L^2(\Omega)$.

1.2 Problème bien posé

1.2.1 Étude statique préliminaire

Cette section a pour but d'établir le résultat suivant :

THÉORÈME 1.2.1 (BIJECTIVITÉ DE $\mathcal{A} + I$) Soit $f \in L^2(\Omega)$. Il existe un unique $u \in \mathcal{D}(\mathcal{A})$ tel que

$$\mathcal{A}u + u = f \tag{1.2}$$

Nous procéderons en trois étapes : on se ramène d'abord à une formulation faible, on établit l'existence et l'unicité de la solution de ce problème, puis on revient au problème initial.

Supposons que u satisfasse (1.2) : dès lors, pour tout élément $v \in H_0^1(\Omega)$,

$$(\mathcal{A}u + u, v)_{L^2(\Omega)} = (f, v)_{L^2(\Omega)}$$

L'appartenance de u à $\mathcal{D}(\mathcal{A})$ nous permet d'appliquer le théorème de Green au terme $(\mathcal{A}u, v)_{L^2(\Omega)}$. Le choix de $v \in H_0^1(\Omega)$ permet d'annuler le terme intégral sur $\partial\Omega$, ce qui nous donne le problème

faible satisfait par u :

$$(\kappa \partial_x u, \partial_x v)_{L^2(\Omega)} + (u, v)_{L^2(\Omega)} = (f, v)_{L^2(\Omega)} \quad (1.3)$$

Définition 1.2.1 (forme bilinéaire a) Soient u et v dans $H_0^1(\Omega)$. On notera

$$a(u, v) = (\kappa \partial_x u, \partial_x v)_{L^2(\Omega)}$$

Lemme 1.2.1 La forme bilinéaire $a(\cdot, \cdot)$ définit un produit scalaire sur $H_0^1(\Omega)$ équivalent au semi-produit scalaire naturel de $H^1(\Omega)$.

Démonstration

La forme bilinéaire $a(\cdot, \cdot)$ est symétrique (ce qui se voit en décomposant le produit scalaire sur chaque domaine Ω_i où κ est constante) : l'équivalence vient de l'hypothèse $0 < \kappa_{\min} \leq \kappa(x) \leq \kappa_{\max} < +\infty$ pour tout $x \in \Omega$. On a directement

$$\kappa_{\min} (\partial_x u, \partial_x u)_{L^2(\Omega)} \leq a(u, u) \leq \kappa_{\max} (\partial_x u, \partial_x u)_{L^2(\Omega)}$$

et sur $H_0^1(\Omega)$, grâce à l'inégalité de Poincaré, $\|\partial_x u\|_{L^2(\Omega)} = 0 \implies u \equiv 0$. D'où $\sqrt{a(u, u)}$ est une norme sur $H_0^1(\Omega)$. □

Employons le théorème suivant, dont on trouvera une démonstration dans Allaire [2012] :

THÉORÈME 1.2.2 (DE LAX-MILGRAM) Soit Ω un domaine connexe, borné, à frontière régulière. Soit $(V, (\cdot, \cdot)_V)$ un espace de Hilbert réel. Soient $\mu(\cdot, \cdot)$ une forme bilinéaire continue coercive, et $L(\cdot)$ une forme linéaire continue sur V . Alors le problème

$$\text{Trouver } u \in V \text{ telle que } \mu(u, v) = L(v) \quad \forall v \in V$$

admet une unique solution dans V , qui dépend continuement de la variable L .

En vertu du lemme 1.2.1, la forme $\mu(\cdot, \cdot) := (\cdot, \cdot)_{L^2(\Omega)} + a(\cdot, \cdot)$ définit un produit scalaire équivalent à la norme naturelle de $V := H_0^1(\Omega)$. Pour $L(\cdot) := (f, \cdot)_{L^2(\Omega)}$, les hypothèses du théorème sont bien vérifiées. Nous sommes assurés de l'existence d'une unique solution $u \in H_0^1(\Omega)$. De plus, la continuité de u par rapport à L implique celle de u par rapport à f : donc \mathcal{A} est continu. Il reste à établir l'assertion suivante :

Lemme 1.2.2 La solution $u \in H_0^1(\Omega)$ du problème (1.3) est dans $\mathcal{D}(\mathcal{A})$.

Démonstration

Exploitons l'inclusion de $\mathcal{D}(\Omega) := \mathcal{C}_c^\infty(\Omega)$, l'espace des fonctions test sur Ω , dans $H_0^1(\Omega)$. Suivant l'abus de langage usuel, notons encore u la distribution définie par $\langle u, v \rangle_{\mathcal{D}'(\Omega), \mathcal{D}(\Omega)} := (u, v)_{L^2(\Omega)}$ pour

$v \in \mathcal{D}(\Omega)$. Remarquons que $\kappa \partial_x u$ est correctement définie, κ étant bornée. On a alors

$$\begin{aligned} & \langle \kappa \partial_x u, \partial_x v \rangle_{\mathcal{D}'(\Omega), \mathcal{D}(\Omega)} + \langle u, v \rangle_{\mathcal{D}'(\Omega), \mathcal{D}(\Omega)} = \langle f, v \rangle_{\mathcal{D}'(\Omega), \mathcal{D}(\Omega)} \quad \forall v \in \mathcal{D}(\Omega) \\ \Leftrightarrow & \langle -\partial_x (\kappa \partial_x u) + u - f, v \rangle_{\mathcal{D}'(\Omega), \mathcal{D}(\Omega)} = 0 \quad \forall v \in \mathcal{D}(\Omega) \end{aligned}$$

donc, au sens des distributions, $-\partial_x (\kappa \partial_x u) = f - u$. Comme $f - u \in L^2(\Omega)$, on en déduit que la distribution notée $-\partial_x (\kappa \partial_x u)$ est définie par la fonction $-\partial_x (\kappa \partial_x u) \in L^2(\Omega)$: ainsi, $\mathcal{A}u \in L^2(\Omega)$, et notre unique solution appartient bien à $\mathcal{D}(\mathcal{A})$.

□

On a donc bien démontré le théorème (1.2.1).

1.2.2 Introduction du temps

Nous allons établir le résultat suivant :

THÉORÈME 1.2.3 (EXISTENCE ET UNICITÉ DE LA SOLUTION FORTE) Si $u_0 \in \mathcal{D}(\mathcal{A})$ et $u_1 \in H_0^1(\Omega)$, alors il existe une unique solution

$$u \in C^2([0, T], L^2(\Omega)) \cap C^1([0, T], H_0^1(\Omega)) \cap C^0([0, T], \mathcal{D}(\mathcal{A}))$$

Pour étudier le problème dynamique, nous allons premièrement le formuler comme un problème d'évolution d'ordre 1 en temps.

Définition 1.2.2 (espace de Hilbert du problème d'ordre 1) Notons $(H, (\cdot, \cdot)_H)$ l'espace défini par $H = H_0^1(\Omega) \times L^2(\Omega)$ et

$$(U, V)_H = a(u_1, v_1) + (u_2, v_2)_{L^2(\Omega)}, \quad \text{pour } U = (u_1, u_2)^t, U \text{ et } V \text{ dans } H$$

Remarque 1.2.1 La forme bilinéaire $a(\cdot, \cdot)$ est un produit scalaire sur $H_0^1(\Omega)$ en vertu du lemme (1.2.1). H est donc un produit d'espaces de Hilbert, donc un espace de Hilbert.

Définition 1.2.3 (opérateur A) On définit $A : D(A) \rightarrow H$ par

$$AU = \begin{pmatrix} -u_2 \\ -\partial_x (\kappa \partial_x u_1) \end{pmatrix} = \begin{pmatrix} 0 & -I \\ \mathcal{A} & 0 \end{pmatrix} U, \quad U = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}$$

Remarque 1.2.2 Le domaine de l'opérateur A est celui pour lequel $-u_2 \in H_0^1(\Omega)$ et $\mathcal{A}u_1 \in L^2(\Omega)$, donc $D(A) = \mathcal{D}(\mathcal{A}) \times H_0^1(\Omega)$.

Proposition 1.2.1 (problème d'ordre 1) Soit $U : [0, T] \rightarrow H$ la fonction donnée par $U(t) = (u(\cdot, t), \partial_t u(\cdot, t))^t$. U est solution du problème

$$\begin{cases} \frac{dU}{dt} + AU = 0 & \text{sur }]0, T[\\ U|_{t=0} = (u_0, u_1)^t \end{cases} \quad (1.4)$$

Démonstration

Le système linéaire se développe en

$$\begin{pmatrix} \partial_t u(\cdot, t) \\ \partial_t^2 u(\cdot, t) \end{pmatrix} + \begin{pmatrix} -\partial_t u(\cdot, t) \\ -\partial_x (\kappa \partial_x u)(\cdot, t) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

qui n'impose à u que la satisfaction de l'équation d'onde considérée. Les conditions initiales se déduisent directement de (1.1). □

Pour énoncer le théorème sur lequel nous nous appuierons, on introduit la définition suivante :

Définition 1.2.4 (opérateur maximal monotone) Soit H un espace de Hilbert. Un opérateur $A : D(A) \rightarrow H$ est dit maximal monotone si

- $A + I$ est surjectif de $D(A)$ dans H ,
- pour tout élément U de $D(A)$, $(AU, U)_H \geqslant 0$.

THÉORÈME 1.2.4 (DE HILLE-YOSIDA [2013]) S'il existe $\lambda \in \mathbb{R}$ tel que $A + \lambda I$ soit maximal monotone, alors pour $U_0 \in D(A)$, le problème (1.4) admet une unique solution dans

$$\mathcal{C}^1([0, T], H) \cap \mathcal{C}^0([0, T], D(A))$$

Proposition 1.2.2 L'opérateur A défini en (1.2.3) est maximal monotone.

Démonstration

Commençons par vérifier la surjectivité. Soit $V = (v_1, v_2)^t \in H$: on cherche $U \in D(A)$ tel que $(A + I)U = V$. Par substitution, on se ramène à la recherche de $u_1 \in \mathcal{D}(A)$ tel que

$$\mathcal{A}u_1 + u_1 = v_2 + v_1 = f \in L^2(\Omega) \quad \text{et} \quad u_2 = u_1 - v_1$$

La bijectivité de $\mathcal{A} + I$ établie en (1.2.1) nous assure l'existence et l'unicité de u_1 : notre opérateur A est donc bien maximal.

Montrons que A est monotone : soit $U = (u_1, u_2)^t \in D(A)$, on a

$$(AU, U)_H = a(-u_2, u_1) + (-(\partial_x(\kappa \partial_x u_1)), u_2)_{L^2(\Omega)} = -a(u_2, u_1) + a(u_1, u_2) = 0$$

d'où notre opérateur est orthogonal, et en particulier, monotone. □

Démontrons alors le théorème (1.2.3). Si $U|_{t=0} \in D(A)$, les propriétés de l'opérateur A permettent d'appliquer le théorème de Hille-Yosida au problème (1.4), nous donnant une unique solu-

tion

$$U = \begin{pmatrix} u \\ \partial_t u \end{pmatrix} \in \mathcal{C}^1([0, T], H) \cap \mathcal{C}^0([0, T], D(A))$$

On en déduit que la fonction u , solution du problème (1.1), existe et est unique dans l'espace

$$u \in \mathcal{C}^2([0, T], L^2(\Omega)) \cap \mathcal{C}^1([0, T], H_0^1(\Omega)) \cap \mathcal{C}^0([0, T], \mathcal{D}(\mathcal{A}))$$

1.3 Propriétés de la solution

1.3.1 Énergie I

Cette section est la première d'une série qui prendra toute son importance dans l'étude de la stabilité numérique. Pour l'instant, on se contente des éléments suivants :

Définition 1.3.1 On définit trois énergies de $]0, T[\rightarrow \mathbb{R}$, paramétrées par u , par

$\mathcal{K}(t) = \frac{1}{2} (\partial_t u(\cdot, t), \partial_t u(\cdot, t))_{L^2(\Omega)}$	l'énergie cinétique
$\mathcal{P}(t) = \frac{1}{2} a(u(\cdot, t), u(\cdot, t))$	l'énergie potentielle
$\mathcal{E}(t) = \mathcal{K}(t) + \mathcal{P}(t)$	l'énergie totale

Proposition 1.3.1 Si u satisfait l'équation d'onde (1.1), l'énergie totale se conserve :

$$\frac{d\mathcal{E}}{dt} = \frac{d\mathcal{K}}{dt} + \frac{d\mathcal{P}}{dt} = 0$$

Démonstration

Le procédé suivant sera le fil directeur de toutes les démonstrations de conservation d'énergie. On le développe ici pour le cas continu : le cas discret se transpose facilement. Les prochaines démonstrations se contenteront d'indiquer les artifices de calcul employés. Si u vérifie l'équation d'onde, on a en particulier

$$(\partial_t^2 u(\cdot, t) + \mathcal{A}u(\cdot, t), \partial_t u(\cdot, t))_{L^2(\Omega)} = 0 \quad \forall t \in]0, T[$$

On utilise premièrement l'identité $\frac{d(f^2)}{dt}(t) = 2f'(t)f(t)$ pour écrire

$$(\partial_t^2 u(\cdot, t), \partial_t u(\cdot, t))_{L^2(\Omega)} = \frac{1}{2} \frac{d}{dt} (\partial_t u(\cdot, t), \partial_t u(\cdot, t))_{L^2(\Omega)}$$

et deuxièmement, avec $\partial_t u(\cdot, t) \in H_0^1(\Omega)$ et le même argument,

$$(\mathcal{A}u(\cdot, t), \partial_t u(\cdot, t))_{L^2(\Omega)} = a(\partial_x u(\cdot, t), \partial_{tx}^2 u(\cdot, t)) = \frac{1}{2} \frac{d}{dt} a(\partial_x u(\cdot, t), \partial_x u(\cdot, t))$$

On reconnaît alors

$$\frac{d\mathcal{K}}{dt}(t) + \frac{d\mathcal{P}}{dt}(t) = \frac{d\mathcal{E}}{dt}(t) = 0$$

□

1.3.2 Expression dans une base spectrale

Définition 1.3.2 (opérateur autoadjoint compact) Soit $(H, (\cdot, \cdot)_H)$ un espace de Hilbert. Un opérateur A défini sur $D(A) \subset H$ est dit

- autoadjoint si pour tout $u, v \in H^2$, $(Au, v)_H = (u, Av)_H$,
- compact si l'image par A de toute partie bornée de $D(A)$ est relativement compacte dans H .

Commençons par nous munir d'une base hilbertienne de $H_0^1(\Omega)$ bien choisie. Pour ceci, on s'appuiera sur le résultat suivant, tiré de Brezis [1983] :

THÉORÈME 1.3.1 (DIAGONALISATION) Soit H un espace de Hilbert séparable, et T un opérateur autoadjoint compact. Alors H admet une base hilbertienne formée de vecteurs propres de T .

Proposition 1.3.2 $L^2(\Omega)$ admet une base hilbertienne formée de fonctions propres de \mathcal{A} .

Démonstration

On sait par les injections de Sobolev que $H_0^1(\Omega)$ s'injecte compactement dans $L^2(\Omega)$, et que $H^1(\Omega)$ est dense dans $L^2(\Omega)$ au sens de la topologie induite par $\|\cdot\|_{L^2(\Omega)}$. De plus, on a établi la continuité de \mathcal{A} comme conséquence du théorème de Lax-Milgram, donc $\mathcal{A} : \mathcal{D}(\mathcal{A}) \subset H_0^1(\Omega) \rightarrow L^2(\Omega)$ est compact. D'autre part, l'annulation au bord des éléments de $H_0^1(\Omega)$ le rend autoadjoint : pour $u, v \in \mathcal{D}(\mathcal{A})$, on a

$$(\mathcal{A}u, v)_{L^2(\Omega)} = a(\partial_x u, \partial_x v) = (u, \mathcal{A}v)_{L^2(\Omega)}$$

□

On notera $(\lambda_n, \omega_n)_{n \in \mathbb{N}^*}$ les éléments propres de \mathcal{A} , qui satisfont $\mathcal{A}\omega_n = \lambda_n \omega_n$.

Remarque 1.3.1 Par la même démarche que dans la section (1.2.1), on montre que $(\lambda_n, \omega_n)_{n \in \mathbb{N}^*}$ est l'ensemble des solutions du problème variationnel spectral « Trouver $\lambda \in \mathbb{R}$ tel qu'il existe $\omega \in V$ non nul satisfaisant

$$a(\omega, v) = \lambda(\omega, v)_{L^2(\Omega)} \quad \forall v \in H_0^1(\Omega) \gg \quad (1.5)$$

En particulier, on a $\lambda_n = \frac{a(\omega_n, \omega_n)}{\|\omega_n\|_{L^2(\Omega)}} > 0$. La suite $(\lambda_n, \omega_n)_{n \in \mathbb{N}^*}$ est ordonnée dans le sens des λ_n croissants.

Proposition 1.3.3 (Décomposition de u) Si $u_0 \in \mathcal{D}(\mathcal{A})$ et $u_1 \in H_0^1(\Omega)$, la solution u admet une expression analytique donnée par

$$u(x, t) = \sum_{n=1}^{\infty} \left((u_0, \omega_n) \cos(\sqrt{\lambda_n} t) + \frac{(u_1, \omega_n)}{\sqrt{\lambda_n}} \sin(\sqrt{\lambda_n} t) \right) \omega_n(x) \quad (1.6)$$

Démonstration

Soit $t \in]0, T[$. On note $u(\cdot, t) = \sum_{n=1}^{\infty} \alpha_n(t) \omega_n$ la décomposition de u dans la base $\{\omega_n\}_{n \in \mathbb{N}^*}$. Comme $u \in H^1(\Omega)$, cette série converge en norme $H^1(\Omega)$: en particulier, on peut écrire

$$\begin{aligned} \frac{\partial^2}{\partial t^2} \left(\sum_{n=1}^{\infty} \alpha_n(t) \omega_n \right) + \mathcal{A} \left(\sum_{n=1}^{\infty} \alpha_n(t) \omega_n \right) &= \sum_{n=1}^{\infty} \frac{\partial^2 \alpha_n}{\partial t^2}(t) \omega_n + \sum_{n=1}^{\infty} \alpha_n(t) \mathcal{A}(\omega_n) \\ &= \sum_{n=1}^{\infty} \left(\frac{d^2 \alpha_n}{dt^2}(t) + \lambda_n \alpha_n(t) \right) \omega_n = 0 \end{aligned}$$

Comme $u_0 \in \mathcal{D}(\mathcal{A})$ et $u_1 \in H_0^1(\Omega)$, les décompositions de ces deux fonctions dans la base $(\omega_n)_n$ convergent également en norme H^1 et l'on peut écrire

$$\sum_{n=1}^{\infty} \alpha_n(0) \omega_n = u_0 = \sum_{n=1}^{\infty} (u_0, \omega_n)_{L^2(\Omega)} \omega_n, \quad \sum_{n=1}^{\infty} (\alpha_n)'(0) \omega_n = u_1 = \sum_{n=1}^{\infty} (u_1, \omega_n)_{L^2(\Omega)} \omega_n$$

La base $(\omega_n)_{n \in \mathbb{N}^*}$ étant orthonormée, les coefficients de u vérifient les problèmes de Cauchy

$$\begin{cases} \frac{d^2 \alpha_n}{dt^2} + \lambda_n \alpha_n = 0 \\ \alpha_n(0) = (u_0, \omega_n)_{L^2(\Omega)}, \quad (\alpha_n)'(0) = (u_1, \omega_n)_{L^2(\Omega)} \end{cases} \quad \text{dans }]0, T[$$

Sachant que $\lambda_n > 0$ pour tout n , les solutions sont données par

$$\alpha_n(t) = (u_0, \omega_n)_{L^2(\Omega)} \cos(\sqrt{\lambda_n} t) + \frac{(u_1, \omega_n)_{L^2(\Omega)}}{\sqrt{\lambda_n}} \sin(\sqrt{\lambda_n} t)$$

ce qui nous donne l'expression analytique de u dans la base des fonctions propres.

□

Remarque 1.3.2 Une première méthode numérique pour approcher u pourrait consister à obtenir des approximations des (λ_n, ω_n) jusqu'à un $n \ll$ suffisamment grand », puis à prendre pour estimation de u la série tronquée.

1.3.3 Aspect en dimension 1

Éléments de $\mathcal{D}(\mathcal{A})$

Les fonctions de $\mathcal{D}(\mathcal{A})$ satisfont les propriétés suivantes :

- elles sont nulles aux bords de Ω (par définition).
- elles sont continues : l'espace $H^1(\Omega)$ s'injecte dans $C^0(\Omega)$. Remarquons que cette propriété n'est valable qu'en dimension 1.
- aux interfaces des Ω_i , elles changent de direction selon une loi assimilable à Snell-Descartes : pour $i \in \llbracket 1, N-1 \rrbracket$, on a

$$\kappa_{i-1} \partial_x u(x_i) = \kappa_i \partial_x u(x_i) \quad \forall u \in \mathcal{D}(\mathcal{A})$$

On s'en convainc en remarquant que $\partial_x(\kappa \partial_x u) \in L^2(\Omega)$, donc $\kappa \partial_x u \in H_0^1(\Omega) \subset C^0(\Omega)$.

Cette troisième condition traduit un saut de la dérivée aux interfaces des domaines $[x_i, x_{i+1}]$. Le lecteur pourra se familiariser avec ces propriétés à l'aide de la figure (1.1).

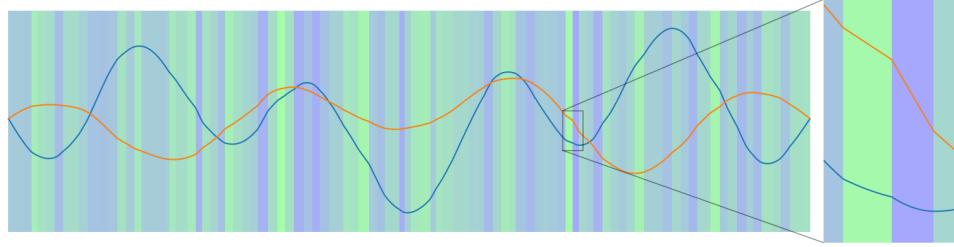


FIGURE 1.1 – éléments de $\mathcal{D}(\mathcal{A})$ à κ non constant

Deux éléments d'un domaine inhomogène sont représentés. Les couleurs du fond indiquent la valeur de la fonction κ . Aux points de discontinuité de κ , le saut entre les dérivées est d'autant plus marqué que la différence de κ est importante.

Fonctions propres de \mathcal{A}

On peut aller plus loin dans l'étude des fonctions propres. Pour ceci, remarquons que le problème variationnel spectral (1.5) est satisfait par la restriction de ω_n à tout ouvert $O \subset \Omega$: l'indicatrice de O est dans $L^2(\Omega)$, donc elle peut être approchée par une suite d'éléments de $\mathcal{D}(\Omega)$. En particulier, pour $O = \Omega_i$, on a

$$\kappa_i \left(\partial_x \omega_n|_{\Omega_i}, \partial_x v \right)_{L^2(\Omega_i)} = \left(\omega_n|_{\Omega_i}, v \right)_{L^2(\Omega_i)} \quad \forall v \in H_0^1(\Omega_i) \implies -\kappa_i \partial_x^2 \omega_n|_{\Omega_i} = \lambda_n \omega_n|_{\Omega_i}$$

et la restriction $\omega_n|_{\Omega_i}$ est une classique sinusoïde de la forme $A \sin \left(\sqrt{\frac{\lambda_n}{\kappa_i}} x + C \right)$. La figure (1.2) permet de se représenter ces fonctions.

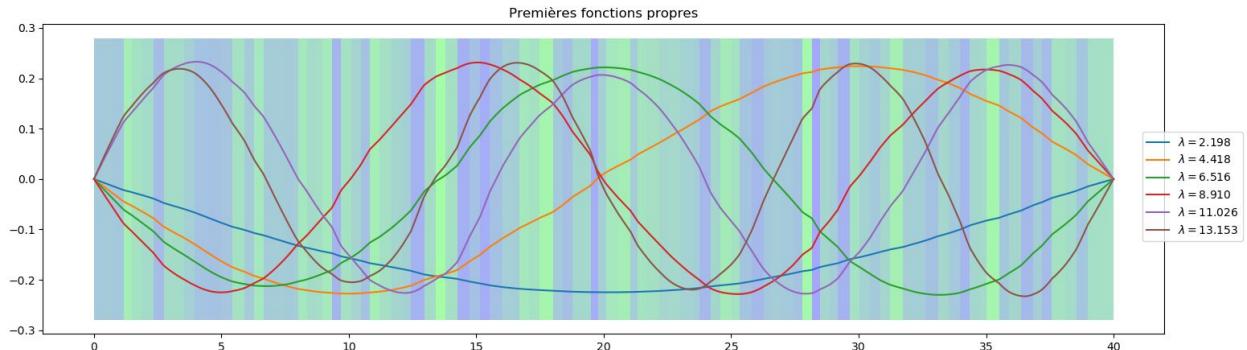


FIGURE 1.2 – premières fonctions propres

Chapitre 2

Méthodes de Galerkin

On appelle « méthodes de Galerkin » la famille des méthodes d'approximation interne, c'est-à-dire l'ensemble des méthodes qui cherchent une solution approchée u_h dans un espace V_h inclus dans V . Pour notre problème, le choix d'une méthode de Galerkin imposera notamment que nos solutions approchées soient continues.

Dans la suite, la notation $m(\cdot, \cdot)$ désignera le produit scalaire $(\cdot, \cdot)_{L^2(\Omega)}$.

2.1 Semi-discrétisation en espace

2.1.1 Formulation faible

La méthode des éléments finis repose sur une formulation faible du problème considéré. En utilisant les notations précédentes, nous chercherons donc $u :]0, T[\rightarrow H_0^1(\Omega)$ vérifiant

$$\begin{cases} \frac{d^2}{dt^2}(u(t), v)_{L^2(\Omega)} + a(u(t), v) = 0 & \forall v \in H_0^1(\Omega), \forall t \in]0, T[\\ u(0) = u_0, \quad \partial_t u(0) = u_1 \end{cases} \quad (2.1)$$

On admettra sans démonstration le résultat suivant :

THÉORÈME 2.1.1 (EXISTENCE ET UNICITÉ DE LA SOLUTION FAIBLE) Si $u_0 \in H_0^1(\Omega)$ et $u_1 \in L^2(\Omega)$, alors le problème faible admet une unique solution dans l'espace

$$u \in C^1([0, T[, L^2(\Omega)) \cap C^0([0, T[, H_0^1(\Omega))$$

On pourra consulter les travaux de Lions and Magenes [1968] et Joly [2001] à ce sujet.

Remarque 2.1.1 La régularité en temps de la solution faible ne permet de vérifier le problème faible qu'au sens des distributions (en temps).

Base de Galerkin

On se donne un espace $V_h \subset V$ de dimension N_h finie, et une base $\Phi = \{\phi_i\}_{i=1}^{N_h}$ de cet espace. Soit un projecteur $\Pi_h : V \rightarrow V_h$.

Définition 2.1.1 (schéma semi-discrétisé en espace) On cherche $u_h \in \mathcal{C}^0 ([0, T[, V_h)$ solution du problème variationnel semi-discrétisé

$$\begin{cases} \partial_t^2(u_h(\cdot, t), v_h)_{L^2(\Omega)} + a(u_h(\cdot, t), v_h) = 0 & \forall v_h \in V_h, \forall t \in]0, T[\\ u(\cdot, 0) = \Pi_h u_0 = u_{0,h} \\ \partial_t u(\cdot, 0) = \Pi_h u_1 = u_{1,h} \end{cases} \quad (2.2)$$

Exprimons ce problème en termes matriciels. On décompose $u_h(\cdot, t)$ en $\sum_{i=1}^{N_h} U_{h,i}(t) \phi_i$, où $U_h(t) \in \mathbb{R}^{N_h}$, et on évalue la formulation faible pour $v_h = \phi_i$, i prenant successivement toutes les valeurs de $[1, N_h]$: on en déduit un système linéaire de la forme

$$\mathbb{M}U''_h(t) + \mathbb{K}U_h(t) = 0, \quad \mathbb{M}_{ij} = (\phi_j, \phi_i)_{L^2(\Omega)} \quad \mathbb{K}_{ij} = a(\phi_j, \phi_i) \quad (2.3)$$

On désignera par la suite \mathbb{M} par **matrice de masse** et \mathbb{K} par **matrice de rigidité**.

THÉORÈME 2.1.2 Pour $V = H_0^1(\Omega)$, le problème (2.2) admet une unique solution.

Démonstration

On sait que sur $H_0^1(\Omega)$, la forme bilinéaire $a(\cdot, \cdot)$ est une norme. En conséquence, les deux matrices \mathbb{M} et \mathbb{K} sont symétriques définies positives : en effet, la symétrie de \mathbb{K} est directement héritée de celle de $a(\cdot, \cdot)$, et pour tout $V \in \mathbb{R}^{N_h}$,

$$V^t \mathbb{K} V = a(v, v) \geq 0, \quad v = \sum_{i=1}^{N_h} V_i \phi_i$$

d'où $a(v, v) = 0 \implies v = 0$. Le raisonnement est identique pour \mathbb{M} avec $(\cdot, \cdot)_{L^2(\Omega)}$.

On forme le vecteur $W(t) = (U(t), U'(t))^t$ qui est solution du problème

$$W'(t) + \mathbb{A}W(t) = 0, \quad \mathbb{A} = \begin{pmatrix} 0 & -\mathbb{I} \\ \mathbb{M}^{-1}\mathbb{K} & 0 \end{pmatrix}$$

Par hypothèse, $\mathbb{M}^{-1}\mathbb{K}$ est inversible, donc \mathbb{A} l'est également, et ce problème admet une unique solution donnée par

$$W(t) = e^{-\mathbb{A}t} \begin{pmatrix} U_0 \\ U_1 \end{pmatrix} = \begin{pmatrix} U(t) \\ U'(t) \end{pmatrix}$$

□

Remarque 2.1.2 Dans le cas général, on ne peut pas affirmer que $u_h = \Pi_h u$. Pour s'en convaincre, considérons l'exemple suivant (illustré figure (2.1)) : soit Ω un matériau homogène, u_0 nulle et u_1 une impulsion à support compact dans Ω . On prend $V_h = \{v \in V \mid v|_{a,c[} \in \mathbb{P}_1 \text{ et } v|_{c,b[} \in \mathbb{P}_1\}$, où c est un point de Ω hors du support de u_1 . On prend pour projecteur Π_h l'interpolateur en $\{a, c, b\}$: on a $\Pi_h u_0 = \Pi_h u_1 \equiv 0$. La solution calculée u_h sera alors identiquement nulle, alors que l'onde se propageant atteindra une valeur non nulle en c .

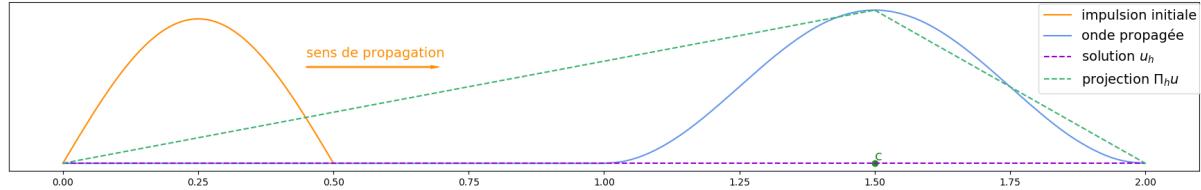


FIGURE 2.1 – exemple où $u_h \neq \Pi_h u$

2.1.2 Énergie II

Définition 2.1.2 Les énergies semi-discrètes sont données par

$$\begin{aligned}\mathcal{K}_h(t) &= \frac{1}{2} (U'(t))^t \mathbb{M} U'(t) && \text{l'énergie cinétique} \\ \mathcal{P}_h(t) &= \frac{1}{2} (U(t))^t \mathbb{K} U(t) && \text{l'énergie potentielle} \\ \mathcal{E}_h(t) &= \mathcal{K}_h(t) + \mathcal{P}_h(t) && \text{l'énergie totale}\end{aligned}$$

Proposition 2.1.1 L'énergie totale se conserve.

Démonstration

Les arguments sont similaires à ceux de la preuve 1.3.1. □

2.1.3 Éléments de convergence

Cette section a pour but d'établir des résultats *a priori* sur les cas particuliers que nous étudierons par la suite : une base spectrale de dimension finie (donc tronquée), et une base éléments finis.

2.1.3.1 Base spectrale

Définition 2.1.3 (méthode de Galerkin spectrale) On choisit la méthode définie par

- une base $\Phi = \{\varphi_i\}_{i=1}^{N_h}$ formée des N_h premiers éléments propres de l'opérateur \mathcal{A} ,
- l'espace d'approximation $V_h = Vect(\Phi) \subset V$,
- le projecteur orthogonal sur Φ :

$$\Pi_h v = \sum_{i=1}^{N_h} (\varphi_i, v)_{L^2(\Omega)} \varphi_i$$

Proposition 2.1.2 Pour la base spectrale tronquée, on a $u_h = \Pi_h u$.

Démonstration

D'une part, u_h est définie comme la solution dans V_h au problème variationnel (2.2). Par le même raisonnement que dans le cas analytique, on peut décomposer $u_h = \sum_{n=1}^{N_h} \alpha_{h,n}(t) \omega_n$ et déduire que les coefficients $\alpha_{h,n}$ satisfont les mêmes problèmes d'évolution que les coefficients α_n .

D'autre part, une base hilbertienne étant par définition orthonormée, la projection de u sur V_h laisse invariants les coefficients associés aux fonctions $\varphi_n = \omega_n$ pour $n \leq N_h$. On a donc coïncidence entre u_h et $\Pi_h u$. □

Proposition 2.1.3 Pour $\|\cdot\|$ désignant indifféremment $\|\cdot\|_{L^2(\Omega)}$ ou $\|\cdot\|_{H^1(\Omega)}$, il existe une constante $C > 0$ telle que

$$\sup_{t \in [0, T]} \|u(\cdot, t) - u_h(\cdot, t)\| \leq C \|u_1 - \Pi_h u_1\|$$

Démonstration

Soit $t \in [0, T]$. Le cas $\|\cdot\| = \|\cdot\|_{L^2(\Omega)}$ s'obtient par majoration directe :

$$\begin{aligned} \|u(\cdot, t) - u_h(\cdot, t)\|_{L^2(\Omega)}^2 &= \left\| \sum_{n=N_h+1}^{\infty} \frac{(u_1, \omega_n)_{L^2(\Omega)}}{\sqrt{\lambda_n}} \sin(\sqrt{\lambda_n} t) \omega_n \right\|_{L^2(\Omega)}^2 \\ &\leq \frac{1}{\lambda_{N_h+1}} \sum_{n=N_h+1}^{\infty} |(u_1, \omega_n)_{L^2(\Omega)}|^2 = \frac{1}{\lambda_{N_h+1}} \|u_1 - \Pi_h u_1\|_{L^2(\Omega)}^2 \end{aligned}$$

Notons $\|\cdot\|_a^2 = a(\cdot, \cdot)$. Pour tout $n \in \mathbb{N}^*$, on a $a\left(\frac{\omega_n}{\sqrt{\lambda_n}}, \frac{\omega_n}{\sqrt{\lambda_n}}\right) = \lambda_n \left(\frac{\omega_n}{\sqrt{\lambda_n}}, \frac{\omega_n}{\sqrt{\lambda_n}}\right)_{L^2(\Omega)} = 1$, ce qui fait de $\left(\lambda_n^{-1/2} \omega_n\right)_{n \in \mathbb{N}^*}$ une base orthonormée de $(H^1(\Omega), a(\cdot, \cdot))$. On a alors

$$\begin{aligned} \|u(\cdot, t) - u_h(\cdot, t)\|_a^2 &= \left\| \sum_{n=N_h+1}^{\infty} (u_1, \omega_n)_{L^2(\Omega)} \sin(\sqrt{\lambda_n} t) \frac{\omega_n}{\sqrt{\lambda_n}} \right\|_{L^2(\Omega)}^2 \\ &\leq \sum_{n=N_h+1}^{\infty} |(u_1, \omega_n)_{L^2(\Omega)}|^2 = \|u_1 - \Pi_h u_1\|_{L^2(\Omega)}^2 \end{aligned}$$

On utilise l'équivalence entre $\|\cdot\|_a^2$ et la semi-norme naturelle de $H_0^1(\Omega)$, établie au lemme (1.2.1), pour obtenir le même type de majoration en norme $H^1(\Omega)$. Dans les deux cas, le passage au sup sur $t \in [0, T]$ donne le résultat attendu. □

2.1.3.2 Base éléments finis

Étude statique

Cette section introduit des résultats de convergence spatiale : la dimension temporelle sera étudiée dans un second temps.

Définition 2.1.4 (maillage conforme) Soit $N_H \in \mathbb{N}^*$ donné. Un maillage conforme K de Ω est défini par une famille finie $\{K_J\}_{J=0}^{N_H-1}$ d'ouverts bornés connexes non vides tels que $\Omega = \bigcup_J \overline{K_J}$, et l'intersection de deux $\overline{K_J}$ soit soit vide, soit un point.

Nous désignerons les K_J par les noms d'arêtes ou de mailles. On note H_J la mesure $\mu(K_J)$ et $H = \max_{J \in \llbracket 0, N_H - 1 \rrbracket} H_J$. Notons également $\{y_J\}_{J=0}^{N_H}$ la discrétisation induite par le maillage, telle que $K_J =]y_J, y_{J+1}[$.

Définition 2.1.5 (méthode de Galerkin éléments finis) On choisit la méthode définie par

- l'espace d'approximation $V_H = \left\{ v \in H_0^1(\Omega) \mid v|_{K_J} \in \mathbb{P}_1(K_J) \quad \forall J \in \llbracket 0, N_H - 1 \rrbracket \right\}$,
- la base $\Phi = \{\varphi_I\}$ définie par $\varphi_I(y_J) = \delta_{IJ}$, $I \in \llbracket 1, N_H \rrbracket$, $J \in \llbracket 1, N_H - 1 \rrbracket$, que l'on s'autorise à appeler **fonctions chapeau**,
- l'interpolateur aux noeuds y_J :

$$\Pi_H u := \mathcal{I}_H u = \sum_{I=1}^{N_H-1} u(y_I) \varphi_I$$

Remarque 2.1.3 En dimension 1, la continuité des éléments de V assure que l'image $\varphi_I(y_J)$ est correctement définie. La famille $\{\varphi_I\}_I$ est bien une base : une manière de le vérifier est de considérer $F : V_H \longrightarrow \mathbb{R}^{N_H-1}$ l'opérateur qui, à $v \in V_H$, associe le vecteur des valeurs de v aux points $\{y_J\}_{J \in \llbracket 1, N_H - 1 \rrbracket}$. F est alors une bijection linéaire, et $\{\varphi_I\}_I$ est l'image réciproque par F de la base canonique de \mathbb{R}^{N_H-1} , donc une base.

Par la suite, nous noterons $|\cdot|_{m,K}$ la semi-norme naturelle de l'espace $H^m(K)$. Établissons premièrement les lemmes suivants :

Lemme 2.1.1 Soit K une arête de mesure H , et $v \in C^\infty(\overline{K})$ s'annulant sur une partie non vide du bord de K . Alors il existe une constante $C > 0$ indépendante de H telle que

$$|v|_{0,K} \leq HC |v|_{1,K}$$

Démonstration

Notons $K =]a, b[$. On suppose sans perte de généralité que v s'annule en a . Soit $x \in K$, on a

$$v(x) - v(a) = v(x) = \int_a^x \frac{dv}{dy}(y) dy, \quad \text{d'où} \quad v^2(x) = \left(\int_a^x 1 \cdot \frac{dv}{dy}(y) dy \right)^2$$

On majore alors le second terme grâce à l'inégalité de Cauchy-Schwarz :

$$v^2(x) \leq \int_a^x 1 dy \int_a^x \left(\frac{dv}{dy}(y) \right)^2 dy \leq (x-a) |v|_{1,K}^2$$

Par intégration sur $[a, b]$, il vient

$$|v|_{0,K}^2 = \int_K v^2(x) dx \leq |v|_{1,K}^2 \int_K (x-a) dx = |v|_{1,K}^2 \frac{(b-a)^2}{2}$$

d'où le résultat avec $C = \frac{1}{\sqrt{2}}$. □

Lemme 2.1.2 Notons $K =]a, b[$ une arête de mesure $H > 0$. Soit $v \in \mathcal{C}^\infty(\overline{K})$, et $\mathcal{I}_H v \in \mathbb{P}_1(K)$ telle que $\mathcal{I}_H v|_{\partial K} = v|_{\partial K}$. Il existe deux constantes $C > 0$ et $C' > 0$ indépendantes de H telles que

$$|v - \mathcal{I}_H v|_{0,K} \leq CH |v - \mathcal{I}_H v|_{1,K} \leq C'H^2 |v - \mathcal{I}_H v|_{2,K} \quad (2.4)$$

Démonstration

La première inégalité est directement issue du lemme (2.1.1). Par le théorème des accroissements finis, il existe $c \in]a, b[$ tel que $v'(c) = \frac{v(b)-v(a)}{b-a} = (\mathcal{I}_H v)'(c)$: on en déduit que sur $]a, c[$ et $]c, b[$, $(v - \mathcal{I}_H v)'$ est \mathcal{C}^∞ et s'annule en une partie du bord. Par le lemme (2.1.1), il existe deux constantes C_1 et C_2 telles que

$$|(v - \mathcal{I}_H v)'|_{0,K_i} \leq HC_i |(v - \mathcal{I}_H v)'|_{1,K_i}, \text{ pour } K_1 =]a, c[\text{ et } K_2 =]c, b[$$

d'où

$$|v - \mathcal{I}_H v|_{1,K}^2 \leq H^2 \max(C_1, C_2)^2 |v - \mathcal{I}_H v|_{2,K}^2$$

et le résultat attendu. □

Définition 2.1.6 On appelle \mathcal{Q} l'espace des fonctions H^2 par morceaux sur le maillage $\{K_J\}$:

$$\mathcal{Q} = \left\{ u \in H_0^1(\Omega) \mid u|_{K_J} \in H^2(K_J) \quad J \in \llbracket 0, N_H - 1 \rrbracket \right\}$$

et l'on note $|\cdot|_{2,H}$ la semi-norme définie sur les éléments de \mathcal{Q} de la manière suivante :

$$|u|_{2,H}^2 = \sum_{J=0}^{N_H-1} |u|_{K_J}^2$$

On a alors le résultat suivant :

THÉORÈME 2.1.3 (CONVERGENCE DE L'INTERPOLATION) Soit $u \in \mathcal{Q}$. Il existe $C > 0$ et $C' > 0$ indépendantes de H telles que

$$\|u - \mathcal{I}_h u\|_{L^2(\Omega)} = CH^2, \quad \|u - \mathcal{I}_h u\|_{H^1(\Omega)} = C'H$$

Démonstration

Soit K_J une arête. On note $\mathcal{D}(K_J) = \mathcal{C}_c^\infty(K_J)$ l'espace des fonctions test sur K_J . Toute fonction test

satisfait les hypothèses du lemme (2.1.2). On sait que $\mathcal{D}(K_J)$ est dense dans $H^2(K_J)$, espace auquel appartient $u - \mathcal{I}_H u$ par hypothèse. Quitte à prendre une suite d'éléments de $\mathcal{D}(K_J)$ tendant vers $u - \mathcal{I}_H u$ et à passer à la limite, les inégalités (larges) suivantes sont valides :

$$|u - \mathcal{I}_h u|_{0,K_J} \leq CH |u - \mathcal{I}_h u|_{1,K_J} \leq C'H^2 |u - \mathcal{I}_h u|_{2,K_J}$$

Remarquons alors que $|u - \mathcal{I}_h u|_{2,K_J} = |u|_{2,K_J}$, car $\partial_x^2 \mathcal{I}_h u$ s'évanouit. En appliquant ceci à toutes les arêtes K_J , on obtient bien

$$\begin{cases} |u - \mathcal{I}_h u|_{0,\Omega} \leq CH^2 |u|_{2,H} \\ |u - \mathcal{I}_h u|_{1,\Omega} \leq C'H |u|_{2,H} \end{cases}$$

□

Le point important de cette étude préliminaire est la condition imposée sur les éléments de \mathcal{Q} pour assurer la convergence. Établissons rapidement le lien avec notre problème d'évolution : supposons que l'on veuille mailler le domaine Ω de manière à ce que pour tout $t \in]0, T[$, la solution $u(\cdot, t)$ appartienne à \mathcal{Q} . On sait depuis la section (1.3.3) que $u(\cdot, t)$ est H^2 seulement sur tout domaine où κ est constant. **Ceci impose donc que le maillage contienne tous les points x_I de discontinuité de κ** : on peut alors imaginer des cas défavorables où la structure du matériau impose un raffinement très coûteux.

Étude dynamique

Choisissons pour espace de Galerkin l'espace V_H défini en (2.1.5), muni de la base $\{\varphi_I\}_{I=1}^{N_H-1}$ des fonctions chapeau. On vérifie que V_H est un sous-espace de V de dimension finie : c'est donc un espace de Hilbert pour le produit scalaire induit par V . En particulier, on peut y définir Π_H^a : $V \rightarrow V_H$ le projecteur orthogonal pour le produit scalaire $a(\cdot, \cdot)$.

Définition 2.1.7 Soit u la solution du problème (1.1), et u_H la solution approchée satisfaisant le problème semi-discrétisé (2.2). Pour tout $t \in]0, T[$, on décompose l'erreur en

$$e(t) = u(\cdot, t) - u_H(\cdot, t) = u(\cdot, t) - \Pi_H^a u_H(\cdot, t) - (u_H(\cdot, t) - \Pi_H^a u(\cdot, t)) = \varepsilon_H(t) - \eta_H(t)$$

Le terme ε_H correspond à la composante de l'erreur orthogonale à V_H pour le produit scalaire $a(\cdot, \cdot)$, tandis que η_H traduit l'écart entre u_H et la meilleure approximation de u dans V_H .

Lemme 2.1.3 Soit M_a la constante de continuité de la forme bilinéaire $a(\cdot, \cdot)$. L'erreur η_H satisfait

$$\sup_{t \in [0, T]} \|\eta_H(t)\|_{H^1(\Omega)} \leq \sqrt{\frac{2}{\alpha}} \left(\frac{1}{\sqrt{2}} F_H(\varepsilon_H, u_0, u_1) + \int_0^T \left\| \frac{d^2 \varepsilon_H}{dt^2}(t) \right\|_{L^2(\Omega)} dt \right)$$

$$\text{où } F_H(\varepsilon_H, u_0, u_1) = \left\| \frac{d\varepsilon_H}{dt}(0) \right\|_{L^2} + \|u_1 - u_{1,H}\|_{L^2} + \sqrt{M_a} (\|\varepsilon_H(0)\|_{H^1} + \|u_0 - u_{0,H}\|_{H^1}).$$

Démonstration

On sait que u et u_H satisfont respectivement le problème variationnel continu (2.1) et semi-discret (2.2). Par linéarité, pour $v = v_H$, l'erreur $u - u_H$ satisfait

$$\begin{cases} \left(\frac{\partial^2(u-u_H)}{\partial t^2}(\cdot, t), v_H \right)_{L^2(\Omega)} + a(u - u_H, v_H) = 0 & \forall v_H \in V_H, \forall t \in]0, T[\\ (u - u_H)_{t=0} = u_0 - u_{0,H}, \frac{\partial(u-u_H)}{\partial t}|_{t=0} = u_1 - u_{1,H} \end{cases}$$

Nous pouvons décomposer $u - u_H$ en $\varepsilon_H - \eta_H$. Par définition de ε_H , le terme $a(\varepsilon_H(t), v_H)$ est nul pour tout $v_H \in V_H$: l'erreur η_H satisfait alors un problème inhomogène donné par

$$\left(\frac{d^2\eta_H}{dt^2}(t), v_H \right)_{L^2(\Omega)} + a(\eta_H(t), v_H) = \left(\frac{d^2\varepsilon_H}{dt^2}(t), v_H \right)_{L^2(\Omega)}$$

Prenons $v_H = \frac{d\eta_H}{dt}(t)$, qui appartient bien à V_H . On omet l'argument (t) pour alléger les notations. On a alors

$$\begin{aligned} & \left(\frac{d^2\eta_H}{dt^2}, \frac{d\eta_H}{dt} \right)_{L^2(\Omega)} + a\left(\eta_H, \frac{d\eta_H}{dt}\right) = \left(\frac{d^2\varepsilon_H}{dt^2}, \frac{d\eta_H}{dt} \right)_{L^2(\Omega)} \\ \implies & \frac{d}{dt} \frac{1}{2} \left(\left\| \frac{d\eta_H}{dt} \right\|_{L^2(\Omega)}^2 + a(\eta_H, \eta_H) \right) \leq \left\| \frac{d^2\varepsilon_H}{dt^2} \right\|_{L^2(\Omega)} \left\| \frac{d\eta_H}{dt} \right\|_{L^2(\Omega)} \end{aligned}$$

Notons $\mathcal{E}_H^\eta = \frac{1}{2} \left\| \frac{d\eta_H}{dt} \right\|_{L^2(\Omega)}^2 + \frac{1}{2} a(\eta_H, \eta_H)$ l'énergie associée à η_H . Le terme $a(\eta_H, \eta_H)$ étant positif, la définition permet d'écrire $\left\| \frac{d\eta_H}{dt} \right\|_{L^2(\Omega)} \leq \sqrt{2\mathcal{E}_H^\eta} \leq 2\sqrt{\mathcal{E}_H^\eta}$. On reconnaît alors la dérivée de $\sqrt{\mathcal{E}_H^\eta}$:

$$\frac{1}{2\sqrt{\mathcal{E}_H^\eta}} \frac{d\mathcal{E}_H^\eta}{dt} \leq \left\| \frac{d^2\varepsilon_H}{dt^2} \right\|_{L^2(\Omega)} \implies \sqrt{\mathcal{E}_H^\eta(t)} - \sqrt{\mathcal{E}_H^\eta(0)} \leq \int_0^t \left\| \frac{d^2\varepsilon_H}{dt^2}(\theta) \right\|_{L^2(\Omega)} d\theta$$

Nous terminons en exploitant l'équivalence entre la norme $a(\cdot, \cdot)$ et la norme $H^1(\Omega)$: on a

$$\frac{\alpha}{2} \|\eta_H(t)\|_{H^1(\Omega)}^2 \leq \frac{1}{2} a(\eta_H, \eta_H) \leq \mathcal{E}_H^\eta(t) \quad \text{et} \quad \mathcal{E}_H^\eta(0) \leq \frac{1}{2} \left\| \frac{d\eta_H}{dt}(0) \right\|_{L^2(\Omega)}^2 + \frac{M_a}{2} \|\eta_H(0)\|_{H^1(\Omega)}^2$$

Pour des raisons d'élégance, on s'autorise à majorer $a^2 + b^2$ par $(a + b)^2$ dans le second terme (tout étant positif). En remarquant que $\eta_H(t) = \varepsilon_H(t) - (u(\cdot, t) - u_H(\cdot, t))$ et en utilisant l'inégalité triangulaire, il vient

$$\sqrt{\mathcal{E}_H^\eta(0)} \leq \frac{1}{\sqrt{2}} \left(\left\| \frac{d\varepsilon_H}{dt}(0) \right\|_{L^2(\Omega)} + \|u_1 - u_{1,H}\|_{L^2(\Omega)} + \sqrt{M_a} \left(\|\varepsilon_H(0)\|_{H^1(\Omega)} + \|u_0 - u_{0,H}\|_{H^1(\Omega)} \right) \right)$$

ce que l'on s'empressera de noter $\frac{1}{\sqrt{2}} F_H(\varepsilon_H, u_0, u_1)$. On obtient

$$\|\eta_H(t)\|_{H^1(\Omega)} \leq \sqrt{\frac{2}{\alpha}} \left(\frac{1}{\sqrt{2}} F_H(\varepsilon_H, u_0, u_1) + \int_0^t \left\| \frac{d^2\varepsilon_H}{dt^2}(\theta) \right\|_{L^2(\Omega)} d\theta \right)$$

et le résultat attendu s'obtient par passage au sup sur $t \in [0, T]$.

□

Proposition 2.1.4 (convergence du schéma éléments finis) Sous l'hypothèse $u \in \mathcal{C}^2([0, T], \mathcal{D}(\mathcal{A}) \cap \mathcal{Q})$ et $u_1 \in \mathcal{Q}$, il existe deux constantes C_T et C'_T indépendantes de H telle que

$$\begin{aligned}\sup_{t \in [0, T]} \|u(\cdot, t) - u_h(\cdot, t)\|_{H^1(\Omega)} &\leq CH \left(H |u_1|_{2, H} + \sup_{t \in [0, T]} \left| \frac{\partial^2 u}{\partial t^2}(\cdot, t) \right|_{2, H} + \sup_{t \in [0, T]} |u(\cdot, t)|_{2, H} \right) \\ \sup_{t \in [0, T]} \|u(\cdot, t) - u_h(\cdot, t)\|_{L^2(\Omega)} &\leq C'H^2 \left(H |u_1|_{2, H} + \sup_{t \in [0, T]} \left| \frac{\partial^2 u}{\partial t^2}(\cdot, t) \right|_{2, H} + \sup_{t \in [0, T]} |u(\cdot, t)|_{2, H} \right)\end{aligned}$$

Démonstration

Décomposons l'erreur $u - u_h$ grâce à l'inégalité triangulaire :

$$\sup_{t \in [0, T]} \|e(t)\|_{H^1(\Omega)} \leq \sup_{t \in [0, T]} \|\eta_H(t)\|_{H^1(\Omega)} + \sup_{t \in [0, T]} \|\varepsilon_H(t)\|_{H^1(\Omega)}$$

Majorons le terme $\varepsilon_H(t)$. Notons α et β les constantes d'équivalence des normes $\|\cdot\|_a := a(\cdot, \cdot)$ et $\|\cdot\|_{H^1(\Omega)}$. On a

$$\alpha \|u(\cdot, t) - \Pi_h^a u(\cdot, t)\|_{H^1(\Omega)} \leq \|u(\cdot, t) - \Pi_h^a u(\cdot, t)\|_a \leq \|u(\cdot, t) - \mathcal{I}_H u(\cdot, t)\|_a \leq \beta \|u(\cdot, t) - \mathcal{I}_H u(\cdot, t)\|_{H^1(\Omega)}$$

Soit $t \in [0, T]$. Sous l'hypothèse $u(\cdot, t) \in \mathcal{Q}$, le lemme (2.1.2) donne l'existence de $C > 0$ et $C' > 0$ indépendantes de H telles que

$$\|u(\cdot, t) - \mathcal{I}_H u(\cdot, t)\|_{H^1(\Omega)} \leq HC |u(\cdot, t)|_{2, H}$$

$$\left\| \frac{\partial^2 u}{\partial t^2}(\cdot, t) - \Pi_h^a \frac{\partial^2 u}{\partial t^2}(\cdot, t) \right\|_{L^2(\Omega)} \leq \left\| \frac{\partial^2 u}{\partial t^2}(\cdot, t) - \mathcal{I}_H \frac{\partial^2 u}{\partial t^2}(\cdot, t) \right\|_{L^2(\Omega)} \leq C'H^2 \left| \frac{\partial^2 u}{\partial t^2} \right|_{2, H}$$

Intéressons-nous maintenant à $\eta_H(t)$. On utilise la majoration du lemme (2.1.3). Dans notre cas, $u_0 = 0$, et les termes $\|\varepsilon_H(0)\|_{H^1(\Omega)} + \|u_0 - u_{0, H}\|_{H^1(\Omega)}$ s'annulent. De plus, pour $u_{1, H} = \mathcal{I}_H u_1$,

$$\left\| \frac{d\varepsilon_H}{dt}(0) \right\|_{L^2(\Omega)} = \|u_1 - \Pi_h^a u_1\|_{L^2(\Omega)} \leq \|u_1 - \Pi_h^a u_1\|_{H^1(\Omega)} \leq \|u_1 - \mathcal{I}_H u_1\|_{H^1(\Omega)} \leq C'H^2 |u_1|_{2, H}$$

ce qui nous donne le résultat en norme $H^1(\Omega) \times \mathcal{C}^\infty([0, T])$. Le résultat en norme $L^2(\Omega) \times \mathcal{C}^\infty([0, T])$ est donné par le lemme (2.1.2) : on a alors

$$\|u(\cdot, t) - \mathcal{I}_H u(\cdot, t)\|_{L^2(\Omega)} \leq CH \|u(\cdot, t) - \mathcal{I}_H u(\cdot, t)\|_{H^1(\Omega)} \quad \forall t \in [0, T]$$

et le passage au sup préserve cette inégalité. □

De cette étude, on retiendra les points principaux suivants :

- Sous réserve que la solution soit suffisamment régulière sur chaque arête du maillage éléments finis, le schéma converge à l'ordre 2 en norme $\mathcal{C}^\infty([0, T]) \times L^2(\Omega)$ et à l'ordre 1 en norme $\mathcal{C}^\infty([0, T]) \times H^1(\Omega)$.

- Notons λ la longueur d'onde de u : on appellera par la suite **échelle macroscopique** l'ordre de grandeur des pas de maillage pour lesquels u est échantillonnée en k points, $k \in \llbracket 10, 20 \rrbracket$. Par constraste, on appellera **échelle de la microstructure** l'ordre de grandeur des domaines $]x_i, x_{i+1}[$ où κ est constant. La condition de régularité devient contraignante lorsque l'échelle micro est inférieure à l'échelle macro : le raffinement imposé par le régime de convergence peut entraîner un surcoût de calcul indésirable.

2.2 Semi-discrétisation en temps

Cette rapide section a pour but de définir les énergies semi-discrètes en temps, qui nous serviront par la suite.

2.2.1 Schéma

La discrétisation en temps emploie des différences finies. On a choisi l'approximation classique de la dérivée seconde, d'ordre 2.

Définition 2.2.1 (schéma semi-discrétisé en temps) Soit $T > 0$ fini, et $\Delta t > 0$ un pas temporel. Notons $u^n \in \mathcal{D}(\mathcal{A})$ l'approximation de $u(\cdot, t_n)$ au temps $t_n = n\Delta t$. La suite $(u^n)_{n\Delta t \in [0, T]} \subset \mathcal{D}(\mathcal{A})$ est l'inconnue du schéma semi-discrétisé donné par

$$\begin{cases} m\left(\frac{u^{n+1}-2u^n+u^{n-1}}{\Delta t^2}, v\right) + a(u^n, v) = 0 & \forall v \in V, \forall n \text{ tq } n\Delta t \leqslant T \\ u^0 = u_0 \\ m(u^1, v) = m(u_0 + \Delta t u_1, v) - \frac{\Delta t^2}{2} a(u_0, v) & \forall v \in H_0^1(\Omega) \end{cases} \quad (2.5)$$

Remarque 2.2.1 (intégration de u_1) Le problème satisfait par u^1 est obtenu en considérant une approximation centrée d'ordre 2 de la dérivée première : ceci permet de conserver l'ordre du schéma. Le théorème de Lax-Milgram (1.2.2) assure qu'il existe une unique solution. Dans le cas particulier $u_0 = 0$, elle est donnée par $u^1 = \Delta t u_1$.

2.2.2 Énergie III

Sans grande originalité, on se donne les fonctions suivantes :

Définition 2.2.2 Les énergies semi-discrétisées en temps sont définies par

$$\begin{aligned} \mathcal{K}^{n+1/2} &= \frac{1}{2} m\left(\frac{u^{n+1}-u^n}{\Delta t}, \frac{u^{n+1}-u^n}{\Delta t}\right) && \text{l'énergie cinétique} \\ \mathcal{P}^{n+1/2} &= \frac{1}{2} a(u^{n+1}, u^n) && \text{l'énergie potentielle} \\ \mathcal{E}^{n+1/2} &= \mathcal{K}^{n+1/2} + \mathcal{P}^{n+1/2} && \text{l'énergie totale} \end{aligned}$$

Proposition 2.2.1 L'énergie totale se conserve.

Démonstration

Évaluons la formulation variationnelle satisfaite par $(u^n)_{n\Delta t \in [0,T]}$ en $v = \frac{u^{n+1} - u^{n-1}}{2\Delta t}$ la vitesse centrée :

$$m\left(\frac{u^{n+1} - 2u^n + u^{n-1}}{\Delta t^2}, \frac{u^{n+1} - u^{n-1}}{2\Delta t}\right) + a\left(u^n, \frac{u^{n+1} - u^{n-1}}{2\Delta t}\right) = 0$$

On utilise la décomposition

$$m\left(\frac{u^{n+1} - 2u^n + u^{n-1}}{\Delta t^2}, \frac{u^{n+1} - u^{n-1}}{2\Delta t}\right) = \frac{1}{2\Delta t} \left(m\left(\frac{u^{n+1} - u^n}{\Delta t}, \frac{u^{n+1} - u^n}{\Delta t}\right) - m\left(\frac{u^n - u^{n-1}}{\Delta t}, \frac{u^n - u^{n-1}}{\Delta t}\right) \right)$$

pour reconnaître le résultat attendu :

$$\frac{1}{\Delta t} (\mathcal{K}^{n+1/2} - \mathcal{K}^{n-1/2}) + \frac{1}{\Delta t} (\mathcal{P}^{n+1/2} - \mathcal{P}^{n-1/2}) = \frac{\mathcal{E}^{n+1/2} - \mathcal{E}^{n-1/2}}{\Delta t} = 0$$

□

2.3 Schéma complètement discréteisé

2.3.1 Schéma

Le schéma complètement discréteisé s'obtient indifféremment en utilisant des différences finies sur le schéma semi-discret en espace (2.1.1), ou la discréteisation dans une base de Galerkin du schéma semi-discret en temps (2.2.1).

Soit V_h l'espace de Galerkin de dimension N_h , Φ une base de V_h et Π_h le projecteur choisi.

Définition 2.3.1 (schéma complètement discréteisé) Pour $T > 0$ le temps final et Δt le pas temporel choisi, on note $U^n \in \mathbb{R}^{N_h}$ le vecteur des coordonnées dans la base de Galerkin de l'approximation de $u(\cdot, n\Delta t)$. La suite $(U^n)_{n\Delta t \in [0,T]}$ est l'inconnue du problème suivant :

$$\begin{cases} \mathbb{M} \frac{U^{n+1} - 2U^n + U^{n-1}}{\Delta t^2} + \mathbb{K} U^n = 0 \\ U^0 = \Pi_h u_0 \\ U^1 = \Delta t \Pi_h u_1 + \left(\mathbb{I} - \frac{\Delta t^2}{2} \mathbb{M}^{-1} \mathbb{K} \right) \Pi_h u_0 \end{cases} \quad (2.6)$$

Remarque 2.3.1 (intégration des conditions initiales) L'intégration de u_1 se fait en utilisant une approximation centrée d'ordre 2 de la dérivée, puis en évaluant le schéma en $n = 0$. Dans le cas $u_0 = 0$, on retrouve $U^1 = \Delta t \Pi_h u_1$.

Au terme de ces différentes discréteisations, nous disposons d'un schéma compréhensible par un ordinateur. Terminons par les propriétés de stabilité de ce schéma.

2.3.2 Énergie IV

Comme précédemment, le schéma implique la conservation d'une énergie totalement discrétisée.

Définition 2.3.2 Les énergies discrètes sont données par

$$\begin{aligned}\mathcal{K}_H^{n+1/2} &= \frac{1}{2} \left(\frac{U^{n+1} - U^n}{\Delta t} \right)^t \mathbb{M} \frac{U^{n+1} - U^n}{\Delta t} && \text{l'énergie cinétique} \\ \mathcal{P}_H^{n+1/2} &= \frac{1}{2} \left(U^{n+1} \right)^t \mathbb{K} U^n && \text{l'énergie potentielle} \\ \mathcal{E}_H^{n+1/2} &= \mathcal{K}_H^{n+1/2} + \mathcal{P}_H^{n+1/2} && \text{l'énergie totale}\end{aligned}$$

Proposition 2.3.1 L'énergie totale se conserve.

Démonstration

La démonstration reprend la structure de la preuve 1.3.1, où l'on emploie l'artifice suivant :

$$\begin{aligned}\left(\mathbb{M} \frac{U^{n+1} - 2U^n + U^{n-1}}{\Delta t^2}, \frac{U^{n+1} - U^{n-1}}{2\Delta t} \right) &= \frac{1}{2\Delta t} \left(\mathbb{M} \left(\frac{U^{n+1} - U^n}{\Delta t} - \frac{U^n - U^{n-1}}{\Delta t} \right), \frac{U^{n+1} - U^n}{\Delta t} + \frac{U^n - U^{n-1}}{\Delta t} \right) \\ &= \frac{1}{2\Delta t} \left(\left(\mathbb{M} \frac{U^{n+1} - U^n}{\Delta t}, \frac{U^{n+1} - U^n}{\Delta t} \right) - \left(\mathbb{M} \frac{U^n - U^{n-1}}{\Delta t}, \frac{U^n - U^{n-1}}{\Delta t} \right) \right)\end{aligned}$$

□

2.3.3 Condition CFL

Le but de cette section est d'établir une condition de Courant-Friedrich-Lowy (CFL) liant le pas temporel Δt à l'espace statique V_h , et sous laquelle le schéma (2.6) soit stable.

THÉORÈME 2.3.1 Le schéma est stable sous la condition CFL

$$\Delta t < \overline{\Delta t}, \quad \overline{\Delta t} = 2 \left(\sup_{v_h \in V_h \setminus \{0\}} \frac{a(v_h, v_h)}{m(v_h, v_h)} \right)^{-1/2}$$

Pour ceci, on se donne le couple de lemmes suivant :

Lemme 2.3.1 S'il existe $\mu(\cdot, \cdot)$ une forme bilinéaire définie positive telle que $\mathcal{E}^{n+1/2} \geq \mu \left(\frac{U^{n+1} - U^n}{\Delta t}, \frac{U^{n+1} - U^n}{\Delta t} \right)$ pour tout $n \in \mathbb{N}$, alors le schéma est stable pour la norme induite par μ .

Démonstration

Supposons que $\mu(\cdot, \cdot)$ soit un produit scalaire, et notons $\|u\|_\mu^2 := \mu(u, u)$. L'hypothèse et l'identité

triangulaire seconde nous donnent

$$\mathcal{E}^{n+1/2} \geq \left\| \frac{u^{n+1} - u^n}{\Delta t} \right\|_{\mu}^2 \geq \frac{1}{\Delta t^2} \left(\|u^{n+1}\|_{\mu} - \|u^n\|_{\mu} \right)^2$$

L'énergie se conservant au cours des itérations, on peut écrire

$$\|u^n\|_{\mu} \leq \|u^{n-1}\|_{\mu} + \Delta t \sqrt{\mathcal{E}^{n-1/2}} \leq \dots \leq \|u^0\|_{\mu} + n \Delta t \sqrt{\mathcal{E}^{1/2}} \leq \|u^0\|_{\mu} + T \sqrt{\mathcal{E}^{1/2}}$$

ce qui traduit la stabilité du schéma pour la norme demandée. □

Lemme 2.3.2 Soient deux formes bilinéaires symétriques définies positives $\alpha(\cdot, \cdot)$ et $\beta(\cdot, \cdot)$, et $\mu(\cdot, \cdot) := \alpha(\cdot, \cdot) - \Delta t^2 \beta(\cdot, \cdot)$. S'il existe $\overline{\Delta t}$ tel que $\forall \Delta t \in]0, \overline{\Delta t}[$, μ soit définie positive, alors les normes induites par μ et α sont équivalentes (au prix de constantes dépendantes de Δt).

Démonstration

Le sens $\mu(u, u) \leq \alpha(u, u)$ est immédiat. Supposons l'hypothèse vérifiée, et soit $\Delta t \in]0, \overline{\Delta t}[$. On a alors

$$\mu(u, u) \geq 0 \implies \alpha(\cdot, \cdot) \geq \overline{\Delta t}^2 \beta(\cdot, \cdot)$$

et

$$\alpha(u, u) = \mu(u, u) + \Delta t^2 \beta(u, u) \leq \mu(u, u) + \left(\frac{\Delta t}{\overline{\Delta t}} \right)^2 \alpha(u, u)$$

d'où

$$\alpha(u, u) \leq \frac{1}{1 - (\Delta t / \overline{\Delta t})^2} \mu(u, u)$$

□

Munis des lemmes précédents, nous pouvons démontrer le théorème (2.3.1).

Démonstration

Commençons par extraire une forme bilinéaire minorant l'énergie $\mathcal{E}^{n+1/2}$. Nous utiliserons l'identité

$$a\left(\frac{u^{n+1} + u^n}{\Delta t}, \frac{u^{n+1} + u^n}{\Delta t}\right) - a\left(\frac{u^{n+1} - u^n}{\Delta t}, \frac{u^{n+1} - u^n}{\Delta t}\right) = \frac{4}{\Delta t^2} a(u^{n+1}, u^n)$$

Le terme $a\left(\frac{u^{n+1} + u^n}{\Delta t}, \frac{u^{n+1} + u^n}{\Delta t}\right)$ étant positif, on a immédiatement

$$\begin{aligned} \mathcal{E}^{n+1/2} &= \frac{1}{2} \left[m\left(\frac{u^{n+1} - u^n}{\Delta t}, \frac{u^{n+1} - u^n}{\Delta t}\right) + a(u^{n+1}, u^n) \right] \\ &\geq \frac{1}{2} \left[m\left(\frac{u^{n+1} - u^n}{\Delta t}, \frac{u^{n+1} - u^n}{\Delta t}\right) - \frac{\Delta t^2}{4} a\left(\frac{u^{n+1} - u^n}{\Delta t}, \frac{u^{n+1} - u^n}{\Delta t}\right) \right] \end{aligned}$$

On définit $\mu(u, v) = m(u, v) - \frac{\Delta t^2}{4} a(u, v)$. Cherchons la condition sous laquelle μ est définie positive,

c'est-à-dire sous laquelle $\forall v_h \in V_h \setminus \{0\}$,

$$\mu(v_h, v_h) > 0 \implies m(v_h, v_h) > \frac{\Delta t^2}{4} a(v_h, v_h) \implies \Delta t < \sqrt{4 \frac{m(v_h, v_h)}{a(v_h, v_h)}} = 2 \left(\frac{a(v_h, v_h)}{m(v_h, v_h)} \right)^{-1/2}$$

ce qui, par passage au sup sur $v_h \in V_h$, nous donne la condition du théorème. Sachant μ définie positive, le lemme (2.3.1) nous donne la stabilité du schéma en norme μ , et le lemme (2.3.2) nous permet d'assurer la stabilité pour la norme $L^2(\Omega)$. □

Remarque 1. Cette condition CFL est indépendante de la base de Galerkin choisie : en particulier, elle pourra servir de point de départ pour les études à venir sur des bases multi-échelles.

Obtention numérique de la CFL

On a le résultat pratique suivant :

Proposition 2.3.2 *Le schéma complètement discréétisé (2.6) est stable sous la condition CFL*

$$\Delta t < \overline{\Delta t}, \quad \overline{\Delta t} \leq 2 \left(\sup_{U \in \mathbb{R}^{N_h} \setminus \{0\}} \frac{U^t \mathbb{K} U}{U^t \mathbb{M} U} \right)^{-1/2} = \frac{2}{\sqrt{\rho(\mathbb{M}^{-1} \mathbb{K})}}$$

La première expression de la CFL correspond à la discréétisation en espace de celle obtenue dans le cas continu. La preuve de stabilité est une transposition exacte de celle fournie précédemment, en utilisant les correspondances

$$U^t \mathbb{M} U \leftrightarrow m(u, u) \quad \text{et} \quad U^t \mathbb{K} U \leftrightarrow a(u, u) \quad \text{pour } u = \sum_I U_I \varphi_I$$

On évite au lecteur la répétition de cet argumentaire. Démontrons alors l'égalité

$$\rho(\mathbb{M}^{-1} \mathbb{K}) = \sup_{U \in \mathbb{R}^{N_h} \setminus \{0\}} \frac{U^t \mathbb{K} U}{U^t \mathbb{M} U}$$

Démonstration

Notons λ le plus petit réel tel que $U^t \mathbb{K} U \leq \lambda U^t \mathbb{M} U$ pour tout $U \in \mathbb{R}^{N_h} \setminus \{0\}$. Soit $\sqrt{\mathbb{M}}$ l'unique matrice symétrique définie positive telle que $\mathbb{M} = \sqrt{\mathbb{M}}^2$. On effectue le changement de base $V = \sqrt{\mathbb{M}} U$ et on divise par $V^t V$ quand ce terme est non nul :

$$\frac{V^t \sqrt{\mathbb{M}}^{-1} \mathbb{K} \sqrt{\mathbb{M}}^{-1} V}{V^t V} \leq \lambda \quad \forall V \in \mathbb{R}^{N_h} \setminus \{0\}$$

On prend alors le sup sur $U \in \mathbb{R}^{N_h}$ (donc, par positivité de $\sqrt{\mathbb{M}}$, sur $V \in \mathbb{R}^{N_h}$). On reconnaît dans l'expression précédente une caractérisation de $\rho(\sqrt{\mathbb{M}}^{-1} \mathbb{K} \sqrt{\mathbb{M}}^{-1})$: or, les matrices $\sqrt{\mathbb{M}}^{-1} \mathbb{K} \sqrt{\mathbb{M}}^{-1}$ et $\mathbb{M}^{-1} \mathbb{K}$ sont semblables, donc partagent le même spectre. □

Remarque 2.3.2 (dépendance au pas spatial) Dans le cadre de la méthode des éléments finis à pas H constant, le calcul de \mathbb{M} et \mathbb{K} montre l'existence de deux matrices \mathcal{M} et \mathcal{K} indépendantes de H telles que $\mathbb{M} = H\mathcal{M}$ et $\mathbb{K} = H^{-1}\mathcal{K}$. La CFL est alors donnée par

$$\overline{\Delta t} = \frac{2}{\sqrt{\rho(H^{-2}\mathcal{M}^{-1}\mathcal{K})}} = H \frac{2}{\sqrt{\rho(\mathcal{M}^{-1}\mathcal{K})}}$$

d'où une dépendance linéaire de $\overline{\Delta t}$ en H . Ce comportement se généralise au pas non constant par majoration : le résultat est alors obtenu pour H_{\max} le pas maximal. **Tout raffinement en espace induit alors un raffinement proportionnel en temps !**

2.3.4 Condensation de masse

Pour conclure cette partie, introduisons une optimisation fondamentale de la méthode des éléments finis. Nous ne développerons que le cas de la dimension 1, mais elle se généralise en dimensions supérieures (pour le bonheur des numériciens).

Le schéma complètement discréteisé (2.6) est propagé selon

$$U^{n+1} = -U^{n-1} + 2(\mathbb{I} - \Delta t^2 \mathbb{M}^{-1} \mathbb{K}) U^n$$

L'inversion de la matrice de masse est une étape critique, pour la raison suivante :

Proposition 2.3.3 *Les matrices de masse et de rigidité associées à la base éléments finis sont tridiagonales.*

Démonstration

Les seules fonctions chapeau dont les supports ne soient pas disjoints sont celles qui partagent une arête, donc qui sont associées à deux y_I consécutifs. Pour la numérotation choisie, ceci implique que tous les produits scalaires $a(\cdot, \cdot)$ et $(\cdot, \cdot)_{L^2(\Omega)}$ entre fonctions de base non consécutives sont nuls : il ne reste que les termes tridiagonaux. □

Définition 2.3.3 On se donne la formule de quadrature des trapèzes

$$\mathfrak{C}_{[a,b]}(f) = \int_a^b f(x) dx \approx \frac{f(a) + f(b)}{2} \cdot (b - a)$$

Soit alors $K = \{K_J\}_{J \in \llbracket 0, N_H \rrbracket}$ un maillage conforme de Ω . On définit un produit scalaire sur l'espace V_H par

$$\langle u, v \rangle_H = \sum_{J=0}^{N_H-1} \mathfrak{C}_{K_J}(uv|_{K_J})$$

Proposition 2.3.4 En substituant à $(\cdot, \cdot)_{L^2(\Omega)}$ son approximation $\langle \cdot, \cdot \rangle_H$ par la formule des trapèzes, la matrice de masse devient diagonale.

Démonstration

Les fonctions chapeau sont définies par $\varphi_I(y_J) = \delta_{IJ}$: aussi, pour φ_I et φ_{I+1} deux fonctions consécutives, on a

$$\langle \varphi_I, \varphi_{I+1} \rangle_H = 0 + 1 \times 0 + 0 \times 1 + 0 = 0$$

Les diagonales supérieure et inférieure s'annulent. La diagonale principale contiendra les termes

$$\langle \varphi_I, \varphi_I \rangle_H = \frac{\mu(K_I) + \mu(K_J)}{2} > 0$$

donc \mathbb{M} est diagonale définie positive. □

Remarque 2.3.3 (convergence du schéma condensé) La question légitime de l'influence de la condensation de masse est traitée en détail dans Raviart and Thomas [1988], Joly [2007] : on admettra que la formule de quadrature employée conserve le même ordre de convergence.

Conclusion sur les méthodes de Galerkin

Au terme de cette partie, nous disposons de deux méthodes numériques dérivant d'une même classe d'approximation interne.

La **méthode spectrale** a pour avantage de ne dépendre que de l'approximation des conditions initiales. Toute la difficulté réside dans l'obtention d'une base spectrale de dimension finie : nous verrons que c'est possible en dimension 1, mais le passage aux dimensions supérieures pour des maillages non structurés n'est pas aussi clément.

La **méthode éléments finis** se généralise plus facilement. Cependant, dans le cas inhomogène, la micro-structure impose un maillage à l'échelle des variations de κ . Ce premier surcoût éventuel est aggravé par la condition CFL, qui fixe une discrétisation en temps dépendant linéairement du pas en espace.

Le chapitre suivant est dédié à cette dépendance quadratique du nombre de coefficients à la microstructure. Nous tenterons d'obtenir une condition CFL ne dépendant que de l'échelle macro.

Chapitre 3

Enrichissement par fonctions bulles

3.1 Quelques définitions

3.1.1 Espace W_h

Soit un maillage conforme $K = \{K_J\}_{J \in \llbracket 0, N_H - 1 \rrbracket}$ de Ω . Nous utiliserons la définition suivante :

Définition 3.1.1 (espace bulle) Soit K_J une arête. Un espace bulle local $\mathcal{B}(K_J)$ est un espace vectoriel de dimension finie inclus dans $H_0^1(K_J)$. On notera l'espace bulle global

$$\mathcal{B} = \left\{ v : \Omega \rightarrow \mathbb{R} \mid v|_{K_J} \in \mathcal{B}(K_J) \right\}$$

Il n'y a aucune raison que les $\mathcal{B}(K_J)$ aient tous la même dimension : par souci de simplicité, dans toute la suite, nous considérerons que $\dim(\mathcal{B}(K_J)) = N_B$ pour tout $J \in \llbracket 0, N_H - 1 \rrbracket$.

Définition 3.1.2 (espace enrichi) Notons V_H l'espace éléments finis associé au maillage K , muni du projecteur $\Pi_H = \mathcal{I}_H$. L'espace enrichi est

$$W_h = V_H \oplus \mathcal{B}$$

où H représente la longueur caractéristique du maillage éléments finis, et h est un paramètre qui tend vers 0 quand la dimension de \mathcal{B} augmente.

Chaque fonction de l'espace enrichi pourra être exprimée comme $u_h = u_H + u_b$, où

- $u_H \in V_H$ est la contribution éléments finis, que l'on appellera *standard* ou *classique*,
- $u_b \in \mathcal{B}$ représente la contribution des bulles.

Proposition 3.1.1 (propriétés de W_h)

1. Les espaces vectoriels V_H et \mathcal{B} sont linéairement indépendants.
2. Soit $\Pi_{\mathcal{B}} : V \rightarrow \mathcal{B}$ un projecteur sur l'espace bulle. Alors $\mathcal{I}_H \circ \Pi_{\mathcal{B}} = 0$.
3. L'espace enrichi W_h est un sous-espace linéaire fermé inclus dans $H_0^1(\Omega)$.

Démonstration

1. Soit $v \in V_H \cap \mathcal{B}$. Sur chaque maille K_J , la restriction $v|_{K_J}$ est un polynôme du premier degré (par définition de V_H), continu sur $\overline{K_J}$ (par continuité des éléments de V_H) qui s'annule aux bords de K_J (par définition de \mathcal{B}). Donc v est identiquement nulle, et $V_H \cap \mathcal{B} = \{0\}$.
2. Soit $v \in V$. La projection $\Pi_{\mathcal{B}}v$ est nulle aux noeuds du maillage, donc l'interpolée $\mathcal{I}_H \Pi_{\mathcal{B}}v$ est identiquement nulle. Remarquons que $\Pi_{\mathcal{B}} \circ \mathcal{I}_H$ n'est pas l'opérateur nul : ceci vient du fait que \mathcal{B} est stable par interpolation aux noeuds du maillage, alors que V_H n'est pas stable par projection sur les bulles.
3. Par définition, $V_H \subset H_0^1(\Omega)$ et $\mathcal{B} \subset H_0^1(\Omega)$ sont deux sous-espaces linéaires fermés, donc leur somme directe également. On vérifie au passage que W_h est un sous-espace vectoriel de dimension finie, donc un espace de Hilbert pour les produits scalaires hérités de $H_0^1(\Omega)$. L'approximation utilisée est donc interne : nous sommes bien dans le cadre des méthodes de Galerkin.

□

3.1.2 Représentation dans l'espace enrichi

Notons $V := H_0^1(\Omega)$. On se donne un projecteur $\Pi_{\mathcal{B}} : V \rightarrow \mathcal{B}$.

Définition 3.1.3 Une fonction de $v \in V$ sera représentée dans W_h par

$$\Pi_h v = \mathcal{I}_H v + \Pi_{\mathcal{B}}(v - \mathcal{I}_H v)$$

Proposition 3.1.2 Π_h est une projection.

Démonstration

On utilise la propriété $\mathcal{I}_H \circ \Pi_{\mathcal{B}} = 0$ pour écrire

$$\mathcal{I}_H \circ \Pi_h = \mathcal{I}_H \circ (\mathcal{I}_H + \Pi_{\mathcal{B}}(id - \mathcal{I}_H)) = \mathcal{I}_H$$

Soit $u \in V$. On a

$$\begin{aligned} \Pi_h \circ \Pi_h u &= \mathcal{I}_H \circ \Pi_h u + \Pi_{\mathcal{B}} \circ (\Pi_h u - \mathcal{I}_H \circ \Pi_h u) \\ &= \mathcal{I}_H u + \Pi_{\mathcal{B}} \circ (\mathcal{I}_H v + \Pi_{\mathcal{B}}(v - \mathcal{I}_H v) - \mathcal{I}_H u) \\ &= \mathcal{I}_H u + \Pi_{\mathcal{B}}(v - \mathcal{I}_H v) \end{aligned}$$

d'où $\Pi_h = \Pi_h \circ \Pi_h$.

□

3.1.3 Cas considérés

Définition 3.1.4 (espaces utilisés)

- Espace bulle **micro éléments finis** : sur chaque maille, on se donne un espace éléments finis de type \mathbb{P}_1 contenu dans $H_0^1(K_J)$.

- Espace bulle **spectral** : sur chaque maille K_J , on résout le problème spectral local donné par
 « Trouver λ tel qu'il existe $u \neq 0$ satisfaisant $a(u, v) = \lambda m(u, v) \quad \forall v \in H_0^1(K_J)$ »
 et l'on constitue une base de $\mathcal{B}(K_J)$ à l'aide des $N_{\mathcal{B}}$ premières fonctions propres solutions.

Pour ces deux cas, on vérifie le résultat suivant :

Proposition 3.1.3 *Supposons que κ soit constante sur chaque maille micro-éléments finis. Pour H fixé, $\Pi_h u$ converge vers u quand $N_{\mathcal{B}}$ tend vers $+\infty$.*

Démonstration

Soit $u \in V$, et $\mathcal{I}_H u$ son interpolée sur le maillage éléments finis macroscopique. Pour $\|\cdot\|_{\Omega}$ désignant indistinctement $\|\cdot\|_{L^2(\Omega)}$ ou $\|\cdot\|_{H^1(\Omega)}$, l'erreur se décompose en

$$\|u - \Pi_h u\|_{\Omega}^2 = \sum_{J=0}^{N_H-1} \|u - \Pi_h u\|_{K_J}^2 = \sum_{J=0}^{N_H-1} \|(u - \mathcal{I}_H u) - \Pi_{\mathcal{B}}(u - \mathcal{I}_H u)\|_{K_J}^2$$

Or, les éléments de convergence étudiés précédemment nous assurent que sur chaque maille K_J , la projection $\Pi_{\mathcal{B}}(u - \mathcal{I}_H u)$ converge vers $u - \mathcal{I}_h u$ (pour les deux cas particuliers d'espaces bulles) : d'où le résultat. □

3.2 Schéma IMEX

Cette section a pour but d'introduire un schéma numérique implicite-explicite (IMEX) basé sur la distinction des contributions éléments finis et bulles.

3.2.1 Problème variationnel associé

On prend pour point de départ le problème variationnel semi-discrétisé en temps 2.5, valable pour toute méthode de Galerkin. La solution u est approchée par la suite $(u^n)_{n \in \mathbb{N}}$ satisfaisant

$$m\left(\frac{u_h^{n+1} - 2u_h^n + u_h^{n-1}}{\Delta t^2}, v_h\right) + a(u_h^n, v_h) = 0 \quad \forall v_h \in W_h$$

Développons le terme $a(u_h^n, v_h)$ selon la décomposition $v_h = v_H + v_b$:

$$m\left(\frac{u_h^{n+1} - 2u_h^n + u_h^{n-1}}{\Delta t^2}, v_h\right) + a(u_H^n, v_H) + a(u_b^n, v_H) + a(u_H^n, v_b) + a(u_b^n, v_b) = 0$$

On introduit alors un θ -schéma sur le terme $a(u_b^n, v_b)$:

Définition 3.2.1 (θ -schéma IMEX) Soit $\theta \in [0, 1]$. On cherche la suite $(u^n)_{n \in \mathbb{N}}$ satisfaisant

pour tout $v_h \in W_h$

$$\begin{cases} m(D_{\Delta t}^2 u_h, v_h) + a(u_H^n, v_H) + a(u_b^n, v_H) + a(u_H^n, v_b) + a((1-\theta)u_b^n + \theta \frac{u_b^{n+1} + u_b^{n-1}}{2}, v_b) = 0 \\ u_h^0 = 0, \quad u_h^1 = \Delta t \Pi_h u_1 \end{cases} \quad (3.1)$$

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

Remarque 3.2.1 On pourra par la suite utiliser l'expression équivalente

$$m\left(\frac{u_h^{n+1} - 2u_h^n + u_h^{n-1}}{\Delta t^2}, v_h\right) + a(u_h^n, v_h) + \theta a\left(\frac{u_b^{n+1} - 2u_b^n + u_b^{n-1}}{2}, v_b\right) = 0 \quad \forall v_h \in W_h$$

Discrétisation totale

On prendra pour base de W_H une famille composée

- d'une base $\Phi = \{\varphi_I\}_{I \in \llbracket 1, N_H - 1 \rrbracket}$ de l'espace V_H ,
- de $N_H - 1$ bases $\Psi_J = \{\psi_{J,j}\}_{j \in \llbracket 1, N_B \rrbracket}$ des espaces bulle locaux $\mathcal{B}(K_J)$.

On vérifie facilement que la famille $\Phi \cup \bigcup_J \Psi_J$ est bien une base de W_h .

Remarque 3.2.2 (numérotation) Dans la suite, on considérera la numérotation des fonctions de base suivante : en premier viennent les fonctions de la base Φ , puis les bases Ψ_J , rangées par J croissant. Ce choix n'est pas le plus optimal pour une implémentation, mais il a l'avantage simplifier les notations.

Définition 3.2.2 (décomposition des matrices) On définit les matrices de masse et de rigidité de la manière suivante :

$$\mathbb{M} = \begin{pmatrix} \mathbb{M}^{H,H} & \mathbb{M}^{H,b} \\ \mathbb{M}^{b,H} & \mathbb{M}^{b,b} \end{pmatrix}, \quad \mathbb{K} = \begin{pmatrix} \mathbb{K}^{H,H} & \mathbb{K}^{H,b} \\ \mathbb{K}^{b,H} & \mathbb{K}^{b,b} \end{pmatrix}, \quad \begin{cases} \mathbb{M}_{IJ}^{H,H} = m(\phi_J, \phi_I) & \mathbb{K}_{IJ}^{H,H} = a(\phi_J, \phi_I) \\ \mathbb{M}_{(J,j),I}^{H,b} = m(\phi_I, \psi_{J,j}) & \mathbb{K}_{(J,j),I}^{H,b} = a(\phi_I, \psi_{J,j}) \\ \mathbb{M}^{b,H} = (\mathbb{M}^{H,b})^t & \mathbb{K}^{b,H} = (\mathbb{K}^{H,b})^t \\ \mathbb{M}_{(I,i),(J,j)}^{b,b} = m(\psi_{J,j}, \psi_{I,i}) & \mathbb{K}_{(I,i),(J,j)}^{b,b} = a(\psi_{J,j}, \psi_{I,i}) \end{cases}$$

On désignera les matrices $\bullet^{H,H}$ par le nom de **standard**, les matrices $\bullet^{H,b}$ par **matrices de couplage**, et les matrices $\bullet^{b,b}$ par **matrice bulle-bulle**.

Définition 3.2.3 (schéma IMEX complètement discréteisé) Soient $0 < T < \infty$ et $\Delta t > 0$. Notons $U_h^n = (U_H^n, U_{\mathcal{B}}^n)^t$ le vecteur des coefficients de u_h^n dans la base de W_h . La suite $(U_h^n)_{n \Delta t \in [0, T]}$

est solution du problème

$$\begin{cases} \mathbb{M} \frac{U_h^{n+1} - 2U_h^n + U_h^{n-1}}{\Delta t^2} + (1-\theta) \begin{pmatrix} \mathbb{K}^{H,H} & \mathbb{K}^{H,b} \\ \mathbb{K}^{b,H} & 0 \end{pmatrix} U_h^n + \theta \begin{pmatrix} 0 & 0 \\ 0 & \mathbb{K}^{b,b} \end{pmatrix} \frac{U_h^{n+1} + U_h^{n-1}}{2} = 0 \\ U_h^0 = \Pi_h u_0, \quad U_h^1 = \Delta t \Pi_h u_1 \end{cases} \quad (3.2)$$

Structure des matrices

Donnons quelques précisions sur la disposition des termes non nuls dans les matrices \mathbb{M} et \mathbb{K} .

Proposition 3.2.1 *Désignons indistinctement les produits scalaires $m(\cdot, \cdot)$ et $a(\cdot, \cdot)$ par (\cdot, \cdot) . La base de W_h possède les propriétés d'orthogonalité suivantes :*

- *Les bases locales Ψ_J sont orthogonales entre elles :*

$$(v_J, v_L) = 0 \text{ pour } v_J, v_L \in (\Psi_J, \Psi_L), \quad J \neq L$$

- *Toute fonction φ_I est orthogonale aux espaces $\mathcal{B}(K_J)$, $J \notin \{I-1, I\}$.*

Démonstration

Ces propriétés découlent des supports des fonctions de base. Par construction, deux fonctions bulles d'espaces différents ont des supports d'intérieur disjoints : elles sont donc orthogonales pour (\cdot, \cdot) . De même, une fonction de base éléments finis a pour support $\overline{K_{I-1} \cup K_I}$, ce qui la rend orthogonale aux espaces bulle associés à d'autres arêtes.

□

En conséquence, les matrices bulle-bulle sont diagonales par blocs, et les matrices de couplages sont creuses. Les matrices standards ont les structures associées au schéma non enrichi : en particulier, on pourra s'aventurer à utiliser une condensation de masse pour diagonaliser $\mathbb{M}^{H,H}$. On peut se faire une idée plus visuelle des agencements à l'aide des schémas suivants :

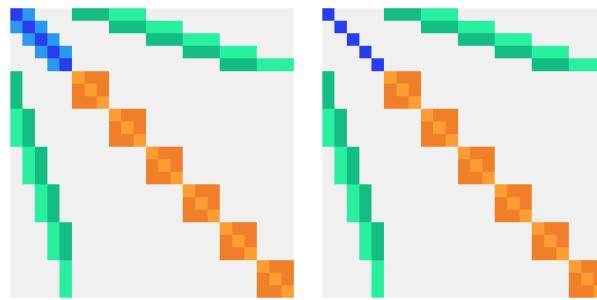


FIGURE 3.1 – termes non nuls des matrices de masse (sans puis avec condensation)

Les termes en bleu représentent $\mathbb{M}^{H,H}$. Les blocs oranges sont les matrices de masse bulle-bulle. Les termes verts représentent les couplages $\mathbb{M}^{H,b}$ et $\mathbb{M}^{b,H}$.

3.2.2 Énergie V

Définition 3.2.4 Les énergies semi-discrétisées associées au schéma (3.1) sont données par

$$\begin{aligned}\mathcal{K}^{n+1/2} &= \frac{1}{2}m\left(\frac{u_h^{n+1}-u_h^n}{\Delta t}, \frac{u_h^{n+1}-u_h^n}{\Delta t}\right) && \text{l'énergie cinétique} \\ \mathcal{P}^{n+1/2} &= \frac{1}{2}a(u_h^{n+1}, u_h^n) + \theta \frac{\Delta t^2}{4}a\left(\frac{u_b^{n+1}-u_b^n}{\Delta t}, \frac{u_b^{n+1}-u_b^n}{\Delta t}\right) && \text{l'énergie potentielle} \\ \mathcal{E}^{n+1/2} &= \mathcal{K}^{n+1/2} + \mathcal{P}^{n+1/2} && \text{l'énergie totale}\end{aligned}$$

Remarque 3.2.3 On remarque que le θ -schéma ajoute un terme par rapport aux énergies correspondantes du schéma non enrichi.

Proposition 3.2.2 L'énergie totale se conserve.

Démonstration

On évalue le schéma en $v_h = \frac{u_h^{n+1}-u_h^{n-1}}{2\Delta t}$ la vitesse centrée. La décomposition du terme de masse est identique à celle utilisée dans la preuve (2.2.2). Le terme de rigidité devient

$$\begin{aligned}&a\left(u_h^n, \frac{u_h^{n+1}-u_h^{n-1}}{2\Delta t}\right) + \theta a\left(\frac{u_b^{n+1}-2u_b^n+u_b^{n-1}}{2}, \frac{u_b^{n+1}-u_b^{n-1}}{2\Delta t}\right) \\ &= \frac{1}{2\Delta t}(a(u_h^n, u_h^{n+1}) - a(u_h^n, u_h^{n-1})) + \theta \frac{\Delta t^2}{2}\left(a\left(\frac{u_b^{n+1}-u_b^n}{\Delta t}, \frac{u_b^{n+1}-u_b^n}{\Delta t}\right) - a\left(\frac{u_b^n-u_b^{n-1}}{\Delta t}, \frac{u_b^n-u_b^{n-1}}{\Delta t}\right)\right) \\ &= \frac{\mathcal{P}^{n+1/2} - \mathcal{P}^{n-1/2}}{\Delta t}\end{aligned}$$

ce qui nous donne le résultat attendu. □

3.2.3 Condition CFL

L'intérêt de la décomposition standard-bulle réside dans les potentielles améliorations de la condition de Courant-Friedrich-Lowy. Cette section a pour but d'établir le résultat encourageant suivant :

THÉORÈME 3.2.1 (STABILITÉ) Le θ -schéma IMEX (3.1) est stable sous la condition CFL

$$\Delta t < \overline{\Delta t}, \quad \overline{\Delta t} = \left(\sup_{v_h \in W_h} \frac{a(v_H, v_H) + (1-\theta)a(v_b, v_b)}{m(v_h, v_h)} \right)^{-1/2}$$

Démonstration

Utilisons les identités

$$\begin{aligned} \frac{4}{\Delta t^2} a(u_h^{n+1}, u_h^n) &= a\left(\frac{u_h^{n+1} + u_h^n}{\Delta t}, \frac{u_h^{n+1} + u_h^n}{\Delta t}\right) - a\left(\frac{u_h^{n+1} - u_h^n}{\Delta t}, \frac{u_h^{n+1} - u_h^n}{\Delta t}\right) \\ a\left(\frac{u_h^{n+1} - u_h^n}{\Delta t}, \frac{u_h^{n+1} - u_h^n}{\Delta t}\right) &\leq 2 \left(a\left(\frac{u_H^{n+1} - u_H^n}{\Delta t}, \frac{u_H^{n+1} - u_H^n}{\Delta t}\right) + a\left(\frac{u_b^{n+1} - u_b^n}{\Delta t}, \frac{u_b^{n+1} - u_b^n}{\Delta t}\right) \right) \end{aligned}$$

pour minorer l'énergie potentielle :

$$\begin{aligned} \mathcal{P}^{n+1/2} &= \frac{1}{2} a(u_h^{n+1}, u_h^n) + \theta \frac{\Delta t^2}{4} a\left(\frac{u_b^{n+1} - u_b^n}{\Delta t}, \frac{u_b^{n+1} - u_b^n}{\Delta t}\right) \\ &\geq -\frac{\Delta t^2}{8} a\left(\frac{u_h^{n+1} - u_h^n}{\Delta t}, \frac{u_h^{n+1} - u_h^n}{\Delta t}\right) + \theta \frac{\Delta t^2}{4} a\left(\frac{u_b^{n+1} - u_b^n}{\Delta t}, \frac{u_b^{n+1} - u_b^n}{\Delta t}\right) \\ &\geq -\frac{\Delta t^2}{4} a\left(\frac{u_H^{n+1} - u_H^n}{\Delta t}, \frac{u_H^{n+1} - u_H^n}{\Delta t}\right) - (1-\theta) \frac{\Delta t^2}{4} a\left(\frac{u_b^{n+1} - u_b^n}{\Delta t}, \frac{u_b^{n+1} - u_b^n}{\Delta t}\right) \end{aligned}$$

L'énergie totale est alors minorée par

$$\mathcal{E}^{n+1/2} \geq \mu \left(\frac{u_h^{n+1} - u_h^n}{\Delta t}, \frac{u_h^{n+1} - u_h^n}{\Delta t} \right), \quad \mu(u, v) = \frac{1}{2} m(u, v) - \frac{\Delta t^2}{4} (a(u_H, v_H) + (1-\theta) a(u_b, v_b))$$

Déterminons sous quelle condition μ est définie positive : on demande à ce que pour $v \in W_h \setminus \{0\}$ arbitraire,

$$\begin{aligned} \mu(v, v) &= \frac{1}{2} m(v, v) - \frac{\Delta t^2}{4} (a(v_H, v_H) + (1-\theta) a(v_b, v_b)) \geq 0 \\ \implies \frac{m(v, v)}{a(v_H, v_H) + (1-\theta) a(v_b, v_b)} &\geq \frac{\Delta t^2}{4} \end{aligned}$$

ce qui nous donne la condition

$$\Delta t \leq \overline{\Delta t}, \quad \overline{\Delta t} = \left(\sup_{v_h \in W_h} \frac{a(v_H, v_H) + (1-\theta) a(v_b, v_b)}{m(v_h, v_h)} \right)^{-1/2}$$

Le lemme (2.3.1) nous assure alors que le schéma IMEX est stable en norme μ , et le lemme (2.3.2) nous donne l'équivalence entre la norme induite par μ et la norme $L^2(\Omega)$. □

Remarque 3.2.4 Pour $\theta = 1$, la condition CFL s'allège d'un terme dépendant des bulles. Cependant, le lecteur attentif aura remarqué que la microstructure est toujours présente dans le terme $m(v_h, v_h)$. La prochaine étape de notre travail va être de supprimer cette influence.

3.3 Orthogonalisation

L'idée de cette section est d'imposer une propriété supplémentaire à l'espace bulle, qui permette d'annuler (ou de négliger) les termes de couplage $\mathbb{M}^{H,b}$ entre fonctions d'échelles différentes. En d'autres termes, nous cherchons à rendre \mathcal{B} orthogonal à V_H pour le produit scalaire $m(\cdot, \cdot)$.

3.3.1 Transformation de l'espace bulle

Dans cette section, K désignera une arête du maillage éléments finis, et $\Phi = \{\varphi_1, \varphi_2\}$ une base de l'espace des polynômes du premier degré. On pourra par exemple prendre la restriction à K des deux fonctions chapeau non nulles sur K .

Définition 3.3.1 (famille orthogonalisée contre Φ) Une famille de fonctions Π sera dite **orthogonalisée contre Φ** (ou orthogonalisée) lorsqu'elle vérifie

$$m(\varphi, \psi) = 0 \quad \forall \varphi \in \Phi, \forall \psi \in \Pi$$

On notera qu'une famille orthogonalisée n'a aucune raison d'être orthogonale au sens usuel.

Définition 3.3.2 (espaces orthogonalisés) On notera par la suite :

- $\mathcal{B}^o(K)$ un espace bulle local engendré par une base orthogonalisée,
- \mathcal{B}^o l'espace bulle global orthogonalisé,
- W_h^o l'espace $V_H \oplus \mathcal{B}^o$.

Notre objectif est de construire des bases bulles orthogonalisées contre la base Φ . Assurons-nous dans un premier temps qu'il est possible de trouver une base bulle orthogonalisée dans un espace raisonnablement dimensionné.

Proposition 3.3.1 Soit Ψ une famille libre de $N_B + 2$ fonctions bulles. Alors il existe une matrice $\Theta \in \mathbb{M}_{N_B+2, N_B}$ de rang plein tel que

$$\Pi := \Psi \Theta$$

soit une famille libre et orthogonale à Φ .

Démonstration

Soit $V_\Psi = Vect(\Psi)$ un espace vectoriel de dimension $N_B + 2$. L'espace $Vect(\Phi)$ est de dimension 2, donc

$$\dim(V_\Psi \cap Vect(\Phi)) \leq 2 \implies \dim(V_\Psi \setminus (V_\Psi \cap Vect(\Phi))) \geq N_B + 2 - 2 = N_B$$

$Vect(\Psi)$ contient donc un espace de dimension au moins N_B : on peut en prendre une base Π et poser Θ la matrice dont la colonne i contient les coefficients de π_i dans la base Ψ . Π étant une base, Θ est de rang au moins N_B .

□

Dans la pratique, nous avons déterminé une matrice Θ de la manière suivante :

Algorithme 3.3.1 (orthogonalisation) Soient trois fonctions $\psi_i, i \in \llbracket 1, 3 \rrbracket$. On forme les deux vecteurs de \mathbb{R}^3 donnés par $V_j = (m(\varphi_j, \psi_i))_{i=1}^3$ pour $j \in \{1, 2\}$. Alors $W := V_1 \wedge V_2$ est

tel que la fonction $\pi = \sum_{i=1}^3 W_i \psi_i$ soit orthogonale à Φ . En effet,

$$m(\varphi_j, \pi) = \sum_{i=1}^3 W_i m(\varphi_j, \psi_i) = W \cdot V_j = 0$$

On calcule alors N_B fonctions π à partir des triplets $\{\psi_i, \psi_{i+1}, \psi_{i+2}\}_{i \in [1, N_B]}$, et l'on remplit la matrice Θ à l'aide des vecteurs W correspondants : les seuls termes non nuls de Θ seront donc ses trois diagonales. On aura pris soin d'ajuster la norme de W pour normaliser la nouvelle base Π .

Dans le cas général, la matrice Θ fournie par cet algorithme n'est pas forcément de rang plein : les vecteurs V_1 et V_2 peuvent être colinéaires, produisant une fonction π nulle. On a cependant la garantie suivante :

Proposition 3.3.2 *Si la numérotation des fonctions de base respecte la propriété*

$$\forall i \in [1, N_B] \quad \det \begin{pmatrix} m(\varphi_1, \psi_{i+1}) & m(\varphi_2, \psi_{i+1}) \\ m(\varphi_1, \psi_{i+2}) & m(\varphi_2, \psi_{i+2}) \end{pmatrix} \neq 0$$

alors le rang de Θ est plein.

Démonstration

Cette condition nous assure que la première coordonnée du vecteur W associé au triplet $\{\psi_i, \psi_{i+1}, \psi_{i+2}\}$ est non nulle : la sous-matrice carrée formée des N_B premières lignes de Θ sera alors triangulaire inférieure à diagonale non nulle, donc inversible. □

Cette condition se vérifie « manuellement » pour les cas particuliers des fonctions bulles micro-éléments finis, et pour les fonctions spectrales dans le cas homogène. Les bases bulles spectrales inhomogènes sont plus compliquées à étudier : pour les besoins pratiques de l'étude, nous nous sommes risqués à employer cet algorithme (qui s'insère très bien dans un code où les matrices de couplage sont déjà calculées, fournissant V_1 et V_2) sans que le cas problématique se produise avant la fin de la période de stage. On pourra visualiser deux cas de bases bulles sur la figure 3.2.

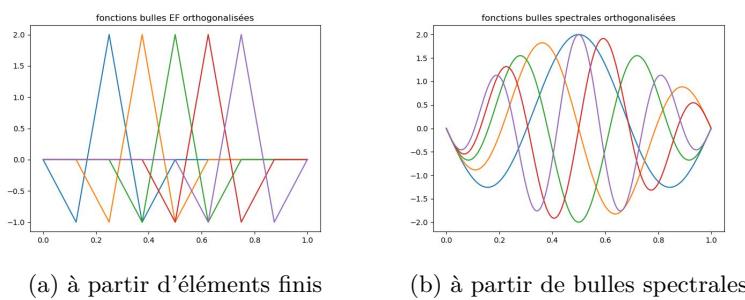


FIGURE 3.2 – Exemples de bases bulles orthogonalisées

3.3.2 Condition CFL

THÉORÈME 3.3.1 (STABILITÉ) Supposons que chacun des espaces $\mathcal{B}(K_J)$ soit orthogonalisé. Le θ -schéma IMEX (3.1) est alors stable sous la condition CFL

$$\Delta t < \overline{\Delta t}, \quad \overline{\Delta t} = \left(\sup_{v_H \in V_H} \frac{a(v_H, v_H) + (1 - \theta) a(v_b, v_b)}{m(v_H, v_H)} \right)^{-1/2}$$

Démonstration

Développons l'énergie cinétique associée au schéma : on note $v_h^{n+1/2} = \frac{u_H^{n+1} - u_H^n}{\Delta t}$ la vitesse décentrée en avant. On a

$$\mathcal{K}^{n+1/2} = \frac{1}{2} m(v_h^{n+1/2}, v_h^{n+1/2}) = \frac{1}{2} m(v_H^{n+1/2}, v_H^{n+1/2}) + 0 + \frac{1}{2} m(v_b^{n+1/2}, v_b^{n+1/2}) \geq \frac{1}{2} m(v_H^{n+1/2}, v_H^{n+1/2})$$

où les termes croisés s'annulent par orthogonalité. En utilisant les résultats de la preuve (3.2.3), on peut minorer l'énergie totale par la forme bilinéaire symétrique $\mu(u_h, v_h) = \frac{1}{2} m(u_H, v_H) - \frac{\Delta t^2}{4} (a(u_H, v_H) + (1 - \theta) a(u_b, v_b))$. Le reste de la démonstration est identique au cas (3.2.3). □

Remarque 3.3.1 Pour $\theta = 1$, la condition CFL n'a plus de termes liés aux bulles : notre objectif est donc atteint.

3.3.2.1 Matrice de masse

Les termes de couplage de la matrice de masse sont rendus nuls par l'orthogonalisation. À l'aide de la méthode de condensation de masse, on dispose donc de l'excellente propriété d'avoir une matrice diagonale par bloc à inverser durant la propagation du schéma. On pourra notamment paralléliser le code en traitant les sous-systèmes associés aux espaces bulles indépendamment.

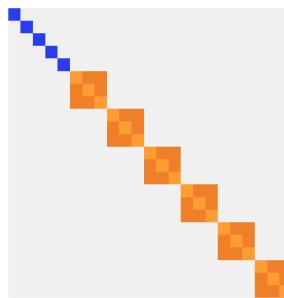


FIGURE 3.3 – termes non nuls de la matrice de masse orthogonalisée

3.3.3 Représentation dans l'espace orthogonalisé

Soit K une arête. Le cadre naturel du contenu de cette section est l'espace $H^1(K)$, sans contrainte de nullité aux bords : on se permettra donc de rétropédaler¹ pour s'y placer.

Notons $\mathbb{P}_1(K)$ l'espace des polynômes d'ordre 1 sur K . C'est un sous-espace vectoriel de dimension finie de $H^1(K)$: il est donc fermé, et admet un complémentaire orthogonal dans $H^1(K)$, que nous noterons $\mathbb{P}_1^\perp(K)$.

THÉORÈME 3.3.2 (ORTHOGONALITÉ DE LA LIMITÉ) Soit $(\mathcal{B}_n)_n$ une suite de sous-espaces vectoriels de dimension finie dont la limite est dense dans $H^1(K)$, et $(\mathcal{B}_n^\circ)_n$ la suite des complémentaires orthogonaux de $\mathbb{P}_1(K)$ dans \mathcal{B}_n : $\mathcal{B}_n^\circ = \mathcal{B}^n \cap \mathbb{P}_1^\perp(K)$. Alors la limite de $(\mathcal{B}_n^\circ)_n$ est dense dans $\mathbb{P}_1^\perp(K)$.

Démonstration

Soit $u \in H^1(K)$. Par hypothèse, il existe $(u_n)_{n \in \mathbb{N}} \subset \mathcal{B}_n$ une suite de limite u . Chaque espace \mathcal{B}_n est de dimension finie, donc est un espace de Hilbert pour le produit scalaire hérité de H^1 : tout u_n se décompose alors en $u_n = \Pi_1 u_n + \Pi_{\mathcal{B}_n^\circ} u_n$. Soit $\varepsilon > 0$: pour n suffisamment grand, on a

$$\|u - u_n\|_{1,K}^2 = \|\Pi_1 u - \Pi_1 u_n\|_{1,K}^2 + \|(u - \Pi_1 u) - \Pi_{\mathcal{B}_n^\circ} u_n\|_{1,K}^2 \leq \varepsilon$$

et $\Pi_{\mathcal{B}_n^\circ} u_n$ tend vers $u - \Pi_1 u$. □

La conséquence de ce résultat est que toute base orthogonalisée contre \mathbb{P}_1 ne pourra correctement approcher que des fonctions elles-mêmes orthogonales à \mathbb{P}_1 . La figure (3.4) présente l'exemple des éléments finis de $H^1(K)$, dont la limite est bien dense dans $H^1(K)$ au sens de la norme $\|\cdot\|_{1,K}$.

Revenons maintenant à $H_0^1(\Omega)$.

Proposition 3.3.3 (fermeture de H_0^1) Soit $(u_n)_{n \in \mathbb{N}^*} \subset H_0^1(K)$. S'il existe $u \in H^1(K) \setminus H_0^1(K)$ telle que $u_n \xrightarrow{n} u$ au sens de la norme $\|\cdot\|_{L^2(K)}$, alors $u_n \not\xrightarrow{n} u$ au sens de la norme $H^1(K)$.

Démonstration

On s'appuie sur les éléments suivants :

- $H_0^1(K)$ est défini comme l'adhérence dans $H^1(K)$ de $\mathcal{C}_c^\infty(K) = \mathcal{D}(K)$, l'espace des fonctions test. C'est donc un fermé pour la topologie induite par la norme $H^1(K)$.
- L'espace $\mathcal{D}(K)$ est dense dans $L^2(K)$, donc dans $H^1(K)$, pour la norme $\|\cdot\|_{L^2(K)}$.

Il est ainsi tout à fait possible qu'une suite de $H_0^1(K)$ ait pour limite un élément de $H^1(K) \setminus H_0^1(K)$ au sens de $\|\cdot\|_{L^2(K)}$: mais la convergence en norme $H^1(K)$ impliquerait $u \in H_0^1(K)$, donc, par contraposée, $u_n \not\xrightarrow{n} u$ dans l'espace $H_0^1(K)$. □

1. Terme hautement technique.



FIGURE 3.4 – représentation des fonctions et étude d’erreur

À gauche, on a approché une fonction f (en bleu) par sa projection sur un espace éléments finis de $H^1(K)$ (en vert). Pour Π_1 le projecteur sur l'espace $\mathbb{P}_1(K)$, on a également calculé $f - \Pi_1 f$ (en orange) et la projection de f sur l'espace orthogonalisé. On constate que les éléments finis orthogonalisés approchent bien la composante de f orthogonale à $\mathbb{P}_1(K)$. Pour s'en convaincre, l'image de droite présente l'erreur que la projection classique commet sur f (en bleu) quand le nombre de fonctions augmente, ainsi que celle que la projection orthogonalisée commet sur $f - \Pi_1 f$ (en violet).

Revenons au cadre de notre étude multi-échelle. Pour les deux types d'espaces bulles considérés, les deux points précédents impliquent le suivant :

Proposition 3.3.4 (convergence du projecteur orthogonalisé) Soit $u \in V$. On définit Π_h^o le projecteur sur W_h^o de la manière suivante :

$$\Pi_h^o u = \mathcal{I}_H u + \Pi_{\mathcal{B}^o} (u - \mathcal{I}_H u)$$

Alors pour H fixé,

$$\Pi_h^o u \xrightarrow[N_B \rightarrow \infty]{} u \Leftrightarrow \forall K \text{ arête du maillage}, \quad \Pi_1 (u - \mathcal{I}_H u)|_K = 0$$

où la convergence est à prendre au sens de la norme $H^1(\Omega)$.

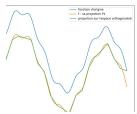
Démonstration

D'après les résultats précédents, la convergence sur Ω impose la convergence locale sur chaque arête. D'après (3.3.2), $\Pi_{\mathcal{B}^o(K)} (u - \mathcal{I}_H u)|_K$ converge vers $(u - \mathcal{I}_H u)|_K$ si et seulement si $(u - \mathcal{I}_H u)|_K$ est orthogonale aux polynômes du premier degré.

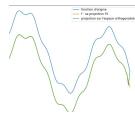
□

Cette condition est très restrictive : *stricto sensu*, elle n'est vérifiée pour H arbitraire que par

la fonction nulle. Dans le cadre de ce stage, nous avons pu constater à quel point la représentation d'une fonction arbitraire dans l'espace orthogonalisé est insatisfaisante.



(a) 20 fonctions éléments finis orthogonalisées



(b) 100 fonctions éléments finis orthogonalisées

Soit une fonction f (en bleu). On observe ici deux phénomènes : premièrement, la projection sur l'espace orthogonalisé (en vert) approche $f - \Pi_1 f$ (en jaune). Deuxièmement, $f - \Pi_1 f$ n'est pas dans H_0^1 : la convergence ne se fait qu'en norme L^2 , ce qui se voit aux oscillations sur les bords dans la figure de droite.

En particulier, la projection de u_1 lors de l'initialisation du schéma ne sera pas améliorée par les bulles. Ceci constitue un frein majeur à l'utilisation des espaces orthogonalisés. On peut toutefois donner deux pistes à suivre pour de futurs développements :

- se placer dans un cadre qui garantisse que la solution u vérifie « à peu près » la condition de convergence, pour des échelles de maillage macro satisfaisantes. On peut penser à des variations très faibles de κ . Il est cependant difficile de construire des fonctions satisfaisant la condition pour des échelles de maillage différentes, et *a fortiori* de trouver des applications intéressantes.
- substituer la projection orthogonale à l'interpolation dans la méthode macroscopique. Les bulles auraient ainsi directement à approcher $u - \Pi_1 u$. Deux problèmes surgissent immédiatement : d'une part, les espaces bulles choisis doivent être dense dans $H^1(K)$ pour les normes désirées. Les bulles deviennent potentiellement non nulles aux interfaces des arêtes, ce qui entraîne d'éventuels flux à prendre en compte dans le schéma numérique. D'autre part, l'approximation fournie par la méthode macroscopique n'a plus de raison d'être continue sur Ω : nous sortons donc des méthodes de Galerkin. Pour citer le colonel Diaz [1937], tout est à recommencer.

3.3.4 Un mot sur la quasi-orthogonalisation

L'idée de cette section est d'imposer aux bases bulles d'avoir des produits scalaires contre Φ qui soient négligeables. L'espoir est de pouvoir négliger les termes de couplage dans le calcul de la CFL, ce qui conserverait l'indépendance par rapport à l'échelle de la micro-structure, tout en permettant aux fonctions bulles de ne pas être parfaitement orthogonales à \mathbb{P}_1 , et d'ainsi améliorer la représentation dans \mathcal{B}^o .

Notons K une arête du maillage.

Définition 3.3.3 (espaces) On note

- Φ une base des polynômes du premier degré sur K .
- $\Psi = \{\psi_1, \dots, \psi_n\}$ une base bulle non orthogonalisée sur K .
- $\mathcal{Z} \subset H^1(K)$ de dimension finie tel que $\Phi \subset \mathcal{Z}$ et $\Psi \subset \mathcal{Z}$. On se donne $Z = \{z_1, \dots, z_M\}$ une base de \mathcal{Z} .

Remarque 3.3.2 La condition imposée à \mathcal{Z} est purement pratique : cet espace servira à représenter les fonctions quasi-orthogonalisées, il doit donc correctement représenter les autres données du problème. Dans notre cas, \mathcal{Z} est un espace « nano-éléments finis ».

On cherche alors la solution au problème suivant :

Définition 3.3.4 (problème de minimisation) Soit $\delta \in [0, 1]$. On cherche $\Pi(\delta) = \{\pi_1, \dots, \pi_N\}$ satisfaisant

$$\begin{aligned} \pi_i &= \underset{u \in \mathcal{Z}}{\operatorname{argmin}} \|u - \psi_i\|_{L^2(\Omega)} \quad \forall i \in \llbracket 1, N \rrbracket \\ m(u, \varphi_j) &= \delta m(\psi_i, \varphi_j) \quad \forall \varphi_j \in \Phi \end{aligned} \tag{3.3}$$

Lemme 3.3.1 Le problème de minimisation admet une solution.

Démonstration

Notons $U \in \mathbb{R}^M$ les coefficients de l'élément $u \in \mathcal{Z}$ dans la base Z . La fonction $U \in \mathbb{R}^M \rightarrow \|\sum_k U_k z_k - \psi_i\|$ est continue et convexe sur \mathbb{R}^M . D'autre part, la condition (3.3) est satisfaite sur un hyperplan de \mathbb{R}^M , donc un fermé : le minimum existe. □

Définition 3.3.5 (termes matriciels) On note

- $\mathcal{M} \in \mathbb{M}_{M,M}$ la matrice de masse de Z , telle que $\mathcal{M}_{i,j} = m(z_i, z_j)$,
- $\mathcal{A} \in \mathbb{M}_{2,M}$ la matrice de couplage entre Φ et Z , telle que $\mathcal{A}_{i,k} = m(\varphi_i, z_k)$,
- $P^i \in \mathbb{R}^M$ le vecteur des coefficients de ψ^i dans Z .

Proposition 3.3.5 (résolution) La solution est donnée par $\pi_i = \sum_{k=1}^M U_k^i z_k$, où

$$U^i = P^i - (1 - \delta) \mathcal{M}^{-1} \mathcal{A}^t \left(\mathcal{A} \mathcal{M}^{-1} \mathcal{A}^t \right)^{-1} \mathcal{A} P^i$$

Démonstration

Soit $\psi \in \Psi$. La condition imposée sur l'espace \mathcal{Z} permet de décomposer exactement $\psi = \sum_{k=1}^M P_k z_k$. Pour $u = \sum_{k=1}^M U_k z_k$, on a ainsi

$$\|u - \psi\|_{L^2(\Omega)} = (U - P)^t \mathcal{M} (U - P) = U^t \mathcal{M} U - 2U^t \mathcal{M} P + P^t \mathcal{M} P$$

Le dernier terme étant constant, on minimisera la fonction $U^t \mathcal{M} U - 2U^t \mathcal{M} P$ sur $U \in \mathbb{R}^M$. La contrainte $m(u, \varphi_j) = \delta m(\psi_i, \varphi_j) \quad \forall \varphi_j \in \Phi$ s'écrit $AU = \delta AP$. Soit $\lambda \in \mathbb{R}^2$. On peut former le Lagrangien $\mathcal{L}(U, \lambda)$ du problème :

$$\mathcal{L}(U, \lambda) = U^t \mathcal{M} U - 2U^t \mathcal{M} P + \lambda \cdot A(U - \delta P)$$

D'après le lemme (3.3.1), le problème admet une solution U^* . En ce point, les conditions de Karush-Kuhn-Tucker nous donnent l'existence de $\lambda^* \in \mathbb{R}^2$ tel que $\nabla \mathcal{L}(U^*, \lambda^*) = 0$. Si $\lambda^* = 0$, la convexité de la fonction coût impose $U = P$: la base obtenue serait exactement égale à Ψ , qui n'est pas orthogonalisée. Donc $\lambda^* \neq 0$, et on a

$$\begin{cases} \nabla_U \mathcal{L}(U^*, \lambda^*) = 2U^t \mathcal{M} - 2P^t \mathcal{M} + \lambda^{*t} A = 0 & (1) \\ \nabla_\lambda \mathcal{L}(U^*, \lambda^*) = (U - \delta P)^t A^t = 0 & (2) \end{cases}$$

Donc, avec \mathcal{M} symétrique définie positive,

$$(1) \implies U = P - \frac{1}{2} \mathcal{M}^{-1} A^t \lambda^* \quad \text{et} \quad (2) \implies A \left(P - \frac{1}{2} \mathcal{M}^{-1} A^t \lambda^* - \delta P \right) = 0 \implies A \mathcal{M}^{-1} A^t \lambda^* = 2(1-\delta) AP$$

La matrice $A \mathcal{M}^{-1} A^t \in \mathbb{M}_{2,2}$ est de rang plein si A l'est : on le vérifie facilement pour les bases bulles considérées, mais le lecteur désireux d'adapter cette idée y prendra garde. On conclut par

$$\begin{aligned} \lambda^* &= 2(1-\delta) (A \mathcal{M}^{-1} A^t)^{-1} P \\ U &= P - (1-\delta) \mathcal{M}^{-1} A^t (A \mathcal{M}^{-1} A^t)^{-1} AP \end{aligned}$$

□

Cette expression explicite nous permet d'une part d'implémenter ce type de base, et d'autre part, de remarquer que la nouvelle base $\Pi(\delta)$ dépend continument de δ . On pourra se faire une idée de l'évolution des bases à l'aide de la figure (3.5).

[évolution base quasi orthogonalisation](#) 

La projection étant également un opérateur continu *en fonction de la base*, on déduit que quand δ évolue de 1 à 0, la projection d'une fonction f sur la base $\Pi(\delta)$ parcourt un chemin continu de sa décomposition d'origine à sa décomposition dans la base complètement orthogonalisée. Ce chemin est illustré sur la figure (3.6).



FIGURE 3.5 – évolution des bases pour $\delta \in \left\{ \frac{3}{4}, \frac{1}{2}, \frac{1}{4}, 0 \right\}$



FIGURE 3.6 – évolution de la projection de $\sin(3\pi \frac{x-a}{b-a}) + \frac{1}{2}$ pour $\delta \in \left\{ \frac{3}{4}, \frac{1}{2}, \frac{1}{4}, 0 \right\}$

évolution fonction quasi orthogonalisation

On sait que la projection sur une base totalement orthogonalisée est insuffisante : la question est donc *à quelle vitesse tend-on vers cette projection*? Toute l'intuition derrière cette méthode est de jouer sur la différence entre la réduction des moments de $\Pi(\delta)$, et la perte de qualité de la projection.

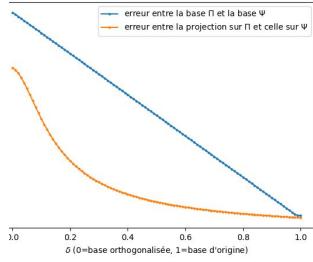


FIGURE 3.7 – Comparaison entre l'évolution de $\Pi(\delta)$ et celle de la projection sur $\Pi(\delta)$

On a affiché en bleu l'erreur relative que Π commet sur Ψ , donnée par $\|\Pi - \Psi\| / \|\Psi\|$. En orange, on a affiché l'erreur relative que la projection $P_\Pi(f)$ commet sur $P_\Psi(f)$, donnée par $\|P_\Pi(f) - P_\Psi(f)\| / \|P_\Psi(f)\|$.

L'interprétation est plutôt favorable à la quasi-orthogonalisation : pour δ partant de 1 vers 0, l'erreur entre les projections reste faible, et n'augmente qu'à la fin. On paie donc moins cher la transformation des projections que celle des bases. Cependant, c'est loin d'être suffisant pour justifier les coûts de calcul engendrés par les nano-éléments finis et le calcul de la solution de minimisation. Nous avons donc écarté cette piste.

Conclusion sur l'enrichissement

Au terme de cette partie, nous disposons de plusieurs paramètres sur lesquels jouer pour construire des modèles. En particulier, on a retenu :

- le **modèle standard** : éléments finis sans bulles. Par défaut, on préfère utiliser une condensation de masse.
- le **modèle IMEX** : on construit un squelette éléments finis sans condensation de masse, et on enrichit avec des bulles (soit micro éléments finis, soit spectrales). Dans tous les modèles enrichis, on emploie le θ -schéma (3.1) avec $\theta = 1$.
- le **modèle LIMEX** (Lumped IMEX)² : la matrice de masse du squelette éléments finis est rendue diagonale.
- le **modèle LOIMEX** (Lumped Orthogonalized IMEX) : les bulles sont orthogonalisées. Les matrices $\mathbb{M}^{H,H}$ et $\mathbb{M}^{b,b}$ étant découplées, on s'autorise une condensation de masse sur le squelette éléments finis.

Le prochain chapitre est dédié à l'implémentation de ces modèles.

2. L'auteure est bien conscience du barbarisme d'un acronyme franco-anglais : on a tout de même préféré ce choix à l'inprononçable IMEXC.

Chapitre 4

Étude numérique

4.1 Obtention d'une solution de référence

Cette section parcourt les différentes méthodes mises en place pour obtenir une solution permettant de valider le modèle éléments finis multi-échelle. L'idée est de s'appuyer sur la solution explicite (1.6) obtenue par l'étude spectrale, que l'on rappelle ici : pour $u_0 \in \mathcal{D}(\mathcal{A})$ et $u_1 \in H_0^1(\Omega)$,

$$u(x, t) = \sum_{n=1}^{\infty} \left(m(u_0, \omega_n) \cos(\sqrt{\lambda_n}t) + \frac{m(u_1, \omega_n)}{\sqrt{\lambda_n}} \sin(\sqrt{\lambda_n}t) \right) \omega_n(x)$$

où la suite $(\lambda_n, \omega_n)_{n \in \mathbb{N}^*}$ correspond au spectre de l'opérateur \mathcal{A} . Dans notre cas, on prendra $u_0 = 0$, et on tronquera la série à $M < \infty$. Une fois une base $(\lambda_n, \omega_n)_{n \in \llbracket 1, M \rrbracket}$ correctement approchée et les coefficients $m(u_1, \omega_n)$, l'estimation des $u(x_i, t_n)$ se fait par simple évaluation, sans dépendance à $\delta = \Delta t$. L'étape délicate est l'approximation des éléments propres.

4.1.1 Première idée : équation non linéaire

On s'appuie sur les remarques de la section (1.3.3) pour caractériser les éléments propres. On appellera **fonction d'ordre n** la fonction ω_n .

Proposition 4.1.1 (paramétrisation) Notons $\mathbb{S} = \mathbb{R}^N \times [0, \pi]^N \times \mathbb{R}$. On peut trouver une fonction $F : \mathbb{R}^{2N+1} \rightarrow \mathbb{R}^{2N+1}$ telle que les solutions dans \mathbb{S} du système $F(X) = 0$, $X \in \mathbb{R}^{2N+1}$ soient en correspondance bijective avec les éléments propres de \mathcal{A} dans $\mathcal{D}(\mathcal{A})$.

Démonstration

Notons $(\lambda_n, \omega_n)_{n \in \mathbb{N}^*}$ l'ensemble des éléments propres de \mathcal{A} dans $\mathcal{D}(\mathcal{A})$, et choisissons pour représentant de chaque sous-espace propre l'unique fonction (encore notée ω_n) de norme 1 et dont la dérivée en a soit strictement positive (cette dernière condition assurant une distinction entre ω_n et $-\omega_n$). Soit $X = (A_{n,0}, \dots, A_{n,N-1}, C_{n,0}, \dots, C_{n,N-1}, \sqrt{\lambda_n})^t \in \mathbb{R}^{2N+1}$. On note $\omega(X)$ la fonction définie par

$$\omega_{n|\Omega_i}(X)(x) = A_{n,i} \sin\left(\sqrt{\frac{\lambda_n}{\kappa_i}}x + C_{n,i}\right) \quad \forall x \in \Omega_i, \quad \forall i \in \llbracket 0, N-1 \rrbracket$$

Il est évident qu'à X fixé, $\omega(X)$ est unique. Soit une fonction propre ω_n : on se rappelle que sur chaque Ω_i , c'est une sinusoïde (voir section 1.3.3). Il existe donc au moins un $X \in \mathbb{R}^{2N+1}$ paramétrant ω_n : on fixe l'unicité en choisissant $X \in \mathbb{S}$. On notera $X(\omega_n)$ cette paramétrisation. Les éléments de $\mathcal{D}(\mathcal{A})$ satisfont tous les contraintes de nullité au bord (2 contraintes), continuité ($N - 1$ contraintes aux x_i , $i \in [1, N - 1]$), saut de la trace normale ($N - 1$ contraintes). On ajoute la normalité pour arriver à $2N + 1$ conditions, qui se mettent sous la forme $F(X) = 0$, $F : \mathbb{R}^{2N+1} \rightarrow \mathbb{R}^{2N+1}$. On sait que tout vecteur $X(\omega_n)$ est solution du système : montrons que toute solution $X^* \in \mathbb{S}$ paramétrise une fonction propre dans $\mathcal{D}(\mathcal{A})$. La construction de la paramétrisation assure que sur chaque Ω_i , $\omega(X^*)|_{\Omega_i}$ satisfait $-\partial_x(\kappa_i \partial_x \omega|_{\Omega_i}) = \lambda \Omega_i$. La condition de continuité fait que la dérivée au sens des distributions de $\omega(X^*)$ est dans $L^2(\Omega)$: la condition de Snell-Descartes rend la fonction $\kappa \partial_x \omega$ continue, donc $\partial_x(\kappa \partial_x \omega)$ sera également dans $L^2(\Omega)$. La nullité au bord achève de montrer que $\omega(X^*)$ est bien une des fonctions de $\mathcal{D}(\mathcal{A})$ recherchée.

□

Remarque 4.1.1 Dans l'implémentation proposée, les solutions sont prises dans \mathbb{R}^{2N+1} et non dans \mathbb{S} , avec un signe de la dérivée en a quelconque.

On utilise alors l'algorithme suivant :

Algorithme 4.1.1 (Newton-Raphson) Soit X_0 une approximation initiale des paramètres de (λ_n, ω_n) . On note $J_{F,X} \in \mathbb{M}_{2N+1, 2N+1}$ la matrice jacobienne de F au point X . Soit $\varepsilon > 0$: on itère le système

$$X_{k+1} = X_k - J_{F,X_k}^{-1} \times F(X_k)$$

(où la matrice jacobienne n'est surtout pas explicitement inversée) tant que $\|F(X_k)\| \geq \varepsilon$.

Cet algorithme possède l'excellente propriété de converger à l'ordre 2, à condition que l'on soit capable de fournir une estimation initiale X_0 suffisamment proche de la solution. Notre première idée a été de calculer la solution analytique dans le cas $\kappa = \kappa_{\text{mean}}$ constant. Dans ce cas, l'élément (λ_n, ω_n) est donné par

$$\omega_n(x) = \sqrt{\frac{2}{b-a}} \sin\left(n\pi \frac{x-a}{b-a}\right), \quad \lambda_n = \kappa_{\text{mean}} \left(\frac{n\pi}{b-a}\right)^2$$

Ce point de départ est relativement efficace si la variance de κ est faible, seulement :

- Quand la variance de κ augmente, la décomposition d'une fonction u_1 très régulière exige de plus en plus de fonctions propres pour atteindre une précision satisfaisante. Pour fixer les idées, dans le cas homogène, 40 fonctions suffisent pour obtenir une précision de 10^{-14} , là où pour une variance de $0.1 \cdot \kappa_{\text{mean}}$, 600 fonctions donnent péniblement une erreur de $[10^{-6}, 10^{-7}]$. On s'est fixé pour objectif de calculer 600 fonctions pour chaque cas inhomogène.
- L'algorithme de Newton-Raphson est très efficace, mais peu contrôlable : il peut converger vers une autre fonction propre que celle désirée. Les basses fréquences sont à peu près correctement atteintes, mais au-delà de $n = 50$, les résultats ressemblent à un tir de Monte-Carlo. De plus, au fil des exécutions, on a acquis la conviction qu'il existe des bassins de convergence très inégaux selon l'ordre n , et que Newton-Raphson ne permettait pas d'obtenir des bases complètes en temps raisonnable.



Vitesse : convergence en moins d'une dizaine d'itérations pour un critère d'arrêt de 10^{-10} .

Contrôlabilité : aucun moyen de forcer la convergence vers un ordre précis. Besoin d'une 2^e idée.



4.1.2 Deuxième idée : prédition-correction

L'idée directrice de cette étape de recherche est de faciliter la convergence de Newton-Raphson en cheminant par bonds successifs du cas homogène au cas d'intérêt. Précisons.

Définition 4.1.1 (chemin perturbateur) Soit κ la fonction discontinue à laquelle on s'intéresse, et κ_{mean} la moyenne de κ sur Ω . On se donne le paramètre $\delta \in [0, 1]$. On note

$$\sqrt{\kappa_\delta}(x) = (1 - \delta) \sqrt{\kappa_{\text{mean}}} + \delta \sqrt{\kappa}(x)$$

On trace ainsi un chemin entre le problème homogène à $\delta = 0$, dont on connaît la solution, et le problème qui nous intéresse, à $\delta = 1$. À chaque δ fixé correspond une base d'éléments propres (λ_n, ω_n) , dont les paramétrisations $(X_n(\delta))$ satisfont le problème $F_\delta(X) = 0$.

Définition 4.1.2 (intégrale première du système) On se donne la variable $Y = (X^t, \delta)^t \in \mathbb{R}^{2N+2}$. Notons F_δ la fonction définie en 4.1.1, dans laquelle on substitue $\sqrt{\kappa_\delta}$ à $\sqrt{\kappa}$. Posons $\tilde{F} : \mathbb{R}^{2N+2} \rightarrow \mathbb{R}^{2N+1}$ la fonction

$$\tilde{F}(Y) = F_\delta(X)$$

Les solutions $Y^* = (X_n^t(\delta), \delta)^t$ sont exactement le noyau de \tilde{F} . On notera \mathcal{C}_n l'ensemble des solutions accessibles par un chemin continu depuis $(X_n^t(0), 0)^t$. L'algorithme suivant est fortement inspiré des travaux de Karkar [2012]. Toute l'idée est de calculer la tangente à \mathcal{C}_n en un point $X_n(\delta)$ connu, puis de se déplacer dans le sens des δ croissants : on fera ainsi une **prédition** de la paramétrisation de la solution à δ +un petit déplacement. Une seconde étape de **correction** vient ramener la paramétrisation sur \mathcal{C}_n , à l'aide de l'algorithme de Newton.

Algorithme 4.1.2 (prédition-correction) Soit $n \in \mathbb{N}^*$ un ordre, qui sera sous-entendu pour alléger les notations. On se donne $Y_0 = \begin{pmatrix} X^0 & 0 \end{pmatrix}$, où X^0 est la solution analytique homogène. On distingue l'étape d'initialisation des étapes suivantes. La première tangente T_0 est choisie parmi les solutions du système

$$\begin{pmatrix} \mathcal{J}_{\tilde{F}, Y_0} \\ (0 \ \dots \ 0 \ \ 1) \end{pmatrix} T_0 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \in \begin{pmatrix} \mathbb{R}^{2N+1} \\ \mathbb{R} \end{pmatrix}$$

où les $2N + 1$ premières lignes du système donnent le noyau de la jacobienne de \tilde{F} , donc la direction de la ligne de niveau de \tilde{F} passant par $(X^0, 0)$: la dernière ligne est là pour fermer le système, et choisit T_0 pointant vers $\delta > 0$.

À chaque étape, on pose $P_{k+1} = X_k + s_k T_k$ une prédition, où $s_k > 0$ est un pas variable qu'il faudra piloter. On calcule ensuite X_{k+1} comme la solution de l'algorithme de Newton initialisé avec P_{k+1} .

La tangente T_k pour $k > 0$ est donnée par le système

$$\begin{pmatrix} \mathcal{J}_{\tilde{F}, Y_0} \\ T_k^t \end{pmatrix} T_{k+1} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \in \begin{pmatrix} \mathbb{R}^{2N+1} \\ \mathbb{R} \end{pmatrix}$$

où la condition fermant le système assure que l'on se déplace toujours dans la même direction. On normalisera les tangentes après chaque résolution de système linéaire. L'algorithme se termine quand la nouvelle prédition a un $\delta > 1$: on ajuste alors le dernier pas pour correspondre à $\delta = 1$ et on calcule la dernière correction.

Remarque 4.1.2 (pilotage du pas) Après quelques essais, nous avons adopté la méthode suivante : à chaque étape, l'ordre de la fonction paramétrée par X_{k+1} est calculé. S'il correspond au n demandé, le pas est multiplié par $1 + \varepsilon$, $\varepsilon > 0$ petit (de l'ordre de 0.005). S'il ne correspond pas, le pas est divisé par 2 et X_{k+1} recalculé à partir de X_k et T_k . Un deuxième mécanisme régule le pas pour que le nombre d'itérations de la méthode de correction reste dans un intervalle défini (en pratique, $\llbracket 3, 6 \rrbracket$).

On pourra se faire une idée plus intuitive du fonctionnement de l'algorithme avec la figure (4.1).



Contrôlabilité : l'algorithme produira invariablement la fonction propre d'ordre demandé.



Vitesse : plusieurs itérations de Newton par étape, temps de calcul pénalisant. Besoin d'une 3^e idée.

POUR ALLER PLUS LOIN...

La justification complète de cet algorithme voudrait que chaque ensemble \mathcal{C}_n soit une courbe de classe C^1 le long de laquelle l'ordre des solutions est invariant. Pour établir ceci, il faudrait prouver que $\mathcal{J}_{\tilde{F}, Y}$ est de rang maximal au voisinage de chaque Y^* solution de $\tilde{F}(Y) = 0$ les ensembles de niveau de \tilde{F} serait alors des variétés d'ordre 1, et admettraient un espace tangent correctement défini.

Une manière de prouver que $\mathcal{J}_{\tilde{F}, Y^*} = \left(\mathcal{J}_{F_\delta, X}, \frac{\partial \tilde{F}}{\partial \delta} \right) \in \mathbb{M}_{2N+1, 2N+1}$ est de rang plein serait d'obtenir que $\mathcal{J}_{F_\delta, X}$ est inversible aux points X^* visés. Or, en tentant d'arriver à ce résultat, l'auteure a démontré que $\mathcal{J}_{F_\delta, X}$ n'est pas inversible en $X(0)$ une solution homogène, et n'a rien pu obtenir sur les cas inhomogènes.

La question est un peu plus profonde : comment varient les éléments propres de \mathcal{A}_δ quand δ augmente ? En particulier, peut-on obtenir la régularité de cette variation ?

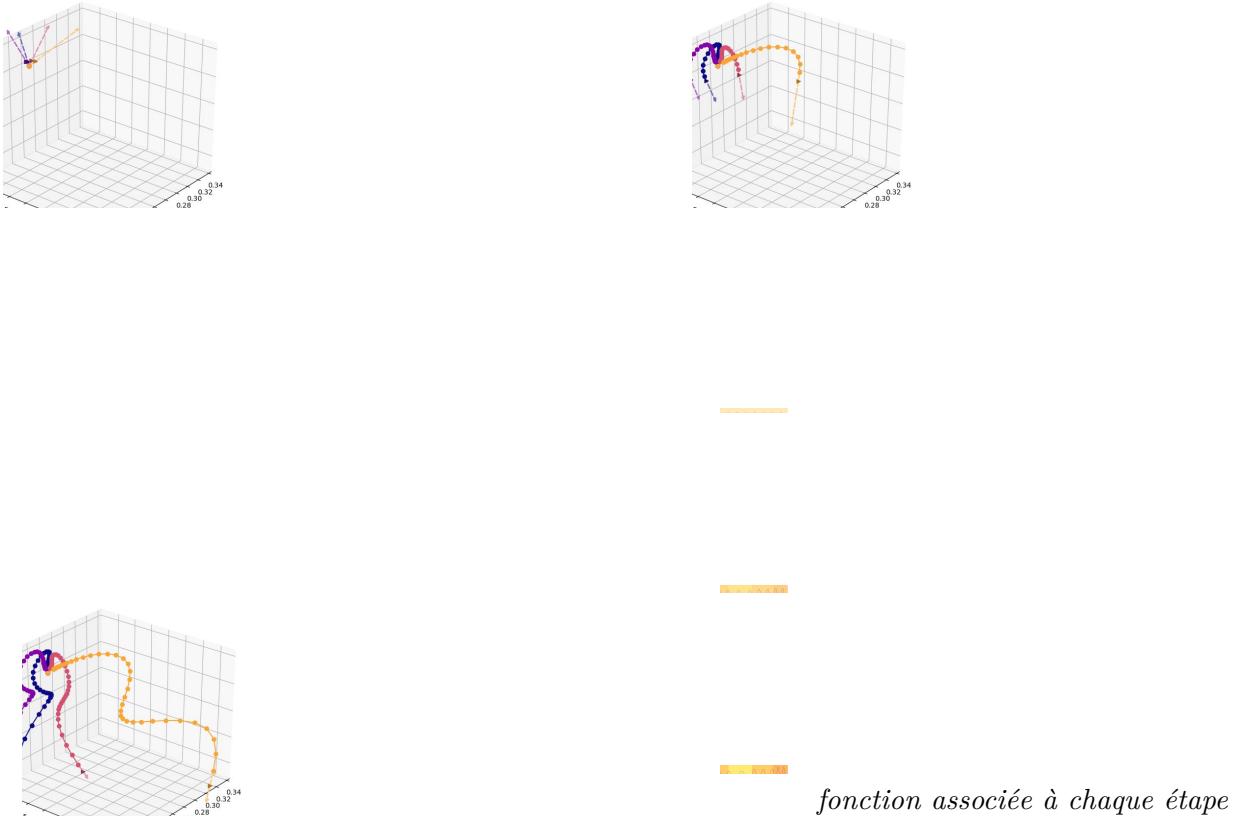


FIGURE 4.1 – Étapes de l'algorithme de prédiction-correction pour $\delta \in \{0, 0.54, 1\}$

On emploie l'algorithme sur un Ω divisé en 4 domaines (Ω_i) sur lesquels κ est constant.

Une fonction $X(\delta)$ est paramétrée par 4 amplitudes, 4 phases et une longueur d'onde. On représente l'évolution de chacun des 4 triplets (A_i, C_i, λ) au fil des itérations par une branche dans l'espace tridimensionnel (amplitude, phase, longueur d'onde). Dans l'ordre, Ω_0 -violet, Ω_1 -bleu, Ω_2 -rose, Ω_3 -jaune.

Les trois premiers graphiques présentent les valeurs successives prises par X_k pour k allant de 0 à respectivement $\{0, 24, 42\}$. On y voit se dessiner la courbe \mathcal{C} qui paramétrise les solutions $X(\delta)$ pour $\delta \in [0, 1]$. Le mouvement le long de cette courbe est guidé par la tangente calculée à chaque itération, représentée ici par un ensemble de 4 flèches pour le dernier point affiché. Le lecteur à la bonne vue aura remarqué un point triangulaire traversé par la tangente : il s'agit de la prédiction faite à l'étape affichée. Pour des raisons de lisibilité, les tangentes sont affichées avec un coefficient multiplicatif permettant de les voir.

La dernière partie de la figure présente la fonction paramétrée par X_k à la dernière étape affichée par chacun des 3 graphes 3D, sur fond de κ_δ .

4.1.3 Troisième idée : résolution numérique

L'idée est de discréteriser le problème spectral (1.5) dans une base éléments finis, puis d'affiner la solution obtenue avec Newton.

Algorithme 4.1.3 (problème spectral discret) Soit V_H un espace éléments finis de type \mathbb{P}_1 sur Ω , muni d'une base Φ de cardinal $N_H - 1$. Soit $u \in V_H$ et $U \in \mathbb{R}^{N_H-1}$ le vecteur des coordonnées de u dans Φ . Le problème spectral discret est donné par

$$\ll \text{trouver } \lambda \in \mathbb{R} \text{ tel qu'il existe } U \neq 0 \text{ satisfaisant } \mathbb{K}U = \lambda \mathbb{M}U \gg \quad (4.1)$$

où \mathbb{M} et \mathbb{K} sont les matrices de masse et rigidité définies en (2.3). \mathbb{M} étant symétrique définie positive, elle admet une décomposition de Cholesky sous la forme $\mathbb{M} = LL^t$, où L est triangulaire inférieure inversible. Par le changement de base $V = L^tU$, le problème devient

$$L^{-1}\mathbb{K}L^{-t}V = \lambda V$$

et l'on peut le résoudre numériquement à l'aide d'un algorithme de puissance itérée. On obtient une famille $\{\tilde{\lambda}_n, \tilde{\omega}_n\}_{n \in \llbracket 1, M \rrbracket}$ d'approximations polynomiales par morceaux : or, la suite exige une représentation en variables $X_n = (A_{n,i}, C_{n,i}, \lambda_n)_{i \in \llbracket 0, N-1 \rrbracket}$. Sur chaque domaine Ω_i , on approche la restriction $\tilde{\omega}_n|_{\Omega_i}$ par $A \sin\left(\sqrt{\frac{\lambda_n}{\kappa_i}}x + C\right)$, où A et C sont déterminés par une méthode des moindres carrés. Enfin, la fonction propre est obtenue par l'algorithme de Newton.

Avec l'aide de Matlab pour le problème spectral, cette dernière stratégie a permis d'atteindre l'objectif de 600 fonctions en un temps raisonnable pour toutes les configurations étudiées. Au terme de cette croisade numérique, on dispose donc d'une solution de référence, et l'on peut se tourner vers les modèles éléments finis multi-échelles. Le lecteur curieux de l'aspect de la solution de référence pourra l'observer sur trois microstructures partageant $\mu(\Omega_i) = 160$ dans le dossier dédié.

référence 160 κ_{var} 

4.2 Choix d'implémentation

4.2.1 Configurations

Le choix des grandeurs est inspiré des travaux de Von Pamel et al. [2015]. Au fil du stage, nous nous sommes progressivement écartés des données de l'équipe de Von Pamel, qui ne nous permettaient pas de suffisamment distinguer les échelles macro et micro.

- **Domaine de simulation :** Le domaine spatial Ω est choisi de longueur 40 mm. Le temps final est égal à $T = 1.0 \mu\text{s}$. Dans toute la suite, on aura pris pour pas de discréétisation en temps celui donné par $\Delta t = 0.9 \times$ la condition CFL.
- **Microstructure :** La longueur caractéristique $\mu(\Omega_i)$ est constante pour tout i . Ce choix permet de faire des maillages éléments finis à pas constant, ce qui facilite grandement l'in-

terprétation des résultats. On présente les résultats pour $\mu(\Omega_i) = 100\mu\text{m}$ et $160\mu\text{m}$, ce qui correspond à respectivement 400 et 100 sous-domaines Ω_i .

- **Paramètre κ** : La moyenne de κ est de $\kappa_{\text{mean}}=28.4019371 \text{ mm}^2.\mu\text{s}^{-2}$. On a pris les caractéristiques de l'Inconel, soit un module de Young $E = 234.6 \text{ GPa}$, et une masse volumique $\rho = 8.260 \text{ g.cm}^3$. La variance de κ suivra une loi normale de variance comprise dans $\{0, 0.05, 0.1\} \times \kappa_{\text{mean}}$.
- **Condition initiale en vitesse** : La fonction u_1 est une sinusoïde de longueur d'onde $\lambda = \frac{\mu(\Omega)}{5} = 8\text{mm}$, modulée par une gaussienne de manière à ne conserver qu'une période.

La figure 4.2 illustre un exemple de configuration.

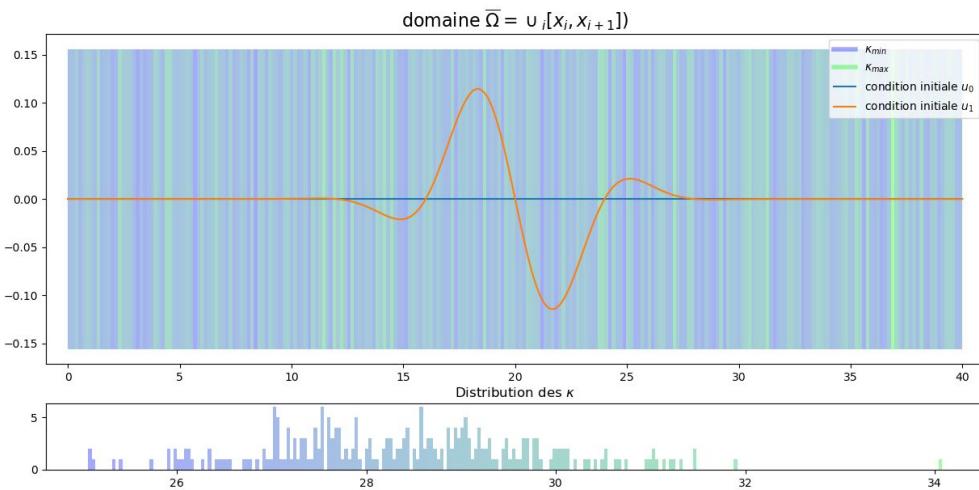


FIGURE 4.2 – Exemple pour $\mu(\Omega_i) = 160\mu\text{m}$ et $\kappa_{\text{var}} = 0.05$

Le plan d'expérience représente 36 cas : l'information contenue dans ce rapport sera donc partielle. On trouvera néanmoins tous les résultats dans le dossier dédié.

4.2.2 Maillages

Définition 4.2.1 (régimes de maillage) On appellera **régime sub-microstructuré** la plage de maillages macroscopiques pour laquelle κ est constante par maille, et **régime supra-microstructuré** la plage de maillages pour laquelle κ varie dans chaque arête.

Après plusieurs essais, la version retenue pour ce rapport est la suivante :

- les éléments finis macroscopiques s'appuient sur un maillage régulier. Quand le nombre de noeuds $N_H + 1$ le permet, on a choisi de prendre κ constante sur chaque maille : ainsi, en régime sub-microstructuré, $N_H + 1$ est choisi parmi les multiples de N .
- le maillage micro-éléments finis est irrégulier : il est construit de manière à minimiser la variance de κ sur chaque maille micro. Pour les faibles nombres de bulles, ceci entraîne

des différences significatives entre h_{mean} et h_{max} , compliquant l'interprétation de certains résultats : malheureusement, il n'est pas possible d'avoir un pas constant à toutes les échelles.

4.3 Résultats

Cette section est centrée autour des études de convergence des schémas quand le pas H macroscopique diminue. On trouvera des visualisations des solutions propagées dans le dossier dédié.

propagation [L][O]IMEX $\mu(\Omega_i)$ κ_{var}

4.3.1 Comportement de la CFL

La figure (4.3) est représentative de l'évolution de la CFL quand le pas de maille macroscopique diminue.

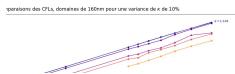


FIGURE 4.3 – variations de la CFL (IMEX éléments finis)

Les données sont représentées en échelle logarithmique. La valeur α indique la pente d'une approximation linéaire des données : on constate que la CFL est linéaire en fonction du pas spatial. Pour les modèles non orthogonalisés, l'ajout de fonctions bulles dégrade la condition CFL : cependant, un modèle enrichi avec des éléments finis dont le pas de maille micro est h sera plus facilement stable qu'un modèle non enrichi de pas macro h . Les CFL des modèles enrichis ont tendance à osciller légèrement pour de grandes valeurs de H : à notre connaissance, ceci est imputable à la construction des maillages.

Pour les modèles avec des espaces bulles orthogonalisés, la CFL ne dépend que des matrices du schéma standard : pour H constant, l'ajout de bulles ne modifie pas sa valeur. Ces modèles atteignent donc l'objectif fixé.

4.3.2 Influence de la microstructure sur le modèle macroscopique

Nous utiliserons les deux normes pour lesquels nous avons des résultats de convergence quand κ est constant sur chaque maille, c'est-à-dire

$$\|u - u_h\|_0 = \frac{\sup_{n \in \llbracket 0, N_T \rrbracket} \|u(\cdot, n\Delta t) - u_h^n\|_{L^2(\Omega)}}{\sup_{n \in \llbracket 0, N_T \rrbracket} \|u(\cdot, n\Delta t)\|_{L^2(\Omega)}} \quad \text{et} \quad \|u - u_h\|_1 = \frac{\sup_{n \in \llbracket 0, N_T \rrbracket} \|u(\cdot, n\Delta t) - u_h^n\|_{H^1(\Omega)}}{\sup_{n \in \llbracket 0, N_T \rrbracket} \|u(\cdot, n\Delta t)\|_{H^1(\Omega)}}$$

Par abus de langage, on désignera par **erreur L^2** l'erreur $\|u - u_h\|_0$, et **erreur H^1** celle donnée par $\|u - u_h\|_1$. On cherche à comparer l'évolution des modèles dans un milieu homogène et dans un milieu inhomogène. La figure (4.4) présente l'évolution de l'erreur L^2 quand H diminue, pour le modèle IMEX enrichi par éléments finis : on se concentre dans un premier temps sur la courbe sans bulles.

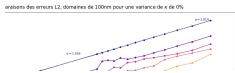


FIGURE 4.4 – différence homogène - inhomogène en norme L^2 (IMEX éléments finis)

On distingue plusieurs phénomènes. Premièrement, dans le cas homogène, le schéma éléments finis sans bulles converge bien à un ordre proche de 2. Dans le cas inhomogène, le comportement du modèle sans bulles distingue clairement les deux régimes de maillage : pour le cas sub-microstructuré

(au moins un point par Ω_i), l'ordre de convergence ne diminue pas beaucoup. Par contre, pour κ non constant, l'ordre est fortement dégradé. On propose un point de vue plus général sur les ordres de convergence dans la figure

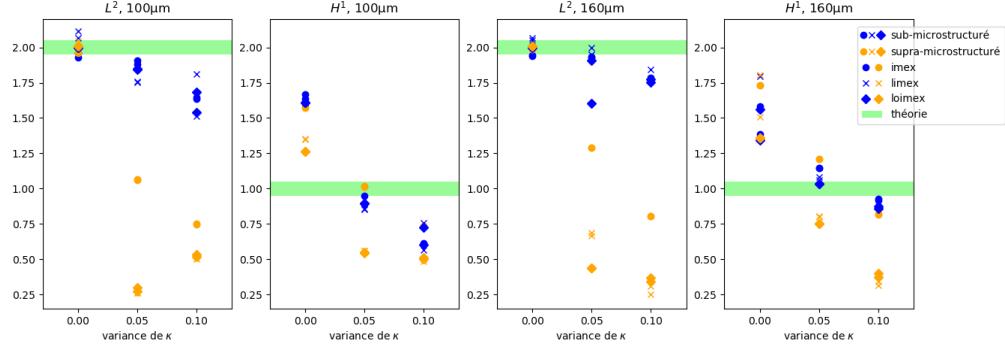


FIGURE 4.5 – Comparaison des ordres de convergence

On constate que dans le cas homogène, les résultats sont soit conformes, soit meilleurs que la théorie. On se doit de faire deux remarques problématiques : dans le cas sub-microstructuré, la variance de κ a un impact sur l'ordre de convergence, et il semblerait que le schéma (IMEX, LIMEX, LOIMEX) influe également. Sans explication alternative, on attribue ces étrangetés à l'implémentation.

Pour le cas supra-microstructuré, la convergence en norme L^2 est très affectée. Celle en norme H^1 est plus erratique. On tire la conclusion prudente que les discontinuités de κ perturbent le modèle élément fini en régime supra-microstructuré.

4.3.3 Influence des fonctions bulles

Comme la figure 4.4 permet de l'appréhender, à H fixé, les erreurs des modèles enrichis sont plus faibles que celles des modèles standard - pour le schéma IMEX. Toutefois, ce type de comparaison à H fixé n'est pas pertinente : on préférerait comparer des modèles de même longueur caractéristique. Pour les espaces bulles éléments finis, on comparera par exemple un modèle à N_B bulles avec un modèle non enrichi de pas $\frac{H}{N_B+1}$. (Notons que les maillages micro ne sont pas réguliers : le pas $\frac{H}{N_B+1}$ ne correspond « qu'en moyenne » au pas h .) La définition d'une longueur caractéristique pour les bulles spectrales est moins évidente : par défaut, on considère la même.

La figure 4.6 rassemble toutes nos données sur le modèle IMEX, en normes L^2 et H^1 .

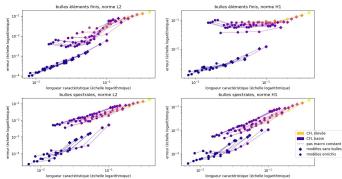


FIGURE 4.6 – Erreurs en fonction de la longueur caractéristique, modèle IMEX

Chaque point représente une moyenne sur les échelles de microstructures et les variances de κ de tous les modèles IMEX disponibles, pour un pas macroscopique, un nombre et un type de bulles fixés. En abscisses, la longueur caractéristique du modèle. En ordonnées, la moyenne des erreurs en norme considérée. Les couleurs indiquent la moyenne des CFL obtenues. Chaque chapelet de modèles commence par un losange pour le modèle sans bulles, puis les modèles enrichis à $N_B \in \{1, 2, 3, 6\}$ bulles. Pour les discrétisations les plus fines, la difficulté de calcul restreint les modèles aux premières valeurs de N_B .

On constate que pour une même longueur caractéristique, les performances des schémas enrichis sont très dépendantes du pas de maille macro. Sur chaque graphique, les régimes sub-microstructurés et supra-microstructurés se distinguent clairement, formant deux cohortes de filaments. On en tire les interprétations suivantes :

- En régime sub-microstructuré, ajouter des bulles éléments finis est similaire à un raffinement du maillage (concentration des chapelets autour d'une droite).
- En régime supra-microstructuré, les bulles éléments finis permettent de gagner un peu en norme L^2 , mais finissent par diverger (même pour 6 bulles, ce qui est très faible) et paraissent moins efficaces en norme H^1 qu'un modèle de même longueur caractéristique sans enrichissement.
- Les bulles spectrales sont plus régulières (pour la longueur caractéristique retenue, qui est peut-être discutable). Sur la fine plage de longueurs caractéristiques pour laquelle des modèles enrichis sont comparés à des modèles sans bulles (autour de 10^{-1} par exemple), on peut voir que l'erreur est un peu réduite par l'enrichissement, et la CFL est meilleure (couleur plus claire). Le comportement des bulles est également plus stable en norme H^1 .

4.3.4 Influence de l'orthogonalisation

Les bulles orthogonalisées n'ont qu'un très faible impact sur la qualité de la décomposition de u_1 , à l'initialisation du schéma. Ceci se retrouve dans la qualité de la solution propagée : la figure

4.7 est caractéristique du comportement du modèle LOIMEX.



FIGURE 4.7 – cas inhomogène en norme L^2 (IMEX spectral orthogonalisé avec condensation)

La comparaison générale des modèles en fonction de la longueur caractéristique est très défavorable aux modèles orthogonalisés :

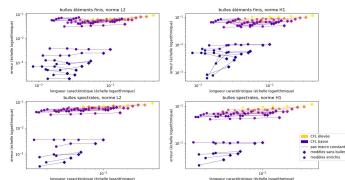


FIGURE 4.8 – Erreurs en fonction de la longueur caractéristique, modèle IMEX

À quelques exceptions près, l'erreur est invariante en fonction du nombre de bulles.

4.3.5 Influence de la condensation de masse

À notre surprise, il semblerait que la condensation de masse sur les éléments finis macroscopiques gêne les bulles quand les couplages de masse sont non nuls. L'étude de convergence suivante illustre

le phénomène :

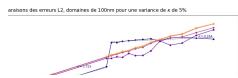


FIGURE 4.9 – cas inhomogène en norme L^2 (IMEX éléments finis avec condensation)

Le premier constat est que les bulles n'améliorent pas forcément l'approximation de la solution : les modèles enrichis sont parfois pires que ceux sans bulles. Le deuxième constat est qu'à H fixé, ajouter des bulles fait *divergir* l'estimation de u . La comparaison entre modèles de même longueur caractéristique a de quoi faire frémir :

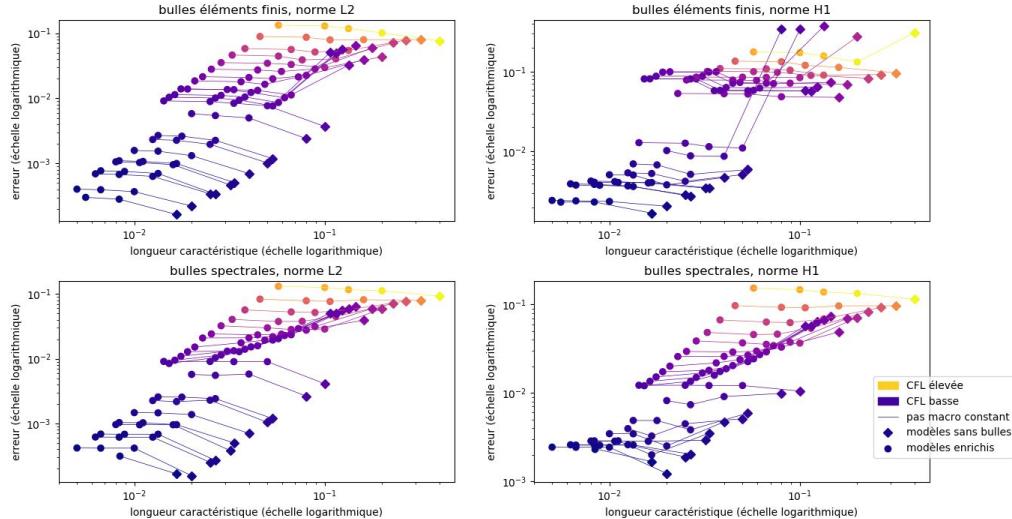


FIGURE 4.10 – Erreurs en fonction de la longueur caractéristique, modèle IMEX

Les modèles enrichis sont donc non seulement plus coûteux numériquement, mais deviennent **nuisibles à la précision du schéma** quand la matrice de masse macroscopique est diagonale. Or, la condensation de masse est fondamentale pour assurer les performances du schéma en dimension supérieure à 1. Sans analyse du phénomène qui permettrait de comprendre à quoi est dû cette

divergence, on est réduit à orthogonaliser les espaces bulles, ce qui, on l'a vu, n'est pas non plus d'une grande aide.

4.3.6 Influence du type de bulles

Les bulles éléments finis ont l'avantage de la simplicité : elles sont faciles à implémenter, les matrices de masse et de rigidité sont bien connues, et les termes de couplages avec l'échelle macroscopique ne demandent que d'intégrer des polynômes. Les bulles spectrales nécessitent un pré-traitement afin de les déterminer, qui peut être pénalisant dès la dimension 1 : par contre, les variations de κ sont directement prises en compte dans l'espace micro.

La figure compare les performances des deux types de bulles sur le modèle IMEX en terme d'amélioration à H constant.

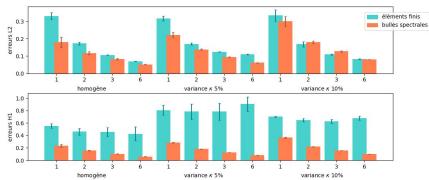


FIGURE 4.11 – Erreurs relatives des modèles enrichis sur le modèle sans bulle

Pour $\|\cdot\|$ désignant la norme L^2 ou H^1 , on a tracé la moyenne sur les pas macroscopiques et les tailles de microstructure de $\|\text{erreur modèle IMEX à } N_B \text{ bulles}\| / \|\text{erreur modèle IMEX sans bulles}\|$, $N_B \in \{1, 2, 3, 6\}$.

On constate qu'en norme L^2 , les modèles sont à peu près similaires. Pour la norme H^1 , les bulles spectrales sont bien plus efficaces : à N_B fixé, elles sont plus précises que les bulles éléments finis, et quand N_B augmente, elles convergent avec plus de fiabilité (sur l'échantillon observé).

4.3.7 Découplage des problèmes

Pour finir, évoquons un phénomène qui se produit quand les termes de couplages entre l'échelle des bulles et l'échelle macroscopique sont quasiment négligeables. Le cas extrême est celui d'un modèle orthogonalisé dans un milieu homogène : la matrice de masse n'a plus de termes de couplage, et la matrice de rigidité non plus. En effet, les bulles sont dans $H_0^1(K)$, et les dérivées des fonctions chapeau sur une arête K sont des constantes : à κ constant, on a donc

$$a(\varphi_I, \psi_j) = \kappa \int_K \partial_x \varphi_I(x) \partial_x \psi_j(x) dx = \kappa C [\psi_j]_{\partial K} = 0$$

La conséquence est que la résolution numérique revient à propager plusieurs problèmes indépendants : le squelette éléments finis se comporte comme dans un modèle non enrichi de même pas H , et sur chaque maille K , les fonctions bulles propagent une onde dont la condition initiale est $(u_1 - \mathcal{I}_H u_1)|_K$. L'ensemble n'a plus rien à voir avec un modèle multi-échelle. On pourra observer ces problèmes découpés sur un cas caricatural.

problèmes indépendants cas homogène 

Malheureusement, il semblerait que ce découplage ne soit pas uniquement confiné au cas homogène orthogonalisé : il se produit avec plus ou moins d'importance dès que les couplages sont faibles. Les couplages de la matrice de rigidité sont directement liés aux points de discontinuité de κ : quand ceux-ci sont peu nombreux, la propagation combine une résolution locale de l'équation d'onde avec des transferts d'énergie entre mailles.

problèmes indépendants cas inhomogène 

Les erreurs engendrées sont difficiles à détecter. En régime supra-microstructuré, les couplages de rigidité sont non nuls : par contre, le régime sub-microstructuré peut peut-être en souffrir.

Conclusion

Au terme de cette étude, nous pouvons dresser les conclusions suivantes.

La combinaison d'un modèle multi-échelle et d'un θ -schéma implicite-explicite permet de partiellement soulager la condition de Courant-Friedrich-Lowy de l'influence de la microstructure. Selon le type d'espace enrichi, cette stratégie peut permettre d'obtenir des modèles plus précis que les éléments finis standards, pour une même longueur caractéristique : en contrepartie, la propagation du schéma nécessite de résoudre un système linéaire plus exigeant qu'une matrice diagonale. L'optimisation classique de la condensation de masse n'est même pas envisageable, car inconciliable avec les bulles.

La contrainte supplémentaire de l'orthogonalisation des espaces bulles contre les polynômes du premier degré permet d'affranchir complètement la CFL de l'échelle micro. Par contre, le choix d'une méthode macroscopique éléments finis utilisant l'interpolation est incompatible avec ce type d'espaces bulles, et le projecteur naturel sur l'espace enrichi n'est plus convergent quand le nombre de bulles augmente. Les problèmes de représentation de fonctions font que l'on n'a pas pu pleinement juger de l'efficacité du schéma numérique, pourtant très prometteur grâce à une matrice de masse diagonale par blocs.

Cette analyse conduit à utiliser une méthode macroscopique discontinue, qui puisse approximer la projection de la solution sur l'espace \mathbb{P}_1 associé à chacune des mailles. L'espoir est d'obtenir une méthode efficace théoriquement et numériquement. Le cadre des méthodes de Galerkin se trouverait dépassé : on laisse cette piste à de futurs chercheurs.

Personnellement, ce travail m'a permis de beaucoup mieux comprendre les bases des schémas numériques utilisés. J'ai pu profiter du contact avec des experts du domaine pour entrevoir l'immensité de ce qu'il me reste à apprendre, tant dans le domaine mathématique que dans les joies du code.

Je tiens à remercier Alexandre pour ses précieux conseils et son énergie inépuisable, ainsi qu'Édouard DEMALDENT pour ses interventions (en particulier dans le domaine des demi-pizzas) et Antoine TONNOIR pour son soutien continual. Je souhaite vivement à de prochains camarades de pouvoir continuer à développer les liens entre INSA et CEA, pour explorer le merveilleux monde des bulles.

Bibliographie

- Hajer Methenni. *Modélisation mathématique et méthode numérique pour la simulation du contrôle santé intégré par ultasons de plaques composites stratifiées*. PhD thesis, 2021. (document)
- Grégoire Allaire. *Analyse numérique et optimisation*. 2012. 1.2.1
- Kösaku Yosida. *Functional analysis*. 2013. 1.2.4
- Haïm Brezis. *Analyse fonctionnelle. Théorie et applications*. 1983. 1.3.2
- Jacques-Louis Lions and Enrico Magenes. *Problèmes aux limites non homogènes et applications*. 1968. 2.1.1
- Patrick Joly. *Analyse et approximation de modèles de propagation d'ondes*. 2001. 2.1.1
- Pierre-Arnaud Raviart and Jean-Marie Thomas. *Introduction à l'Analyse Numérique des Equations aux Dérivées Partielles*. Collection Mathématiques Appliquées pour la Maitrise. 1988. 2.3.3
- Patrick Joly. *Waves in Nonlinear Pre-Stressed Materials*. CISM Courses and Lectures. 2007. 2.3.3
- Hergé. *L'Oreille cassée*. 1937. 3.3.3
- Sami Karkar. *Méthodes numériques pour les systèmes dynamiques non linéaires. Applications aux instruments de musique auto-oscillants*. PhD thesis, 2012. 4.1.2
- Anton Von Pamel, Colin R. Brett, Peter Huthwaite, and Michael J. S. Lowe. Finite element modelling of elastic wave scattering within a polycrystalline material in two and three dimensions. *Journal of Acoustical Society of America*, 2015. 4.2.1

GRILLE DE DÉROULEMENT DU STAGE

Légende :  = télétravail,  = présentiel,  = congés

Semaine	Mode de travail	Activités
amont		Préparation théorique
31/05	    	Découverte des locaux et des namespaces en C++, reprise du code écrit en amont
07/06	    	Méthode de prédiction-correction, configurations tirées de Von Pamel <i>et al</i> , algorithme naïf d'orthogonalisation
14/06	    	Minage des grandes configs (400µm), début des bulles
21/06	    	Semaine infernale de compilation sous Windows (problèmes d'édition des liens avec Boost et MinGW32)
28/06	    	Affichage python et calculs Matlab, le minage des bases de référence est enfin bien rodé, début des bulles spectrales
05/07	    	Tests d'un type d'orthogonalisation (finalement non retenu) et soutenance de mi-stage
12/07	    	Études de convergence, problèmes chez les bulles spectrales (le pas H constant sur des maillages à $\mu(\Omega_i)$ variable produit des sous-domaines de taille arbitrairement petite, faisant exploser la taille du problème spectral discret)
19/07	    	Passage à pas variable, études de convergence difficilement interprétables (soit en h_{mean} , pas de résultats théoriques, soit en h_{max} , trop peu de valeurs différentes)
26/07	    	Passage des configurations en pas constant, export des bases finales, études de convergence définitives
02/08	    	Point final, rédaction du rapport, adieu aux locaux
09/08	    	
16/08	    	Rédaction du présent rapport
23/08	    	Remise des badges et fin du stage.